

Table of Contents

ICM-Pro User Guide v.3.8.....	1
1 Introduction.....	3
Background.....	3
Minimum Specifications.....	3
Requesting an ICM license and Installation.....	3
Install ICM.....	3
How to Start ICM.....	3
2 Reference Guide.....	5
2.1 ICM-Browser Reference Guide.....	5
2.1.1 Download and Install ICM-Browser.....	5
2.1.2 How to use the Graphical Display.....	5
2.1.3 How to make Graphical Selections.....	6
2.1.4 How to Convert Proteins, Display Hydrogens and Ligand Binding Pocket.....	7
2.1.5 How to change Graphics Effects.....	7
2.1.6 How to add Labels and Annotations.....	7
2.1.7 How to Make High Quality Publication Images.....	8
2.1.8 How to Superimpose Protein Structures.....	8
2.1.9 How to Measure Distances and Angles.....	8
2.2 ICM-Browser-Pro Reference Guide.....	8
2.2.1 Download and Install ICM-Browser-Pro.....	8
2.2.2 Graphics.....	9
2.2.3 Protein Structure Analysis.....	9
2.2.4 Surfaces.....	9
2.2.5 Superimpose Proteins.....	9
2.2.6 Crystallographic Tools.....	9
2.2.7 Sequence Analysis.....	10
2.2.8 Plotting Tools.....	10
2.3 ActiveICM Reference Guide - Create 3D Molecular Documents for the Web and PowerPoint.....	10
Creating 3D Documents Is Straightforward.....	10
2.3.1 Getting Started.....	11
2.3.2 How to Create a Series of Fully-Interactive 3D Slides.....	11
2.3.3 How to Create Molecular Documents.....	11
2.3.4 How to Display Molecular Documents in PowerPoint.....	11
2.3.5 How to Display Molecular Documents on the Web.....	12
2.4 ICM-Chemist Reference Guide.....	12
2.4.1 How to Import, Sketch, and Edit Chemicals.....	12
2.4.2 How to Work with Chemical Spreadsheets.....	12
2.4.3 How to Undertake a Chemical Search.....	13
2.4.4 How to Work with Pharmacophores.....	14
2.4.5 How to Perform Chemical Clustering.....	14
2.4.6 How to Generate Stereoisomers and Tautomers.....	14
2.4.7 How to Generate Combinatorial Libraries.....	14
2.4.8 How to Generate Plots and Histograms.....	15
2.5 ICM-Chemist-Pro Reference Guide.....	15
2.5.1 How to use the ICM 3D Ligand Editor.....	15
2.5.2 How to Convert Chemicals to 3D.....	16
2.5.3 Ligand Energetics.....	16
2.5.4 How to Superimpose Chemicals.....	16
2.5.5 How to Perform QSAR.....	17
2.6 Menu Option Guide.....	17
2.6.1 File Menu.....	17
2.6.2 Edit Menu.....	19
2.6.3 View Menu.....	21
2.6.4 Bioinfo Menu.....	26
2.6.5 Tools Menu - Xray.....	26
2.6.6 Tools Menu - 3D Predict.....	26
2.6.7 Tools Menu - Analysis.....	26
2.6.8 Tools Menu - Superimpose.....	26
2.6.9 Tools Menu - Extras.....	26
2.6.10 Tools Menu - Table.....	26
2.6.11 Tools Menu - Chemical Search.....	28
2.6.12 Tools Menu - Molecular Editor.....	28

Table of Contents

2 Reference Guide

2.6.13 Homology Menu.....	28
2.6.14 Chemistry Menu.....	28
2.6.15 Docking Menu.....	28
2.6.16 MolMechanics Menu.....	28
2.6.17 Windows Menu.....	28
2.7 Tab Guide.....	29
2.7.1 Display Tab.....	30
2.7.2 Light Tab.....	30
2.7.3 Labels Tab.....	30
2.7.4 PDB Search Tab.....	30
2.7.5 Ligedit Tab.....	30
2.7.6 Meshes Tab.....	31

3 Getting Started.....33

3.1 The Basics of the Graphical User Interface.....	33
3.1.1 The Components of the GUI.....	33
3.1.2 How to load a PDB Structure.....	33
3.1.3 How to Move a Structure in the Graphical Display.....	34
3.1.4 How to Display Molecules using the ICM Workspace Panel.....	35
3.1.5 How to Make Selections.....	36
3.1.6 How to Change Protein Representation.....	37
3.1.7 How to Color.....	38
3.1.8 How to Display a Binding Pocket Surface.....	39
3.1.9 How to Save an ICM Object.....	40
3.1.10 How to Save an ICM Project File.....	41
3.1.11 How to Drag and Drop.....	41
3.1.12 How to: Right Click Options.....	41
3.1.13 How to Move Windows.....	41
3.1.14 How to Make a Picture.....	46
3.2 How to search and download Protein Structure, Sequences, and Chemicals.....	46
3.2.1 Search the PDB.....	47
How to Search the Protein Databank and Download.....	47
3.2.2 Search Pocketome.....	51
3.2.3 ChEMBL Search.....	53
3.2.4 Search SureChEMBL.....	54
3.2.5 BLAST Search.....	54
3.2.6 Search UniProt.....	55
3.2.7 Search PDB by Ligand Code.....	56
3.2.8 Search Drug Bank.....	56
3.2.9 Search PubChem.....	57
3.3 Create New Objects.....	58
3.4 Open and Read Files.....	59
3.4.1 Open with Password.....	59
3.4.2 Extract from icb file.....	59
3.5 Saving Files.....	60
3.6 Making Selections.....	61
3.6.1 Graphical Selection Tools.....	62
3.6.2 Quick Selection.....	62
3.6.3 How to Change the Selection Level and Mode.....	63
3.6.4 How to check what is selected.....	63
3.6.5 Orange Selection.....	64
3.6.6 Clear Selection.....	65
3.6.7 Changing a Selection.....	65
3.6.8 Filter Selection.....	66
3.6.9 Workspace Selections.....	67
3.6.10 Workspace Navigation.....	68
3.6.11 How to Select an Object.....	68
3.6.12 How to select a Molecule.....	68
3.6.13 How to select Residues.....	69
3.6.14 Select All.....	71
3.6.15 Selecting Neighbors.....	71
3.6.16 Selecting Neighbors: Graphical.....	71
3.6.17 Selecting Neighbors: Workspace.....	73
3.6.18 Select by Residue Number.....	73

Table of Contents

3 Getting Started

3.6.19 Alignment and Table Selections.....	73
3.6.20 Making Links.....	74
3.6.21 Tags.....	74
3.7 Preferences.....	77
3.7.1 Bonds Preferences.....	77
3.7.2 Directories Preferences.....	78
3.7.3 Graphics Preferences.....	79
3.7.4 GUI Preferences.....	81
3.7.5 GUI Preferences.....	82
3.7.6 Image Preferences.....	83
3.7.7 Label Preferences.....	84
3.7.8 Plot Preferences.....	85
3.7.9 Ribbon Preferences.....	85
3.7.10 Shell Preferences.....	86
3.7.11 System Preferences.....	86

4 Protein Structure.....89

4.1 Convert to ICM Object.....	89
4.1.1 Load a Protein Structure.....	89
4.1.2 Converting PDB Files Into ICM Objects.....	89
4.1.3 How to Edit and Convert a Chemical from the PDB into an ICM Object.....	90
4.2 Pocket Display.....	97
4.2.1 Receptor Pocket Surface.....	97
4.2.2 Ligand Surface.....	97
4.2.3 How to Display Hydrogen Bonds.....	98
4.2.4 Ligand Pocket Interactions.....	101
4.3 Crystallographic Analysis.....	102
4.3.1 Crystallographic Neighbor.....	102
4.3.2 Crystallographic Cell.....	104
4.3.3 Biomolecule Generator.....	105
4.3.4 Get Electron Density Map.....	106
4.3.5 Map's Original Cell.....	108
4.3.6 Contour Electron Density Map.....	108
4.3.7 Convert Xray Density to Grid.....	109
4.4 Protein Superposition.....	109
4.4.1 Select Proteins for Superposition.....	110
4.4.2 Superimpose Button.....	111
4.4.3 Superimpose by 3D.....	111
4.4.4 Superimpose Multiple Proteins.....	112
4.4.5 Arrange as Grid.....	113
4.4.6 Superimpose Sites by Atomic Property Fields.....	113
4.5 Protein Structure Analysis.....	119
4.5.1 Find Related Chains.....	119
4.5.2 Calculate RMSD.....	120
4.5.3 Contact Areas.....	120
4.5.4 Identify Closed Cavities.....	121
4.5.5 Surface Area.....	122
4.5.6 Measure Distances.....	122
4.5.7 Planar Angle.....	123
4.5.8 Dihedral Angle.....	124
4.5.9 Ramachandran Plot Interactive.....	125
4.5.10 Export Ramachandran Plot.....	125
4.6 3D Predict.....	125
4.6.1 Assign Helices and Strands.....	126
4.6.2 Protein Health.....	126
4.6.3 Local Flexibility.....	126
4.6.4 Protein-Protein Interface Prediction.....	127
4.6.5 Identify Ligand Binding Pockets.....	127

5 Molecular Graphics.....131

5.1 Molecule Representation.....	131
5.1.1 Wire Representation.....	132
5.1.2 Stick and Ball (Xstick) Representation.....	133
5.1.3 Ribbon Representation.....	134

Table of Contents

5 Molecular Graphics

5.1.4 Skin Representation.....	134
5.1.5 CPK Representation.....	135
5.1.6 Surface Representation.....	135
5.1.7 Display and Undisplay Hydrogens.....	136
5.1.8 Display Hydrogen Bond.....	137
5.1.9 Display Formal Charges.....	138
5.1.10 Ca (carbon alpha) Trace.....	138
5.2 Multi-Windows.....	139
5.3 Meshes - Surface - Grobs.....	140
5.3.1 Surfaces.....	141
5.3.2 MolSkin.....	142
5.3.3 Color Surface by Proximity.....	143
5.3.4 Color Surface by Selection.....	144
5.3.5 Meshes.....	146
5.3.6 Macroshape.....	146
5.3.7 Google 3D Objects (Sketchup).....	146
5.3.8 Display or Undisplay Meshes or Surfaces.....	147
5.3.9 Mesh Options.....	147
5.3.10 Mesh Clipping.....	149
5.3.11 Save Mesh.....	150
5.3.12 Occlusion Shading.....	150
5.4 Coloring.....	150
5.4.1 Coloring.....	151
5.4.2 Color Background.....	151
5.4.3 Background Image.....	151
5.5 Lighting.....	153
5.6 3D Stereo.....	154
5.7 3D Printing.....	154
5.8 Labeling and Annotation.....	154
5.8.1 Labeling.....	155
5.8.2 Labeling Atoms.....	155
5.8.3 Labeling Residues.....	155
5.8.4 Move Residue Label.....	156
5.8.5 Label Variables.....	156
5.8.6 Sites and Annotation.....	157
5.8.7 Changing Label Colors.....	160
5.8.8 Customized Label 2D or 3D.....	160
5.8.9 Undisplay Customized Label.....	161
5.8.10 Labeling Distances.....	161
5.8.11 (Un)display Origin.....	162
5.8.12 Displaying Tethers.....	163
5.8.13 Displaying Distance Restraints.....	163
5.8.14 Display Clash.....	163
5.8.15 Display Rainbow, Box, Ruler.....	164
5.8.16 Display Gradient.....	164
5.9 Display Distances and Angles.....	164
5.9.1 Display Distance Between Two Atoms - the quick way.....	164
5.9.2 Display Planar Angle.....	164
5.9.3 Display Dihedral Angle.....	165
5.9.4 Delete Label.....	165
5.10 Graphics Effects.....	165
5.10.1 Fog.....	165
5.10.2 Shadows.....	165
5.10.3 Sketch Accents.....	166
5.10.4 Elegant Ribbon & Ligand Sketch.....	166
5.10.5 Occlusion Shading.....	166
5.10.6 Perspective.....	167
5.10.7 Animate View.....	167
5.11 Graphics Shortcuts.....	169
5.12 Molecule Move Buttons.....	170
5.12.1 Rotation.....	170
5.12.2 Custom Rotation.....	170
5.12.3 Translation.....	171
5.12.4 Zoom.....	172

Table of Contents

5 Molecular Graphics	
5.12.5 Center.....	172
5.12.6 Torsion Angles.....	172
5.12.7 Connect (Move).....	173
5.13 Clipping Tools.....	174
5.13.1 Mesh Clipping.....	175
5.14 Graphic Layers.....	175
5.15 High Quality Publication Images.....	175
5.15.1 High Quality Image.....	175
5.15.2 Quick Image.....	176
5.15.3 Image Options.....	177
5.16 Movie Making.....	177
5.16.1 Making a Movie from a Set of Slides.....	178
5.16.2 Screenshot Movie.....	180
6 Slides & ActiveICM.....	183
6.1 Making Slides.....	183
6.2 Make a Movie from a Set of Slides.....	187
6.3 How to View and Navigate Slides.....	187
6.3.1 View Slide Show.....	187
6.3.2 Slide Navigation.....	188
6.4 How to Edit Slides.....	189
6.4.1 Edit Slide.....	189
6.4.2 Move Slide.....	190
6.5 How to Add Smooth Blending and Transition Effects Between Slides.....	191
6.6 How to Make Molecular Documents - Link HTML Text to Slides.....	191
6.6.1 How to Add Text or Edit a Molecular Document.....	192
6.6.2 How to Make a Hyperlink Between Text and a Slide.....	194
6.6.3 Insert Image.....	195
6.6.4 Insert Script.....	196
6.6.5 Insert a Dialog Box.....	197
6.6.6 Document Navigation.....	198
6.6.7 Protect Shell Objects From Deletion.....	199
6.7 ActiveICM.....	199
6.8 How to Embed in Microsoft PowerPoint 2003.....	199
6.9 How to Embed in Microsoft PowerPoint 2007.....	200
6.10 How to Embed in Microsoft PowerPoint 2010.....	202
6.11 Embed in Web Browser.....	202
6.12 How to Use ActiveICM in PowerPoint.....	203
6.13 How to Change ActiveICM Component Properties.....	204
6.14 Advanced use of activeICM: Macros to direct visualisation changes.....	206
6.14.1 PowerPoint Cache Errors.....	208
6.15 Background Images.....	208
6.16 ICM JavaScript (IcmJS).....	209
7 Working with Sequences and Alignments.....	211
7.1 Read and Save Sequences.....	211
7.1.1 Read a Sequence from UniProt.....	211
7.1.2 Cut and Paste a Sequence.....	211
7.1.3 Extract a Sequence from a PDB File.....	212
7.1.4 Read directly from a Sequence File.....	212
7.1.5 Save Sequence.....	212
7.2 Sequence Secondary Structure.....	212
7.3 Bioinfo Menu.....	213
7.3.1 Residue Content.....	214
7.3.2 Predict Secondary Structure.....	214
7.3.3 Six Frame Translation.....	214
7.3.4 Set Sequence Type.....	214
7.3.5 Align Two Sequences.....	215
7.3.6 Sequence to Structure alignment.....	215
7.3.7 Multiple Sequence Alignment.....	215
7.3.8 Alignment Strength Profile.....	215
7.3.9 Link to Structure.....	216
7.3.10 Extract Sub-Alignment As Is.....	217
7.3.11 Cut Vertical Alignment Block.....	217

Table of Contents

7 Working with Sequences and Alignments	
7.3.12 Reorder Sequences.....	218
7.3.13 Extract Unique Sequences.....	218
7.3.14 Load Example Alignment.....	218
7.4 BLAST Search.....	219
7.5 Sequence Editor.....	219
7.6 Sequence Alignments.....	221
7.6.1 Default Alignment.....	221
7.6.2 Align Two Sequences.....	221
7.6.3 Sequence to Structure alignment.....	223
7.6.4 Align DNA vs Protein.....	223
7.6.5 Multiple Sequence Alignment.....	224
7.6.6 Drag and Drop.....	225
7.7 Alignment Editor.....	225
7.7.1 Edit an Alignment.....	226
7.7.2 Save Alignment Image.....	227
7.7.3 Save Alignment.....	227
7.7.4 Print Alignment.....	227
7.7.5 Delete Alignment.....	227
7.7.6 Annotate Alignment.....	227
7.7.7 Box, Shade, Change Font.....	228
7.7.8 Phylogenetic Trees.....	229
7.7.9 Coloring an Alignment.....	230
7.7.10 Alignment View Options.....	231
7.7.11 Alignment Gaps.....	233
7.7.12 Searching an Alignment.....	235
7.7.13 Making Alignment Selections.....	235
7.7.14 Select by Consensus.....	237
7.7.15 Calculate Sequence Identity and Similarity.....	238
7.7.16 Alignment Secondary Structure.....	240
8 Protein Modeling.....	243
8.1 Homology Modeling.....	243
8.1.1 Homology Modeling Introduction.....	243
8.1.2 Quick Model.....	244
8.1.3 Full Model Builder.....	246
8.1.4 Multi-Template Model Editor.....	247
8.2 Loop Modeling.....	250
Background.....	250
8.3 Design Loop.....	251
8.4 Graft Loop.....	255
8.5 Loop Preferred Residues.....	260
8.6 Find Loops in PDB.....	260
8.7 Regularization.....	261
8.8 Refine Side Chains.....	261
8.9 Making a disulfide bond.....	262
8.10 Making a Mutation.....	262
8.11 Predicting Change in Binding Free Energy upon Mutation.....	264
8.12 Predicting Change in Protein Stability upon Mutation.....	267
8.13 Effect of Mutation on Protein-Peptide.....	270
8.14 Effect of Mutation on Protein Ligand Binding.....	273
8.15 Predict Disulfide Bond.....	273
8.16 Protein Sculpting.....	274
8.17 Sample Protein.....	276
8.18 Sample Peptide.....	276
8.19 Molecular Mechanics.....	276
8.19.1 ICM Convert.....	276
8.19.2 Optimize H,His,Asn,Gln,Pro.....	276
8.19.3 Regularization.....	277
8.19.4 Impose Conformation.....	277
8.19.5 Edit Structure.....	277
8.19.6 MMFF.....	278
8.19.7 Torsion Scan.....	278
8.19.8 Minimize.....	278
8.19.9 Sample Loop.....	279

Table of Contents

8 Protein Modeling

8.19.10 Design Loop.....	279
8.19.11 Sample Protein.....	279
8.19.12 Sample Peptide.....	279
8.19.13 Internal Coordinates Table.....	279
8.19.14 Generate Normal Mode Stack.....	282
8.19.15 Stack.....	282
8.19.16 GAMESS.....	283
8.19.17 Waters.....	283
8.19.18 Energy Terms.....	284

9 Cheminformatics.....285

9.1 Read Chemical Structures.....	285
9.1.1 Extract 2D structure from PDB Chemical.....	286
9.1.2 Extract 3D coordinates from 3D Chemical.....	286
9.2 Save Chemical Structures and Images.....	286
9.2.1 Saving from a chemical table.....	286
9.2.2 Saving Chemicals in the Molecular Editor.....	288
9.2.3 Saving Chemicals in the ICM Workspace.....	288
9.2.4 Saving Chemical Images.....	288
9.2.5 IUPAC Name.....	288
9.3 Working with Chemical Spreadsheets.....	288
9.3.1 Molecular Table Display.....	289
9.3.2 How to add columns into a chemical spreadsheet.....	290
9.3.3 How to sort a column(s) in a chemical spreadsheet.....	291
9.3.4 How to copy, cut and paste columns and rows in a chemical spreadsheet.....	291
9.3.5 How to show and hide columns and rows in a chemical spreadsheet.....	292
9.3.6 How to save a chemical spreadsheet in sdf format.....	293
9.3.7 How to export your chemical spreadsheet into Excel.....	293
9.3.8 How to print a chemical spreadsheet.....	294
9.3.9 How to filter columns in a chemical spreadsheet.....	294
9.3.10 How to use find and replace in a chemical spreadsheet.....	295
9.3.11 How to mark and label rows in a chemical spreadsheet.....	296
9.3.12 How to insert hyperlinks to the PDB, PubMed, and Uniprot.....	296
9.3.13 How to copy and paste 2D chemicals.....	297
9.3.14 How to edit data inside a chemical spreadsheet.....	297
9.3.15 How to remove salts, explicit hydrogens and standardize chemical groups.....	298
9.3.16 How to calculate chemical properties in a chemical spreadsheet.....	298
9.3.17 How to identify duplicate chemicals in a chemical spreadsheet.....	299
9.3.18 How to compare two chemical spreadsheets.....	299
9.3.19 How to merge two chemical spreadsheets.....	300
9.3.20 Display and Convert Molecule.....	300
9.3.21 Copy Molecule.....	300
9.3.22 Edit Molecule.....	301
9.3.23 Color Table Column.....	301
9.3.24 Chemical View Options.....	301
9.3.25 Chemical Table Side-by-Side View.....	302
9.3.26 Zoom, Translate and Z-rotate a Chemical in a table.....	303
9.3.27 Set Chemical Table 3D Browse Mode.....	303
9.3.28 Chemical Find and Replace.....	304
9.3.29 Split Chemical(s) into Fragments.....	306
9.3.30 Rotate Chemical for Best-Fit.....	307
9.3.31 Color Chemical Structure.....	307
9.4 Molecular Editor.....	308
9.4.1 Drawing a New Chemical Structure.....	309
9.4.2 Stereo Bond.....	310
9.4.3 Right Click Options.....	310
9.4.4 Chemical Dictionary.....	311
9.4.5 Dictionary of Chemical Groups.....	312
9.4.6 Template Menu.....	312
9.4.7 Set Free Radical Type.....	313
9.4.8 Adding and rotating a fragment in molecular editor by clicking-holding-and-dragging.....	314
9.4.9 Property Monitor.....	314
9.4.10 Editing structure using keyboard.....	315

Table of Contents

9 Cheminformatics

9.4.11 Chemical Editing Options.....	316
9.4.12 Paste SMILES or InChi String.....	316
9.4.13 Save Chemical in Mol/RXN Format.....	316
9.4.14 Save to Chemical Spreadsheet.....	316
9.4.15 Save as SMILES or InChi.....	317
9.4.16 Convert Chemical to 3D or Generate Conformers.....	317
9.4.17 Isis Draw Copy and Paste.....	317
9.5 Chemical Search.....	318
9.5.1 Query Setup.....	318
9.5.2 Filter Search.....	327
9.5.3 Query Processing.....	328
Data Source.....	328
Query Options.....	329
Results.....	329
Search.....	329
9.5.4 Search a Database by Text.....	329
9.5.5 Pharmacophore Drawing and Searching.....	330
9.6 2D Ligand Interaction Diagram.....	336
9.7 Convert Chemicals to 3D.....	337
9.7.1 Converting a Chemical from the PDB.....	339
9.7.2 Converting a Chemical from the PDB using the ICM Workspace.....	339
9.7.3 Converting a Chemical from the PDB using the Graphical Display.....	339
9.8 ToxScore.....	340
How the ToxScore is Calculated.....	340
How to Calculate ToxScore.....	340
9.9 Generating Chemical Fragments.....	340
9.10 Molcart.....	341
9.10.1 Molcart Installation.....	341
9.10.2 Molcart Getting Started.....	342
9.10.3 Creating Databases and Importing Data.....	342
9.10.4 Molcart Search.....	343
9.10.5 Molcart Administration.....	347
9.11 Calculate Properties.....	349
9.12 Standardize Table.....	349
9.13 Annotate By Substructure.....	350
9.14 Align/Color by 2D Scaffold.....	353
9.15 Set Formal Charges.....	354
9.16 Enumerate Formal Charge States.....	354
9.17 Convert.....	355
9.17.1 Convert 2D to 3D.....	355
9.17.2 3D to 2D Depiction.....	355
9.17.3 Convert Smiles to 2D.....	355
9.17.4 Convert Structure to Smiles.....	356
9.17.5 Convert to Racemic.....	356
9.18 Build Prediction Model.....	357
9.19 Predict.....	357
9.20 Generate 3D Conformers.....	357
9.21 Generate Tautomers.....	358
9.22 Generate Stereoisomers.....	360
9.23 Prodrug.....	362
9.24 Ligand Energetics.....	363
9.24.1 Conformational Entropy.....	363
9.24.2 Strain.....	363
9.25 Cluster Set.....	363
9.25.1 How to perform chemical clustering.....	363
9.25.2 How to select representative centers from a tree.....	364
9.25.3 How to reorder branches and change the distance of trees.....	365
9.25.4 How to edit the tree - labels, spacing and coloring.....	365
9.25.5 Auto Close Tree Mode.....	366
9.25.6 Calculate Maximum Common Substructure for Nodes.....	366
9.25.7 R Group Decomposition from Clustering Node.....	366
9.26 PCA Analysis.....	367
9.27 Compare Two Sets.....	367
9.28 Merge Two Sets.....	368

Table of Contents

9 Cheminformatics

9.29 Select Duplicates.....	368
9.30 Combinatorial Chemistry.....	369
9.30.1 Create/Modify Markush.....	369
9.30.2 Enumerate by Scaffold.....	371
9.30.3 R-Group Decomposition.....	373
9.30.4 Enumerate by Reaction.....	375
9.31 Substructure Superposition.....	377
9.31.1 Rigid Superposition of Compounds in a Table onto a Template in The Graphical Display.....	378
9.31.2 Rigid Substructure Superposition.....	379
9.31.3 Flexible Substructure Superposition.....	379
9.32 Atomic Property Field Superposition.....	379
9.32.1 Pairwise APF Alignment/Superposition.....	380
9.32.2 Flexible APF Superposition to Template.....	380
9.32.3 Multiple APF Alignment of Compounds in a Table.....	382
9.33 APF Tools.....	382
9.33.1 Consensus Pharmacophore.....	383
9.33.2 Pairwise APF Score.....	386
9.33.3 APF Ligand Based Screen - Screen SD file against template molecule(s).....	389
9.33.4 APF 3D QSAR.....	389

10 Learn and Predict.....391

10.1 Learn.....	391
10.1.1 About the Model.....	392
10.1.2 Model Weights.....	392
10.1.3 Improve Model.....	393
10.2 Predict.....	393
10.3 Fingerprint Methods.....	393
10.4 A Little Theory on Learning.....	393

11 Docking.....395

11.1 Small Molecule Docking.....	395
11.1.1 Docking Introduction.....	395
11.1.2 Docking Steps.....	397
11.1.3 Docking Project Setup.....	397
11.1.4 Review and Adjust Binding Site.....	402
11.1.5 (Re) Make Receptor Maps.....	402
11.1.6 Check Docking Preferences.....	402
11.1.7 Interaction Restraints.....	404
11.1.8 Begin the Docking Simulation.....	405
11.1.9 Interactive Docking.....	405
11.1.10 Batch Docking.....	407
11.1.11 Run Docking Batch.....	409
11.1.12 Viewing Your Docking Results.....	410
11.1.13 Make a HIT LIST.....	412
11.1.14 Hitlist Export.....	418
11.1.15 Dock Chemical Table.....	419
11.1.16 Reload a Docking Project.....	422
11.2 Induced Fit (Flexible) Docking.....	422
11.2.1 Ligand and Receptor Complex Refinement.....	422
11.2.2 Multiple Receptor 4D Docking.....	423
11.2.3 Explicit Group Docking.....	429
11.2.4 SCARE - Dual Alanine Scanning and Refinement.....	429
11.3 Template Docking.....	439
11.4 Covalent Docking.....	439
11.4.1 Sketch a Reaction.....	440
11.5 ICM X-Ray AutoFit - Automated Model Building into Density.....	443
11.6 Protein-Protein Docking.....	445
11.6.1 Determination of Protein-Protein Docking Interface.....	445
11.6.2 Protein-Protein Docking Procedure.....	445
11.6.3 Protein-Protein Set Project.....	445
11.6.4 Protein-Protein Receptor Setup.....	446
11.6.5 Protein-Protein Ligand Setup.....	446
11.6.6 Epitope Selection.....	447

Table of Contents

11 Docking	
11.6.7 Protein-Protein Make Receptor Maps.....	448
11.6.8 Protein-Protein Docking Batch.....	448
11.6.9 Display Grid Docking Results.....	449
11.6.10 Complex Refinement.....	450
12 Virtual Ligand Screening.....	451
12.1 Virtual Ligand Screening.....	451
12.1.1 Virtual Ligand Screening.....	451
12.1.2 VLS Getting Started.....	451
12.1.3 Database File Format.....	451
12.1.4 VLS Preferences.....	452
12.1.5 Docking Project Table File.....	453
12.1.6 Run VLS in the Graphical User Interface.....	453
12.1.7 Running VLS Jobs in PBS UNIX Cluster Environment.....	453
12.1.8 Parallelization.....	453
12.1.9 VLS Results.....	454
12.1.10 Post Screen Analysis.....	454
12.1.11 Post Screen Visualization.....	459
12.2 Fragment Screening.....	460
12.3 Ligand-Based Screening using APF 3D Pharmacophores.....	468
13 MolScreen.....	473
13.1 Model Types.....	473
13.1.1 Kernel Chemical Classification/Regression (kcc) models.....	473
Hybrid Fingerprints.....	475
Dataset.....	477
Training.....	479
Predicted Value.....	481
13.1.2 Docking to Ligand Field Z-Score (dfz) models.....	481
APF.....	483
Dataset.....	485
Training.....	487
Predicted Value.....	489
13.1.3 Docking to Ligand Field Classification/Regression (dfa) models.....	489
Hybrid 4D/2D.....	491
Dataset.....	493
Training.....	495
Predicted Value.....	497
13.1.4 Docking to Protein Pocket Classification/Regression (dpc) models.....	497
Hybrid Docking.....	499
Dataset.....	501
Training.....	503
Predicted Value.....	505
13.2 Brief Panel.....	505
13.3 Full Model Panel.....	505
13.4 Custom Model Panel.....	505
13.5 Results Table.....	506
13.6 Make Chemical Classification Model.....	506

Table of Contents

Predicted Value	
13.7 Make APF Docking SAR Model.....	507
14 The 3D Fully Interactive Ligand Editor.....	509
14.1 Setup Ligand and Receptor.....	509
14.2 Ligand-Editor-Preferences.....	513
14.3 Display Options.....	513
14.3.1 Display the receptor pocket surface.....	513
14.3.2 Display the ligand pocket surface.....	514
14.3.3 Display Hydrogen Atoms.....	515
14.3.4 Display Hydrogen Bonds.....	515
14.3.5 Display Atomic Energy Circles.....	516
14.3.6 Display Pretty View.....	517
14.3.7 2D Interaction Diagram.....	518
14.3.8 Unsatisfied Hydrogen Bonds.....	518
14.3.9 Display Relaxed Ligand.....	519
14.3.10 Visualize Ligand Strain.....	520
14.3.11 Identify Ligand-Receptor Contacts.....	521
14.3.12 Center on Ligand.....	522
14.4 Evaluate Ligand Score and Strain.....	523
14.5 Edit Ligand.....	524
14.5.1 Edit Ligand in 2D.....	525
14.5.2 Multiple Substituents.....	527
14.5.3 User Defined Modifiers.....	528
14.5.4 Undo and Redo.....	529
14.5.5 Modification History.....	529
14.5.6 Save to Chemical Spreadsheet.....	530
14.5.7 After Editing the Ligand.....	532
14.6 Virtual Screen: Find Best Replacement Group.....	532
14.7 Multiple Position Group Scan.....	533
14.8 Dock or Minimize Ligand.....	533
14.8.1 Change the Size of the Docking Region.....	534
14.9 Dock from Table.....	534
14.9.1 Covalent Docking from Table.....	535
14.10 Fragment Linking.....	540
14.11 Impose Tethers or Distance Restraints.....	543
14.11.1 Impose Ligand Atom Tether.....	544
14.11.2 Impose Distance Restraint.....	544
14.12 Defined Flexible Receptor Groups.....	546
14.13 Side Chain Refinement.....	546
14.14 Multiple Receptor Docking.....	546
14.15 Covalent Docking in the Ligand Editor.....	551
14.16 Dock to Atomic Property Fields.....	556
14.17 How to Save Ligand Receptor Complex as PDB.....	561
14.18 Export Docking Project.....	562
14.19 Close and Clean Up Project.....	562
15 Working with Tables.....	565
15.1 Standard ICM Tables.....	565
15.1.1 Generate New Table.....	565
15.1.2 Reading a Table.....	566
15.1.3 Saving a table.....	566
15.1.4 Basic Table Navigation.....	567
15.1.5 Table View (Grid Layout).....	567
15.1.6 Change Row Height.....	567
15.1.7 Table View Save.....	568
15.1.8 Table Search.....	568
15.1.9 Table Color.....	568
15.1.10 Table Font.....	568
15.1.11 Table Alignment.....	569
15.1.12 Mark a Row.....	569
15.1.13 Table right click options.....	569
15.1.14 Rename a Table.....	569
15.1.15 Clone a Table.....	570
15.1.16 Delete a Table.....	570

Table of Contents

15 Working with Tables	
15.1.17 Page Setup	570
15.1.18 Print a Table	570
15.1.19 Export to Excel	570
15.1.20 Save a Table	570
15.1.21 Change Column and Row Width	570
15.1.22 Making Table Selections	571
15.1.23 Editing a Table	573
15.1.24 Inserting Columns	573
15.1.25 Column Statistics	574
15.1.26 Inserting Rows	574
15.1.27 Copy Cut and Paste Row	575
15.1.28 Copy Cell	575
15.1.29 Copy Selection to an ICM Table	575
15.1.30 Deleting Columns and Rows	576
15.1.31 Hide and Show Columns	576
15.1.32 Column Format and Custom Actions	577
15.1.33 Table Sorting	579
15.1.34 Table Filtering and Appending	580
15.1.35 Mark and Select Rows	582
15.1.36 Group View by Column	583
15.1.37 Mouse and Cursor Actions on a Table	584
15.2 Molecular Tables	584
15.3 Insert Interactive Objects into Table Cell	585
15.4 Plotting Table Data	585
15.4.1 Column Histogram	586
15.4.2 Histogram Options	587
15.4.3 Histogram Bins	588
15.4.4 Plots	588
15.4.5 Heatmap	589
15.4.6 Add a title to a plot	590
15.4.7 Axis Options	590
15.4.8 Change Axis Data	590
15.4.9 Logarithmic Plots	590
15.4.10 Change Mark Shape or Size	591
15.4.11 Change Mark Color	591
15.4.12 Grid and Axis Display	592
15.4.13 Least Squares Fitting	592
15.4.14 Zoom, Translate and Center	593
15.4.15 Plot Selection	594
15.4.16 Print Plot	595
15.4.17 Saving a Plot Image	595
15.4.18 Table Inline Plots	595
15.4.19 Annotate Plot	596
15.5 Principal Component Analysis	596
15.6 Learn and Predict	597
15.7 Data Clustering	597
15.7.1 Cluster	598
15.7.2 Tree Selection	599
15.7.3 Save and Print Tree	600
15.7.4 Tree View	601
16 Working with Local Databases	605
16.1 How to make a local database	605
16.2 Browse Database	605
16.3 Edit Database	607
16.4 Query Local Database	609
17 ICM-Scrab - Data Capture, Integrator and Miner	611
A Flexible LIMS	611
Unique Flexible Data Capture and Mining Interface	611
A One-Stop-Shop	611
17.1 Scrab Installation	611
17.2 Upload Data	612
17.3 Query Metadata	615

Table of Contents

17 ICM-Scarab - Data Capture, Integrator and Miner	
17.4 Browse and Export Data.....	618
17.5 Add a New User.....	619
17.6 Create a New Project and Assign User.....	621
17.7 Pages.....	623
18 ICM in KNIME.....	629
Available Nodes.....	629
To use ICM in KNIME.....	629
19 Tutorials.....	635
19.1 Graphics Tutorials.....	636
Introduction.....	636
19.1.1 Change Molecule Representation and Color.....	636
19.1.2 Annotation.....	639
19.1.3 Labels.....	640
19.1.4 2D and 3D Labels.....	643
19.2 Creating Fully Interactive Slides for PowerPoint and the Web Tutorial.....	646
Introduction.....	646
Tutorial.....	646
19.3 Protein Structure Tutorials.....	652
Introduction.....	652
Tutorial.....	653
19.3.1 PDB Search.....	653
19.3.2 Convert PDB to ICM Object.....	653
19.3.3 Protein Superposition.....	654
19.3.4 Protein Structure Analysis.....	655
19.3.5 Contact Areas.....	656
19.3.6 Hydrogen Bond.....	658
19.3.7 Pocket Identification.....	658
19.3.8 Protein Preparation and Crystallographic Analysis Tutorial.....	659
19.3.9 PDB Preparation - Symmetry.....	659
19.3.10 PDB Preparation - Occupancy and B-Factors.....	660
19.3.11 PDB Preparation - Residue Alternative Orientation.....	661
19.3.12 Biomolecule Generator.....	662
19.4 Sequence and Alignment Tutorials.....	663
19.4.1 Display and Superimpose Protein Kinase Structures in Preparation for the Tutorial.....	664
19.4.2 Linking Sequence Alignment to Structure.....	668
19.4.3 Identify Sequence Conservation in Ligand Binding Pocket.....	670
19.4.4 Alignment Annotation.....	673
Introduction.....	673
19.5 Protein Modeling Tutorials.....	677
19.5.1 GPCR Modeling Example.....	677
19.5.2 Loop Modeling Tutorial.....	682
19.5.3 Effect of Mutation on Protein Binding Tutorial.....	688
19.5.4 Effect of Mutation on Protein Stability Tutorial.....	691
19.6 Cheminformatics Tutorials.....	694
19.6.1 Chemical Draw and Spreadsheet Tutorial.....	694
Background.....	694
Tutorial.....	694
19.6.2 Chemical Searching.....	702
Background.....	702
Tutorial.....	702
19.6.3 Chemical Clustering.....	708
Background.....	708
Tutorial.....	708
19.6.4 3D Pharmacophore Searching.....	710
19.6.5 2D Pharmacophore Searching.....	711
19.6.6 How to create a Markush structure.....	712
19.6.7 How to enumerate a Markush library.....	715
19.6.8 How to decompose a library based on a Markush structure.....	717
19.6.9 How to enumerate a chemical library by reaction.....	718
19.7 3D Ligand Editor Tutorial.....	720
19.8 Atomic Property Field Tutorials.....	731

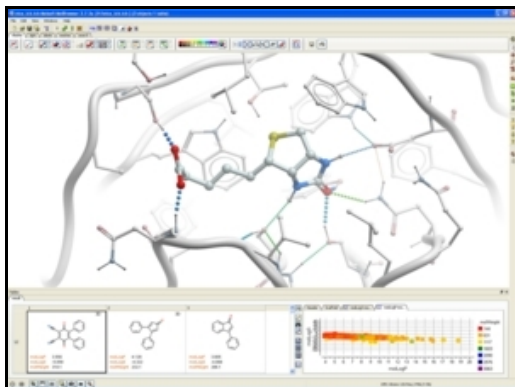
Table of Contents

19 Tutorials

19.8.1 APF Superposition.....	732
-------------------------------	-----

ICM-Pro User Guide v.3.8

by Ruben Abagyan, Andrew Orry, Eugene Raush, and Maxim Totrov
Copyright © 2015

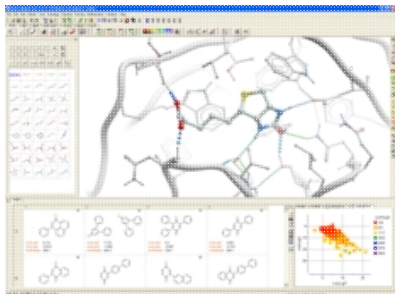


Oct 28 2016

Feedback.

1 Introduction

Background



The ICM Suite of Software provides an easy to use general environment for a biologist or chemist who is curious about protein structure. In just a few seconds you can browse hundreds of structures of interest, analyze and visualize sequences, alignments and binding sites. You can also perform molecular modeling, fully-flexible ligand and receptor docking, virtual ligand screening, chemical similarity searching, chemical clustering and much more... This book describes how to use the program via the Graphical User Interface (GUI) without the knowledge of the commands and functions running through your terminal window. After reading this book you may read the full

ICM Language Reference Manual (www.molsoft.com/man) for ICM command line users and programmers. This is a separate document which is provided with a full ICM distribution or can be obtained from Molsoft in a printed form.

ICM is constantly updated with new features and so it is always a good idea to keep an eye on the [release notes](#). Most of the things you will read in this manual are sort of natural or can be figured out by common sense and trial-and-error. Since this book is intended for basically anyone who is even remotely interested in molecules, some basic knowledge of biology and chemistry is implied. Do not try to find definitions of "atom" or "sequence" here, but most terms beyond that will be explained. A keyword search of this manual is available on the online version which is located at (www.molsoft.com/gui). For detailed information regarding ICM programming please consult the separate ICM language reference guide (www.molsoft.com/man). If you want to have a more hands-on introduction to ICM you are always more than welcome to join us at one of our ICM workshops which are held periodically throughout the year ([click here for training information](#)).

Minimum Specifications

The minimum specifications are listed here: <http://www.molsoft.com/min-specs.html>

Requesting an ICM license and Installation

Please [contact us](#) to request a license key.

Install ICM

Full installation instructions are provided at our support site www.molsoft.com/support

How to Start ICM

Starting ICM in Windows

There are several ways to start ICM under Windows, including the following:

1. Click the **Start** button on the taskbar, then select **Programs/Molsoft/ICM**
2. Double-click the file with one of the ICM extensions, including pdb files (`--*.pdb`) icm projects and binaries (`*.icb`), and other ICM file types.
3. Create a shortcut to the ICM binary.
4. Start a command prompt window and type the path the ICM binary, usually it is `C:\Program Files\Molsoft LLC\ICM\icm.exe -g`. In this case you can start ICM with different options.

Starting ICM in Unix or Mac

Mac users can start icm as described below in unix or click on the ICM icon.

Under a UNIX platform your executables will reside in the `$ICMHOME` directory.

`$ICMHOME` is an environmental variable of your UNIX shell and it needs to be set to the actual location of the icm files. The installation procedure does tell you to what value the ICMHOME variable needs to be set.

Examples:

```
/usr/icm/icm -X # returns
/usr/icm/icm -g # -g GUI interface will be displayed
/usr/icm/icm -h # help
/usr/icm/icm -bio -g # starts ICM-bio program with gui
/usr/icm/icm -G # GUI interface will be displayed with a separate window for the ICM command
```

Once you are in ICM you can spawn another window by choosing **File/New ICM Window** .

In this case if you close the main ICM window, all the children will be closed too.

Activating the Graphical User Interface

If you are running ICM in Windows then the graphical user interface will be displayed automatically.

However in Unix the GUI version of ICM can be activated by typing `icm -g` or `icm -G` and hitting RETURN. Or, to start the graphical user interface from the ICM command line, simply type `gui`.

2 Reference Guide

Chapter Contents:

ICM-Browser Reference Guide

ICM-Browser-Pro Reference Guide

ActiveICM Reference Guide - Create 3D Molecular Documents for the Web and PowerPoint

ICM-Chemist Reference Guide

ICM-Chemist-Pro Reference Guide

Menu and Tab Reference Guide:

Menu Option Guide

Tab Guide

2.1 ICM-Browser Reference Guide

For instructions on how to use ICM-Browser to make fully-interactive 3D slides and publish them in PowerPoint and the web please see the [ActiveICM User Guide](#). ActiveICM is a free plugin for Windows PowerPoint and web browsers. Other related tutorials include:

- Graphical Display: Molecule Representation, Coloring, Labeling and Annotation
- Graphical Selections Tutorial
- Creating Fully Interactive Slides for PowerPoint and the Web Tutorial

2.1.1 Download and Install ICM-Browser

Getting Started: Download and Install ICM-Browser and ActiveICM.		
Download ICM-Browser Distribution.	Download	video
Install ICM-Browser Instructions.	Windows Linux Mac	
Download ActiveICM Distribution.	Download	video
Install ActiveICM.	Windows Linux Mac	

2.1.2 How to use the Graphical Display

```
}; winRef.document.write(str); }
```

How to use the Graphical Display		
How to search the PDB.	HTML GUI Manual	video

How to Move a Structure in the Graphical Display.	HTML GUI Manual	video
How to use the Graphics window controls.	HTML GUI Manual	
How to use the ICM Workspace Panel	HTML GUI Manual	video
How to Display a Molecule.	HTML GUI Manual	video
How to Change Protein Representation.	HTML GUI Manual	video
How to Change Ribbon Representation.	HTML GUI Manual	video
How to color wire or xstick carbon atoms.	HTML GUI Manual	video
How to Display the Residues Surrounding the Ligand Binding Pocket.	HTML GUI Manual	video
How to remove chain breaks (dotted lines).	HTML GUI Manual	video
How to Color.	HTML GUI Manual	video
How to Change the Color of Molecule Representations.	HTML GUI Manual	video
How to Change the Background Color.	HTML GUI Manual	video
How to Display a Binding Pocket Surface	HTML GUI Manual	video
How to (Un)Display Hydrogens.	HTML GUI Manual	video
How to Save an ICM Object.	HTML GUI Manual	video
How to Save an ICM Project File.	HTML GUI Manual	video
How to Drag and Drop.	HTML GUI Manual	
How to: Right Click Options.	HTML GUI Manual	
How to Move Windows.	HTML GUI Manual	video
How to Arrange Windows	HTML GUI Manual	video

2.1.3 How to make Graphical Selections

How to Make Selections.	HTML GUI Manual	
How to Select an Object	HTML GUI Manual	video
How to Select a Molecule	HTML GUI Manual	video
How to Select Residues	HTML GUI Manual	video
How to Select Atoms	HTML GUI Manual	video
How to Make a Spherical Selection.	HTML GUI Manual	video

How to Invert a Selection.	HTML GUI Manual	video
How to Remove a Selection.	HTML GUI Manual	video
How to Change the Selection Level and Mode.	HTML GUI Manual	video
How to Check What is Selected.	HTML GUI Manual	

2.1.4 How to Convert Proteins, Display Hydrogens and Ligand Binding Pocket.

Convert Protein, Display Hydrogens and Ligand Binding Pocket.		
How to Convert a PDB Structure into an ICM Object.	HTML GUI Manual	video
How to Display Ligand Binding Pocket.	HTML GUI Manual	video
How to Display Hydrogen Bonds.	HTML GUI Manual	video

2.1.5 How to change Graphics Effects

How to change Graphics Effects		
How to display the FOG effect.	HTML GUI Manual	video
How to display side-by-side stereo.	HTML GUI Manual	
How to display Anaglyph stereo.	HTML GUI Manual	
How to toggle full screen mode.	HTML GUI Manual	video
How to adjust perspective.	HTML GUI Manual	
How to change the lighting.	HTML GUI Manual	video
How to display sketch accents.	HTML GUI Manual	video
How to display elegant ribbon and ligand sketch.	HTML GUI Manual	video

2.1.6 How to add Labels and Annotations

How to add Labels and Annotations		
How to Label Residues.	HTML GUI Manual	video
How to Label Atoms.	HTML GUI Manual	video
How to Label Variables.	HTML GUI Manual	video
How to Display and Undisplay Sites.	HTML GUI Manual	

How to Make and Display Sites (Annotations).	HTML GUI Manual	
How to Make and Display 2D and 3D Labels.	HTML GUI Manual	video

2.1.7 How to Make High Quality Publication Images

How to Make High Quality Publication Images		
How to Toggle High Quality Display	HTML GUI Manual	video
How to Toggle Antialiasing.	HTML GUI Manual	video
How to Copy Image to ClipBoard	HTML GUI Manual	video
How to Write an Image.	HTML GUI Manual	video
How to Use the Advanced Write Image Options.	HTML GUI Manual	video
How to Add an Image to the ICM Photo Album.	HTML GUI Manual	video

2.1.8 How to Superimpose Protein Structures.

How to Superimpose Protein Structures.		
How to Superimpose Two or More Protein Structures.	HTML GUI Manual	video

2.1.9 How to Measure Distances and Angles.

How to Measure Distances and Angles.		
How to Measure Distances Between Two Atoms.	HTML GUI Manual	video
How to Measure Distances From One Atom to Many.	HTML GUI Manual	video
How to Show Corresponding Distances in Two Objects.	HTML GUI Manual	video
How to Display the Ruler Bar.	HTML GUI Manual	video

2.2 ICM-Browser-Pro Reference Guide

NOTE: ICM-Browser-Pro contains all the features in ICM-Browser. [Click here](#) for the ICM-Browser Reference Guide.

2.2.1 Download and Install ICM-Browser-Pro

Getting Started: Download and Install ICM-Browser-Pro	
Download ICM-Browser-Pro Distribution.	Download
Install ICM-Browser-Pro.	Windows Linux Mac

2.2.2 Graphics

Graphics and Movie Making	
How to generate the shadow effect.	HTML GUI Manual
How to make a screenshot movie.	HTML GUI Manual
How to make a movie from a set of slides.	HTML GUI Manual
How to move a molecule independently of the other display objects (Connect).	HTML GUI Manual

2.2.3 Protein Structure Analysis

Protein Structure Analysis	
How to calculate contact areas between molecules.	HTML GUI Manual
How to identify closed cavities.	HTML GUI Manual
How to calculate surface area.	HTML GUI Manual
How to generate interactive Ramachandran plots.	HTML GUI Manual

2.2.4 Surfaces

Surfaces	
How to generate electrostatic and binding property surfaces.	HTML GUI Manual
How to connect and rotate/translate surface (mesh).	HTML GUI Manual
How to crop a mesh/surface.	HTML GUI Manual
How to save a mesh/surface.	HTML GUI Manual

2.2.5 Superimpose Proteins

Superimpose Proteins	
How to superimpose proteins based on 3D by visible atoms, C-alpha, backbone or heavy atoms.	HTML GUI Manual
How to superimpose multiple proteins based on aligned residues..	HTML GUI Manual

2.2.6 Crystallographic Tools

Crystallographic Tools	
How to contour electron density.	HTML GUI Manual

How to display crystallographic cell.	HTML GUI Manual
How to display crystallographic symmetry.	HTML GUI Manual
How to convert x-ray density to a grid.	HTML GUI Manual

2.2.7 Sequence Analysis

Sequence Analysis	
How to annotate an alignment - box and shade.	HTML GUI Manual
How to annotate an alignment with text.	HTML GUI Manual
How to display secondary structure in an alignment.	HTML GUI Manual
How to extract sequences from pdb files.	HTML GUI Manual
How to assign secondary structure.	HTML GUI Manual
How to link sequence, alignments, and structures.	HTML GUI Manual
How to save as image, print, and delete sequences and alignments.	HTML GUI Manual

2.2.8 Plotting Tools

Plotting Tools	
Make fully interactive colorful X-Y plots and histograms with up to 4 dimensions.	HTML GUI Manual
Save plot and histogram as image.	HTML GUI Manual

2.3 ActiveICM Reference Guide - Create 3D Molecular Documents for the Web and PowerPoint

This guide is focused on how to make fully interactive 3D documents for Windows PowerPoint and the Web. For more information on the other features in ICM-Browser please see the ICM-Browser User Guide.

Creating 3D Documents Is Straightforward

Creating fully interactive 3D documents for PowerPoint, the web, and standalone browser is straightforward.

1. Download ICM-Browser and the ActiveICM plugin. They are completely free! [video]
2. Open the ICM-Browser and make a series of animated fully-interactive slides showing different colored and rendered views of your molecules. [video]
3. Add hyperlinked HTML text to annotate and link to your slides. [video]
4. Save your file in ICM-Browser and then insert into PowerPoint or the web using the ActiveICM plugin. You can also share your documents in the standalone ICM-Browser. [video - powerpoint] [video -web browser]

```
}; winRef.document.write(str); }
```

2.3.1 Getting Started

Getting Started: Download and Install ICM-Browser and ActiveICM.		
Download ICM-Browser Distribution.	Download	video
Install ICM-Browser Instructions.	Windows Linux Mac	
Download ActiveICM Distribution.	Download	video
Install ActiveICM.	Windows Linux Mac	

2.3.2 How to Create a Series of Fully-Interactive 3D Slides.

Creating Slides How to Create a Series of Fully-Interactive 3D Slides.	video
How to Make Fully Interactive 3D Slides	HTML GUI Manual
How to Animate Slides	HTML GUI Manual
How to View and Navigate Slides in the ICM-Browser.	HTML GUI Manual
How to Edit Slides.	HTML GUI Manual
How to Add Smooth Blending and Transition Effects Between Slides.	HTML GUI Manual

2.3.3 How to Create Molecular Documents

How to Create Molecular Documents: Linking Slides to HTML Text.	video
How to Create an HTML Document.	HTML GUI Manual
How to Edit an HTML Document.	HTML GUI Manual
How to Make a Hyperlink Between HTML Text and a Slide.	HTML GUI Manual

2.3.4 How to Display Molecular Documents in PowerPoint

How to Display Molecular Documents in PowerPoint	video
How to Embed in Microsoft PowerPoint 2003	HTML GUI Manual
How to Embed in Microsoft PowerPoint 2007	HTML GUI Manual
How to Use ActiveICM in PowerPoint	HTML GUI Manual
How to Change ActiveICM Component Properties in PowerPoint	HTML GUI Manual
Advanced use of ActiveICM: Macros to direct visualisation changes.	HTML GUI Manual

2.3.5 How to Display Molecular Documents on the Web

How to Display Molecular Documents in Web Browsers	video
How to Display Molecular Documents in Web Browsers	HTML GUI Manual

2.4 ICM-Chemist Reference Guide

2.4.1 How to Import, Sketch, and Edit Chemicals

How to Import, Sketch, and Edit Chemicals		
How to load the ICM Molecular Editor.	HTML GUI Manual	video
How to sketch chemicals in the ICM Molecular Editor.	HTML GUI Manual	video
How to save a 2D sketch into a chemical spreadsheet.	HTML GUI Manual	video
How to save a 2D sketch in mol format.	HTML GUI Manual	video
How to use SMILES strings to sketch a chemical.	HTML GUI Manual	video
How to load a mol, sdf or mol2 file.	HTML GUI Manual	video
How to extract a 2D sketch of a ligand in complex with a PDB structure.	HTML GUI Manual	video

2.4.2 How to Work with Chemical Spreadsheets

Working with Chemical Spreadsheets		
How to add columns into a chemical spreadsheet.	HTML GUI Manual	video
How to sort a column(s) in a chemical spreadsheet.	HTML GUI Manual	video
How to change the view of a chemical spreadsheet - form, table and grid.	HTML GUI Manual	video
How to copy, cut and paste columns and rows in a chemical spreadsheet.	HTML GUI Manual	video
How to show and hide columns and rows in a chemical spreadsheet.	HTML GUI Manual	video
How to save a chemical spreadsheet in sdf format.	HTML GUI Manual	video

How to export your chemical spreadsheet into Excel.	HTML GUI Manual	video
How to print a chemical spreadsheet.	HTML GUI Manual	video
How to filter columns in a chemical spreadsheet.	HTML GUI Manual	video
How to use find and replace in a chemical spreadsheet.	HTML GUI Manual	video
How to mark and label rows in a chemical spreadsheet.	HTML GUI Manual	video
How to insert hyperlinks to the PDB, PubMed, and Uniprot.	HTML GUI Manual	video
How to copy and paste 2D chemicals.	HTML GUI Manual	video
How to edit data inside a chemical spreadsheet.	HTML GUI Manual	video
How to remove salts, explicit hydrogens and standardize chemical groups.	HTML GUI Manual	video
How to calculate chemical properties in a chemical spreadsheet.	HTML GUI Manual	video
How to identify duplicate chemicals in a chemical spreadsheet.	HTML GUI Manual	video
How to compare two chemical spreadsheets.	HTML GUI Manual	video
How to merge two chemical spreadsheets.	HTML GUI Manual	video

2.4.3 How to Undertake a Chemical Search

How to Perform Chemical Searching		
How to setup a chemical search.	HTML GUI Manual	video
How to draw a chemical search query.	HTML GUI Manual	video
How to add conditions to your chemical search.	HTML GUI Manual	video
How to search chemical spreadsheets, local databases and MolCart.	HTML GUI Manual	video
How to send a chemical search query.	HTML GUI Manual	video

2.4.4 How to Work with Pharmacophores

How to Work with Pharmacophores		
How to draw 2D pharmacophore.	HTML GUI Manual	video
How to search a 2D pharmacophore.	HTML GUI Manual	video
How to draw 3D pharmacophore.	HTML GUI Manual	video
How to extract a 3D pharmacophore from a ligand.	HTML GUI Manual	video
How to edit a 3D pharmacophore.	HTML GUI Manual	video
How to send a 3D pharmacophore search query.	HTML GUI Manual	video
How to color a 2D chemical sketch by pharmacophore feature.	HTML GUI Manual	video

2.4.5 How to Perform Chemical Clustering

How to Perform Chemical Clustering		
How to perform chemical clustering.	HTML GUI Manual	video
How to select representative centers from a tree.	HTML GUI Manual	video
How to reorder branches and change the distance of trees.	HTML GUI Manual	video
How to edit the tree - labels, spacing and coloring.	HTML GUI Manual	video

2.4.6 How to Generate Stereoisomers and Tautomers

How to Generate Stereoisomers and Tautomers		
How to generate stereoisomers.	HTML GUI Manual	video
How to generate tautomers.	HTML GUI Manual	video

2.4.7 How to Generate Combinatorial Libraries

How to Generate Combinatorial Libraries		
How to enumerate a Markush library.	HTML GUI Manual	video
How to decompose a library based on a Markush structure.	HTML GUI Manual	video
How to create a Markush structure.	HTML GUI Manual	video
How to enumerate a chemical library by reaction.	HTML GUI Manual	video

2.4.8 How to Generate Plots and Histograms

How to Generate Plots and Histograms		
How to make a histogram.	HTML GUI Manual	video
How to make an X-Y scatter plot.	HTML GUI Manual	video

2.5 ICM-Chemist-Pro Reference Guide

2.5.1 How to use the ICM 3D Ligand Editor

ICM Chemist-Pro Contains All the Tools in ICM-Chemist	ICM-Chemist Reference Guide	
ICM 3D Ligand Editor: Setup		
How to setup the ligand in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to setup the receptor in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to change the 3D Ligand Editor preferences.	HTML GUI Manual	video
ICM 3D Ligand Editor: Display		
How to display and undisplay the ligand surface representation in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to display hydrogen bonds in the ICM 3D ligand editor.	HTML GUI Manual	video
How to display energy atomic circles in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to display and undisplay hydrogen atoms in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to display unsatisfied hydrogen bonds in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to center on a ligand in the ICM 3D Ligand Editor.	HTML GUI Manual	video
ICM 3D Ligand Editor: Edit Ligand		
How to begin editing your ligand in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to undo and redo changes in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to add and sample new substituents to your ligand in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to sample more than one substituent at a time in the ICM 3D Ligand Editor.	HTML GUI Manual	video

How to edit the ligand in 2D in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to evaluate the SCORE and ligand strain.	HTML GUI Manual	video
How to add an edited ligand to a chemical spreadsheet (table).	HTML GUI Manual	video
How to change the size of the ligand binding pocket - change purple box size.	HTML GUI Manual	video
ICM 3D Ligand Editor: Docking and Minimization		
How to perform ligand minimization in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to re-dock a ligand in the ICM 3D Ligand Editor.	HTML GUI Manual	video
How to restrain (tether) atoms during docking.	HTML GUI Manual	video
How to screen databases of chemical substituents.	HTML GUI Manual	video
How to sample linkers between two chemical fragments.	HTML GUI Manual	video

2.5.2 How to Convert Chemicals to 3D

Convert Chemicals to 3D		
How to convert 2D chemical sketches to 3D.	HTML GUI Manual	video
How to convert 3D chemicals to 2D.	HTML GUI Manual	video
How to generate 3D ligand conformers.	HTML GUI Manual	video

2.5.3 Ligand Energetics

Ligand Energetics		
How to calculate conformation entropy.	HTML GUI Manual	video
How to calculate conformational strain.	HTML GUI Manual	video

2.5.4 How to Superimpose Chemicals

How to Superimpose Chemicals		
How to Perform Rigid Chemical Substructure Superposition.	HTML GUI Manual	video
How to Perform Flexible Chemical Substructure Superposition.	HTML GUI Manual	video
How to use Atomic Property Fields for Chemical Superposition	HTML GUI Manual	video

2.5.5 How to Perform QSAR

QSAR		
How to build a QSAR prediction model.	HTML GUI Manual	video
How to apply a QSAR prediction model.	HTML GUI Manual	video

2.6 Menu Option Guide

Note: Click **Next** (top right hand corner) to navigate through this chapter. Headings are listed on the left hand side (web version) or by clicking the **Contents** button on the left-hand-side of the help window in the graphical user interface.

Here we describe all the options in the drop down graphical user interface menus.



2.6.1 File Menu

2.6.1.1 New

This option allows you to create new peptides, dna, sequences etc... and is described in the [Create New Objects](#) chapter.

2.6.1.2 Open

This option is described in the [Open](#) and [Read](#) chapter.

2.6.1.3 Open with Password

This option is described in the [Open](#) and [Read](#) chapter.

2.6.1.4 Extract from ICB

This option is described in the [Open](#) and [Read](#) chapter.

2.6.1.5 Convert to Local Database

Please see the [Local Databases](#) chapter for more information about this option.

2.6.1.6 Load

Options contained within the menu [File/Load](#)

PDB - read PDB from FTP, http, and local PDB

From Multiple Object File - A multiple object file will have a file extension *.ob and you can select which member of the multiple object is displayed.

PFam Alignment - PFam is a collection of multiple sequence alignments - enter FASTA ID

SwissProt - Download SwissProt sequence.

All Images from Dir - Read into ICM multiple image files png or jpg.

Electron Density Map - Download electron density map from Uppsala electron density server
<http://eds.bmc.uu.se/eds/>

3D Mesh in KMZ or COLLADA Format from Google - see
<http://sketchup.google.com/3dwarehouse/> to download KMZ or COLLADA.

2.6.1.7 Save Project

This is described in the `Save File` chapter.

2.6.1.8 Save Project As

This is described in the `Save File` chapter.

2.6.1.9 Save Project Compatible with ICM 3_5

File/Save Project Compatible with ICM 3_5

Use this option to save a version of your ICM project compatible with an older version of ICM. Version 3.5 or older. If you have an ICM license you can update your version of ICM by visiting our support site at www.molsft.com/support

2.6.1.10 Save with Password

To save a project which is protected by a password:

- File/Save with Password
- Enter a file name or browse for a previously saved project.
- Enter a password
- Determine whether you want the file to be **Fully Protected, read only** or **Read Only and Allow Comments**.

2.6.1.11 Export as ActiveICM Html

To embed in a web browser.

1. Download ActiveICM from here
<http://www.molsoft.com/getbrowser.cgi?product=activeicm&act=list> (it is free!).
2. Create an HTML page in ICM (File/New/Html).
3. Add a series of slides.
4. File/Export As ActiveICM Html..

2.6.1.12 Close Project

To close a project:

File/Close Project

2.6.1.13 Quick Image

See the `High Quality Publication Image` chapter.

2.6.1.14 Write Image

See the `High Quality Publication Image` chapter.

2.6.1.15 Preferences

Preferences are described in the `Preferences` chapter.

2.6.1.16 Recent Files

Recently viewed projects and files can be easily downloaded from the "Recent Files" option. To access this:

- Select `File/Recent Files`.
- Select the desired project by clicking on it once.

2.6.1.17 Recent PDB Codes

Quickly retrieve and display PDB structures that have recently been viewed.

- Select `File/Recent PDB Codes`
- Select desired PDB code by clicking on it once and it will be loaded into the graphical display.

2.6.1.18 Quit

Need to close down ICM - no problem. You do one of the following:

1. Select **File/Quit**. ICM will quit without saving files.
2. Save and Click **X** at the upper right corner of the ICM window.
3. Type `quit` in the terminal window.

NOTE: You may want to save the `icm` session as an `ICM Project` file before quitting.

2.6.2 Edit Menu

2.6.2.1 Delete

This option will delete anything that is selected.

2.6.2.2 Delete All

This option will delete everything e.g. sequences, structures, tables ... Use with care!

2.6.2.3 Select All

This option will select everything e.g. sequences, structures, tables...

2.6.2.4 Search in Workspace

This option allow you to search for a particular text in the workspace

2.6.2.5 Selection

This option allows you to make a precise selection either by neighbors or specifying a particular atom or neighbor. Click on the tabs to jump between selection levels.

2.6.2.6 Invert Selection

This option will select everything that is not currently selected.

2.6.2.7 Clear Selection

This option will remove all selections. For more information on selections see the [Making Selections Chapter](#).

2.6.2.8 Neighbor Selection

This option will allow you to select neighboring atoms. For more information see the [Select Neighbors](#) section in the [Selections Chapter](#).

2.6.2.9 Undo

Due to the complexities of working in an internal coordinates environment not everything can be undone or redone. Certain things like coloring and representations can be undone or redone.

2.6.2.10 Redo

Due to the complexities of working in an internal coordinates environment not everything can be undone or redone. Certain things like coloring and representations can be undone or redone.

2.6.2.11 Restore Recent Backup

ICM periodically makes a backup of your ICM project. If for whatever reason you lose an ICM session and you want to load the backup for the file use:

Edit/Restore Recent Backup

2.6.2.12 PDB Search

See [PDB Search Tab](#)

2.6.2.13 PDB Search by Field

See [PDB Search Tab](#)

2.6.2.14 PDB Search by Identity

See [PDB Search Tab](#)

2.6.2.15 PDB Search by Homology

See [PDB Search Tab](#)

2.6.2.16 PDB Search with External Sequence

See [PDB Search Tab](#)

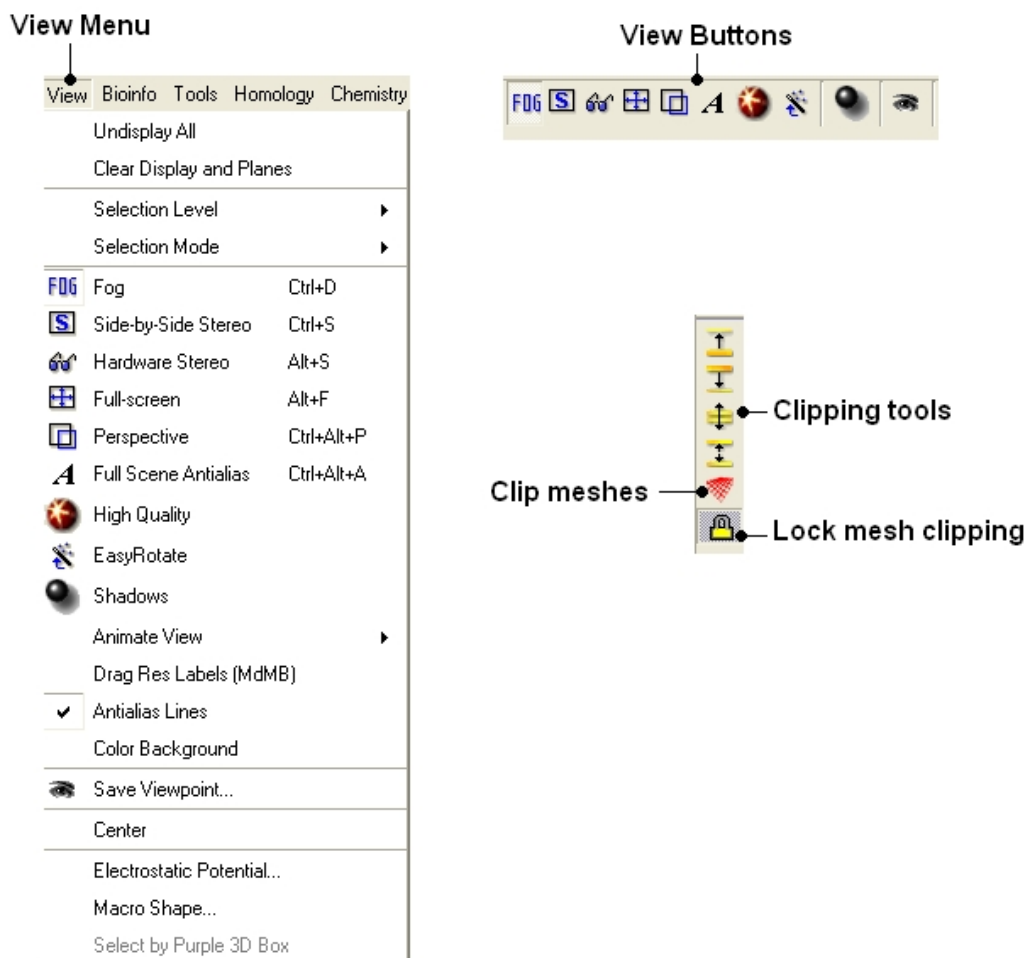
2.6.2.17 Ligand Tools

See the `ligand editor` section of the manual.

2.6.2.18 Ligand Editor Preferences

See the `ligand editor` section of the manual.

2.6.3 View Menu



2.6.3.1 Undisplay All

To undisplay everything currently displayed in the graphical display

- View/Undisplay All

Note For more details on displaying structures please see the GUI Overview chapter.

2.6.3.2 Clear Display & Planes

To clear the display and planes

- View/Clear Display and Planes

NOTE: For more details on planes please see the sections on clipping tools and mesh clipping.

2.6.3.3 Selection Level

There are four levels of selection - atom, residue, molecule and object. For more details on selections please see the `Making Selections` section.

2.6.3.4 Selection Mode

There are four different ways to make selections - new, add, remove and toggle. For more details on selections please see the `Making Selections` section.

2.6.3.5 Fog

Fog Toggle(Ctrl + D) : this feature creates a fog-like environment for your object, so that the part of your structure that is closer appears clear and the distant parts are faded as if they are in fog. The `clipping planes` control the point at which the fog begins.

- View/Fog

2.6.3.6 Side-by-Side Stereo

Side-by-side stereo toggle(Ctrl + S) : this feature allows you to view your structure in 3D form without any 3D goggles.

- View/Side-by-Side Stereo

2.6.3.7 Hardware Stereo

Hardware stereo toggle(Alt + S) - if you have 3D goggles and you wish to view your structure in 3D form, this feature will allow you to do so.

- View/Hardware Stereo

2.6.3.8 Full Screen

Full screen toggleAlt_F - this makes your graphical display fill the entire screen. If you wish to exit this mode, press escape.

- View/Full Screen

2.6.3.9 Perspective

Toggle perspective Ctrl_P this will add perspective to your structure, enhancing depth in the graphical display.

- View/Perspective

2.6.3.10 Full Scene Antialias

Anti-aliasing is the technique of minimizing the distortion artifacts known as aliasing when representing a high-resolution signal at a lower resolution. Always use this option before making high resolution images.

- View/Full Scene Antialias

2.6.3.11 High Quality

Toggle High Quality: this option will give your ICM object better resolution and higher quality. The change in quality is most visible at a high magnification. However, if your object is very large, this feature could slow down your program.

Always use this option before making high resolution images.

- View/High Quality

2.6.3.12 Easy Rotate

Toggle easy rotation: this feature is necessary if your structure is very large or perhaps your computer cannot quickly rotate it. It will prevent your structure from fully loading each time you rotate it, therefore speeding up the process.

- View/Easy Rotate

2.6.3.13 Shadows

- View/Shadows

See `Graphics Effects` chapter.

2.6.3.14 Sketch Accents

To make images as shown below use:

- View/Sketch Accents

See `Graphics Effects` chapter.

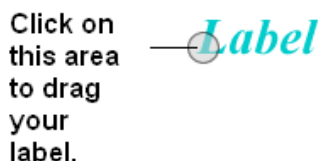
2.6.3.15 Animate View

This tool is described in more detail in the `Molecular Animations and Transitions` section.

2.6.3.16 Drag Res Labels

To change the location of your residue label:

- Select View/Drag res labels.
- If your mouse has a middle mouse button, then click on handle (as shown) of the label you wish to move, and drag it to your desired area.



- If your mouse has no middle mouse button, then click on the Translation icon on the toolbar, and click on the handle (as shown) of the label you wish to move, and drag it to your desired area.

The +/- buttons on the side of the Residue and Atom buttons will shift the label. There are also other **residue label move** options available when you click and hold the residue label button. These options include **Shift to Sidechain Tips**, **Shift to Calphas**, and **Restore Positions**

2.6.3.17 Antialias Lines

Use this option to activate antialias lines. It is recommended to leave this option selected.

- View/Antialias Lines

2.6.3.18 Color Background

To change the background color

- View/Color Background
- Select a color from the panel and press OK.

This option is also in the more convenient `display` tab.

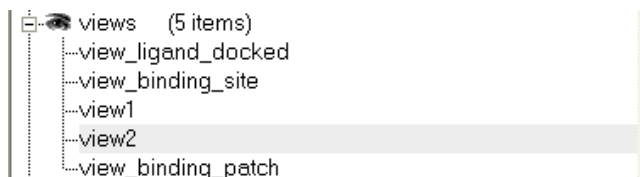
2.6.3.19 Save Viewpoint

It is possible to store a current view using the button shown below.



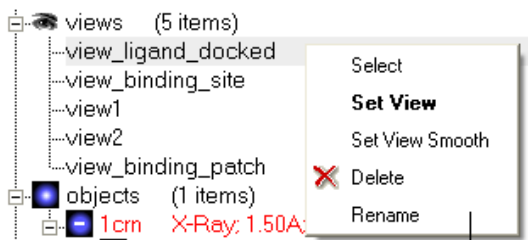
Store Current View

Click on the button and the current view will be stored so that you can view it later. A data entry box will be displayed asking you to name the view. All stored views can be found in the ICM workspace as shown below.



- Double click on the view in the ICM Workspace to display it.

A number of view display options are available by right clicking on the view in the ICM workspace as shown below.



Store current view right click menu

The option in the right click menu called "set view smooth" returns to the view slowly showing the trajectory between the original view and the current one.

2.6.3.20 Center

To center on an object displayed in the graphical display

- Make a selection on the region on which you wish to center on.

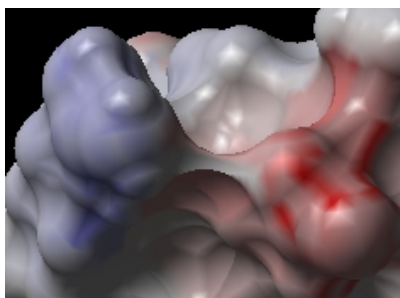
- Tools/Center (or use the `center` button on the right hand-side of the graphical display).

2.6.3.21 Electrostatic potential

This option generates the skin representation of the molecular surface colored according to the electrostatic potential calculated by the REBEL method (hydrogen atoms are ignored). REBEL is a method to solve the Poisson equation for a molecule. REBEL is a powerful implementation of the boundary element method with analytical molecular surface as dielectric boundary. This method is fast (takes seconds for a protein) and accurate. REBEL stands for Rapid Exact-Boundary ELectrostatics. The energy calculated by this method consists of the Coulomb energy and the solvation energy

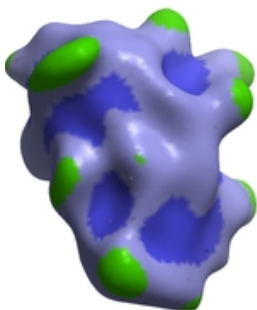
In order to color the skin of your molecule by electrostatic potential:

- Select **View/Electrostatic potential**.
- Enter the potential scale value. This is the local electrostatic potential in kcal/e.u.charge units at which the surface element is colored by extreme red or extreme blue. All higher values will have the same color. This absolute scaling is convenient to develop a feeling of electrostatic properties of molecular surfaces.
- Areas colored blue represent positive areas and red represents negative areas.



2.6.3.22 Macro Shape

A macroshape allows easy viewing and manipulation of a structure. A macroshape representation is ideal for large structures which allows the user to easily identify important regions of the structure and facilitate the return to the 'standard' view of a particular molecule. The level of detail displayed in the macroshape can be controlled by changing the number of harmonics, gridStep, and, contour level.



- View/Macro Shape

2.6.3.23 Select by Purple 3D Box

An alternative way to make a `make-selection{selection}` is to use the purple 3D box. To do this:

- Select the **display** tab and the purple box button
- View/ Select by Purple 3D Box
- The atoms contained within the purple box will be selected.

2.6.4 Bioinfo Menu

The tools in this menu are described in the `Bioinfo Menu` chapter.

2.6.5 Tools Menu - Xray

The options in the menu are described in the `Crystallographic Analysis` chapter.

2.6.6 Tools Menu - 3D Predict

The options in this menu are described in the `3D Predict` chapter.

2.6.7 Tools Menu - Analysis

The options in this menu are described in the `Protein Structure Analysis` chapter.

2.6.8 Tools Menu - Superimpose

The options in this menu are described in the `Protein Superposition` chapter.

2.6.9 Tools Menu - Extras

2.6.9.1 Plot Function

To plot a function:

- Tools/Extras/Plot Function
- Enter the Function(x) eg $\sin(x)$
- Enter the starting value of x (From).
- Enter the end point of x (To).
- Enter the number of points (N points).
- Click OK and your plot will be displayed next to a table of values for your function.

2.6.10 Tools Menu - Table

2.6.10.1 Build Prediction Model

Learn and Predict tools are described [here](#).

2.6.10.2 Predict

Learn and Predict tools are described [here](#).

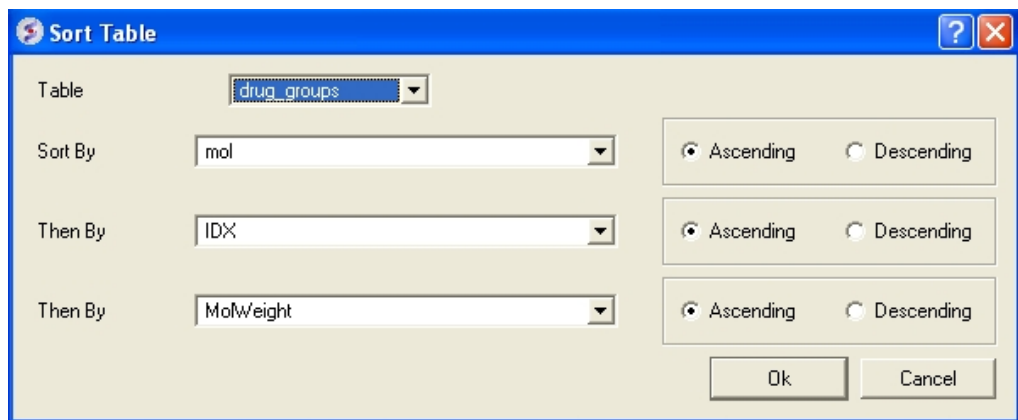
2.6.10.3 Cluster Set

This is described in the `cluster` section of the `Working with Tables` Chapter.

2.6.10.4 Sort Table

There are a couple of ways to sort a chemical table. You can right click on the a column header and select `sort` or you can use the option in the menu `Chemistry/Sort Table`.

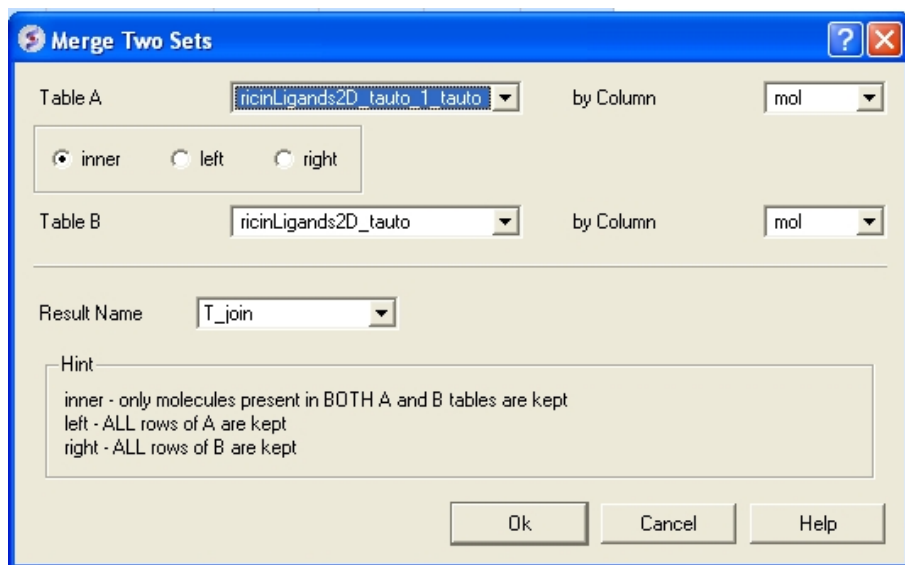
- Read a chemical table into ICM.
- Select the columns by which you wish to sort by as shown below.
- Select **Ascending** or **Descending** and for each **sort by** option and then click OK



2.6.10.5 Merge Two Sets

To merge two tables:

- Read the two tables into ICM.
- Tools/Table/Merge Two Sets
- Select the first table from the drop down list (Table A) and the column you wish to use to merge the table by.
- Select merge method 1. **inner** - only molecules present in BOTH A and B tables are kept; or 2. **left** ALL rows of A are kept ; or 3. **right** ALL rows of B are kept.
- Select the second table from the drop down list (Table B) and the column you wish to use to merge the table by.
- Enter a name for the output table.
- Click OK and a new table will be displayed.



2.6.10.6 Add External Columns

To add external columns to a table:

- Read at least two tables into ICM - the table you want to add to and the table you want to add the column from.
- Tools/Table/Add External Columns
- Enter the target table name and the column you wish to match each table by.
- Enter the source of the new column (Other table and column name)
- Choose to add "all the columns" from the source or "overwrite matching columns" or select the columns you want to add by selecting the "choose column" option.

2.6.10.7 Append Rows

To append rows from one table to another one:

- Read at least two tables into ICM - the table you want to add to and the table you want to add the column from.
- Tools/Table/Append Rows
- Enter the name of the Target Table (where you will append).
- Enter the name of the Source Table (where you will append from).

2.6.11 Tools Menu - Chemical Search

Chemical searching is described in the Chemistry chapter [here](#).

2.6.12 Tools Menu - Molecular Editor

The molecular editor is described in the Chemistry chapter [here](#).

2.6.13 Homology Menu

The options in this menu are described in the Homology Modeling Chapter.

2.6.14 Chemistry Menu

The tools in the **Chemistry** menu are described [here](#).

2.6.15 Docking Menu

The tools in the **Docking** menu are described in the Docking chapter.

2.6.16 MolMechanics Menu

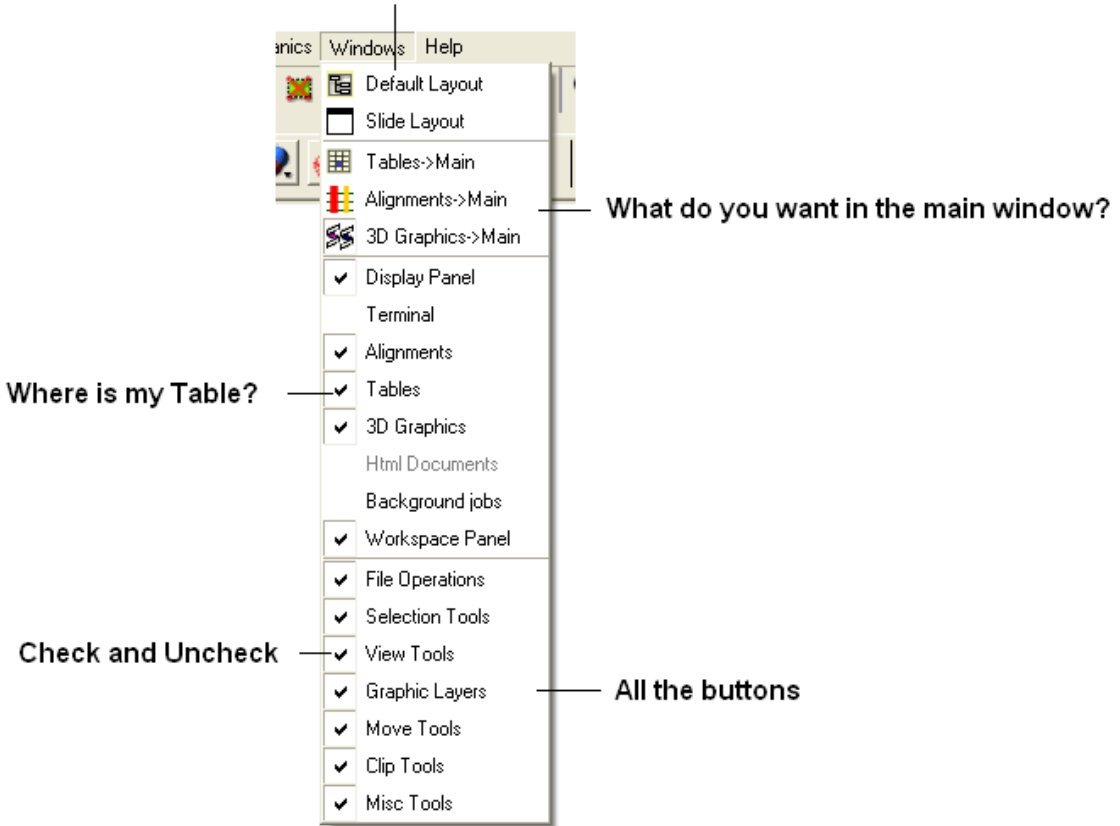
The options in this menu are described in the Molecular Mechanics chapter.

2.6.17 Windows Menu

This menu allows you to choose the windows you wish to display. The windows which open automatically when you first open GUI are shown in the [Getting Started](#) section. Other windows can be displayed by selecting the windows menu. For example, if you have loaded a table but cannot see it in the GUI it may be because the Tables option in the window menu hasnt been selected.

To add or remove windows from the GUI display select the 'window menu'. Other windows not included in the default display such as tables and alignments can be added.

I have windows open everywhere - Please bring some order.



To return to the default display option select the 'Default layout' option in the windows menu.

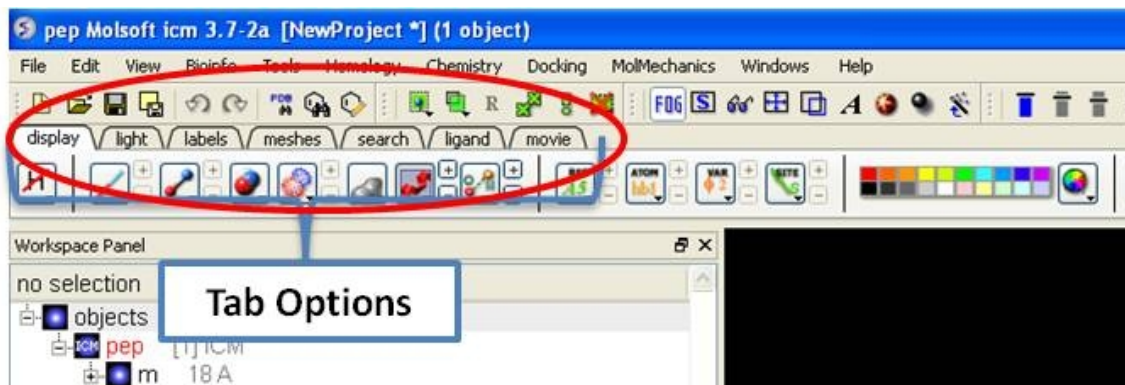
OR

Click the default layout icon.



2.7 Tab Guide

In this section we describe the contents of the tabs in the graphical user interface.

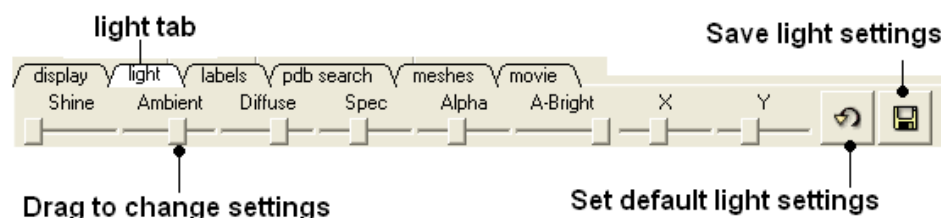


2.7.1 Display Tab

The display tab contains tools for a variety of functions including - structural representations, coloring, labeling and superposition. This tab is shown below.

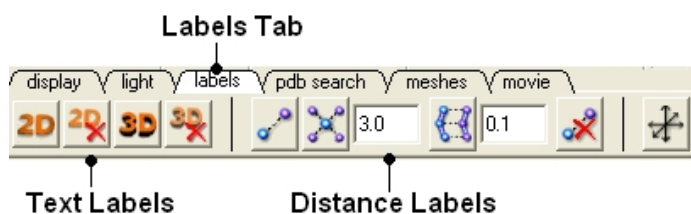


2.7.2 Light Tab



The options in this tab are described in the [Lighting](#) Section.

2.7.3 Labels Tab



The options in this tab are described in the [labels](#) section of this manual.

2.7.4 PDB Search Tab

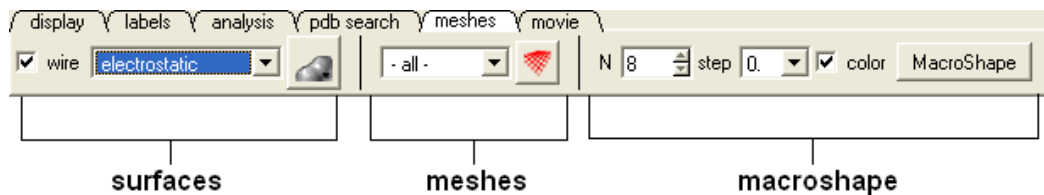
Instructions on how to use this tab can be found in the [Search PDB](#) section.

2.7.5 Ligedit Tab

Instructions on how to use the options in this tab can be found in the [How to use the 3D Interactive Ligand Editor](#) section.

2.7.6 Meshes Tab

Click on the tab button entitled '**meshes**' and three different graphical display tools are available for you to use. The three displays are surface, meshes and macroshape and are collectively referred to as meshes.



The benefits and applications of each display are described in the [section](#).

3 Getting Started

Chapter Contents:

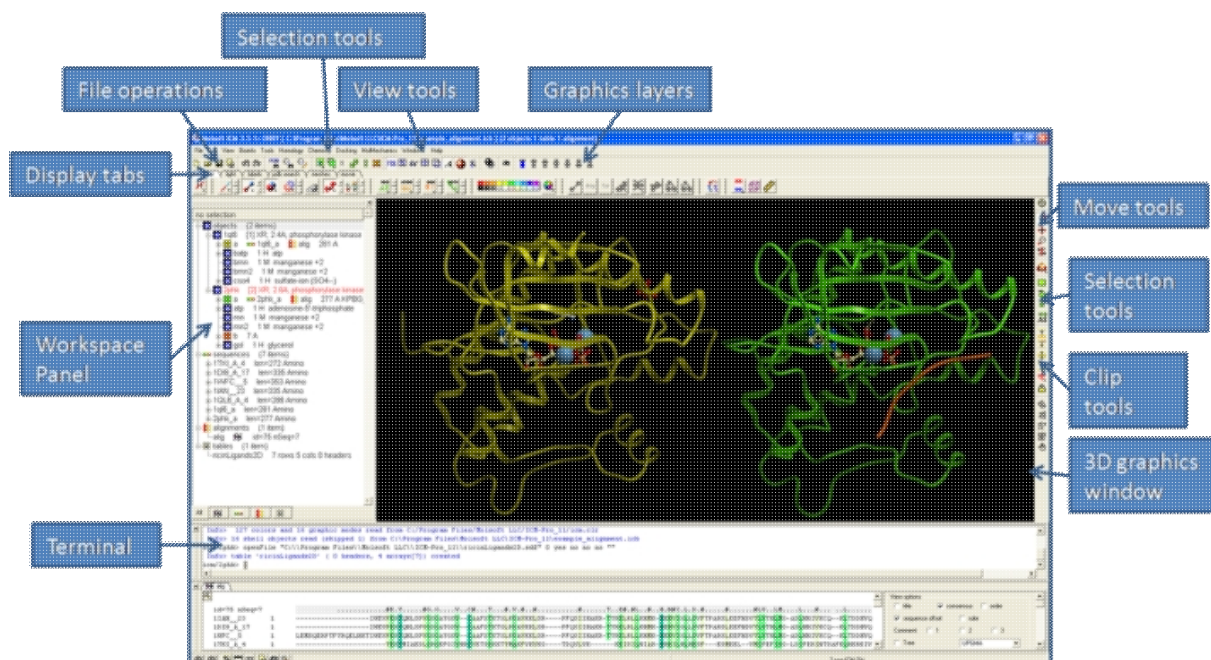
- The Basics of the Graphical User Interface
- How to search and Download Protein Structure, Sequences, and Chemicals
- How to create new Objects
- How to open files
- How to save files
- How to Make Selections
- Changing Preferences

3.1 The Basics of the Graphical User Interface

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro | ICM-Chemist

3.1.1 The Components of the GUI

The **Graphical User Interface** (GUI) has many components. When you first use the GUI the default window layout is displayed as shown below.

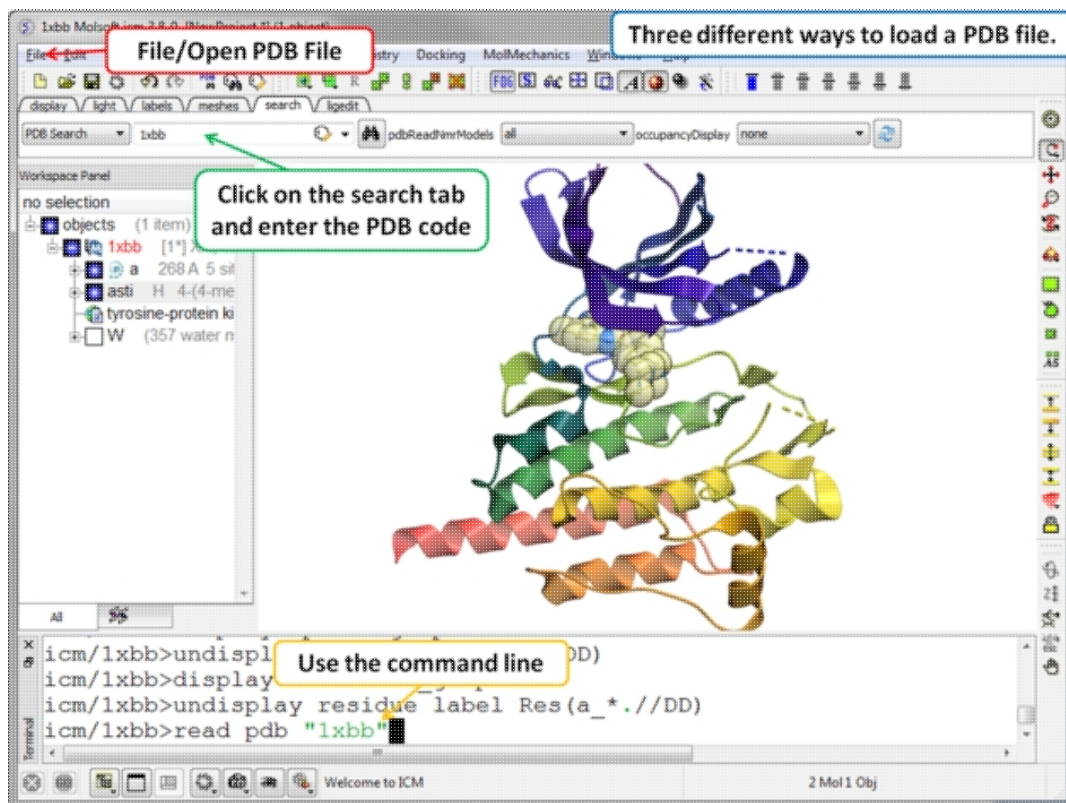


3.1.2 How to load a PDB Structure

There are three main ways to read in a PDB file.

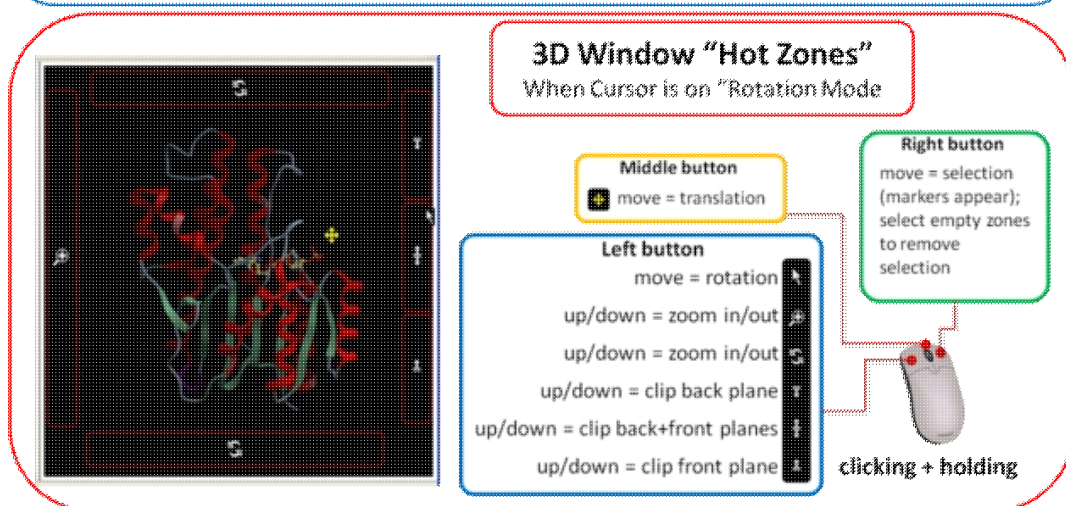
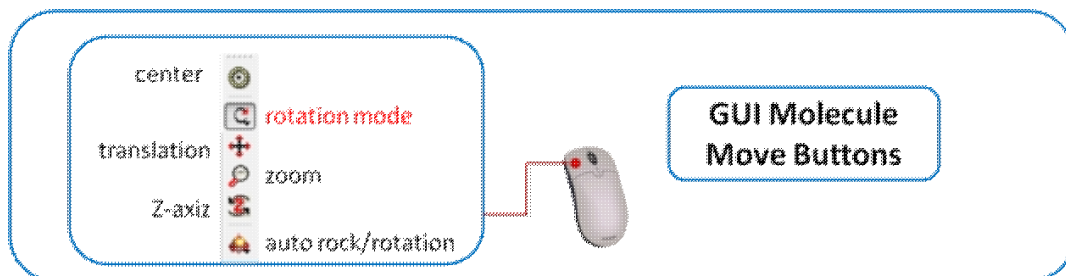
1. Using the command line.
2. Using File/Open button
3. Using the PDB Search tab

Other PDB search options are described in more detail in the PDB Search section of this manual.



3.1.3 How to Move a Structure in the Graphical Display

Available buttons and options for moving molecules around the graphical display window. This is described in more detail in the section entitled *Move Buttons*.



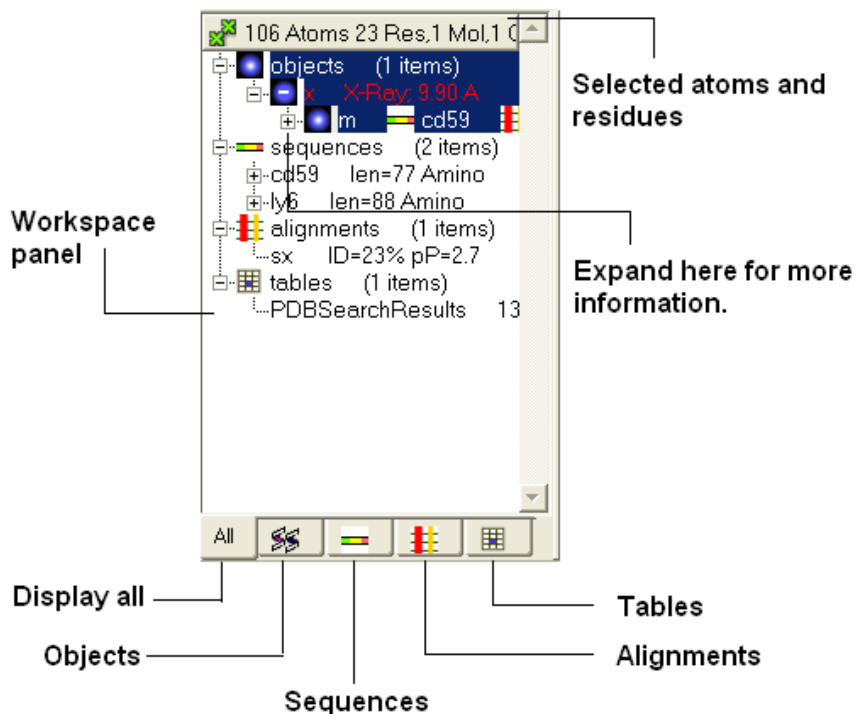
How to use the Graphics window controls

In the graphics window you can use various `tools` described elsewhere but it is helpful to know the following things:

- **Picking a tool:** the left mouse button will function according to the selected tool
- **Popup menus:** right click on an atom gives a pop-up menu
- **Selecting in the rotation mode:** the right mouse button will select atoms
- **Translating in the rotation mode:** the middle mouse button will translate the scene
- **Zooming and moving clipping planes in the rotation mode:** the left, top and right margins of the graphics window are reserved for other actions, zoom, z-rotation, and clipping planes, respectively. That means that even if you are picking atoms, by pressing control you can still rotate your molecule with the left-mouse-button.
- **Rotating in any non-rotation mode:** if you press **Control** in any mouse mode, e.g. zoom, pick etc., it will temporarily switch to **rotation**
- **Escaping from the connect and continuous movement modes:** pressing **Escape** helps to get out of certain modes, such as Full Screen, Continuous rotation or rocking, the Connect mode.
- **Global rotation in the Connected mode:** pressing **Shift** will temporarily switch to the global rotation/translation mode.

3.1.4 How to Display Molecules using the ICM Workspace Panel

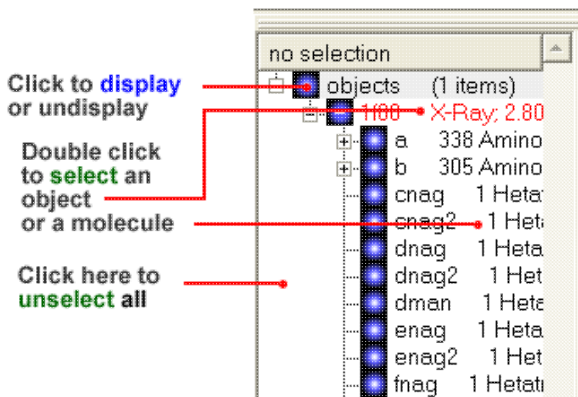
The workspace panel (located on the left hand side panel of the gui) is an important place within the graphical user interface because it displays which sequences, structures, objects, tables and alignments are currently loaded into ICM. From the ICM Workspace panel you can make graphical selections and display and undisplay molecules.



Once a structure has been loaded into ICM the individual components of that structure (i.e. amino acids, metal ions, binding sites etc) are listed in the ICM workspace.

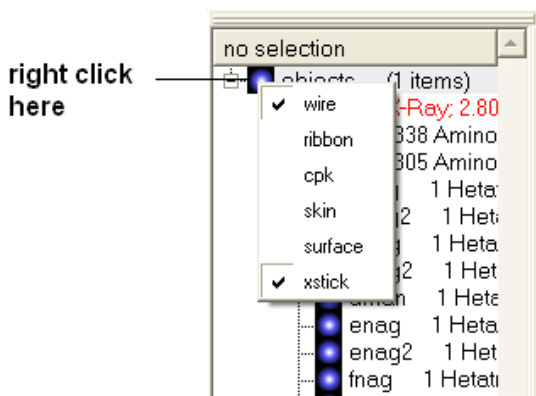
To display every component of the object except for binding sites and water atoms:

- Click on the white box next to the word object at the top of the ICM workspace. This box will be colored blue once the structure is displayed



To display the whole structure in wire, ribbon, cpk, skin, surface and xstick representations:

- Right click on the blue box next to the word object. A menu will be displayed.



- Select which representation you desire for your structure by clicking on the appropriate word. A check mark indicates the representation currently displayed. To un-display a particular representation click on the word again.

In order to clear your graphical display:

- Select **View/Undisplay All**

If you only wish to display part of the structure then click in the boxes further down the tree in the ICM workspace.

To display selected regions of the molecule in wire, ribbon, cpk, skin, surface and xstick representations:

- Right click on the appropriate box in the **ICM** workspace. A menu will be displayed and select the representation you would like to use (e.g. wire, ribbon etc...)

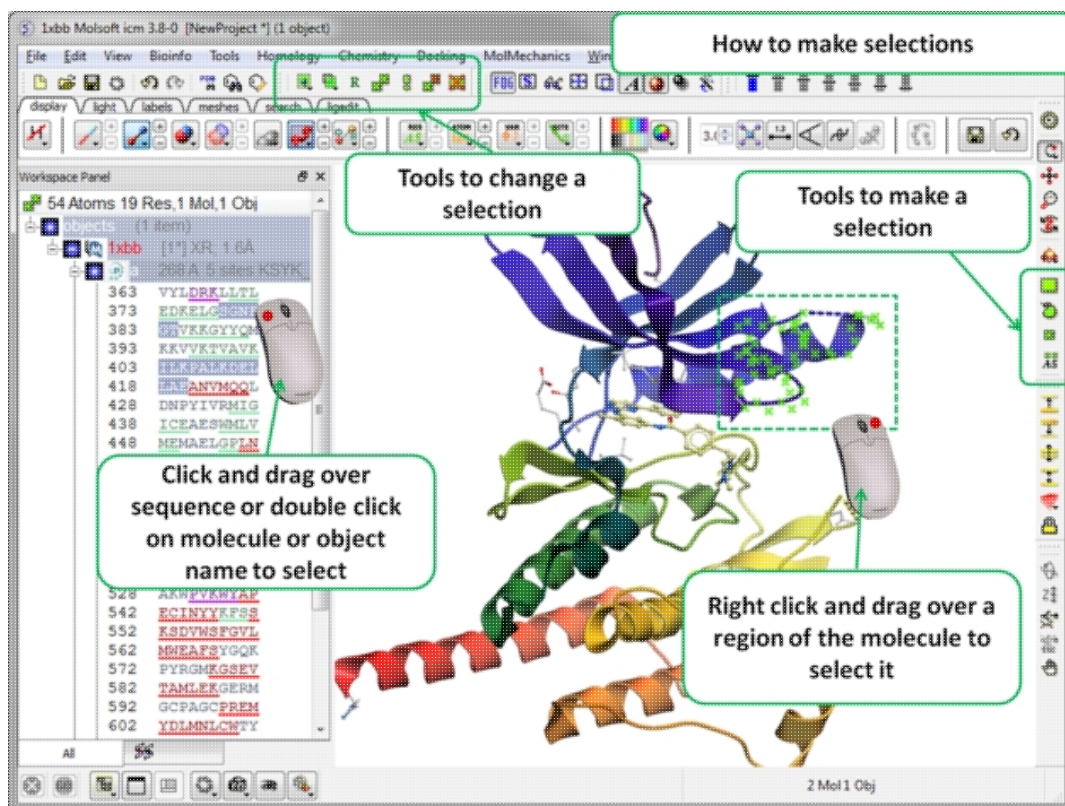
3.1.5 How to Make Selections

Making selections in ICM is an important skill to master (e.g. you may want to select a binding pocket for docking or a region of a molecule for coloring). The four levels of selection are:

1. Atoms
2. Residues
3. Molecules
4. Objects (multiple molecules comprising a PDB entry)

There are several ways of making selection in ICM. The simplest is to interact directly with the graphics window - **right-click, hold and drag** around the area of the screen you want to select. Alternatively, in the workspace window, expand the tree of molecules and chains until the relevant protein sequences is displayed. Then left click and drag to mark residues to form a selection.

See the chapter entitled Making Selections for more information.

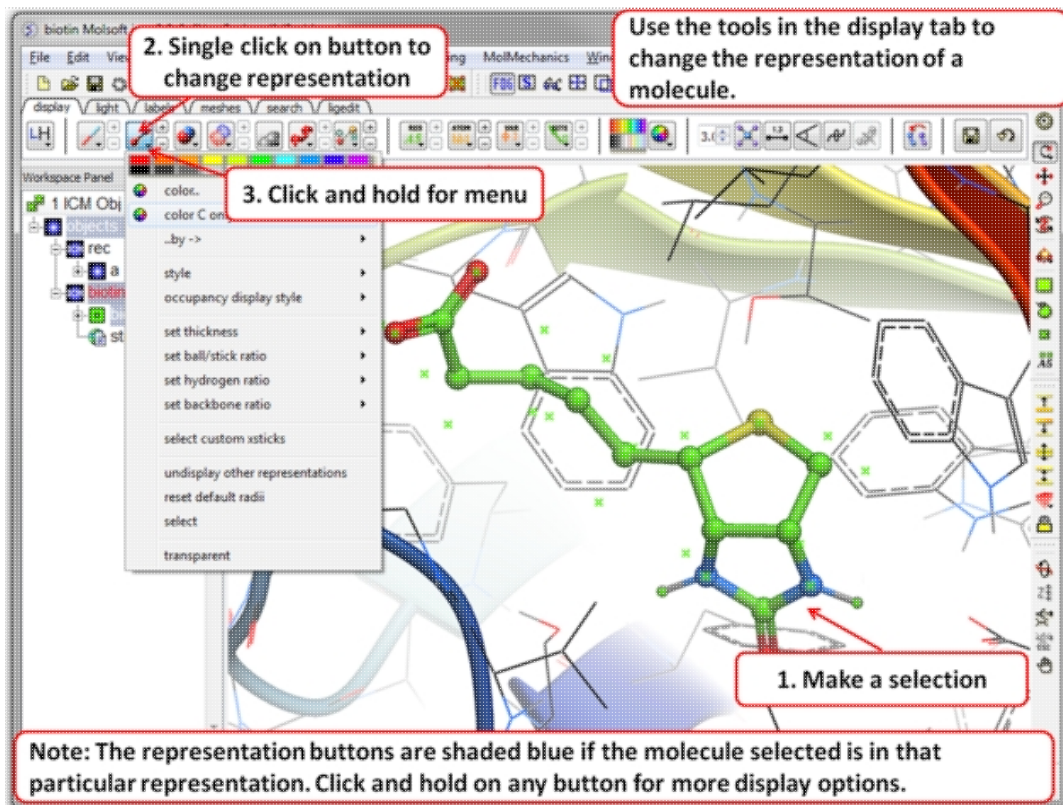


3.1.6 How to Change Protein Representation

To change the representation of the protein, make a selection and then use the tools in the display tab.

There are 6 main types of representation:

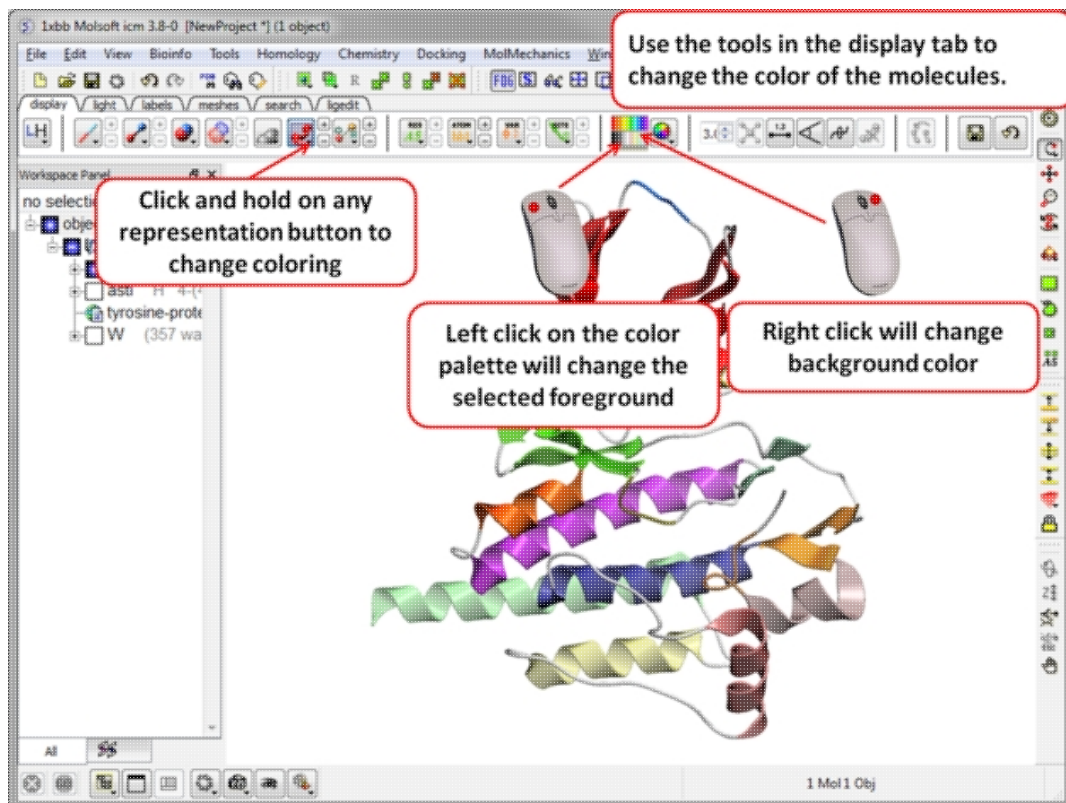
- **Wire**: Wires connecting covalently bound atoms of a molecule. This representation has no defined thickness as such will not make shadows. Useful for showing the chemical structure of a small molecule.
- **Xstick**: Covalent bonds are represented as cylinders whilst atoms are represented as small spheres.
- **CPK**: Atoms are represented as spheres with their respective van der Waals radius and coloured according to a standard defined by Corey, Pauling and Koltun.
- **Surface**: Solvent accessible surface. This is the center of water sphere as a water probe rolls over the molecule.
- **Skin**: A Connolly molecular surface over the selection. This is a smooth envelope touching the van der Waals surface of atoms as a water probe rolls over the molecule.
- **Ribbons**: Cartoon representation of protein and DNA secondary structure. Protein residues marked as alpha-helices ('H') are shown as a flat, helical ribbon, those marked as beta-sheets ('E') are shown as a flat ribbon with an arrow-head, and the rest are shown as a cylindrical "worm". If secondary elements are not defined everything will be shown as a cylindrical worm. ICM can automatically assign secondary structure: Tools/3D predict /Assign Helices and Strands



3.1.7 How to Color

To change the color of the representation you need to use the buttons in the `display` tab.

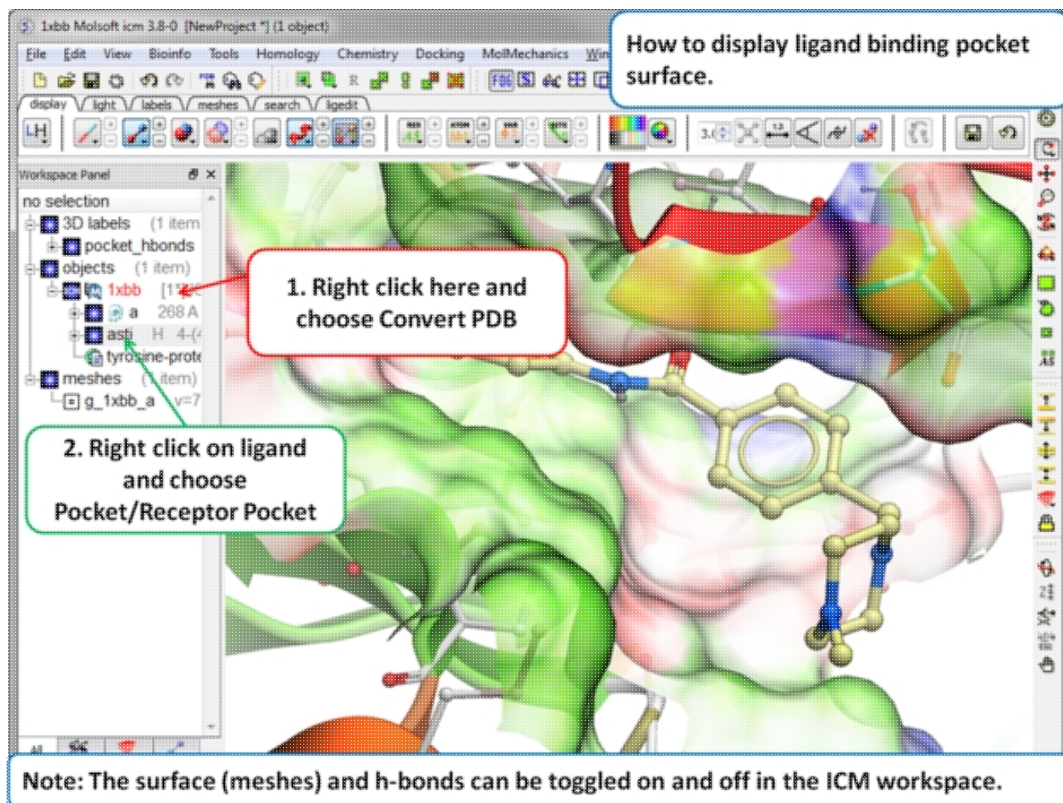
Changing the colour of a representation works in much the same way as displaying the representation itself. The selection rationale is the same followed by clicking on a colour in the palette in the `display` tab. It is also possible to colour different representations of the same selection independently by clicking and holding on the representation buttons in the `display` tab.



3.1.8 How to Display a Binding Pocket Surface

To display the surface of a small molecule ligand or peptide binding pocket:

- Load the PDB of interest.
- Convert PDB to ICM object. If you do not convert you will not get the properties of the pocket displayed on the surface.
- Right click on the small molecule or peptide in the ICM Workspace and select **Ligand Pocket**.

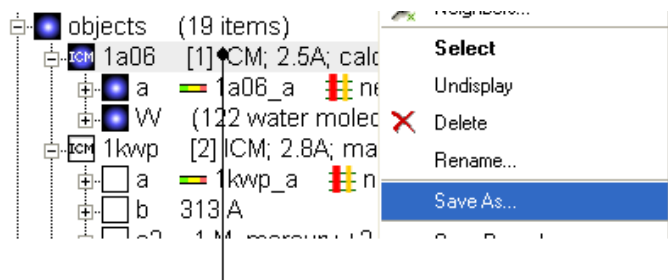


3.1.9 How to Save an ICM Object

Any ICM object such as a structure, sequence, or alignment, can be saved for use at a later time.

To save an object:

- Right click on the object name in the ICM workspace or ICM alignment editor and a menu will be displayed.
- Click on the **Save As...** option.
- Enter the unique name you wish to call your object in the box labeled **File name:**
- Choose which folder or directory you wish to save your object by clicking scrolling down in the box labeled **Save in:**
- Choose which file type you would like to save your object as by scrolling down in the box labeled **Save as type** . ICM structure objects should have the file ending yourfilename.ob and alignments yourfilename.ali
- Once the appropriate information has been entered click on the **Save** button in the bottom right hand section of the window.
- The object is now saved.



To save an ICM object or PDB file right click and select SaveAs..

3.1.10 How to Save an ICM Project File

All objects contained within an ICM session can be saved in a single file with the extension .icb. The file can then be read into ICM and the exact layout of the file will be preserved. To save a project file go to the **File** menu and select **Save Project**.

3.1.11 How to Drag and Drop

NOTE: "Drag and Drop" is a useful way of moving objects and sequences around the graphical user interface.

Sequences and objects can be moved around the graphical user interface by dragging and dropping them. All loaded sequences and objects are always displayed in the workspace panel. Select the desired object or sequence from the workspace panel by clicking and holding, move the selection to the desired location and release.

This is a useful application in the graphical user interface. For example, you may have an alignment displayed and you wish to add another sequence to the alignment. This can simply be accomplished by dragging a loaded sequence from the workspace panel into the alignment display panel. Or, you can quickly view an object by dragging and dropping it from the workspace panel into the 3D graphics window.

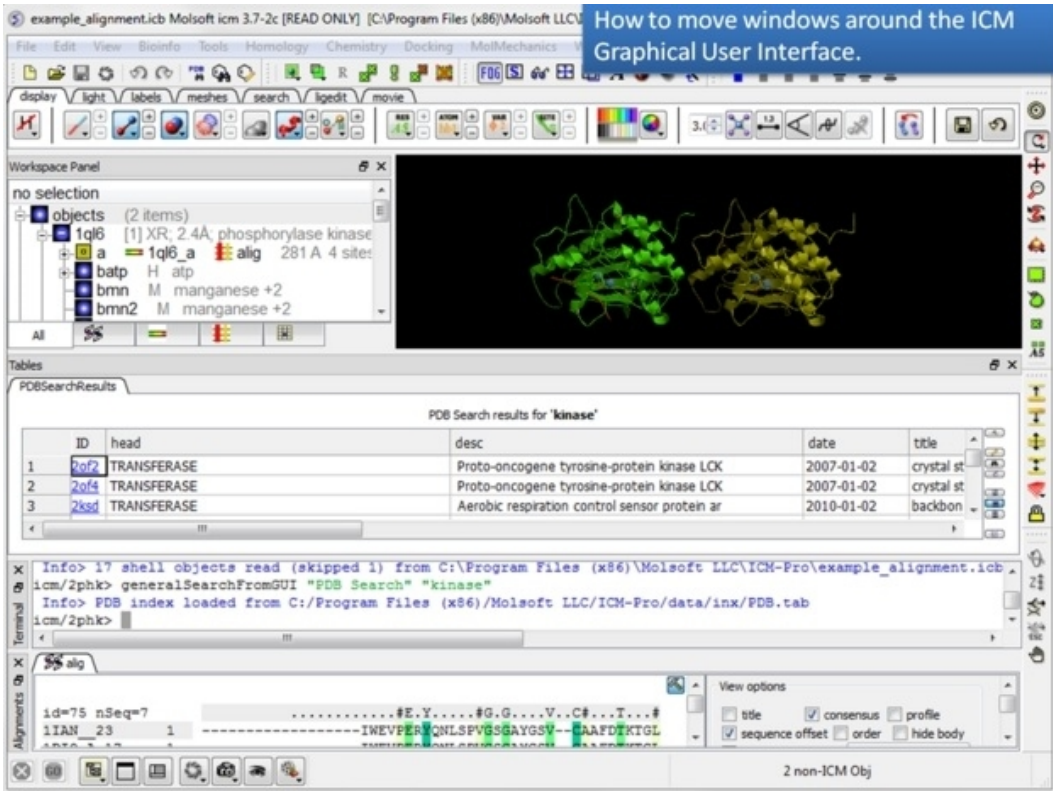
3.1.12 How to: Right Click Options

NOTE: If you right click on any object you will see a new menu of options related to that object.

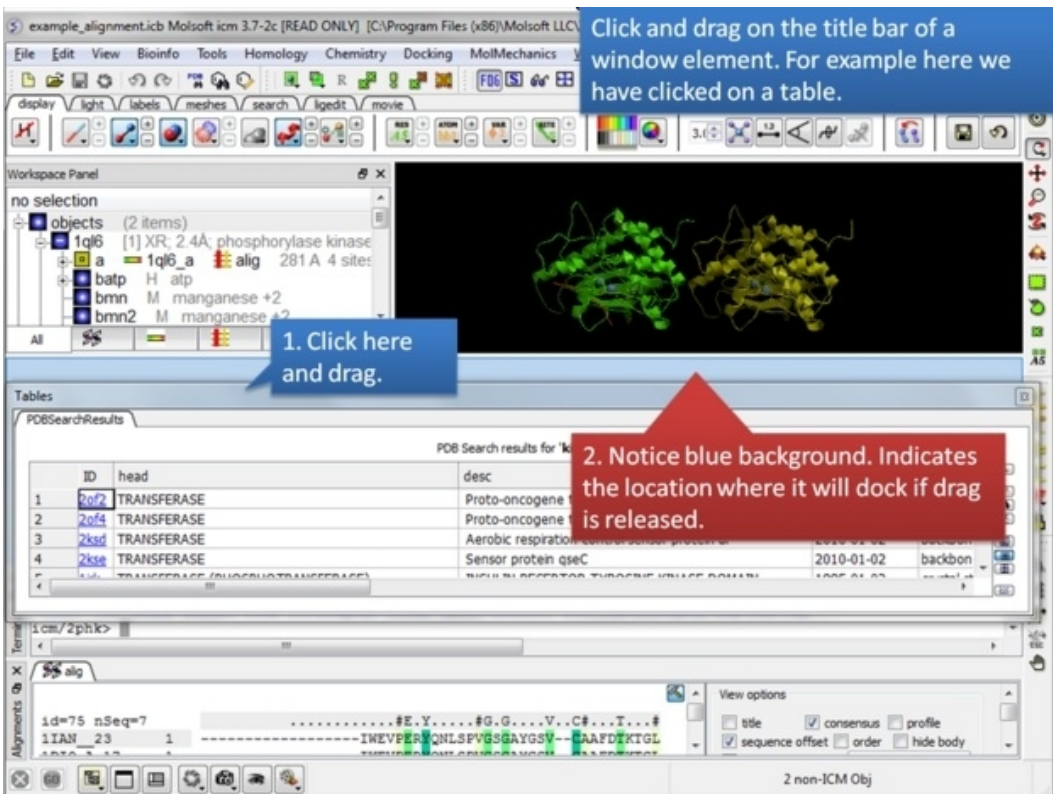
The right click mouse option can be used throughout the graphical user interface. It is a very useful means of opening up a whole new world of menus and options. Most of these options are described in this book. However, when using the graphical user interface it is always a good idea to try right clicking the mouse on an object and seeing which extra options that are available for you to use.

3.1.13 How to Move Windows

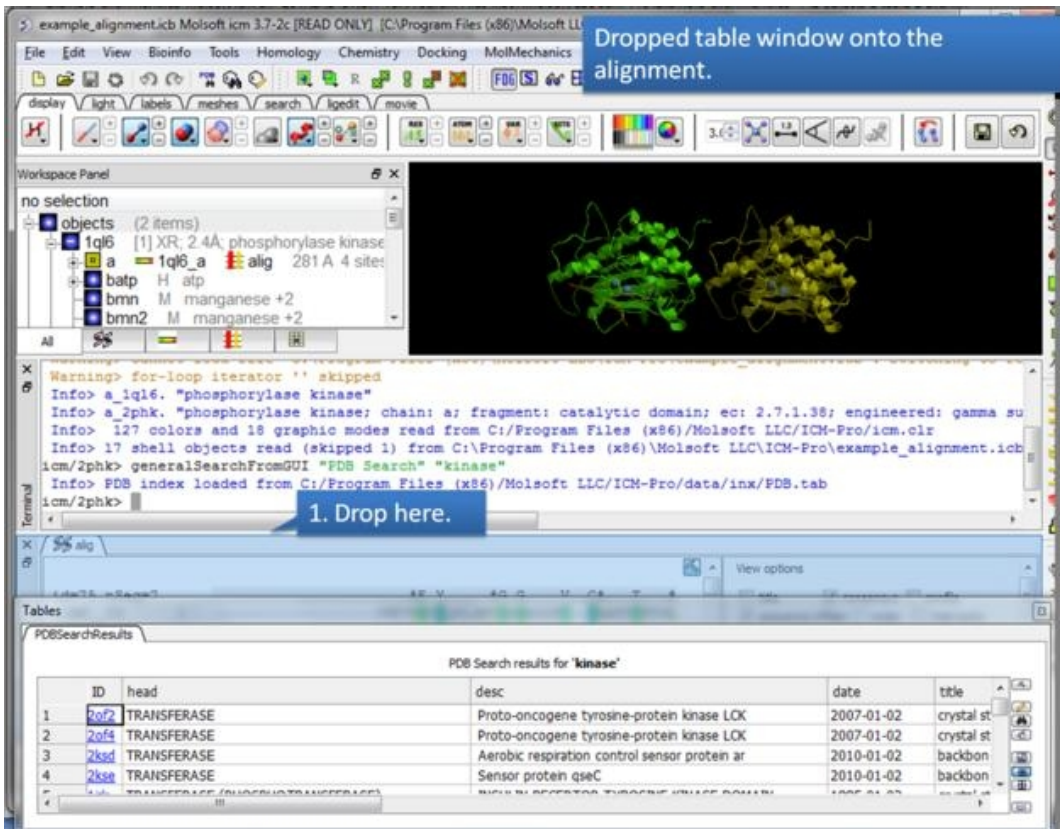
Is your graphical user interface looking a bit messy? Do you have tables, alignments, plots all over the place? Here we show you how to arrange everything to make a clearer GUI environment.



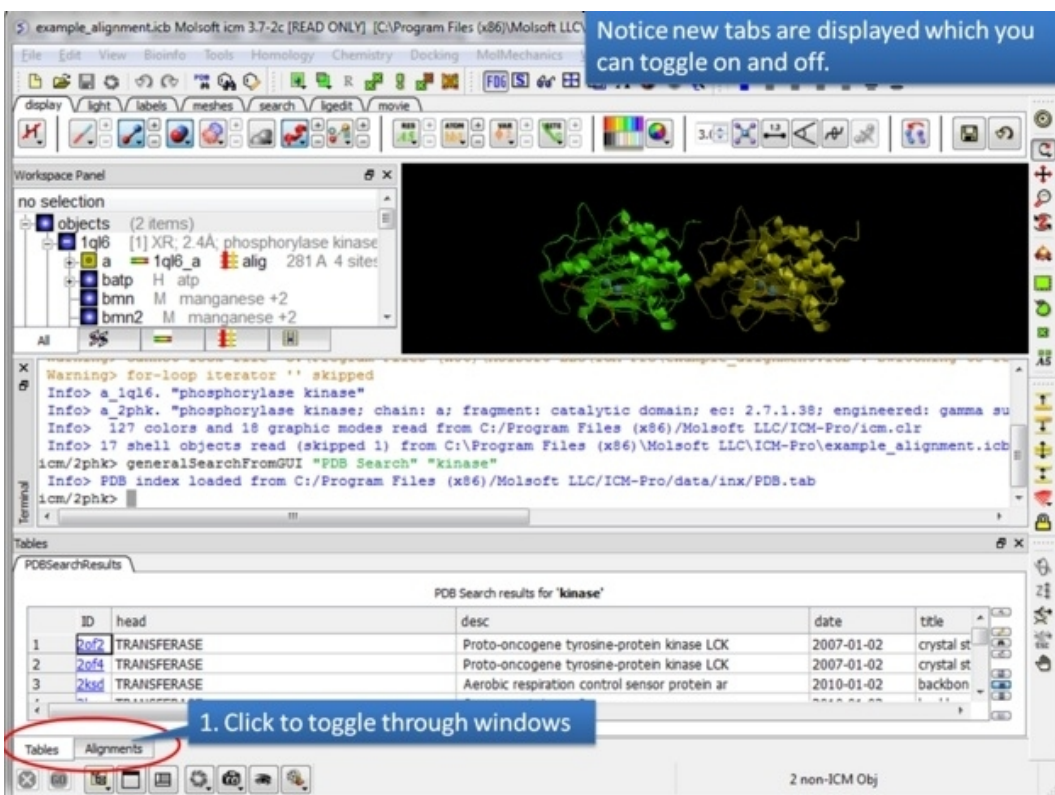
1. Too many windows in your GUI? You can move them around to make viewing easier.



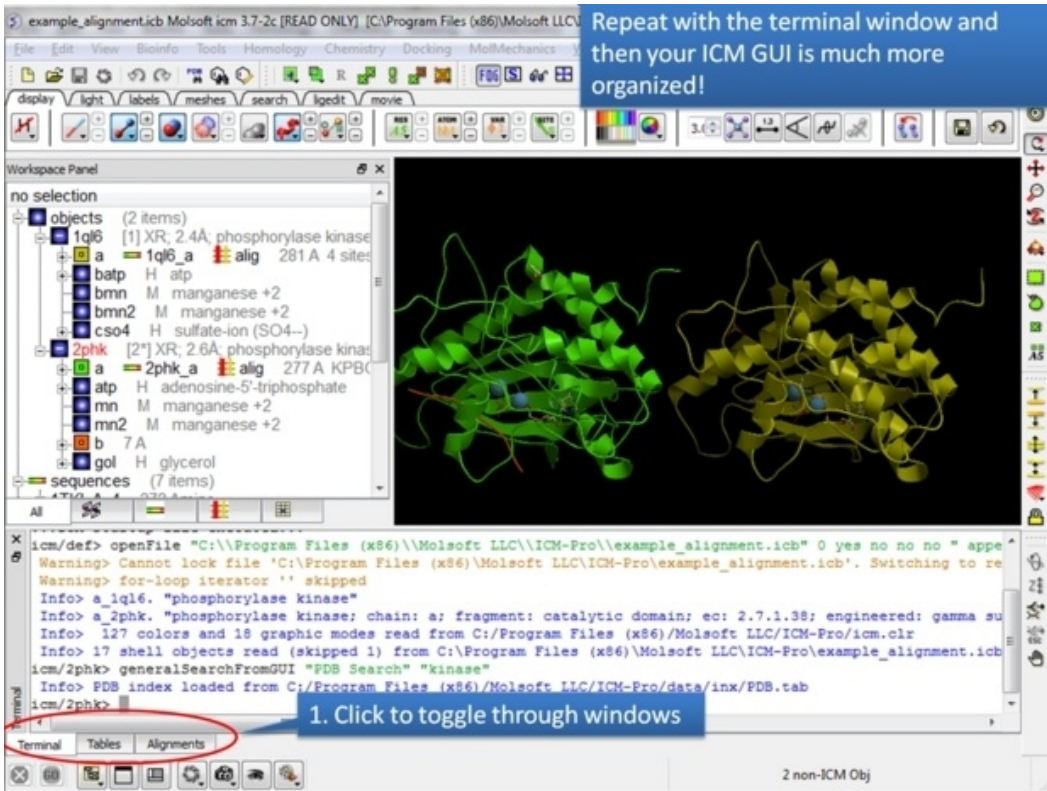
2. Click and drag on a window title bar to move it.



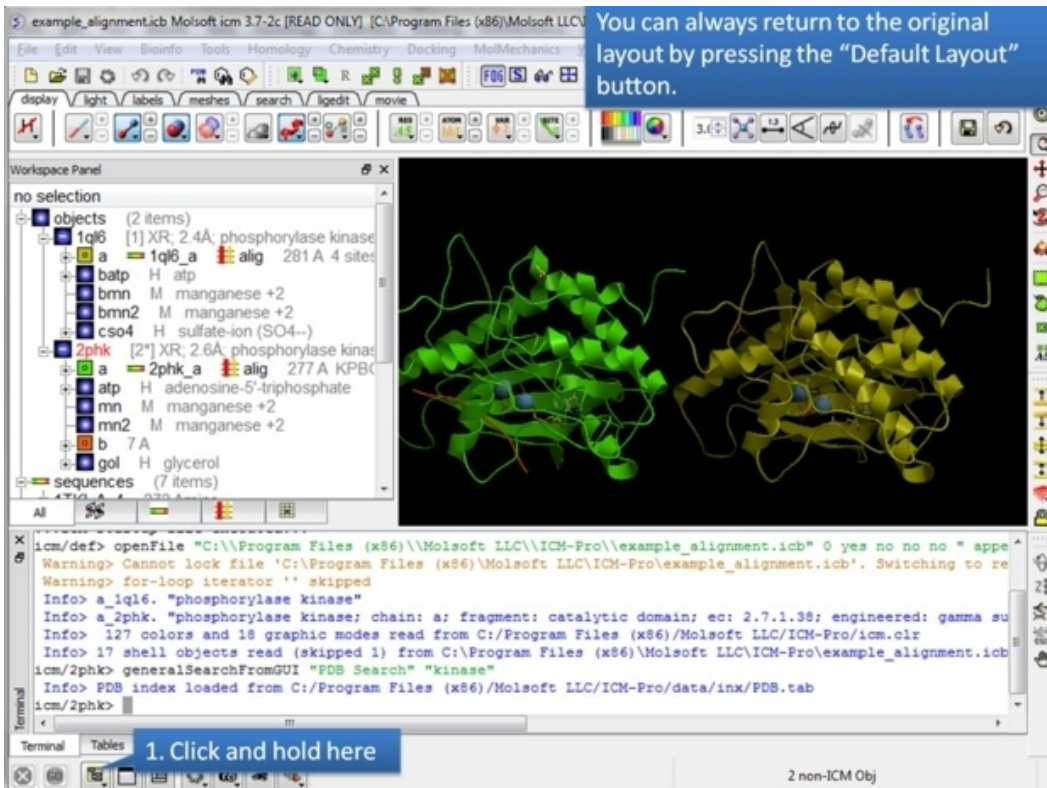
3. Drop one window on another one to dock it.



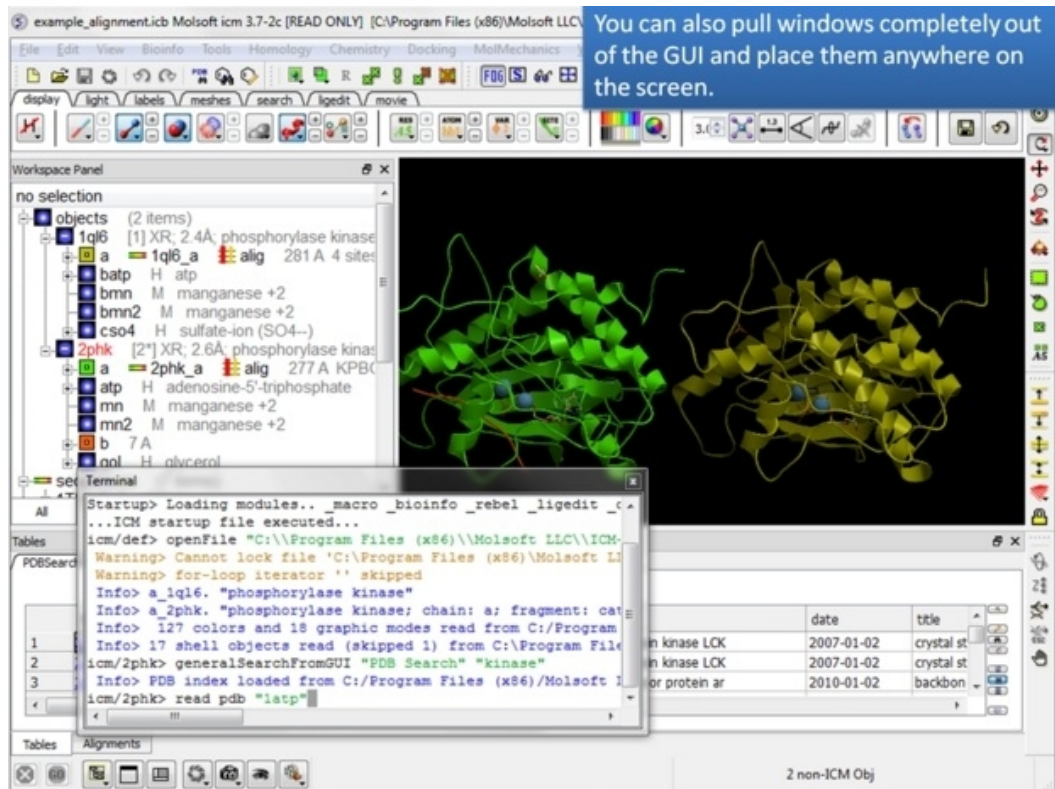
4. You can access each window by clicking on the tabs.



5. In this example we have the table, alignment, and terminal all in the same panel



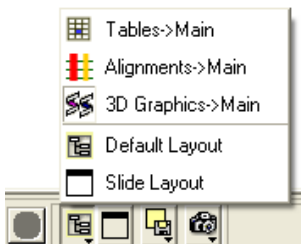
6. There is a quick way to return the windows to their original location.



7. You can also move a window to any part of your screen.

NOTE: To return to the default display option select the 'Default layout' option in the windows menu.
OR

Click the default layout icon.



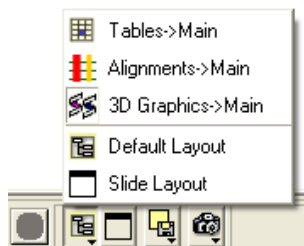
Some options can be accessed at the bottom of the GUI.

OR

Double click on the window header.

3.1.13.1 How to Arrange Windows

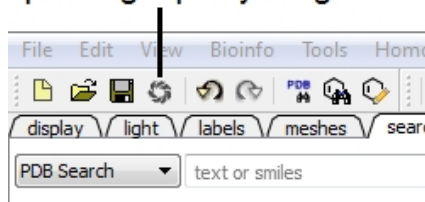
Sometimes when using ICM you may have many items displayed such as structures, alignments and tables. As a default the graphical display is the largest and centered in the middle of the ICM graphical user interface. However if you wish to work on an alignment or table you can place the alignment or table as the main display by clicking on the buttons shown below. The larger display generally makes it easier to manipulate the alignment or table. There are ofcourse other ways to alter the layout such as `tier` the windows but this option is just a simple click and can sometimes come in useful.



3.1.14 How to Make a Picture

There are several ways of taking a picture of the contents of the 3D graphical display window see the write image section. However the easiest way is to simply click on the button in the view tools panel (see image below).

quick high quality image



Or select /File/Quick Image

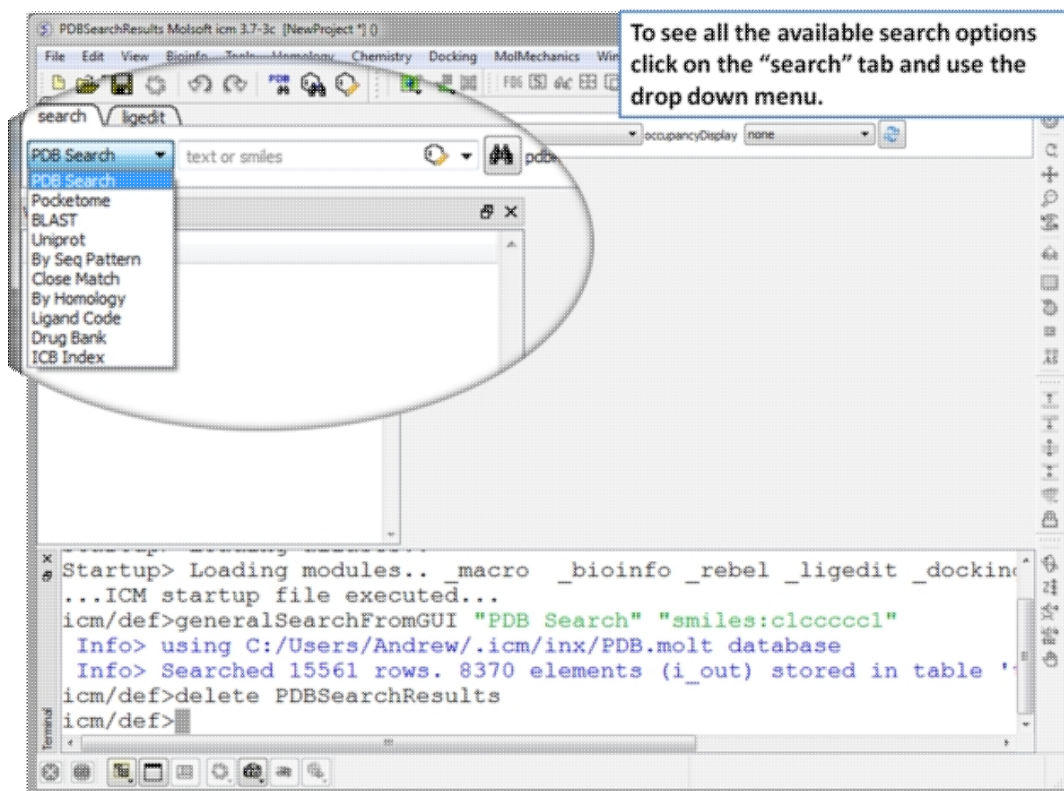
The picture will be automatically saved as a PNG file in the directory from which you loaded ICM. The default picture name is icm[n].png, where n is the number of pictures taken in one ICM session. To save in other picture formats and to change the file name see the write image section.

3.2 How to search and download Protein Structure, Sequences, and Chemicals

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro | ICM-Chemist

The **Search** tab allows you to search and download content from the following databases:

- The Protein Databank www.wwpdb.org - search for protein and chemical
- The Pocketome Database <http://pocketome.org/>
- BLAST search the NCBI sequence database
- The UniProt Sequence database <http://www.uniprot.org/>
- Search the PDB by ligand code.
- Search Drug Bank <http://www.drugbank.ca/>
- Search PubChem <https://pubchem.ncbi.nlm.nih.gov/>
- Search ChEMBL <https://www.ebi.ac.uk/chembl/>
- Search SureChEMBL <https://www.surechembl.org/search/>
- Search the Crystallography Open Database <http://www.crystallography.net/>



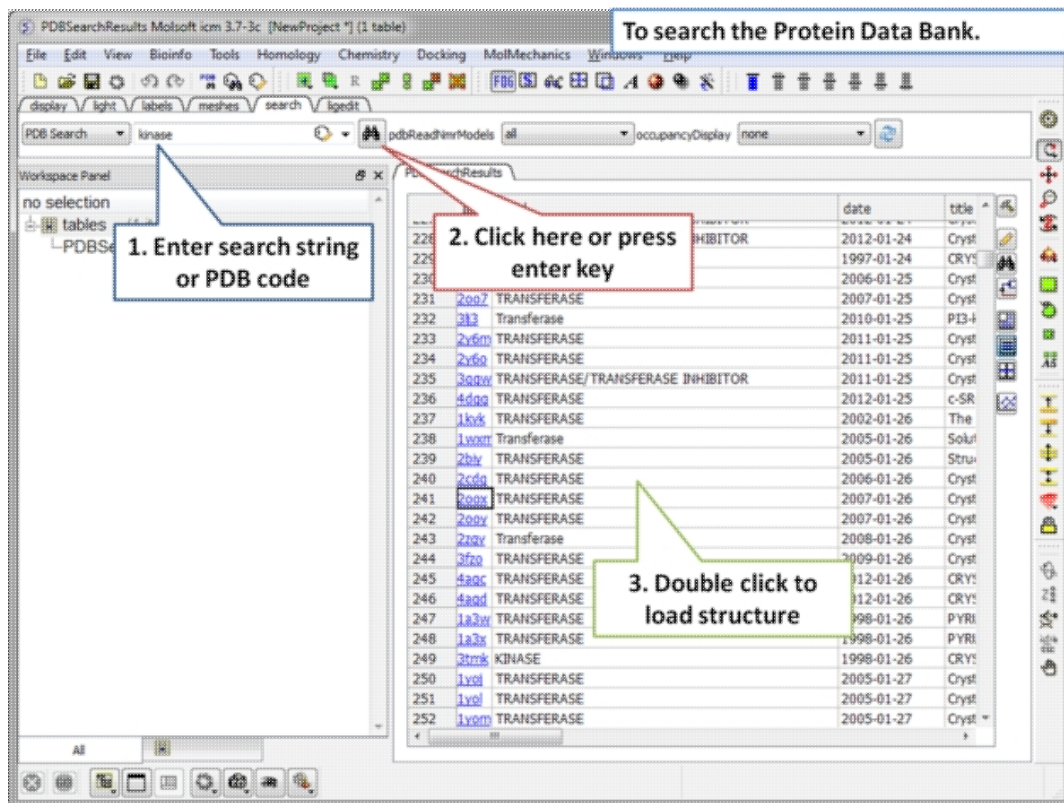
Click on the search tab and then use the drop down menu to identify the database you would like to search and download from.

3.2.1 Search the PDB

How to Search the Protein Databank and Download

The **PDB search tab** provides easy access to the PDB database. You can use keyword searching or type in the PDB code you are interested in. An asterisk (*) wildcard can be used to list all the pdb files currently available in the protein databank. Different fields can be searched by using the drop down arrow as shown below. More advanced PDB search tools and how to use the PDB search result table are described in the section entitled *Searching the PDB*.

Once a search is complete a table of PDB files relating to your search query will be displayed. To view the PDB file in 3D in the graphical display double click on a row in the PDBSearchResults table.



To Search the PDB and Download a Structure:

- Click on the search tab and enter a search string or PDB code.
- A table containing the results will be displayed.
- Double click on a row to load the PDB file. Read more about PDB search [here](#).

NOTE: If you have a PDB structure already saved you can read it into ICM by going to the File Menu and selecting Open. PDB files that have been viewed previously can be loaded using File/ Recent PDB Codes.

3.2.1.1 How to Search the Protein Databank by Chemical Substructure

To search the Protein Data Bank by chemical substructure.

1. Click on the molecule editor button

2. Sketch the substructure

3. Close the window

4. Click here or press enter key

1	Formula	
2	Smiles	
3	ILPAC	
4	InChI	
5	InChIKey	
6	MolWeight	
7	HBA	
8	HBD	
9	RotB	
10	DrugLikeness	
11	MolArea	
12	MolDf	
13	MolLogP	
14	MolLogS	
15	MolPSA	
16	Volume	
17	Bad Groups	
18	Groups	

Step 1: Click on the molecule editor button inside the search tab. Sketch the substructure you are interested in.

Double click on the table to load the PDB structure.

1. Double click to load PDB structure

HETID	mol	ID	date	head
VHL		2zhc	2005-07-27	LYASE
VHL		1xt	2004-09-30	LYASE
2PB				HYDROLASE

Step 2: A table containing the results will be displayed. Double click on a row to load the PDB file.

3.2.1.2 How to Query the PDB by Sequence

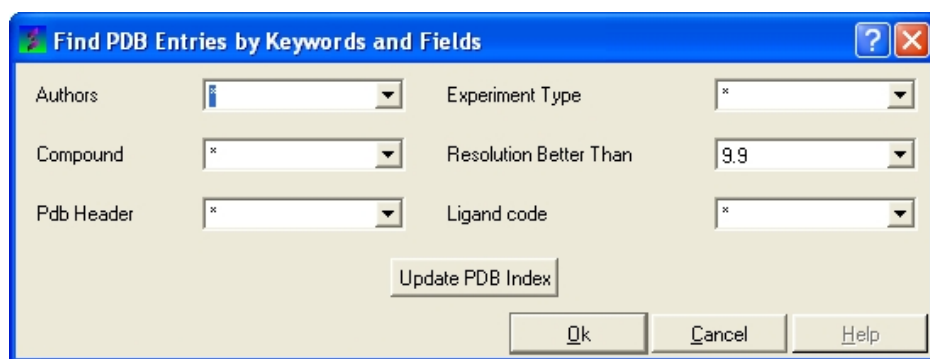
There are a number of ways to search the PDB by sequences.

1. BLAST search the NCBI database using the option "Sequences from PDB".
2. Search by **Sequence Pattern** using the option in the Edit menu.
3. Search by **Identity** using the option in the Edit menu.
4. Search by **Homology** using the option in the Edit menu.

3.2.1.3 How to Query PDB by PDB Field

To query the PDB by field (Author, Compound, PDB Header, Experiment Type, Resolution or Ligand Code

- Select **Edit/PDB** search by field.
- Enter the search string or value
- Click **OK** and a list of related PDB entries based on your search will be displayed in the PDBSearchResults table of the graphical user interface.



3.2.1.4 Load and Display NMR Structures

Use the PDB Search tab to load NMR structures from the PDB. You can use the drop down button shown below to determine how you want to display your NMR structure. You can choose to display and download the first NMR model, all models in the PDB file or all models in the PDB in a stack. If you choose the stack option the the stack will be stored in the object as described [here](#).



3.2.1.5 Occupancy Display

You can use the options in the PDB Search tab to control if and how the partial or zero atom occupancies are displayed. You can choose to circle or label the poor occupancy atoms.



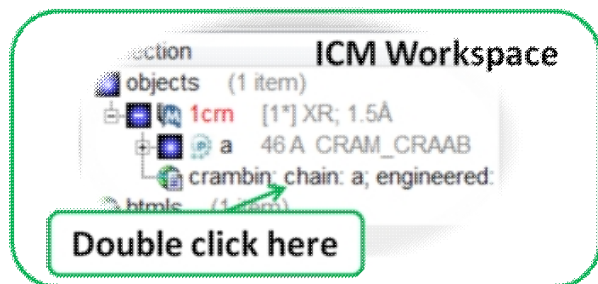
```
>>load-pdb-hyperlinks{pdb search hyperlinks} h4-- Hyperlinks to PDB Website and UniProt {Hyperlinks to Databases}
```

In the PDB Search Results Table you will see blue hyperlinks that will take you directly to the PDB website or Uniprot website.

3.2.1.6 Display PDB Header

To display the PDB Header for a PDB file.

- First load a PDB file into ICM (see Search PDB)
- Double click on the word header in the ICM Workspace.

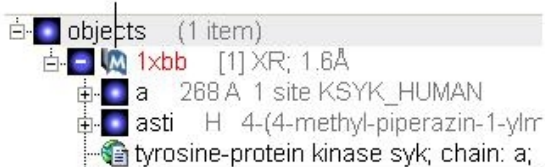


- The PDB Header information will be displayed.
- Click on the blue hyperlinked text to link to external web pages for additional information if needed.

3.2.1.7 Direct link to PubMed

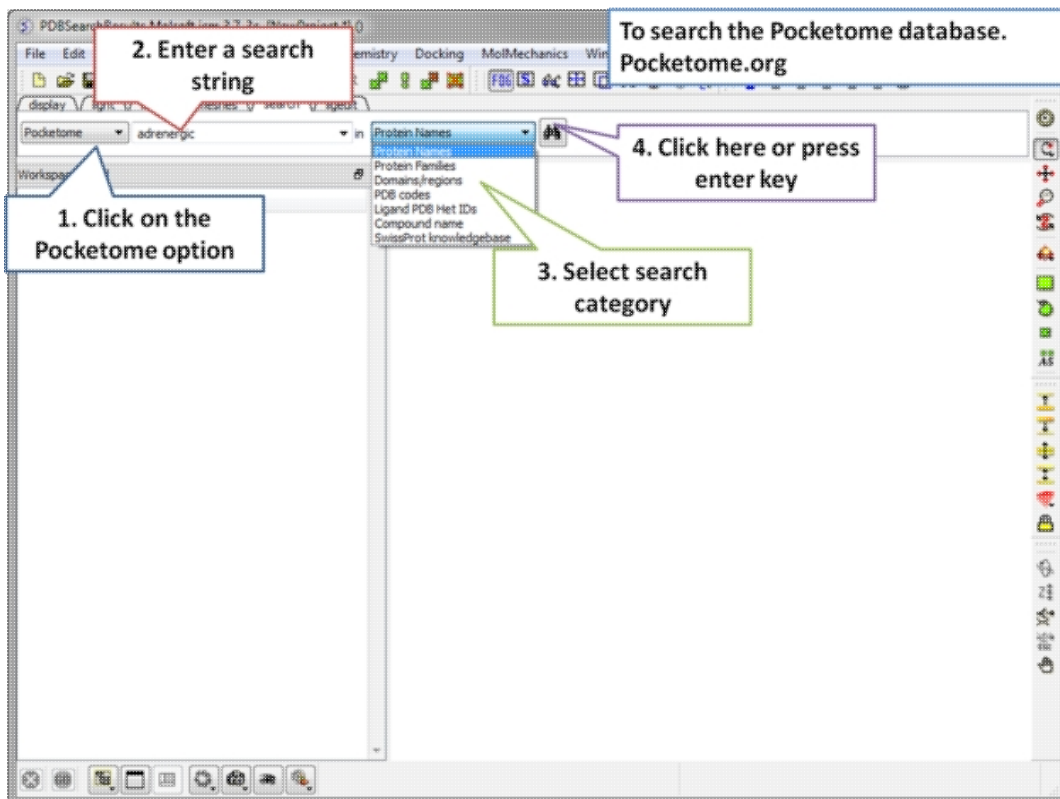
When you search for a PDB file and load it into ICM you will see an icon (shown below) next to your protein name in the ICM Workspace. Click the icon and you will be taken directly to the PubMed primary reference relating to the structure.

Direct access to PubMed

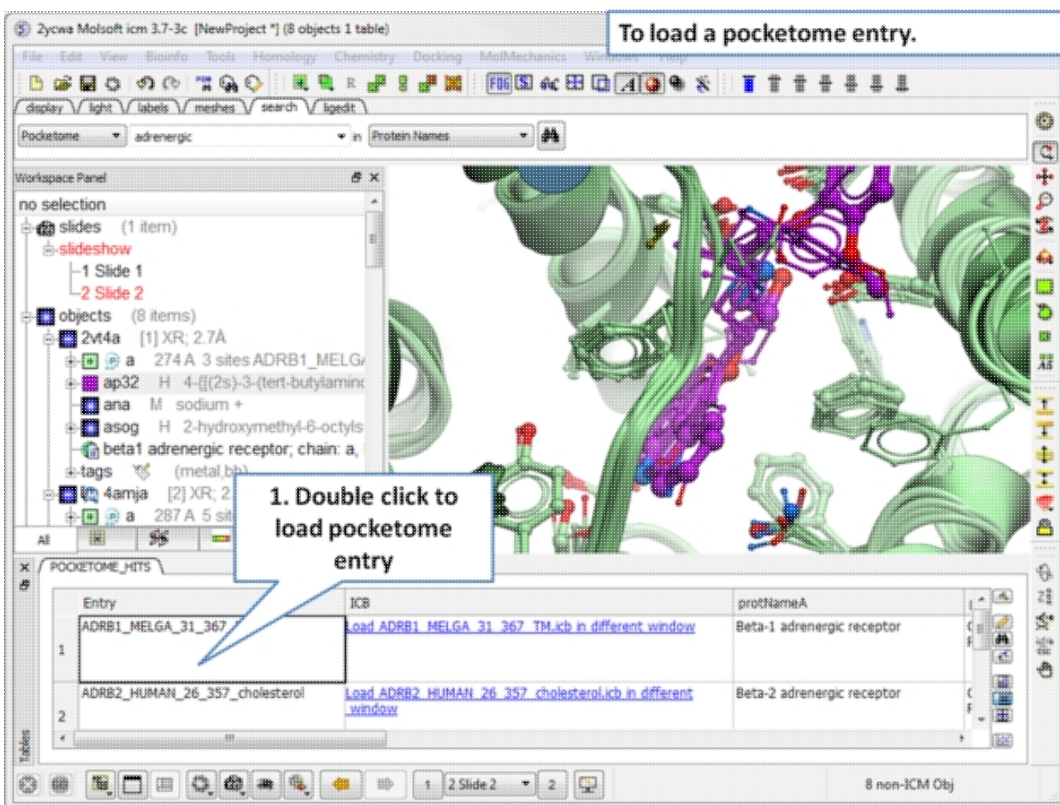


3.2.2 Search Pocketome

The Pocketome (www.pocketome.org) is an encyclopedia of conformational ensembles of all druggable binding sites that can be identified experimentally from co-crystal structures in the Protein Data Bank.



Step 1: Click on the search tab and choose the Pocketome option. Check a field you would like to search e.g. Protein Name, Family, domain...



Step 2: A table of pocketome hits will be displayed. Browse the results and double click to load a pocketome entry.

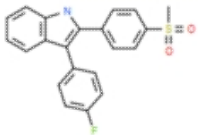
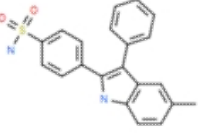
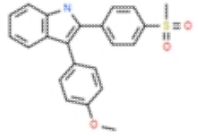
3.2.3 ChEMBL Search

ChEMBL (<https://www.ebi.ac.uk/chembl/>) is a manually curated chemical database of bioactive molecules with drug-like properties. It is maintained by the European Bioinformatics Institute (EBI), of the European Molecular Biology Laboratory (EMBL), based at the Wellcome Trust Genome Campus, Hinxton, UK.

To search and download data from ChEMBL:

- Click on the **Search** tab.
- Choose ChEMBL from the drop down button.
- Enter a search string or search by chemical sketch by clicking on the editor button in the panel.
- Select a field to search e.g. Protein Name.
- Select whether you wish to return just activity data or all data.
- Click the search button.

A table will be displayed as shown below.

mol	chembl_id	synonyms	uniprot_id	fullname
	CHEMBL501208		PGH2_HUMAN	Prosta
	CHEMBL525247		PGH2_HUMAN	Prosta
	CHEMBL499069		PGH2_HUMAN	Prosta
	CHEMBL501208		PGH2_HUMAN	Prosta

The table contains the original activity data without merging or aggregation. For instance,
Chemical 1, Protein 1, Activity 1
Chemical 1, Protein 1, Activity 2
Chemical 1, Protein 2, Activity 1
Chemical 2, Protein 1, Activity 1

Click on the button **Get Best Activity** to display the best activity for each chemical, protein pair. For instance,

Chemical 1, Protein 1, Best(Activity 1, 2, 3, ...)
Chemical 1, Protein 2, Best(Activity 1, 2, 3, ...)
Chemical 2, Protein 1, Best(Activity 1, 2, 3, ...)

Click on the button **Get Ligand Summary** to display a list of chemicals with all activities found in ChEMBL. For instance,

Chemical 1, Protein 1, Activity 1

Protein 1, Activity 2
Protein 2, Activity 1

Chemical 2 Protein 1, Activity 1

Protein 2, Activity 1

Protein 2, Activity 2

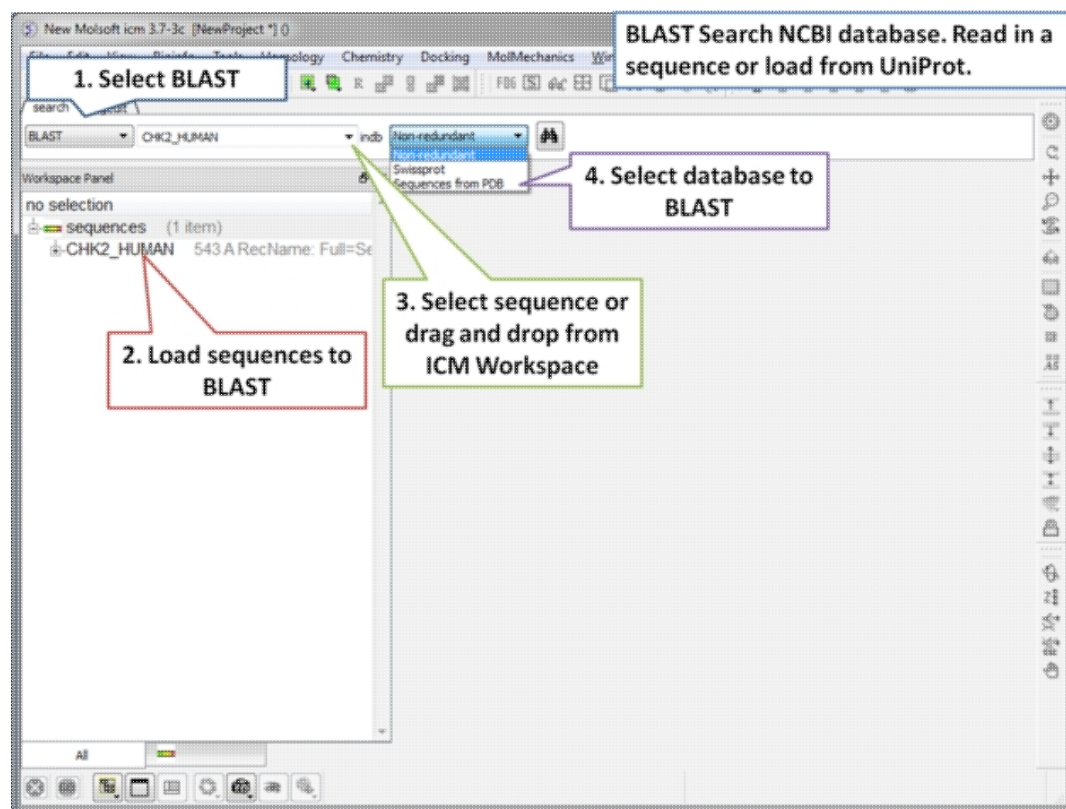
3.2.4 Search SureChEMBL

SureChEMBL (<https://www.surechembl.org/search/>) provides free access to chemical data extracted from the patent literature. To search SureChEMBL:

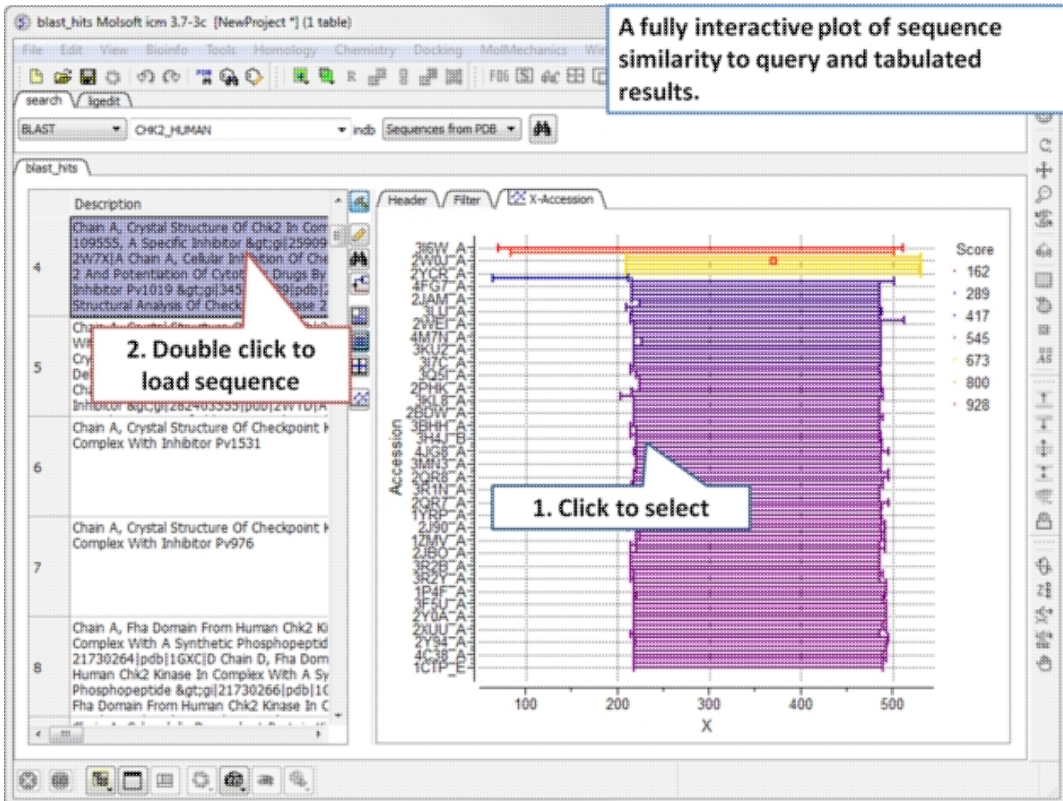
- Click on the **Search** tab.
- Choose SureChEMBL from the drop down button.
- Enter a search string or search by chemical sketch by clicking on the editor button in the panel.
- Select a field to search e.g. Protein Name.
- Select whether you wish to return just activity data or all data.
- Click the search button.

3.2.5 BLAST Search

To BLAST search the NCBI sequence database:

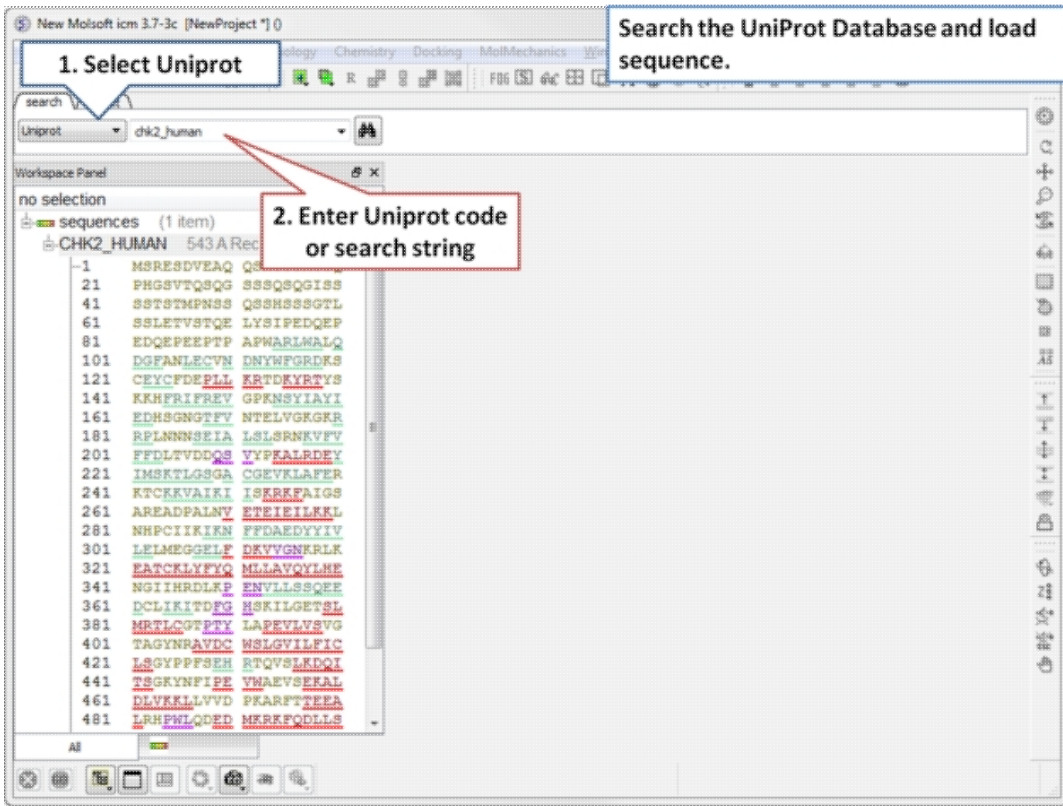


Step 1: Load a sequence into ICM. Select the Search tab and choose the BLAST option. Drag and drop the sequence into the search field, use the drop down menu or type the sequence name.



Step 2: A fully interactive table and plot of sequence conservation will be displayed. Double click to load a sequence.

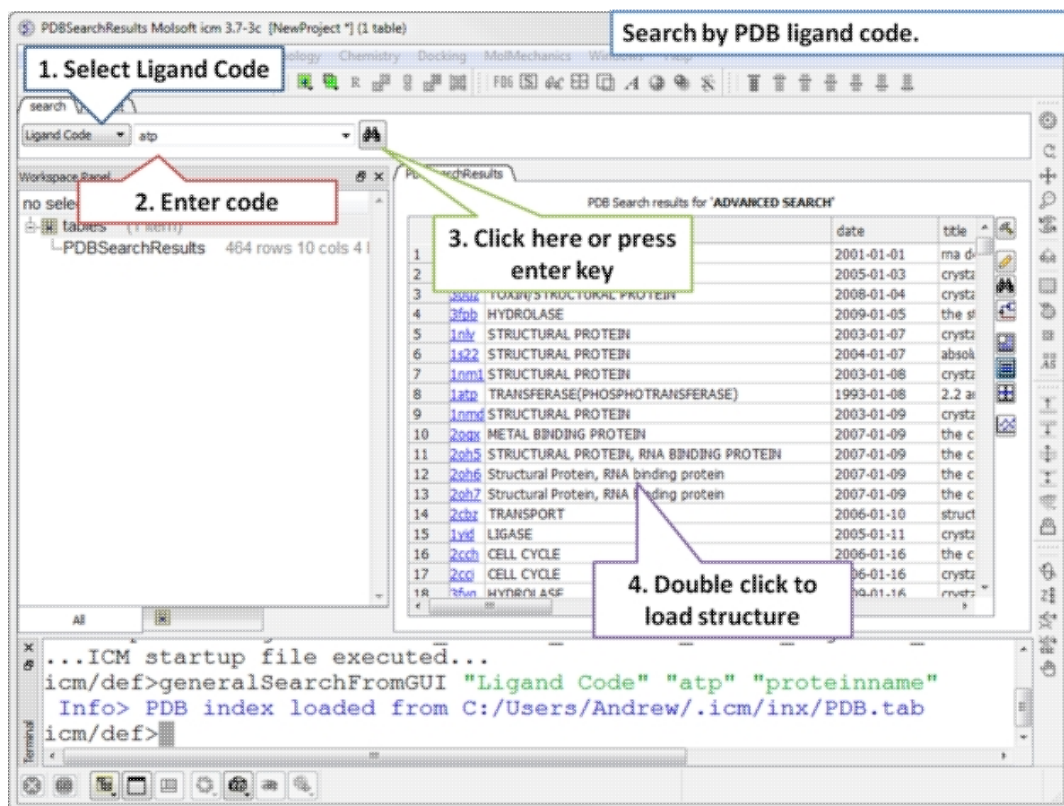
3.2.6 Search UniProt



Step 1: Click on the search tab and select **Uniprot** from the drop down menu. Enter the UniProt code. The sequence will be loaded directly into ICM and you will see it in the ICM workspace.

3.2.7 Search PDB by Ligand Code

To find structures containing a particular ligand in the PDB.



Step 1: Click on the search tab and select **Ligand Code** from the drop down menu. Enter the Ligand code and a table of hits will be displayed.

3.2.8 Search Drug Bank

To search Drug Bank (<http://www.drugbank.ca/>)

Search Drug Bank by key word.

1. Select Drug Bank

2. Enter search text

3. Click here or press enter key

4. Double click to load structure

ID	DrugName	Cpd Synonym
1	Imetit	
739	Clobenpropit	
747		
3		

Step 1: Click on the search tab and select **Drug Bank** from the drop down menu. Enter a search string and a table of results will be displayed.

The table contains the original activity data without merging or aggregation. For instance,
 Chemical 1, Protein 1, Activity 1
 Chemical 1, Protein 1, Activity 2
 Chemical 1, Protein 2, Activity 1
 Chemical 2, Protein 1, Activity 1

Click on the button **Get Ligand Summary** to display a list of chemicals with all activities found in Drug Bank. The button is located in the extra panel (click on hammer - top right of table panel). For instance,
Chemical 1, Protein 1, Activity 1

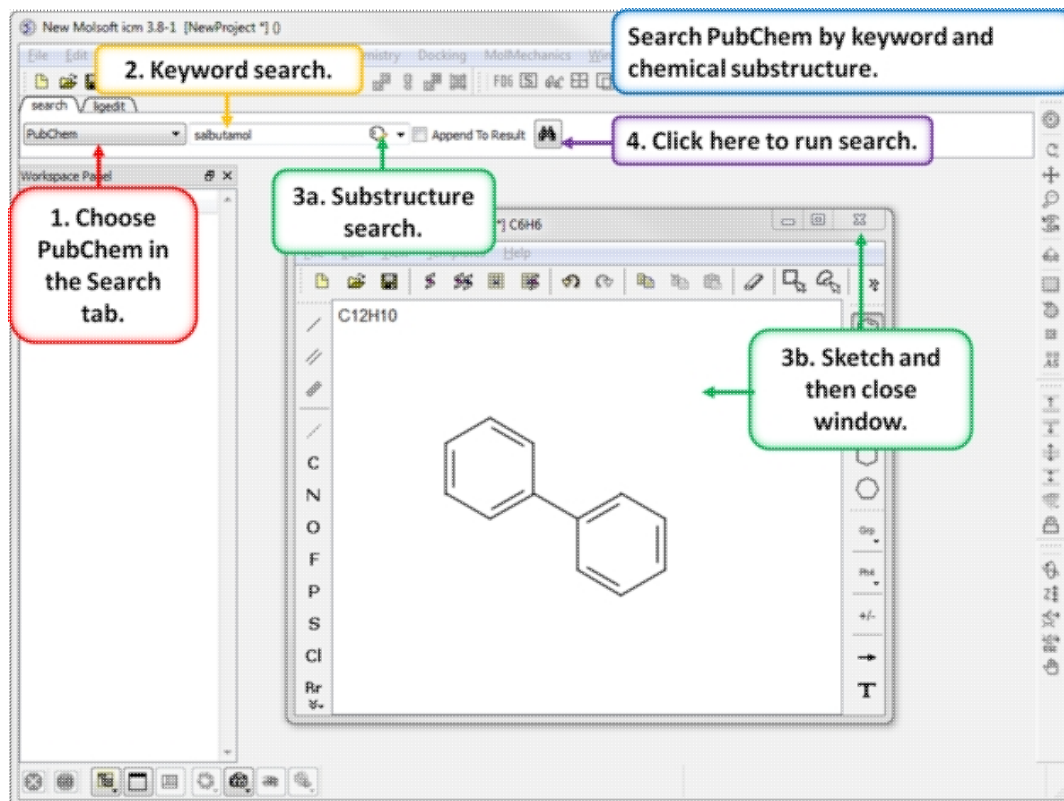
Protein 1, Activity 2
 Protein 2, Activity 1
 Protein 3, Activity 1

Chemical 2 Protein 1, Activity 1

Protein 2, Activity 1
 Protein 2, Activity 2

3.2.9 Search PubChem

To search PubChem (<https://pubchem.ncbi.nlm.nih.gov/>):



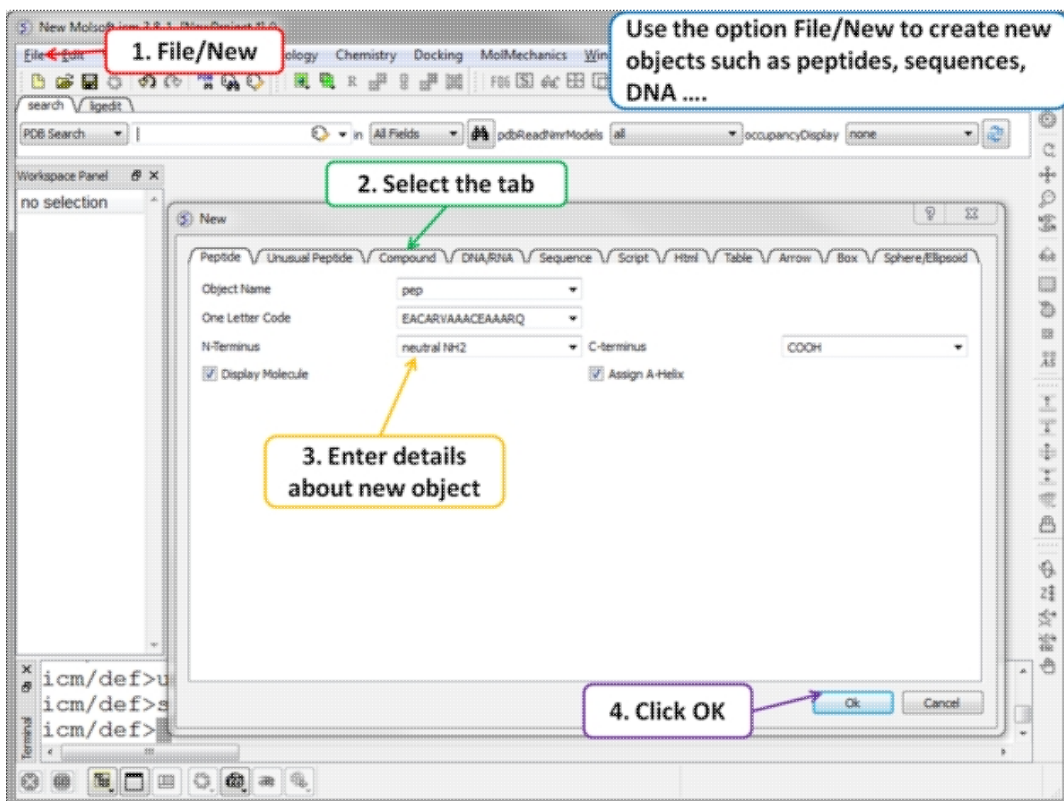
- Click on the search tab and select **PubChem** from the drop down menu.
- Enter a search string and a table of results will be displayed or click on the chemical sketch button and sketch a substructure. When you close the chemical sketch window the smiles string will be added to the keyword search panel.
- Click on the "binocular" button and run the search.
- A table of hits will be displayed.

3.3 Create New Objects

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro | ICM-Chemist

In the File menu there is an option called New. This option can be used to create new objects of the following types:

- Peptide
- Unusual Peptide
- Compound
- DNA and RNA
- Sequence
- Script
- HTML
- Table
- Arrow
- Box
- Sphere/Ellipsoid



3.4 Open and Read Files

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro | ICM-Chemist

Any file that ICM can understand can be opened by:

- Selecting File/Open.

A file with an extension .icb is an ICM binary file and can be viewed in the GUI. A .icb file can contain many objects such as sequences, meshes, protein objects, alignments, tables etc...

3.4.1 Open with Password

To open a file that is password protected:

- File/Open with Password

3.4.2 Extract from icb file

An **icb** file is an `icm project` file, in some instances you may want to take objects saved in an icb file and load it in your current ICM session. This option allows you to view a tabulated list of what a icb file contains and load individual object files from it.

- File/ Extract from ICB.
- Locate the saved icb file.
- A table as shown below will be displayed
- Double-click on any of the entry to extract that object from the icb file.

	name	type	size
2	openFilePRJNAME	string	9
3	1TKI_A_4	sequence	491
4	1DI9_A_17	sequence	509
5	1wFC_5	sequence	538
6	1IAN_23	sequence	415
7	1QL6_A_4	sequence	413
8	1ql6_a	sequence	907
9	2phk_a	sequence	786
10	alig	alignment	10927
11	1ql6	object	231820
12	2phk	object	235471

3.5 Saving Files

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro | ICM-Chemist

Anything you have displayed in an ICM graphics session can be conveniently stored in a single file called an .icb file. To do this:

- File/Save Project

or

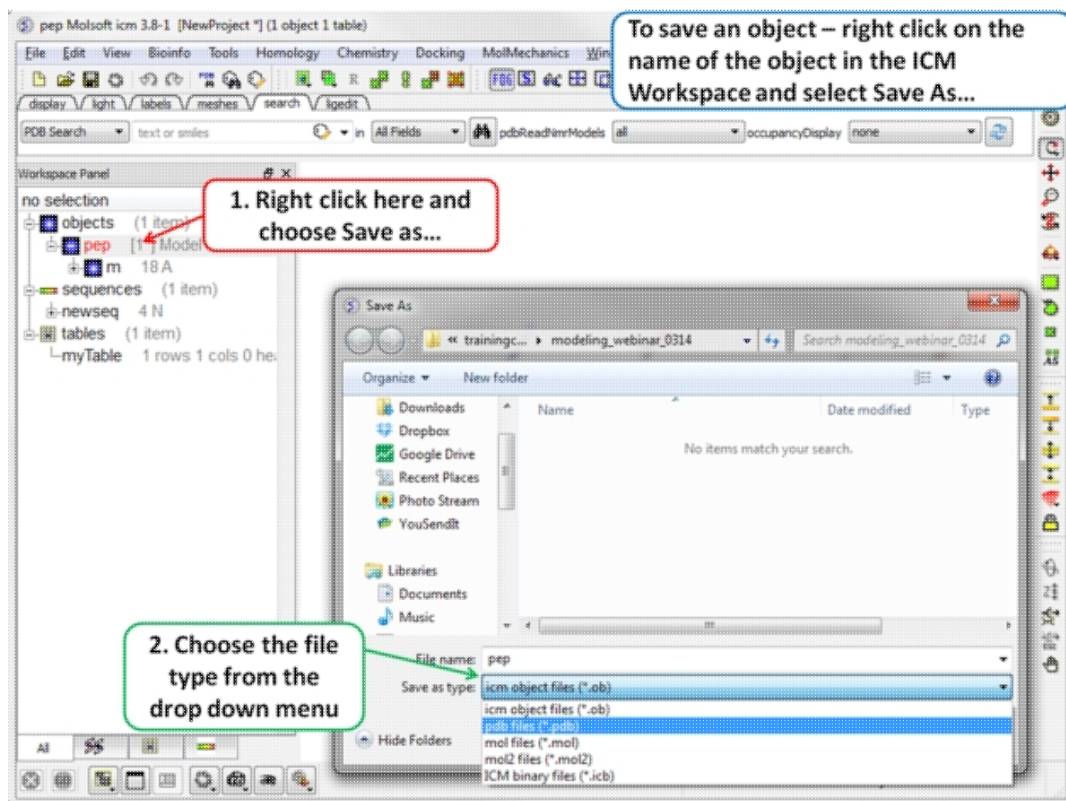
- File/Save Project As

A .icb file can be password protected by:

- File/Save with Password
- Enter a file name or browse for a previously saved project.
- Enter a password
- Determine whether you want the file to be Fully Protected, read only or Read Only and Allow Comments .

Any object (e.g. protein,table, alignment...) inside the ICM GUI can be saved by:

- Right click on the name of the object in the ICM Workspace.
- Select "Save As"
- Choose the file format from the "Save as Type" drop down dialog.



3.6 Making Selections

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro | ICM-Chemist

Note: Click **Next** (top right hand corner) to navigate through this chapter. Headings are listed on the left hand side (web version) or by clicking the **Contents** button on the left-hand-side of the help window in the graphical user interface.

There will be many occasions when you will have to make selections. For example, if you want to display a particular region or molecule contained within your protein structure or if you want to select residues around a binding pocket. If you have a molecule displayed in the graphics window, then selections will be displayed as green crosses. The selection you have made is also displayed at the top of the ICM Workspace. It is always a good idea to keep an eye on what is selected and what isn't.

There are four basic levels of selection

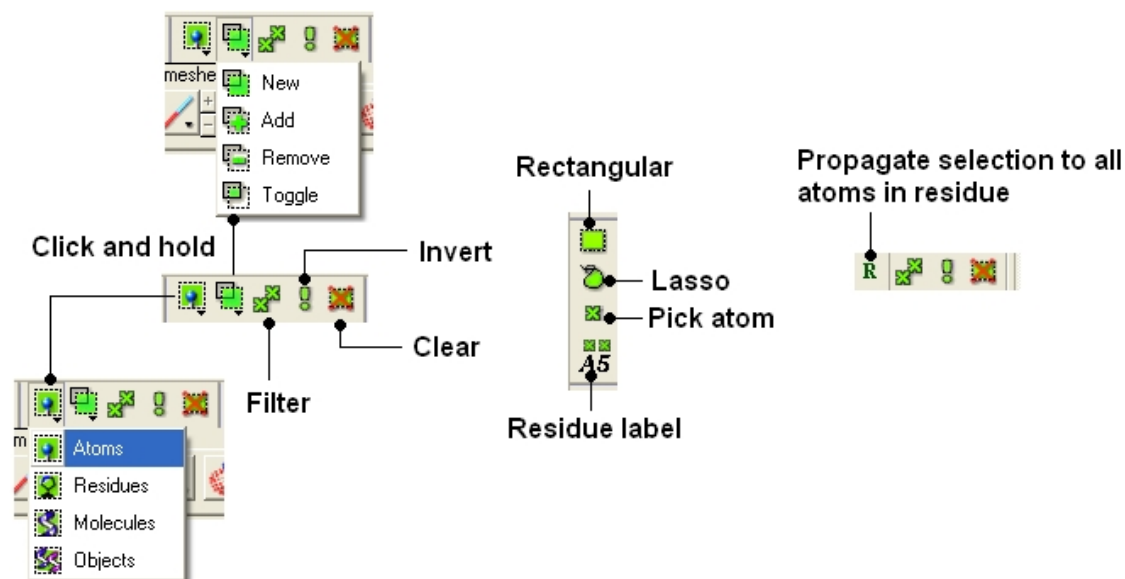
1. Object (eg a PDB structure or ICM object)
2. Molecule
3. Residue
4. Atom

You can make selections in:

- The Graphics Display
- The ICM Workspace (Selections are highlighted in blue)
- Tables
- Sequences
- Plots
- Alignments

3.6.1 Graphical Selection Tools

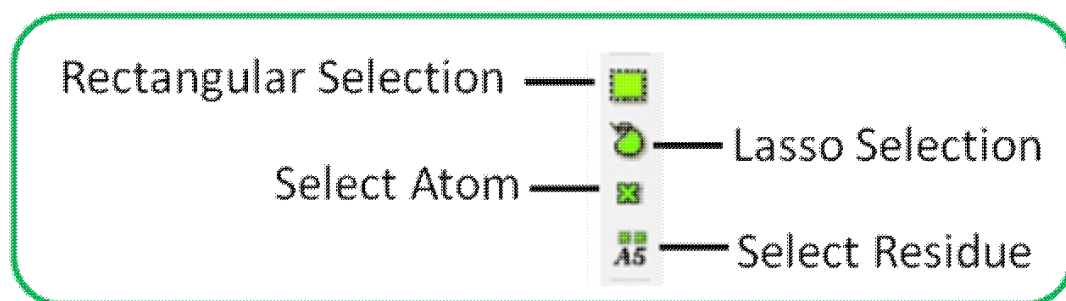
The following buttons can be used to make a selection once a structure is displayed.



NOTE: All selection tool buttons are colored green. Graphical selections are represented as green crosses.

3.6.2 Quick Selection

To make a quick selection the following buttons can be used.



To select parts of your structure:

- Click on the **Rectangular selection icon** and click and drag around the part of the structure you wish to select.

OR

- Click on the **Lasso selection icon** and click and drag your mouse around the area of the structure you wish to select, forming a lasso around it.

To pick individual atoms:

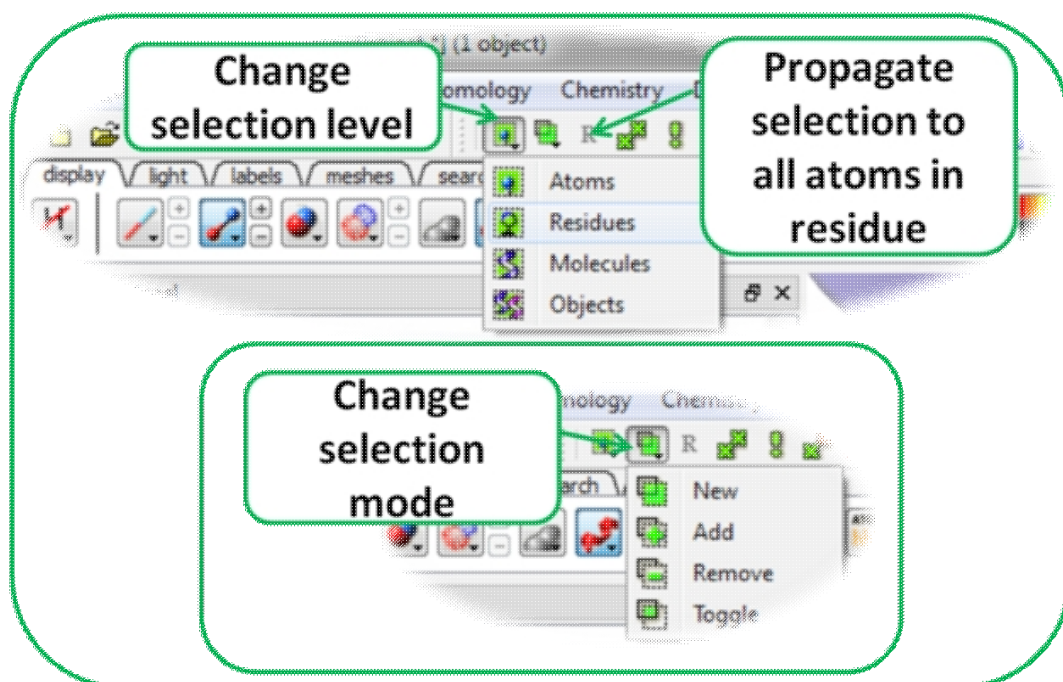
- Click on the 'pick atom' button

3.6.3 How to Change the Selection Level and Mode

It is possible to change the level of selection before or during the building of a selection. The selection level drop-down button can be used to do this (see image below).

For example, a C-alpha of a residue is selected but one would like to select all atoms in the residue. You can change the level to **Residues**. This selection can then be changed into all atoms of the residue by then selecting the **Atoms** level again. Or you can use the **Propagate Selection to all Atoms** button (see image below).

- Click on the **Select objects**, **Select molecules**, **Select residues**, or **Select atoms** icon, depending on which part of the structure you wish to be highlighted.



It is also important to observe the selection mode that is being used. There are four modes:

- **New:** new selection replaces everything selected before
- **Add:** new selection is added to previous selection(s), if any
- **Remove:** previously selection (part or whole of it), if included in the new selection will be unselected.
- **Toggle:** within the new selection, everything that has been selected is unselected and everything that hasn't been selected, will be selected

3.6.4 How to check what is selected.

All selections are displayed as green crosses in the graphical display and blue in the ICM Workspace. All green selections are returned to a variable called as `_graph`.

Selections are displayed as green crosses in the graphical display and highlighted in blue in the ICM Workspace.

Here the selection information is reported

Selection are highlighted in blue in the ICM workspace

Selection shown as green cross in graphics window

```

icm/lxbb>display xstick as_graph
icm/lxbb>as_graph = Atom( Res(Sphere( a_lxbb.asti/1 (Mol( {a_.*}/D
icm/lxbb>undisplay store a_l
icm/lxbb>display wire a_lx
icm/lxbb>undisplay store a_l
icm/lxbb>display xstick a_l
icm/lxbb>as_graph
  
```

All green graphical selections are returned to a variable called as_graph

3.6.5 Orange Selection

Sometimes it is necessary to have two different selections. The Orange selection allows you to do this it is useful for such operations as superposition and more technical procedures such as designing a protein loop. The orange selection is returned to a variable called as2_graph.

Orange selection allows you to copy a green selection to another selection. This allows you to have two different selections.

Click here to transfer a green selection to an orange selection

Orange selection

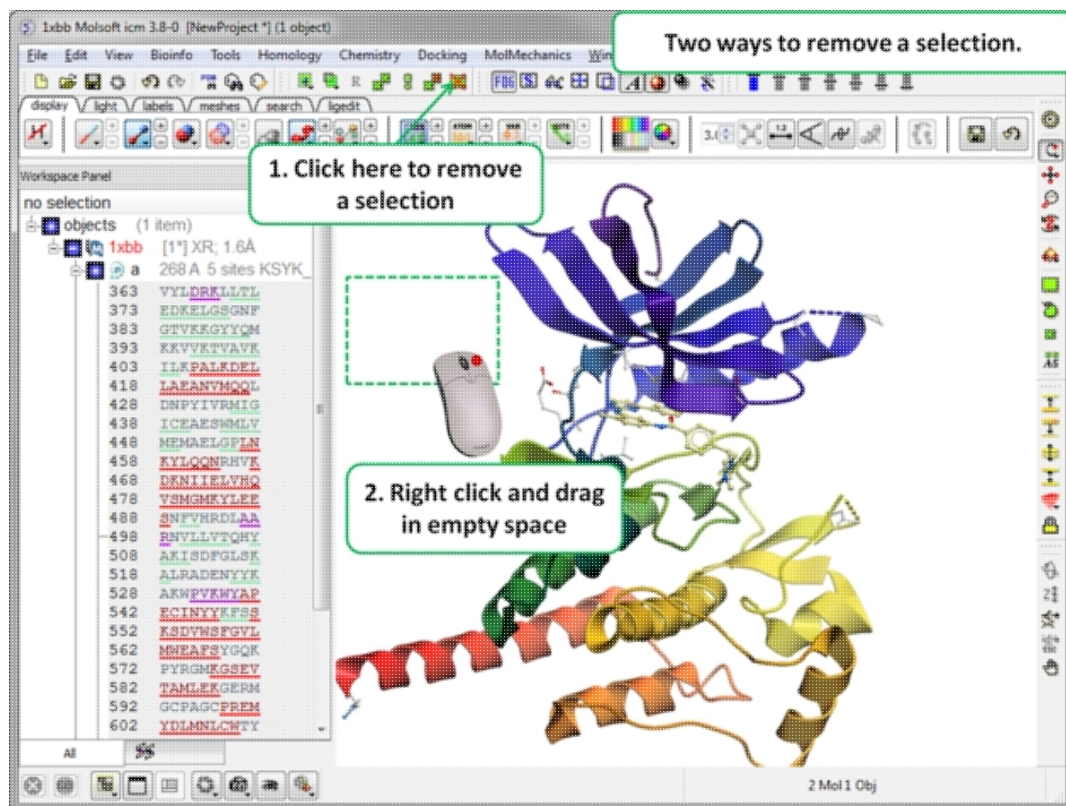
3.6.6 Clear Selection

To unselect everything you have previously selected:

- Simply click on the **Clear Selection** button on the selection toolbar.

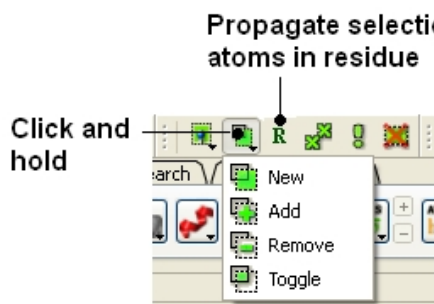
OR

- Right click and drag away from the displayed structure.



3.6.7 Changing a Selection

Once you have made a selection you may wish to add or remove parts of the selection. The buttons shown below allow you to accomplish this.



To add or remove from your current selection:

- Click on the **Selection mode: add** or **Selection mode: remove** icon on the toolbar.
- Click and drag around the part of your structure you wish to add or remove.

You may also wish to invert your selection in a specific part of the structure.

The parts that are currently selected will become unselected, and the unselected parts will become selected.

In order to invert a selection:

- Click on the **Invert** icon on the toolbar.

If you wish to select and unselect certain regions of a selection the toggle selection button is very useful.

- Click on the Toggle selection button.
- Right click around the selections you wish to select or unselect.

NOTE: The selection you have made is recorded at the top of the ICM workplace. Any selection is stored in the variable as_graph.

3.6.8 Filter Selection

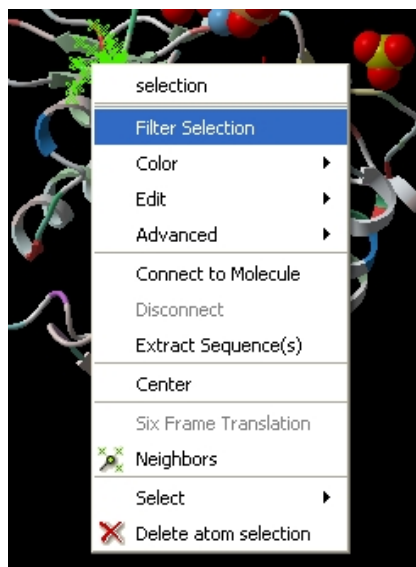
You may want to be very specific about a selection you want to make. For example you may only want to select protein backbone atoms.

The button shown below enables you to filter your selection:

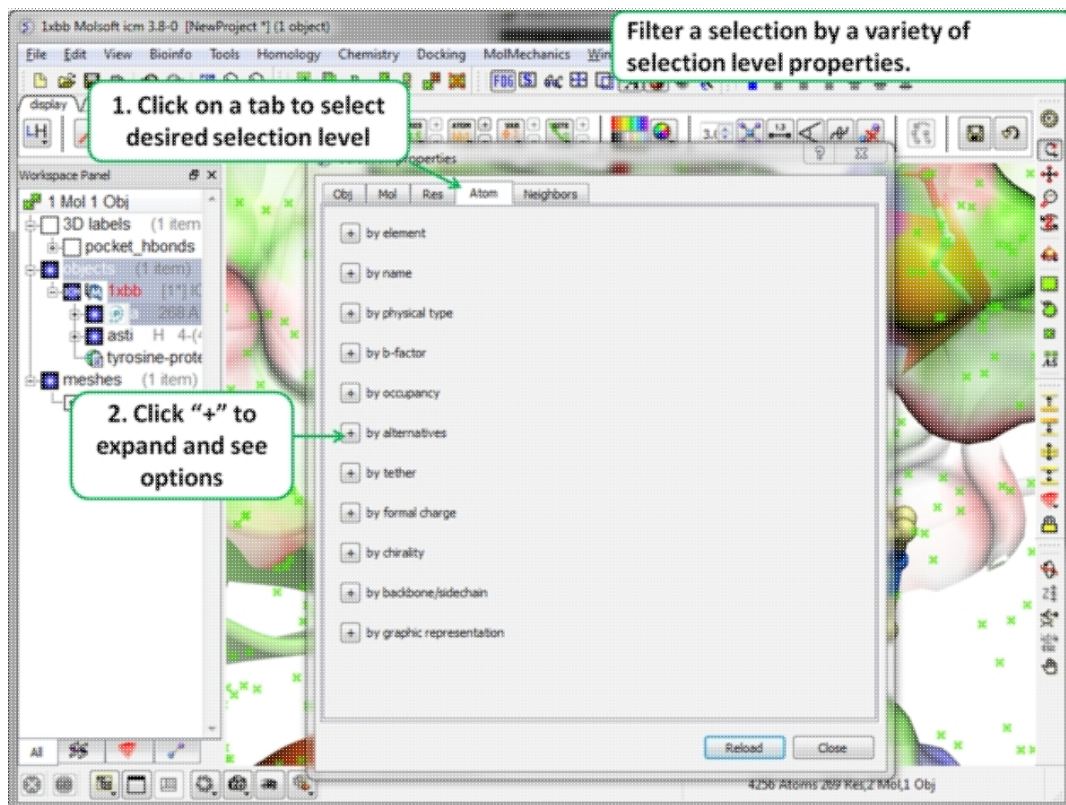


Or

Right click on a selection and a menu as shown below will be displayed.

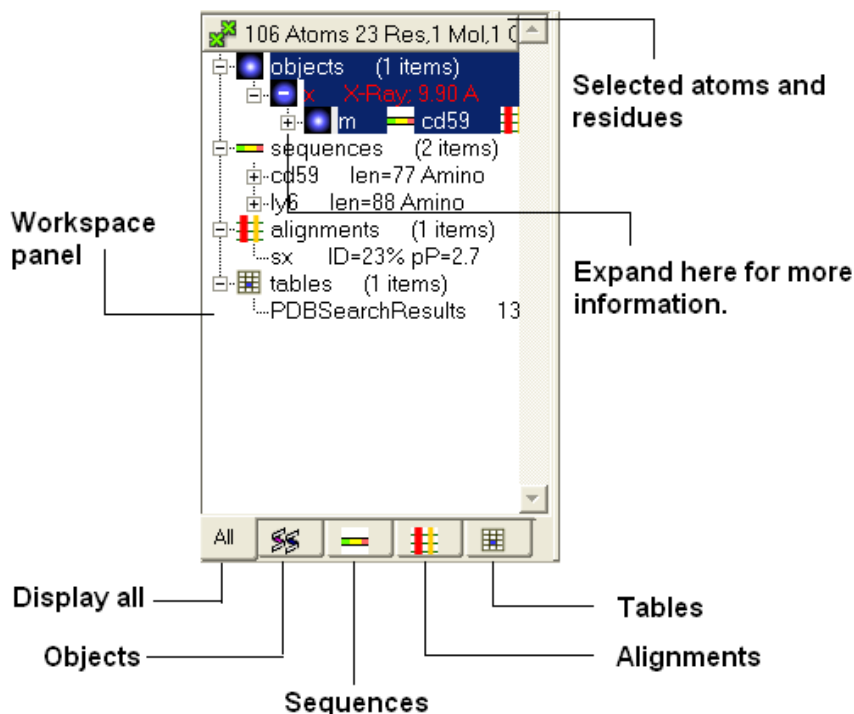


- Select the Filter Selection option.
- A dialog box will be displayed as shown below.



3.6.9 Workspace Selections

In the default GUI layout the workspace panel is located to the left of the 3D graphics display. It is a great tool for keeping track of all your sequences, pdb structures, objects, tables and alignments. As you will see in this section it also provides a way of making selections.



3.6.10 Workspace Navigation

Once you have mastered how to navigate the ICM workspace making a selection will become easier. Each object is divided into 3 levels:

1. Object Level - Shown in red if it is the current object. Holds details about the structure - name, X-ray, NMR, resolution etc. Importantly it will state whether the structure is an ICM object or a structure straight from the PDB. To learn how to convert a PDB into an ICM object go to the section on converting a PDB.
2. Molecular Level - Shows the individual subunits, ligands and hetatoms of a molecule.
3. Residue Level - Shows the sequence.

The screenshot shows a tree view of the ICM workspace. The root is 'no selection'. Underneath is 'objects (3 items)'. The first item is 'x1 ICM' (highlighted in red). It has three sub-items: 'm cd59 sx 77 Amino', 'x X-Ray: 9.90 A', and '1cm X-Ray: 1.50 A; crambin'. The '1cm' item has a sub-item 'm 46 Amino', which in turn has a sub-item '1' showing a sequence: '1 ITCCPSIUAR SNFNUCRLPG', '21 TPEAICATYI GCIIIPGATC', and '41 PGDYAN'. Annotations include: 'Current object in red' pointing to the 'x1 ICM' item; 'Object level' pointing to the '1cm' item; 'Molecule level' pointing to the 'm 46 Amino' item; 'Residue level' pointing to the sequence lines; and 'Object information' pointing to the details of the 'm 46 Amino' item. A note at the bottom says 'Click to expand tree'.

NOTE: You can expand each level of the ICM workspace by clicking the "+" button as shown above.

3.6.11 How to Select an Object

To select the whole object:

- Double click on the object level.

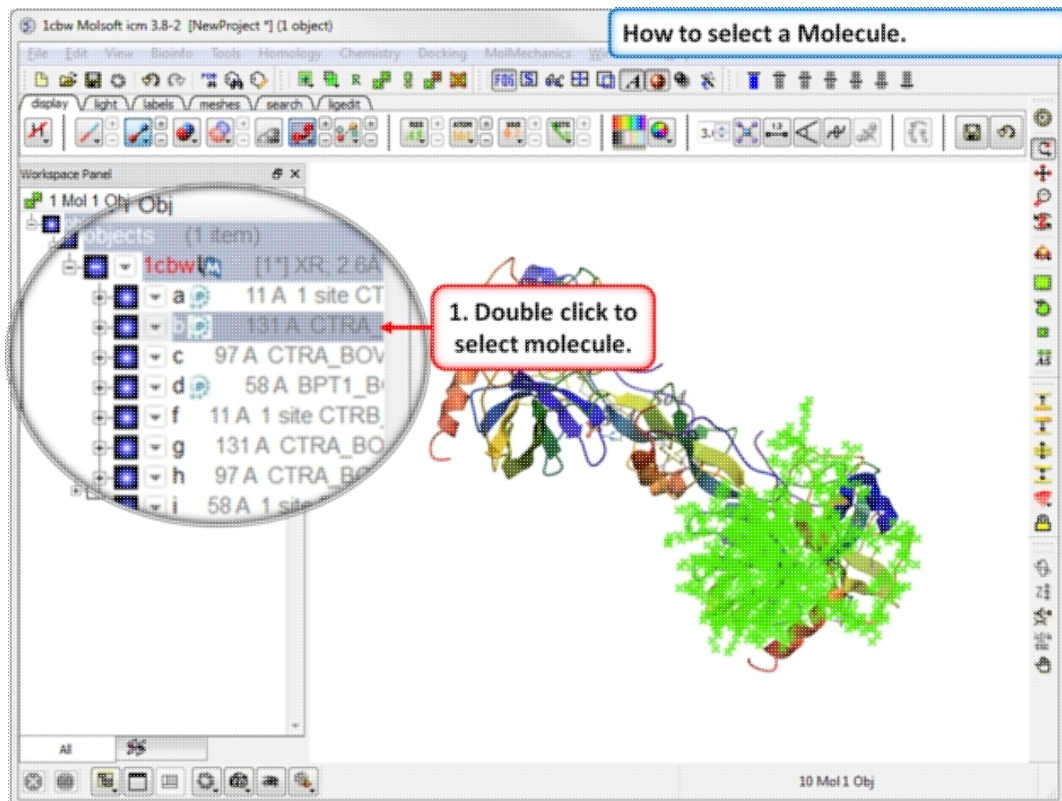
This screenshot is similar to the previous one but with annotations for selection. A line points to the 'objects (3 items)' header with the text 'Double click here to select ALL currently loaded objects.' Another line points to the '1cm X-Ray: 1.50 A; crambin' item with the text 'Double click here at the object level to select the whole object'.

Use the CTRL button to select multiple non-contiguous objects or if they are continuous you can use double click and hold the tab button.

3.6.12 How to select a Molecule

To select a molecule(s):

- Double click on the molecule in the ICM workspace.



Use the CTRL button to select multiple non-contiguous molecules or if they are continuous you can use double click and hold the tab button.

3.6.13 How to select Residues

There are different options to select residues:

OPTION 1:

- Click and drag over the residues you wish to select in the ICM workspace. Selected residues will be highlighted in dark blue in the workspace and with green crosses in the graphical display.

Selection information is recorded here

Click here to expand tree to show amino acid residues.

Click and drag over residues you wish to select.

Selected residues will be highlighted in the workspace and graphics window

184 Atoms 35 Res, 1 Mol, 1 Obj

objects (1 items)

1BB X-Ray, 2.80 Å: rhodopsin

a 338 Amino OPSPD_BOVIN

1	MNGTEGPNFY	UPFSNKTGVU
21	RSPFEAPQVY	LAEPVQFSML
41	RAYMFLIML	GFPINFLTLY
61	UTUQHKKLRT	PLNYILLNLA
81	UADLFHVFGE	FTTTLTYSLH
101	GVFVEGPTGC	NLEGFFATLG
121	GEIALWLSLV	LAIERYVUUC
141	KPMSNFRFGE	NHAIMGUAFI
161	WUMALACAAP	PLUGWSRYIP
181	EGHQSCGID	YYTPHEETNN
201	ESFUIYMFUV	HFIIPLIUIF
221	FCYGLUFTU	KEAASATTQ
241	KAEKEUTRMU	IIMVIAFLIC
261	WLPYAGUAFY	IFTHQGSDFG
281	PIFNTIPAFF	AKTSAUVNPU
301	IYIMMNKQFR	NCHUTLCCG
321	KNPSTTUSKT	ETSQVAPA

Any selection is highlighted in the workspace as well as in the 3D graphics window if the structure is displayed.

OPTION 2:

- Click on the **rectangular selection** icon or **lasso selection** icon on the toolbar.
- Click and drag around the residues you wish to select. Selected residues will be displayed by green crosses on the graphical display and blue in the ICM workspace.
- Use the propagate selection to residue level button.

OPTION 3:

- Right click on the a residue in the graphical display and a menu as shown here will be displayed.

- selection
- Selection Dialog
- Advanced
- Residue atoms
- Open with MolEdit
- Connect to Molecule
- Disconnect
- Extract Sequence(s)
- Center
- Annotate selection
- Neighbors
- Closed Cavities
- Select
- Delete residue selection

- Click on **Select** and a further menu will be displayed.
- Click on **Residue, Molecule** or **Object**.

- Closed Cavities
- Select
- Delete residue selection

- Residue
- Molecule
- Object

OPTION 4:

Use the `select residue` button.

3.6.14 Select All

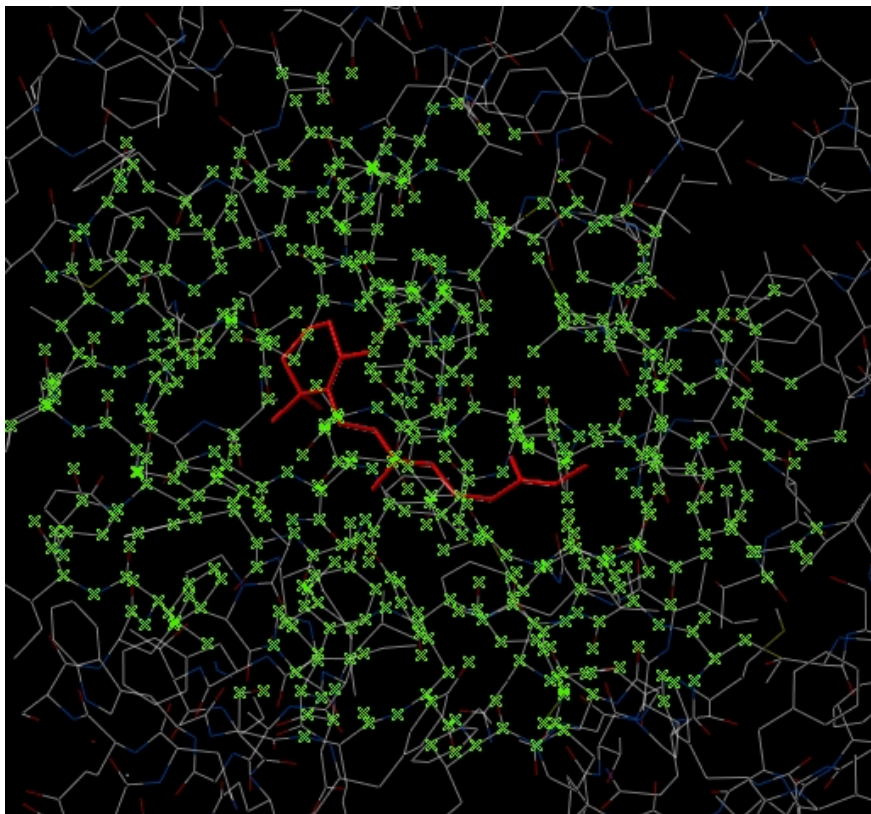
Ctrl + A will select everything in the ICM workspace, and Ctrl + Shift + A will unselect your objects.

NOTE: The selection you have made is always recorded at the top of the ICM workplace. If you are familiar with using the ICM terminal (See language manual) the atoms, residues, molecules or objects selected interactively in the graphics window are automatically s

3.6.15 Selecting Neighbors

In some instances you may only want to display or select only a subset of a structure. For example you may only wish to display the residues surrounding a ligand (as shown below (ligand red; graphical selection green crosses)). The "Selecting Neighbors" option selects the residues within a sphere of a defined radius.

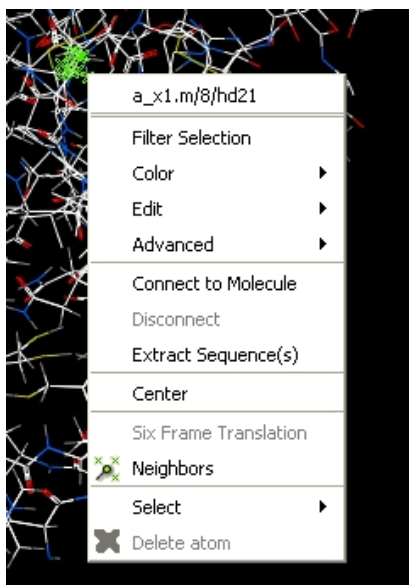
There are two ways of selecting neighbours to a particular atom or residue in ICM. Either by right clicking on the atom or residue in the graphical display or by right clicking in the ICM workspace.



3.6.16 Selecting Neighbors: Graphical

To select neighboring atoms or residues around a sphere of a certain radius:

- First select the residue(s) or atom(s) around which you wish to select neighbors. (See the Selection Toolbar Section)
- Right click on the selection and a menu as shown below will be displayed or choose Tools/Geometry/Neighbors.



- Select the Neighbors option and a data entry box as shown below will be displayed.

This option will allow you to make a spherical selection.

The window will open as displayed as below:



- Select the molecule you wish to select neighbors around. For example you can select a ligand in the ICM Workspace and then choose the **Graphical Selection** option in the "Select Neighbors For" dialog entry box. Or alternatively you can select the object by clicking on the drop down button next to the "Select Neighbors For" dialog entry box.
- Enter the radius in Angstroms for the neighbor selection. e.g. 5.
- **Type** - this option is **important**. This option relates to what is going to be selected. For example if you leave this option as **visible** and you only have ribbon representation displayed for your receptor (e.g. when selecting neighbors for a ligand) then only backbone atoms will be selected.

Selection **Type** option includes:

- **visible** - will select all atoms displayed within the radius selected.
- **visible sidechains** will select all visible side-chains - not backbone atoms.
- **same_object_other_chains** will select all atoms in other chains in the same object as the original selection.
- **other objects** will select atoms in objects other than the original selection.
- **same object** will select atoms in the same object as the original selection.
- **all_objects** will select atoms in all objects
- **choose_from_list** will allow you to select the object you wish to include in the neighbors selection.
- **exclude source** if checked will not include your original selection in the spherical selection.

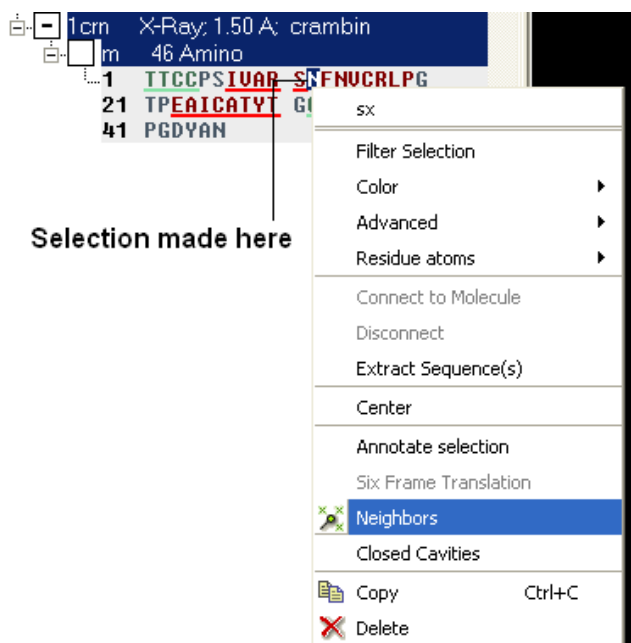
- **unselect water** if checked will not select water molecules
- **Undisplay Beyond Selection** will only display the atoms selected.

NOTE: The selection you have made is always recorded at the top of the ICM workplace. If you are familiar with using the ICM terminal (See language manual) the atoms, residues, molecules or objects selected interactively in the graphics window are automatically saved in the variable as_graph. Graphical selections are shown in green (crosses) or highlighted in blue in the ICM Workspace.

3.6.17 Selecting Neighbors: Workspace

To select neighboring atoms or residues around a sphere of a certain radius from a residue in the ICM workspace:

- First select the residue in the ICM workspace around which you wish to select neighbors. (See the Residue Selection)
- Right click on the selection and a menu as shown below will be displayed.



- Select the Neighbors option and a data entry box as shown below will be displayed.
- Follow the instructions in the previous section.

NOTE: The selection you have made is always recorded at the top of the ICM workplace. If you are familiar with using the ICM terminal (See language manual) the atoms, residues, molecules or objects selected interactively in the graphics window are automatically s

3.6.18 Select by Residue Number

To select by residue number:

- Right click on the molecule you wish to make the selection in and choose Select/Res Number...
- Enter the residue numbers e.g. 3,7:12,30 Colon indicates range.

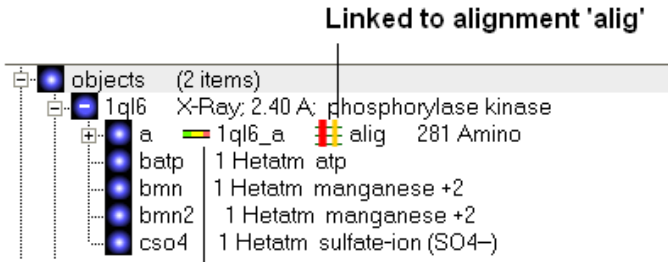
3.6.19 Alignment and Table Selections

Descriptions on how to make selections in Alignments and Tables are in the sections entitled Making Selections in Alignments and Making Table Selections.

3.6.20 Making Links

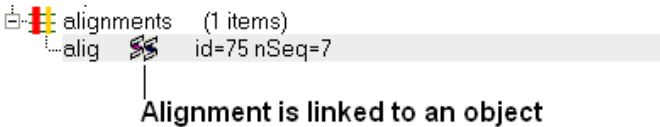
It is sometimes necessary to make links between sequences objects and alignments. A link enables you to make selections in one environment such as an alignment and then these selections are transferred to the object such as the PDB structure displayed.

If a link is made then a symbol will be displayed next to the object in the ICM workspace. In the example shown below subunit_a of the X-ray structure 1ql6 is linked to the sequence 1ql6_a and the alignment called 'alig'.



Linked to sequence 1ql6_a

If an object is linked to an alignment a symbol as shown below will be displayed.



Alignment is linked to an object

To link a sequence from an object - extract the sequence from the object.

- Right click on the object in the ICM workspace.
- Select extract sequence.

To link a sequence and object to an alignment.

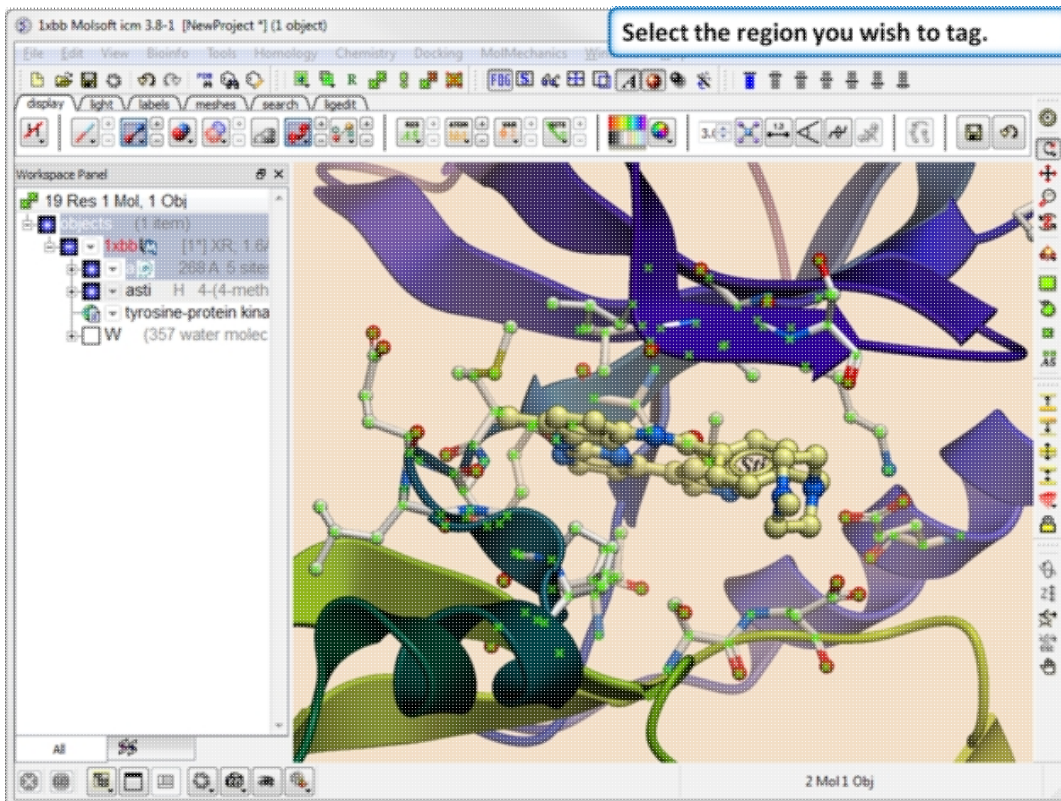
Use the extracted sequence as described above to build your alignment.

In addition a link can be made between a structure and alignment by:

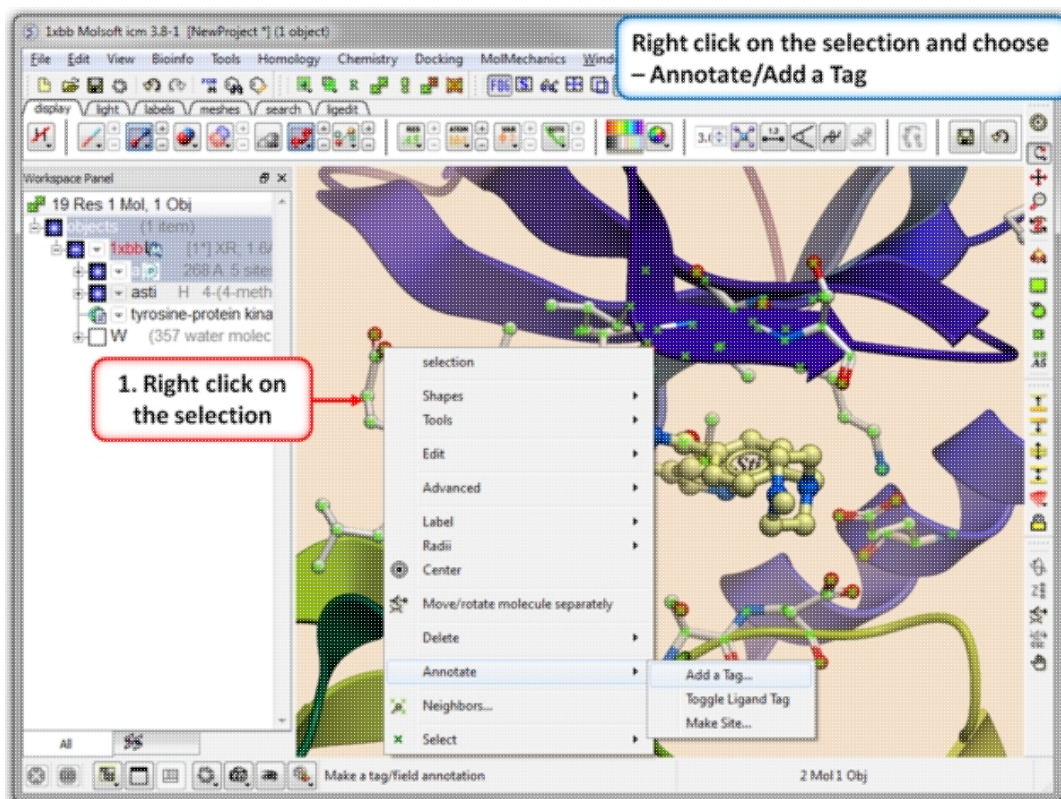
- Bioinfo/Link to Structure.
- Enter alignment name.
- OK

3.6.21 Tags

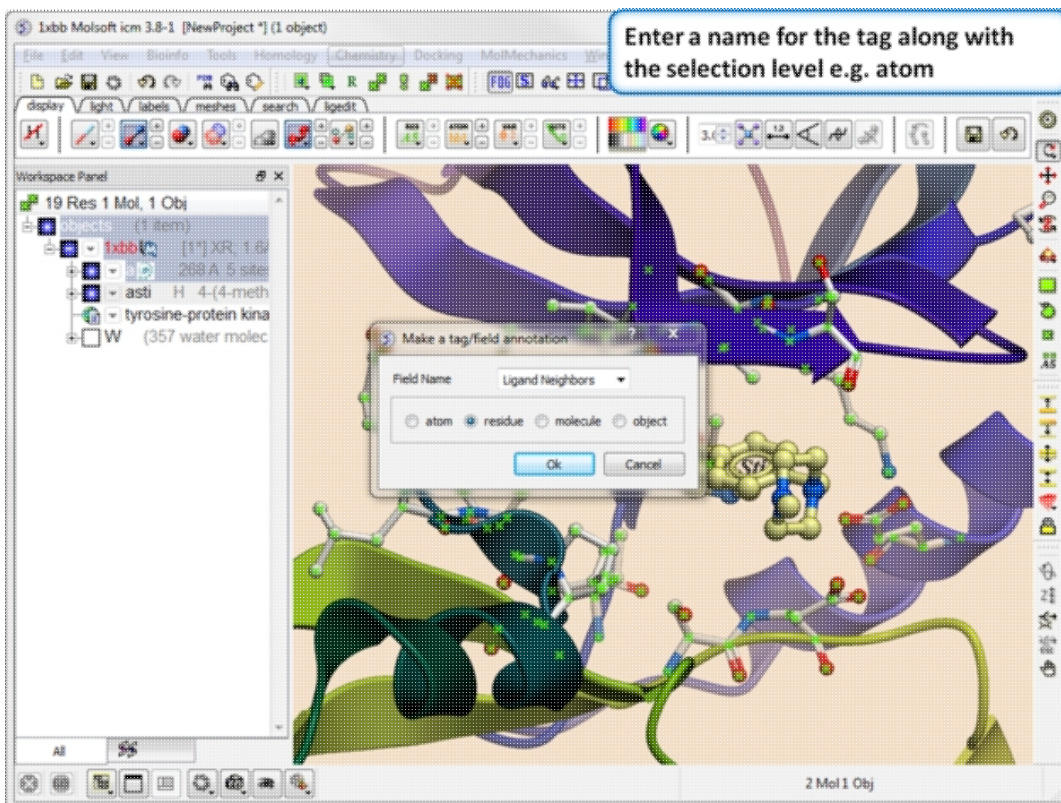
It is sometimes convenient to tag selections so you can come back and use them at a later date. To do this:



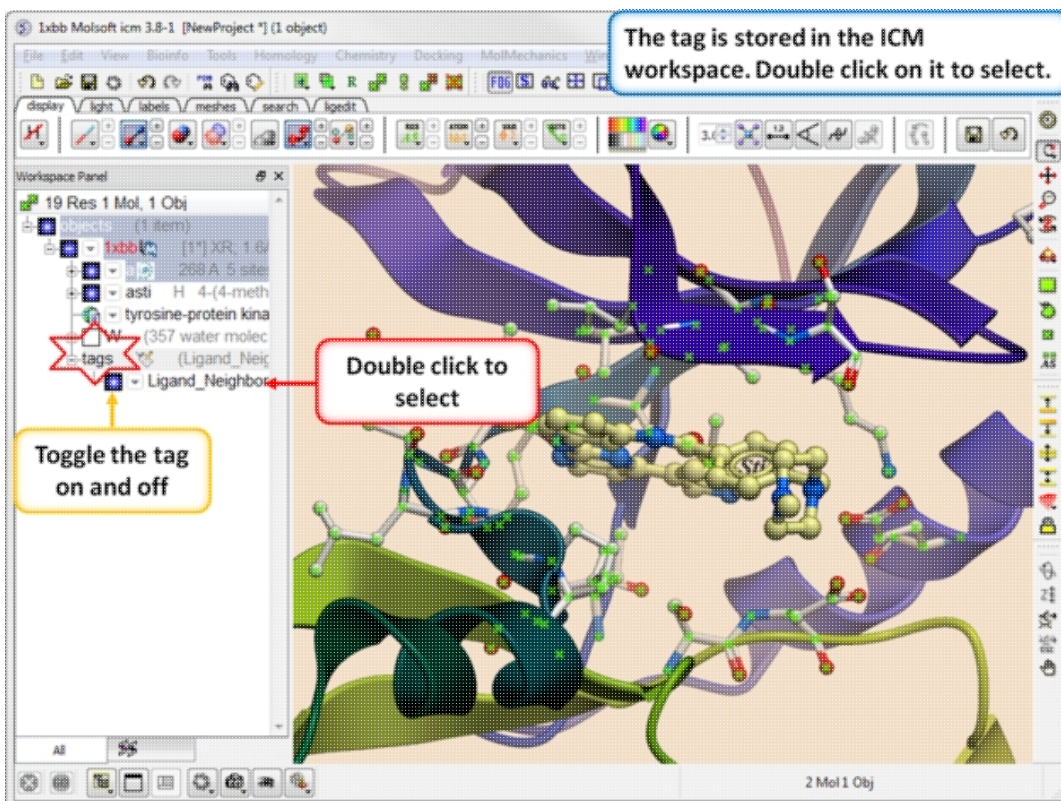
Step 1. Select the region you wish to tag.



Step 2. Right click on the selection and choose Annotation/Add a Tag.



Step 3. Give the tag a name and choose the selection level (atom, residue, molecule, or object).



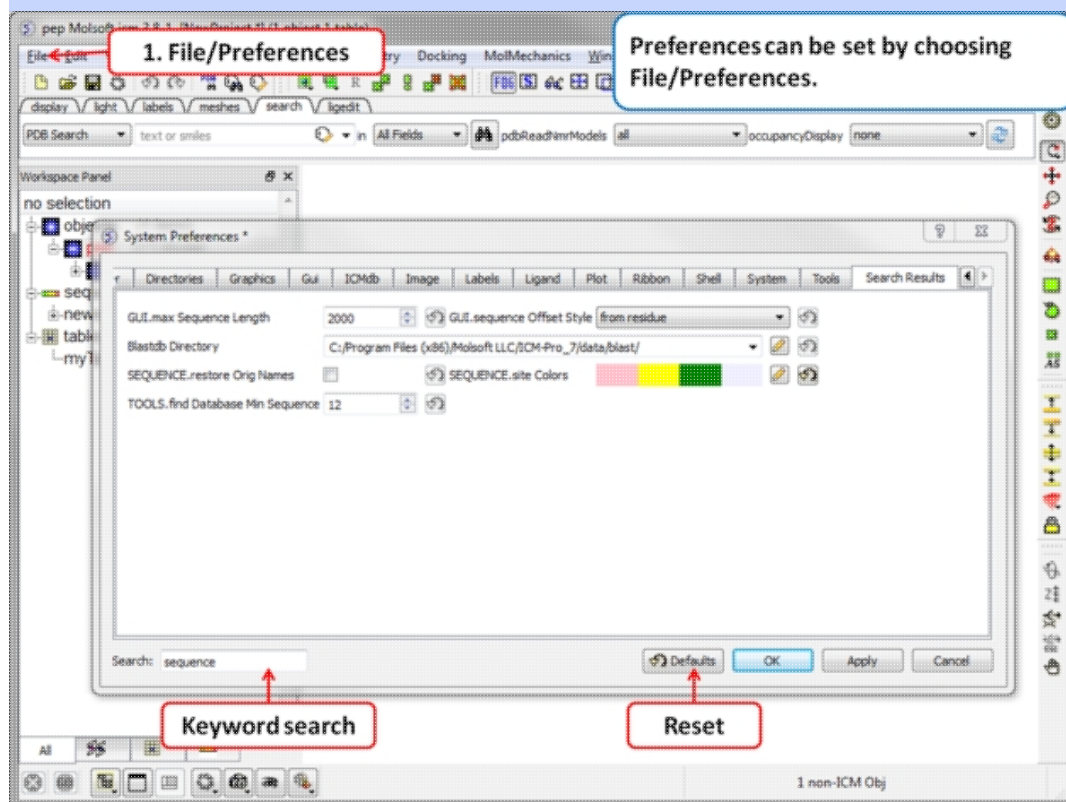
Step 4. You will see the tag in the ICM Workspace. You can toggle the display on or off or select the tag by double clicking on it in the ICM Workspace.

3.7 Preferences

Your ICM preferences can be changed by:

- Select File/Preferences.

NOTE: There is a "Reset to Default" button in case you make any changes you are not happy with and also a search option.



3.7.1 Bonds Preferences

To change Bond Preferences:

- Select File/Preferences.
- Choose the **Bonds** tab.

GRAPHICS.ballStickRatio - A default ratio of ball and stick radii. This ratio is applied when the styles are switched from the GUI xstick toolbar. Default (1.4)

GRAPHICS.hbond Ball Period - Default (3)

GRAPHICS.hbondMinStrength - parameter determines the hbond strength threshold for hbond display. The strength value is between 0. and 2. By changing 1. to 0.2 you will see more weak hydrogen bonds. Default: (1).

GRAPHICS.hbondStyle - determines the style in which hydrogen bonds are displayed. Here hbond-Donor, Hydrogen, and hbond-Acceptor atoms will be referred to as D, H and A, respectively,

GRAPHICS.hetatmZoom - The default ball and stick radii of a ligand can be different by the GRAPHICS.hetatmZoom factor. This makes a better ligand view since the ligand stands out from the surrounding protein atoms.

GRAPHICS.stickRadius - radius (in Angstroms) of a cylinder displayed as a part of stick or xstick graphical representation of a molecule. Individual (residue-wide) control of stick radii.

GRAPHICS.xstick Backbone Ratio - Default (1.2)

GRAPHICS.xstick Style - `xstick` style

wireBondSeparation the distance between two parallel lines representing a chemical double bond if `wireStyle = "chemistry"`. Default (0.2 Angstroms).

GRAPHICS.distance Label Drag - enable distance label dragging

GRAPHICS.hbondAngleSharpness determines how the strength depends on the D-H...A(lone pair) angle. The preference can be found the general Preferences menu Default (1.7)

GRAPHICS.hbond Ball Style even, by atom size, by energy or telescopic

GRAPHICS.hbond Rebuild

GRAPHICS.hbondWidth relative width of a displayed hbond .

GRAPHICS.hydrogenDisplay determines the default hydrogen display mode for the display command.

```
GRAPHICS.hydrogenDisplay = "polar"
1 = "all" # all hydrogens are shown
2 = "polar" <-- current choice # polar displayed, the non-polar hidden
3 = "none" # no hydrogens are displayed
```

GRAPHICS.wire Width - relative width of wire Default (1)

GRAPHICS.xstick Hydrogen Ratio - Default (0.5)

GRAPHICS.xstick Vw Ratio - Default (0.6)

Wire Style - change the default `wire` style

```
GRAPHICS.hbondStyle = "dash"
1 = "wire" # Just a line
2 = "chemistry" # shows different types of chemical bonds.
3 = "tree" # shows a directed graph of the ICM-molecular tree
4 = "aromatic" #
```

3.7.2 Directories Preferences

DIRECTORIES TAB:

To change Directory Preferences:

- Select File/Preferences.
- Choose the **Directories** tab.

Within this tab you can select the default directories for:

FILTER.gz, **FILTER.uue**, **FILTER.Z**, **Filter.zip** allows you to read compressed files `.gz`, `.uue`, `.Z`, and `.zip` files automatically leaving the compressed file intact.

PDB Directory Style - The style of your Protein Data Bank directory/directories. ICM will understand all of the listed styles, including distributions with compressed `*.gz`, `*.bz2` and `*.Z` files

BlastDB Directory - return directory with Blast-formatted sequence files for ICM sequence searches. You can download Blast formatted databases from here <ftp://ftp.ncbi.nih.gov/blast/db/>

Dock Directory - Default directory for storing docking files.

CCP4 Directory

Editor - Select a default text editor

Inx Directory - location of stored index (`*.inx`) files.

Log Directory - when you quit an icm-session, a `_seslog.icm` file is automatically stored. If the `s_logDir` variable is empty, it is stored to the `s_userDir + "/log/"` directory. However one can redirect it to the current working directory (`."`) or any other directory.

Output Directory -

PDB Directory - directory containing the PDB database of 3D structures. These files can also be easily downloaded directly from the PDB site if the variables are set as in the example below. PDB distributions can exist in several styles (all files in the same directory, or divided etc.).

PDB Directory FTP

PDB Directory Web

Projects Directory - Select the default location for storing ICM projects. Save your data in an ICM project. It is a convenient way of keeping all your structures, alignments, tables, docking results etc... in one place. A description on how to save an ICM project is described in the `GUI Basics` section of this manual.

Prosite Dat - location of the `prosite.dat` file a dictionary of protein sites and patterns, (Copyright by Amos Bairoch, Medical Biochemistry Department, University of Geneva, Switzerland).

Ps Viewer - Select a postscript viewer

Swissprot Dat - location of `swissprot.dat` file

Temp Directory - scratch directory for temporary files (some montecarlo files will be saved there).

Uniprot Dat - location of `uniprot.dat` file

XPDB Directory - Path to the ICM XPDB database of compact binary ICM objects which are annotated with the site information. The advantage of the XPDB database is the speed of reading and smaller size than PDB. XPDB entries are read about 80 times faster!

TOOLS.default ChemDB

TOOLS.eds Directory

TOOLS.pdb Read Nmr Models

1. = "first" : reads only one model from a multi-model (e.g. NMR) pdb file
2. = "all" : reads all models from a multi-model (e.g. NMR) pdb file and creates a separate
3. = "all stack" : creates one object and loads all other models as a stored cartesian stack

3.7.3 Graphics Preferences

To change Graphics Preferences:

- Select File/Preferences.
- Choose the **Graphics** tab.

Atom Single Style - display style of isolated atoms in the wire mode.

1. "tetrahedron"
2. "cross"
3. "dot"

GRAPHICS.clash Style - choose clash length, strain or length.

GRAPHICS.clip Grobs - enable grob clipping.

GRAPHICS.clip Static -

GRAPHICS.grobLineWidth - relative width of displayed lines of 3D meshes (grobs). Also affects the interatomic distance display.

GRAPHICS.lightPosition - X ,Y and Z position of the light source in the graphics window. The X and Y coordinates are usually slightly@@ beyond the [-1. 1] range where [-1.,1.] is the size of the window, and

the Z position is perpendicular to the screen and is set to 2. (do not make it negative).

GRAPHICS.occupancyDisplay preference controlling if and how the partial or zero atom occupancies are displayed. The abnormal occupancies are shown as circles around atoms. These following values are allowed.

```
1. = "none" # nothing is displayed
2. = "circle" # a circle is displayed
3. = "label" # a circle and a label with the value (zero values are not shown)
```

GRAPHICS.quality - integer parameter controlling quality (density of graphical elements) of such representations as cpk, ball, stick, ribbon . Do not make it larger than about 20 or smaller than 1.

GRAPHICS.ruler Style - change ruler from center to side

GRAPHICS.stereoMode - 1. "up-and-down", 2. "line interleaved" 3. "in-a-window"

- a simple hardware stereo mode for workstations with a horizontal frame splitter.
- In the "up-and-down" mode a longer frame with two stereo images on top of each other is generated and the two halves are then superimposed with the splitter. This mode does not require anything from a graphics card, but does require a frame splitter. A frame splitter box was connected between a monitor and a graphics card output. This mode has an unpleasant side effect, the rest of the screen (beyond the OpenGL window) becomes stretched and the lower part of the screen is superimposed on the top half.
- The "line interleaved" mode can be used with a new type of frame splitter at the line level. In this case the odd lines from one stereo-image are interleaved with the even lines of another. The side-effect of this mode is that the intensity is reduced in half since at each moment one sees only one half of the lines. The splitter device for this mode can be purchased from Virex (www.virex.com). This mode produces a dark stereo image but is easily available (requires stereo goggles, e.g. from Virex).
- The "in-a-window" mode is used in SGI workstations and in a Linux workstation with an advanced graphics card supporting a quad graphics buffer. In this mode the hardware stereo regime applies only to an OpenGL window. This is the best mode but it requires an expensive graphics card (plus the stereo goggles).

GRAPHICS.surfaceDotDensity - Determines the number of dots per square Angstrom on the graphical solvent accessible surface.

GRAPHICS.surfaceProbeRadius - An increment to the van der Waals radii of atoms at which the dotted atomic surface is calculated. It is used by the display surface command to display dotted van der Waals surface. If the GRAPHICS.surfaceProbeRadius is set to 1.4 the surface becomes equivalent to the solvent accessible surface with a probe of 1.4Å

GROB.arrowRadius - a real arrow radius in Angstroms used by the Grob("ARROW", R_) function. Default: 0.5.

GROB.contourSigmaIncrement - a real increment in the sigma level used to re-contour an electron density map using the make grob m_ed add r_increment command. This parameter is used in the GUI when plus and minus are pressed.

GROB.relArrow Size - a real ratio of the arrow head radius to the arrow radius. This parameter is used by the Grob("ARROW", R_) function. Default: 3.0.

shineStyle - defines how solid surfaces of cpk , skin and grobs reflect light. Possibilities:

```
1. "white" <- default
2. "color"
```

The first option gives a more shiny and greasy look.

GRAPHICS.center Follows Clipping - determine the function of center button.

GRAPHICS.clashWidth - relative width of a displayed clash .

GRAPHICS.clip Skin - enable skin clipping.

GRAPHICS.displayMapBox - controls if the bounding box of a map is displayed

GRAPHICS.light - a rarray of 13 elements between 0. and 1. which controls the main properties of lighting model in GL.

GRAPHICS.mapLineWidth - relative width of lines and dots of a displayed map.

GRAPHICS.occupancy Radius Ratio - preference controlling the radius of the partial or zero atom occupancies

GRAPHICS.resize Keep Scale

GRAPHICS.selectionStyle - preference for the style in which the graphical selection is shown. The preference may have the following values.

GRAPHIC.store Display - maintains representation and coloring for an object.

GRAPHICS.surfaceDotSize - Determines the size of the dot on the solvent accessible graphical surface.

GRAPHICS.transparency - Two parameters regulating the transparency of grobs.

GROB.atomSphereRadius - default radius (in Angstroms) which is used to select a patch on the surface of a grob.

GROB.relArrowHead - a real ratio of the arrow head radius to the arrow radius.

lineWidth - the real width of lines used to display the wire representation of chemical bonds.

3.7.4 GUI Preferences

GUI TAB:

The options contained within the Preferences/Gui tab are described below.

GRAPHICS.alignment Rainbow - This option controls how alignments are colored by default.

GRAPHICS.NtoC Rainbow - Controls the coloring of structural representation from the N-terminal to the C-terminal

GRAPHICS.rocking - Controls default `rocking motion`.

GRAPHICS.rocking Speed - Controls rocking or rotation speed.

GUI.auto Save Interval - Controls auto save period (minutes)

GUI.table Row Mark Colors - Controls colors used for marking tables.

GUI.workspaceTabStyle - Controls the style of ICM-object tabs created in the `workspace panel` of ICM GUI.

Movie.fade Nof Frames - Controls number of frames for the fade out option in `screenshot movie` making.

Movie.quality - Controls the resoluion of the movie

SEQUENCE.site Colors - Controls coloring of squence sites.

SLIDE.ignore Fog - Fog representations can be ignored in `slide` preparation if desired.

GRAPHICS.discrete Rainbow -

GRAPHICS.rainbow Bar Style - determines if and where the color bar will appear after a molecule is colored by an array.

GRAPHICS.rocking Range - real value of rocking range.

GUI.auto Save - auto save on or off

GUI.max Sequence Length - maximum sequence length displayed in ICM

GUI.workspace Folder Style - Workspace folder style.

MOVIE.frame Grab Mode - with `screenshot movie` making you can choose either fixed frame time or real time.

Movie.quality Auto - with `screenshot movie` making you can allow ICM to control the movie resolution.

SLIDE.ignore Background Color - Ignore background color when you are making a `slide`.

3.7.5 GUI Preferences

To change GUI Preferences:

- Select File/Preferences.
- Choose the **GUI** tab.

Quality - controls the quality (density of graphical elements) of such representations as `cpk`, `ball`, `stick`, `ribbon`. Do not make it larger than about 20 or smaller than 1. We recommend to make this parameter at least 15 if you want to make a high quality image. You can also increase the number of image resolution by making the image window 2,3,4 times larger (in the example below it is 2 times larger) than the displayed window.

Wire Style - Four different wire styles are available.

Hydrogen Display - Select whether you always want all hydrogens displayed or just-polar hydrogens or no hydrogens at all.

Rainbow Scale - determines if and where the color bar will appear after a molecule is colored by an array. Coloring by an array is one of the options of the `display` and `color` commands.

```
1. = "left" <- default choice
2. = "right"
3. = "no text"
4. = "no bar"
```

Ball Ratio - The ratio of ball and stick radii. This ratio is applied when the styles are switched to `xstick` from the GUI `xstick` toolbar.

Selection Style - Change the graphical display of your selections. Default is a green cross.

Clash Threshold - a clash is defined as an interatomic distance less than a sum of van der Waals radii of two atoms of interest multiplied by the `clashThreshold` parameter. For hydrogen bonded atoms, the distance threshold is additionally reduced by 20%. Default = 0.82

DotSurfaceRadiusIncrement - radius of a probe sphere used to display a dotted surface of a molecule. All van der Waals radii are expanded by this value. `vwExpand=0` corresponds to the CPK surface, `vwExpand=1.4` corresponds to the water-accessible surface. Be aware of the difference between the `waterRadius` and `vwExpand` parameters: `waterRadius` is used in

- show energy "sf"
- show [arealvolume] skin
- display skin while `vwExpand` is used in
- show [arealvolume] surface
- display surface

Default (1.4).

H Bond Style - How do you wish your H-Bonds to be displayed by default? Dashes, Bond Length, Bond Length and Angle.

grobLineWidth - relative width of displayed lines of 3D meshes (`grobs`). Also affects the interatomic distance display.

general line width - the real width of lines used to display the wire representation of chemical bonds. See also `IMAGE.lineWidth` parameter which controls line thickness in molecular images generated by the `write postscript` command, and the `PLOT.lineWidth` which controls the width for the `plot` command. Default (1.0)

single atom as - display style of isolated atoms in the wire mode.

1. "tetrahedron"
2. "cross"
3. "dot"

The size of the first two representation is controlled by the `GRAPHICS.ballRadius` parameter and

xstickhetatomzZoom - The default ball and stick radii of a ligand can be different. This makes a better ligand view since the ligand stands out from the surrounding protein atoms.

solid shine style - choose either white or color

Stick Radius - radius (in Angstroms) of a cylinder displayed as a part of stick or `xstick` graphical representation of a molecule. Individual (residue-wide) control of stick radii.

Stereo Mode - Select a default stereo mode

Display Style - A default display style can be chosen using a combination of styles.

Water Radius - radius of water sphere which is used to calculate an analytical molecular surface (referred to as skin) as well as the solvent-accessible surface (centers of water spheres).

clashWidth - relative width of a displayed clash.

hbondWidth - relative width of hydrogen bond display

mapLineWidth - relative width of lines and dots of a displayed map.

3.7.6 Image Preferences

To change Image Preferences:

- Select File/Preferences.
- Choose the **Image** tab.

IMAGE.color - logical to save color or `black_and_white` ('bw') images.

IMAGE.gammaCorrection - real variable to to lighten or darken the image by changing the gamma parameter. A gamma value that is greater than 1.0 will lighten the printed picture, while a gamma value that is less than 1.0 will darken it.

IMAGE.lineWidth - this real parameter specifies the default line width for the postscript lines.

IMAGE.orientation - image orientation.

IMAGE.previewer - a string parameter to specify the external filter which creates a rough binary (pixmap) postscript preview and adds it to the header of the ICM-generated high resolution bitmap or vectorized postscript files saved by the `write image postscript`, and `write postscript`, respectively.

IMAGE.print - unix command for printer.

IMAGE.scale - real variable. If non zero, controls the image scale with respect to the screen image size.

IMAGE.stereoBase - real variable to define the stereo base (separation between two stereo panels) in the `write image postscript` and `write postscript` command.

IMAGE.writeScale - an integer parameter used to increase the image resolution in the `Quick Image Write tool`.

IMAGE.bondLength2D - real length of a chemical bond (in inches) in chemical 2D drawings upon the `Copy Image` command.

IMAGE.compress - logical to toggle simple lossless compression, standard for .tif files. This compression is required to be implemented in all TIFF-reading programs.

IMAGE.generateAlpha - logical to toggle generation of the alpha (opacity) channel for the SGI rgb, tif and png image files to make the pixels of the background color transparent.

IMAGE.lineWidth2D - integer thickness of bonds in chemical 2D drawing upon the Copy Image command. This is useful for cutting and pasting from ICM to external documents.

IMAGE.paper Size - specify paper size.

IMAGE.previewResolution - integer resolution of the rough bitmap preview added to the vectorized postscript file in lines per inch.

IMAGE.printerDPI - this integer parameter the printer resolution in Dot Per Inch (DPI). Important for the write image postscript command.

IMAGE.stereoAngle - real variable to define stereo angle (relative rotation of two stereo images) in the write image postscript and write postscript command.

IMAGE.stereoText - logical to make text labels for only one panel or both panels of the stereo diagram.

3.7.7 Label Preferences

To change Label Preferences:

- Select File/Preferences.
- Choose the **Labels** tab.

atomLabelStyle style of atom labels invoked by clicking on an the `atom label button`.

GRAPHICS.displayLineLabels - enables/disables the display of edge lengths (inter-point distances) of a grob generated with the `Grob("distance" ..)` function.

GRAPHICS.font Line Spacing - Change the spacing between lines in labels.

GRAPHICS.resLabelDrag - if yes, enables dragging of the displayed residue labels with the middle mouse button.

GRAPHICS. site Arrow - Highlight sites with an arrow yes or no.

Show Res Code In Selection - When you make a selection the icm selection language will be displayed when you right click on the selection.

Res Label Style - Default `residue label style`.

SITE.label Style - Default `label sites style`.

Var Label Style - Default `label variable style`.

GRAPHICS.atomLabelShift - a non-negative integer number of spaces preceding an atom label. This parameter is useful for displaying labels next to a solid representation,

GRAPHICS.fontColor - set font color

GRAPHICS.font Scale - set font size

GRAPHICS.site Label Shift - GRAPHICS.resLabelShift a non-negative integer number of spaces preceding a site label.

GRAPHICS. site Label Drag - if yes, enables dragging of the displayed site labels with the middle mouse button.

Res Label Shift - a non-negative integer number of spaces preceding a residue label. This parameter is useful for displaying residue labels next to a solid

SITE.labelOffset - (default 5. A) the real offset of the site label with respect to the residue label atom.

SITE.wrap Comment - Number of characters per comment line.

3.7.8 Plot Preferences

To change Plot Preferences:

- Select File/Preferences.
- Choose the **Plot** tab.

PLOT.color - logical to generate a color plot. Usually it does not make sense to switch it off because your b/w printer will interpret the color postscript just fine anyway.

PLOT.draw Tics logical yes or no

PLOT.fontSize real font size. Any reasonable number from 3. (1 mm, use a magnifying glass then) to 96.

PLOT.lineWidth - real line width for graphs (not the frame and tics)

PLOT.markSize - real mark size in points. Allowed mark types: line, cross, square, triangle, diamond, circle, star, dstar, bar, dot, SQUARE, TRIANGLE, DIAMOND, CIRCLE, STAR, DSTAR, BAR. Uppercase words indicate filled marks.

PLOT.paper Size - preference to specify plot paper size

PLOT.rainbowStyle - preference defining the color spectrum used by the plot area command.

PLOT.Yratio - real aspect ratio of the ICM plot frame. Using link option of the plot command is equivalent to setting this variable to 1.0. If PLOT.Yratio is set to 0. , the ratio will be set automatically to fill out the available box optimally.

[PLOT.date] - display date on plot

PLOT.font - preference for the title/legend font.

PLOT.labelFont - preference for the data point label font.

PLOT.logo - logical switch for the ICM-logo on the plot.

PLOT.orientation - preference for the plot orientation.

PLOT.previewer - command to local ps viewer

PLOT.seriesLabels - preference to indicate position of a series/color legend inside the plot frame.

3.7.9 Ribbon Preferences

To change Ribbon Preferences:

- Select File/Preferences.
- Choose the **Ribbon** tab.

Combo Display Style - select ribbon-cpk, atoms, ribbon-ligand, chemical

GRAPHICS.dnaRibbonRatio - real ratio of depth to width for the DNA ribbon .

GRAPHICS.dnaRibbonWorm - logical which, if yes, makes the DNA backbone ribbon round, rather than rectangular. Default: no

GRAPHICS.dnaWormRadius - real radius of the worm representing bases in DNA ribbon .

GRAPHICS.ribbonWidth - real width of the protein ribbon .

GRAPHICS.wormRadius - radius of coiled segments (i.e. those where the secondary structure is marked as "_") of a polypeptide chain in ribbon representation. Default (0.3).

Ribbon Style - specifies default style when ribbon is displayed.

GRAPHICS.dnaBallRadius - DNA bases in ribbon representation are shown as balls controlled by this real parameter.

GRAPHICS.dnaRibbonWidth - real width (in Angstroms) of the DNA ribbon .

GRAPHICS.dnaStickRadius - real radius of the sticks representing bases in DNA ribbon .

GRAPHICS.ribbonRatio - real ratio of depth to width for the protein ribbon .

GRAPHICS.ribbonWorm - logical parameter, if yes, makes the ribbon round, rather than rectangular.

ribbonColorStyle -

- sets the ribbon coloring scheme.

1 = "type" default. colors by secondary structure type or explicit color

2 = "NtoC" colors each chain gradually blue-to-red from N- to C- (or from 5' to 3' for DN

3 = "alignment" if there is an alignment linked to a protein, color gapped backbone regions gr

4 = "reliability" 3D gaussian averaging with selectSphereRadius of alignment strength in

If ribbonColorStyle equals to 4, the conserved areas will be colored blue, while the most dive

3.7.10 Shell Preferences

To change Shell Preferences:

- Select File/Preferences.
- Choose the **Shell** tab.

Clash Threshold - a clash is defined as an interatomic distance less than a sum of van der Waals radii of two atoms of interest multiplied by the clashThreshold parameter.

Map Sigma Level - (in Rmsd values over the mean value). Margin value used for making graphical objects contouring the 3D density map .

Mnconf - maximal number of conformations in the conformational stack . The stack stops growing after this number is achieved and starts replacing representative conformations with higher energy values by new conformations with superior energies, if the latter are found.

Icm Prompt - defines the ICM-prompt string.

Select Min Grad - default minimal gradient vector length for gradient atom selection (a_//G). This parameter is also used by the montecarlo fast command, which requires a value of 2. to 10. for optimal performance.

Map Atom Margin - Margin in Angstroms around selected atoms. The margin is added to the positional boundaries to define a submap index box in the Map (map_source , as_) function.

maxColorPotential - local electrostatic potential in kcal/e.u.charge units at which the surface element is colored by extreme red or extreme blue. All higher values will have the same color. This absolute scaling is convenient to develop a feeling of electrostatic properties of molecular surfaces.

mnSolutions - this parameter limits the number of hits retained by the program after a search.

Real Format - format of real numbers

Water Radius - radius of water sphere which is used to calculate an analytical molecular surface

3.7.11 System Preferences

To change System Preferences:

- Select File/Preferences.

- Choose the **System** tab.

FTP.createFile -

FTP.proxy - string path to the proxy server for connections through firewall. Default: "" (empty string).

GUL.max Nof Recent Files - maximum number of `recent files` stored.

GUL.splash Screen Image - path to splash image displayed on startup

HTTP.support Cookies - http support cookies yes or no

HTTP.user Agent - client application used within a particular network protocol for www

Beep - warning beep yes or no

Max File Size Mb - Maximum file size in MegaBytes that can be loaded into ICM.

USER.friends

USER.organization

FTP.keep File - (default no). If yes, the temporary file is kept in the `s_tempDir` directory. Otherwise the file is deleted.

GUL.enumeration Memory Limit - memory limit for enumeration operations.

GUL.splash Screen Delay

HTTP.proxy - string for HTTP server for connection through firewall

HTTP.timeout - timeout in seconds

Http Read Style `icm` or `lynx`

Force Auto Bond Typing - yes/no

USER.email, USER.full Name, USER.phone

4 Protein Structure



Chapter Contents:

- How to Convert Proteins and Chemicals to ICM Objects.
- Binding Site Display Tools.
- Crystallographic Analysis Tools.
- Protein Superposition.
- Protein Structure Analysis.
- Protein Structure Prediction Tools e.g. ICM Pocket Finder.

4.1 Convert to ICM Object

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro | ICM-Chemist

4.1.1 Load a Protein Structure

There are a couple of ways to read into ICM a protein structure.

1. Search the PDB using the [Search Tab](#).
2. Load in a PDB file that you have saved on your computer using [File/Open](#).

4.1.2 Converting PDB Files Into ICM Objects

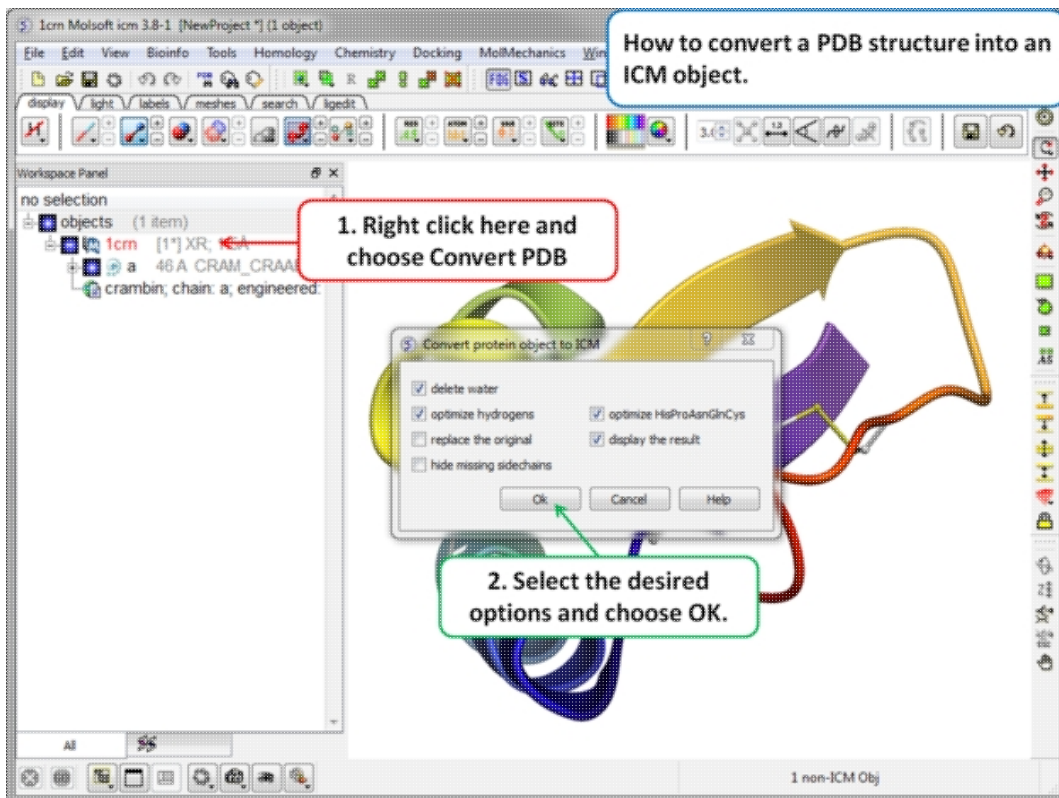
If you are going to make any type of energy calculation in ICM (eg docking, display H-bonds, display electrostatic and binding property surfaces etc..) it is necessary to convert a protein or chemical into an ICM object. Be aware that upon conversion ICM adds missing side-chain atoms (but wont try to build missing loops) due to the nature of the internal-coordinate system. The list of residues/ atoms added is presented in the command line shell and can be reviewed. For reference, the original PDB entry is kept in the system. See the command line manual for a more complete description of what the conversion process does.

NOTE: Before converting a protein structure to an ICM object make sure that the chemicals contained within the structure (e.g. ligands) are correct. If an error is found you can edit the ligand as described [here](#).

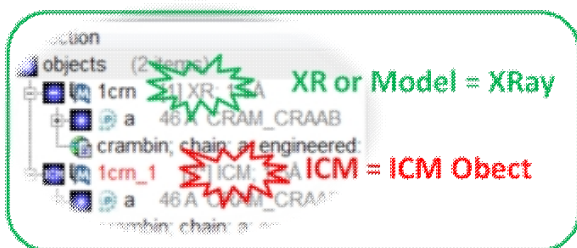
To convert a PDB structure into an ICM object follow the steps shown below:

- Right click on the name of the protein you wish to convert in the ICM [Workspace](#).
- A dialog box will be displayed as shown below.
- If you want to delete the water molecules select **Delete Waters**
- If you want to optimize hydrogen atoms (recommended for important work) select **Optimize Hydrogens**. This option performs global optimization of hydrogens to find the best hydrogen bonding network.

- If you want to optimize the orientation of His, Pro, Asn, Gln, Cys residues then choose **optimizeHisProAsnGlnCys**. The following residues will be further optimized: His - three protonation states and two rotations will be tried and the residue will be renamed according to its subtype: hie (epsilon tautomer) or hip (+). Asn and Gln - (a 180 deg. flip will be tried). Cys - in the vicinity of Zn, Cu, Fe and Co to cym.
- If you want to keep a copy of your PDB file uncheck the option **replace original**.
- The converted structure can be displayed immediately by checking **display the result**
- **Uncheck** the box **hide missing side chains** if you want ICM to build missing heavy atoms that are not reported in the PDB (due to the lack of density), they will be added according to the residue name and assigned zero occupancies. **Check** this box if you want residues missing heavy atoms to be hidden.



If your object is an ICM object it will display **ICM** next to the molecule in the ICM Workspace.



4.1.3 How to Edit and Convert a Chemical from the PDB into an ICM Object

The protein data bank has not been storing any information about covalent bond types and formal charges of the chemical compounds interacting with proteins! This oversight makes it impossible to automatically convert those molecules to anything sensible and requires your manual interactive assignment of bond types and formal charges for each compound in a pdb-entry. Therefore, if you apply the convert command to a pdb-entry with ligands, the ligands will just become some crippled incomplete molecules that can not be further conformationally optimized.

Therefore, follow these steps to convert a chemical properly from a pdb form to a correct icm object. There are two ways to do this either via the ICM Workspace (recommended) or via the Graphical Display.

4.1.3.1 Converting a Chemical from the PDB using the ICM Workspace

- **File/Open PDB**
- View the ligand in the ICM Workspace by expanding the molecule tree (see below).

ICM Workspace

The screenshot shows the ICM Workspace interface. At the top, there is a status bar that says "no selection". Below it is a tree view of objects:

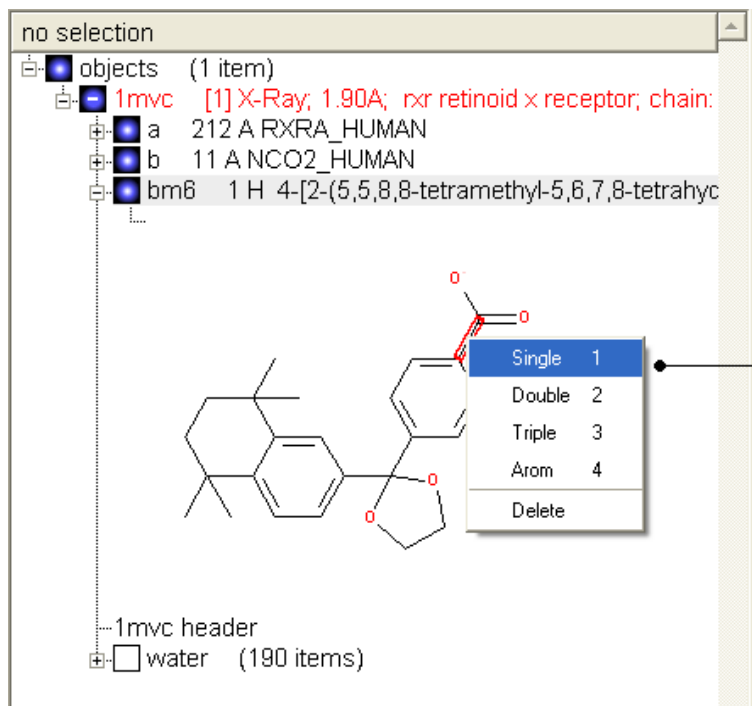
- objects (1 item)
 - 1mvc [1] X-Ray, 1.90Å; rxr retinoid x receptor; chain:
 - a 212 A RXRA_HUMAN
 - b 11 A NCO2_HUMAN
 - bm6 1 H 4-[2-(5,5,8,8-tetramethyl-5,6,7,8-tetrahydropyridin-2-yl)phenyl]propanoic acid

Below the tree view, there is a 3D ball-and-stick model of the chemical structure of the ligand. A red arrow points from the text "Click here to expand" to the expand icon next to the "bm6" object in the tree view.

At the bottom of the workspace, there is a "1mvc header" and a "water (190 items)" entry.

Change bond orders:

- Change the bond orders by selecting the bond (highlighted in red).
- Right click and select the desired bond as shown below.

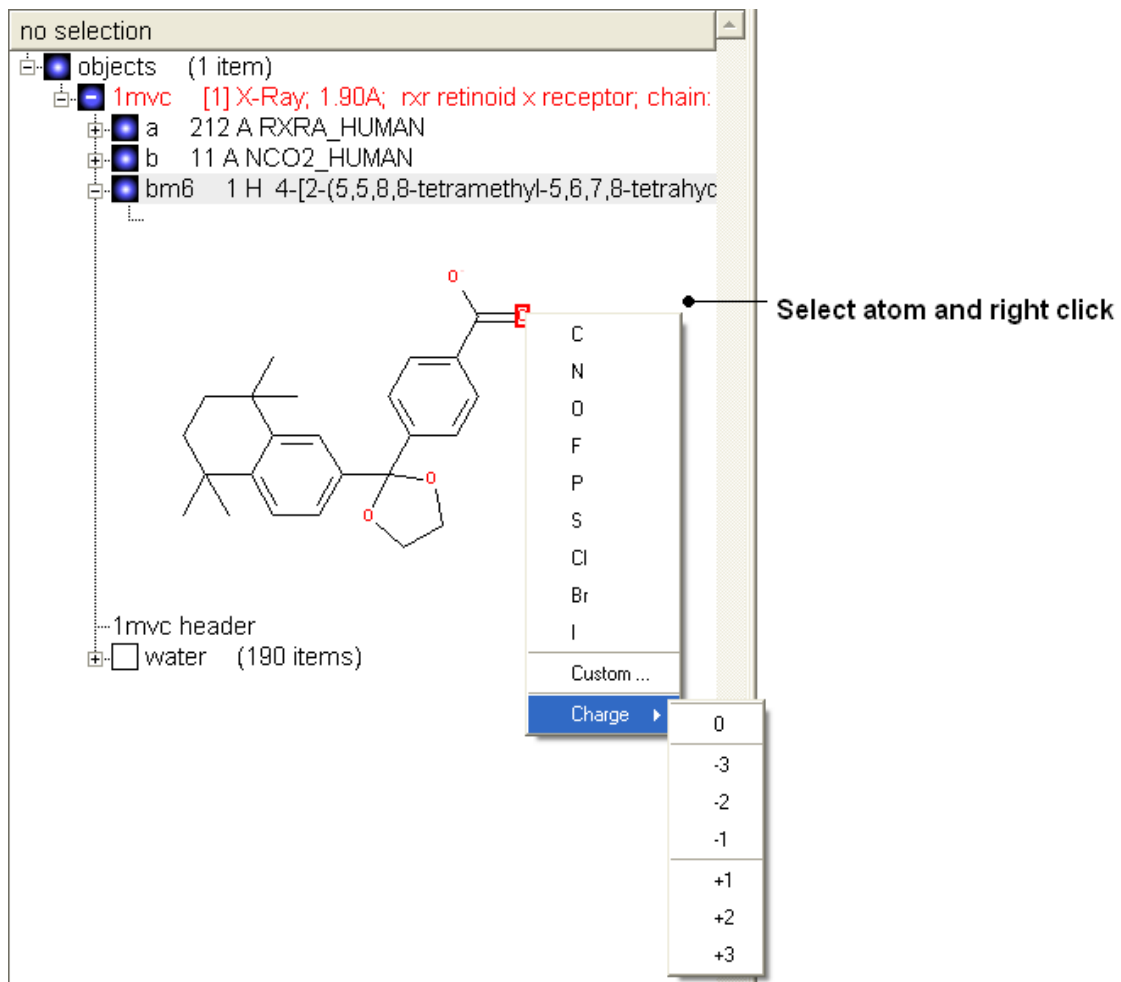


Select bond and right click

NOTE: Keyboard shortcuts are provided to make editing faster.

Change atom and charge:

- Change the atom or charge by selecting the atom (highlighted in red).
- Right ckick and select the desired atom or charge as shown below.



NOTE: Keyboard shortcuts are provided to make editing faster.

Convert to 3D in MMFF force field:

- Once you have made the changes to the ligand - right click on the name of the ligand in the ICM Workspace and select **Move from Object**.

no selection

objects (1 item)

- 1mvc [1] X-Ray, 1.90Å; rxr retinoid x receptor; chain:
- a 212 A RXRA_HUMAN
- b 11 A NCO2_HUMAN
- bm6 1 H 4-[2-(5,5,8,8-tetramethyl-5,6,7,8-tetrahydropyridin-2-yl)phenyl]butanoic acid

1mvc header

water (190 items)

Right click and select "Move from Object"

- Select the ligand by double clicking on it in the ICM Workspace.
- Select **MolMechanics/ICM-Convert/Chemical**

File Edit View Bioinfo Tools Homology Docking MolMechanics Windows Help

ICM-Convert

- Protein...
- Chemical...
- Re-root Compound...

Optimize H,His,Asn,Gln,Pro

Regularization...

Edit Structure

MMFF

Minimize

Sample Loop...

View Stack

Energy Terms...

1 non-ICM Obj

objects (2 items)

- 1mvc [1] X-Ray, 1.90Å; rxr retinoid x receptor; chain:
- a 212 A RXRA_HUMAN
- b 11 A NCO2_HUMAN

1mvc header

water (190 items)

- 1mvc1 [2] X-Ray, 1.90Å; rxr retinoid x receptor; chain:
- bm6 1 H 4-[2-(5,5,8,8-tetramethyl-5,6,7,8-tetrahydropyridin-2-yl)phenyl]butanoic acid

1mvc1 header

Double click here to select ligand

NOTE: If you need to add an extra bond you will need to use the full molecular editor. Right click on the name of the ligand in the ICM Workspace and select **Edit/Edit Compound**.

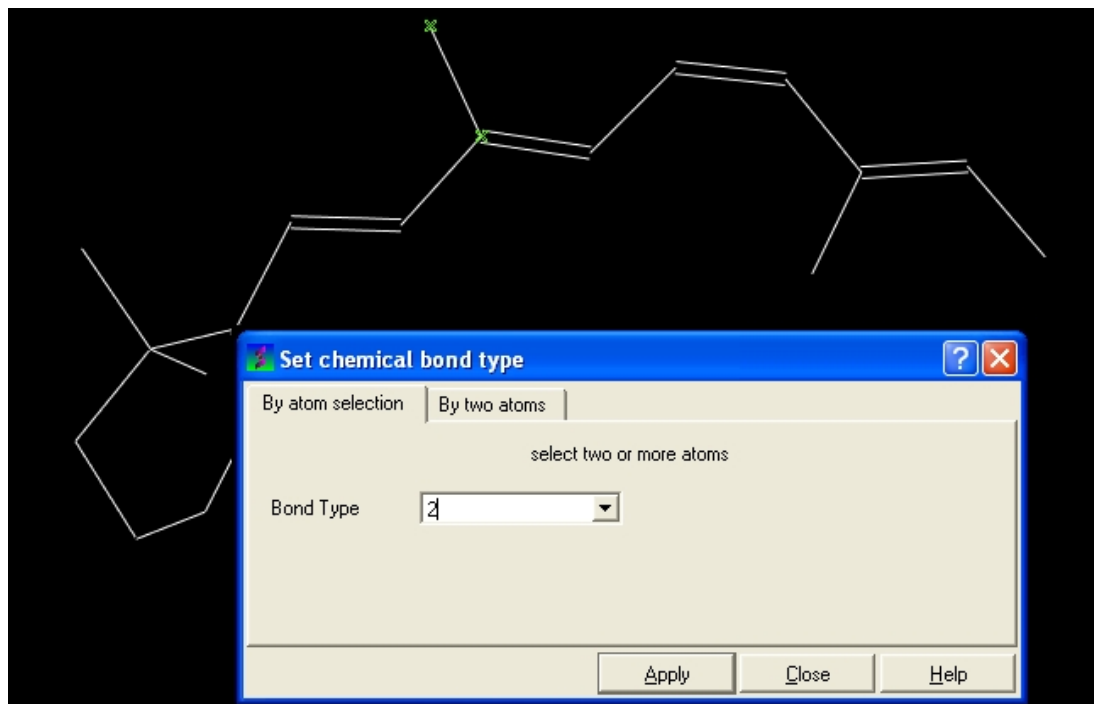
4.1.3.2 How to Convert a Chemical from the PDB using the Graphical Display

- Display the molecule in wire chemistry style mode by right clicking on the Wire Representation button (see Wire Representation section).

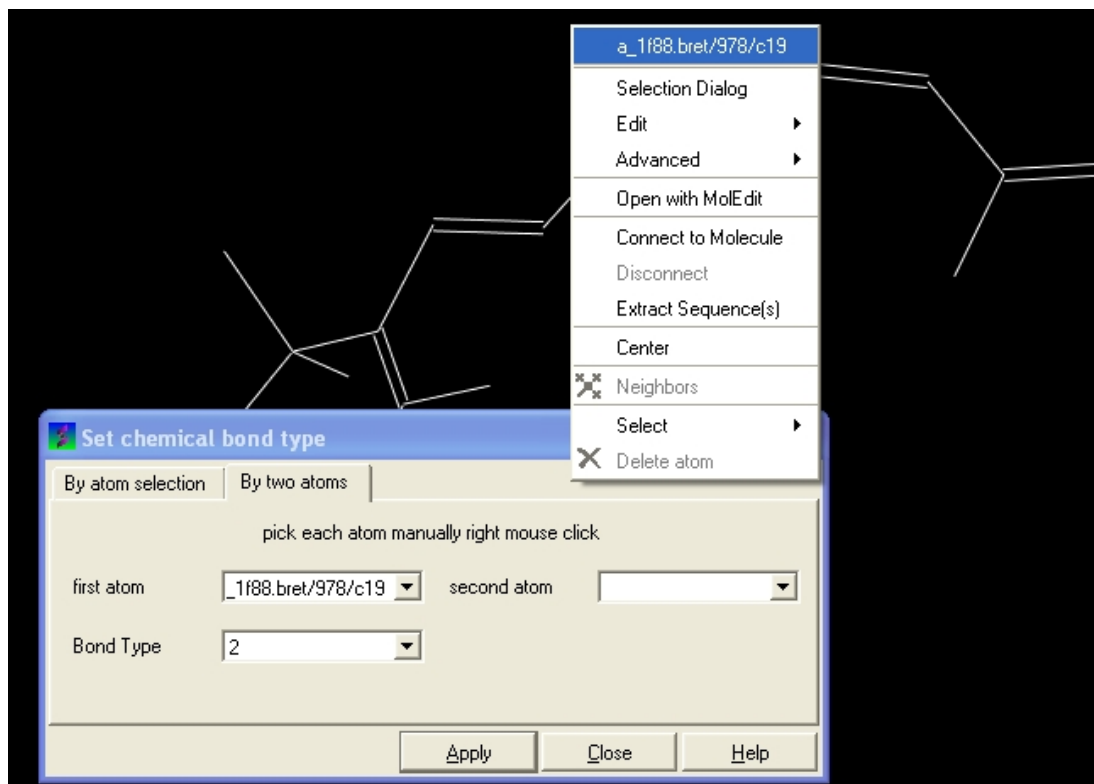
To change the bond types in your ligand:

- Click on MolMechanics/Edit Structure/Set Bond Type and the Set chemical bond type data entry box will be displayed.

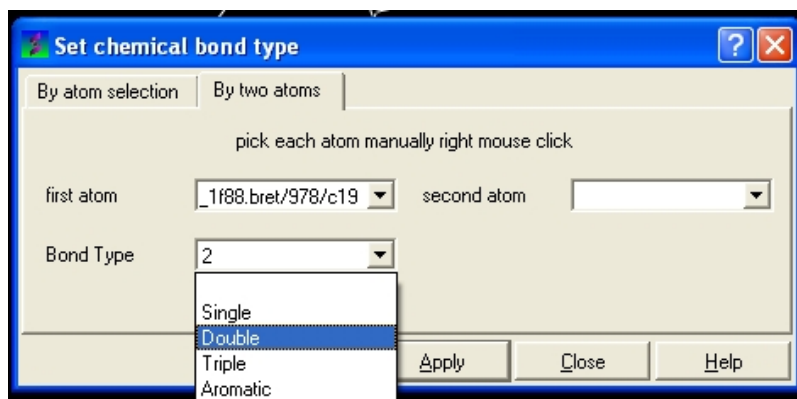
You can either select (see selection menu section) the atoms you wish to change graphically using the rectangular or lasoo selection button OR



You can select the **By two atoms** tabs and right click on the atoms you wish to change and then selecting the atom descriptor with the left mouse button as shown below.

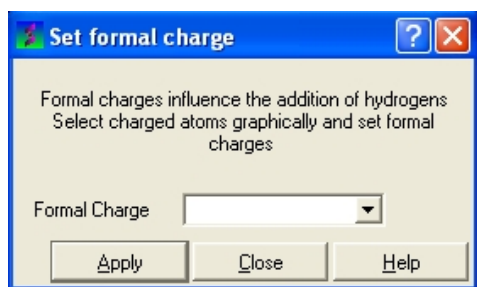


- Select the desired bond type either single, double, triple or aromatic.



To set the formal charge of a compound:

Click on MolMechanics/Edit Structure/Set Formal Charge and then select the appropriate charge.



The final step is to convert the compound into an ICM object:

- Select the chemical (green crosses in graphical display).

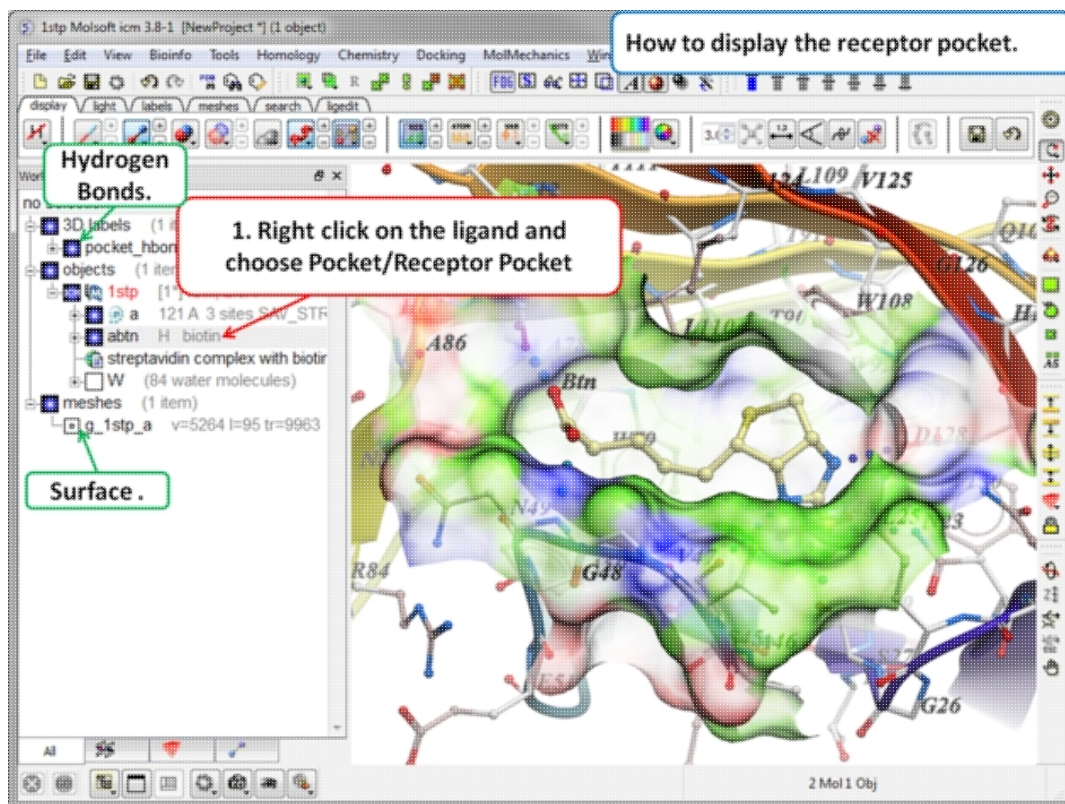
4.2 Pocket Display

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

4.2.1 Receptor Pocket Surface.

To display the Receptor Pocket Surface:

- As an example we will use the PDB structure 1STP. Type 1STP in the `pdb search` tab and press return.
- Convert the protein to an ICM object. If you do not convert a generic surface will be displayed that is not colored by binding property.
- Right click on the ligand "abtn" and select **Pocket/Receptor Pocket**.
- Select whether you would like to display side-chain hydrogen bonds and label.
- The receptor pocket will be displayed colored by binding property - White=neutral surface Green=hydrophobic surface Red=hydrogen bonding acceptor potential Blue=hydrogen bond donor potential.
- The surface can be toggled on and off by selecting in the ICM Workspace in the **meshes** section.

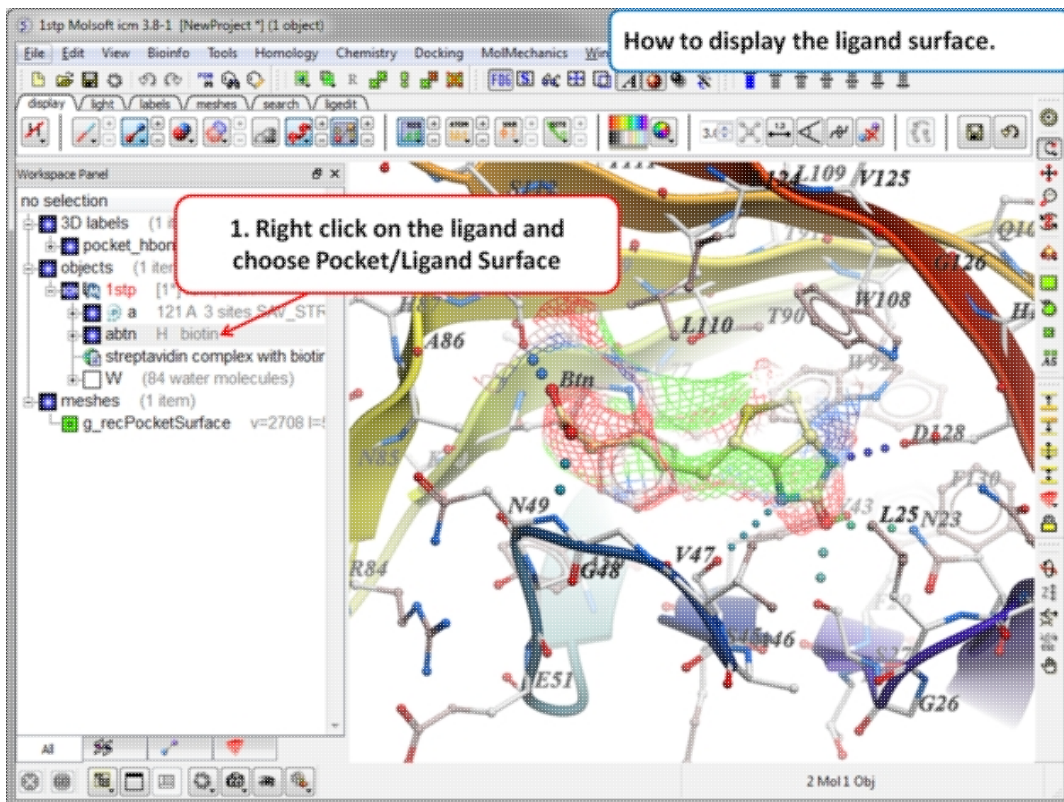


4.2.2 Ligand Surface.

The Ligand Surface option allows you to visualize cavities that are open for ligand modifications. To display the Ligand Surface:

- As an example we will use the PDB structure 1STP. Type 1STP in the `pdb search` tab and press return.
- Convert the protein to an ICM object. If you do not convert a generic surface will be displayed that is not colored by binding property.
- Right click on the ligand "abtn" and select **Pocket/Ligand Surface**.

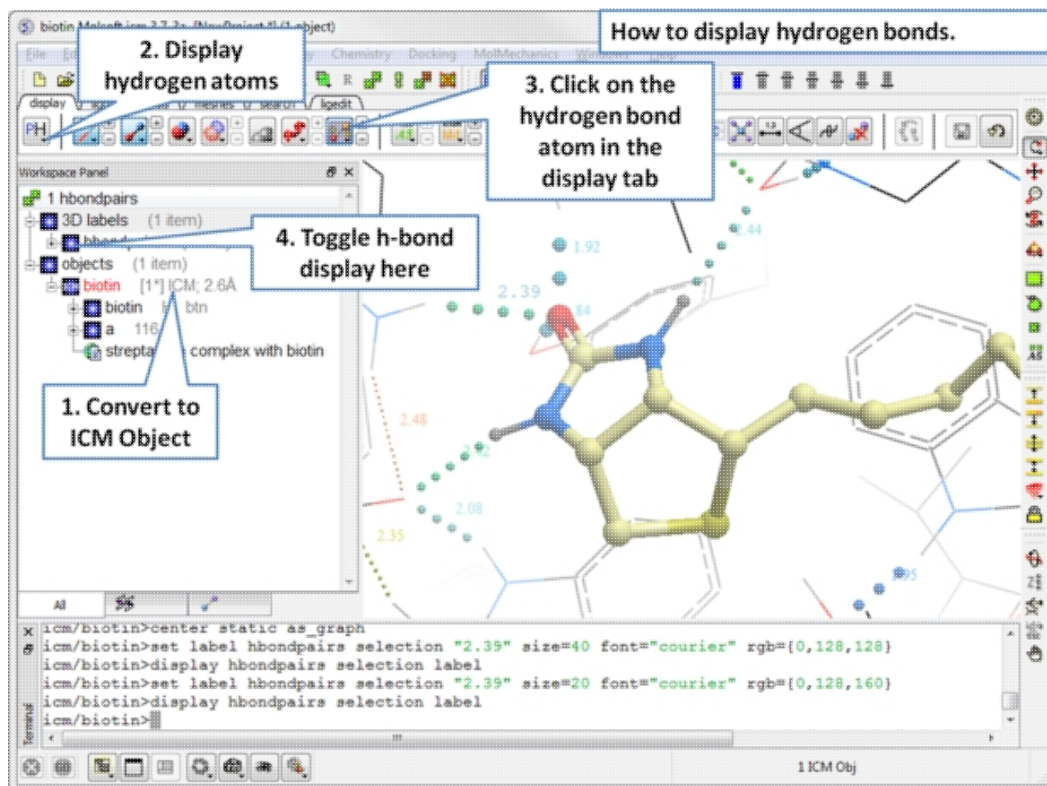
- The Ligand Surface will be displayed colored by binding property - White=neutral surface Green=hydrophobic surface Blue=hydrogen bonding acceptor potential Red=hydrogen bond donor potential.



4.2.3 How to Display Hydrogen Bonds

NOTE: The method by which hydrogen bonds are calculated is described [here](#) in the command line manual. The GRAPHICS.hbondMinStrength parameter determines the hbond strength threshold for hbond display. The strength value is between 0. and 2. By changing 1. to 0.2 you will see more weak hydrogen bonds.

- In order to display energy related properties we need to convert the PDB file into an ICM object. Convert 1STP into an ICM object. In this example, the option "Replace the Original" was selected.
- Display the receptor in wire format and the ligand in xstick.
- Right click on the ligand and select "Neighbors" - Enter 3 Angstroms and Type = Visible. Do not exclude source (the ligand) therefore remove tick from box entitled "exclude source".
- Select the display tab and then select the Display H-Bond button.

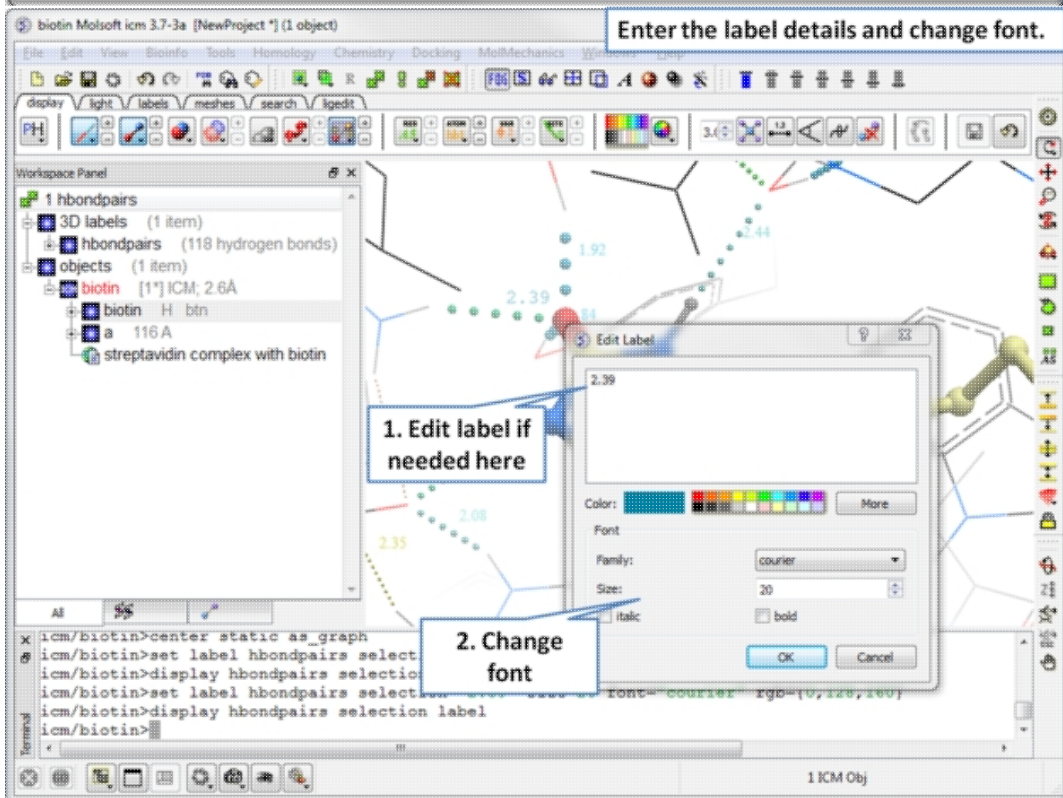
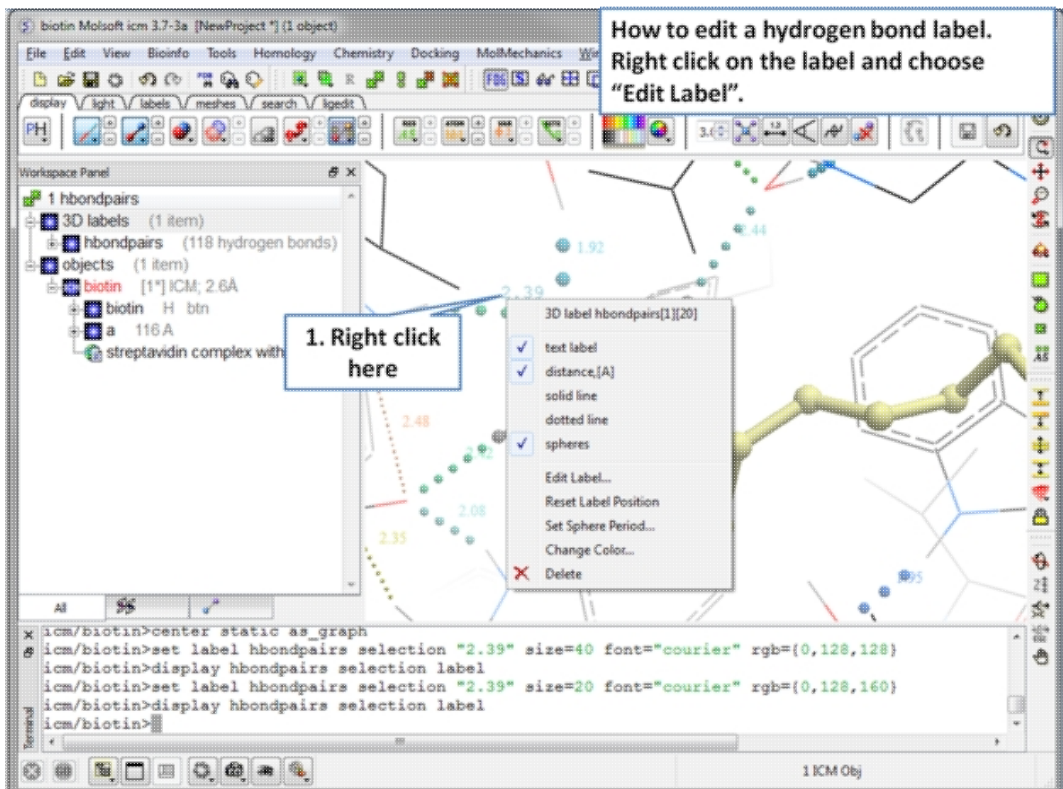


NOTE: Different options for displaying the H-bond can be accessed by clicking and holding on the H-bond button in the "Display" tab. The coloring of the H-bonds are red (strong - thick spheres) to blue (weak - thin spheres). Once the hydrogen bonds have been displayed they can be displayed and undisplayed in the 3D labels section of the ICM Workspace (left hand side of graphical window).

4.2.3.1 Edit Hydrogen Bond Label

To edit a hydrogen bond label

- Right click on the label.
- Choose Edit Label.
- Make changes in the dialog box and press OK.

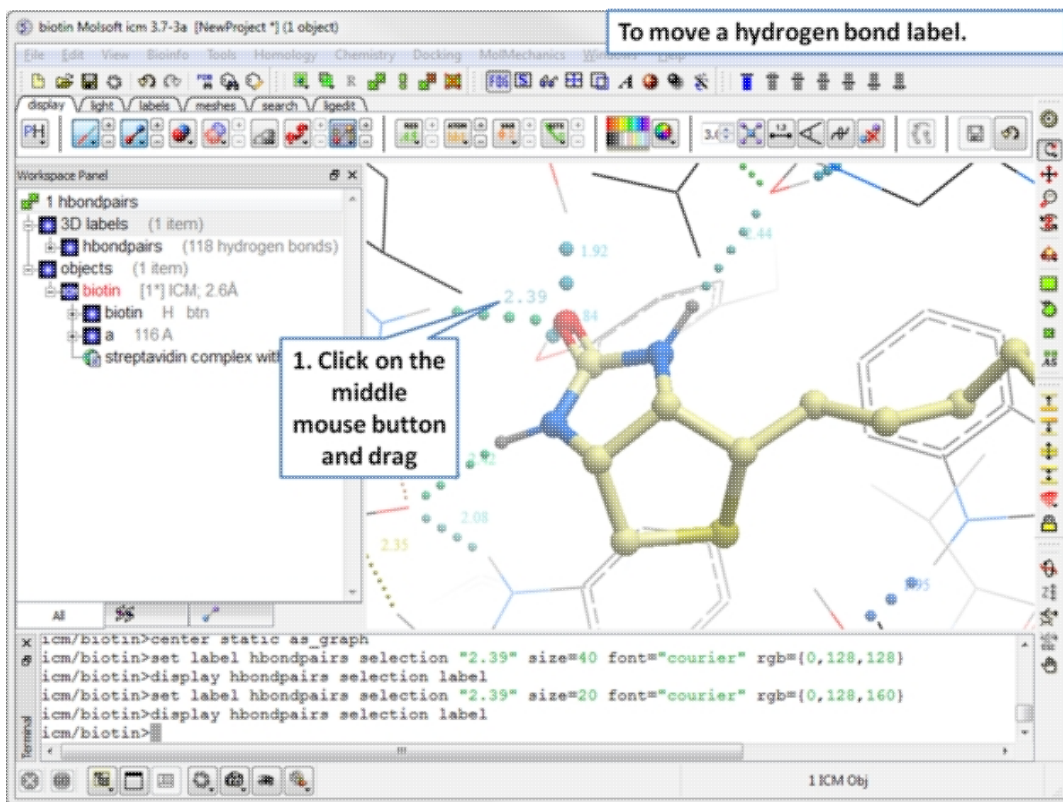


4.2.3.2 Move Hydrogen Bond Label

To move a hydrogen bond label

- Right click on the label.
- Click on the label using the middle mouse button.

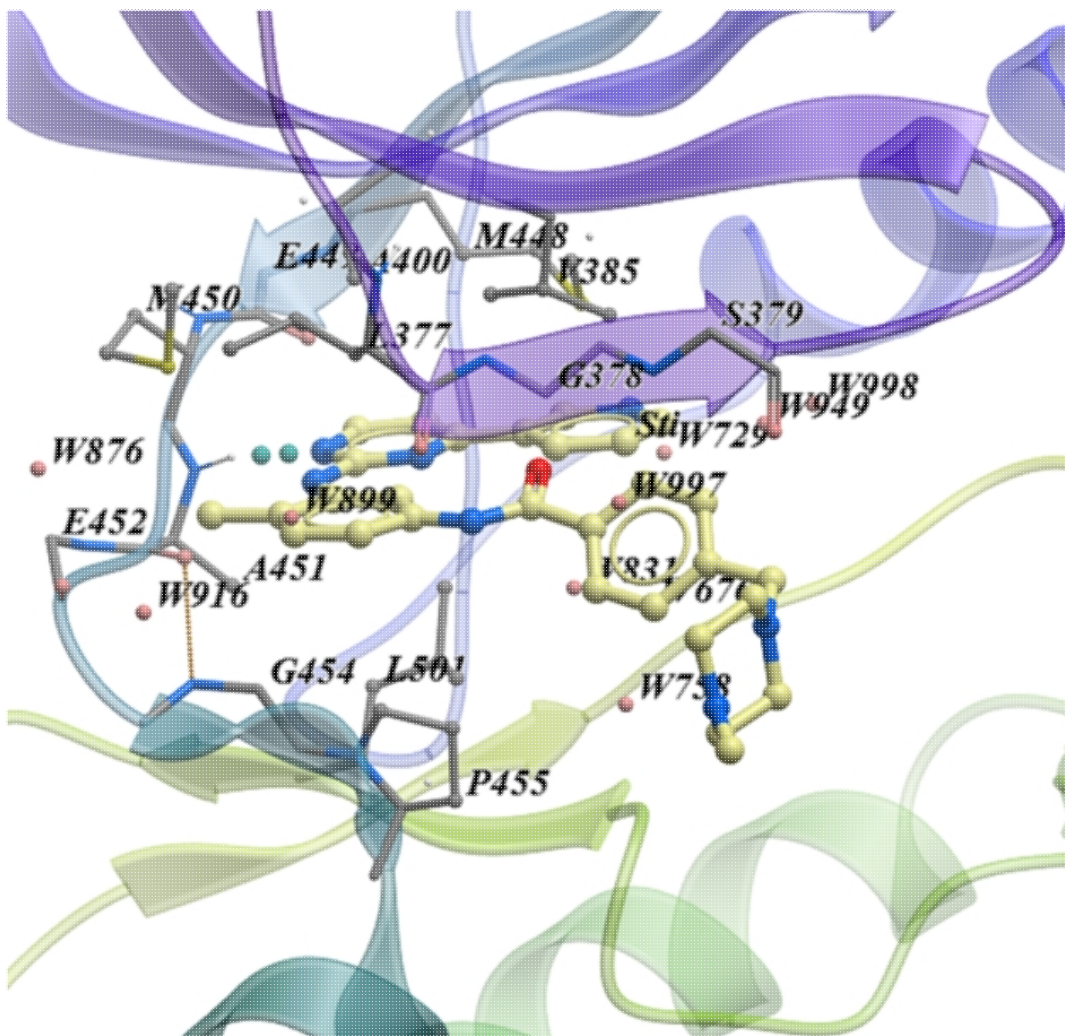
- Hold the middle mouse button down and drag.



4.2.4 Ligand Pocket Interactions

To view the ligand pocket interactions.

- Right click on the ligand and choose Pocket.
- Choose Ligand Pocket Interactions. Hydrogen bonds, key waters and neighboring residues will be displayed.



4.3 Crystallographic Analysis

Available in the following product(s): ICM-Browser-Pro | ICM-Pro

4.3.1 Crystallographic Neighbor

Theory

Molecular objects and 3D density maps may contain information about crystallographic symmetry. It consists of the following parameters:

1. Crystallographic group eg. P2121 that determine N (depends on a group) transformations for the atoms in the asymmetric unit.
2. Crystallographic cell parameters A, B, C, Alpha, Beta and Gamma

To generate the coordinates within one cell one needs to apply N transformations and then to generate neighboring cells the content of one cell needs to be translated in space according to the cell position.

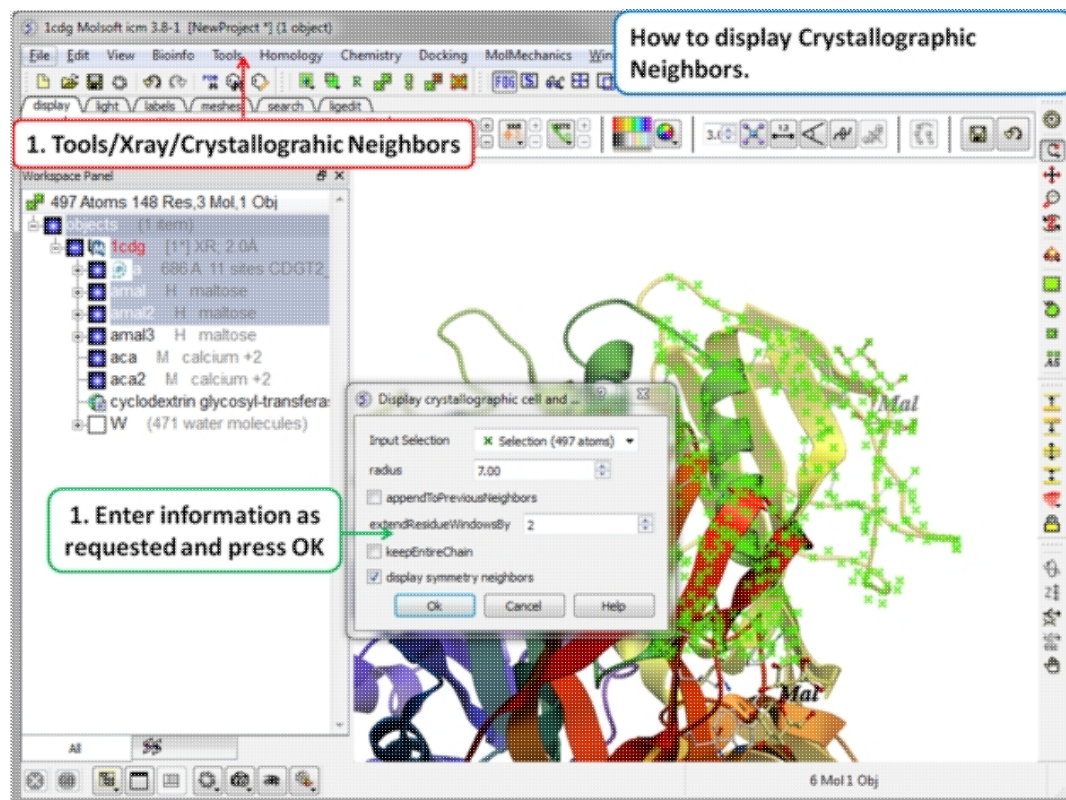
ICM has a function which generates crystallographic neighbors for the selected atoms. For large proteins it is impractical to generate neighbors for the whole molecule due to the high number of atoms in all neighboring molecules.

This information allows to generate symmetry related parts of the density or molecular objects.

To generate symmetry related molecules around a selection of atoms:

- Read a PDB file into ICM. For instruction see the section entitled *Search PDB*.
- Display the structure and select the residues around which the symmetry will be generated. For information on how to select residues see the *Making Graphical Selections* section.
- Select the menu *Tools/Xray/Crystallographic Neighbors*.

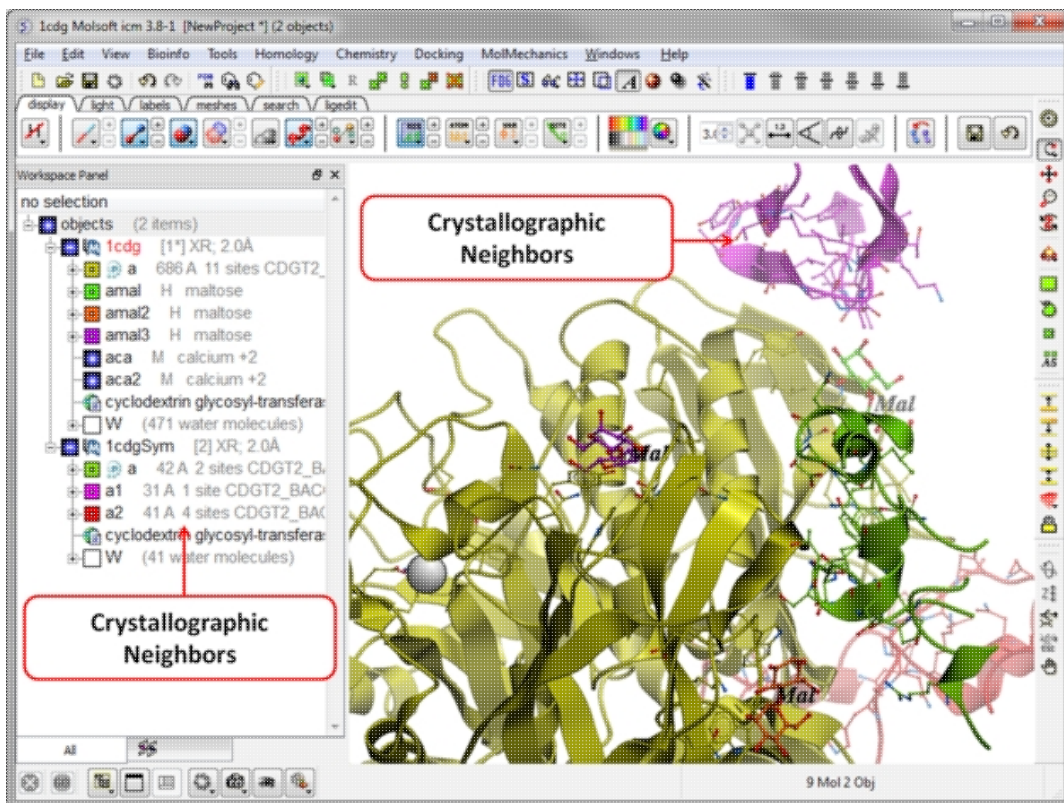
A data entry box as shown below will be displayed.



- Check that the selection is correct in the **Input Selection** box.
- Enter the radius around your selection from which you wish to construct the symmetry related molecules.
- If you have made symmetry related molecules previously you can select **appendToPreviousNeighbors** otherwise leave unchecked.
- The **extendResidueWindowsBy** option will allow a window of residues outside of the selection radius selected above to be displayed
- If you leave the **keepEntireChain** unchecked then a fragment of each neighbor will be created. If you check this box the full neighbor will be generated
- Check **display symmetry neighbors** to display them in the graphics window. The nearest neighbor residues will be displayed in xstick representation and the each neighbor colored by molecule.
- Click OK.

The crystallographic symmetry neighbors will be displayed in the Workspace. By default the object will have the object name + "Sym" and each of the neighbors will be individual molecules.

For packing analysis and display you can color each symmetry unit a different color as described in the *Structural Representations Color* section. This is shown in the picture below.



4.3.2 Crystallographic Cell

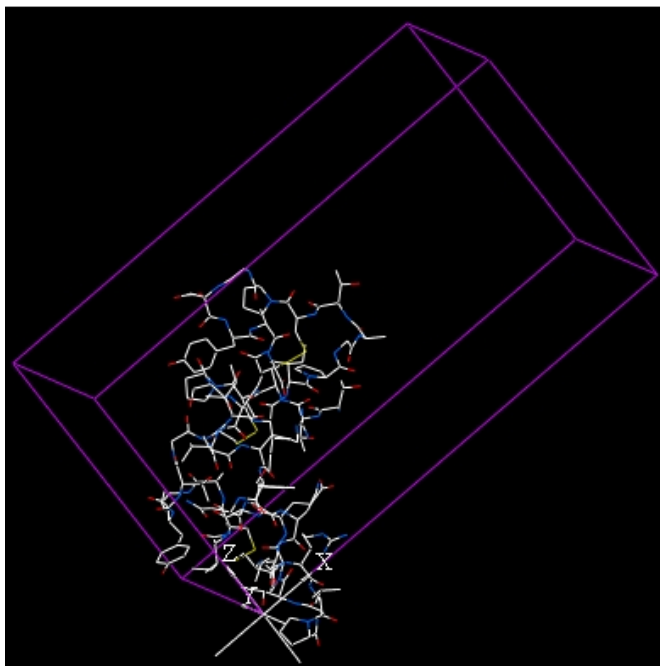
Theory

The crystal structure of a protein is often discussed in terms of its unit cell. The unit cell is a box containing one or more motifs, a spatial arrangement of atoms. The units cells are tiled in three-dimensional space to describe the crystal. The unit cell is given by its lattice parameters, the length of the cell edges and the angles between them, while the positions of the atoms inside the unit cell are described by the set of atomic positions measured from a lattice point.

To display the crystal cell of a PDB structure:

- Read a PDB file into ICM. For instruction see the section entitled `Search PDB`.
- Select the whole object. You can do this by double clicking on the name of the structure in the ICM Workspace (a selection is highlighted blue in the ICM Workspace and green crosses in the graphical display) or you can use the right-click button and drag it over the whole structure in the graphical display.
- Select the menu `Tools/Xray/Crystallographic Cell` and a data entry box will be displayed.
- Click OK

The crystallographic cell will be displayed as a box as shown below.



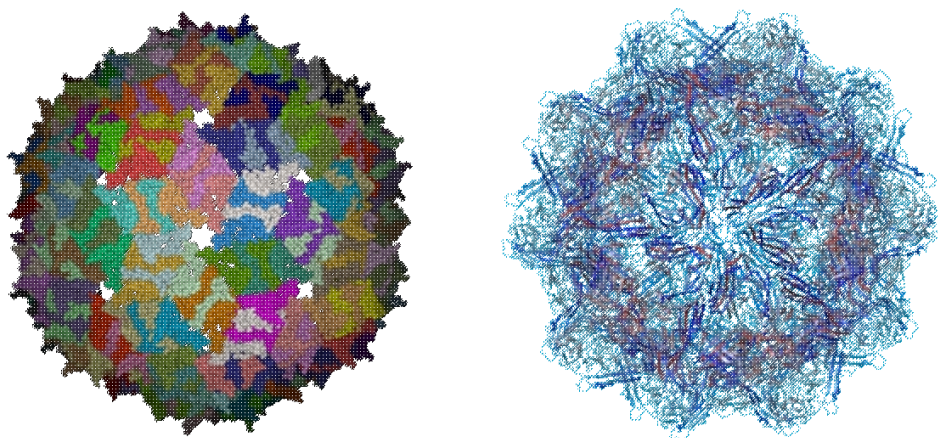
4.3.3 Biomolecule Generator

Theory

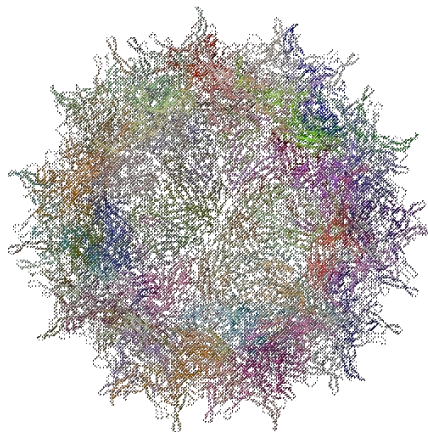
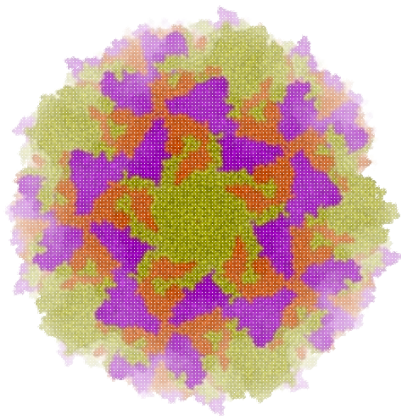
It is very useful to know how a protein from the PDB may look in a biological environment. The PDB entries solved by X-ray crystallography and deposited in the PDB contain the information about the crystal structure rather than the biologically relevant structure. For example, for a viral capsid only one instance of capsid protein complex will be deposited and only one or two molecules of haemoglobin that is a tetramer in solution maybe deposited.

In some other cases the asymmetric unit may contain more than one copy of a biologically monomeric protein. ICM reads the biological unit information and has a tool to generate a biological unit. Not every PDB entry has the biological unit information.

A gallery of images created using the ICM Biomolecule generator is shown below:



Left: PDB: 1DWN Bacteriophage Pp7 From Pseudomonas Aeruginosa At 3.7 A Resolution **Right:** PDB: 1C8E Feline Panleukopenia Virus Empty Capsid Structure At 3.0 A Resolution



Left: PDB: 1AL2 P1/Mahoney Poliovirus, Single Site Mutant V1160I At 2.9 Å Resolution **Right:** PDB: 1LP3 Adeno-Associated Virus (Aav-2), A Vector For Human Gene Therapy At 3.0 Å Resolution

NOTE: Right click on a PDB structure in the ICM workspace to determine whether a structure from the PDB has biological unit information. If it does have this information then there will be an option in the menu entitled "Generate Biomolecules" if not the option will be blanked out.

To generate a biological unit with ICM:

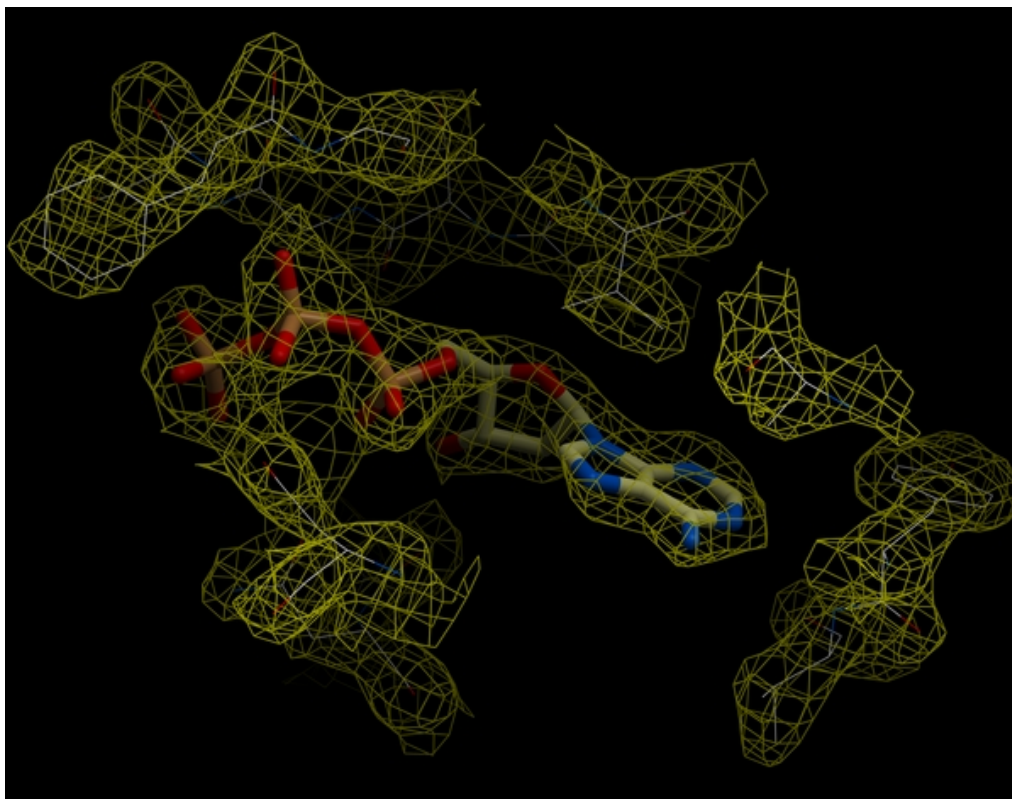
- Select the object or PDB file.
- Select the menu Tools/Xray/Biomolecule Generator.
- Tick the **makeAllBiomolecules** box.
- Click OK - with very large molecules the biomolecule generation may take some time.

4.3.4 Get Electron Density Map

Theory

An electron density map is a representation of a crystal structure based on the diffraction data. The map is constructed by a summation of waves of known phase, amplitude and frequency using Fourier transform. The electron density map of a protein can be viewed along with the pdb structure. The easiest way to view the electron density map is to contour and convert it into a graphical object (mesh).

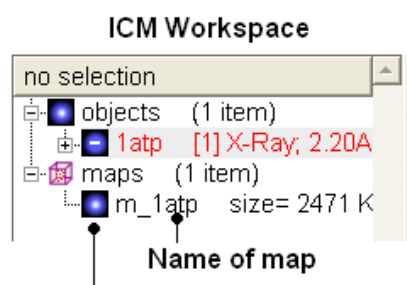
A figure showing the electron density contours surrounding the ATP molecule in pdb entry 1ATP.



To load an electron density map:

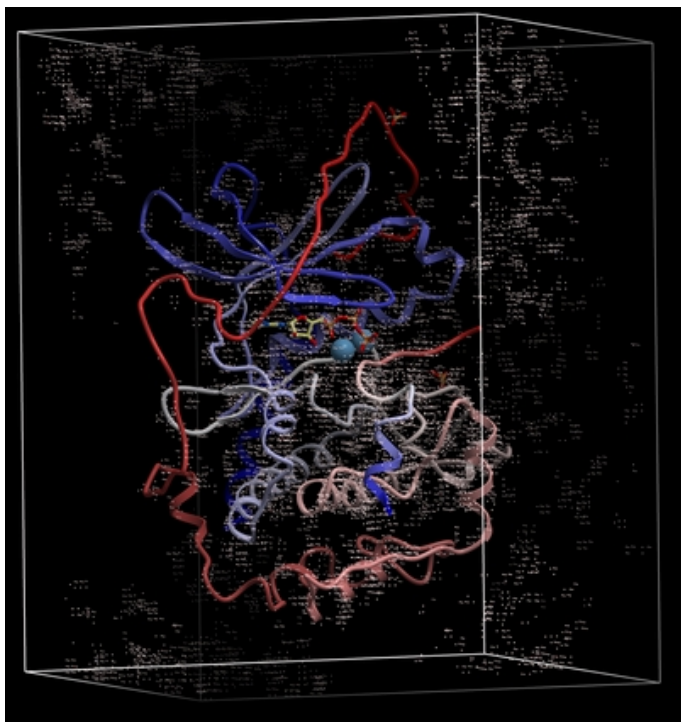
- Tools/Xray/Get Electron Density Map
- Enter the PDB code of the map you would like to view.
- Click OK and the map will be downloaded from the Uppsala Electron Density Server.

The map will be represented in the ICM Workspace as shown below.



Display and undisplay map here

The map can be displayed as shown below however a clearer way of representing the density is to contour the map into a graphical object (mesh) as described in the following section.



4.3.5 Map's Original Cell

To display the original **crystallographic cell** of an electron density map:

- Tools/Xray/Map's Original Cell
- Enter the name of the map or use the drop-down button to locate it. If you do not know the name of the map the name can be located in the ICM Workspace.
- Click OK and the cell will be displayed. The map can be displayed and undisplayed in the **meshes** section of the ICM Workspace.

Generated cell

meshes (1 item)

gCellMap v=8 l=1

maps (1 item)

m_2bdw size= 94

Name of map

Map's Original Cell

map m_2bdw

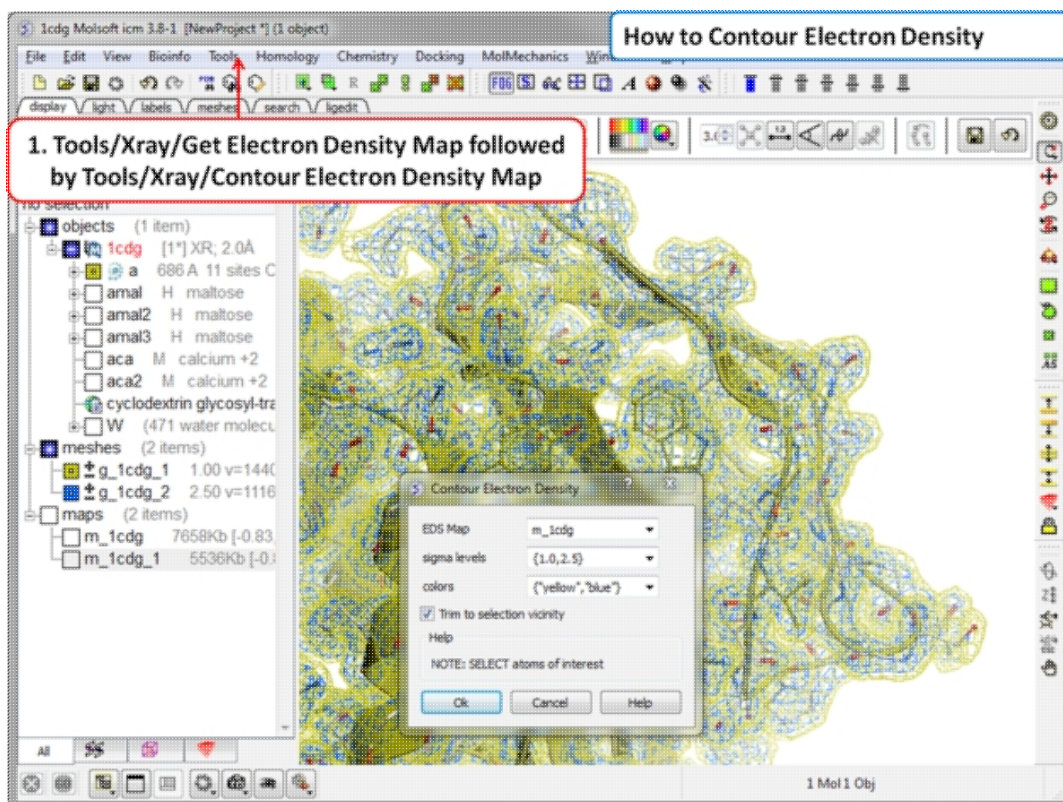
Ok Cancel

4.3.6 Contour Electron Density Map

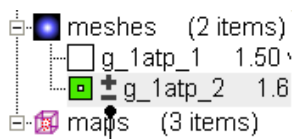
To contour an electron density map and display as a graphical object:

- Load an electron density map as described earlier in the Load Map section.
- Read in the PDB file - File/Load PDB or use the PDB search tab.
- Tools/Xray/Contour Electron Density.
- Enter the name of the map e.g. m_1cdg - the name of the map is displayed in the ICM Workspace or use the drop down arrow to locate it.
- If nothing is displayed then the whole map will be contoured. If you only want to contour a particular region of the map then you need to select the region to be contoured (e.g.the binding pocket) and then select the option **Trim to Selection Vicinity**.
- Enter a sigmaLevel value for more information see: <http://www.molsoft.com/man/reals.html#mapSigmaLevel>. Once the contoured

map has been created the sigma level can be changed manually using the +/- buttons in the ICM workspace.



The sigma level can be changed interactively in the ICM workspace as shown below.



Click here to increase or decrease the sigma level of the contouring

NOTE: Meshes can be cut away using the mesh clipping tools.

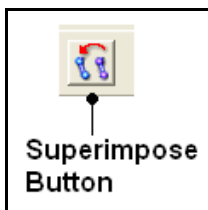
4.3.7 Convert Xray Density to Grid

For some applications, such as trying to fit a structure to a density map, you may want to extract a sub map and convert to a grid. You can do this by

- First read into ICM a map (eg File/Open or Tools/X-ray/Get Electron Density Map)
- Tools/X-ray/Convert Xray Density to Grid
- Enter the map name or use the drop down list
- Enter a grid size
- Click OK

4.4 Protein Superposition

Available in the following product(s): ICM-Browser-Pro | ICM-Pro



One or more proteins can be superimposed. Simply select the molecules or parts of the molecules you wish to superimpose and then use the selection of protein superimpose tools described in this section. A convenient superimpose button can be found in the `Display` tab (see image of button (left)).

Chapter Contents:

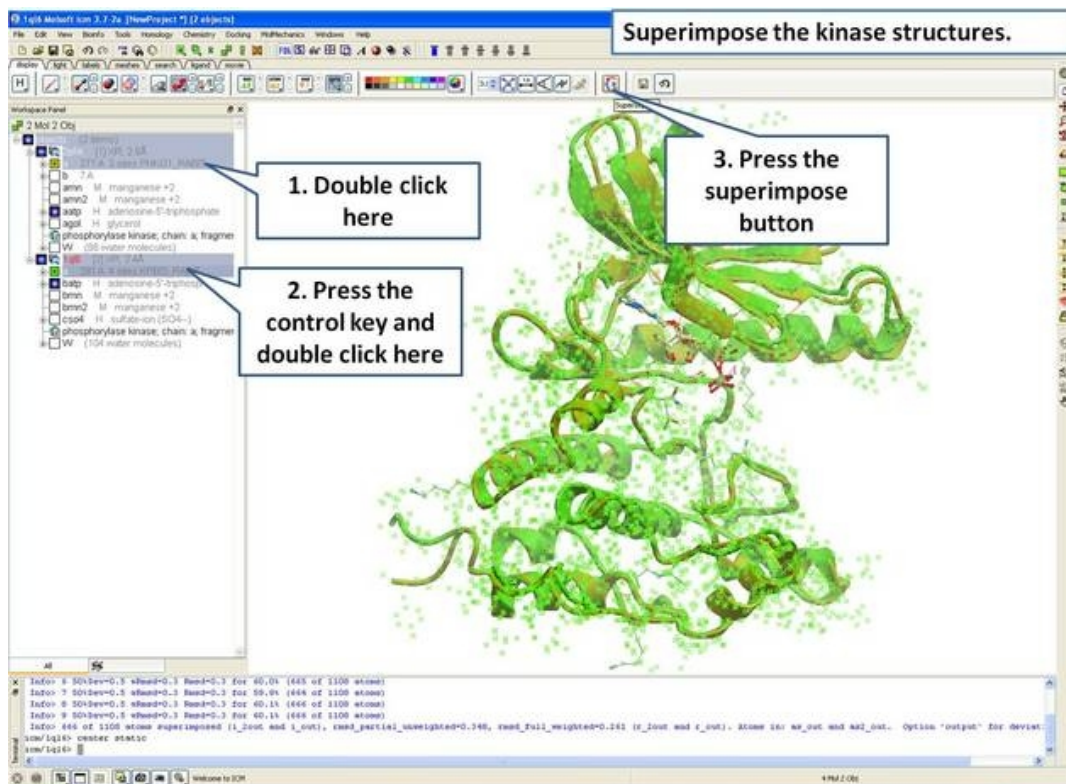
- Select Proteins for Superposition
- Superimpose Button
- Superimpose by 3D
- Superimpose Multiple Proteins
- Arrange as Grid
- Superimpose Sites by Atomic Property Fields

4.4.1 Select Proteins for Superposition

Before any superposition operation can be undertaken you need to select the protein structures you wish to superimpose.

One way to do this is by selecting in the ICM workspace. For other selection tools please see the `Making Selections` section of the manual.

- Select both receptors by double clicking on the name of the molecule in the ICM Workspace. To select two molecules use the `Ctrl` button or use the `shift` button to select a range of objects in the ICM Workspace. A receptor which is selected will be highlighted in blue in the ICM Workspace and with green crosses in the graphical display.



Once the molecules are selected you can then superimpose them using the options described in the next section of this manual.

4.4.2 Superimpose Button

A convenient way to superimpose two molecules is by using the superimpose button in the **display** tab, ICM will calculate the Ca-atom, backbone atom and heavy atom differences between the two structures. More advanced superimpose options can be found in the **Tools/Superimpose** menu.

To superimpose:

- First load the two structures into ICM.
- Select which parts or all of the two structure you wish to superimpose (see the chapter on [Selections](#) or the description [protein-superposition-select{here}](#)).
- Select the **display** tab (previously called Advanced tab) at the top of the GUI.
- Select the superimpose button.



Superimpose
Button

The rmsd will be displayed in the terminal window as shown below:

```
Info> 64 atoms superimposed, rmsd=1.381643
icm/ly6> superimpose ( Res( as_graph ) & a_./ca,c,n,o ) & Obj( as_graph )[1]
Warning> [110] skipped 4 atom pairs with zero occupancies
Info> 64 atoms superimposed, rmsd=1.381643
icm/ly6>
```

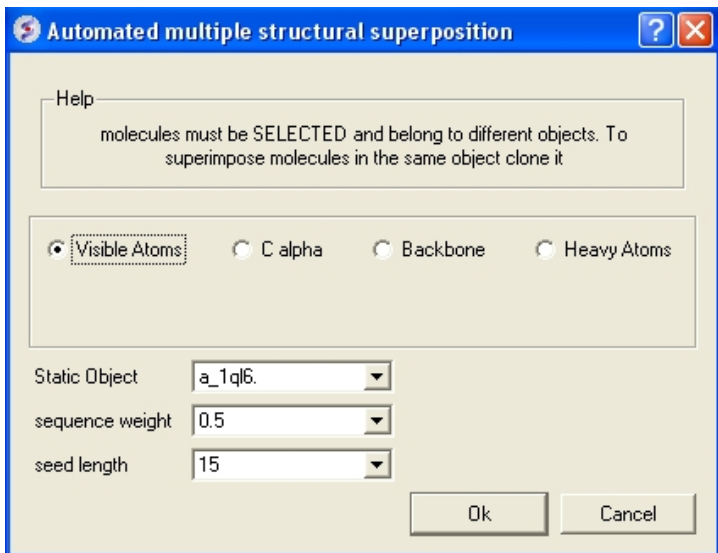
RMSD displayed here

NOTE: You do not need to select the whole molecule, the superimpose button will work on small selections e.g the loop regions or domains.

4.4.3 Superimpose by 3D

To superimpose proteins by 3D:

- First **display** and select the proteins you wish to superimpose by 3D.
- Tools/Superimpose/Proteins by 3D
- A window as shown below will be displayed.

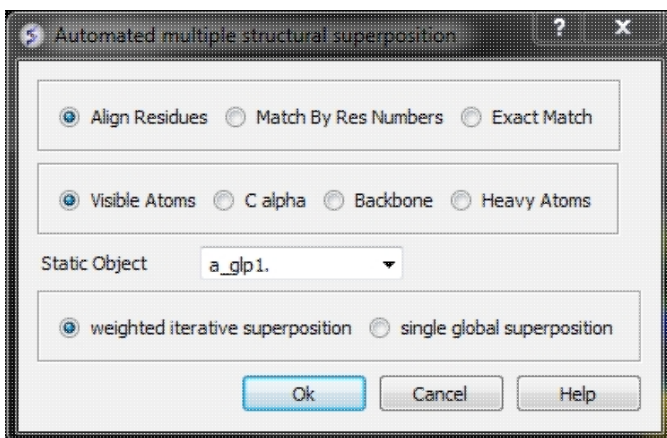


- Select by which atoms you wish to superimpose.
- Enter the ICM selection language description for the protein structure you wish to remain static. You can also use the drop down arrow button to select it.
- Enter the **sequence weight** Average local sequence alignment score.
- Enter the **seed length** This is the similarity window size.

4.4.4 Superimpose Multiple Proteins

To superimpose multiple proteins:

- First display and select the proteins you wish to superimpose by 3D.
- Tools/Superimpose/Multiple Proteins
- A window as shown below will be displayed.



- Select by which method you would like to superimpose

Align Residues - Residue correspondence is established by sequence alignment using the ICM ZEGA alignment Abagyan, Batalov, 1997. Atom alignment: by atom name.

Match by Res Numbers - Residue alignment by residue number. Atom alignment: by atom name for pairs of identical residues or pairs of close residues (F with Y; B with D,N; D with N; E with Q or Z, Q with Z), for other residue pairs only the backbone atoms ca,c,n,o,hn,ha are aligned.

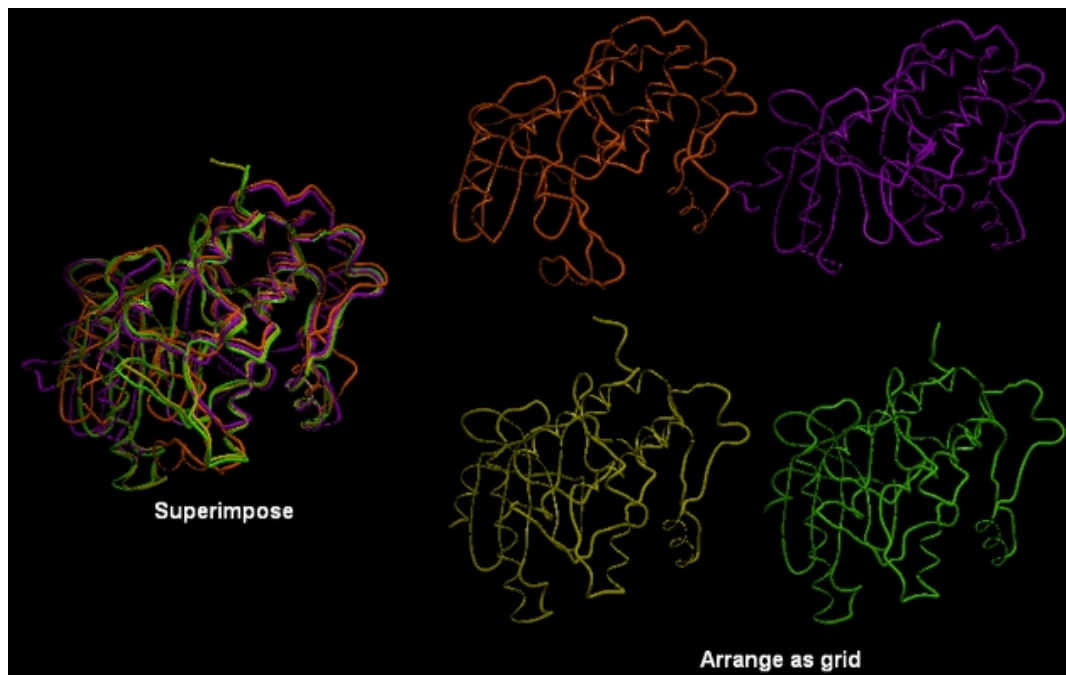
Exact Match - Residue alignment is by the Needleman and Wunsch method. Inside residue atoms are aligned sequentially and regardless of the name.

- Select which atoms you would like to superimpose. Visible Atoms, C alpha, Backbone, or Heavy Atoms.
- Select whether you would like to use **weighted iterative superposition** as described here <http://www.molsoft.com/man/icm-commands.html#superimpose-minimize> or non-iterative **single global superposition**.

4.4.5 Arrange as Grid

To separate superimposed proteins:

- Tools/Superimpose/Arrange as Grid

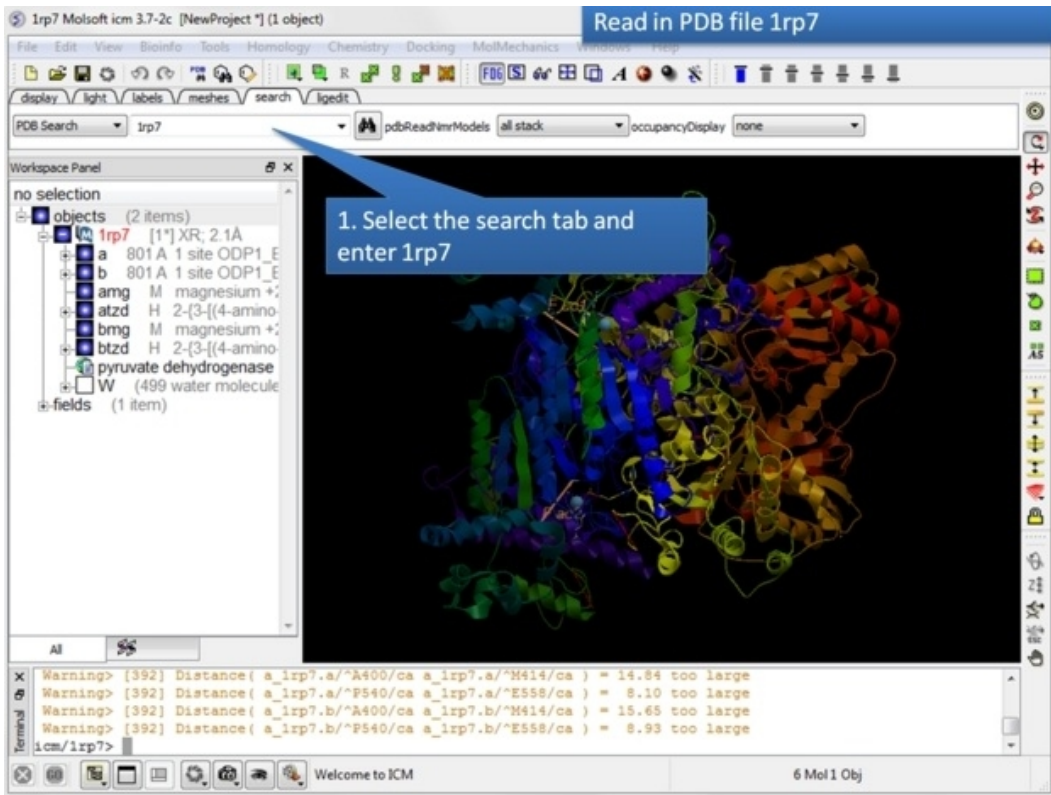


4.4.6 Superimpose Sites by Atomic Property Fields

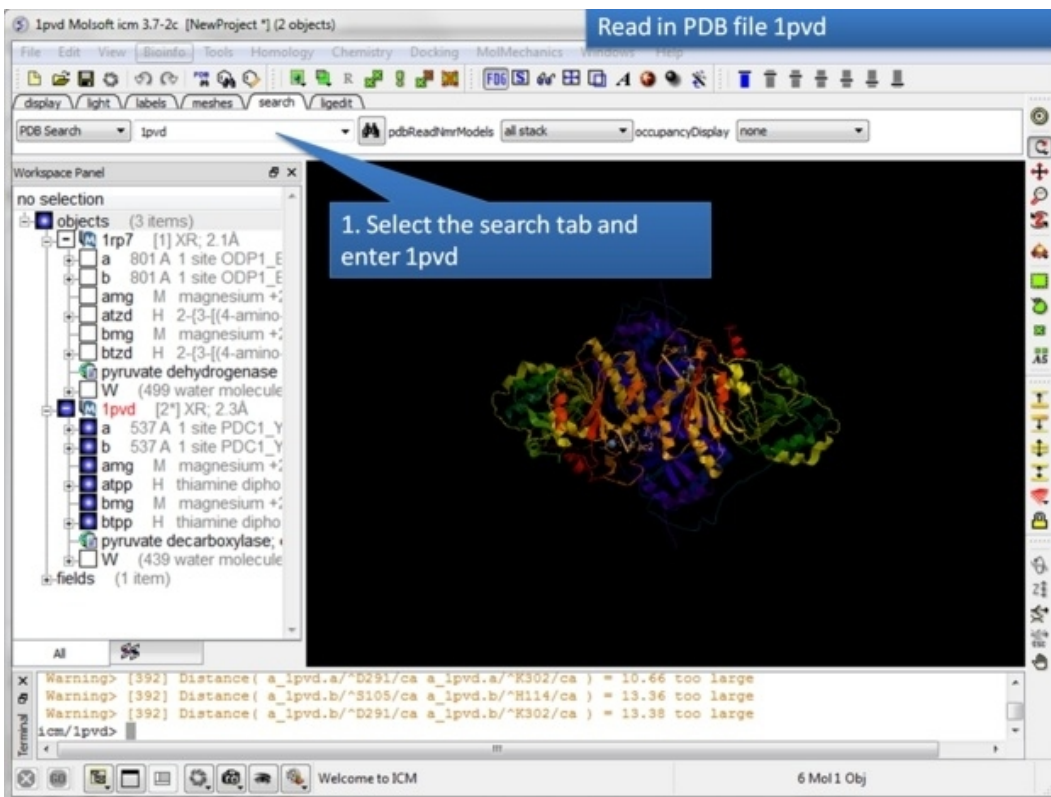
Here we describe how to superimpose and compare a ligand binding site using Atomic Property Fields. This method is described in more detail in this [publication](#).

In this example we superimpose the ligand binding pocket of thiamine diphosphate in the binding sites of pyruvate dehydrogenase (pdb code: 1rp7) and pyruvate decarboxylase (pdb code: 1pvd). Even though the sequence identity between both proteins is very low (19%) and the secondary structure surrounding the ligand undergoes considerable displacement you will see that the pockets can still be superimposed very well using the APF method.

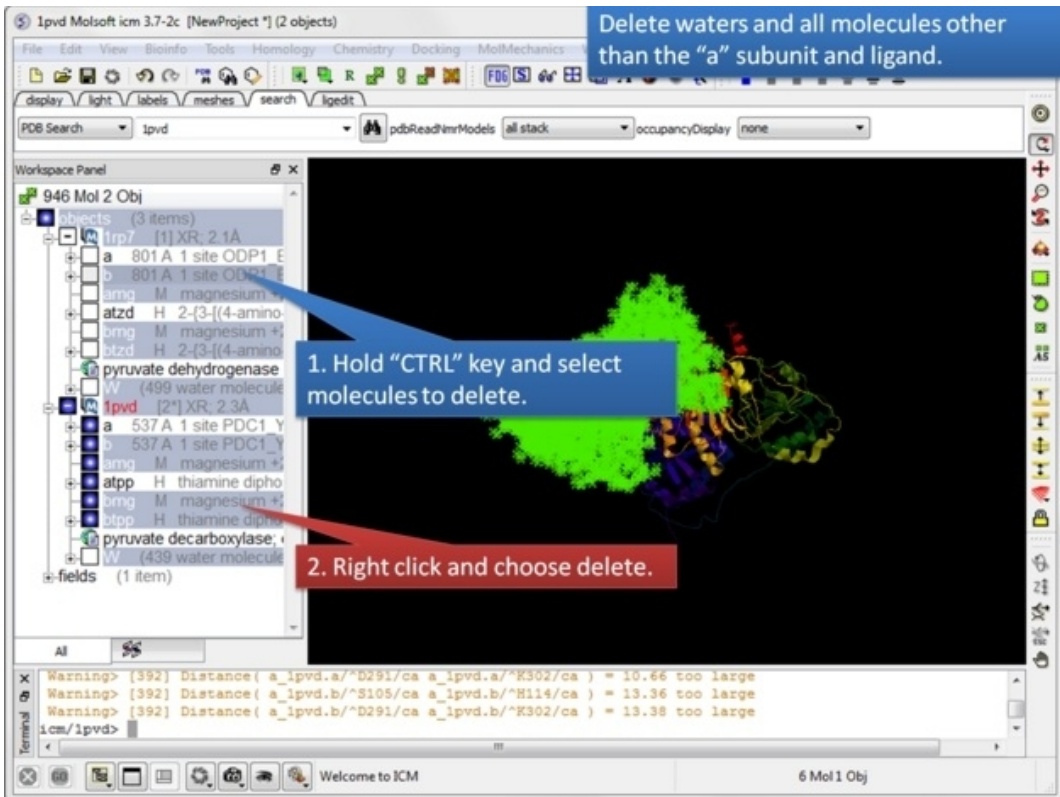
Some Comments about Selecting the Pocket APF pocket overlay works by superimposing a layer of receptor atoms around the pocket using APF chemical property fields of these atoms "pocket" of course is something that is not uniquely defined and is somewhat subjective. One way to define it is to use the ligand inside, in which case a layer of receptor atoms within a certain cutoff distance from the ligand is considered in superposition. This is "around selected ligand" mode. It focuses on atoms that directly interact with the ligand while largely disregarding backbone fold, which can be helpful when comparing different pockets binding similar ligands. The other mode uses input selections directly, leaving it up to the user to define the sets of atoms that comprise the two pockets. Depending on the system and degree of similarity, different selections may work better or worse. For instance in the absence of ligand one may want to detect pocket with `pocket finder` and then select residues around the pocket "blob". Regarding the influence of ligand on superposition, if pocket is defined via the ligand it will of course matter whether the ligands extend into certain regions or not. If external pocket definition is used, ligands will have no direct effect on superposition process.



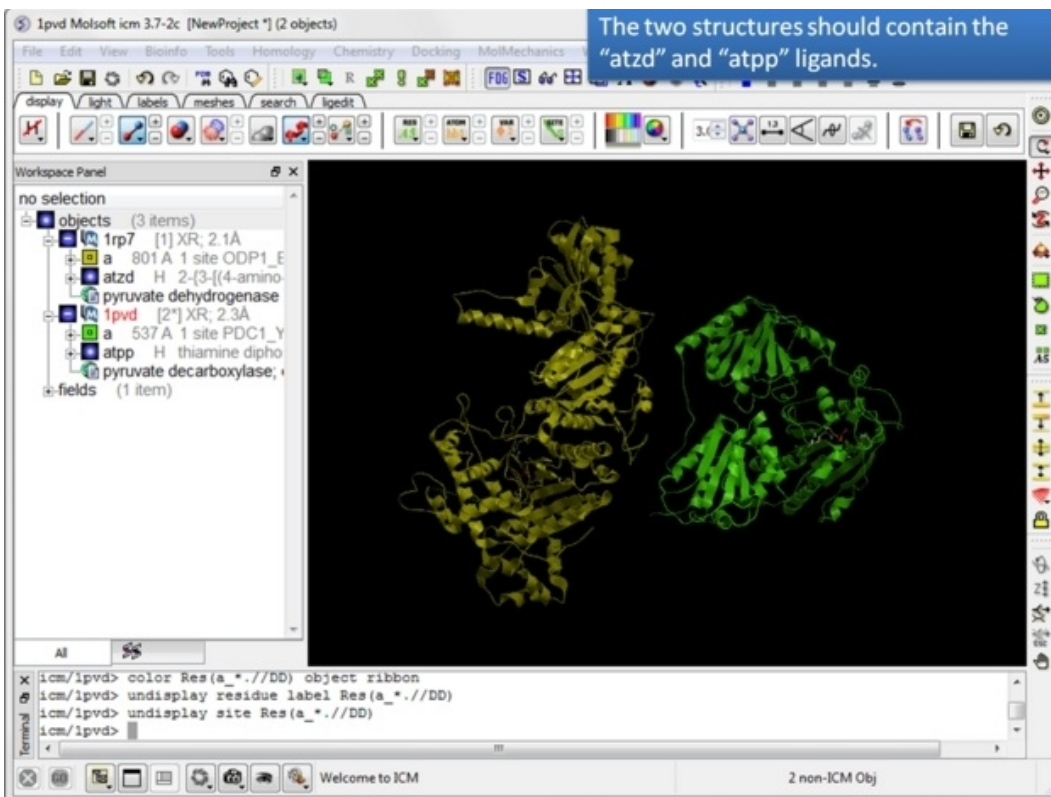
Step 1: Read in the pyruvate dehydrogenase (pdb code: 1rp7) structure.



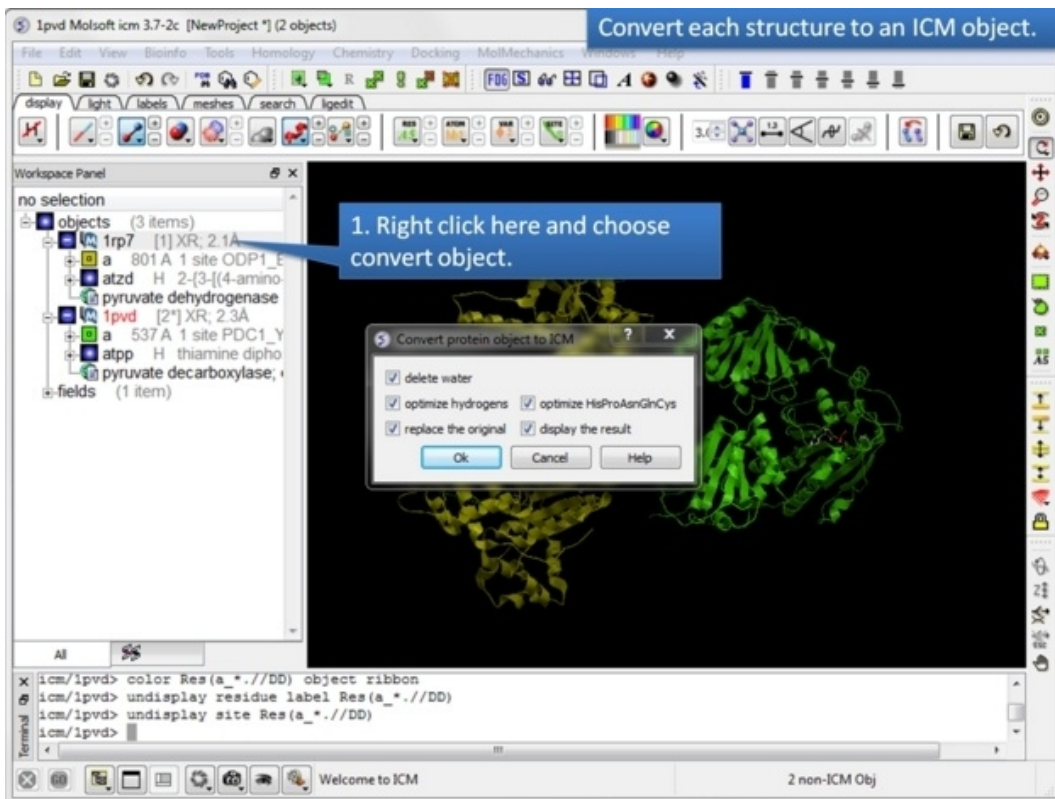
Step 2: Read in the pyruvate decarboxylase (pdb code:1pvd).



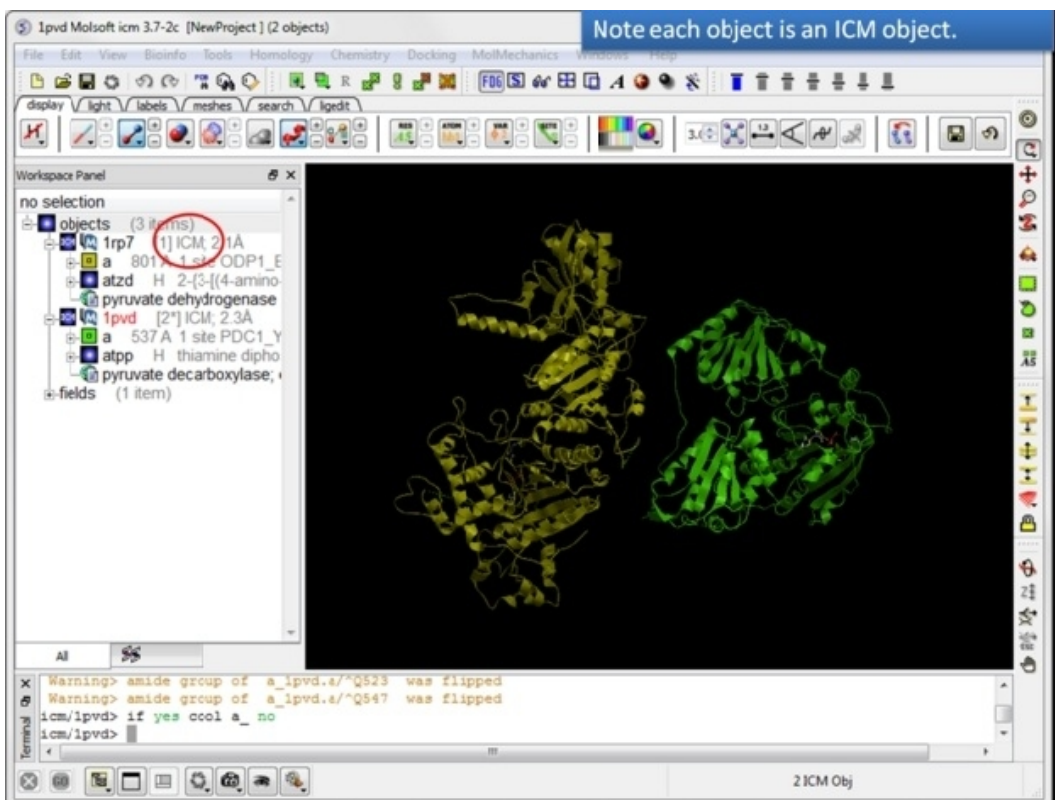
Step 3: Delete unwanted molecules in the b chain and waters.



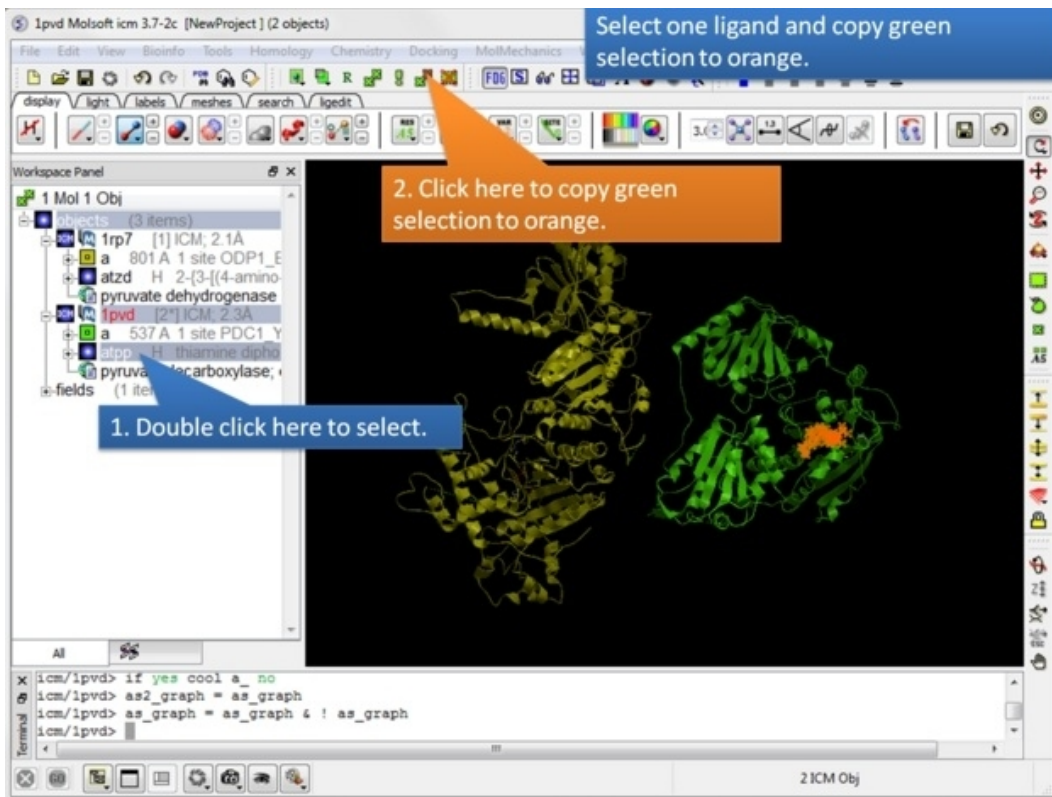
Step 4: You should have two objects containing the protein and ligand.



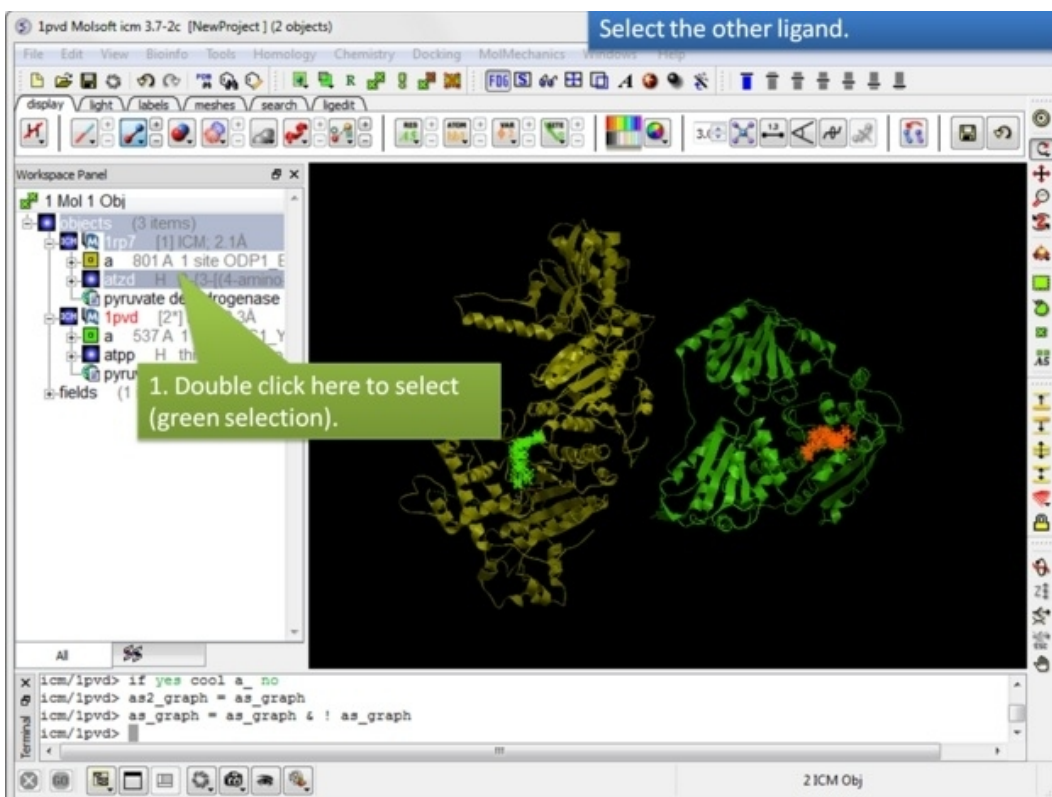
Step 5: Convert both structures to an ICM object.



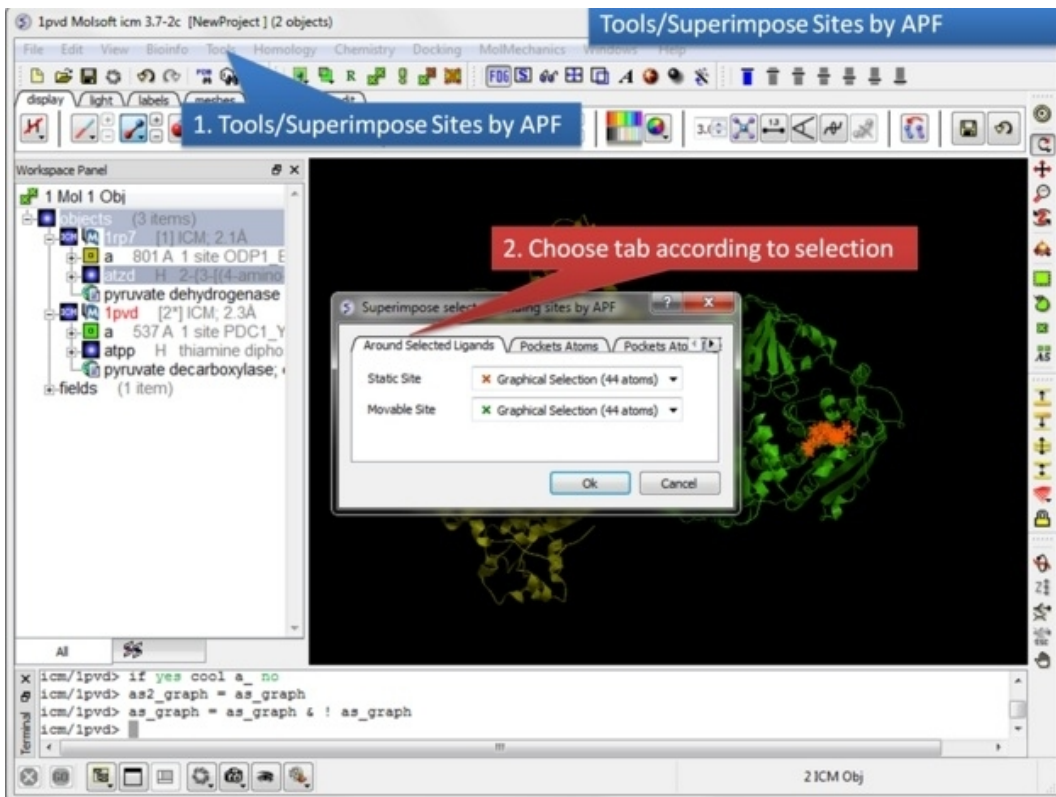
Step 6: If the pdb is converted you will see "ICM" in the workspace.



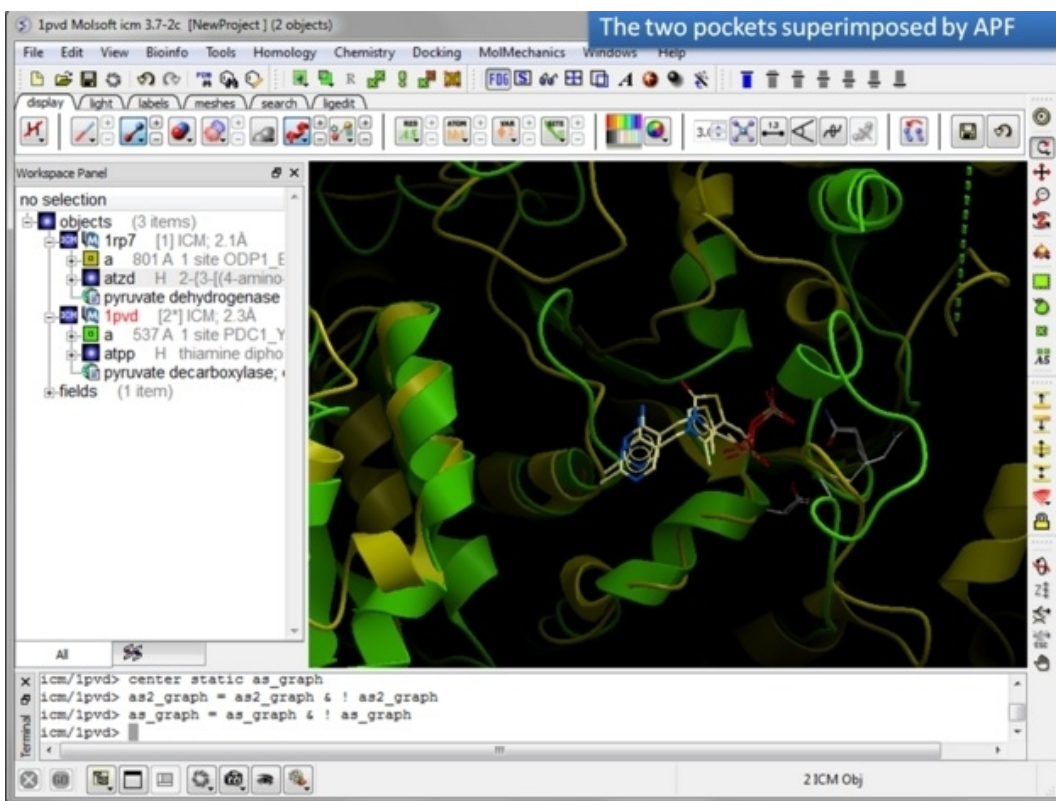
Step 7: Select one ligand and copy the green selection to orange.



Step 8: Select the other ligand (regular green selection).



Step 9: Select the option Tools/Suerpimpose/Sites by APF. Note you can also superimpose pockets - choose the appropriate tabs in the dialog box.



Step 10: Observe the superimposed site.

4.5 Protein Structure Analysis

Available in the following product(s): ICM-Browser-Pro | ICM-Pro

In this chapter we describe the tools available for analyzing protein structure. These tools include calculating RMSD, identifying closed cavities, calculating contact and surface area, measuring angles and distances, and generating Ramachandran plots.

Chapter Contents:

- Find Related Chains
- Calculate RMSD
- Contact Areas
- Identify Closed Cavities
- Surface Area
- Measure Distances
- Planar Angle
- Dihedral Angle
- Ramachandran Plot Interactive
- Export Ramachandran Plot

4.5.1 Find Related Chains

This option allows you to search the currently loaded PDB files or ICM objects and identify chains which are similar and/or related.

You can do this by:

- Select the objects or pdb files you want to compare.
- Tools/Analysis/Find Related Chains
- Click OK to confirm the selection you made
- A table as shown below will be displayed.

	name1	name2	len1	len2	seqid	rmsd	consensus
1	2jc6.a	2jc6.c	278	277	100	0.12	SWKKQAEDIKKIFEFKETLGTGAFSEVLAEEKATGKLFAY IENEIAVLRKIKHENIVALEDIYESPNHLYVMQLVSGGELF
2	2jc6.a	3bhh.a	278	289	39	0.42	# # # E # G # V L # T G # A # K # I # + # # E E # # # R # K H N I V L D # # # Y L V # L V G G E L F - I V + # Y E D A S # I Q H L - A V # # # # M G # V H R D L K P E N L L # # S # # P G Y # P E V L # Y K # V D # W # G V I # Y I L L # G Y P P F # D E
3	2jc6.a	3bhh.b	278	289	38	0.41	# # # E # G # # # # V # T G # A # K # I # + # # E E #
4	2jc6.a	3bhh.c	278	285	37	0.44	# # # E # # # # E A # K # # + # # E E # # # R # K H
5	2jc6.a	3bhh.d	278	286	37	0.43	# # # E # # # # E A # K # # + # # E E # # # R # K H
6	2jc6.c	3bhh.a	277	289	39	0.44	# # # E # G # V L # T G # A # K # I # K # E E # # # R
7	2jc6.c	3bhh.b	277	289	39	0.43	# # # E # G # # # # V # T G # A # K # I # K # E E #
8	2jc6.c	3bhh.c	277	285	38	0.44	# # # # # # V # # T G # A # K # I # K # E E # # # R

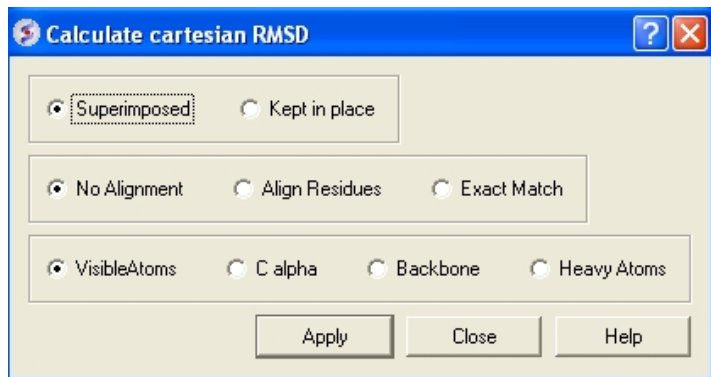
name1 = Name of query structure molecule **name2** = Name of hit **len1** = length of query **len2** = length of hit **seqid** = Sequence identity percentage **consensus** = Consensus sequence

4.5.2 Calculate RMSD

NOTE: This option is for protein structures only not for chemical compounds. You can use the command line options RMSD (<http://molsoft.com/man/icm-functions.html#Rmsd>) and Srmsd (<http://molsoft.com/man/icm-functions.html#Srmsd>) for chemicals.

To calculate RMSD between two protein structure:

- Read into ICM the two structures (File/Open or PDB Search or Read in Chemical) you wish to compare.
- Select one of the two molecules you wish to compare, you can do this by double clicking on the name of the structure in the ICM Workspace. Convert this selection to an Orange Selection.
- Select the second molecule, and then you should have one orange and one green selection in the graphics display.



- Select whether you wish the atoms to be superimposed onto one another or kept in place. The kept in place option would be ideal for comparing docked structures.
- Choose whether you wish to make the superposition by alignment or exactly matching the atom names.
- Select which atom types you wish to superimpose.

The RMSD value will be displayed in the terminal window.

4.5.3 Contact Areas

- Read in a protein structure (File/Open or PDB Search).
- Select the region you wish to analyse.
- Tools/Analysis/Contact Areas
- The xstick display in the region will be scaled according to the atom/residue contact area. For example, residues making large contacts with a ligand will be displayed in thicker xstick representation (and colored yellow) than those making less significant contacts.
- A table as shown below will be displayed. The table lists the contact area, exposed area, percentage of contact area compared to exposed, the nearest atom of a residue and the distance.

How to calculate contact areas.

2. Tools/Analysis/Contact Areas

1. Make a selection

The screenshot shows the ICM software interface with a protein structure and a table of contact areas. The table is as follows:

A	ContactArea	ExposedArea	Percent	Closest Atom	Closest Dist
1 1xbb.astv/1	205.1	579.8	35	a_1xbb.a/ab/n	3.02
2 1xbb.a/^L377	28.1	110.2	25	a_1xbb.astv/stv/c17	3.796
3 1xbb.a/^V385	24.2	71.4	34	a_1xbb.astv/stv/o29	3.484
4 1xbb.a/^H448	22.6	89.5	25	a_1xbb.astv/stv/c4	3.168
5 1xbb.a/^L501	20.8	90.6	23	a_1xbb.astv/stv/c7	3.26

4.5.4 Identify Closed Cavities

This tool will identify cavities within a molecule which are completely closed. If you are looking for buried and open pockets then use `icmPocketFinder`.

- Read in a protein structure (File/Open or PDB Search).
- Convert the protein structure to an ICM object.
- Tools/Analysis/Closed Cavities
- Use the drop down arrow to locate the molecule you are interested in.
- Enter the minimum volume of the cavities you wish to identify.
- Click OK
- The closed cavities will be displayed in the meshes section of the ICM Workspace and a table of the cavities will be displayed. Double click on a row in the table to jump to a particular closed cavity and select the residues surrounding it.

How to predict Closed Cavities.

2. Tools/Analysis/Closed Cavities

1. Right click and choose convert to ICM object

3. Toggle cavity display on/off here

4. Double click here to select the residues surrounding the cavity

i	Volume	Area	Radius	Type
1	2	72.83	91.92	2.377 cavity2
2	3	30.11	48.89	1.848 cavity3
3	4	31.51	49.67	1.902 cavity4
4	5	23.17	40.27	1.726 cavity5
5	6	21.22	38.32	1.661 cavity6
6	7	17.3	32.99	1.573 cavity7

Header Filter Volumes an...

Volumes and areas

Area

Volume

2 Mol 1 Obj

What is the difference between Closed Cavity and ICMPocketFinder? Closed pocket (cavityFinder) is purely geometrical/topological - it is a part of molecular surface that is completely disconnected from the exterior surface. It means that a probe sphere of 1.4Å radius (representing a water molecule) can not pass in or out of the cavity (considering protein as completely rigid of course). icmPocketFinder identifies pockets that are likely to contain ligands (not specifically open or closed pockets). Pockets are defined based on physical interaction rather than geometric criterion. Blobs of 'pocket density' generated by icmPocketFinder represent continuous regions of space where there is significant favorable van der Waals interaction with the receptor.

4.5.5 Surface Area

This option calculates solvent accessible area of each selection in multiple objects and stores it in a table. If a molecule is specified in a multi-molecular object, the surface area of an isolated molecule is calculated and other molecules are ignored. The area is reported in square Anstroms and the probe radius is assumed to be the value set in the variable waterRadius.

Output: the macro creates table AREA . The empty comment field is added for user's future use. If the table exists, new rows are appended.

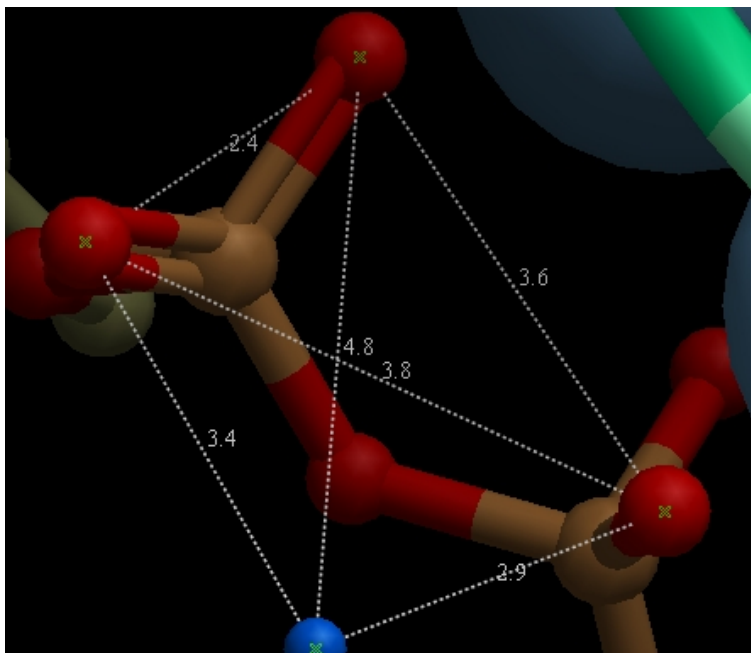
To calculate a surface area:

- Read in a protein structure (File/Open or PDB Search).
- Select the region you wish to analyse.
- Tools/Analysis/Surface Area
- A table will be displayed listing the residues in the selection along with the corresponding total surface area.

4.5.6 Measure Distances

There are two approaches to calculating and displaying distances between atoms. You can either use the options in the Labels tab or use Tools/Analysis/Distance.

To display **all to all** distances:



- Select the atoms between which you would like to find the distance. (See selection toolbar)
- Tools/Analysis/Distance
- Select all to all

To display **intermolecular** distances

- Select the atoms between which you would like to find the distance. (See selection toolbar)
- Tools/Analysis/Distance
- Select **intermolecular**

To display the distances between the same atoms in two objects.

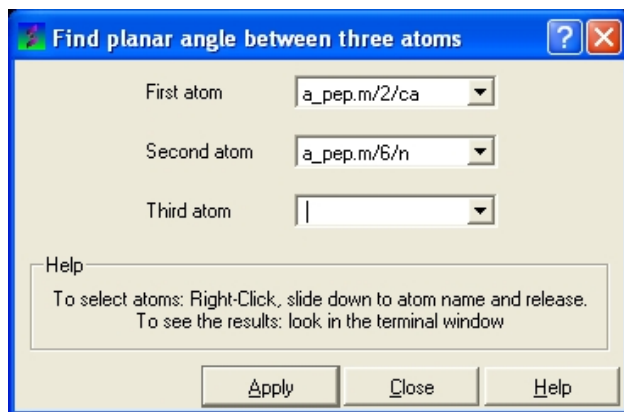
- Select the atoms between which you would like to find the distance. (See selection toolbar)
- Tools/Analysis/Distance
- Select **same atoms in two objects**

NOTE: Distances can be displayed and undisplayed in the 3D labels section of the ICM Workspace. You can change the color of a distance label by right clicking on it in the ICM Workspace. You can also export the distance to a table.

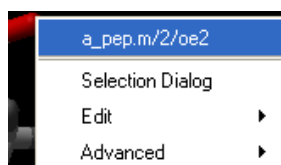
4.5.7 Planar Angle

If you wish to find the planar angle between three atoms:

- Select **Tools/Analysis/PlanarAngle**



- Right click on the each of the three atoms which you wish to use, and select their name. The spaces next to **First atom**, **Second atom**, and **Third atom** should now contain the name of your atoms.



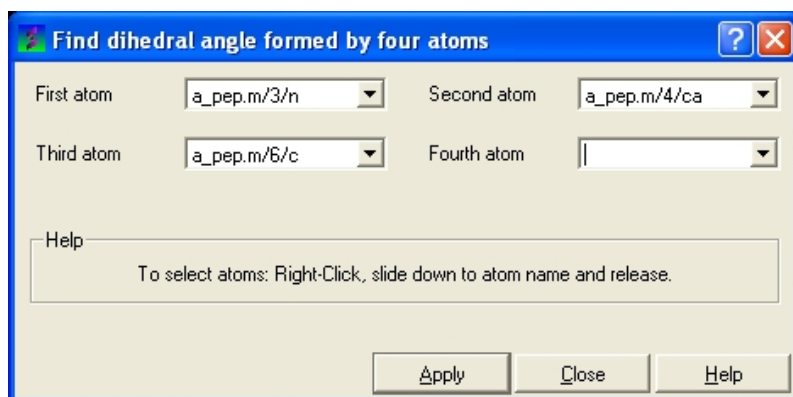
- Click **Apply** to display the angle measure in the terminal window.

Angle (a_pep.m/6/hh21 a_pep.m/2/oe2 a_pep.m/3/o) = 74.72 deg.

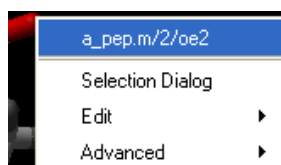
4.5.8 Dihedral Angle

In order to find the angle dihedral angle between two sets of atoms:

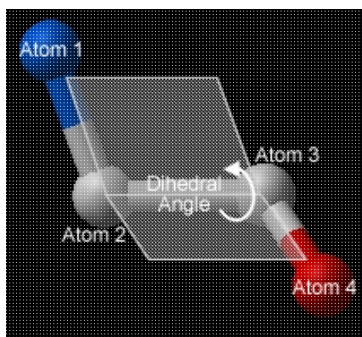
- Select **Tools/Analysis/Dihedral Angles**.



- Right click on each of the four atoms which you wish to use, and select the name of the atoms. The spaces next to **Atom 1**, **Atom 2**, **Atom 3**, and **Atom 4** should now contain the names of your atoms.



- To find the correct angle, select your atoms according to the following diagram:



- Click **Apply** to display your dihedral angle measure in the terminal window.

4.5.9 Ramachandran Plot Interactive

To make an interactive ramachandran plot:

- Read in a protein structure (File/Open or PDB Search).
- Select the structure you wish to build the plot for. You can do this by double clicking on the name of the structure in the ICM Workspace (a selection is highlighted blue in the ICM Workspace and green crosses in the graphical display) or you can use the right-click button and drag it over the whole structure in the graphical display.
- Tools/Analysis/Ramachandran Plot Interactive
- The interactive ramachandran plot will be displayed in table called **RAMA**.
- You can view the **Omega, Phi/Psi** (Gly) or Phi/Psi angles by clicking on the tabs at the top of the plot. Each point is linked to the data in the table **RAMA** and also to the graphical display. Clicking on a point in the plot will highlight the corresponding angles in the table and also center on this region in the 3D display.

4.5.10 Export Ramachandran Plot

- Read in a protein structure (File/Open or PDB Search).
- Select the structure you wish to build the plot for. You can do this by double clicking on the name of the structure in the ICM Workspace (a selection is highlighted blue in the ICM Workspace and green crosses in the graphical display) or you can use the right-click button and drag it over the whole structure in the graphical display.
- **Tools/Analysis/Ramachandran Plot Export**

A postscript viewer needs to be downloaded onto your machine in order to view the plot. This can be downloaded from <http://www.cs.wisc.edu/~ghost/>. Once this software is downloaded you need to tell ICM where it is located by typing the pathname into File/Preferences.

NOTE: You can always export the plot as an image directly in ICM without exporting. You can do this by right clicking on the plot and select **save as image**. Another approach could be to export the RAMA table to Excel and use the plotting tools there. You can do this by right clicking on the table name tab and selecting "Export to Excel" or save as ".csv".

4.6 3D Predict

Available in the following product(s): ICM-Pro

Chapter Contents:

- Assign Helices and Strands
- Protein Health
- Local Flexibility
- Protein-Protein Interface Prediction
- Identify Ligand Pockets

4.6.1 Assign Helices and Strands

Theory

The **Assign helices and Strands** option will manually reassign secondary structure to a protein structure. This command does not change the geometry of the model, it only formally assigns secondary structure symbols to residues. The method uses a modification of the DSSP algorithm of automatic secondary structure assignment (Kabsch and Sander, 1983) based on the observed pattern of hydrogen bonds in a three dimensional structure. The DSSP algorithm in its original form overassigns the helical regions. For example, in the structure of T4 lysozyme (PDB code 103l) DSSP assigns to one helix the whole region a_/93:112 which actually consists of two helices a_/93:105 and a_/108:112 forming a sharp angle of 64 degrees. ICM employs a modified algorithm which patches the above problem of the original DSSP algorithm.

To assign secondary structure:

- Read in a protein structure (`File/Open` or `PDB Search`).
- Select the structure. You can do this by double clicking on the name of the structure in the ICM Workspace (a selection is highlighted blue in the ICM Workspace and green crosses in the graphical display) or you can use the right-click button and drag it over the whole structure in the graphical display.
- Tools/3D Predict/Assign helices and Strands.

The secondary structure sequence string (e.g. EEEEEEEEE__HHHH) will be displayed in the terminal window. Assigned secondary structure types are the following: "H" - alpha helix, "G" - 3/10 helix, "I" - pi helix, "E" - beta strand, "B" - beta-bridge, "_" or "C" - coil. If the sequence is viewed in the ICM workspace or in an alignment the secondary structure is colored as follows alpha helix (red bar), beta sheet (green bar), non-canonical helix (purple bar).

4.6.2 Protein Health

Theory

The protein health option calculates the energy strain of a structure in ICM. It is generally a good idea to investigate the energy strain of any protein structure before undertaking such processes as docking. It is also essential to use this tool after making a model (see `Molecular Modeling`) to identify strained regions within your model and then some optimization procedure can be undertaken to rectify the problems.

The protein health option calculates the relative energy of each residue for a selection and colors the selected residues by strain. This macro uses statistics obtained in the following paper Maiorov, V.N. and Abagyan, R.A. (1998) Energy strain in three-dimensional protein structures *Folding and Design*, 3 , 259-269.

To use the Protein Health option:

- Read in a protein structure (`File/Open` or `PDB Search`).
- Convert your PDB structure into an ICM object.
- Make a selection of the residues you wish to analyze.
- Tools/3D Predict/Protein Health and a window as shown below will be displayed.
- The scale of the coloring can be changed by altering the value within the `trimEnergy` data entry box.
- Click OK and the structure will be colored according to energy strain (red - high) and a table of residue energy will be displayed in a table.
- A table and plot of Normalized energies for each amino acid in the selection will be displayed. The table is ranked and colored by residues with poor normalized energies. Click on a row in the table or plot to center in on the residue.

4.6.3 Local Flexibility

This option systematically samples rotamers for each residue side-chain in the input selection and uses resulting conformational ensembles to evaluate energy-weighted RMSDs for every side-chain atom. These are stored in the 'field' values on atoms and can be used for example to color the structure by side-chain flexibility. Conformational entropy for each residue side-chain is also calculated and stored in a table. If

If the `_entropyBfactor` flag is on, the atom rmsds are normalized within the residue to reflect its total conformational entropy. If the `_bfactor` flag is set, the b-factors are reset to the same values that are placed in the atom 'field', and occupancy is set to be inversely proportional to it ($O=1/(1+2*rmsd)$)

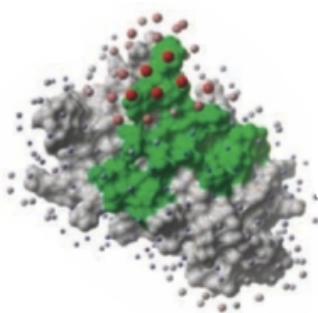
- Read pdb file (File/Open or PDB Search Tab).
- Convert to an ICM Object.
- Tools/3D Predict/Local Flexibility

4.6.4 Protein-Protein Interface Prediction

The ICM Optimal Docking Area method is a useful way of predicting likely protein-protein interaction interfaces. If you do not have mutational data or other experimental data which indicates the likely protein-protein docking site this method will be useful. This procedure can save you time during the docking procedure by focusing your docking only on areas on the receptor and ligand most likely to interact.

Theory

ODA (Optimal Docking Areas) is a new method to predict protein-protein interaction sites on protein surfaces. It identifies optimal surface patches with the lowest docking desolvation energy values as calculated by atomic solvation parameters (ASP) derived from octanol/water transfer experiments and adjusted for protein-protein docking. The predictor has been benchmarked on 66 non-homologous unbound structures, and the identified interactions points (top 10 ODA hot-spots) are correctly located in 70% of the cases (80% if we disregard NMR structures). For a description of the method see Fernandez-Recio et al *Proteins* (2005) 127: 9632.



To display the optimal docking area.

- Convert the PDB file to an ICM object.
- Tools/3D Predict/Protein Interface by ODA
- If you select the **Residue Table** option the average ODA score for each residue will be displayed in a table. The lower the number the higher the chance the residue will be involved in protein-protein interactions. Regions colored red represent low ODA score and blue represents a high score.

4.6.5 Identify Ligand Binding Pockets

Theory

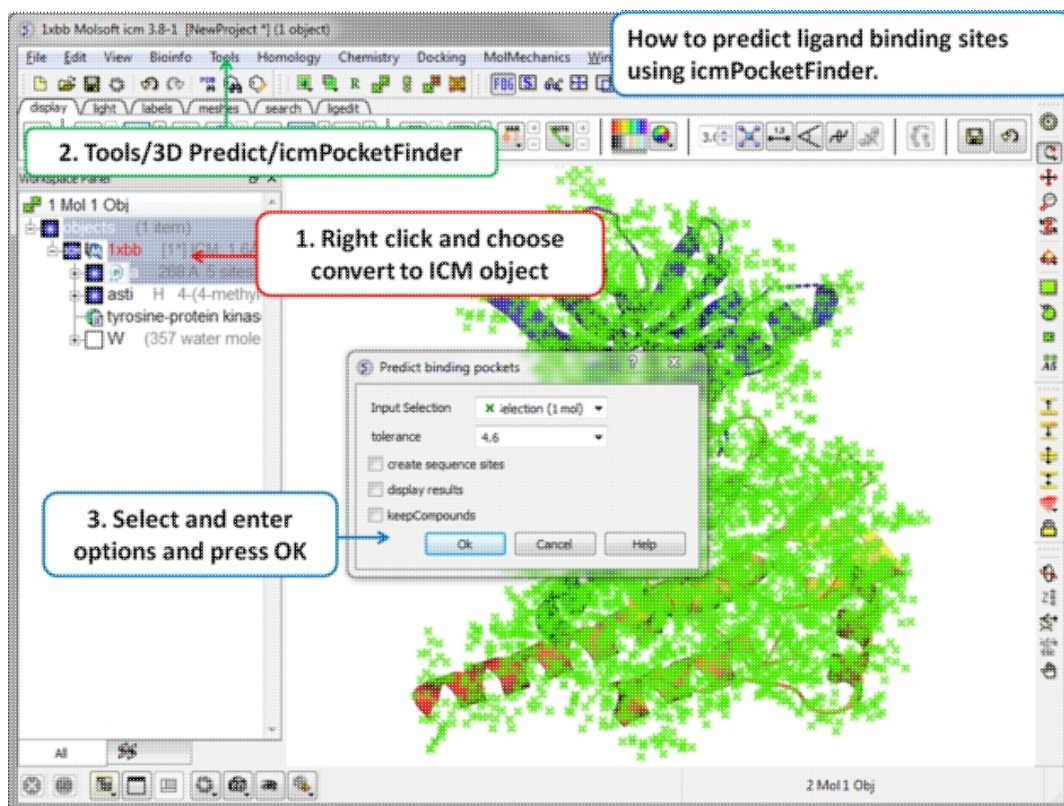
The ICM Pocket Finder method (1-2) uses only the protein structure for the prediction of cavities and clefts. No prior knowledge of the substrate is required. The position and size of the ligand-binding pocket are determined based on a transformation of the Lennard-Jones potential by convolution with a Gaussian kernel of a certain size, a grid map of a binding potential and construction of equipotential surfaces along the maps. The pockets are displayed graphically as a surface and the dimensions of each pocket are presented in an interactive table and plot.

Factors that can influence ligand binding to a pocket include the pocket volume and area, buriedness, hydrophobicity, and how compact the pocket is. All these properties are calculated using ICMPocketFinder and tabulated. Scientists at Merck used MolSoft's ICMPocketFinder algorithm to define a way for quantifying "drugability" of a protein target (3). The metric they used is called Drug-Like-Density (DLID), and this score is also provided in the results table. A good example of the use of the icmPocketFinder method is in the database called Pocketome. The Pocketome (www.pocketome.org) is an encyclopedia of

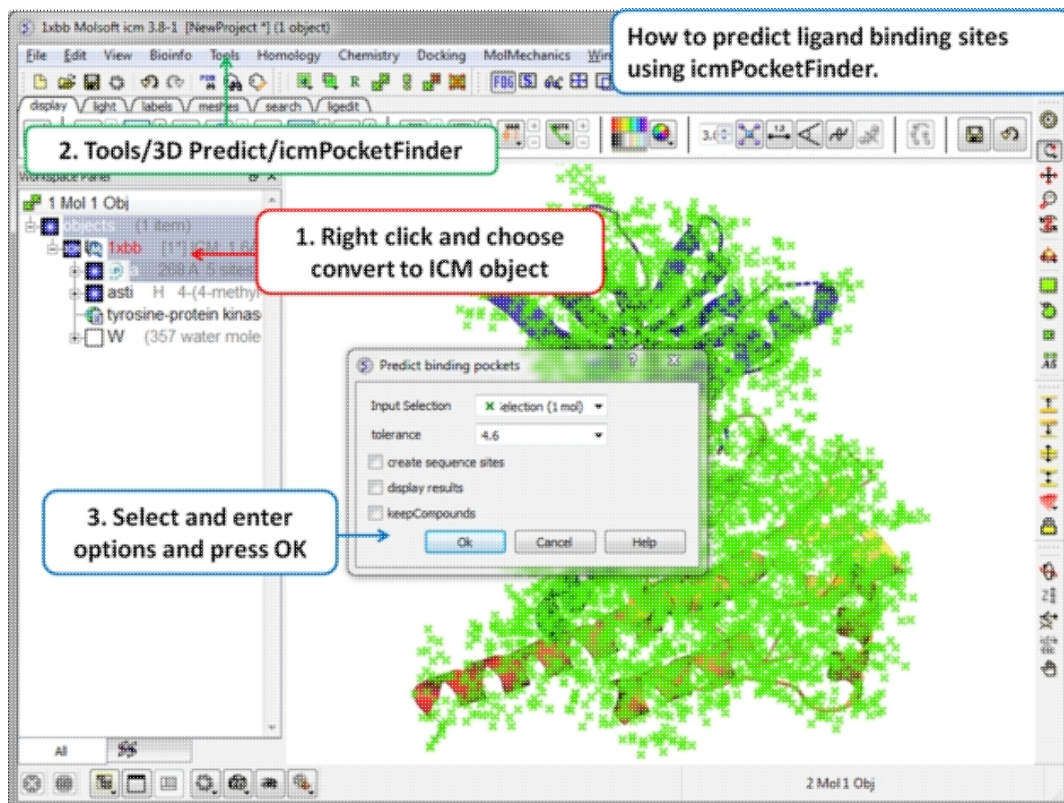
conformational ensembles of all druggable binding sites that can be identified experimentally from co-crystal structures in the Protein Data Bank (4).

1. An, J., Totrov, M. & Abagyan, R. Pocketome via comprehensive identification and classification of ligand binding envelopes. *Mol. Cell. Proteomics* 4, 752 (2005).
2. Abagyan, R. & Kufareva, I. The flexible pocketome engine for structural chemogenomics. *Methods Mol. Biol. Clifton NJ* 575, 249.279 (2009).
3. Sheridan, R. P., Maiorov, V. N., Holloway, M. K., Cornell, W. D. & Gao, Y.-D. Drug-like density: a method of quantifying the .bindability. of a protein target based on a very large set of pockets and drug-like ligands from the Protein Data Bank. *J. Chem. Inf. Model.* 50, 2029.2040 (2010).
4. Kufareva, I., Ilatovskiy, A. V. & Abagyan, R. Pocketome: an encyclopedia of small-molecule binding sites in 4D. *Nucleic Acids Res.* 40, D535.540 (2012).

To predict pockets:



- Read in a protein structure (File/Open or PDB Search).
- Convert the protein structure to an ICM object.
- Tools/3D Predict/icmPocketFinder
- Enter a tolerance level (4.6 is the default value and we recommended you to use this). The lower the tolerance value the more pockets predicted and the higher the tolerance the less pockets predicted.
- Check the box **create sequence sites** if you wish the site to be labeled.
- Check the box **display results** to see the predicted pockets as grobs in the display panel.
- Check the box **keep compounds** if you wish the compounds (ligands) in the receptor to be included in the prediction. If you dont check this box the pockets will be calculated based on the receptor without ligands.
- Click OK to run icmPocketFinder and the results will be displayed in a table.



Results

- Click on a row of the results table to select the residues surrounding the pocket.
- The pocket view can be toggled on or off in the meshes section of the ICM Workspace. Right click on the name of the mesh in the ICM workspace to change properties such as color.
- A fully interactive plot (Area vs Volume) is also provided, pockets that fall within the blue shaded region have a "drug-like" volume and size.

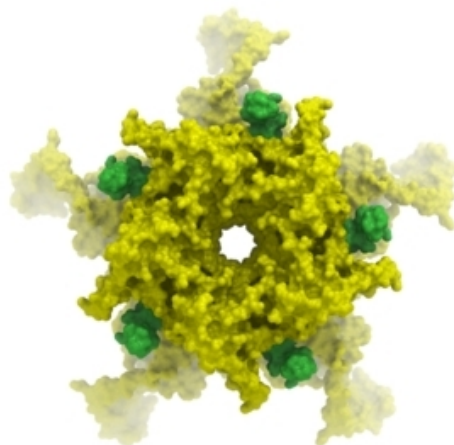
About the icmPocketFinder Results Table

Factors that can influence ligand binding to a pocket include the pocket volume and area, buriedness, hydrophobicity, and how compact the pocket is. These values are reported in the results table:

- **Volume** of the pocket in \AA^3 .
- **Area** of the pocket in \AA^2 .
- **Hydrophobicity** - represents the percentage of the pocket surface s in contact with hydrophobic protein residues (values can range from 0-1)
- **Buriedness** - The buriedness parameter is calculated as follows: One measures the solvent accessible surface area of the pocket (probe radius, 1.4) in isolation. Then one measures the solvent accessible surface area of the pocket covered by its shell. The ratio of the second number to the first is the fraction buried. The lowest possible value is 0.5; i.e. the pocket is completely open and the surface flat. The highest is 1.0, i.e., completely buried.
- **DLID** - Merck's Drug-like density score (see Sheridan et al JCI 2010). >0.5 is considered "druggable".
- **Radius** and **Nonsphericity** give an indication of how spherical the pocket is.
- **Type** - returns the ICM selection for the residues forming the pocket.

5 Molecular Graphics

In this chapter we describe how to make beautiful graphical representations of molecules and manipulate them in the 3D graphics window. This includes how to change color, light, representations, clipping planes, and how to use built in graphics effects. We also teach how to label and annotate molecules displayed in the graphical user interface.



Chapter Contents:

- Molecule Representation
- Meshes-Surface-Grobs
- Coloring
- Lighting
- Labeling and Annotation
- Display Distances and Angles
- Graphics Effects
- Graphics Shortcuts
- Molecule Move Buttons
- Clipping Tools
- Graphic Layers
- High Quality Publication Images
- Movies

5.1 Molecule Representation

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

To change the molecule display representation:

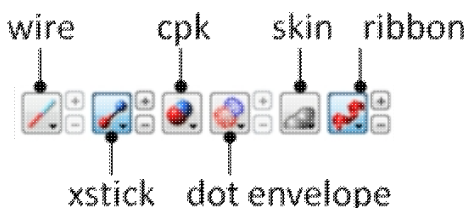
- Select the atoms, residues, molecules, or objects you wish to change in the graphical display or in the ICM Workspace.
- Then use the molecule representation (e.g. wire, ribbon) options in the `Display` Tab.

The display tab contains tools for a variety of functions including - structural representations, coloring, labeling and superposition.

There are six main types of structural representation in ICM. They are wire, ball and stick (Xstick), ribbon, skin, CPK and dot envelope (surface).

To display one of these representations:

- Click on the representation button you desire in the **display** tab.



To remove a displayed representation or to toggle between display and undisplay:

- Click on the corresponding representation button in the **display** tab.

Identify which representations are displayed: The button display will change appearance (shaded blue) when selected. This makes it easier to identify which representations are currently being displayed.

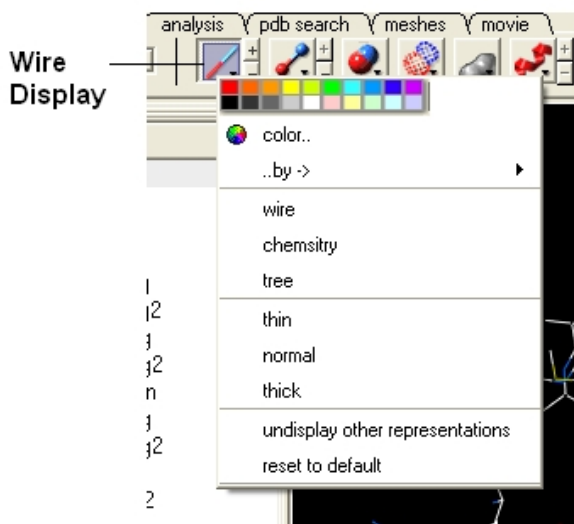


Change the display and size of the representations Many characteristics of the graphical representation such as color can be changed by clicking and **holding** on the button. The size can be changed by clicking the plus(+) and minus(-) buttons next to them.



5.1.1 Wire Representation

Click and hold on the **wire representation** button. A menu will be displayed as shown below.



To change the wire style:

- Click and hold on the **wire representation** button and then click on **wire**, **chemistry** or **tree**.

To change the size of the wire representation:

- Click and hold on the **wire representation** button and then click on **thin**, **normal** or **thick**.

NOTE: Clicking on the +/- next to the **wire representation** button also changes the thickness of the wire representation.

To undisplay representations other than wire:

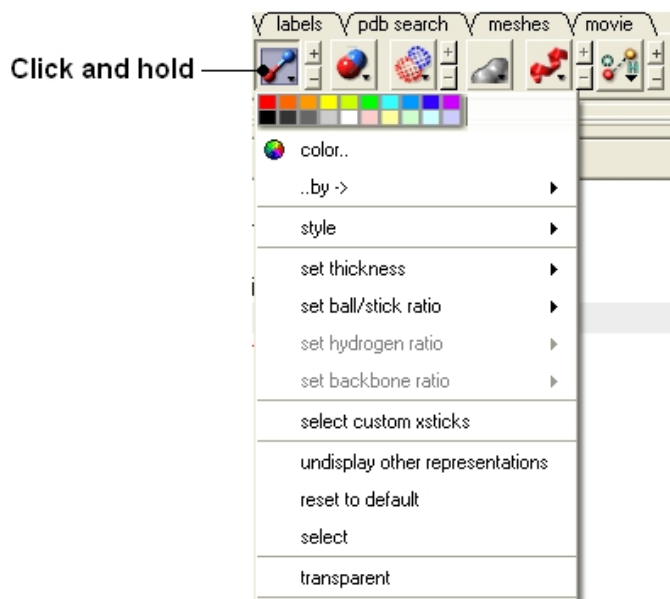
- Click and hold on the **wire representation** button and then click on undisplay other representations.

If you make a mistake or you are not happy with the way your structure is displayed with the wire representation:

- Click and hold on the **wire representation** button and then click on reset to default.

5.1.2 Stick and Ball (Xstick) Representation

Click and hold on the **stick and ball representation** button. A menu will be displayed as shown below.



To change the style of the Xstick representation:

- Click and hold on the **stick and ball representation** button and then click on style and choose either **plain**, **chemistry**, **Chem Plus**, or **Chem Plus Aromatic**.

To change the size of the Xstick representation:

- Click and hold on the **stick and ball representation** button and then click on **set thickness**, **set ball/stick ratio**, **set hydrogen ratio**, and **set backbone ratio**.

NOTE: Clicking on the +/- next to the **xstick representation** button also changes the thickness of the xstick representation.

In order to make some parts of your picture clearer, the xstick representation can be set to transparent:

- Click and hold on the **stick and ball representation** button and then click on transparent.

To undisplay representations other than xstick:

- Click and hold on the **stick and ball representation button** and then click on undisplay other representations.

If you make a mistake or you are not happy with the way your structure is displayed with the xstick representation:

- Click and hold on the **stick and ball representation** button and then click on reset to default.

5.1.3 Ribbon Representation

Click and hold on the ribbon representation button and a menu will be displayed.

To change the style of the Ribbon representation:

- Click and hold on the **ribbon representation button** and then click on a style option: **smooth, wide, wide smooth, cylinders, protein worm** .

To assign secondary structure:

To accurately represent the secondary structure of the molecule in ribbon representation you may wish to assign secondary structure:

- Click and hold on the **ribbon representation** button and then click on **assign sec. structure**.

To make some parts of your picture clearer, the ribbon representation can be set to transparent:

- Click and hold on the **ribbon representation button** and then click on **transparent**.

To undisplay representations other than ribbon:

- Click and hold on the **ribbon representation** button and then click on undisplay other representations.

If you make a mistake or you are not happy with the way your structure is displayed with the ribbon representation:

- Click and hold on the **ribbon representation button** and then click on **reset to default**.

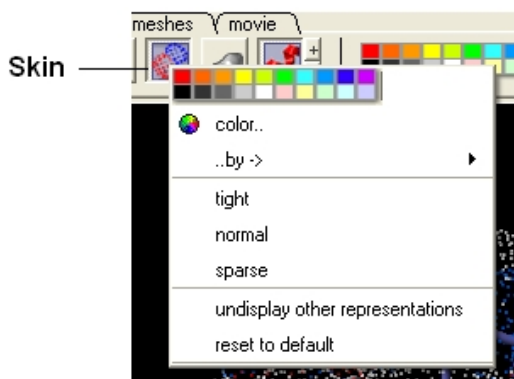
NOTE: Always use the **ICM assign sec.**** structure tool in the ribbon right click menu to get accurate secondary structure assignment. This is particularly important when studying helices which may have non-cannonical elements within them such as 3/10 or pi. To view non-cannonical helix segments use the segment option in the ribbon right click menu.

To change the display of chain breaks (dotted lines):

- Click and hold on the ribbon representation button.
- Select the options Display Chain Breaks or Display Chain Break label.

5.1.4 Skin Representation

Click and hold on the **skin representation button**. A menu will be displayed as shown below.



To make some parts of your picture clearer, the skin representation can be set to tight, normal or sparse:

- Click and hold on the **skin representation button** and then click on either tight, normal or sparse.

To undisplay representations other than skin:

- Click and hold on the **skin representation button** and then click on **undisplay other representations**.

If you make a mistake or you are not happy with the way your structure is displayed with the skin representation:

- Click and hold on the **skin representation button** and then click on **** reset to default****.

NOTE: Sometimes due to singularity problems holes may appear within the skin surface. To cure this infliction select atoms nearby and right click select Advanced->RandomizeAtoms

5.1.5 CPK Representation

Click and hold on the **CPK representation button**. A menu will be displayed as shown below.



To undisplay representations other than CPK:

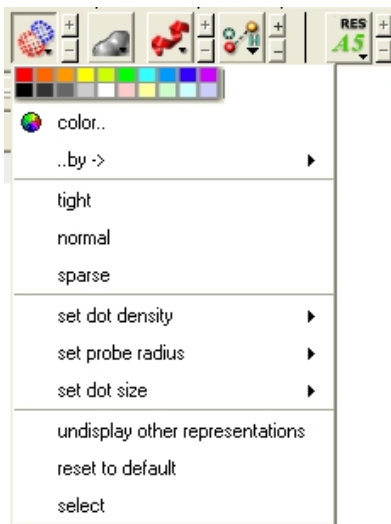
- Click and hold on the **CPK representation button** and then click on **undisplay other representations**.

If you make a mistake or you are not happy with the way your structure is displayed with the cpk representation.

- Click and hold on the **CPK representation button** and then click on **reset to default**.

5.1.6 Surface Representation

Click and hold on the surface representation button. A menu will be displayed as shown below.



To change the style of the surface representation:

- Click and hold on the **surface representation button** and then click on **tight, normal**, or **surface**.

To undisplay representations other than surface:

- Click and hold on the **surface representation button** and then click on undisplay other representations.

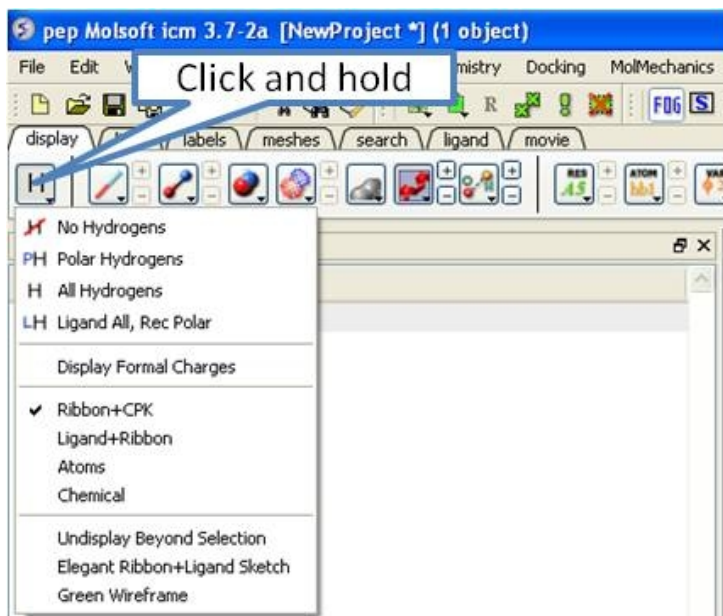
If you make a mistake or you are not happy with the way your structure is displayed with the surface representation:

- Click and hold on the **surface representation button** and then click on **reset to default**.

5.1.7 Display and Undisplay Hydrogens

To display and undisplay hydrogens. Click and hold on the "**Change Hydrogen Display**" button shown below. Multiple single clicks will toggle through the hydrogen display options.

- Display Tab
- Click and hold on the "**Change Hydrogen Display**" button shown below.

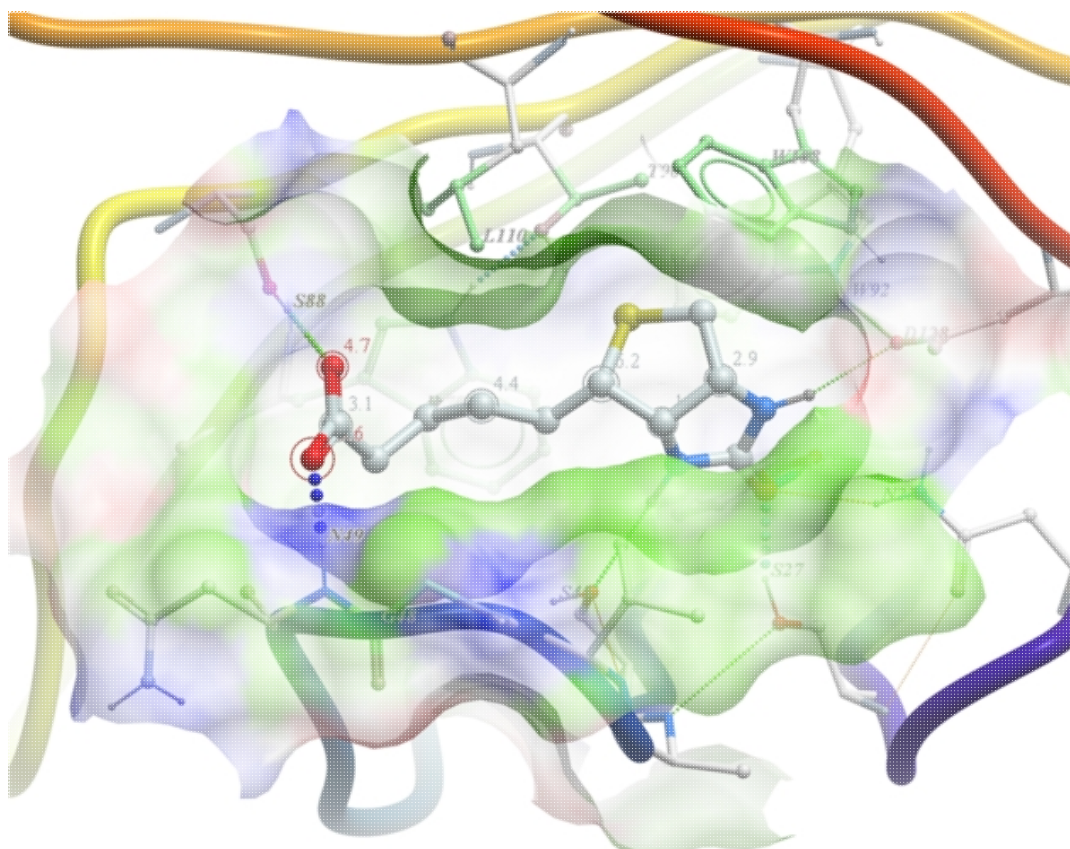
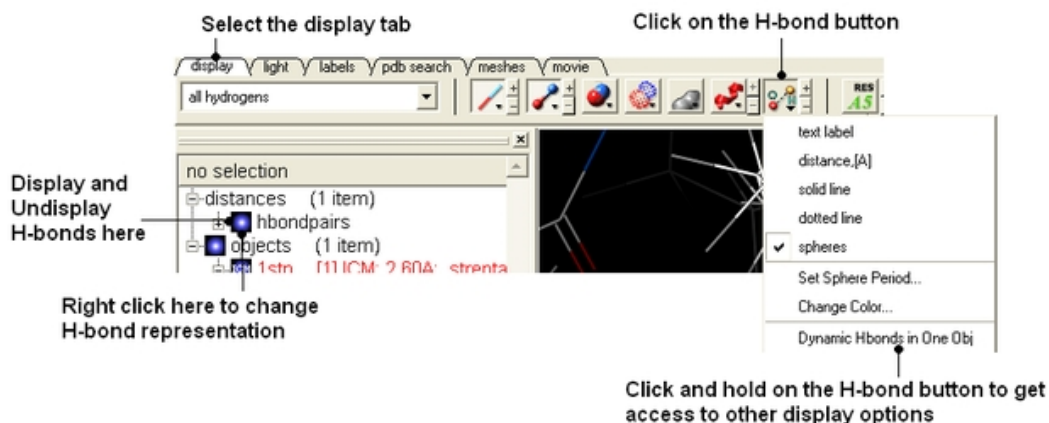


5.1.8 Display Hydrogen Bond

NOTE: The method by which hydrogen bonds are calculated is described [here](#) in the command line manual. The GRAPHICS.hbondMinStrength parameter determines the hbond strength threshold for hbond display. The strength value is between 0. and 2. By changing 1. to 0.2 you will see more weak hydrogen bonds.

In order to display potential hydrogen bonds in your structure:

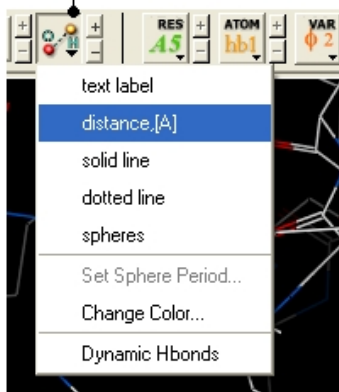
- Convert to an ICM Object
- Make a selection - if you are trying to display the H-bonds between a ligand and the receptor make sure the ligand is part of the selection.
- Click the Display Tab.
- Click on the **Toggle H-bonds** icon in the **display** tab.



- Click the +/- on the right of the H-Bond button to change thickness of H-bond representation.

- Click and hold the button to change representation or use the **hbondpairs** option in the ICM Workspace.

Click and hold



What do the default coloring of the H-bond represent?

Longer and shorter H-X distances in the hydrogen bond are color-coded, from red to blue, respectively.

NOTE Dynamic hydrogen bonds can be set by clicking and holding on the **H-bond toggle** button in the **Display** tab. Hydrogen bonds will then respond to any changes made to the ligand.

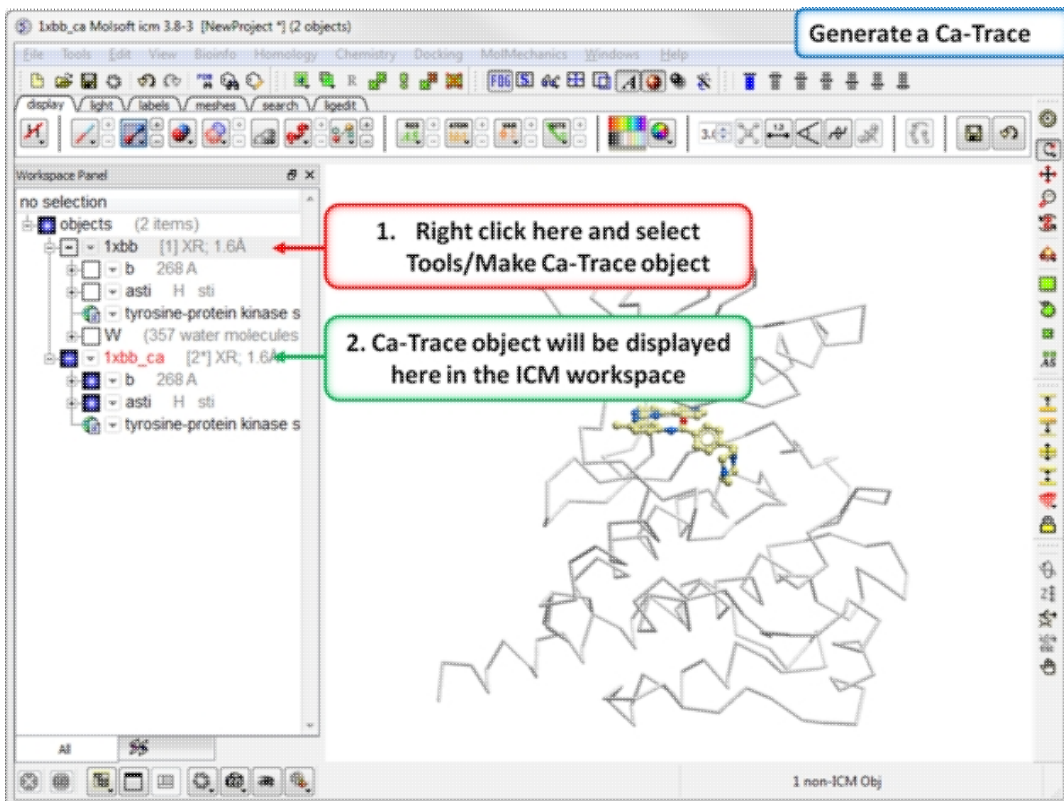
5.1.9 Display Formal Charges

You can display formal charges by clicking and holding on the "**Change Hydrogen Display**" button in the **Display** tab.

5.1.10 Ca (carbon alpha) Trace

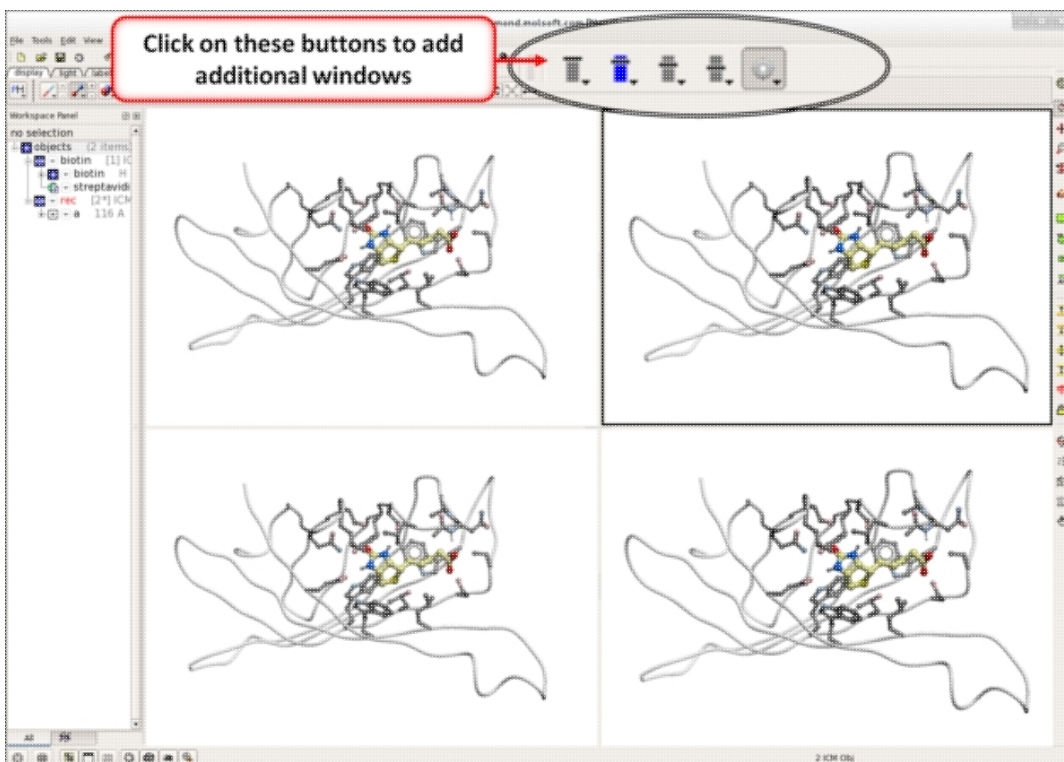
To generate a carbon alpha (Ca) trace:

- Right click on the object in the ICM workspace and select Tools/Make Ca-Trace object.
- The Ca-Trace object will be displayed in the ICM workspace.

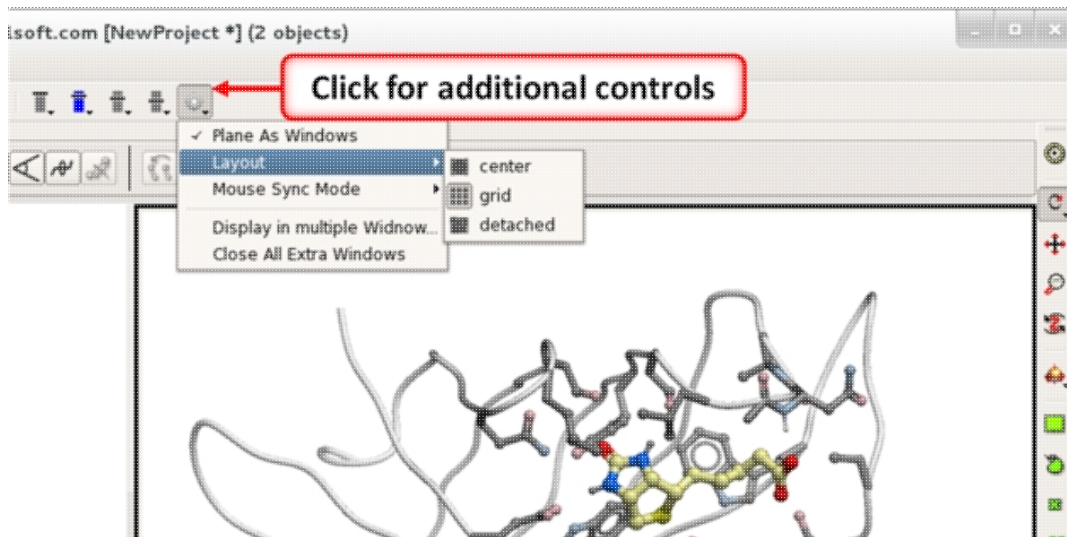


5.2 Multi-Windows

To display more than one window click on the buttons shown below.



To control the layout and mouse sync mode click on the button shown below.



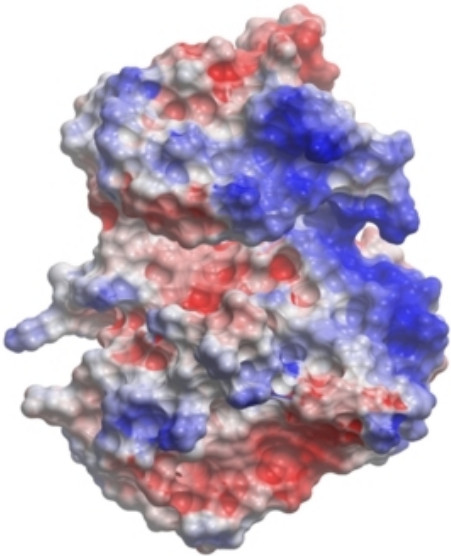
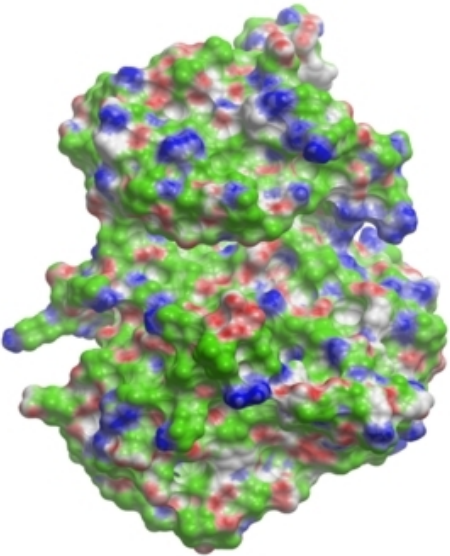
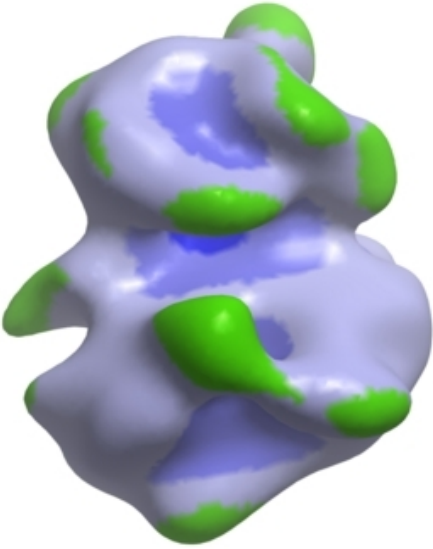
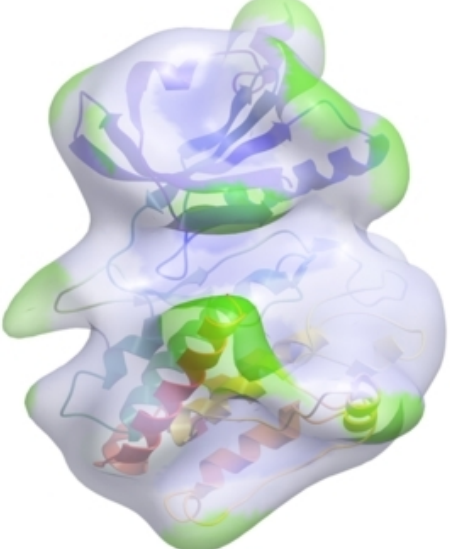
5.3 Meshes - Surface - Grobs

Available in the following product(s): ICM-Browser-Pro | ICM-Pro

Click on the tab button entitled '**meshes**' and more graphics tools for surfaces are available. In ICM surfaces are sometimes referred to as meshes or graphical objects (Grobs).



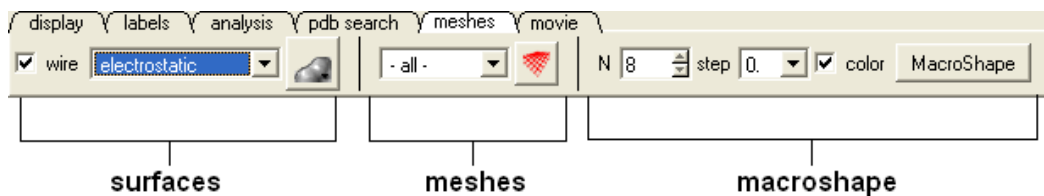
Surface	Plain Solid

	
Electrostatic Surface	Binding Property Surface
	
Macroshape Solid	Macroshape Transparent

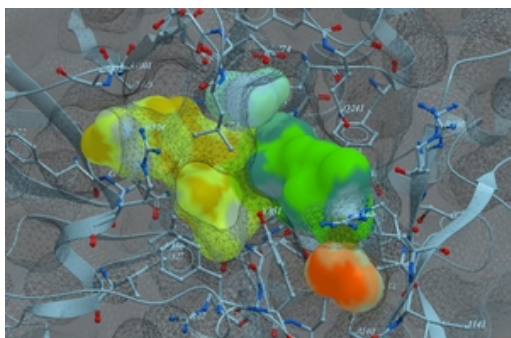
5.3.1 Surfaces

The surface of your structure can be displayed and colored by **electrostatics** or **binding** properties. To do this:

- Load a structure into ICM File/Open or tab-pdb{PDB Search}
- Convert the structure into an ICM object.
- Select the 'meshes' tab button.
- Click on the drop down arrow menu shown below and select which surface you wish to generate.
- Click on the generate surface button next to the drop down arrow.



5.3.2 MolSkin



MolSkin is a new macro that instantly makes publication quality graphics and it gives each molecule a color and colors contact patches. It undertakes the following:

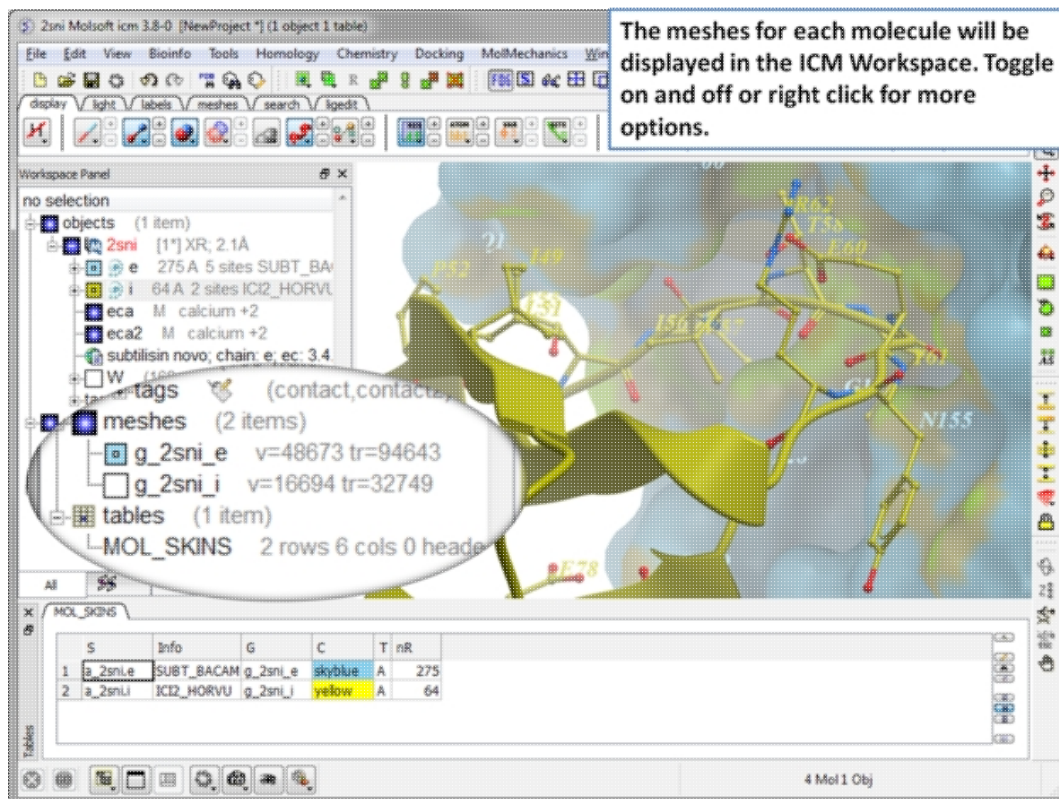
- molecular surfaces are generated with those assigned colors and occlusion shaded
- carbons of each molecule are colored by its own consistent color
- labels, ribbons, are colored accordingly
- surfaces are gradually cross-colored by each other to mark the contact patches
- the labels are brought to the surfaces

To use MolSkin:

The screenshot shows the ICM software interface with a right-click context menu open over a molecule. The menu options are: Skin Mesh (with sub-option Publication View), Tools, Edit, Advanced, Center, Default Display, Connect to Object, Extract Sequence(s), and Update. A callout box points to the 'Skin Mesh/Publication View' option with the text: '1. Right click here. Skin Mesh/Publication View'. Another callout box points to the 'Publication View' sub-option with the text: 'Load the molecule into ICM and then right click on the object name and choose Skin Mesh/Publication View'. Below the menu, a table shows the following data:

S	Info	G	C	T	nR
1	a_2sniLe SUBT_BACAM g_2sni_e	skyblue	A	275	
2	a_2sniL ICI2_HORVU g_2sni_L	yellow	A	64	

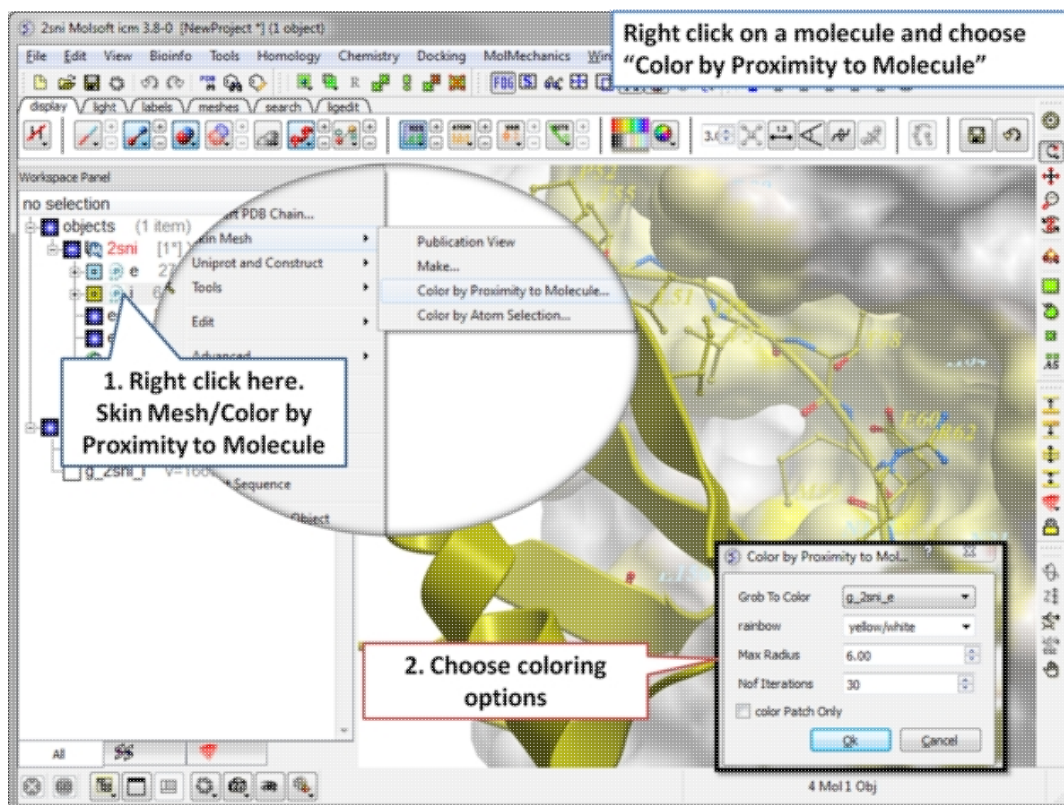
Read into ICM a PDB file. Next, right click on the object name and choose Skin Mesh/Publication View .



The surfaces will be displayed in the meshes section of the ICM Workspace.

5.3.3 Color Surface by Proximity

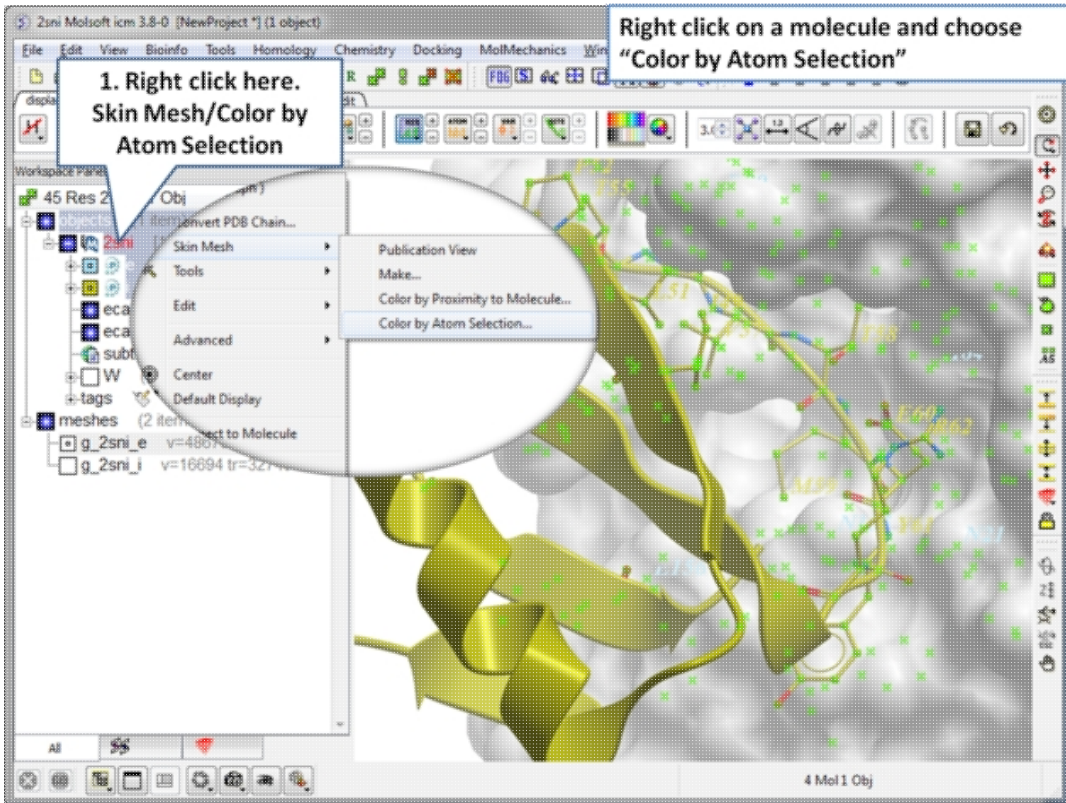
To color a surface by proximity to neighboring molecules:



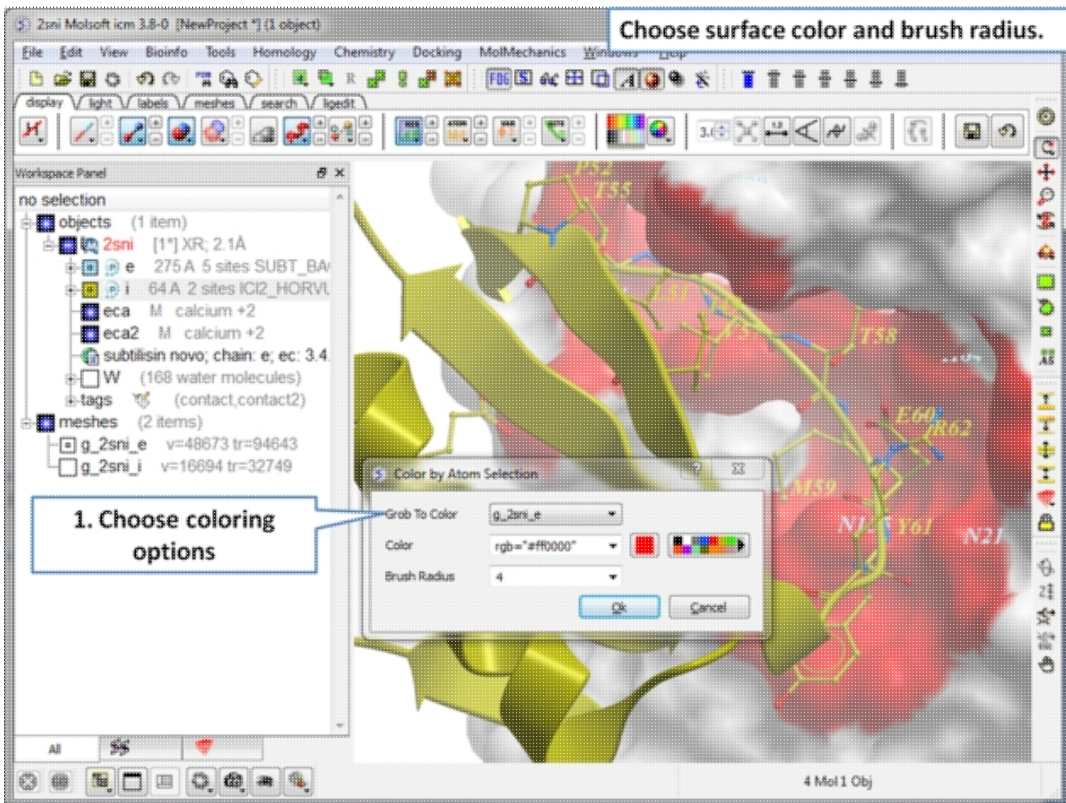
Create a surface then right click on the molecule and choose "Skin Mesh/Color by Proximity to Molecule".

5.3.4 Color Surface by Selection

To color a surface by a selection:



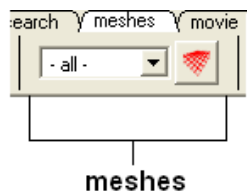
Create a surface and then make a selection. Right click on the molecule and choose Skin Mesh/Color by Selection.



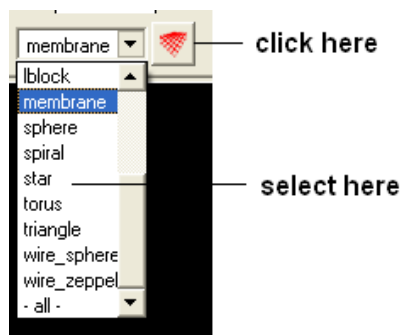
Choose the color and brush radius.

5.3.5 Meshes

A variety of shapes can be constructed automatically using ICM. These shapes are referred to as meshes. The types of shapes you can build are shown in the drop down option in the meshesh tab

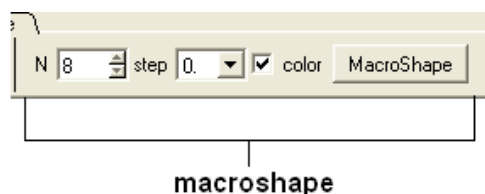


To make a shape select it from the menu by clicking on the down arrow and then click the button next to the menu. The shape will then be displayed in the 3D graphics window.



5.3.6 Macroshape

A macroshape can be constructed and allows easy viewing and manipulation of the structural representation. A macroshape representation is ideal for large structures which allows the user to easily identify important regions of the structure and facilitate the return to the 'standard' view of a particular molecule. All the buttons needed to display a macroshape structure are shown below in the 'meshes' tab.



To construct a macroshape:

- Load a molecule into ICM File/Open or tab-pdb{PDB Search}
- Select the amount of detail required in the shape by increasing the values in 'N' or 'step' data entry box (note the default values are usually sufficient).
- Check the 'color' if you wish your molecule to be colored.
- Click the button labeled 'MacroShape'.

Macroshape can also be used from the View menu: View/Macro Shape

5.3.7 Google 3D Objects (Sketchup)

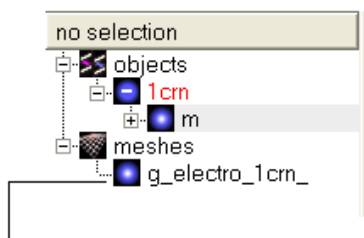
To read in a 3D Mesh from Google in KMZ or COLLADA format:

- File/Load/ 3D Mesh in KMZ or COLLADA Format from Google
- Search for the object you would like to view and download it.
- To read the file go to File/Open

5.3.8 Display or Undisplay Meshes or Surfaces

To display or undisplay the surface click in the box in the ICM workspace as shown below:

ICM Workspace

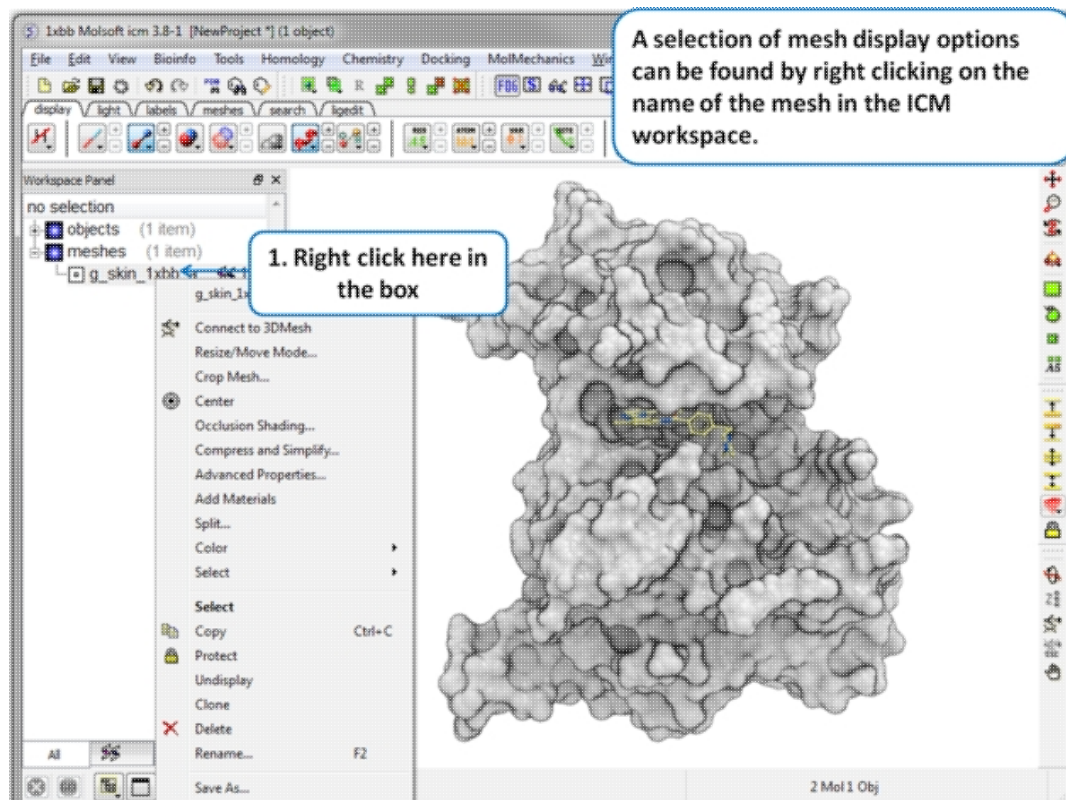


check here to display
or undisplay surface

NOTE: All surfaces, meshes and macroshapes come under the one heading of **meshes** in the workspace panel.

5.3.9 Mesh Options.

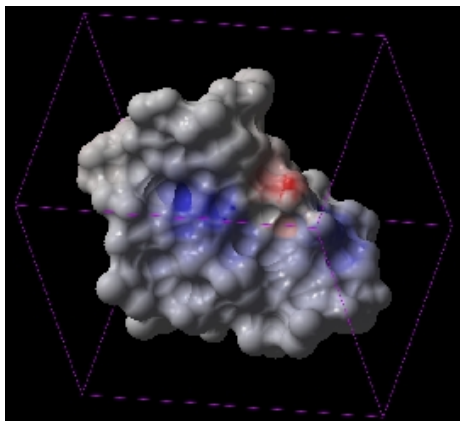
A number of options relating to meshes can be found by right clicking on the mesh in the ICM Workspace. This section describes some of these options.



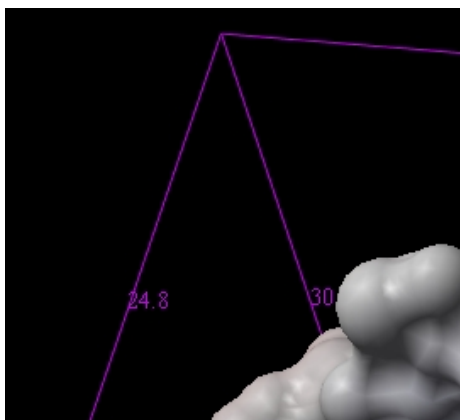
5.3.9.1 Resize Mesh

Once a mesh has been created you can move it and resize it. To do this:

- Locate the mesh in the ICM Workspace and right click on it
- Select the **Resize/Move Mode** in the menu.
- A purple box as shown below will surround the molecule.



To resize the mesh click on one of the corners of the box and drag to the required size. The number displayed on the edges of the box represent the dimensions.



5.3.9.2 Move Mesh

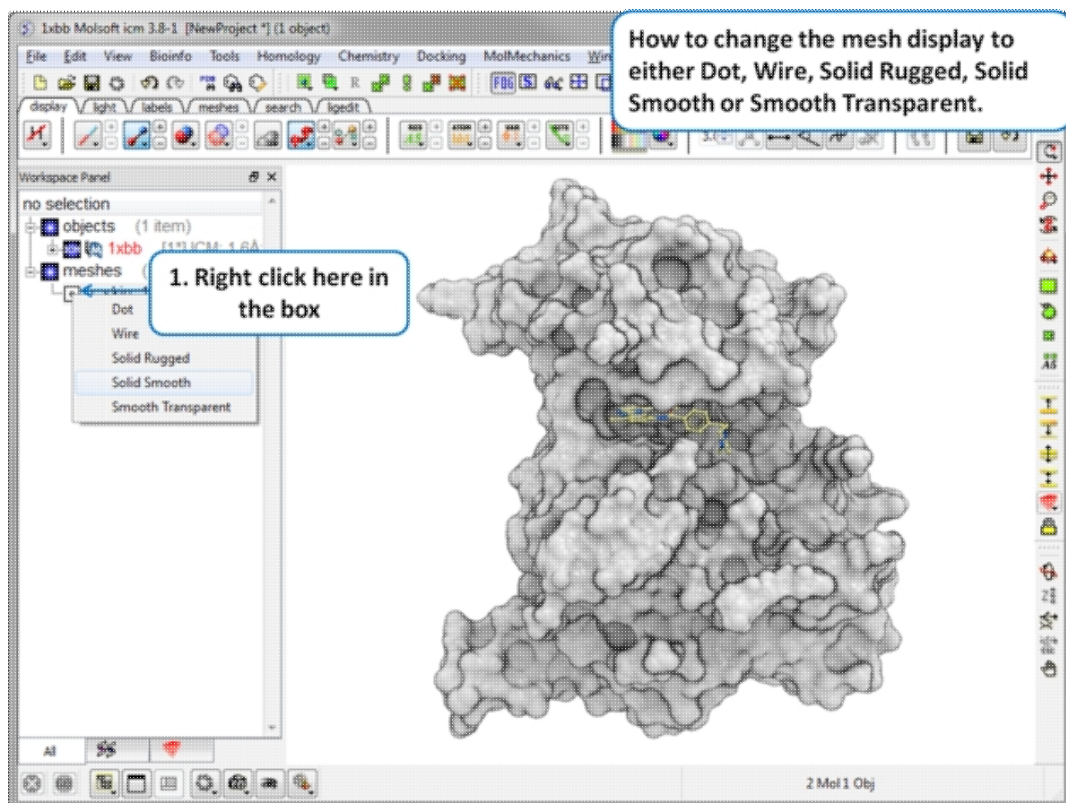
To move a mesh:

- Locate the mesh in the ICM Workspace and right click on it
- Select the **Resize/Move Mode** in the menu.
- Click on the mesh in the graphical display with the **middle** mouse button.

5.3.9.3 Mesh Representation

There are five different display modes:

- Dot
- Wire
- Solid Rugged
- Solid Smooth
- Smooth Transparent.



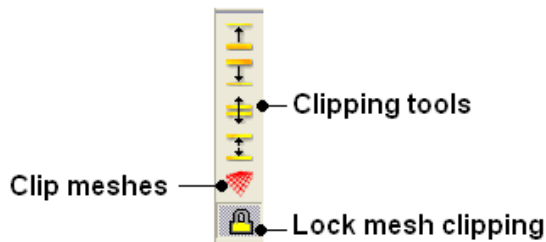
5.3.9.4 Mesh Color and Lighting

There are a number of options to color and change the display of the mesh. These options can be accessed simply by right clicking on the mesh name in the ICM Workspace as shown below.

5.3.10 Mesh Clipping

Clipping tools can be used to adjust the frames of the mesh independently of other objects.

The buttons shown below can be used for this purpose.



The buttons used for clipping are described in the section entitled [Clipping Tools](#).

To clip the skin independently of the object as shown in the image above:

- Display object and mesh.
- Click and hold on clip meshes button (red button) and select Reset Clipping to Default and check the option Clipping enabled
- Click on the front clipping plane button.
- (Optional) To lock the mesh clipping click on the yellow lock mesh clipping button.

5.3.11 Save Mesh

You can save a mesh as a wavefront object by right clicking on the mesh in the ICM Workspace and selecting **SaveAs**.

5.3.12 Occlusion Shading

The occlusion shading option provides better representation of depth within a cavity. The color of each surface element of a grob (mesh) is changed by mixing its own color with the background depending on the burial of the surface element.

To add occlusion shading:

- Right click on the mesh in the ICM Workspace and select **Occlusion Shading**. The occlusion shading value can also be changed before generating the mesh in the meshes tab.
- Enter a depth value - default is 0.8. Higher values will generate a more dramatic shading.

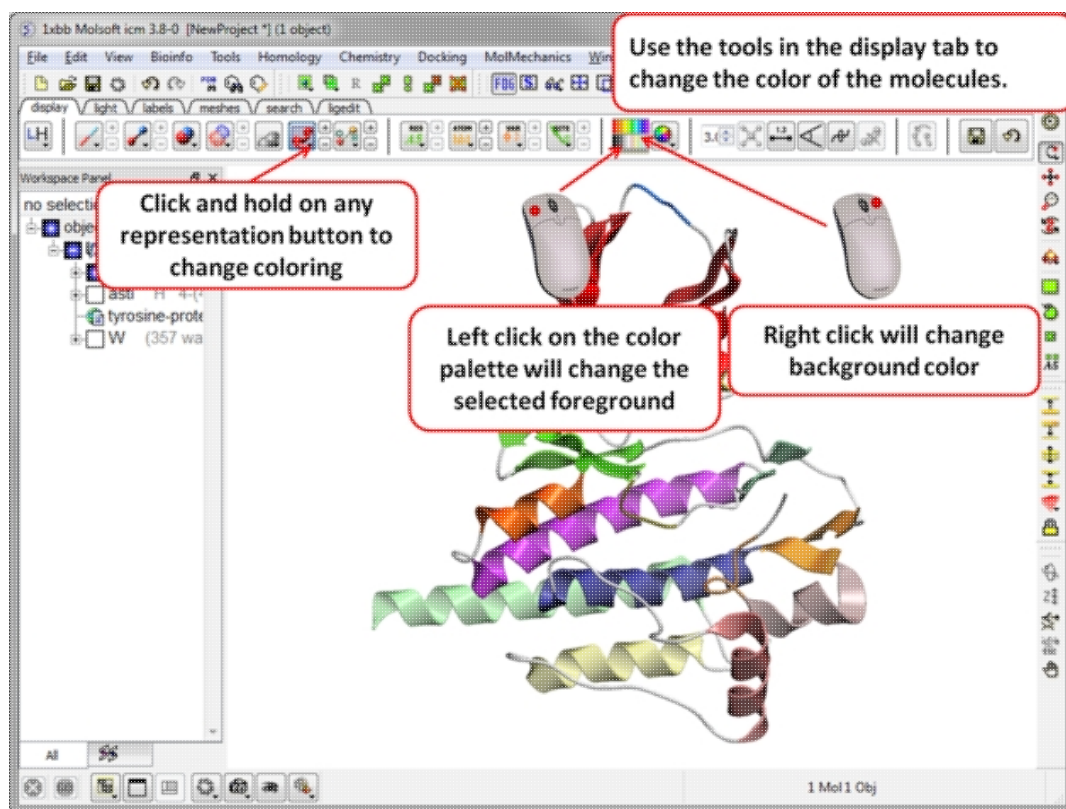
See example here.

5.4 Coloring

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

To change the coloring of molecules:

- Select the atoms, residues, molecules, or objects you wish to color in the graphical display or in the ICM Workspace.
- Then use the color options in the Display Tab.



5.4.1 Coloring

To change the color of a structural representation such as CPK, Xstick, wire or ribbon.

- Select the atoms, residues, molecules, or objects you wish to color.
- Click and hold on the structural representation button for the representation you wish to color (e.g. wire, ribbon etc...) in the **Display** tab.
- Select a color by clicking color.

To color by a particular parameter such as atom type, b-factor, secondary structure etc...

- Click and hold on the structural representation button for the representation you wish to color (e.g. wire, ribbon etc...) in the **Display** tab.
- Select .by-> option

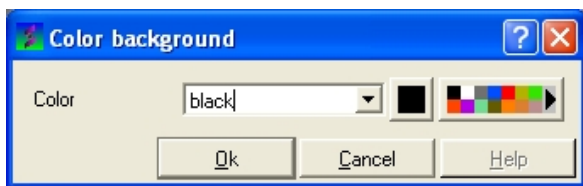
To change the color of everything displayed :

Click on the color palette in the `Display` Tab.

5.4.2 Color Background

To change the color of the background:

- Select **View/Color** background.



- Click on the square of your desired color. If you are not satisfied with the color palate, click on the arrow next to the colors to customize a color.

OR

- Right click on a color in the colors panel in the display tab.



Right click here in the display panel to set the background color.

5.4.3 Background Image

A background image can be added to the graphical display. This can be useful for making cool images or for comparing structures (e.g. compare displayed object with background image of object).

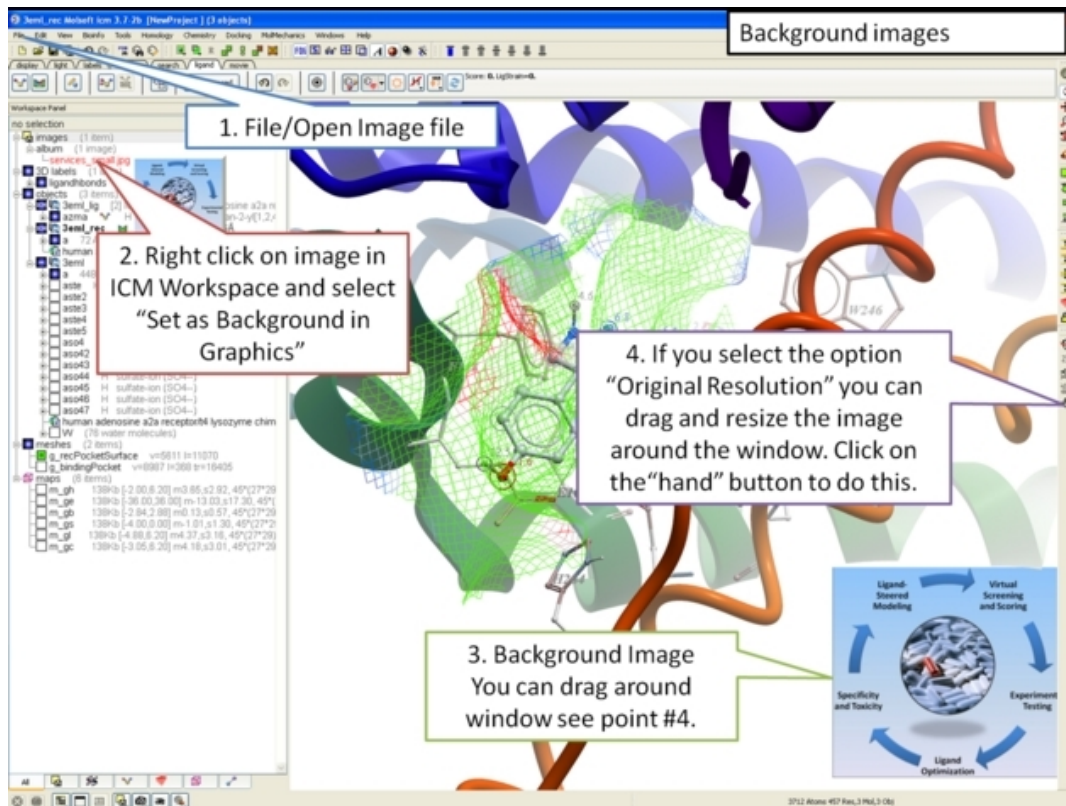
To add a background image from an image file (png or jpeg):

- File/Open Image
- Right click on the image in the ICM Workspace and select "Set as Background in Graphics."
- Choose one of the following: (1) Original Resoution, (2) Original Resolution Centered, or (3) Scale to Fit
- Enter the Scale (%)
- Enter whether you want to keep existing background.

To move and resize a background image: Version 3.7-2b and higher.

- File/Open Image

- Right click on the image in the ICM Workspace and select "Set as Background in Graphics."
- Choose the option **Original Resolution**.
- Click on the drag atom button (looks like a hand).
- Click and drag image.
- Remember to deselect drag atom button.

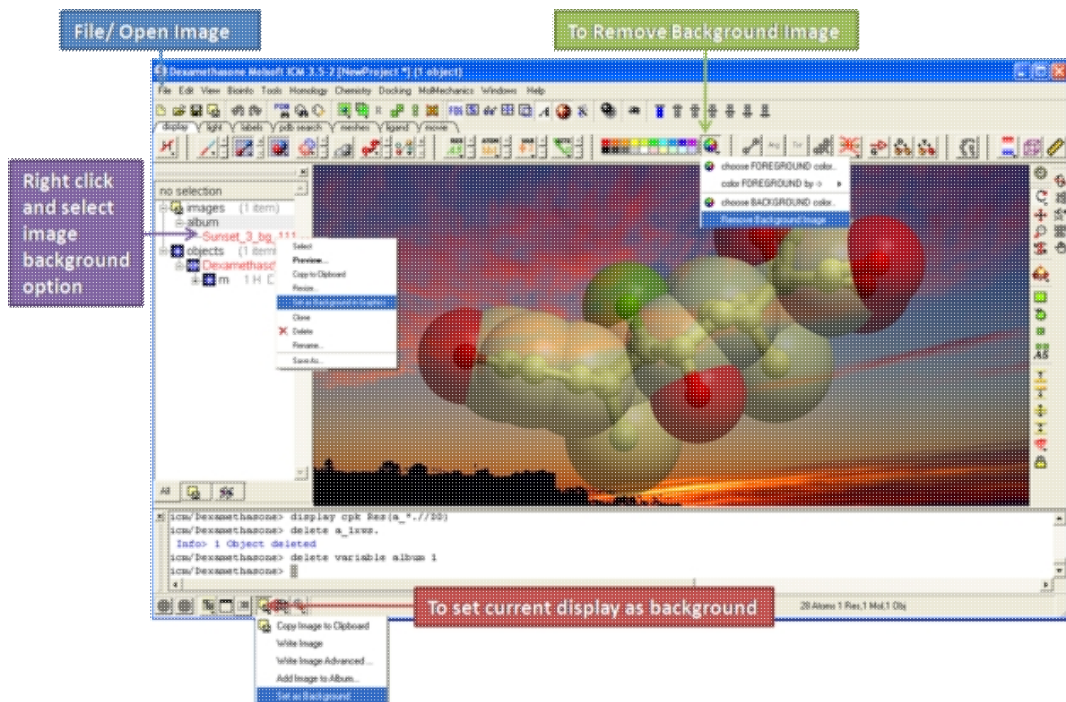


To set current display as background image:

- Click and hold on the "Copy Image to Clipboard" button at the bottom of the gui and select the "Set as Background" option.

To remove a background image:

- Select the **display** tab and then click and hold on the color sphere button and select "Remove Background Image".



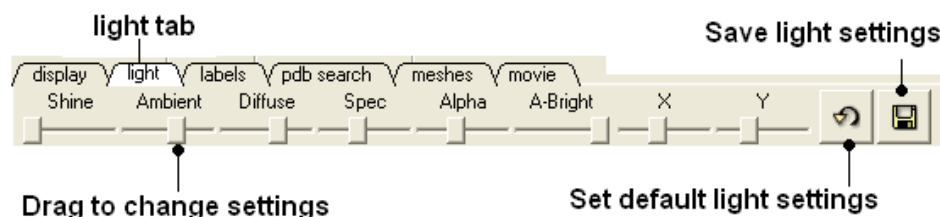
OR

- Right click on the image in the ICM Workspace and select "Set as Background in Graphics."
- Choose the option **Remove From Background**

5.5 Lighting

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

These options are in the `light` tab



Click and drag the sliders to change the lighting. You can also save your preferred lighting settings and return to default.

Shine - shininess property of the solid material

Ambient - ambient light intensity of RGB for ambient light

Diffuse - diffuse light intensity of RGB for diffuse light

Spec - specular light intensity of RGB for specular light

Alpha - transparency setting for grob

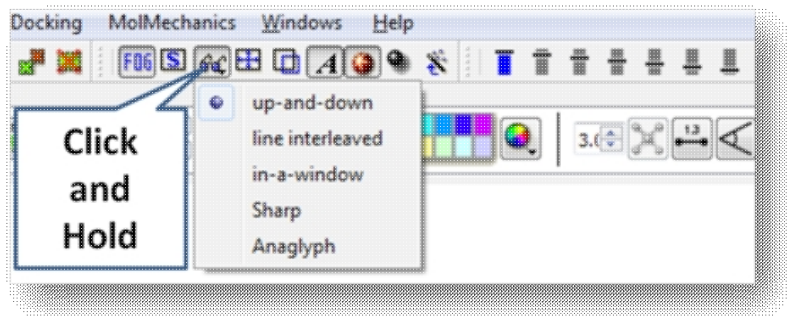
A-Bright - light intensity shining on grob

X and Y - Change the position of the light source in the graphics window

5.6 3D Stereo

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

Click and hold on the stereo hardware button (see image below).

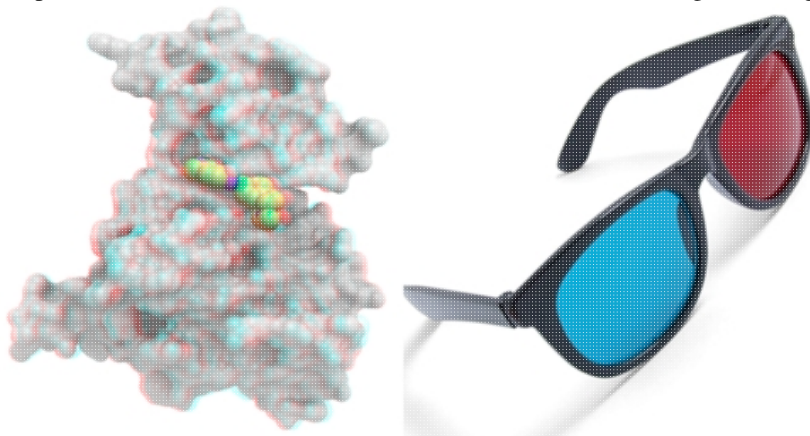


There are 5 different options available which are described here

<http://molsoft.com/man/icm-table.html#GRAPHICS.stereoMode>.

1. up-and-down
2. line interleaved
3. in-a-window
4. Sharp
5. Anaglyph

The Anaglyph option is the easiest to use with inexpensive 3D glasses and without any expensive 3D compatible hardware or monitors. The 3D effect is better with a lighter background.



5.7 3D Printing

You can save surfaces in the STL format for 3D printing. To do this:

- Make a mesh e.g. go to the meshes tab - plain solid.
- The mesh (surface) will then be displayed in the ICM workspace (left hand side).
- Right click on the mesh and choose Save as.. and choose stl from the format option list

5.8 Labeling and Annotation

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

To add labels or display or undisplay pre-defined annotation:

- Select the atoms, residues, molecules, or objects you wish to label in the graphical display or in the ICM Workspace.
- Then use the label options in the `Display Tab`.

To add new user-defined annotation:

- Select the atoms, residues, molecules, or objects you wish to label in the graphical display or in the ICM Workspace.
- Right click on the selection and choose "Annotate Selection".

5.8.1 Labeling

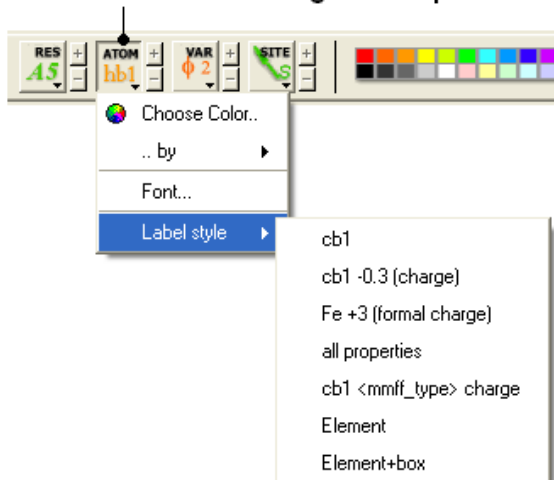
Labeling options are contained within the `Labels` or `Display Tab`. In many cases clicking and holding a label button will allow you to view more options.

5.8.2 Labeling Atoms

Select the atoms you wish to label (see `display structure` or `selection toolbar`).

- Select the **display** tab.
- Click the **label ATOM** button.

Click and hold to change label options



To change the level of label detail:

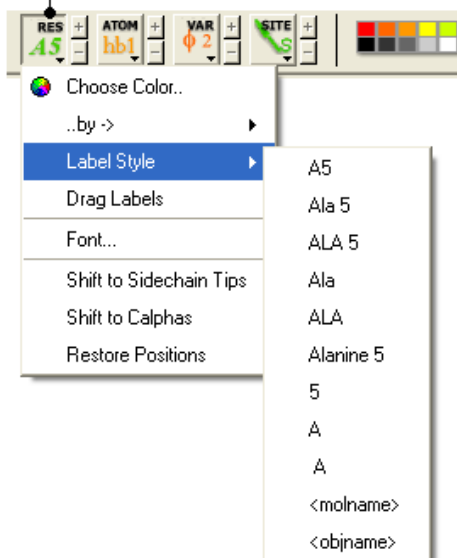
- Click and hold the **label ATOM** button and select the desired level of label detail, color or style.

5.8.3 Labeling Residues

To label residues:

- Select the **display** tab.
- Select the residues you wish to label (see `display structure` or `selection toolbar`).
- Click the **label RES** button.

Click and hold for more options



To change the level of label detail:

- Click and hold the **label RES** button and select the desired level of label detail or style.

5.8.4 Move Residue Label

To change the location of your residue label:

- Select View/Drag res labels.
- If your mouse has a middle mouse button, then click on handle (as shown) of the label you wish to move, and drag it to your desired area.

Click on
this area
to drag
your
label.



- If your mouse has no middle mouse button, then click on the Translation icon on the toolbar, and click on the handle (as shown) of the label you wish to move, and drag it to your desired area.

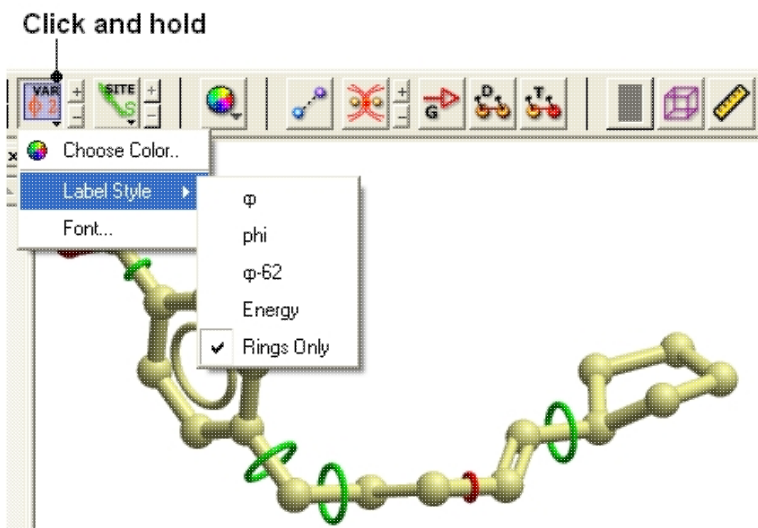
The +/- buttons on the side of the Residue and Atom buttons will shift the label. There are also other **residue label move** options available when you click and hold the residue label button. These options include **Shift to Sidechain Tips**, **Shift to Calphas**, and **Restore Positions**

5.8.5 Label Variables

To label variable angles (dihedral-torsion, planar and phase angle) the molecule needs to be converted into an ICM object.

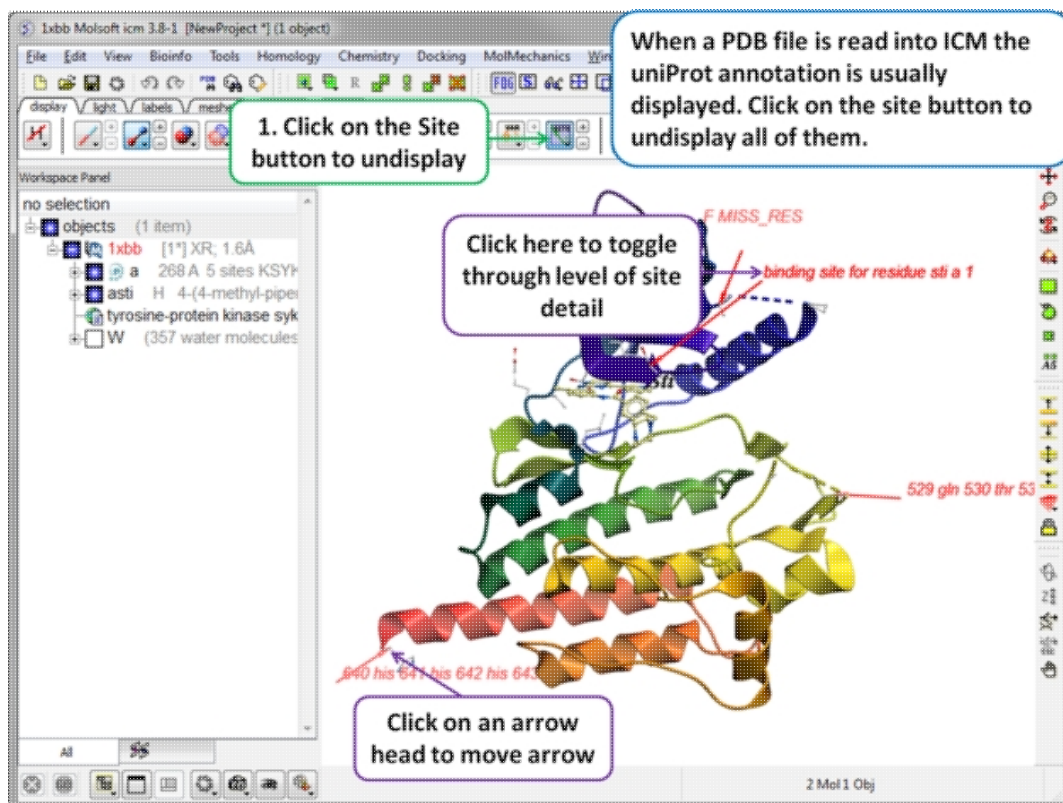
- Convert the molecule to an ICM object.
- Select the atoms for which you would like to display the variables (see `display structure` or `selection toolbar`).
- Click on the **toggle variable label** button shown above located in the **display** tab.
- Change the font size by using the +/- buttons.
- Change the font and color by clicking and holding on the variable atom label button.

To change the variable label style click and hold the variable atom label button as shown below.

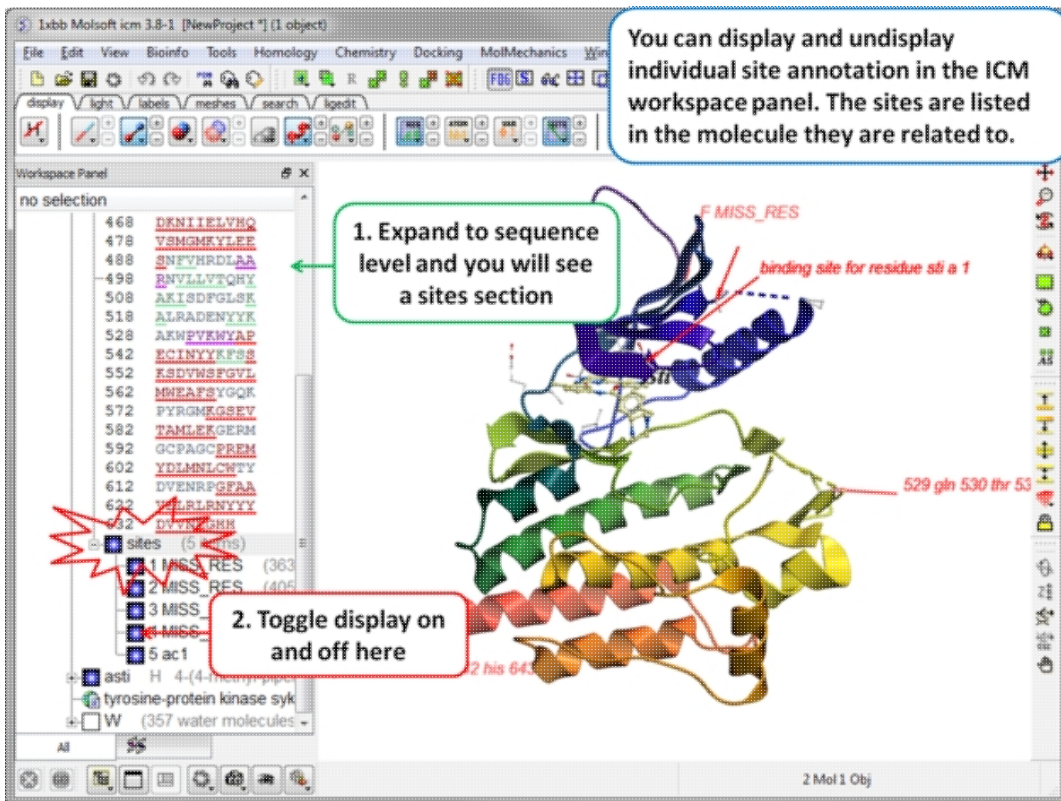


Rings of varying diameter and color are superimposed on rotatable bonds. Green rings with large diameter are considered less constrained than rings with small green rings. Red rings are highly constrained. When the **Label Style/Energy** option is selected the first number displayed represents the bond angle, the second the energy and the third the worst energy that could be achieved by rotating the bond.

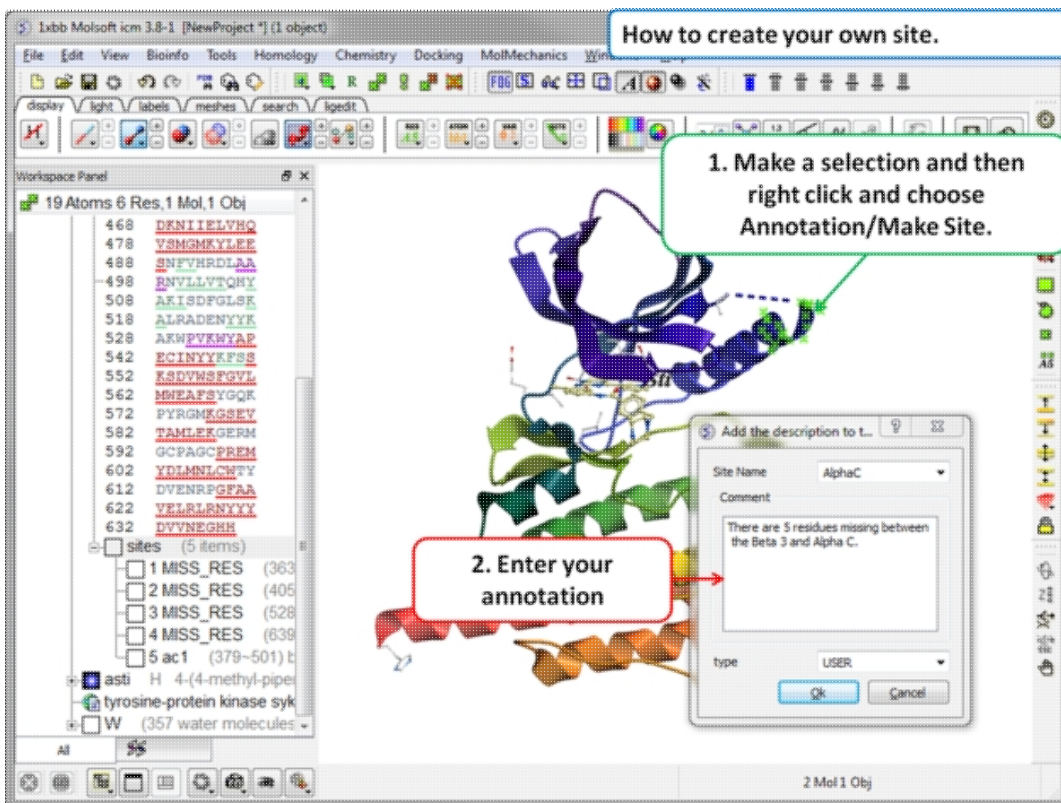
5.8.6 Sites and Annotation



When a PDB file is read into ICM the sequence functional/mutation sites listed in Uniprot are automatically displayed. To undisplay all displayed sites click on the **Sites** button in the `display` tab. Click in the bottom left hand corner of the annotated site to change the level of detail or click on an arrow head to move the location.



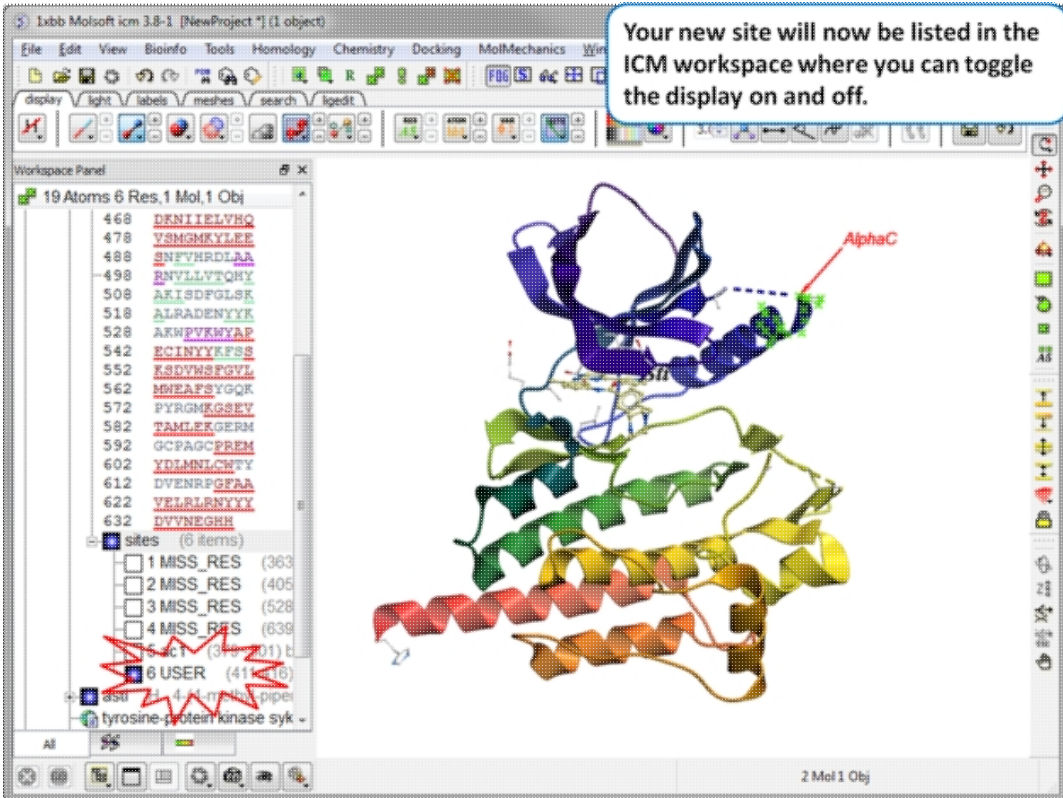
Individual sites can be displayed or undisplayed in the ICM workspace.



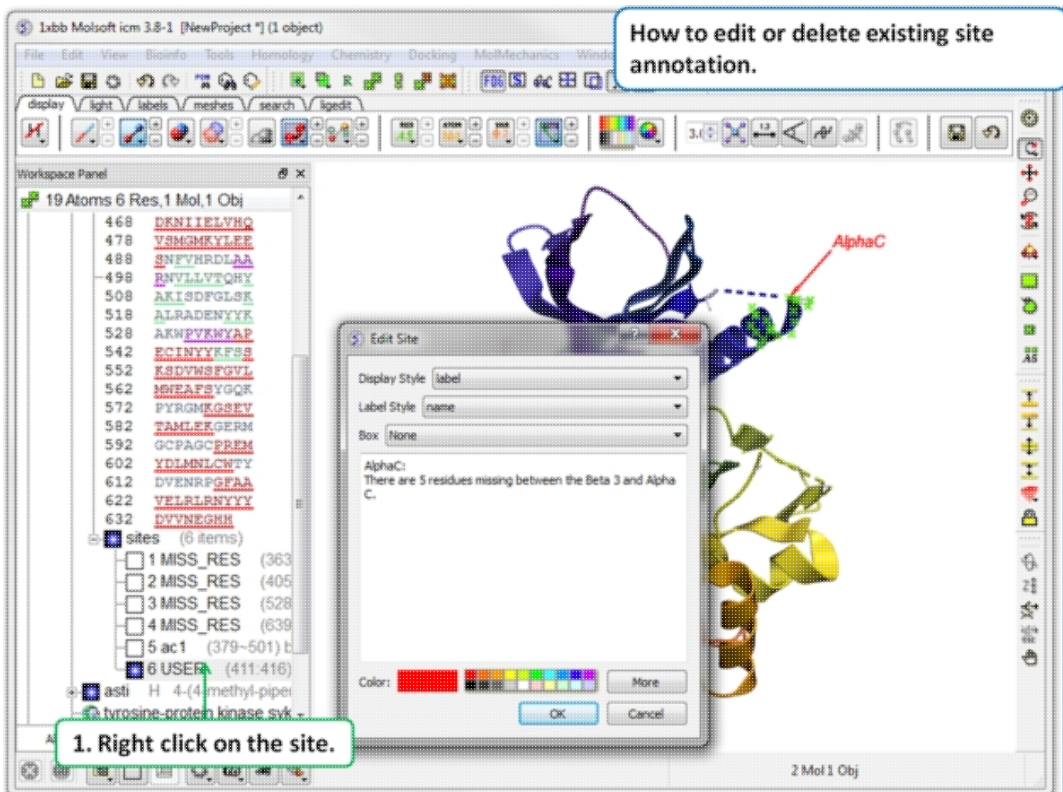
To create your own annotated site:

- Select the region you wish to annotate as a Site.
- Right click on the selection.

- Choose the option Annotate/Make Site
- Enter the annotation into the text box and select ok



Your new site will now be listed in the ICM Workspace.



To edit or delete a site:

- Right click on the site in the ICM workspace.

5.8.7 Changing Label Colors

To change the color of any label:

- Click and hold down the required label button and a menu as shown below will be displayed.

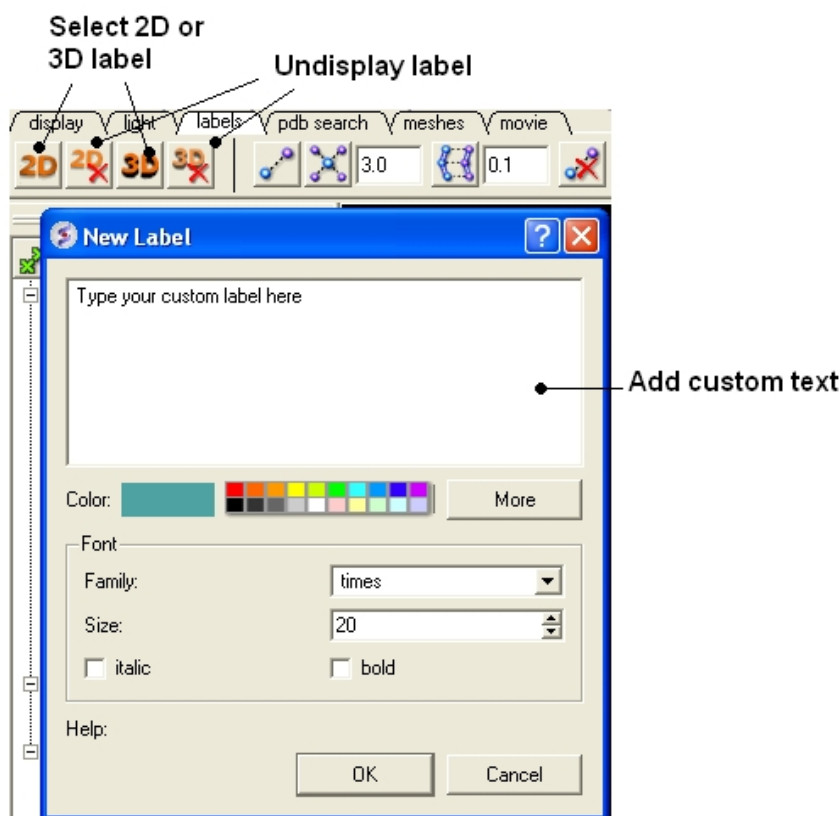


- Select **color**.

5.8.8 Customized Label 2D or 3D

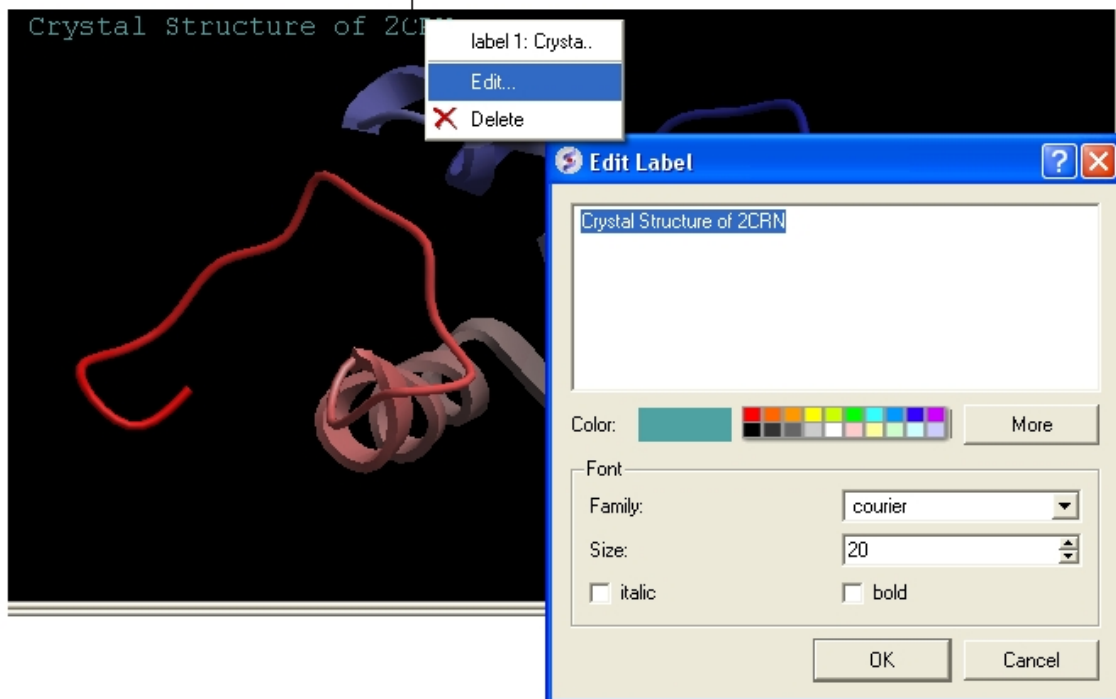
To generate a customized a label:

- Select the **labels** tab.
- Select either 2D or 3D button.
- Enter your label and select the desired color, font and size.



To **edit** or **delete** a label - right click on the label in the graphical display as shown below.

Right click here to Edit or Delete label



5.8.9 Undisplay Customized Label

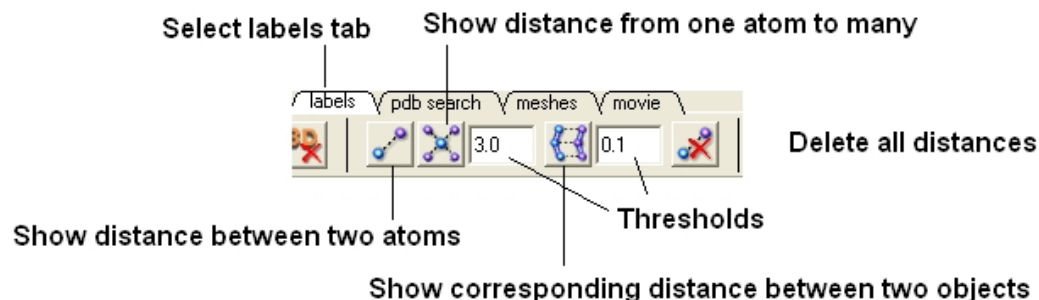
Undisplay Residue, Atom, and Variable Label Any label that is displayed can be undisplayed by selecting the region of the molecule related to the label and clicking on the corresponding label button in the labels tab. For example if you wish to undisplay an atom label - click the atom label button. If a label is displayed the corresponding button in the **display** tab will be shaded blue. When you delete the button will return to grey. 2D and 3D labels have an undisplay button (red cross on the button see customized label section).

Undisplay 2D or 3D label Click on the undisplay label button in **labels** tab.

NOTE: A label can also be deleted by right clicking on the label in the graphical display and selecting **delete**.

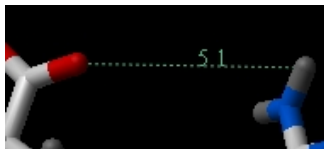
5.8.10 Labeling Distances

Within the **labels** tab there are tools for calculating and displaying distances. These tools can also be found in the Tools/Analysis menu.



To display distance between two atoms:

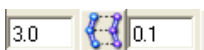
- Click on the labels tab (previously called advanced tab).
- Select the atoms between which you would like to find the distance. (See `selection toolbar`)
- Click on the 'Show Distances Between Two Atoms' Button
- The distance will be displayed in angstroms, in green.



To find the distance from one atom to many:

- Click on the labels tab (previously called advanced tab).
- Select the atom from which you wish to measure the distance from (See `selection toolbar`)
- Click on the 'Show Distances From One Atom To Many' button.
- The distances will be displayed in green.

The maximal and minimal distances can be selected by entering values in the boxes shown here (below) in the labels tab (previously called Advanced tab).



NOTE: Distances can be displayed and undisplayed in the 3D labels section of the ICM Workspace. See image below.

To change the color of the distance label

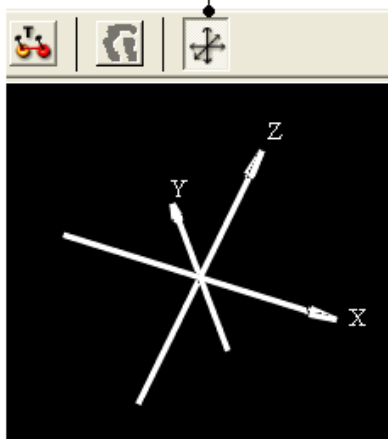
- Right click on the **distpairs** under the **3D labels** section of the ICM workspace and select **Change Color**.

5.8.11 (Un)display Origin

To display and undisplay the axis of the coordinate frame (origin):

- Select the **labels** tab and select the **toggle origin** button.

Display or undisplay origin button - located in the labels tab



5.8.12 Displaying Tethers

Theory

A tether is a harmonic restraint pulling an atom in the current object to a static point in space. This point is represented by an atom in another object. Typically, it is used to relate the geometry of an ICM molecular object with that of, say, an X-ray structure whose geometry is considered as a target. Tethers can be imposed between atoms of an ICM-object and atoms belonging to another object, which is static and may be a non-ICM-object. You cannot create tethers in ICM-Browser, however, if the project that you have loaded contains tethers between two objects, then they can be displayed:

- Click on the **display tab** (previously called advanced tab).
- Click on the 'Toggle Tethers' button.

5.8.13 Displaying Distance Restraints

Theory

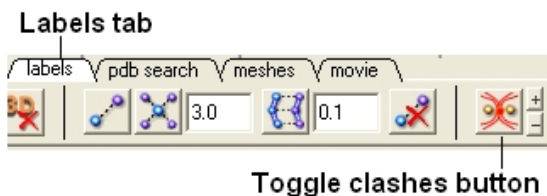
A distance restraint imposes a penalty function on the distance between two atoms in the same object. You cannot create distance restraints in ICM-Browser, however, if the project that you have loaded contains distance restraints, then they can be displayed:

- Click on the **display tab** (previously called advanced tab).
- Click on the 'Toggle distance restraints' button.

5.8.14 Display Clash

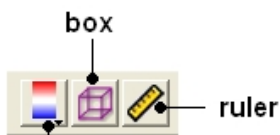
To display a clash the file needs to be an ICM Object.

- Select the region around which you would like to identify clashes.
- Select the display tabs and the "toggle clashes" button shown below.



5.8.15 Display Rainbow, Box, Ruler

To (un)display a rainbow scale, box or ruler use the buttons shown below located in the Labels Tab.

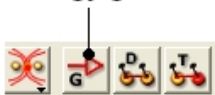


rainbow (click and hold to change colors)

5.8.16 Display Gradient

This button is located in the **display** tab.

Toggle energy gradient button



This option is described in detail in the language manual

<http://www.molsoft.com/man/icm-commands.html#display-gradient>

5.9 Display Distances and Angles

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

5.9.1 Display Distance Between Two Atoms - the quick way

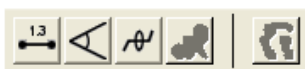
- Click on the **Display** tab
- Click on the **Distance between two atoms** button shown below.
- Click on the atoms you wish to measure.
- Distance will be displayed in the graphical display. You can turn this on and off in the ICM Workspace panel under the heading **3D labels**.



Display distance

5.9.2 Display Planar Angle

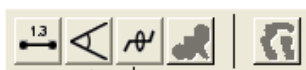
- Select the **display** tab.
- Select three atoms.
- Select the button shown below.



Display planar angle

5.9.3 Display Dihedral Angle

- Select the **display** tab.
- Select four atoms.
- Select the button shown below.



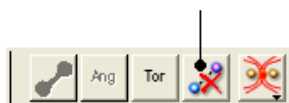
Dihedral angle

5.9.4 Delete Label

To delete distance or angle labels

- Select the **display** tab.
- Select the delete distance or angle label button shown below.

Delete distance or angle labels button in display tab



5.10 Graphics Effects

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

All the visual effects tools can be accessed by the **View** Menu or click on the corresponding button in the **View Tools** panel shown below.



5.10.1 Fog

Fog Toggle(Ctrl + D) : this feature creates a fog-like environment for your object, so that the part of your structure that is closer appears clear and the distant parts are faded as if they are in fog. The **clipping** planes control the point at which the fog begins.

- View/Fog

5.10.2 Shadows

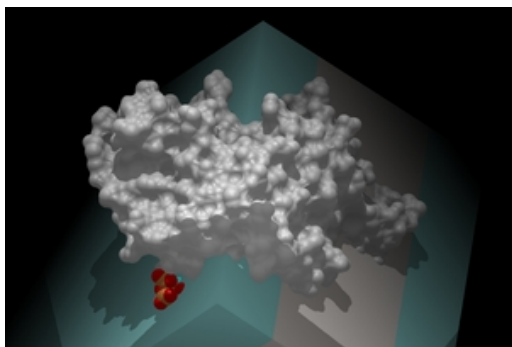
- View/Shadows

OR

select the shadow button shown below.

Toggle shadow

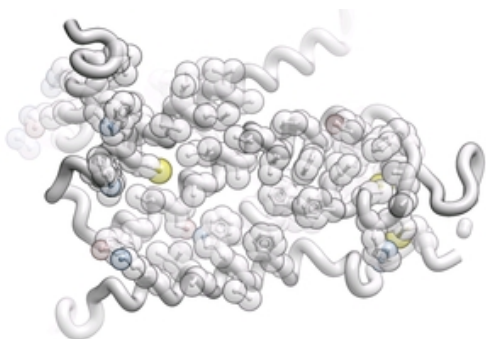




5.10.3 Sketch Accents

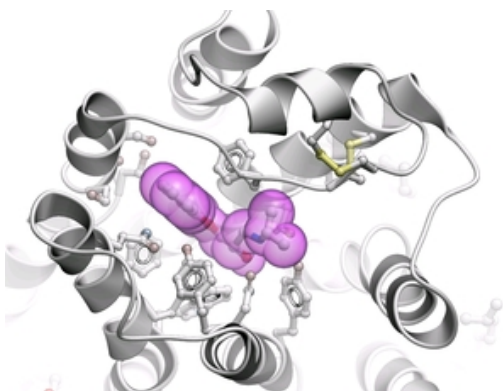
To make images as shown below use:

- View/Sketch Accents



5.10.4 Elegant Ribbon & Ligand Sketch

- Display Tab
- Click and hold Hydrogen button
- Select Elegant Ribbon+Sketch

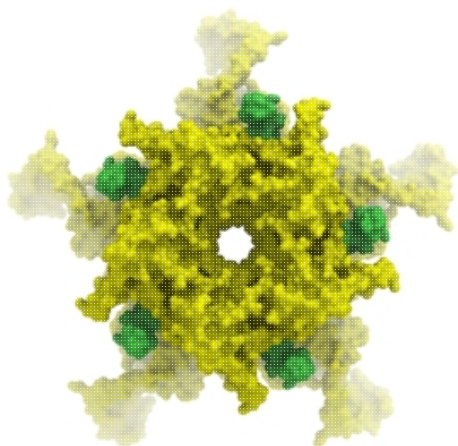


5.10.5 Occlusion Shading

Occlusion shading is for surfaces and meshes to give a perception of depth.

To add occlusion shading:

- Right click on the mesh in the ICM Workspace and select Occlusion Shading. The occlusion shading value can also be changed before generating the mesh in the meshes tab.
- Enter a depth value - default is 0.8. Higher values will generate a more dramatic shading.



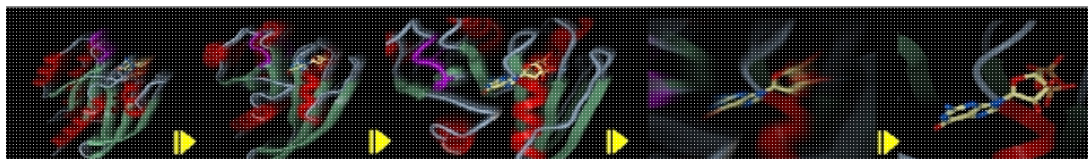
5.10.6 Perspective

Toggle perspective Ctrl_P this will add perspective to your structure, enhancing depth in the graphical display.

- View/Perspective

5.10.7 Animate View

Learn how to build fully interactive and interruptable animations.

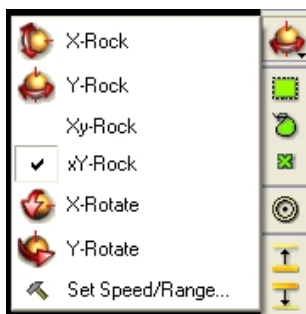


Smooth Animated Transitions

5.10.7.1 Make Animation

To quickly produce an ICM Molecular Animation:

- Click and hold down the "Begin rocking/rotation" button shown in the picture below.
- Choose from the following options - X-Rock, Y-Rock, Xy-Rock, xY-Rock, X-Rotate, and Y-Rotate.



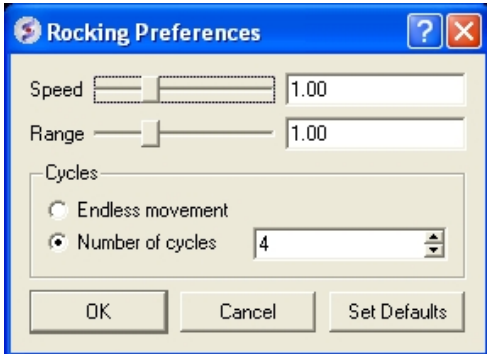
Click and HOLD DOWN the Begin rocking/rotation button

NOTE: Default rocking representation can be changed in the File/Preferences/Gui menu.

5.10.7.2 Change Speed, Range and Cycle Length of Animation

To change the speed, range and cycle length of the animation:

- Click and hold down the "Begin rocking/rotation" button shown in the picture above.
- Choose the set speed range option and change the speed and range using the drag bars. Any change will appear in the graphical display behind this box.
- If desired you can change the number of cycles of the animation. This is an ideal tool for screen-shot movie making.



NOTE: There is a return to default button in the Rocking Preferences dialog box shown above and default values can be changed in `File/Preferences/Gui`.

NOTE: Default rocking speed can be changed in the `File/Preferences/Gui` menu.

5.10.7.3 Interrupt Animation

An ICM Animation or Transition is fully interactive and is interrupted by a single click of the mouse.

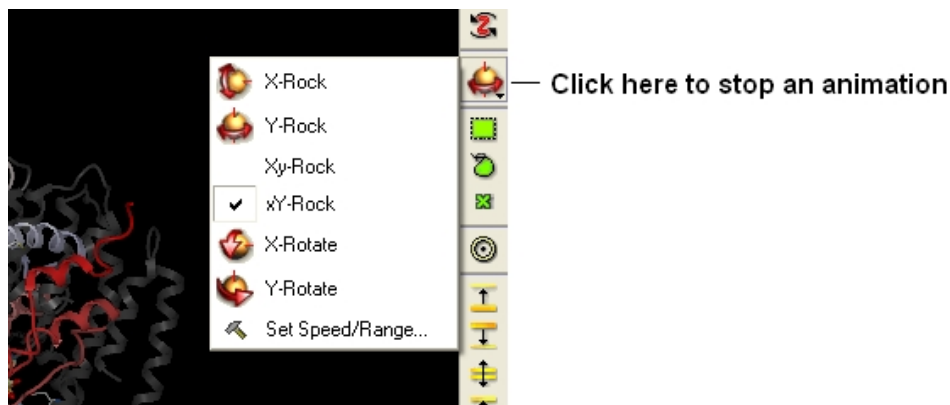
To stop or change an animation or transition:

- Click the "Begin rocking/rotation" button shown in the picture below.

To temporarily halt an animation or transition:

- Click in the graphical display. Once you release the mouse button the animation will start again.

NOTE: If you click on the graphical display during an animation the animation will be interrupted. Whilst clicking and holding the mouse button other operations can be performed such as zooming and selections.



Click in the graphical display window to temporarily interrupt an animation

5.10.7.4 Saving an Animation

An animation can be saved in an ICM project:

File/Save Project

Or

as a slide.

5.11 Graphics Shortcuts

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

The left mouse button can be mapped onto different graphics tools which can be selected from the right hand tool bar.

Note: (1) You can access many non-rotation modes directly from the rotation mode by using Middle and Right-mouse buttons, as well as by using the right, top and left margins of the graphics window. (2) You can access the rotation mode from non-rotation modes by pressing Ctrl.

- rotation (**the default** , press Ctrl if you in the non-rotation modes)
- translation (the middle mouse button in the rotation mode)
- zooming in and out by dragging the mouse up and down (the left margin in the rotation mode, or use the mouse **wheel**)
- Z-rotation (the top margin in the rotation mode)
- selecting by box (the right mouse click in the rotation mode)
- selecting by lasso (Ctrl-draw lasso in the rotation mode)
- picking out atoms (a toggle)
- picking out and labeling residues (a toggle)
- moving the front clipping plane (the top section of the right margin in the rotation mode)
- moving the rear clipping plane (the bottom section of the right margin in the rotation mode)
- moving the slab (the middle section of right margin in the rotation mode)
- unclipping (Ctrl-U)
- rotating torsions (Ctrl-left-mouse-click in the rotation mode)
- connect and unconnect separate molecules to movement controls

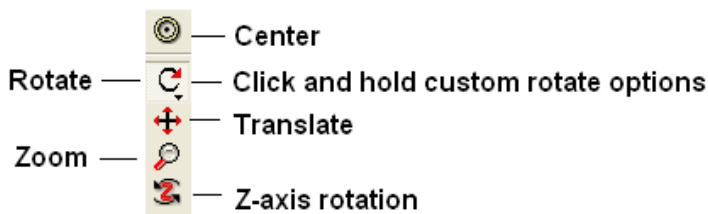
Many useful graphics tips are summarized [here](#).

NOTE: Key mouse controls are summarized in the command line manual [here](http://www.molsoft.com/man/graphics-controls.html)

5.12 Molecule Move Buttons

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

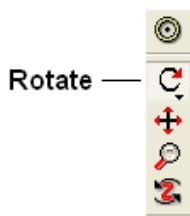
To move your structure it must first be displayed in the graphics window (for instructions on how to display a structure see the `Display` Tab). All of the following options are displayed in the Move Tools toolbar (shown below).



5.12.1 Rotation

In order to achieve the best pose for a picture or to enable the study of a certain region of your structure in more detail you may need to rotate the structure:

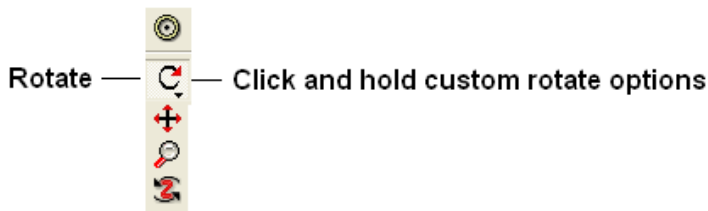
- Click on the **rotation** icon on the toolbar.



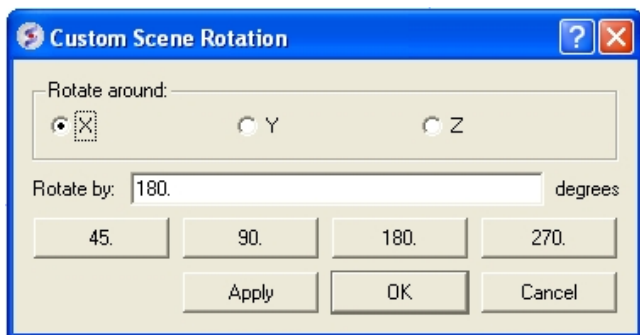
- Click and drag on your structure in the display window until it is in the desired position.

5.12.2 Custom Rotation

An option is provided to customize the rotation of the molecule. This allows exact rotation by a specified number of degrees.



- Click and hold down the rotation button and a data entry box as shown below will be displayed.
- Enter the number of degrees of rotation you require and in which X, Y or Z coordinate.



To continuously rotate the picture:

- Click on the **continuous rotation** icon on the toolbar.
- Click, hold, and slightly move your mouse anywhere on the graphical display window. The point at which you hold your mouse, is the direction to which the object will turn.
- Positioning the mouse towards the center of the display will move the object slower than if the mouse is positioned towards the edge of the graphical display.

In order to rotate your picture around the Z-axis:

- Click on the **Z-axis rotation** icon on the toolbar.

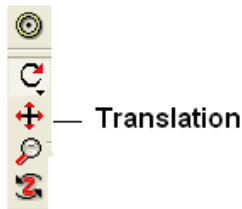


- Click and drag your object around the Z-axis until it is in the desired position.

5.12.3 Translation

To translate your structure up, down, left, or right:

- Click on the **translation** icon on the toolbar.



- Click and drag on your structure in the display window until it is in the desired position.

When you are displaying more than one object and you wish to translate one object in relation to the other on the Z-axis:

- Right click on the name of the object you wish to move in the ICM workspace and select connect to object. This object is now independent from the other object and can now be manipulated separately.
- Click on the **Z translate** icon on the toolbar.
- Click and drag your structure along the Z-axis, moving it closer or further from your unconnected structure.
- Once you are finished, right click on the name of the object which is connected, and click on disconnect.

5.12.4 Zoom

To zoom in or out of your structure:

- Click on the **zoom** icon on the toolbar.



- Click and drag your mouse up to zoom in and down to zoom out.

You can also zoom in and out directly with the right-mouse-button *without* explicitly switching to the zoom tool, if you use the **left 5%-margin** of the graphics window.

5.12.5 Center

To restore your picture to the center of the graphical display window or to center on a selection:

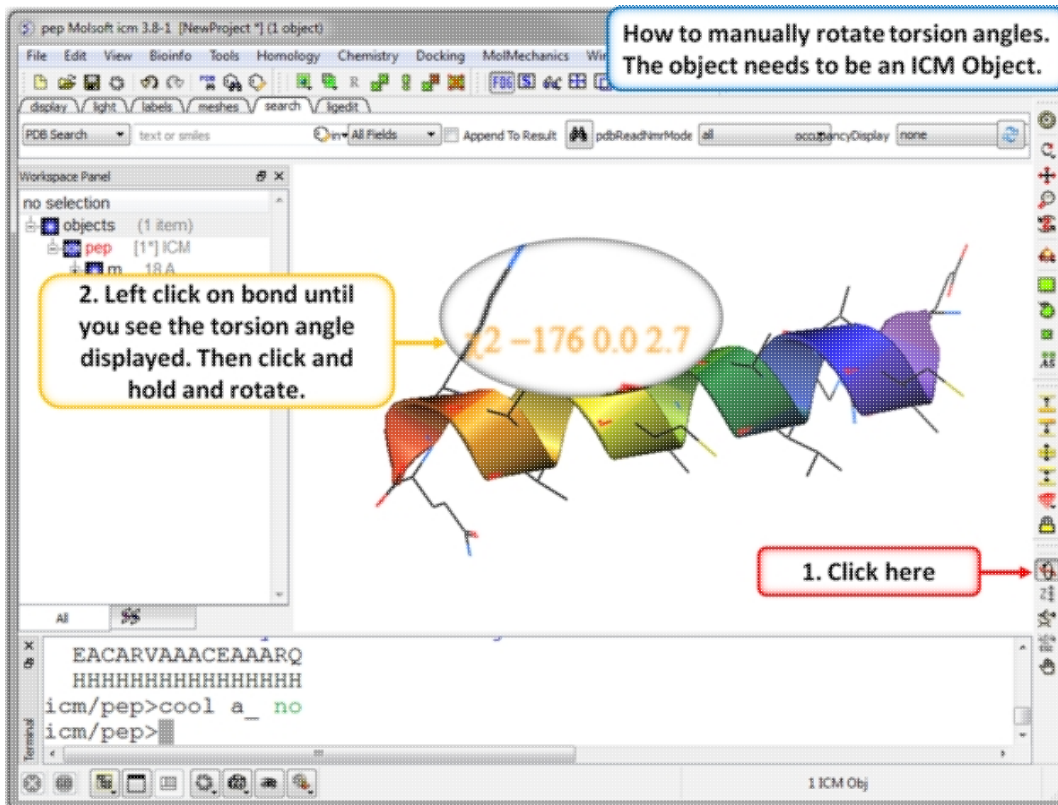
- Make a *selection* of the region you wish to zoom into - if no selection is made the whole structure will be centered.
- Click on the **center** icon on the toolbar.



5.12.6 Torsion Angles

To alter the torsion angle of certain residues of your structure manually:

- Convert your pdb structure into an ICM object.
- Click on the **change torsion angles** icon on the toolbar (see image below).
- Click on the bond until you see the torsion angle displayed in yellow (see image below). Click and hold on the bond around which you wish to rotate a residue. The changing torsion angle will be displayed in orange.



NOTE: This option can be used more effectively in conjunction with the `variable label` option.

To alter torsions by entering specific angle values:

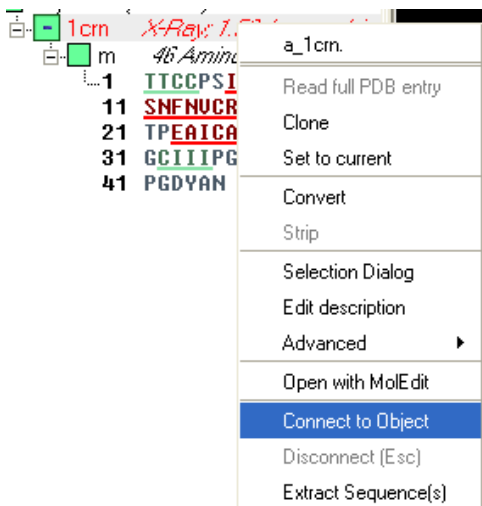
- You can edit torsions by specifically defining the exact angle as described [here](#).

5.12.7 Connect (Move)

When there is more than one object displayed in the graphical display window the objects are connected to one another. If you wish to move or manipulate one object independently from the others you need to **connect** to it

To do this from the ICM Workspace:

- Right click on the name of the object you wish to move in the ICM workspace and select **Connect to Object**. The object will now be colored yellow.



- The object is now controlled separately from the rest of your objects by your mouse.
- Disconnect your object by once again right clicking on the name of the object in the ICM Workspace and selecting disconnect in the drop down menu or **Press the ESCAPE key**.

Note: you can temporarily switch to the global rotation in the connected state if you press Shift

Note: use the Escape button to disconnect

5.13 Clipping Tools

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro



The clipping tools allow you to adjust the frames of the ICM window, changing the clipping planes.

Clipping planes can also be moved *without* switching to the clipping tool, if you click the right hand margin of the graphics window:

- The top section of the right 5% margin of the graphics window: moves the back clipping plane
- The middle section of the right 5% margin of the graphics window: moves the slab (both clipping planes)
- The bottom section of the right 5% margin of the graphics window: moves the front clipping plane

In order to move the front or rear clipping planes of your screen:

- Click on the **Move front clipping plane** or **Move rear clipping plane** icons on the toolbar.
- Click and drag the respective plane forward or backward, depending on how you wish to clip it.

You can also move the **slab** of viewing window, keeping the distance between the front and back clipping planes. In order to adjust the area of the structure where your viewing window is located:

- Click on the **Slab** icon on the toolbar.
- Click and drag the slab forward or backward, depending on the desired area of the structure you wish to see.

If you have made changes to the clipping planes which you do not wish to keep or you wish to automatically fit your entire structure within the clipping planes:

- Click on the **Unclip** icon on the toolbar. This will automatically set the clipping planes to fit your object.

5.13.1 Mesh Clipping

Clipping tools can be used to adjust the frames of the mesh independently of other objects. This is described [here](#).

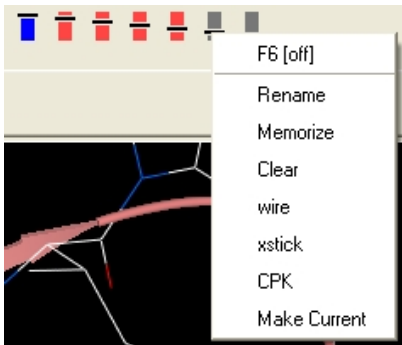
5.14 Graphic Layers

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

To display and undisplay layers of a structure you can use the buttons shown below. Seven layers can be created and within each layer different structural representations can be displayed.



Right click on one of the layer buttons and a number of options can be chosen as shown below.



To change the display in one of the layers:

- Right click on one of the layer buttons.
- Select a representation wire, xstick or CPK.
- You can do this for each of the seven layer buttons.
- Click on the layer button to display and undisplay. If the layer button is shaded red then the layer is not displayed. If the layer button is shaded light blue then it is displayed. You can switch between layers by clicking on the button or using the **F6** key. You can use the **memorize** button to store a particular representation and **clear** to remove a memorized representation.

5.15 High Quality Publication Images

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

5.15.1 High Quality Image

To make high quality publication images:

The first step is to improve the quality of the image using the **High Quality Image** and **Antialias** buttons shown below.

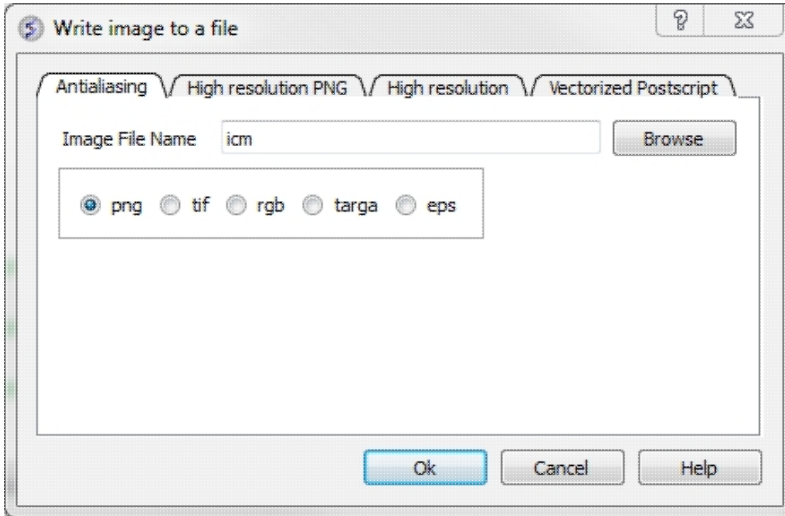
antialias



high quality image

To save and write an image:

- Select **File/Write Image** and the following window will be displayed:



Choose one of the tabs depending on the image type you require:

- **Antialiasing** allows you to make png, tif, targa, eps or rgb file types.
- **High Resolution PNG** Allows you to specify a resolution increase and transparent background.
- **High Resolution** Allows you to specify a resolution increase for tif, targa and rgb files.
- **Vectorized Postscript** This creates a vectorized postscript model of the screen image. Instead of the bitmap snapshot this command generates lines, solid triangles and text strings corresponding to the displayed objects. Since the postscript language is directly interpreted by high-end printers, the printed image may be even higher quality than the displayed image. The final resolution is limited only by the printer since the original image is not pixelized. Warning: there may be inevitable side-effects for some types of solid images at the intersection lines of solid surfaces (i.e. large scale cpk representation, hint: use display skin instead).

****NOTE:** With some graphics drivers the high resolution image may be distorted or multiple representations of the image in one frame. Please see the [FAQs](#) for a solution to this.

5.15.2 Quick Image

A quick image can be saved using this option. The image will be saved as `icm1.png` in the current directory in which you are working. Each subsequent image produced will be incrementally numbered.

This option is also available via a button as shown below:

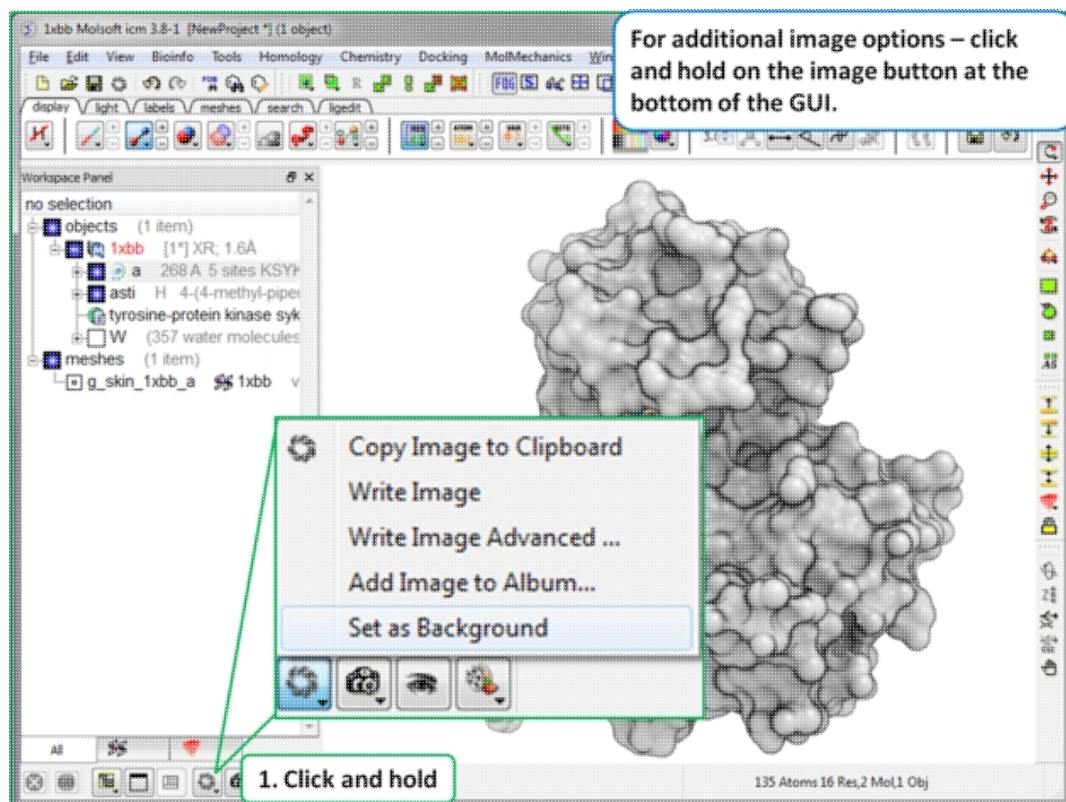


Quick Image Button

5.15.3 Image Options

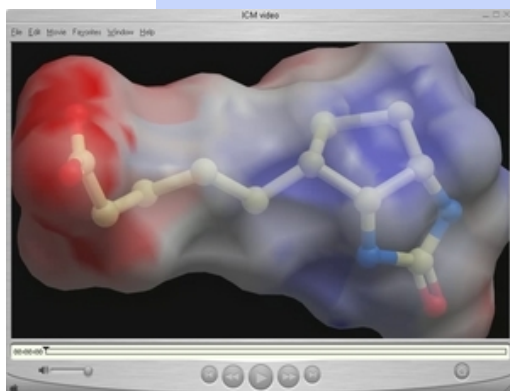
Click and hold on the **image** button at the bottom of the GUI to:

- Copy image to clipboard (single click on the button also does this).
- Resize image (Write Image Advanced)
- Transparent Background (Write Image Advanced)
- Store an image in ICM (Write Image Advanced)
- If you are making an ICM document you may want to store images inside ICM (Add Image to Album).



5.16 Movie Making

Available in the following product(s): ICM-Browser-Pro | ICM-Pro



A movie is an excellent way of communicating 3D structural data. The resulting movie can easily be transferred into other applications such as Microsoft Powerpoint.

There are two ways to make a movie:

- Make a movie directly from a series of slides.

- Make a movie directly from screenshots.

An alternative to a movie is a fully interactive 3D document embedded in Windows PowerPoint or a Web Browser, this can be undertaken using slides and the ActiveICM plugin.

5.16.1 Making a Movie from a Set of Slides

To make a movie from a set of slides:

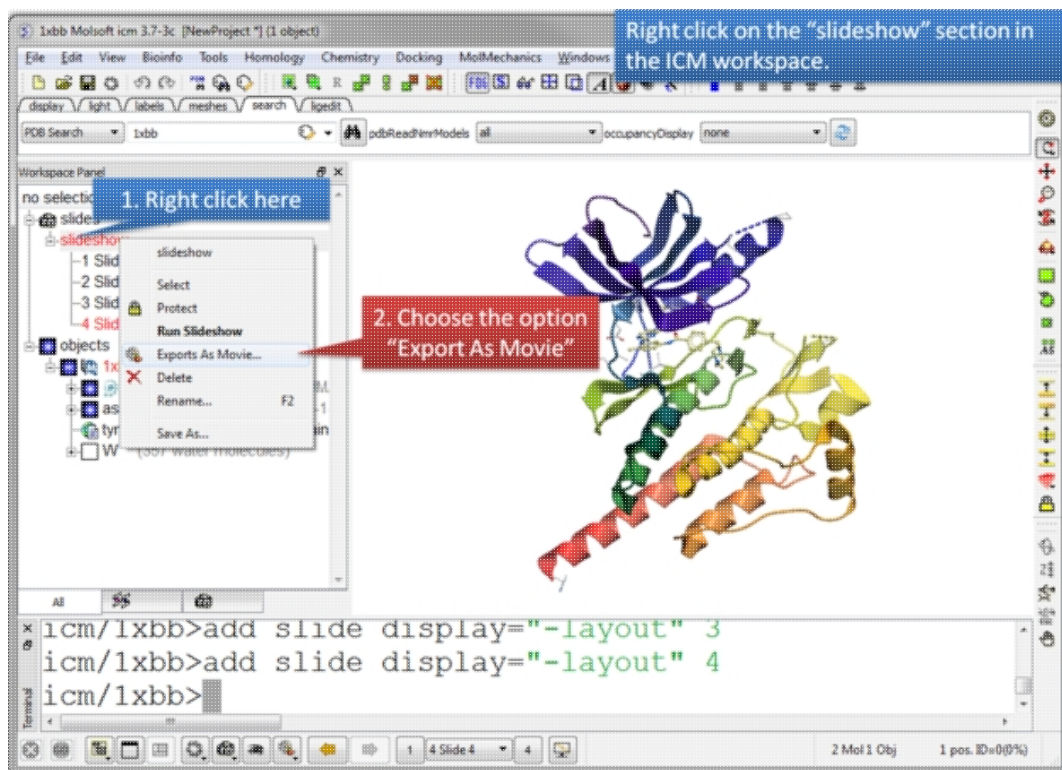
Make a series of slides using the camera button at the bottom of the screen.

2. Slides are listed here

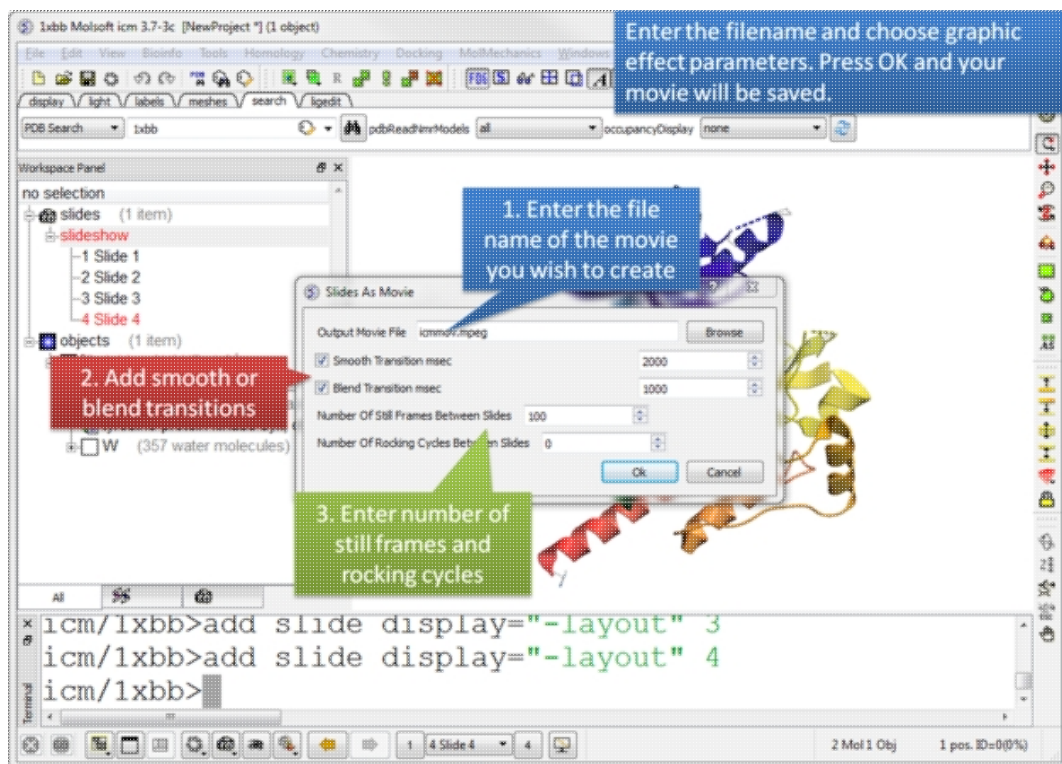
1. Make Slides

```
icm/lxbb>add slide display="-layout" 3
icm/lxbb>add slide display="-layout" 4
icm/lxbb>
```

Make a series of slides using the camera button.



Right click on the slideshow in the ICM Workspace and choose the option "Export as Movie".



Enter the filename for your movie and set the movie effect parameters. Press OK and the movie will be written to a file.

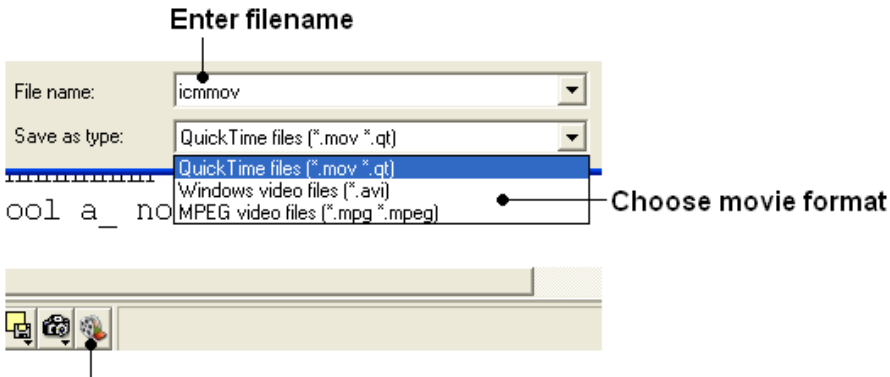
5.16.2 Screenshot Movie



— Screenshot movie making

To make a **Screen-grabbing Movie** follow these steps:

- Resize the graphical display to the screen size/resolution you need. You may also want to select the high quality image button and antialiasing to improve the quality of the movie or add visual effects such as shadows.
- To begin making a movie click on the movie making button at the bottom of the graphical user interface (as shown below).
- Enter a file name for your movie and select the movie format (.mov, .avi, mpeg).



— Screenshot movie making button

NOTE: If you want to make a movie to include in a PowerPoint presentation you need to save the movie in AVI format.

- To begin recording the screenshot movie click on the red **Record video** button. Anything displayed in the graphical display will be recorded, for example you can record animations and transitions. Specifying the number of cycles in the animation (rocking, rotation) is an ideal tool for screen-shot movie making. If you have a fast computer you can use **Realtime screen grabbing** which can be selected by clicking and holding the **Record video** button. The real time option can also be set in File/Preferences/Gui menu.



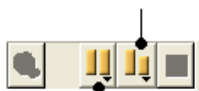
— Record video button

- The length of the movie in minutes, seconds and milliseconds is displayed in the top right hand corner of the graphical display.



- You can pause the movie and fade out by clicking on the button shown below. The number of frames for the fading out option can be controlled using the option in File/Preferences/Gui

Pause and fade out



Pause recording

- You can record a smooth transition from a previous frame by clicking on the button shown below.

Smooth transition from previous frame



NOTE: Anything you do in the graphical display will be recorded in the movie. For example you can change representations, lighting, add new molecules etc. This can be achieved in a more controlled manner using the **pause** and **record smooth transition** button.

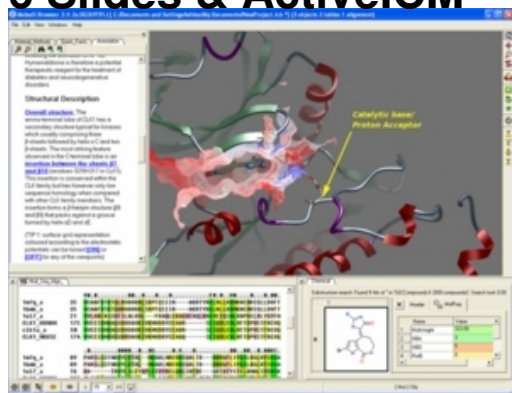
- Once you have paused the recording the viewpoint and representation of the molecules can be changed and a smooth transition from the previous frame can be generated by selecting the **Record smooth transition from previous frame** button.
- To stop recording a video press the button shown below.

Record smooth transition from previous frame



Stop recording

6 Slides & ActiveICM



www.molsoft.com/activeicm.html

In this chapter you will find a description of the tools available to create files (.icb) containing fully-interactive three-dimensional (3D) molecules and two-dimensional (2D) data. These files can contain multiple interactive views and animations of molecular structures and objects in conjunction with related hyperlinked text, chemical, biological sequence, alignment and data views. The files are small and easily transferable and downloadable. The files can be used for Molecular Presentation and Documents inside the ICM browser or displayed on the web and in PowerPoint using the ActiveICM plugin. For examples of ICM Molecular Documents please see MolSoft's ActiveICM product page at

Chapter Contents:

- Making Molecular Slides
- How to View and Navigate Slides
- How to Edit Slides
- How to Add Smooth Blending and Transition Effects Between Slides
- How to Make Molecular Documents - Link HTML Text to Slides

6.1 Making Slides

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

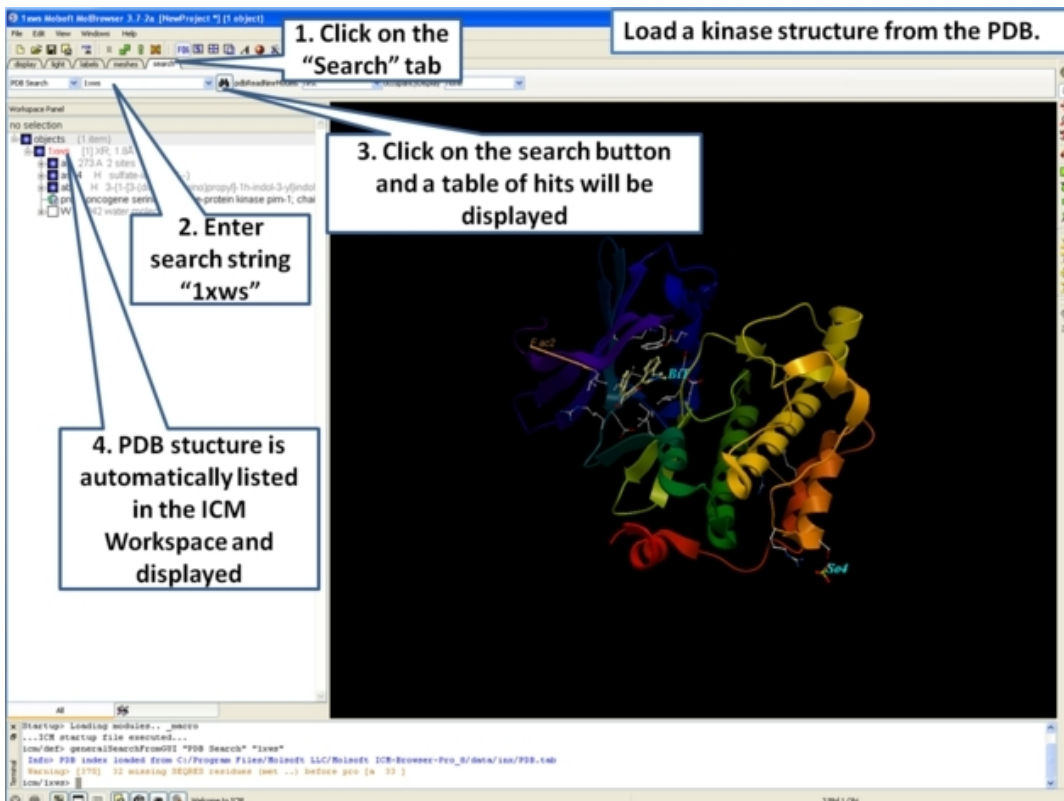
A slide enables you to store a large number of different 3D visualization along with text and window layout. The following information can be stored in a slide:

- Viewpoint (e.g. rotation, translation, zoom, lighting and depth effects)
- A set of atom-specific graphical representations such as surfaces, which can be represented as smooth, transparent or wireframe, ball-and-stick models, CPK (space-filling) models.
- A set of atom, residue or distance labels on any of the atomic items
- A set of arbitrary 3D textual annotations assigned to a point in space
- A set of arbitrary 2D annotations assigned to specific 2D coordinates on a screen
- Parameters of the parametric animation
- Window layout - for when the slides are viewed in the browser.
- Current table(s)
- Sequences and Alignments
- HTML text with hyperlinks
- 2D images
- For each grob (mesh): representation and colors.

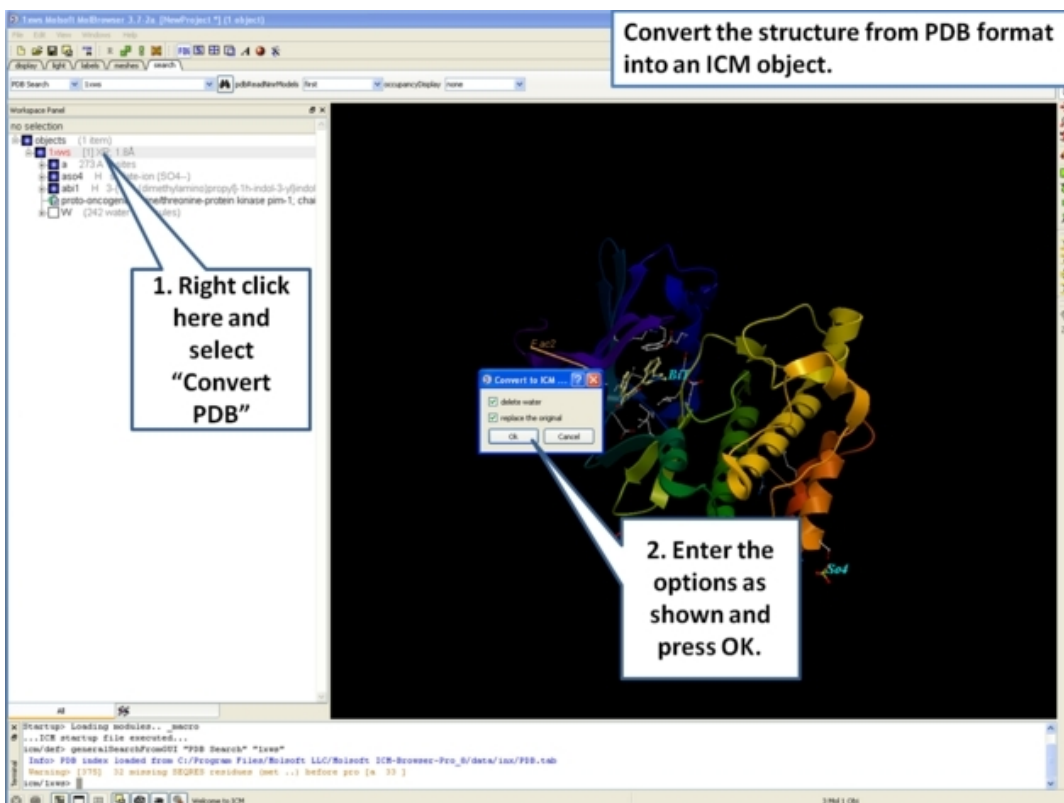
This tutorial takes you through the steps to create a series of fully interactive 3D slides. The slides can then be embedded into the web, or PowerPoint using ActiveICM or viewed in ICM-Browser (or ICM-Pro).

To begin making ICM Molecular Slides:

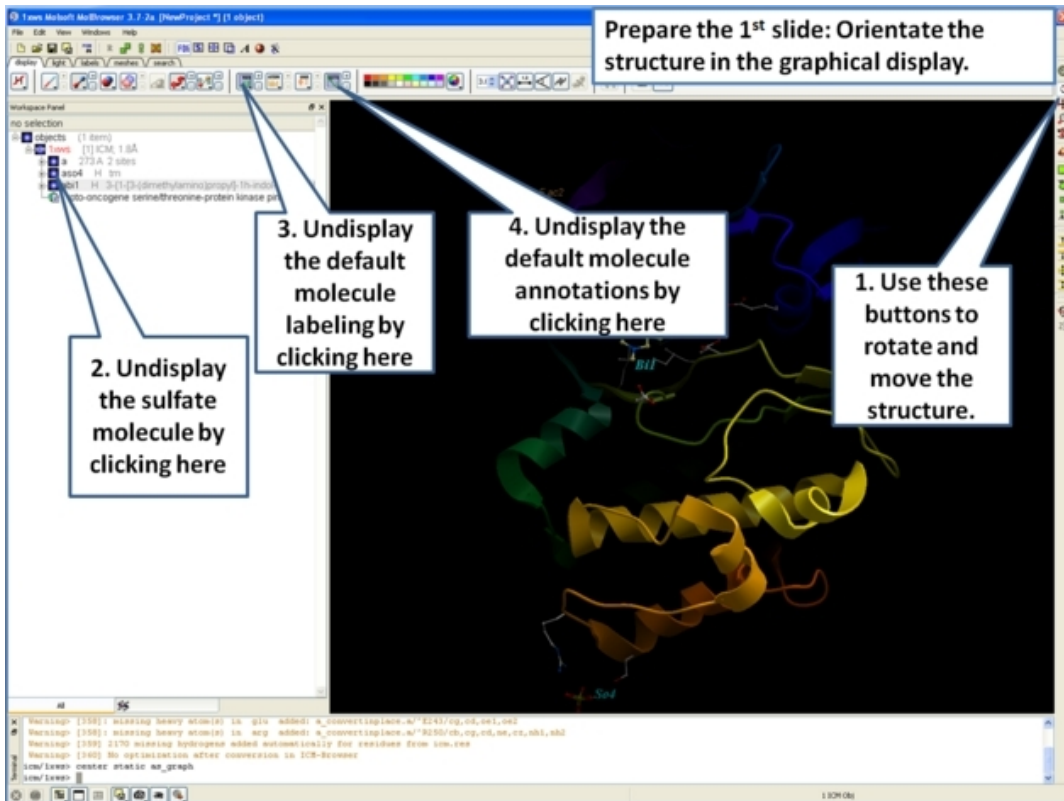
- First load the structure or structures you wish to display in your first slide. Additional structures, labels etc and text can be added at any point during the slide making process. In this example we will load the PDB file 1XWS a PIM1 kinase.



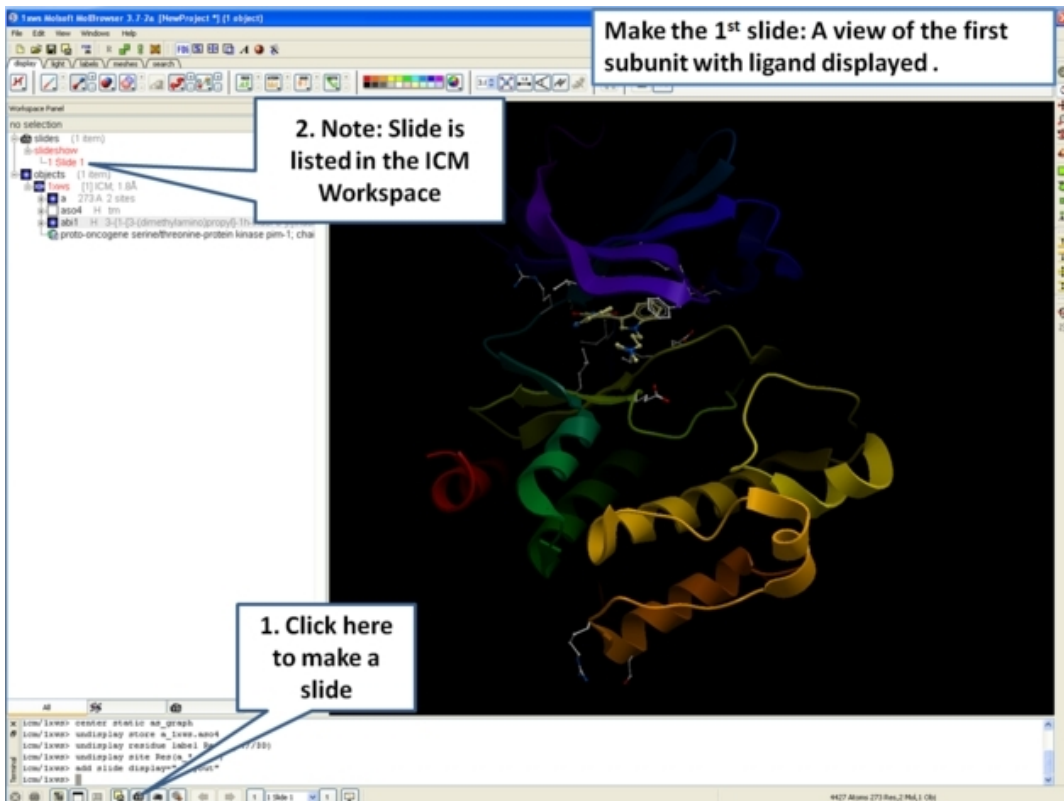
- Next, we will convert the PDB file to an ICM object so we can make slides of the ligand-receptor hydrogen bonds and binding pocket surface.



- Now we are going to prepare the first slide by rotating the protein structure to an orientation which allows the viewer to see the key features of the kinase. For example the bulge in the hinge region (between the N- and C- lobes) which is unique to PIM proteins.



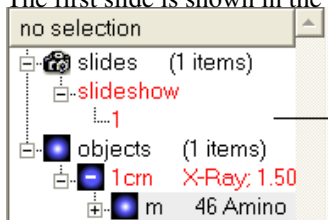
- Next, make the first slide by clicking on the camera button at the bottom of the graphical user interface.





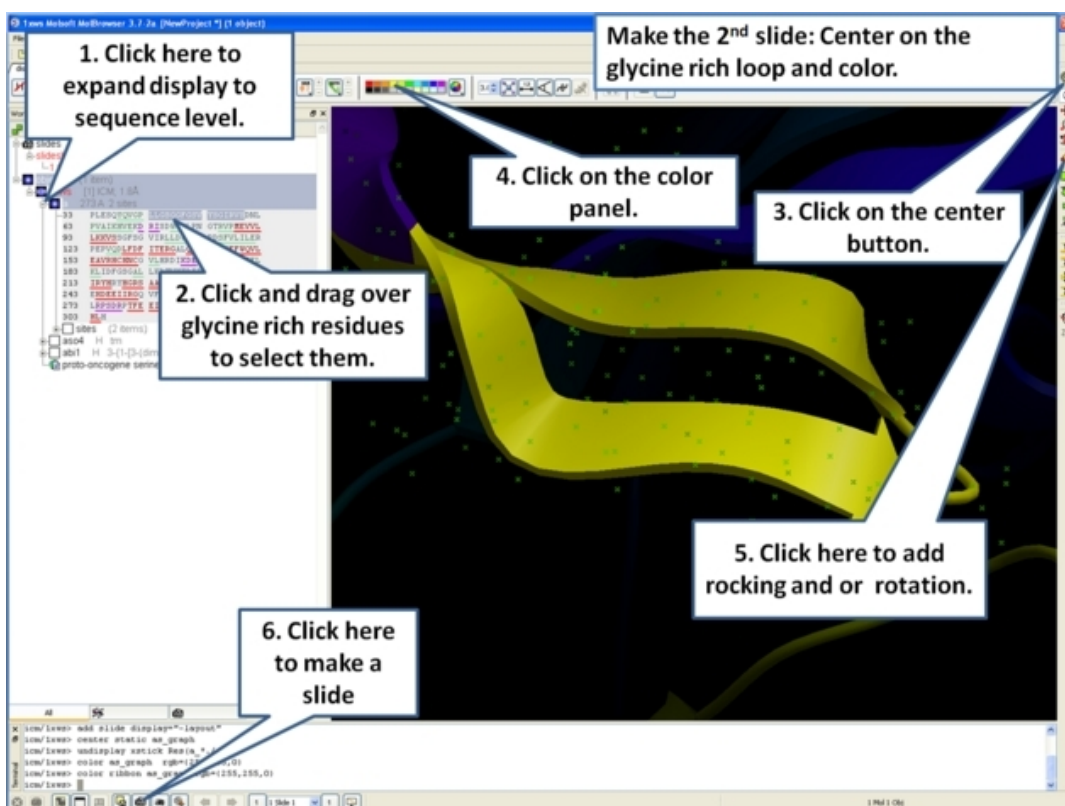
Click to add slide

- Once you have clicked on the camera button you will see that the first slide has been generated. The first slide is shown in the ICM Workspace window as shown below.

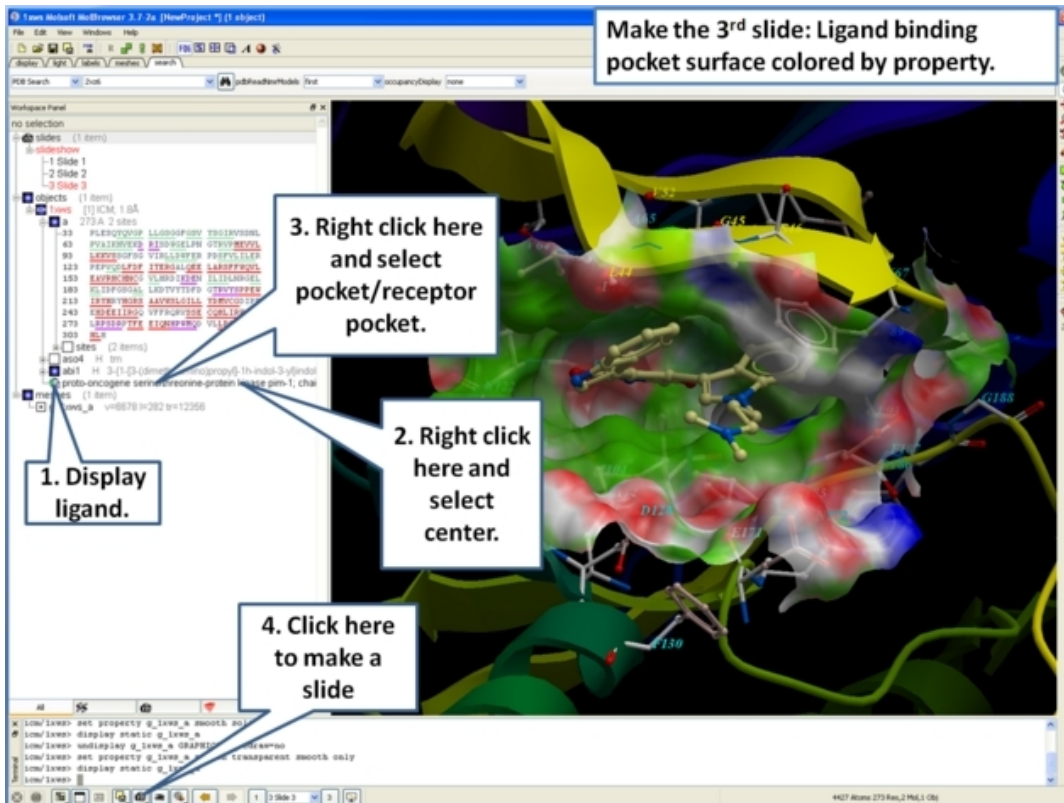


The number and name of the first slide is displayed in the ICM Workspace

- Slides can consist of Static views or Transitions and Animations. Here we will zoom into the flexible glycine rich region of the kinase which lays across the roof of the ATP-binding pocket. Click on the camera button and make the second slide



- Next, we will make a slide of the surface of the ligand binding pocket colored by binding property.



- Now save the document as an icb file. Go to File/Save as...

6.2 Make a Movie from a Set of Slides

Available in the following product(s): ICM-Browser-Pro | ICM-Pro

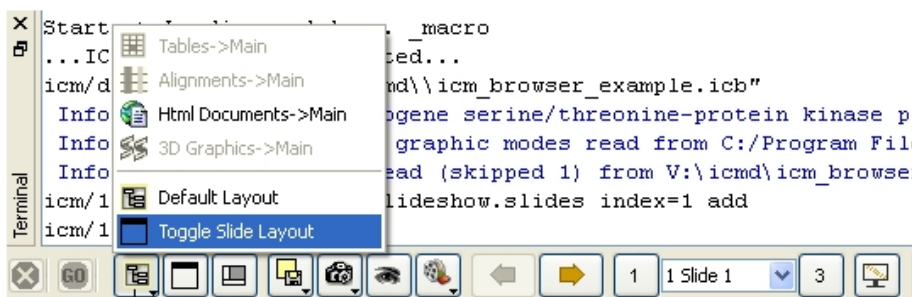
Please see the movie chapter on how to make a movie from a set of slides.

6.3 How to View and Navigate Slides

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

6.3.1 View Slide Show

To view a slide show select the buttons shown below:



Click and hold and select Toggle Slide Layout

Run Slideshow

NOTE: Slides are associated with the objects currently loaded into ICM. Therefore if you delete an object then the slides will not work. However if you delete an object and then re-read the same object with the same name and structure the slides will be ok.

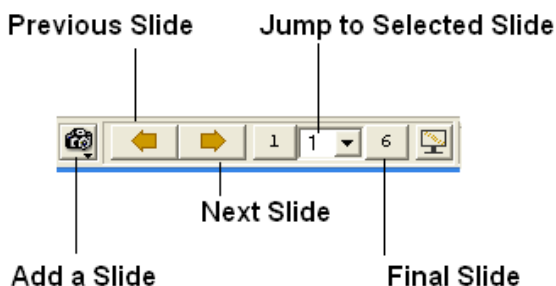
To save a slide show

- File/Save Project

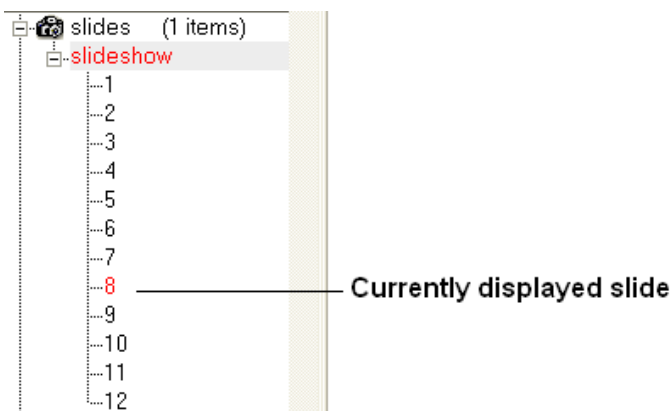
6.3.2 Slide Navigation

You can make as many slides as you wish as described in the *Making Molecular Slides* section.

To navigate through the slides you can use the buttons shown below, the cursor keys for some operations or the right click options in the ICM Workspace.

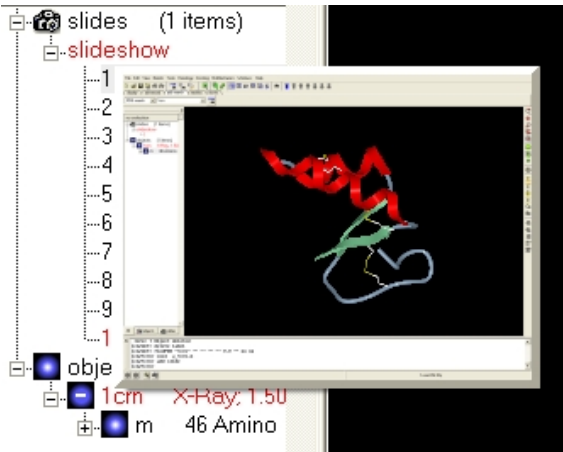


The slide currently displayed is highlighted in red in the ICM Workspace.



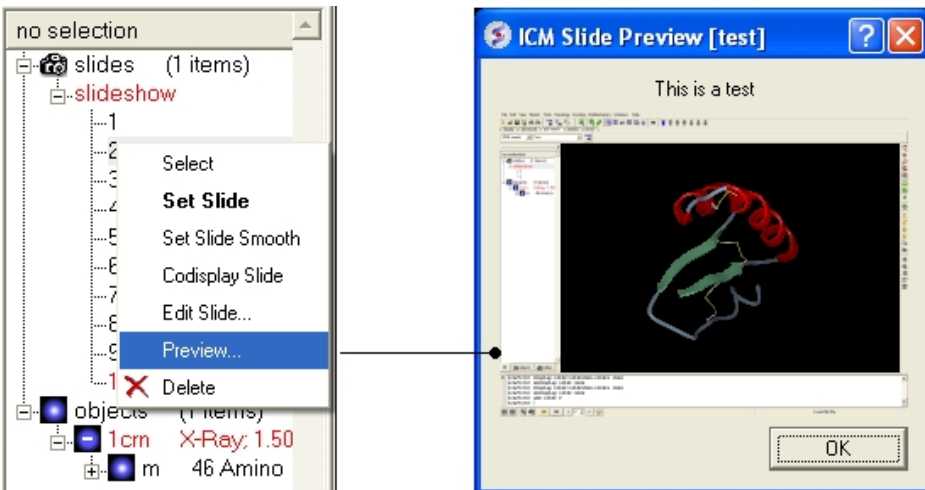
To jump to another slide right click and select "Set Slide".

All slides are displayed in the ICM Workspace. You can hover the mouse over a slide name in the ICM Workspace and a thumbnail sketch of the slide is displayed as shown below. This can be used for slide navigation purposes.



Hover mouse over slide name in the ICM Workspace and a thumbnail sketch of that slide will be displayed.

Or you can right click on the name of the slide in the ICM Workspace and select the option "Preview".



6.4 How to Edit Slides

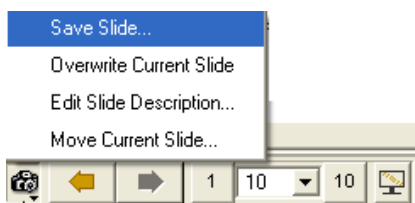
Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

You can jump to the slide you wish to edit by following the slide navigation instructions.

6.4.1 Edit Slide

Edit slide contents: To edit the content of a slide the procedure is to add a new slide and then delete the old one or use the "overwrite current slide" option as shown below:

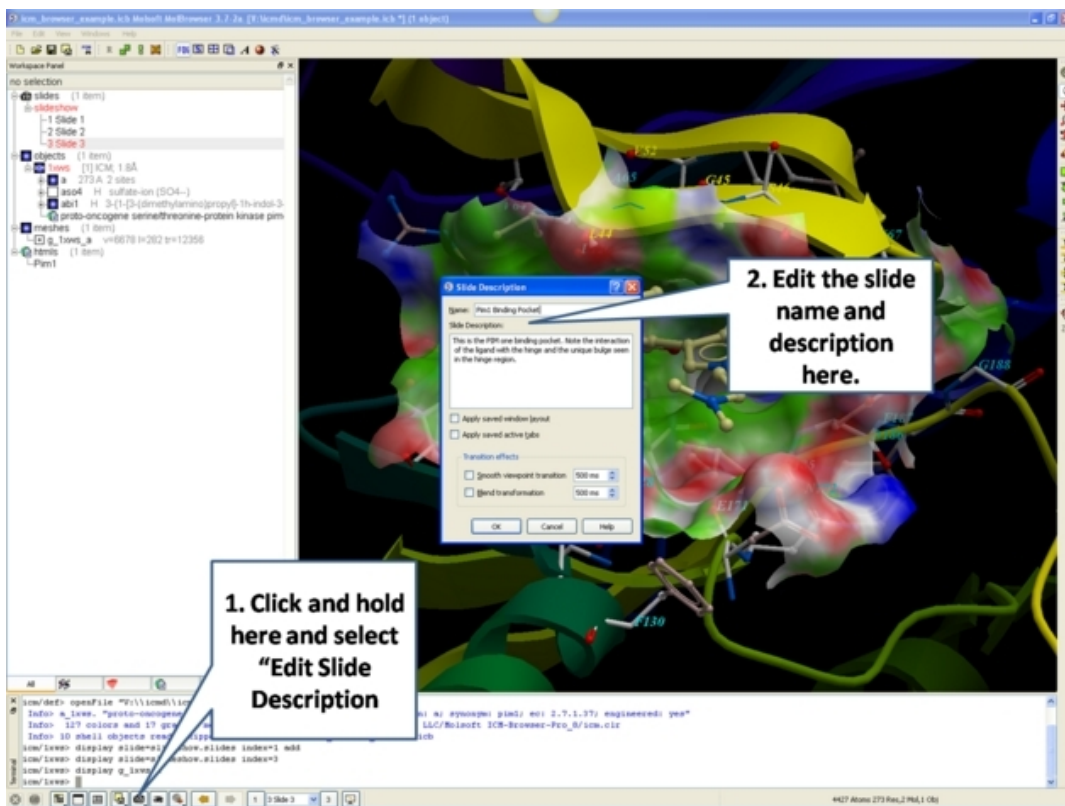
- Click and hold down on the camera button.



Click and hold

To edit a slide description.

- Click and hold down on the camera button and select the option "Edit Slide Description".



- Enter the name of the slide
- Enter a description of the slide.
- If you wish to keep the current window layout or active tabs check the boxes provided

To delete a slide:

- Right click on the name of the slide in the ICM Workspace and select Delete.

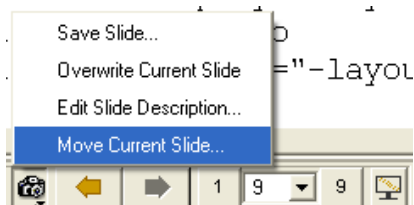
To change the name of a slide

- Right click on the name of the slide in the ICM Workspace and select Edit Slide.

6.4.2 Move Slide

To change the slide's position in the slideshow use the Move Current Slide option and select the new position from the list.

- Click and hold on the "make slide button".
- Select Move Current Slide.



Click and hold

- Select the position in the slide show where you want to move the slide to.



Co-display more than one slide

- Right click on the name of the slide in the ICM Workspace you wish to co-display with the currently displayed slide.
- Select the option co-display slide.

6.5 How to Add Smooth Blending and Transition Effects Between Slides

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

How to add smooth and blend transitions to a slide.

- Right click on the name of the slide in the ICM Workspace.
- Select **Edit Slide**.
- Select the desired transition effect - **smooth** or **blend** as shown below.
- Select the length of the transition in milliseconds.

Add smooth transitions and transformations between slides.

1. Right click here and choose Edit Slide

2. Check the Smooth and/or Blend option

3. Enter the transformation time in milliseconds

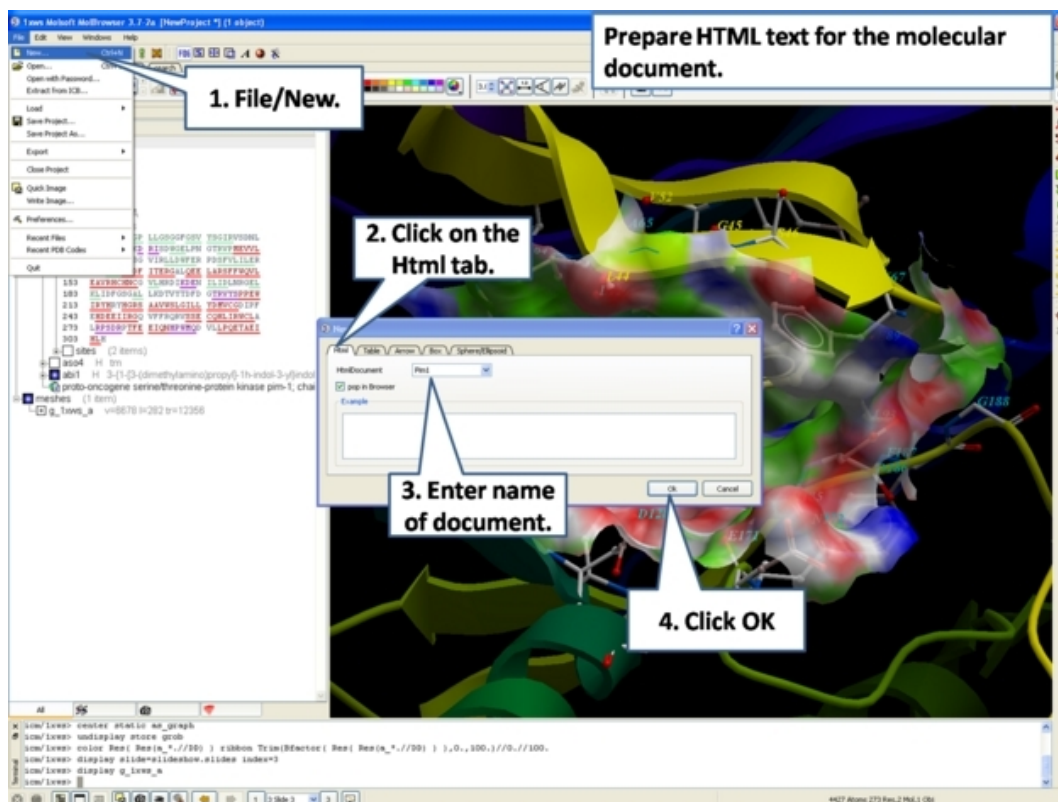
6.6 How to Make Molecular Documents - Link HTML Text to Slides

Available in the following product(s): ICM-Browser | ICM-Browser-Pro | ICM-Pro

An ICM Molecular Document contains text and images which can be hyperlinked to the graphical display. Click on the hyperlinked text and then a fully-interactive 3D slide will be displayed. The hyperlinks are usually linked to a set of slides but can also be linked to a series of commands in a script, a web page, a table or alignment. Once a molecular document has been made you can view it in the ICM-Browser (File/Save Project .icb file) or download ActiveICM and view it in a web page or Powerpoint.

To begin creating an ICM document

- File/New/ and click on the HTML tab.
- Enter some text. E.g the Name of the HTML document. Formatting can be changed as described in the edit section below.
- Click OK



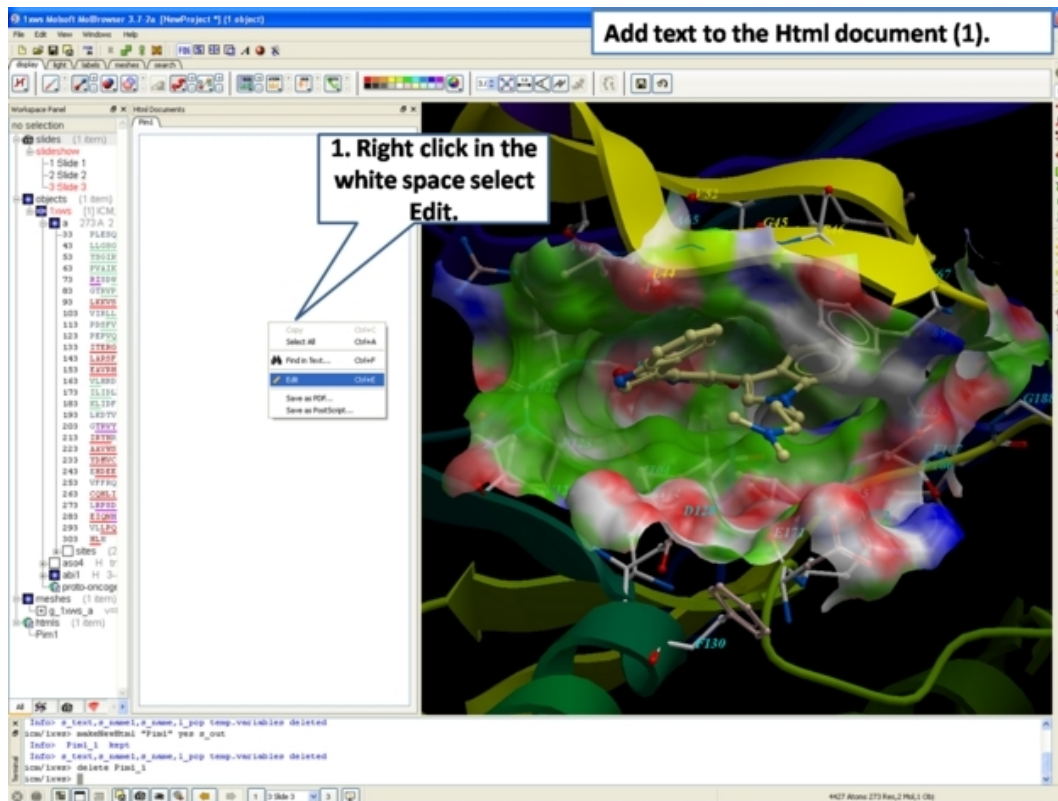
- A HTML text panel will be displayed in the graphical user interface.

NOTE: You can add multiple documents into a single file. The documents will be accessible via tabs at the top of the HTML panel.

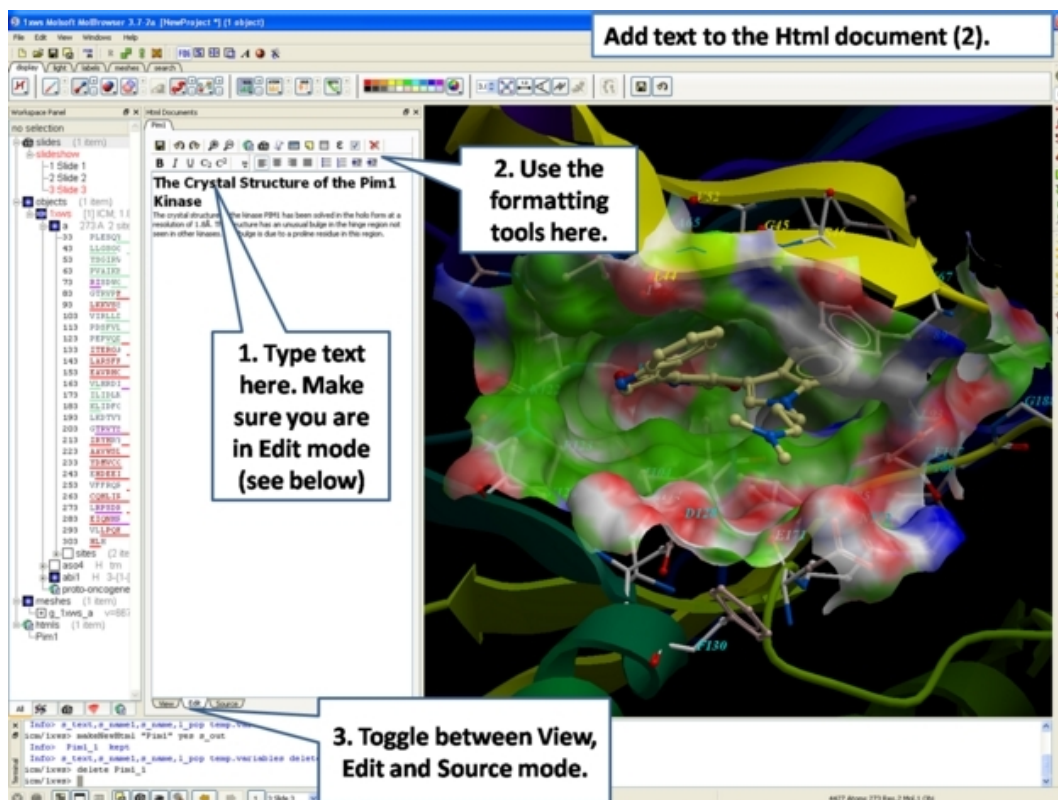
6.6.1 How to Add Text or Edit a Molecular Document

To edit the HTML text in the graphical display

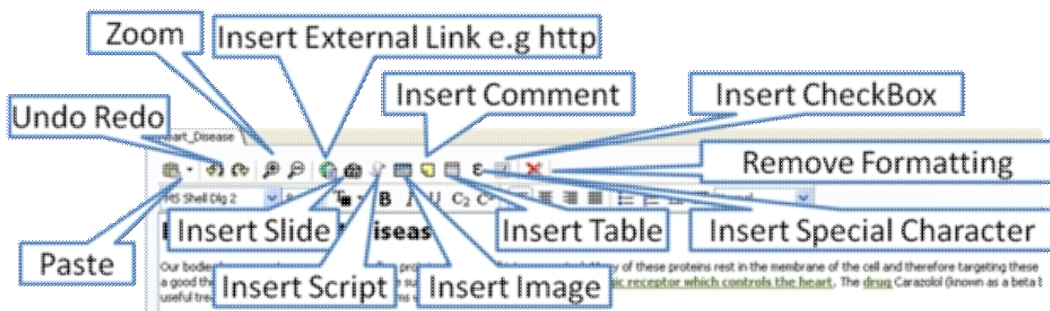
- First create an HTML document and the text panel will be displayed in the graphical user interface.
- Right click in the body of the text display panel and select **Edit**.



- Enter text and use the formatting tools provided in the panel above the text editor. Make sure you have selected the **Edit** tab in the HTML editor. You can see your page in the **View** tab or write directly in HTML in the **Source** tab.



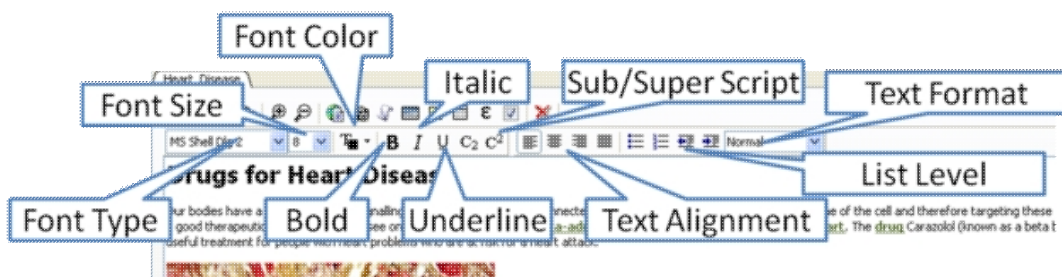
The key formatting tools in the HTML editor are shown below.



Check appearance using the **View Tab**



Write in HTML using the **Source Tab**

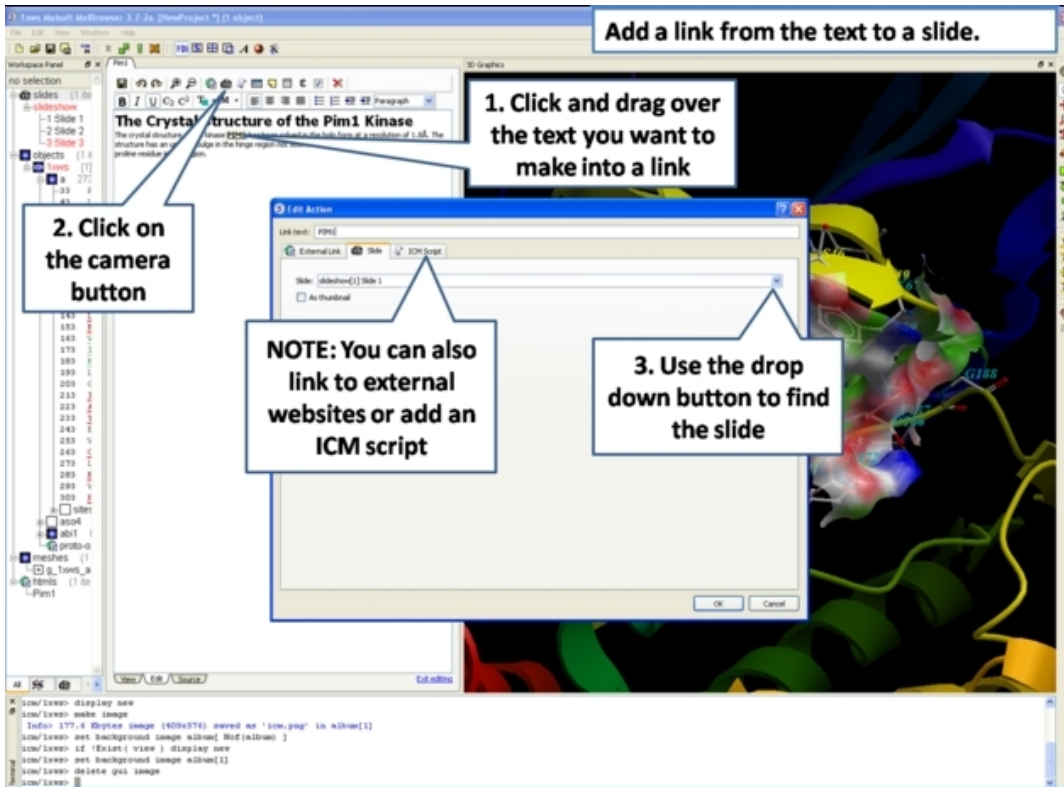


6.6.2 How to Make a Hyperlink Between Text and a Slide

To make a hyperlink between the text and the graphical display (slide)

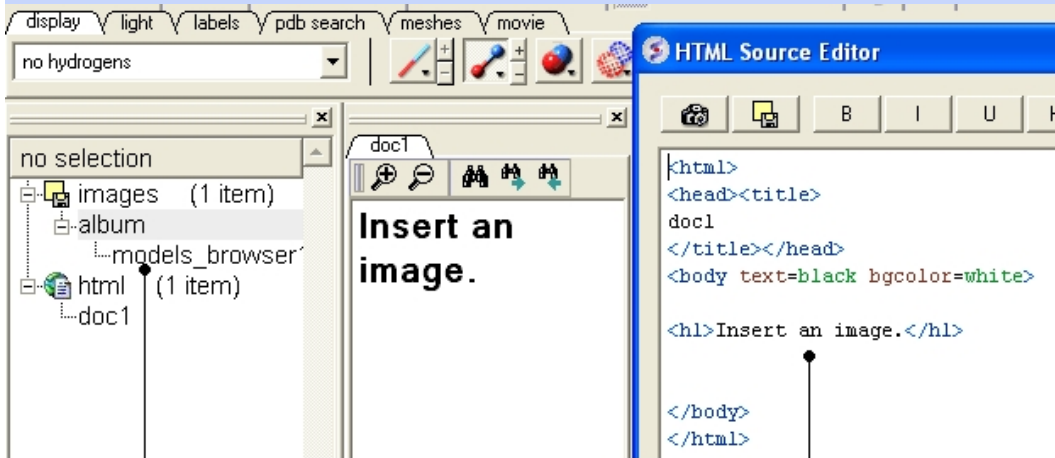
Make a slide or set of slides of the graphical display you wish to link to. See Making Molecular Slides for help on this. Once slides have been created:

- File/New/Html
- Right click in the body of the text display panel.
- Select Edit.
- Highlight the text you wish to link to a graphical display - you can do this by left clicking and dragging over the text (selected text will be highlighted in blue).
- Click on the "Camera button" in the HTML editor formatting tool panel.
- Select the **Slide** tab.
- Select which number slide you wish the text to be linked to from the drop down menu.
- There is an option to display the slide as a thumbnail image in the text document panel. Check if appropriate.



6.6.3 Insert Image

NOTE: The easiest way to add images (PNG or JPEG) into an ICM Document is to use drag and drop. You can drag and drop the image into the ICM Workspace or go to File/Open. Once the image is in the album in the ICM Workspace you can then drag it from the ICM workspace into the HTML editor.

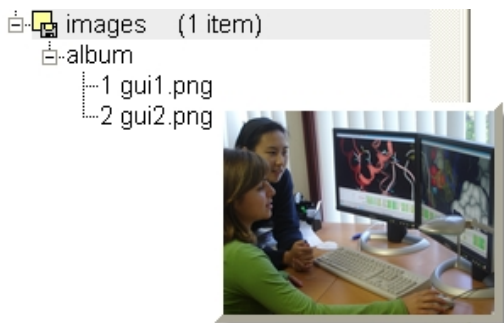


Drag and Drop from the ICM Workspace to the HTML source editor

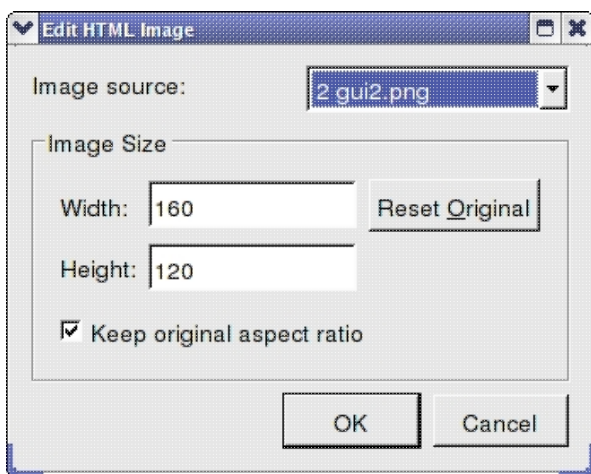
Another way to insert a picture into the HTML text panel

- First read the image into the ICM photo album File/Open OR Drag and Drop from directory into the ICM Workspace.

The image name and preview will then be displayed in the ICM Workspace.



- Create HTML text - File/New/HTML. Add text.
- Right click in the HTML window and select 'Edit Source'.
- Right click on the position in the ICM Script Editor where you would like to insert the image.
- Select 'Insert Image'



- Select the image name source.
- Choose the desired Width and Height.
- Click OK.
- Click Save in the ICM Script Editor.

6.6.4 Insert Script

How to insert a script to the text panel

There are 3 ways to add a script - described in more detail below

1. Drag and drop script from ICM Workspace
2. In the HTML Source Editor - right click and select Insert Slide or Action
3. Create an "inline" script

These methods are described below:

Drag and Drop Method

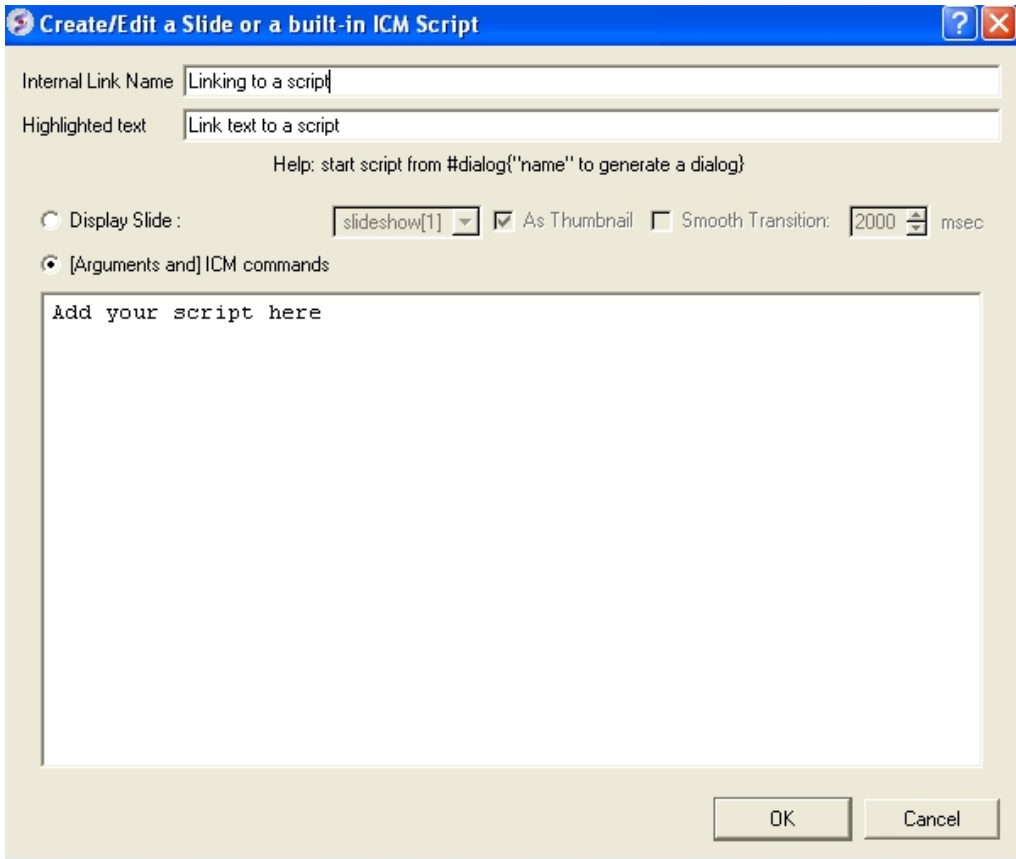
- Create a script File/New/Script
- The script will be displayed in the ICM Workspace.
- Right click in the HTML Text Panel (for instructions on how to create this panel see create molecular document) and select edit source and the HTML Source Editor will be displayed.
- Click-Drag and Drop the script into the HTML Source Editor

A line as shown below will be added.

```
<a href="#icm/script/script1">text placed here will be displayed as a link in the document</a>
```

Another way to add a script to the document is to Insert Action:

- Right click in the body of the text display panel.
- Select Edit Source
- Highlight the text you wish to link to a graphical display - you can do this by left clicking and dragging over the text (selected text will be highlighted in blue).
- Right click and select 'Insert Slide or Action' or select the button in the HTML Source Editor and a window as shown below will be displayed.



- Select the option [Arguments and] ICM commands
- Add script in the editor provided
- Select ok

Inline Script

A script can be added to the HTML text in the following way

- Right click in the body of the text display panel.
- Select Edit Source
- Enter script in the format as shown below.

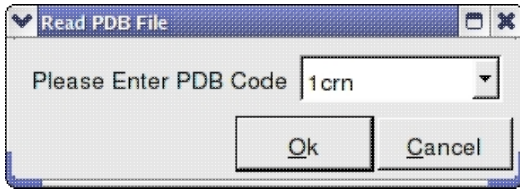
```
<!--icmscript name="script2"  
#dialog{"Test"}  
# i_number1 (2)  
# i_number2 (3)  
print $1 + $2  
--><a name="script2" href="#_">script2</a>
```

6.6.5 Insert a Dialog Box

Dialog boxes are provided to enable a viewer to interact with a presentation or document file. The dialog box will be a gui data entry box. For an example here is a script to prompt the user of the file to enter a pdb

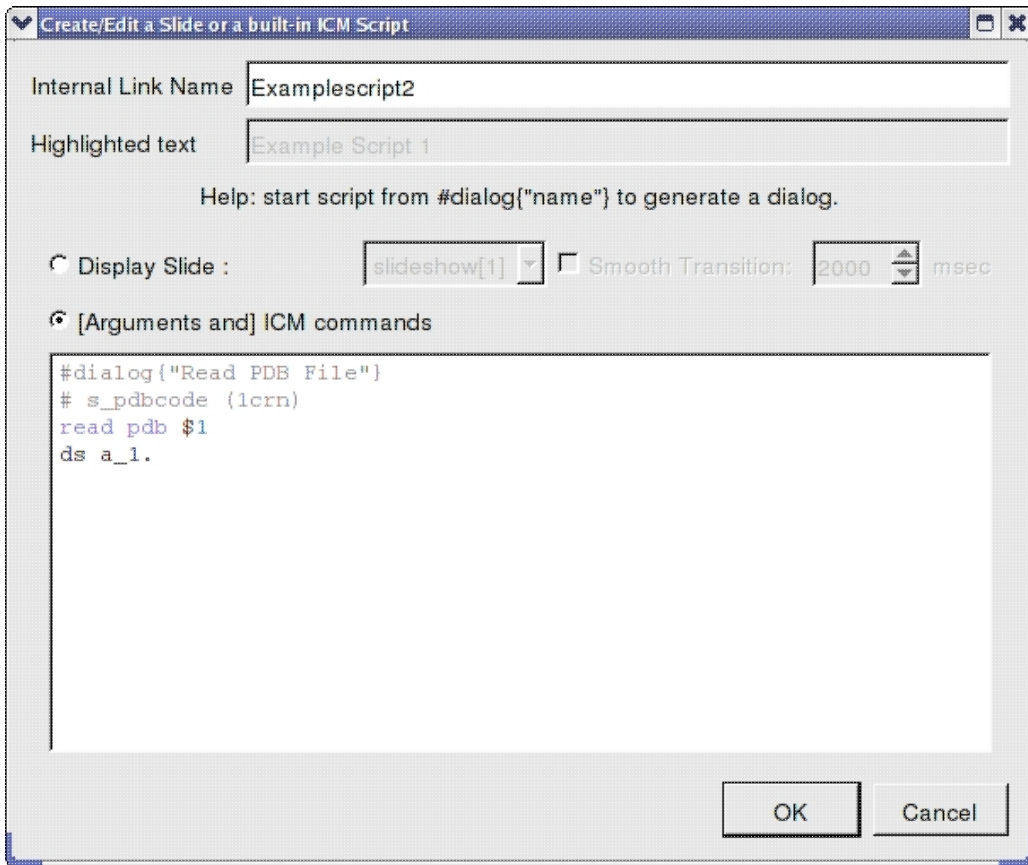
code:

```
#dialog{"Read PDB File"}  
# s_pdbcode {1crn}  
read pdb $1  
ds a_1.
```



The code above can be saved as a script or inside the html text. To do this:

1. Right click on the HTML text display and select "Edit Source".
2. Highlight the text you wish to link to a dialog box and then select the right click and select 'Insert Slide or Action' or select the button in the HTML Source Editor and a window as shown below will be displayed.



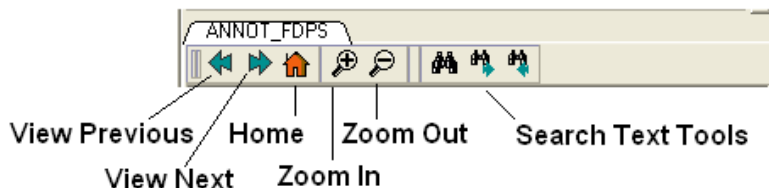
OR.

1. Right click on the HTML text display and select "Edit Source".
2. Add a link to a script as shown below.

```
<a href="#icm/script/script1">Example Script 2</a>
```

6.6.6 Document Navigation

The following buttons shown below aid document navigation. Also remember that more than one document can be stored and the header of each document file will be displayed in multiple tabs in the text panel window.

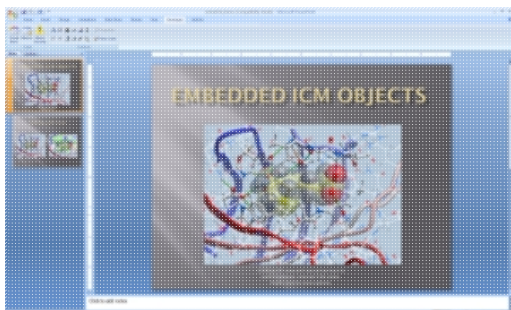


6.6.7 Protect Shell Objects From Deletion

When making a molecular document you can protect objects from deletion by the person who reads your document by:

- Right click on the object in the ICM Workspace.
- Select the **Protect** option.

6.7 ActiveICM



ActiveICM enables you to view and display ICM graphical slides and animations interactively inside Windows Microsoft PowerPoint and web browsers such as Internet Explorer and Mozilla Firefox.

Chapter Contents:

- How to Embed in Microsoft PowerPoint 2003
- How to Embed in Microsoft PowerPoint 2007
- How to Embed in Microsoft PowerPoint 2010
- How to Embed in a Web Browser
- How to Use ActiveICM in PowerPoint
- How to Change ActiveICM Component Properties
- Advanced use of activeICM: Macros to direct visualisation changes
- Background Images

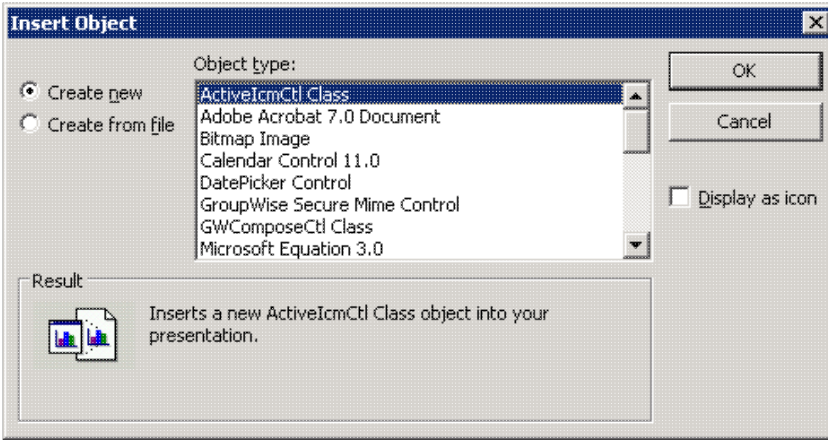
6.8 How to Embed in Microsoft PowerPoint 2003

Setup

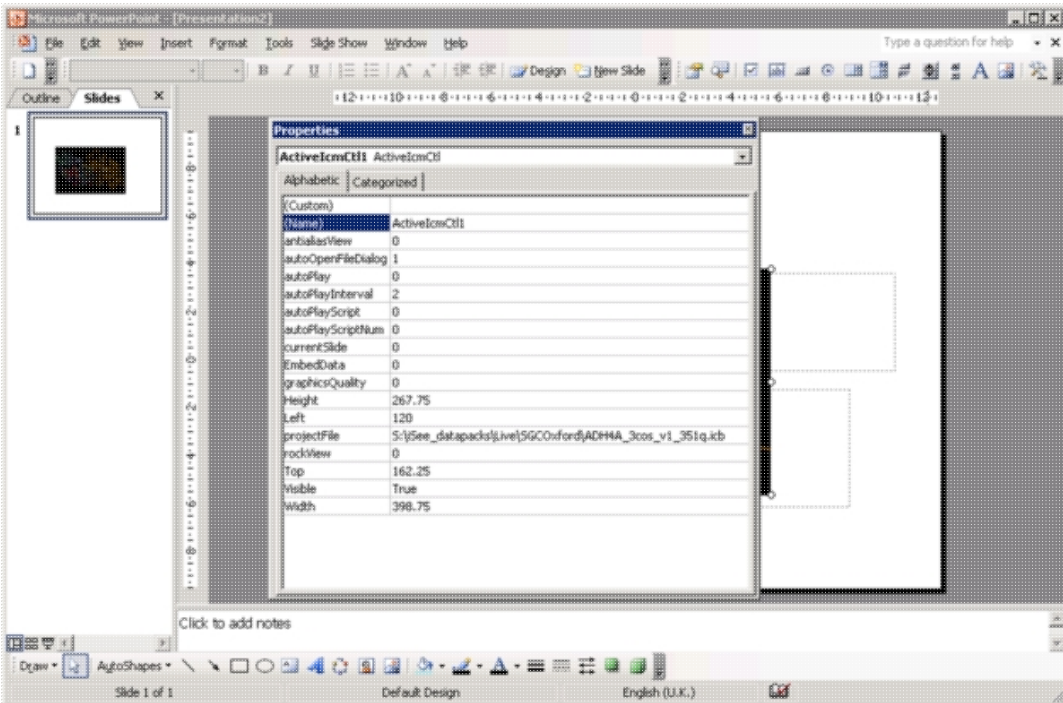
- Download ActiveICM from www.molsoft.com/support
- Save an ICM file (.icb) containing slides. [Click here](#) to see how to make slides.

Embed icb file

- Open the Insert menu from the top bar of PowerPoint and select **Object**
- This opens up the Object dialogue. Select **ActiveIcmCtlClass**:



- Click on OK. A file dialogue will then be opened. Open the ICB file you wish to use via this dialogue. IMPORTANT: To avoid later problems, make sure the ICB file is in the same folder as the PowerPoint file.
- A low-resolution snapshot of the first slide in the ICB file will be shown in the activeICM control you created. You can change the shape of the control by dragging the corners of the control with the mouse, once selected.
- Right-click on the activeICM control and select the Properties menu item



- Save the PowerPoint presentation

6.9 How to Embed in Microsoft PowerPoint 2007

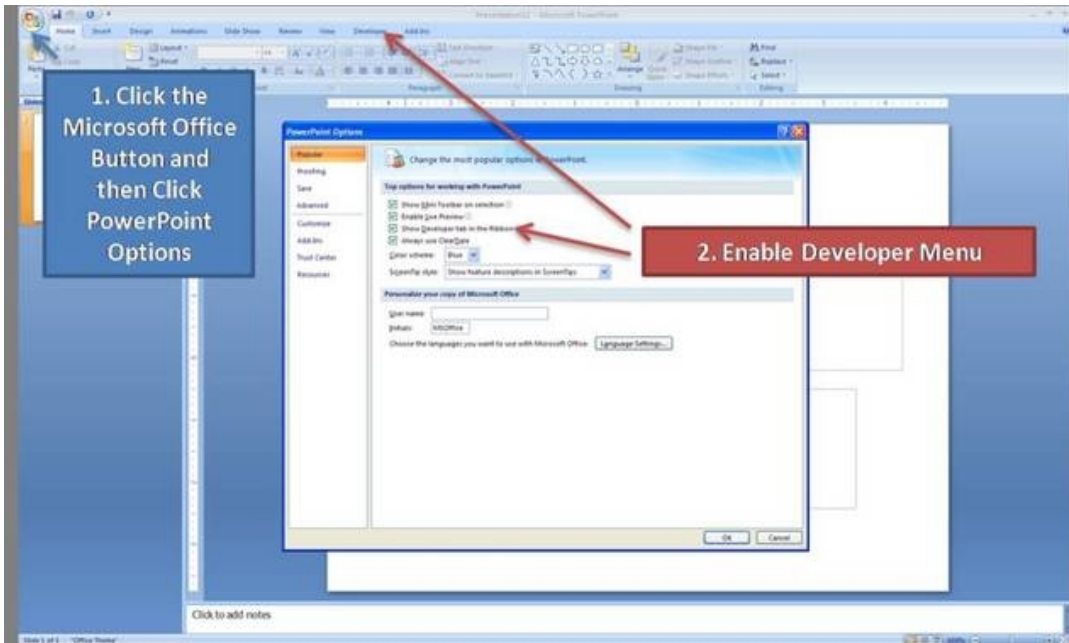
Setup

- Download ActiveICM from www.molsoft.com/support
- Save an ICM file (.icb) containing slides. [Click here](#) to see how to make slides.

NOTE: Here are the instructions for ActiveICM in Microsoft Office 2007, for older versions of PowerPoint see [here](#).

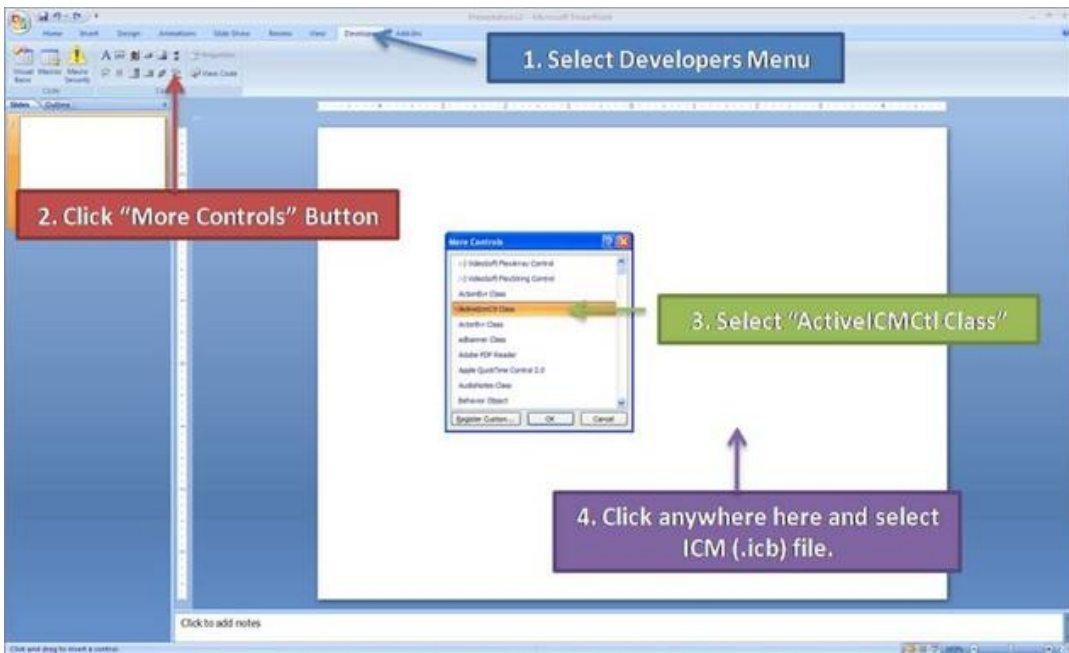
Enable the Developer Menu:

- Click the **Microsoft Office Button** (button top left), and then click **PowerPoint Options**.
- In the **PowerPoint Options** dialog box, click **Popular**.
- Under **Top options for working with PowerPoint**, set the **Show Developer tab in the Ribbon** check box, and then click **OK**.



Insert ActiveICM into PowerPoint:

- Select the **Developer** menu.
- Select the **More Controls** button in the **Controls** field.
- Select **ActiveICMctl Class** from the list of controls and click **OK**.
- Click the mouse anywhere in the white PowerPoint space and a dialog box will be displayed asking you to select your ICM (.icb) file.
- Click and drag at the corners of the image to resize the normal way you would resize an object in PowerPoint.



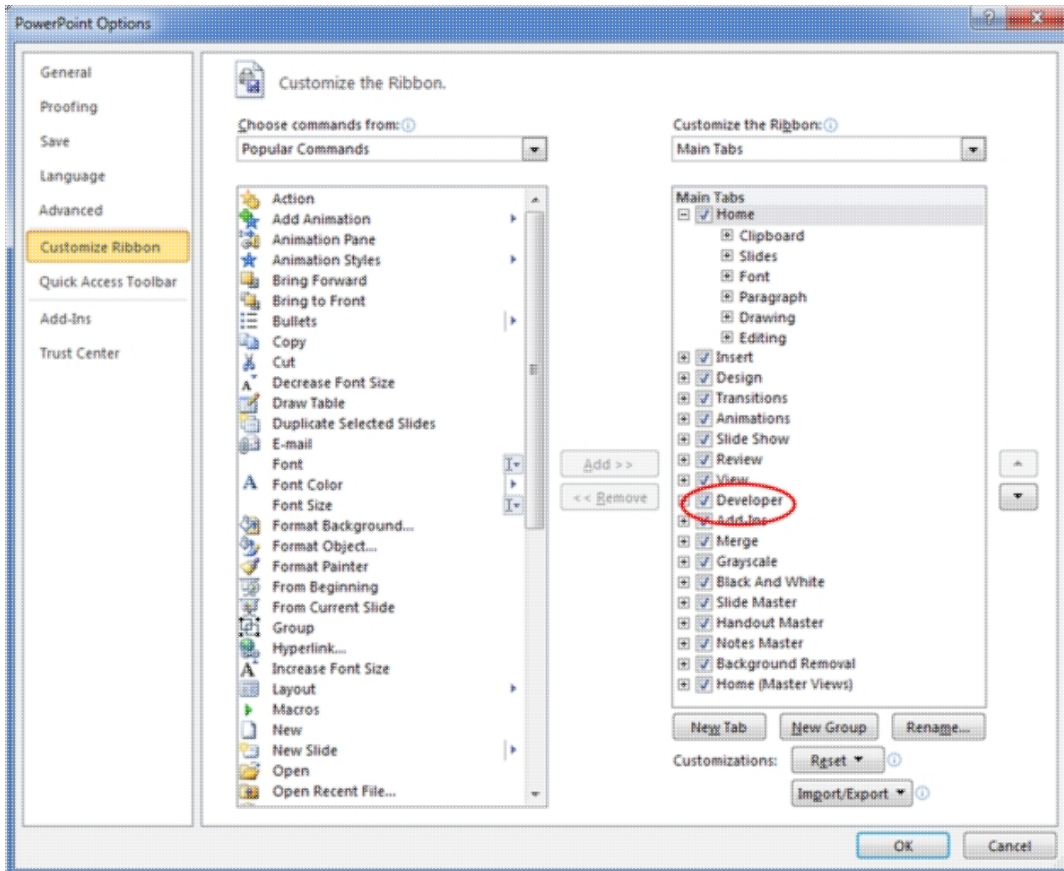
6.10 How to Embed in MicroSoft PowerPoint 2010

Setup

- Download ActiveICM from www.molsoft.com/support . Please always check you are using the latest version.
- Save an ICM file (.icb) containing slides. Click [here](#) to see how to make slides.

Enable the Developer Menu:

- Click on the File tab, and then click Options menu.
- Select the **Customize Ribbon** option and then check the **Developer** option in the right hand side panel (see image below).
- Click on the **Developer Menu**.
- **IMPORTANT!** Click OK and then you will see the Developer Menu in the Ribbon at the top of PowerPoint.
- Select the Developer menu.
- Select the **More Controls** button in the **Controls** field.
- Select **ActiveICMctl Class** from the list of controls and click **OK**.
- Click the mouse anywhere in the white PowerPoint space and a dialog box will be displayed asking you to select your ICM (.icb) file.
- Click and drag at the corners of the image to resize the normal way you would resize an object in PowerPoint.



6.11 Embed in Web Browser

To embed in a web browser.

1. Download ActiveICM from [here](http://www.molsoft.com/getbrowser.cgi?product=activeicm&act=list) <http://www.molsoft.com/getbrowser.cgi?product=activeicm&act=list> (it is free!).

2. Create an HTML page in ICM (File/New/Html).
3. Add a series of slides.
4. File/Export As ActiveICM Html..

NOTE. There is an issue with FireFox21 because it disables the ActiveICM plugin. The workaround at the moment is:

- type about:config in the location (address) bar and press the "Enter" key to open the about:config page, just like you open a website by typing the URL in the location bar.
- if you see a warning message then you can confirm that you want to access the about:config page.
- in the Search bar at the top of the about:config page type 'load_appdir_plugins'
- double click on the preference to set value to true
- restart FF

6.12 How to Use ActiveICM in PowerPoint

****IMPORTANT There are two ways to open a presentation:**

- Double click on the ppt file in windows folder. (in this case PowerPoint will set the current directory to the one which contains the file and there should be no problems with both relative and absolute paths)
- Open ppt through the "File-Open" or recent files. (in this case PowerPoint DOES NOT SET the current directory to the one which contains the file -> relative path might not work and user will be prompted to locate the ICB file unless file is found in absolute location)

To view the slides you must be in Slide Show mode

- Press the **F5** button to start the **Slide Show**. In edit mode (i.e. not presentation mode), the control is shown as a static image $\frac{1}{2}$ it is not possible to interact with the ICB file. Therefore, to prepare the presentation so that the control shows the correct initial visualisations it is necessary to run the PowerPoint slide(s) in presentation mode

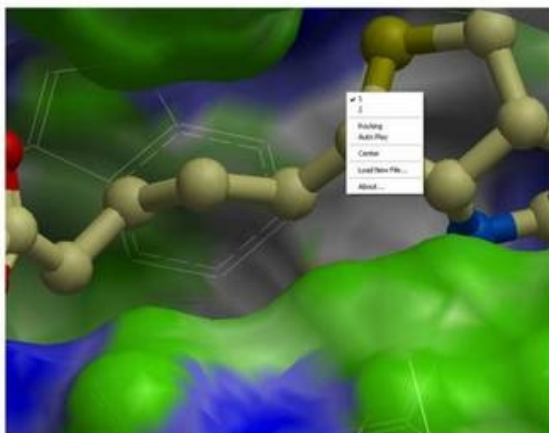
Change Slides

- Use the left and right cursor keys to change slides.

A number of other options can be accessed by right clicking on the slide. These options include:

- Select Slide
- Auto Play
- Set on/off rocking
- Center
- Load a new ICM File

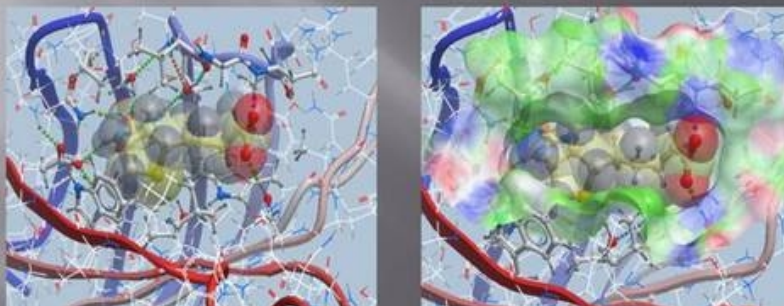
ICM Embedded Into PowerPoint in Slide Show Mode (F5 button)



You can also add multiple ActiveICM 3D displays in one slide:

- To display multiple ActiveICM 3D displays in one slide just copy the original display or repeat the steps described above. All powerpoint slides should point to the same ICM file (.icb) but they can point to different slides.

Side-by-side view



6.13 How to Change ActiveICM Component Properties

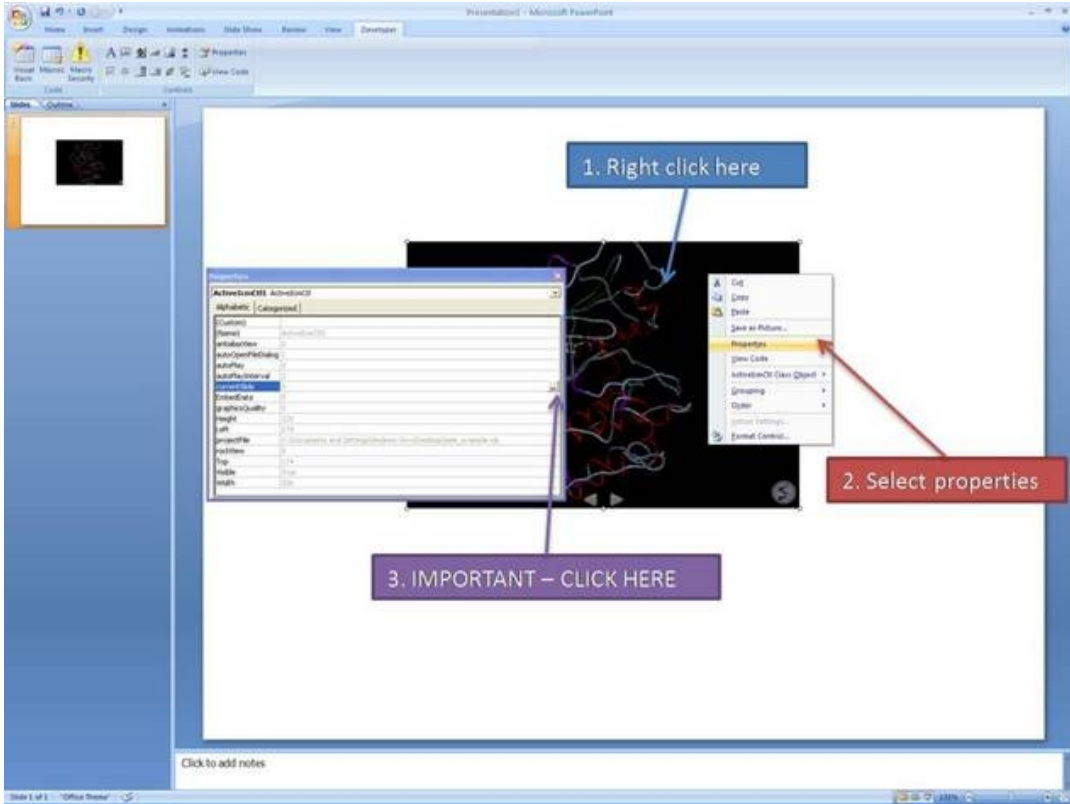
A number of properties of ActiveICM can be changed once embedded in powerpoint. The options include:

- Select the first slide to be displayed.
- Set slide auto play.

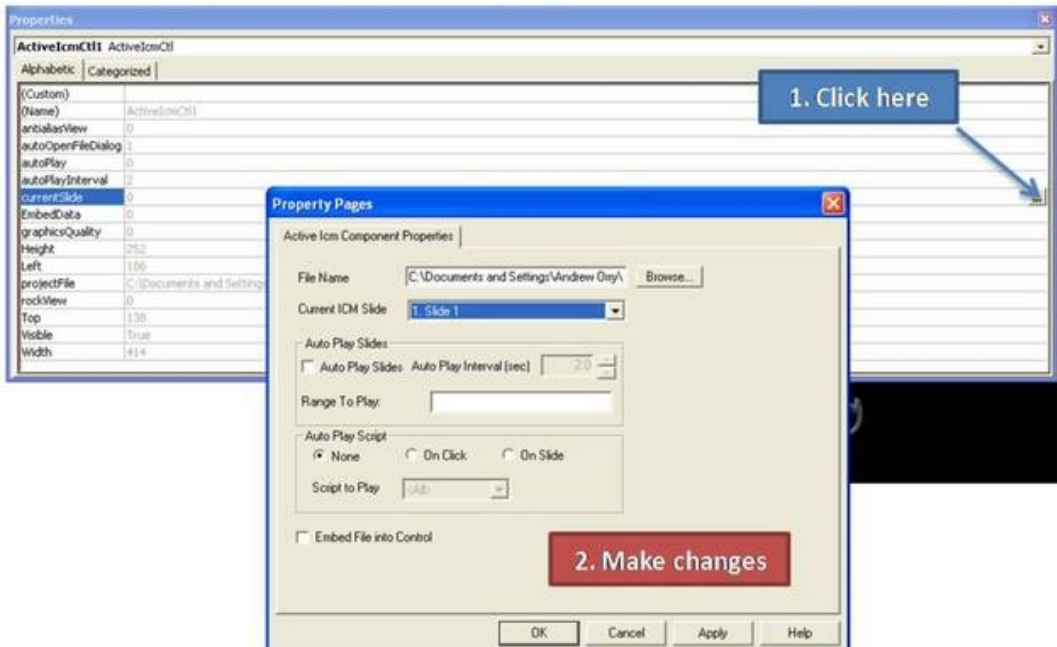
- Set auto play of a script.
- Embed the powerpoint file and the icb file all into one file.

To change these options:

- Right click on your embedded activeICM in Powerpoint.
- Select **Properties** and click on the button shown below.



- A **Property Pages** window will then be displayed as shown below.



To change the file name of the icb file linked to activeICM: Simply type in the path to the file or use the browse option.

To change the current ICM slide: Use the drop down button next to **Current ICM Slide** to select the slide you wish to display first in your presentation.

To auto play slides: Check the **Auto Play Slides** box and select the interval between slides option. A range of slides can be played by entering the number of the slides separated by a comma.

To auto play a script: Select whether you want the script to run **On Click** or **On Slide** then select the script from the **script to play** drop down button. You should first save your script in the icb file.

To embed the icb file in the ppt file Click the **Embed File into Control** option. **Important** - Please save your PowerPoint file in the t 1997-2003 ppt format not pptx.

6.14 Advanced use of activeICM: Macros to direct visualisation changes

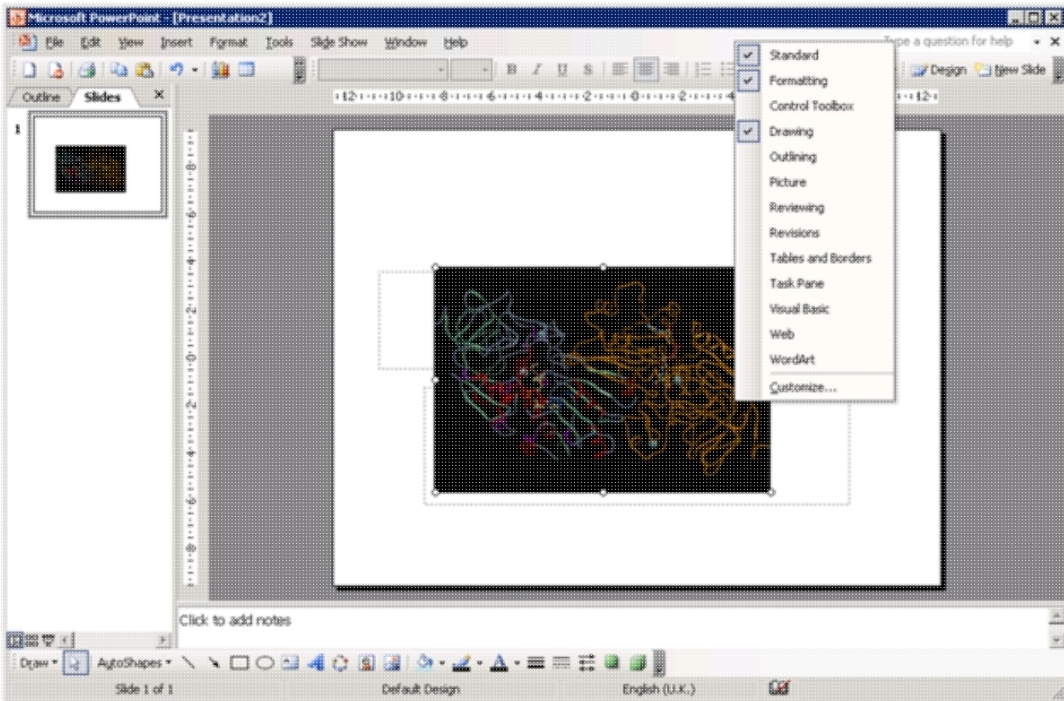
Documentation kindly provided by Dr. Brian Marsden (SGC Oxford
<http://www.sgc.ox.ac.uk/people/brian/>)

It is possible to write simple VisualBasic scripts to avoid having to use the right-click menu approach to changing activeICM control slides within the control itself. This allows one to place buttons outside of the activeICM control, but in the same PowerPoint slide, which controls the control's behaviour. Below are a couple of useful examples of this approach.

Creating a button to set the control's active slide:

Insert a button Office 2003

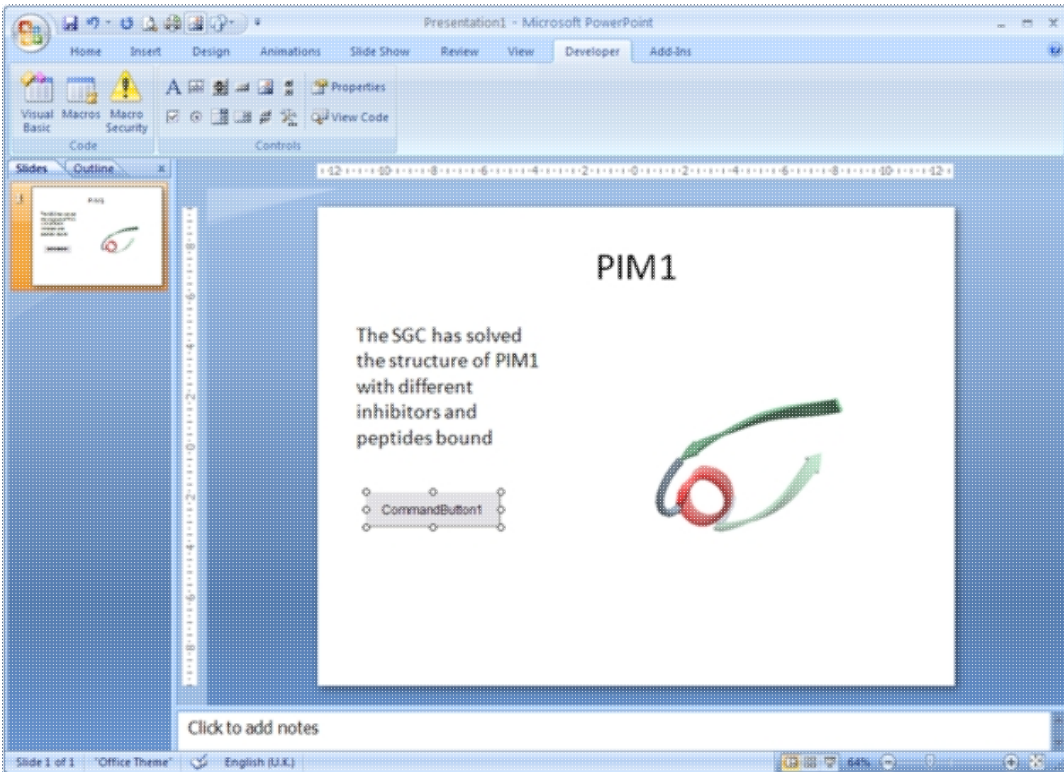
- In edit mode, make sure the control toolbox toolbar is shown by right-clicking the blank area at the top of the top bar and ensuring **Control Toolbox** is ticked.



- Click on an icon in the Control Toolbox which corresponds to the sort of button you wish to use. Then click and drag in the PowerPoint slide to generate the button.

Insert a button Office 2007:

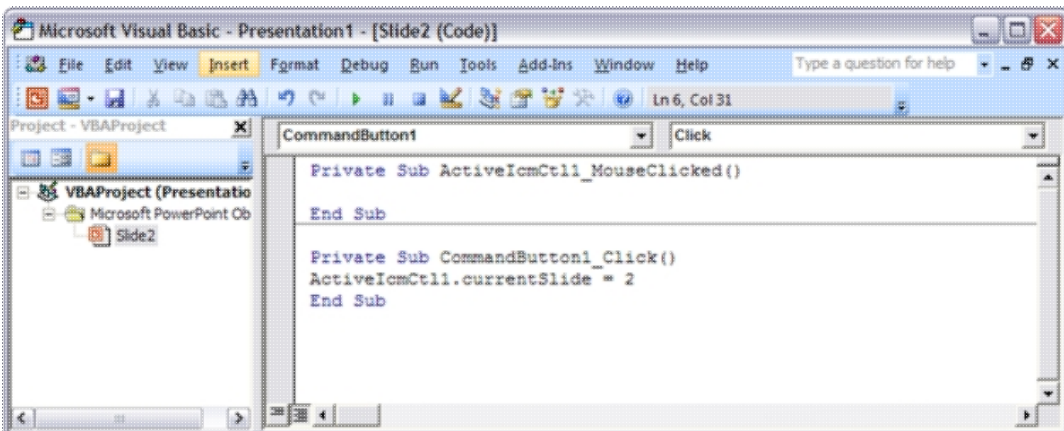
- In edit mode, click on an icon in the Developer menu or ribbon which corresponds to the sort of button you wish to use. Then click and drag in the PowerPoint slide to generate the button.



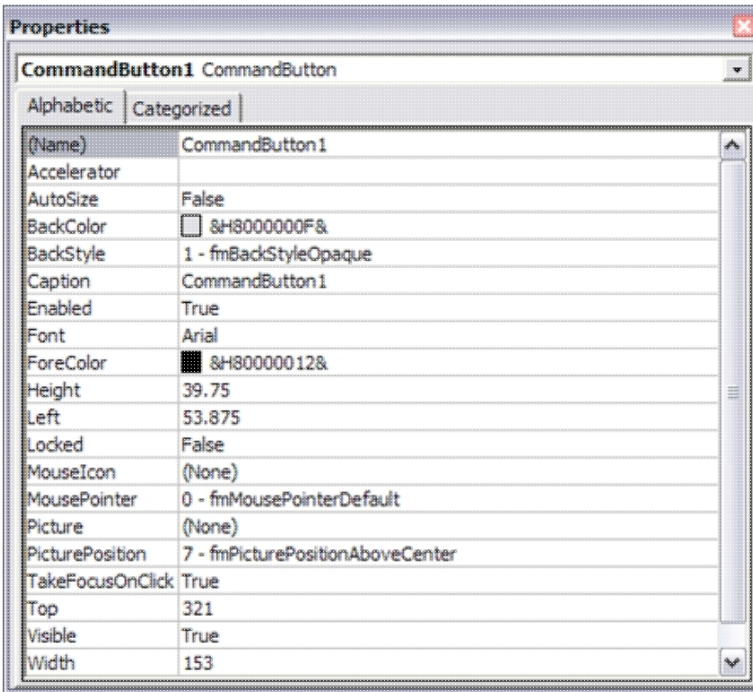
- Double-click on the new button to open the VisualBasic editor with two empty functions pre-defined. The first one pertains to the control itself and can be ignored in this context *For the second function (which is for the newly-created button), copy the following into the editor, between the two lines of function code:

```
ActiveIcmCtl1.currentSlide = 2
```

- This sets the current activeICM control's slide to be number 3 **note** that the value placed in this code needs to be 1 less than the actual slide number (confusing, no?). Obviously, use a value here that makes sense in the context of your ICB file.
- This should leave the editor looking like this:



- Close the Visual Basic editor
- To change the physical properties of the button e.g. text, colour e.t.c.right-click on the button and select the Properties menu option. This opens up a dialogue as below, where many properties of the button can be changed:



- Using this dialogue, it should be possible to disguise the button to look like normal text (for example) which can be clicked on during the presentation to change the visualisation of the control, apparently magically. Note that the button will only work in presentation mode.
- **IMPORTANT:** In Office 2007, remember to save the PowerPoint presentation now as a pptm file that is, a macro-enabled PowerPoint file otherwise the macros will not work next time you load the presentation.

Other code examples: Just copy and paste the example of interest inside the function for the button in the Visual Basic editor. Code that enables a button to cycle through the ICB files slides in order (including wrap-around)

```
currentSlide = ActiveIcmCtl1.currentSlide
numSlides = ActiveIcmCtl1.nofSlides
If currentSlide = numSlides - 1 Then
ActiveIcmCtl1.currentSlide = 0
Else
ActiveIcmCtl1.currentSlide = currentSlide + 1
End If
```

6.14.1 PowerPoint Cache Errors

PowerPoint caches some information about active controls. Sometimes after an ActiveICM upgrade you may get an error when trying to access some property or method: "Wrong number of arguments or invalid property assignment" or something similar.

In this case you need to close PowerPoint and remove all files from the location below:

C:\Documents and Settings\seva\Local Settings\Temp\PPT11.0

6.15 Background Images

In ActiveICM version 1.1-5 and higher you can add background images to your icb file and display in PowerPoint and the web using ActiveICM. The documentation on how to insert background images can be found [here](#).

6.16 ICM JavaScript (IcmJS)

IcmJS (formerly known as ActiveIcmJS) is a JavaScript/HTML5 version of the ActiveICM plugin. You can read more about IcmJS [here](#).

To make a fully interactive 3D molecular web page from ICM and embed using IcmJS:

- Use ICM version 3.8-5 or higher. If you do not have access to this please contact MolSoft to request it.
- Make a set of *slides*.
- Go to File/Export as ActiveICM Html and then select the IcmJS check box.
- The IcmJS file will then be saved to your computer which you can open in any modern browser e.g. Chrome, FireFox or Safari.

7 Working with Sequences and Alignments

```
a 2phk_a align 277 Amino KPBG_RABBIT
1  GFYENYEPKE ILGRGUSSUU RRCIHKPTCK EYAUKIIDUT GGGSFSAEEU
51  QELREATLKE UDILRKUSGH PNIIQLKDTY ETNTFFFLUF DLMKKGELFD
101 YLTEKVTLSE KETRKIMRAL LEVICALHKL NIUHRDLKPE NILLDDDMNI
151 KLTDFGFSCQ LDPGEKLREU CGTPSYLAPE IIEGSMNDNH PGYGKEUDMW
201 STGUIMYTLL AGSPPFWHRK QMLMLRMIMS GNYQFGSPEW DDYSDTUKDL
251 USRFLVUQPQ KRYTAEERAL HPFFQQY
```

Many powerful sequence manipulation tools are contained within ICM. In this chapter we will describe some of the bioinformatic manipulations you can perform with sequences and alignments.

Chapter Contents:

- Read and Save Sequences
- Sequence Secondary Structure
- Bioinfo Menu Options
- BLAST - Sequence Search and Align
- Sequence Alignments
- Alignment Editor

7.1 Read and Save Sequences

Available in the following product(s): ICM-Pro

There are a number of different ways to load a sequence into ICM via the Graphical User Interface.

7.1.1 Read a Sequence from UniProt

Uniprot is a curated protein sequence database which provides a high level of annotation (such as the description of the function of a protein, its domains structure, post-translational modifications, variants, etc.).

To read a sequence from UniProt:

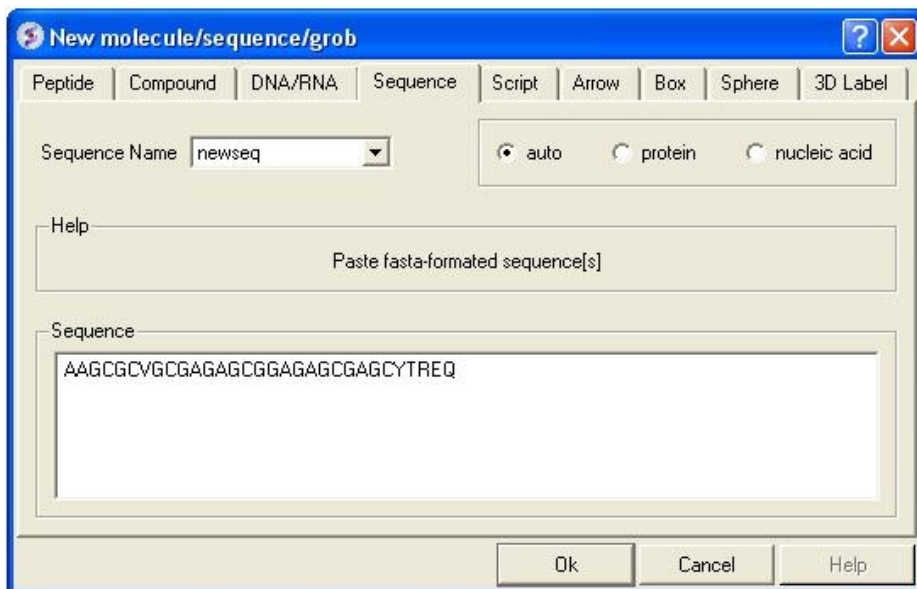
- File/Load/UniProt
- Enter UniProt ID

or use the option in the Search Tab.

7.1.2 Cut and Paste a Sequence

Paste your own sequence into ICM

- File/New/Sequence
- Paste the Sequence into the box provided.



7.1.3 Extract a Sequence from a PDB File

Extract a Sequence from a PDB file

- Right click on a molecule in a loaded PDB file in the ICM Workspace.
- Select Extract Sequence(s) and the sequence will be loaded into the Sequences section of the ICM Workspace.

7.1.4 Read directly from a Sequence File

To read a sequence from a file:

If you have a sequence file saved in FASTA format it can be read into ICM by:

- File/Open

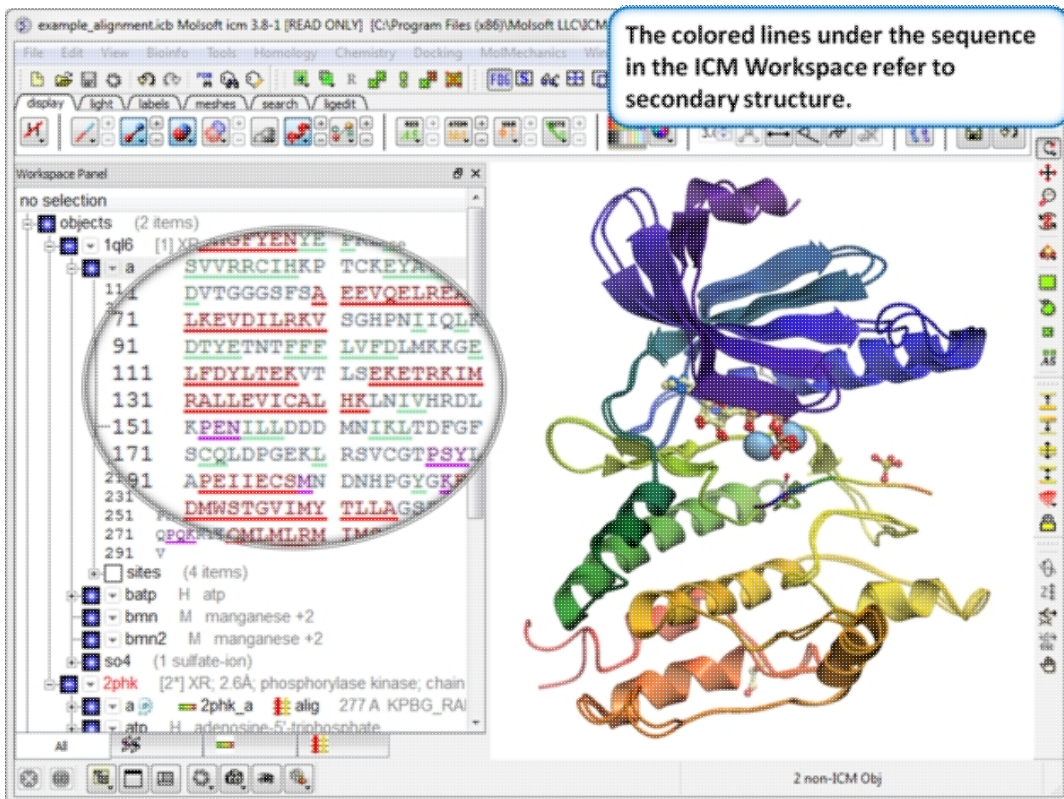
7.1.5 Save Sequence

To save a sequence in FASTA format:

- Right click on the sequence name in the ICM workspace.
- Select the option Save As.
- Choose FASTA format from the Windows "Save as type:" drop down menu option.

7.2 Sequence Secondary Structure

If a sequence has assigned secondary structure then colored lines will appear under the sequence in the ICM Workspace. Regions underlined in red = alpha helix, green = beta sheet, blue = pi helix, magenta = 3/10 helix.

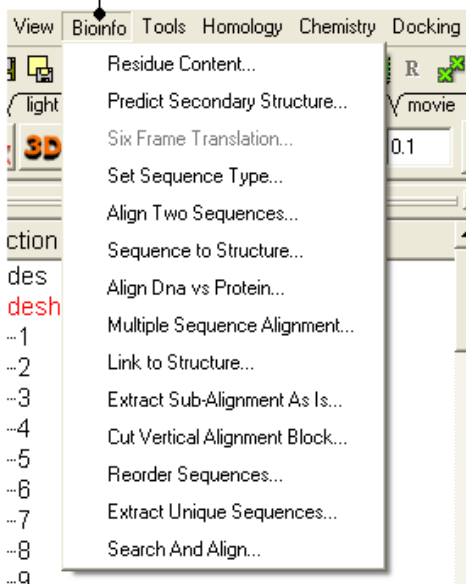


7.3 Bioinfo Menu

Available in the following product(s): ICM-Pro

Note: Click **Next** (top right hand corner) to navigate through this chapter. Headings are listed on the left hand side (web version) or by clicking the **Contents** button on the left-hand-side of the help window in the graphical user interface.

Bioinformatics Menu



7.3.1 Residue Content

To determine the residue content of a sequence.

- Bioinfo/Residue Content and a data entry box as shown below will be displayed.
- Enter the sequence name. See the Load Sequence section for more information on how to load a sequence into ICM using the Graphical User Interface.
- A table and graph of residue frequencies will be displayed.



7.3.2 Predict Secondary Structure

To predict the secondary structure of a sequence:

- Bioinfo/Predict Secondary Structure
- Enter the sequence name. See the Load Sequence section for more information on how to load a sequence into ICM using the Graphical User Interface.
- An option is provided to ignore currently assigned secondary structure.

To view the secondary structure prediction click on and expand the sequence in the ICM workspace. Regions underlined in red = alpha helix, green = beta sheet, blue = pi helix, magenta = 3/10 helix.

```
1  MNGTEGPNFY  UPFSNKTGVU  RSPFEAPQYY  LAEPWQFSML
41  AAYMFLIIML  GFPINFLTLY  UTUQHKKLRT  PLNYILLNLA
81  UADLFMVFGG  FTTTLYTSLH  GYFVFGPTGC  NLEGGFATLG
121 GEIALWLSLUV  LAIERVUUUC  KPMSNFRFGE  NHAIMGUAFI
161 WUMALACAAP  PLUGWSRYIP  EGMQCSCGID  YYTPHEETNN
201 ESFUIYMFVV  HFIIPLIUIF  FCYQQLUFTV  KEAASATTQ
241 KAEKEUTRMV  IIMUIAFLIC  WLPYAGUAFY  IFTHQGSDFG
281 PIFMTIPAFF  AKTSAUYNPU  IYIMMNKQFR  NCHUTTLCCG
321 KNPSTTVSRT  ETSQVAPA
```

7.3.3 Six Frame Translation

This options returns the translated DNA or RNA sequence ('-' for a Stop codon, 'X' for an ambiguous codon) using the standard genetic code.

- Read into ICM a DNA sequence from a file (eg File/Open FASTA) or use the File/New option and cut and paste a DNA sequence.
- Bioinfo/Six Frame Translation
- Translate all frames or use start codon.

7.3.4 Set Sequence Type

This option allows you to define whether a sequence that is read into ICM is a protein or nucleotide sequence.

- Read into ICM a sequence (eg File/New and cut and paste sequence or File/Open FASTA)
- Bioinfo/Set Sequence Type
- Select the sequence name using the drop down button
- Select sequence type protein or DNA.

7.3.5 Align Two Sequences

Please see the Alignment Chapter.

7.3.6 Sequence to Structure alignment

Please see the Alignment Chapter.

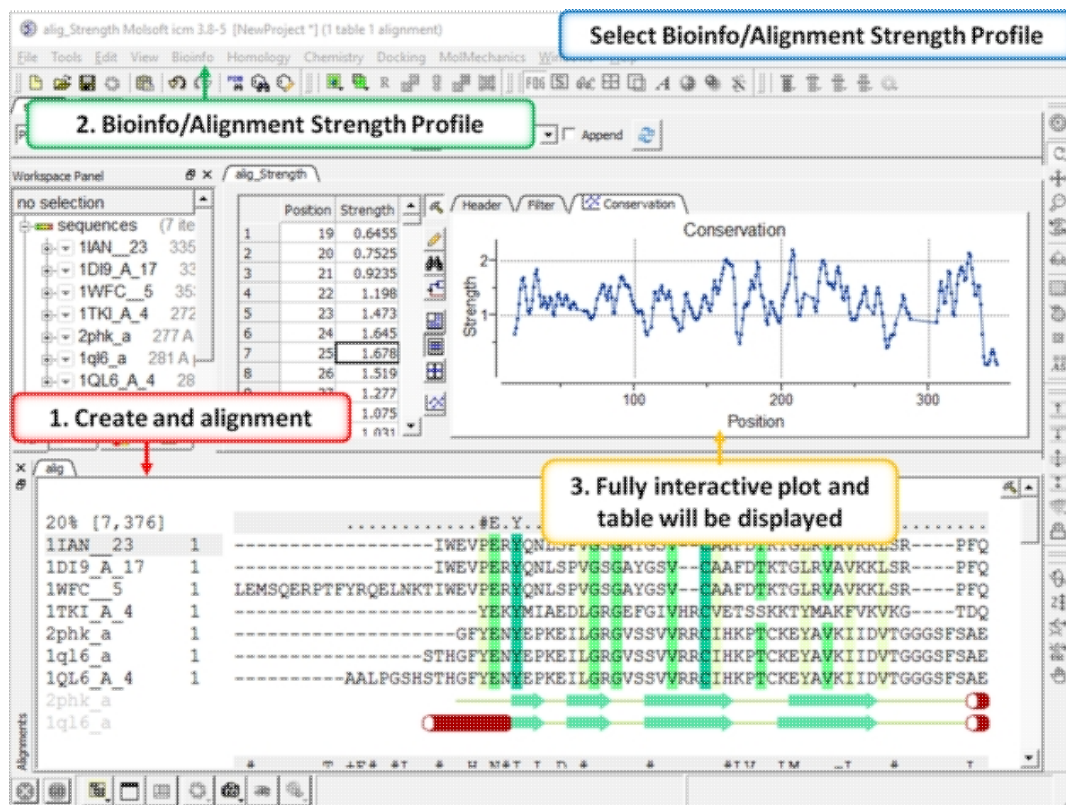
7.3.7 Multiple Sequence Alignment

Please see the Alignment Chapter.

7.3.8 Alignment Strength Profile

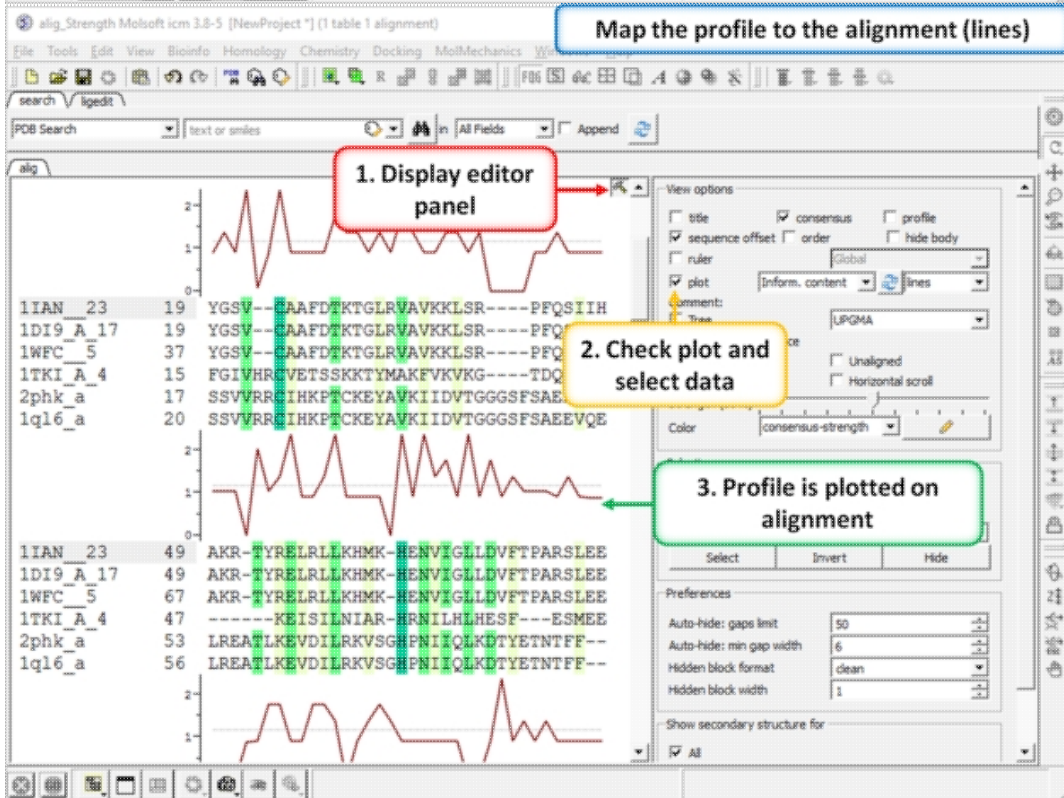
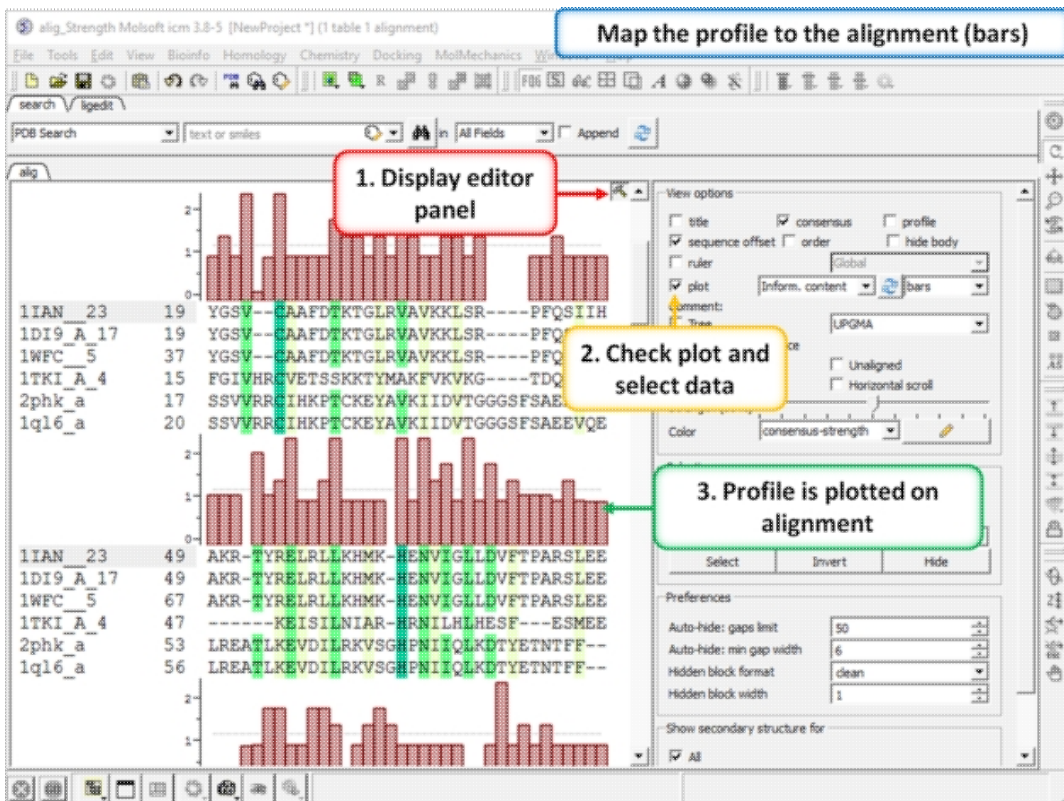
To view a fully interactive alignment strength profile:

- Create an alignment.
- Bioinfo/Alignment Strength Profile
- A fully interactive table and plot will be displayed.



The profile can be plotted onto the alignment:

- Open the alignment editor panel.
- Select the plot checkbox and choose either bars or line plot.



7.3.9 Link to Structure

To link a structure to an alignment:

- Double click on the structure in the ICM workspace to select it.

- Bioinfo/Link to Structure

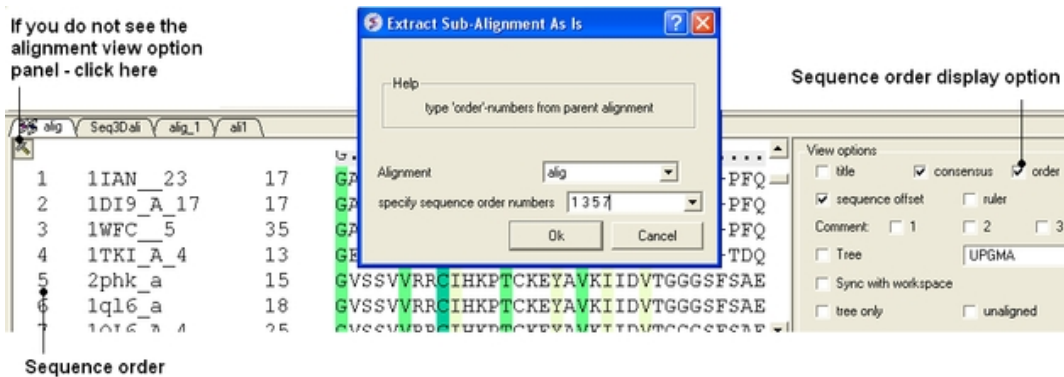
NOTE Links are described in more depth in the Making Links Section of the manual.

7.3.10 Extract Sub-Alignment As Is

On occasion you may want to extract a sub alignment from a bigger alignment. For example you may only want to extract the alignment for the sequences linked to a structure.

To extract a sub-alignment:

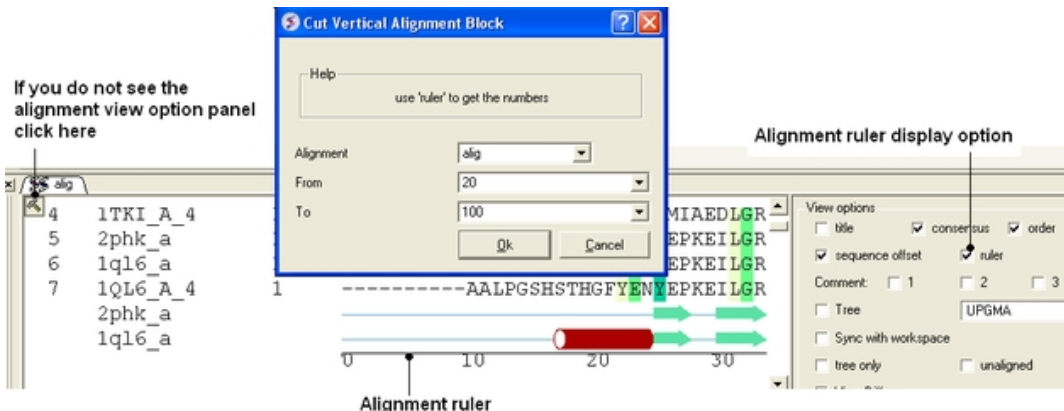
- An initial multiple sequence alignment must first be displayed in the graphical user interface.
- Bioinfo/Extract Sub-Alignment As Is
- Enter the name of the alignment from which you wish to extract a sub-alignment from.
- Specify the sequence order numbers you wish to extract - enter each number separated by a space. You can see the sequence order alignment number by selecting the **order** option in the alignment view options panel. See image below.
- Click OK and the extracted sequence alignment will be displayed in a separate alignment tab.



7.3.11 Cut Vertical Alignment Block

To cut a vertical alignment block:

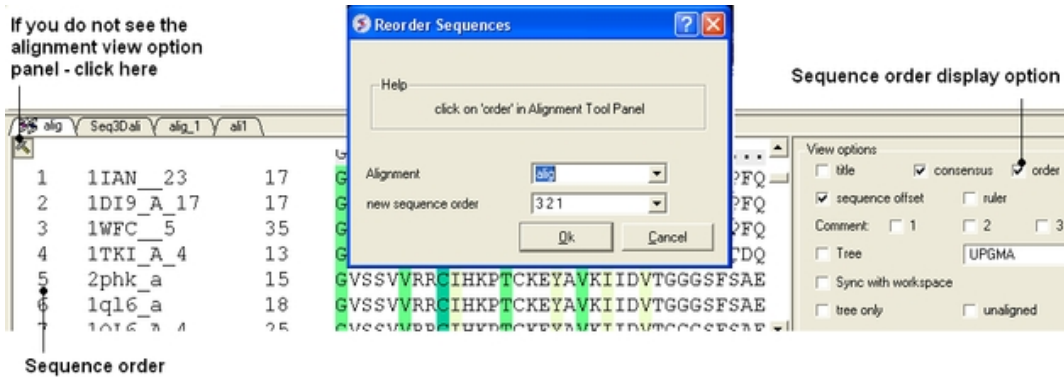
- An initial alignment must first be displayed in the graphical user interface.
- Bioinfo/Cut Vertical Alignment Block
- Enter the alignment from which you wish to cut from.
- Enter the region of the alignment you wish to cut (from: to:). The easiest way to determine the region to cut is to display the ruler in the alignment. This is an option in the alignment view panel - see image below.
- Click OK and the cut section will be displayed in a new alignment.



7.3.12 Reorder Sequences

To reorder sequences in an alignment

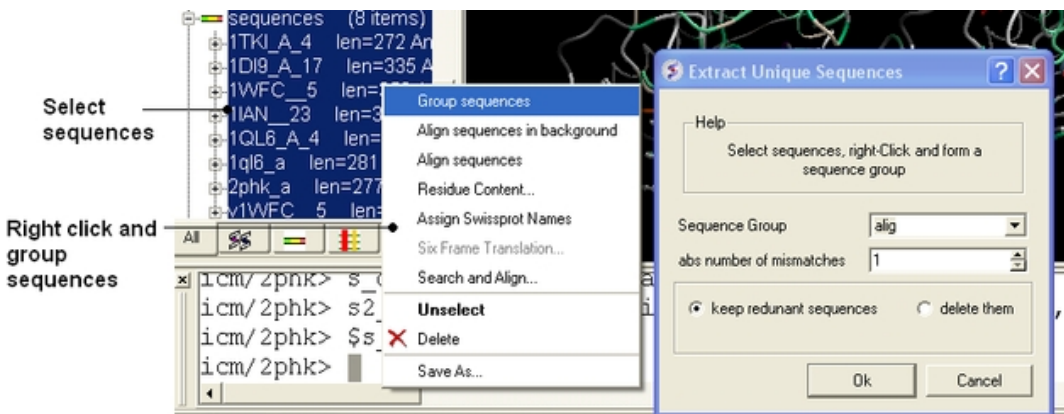
- An initial multiple sequence alignment must first be displayed in the graphical user interface.
- Bioinfo/Reorder Sequences
- Enter the alignment name
- Enter the new sequence order. You can see the sequence order alignment number by selecting the **order** option in the alignment view options panel. See image below below.



7.3.13 Extract Unique Sequences

To extract unique sequences from a group of sequences:

- Read into ICM the sequences you wish to make unique.
- Select the sequences. A sequence can be selected by double clicking (highlighted blue in ICM workspace) - a range of sequences in the ICM Workspace can be selected by holding down the SHIFT button and double clicking. A non-contiguous selection can be made by holding down the CTRL button and double clicking.
- Right click on the sequence selection in the ICM Workspace and select Group sequences
- Bioinfo/Extract Unique Sequences
- Enter the name of the sequence group.
- Enter the number of residue mismatches necessary to determine that a sequence is unique or not.
- Select whether you want to keep the redundant sequences or delete them from ICM.



7.3.14 Load Example Alignment

To see an example of an alignment select:

- Bioinfo/Load Example Alignment

7.4 BLAST Search

Available in the following product(s): ICM-Pro

To BLAST the NCBI sequence database (including PDB sequences):

- Go to the Search tab
- Choose the drop down option for BLAST and follow the instructions [here](#).

To BLAST search a local BLAST formatted file:

ICM provides a fast tool to search any Blast-formatted database with a query sequence and generate an alignment on the fly. For example you can use this option to find template structures for homology modeling.

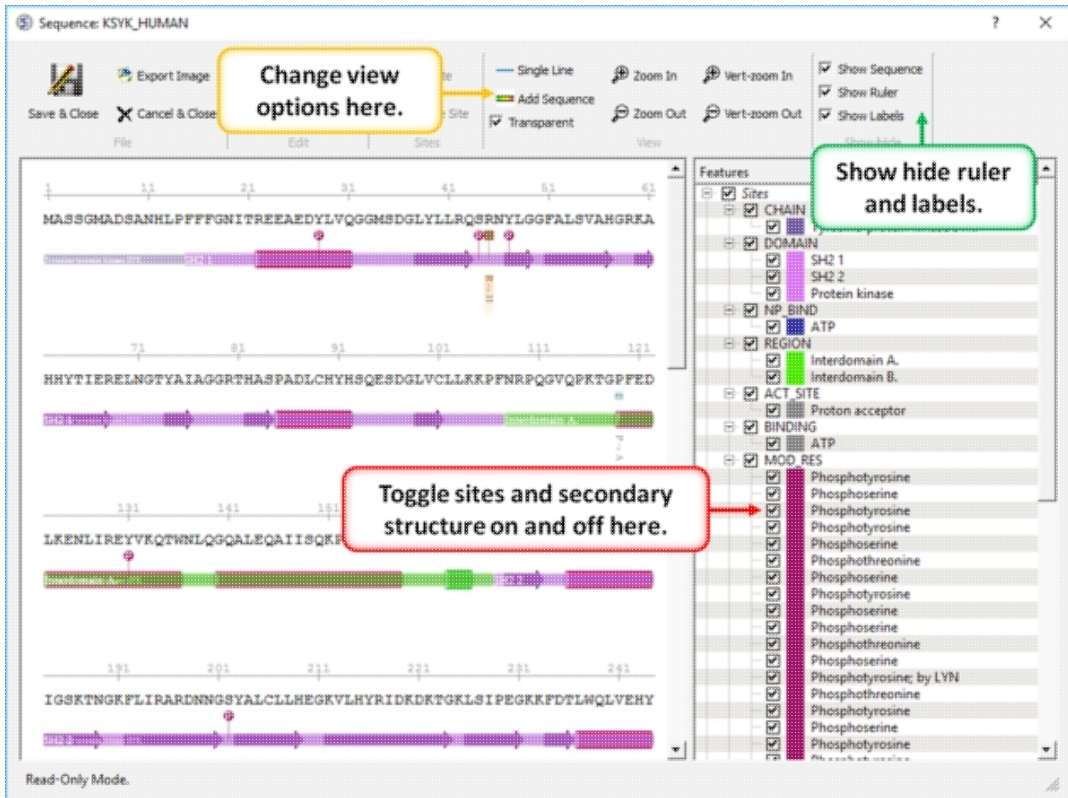
1. Select Bioinfo/Search and Align (or right click on a sequence)
2. Enter the sequence name you wish to search against the database with.
3. Locate the blast-formatted database to search. You can download Blast formatted databases from [here ftp://ftp.ncbi.nih.gov/blast/db/](ftp://ftp.ncbi.nih.gov/blast/db/) eg. pdbaa - PDB sequence database.
4. Specify the speed parameter (0 is the slowest, but the most detailed, 100 is the fastest search for nearly identical sequences).
5. Enter the number of hits you would like to view
6. Enter an identity threshold filter to narrow down the number of hits.
7. The top hits will be displayed in a table.

7.5 Sequence Editor

The sequence editor allows you to easily display and create sequence sites. To open the editor:

- Read in a sequence into ICM.
- Right click on the sequence in the ICM workspace and choose "Sequence Editor".

To toggle site display on and off and change view options:



To edit a sequence or site:



To save your changes or export an image:



7.6 Sequence Alignments

Available in the following product(s): ICM-Pro

7.6.1 Default Alignment

To make an alignment using the default parameters:

example_alignment.lib Molsoft icm 3.8-1 [READ ONLY] [C:\Program Files (x86)\Molsoft LLC\ICM

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics Win

Workspace Panel

1 ID19_A 17,1 1IAN 23,1 1QL6

objects (2 items)

- 1ql6 [1] XR; 2 4A; phospho
- 2phk [2] XR; 2 6A; phospho

sequences (7 items)

- 1TKI_A 4 272A
- 1DG_A 1 335
- 1WFC_5 353A
- 1IAN_23 335A
- 1QL6_A 4 286A
- 1ql6_a 281 A phosphorylase
- 2phk_a 277 A phosphatase

alignments (1 item)

- align id=75 nSeq=7

2phk_a

1ql6_a

13 PVGSGAYGSV--CAAFDKTGLRVAVKKLS

13 PVGSGAYGSV--CAAFDKTGLRVAVKKLS

31 PVGSGAYGSV--CAAFDKTGLRVAVKKLS

9 DLGRGEFGIVHRCVETSCKKTYMAKFVKVK

24 ILGRGVSSVVRRCIHKPTCKEYAVKIIDVT

24 ILGRGVSSVVRRCIHKPTCKEYAVKIIDVT

21 ILGRGVSSVVRRCIHKPTCKEYAVKIIDVT

41 R----PFQSI I HAKR-TYRELRLLLKHKM-E

41 R----PFQSI I HAKR-TYRELRLLLKHKM-E

59 R----PFQSI I HAKR-TYRELRLLLKHKM-E

39 G----TDQVLVK-----KEISILNIAR-E

54 GGGFSFAEEVQELREATLKEVDILRKVSGE

54 GGGFSFAEEVQELREATLKEVDILRKVSGE

51 GGGFSFAEEVQELREATLKEVDILRKVSGE

2phk_a

1ql6_a

2 non-ICM Obj

How to make a sequence alignment.

1. Select the sequences you wish to align. Hold CTRL key and click.

2. Right click on the selection and choose "Align Sequences".

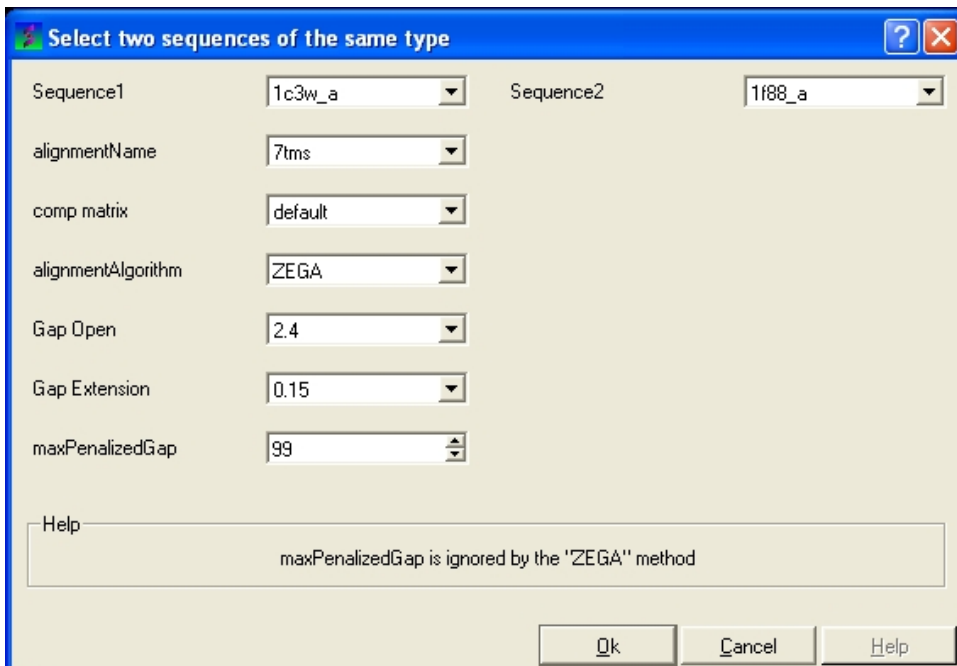
- Read in the sequences you wish to align.
- Select the sequences in the ICM workspace.
- Right click and choose "Align Sequences".

Alignment parameters (such as alignment method) can be changed using the options in the Bioinfo menu (e.g. Align Two Sequences and Multiple Sequences).

7.6.2 Align Two Sequences

To align two sequences:

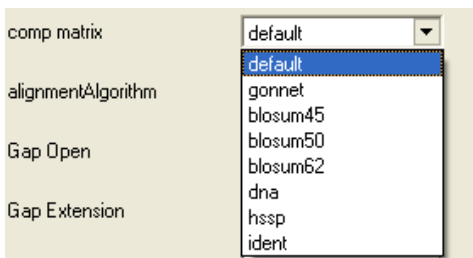
- Read into ICM two or more sequences.
- Bioinfo/Align Two Sequences



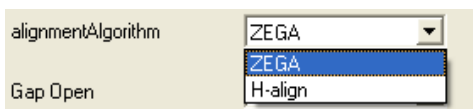
- Enter the name of your first sequence in the 'Sequence 1' data entry box.
- Enter the name of your second sequence in the 'Sequence 2' data entry box.

NOTE: Any sequences already loaded into ICM can be seen by clicking on the down arrow next to the 'Sequence 1 and 2' data entry boxes. This can save typing and trying to remember what you called your sequence.

- Enter a unique alignment name in the 'alignmentName' data entry box.
- Select a comparison matrix from the list shown below by clicking on the arrow next to the 'comp matrix' data entry box.



- Select the alignment algorithm you wish to use from the list shown below by clicking on the arrow next to the 'alignmentAlgorithm' data entry box.



ZEGA - a Zero End-gap Global Alignment, that is a pairwise alignment method based on the Needleman and Wunsch algorithm modified to use zero gap end penalties. This type of alignment was first described by Michael Waterman, who called it the "fit" alignment. The paper of Abagyan and Batalov, 1997 describes the statistics of the structural significance of the alignment score and optimization of the alignment parameters for the best recognition of structurally related proteins.

H-Align - alignment method used in the Align and Score functions and find database command (as described in Batalov and Abagyan, 1999)

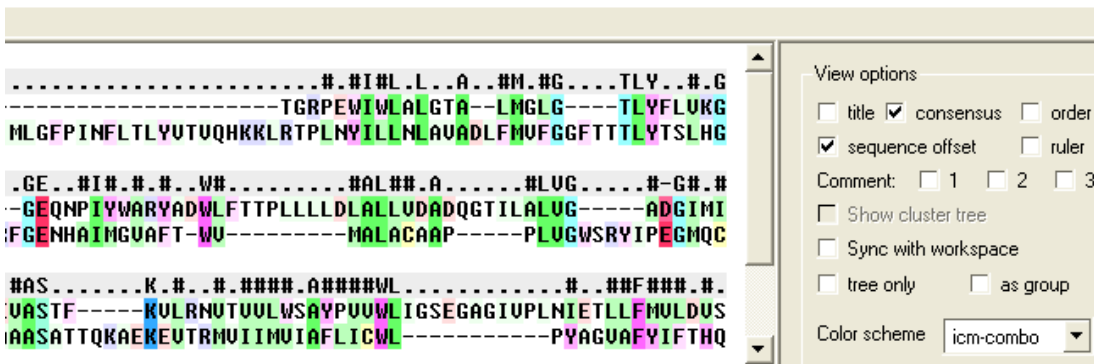
- Enter the values you wish to use for Gap Open, Gap Extension and the maximum penalized gap penalty.

Gap Open The absolute gap penalty is calculated as a product of gapOpen and the average diagonal element of the residue comparison table. You may vary gapOpen between 1.8 and 2.8 to analyze dependence of your alignment on this parameter. Lower pairwise similarity may require somewhat lower gapOpen parameter. A value of 2.4 (gapExtension=0.15) was shown to be optimal for structural similarity recognition with the Gonnet et. al.) matrix, while a value of 2.0 was optimal for the Blosom50) matrix (Abagyan and Batalov, 1997).

Gap Extension The absolute gap penalty is calculated as a product of gapExtension and the average diagonal element of the residue comparison table.

maxPenalizedGap The maximum penalized gap which is used for Gap Open and Extension

- Click OK and the alignment will be displayed in the alignment editor window at the bottom of the graphical user interface.
- Remember to save the project or write the alignment if you wish to keep the alignment for use at another time.



7.6.3 Sequence to Structure alignment

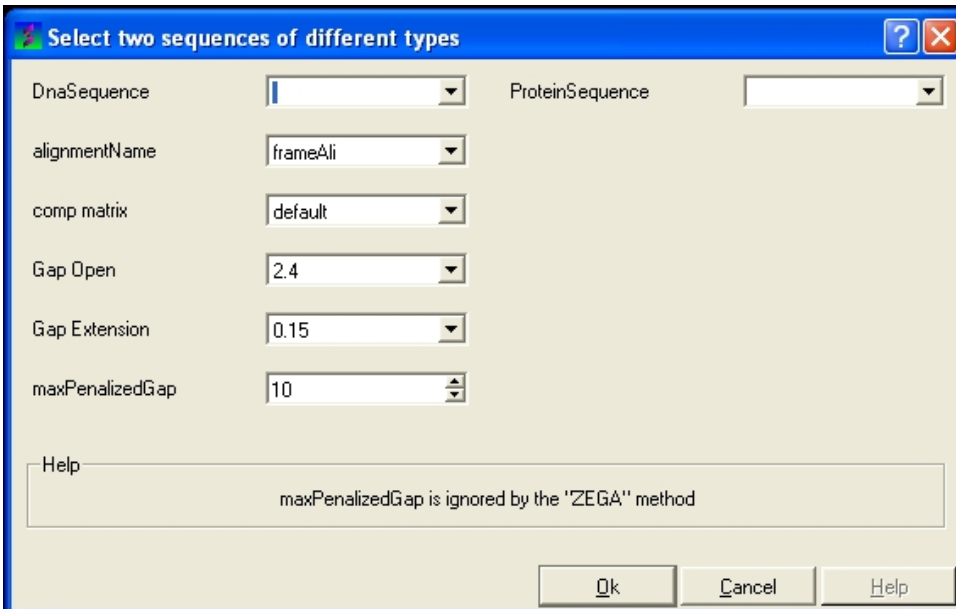
This option allows you to align a sequence to a template structure sequence using secondary structure weighting.

- Read into ICM the sequence (ModelSeq) you wish to align to the template sequence.
- Read in the template (TemplateSeq) structure and extract the sequence from this structure
- Bioinfo/Sequence to Structure Alignment
- Enter the ModelSeq and the TemplateSeq name
- Enter the name you wish to call the alignment.
- Enter the weights you wish to use for alpha and beta secondary structure. The default values have been very well tested.
- This function uses a dynamic algorithm to find the alignment of the locally structurally similar backbone conformations. The RMSD is calculated within a certain residue window. The default is 3.
- Press OK and the alignment will be displayed in the bottom of the gui interface.

7.6.4 Align DNA vs Protein

To align DNA to protein:

- Select the 'Bioinfo' menu.
- Select the option Align DNA vs Protein
- Follow the data entry instructions shown in the previous section entitled "align two sequences" but enter one DNA sequence and one protein sequence.



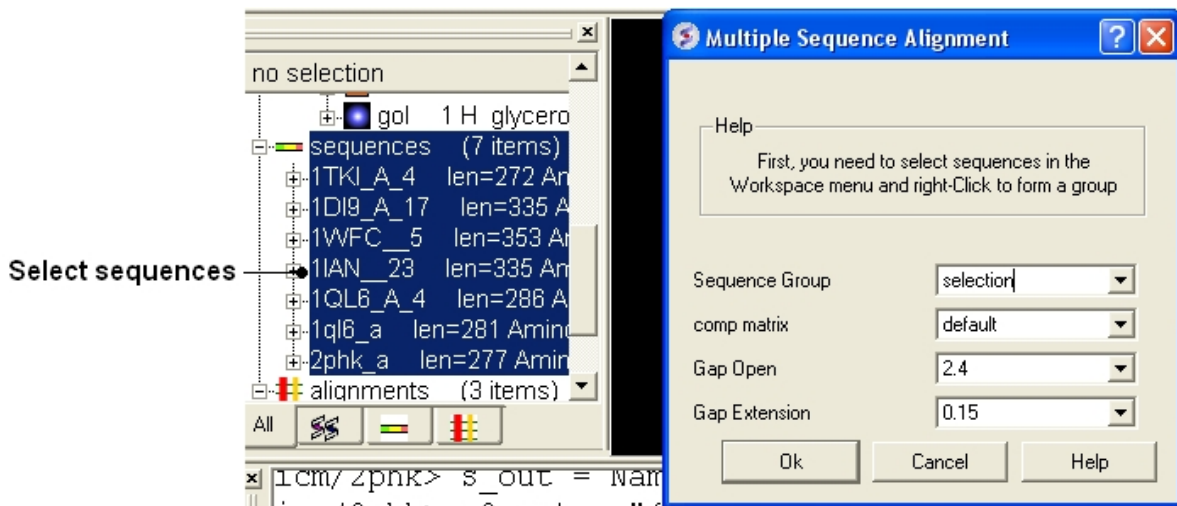
7.6.5 Multiple Sequence Alignment

To align more than 2 sequences:

- Read into ICM the sequences you wish to align.
- Select the sequences you wish to align in the ICM workspace. A sequence can be selected by double clicking (highlighted blue in ICM workspace) - a range of sequences in the ICM Workspace can be selected by holding down the SHIFT button and double clicking. A non-contiguous selection can be made by holding down the CTRL button and double clicking.
- Bioinfo/Multiple Sequence Alignment
- Enter the name of the sequence group. If you selected the sequences as described above then the name of the group is **selection**. Other named groups of sequences can be made by right clicking on the sequence selection.
- Select the comparison matrix you would like to use.
- Enter Gap open and extension values.

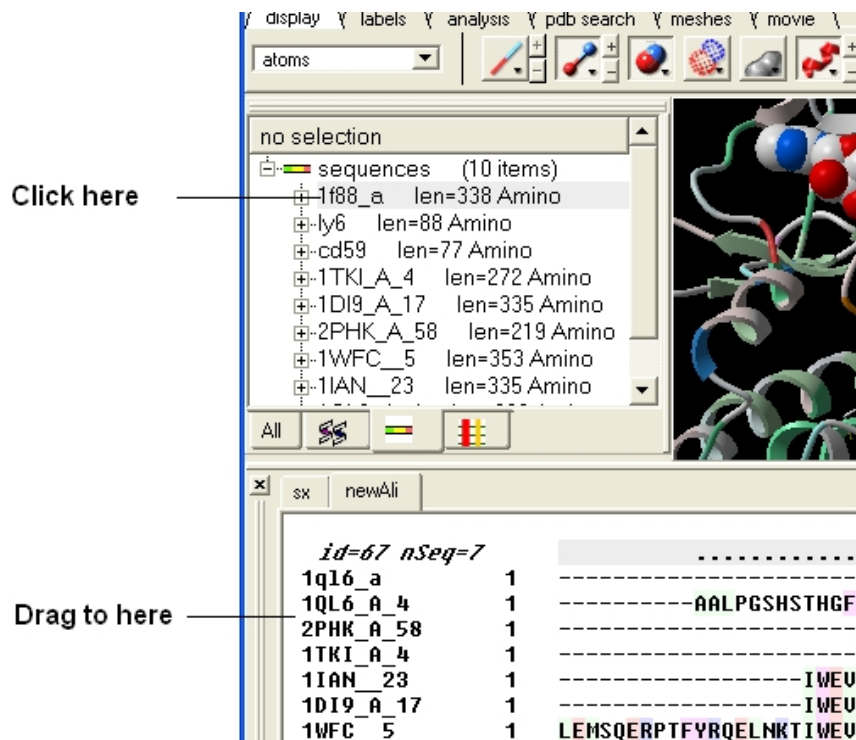
Gap Open The absolute gap penalty is calculated as a product of gapOpen and the average diagonal element of the residue comparison table You may vary gapOpen between 1.8 and 2.8 to analyze dependence of your alignment on this parameter. Lower pairwise similarity may require somewhat lower gapOpen parameter. A value of 2.4 (gapExtension=0.15) was shown to be optimal for structural similarity recognition with the Gonnet et. al.) matrix, while a value of 2.0 was optimal for the Blosum50) matrix (Abagyan and Batalov, 1997).

Gap Extension The absolute gap penalty is calculated as a product of gapExtension and the average diagonal element of the residue comparison table.



7.6.6 Drag and Drop

An easy way to add another sequence to an alignment is to drag and drop a loaded sequence from the ICM workspace panel to the alignment window. The sequence automatically becomes part of the alignment.



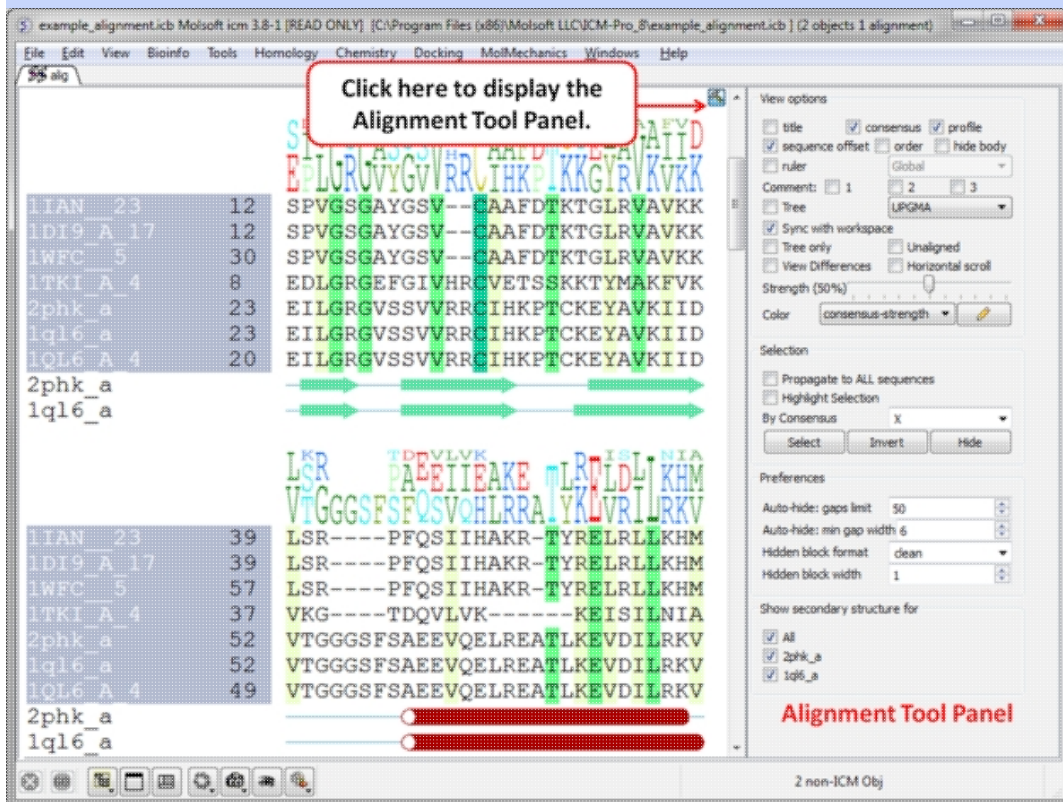
7.7 Alignment Editor

Available in the following product(s): ICM-Pro

The default position for the alignment editor is at the bottom of the graphical user interface. If you have made an alignment and you cannot see the alignment you can select Window/Alignments (See Window Menu section of this manual) and it will be displayed.

ICM has an easy to use editor for pairwise and multiple alignments. ICM alignment editor is robust and always protects the integrity of your alignment by protecting you from making unintended changes in the alignment.

NOTE: To increase or decrease the size of the font in the Alignment Editor press the CTRL key and the '+' or '-' keys.



7.7.1 Edit an Alignment

To edit an alignment one only needs four types of operations:

- select a block with one or several sequences to be moved (press Ctrl to add blocks). **Important:** since you can only move the selection **to the gapped space**, the moving front of the selection must be next to the gaps.
- (optional) create space on both sides around a vertical section of the alignment
- use the keyboard *arrows* to move the selected block with respect to the other sequences
- squeeze out the excessive gaps (an item in the alignment popup menu)

OPERATION

KEYS

set a vertical selection for ALL sequences in the alignment

Double-Click

add white space by hitting the Space bar

SpaceBar

remove white space

Backspace

select a sub-block for shifting

Drag Left-Mouse-Button

shift the selected block next to a gapped area

Right and Left Arrows

7.7.2 Save Alignment Image

To save your alignment as an image file:

- Right click on the alignment and select the Save/Export image option.
- Choose the option save as image.
- Enter the filename you wish to call your alignment. We advice you to keep the .png file extensions.
- Select the style you desire from full-width, as is, or 60.
- Select the resolution for the image. We recomend 3.0.
- Select the browse button if you wish to save the picture in a directory other than the one you are running ICM in. If you decide to change directories you will have to reenter the desired file name and click ok. The path of the file will then be entered in the save options data entry box.
- Click OK.

To save your alignment to the clipboard:

- Right click on the alignment and select the Save/Export image option.
- Choose the Save to clipboard option.

7.7.3 Save Alignment

To save an alignment or tree:

- Right click on the alignment or tree and a menu will be displayed.
- Select the save as option.
- In the windows "Save as Type" dialog box you can choose icb (ICM format), ali or FASTA.

7.7.4 Print Alignment

To print an alignment or tree:

- Right click on the alignment or tree and a menu will be displayed.
- Select the print option.

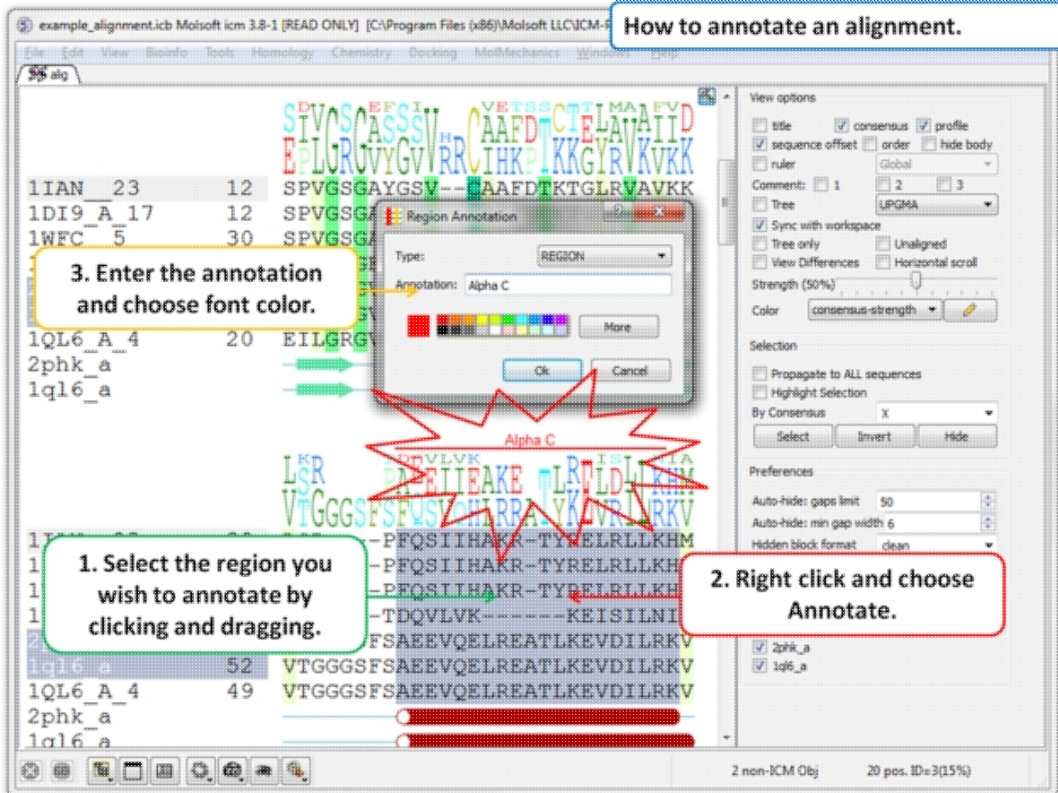
7.7.5 Delete Alignmnet

To delete an alignment or tree:

- Right click on the alignment or tree and a menu will be displayed.
- Select the delete option.

7.7.6 Annotate Alignment

To annotate an alignment:



- Click and drag over the region of the alignment you wish to annotate. The selected region will be highlighted in blue.
- Right click on the blue selected region and select 'Annotate'.
- Enter your comment text and color of font.
- Unselect by clicking away from the selection.

To Edit an annotation:

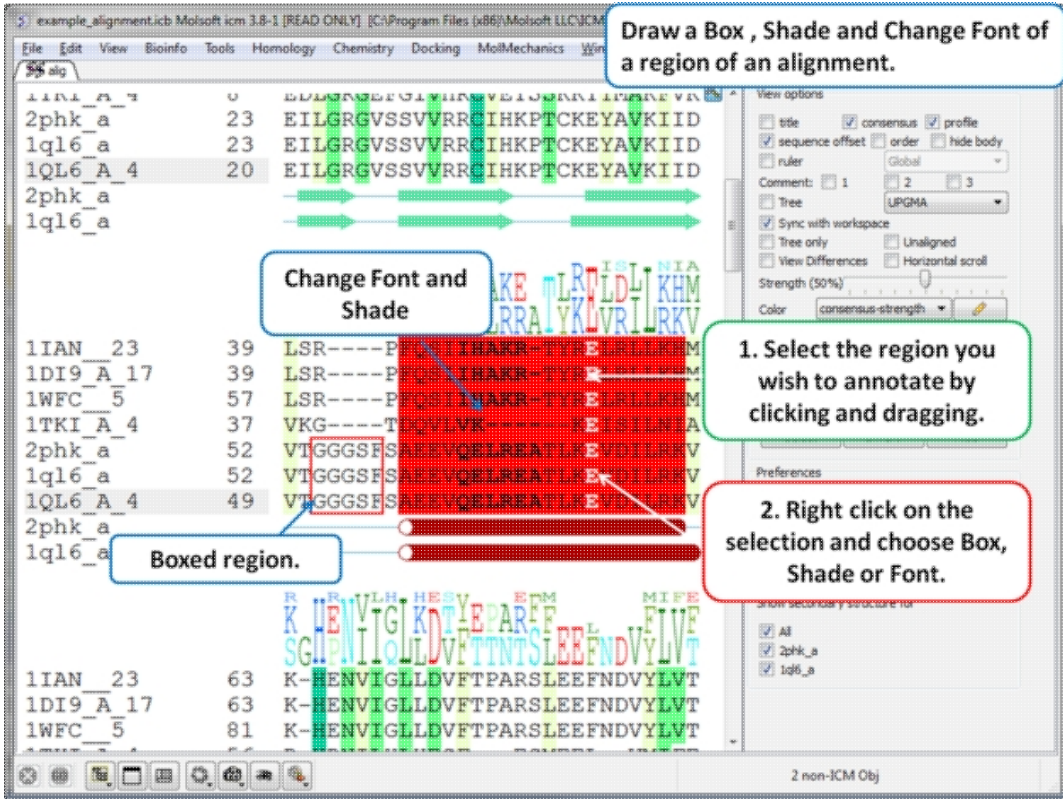
- Right click on the annotation and choose **Edit** }

To Delete an annotation:

- Right click on the annotation and choose **Delete** }

7.7.7 Box, Shade, Change Font

To draw a box, shade or change the font around a segment of the alignment:



- Click and drag over the region of the alignment you wish to box or shade. The selected region will be highlighted in blue.
- Right click on the blue selected region and select either font, shade or box.
- Unselect by clicking away from the selection.

7.7.7.1 Undisplay box, font or shading

To undisplay a box, font or shading:

- Click and drag over the boxed or shaded region of the alignment.
- Right click and choose the option "Clear Font/Color/Border".

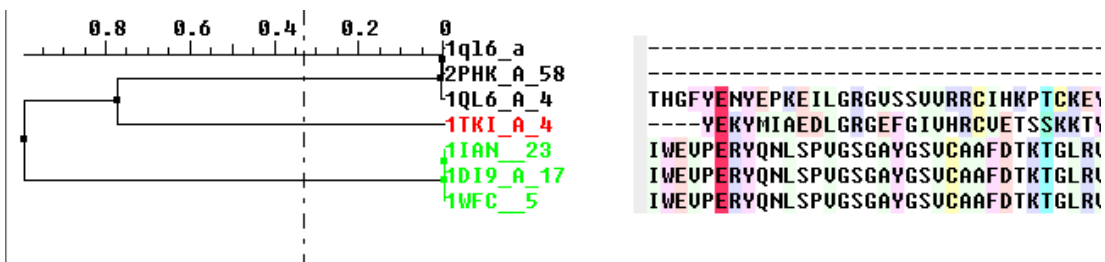
7.7.8 Phylogenetic Trees

NOTE: Before constructing a phylogenetic tree you need to align the sequences as described in the alignment section

To view a tree:

- Check the 'Tree' option in the alignment editor.

The tree will be displayed in the editor as shown below:



To display the tree alone without the alignment:

- Check the 'tree only' option in the alignment editor.

Tree functionality:

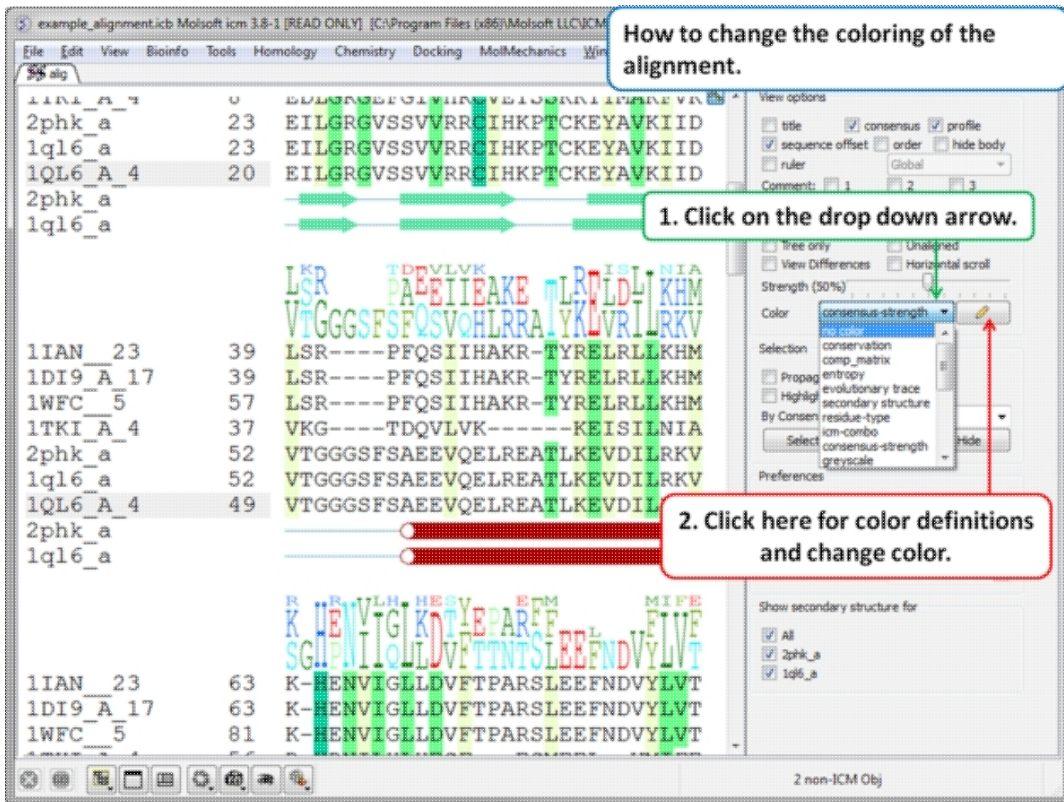
- the tree-section in the alignment is **resizeable**, just grab the rightmost end of the top ruler and drag it
- **branch swapping** : the sequences can be reordered by swapping the tree branches. Just do the following:
 - ◆ **Right-click** on a tree-node to get a popup-menu
 - ◆ **swap** the branches
- **selecting a branch**: double click on a tree-node to select the sequences belonging to it.

7.7.9 Coloring an Alignment.

How to color an alignment:

There are a wide range of default coloring schemes.

- Click on the drop down arrow beside the "Color" data entry box and a number of color schemes will be displayed.
- Select the color scheme.



NOTE: The colors are shaded from pale to bright where the brighter color represents higher conservation at that point in a multiple alignment.

To view the alignment color key:

- Click the pencil icon next to the color scheme selection tools and a table will be displayed.

To edit the alignment color key:

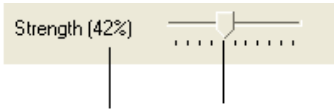
- Click the pencil icon next to the color scheme selection tools and a table will be displayed.
- Click on the "Edit Cells by Double Click" button on the right hand side of the table.

- Double click on a color panel in the table to change the color.

To color by strength of consensus:

To color your multiple alignment by the strength of consensus at each point in an alignment:

- Choose the Color option "Consensus Strength"
- Click and drag on the consensus strength button shown below:

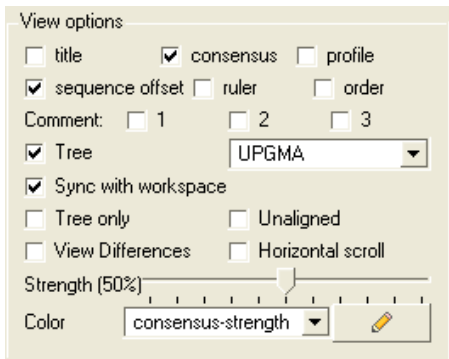


Consensus strength is displayed here

Click and drag for desired consensus strength

7.7.10 Alignment View Options

The alignment view options are located on the right hand side of the alignment editor.

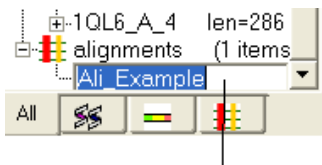


To add or remove the alignment title:

- Check the title box in the view options.

To rename an alignment:

- Right click anywhere in the alignment or on the alignment tab and a menu will be displayed.
- Select the 'Rename' option.
- Type the new name for your alignment in the data entry box which becomes activated in the ICM workspace (See below).



Rename your alignment here

To add or remove the alignment consensus display.

- Check the box labeled 'consensus' in the view options.

id=67 nSeq=7 %-%.....%g.g.....%-%...t.....%+%.....%-%... .%+E#.##L+

The alignment consensus line

If you have a large alignment it may be convenient to show the number of each sequence

To number your alignment:

- Check the 'order' box in the view options.

```
id=67 nSeq=7
1 1q16_a
2 2PHK_A_58
3 1QL6_A_4
4 1TKI_A_4
5 1IAN_23
6 1DI9_A_17
7 1WFC_5
```

└── Alignment order number
displayed here

Horizontal Scroll

To view the alignment in **Horizontal scroll** click on the "Horizontal scroll" button in the **View Options** panel in the Alignment tool bar.

To view the sequence offset number for each of your sequences in an alignment:

- Check the 'offset' box in the view options.

```
      ..K.
1q16_a      106 TEK
2PHK_A_58   46  TEK
1QL6_A_4    113 TEK
1TKI_A_4    91  TSAF
1IAN_23     103 CQK-
1DI9_A_17   103 CQK-
1WFC_5      121 CQK-
```

└── Sequence offset number

To view the sequence ruler:

- Check the 'ruler' box in the view options.

```
-----IWEUPERYQNLSPUGSGAYGSUCAAF
-----IWEUPERYQNLSPUGSGAYGSUCAAF
LEMSQERPTFYRQELNKT IWEUPERYQNLSPUGSGAYGSUCAAF
0          10          20          30          40
```

└── Alignment ruler

To view the sequence consensus.

- Check the consensus option in the view options.

```
·R·V·L·H·E·S·
·E·N·V·I·G·L·K·D·T·
·P·I·Q·L·L·D·V·
·H·E·N·V·I·G·L·L·D·V·
·H·E·N·V·I·G·L·L·D·V·
·H·E·N·V·I·G·L·L·D·V·
```

To view secondary structure.

If one of the sequences of the alignment is linked to a structure then you can display the secondary structure by:

- Check the "show secondary structure for" box.

The secondary structure will be displayed at the bottom of the alignment. The color code is: red cones = alpha helix, green arrows = beta sheet, blue cones = pi helix, magenta cones = 3/10 helix.

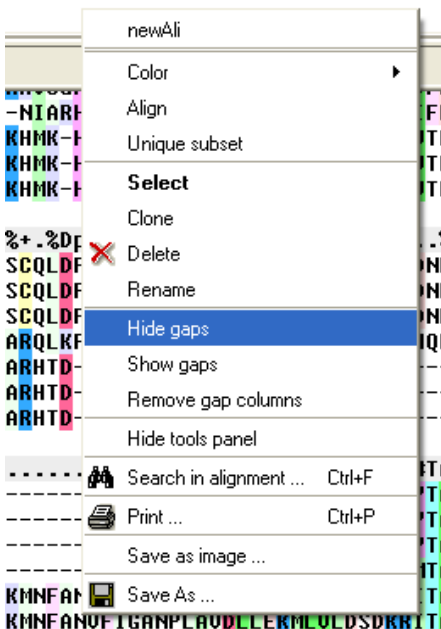


7.7.11 Alignment Gaps

To make an alignment clearer you may wish to HIDE gap regions.

To hide all gap regions:

- Right click on the alignment and a menu as shown below will be displayed.



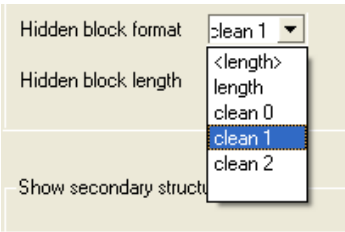
- Select the "Hide gaps" option.

The gaps in your alignment will be hidden according to the preference made in the alignment tools panel shown below. Click on the drop down arrow in the "Hidden block format" data entry box.

Two parameters can be specified directly from the Tools Panel in the alignment window:

1. the "Hidden Block Format" can use the following special symbols:
 - ◆ %l number of hidden chars
 - ◆ %L length of the hidden block
 - ◆ %f hidden from
 - ◆ %t hidden to

- e.g. " %f . . %t " or " %L "
- the "Hidden block width" which defines the total length of the hidden section.



Some predefined hidden block formats are shown here:

length: displays the length of the gap

```
S . . 27 #%D
SDT   UKD
SDT   UKD
SDT   UKD
SIE   AMD
SSE 27 AUD
SSE 27 AUD
SSE 27 AUD
```

length: displays the length of the gap in .

```
IS . . < 27 >#%D
'SDT<   >UKD
'SDT<   >UKD
'SDT<   >UKD
SIE<    >AMD
SSE< 27 >AUD
SSE< 27 >AUD
SSE< 27 >AUD
```

clean0: displays no indication of a gap

```
IS . . #%D
'SDTUKD
'SDTUKD
'SDTUKD
SIEAMD
SSEAUD
SSEAUD
SSEAUD
```

clean1: displays grey panel in the gap position.

```
IS . . #%DI
'SDT UKDI
'SDT UKDI
'SDT UKDI
SIE AMDI
SSE AUDI
SSE AUDI
SSE AUDI
```

clean2: displays a wider grey panel in the gap position

```
IS . . #%DI
'SDT UKDI
'SDT UKDI
'SDT UKDI
SIE AMDI
SSE AUDI
SSE AUDI
SSE AUDI
```

NOTE: The width of the hidden panel can be changed as shown below.

Hidden block length 3

Enter length of hidden block here

If you have hidden all the gaps individual gaps (or blocks) can be displayed by:

- Right clicking on the gap and select "Show hidden block" option.

```
IS..  #DL#.#LU#...KRI
'SDT  UKDLUSRFLUQPQR
'SDT
'SDT  Show hidden block
SIE   AMDFUDRLLUKERKSR
SSE   AUDLLEKMLULDSDKR
SSE   AUDLLEKMLULDSDKR
SSE   AUDLLEKMLULDSDKR
```

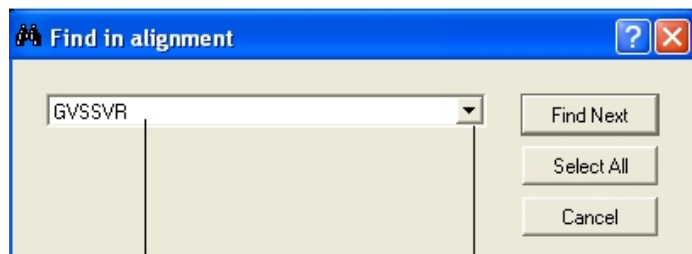
To show all gaps:

- Right click on the alignment away from a gap region and a menu will be displayed.
- Select the "Show gaps option".

7.7.12 Searching an Alignment

If you have a large alignment and you wish to find a specific group of amino acids within that residue the you can use the Alignment search tool.

- Right click on the alignment away from any hidden gaps and a menu will be displayed.
- Select the "Search in alignment option" and a data entry box as shown below will be displayed.



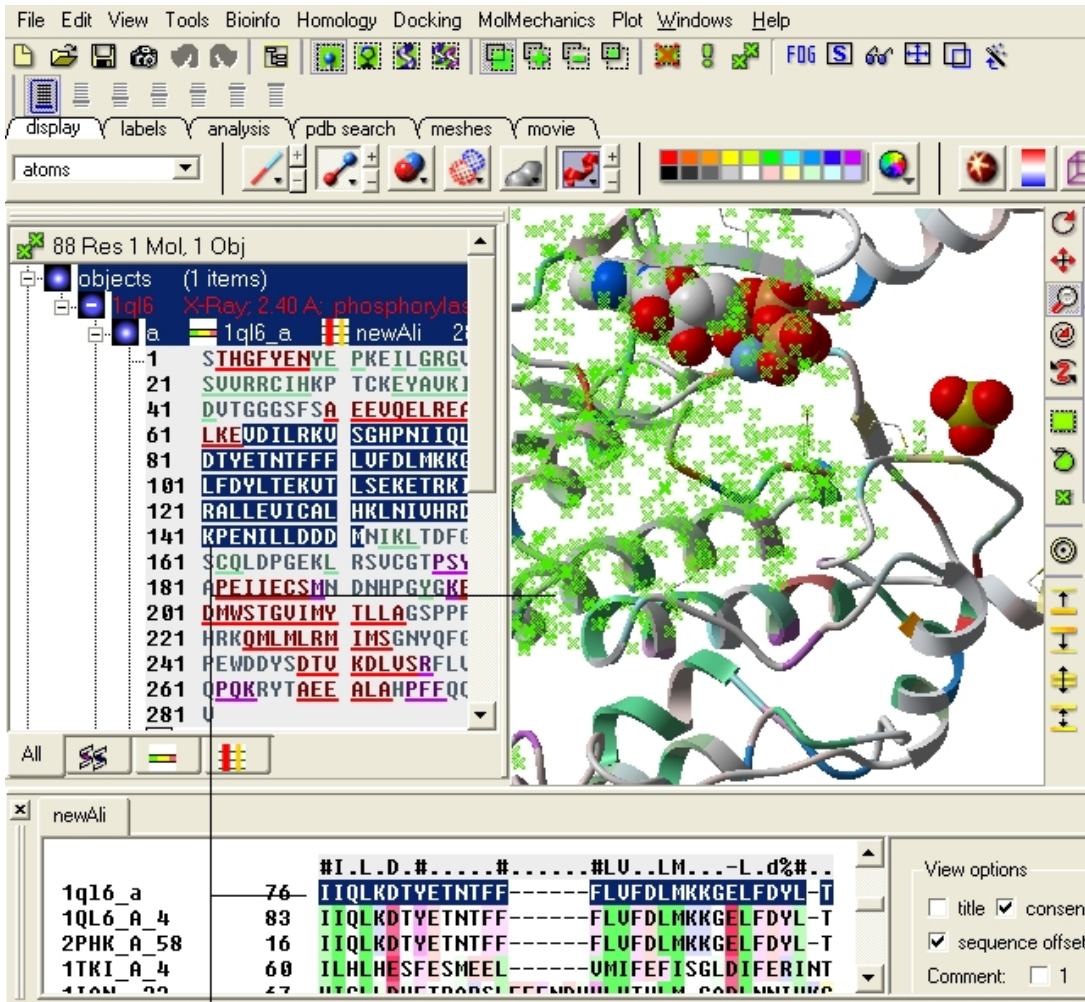
Type search sequence here

Previous search queries can be found here

NOTE: Another way of searching an alignment is to use the alignment selection tools which are linked to the ICM workspace and 3D graphical window. This is described in the section entitled Making Selections in Alignments.

7.7.13 Making Alignment Selections

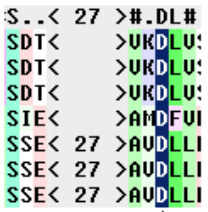
ICM has a very powerful alignment selection tool which enables sequences and structures to be interlinked with the 3D graphical window, the alignment window and the ICM workspace.



Selection is displayed in the ICM workspace (blue) the alignment window (blue) and the 3D graphics display (green cross).

To select a single column of an alignment:

- Double click.



One column selected by double clicking

To select parts of an alignment:

- Click and drag over the region you wish to select.

```

S..< 27 >#.DL#.#+#LU!
SDT< >UKDLUSRFLU!
SDT< >UKDLUSRFLU!
SDT< >UKDLUSRFLU!
SIE< >AMDFUDRLLU!
SSE< 27 >AUDLLEKMLU!
SSE< 27 >AUDLLEKMLU!
SSE< 27 >AUDLLEKMLU!

```

Click and drag over to select

To select multiple discontinuous parts of an alignment:

- Click and drag whilst holding down the control key.

```

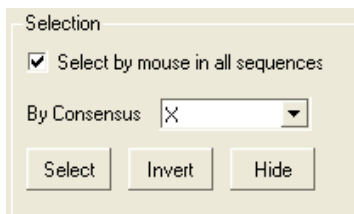
#.DL#.#+#LU%...KR#TA..ALAH#
UKDLUSRFLUQQPKRYTAEELAHPI
UKDLUSRFLUQQPKRYTAEELAHPI
UKDLUSRFLUQQPKRYTAEELAHPI
AMDFUDRLLUKERKSRMTASEALQHP!
AUDLLEKMLULDSDKRITAAQALAH!
AUDLLEKMLULDSDKRITAAQALAH!
AUDLLEKMLULDSDKRITAAQALAH!

```

Click and drag whilst holding down the control key.

To enable the easy selection of all sequences in an alignment:

- Check the box labeled "Select by mouse in all sequences".



```

S..< 27 >#.DL#.#+#LU%.
SDT< >UKDLUSRFLUUC
SDT< >UKDLUSRFLUUC
SDT< >UKDLUSRFLUUC
SIE< >AMDFUDRLLUKE
SSE< 27 >AUDLLEKMLULI
SSE< 27 >AUDLLEKMLULI
SSE< 27 >AUDLLEKMLULI

```

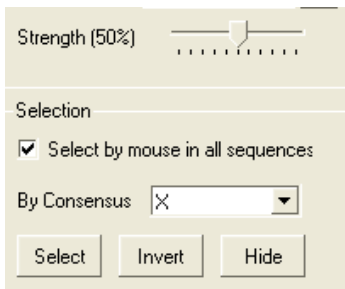
All sequences are selected using the mouse

NOTE: All selections made in the alignment window are linked to the 3D graphics window and the ICM workspace if a structure is in the alignment.

7.7.14 Select by Consensus

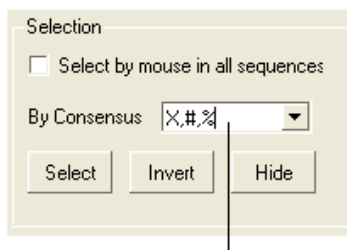
This is a very useful tool, for example, you may want to color the conserved regions of your structure in the 3D display window a different color to the rest of the structure. This tool allows you to select the conserved regions in the sequence alignment. Once the selection has been made it can be used for a number of different ICM operations such as coloring and displaying secondary structure.

A selection can be made based on the alignment consensus. The buttons relating to this are in the alignment tool panel.



Before selecting by consensus you first need to define a consensus strength:

- Click and drag on the bar labelled "Strength" and select your desired percentage.
- Enter which elements of the consensus you wish to select separated by comma. Refer to the language manual for definition of each consensus symbol.



Enter which elements of the consensus you wish to select seperated by a comma.

- Click the Select button and your selection will be highlighted in blue in the alignment window and ICM workspace and as green crosses in the 3D graphical display window.

Once the selection has been made it can be used for a number of different ICM operations such as coloring and displaying secondary structure.

To invert a selection:

- Click on the invert button.

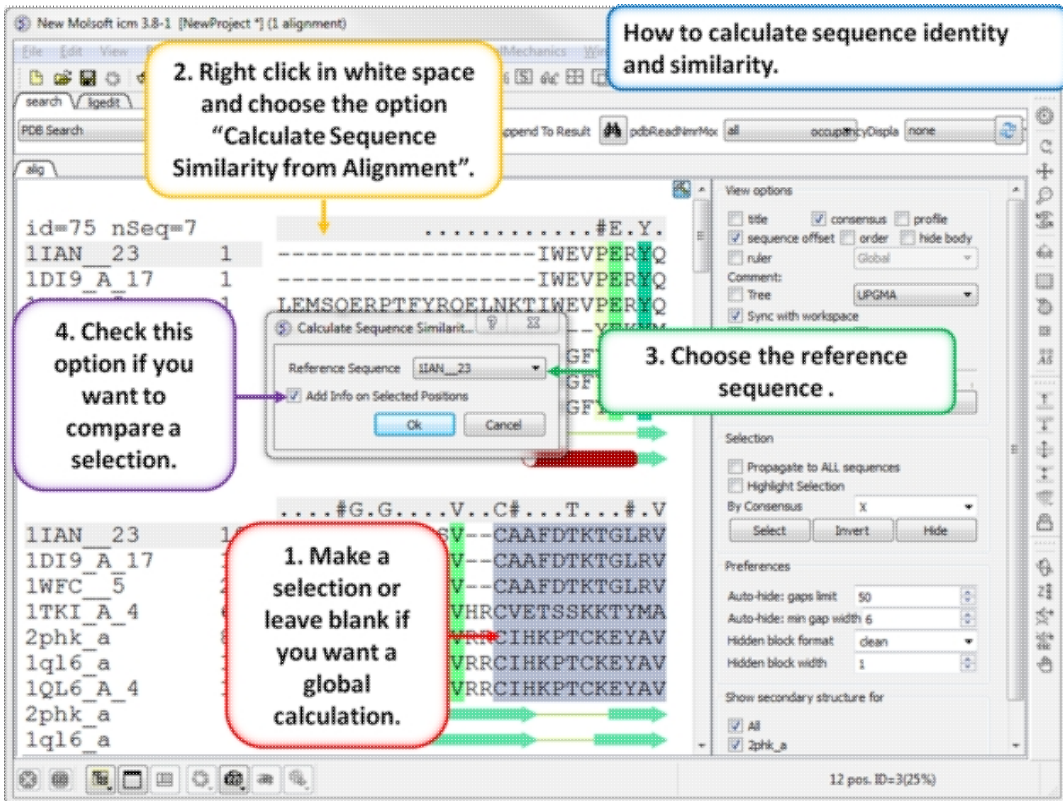
To hide a selection.

- Click on the Hide button

NOTE: All selections made in the alignment window are linked to the 3D graphics window and the ICM workspace if a structure is in the alignment.

7.7.15 Calculate Sequence Identitiy and Similarity

To calculate the sequence identity and similarity between one sequence and other sequences in an alignment:



- Make an alignment.
- If you wish to calculate local sequence identity you need to select the region of interest.
- Right click in white space in the alignment and choose the option "Calculate Sequence Similarity from Alignment".
- Choose the reference sequence from the dialog box.
- If you wish to calculate local sequence identity you need to select the option "Add info on Selected Positions".

SeqSimilarity Molsoft icm 3.8-1 [NewProject *] (1 table alignment)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics

search All Fields Append To Result

align

.....#G.G....V..C#...T...#.V

```

1 IAN_23      10  NLSPVGS1GAYGSV2--CAAFD3T4TKT5GLRV6
1 DI9_A_17   10  NLSPVGS1GAYGSV2--CAAFD3T4TKT5GLRV6
1 WFC_5      28  NLSPVGS1GAYGSV2--CAAFD3T4TKT5GLRV6
1 TKI_A_4    6   IAEDLGRGE1FGIVHRCVET2SSK3KTYMA4
2 phk_a      8   PKEILGRGV1SSV2VRR3CIHKPT4CKEYAV5
1 ql6_a     11  PKEILGRGV1SSV2VRR3CIHKPT4CKEYAV5
1 QL6_A_4   18  PKEILGRGV1SSV2VRR3CIHKPT4CKEYAV5
2 phk_a
1 ql6_a

```

View options

- title
- consensus
- profile
- sequence offset
- order
- hide body
- ruler
- Global
- Comment:
- Tree
- UPGMA
- Sync with workspace
- Tree only
- Unaligned
- View Differences
- Horizontal scroll
- Strength (50%)
- Color consensus-strength

Selection

- Propagate to ALL sequences
- Highlight Selection
- By Consensus X

SeqSimilarity

Ref	Seq	Ident	Sim	i	IdSel	SimSel	p43	p44	p45	p46	p47	p48	p49	p50	p51	p52	p53	p54
1	IAN_23	100	99.97	1	100	100	C23	A24	A25	F26	D27	T28	K29	T30	G31	L32	R33	V34
2	IAN_23	100	99.97	2	100	100	C23	A24	A25	F26	D27	T28	K29	T30	G31	L32	R33	V34
3	IAN_23	100	99.97	3	100	100	C41	A42	A43	F44	D45	T46	K47	T48	G49	L50	R51	V52
4	IAN_23	1TKI_A_4	25.73	30.49	4	16.66	C21	R22	E23	T24	S25	S26	K27	K28	T29	F30	K31	A32
5	IAN_23	2phk_a	27.43	33.6	5	25	C23	E24	H25	K26	P27	T28	C29	K30	R31	Y32	A33	V34
6	IAN_23	1ql6_a	27.04	32.95	6	25	C26	I27	H28	K29	P30	T31	C32	K33	E34	Y35	A36	V37
7	IAN_23	1QL6_A_4	26.22	31.92	7	25	C33	I34	H35	K36	P37	T38	C39	K40	E41	Y42	A43	V44

12 pos. ID=3(25%)

A table containing the sequence identity and similarity for the whole sequence and selected sequences will be displayed.

A table of sequence similarity and identity will be displayed. A guide to the columns in the table:

- **Ref** = name of reference sequence.
- **Seq** = name of comparison sequence.
- **Ident** = identity percentage
- **Sim** = similarity percentage
- **i** = order
- **IDSel** = identity percentage for selection
- **SimSel** = similarity percentage for selection
- **p..** = residues in selection.

7.7.16 Alignment Secondary Structure

If the one or more of the sequences contain secondary structure annotation then you will see colored cylinders and arrows under the alignment. Red Cylinder = alpha helix, green arrow = beta sheet, blue cylinder = pi helix, magenta cylinder = 3/10 helix.

The colored cylinders and arrows under the sequence alignment refer to secondary structure.

1q16_a
..#L..##H.#.I.HRDLKP.N#.D..#K#.DFG.....
YQILRGLKYI SADIIHRDLKPSNL--AVNEDCELKILDFGLARH--T
1DI9_A_17 YQILRGLKYI SADIIHRDLKPSNL--AVNEDCELKILDFGLARH--T
1WFC_5 YQILRGLKYI SADIIHRDLKPSNL--AVNEDCELKILDFGLARH--T
1TKI_A_4 HQVCEALQFLSHNIGHFDIRFENIIYQTRRSSTIKIIIEFGQARQLKP
2phk_a RALLEVICALKLNIVHRDLKPSNILEN--LDDMNIKLTDFGFSCQLDP
1q16_a RALLEVICALKLNIVHRDLKPSNILEN--LDDMNIKLTDFGFSCQLDP
1QL6_A_4 RALLEVICALENI VHRDLKPSNILEN--LDDMNIKLTDFGFSCQLDP
2phk_a
1q16_a
.-.##.###.T...#...G...IM#.LL.
1IAN_23 DDEMTGYVATRWYRAPEIMLNWMH-----NQTVDIWSVGCIMAELLT
1DI9_A_17 DDEMTGYVATRWYRAPEIMLNWMH-----NQTVDIWSVGCIMAELLT
1WFC_5 DDEMTGYVATRWYRAPEIMLNWMH-----NQTVDIWSVGCIMAELLT
1TKI_A_4 GDNFRLFLTPEY YAPEVHQHDV-----VSTATDMWSLGLTVYVLLS
2phk_a GEKLRVCGTPSYLAPEIECSMNDNHPG GKEVDMWSTGVIMYTLLA
1q16_a GEKLRVCGTPSYLAPEIECSMNDNHPG GKEVDMWSTGVIMYTLLA
1QL6_A_4 GEKLRVCGTPSYLAPEIECSMNDNHPG GKEVDMWSTGVIMYTLLA
2phk_a

Use the alignment tool panel to toggle the secondary structure display on and off.

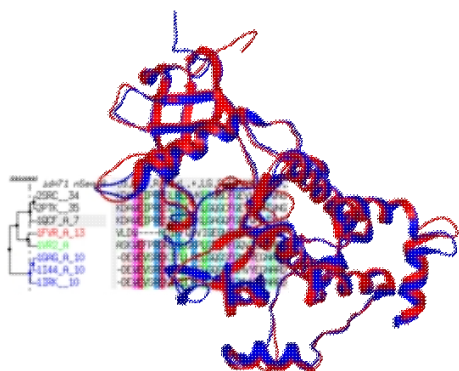
To toggle secondary structure display on and off use the options in the alignment tools panel.

1. Click here to see the tools panel.

2. Toggle on and off

1q16_a
+E#.#L..##.H.N#I.L.D.#....
1IAN_23 RELRLKHKM--ENVIGLLDVFTPAR
1DI9_A_17 RELRLKHKM--ENVIGLLDVFTPAR
1WFC_5 RELRLKHKM--ENVIGLLDVFTPAR
1TKI_A_4 KEISILNIAR--RNILHLESF---E
2phk_a KEVDILRKVSGFPNIIQLKDTYETNT
1q16_a KEVDILRKVSGFPNIIQLKDTYETNT
1QL6_A_4 KEVDILRKVSGFPNIIQLKDTYETNT
2phk_a
1q16_a
.#.....#LV..LM...-L...#..
1IAN_23 SLEEFNDVYLVTHLMG-ADLNIVK
1DI9_A_17 SLEEFNDVYLVTHLMG-ADLNIVK
1WFC_5 SLEEFNDVYLVTHLMG-ADLNIVK
1TKI_A_4 SMEEL--VMIFEFISG-LDIFERINT
2phk_a FF-----F
1q16_a FF-----F
1QL6_A_4 FF-----FLVFDLMMKKGELFDYLTE
2phk_a

8 Protein Modeling



In this Chapter:

- Homology Modeling
- Interactive Modeling
- Loop Modeling
- Design Loop
- Graft Loop
- Regularization
- Refine Side Chains
- Making a Disulfide Bond
- Make a Mutation
- Effect of Mutation on Protein Binding
- Effect of Mutation on Protein Stability
- Protein Sculpting
- Molecular Mechanics Menu

8.1 Homology Modeling

Available in the following product(s): Homology

8.1.1 Homology Modeling Introduction

8.1.1.1 Modeling Options

There are three main modeling options to choose from:

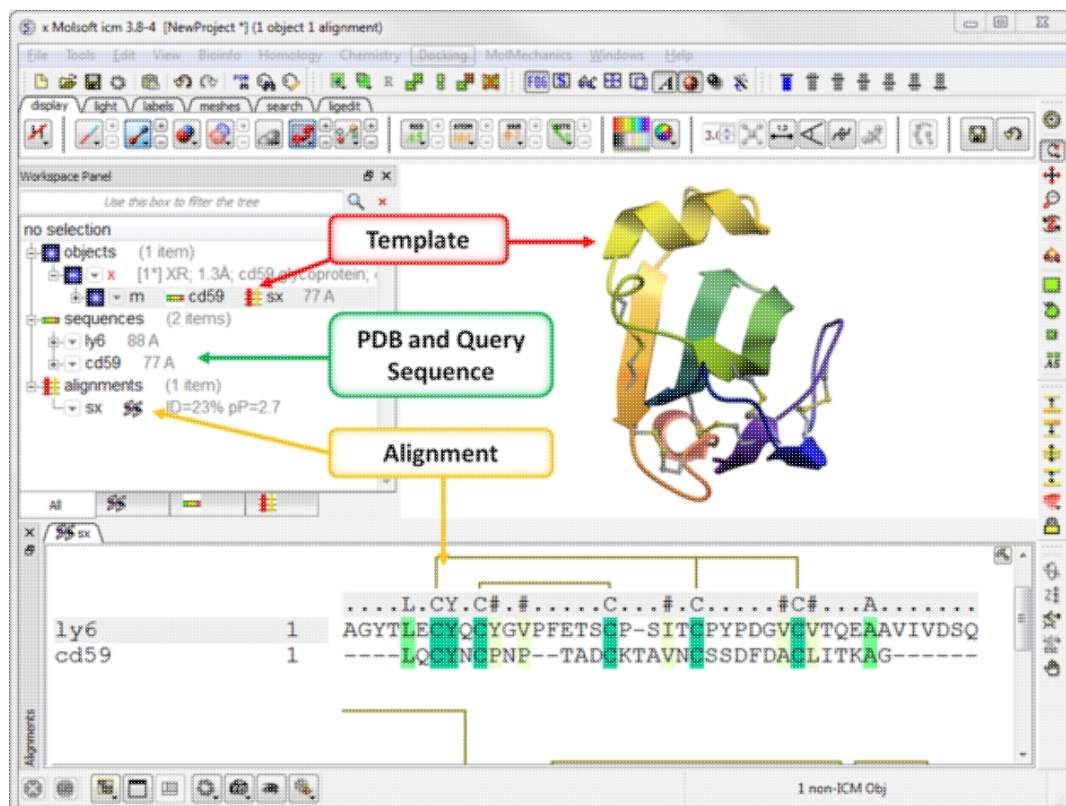
- **Quick Model.** This option will provide a model by placement of the aligned polypeptide chain onto the template structure. The method is fast and is ideal if you want a quick model for visualization purposes.
- **Full Model Builder.** This option will give you a fully refined model and can take many hours to complete the model building process. There are options to fully refine side-chains, refine side-chains and loops or undertake a full refinement. The refinement is undertaken using the ICM BPMC modeling method in the ICM force field which includes surface terms, electrostatics with the boundary element solution of the Poisson equation and side chain entropy terms.
- **Multi-Template Model Editor.** This option is ideal if you want more control over your model - particularly in loop regions or if you want to use multiple templates. It provides an interface to add new templates to a model and browse through multiple conformations of loop regions. It also provides an interactive visualization interface to see how well the residues in your model falls into the Ramachandran Plot and how good the omega angles are.

8.1.1.2 Homology Modeling Getting Started

The items you need to begin making an Homology Model are:

1. A template structure from the PDB converted into an ICM object.
2. Extract the sequence of the template structure from the PDB file.
3. The sequence of your query molecule (e.g. from UniProt or FASTA).
4. An alignment (see alignment section) of your query sequence against the template sequence from the PDB.

Your graphical user interface window should look something like this:



8.1.1.3 Older Versions of Homology Modeling in ICM

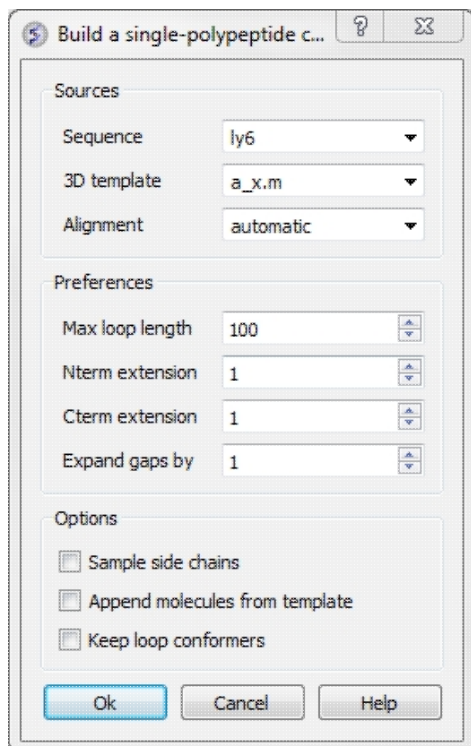
If you are using a version of ICM 3.8-3 or earlier you will see different options in the Homology Menu. We recommend updating your ICM version to 3.8-4 or newer at www.molsoft.com/support If you cannot update your ICM version then Homology/Build model is similar to Quick Model, Homology

8.1.2 Quick Model

8.1.2.1 Single Chain from Sequence

To build a 'Quick' homology model from a single chain:

- Check that you have your query sequence, template structure and alignment loaded in ICM.
- Click on the 'Homology' menu at the top of the graphical user interface.
- Select Quick Model and a dialog box will be displayed as shown below.



This data entry box is split into 3 sections, the first is 'sources' where you need to specify your query sequence, template and alignment. The second section is called 'preferences' where penalty information for the model needs to be entered and the third section is called 'Options'. Each will be described in detail below.

To construct your model follow these steps:

- Enter the name of your QUERY (ie the sequence of the model you wish to build - NOT the template) sequence in the data entry box labeled **Sequence**. If you click on the arrow next to this box a list of sequences loaded in ICM will be displayed. The names of the sequences are also listed in the workspace panel on the left of the graphical user interface.
- Enter the name of your template structure in the '3D template' data entry box. Once again the name of your template structure can be found by clicking the down arrow or in the workspace panel.
- Enter the name of your alignment in the 'Alignment' data entry box.

You could build your model now as ICM has enough information but it may be wise to take a look at some of the preferences and change them accordingly. However in most cases the default values provided are sufficient to produce a good quality model.

To change the preferences either type the number you wish or use the up and down arrows next to the data entry boxes.

- **Max loop length** (default= 999) - loops longer than this value are not modeled
- **Nterm extension** (default=1) - the maximal length of the N-terminal model sequence which extends beyond the template
- **Cterm extension** (default=1) - the maximal length of the C-terminal model sequence which extends beyond the template
- **Expand gaps by** (default=1) - additional widening of the gaps in the alignment. End gaps are not expanded

Now all you need to do to build your model is to select some options. The options are:

- **Sample side chains** - performs monte-carlo optimization on the side chains
- **Append molecules from template** - this option will transfer all hetatoms and molecules in the template structure to your model.
- **Keep loop conformers** - this option will save alternative loop conformations in the ICM workspace.

To build your model:

- Click OK

Once your model is built a new object containing your model will be available in the ICM workspace panel (left hand side).

8.1.2.2 Multi-Chain from Linked Alignment

To make a multi-chain model you must first link all the chains to a template via an alignment. For example, in an antibody you have light chains and heavy chains so you would make two alignments with the chains in the template structure.

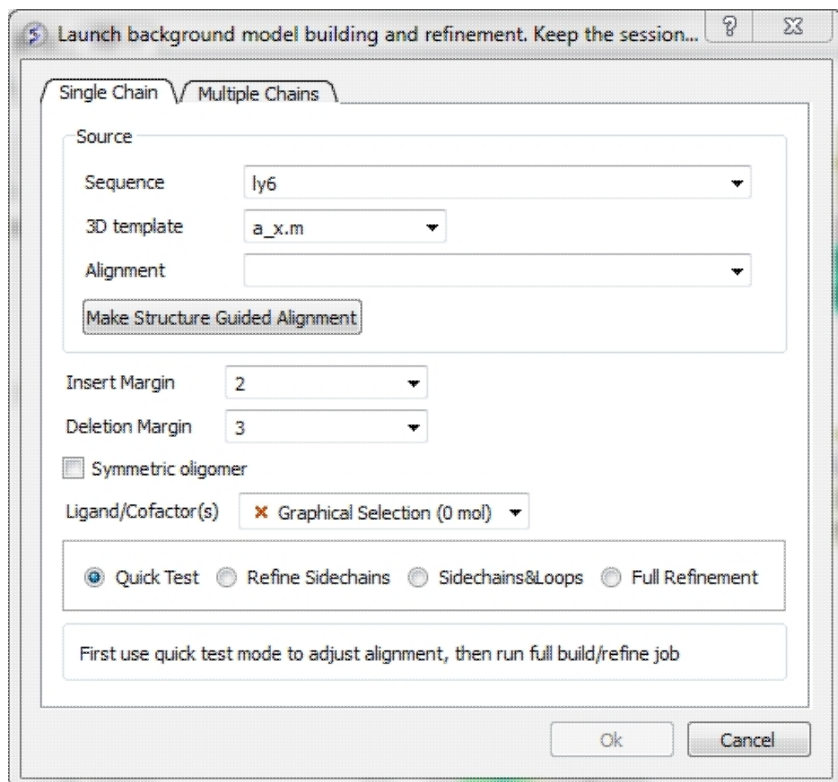
To do this:

- Read in the template PDB structure.
- Read in the query sequence.
- Extract the sequence from the chains in the PDB you wish to model.
- Make alignments of the query and chain sequences.
- Go to Homology/Quick Model/Multi-Chain from Linked Alignment
- Enter the name of the template structure.
- Enter a name for the model you are going to build .
- Click "sample side-chains" if required.
- When the model is ready it will be displayed in the ICM Workspace.

8.1.3 Full Model Builder

The **Full Model Builder** will give you a fully refined model and can take many hours to complete the model building process. There are options to fully refine side-chains, refine side-chains and loops or undertake a full refinement. The refinement is undertaken using the ICM BPMC modeling method in the ICM force field which includes surface terms, electrostatics with the boundary element solution of the Poisson equation and side chain entropy terms.

- Check that you have your query sequence, template structure and alignment loaded in ICM.
- Click on the 'Homology' menu at the top of the graphical user interface.
- Select **Full Model Builder** and a dialog box will be displayed as shown below.



- Enter the name of your QUERY (ie the sequence of the model you wish to build - NOT the template) sequence in the data entry box labeled **Sequence**. If you click on the arrow next to this box a list of sequences loaded in ICM will be displayed. The names of the sequences are also listed in the workspace panel on the left of the graphical user interface.
- Enter the name of your template structure in the '3D template' data entry box. Once again the name of your template structure can be found by clicking the down arrow or in the workspace panel.
- Enter the name of your alignment in the 'Alignment' data entry box.
- Enter a value for **Insert Margin**. Insert margin is the number of residues that will be untethered around an insertion in the modeled sequence in sequence-structure alignment.
- Enter a value for **Deletion Margin**. Delete margin is the number of residues around the deletion in the query (insertion in the template) that are untethered.
- Select oligomeric if you want to build a model of an oligomer provided in the template structure.
- You can choose to include Ligands and Co-factors by selecting them using the orange selections
- First use **Quick Test** to check the model and make any necessary adjustments to the alignment. When you are happy proceed to refine side chains, or refine side-chains and loops, or full refinement.

8.1.3.1 Full Model of Multi-Chain from Linked Alignment

To make a multi-chain model you must first link all the chains to a template via an alignment. For example, in an antibody you have light chains and heavy chains so you would make two alignments with the chains in the template structure.

To do this:

- Read in the template PDB structure.
- Convert the PDB file into an ICM object.
- Read in the query sequence.
- Extract the sequence from the chains in the PDB you wish to model.
- Make alignments of the query and chain sequences.
- Go to Homology/Full Model Builder
- Click on the **Multiple Chain** tab.
- Use the check boxes in the sequence drop down button to select multiple sequences for the chains you are modeling.
- Use the check boxes in the alignment drop down button to select the sequence alignment for each chain.

8.1.4 Multi-Template Model Editor

Available in the following product(s): ICM-Homology

What you need before you can use the **Multi-Template Model Editor**:

Build a New Model

- A template structure from the PDB converted (See `convert` object) into an ICM object.
- A sequence file of your query sequence.

You can use an alignment you have constructed yourself or allow ICM to generate one (referred to as automatic)

- An alignment (see `alignment` section) of your query sequence against the template structure from the PDB. This is a *.ali file in ICM.
- Homology/Multi-Template Model Editor
- Enter the name of your loaded sequence.
- Enter the name of your loaded 3D template
- Enter the name of your alignment or allow ICM to generate an alignment by selecting automatic
- Check whether you wish ICM to sample the loop regions of your model or not

An Interactive Modeler's View of the alignment will then be displayed.

Pre-Existing Model

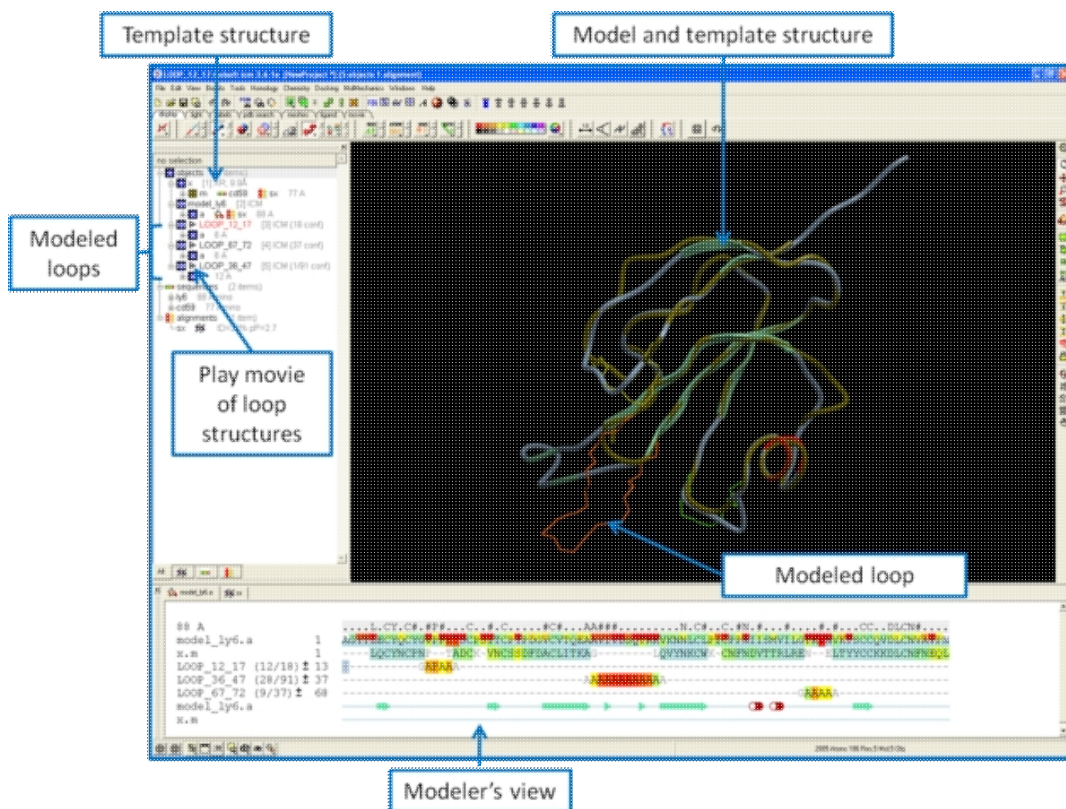
If you have built a model using Quick Model or the Full Model Builder you can edit this model in the editor by:

- Homology/Multi-Template Model Editor/Setup from Existing Model
- Enter the name of the Model and The Template you wish to use for multi-template modeling.

A Interactive Modeler's View of the alignment will then be displayed.

8.1.4.1 Modeler's View

Once you have made your interactive model your graphical user interface should look something like this:



What do all the elements of the **Modeler's View** mean?

Number of residues in the model structure

3D model sequence:
Top color represents how well the residue falls into the Ramachandran Plot.
Bottom color represents the omega angles. Red (poor) – Yellow – Blue (good)

Secondary structure of template and model

Loops:
 Use the +/- button to browse through different loop conformations – keep an eye on the coloring of the sequence. Red (poor omega angles) – yellow – blue (good omega angles)

3D template sequence:
 Colored from Red (poor) - Yellow – Red – Blue (good) depending on how well the model follows the template structure.

8.1.4.2 Interactive Loop Modeling

You can browse a number of different loop conformations by clicking on the +/- button next to the loop in the `modeller view`'s window. Keep an eye on the colors of the loop residues in the alignment - they are colored red (poor) -yellow - blue (good) depending on how good the omega angles are. The residues in the loops are represented by Alanines for every residue other than Glycine or Proline.

If you want to remodel the loop.

- Set tethers at the residue at the start and end of the loop. To set a tether double click on the two residues whilst holding down the CTRL key. Right click and select **Set tethers**. Select how you want to tether to the template either by **alignment**, **residue numbering** or **selection**. Select the **alignment** - automatic represents the **Modeller's View**. Select the **Template Molecule** which is the object you built the model on.
- Select the loop you wish to model including the tethered residues at the start and end of the loop and then right click and select **Create Loop Template**
- Once the simulation has finished you can once again browse the solutions by clicking on the +/- button next to the loop in the `modeller view`'s window.
- When you have identified a reasonable conformation of the loop you can re-thread it onto the model structure by selecting the loop region in the **Modeller's View**. You can do this by clicking and dragging over the loop. Right click and select **Thread Through Tethers**.

1. Double click on the first residue of the loop - hold the CTRL - key and double click on the end residue.

2. Right click on one of the selected residues and select "Set Tethers..."

3. Right click again on the selected residues and select "Create Loop Template ..."

4. Browse the loops. When you find a reasonable model of the loop – select the loop – right click - and then choose the option "Thread Through Tethers..."

8.1.4.3 Multi Template Model

To use more than one template for your model.

- First make an initial model and setup the modeler's view alignment (Homology/Multi-Template Model Editor).
- Make an alignment between the query sequence and the new template.
- Select in the alignment (right click and drag) where you would like to use the additional template.
- Right click on the alignment selection and choose "Set Tethers". Select the alignment and then select the template object from the drop down dialog boxes. Click OK.
- Right click again on the alignment and choose "Thread Through Tethers" and the model will be built.

8.2 Loop Modeling

Available in the following product(s): Homology

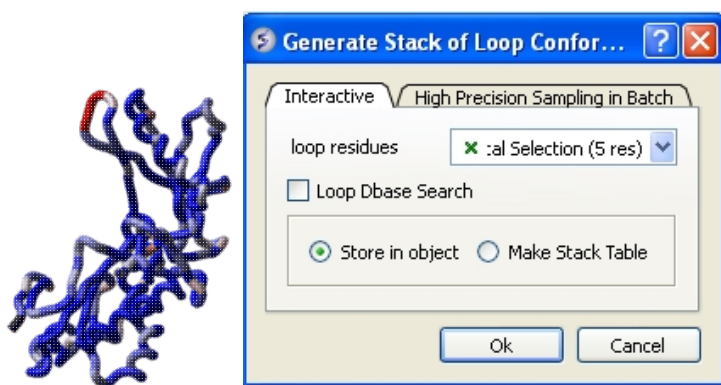
Other resources: [Tutorial](#)

Background

Building an accurate model of a loop is very tough. However with small loops ICM has been very succesful. ICM was used to design two new 7 residue loops and in both cases the designs were successful. Moreover, the predicted conformations turned out to be exactly right (accuracy of 0.5Å RMSD) after the crystallographic structures of the designed proteins were determined in Rik Wierenga's lab (Thanki et al Protein Eng. 1997 10, 159-167). In 2010, scientists at MolSoft developed a new physics-based internal coordinate mechanics force field (Arnautova et al Proteins. 2010, 79 : 477) and was tested on loop structures. The new force-field contains new parametrization for the dielectric constant, an improved hydrogen bond determination method, and implementation of novel backbone atom torsional potentials which include bond angles of the carbon (alpha) atoms into the internal variable set.

There are two loop modeling options:

1. **Interactive** This option will re-build a loop by searching icm.lps contains loop conformations from the PDB database. It searches for suitable loops with matching loop ends and as close loop sequence as possible, inserts them into the model and modifies the side-chains according to the model sequence.
2. **High Precision Sampling in Batch** This option does a full simulation of the loop in the latest ICM force-field as benchmarked in this paper Arnautova et al Proteins. 2010, 79 : 477. This is a more rigorous simulation than the interactive approach and therefore takes longer to run (see [Tutorial](#) here).



Interactive - fast loop sampling:

- Read your modeled structure into ICM. Or continue immediately after using build model.
- Select the loop region you wish to model (green crosses in the graphical display). Make sure the object you are selecting is the **current object**. You can make an object current by right clicking on the object in the ICM Workspace and select "Set to Current"
- MolMechanics/Sample Loop.
- Select the tab labeled **Interactive**
- You should see graphical selection in the "loop residue" panel or you can enter the selection here using the icm selection language.

- If you want to perform a fast knowledge-based loop search (searching a database of loops) check **Loop Dbase Search**. If you would like to perform an *ab initio* search that takes into account the surrounding environment e.g. a ligand leave the box unchecked. The *ab initio* search is slower.
- Check whether you would like the stack of solutions saved in an object or in a table. If you choose **Make Stack Table** a table will be displayed with the loop structures ranked by energy. To view each structure double click on the table. The first row is the loop with the best energy, you can sort a column by right clicking on the column header and choosing **sort**. If you choose store in an object you can view each member of the stack by clicking on the "+" button in the **stack** level of the object in the ICM Workspace.

High Precision Sampling in Batch

- Read your modeled structure into ICM. Or continue immediately after using build model.
- Select the loop region you wish to model (green crosses in the graphical display). Make sure the object you are selecting is the **current object**. You can make an object current by right clicking on the object in the ICM Workspace and select "Set to Current"
- MolMechanics/Sample Loop.
- Select the tab labeled **High Precision Sampling in Batch**
- Enter a value for effort factor (former 'thoroughness'). This will determine how long the simulation will last.
- Check whether or not you want to sample surrounding sidechains.
- Check whether or not you want to truncate remote parts of the protein.

8.3 Design Loop

Available in the following product(s): Homology

To design a loop between a gap in your protein structure use the **MolMechanics/Design Loop** option. This option will search for loops in the PDB which have a similar gap distance and geometry and output energetically favorable solutions detailing which PDB the loop was taken from. If you already have a loop and want to sample it use the `loop model` option.

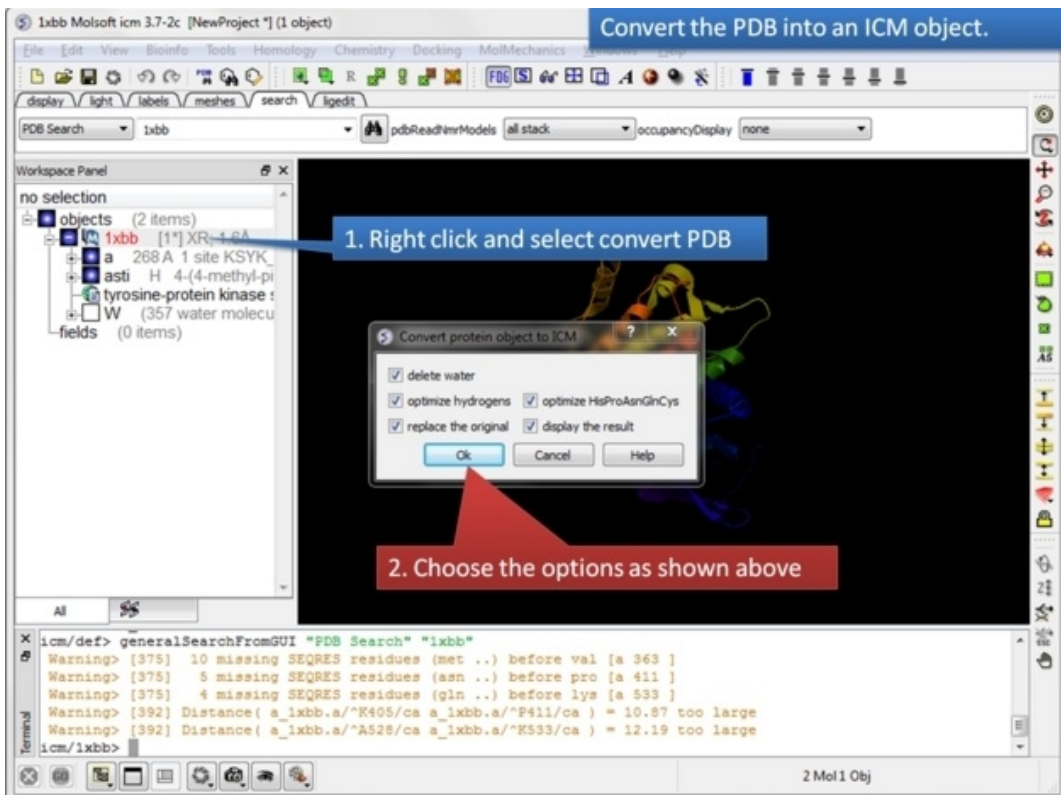
The screenshot shows the ICM software interface. At the top, a blue callout box reads "Read in the PDB file of interest. In this example we are using 1xbb." Below this, the "PDB Search" dropdown menu is set to "1xbb". A blue callout box points to the "search" tab in the top toolbar with the text "1. Select the search tab and enter 1xbb." The main workspace displays a protein structure with a loop highlighted in yellow and orange. The terminal window at the bottom shows the following output:

```

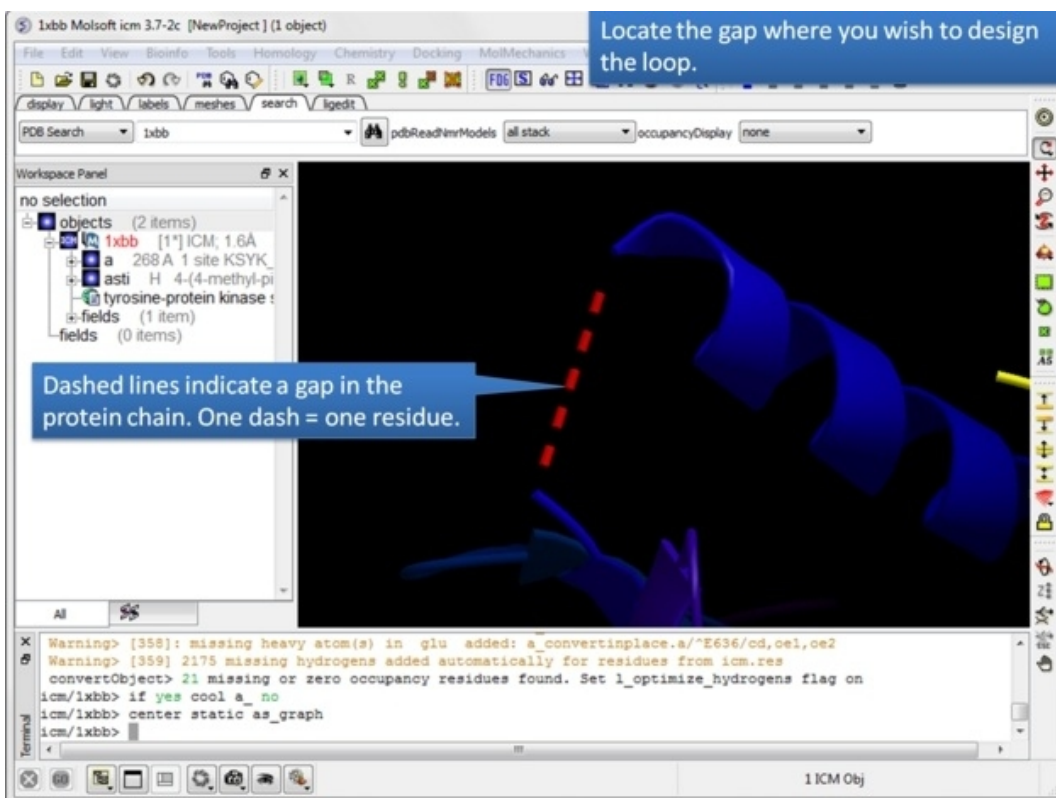
icm/def> generalSearchFromGUI "PDB Search" "1xbb"
Warning> [375] 10 missing SEQRES residues (met ..) before val [a 363 ]
Warning> [375] 5 missing SEQRES residues (asn ..) before pro [a 411 ]
Warning> [375] 4 missing SEQRES residues (gln ..) before lys [a 533 ]
Warning> [392] Distance( a_1xbb.a/~K405/ca a_1xbb.a/~F411/ca ) = 10.87 too large
Warning> [392] Distance( a_1xbb.a/~A528/ca a_1xbb.a/~K533/ca ) = 12.19 too large
icm/1xbb>

```

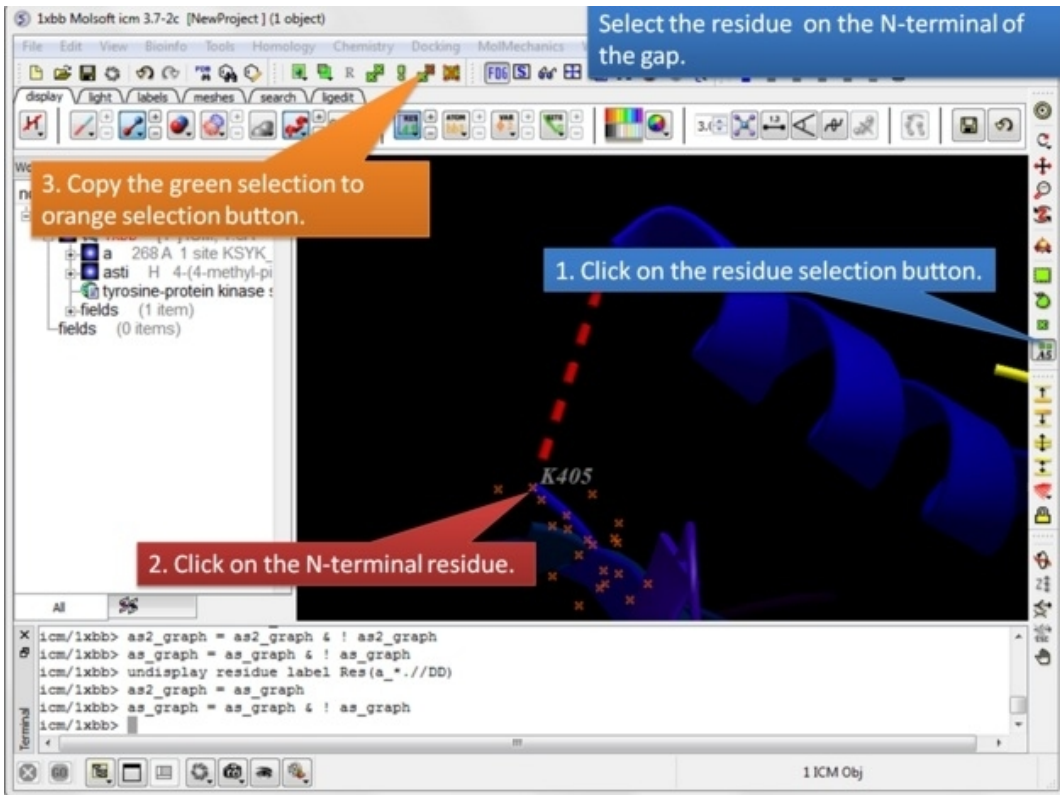
Step 1: Read in the PDB structure.



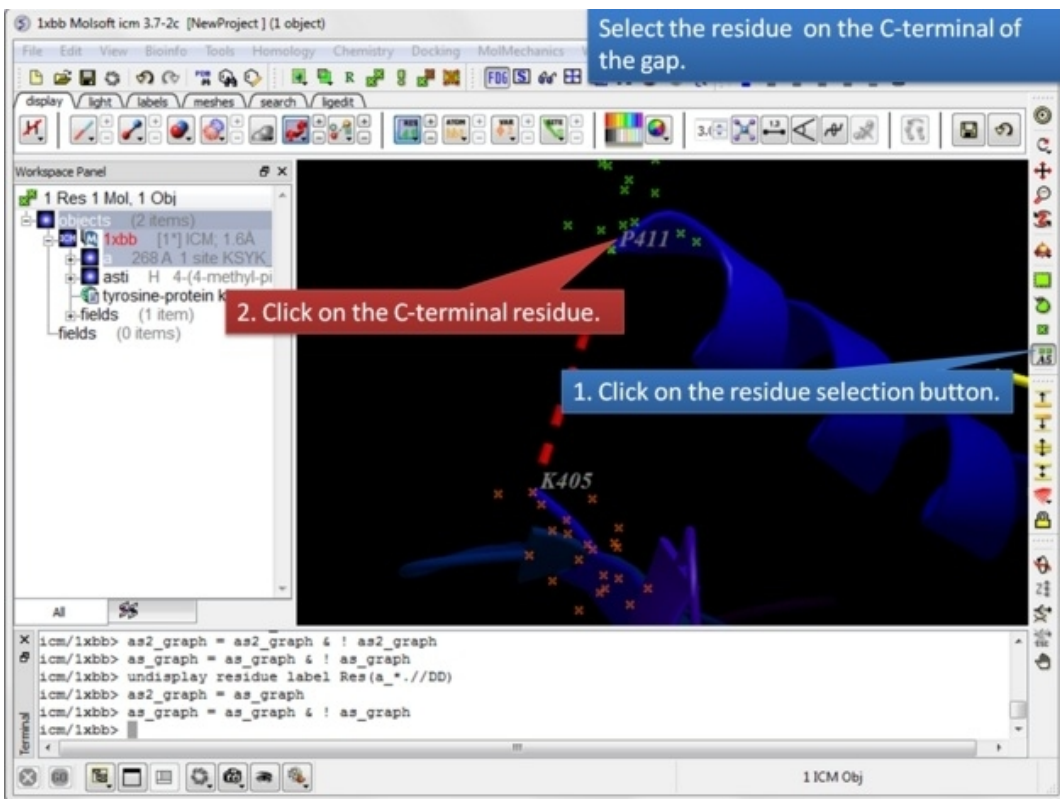
Step 2: Convert the PDB structure to an ICM object.



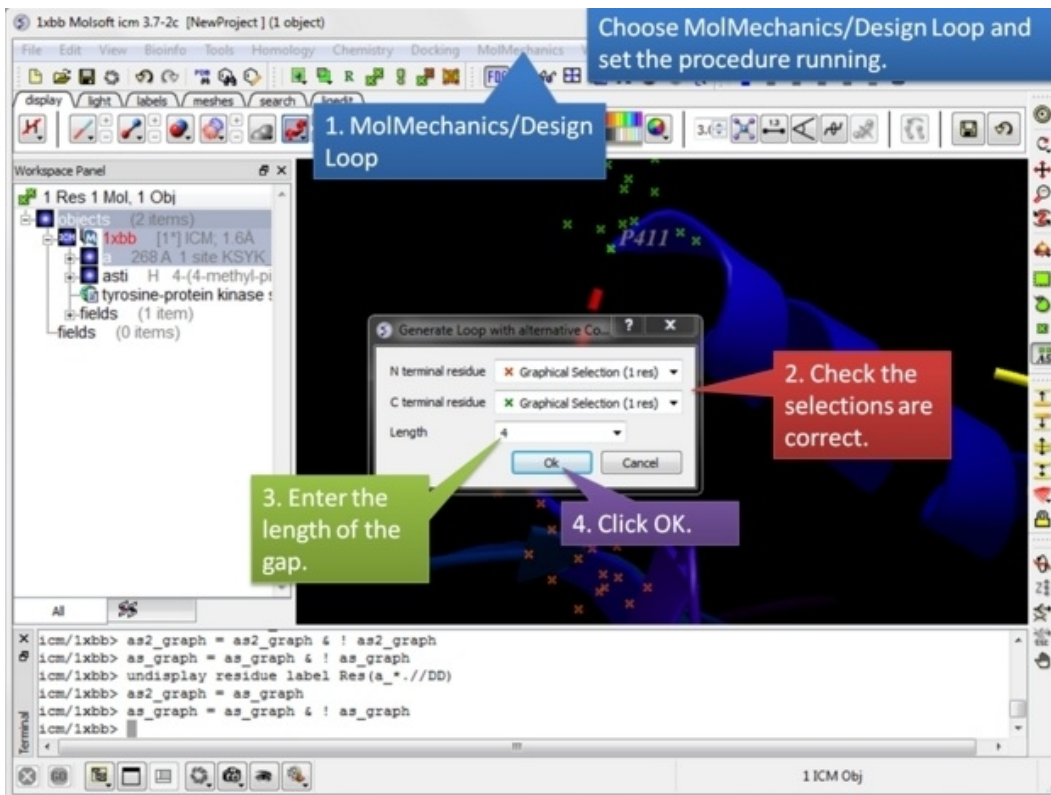
Step 3: Locate the gap where you want to design the loop. Gaps are represented by dashed lines when the structure is displayed in ribbon format.



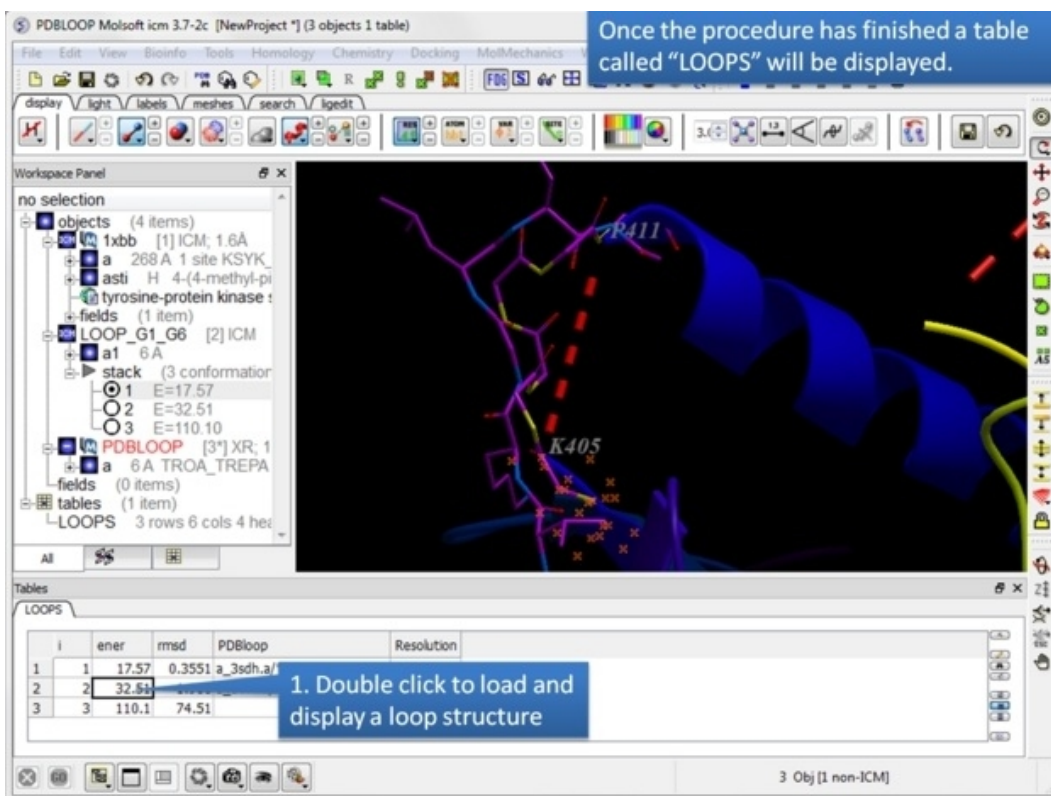
Step 4: Select the N-terminal residue and copy the green selection to orange.



Step 5: Select the C-terminal residue.



Step 6: Choose MolMechanics/Design Loop menu option.



Step 7: A table called **LOOPS** with the loop results will be displayed. Double click on the table to load the loop or use the `stack display` options in the ICM workspace. The loops are ranked by energy in the table and the column names refer to:

- **i** rank in stack
- **ener** Energy kcal/mol
- **rmsd** the distance (either Cartesian or angular RMSD) between the current conformation
- **PDBloop** PDB code and residues numbers that the loop was designed on
- **Resolution** Resolution of PDB template listed in PDBloop.

The loops are ranked by energy and the PDB structure upon which the loop is based is listed.

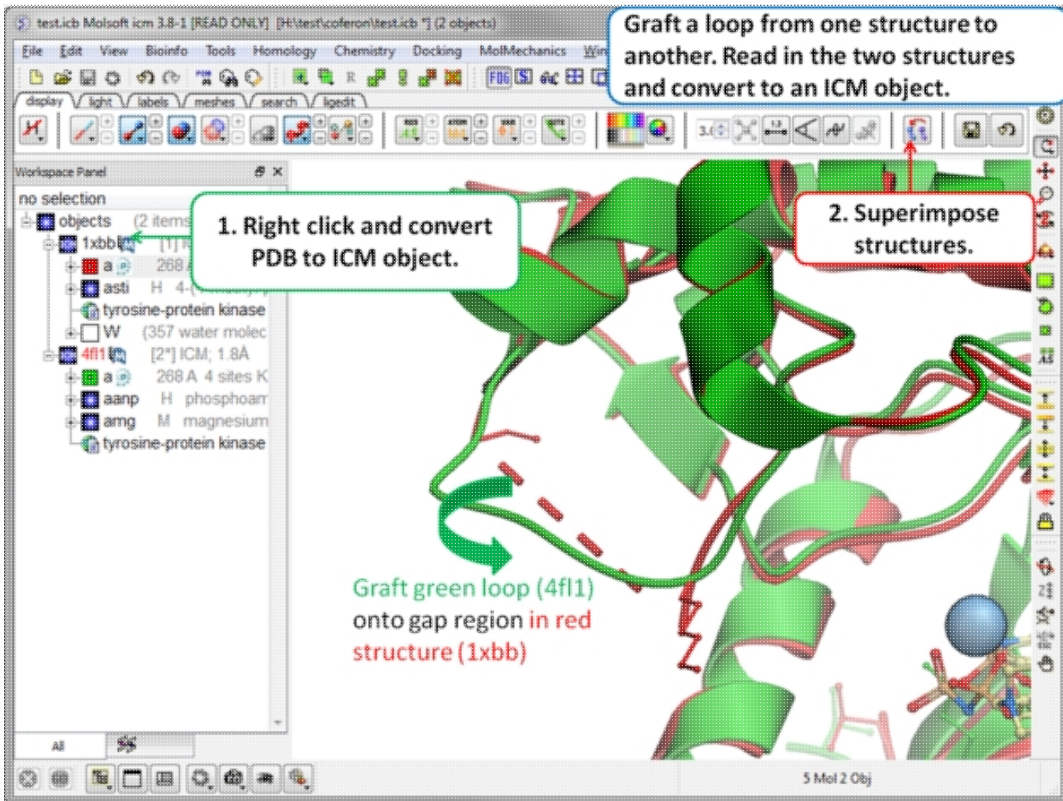
i	ener	rmsd	PDBloop	Resolution
1	17.57	0.3951	a_3sdh.a/* [54:59]	1.4
2	32.51	1.088	a_1toa.a/* [182:187]	1.8
3	110.1	74.51		0

Step 8: Note the PDB upon which the loop design is based is listed in the table.

8.4 Graft Loop

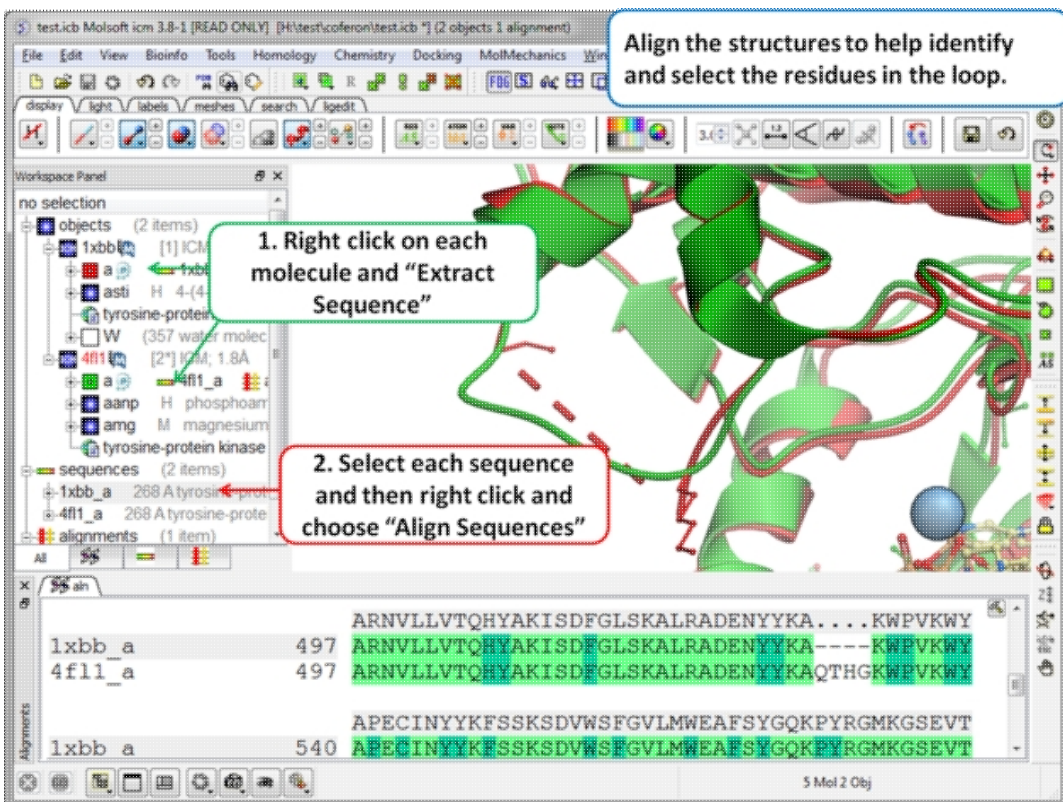
Available in the following product(s): Homology

To transfer a loop from one object to another:



Step 1: Read in the structures and convert to ICM objects.

- Read in the two structures into ICM.
- Convert the structures to an ICM Object.
- Superimpose the two structures.



Step 2 (Optional): To help identify and select the gap region you can align the PDB sequences.

- Right click on the protein molecule in each object and choose the option "Extract Sequences"
- Select the sequences in the ICM Workspace by double clicking on one sequence- hold CTRL- and click on the other sequence.
- Right click on the selected sequences and choose the Align option.

The screenshot shows the Molsoft ICM 3.8-1 interface. The top menu bar includes File, Edit, View, Bioinfo, Tools, Homology, Chemistry, Docking, MolMechanics, and Win. The Workspace Panel on the left lists objects: 6 Res, 1 Mol, 1 Obj. Under 'Objects (2 items)', there are two ICM objects: 1xbb_a (1 ICM, 1 6A) and 4f11_a (2 ICM, 1 6A). The 'sequences' section shows 1xbb_a (268 A tyrosine-prot) and 4f11_a (268 A tyrosine-prot). The 'alignments' section shows one alignment. The main 3D view displays a protein structure with a green loop highlighted. A callout box with a blue border says "Select the loop you wish to graft. Remember to include the termini residues." The alignment window at the bottom shows the following sequences:

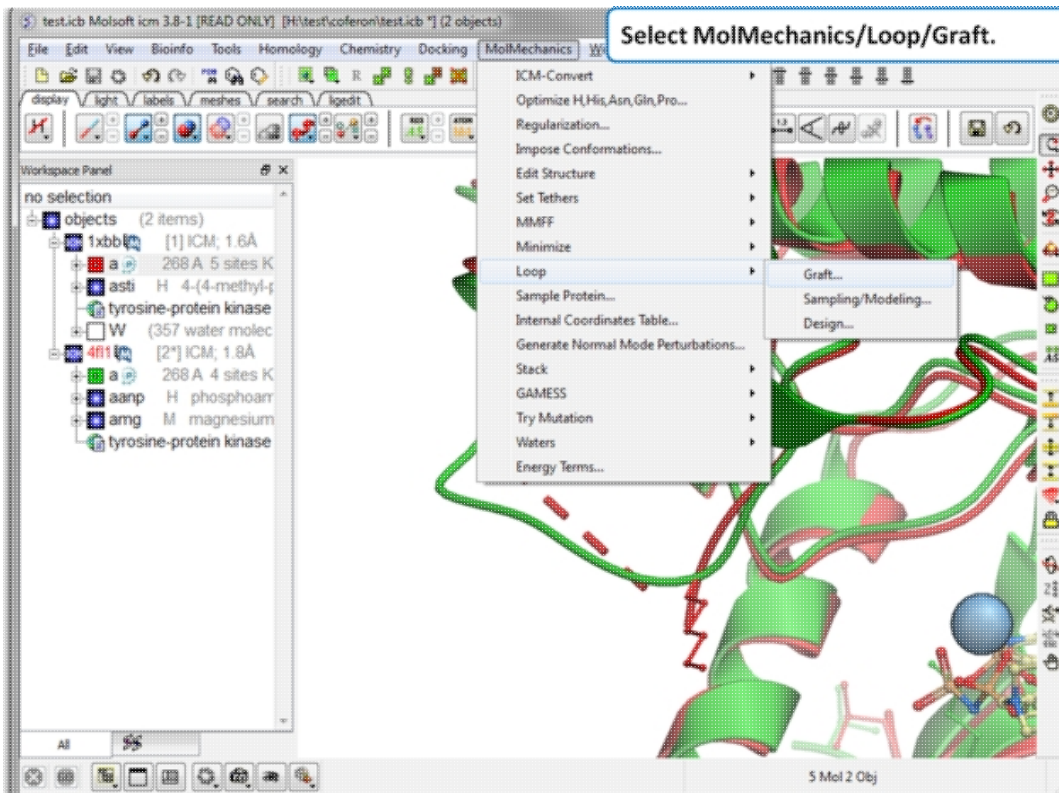
```
ARNVLLVTQHYAKISDFGLS
1xbb_a 497 ARNVLLVTQHYAKISDFGLSKALRADENYYKA----KHPVKWY
4f11_a 497 ARNVLLVTQHYAKISDFGLSKALRADENYYKAQTHGKHPVKWY

APECINYYKFSSKSDVWSFGVLMWEAFSYGQKPYRGMKGSEVT
1xbb_a 540 APECINYYKFSSKSDVWSFGVLMWEAFSYGQKPYRGMKGSEVT
```

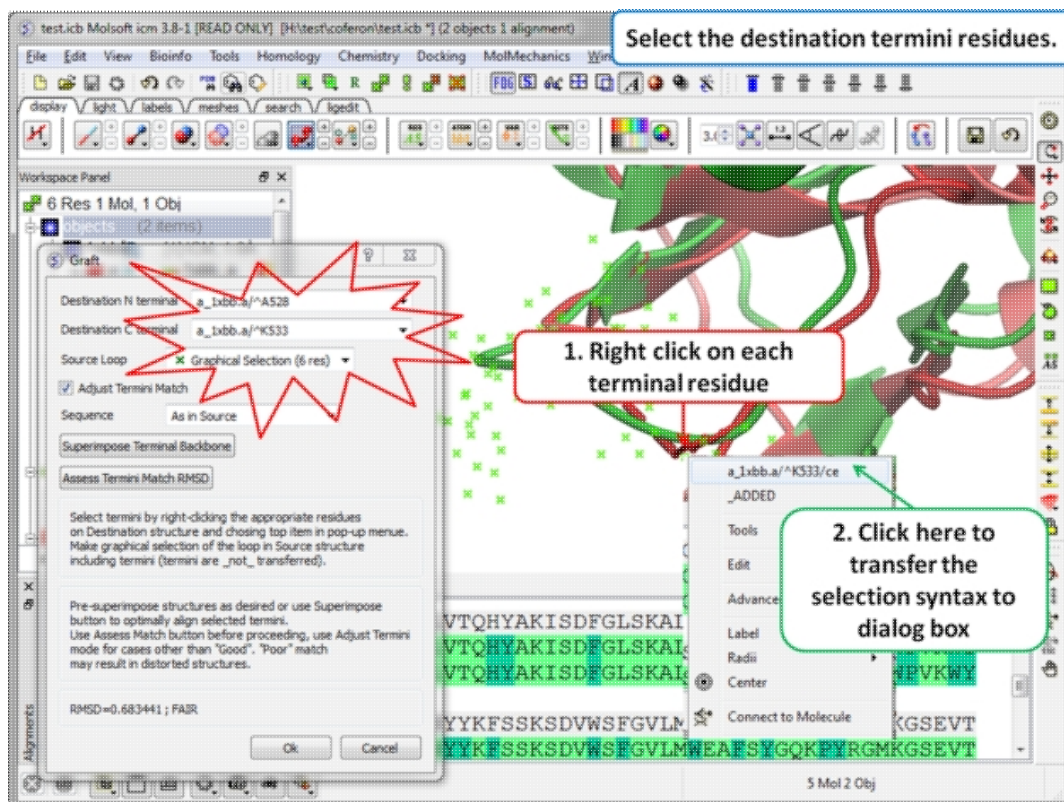
A red dashed box highlights the sequence "ARNVLLVTQHYAKISDFGLSKALRADENYYKA" in the alignment, with a red arrow pointing to it from a callout box that says "1. Click and drag to select".

Step 3: Select the source loop.

- Select the loop and the **termini residues** (e.g. click and drag over them in the alignment).



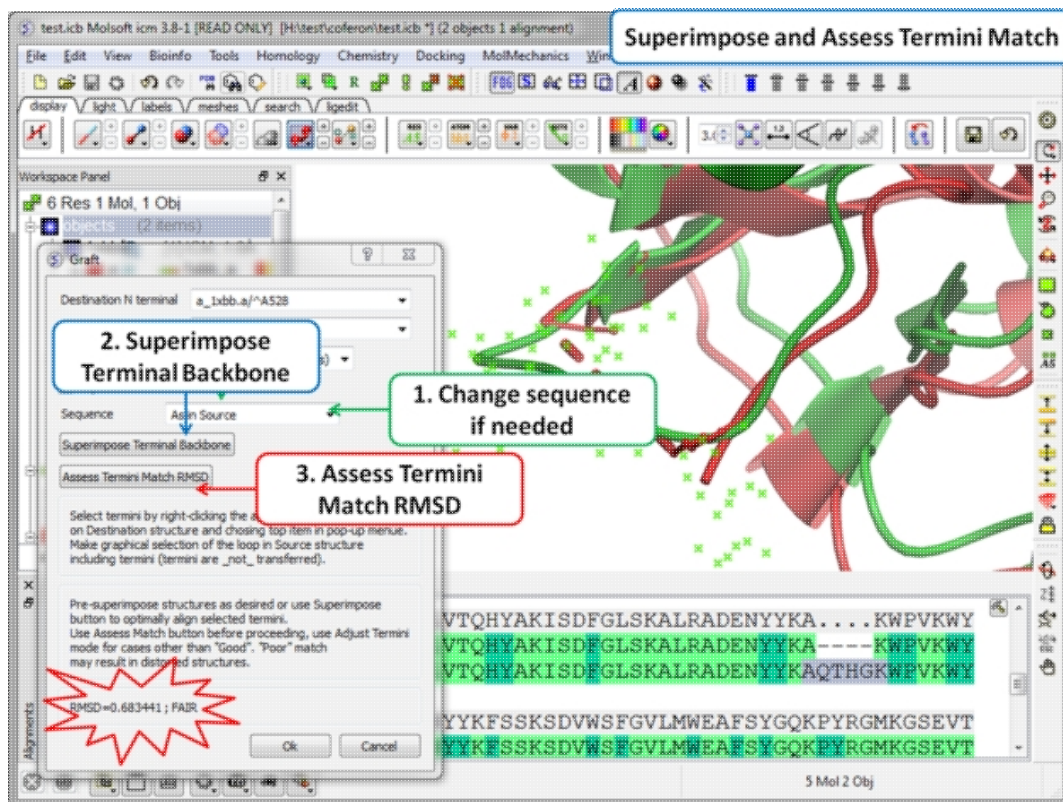
Step 4: Select MolMechanics/Loop/Graft



Step 5: Select the termini residues of the structure with the gap (destination).

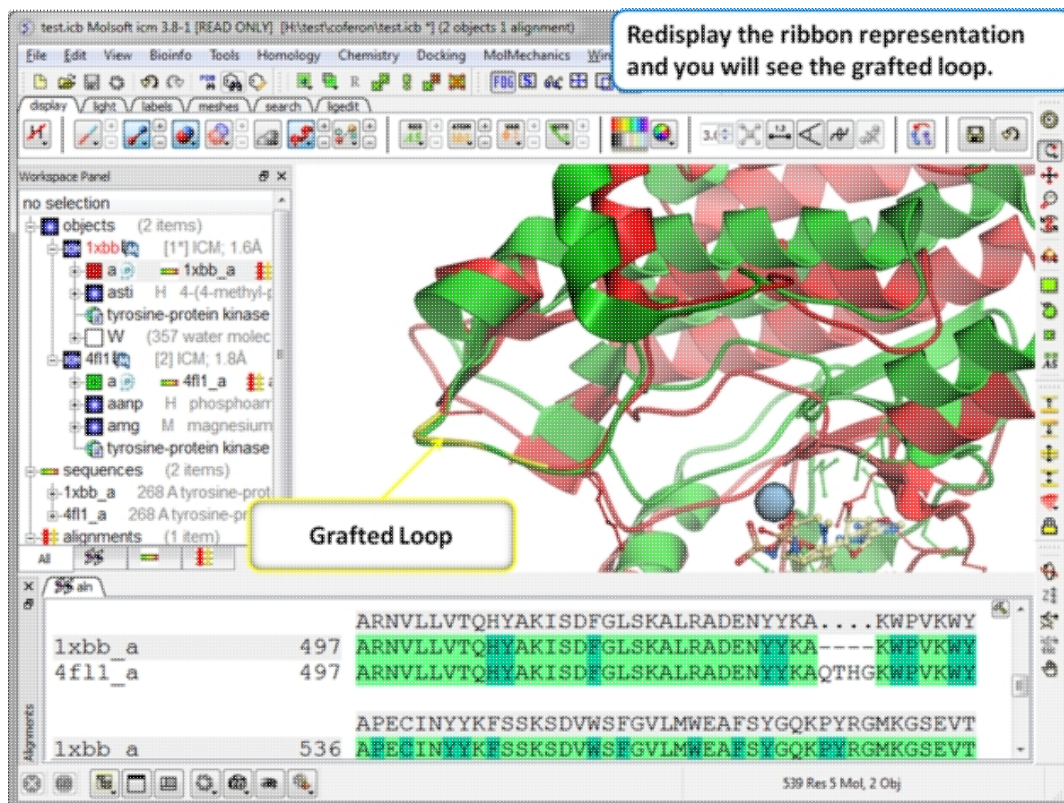
- Right click on the N-terminal residue flanking the gap and then left click on the top item of the menu. This will transfer the selection to the dialog box.

- Right click on the C-terminal residue flanking the gap and then left click on the top item of the menu. This will transfer the selection to the dialog box.



Step 6: Superimpose termini backbone and assess the quality of the match.

- You can change the sequence of the residues in the grafted gap (click on the drop down button in the Sequence dialog box).
- Click on the superimpose terminal backbone button.
- Use "Assess Match" button before proceeding, the RMSD and report will be displayed in the terminal box. Use adjust termini mode for cases other than Good.
- Click OK.



Step 7: Display the grafted structure. Sometimes you may need to refresh the ribbon display to view the new loop.

8.5 Loop Preferred Residues

For a given segment of polypeptide chain ("loop") this method scans the PDB database for segments that have similar conformations (within "vicinity") and calculates relative frequencies of residues at each position. Thus you can find out if certain residues might be preferred in some positions, for instance in beta-turns depending on the type a glycine or a proline might be preferred.

- Read in a protein structure and convert to an ICM object.
- Select the residues in the loop.
- MolMechanics/Loop/Preferred residues
- Enter a vicinity value (see: <http://www.molsoft.com/man/reals.html#vicinity>).
- Click OK.
- A table of the relative frequencies of residues at each position are reported.

8.6 Find Loops in PDB

To find loops in the PDB that have similar conformation to a loop displayed in ICM:

- Read in a protein structure and convert to an ICM object.
- Select the residues in the loop.
- MolMechanics/Loop/Find PDB Segments
- Enter a vicinity value (see: <http://www.molsoft.com/man/reals.html#vicinity>).
- Enter the number of hits it should return.
- Click OK.
- A table of the hits containing the PDB codes containing a similar loop conformation will be displayed. The column aRMSD is angular (torsional) RMSD.

8.7 Regularization

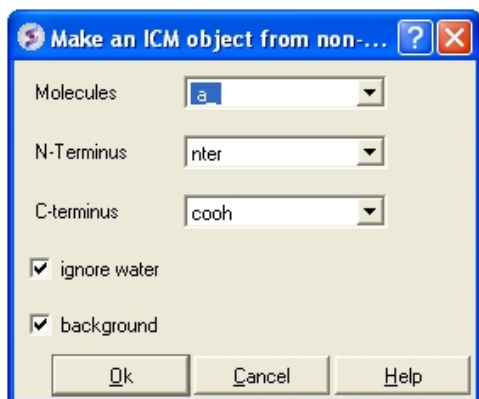
Available in the following product(s): ICM-Pro

Regul stands for Regularization which is a procedure for fitting a protein model with the ideal covalent geometry of residues (as represented in the `icm.res` residue library) to the atom positions of a target PDB structure.

Regularization is a procedure for fitting a protein model with the ideal covalent geometry of residues (as represented in the `icm.res` residue library) to the atom positions of a target PDB structure (usually provided by X-ray crystallography or NMR). Regularization is required because the experimentally determined PDB-structures often lack hydrogen atoms and positional errors may result in the unrealistic van der Waals energy even if these structures were energetically refined (since the refinement of the crystallographic structures typically ignores hydrogen atoms and employs different force fields). The following steps are required to create the regularized and energy refined ICM-model of an experimental structure.

To use the regularization protocol:

- MolMechanics/Regularization and a data entry box as shown below will be displayed.
- Enter the molecule you wish to refine.
- Choose which kind of N and C- terminus
- Choose whether you wish to include water molecules.
- Choose to run in background if the structure is very large.



Once the refinement is complete a new ICM object will be displayed in the ICM workspace called `object_name_reg`

8.8 Refine Side Chains

Available in the following product(s): ICM-Pro

Once a molecular model has been constructed it is generally a good idea to analyze it using the `Protein Health` macro. The results from protein health will indicate which part of the molecule is strained or has unusual geometry. A way of solving these problems is to refine or optimize the side-chains. Other ways to minimize or optimize the structure are available in the `MolMechanics/Minimize` menu.

To refine or optimize a selection of side chains the structure needs to be an ICM object.

- Make a selection of the side chains you wish to optimize. See `how to make selections` section.
- Right click on the selection in the graphical display. A selection will be displayed as green crosses.
- Select from the right click menu `Advanced/Optimize Side Chains` and a data entry box as shown below will be displayed.
- Enter the number of calls per variable you wish to use for the simulation. For more details on this please see the ICM language manual.



A stack of energy conformations will be displayed in a table. Each conformation can be viewed simply by clicking in the table.

8.9 Making a disulfide bond.

Available in the following product(s): ICM-Pro

To make a disulfide bond:

- MolMechanics/Edit Structure/Set Disulfide Bond...
- Click in the "first cys" data entry box and right click on the cys residue or type in the ICM selection language for the first cysteine.
- Click in the "second cys" data entry box and right click on the cys residue or type in the ICM selection language for the second cysteine.
- Click on the **Apply** button

1. Click in the data entry box

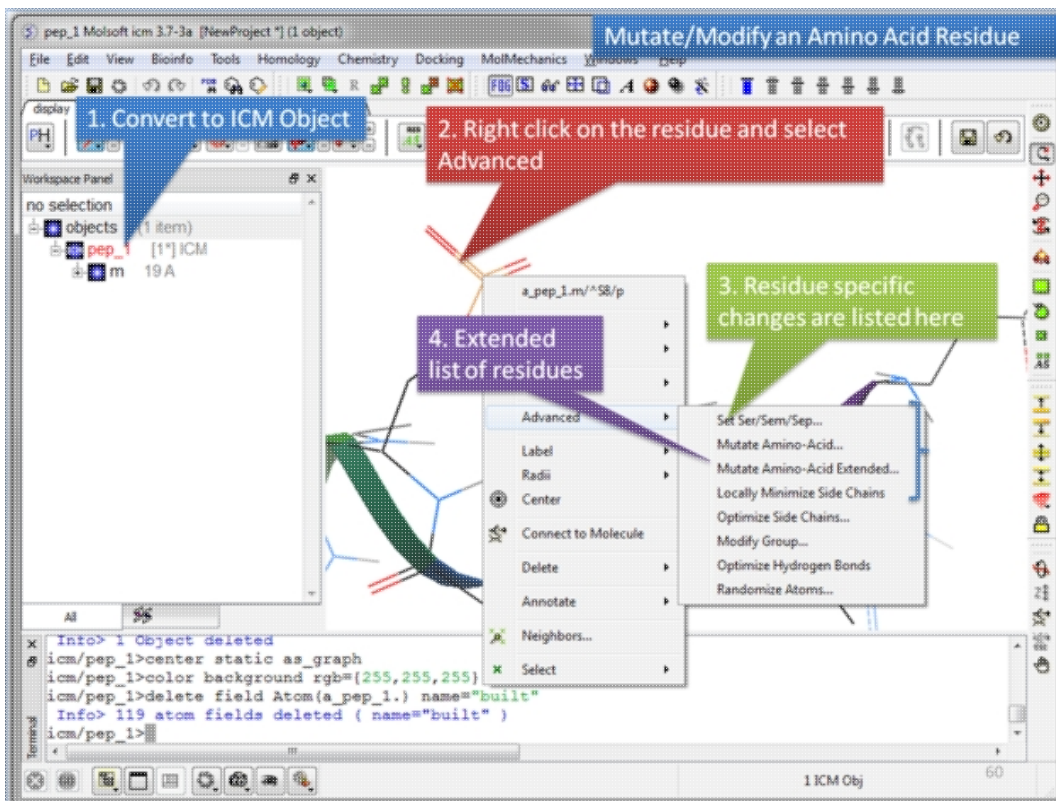
2. Right click on the CYS residue and select the selection language description of the atom. This will then be added to the Set Disulfide Bond data entry box.

An alternative method is to type the selection language directly.

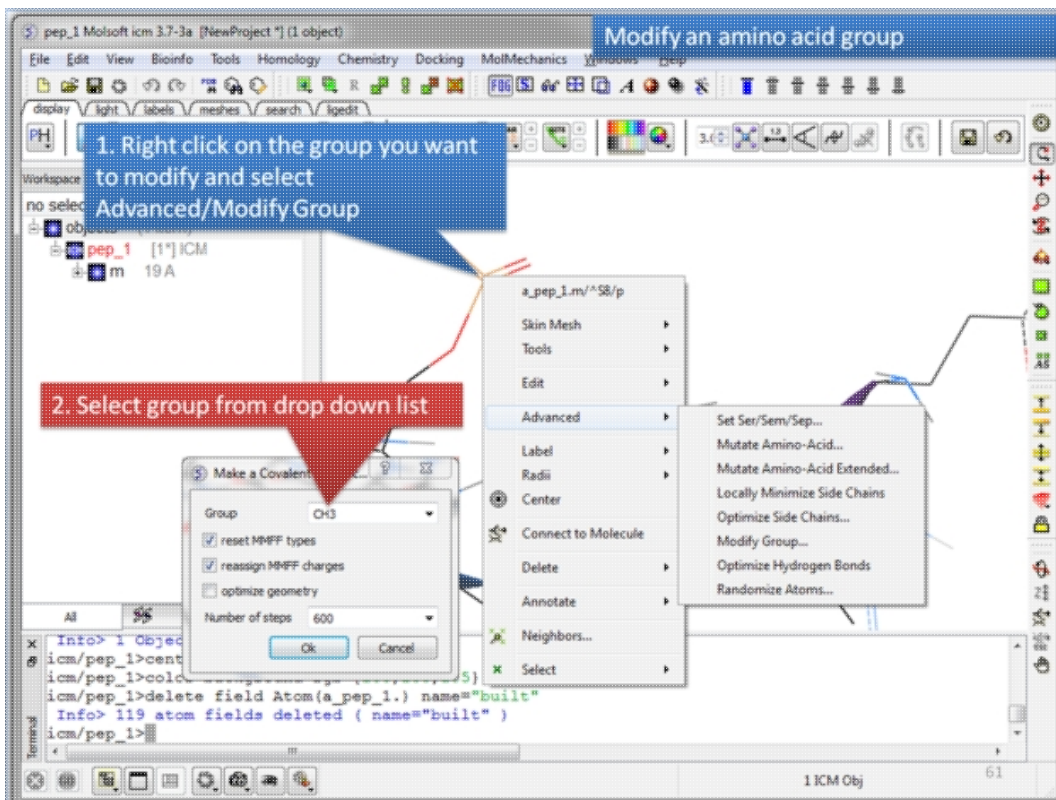
8.10 Making a Mutation

Available in the following product(s): ICM-Pro

How to Mutate or Modify an Amino Acid Residue



In order to make a mutation the molecule needs to be an ICM Object. Right click on the amino acid residue you wish to mutate and choose Advanced and then select one of the mutation options.
How to Modify an Amino Acid Group



In order to modify a group you need to convert your molecule into an ICM Object. Right click on the amino acid residue you wish to modify and choose Advanced/Modify Group.

8.11 Predicting Change in Binding Free Energy upon Mutation

|Tutorial|

Available in the following product(s): ICM-Pro

This method computes the change in binding free energy of a protein complex upon mutation of a single residue.

Getting Started A PDB structure or ICM object containing the protein complex is needed. A graphical selection of the residue to be mutated is then made, the mutant amino acid (e.g., "ala" or "all" for calculation of the energy for all natural amino acids) is then selected. The energy is computed as interaction energy between two parts (each part may contain more than one protein chain). One part has to be specified, e.g., "a" or "a,b".

Calculation The binding free energy change, ΔG_{bind} , is computed as a difference between the free energy of mutant and wild type:

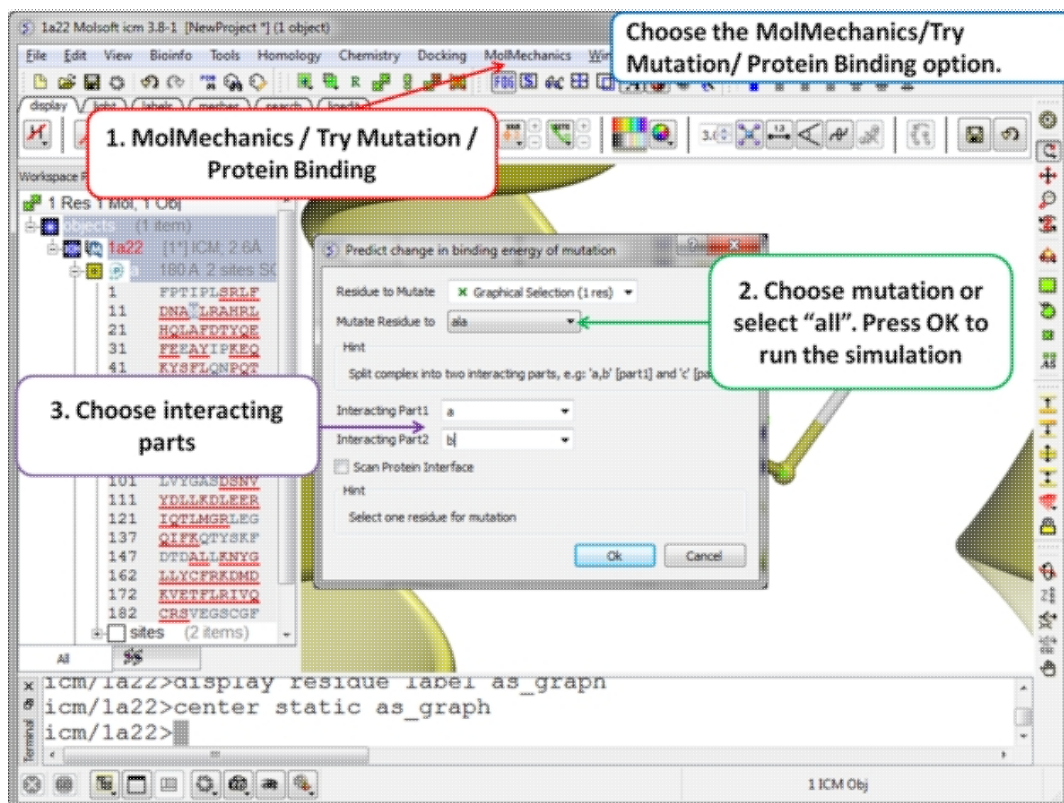
$$\Delta\Delta G_{bind} = \Delta G_{bind}^{mut} - \Delta G_{bind}^{wt}, \text{ where}$$
$$\Delta G_{bind} = (E_{intra}^{comp} - E_{intra}^{parts}) + (\Delta G_{solv}^{comp} - \Delta G_{solv}^{parts})$$

where E^{comp} represents the energy of the *complex*, and E^{parts} represents the sum of the energies, internal or solvation, for each part. The energy is calculated for fixed backbone and all the side chains except those in the vicinity of the mutable residue. Monte Carlo simulations are carried out to relieve possible atomic clashes created as a result of mutations to larger amino acid residues. "Scan Protein Interface" option allows to mutate all residues (one by one) of the Interacting Part in close contact with the second part of the complex.

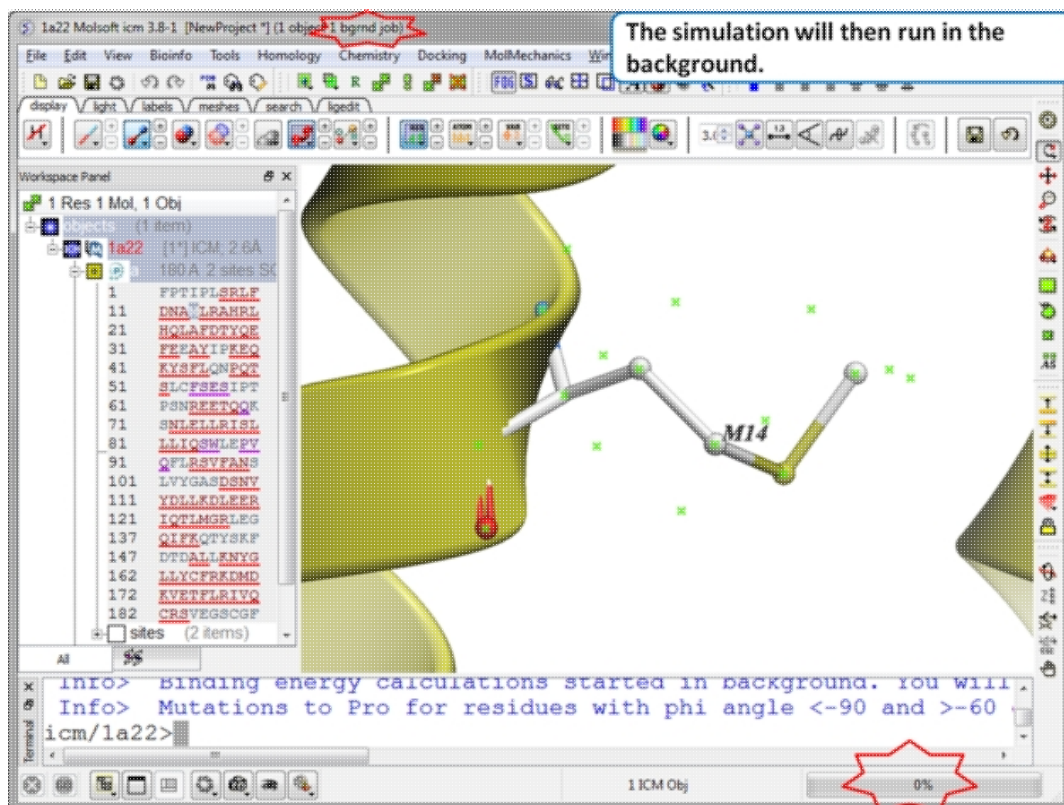
The screenshot shows the ICM-Pro software interface. The main window displays a protein structure with a yellow highlighted region. A red box with the text "1. Use the residue selection button to select the residue to mutate" points to a button in the right-hand toolbar. A blue box with the text "Read in a protein structure and select the residue you wish to mutate." is located at the top right. The left panel shows a list of residues with their sequences, and the bottom panel shows a terminal window with the following commands:

```
icm/la22>display xstick as_graph
icm/la22>display residue label as_graph
icm/la22>
```

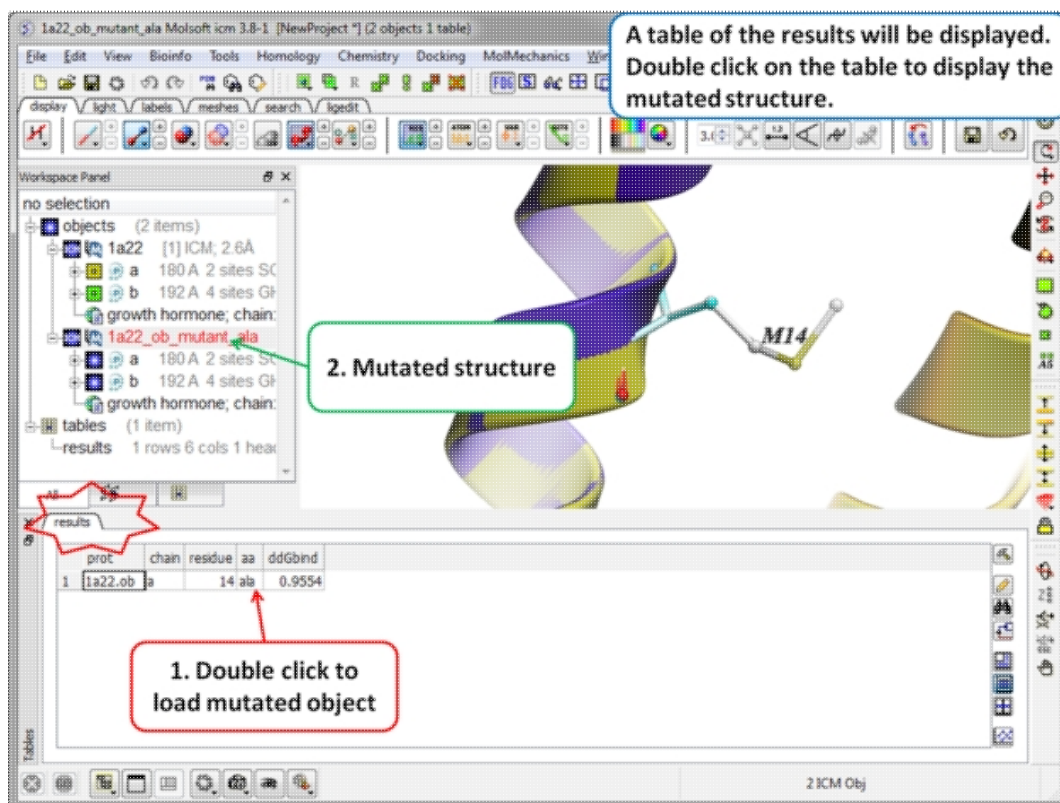
Step 1 Read in the PDB file of interest and select the residue you wish to mutate.



Step 2 Choose the option MolMechanics/Try Mutation/ Protein Binding and a dialog box as shown above will be displayed. If you have selected a residue it should be reported in the "Residue to Mutate" box. Select a residue to mutate to from the drop down list or select "all" to try all 20 natural amino acids. One or more interacting part(s) of the protein need to be specified e.g. if chain a interacts with chain b - enter "a" as interacting chain 1 and "b" as interacting chain 2. The "Scan Protein Interface" option allows to mutate all residues (one by one) of the Interacting Part1 in close contact with the second part of the complex



Step 3 The simulation will run in the background and may take a while to finish if the option "all" or "scan protein interface" is selected.



Step 4 A table reporting the binding free energy in kcal/mol (ddGbind) will be displayed. Double click a row of the table to load and display the mutated structure.

8.12 Predicting Change in Protein Stability upon Mutation

|Tutorial|

Available in the following product(s): ICM-Pro

This method computes the change in protein stability upon mutation of a single residue.

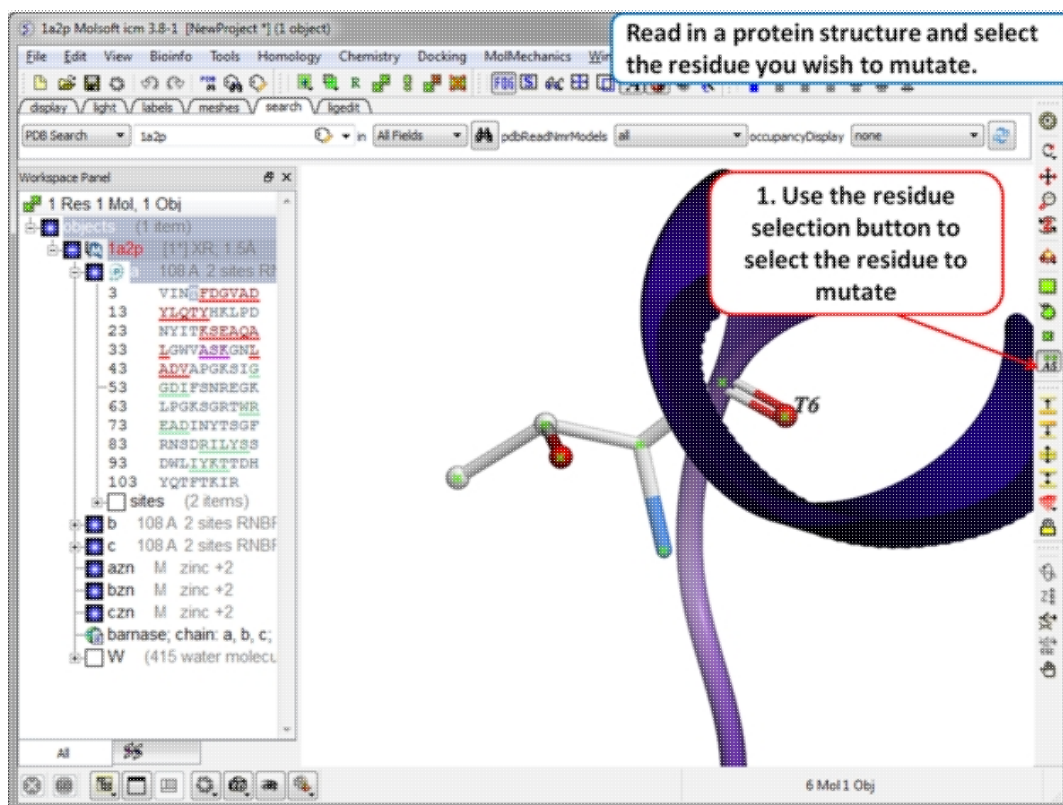
Getting Started: A PDB structure or ICM object containing the protein complex is needed. A graphical selection of the residue to be mutated is then made, the mutant amino acid (e.g., "ala" or "all" for calculation of the energy for all natural amino acids) is then selected.. "Scan Sequence" allows you to mutate all the residues in the sequence to the specified amino acid.

The free energy change in protein stability is computed as follows:

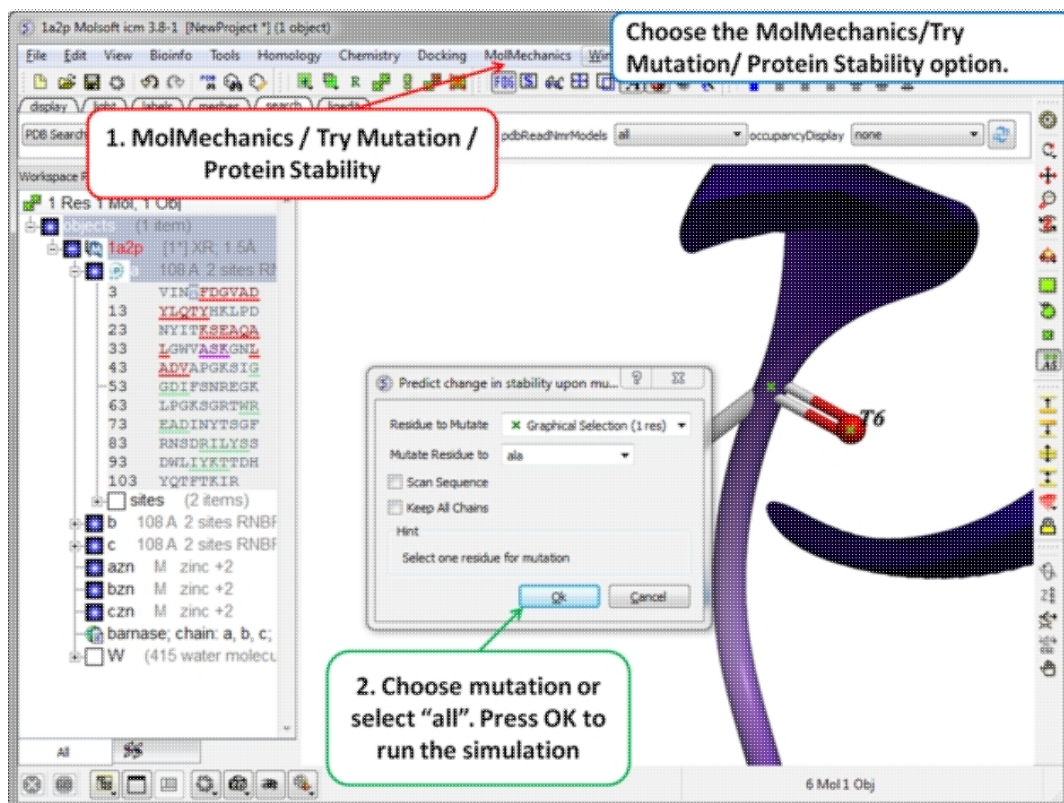
$$\Delta\Delta G = \Delta G^{\text{mutant}} - \Delta G^{\text{wt}}$$

$$\Delta G = \Delta G_{\text{folded}} - \Delta G_{\text{unfolded}}$$

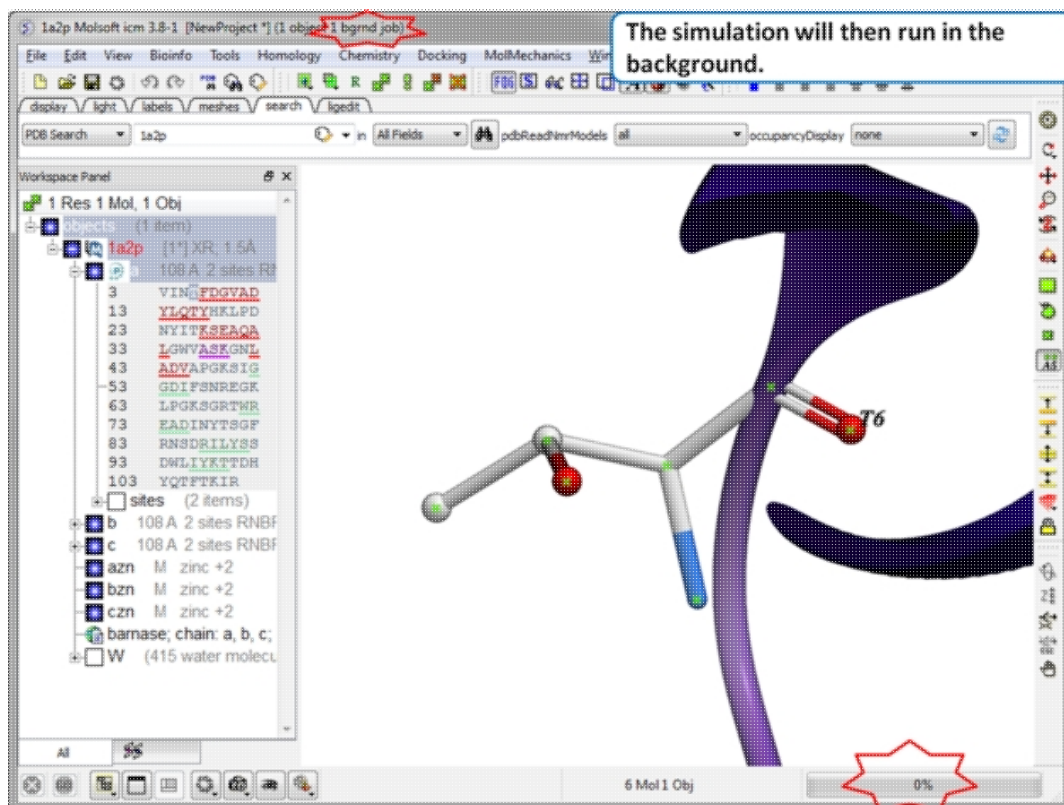
Calculation The free energy of the unfolded and misfolded states is approximated by a sum of the residue-specific energies. The residue-specific energies were derived empirically using a large set of experimental data. Mutation of a given residue is followed by Monte Carlo simulations with flexible side chains for the mutated residue and its neighboring residues. The rest of the protein structure is considered rigid. A positive energy value indicates that the mutation is likely to be destabilizing.



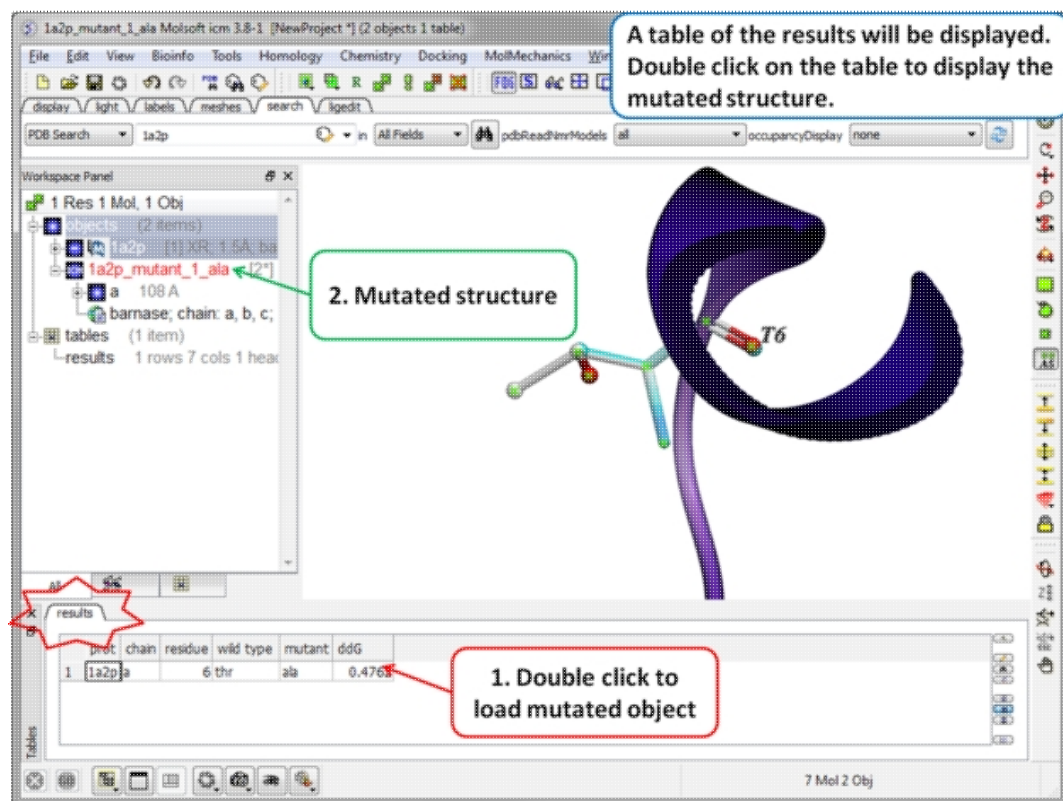
Step 1 Read in the PDB file of interest and select the residue you wish to mutate.



Step 2 Choose the option MolMechanics/Try Mutation/ Protein Stability and a dialog box as shown above will be displayed. If you have selected a residue it should be reported in the "Residue to Mutate" box. Select a residue to mutate to from the drop down list or select "all" to try all 20 natural amino acids. The "Scan Sequence" option allows you to mutate all the residues in the sequence to the specified amino acid. The "Keep All Chains" keeps all the chains in the mutated structure, otherwise, all chains except the one containing the mutable residue will be deleted".



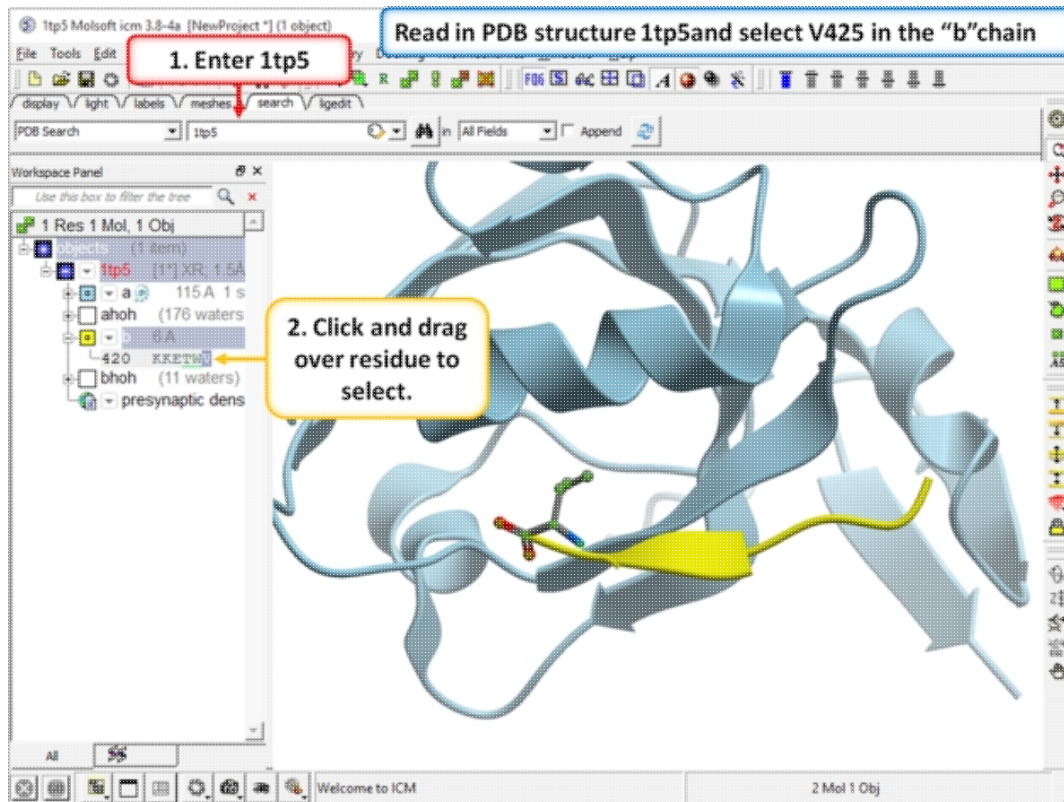
Step 3 The simulation will run in the background and may take a while to finish if the option "all" or "Scan Sequence" is selected.



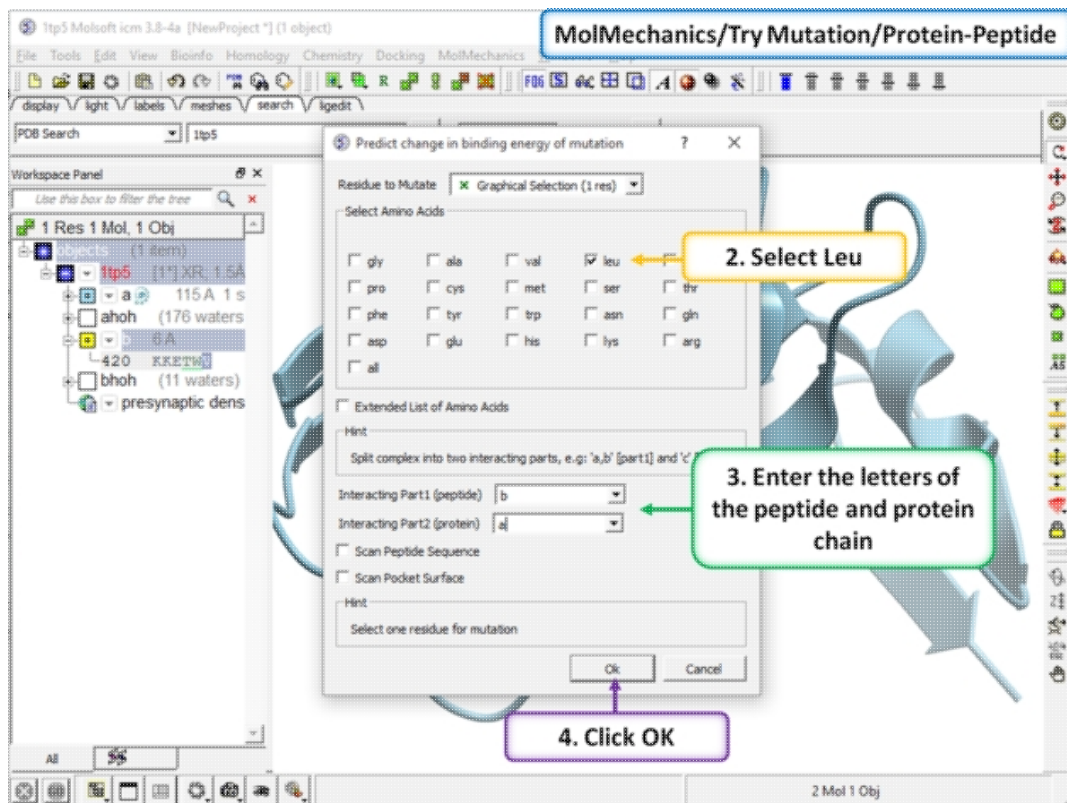
Step 4 A table reporting the free energy change in protein stability will be displayed (ddG kcal/mol). Double click a row of the table to load and display the mutated structure.

8.13 Effect of Mutation on Protein-Peptide

This option uses the same methodology as the effect of mutation on Protein-Protein interaction except that backbone in the peptide is flexible.

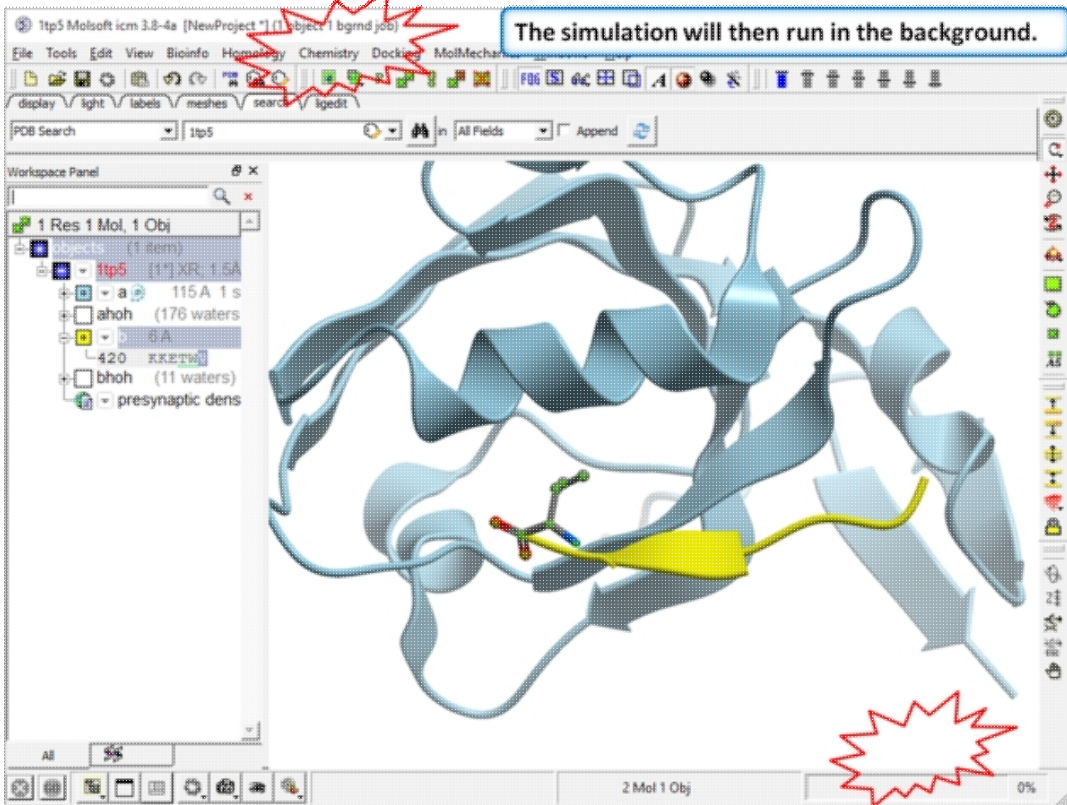


Step 1 Read in your PDB structure. In this example we will use PDB 1tp5 and select V425 in the peptide (b chain).

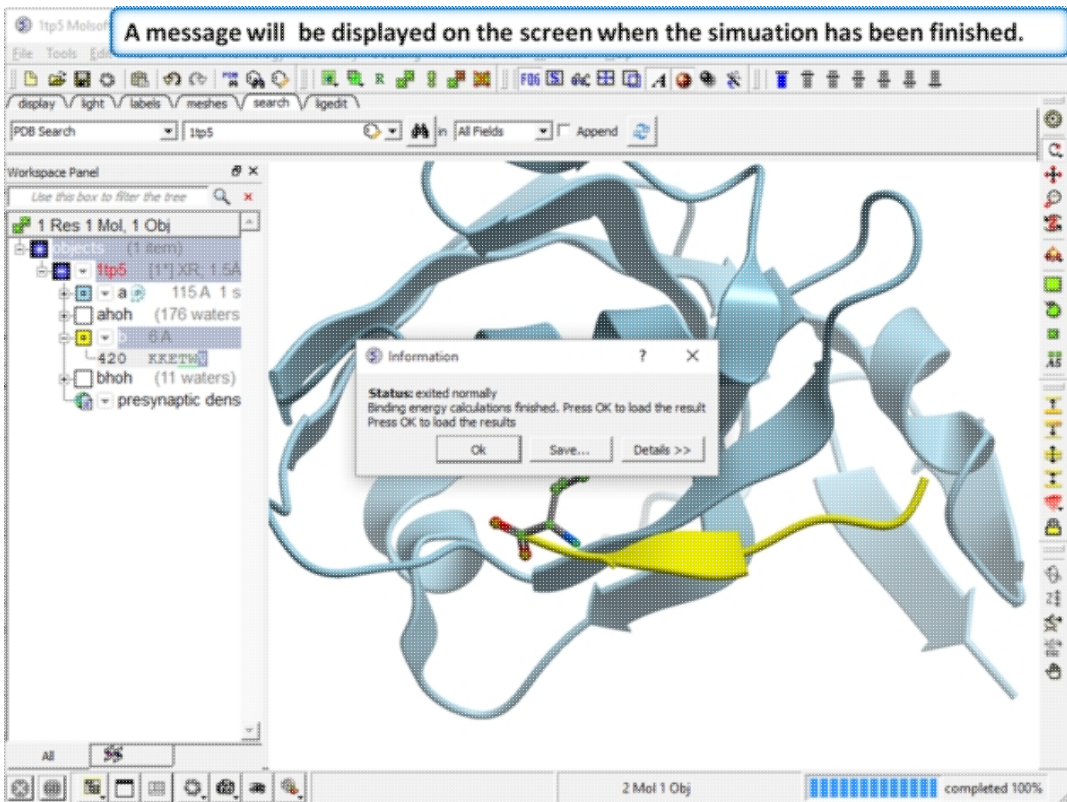


Step 2 In this example we will choose to see the effect of mutation of Val > Leu. Enter the letter of the chain for the interacting part 1 (peptide) and part 2 (protein). In addition you can choose:

- **Scan Peptide Sequence** This option will try a single residue side-chain mutation at each position in the peptide.
- **Scan Pocket Surface** This option will try a single residue side-chain mutation at each position in the protein that is in contact with the peptide.



Step 3 The simulation will run in the background and may take a while to finish if the option "all" or "Scan Sequence" is selected.



Step 4 A message dialog box will let you know when your simulation has finished.

A results table will be displayed reporting the ddG value.

prot	chain	residue	aa	ddGbind
1tp5	b	425	LEU Leucine	0.4787

1. Double click to load mutant structure

Step 5 A table reporting the free energy change in peptide will be displayed (ddG kcal/mol). Double click a row of the table to load and display the mutated structure.

8.14 Effect of Mutation on Protein Ligand Binding

To predict the effect of mutation on Protein Ligand Binding:

- Read in your PDB structure. The ligand and receptor have to be in the same object.
- Select the residue you wish to mutate
- Go to MolMechanics/Try Mutation/Protein Ligand Binding
- Select the ligand in the drop down menu.
- Select the amino acids you would like to try.
- **Scan Pocket Surface** This option will try a single residue side-chain mutation at each position in the protein that is in contact with the receptor.
- **Keep All Chains** will maintain the full PDB file in the results table.

Results

- A table will be displayed with the calculated ddG(kcal/mol) for each mutation selected.

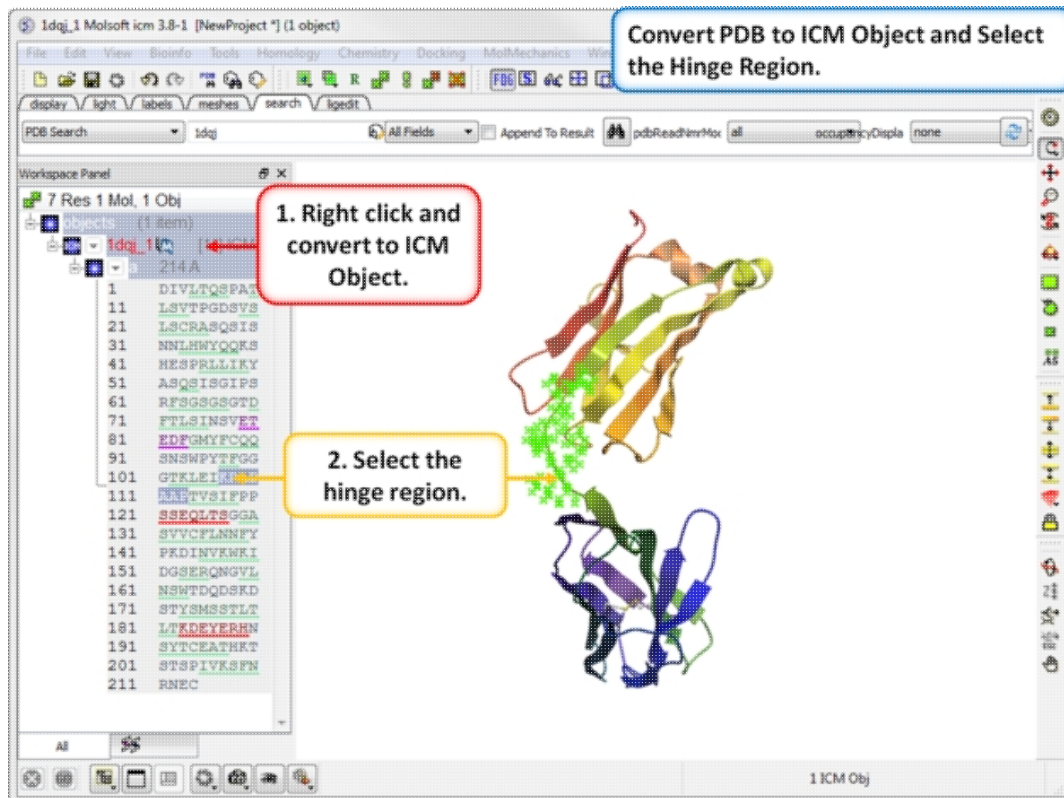
8.15 Predict Disulfide Bond

This option allows you to scan pairs of residues in two input residue selections that can be replaced by Cysteines forming a disulfide bridge. To do this:

- Read in a PDB structure.
- Select one or more chains using the regular green selection as well as the orange selection. If the second argument is empty it is assumed equal to the first selection.
- Tools/Design/Disulfide Bridges
- The procedure returns a table called CysBridges with possible pairs sorted by the strain associated with the bridge.

8.16 Protein Sculpting

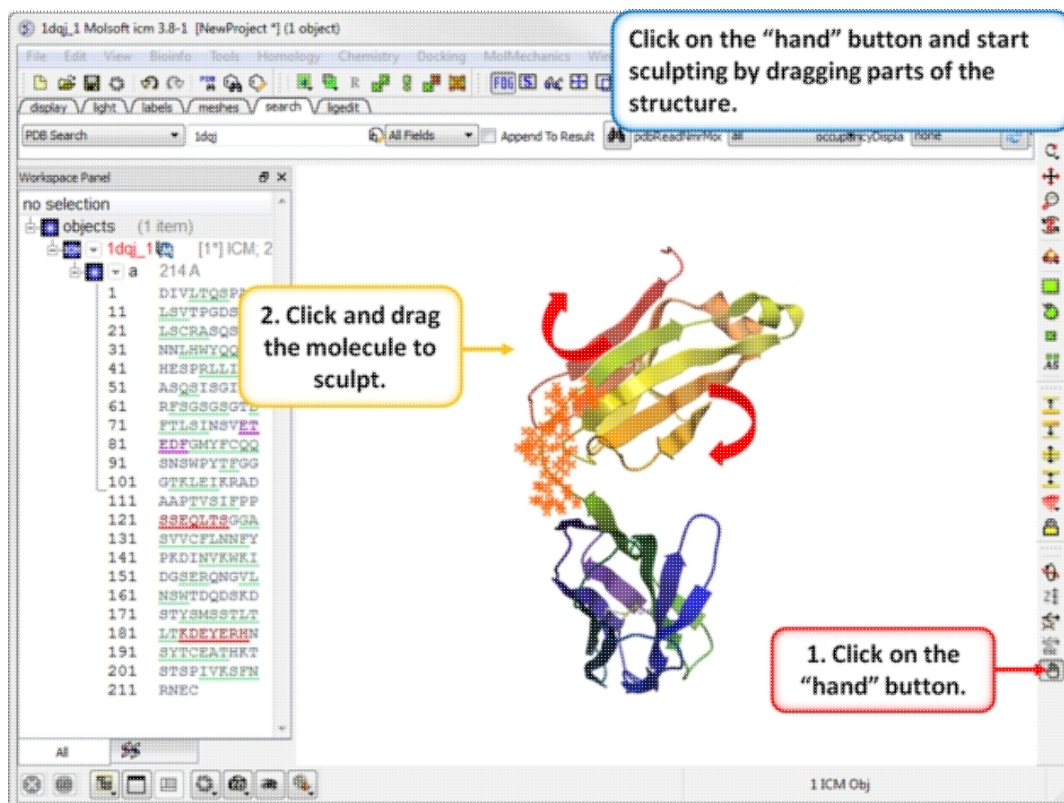
Protein sculpting can be used to model the interaction of two connected domains. You simply drag the molecules around a hinge region followed by a minimization.



Step 1. Convert your protein to an ICM Object and select the hinge region.



Step 2. Change the selection to an orange selection.



Step 3. Select the "hand" button and then click and drag the protein. Wait for the minimization to finish before making another change.

8.17 Sample Protein

To generate a stack of energetically favorable protein conformers using the ICM Biased Probability Monte Carlo method:

- Read in a protein structure and `convert` to an ICM object.
- Select the residues you wish to sample.
- MolMechanics/Sample Protein
- Enter a value for Effort. This represents the length of the simulation - values between 1 and 10 will give a thorough sampling. A larger effort value will be needed for larger proteins.
- Enter a value for the Angular Vicinity (default 15). Please see <http://www.molsoft.com/man/reals.html#vicinity>
- Enter a Temperature value. Please see <http://www.molsoft.com/man/reals.html#temperature>
- Click OK.
- When the simulation is finished a stack of conformations will be displayed in the ICM workspace.

8.18 Sample Peptide

To model a peptide using the ICM Biased Probability Monte Carlo method:

- MolMechanics/Sample Peptide
- Enter a name for the simulation.
- Enter the one letter amino acid code for the peptide.
- Choose whether the peptide is cyclic or not.
- Determine the state of the N-terminus and C-terminus.
- Add a value for the effort. This represents the length of the simulation - values between 1 and 10 will give a thorough sampling. A larger effort value will be needed for longer peptides.
- When the simulation is finished a stack of conformations will be displayed in the ICM workspace.

8.19 Molecular Mechanics

Available in the following product(s): ICM-Pro

8.19.1 ICM Convert

To calculate energy, build a molecular surface and for all energy operations you need to convert a PDB file into an ICM object.

- MolMechanics/ICM-Convert/Protein
- Select the object you want to convert from the drop down list.
- Check or uncheck the options, delete water, optimize hydrogens, replace the original, and/or display the result.

To convert a small molecule into an ICM object.

- MolMechanics/ICM-Convert/Chemical
- Select the ligand.
- Choose whether you want to keep the current geometry of the ligand or not.
- Check or uncheck the options build hydrogens, fix amide bonds, and/or overwrite geometry.

8.19.2 Optimize H,His,Asn,Gln,Pro

This option optimizes H, His, Asn, Gln, and Pro by maximizing hydrogen bonds and other interactions with the rest of the protein and/or with the ligand.

To perform this optimization

- Convert your protein to an ICM Object.
- Select MolMechanics/Optimize, H, His,Asn,Gln and Pro

- Choose whether you want to sample rotatable hydrogens or optimize His, Asn and Gln.

8.19.3 Regularization

This option is described in detail in the modeling chapter [here](#).

8.19.4 Impose Conformation

If you have two protein structures with the same atom names and ALTER records but with different conformations you can impose the conformation of one of the protein structures onto the other.

You can do this by:

- Convert the structures into ICM Objects.
- MolMechanics/Impose Conformation
- Select the source molecule (the conformation you wish to impose).
- Optional: superimpose the structures
- Optional: re-optimize hydrogens

8.19.5 Edit Structure

Set Bond Type

- Select the two atoms forming the bond.
- MolMechanics/Edit Structure/Set Bond Type
- Choose the **Bond Type** from the drop down arrow.
- Press Apply button

Set Formal Charge

- Select the atom.
- MolMechanics/Edit Structure/Set Formal Charge
- Choose the **Charge** from the drop down arrow.
- Press Apply button

Set Chirality

- Select the atom.
- MolMechanics/Edit Structure/Set Chirality
- Choose the **Chirality** from the drop down arrow.
- Press Apply button

Build Hydrogens

- Select the atoms.
- MolMechanics/Edit Structure/Build Hydrogens
- Press Apply button

Set Tether

Theory

A tether is a harmonic restraint pulling an atom in the current object to a static point in space. This point is represented by an atom in another object. Typically, it is used to relate the geometry of an ICM molecular object with that of, say, an X-ray structure whose geometry is considered as a target. Tethers can be imposed between atoms of an ICM-object and atoms belonging to another object, which is static and may be a non-ICM-object. You cannot create tethers in ICM-Browser, however, if the project that you have loaded contains tethers between two objects, then they can be displayed:

- Convert the two structures you wish to tether to an ICM object.
- MolMechanics/Edit Structure/Set Tether
- Right click on the first atom you wish to tether and click on the first option which is the selection language for the atom. This information will automatically be placed in the **first atom** dialog box. Alternatively you can type in the ICM selection language into the dialog box.

- Right click on the second atom (in a different object) you wish to tether and click on the first option which is the selection language for the atom. This information will automatically be placed in the **second atom** dialog box. Alternatively you can type in the ICM selection language into the dialog box.
- Press Apply button

Delete Tether

- Select the atoms you wish to delete the tethers from.
- MolMechanics/Edit Structure/Delete Tether

8.19.6 MMFF

Set Types This option is described in detail here in the command line manual
<http://www.molsoft.com/man/icm-commands.html#set-type-mmff>

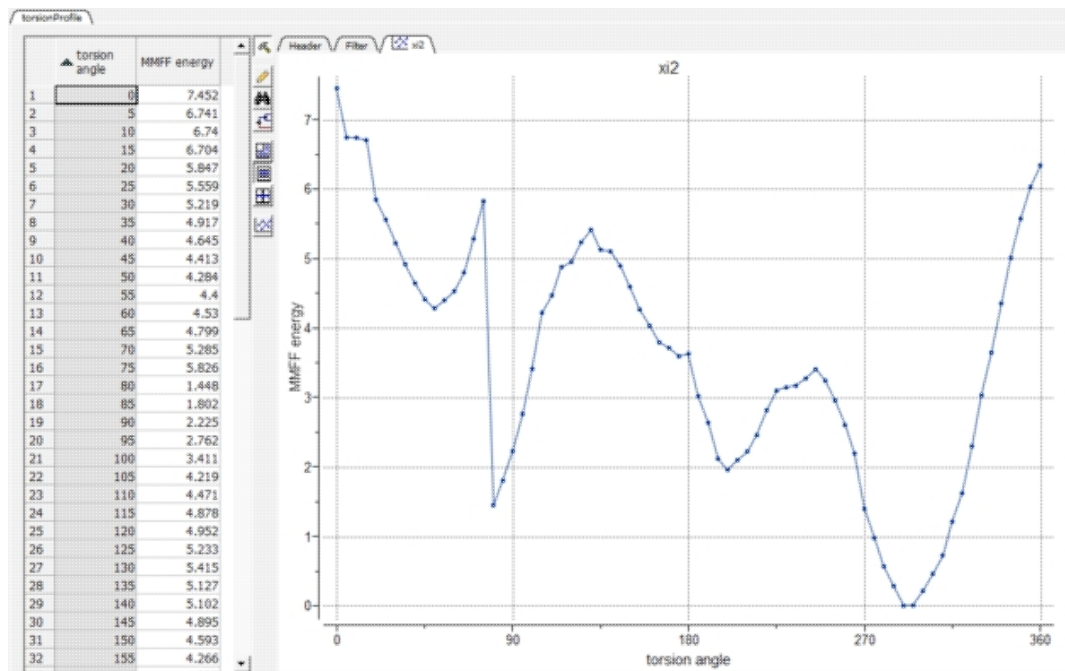
Set Charges This option is described in detail here in the command line manual
<http://www.molsoft.com/man/icm-commands.html#set-chargemmff>

Read Libraries This option is described in detail here in the command line manual
<http://www.molsoft.com/man/icm-commands.html#read-librarymmff>

8.19.7 Torsion Scan

To sample the torsion angles.

- Read in or convert your molecule to an ICM object.
- Select the torsion angle. A convenient way of doing this is with the `select atom` selection tool.
- MolMechanics/Torsion Scan
- Enter the step in degrees you would like to sample.
- A table and plot will be displayed with the MMFF energy and angle.



8.19.8 Minimize

Cartesian This option is described in detail here in the command line manual
<http://www.molsoft.com/man/icm-commands.html#minimize-cartesian>

Local This option is described in detail here in the command line manual
<http://www.molsoft.com/man/icm-commands.html#minimize>

Global This option is described in detail here in the command line manual
<http://www.molsoft.com/man/icm-commands.html#montecarlo>

8.19.9 Sample Loop

This option is described in the `Loop Modeling` section.

8.19.10 Design Loop

This option is described in the `Design Loop` section.

8.19.11 Sample Protein

This option is described earlier in this chapter here.

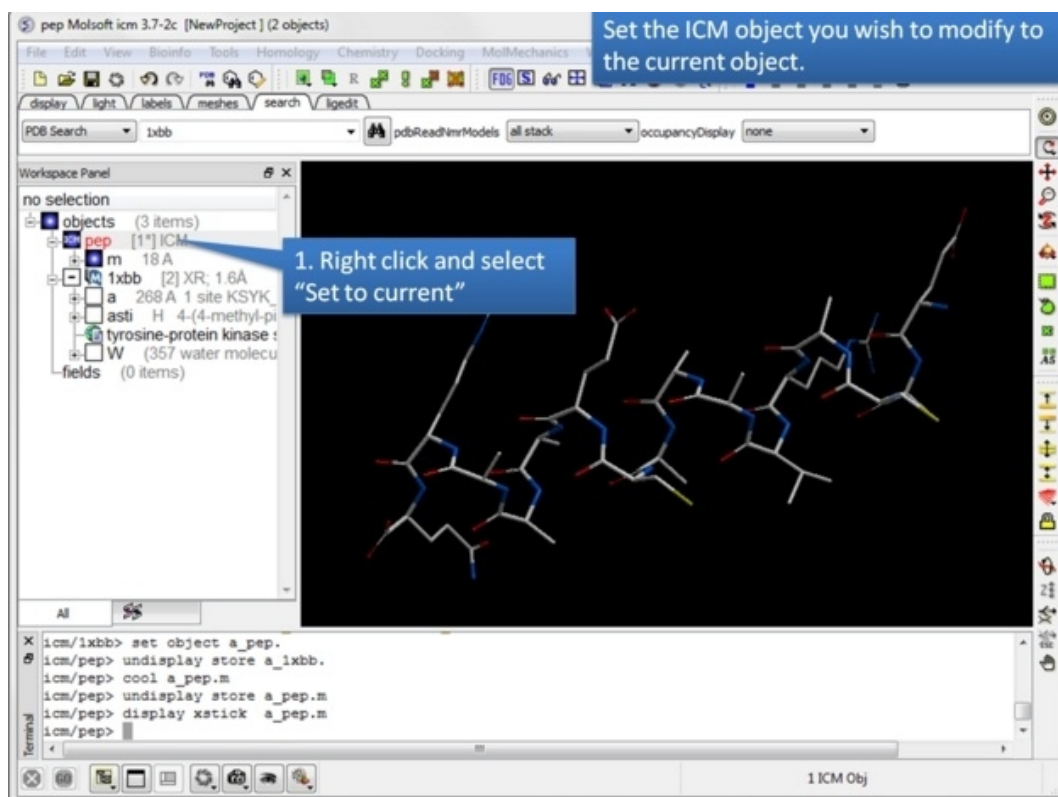
8.19.12 Sample Peptide

This option is described earlier in this chapter here.

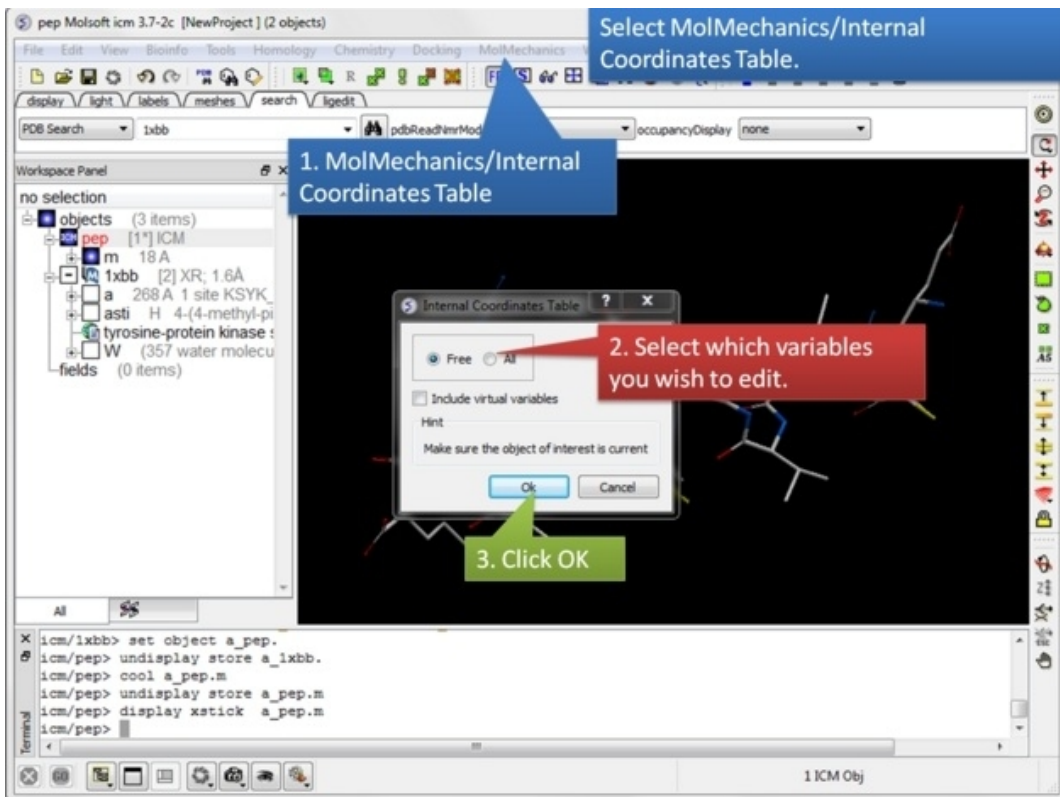
8.19.13 Internal Coordinates Table

An easy way to set values of internal coordinates is to edit the Internal Coordinates Table directly in the GUI. Variables are discussed in more detail here:

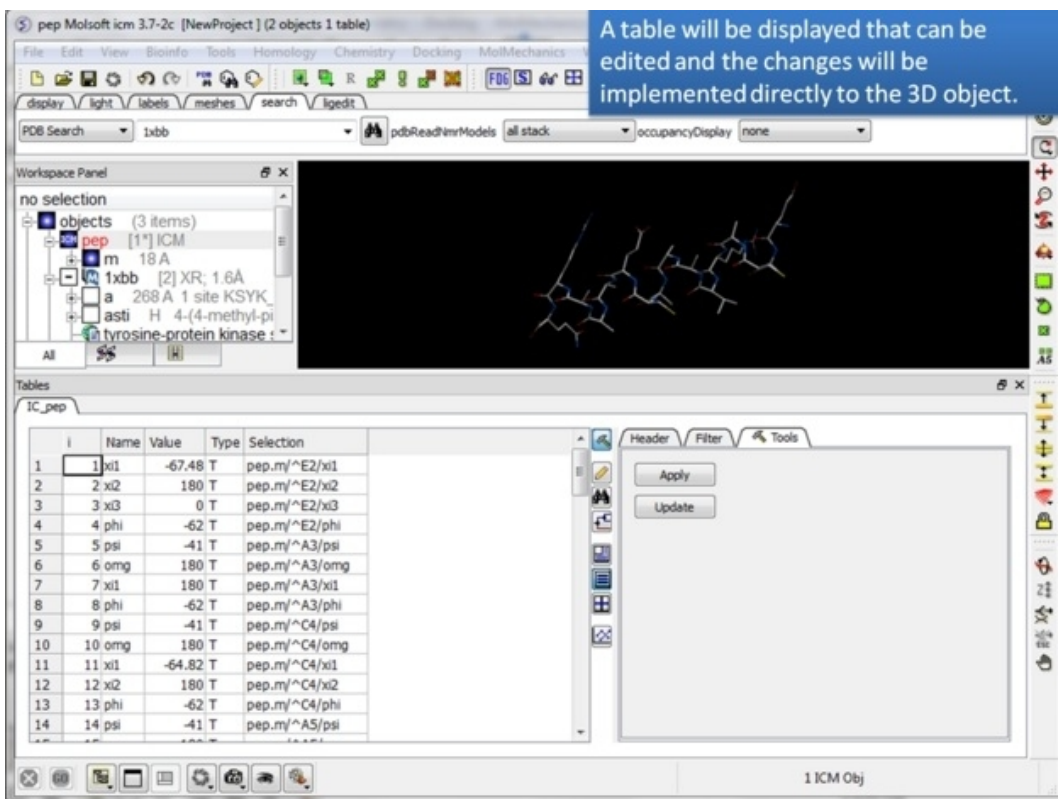
http://www.molsoft.com/man/selection.html#vs_



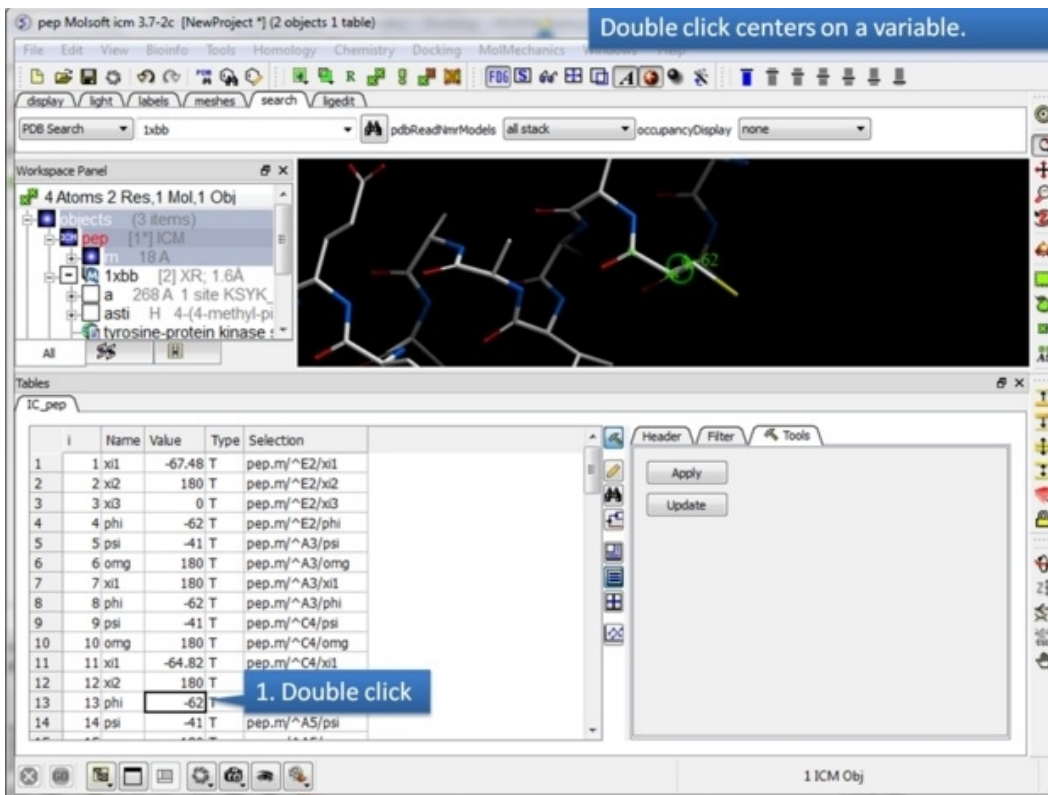
Step 1: The object you wish to edit has to be the **Current Object**.



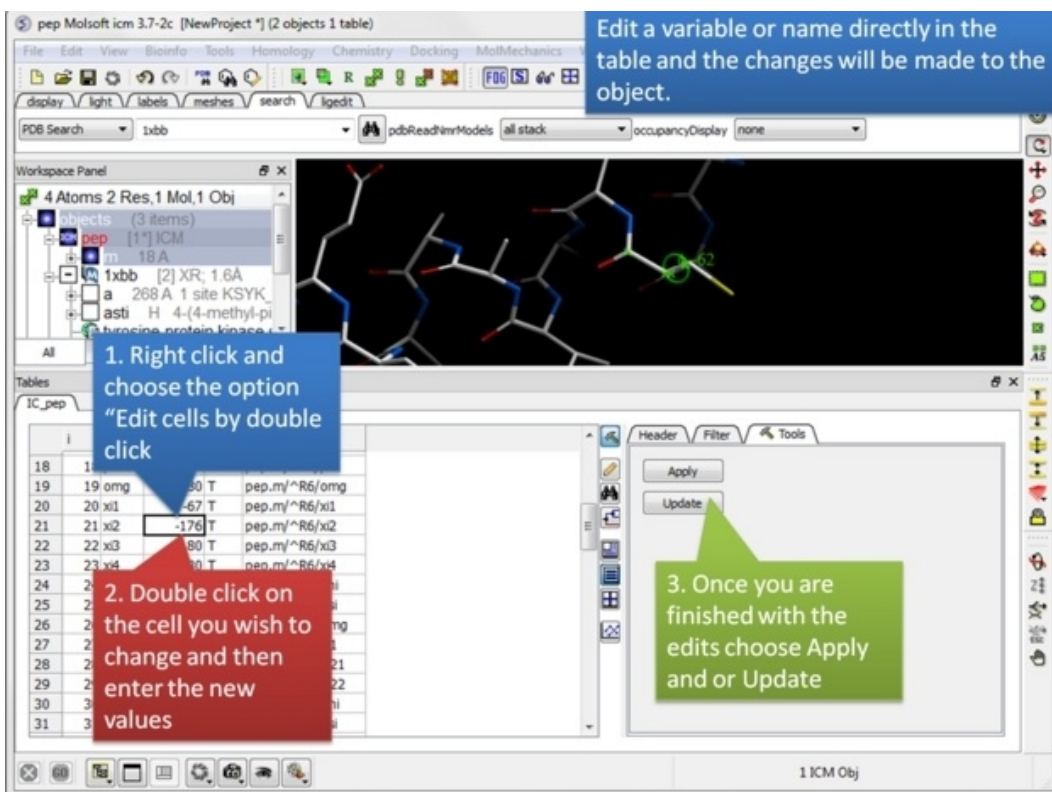
Step 2: Select MolMechanics/Internal Coordinates Table. Choose whether you wish to edit just the Free or All and include virtual variables.



Step 3: A table will be displayed containing the variable Name, Value, Type and Selection syntax.



Step 4: To center on a variable - double click on a row in the table.



Step 5: Edit a variable directly in the table and apply the change.

8.19.14 Generate Normal Mode Stack

Normal modes can be used to generate an ensemble of protein structures. For example the method can be used to represent flexibility in the pocket.

To generate an ensemble of structures using normal modes.

- Convert your protein to an ICM object.
- MolMechanics/Generate NM stack
- Enter the number of normal modes to sample
- Enter the relative amplitude of the normal modes.
- Optional: select to make random combination of modes.
- Optional: select Fast GAP model only.
- Optional: run the normal mode generation locally.

8.19.15 Stack

Operations which use the ICM Biased Probability Monte Carlo method e.g. docking and loop modeling generate a stack of energy conformations.

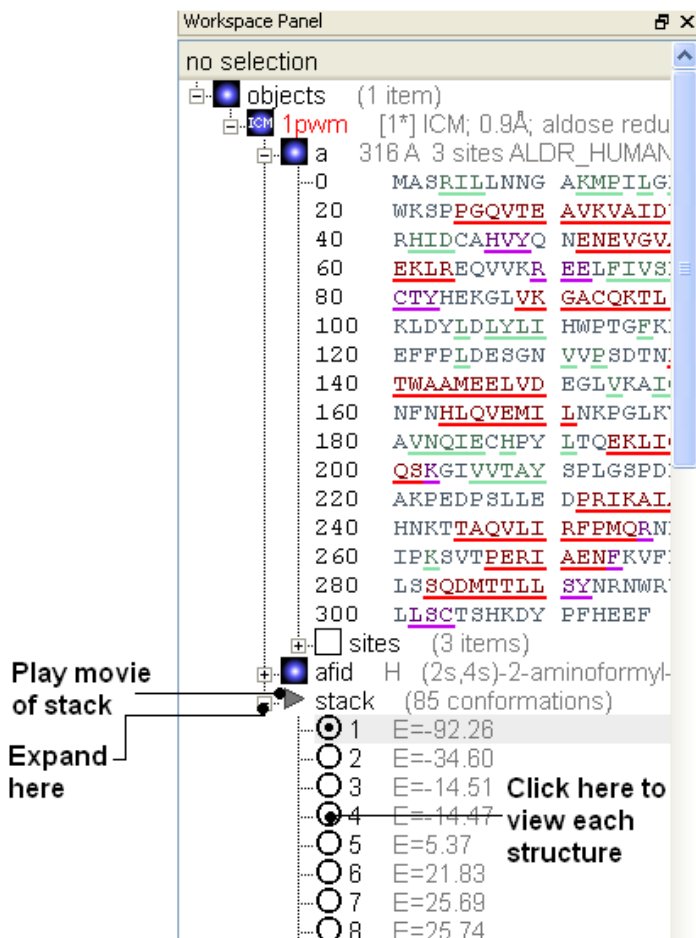
MolMechanics/Stack/View will display the conformations of a stack in a table ranked by energy. Each conformation can be viewed by double clicking on the table. A stack file will have the extension `.cnf`. For example, after running the `sample loop` algorithm a stack of different loop conformations will be generated.

MolMechanics/Stack/Play This option will play the elements of the stack as a movie. You can set the number of frames for the movie and also select whether you would like ICM to interpolate between each frame. You can save this movie in avi, mpeg format using the `Screen-Grabbing Movie` options.

MolMechanics/Stack/Add current conformation This option will add the currently displayed conformation to the stack. This is useful for experiments such as `multiple receptor docking` whereby you dock to a stack of conformations.

MolMechanics/Stack/Store Stack in Object This option takes the current stack and stores it in a compressed form inside the specified object. The compressed stack can then be extracted with the `load stack object` command. Option `stack` of the `montecarlo` command stores the generated stack inside the current object automatically.

To view the stack in gui:



MolMechanics/Stack/Delete Deletes the current stack.

MolMechanics/Stack/Set conf Comparison This option compares the stack as described here: <http://www.molsoft.com/man/preference.html#compareMethod> and <http://www.molsoft.com/man/icm-commands.html#compare>

MolMechanics/Stack/Recalculate Energies Recalculates the energy of a current stack if changes have been made.

8.19.16 GAMESS

This option is described in detail here in the command line manual <http://www.molsoft.com/man/E-H.html#gamess>

8.19.17 Waters

'Flood' procedure populates a box around the region of interest on the protein with water molecules, generating a stack of low-energy water configurations. 'Quick flood' performs the procedure within the interactive ICM session. For more thorough sampling, it is recommended to use 'Sample Flood' that runs the same procedure in the background.

Try an example:

- MolMechanics/Waters/Load Example
- Click on the html link to generate a box around the region of interest. You should see a purple box - the size of it can be changed by dragging on the corners of the box. In a real example you can make a selection and then in the MolMechanics/Waters/Sample Flood dialog box there is a button "Make Box Around Selection".
- MolMechanics/Sample Flood

- Enter the name of the protein.
- Enter the number of water configurations you would like to generate.
- Choose whether to make the water density map and/or contour.
- Click OK.
- A stack of low energy water molecule configurations will then be displayed in the ICM Workspace.

8.19.18 Energy Terms

The energy function calculated for any conformation of an ICM molecular object consists of individual terms described which can be turned on and off using **MolMechanics/Energy Terms**. These terms are described in more detail here <http://www.molsoft.com/man/terms.html>

9 Cheminformatics

Chapter Contents:

- Read Chemical Format Files
- Save Chemical Structures and Images
- Chemical Spreadsheets (Chemical Tables)
- Sketch Chemicals – Molecular Editor
- Chemical Search
- Convert Chemicals to 3D
- ToxScore
- Generate Chemical Fragments
- MolCart – Chemical Cartridge
- Calculate Properties
- Standardize Table
- Annotate Chemical by Substructure
- Align and Color by 2D Scaffold
- Set Formal Charges
- Enumerate Formal Charge States
- Convert
- Build Prediction Model
- Predict using Model
- Generate 3D Conformers
- Generate Tautomers
- Generate Stereoisomers
- Calculate Ligand Energetics
- Chemical Clustering
- PCA Analysis
- Compare Two Sets (tables)
- Merge Two Sets (tables)
- Select Duplicates
- Combinatorial Chemistry
- Chemical Superposition
- APF Superposition
- APF Tools

9.1 Read Chemical Structures

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

Chemical structures can be read into ICM from MOL/MOL2, SMILES, and SDF files or you can construct your own structures by drawing them in the ICM Molecular Editor. Chemicals can also be searched and downloaded via the Search Tab using the PDB Chemical Substructure Search, search by PDB Ligand Code, or by searching DrugBank.

To read a chemical format file:

- Select File/Open.
- Select the chemical structure file you wish to open: MOL, MOL2 or SDF.
- Once selected the file will be displayed as a chemical table (See ICM molecular tables section).

If you know the chemical smiles string for the compound you can build the 2D molecule by:

- Select File/New.
- Click the **Compound** tab at the top of the window.
- Enter a name for the compound.
- Type in the **Smiles String** in the Smiles String data entry box. Remember to delete the previous string.
- Check the boxes Display Molecule Delete Other Objects according to your preference.
- Click the OK button.

Smiles can be read from a text file into a chemical table by:

- **File/Open** and select **Files of type:** Smiles format

Smiles can be read directly into the **ICM Molecular Editor**:

- Open the **ICM Molecular Editor** window.
- Select **Edit/Add Smiles**

Chemicals can be read directly from **ChEMBL, SureChEMBL, PubChem, and DrugBank**

- Click on the `search` tab and find the database you wish to search in the drop down dialog box.

9.1.1 Extract 2D structure from PDB Chemical

To extract a 2D chemical representation from a PDB file into a chemical spreadsheet.

- Read the PDB file.
- Right click on the ligand and choose **Extract Ligand**.
- Choose whether you want to extract the 3D coordinates.
- Choose whether you want to append to an existing table (chemical spreadsheet).

9.1.2 Extract 3D coordinates from 3D Chemical.

To extract the 3D coordinates of a chemical into a chemical spreadsheet e.g for 3D QSAR or after chemical superposition.

- Select the chemicals in the ICM workspace. One or more can be selected by clicking holding the CTRL key and clicking. A range of chemicals can be selected by holding the keyboard tab and clicking.
- Right click on any of the selected chemicals in the ICM workspace and choose **Extract Ligand**.
- Select "keep 3D coordinates."
- Append to an existing table if needed.

9.2 Save Chemical Structures and Images

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

There are a variety of ways to save chemical structures. Chemicals can be saved in mol, sdf and smiles format from a chemical table (spreadsheet), molecular editor or from the ICM-Workspace. An image of the 2D sketch can be saved as an image from a chemical table.

9.2.1 Saving from a chemical table.

To save all the chemicals in a table as an SDF file:

- Right click on the chemical table header.
- Select **Save as..**
- Select **Save as type:** (SD file or Mol file_

Right click here

	mol	MolWeight	MolPSA	MolLogS
1		7.226	157.205	-1.44554
2		312.29	114.926	-4.43001

To save selected chemicals in a table as an SDF file:

- Select the row(s) of the chemicals you wish to save in SDF format. Row selections in tables is described in the Tables chapter.
- Right click on any of the selected rows and select **Save Selection As...**

Select rows here

	mol	MolWei
1		chiral 347.226
2		312.29
3		chiral 253.219

Right click here

9.2.2 Saving Chemicals in the Molecular Editor

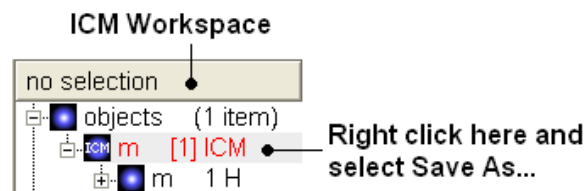
Chemicals drawn in the ICM Molecular Editor can be saved:

- **File/Save** or **File/Save As...** or **Edit/Copy as SMILES**

9.2.3 Saving Chemicals in the ICM Workspace

If you have converted a chemical sketch into 3D, the 3D structure will be displayed in the ICM Workspace. To save this structure in mol format:

- Right click on the name of the chemical in the ICM Workspace.
- Select **Save As..**

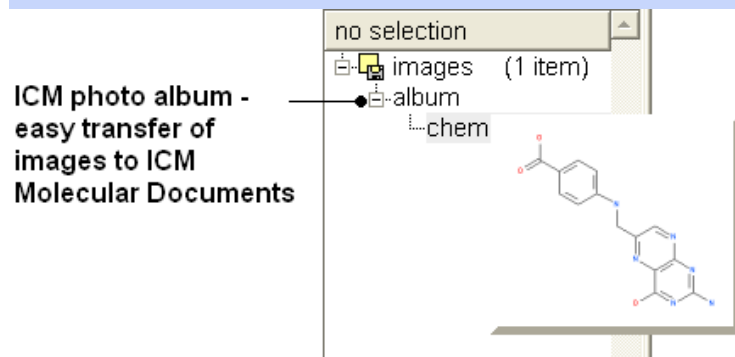


9.2.4 Saving Chemical Images

To save an image of a chemical sketch:

- The chemical needs to be displayed in a chemical table.
- Right click on the 2D image
- Select **Save Molecule As Image**

NOTE Using the right click options you can also save the image to the **clipboard** or copy the molecule to the ICM Image Album for use in Molecular Documents.



9.2.5 IUPAC Name

The IUPAC nomenclature for a compound can be generated on the fly for a chemical. You can view the IUPAC name in the Molecule Editor or you can insert a column into a chemical spreadsheet with the IUPAC name.

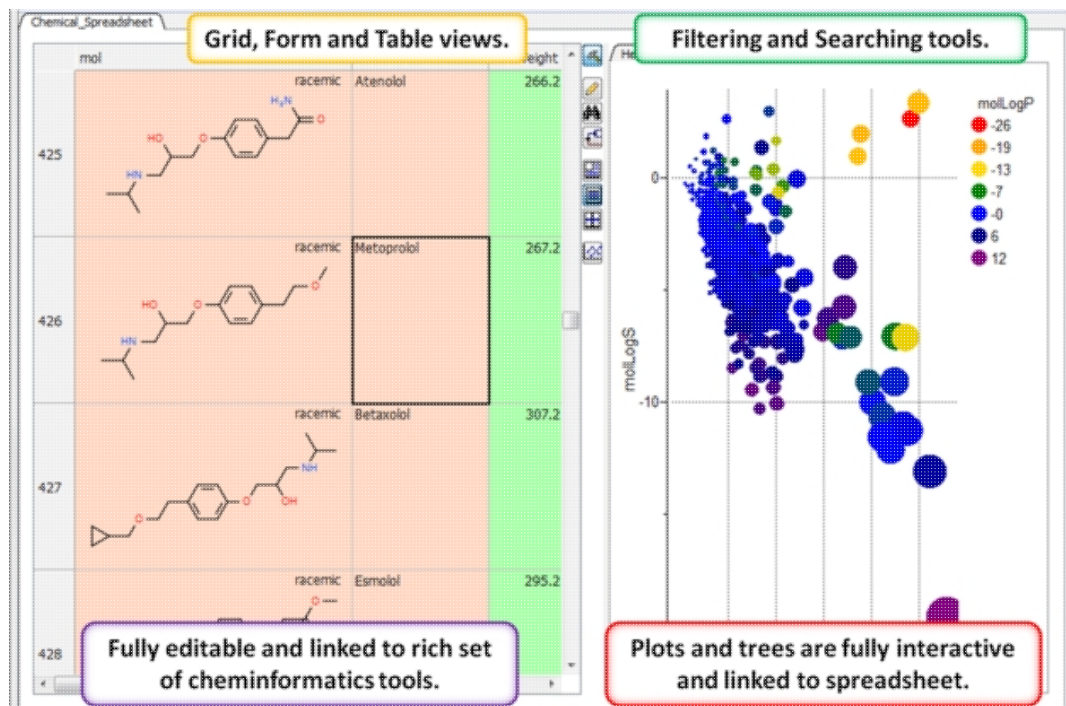
9.3 Working with Chemical Spreadsheets.

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

When an sdf file is read into ICM it is displayed as a chemical spreadsheet. Many of the operations you can perform on chemical spreadsheets (Molecular Tables) are described in the table section of this manual.

Some useful chemical-only options are described in this section.

An example of an ICM molecular table:



9.3.1 Molecular Table Display

There are three built-in table views.

- Table View (Default)
- Grid View - select the columns you wish to display in a grid.
- Form View

You can toggle between each view using the buttons on the right hand side of the table shown below.

Toggle View

molLogS molPSA molArea

2	444.9	2.659	-4.313	75.59	498
5	410.7	0.3258	-3.092	106	452.9

Form
Grid
Table

Table View Default

mol	Molecule Name	molLogS	molPSA	molArea	molWeight
355	Trandapt	439.2	9	2	444.9
356	Lanzop	403.2	10	5	410.7
357	Quemp	439.2	9	2	443

Grid View

Form View

9.3.2 How to add columns into a chemical spreadsheet.

How to add columns into a chemical spreadsheet.

1. Right click on a column header and select "insert column".
2. Click the main type of argument you want to add. E.g. mol for chemical functions.
3. Press the "Add To List" button
4. Select the argument (s)

9.3.3 How to sort a column(s) in a chemical spreadsheet.

How to sort a column(s) in a chemical spreadsheet.

1. Right click on a column header and select "Sort"

mol	NAME	molwt	molclogp	molrd	molh	molhlogp	molhlogp2	molhlogp3	molhlogp4
1	<chem>NC(=O)C1=CC=CC=C1</chem>	177.15	1.41	1.91	1.91	1.91	1.91	1.91	1.91
2	<chem>NC(=O)C1=CC=CC=C1C</chem>	193.16	1.56	2.06	2.06	2.06	2.06	2.06	2.06
3	<chem>NC(=O)C1=CC=CC=C1N</chem>	179.16	1.56	2.06	2.06	2.06	2.06	2.06	2.06
4	<chem>NC(=O)C1=CC=CC=C1C=O</chem>	191.17	1.56	2.06	2.06	2.06	2.06	2.06	2.06
5	<chem>NC(=O)C1=CC=CC=C1C(=O)O</chem>	179.17	1.56	2.06	2.06	2.06	2.06	2.06	2.06
6	<chem>NC(=O)C1=CC=CC=C1C(=O)O</chem>	179.17	1.56	2.06	2.06	2.06	2.06	2.06	2.06

Info: temp_a_totl_temp_a_molclogp_max_a_totl_temp_a_totl_disp_l_1_tableOutput_1_joins_1_selection_1_selection2Term_1_display_1_delete_h_as_condit_1_as_attach temp_variables deleted
 low/low delete all
 low/low undisplay window
 low/low _currentProject = ""
 low/low openFile "C:\vswd\molstructure.mdf"
 Info: table "substructure" (0 headers, 8 arrays(10000)) created
 low/low

9.3.4 How to copy, cut and paste columns and rows in a chemical spreadsheet.

How to copy, cut, and paste columns and rows in a chemical spreadsheet.

1. Select a column or row by clicking on the header. (click and hold CTRL key for non-contiguous selections or SHIFT key for range)

2a. To copy, cut and paste rows – right click here on a selected row(s)

2b. To copy, cut and paste rows – right click here on a selected column(s)

mol	NAME	molwt	molclogp	molrd	molh	molhlogp	molhlogp2	molhlogp3	molhlogp4
1	<chem>NC(=O)C1=CC=CC=C1</chem>	177.15	1.41	1.91	1.91	1.91	1.91	1.91	1.91
2	<chem>NC(=O)C1=CC=CC=C1C</chem>	193.16	1.56	2.06	2.06	2.06	2.06	2.06	2.06
3	<chem>NC(=O)C1=CC=CC=C1N</chem>	179.16	1.56	2.06	2.06	2.06	2.06	2.06	2.06
4	<chem>NC(=O)C1=CC=CC=C1C=O</chem>	191.17	1.56	2.06	2.06	2.06	2.06	2.06	2.06
5	<chem>NC(=O)C1=CC=CC=C1C(=O)O</chem>	179.17	1.56	2.06	2.06	2.06	2.06	2.06	2.06
6	<chem>NC(=O)C1=CC=CC=C1C(=O)O</chem>	179.17	1.56	2.06	2.06	2.06	2.06	2.06	2.06

Info: temp_a_totl_temp_a_molclogp_max_a_totl_temp_a_totl_disp_l_1_tableOutput_1_joins_1_selection_1_selection2Term_1_display_1_delete_h_as_condit_1_as_attach temp_variables deleted
 low/low delete all
 low/low undisplay window
 low/low _currentProject = ""
 low/low openFile "C:\vswd\molstructure.mdf"
 Info: table "substructure" (0 headers, 8 arrays(10000)) created
 low/low

9.3.5 How to show and hide columns and rows in a chemical spreadsheet.

How to hide rows in a chemical spreadsheet.

1. Select row(s) (click and hold CTRL key for non-contiguous selections or SHIFT key for range)

2. Right click here and select "Hide Rows"

NAME	mol	mol(MW)	mol(MW)	mol(MW)	mol(MW)	mol(MW)	mol(MW)
C13H18N2O	104801	201.1	6.529	-3.096	34.46	N(C1=CC=C(C=C1)C(=O)CC=C(C=C1)C	2
C13H18N2O	51052	212.1	1.912	-0.61	27.46	C1C=ANC=1C1C=C(C=C1)C(=O)CC=C(C=C1)C	1
C13H18N2O	153405	214.1	1.896	-3.367	28.46	N(C1=CC=C(C=C1)C(=O)CC=C(C=C1)C	2
C13H18N2O	104801	214.1	2.026	-3.689	28.13	C1C=ANC=1C1C=C(C=C1)C(=O)CC=C(C=C1)C	2

How to hide columns in a chemical spreadsheet.

1. Select column(s) (click and hold CTRL key for non-contiguous selections or SHIFT key for range)

2. Right click here and select "Hide Column(s)"

NAME	mol	mol(MW)	mol(MW)	mol(MW)	mol(MW)	mol(MW)	mol(MW)
C13H18N2O	104801	201.1	6.529	-3.096	34.46	N(C1=CC=C(C=C1)C(=O)CC=C(C=C1)C	2
C13H18N2O	51052	212.1	1.912	-0.61	27.46	C1C=ANC=1C1C=C(C=C1)C(=O)CC=C(C=C1)C	1
C13H18N2O	153405	214.1	1.896	-3.367	28.46	N(C1=CC=C(C=C1)C(=O)CC=C(C=C1)C	2
C13H18N2O	104801	214.1	2.026	-3.689	28.13	C1C=ANC=1C1C=C(C=C1)C(=O)CC=C(C=C1)C	2

9.3.6 How to save a chemical spreadsheet in sdf format.

How to save a chemical spreadsheet in sdf format.

1. Right click on the spreadsheet header tab

2. Select "Save As.."

NAME	molwt	molhlogp	molhlogp2	molhlogp3	molhlogp4	molhlogp5	molhlogp6
C13H12N2	202.96	172.1	1.91	-0.57	28.11	NC1=CC=C(OB1=CC=CC=C1)C=C1	2
C13H12N2	168.99	200.1	1.625	-0.214	28.13	C1=CC=C(C=C1)C=C1C=C1	2
C13H12N2	198.92	212.1	1.912	-0.61	27.46	C1=CC=C(C=C1)C=C1C=C1C=C1	2
C13H14N2	158.95	214.1	1.886	-0.307	28.48	NC1=CC=C(C=C1)C=C1C=C1	2
C13H14N2	168.97	214.1	2.026	-0.603	28.13	C1=CC=C(C=C1)C=C1C=C1C=C1	2

9.3.7 How to export your chemical spreadsheet into Excel.

How to export a chemical spreadsheet into Excel.

1. Right click on the spreadsheet header tab

2. Select "Export To Excel"

NAME	molwt	molhlogp	molhlogp2	molhlogp3	molhlogp4	molhlogp5	molhlogp6
C13H12N2	202.96	172.1	1.91	-0.57	28.11	NC1=CC=C(OB1=CC=CC=C1)C=C1	2
C13H12N2	168.99	200.1	1.625	-0.214	28.13	C1=CC=C(C=C1)C=C1C=C1	2
C13H12N2	198.92	212.1	1.912	-0.61	27.46	C1=CC=C(C=C1)C=C1C=C1C=C1	2
C13H14N2	158.95	214.1	1.886	-0.307	28.48	NC1=CC=C(C=C1)C=C1C=C1	2
C13H14N2	168.97	214.1	2.026	-0.603	28.13	C1=CC=C(C=C1)C=C1C=C1C=C1	2

9.3.8 How to print a chemical spreadsheet.

How to print a chemical spreadsheet.

1. Right click on the spreadsheet header tab

2. Select "Print Table"

NAME	molwt	molHeight	molLogP	molLogD	molPSA	index	ref ID#
C13H2N2O	201.96	172.1	1.91	-0.21	28.11	163	C16-09H-C16-C16-C16-C16
C13H2N2O	168.09	200.1	1.625	-0.214	28.13	116	4NE-10C16-C16-C16-C16
C13H2N2O	168.09	212.1	1.912	-0.61	27.46	116	4NE-10C16-C16-C16-C16
C13H2N2O	158.05	214.1	1.86	-0.367	28.48	116	4NE-10C16-C16-C16-C16
C13H2N2O	168.09	214.1	2.02	-0.68	28.13	116	4NE-10C16-C16-C16-C16

9.3.9 How to filter columns in a chemical spreadsheet.

How to filter columns in a chemical spreadsheet (1).

1. Right click on the column header you wish to filter

2. Select Filter Custom or From List...

mol	molwt	molHeight	molLogP	molLogD	molPSA	index	ref ID#
6.376	300	1.925	-4.68	52.96	202.2	96.4	101116
6.525	373.1	4.258	5.809	42.92	268.4	79.3	6397
6.6	317	2.206	-4.619	43.16	288.4	99.96	0.129
6.676	339	2.759	-5.176	42.62	288.4	99.96	0.2019

How to filter columns in a chemical spreadsheet (2).

2. This symbol indicates the column is filtered

1. Enter filters

To remove a filter right click on the column that is filtered and select Filter/Clear

mol	molwt	molweight	molLogP	molLogP	molSP					
1		471	4.775	-7.136	42.74	336	134.4	0.00017	3	
2		380.1	3.924	-6.126	42.74	314	128.1	0.00059	3	
3		6.375	301	1.925	-4.660	52.56	222.2	98.47	0.00176	3
4		6.525	317	2.206	-4.619	43.16	380.4	95.96	0.129	3
5		6.675	331	2.759	-5.176	42.43	380.4	95.96	0.2019	3
6		7.275	365.1	3.102	-5.125	44.16	380.4	95.96	0.2019	3

9.3.10 How to use find and replace in a chemical spreadsheet.

How to use find and replace in a chemical spreadsheet.

1. Right click on the mol column

2. Select Chemical Find and Replace

3. Use the editor to sketch the group you want to find

4. Use the editor to sketch the group you want to replace with

You can also do text find and replace by right clicking anywhere on the table and select Find and Replace.

mol	molwt	molweight	molLogP	molLogP	molSP					
1		471	4.775	-7.136	42.74	336	134.4	0.00017	3	
2		380.1	3.924	-6.126	42.74	314	128.1	0.00059	3	
3		6.375	301	1.925	-4.660	52.56	222.2	98.47	0.00176	3
4		6.525	317	2.206	-4.619	43.16	380.4	95.96	0.129	3
5		6.675	331	2.759	-5.176	42.43	380.4	95.96	0.2019	3
6		7.275	365.1	3.102	-5.125	44.16	380.4	95.96	0.2019	3

9.3.11 How to mark and label rows in a chemical spreadsheet.

How to mark and label rows in a chemical spreadsheet.

1. Select the row(s) you wish to label by clicking on the row number. Click and hold CTRL key for non-contiguous selections or SHIFT key for range.

2. Right click on the selected row and choose "Mark Selected Rows" Select a color label.

3. You can select marked (labeled) rows here.

row	molwt	molweight	molLogP	molTPSA	molHBA	molHBD	molTPSA	molHBA	molHBD
1	471	4.775	7.196	42.74	306	134.4	0.05017	3	1
2	6136	42.74	314	126.1	0.07900	3	1		
3	3.533	53.02	183.3	98.27	0.2883	3	2		
4	4.880	52.76							
5	2.206	4.615	43.16	281.4	95.96	0.125	3	1	
6	6.675	331							

9.3.12 How to insert hyperlinks to the PDB, PubMed, and Uniprot.

How to insert hyperlinks.

1. Right click on the column header and select "Format"

2. Choose hyperlink type here

3. Click OK and the hyperlink will be generated

row	molwt	molweight	molLogP	molTPSA	molHBA	molHBD	molTPSA	molHBA	molHBD	pdb
1	471	4.775	7.196	42.74	306	134.4	0.05017	3	1	pdb
2	6136	42.74	314	126.1	0.07900	3	1			pdb
3	3.533	53.02	183.3	98.27	0.2883	3	2			pdb
4	4.880	52.76								pdb
5	2.206	4.615	43.16	281.4	95.96	0.125	3	1		pdb
6	6.675	331								pdb

9.3.13 How to copy and paste 2D chemicals.

How to copy and paste 2D chemicals.

1. Right click on the 2D chemical.

2. Select "Copy 2D Molecule"

mol	mw	molweight	mollogP	mollogS	molPSA	molKow	molBP	fragKow	molHBA	molHBD	job	
1		6	471	4.775	-7.136	42.74	336	134.4	0.05017	3	13abc	
2		6	393.1	3.924	-6.126	42.74	314	126.1	0.07969	3	13abc	
3		6	311	1.376	-5.533	53.62	183.3	98.27	0.3593	3	23abc	
4		6	303	1.925	-4.866	52.56	222.2	98.47	0.007176	3	23abc	
5		6	311	4.295	-6.889	42.92	310.4	79.3	0.6937	3	13abc	
6		6	317	2.206	-4.619	43.16	265.4	99.96	0.129	3	13abc	
7		6	6475	331	2.759	-5.176	42.63	265.4	95.96	0.2819	3	13abc

9.3.14 How to edit data inside a chemical spreadsheet.

How to edit data inside a chemical spreadsheet.

1. Click here

2. Edit data by double clicking on the data.

mol	mw	molweight	mollogP	mollogS	molPSA	molKow	molBP	fragKow	molHBA	molHBD	job	
1		6	471	4.775	-7.136	42.74	336	134.4	0.05017	3	13abc	
2		6	393.1	3.924	-6.126	42.74	314	126.1	0.07969	3	13abc	
3		6	311	1.376	-5.533	53.62	183.3	98.27	0.3593	3	23abc	
4		6	6375				98.47	0.007176	3	23abc		
5		6	6525	373.1	4.295	-6.889	42.92	310.4	79.3	0.6937	3	13abc
6		6	66	317	2.206	-4.619	43.16	265.4	99.96	0.129	3	13abc
7		6	6475	331	2.759	-5.176	42.63	265.4	95.96	0.2819	3	13abc

9.3.15 How to remove salts, explicit hydrogens and standardize chemical groups.

How to remove salts, explicit hydrogens and standardize chemical groups.

1. Chemistry/Standardize

2. Select here

3. Select Files tab for batch mode

The screenshot shows the Chem3D interface with a spreadsheet of chemical data. The 'Standardize' dialog box is open, showing options for 'Remove Salts' and 'Remove Explicit Hydrogens'. The 'Files' tab is selected in the dialog. A callout box points to the 'Standardize' option in the 'Chemistry' menu.

9.3.16 How to calculate chemical properties in a chemical spreadsheet.

How to calculate chemical properties in a spreadsheet.

1. Chemistry/Calculate Properties

2. Select properties

3. Properties will be displayed in new columns

The screenshot shows the Chem3D interface with a spreadsheet of chemical data. The 'Calculate Chemical Properties' dialog box is open, showing a list of properties to be calculated. Callouts indicate the menu path and the selection of properties.

9.3.17 How to identify duplicate chemicals in a chemical spreadsheet.

How to identify duplicate chemicals in a spreadsheet.

1. Chemistry/Select Duplicates...
2. Enter table/spreadsheet name
3. Duplicate compounds will be selected and highlighted in blue

Table	Index	molwt	mollogP	molTPSA	molMR	molTPSA	molMR	molTPSA	molMR	
1	6	471	4.775	-7.196	42.74	336	134.4	0.00017	3	
2	6	283.1	3.924	-6.126	42.74	314	128.1	0.07969	3	
3	6	211.1	1.376	-3.933	63.02	183.3	88.27	0.000	3	
4	6	6.526	373.1	4.255	6.809	42.92	354.4	75.3	0.6907	3
5	6	381.1								3
6	6.6	317	2.206	-4.619	43.16	240.4	99.96	0.129	3	

9.3.18 How to compare two chemical spreadsheets.

How to compare two chemical spreadsheets.

1. Chemistry/Compare Two Sets...
2. Choose the tables you want to compare
3. Select the comparison method and output

Table	Index	molwt	mollogP	molTPSA	molMR	molTPSA	molMR	molTPSA	molMR	
1	6	471	4.775	-7.196	42.74	336	134.4	0.00017	3	
2	6	283.1	3.924	-6.126	42.74	314	128.1	0.07969	3	
3	6	211.1	1.376	-3.933	63.02	183.3	88.27	0.000	3	
4	6	6.526	373.1	4.255	6.809	42.92	354.4	75.3	0.6907	3
5	6	381.1								3
6	6.6	317	2.206	-4.619	43.16	240.4	99.96	0.129	3	

9.3.19 How to merge two chemical spreadsheets.

The screenshot shows the ChemBio3D interface with a table of chemical data. A 'Merge Two Sets' dialog box is open, showing 'Table A' and 'Table B' both set to 'binding_data'. The 'Merge' dropdown is set to 'left'. A 'Chemistry' menu is open, with 'Merge Two Sets' highlighted. Three callout boxes provide instructions: 1. 'Chemistry/Merge Two Sets...', 2. 'Choose the tables you want to compare', and 3. 'Select the way you want to merge them'. The bottom console shows commands like 'select index = Index(binding_data.tbl) exact stereo'.

9.3.20 Display and Convert Molecule

To display and convert a molecule from a molecular table in the 3D graphics display window:

- Select the molecule image or images in the molecular table.
- Right click and select the Chemistry/ Convert to 3D option.

The screenshot shows a 3D graphics window with a molecule selected in a table. A context menu is open, showing options like 'Edit DB Record...', 'Query Molecule', 'Chemistry', 'Copy Row(s)', 'Paste', and 'Print Table...'. The 'Chemistry' menu is expanded, showing options like 'Enumerate Stereoisomers', 'Enumerate Tautomers', 'Remove Salt', 'Convert to 3D and Optimize', 'Load and Preserve Coordinates', 'Conformation Generator', and 'Chemical Template Superposition'. The 'Convert to 3D and Optimize' option is highlighted.

9.3.21 Copy Molecule

To copy a molecule to paste into another application or into the ICM Molecular Editor:

- Right click on the molecule and a menu will be displayed.
- Select the option "Copy Molecule"

To copy a molecule or image to paste into another row within an ICM table or into the ICM Molecular editor:

- Right click on the molecule and a menu will be displayed.
- Select the option "Copy Molecule"
- Right click in the cell into which you wish to paste the molecule.
- Select the option "Paste Molecule"

NOTE: To learn how to insert a row read the `insert row` section.

Molecules drawn in ICM can be cut and pasted into ISIS-Draw and molecules from ISIS-Draw can be cut and pasted into ICM.

To perform these functions the correct settings need to be turned on in ISIS-Draw so that the compound drawing is saved in the clipboard.

IN ISIS-DRAW - Go to Options/Settings/General/Copy Mol Rxn file to the clipboard.

Compounds drawn in ISIS Draw can be cut and pasted into the ICM Molecular Editor and into ICM tables. Compounds can be copied in ICM by:

- Right click on the compound in the chemical table and select Copy Molecule.

9.3.22 Edit Molecule

To edit a molecule:

- Right click on the molecule and a menu will be displayed.
- Select the option **Edit Molecule** and the ICM Molecular Editor will be displayed.
- Edit the molecule.
- Click Exit in the ICM molecular editor.

9.3.23 Color Table Column

You can color your table based on values within a column by: You can also color the compound according to specific values see `Color Chemical Structure`.

- Selecting the column.
- Right click on the column header and a menu will be displayed.
- Select the option "Color By"

NOTE: You can remove the color from the table by right clicking on the column header selecting **Clear Selection Color**.

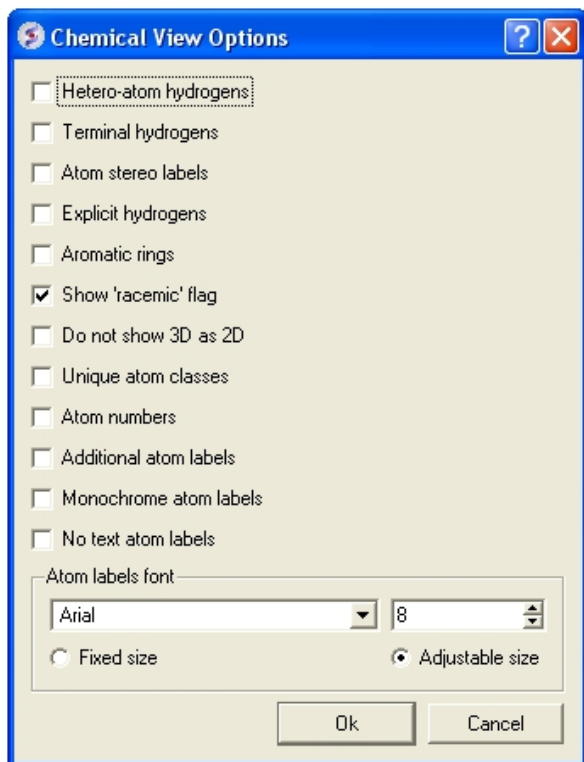
Links to other coloring options:

- Rows can be colored by marking them as described [here](#).
- A 2D chemical sketch can be colored according to specific values (e.g. pharmacophore features) see `Color Chemical Structure`.

9.3.24 Chemical View Options

Different chemical view options in the ICM molecular table can be set.

- Right click on the "mol" column header.
- Select Chemical view options... and the following data entry box will be displayed



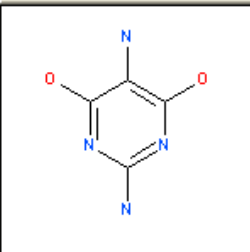
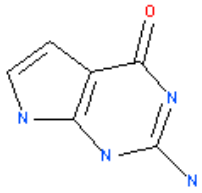
Options can be changed by checking the appropriate boxes or by entering the desired font and size.

9.3.25 Chemical Table Side-by-Side View

Chemical tables can be visually compared by placing them side-by-side. This can be done by:

- Double click on the table header for side-by-side view.
- Double click on the table header again to remove side-by-side view.

Double click here for "side-by-side" view

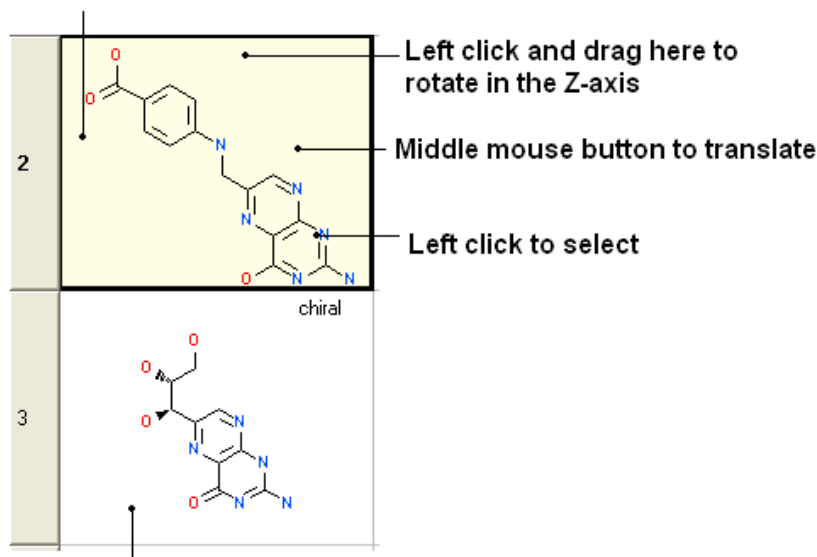
ricinLigands2D		ricin_inhibitors		
	mol	MolWeight	MolLogP	MolLogS
1		142.118	-0.589686	-1.70292
2		150.141	-0.070759	-2.11729

9.3.26 Zoom, Translate and Z-rotate a Chemical in a table.

Sometimes you may want to get a better view of a chemical in an ICM molecular table you can do this inside the chemical table by:

- Double clicking on the chemical drawing and the background will turn yellow.

Left click and drag here to zoom



Double click for yellow background to enable changes in chemical display

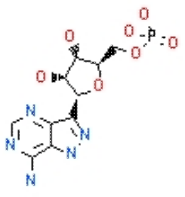
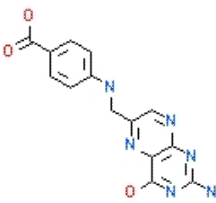
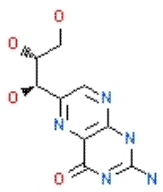
NOTE: You can also browse your structures in 3D. See section entitled "Set Chemical Table 3D Browse Mode".

9.3.27 Set Chemical Table 3D Browse Mode

To browse the chemicals contained within a Chemical Spreadsheet in the graphical display:

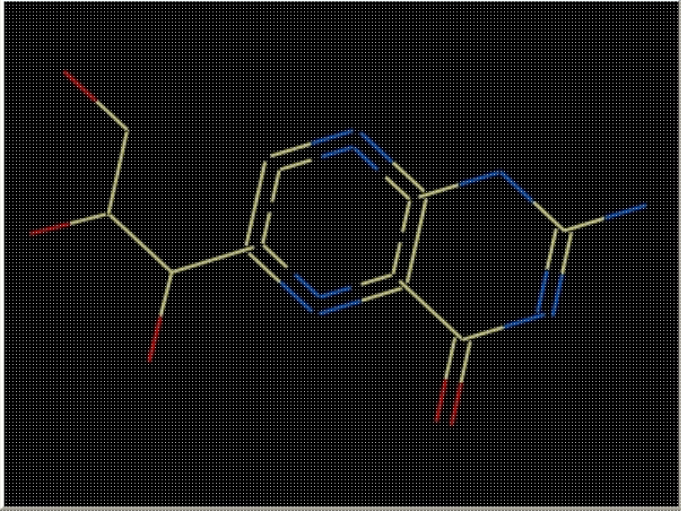
- Right click on the "mol" column header.
- Select the option "Set 3D Browse Mode".
- Use the box in the column "L" to toggle the 3D display on and off.
- Multiple chemicals can be displayed by selecting the rows and then checking the box.

ricinLigands2D TAUTOMERS

	mol	L	NAME
1	 <p>chiral</p>	<input checked="" type="checkbox"/>	
2		<input type="checkbox"/>	
3	 <p>chiral</p>	<input checked="" type="checkbox"/>	

Double click here to rotate the molecule in 3D in the cell.

3D Graphics



Toggle the display of the chemical in the 3D graphics window using the check box.

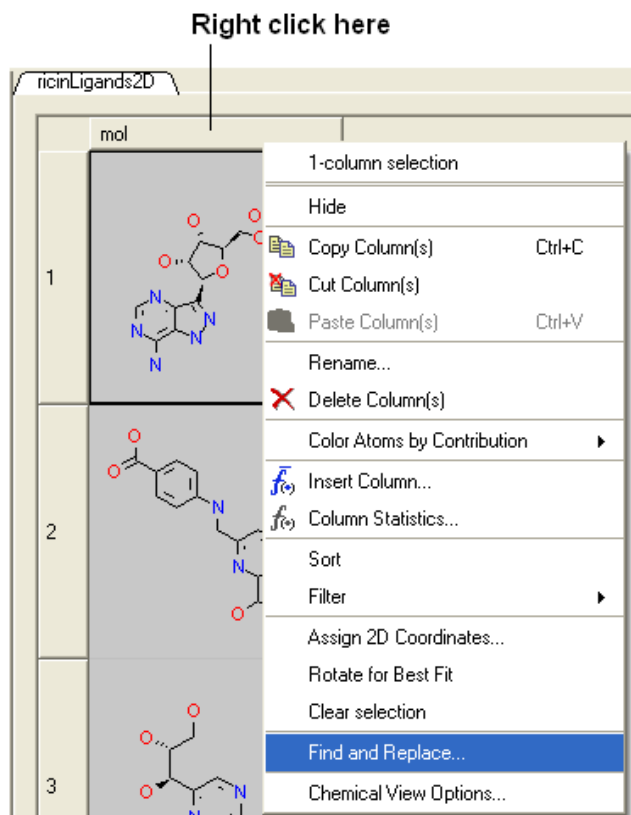
To remove 3D browse:

- Right click on the "mol" column header.
- Select UnSet 3D Browse Mode.

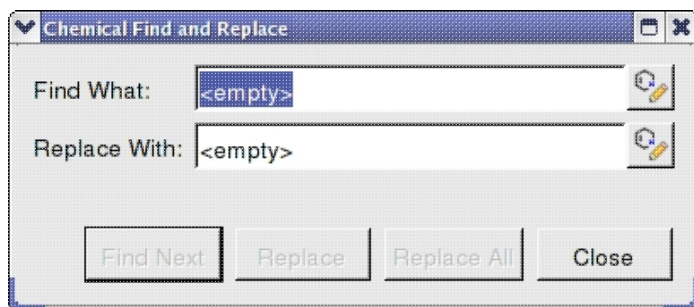
9.3.28 Chemical Find and Replace

Chemical Find&Replace tool allows you to find an arbitrary chemical fragment with one or more attachment point(s) and replace it with another fragment with the same number of attachment points.

To find a substructure and replace it with something else:



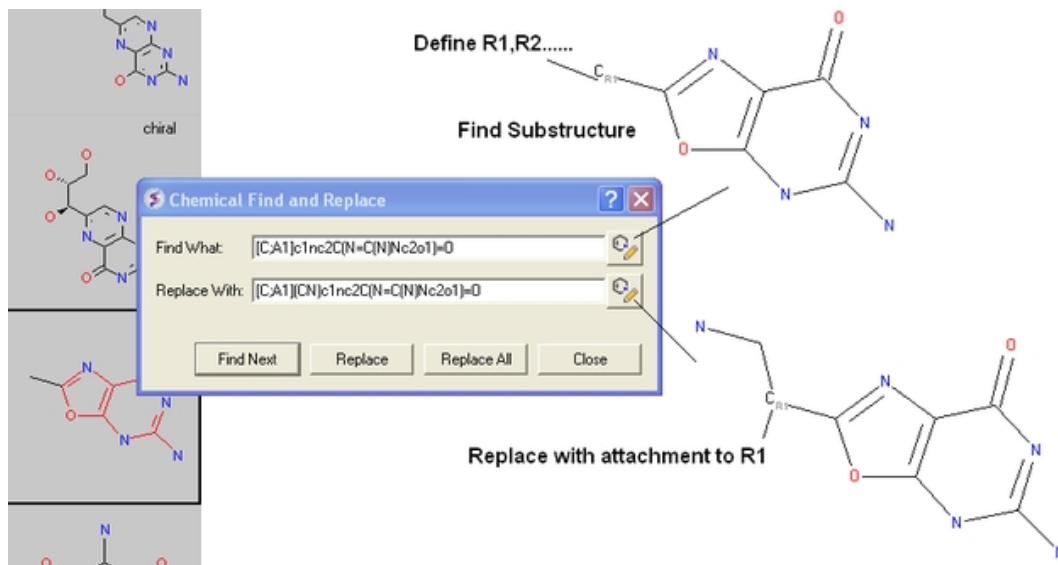
- Select the column in which the molecular structures are displayed. The column is usually called "mol".
- Right click on the "mol" column header and select Find and Replace. A data entry box as shown below will be displayed.



- Click on the Molecular Editor button at the end of the Find What: data entry box.
- The ICM Molecular Editor will be displayed. Draw the substructure you wish to search for and replaced.
- Draw the pattern and mark attachment points with R1,R2.... R-groups can be added by right clicking at the attachment point and selecting the R-group from the drop down options.
- Close the ICM Molecular Editor and the string will be displayed.
- Repeat with the "Replace With:" data entry box. Make sure the same number of R1,R2... labels are drawn.
- Click the Find Next button and then Replace or Replace All. When a substructure to replace is identified it will be colored red.

NOTE: There are a number of keyboard shortcuts which can be used to draw chemicals. Also please note that an aromatic bond in the source molecule will not match a double bond in the replacement pattern.

Here is an example:

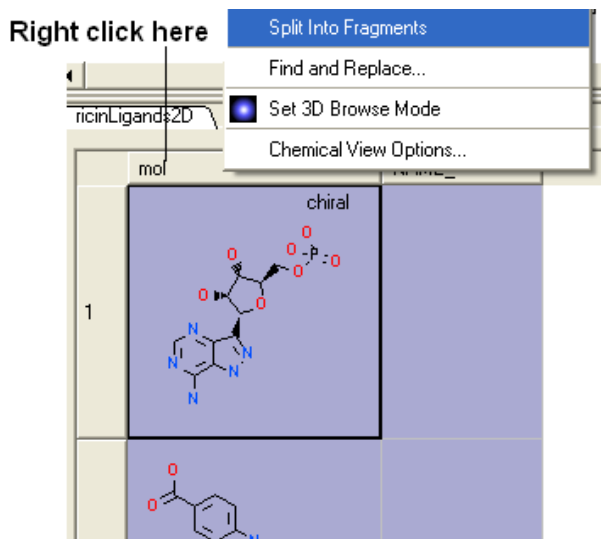


9.3.29 Split Chemical(s) into Fragments

Chemicals displayed in an ICM Molecular Table can be split into fragments. This is useful for generating a series of R-groups to be added to a scaffold (See section describing reactions).

To generate fragments:

- Select the column or row(s) you wish to generate the fragment from.
- Right click on the "mol" column header and select "Split Into Fragments".
- A new table of chemical fragments will be displayed. Each fragment is assigned an attachment point which is flagged with an asterisk (*).



New table is generated containing fragments

mol	cnt	n_apo
	1	1
	1	2

9.3.30 Rotate Chemical for Best-Fit

To improve the display of a chemical within an ICM `molecular table` you can choose an option called "Rotate for best fit".

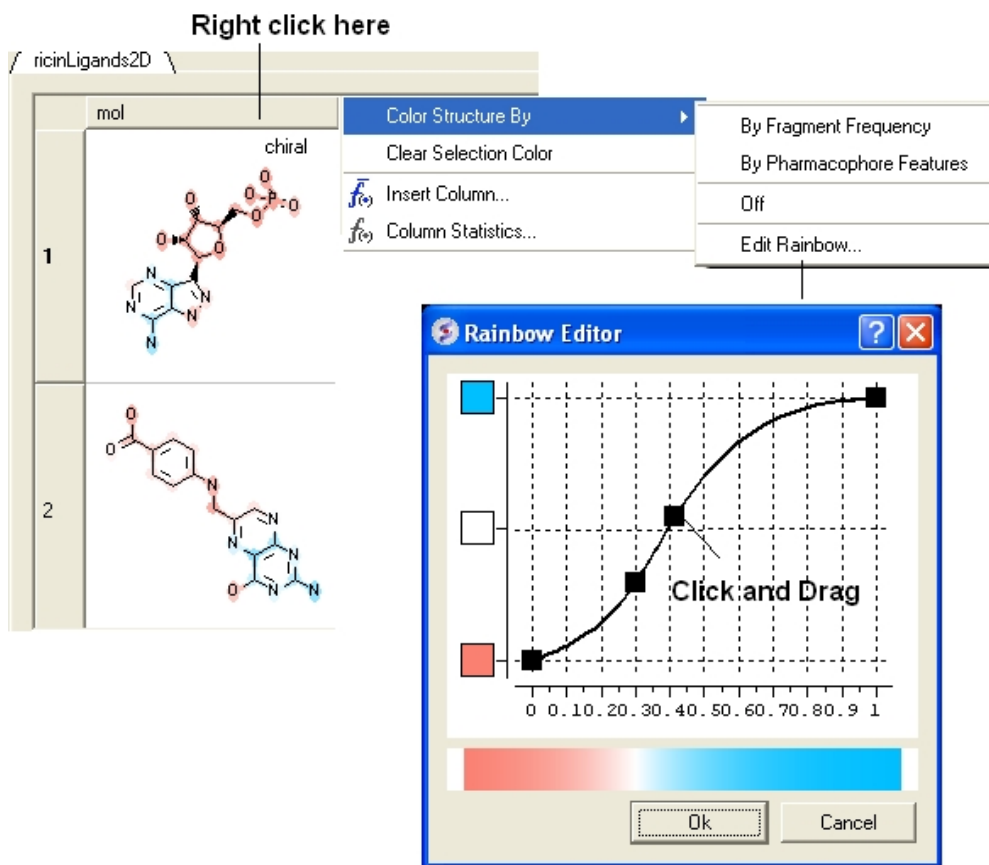
This option can be found by:

- Right click on the "mol" column header.
- Select "Rotate for Best Fit"

9.3.31 Color Chemical Structure

To color the structure of a chemical in an ICM `molecular table` by fragment of pharmacophore frequency:

- Right click on the "mol" column header.
- Select Color Structure By
- Select "By Fragment Frequency" or "By Pharmacophore Features"



NOTE: The coloring can be controlled using the "Edit Rainbow" option and the coloring can be removed using the "Off" option

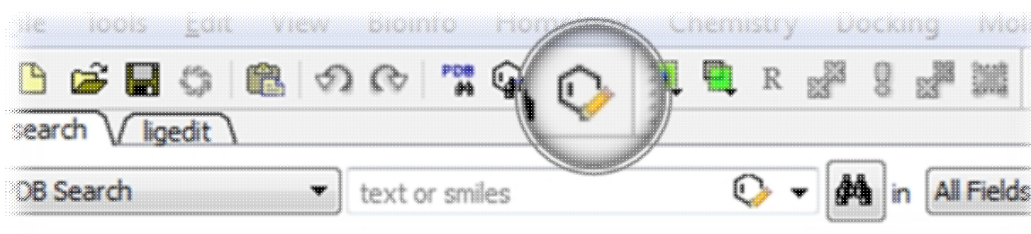
9.4 Molecular Editor

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS
| Video |

Draw new compounds by using the ICM Molecular Editor.

The molecular editor can be activated by:

- Clicking on the Open ICM Molecular Editor button shown below.



OR

- Select Chemistry/Molecular Editor and the editor as shown below will be displayed.

9.4.1 Drawing a New Chemical Structure

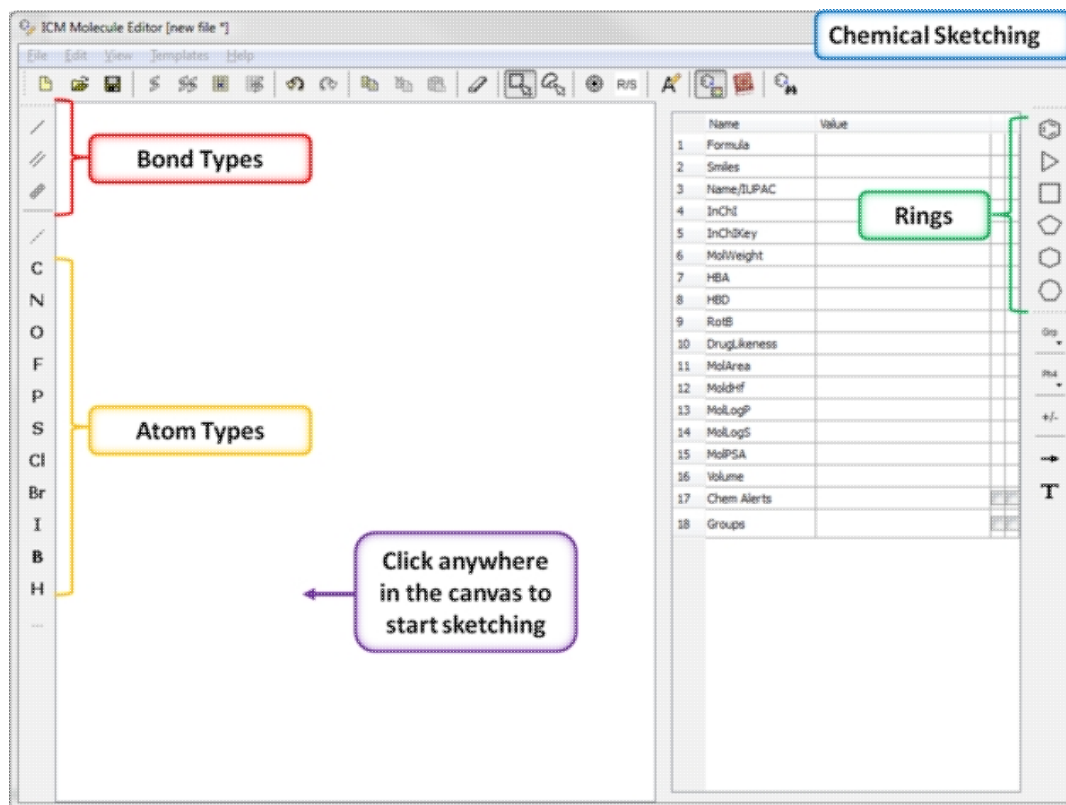
To draw a new chemical structure the ICM Molecular Editor should be loaded.

To do this:

- Select Chemistry/Molecular Editor

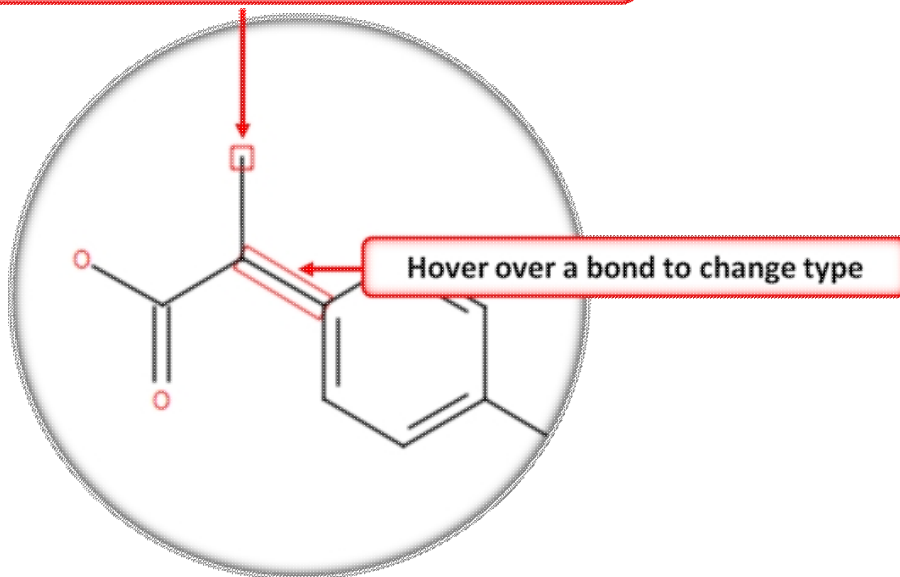
Now you can start drawing your structure.

- First select an atom, bond or ring and then click in the white Molecular Editor Workspace to start sketching.



- You can extend your structure by selecting another atom, bond or ring.
- Select where on the structure you would like to add the new group by hovering the mouse over the desired position. The position you will add to will be highlighted in a red box.

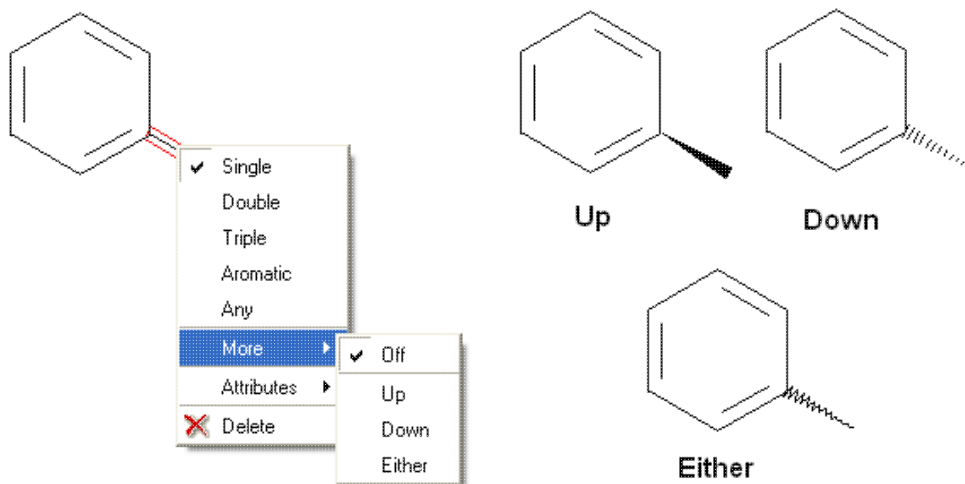
Hover over an atom and new bonds or atoms will be added from the position highlighted by a red box



9.4.2 Stereo Bond

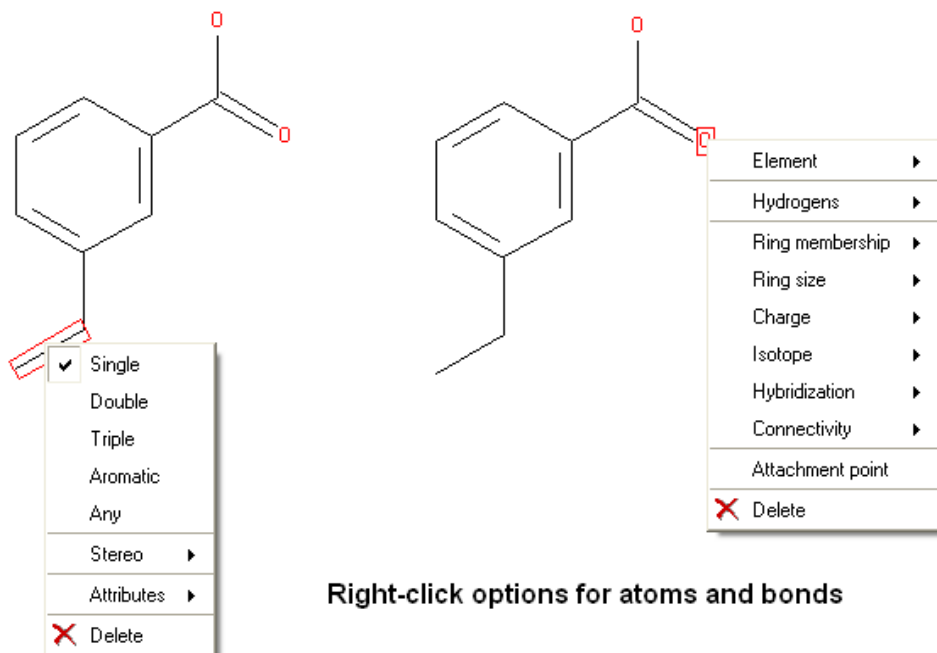
To set the stereo bond type:

- Right click on the bond and a menu as shown below will be displayed.
- Select which bond direction you desire from up, down or either.



9.4.3 Right Click Options

There are a number of different options available when you **right click** on a **bond** or **atom**. These can be used for editing bond and atom types and there are additional options for Chemical Searching.



Right-click options for atoms and bonds

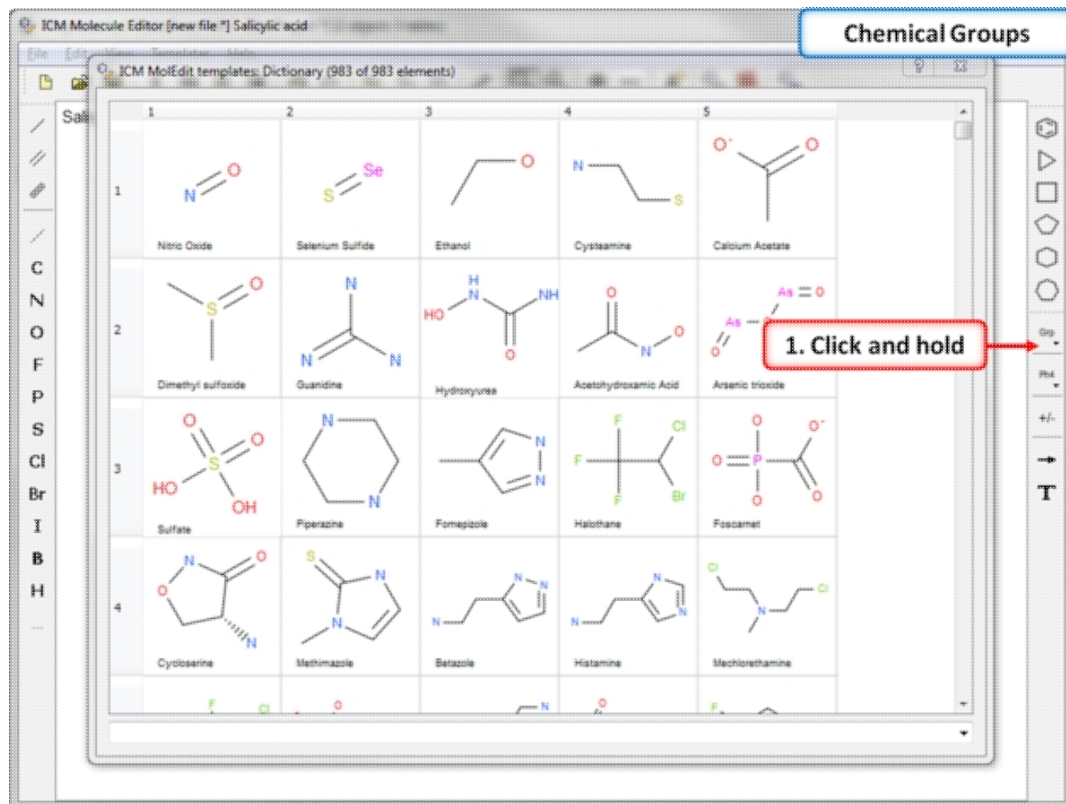
9.4.4 Chemical Dictionary

There is an inbuilt chemical dictionary which enables you to import a 2D sketch quickly. To use it:

- Click on the chemical dictionary button.
- Start typing the name of the chemical.
- Select the chemical from the drop down list.

9.4.5 Dictionary of Chemical Groups

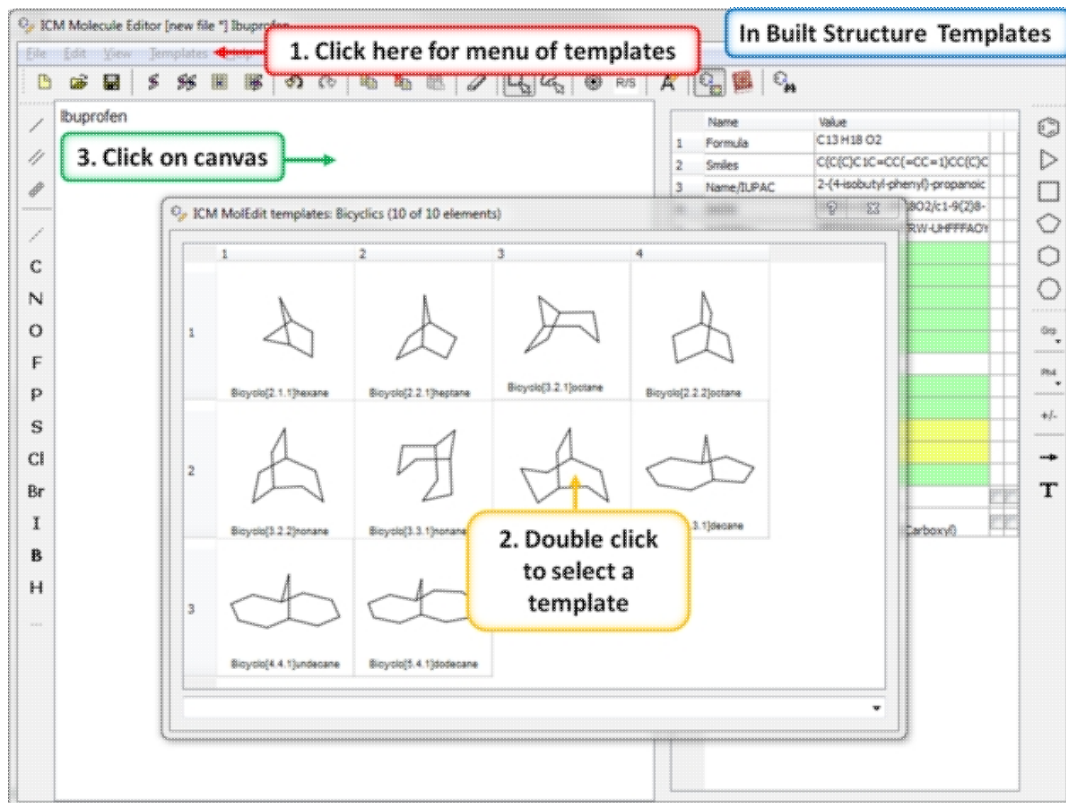
The ICM Chemical Editor has a dictionary of chemical groups. These groups are stored in a table stored in `CHEM_GROUPS.csv` file. This file can be manually edited to remove or add new definitions. The group can be selected using the `Grp` button of the Editor.



9.4.6 Template Menu

A set of chemical templates are available for:

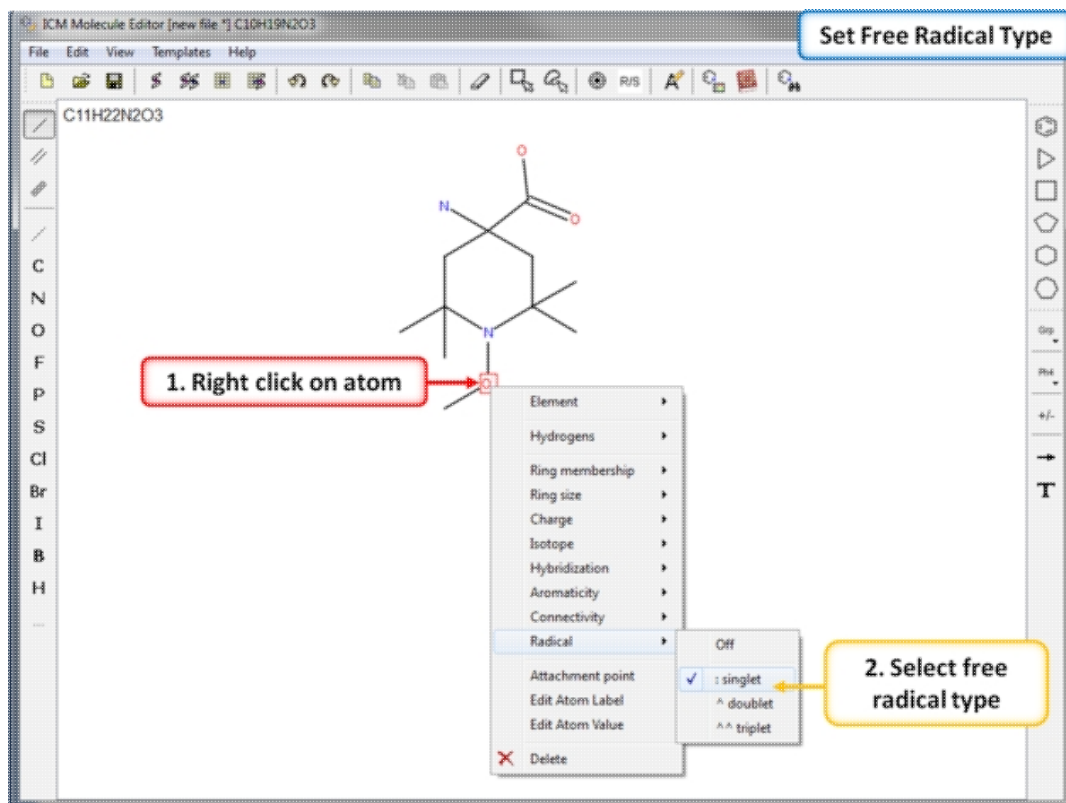
- amino acids
- bases
- bicyclics
- rings



9.4.7 Set Free Radical Type

To set a free radical type (singlet, doublet or triplet) to an atom:

- Sketch the ligand in the Molecular Editor.
- Right click on the atom and select the free radical type.
- Convert to 3D or add to a chemical table if needed for docking.



9.4.8 Adding and rotating a fragment in molecular editor by clicking-holding-and-dragging.

The Chemical Editor allows one to add bonds (click the **bond** button), or chemical groups (the **Grp** button). Often one needs to re-orient the added bond or fragment. In this case press on the atom to which you are going to append the bond or the fragment and **hold-and-drag** the mouse it until you see added fragment rotating around the attachment point. While you keep the mouse button pressed you can **rotate** by dragging in the preferred direction.

If you add templates (the **Templates** menu, or **paste** fragments, the mechanism is somewhat different because the group is added in its default orientation. In this case you can press **Ctrl** to rotate the fragment to be able to attach the fragment in the desired orientation.

9.4.9 Property Monitor

When drawing a compound in ICM you can monitor important ADME-Tox and drug-likeness properties.

To view the monitor

- ICM Molecular Editor/View/Chemical Monitor or click on the Chemical Monitor button and a window as shown below will be displayed.

1. Click here to open chemical monitor panel

3. The properties are calculated on the fly as you sketch

2. Chemical monitor panel will be displayed here

Name	Value
1 Formula	C ₁₃ H ₁₈ O ₂
2 Smiles	C(C(C)C)C(=O)C=C(C)C
3 Name/IUPAC	2-(4-isobutyl-phenyl)-propanoic
4 InChI	InChI=1S/C ₁₃ H ₁₈ O ₂ /c ₁₋₉ (2)8-
5 InChIKey	HEPHEWSOOOATRW-UHFFFAOY
6 MolWeight	206.1307
7 HBA	3
8 HBD	1
9 RotB	4
10 DrugLikeness	0.958908
11 MolArea	266.78
12 MolHf	-92.8134
13 MolLogP	3.33837
14 MolLogS	-2.96525
15 MolPSA	28.6465
16 Volume	214.268
17 Chem Alerts	
18 Groups	Phenyl Carboxylic acid (Carboxy)

The following properties are monitored.

- Chemical Formula
- SMILES string
- Name/IUPAC Name
- InChI
- InChI Key
- Molecular Weight (MolWeight)
- Number of Hydrogen Bond Acceptors (HBA)
- Number of Hydrogen Bond Donors (HBD)
- Number of Rotatable Bonds (RotB)
- DrugLikeness value Druglikeness is not based on other properties, this is separate model built using Bayesian Classifier (See: <http://www.molsoft.com/mpropdesc.html>). Normally Bayesian classifier returns either 1 or -1 (positive or negative) therefore positive values means drug-like, negative - not.
- Molecular Area (MolArea)
- Prediction model build for 'delta Hf in gas' property using public NIST database (description can be found here: <http://webbook.nist.gov>). A low dhf value means that the compound is more 'stable'.
- LogP (MolLogP)
- LogS (MolLogS)
- Polar Surface Area (PSA)
- Volume
- Chem Alert will flag any reactive group, PAINS or Bad ADME-TOX groups.
- Groups option highlights chemical groups in the sketch.

The rows in the monitor window are colored from green (good) to red (poor). You can right click on the table to copy a value or add your own prediction model.

9.4.10 Editing structure using keyboard

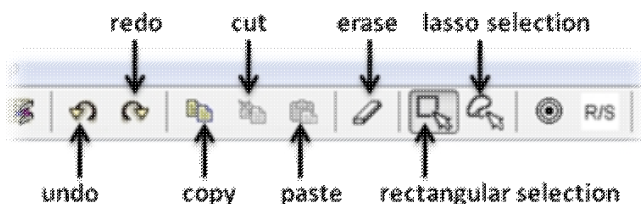
You can select a fragment or simply position your mouse cursor over a bond or atom and use the following keystrokes for editing:

- **Changing atom properties**

- d set/unset heavy *atom connectivity* as drawn to avoid additional branches in chemical searches. Toggle.
- r set/unset the "**in-ring**" property of *atoms*. Toggle.
- y set/unset *atom hybridization* type (*sp1,sp2,sp3*). Toggle.
- A set *atomic* property to be **aliphatic**
- a set *atomic* property to be **aromatic**
- * (asterisk) set *atomic* property to be **any atom**
- C Carbon
- H Hydrogen
- N Nitrogen
- O Oxygen
- F Fluorine
- P Phosphorus
- S Sulfur
- I Iodine
- **Changing chemical bonds**
- 1,- (dash) **single bond**.
- 2,= (equal) **double bond**.
- 3,# (hash) **triple bond**.
- 4,: (column) **aromatic bond**.
- 0,~ (tilde) **any bond**
- **Changing R-Groups**
- Press 1 to mark an atom as R1
- Press Ctrl-1 to preserve its type and create [C;R1]
- Press Ctrl-0 to remove the mark.

9.4.11 Chemical Editing Options

Tools are provided to help editing your chemical sketch. These include buttons for selection, erasing, copy and paste.



9.4.12 Paste SMILES or InChi String

To paste a SMILES string into the Molecular Editor:

- Open the Molecular Editor.
- In the Molecular Editor window select Edit/Add Smiles/InChi.
- Click on the canvas.

9.4.13 Save Chemical in Mol/RXN Format

To save the structure as a MOL or SDF file on your machine or server:

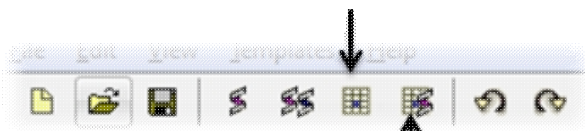
- Select File/Save As... in the Molecular Editor window
- Enter a filename and select where you wish to save the file.
- Save as type mol or rxn

9.4.14 Save to Chemical Spreadsheet

To save the chemical to a chemical spreadsheet:

- Click on one of the buttons shown below. You can append to a new table or an existing one.

append molecule to a chemical spreadsheet



append molecule to a chemical spreadsheet
and convert to 3D

9.4.15 Save as SMILES or InChi

To save the chemical in SMILES format

- Edit/Copy as Smiles

To save the chemical in InChi format

- Open the Chemical Monitor.
- Right click on the InChi row and choose copy.

9.4.16 Convert Chemical to 3D or Generate Conformers

To convert the chemical to 3D or to generate conformers:

- Click on the buttons shown below to convert to 3D or generate conformers in the Merck Molecular Force Field (MMFF)
- The 3D chemical will be listed in the ICM workspace. If conformers are generated a table showing the conformers will be created.

Convert to 3D



Generate Conformers

9.4.17 Isis Draw Copy and Paste

Molecules drawn in ICM can be cut and pasted into ISIS-Draw and vice-versa molecules from ISIS-Draw can be cut and pasted into ICM.

To perform these functions the correct settings need to be turned on in ISIS-Draw so that the compound drawing is saved in the clipboard.

IN ISIS-DRAW - Go to Options/Settings/General/Copy Mol Rxn file to the clipboard.

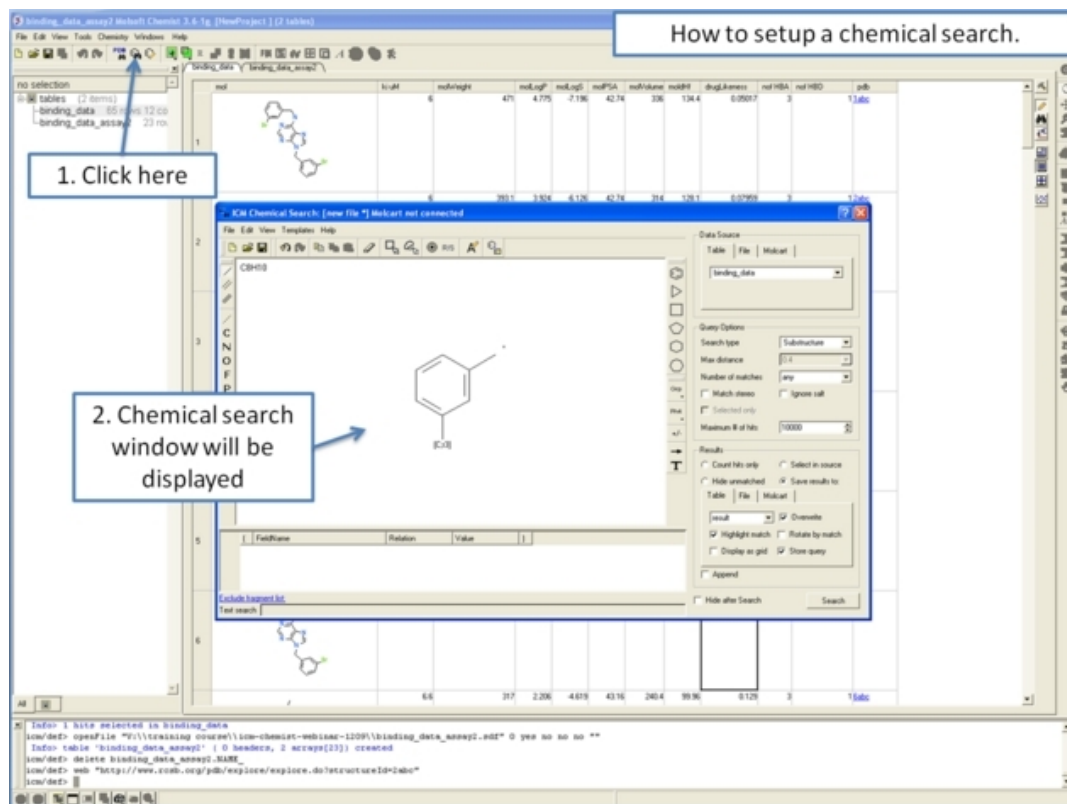
Compounds drawn in ISIS Draw can be cut and pasted into the ICM Molecular Editor and into ICM tables. Compounds can be copied in ICM by:

- Right click on the compound in the chemical table and select Copy Molecule.

9.5 Chemical Search

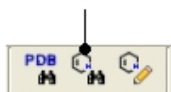
Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

Chemical similarity searching can be used to screen a database of compounds for structural similarity to a query chemical structure. The chemical similarity search window is shown below.



To access this window use the Tools/Chemical Search menu or select the button shown below.

Chemical search



9.5.1 Query Setup

To set up a query first you must have either drawn or loaded a chemical structure into ICM. Instructions for this are described in the Reading Chemical Structures and Molecular Editor sections of this manual. If for example the query molecule you want to search is in a chemical spreadsheet you can right click on chemical in the spreadsheet and select **Query Molecule**.

At this point your query can be modified as described in the Molecular Editor sections of this manual. **However**, there are a number of ways to specifically modify your query and filter your search. The way to accomplish this is to right click on a bond or atom and a menu as displayed below will be displayed. The menus differ depending on whether you right click on a bond or atom.

If you wish to specify a filter at an atom.

- Right click on the atom and the menu shown below will be displayed.

The screenshot shows the ICM Chemical Search interface. The main window displays a chemical structure of 1-phenylpropene (C₁₁H₁₄). A context menu is open over a carbon atom, with the text "Right click on atom" above it. The menu options are:

- Element
- Hydrogens
- Ring membership
 - Off
 - As Drawn R
 - R0
 - R1
 - R2
 - R3
- Ring size
- Charge
- Isotope
- Hybridization
- Aromaticity
- Connectivity
- Attachment point
- Delete

The right-hand panel contains the following sections:

- Data Source:** Table | File | Molcart
- Query Options:**
 - Search type: Substructure
 - Max distance: 0.4
 - Number of matches: any
 - Match stereo
 - Ignore salt
 - Selected only
 - Maximum # of hits: 10000
- Results:**
 - Count hits only
 - Select in source
 - Hide unmatched
 - Save results to:
 - Table
 - File
 - Molcart
 - result
 - Overwrite
 - Highlight match
 - Rotate by match
 - Display as grid
 - Store query
 - Append
 - Hide after Search
 - Search

At the bottom, there is a table with columns: FieldName, Relation, Value. Below the table is a "Text search" field and a link for "Exclude fragment list".

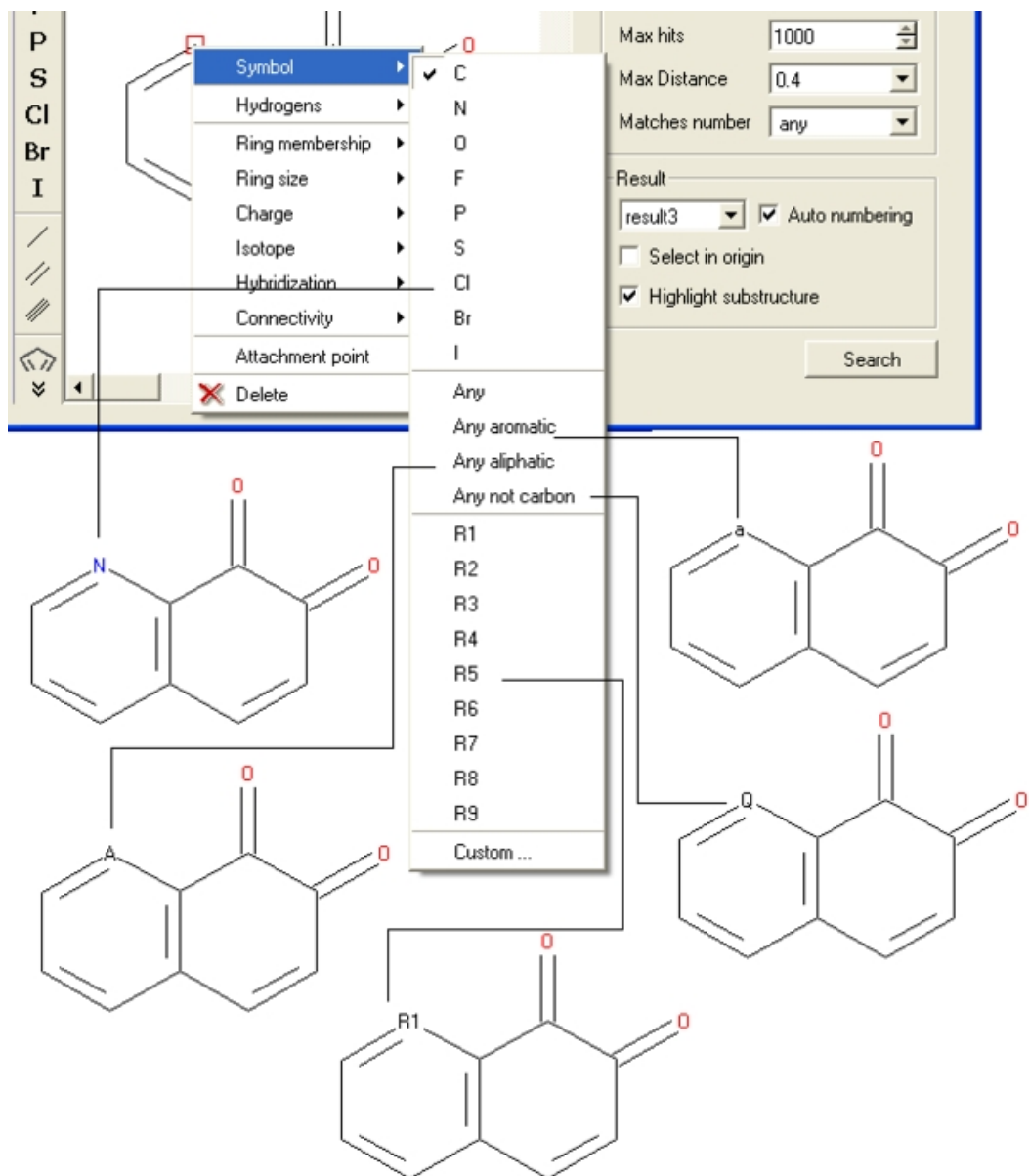
If you wish to specify a filter at a bond.

- Right click on the bond and the menu shown below will be displayed.

The screenshot shows the ICM Chemical Search interface. The main window displays a benzene ring with a bond highlighted in red. A context menu is open over this bond, listing options: Single, Double (checked), Triple, Aromatic, Any, Stereo, Attributes, and Delete. The text "Right click on bond" is written next to the menu. The software interface includes a menu bar (File, Edit, View, Templates, Help), a toolbar, a vertical element list (C, N, O, F, P, S, Cl, Br, I, B, H), a search results table with columns for FieldName, Relation, and Value, and a right-hand panel with "Data Source", "Query Options", and "Results" sections.

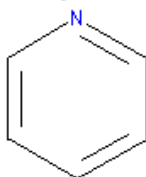
To specify a particular atom type, aromatic, aliphatic or R-group at a particular atom site.

- Right click on the atom and select the "symbol" option as shown below.
- Select the desired atom type, aromatic, aliphatic or R-group and a symbol will be displayed as shown below.



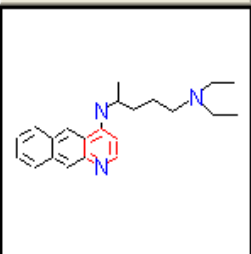
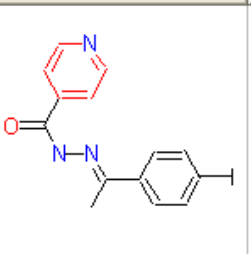
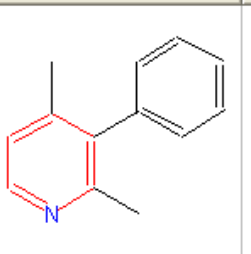
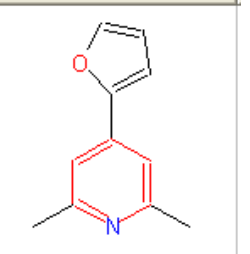
For example:

Query:



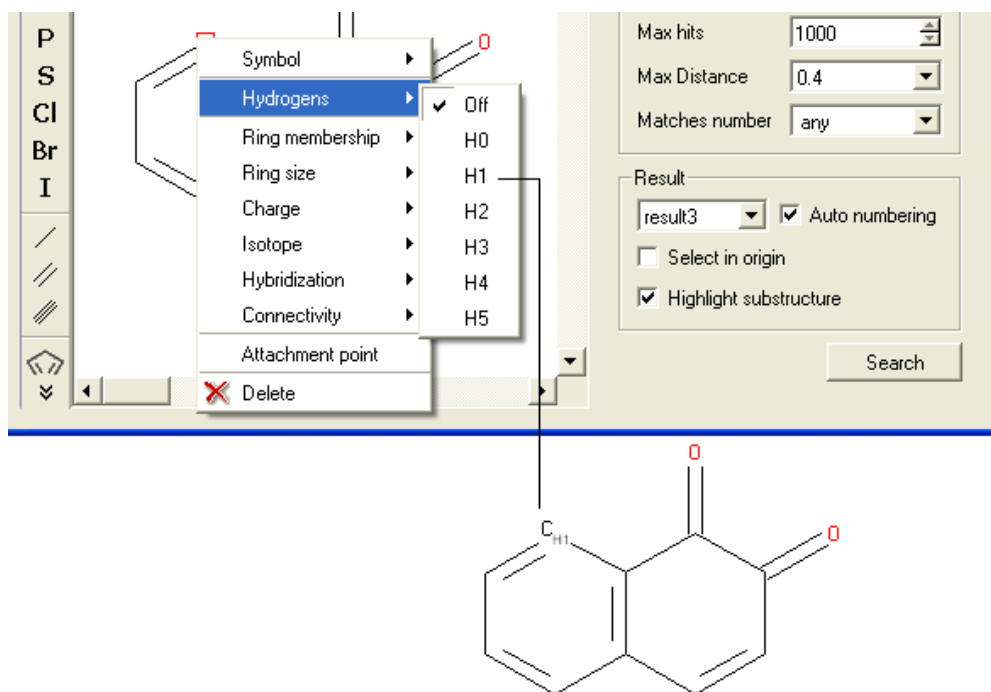
Right click atom - Symbol - N

Selection of chemical substructure search results

	1	2	3	4
0				

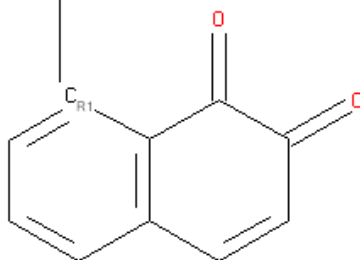
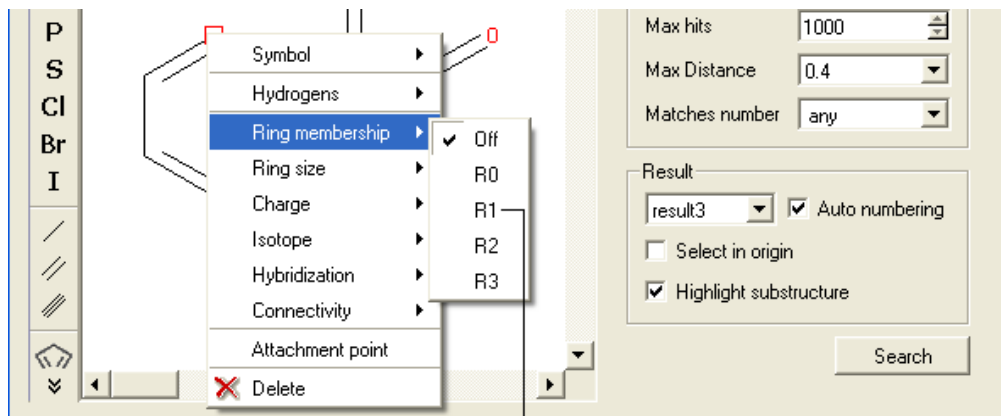
To specify a particular number of hydrogen atoms at a particular site:

- Right click on the atom and select the "Hydrogens" option as shown below.
- Select how many hydrogens you wish to specify and a symbol will be displayed as shown below.



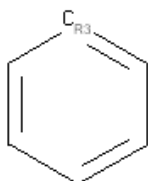
To specify the number of rings a particular atom will be a member of:

- Right click on the atom and select the "Ring membership" option as shown below.
- Select whether the atom should be part of 1, 2 or 3 rings.



For example:

Query:



Right click atom - Ring Membership - R3

Selection of chemical substructure search results

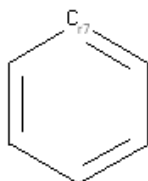
	1	2	3	4
24				

To specify the ring size connected to an atom:

- Right click on the atom and select the "Ring size" option as shown below.
- Select the size of the ring the atom should be connected to and a symbol will be displayed as shown below.

For example:

Query:



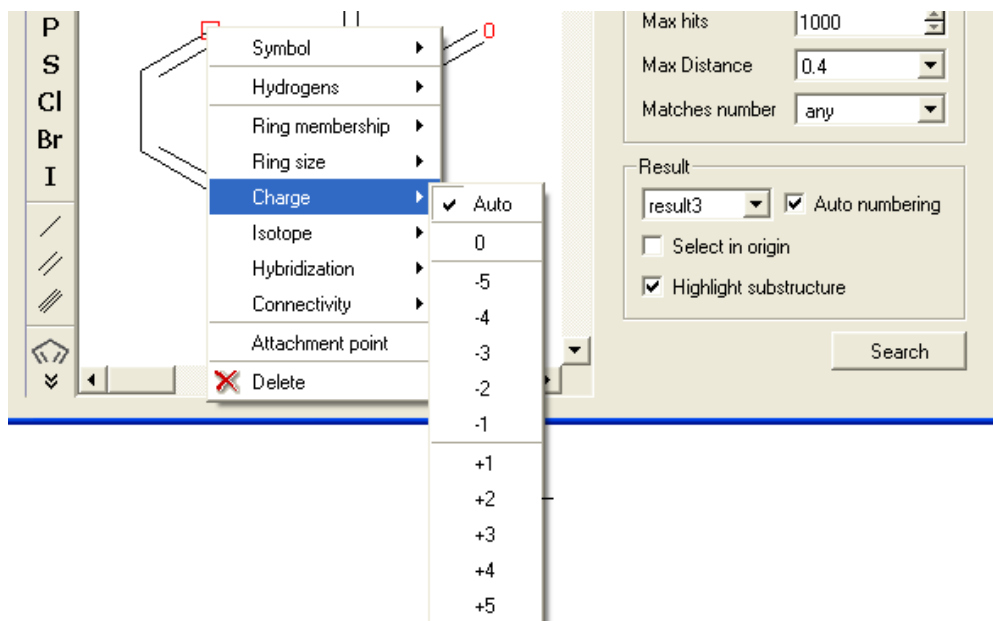
Right click atom - Ring size - R7

Selection of chemical substructure search results

	1	2	3	4
0				

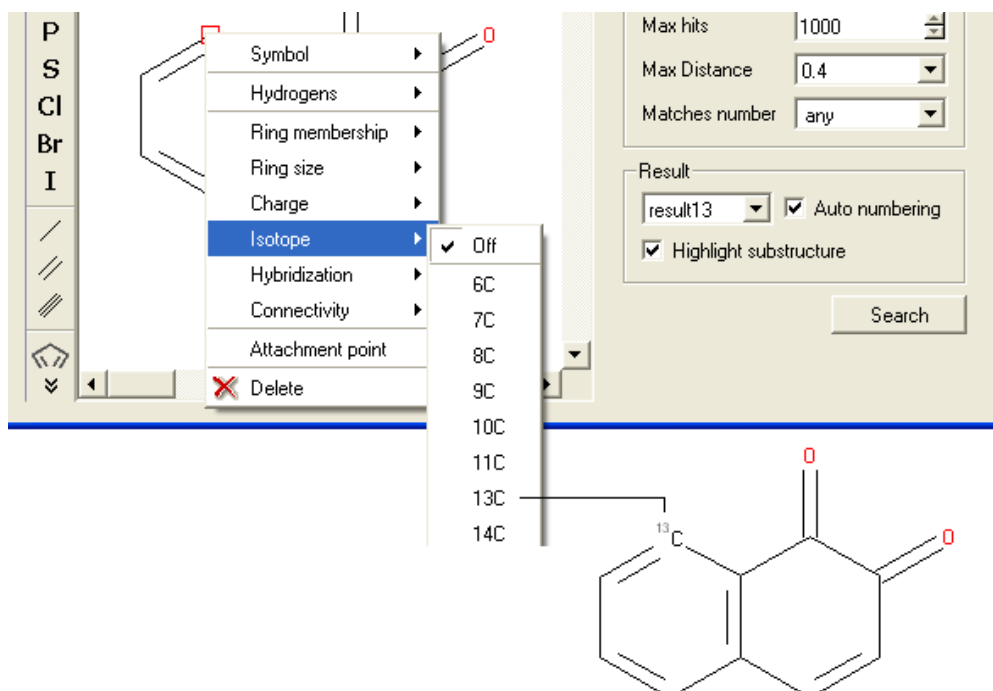
To specify the charge at a particular point:

- Right click on the atom and select the "Charge" option as shown below.
- Select the desired charge and a symbol will be displayed as shown below.



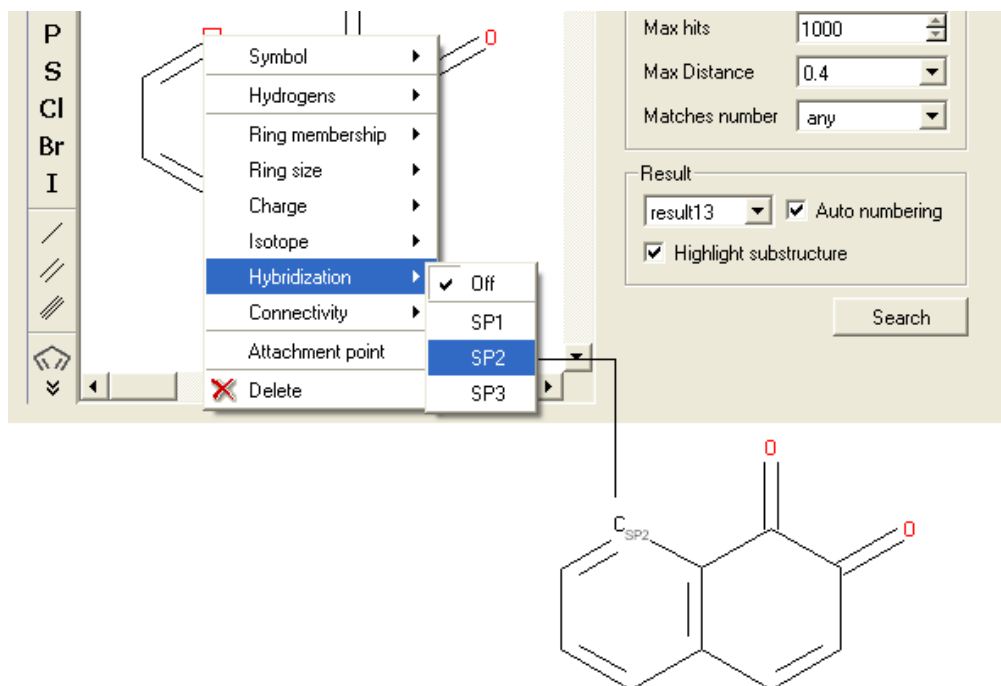
To specify an isotope at a particular atom

- Right click on the atom and select the "Isotope" option as shown below.
- Select the desired isotope from the list and a symbol as shown below will be displayed.



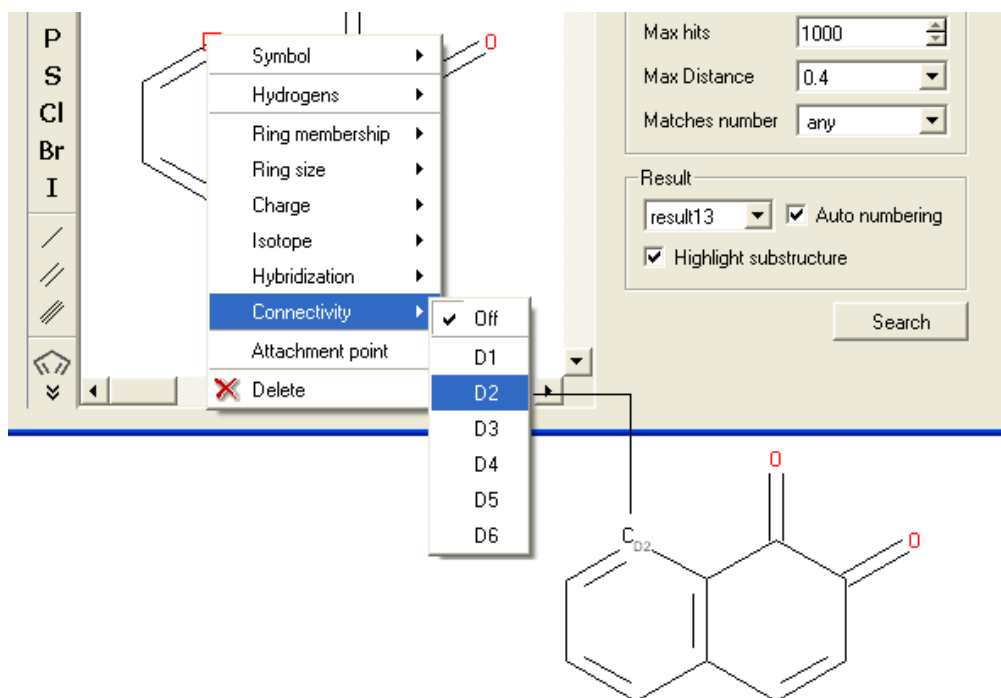
To specify the hybridization state of the atom:

- Right click on the atom and select the "Hybridization" option as shown below.
- Select the desired hybridization state and a symbol will be displayed as shown below.



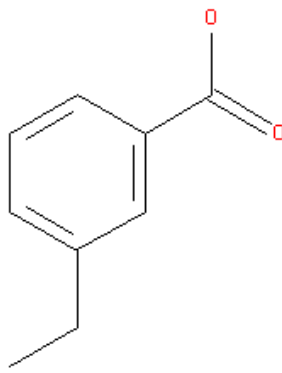
To specify the number of atoms you wish an atom to be connected to:

- Right click on the atom and select the "Connectivity" option as shown below.
- Select the number of atoms you wish the atom to be connected to and a symbol will be displayed as shown below.



To specify an attachment point - ie the position at which substituents will be added

- Right click on the atom and select the "Attachment point" option as shown below.
- An asterisk representing the attachment point will be displayed next to the atom.



Attachment point ● *

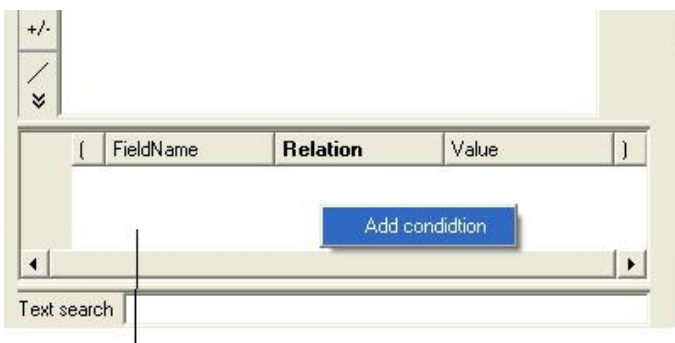
An attachment point means that the atom can be attached to zero or more bonds to heavy atoms.

9.5.2 Filter Search

How to filter your query

To filter your query:

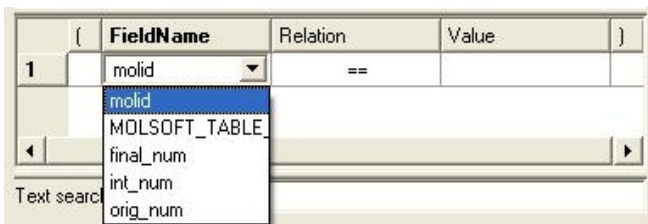
- Right click in the box shown below and select 'Add Condition'



Right-click here and select Add Condition

To add conditions to your filter:

- Double click in the fields labeled "Name" and "Relation" and select the options from the drop down arrow or type in values.

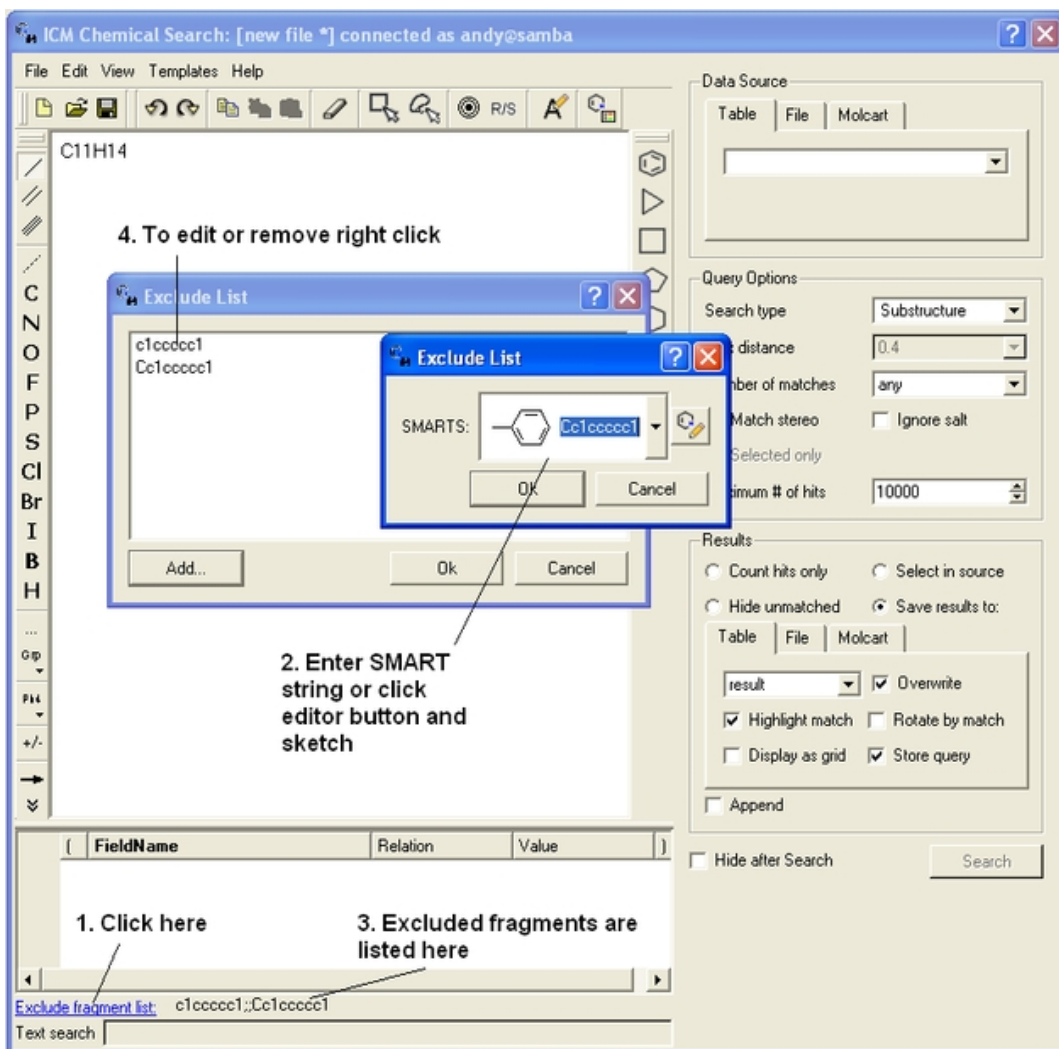


To remove a filter, right click on the filter and select 'Remove Filter'.

To exclude a fragment from your search

Click on the option **Exclude fragment list**:

- Enter the SMART string of the fragment or use the Molecular Editor option to sketch the fragment.
- Click OK and the fragment will be listed at the bottom of the chemical search window as shown below.



9.5.3 Query Processing

To begin processing your query first you need to decide which database to search. The options are listed in the **Data Source** section of the ICM chemical search window.

Data Source

You can either search a **Table** - Chemical Spreadsheet a **File** - Local Database or **MolCart**.

If you select **MolCart** you first need to setup the link to the correct database.

- Enter the Server Name in which the database is stored.
- Enter the database name.
- Enter your username and password for the server.
- You can save these details so you don't have to re-enter this information each time you use the chemical similarity search.

If you are searching a **Table**, click on the **Table** tab and then select the drop down button where the names of your currently loaded tables are stored.

If you are searching a **File** click on the **File** tab and then locate your local database file .molt or if it is already loaded into ICM you can locate it with the drop down button.

Query Options

Now select a search type:

- Click on the drop down arrow next to the "Search Type" option in the Query Options panel.
- Select the search type you want to use.

A **substructure** search is a search whereby only the defined molecule in the query will be searched against the database. Whereas, a **FP similarity** search which stands for fingerprint search enables any fingerprint within a structure to be searched for in the database.

The **Max distance** option is available for use with the FP search and the **Matches number** option is for use with the substructure search. The option you do not require based on your search method will be blanked out. A "Max distance" value of 0 means that the search will only identify matches exactly the same as the fingerprint - the default is 0.4. The "Matches number" option allows you to stipulate how many times within a structure in the database your query can be found.

You can match stereo by selecting the **Match stereo** option and **ignore salts**. If you make a selection of your query ICM can use that selection to search. How to make selections in the Molecular Editor are described here. Enter the Maximum number(#) of hits you would like returned.

Results

Before processing the query determine how you would like your results displayed in the **Results** section of the **Chemical Search** window.

Count hits only - this option will count the number of hits and display this number in a window once the searching has been completed.

Select in source - If you are searching a table you can select and highlight the query in the source table that you are searching.

Hide unmatched - Hide unmatched will hide the compounds that were not matched from view.

Save results to: - this option gives you the option to save the output results to a Table- Chemical Spreadsheet a File -Local Database or MolCart.

Append - this option will allow you to append to current table, file or Molcart database.

Search

- Click on the **Search** button to execute the search. You can choose to hide the window after the search.

9.5.4 Search a Database by Text

To search a database by text enter the text you wish to search in the **Text Search** data entry box at the bottom of the ICM Chemical Search window.

ICM Chemical Search: [new file *] Molcart not connected

File Edit View Templates Help

C11H14

[C,R1]

Single
 Double
 Triple
 Aromatic
 Any
 Stereo ▶
 Attributes ▶
 Delete

Data Source
 Table | File | Molcart

Query Options
 Search type: Substructure
 Max distance: 0.4
 Number of matches: any
 Match stereo Ignore salt
 Selected only
 Maximum # of hits: 10000

Results
 Count hits only Select in source
 Hide unmatched Save results to:
 Table | File | Molcart
 result Overwrite
 Highlight match Rotate by match
 Display as grid Store query
 Append

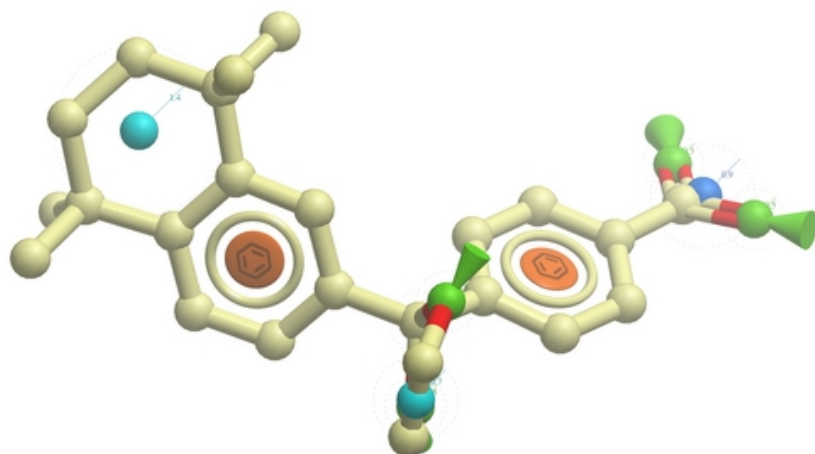
Hide after Search Search

Exclude fragment list
 Text search

Enter text here

9.5.5 Pharmacophore Drawing and Searching

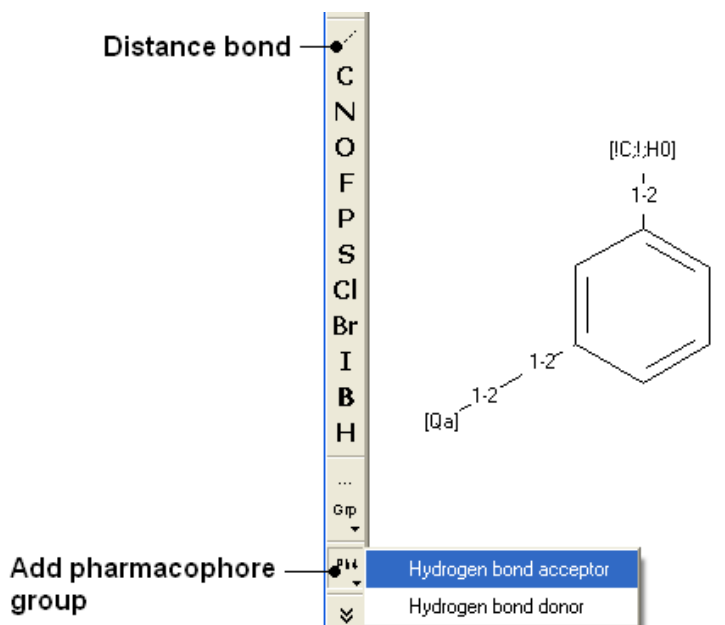
Pharmacophores can be drawn in 2D in the ICM Molecular Editor or in 3D in the Graphical Display. 2D pharmacophore sketches can be used to search chemical tables (spreadsheets) containing 2D or 3D coordinates. A 3D pharmacophore can be used to search chemical tables containing 3D coordinates only.



9.5.5.1 Pharmacophore Draw 2D

A 2D pharmacophore can be drawn using the ICM Molecular Editor or if you are going to use the drawing to search it is more efficient if you draw it in the Chemical Search window.

Use the distance bond button and the add pharmacophore group button to sketch the pharmacophore. The distance bond button represents the number of bonds between each pharmacophore point. You can edit the distance by right clicking on the bond and selecting edit. Other groups such as aromatic can be added using the standard molecular editor buttons.



NOTE: Do not mix the 2D and 3D pharmacophore environment. For example do not edit a 3D pharmacophore in the 2D editor.

9.5.5.2 Pharmacophore Draw 3D

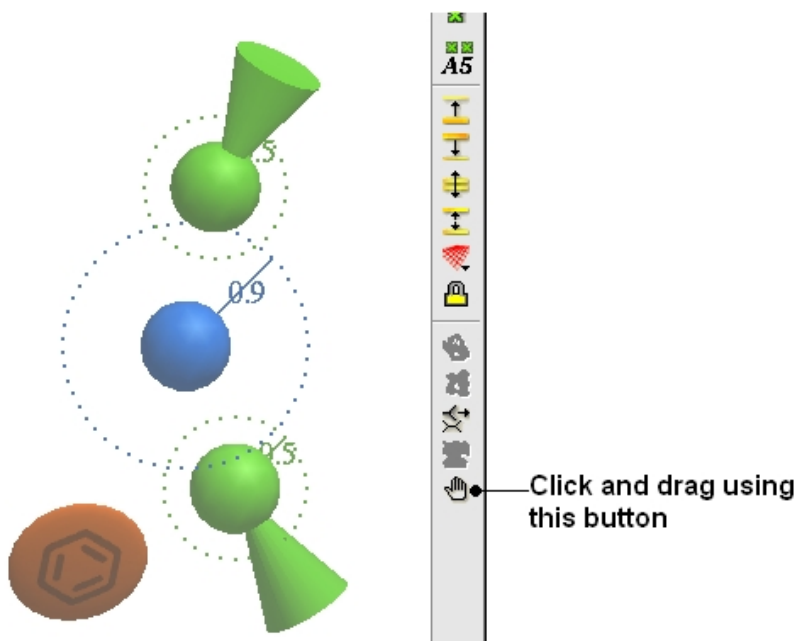
The easiest way to begin drawing a 3D pharmacophore is to draw a chemical in the ICM Molecular Editor which contains the key pharmacophore groups you want and then convert to 3D and extract the pharmacophore groups.

To draw a 3D pharmacophore this:

- Once the ligand is converted to 3D, right click on the ligand in the ICM Workspace.
- Select the Tools/Copy as Pharmacophore and choose the **pharm centers** option.
- The pharmacophore groups or centers will then be displayed in the graphical display and can be displayed and undisplayed in the ICM workspace.

To move a pharmacophore group:

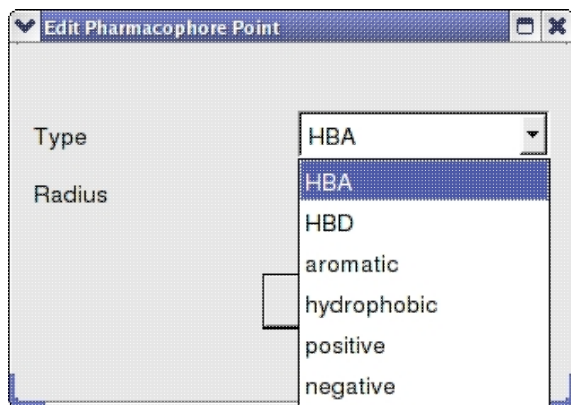
- Use the **drag atom** button (picture of a hand. See below)
- Click on the pharmacophore group and then drag.



NOTE: Distances between groups can be monitored using the atom distance measurement tool. See Calculating the distance between two atoms.

To change a pharmacophore group:

- Right click on the pharmacophore group in the ICM Workspace or in the Graphical Display.
- Select **Pharmacophore/Edit Point**
- Choose the group from the drop-down list shown below.
- Enter the desired radius.



To make a new pharmacophore group:

The easiest way to make a new pharmacophore group is to clone a pre-existing one. To do this:

- Right click on the pharmacophore group in the ICM Workspace or in the Graphical Display.
- Select **Pharmacophore/Clone Point**
- You can then move the new group as described above.

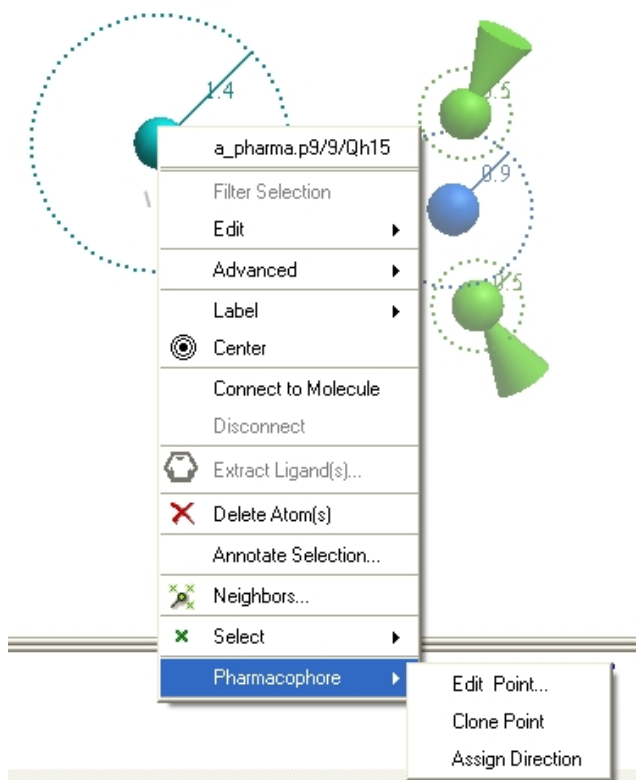
To change the direction of a pharmacophore group:

- Right click on the pharmacophore group in the ICM Workspace or in the Graphical Display.
- Select **Pharmacophore/Assign Direction**
- You can then move the new group using the drag atom button described above.

To remove the direction of a pharmacophore group:

- Right click on the pharmacophore group in the ICM Workspace or in the Graphical Display.
- Select **Pharmacophore/Remove Direction**

Pharmacophore right click options

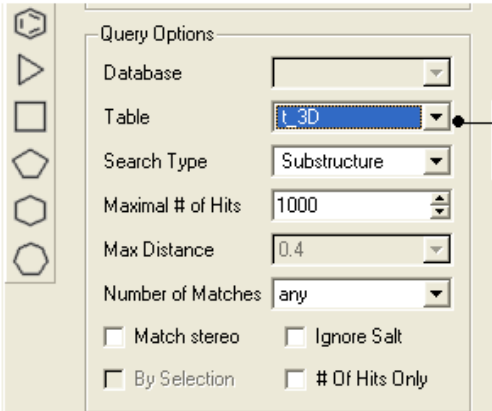
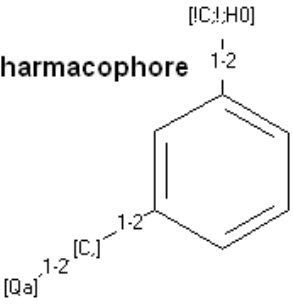


9.5.5.3 Pharmacophore Search

To perform a pharmacophore search using a 2D pharmacophore:

- Draw the 2D pharmacophore as described earlier in the Chemical Search window.
- Read in a molecular table to search or search a table in MOLCART.
- Select the chemical search options as shown below.
- Once the search has completed a new table with the results will be displayed.

Draw pharmacophore



Select the table you wish to search here

To perform a pharmacophore search using a 3D pharmacophore:

- Right click on the pharmacophore in the ICM Workspace
- Select **Search** Pharmacophore
- Use the drop-down button to select the table you wish to search. The table must contain 3D coordinates.

Right click here



Enter name of table containing 3D coordinates

How to send a 3D pharmacophore search query.

1. Right click on the pharmacophore points in the ICM Workspace and select "Search Pharmacophore"

2. Search a table with chemicals that have 3D coordinates

mol	name	molwt	count	L	D	T
1			430	0	0	
2			440	0	0	

9.5.5.4 How to extract a 3D pharmacophore from a ligand.

How to extract a 3D pharmacophore from a 3D ligand.

1. File/Open 3D ligand or read in PDB

2. Right click on 3D ligand and select Tools/Copy as pharmacophore

3. Display and undisplay pharmacophore types

```

>copy pharmacore display with a_ligand=H011
>copy pharmacore undisplay error xactick a_ligand=H011
>copy pharmacore display xactick a_ligand=H011
>copy pharmacore mouse background rgb=(255,255,255)
>copy pharmacore
  
```

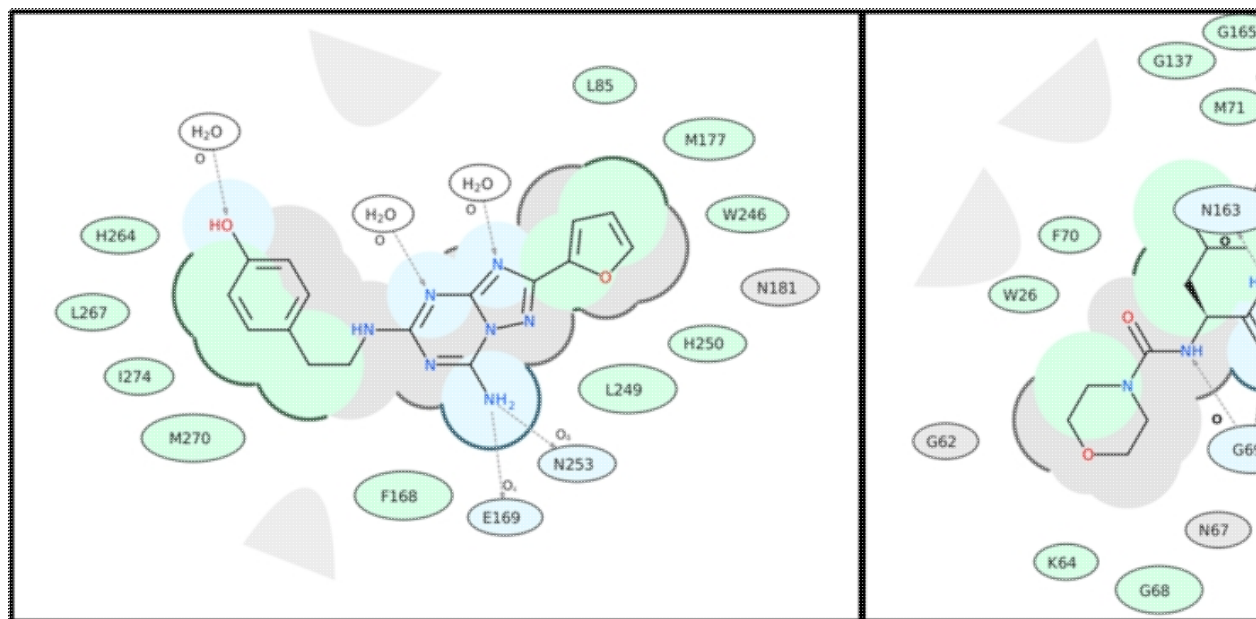
9.5.5.5 How to color a 2D chemical sketch by pharmacophore feature.

How color a 2D chemical sketch by pharmacophore feature.

1. Right click on column header

2. Select Color Structure By/ By Pharmacophore Features

9.6 2D Ligand Interaction Diagram



To create a 2D interaction diagram between a ligand and receptor:

- Read into ICM a PDB file containing the ligand and receptor. For correct hydrogen bond assignment you will need to convert to an ICM object.
- Right click on the ligand and choose Tools/Ligand 2D Diagram.
- A dialog box will be displayed. Select the receptor object from the drop down button and allow ICM to Auto Detect the ligand if needed.

- The **Hydrophobic Distance Cutoff** is \leq to the minimum distance between ligand and the closest atom in the residue. The **Hydrogen Bond Strength** is based on the GRAPHICS.hbondMinStrength parameter so lowering the value will identify weaker hydrogen bonds.
- An option is available to match the 2D representation to a pre-calculated one saved in a chemical table loaded into ICM.

A guide to the coloring and representation of the 2D display:

- Green shading represents hydrophobic region.
- Blue shading represents hydrogen bond acceptor.
- White dashed arrows represents hydrogen bonds.
- Grey parabolas represents accessible surface for large areas.
- Broken thick line around ligand shape indicates accessible surface.
- Size of residue ellipse represents the strength of the contact.
- 2D distance between residue label and ligand represents proximity.
- Covalently bound residue represented with thick black line around ellipse. (C25 in the picture above)

To edit the 2D image (e.g. rotate a group):

- Right click on the image and choose Edit Molecule and edit using the Molecular Editor.

To save the 2D interaction diagram to an image file:

- Right click on the image and choose Save Export/Molecular Image

9.7 Convert Chemicals to 3D

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

To convert a chemical structure from 2D to 3D:

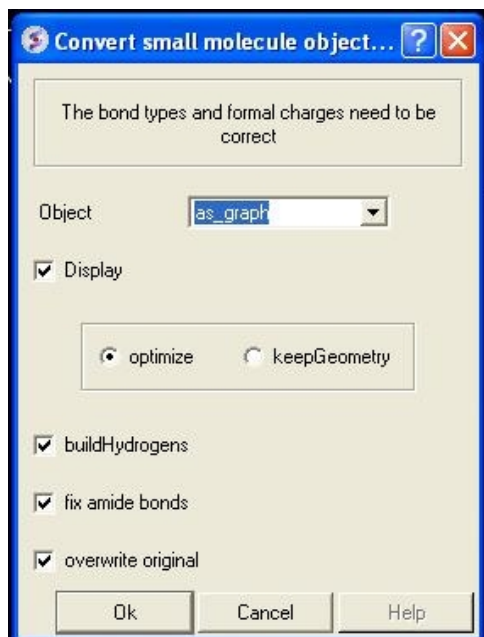
There are three ways in which to do this depending on whether you have a chemical in a chemical table or in the ICM workspace.

From the Chemistry/Convert/2D to 3D option menu option - convert a chemical in a chemical table.

- Read in chemical spreadsheet table.
- Chemistry/Convert/2D to 3D
- Select the table from the drop down list.
- Select to keep hydrogens and/or fix amide bonds.
- Keep current table or overwrite.

If the compound is in the ICM Workspace:

- Select all of the structure to be converted by double clicking on it in the ICM workspace or by using other selection tools described in the Selection Toolbar Section of this manual.
- MolMechanics/ICM-Convert/Chemical and a data entry box as shown below will be displayed.



If you have selected the compound as described above the "as_graph" option in the Object data entry box will suffice. You can decide whether you wish to keep the chemical geometry or optimize it in a force-field. Other options include whether you wish to add hydrogens and fix amide bonds.

From an ICM chemical table:

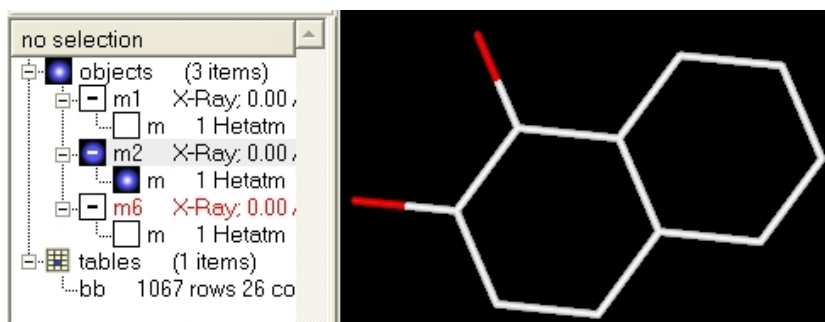
- Select which structures you wish to convert in the molecular table. For instructions on making selections within tables see the Making Table Selections part of this manual.
- Right click on one of the selections you have made and a menu as shown below will be displayed. Or use the Chemistry menu and select **Convert To 3D**.

mol	chiral	253.219	-3.01583	-2.79511
	chiral			
	chiral			

Right click on selected rows

- Select the **Chemistry/Convert to 3D and Optimize** option and you will see the compounds being converted and minimized in the 3D graphical display window.

Once converted the compounds will be displayed in the 3D graphical display window and also in the ICM workspace.



Converted chemical structures are displayed in the 3D graphical window and the ICM workspace

Another way to convert all the ligands contained within a table (or a selection) into 3D coordinates :

- Chemistry/Convert to 3D..
- Use the drop down list to select the Molecular Table.
- Select whether you want to **Keep Hydrogens**, **Fix Omegas** and **In Place**. Select **In Place** if you want to overwrite current table.
- Click OK
- If you wish to run in **Batch** mode select the **Files** option.

NOTE: Use the 3D-Browse mode to view the chemicals in the graphical display.

To convert 3D representation in a molecular table back to 2D:

- Chemistry/2D depiction
- Use the drop down list to select the Molecular Table.
- Select **In Place** if you want to overwrite the current table.
- Click OK.

9.7.1 Converting a Chemical from the PDB

The protein data bank has not been storing any information about covalent bond types and formal charges of the chemical compounds interacting with proteins! This oversight makes it impossible to automatically convert those molecules to anything sensible and requires your manual interactive assignment of bond types and formal charges for each compound in a pdb-entry. Therefore, if you apply the convert command to a pdb-entry with ligands, the ligands will just become some crippled incomplete molecules that can not be further conformationally optimized.

Therefore, follow these steps to convert a chemical properly from a pdb form to a correct icm object. There are two ways to do this either via the ICM Workspace (recommended) or via the Graphical Display.

9.7.2 Converting a Chemical from the PDB using the ICM Workspace

This is described in the `Convert Protein and Chemicals` chapter.

9.7.3 Converting a Chemical from the PDB using the Graphical Display

This is described in the `Convert Protein and Chemicals` chapter.

9.8 ToxScore

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

The **ToxScore** gives a prediction for how reactive or toxic a chemical is.

How the ToxScore is Calculated

Around 1000 SMARTS strings associated with toxicity/reactivity were collected from various sources. They were assigned demerit score based on their perceived toxicity and frequency of appearance in approved drugs. The toxscore of any compound was calculated by summing up all the matching SMARTS strings. A toxscore ≥ 1 . indicates likely toxicity based on substructure match.

How to Calculate ToxScore

- Read into ICM an SDF file. A chemical spreadsheet will be displayed in the GUI.
- insert a column into the chemical spreadsheet. Right click on a column header > choose insert column > choose the chemical function > select Tox Score.
- A column containing a Tox Score and Tox Names will be displayed.

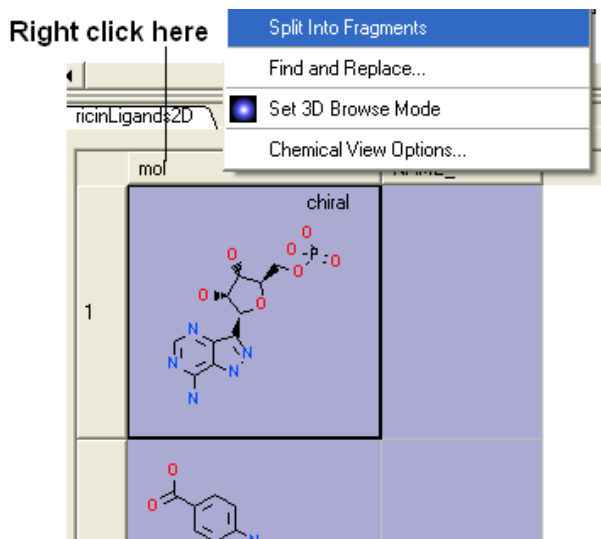
9.9 Generating Chemical Fragments

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

Chemicals displayed in an ICM Molecular Table can be split into fragments. This is useful for generating a series of R-groups to be added to a scaffold (See section describing reactions).

To generate fragments:

- Select the column or row(s) you wish to generate the fragment from.
- Right click on the "mol" column header and select "Split Into Fragments".
- A new table of chemical fragments will be displayed. Each fragment is assigned an attachment point which is flagged with an asterisk (*).



New table is generated containing fragments

mol	cnt	n_apo
	1	1
chiral	1	2

9.10 Molcart

Available in the following product(s): MolCart

Molcart is an enterprise wide chemical management system. Compound databases of any size can be stored in MolCart and analyzed and searched using ICM cheminformatic tools.

9.10.1 Molcart Installation

In order to run MolCart it is necessary to install the Free Open Source MySQL database on your machine. Please see your systems administrator or see <http://dev.mysql.com/downloads/>

Linux Installation

MySQL: Check if mysql daemon is running :

```
/etc/init.d/mysql status
```

If you see that the MYSQL service is unused (not running), you need to start the mysql daemon. Become root and do the following:

```
/etc/init.d/mysql start
```

Download and install Molcart files.

The MolCart package is a self-extracting executable file.

Installation Instructions:

- Download the MolCart file (`molcart-version-platform.sh`) from the Molsoft website.
- Type in a shell window: `./molcart-version-platform.sh -p=THE_PATH_YOU_WANT_TO_DOWNLOAD` (NOTE: You must be logged in as 'root' to install the 'molcart-version-platform.sh` to 'usr/molcart-version-platform'
- The following question will be displayed: Do you want to install the molcart-1.6-6 to "usr/molcart-1.6-6" now? (y/n) [y]
- Answer YES and the unpacking process will begin
- You will now be prompted for a password.
- Select which default databases you wish to install.
- Make a note of the HOSTNAME, DATABASE NAME and USER NAME
- MolCart is now fully installed.

Mac Installation

System requirements: * Mac OS 10.3 * MySQL server for Mac OS 10.3

To install Molcart on the Mac just run this:

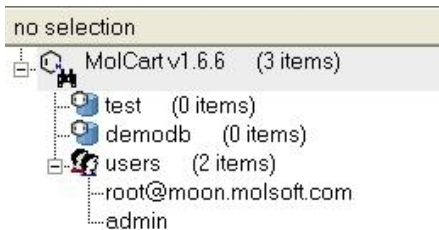
```
sudo /some/path/molcart-1.6-6-darwin.sh
```

9.10.2 Molcart Getting Started

To start Molcart:

Tools/Connect Molcart

Once you have activated MolCart the loaded databases and users will be shown in the ICM Workspace as shown below.

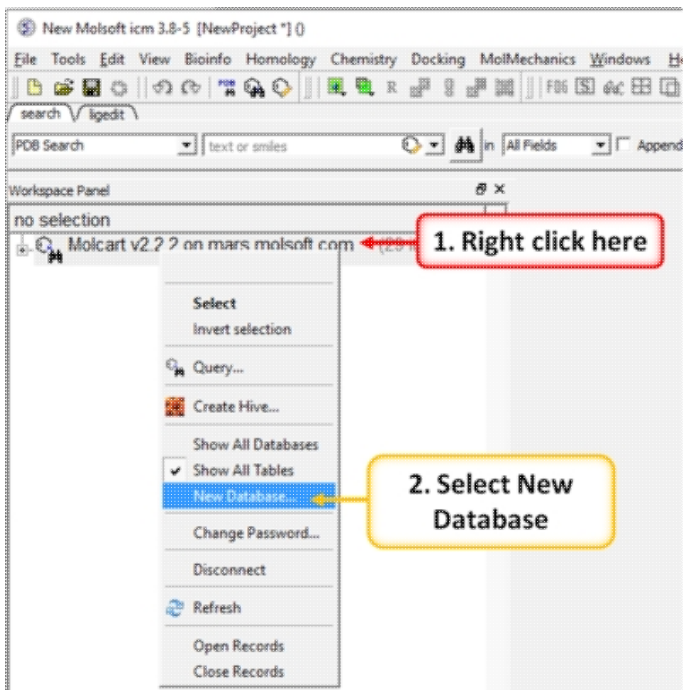


All the records and fields contained within each database can be viewed by expanding the tree structure in the ICM Workspace.

9.10.3 Creating Databases and Importing Data

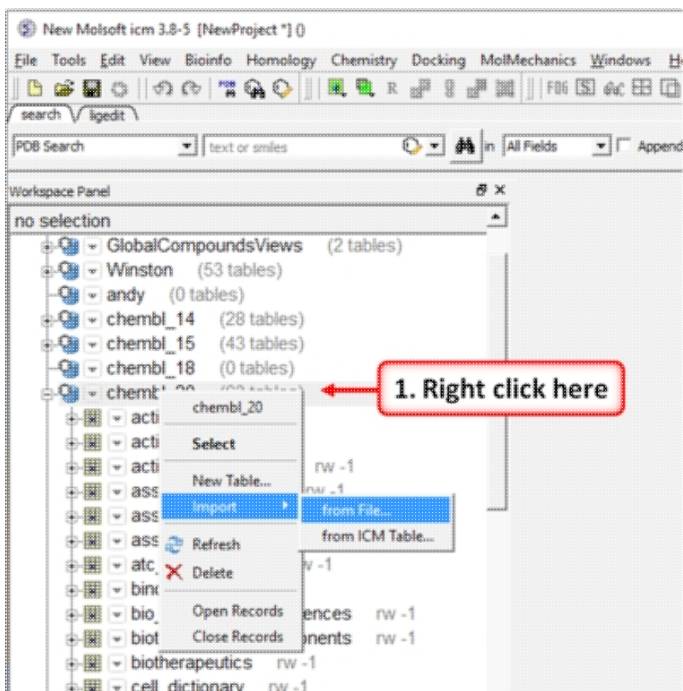
To make a new database:

- Right click on the top of the MolCart tree and then choose **New Database**



To upload SDF, csv, SMILES or MOL files or an ICM table:

- Right click on the database where you wish to add the file.
- Select Import.



9.10.4 Molcart Search

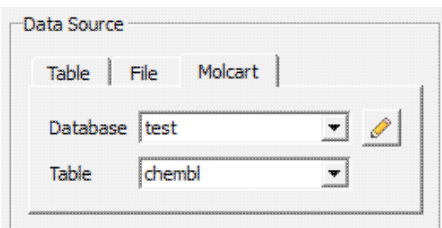
How to search the databases contained within MolCart

Click on the button shown below.



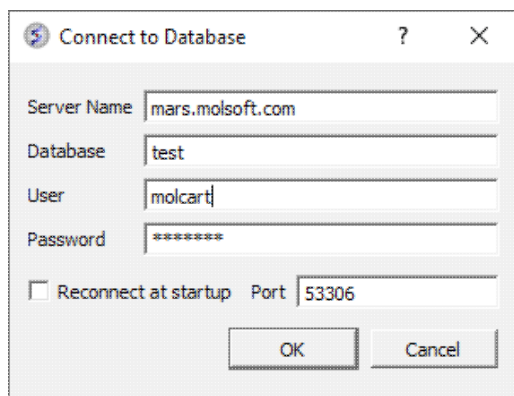
MolCart and Chemical Search Tools

To begin processing your query first you need to decide which database to search. The options are listed in the "Data Source" section of the ICM chemical search window.



You can either search a local table (molecular table) or you can search MolCart.

If you select MolCart you first need to setup the link to the correct database - described earlier. Click on the button shown above (yellow pencil) and the Connect to Molcart window will be displayed as shown below.



- Enter the Server Name in which the database is stored.
- Enter the database name.
- Enter your username and password for the server.
- You can save these details so you don't have to re-enter this information each time you use the chemical similarity search.

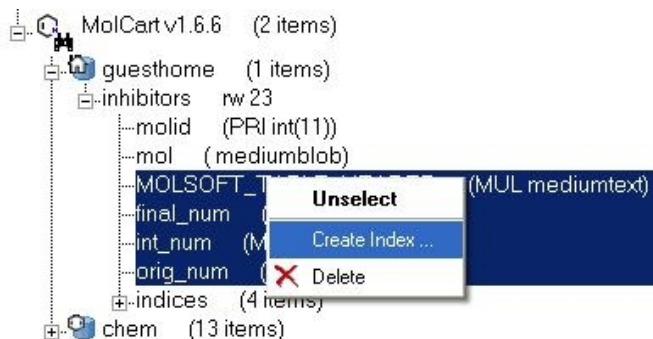
See the Chemical Search section of this manual on the many different search procedures.

How to perform a text search

To perform a text search on one of the databases contained within Molcart you first need to index the text within the database and then search using the query option.

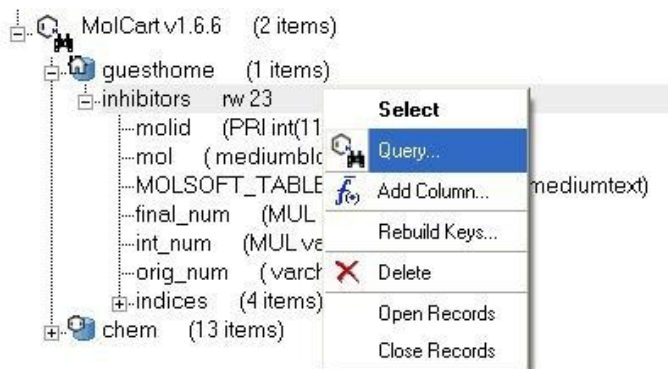
To make the database index, see text and picture below:

- Expand the tree of the database in the ICM Workspace.
- Select the column headers you wish to search which contain Full Text or Partial Text (the data type for each column is listed next to the column name). Multiple column headers can be selected by clicking and holding down the CTRL key. A range of column headers can be selected by holding down the shift key and clicking to select.
- Next, right click and select Create Index.
- Select 'Full Text' and you will notice an additional header in the ICM Workspace called 'indices'. The value in the items category represents the number of columns you have chosen to text search.

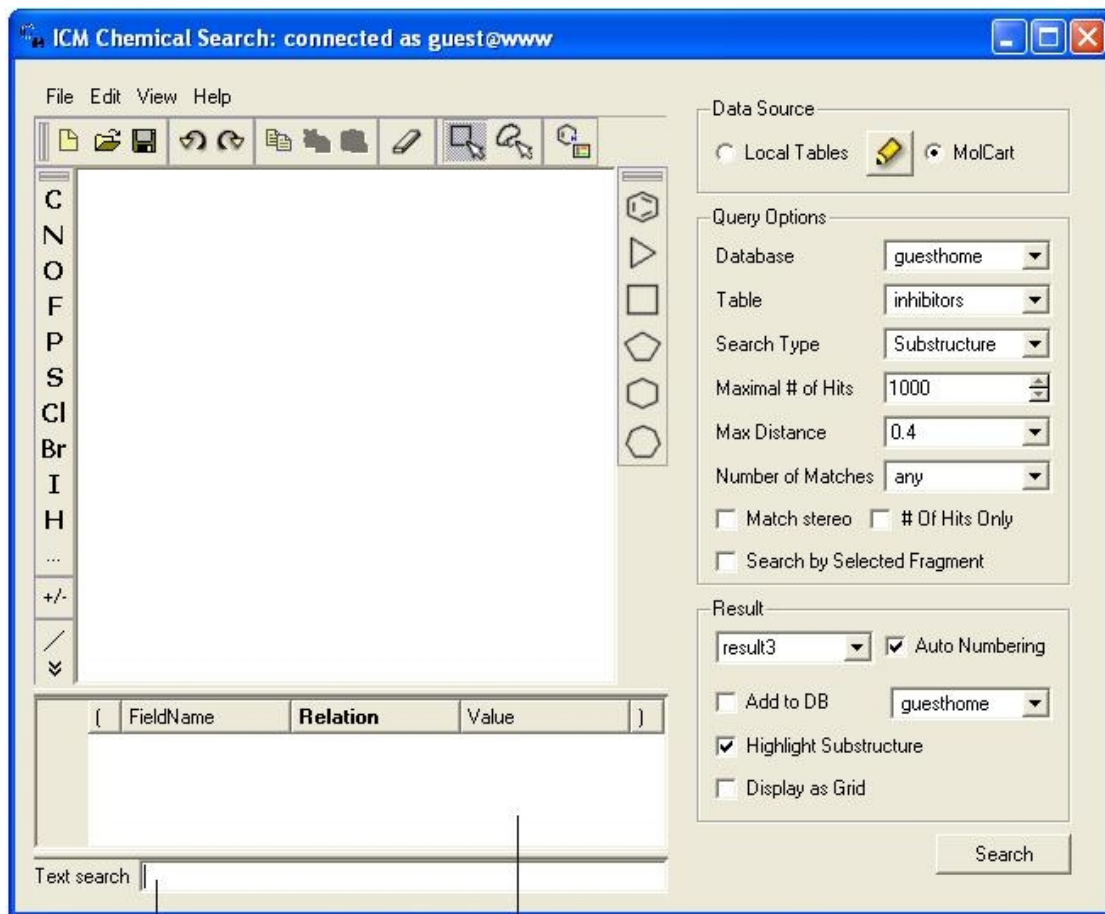


To perform the text search:

- Right click on the database name in the ICM Workspace.
- Select 'Query' as shown below.



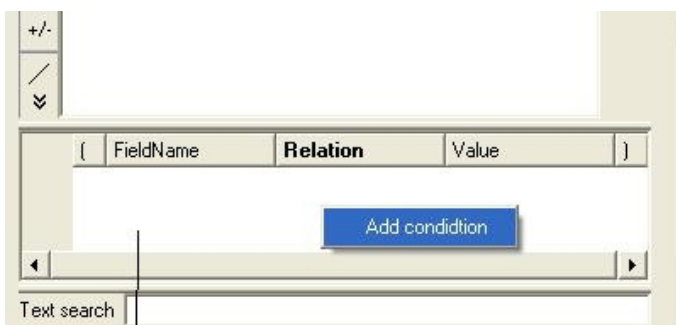
The ICM Chemical Search query window will be displayed as shown below. Type your query text in the space provided and hit the enter key or click on the search button.



Type text search query here Filter your search here

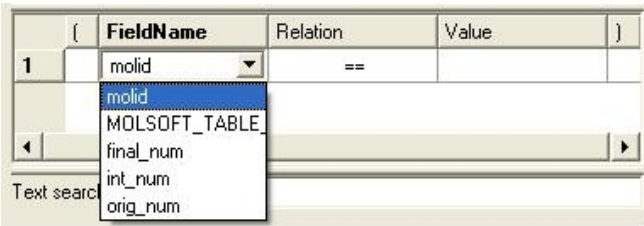
How to add conditions to your query

- Right click in the box shown below and select 'Add Condition'. You can add as many conditions as you like.



Right-click here and select Add Condition

- Double click in the fields labeled "Name" and "Relation" and select the options from the drop down arrow or type in values.



To remove a condition, right click on the filter and select 'Remove Filter'.

9.10.5 Molcart Administration

Edit Password

- Right click on the MolCart header in the ICM Workspace and select Change Password.

Add a New MolCart User - Root Only

- Right click on the User Section of MolCart in the ICM Workspace as shown below.



- Select the New User Option.
- Add new username and password in the data entry box.
- New user will be displayed in the ICM Workspace.

Edit User Privileges - Root Only

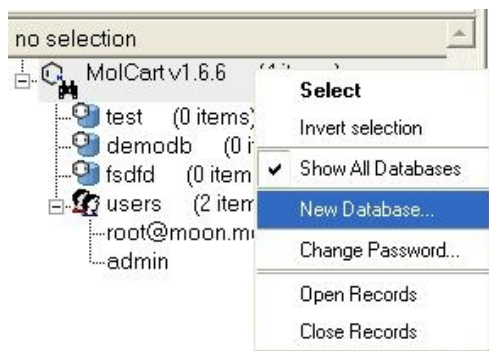
- Right click on the user in the ICM Workspace.
- Select Edit Privileges and a data entry box as shown below will be displayed.



- Select the database name.
- Edit the privileges by checking or unchecking the appropriate boxes.

Add a New Database

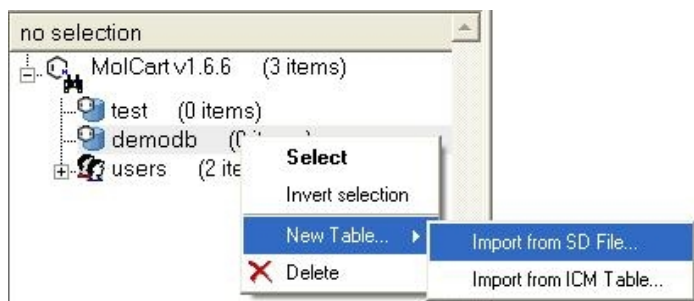
- Right click on the MolCart Header in the ICM Workspace as shown below.



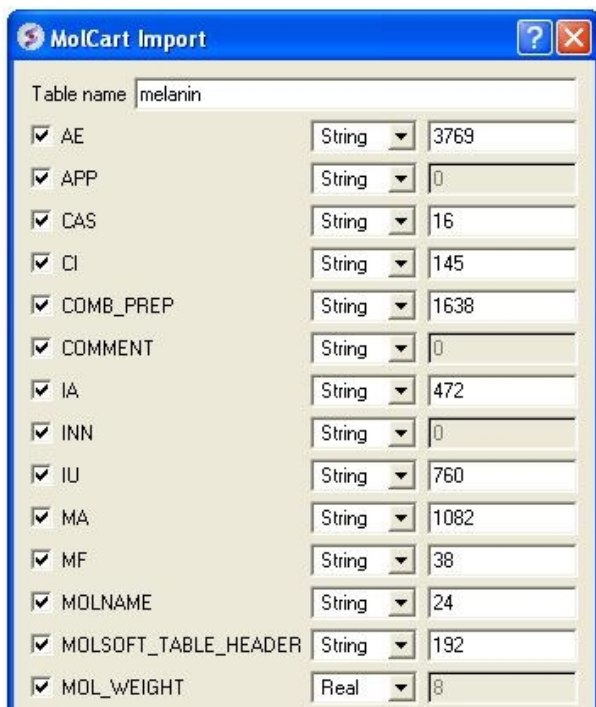
- Select the New Database option.
- Enter a unique name for your new database.
- The new database name will appear in the ICM Workspace.
- Now you need to add data to your new database (See Instructions Below).

Add New Data to Database

- Right click on the database name in the ICM Workspace as shown below.



- Select New Table.
- Select either Import from SD file or Import from ICM table.
- Select the appropriate file and the records structure of your sdf or ICM table will be displayed as shown below.



- The database name can be changed at this point and the fields contained within the database can be altered. Certain fields can be excluded by checking the boxes - this will help in minimizing the size of a database. Caution must be taken if you want to change the field type or length.
- Click OK and your sdf file or ICM table will be added to the database. This can be seen by expanding the tree structure in the ICM Workspace.

Delete a Database or User

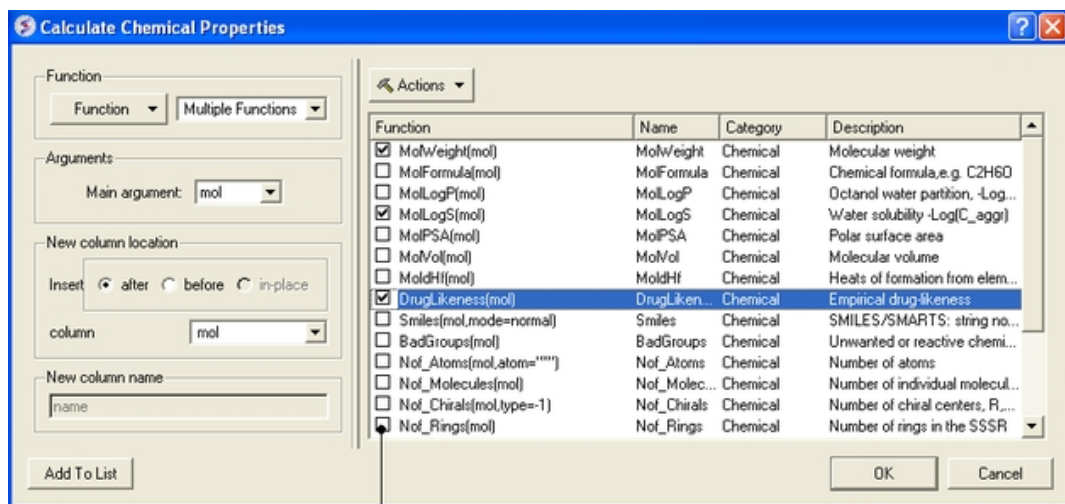
- Right click on either the database or user in the ICM Workspace.
- Select Delete.

9.11 Calculate Properties

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To calculate chemical properties for compounds within a chemical table:

- Read in the chemical table.
- Select **Chemistry/Calculate Properties** and a window as shown below will be displayed.
- **Arguments** - usually the column with the chemical structure is called mol. If this is not the case click on the drop down button and select the correct column.
- **New column location** - select where you would like to insert the new columns.
- **New column name** - This option is usually greyed out but the column name will be the same as the property name.
- Select the properties you wish to calculate using the 'tick' check boxes.
- Click OK and the properties will be added as new columns in the chemical table.



Select properties here

Actions Button Helps with Selections

If you click on the drop down button labeled **Actions** you will see 3 selection options:

- Select All properties.
- Uncheck selected options.
- Remove unchecked.

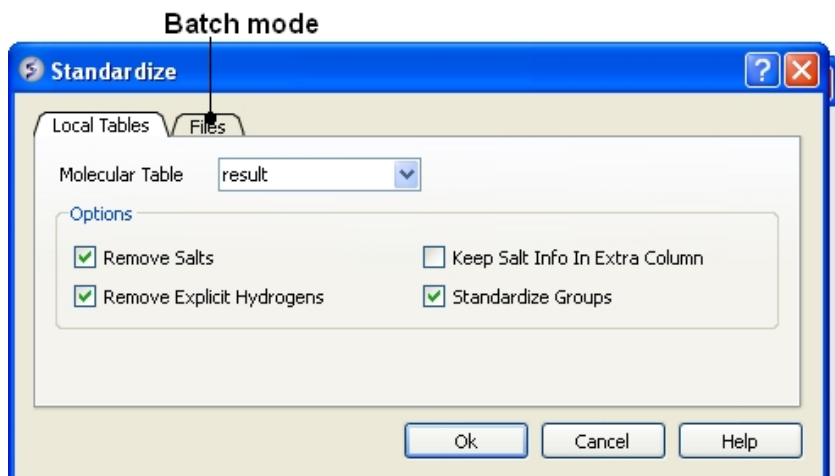
9.12 Standardize Table

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

NOTE: Before standardizing a chemical you may want to make a copy of the original so you do not lose any information. You can do this by right clicking on the name tab of the table and selecting **clone** or **save as**.

To remove salts, explicit hydrogens and standardize groups in a chemical table:

- Chemistry/Standardize
- Select the table from the drop-down list. This will also work for selections made on the table.
- Select option - Remove Salts, Keep Salt Info in Extra Column (NOTE: option available in 3.7-2b and higher), Remove Explicit Hydrogens or Standardize Groups
- Click OK.



Standardize chemical groups will apply rules from ICMHOME/CHEMNORMRULES.tab For example some chemical groups may have different representations e.g. [N+] (=O)O versus N(=O)=O

To run in batch mode:

- Chemistry/Standardize
- Select the **Files** tab
- Enter the path and name of the sdf file you wish to standardize or use the browse button.
- Enter the path and name of the output file or use the browse button.

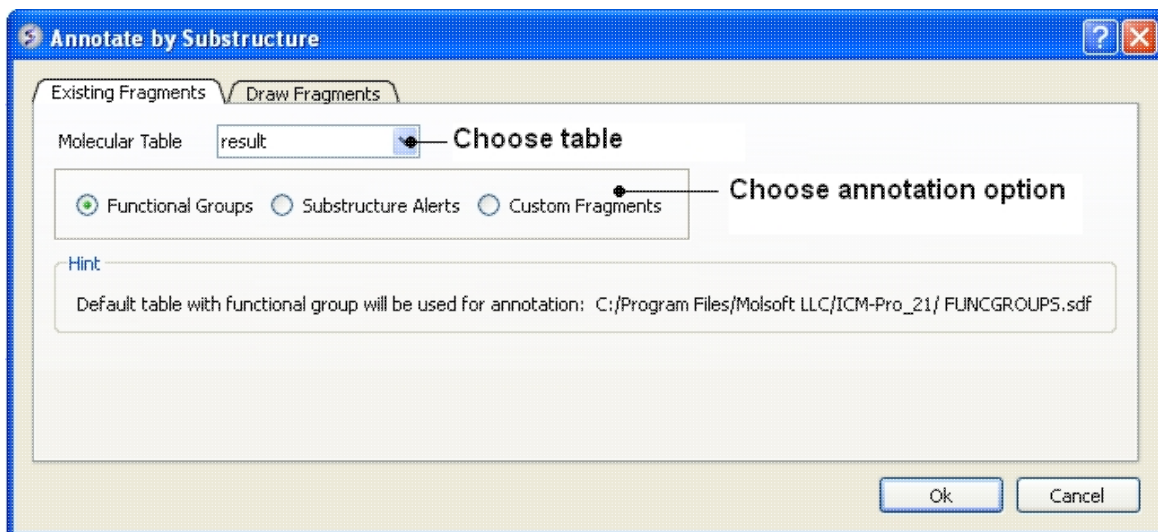
To remove a salt from an individual row in a chemical table:

- Select the row or rows.
- Right click - Chemistry/Remove Salt (Selected Row)

9.13 Annotate By Substructure

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

This feature allows you to annotate a chemical spreadsheet according to functional group. It also allows you to flag substructures which may have poor ADME properties.



First read in a chemical spreadsheet or sdf file you wish to annotate. To do this:

- Information on how to read into ICM chemical spreadsheets and structure can be found [here](#).

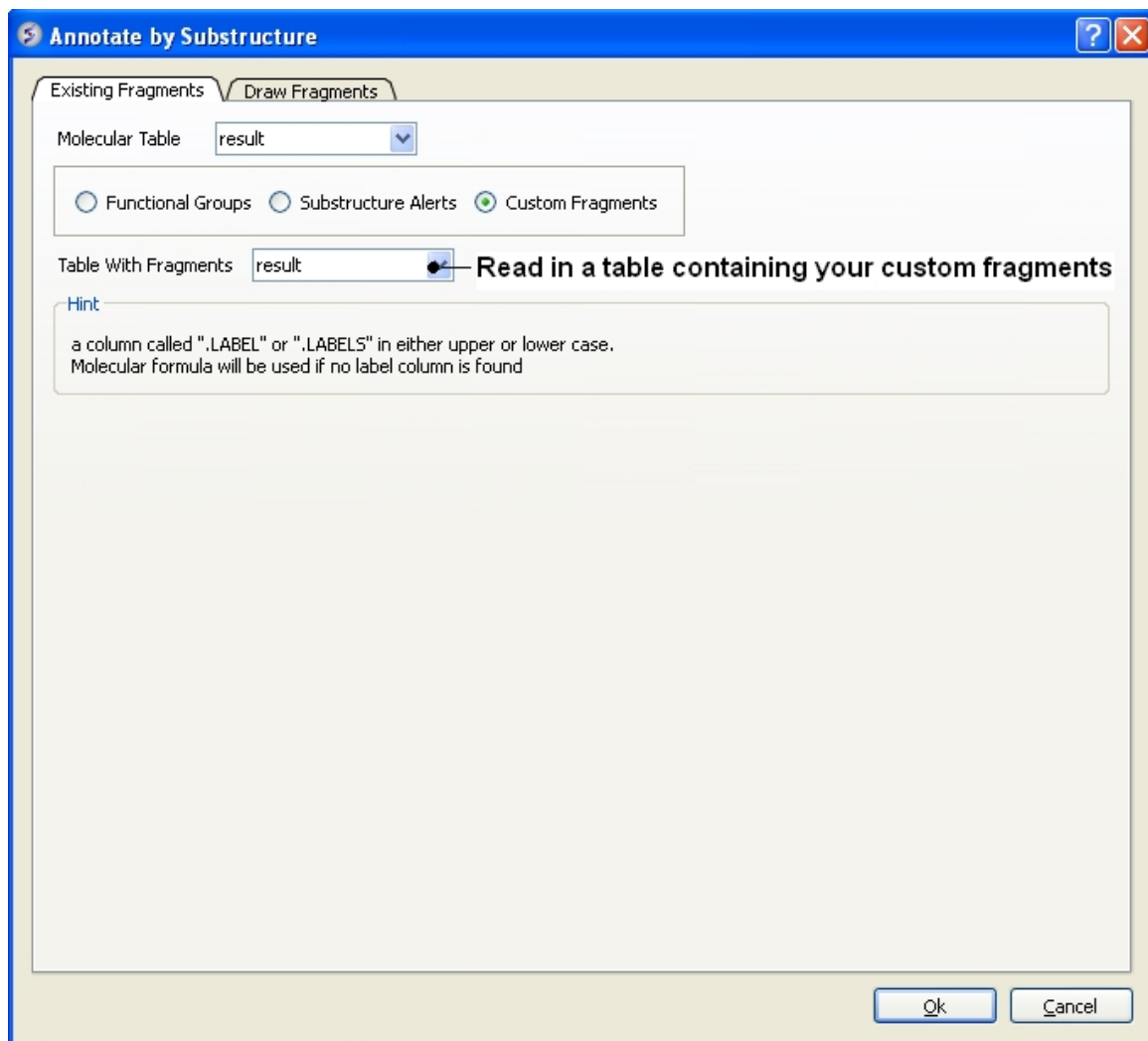
To annotate functional groups:

- **Chemistry/Annotate** by Substructure.
- Enter the name of the Molecular Table (Chemical Spreadsheet) or use the drop down button to locate it.
- Check the **Functional Groups** option.
- The functional groups will be listed in a new column in your chemical spreadsheet called **funcgroup**. The default table with functional group will be used for annotation called FUNCGROUPS.sdf in ICMHOME.

To annotate potentially poor ADME groups (Substructure Alert).

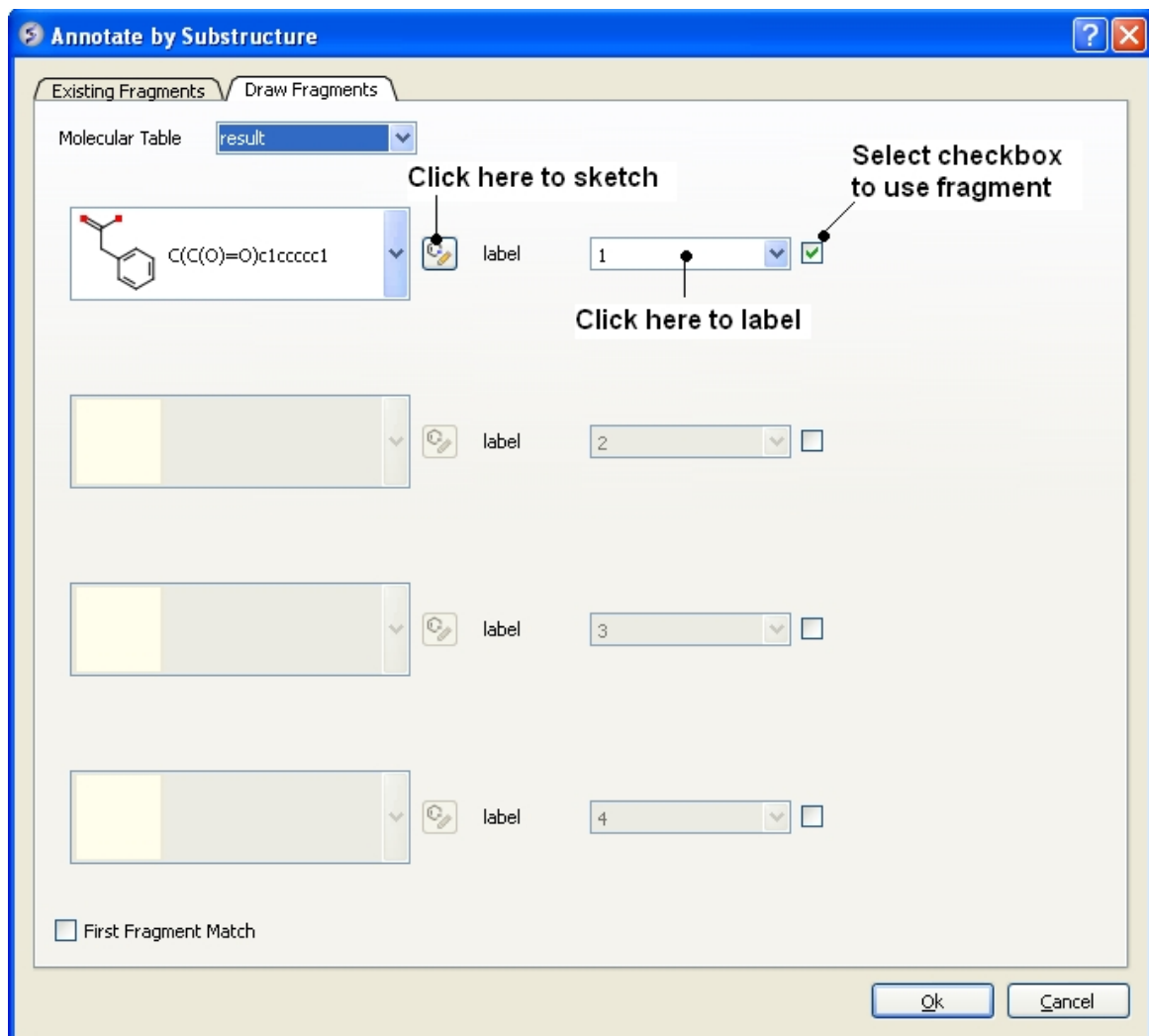
- **Chemistry/Annotate** by Substructure.
- Enter the name of the Molecular Table (Chemical Spreadsheet) or use the drop down button to locate it.
- Check the **Substructure Alerts** option.
- The alerts will be listed in a new column of your chemical spreadsheet called **alerts**. The default table with substructure alerts will be used for annotation called CHEMFILTER.sdf

To annotate by custom fragments



- Read into ICM a table with annotated fragments. The fragments should be labeled in a column called ".LABELS" or ".LABEL" in either upper or lower case. If not "LABELS" column is found the molecular formula will be used instead.
- **Chemistry/Annotate** by Substructure.
- Check the **Custom Fragments** button

To draw a fragment and annotate.



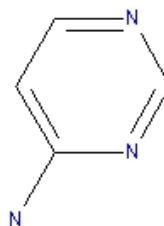
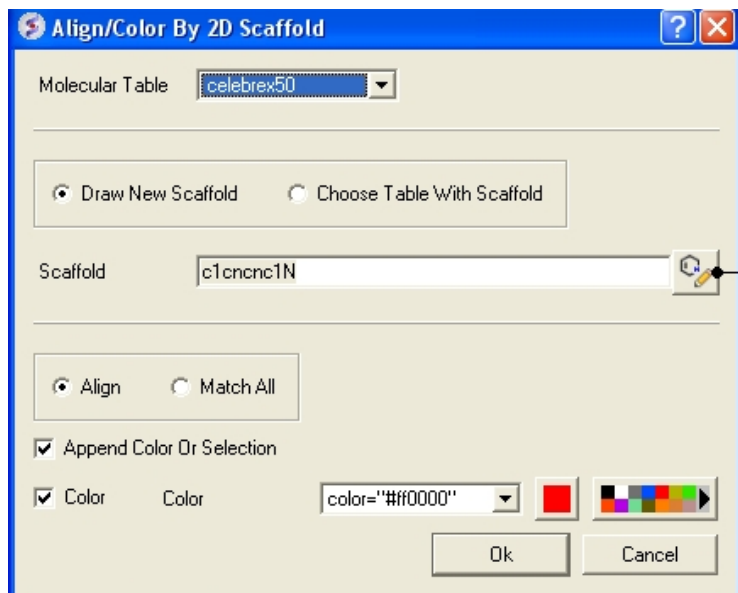
- Click on the **Draw Fragments** tab.
- Enter the SMILES string of the fragment or click on the ligand editor button in the dialog window to sketch the fragment.
- Enter a label name.
- Click the check box next to the label name.
- Click OK

9.14 Align/Color by 2D Scaffold

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

This option **aligns** a set of sketches in a chemical table in the same orientation according to a defined scaffold or **color** by a common substructure.

- Chemistry/Align-Color by 2D Scaffold
- Choose a loaded molecular table from the drop-down arrow.
- Draw a new scaffold using the molecular editor or choose the scaffold from a table (Index = row number).
- You can then **Align** or **Match All** the substructure and color.
- If coloring has already applied to the molecule then this new coloring by scaffold can be appended.
- Select **color** for common scaffold.



Draw scaffold using molecular editor

9.15 Set Formal Charges

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To set formal charge:

- Read in a chemical spreadsheet **File/Open SDF**
- **Chemistry/Set Formal Charges**
- Enter the name of the chemical spreadsheet (table).
- Choose the option to either make **Charge, Make Neutral**, or Charge according to ICM pKa Model
- Select the groups you want to charge or neutralize.
- The charges will automatically be added to your table.

9.16 Enumerate Formal Charge States

To enumerate formal charge states.

- Read in a chemical spreadsheet **File/Open SDF**
- Choose **Chemistry/Enumerate Formal Charge States**
- Enter the name of the chemical spreadsheet (Molecular Table)
- Choose the pH at which you wish to evaluate the charge
- Choose a window value.
- Click OK and a new table "_charge" will be displayed with the enumerated charge states.

About the Method

For basic charges the logic is following:

- atoms with $pK_a > pH + \text{window}$ is always charged
- atoms with pK_a inside the range $[pH - \text{window} : pH + \text{window}]$ will be enumerated as charged and uncharged.

For example if you have two centers with pK_a within the range then four states will be generated ([0,0] [+1,0] [0,+1] [+1,+1])

For acids it is similar.

9.17 Convert

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

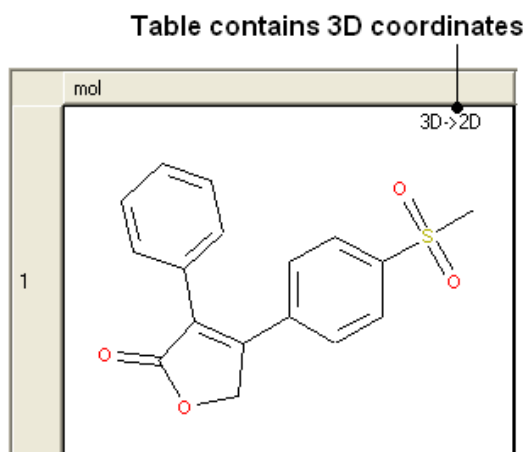
9.17.1 Convert 2D to 3D

- Read a chemical table (sdf file) into ICM.
- Chemistry/Convert/2D to 3D
- Select the table from the drop down list.
- Select to keep hydrogens and/or fix amide bonds.
- Keep current table (In Place) or overwrite.

9.17.2 3D to 2D Depiction

If you have a chemical table displayed containing 3D coordinates or you wish to reassign the 2D coordinates in an sdf file you can use this option.

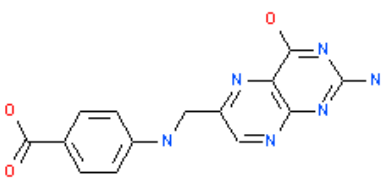
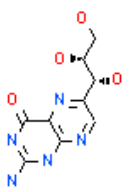
- Read a chemical table (sdf file) into ICM.
- Chemistry/Convert/2D Depiction
- Enter name of loaded chemical table.
- Choose **In Place** if you want to overwrite the table.
- Choose the **Files** tab to run in batch mode.



9.17.3 Convert Smiles to 2D

To convert smiles strings to 2D sketches

- Read in a table containing the smiles strings in separate rows. For example the smiles strings maybe in an Excel file and you can load this into ICM by saving the Excel file as comma-separated (csv).
- Select **Chemistry/Convert/Smiles to 2D**.
- Select the table you want to convert using the drop down arrow and the name of the column containing the smiles string.
- Select whether you wish to keep the smiles column in the new table.
- Click OK and a table will be displayed containing the 2D structure.

ricinLigands2D	
mol	smiles
2	 C1 (C=NC(=C(N=1)C1O)N=C(N=1)N)CNC(=CC=C(C1)C(O)=O)C =1
3	chiral  C(C1=O)(=C(NC(=N1)N)N=CC1[C@H]([C@@H] (CO)O)O)N=1

9.17.4 Convert Structure to Smiles

To convert an sdf file of 2D or 3D chemical coordinate in Smiles:

- Read a chemical table (sdf file) into ICM.
- Select **Chemistry/Convert Structure to Smiles**.
- Select the table you want to convert using the drop down arrow and the name of the column containing the 2D sketch.
- Select whether you wish to keep the 2D sketch column in the new table.
- Click OK and a table will be displayed containing the smiles string.

9.17.5 Convert to Racemic

To remove stereo bonds and make all chemical centers R/S in a chemical table:

- **Chemistry/Convert to Racemic**
- Select the Molecular Table
- Select **In Place** if you wish to overwrite the table.



NOTE: To reassign stereo bonds use the `Generate Stereoisomers` option

9.18 Build Prediction Model

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

Structure-Activity Relationship (SAR) is a process by which the activity of a molecule is related to its molecular structure. If a significant amount of structural and activity data is available a model can be made which can be used to predict the activity of a molecule or series of molecules.

In ICM SAR is undertaken using the Learn and Predict tools in a Molecular Table.

9.19 Predict

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

To make a prediction using a model you need either ICM-Pro + Chemistry or ICMChemistPro.

Read the table of data into ICM from which you wish to predict. Make sure the table contains the same columns used for the learn model.

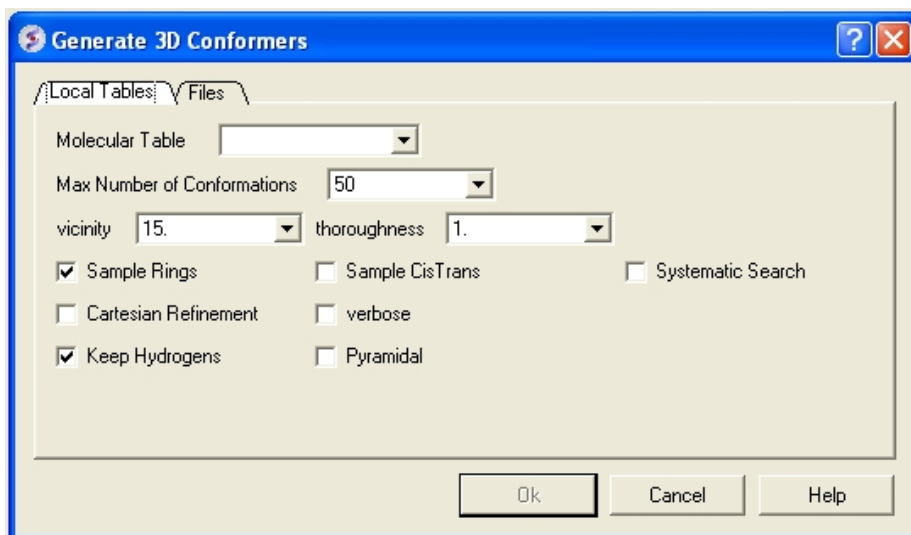
- Tools/Table/Predict
- Select which table you wish to make the prediction on.
- Select which model you wish to use.
- Check that the required columns are in the table. If they are absent a red mark will appear against the column that is missing.
- Click Predict.

9.20 Generate 3D Conformers

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

To generate a series of conformers for a ligand(s):

- Select the compounds (row(s)) you wish to generate conformers for in an ICM Molecular Table . Or to convert a whole table of compounds select **Chemisty/Generate 3D Conformers** menu.
- Right click on the selected row(s) and Chemistry/Conformation Generator (selected rows) and a data entry window as shown below will be displayed.



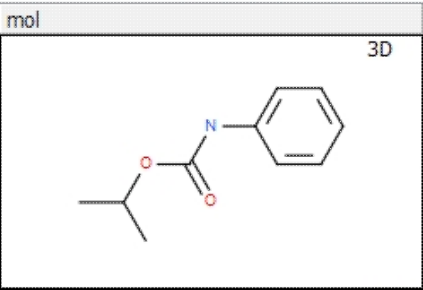
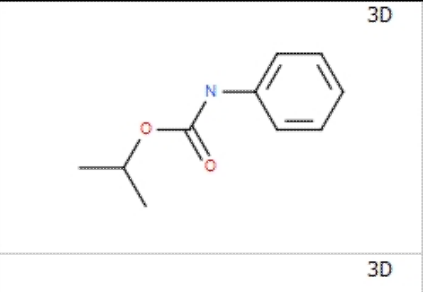
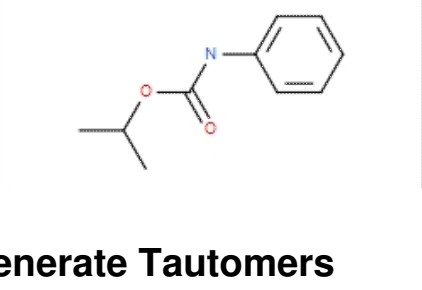
- Enter the maximum number of conformations you wish to generate.
- Enter a vicinity value. For more information on vicinity please see the command language manual <http://www.molsoft.com/man/reals.html#vicinity> . Density of conformers is

controlled by vicinity= parameter. Currently it always operates in terms of torsion RMSD, default is 30 degrees angular RMSD. Note that we only generate local minima, therefore reducing vicinity below some level will typically not produce any new conformers.

- Enter a effort value (thoroughness in releases before 3.8-2) . This relates to the length of sampling time. Effort multiplies the sampling length determined by default heuristic based on size and number of torsions.
- Check boxes for Sampling Rings, Systematic Search, Cartesian Refinement (<http://www.molsoft.com/man/real.html#vicinity>), Sample Cis and Trans, sample Pyramidal and Verbose (Display Warnings). The 'Cartesian' flag enables cartesian minimization/refinement of conformations (produced by torsion-only sampling). This should be checked whenever energetics is of importance. Torsion-only sampling is fast but only gives qualitative sample of accessible conformational space, relative energies of different conformers may not be accurate.
- Click OK and the sampling will be undertaken in the background - see Windows/Background Jobs
- Once the sampling has finished a table as shown below will be displayed. Click in the toggle boxes in the column labeled "L". To display more than one 3D structure at a time use the row selection options.

Click in box to display ligand in 3D Row number from original table

Original table Results table Unique name Energy rank Energy value

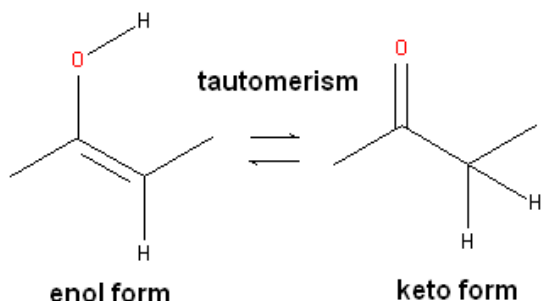
celebrex50		conformers						
	mol	L	Molecule	MOL NUM	CONF NUM	ENERGY		
1		<input checked="" type="checkbox"/>	sdmol_5023d	1	1	0		
2		<input type="checkbox"/>	sdmol_5023d	1	2	0.4241		
3		<input type="checkbox"/>	sdmol_5023d	1	3	6.29		

9.21 Generate Tautomers

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

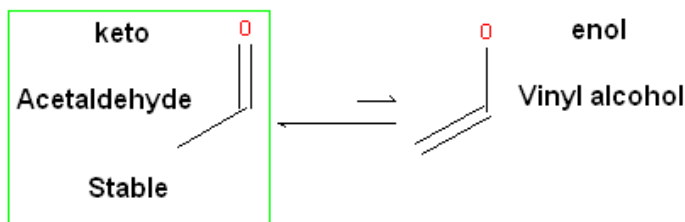
Theory

Tautomers are formed by an interconvertible reaction called tautomerization whereby there is a formal migration of a hydrogen atom along with a switch of a single bond and an adjacent double bond. A common example is the keto to enol tautomerism:



During tautomerization a chemical equilibrium of the tautomers will be reached based on several factors, including, pH, temperature and solvent. Tautomerizations are catalyzed by: bases (deprotonation, formation of a delocalized anion, and, protonation at a different position of the anion; and acids (protonation, formation of a delocalized cation, and deprotonation at a different position adjacent to the cation).

ICM will only generate energetically favorable tautomers. Generally tautomers that have a change in hybridization state are less stable and so ICM will not generate these thus reducing the number of scaffolds generated. For example the keto form shown below is more stable by ~14 kcal.mol than the enol therefore ICM will not generate the enol form.



To generate tautomeric conformations of your compound:

- Select the compound(s) in the molecular table. Selected compounds are highlighted in blue.
- Right click in the table and select the option Chemistry/ Enumerate Tautomers. Or select the Chemistry menu/Generate Tautomers.
- If you select Chemistry/Generate Tautomers a dialog box will be displayed. There is also an option to run in batch mode (click the **Files** Tab).
- Choose the table containing the compounds using the drop-down list.
- **Filter Unwanted Groups** option will filter results from patterns in the TAUTOFILTER.tab file provided in the distribution. If results match any row from that table then it will be excluded.
- **Preserve Hybridisation** Although generally a change in hybridisation state will generate less stable compounds in some cases this is not the case and so you can choose to change hybridisation for a single atom.
- **Group Rows With Color** option will color tautomers from the same compound with the same color to visually highlight groups.
- The compounds will be displayed in a separate molecular table called TAUTOMERS.

How to generate tautomers.

The screenshot shows the ICM-Chemist interface with a table of chemical structures. A menu is open on the left, highlighting 'Generate Tautomers'. A dialog box titled 'Generate Tautomers' is open, showing a dropdown menu for 'Molecule Table' with 'my_database' selected. Callout boxes point to the menu item and the dialog box.

mol	NAME	molid	version	molweight mol	mol_log mol	mol	id		
1	vacant	C2H6O2	129	Chembridge 1140002	Bechemicals F20140009	301.2	2.907	1	1
2	vacant	C2H6O2	501320	Bechemicals F20140201		506.2	4.137	2	2
3	vacant	C2H6O2	501320	Bechemicals F20140105		436.2	2.80	3	2
4	vacant	C2H6O2	501320	Bechemicals F20140105		301.2	2.938	4	2
5	vacant	C2H6O2	501320	Bechemicals F20140105		474.2	2.97	5	3
6	vacant	C2H6O2	501320	Bechemicals F20140105		434.2	3.284	6	4

9.22 Generate Stereoisomers

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

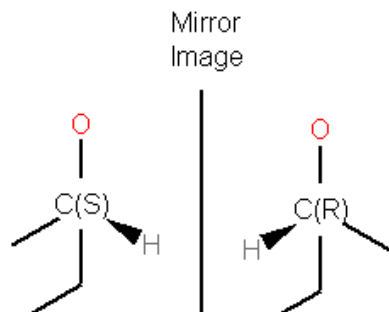
Theory

Isomers are molecules which have the same chemical formula and sometimes the same kind of bonds but in which the atoms are arranged differently.

Structural isomers have different atom-to-atom connections e.g. propanol (C₃H₈O or C₃H₇OH) has two isomers Propan-1-ol and Propan-2-ol.

Diastereomers are not mirror images and have different internal dimensions (e.g. dihedral angles and distances between non-bonded atoms). They can be configurational diastereomers (which can be interconverted only by breaking bonds or by changing the configurations of stereocenters) or conformational diastereomers (which can be interconverted by rotation about bonds - including chair flips or by lone pair inversion).

Enantiomers have identical internal dimensions, and are nonsuperimposable mirror images. Enantiomers can be configurational and conformational.



Enantiomers are distinguished based on the Latin terms for left (sinister) and right (rectus). In some cases where the handedness is unknown a chiral center can be labeled "RS" or unknown.

To enumerate and display in a separate table the stereoisomers of selected compounds.

- Select the compound(s) in the molecular table. Selected compounds are highlighted in blue.
- Right click in the table and select the option **Chemistry/ Enumerate stereoisomers**. Or select the Chemistry menu and choose **Generate Stereoisomers**. If you generate stereoisomers via the Chemistry menu you will get a dialog box whereby you can run the process in batch mode. There is also an option to color stereoisomers from the same compound with the same color.
- **Group Rows With Color** option will color stereoisomers from the same compound with the same color to visually highlight them.
- **Skip if Number of Centers Greater than** option can be used to filter out compounds which have a defined number of stereo centers. If you do not wish to filter just use the default **No limit** option.
- The compounds will be displayed in a separate molecular table called STEREOISOMERS.

From the Chemistry Menu:

How to generate stereoisomers.

1. Chemistry/Generate Stereoisomers

Choose "Files" for batch mode

2. Enter name of chemical table

name	formula	weight	vendor	noRight	noLeft	no	id
	C10H16O	152	chembridge 314002 BiochemLab F3214-0008	301.2	2.80	1	1
	C10H16O	152	BiochemLab F3214-0201	306.2	4.917	2	2
	C20H32O	284		436.2	2.80	3	2
	C20H32O	284		301.2	2.938	4	2
	C20H32O	284		474.2	2.97	5	3
	C20H32O	284	BiochemLab F3214-0103	434.2	3.288	6	4

```

low/de> delete hydrogen my_database.mol
Info: 3297 hydrogen removed
low/de> modify my_database.mol auto
Info: 805 replacements done
low/de> delete my_database.cluster(1)
low/de>
  
```

From a chemical spreadsheet:

ricinLigands2D STEREOISOMERS

mol	NAME_
1	chiral ricinLigands2D
2	Select one or multiple compounds and right click on the drawing.
3	chiral
4	

Results will be in the table entitled "STEREOISOMERS"

- Edit Cells by Double-click
- Table View
- Find in Table Ctrl+F
- Color By
- Clustering...
- Analysis
- Copy Molecule Image
- Copy Molecule Image To Album
- Save Molecule As EPS...
- Save Molecule As Image...
- Edit Molecule...
- Query Molecule
- Chemistry
 - Enumerate Stereoisomers (current row)
 - Enumerate Tautomers (current row)
 - Remove Salt (current row)
 - Convert to 3D and Optimize
 - Load and Preserve Coordinates
 - Conformation Generator (selected rows)
 - Chemical Template Superposition (selected rows)
- Cell mol[1]
- Copy Molecule Ctrl+C
- Paste Ctrl+V
- Print Table... Ctrl+P

NOTE: Only centers with unknown chirality will be enumerated.

9.23 Prodrug

To generate a drug from a prodrug:

- Read in a chemical table (sdf file) into ICM.
- Choose the menu Chemistry/Generate Drug from Prodrug and a dialog box will be displayed.
- Enter the name of the Chemical Table.
- Select the reactions you wish to use to generate the drug. You can choose "Prodrug", "All", "Hydrolysis" "Reduction". You can choose the option "all" to use all the available reactions. You can see the reactions in each subset by clicking on the tab labeled "Load Reaction Panel". A chemical table containing the reactions will be displayed and the reaction is labeled in a column called "subset".
- You can choose to output the results in the "in-addition" mode where the converted drug will be in the row below the prodrug or "replace original" whereby the prodrug will be replaced with the converted drug.

To generate a prodrug from a drug:

- Read in a chemical table (sdf file) into ICM.
- Choose the menu Chemistry/Generate Prodrug from drug and a dialog box will be displayed.
- Enter the name of the Chemical Table.
- Select the reactions you wish to use to generate the prodrug. You can choose "carboxylate", "hydroxyl", "amino", "diol", "sulfide". You can choose the option "all" to use all the available

reactions. You can see the reactions in each subset by clicking on the tab labeled "Load Prodrug Panel". A chemical table containing the reactions will be displayed and the reaction is labeled in a column called "subset".

- You can choose to output the results in the "in-addition" mode where the converted prodrug will be in the row below the drug or "replace original" whereby the drug will be replaced with the converted prodrug.

9.24 Ligand Energetics

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

9.24.1 Conformational Entropy

To calculate the conformational entropy of chemicals in an SDF file or a chemical spreadsheet:

- **Chemistry/Ligand Energetics/Conformational Entropy**
- Choose either the **Local Tables** (loaded chemical spreadsheet) or {Files} (SDF file) tab.
- Enter the maximum number of conformations you wish to generate.
- Enter a **vicinity** value. For more information on vicinity please see the command language manual <http://www.molsoft.com/man/reals.html#vicinity>
- Enter a **effort** value. This relates to the length of sampling time.
- Check boxes for Sampling Rings or Systematic Search (instead of Monte Carlo).
- Click OK.

The calculation will then run and you will see a progress bar in the bottom right hand corner of the GUI or you can check the progress of the job go to **Windows/Background**. When the job has finished:

- A table will be displayed called **entropies** with a column called **Conf Entropy**. The energy is in kcal/mol.

9.24.2 Strain

To calculate the conformational strain of chemicals in an SDF file or a chemical spreadsheet:

- **Chemistry/Ligand Energetics/Strain**
- Choose either the **Local Tables** (loaded chemical spreadsheet) or {Files} (SDF file) tab.
- Enter a **effort** value. This relates to the length of sampling time.
- Check the **Systematic Search** option if you do not want to use Monte Carlo search.

The calculation will then run and you will see a progress bar in the bottom right hand corner of the GUI or you can check the progress of the job go to **Windows/Background**. When the job has finished:

- A table will be displayed called **strain** with a column called **Strain**. The energy is in kcal/mol.

9.25 Cluster Set

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

Clustering is described in more detail in the `Tables Clustering` section of this manual. To undertake chemical clustering choose:

- Chemistry/Cluster Set

9.25.1 How to perform chemical clustering.

How to perform chemical clustering.

1. Click on the "Run Clustering" button

ord	name	mol	vendor	subInq
1	C2H3O2N2O2	129 ChemBridge 914002	Bechemcal F334-0008	30
2	C2H3O2N2O2	501323 Bechemcal F334-0190		43
3	C2H3O2N2O2	501320 Bechemcal F334-0190		35
4	C2H3O2N2O4	501322 Bechemcal F334-0190		43
5	C2H3O2N2O2	501324 Bechemcal F334-0193		43

```

low/de> make tree my_database full "OPERA" split="cl" label="SHARE_" name=""
Info: Column 'ord' (record number in the tree order) has been appended to table my_database.
Info: Column 'cl' (cluster number) has been appended to table my_database.
Info: Tree 'my_database.cluster[1]' has been successfully created.
low/de> sort my_database.ord
low/de>
  
```

9.25.2 How to select representative centers from a tree.

How to select representative centers from a tree.

Right click in white space for all centers or on a node for centers of a branch

ord	name	mol	vendor	subInq
1	C2H3O2N2O2	129 ChemBridge 914002	Bechemcal F334-0008	30
2	C2H3O2N2O2	501323 Bechemcal F334-0190		43
3	C2H3O2N2O2	501320 Bechemcal F334-0190		35
4	C2H3O2N2O4	501322 Bechemcal F334-0190		43
5	C2H3O2N2O2	501324 Bechemcal F334-0193		43

```

low/de> make tree my_database full "OPERA" split="cl" label="SHARE_" name=""
Info: Column 'ord' (record number in the tree order) has been appended to table my_database.
Info: Column 'cl' (cluster number) has been appended to table my_database.
Info: Tree 'my_database.cluster[1]' has been successfully created.
low/de> sort my_database.ord
low/de>
  
```

9.25.3 How to reorder branches and change the distance of trees.

How to re-order branches and change distance of tree.

To change distance – click and drag here

To re-order branches – right click in white space

Close Clusters
Open All
Select Centers
Select Centers by:
Reorder Branches by Field
Reorder Table by Cluster
Print...
Save/Export Image
Save As...
Change Record Labels
Preferences

id	name	mol	vendor
1	C2H4O2N2	129 chembridge 714002	Bechemicals 7204-0008
2	C2H4O2N2	561320 Bechemicals 7204-0100	
3	C2H4O2N2	561330 Bechemicals 7204-0100	
4	C2H4O2N2	561330 Bechemicals 7204-0100	
5	C2H4O2N2O4	561322 Bechemicals 7204-0100	
6	C2H4O2N2	561324 Bechemicals 7204-0103	

9.25.4 How to edit the tree - labels, spacing and coloring.

How to edit the tree – labels, spacing and coloring.

To edit labels, spacing and coloring: right click on the tree

To change record labels: right click in white space and select "Change Record Labels"

Close Clusters
Open All
Select Centers
Select Centers by:
Reorder Branches by Field
Reorder Table by Cluster
Print...
Save/Export Image
Save As...
Change Record Labels
Preferences

id	name	mol	vendor
1	C2H4O2N2	129 chembridge 714002	Bechemicals 7204-0008
2	C2H4O2N2	561320 Bechemicals 7204-0100	
3	C2H4O2N2	561330 Bechemicals 7204-0100	
4	C2H4O2N2	561330 Bechemicals 7204-0100	
5	C2H4O2N2O4	561322 Bechemicals 7204-0100	
6	C2H4O2N2	561324 Bechemicals 7204-0103	

9.25.5 Auto Close Tree Mode

To make working with large cluster trees a little easier you can activate the auto close mode which will close downstream clusters and make them more compact.

To do this:

- Right click in white space on the cluster and choose Preferences.
- Select "Auto Close Clusters Downstream".

9.25.6 Calculate Maximum Common Substructure for Nodes

To calculate the Maximum Common Substructure (MCS) for nodes in a tree:

- Chemistry/Cluster Set
- Select the option Calculate Maximum Common Substructure for Nodes
- Hover over a node to view the MCS.
- Right click on a node and choose "Extract Maximum Common Substructure" to extract it to a table.

The image shows two parts of a software interface. On the left is the 'Clustering Parameters' dialog box. A yellow callout box labeled '2. Select here' points to the 'Calculate Maximum Common Substructure for Nodes' checkbox, which is checked. On the right is a cluster tree. A green callout box labeled '3. Hover over node to view MCS' points to a node in the tree. A purple callout box labeled '4. Right click on a node to extract MCS to a table' points to the same node. A red callout box labeled '1. Chemistry/Cluster Set' is positioned above the dialog box. The tree shows a chemical structure for a node, with R1, R2, R3, R4, and R5 labels indicating substructures.

9.25.7 R Group Decomposition from Clustering Node

To perform R-group decomposition directly from a cluster node:

- Chemistry/Cluster Set
- Select the option Calculate Maximum Common Substructure for Nodes.
- Hover over a node to view the MCS.
- Right click on a node and choose "R-group decomposition for node".

The resulting table can be used for R-Group Enumeration or SAR.

1. Chemistry/Cluster Set and choose the calculate Maximum Common Structure option.

2. Right click on node

3. New chemical table

9.26 PCA Analysis

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

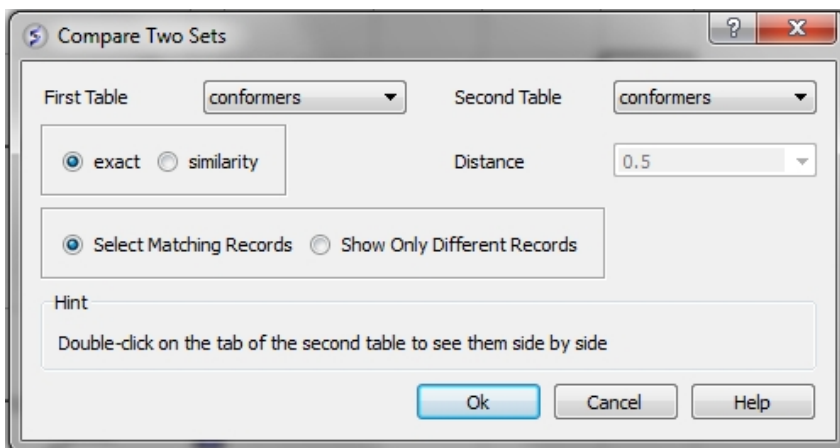
PCA analysis is described in the tables section [here](#).

9.27 Compare Two Sets

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To compare two chemical tables for similar compounds:

- Read the two tables into ICM.
- **Chemistry/Compare Two Sets...**
- Select the first table from the drop-down list and then select the second table.
- Choose whether you want to use **exact** or **similarity** comparison. If the **similarity** option is selected a **Distance** value needs to be entered. The distance value should be decreased below 0.5 (0.2 or 0.3 or even smaller if you want to find "strong" similarity differences. The distance relates to the Tanimoto distance as described [here](#)
- Choose whether you want to **Select Matching Records** or **Show Only Different Records**.
- Select OK
- Similar compounds will be highlighted blue (selected). A selection can be transferred to a new table by right-clicking on the table and select **Copy Selection to ICM Table**.

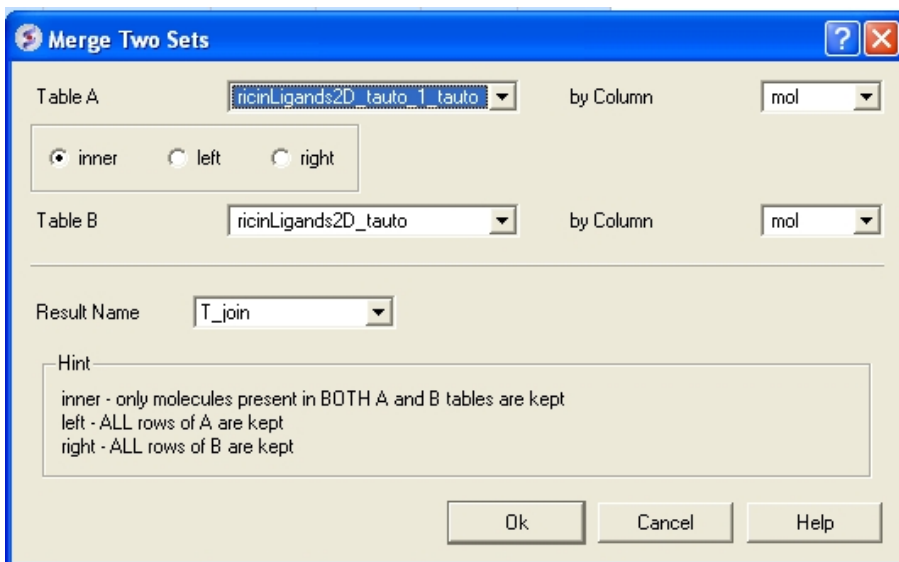


9.28 Merge Two Sets

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To merge two tables:

- Read the two tables into ICM.
- Chemistry/Merge Two Sets
- Select the first table from the drop down list (Table A) and the column you wish to use to merge the table by.
- Select merge method 1. **inner** - only molecules present in BOTH A and B tables are kept; or 2. **left** ALL rows of A are kept ; or 3. **right** ALL rows of B are kept.
- Select the second table from the drop down list (Table B) and the column you wish to use to merge the table by.
- Enter a name for the output table.
- Click OK and a new table will be displayed.



9.29 Select Duplicates

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

NOTE: Gui option is available in versions 3.5-1o and higher. The command line options for this function are described in the ICM Command Language manual at <http://www.molsoft.com/man/icm-functions.html#Index-chemical>

This option allows you to select and remove duplicate chemicals in a table.

- Read a chemical table into ICM.
- Chemistry/Select Duplicates
- Enter the table name you want to check for duplicates
- Enter whether you want chirality or the salts included in the analysis.
- Press OK
- Duplicate compounds will be **highlighted** blue in the table. You can delete them by right clicking on the row header and selecting **Delete Row(s)**



Right click and select Delete Row(s)

Duplicate rows are highlighted in blue

9.30 Combinatorial Chemistry

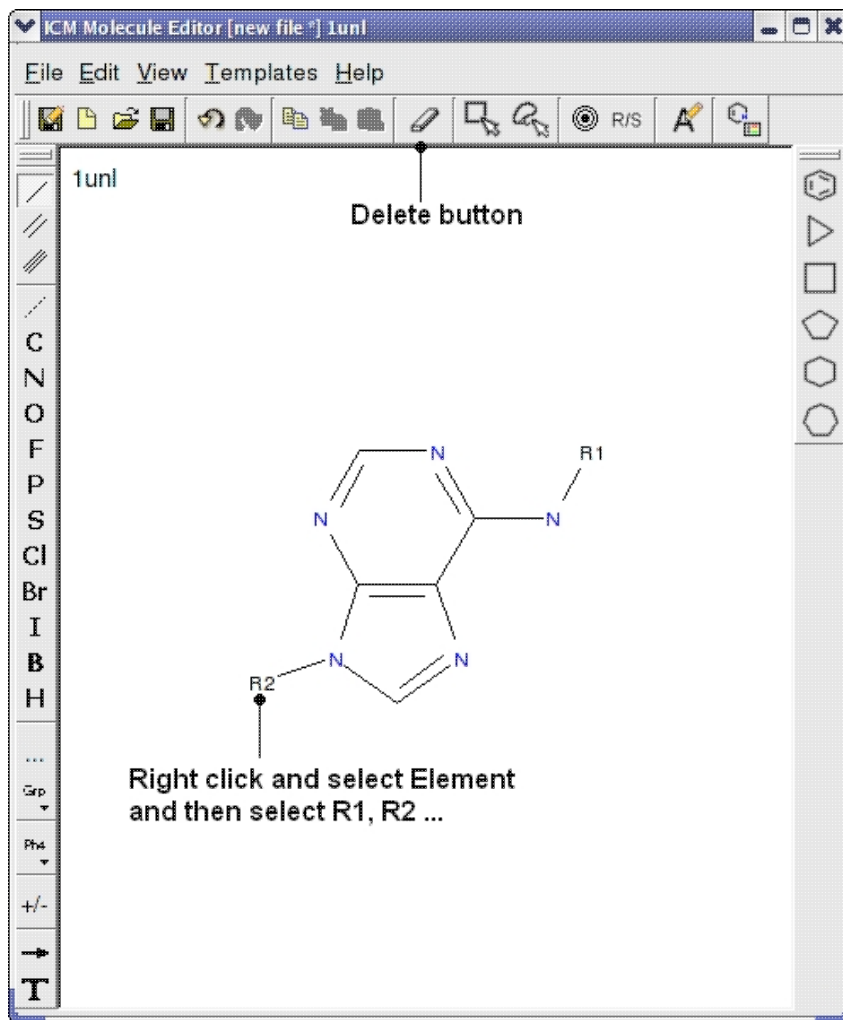
Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

9.30.1 Create/Modify Markush

Tutorial | Video

To create or modify a Markush Structure:

- Use the Molecular Editor to edit the scaffold as shown below.



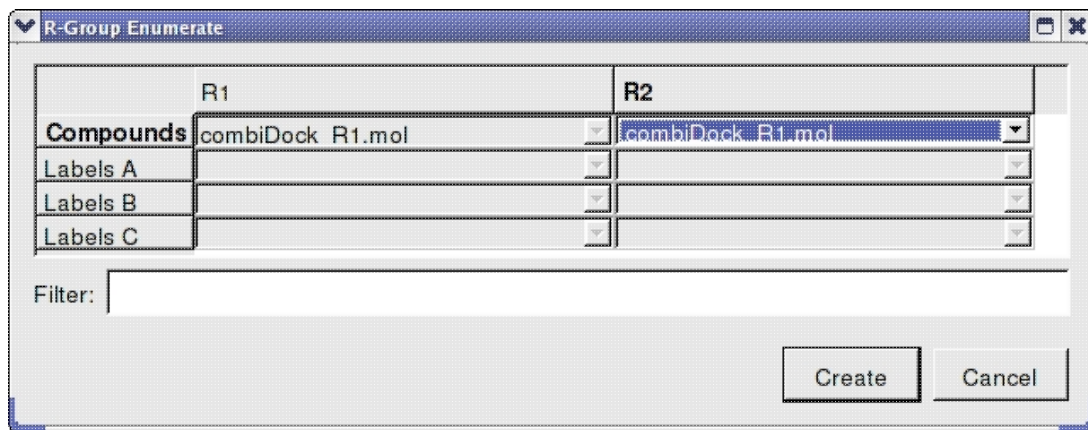
- Close the Molecular Editor window by clicking on the cross in the top right hand corner and the changes will be submitted to the table.
- The sketch in the chemical spreadsheet is named "chem" by default. For this example we will rename it "scaffold". You can rename it by right clicking on the table tab and selecting rename.

Step 4: Create Markush Combinatorial Library

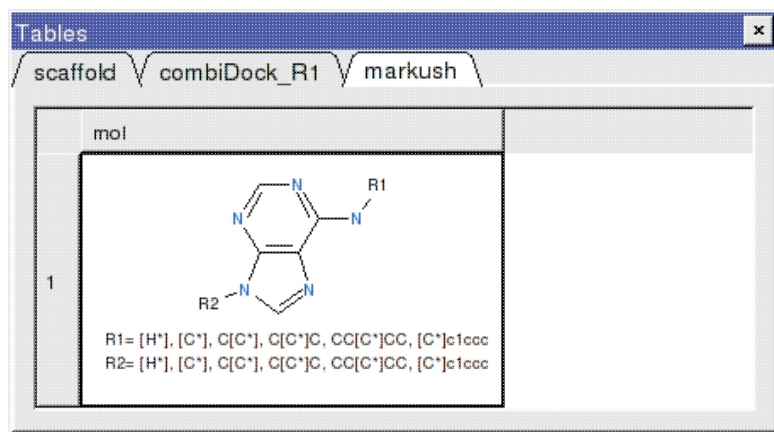
- Read in a table of substituents. For this example we will use an sdf file called combiDock_R1.sdf - this can be found in the ICM distribution (File/Open). If you cannot find this file please E mail support@molsoft.com and we will send it to you.
- Chemistry/Create Modify Markush and enter the data as shown below and press next.

The screenshot shows the "Create/Modify Markush" dialog box. It has three main input fields: "Scaffold Or Markush" with a dropdown menu showing "scaffold", "Index" with a numeric input field showing "1", and "Result Name" with a dropdown menu showing "markush". At the bottom right, there are two buttons: "Next" and "Cancel".

- Enter the name of the table containing substituents for R1 and R2. In this example we will use the same table **combiDock_R1** for R1 and R2 as shown below. You can use the drop down arrows to select the table you require.



- Once the tables are selected press **Create** and a new chemical table will be displayed with the markush structure annotated with the substituents for R1 and R2 as shown below.

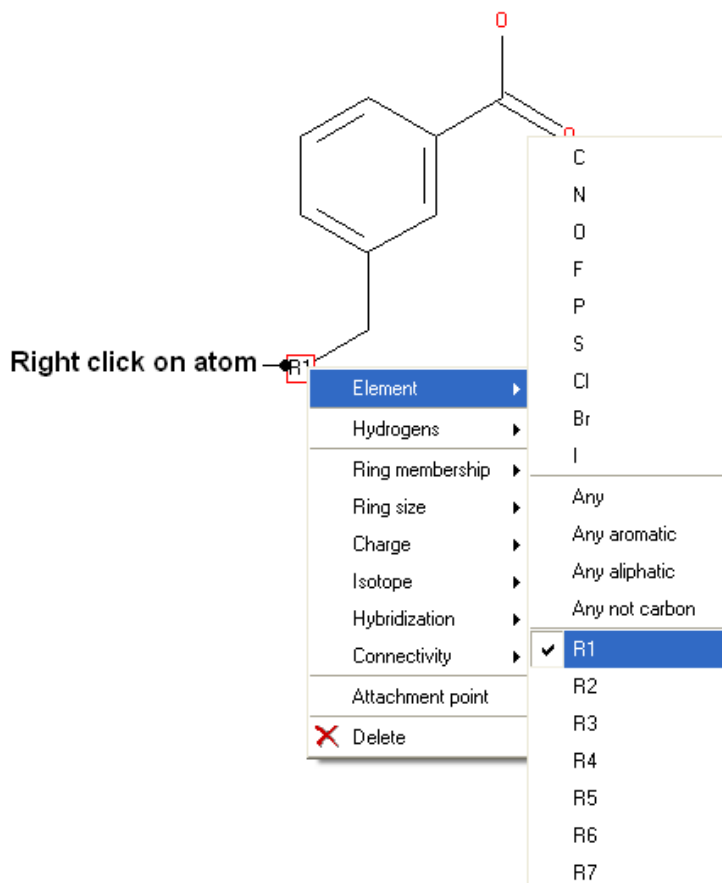


9.30.2 Enumerate by Scaffold

Tutorial | Video

To enumerate a library based on R-groups you first need to draw a sketch of the structure and display it in a chemical spreadsheet. To do this:

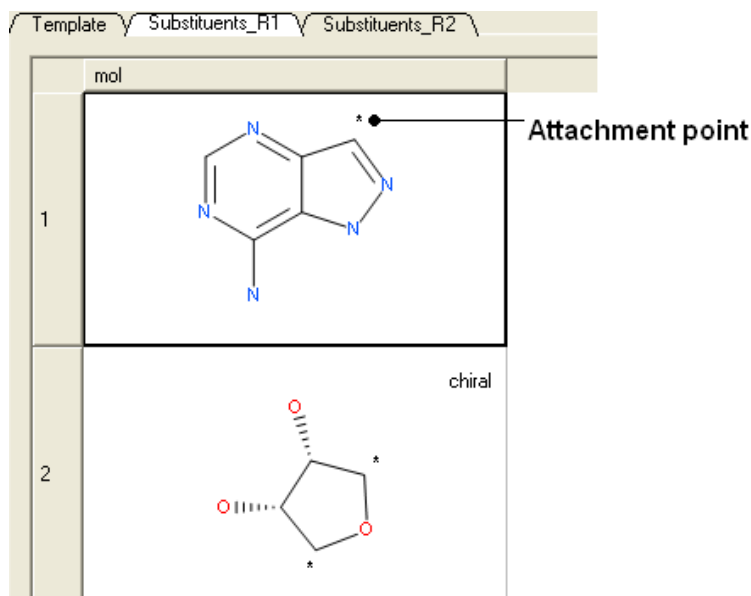
- Open up the ICM Molecular Editor.
- Draw the template structure with R-groups attached. Right click on an atom and select Element/R1, R2 ...



- In the Molecular Editor select File/Append to Table/New

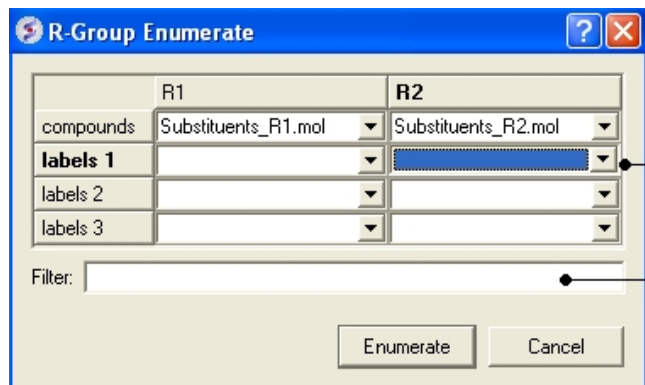
The next step is to read into ICM or construct a table of substituents. You can read in an SDF, mol, smiles file or extract fragments. If you do not want the first atom of the substituents to be the attachment point you need to define the attachment point. Attachment points are automatically assigned when you extract fragments or you can define them manually by:

- Right click on the substituent sketch and select **Edit Molecule**
- Right click on the atom and select **Attachment point**.



Next enumerate the library

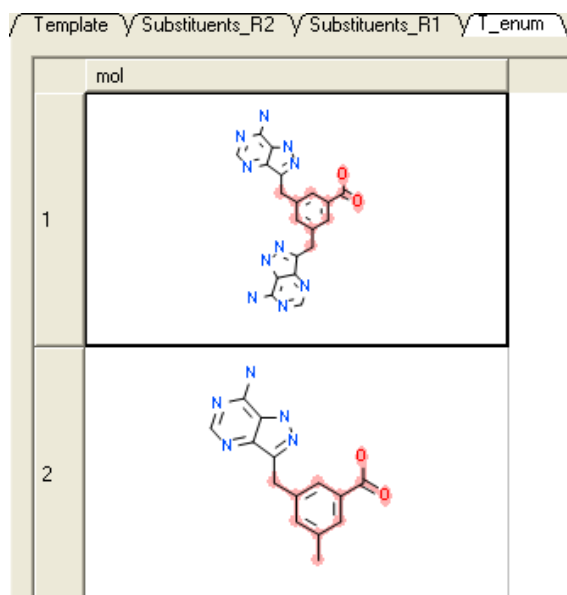
- Select the template structure (highlighted blue).
- Right click on the structure and select Chemistry/Enumerate R-groups or use the Chemistry menu and select Enumerate by Markush. If you use the menu option you will need to choose the table containing the scaffold from the drop down list of currently loaded tables. The index number refers to the row number in the scaffold table. In this example we only have one row containing the scaffold so the index number is 1.
- Select the R1, R2... table , labels and filters if necessary.



Bring over important labels into enumerated library eg compound ID number

Type condition here:
eg R1 != R2
MolWeight < 500

A new table will be produced called **T_enum** with the Template structure highlighted in red.

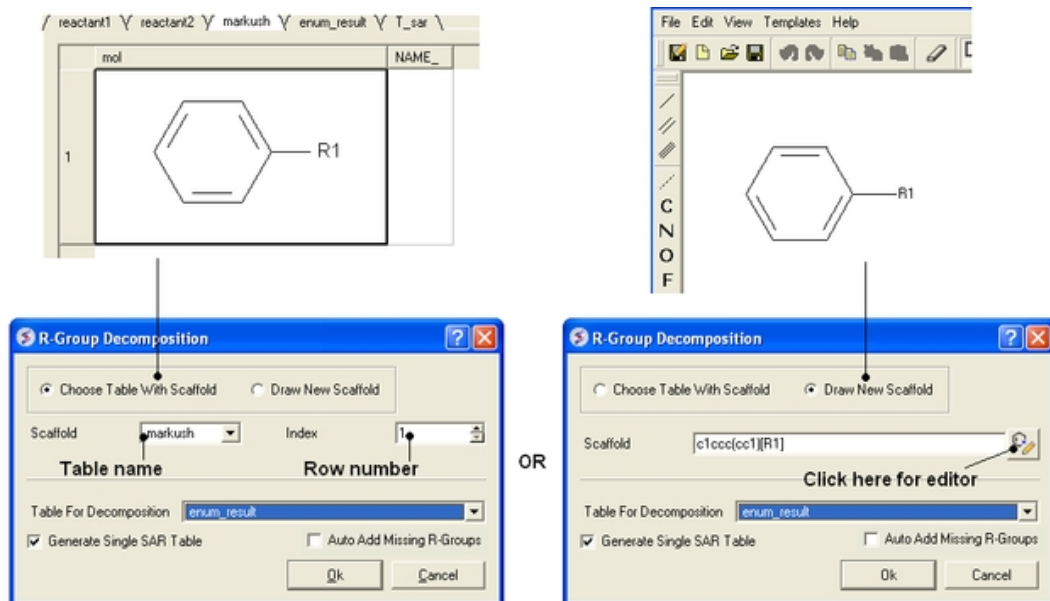


9.30.3 R-Group Decomposition

Tutorial Video

To decompose a library into fragments based on a Markush scaffold (opposite of R-group (Markush) enumeration):

- Read the sdf file you wish to decompose into ICM and it will be displayed as a molecular table.
- **Chemistry/Combinatorial Chemistry/R-Group Decomposition** and a window as shown below will be displayed.
- You now have two options on how to define the Markush scaffold. You can either 1). Draw it using the molecular editor and the smiles string will be added to the window shown below or 2). select a row of a preexisting table.



- Use the drop-down option to select the table you wish to decompose.
- If you have more than one R-group ICM can either generate a different table for each R-group or it can merge it into one single table whereby column will represent R1 and column two R2 This option is useful if you want to generate a SAR table with a column of activity data next to the R1 and R2 columns (see below).
- If you check the box "Auto Add Missing R Groups" then unique R-groups will be extracted from the scaffold where hydrogens can be attached.

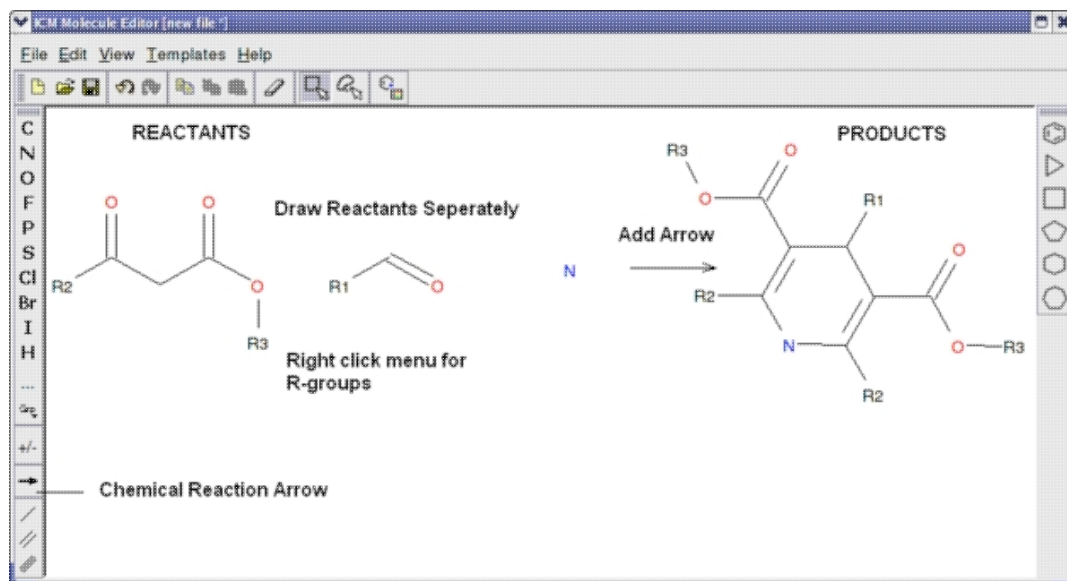
Single SAR table

	reactant1	reactant2	markush	enum_result	T_star
	R1	R2			Activity
67					racemic 78
68					77
69					77

9.30.4 Enumerate by Reaction

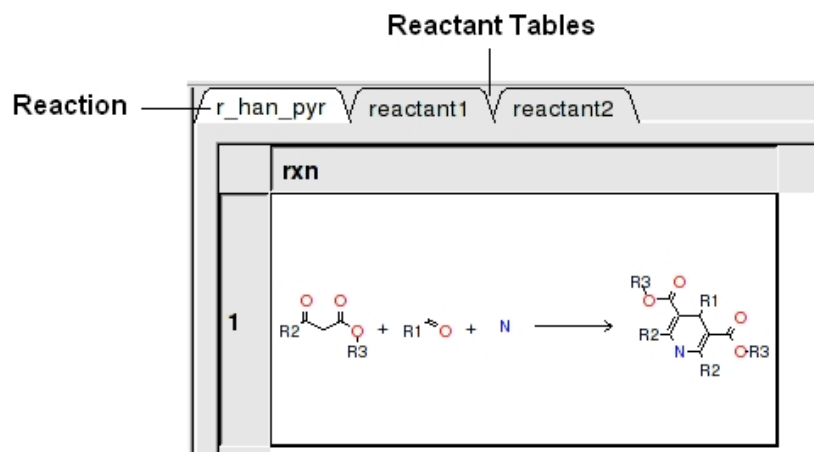
Additional Resources: [Tutorial](#) | [Video](#) | [Command Line Docs](#)

Reactions can be drawn using the ICM Molecular Editor. Reactants should be drawn side-by-side (no + sign is necessary) and separated from the product using the arrow. See example shown below:

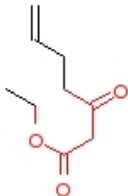
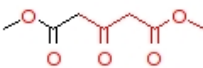

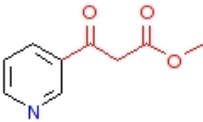
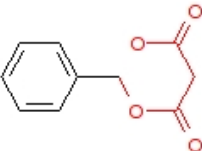


This example is available in the ICM distribution as `example_reaction1.icb`. The reaction is the Hantzsch Dihydropyridine (Pyridine) Synthesis. This reaction allows the preparation of dihydropyridine derivatives by condensation of an aldehyde with two equivalents of a α,β -ketoester in the presence of ammonia. Subsequent oxidation (or dehydrogenation) gives pyridine-3,5-dicarboxylates, which may also be decarboxylated to yield the corresponding pyridines.

In this example we have two reactants therefore it is necessary to have two reactant substructure tables loaded into ICM. ICM will match the substructure drawn in the reaction with the chemicals in the reactant table.



reactant 1 table:

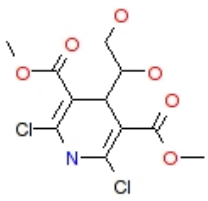
r_han_pyr	reactant1	reactant2	Substructure search: Found 88 hits of 'O=C([C;D2]C(=O)O[R3])[
	molid	MolW	mol	vendors
18	65236	170.094		asdi:500028335
19	87804	174.053		apolloscientific:12582 asdi:500014701
20	65360	182.058		apolloscientific:13718 keyorganics:11X-0925
21	87822	179.058		apolloscientific:2965 interchim:616
22	3341	194.058		asdi:500016383

To apply a reaction:

- Chemistry/Enumerate by Reaction.
- In this example (example_reaction1.icb) we already have the reaction drawn in a chemical table. Therefore select the **Choose Table With Reaction**. If you would like to draw a new reaction select **Draw New Reaction**.
- Enter the name of the table containing the reaction. If you have more than one reaction drawn you can select the row using the index option.
- Click OK and then you will be asked to enter the Reactants. Select the reactant tables from the drop down arrow for Reactant 1 and Reactant 2.
- You can transfer information to the reactant table by selecting columns in the **Labels** section.
- Unused reactants can be marked.
- Select what you want to do with multiple matches.

A table of Products will be then displayed in a table called T_out. Columns in T_out labeled "rct" display which reactants were used to build the product.

Table of Products

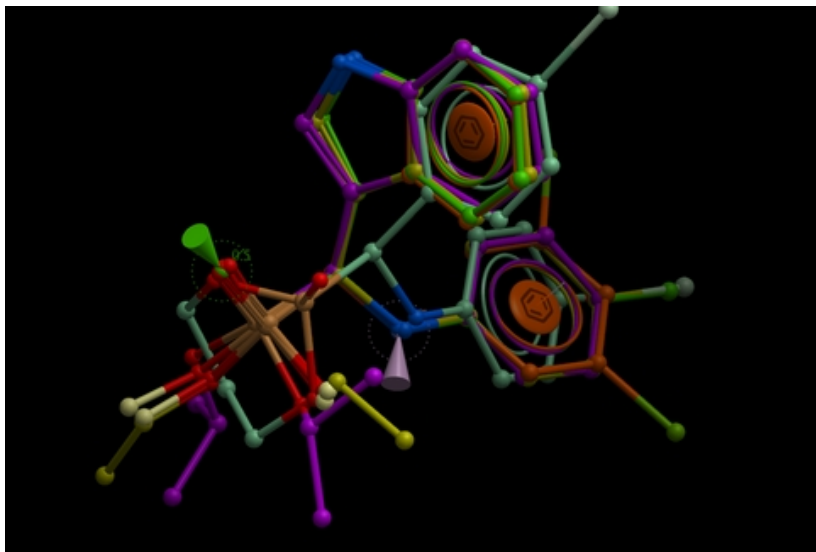
r_han_pyr	reactant1	reactant2	T_out
	mol	rct1	rct2
1		1	1
		Numbers indicate which reactant was used for the product	

9.31 Substructure Superposition

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

Chemical superposition can be undertaken in the following ways.

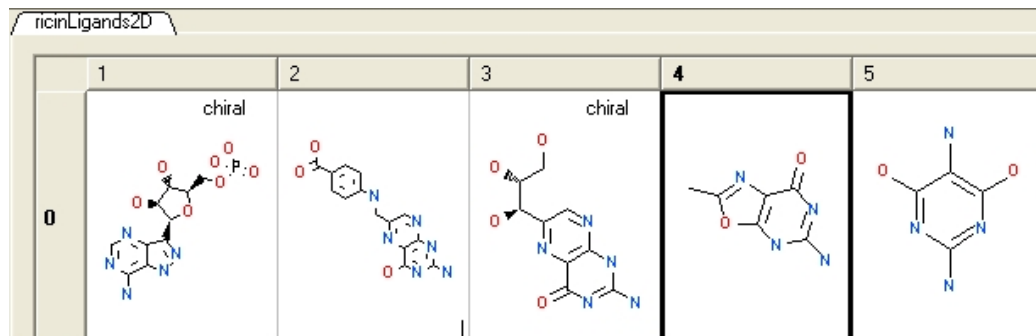
- Rigid Superposition of Compounds in a Table onto a Template in The Graphical Display
- Rigid Substructure Superposition of Chemicals in the Graphical Display
- Flexible Substructure Superposition
- See next section Atomic Property Field Superposition



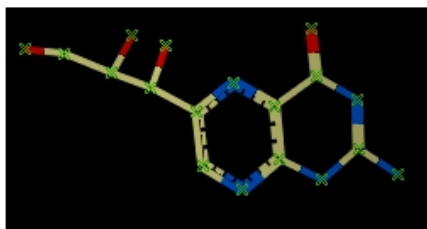
NOTE: The substructure superposition requires topologically exact match with the template. As long as there is a large consistent scaffold then the substructure superposition is the best approach. In cases where the structures differ topologically then the APF methods should be used to superimpose moieties which have similar properties.

9.31.1 Rigid Superposition of Compounds in a Table onto a Template in The Graphical Display

Here we describe how to superimpose chemicals from an ICM Molecular Table onto a 3D template displayed in the graphical display.

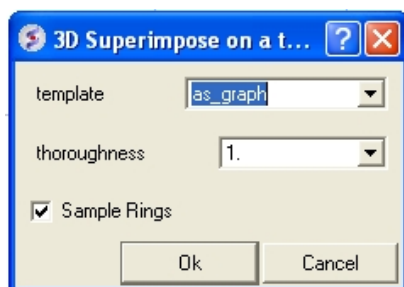


Superimpose from a molecular table onto a 3D template



Select chemical template in 3D display

- Load the template chemical into the 3D display. It needs to be a non-ICM object. To convert an ICM object into a non-icm object: right click on the name of the object in the ICM workspace and select convert to non-ICM. You should see that the object is now labeled as a model in the ICM workspace.
- Select the chemical template. One way to do this is to double click on the chemical name in the ICM Workspace (selected=blue in ICM Workspace and green crosses in graphical display).
- Select the chemical(s) (row(s)) in an ICM Molecular Table.
- Right click on the table and select Chemistry/Chemical Template Superposition and a table a data entry window as shown below will be displayed.



- Enter the name of the template or use as_graph if you selected the template as described above.
- The **effort** value (formerly 'thoroughness') represents the sampling length. The higher the value the longer the sampling takes.
- Select whether or not you wish the rings to be sampled.
- Click OK and the selected chemicals will be superimposed on the template in the chemical display.

9.31.2 Rigid Substructure Superposition

Here we describe how to perform a rigid-superposition of chemical structures in the graphical display:

- Select the chemicals you wish to superimpose. One way to do this is to double click on the chemical names in the ICM Workspace whilst holding down the control button (selected=blue in ICM Workspace and green crosses in graphical display) or hold the right mouse button and drag over the chemicals in the graphical display.
- Chemistry/Rigid Substructure Superimpose
- A window will be displayed. Enter the name of the template structure using the ICM selection language. The ICM selection language can be found by right clicking on the molecule in the ICM Workspace - first line of right click menu.
- Click OK

NOTE: Superimposed chemicals can be separated easily using the Arrange as grid option (Tools/Superimpose).

9.31.3 Flexible Substructure Superposition

Here we describe how to perform a flexible-superposition of chemical structures in the graphical display:

- Select the chemicals you wish to superimpose. One way to do this is to double click on the chemical names in the ICM Workspace whilst holding down the control button (selected=blue in ICM Workspace and green crosses in graphical display) or hold the right mouse button and drag over the chemicals in the graphical display.
- Chemistry/Flexible Substructure Superimpose
- A window will be displayed. Enter the name of the template structure using the ICM selection language. The ICM selection language can be found by right clicking on the molecule in the ICM Workspace - first line of right click menu.
- Click OK

NOTE: Superimposed chemicals can be separated easily using the Arrange as grid option (Tools/Superimpose).

9.32 Atomic Property Field Superposition

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

NOTE: The APF superposition method should be used when there is no common substructure between the chemicals that are being superimposed. If a common substructure is present then the substructure superposition methods described earlier should be used. The APF method will superimpose moieties that similar properties.

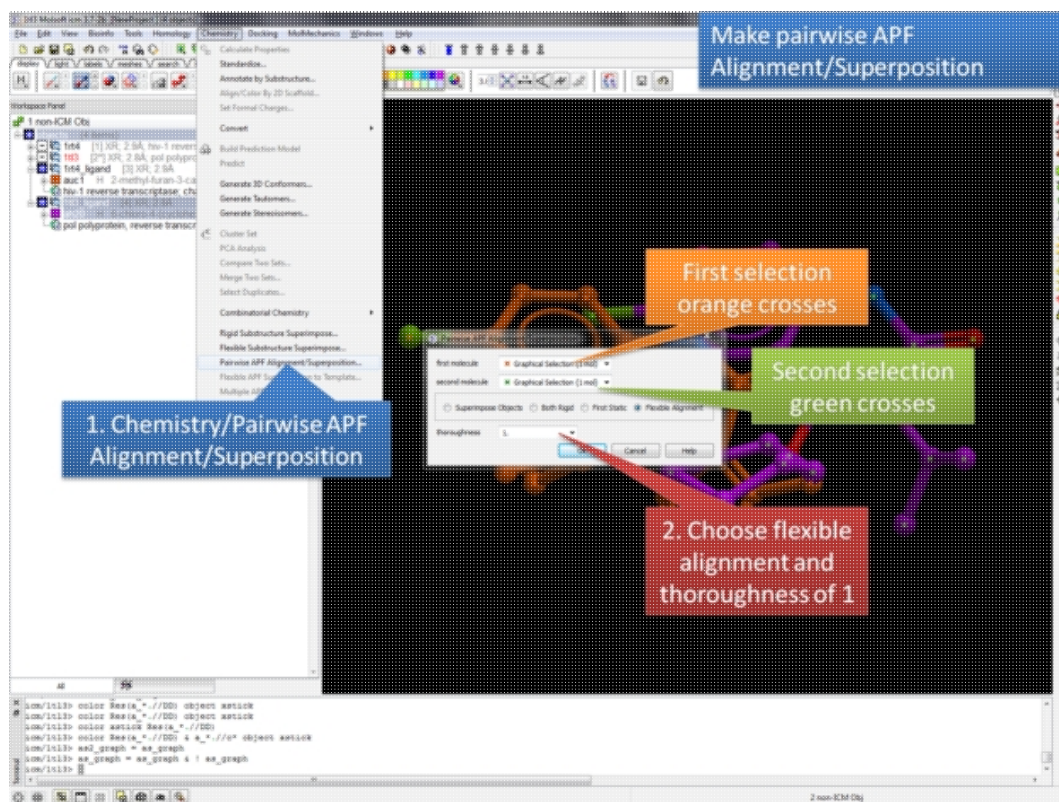
The Atomic Property Fields (APF) superposition/alignment method was first reported by Maxim Totrov PhD (Principal Scientist - MolSoft) at the 2007 233rd American Chemical Society National Meeting, Chicago, IL USA and then published [here](#).

APF is a 3D pharmacophoric potential implemented on a grid. APF can be generated from one or multiple ligands and seven properties are assigned from empiric physico-chemical components (hydrogen bond donors, acceptors, Sp² hybridization, lipophilicity, size, electropositive/negative and charge). Here we describe template APF superposition whereby the APF is generated from a single or multiple template and is then globally optimized with the internal force-field energy of the ligand.

9.32.1 Pairwise APF Alignment/Superposition

This option allows you to superimpose two chemicals by APF.

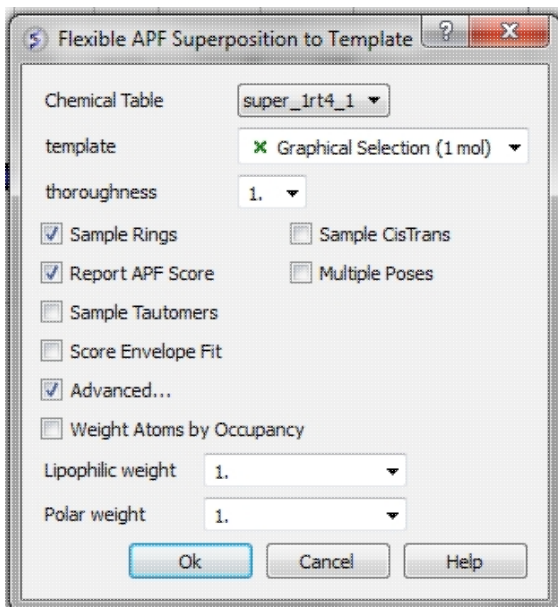
- Read in two 3D chemical file or convert a chemical sketch to 3D.
- Select the first chemical by double clicking on it in the ICM workspace. It should be highlighted in blue in the ICM workspace and green crosses in the graphical display. Next, copy this selection to orange (as2_graph)selection to save it - your selection should now have orange crosses.
- Select the second chemical by double clicking on it in the ICM workspace. It should be highlighted in blue in the ICM workspace and green crosses in the graphical display.
- Chemistry/Pairwise APF Superposition/Alignment
- Double check that the first molecule and second molecule selections are correct.
- Choose whether you wish to superimpose both chemical by APF, have both chemicals rigid, first static, or flexible alignment.
- The results will output a new object of the superimposed ligands in the ICM workspace.



9.32.2 Flexible APF Superposition to Template

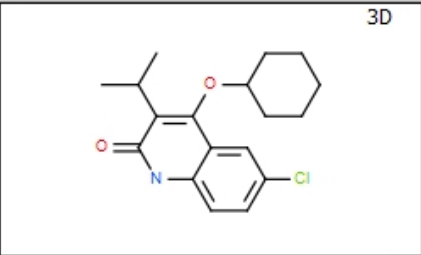
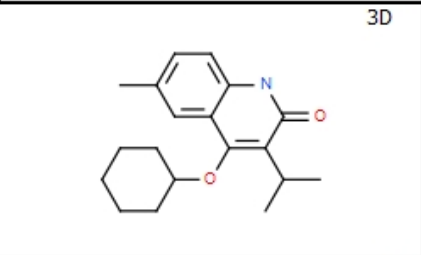
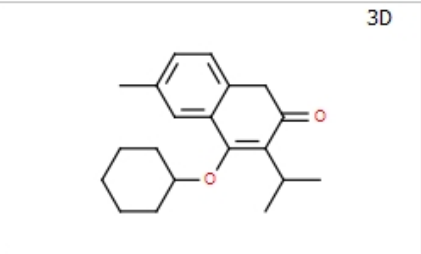
To perform Flexible APF Superposition to a template:

- Read a chemical table into ICM containing the compounds you wish to superimpose.
- Display in 3D the template structure you wish to superimpose on. See convert to 3D for instructions on how to generate a 3D template structure.
- Select **Chemistry/Flexible APF Superposition to Template** and a window as shown below will be displayed.



- Use the drop-down arrow to select the chemical table containing the chemicals you wish to superimpose on the template.
- * Select the 3D chemical **template** by double clicking on it in the ICM workspace. It should be highlighted in blue in the ICM workspace and green crosses in the graphical display.
- Enter an **effort** value. This represents how long the simulation will run for. A value of 1 has been validated as being a suitable length for this kind of superposition.
- Select whether you want flexible rings to be sampled by checking the appropriate box.
- Select whether you want cis and trans conformations of double bonds to be sampled by checking the appropriate box.
- Select whether you want the superposition to be scored in order to rank solutions by checking the appropriate box. About the APF SCORE - the lower the score the better the fit with the template. No one score or threshold can be stated as being a "good" score it depends on the chemical types being superimposed. One way to tell whether an APF score is good or bad is to compare it to the score obtained by self superimposing the template.
- Select whether you would like to display multiple poses of the superpositions.
- Select whether you would like to sample tautomers.
- Select the **Advanced** button if you want the superposition to be weighted by occupancy of the atoms by checking the appropriate box. It is often desirable to preferentially superimpose parts of a ligand while ignoring other regions. This can be achieved by setting the occupancy to zero for regions you are not focusing on. See: command line manual <http://www.molsoft.com/man/icm-commands.html#set-occupancy> . You can also define the weights for **lipophilic** and **polar** matches.
- Click OK and the simulation will run in the background. Once the superposition is complete the molecules will be displayed in the graphical display.

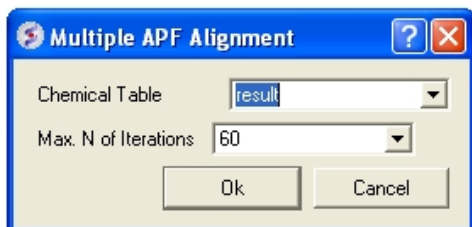
Click here to view the chemical superimposed in 3D

mol	L	Molecu Name	name	comment	Score
1 	<input type="checkbox"/>	1t13	ah20	6	-142.6
2 	<input type="checkbox"/>	1t13	ah20	6	-172
3 	<input type="checkbox"/>	1t13	ah20	6	-149

9.32.3 Multiple APF Alignment of Compounds in a Table

To superimpose multiple chemicals in a chemical table by the APF method:

- Read a chemical table into ICM containing the compounds you wish to superimpose.
- Select Chemistry/Multiple APF Alignment and a window as shown below will be displayed.
- Use the drop down arrow to select the chemical table.
- Select the number of iterations for the simulation. This represents how long the simulation will run for. A value of 60 has been validated as being a suitable length for this kind of superposition.
- Click OK and the simulation will run in the background. Once the superposition is complete the molecules will be displayed in the graphical display.



The 3D coordinates can be extracted to a chemical spreadsheet as described here.

9.33 APF Tools

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

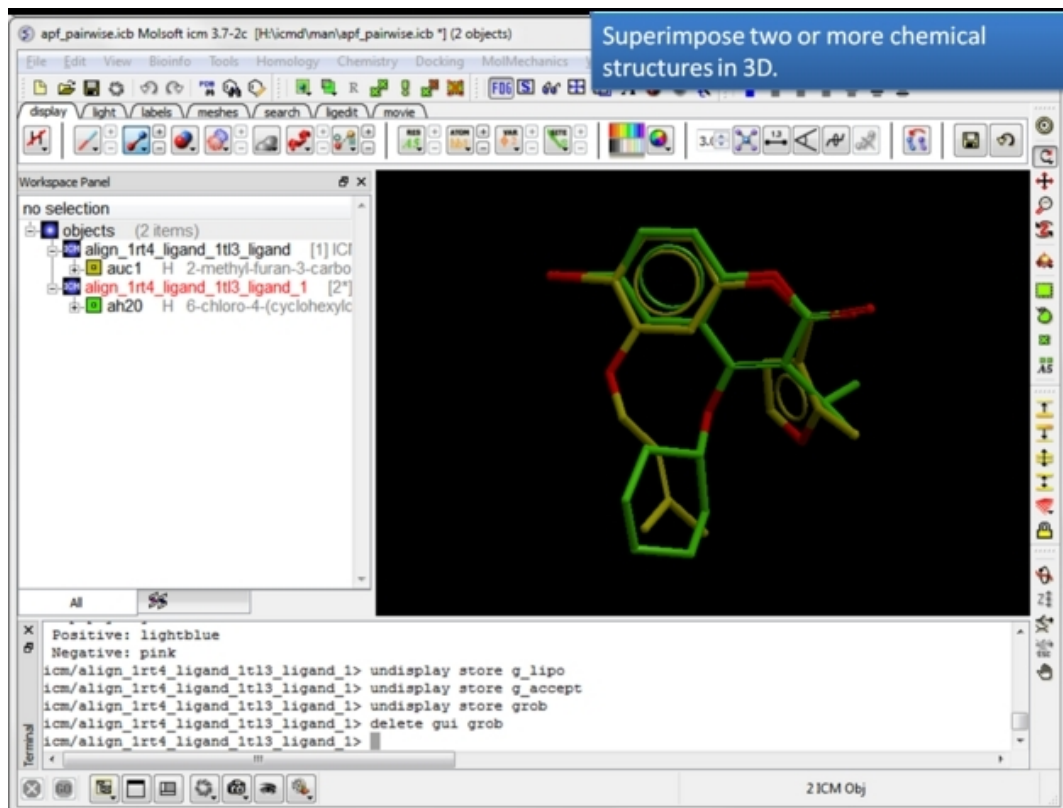
The Atomic Property Fields (APF) superposition/alignment method was first reported by Maxim Totrov PhD (Principal Scientist - MolSoft) at the 2007 233rd American Chemical Society National Meeting, Chicago, IL USA and then published [here](#). This section shows you how to:

- Generate a consensus pharmacophore from the APF fields.
- Calculate a pairwise APF score.
- Perform ligand-based screening using APF fields.

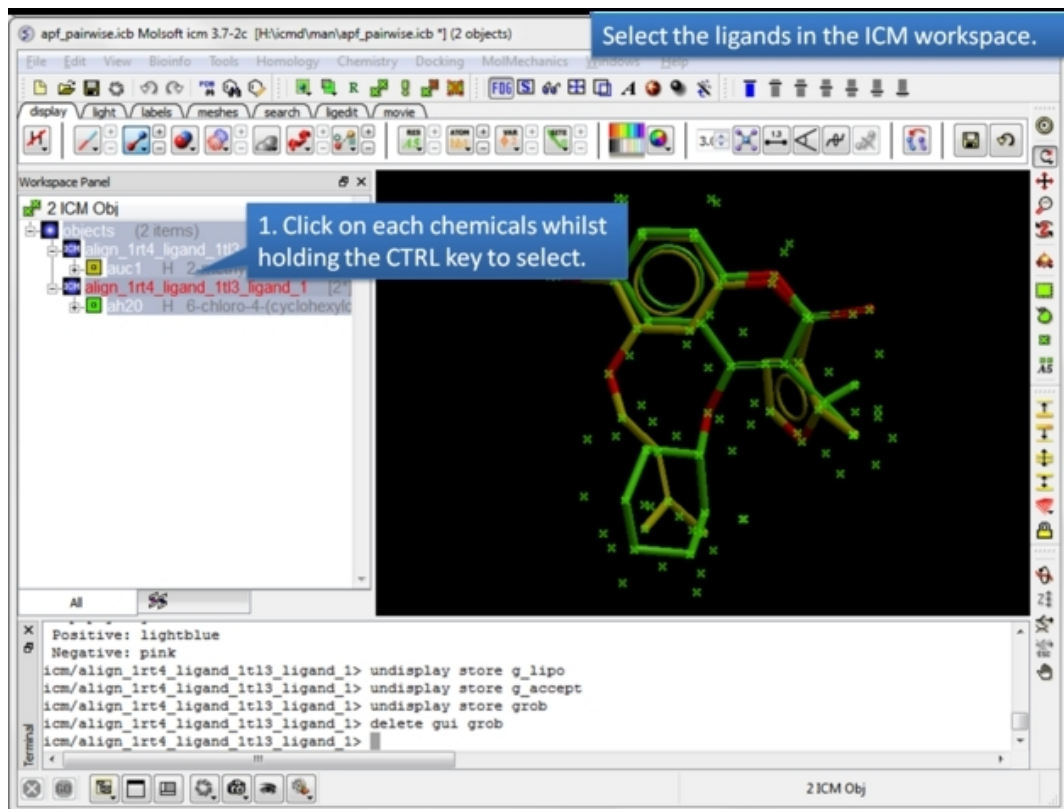
9.33.1 Consensus Pharmacophore

To generate a Consensus Pharmacophore based on APF fields use the Chemistry/APF Tools/Consensus Ph4 menu option.

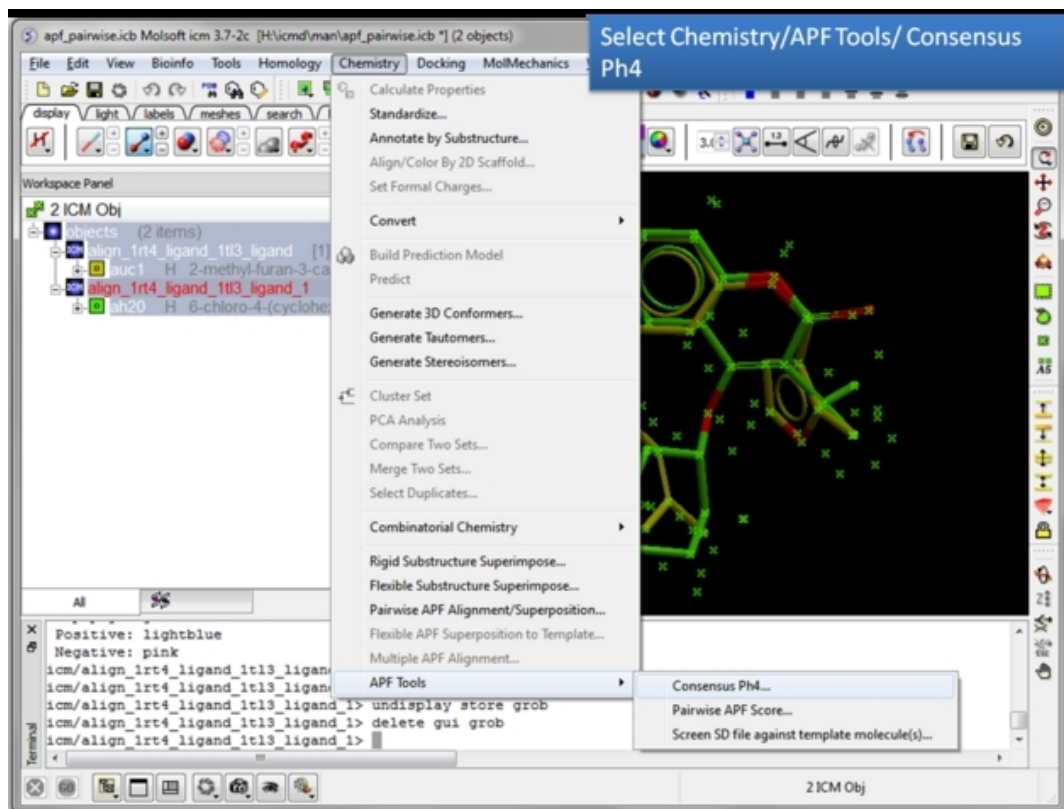
First convert your chemical into an ICM object.



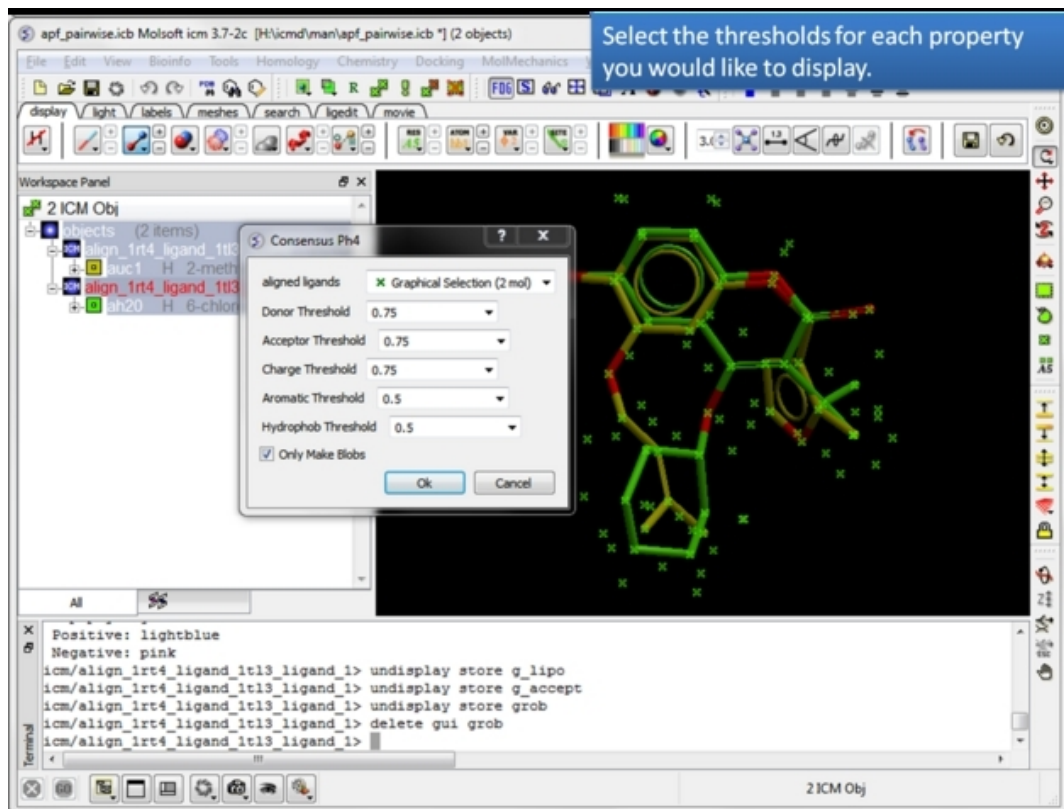
Load and display two or more superimposed chemicals. You can superimpose the chemicals using APF fields as described [here](#) or using substructure superposition.



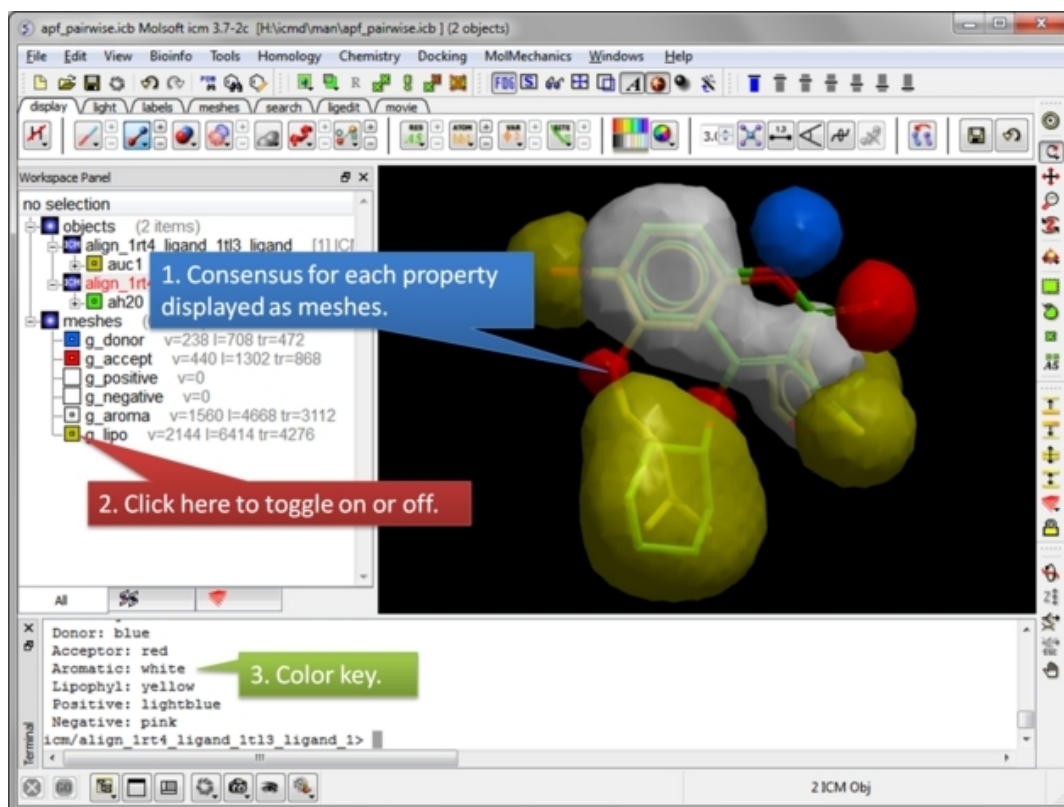
Select the ligands. You can select the ligands by clicking on them in the ICM workspace whilst holding down the CTRL key. Alternative selection modes are described [here](#).



Choose the Consensus Ph4 menu option. Chemistry/APF tools/Consensus Ph4.



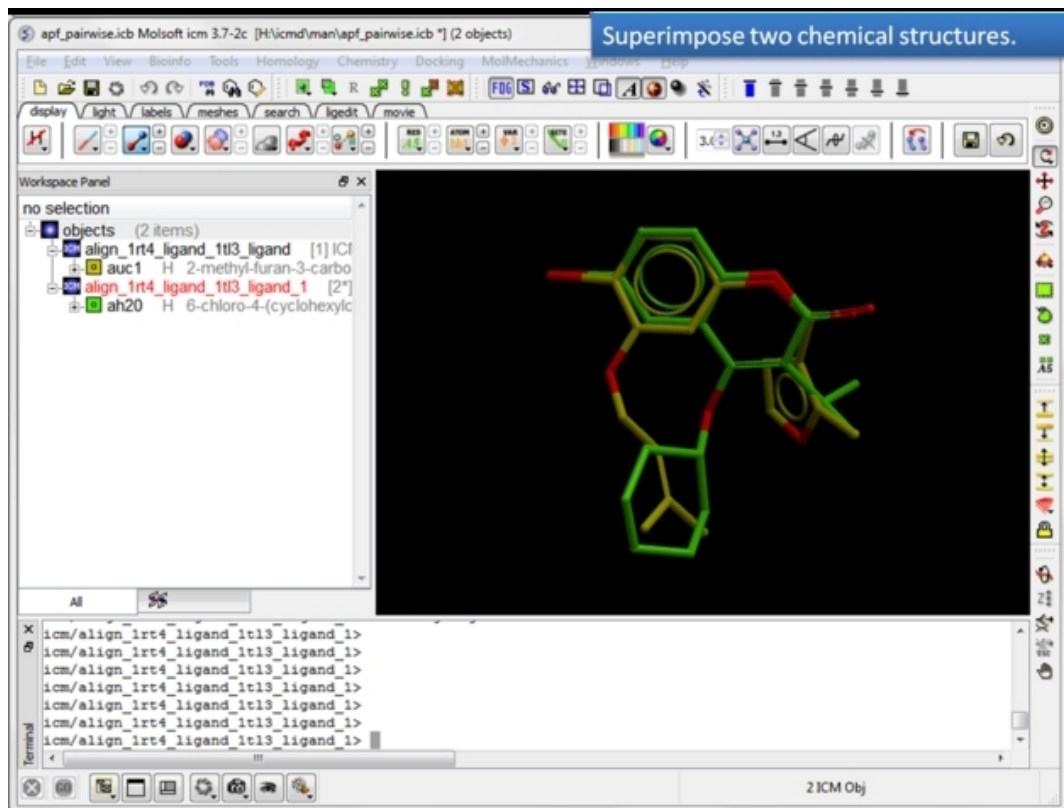
Select the threshold for each property. For example, if the setting is 0.75 then the pharmacophore will be displayed if the property is found in 75% or more of the ligands.



Consensus is displayed as meshes. Each mesh can be displayed or undisplay in the ICM workspace.

9.33.2 Pairwise APF Score

To determine how well the APF fields of two chemicals match you can determine the Pairwise APF Score using the Chemistry/APF Tools/Pairwise APF Score menu option.



Load and display two superimposed chemicals. You can superimpose the chemicals using APF fields as described here.

apf_pairwise.icb Molsoft icm 3.7-2c [H:\icmd\man\apf_pairwise.icb] (2 objects)

Select the first chemical structure.

2. Click here to copy green selection to orange.

1. Double click here to select.

Workspace Panel

- 1 ICM Obj
 - objects (2 items)
 - align_1rt4_ligand_1t13_ligand_1
 - auc1 H 2-methyl-furan-3-carbo
 - align_1rt4_ligand_1t13_ligand_1 [2*]
 - ah20 H 6-chloro-4-(cyclohexylc

Terminal

```
icm/align_1rt4_ligand_1t13_ligand_1>
icm/align_1rt4_ligand_1t13_ligand_1>
icm/align_1rt4_ligand_1t13_ligand_1>
icm/align_1rt4_ligand_1t13_ligand_1>
icm/align_1rt4_ligand_1t13_ligand_1> as2_graph = as_graph
icm/align_1rt4_ligand_1t13_ligand_1> as_graph = as_graph & ! as_graph
icm/align_1rt4_ligand_1t13_ligand_1>
```

2 ICM Obj

Select the first chemical. Use the green selection to orange button to select the first chemical.

apf_pairwise.icb Molsoft icm 3.7-2c [H:\icmd\man\apf_pairwise.icb] (2 objects)

Select the second chemical structure.

1. Double click here to select.

NOTE: Orange and green selections

Workspace Panel

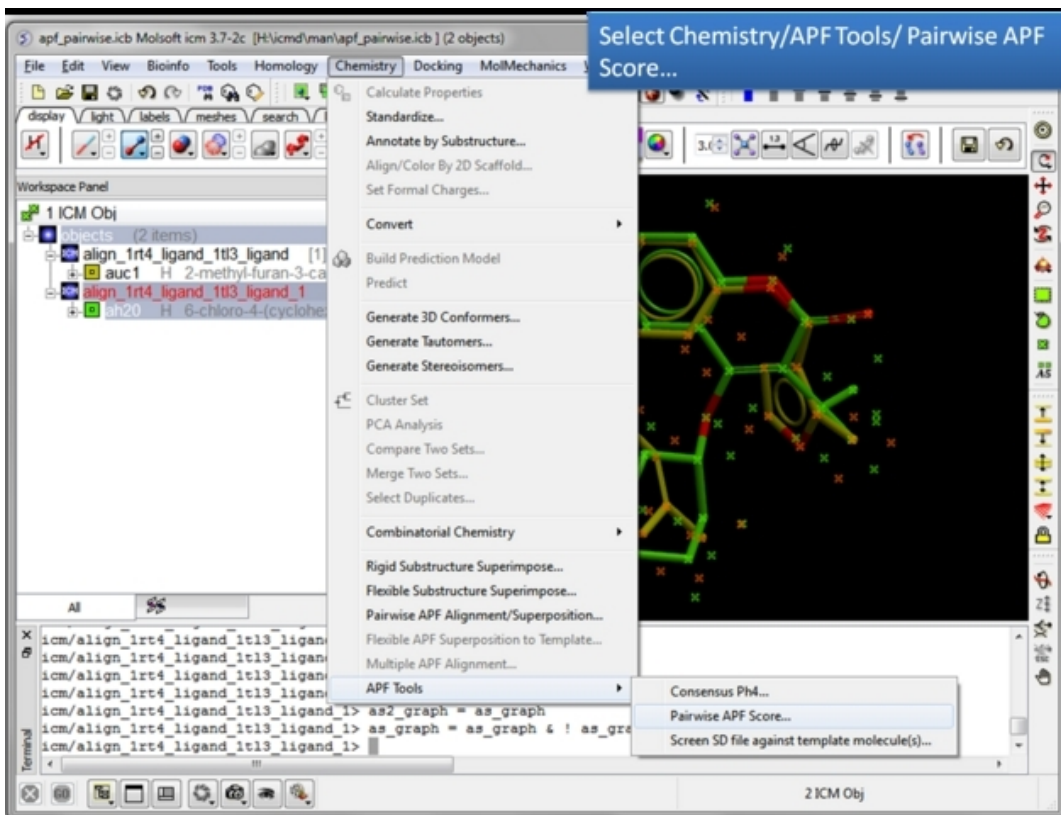
- 1 ICM Obj
 - objects (2 items)
 - align_1rt4_ligand_1t13_ligand_1
 - auc1 H 2-methyl-furan-3-carbo
 - align_1rt4_ligand_1t13_ligand_1 [2*]
 - ah20 H 6-chloro-4-(cyclohexylc

Terminal

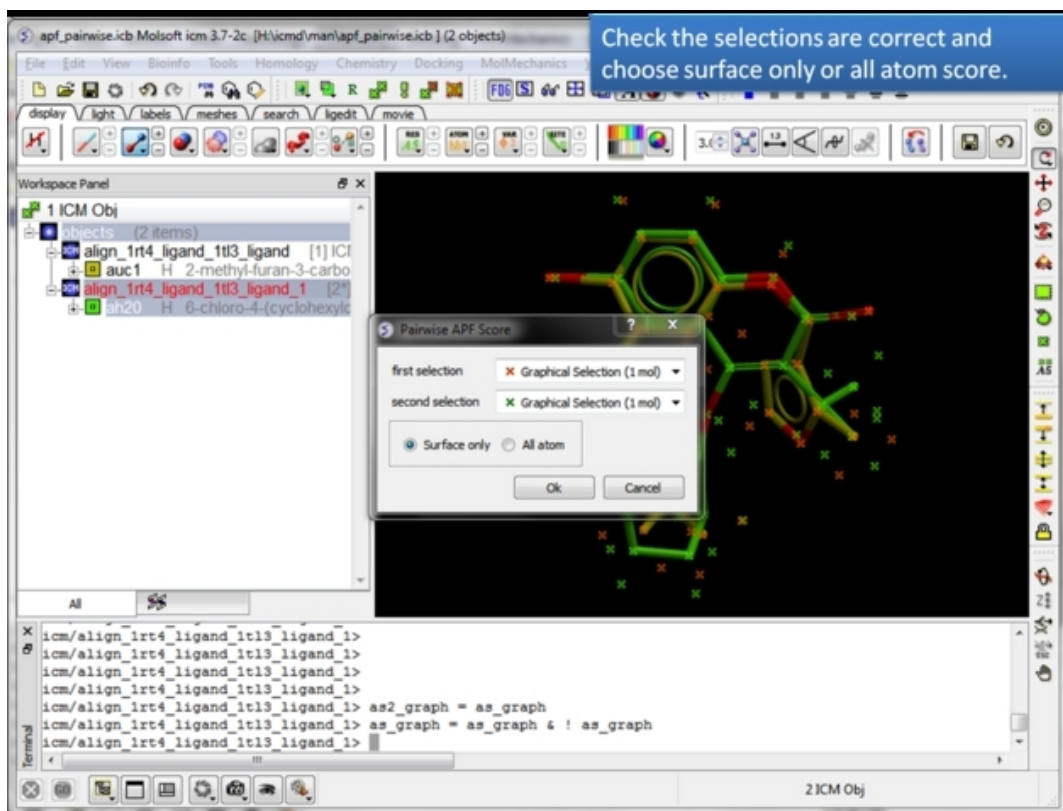
```
icm/align_1rt4_ligand_1t13_ligand_1>
icm/align_1rt4_ligand_1t13_ligand_1>
icm/align_1rt4_ligand_1t13_ligand_1>
icm/align_1rt4_ligand_1t13_ligand_1>
icm/align_1rt4_ligand_1t13_ligand_1> as2_graph = as_graph
icm/align_1rt4_ligand_1t13_ligand_1> as_graph = as_graph & ! as_graph
icm/align_1rt4_ligand_1t13_ligand_1>
```

2 ICM Obj

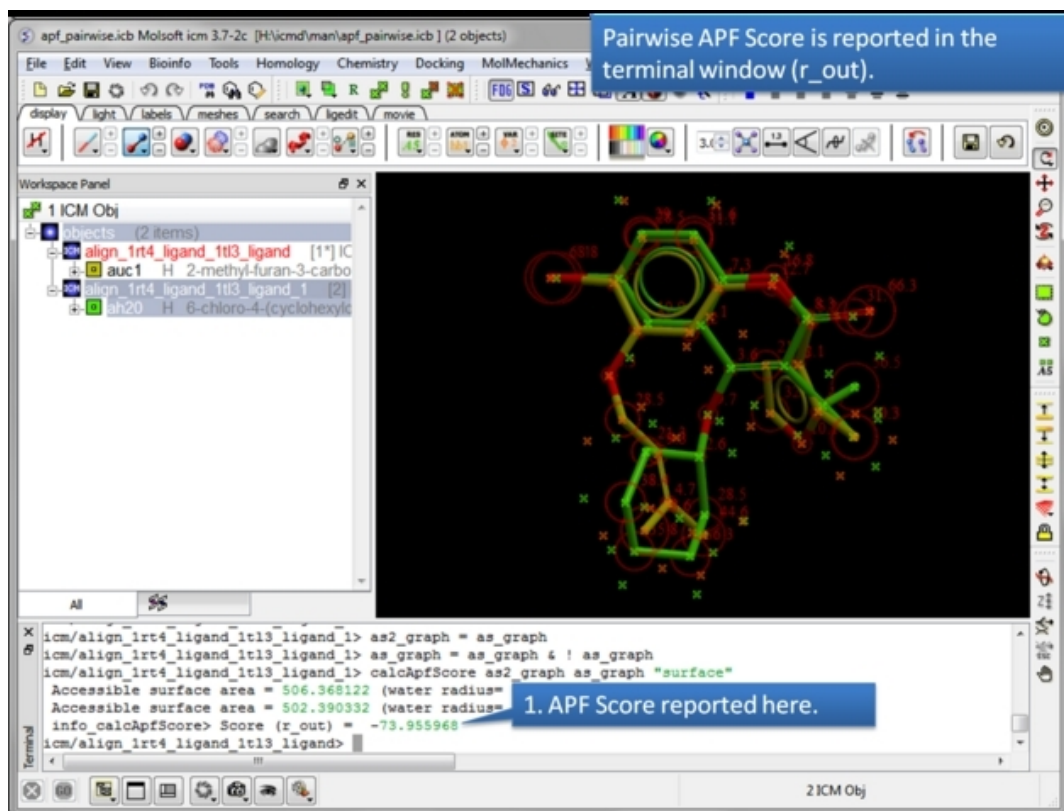
Select the second chemical. Double click the chemical in the ICM workspace to select it.



Choose the **Pairwise APF Score** menu option. Chemistry/APF tools/Pairwise APF Score



Check the selections are correct. Choose whether you wish to calculate the score based on all atoms or just the surface atoms.



The APF Pairwise Score is displayed in the terminal window. The score is also returned to r_out. Do not see the Terminal window? Go to Windows menu and check the Terminal box.

9.33.3 APF Ligand Based Screen - Screen SD file against template molecule(s)

This option allows you to perform a ligand-based screen using the Atomic Property Fields of superimposed chemicals. SD files are screened and the ligands are scored according to their fit into the APF field. The documentation for this option is in the Screening Chapter.

9.33.4 APF 3D QSAR

This method uses 3D pharmacophores Atomic Property Fields.

There are three steps to undertake APF 3D QSAR:

Step 1: Chemical Superposition

- The APF 3D QSAR method relies on a set of superimposed chemicals. You can superimpose chemicals using the tools [here](#). The 3D chemicals in the ICM workspace then need to be extracted to a chemical spreadsheet as described [here](#). Both the chemicals used for training and the chemicals that you wish to predict need to be superimposed.

Step 2: Learning

- Use the chemical table containing the superimposed chemicals (Step 1). This table should contain a mol column and a column containing activity data.
- Chemistry/APF Tools/APF 3D QSAR and select the learn tab.
- Choose the name of the chemical table from the drop down list.
- Choose the column containing the activity data.
- The statistics related to the model can be seen in the model section of the ICM workspace - left hand side.
- A table called **BASIS** will be created containing a column with the predicted value and cross-predicted value for each element of the training set.

Step 3: Prediction

- Read in a chemical table containing data you wish to predict. The column formats must match the model table and the chemicals must be superimposed on the training set (See Step 1).
- Chemistry/APF Tools/APF 3D QSAR/ and select the predict tab.
- Enter the name of the chemical table you wish to predict along with the activity column name.
- The table will be appended with a prediction column called **MolApfQsar**.

NOTE: The prediction model might be improved by removing some chemicals from the superposition in Step 1. It might be a subjective call as to which chemicals are good for the model and which ones should be removed. In some cases a crystal structure of a protein-ligand complex is available and the ligand pose would be a good `template` to superimpose onto.

10 Learn and Predict

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

There are four learning algorithms built into ICM

- Partial Least Squares (`plsRegression`)
- Principal Component Regression(`pcRegression`)
- Bayesian Classifier (`bayesianClassifier`)
- Random Forest (`randomForest`)

All 2D molecular property predictors are calculated using fragment-based contributions and 3D models use Atomic Property Fields.

For 2D fingerprints we developed an original method for splitting a molecule into a set of linear or non-linear fragments of different length and representation levels and then each chemical pattern found is converted into a descriptor.

In order to perform 'learn and predict' in ICM information must be stored in a table, molecular table or csv file. See the `tables` chapter for more information on ICM tables. Both chemical compounds and numeric data can be source for building prediction models.

In this chapter:

- Model building using the `Learn` method.
- Use a model to make a `Prediction`
- Fingerprint Methods
- Learning Theory

10.1 Learn

To Learn from a data set and make a model:

First load in a table of data on which you wish to perform the learn and predict functions. See the `tables` chapter for more information on ICM tables.

- Read in your data e.g. chemical spreadsheet (e.g. File/Open SDF or .csv).
- Select Tools/Table/Build Prediction Model (Learn) and a window or use the Chemistry/Build Prediction Model option.
- Enter the name of table with which you want to perform the predictions. You may locate your table from the drop down arrow menu.
- Select the **Column** from which you wish to learn. Use the drop down arrow to select. **NOTE** If the table does not contain any numeric (integer or real) columns, there is nothing to predict, so the "Build Prediction Model - Learn" button will be disabled.
- Enter a name for the learn model.
- Select which learning method you wish to use from the drop down menu (PLS, PCRegression, Bayesian Classifier, Random Forest).
- Select whether you would like ICM to estimate the number of latent variables or you can specify a range.
- Select **No Free Term Constraint:f(0)=0** if you would the PLS method to find a constant and add it to the model.

Descriptors




- Select which columns (descriptors) of your table you wish to use to 'learn'. You can select just the **mol** column or all numerical columns.

Chemical Fingerprints

- If you are using chemical descriptors select the `fingerprint` method want to use **Linear** or **Extended Connectivity Fingerprints (ECFP)**
- Select **Binary** if you want to use Binary fingerprints - do not check if you want to use counted.

- **Minimal chain length** the minimal length of the chain of atoms, enumerated in each compound, usually it is 1 which means that you will consider an atomic composition numbers as descriptors. The typing of atoms can be customized with the map argument. .
- **Maximal chain length** the maximal length of the chain atoms and bonds enumerated in each compound Usually it is 3 or 4. Larger values of this parameter will lead to a large number of possible combinations. To overcome that, either the typing needs to be simple (e.g. only sp² and sp³ property, regardless of the atom number), or the data set needs to be really large.
- **Length** - this is the total number of bins in a final fingerprint. Typically the size (either explicit or automatically estimated) is in a hundreds to thousands range. We recommend to check the **auto** option.

Customizing the fingerprint to fit the predicted property

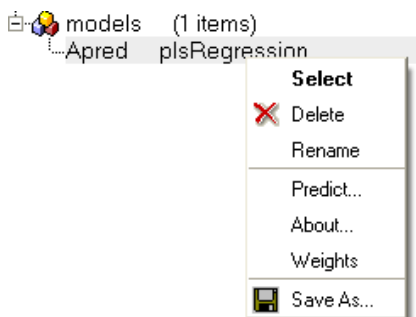
	lev	Atom properties	Bond properties	
1	1	element number,nHydrogens		
2	2	element number,nHydrogens	bond order	
3	3	element number,nHydrogens	bond order	

- In this dialog you customize the atom properties for every chain length or ecfp fragment size. Click on the "pencil button" to add additional properties to a chain or fragment size (lev). Once the model is made you will see the weights of these properties in the model (right click on model in the ICM workspace and choose Weights).

Test

- Select the number of cross-validation groups you wish to use or selected rows can be used for cross validation. The number of iterations will impact the speed of the calculation. 5 is the default number of groups but 2 would be the least rigorous and selecting the 'Leave-1-out' would be the most rigorous calculation.
- **Bootstrapping** generates statistics (R², RMSD) for a random prediction and you should aim for there to be a significant difference between the random R² and the model R².
- Click on the learn button and a model will be created and placed in the ICM Workspace panel (left hand side).

All models are then stored in the ICM workspace as shown below. A number of options are displayed in the right click menu.



10.1.1 About the Model

- To see more details about your model (e.g. R², RMSE etc....)
- Right click on your model in the ICM workspace and choose **About**.

10.1.2 Model Weights

The **Weights** dialog can be helpful to see what is driving the correlation in your model or for troubleshooting. The key column is "**w**" which represents the weight of the fragment in the regression. You may see a particular fragment providing a high weight to the model so you may want to **add** similar fragments to your training set to see if it will improve the model. Removing fragments from the training set based on their weight is not recommended because each fragment is part of a multiple component and it is difficult to know precisely its importance.

- Right click on the model in the ICM Workspace and choose **Weights**
- A table will be displayed containing the following columns:

BITS = the number of the chain or fragment.

name = smiles string of fragment. You can copy and paste this into the molecular editor to view it in 2D.

w = is actual weight of the fragment in the regression (used directly to get the result value).

mean = indicates the mean occurrence of the fragment in the model . For example, a value of 6 means the fragment is used 6 times in the regression.

rmsd = a high value indicates high occurrence of the fragment and influences the relative weight parameter.

wRel = relative 'importance' of the fragment based on the RMSD, mean and CorrY values.

10.1.3 Improve Model

The model can be improved by:

- Right click on the model in the ICM Workspace and choose **Improve**
- A dialog box similar to the one you used to make the model in `Learn` will be displayed.

10.2 Predict

To make a prediction using a `created` model.

Read the table of data into ICM from which you wish to predict. Make sure the table contains the same columns used for the `learn` model.

- Tools/Table/Predict or Chemistry/Predict
- Select which table you wish to make the prediction on.
- Select which model you wish to use.
- Check that the required columns are in the table. If they are absent a red mark will appear against the column that is missing.
- Click Predict.

10.3 Fingerprint Methods

MolSoft's Fingerprints MolSoft has developed a proprietary method for generating hashed chemical fingerprints used in chemical search, similarity calculations and QSAR. ICM fingerprints do not use any pre-defined set of fragments. Instead, ICM enumerates all possible linear and some non-linear fragments (circular and triangular) up to a certain size and hash them into a result vector. Equal fragments always get the same position in the vector and fingerprints can be either binary (presence or absence of a fragment) or counted (number of fragment occurrences is counted). The method includes 2D pharmacophore fingerprints whereby Euclidian 3D distance is replaced by number of bonds between atoms to make it applicable to the 2D structures directly.

Additional information on Chemical Fingerprint properties can be read here:

<http://www.molsoft.com/man/fingerprints.html>

Extended Connectivity FingerPrints (ECFP) We also support Extended Connectivity FingerPrints (ECFP).

10.4 A Little Theory on Learning

For a more detailed explanation of the theory behind Partial Least Squares (PLS) we suggest you read Geladi et al *Analytica Chimica Acta* (1986) 1-17.

PLS (Partial Least Squares) Regression PLS regression algorithm builds linear prediction model: in format $y = (w, x) + b$, where b is the **bias** - a real number, and w is the weights vector, which is scalarly multiplied by the data vector x . PLS uses the given learning y values very actively which allows it to

produce fairly good models with respect to constraint of being linear. Although linear regression models have an advantage of weights for each descriptor which gives a useful information and allows feature selection in many cases.

The linear model simply is not able to predict higher order dependencies.

There are different ways to deal with it. By adding the second order columns into the descriptor set you can let PLS predict them. Actually if you have a lot of columns derived from basic data, the linear model built will be able to make a high-quality linear approximations of the actual functions. ICM has a powerful tool for automatical generation of such descriptors based on compound data -- molecule fingerprints generation algorithm. It generates hundreds of columns based on initial data. The withdraw is that analysing the weights given by PLS to generated descriptors is almost senseless. You will need a *mol* column in your table to use this feature.

ICM has built-in models for prediction of several significant molecule properties, like logP, logS, PSA based on fingerprints+PLS symbiosis, which have proven their quality.

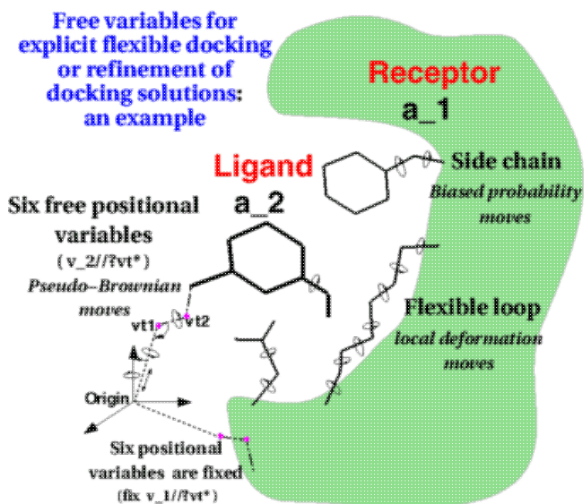
PC (Principal Component) Regression

PCR also builds linear model in its simplest form, as PLS does, though it sets other weights to descriptors, and built models are usually worse in sense of predicting, because PCR uses *value* information of the *training data* only in secondary way. We recommend you to use PCR, when you want to build an ordinary regression (MLR - Multiple Linear Regression) model by using only some number of first principal components of X data matrix (ordered by decreasing eigenvalues) or even building the full MLR model (by setting the number of PCs to value higher than the number of rows in matrix).

11 Docking

Chapter Contents

- Small Molecule Docking
- Flexible Receptor Docking
- Template Docking
- Covalent Docking
- Automated Model Building into Density
- Protein-Protein Docking



11.1 Small Molecule Docking

Available in the following product(s): ICM-Pro | ICM-VLS |

Additional Resources: Tutorial: Dock Biotin to Streptavidin Receptor | Tutorial: Dock a Ricin Inhibitor | Command Line Docs | Interactive Docking Tutorial Video | Re-dock Tutorial Video

11.1.1 Docking Introduction

This section is concerned with predictions of interactions of drugs or small biological substrates (less than about 600-700 Da) to pockets of larger, more rigid, receptors (typically, protein molecules, DNA or RNA).

For accurate ligand docking, the goal is to have an adequate three-dimensional model of the receptor pocket you are planning to dock ligands to. If this is the case then ICM docking has been shown to be very accurate in a number of independent assessments.

However, there are a number of pitfalls which need to be overcome to achieve accurate ligand docking. The pitfalls are that your model is not accurate overall, does not reflect the induced fit, or alternative conformations of the receptor binding pocket are missed.

Some key points about ICM Ligand Docking:

- An average docking time is 0.5 seconds to 30 seconds per ligand per processor. The time per ligand was chosen to be the smallest possible to allow screening of very large data sets. To increase the time spent per ligand, change the Docking_thoroughness parameter.
- ICM docking is probably the most accurate predictive tool of the binding geometry today. ICM docking has consistently ranked first place compared to other leading docking software in terms of accuracy. ICM has been successful in many drug design applications by scientists in academia and industry.

- ICM ligand docking procedure performs docking of the fully flexible small-molecule ligand to a known receptor 3D structure.
- The goal of the flexible docking calculation is prediction of correct binding geometry for each binder.
- ICM stochastic global optimization algorithm attempts to find the global minimum of the energy function that includes five grid potentials describing interaction of the flexible ligand with the receptor and internal conformational energy of the ligand.
- During the docking process a stack of alternative low energy conformations is saved.

Please see this link <http://www.molsoft.com/docking.html> to read more about ICM Docking.

11.1.1.1 Receptor Considerations

If you have only a single PDB entry for your receptor, convert the protein to an ICM object, delete water molecules and irrelevant chains. However, if you have a choice between several templates, take the following into account:

- X-ray structure is preferable to an NMR structure.
- High resolution X-ray structure (less than 2.1Å) is much better than, say 2.5Å .
- Watch out for high-B-factor regions and avoid them; sometimes crystallographers deposit fantasy coordinates with high-B-factors.
- Check the pocket for correct placement of polar hydrogens and choose correct form of histidine.
- A bound conformation of the receptor is preferable, however if you use an apo-model, an NMR structure or a model by homology, the side-chains in a pocket may be incorrect. Frequently they stick out and prevent a ligand from binding. Those stubborn side-chains can be 'tamed', (i) manually; (ii) by a side chain simulation with elevated surfaceTension; or (iii) using a method such as Dual Alanine Scanning and Refinement (SCARE).
- A model by homology can be built with the build model command (see molecular modeling section of this manual) and used for docking.

11.1.1.2 Ligand Considerations

Usually a good place to start is to try to dock the known ligand(s) to the receptor model. You may also want to dock a library of compounds in order to identify lead candidates. In this case the main pitfalls are that the library is too restricted, molecules are not chemically feasible or not drug-like. For peptide docking please use the `protein-protein docking` protocol.

There is no need to convert the ligands to 3D, this is done "on-the-fly" during the docking process.

NOTE: If you are docking a ligand directly from the PDB please check the bond types and formal charges of the ligand. This is discussed in the section entitled `Converting a Chemical from the PDB`

11.1.1.3 Chirality Considerations

To sample linear chiral centers during docking you can check the option "Sample Racemic Centers" in `Docking/Preferences/General`. To sample chiral centers in rings it is best to generate all 2D enantiomers using `Chemistry/Generate Stereoisomers`.

11.1.1.4 Protonation Consideration

The protonation state can be set in the `Docking/Preferences/General` dialog box. Choose the drop down option in the field labeled **Charge Groups**. This option charges some ionizable basic groups, it currently accepts the following values: space-separated NH₂, NH, NT for primary secondary and tertiary aliphatic amines. There are also options for imidazole and amidine. An alternative is **auto**, which uses built-in prediction of K_a/K_b to charge and protonate/deprotonate appropriate groups. The **none** option still charges acids unless `neutralAcids` flag is set to **no** in the docking table file `.dtb` or `.tab`.

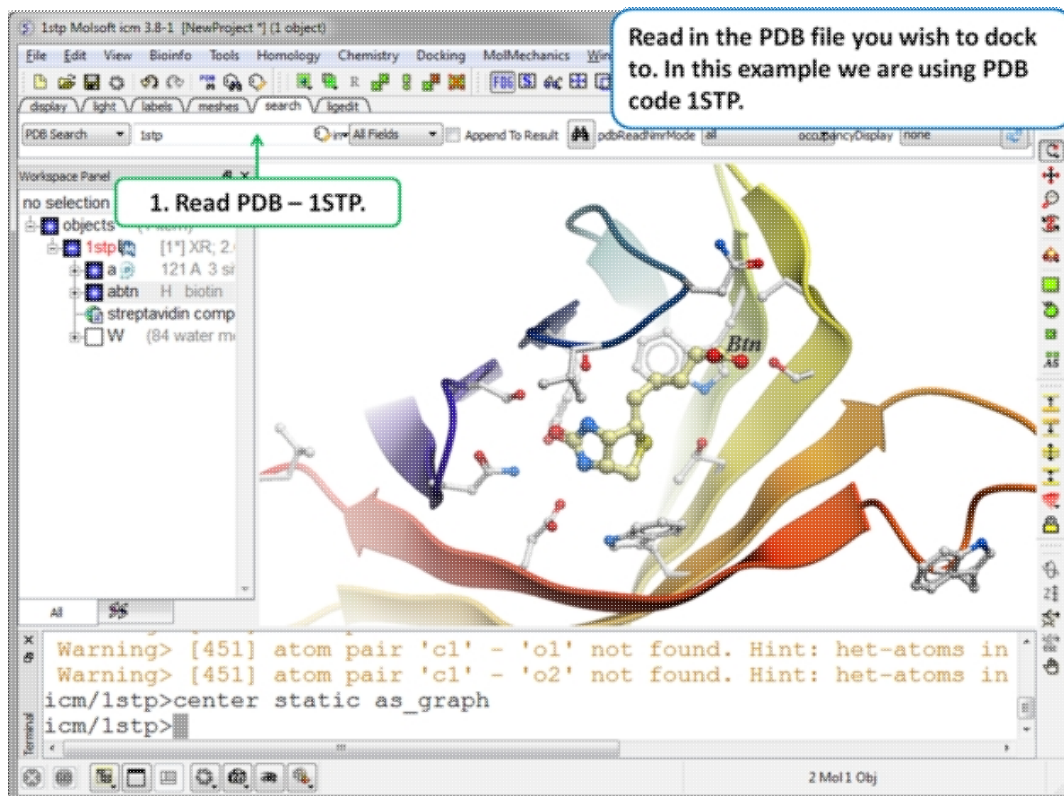
To determine protonation at a range of pH values you can use the options in `Chemistry/Set Formal Charges/` and choose the option **Auto using Pka Model**.

11.1.2 Docking Steps

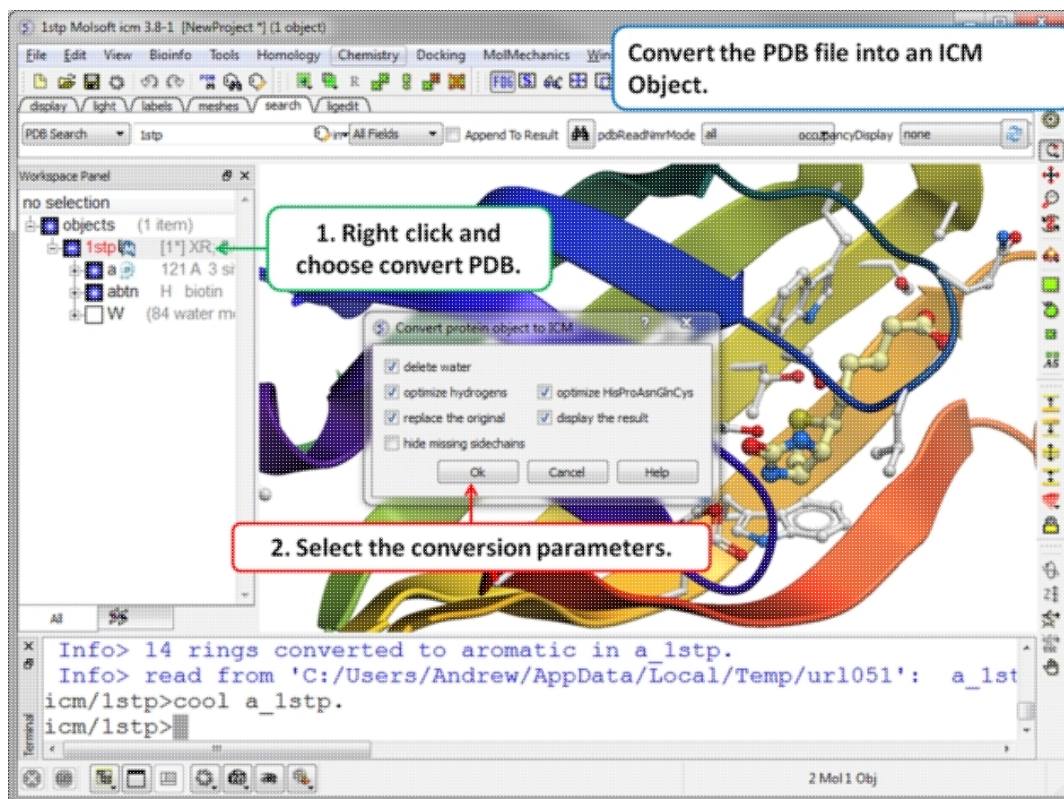
The key docking steps are:

1. Setup the Docking Project.
2. Check Docking Preferences.
3. Choose Ligand or Database and Start Docking Simulation
4. View Docking Results

11.1.3 Docking Project Setup

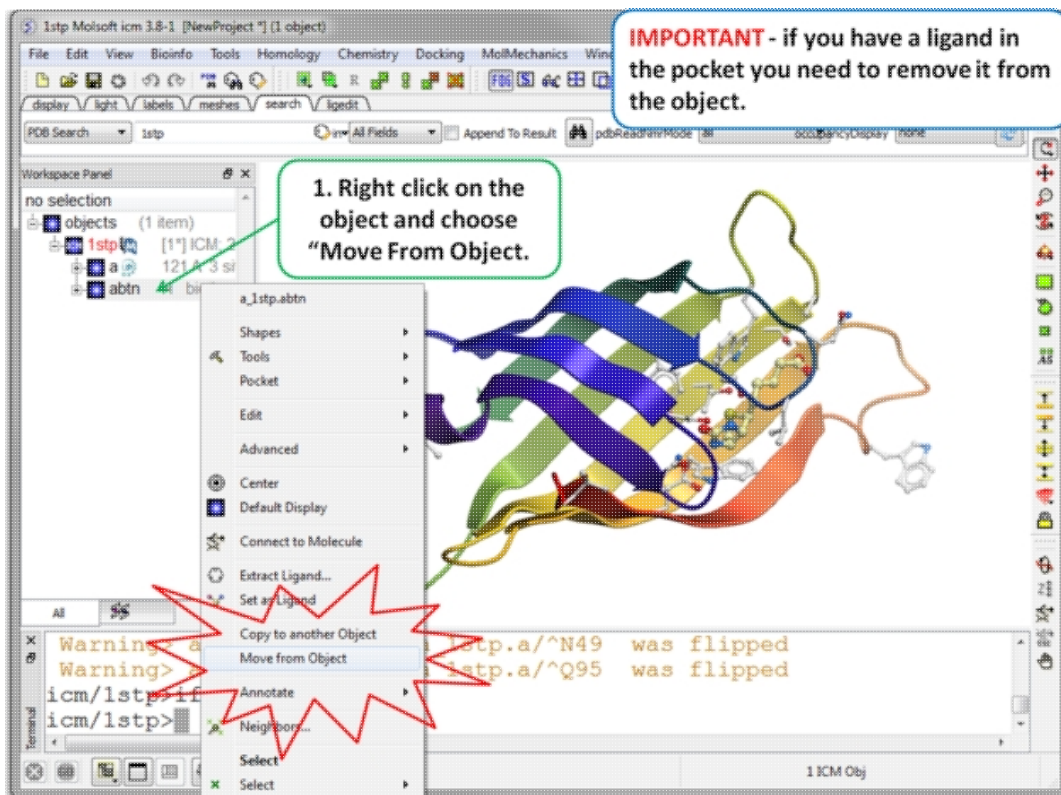


Step 1. Read in the PDB file. In this example we will re-dock the biotin to streptavidin using PDB code 1STP.

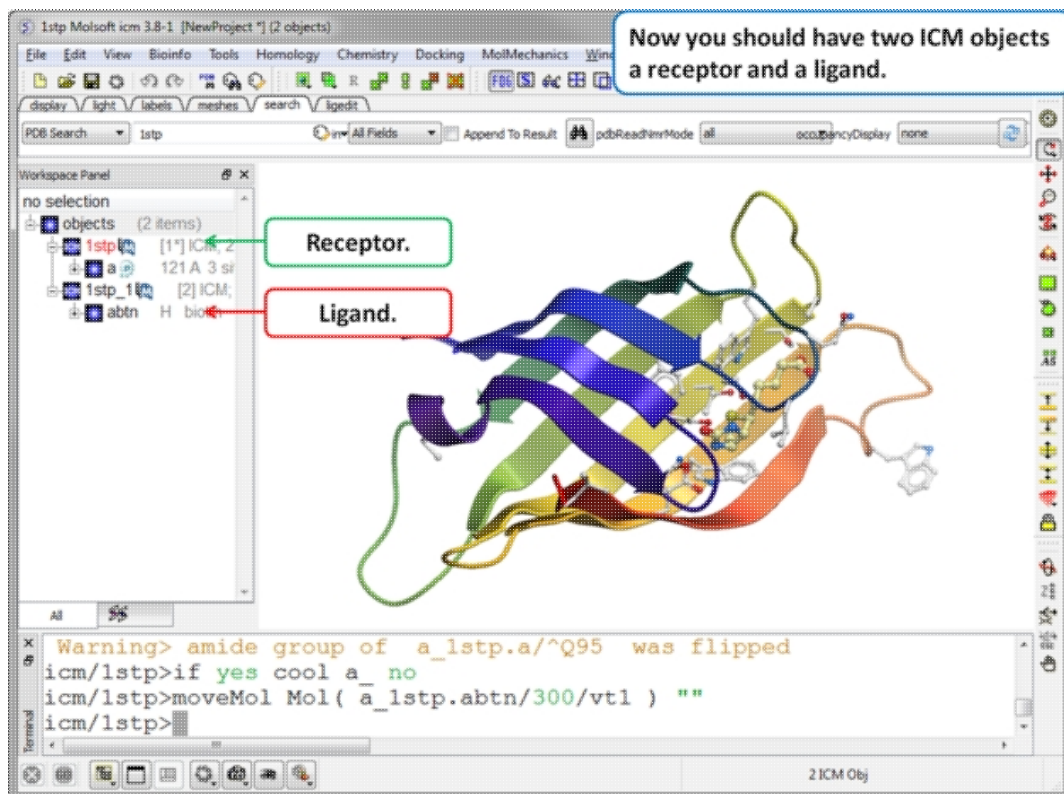


Step 2. Convert PDB

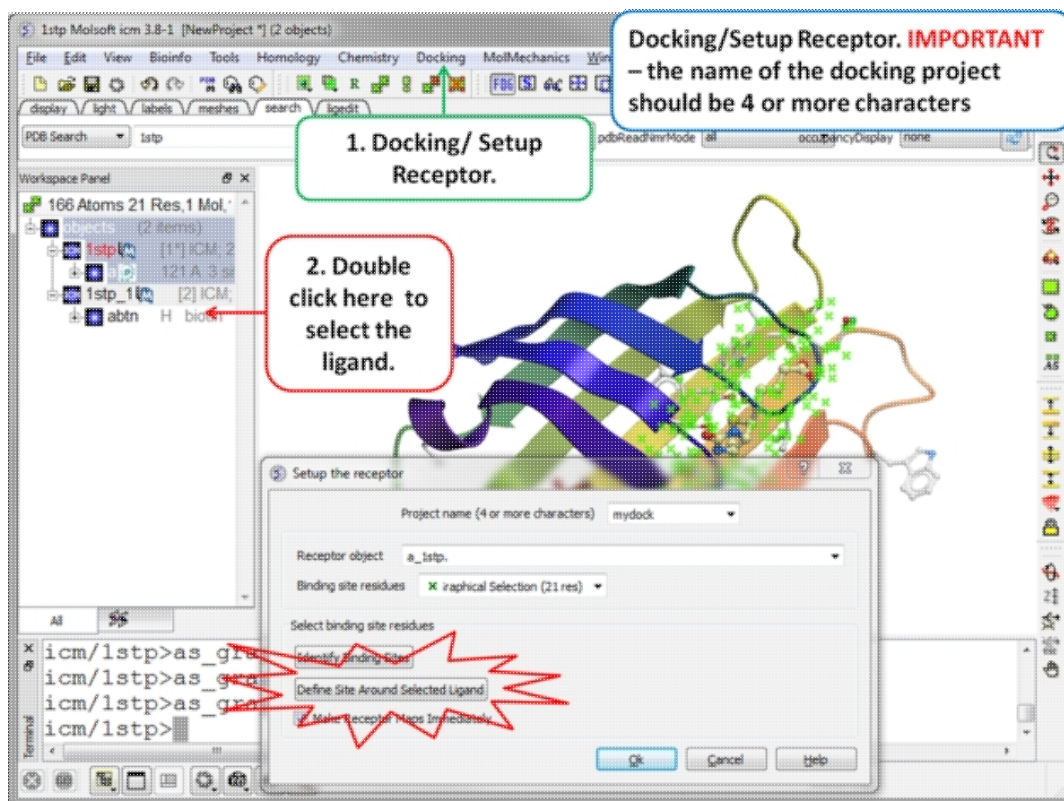
- Convert the PDB file into an ICM object. Remember to select the option to optimize all hydrogens and Optimize HisProAsnGlnCys.
- It is recommended that "optimize hydrogens" option is selected. To accelerate the procedure, disable the 3D graphics window (type in the terminal window `unds` window) When the procedure finishes, converted object is the 'current' object in icm. You can check the results by displaying the converted structure.
- If you do not select "replace original" just make sure you understand which of your objects is an ICM Object and which one is in PDB format.
- In this example we delete all water molecules but for some cases you may need to keep one or two waters in the pocket. We usually recommend to keep a water molecule if it is tightly bound (e.g. 2-3 hydrogen bond coordination).
- Uncheck the box hide missing side chains if you want ICM to build missing heavy atoms that are not reported in the PDB (due to the lack of density), they will be added according to the residue name and assigned zero occupancies. Check this box if you want residues missing heavy atoms to be hidden.



Step 3 Move the Ligand out of the Pocket IMPORTANT! If you are redocking a ligand please remember to remove the ligand from the ligand binding pocket otherwise the ligand will be included in the docking maps and you will not be able to re-dock it correctly. To remove a ligand from an object - right click on the ligand in the ICM Workspace and select "move from object". **Simply undisplaying the ligand is NOT sufficient.**



Step 4. Check the Ligand. You should have two objects in the ICM Workspace - the receptor (e.g. 1stp) and the ligand (1stp_1). For clarity you can rename them by right clicking on the name in the ICM workspace and choose rename. This is a good opportunity to double check that the bond types and formal charge of the ligand are correct. You can edit them as described [here](#).



Step 5. Docking/Setup Receptor

- Docking/Setup Receptor.
- Enter a unique alphanumeric name (>3 characters) into the Project name data entry box. Avoid using spaces and leading digits in the name and do not include unusual characters such as hyphens etc... **NOTE:** To load a project that has already been setup go to Docking/Open Project.
- Enter the name of the receptor molecule(s) in the **Receptor object** data entry box (e.g. a_1stp. - you can use the drop down arrow to locate the name if more than one objects are loaded). In most cases a_* will do - all molecules in the current object will be included, this is useful if your pocket contains 2 chains, hetatoms or waters.

There are different ways to enter the binding site residues:

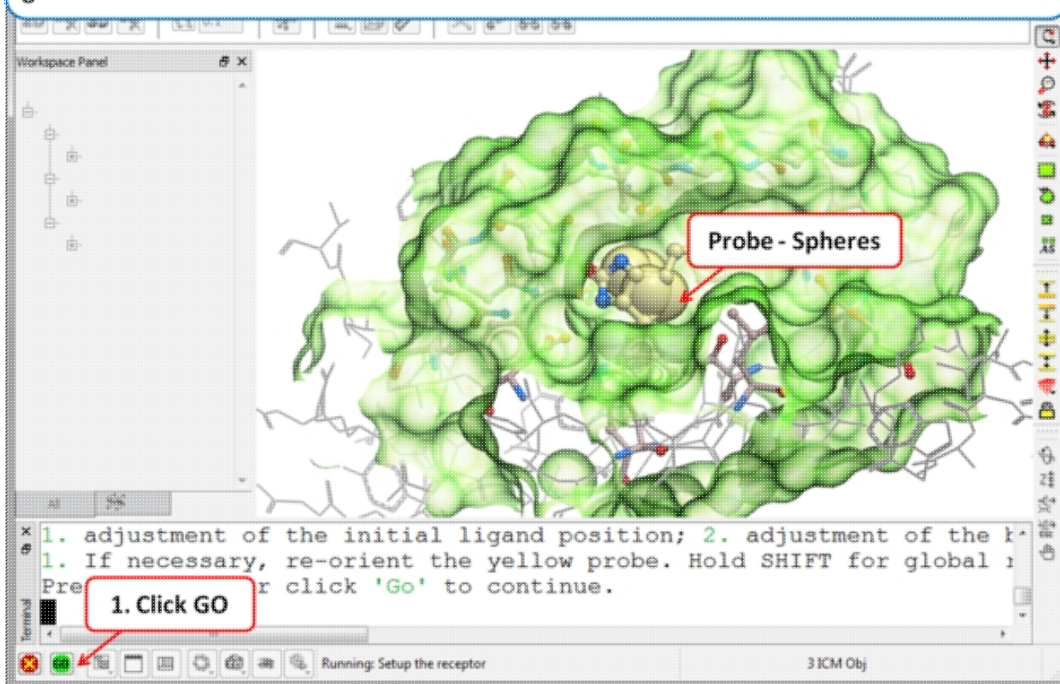
1. Define the binding site residues manually e.g. by entering the selection in the **Binding site residues** data entry box e.g. a_/123,144,152 for selection by residue numbers.
2. Select the site using the graphical selection tools such as the lasso tool.
3. Use the `icmPocketFinder` function option or press the **Identify Binding Sites** button.
4. **Easiest Way!** - as shown in image If you have a ligand in the binding pocket already - select the ligand in the ICM Workspace by double clicking on it. Then press the "Define Site Around Selected Ligand" button. This will make a graphical selection (green crosses) of the residues surrounding the ligand. If you use this option make sure the ligand is in a separate object and not in the same object as the receptor (See Step 3).

Note if you have made your selection correctly the pocket will be surrounded by green crosses.

- Choose the option **Make receptor maps immediately**.
- Click on the OK button.

NOTE: At this stage of the docking setup it is a good idea to keep an eye on the terminal window. Instructions and any error messages will be displayed in the terminal window. If you do not see the terminal window select **Windows/Terminal Window**.

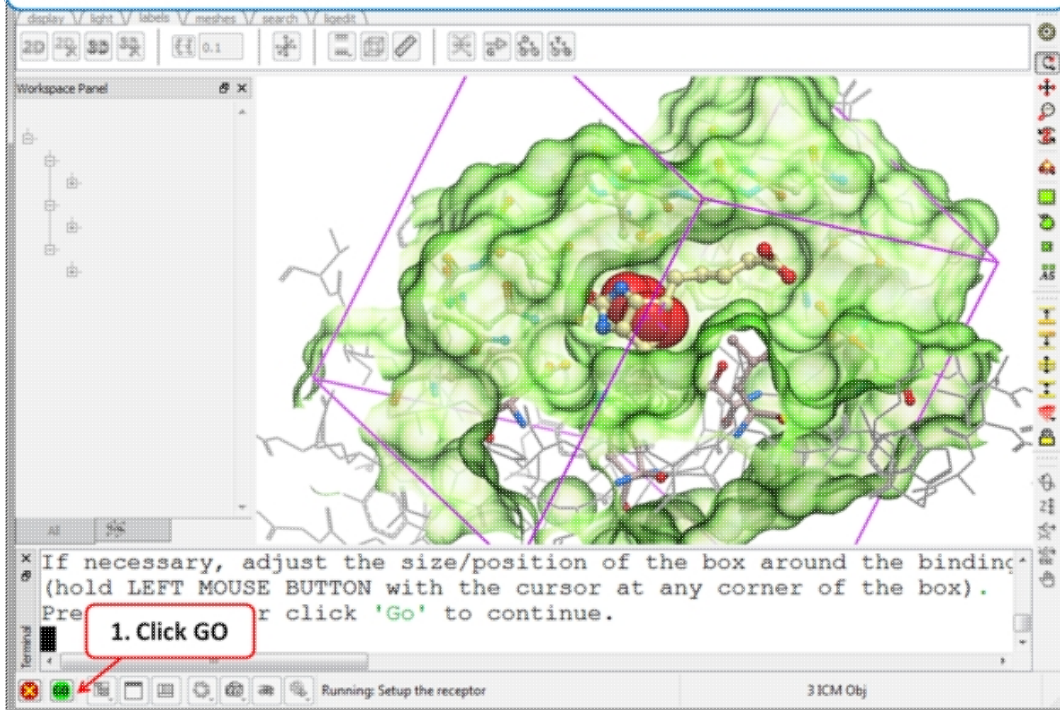
The probe represents the docking starting position. Generally you can keep it in the default position. If you want to move it use the middle mouse button and hold the SHIFT button for global rotation.



Step 6: Adjusting the Position of the Probe The position of the probe (usually represented as 4 spheres in the center of the pocket) represents the initial position where sampling will begin. The default probe position is generally OK for most purposes but if you would like to move it to a critical part of the receptor so that sampling initially concentrates in that region you can do so using the middle mouse button and holding the SHIFT button for global rotation. Once you are happy with the position of the box press the enter key or click on "GO".

NOTE The probe position can be changed again if needed using the **Docking/Review/Adjust Ligand/Box..** option.

The purple box represents the region in which maps will be generated. To change the box size use the left mouse button with the cursor at any corner of the purple box to change it.



Step 7: Adjusting the Size of the Box The purple box represents the region in which maps will be generated. The box needs to be large enough to encompass the binding pocket but not too large and including regions of the receptor which are not relevant for the ligand to bind. If the binding site is correctly defined in the earlier **Receptor setup** then the default box size is usually fine. If it is necessary to change the box size you can use the left mouse button with the cursor at any corner of the purple box to change it.

NOTE The size of the box can be changed again if needed using the `Docking/Review/Adjust Ligand/Box...` option but if you do change something please remember to re-make the maps **Docking/(Re) Make Receptor Maps**.

11.1.4 Review and Adjust Binding Site

If at any point you want to change the binding site use the option `Docking/Review Adjust Ligand/Box...` Here you can change the initial position of the probe (default center or box) and the size of the box. If you change the size of the box you need to re-make the maps using the option `Docking/(Re) Make Receptor Maps`. You can use this option to also change the default map parameters such as the grid cell size and map weighting by occupancy.

11.1.5 (Re) Make Receptor Maps

If you change the size of the binding site box you will need to re-make the maps. To do this:

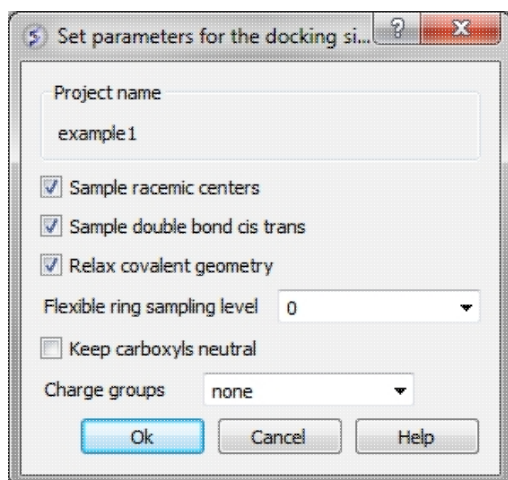
- `Docking/(Re) Make Receptor Maps`

11.1.6 Check Docking Preferences

The docking preferences options can be found here `Docking/Preferences`.

11.1.6.1 General Preferences

Docking/Preferences/General



You can choose to:

- Sample racemic centers.
- Sample double bonds as cis trans.
- Relax covalent bond geometry (set as yes by default).
- Flexible ring sampling level. This option can be used for macrocycles. If the value is set to "1" then the ring is only flexible in pre-sampling. If the value is set to "2" then the ring is flexible throughout the simulation.
- Keep carboxyls neutral (set as no by default).
- Charge groups: provides simple interface to charge some ionizable basic groups, it currently understands following values: NH₂, NH, NT for primary secondary and tertiary aliphatic amines, also it understands 'imidazole' and 'amidine'. An alternative is 'auto', which uses built-in prediction of Ka/Kb to charge and protonate/deprotonate appropriate groups.

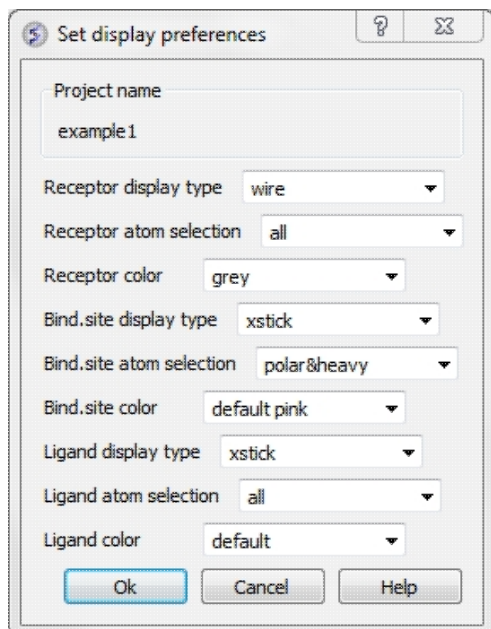
11.1.6.2 Database Scan Preferences

Docking/Preferences/Database Scan

These options are required when undertaking large scale docking and scoring using Virtual Ligand Screening (VLS). They are described in the VLS chapter of this manual.

11.1.6.3 Display Preferences

A number of docking display options can be changed in Docking/Preferences/Display

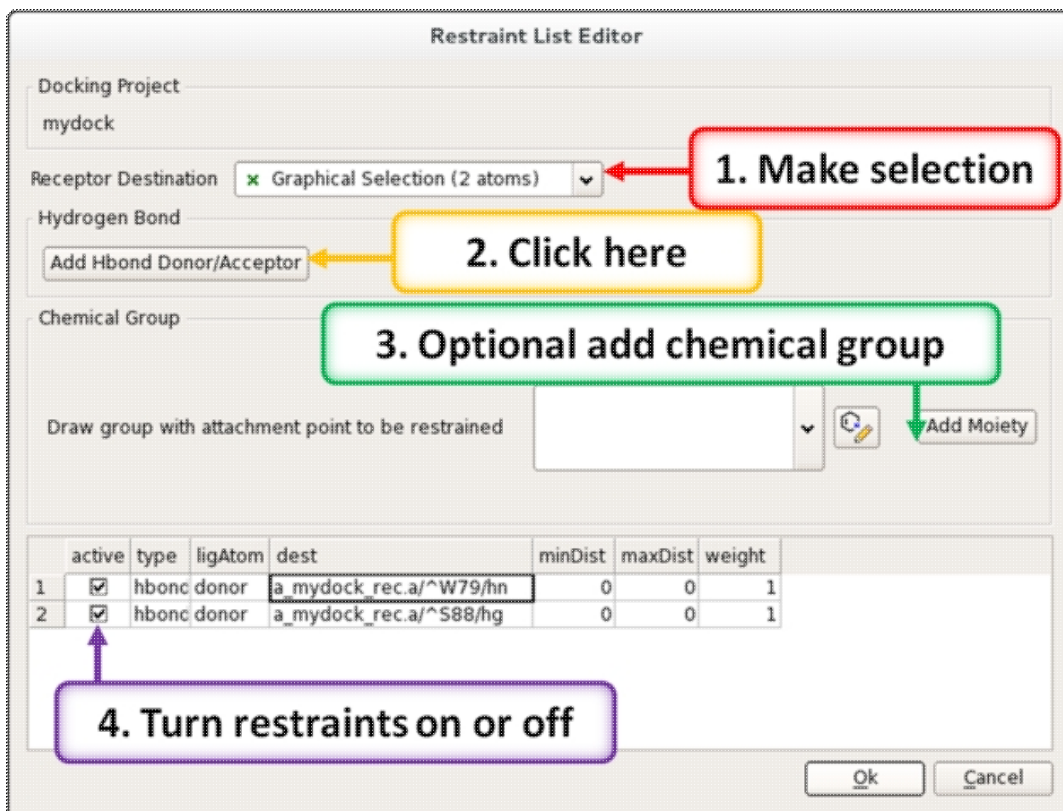


Now begin the docking procedure.

11.1.7 Interaction Restraints

To restrain the docking to a particular hydrogen bond donor or acceptor in the receptor.

- Docking/Interaction Restraints.
- Select the atoms in the docking project receptor that you would like the ligand(s) to interact with and click on the "Add Hbond Donor/Acceptor" button.
- You can also define a chemical moiety that you would like to interact with the receptor. To do this click add the SMILES string or click on the button to sketch the chemical group. An attachment point needs to be defined (right click on atom and choose Attachment Point from the menu).
- A table of the restraints will be displayed. You can check the restraint on and off using the check marks in the "Active" column.



11.1.8 Begin the Docking Simulation

Once the receptor and maps have been correctly set up then the docking procedure can begin.

There are three possible options to choose when INTERACTIVE, BATCH, or Dock Chemical Table docking (Please note some of the options may be limited for users without ICM-VLS)

- INTERACTIVE docking should be used when you have a single ligand and you simply want to re-dock the ligand to the receptor.
- BATCH docking will run in the background and is an ideal choice for screening a database of compounds e.g. vls.
- Dock Chemical Table will allow you to easily dock a chemical spreadsheet that is loaded into ICM.

11.1.9 Interactive Docking

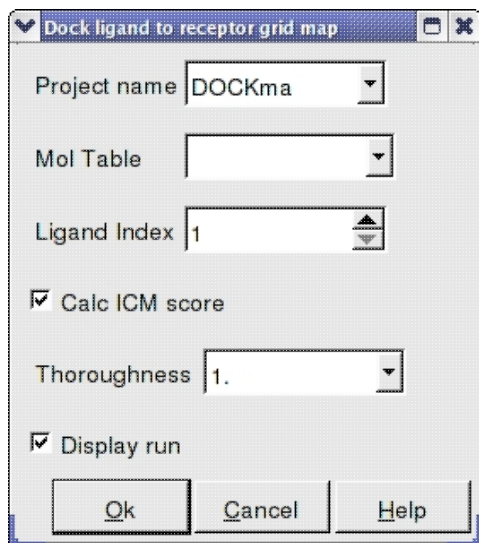
Use interactive docking to dock one ligand at a time in the foreground. It is ideal to use this option for small-scale docking.

- Click on the menu Docking/Interactive Docking

Choose either Mol Table Ligand or Loaded Ligand

11.1.9.1 Interactive Docking: Mol Table Ligand

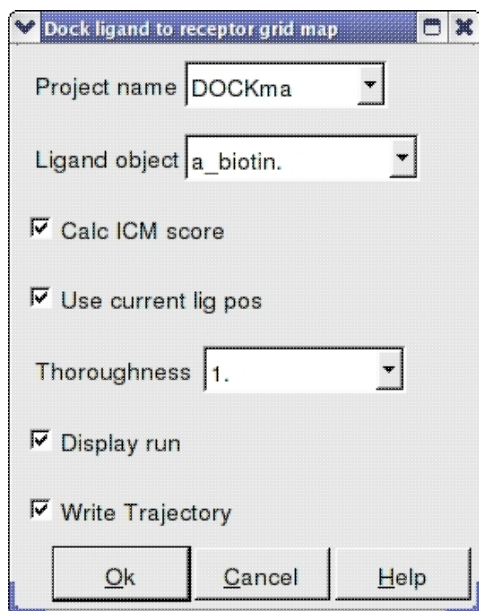
If you have a chemical table already loaded into ICM you can use this option to dock chemicals in it. You can read mol/mol2 or sdf files into ICM by using File/Open. They will be displayed in a table.



- Enter the Docking Project Name
- Use the drop down arrow to find the table of ligands you wish to dock.
- Enter the number of the ligand in the table you wish to dock. Eg if the ligand is in row 6 enter 6.
- If you have ICM-VLS you can retrieve an ICM docking score for the docked ligand.
- 'Effort' or thoroughness represents the length of the simulation. Generally 1 is a reasonable value for buried hydrophobic pockets. If you are docking to solvent exposed pockets or pockets containing metal ions you may wish to increase this parameter.
- Display run will display the ligand sampling the energy in the ligand binding pocket. Although this is fun to watch this significantly slows down the docking operation.

11.1.9.2 Interactive Docking: Loaded Ligand

If you have a ligand as an ICM object you can use this option.



- Enter the Docking Project Name
- Use the drop down arrow to find the ligand.
- If you have ICM-VLS you can retrieve an ICM docking score for the docked ligand.
- If the ligand is already located in the pocket you can use this option. However by default the ligand will start sampling in the center of the pocket so this option does not need to be used.
- Thoroughness represents the length of the simulation. Generally 1 is a reasonable value for buried hydrophobic pockets. If you are docking to solvent exposed pockets or pockets containing metal

ions you may wish to increase this slightly.

- Display run will display the ligand sampling the energy in the ligand binding pocket. Although this is fun to watch this significantly slows down the docking operation.
- You can write the docking simulation to a trajectory file. Please see the command language manual for more information on this.

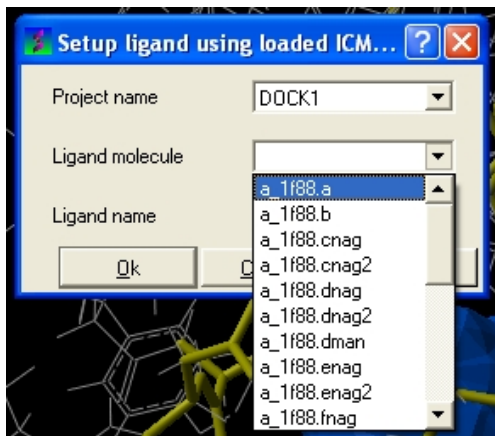
11.1.10 Batch Docking

Batch Docking is used for running docking jobs in the background. It is ideal for large-scale docking jobs.

- Use the menu option Docking/Batch Ligand Setup to setup the docking run
- Once setup use the menu option Docking/Run Docking Batch to run the batch docking.

11.1.10.1 Batch Docking: From Loaded ICM Object

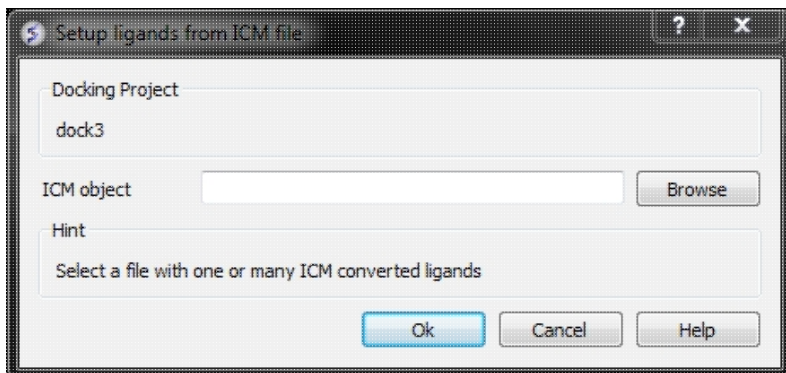
From Loaded ICM Object Your ligand needs to be an ICM object and loaded into ICM (File/Open). Your object will be displayed in the ICM Workspace.



- Enter the name of the docking project.
- **Ligand Molecule** Use the drop down button to find the loaded ligand or enter the the ICM selection language.
- **Ligand Name** A docked ligand will be given a name. Change this if you wish or use the default name.
- Once setup use the menu option Docking/Run Docking Batch to run the batch docking.

11.1.10.2 Batch Docking: From File

From File: ICM If your ligand (s) is saved and converted to an ICM object but is not loaded into ICM then you can use this option.

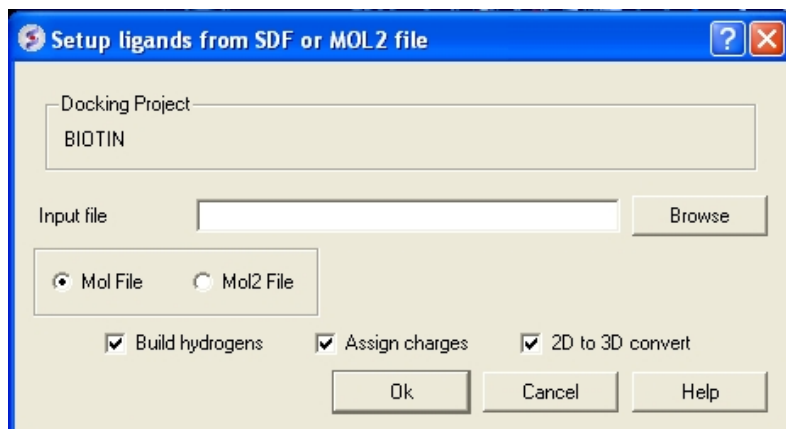


- Browse for the ICM Object.
- Click OK

11.1.10.3 Batch Docking: From Mol/Mol2

From File:MOL/MOL2

If your ligand is a MOL or MOL2 file then



- Browse for your MOL/MOL2 file.
- Select whether your ligand is in MOL or MOL2 format.
- If you wish hydrogens to be added to your compound or charges to be assigned then click on the appropriate boxes in the display panel.
- Click OK

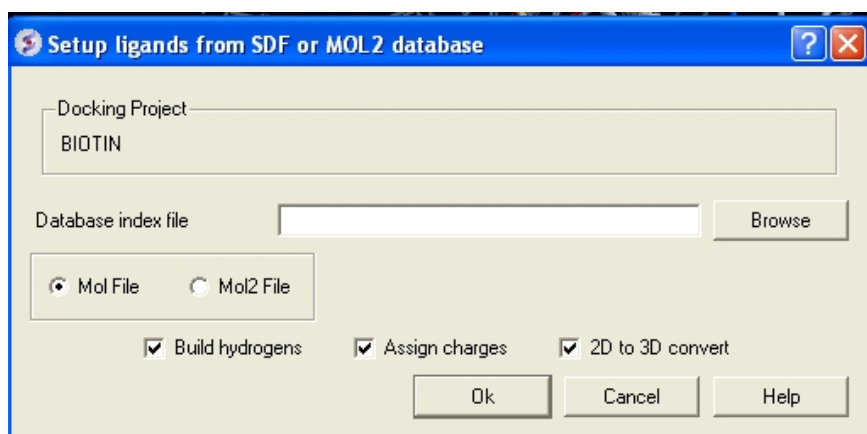
11.1.10.4 Batch Docking: From Indexed Database

From Indexed Database - only available with ICM-VLS

In most cases the ligand input file will be an SDF or MOL2 file. These files need to be indexed by ICM before they can be used in VLS runs (see next section of this manual). The index is used to allow fast access to an arbitrary molecular record in a large file such as an SDF file which in some cases contains over one million compounds.

To index an sdf file:

- Click on the menu Docking/Tools/Index Mol/Mol2 File/Database to generate the index. The following data entry box will be displayed.



- Enter the name of your Mol/Mol2 file and enter the name you wish to call your index file.
- Select whether your file is in Mol or Mol2 format.
- Browse for your Index file.
- Select whether your ligand is in MOL or MOL2 format.
- If you wish hydrogens to be added to your compound or charges to be assigned then click on the appropriate boxes in the display panel.

- Click OK

11.1.10.5 Batch Docking: From MolCart Database

From MolCart - only available with ICM-VLS

NOTE A separate license is required for MolCart

- Enter docking project name.
- Enter the MolCart server
- Enter your username
- Enter Password
- Enter the Database Name
- Enter the name of the MolCart table within the Database
- Select whether you would like to build hydrogens or convert the compounds from 2D to 3D.

11.1.11 Run Docking Batch

IMPORTANT Once you have setup the docking you need to set it running using the menu option Docking/Run Docking Batch

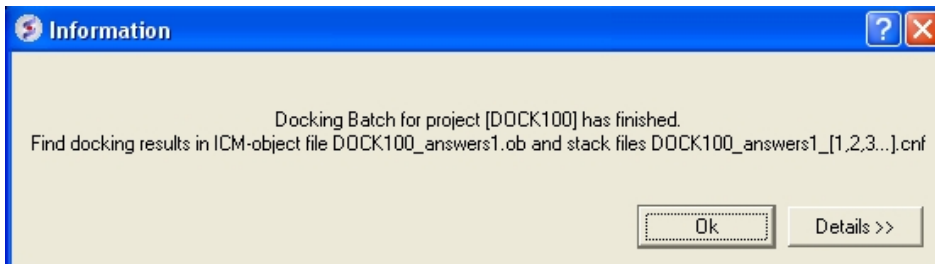
- Docking/ Run Docking Batch
- Enter the Docking Project Name
- Batch docking will generate an output file - enter the name you wish to call this file here.
- Thoroughness represents the length of the simulation. Generally 1 is a reasonable value for buried hydrophobic pockets. If you are docking to solvent exposed pockets or pockets containing metal ions you may wish to increase this slightly.

- Selecting Store Alternative Conf will allow you to look at all conformations in the energy stack.
- Score All Stack Conf. will allow you to determine an ICM docking Score for all members of the stack - this will slow the docking down.
- License type: For nearly all license (e.g. standard ICM-VLS licenses) types you need to **leave this entry blank**. If you have a -vlscluster license or a molvls license (-vls) then select the option from the drop down list. Any questions first check your license.dat file or Email support@molsoft.com.
- Click **OK IMPORTANT - IF THE JOB IS RUNNING IT WILL TELL YOU bgrnd job AT THE TOP OF THE GUI - SEE BELOW**

Docking simulation is running

(3 objects 1 table 1 bgrnd job)

Docking simulation has ended message



To check the status of your docking simulation

- Windows/Background Jobs

11.1.12 Viewing Your Docking Results

Docking results can be visualized and browsed in one of the following ways.

- Docking/Browse/Scan Hits - If you have docked a multi-ligand file (eg SDF) or if you want to go back and see a single docked structure you can scan the structures using this option.
- Docking/Browse/Stack Conformation - View other possible conformations for the docked ligand ranked by energy.
- Docking/Make Hit List - Rank docked compounds by ICM Docking Score **Only available with ICM-VLS**

The results of the docking are saved in the following files

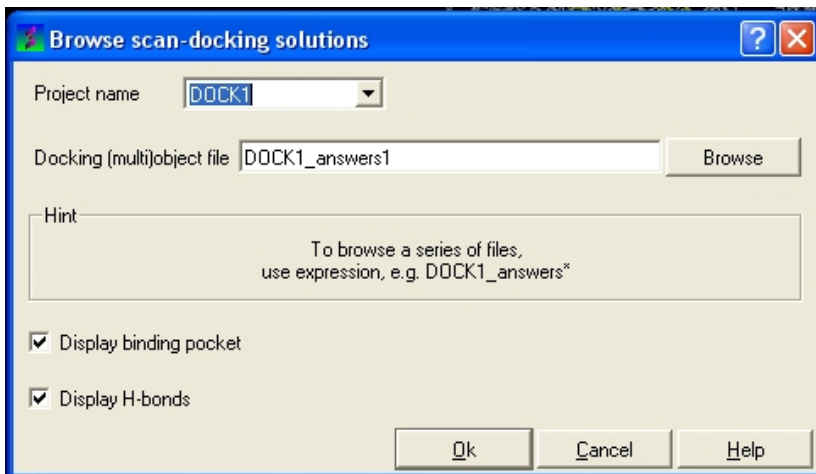
PROJECTNAME_LIGANDNAME.ob #icm-object file with best solutions for each ligand
 PROJECTNAME_*.cnf # icm conformational stack files with multiple docked conf.

The results of the docking job using ICM-VLS (separate license required) are saved in the following files:

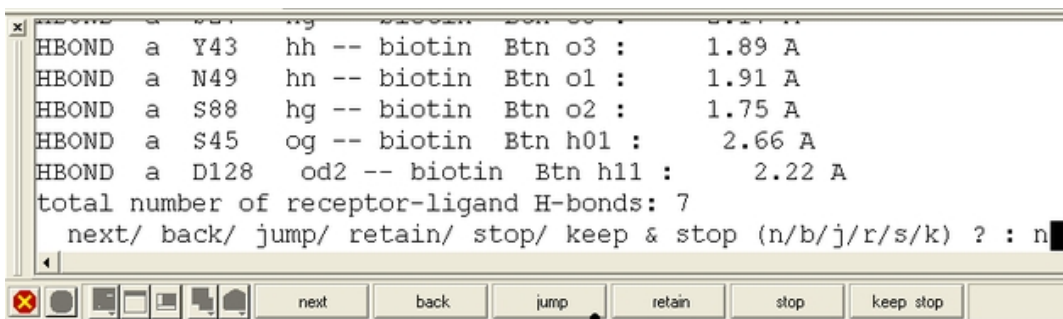
PROJECTNAME_answers*a.ob #icm-object file with best solutions for each ligand
 PROJECTNAME_*.cnf # icm conformational stack files with multiple docked conf.
 PROJECTNAME_*.ou # output file where various messages are stored eg.SCORE

11.1.12.1 Docking Results - Scan Hits

- Select **Docking/Browse/Scan Hits** and a data entry box as shown below will be displayed.



- Select the correct project name for the docking simulation results you wish to browse.
- Enter the name of the icm object file in the Docking (multi)object data entry field. This file will be called PROJECTNAME_answers*.ob or LIGAND_NAME.ob The browse button can be used to search for the correct file.
- You can display the binding pocket or the H-bonds by selecting the appropriate boxes in the Browse scan-solutions data entry window (shown above).
- Use the buttons at the bottom of the graphical user interface to browse the docked conformations. **NEXT**(or type "n"), **BACK** (or type "b"), **JUMP** (or type "j"), **RETAIN** (or type "r"), **STOP** (or type "s"), **KEEP_STOP** (or type "k").



Use these buttons to browse the docked conformations

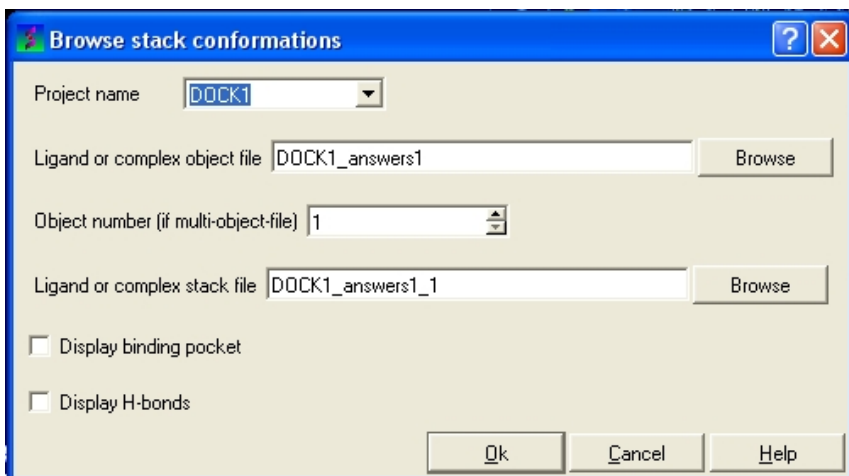
- The options **keep and stop** and **retain** will retain the displayed ligand in the graphical user interface. If you want to export the docked complex as a PDB file you will need to move the ligand and receptor into one ICM object. Moving objects is described in the FAQ section entitled How can I merge two objects into one?

11.1.12.2 Docking Results - View Stack Conformations

To view the multiple positions of a single ligand in the docking simulation ranked by energy.

- Select menu Docking/Browse/Stack Conformations

The Browse Stack Conformation data entry window will be displayed.



- Select the correct project name for the docking simulation results you wish to browse.
- Enter the name of the icm object file in the Docking (multi)object data entry field. This file will be called PROJECTNAME_answers*.ob .The browse button can be used to search for the correct file.
- Enter the name of the icm conformational stack files with multiple docked conformations into the Ligand or complex stack file data entry box. This file will be called PROJECTNAME1_1.cnf . The browse button can be used to search for the correct file. The second solution in the stack can be viewed by changing the number 1 at the end of the file name to 2 (PROJECTNAME1_2.cnf) and so on for each solution in the stack.
- You can display the binding pocket or the H-bonds by selecting the appropriate boxes in the Browse scan-solutions data entry window (shown above).
- Use the buttons at the bottom of the graphical user interface to browse the docked conformations. **NEXT**(or type "n"), **BACK** (or type "b"), **JUMP** (or type "j"), **RETAIN** (or type "r"), **STOP** (or type "s"), **KEEP_STOP** (or type "k").
- The options **keep and stop and retain** will retain the displayed ligand in the graphical user interface. If you want to export the docked complex as a PDB file you will need to move the ligand and receptor into one ICM object. Moving objects is described in the FAQ section entitled How can I merge two objects into one?

Columns in the Stack Table

i rank in stack

ener Energy kcal/mol

gvw van der Waals grid potential

gb hydrogen bonding grid potential

ge electrostatic grid potential

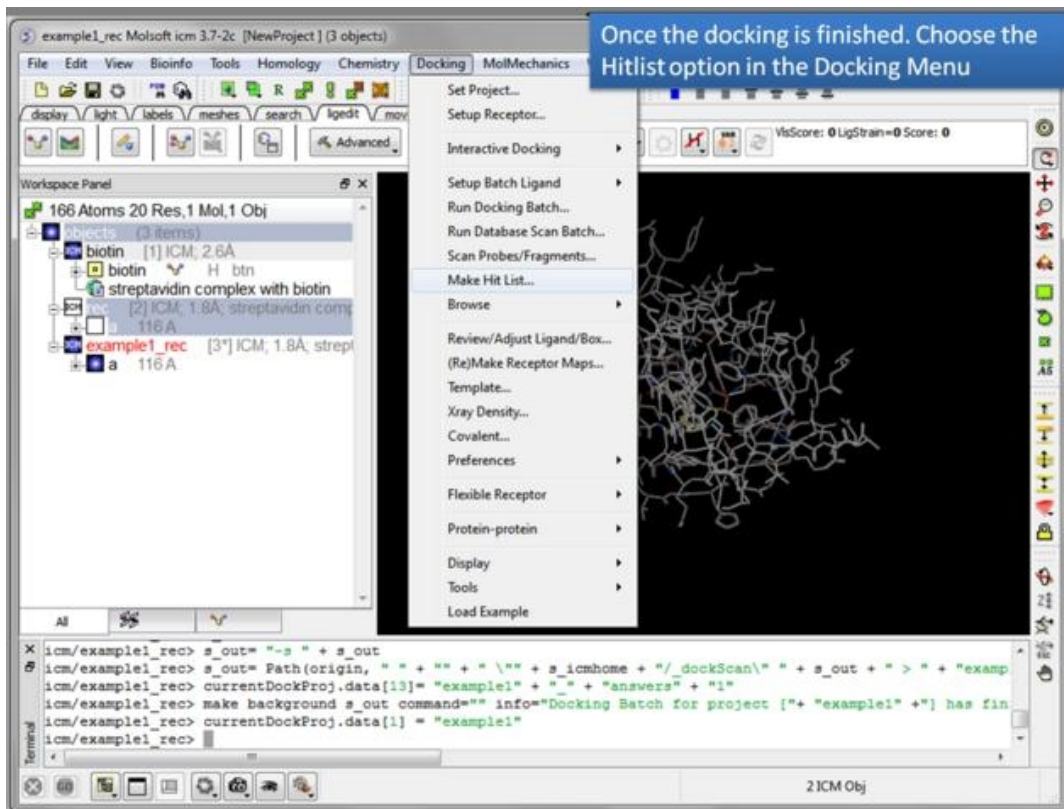
gs hydrophobic grid potential

Einternal is internal conformation energy of the ligand

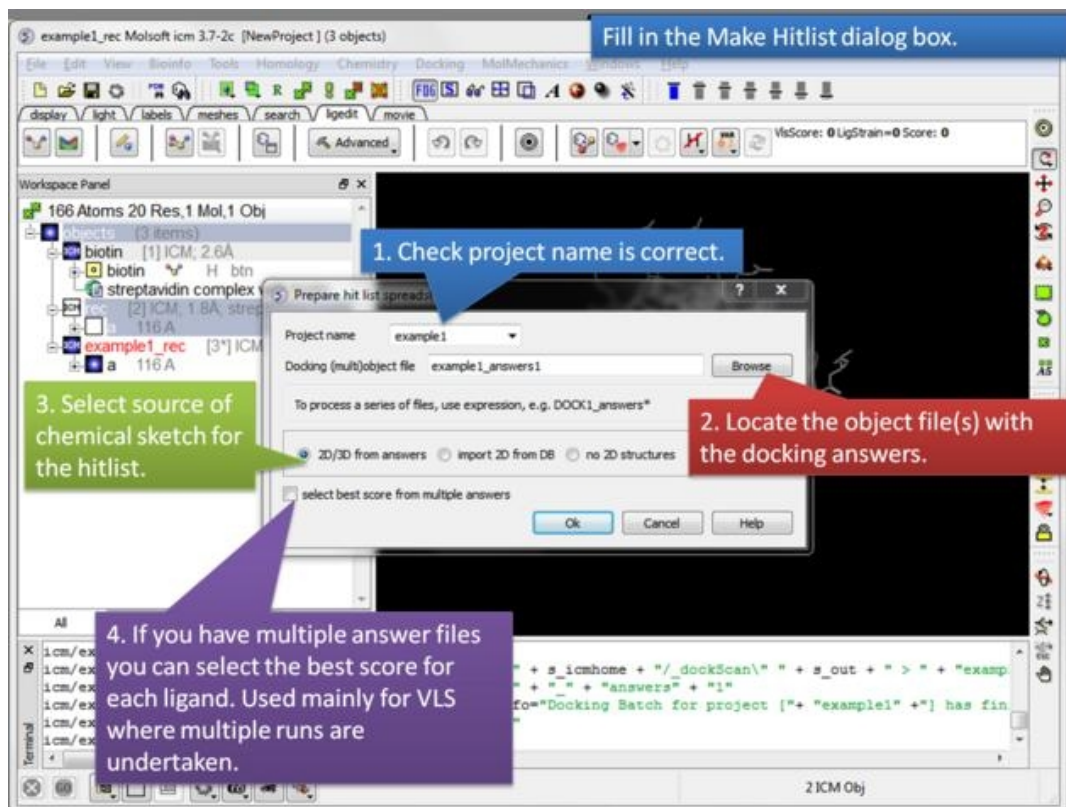
11.1.13 Make a HIT LIST

A hilit is a convenient way to view your docking results in a chemical spreadsheet. To make a hitlist

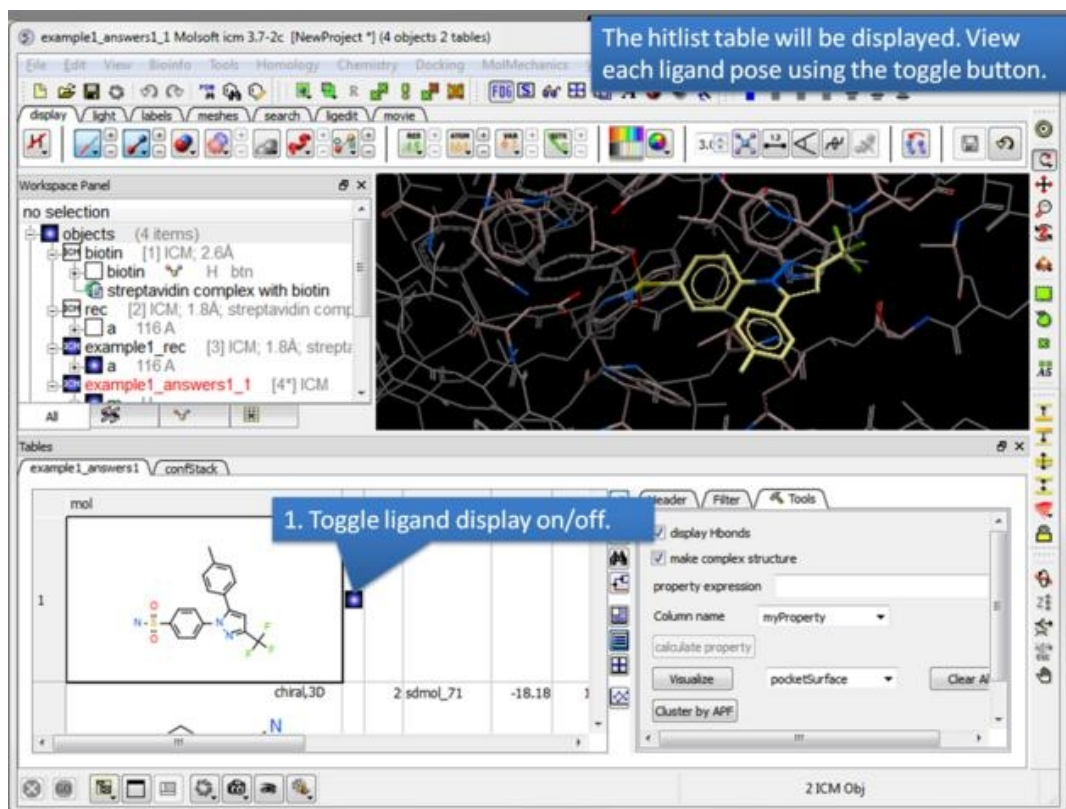
- Docking/Make Hit List



- Enter project name
- Use the browse button to locate DockingProjectName_answers.ob file
- You can include a 2D image into the HITLIST
- Select **Unique** if you have made multiple docking runs the best docking score will be taken to make the hitlist unique.
- A HITLIST table will be displayed. Each docked ligand can be viewed by double clicking in the HITLIST table. A stack of conformations for each ligand will also be displayed in a table.



There is a toggle button to easily undisplay or display each ligand pose.



Columns in the HitList Table

IX is the index number from the docked database

Score is the ICM score -32 and lower are generally considered good scores - but depends on the receptor (e.g. exposed pockets or pockets with metal ions may have higher scores than -32)

Natom is the number of atoms in docked ligand

Nflex is the number of rotatable torsions

Hbond is Hydrogen Bond energy

Hphob is the hydrophobic energy in exposing a surface to water

VwInt is the van der Waals interaction energy (sum of gc and gh van der waals). Current version of the score uses explicit van der Waals interaction energy calculation (no grids)

Eintl is internal conformation energy of the ligand

Dsolv is the desolvation of exposed h-bond donors and acceptors

Solel is the solvation electrostatics energy change upon binding

mfScore is the potential of mean force score

RecConf - if multiple receptor conformations was used Docking/Flexible Receptor/Setup 4D grid and represents the receptor conformation number

NOTE There are a number of post-screening analysis built into the **Tools** panel on the right hand side of the hitlist. These are described [here](#).

example1_answers1_1 Molsoft icm 3.7-2c [NewProject *] (4 objects 2 tables)

Sort a column (e.g. Score) right click on column header.

1. Right click here and choose "Sort".

mol	L	IX	NAME	Score	Natom
1		m		-24.67	
		2	sdmol_71	-18.18	1

2 ICM Obj

How to sort the hitlist. Right click on a column header and select "Sort". For example you may want to sort by docking score.

example1_rec Molsoft icm 3.7-2c [NewProject *] (3 objects 1 table)

Single Click to make Ligand and Receptor complex in ONE object.

Workspace Panel

no selection

- objects (3 items)
 - biotin [1] ICM, 2.6Å
 - biotin H btn
 - streptavidin complex with biotin
 - rec [2] ICM, 1.8Å, streptavidin co
 - a 116 Å
 - example1_rec [3] ICM, 1.8Å
 - a 116 Å
 - m H

2. Ligand and receptor complex in ONE object displayed here

1. Single click here

mol	L	DX	NAME	Score	Natoc
1	30	1	m	-24.67	
			chr1,3D	-18.18	
			2 sdmol_71	-18.18	

Header Filter Tools

- display Hbonds
- make complex structure

property expression

Column name myProperty

calculate property

Visualize pocketSurface Clear All

Cluster by APF

1 ICM Obj

Single click. A single click on a row in the hitlist will load the ligand and receptor complex in one object.

example1_rec Molsoft icm 3.7-2c [NewProject *] (3 objects 1 table)

If the ligand and receptor are in the same object you can save it as a PDB file.

Workspace Panel

no selection

- objects (3 items)
 - biotin [1] ICM, 2.6Å
 - biotin H btn
 - streptavidin complex with biotin
 - rec [2] ICM, 1.8Å, streptavidin co
 - a 116 Å
 - example1_rec [3] ICM, 1.8Å
 - a 116 Å
 - m H

1. Right click here and choose "Save As" -> then choose Type PDB

mol	L	DX	NAME	Score	Natoc
1	30	1	m	-24.67	
			chr1,3D	-18.18	
			2 sdmol_71	-18.18	

Header Filter Tools

- display Hbonds
- make complex structure

property expression

Column name myProperty

calculate property

Visualize pocketSurface Clear All

Cluster by APF

1 ICM Obj

Save a complex as a PDB file. If the ligand and receptor are in the same object you can save them both as a PDB file. Right click on the name of the complex in the ICM Workspace. Select "Save as" and then a Windows dialog window will be displayed. Use the drop down button to select Type PDB (.pdb .ent).

example1_rec Molsoft icm 3.7-2c [NewProject *] (3 objects 2 tables)

Workspace Panel

no selection

- objects (3 items)
 - biotin [1] ICM; 2.6Å
 - biotin H btn
 - streptavidin complex with biotin
 - rec [2] ICM; 1.8Å; streptavidin comp
 - a 116 A
 - example1_rec [3] ICM; 1.8Å; strep
 - a 116 A
 - m H

Tables

example1_answers1 confStack

1. Double click here

2. New table of stack conformations

mol	L	DX	NAME	Score	Nato
3D	1	m		-24.67	
1			chra1,3D	2	sdmo1_71
				-18.18	

Header Filter Tools

- display Hbonds
- make complex structure

property expression

Column name myProperty

calculate property

Visualize pocketSurface Clear All

Cluster by APF

1 ICM Obj

Display stack of alternative ligand conformations. Double click on a row generates a new stack conformation table.

example1_rec Molsoft icm 3.7-2c [NewProject *] (3 objects 2 tables)

Workspace Panel

no selection

- streptavidin complex with biotin
- rec [2] ICM; 1.8Å; streptavidin comp
- a 116 A
- example1_rec [3] ICM; 1.8Å; strep
- a 116 A
- m H
- tables (2 items)
 - example1_answers1 50 rows 18 cols
 - confStack 25 rows 9 cols 4 headers

Tables

example1_answers1 confStack

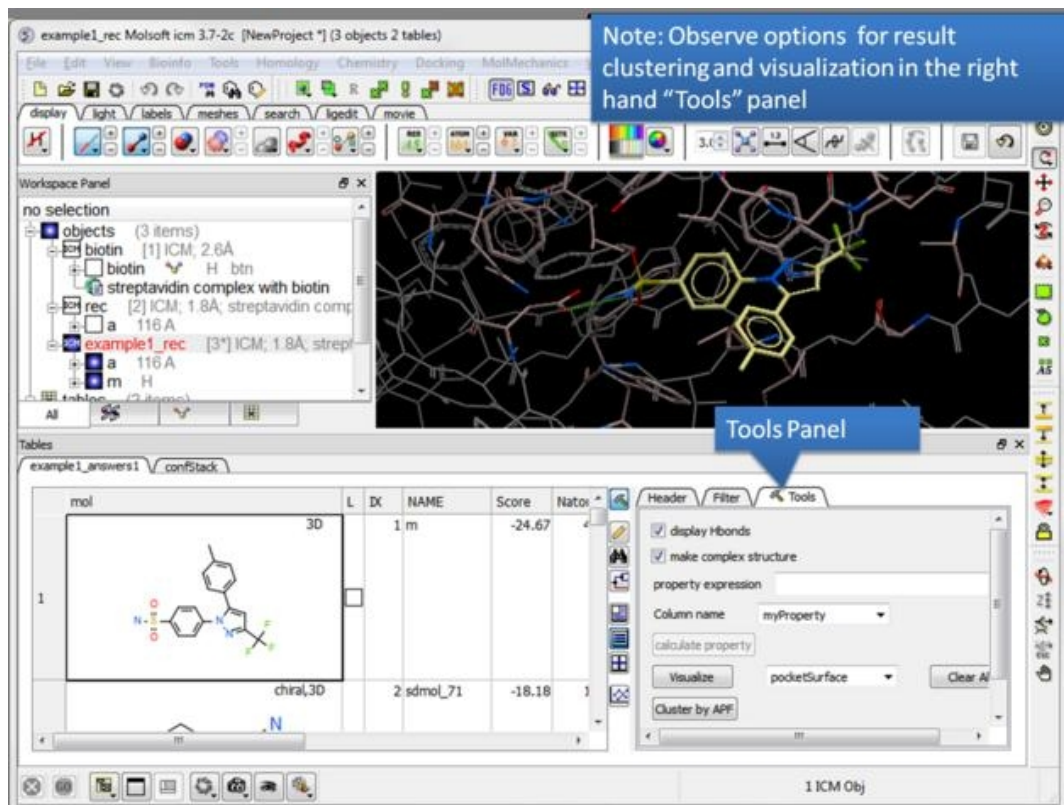
1. Single click allows you to browse stack conformations

2. Double click loads stack conformation and receptor complex into the ICM workspace

i	ener	rmsd	gww	gb
1	1	-67.79	0.25e-006	-53.98
2	2	-64.46	1.058	-54.33
3	3	-63.68	1	-54.45
4	4	-62.06	6.219	-50.98
5	5	-60.08	6.219	-50.47
6	6	-60		-50.06
7	7	-56.84		-50.06
8	8	-55.55		-50.06
9	9	-55.55		-50.06
10	10	-55.55		-50.06
11	11	-55.55		-50.06
12	12	-53.15	5.668	-47.03

1 ICM Obj

Browse stack conformations. Single click allows you to browse (or use up/down arrow keys) and double click generates complex in the ICM workspace.

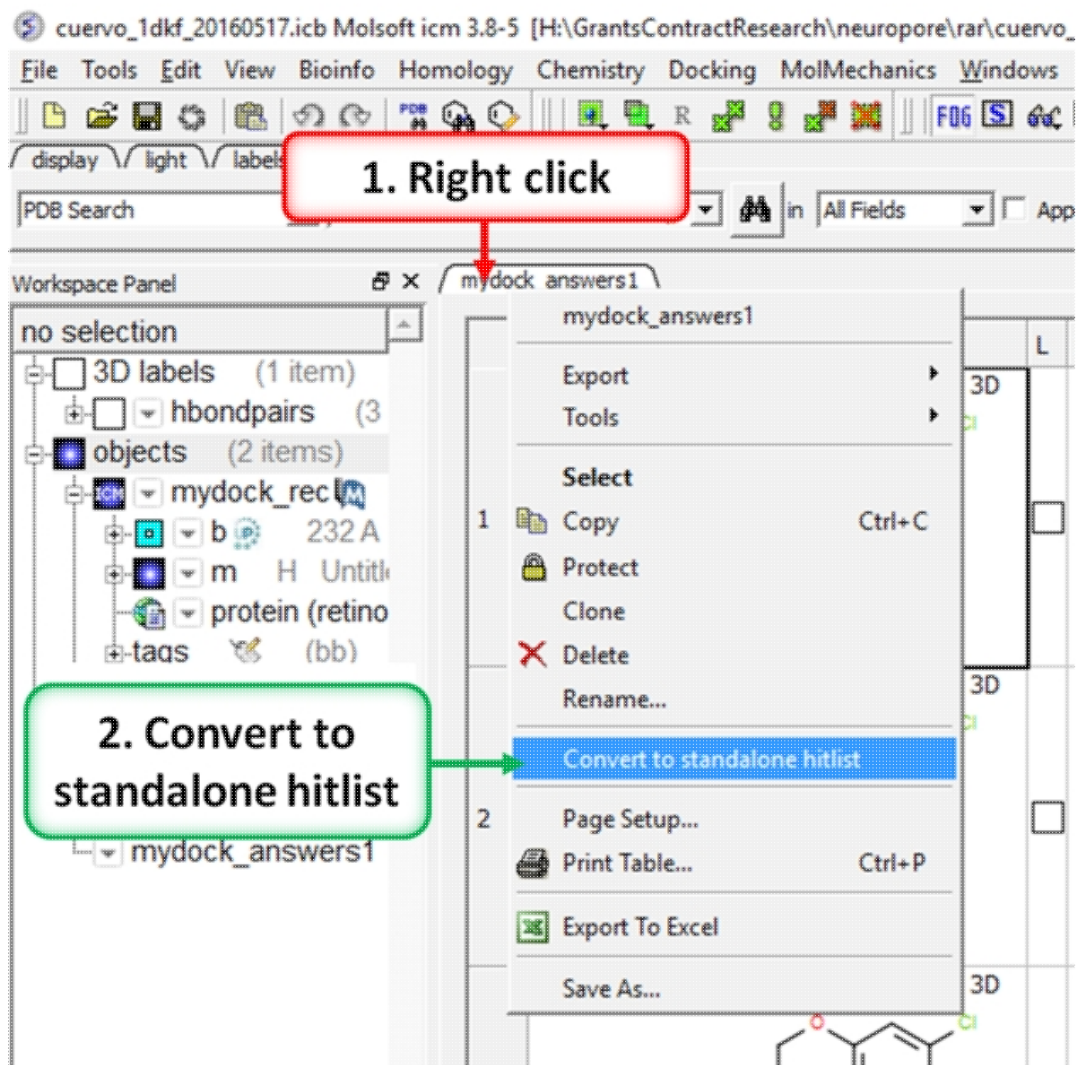


Display stack of alternative ligand conformations. Double click on a row generates a new stack conformation table.

11.1.14 Hitlist Export

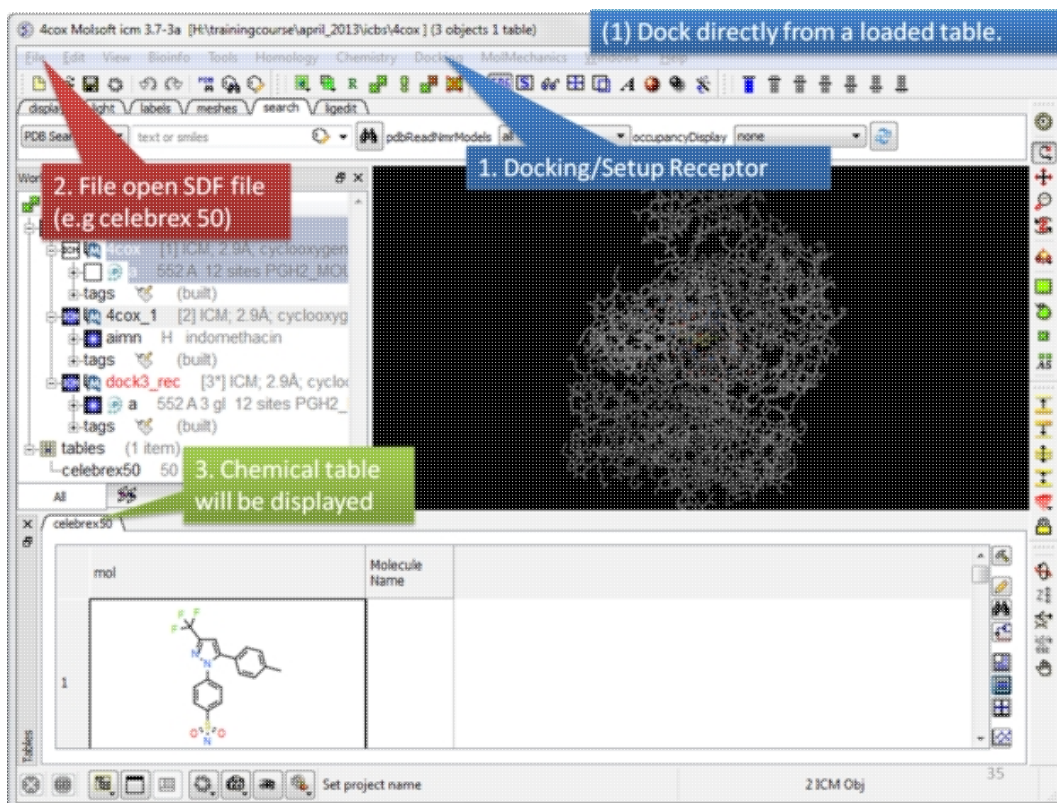
A docking hitlist can be made into a standalone docking project file to share with others and open in locations away from the docking directory. This can be done by:

- Make the docking hitlist Docking/Make Hitlist.
- Right click on the hitlist header tab (see image below).
- Select Convert To Standalone Hitlist
- File/ Save Project and save the project as an .icb file that others can view using ICM or the free ICM-Browser.

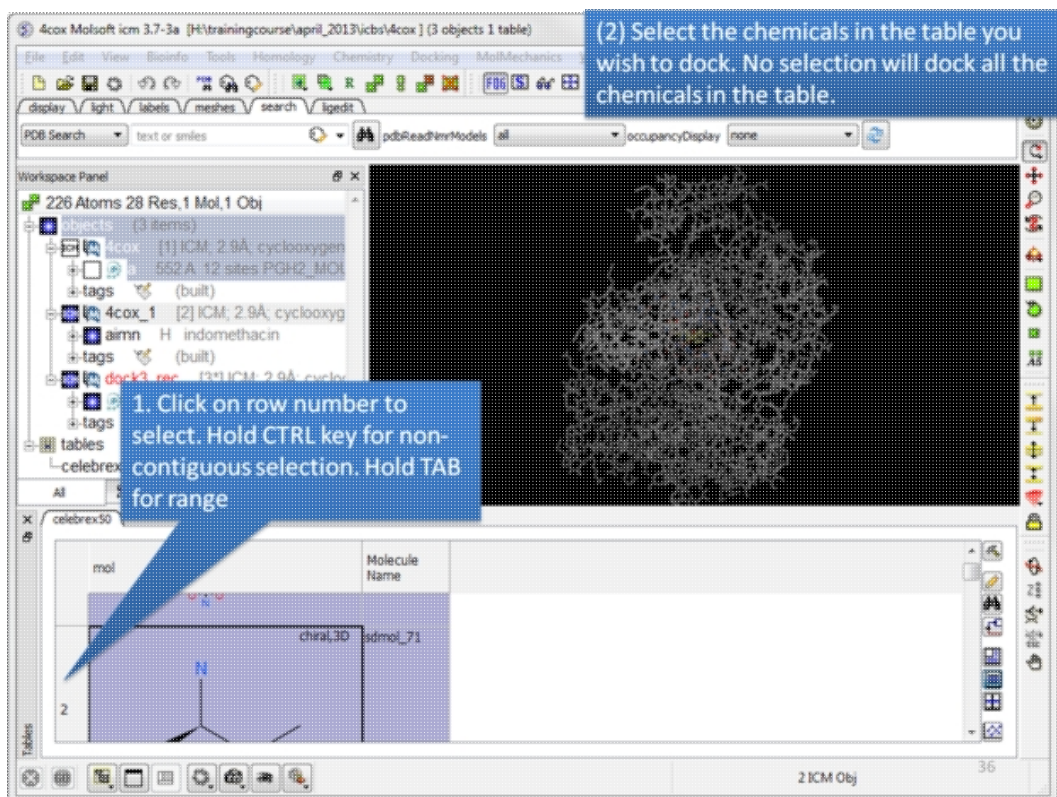


11.1.15 Dock Chemical Table

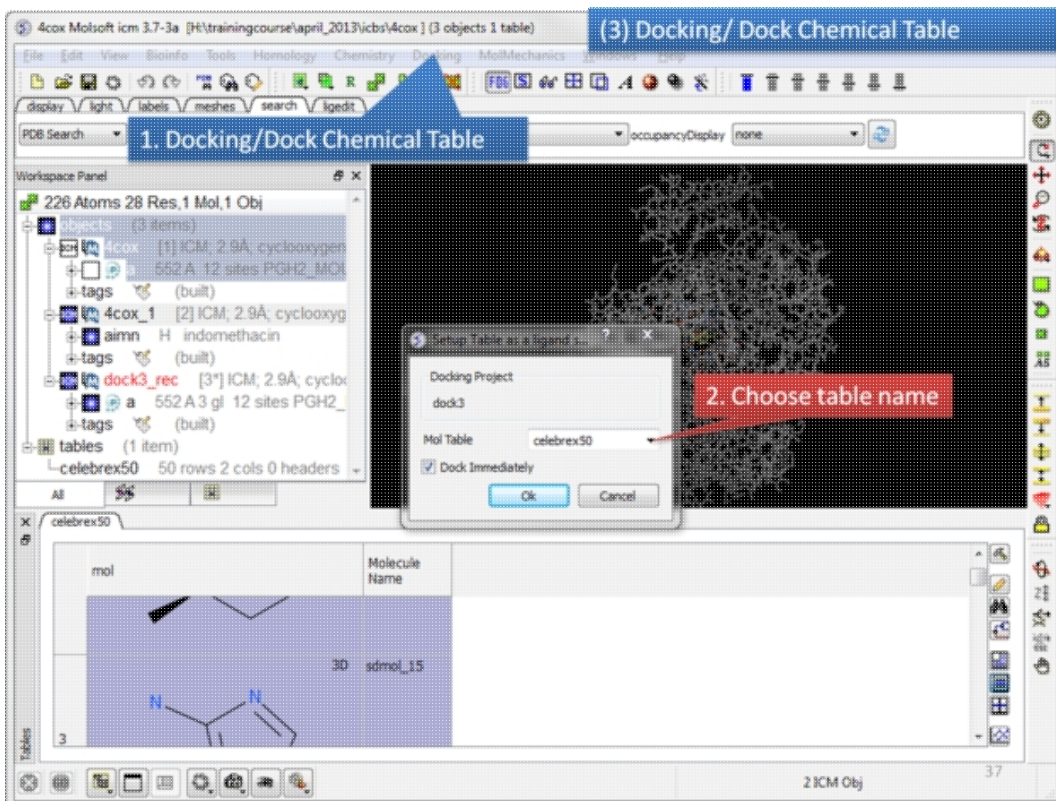
To dock directly from a chemical table:



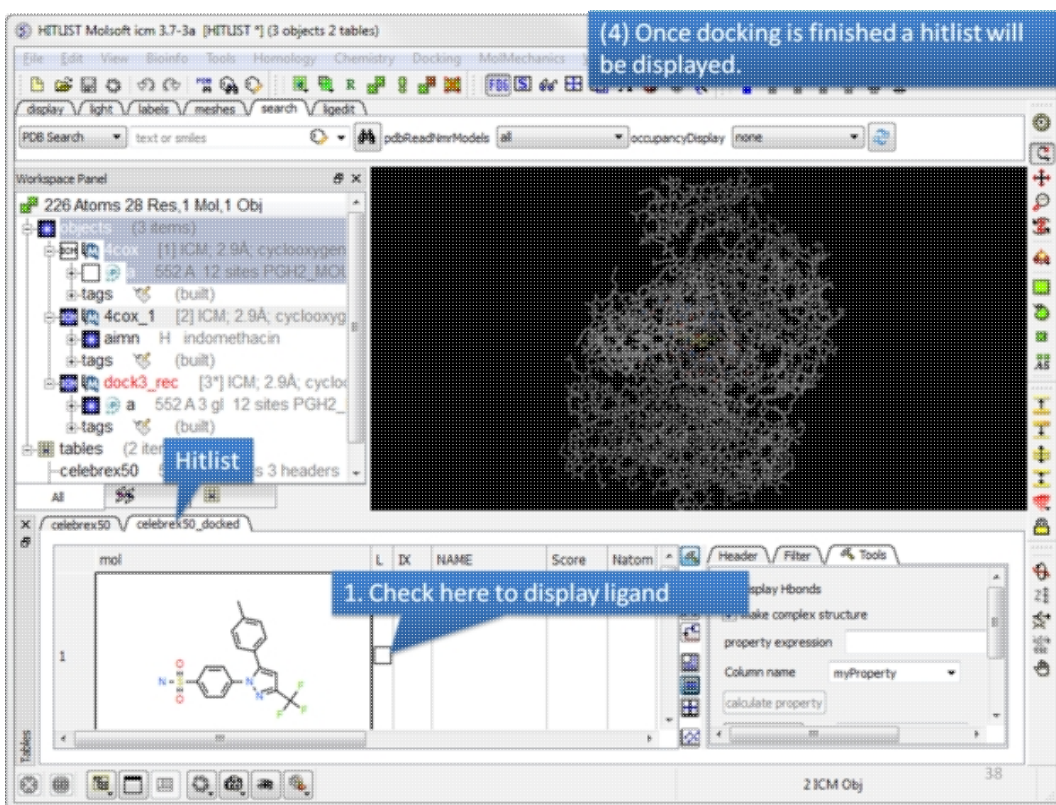
Setup the receptor and read in the chemical table.



Select the chemicals in the table you wish to dock. If no selection is made ICM will dock the whole table. You can make selections by clicking on the row numbers whilst holding CTRL (non-contiguous selection) or TAB (contiguous selection).



Click on the Docking menu and select Dock Chemical Table.



Once the docking has finished a hitlist will be displayed.

11.1.16 Reload a Docking Project

To reload a docking project in version 3.7-2 and earlier.

/Docking/Set Project - Type in the Docking Project Name (Case Sensitive)

or in versions above 3.7-2

/Docking/Open Project

Now you can browse scan solutions etc.... and use the maps to dock another ligand.

11.2 Induced Fit (Flexible) Docking

Available in the following product(s): ICM-Pro | ICM-VLS |

ICM contains a selection of tools to account for pocket flexibility (induced fit) in docking and virtual screening. You can read more about the methods here: <http://molsoft.com/induced-fit.html>

Currently there are four ways to incorporate induced fit into docking:

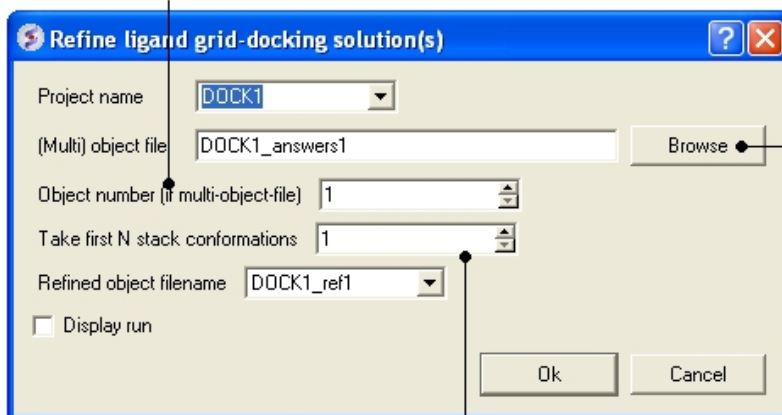
- **Ligand and receptor complex refinement.** This approach allows the side-chains of the receptor to be fully flexible.
- **Multiple receptor conformation docking.** Multiple receptor conformations can be obtained from modeling alternative conformations or from an ensemble of crystal structures.
- **Explicit group docking** allows certain defined residues to be flexible during the docking procedure. For example the hydroxyls of Ser, Thr, and Tyr can be treated explicitly during docking and the rest of the atoms will be treated as maps.
- The SCARE method allows you to take into account induced fit of the receptor upon ligand docking. The method systematically scans pairs of neighboring side-chains and replaces them with Alanine and docks a ligand to each of the "gapped" models.

11.2.1 Ligand and Receptor Complex Refinement

To undertake flexible ligand and receptor refinement docking you first need to dock the ligand into the receptor using the procedures described in the previous section entitled *Small Molecule Docking*. The next step is to select the pose from your docking experiment you wish to use for flexible docking.

- Docking/Flexible-Receptor/Refinement and a window as shown below will be displayed.
- Enter the name of the initial docking project.
- Select the answers.ob file from the docking experiment.
- If you docked more than one ligands you need to select which ligand you want to use for the flexible receptor docking. For example if you want to refine the 3rd ligand in your docked database you would enter 3 in the **Object number** data entry box.
- Select which member of the stack of docking solutions you wish to use. In most cases it will be the first member of the stack and therefore you will enter "1".
- Enter a name for the refined complex.
- You can display the run in the graphical display but this will slow the process down.

Select which ligand from your docking run you would like to use for flexible receptor docking



Select _answers.ob file

Select which member of the stack of conformations (*.cnf) you would like to use

To view the results: The refinement procedure creates two new files in your working directory (see Tools/Change Working Directory for location). The files are the object with docking_project_name_ref1.ob and a stack of conformations called docking_project_name_ref1.cnf.

- File/Open docking_project_name_ref1.ob
- File/Open docking_project_name_ref1.cnf
- Select MolMechanics/Stack View and a table of conformations will be displayed.
- Click on each row to load the conformation and view it.

11.2.2 Multiple Receptor 4D Docking

Additional Resources: Tutorial: 4D Docking to Multiple Conformations of a Flexible Loop in the Aldose Reductase Ligand Binding Pocket

To dock to more than one conformation of a receptor (4D docking) you must first generate a stack of conformations. Your starting point might be one of two scenarios:

1. **Ideal Scenario** - you have an ensemble of crystal structures of the same protein.
2. You have a single crystal structure and you wish to generate an ensemble of possible conformations of a particular region of the binding pocket. See the Aldose Reductase example which contains a flexible loop in the ligand binding pocket.

How to generate a stack of conformations for a single crystal structure:

One way to do this is to select a number of side-chains and right click on the graphical selection and choose **Advanced/Optimize Side Chains** or right click on the ligand in the binding pocket and select **Advanced/Optimize Ligand Vicinity**. Or you can model the loop regions using MolMechanics/Sample Loop.

How to generate a stack from an ensemble of crystal structures for Multiple Receptor Docking: The steps to undertake multiple receptor docking (4D docking) are described below. Before you can start multiple receptor docking all your structures need to be in a stack.

1. Convert the Main Object to an ICM object (Right Click / Convert PDB)

Multiple receptors

Ligand from Main Object

Read in the conformations you wish to use into ICM. Choose one to be the main object onto which you will superimpose.

Ligand from main object has been removed – this makes pocket selection easier. But if you do not have a ligand you can choose Pocketfinder or select manually in the receptor setup stage.

Step 1: Read in the conformations of the receptor you would like to dock to. Convert the Main Object to an ICM Object.

1. MolMechanics/Impose Conformation

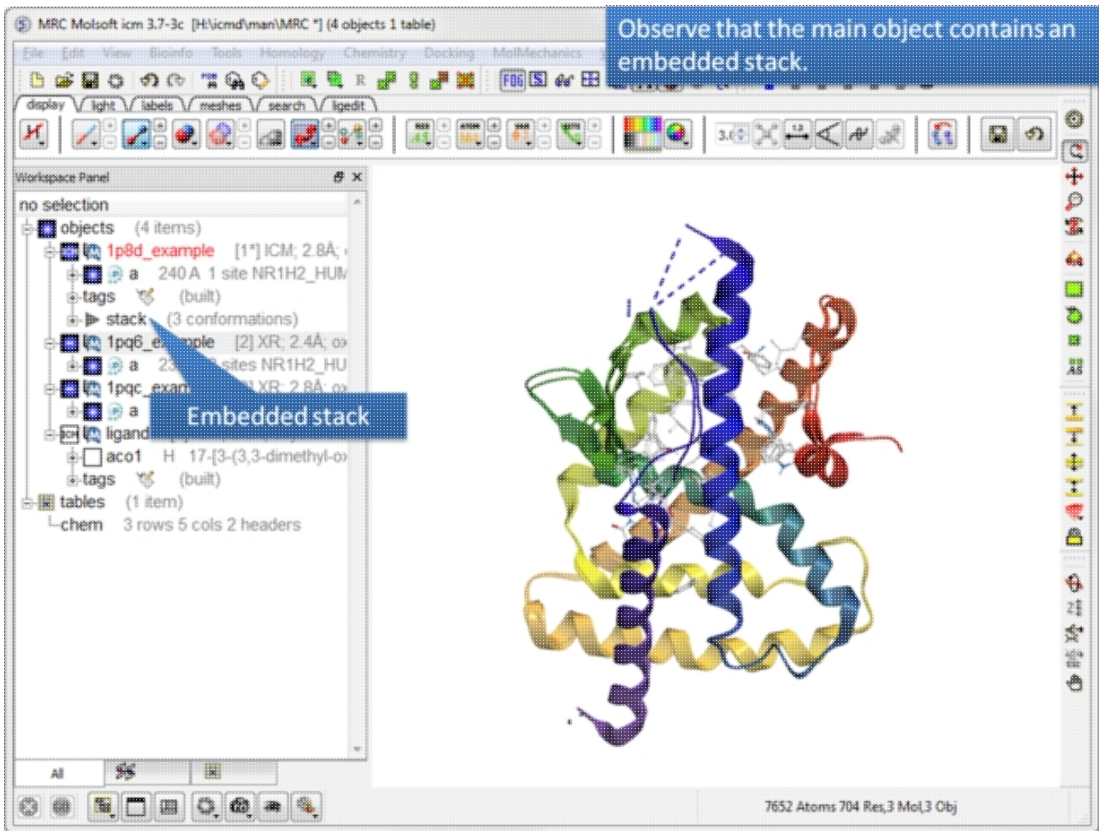
To place all the conformations into a stack – choose MolMechanics/Impose Conformation.

2. Select the other receptors you wish to add to the stack in the main object. Hold CTRL key and click to select

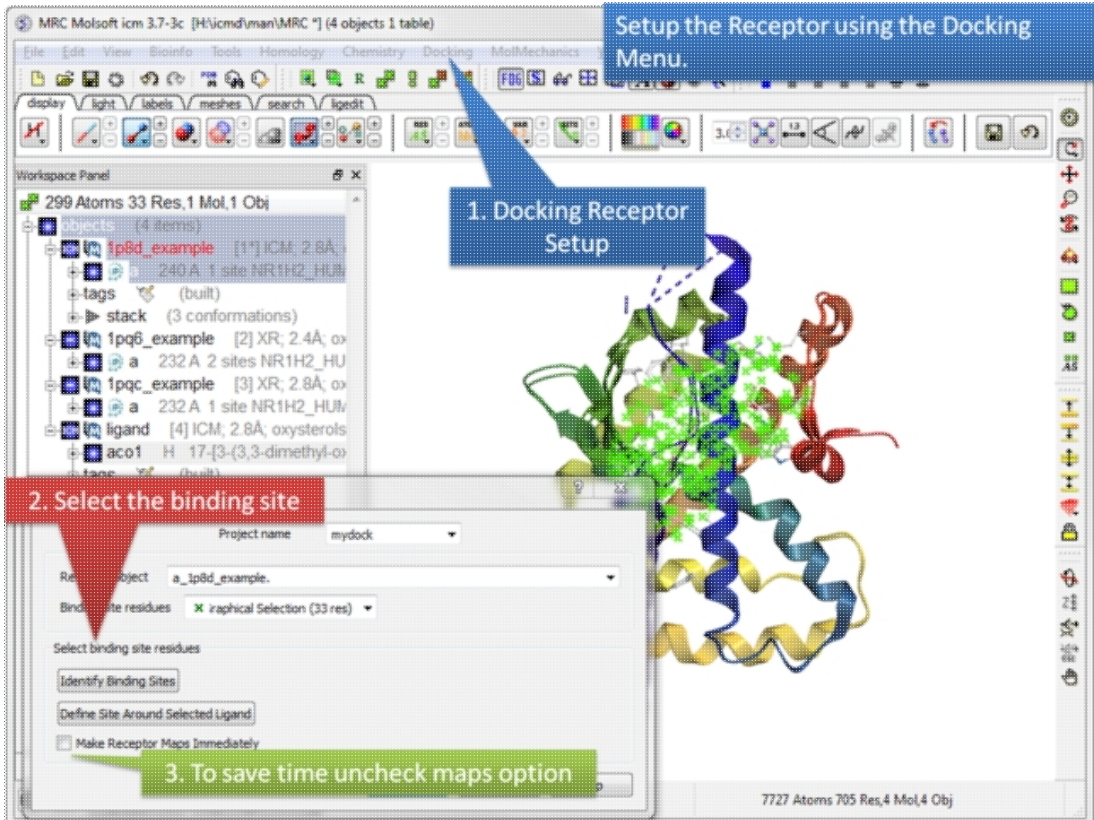
3. Fill in as shown

4. Select superimpose and optimize hydrogens

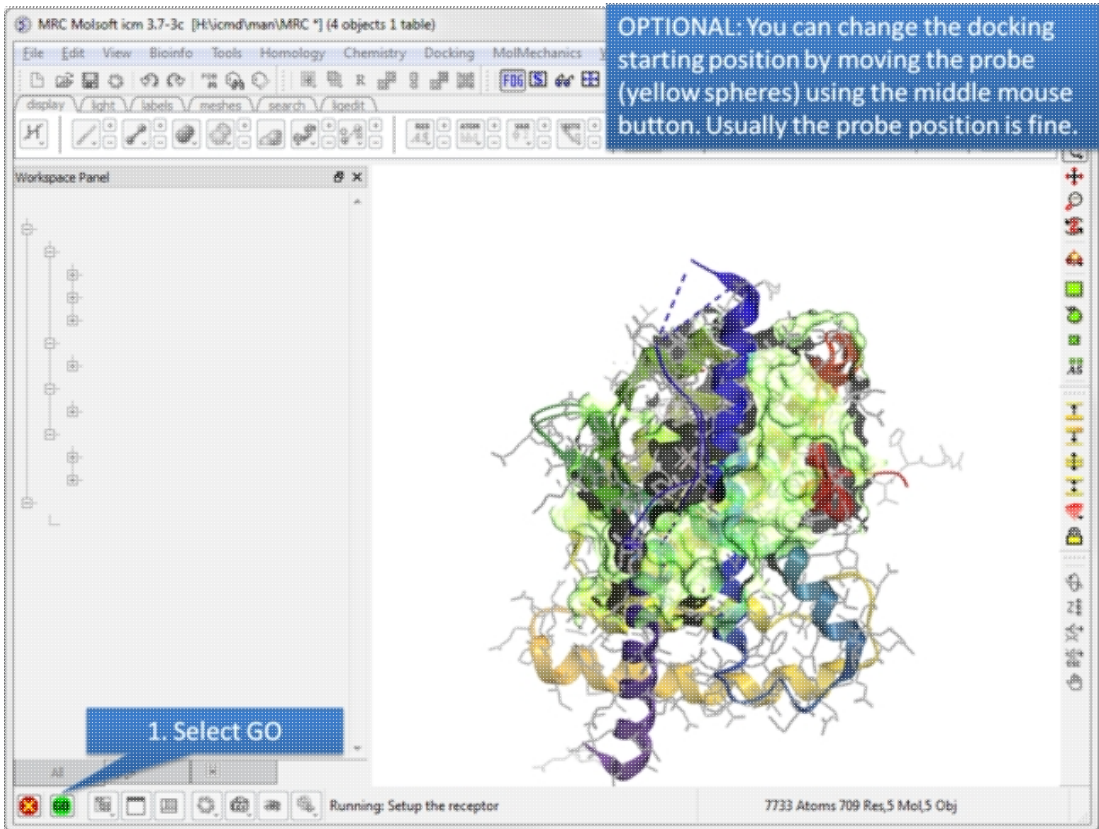
Step 2: Select MolMechanics/Impose Conformation.



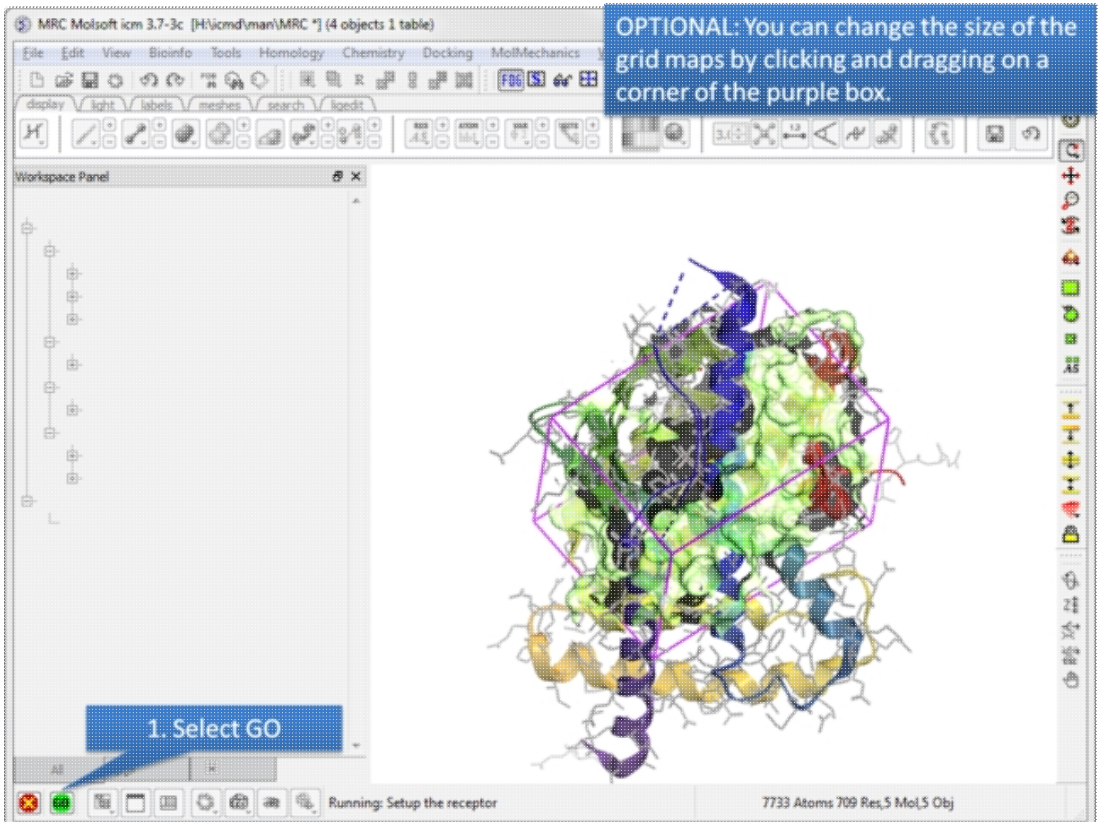
Step 3: Observe the stack is embedded in the main object.



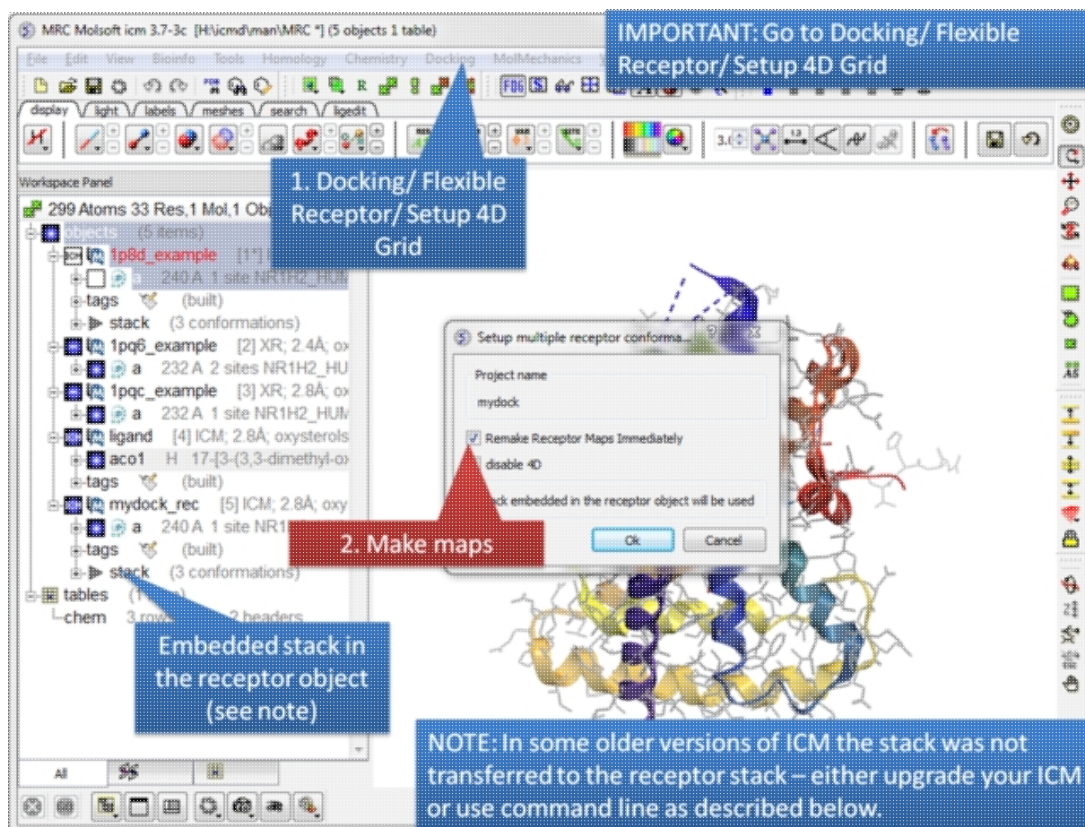
Step 4: Docking/Receptor Setup .



Step 5: Optional: Change simulation starting point.



Step 6: Optional: Change map size.



Step 7: IMPORTANT: Setup the 4D maps Docking/ Flexible Receptor/ Setup 4D Grid.

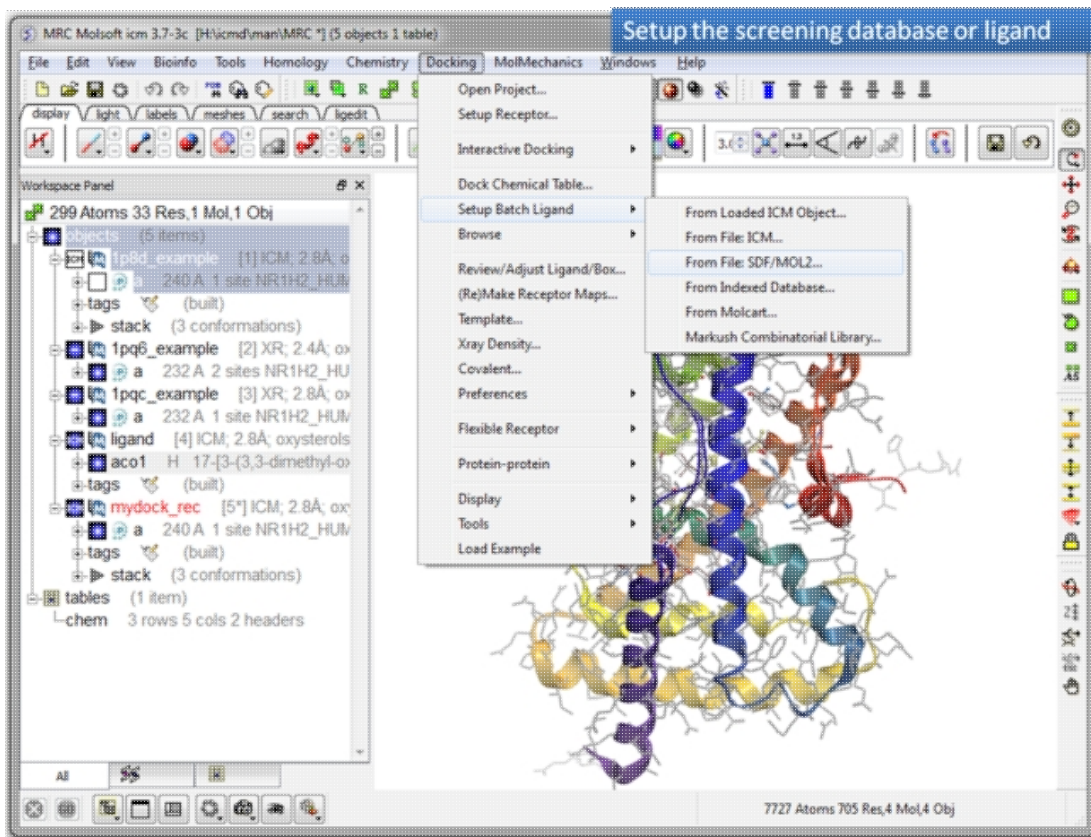
NOTE: In ICM version 3.7-3b and below it was necessary to move the stack from the main object to the receptor object. To do this:

- Move the embedded stack in the original object to a global stack

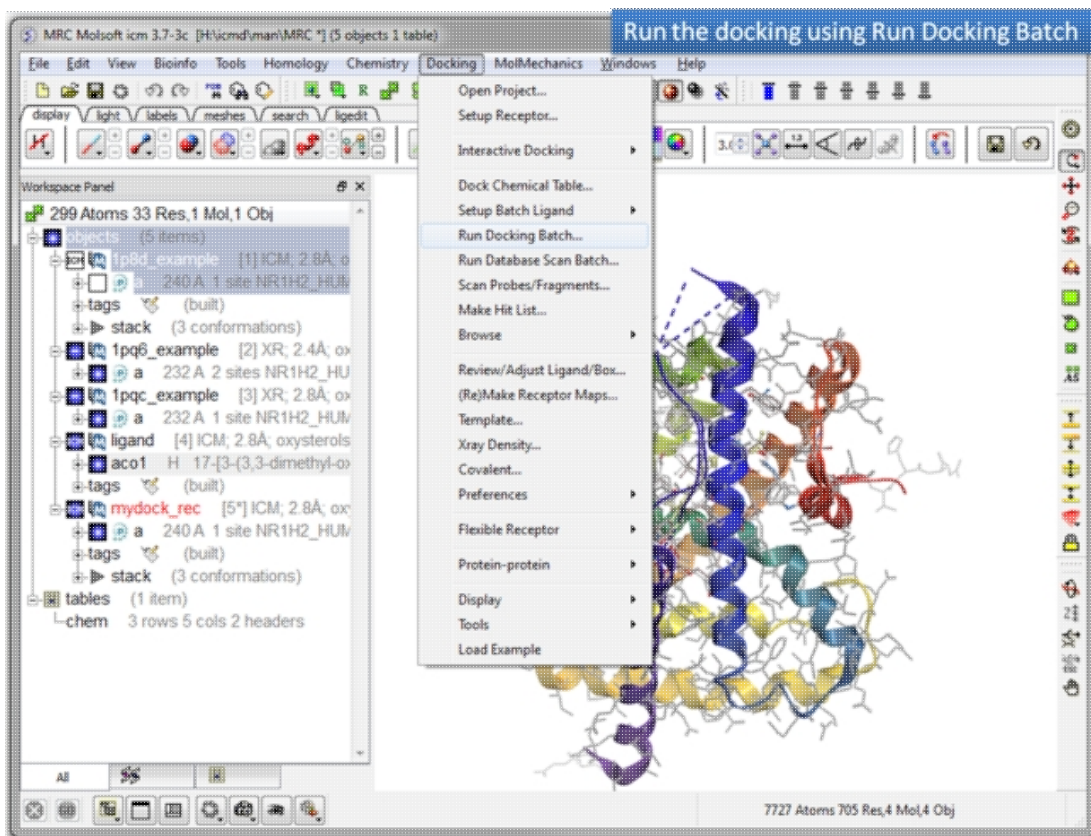
```
load stack a_
```

- Move the global stack into the docking receptor object.

```
store stack a_3. #if the _rec object is the third object.
```



Step 8: Setup the screening database.



Step 9: Run the docking.

When the docking has finished make a hitlist

1. Docking/ Make Hitlist

Column in Hitlist reports which conformation the ligand is docked to

Single click to view confs in 3D

Score	Natom	Nflex	Hbond	Hphob	VwAnt	Ent	Dsol	SoE	...	RecConf
-28.01	40	2	0	-7.775	-32.48	1.407	6.762	5.172	-164.1	2
-12.42	16	2	-3.285	-2.744	-10.71	0.03347	6.665	3.399	-25.32	2

Step 10: Make a hitlist.

11.2.3 Explicit Group Docking

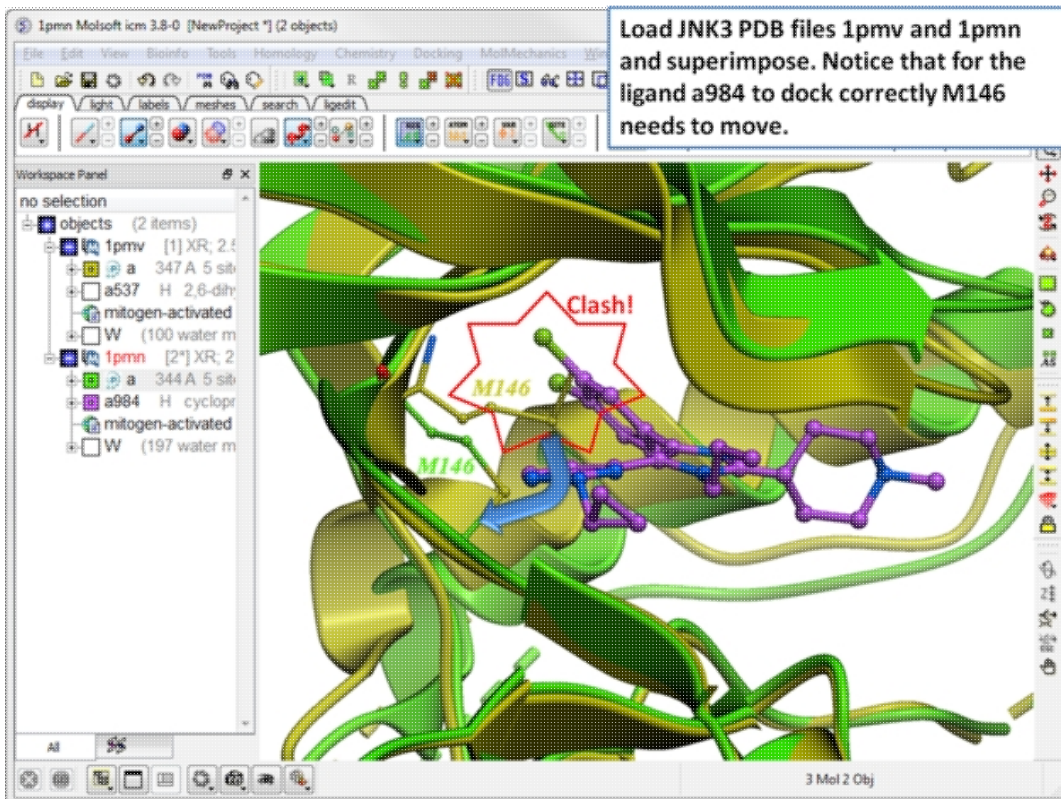
Additional Resources: Tutorial

This option allows you to keep certain residues as explicit during grid docking. For example Hydroxyls of Ser, Thr, and Tyr can be treated explicitly during docking.

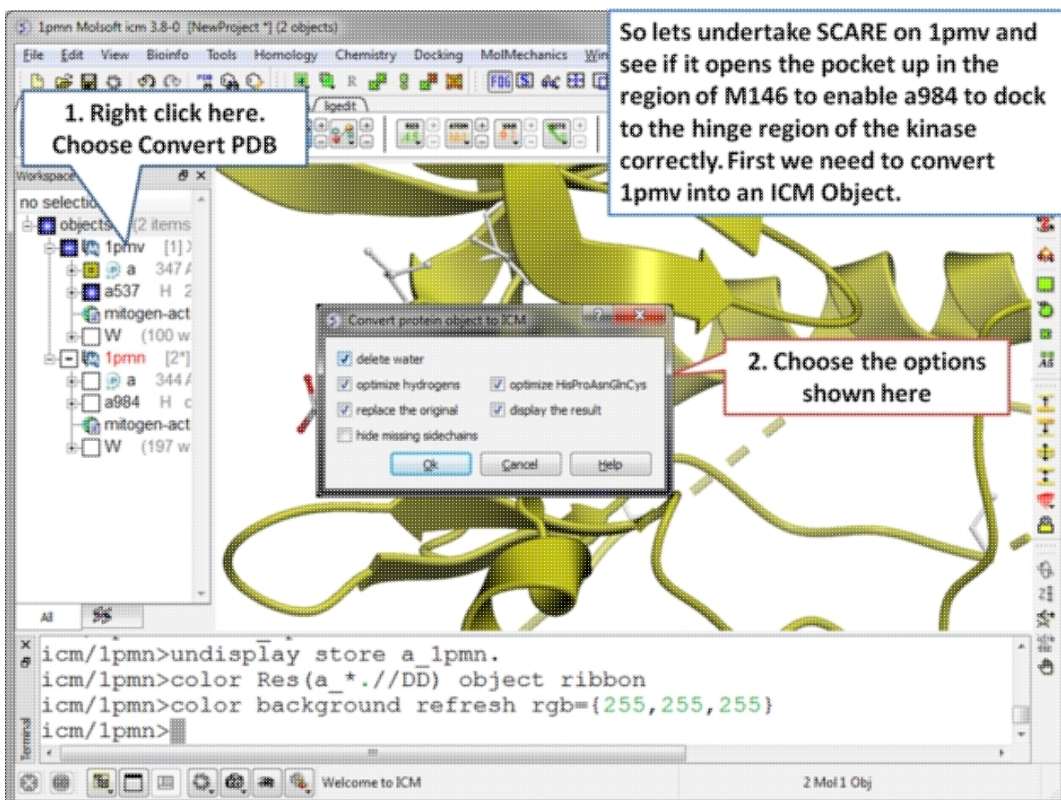
- Setup the docking as described in the Small Molecule Docking section.
- Before making maps choose Docking/Flexible Receptor/Explicit Groups

11.2.4 SCARE - Dual Alanine Scanning and Refinement.

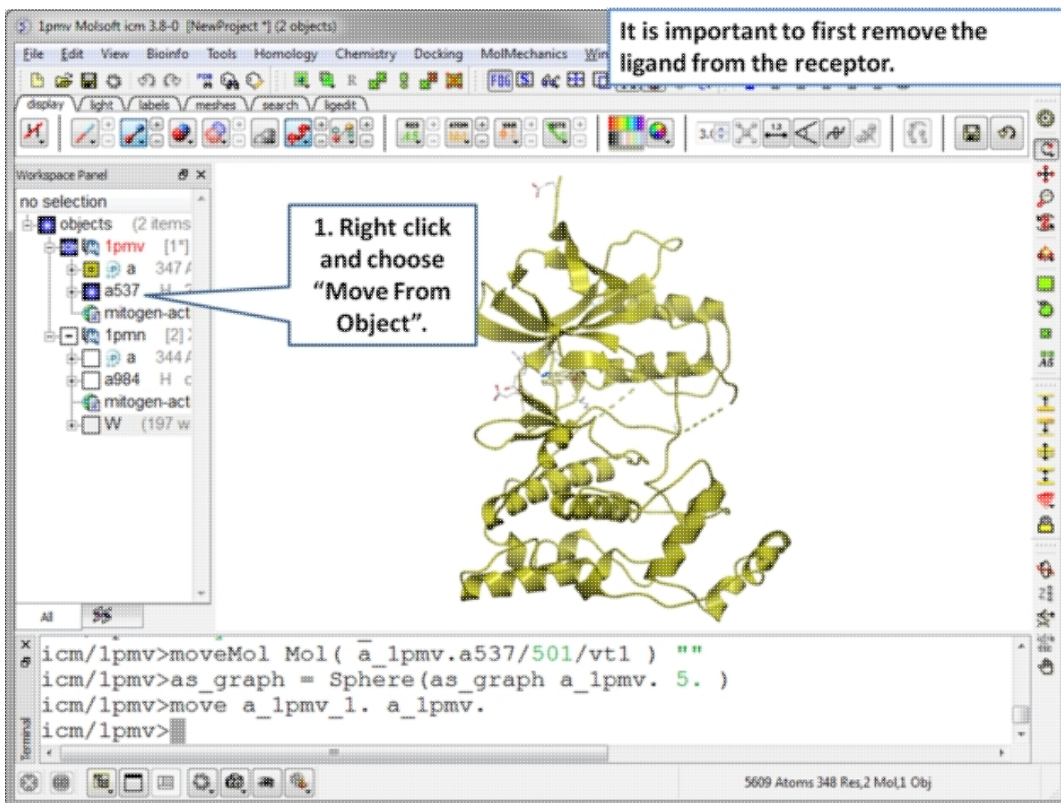
The SCARE method allows you to take into account induced fit of the receptor upon ligand docking. The method systematically scans pairs of neighboring side-chains and replaces them with Alanine and docks a ligand to each of the "gapped" models. A full description of this method and validation is published [here](#). To run the method you will need a single ligand-receptor complex. In the example below we have two JNK3 ligands, one ligand a984 cannot bind at the ATP binding pocket hinge region due to the orientation of M146. The example shows that the pocket can be opened up using SCARE to allow a984 to dock correctly.



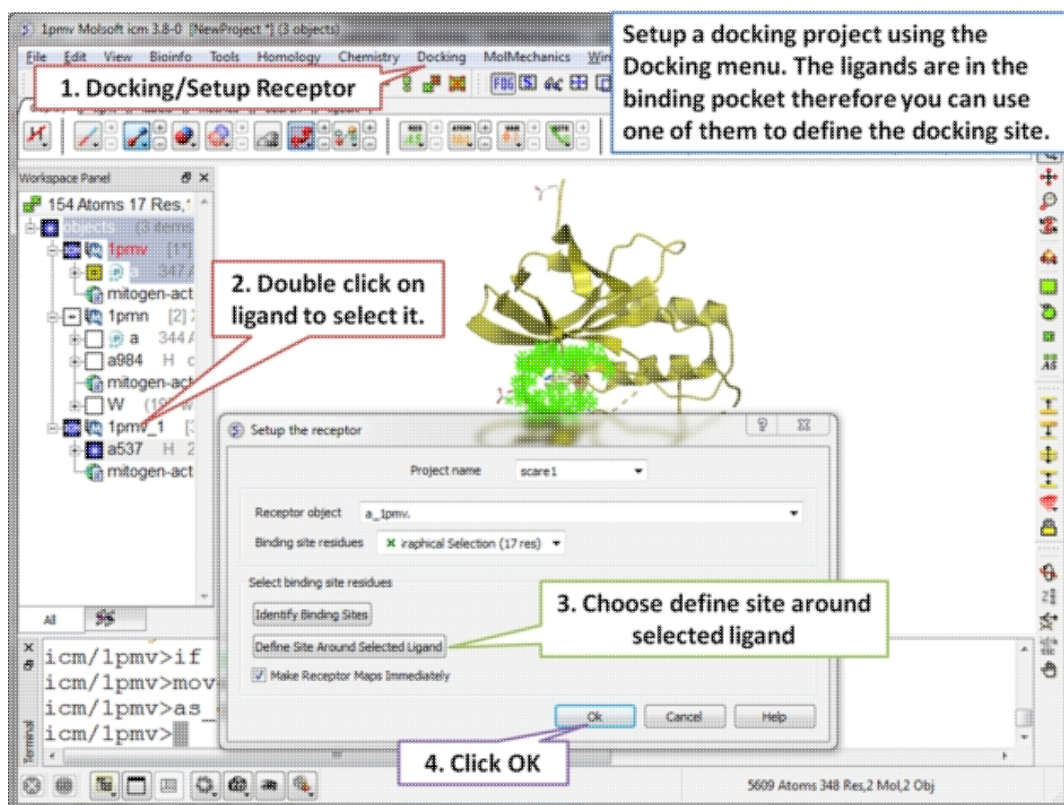
Step 1 In this example we have two JNK3 receptor PDB files. To load them go to the Search tab and enter 1pmv and 1pmn. Select each receptor and then superimpose. Notice that for the ligand a984 in PDB 1pmv to dock correctly M146 needs to move. The SCARE method should be able to open up the pocket so that a984 ligand can dock correctly.



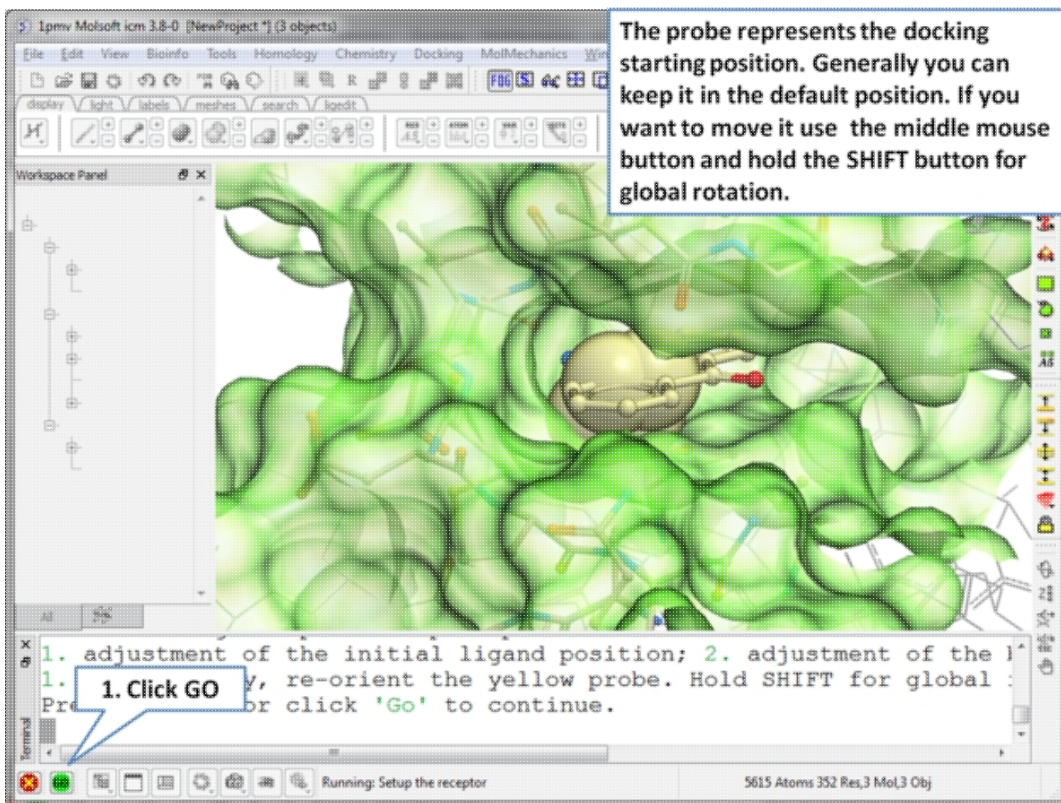
Step 2 Convert 1pmv to an ICM object.



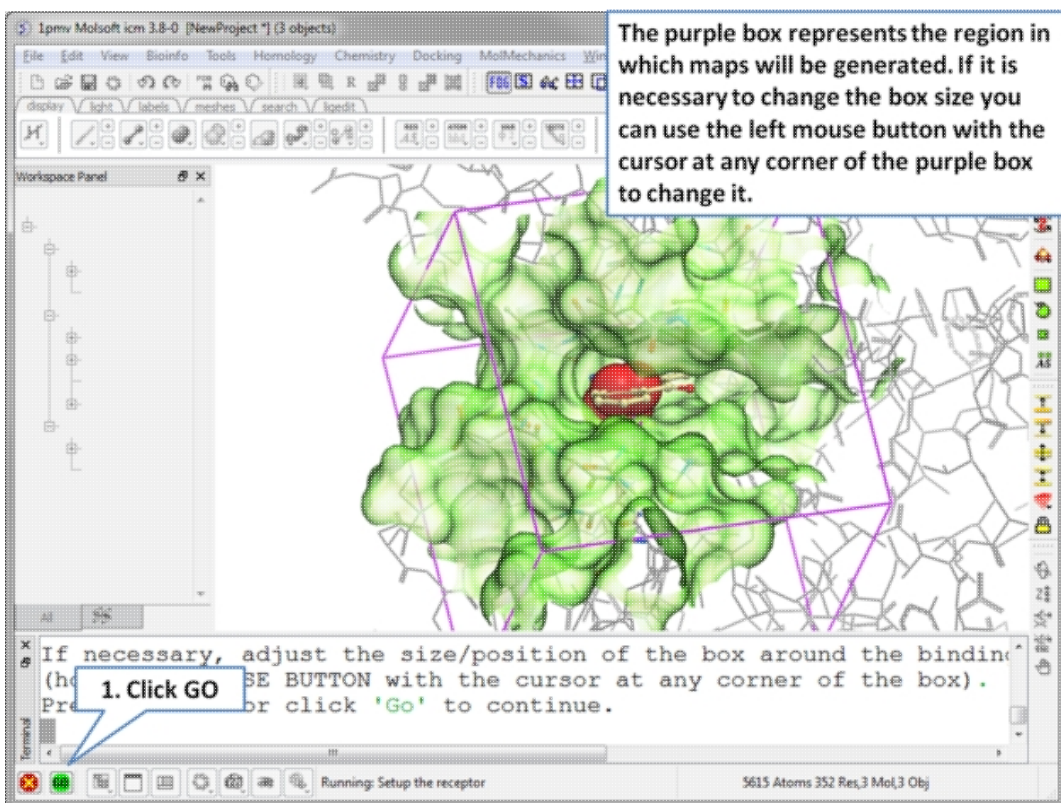
Step 3 Move the ligand out of the 1pmv structure. This prevents the ligand being included in the maps.



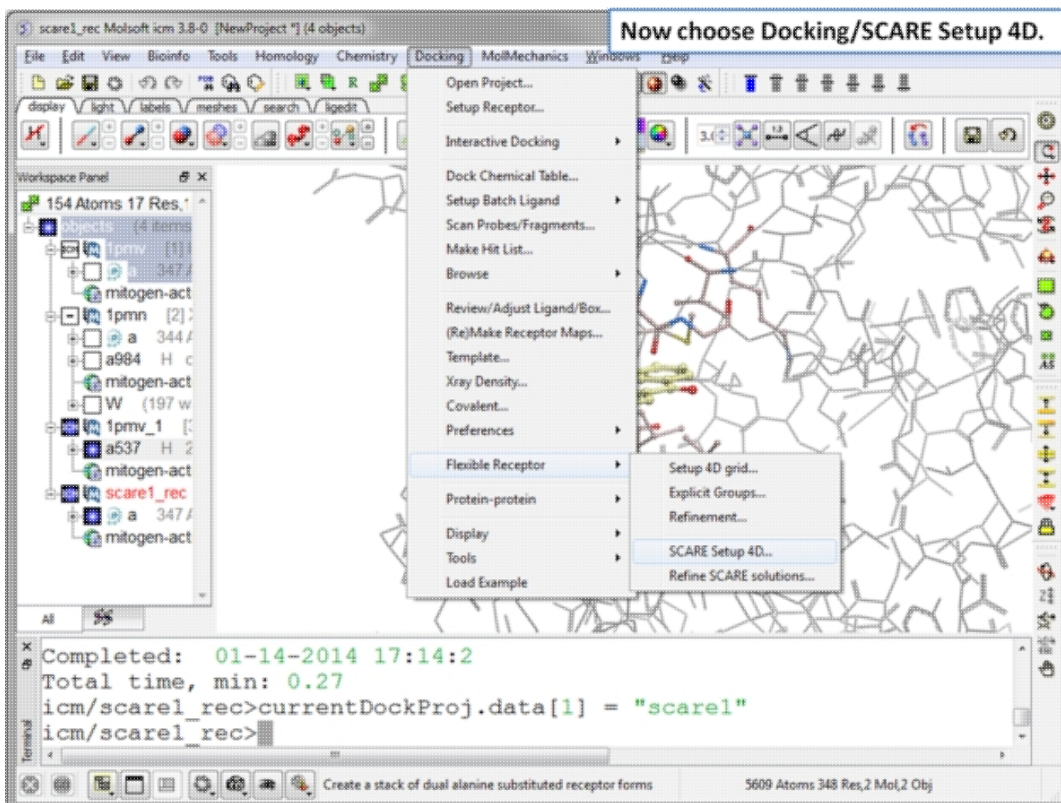
Step 4 Setup the docking project by choosing the menu option Docking/Setup Receptor.



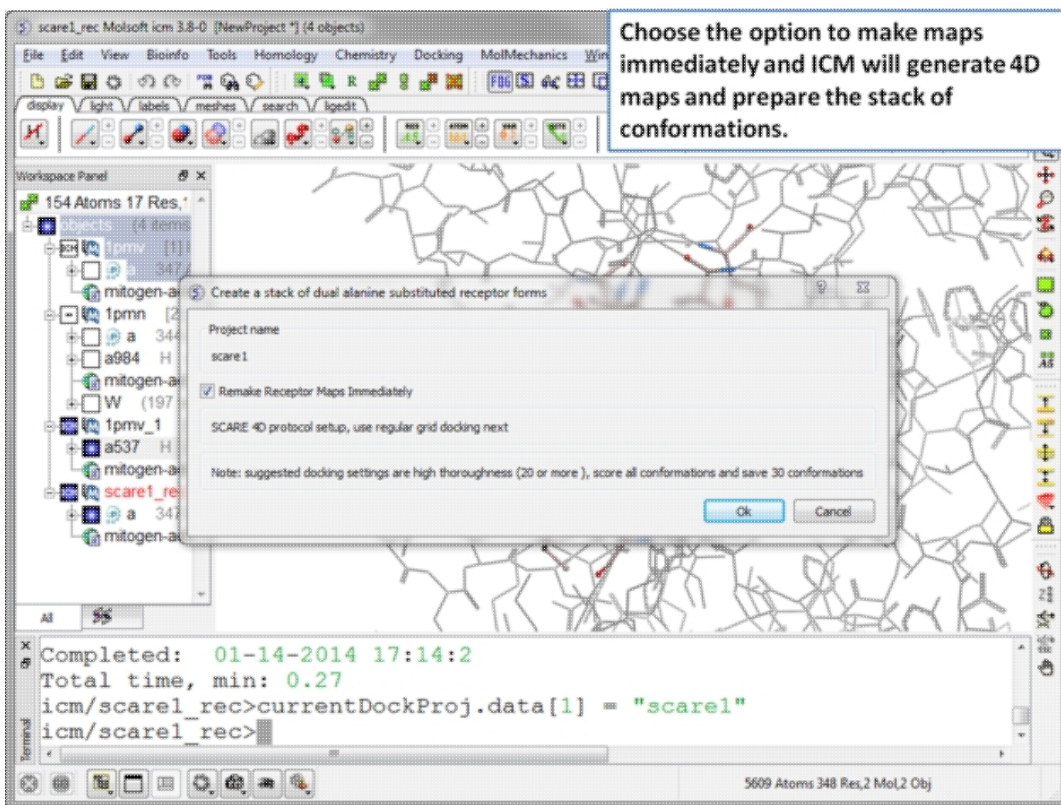
Step 5 The probe represents the starting point for the docking simulation. By default the probe is in the center of the pocket.



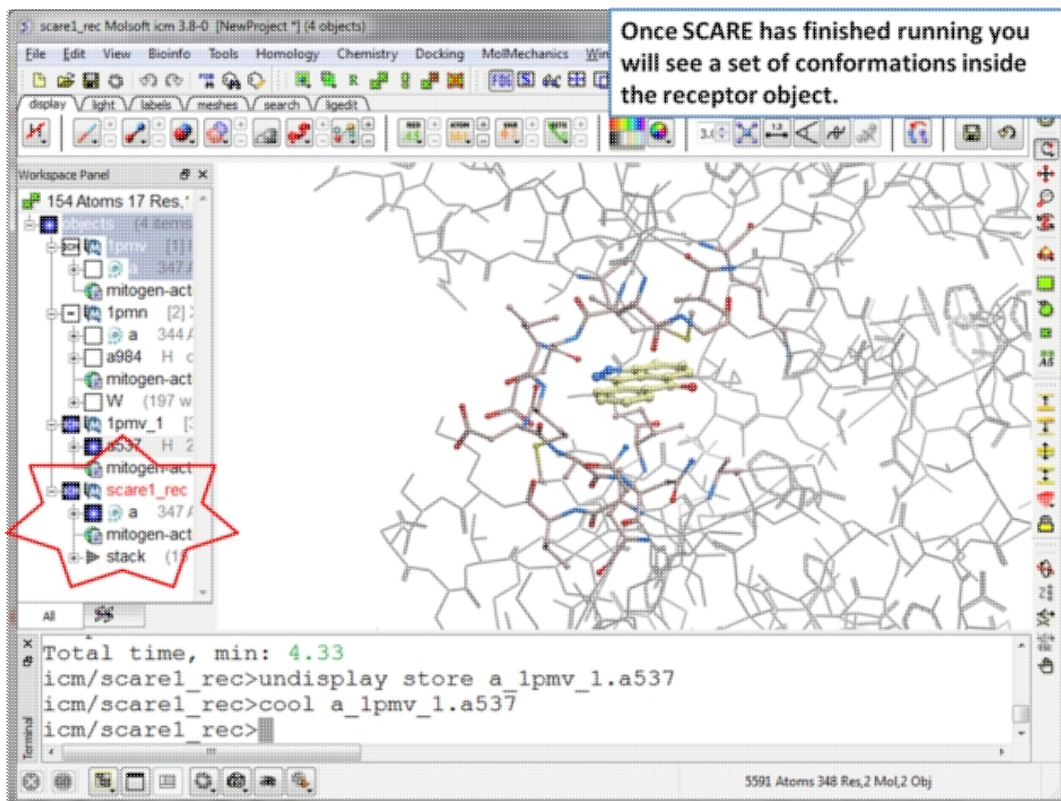
Step 6 The dimensions of the box can be changed. .



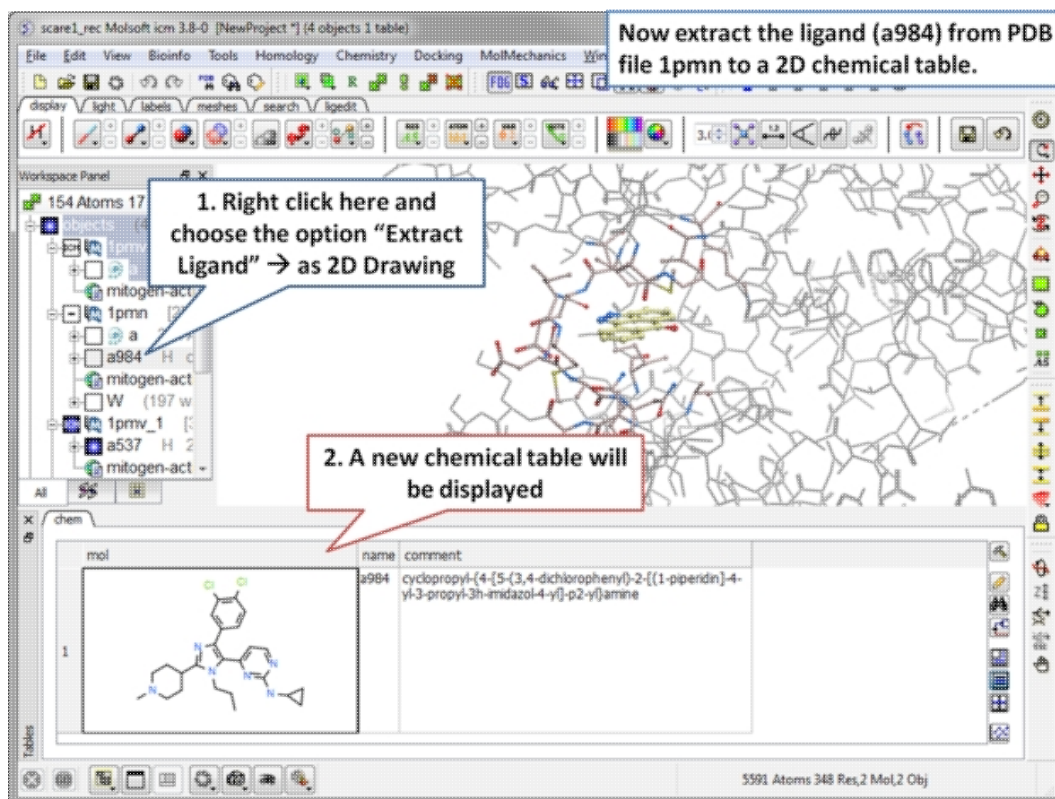
Step 7 Choose the menu option Docking/Flexible Receptor/Scare Setup 4D.



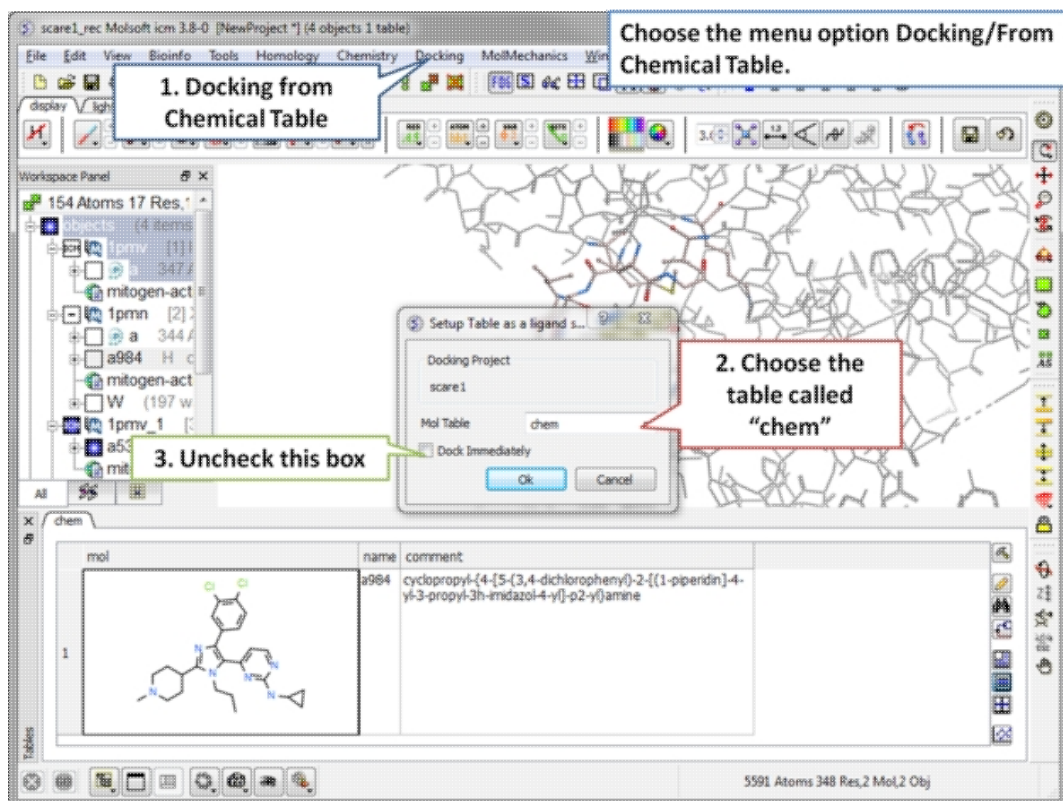
Step 8 A dialog window will be displayed - click OK.



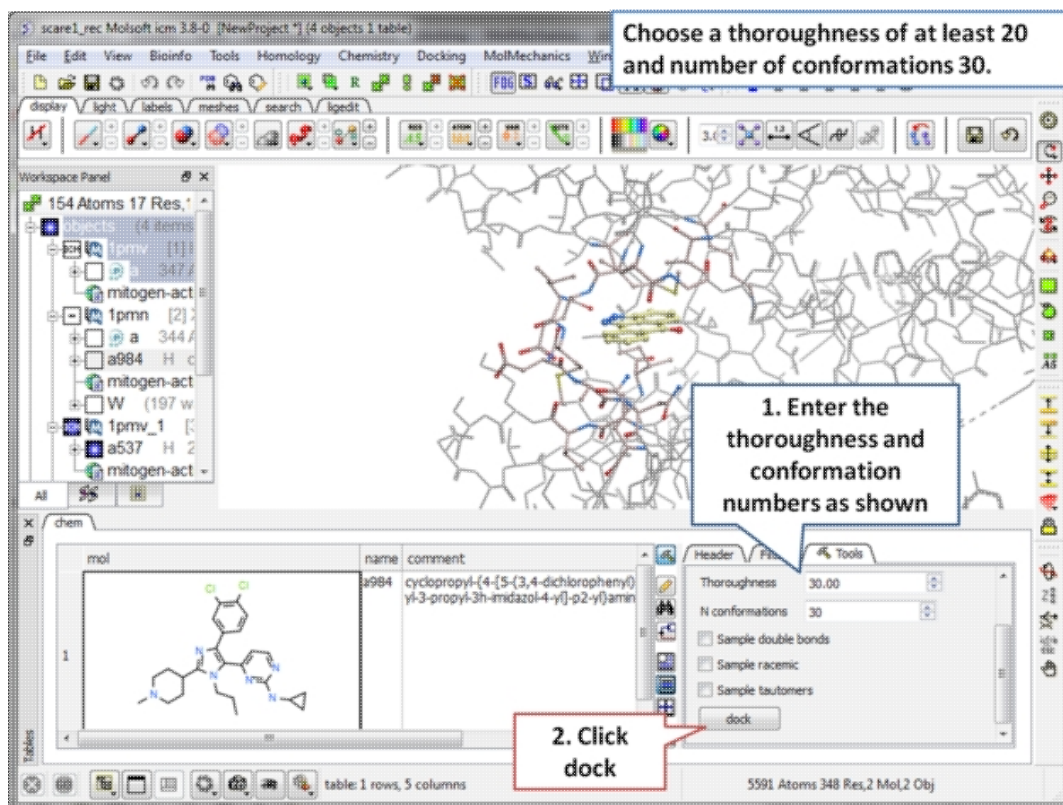
Step 9 Once the SCARE simulation has finished running you will see a stack of receptor conformations in the ICM Workspace.



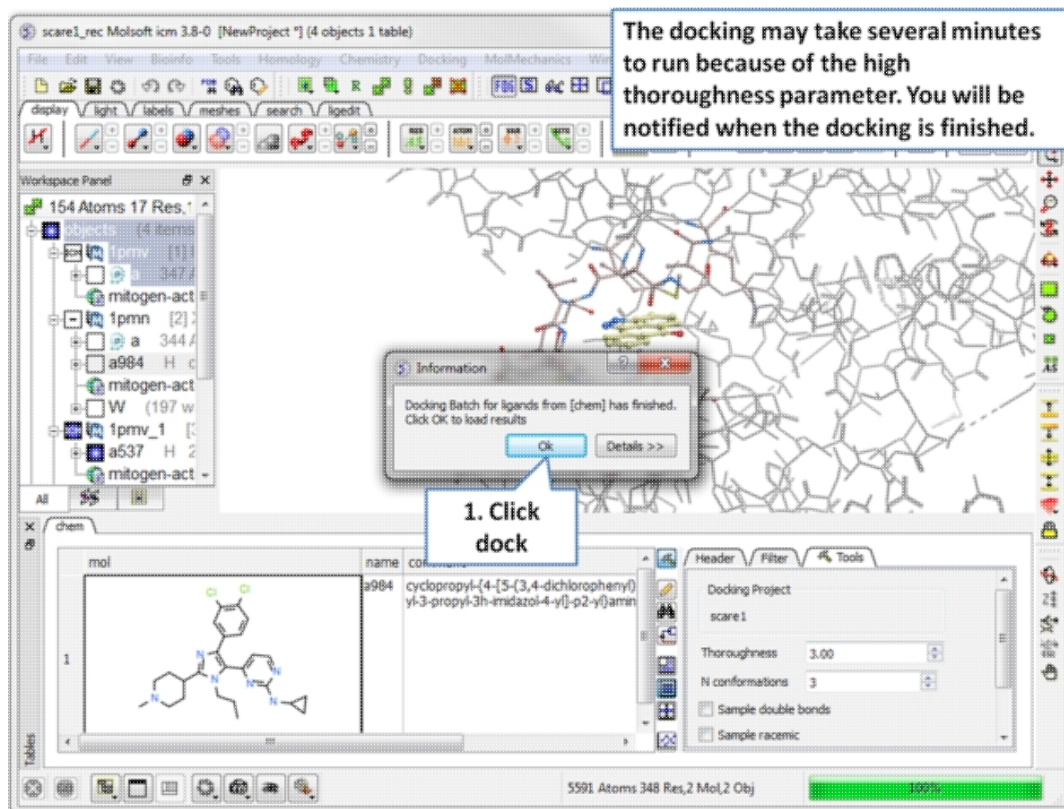
Step 10 Now in preparation for ligand docking we will extract the ligand a984 from the PDB file 1pmv in the ICM Workspace.



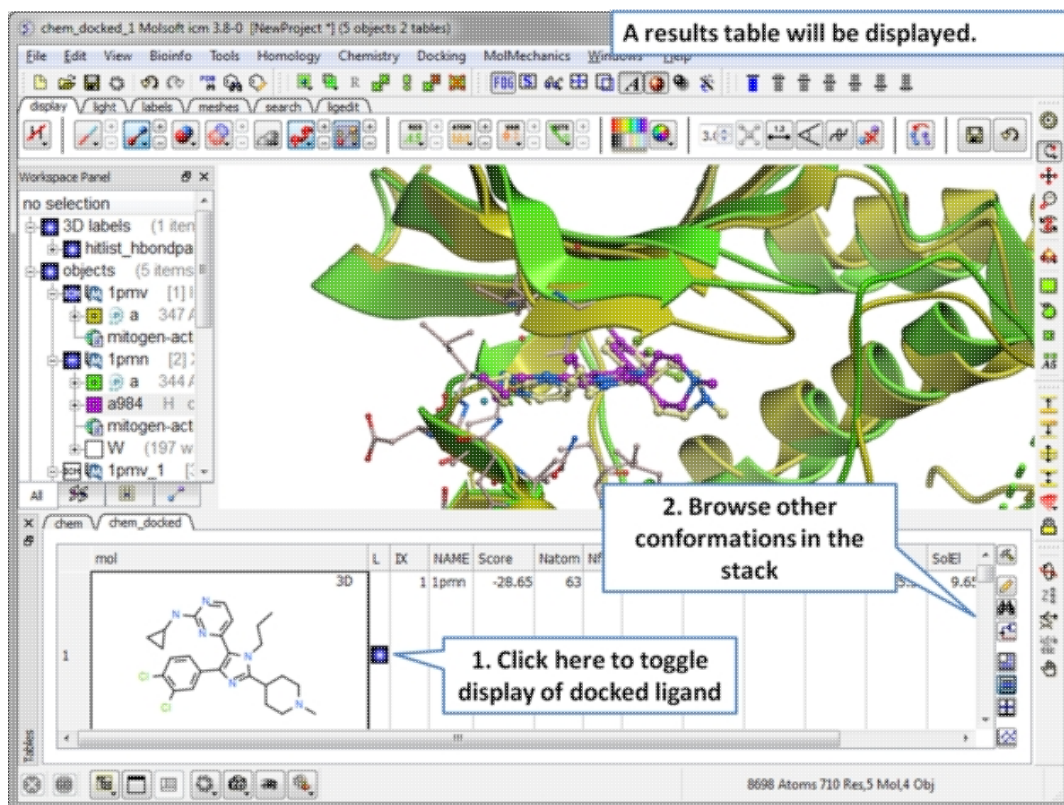
Step 11 The ligand will be docked directly from the table.



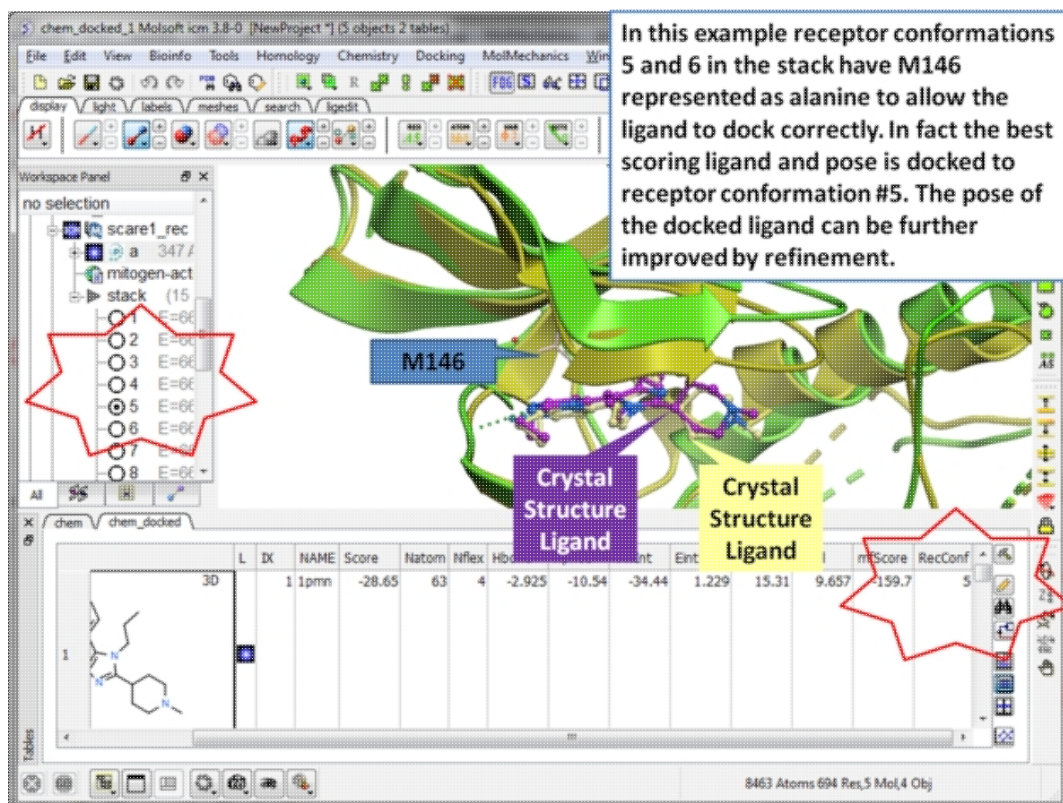
Step 12 Choose a thoroughness of at least 20 and number of conformations 30. .



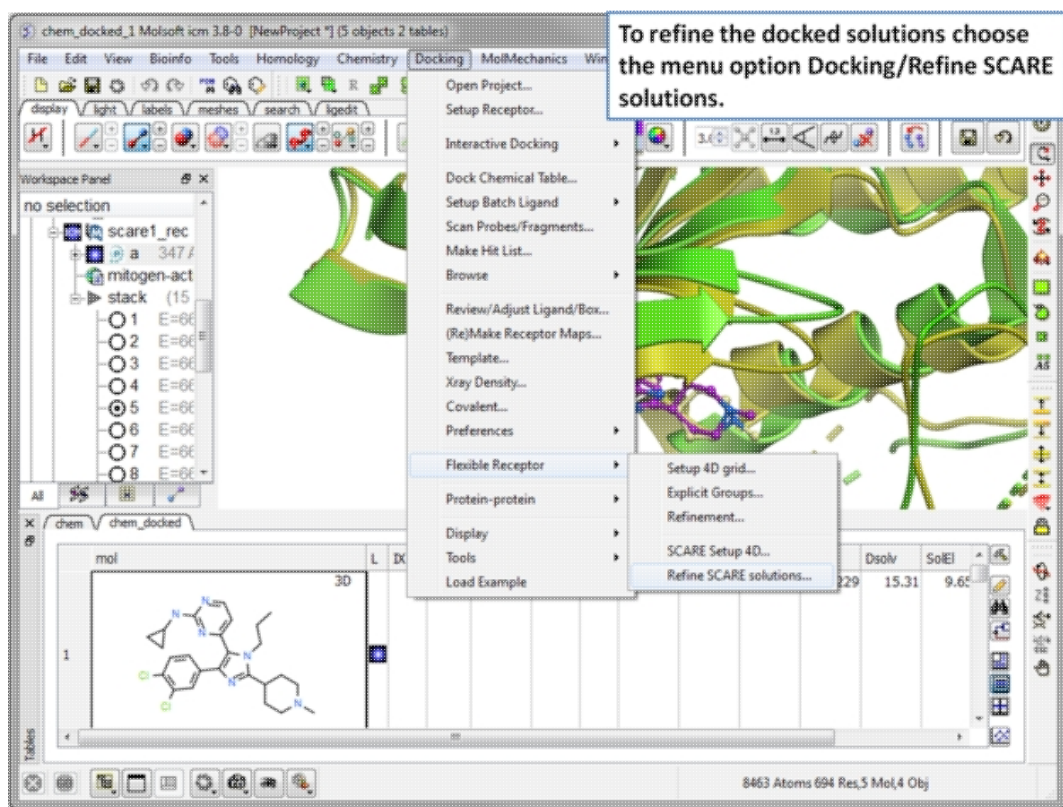
Step 13 Run the docking, it may take a few minutes because the thoroughness value is high.



Step 14 A results table will be displayed at the bottom of the gui when the docking has finished.



Step 15 Toggle through the different conformations and observe the best pose and receptor conformation.



Step 16 The docked ligand usually requires refinement.

The refinement will use all members of the stack. The refinement may take a few hours to run to completion.

1. Browse for the docked object file

2. If more than one ligand was docked in the first step choose which ligand you wish to refine

L	DK	NAME	Score	Natom	Nflex	Hbond	Hphob	VwInt	EntI	Dsolv	SoIEI
1		1prmn	-28.65	63	4	-2.925	-10.54	-34.44	1.229	15.31	9.62

8463 Atoms 694 Res, 5 Mol, 4 Obj

Step 17 The refinement will use all members of the stack. The refinement may take a few hours to run to completion. .

To see the refinement results go to File/Open and load the docking_project_name_refine.ob file.

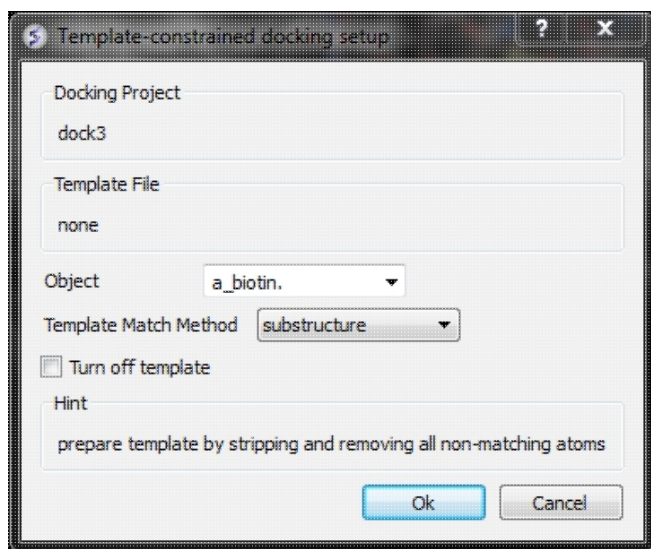
1 ICM Obj

Step 18 To see the refinement results go to File/Open and load the docking_project_name_refine.ob file.

11.3 Template Docking

Available in the following product(s): ICM-Pro | ICM-VLS |

To perform constrained docking to a template structure you need to position the template in the desired position in the receptor - this can be done by docking or using the `connect` option if the template is not already in the correct position. Make sure the template contains only the atoms you wish to match - a template can be edited by right clicking on it in the ICM Workspace and select Edit/ Edit Compound. It is also important to delete hydrogen atoms on the template particularly in places where substituent groups will be placed.



- Setup the docking project as described in the `Small Molecule Docking` chapter.
- Before starting the docking select Docking/Template
- Read in the ICM object that you wish to use as a template. Select the template object name in the Object data entry box using the drop down button.
- Select the template match method. Match by atom name, substructure, fuzzy or APF.
- Click OK
- Set the docking job running - Docking/Interactive or Docking/Run Docking Batch

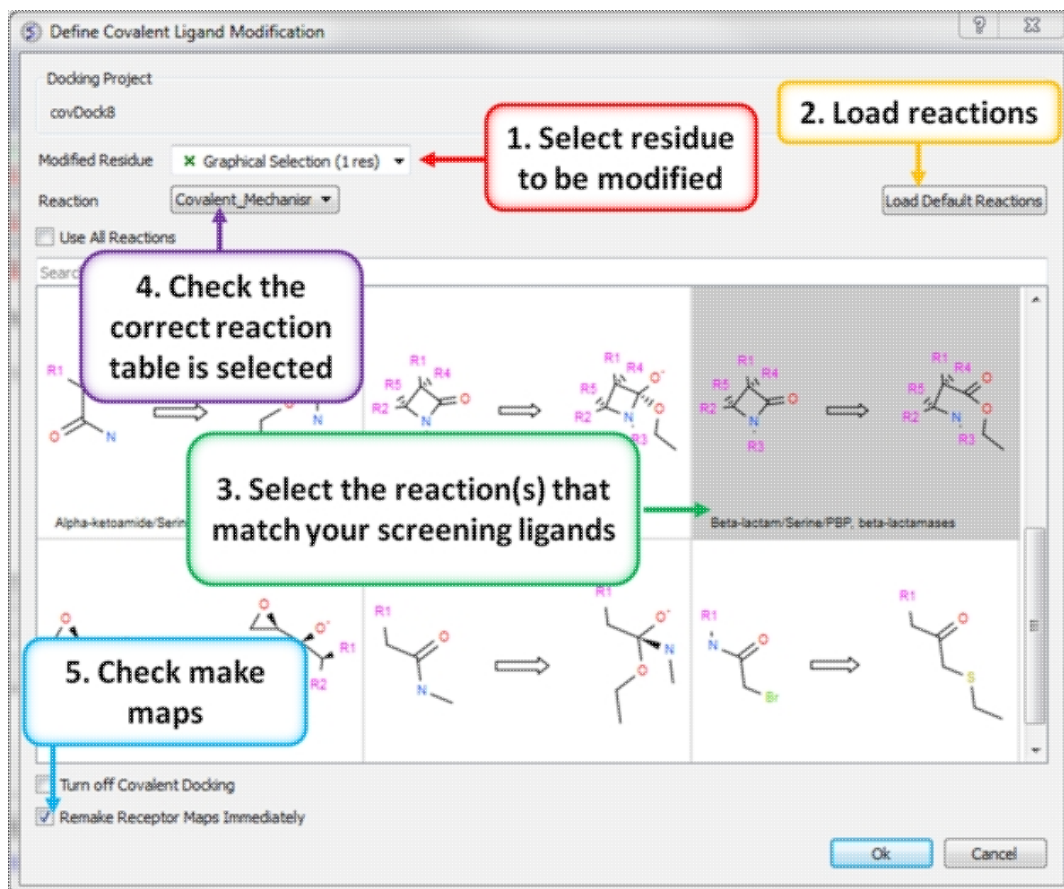
11.4 Covalent Docking

Tutorial | Video Tutorial

Available in the following product(s): ICM-Pro | ICM-VLS |

To undertake covalent docking you should:

- Setup the docking project (Docking/SetProject).
- Setup the receptor (Docking/Setup Receptor) - you can remove the check mark in the "Make Receptor Maps Immediately" option because for covalent docking the maps are made at a later stage.
- Select the residue that will be modified (makes covalent bond) in the docking project receptor (e.g. `mydock_rec`).
- Choose the menu option Docking/Covalent and a dialog window as shown below will be displayed.



- Make sure you have selected the modification residue in the docking project receptor (e.g. mydock_rec).
- You can use the in built reactions by clicking on "Load Default Reactions" button.
- Select the reactions you need by clicking on them or choose "Use All Reactions". You will notice a new chemical table will be displayed called "Covalent Mechanisms".

NOTE: If the reaction you need is not included in our default selection you will need to sketch the reaction as described [here](#). The docked ligands need to match the reactants (e.g. lactam ring).

- Next read in an SDF file containing the chemicals you wish to dock that match the reactant of the reaction.
- To dock the ligands choose `Docking/Dock from Table`.

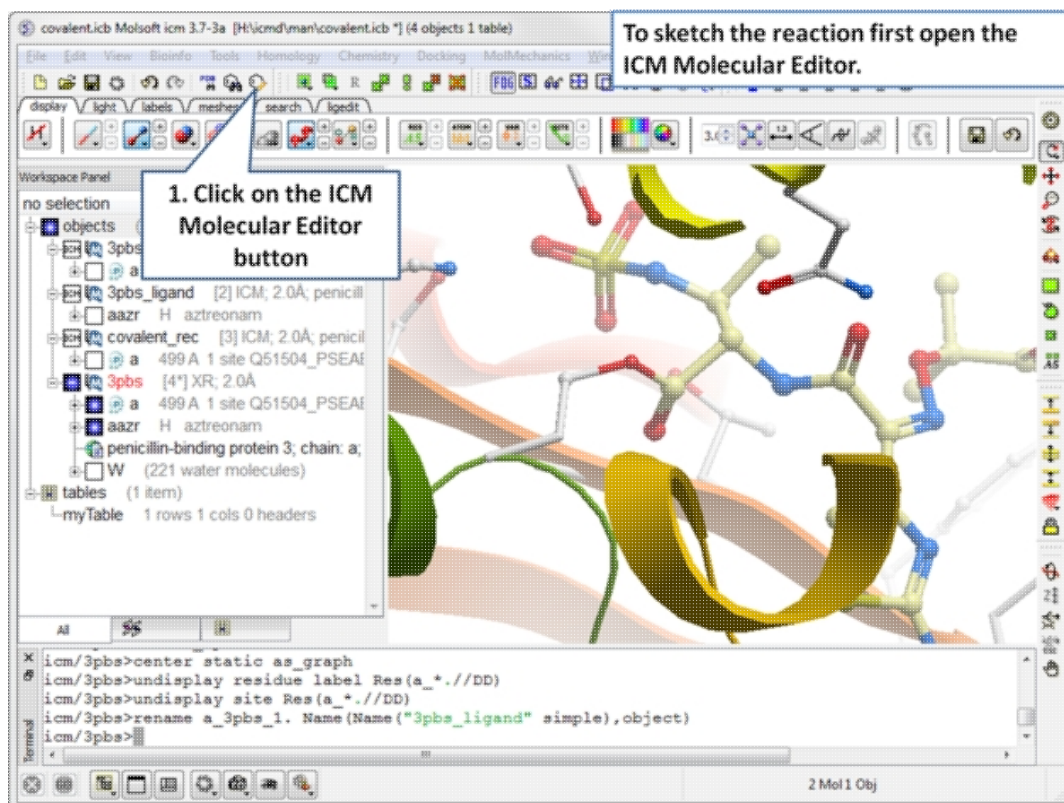
To turn off covalent docking:

- `Docking/Covalent docking` and check the option "Turn off Covalent Docking".

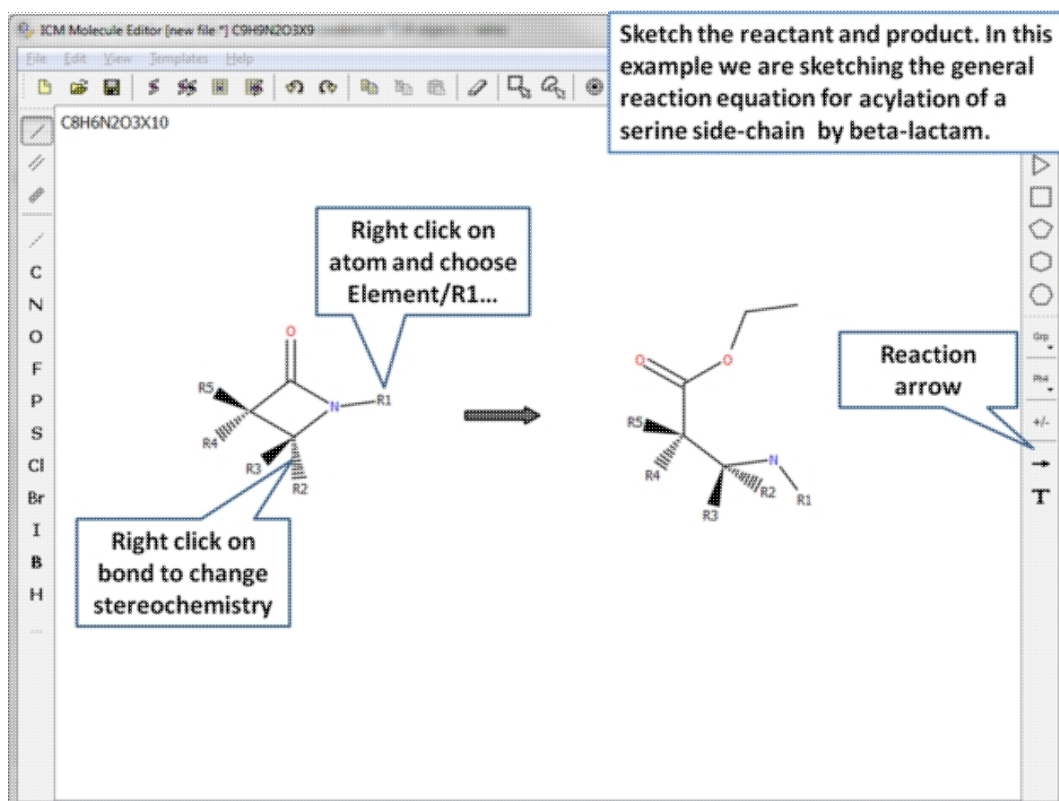
11.4.1 Sketch a Reaction

NOTE: Save time! Please check the in-built reactions in ICM to check if your reaction has already been drawn.

In this example we are sketching the general reaction equation for acylation of a serine side-chain by beta-lactam:

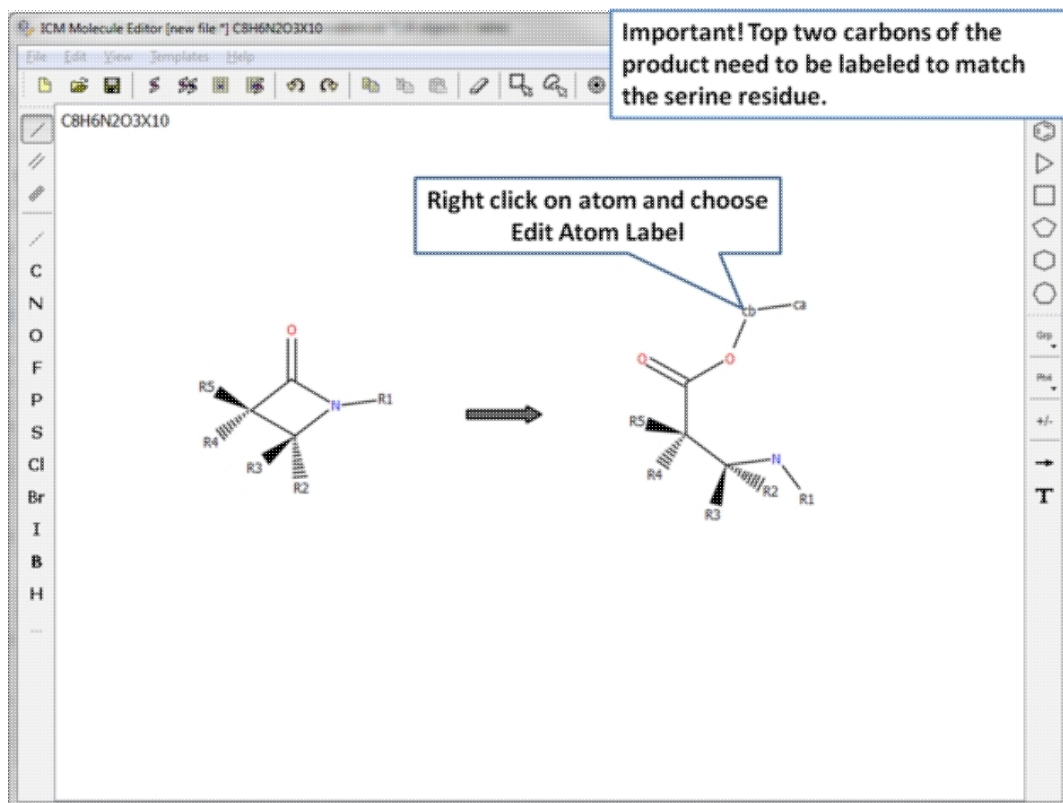


Open the ICM Molecular Editor.

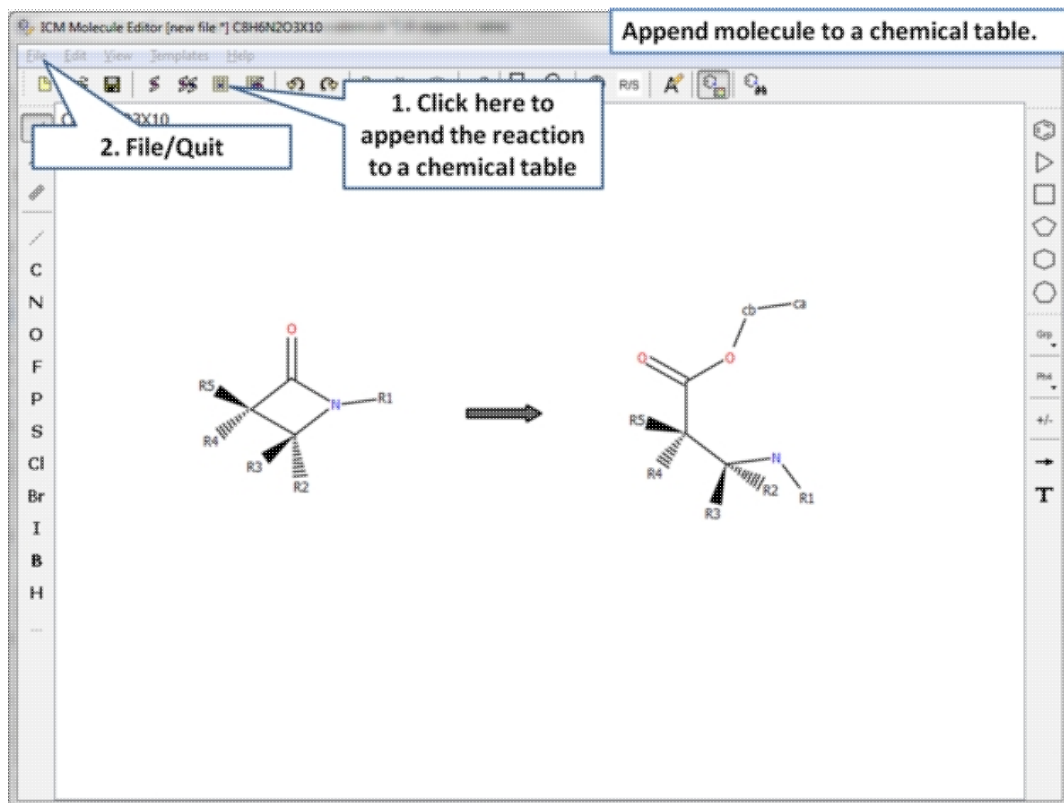


Sketch the reactant and product.

Note: The reaction table may contain multiple reaction mechanisms, in which case the system will try to apply them sequentially to the ligand, until first matching reaction is found. This is useful if ligands with different types of 'warheads' (reactive groups) will be considered.



Important! The top two atoms in the product which correspond to the modified serine residue need to be labeled to match the residue. In this example the atoms are labeled ca and cb.

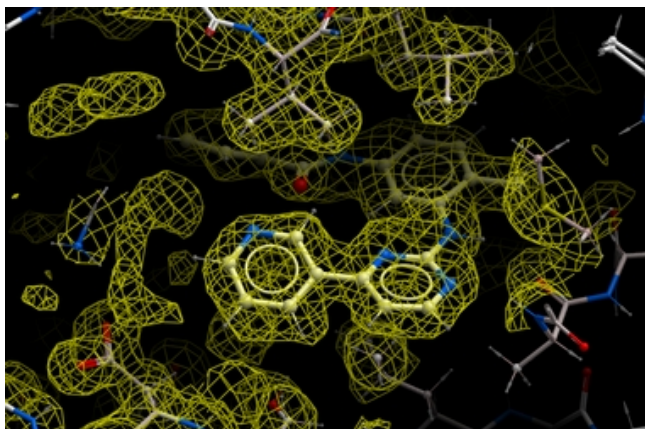


Append the reaction to a chemical spreadsheet (table). If you want to view the atom labels (ca, cb) please select the mol column and choose "Chemical View Options" > "Show Full Atom Names". Once the reaction has been sketched you can start the covalent docking procedure.

11.5 ICM X-Ray AutoFit - Automated Model Building into Density

Available in the following product(s): ICM-Pro | ICM-VLS |

The **ICM X-Ray AutoFit** is an automated method to fit a ligand into electron density. The tool combines the powerful ICM docking algorithm with an electron density fitting function.

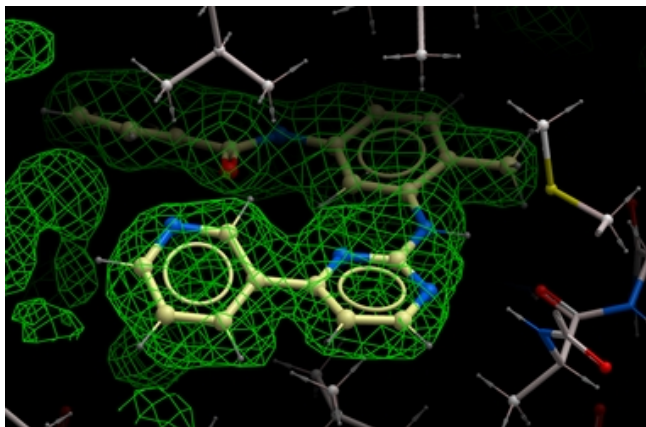


The input for **ICM X-Ray AutoFit** is an electron density map in CCP4 format, the protein receptor and ligand which can either be drawn or imported into ICM.

Theory

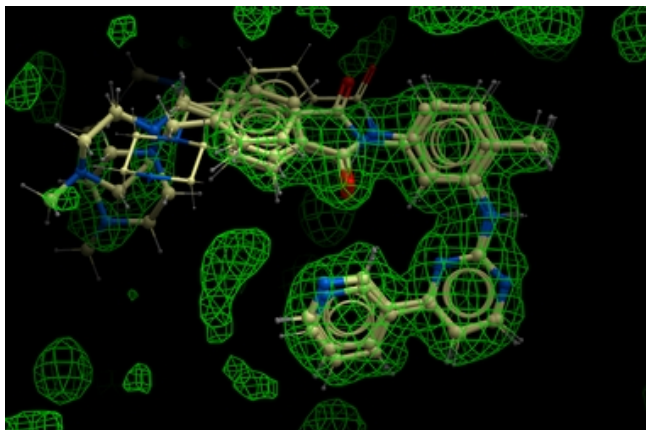
The **ICM X-Ray AutoFit** method includes the following features:

- Soft docking energy function.
- Intra and inter ligand interaction energy function.
- Weighted electron density contributions.
- The electron density for the fit function is filtered to exclude areas occupied by the protein receptor atoms.

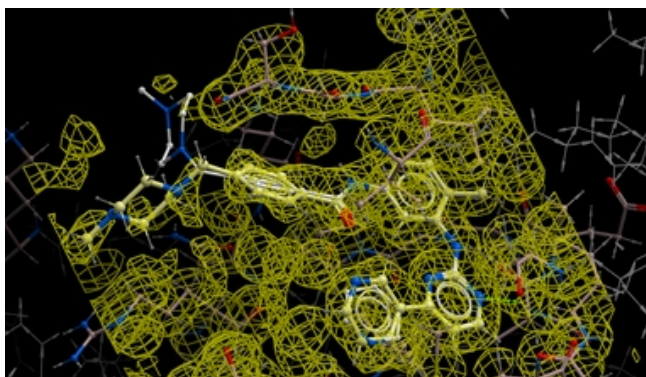


NOTE: Density from the receptor is automatically filtered out from the analysis. In the picture shown above the green map represents attractive potential.

The method generates multiple hits for each ligand with a score assigned. It has been demonstrated that improved ligand receptor interactions can be determined by the **ICM X-Ray AutoFit** method compared to published crystal structures.



In the figure shown below the interaction between Gleevec and Syk kinase is shown. The white carbon atoms are the published ligand pose and the yellow carbon ligand is the result of ICM which gives a better fit to the density.



Instructions

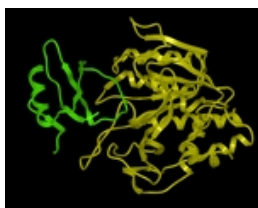
How to run the **ICM X-Ray AutoFit**.

- Load the receptor structure and `convert` to an ICM object.
- Load the CCP4 map (File/Open)
- Load the ligand and `convert` to an ICM object.
- Follow the small molecule docking procedure but **after** generating maps select Docking/X-Ray Density and then undertake docking in the standard way.

11.6 Protein-Protein Docking

Available in the following product(s): ICM-Pro

Here we describe the steps for protein-protein docking. An example is described using a complex of subtilisin and chymotrypsin (PDB code:2sni). The example will re-dock the ligand (PDB code entry 2ci2) into the receptor molecule (PDB code 2st1) and then determine how accurately the molecules are docked by comparison with the complex 2sni. The structure of 2sni is shown below with the ligand displayed in green and the receptor in yellow.



11.6.1 Determination of Protein-Protein Docking Interface

The ICM `Optimal Docking Area` method is a useful way of predicting likely protein-protein interaction interfaces. If you do not have mutational data or other experimental data which indicates the likely protein-protein docking site this method will be useful. This procedure can save you time during the docking procedure by focusing your docking only on areas on the receptor and ligand most likely to interact.

11.6.2 Protein-Protein Docking Procedure

To begin the protein-protein docking procedure:

1. Read in the PDB files for 2ci2 (ligand) 2st1 (receptor) and 2sni (complex for comparison). For instructions on how to load a PDB structure into ICM please [click here](#).
2. Convert all three PDB files into ICM objects.
3. Delete all waters and sulfate ions, you can keep the calcium ions if you wish.

Now go onto the first step of the protein-protein docking protocol which is to `Set Project name`.

NOTE: All the protein-proteing docking options can be found in the GUI menu `Docking/Protein-Protein`.

11.6.3 Protein-Protein Set Project

`Docking/Protein-protein/Set Project`

Start the protein-protein docking project setup by defining the project name:

- Click on `Docking/Protein-protein/Set project`
- Enter a unique name into the `Project name` data entry box. Avoid spaces and leading digits in the name - the name must be alphanumeric - no hyphens etc... All files related to the docking project will be stored under names, which start from the project name.

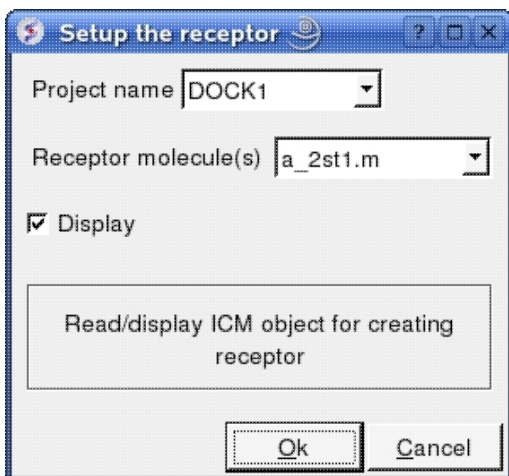


Now setup the receptor.

11.6.4 Protein-Protein Receptor Setup

Docking/Protein-protein/Receptor setup

- Enter the Docking Project name e.g. DOCK1
- Enter the receptor molecule e.g. a_2st1.m (use a_2st1.* if you want to include all molecules such as Calcium ions)
- Click OK

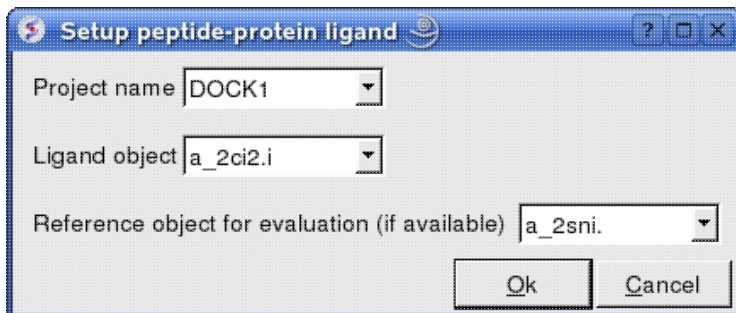


Now setup the ligand.

11.6.5 Protein-Protein Ligand Setup

Docking/Protein-protein/Ligand setup

- Enter the project name e.g. DOCK1
- Enter the ligand molecule e.g. a_2ci2.i
- If you wish to compare your docking data with a solved structure enter the name of the converted reference object in the "Reference Object" data entry box e.g. a_2sni.
- Click OK



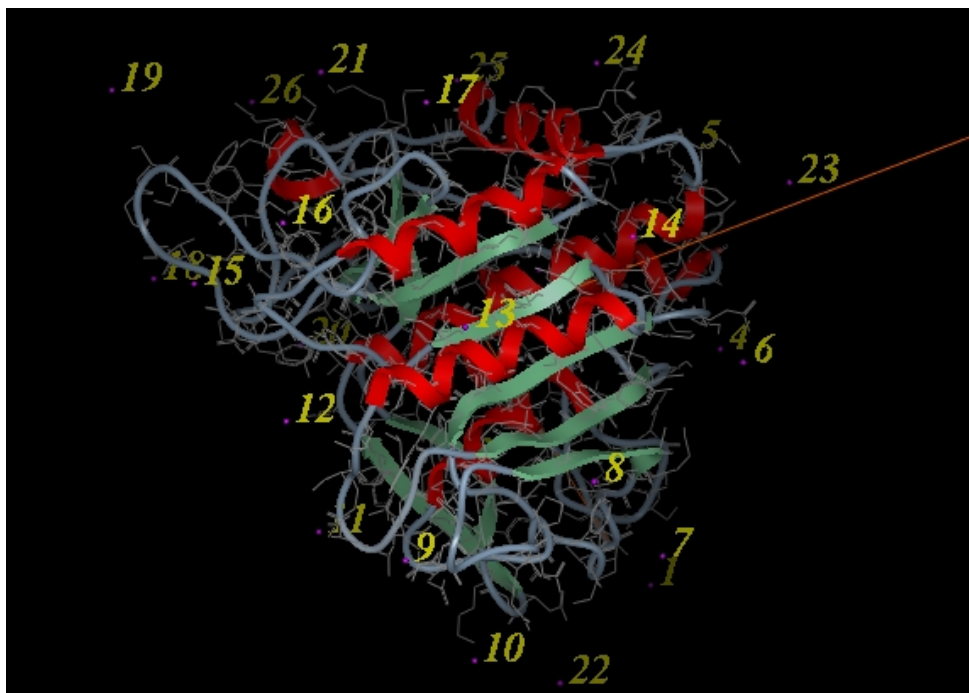
Now select an initial point of interest on the receptor referred to as `epitope` selection (NOTE: This step is optional. If you do not wish to select an initial point of interest jump to the `make maps` section. It is recommended to narrow down the docking interaction site as much as possible by selecting epitopes. Protein-protein interaction sites can be predicted by ODA or experimental data or based on a reference complex.

11.6.6 Epitope Selection

Docking/Protein-protein/Epitope selection

Select an initial point of interest on the receptor for the docking simulation. You may want to check biological data or a reference complex or predict interaction sites using ODA before doing this step. If you do not know the docking interface location you should setup multiple docking runs for different epitopes on the receptor.

- You can make selections on either the ligand or receptor or both the ligand and receptor. Check the appropriate box(es).
- A display as shown below will be displayed.
- Using the right button of the mouse select a numbered sphere surrounding the receptor or ligand that you wish to dock to by clicking and dragging the mouse over the sphere. The spheres are numbered and change color from purple to yellow when they are selected. If you are happy with the selection type 'a' or press the apply button. The selected numbered regions will change from purple to yellow. The easiest way to select multiple epitopes is to use the `pick atom` button (green cross button).
- When you have finished selecting the epitopes type 'q' or select the quit button in the terminal window.



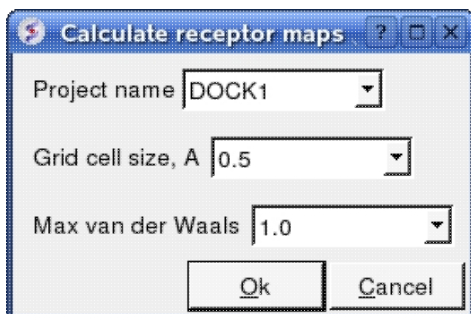
NOTE: If you are unsure which epitopes you have selected they are listed in the DOCKING_PROJECT_NAME.tab file in the first two fields called I_selLigPos (receptor epitopes) and I_selLigRot (ligand epitopes).

The next step is to make the maps of the receptor.

11.6.7 Protein-Protein Make Receptor Maps

Docking/Protein-protein/Make Receptor Maps

- Enter the Project name e.g. DOCK1
- Enter the grid size e.g. how detailed you want your maps - the default value of 0.5 is generally ok.
- Enter the Max van der Waals value which gives the receptor an element of 'softness' to incorporate some induced fit - the default value of 1.0 is generally ok.



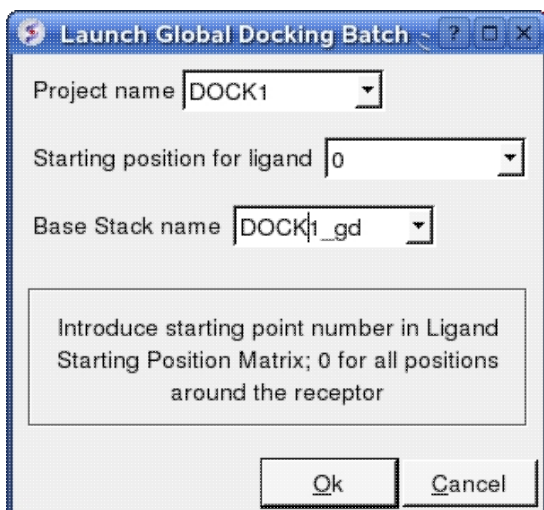
Now run the docking simulation.

11.6.8 Protein-Protein Docking Batch

The docking can be run on your local machine or in PBS.

To run on your local machine:

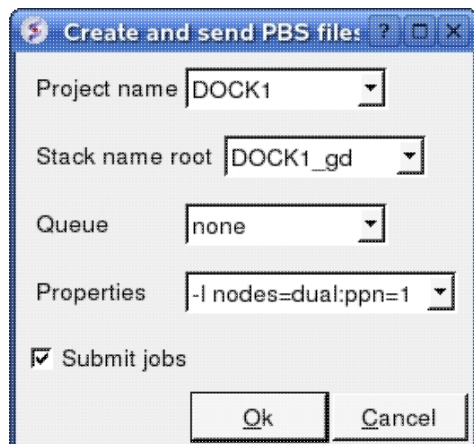
Docking/Protein-protein/Docking Batch/Local Machine



- Enter the Project Name
- Starting Position for Ligand - if you select 0 it will sample all the points on the receptor you selected in the epitope step of the docking project setup. If however you want to break your jobs down into smaller chunks you can enter the number of a position on the receptor you chose in the epitope selection step and it will sample that point.
- Enter a name for the conformational stack which will be saved.

To run in PBS:

Docking/Protein-protein/Docking Batch/PBS



You can check the docking progress by going to Windows/Background Jobs. Once the docking has finished a dialog window will be displayed telling you that the docking job is complete - **Click OK** and a results table will be displayed. Read the next section on how to view the results.

Error messages are reported at the end of the DOCKING_Project.out output file. You can also see details about any errors by clicking on the **Details** button in the dialog window which is displayed once the docking run is finished.

11.6.9 Display Grid Docking Results

Once the docking is finished you should see a table. A table as shown below will be displayed. You can sort the table by Energy (ener) by right clicking on the column header and select sort.

i	ener	rmsd	naft	nvis	v1	v2	v3	v4	v5	v6	Ey_gh	Ey_gc	Ey_gb	Ey_ge	Ey_gs	Ey_sfl
1	-41.391	0	2	6	0.04483	90.9106	524.007	92.0525	49.7742	83.8696	-7.8806	-35.857	-5.8146	-1.2662	-6.2756	11.166
2	-40.565	0	2	2	0.19569	91.2981	520.773	-8.9542	86.9122	79.9324	-9.0676	-34.638	-3.9748	-1.7934	-11.620	13.840
3	-38.606	0	2	4	1.02538	91.1309	522.608	171.737	101.824	81.2009	-5.5076	-36.427	-3.9280	0.35496	-7.8400	11.036
4	-37.223	0	2	2	-0.2911	90.6959	527.715	-168.89	106.47	16.6569	-7.7926	-26.602	-6.7522	-0.8824	-7.4796	8.4756
5	-36.899	0	2	5	-1.8815	88.37	522.204	178.837	103.21	40.3486	-4.1510	-33.382	-6.2505	-1.2081	-8.5090	11.532
6	-34.038	0	2	8	-1.6260	89.4907	528.356	72.7411	74.769	66.6841	-1.4804	-29.692	-6.3669	-1.0912	-8.6190	9.6204
7	-33.743	0	2	3	-0.9527	90.0158	528.304	-160.48	89.2895	72.3303	-6.0312	-27.820	-3.3335	-2.1168	-8.5995	10.073
8	-33.086	0	2	2	-0.9266	91.388	521.931	-36.119	116.206	77.6997	2.43218	-39.036	-5.2885	-0.6586	-11.599	15.206
9	-30.652	0	2	2	-1.7173	88.3469	523.435	153.731	124.097	68.3627	-5.8456	-27.796	-3.8465	-1.4351	-5.8874	9.8432

To display the complexes:

- **Single** click on a row of the table shown above. The ligand will be displayed in the ICM workspace and named according to the project name followed by "_lig" (e.g. DOCK1_lig).
- Check the R_Srmsd column for the difference between docked and the crystal structure complex for comparison (if selected).
- To view interactions between the receptor and the ligand each molecule needs to be in the same object. See the FAQ section: How do I merge two separate objects into one.

If no docking table is displayed you can process the results by:

- Docking/Protein-protein/Docking Batch/Process Global Docking Solutions.

The output columns represent:

- i - a slot number in the stack of conformations
- ener - total energy as calculated before the conformation was stored

- rmsd - the distance (either Cartesian or angular RMSD) between the current conformation of the object and the stack conformation calculated according to the compare command.
- naft - the number of visits AFTER the last improvement of energy
- nvis - the total number of visits to this slot; since new conformations are only compared with the last stack conformation the conformations may drift and cover a large area than described by the vicinity parameter
- v1 to v6 - are the virtual variables defining position and rotation of the ligand molecule.
- ey gh - van der Waals grid potential - hydrogen probe
- ey gc - van der Waals grid potential - carbon probe
- ey gb hydrogen bonding grid potential
- ey ge electrostatics grid potential
- hydrophobic potential
- ey sfPola - polar terms of the solvation energy
- Ey_sfAl - aliphatic terms of the solvation energy
- Ey_sfAr - aromatic terms of the solvation energy
- Ey_compSol - weighted total of the solvation energy terms

All your docking data is saved so if you want to return to view the table at a later date please use the following commands in the command line.

```
read object "DOCK1_rec" # read receptor (if not read yet)
display a_DOCK1_rec. # display receptor (if not displayed yet)
read object "DOCK1_lig" # read ligand object
display a_ # display this ligand
read table "DOCK1_gd.Var" # read table of the ligand conformations. Click on table rows to vi
```

11.6.10 Complex Refinement

You can refine a complex by:

- Right click on the conformation you wish to refine in the `results` table.
- Select the option **Refine Conformation**.
- Click the green **GO** button located at the bottom of the graphical user interface.

Once the simulation has finished you can view the results by:

- MolMechanics/Stack View
- A table will be displayed called **confStack** ranked by energy.
- Click on each row of the table to view the solution.

12 Virtual Ligand Screening

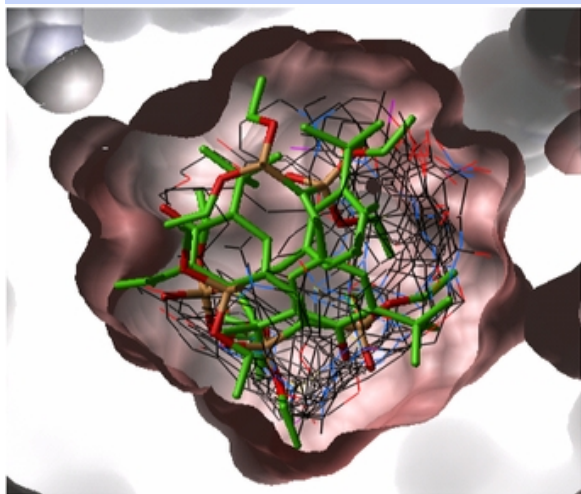
Chapter Contents

- Structure-Based Virtual Ligand Screening (VLS)
- Fragment Screening
- Ligand-Based Screening using Atomic Property Fields (APF) 3D Pharmacophores
- MolScreen >360 prepared 2D FingerPrint and 3D Pharmacophore Models

12.1 Virtual Ligand Screening

Available in the following product(s): ICM-VLS |

Additional Resources: Tutorial: Screen to Ricin Receptor | Tutorial: Screen to Cyclooxygenase | Tutorial: Screening a Markush Library | Command Line Docs | Video



12.1.1 Virtual Ligand Screening

Virtual Ligand Screening (VLS) in ICM is performed by docking a database of ligands to a receptor structure followed by an evaluation of the docked conformation with a binding-score function. Best-scoring ligands are then stored in the multiple icm-object file. The set-up of the VLS process is largely identical to the set-up for the small molecule docking simulation (see Small Molecule Docking section).

12.1.2 VLS Getting Started

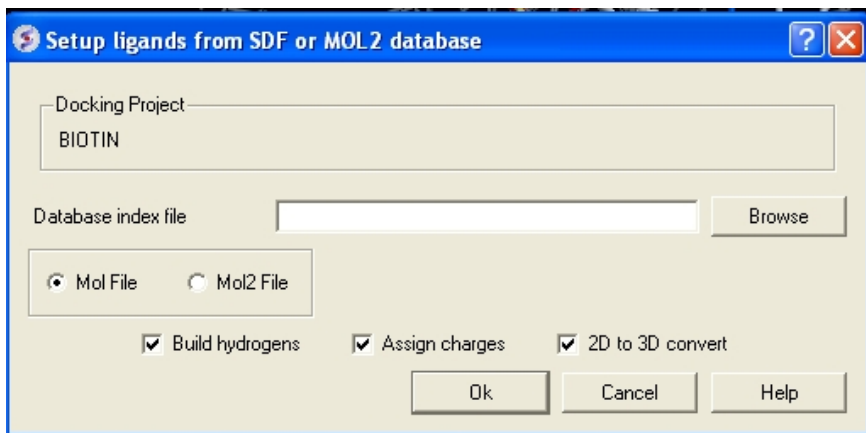
Follow the instructions in the small molecule docking section manual from docking project setup option to the calculate maps option. Use Docking/Setup Batch Ligand option to select the database you wish to dock.

12.1.3 Database File Format

In most cases the ligand input file will be an SDF or MOL2 file. These files need to be indexed by ICM before they can be used in VLS runs. The index is used to allow fast access to an arbitrary molecular record in a large file such as an SDF file which in some cases contains over one million compounds.

To index an sdf file:

- Click on the menu Docking/Tools/Index Mol/Mol2 File/Database to generate the index. The following data entry box will be displayed.



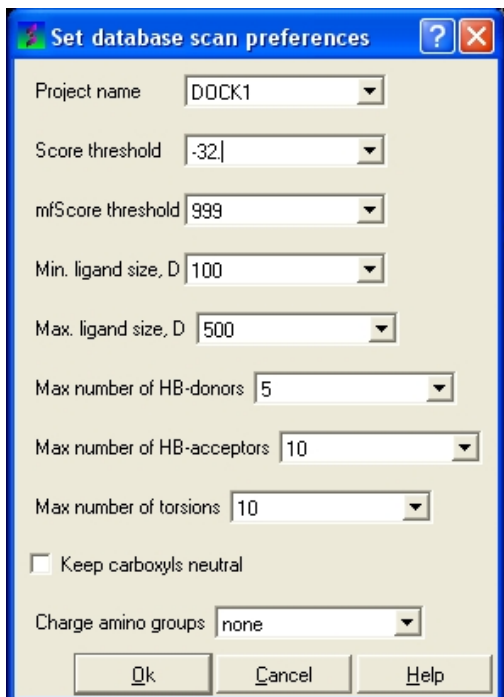
- Enter the name of your Mol/Mol2 file and enter the name you wish to call your index file.
- Select whether your file is in Mol or Mol2 format.

12.1.4 VLS Preferences

NOTE: It is important to setup the VLS preferences before undertaking VLS run. Additional preferences can be changed in the `docking_project_name.tab` file. The fields in this file are described here: <ftp://ftp.molsoft.com/man/dockTable.pdf>

VLS preferences can be setup by:

- Selecting the menu Docking/Preferences/Database Scan



Different options are available to select by clicking the down arrow next to the data entry field. These options are described here:

Score Threshold:

An important parameter of the VLS run is the score threshold. Docked conformations for a particular ligand will only be stored by ICM VLS procedure if its binding-score is below the threshold. The choice of the threshold can be done in two ways: based on the scores calculated by docking known ligands. Generally, a value somewhat above typical score observed for known ligands is a good guess. If no ligands are known, a pre-simulation can be run using ~1000 compounds from the target database. Using the resulting statistics for the scores, the threshold should be set to retain ~1% of the ligands.

Potential of mean force score:

Potential of mean force calculation (pmf) provides an independent score of the strength of ligand-receptor interaction. The pmf-parameters are stored in the icm.pmf file.

Other selection criteria:

Other selection criteria which can be changed include

Minimum/Maximum Ligand Size you wish to be screened.

Maximum number of H-bond donors

Maximum number of H-bond acceptors

Maximum number of torsions

12.1.5 Docking Project Table File

ICM VLS uses a number of criteria to pre-select compounds before docking. Edit the project .tab file to change their defaults. A full description of each field in the .tab file can be found here

<ftp://ftp.molsoft.com/man/dockTable.pdf>

```
#>i DOCK1.i_maxHdonors
5
#>i DOCK1.i_maxLigSize
500
#>i DOCK1.i_maxNO
10
#>i DOCK1.i_maxTorsion
10
#>i DOCK1.i_minLigSize
100
```

12.1.6 Run VLS in the Graphical User Interface

First setup the docking project (From Set Project to Setup Batch Ligand)

To start the vls job:

- Docking/Run Docking Batch

12.1.7 Running VLS Jobs in PBS UNIX Cluster Environment

Before VLS jobs can be run make sure you follow the instructions in the manual entitled Small Molecule Docking from docking project setup menu to the calculate maps menu. Select the "From indexed database", "From MolCart" or "From File: SDF/Mol2" option in the Setup Batch Ligand. Docking setup can be scripted see the terminal output from the GUI options to view the commands. More information on Docking and Virtual Screening using the command line can be found here <http://www.molsoft.com/man/ligand-fit.html>

12.1.8 Parallelization

If the database size exceeds several thousand compounds, it is desirable to run a number of VLS jobs in parallel to speed up the calculations. You can do this by choosing the Docking/Run Database Scan Batch option.

- Select Docking/Run Database Scan Batch option
- Select the **local** tab
- Select the docking project name.
- Enter the compound number you would like to dock from e.g. 8001
- Enter the compound number you would like to dock to e.g. 10000
- License type: For nearly all license (e.g. standard ICM-VLS licenses) types you need to **leave this entry blank**. If you have a -vlscluster license or a molvls license (-vls) then select the option from the drop down list. Any questions first check your license.dat file or Email support@molsoft.com.

12.1.9 VLS Results

The easiest way to view the results of a VLS run is to make a hitlist. This was described earlier in the `hitlist` section of the small molecule docking chapter of the manual. Other ways of manipulating VLS data are described here:

12.1.10 Post Screen Analysis

12.1.10.1 Sorting the compounds in your HITLIST

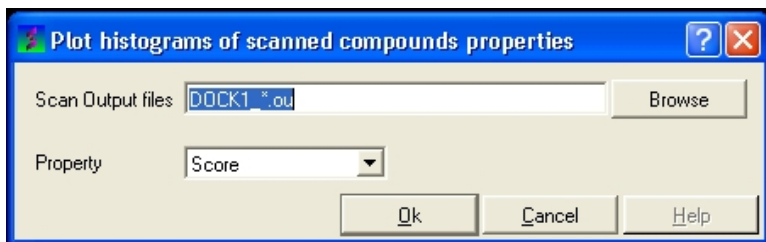
Compounds can be sorted according to their SCORE etc. See the `tables` section of this manual for more information about manipulating tables.

12.1.10.2 How to Plot Histograms and Scatterplots of VLS Data

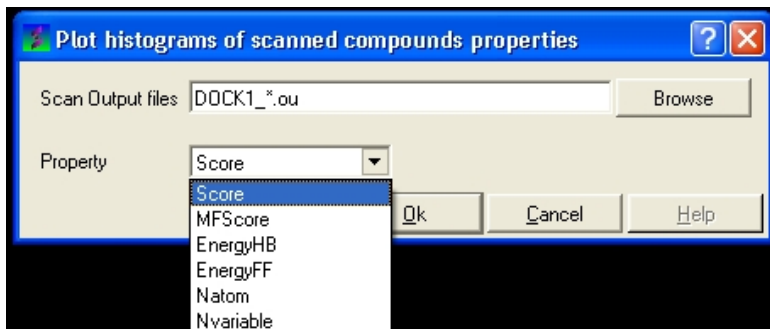
The hitlist contains many columns with numerical data. ICM can build interactive plots with the table columns (See `Tables` section). However, there are some easy to use plotting options in the docking menu which is described here.

12.1.10.3 To construct a histogram of your VLS data

- Select the menu Docking/Tools/Scan Results Histogram



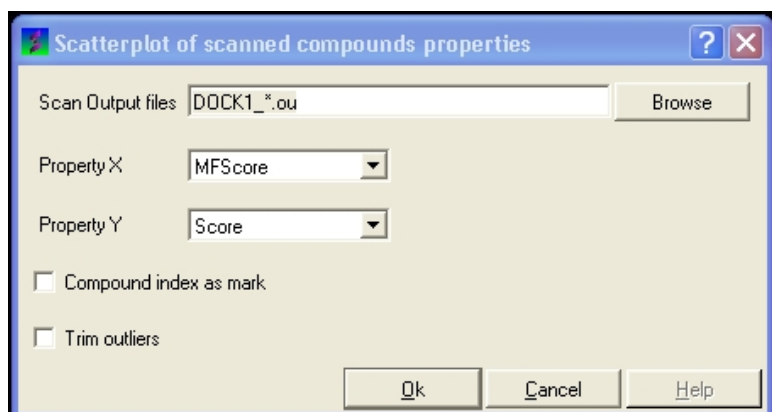
- Enter the name of the VLS output file (*.ou) you wish to construct a histogram for.
- Select which parameter you wish to plot against frequency (see below).



- Click OK and a def.eps file will be saved with a picture of your histogram.

12.1.10.4 To construct a scatterplot of your VLS data

- Select the menu Docking/Tools/Scan Results Scatterplot

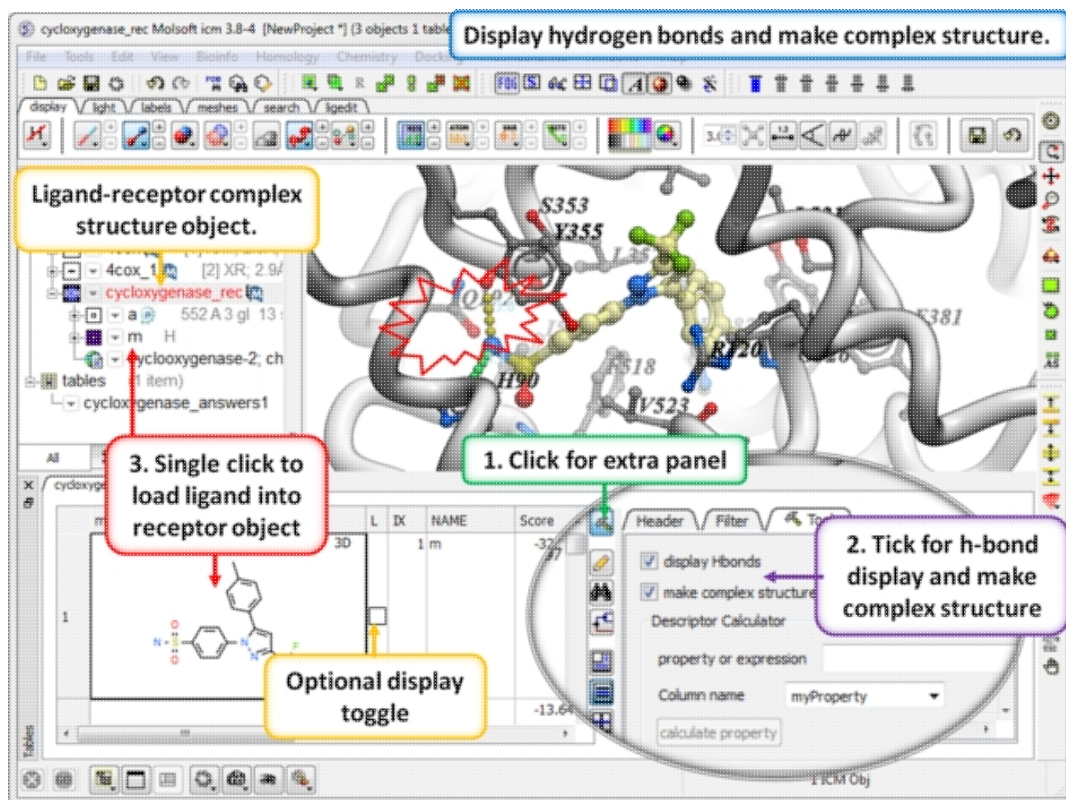


- Enter the name of the VLS output file (*.ou) you wish to construct a scatterplot for.
- Select which parameter you wish to on the X axis.
- Select which parameter you wish to on the Y axis.
- Click OK and a def.eps file will be saved with a picture of your scatterplot.

12.1.10.5 Hitlist - Make Complex and Display Hydrogen Bond

When you make the hitlist you will see a panel next to the chemical table called **Tools**. The options here can be used to:

- Check the option **display Hbonds** if you would like to display hydrogen bonds when you browse the hitlist.
- Check the option **make complex structure** if you would like the ligand-receptor complex to be displayed when you browse the hitlist.



12.1.10.6 Post Screening Tools - Calculate Properties

There are a selection of **property expressions** associated with the ligands in the hitlist that can be calculated. To use these press the drop down button next to the **Property Expression** button. The options include:

- Calculate charge of the ligands in the hitlist.
- Calculate buried surface of the ligands in the hitlist.
- Calculate volume change of the ligands in the hitlist.
- Calculate whether the ligand makes a hydrogen bond with a selection. Make a selection of the residue in the receptor first.
- Calculate pharmacophore RMSD. This property expression needs an object called 'pharma' loaded in the workspace, with the pharmacophore in it.
- Calculate cisTransAmide Angle of the ligands in the hitlist.
- Calculate distance to selection. Make a selection of the residue or atom in the receptor first.
- Calculate non-overlap with selection. Make a selection of the residue in the receptor first.
- APF pose similarity to a selected ligand. Make a selection of a ligand first.
- Calculate and cluster the interaction fingerprints.

Each of these properties will be placed in the hitlist table as a new column. You can choose a column name before calculating using the **Column name** dialog box. Click on the **Calculate property** button to run the calculation.

mol	L	IX	NAME	Score
1	racemic	3	C14H22N3O3	-37.23
	chiral	24	91536	-36.55

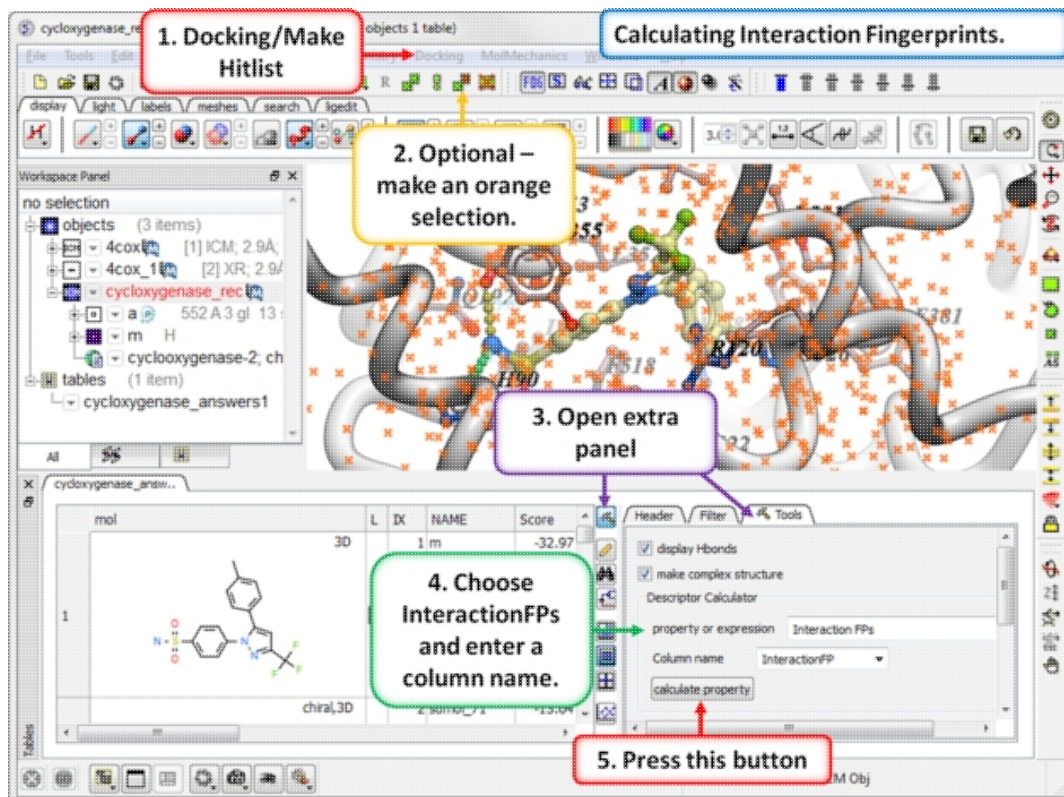
12.1.10.7 Cluster by APF Ligand Pose Similarity

Screening results in the hitlist can also be clustered by 3D pharmacophore chemical pose similarity using Atomic Property Fields. To do this:

- Make the hitlist (Docking/Make Hitlist)
- Press the **Cluster by APF** button on the Tools panel (right hand side of hitlist table).
- A tree will be displayed clustering the pose similarity of the compounds by APF fit.
- The distance matrix is added to the hitlist.

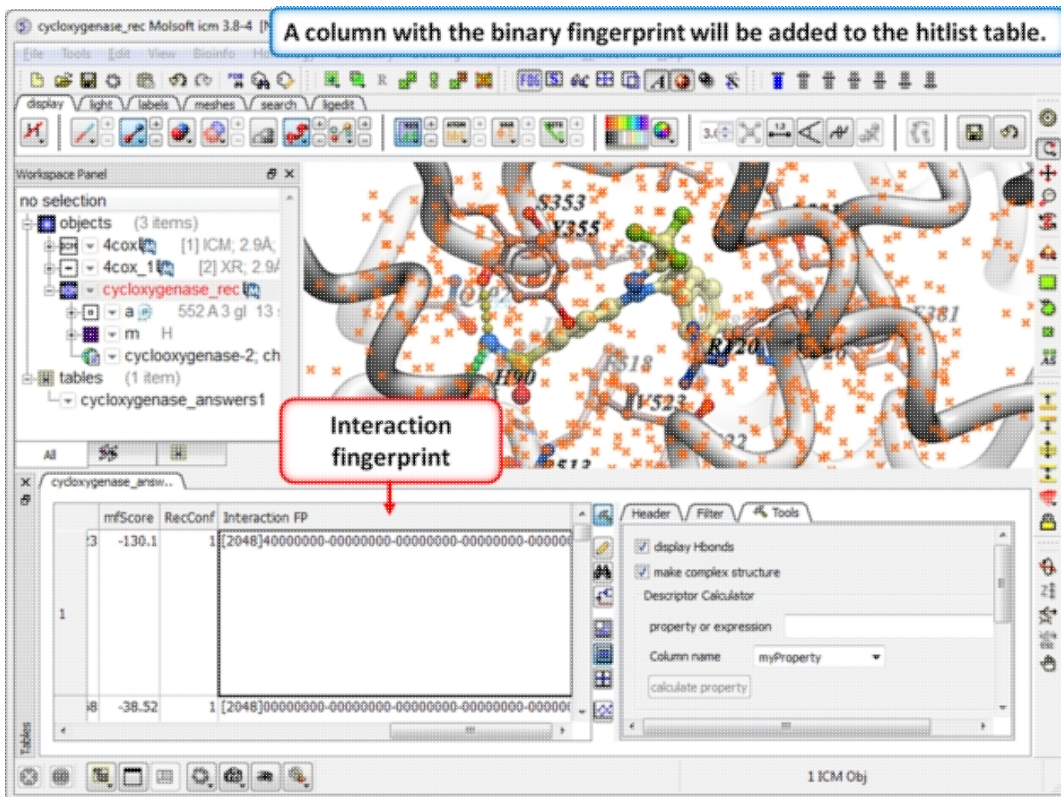
12.1.10.8 Interaction Fingerprints

An interaction fingerprint can help choose docked ligands with common interactions or a diverse set of interactions. The fingerprint is a bit string representing a contact with each atom. For specificity there is a special subset of bit strings which represent hydrogen bonding. An alternative to interaction fingerprints is to cluster by 3D pharmacophoric similarity using APF.

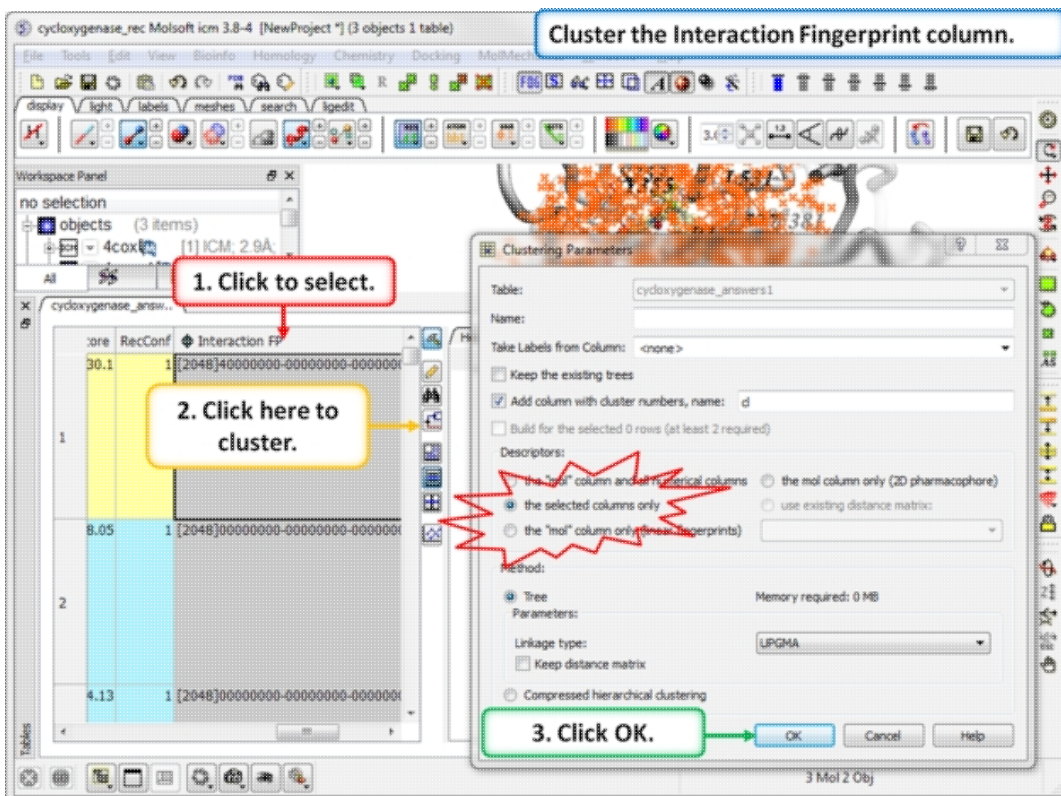


Step 1: Calculate Interaction Fingerprints

- Setup and run the docking and then make a hitlist.
- **Optional:** Make an orange selection of the ligand binding pocket. If no selection is made ICM will use the binding pocket as defined in the docking project.
- Open the extra panel in the hitlist table.
- Choose the option InteractionFPs in the Descriptor Calculator section of the panel.
- Enter a name for the column.
- Click **calculate property**.

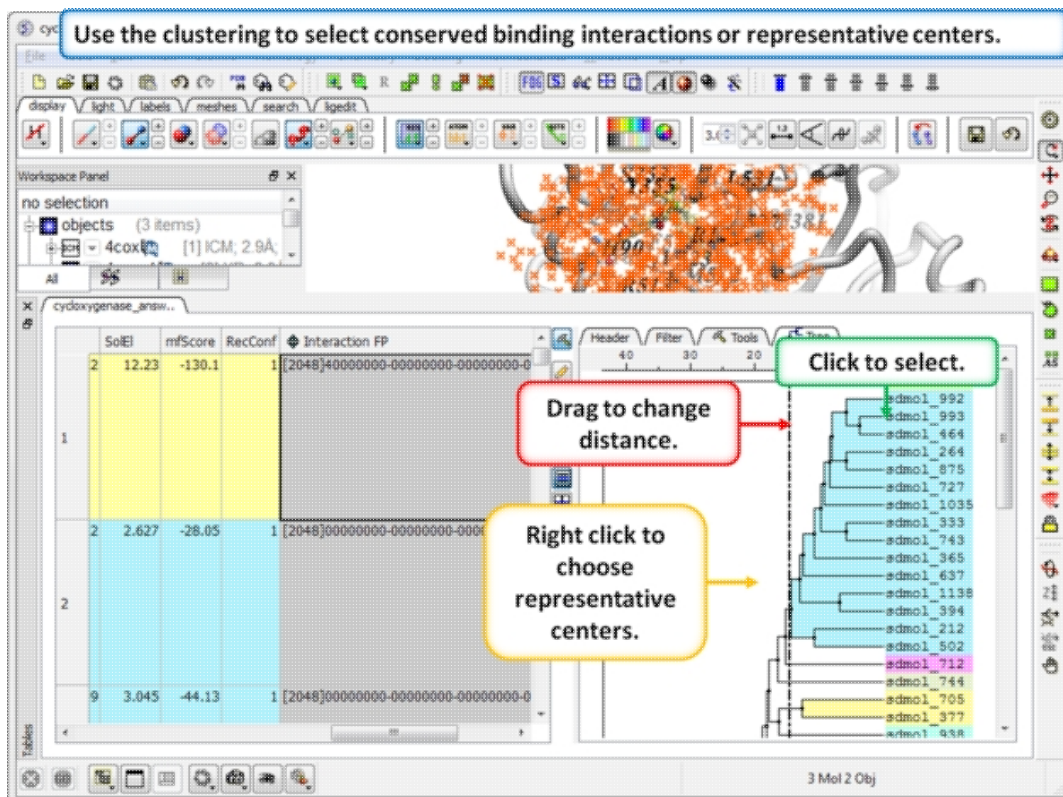


Step 2: Once the calculation has finished a column will be added to the hitlist table.



Step 3: Cluster the Fingerprints

- Select the fingerprint column.
- Click on the cluster button.
- Choose the options shown in the dialog window above.



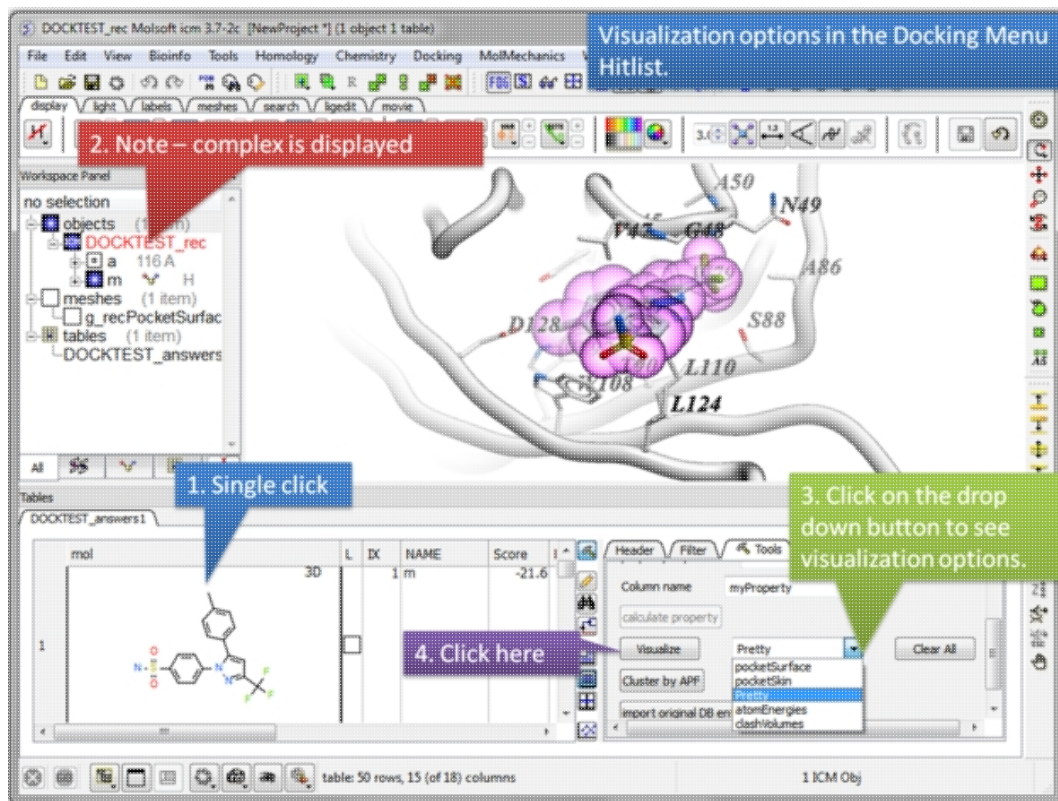
Step 4: Select representative centers or conserved binding interactions.

- Drag the distance bar to change the tree.
- Click on the branches of the tree to select (use CTRL for non-contiguous selections).
- Right click in "white space" and choose representative centers.

12.1.11 Post Screen Visualization

Once you have made a hitlist of the screening results you can use the panel on the right hand side of the hitlist to change visualization options. The available options are:

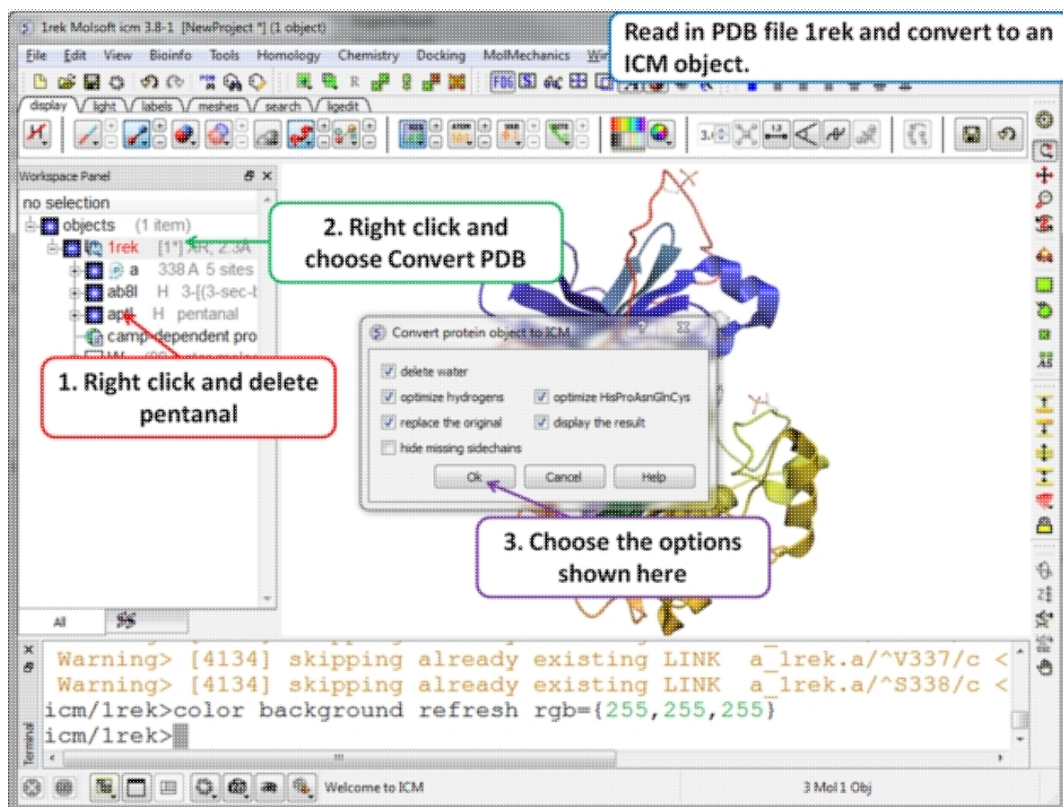
- pocketSurface = The surface of the ligand binding pocket. White=neutral surface Green=hydrophobic surface Red=hydrogen bonding acceptor potential Blue=hydrogen bond donor potential
- pocketSkin = A skin representation of the neighboring residues.
- Pretty = A default representation for ligand and receptor (see image below).
- atomEnergies = An energy value for each atom in the ligand is displayed. Green and low numbers indicate a positive contribution to binding.
- clashVolumes = Highlights any clashes between ligand and receptor.



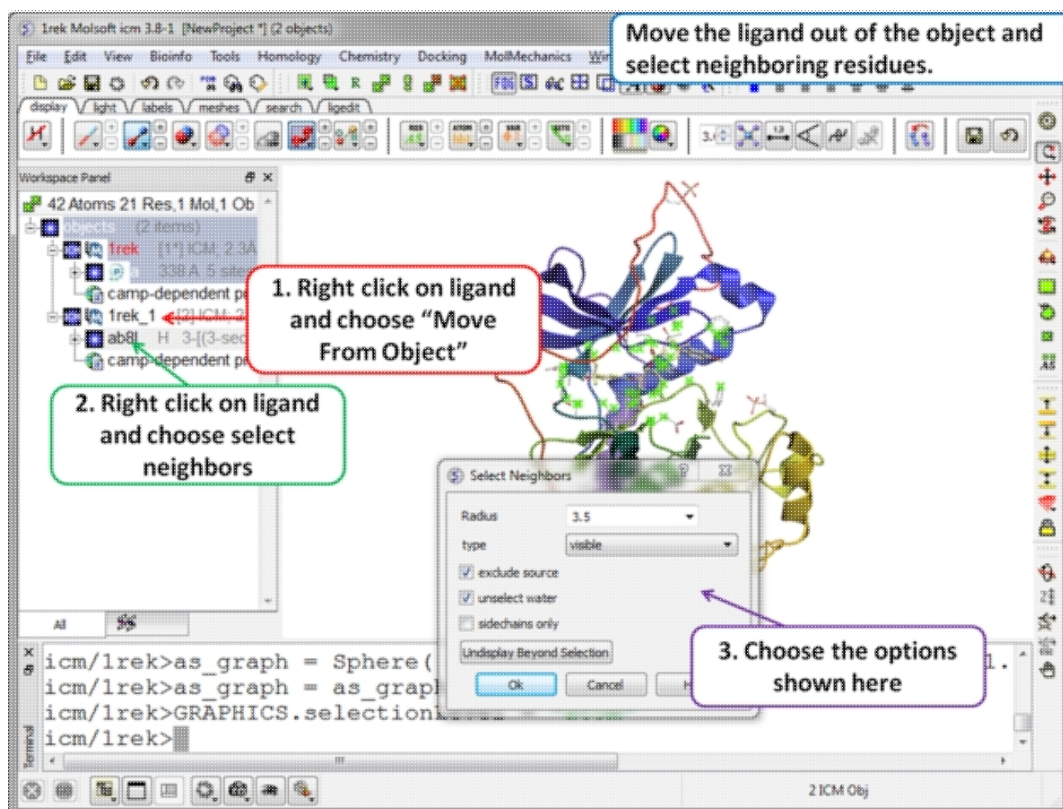
12.2 Fragment Screening

Available in the following product(s): ICM-VLS I

Fragment screening allows you to screen a database of small chemical fragments to a protein receptor. The fragments are then ranked by p-value providing a quantitative measures of confidence that a fragment is a true active and binds in the specific pose in absolute terms, instead of just ranking it versus other fragments. The fragment hits can then be "grown" or linked to produce a lead with high affinity. In this example we screen a database of fragments to cAMP-dependent protein kinase ATP binding pocket.

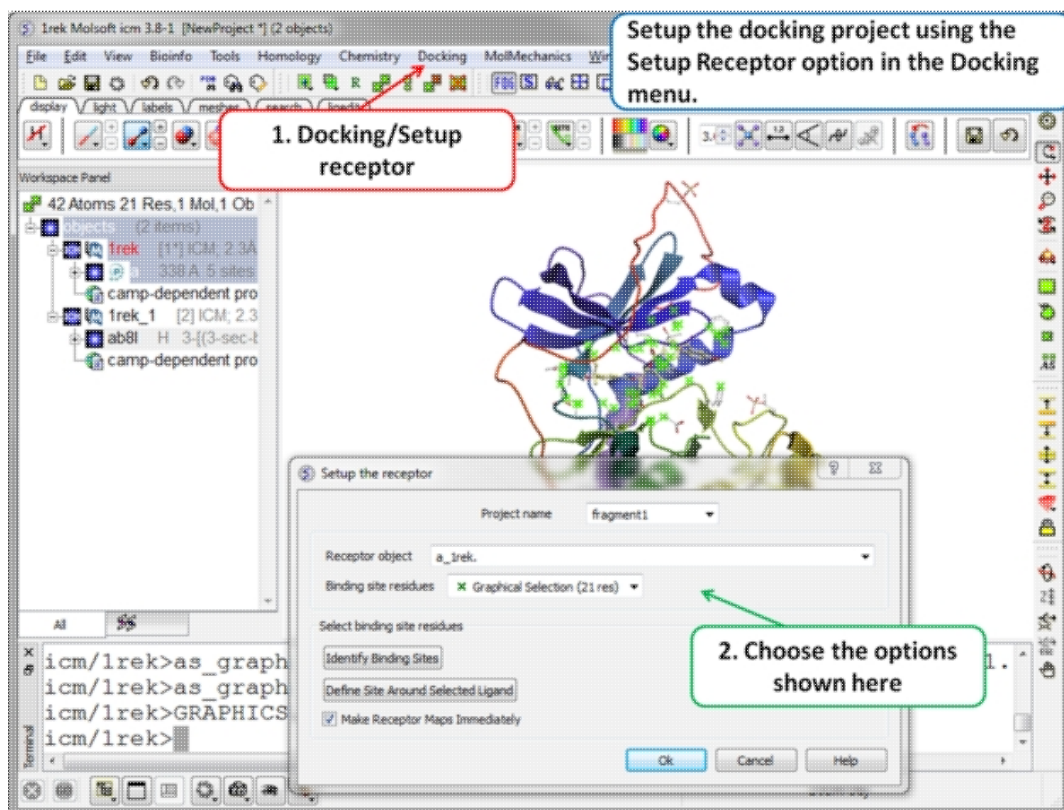


Step 1 For this example we will use PDB file 1REK. It needs to be converted to an ICM Object and the waters and pentanal should be deleted.

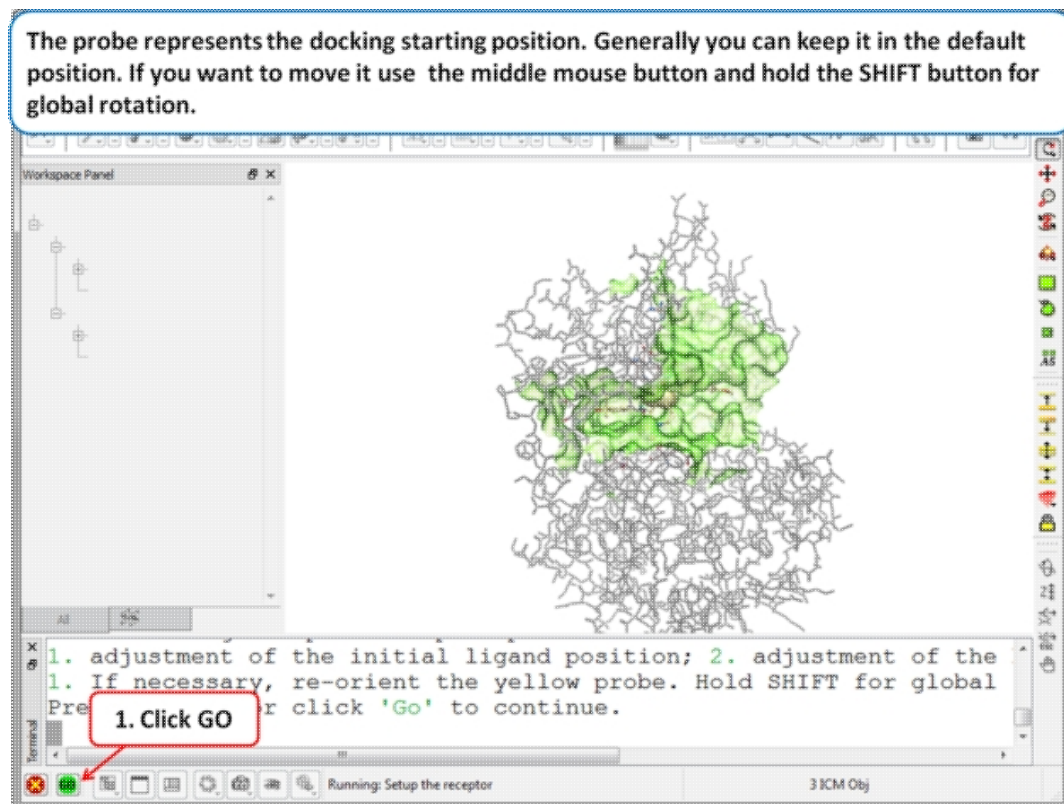


Step 2 Move the ligand out of the object into a separate object so it is not included in the docking maps. Select the binding site by right clicking on the ligand and choose "Select Neighbors" in this example a radius of 3.5Å is sufficient to select the pocket. Be careful not to select residues outside of the pocket

because this could potentially include smaller cavities outside the pocket where fragments could bind.



Step 3 Setup the docking project using the docking menu and the option "Setup Receptor".



Step 4 The probe represents the docking starting position. Generally you can keep it in the default position. If you want to move it use the middle mouse button and hold the SHIFT button for global

rotation. .

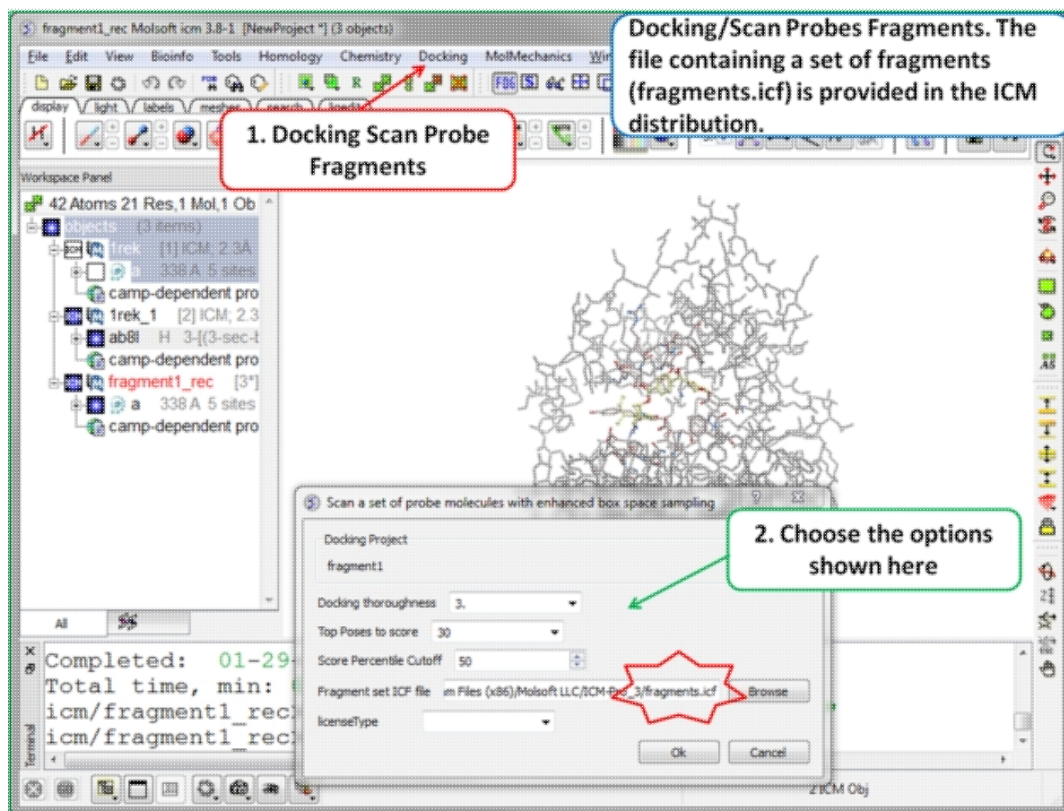
The purple box represents the region in which maps will be generated. To change the box size you can use the left mouse button with the cursor at any corner of the purple box to change it. It is important for fragment docking to minimize the number of cavities outside the pocket so to prevent fragments from docking outside of the pocket. Make sure the purple box is as close to the pocket as possible.

If necessary, adjust the size/position of the box around the binding pocket (hold LEFT MOUSE BUTTON with the cursor at any corner of the box). Press 'GO' or click 'Go' to continue.

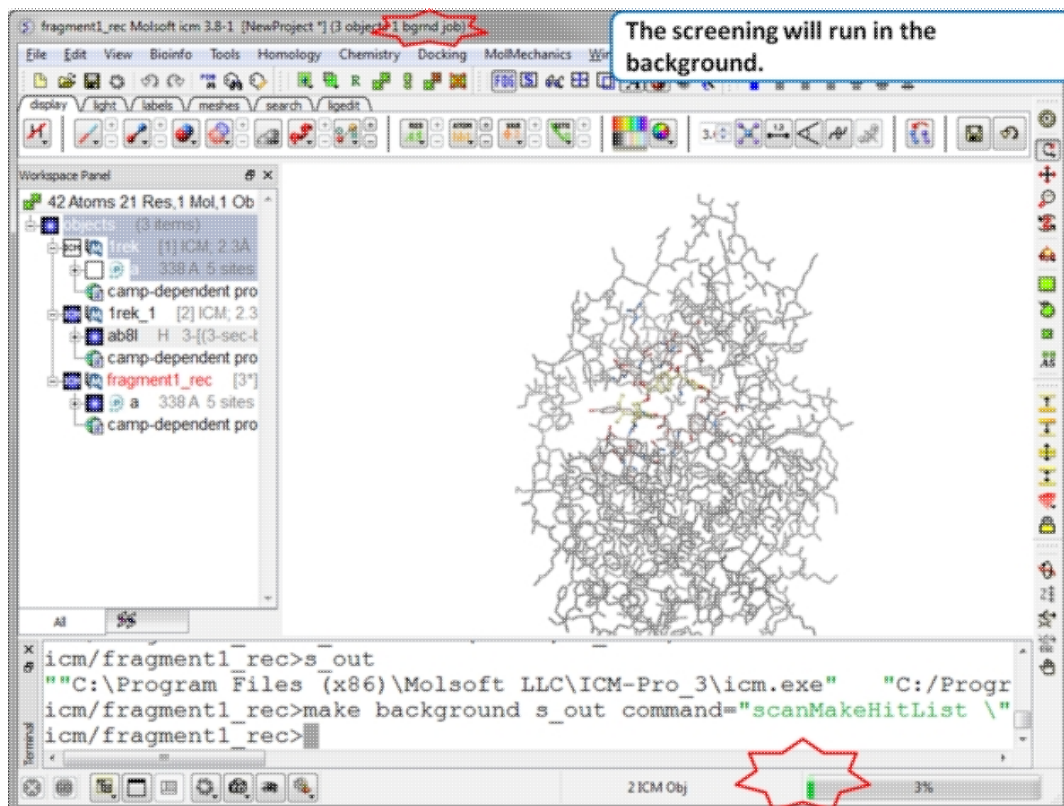
1. Click GO

Running: Setup the receptor 3 ICM Obj

Step 5 The purple box represents the region in which maps will be generated. To change the box size you can use the left mouse button with the cursor at any corner of the purple box to change it. It is important for fragment docking to minimize the number of cavities outside the pocket so to prevent fragments from docking outside of the pocket. Make sure the purple box is as close to the pocket as possible. .



Step 6 To start the screening go to Docking/Scan Probes Fragments. A database of fragments is provided with the database distribution - the file is called fragments.icf.



Step 7 The screen will take several minutes to run in the background.

fragment1_fragments_1 Molsoft icm 3.8-1 [NewProject ""] (4 objects: 1 table)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics Win

display light labels meshes search ligedit

Workspace Panel

42 Atoms 21 Res, 1 Mol, 1 Ob

- 1rek_1 [2] ICM, 2.3
- ab8l H 3-[(3-sec-t
- camp-dependent pro
- fragment1_rec [3]
- a 338 A 5 sites
- camp-dependent pro
- fragment1_fragments1
- m H C5H8N2
- fragment1_fragment
- stack (1 conformation
- tables (1 item)
- fragment1_fragments1 2

Once the screen has finished a table of results will be displayed.

1. Toggle here to view the docked fragment

mol	L	IX	NAME	Score	Natom	Nflex	Hbond	Hphob	VwInt	Eintl	Dsolv	So
1	<input type="checkbox"/>	1	C5H8N2	-21.69	15	0	-5.329	-3.204	-13.25	2.008	4.492	

3 ICM Obj

Step 8 Once the screening has finished a table of the results will be displayed. The docked fragments can be viewed by toggling the view on and off using the buttons in the "L" column. Guide to the table columns:

- **IX** is the index number from the docked database
- **Score** is the ICM score -32 and lower are generally considered good scores - but depends on the receptor (e.g. exposed pockets or pockets with metal ions may have higher scores than -32).
- **Natom** is the number of atoms in docked ligand
- **Nflex** is the number of rotatable torsions.
- **Hbond** is Hydrogen Bond energy
- **Hphob** is the hydrophobic energy in exposing a surface to water
- **VwInt** is the van der Waals interaction energy (sum of gc and gh van der waals). Current version of the score uses explicit van der Waals interaction energy calculation (no grids)
- **Eintl** is internal conformation energy of the ligand
- **Dsolv** is the desolvation of exposed h-bond donors and acceptors.
- **Solel** is the solvation electrostatics energy change upon binding.
- **mfScore** is the potential of mean force score
- **pValue** and **Percentile**. Because the fragment database is relatively small it has been possible to pre-validate the scoring individually for each member of the library. This allows us to estimate sensitivity and specificity (true positive/true negative rates) at any given score cutoff. As a result, expected number of false positive binding poses below the cutoff for each particular fragment can be generated. This provides a quantitative measure of confidence that a fragment is a true active and binds in the specific pose in absolute terms, instead of just ranking it versus other fragments (see Step 9).
- **examplePDB** Each fragment in the library is linked back to the PDB entries with complexes of the fragment-containing ligands, and appropriate rotation/translation matrices for instant superposition. Click on the hyperlink to display the fragments.
- **Location** these values indicate the location of the ligand in the pocket. The values can be clustered to select fragments in different regions of the pocket (Step 10-12).

Sort the table by pValue.

1. Right click and choose sort

VwInt	EntI	Dsolv	SoEI	mfScore	pValue	Percentile	RecConf	examplePDB
-14.36	0.5203	5.887	-0.9164	-47.54	<0.006133	0	1	2et1_a5bp_3np_a3nw_3unj_a0bx_3unz_a0bz_2xf1a16x_1afk

Step 9 Sort the table by pvalue to rank the fragments.

To cluster the fragments by location in the pocket. Select the "Location" column and click on the cluster button.

1. Single click on column header will select the column

2. Click on the cluster button

3. Drag here to change clustering distance

Location	ord	cl
2xf1a16x_1afk	15.08	37.98, 19.38

Step 10 To cluster the fragments by location in the pocket. Select the .Location. column and click on the cluster button.

fragment1_rec Molsoft icm 3.8-1 [NewProject *] (4 objects 1 table)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics Win

display light labels meshes search ligedit

Workspace Panel

no selection

- 1rek [1] ICM, 2.3A
- a 338 A 5 sites
- camp-dependent pro
- 1rek_1 [2] ICM, 2.3
- camp-dependent pro
- fragment1_fragments1
- m H C5H7NO
- fragment1_fragment:

All

fragment1_fragments1

mol

24

Filter Tree Tools

1211109876543210

C5H7NO

C6H14H

C6H15H2

C6H14H

C6H6H2

C2H4H4

C6H13NO

C6H5H5

413 Atoms 23 Res,3 Mol,3 Obj

2. Enter values as shown

1. Right click here and choose "select in cluster"

Select representative centers from the cluster tree. Selections will be highlighted in blue in the tree and the table.

Step 11 Select representative centers from the cluster tree by lowest p-value. Selections will be highlighted in blue in the tree and the table.

fragment1_fragments1_125 Molsoft icm 3.8-1 [NewProject *] (8 objects 1 table)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics Win

display light labels meshes search ligedit

Workspace Panel

1 Atoms 1 Res,1 Mol,1 Obj

- m H C7H7NO
- fragment1_fragment:
- stack (7 conformat
- fragment1_fragments1
- m H C9H11NO
- fragment1_fragment:
- stack (4 conformat
- fragment1_fragments1
- m H C9H11NO
- fragment1_fragment:
- stack (4 conformat
- fragment1_fragments1

All

Crystal structure ligand

Fragment 1

Fragment 2

Fragment 3

Header Filter Tree Tools

1211109876543210

C6H9H3

C7H8H5

C8H10

C7H8H0

C7H7H0

C8H10O2

C7H8H0

C5H8H2

C6H9H5

6 ICM Obj

Toggle through the selected fragments to choose lead candidates for linking.

Step 12 Toggle through the selected fragments to choose lead candidates for linking using the Ligand Editor.

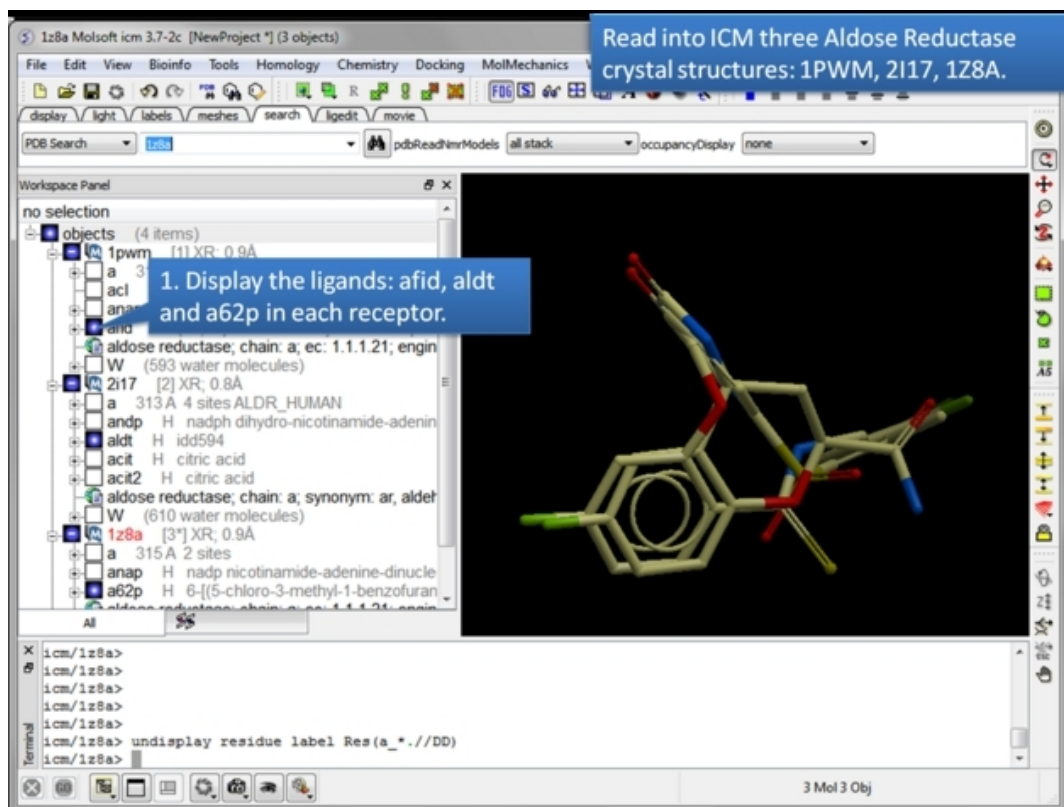
Reload Fragment Screening Results To reload the fragment screening results:

- Docking/Open Project and locate the directory where your fragment screening results are located.
- Select Docking/Make Hitlist and browse for the Docking_Project_Name_Fragments.ob file which contains the results of the screen.
- A table as shown in Step 8 will be displayed.

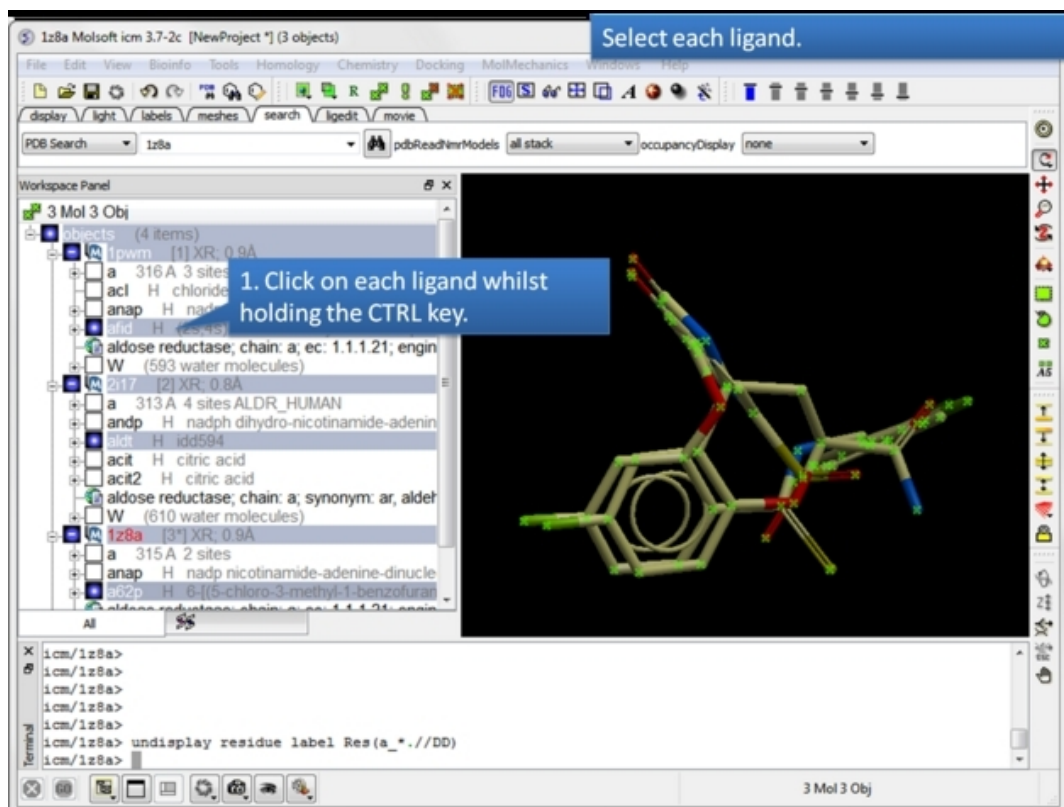
12.3 Ligand-Based Screening using APF 3D Pharmacophores

Available in the following product(s): ICM-VLS I

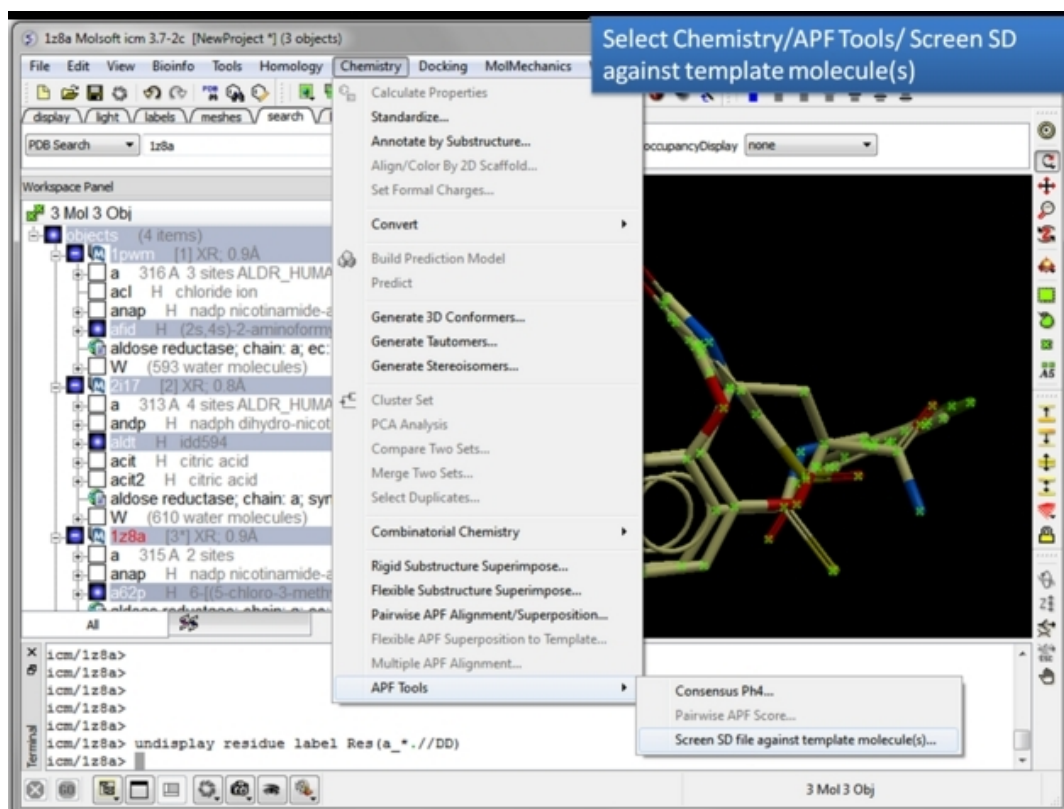
This option allows you to perform a ligand-based screen using the Atomic Property Fields (APF) of superimposed chemicals. SD files are screened and the ligands are scored according to their fit into the APF field. Read more about APF here: <http://www.molsoft.com/apf.html>



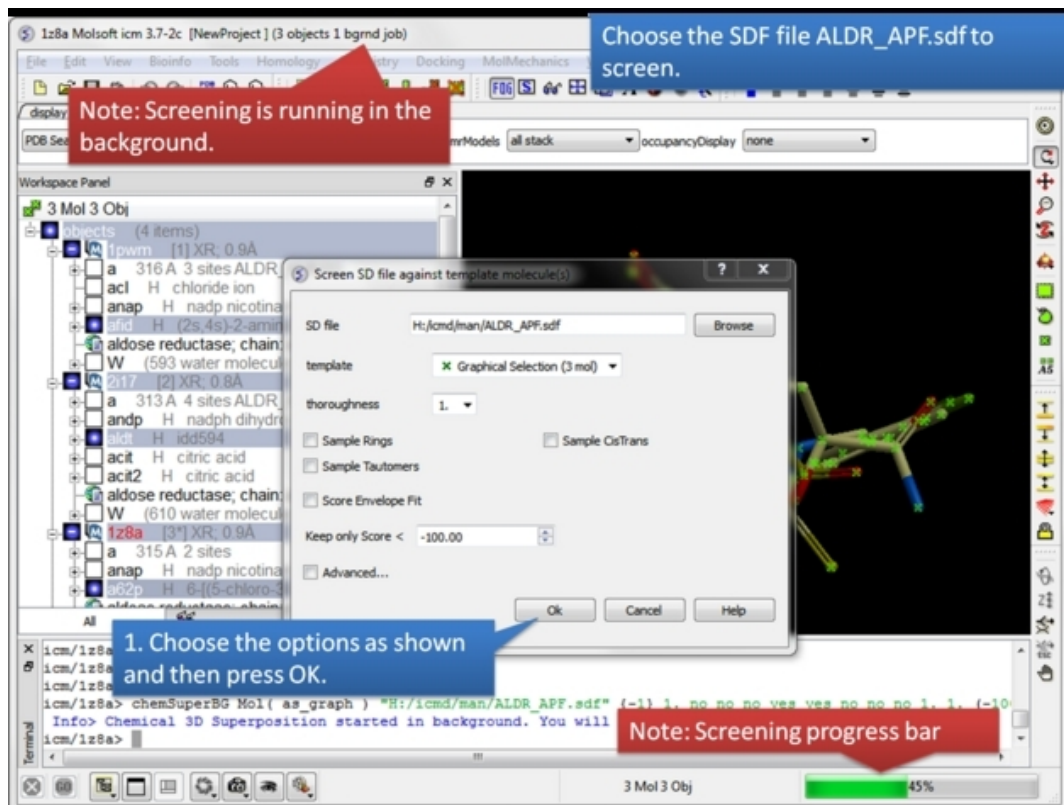
Read in and display the superimposed template chemicals. In this example we will use three aldose reductase crystal structures 1PWM, 2117, and 1Z8A. **IMPORTANT:** Convert each PDB file to an ICM object.



Select each chemical you wish to contribute to the APF fields. Select by double clicking on the chemical in the ICM workspace and hold down the CTRL key.



Choose the Screen SD file against template molecule(s) option. Chemistry/APF tools/Screen SD file against template molecule(s)



Select the SD database you wish to screen.

- Select the database in SD format you wish to screen and
- Select whether you want flexible rings to be sampled by checking the appropriate box.
- Select whether you want cis and trans conformations of double bonds to be sampled by checking the appropriate box.
- Select whether you would like to sample tautomers.
- A score for how well the ligand fits into the APF envelope can be generated. Select if desired.
- Scores Keep only Score < option.
- Select the **Advanced** button if you want the superposition to be weighted by occupancy of the atoms by checking the appropriate box. It is often desirable to preferentially superimpose parts of a ligand while ignoring other regions. This can be achieved by setting the occupancy to zero for regions you are not focusing on. See: command line manual <http://www.molsoft.com/man/icm-commands.html#set-occupancy> . You can also define the weights for lipophilic and polar matches.

You will be notified when the screening has completed.

Information

Chemical 3D Superposition finished. Press OK to load the result

Ok Details >>

1. Click OK

```

icm/1z8a>
icm/1z8a>
icm/1z8a> undisplay residue label Res(a.*./DD)
icm/1z8a> chemSuperBG Mol( as_graph ) "H:/icmd/man/ALDR_APPF.sdf" {-1} 1. no no no yes yes no no no 1. 1. (-10
Info> Chemical 3D Superposition started in background. You will be notified when the job is complete.
icm/1z8a>
  
```

3 Mol 3 Obj 100%

A notification will be displayed when the screen is completed. Click OK and a hitlist table will be displayed.

A fully interactive results table will be displayed.

1. Right click and sort by Score. The lower the score the better.

2. Browse the hitlist to find other Aldose Reductase inhibitors.

mol	L	Molecule Name	pdb	Score
1	3D	aldr	1us0	-135.2

4 Mol 4 Obj

A hitlist table will be displayed. You can sort the table by score and toggle the display on and off using the buttons in the L column.

13 MolScreen

Available in the following product(s): ICM-VLS I

MolScreen is a set of high quality 2D fingerprint and 3D pharmacophore models for a broad range of pharmacology and toxicology targets. The models can be used for lead discovery or counter screening. The models use MolSoft's 2D QSAR/Fingerprint and 3D Atomic Property Fields methods. The models can be screened directly using MolSoft's ICM-Pro + VLS software or via our contract research services. For more information please see: www.molsoft.com/molscreen.html

There are two screening panels in the Chemistry/MolScreen menu:

1. Brief Model Panel. This selection of models is provided free of charge from MolSoft.
2. Full Model Panel. Please contact MolSoft to use this this panel containing all available models.

In this chapter:

- MolScreen Model Types
- Brief Panel
- Full Model Panel
- Custom Model Panel
- Results Table
- Make Classification Model
- Make APF/SAR Model

13.1 Model Types

13.1.1 Kernel Chemical Classification/Regression (kcc) models

Hybrid Fingerprints

Dataset

IC50, EC50, Ki, Kd data was downloaded from ChEMBL18, they were combined and converted to pKd value. Compounds with pKd > 5, (i.e. 10 micro Molar) were classified as positives. 90% of ChEMBL18 compounds were assigned to training set. The remaining 10% compounds were assigned to external test set. To the training set, a subset of compounds with pKd>7. against any targets from ChEMBL18 was added as decoy. To the external test set, the approved drugs were added as decoy

Training

A Kernel Chemical Classification/Regression (kcc) model was trained using the training set compounds. The performance of the model was evaluated using the external test set

Predicted Value

The kcc model returns two scores:

The kcc score is the classification score: The positives have a median kcc score of 1. The decoys have a median kcc score of 0. The kcc score is renamed to MolClass in the pairwise table.

The kca score is the regression score of pKd value. A kca score of 6. indicates microMolar activities. The kca score is renamed to MolpKd in the pairwise table.

In the pairwise table, a pPvalue is calculated for both the kcc and kca score. The maximum pPvalue was taken as the final pPvalue for that model. The pPvalue indicates -Log of probability that the compound belongs to the random decoy. A pPvalue of 1 indicates the compound is comparable to the top 90 percentile decoy. A pPvalue of 2 indicates the compound is comparable to the top 99 percentile decoy.

13.1.2 Docking to Ligand Field Z-Score (dfz) models

APF

Dataset

If there is available data from ChEMBL18: IC50, EC50, Ki, Kd data was downloaded from ChEMBL18, they were combined and converted to pKd value. Compounds with pKd > 5. (i.e. 10 micro Molar) were classified as positives. 80% of ChEMBL18 compounds were assigned to training set. The remaining 20% compounds were assigned to external test set. To the training set, a subset of compounds with pKd>7. against any targets from ChEMBL18 was added as decoy. To the external test set, the approved drugs were added as decoy

If ChEMBL data is not available, approved drugs were used as decoy to calculate Z-Score.

Training

The dfz model was training in the following way:

1. For any target, all of its associated mammalian pocketome entries were used.
2. If ChEMBL data is not available, the pocketome entry with the highest number of co-crystallized ligand was used.
3. If ChEMBL training set is available, it was docked to each pocketome entries using APF method.
4. The best combination of template clusters were selected to maximize differentiation between actives and decoys.
5. Z-Score was calculated using approved drug's mean and standard deviation of APF score.

Predicted Value

The dfz model returns one score:

The dfz score is the Z-Score: A score of 1 means the compound is 1 standard deviation above the mean score of approved drugs decoy. The dfz score is renamed to MolZScore in the pairwise table.

In the pairwise table, a pPvalue is calculated for both the Z-Score. The pPvalue indicates -Log of probability that the compound belongs to the random decoy. A pPvalue of 1 indicates the compound is comparable to the top 90 percentile decoy. A pPvalue of 2 indicates the compound is comparable to the top 99 percentile decoy.

13.1.3 Docking to Ligand Field Classification/Regression (dfa) models

Hybrid 4D/2D

Dataset

IC50, EC50, Ki, Kd data was downloaded from ChEMBL18, they were combined and converted to pKd value. Compounds with pKd > 5, (i.e. 10 micro Molar) were classified as positives. 90% of ChEMBL18 compounds were assigned to training set. The remaining 10% compounds were assigned to external test set. To the training set, a subset of compounds with pKd>7. against any targets from ChEMBL18 was added as decoy. To the external test set, the approved drugs were added as decoy

Training

The dfa model is trained in the following steps:

1. Either: a. Training set was docked to the 4D maps of all the pocketome entries associated with the target. b. If pocketome entry is not available. The training set compounds were aligned in 3D.
2. Combining the training set compounds with pocketome co-crystallized ligands (if available), cluster in APF.
3. Subset of ligands were selected from each cluster as APF template.
4. All training set compounds were docked using APF method to all clusters.
5. The best combination of clusters were selected to maximize recognition of actives from decoys.
6. For each selected cluster, a pKd regression model was trained using the 3D poses of the ligands above a certain APF score cutoff.

Predicted Value

Any compound will be predicted using the dfa models in the following way: The compound will be docked using APF method to each of the template clusters. The compound will be assigned to the cluster that gives the highest normalized APF score. The 3D regression model of that cluster will then be used to predict the pKd value of that compound if the APF score is within the score cutoff.

The dfa model returns two scores:

The dfc score is the classification score: The positives have a median dfc score of 1. The decoys have a median dfc score of 0. The dfc score is renamed to MolClass in the pairwise table.

The dfa score is the regression score of pKd value. A dfa score of 6. indicates microMolar activities. The dfa score is renamed to MolpKd in the pairwise table.

In the pairwise table, a pPvalue is calculated for both the dfc and dfa score. The maximum pPvalue was taken as the final pPvalue for that model. The pPvalue indicates -Log of probability that the compound belongs to the random decoy. A pPvalue of 1 indicates the compound is comparable to the top 90 percentile decoy. A pPvalue of 2 indicates the compound is comparable to the top 99 percentile decoy.

13.1.4 Docking to Protein Pocket Classification/Regression (dpc) models

Hybrid Docking

Dataset

IC₅₀, EC₅₀, K_i, K_d data was downloaded from ChEMBL18, they were combined and converted to pK_d value. Compounds with pK_d > 5, (i.e. 10 micro Molar) were classified as positives. 90% of ChEMBL18 compounds were assigned to training set. The remaining 10% compounds were assigned to external test set. To the training set, a subset of compounds with pK_d>7. against any targets from ChEMBL18 was added as decoy. To the external test set, the approved drugs were added as decoy

Training

The dpc model is trained in the following steps:

1. The pocketome entry with the most co-crystallized ligands associated with the target is selected
2. The residues around the pocket were clustered, selected representative PDBs were retained
3. All training set compounds were docked to the 4D maps of the pocket in the presences of the co-crystallized ligands in the form of APF template
4. A score cutoff to maximize sensitivity and accuracy was selected
5. All compounds within that score cutoff were used to train a pKd prediction model using their 3D poses.

Predicted Value

Any compound will be predicted using the dpc models in the following way: The compound will be docked to the 4D maps of the pocket in the presence of APF template. The 3D regression model of the pocket will then be used to predict the pKd value of that compound if the docking score is within the score cutoff.

The dpc model returns two scores:

The dpc score is the classification score: The positives have a median dpc score of 1. The decoys have a median dpc score of 0. The dpc score is renamed to MolClass in the pairwise table.

The dpa score is the regression score of pKd value. A dpa score of 6. indicates microMolar activities. The dpa score is renamed to MolpKd in the pairwise table.

In the pairwise table, a pPvalue is calculated for both the dpc and dpa score. The maximum pPvalue was taken as the final pPvalue for that model. The pPvalue indicates -Log of probability that the compound belongs to the random decoy. A pPvalue of 1 indicates the compound is comparable to the top 90 percentile decoy. A pPvalue of 2 indicates the compound is comparable to the top 99 percentile decoy.

13.2 Brief Panel

To use the Brief Model Panel:

- Read in a set of chemicals you wish to screen against the MolScreen panel (e.g File/Open sdf).
- Select Chemistry/MolScreen/Brief Model Panel
- Select the `model` type you would like to screen from the tab menu.
- Use the dropdown button to locate the chemical table you wish to screen
- Use the check boxes to select the targets you would like to screen or choose select all.
- There are some options which can be selected: **Color** will color the table by a rainbow scaled to the value range in the table. **LookupOnly** will search your chemicals to see if they are contained in the training and test set. **deleteEmptyCols** will clean up your results table by removing columns that do not have data. **PairwiseTable** will generate a model and target pairwise table.
- The `results` will be displayed in a separate table.

13.3 Full Model Panel

To run a full panel screen:

- Read in a set of chemicals you wish to screen against the MolScreen panel (e.g File/Open sdf).
- Select Chemistry/MolScreen/Full Model Panel
- Select the `model` type you would like to screen from the tab menu.
- Use the dropdown button in the table **Tools** panel (right hand side) to locate the chemical table you wish to screen.
- There are some options which can be selected: **Color** will color the table by a rainbow scaled to the value range in the table. **LookupOnly** will search your chemicals to see if they are contained in the training and test set. **deleteEmptyCols** will clean up your results table by removing columns that do not have data. **PairwiseTable** will generate a model and target pairwise table.
- Select the targets you wish to screen against by selecting the rows in the table. Click on the row number to select the row (the row will then be shaded blue). Use the CTRL key and click to select a non-contiguous range or use Tab to select a contiguous range of models.
- Click **Apply Selected Models**
- The `results` will be displayed in a separate table.

To view the performance of a model select the model by selecting the row number and then choose **Check Model Performance** in the Tools panel of the table.

13.4 Custom Model Panel

MolScreen has some preprepared model panels available. To select a preprepared custom model panel:

- Chemistry/MolScreen/Select Custom Model Panel
- Choose the panel from the drop down Panel Name dialog box.

- Select the path where your MolScreen models are located (usually in .icm/models).

To create your own custom model panel:

- Chemistry/MolScreen/Make Custom Model Panel
- Give your panel a unique name.
- Enter a uniprot prefix (e.g. ESR1, ANDR)
- Enter a keyword (e.g. nuclear)
- Press OK and a table will be displayed containing the models.
- You can repeat the steps above to append to the custom model panel.
- Save your custom panel in your models directory e.g. (.icm/model)

To load your custom panel:

- Chemistry/MolScreen/Full Model Panel and check the option **Read Custom Panels**

13.5 Results Table

The results table will contain the columns from the chemical table that you screened plus the following results columns: Depending on your selection you might have two tables **Model Pairwise** and **Target Pairwise**

- **mol** - your query chemical
- **target** - the target model you screened against
- **ppValue** - kcc/hybrid4D+2D models only: -Log(probability) relative to approved drug decoys capped at 5: 1: Top 10 percent 2: Top 1 percent 3: Top 0.1 percent
- **MolpKd** - kcc/hybrid4D+2D models only: Prediction of pKd, pKi, pIC50, pEC50, etc >6. indicates submicromolar value, >9. indicates nanomolar, ND indicates dissimilarity to known binders
- **MolpKd Error** - kcc/hybrid4D+2D models only: Estimated total uncertainty/error for pKd prediction, including experimental error of ~1. pKd unit
- **MolClass** - kcc/hybrid4D+2D models only: Classification Score: for kcc models: nProbability of being a positive, assume 0.5% hitrate for others: 1. Median value of actives 2. Median value of inactives
- **TANsim** - kcc models only: Tanimoto distance from model nearest compound
- **modelnearestCpd** - kcc models only: nearest compound to the query in the model
- **MolZScore** - apf models only: Z-Score 3. indicates 3 standard deviation above approved drugs decoy
- **predictType** - model type
- **modelAUC** - Area Under Curve (AUC) for model
- **modelpKdAUC** - kcc models only: AUC for pKd model
- **modelRMSE** - kcc models only: RMSE for model
- **modelQ2** - kcc models only: Q2 for model
- **MolDockScore** hybrid docking only - ICM docking score, Score < -32. are in general significant, but may vary from structure to structure.

Fold Columns to Original Table

In the **Tools** panel of the table there is a convenient way of packaging your results into your original query table.

13.6 Make Chemical Classification Model

This option allows you to make your own kcc models either on your own data (local table) or from ChEMBL.

From your own data:

- Read in a chemical spreadsheet containing your chemical 2D structure (mol column) and an activity column containing data such as pKd, mM, uM or nM.
- Select Chemistry/MolScreen/ Make Chemical Classification Model
- Select the **Local Table** tab.
- Give your model a unique name.
- Find your chemical table from the drop down list.
- Find the column name containing **activity** data from the drop down list.

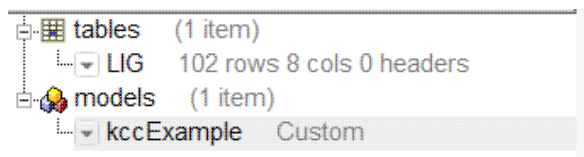
- Select the activity unit pKd, mM, uM or nM.
- Select whether you would want pKa charge prediction.
- Click OK.

From ChEMBL data:

- Give your model a unique name.
- Enter a Uniprot ID (e.g. 5ht1f)
- Select whether or not you want to include all mammalian data or just human.
- Check makemodel to make the model immediately.

How to use the model to predict:

- When the model has been built you will see the model listed in the ICM workspace as shown below.



To run the model against a set of chemicals:

- Read in a chemical spreadsheet containing your chemical data.
- Right click on the model in the ICM workspace and choose **Predict**.
- Enter the table name and prediction model you would like to use.
- Once the prediction has finished you will see three additional columns: the *k_{cc}* value, *k_{ca}* value and the tanimoto distance to the nearest compound in the model dataset.

13.7 Make APF Docking SAR Model

This option allows you to make your own *dfa* model either on your own data (local table) or from ChEMBL. If the chemicals are not superimposed then this option will automatically superimpose them. If you want to superimpose yourself before running then the chemical superposition options are described [here](#).

From your own data:

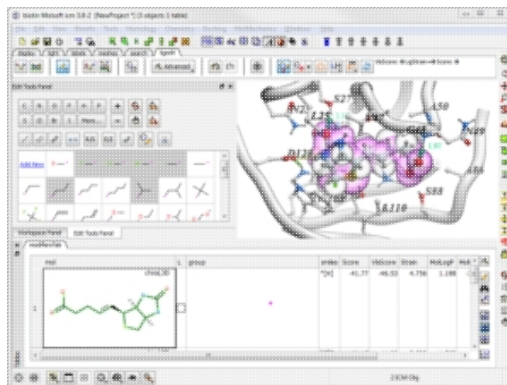
- Read in a chemical spreadsheet containing your chemical 2D structure (mol column) and an activity column containing data such as pKd, mM, uM or nM.
- Select Chemistry/MolScreen/ Make Chemical Classification Model
- Select the **Local Table** tab.
- Give your model a unique name.
- Find your chemical table from the drop down list.
- Find the column name containing **activity** data from the drop down list.
- Select the activity unit pKd, mM, uM or nM.
- Select whether you would want pKa charge prediction.
- Click OK.

From ChEMBL data:

- Give your model a unique name.
- Enter a Uniprot ID (e.g. 5ht1f)
- Select whether or not you want to include all mammalian data or just human.
- Check makemodel to make the model immediately.

14 The 3D Fully Interactive Ligand Editor

The ligand editor is a powerful tool for the interactive design of new lead compounds in 3D. It allows you to make modifications to the ligand and see the affect of the modification on the ligand binding energy and interaction with the receptor.



Watch a webinar about the 3D ligand editor [here](#).

There is a tutorial on how to use the 3D ligand editor [here](#).

Chapter Contents:

- Setup Ligand and Receptor
- Ligand Editor Preferences
- Display Options
- Evaluate Ligand Score and Strain
- Edit Ligand
- Screen Replacement Group
- Dock or Minimize Ligand
- Dock Chemical Table
- Fragment Linking
- Impose Tethers and Distance Restraints
- Defined Flexible Receptor Groups
- Side Chain Refinement
- Dock to Multiple Receptor Conformations
- Covalent Docking
- Dock to Atomic Property Fields
- Save Docked Ligand as PDB
- Export Docking Project
- Close and Clean Up Ligand Editor Project

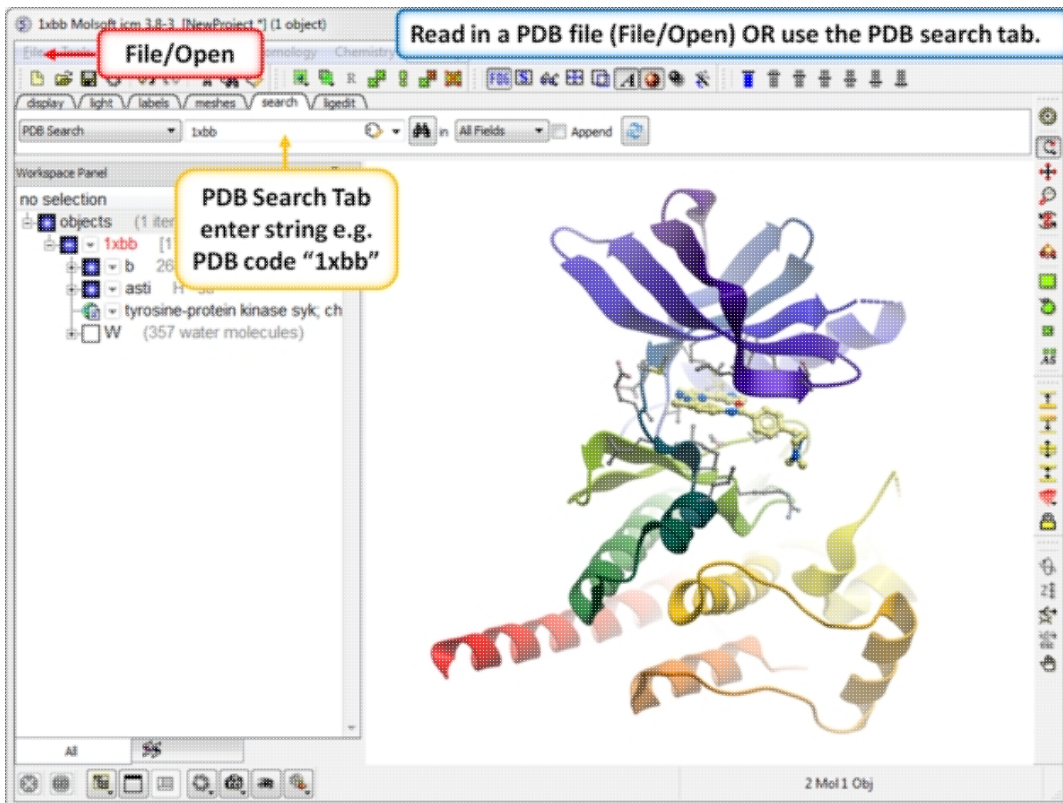
14.1 Setup Ligand and Receptor

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

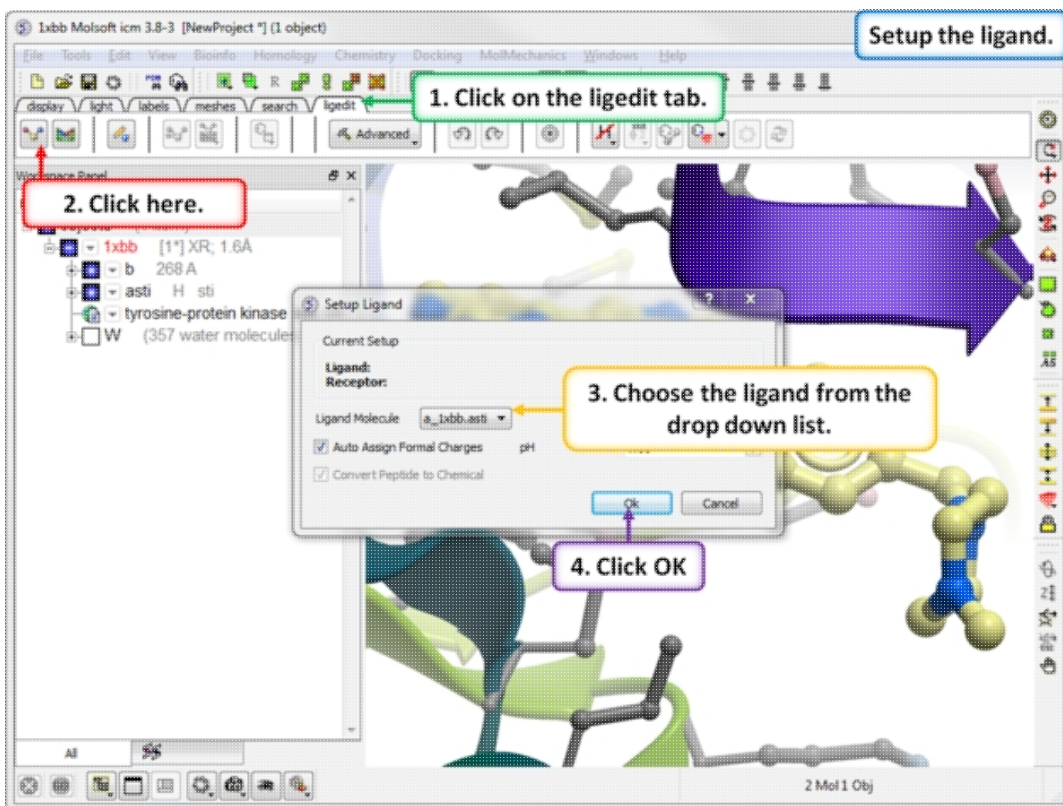
Video |

As an example we will use the kinase structure 1xbb.

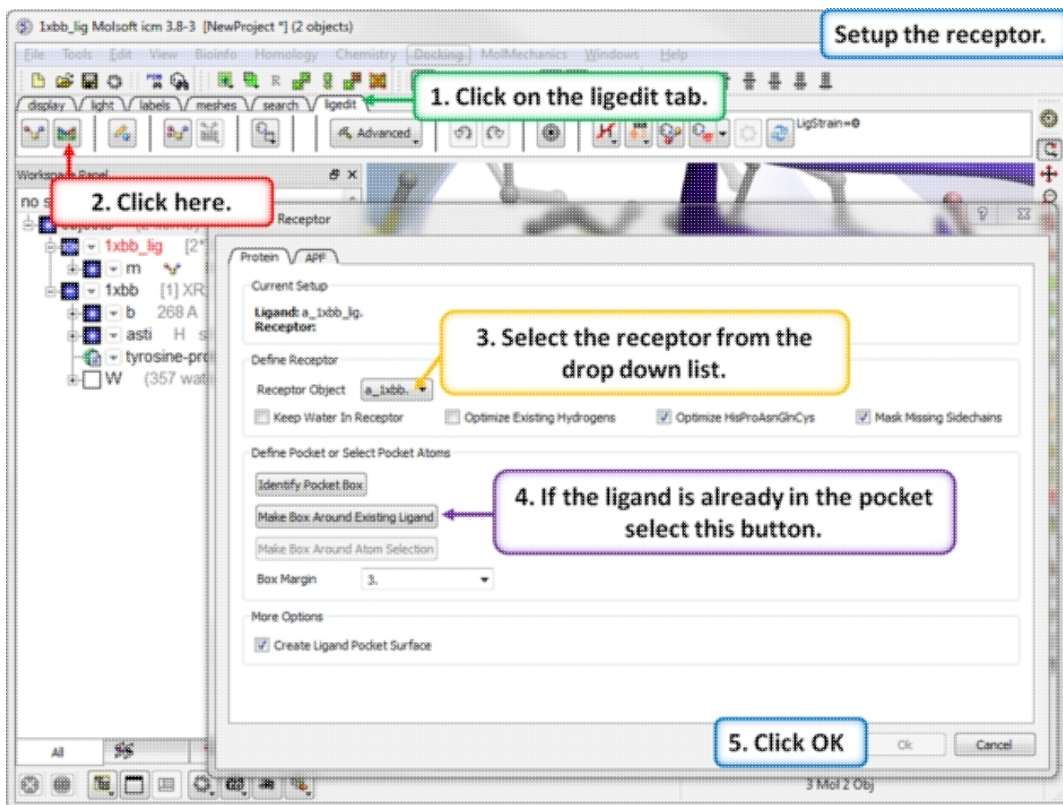
- Click on the search tab and type 1xbb.



- Click on the **ligedit** tab
- Click on the **Setup Ligand** button.
- Enter the ICM selection language for the **Ligand Molecule** (a_1xbb_ast) or use the drop down button to locate it.
- Select whether or not you would like ICM to automatically assign formal charges and set the pH.



- Select the **Receptor Setup** button.
- Enter the ICM selection language for **Receptor Object** (a_1xbb.) or use the drop down button to locate it.
- Select whether or not you wish to keep all waters in the receptor, keep tight waters, delete all waters, optimize existing hydrogens and optimize His, Pro, As, Gln, and Cys. The "mask" option will hide atoms not reported in the crystal structure (e.g. occupancy 0). If this box is not checked ICM will build the missing atoms of the side-chain.
- Click on the option to select **Box Around Existing Ligand**. There are other options: **Identify Pocket Box** will run ICMpocketFinder and return a table of pockets. Click on the table to select the pocket you want and then press OK. You can also **Make Box Around Atom Selection**.
- Enter a box margin of 3. This option defines the size of the energy maps around the ligand. The value of 3. should encompass the whole site but if you have a binding pocket that is very elongated or unusual in any way it is recommended that you check that the purple box covers the site you are interested in.



- At this stage you can change the size of the box by dragging on the corners of the purple box in the graphics display. Generally the default box is the appropriate size so no changes are needed.
- Click OK and the energy maps will be generated.

The energy maps are made inside the purple box.

Prior to pressing OK in the receptor setup step. You can adjust the size of the box by dragging on the corners.



- Once everything is setup you will see the ligand and receptor objects along with the original pdb file in the ICM workspace.

Everything is now setup and ready to start editing.

Click here to undisplay.

Ligand to be edited.

Receptor.

Ligand pocket surface.

Original PDB structure

Ligand to be edited.

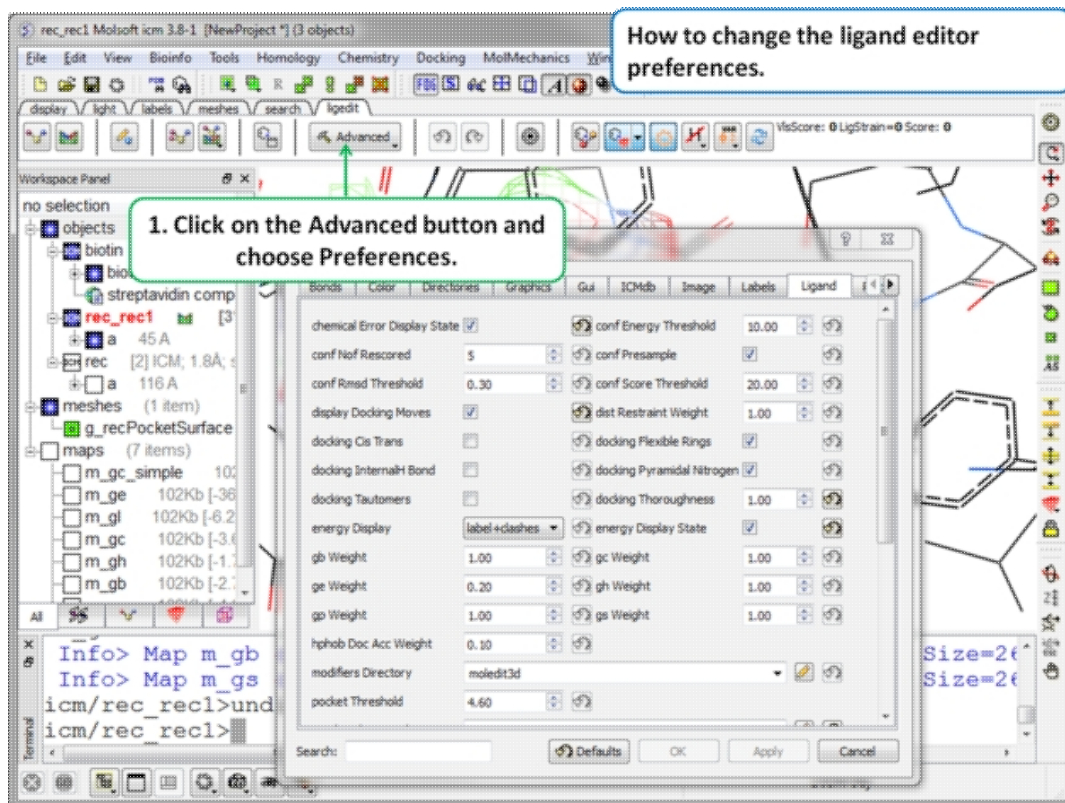
Ligand pocket surface.

- If the ligand-receptor complex is a PDB structure and hydrogens have been added and optimized then it is important to run a minimization after setup.

14.2 Ligand-Editor-Preferences

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

This step is optional but you may want to tweak the default preferences a bit or add new modifiers and linkers. You can change the preferences by clicking on the **Advanced** button in the ligedit tab and then choose "Setup Ligand Editor Preferences".

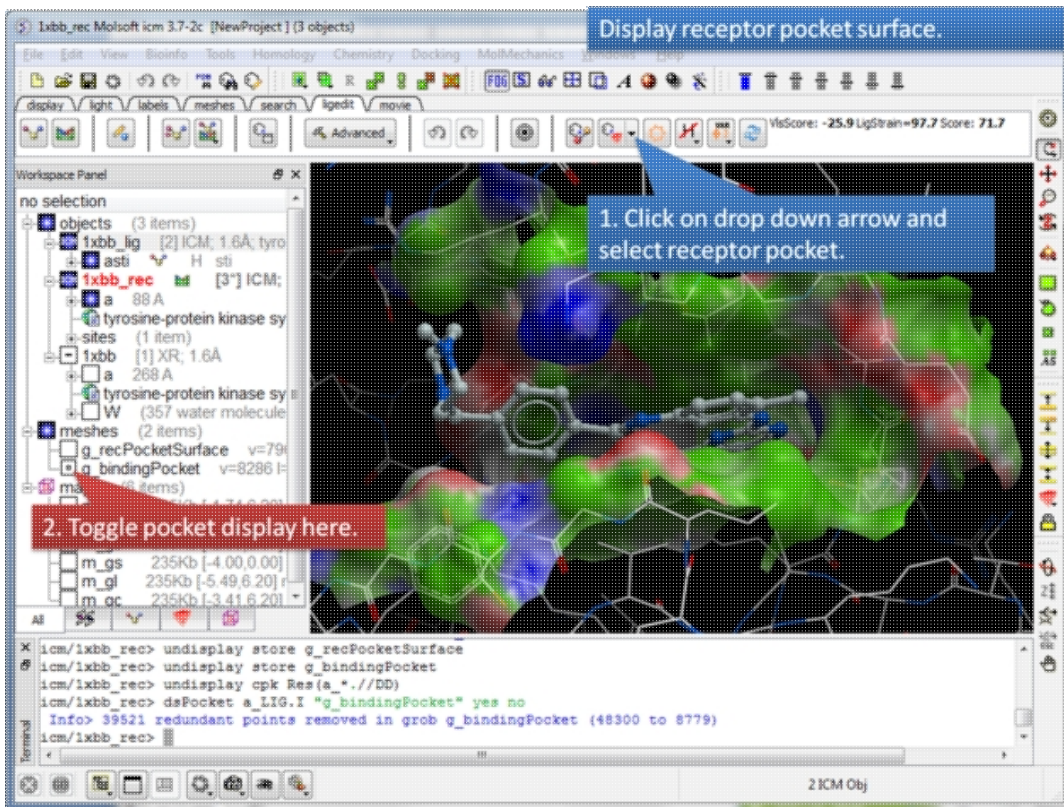


14.3 Display Options

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

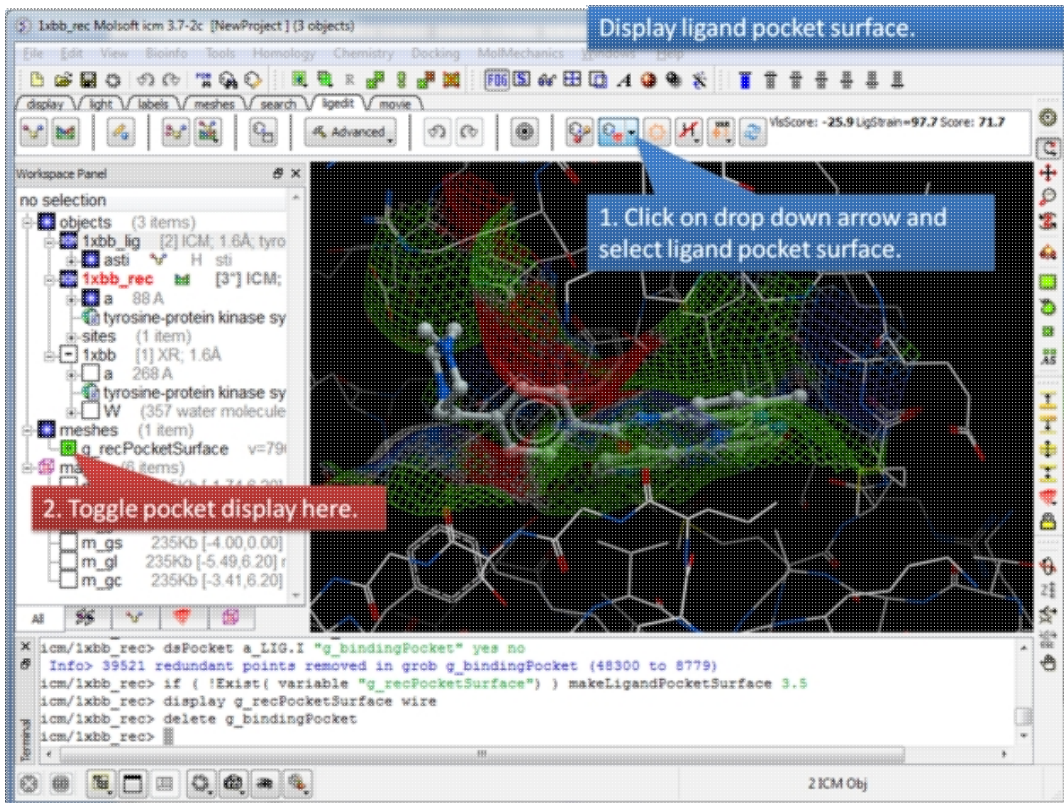
Video |

14.3.1 Display the receptor pocket surface



About the Receptor Pocket Surface: White=neutral surface Green=hydrophobic surface Red=hydrogen bonding acceptor potential Blue=hydrogen bond donor potential

14.3.2 Display the ligand pocket surface

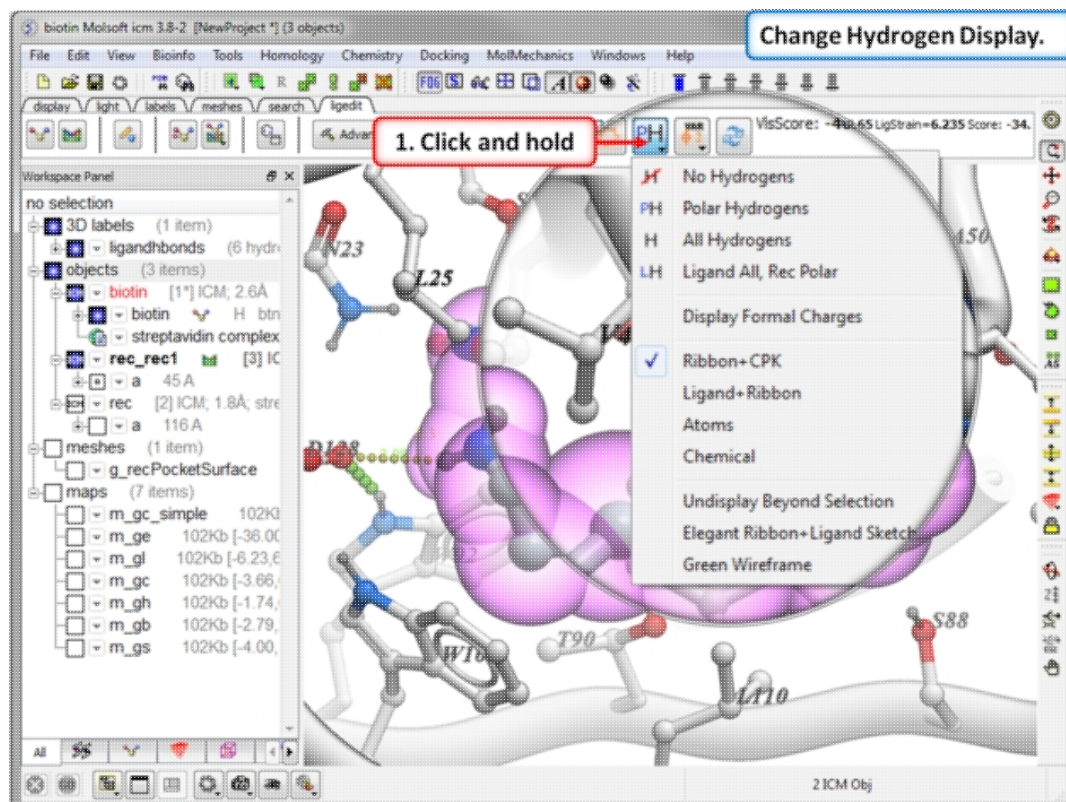


About the Ligand Pocket Surface: White=neutral surface Green=hydrophobic surface Red=hydrogen bond acceptor potential Blue=hydrogen bond donor potential

14.3.3 Display Hydrogen Atoms

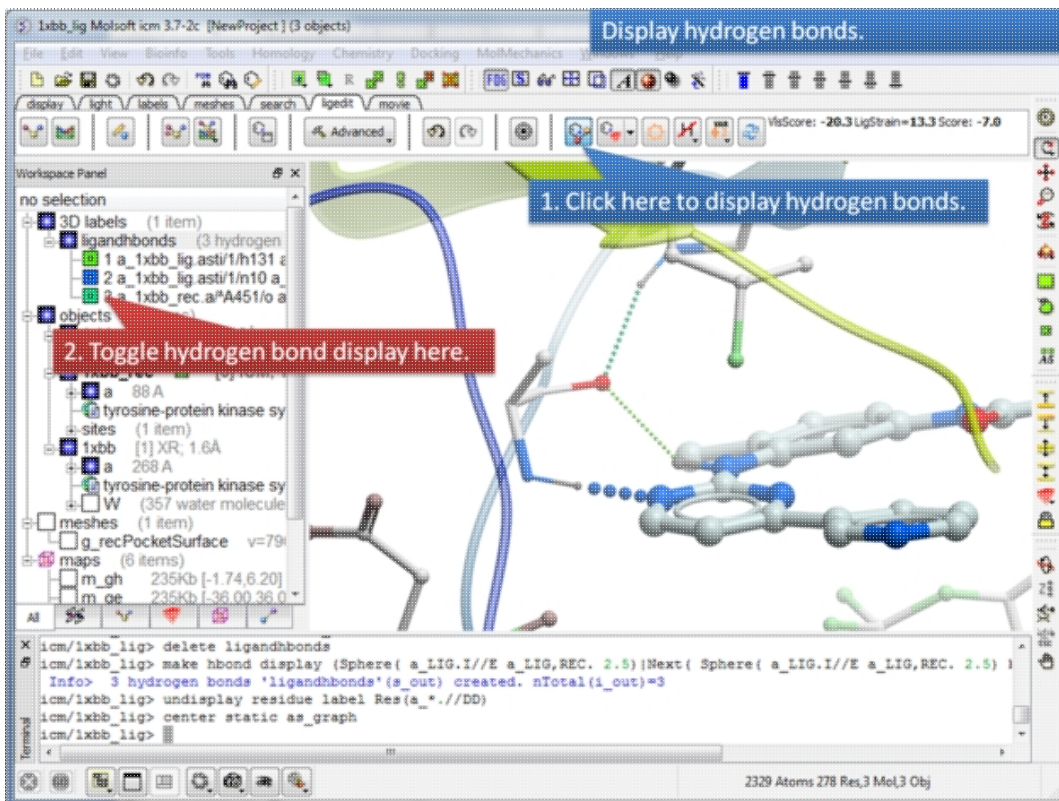
To display hydrogen atoms:

- Single click on the hydrogen atom display button will toggle the display between: no hydrogens, polar hydrogens, all hydrogens, all hydrogens in ligand and polar hydrogens in receptor.
- Click and hold on the button and you can choose the option rather than using the toggle.



14.3.4 Display Hydrogen Bonds

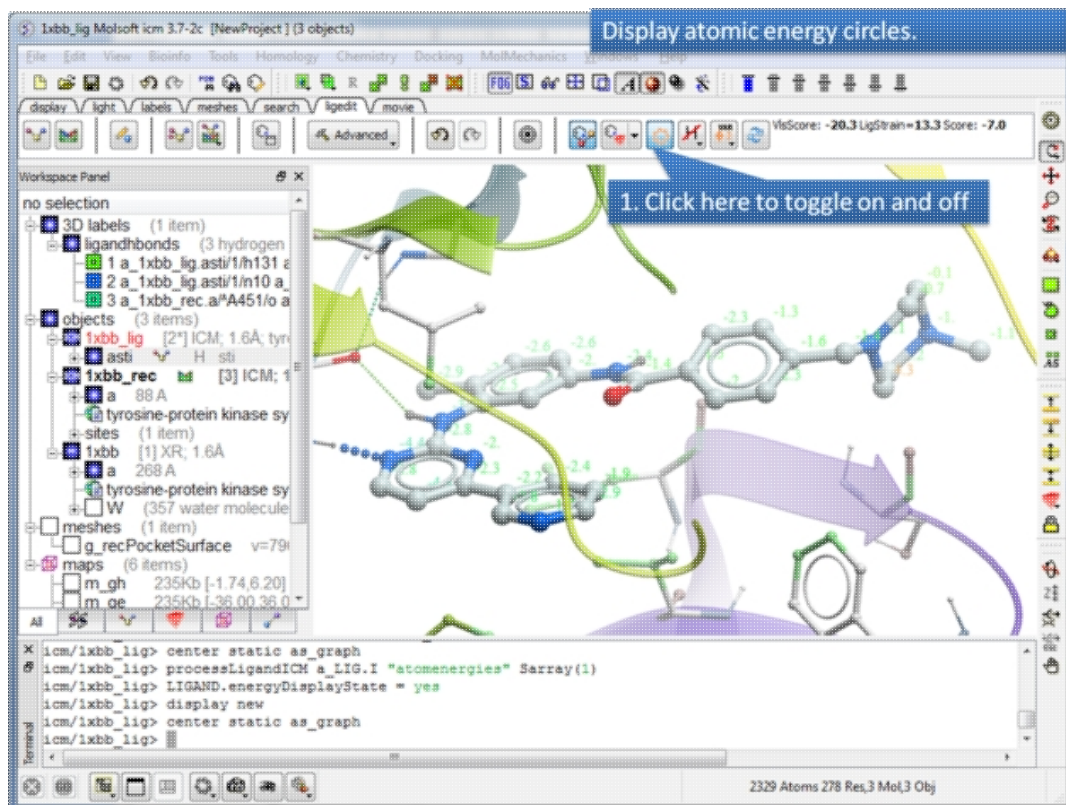
To display hydrogen bonds:



About Hydrogen Bonds: The coloring of the H-bonds are red (strong - thick spheres) to green (intermediate) to blue (weak - thin spheres). Once the hydrogen bonds have been displayed they can be displayed and undisplayed in the **3D labels** section of the ICM Workspace (left hand side of graphical window).

14.3.5 Display Atomic Energy Circles

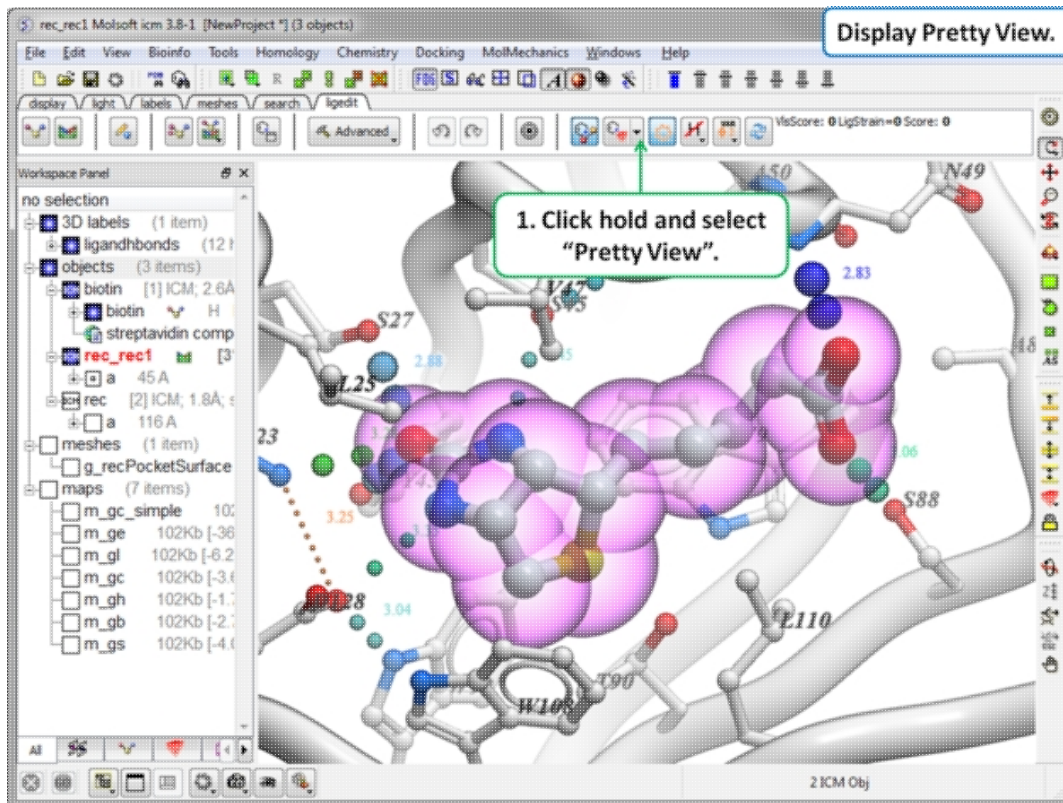
To display atomic energy circles:



About the Atomic Energy Circles: Good ligand-receptor interactions are highlighted by green spheres. Poor energy interactions are displayed as orange-->red stars - red being a major clash and a very poor energy contribution. Each atom is given an energy value relating to its contribution to the total receptor-ligand interaction energy. Low values colored green are considered favorable.

14.3.6 Display Pretty View

To display the default "Pretty View":



- Set up the ligand and receptor.
- Click and hold on the button shown below and choose "Pretty View".

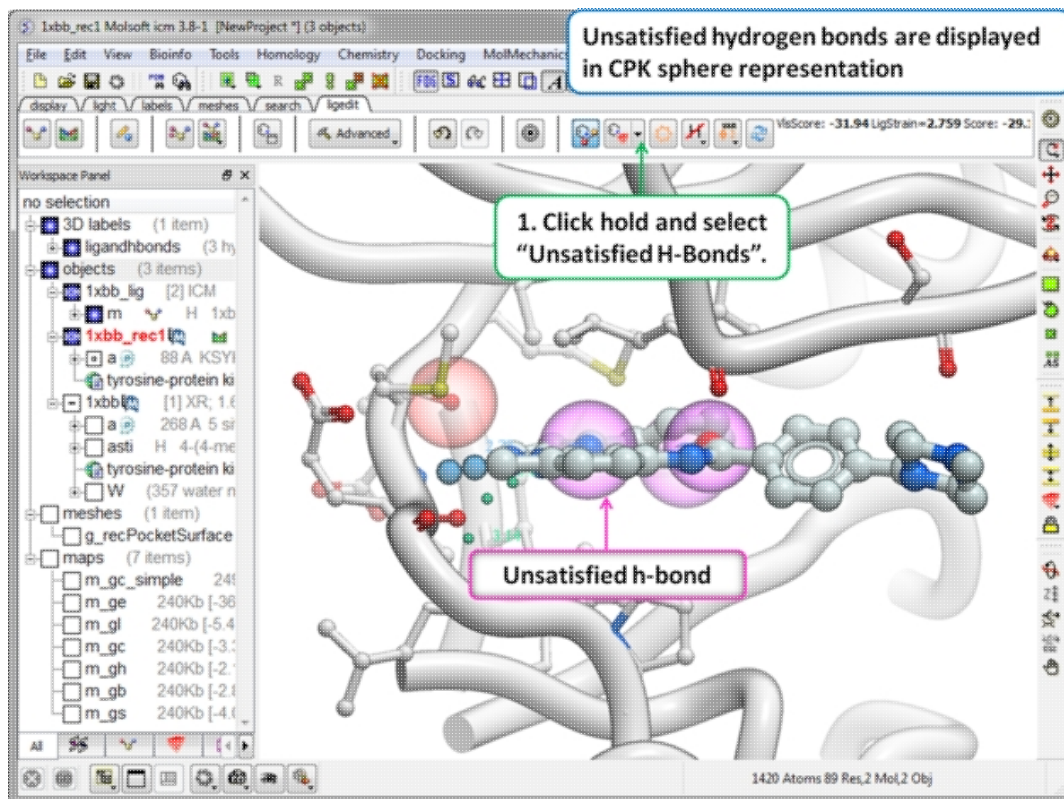
14.3.7 2D Interaction Diagram

To create a 2D interaction diagram - click on the 2D interaction diagram button shown below. More details regarding this option can be found [here](#).



14.3.8 Unsatisfied Hydrogen Bonds

To display unsatisfied hydrogen bonds:

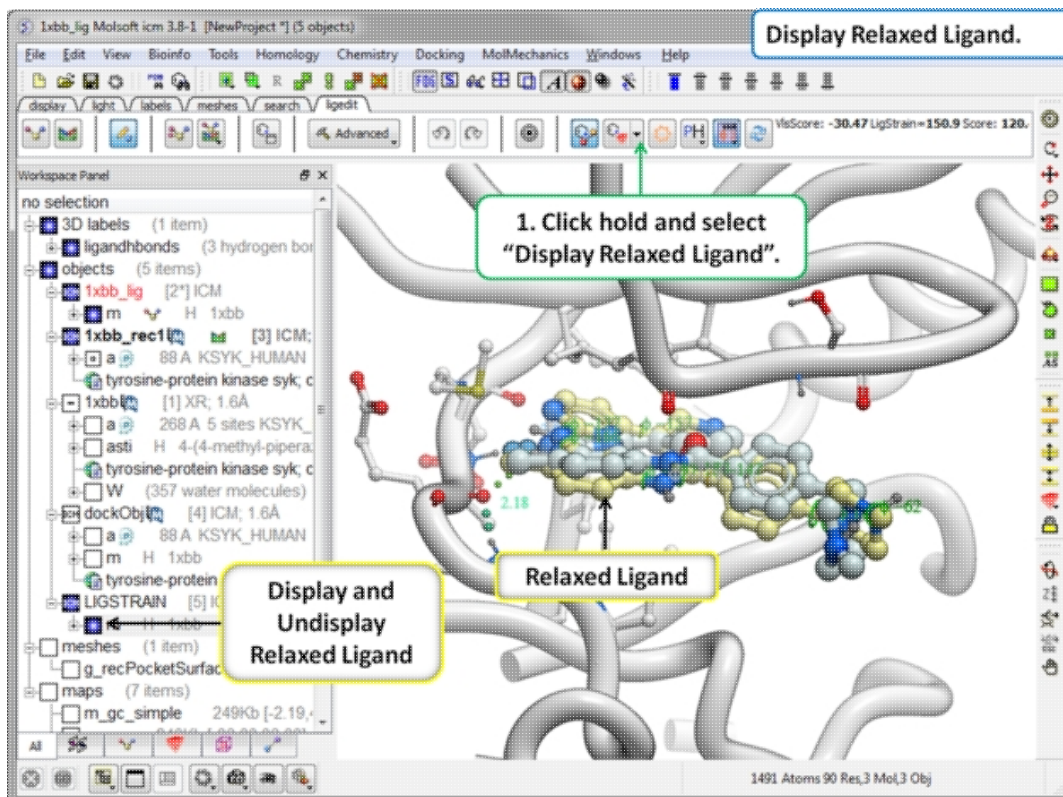


- Setup the ligand and receptor.
- Click and hold on the button shown below and choose "Unsatisfied H-bonds".

14.3.9 Display Relaxed Ligand

The relaxed ligand is the closest local minimum after minimization in the Merck Molecular Force Field (MMFF) outside of the receptor.

To display the relaxed ligand:



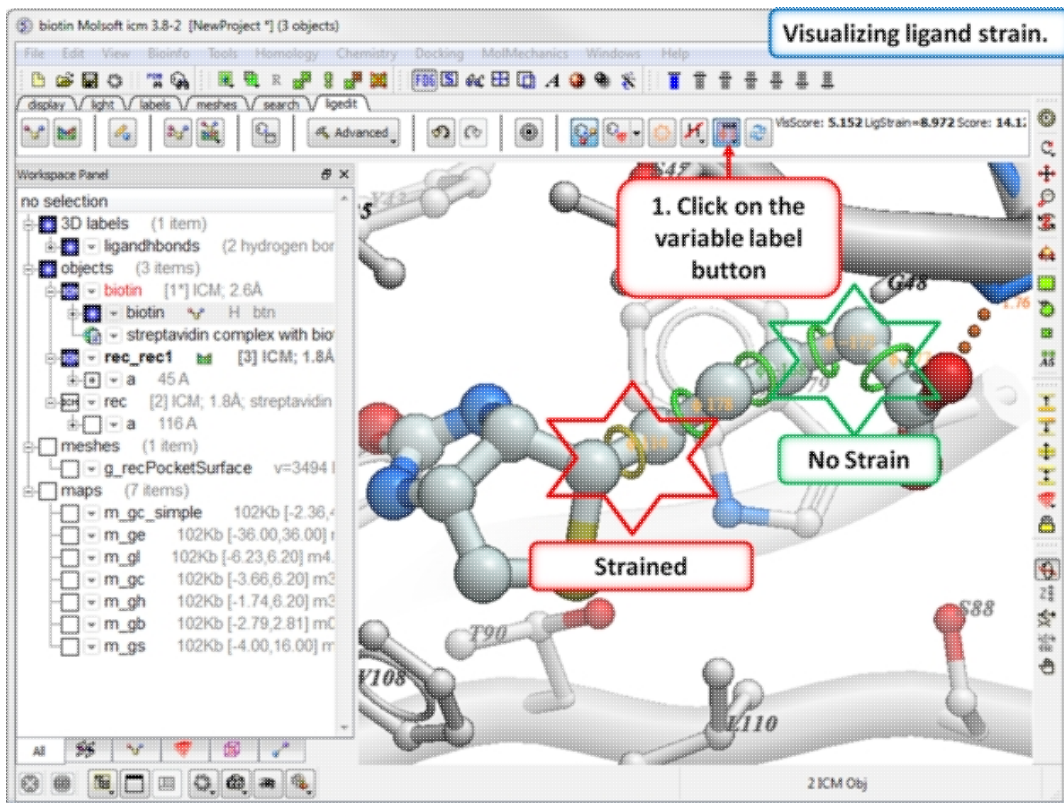
- Setup the ligand and receptor.
- Click and hold on the button shown below and choose "Display Relaxed Ligand".

14.3.10 Visualize Ligand Strain

If you are seeing a high ligand strain value when you calculate the score you can visualize the cause of this by viewing the variable angles.

To view the variable angles:

- Click on the variable angle button in the ligand editor tool panel.



Rings of varying diameter and color are superimposed on rotatable bonds. Green rings with large diameter are considered less constrained than rings with small green/yellow/red rings. Red rings are highly constrained. You can also compare your strained ligand with the relaxed ligand.

14.3.11 Identify Ligand-Receptor Contacts

To display ligand-receptor contacts:

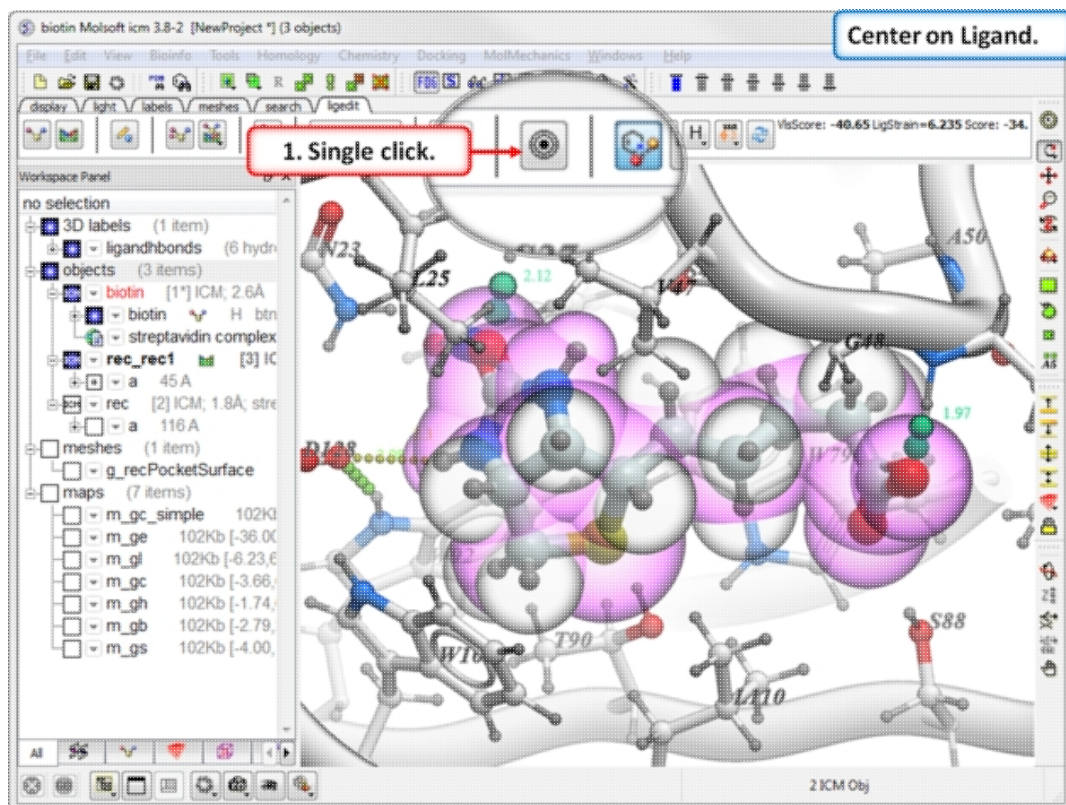
- Click on the button shown below.
- Select "Analyze Contacts".
- The xstick display in the region will be scaled according to the atom/residue contact area. For example residues making large contacts with a ligand will be displayed in thicker xstick representation (and colored yellow) than those making less significant contacts.
- A table as shown below will be displayed. The table lists the contact area, exposed area, percentage of contact area compared to exposed, the nearest atom of a residue and the distance.

A	ContactArea	ExposedArea	PercentBuried	Closest Atom	Closest Dist
1	biotin.biotin/300	260.7	340.7	77 a_rec_rec1.a/vai/co2	2.099
2	rec_rec1.a/^W79	31.1	74.5	42 a_biotin.biotin/btn/cs1	3.165
3	rec_rec1.a/^L110	25.8	60.1	43 a_biotin.biotin/btn/cs	3.042
4	rec_rec1.a/^V47	25.0	86.5	29 a_biotin.biotin/btn/c4	2.27
5	rec_rec1.a/^H49	18.4	94.7	19 a_biotin.biotin/btn/o1	1.834
6	rec_rec1.a/^W108	17.8	96.0	19 a_biotin.biotin/btn/c6	3.297
7	rec_rec1.a/^L25	16.0	102.9	16 a_biotin.biotin/btn/n1	2.766
8	rec_rec1.a/^T90	14.5	19.6	74 a_biotin.biotin/btn/s1	2.076
9	rec_rec1.a/^G48	12.0	66.9	18 a_biotin.biotin/btn/c10	2.93

14.3.12 Center on Ligand

To center on a ligand in the graphical display:

- Click on the center button in the ligedit tab.



14.4 Evaluate Ligand Score and Strain

Video |

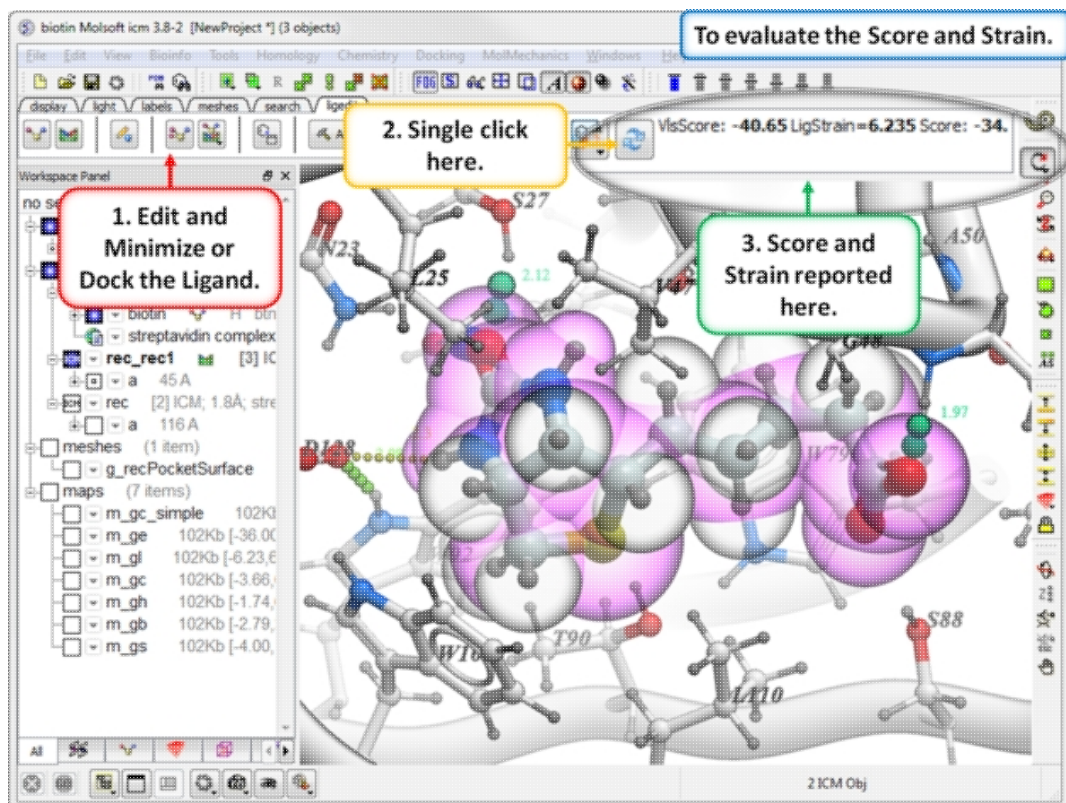
After you have edited a ligand and re-docked or minimized it you can evaluate the Score and Strain for the ligand. Ideally a good modification to your ligand will lower the score.

About the Score

- **VlsScore** = the ICM-VLS score used for ligand screening. The lower the Score the better the predicted interaction. VLS score is available with ICM-Pro + VLS and ICM-Chemist-Pro products.
- **LigStrain** = The strain of the ligand. Lower values = less strain.
- **Score** = VlsScore + LigStrain. The lower the Score the better the predicted interaction.

To evaluate the Score and Strain:

- Click on the "Recalculate Binding Score Button".



14.5 Edit Ligand

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

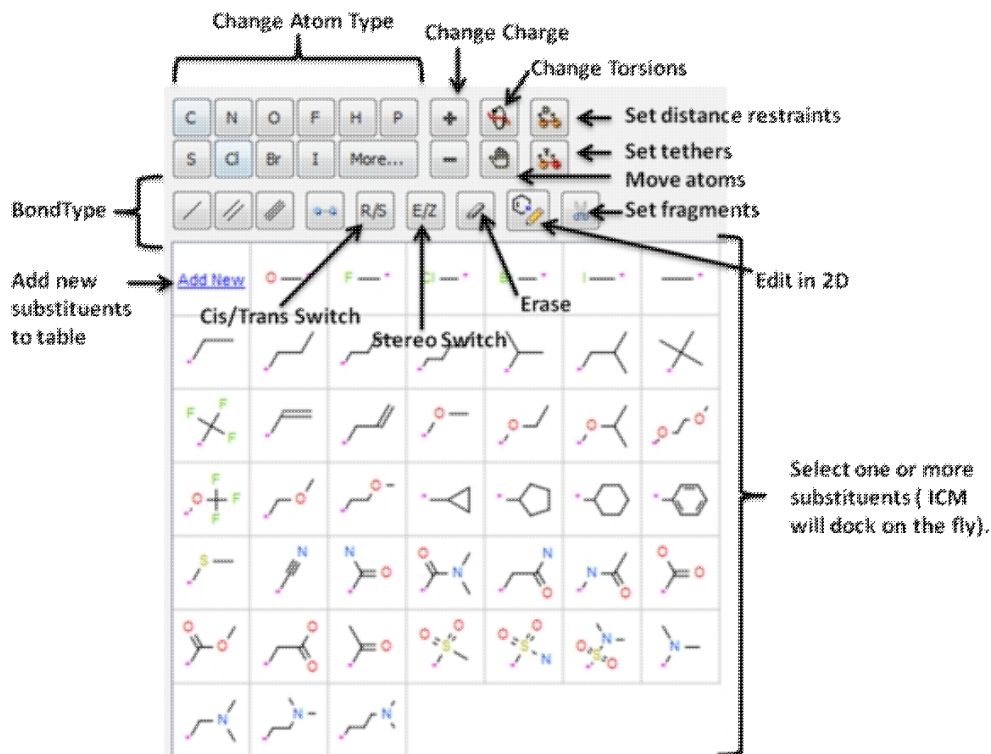
Video |

To edit a ligand:

- Setup the ligand and receptor.
- Display hydrogens using the hydrogen display button.
- Click on the **Edit Tools Button** and a panel of buttons will be displayed as shown below.
- To edit an atom or bond, first click on the desired atom, group or bond in the panel and the click on the atom or bond which you want to modify in the graphical display.
- Once you have made a modification you can evaluate the score and judge how good the modification was.



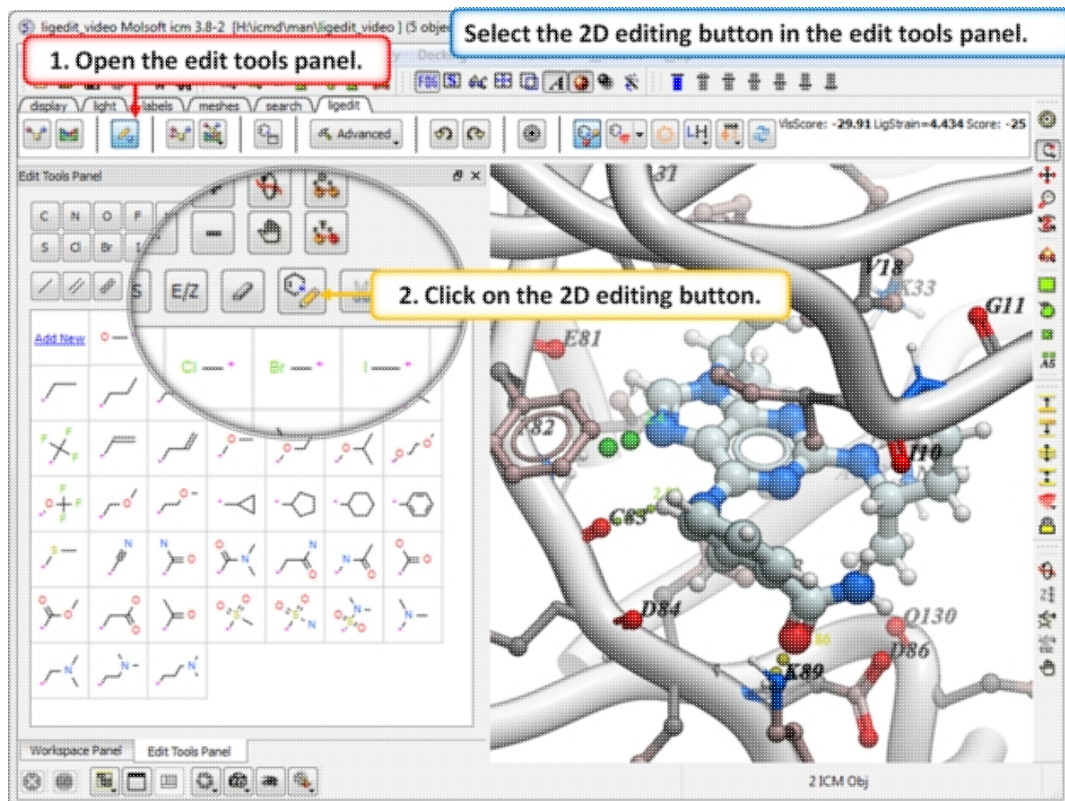
- See graphic below to understand what each button does.



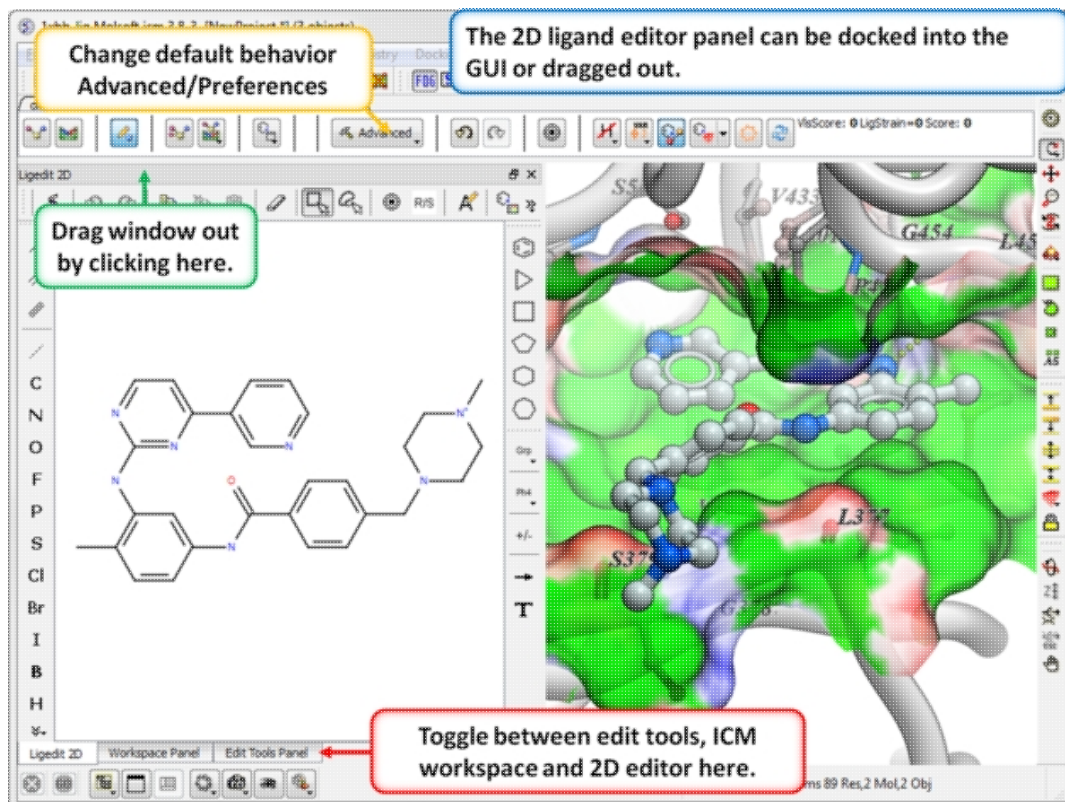
14.5.1 Edit Ligand in 2D.

To edit the ligand in 2D:

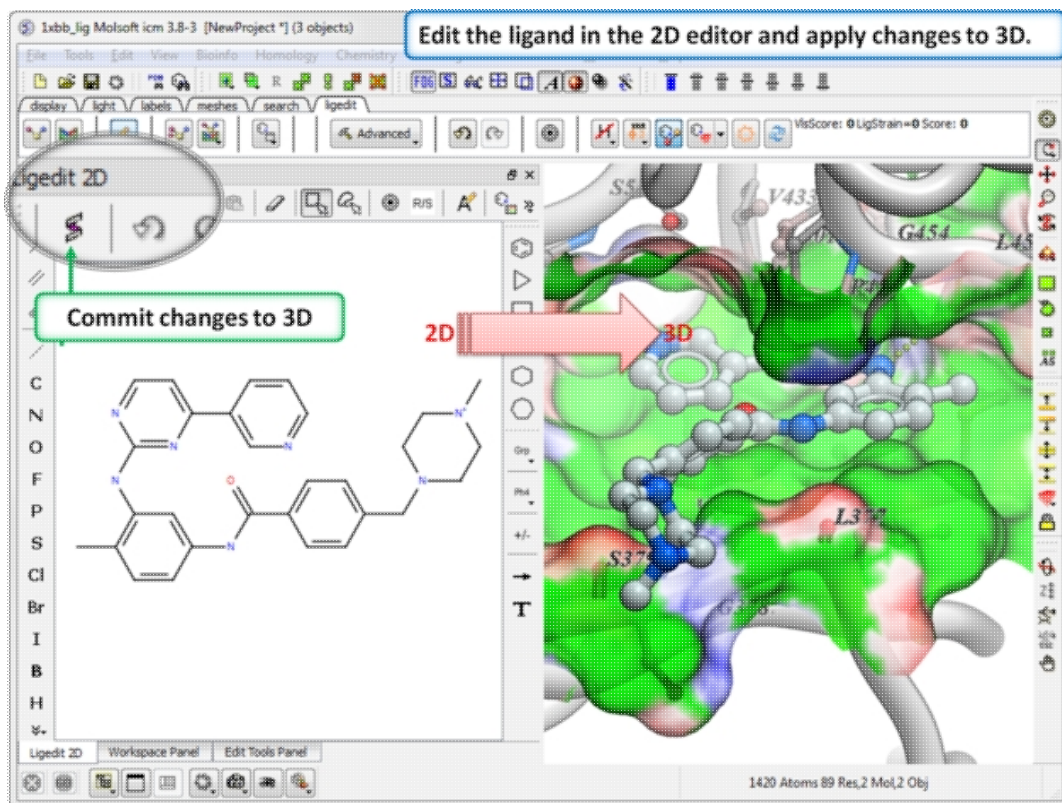
- Setup the ligand and receptor.
- Click on the **Edit Tools Button** and a panel of buttons will be displayed as shown below.
- Click on the 2D editor button and edit using the tools in the molecule editor.



- The 2D ligand editor can be displayed in the panel on the left hand side docked into the GUI or you can change the location by dragging it. The default location can be controlled by clicking on the Advanced button and choosing Preferences/Ligedit2D Position.



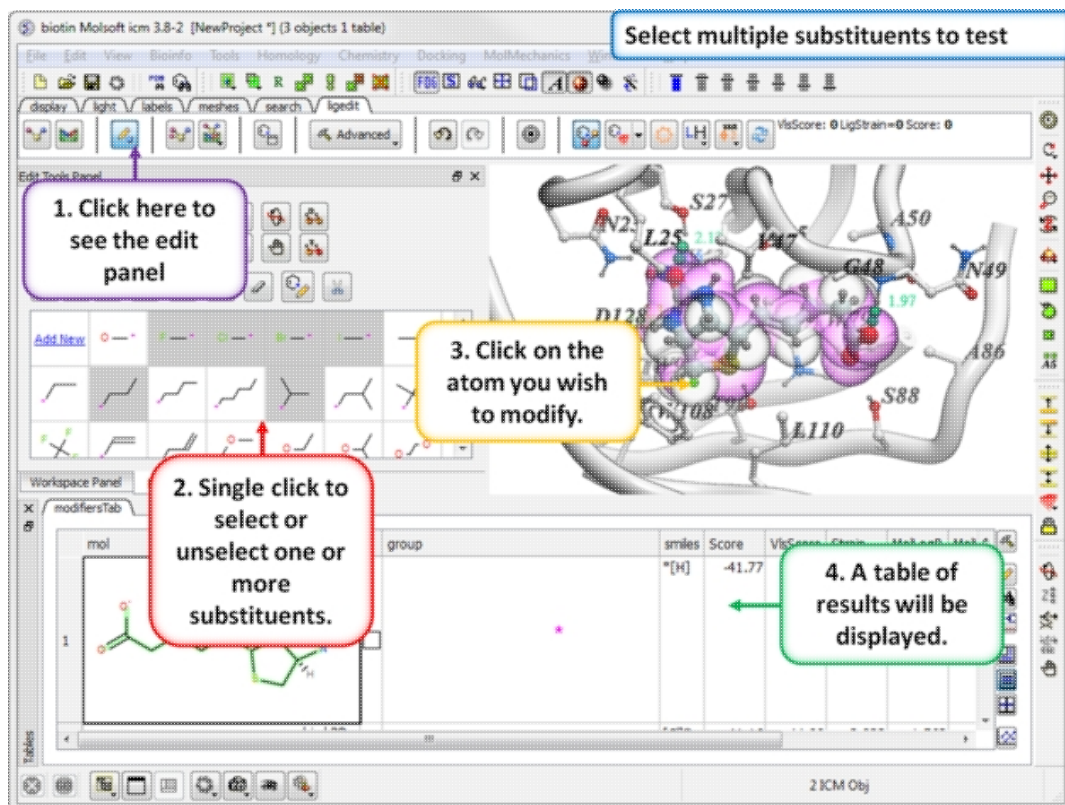
- Make changes to the ligand in the 2D editor and commit them to 3D as shown below.



14.5.2 Multiple Substituents

You can select more than one group. ICM will sample the energy of each group and return a table of the results ranked by binding score. The table columns are described below.

- Click on one or more substituents (shaded grey when selected).
- Click on the atom you wish to modify.
- The results will be reported in the modifier table.



About the modifiers table

mol = 2D sketch of ligand with core substructure highlighted in green. Modifier group is not highlighted in green.

L = Click in box to display ligand with modifier group.

smiles = smiles string of modifier group

SubstScore = Score for modifier group only

Score = Score of whole ligand including modifier group

MolLogP = Predicted LogP

MolLogS = Predicted LogPredicted LogSS

MoldHf = Prediction model build for 'delta Hf in gas' property. using public NIST database. Description can be found: <http://webbook.nist.gov> A low dHf value means that the compound is more 'stable'.

MolPSA = Polar Surface Area

Volume = Volume of ligand.

14.5.3 User Defined Modifiers

The default set of modifiers LIGAND.modifiers is just an array of smiles and it is hard-coded into ICM. You can add your own groups by clicking on the "Add New" button in the editor

<http://www.molsoft.com/gui/edit-ligand.html> When sketching you need to add an attachment point - right click on the attachment point atom in the 2D editor and select attachment point.

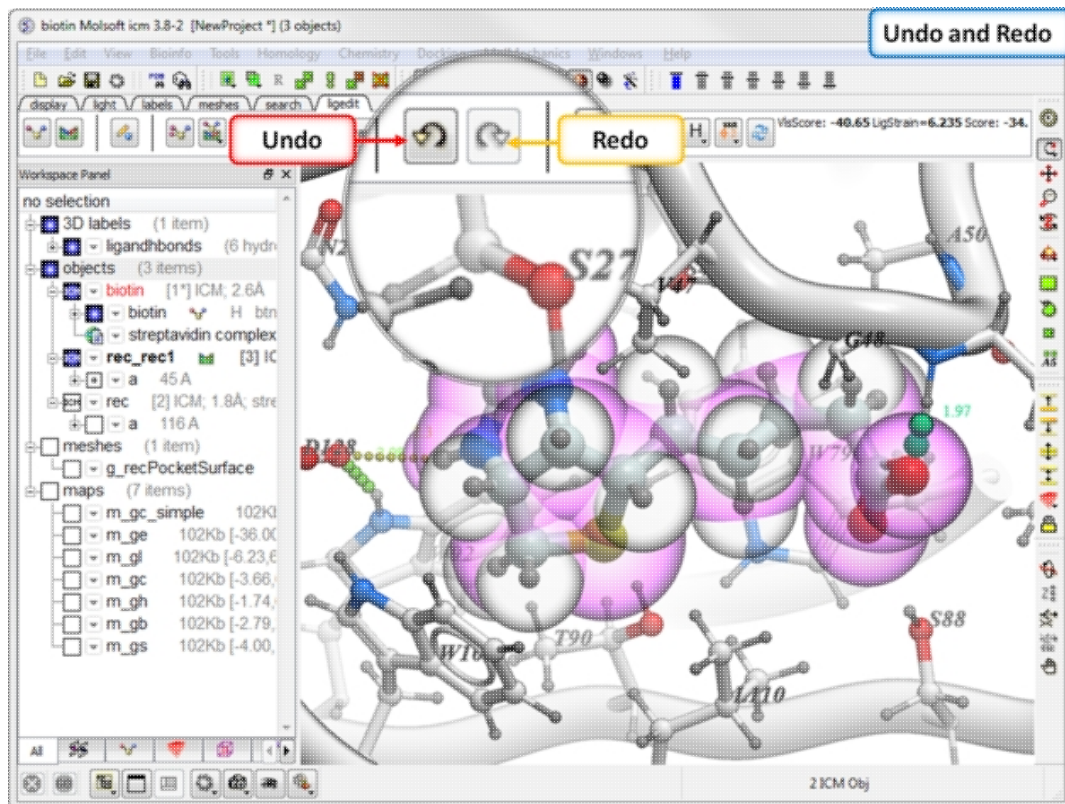
You can also add manually through the command line:

```
LIGAND.userModifiers = LIGAND.userModifiers // "c1[c*]nc(N)s1"
write system preference
```


14.5.4 Undo and Redo

Every edit (even after docking) can be undone or redone.

To undo or redo:



14.5.5 Modification History

You can view a table with a log of all modifications by clicking and holding on the "undo" button and choose "Modification history". This will produce a chemical table with a column (operation) describing the modification. Double click to load the modification into the receptor or use the "L" column to toggle the display on and off.

Save 3D Ligand to Chemical Spreadsheet.

1. Click here to save the modified ligand to a chemical spreadsheet.

2. A new table will be generated.

3. Toggle the 3D display on and off.

N	mol	rec	RecConf	Molecule Name	L	Score	VisScore	Strain	Steric	Torsion	Elect.
1		rec_rec1	rec_rec1:1	biotin		-3.195	-44.09	40.89	-21.25	4	3.

You can also tag a chemical with a unique name which is stored in the spreadsheet. To do this click and hold on the "Add Current Ligand to Table" button and enter a tag name. The ligand table can be grouped by this tag column (Right Click on the column and choose "Group View By").

Modifications can be saved directly to a chemical spreadsheet and tagged if needed.

Single click to add to spreadsheet. Click and Hold to tag.

Enter tag name.

N	mol	tag	RecConf	rec	Molecule Name	L	Score	VisScore	Strain	Steric
1		3D Series 1	1xbb_rec1:1	1xbb_rec1	1xbb_lig		-27.47	-30.31	2.841	-32.22

14.5.7 After Editing the Ligand

After editing the ligand you can:

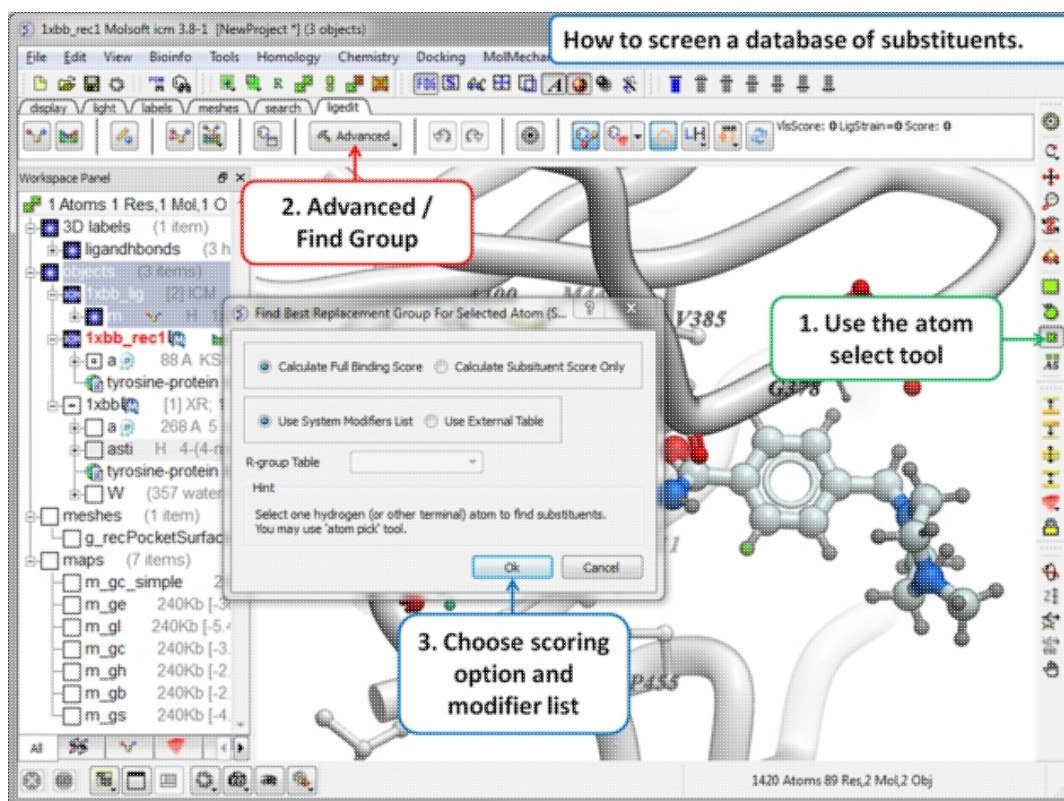
- Re-dock or minimize the ligand and evaluate the energy and score or view interactions such as hydrogen bonds and energy circles.
- Before docking if certain groups or distances need to be retained you can set distance restraints or tether atoms.
- You can try additional changes by screening a database of substituents.
- Add the ligand to a chemical spreadsheet.
- Save the docked ligand and receptor complex as a PDB file .

14.6 Virtual Screen: Find Best Replacement Group

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

This option allows you to screen a database of substituents. You can use the ICM "in built" substituents or a database of your own.

To find best replacement group for selected atom:

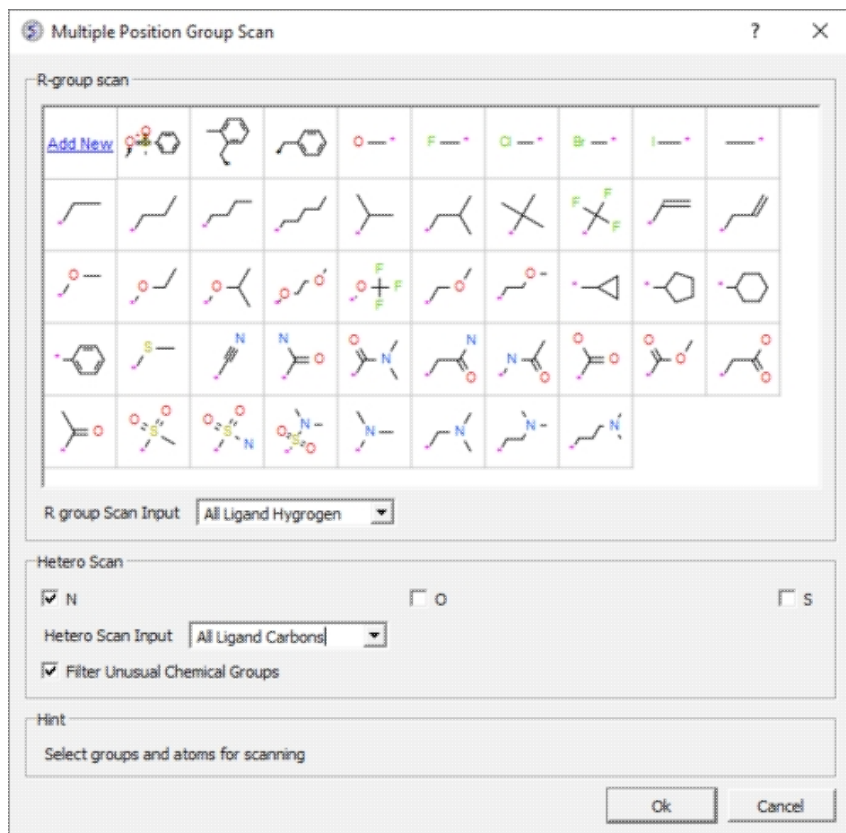


- Setup the ligand and receptor.
- Select the atom you want to add a new replacement group to.
- Click on the **Advanced/Find Group** button.
- A dialog box as shown below will be displayed. Select whether you want the **substituent score only** to be evaluated (quick) or the **full binding** score.
- Select whether you want to screen the modifying groups built into ICM (see sarray of smiles called LIGAND.modifiers) or a table of your own modifier groups. If you choose your own table you will need to load the table (sdf file) into ICM and enter the name of the table into this dialog box or you can add modifiers to the sarray of smiles called LIGAND.userModifiers.
- ICM will add each fragment to the target atom and sample the energy and return a table ranked by score.

14.7 Multiple Position Group Scan

This option allows you to scan groups and hetero atoms at multiple locations on the ligand. To do this:

- Select the atoms on the ligand you wish to scan against.
- Select the groups from the R-group scan table (single click selects) and/or select the hetero atom scan.

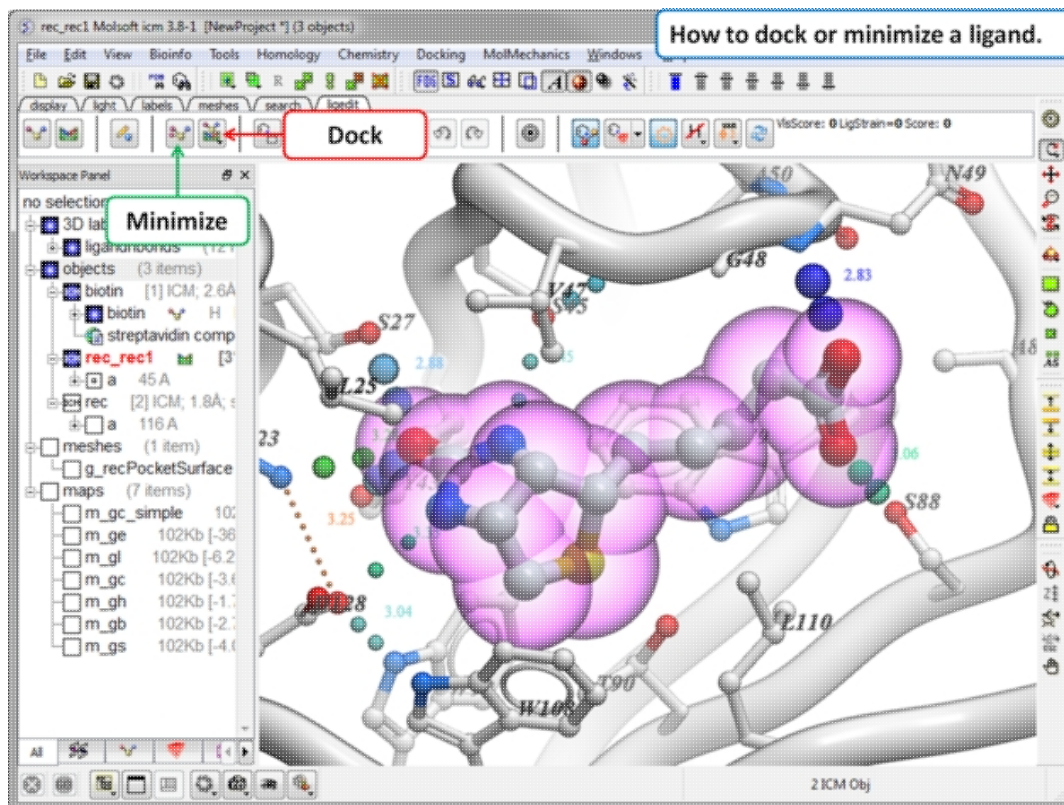


14.8 Dock or Minimize Ligand

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

Video |

How to Dock or Minimize a Ligand in the Grid maps:



- Setup the ligand and receptor.
- Click on the dock or minimize button (see image above).
- Check the Score and Strain in the top right panel.

How to undertake an all-atom minimization in cartesian coordinates:

- Note this option is available in ICM 3.8-4 and higher.
- Click and hold on the **Minimize** button.
- Choose **Minimize in Receptor**.

14.8.1 Change the Size of the Docking Region

How to change the size of the docking region:

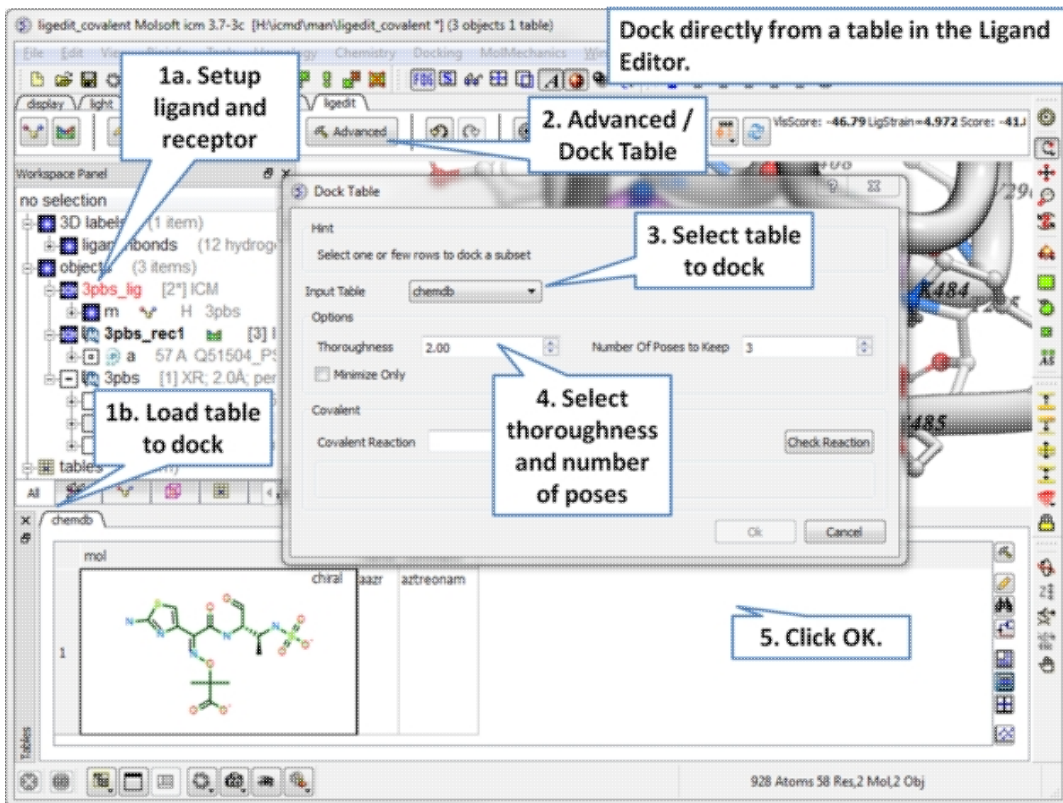
The purple box represents the region in which the energy maps are generated. If you want to change the size of this region you can do so by displaying the box (go to Advanced menu) and clicking and dragging on the corners of the purple box. The maps will then be remade on the fly.

14.9 Dock from Table

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

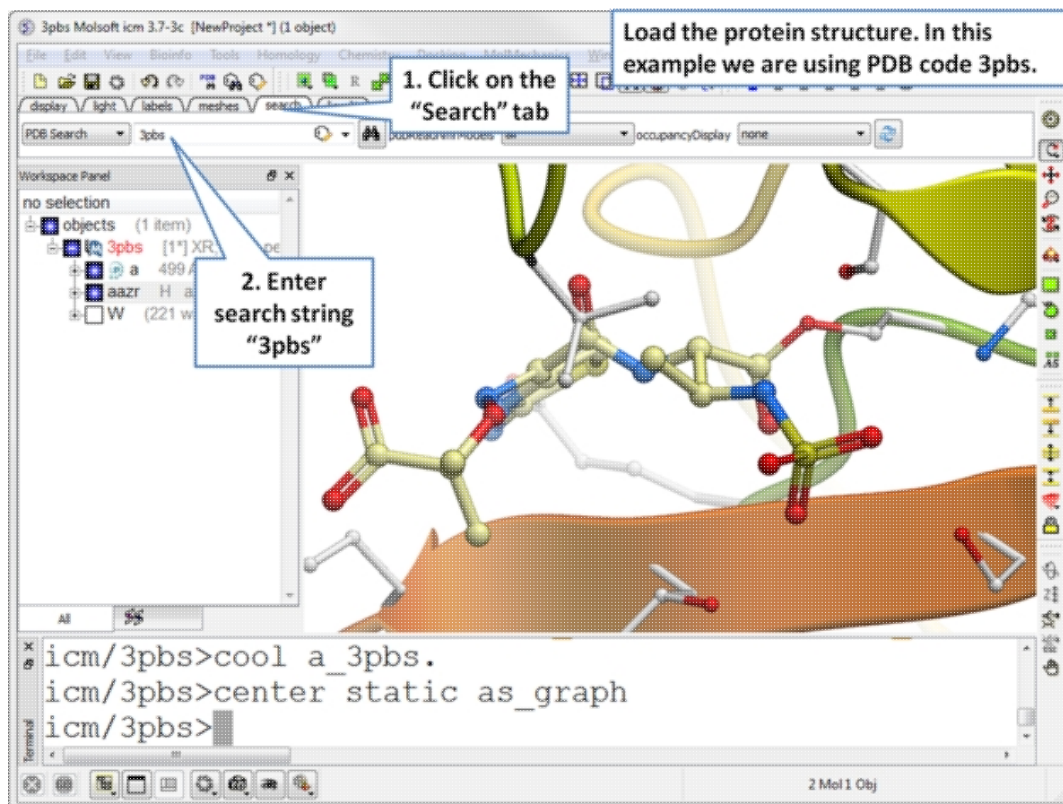
To dock from a table in the ligand editor.

- Setup the ligand and receptor in the ligand editor.
- Read in the sdf file or chemical spreadsheet which contains the chemicals you wish to dock.
- Select the **Advanced** button in the ligand editor and select Dock Table.
- In the dialog box select the thoroughness (length of simulation) and the number of conformations.

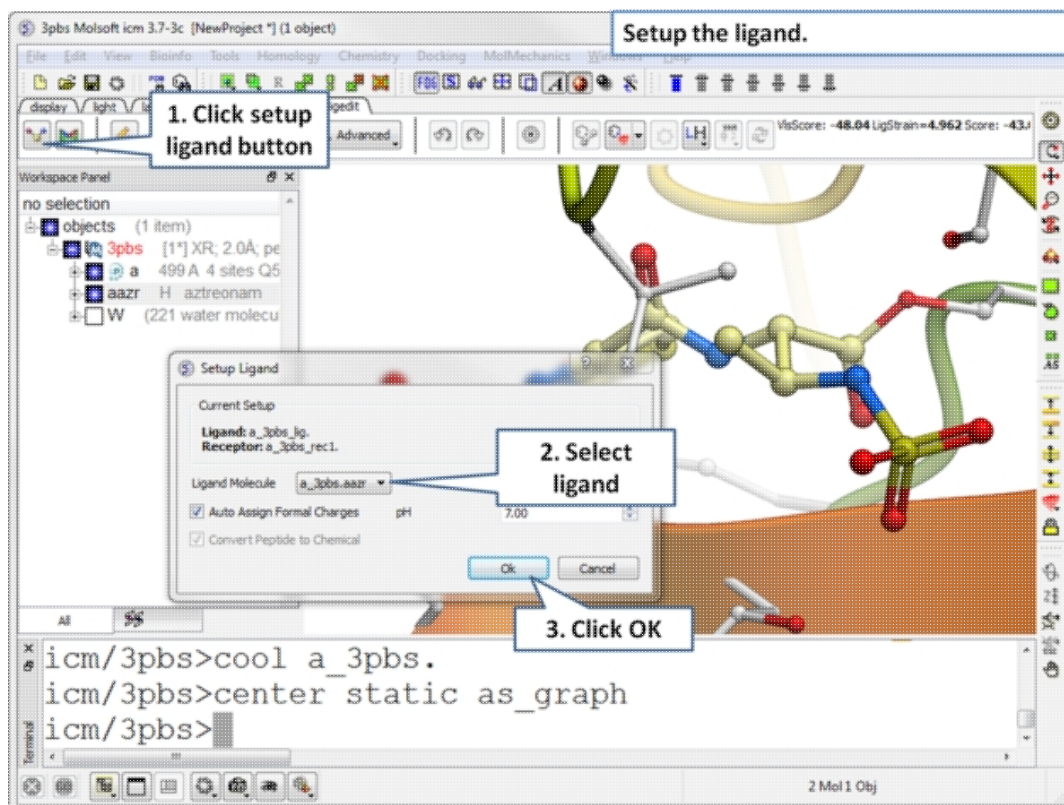


14.9.1 Covalent Docking from Table

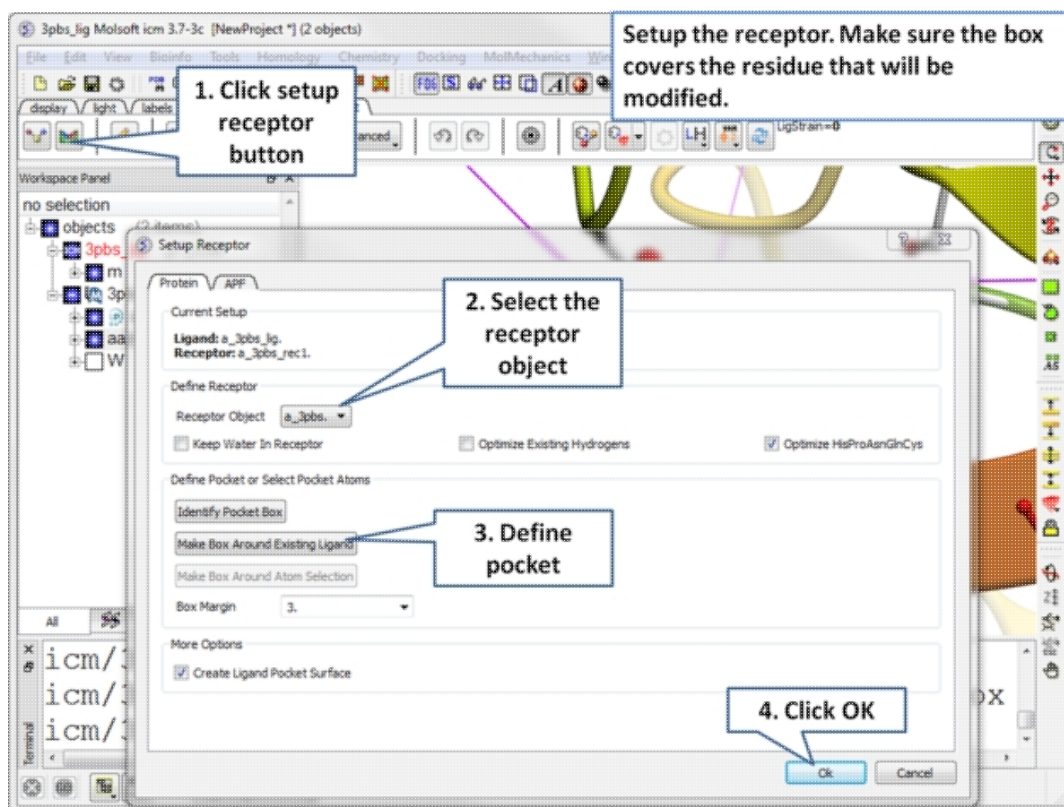
To undertake covalent docking from a table:



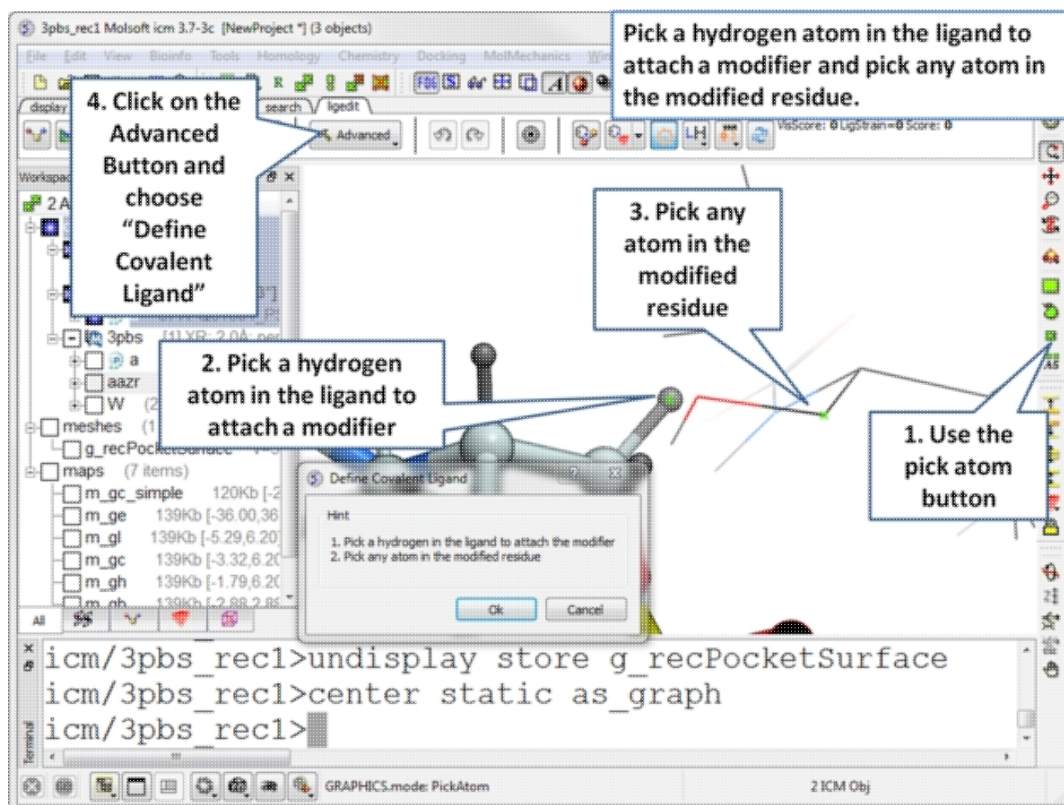
Read in the PDB file you wish to dock to containing the covalently linked ligand.



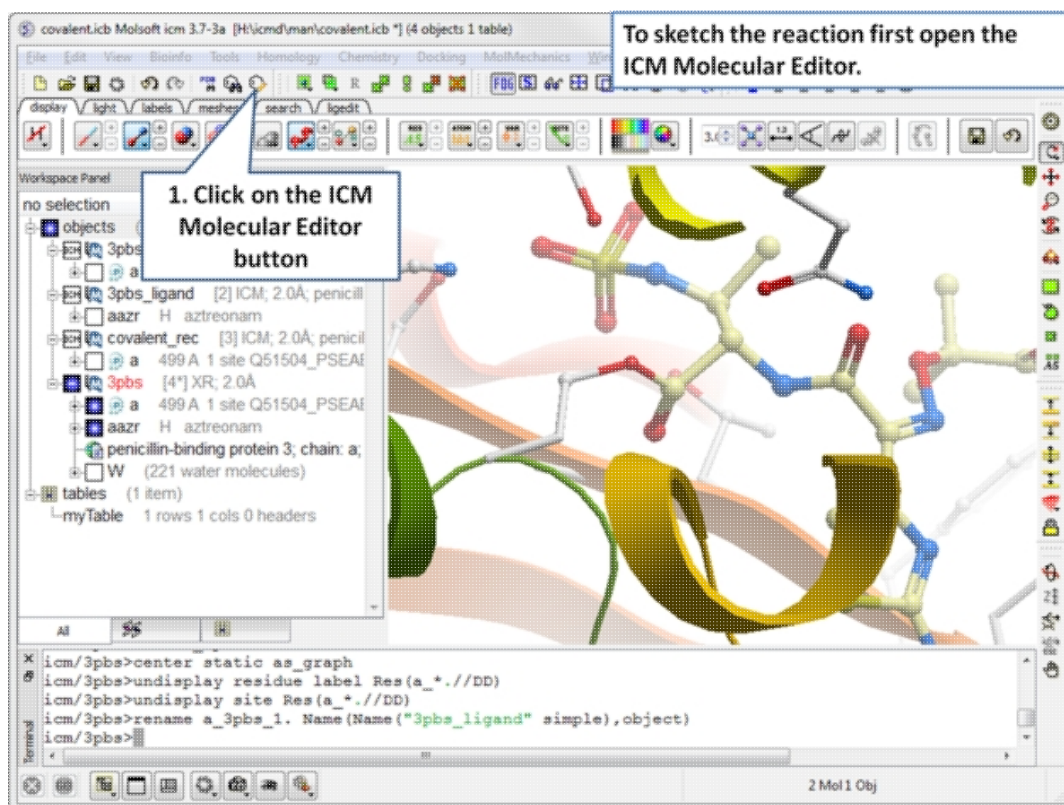
Setup the ligand in the ligand editor.



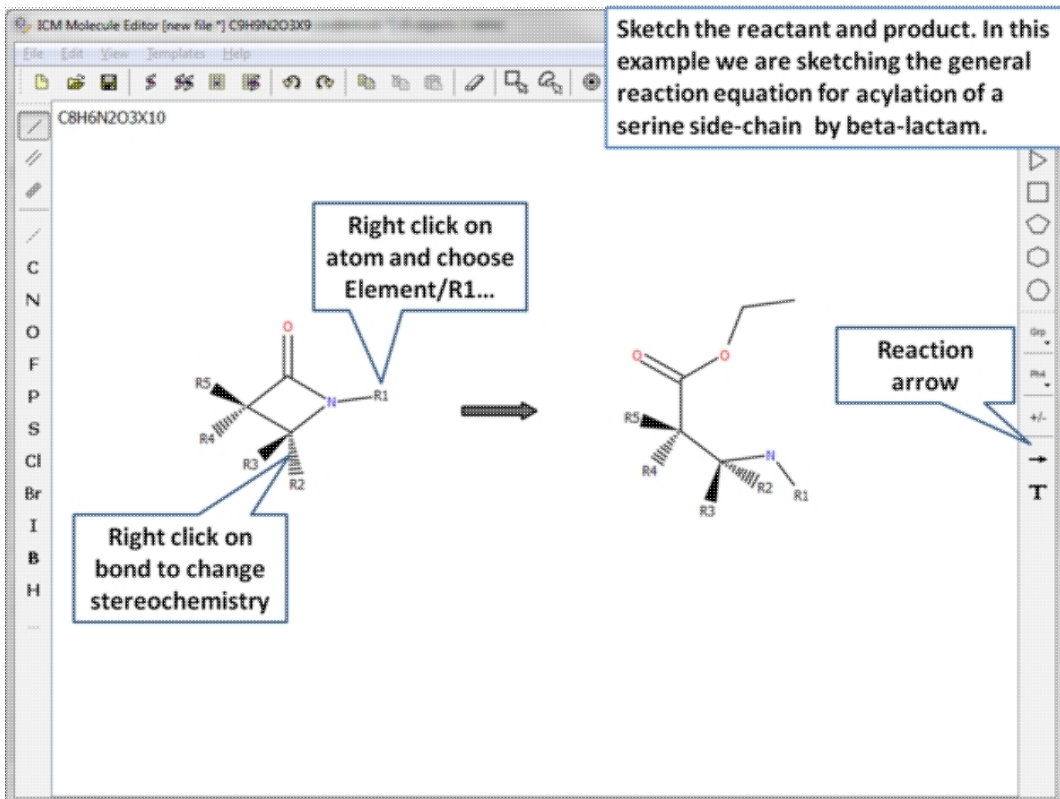
Setup the receptor in the ligand editor.



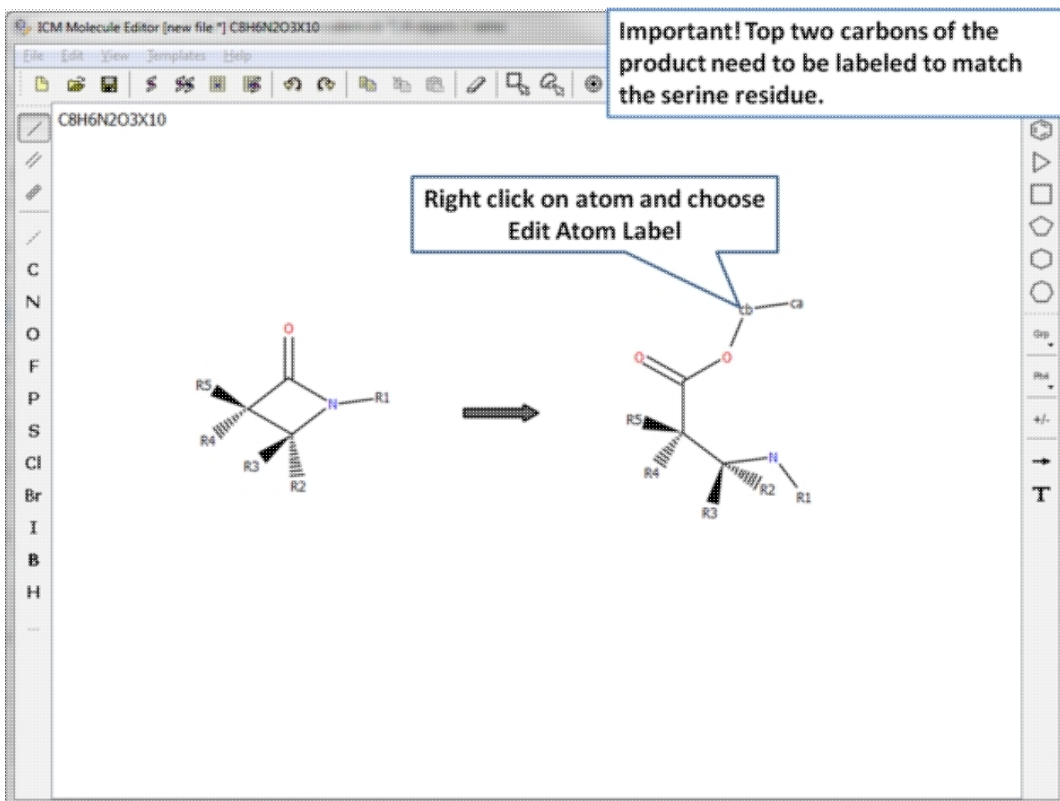
Pick a hydrogen atom in the ligand to attach a modifier and pick any atom in the modified residue. Next, click on the Advanced button and choose "Define Covalent Ligand".



Open the molecular editor. Note you may first need to click on the display tab to see the Molecular Editor button.



Sketch the reactant and product.



Two of the atoms need to be labeled to match the residue in the receptor.

1. Click here to append the reaction to a chemical table

2. File/Quit

Append molecule to a chemical table. Note: more than one reaction can be appended to the table to match the chemistry of the compounds you wish to dock.

Append the reaction to a chemical table. Note more than one reaction can be added to the table to match the ligands in your chemical table that you wish to dock.

Load the chemical table containing the chemicals for docking which match the reaction.

Reaction(s)

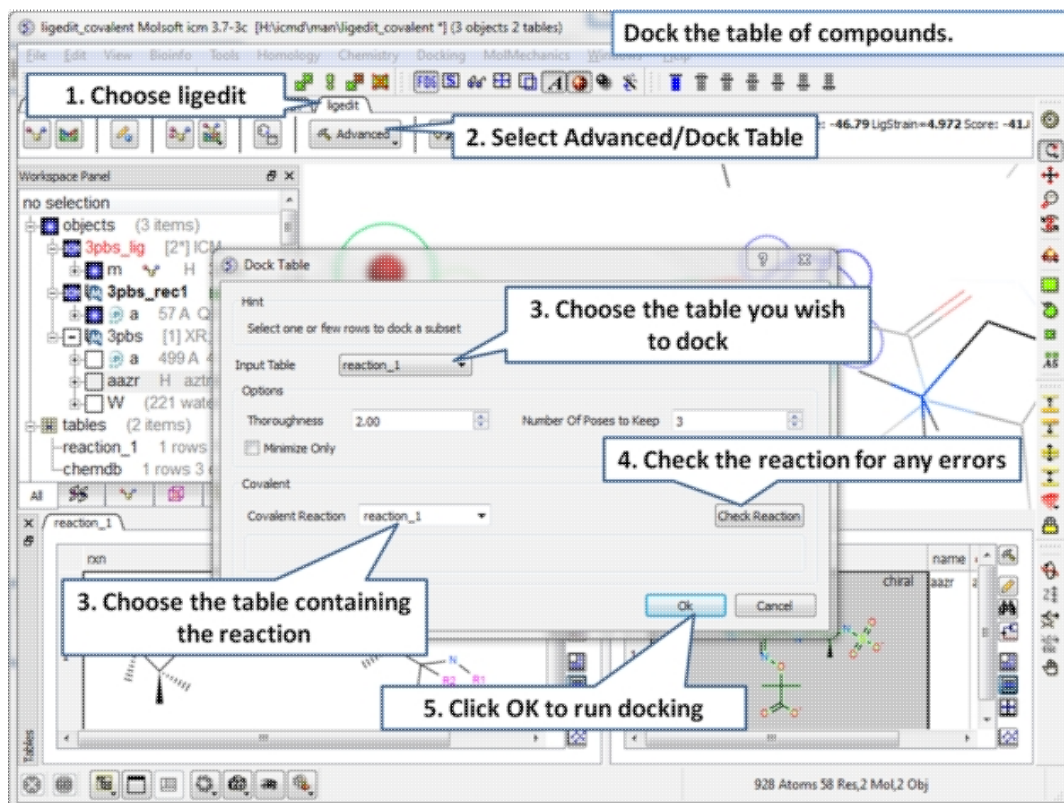
Compounds to dock

Matching Chemistry!

mol	name
chiral	aazr

928 Atoms 58 Res, 2 Mol, 2 Obj

Load the chemical table containing the ligands you wish to dock.

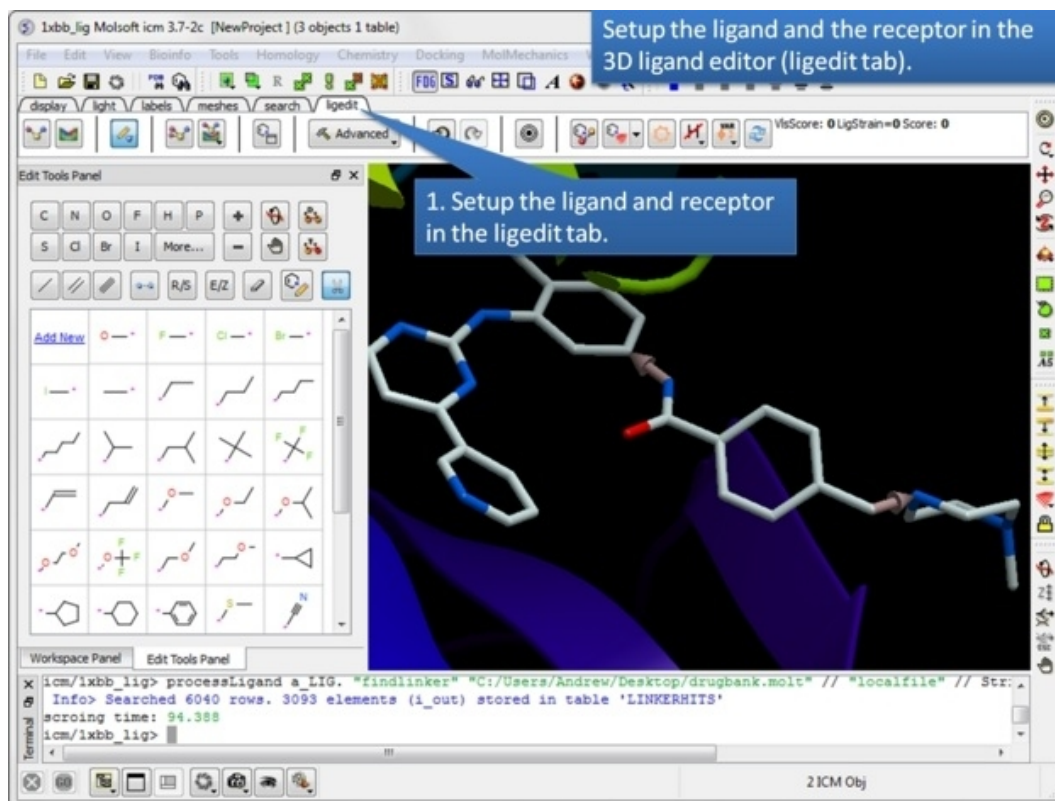


Choose Advance/Dock Table and select the covalent options shown.

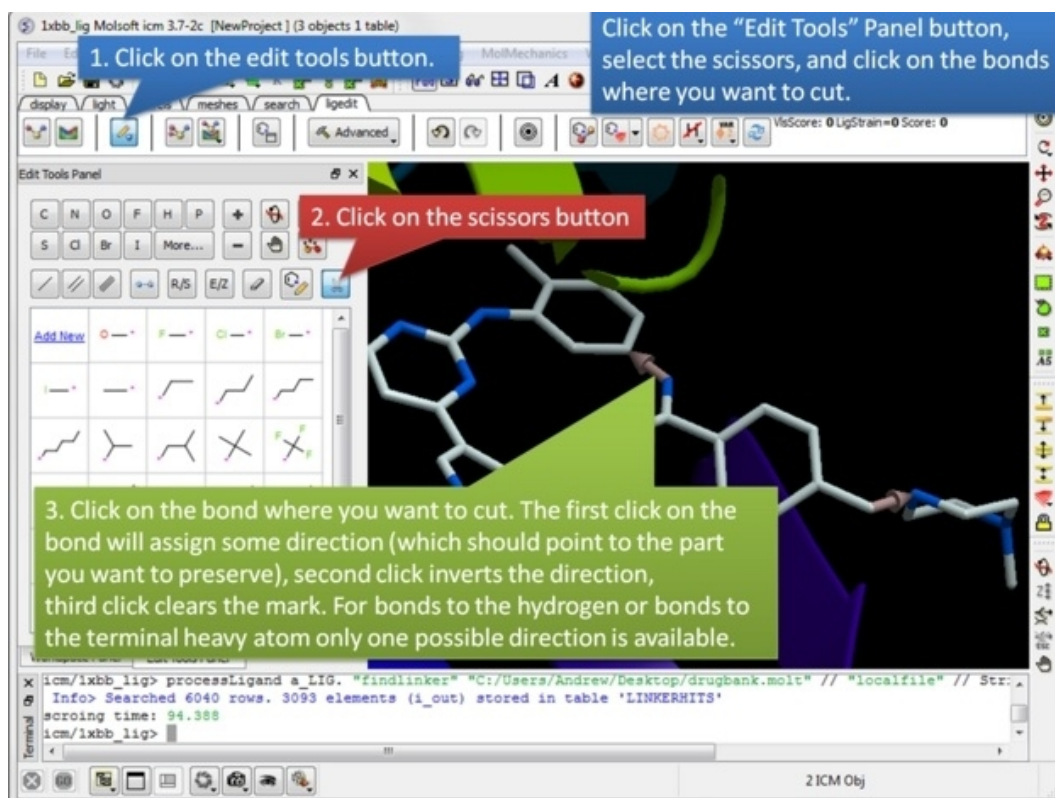
14.10 Fragment Linking

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

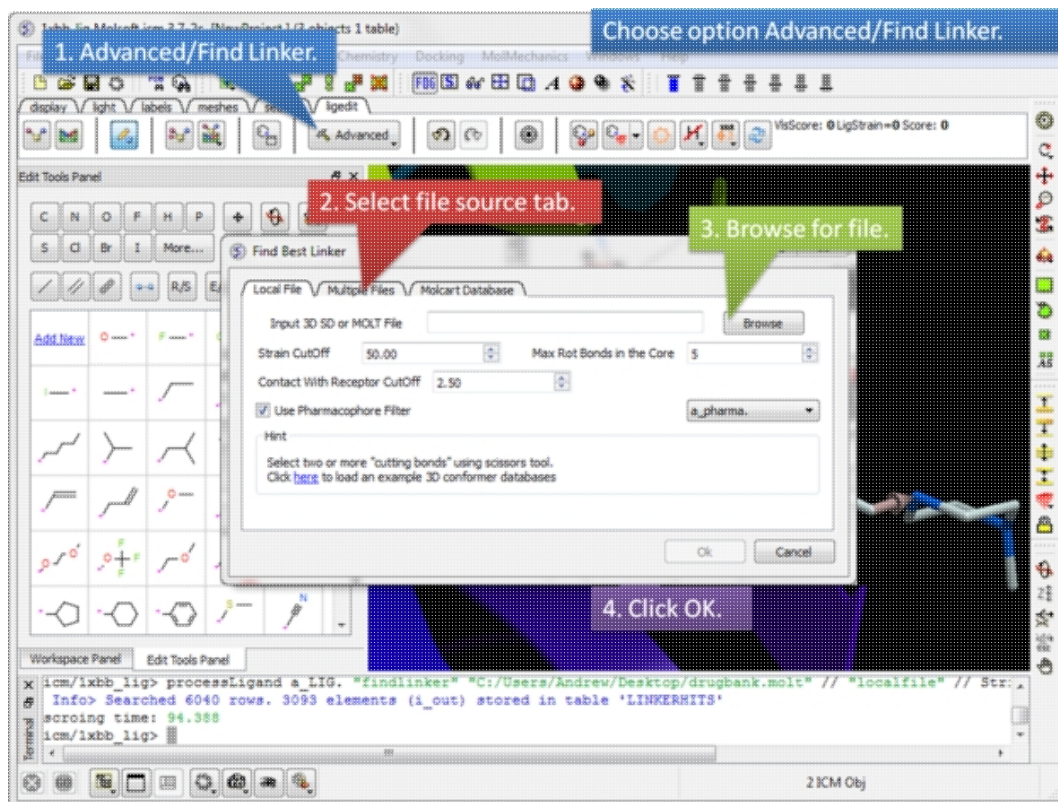
This feature finds the best scaffold replacement or linker fragment from a database. It lets you interactively select two or more points of connection or replacement and search for possible fragments in the 3D database (SDF or Molsoft local database MOLT). The cutting bonds used to define scaffold are matched against a 3D database with a certain RMSD cutoff (usually very small). The best superposition is taken for the original compound and the chemical is minimized. The result is returned as a standard Ligand Editor hit list where results can be easily browsed, sorted and compared. Linker prediction can be made in the presence or absence of the receptor. Filters such as energy strain, maximum number of rotatable bonds and contact distance to the receptor can be used.



Step 1: Setup the ligand and/or the receptor in the 3D ligand editor. If you have fragments already made you should select the fragments and then choose "Selection" in the drop down menu in the ligand setup step. If you do not wish for the receptor to be considered when sampling linkers then just setup the ligand only. Once the ligand and/or receptor are setup it is generally a good idea to minimize the ligand to remove any clashes and optimize the geometry.



Step 2: Once the ligand and/or the receptor has been setup click on the edit tools button. A panel of editing tools will be displayed, click on the scissors button. Mark two or more bonds using the scissors tool. The first click on the bond will assign some direction (which should point the part you want to preserve), second click inverts the direction, third click clears the mark. For bonds to the hydrogen or bonds to the terminal heavy atom only one possible direction is available.



Step 3: Select the **Advanced** button and choose the option **Find Linker**. A dialog window will be displayed as shown above. You can choose to screen one or more SDF files containing 3D coordinates or a `molt` file that is stored locally on your machine or on MolCart.

- Provide a path to the 3D SDF or Molt file. We have provided some examples of 3D conformer databases here <ftp://ftp.molsoft.com/molcartdb/3D>
- **Strain Cutoff** - any ligand exceeding this Strain value will not be placed in the results table. If you are using a ligand directly from the PDB it is a good idea to minimize the ligand and check the initial strain or increase this cutoff.
- **Max Rot Bonds in the Core** - any ligand exceeding this number of rotatable bond value will not be placed in the results table.
- **Contact with Receptor Cutoff** - any ligand exceeding this distance will not be placed in the results table. This value is significant if the receptor is included in setup.
- **The use pharmacophore filter** allows you to match linkers with a pharmacophore object. Learn how to make a 3D pharmacophore object [here](#).

Results table will be displayed.

1. Table ranked by strain

2. Display and undisplay

3. Hit is displayed

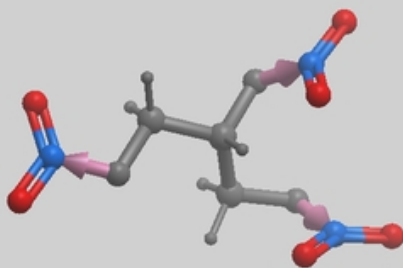
mold	mol	Score2	Strain	L	rmsd	3D	name
1	3193	10	10	<input checked="" type="checkbox"/>	3.003275	3D	3-Carboxy-N,N,N-Trimethyl-2-(Octanoyloxy)Propan-Aminium

3 Obj [1 non-ICM]

Step 4: Once the fragment fitting has completed a table of results will be displayed. The table is ranked by strain with the top hit in row 1. You can browse the hits using the toggle buttons in the "L" column.

- **Strain** is the energy strain of the ligand.
- **Score2** is the same value as strain (possible place holder for future release).
- **RMSD** - is the RMSD to the original fragment before linker connection.

Three Point Core Replacement Three cores can be linked by selecting them with the scissors button as shown in Step 1 above. Number of hits will be significantly lower than in 2 point replacement because the extra anchor adds some restraint. We recommend to use larger database for the search (e.g. chembridge100K.molt) and make sure that ligand is minimized before searching.



14.11 Impose Tethers or Distance Restraints

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

14.11.1 Impose Ligand Atom Tether

To impose tethers to selected atoms of the ligand before redocking.

- Setup the ligand and receptor.
- Click on the edit tools panel button.
- Click on the impose positional restraints button.
- Select the atoms you wish to tether in the ligand.
- A table listing the tethers will be displayed and red crosses through the atom indicates a tether.
- Click on the re-dock button and the atoms selected will remain tethered in place.

1. Select edit tools panel

2. Select impose positional restraints button

3. Select one or more atoms

4. Table values can be edited

	use	L	type	atom1	atom2	max dist	min dist	weight	rm
1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	position	a_biotin.biotin/300/n1	35.31,10.86,-4.22	0	0	10	delete
2	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	position	a_biotin.biotin/300/n2	34.88,12.57,-5.55	0	0	10	delete
3	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	position	a_biotin.biotin/300/c5	34.28,10.36,-5.09	0	0	10	delete
4	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	position	a_biotin.biotin/300/c4	34.09,11.48,-6.11	0	0	10	delete

About the table: The tethers can be edited and deleted using the table. Double click on a value in the table to edit it.

- Use Check "use" to use the tether in docking.
- L Toggle the tether display on or off.
- **Type** Type of tether "position" indicates a positional restraint (tether).
- **Atom1** ICM selection language for the tethered atom. Atom 1 can also be SMARTS expression which define the atom (for both distance and position) The matched atom should be marked with '*'. For example oxygen in the double bonded to the 5-member ring in biotin can be defined as: NC(=[O*])N
- **Atom2** xyz position (3 comma separated numbers)
- **max dist** Sets the maximum distance for the tether.
- **min dist** Sets the minimum distance for the tether.
- **weight** the overall weighting factor for the tether penalty term.
- **rm** Delete the tether.
- **err** Deviation from the position after docking.

14.11.2 Impose Distance Restraint

There are three types of distance restraint you can make:

1. lig - static rec : fixed side-chain (weighted tether)
2. lig - flex rec : flexible side-chain (drestraint)

3. lig - lig : drestraints

To impose distance restraints before redocking.

- Setup the ligand and receptor.
- Click on the edit tools panel button.
- Click on the distance restraint button.
- Select two atoms to form the distance restraint. The distance restraint can be between two atoms in a ligand or between the ligand and receptor
- A table detailing the distance restrints will be displayed and red crosses through the atom indicates a tether.
- Click on the re-dock button and the atoms selected will remain tethered in place.

The screenshot shows the Docking Tools software interface. The main window displays a 3D molecular model of a ligand (red and white) docked into a receptor (grey). A distance restraint is shown as a blue dashed line between two atoms, labeled '3.28'. The 'Edit Tools Panel' is open, showing various restraint options. A table at the bottom, titled 'DOCK_RESTRAINTS', lists the current restraint. The table has columns for 'use', 'L', 'type', 'atom1', 'atom2', 'max dist', 'min dist', 'weight', and 'rm'. The first row shows a distance restraint between 'a_biotn.biotin/300/o3' and 'a_rec_rec1.a/N23/nd2' with a maximum distance of 3.277 and a minimum distance of 3.277. A red 'X' is visible over the 'atom2' cell, indicating a tether. A callout points to the 'rm' column, stating '4. Table values can be edited'.

	use	L	type	atom1	atom2	max dist	min dist	weight	rm
1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	distance (lig-rec)	a_biotn.biotin/300/o3	a_rec_rec1.a/N23/nd2	3.277	3.277		3 delete

About the table: The distance restraints can be edited and deleted using the table. Double click on a value in the table to edit it.

- **Use** Check "use" to use the distance restraint in docking.
- **L** Toggle the distance restraint display on or off.
- **Type** Type of restraint. There are three types of distance restraint
 1. lig - static rec : fixed side-chain (weighted tether)
 2. lig - flex rec : flexible side-chain (drestraint)
 3. lig - lig : (drestraint)
- **Atom1** ICM selection language for the first atom in the distance restraint.
- **Atom2** ICM selection language for the second atom in the distance restraint.
- **max dist** Sets the maximum distance for the restraint.
- **min dist** Sets the minimum distance for the restraint.
- **weight** the overall weighting factor for the distance restraint penalty term.
- **rm** Delete the distance restraint.
- **err** Deviation from the position after docking.

Note: The area within min-max distance is zero (flat-bottomed). With a weight of 1. you will have fair coverage of conformations with minor violations, going up to 5 or 10 will pretty much force the system to always stay there, below 1. you will only have a sort of weak steering force.

14.12 Defined Flexible Receptor Groups

You can allow certain defined residues in the receptor to be flexible during docking. To do this:

- Setup the ligand and receptor.
- Select the residues you wish to be flexible using the `residue selection` button.
- Select `Advanced/Defined Flexible Receptor Groups`
- Choose full side-chains or hydroxy and water.
- The side chains you have selected will then be displayed in thicker stick and will be flexible when you dock or edit a ligand.

NOTE: It is recommended to not select more than 3 residues to be flexible.

14.13 Side Chain Refinement

You can refine the side-chains around a docked ligand by:

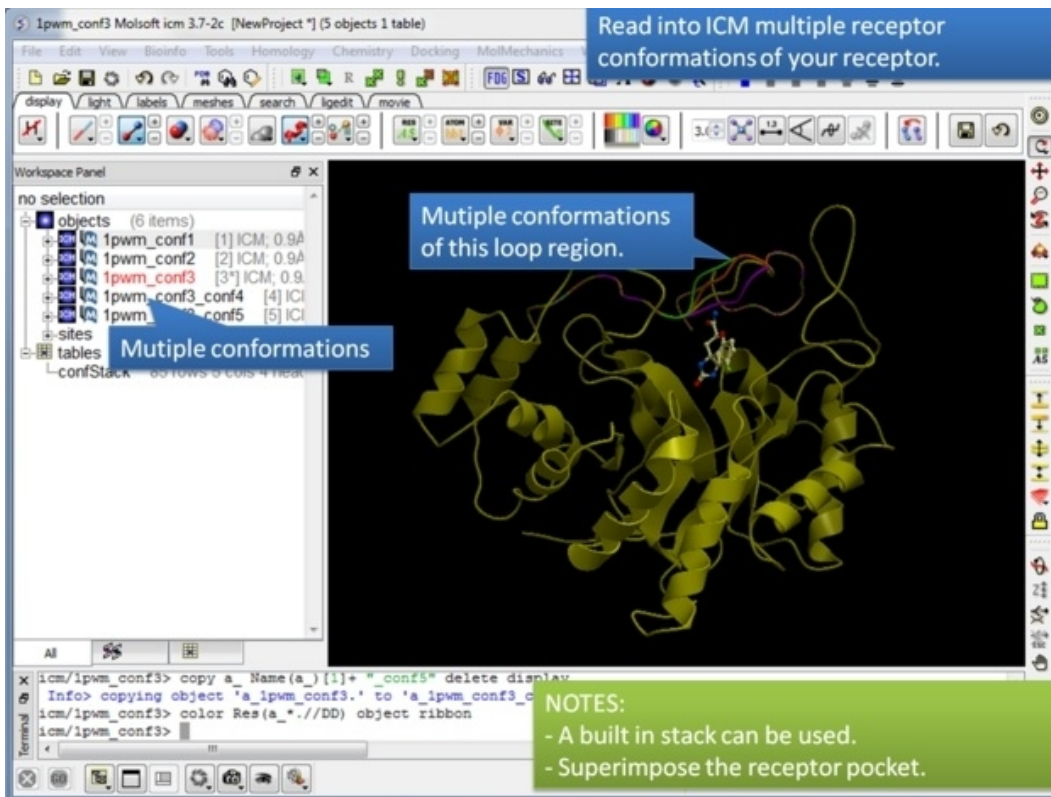
- Setup the ligand and receptor.
- Select the residues you wish to be flexible using the `residue selection` button.
- `Advanced/Receptor Side Chain Refinement`.

14.14 Multiple Receptor Docking

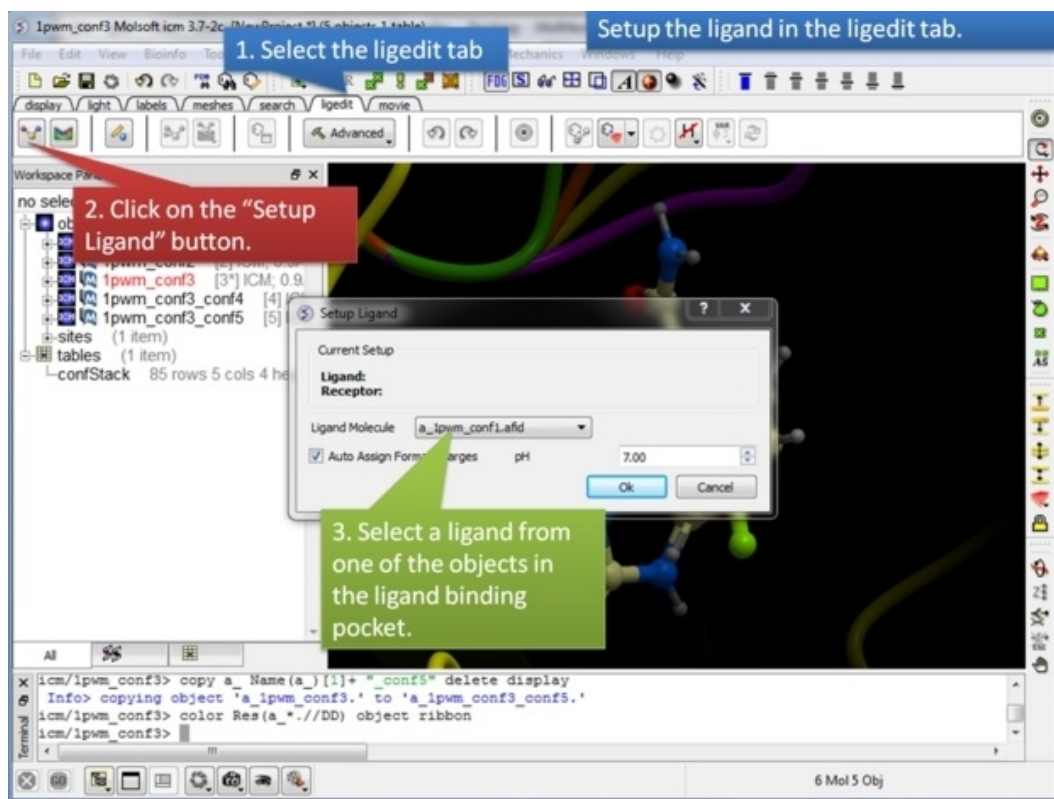
Available in the following product(s): `ICM-Pro` | `ICM-VLS` | `ICM-Chemist-Pro`

| Video |

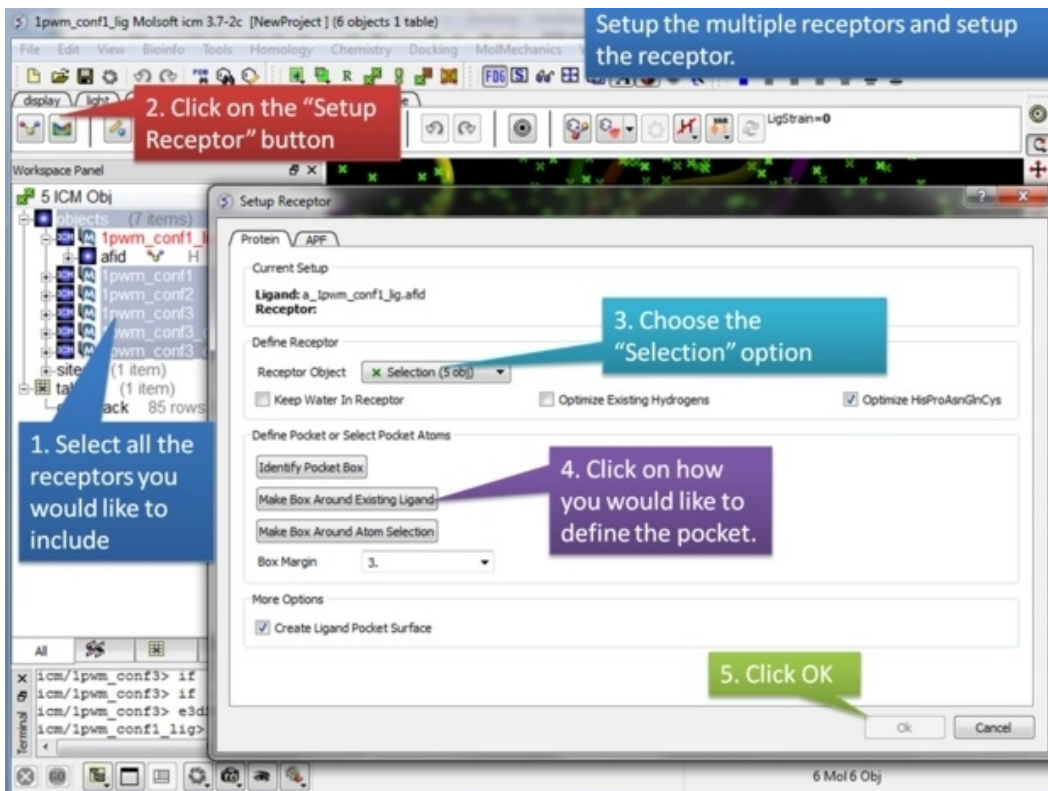
Multiple receptor conformation docking has been shown to be an efficient way of incorporating receptor flexibility into ligand docking (see Bottegoni et al *J.Med.Chem* 2010). You can dock to multiple receptor conformations in the 3D ligand editor by selecting the superimposed structures receptor setup stage. 4D maps will be generated encompassing all the receptors selected.



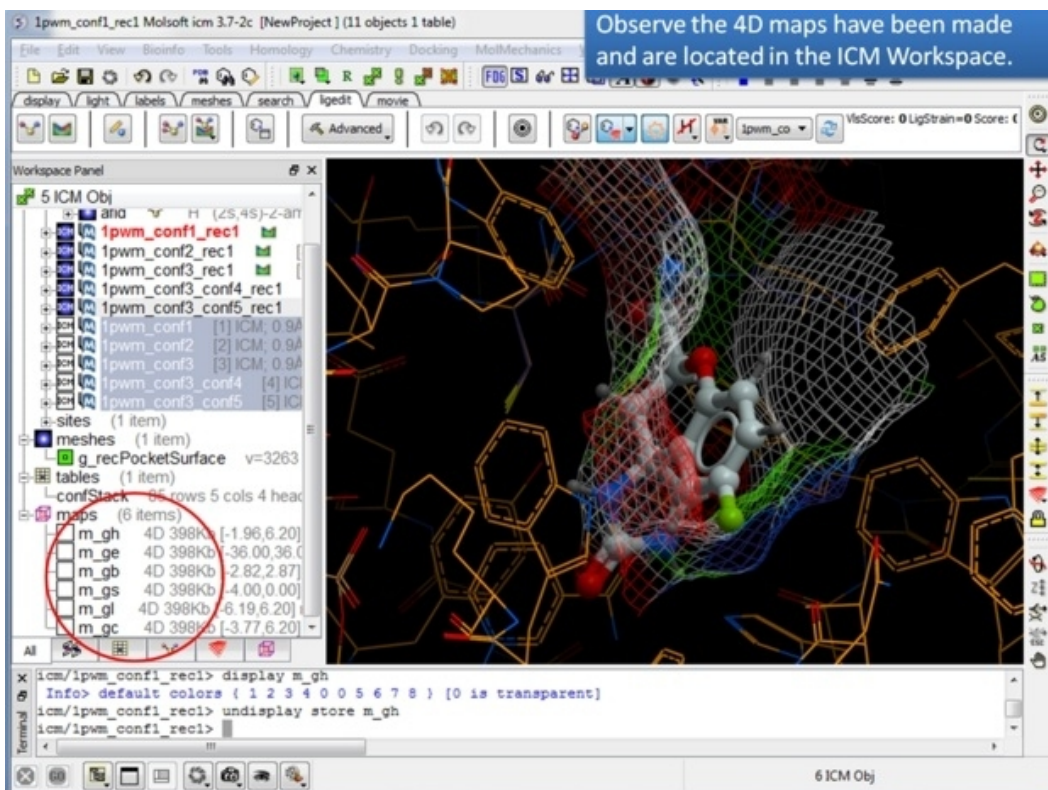
Step 1: Read into ICM the receptor conformations you would like to dock to. If you have a conformation stack already loaded into ICM (e.g. after modeling a loop region) ICM can use that.



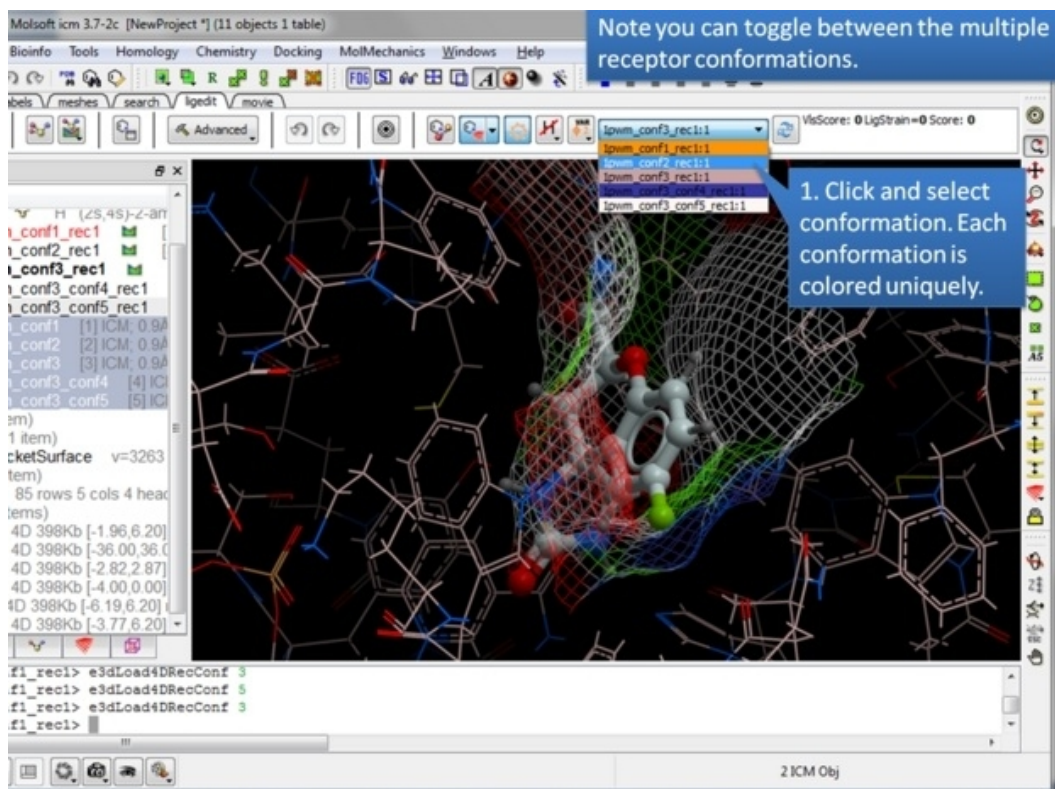
Step 2: Setup the initial ligand you would like to dock.



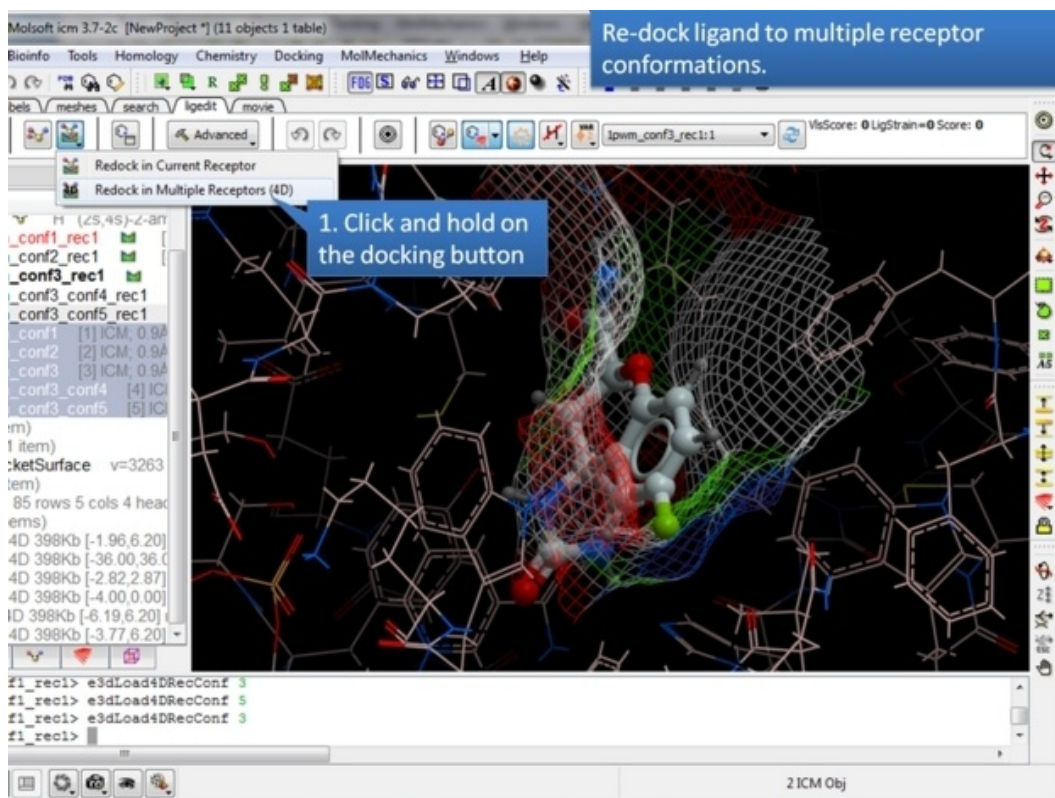
Step 3: Select the receptors you would like to dock to. Make sure they are correctly superimposed.



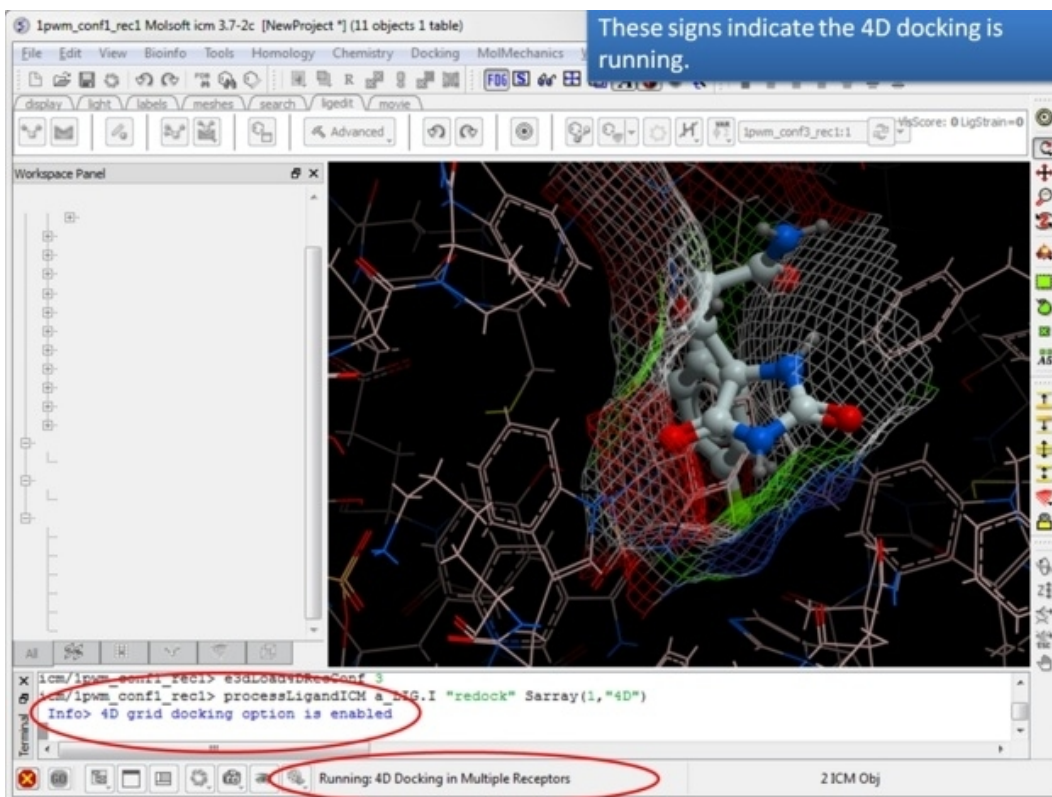
Step 4: You will see that 4D maps have been generated and stored in the ICM Workspace.



Step 5: You can easily toggle between each receptor to display them.



Step 6: Click and hold on the docking button to select "Re-dock to multiple Receptors (4D)".



Step 7: The docking will start running.

Best docked ligand conformation is listed in a table ranked by Score.

1. Click here to display the ligand.

ConfNum	L	Score	VisScore	Strain	RecConf	Steric	Torsion	Electro	Hbond	Hvdroop	Surface
1	14	-4.268	-15.61	11.34	1pwm_conf1_rec1:1	-14.4	1	3.216	-3.671	-3.729	10.47
2	2	-4.008	-22.89	18.88	1pwm_conf1_rec1:1	-25.52	1	1.871	-2.887	-3.333	13.78
3	12	-1.764	-15.32	13.55	1pwm_conf1_rec1:1	-18.18	1	4.072	-3.654	-3.406	14.8
4	13	1.824	-11.06	12.88	1pwm_conf1_rec1:1	-20.53	1	4.881	-0.4725	-3.77	12.5

```

icm/1pwm_conf1_lig> updateLigandScore a_LIG.I 0. 0. yes
icm/1pwm_conf1_lig> e3dLoad4DRecConf 1
icm/1pwm_conf1_lig> updateLigandScore a_LIG.I 0. 0. yes
icm/1pwm_conf1_lig>
  
```

Step 8: Use the interactive table to view the results.

The screenshot shows the ICM software interface. At the top, a blue banner reads "Toggle through each conformation and calculate the score." Below this, a 3D molecular model is displayed with a callout box stating "1. Toggle through each conformation here." pointing to a dropdown menu. Another callout box states "2. Click here to calculate the score." pointing to a button. The interface also shows a table of docking results and a command window.

core	Strain	RecConf	Steric	Torsion	Electro	Hbond	Hydroph	Surface
15.61	11.34	1pwm_conf1_rec1:1	-14.4	1	3.216	-3.671	-3.729	10.47
12.89	18.88	1pwm_conf1_rec1:1	-25.52	1	1.871	-2.887	-3.333	13.78
15.32	13.55	1pwm_conf1_rec1:1	-18.18	1	4.072	-3.654	-3.406	14.8
11.06	12.88	1pwm_conf1_rec1:1	-20.53	1	4.881	-0.4725	-3.77	12.5
12.65	12.31	1pwm_conf1_rec1:1	-18.38	1	10.76	0.2644	3.072	12.88

Command window output:

```

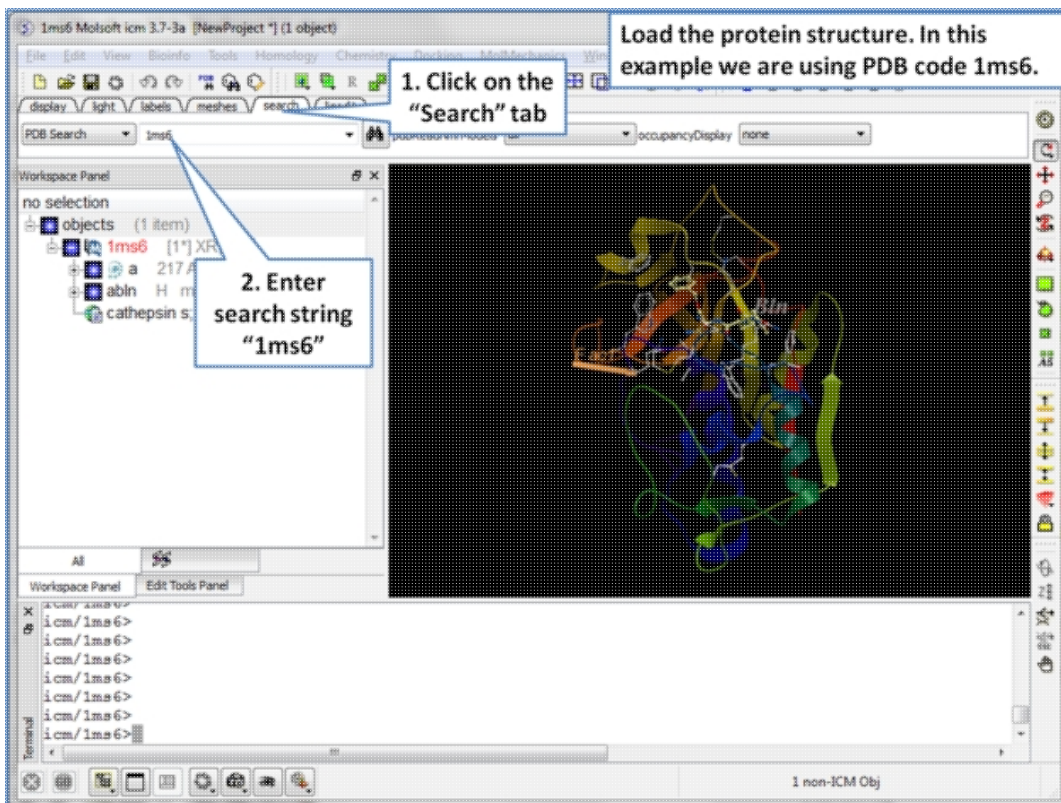
dateLigandScore a_LIG.I 0. 0. yes
!Load4DRecConf 1
dateLigandScore a_LIG.I 0. 0. yes
  
```

Bottom status bar: 5 Obj (3 non-ICM)

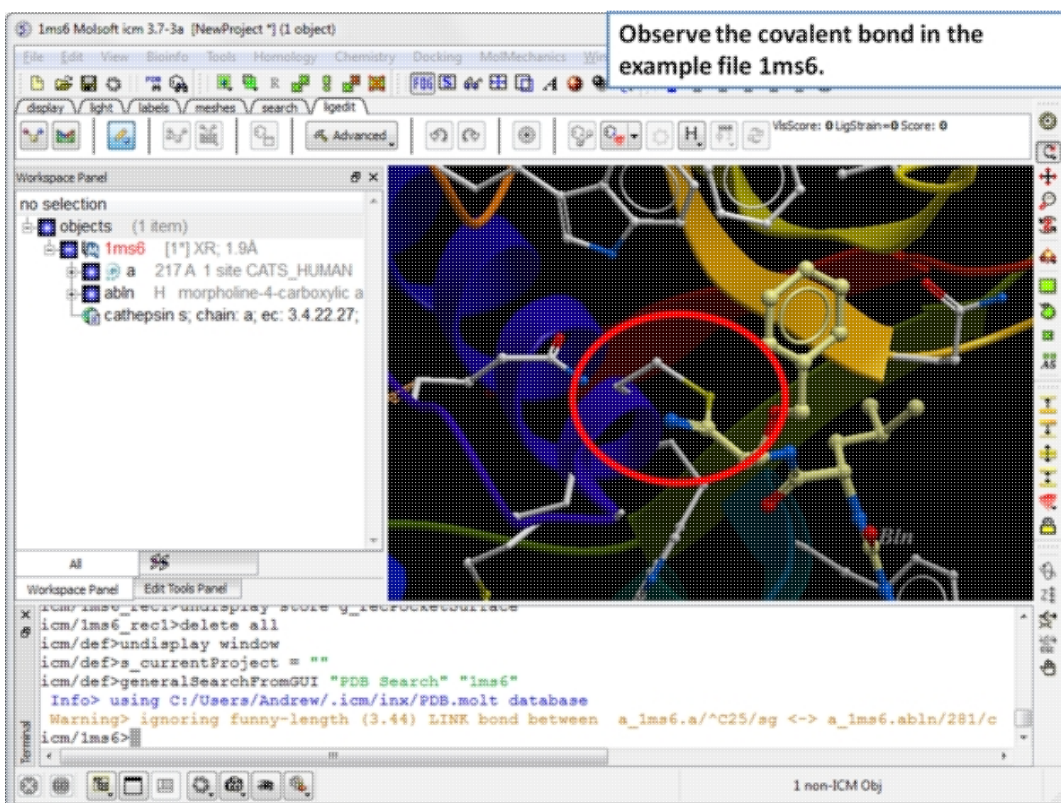
Step 9: Toggle through the results and view the score.

14.15 Covalent Docking in the Ligand Editor

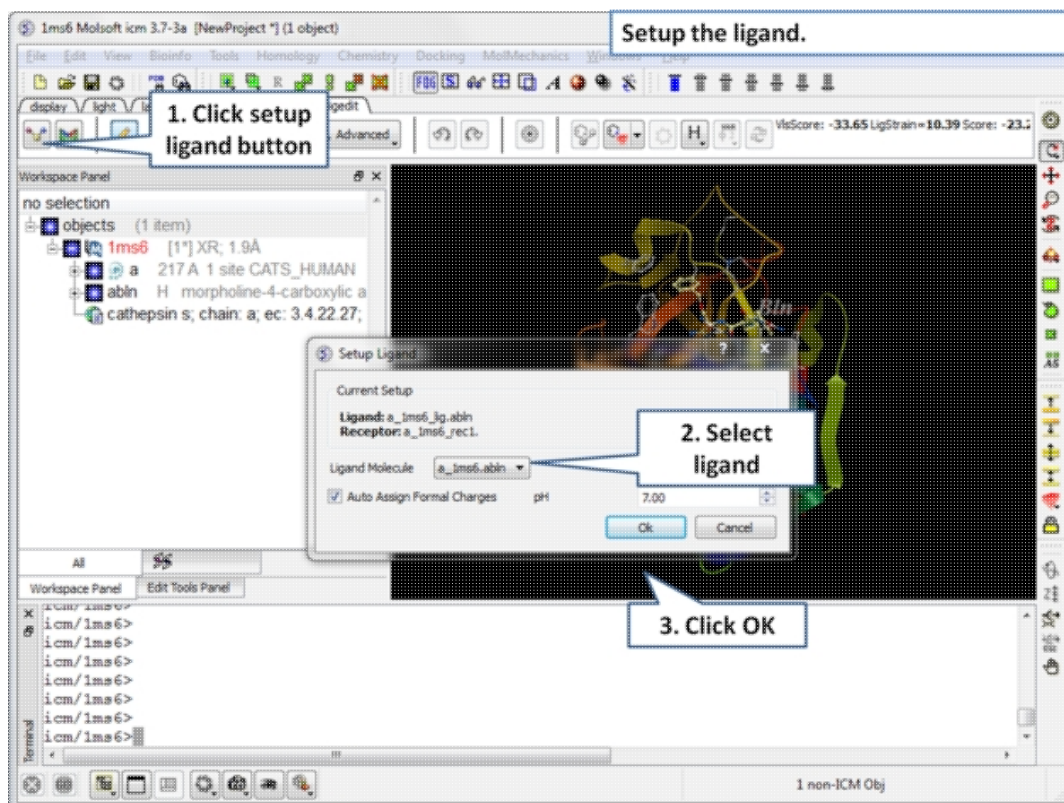
Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro



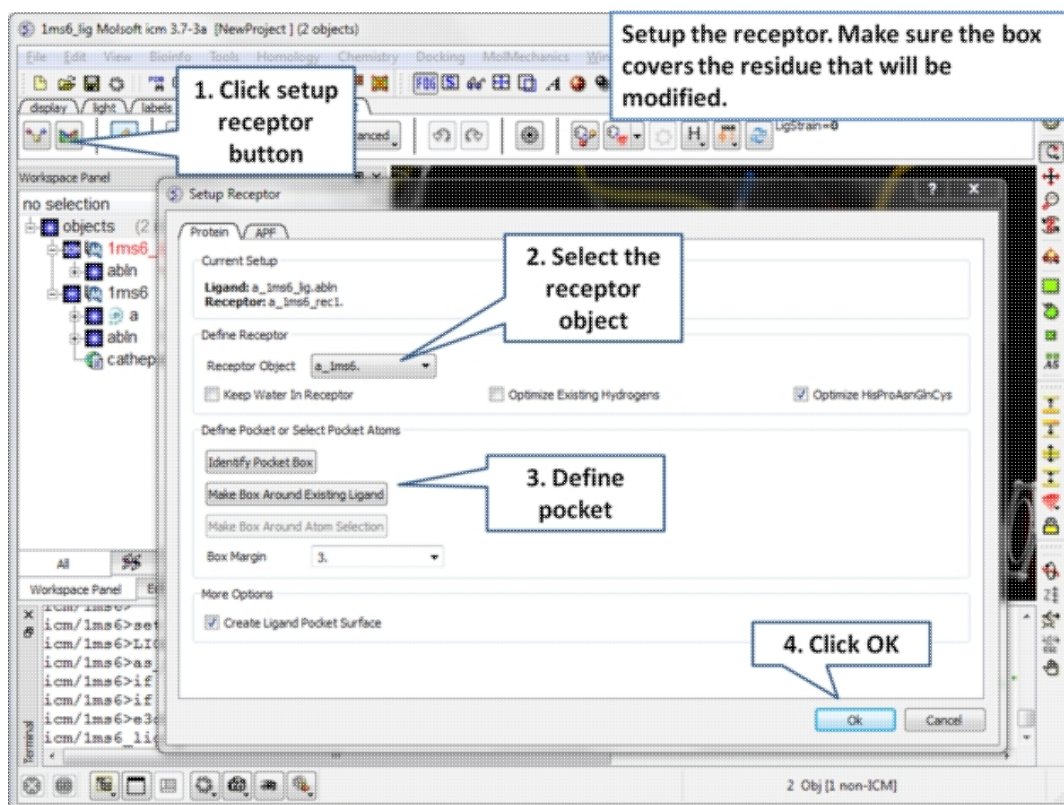
Step 1: Load the protein structure via the Search tab. In this example we will use PDB structure 1ms6.



Step 2: The structure 1ms6 has a covalent bond between the ligand "abln" and the receptor residue C25.



Step 3: Choose the ligand "abln" and setup the ligand in the regular way.



Step 4: Setup the receptor and make the maps. It is important to make sure the box encompasses the modified residue C25.

1. Click the edit tools button

2. Click the double bond button

3. Click on the nitrile bond

In this example we need to modify the nitrile in the ligand to obtain a terminal hydrogen to define connection to the receptor. If compound is represented in unbound state then one needs to make the necessary modifications. In this particular example: triple bond (from unbound state) -> double

```

icm/lms6_lig>processLigand a_lms6_lig.abln/281/c' "makebond" + "atom" Sarray( 1, "1" )
Warning> cannot make bond of this type
icm/lms6_lig>processLigand a_lms6_lig.abln/281/c' "makebond" + "atom" Sarray( 1, "1" )
Warning> cannot make bond of this type
icm/lms6_lig>processLigand a_lms6_lig.abln/281/c' | a_lms6_lig.abln/281/n40 "makebond" + "bond" Sarray( 1, "1" )
icm/lms6_lig>processLigand a_LIG.I "undo" Sarray()
icm/lms6_lig>
  
```

Step 5: Modify the nitrile in the ligand to obtain a terminal hydrogen to define the connection to the receptor.

1. Use the pick atom button

2. Pick a hydrogen atom in the ligand to attach a modifier

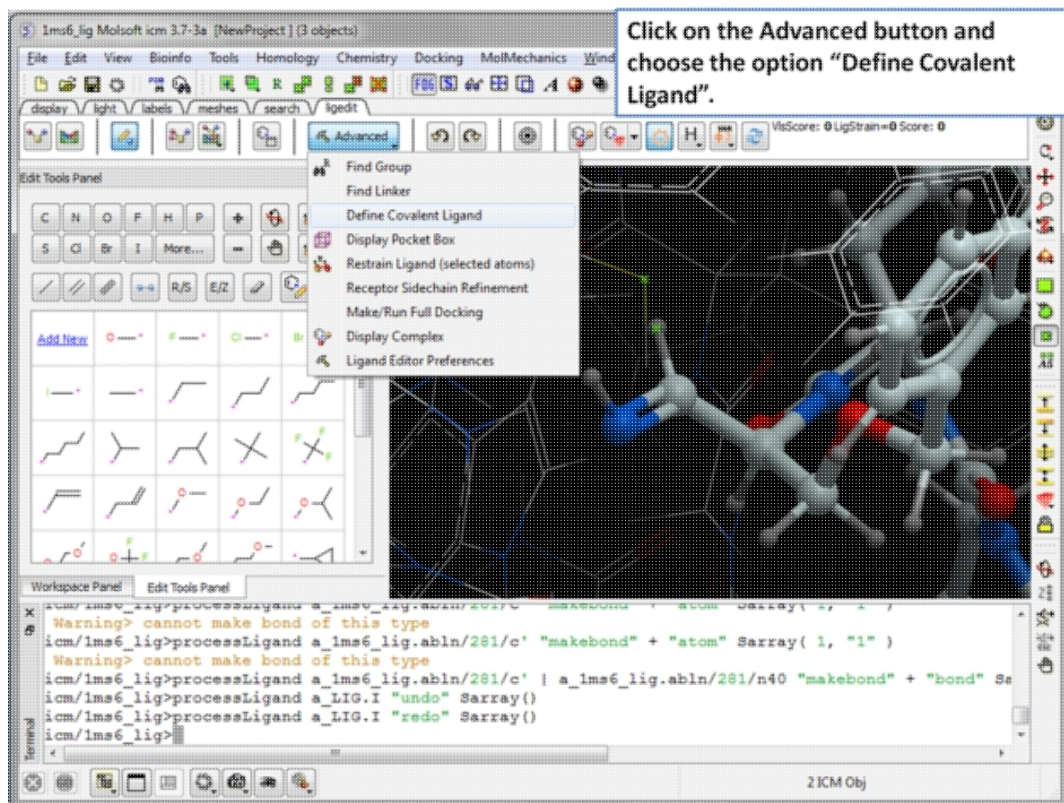
3. Pick any atom in the modified residue

Pick a hydrogen atom in the ligand to attach a modifier and pick any atom in the modified residue.

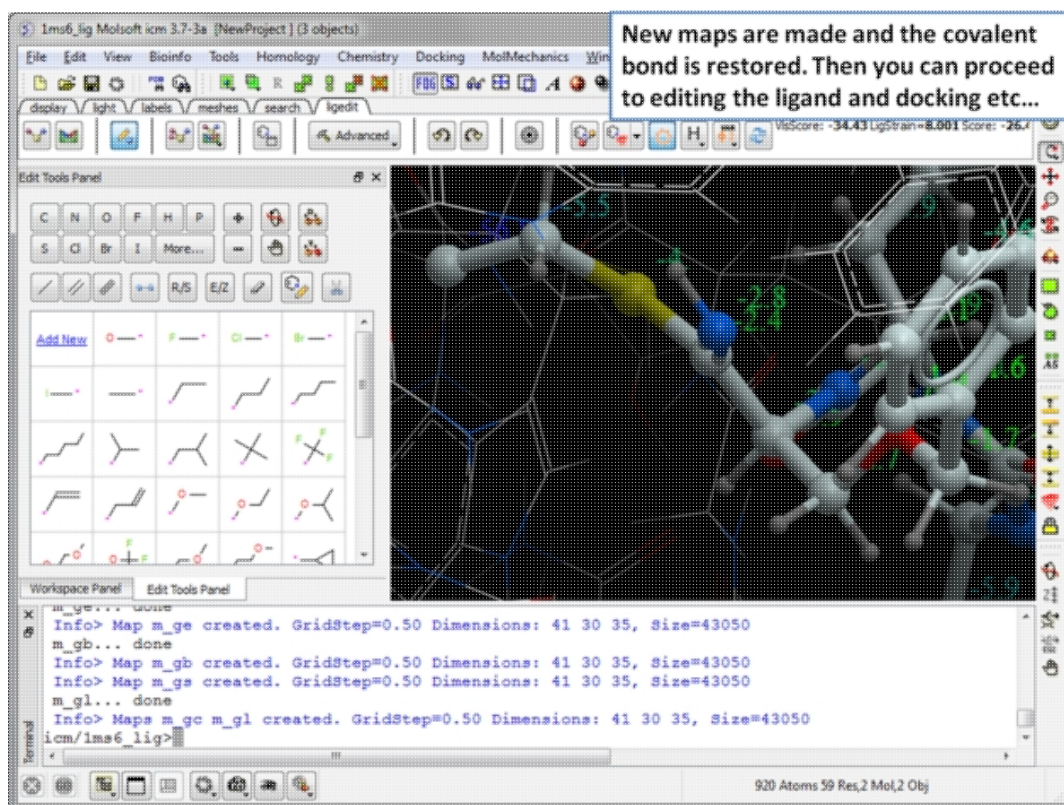
```

icm/lms6_lig>processLigand a_lms6_lig.abln/281/c' "makebond" + "atom" Sarray( 1, "1" )
Warning> cannot make bond of this type
icm/lms6_lig>processLigand a_lms6_lig.abln/281/c' | a_lms6_lig.abln/281/n40 "makebond" + "bond" Sarray( 1, "1" )
Warning> cannot make bond of this type
icm/lms6_lig>processLigand a_LIG.I "undo" Sarray()
icm/lms6_lig>processLigand a_LIG.I "redo" Sarray()
icm/lms6_lig>
  
```

Step 6: Pick the hydrogen atom in the ligand where the modifier will be attached and then pick any atom in the modified residue.



Step 7: Click on the Advanced button and choose the option "Define Covalent Ligand".



Step 8: New maps will be made and the covalent bond is restored. Then you can re-dock, minimize and edit the ligand in the usual way inside the ligand editor.

14.16 Dock to Atomic Property Fields

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

The Atomic Property Fields (APF) superposition/alignment method was first reported by Maxim Totrov PhD (Principal Scientist - MolSoft) at the 2007 233rd American Chemical Society National Meeting, Chicago, IL USA and then published [here](#).

APF is a 3D pharmacophoric potential implemented on a grid. APF can be generated from one or multiple ligands and seven properties are assigned from empiric physico-chemical components (hydrogen bond donors, acceptors, Sp² hybridization, lipophilicity, size, electropositive/negative and charge). Here we describe template APF superposition whereby the APF is generated from a single or multiple template and is then globally optimized with the internal force-field energy of the ligand.

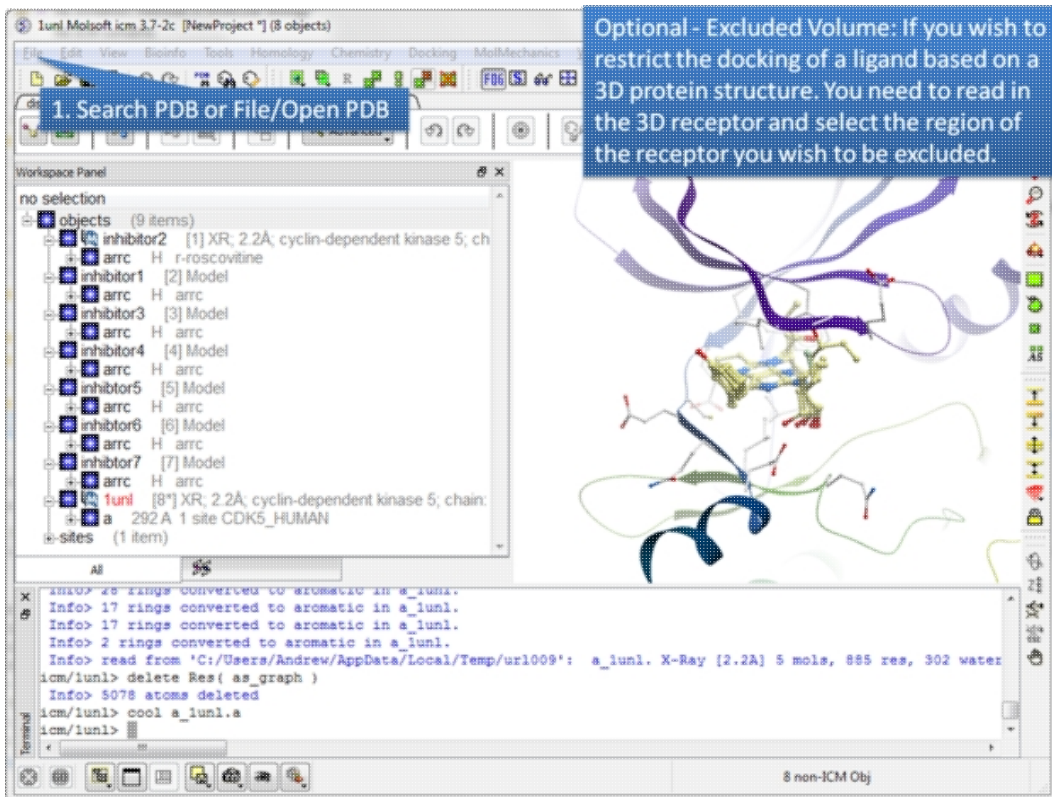
Read in 3D chemicals for APF template. Superimpose using the tools in the Chemistry menu. These chemicals will act as the template for the APF field.

1. Tools to superimpose chemicals can be found in the Chemistry menu.

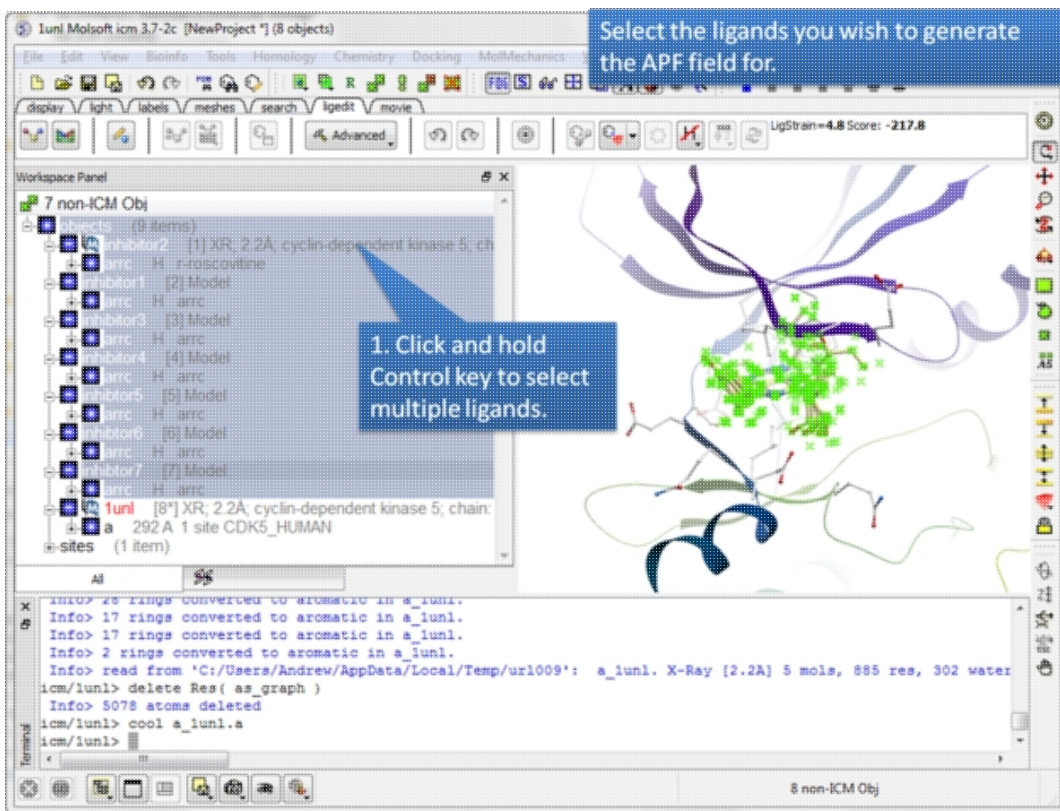
```
icm/inhibitor7> read a_inhibitor7_copy. NAME{NAME{ "inhibitor7" "simple", Object}}
icm/inhibitor6> editSelection a_inhibitor7.arcc yes no no
icm/inhibitor7> delete a_lunl.
Info> 1 Object deleted
icm/inhibitor7> color Res(a_*/DD) s a_*/c* object xstick
icm/inhibitor7> color background rgb=(255,255,255)
icm/inhibitor7> web "http://www.ncbi.nlm.nih.gov/pubmed/15689152"
icm/inhibitor7> color xstick Res(a_*/DD)
icm/inhibitor7>
```

7 non-ICM Obj

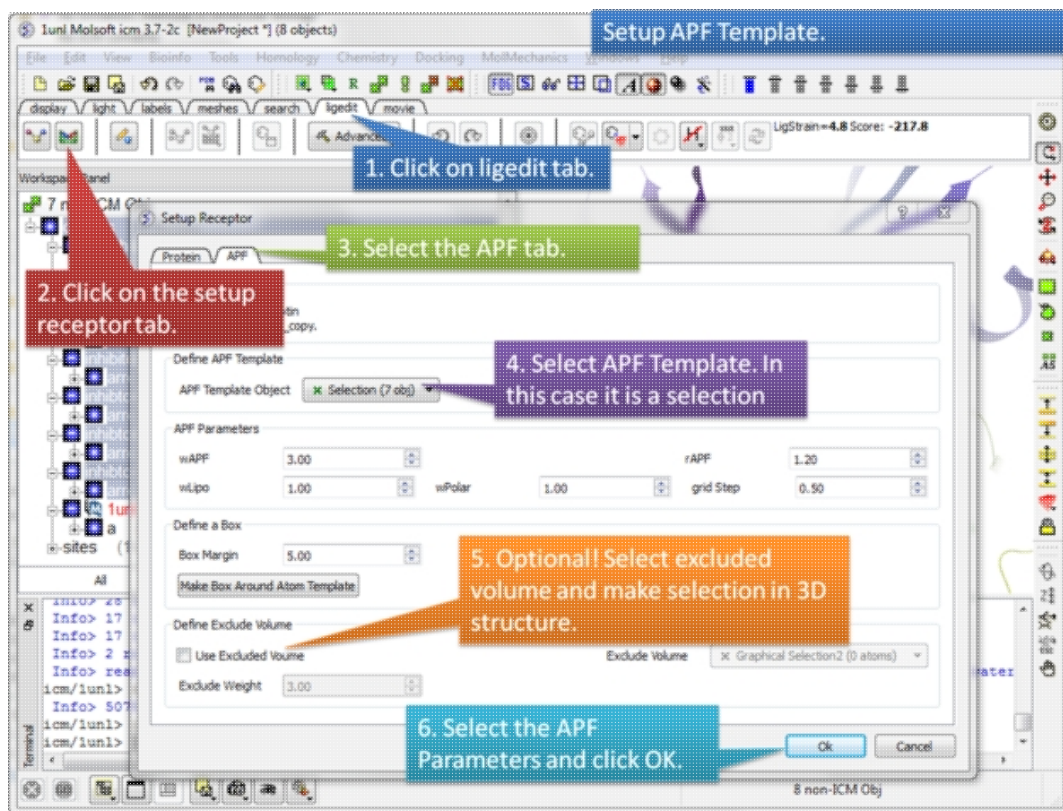
Step 1: Read in the chemicals that will make up the Atomic Property Field you wish to dock to.



Step 2: Certain regions of the pocket can be excluded from docking by selecting an excluded volume.

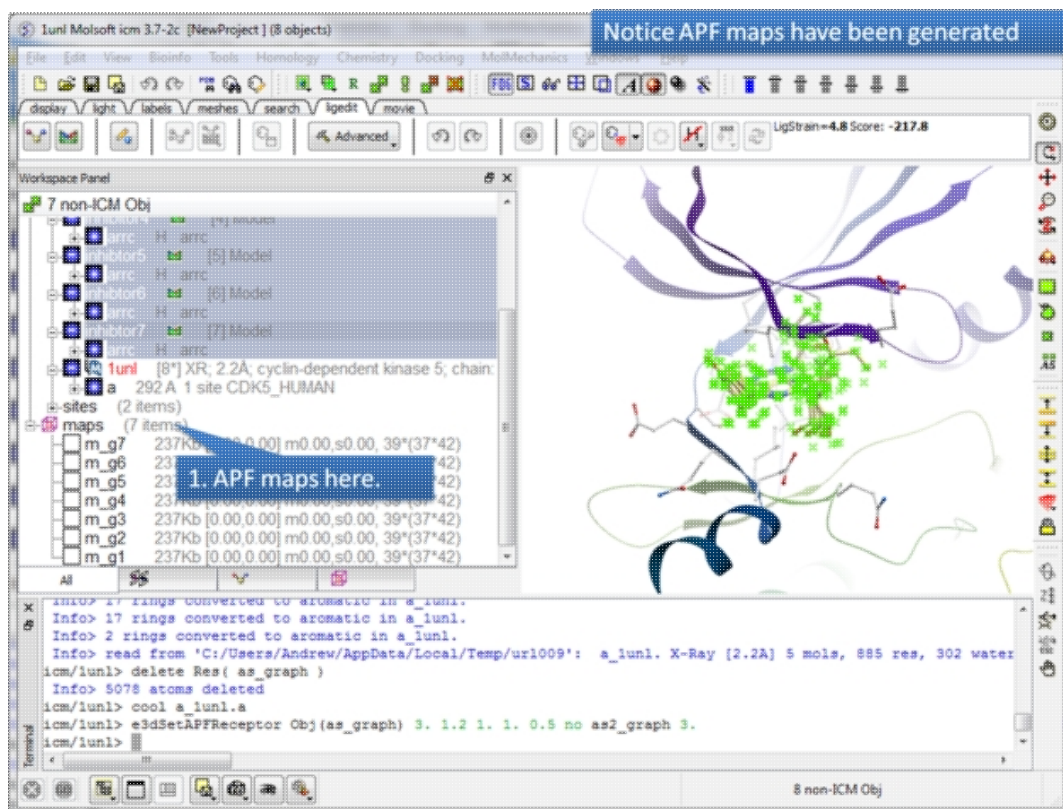


Step 3: Select the ligands you wish to generate an APF field for.

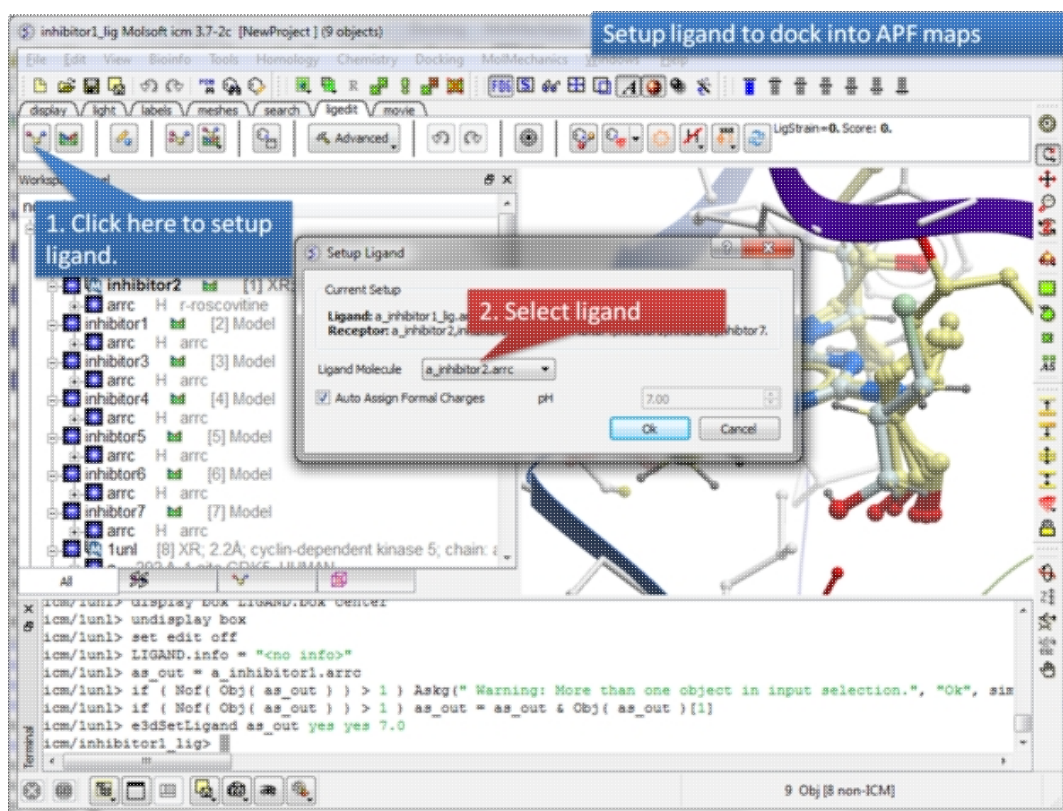


Step 4: Setup the APF template. The APF parameters are:

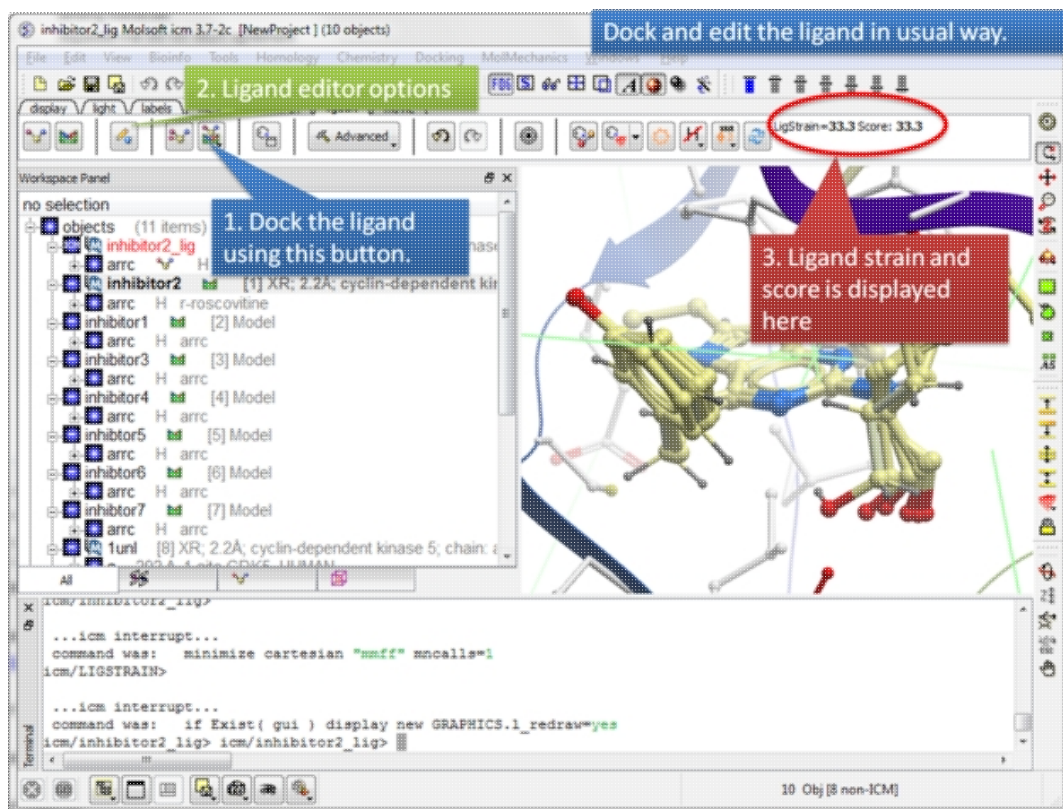
- wAPF = relative weight of APF with respect to internal force-field energy. Adjusts tolerance to strain relative to enforcement of chemical alignment.
- wLipo = relative weight of lipophilic atoms.
- wPolar = relative weight of polar atoms.
- rAPF = radius of APF gaussians. Can be used to adjust 'softness'/tolerance of molecular alignment.
- grid step = use with caution, potentially can be reduced to allow very 'sharp' alignments.



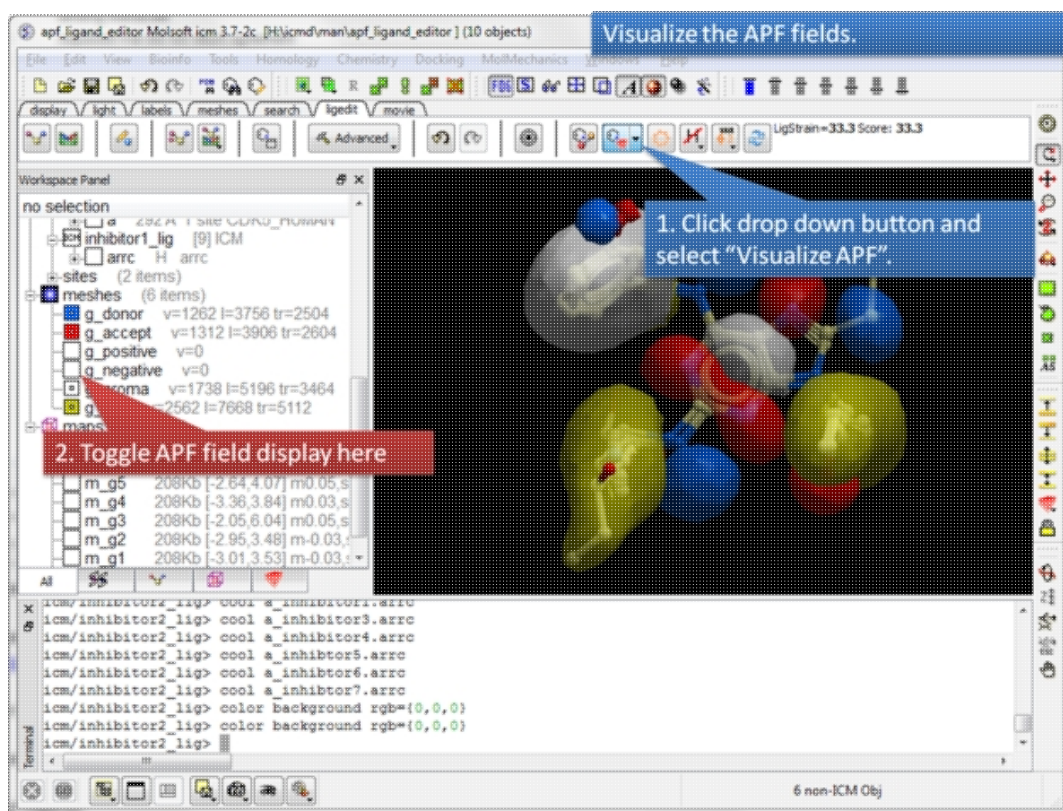
Step 5: The APF maps are listed in the ICM Workspace left hand side.



Step 6: Select the ligand you wish to dock into the APF maps.



Step 7: Dock the ligand using the ligand docking button and note the Score.



Step 8: Visualize the APF maps.

14.17 How to Save Ligand Receptor Complex as PDB

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

How to save docked complex as a PDB file.
Drag and drop ligand into receptor object.

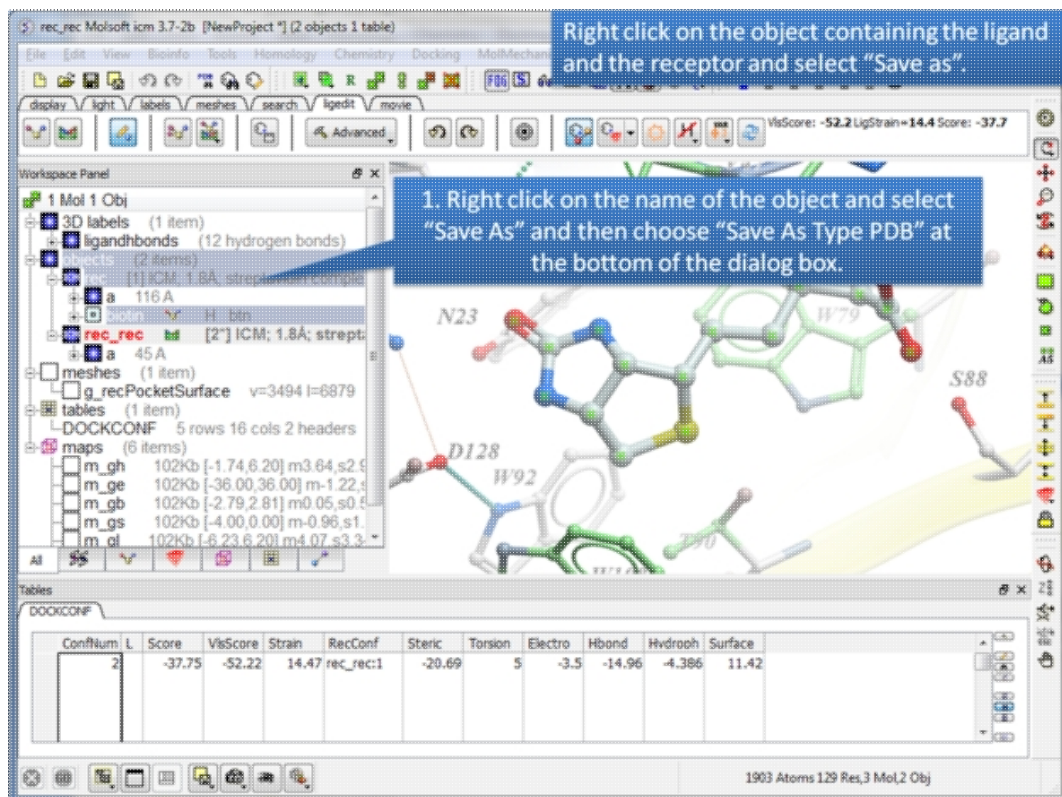
1. Double click on the ligand to select.

2. Drag the selected ligand to the receptor object and drop.

Conform	L	Score	VisScore	Strain	RecConf	Steric	Torsion	Electro	Hbond	Hydroph	Surface
	2	-37.75	-52.22	14.47	rec_rec1	-20.69	5	-3.5	-14.96	-4.386	11.42

1903 Atoms 129 Res, 3 Mol, 3 Obj

- Select the ligand and drag and drop it into the receptor to make a single object.



- Save the complex as a PDB file by right clicking on the complex in the ICM workspace and choosing "Save As".

14.18 Export Docking Project.

Available in the following product(s): ICM-Pro | ICM-VLS | ICM-Chemist-Pro

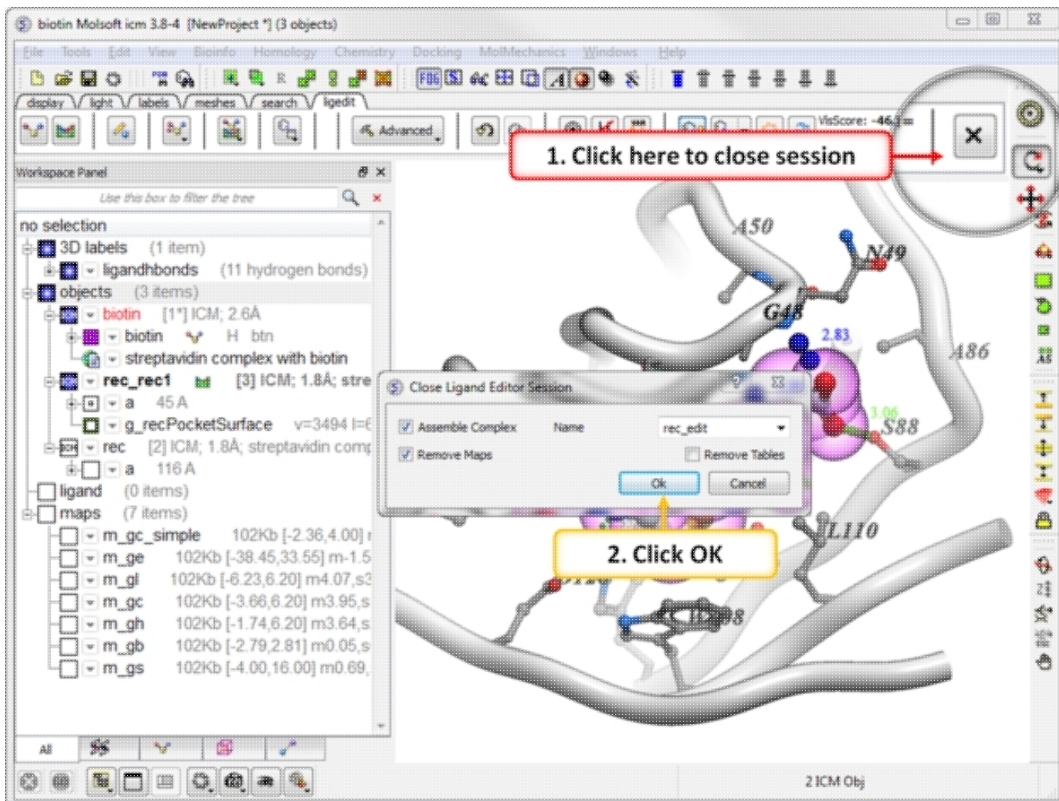
This option makes it easy to switch from the 3D ligand editor environment (ligedit tab) to the docking menu. All the docking setup steps you have created in the 3D ligand editor are converted so you can use the docking and screening options in the docking menu.

- Setup the ligand and receptor in the ligedit tab. Perform edits to the ligand etc...
- Click on the **Advanced** button in the ligedit tab
- Click on Export to Docking Project
- You can choose to run a docking job immediately with the current ligand or just create the docking files. Check or uncheck the **Run Docking Job** button.
- The docking files will be placed in your working directory (Tools/Change Working Directory). The files will have the docking project name dock_name_of_receptor and be stored in a directory called dock_name_of_receptor
- You can use these files to perform screening using the docking menu. For example you would set the docking project name as dock_name_of_receptor.

14.19 Close and Clean Up Project

The Close Session option allows you to close and clean up your ligand editor environment. Click on the close button and you can:

- Assemble the complex
- Remove maps
- Remove tables

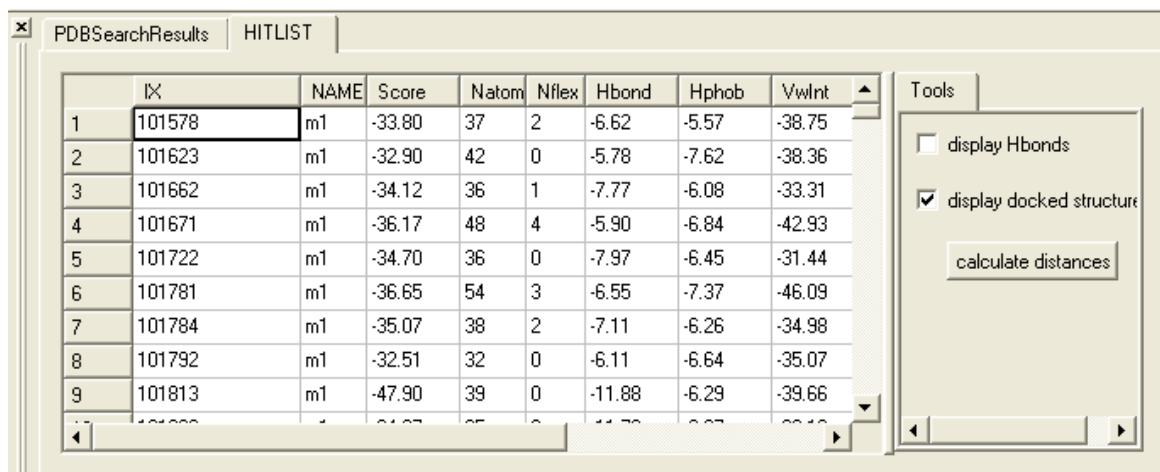


15 Working with Tables

One of the easiest ways to store, sort and display data in ICM is by the use of a table. In most cases tables are automatically created, for example, if you search for a PDB file or when you load a compound database (SDF file). It is also possible for you to create your own table. Once a table is created, ICM provides easy to use tools to sort, add, edit and plot data.

Here we will concentrate on describing the actions you can perform on a table once it has been read into ICM. We will start by describing a simple table. Actions which can be performed on chemical tables are described in the section entitled *Working with Chemical Spreadsheets*.

A standard ICM table:



The screenshot shows a window titled 'PDBSearchResults' with a sub-tab 'HITLIST'. It contains a table with the following data:

	IX	NAME	Score	Natom	Nflex	Hbond	Hphob	VwInt
1	101578	m1	-33.80	37	2	-6.62	-5.57	-38.75
2	101623	m1	-32.90	42	0	-5.78	-7.62	-38.36
3	101662	m1	-34.12	36	1	-7.77	-6.08	-33.31
4	101671	m1	-36.17	48	4	-5.90	-6.84	-42.93
5	101722	m1	-34.70	36	0	-7.97	-6.45	-31.44
6	101781	m1	-36.65	54	3	-6.55	-7.37	-46.09
7	101784	m1	-35.07	38	2	-7.11	-6.26	-34.98
8	101792	m1	-32.51	32	0	-6.11	-6.64	-35.07
9	101813	m1	-47.90	39	0	-11.88	-6.29	-39.66

To the right of the table is a 'Tools' panel with the following options:

- display Hbonds
- display docked structure
- calculate distances

Chapter Contents:

- Standard ICM Tables
- Molecular Tables
- Insert Interactive Objects into Table Cell
- Plotting Table Data
- Principal Component Analysis
- Learn and Predict
- Cluster

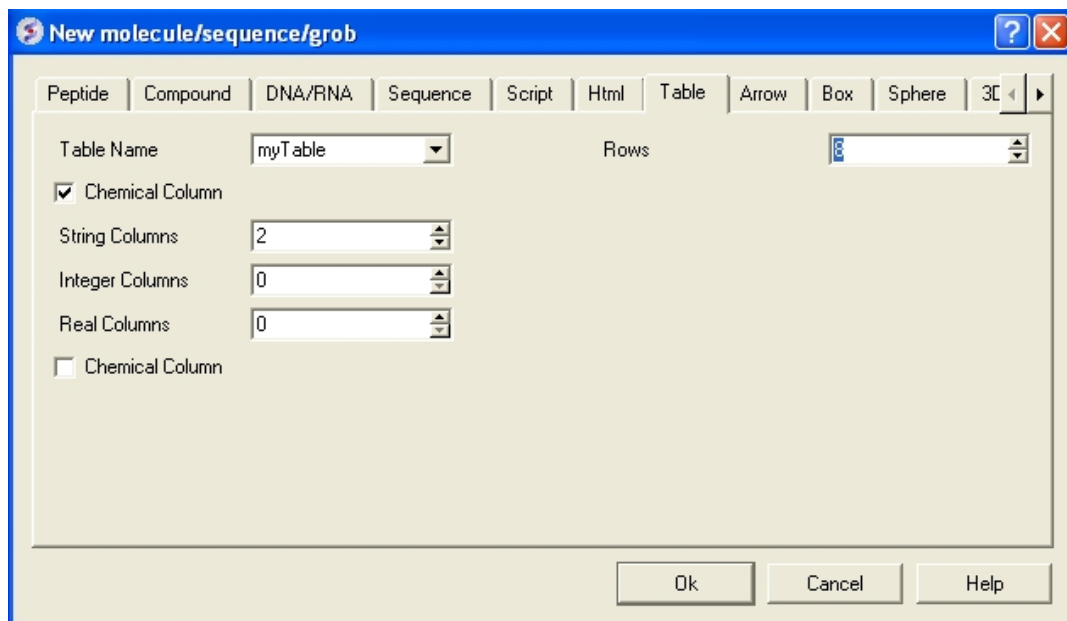
15.1 Standard ICM Tables

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

15.1.1 Generate New Table

To generate a new empty table:

- File/New and select the **Table** tab and a window as shown below will be displayed.
- Enter the number of rows and columns you wish to include in your table and whether you wish to add a column with chemical data.
- If you wish to make a chemical table (chemical spreadsheet) select the **Chemical Column** box.



15.1.2 Reading a Table

A table can be read and saved as a *.csv , *.tsv or a .tab file. Saving or reading your table as a csv (comma separated value) file enables the table to be transferred or loaded from other applications such as Microsoft Excel. A compound database such as an .sdf file can also be viewed as a table in ICM, additional details on how to manipulate a molecular table is explained in the next section.

A table can be read into ICM by selecting:

- File/Open and then selecting the table you have saved.

OR

Sometimes data is naturally stored and displayed in a table - e.g. PDB data. A common use of tables is for compound data. An explanation of how to use compound molecular tables is in the next section entitled ICM Molecular Tables.

For an example of a table try the following:

- Select PDB search tab.
- Type * into data entry box.
- Click on the button next to the data entry box.

A table of all the PDB structures will be displayed at the bottom of the GUI.

NOTE: If you have loaded a table and it is not displayed it may be because the table window is hidden. To display the table, select the window menu and select table see the [Window Menu Section](#).

15.1.3 Saving a table

To save the whole table:

- To save a table right click on the table header tab and select Save As..

To save a row selection:

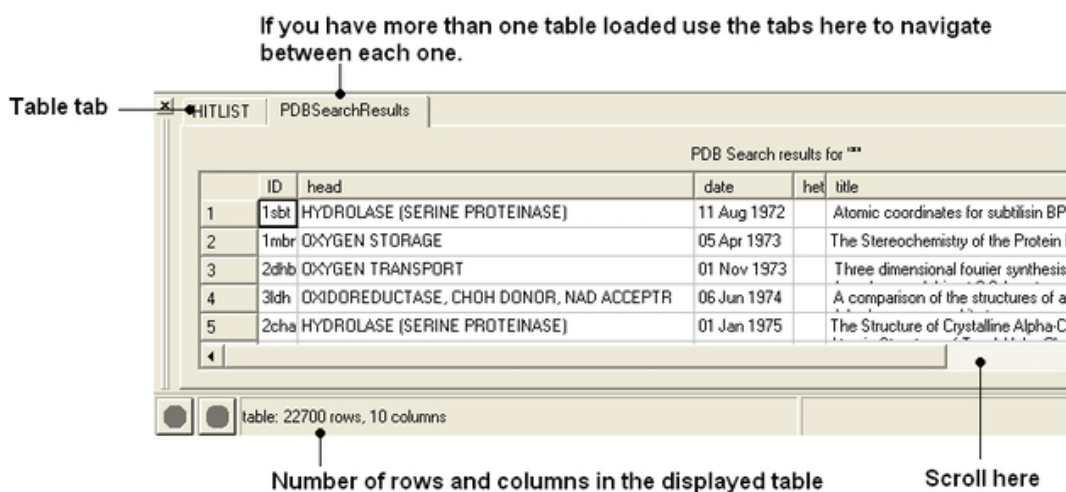
- Select a row(s)
- Right click and choose **Save Selection As** or **Save Selection As Csv + Headers**

15.1.4 Basic Table Navigation

To view the contents of a table you can move the table up and down using the scroll bars on the side and bottom of the display.

NOTE: If you have loaded a table and it isn't displayed it may be because the table display isn't selected. To select the table display, select the window menu and select table (See Window Menu Section).

If you have read more than one table in ICM you can select a table by clicking the tab on the top of the table (See Below).



NOTE: Double clicking on the tab allows two tables to be displayed at once. Double clicking again returns to the default table layout.

NOTE: Information regarding the number of rows and columns within a table is displayed at the bottom of the table.

If you would like the table to be the main window in the graphical user interface:

- Select Windows/Table->Main

15.1.5 Table View (Grid Layout)

To change the table view (layout):

- Select the columns you wish to display in grid view. No selection will place all columns in grid view
- Right click on a table row and select **Table View**
- You can view the table in **Grid View** and toggle between grid and standard view. You can define your own grid using the **Custom Grid** option or display the table in **Form View**.

NOTE: You can save a table view.

15.1.6 Change Row Height

To change the row height:

- Click and drag on the row separator
- Hold the Ctl key to change the height of all rows.

15.1.7 Table View Save

Once you have a table view that you want to keep. You can save it by:

- Right click on a table row and select **Store Views**
- Select **Save Current View**
- Enter a name for the table view and you can return to that view by repeating the first two steps above.
- You can rename, delete or restore view by right clicking on the name of the table view.

15.1.8 Table Search

To search a table:

- Right click on a table row and select **Find and Replace**. You can also use CTRL F.
- Enter a search string.
- Press the **Find** button.

15.1.9 Table Color

You can color your table based on values within a column by:

- Right click on the column header and select **Format**.
- In the **Background** panel select the color you desire eg **Single Color** or you can by a rainbow according to the data in the column. To edit the range of values relating to each color click on the pencil (edit) button as shown below.

The image shows a screenshot of a software interface with a table of data. The table has columns for 'Author', 'Method', 'Resolution', and 'PDB ID'. The 'Resolution' column is highlighted in blue. A callout box labeled '1. Right click and select format' points to the right-click context menu. A callout box labeled '2. Select rainbow' points to the 'Rainbow' option in the 'Background' panel of the 'Format' dialog box. A callout box labeled '3. Click here to display color editor' points to the 'Edit' button in the 'Background' panel. A callout box labeled '4. Edit gradient by clicking and dragging line' points to the gradient editor in the 'Background' panel. A callout box labeled '5. Right click here to edit color' points to the right-click context menu over the 'Resolution' column header.

15.1.10 Table Font

- Right click on the column header and select **Format**.
- Change the font using the options in the **Font** panel.

15.1.11 Table Alignment

- Right click on the column header and select **Format**.
- Change the font using the options in the **Alignment** panel.

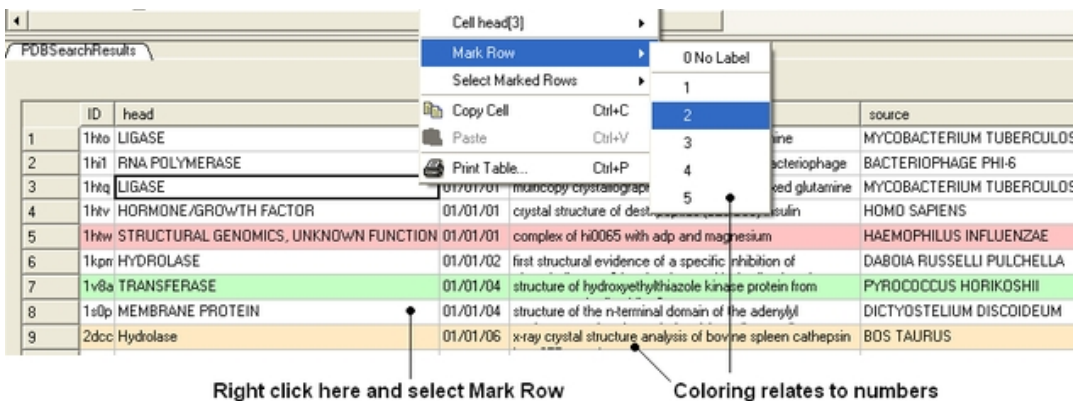
Rows can be colored by marking them as described here

15.1.12 Mark a Row

A row in a table can be marked and grouped by a label which enables the row(s) to be selected easily at a later time.

To mark a row

- Right click on the row in the table you wish to **mark**. Or select multiple rows and then right click.
- Select **Mark Row/** and then choose a number. In the GUI the number of rows that can be marked is limited to 5 but this can be increased using the command line command.
- A row that is marked will be colored - each number is assigned a color. The coloring can be changed in the gui tab in preferences.



To select marked rows

- Right click on the table and choose **Select Marked Rows** and choose a number which relates to the marked rows as described earlier.
- Selected rows will be highlighted blue - once rows are selected a number of right click options are activated such as copy selection to new ICM table.

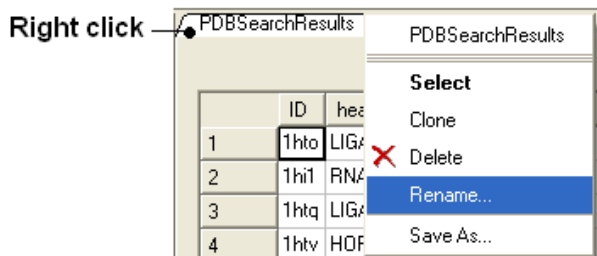
15.1.13 Table right click options

Right-click options vary according to where you click and what is selected. The options are intuitive, for example options that are performed on the whole table (eg Save and Delete) are performed by right-clicking on the Table tab. Other right-click options vary according to whether the row or column is selected or not.

15.1.14 Rename a Table

To rename a table:

- Right click on the table tab and select rename.
- Enter a new name and select OK.



15.1.15 Clone a Table

- Right click on the table tab and select clone.

15.1.16 Delete a Table

- Right click on the table tab and select delete.

15.1.17 Page Setup

Before printing a table you can change the orientation and scale.

To do this:

- Right click on the table header and select **Page Setup**.

15.1.18 Print a Table

A table can be printed by:

- Right click on the table tab and a menu will be displayed.
- Select the "Print" option. You may want to change the setup of the table (eg orientation and scale). You can do this using Page Setup option.

15.1.19 Export to Excel

To export a table to excel.

- Right click on the table header.
- Select the option to **Export to Excel**.

15.1.20 Save a Table

- Right click on the table tab and select Save As..

NOTE: You can save your table in comma separated format if you want to read it into another program such as Microsoft Excel.

15.1.21 Change Column and Row Width

To change the width of column and rows:

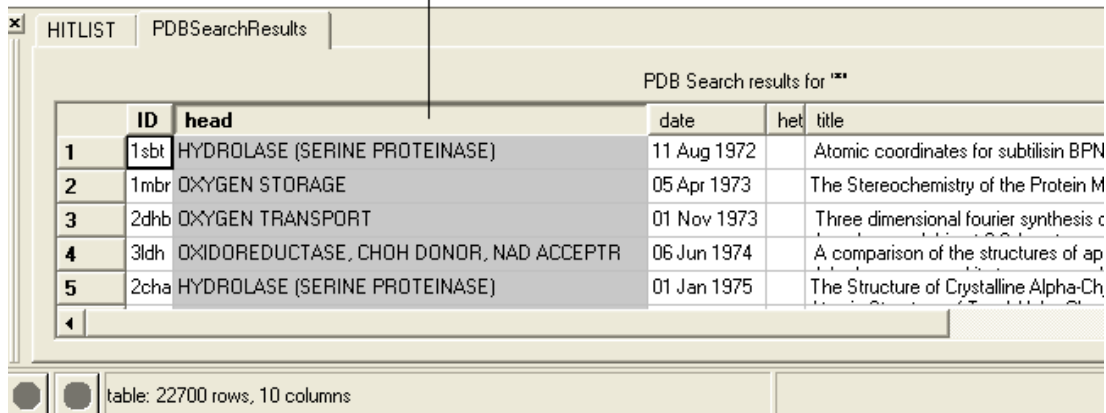
You can change the width of a row or column by clicking on the separating line and dragging. You can make each row the same width by holding down the **Shift** key and dragging one of the row edges.

15.1.22 Making Table Selections

To select one column of a table:

- Click on the column header

Click here to select a column



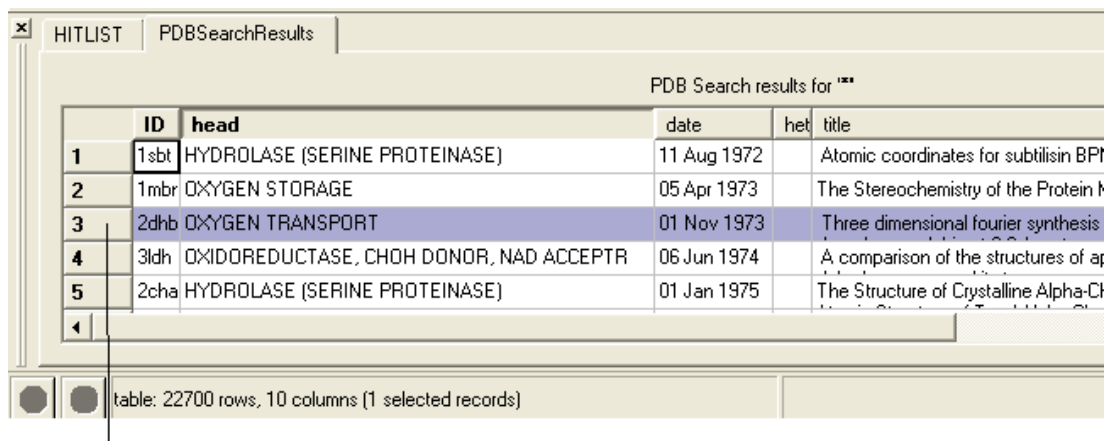
The screenshot shows a web interface with a table titled "PDB Search results for '***'". The table has five columns: ID, head, date, het, and title. The first five rows are visible. The 'head' column header is highlighted with a mouse cursor, and a callout box points to it with the text "Click here to select a column". Below the table, a status bar indicates "table: 22700 rows, 10 columns".

	ID	head	date	het	title
1	1sbt	HYDROLASE (SERINE PROTEINASE)	11 Aug 1972		Atomic coordinates for subtilisin BPN
2	1mbr	OXYGEN STORAGE	05 Apr 1973		The Stereochemistry of the Protein M
3	2dhh	OXYGEN TRANSPORT	01 Nov 1973		Three dimensional fourier synthesis c
4	3ldh	OXIDOREDUCTASE, CHOH DONOR, NAD ACCEPTR	06 Jun 1974		A comparison of the structures of ap
5	2cha	HYDROLASE (SERINE PROTEINASE)	01 Jan 1975		The Structure of Crystalline Alpha-Ch

table: 22700 rows, 10 columns

To select one row of a table:

- Click on the row header



The screenshot shows the same PDB search results table as above. The third row (ID: 2dhh) is highlighted in blue. A mouse cursor is positioned over the 'ID' header of this row. Below the table, a status bar indicates "table: 22700 rows, 10 columns (1 selected records)".

	ID	head	date	het	title
1	1sbt	HYDROLASE (SERINE PROTEINASE)	11 Aug 1972		Atomic coordinates for subtilisin BPN
2	1mbr	OXYGEN STORAGE	05 Apr 1973		The Stereochemistry of the Protein M
3	2dhh	OXYGEN TRANSPORT	01 Nov 1973		Three dimensional fourier synthesis c
4	3ldh	OXIDOREDUCTASE, CHOH DONOR, NAD ACCEPTR	06 Jun 1974		A comparison of the structures of ap
5	2cha	HYDROLASE (SERINE PROTEINASE)	01 Jan 1975		The Structure of Crystalline Alpha-Ch

table: 22700 rows, 10 columns (1 selected records)

To select a row click here

To select more than one row or column:

- Click on one row or column whilst pressing the Ctrl key
- Select multiple number of rows or columns whilst still pressing the Ctrl key

	ID	head	date	het	title
1	1sbt	HYDROLASE (SERINE PROTEINASE)	11 Aug 1972		Atomic coordinates for subtilisin BPN
2	1mbr	OXYGEN STORAGE	05 Apr 1973		The Stereochemistry of the Protein M
3	2dhb	OXYGEN TRANSPORT	01 Nov 1973		Three dimensional fourier synthesis
4	3ldh	OXIDOREDUCTASE, CHOH DONOR, NAD ACCEPTR	06 Jun 1974		A comparison of the structures of ap
5	2cha	HYDROLASE (SERINE PROTEINASE)	01 Jan 1975		The Structure of Crystalline Alpha-Cr

table: 22700 rows, 10 columns (3 selected records)

Select multiple rows and columns by clicking and selecting whilst pressing the Ctrl key.

NOTE: The Ctrl key acts as a toggle enabling select and unselect.

To select a range of columns or rows:

- Click on the first row or column in the range whilst pressing the Shift key.
- Click on the last row or column in the range whilst pressing the Shift key.

To select a range of columns or rows - click on the first member of the range and the last whilst pressing the shift key.

	ID	head	date	het	title
7	3lyz	HYDROLASE (O-GLYCOSYL)	01 Feb 1975		Real-space refinement of the struct
8	1lyz	HYDROLASE (O-GLYCOSYL)	01 Feb 1975		Real-space refinement of the struct
9	6lyz	HYDROLASE (O-GLYCOSYL)	01 Feb 1975		Real-space refinement of the struct
10	5lyz	HYDROLASE (O-GLYCOSYL)	01 Feb 1975		Real-space refinement of the struct
11	2lyz	HYDROLASE (O-GLYCOSYL)	01 Feb 1975		Real-space refinement of the struct
12	1chg	HYDROLASE ZYMOGEN (SERINE PROTEINASE)	01 Mar 1975		Chymotrypsinogen: 2.5-angstrom cr
13	2cna	LECTIN (AGGLUTININ)	01 Apr 1975		The covalent and three-dimensiona
14	1hip	ELECTRON TRANSFER (IRON-SULFUR PROTEIN)	01 Apr 1975		Two-Angstrom crystal structure of o
15	1gpd	OXIDO-REDUCTSE(ALDEHYDE/DONR,NAD/ACCP)	01 Jul 1975		Studies of asymmetry in the three-di

table: 22700 rows, 10 columns (6 selected records) 1 non-ICM

Click here hold the shift key

Click here hold the shift key

To invert a selection:

- Right click on the original selection and a menu will be displayed.
- Select the Row Selection/Invert selection option.

NOTE: Invert selection can only be used on rows.

To select the whole table:

- Right click in the table and a menu will be displayed.

- Select the Row Selection/Select All option.

To remove a selection:

- Click anywhere within the table.

A selection can also be made from a plot select(^ table-plot{See Select plot section}).

15.1.23 Editing a Table

To edit the contents of a table column:

- Select the column and then right-click on a column header and a menu will be displayed.
- Select the "Edit Mode" option. A tick will be displayed if it is selected.

OR

To edit the text or values within a cell:

- Right click on the table and select **Edit Cells by Double-click** .

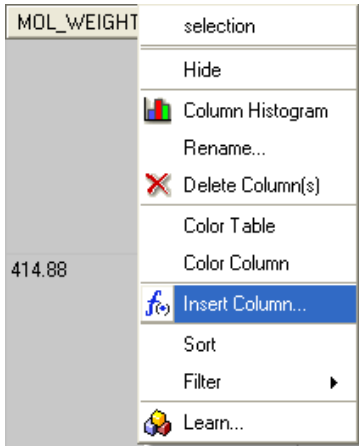
To edit the name of a column:

- Right click on the column header and a menu will be displayed.
- Select the option "Rename Column..." and enter the appropriate new text.

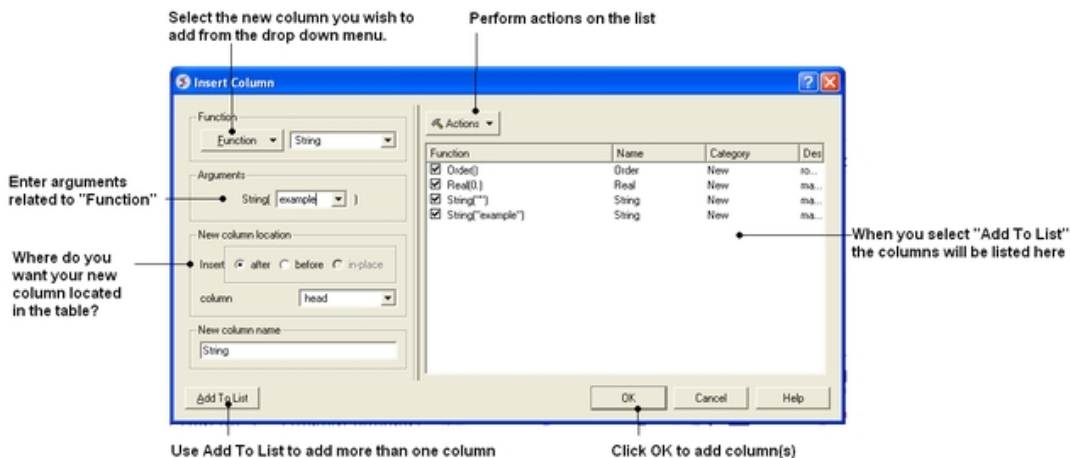
15.1.24 Inserting Columns

To insert a column:

- Identify the position within the table where you wish the column to be inserted.
- Right click on the column header and a menu will be displayed.
- Select "Insert Column"



A dialog box will then be displayed as shown below.



- Select the function you wish to add to the new column. Functions can be applied to many columns e.g. add etc..
- A set of arguments related to the function selected will then be displayed.
- Enter the appropriate arguments related to the function selected.
- Select where you want the new column to be located in the table.
- Enter the new column name
- If you wish to add multiple columns then use the **Add to List** option.

Many different functions are available:

- **New** Add a new column containing a real number, integer, string, or random number.
- **Transformations** A number of transformations can be selected and applied to a table column as shown below.
- **Mathematical** A number of mathematical functions
- **Text** Apply a number of different functions to the text in a column.
- **Chemical** Calculate a number of different chemical properties.
- **Convert Units** Radian to Degrees and Degree to Radian

Once the function and the correct arguments have been entered:

- Select whether you wish the new column to be added before, after or in place of this column.
- Enter the name of the new column.

NOTE: If you want to add more than one column choose **Add to List** and the action will be added to a list on the right hand side of the dialog box.

15.1.25 Column Statistics

To calculate various statistics describing columns and inter-column relationship:

- Right click on the column header and a menu will be displayed.
- Select "Column Statistics"

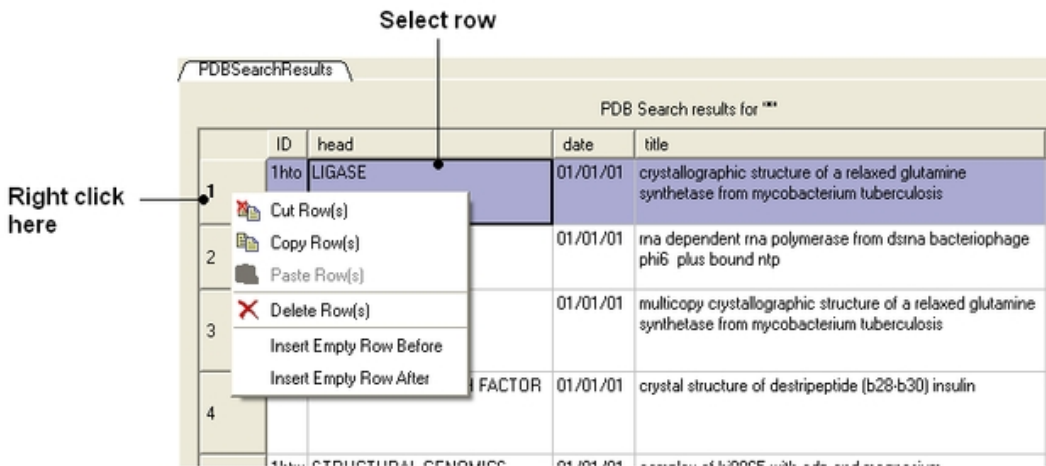
The output is printed into the ICM Terminal window and the Column Statistics Window.

15.1.26 Inserting Rows

To insert a row:

- Identify the position within the table where you wish the row to be inserted and select the row.
- Right click on the row name (eg the number of the row) and a menu will be displayed.
- Select **Insert Row Before** or **Insert Row After**.

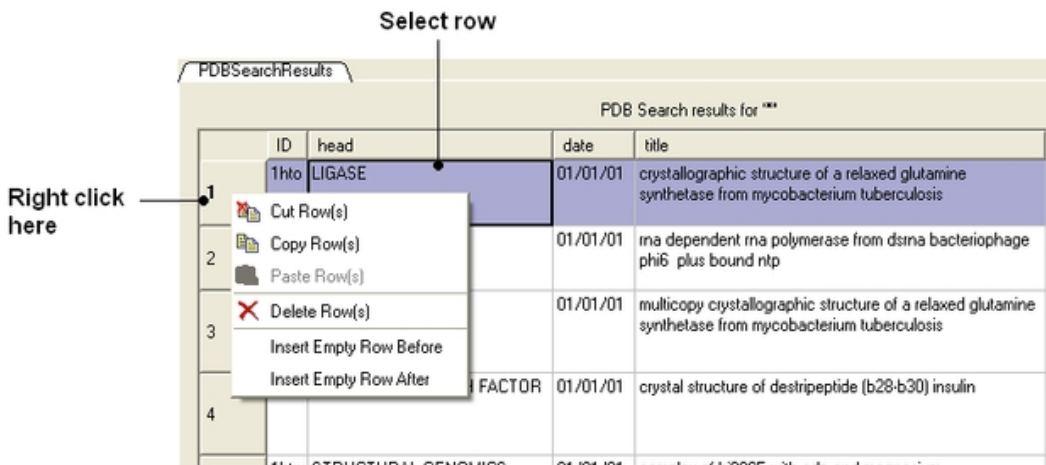
A blank row will be inserted. You can add data to this row by following the instructions in the edit table section.



15.1.27 Copy Cut and Paste Row

Copy, Cut and Paste Row:

- Select the row(s) See table selection section.
- Right click on the row header
- Select **Copy Row(s)**.
- To paste a row select the row header under which you wish to paste the row. Right click and select **Paste Row(s)**



15.1.28 Copy Cell

To copy a table cell:

- Right click on cell.
- Select Copy Cell - you can then paste it into a new table.

15.1.29 Copy Selection to an ICM Table

To copy a selection to a new table:

- Select the row(s) See table selection section.
- Right click on the row header
- Select **Copy Selection to ICM Table** and then choose Auto (ICM will name the table or New and you can enter a new table name).

15.1.30 Deleting Columns and Rows

To delete a column or row:

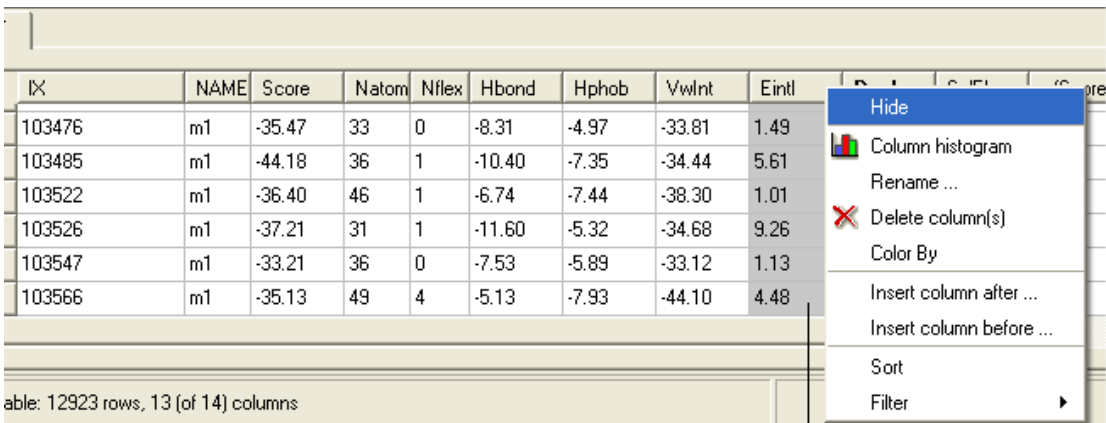
- Select the column(s) or row(s) you wish to delete. See the `select table` section for information on how to make table selections.
- Right click on the row to delete a row or right click on the column header to delete a column and select the delete option from the menu.

15.1.31 Hide and Show Columns

If you have a large table you may wish to only show and display certain columns and hide others. By default any loaded table will have all the columns displayed.

To select which columns you wish to hide:

- Select the column(s) you wish to hide. See the `select table` section for information on how to make table selections.
- Right click and select the hide option from the menu.



IX	NAME	Score	Natom	Nflex	Hbond	Hphob	Vwlnt	Eintl
103476	m1	-35.47	33	0	-8.31	-4.97	-33.81	1.49
103485	m1	-44.18	36	1	-10.40	-7.35	-34.44	5.61
103522	m1	-36.40	46	1	-6.74	-7.44	-38.30	1.01
103526	m1	-37.21	31	1	-11.60	-5.32	-34.68	9.26
103547	m1	-33.21	36	0	-7.53	-5.89	-33.12	1.13
103566	m1	-35.13	49	4	-5.13	-7.93	-44.10	4.48

able: 12923 rows, 13 (of 14) columns

Select column(s), right click and then select the hide option.

To show hidden columns:

- Right click on the column header and a menu will be displayed.
- Select the **Show Columns** options.
- Select which column you wish to show from the drop down list.

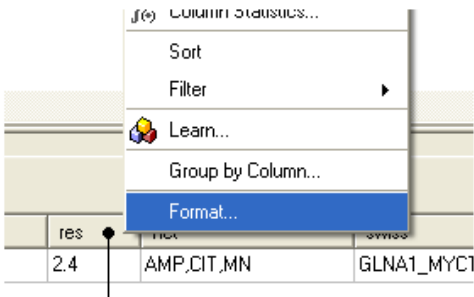
IX	NAME	Score	Natom	Nflex	Hbond	Hphob	Vv			
103476	m1	-35.47	33	0	-8.31	-4.97	-33			
103485	m1	-44.18	36	1	-10.40	-7.35	-34			
103522	m1	-36.40	46	1	-6.74	-7.44	-38			
103526	m1	-37.21	31	1	-11.60	-5.32	-34			
103547	m1	-33.21	36	0	-7.53	-5.89	-33			
103566	m1	-35.13	49	4	-5.13	-7.93	-44			
103592	m1	-36.01	47	2	-4.39	-7.99	-49			
103614	m1	-36.25	41	5	-8.71	-5.47	-29			
103615	m1	-34.95	49	1	-7.84	-7.68	-38			
103621	m1	-34.46	31	0	-6.87	-5.01	-39.21	2.08	9.51	
103626	m1	-32.70	34	0	-7.00	-4.72	-35.59	2.15	6.61	
103648	m1	-36.11	31	3	-7.40	-5.63	-35.08	4.93	2.93	
103707	m1	-32.11	50	4	-5.91	-7.89	-49.16	2.12	16.18	

table: 12923 rows, 10 (of 14) columns

15.1.32 Column Format and Custom Actions

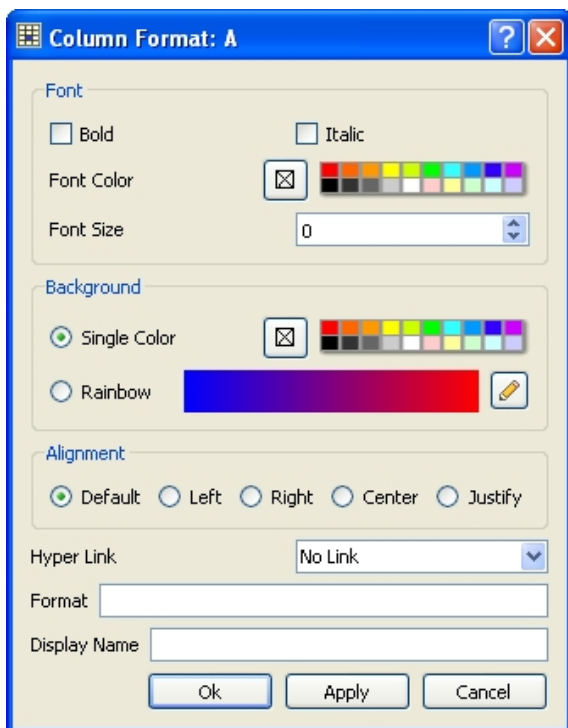
To change the **font** color or size, the **alignment** of the column data, the background color, add hyperlink, or script:

- Right click on the column header and select **Format**



Right click on the column header

- A window as shown below will be displayed.
- Make the desired changes and click **Apply**



To add a hyperlink to PubMed, PDB or Uniprot:

- Click on the drop down hyper link button.
- If the data in the column is a **PDB**, **Uniprot/SwissProt** or **PubMed** code then choose the built in format from the menu.
- Click Apply and the data in the column will become blue hyperlinks.

To add a user-defined hyperlink:

- Click on the drop down hyperlink button.
- Choose "Simple link" See below for details on how to access cell data.
- Click Apply and the data in the column will become blue hyperlinks.

To add an internal ICM link:

- Click on the drop down hyperlink button.
- Choose "Internal ICM link" See below for details on how to access cell data.
- Enter ICM scripting language in the panel that opens (see image below).
- Click Apply and the data in the column will become blue hyperlinks.

The value of the clicked cell is accessed as %1. To refer to the other cells and/or table itself the following shortcuts can be used:

- %@ # table name
- %# # clicked row number
- %^ # clicked column number

Example:

To add a an internal ICM link that reads a pdb file and then displays it you could use the following ICM commands in the link:

Column Format: ID

Font

Bold Italic

Font Color [Color Picker]

Font Size 0

Background

Single Color [Color Picker]

Alignment

Default Left Right Center Justify

Hyper Link Internal ICM link

Format `%1 `

```
read pdb %@.%^[%#]
cool
```

Display Name

Ok Apply Cancel

	ID	head
1	1hi1	RNA
2	1hto	LIGAS
3	1htg	LIGAS
4	1hty	HORM
5	1htw	STRU
6	1kpm	HYDF
7	1sOp	MEMB
8	2dca	Hydro

15.1.33 Table Sorting

To sort a table by a column value:

- Right click on the column header.
- Select the **Sort** option.

ID	head	date	help	title
1uot	REGULATOR OF COMPLEMENT PATHWAY	23 Sep 2003		
1r1c	ELECTRON TRANSPORT	23 Sep 2003		
1o5j	UNKNOWN FUNCTION	19 Sep 2003		
1o5h	STRUCTURAL GENOMICS, UNKNOWN FUNCTION	7 Sep 2003		
1o5i	OXIDOREDUCTASE	7 Sep 2003		
1qzr	ISOMERASE	7 Sep 2003		
1qvq	LUMINESCENT PROTEIN	11 Sep 2003		
1qvo	LUMINESCENT PROTEIN	11 Sep 2003		
1qvu	IMMUNE SYSTEM	10 Sep 2003		

Right click in the column header

The arrow represents ascending or descending order.

15.1.34 Table Filtering and Appending

There are two ways to filter table content one is via the table panel or by using the right click column header options.

To filter a table using the extra panel sliders:

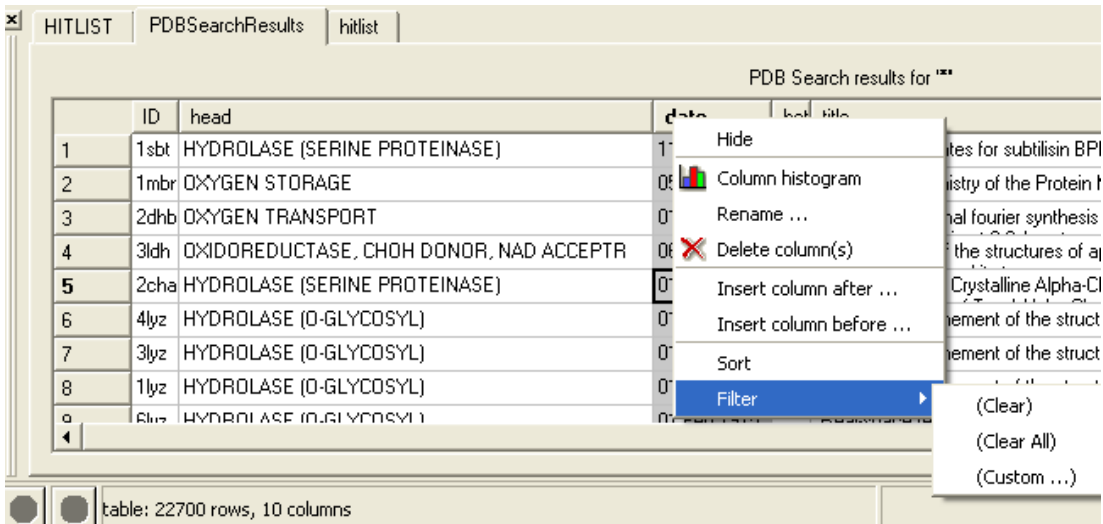
How to apply filters to a table.

1. Click here to view extra panel (right)
2. Click on the Filter tab
3. Click here to enter filters
4. Drag to make changes
5. Click to apply or remove filters
6. Symbol indicates column has been filtered

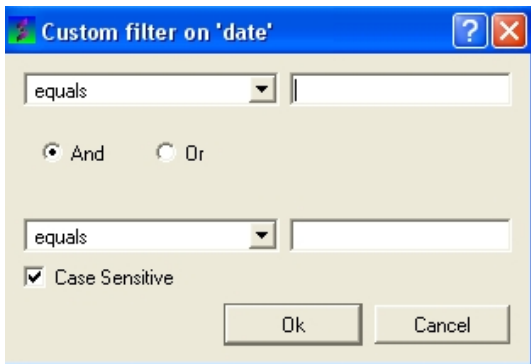
Open the extra panel and choose the filter tab. Click and drag on the sliders to conveniently choose the filtering values or enter the values directly. Click on the green 'tick' to apply filters.

To filter a table using the column header:

- Select the column you wish to filter. See the select table section for information on how to make table selections.
- Right click on the column header.
- Select the **Filter** option.



- Select the "Custom" option and a data entry box as shown below will be displayed.
- Enter the appropriate operations and filter values for your search.
- Click OK.



NOTE: When a column has been filtered a symbol as shown below will appear in the header of the column.

The screenshot shows a table titled "PDB Search results for '12'" with columns: ID, head, date, het, and title. The 'date' column header has a small downward-pointing triangle icon. The table contains several rows of search results, with the row for ID 21264 highlighted. At the bottom of the window, a status bar indicates "table: 1441 (of 22700) rows, 10 columns".

	ID	head	date	het	title
	21260	1njp	RIBOSOME	02 Jan 2003	PPL THE CRYSTAL STRUCTURE OF '...
	21261	1njt	HYDROLASE	02 Jan 2003	ACE COMPLEX STRUCTURE OF HCM'...
	21262	1njs	TRANSFERASE	02 Jan 2003	HUMAN GAR TFASE IN COMPLE>...
	21263	1njq	METAL BINDING PROTEIN	02 Jan 2003	ACE NMR STRUCTURE OF THE SING...
	21264	1njp	RIBOSOME	02 Jan 2003	PPL THE CRYSTAL STRUCTURE OF '...
	21265	1nju	HYDROLASE	02 Jan 2003	DNI COMPLEX STRUCTURE OF HCM'...
	21266	1oa7	HYDROLASE	02 Jan 2003	STRUCTURE OF MELANOCARPL...
	21267	1oa6	HYDROLASE INHIBITOR	02 Jan 2003	THE SOLUTION STRUCTURE OF...
	21268	1oa5	HYDROLASE INHIBITOR	02 Jan 2003	THE SOLUTION STRUCTURE OF...

This symbol means that the table has been filtered according to data within this column.

To append the filtered information into a new table:

- Select the whole table either by right clicking or pressing Ctrl A.
- Right click on the table and select "Append to other table".
- Enter a new name for the table you are appending with your filter results.

OR

Selected rows can be appended to a new table by:

- Right clicking on the selected rows and a menu will be displayed.
- Selecting the "copy selection to ICM table" option.

A table can be filtered by a cell value:

- By clicking once in a cell.
- Right click and a menu will be displayed.
- Select the option "Filter by cell value".

A filter can be cleared by:

- Right clicking on the column selection and selecting Filter/Clear or Filter/Clear All

15.1.35 Mark and Select Rows

A row in a table can be marked and grouped by a label which enables the row(s) to be selected easily at a later time.

To mark a row

- Right click on the row in the table you wish to **mark**.
- Select **Mark Row/** and then choose a number. In the GUI the number of rows that can be marked is limited to 5 but this can be increased using the command line command.
- A row that is marked will be colored - each number is assigned a color. The coloring can be changed in the gui tab in preferences.

Right click here and select Mark Row

Coloring relates to numbers

ID	head	source
1	1hxo LIGASE	MYCOBACTERIUM TUBERCULO...
2	1hi1 RNA POLYMERASE	BACTERIOPHAGE PHI-6
3	1hmq LIGASE	MYCOBACTERIUM TUBERCULO...
4	1hkv HORMONE/GROWTH FACTOR	HOMO SAPIENS
5	1hvw STRUCTURAL GENOMICS, UNKNOWN FUNCTION	HAEMOPHILUS INFLUENZAE
6	1kpi HYDROLASE	DABOIA RUSSELLI PULCHELLA
7	1v8a TRANSFERASE	PYROCOCCUS HORIKOSHII
8	1s0p MEMBRANE PROTEIN	DICTYOSTELIUM DISCOIDEUM
9	2dcc Hydrolase	BOS TAURUS

To select marked rows

- Right click on the table and choose **Select Marked Rows** and choose a number which relates to the marked rows as described earlier.
- Selected rows will be highlighted blue - once rows are selected a number of right click options are activated such as copy selection to new ICM table.

15.1.36 Group View by Column

To group a column by value:

- Right click on the column header and choose "Group View by Column". Chemicals in the same group are grouped in the same row. You can expand the view by clicking on the "+" sign".

Group by column.

1. Right click and choose "Group view by column".

2. Each group can be expanded here to see all members of the group.

mol	Molecule Name	class
<chem>CCNCC1=CN=CN=C1</chem>	Histamine	
<chem>CC1=CC=C(C=C1)C2=NC(=NC=C2)N</chem>	Lamotrigine	32
<chem>NC(=O)N1C=NC=C1</chem>	Pyrazinamide	33

To remove grouping:

- Right click on the column header and choose "Undo the Group View"

15.1.37 Mouse and Cursor Actions on a Table

The actions resulting from a mouse click or cursor on a table can be changed by:

- Right click on a table and select Table View/Show Extra Panel
- A panel as shown below will be displayed.

The screenshot shows a table titled "PDB Search results for 'kinase'". The table has two columns: "Name" and "Value". The "Name" column contains the following entries: "tableTitle", "doubleClick", "cursor", and "separateTab". The "Value" column contains the following entries: "PDB Search results for 'kinase'", "nice \"%1\" no no no no delete a_\"MiniObj.", "cursorFindPDB %# \"%@\"", and "yes".

Annotations in the image:

- A right-click arrow points to the "thiazole kinase protein from" row in the main table, with the text: "Right click here and select Table View to display and undisplay extra panel".
- A double-click arrow points to the "doubleClick" row in the configuration table, with the text: "Double click here to edit actions".

- Double click in the **Value** column and the column can be edited. Add ICM commands for the action you want. A value in a column can be referred to using "%" e.g. column two would be referred to as "%2". In the example shown above the function nice is acting on the contents of column one for the double click action.

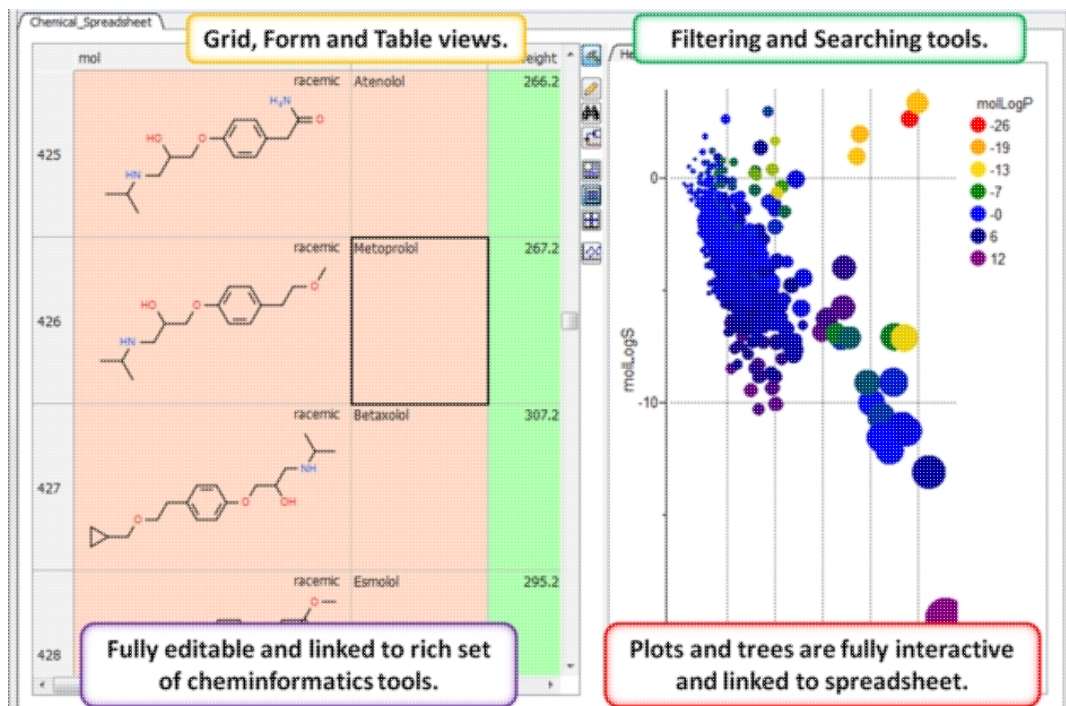
NOTE: The action associated with cursor and double click is placed in a variable name TableName.cursor and TableName.doubleClick

15.2 Molecular Tables

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

An ICM molecular table is created when an SDF or Mol file is read into ICM. To read and open a mol or sdf file go to File/Open (See Open an ICM file section) All of the table functions described in the previous section Standard ICM Table can be applied to molecular tables. Molecular tables are described in more detail in the Cheminformatics chapter.

An example of an ICM molecular table:



15.3 Insert Interactive Objects into Table Cell

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

NOTE: This option is available in Version 3.7-2b and above.

To store objects in a table please use the following:

Example:

```
read pdb "1crn"
display a_
add column t Parray( object a_ preview )
```

Viewing Objects from the table can be dragged and dropped into the ICM Workspace to display them again.

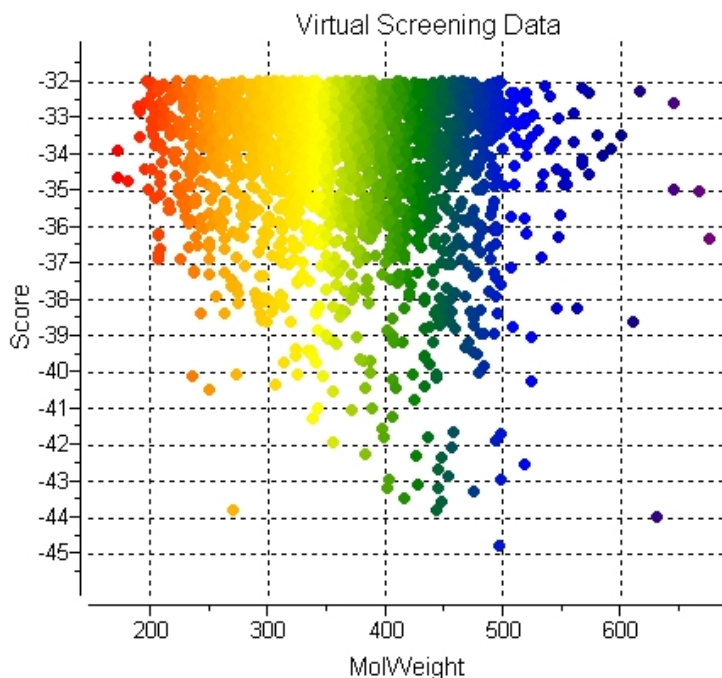
How to make the objects in the table fully interactive

- Right click on the column header and select "Live Rotatable View"

15.4 Plotting Table Data

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

The data within a table can be plotted graphically. A histogram can be made for the data within one column or a plot can be constructed for the data within two columns.



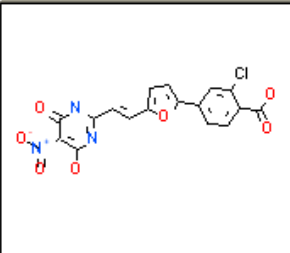
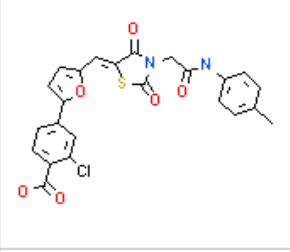
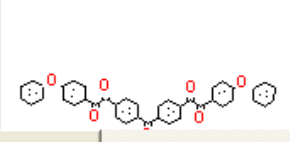
15.4.1 Column Histogram

To plot a histogram of the data within one column:

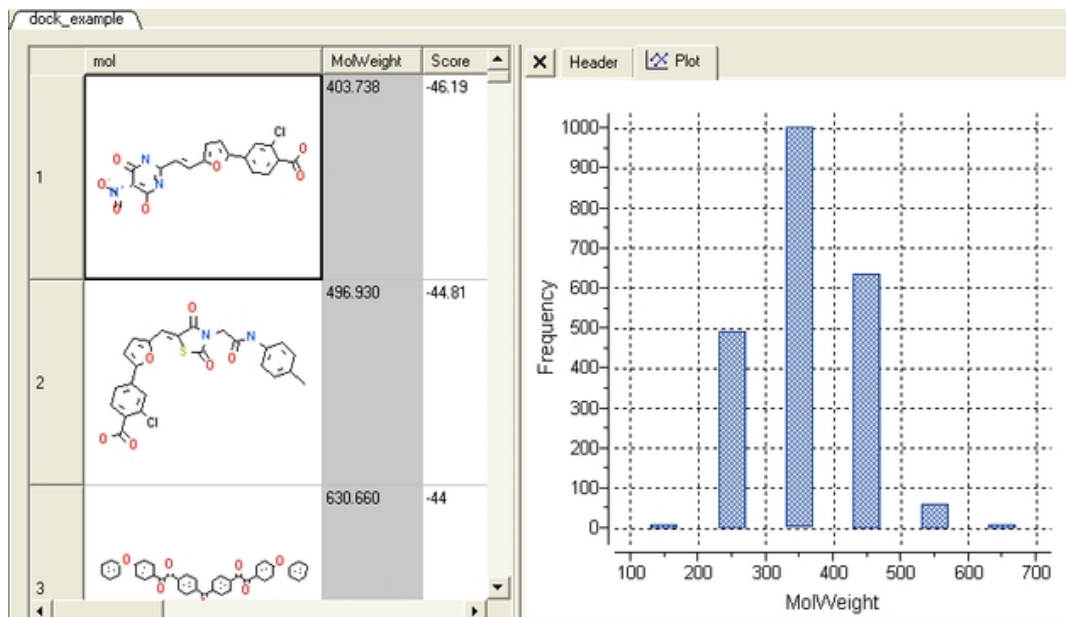
- Select the column by clicking on the column header.
- Right click on the column header.
- Select the **Column histogram** option.

Click here to select the column and then right click and select column histogram option

dock_example

mol	MolWeight	Score
	403.738	-46.19
	496.930	-44.81
	630.660	-44

A plot will then be displayed next to the table.

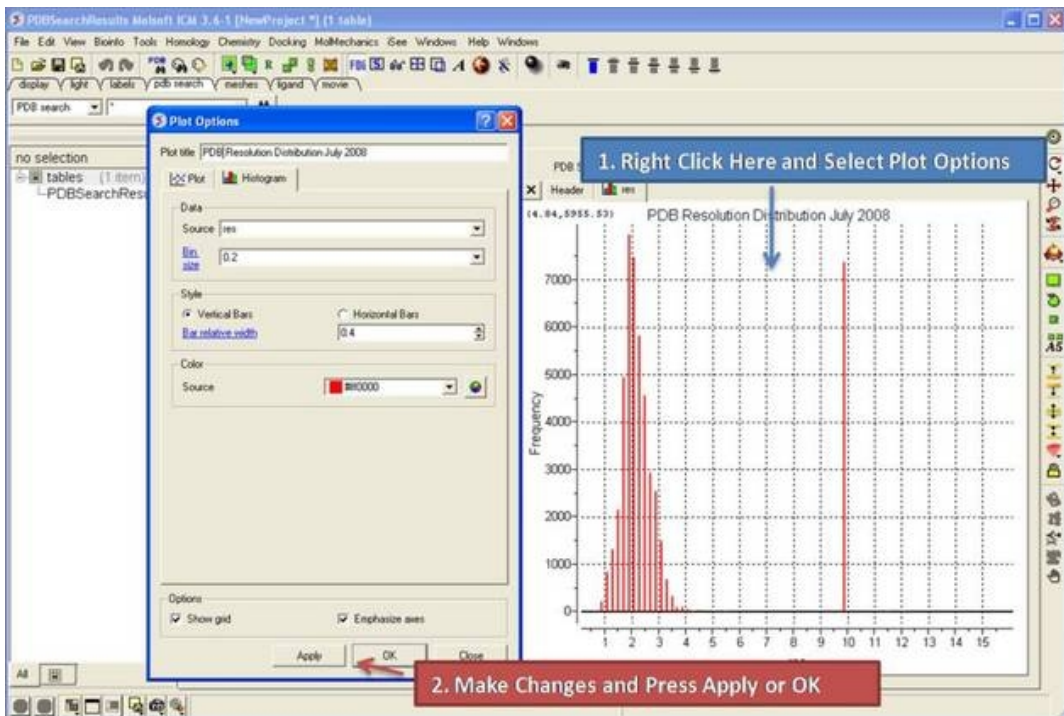


15.4.2 Histogram Options

Once you have created a histogram you can change the following parameters by right clicking on the plot and selecting options

Options:

- Change plot title.
- Change the data source using the drop down button and select another column in the table.
- Change the histogram bin size.
- Change the bars positioning from vertical to horizontal.
- Change the bar relative width compared to bin size. Bigger values give thicker bars.
- Color the bars



15.4.3 Histogram Bins

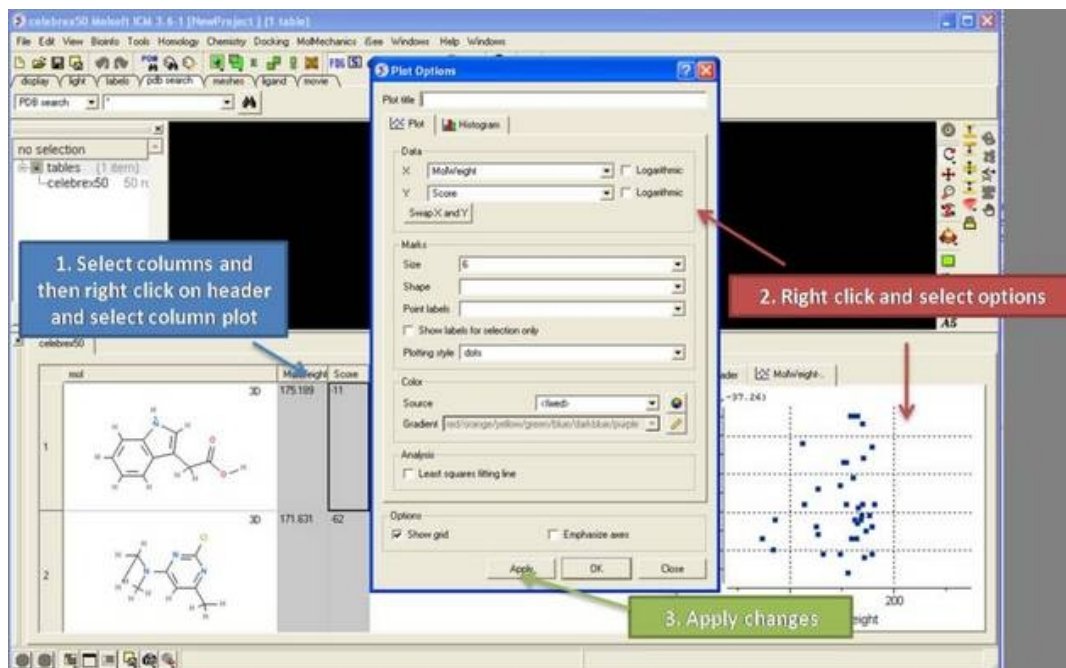
There are two ways to change the bin size. 1. Using the **options** dialog box or 2. interactively by left clicking and dragging at the top of the plot as shown below - this will allow you to find the best density estimation picture.



15.4.4 Plots

To construct a scatter plot from data within two columns:

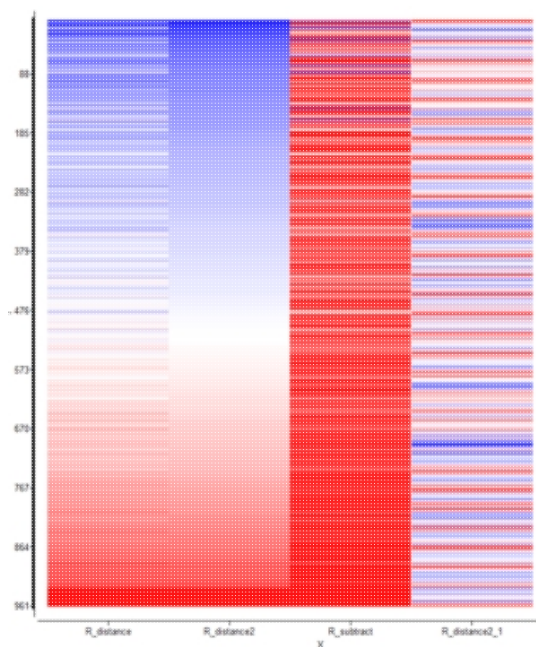
- Select the two columns.
- Right click on the column header.
- Select the **Columns** plot option.



15.4.5 Heatmap

To generate a heatmap:

- Read in a table containing the data.
- Click on the Plot/Histogram button on the side of the table.
- Click on the heatmap tab at the top.
- Add the data sources you want in the heatmap based on the column headers in the table.
- Check how you want to label the rows.
- Edit the color gradient if necessary.
- Click OK.



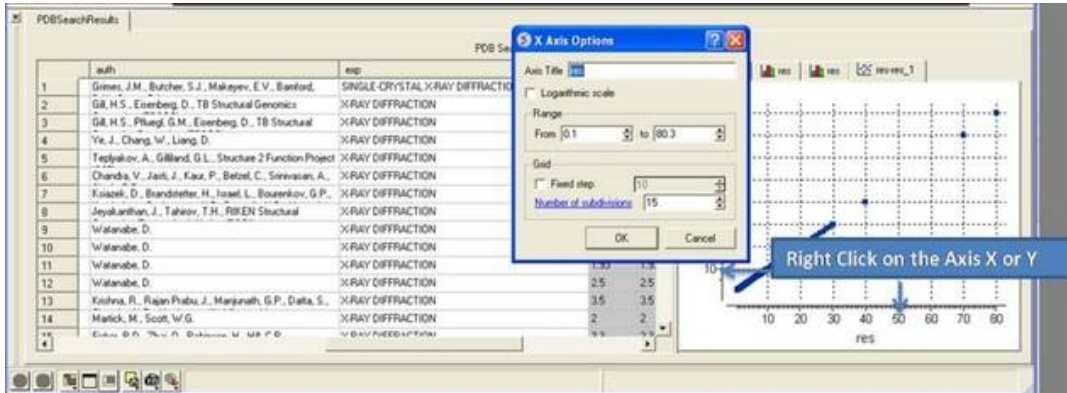
15.4.6 Add a title to a plot

To add a title to a plot:

- Right click on the plot and select **Edit Title** or choose **Options**

15.4.7 Axis Options

Each axis has a set of options which can be accessed by right clicking on the axis and selecting **Options**.



To change the title of the X or Y Axis:

- Right click on the axis and select **options**

To change the data range:

- Right click on the axis and select **options**
- Change the **From** and **to** values in the **Range** box

To change the Grid steps (ticks) on the X or Y axis:

- Select either a fixed step e.g. 10 and you can define the number of subdivisions (ticks) in each step. Choosing 1 will display zero ticks between divisions.

To change the axis to logarithmic

- Right click on the axis and select **options**
- Select the **Logarithmic scale** check box

15.4.8 Change Axis Data

To swap the X and Y axis:

- Right click on the plot and select **Options**.
- Select the **Swap X and Y** button.
- Click **OK**.

To change the data source for either the X or Y axis:

- Right click on the plot and select **Options**.
- Select the drop down arrow as shown below and select a different column from the table.
- Click **OK**.

15.4.9 Logarithmic Plots

To change the scale of the axis to logarithmic:

- Right click on the plot and select **Options**.
- Select the **Logarithmic** check box.

15.4.10 Change Mark Shape or Size

To change the plot mark, shape, style or label:

- Right click on the plot and select **Options**.
- Select the desired size and shape using the drop-down buttons in the **Marks** section of the window.

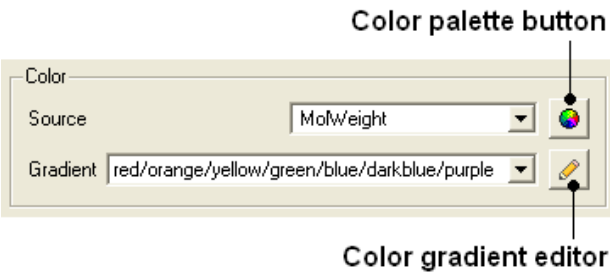
To add point labels:

- Right click on the plot and select **Options**.
- Select the drop down arrow in the **Point labels** dialog box
- If you only want to label selected points check the **Show labels for selection only** option. Making plot selections is described [here](#).

15.4.11 Change Mark Color

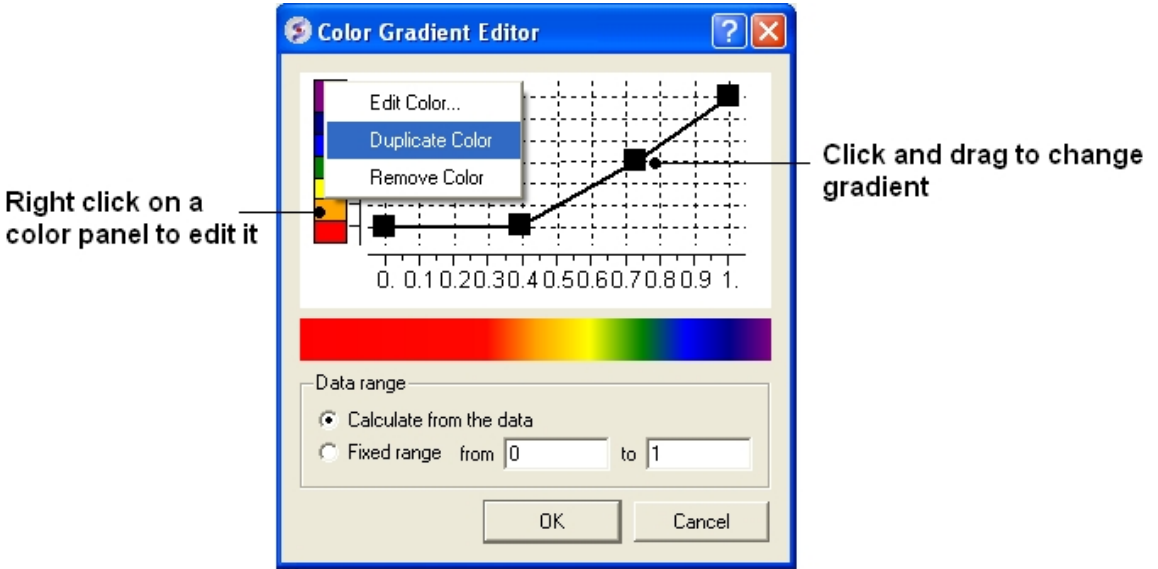
To change the color of the plot marks:

- Right click on the plot and select **Options**.
- In the **Color** section of the window select the **Source** (column name plotted as X or Y) you wish to color.
- Select the color palette and choose the desired color or you can choose a Gradient of colors.



To edit the color gradient

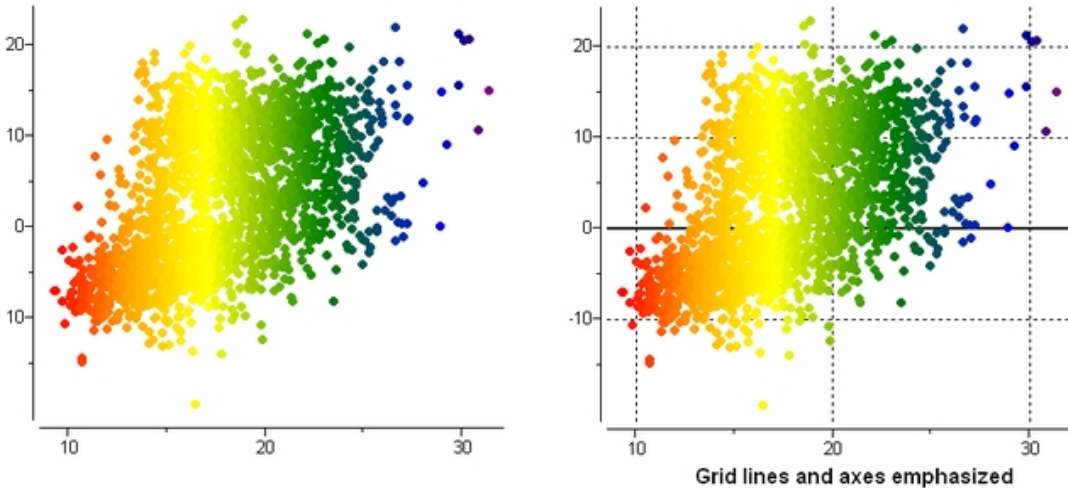
- Click on the **Color gradient editor** button and a window as shown below will be displayed.
- Click and drag on a mark in the gradient plot to change the color gradient.
- Right click on a color in the Y-axis to - **Edit, Duplicate or Remove Color**.
- The color gradient can be applied to all points in the data or for a fixed range.



15.4.12 Grid and Axis Display

To remove the grid display and/or highlight the axes:

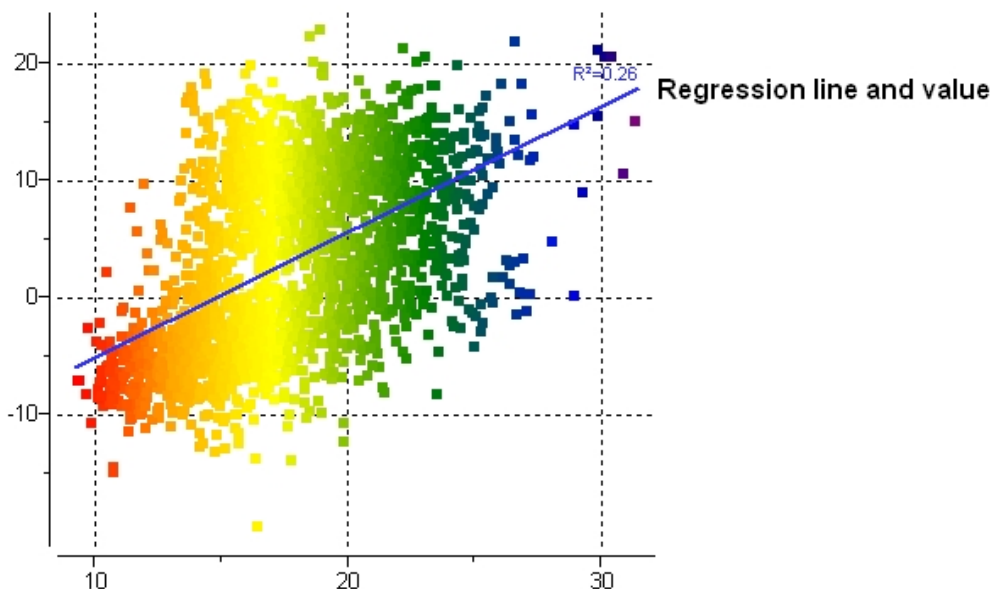
- Right click on the plot and select **Options**.
- Check the **Show grid** or **Emphasize axes** options.



15.4.13 Least Squares Fitting

To fit the data to a straight line using least square fitting

- Right click on the plot and select **Options**.
- Select the check box for **Least squares fitting line**.



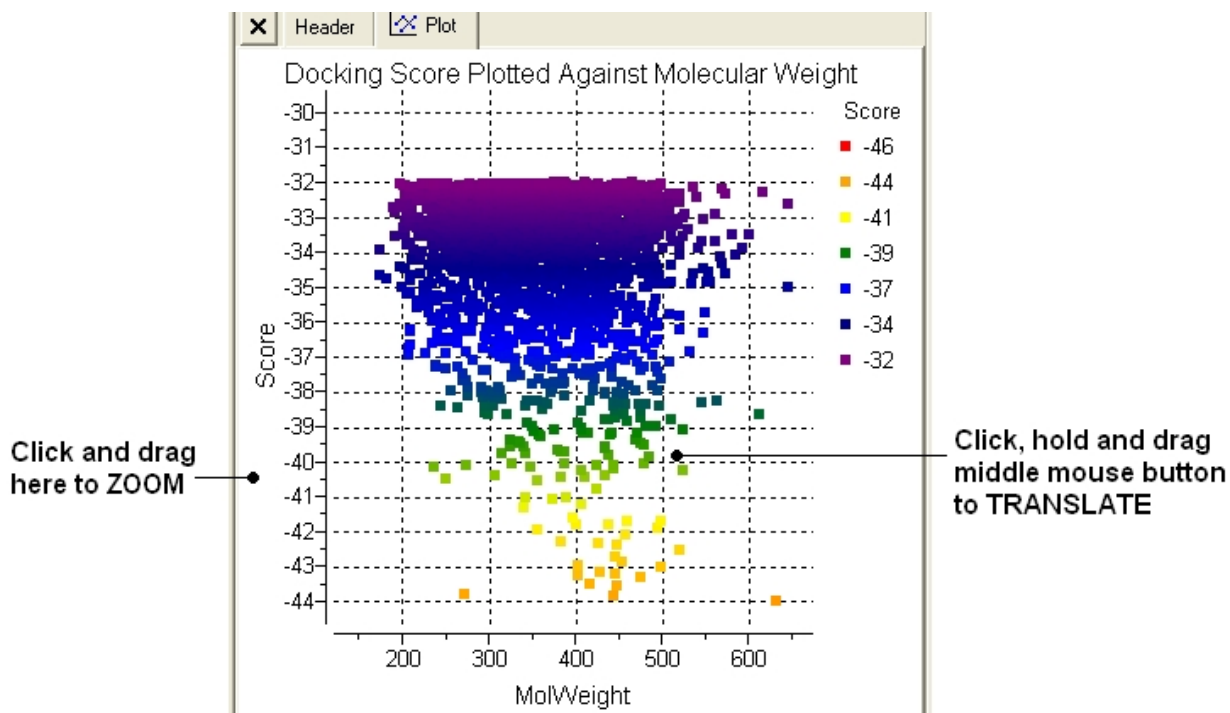
15.4.14 Zoom, Translate and Center

To zoom into a plot:

- Click outside the plot on the left-hand-side and drag the mouse or use the middle mouse wheel to zoom in and out.

To translate a plot

- Click, hold and drag using the middle mouse button on the plot.



To center onto a plot

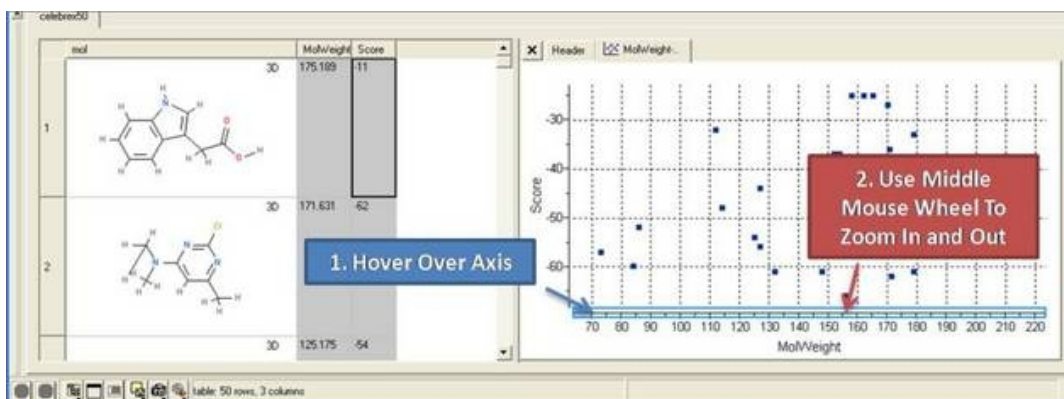
- Right click on the plot and select **Center all** or **Center Selection**. Making selections in a plot is described in the next section.

To center into an axis

- Right click on the axis and select **center**.

To zoom into an axis

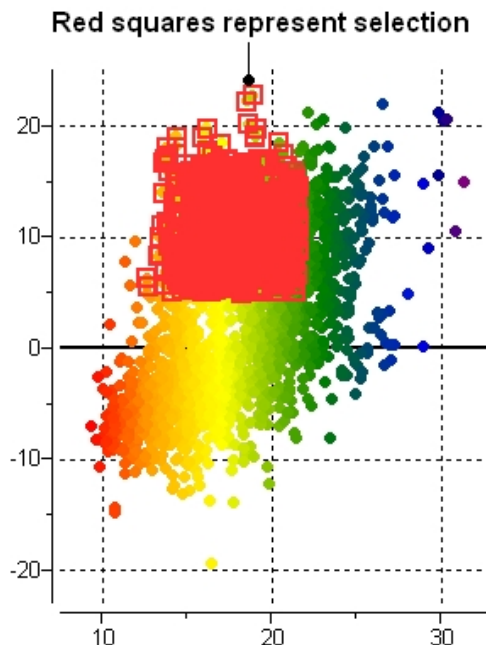
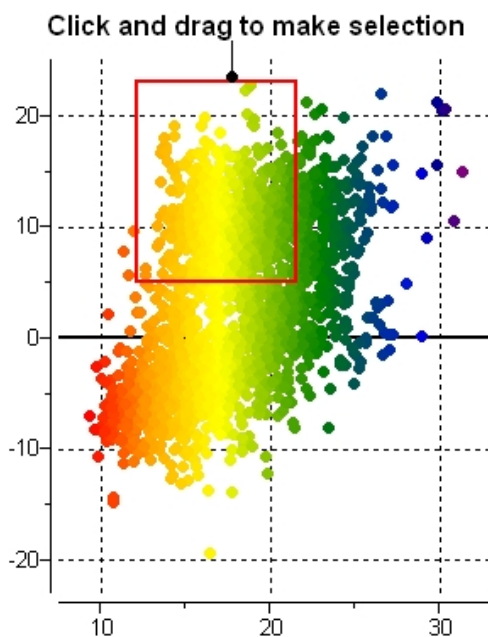
- Hover the mouse over the axis until you see a blue rectangle surrounding the axis.
- Use the middle mouse wheel to zoom in and out as shown below.



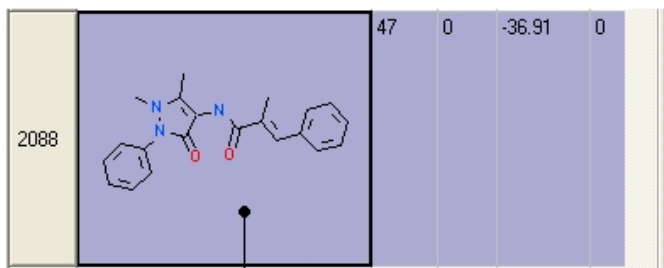
15.4.15 Plot Selection

To make a selection in a plot:

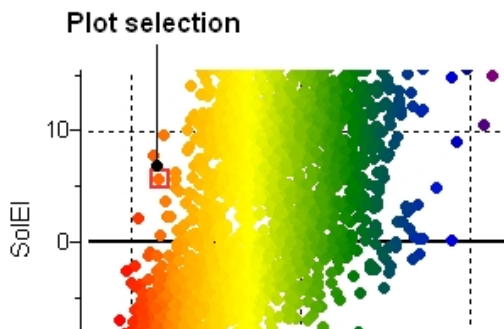
- Click and drag in the plot to make a selection. Individual points can be selected with a single click.



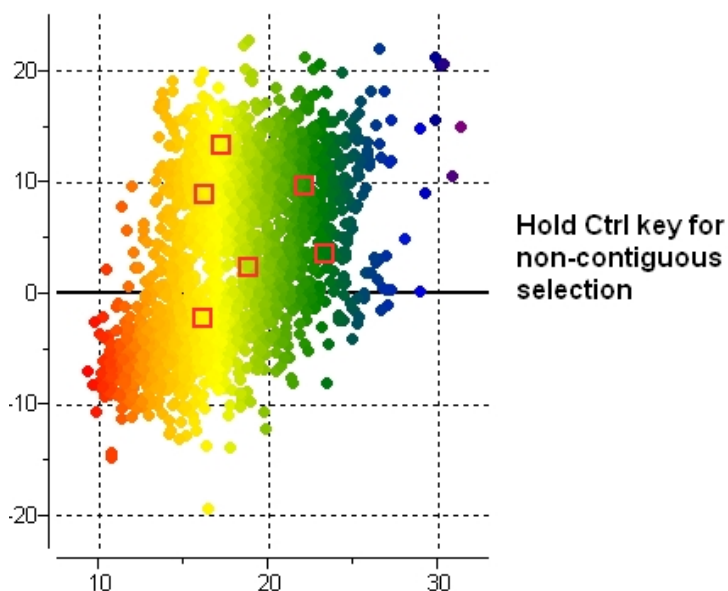
All selections are directly linked to the table from which the plot was made. Selections in the table are highlighted in blue.



Selection highlighted in blue in a table



Non-contiguous selections in the plot can be made by holding the CTRL key.



15.4.16 Print Plot

To print a plot:

- Right click on the plot and a menu will be displayed.
- Select the print option.

15.4.17 Saving a Plot Image

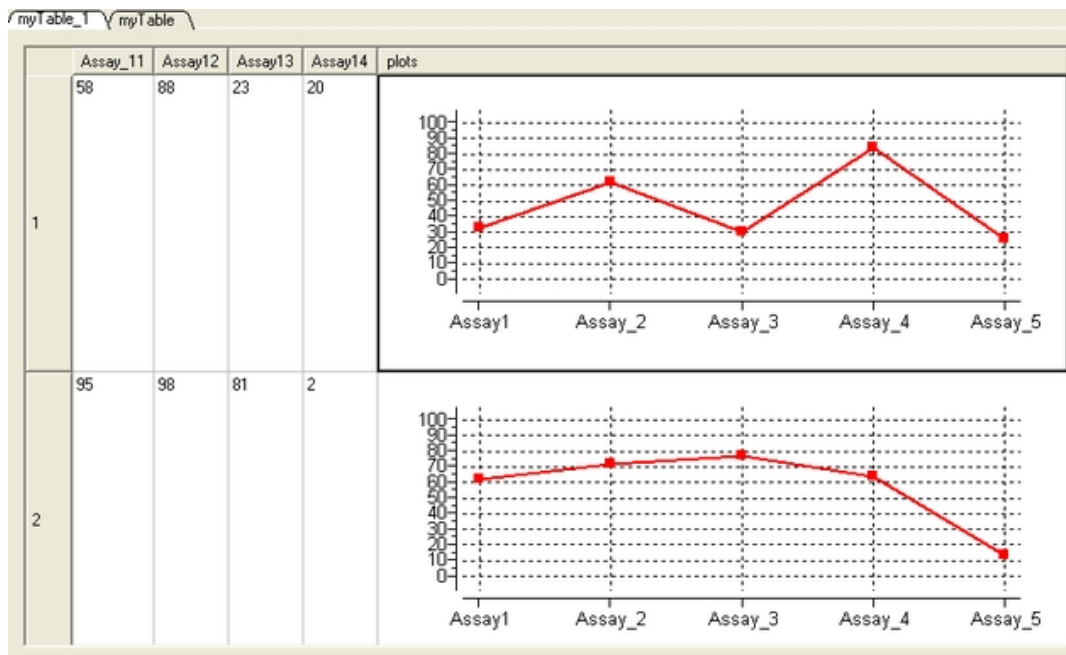
To save a plot image or copy to clipboard:

- Right click on the plot and a menu will be displayed.
- Select the **Save/Export Image** option.

15.4.18 Table Inline Plots

Plots can be inserted into a table row by:

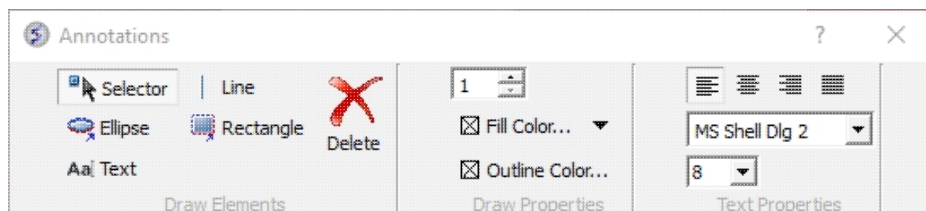
- Select the columns you wish to plot.
- Right click on the column header and select **Inline Plots**
- The plot will then be displayed in each row of the table.



15.4.19 Annotate Plot

You can add text, lines, rectangles, and ellipse to a plot. To annotate a plot:

- Right click on a plot and select Annotate.
- Select the draw element and then click on the plot.
- Change the draw element properties if needed (e.g. font size).



15.5 Principal Component Analysis

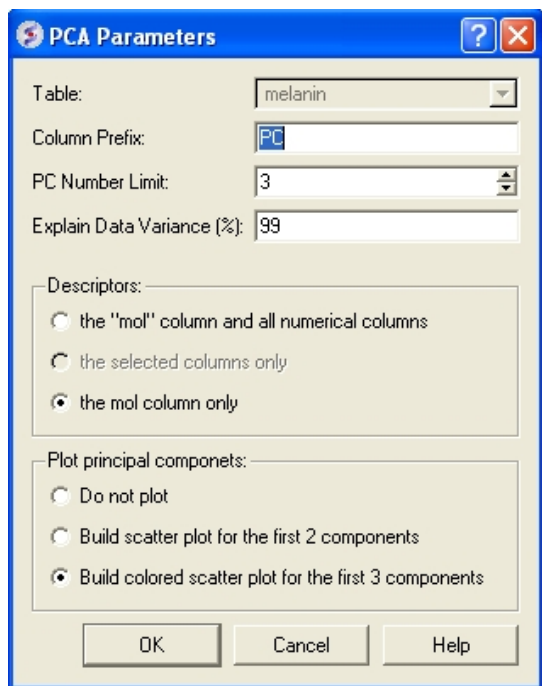
Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

Principal Component Analysis (PCA) is the younger brother of ICM's more powerful data analysis tools, like property prediction and clustering, though it still may give a good description of the data with a few columns or even chemical compounds. PCA is a mathematical procedure that transforms a number of correlated variables into a number of smaller uncorrelated variables known as Principal Components. The first component accounts for as much of the variability as possible with the rest of the components accounting for the remainder. PCA may be very helpful when you believe the data actually contains only a few meaningful components. Principal components are linear combinations of the provided data columns.

To perform a PCA analysis a table (either chemical or standard ICM table) needs to be loaded into ICM. For information regarding ICM Tables and ICM Chemical Tables please follow these links.

To begin the PCA procedure

- Right click on a ICM Tables and ICM Chemical Table and select the PCA option. It is important to right click inside the data table and not on a column or row header in order to see the correct menu on which pca is listed.
- Select which columns you wish to incorporate into the PCA analysis.



- Enter the table name on which you wish to perform the PCA analysis. If only one table is loaded this option will be greyed out.
- Enter the number of Principal Components (PC number limit) you wish to generate. Generally 3 principal components may be effectively visualized and it will be enough often to fulfil the data variance percentage requirement (see next option). The value displayed in the terminal window under the heading "cumulative explained data variance" will show what percentage of data relates to each PC.
- Enter a value in the "Explain Data Variance (%)" data entry box (99% is the default value) if you prefer this indirect way of limiting number of PC. The algorithm will stop when either PC number or explained variance limit is reached, so if you want only one of this criteria to work, make sure that the other limit is weak (by assigning accordingly the number of PC limit a high value, e.g. 50, or setting data variance to 100%).
- Select which descriptors you would like to include in the PCA analysis.
- Select which plot you would like to display. If you choose to display a plot use the color key on the side of the plot and the information contained within the ICM terminal window to relate which axes and points relates to which PC. PC3 is usually the color in the plot with the values displayed in the plot key.
- Click OK and if selected a plot will be displayed on the right-hand-side of the table. Points within a plot are linked to the table and can manipulated as other plots contained within a table.

15.6 Learn and Predict

Available in the following product(s): ICM-Chemist-Pro | ICM-VLS

Please click [here](#) to go to the Learn and Predict chapter.

15.7 Data Clustering

ICM allows you to create hierarchical clusters for chemical and other objects. Cluster trees can be used for:

- Navigation through large data sets.
- Selecting group representatives (taxons).
- Filtering tables to exclude redundancy.
- Finding similar elements, and more.
- Creating hierarchical views of data sets in many different styles, with subsequent image export/printing ability.

15.7.1 Cluster

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To perform clustering based on chemical similarity or any other data you must have an SDF file or table loaded into ICM.

- Right click on the table and select the Clustering option OR select the Chemistry menu and choose the option entitled **Cluster Set**.

OR

- Tools/Table/Clustering.

A data entry box as shown below will be displayed.

Clustering Parameters

Table: xdict

Name:

Take Labels from Column: NAME_

Keep the existing trees

Add column with cluster numbers, name: c

Build for the selected 0 rows (at least 2 required)

Descriptors:

the "mol" column and all numerical columns

the mol column only (2D pharmacophore)

the selected columns only

the seq column only

the "mol" column only (fingerprints)

use existing distance matrix: dimt_UPGMA

Method:

Tree

Memory required: 1 MB

Parameters:

Linkage type: UPGMA

Keep distance matrix

Compressed hierarchical clustering

OK Cancel Help

- Select the table you want to cluster from the drop down menu.
- Give the tree a name.
- Select the column you would like to use to label the tree (this can be changed or appended to once the tree has been made - right click on tree and choose "Change Record Labels").
- You can choose to keep the existing tree.
- The cluster number names will be added to your table - you can change the label of the columns (default = c).

Now select the descriptors you wish to use for clustering.

- The "mol" column and all numerical columns.
- The "mol" column only (2D pharmacophore)
- The selected columns only.
- The sequence column only.
- The "mol" column only (fingerprints).
- Use existing distance matrix. UPGMA is a linkage method which is used to group tree nodes together. (they are described in the link below). dimt_UPGMA is used when 'use exiting matrix' mode is selected. It can be any exiting matrix. This usually comes from previous clustering if

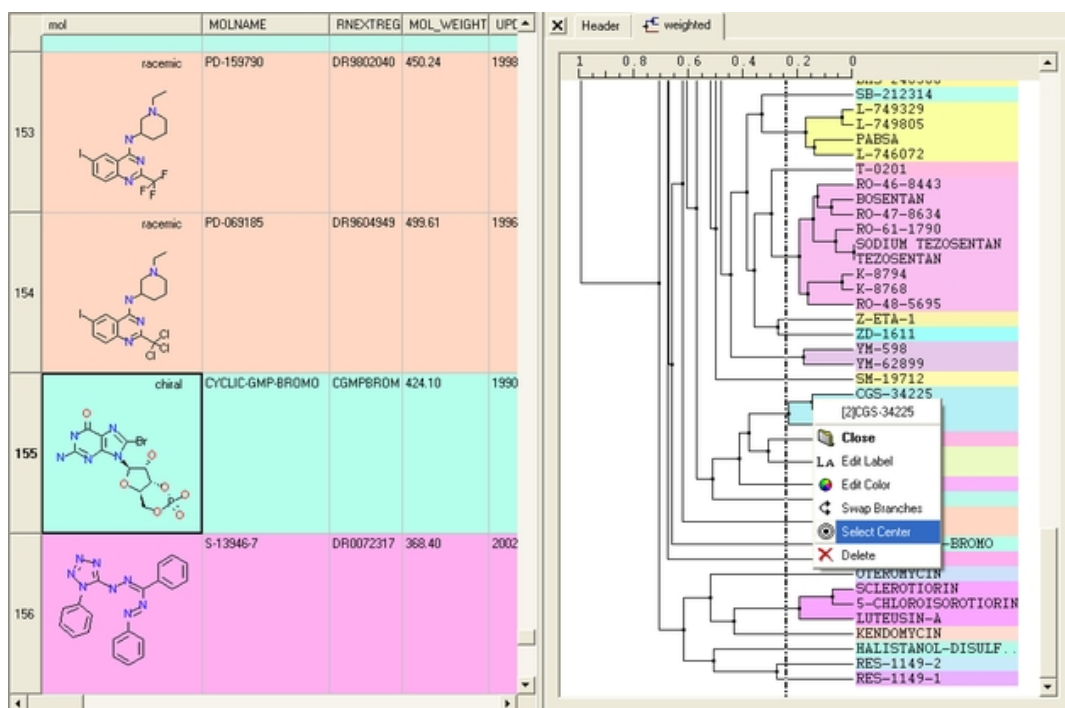
'Keep distance matrix' option was selected

Select the method you wish to use for clustering.

- **UPGMA** (unweighted pair group method using averages): Distance calculated is the average of all elements (recommended). <https://en.wikipedia.org/wiki/UPGMA>
- **Single linkage**: Nearest neighbour linkage - https://en.wikipedia.org/wiki/Single-linkage_clustering
- **Complete linkage**: Furthest neighbour linkage - https://en.wikipedia.org/wiki/Complete-linkage_clustering
- **WPGMA**: (weighted pair group method using averages): Rough approximation of weighted (slightly faster)
- **Farthest** https://en.wikipedia.org/wiki/Complete-linkage_clustering
- **Ward**: https://en.wikipedia.org/wiki/Ward%27s_method
- **UPGMC**: Centroid linkage clustering
- **UPGMC-corr**: Correlation clustering provides a method for clustering a set of objects into the optimum number of clusters without specifying that number in advance.

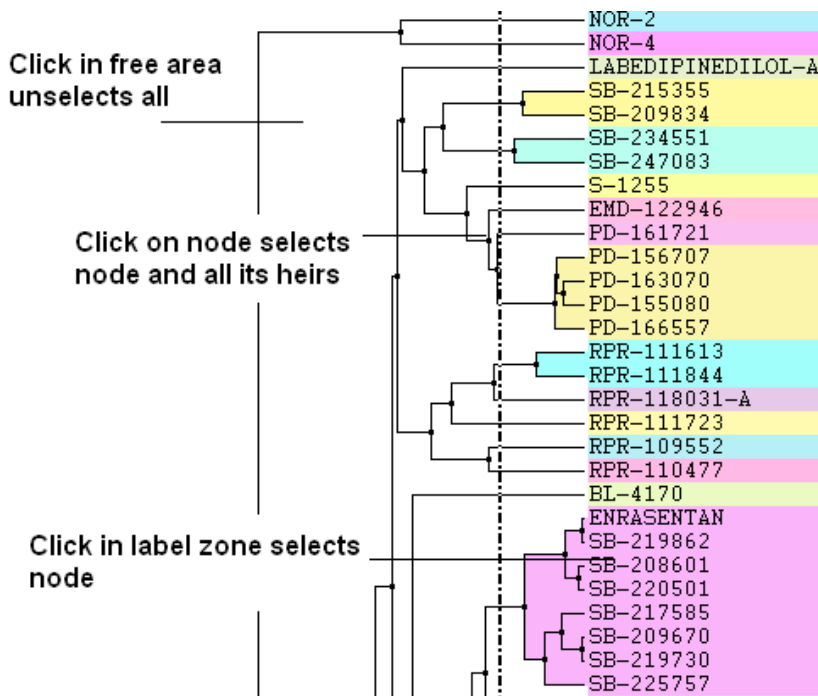
NOTE If you wish to cluster a selection of a table; first select the compounds you wish to cluster (Ctrl A will select all or see the section of the manual entitled making table selections).

Check the option "Keep Distance Matrix" if you wish the Distance Matrix to be saved.



15.7.2 Tree Selection

A number of different selections can be made once the tree has been created.



- Click in free area unselects all.
- Click on node selects node and all its heirs.
- Ctrl-click adds to (if not selected yet) or removes from (if already in selection) current selection.
- Shift plus click allows a range to be selected.
- Click in label zone selects node.
- Keyboard "Up"/"Down" cursor keys move selected node up/down in the tree. With Shift held the selection expands in the according direction.
- Keyboard "Escape" unselects all.

Select Center Representatives From a Node.

To select the "center" representative compound or value from a cluster.

- Right click in the free area of the tree and select the option "Select Centers".

NOTE: Selections in the tree will be highlighted in blue in the tree and in the table.

SB-234551
 SB-247083
 S-1255
 EMD-122946
 PD-161721
 PD-156707

Copy selection to new table.

All selections can be copied to a new table by:

- Right clicking on the selected rows in table and a menu will be displayed.
- Selecting the "copy selection to ICM table" option.

15.7.3 Save and Print Tree

To save a tree

Option 1:

- Save the whole session as an ICM project. See Saving an ICM project

Option 2:

- Save the table as an .icb file. Right click on table header and select "save as".

To save a tree as a picture

- Right click in the "free area".
- Select "Save Image..."

The image of the tree can also be saved to the clipboard.

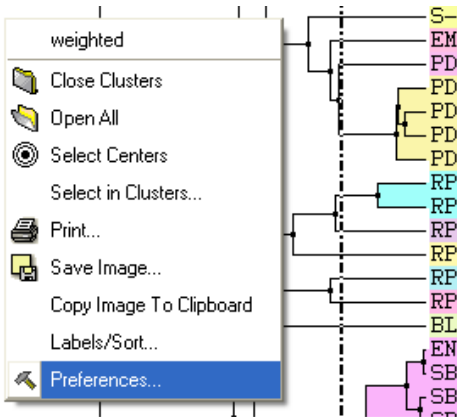
- Right click in the "free area".
- Select "Copy Image to Clipboard"

To print the tree

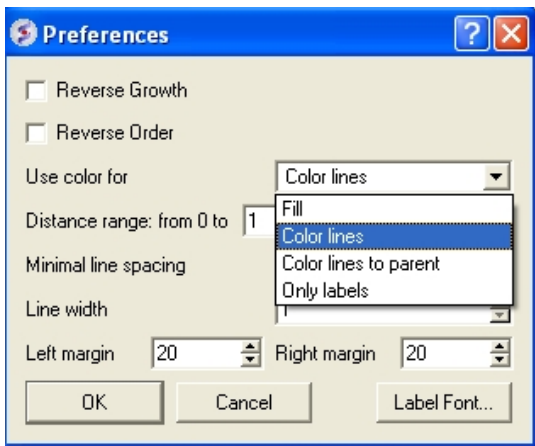
- Right click in the "free area".
- Select "Print"

15.7.4 Tree View

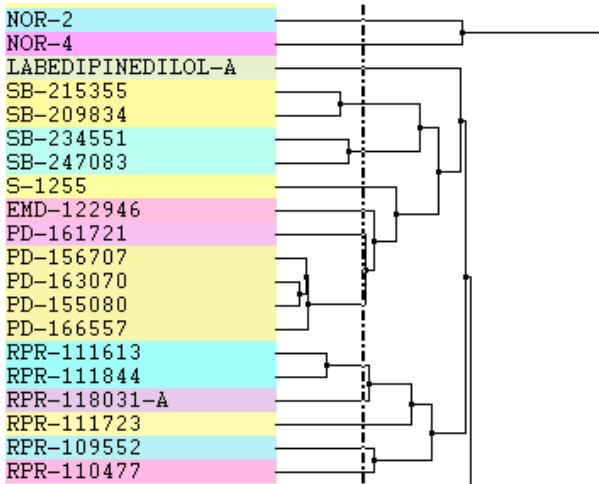
A number of changes to the tree view can be made by right clicking in the tree "free area" and selecting preferences as shown below.



A preferences window as shown below will be displayed.

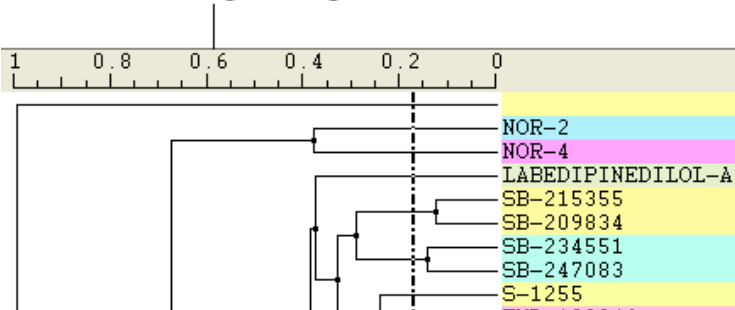


Reverse Growth:

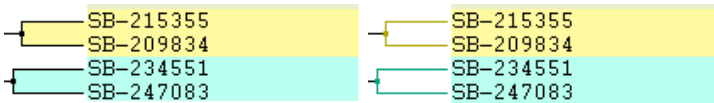


Distance Range:

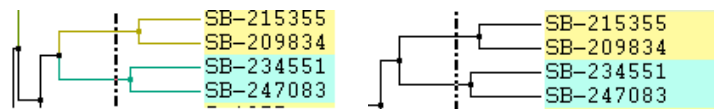
Distance range changes scale here



Color Preferences:

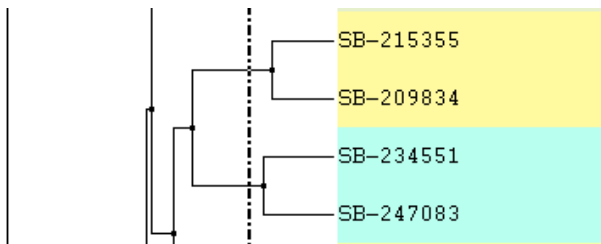


Left: Fill **Right:** Color Lines

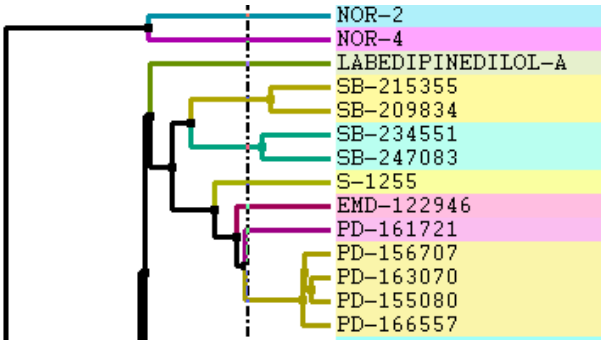


Left: Color lines to parent **Right:** Only labels

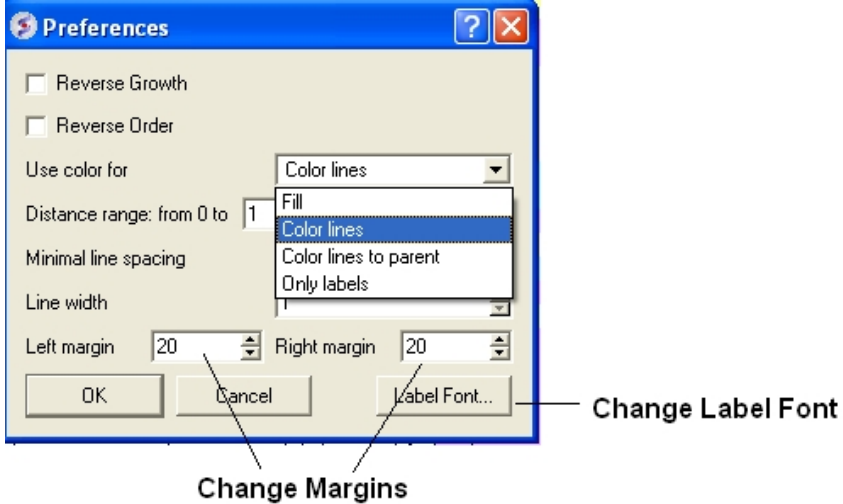
Increase or Decrease Line Spacing:



Increase or Decrease Line Width:



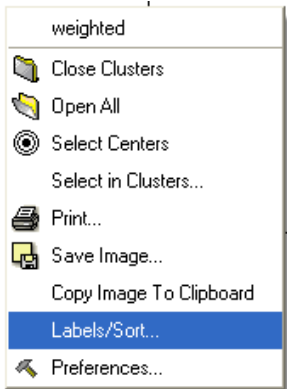
Change Font and Margins:



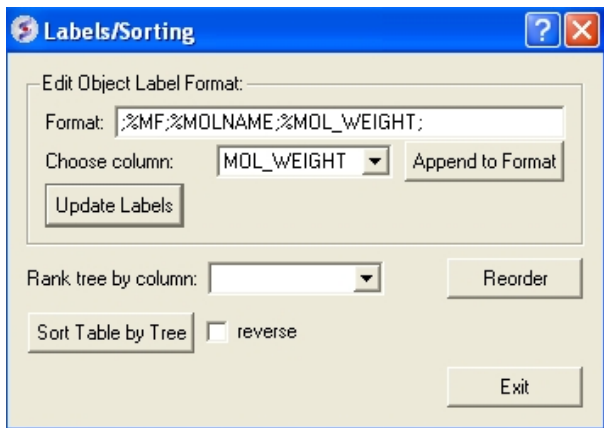
Change Label

To change the tree labeling (GLOBAL):

* Right click in the "free area" and select the Label/Sort option.



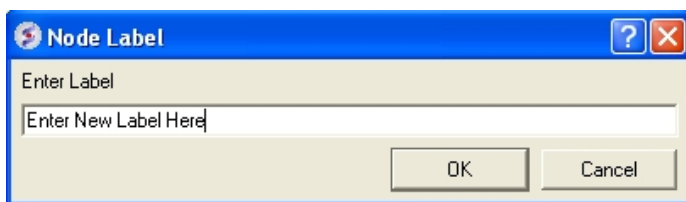
A window as shown below will be displayed:



- Choose the column name you wish to label the elements of your tree.
- Click the Append to Format button. You may wish to delete some of the information in the Format data entry box if you do not wish that column header to be displayed. Note more than one column header can be used as a label.
- Click the update labels button to view the changes.

To change the tree labeling (NODE):

- Right click on the node you wish to change the label of and select "Edit Label" and a data entry box as shown below will be displayed.



- Enter the new label.

Change Node Coloring

To change the color of a node:

- Right click on the node you wish to change the color of and select "Edit Color".
- Select the desired color and click OK.

16 Working with Local Databases

ICM tables are capable of storing tens of thousands records. However, some problems operate with data sets so large that they do not fit in computer's memory. To work with such large amounts of data ICM uses the concept of Molsoft database (MOLT) files. Unlike many other table file formats, such as SDF, CSV and others, database files are optimized for fast search and other operations, like unique entry addition and diverse subset selection.

Database files do not provide all the functionality available for tables but they allow the user to organize large amounts of data, search data using various advanced criteria and share created data collections with other users. For large amounts of chemical data database files provide specialized chemical functionality.

ICM provides the following database file operations:

- create database files from SDF, CSV/TSV, SMILES files and from ICM tables;
- impose unique constraints on certain columns upon table creation to avoid redundancy;
- store multiple tables in a single file; rename, delete tables in a database file;
- search fast using advanced conditions, including advanced chemical search;
- select diverse subsets from chemical database files;
- browse database tables using flexible filtering and sorting conditions;
- directly edit/delete/insert entries in the database;
- export in popular formats, such as SDF (for chemistry) and CSV;
- export and import tables to and from Molcart.

Chapter Contents:

- How to make a local database
- Browse Database
- Edit Database
- Query Local Database

16.1 How to make a local database.

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To convert a file into MOLT format.

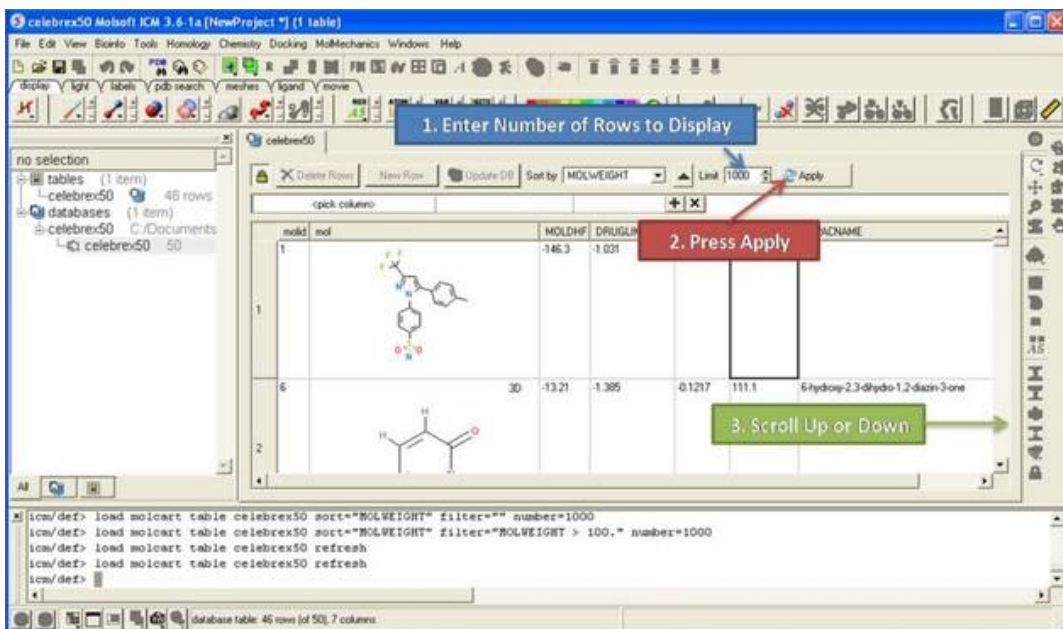
- File/Convert large SDF, CSV to Database (molt).
- A **Database Import** dialog box will be displayed. Select the fields you wish to import.
- Once the file has been converted to MOLT and imported then it will be displayed in the ICM Workspace.
- Double click on the file name in the ICM workspace and the database browse mode will activated.

16.2 Browse Database

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To browse a database first decide how many rows of your database you wish to view

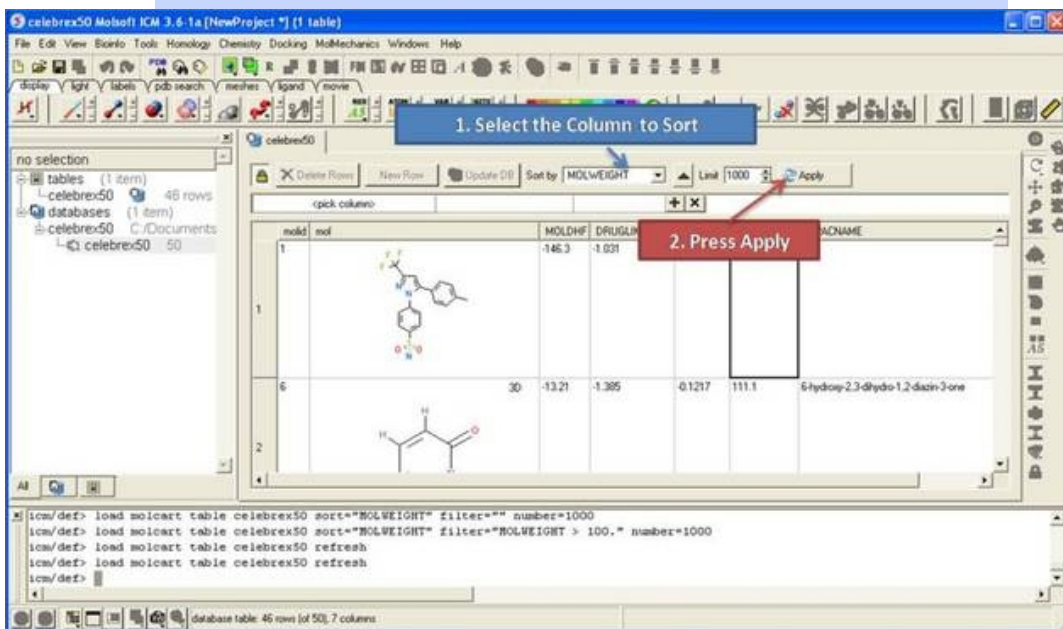
- Double click on the file name in the ICM workspace and the database browse mode will activated.
- Select the number of rows you wish to be displayed in the **Limit** data entry box.
- Click on the **Apply** button.
- Scroll up and down the displayed rows using the scroll bar on the right hand side of the table.



To sort a database by a value in a column

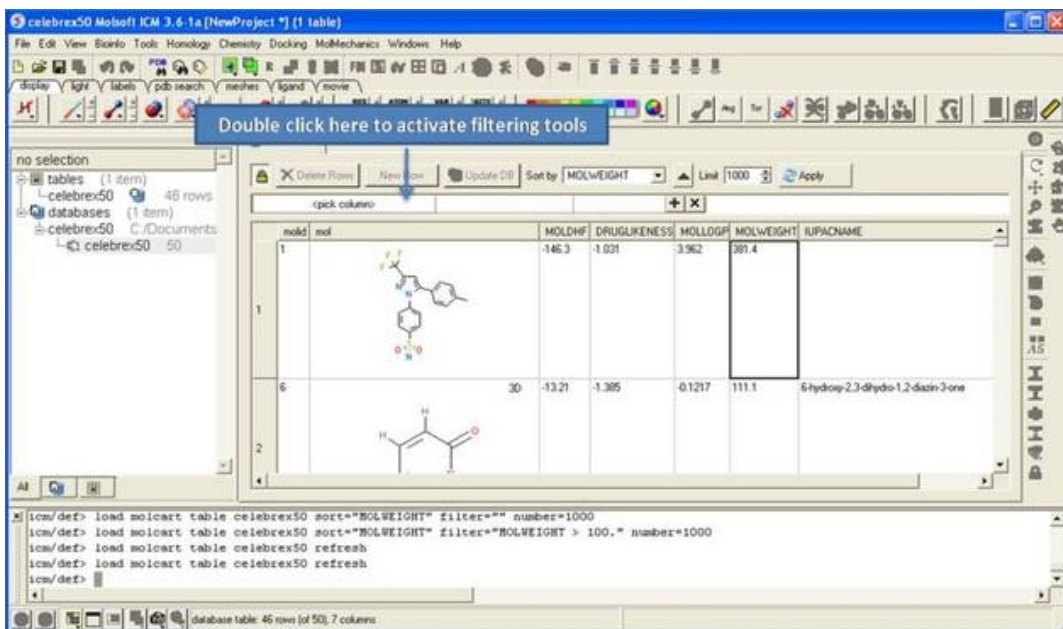
- Click on the drop down arrow next to the **Sort by** data entry box and select a column.
- Click on the **Apply** button.

NOTE: The database will be sorted globally not just the displayed rows.

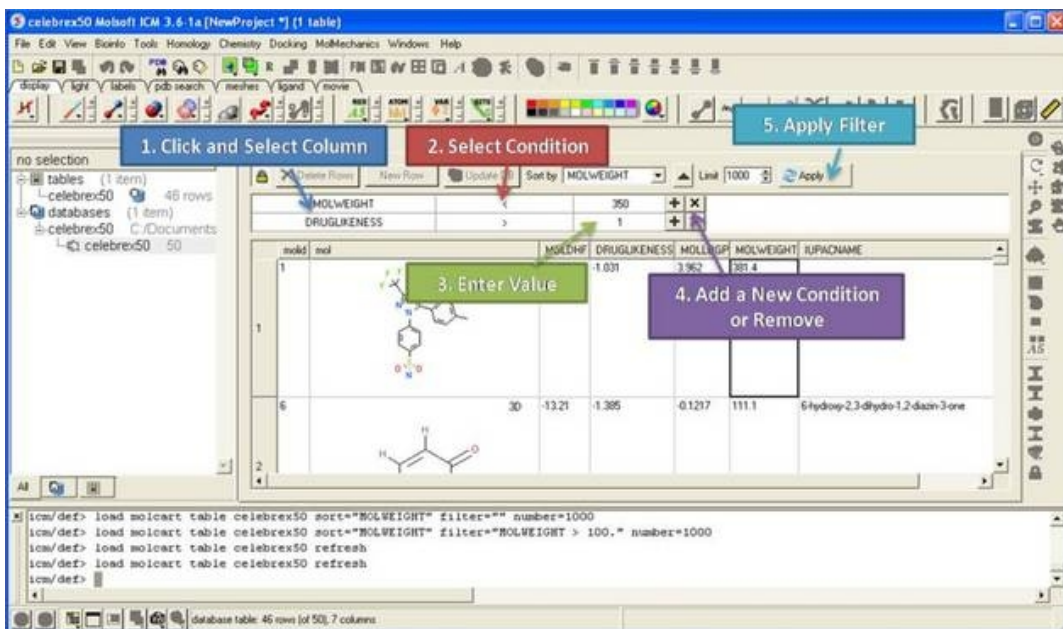


To filter a database:

- Double click on the panel labeled and the filtering tools will be activated.



- Click and select a column to filter (as shown below).
- Enter a condition.
- Enter a value.
- Add or remove another condition
- Click the **Apply** button



To view a diverse set of rows from your database:

- Right click on the database in the ICM Workspace and select **Select Diverse Set**.

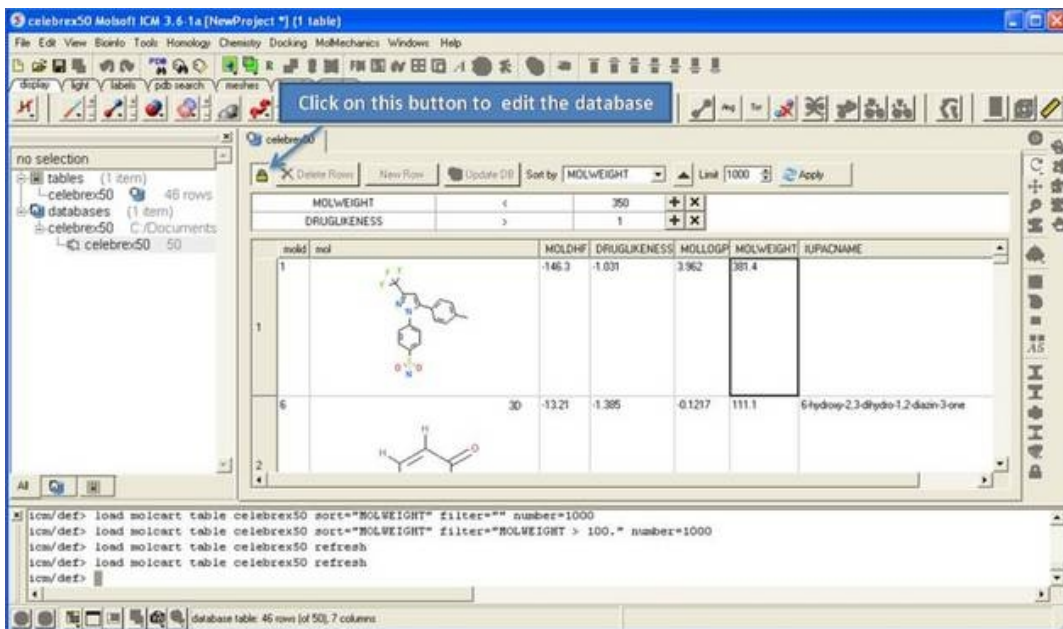
16.3 Edit Database

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To make any changes to a database you must first unlock it:

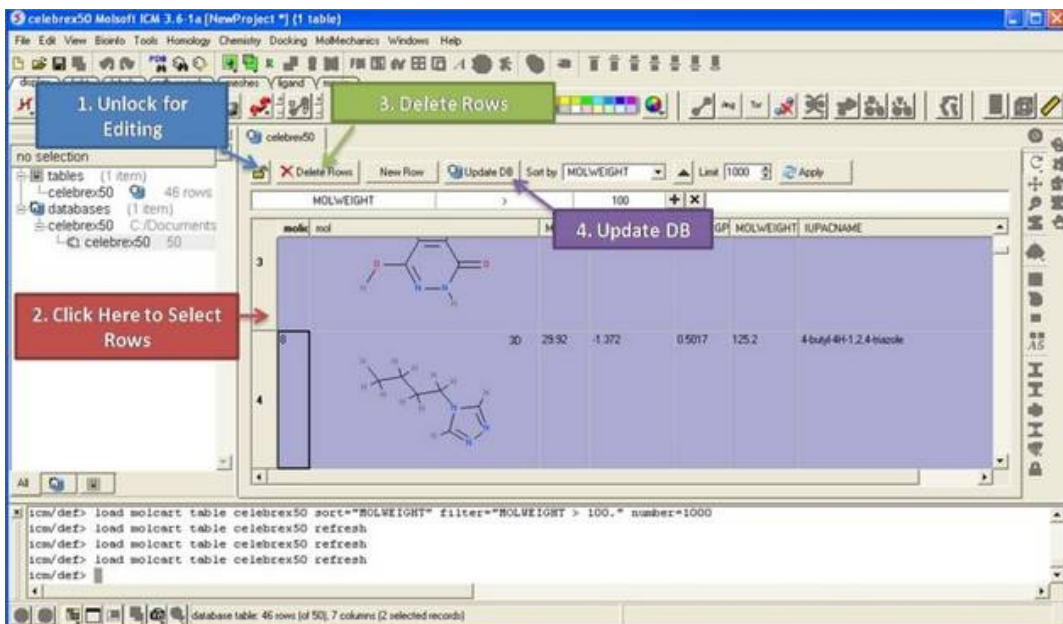
- Click on the **Allow editing rows** button which has a picture of a lock on it.

- Once this button has been selected the **Delete Rows**, **New Row**, and **Update DB** buttons will become activate.



To delete rows

- Select the row(s) by clicking on the row numbers. A range of rows can be selected by holding the shift key and clicking a non-contiguous set of rows can be selected by holding down the control key.
- Click the **Delete Rows** button.



To insert a new row

- Select a row.
- Click on the **New Row** button. A new row will be inserted underneath the selected row.
- Click **Update DB**

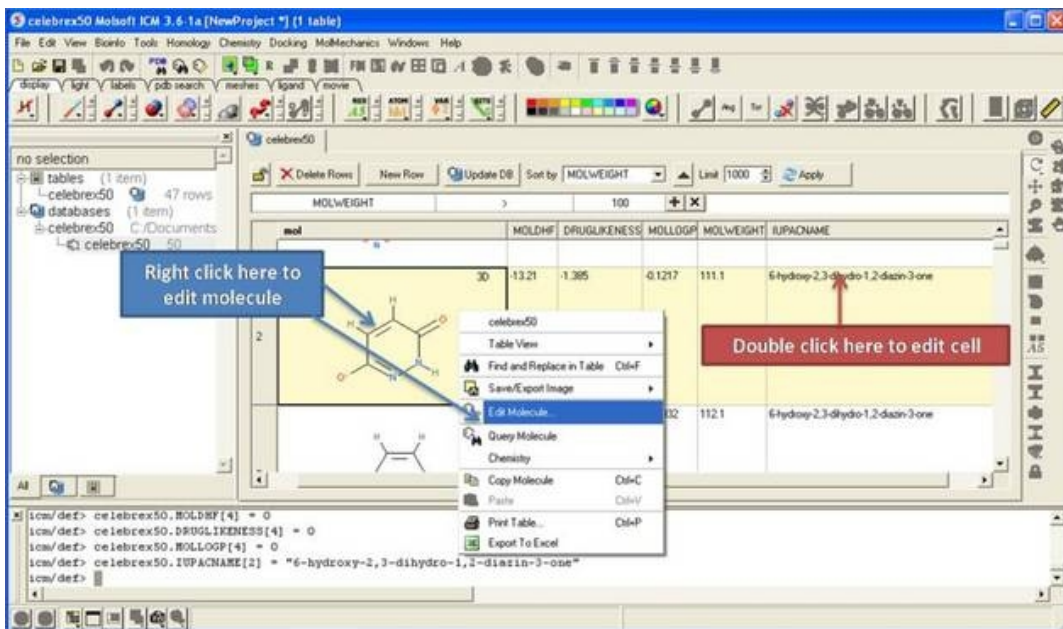
To edit strings and numbers in a row

- Double click on the cell you wish to edit and then enter a new value.

- Click **Update DB**

To edit a chemical (2D sketch):

- Right click on the cell and select **Edit Molecule**
- Make changes using the Molecular Editor
- Click **Update DB**

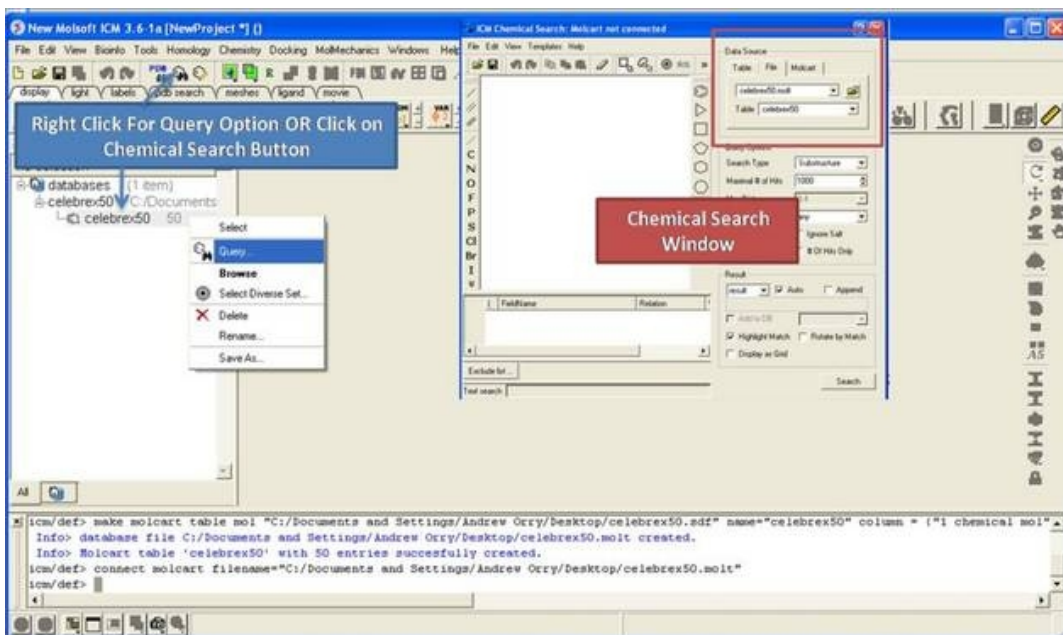


16.4 Query Local Database

Available in the following product(s): ICM-Chemist | ICM-Chemist-Pro | ICM-VLS

To query a local database:

- Right click on the database in the ICM Workspace and select **Query**.
- Query using the chemical search tools.



17 ICM-Scarab - Data Capture, Integrator and Miner

Developed by MolSoft in Collaboration with Scientists at the Structural Genomics



Chapter Contents:

- Scarab Installation
- Upload Data
- Query Metadata
- Browse and Export Data
- Add a New User
- Create a New Project and Assign User
- Pages - Electronic Lab Notebook

A Flexible LIMS

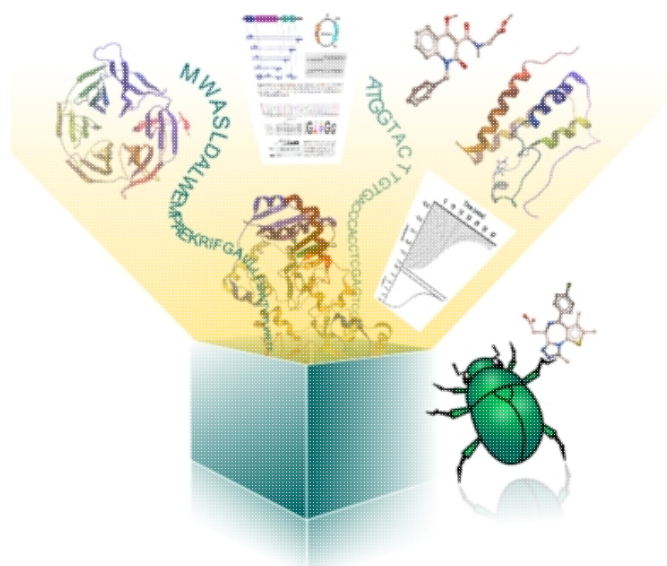
- Capture structured and unstructured data at the same time in the same place without losing data mining flexibility. ICM-Scarab contains an Electronic Lab Notebook (ELN) component.

Unique Flexible Data Capture and Mining Interface

- No need to have reports built for you by your administrator - build them yourself
- Present databases to users in an intuitive fashion without needing to modify the underlying database(s) structure
- Intelligently interfaces with any Oracle & MySQL databases
- Seamlessly build and run complex queries across database architectures
- Built-in compound searching capabilities
- Fine-grained editing of data across tables
- Bulk uploading of data

A One-Stop-Shop

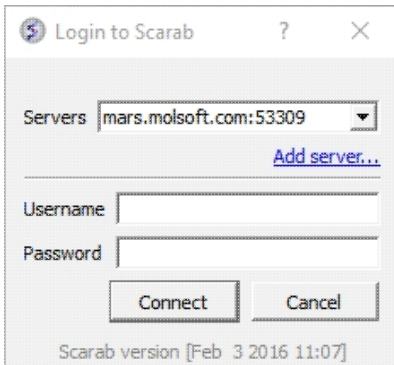
- Integration with MolSoft's standard functions for structural biology, bioinformatics and cheminformatics



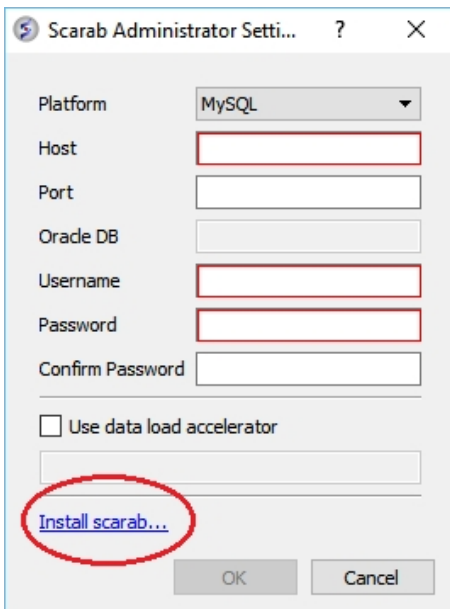
17.1 Scarab Installation

How to install Scarab:

- Install MolCart on your Linux server - download from www.molsoft.com/support
- Install mysql server version (v 5.5.3) on the server.
- Install and unpack Scarab on your Windows machine (contact MolSoft to get download link).
- Click on **Add Server**



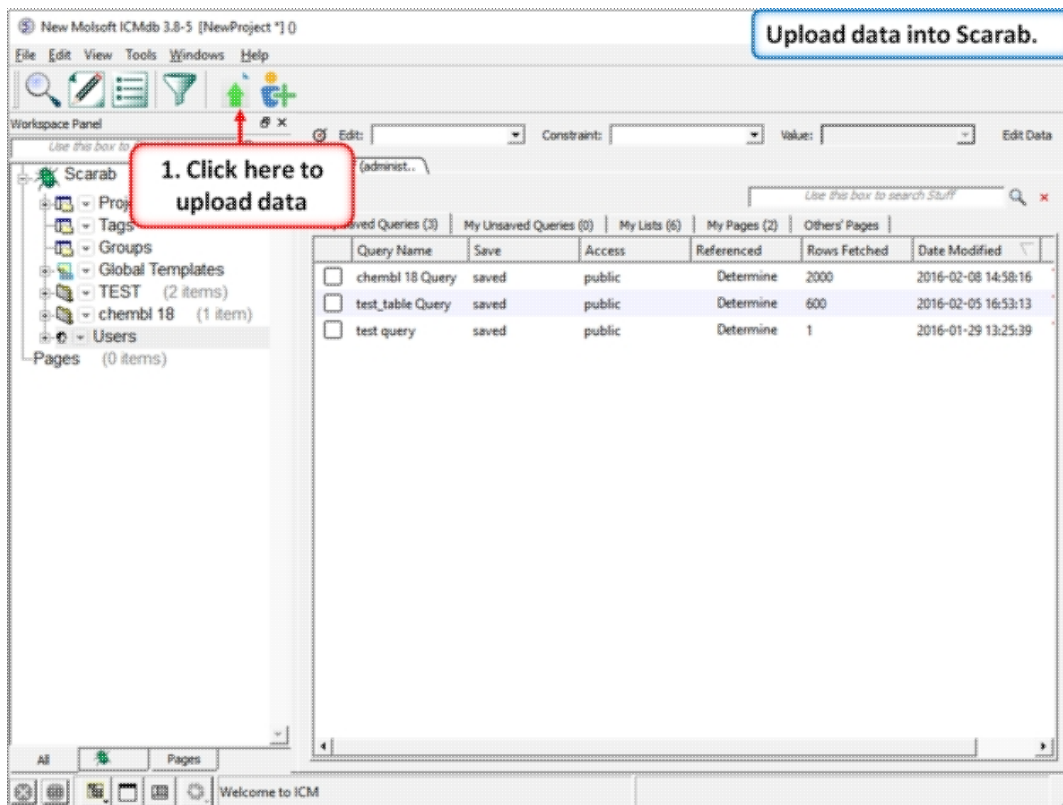
- After the package is installed , the user should install the Scarab metadata into a mysql server (where Molcart is already installed) from here (this opens from Add server... on the login to Scarab window). Click on Install scarab...



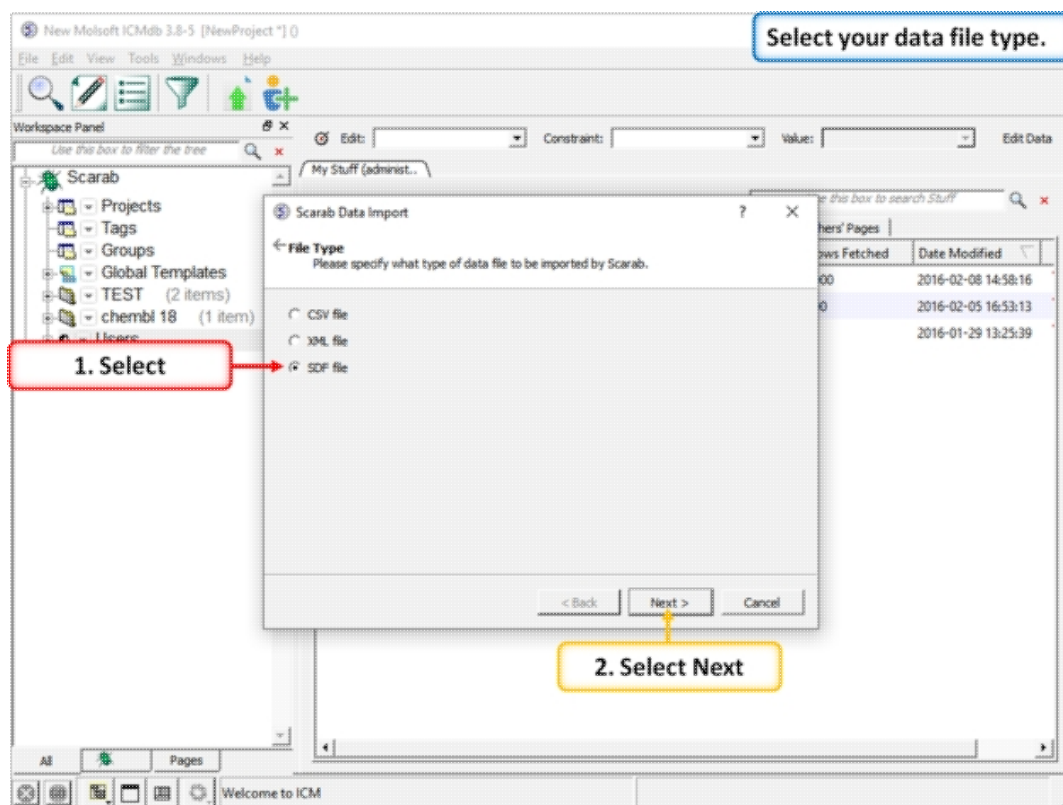
- Enter name of server.
- Leave port empty (default 3306)
- Enter mysql root name and password
- Import data.

17.2 Upload Data

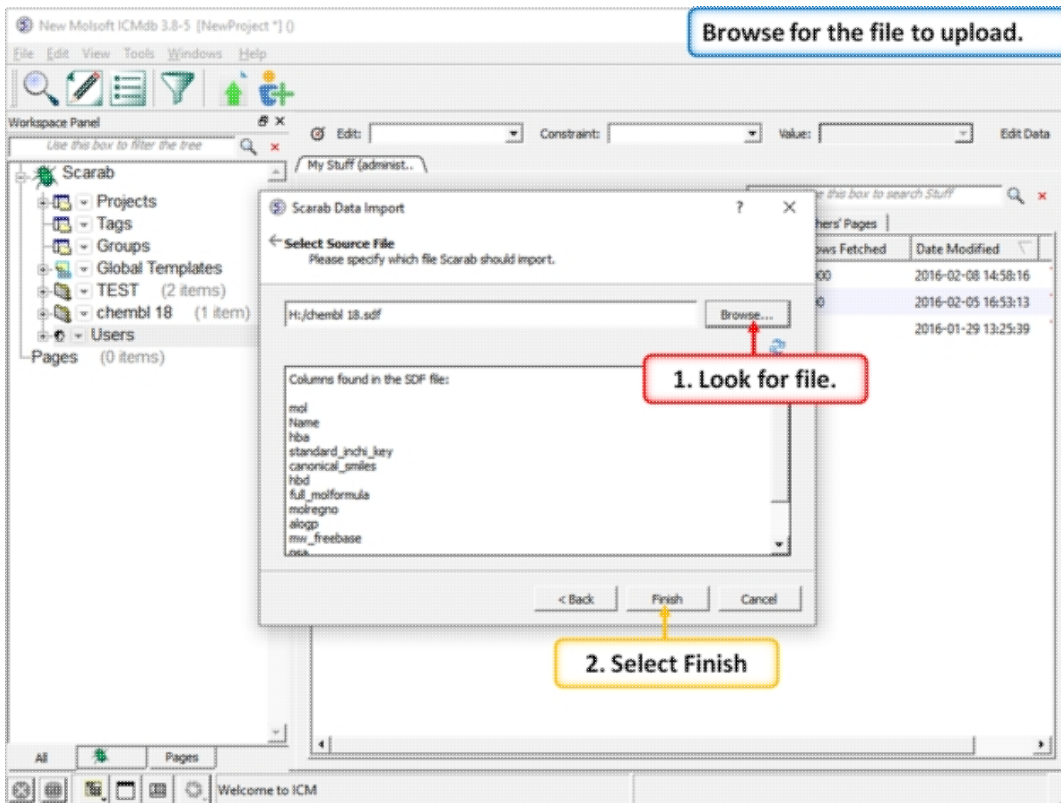
You can upload sdf, csv or xml files into Scarab.



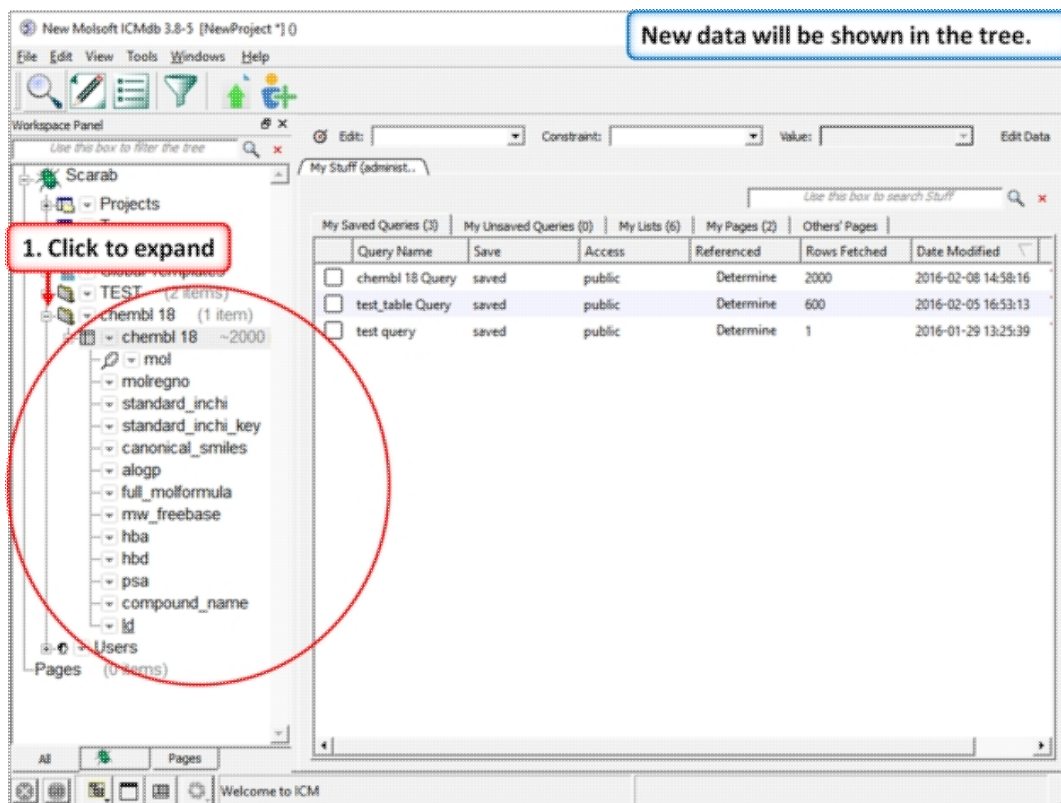
Click on the "Import Existing File" button.



Select the file type sdf, csv or xml.



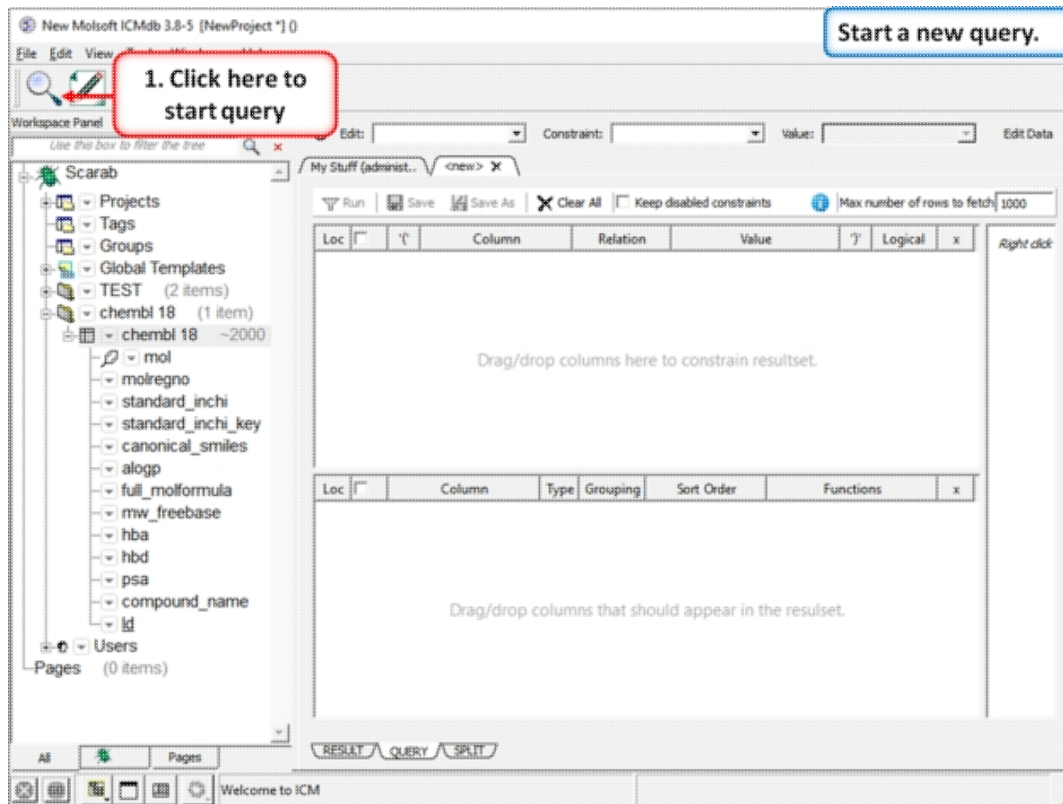
Browse for the file on your file system.



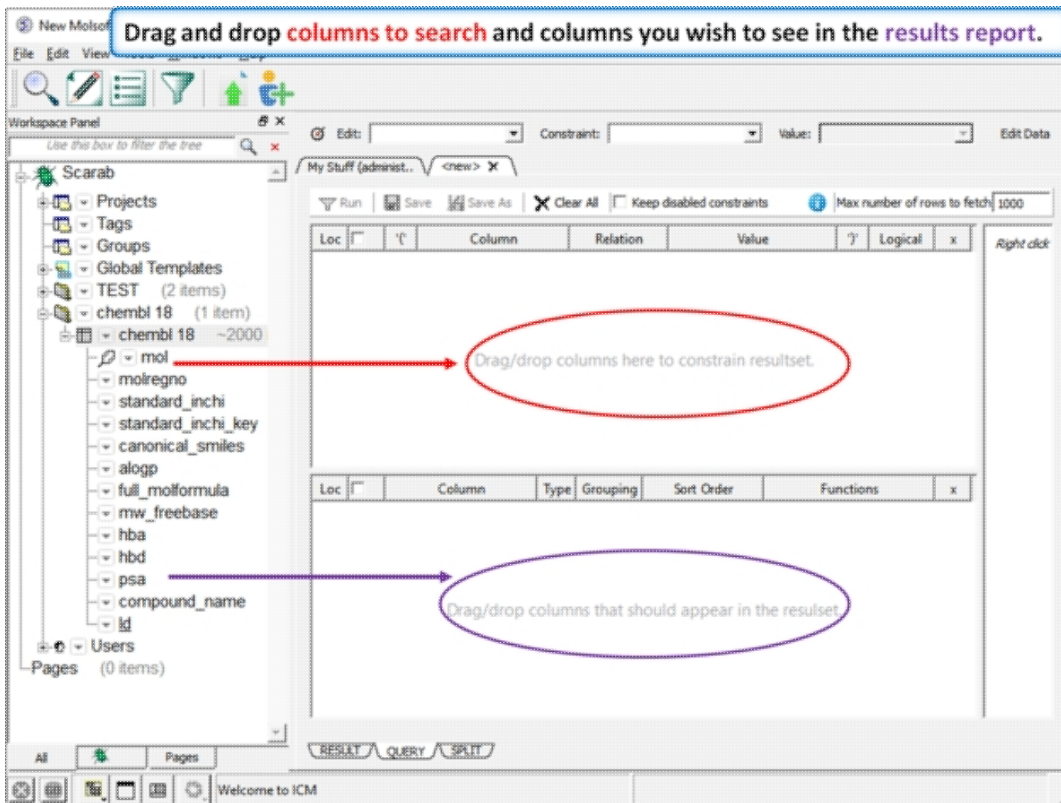
When the file is fully uploaded you will see the metadata in the tree. The tree can be expanded to view the metadata hierarchy.

17.3 Query Metadata

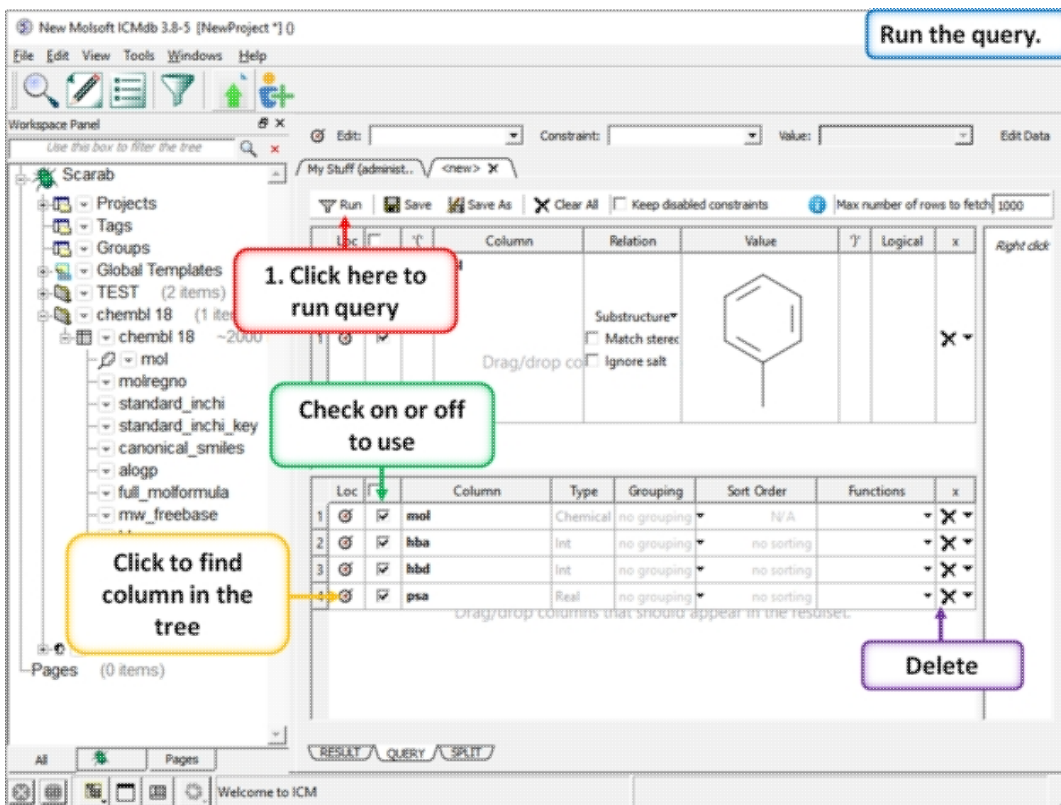
Scarab provides a convenient window to setup your query and report using drag and drop directly from the metadata tree. You can also save and share queries with other users.



Click the "Start a Query" button and the query



Drag and drop columns to search into the top panel and columns you wish to see in the results report into the bottom panel.



Click the "Run Query" button to start the search.

Query results.

Save query for future use

Loc	Column	Relation	Value	Logical	x
1	mol	Substructure*	<chem>C1=CC=CC=C1</chem>		X

Loc	Column	Type	Grouping	Sort Order	Functions	x
1	mol	Chemical	no grouping	N/A		X
2	hba	Int	no grouping	no sorting		X
3	hbd	Int	no grouping	no sorting		X
4	psa	Real	no grouping	no sorting		X

mol	hba	hbd	psa
	6	2	77.38

Change result view

When the query has finished the results will be shown in the bottom panel. You can change the view using the tabs and save the query.

View queries.

Click here to view saved and unsaved queries

My Saved Queries (3) My Unsaved Queries (1)

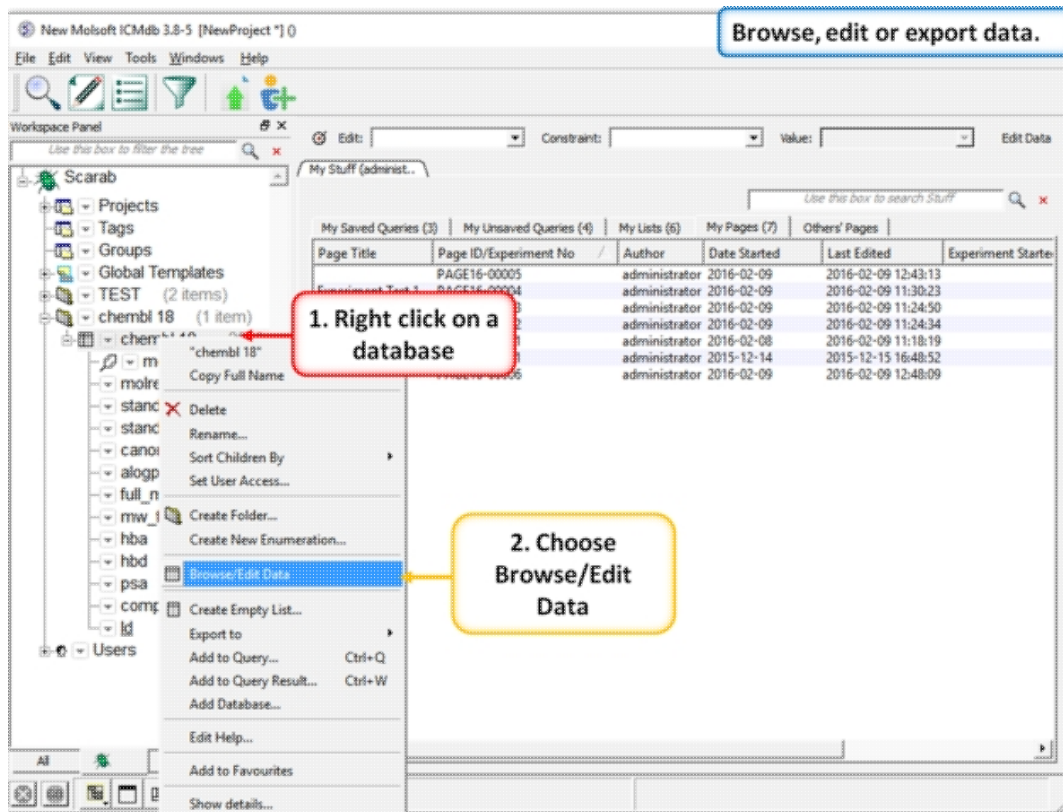
Query Name	Saved	Access	Referenced	Rows Fetched	Date Modified
chembl 18 Query	saved	public	Determine	2000	2016-02-08 14:58:16
test_table Query	saved	public	Determine	600	2016-02-05 16:53:13
test		public	Determine	1	2016-01-29 13:25:39

Double click to run query again

Your queries will be saved in the "My Stuff" tab.

17.4 Browse and Export Data

You can browse and export the metadata by right clicking on a database in the tree which then opens up the tree.



Right click on the metadata and choose the option "Browse/Edit Data"

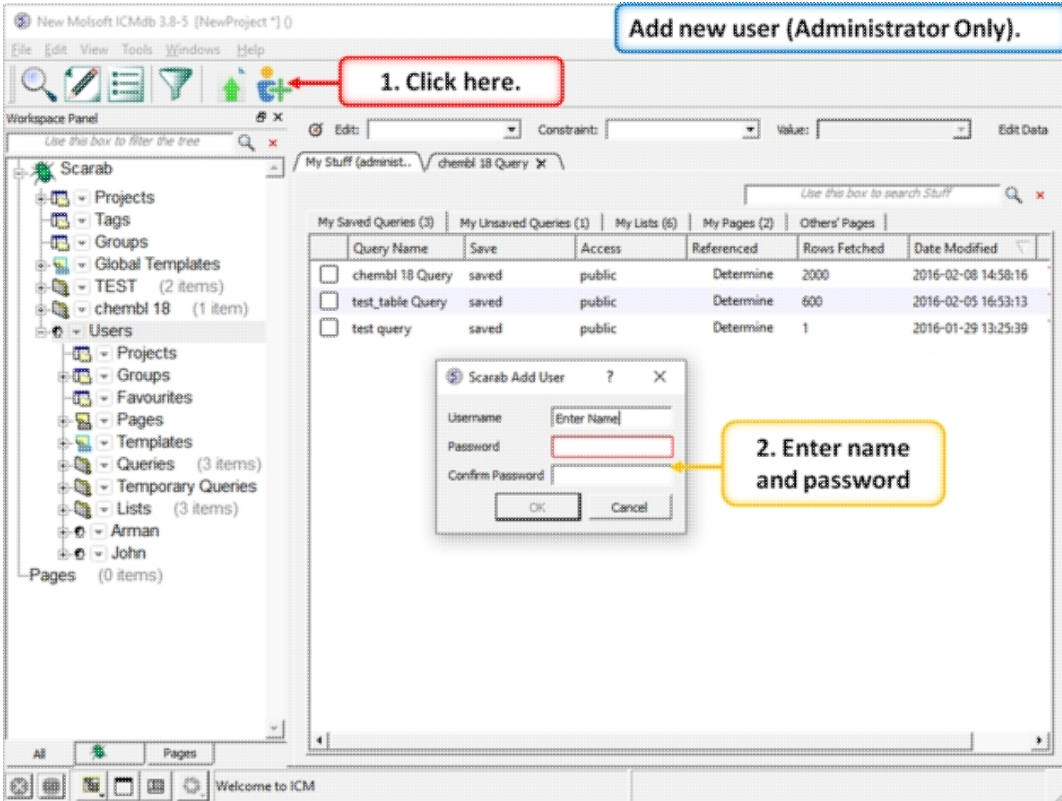
The screenshot shows the Molsoft ICMdb 3.8-5 interface. The main window displays a table with columns for 'mol', 'molregno', and 'standard_inchi'. The table contains data for years 1998, 1999, 2000, and 2001. The 2001 row is highlighted in pink, indicating it is selected for editing. A yellow callout box labeled '2. Double click cell to edit' points to the 'mol' cell of the 2001 row. A red callout box labeled '1. Click here to edit' points to the 'Save Format' button in the toolbar. A green callout box labeled '3. Save changes' points to the 'mol' column header. A purple callout box labeled '4. Export or print' points to the 'More Actions' button in the toolbar. A blue callout box labeled 'Browse, edit or export.' is located in the top right corner. The left sidebar shows a tree view of the project structure, including 'Projects', 'Tags', 'Groups', 'Global Templates', 'TEST', and 'chembl 18'.

Year	mol	molregno	standard_inchi
1998		625038	InChI=1S/C27H22BrN3O2/c28-21-12-6-4-11-19(21)26-25-20... WJSP
1999		625037	InChI=1S/C27H22BrN3O2/c28-21-12-6-4-11-19(21)26-25-20... WJSP
2000		625007	InChI=1S/C27H22BrN3O2/c28-21-12-6-4-11-19(21)26-25-20... WJSP
2001	<Click to edit>	<Text to add>	<Text to add>

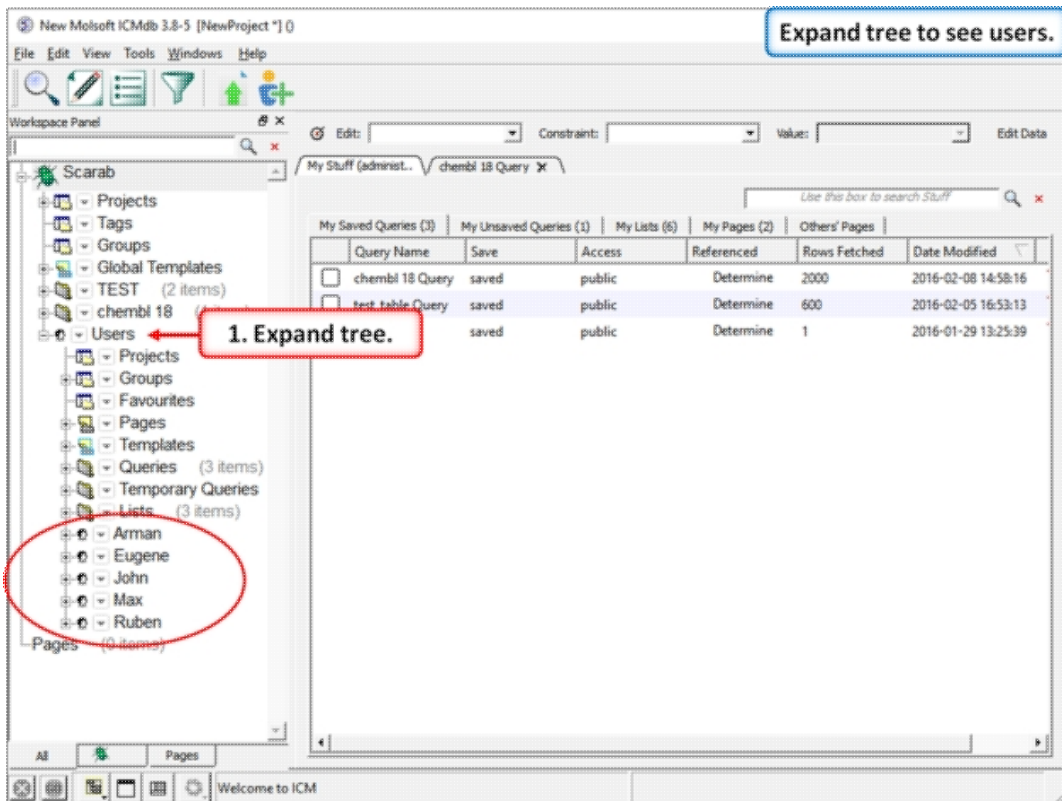
Click on the "Enable Data Editing" button and then double click to edit a cell. Pink coloring in the row indicates that the cell has to be filled and grey indicates the cell will be automatically populated. Click on the "Save Format" button when you are finished.

17.5 Add a New User

New users can be added in Administrator mode.



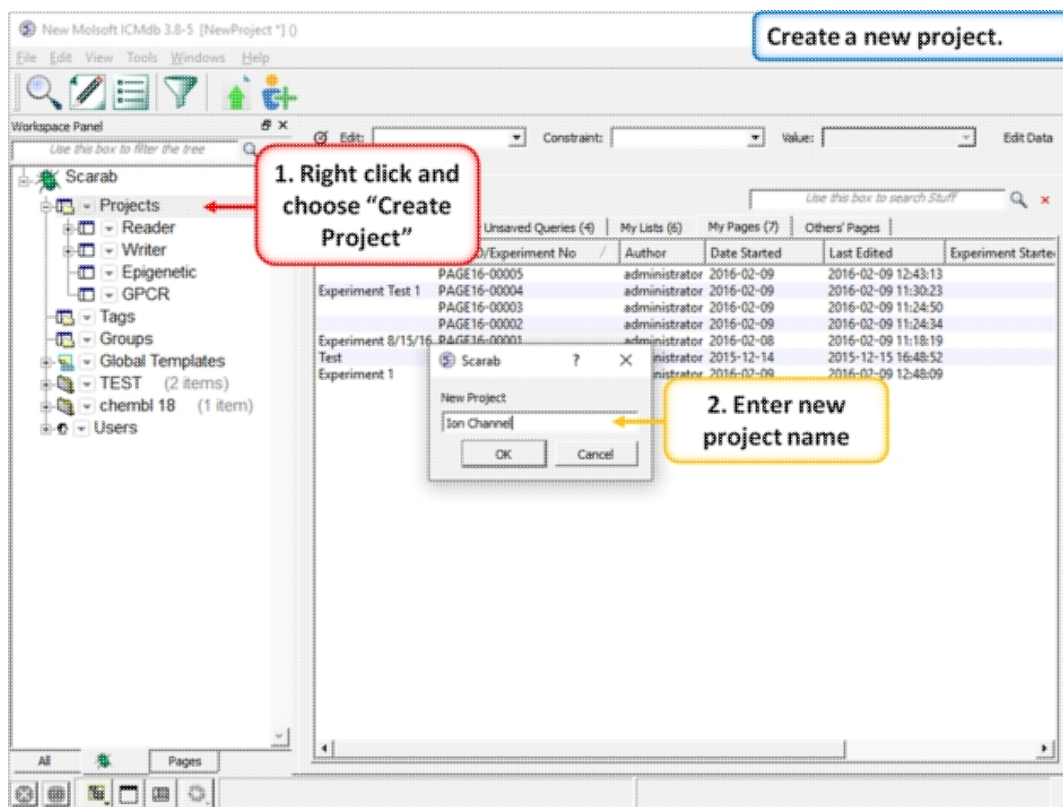
To add a new user click on the "Add New User" button and a dialog window will be displayed where you can add the username and password.



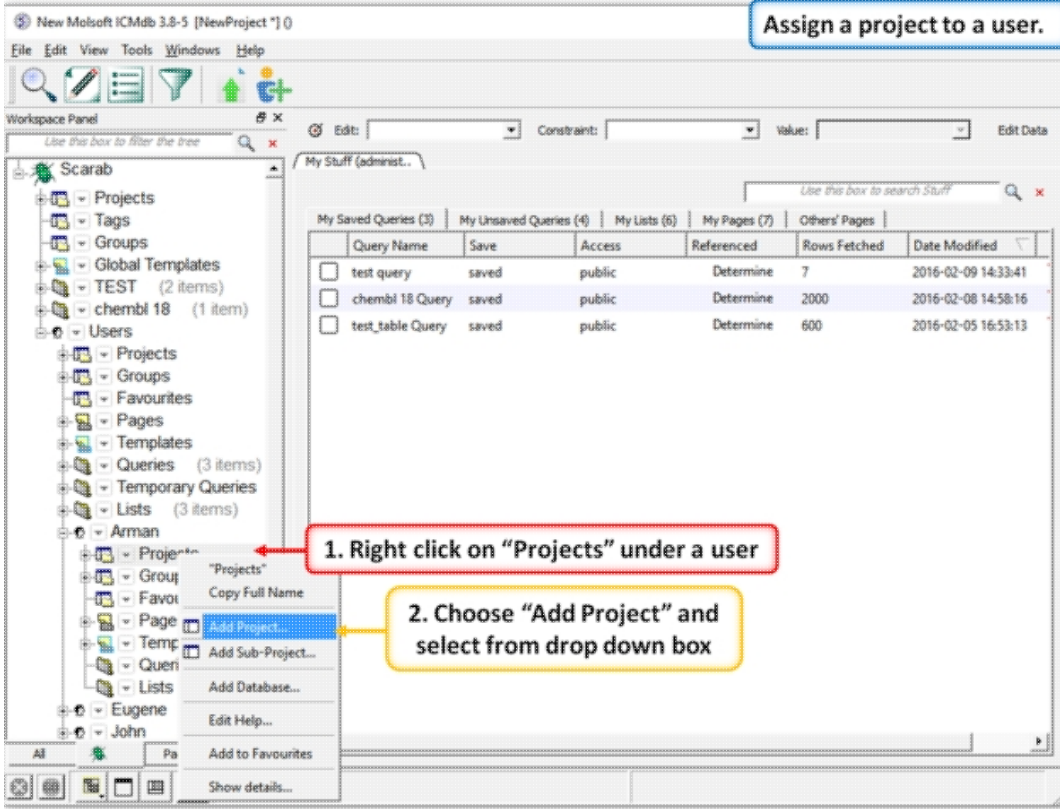
New users will be displayed in the tree along with any associated Pages, Projects and Saved Queries.

17.6 Create a New Project and Assign User

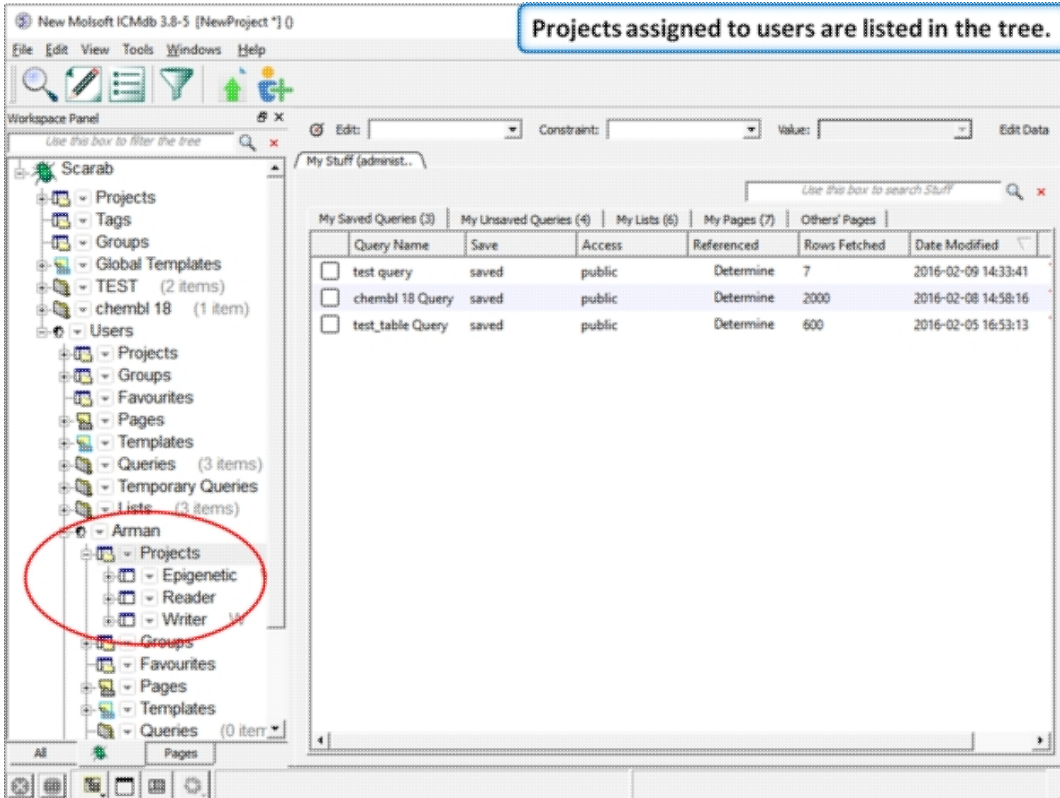
New projects can be created and assigned to a user.



To create a new project right click on the "Projects" branch of the tree and select "Create Project".



To assign a project to a User - right click on the Projects branch of the tree under the users and then choose "Add Project".



Projects that are assigned to users are listed in the tree.

17.7 Pages

Pages represents the Electronic Lab Notebook component of Scarab. Pages allows you to:

Intuitively write up your experiments

- Embed & annotate images
- Embed Excel spreadsheets
- Attach files
- Embed PDF files

Share your experiments

- Permissions at Project level
- Named users may write collaboratively

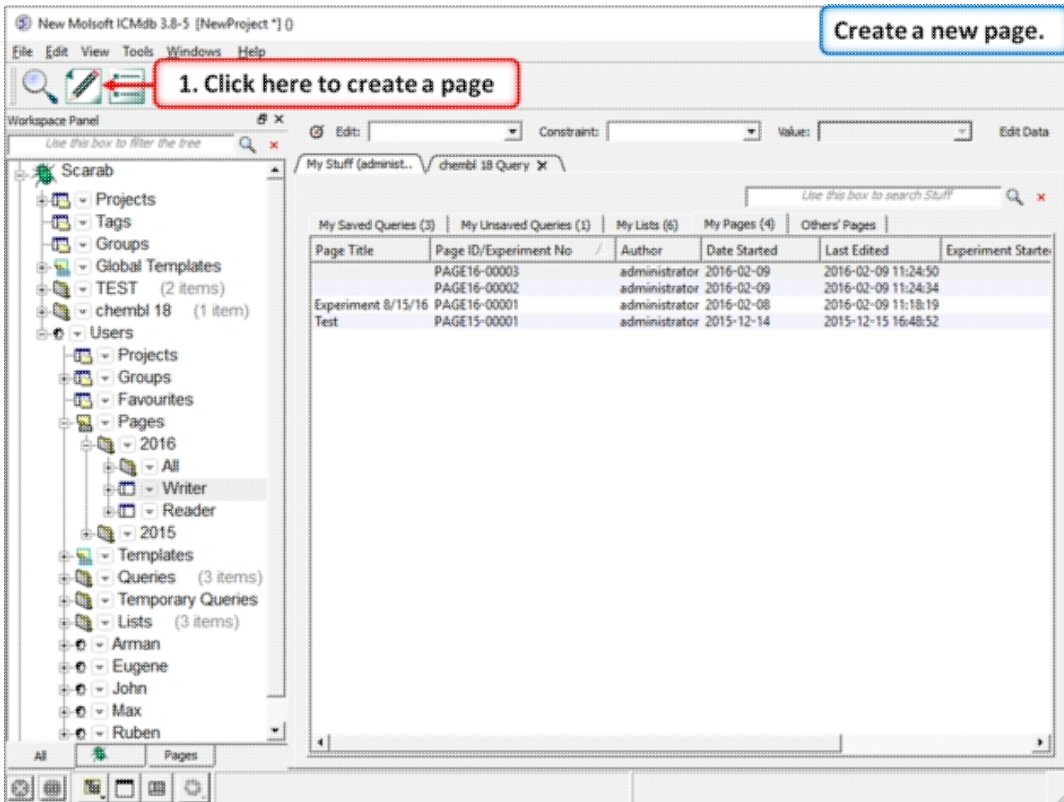
Export your Experiments

- Save as PDF
- Print experiments

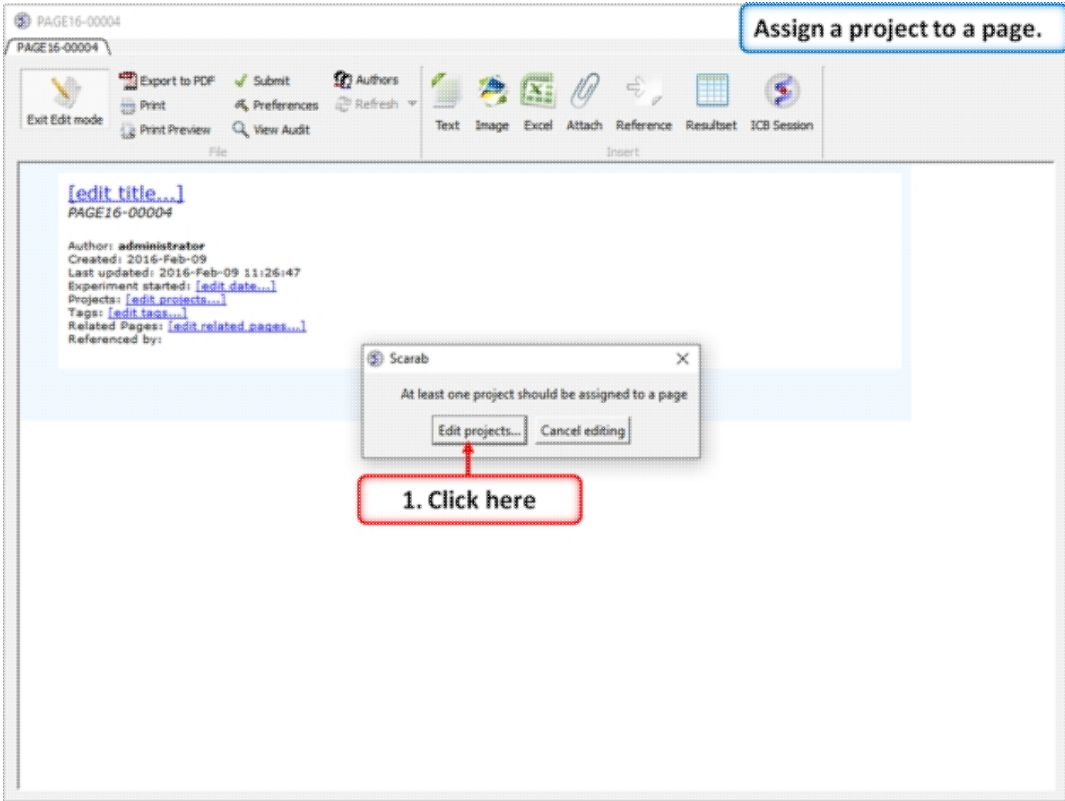
Upload Data

- Directly Import Data
- Type in Data
- Modify Data Within spreadsheet or form view environments
- Validate against business rules

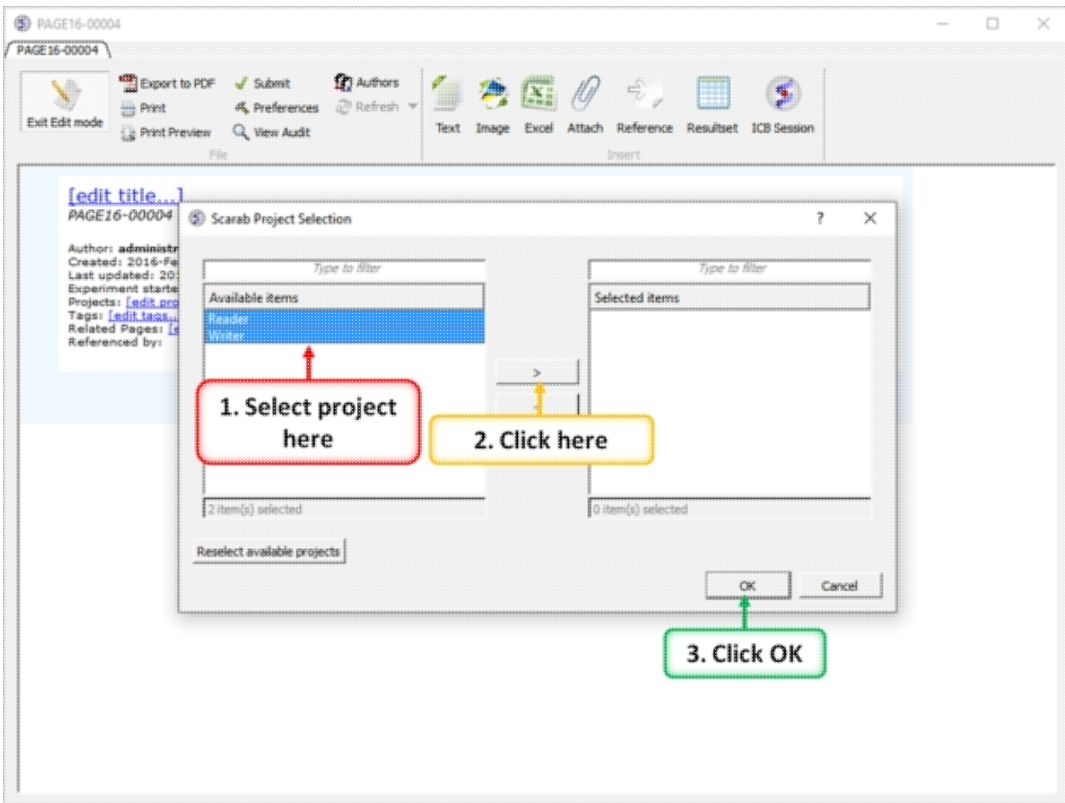
To create a new page:



Click on the "Create a Page" button.



First you need to assign a project to a page. The default projects are writer and reader but you can create new projects as described here.



Select the projects and move them using the buttons shown above. Multiple projects can be holding CTRL and clicking.

Click on hyperlinks to edit page title, date , associated project, tags and related pages...

Exit Edit mode

1. Make sure you are in edit mode

Text Image Excel Attach Reference Resultset ICB Session

[edit title...]
PAGE16-00005

Author: administrator
Created: 2016-Feb-09
Last updated: 2016-Feb-09 12:38:53
Experiment started: [edit_date...]
Projects: Reader:Writer [edit_projects...]
Tags: [edit_tags...]
Related Pages: [edit_related_pages...]
Referenced by:

2. Click to edit

Insert - Text, Image, Excel, Attachment, Reference, Resultset or MolSoft ICM file into the page.

Export to PDF Submit Authors
Print Preferences Refresh
Print Preview View Audit

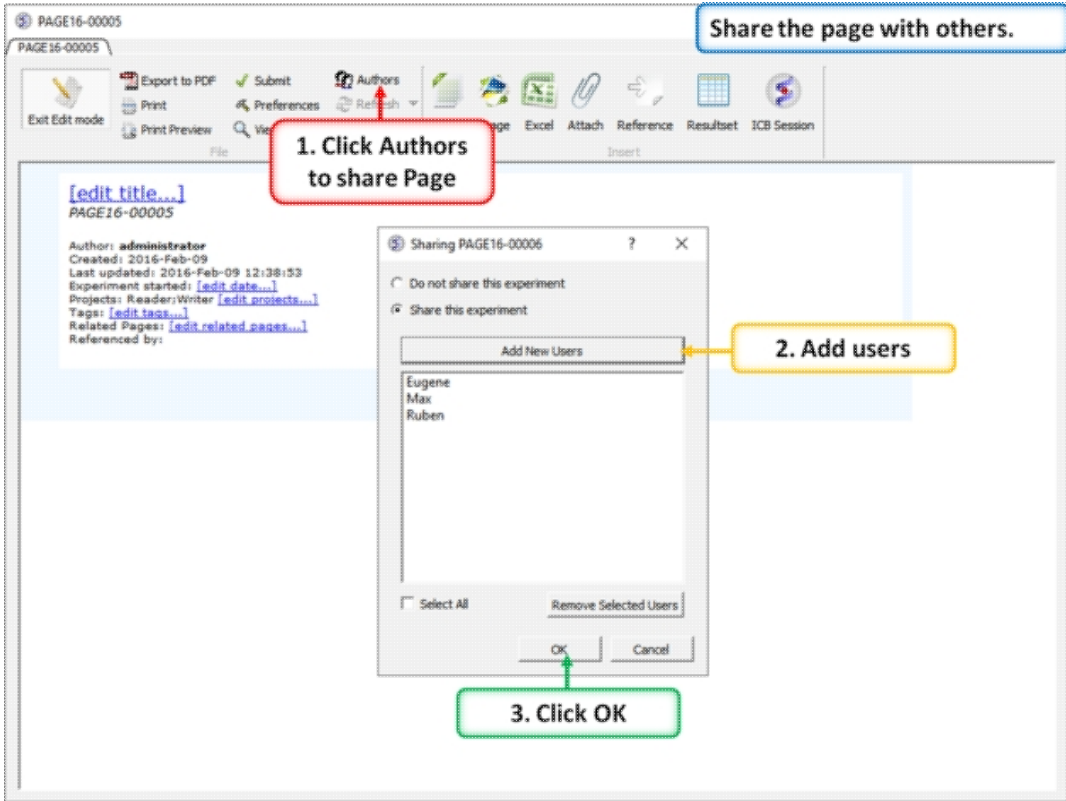
Text Image Excel Attach Reference Resultset ICB Session

1. Select the data type

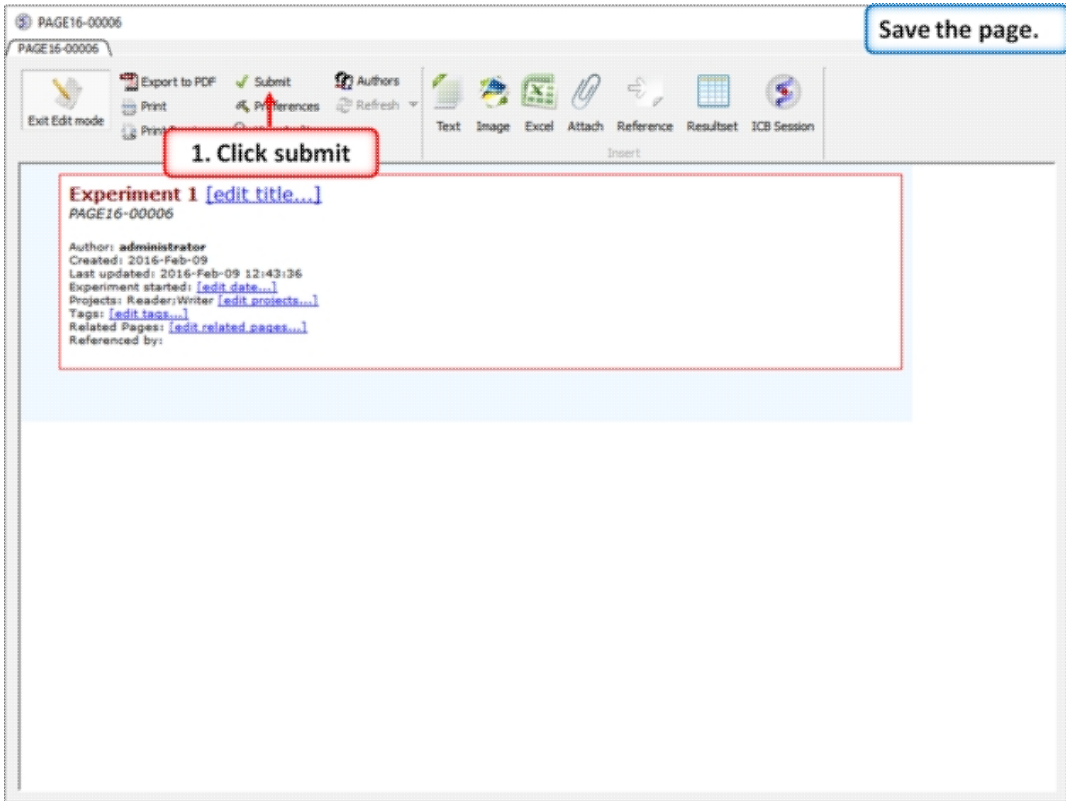
[edit title...]
PAGE16-00005

Author: administrator
Created: 2016-Feb-09
Last updated: 2016-Feb-09 12:38:53
Experiment started: [edit_date...]
Projects: Reader:Writer [edit_projects...]
Tags: [edit_tags...]
Related Pages: [edit_related_pages...]
Referenced by:

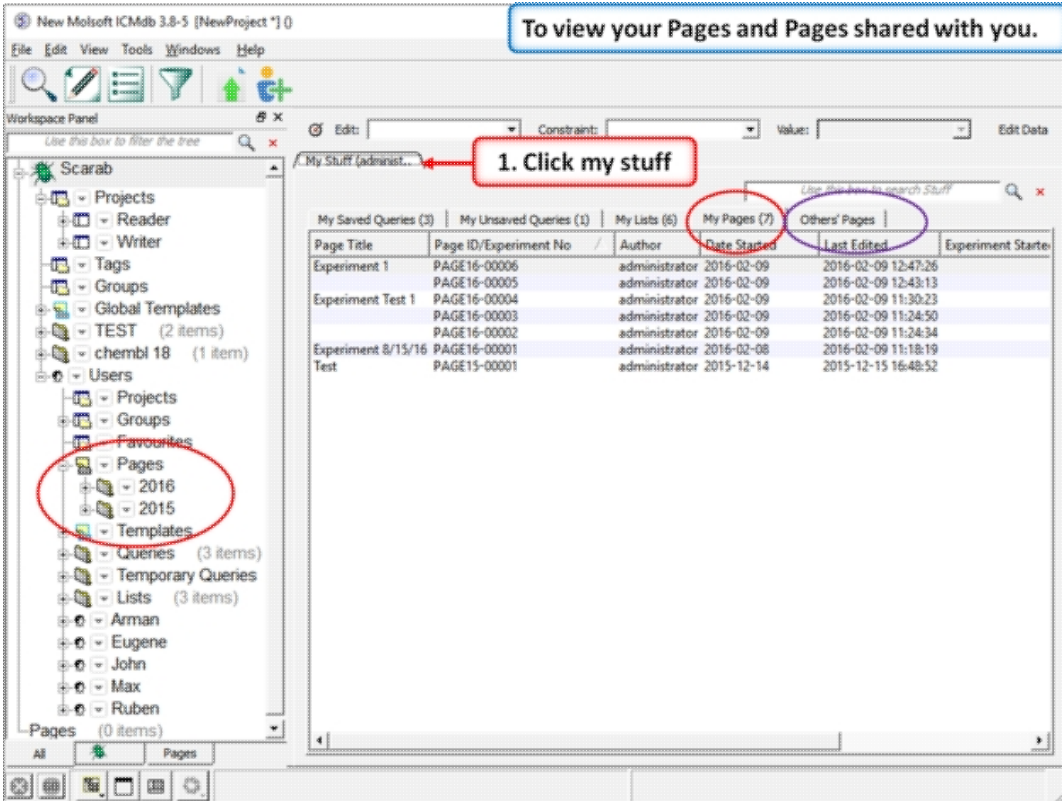
Insert - Text, Image, Excel, Attachment, Reference, Resultset or MolSoft ICM file into the page.



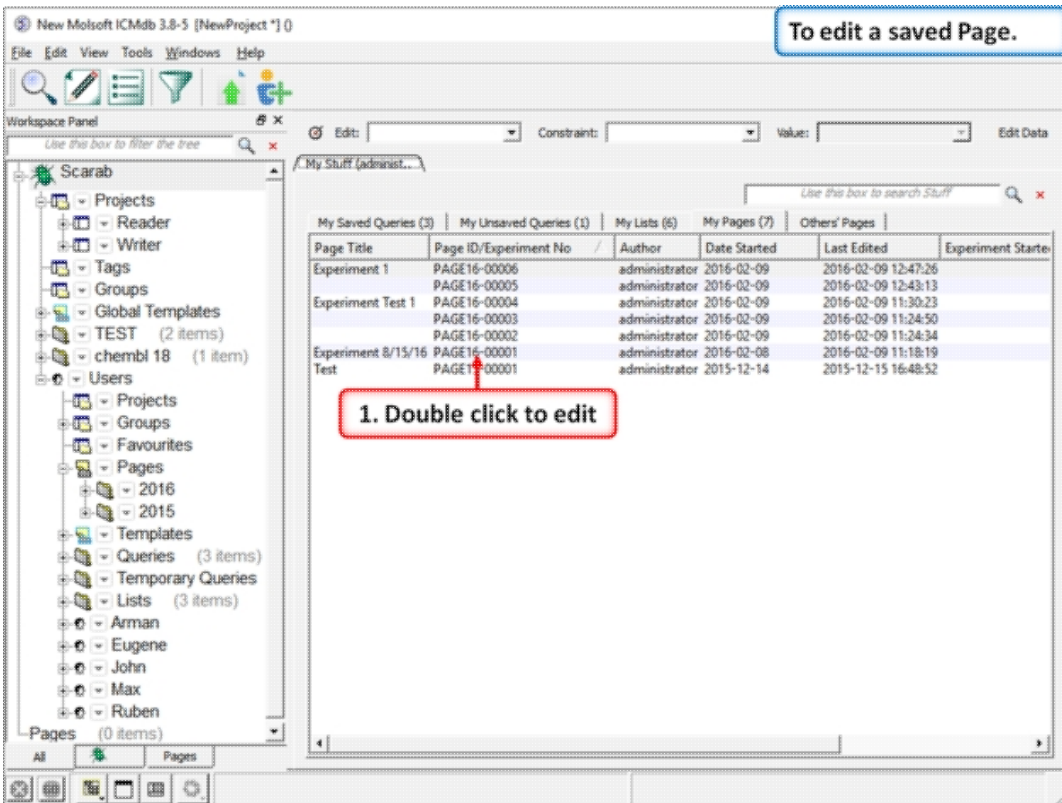
You can share a page with other users.



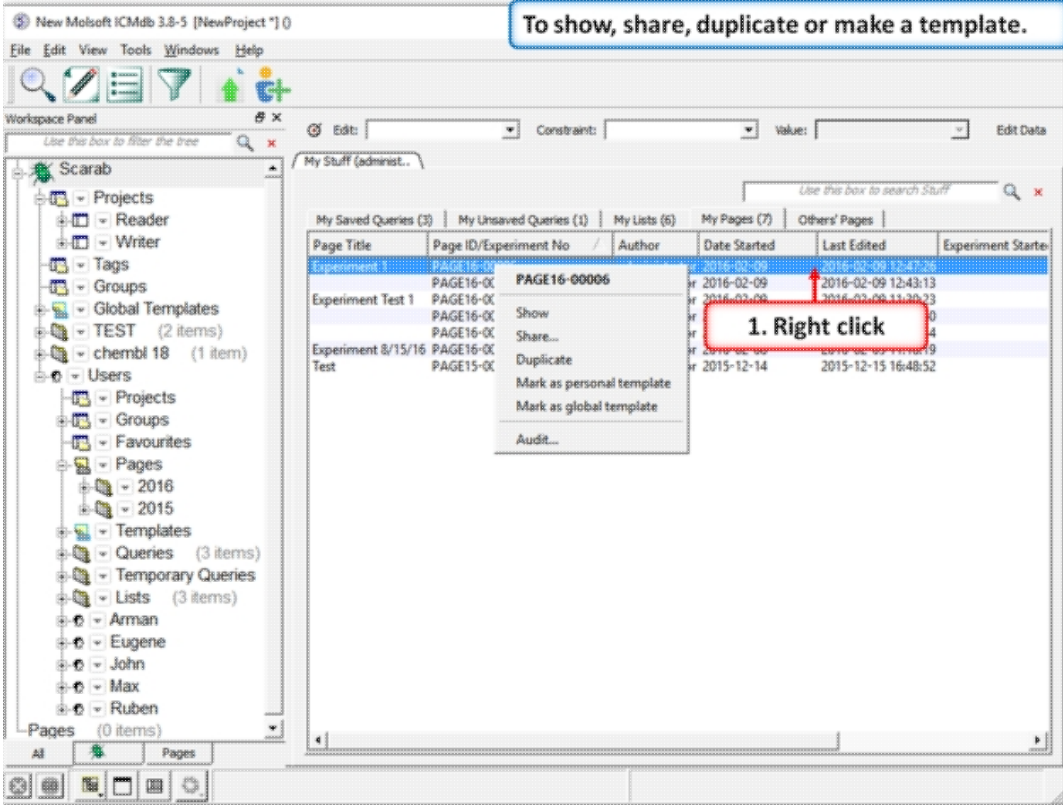
Save the page by clicking "Submit".



Click on "My Stuff" tab to see your pages and others' pages.



Right click on a page in the "My Stuff" tab to edit a saved page.



Right click on a page in the "My Stuff" tab to show, share, duplicate or make a template.

18 ICM in KNIME

|Video - Installation|Video - MolCalcProp Node|

KNIME (Konstanz Information Miner) is a user-friendly and comprehensive open-source data integration, processing, analysis, and exploration platform. Currently we have KNIME nodes for the following ICM functions:

Available Nodes

Docking and Screening

- **MolIcmDockPrep** - Prepare receptor and ligand for docking. Covalent docking support for MolIcmDockPrep node (ICM 3.8-1 or higher is required)
- **MolIcmDock** - Ligand Docking and Screening (`_dockScan` and `_dockProjPrep`). Recently we have added "Output Type" for MolIcmDock node. It allows to choose between SDF and interactive ICM HitList which can be viewed using the MolView node.
- **MolScreen** - MolScreen is a set of high quality 2D fingerprint and 3D pharmacophore models for a broad range of pharmacology and toxicology targets.

Cheminformatics

- **MolConverTo3D** - Chemical 2d to 3d conversion.
- **MolConfGen** - Conformation generator (`_confGen`).
- **MolAPFSuperimpose** - Flexible APF superposition (`_chemSuper`).
- **MolCalcProp** - Allows to calculate various chemical properties and descriptors (e.g. Calculate MolArea, MolWeight, InChi...). For a complete list of available functions please click [here](#).
- **MolStandardize** - Chemical Table/Spreadsheet standardization - remove salt, remove hydrogen, standardize groups, apply 2D depiction, assign formal charges.
- **MolUnique** - removes duplicate molecules.
- **MolCenters** - Performs chemical clustering and selects a specified number of representatives.
- **MolCluster** - Cluster either by fingerprints or by numerical column selection.

Visualization

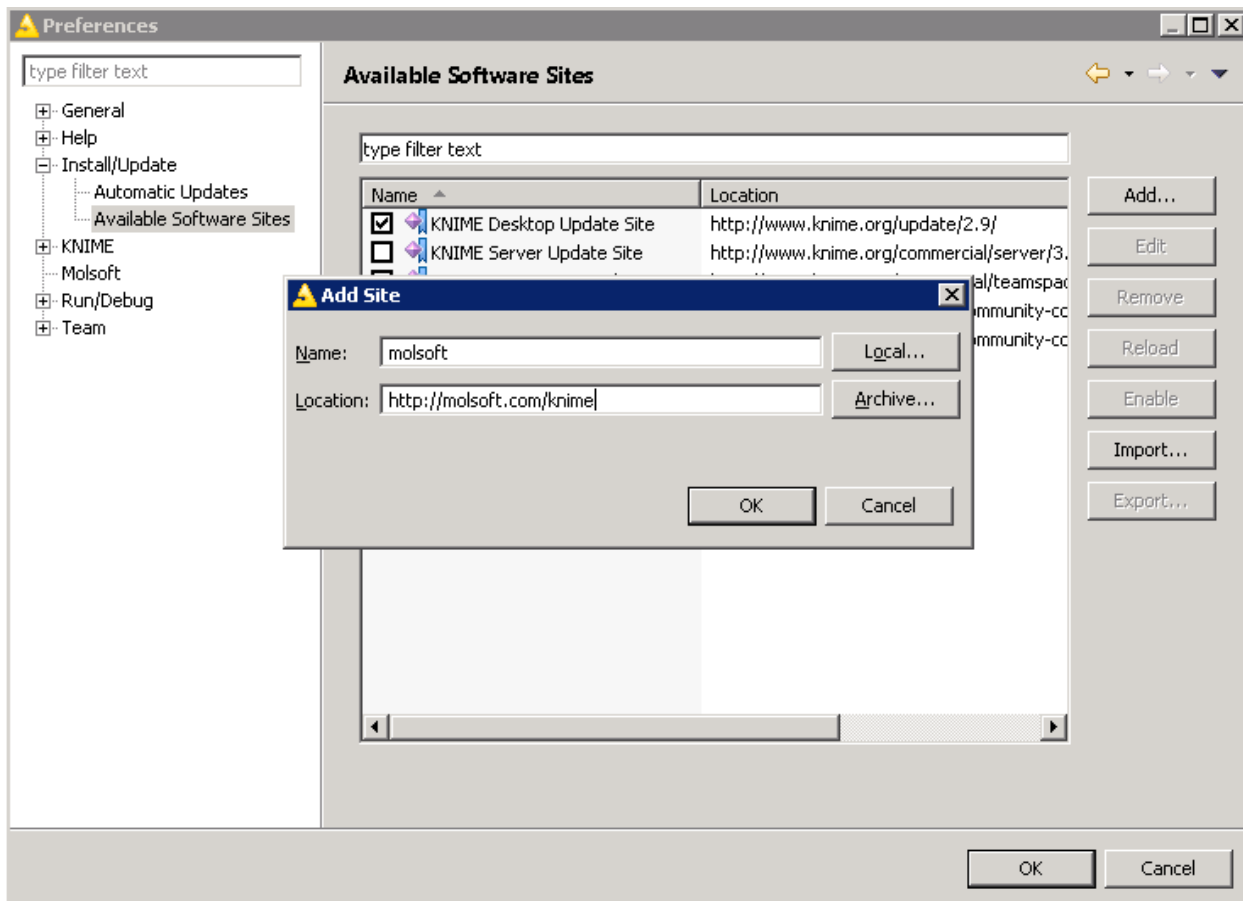
- **MolView** - Connect this node to the output of other Molsoft nodes. (e.g. MolScreen, MolIcmDock...).

To use ICM in KNIME

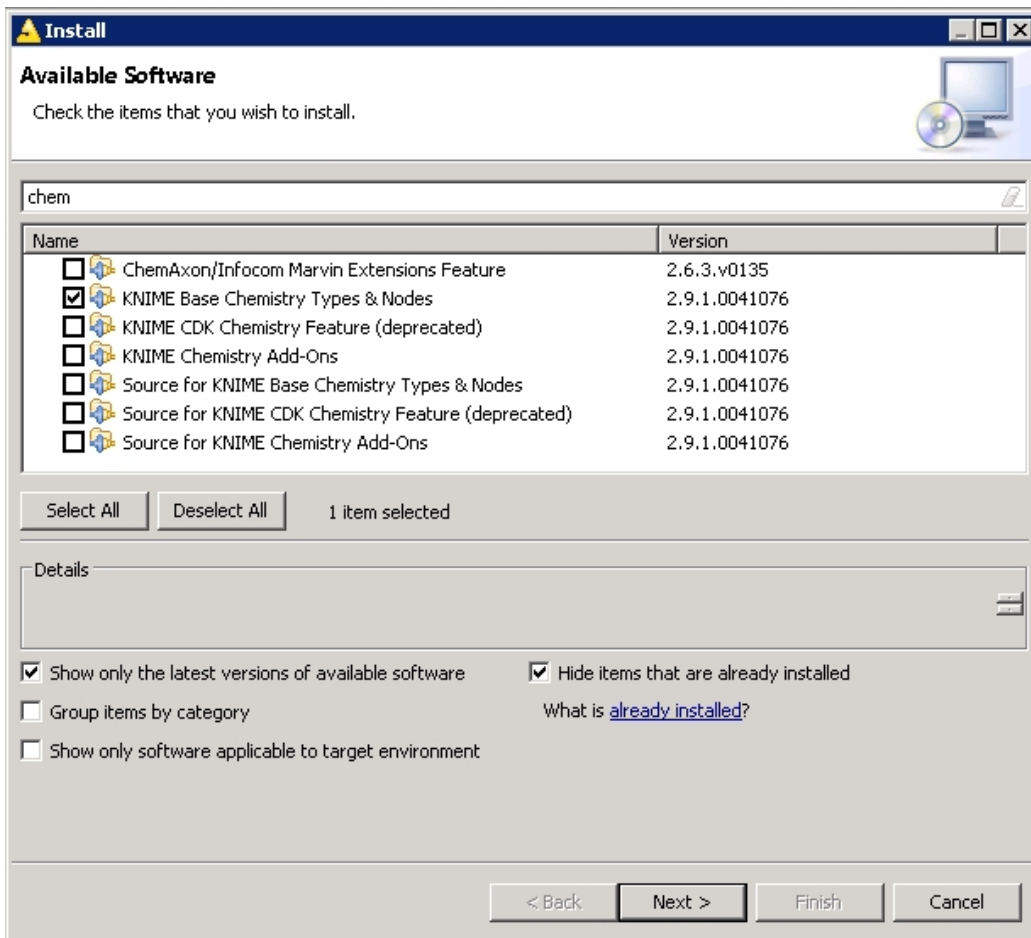
1. Download KNIME (Analytics Platform) from www.knime.org
2. In KNIME go to File/Install KNIME extensions and search for and select the KNIME Base Chemistry Types and Nodes. When you do this you will see a Chemistry directory in the Node Repository.
3. Download MolSoft's KNIME Java (see instructions below). **Note, that if you have an old version of Molsoft Nodes installed manually you need to remove the jar file and restart KNIME with "-clean" option.**

MolSoft Knime

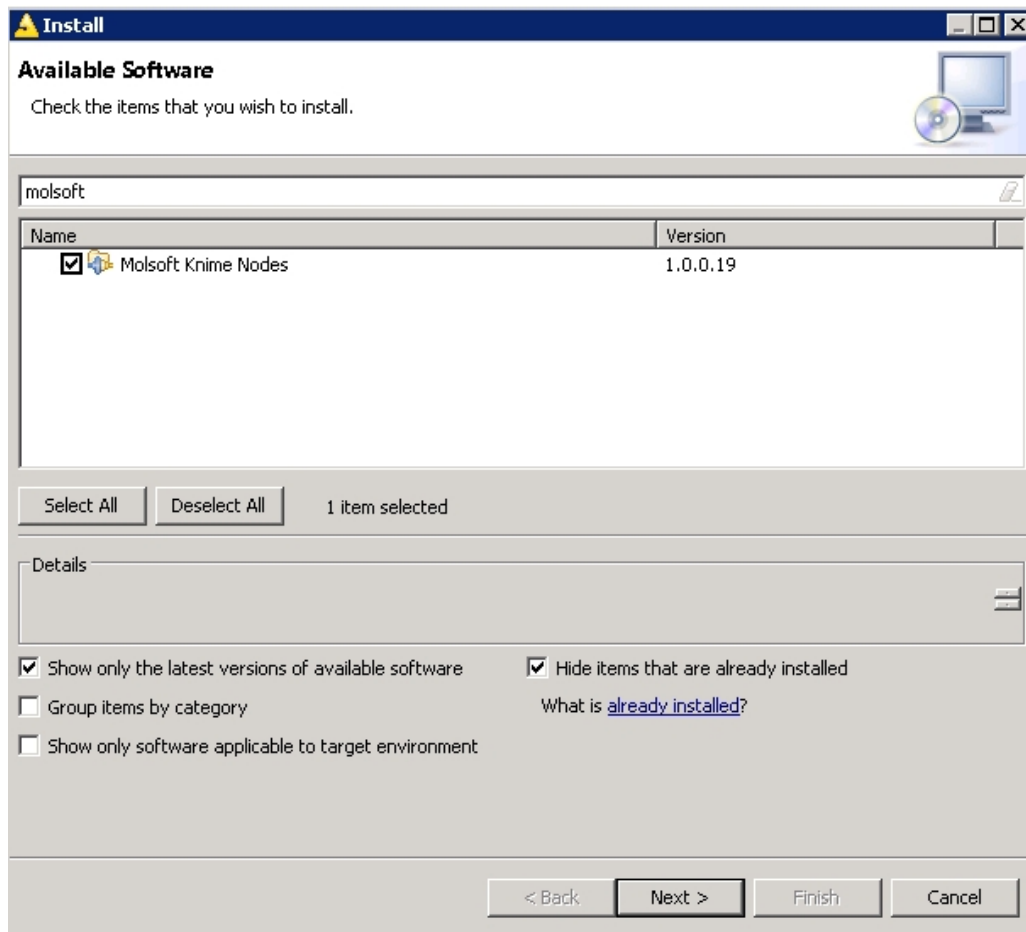
- To enable automatic update of molsoft nodes: go to File/Preferences. Open Install/Update - Available Software Sites, Click on "Add..." button and enter information below:



- Then, goto "File/Install update extensions", uncheck "Group items by category" If you don't have "Base Chemistry" package installed, then install it first (type 'chem' to locate package)



- Next install 'Molsoft' nodes. Please uncheck the option "Group items by category" otherwise it will not list the nodes.



- After this point everything is setup. You can either go to File/Update KNIME to get the updates manually or setup automatic update.

How to run MolSoft in Knime:

Setting Path to ICM Executable Usually KNIME can figure out where the Molsoft ICM executable is located. In some cases when there are multiple versions of ICM it might return an error as shown below:

Execute failed: Cannot run program "/icm": CreateProcess error=2, The system cannot find the f

To fix this:

- File/Preferences
- Select MolSoft and then browse for the location of ICM e.g.C:\Program Files (x86)\Molsoft LLC\ICM-Pro

Here is an example for Structure-based virtual ligand screening in KNIME:

Example: Structure-based virtual ligand screening nodes.

1. Drag ICM Nodes from here

2. Click to configure and run

3. Documentation on each node here

4. SDF writer node is in the Chemistry Node Directory in the Node Repository

SDF Writer

This node writes SDF or Mol cells into a continuous SDF file. You may also add additional properties from columns to the SDF molecules while they are written out.

KNIME Console

```
*** Welcome to KNIME v2.5.4.0032975 - the Konstanz Info
*** Copyright, 2003 - 2012, Uni Konstanz and KNIME Gmb
Log file is located at: C:\Program Files (x86)\Molsoft Li
```


19 Tutorials

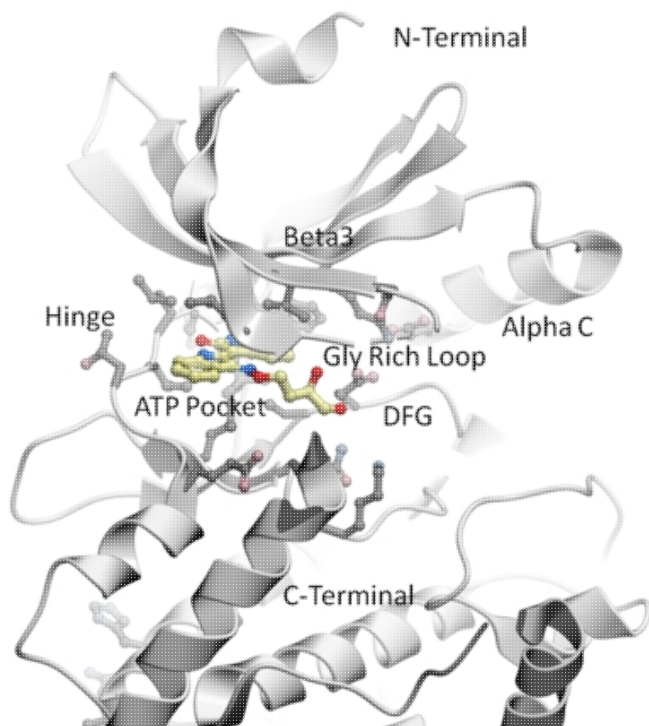
Graphics Tutorials		
Molecule Representation, Coloring, Labeling and Annotation	HTML	
Creating Fully Interactive Slides for PowerPoint and the Web	HTML	Reference Guide Webinar
Sequence and Alignment Tutorials		
Read Sequences and Align	HTML	Video
Link Sequence to Structure	HTML	Video
Alignment Annotation	HTML	
Protein Structure Tutorials		
PDB Search	HTML	
Convert PDB to ICM Object	HTML	
Superimpose Proteins	HTML	
Structure Analysis	HTML	
Calculate contacts	HTML	
Display hydrogen bonds	HTML	
Identify ligand binding pockets	HTML	
PDB Preparation: Symmetry	HTML	
PDB Preparation: Occupancy & B-Factors	HTML	
PDB Preparation: Alternative Orientation	HTML	
PDB Preparation: Biomolecule	HTML	
Protein Modeling Tutorials		
Building an Homology Model	HTML	
GPCR Modeling Example	HTML	
Loop Modeling	HTML	
Predicting the Effect of a Mutation on Binding	HTML	
Predicting the Effect of a Mutation on Stability	HTML	
Cheminformatics Tutorials		
Chemical Sketching using the Molecular Editor	HTML	Reference Guide
Chemical Search	HTML,	Reference Guide
Chemical Clustering	HTML,	
3D Pharmacophore Search	HTML	Video
2D Pharmacophore Search	HTML	Video
How to Create a Markush Structure	HTML	Video
How to Enumerate a Markush Library	HTML	Video
How to Decompose a library based on a Markush Library	HTML	Video
How to Enumerate a library by reaction	HTML	Video
Atomic Property Fields (APF)		
APF Superposition, Score, Consensus, Screening and Clustering.	HTML	
Ligand Editor Tutorials		
Working with the ICM Interactive 3D Ligand Editor	HTML	Reference Guide.
Ligand Docking Tutorials		
Dock Biotin to the Streptavidin Receptor	HTML	Video
Re-Dock an Inhibitor to Ricin Crystal Structure	HTML	Video
Covalent Docking	HTML	Video
Protein-Protein Docking Tutorial		
Protein-Protein Docking using FFT Method	HTML	
Virtual Screening Tutorials		
Virtual Ligand Screening to Ricin Receptor	HTML	
Virtual Ligand Screening to Cyclooxygenase	HTML	Video-1 Video-2
Docking a Markush Library	HTML	
Induced Fit Docking Tutorials		

19.1 Graphics Tutorials

Introduction

The following examples are focused on the structure of the kinase catalytic domain. The catalytic domain is nestled between the N- and C- terminal and has high sequence conservation between kinase families. The adenine moiety of ATP interacts with the hinge region which links the N- and C-terminals of the catalytic domain. The ATP sugar binding region is on the "floor" of the pocket where ribose makes a hydrogen bond with a polar residue. A conserved lysine residue on the beta3 strand is part of the "roof" of the pocket where the triphosphate group of ATP binds. A flexible glycine-rich loop moves in and out of the pocket depending on the ligand bound state of the PK and is regulated to some extent by the movement in and out of the pocket by the AlphaC helix. A buried region at the "back" of the pocket is protected by a "gatekeeper residue" forming a variable hydrophobic cavity. The hydrophobic cavity along with the DFG motif region are of interest for drug design because it opens up regions of the pocket which are not conserved and do not bind ATP.

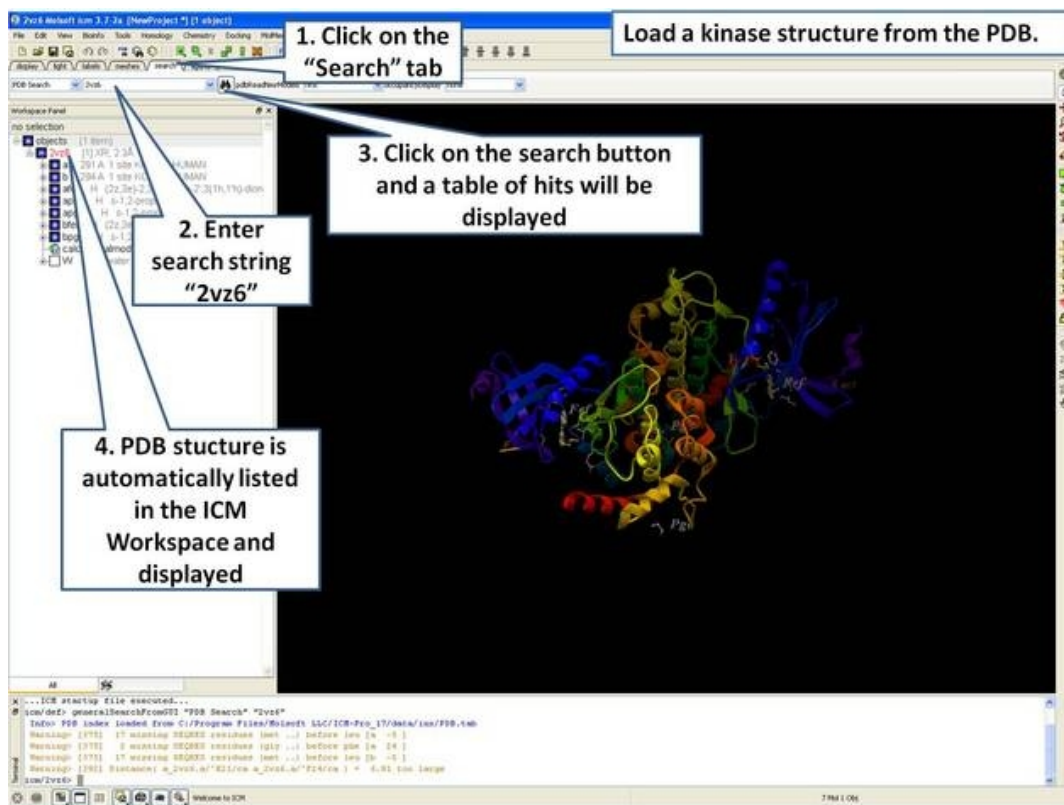
To learn the basics of the graphical user interface we will annotate the key structural regions of a protein kinase.



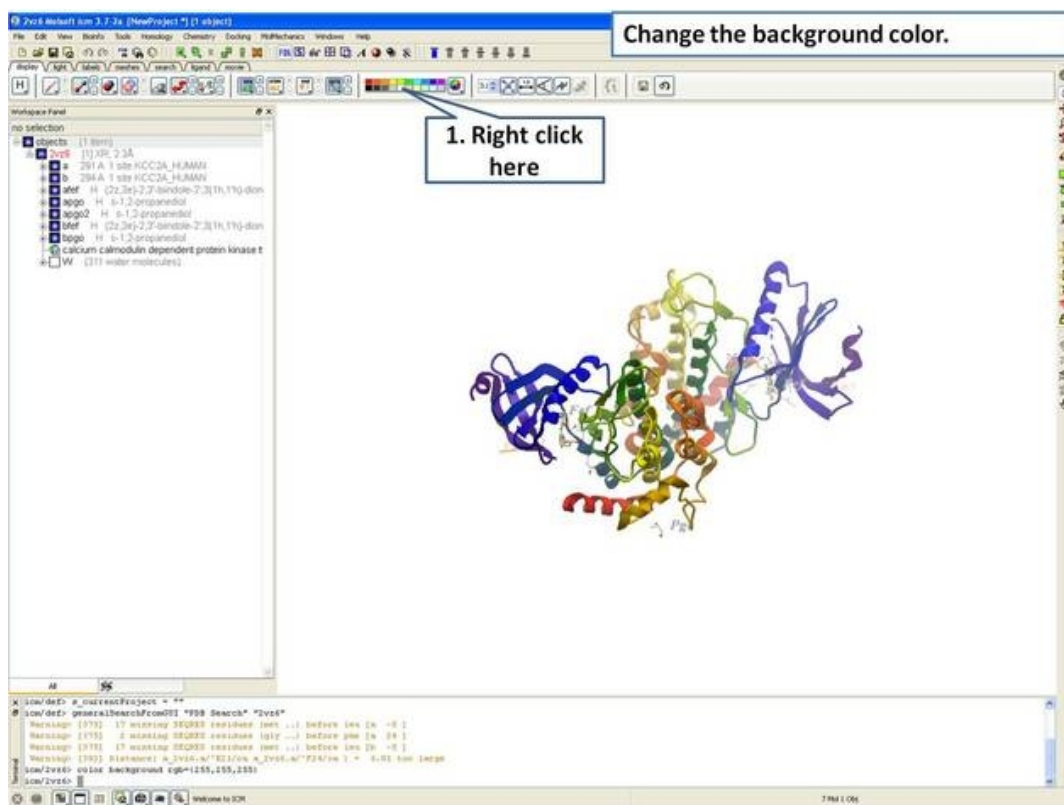
19.1.1 Change Molecule Representation and Color

This tutorial shows you how to:

- How to load a PDB structure into ICM.
- How to change the background color.
- How to display and undisplay a molecule.
- How to color ribbon representation.



Step 1: Click on the search tab to load and display the PDB structure 2vz6



Step 2: Change the background color by right clicking on the color palette.

Display the "a" subunit and ligand.,

1. Click in the boxes to display and undisplay

2. Use these buttons to rotate molecule

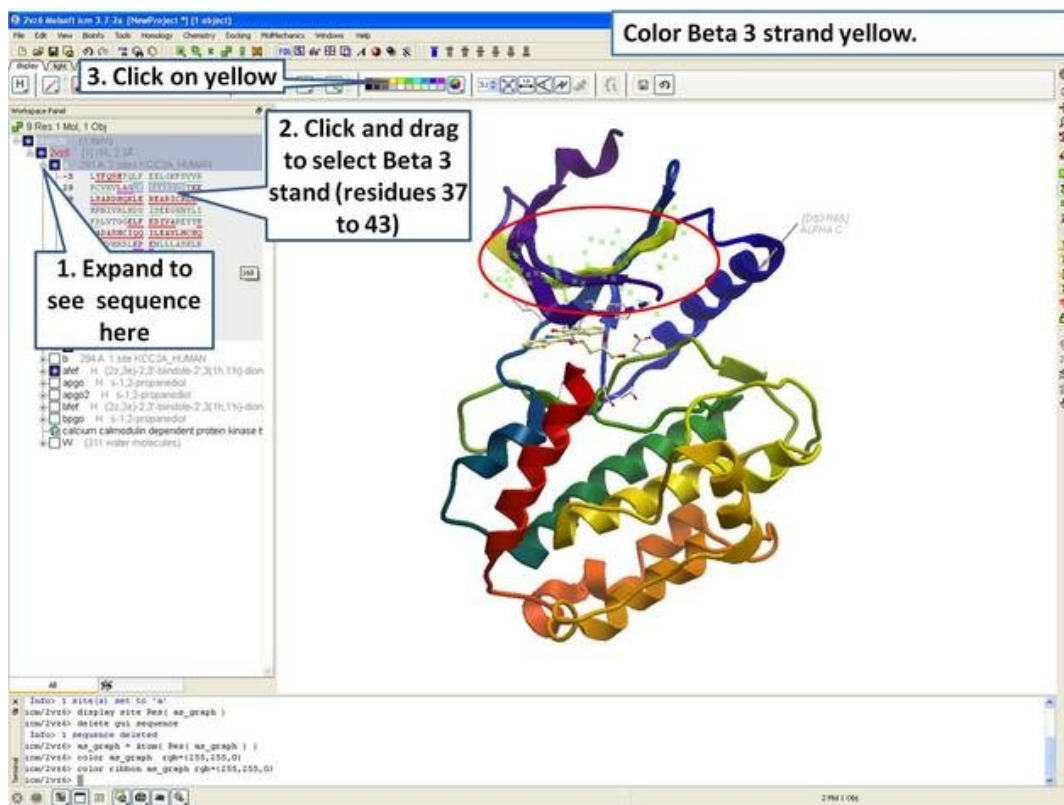
Step 3: Display the "a" subunit of the kinase.

Color ribbon by secondary structure segment.

1. Click and hold

Note: All representation buttons have similar coloring options

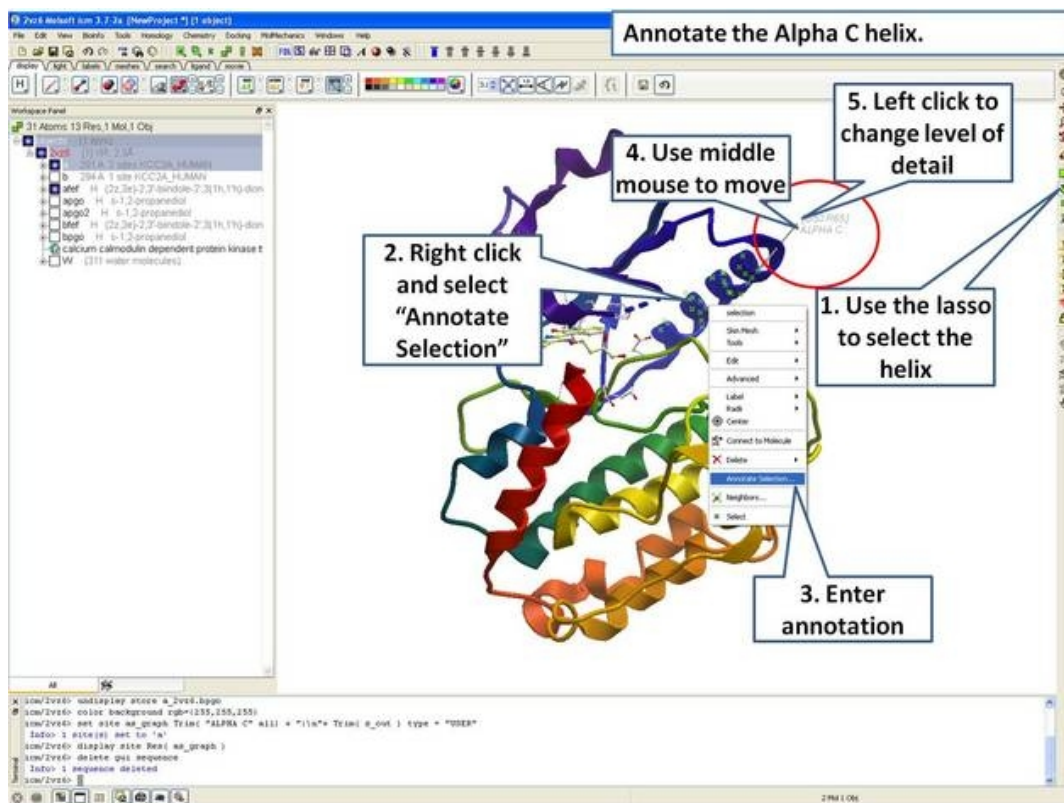
Step 4: Use the options in the "Display Tab" to color the ribbon.



Step 5: Select the beta 3 strand (residues 37:43) and color yellow.

19.1.2 Annotation

This tutorial shows you how to add user defined annotation to a particular region of a protein structure.

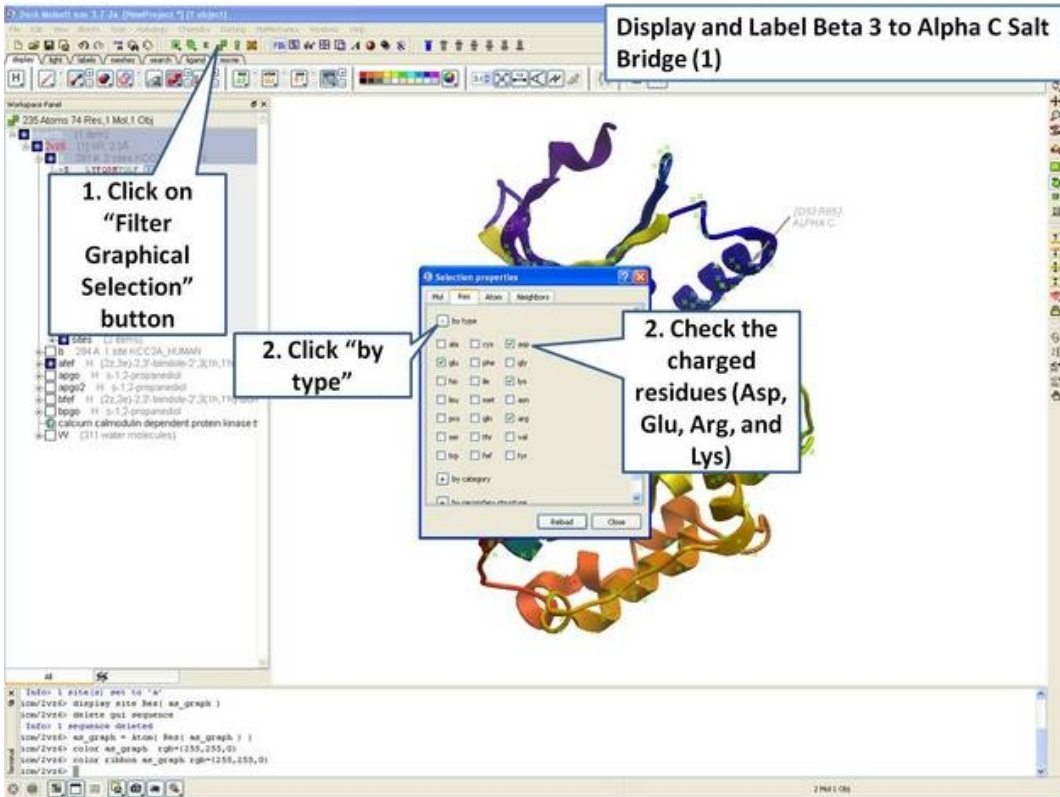


Step 6: Select the region you wish to annotate and right click on the selection for options.

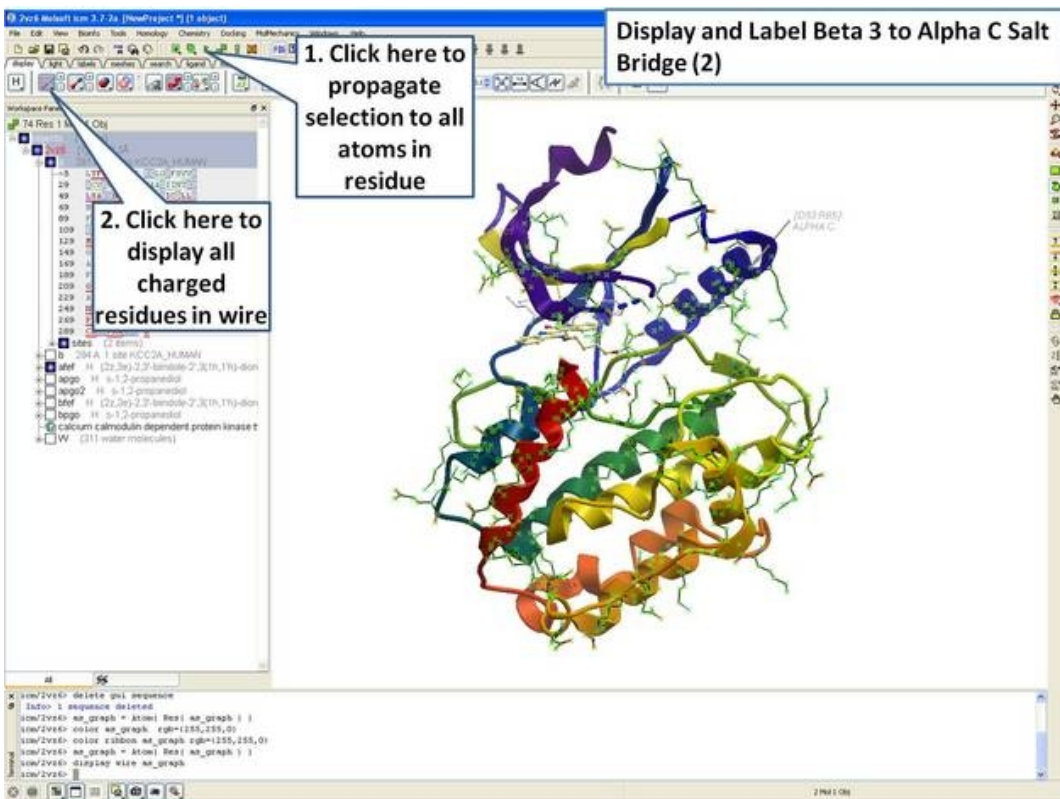
19.1.3 Labels

This tutorial is a continuation from the previous tutorials in this section and shows you how to:

- How to select individual residue types.
- How to propagate a selection from one atom to all.
- How to use the residue label button.
- How to make a spherical selection.
- How to label a residue selection.



Step 7: Use the "Filter Graphical Selection" button to select charged residues.



Step 8: Propagate selection to all taoms to display all atoms of the side chain.

Display and Label Beta 3 to Alpha C Salt Bridge (3)

1. Undisplay all xstick representations for clarity

2. Use the zoom in button to locate alpha C to beta 3 salt bridge

3. Click on the residue label button

4. Click on each residue to label

5. Click and hold on the residue label button for label coloring and style options

Step 9: Locate the B3 to AlphaC salt bridge (K42- E60) and label using residue label button. **Note:** Once you have selected the two residues you can invert selection (click on invert selection button) and then undisplay all the other charged residues.

Display and label the residues near the ligand binding pocket (1).

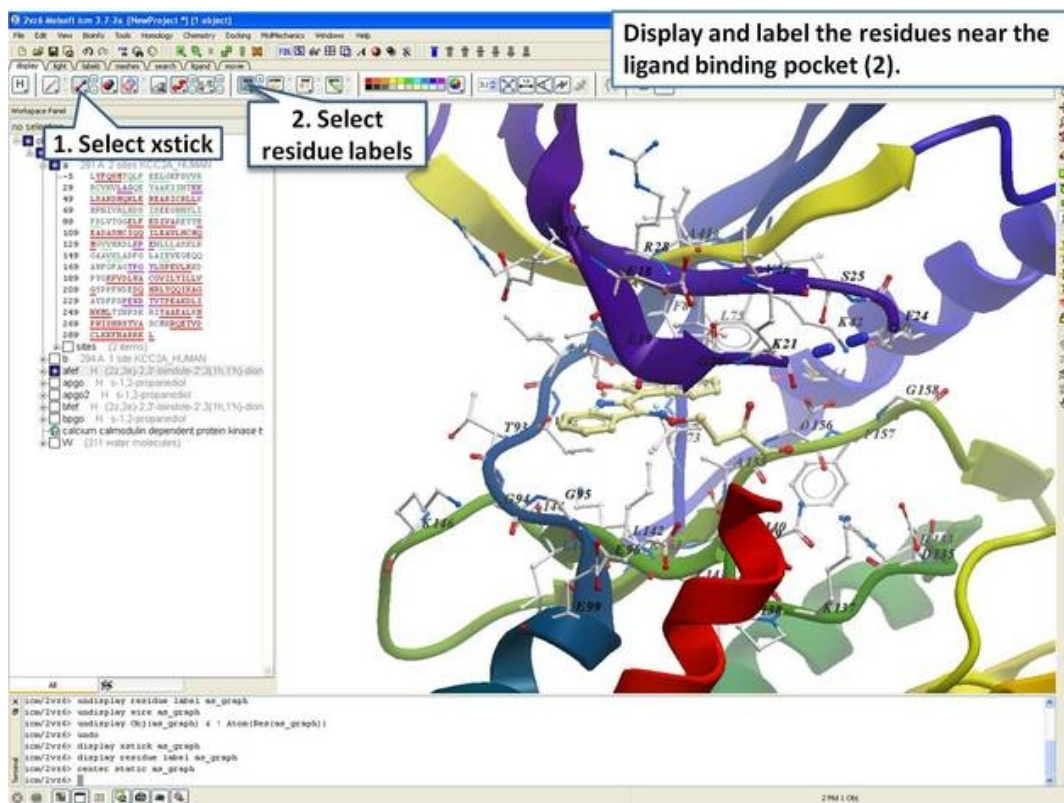
1. Undisplay wire and other label annotations from previous examples

2. Right click on the "a" subunit ligand and choose "Select Neighbors"

3. Enter the parameters as shown

4. Propagate selection to all atoms

Step 10: Select neighboring residues to the ligand.

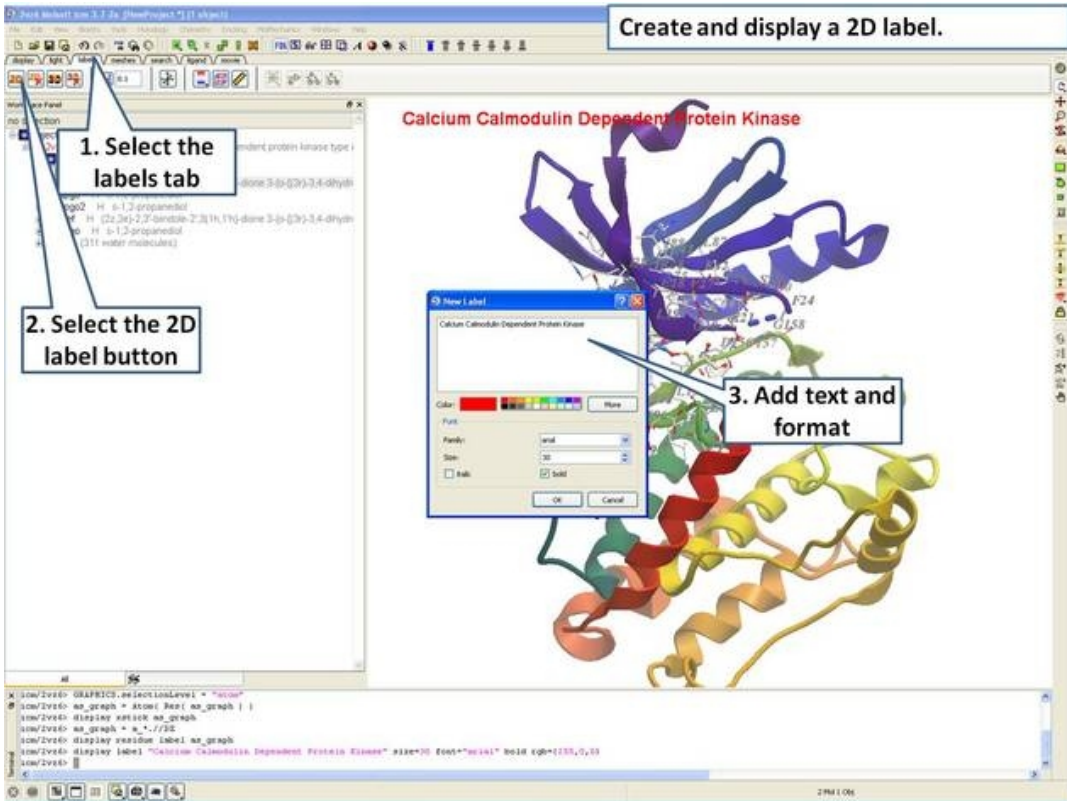


Step 11: Display the residues and label them.

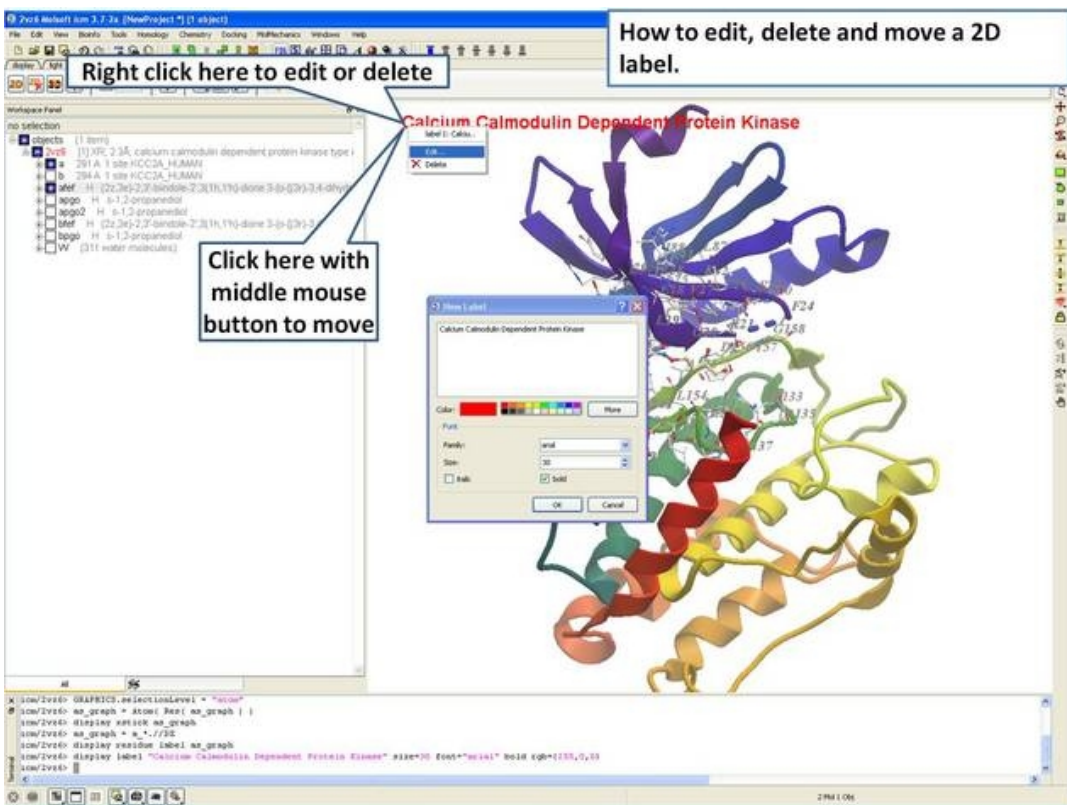
19.1.4 2D and 3D Labels

This tutorial is a continuation from the previous tutorials in this section and shows you how to:

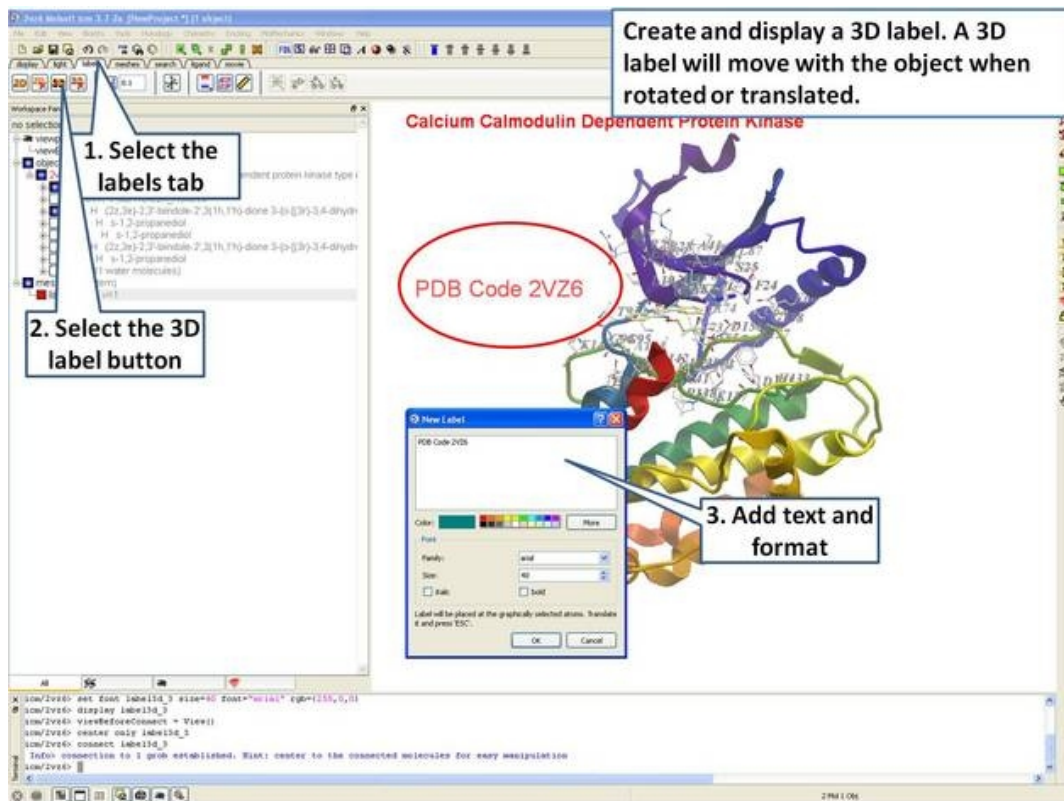
- Create and display a 2D label.
- Delete and move a 2D label.
- Create and display a 3D label.
- Delete and move a 3D label.



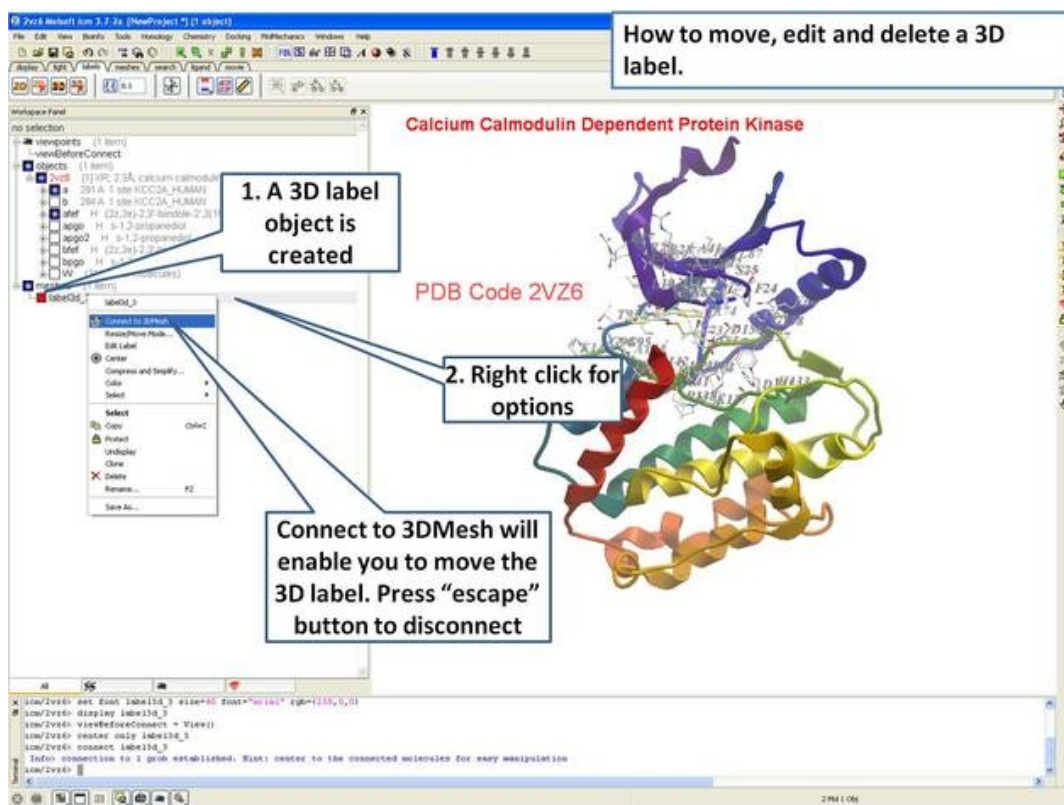
Step 12: Create a 2D label using the options in the "labels" tab.



Step 13: Right click on the 2D label to edit or delete it.



Step 14: Create a 3D label using the options in the "labels" tab.



Step 15: Right click on the 3D label to edit or delete it.

19.2 Creating Fully Interactive Slides for PowerPoint and the Web Tutorial

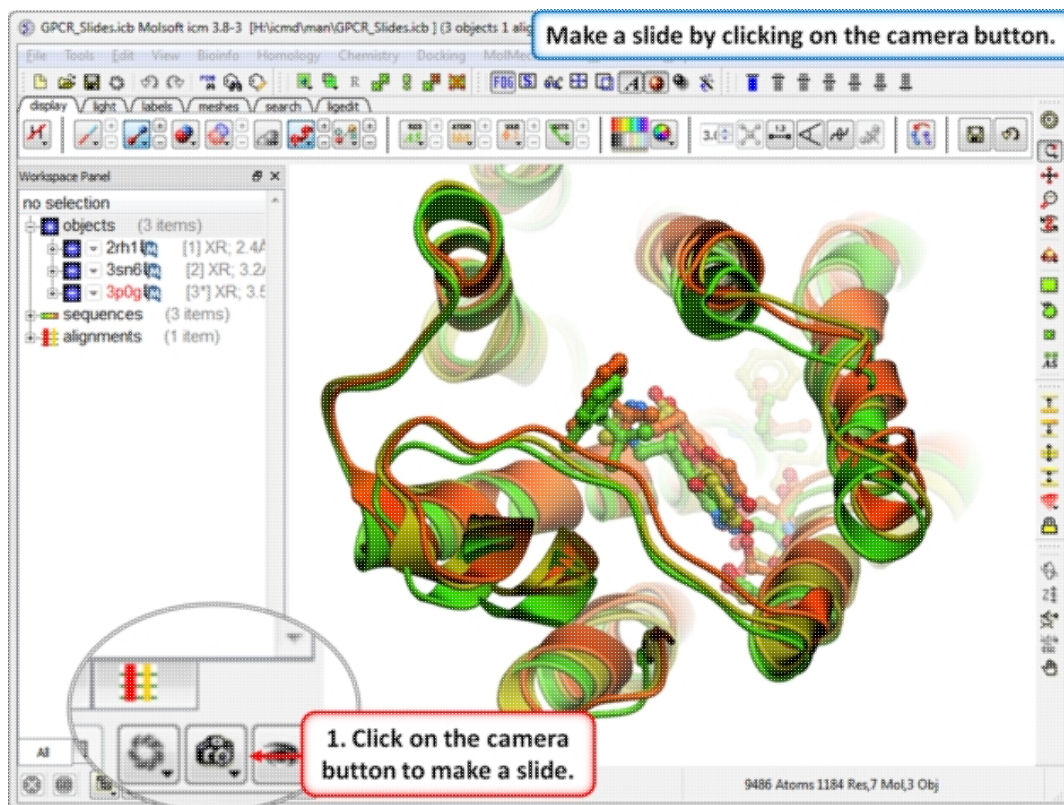
Introduction

A slide enables you to conveniently store a large number of different 3D visualization properties (e.g. color, viewpoint, representation etc...) along with text and window layout. ActiveICM enables you to view and display ICM graphical slides and animations interactively inside Windows Microsoft PowerPoint and web browsers such as Internet Explorer and Mozilla Firefox.

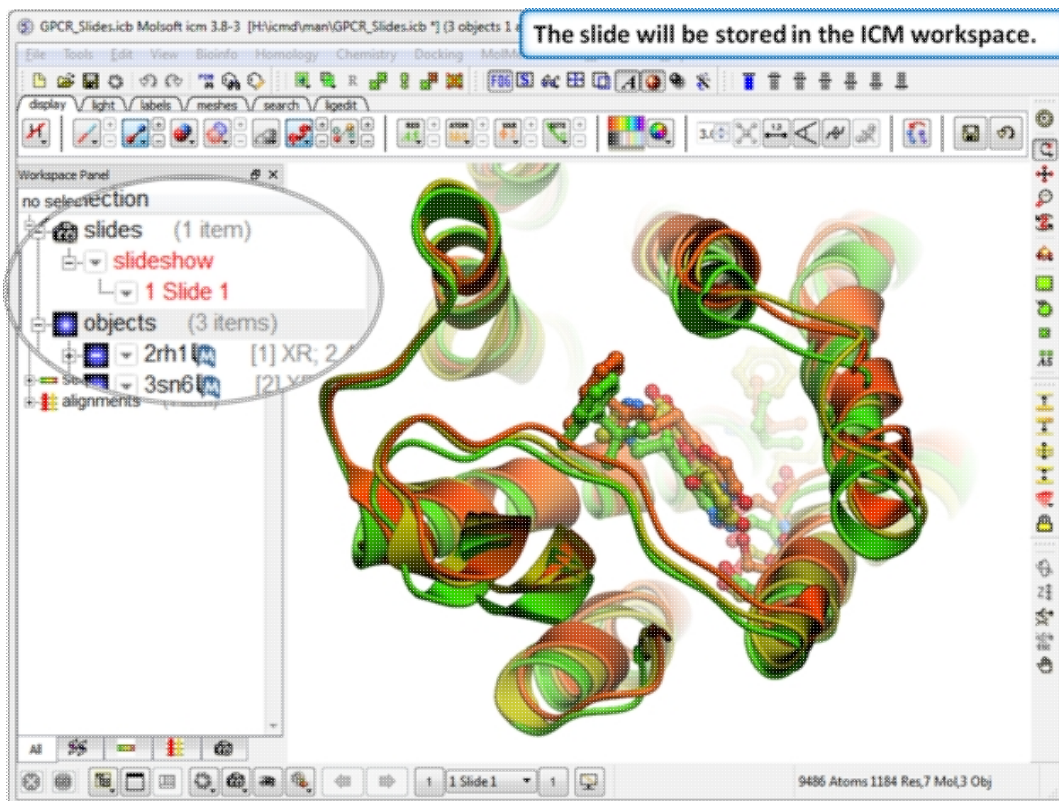
In this example we are going to look at the effect of an agonist binding to the Beta-2-Adrenergic GPCR (B2AR) and prepare a molecular document. A prepared ICM file containing three B2AR GPCRs superimposed can be downloaded here <ftp://ftp.molsoft.com/pub/workshop/> (1. cut and paste the address into a web browser. 2. Right click on GPCR_slides.icb and choose "Save Link As". 3. Open the .icb in ICM File/Open). The file contains pdb files 2RH1 (R inactive state), 3POG (R* active state, with G-alpha mimic), and 3SN6 (R*G G-protein signalling state). The agonist ligand causes Helix 5 to shift inwards and Helix 6 to swing outwards. This agonist binding poses were predicted using ICM a year before the crystal structures were solved, you can read about this [here](#) and a good review of the structure and function of GPCRs [here](#).

Tutorial

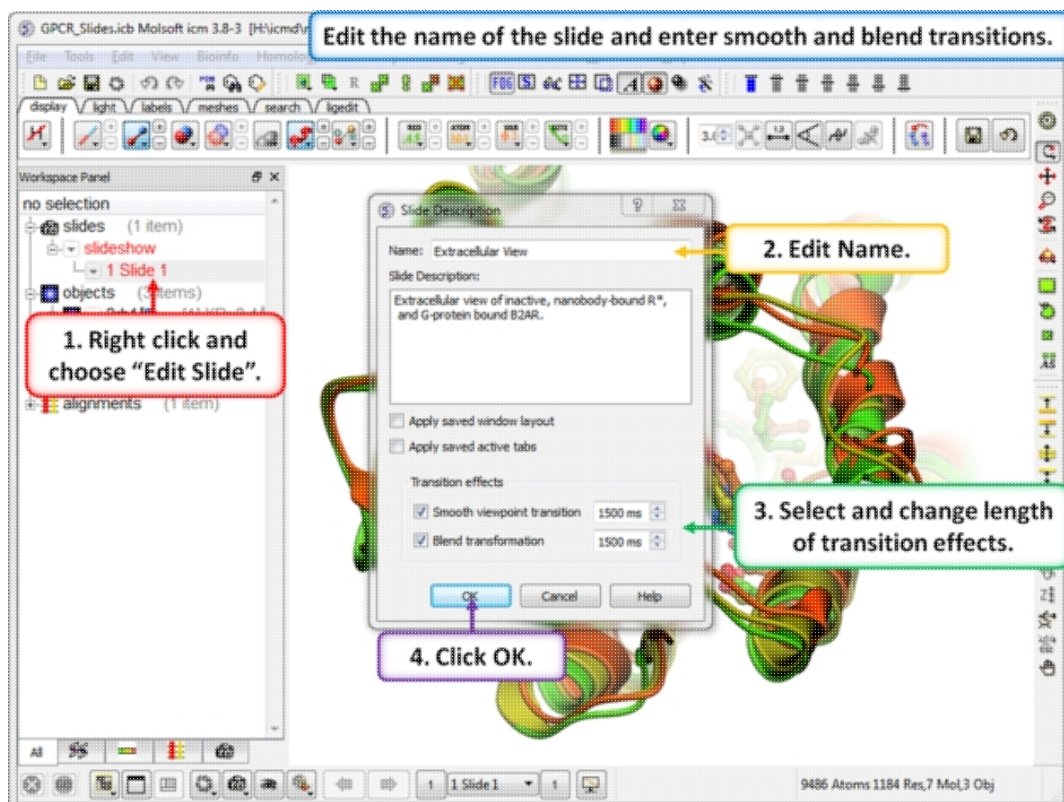
In this tutorial we will build a series of slides and then annotate them using HTML to create hyperlinks to the slides. The resulting molecular document can be exported in movie format or in fully interactive 3D files.



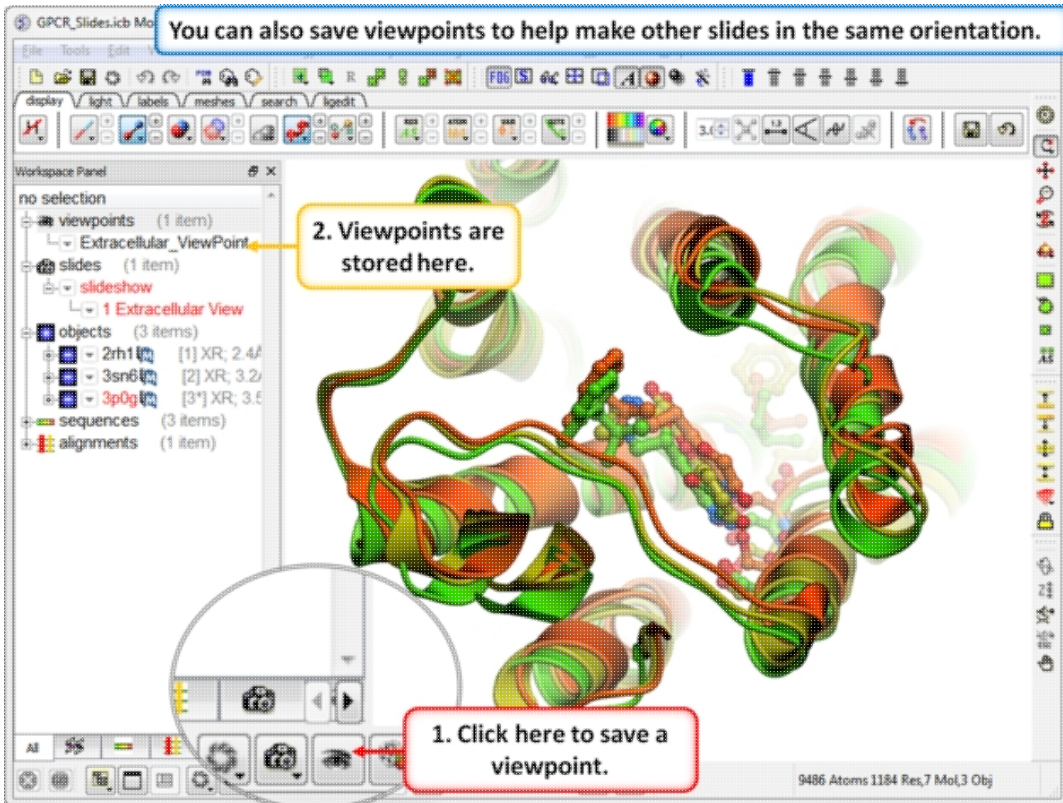
Step 1: Download the example file here <ftp://ftp.molsoft.com/pub/workshop/> (1. cut and paste the address into a web browser. 2. Right click on GPCR_slides.icb and choose "Save Link As". 3. Open the .icb in ICM File/Open). Make a slide by clicking on the camera button at the bottom of the GUI.



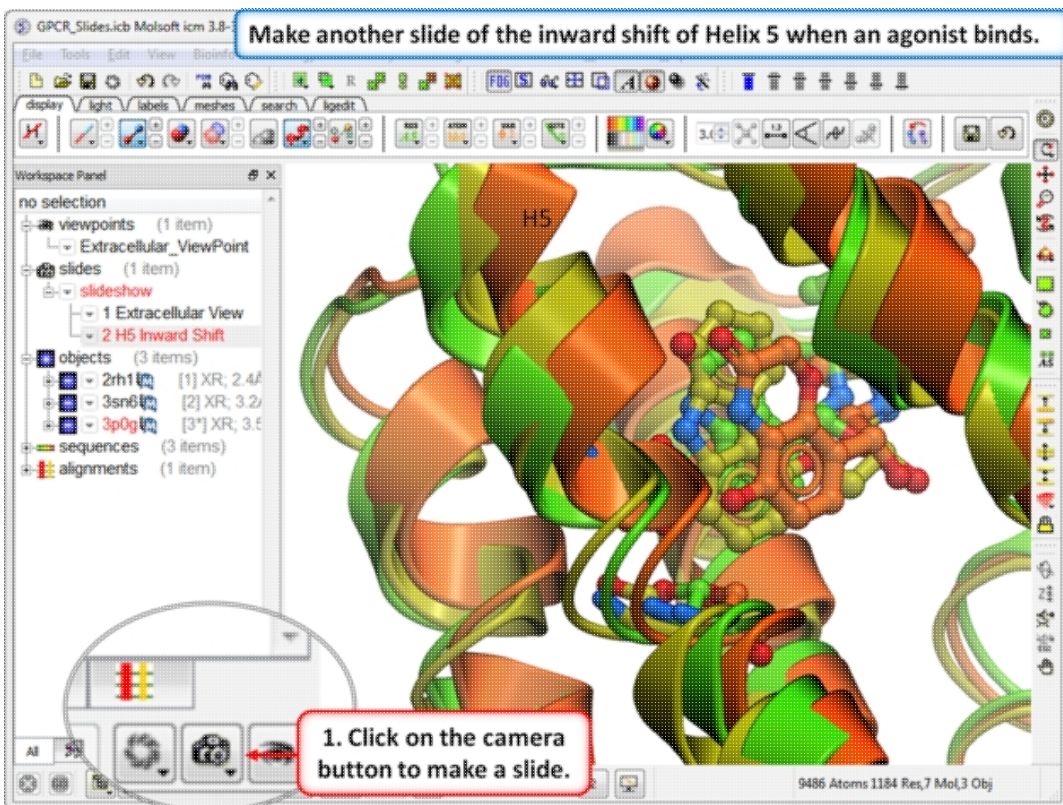
Step 2: The slide is stored in the ICM Workspace panel (left hand side of the GUI). You can make changes to the 3D display and return to the display in the slide by clicking on the slide in the ICM Workspace.



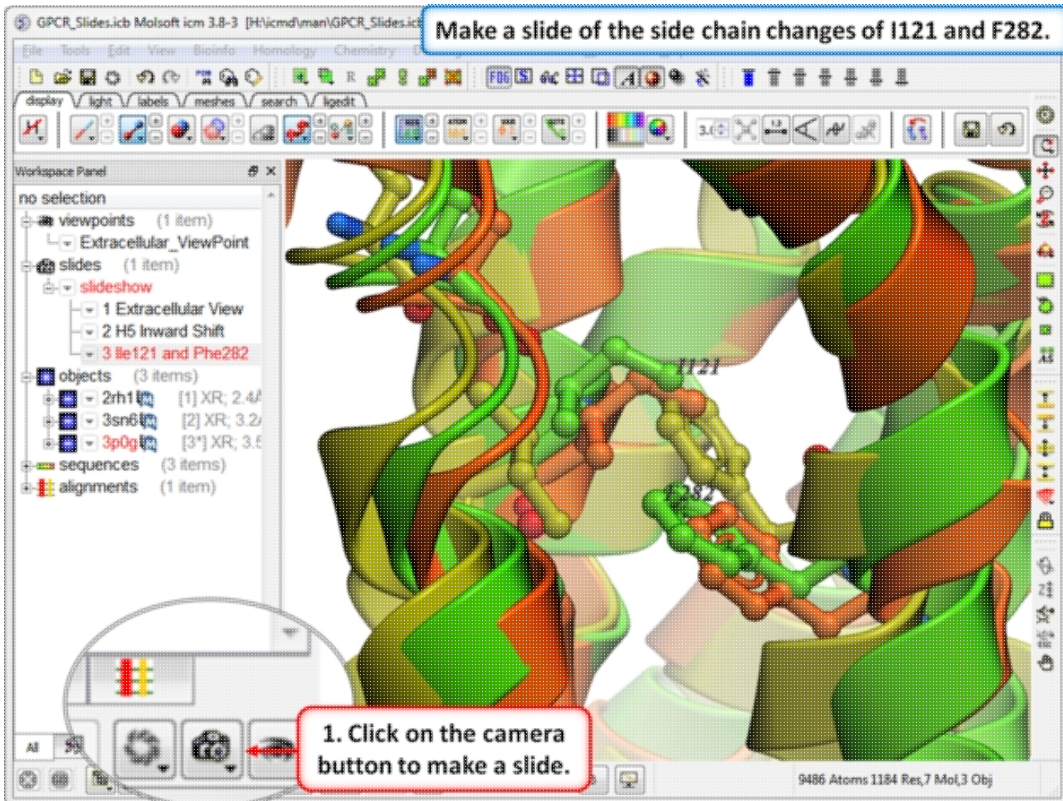
Step 3: Properties of the slide can be edited by right clicking on the slide in the ICM workspace.



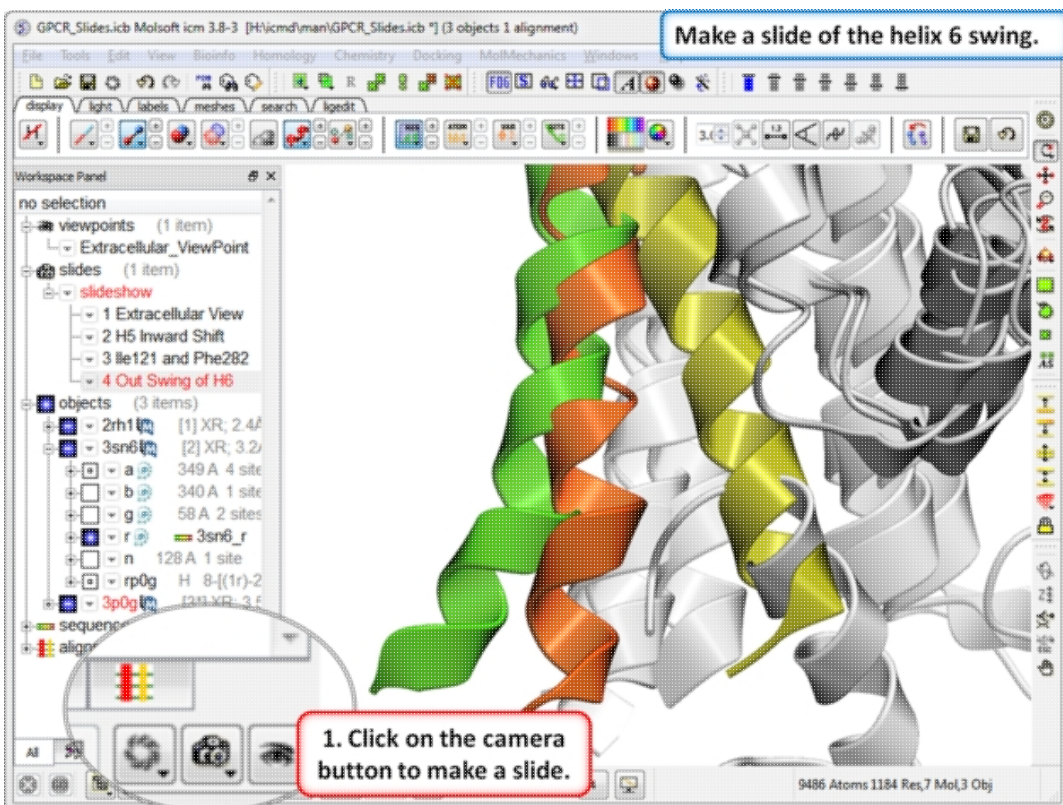
Step 4: Saving a viewpoint is useful if you are making many slides from the same viewing position.



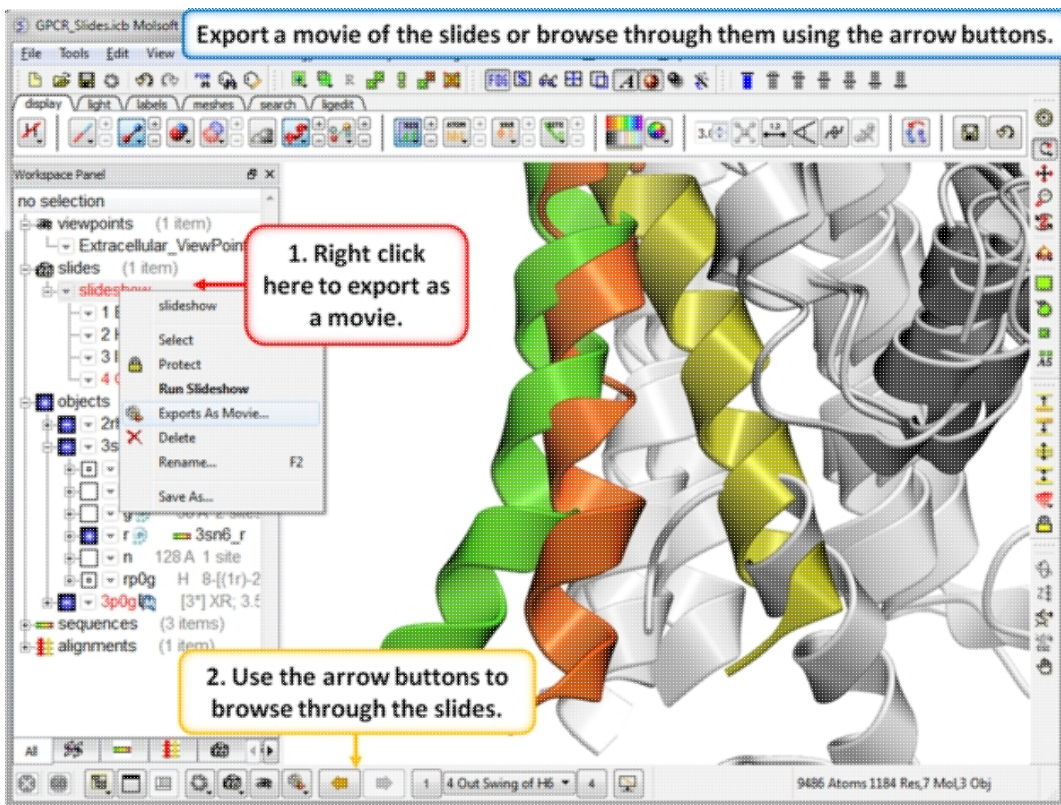
Step 5: In this slide we show that Helix 5 makes an inward shift when an agonist binds.



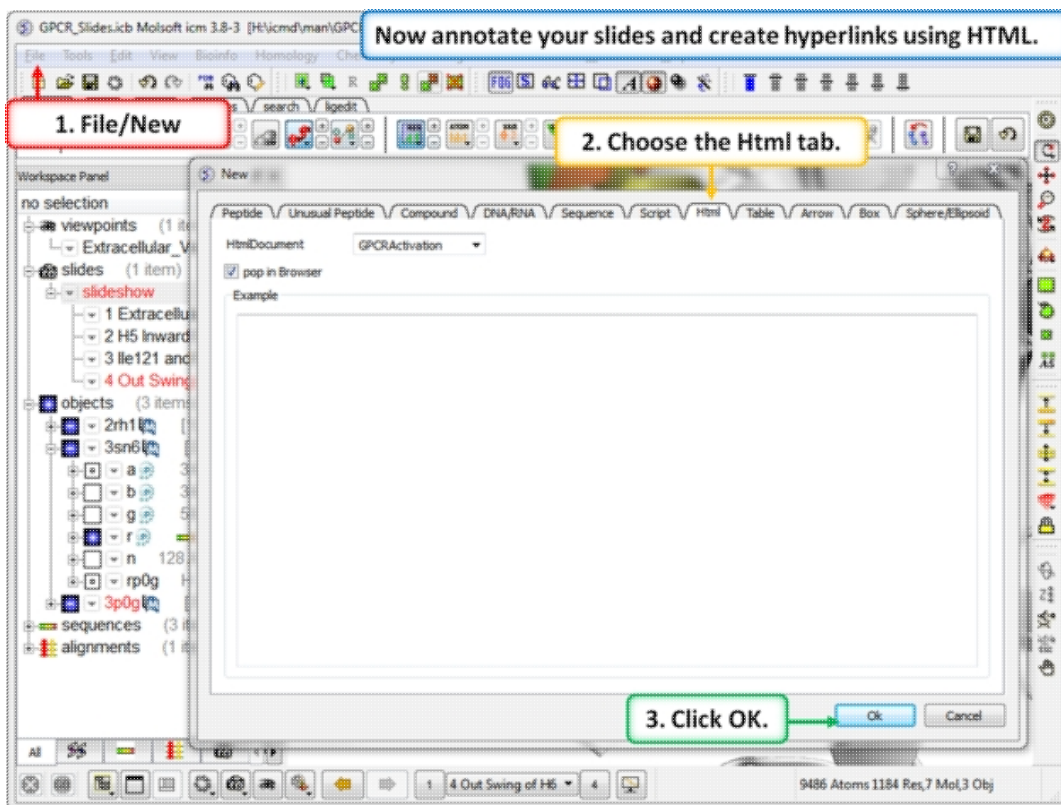
Step 6: In this slide we show that upon binding an agonist Ile121 switches its rotameric state and Phe 282 moves.



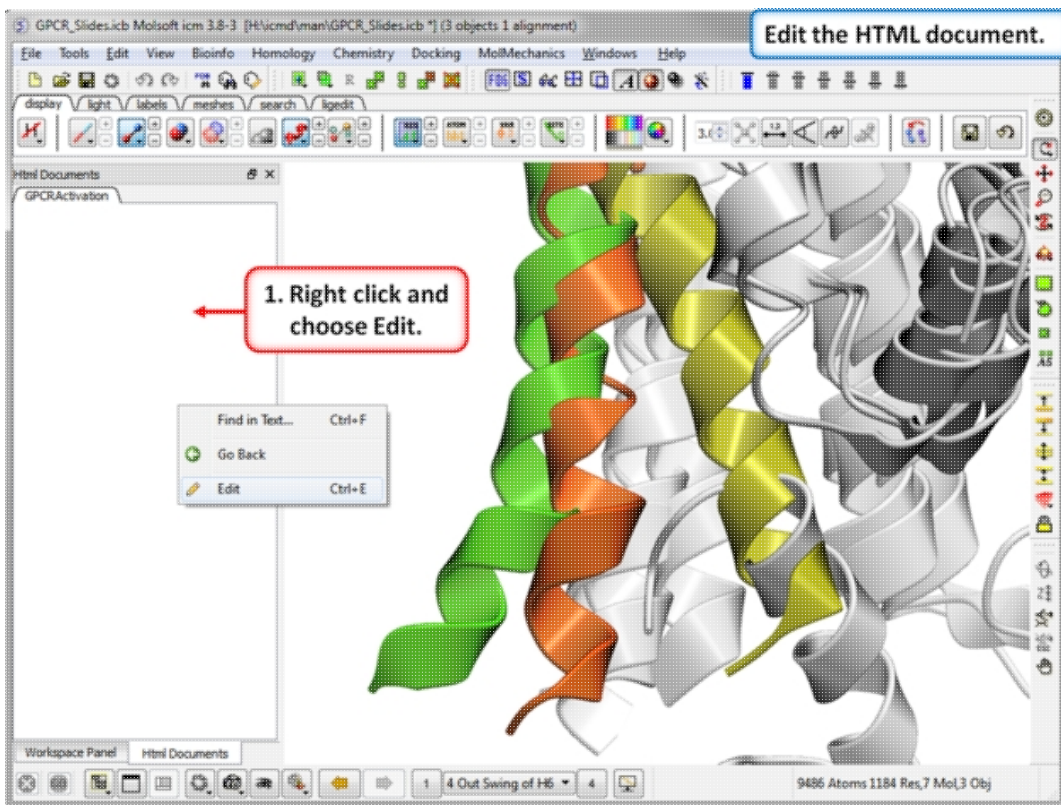
Step 7: In this slide we show that agonist binding causes Helix 6 to swing out.



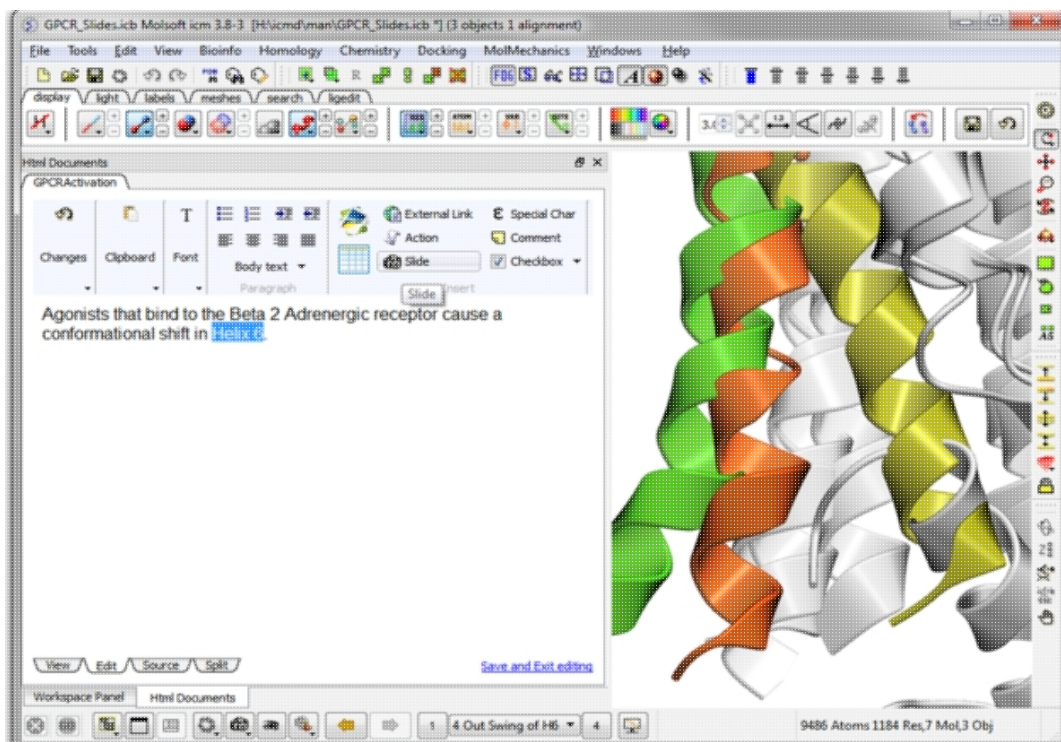
Step 8: You can export the slides as a movie or you can browse through them using the arrow buttons.



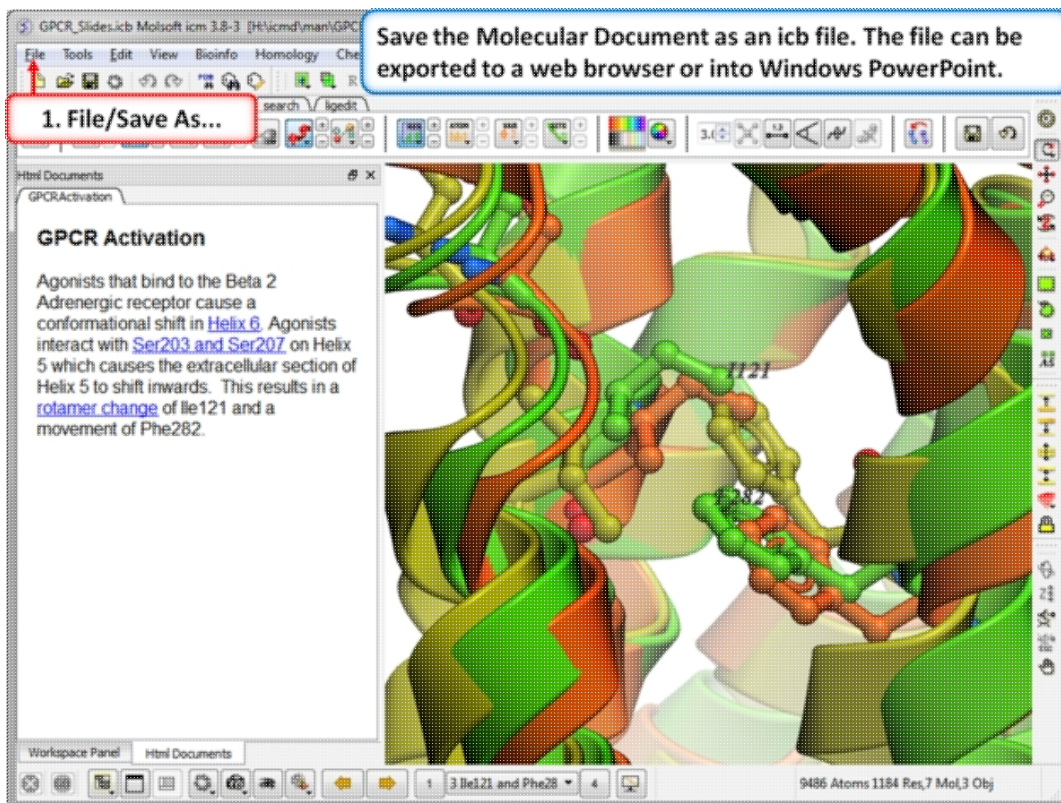
Step 9: To build a molecular document you can annotate the slides using HTML.



Step 10: Edit the document by right clicking on the HTML panel and choosing **Edit**.



Step 11: Hyperlinks can be made to the slide. Other functions such as adding check boxes, ICM scripts and external links are also available.

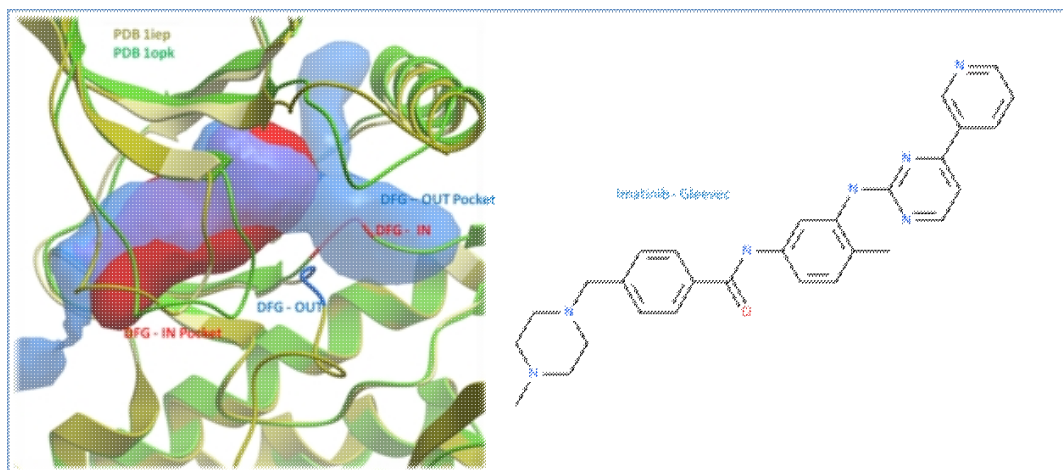


Step 12: Save the molecular document in icb format and view in ICM-Pro or ICM-Browser. You can also export to a web browser or view the fully interactive slides in Windows Powerpoint using a plugin called ActiveICM.

19.3 Protein Structure Tutorials

Introduction

The positioning of the DFG motif in the activation loop of a protein kinase has an effect on the size and properties of the ATP binding pocket. Most kinase inhibitors target the kinase with the DFG inwards to the ATP binding site. Type-II inhibitors target the site where the DFG motif is in the out position which opens up the pocket and provides additional hydrophobic binding sites. Targeting DFG-out conformations can improve inhibitor specificity and slower-off rates.

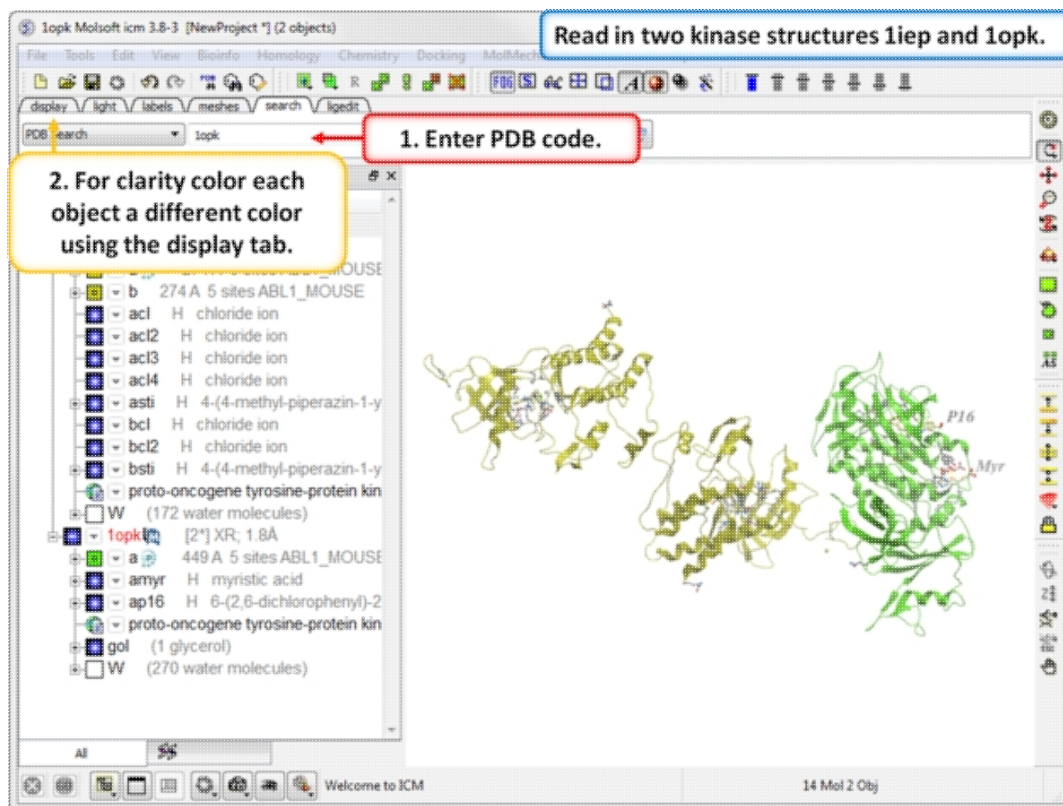


Comparison of the crystal structures of Abl kinase with DFG-IN and DFG-OUT (Left). Imatinib (Gleevec - right) binds to the DFG-OUT conformation.

Tutorial

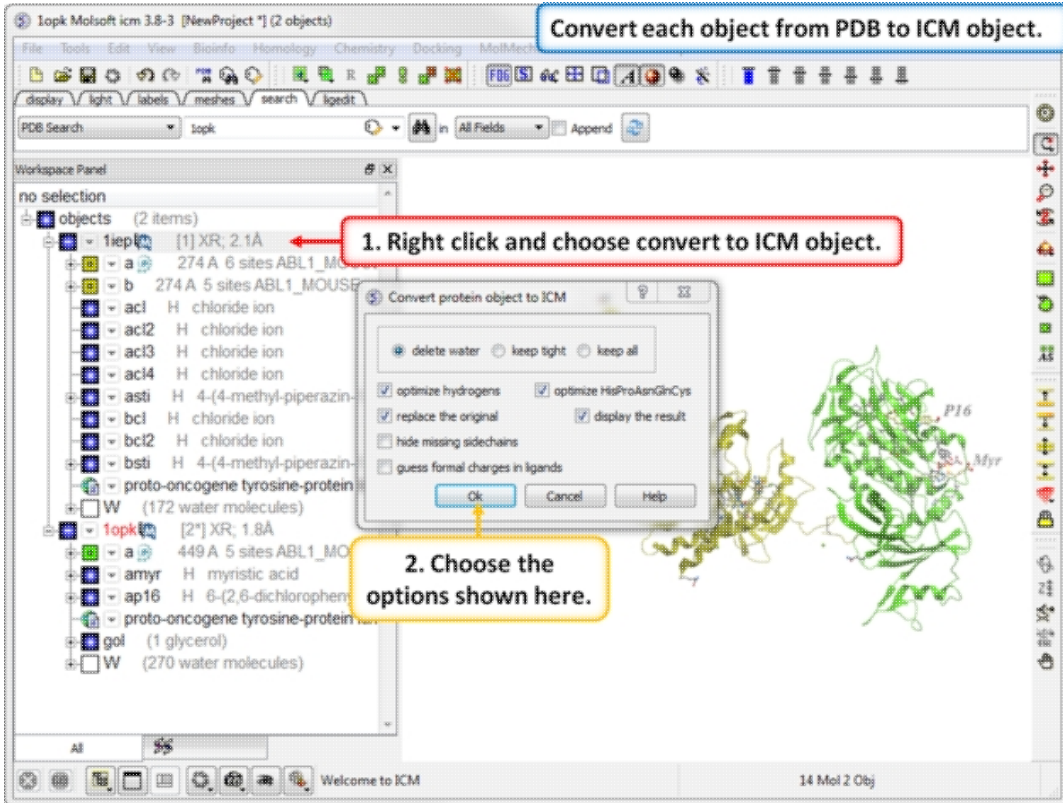
In this example we will look at two kinase structures and compare their binding pockets. One of the structures is bound to the Type-II inhibitor Imatinib (Gleevec) a drug used to treat Chronic Myelogenous Leukemia. We will use ICM tools to look at the side-chains that interact with the inhibitor and compare the pocket with a DFG-in structure using `ICMPocketFinder`. Please click on the links provided to learn more about the methods used.

19.3.1 PDB Search



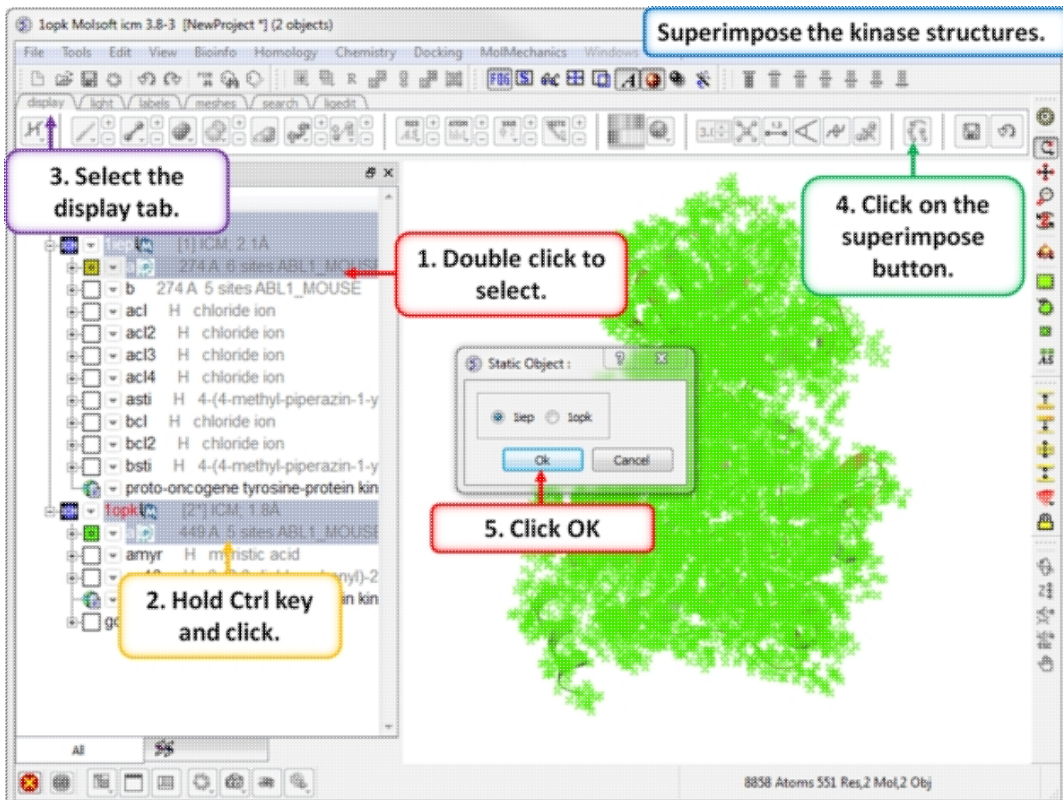
Step 1: Click on the `Search` Tab and choose "PDB Search" and load the PDBs structures 1iep and 1opk. You can color each object a unique color by clicking on the `Display` Tab and clicking and holding on the `Color by>Object`.

19.3.2 Convert PDB to ICM Object



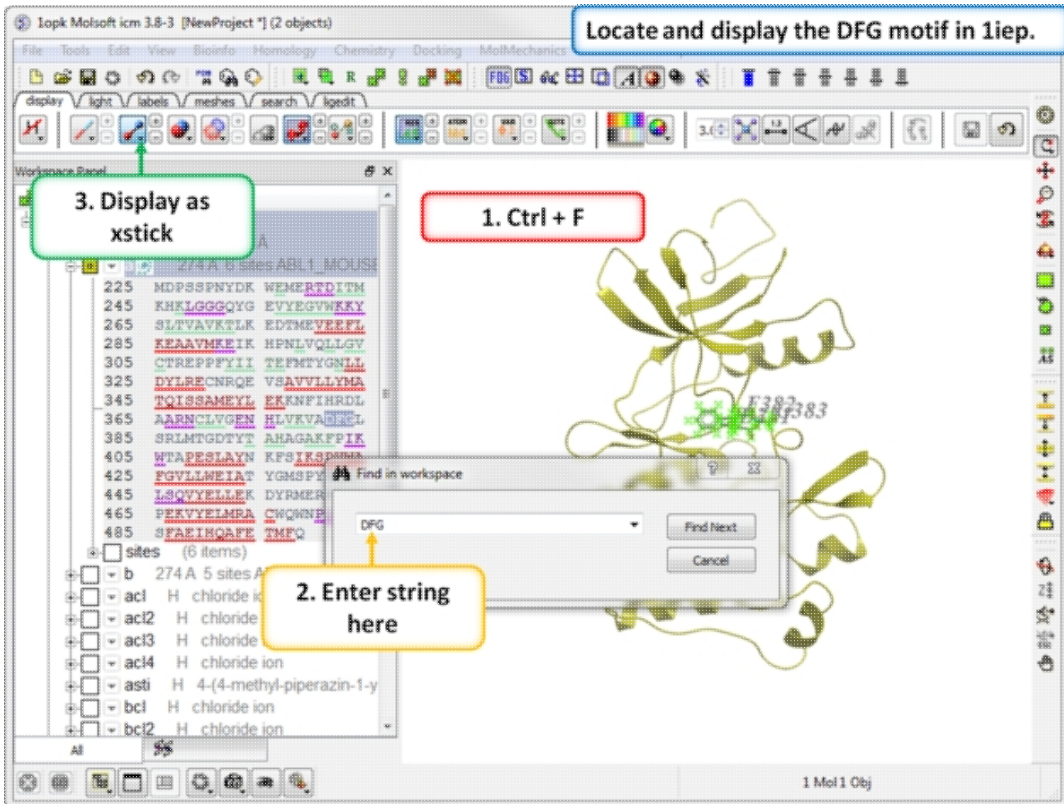
Step 2: Convert each PDB file into an ICM object.

19.3.3 Protein Superposition

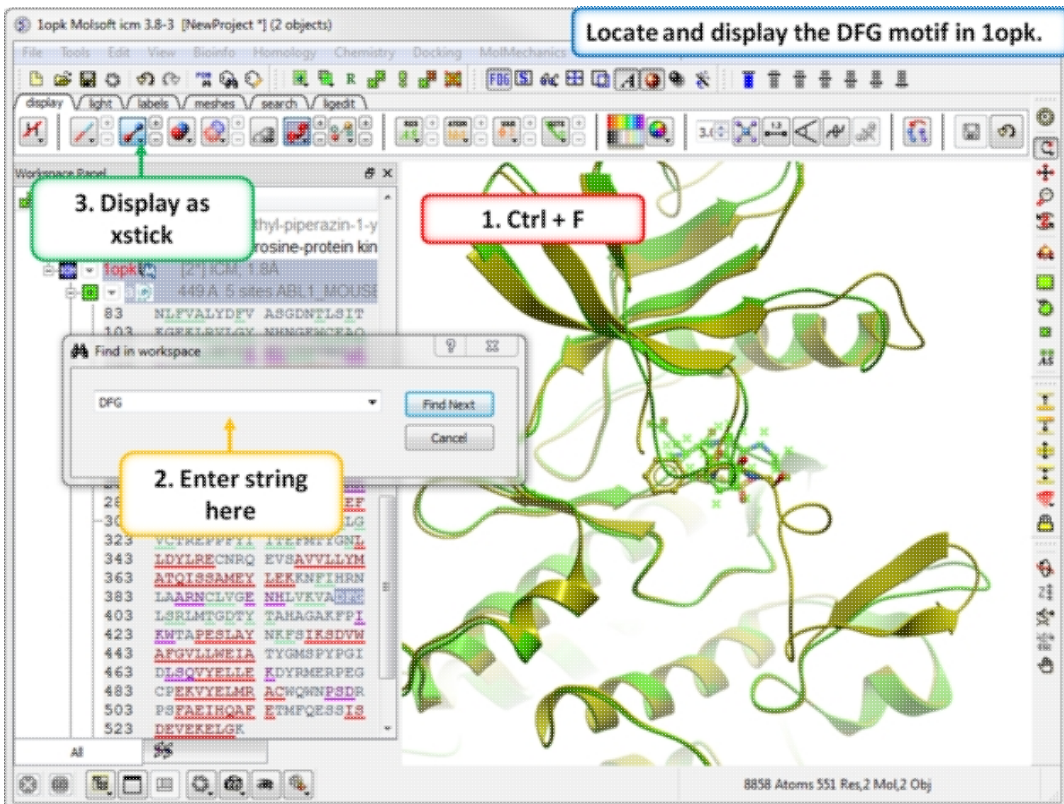


Step 3: Superimpose the two kinase structures.

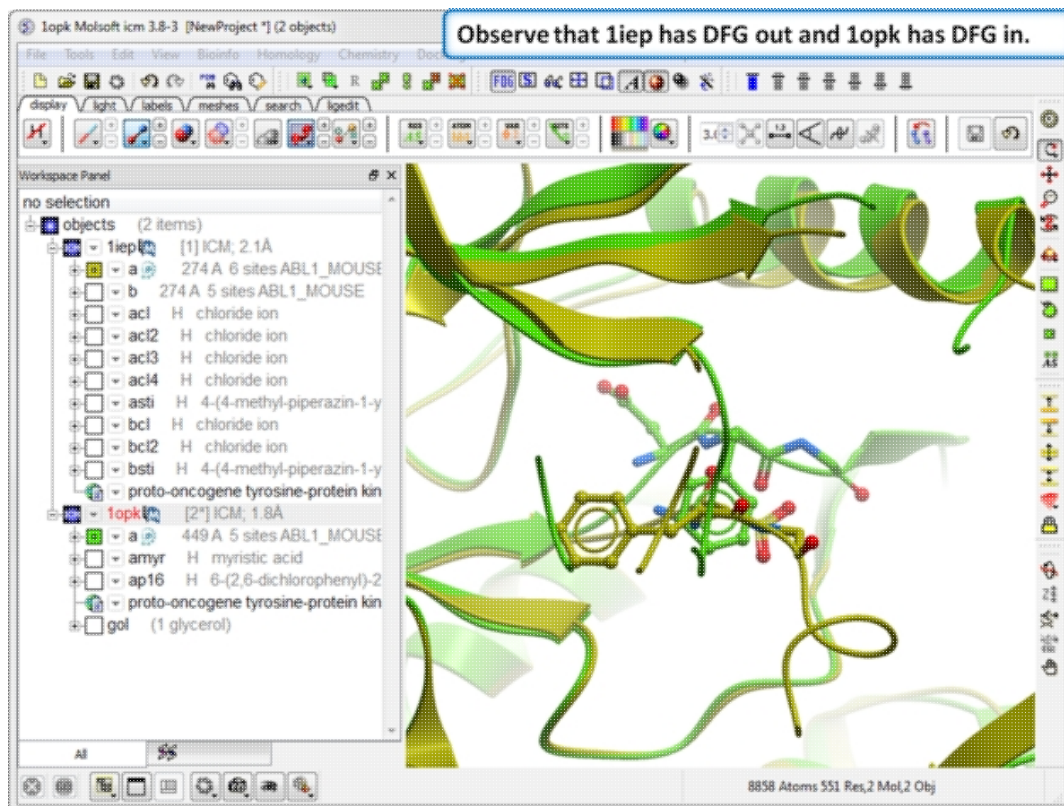
19.3.4 Protein Structure Analysis



Step 4: Search the "a" subunit of 1iep for the "DFG" motif. You can press Ctrl F and expand the sequence in the ICM workspace. Display the DFG region in xstick.

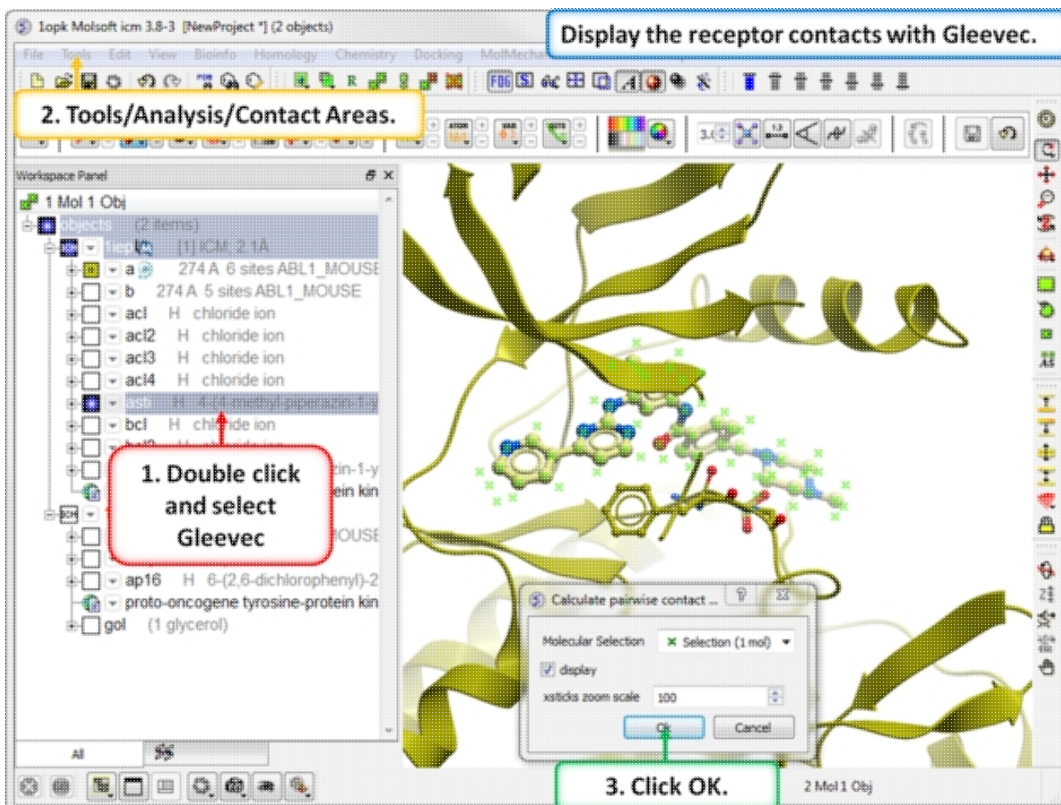


Step 5: Search the "a" subunit of 1opk for the "DFG" motif. You can press Ctrl F and expand the sequence in the ICM workspace. Display the DFG region in xstick.

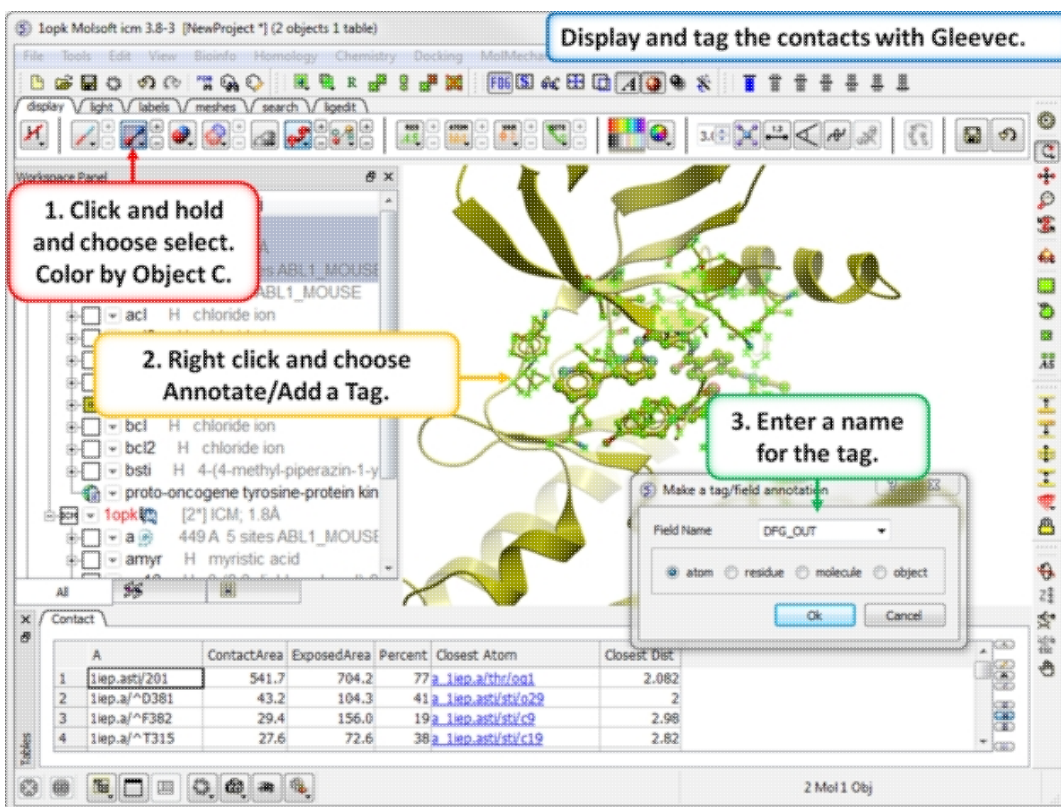


Step 6: Observe the DFG location in both structures.

19.3.5 Contact Areas

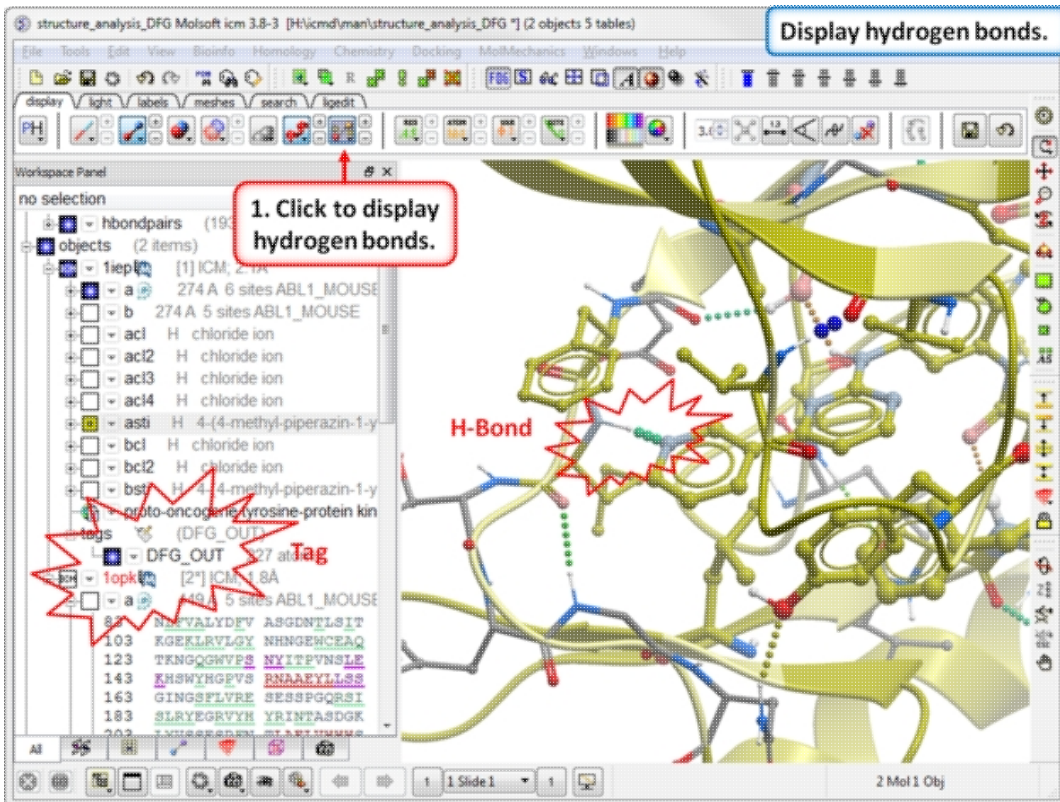


Step 7: Use Tools/Analysis/Contact Areas to display the contacts with the Gleevec ligand.



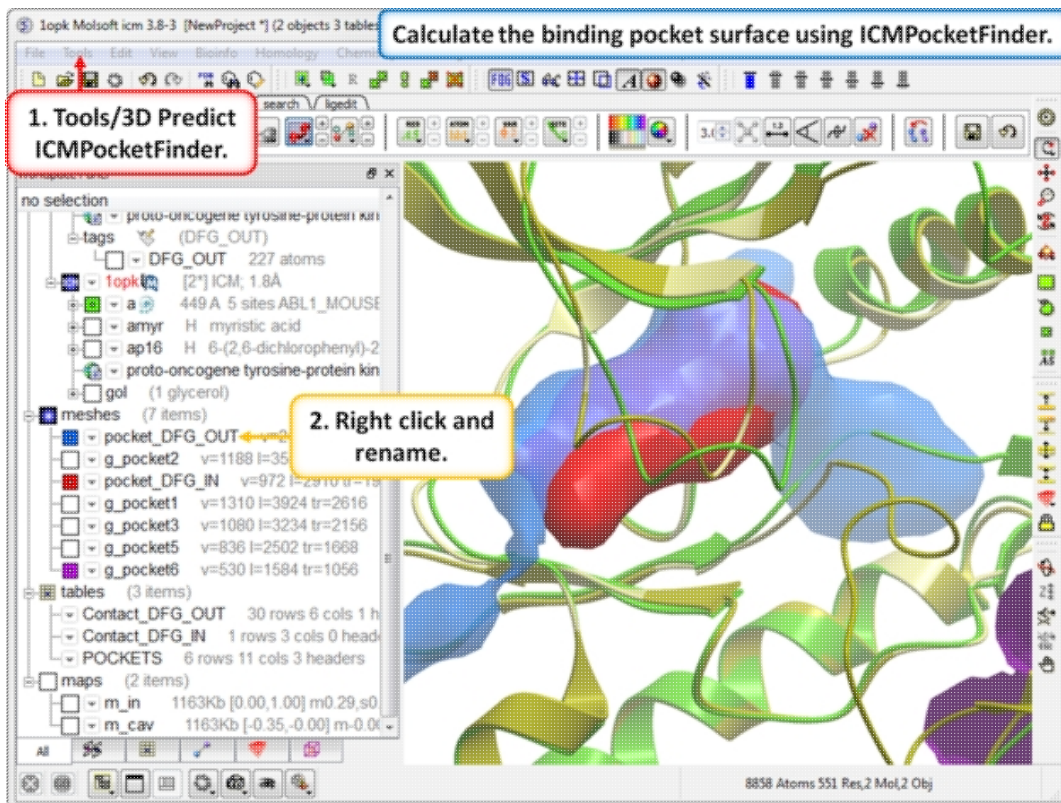
Step 8: Tag the residues surrounding the ligand. Repeat Steps 7 and 8 for the 10pk structure and ligand.

19.3.6 Hydrogen Bond



Step 9: Compare the hydrogen bond interaction of the Type I and Type II inhibitors with the hinge region (1lep - residues 317:322) of the kinase.

19.3.7 Pocket Identification



Step 11: Use ICMPocketFinder to compare the pockets of DFG-IN and DFG-OUT. Run ICMPocketFinder for each structure - rename the pocket so as not to overwrite it.

19.3.8 Protein Preparation and Crystallographic Analysis Tutorial

19.3.9 PDB Preparation - Symmetry

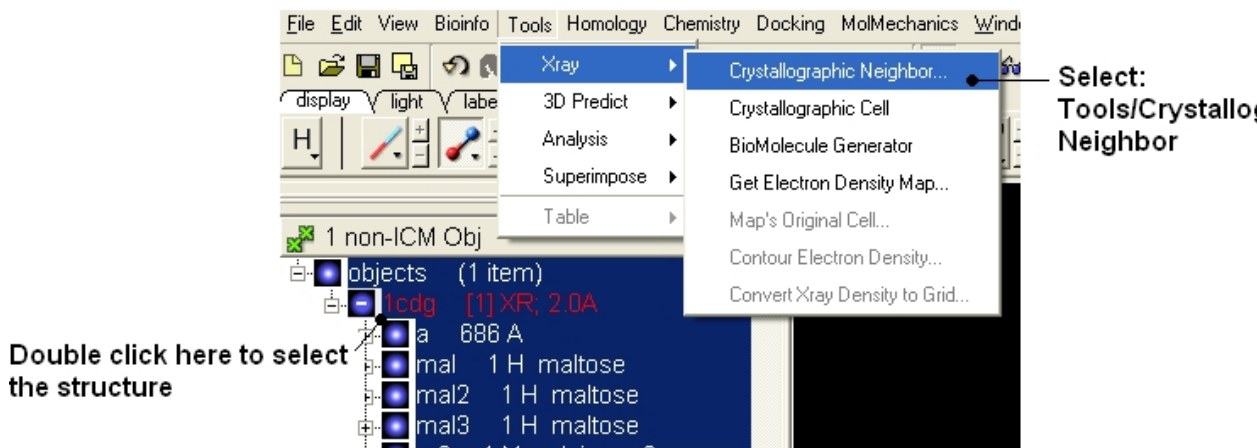
Background When inspecting a ligand binding pocket it is important to check that the true pocket is formed by chains which are not explicitly present in a PDB entry. Therefore it is necessary to use **Tools/X Ray/Crystallographic Neighbor** to find all molecules/subunits or chains involved in the interaction with the ligand. Molecular objects and 3D density maps may contain information about crystallographic symmetry. It consists of the following parameters:

1. Crystallographic group eg. P2121 that determine N (depends on a group) transformations for the atoms in the asymmetric unit.
2. Crystallographic cell parameters A, B, C, Alpha, Beta and Gamma

To generate the coordinates within one cell one needs to apply N transformations and then to generate neighboring cells the content of one cell needs to be translated in space according to the cell position.

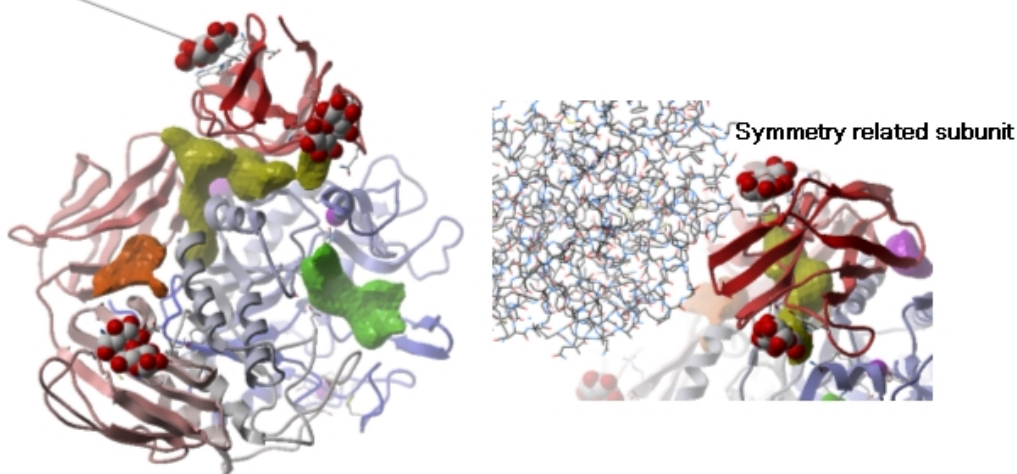
Example As an example let us look at Cyclodextrin glycosyltransferase (PDB Code: 1CDG). The problem with docking to this receptor is that the true pocket is formed by chains which are not explicitly present in the PDB entry. Site mb1 includes serine 382. This cannot be predicted just by looking at the structure. Therefore we need to identify symmetry related molecules to this protein.

- Use the PDB search tab to load the crystal structure 1cdg.
- Inspect the ligand binding pocket of **maltose** (mal)
- To identify if there are any other chains involved in the interaction with the ligand select the whole structure in the ICM Workspace.



- Tools/Crystallographic Neighbor
- Select a 7Å radius
- Check "create symmetry related molecules" and "display symmetry neighbors".
- Inspect the neighbors surrounding **maltose(mal)**. Each symmetry related subunit can be colored by object by clicking and holding the representation button in the display tab and selecting color-by.

site mb1 includes residue ser 382 for symmetry-related molecule. site mb3 includes the following residues for symmet



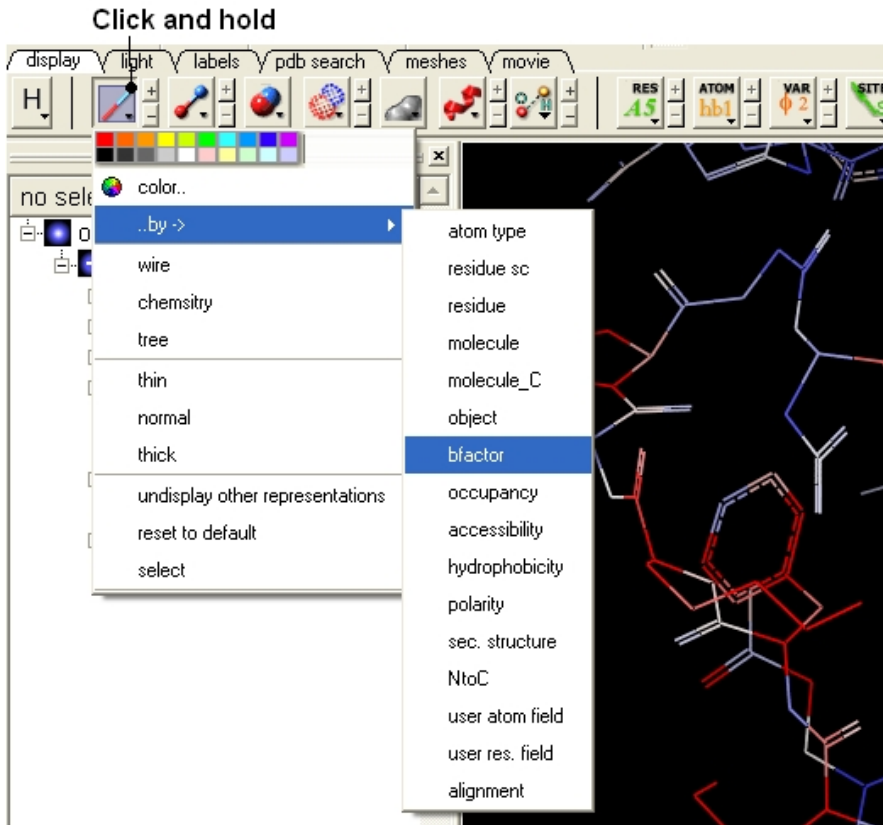
19.3.10 PDB Preparation - Occupancy and B-Factors

Background When preparing a PDB for analysis (eg docking or modeling) it is important to check the reported occupancies and b-factors. The occupancy is a fraction of atomic density at a given center. If there are two equally occupied conformers both will have an occupancy of 0.5 - the normal value is 1 range 0-1. The $\{B\text{-Factor}\}$ is the mean-square displacement of atom from its position in the model - the normal range is 5-50.

One way of visualizing the occupancy and b-factor is by coloring the structure by these values. You can do this by clicking and holding on a representation button in the **display** panel and selecting Color-by.

As an example let us look at the crystal structure 1ATP

- Type in the PDB search tab 1atp and the structure will be displayed in the graphical display.
- Use the ICM workspace to undisplay everything except for the "e" subunit. You can do this by clicking in the blue boxes in the ICM Workspace.
- Display the "e" subunit in wire representation using the wire button in the **display** tab.
- Click and hold on the wire button and select Color-by B-Factor. Regions of high B-factor are colored red.



Residues with high B-Factor are colored red

19.3.11 PDB Preparation - Residue Alternative Orientation

For some very high resolution structures two alternative conformations for a residue are provided. Therefore for docking you need to decide to use one conformation of the residue or generate several separate docking models. This could be performed using multiple receptor conformation docking.

Here is an example of alternative residue orientations found in a crystal structure of a Fatty Acid Binding protein in complex with stearic acid.

Load and display the PDB structure 1HMT.

1. PDB search 1HMT

2. Undisplay xstick, residue and variable labels

```

[com/daZ] undisplay window
[com/daZ] _m_moltenProject = ""
[com/daZ] generalSearchFromDB "PDB Search" "1hmt"
[com/1hmt] undisplay xstick Res(n_"/10)
[com/1hmt] undisplay residue label Res(n_"/10)
[com/1hmt] undisplay site Res(n_"/10)
[com/1hmt]

```

Display alternative conformations for V32 and T36.

1. Click here to expand residue level

2. Hold the control key and click and drag over V32 and T36

3. Click the xstick button

4. Click center button

```

[com/1hmt] set stick residue label size=10
[com/1hmt] set stick residue label size=10
[com/1hmt] set stick residue label size=10
[com/1hmt] set stick residue label size=10
[com/1hmt] set stick residue label size=10
[com/1hmt] set stick residue label size=10
[com/1hmt]

```

19.3.12 Biomolecule Generator

Objective

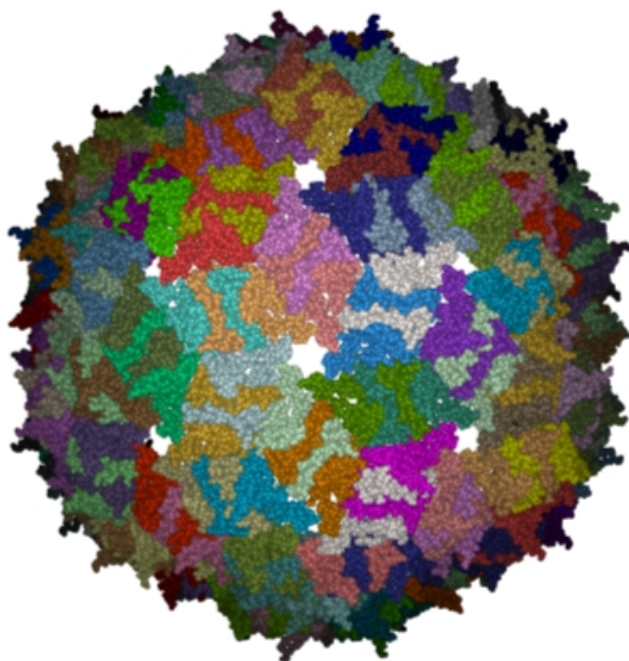
Here we will investigate the biological environment of a virus protein . PDB code 1DWN.

Background

It is very useful to know how a protein from the PDB may look in a biological environment. The PDB entries solved by X-ray crystallography and deposited in the PDB contain the information about the crystal structure rather than the biologically relevant structure. For example, for a viral capsid only one instance of capsid protein complex will be deposited and only one or two molecules of haemoglobin that is a tetramer in solution maybe deposited. In some other cases the asymmetric unit may contain more than one copy of a biologically monomeric protein. ICM reads the biological unit information and has a tool to generate a biological unit. Not every PDB entry has the biological unit information.

Instructions

- Read and load the PDB file 1DWN
- Tools/Xray/Biomolecule Generator
- Tick the makeAllBiomolecules box (Warning this may take a few minutes to generate)
- The generated molecules will be listed in the ICM Workspace. Each one can be selected and displayed. The biomolecule is shown below.



NOTE: Please note that right clicking on a PDB file in the ICM Workspace will tell you whether there is any **Biomolecule** information available for the structure. If this information is not present then the option will be greyed out.

Manual References (Web Links)

Biomolecule Generator

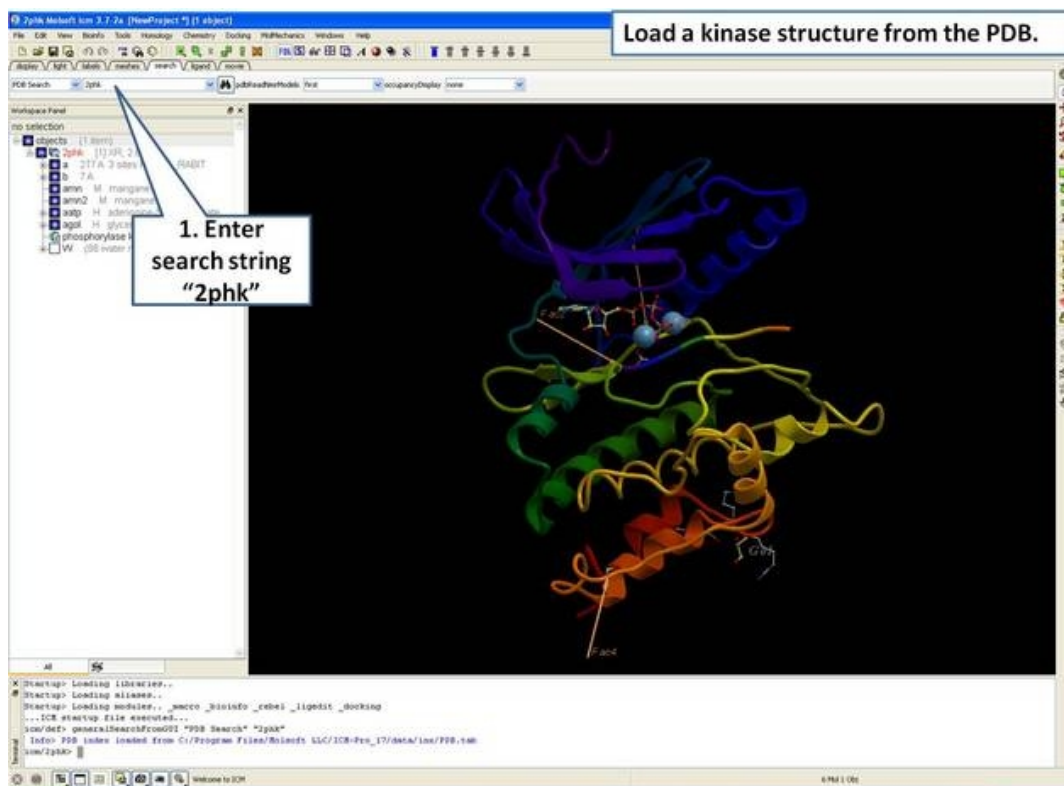
19.4 Sequence and Alignment Tutorials

The key topics in this tutorial are:

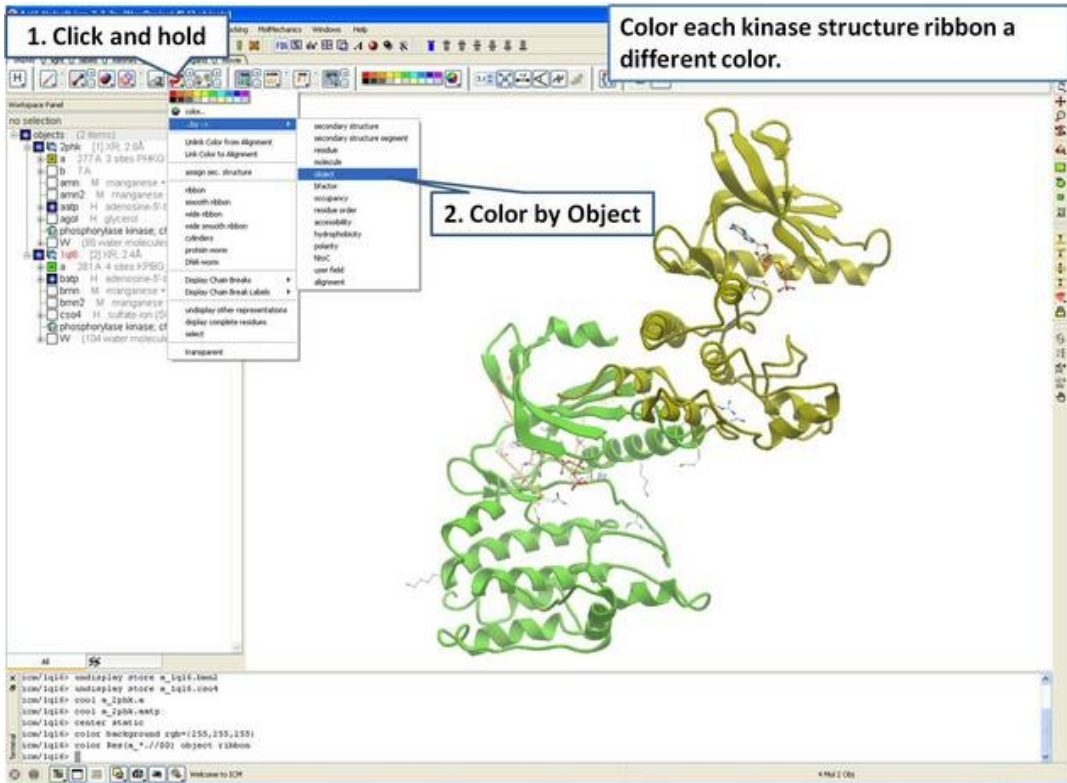
- How to extract sequences from PDB structures.
- How to load sequences from UniProt.
- How to make an alignment.
- How to link the alignment to 3D structure.
- How to display sequence conservation in ligand binding pocket.
- How to color 3D structure by alignment.

19.4.1 Display and Superimpose Protein Kinase Structures in Preparation for the Tutorial

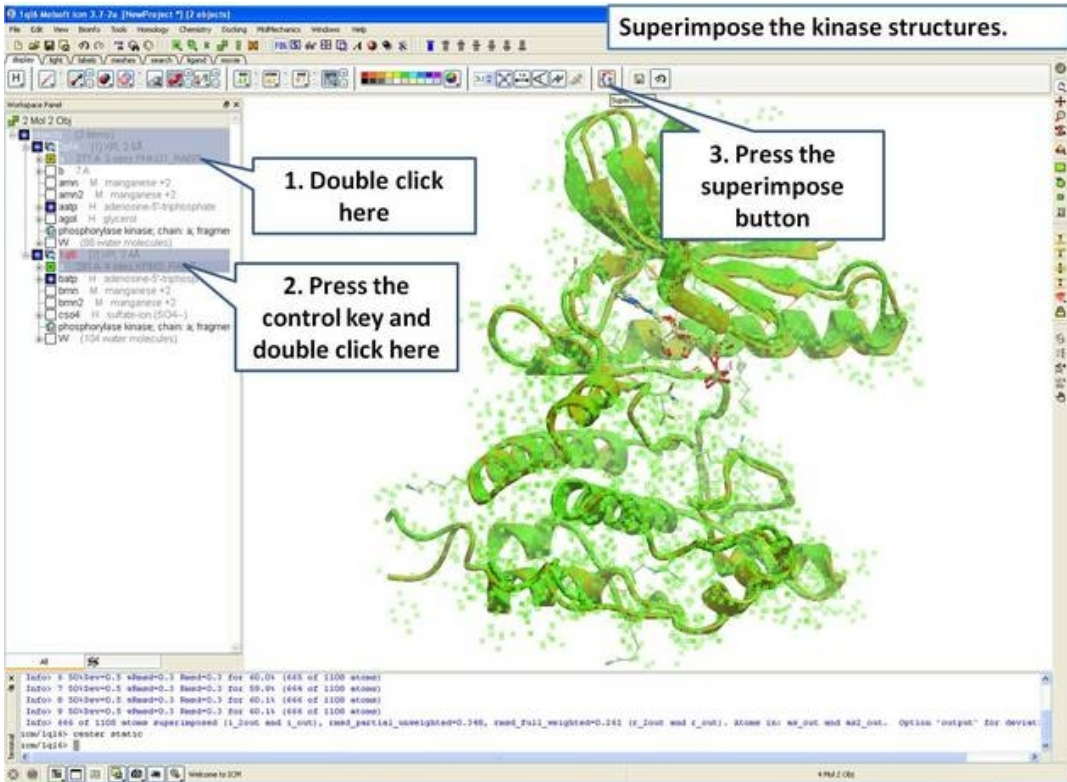
In preparation for the tutorial we will read in two PDB file and superimpose them. We will then extract sequences from the PDB structures and read in additional kinase sequences from Uniprot.



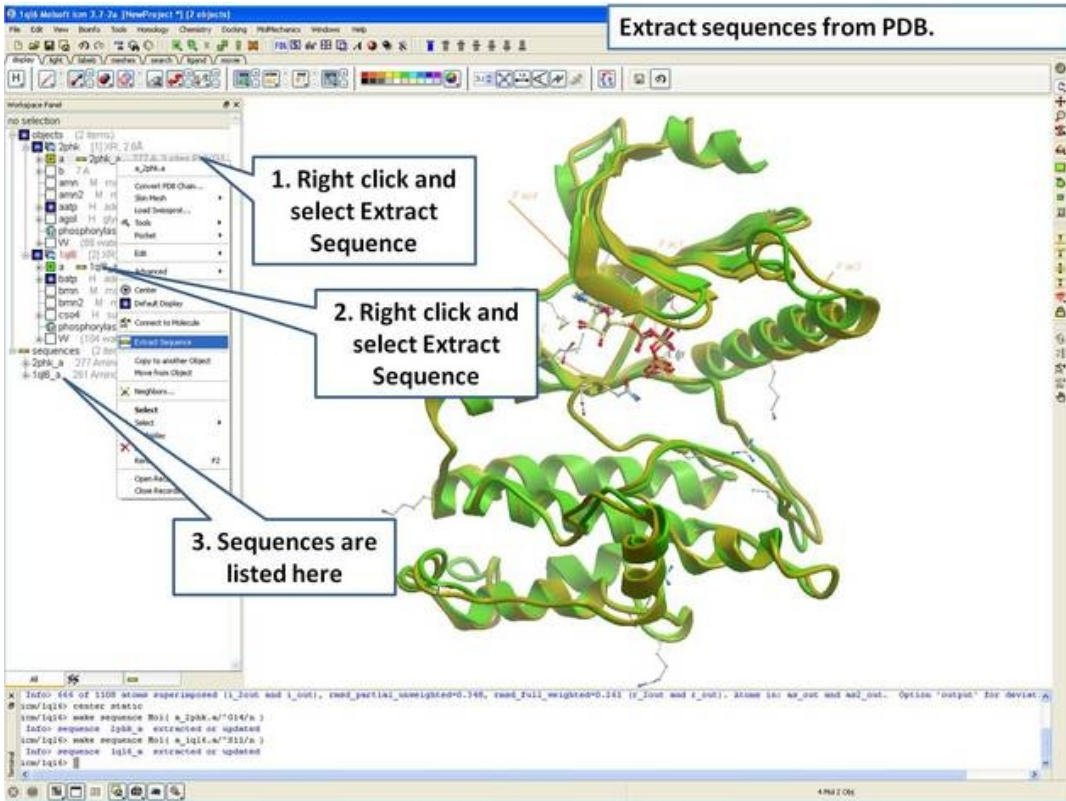
Step 1: Load and display the kinase structure 2PHK.



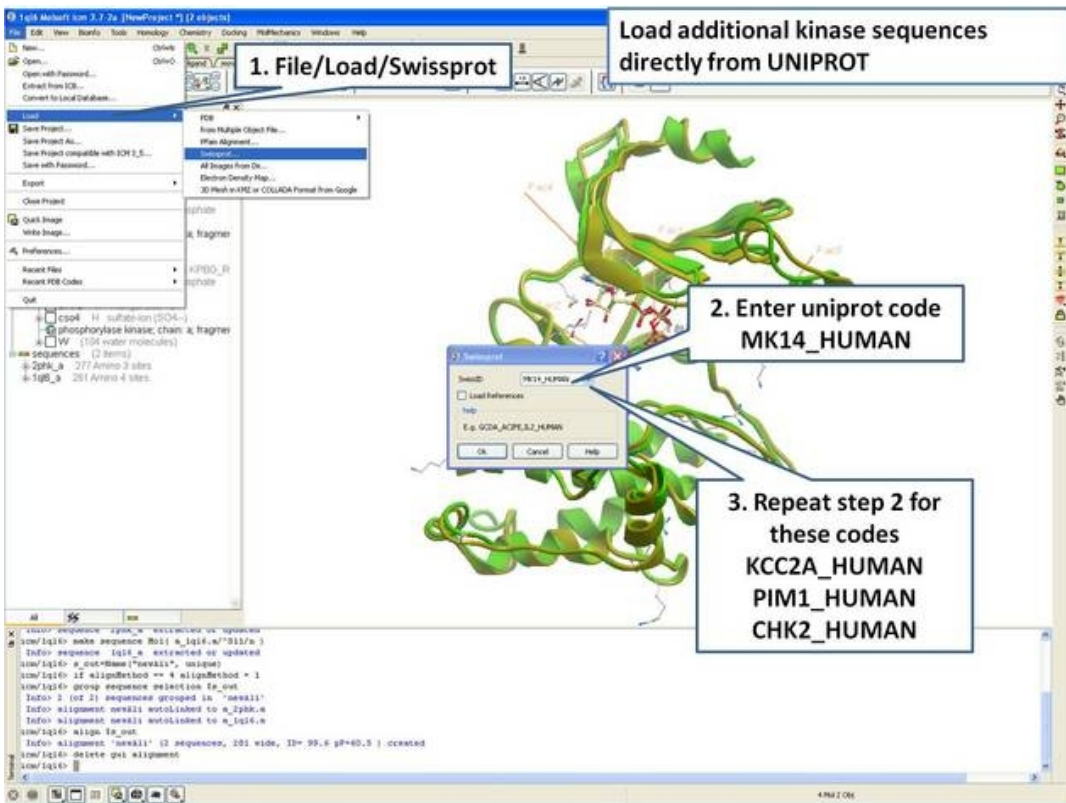
Step 4: For clarity, color each receptor a different color.



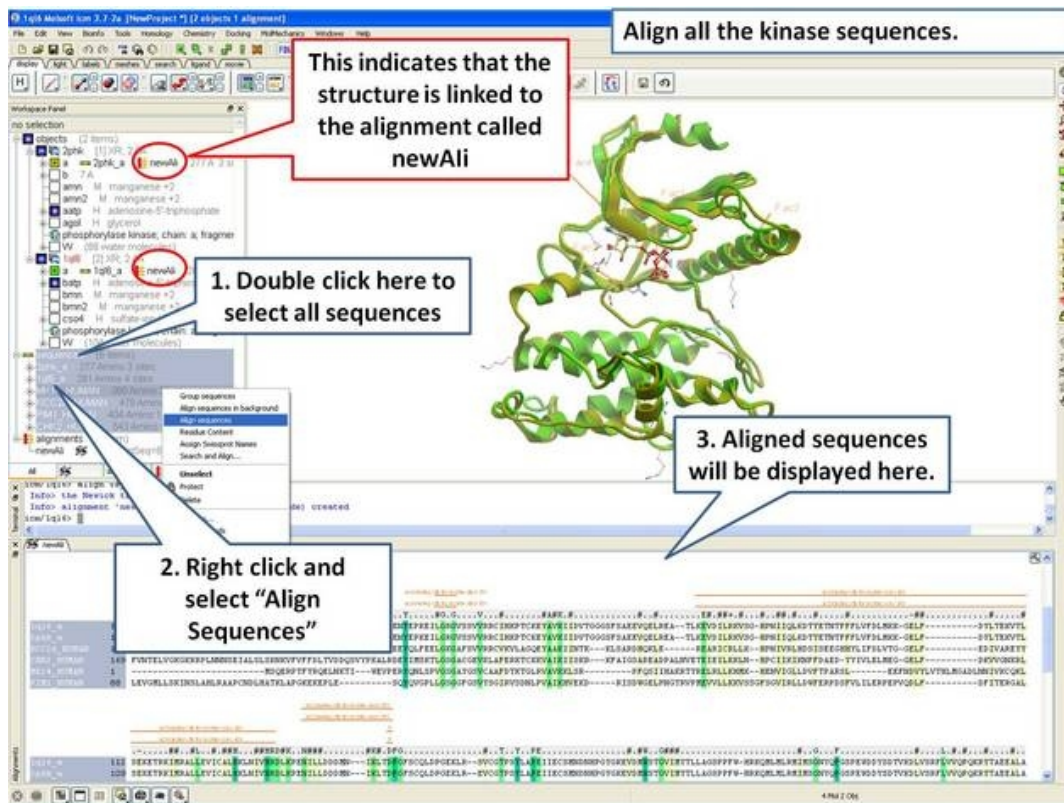
Step 5: Superimpose both kinase structures using the superimpose tab.



Step 6: Extract the sequences from the PDB structure.



Step 7: Load UniProt sequences.



Step 8: Make a sequence alignment.

19.4.2 Linking Sequence Alignment to Structure

An alignment can be directly linked to the 3D structure. If the sequence has been extracted from a displayed PDB file this will be done automatically. If multiple alignments are loaded you can change the link using Bioinfo/Link to Structure... menu. This examples continues from the previous section.

Color alignment by sequence conservation.

1. Click here to see alignment tools panel.

2. Select Color:conservation.

Step 9: Color the alignment by sequence conservation.

1. Click and hold

Color structure ribbons by alignment. Dark green = fully conserved.

2. Color by alignment

Step 10: Color the 3D structure by the alignment color.

19.4.3 Identify Sequence Conservation in Ligand Binding Pocket

In this example we show how to display sequence conservation in the ligand binding pocket. This examples continues from the previous section.

The screenshot displays the MolScribe software interface. At the top, a menu bar includes File, Edit, View, Reports, Tools, Homology, Chemistry, Docking, Multitasking, Windows, and Help. Below the menu is a toolbar with various icons. The main window is divided into several panels:

- Left Panel:** A tree view showing the project structure, including sequences and alignments.
- Center Panel:** A 3D ribbon representation of a protein structure with an ATP ligand bound in a pocket. Green crosses are overlaid on the protein surface, indicating selected residues. A red circle highlights one of these crosses.
- Bottom Panel:** A sequence alignment window showing multiple protein sequences. A vertical column of green crosses is visible, corresponding to the residues selected in the 3D model. A red circle highlights a specific residue in this column.
- Right Panel:** A 'View options' panel with various checkboxes and settings for the visualization.

Annotations and callouts:

- 1. Right click on ATP ligand in the first object:** A callout box pointing to the ATP ligand in the 3D model.
- 2. Select/Neighbors:** A callout box pointing to the 'Select/Neighbors' option in the context menu.
- Green crosses = graphical selection:** A callout box pointing to the green crosses on the protein surface.
- Notice selection is returned to alignment – highlighted in blue:** A callout box pointing to the highlighted residue in the sequence alignment window.

Step 11: Select the residues surrounding the ligand binding pocket (or any other selection of interest).

Propagate selection to all sequences in alignment.

Notice selection is propagated to all sequences in the alignment – highlighted in blue

1. Propagate to all sequences

4 Nu 2 06

Step 12: The 3D structure and alignment are linked, therefore the selections made in 3D are transferred to the alignment. Only the sequences that have 3D structures will be selected in the alignment, therefore we need to propagate the selection to all sequences in the alignment.

Display only the residues in the ligand binding pocket.

Residue conservation in the ligand binding pocket

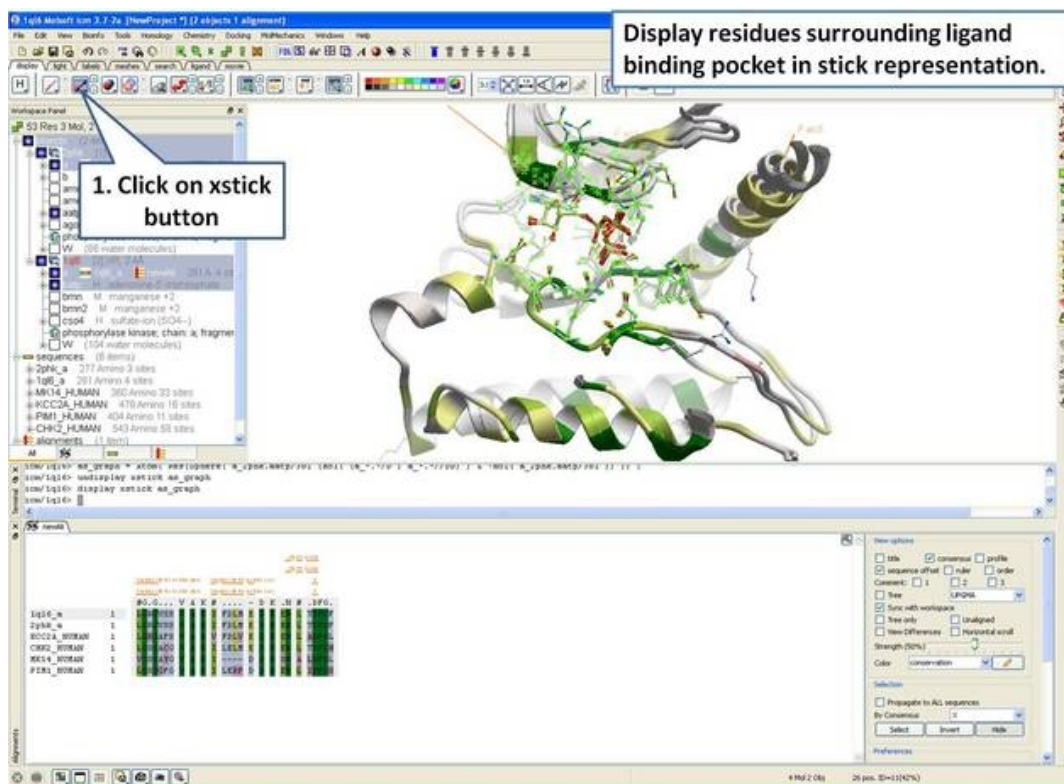
3. Click again to restore view

2. Hide selection

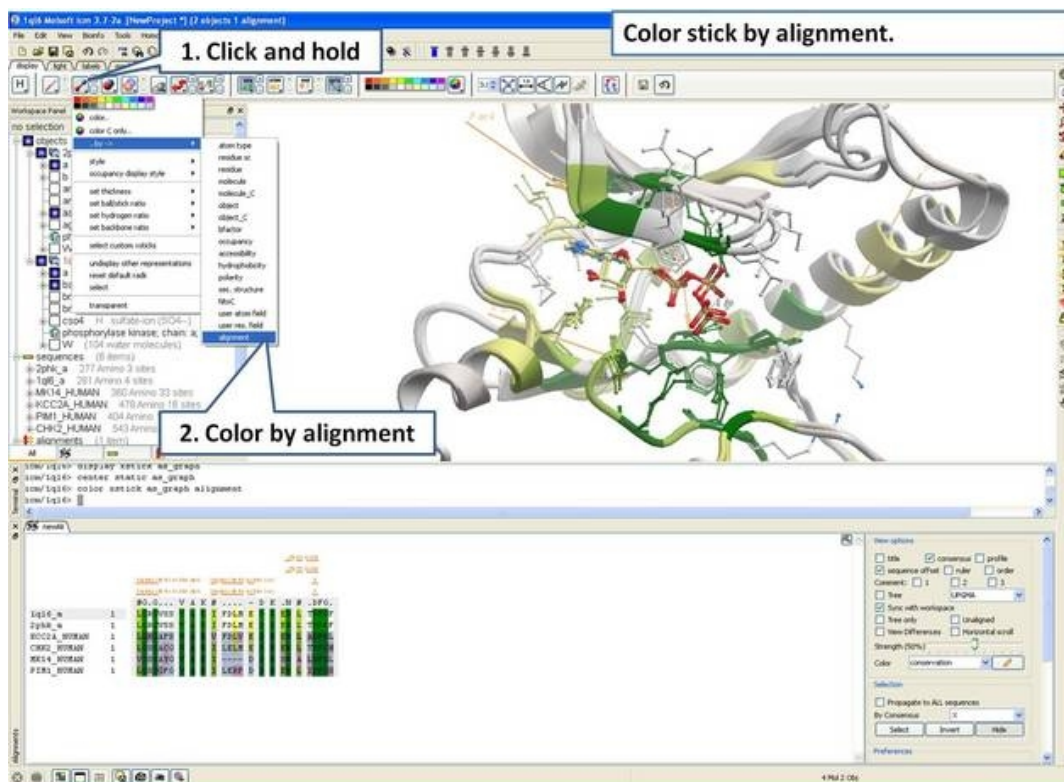
1. Invert selection

4 Nu 2 06

Step 13: Display only the residues in the pocket in the alignment.



Step 14: Display the residues surrounding the pocket in stick representation.



Step 15: Color stick by the alignment color.

19.4.4 Alignment Annotation

Introduction

G-Protein Coupled Receptors (GPCRs) all share a common structural core of seven transmembrane helices but they lack significant sequence homology between subfamilies. When modeling GPCRs it is important to get a good alignment between the query and template structure. Each helix has one or more conserved motifs:

Helix 1: GX₃N or GN

Helix 2: N(S,H)LX₃DX_{7,8,9}P

Helix 3: SX₃LX₂IX₂D(E,H)RY

Helix 4: WX_{8,9}P

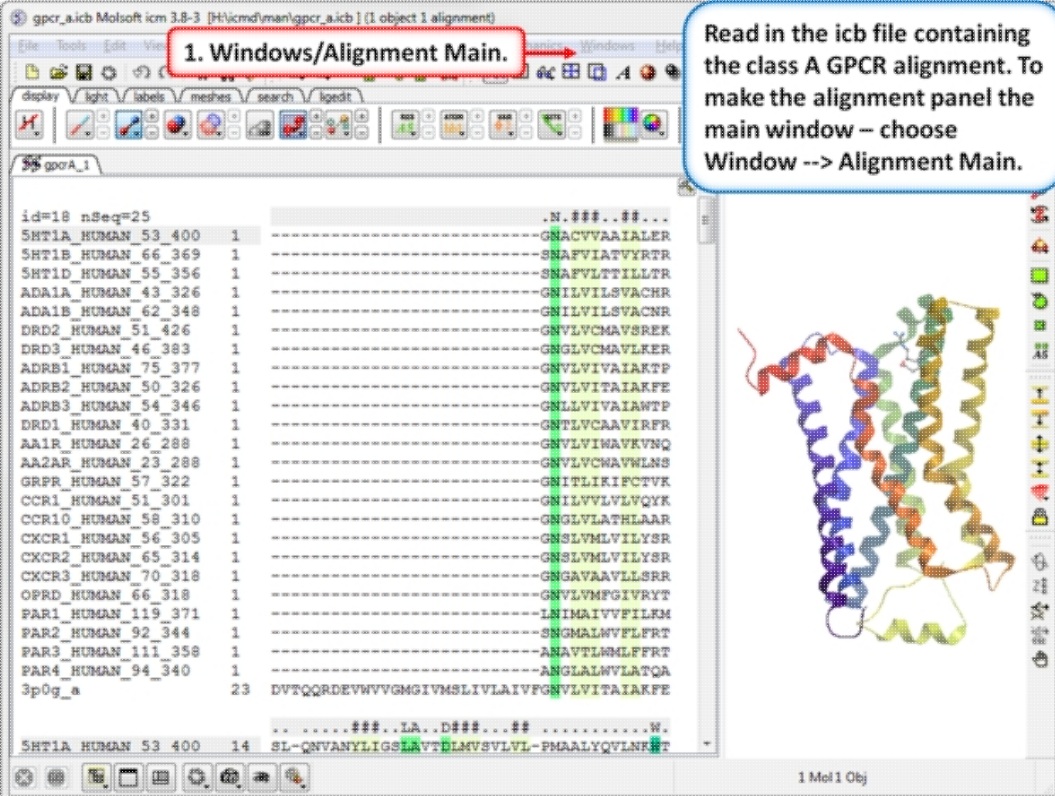
Helix 5: FX₂PX₇Y

Helix 6: FX₂CW(Y,F)XP

Helix 7/Helix 8: LX₃NX₃N(D)PX₂YX_{5,6}F

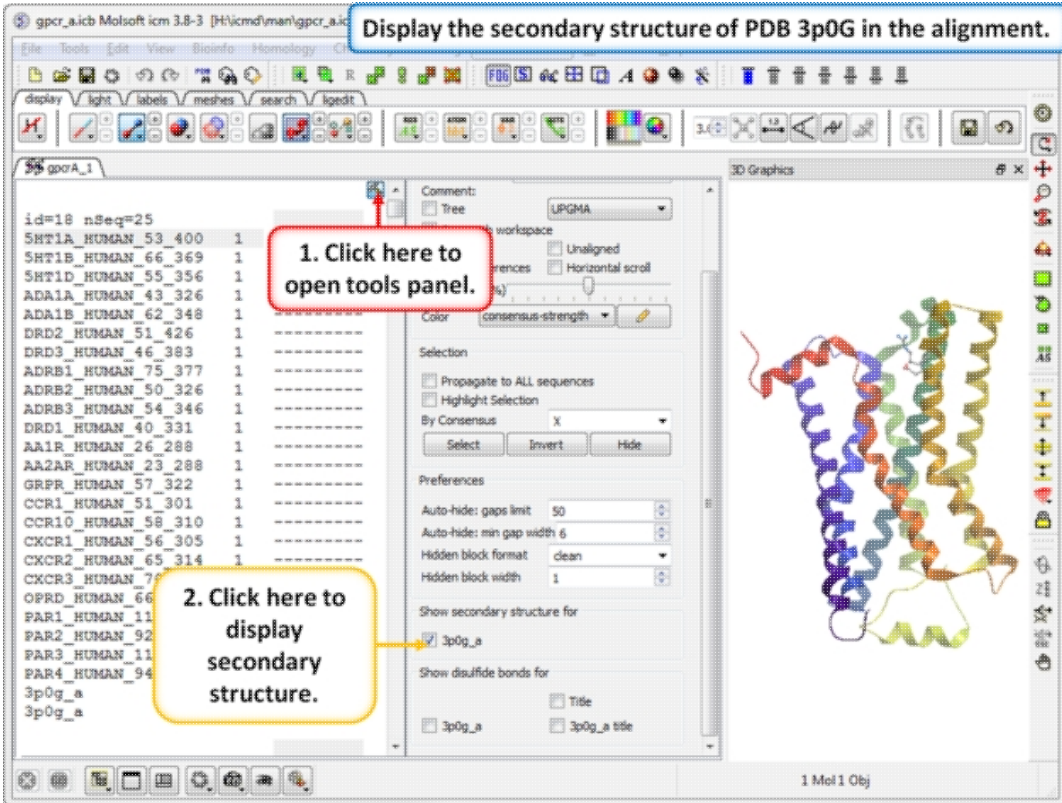
The ProSite class A alignment <http://prosite.expasy.org/PDOC00210> can be used to guide GPCR alignments. For this example we have prepared an icb file containing a subset of the class A GPCR sequences from ProSite. We will use this file to annotate the key conserved motifs this will help with the GPCR modelling example. The file can be downloaded here:

<ftp://ftp.molsoft.com/pub/workshop/> (1. cut and paste the address into a web browser. 2. Right click on gpcr_ClassA and choose "Save Link As". 3. Open the .icb in ICM File/Open).

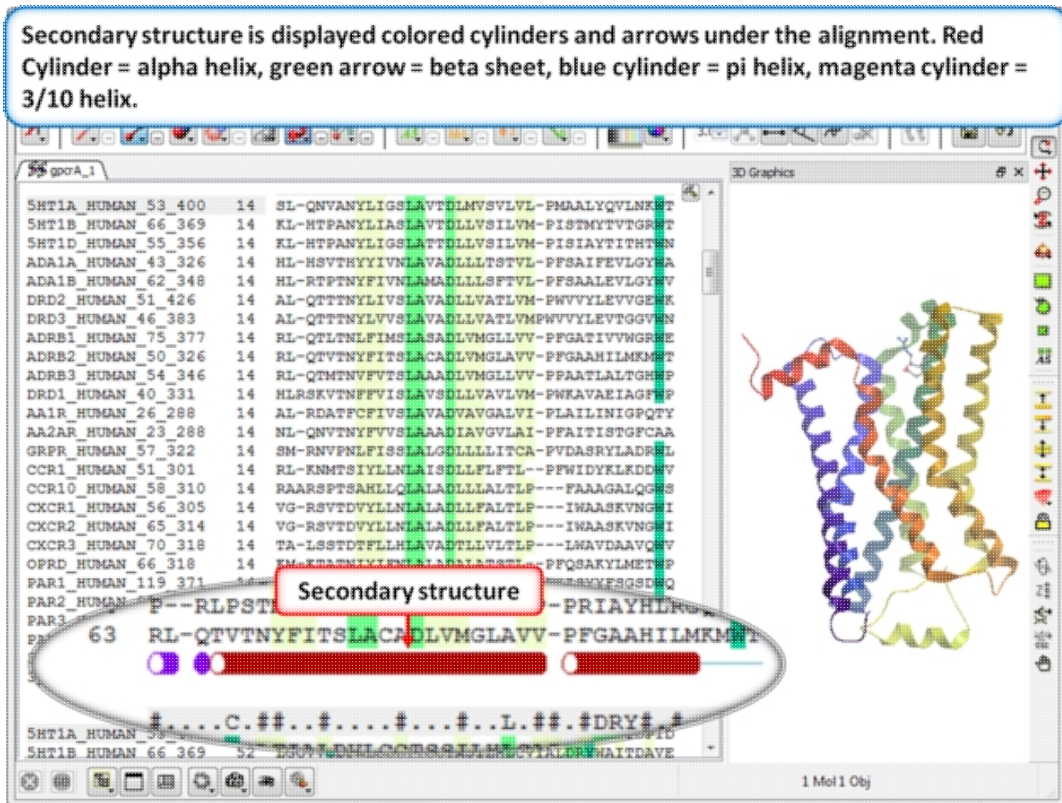


The screenshot shows the Molsoft ICM software interface. The main window displays a sequence alignment of various GPCR sequences. A red box highlights the menu path "1. Windows/Alignment Main." with an arrow pointing to the "Alignment Main" option in the "Windows" menu. A blue callout box contains the following text: "Read in the icb file containing the class A GPCR alignment. To make the alignment panel the main window – choose Window --> Alignment Main." The alignment panel shows a list of sequences with their corresponding amino acid residues. A ribbon diagram of a GPCR structure is visible on the right side of the window.

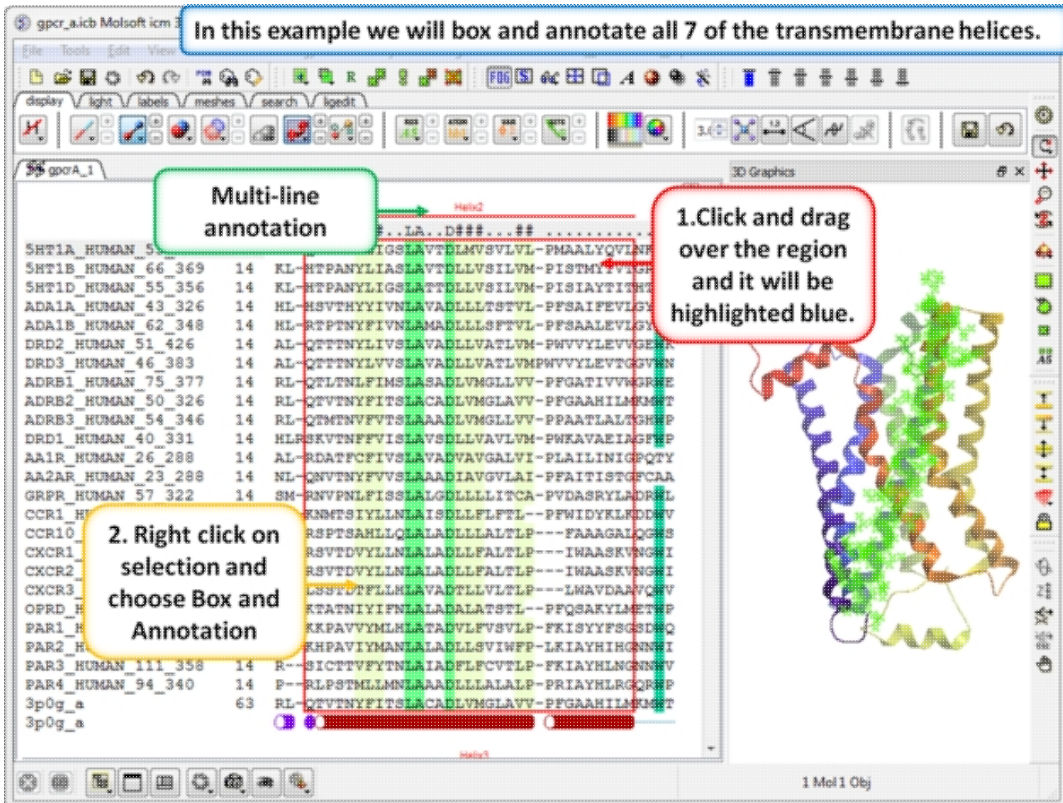
Step 1: Read in the class A GPCR alignment (gpcr_ClassA.icb - see link above). Make the alignment the main window (Windows/Alignment Main).



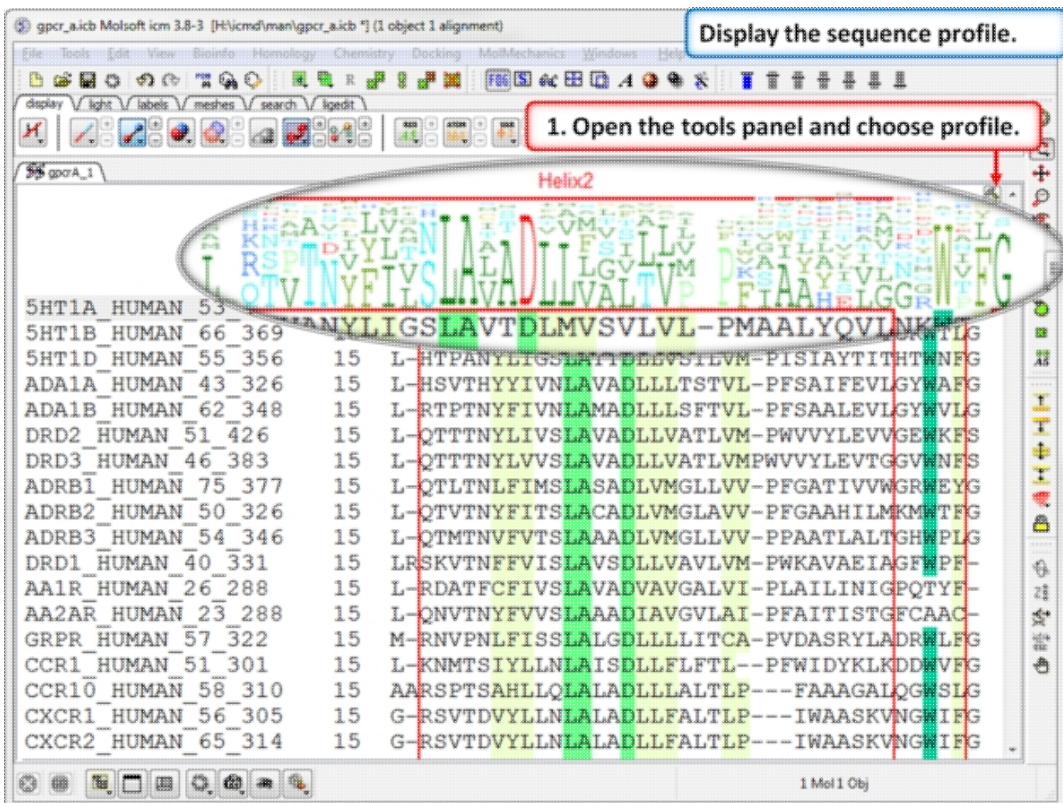
Step 2: The sequence of the beta 2 adrenoceptor is appended to the alignment. This allows us to display the secondary structure of the protein in the alignment.



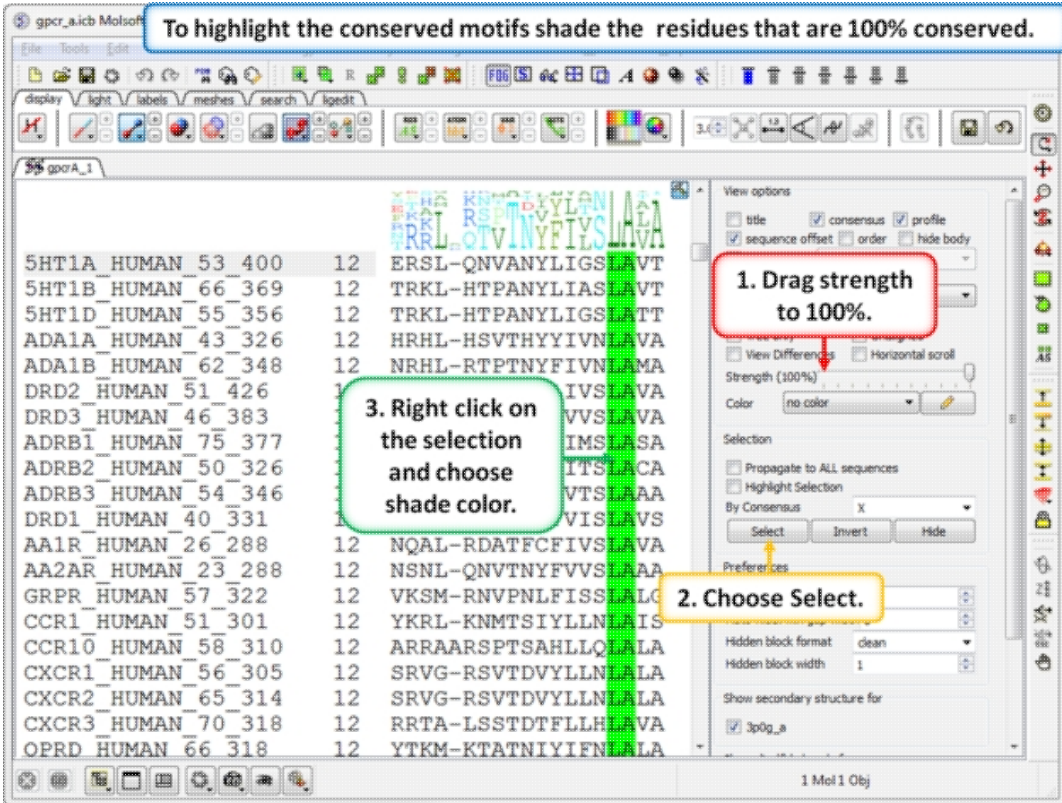
Step 3: Secondary structure is displayed as colored cylinders and arrows under the alignment. Red Cylinder = alpha helix, green arrow = beta sheet, blue cylinder = pi helix, magenta cylinder = 3/10 helix.



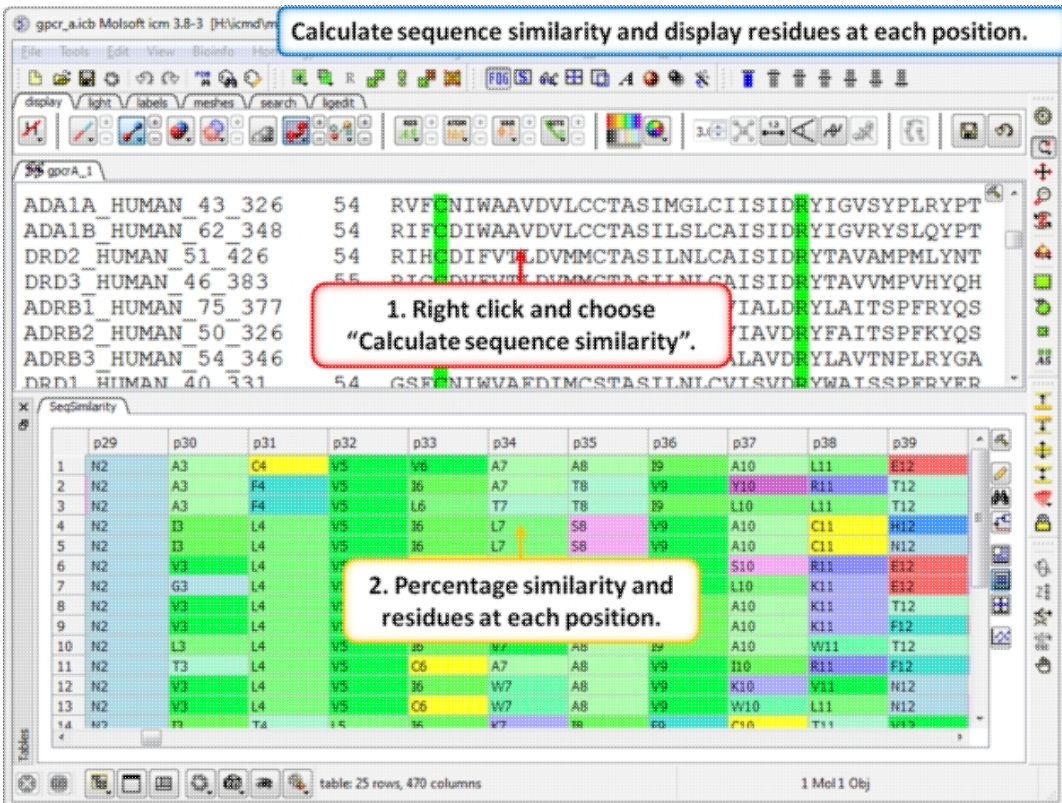
Step 4: Box and annotate the seven transmembrane helices.



Step 5: A sequence profile helps to get a better understanding of the sequence conservation at each point.



Step 6: To highlight the conserved sequence motifs - set the consensus strength to 100% and then color the fully conserved residues. You may need to lower the conservation strength to pick out all the motifs.



Step 7: Calculate the sequence similarity and display the alignment residue profile.

19.5 Protein Modeling Tutorials

Overview

This tutorial will take you through the basics of protein modeling. Topics include:

- GPCR modeling example
- Loop Modeling
- Predicting the Effect of Mutation on Binding
- Predicting the Effect of Mutation on Protein Stability

Background

ICM has an excellent record in building accurate models by homology. The ICM modeling procedure builds the framework, shakes up the side-chains and loops by global energy optimization. You can also color the model by local reliability to identify the potentially incorrect regions of the model. ICM also offers a fast and completely automated method to build a model by homology and extract the best fitting loops from a database of all known loops. It just takes a few seconds to build a complete model by homology with loops. Some selected publications related to modeling and structure determination are listed here.

Abagyan, R.A., and Totrov, M.M. (1994). Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins. *J. Mol. Biol.*, 235, 983-1002

Cardozo, T., Totrov, M., and Abagyan, R. (1995). Homology modeling by the ICM method. *Proteins: Structure, Function, Genetics*, 23, 403-414

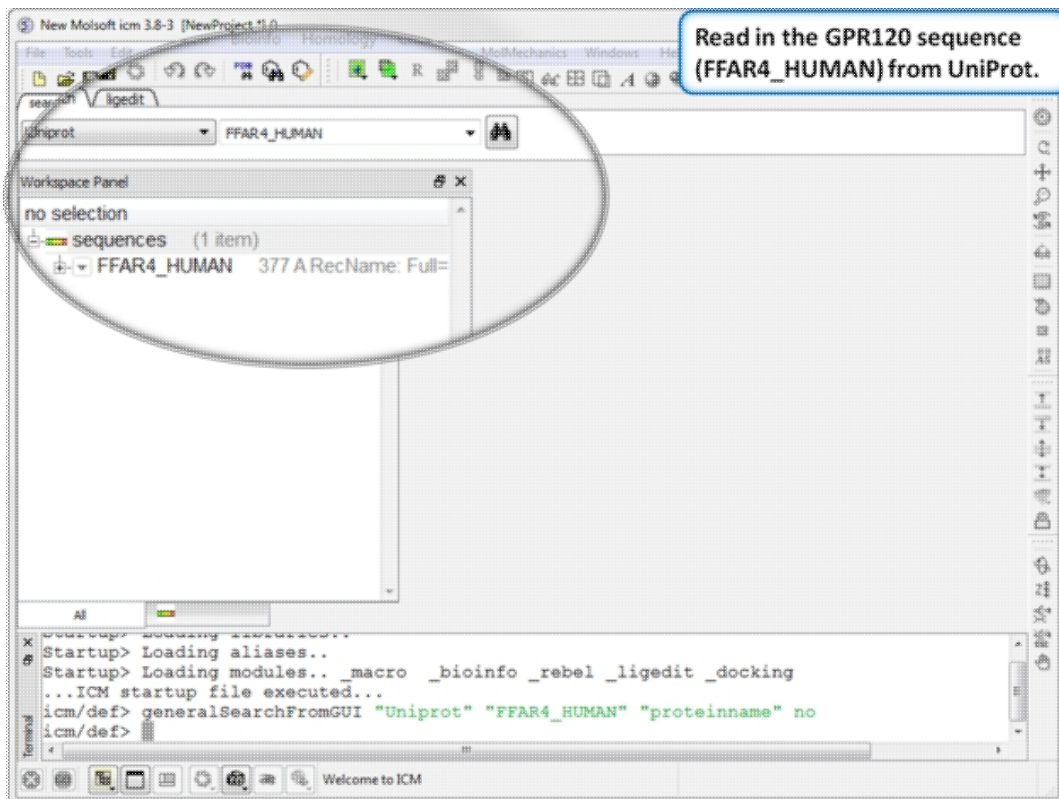
Abagyan, R., and Totrov, M. (1999). Ab initio folding of peptides by the optimal-bias Monte Carlo minimization procedure. *Journal of Computational Physics*, 151, 402-421

Maiorov, V.N., and Abagyan, R.A. (1997). A new method for modeling large-scale rearrangements of protein domains. *Proteins*, 27, 410-424

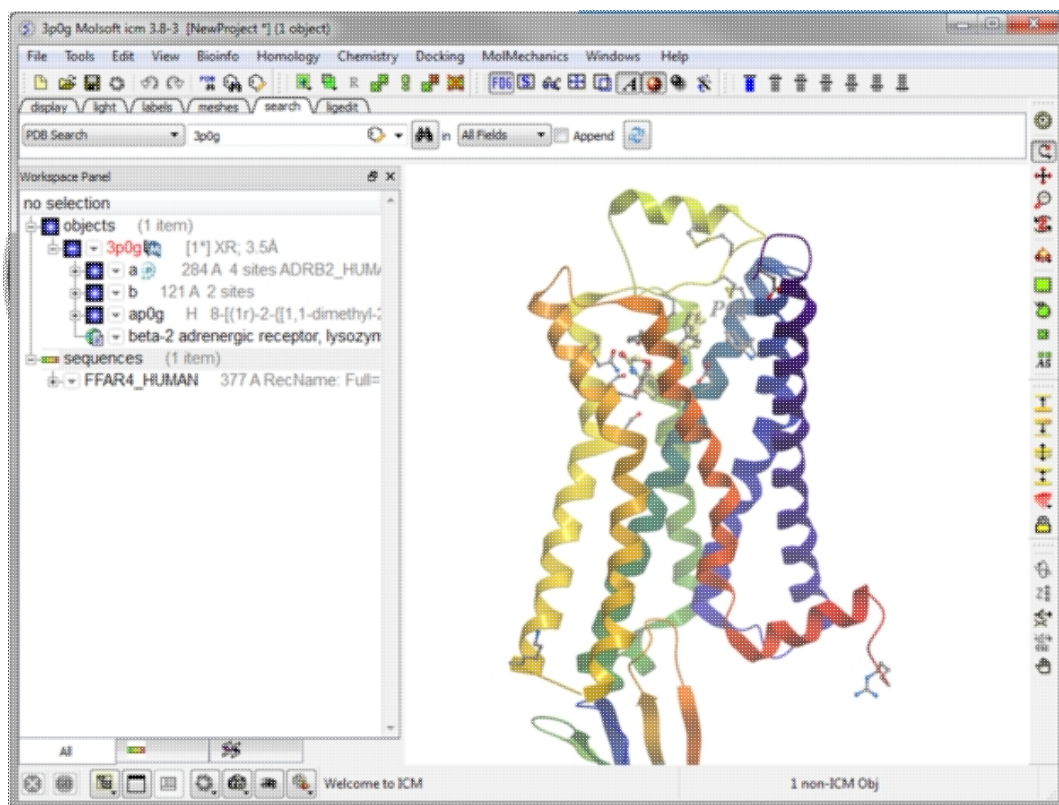
Schapira, M., Totrov, M. and Abagyan, R. (2002). Structural Model of Nicotinic Acetylcholine Receptor Isotypes Bound to cetylcholine and Nicotine. *BMC Structural Biology* 2:1

19.5.1 GPCR Modeling Example

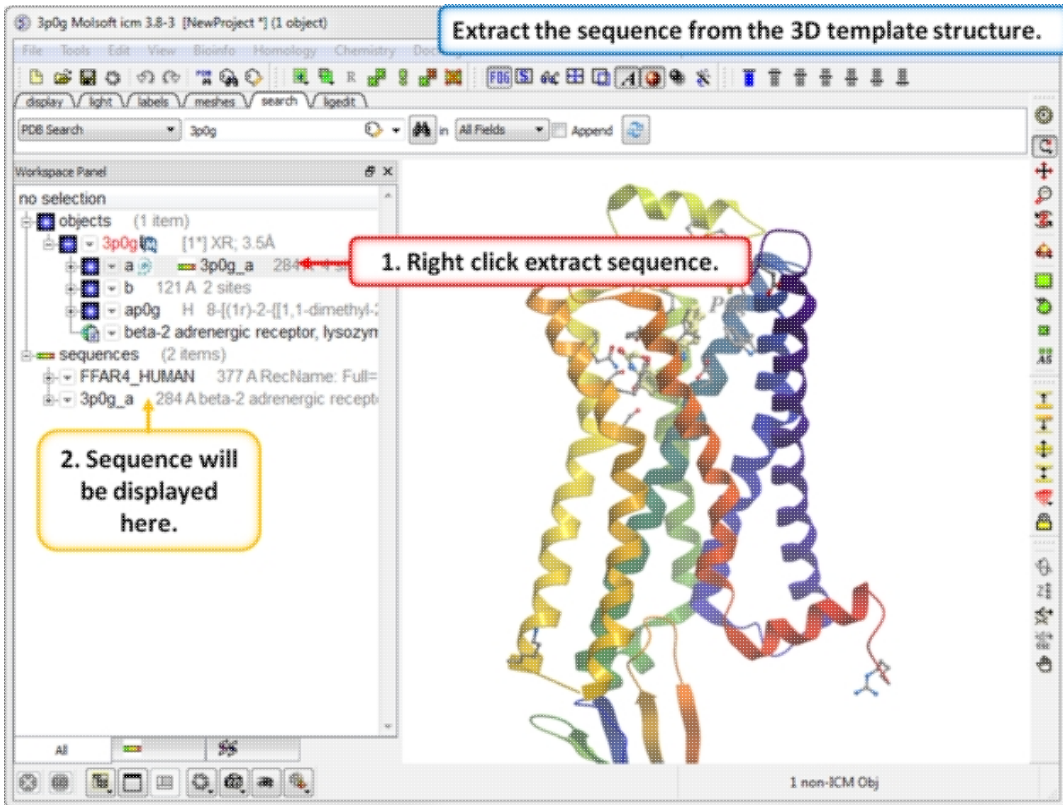
In this example we will build a homology model of GPR120 which is a member of the Class A GPCRs. GPR120 is gaining particular interest because it has been shown to be involved in the insulin-sensitizing effects of omega 3 fatty acids as well as inflammation. GPR120 dysfunction is responsible for reduced fat metabolism, thereby leading to obesity.



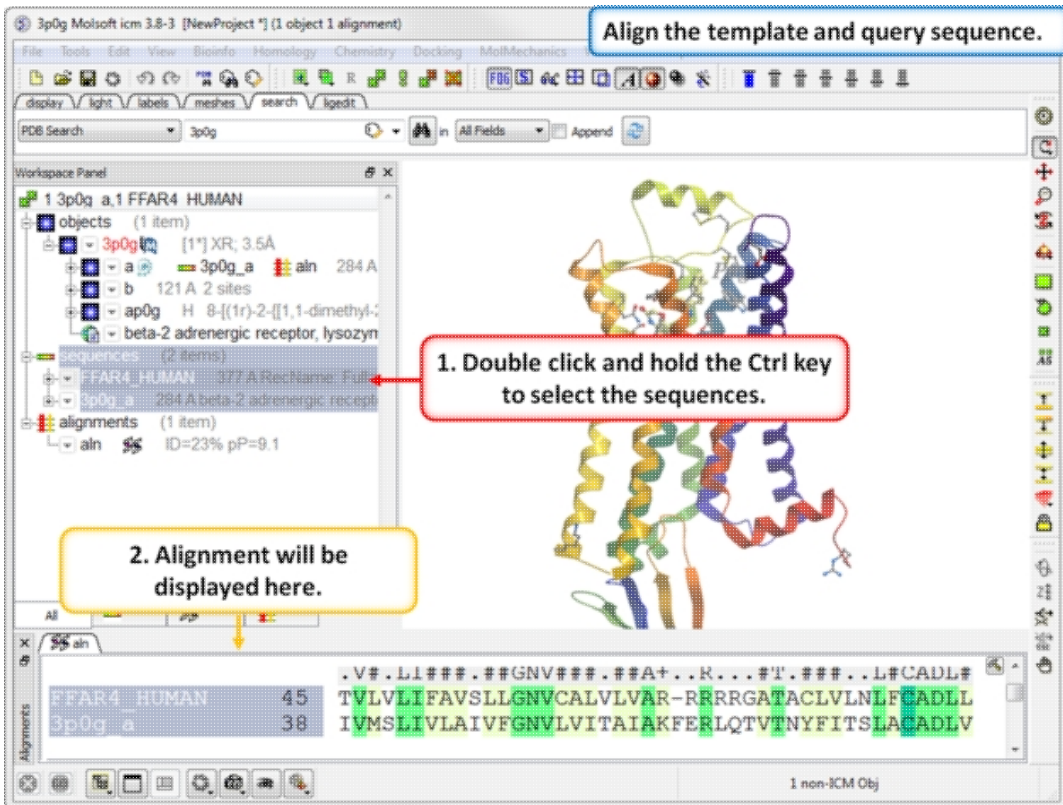
Step 1: Read in the GPR120 sequence (FFAR4_HUMAN) from UniProt.



Step 2: Read in the the template structure 3P0G.



Step 3: Extract the sequence from the 3D template structure.



Step 4: Align the template and query sequence.

Use the cursor keys to adjust the alignment so that the key conserved residues are aligned and there are no gaps in the helices.

```

.....DV.....V###.V#.LI###.##GNV#
SLEQANRTRFPFFSDVKGDRHLVLAAVETTIVLVLIFAVSLLGNVC
-----DVTQQRDEVVVVGMGIVMSLIVLAIVFGNVL
#####L#CADL#
FFAR4_HUMAN 61 ALVLVAR-RRRRGATACLVLNLFCADLLFISAIPLVLAVRWTE-AWLLGVPVACHLLFYVM
3p0g_a 54 VITAIKAKFERLQTVTNYFITSLACADLVMLGLAVVFGAAHILMKMWTFGNFWCFEFTSID
3p0g_a

.L...#.I.TL###.#-R###I..#+...#+...V#####W.#S..#####.##.
FFAR4_HUMAN 119 TLGSGVTILTLAAVSLERMVCIVHLQRGVGRGRARAVLLALIWGYSAVAALPLCVFFR
3p0g_a 114 VLCVTASIEITLCVIAVDRYFAITSPPKYQSLLTKNKARVILMVMVIVSGLTSFLPIQMHW
3p0g_a

..#.#Q.##.#.E.#.C.###.....#.###.F#VP.###V#.YS##Q....
FFAR4_HUMAN 179 --VVPQRLPGADQEISICTLIWPTIPGEISWDVSVFTLNFLVPLVIVISYSKILQTSHE
3p0g_a 174 YRATHQEAINCYAETCCDDF-----TNQAYAIASSIVSFVPLVMVIFYSRVFFQEAK-
3p0g_a

.....-#####TL.##M..F.##W.P#
FFAR4_HUMAN 237 LLDARAVVTHSEITKASRKRLTVSLAYSESHQIRVSCQDFRLFRFTLFLLMVSFFIMWSPFI
3p0g_a 266 -----LKEHKALKTLGIIMGTFTLCWLFP
3p0g_a

#I.#####IQ.....I.....#####.###NS.#NP#Y..#.....#.#C#...
FFAR4_HUMAN 297 IITILLILIQNFKQDLVIVPSLFFWVVAFTFANSALNPILYNMILCRNEWKKIFCCFWFP
3p0g_a 290 FIVNIVHVIQD-----NLIRKEVYVILLNWIGVYVNSGFNPLIYCRSPDFRIAFQELLCLRR-
3p0g_a

```

Step 5: Use the cursor keys to adjust the alignment so that the key conserved residues are aligned and there are no gaps in the helices.

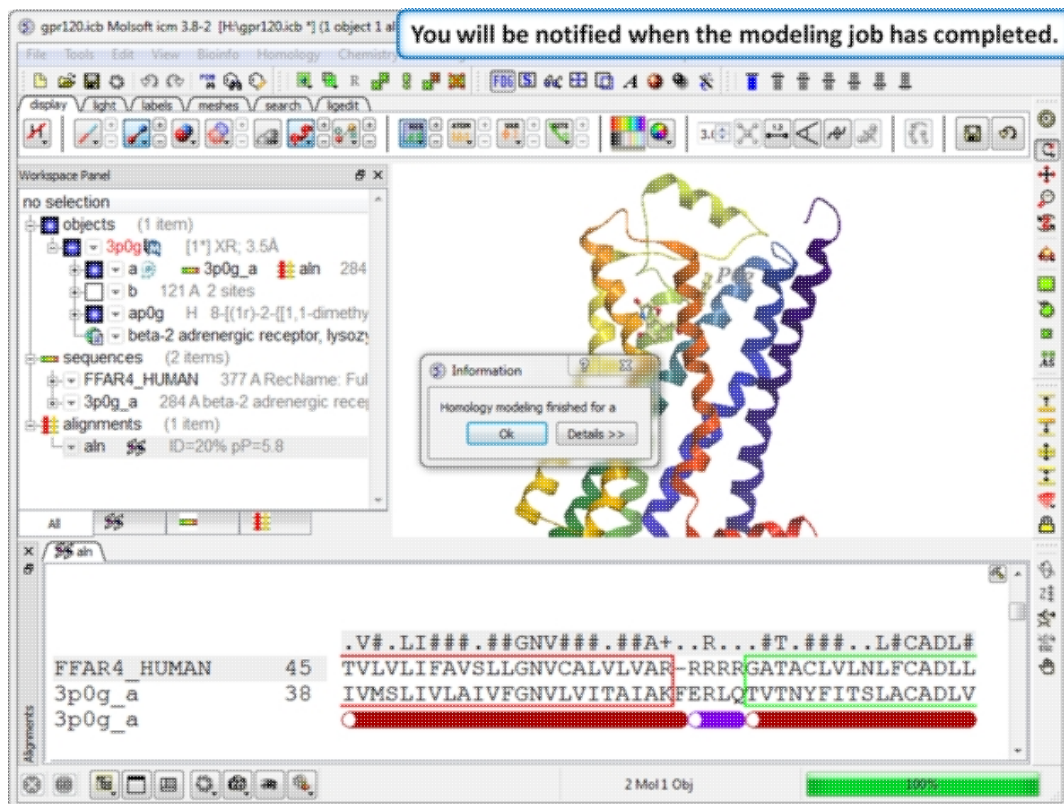
Homology/Full Model Builder. Use the Quick Test method first.

Workspace Panel

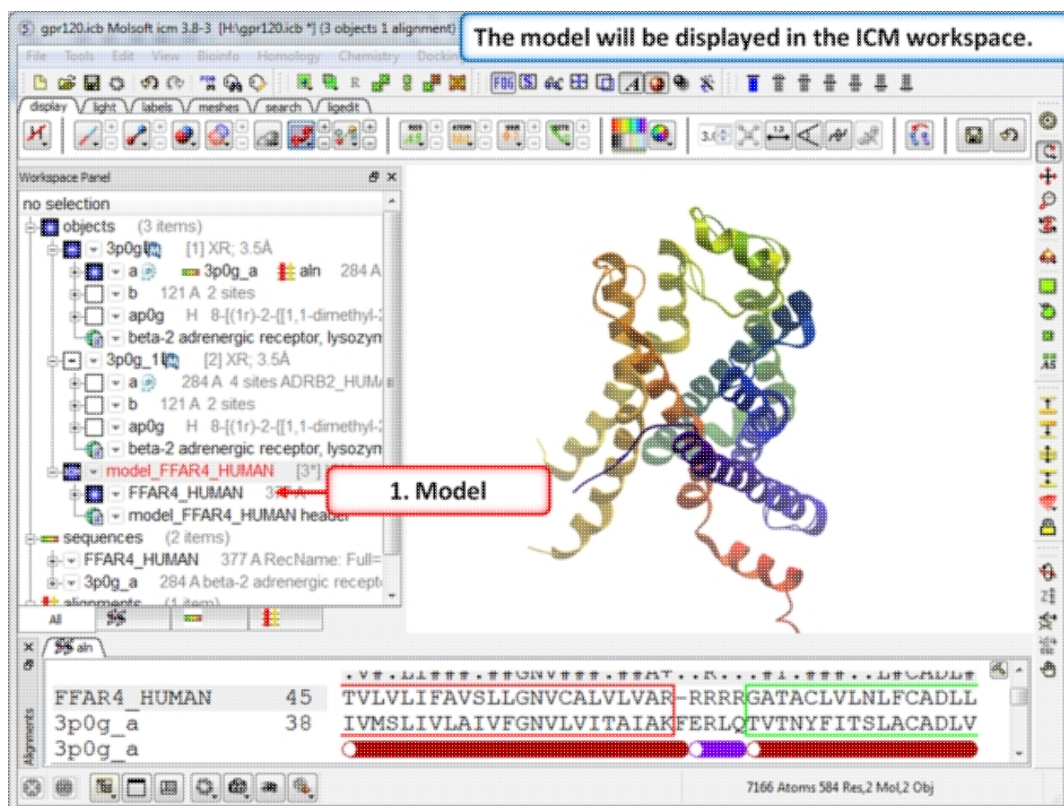
- no selection
- objects (1 item)
 - 3p0g [1] XR, 3.5Å
 - a 3p0g_a
 - b 121 A 2 sites
 - ap0g H 8-((1r)-2)
 - beta-2 adrenergic re
- sequences (2 items)
 - FFAR4_HUMAN 377 A R
 - 3p0g_a 284 A beta-2 adf
- alignments (1 item)
 - aln ID=20% pP=5

FFAR4_HUMAN 45 TVLVLI FAVSLLGNVCALVLVAR-RRRRGATACLVLNLFCADLL
3p0g_a 38 IVMSLIVLAI VFGNVLVITAIKAKFERLQTVTNYFITSLACADLV
3p0g_a

Step 6: Choose the option Homology/Full Model Builder. Use the Quick Test method first to inspect the quality of the model-sequence alignment.



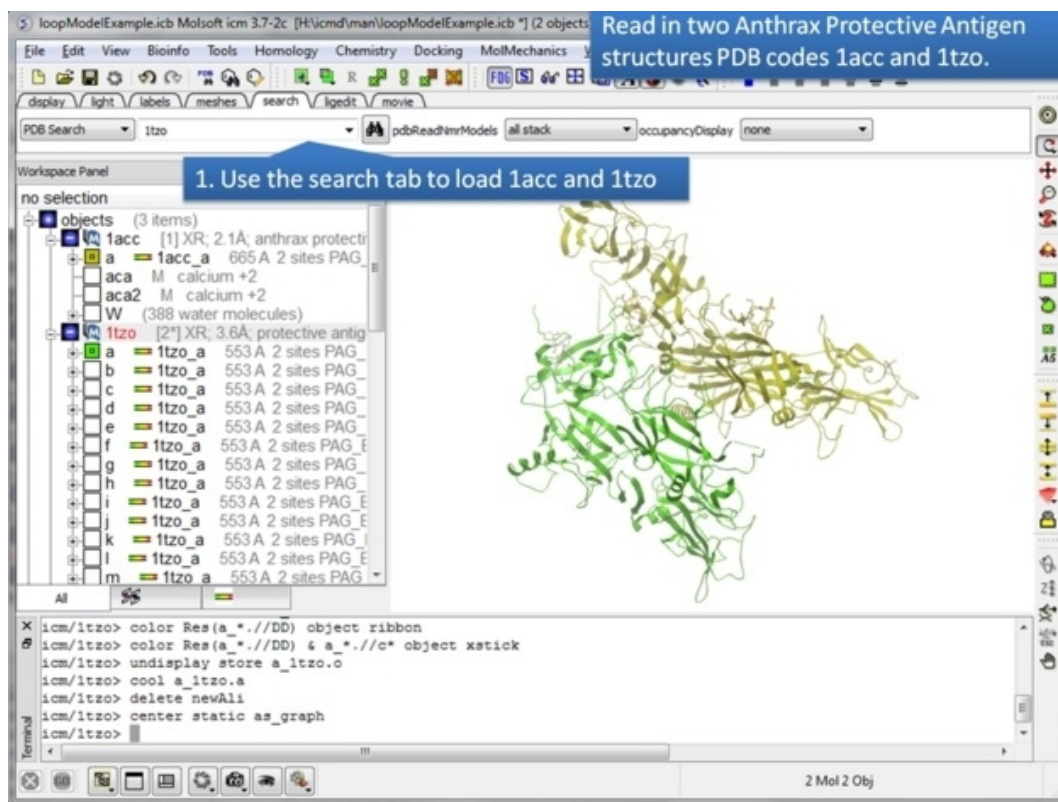
Step 1: A dialog box will be displayed when the modeling job has finished.



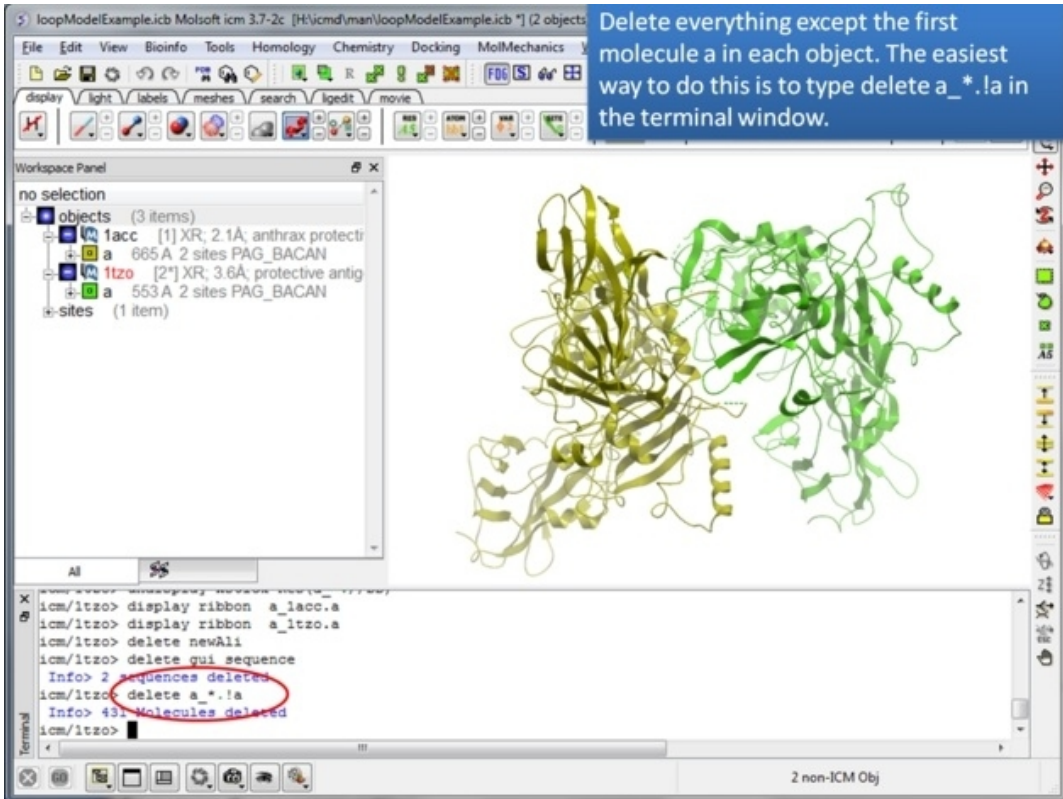
Step 8: The model will be displayed in the ICM Workspace.

19.5.2 Loop Modeling Tutorial

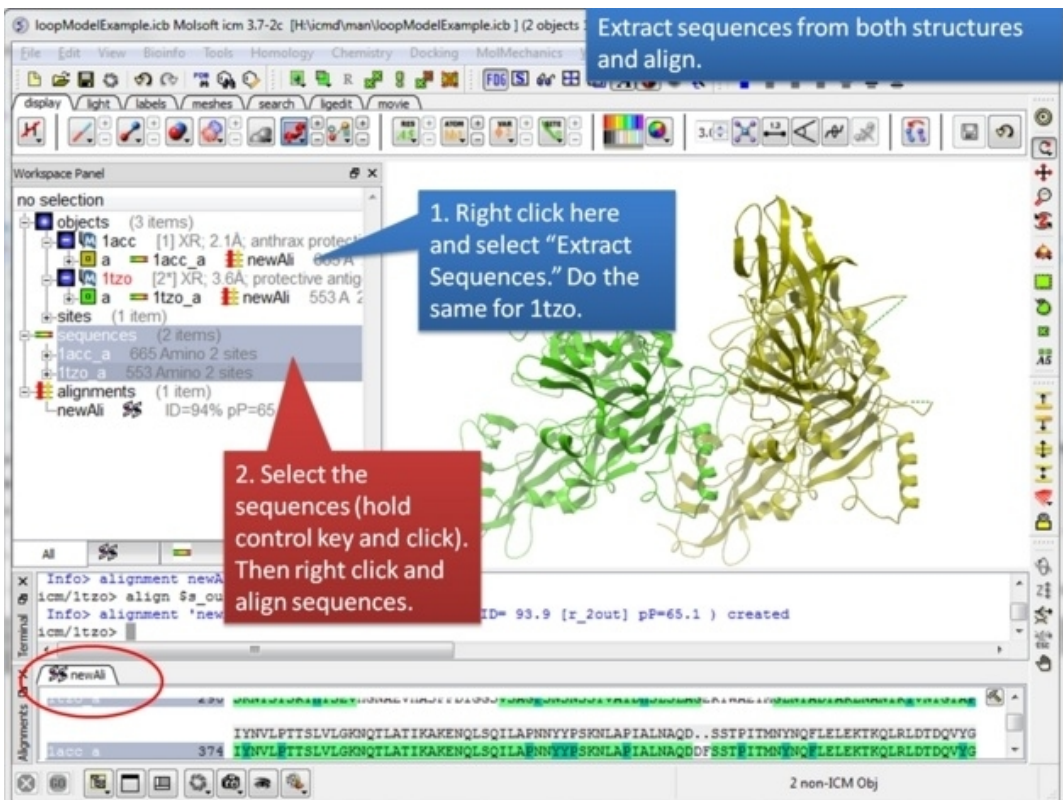
In this example we will use the high precision loop modeling method Arnautova et al Proteins. 2010, 79:477 to model a loop region in Anthrax Protective Antigen which contains a proline residue where the cis/trans conformation is not clear from the available X-ray crystal structures. We will determine the most energetically favorable conformation of the proline and the neighboring loop residues.



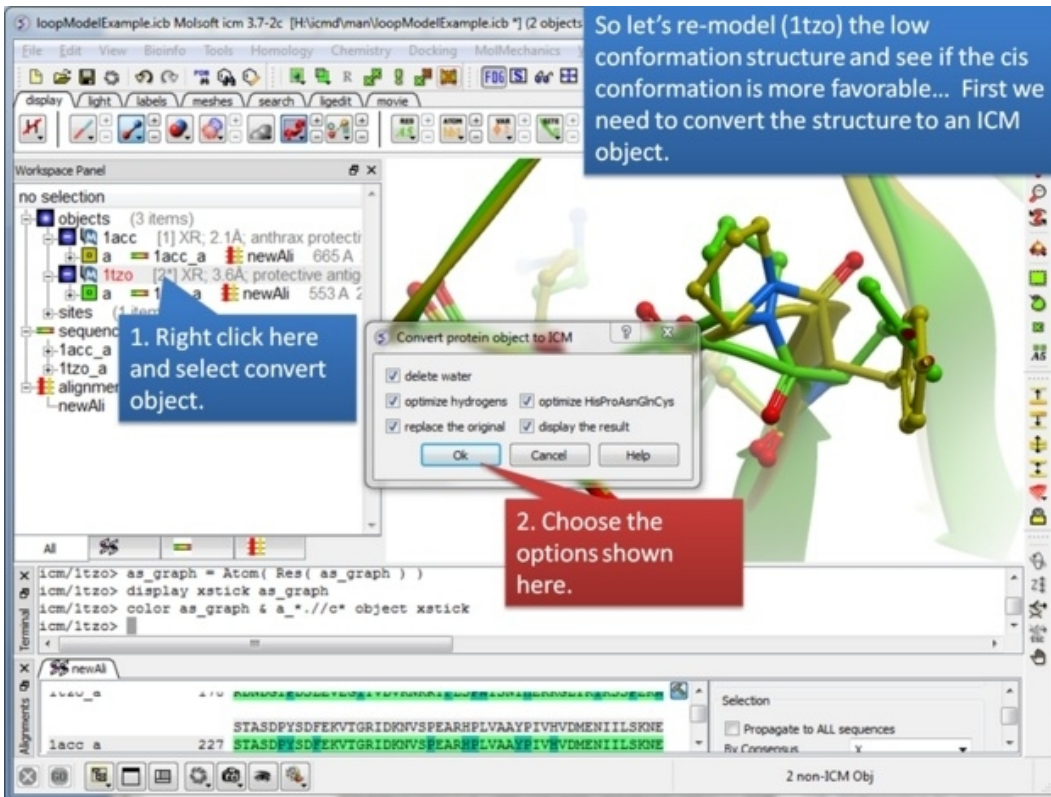
Step 1: Use the search tab to read in two pdb structures 1acc and 1tzo.



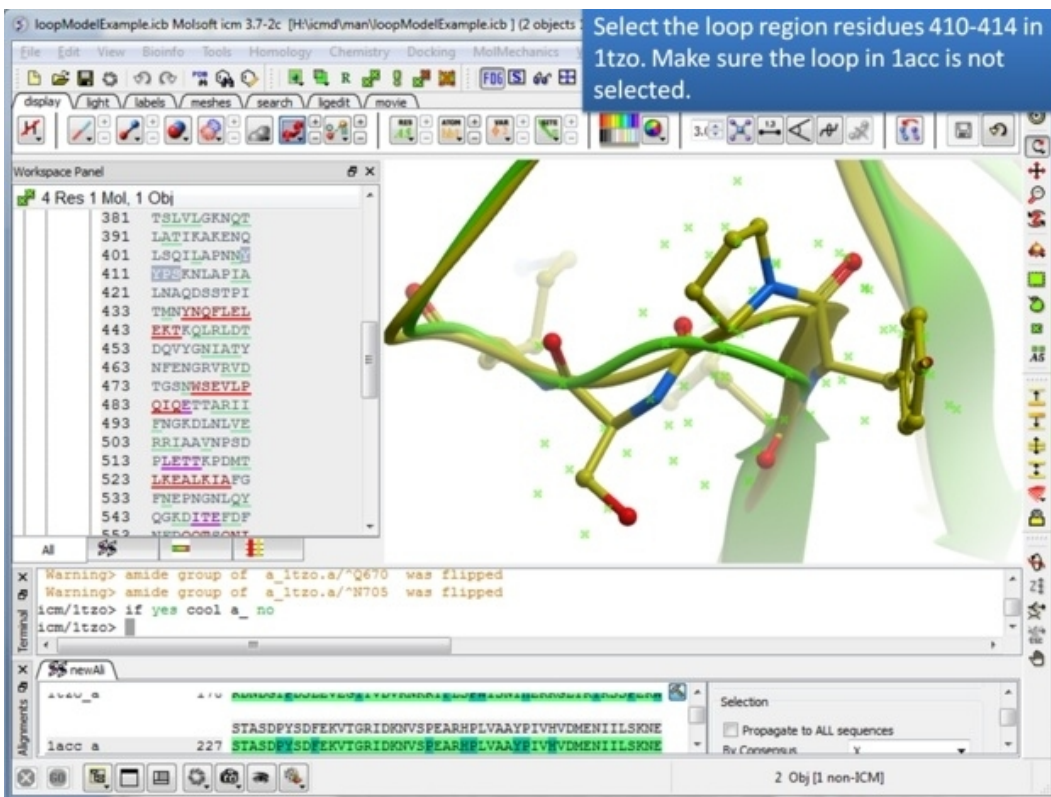
Step 2: Delete everything except for the first molecule of each protein. You can use the command line as suggested to delete the other molecules or you can: -use the CTRL button and click on them - right click and select delete.



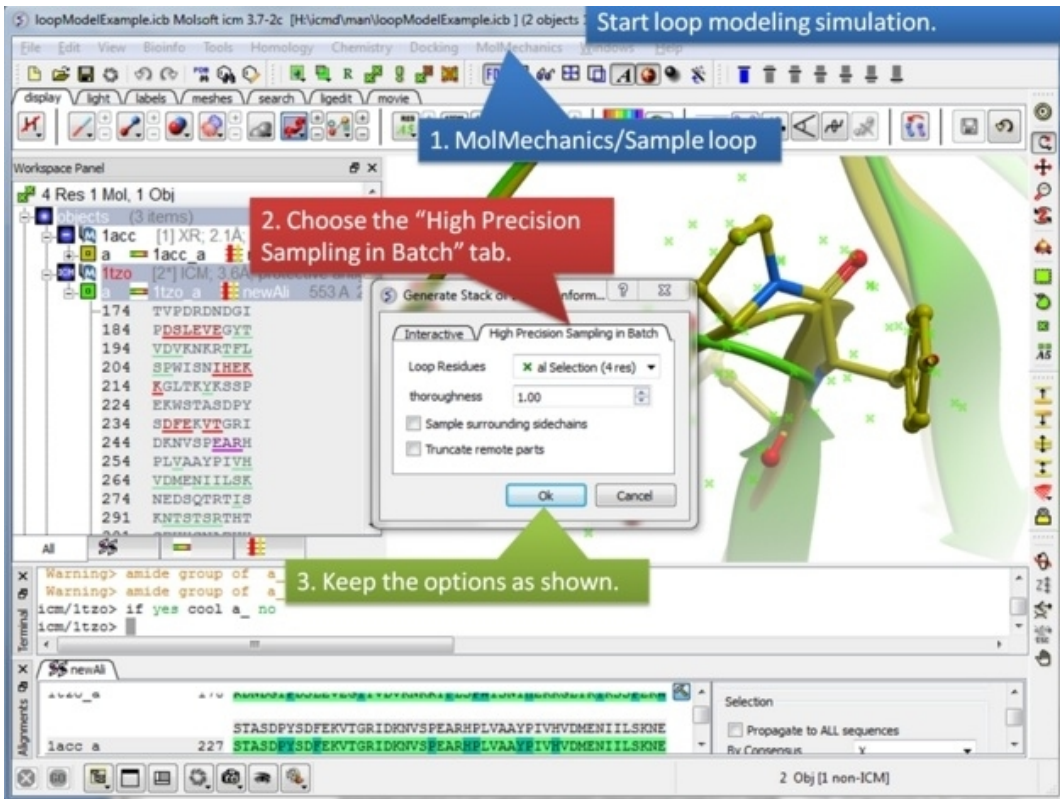
Step 3: Extract the sequences from both receptor by right clicking on the molecule in the ICM workspace. The alignment makes it easier to superimpose the loop and select the loop in both structures.



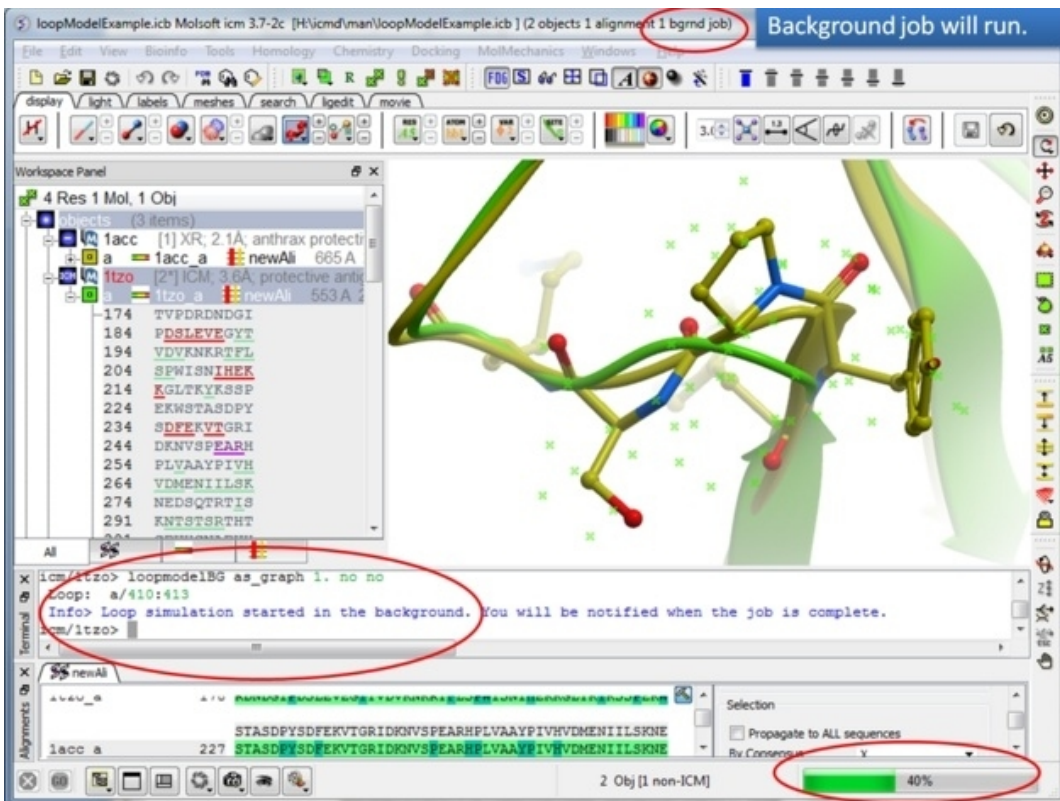
Step 6: We will re-model the lower resolution structure and so we need to convert it into an ICM object.



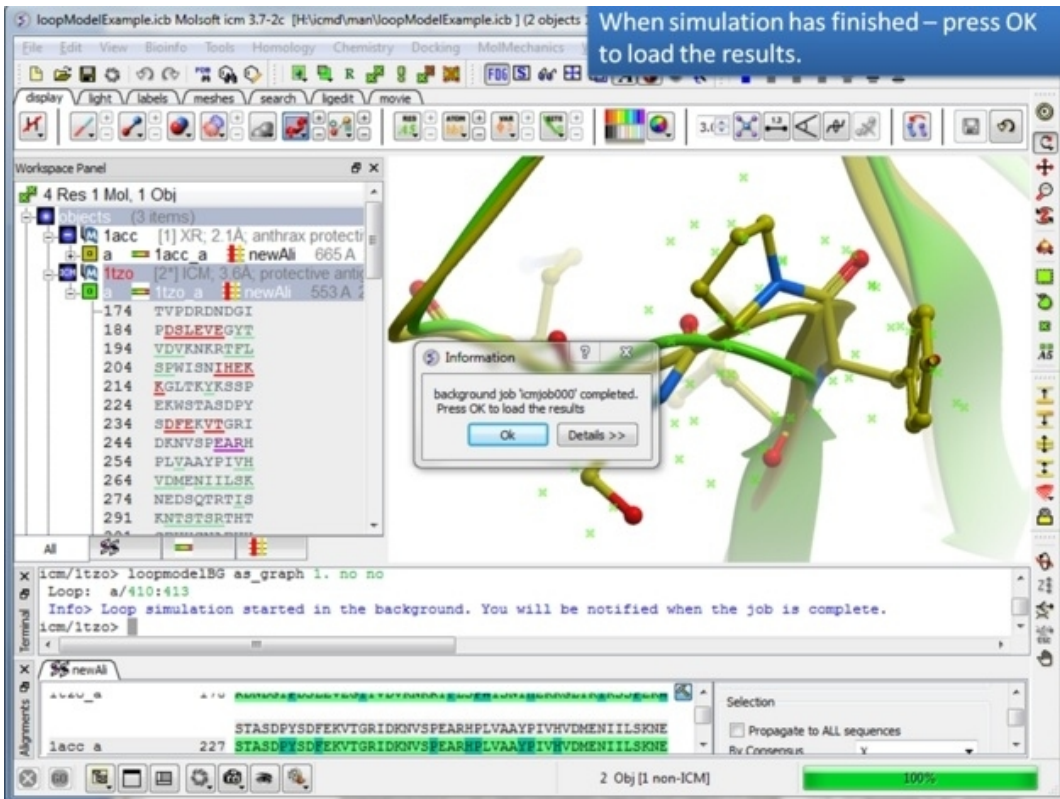
Step 7: Select the loop region in 1tzo only - do not select the loop in both proteins.



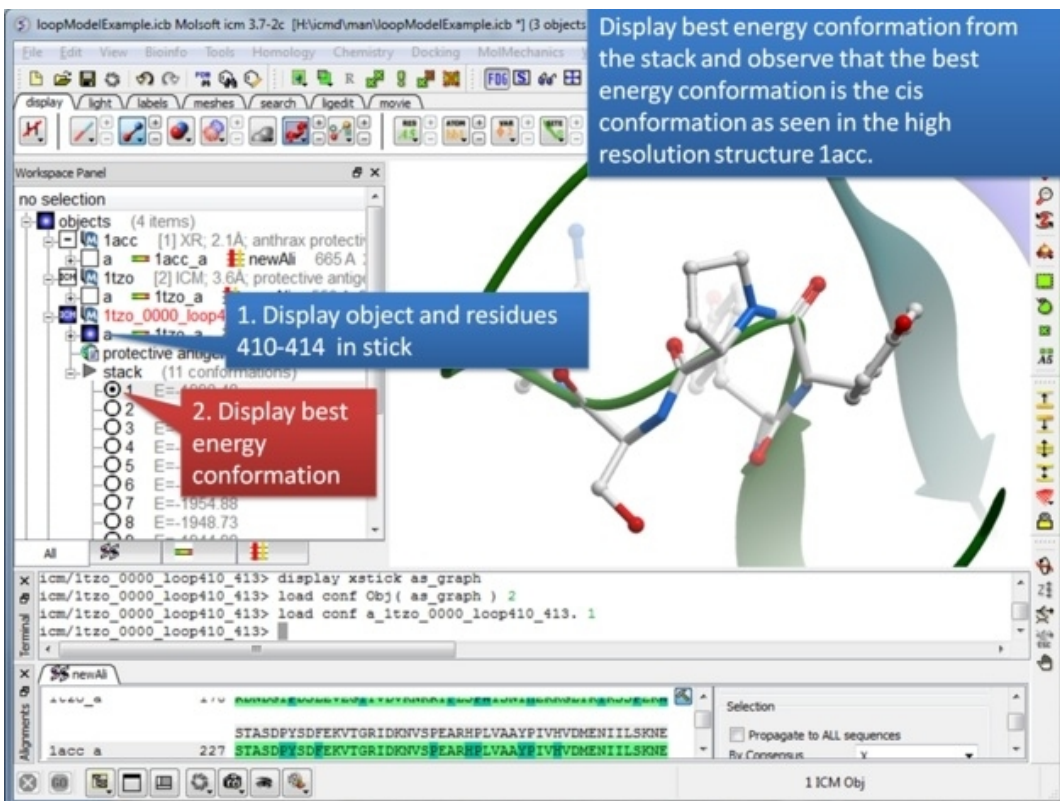
Step 8: Start the loop modeling using the option in the MolMechanics menu.



Step 9: The docking will run in the background.



Step 10: You will be notified when the simulation is finished.



Step 11: The best energy conformation of the proline in the loop is the cis conformation.

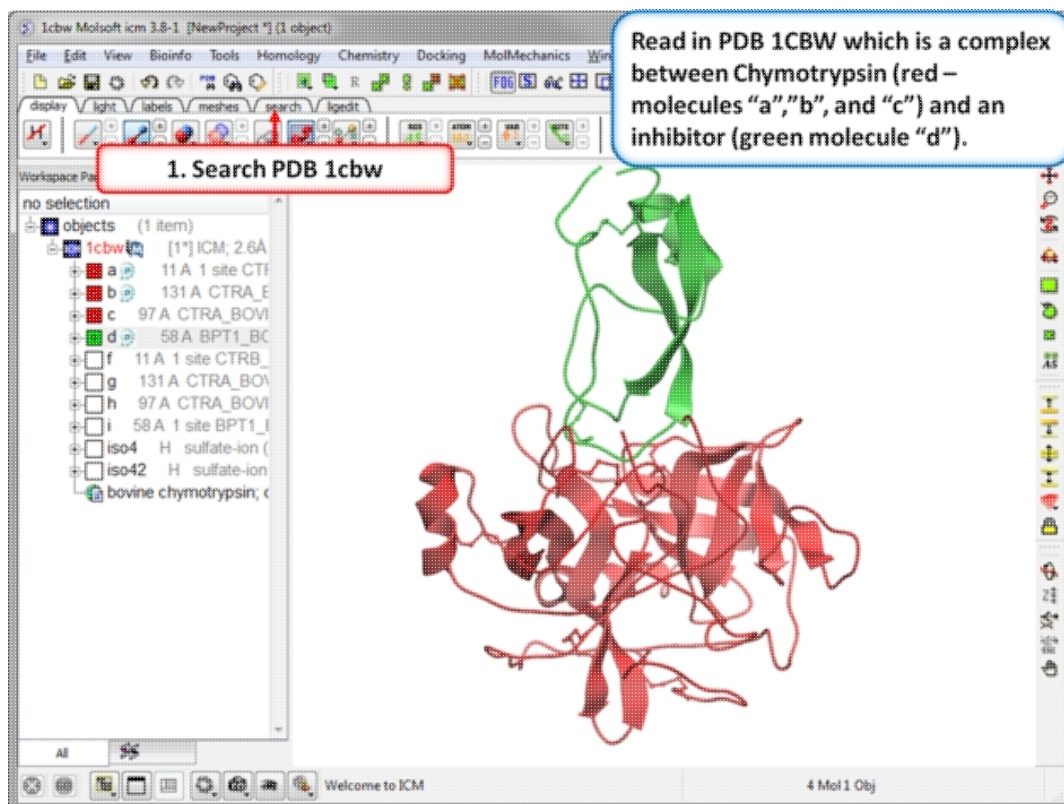
19.5.3 Effect of Mutation on Protein Binding Tutorial

| Documentation |

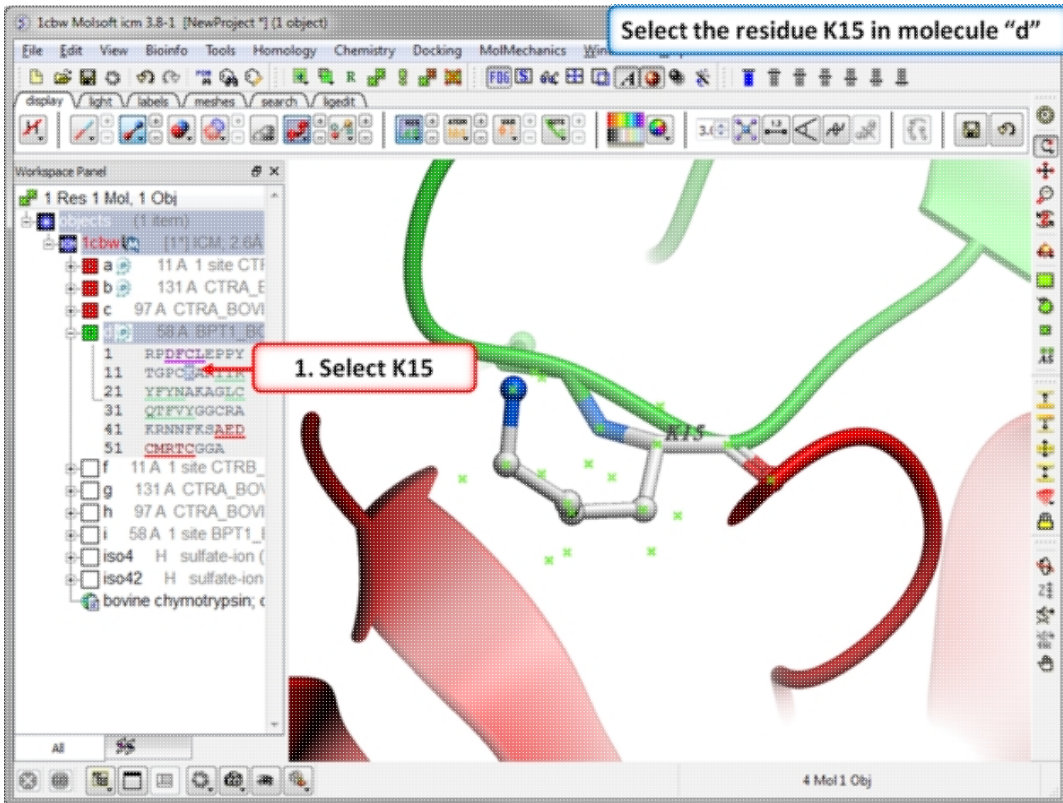
The effect of mutation on protein binding can be predicted.

19.5.3.1 Effect of Mutation between Chymotrypsin and Inhibitor

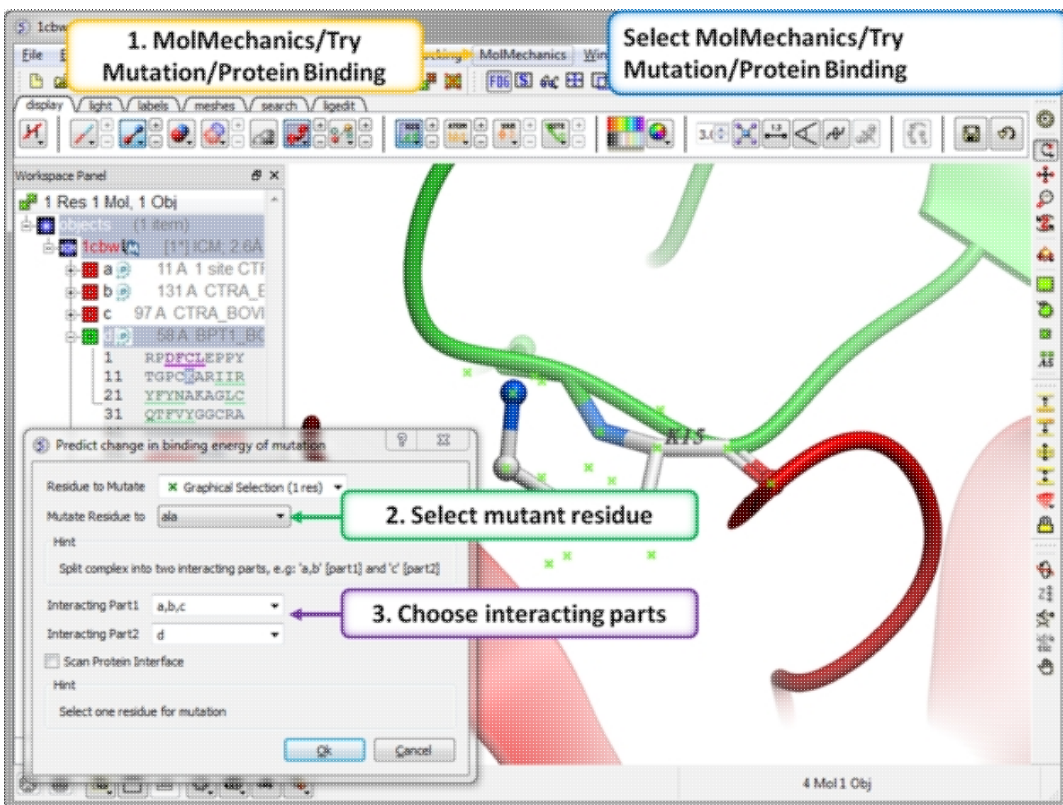
In this example the effect of mutation on binding between chymotrypsin and an inhibitor is predicted (using PDB 1CBW). LYS15 in the inhibitor makes a key binding interaction with the protease and mutation to alanine leads to a significant drop in the binding free energy.



Step 1: Read in the Chymotrypsin-Inhibitor complex (PDB: 1CBW). Notice that the inhibitor (green) interacts with three molecules of chymotrypsin (red).



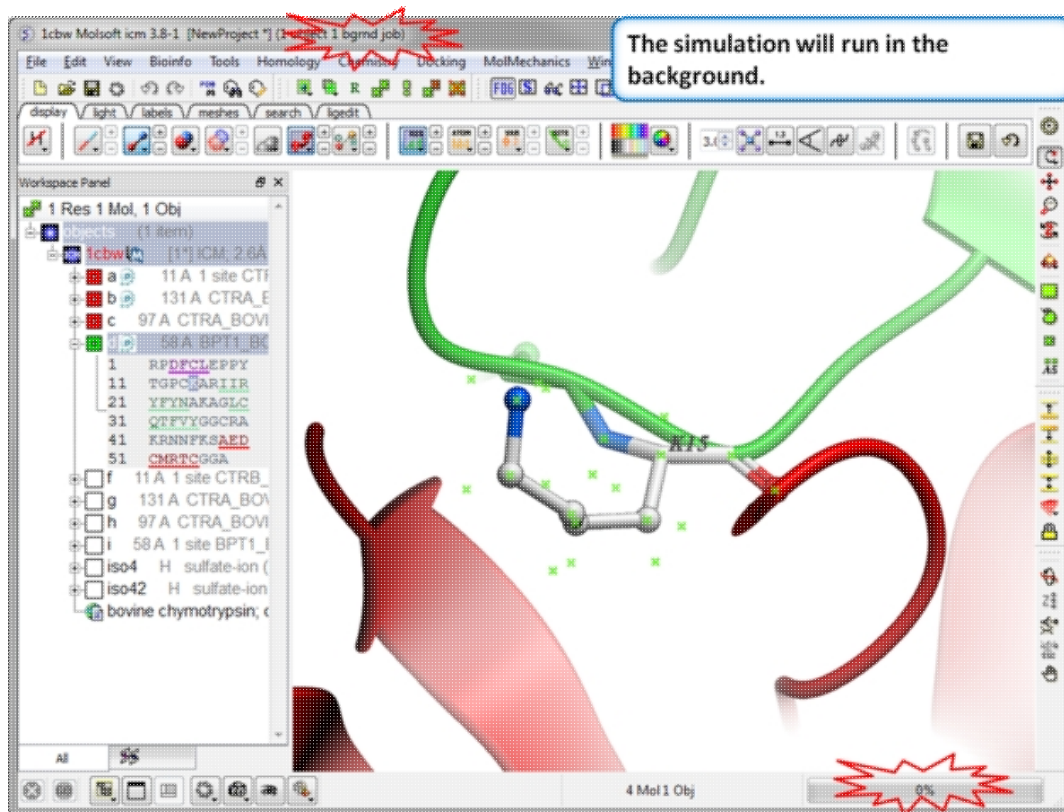
Step 2: Select the residue to mutate. Select residue K15 in the inhibitor - molecule "d".



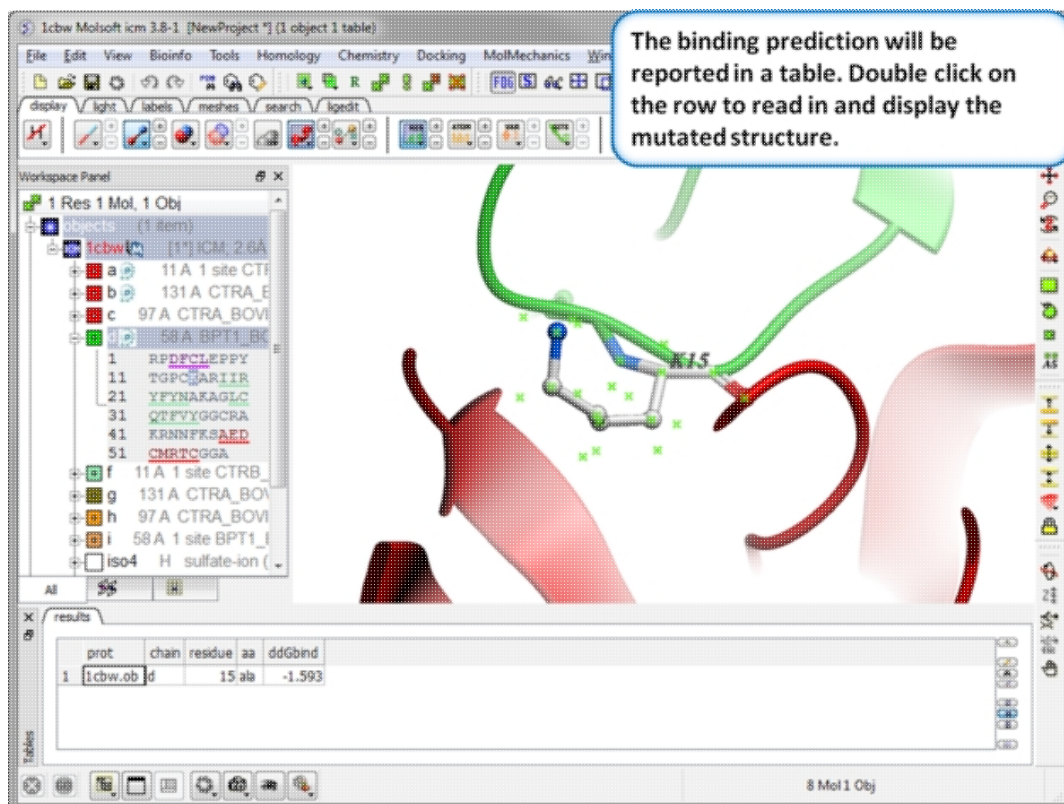
Step 3: MolMechanics/Try Mutation/Protein Binding.

- Check that the residue to mutate data entry box matches the selection (e.g. (1 res)).
- Interacting part 1 = molecules a,b,c

- Interacting part 2 = molecule d
- Click OK.



Step 4: Prediction Running. Prediction will run in the background.

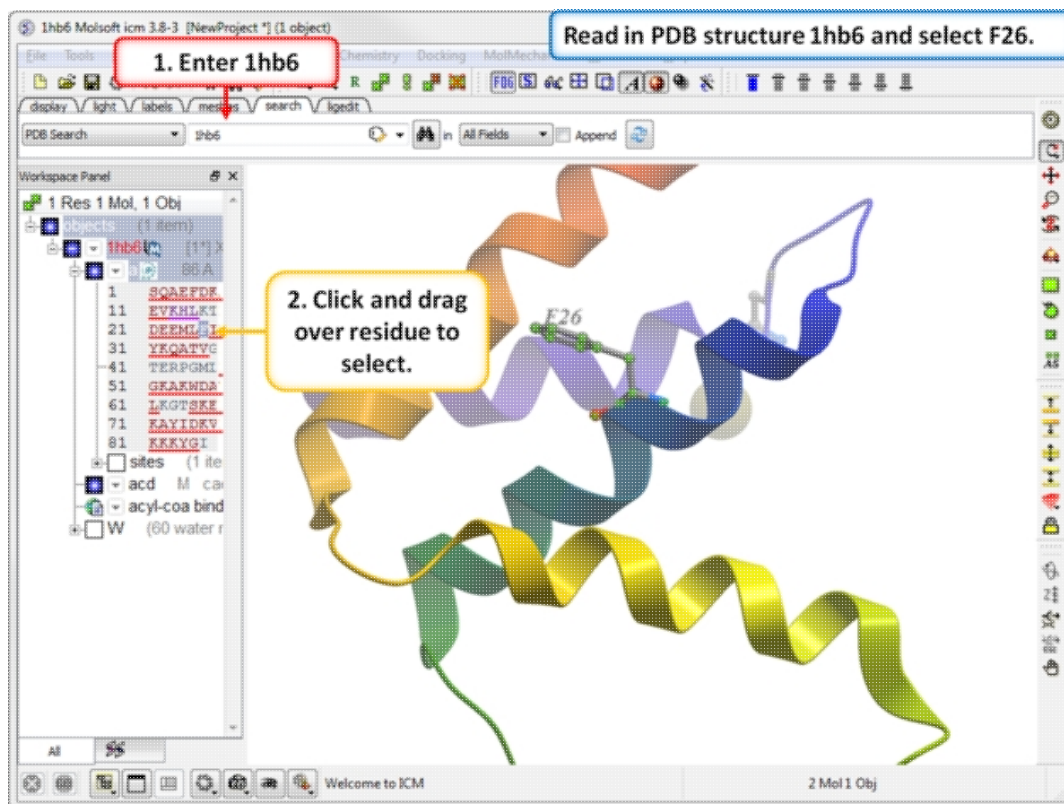


Step 5: Results. A table reporting the binding free energy in kcal/mol (ddGbind) will be displayed. Double click a row of the table to load and display the mutated structure.

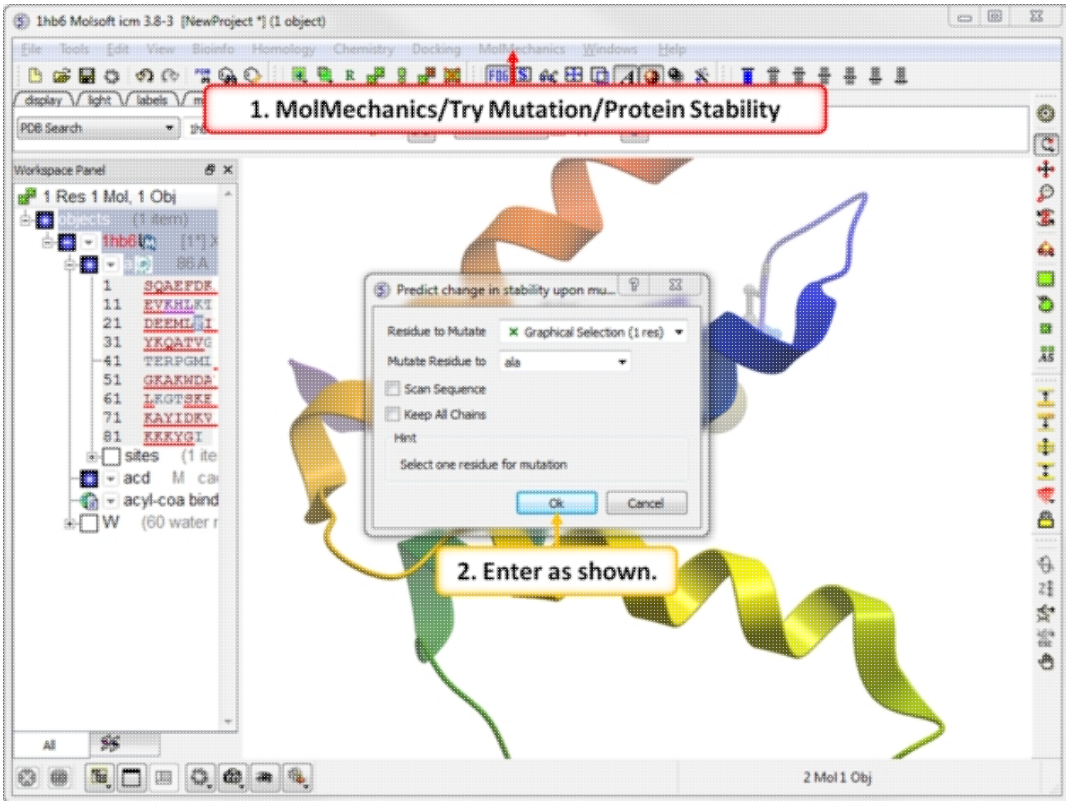
19.5.4 Effect of Mutation on Protein Stability Tutorial

| Documentation |

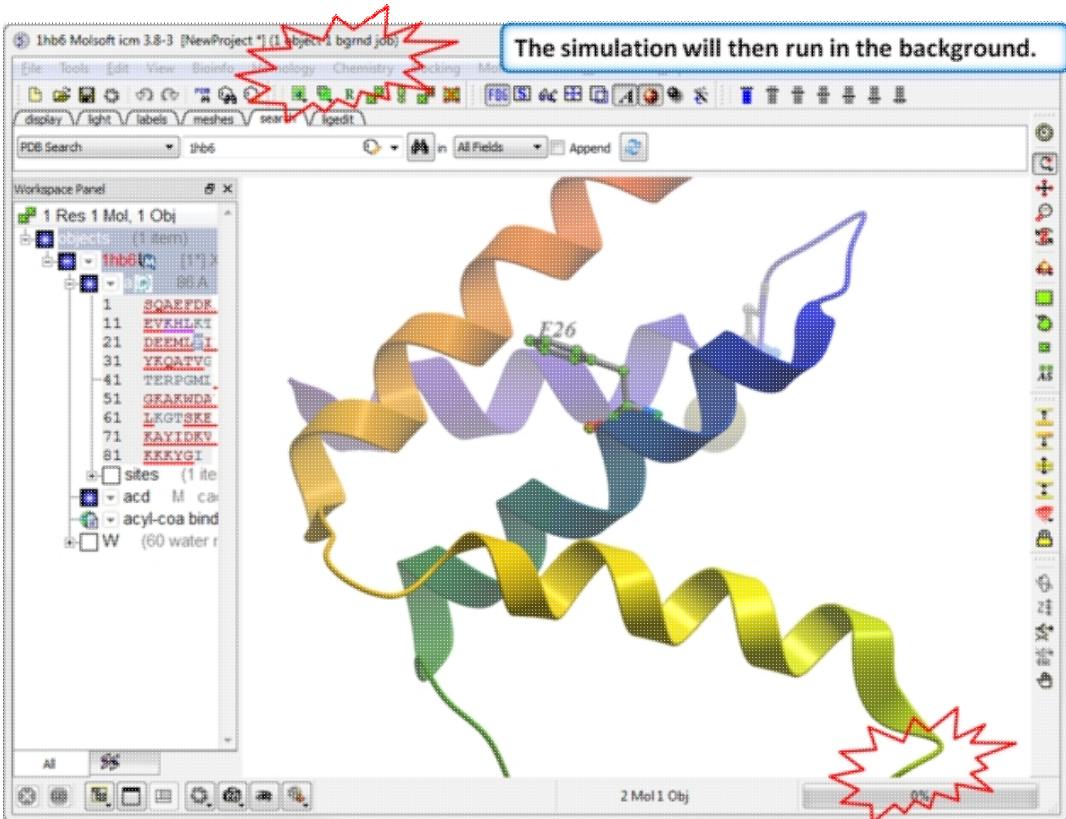
The effect of mutation on protein stability can be predicted. In this example we mutate residue F26 to Ala in the crystal structure of bovine Acyl-CoA binding protein. The experimentally calculated value of this mutation is 1.54 kcal/mol (Kragelund et al. NSB 6, 594:601 (1999)).



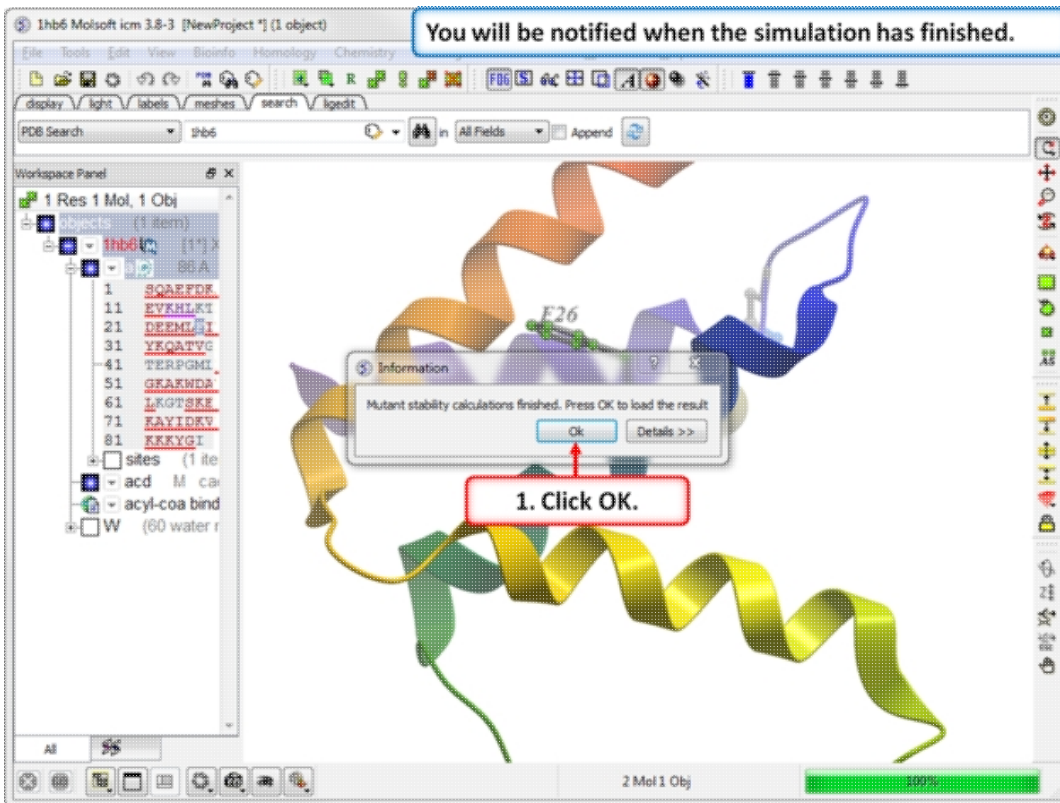
Step 1: Read in PDB 1hb6 and select residue F26.



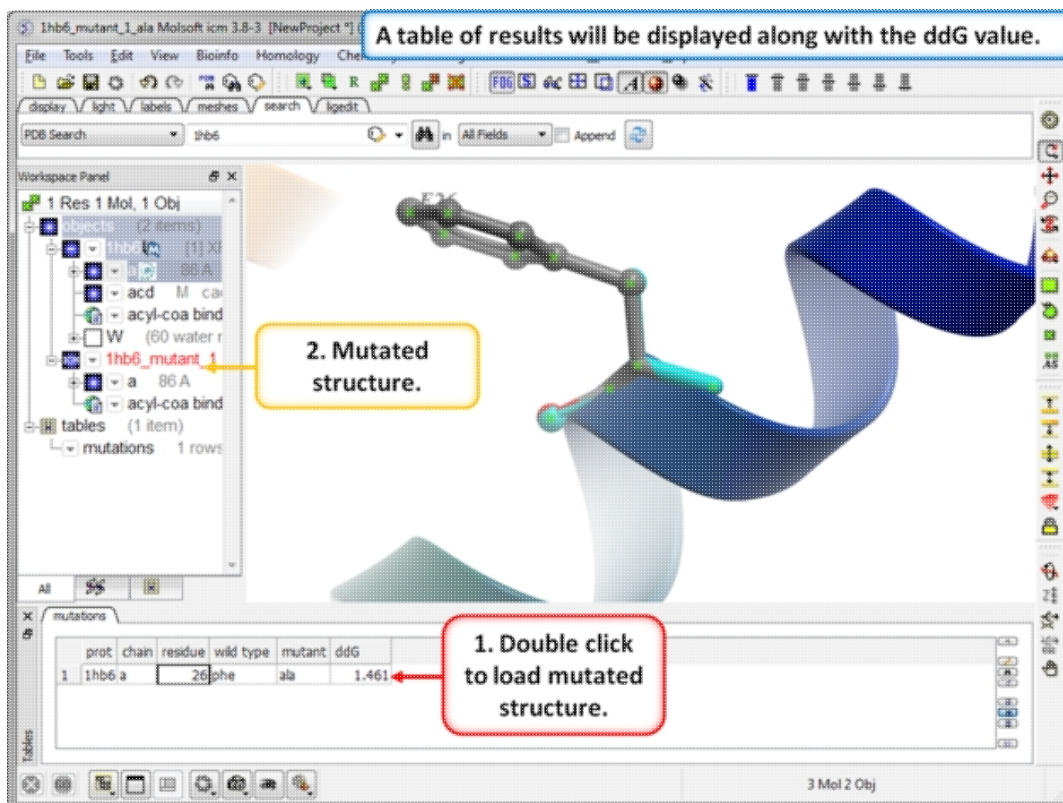
Step 2: Choose the menu option MolMechanics/Try Mutation/Protein Stability and choose Ala mutation.



Step 3: The simulation will run in the background.



Step 4: You will be notified when the simulation is complete.



Step 5: The predicted ddG value should be close to the experimental value of 1.54kcal/mol. Double click on the results table to load and view the mutated structure.

1. Right click "Insert New Row"

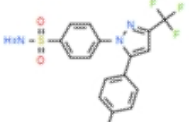
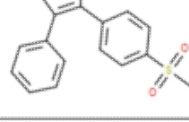
2. Double click and sketch Vioxx shown (right)

3. Close editor and the chemical will be added to the spreadsheet

Add a new row to the chemical spreadsheet and add Vioxx.

Step 4: Add a new row. Add a new row to the spreadsheet and then sketch the Vioxx (Rofecoxib) drug (shown).

Now you should have two chemicals in your spreadsheet.

mol	Molecule Name
	Celebrex
	Vioxx

Step 5: Observe Chemical Spreadsheet. Now you should have 2 chemicals in your spreadsheet.

Add the COX-2 inhibitor Bextra to the spreadsheet using SMILES.

1. Click on the Ligand Editor button
2. Edit/Add Smiles
3. Paste SMILES string
4. Export chemical to spreadsheet

Enter SMILES string (e.g. C(=O)O)
O=S(=O)(N)c3ccc(c2c(oc2c1cccc1)C)cc3

	Value
1 Formula	C16 H14 N2 O3
2 Smiles	O=S(=O)(N)C(=O)C1=CC=C(C=C1)C2=CN(C=C2)C3=CC=CC=C3
3 BLPAC	4-(5-methyl-3-pyridin-2-ylphenyl)isoxazole-3-sulfonamide
4 InChI	InChI=1S/C16H14N2O3
5 InChIKey	LNPTQAFDNC
6 MolWeight	314.0725
7 HBA	5
8 HBD	2
9 RotB	3
10 DrugLikeness	0.0330088
11 MolArea	316.735
12 MolChf	-14.1406
13 MolLogP	3.06632
14 MolLogS	-4.98608
15 MolPSA	72.5302
16 Volume	273.547
17 Bad Groups	
18 Groups	Phenyl Isoxazole Sulfonamide

Step 6: Using SMILES. A convenient way to import chemical data is by using SMILES. Bextra (Valdecoxib) is a COX-2 inhibitor that was removed from the market by Pfizer in 2005 due to concerns of an increased risk of heart attack and stroke. Bextra's SMILES string is:
O=S(=O)(N)c3ccc(c2c(oc2c1cccc1)C)cc3

Use the chemical dictionary to read in the drug Relafen (Nabumetone).

1. Click and start typing "Nabumetone"
2. Export to Chemical Spreadsheet

Enter SMILES string (e.g. C(=O)O)
CC(=O)CCc1ccc2cc(OC)ccc2c1

	Value
1 Formula	C15 H16 O2
2 Smiles	CC(=O)CC(=O)C1=CC=C(C=C1)C2=CC=CC=C2OC
3 BLPAC	4-(8-methoxy-1-naphthyl)butan-1-one
4 InChI	InChI=1S/C15H16O2
5 InChIKey	BLUXXJDCDQOQ
6 MolWeight	228.1150
7 HBA	2
8 HBD	0
9 RotB	4
10 DrugLikeness	-0.746927
11 MolArea	276.562
12 MolChf	-51.083
13 MolLogP	3.82632
14 MolLogS	-4.57031
15 MolPSA	20.8457
16 Volume	240.325
17 Bad Groups	
18 Groups	Ketone (Carbonyl) Ether

Step 7: Using the Chemical Dictionary. Another convenient way to import chemical data is to use the built chemical dictionary. Relafen is Nabumetone is an NSAID that is rapidly metabolized in the liver to a

major active metabolite, 6-methoxy-2-naphthyl acetic acid.

4cox Molsoft icm 3.8-1 [NewProject *] (1 object 1 table)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics Win

display light labels meshes search ligedit

Workspace Panel

no selection

objects (1 item)

- 4cox [1] XR; 2.9A
- a 552 A 3 gl 13 sites PGH2_I
- b 552 A 3 gl 9 sites PGH2_MOU
- c 552 A 3 gl 8 sites PGH2_MOU
- d 552 A 3 gl 8 sites PGH2_MOU
- ahem H protoporphyrin ix conta
- amrn H indomethacin
- bhem H protoporphyrin ix conta

tables (1 item)

- Cox2_inhibitors 4 rows 2 cols 0 header

Extract ligand into chemical table

as 2D drawing keep 3D coordinates

append To Existing Table Cox2_inhibitors

Ok Cancel

3. Enter as shown

1 non-ICM Obj

Step 8: Extracting Chemicals from the PDB. You can extract a chemical from a PDB file into a chemical spreadsheet by right clicking on the ligand in the ICM workspace and choosing the option "Extract Ligand".

4cox Molsoft icm 3.8-1 [NewProject *] (1 object 1 table)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics Win

display light labels meshes search ligedit

Workspace Panel

no selection

objects (1 item)

- 4cox [1] XR; 2.9A
- a 552 A 3 gl 13 sites F
- b 552 A 3 gl 9 sites PGH
- c 552 A 3 gl 8 sites PGH

mol

Molecule Name

Celebrex

Calculate Chemical Properties

Function:

Arguments: Main argument: mol

New column location: mol

Insert: after before in-place at end

New column name: name

Add To List

Function	Name	Category	Description
<input checked="" type="checkbox"/> MolWeight(mol)	molWeight	Chemical	Molecular w...
<input type="checkbox"/> MolFormula(mol)	molFormula	Chemical	Chemical fo...
<input type="checkbox"/> IupacName(mol)	iupacName	Chemical	IUPAC nom...
<input type="checkbox"/> MolLogP(mol)	molLogP	Chemical	Octanol wat...
<input checked="" type="checkbox"/> MolLogS(mol)	molLogS	Chemical	Water solub...
<input type="checkbox"/> MolPSA(mol)	molPSA	Chemical	Polar surfac...
<input type="checkbox"/> MolVolume(mol)	molVolume	Chemical	Molecular v...
<input type="checkbox"/> MolHf(mol)	molHf	Chemical	Heats of for...
<input type="checkbox"/> DrugLikeness(mol)	drugLikeness	Chemical	Empirical dr...
<input type="checkbox"/> Smiles(mol, mode="Unique")	smiles	Chemical	SMILES/SM...
<input type="checkbox"/> BadGroups(mol)	badGroups	Chemical	Unwanted o...
<input type="checkbox"/> Nof_Atoms(mol, atom="")	nof_Atoms	Chemical	Number of ...
<input type="checkbox"/> Nof_Molecules(mol)	nof_Molecules	Chemical	Number of l...
<input type="checkbox"/> Nof_Fragments(mol, fragment="")	nof_fragments	Chemical	Number of f...
<input type="checkbox"/> Nof_Chirals(mol, type="any")	nof_Chirals	Chemical	Number of c...
<input type="checkbox"/> Nof_Rings(mol)	nof_Rings	Chemical	Number of r...
<input type="checkbox"/> Max_Sim_Sim(mol)	max_Sim_Sim	Chemical	1-similarity...

Step 9: Calculate Properties and Insert Columns. Use the Chemistry/Calculate Properties to calculate properties (e.g. logP, logS, MW) for the drugs in the spreadsheet.

In this example we will import more COX-2 inhibitor information from ChEMBL. ChEMBL is a database of bioactive drug-like small molecules curated from the primary scientific literature.

Note: In version 3.8-4 and higher you can search ChEMBL directly from the GUI interface.

The screenshot shows the ChEMBL website interface in a Mozilla Firefox browser window. The address bar displays 'https://www.ebi.ac.uk/chembl/'. The page header includes the ChEMBL logo and the Wellcome Trust logo. The main navigation menu includes 'Compounds', 'Targets', 'Assays', 'Documents', and 'Activity Sources'. The 'Targets' tab is selected, and the search bar contains the text 'cox2'. A red box highlights the 'Targets' tab with the text '1. Search by Target "cox2"'. The search results section is titled 'List Search' and includes options for 'SMILES Search', 'ChEMBL ID Search', and 'Keyword Search'. The 'ChEMBL ID Search' option is selected. Below the search options is a text input field with the placeholder text 'Please enter a list of Compound IDs, keywords, or SMILES separated by newlines'. A 'Fetch Compounds' button is located below the input field. The left sidebar contains various navigation links such as 'Downloads', 'Malaria Data', 'ChEMBL-NTD', 'Kinase SARfari', 'GPCR SARfari', 'DrugEBiRty', 'Web Services', and 'FAQ'. The bottom of the page shows the text 'Transferring data from www.ebi.ac.uk...' and a Zotero logo.

Step 10: Search for COX-2 in ChEMBL.

ChEMBL - Mozilla Firefox

Open up ChEMBL record ID 230.

ChEMBL

Search ChEMBL...

Compounds Targets Assays Documents Activity Sources

ChEMBL Target Search Results: 11

10 records per page

ChEMBL ID	Preferred Name	UniProt Accession	Target Type	Organism	Compounds	Bioactivity
CHEMBL230	Cyclooxygenase-2	P35354	SINGLE PROTEIN	Homo sapiens	5282	8829
CHEMBL4102	Cyclooxygenase-2	P79208	SINGLE PROTEIN	Ovis aries	878	1267
CHEMBL2094253	Cyclooxygenase	P23218	PROTEIN	Homo sapiens	863	1146

ChEMBL Statistics

- DB: ChEMBL_18
- Targets: 9,414
- Compound records: 1,566,998
- Distinct compounds: 1,359,508
- Activities: 12,419,715
- Duplications: 61,308

Step 11: Select Human COX-2 entry.

ChEMBL230 Target Report Card - Mozilla Firefox

Open up the records containing IC50 data

ChEMBL230 Target Report Card

CHEMBL622	ETOUJULAL	Cyclooxygenase-2 inhibitor	PubMed
CHEMBL699	MELOXICAM	Cyclooxygenase-2 inhibitor	PubMed
CHEMBL1070	NABUMETONE	Cyclooxygenase-2 inhibitor	DailyMed
CHEMBL122	ROFECOXIB	Cyclooxygenase-2 inhibitor	FDA
CHEMBL865	VALDECOXIB	Cyclooxygenase-2 inhibitor	DailyMed

Target Associated Bioactivities

ChEMBL Activity Types for Target CHEMBL230

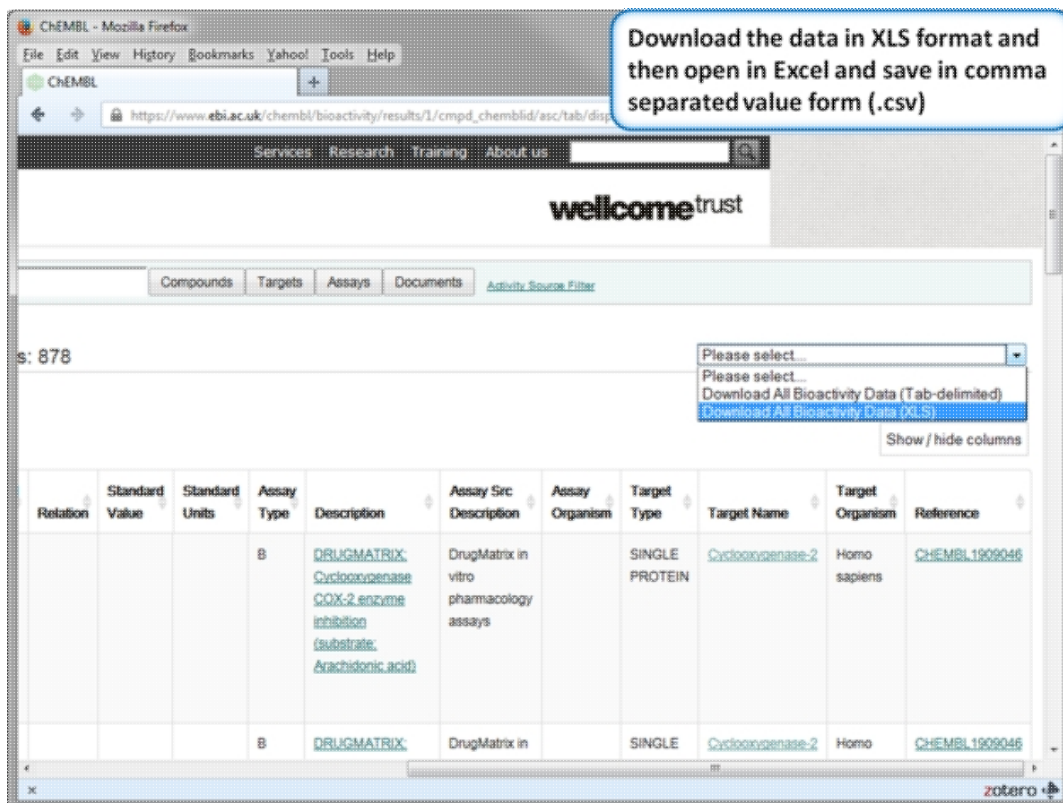
Total: 8829

Target Associated Assays

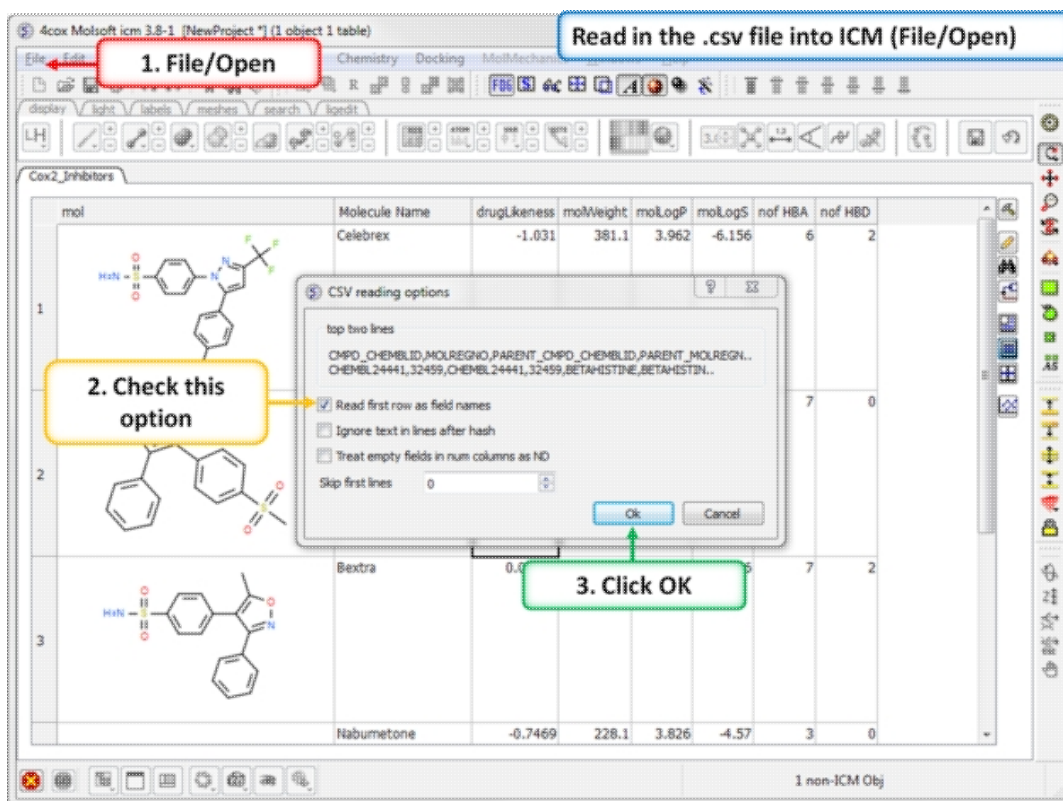
ChEMBL Assays for Target CHEMBL230

Transferring data from www.ebi.ac.uk...

Step 12: Select chemicals annotated with Ki values.



Step 13: Download in XLS format.



Step 14: Convert to csv format. Open the .xls file in Excel and save in comma separated value format (.csv). Read in the .csv file into ICM using File/Open. Rename the table Cox2_inhibitors - right click on table header tab and choose rename.

4cox Molsoft icm 3.8-1 [NewProject *] (1 object 2 tables)

Convert the SMILES string into 2D sketch.

1. Chemistry/Convert/Smiles to 2D

.LOGP	PSA	NUM ROS VIOLATIONS	CANONICAL SMILES	ACTIVITY ID	STANDARD TYPE	REL
1	.68	24.92	0	CNCCc1ccccc1	7613656	KI
2	.63	64.79	0	CCCCc1ccc(cc1N)C(=O)OCCN(C)CC	7732956	KI
3	.9	138.53	0	Cc1cn(c(n1)C(=O)NCCCc2ccc(cc2)S(=O)(=O)C	7668268	KI
4	.98	202.26	1	CC(C)[C@H](N(C(=O)N(C)C)C1csc(n1)C(C)C)C(=O)N[C@H]	7809126	KI
5	.4	78.49	0	CN1C(=O)CN=C(C2CCCC2)C3CC(CCC13)[N+](=O)[O-]	7722813	KI
6	.99	62.66	0	CN(C1CC(CCC1)C(=O)N)CCC(O)C2CC(O)CCC2	7688611	KI
7						
8	.78	40.46	0			
9	.9	36.86	0			
10	.75	91.76	1			
11	.1	123.79	1			
12	.94	56.73	0			
13						
14	.61	37.29	0	CC(C)Cc1ccc(cc1)C(C)C(=O)O	7662384	KI
15	.21	64.09	0	COH1C=C(C(=O)O)C(=O)C2C	7727701	KI
16	.88	29.54	0	COc1ccccc1C(=O)N2CCC(Cc3	7636680	KI
17	.75	63.32	0	NC(C@@H)1CC(C@@H)(CC1)C(=O)O	7785816	KI
18	.5	26.02	0	NC12CC3CC(C(C)3)C1)C2	7625398	KI
19	.41	72.88	0	CC1OH(COH1)C2CC3N(C=C(C(=O)O)C(=O)C3c(C)C2F)O4C	7676976	KI
20	.54	68.36	0	C[C@@H](CO)NC(=O)[C@H]1CN(C)	7667088	KI
21	.31	63.4	0	NC(=O)CN1CCCC1=O	7713348	KI
22				CN(C@H)(CC(C)C)C(=O)N[C@@H]1[C@H]	7787337	KI
23	.13	65.84	0	O=C1N=C1=C2NHC=C12	7623701	KI
24	.22	72.83	0	CC(C@@H)(C)C(=O)O[C@H]1C[C@@H](C)C=C2C=C[C@H]	7753250	KI
25	.2	118.21	1	CN1C[C@H]	7698463	KI

Smiles to 2D

Table: chembl_cox2 Smiles Column: CANONICAL_SMILES

Keep Original Column:

2. Enter as shown

Step 15: SMILES to 2D. Convert the SMILES column into 2D. This may take a little bit of time because the table is quite large. We can use this table in the following tutorial.

19.6.2 Chemical Searching

Background

This tutorial follows on from the previous tutorial where we made a database of COX-2 inhibitors (Cox2_Inhibitors) and read in a chemical spreadsheet containing IC50 data from ChEMBL (chembl_cox2). In this tutorial we will perform chemical spreadsheet comparison and explore some of the chemical search options inside ICM.

Tutorial

Let us compare our small spreadsheet of COX-2 drugs with the larger set of inhibitors from ChEMBL.

1. Chemistry/Compare two sets

Compare Two Sets

First Table: Cox2_inhibitors Second Table: chembl_cox2

Select Similarity

Distance: 0.5

Select Matching Records Show Only Different Records

Hint: Double-click on the tab of the second table to see them side by side

2. Enter as shown

Step 1: Similarity search - compare two spreadsheets. Here we can compare our two chemical spreadsheets using Chemistry/Compare two sets using a similarity search.

Let us make a new spreadsheet from the comparison hits. Selected compounds will be highlighted in blue.

Table

New table name: similar_0_5

1. Right click and choose Copy rows to ICM Table

2. Enter a new table name for the similar compounds

Step 2: Output the similar chemicals to a new spreadsheet. Similar compounds are selected (shaded blue) we can extract them to a new table using the right click menu.

Save the spreadsheets in SDF format and delete Cox2_Inhibitors and similar_0_5

mol	Molecule Name	drugLikeness	molWeight	molLogP	molLogS	nof HBA	nof HBD
1		-1.031	381.1	3.962	-6.156	6	2
2	Voxx	-0.7455	314.1	2.205	-4.957	7	0
3	Bextra	0.03301	314.1	3.066	-4.986	7	2
	Nabumetone	-0.7469	228.1	3.826	-4.57	3	0

Step 3: Save as SDF. You can save any chemical spreadsheet by right clicking on the table header tab.

Chemical Searching

CMPD CHEMBL	MOLREGNO	PARENT CMPD CHEMBLID	PARENT MOLREGNO	MOL PREI
1		32459 CHEMBL24441		32459 BETAHIST
CHEMBL1196	209007	CHEMBL1196		209007 PROPARA
CHEMBL1073	139286	CHEMBL1073		139286 GLIPIZIDE

Step 4: Chemical Searching. Now let us explore the chemical search options.

We will start by sketching a sulfonamide group and setting ring membership as R1 and an Attachment Point on the Carbon atom.

1. Right click on atom

2. Select Attachment point

3. Select ring membership R1

Step 5: Sketch search query. We will start by sketching a sulfonamide group and setting ring membership as R1 and an Attachment Point on the Carbon atom.

Your query should look like this.

1. Our query is the sketch

2. Table to be searched

3. Select substructure search

4. Results will return to a new table called "result"

5. Click Search

Step 6: Chemical search window. Choose the options as shown in the screenshot and press search.

Results are shown with the query highlighted in green. We are finding chemicals with sulfonamide linked between two substituents e.g Phenyl-Sulfonamide-Thiazole shown - let us filter these out (next step).

CMPD	CHEMBLID	mol	MOLREGNO
1	CHEMBL21		
2	CHEMBL20		1125 CHEI
3	CHEMBL437		2095 CHEI

Step 7: Search query results. The results will be displayed in a table as shown.

1. Right click on the Nitrogen atom and choose Connectivity "As Drawn"

By changing the connectivity of the Nitrogen we can find only terminal sulfonamide groups similar to the drug Celebrex.

ICM Chemical Search: [new file *] Molcart not connected

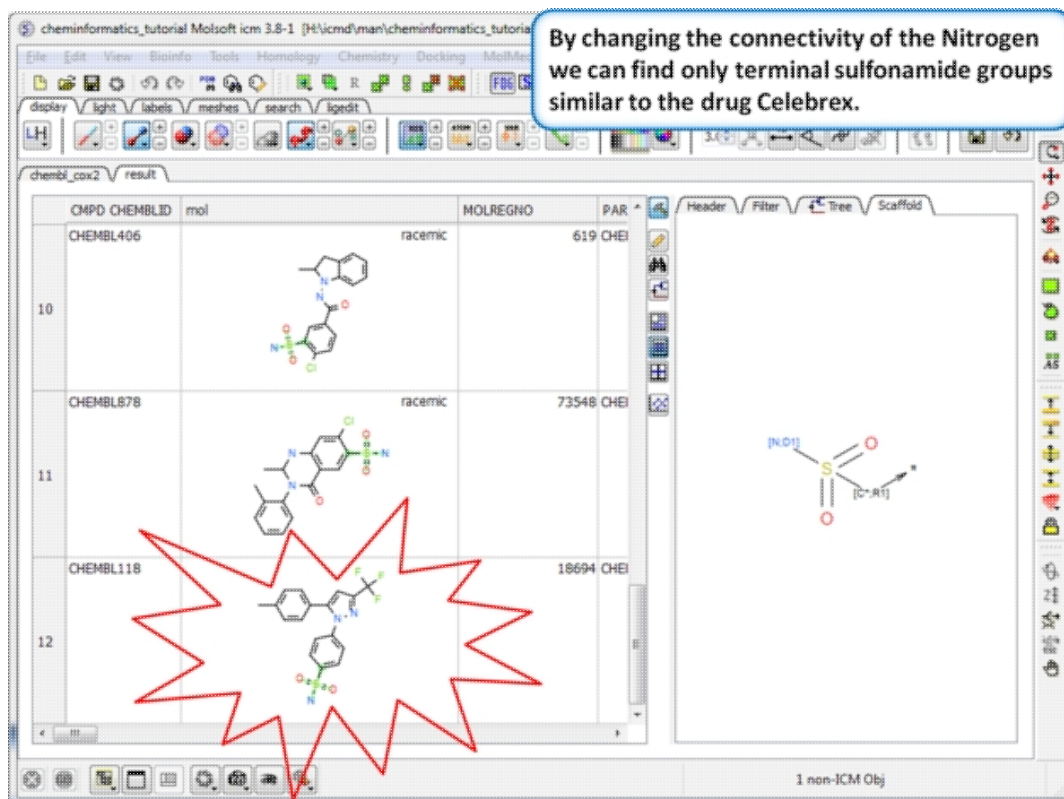
CH5NO2S

Data Source: Table File Molcart
chembl_cox2

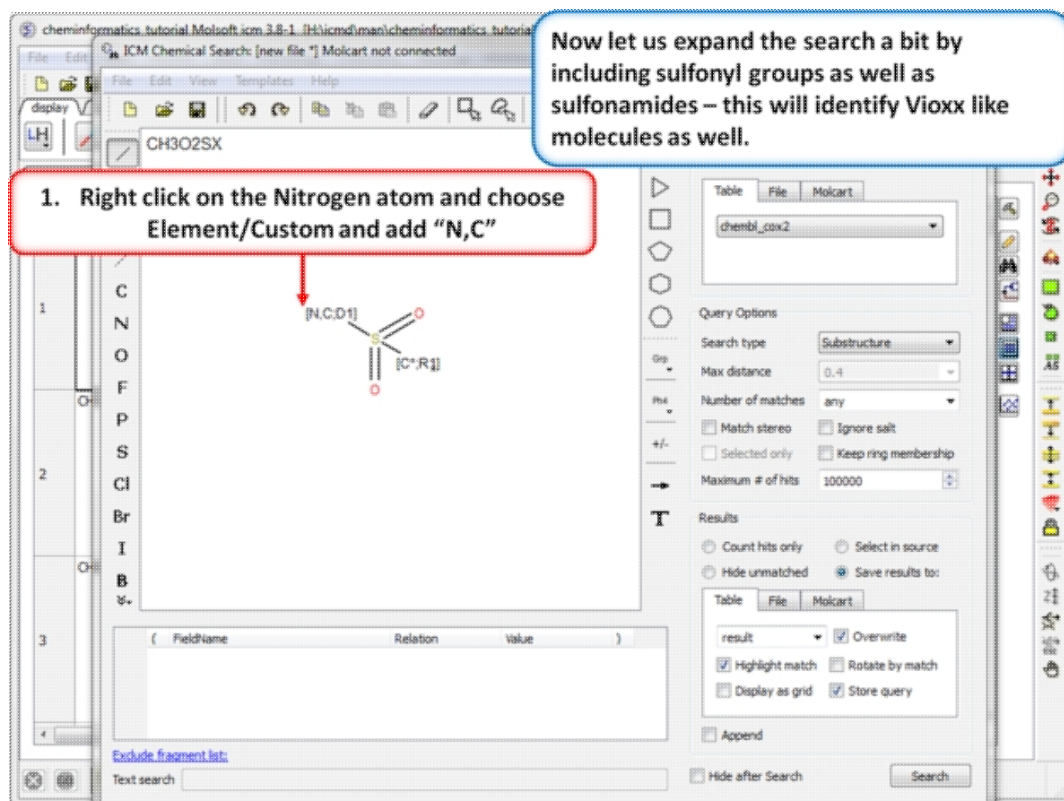
Query Options
Search type: Substructure
Max distance: 0.4
Number of matches: any
Match stereo: Ignore salt:
Selected only: Keep ring membership:
Maximum # of hits: 100000

Results
Count hits only: Select in source:
Hide unmatched: Save results to:
Table File Molcart
result Overwrite
 Highlight match Rotate by match
 Display as grid Store query
 Append
 Hide after Search

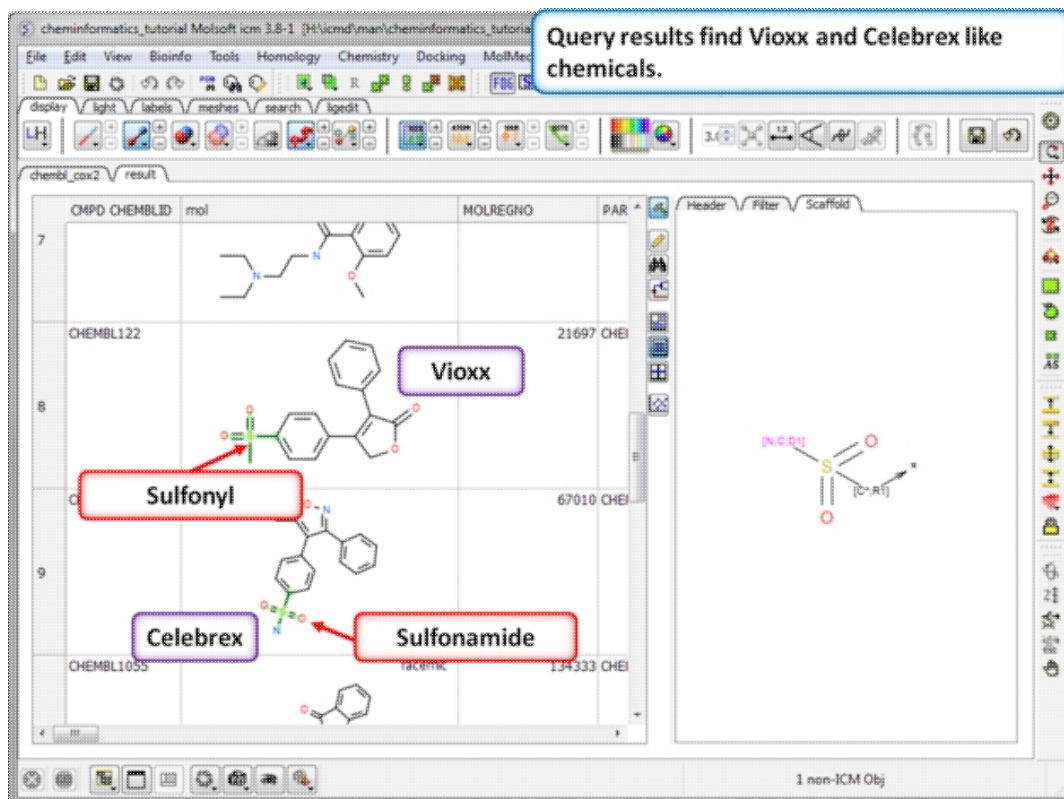
Step 8: Query filter - connectivity. By changing the connectivity of the Nitrogen we can find only terminal sulfonamide groups similar to the drug Celebrex.



Step 9: Query filter results By changing the connectivity of the Nitrogen we can find only terminal sulfonamide groups similar to the drug Celebrex.



Step 10: Query filter - Custom Element. Now let us expand the search a bit by including sulfonyl groups as well as sulfonamides - this will identify Vioxx like molecules as well.



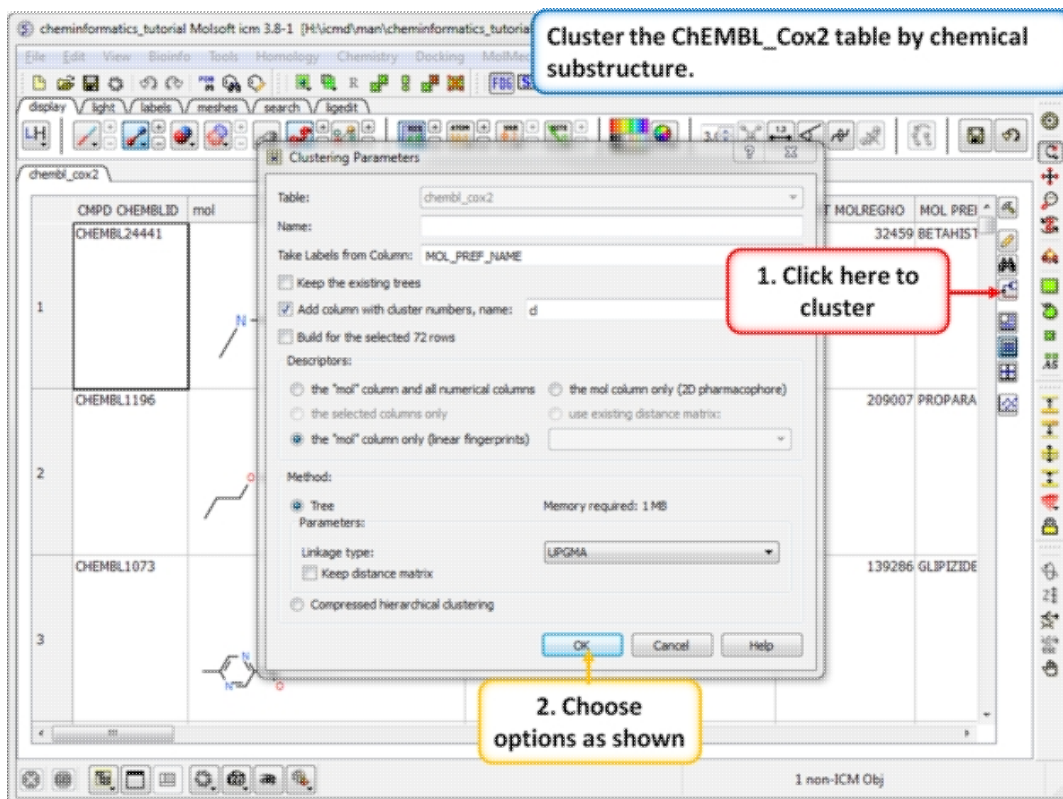
Step 11: Query filter results. By adding to the filters we can identify Vioxx-like chemicals containing sulfonyls and Celebrex-like chemicals containing sulfonamide groups.

19.6.3 Chemical Clustering

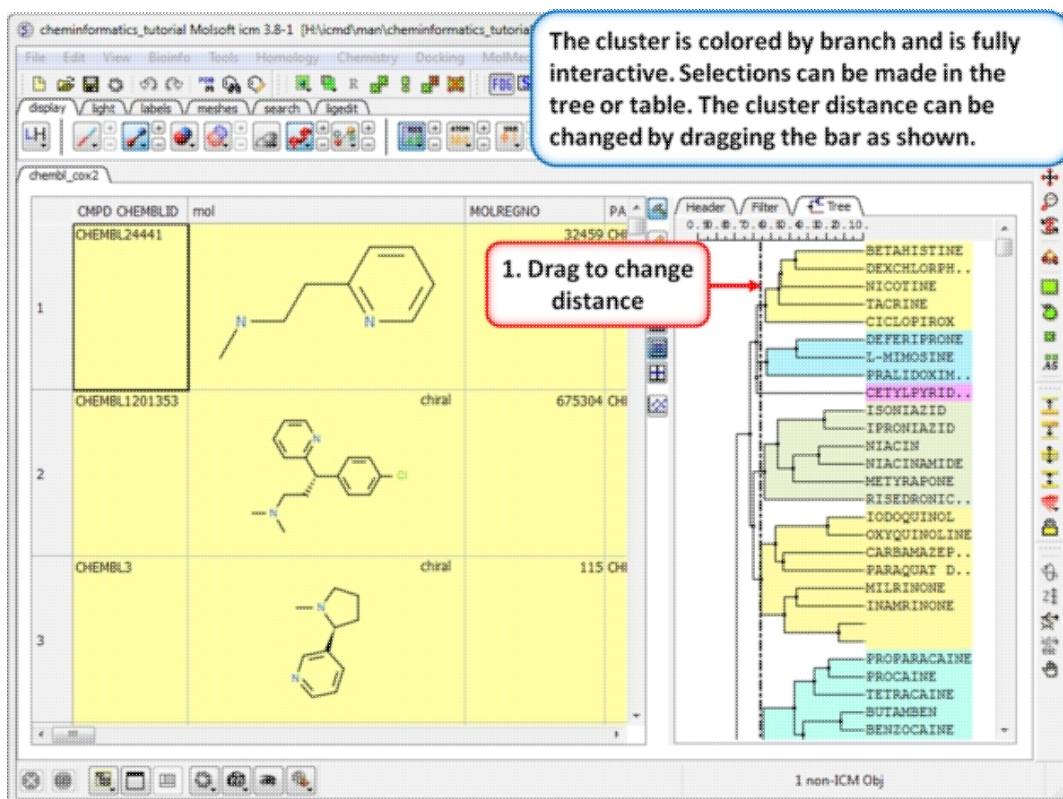
Background

This tutorial follows on from the previous tutorial where we made a database of COX-2 inhibitors (Cox2_Inhibitors) and read in a chemical spreadsheet containing IC₅₀ data from ChEMBL (chembl_cox2). In this tutorial we will perform chemical clustering on the ChEMBL chemical spreadsheet.

Tutorial

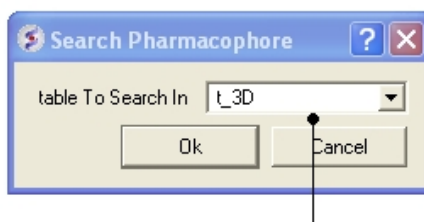
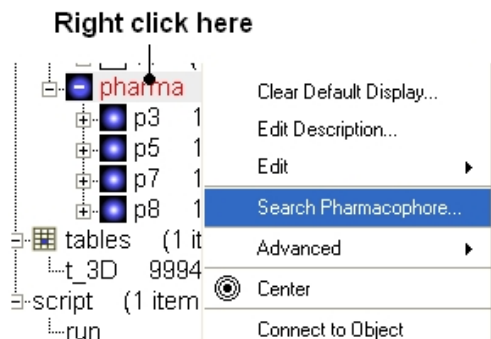


Step 1: Cluster Cluster the ChEMBL_Cox2 table by chemical substructure using the options shown in the screenshot.



Step 2: Change the distance of the tree. The cluster is colored by branch and is fully interactive. Selections can be made in the tree or table. The cluster distance can be changed by dragging the bar as shown.

- In this example the 3D pharmacophore has already been extracted from a ligand. To find out how to generate a 3D pharmacophore see the section entitled *Pharmacophore Draw 3D*.
- In this example a table containing 3D coordinates is already provided containing 3D coordinates. The table is called *t_3D*.
- To run 3D pharmacophore searching right click on the name of the pharmacophore object in the ICM Workspace and select **Search Pharmacophore**.
- Select the table *t_3D* from the drop down list and click OK.
- A table of search results will be displayed.



Enter name of table containing 3D coordinates

- You can browse the results by clicking on the table and the ligand will be displayed in the graphical display.
- Remember you can use the check boxes in column *L* to lock compounds and overlay them.

Browse and lock compound in 3D display

mol	NAME_	molid	confid	L	rmsd	score
1	chiral,3D->2D	822	189664	<input checked="" type="checkbox"/>	0.00469	100
	chiral,3D->3D	824	190184	<input type="checkbox"/>	0.007595	100

Double click to view in 3D

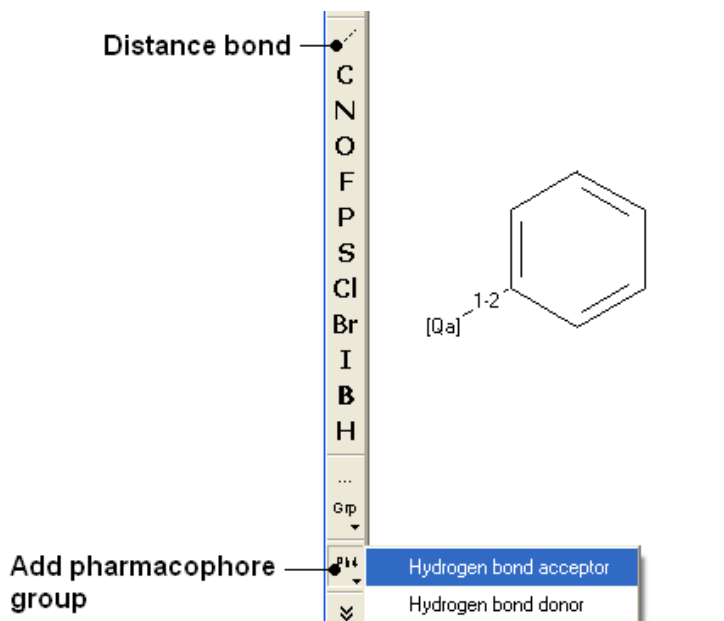
Manual References (Web Links)

3D Pharmacophore Search

19.6.5 2D Pharmacophore Searching

Objective Undertake a 2D pharmacophore search of a chemical spreadsheet.

- File/Open **example_ph4.icb** (this file is provided in the ICM distribution and therefore can be found in \$ICMHOME or in Windows Program Files/MolSoft)
- Chemistry/Chemical Search
- Draw the query as shown below using the



- Fill in the query and results option as shown below.

Data Source

Local Tables MolCart

Query Options

Database

Table t_3D

Search Type Substructure

Maximal # of Hits 1000

Max Distance 0.4

Number of Matches any

Match stereo Ignore Salt

By Selection # Of Hits Only

Result

sult_new Auto Append

New Table Select in Source

Add to DB

Highlight Match Rotate by Match

Display as Grid

Search

Relation	Value

Manual References (Web Links)

2D Pharmacophore Search

19.6.6 How to create a Markush structure.

How to create a Markush structure. (1)

1. Sketch a Markush structure in the molecule editor and save it as a chemical spreadsheet

Remember to add R-groups to your scaffold.

How to create a Markush structure. (2)

Read in a set of chemical substituents for each of your R-groups in your Markush scaffold

ICM comes with some built in substituents if you want to use those.

How to create a Markush structure. (3)

1. Chemistry/Create/Modify Markush

2. Enter the name of the table and the row number containing your Markush structure (scaffold)

Info: table markush (0 headers, 1 rgrp) has been created
 Headers:
 Access: 1 markush.mol
 low/dee: delete markushName
 low/dee: delete markush
 low/dee: []

How to create a Markush structure. (4)

1. Enter the table names for each R-group

2. Click "Create"

```
low/dee: delete markushName
low/dee: delete markush
low/dee: molid = molFold.mol( 1 )
low/dee: markushName = Show( "markush" unique )
low/dee: a_out2 = Ahq( molid, enumerate link )
```

How to create a Markush structure. (5)

```

lsm/de> if ( #_out2 != "" ) group table $molname molid "mol"
Info: table molmarkush (0 headers, 1 array) has been created
Headers:
Array: molmarkush.mol
lsm/de> delete molname$mol
lsm/de>
  
```

19.6.7 How to enumerate a Markush library.

How to enumerate a Markush chemical library. (1)

1. Sketch a Markush structure in the molecule editor and save it as a chemical spreadsheet

```

Info: 1 table written to file V:\training course\lcm-chemsit-web\lcm-1209\R1.mol
lsm/de> write table mol compress R2 "V:\training course\lcm-chemsit-web\lcm-1209\R2.mol"
Info: 1 table written to file V:\training course\lcm-chemsit-web\lcm-1209\R2.mol
lsm/de> write table mol compress R3 "V:\training course\lcm-chemsit-web\lcm-1209\R3.mol"
Info: 1 table written to file V:\training course\lcm-chemsit-web\lcm-1209\R3.mol
lsm/de>
  
```

Remember to add R-groups to your scaffold.

How to enumerate a Markush chemical library. (2)

Read in a set of chemical substituents for each of your R-groups in your Markush scaffold

ICM comes with some built in substituents if you want to use those.

How to enumerate a Markush chemical library. (3)

1. Chemistry/Enumerate by Scaffold

2. Enter the name of the table and the row number containing your Markush structure (scaffold)

How to enumerate a Markush chemical library. (4)

1. Enter the table names for each R-group

2. Click "Enumerate" and a new table of compounds will be displayed called "enum_result"

```

low/de> molid = molEtoId.mol[ 1 ]
low/de> a_out2 = Anq[ molid, enumerate ]
low/de> if ( a_out2 != "" ) enumerate library molid fa_out2 name = Name( "enum_result" unique )
low/de> molid = molEtoId.mol[ 1 ]
low/de> a_out2 = Anq[ molid, enumerate ]
  
```

19.6.8 How to decompose a library based on a Markush structure.

How to decompose a library based on a Markush structure. (1)

1. Sketch a Markush structure in the molecule editor and save it as a chemical spreadsheet

2. Read in the table you wish to decompose

```

low/de> if ( a_out2 != "" ) enumerate library molid fa_out2 name = Name( "enum_result" unique )
low/de> delete R1
low/de> delete R2
low/de> delete R3
low/de> create enum_result Name(Name("binding_data" a_inq), unique)
low/de>
  
```

How to decompose a library based on a Markush structure. (2)

1. Chemistry/R-Group Decomposition

2. Choose table with scaffold and the row number

3. Choose the table for decomposition and the name of the column containing the 2D sketch

4. Click OK and a new table called "T_sar" will be displayed

```

[cmd/de> delete R2
[cmd/de> remove name_result from (name ("binding_data" *single), unique)
[cmd/de> merge mol1 = molE016.mol1 1
[cmd/de> make minor binding_data.mol1 mol18 sort auto append
[cmd/de> sort T_sar.R1 T_sar.R2 T_sar.R3
[cmd/de>

```

19.6.9 How to enumerate a chemical library by reaction.

How to enumerate a chemical library by reaction. (2)

1. Read in a table of your reactants

Hanfzisch Dihydropyridine (Pyridine) Synthesis

This reaction shows the preparation of dihydropyridine derivatives by condensation of an aldehyde with two equivalents of β -ketoesters in the presence of ammonia. Subsequent reduction (or dehydrogenation) gives pyridine 3,5-dialkylolates, which may also be dealkylated to yield the corresponding pyridines.

Click here for a reaction.

ID	mol1	mol2	mol3	mol4
0798	116			
0799	118			
08213	150			
08250	132			
0798	116			
48281	1421			

```

[cmd/de> undisplay window
[cmd/de> a_sourcelist reactants = **
[cmd/de> openfile "C:\Program Files\MOE\LLCCM-Pro\example_reactions.icb" 1 yes no no **
Info: 6 shell objects read (skipped 3) from C:\Program Files\MOE\LLCCM-Pro\example_reactions.icb
[cmd/de>

```

example_reacton1.icb Molsoft Chem3D 3.6.1g [C:\Program Files\Molsoft\LLCGM-Pro\example_reacton1.icb] (table)

How to enumerate a chemical library by reaction. (1)

1. Sketch reactants and products in the molecule editor

File Edit View Tools Chemistry Windows Help

Hantzsch Dihydropyridine (Pyridine) Synthesis

This reaction allows the preparation of dihydropyridine derivatives by condensing an aldehyde with two equivalents of a 2-ketone in the presence of ammonia. Subsequent oxidation (or dehydrogenation) gives pyridine 2,5-dicarbonyls, which may also be decarboxylated to yield the corresponding pyridine.

Click here to try a reaction.

```

c:\m\de> volapijey window
c:\m\de> _m_currentProject = **
c:\m\de> openFile "C:\Program Files\Molsoft\LLCGM-Pro\example_reacton1.icb" 1 yes no no **
Info: 6 shell objects read (skipped 1) from C:\Program Files\Molsoft\LLCGM-Pro\example_reacton1.icb
c:\m\de>
  
```

Table 1 row, 1 column

example_reacton1.icb Molsoft Chem3D 3.6.1g [C:\Program Files\Molsoft\LLCGM-Pro\example_reacton1.icb] (table)

How to enumerate a chemical library by reaction. (3)

1. Chemistry/Enumerate by Reaction

2. Enter the name of the table and the row number containing your reaction sketch (Reaction)

File Edit View Tools Chemistry Windows Help

Hantzsch Dihydropyridine (Pyridine) Synthesis

This reaction allows the preparation of dihydropyridine derivatives by condensing an aldehyde with two equivalents of a 2-ketone in the presence of ammonia. Subsequent oxidation (or dehydrogenation) gives pyridine 2,5-dicarbonyls, which may also be decarboxylated to yield the corresponding pyridine.

Click here to try a reaction.

Enumerate by Reaction...

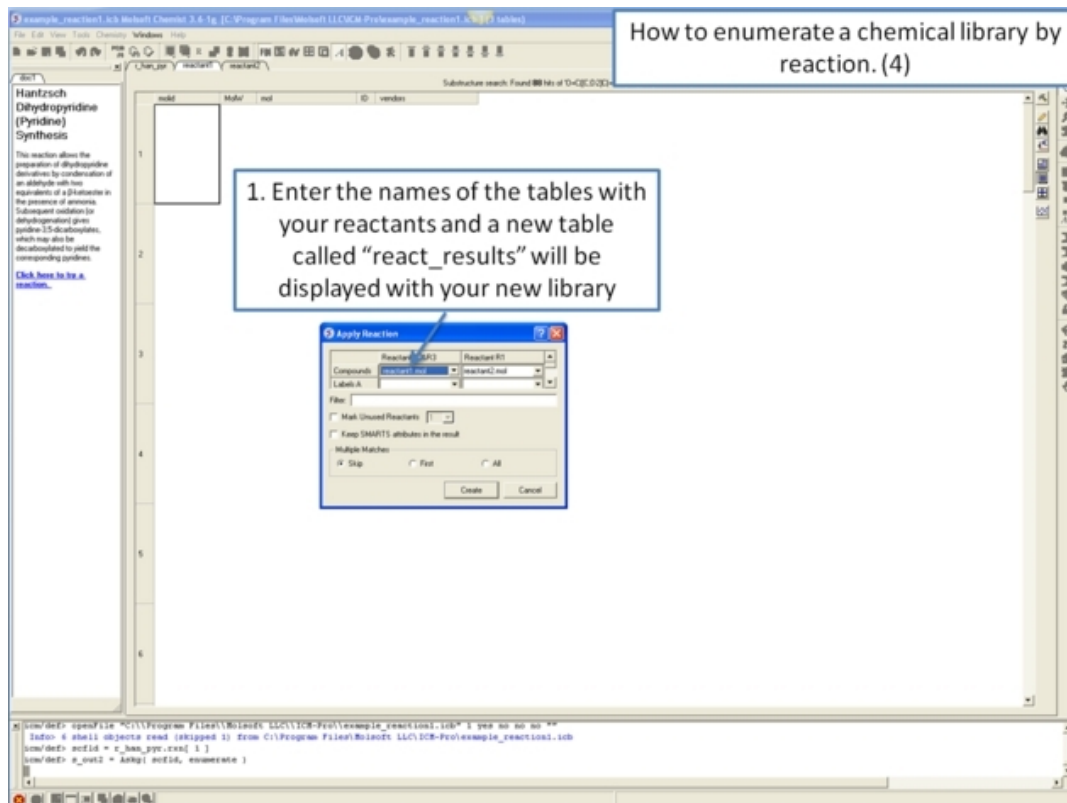
Enumerate by Reaction

Choose Table With Reaction:

MolID	mol	vender
136		apocentric 2039 mol 50008171 www.3751781-0
138		www.36279-09-5
6520	132	mol 700
6765	138	
4021	1421	mol 1

```

c:\m\de> volapijey window
c:\m\de> _m_currentProject = **
c:\m\de> openFile "C:\Program Files\Molsoft\LLCGM-Pro\example_reacton1.icb" 1 yes no no **
Info: 4 shell objects read (skipped 1) from C:\Program Files\Molsoft\LLCGM-Pro\example_reacton1.icb
c:\m\de>
  
```

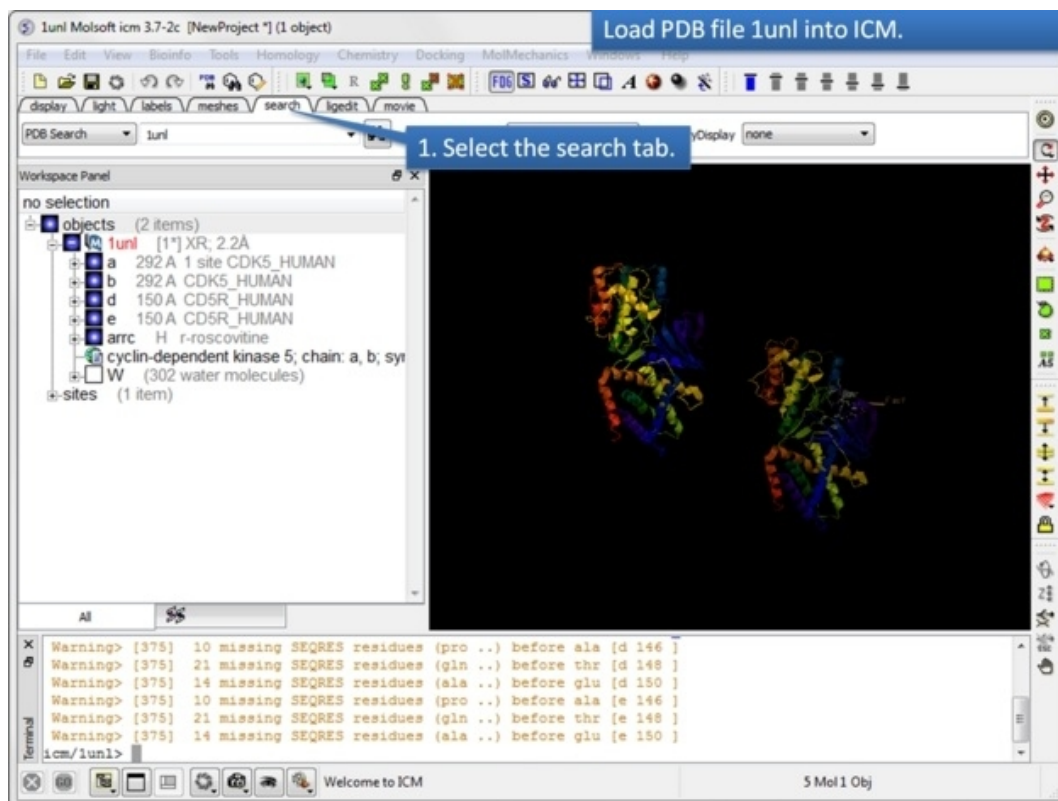


19.7 3D Ligand Editor Tutorial

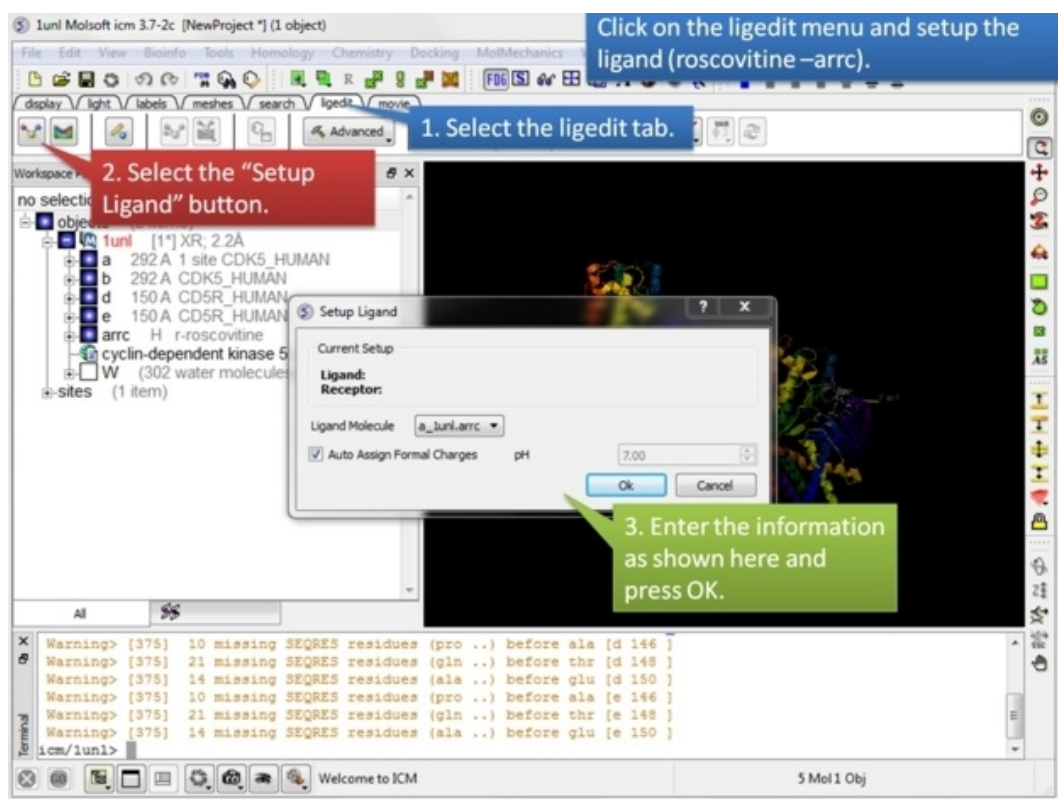
In this example we will use the 3D interactive editor to modify a cyclin-dependent kinase 5 inhibitor. A hydrophobic cavity not occupied by the crystal structure ligand can be exploited to improve the interaction between the ligand and the receptor.

This tutorial teaches you how to:

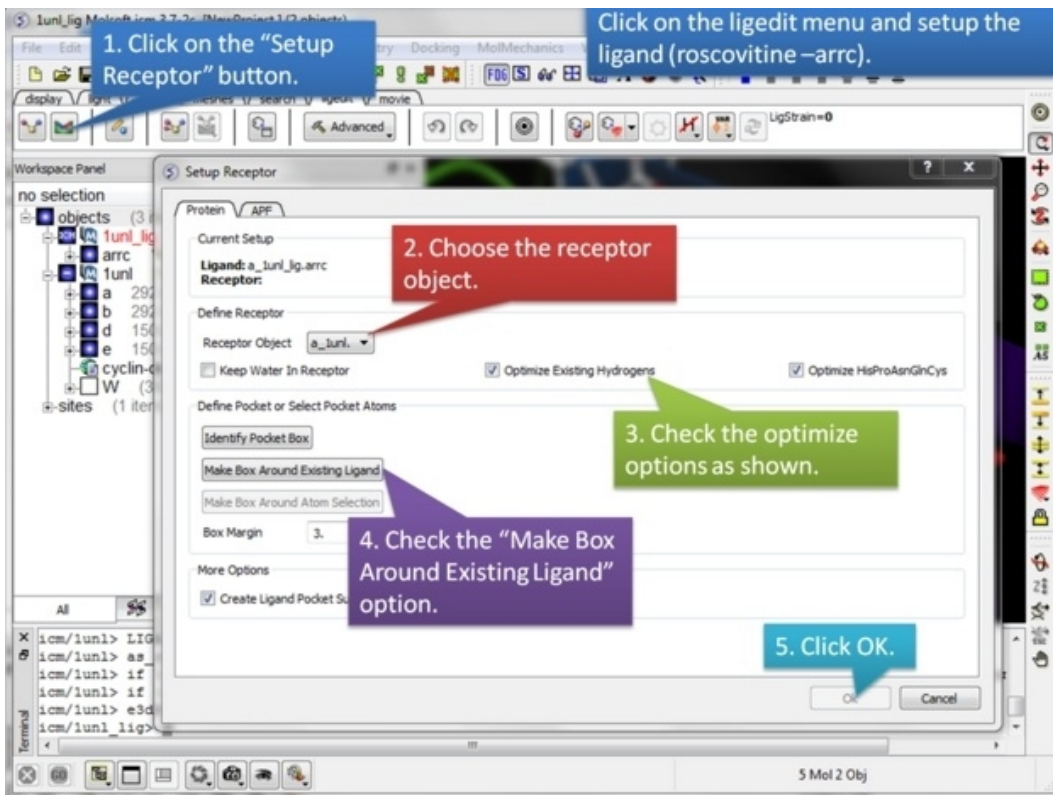
- How to setup a ligand in the editor.
- How to setup the receptor in the editor.
- How to display surfaces and hydrogen bonds.
- How to minimize the ligand.
- How to make modifications to the ligand.
- How to tether atoms.
- How to dock.
- How to screen a database of substituent groups.
- How to minimize side-chains.



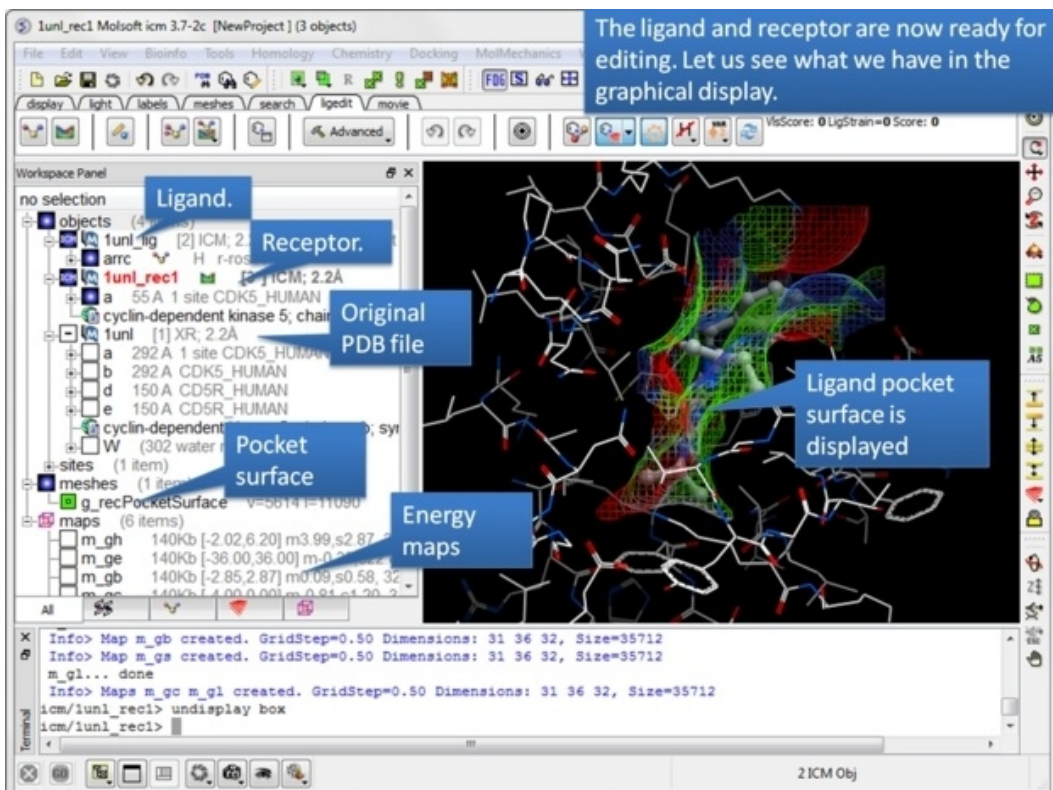
Step 1: Read in the PDB file 1UNL.



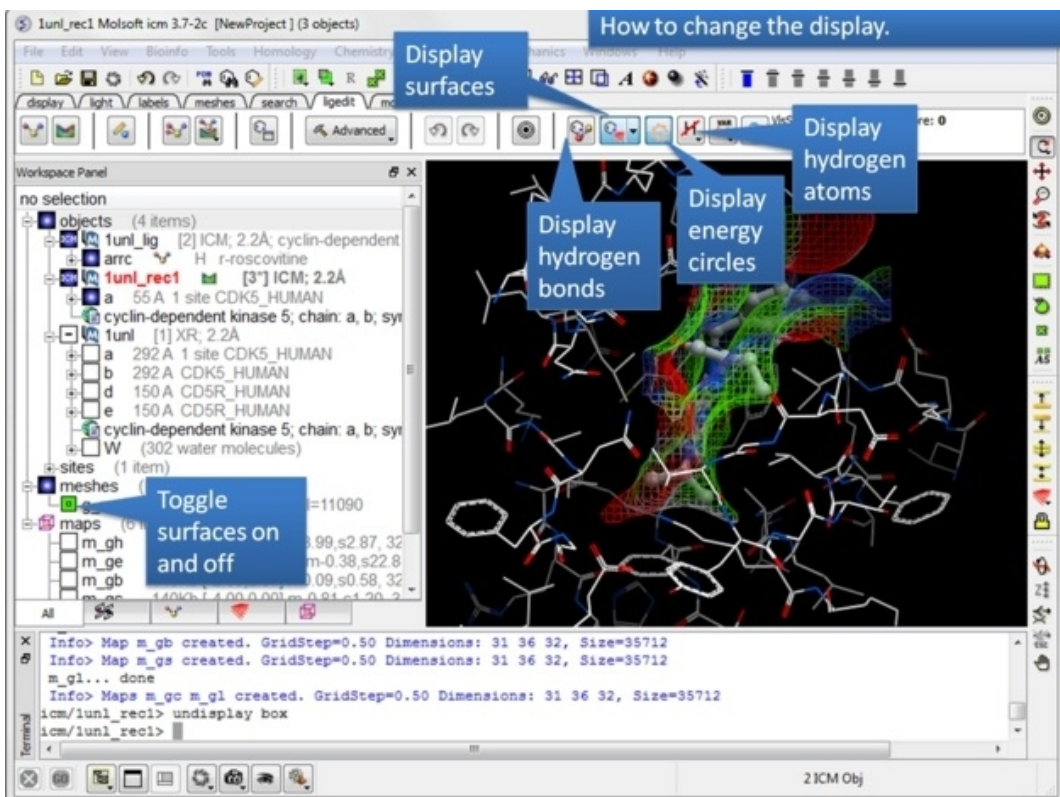
Step 2: Setup the ligand in the 3D ligand editor.



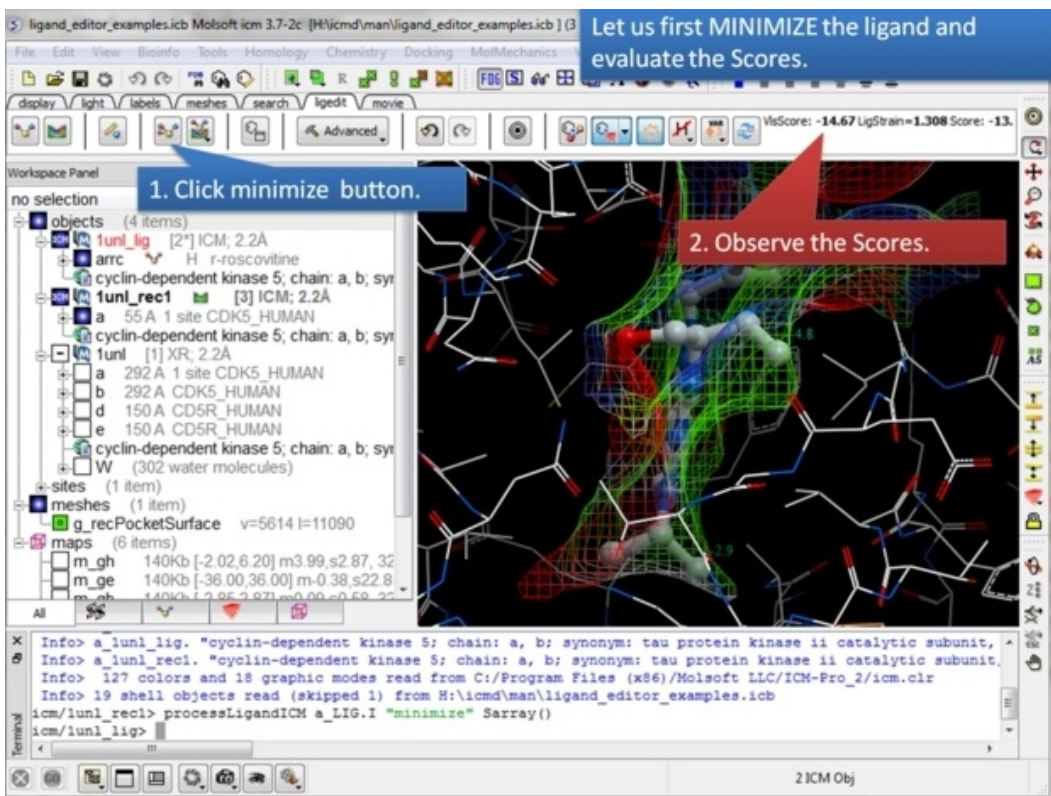
Step 3: Setup the receptor in the 3D ligand editor.



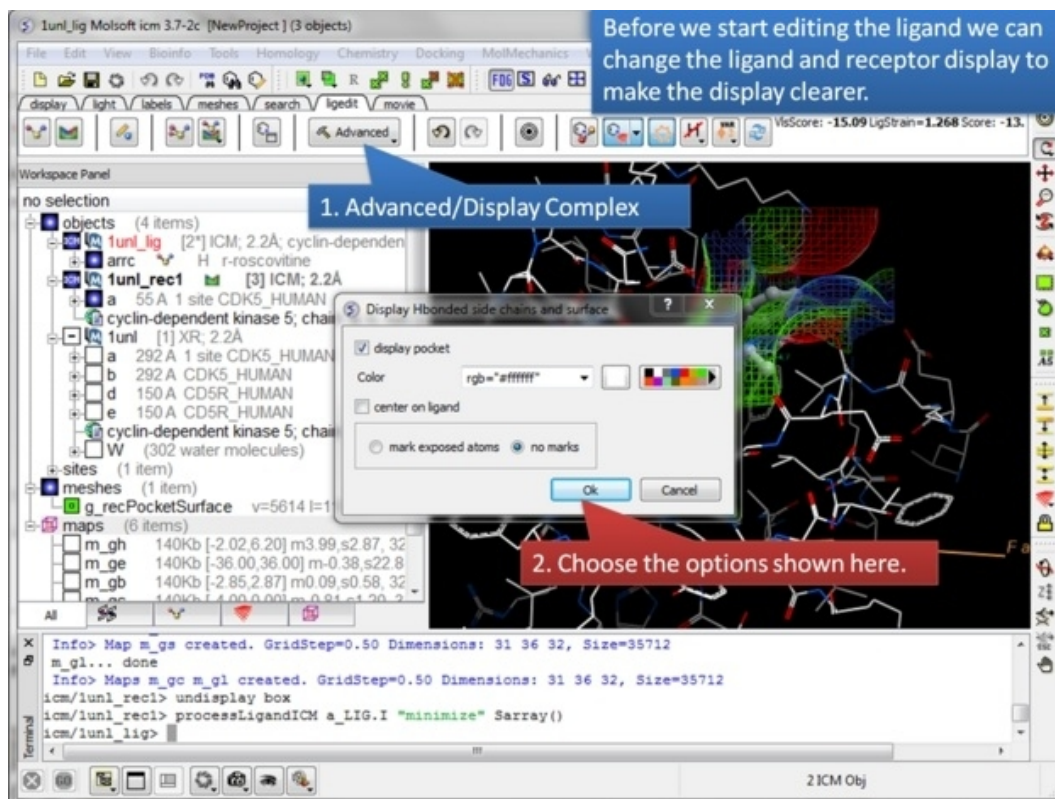
Step 4: Observe the new objects, meshes and maps that have been created.



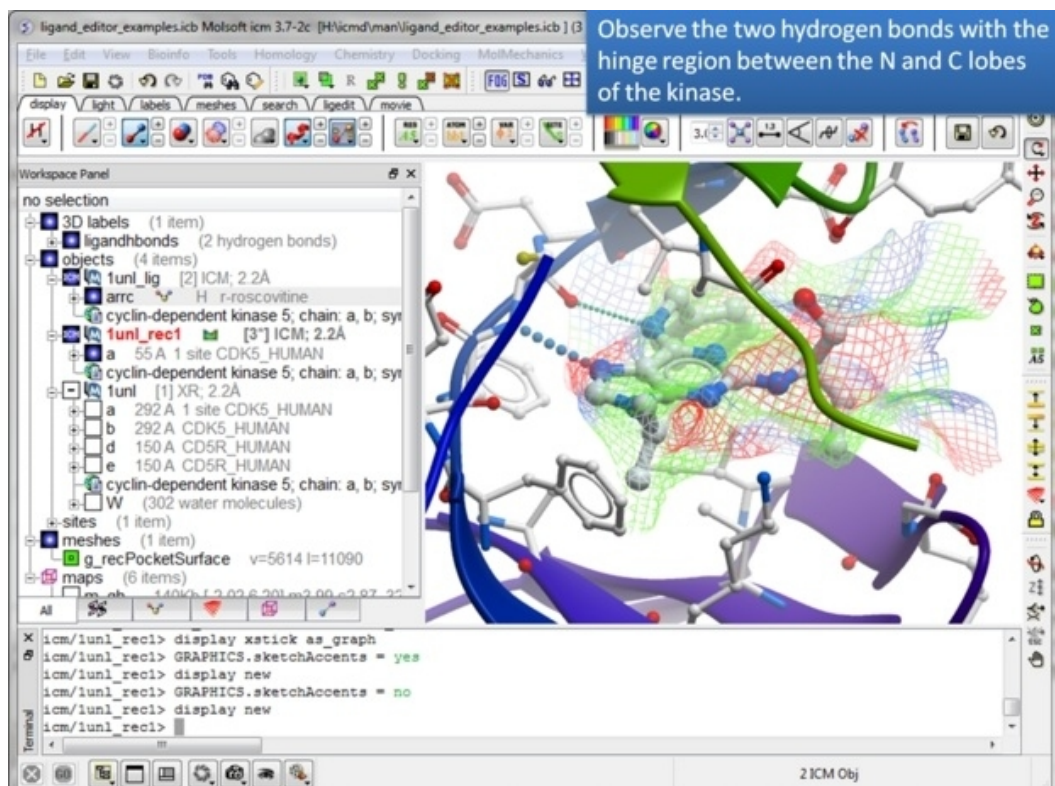
Step 5: You can display and undisplay, hydrogen atoms, surfaces, and hydrogen bonds.



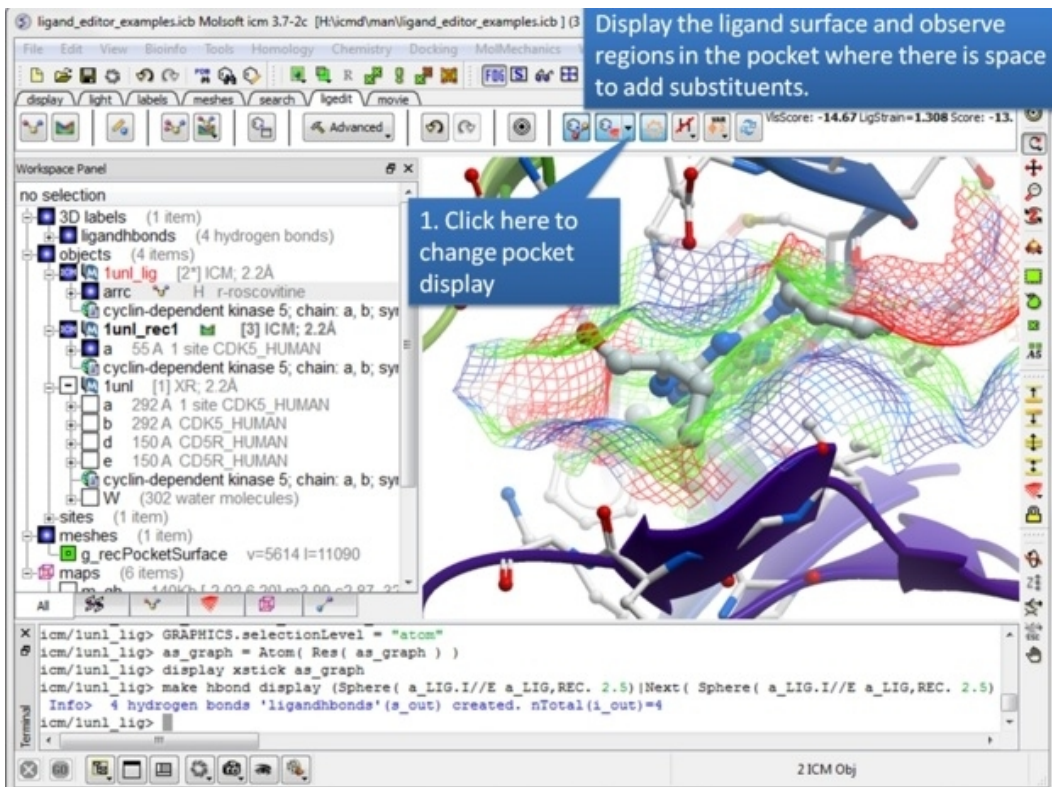
Step 6: First minimize the ligand and observe the Scores.



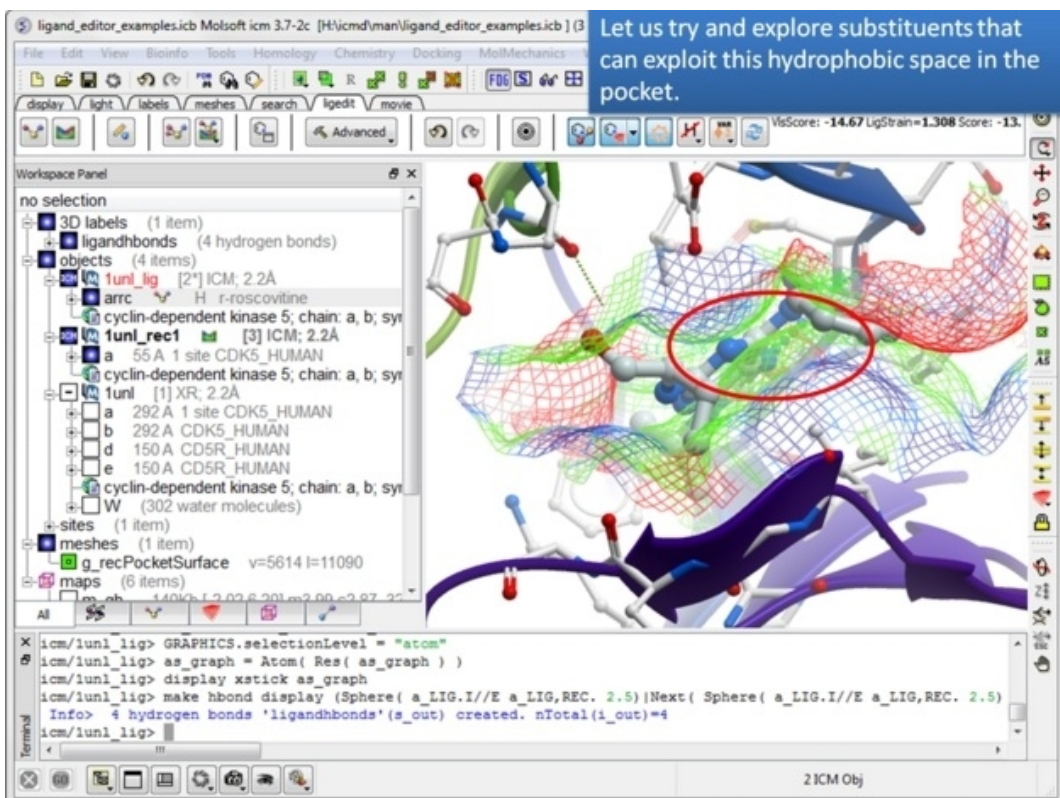
Step 7: In the Advanced menu there is an option to make a clearer view of the complex.



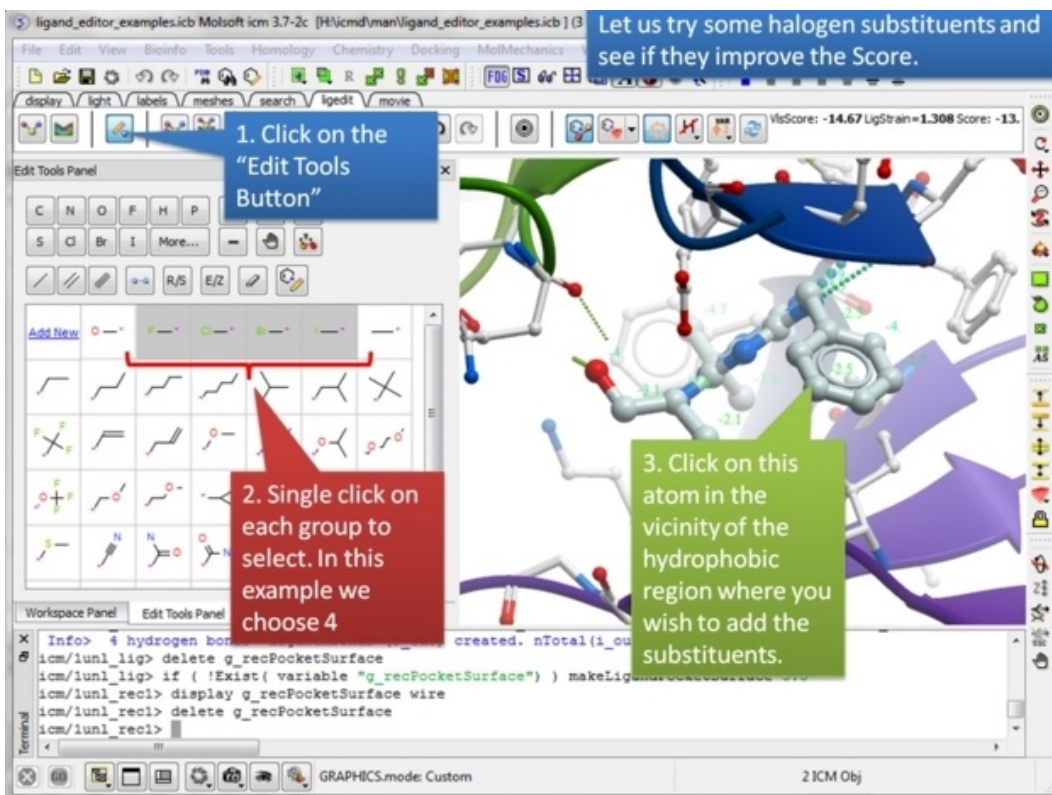
Step 8: Observe that the ligand makes key hydrogen bond interactions with the hinge region of the kinase. It is important to conserve these interactions when making modifications.



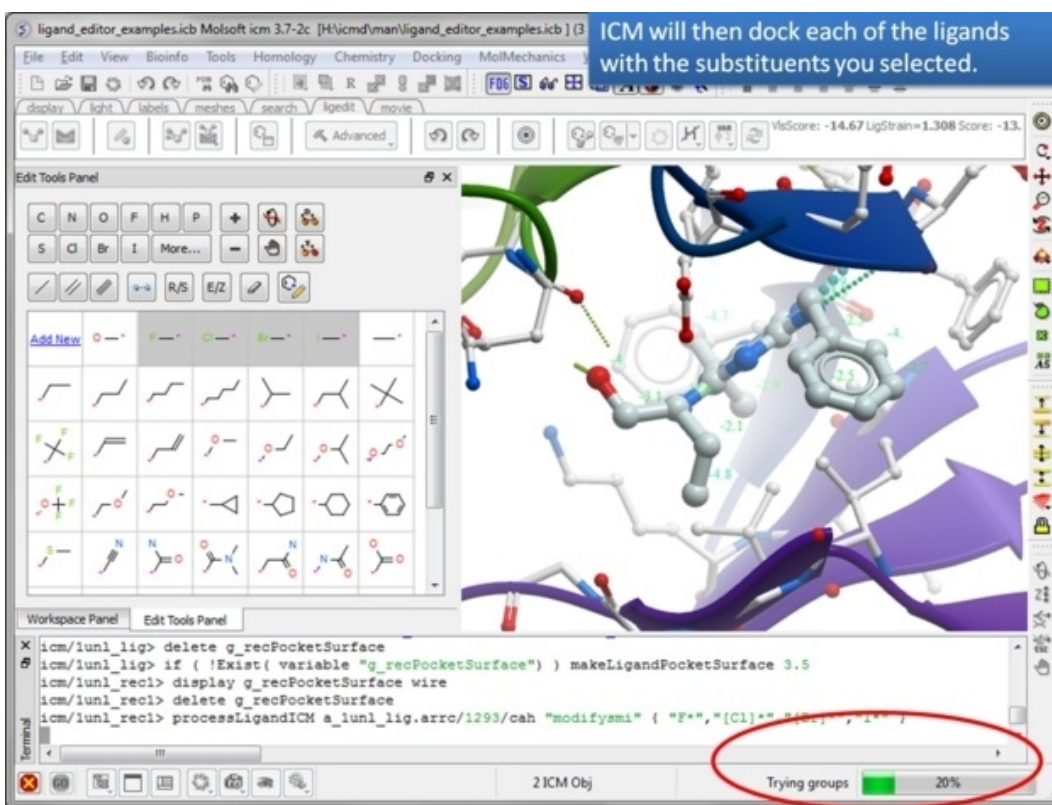
Step 9: The ligand surface representation is a useful way of identifying space in the pocket where you can add groups to the ligand. The surface is colored: White=neutral surface Green=hydrophobic surface Red=hydrogen bond acceptor potential Blue=hydrogen bond donor potential. The coloring of the surface helps you determine which modifications to make, for example if the surface is colored red you ideally would add a hydrogen bond acceptor in that region.



Step 10: There seems to be a hydrophobic region in this pocket that we might try to explore.



Step 11: You can try any modification group you wish from the "edit tools panel". We will try some halogen substituents and see if they improve the score.



Step 12: The editor will add each substituent and evaluate the score. This may take some time depending on the number of groups you chose but you can monitor the progress in the bottom right hand corner.

A table of the docking results from each of the substituents will be displayed.

Docking results

mol	chiral,3D	L group	smiles	Score	VisScore	Strain	MoLoop	MoLoop
1			[Br]*	-14.92	-16.16	1.233	3.928	-6.46

2 ICM Obj

Step 13: A table of ranked hits will be displayed. In this case the bromine substituent scored best.

Double click on the table to load the modified ligand and look at new score.

2. Observe Score has improved

1. Double click here to load the ligand

mol	chiral,3D	L group	smiles	Score	VisScore	Strain	MoLoop	MoLoop
1			[Br]*	-14.92	-16.16	1.233	3.928	-6.46

2 ICM Obj

Step 14: Double click on the table row to display the complex and observe the improved score.

We can now try to RE-DOCK the ligand. First select the atoms that are making the key hydrogen bonds with the hinge region.

1. Click on the select atom button.

2. Click on the atoms near the hinge.

Note: Selected atoms are highlighted with a green cross.

mol	chiral	3D	L	group	smiles	Score	VlsScore	Strain	MoLoop	MoLoop
1				Br	[Br]*	-14.92	-16.16	1.233	3.928	-6.46

Step 15: We can now re-dock this ligand to make sure the interactions are optimal. We must remember to tether the atoms making a hydrogen bond with the hinge and so first we must select these atoms.

1. Advanced/ Restrain Ligand

Tether the selected atoms. These atoms will maintain their position during docking.

Tethered atoms will have a red cross through them

mol	chiral	3D	L	group	smiles	Score	VlsScore	Strain	MoLoop	MoLoop
1				Br	[Br]*	-14.92	-16.16	1.233	3.928	-6.46

Step 16: Tether the atoms.

1. Click here to re-dock

2. Observe Score has improved

Re-dock the ligand and the Scores should improve.

ConflNum	L	Score	VsScore	Strain	RecConf	Steric	Torsion	Electro	Hbond	Hvdrooh	Surface
1	2	-22.77						.64	-4.669	-8.51	17.55
2	3	-17.31						.95	-4.65	-8.446	17.43
3	4	-16.23						.23	-3.829	-8.457	19.05
4	5	-16.23						.44	-3.785	-8.194	18.9
5	1	-14.96	-16.2	1.233	1unl_rec1:1	-30.17	7	21.23	-5.148	-7.983	18.85

Step 17: Click on the re-dock button and a stack of the best conformations will be displayed in a table called DOCKCONF. Observe the score has improved significantly for this modified ligand.

Note: You can undo and redo any ligand modification. You can also save any conformation to a table.

Add current conformation to a table

Undo/Redo

ConflNum	L	Score	VsScore	Strain	RecConf	Steric	Torsion	Electro	Hbond	Hvdrooh	Surface
1	2	-22.77	-26.74	3.974	1unl_rec1:1	-38.9	7	18.64	-4.669	-8.51	17.55
2	3	-22.2	-26.6	4.396	1unl_rec1:1	-38.95	7	18.95	-4.65	-8.446	17.43
3	4	-17.31	-20.35	3.048	1unl_rec1:1	-36.18	7	19.23	-3.829	-8.457	19.05
4	5	-16.23	-19.35	3.114	1unl_rec1:1	-36.14	7	20.44	-3.785	-8.194	18.9
5	1	-14.96	-16.2	1.233	1unl_rec1:1	-30.17	7	21.23	-5.148	-7.983	18.85

Step 18: Never worry about modifying a ligand because everything is saved. Try anything you want because you can always undo and redo. There is also a convenient way of storing a modification you like in a table.

1. Click on the "Edit Tools Button"

Other options to try: There are a variety of editing tools.

Delete atoms

Manually change torsion angles

Edit in 2D editor

Edit bond type

Stereochemistry

ConfNum	L	Score	VisScore	Strain	RecConf	Steric	Torsion	Electro	Hbond	Hvdrooh	Surface
1	2	-22.77	-26.74	3.974	1unl_rec1:1	-38.9	7	18.64	-4.669	-8.51	17.55
2	3	-22.2	-26.6	4.396	1unl_rec1:1	-38.95	7	18.95	-4.65	-8.446	17.43
3	4	-17.31	-20.35	3.048	1unl_rec1:1	-36.18	7	19.23	-3.829	-8.457	19.05
4	5	-16.23	-19.35	3.114	1unl_rec1:1	-36.14	7	20.44	-3.785	-8.194	18.9
5	1	-14.96	-16.2	1.233	1unl_rec1:1	-30.17	7	21.23	-5.148	-7.983	18.85

2 ICM Obj

Step 19: Explore the "Edit Tools Panel".

Other options to try: Screen a large database of substituent groups.

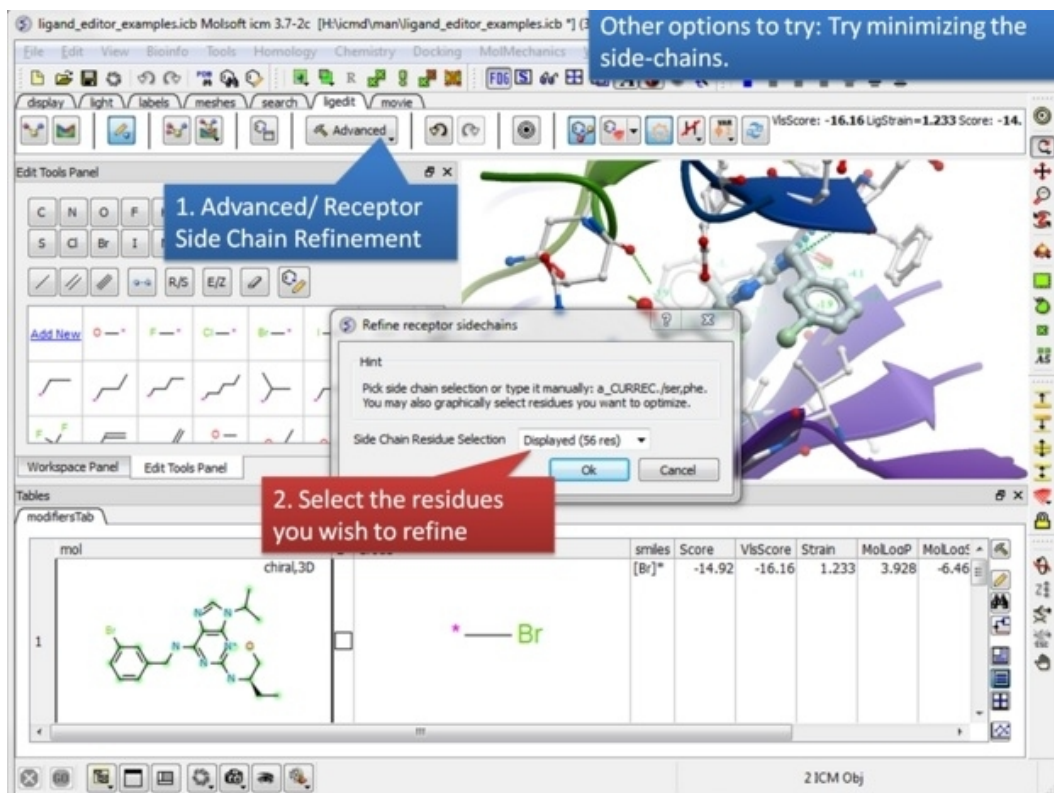
2. Advanced/Find Group.

1. Select one hydrogen atom or terminal atom.

ConfNum	L	Score	VisScore	Strain	RecConf	Steric	Torsion	Electro	Hbond	Hvdrooh	Surface
1	2	-22.77	-26.74	3.974	1unl_rec1:1	-38.9	7	18.64	-4.669	-8.51	17.55
2	3	-22.2	-26.6	4.396	1unl_rec1:1	-38.95	7	18.95	-4.65	-8.446	17.43
3	4	-17.31	-20.35	3.048	1unl_rec1:1	-36.18	7	19.23	-3.829	-8.457	19.05
4	5	-16.23	-19.35	3.114	1unl_rec1:1	-36.14	7	20.44	-3.785	-8.194	18.9
5	1	-14.96	-16.2	1.233	1unl_rec1:1	-30.17	7	21.23	-5.148	-7.983	18.85

2 ICM Obj

Step 20: Perform a virtual screen of a substituent database.



Step 21: You can also minimize side-chains inside the ligand editor. This may remove unwanted clashes between your receptor and ligand.

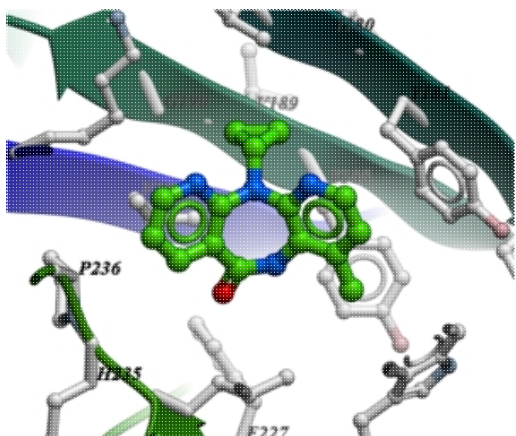
19.8 Atomic Property Field Tutorials

APF Background

Atomic Property Fields (APF) is a 3D pharmacophoric potential implemented on a grid. APF can be generated from one or multiple ligands and seven properties are assigned from empiric physico-chemical components (hydrogen bond donors, acceptors, Sp² hybridization, lipophilicity, size, electropositive/negative and charge). Here we describe template APF superposition whereby the APF is generated from a single or multiple template and is then globally optimized with the internal force-field energy of the ligand. You can read more about the APF method [here](#).

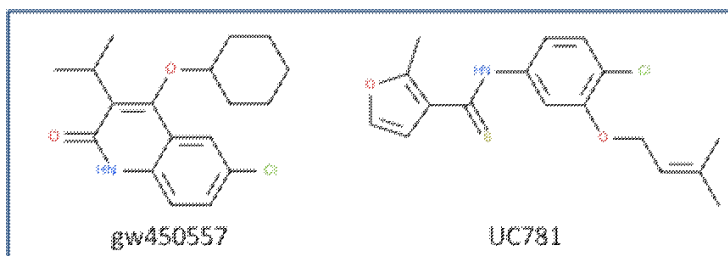
About this Example In this example we will use Atomic Property Fields to superimpose, screen and cluster a set of Non-nucleoside reverse-transcriptase inhibitors (NNRTIs) (antiretroviral drugs) used in the treatment of human immunodeficiency virus (HIV). We will use a non-trivial superposition example of two NNRTIs which have similar pharmacophoric properties but are significantly different substructure. We will then read in additional PDBs which contain first generation NNRTIs approved by the FDA (Nevirapine, Delavirdine, Efavirenz) we will then use the combined fields to screen an sdf library to the APF fields

NNRTIs are chemically diverse but bind in the same allosteric pocket in the palm domain of the p66 subunit and part of the p51 subunit. Many NNRTIs bind in a "butterfly" conformation (see image below of Nevirapine) making pi-pi interaction with the rings in the pocket.



19.8.1 APF Superposition

This is an example of a non-trivial chemical superposition, we will superimpose inhibitor UC781 to gw450557. A prepared ICM file containing the two inhibitors can be downloaded here <ftp://ftp.molsoft.com/pub/workshop/> (1. cut and paste the address into a web browser. 2. Right click on apf_superposition_example.icb and choose "Save Link As". 3. Open the .icb in ICM File/Open).



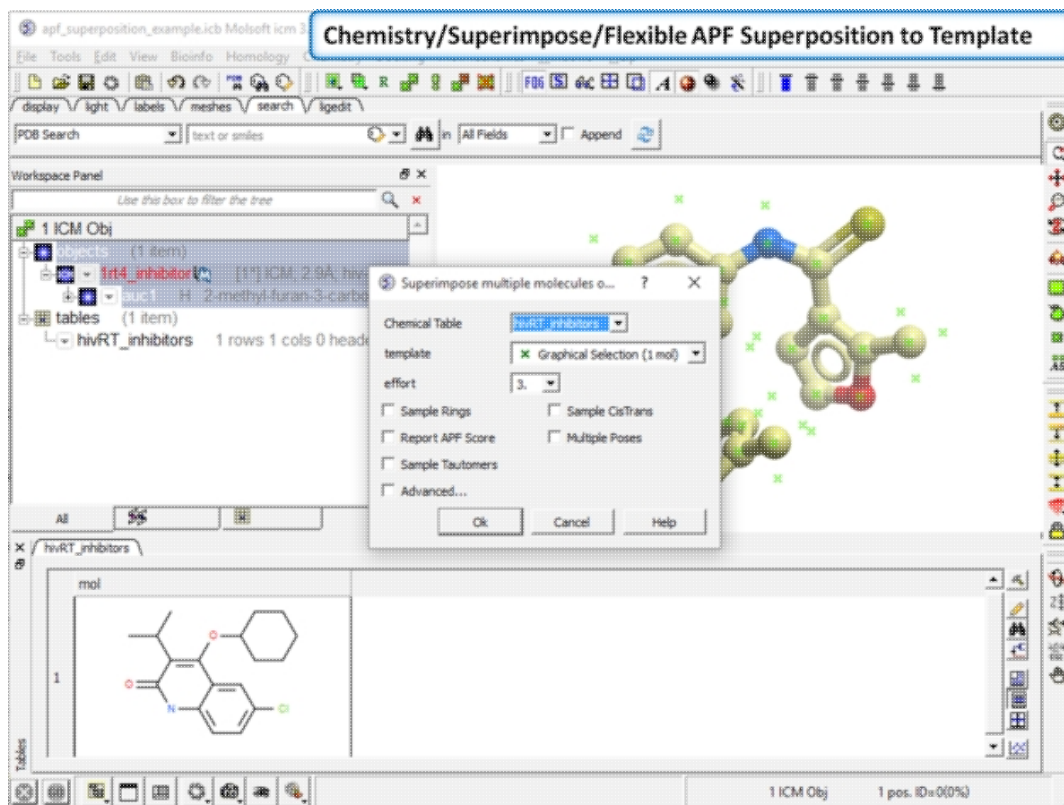
Read in the apf_superposition_example.icb file.

Template

For superposition

1 ICM Obj 1 pos. ID=0(0%)

Step 1: Download and open the prepared ICM file. The file can be downloaded here <ftp://ftp.molsoft.com/pub/workshop/> (1. cut and paste the address into a web browser. 2. Right click on `apf_superposition_example.icb` and choose "Save Link As". 3. Open the `.icb` in ICM File/Open).



Step 2: Superimpose the chemical from the table onto the 3D structure using Chemistry/Superimpose/Flexible APF Superposition to Template. Enter the parameters as shown in the screenshot.

Click on table to view superposition result.

2. Superimposed molecule is displayed in the workspace. Right click and rename it

1. Click here to display superimposed chemical

mol	L	Molecu Name	MOL NUM
1		1t13	1

Step 3: When the superposition has finished a results table will be displayed. You can toggle the superimposed molecule display on or off in the table.

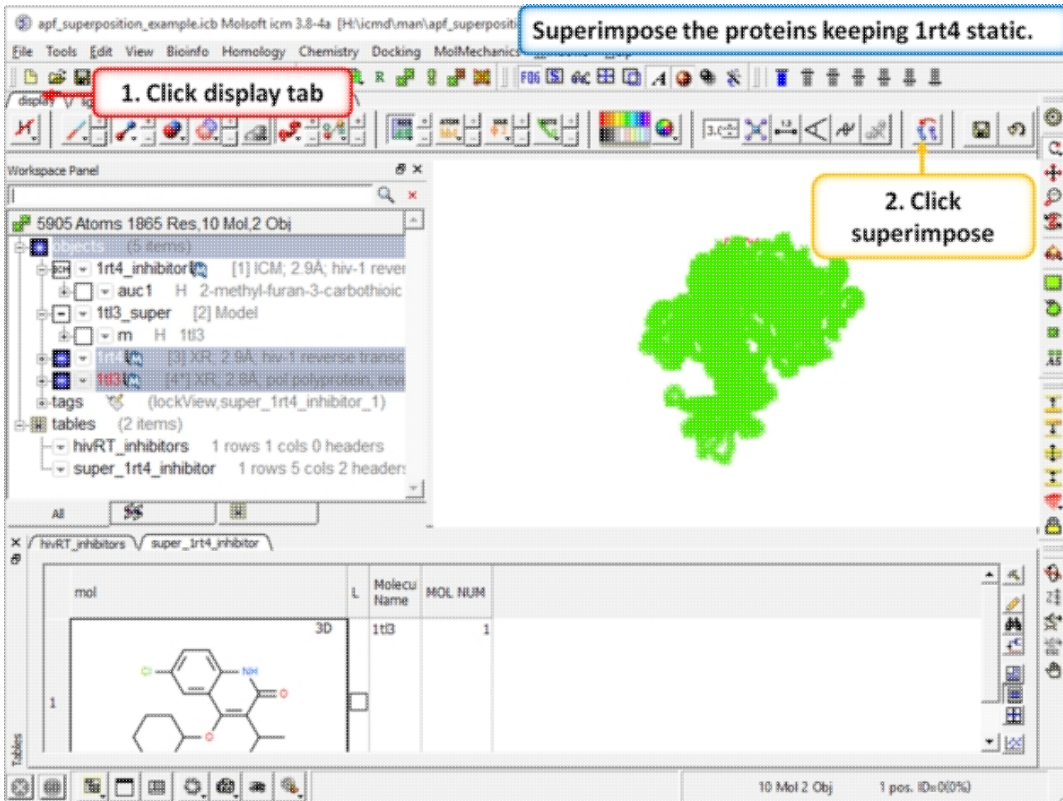
Read in PDB structures 1rt4 and 1t13

1. Read in 1rt4 and 1t13

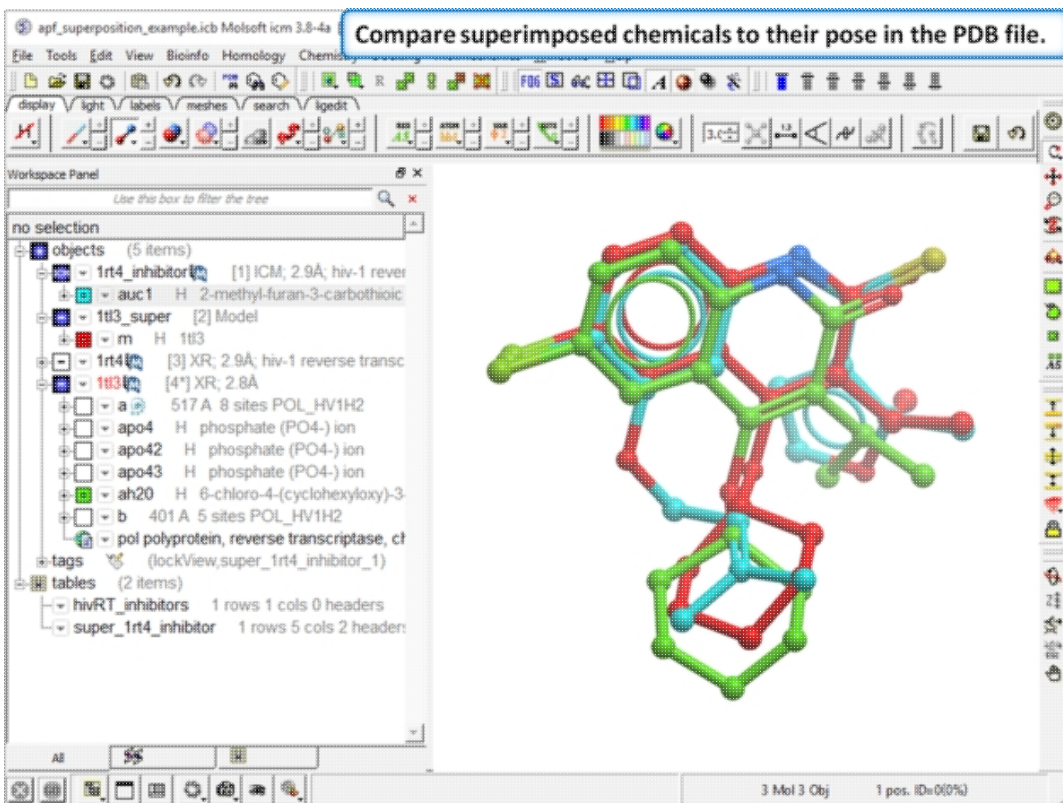
2. Select both proteins

mol	L	Molecu Name	MOL NUM
1		1t13	1

Step 4: Now lets compare the docked pose of UC781 with the crystal structure pose. First we need to read in the PDB structures contaning UC781 (1t13) and gw450557 (1rt4).



Step 5: Superimpose the crystal structures keeping 1rt4 static. You need to select the two pdbs and use the superimpose button in the display tab.



Step 6: Compare the APF superimposed structure with the crystal structure. The RMSD should be

