From Internal Coordinate Mechanics to ICM

Modeling

Docking

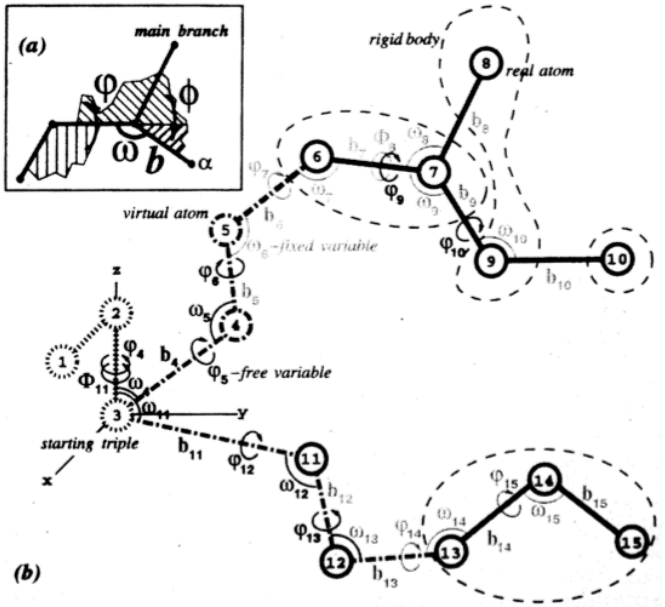
Cheminformatics

Graphical and Scripting Environment

Big Structural Data

Internal Coordinate Mechanics

- ICM (ϕ,α,b)
  - Full set of Internal Variables: ICM (ϕ,α,b)
  - **ICM stochastic optimizer** (1992-99)
    - Any subset of molecules, svariables
    - Collective variables
    - Square root (biased) sampling
    - Restraints and grids to modify the energy function



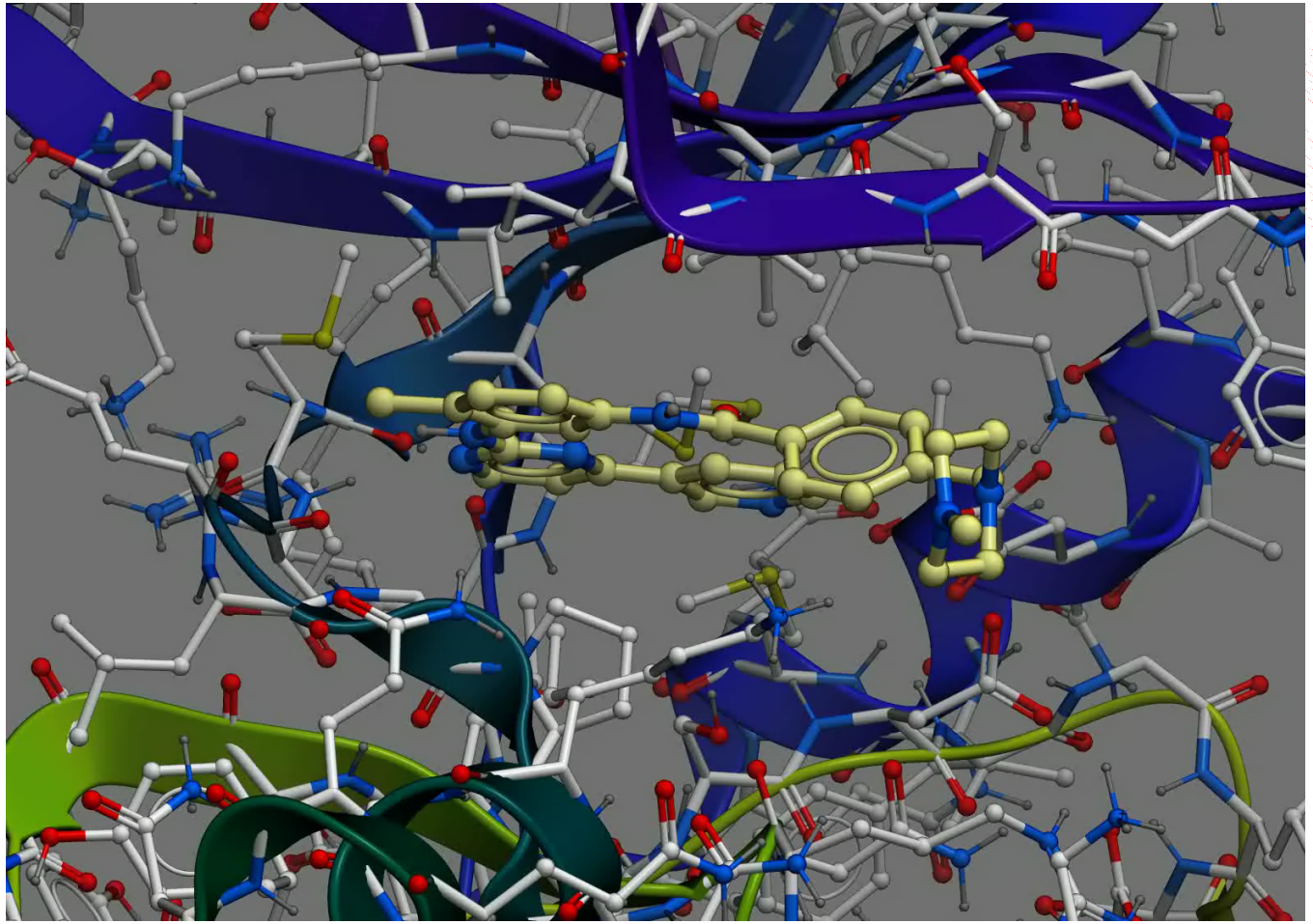
Mazur, Abagyan (1989) "**General Equations for multiple branched polymers of motion in internal coordinates**", *JBSD*..

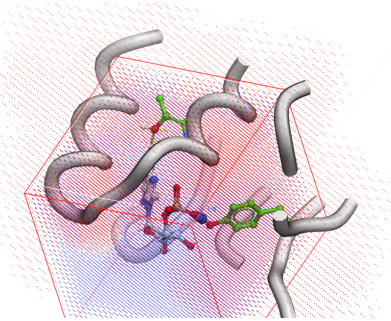
Abagyan, Totrov, Kuznetsov (1994) "**ICM - a new method for protein modeling and docking**" *J. Comp. Chem*. 15, 488-50

Abagyan, and Totrov, (1994)."**Biased Probability Monte Carlo searches and Electrostatics**" *J. Mol. Biol.*

Totrov M, Abagyan R (1994) **Detailed ab initio prediction of lysozyme-antibody complex with 1.6 A accuracy**. *Nat Struct Biol*, 1994

ICM Ligand Docking



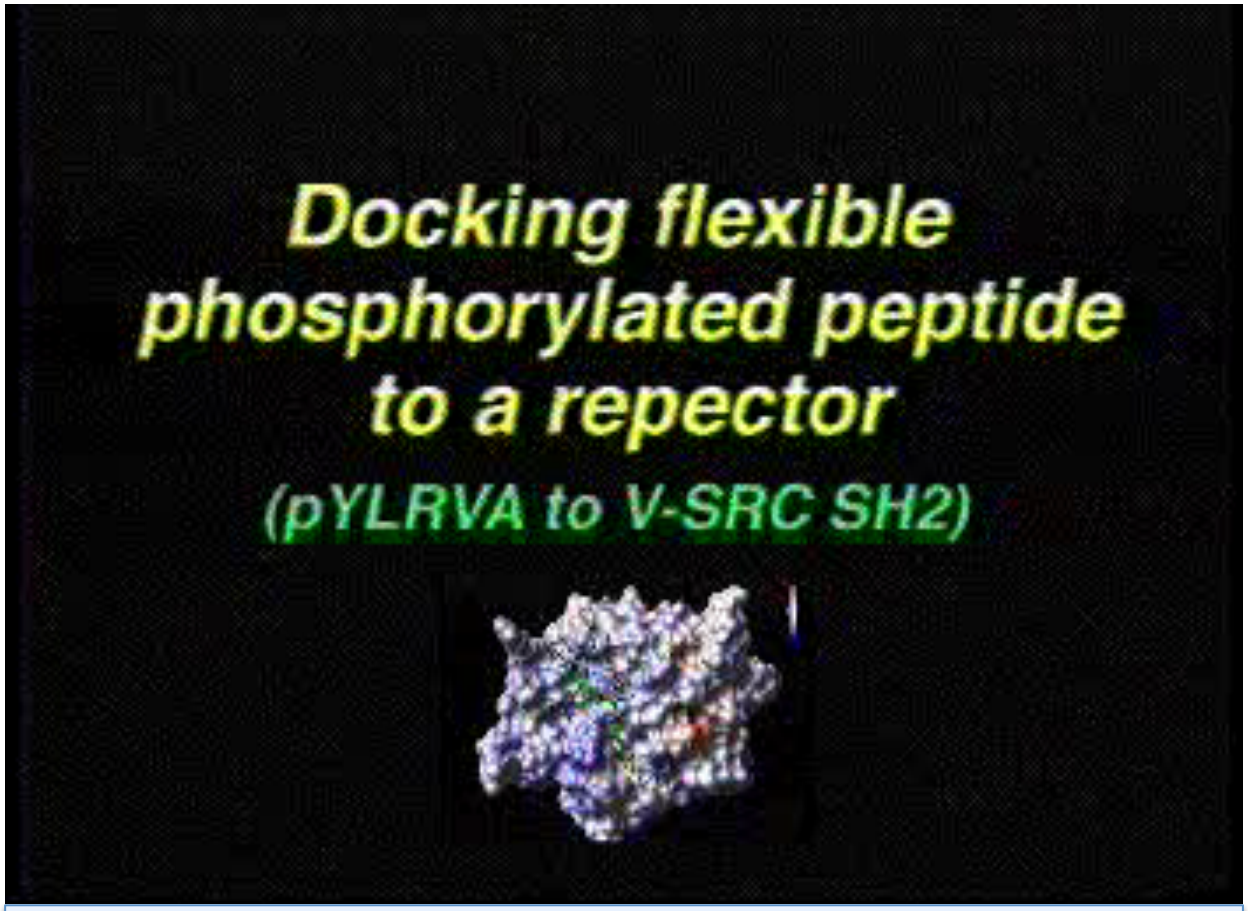


Grids

with second derivatives

Totrov, Abagyan (1997) **Flexible Ligand Docking.** *Proteins*

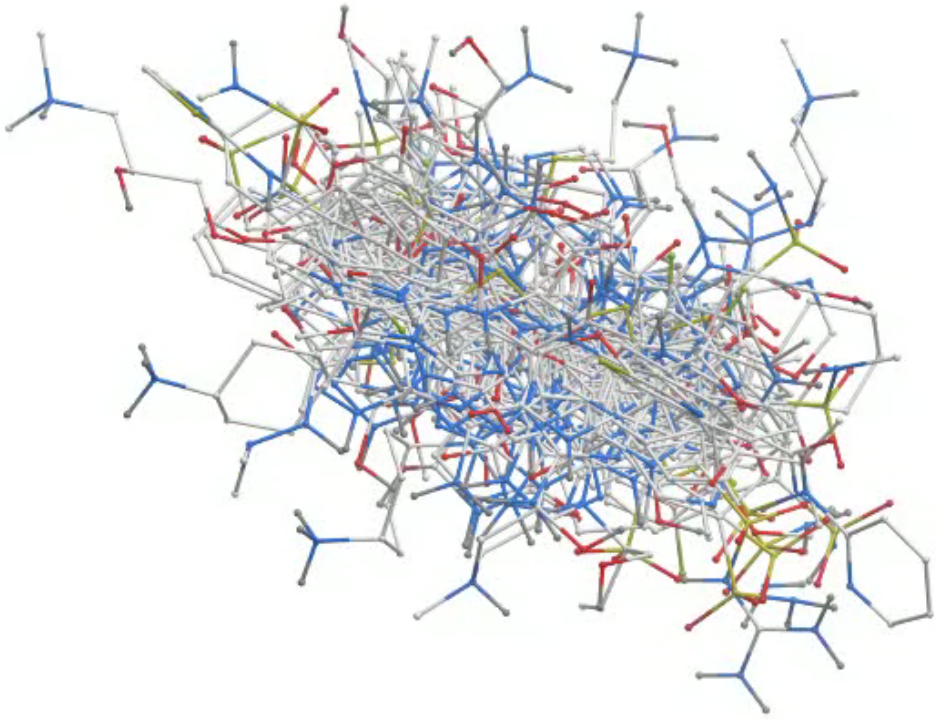
Tools: Protein and Peptide Docking

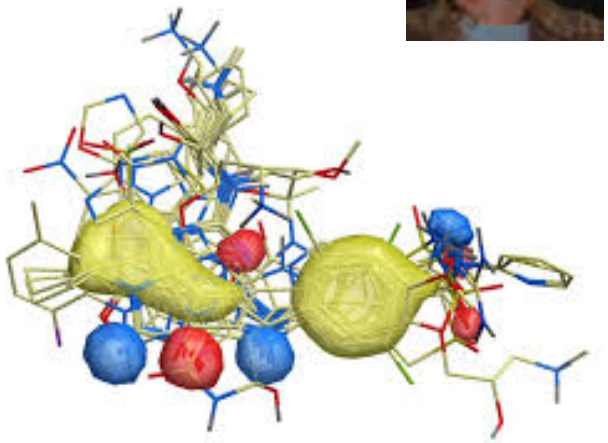


New:
Protein-Protein
FFT Docking

Zhou, Abagyan, *Folding & Des* 1998. **How and why pTyr-containing peptides bind to the SH2 and PTB domains.**
Bordner AJ, Abagyan R. **Ab initio prediction of peptide-MHC binding geometry for diverse class I MHC allotypes**. *Proteins*. 2006 May 15;63(3):512-26

APF : Atom Property Fields





Tools: **_chemAlign** and **_chemSuper** (screen)

Totrov M. **Atomic property fields: generalized 3D pharmacophoric potential for automated ligand superposition, pharmacophore elucidation and 3D QSAR** *Chem Biol Drug Des.* 2008

Independent APF Validation

Giganti *et al* **Comparative Evaluation of 3D Virtual Ligand Screening Methods: Impact of the Molecular Alignment on Enrichment** JCIM (2010), *Pasteur Inst., Paris*

- ICM, Surflex-dock/Surflex-sim, FlexX/FlexS, OMEGA-FRED/OMEGA-ROCS

Chemical Super-position

Table 3. Molecular Alignment of the DUD-own Active Compounds[a]

		ADA (39)	CDK2 (72)	DHFR (410)	ER (39)	FXA (146)	HIVRT (43)	NA (49)	P38 (454)	THR (72)	TK (22)	TRP (49)	mean (1100)
Surflex-sim	2	12.82	12.5	44.39	56.41	4.11	18.6	18.37	9.69	4.17	68.18	40.82	23.15
	1	35.9	51.39	53.66	43.59	33.56	72.09	75.51	69.6	93.06	31.82	59.18	59.07
	0	51.28	36.11	1.95	0	62.33	9.3	6.12	20.7	2.78	0	0	17.78
ROCS	2	12.82	43.06	74.15	41.03	14.38	30.23	79.59	9.47	2.78	86.36	8.16	35.63
	1	20.51	36.11	14.39	56.41	28.77	34.88	14.29	41.19	69.44	9.09	81.63	32.83
	0	66.67	20.83	11.46	2.56	56.85	34.88	6.12	49.34	27.78	4.55	10.2	31.54
FlexS	2	15.38	25	56.1	48.72	35.62	16.28	36.73	14.98	30.56	81.82	18.37	33.48
	1	20.51	19.44	11.71	43.59	13.7	46.51	57.14	74.01	5.56	13.64	2.04	35.77
	0	64.1	55.56	32.2	7.69	50.68	37.21	6.12	11.01	63.89	4.55	79.59	30.75
ICMsim	2	46.15	12.5	86.83	51.28	70.55	18.6	75.51	20.04	88.89	90.91	69.39	54.48
	1	23.08	68.06	11.95	46.15	16.44	46.51	14.29	68.28	9.72	9.09	28.57	36.49
	0	30.77	19.44	1.22	2.56	13.01	34.88	10.2	11.67	1.39	0	2.04	9.03

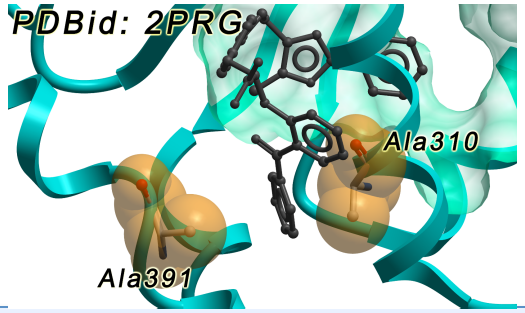
Docking/ Screening Enrichment

Table 4. Early (1%) and Late (10%) Enrichments for the DUD-own Active Compounds[a]

enrichment	Surflex-sim		ROCS		FlexS		ICMsim		DOCK	
all actives	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%
ADA	5.13	23.08	7.69	33.33	15.38	30.77	10.26	28.21	15	40
CDK2	2.78	8.33	22.22	38.89	5.56	11.11	9.72	30.56	10	25
DHFR	4.88	25.61	19.02	61.95	8.05	27.8	20.24	80.73	25	60
ER	10.26	82.05	10.26	89.74	15.38	71.79	17.95	76.92	10	20
FXA	1.37	2.74	4.11	4.11	5.48	19.86	11.64	46.58	5	35
HIVRT	9.3	18.6	20.93	25.58	27.91	58.14	18.6	39.53	5	25
NA	16.33	55.1	34.69	89.8	20.41	69.39	16.33	73.47	10	60
P38	9.03	21.15	8.81	15.42	12.56	26.21	11.23	17.84	2	25
thrombin	1.39	15.28	0.75	15.28	2.78	12.5	2.78	76.39	5	35
TK	22.73	50	22.73	50	9.09	54.55	22.73	63.64	0	25
trypsin	0	59.18	4.08	34.69	18.37	18.37	10.2	95.92	0	30
mean	7.56	32.83	14.12	41.71	12.82	36.41	13.79	57.25	7.91	34.55
st. dev.	6.99	24.91	10.48	28.79	7.57	22.72	5.85	25.77	7.35	13.87
median	5.13	23.08	10.26	34.69	12.56	27.8	11.64	63.64	5	30

50 Shades of Pocket Flexibility

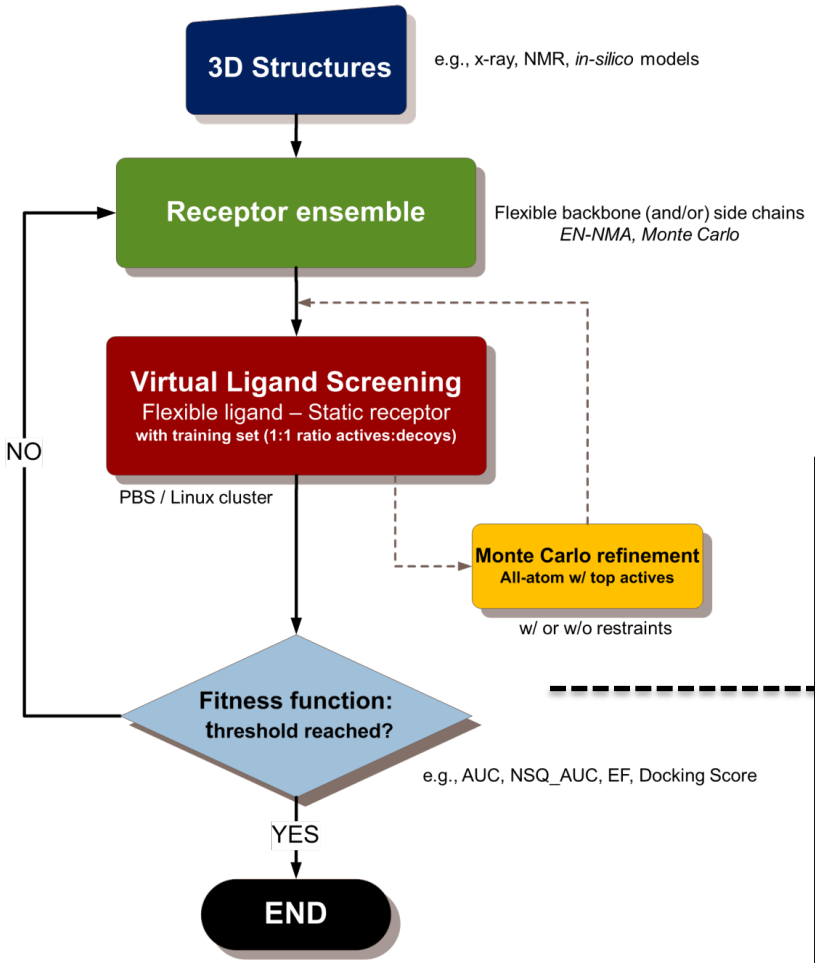
- Explicit Side Chains in LigEdit and _dockScan
- 4D Docking for fast Ensemble Docking
- SCARE (SCan Alanines & Refine)
- NMA or ICM-Generated backbone variations
- Full ICM simulations with custom defined space and restraints



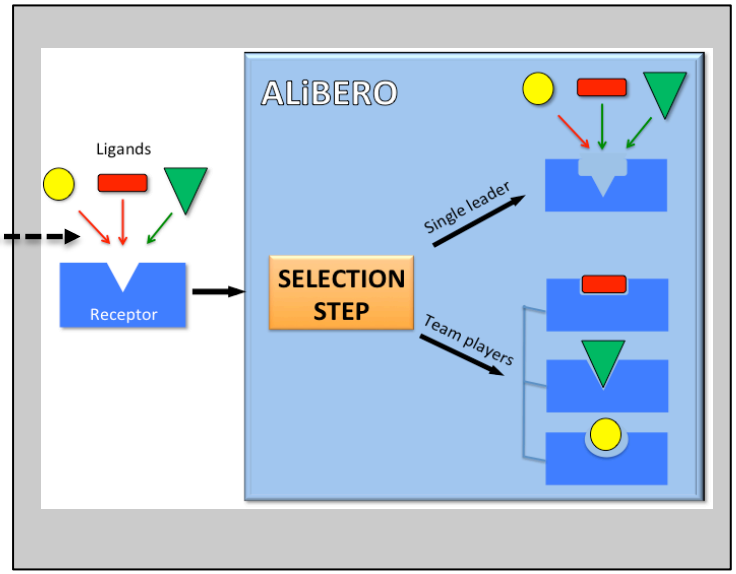
Bottegoni et al. 2008 "**SCARE..**" *JCAMD*

Bottegoni et al. 2009 **"4D docking: a fast and accurate account of discrete receptor flexibility ..",** *JMC*

Ligand Guided Model Building



- ALibero
- Extension of _dockScan



Bisson, Cheltsov et al. 2006, PNAS
Katritch et al. 2008, 2011, ++ **GPCR agonist binding revealed by modeling**..,
Rueda et al. **ALiBERO: Evolving a team of complementary pocket conformations rather than a single leader** (2012) *J Chem Inf Mod*

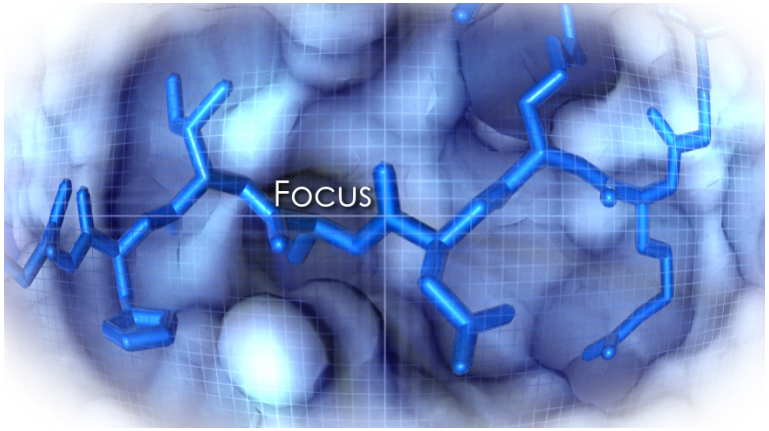
JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

FOCUS — Development of a Global Communication and Modeling Platform for Applied and Computational Medicinal Chemists

Nikolaus Stiefl,*[,†] Peter Gedeck,[‡] Donovan Chin,[§] Peter Hunt,[∥] Mika Lindvall,[⊥] Katrin Spiegel,[∥] Clayton Springer,[§] Scott Biller,[§] Christoph Buenemann,[#] Takanori Kanazawa,[▽] Mitsunori Kato,[§,▽] Richard Lewis,[†] Eric Martin,[⊥] Valery Polyakov,[⊥] Ruben Tommasi,[§] John van Drie,[§] Brian Vash,[§] Lewis Whitehead,[§] Yongjin Xu,[⊥] Ruben Abagyan,[○] Eugene Raush,[○] and Max Totrov[○]

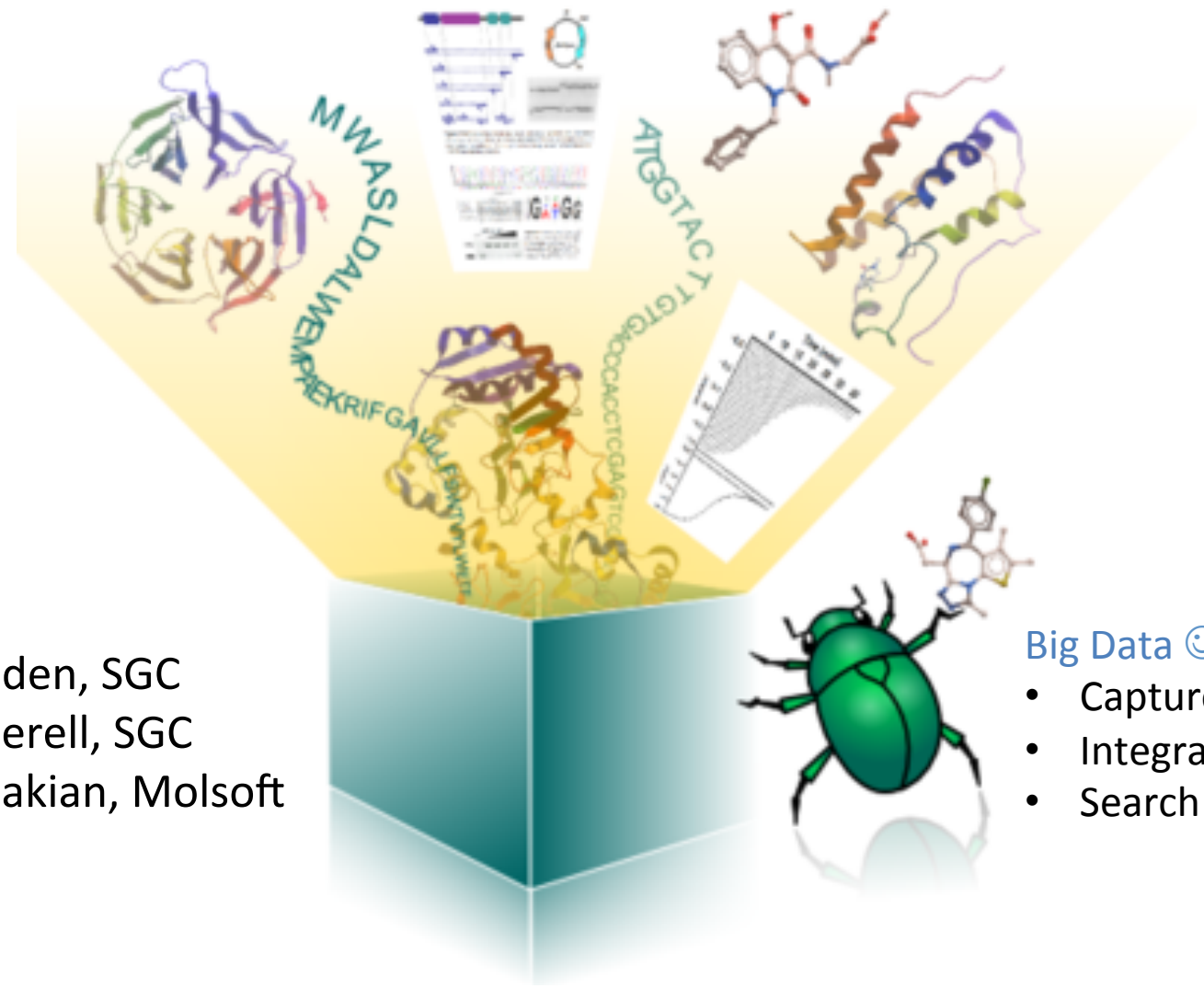


NOVARTIS

specifically Kenji Namoto, David Wayne Hughes, Paul Butler, Gregory Landrum, Sandra Mueller, Joerg Muehlbacher, Finton Sirockin, Simona Cotesta, Marc Litherland, Rajeshri Karki, Yipin Lu, Robert Pulz, Catherine Leblanc, Sabina Pecchi, Gabriel Gamber, Jay Knowles, Michael Dechantstreiter, and Thomas Dice. We also thank Jose Duca for supporting the project during the last three years. On the Molsoft side, we thank Andrew Orry, Crystal Crawford, Levon Budagyan, Polo Lam, Irina Kufareva, and Elena Arnauto y

Stiefl et al. JCIM, 2015

Scarab



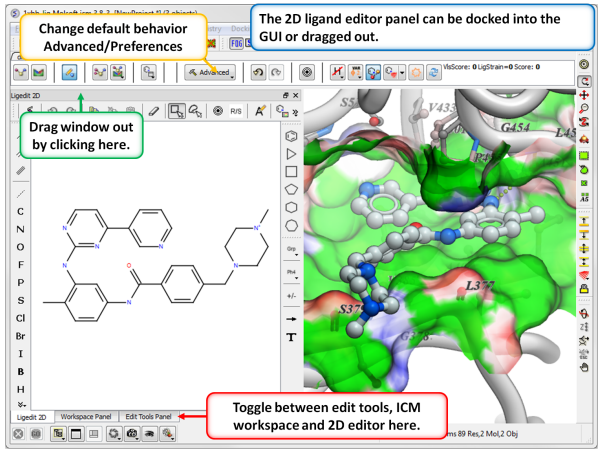
Brian Marsden, SGC
David Damerell, SGC
Arman Sahakian, Molsoft

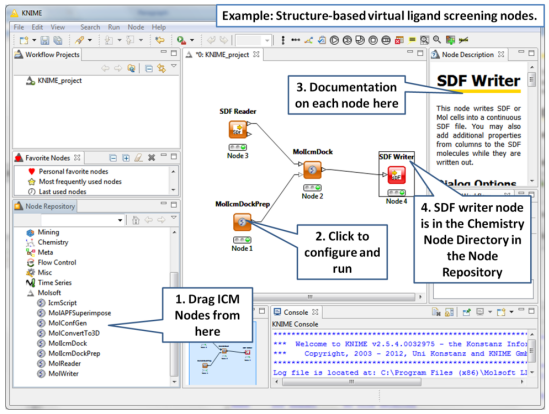
Big Data ☺ :

- Capture and Create
- Integrate
- Search and retrieve

Three in One

- Workstation GUI Client (eg ICM-pro, Focus, Scarab)
- ICM based backend units (models, screens, homology)
- Mobile or Web dissemination





Microsoft Windows

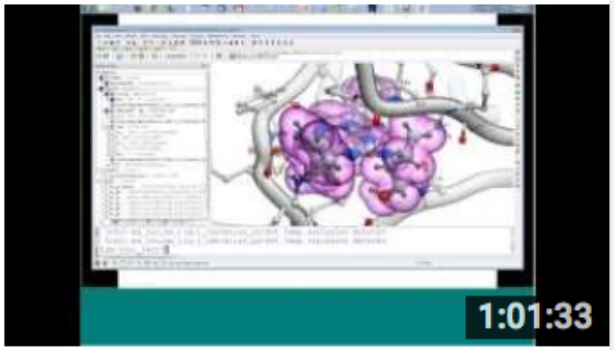
Linux

iPhone

iPad

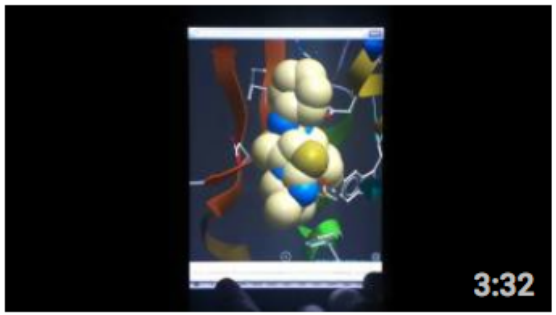
amazon

Training Videos



MolSoft Webinar: Ligand Design using ICM 3D Interactive Ligand…

1,695 views • 2 years ago



iMolview

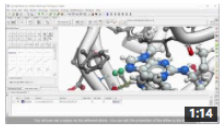
3,743 views • 4 years ago

YouTube

MolSoft Molecules in Silico

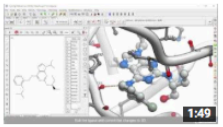
Home | Videos | Playlists | Channels | Discussion | About

Uploads



ICM 3D Ligand Editor: Distance Restraints and Tethers in Docking

17 views • 3 weeks ago

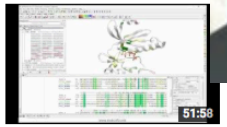


Ligand Editor - Ligand editing in 2D and commit changes to 3D.

6 views • 3 weeks ago

Explicit / Hybrid Docking: Fully flexible side-chains in the ICM…

17 views • 3 weeks ago

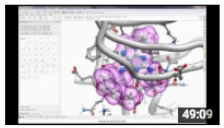


Webinar: Linking Protein Structure, Sequences and…

161 views • 1 month ago

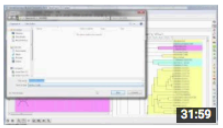
Webinar: Molecular Graphics, Movies, and Fully Interactive 3…

293 views • 2 months ago



Webinar: Ligand Design using the ICM 3D Fully Interactive Ligand…

223 views • 2 months ago

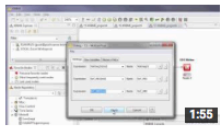


ICM Chemist Webinar - July 2015

176 views • 3 months ago

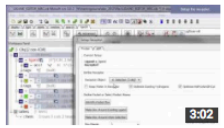
Install MolSoft ICM Nodes for KNIME

88 views • 11 months ago



MolSoft MolCalcProp in KNIME

59 views • 11 months ago

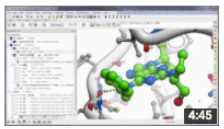


4D Docking in the Ligand Editor.

169 views • 1 year ago

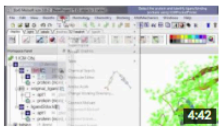
ICM Molecular Editor Tutorial

166 views • 1 year ago



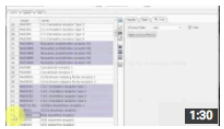
MolSoft - Fully Interactive 3D Ligand Editor v3.8

254 views • 1 year ago



How to re-dock a ligand using MolSoft's ICM-Pro desktop…

457 views • 1 year ago

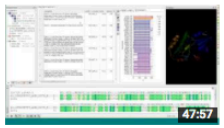


Introducing MolScreen from MolSoft.

49 views • 1 year ago

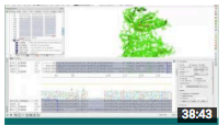
Covalent Docking Tutorial in ICM-Pro

486 views • 1 year ago



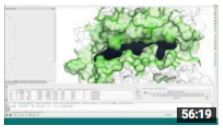
Protein Structure Modeling and Analysis Webinar

1,872 views • 1 year ago



Linking Protein Sequence to 3D Structure using the ICM…

778 views • 1 year ago

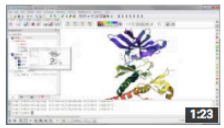


Introducing ICM version 3.8

563 views • 1 year ago

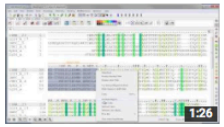
Molecular Graphics, Movies, and Fully Interactive 3D Documents…

964 views • 1 year ago



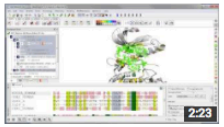
How to make a movie from a series of slides.

69 views • 1 year ago



Alignment Annotation in ICM

76 views • 1 year ago

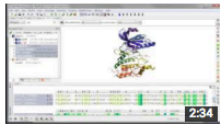


Highlighting Sequence Identity in the Ligand Binding Pocket usin…

140 views • 1 year ago

Editing an Alignment in ICM

80 views • 1 year ago

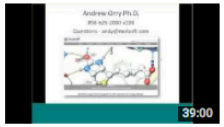


Making Sequence Alignments in ICM - Linking Sequence and…

301 views • 1 year ago

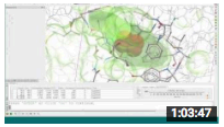
How to dock to two molecules.

174 views • 1 year ago



MolSoft Webinar: Ligand Based Lead Discovery using Atomic…

488 views • 2 years ago

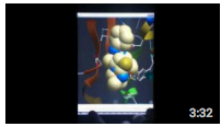


Webinar: Structure Based Lead Discovery using ICM Virtual…

1,213 views • 2 years ago

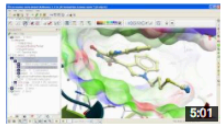
MolSoft Webinar: Ligand Design using ICM 3D Interactive Ligan…

1,695 views • 2 years ago



iMolview

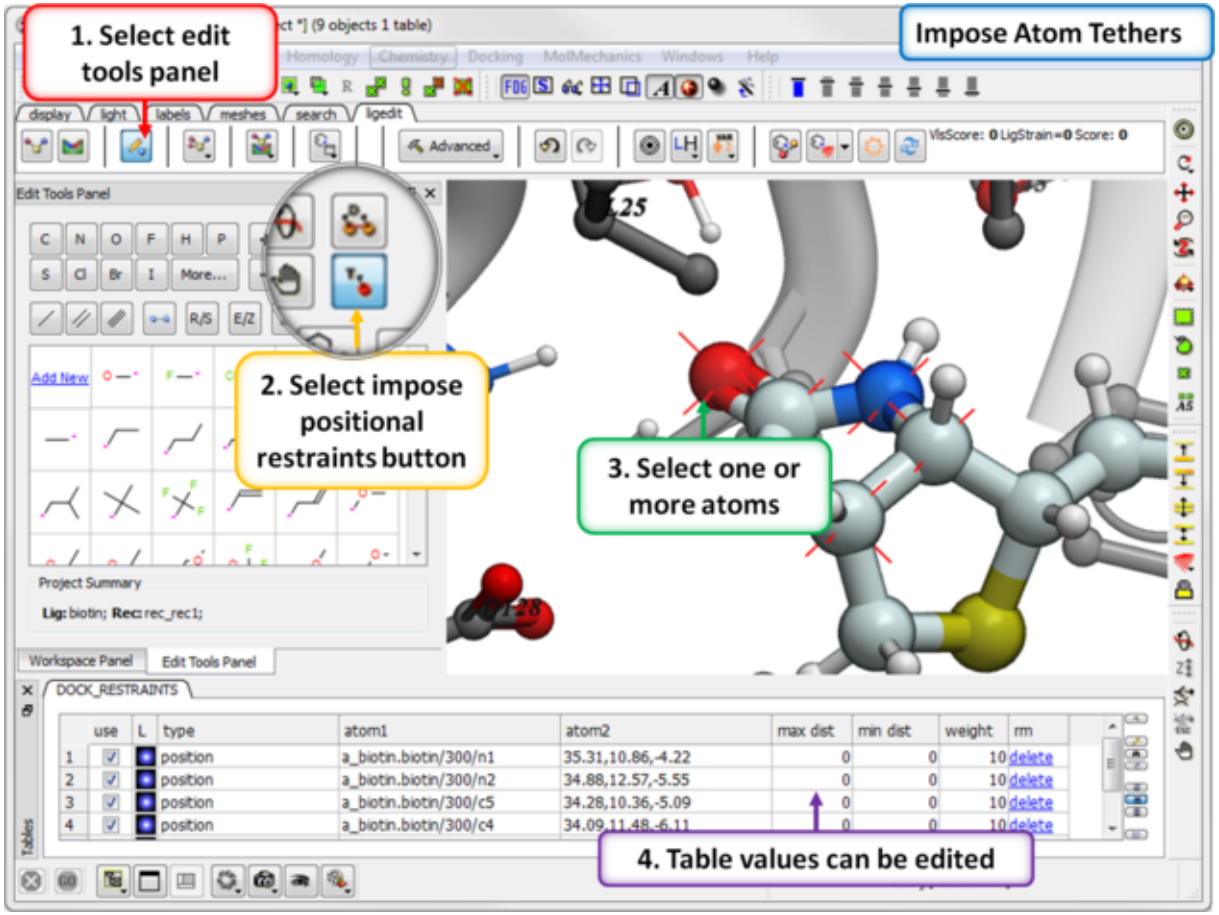
3,743 views • 4 years ago

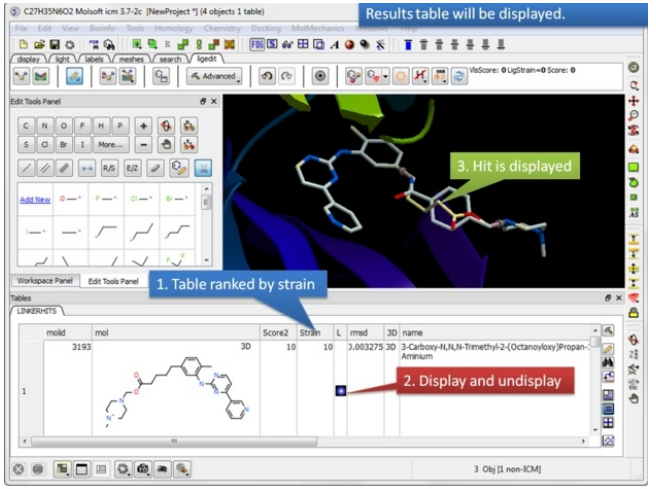


ICM-Browser & ActiveICM

698 views • 4 years ago

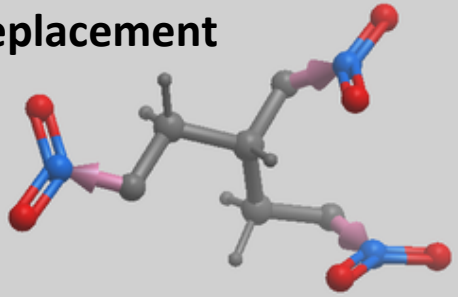
Ligand Editor





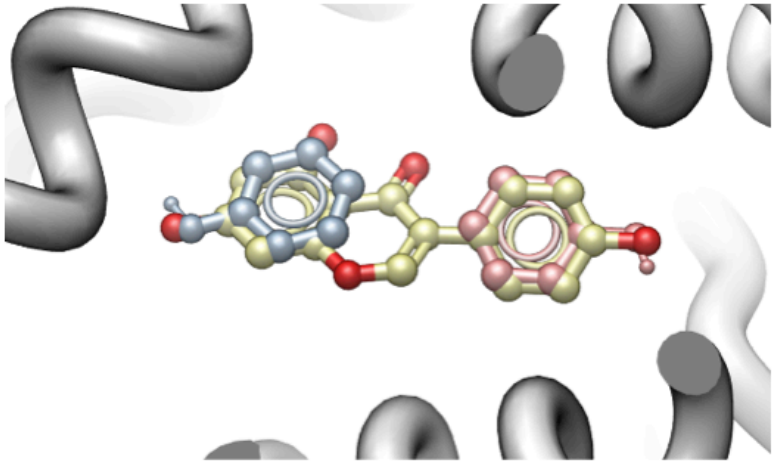
Fragment Linking

Three Point Core Replacement

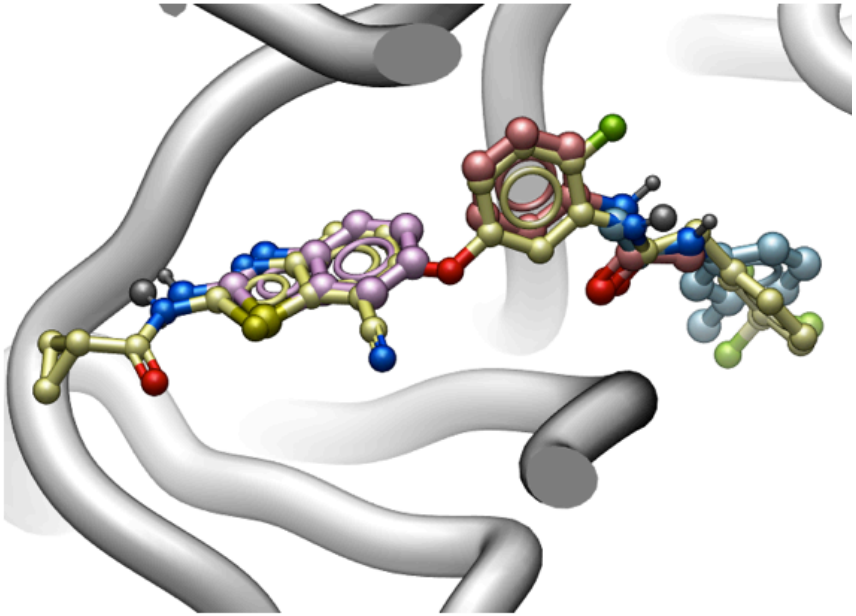


Computational Fragment Screen:

Examples of high-confidence fragment poses versus real ligands



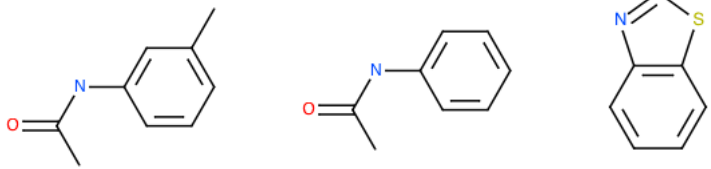
Bluish white -18.8 Red -22.7
p = 0.055 p = 0.011



Bluish white Purple Red
p = 0.2 p = 0.047 p = <0.005

Phenol fragments in PDB 1QKM (ER)

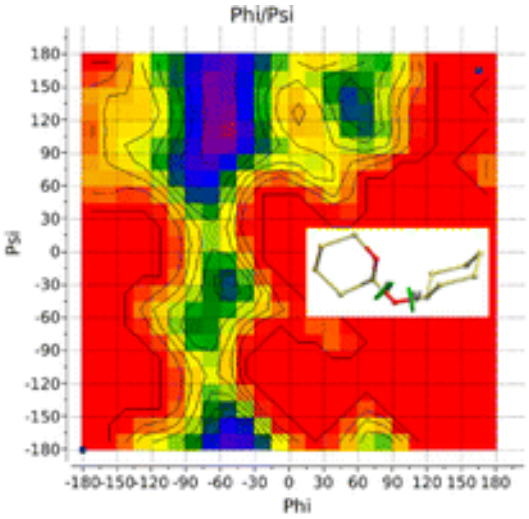
PDB 4KSP (b-RAF kinase)

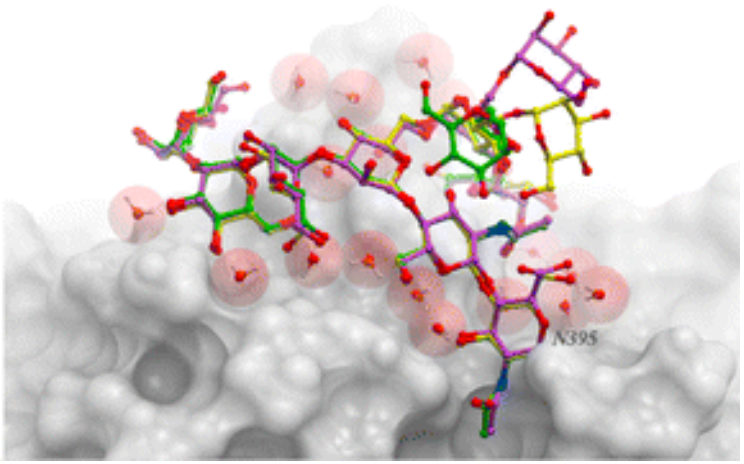


New ICM Force Field: ICMFF

- First ideas : ICFF *Seva Katrich et al. 2003*
- Current QM-based ICMFF:
  - Softer flexibility model, better VW, ε=2, accurate torsion profiles, ..

ICM Scripts:
- _loopmodel
- _dockScan
- _mutant*
- _protDesign





Arnautova, Abagyan, Totrov. RNA; 2015 (glycoproteins) and 2011 (loops, peptides). **All-Atom Internal Coordinate Mechanics(ICM) Force Field for Hexopyranoses and Glycoproteins**. *J Chem Theory Comput*. 2015

First morning after Sweden changed from driving on the left side to driving on the right, 1967 (themetapicture.com)

- Switching to the new force field needs to be done very carefully

Change is difficult

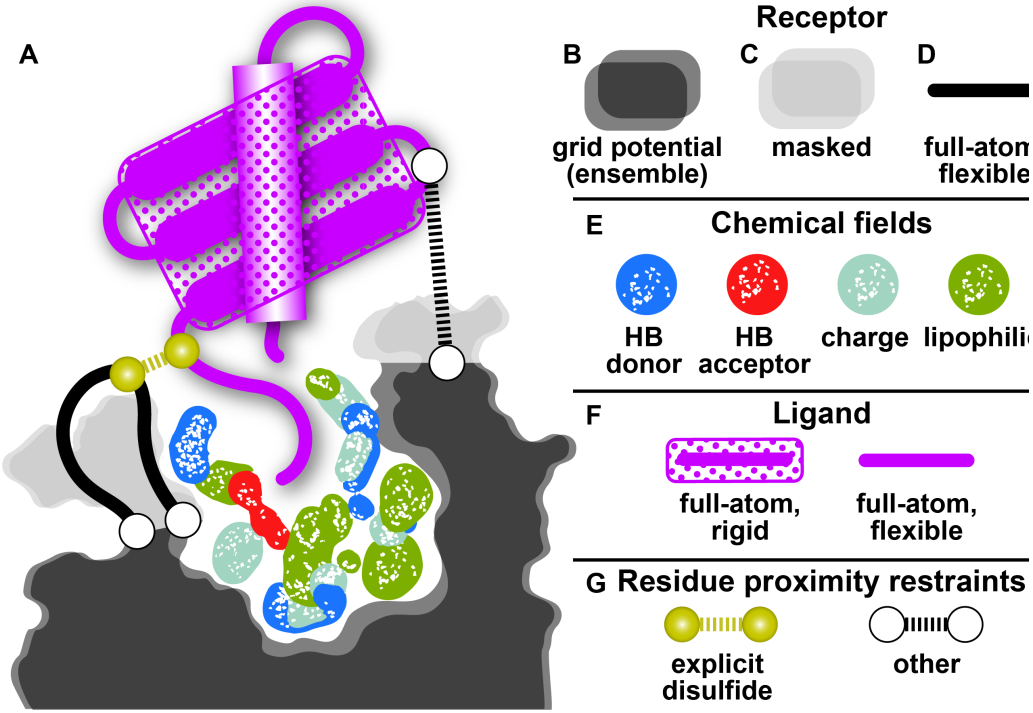
Expanding from Xtal to Models

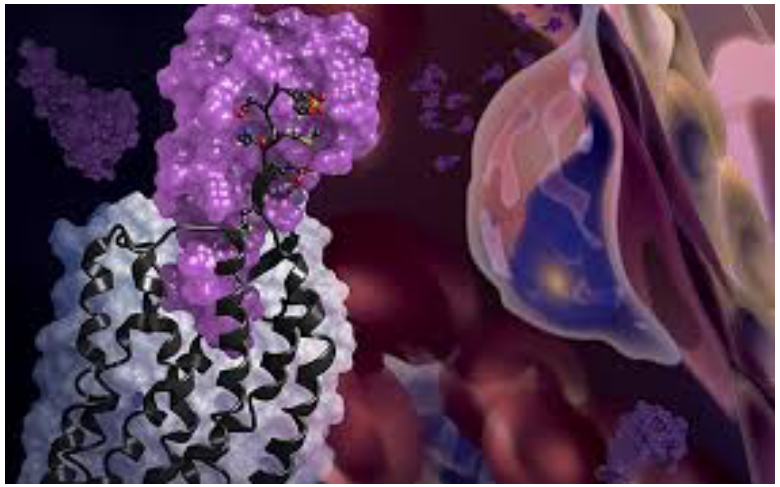
- 25% protein coverage
- 10% amino acid coverage
- 3% GPCR coverage

Modeling
- Ab initio
- Homology
- Loops and Tails

Docking
- Small molecules
- Peptides
- Proteins

Custom modeling protocols



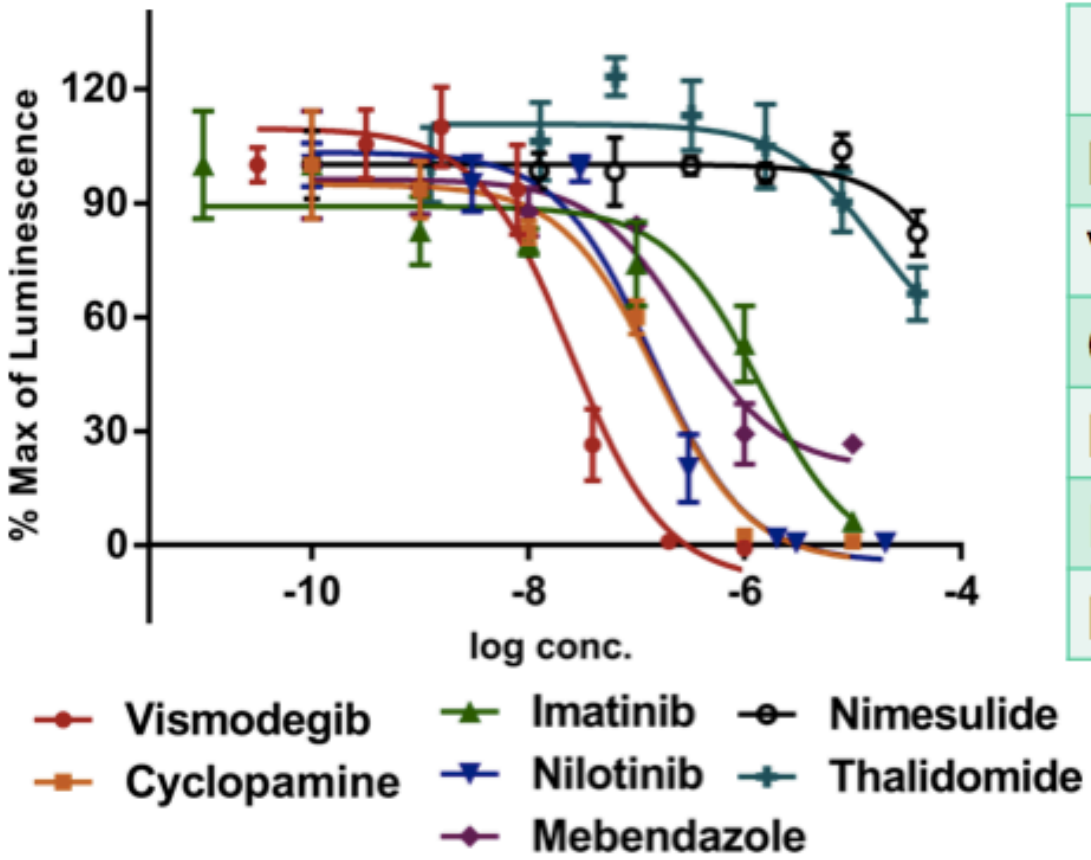


Application to chemokine receptors

Kufareva et al. (2015) ***Experiment-guided molecular modeling of protein-protein complexes involving GPCRs***, *Meth Mol Biol, Wiley, 1335, 295-311, chapter*

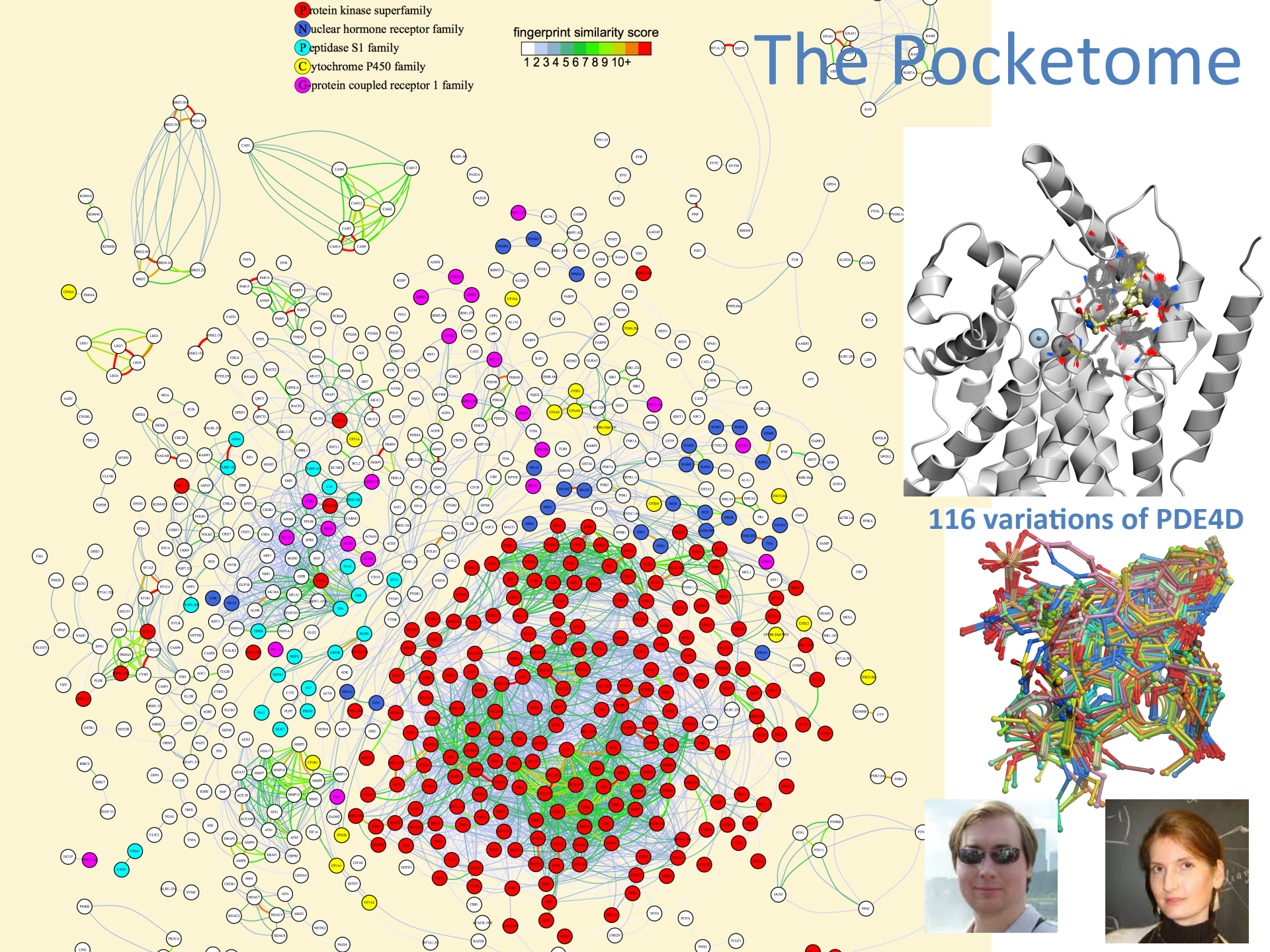
Qin, …,Handel (2015) **Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine**, *Science*.

Three drugs identified as SMO modulators



Effect of Drugs on Hedgehog (Hh) Pathway Activity (Gli- assay)

Table 1	
Drug	**IC50 (nM)**
Vismodegib	25
Cyclopamine	153
Nilotinib	140
Imatinib	1402
Mebendazole	287


Protein kinase superfamily
Nuclear hormone receptor family
Peptidase S1 family
Cytochrome P450 family
G-protein coupled receptor 1 family
fingerprint similarity score
1 2 3 4 5 6 7 8 9 10+
The Pocketome
116 variations of PDE4D

MolScreen: The Five Model Types 2000 Models

- Docking Models
  - **.dfa** docking and multicluster pKd/pIC50 prediction
  - **.dfz** APF docking and Z-scoring
  - **.dpc** 4D Pocket docking and Classification + pKd
- QSAR models
  - **.kcc**
- Property models : Caco2, T1/2, hERG, LD50, PAINS, PAMPA, PgPinh, PgPsubstrate
  - **.mpc**

Molsoft: Polo Lam, Eugene Raush, Max Totrov

On going collaboration with Novartis Group:
Christine Hajdin, Donovan Chin, Eric Martin

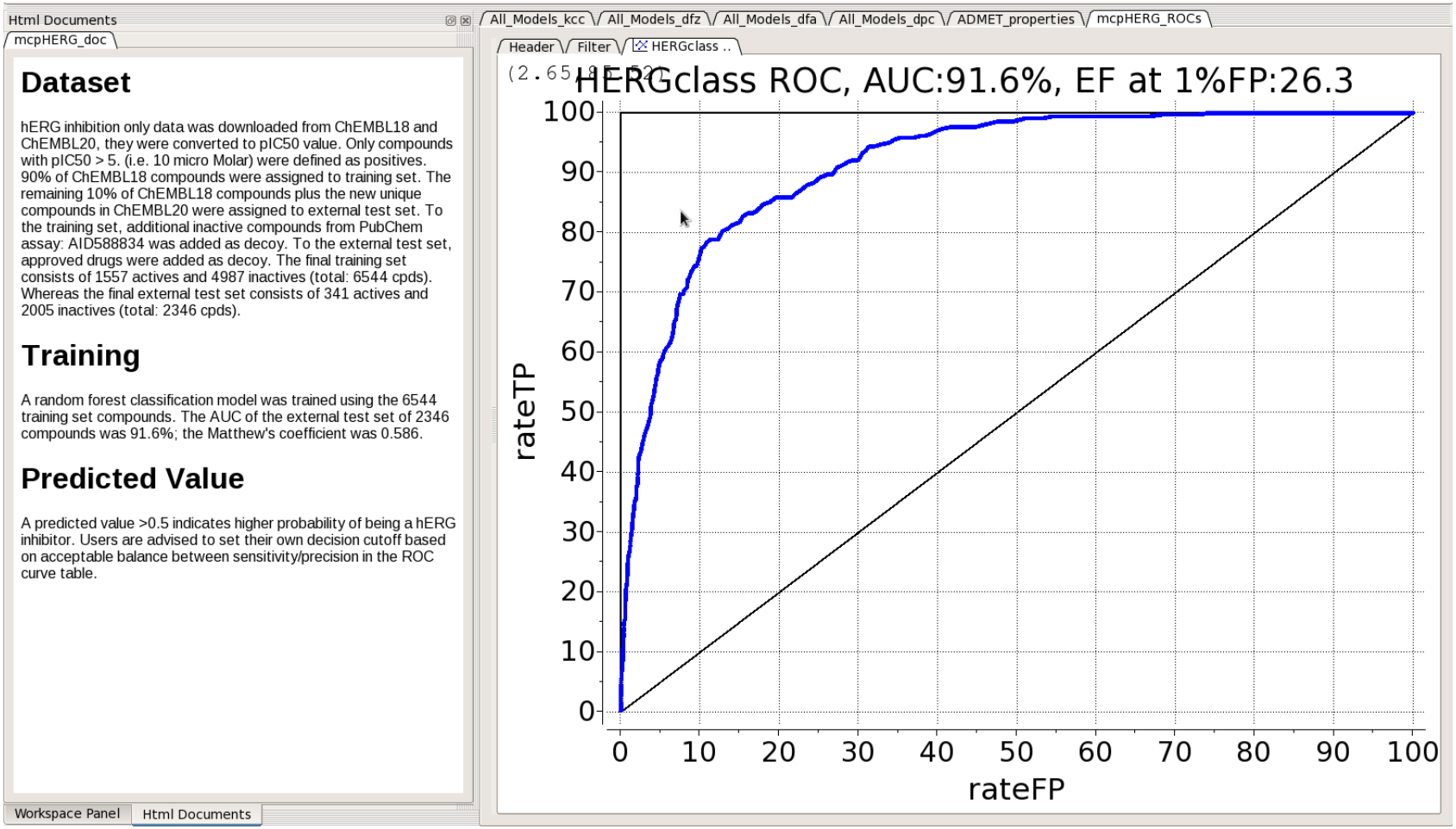
Activity Models: ADRs, Target Screen, Properties

	ADR	Category	Disease	Function	Drug Mechanism
+ 1601 (1 of 2)	Biliary colic	Family A G protein-coupled receptor	Analgesics	Inhibits neurotransmitter release by reducing	Mu opioid receptor agonist: A
+ 1603 (1 of 4)	Biliary colic	Family A G protein-coupled receptor	Alcohol dependence	Inhibits neurotransmitter release by reducing	Kappa opioid receptor agonis
+ 1607 (1 of 3)	Biliary colic	Family A G protein-coupled receptor	Analgesics	Inhibits neurotransmitter release by reducing	Opioid receptors; mu/kappa/
1610	Biliary colic	Membrane receptor	Neuropsychiatric disorders		Sigma non-opioid intracellular
+ 1611 (1 of 2)	Bradycardia	Family A G protein-coupled receptor	Heart failure	Alpha-2 adrenergic receptors mediate the	Adrenergic receptor alpha-2
+ 1613 (1 of 2)	Bradycardia	Family A G protein-coupled receptor	Neuropsychiatric disorders	Alpha-2 adrenergic receptors mediate the	
+ 1615 (1 of 2)	Bradycardia	Family A G protein-coupled receptor	Heart failure	Alpha-2 adrenergic receptors mediate the	Adrenergic receptor alpha ag
+ 1617 (1 of 2)	Bradycardia	Hydrolase	Alzheimer's disease		Cholinesterase activator: Pral
1619	Cardiac arrest	Voltage-gated ion channel	Analgesics	Mediates the voltage-dependent sodium ion	Sodium channel protein type
+ 1620 (1 of 2)	Decreased appetite	Kinase	Breast cancer	Receptor for egf, but also for other members of the	Epidermal growth factor rece
+ 1622 (1 of 4)	Diarrhoea	Family A G protein-coupled receptor	Peripheral Vascular Disease		Prostaglandin E2 receptor EP
+ 1626 (1 of 6)	Diarrhoea	Phosphodiesterase	Chronic lymphocytic leukemia		cAMP-specific 3',5'-cyclic pho
+ 1632 (1 of 4)	Diarrhoea	Hydrolase	Alzheimer's disease	Rapidly hydrolyzes choline released into the	Acetylcholinesterase activato
+ 1636 (1 of 2)	Dizziness	Family A G protein-coupled receptor	Cardiovascular disease, unspecified	Receptor for angiotensin II. Mediates its action by	Type-1 angiotensin II recepto
1638	Dry mouth	Family A G protein-coupled receptor	Anxiety disorder, unspecified	This is one of the several different receptors for 5-	5-hydroxytryptamine recepto
1639	Dry mouth	Electrochemical transporter	Parkinson's disease	Amine transporter, terminating the action of	Dopamine transporter inhibito
+ 1640 (1 of 3)	Dry mouth	Oxidoreductase	Depression	Catalyzes the oxidative deamination of biogenic and	
1643	Dyskinesia	Family A G protein-coupled receptor	Migraine	This is one of the several different receptors for 5-	5-hydroxytryptamine recepto
+ 1644 (1 of 4)	Dyspepsia	Nuclear receptor	Not Available		Peroxisome proliferator-activ
+ 1648 (1 of 4)	Dyspepsia	Phosphodiesterase	Asthma	Regulates the levels of camp in the cell.	
+ 1652 (1 of 2)	Dyspepsia	Phosphodiesterase	Colorectal cancer	Hydrolyzes both cyclic amp (camp) and cyclic gmp	cGMP-dependent 3',5'-cyclic
+ 1654 (1 of 2)	Dyspepsia	Phosphodiesterase	Bronchial asthma	Hydrolyzes both cyclic amp (camp) and cyclic gmp	Phosphodiesterase 3A inhibit
1656	Ejaculation disorder	Family A G protein-coupled receptor	Neurologic and psychiatric diseases	This is one of the several different receptors for 5-	5-hydroxytryptamine recepto
- 1657 (2)	Electrocardiogram QT	Voltage-gated ion channel	Cardiac arrhythmias	Pore-forming (alpha) subunit of voltage-gated	HERG blocker: Amiodarone, D
1658	Electrocardiogram QT	Voltage-gated ion channel	Cardiac arrhythmias	Pore-forming (alpha) subunit of voltage-gated	HERG blocker: Amiodarone, D
+ 1659 (1 of 2)	Electrolyte imbalance	Lyase			Carbonic anhydrase 13 inhibit
1661	Extrapyramidal disorder	Voltage-gated ion channel	Analgesics		Voltage-dependent T-type ca
+ 1662 (1 of 4)	Flushing	Family A G protein-coupled receptor	Cancer, unspecific	Receptor for gonadotropin releasing hormone	Gonadotropin-releasing horm
+ 1666 (1 of 2)	Flushing	Voltage-gated ion channel	Analgesics	Voltage-sensitive calcium channels (vscc) mediate	Voltage-gated N-type calcium
+ 1668 (1 of 2)	Gastrointestinal disorder	Enzyme	Analgesics	Catalyzes the nadph-dependent reduction of a wide	Aldose reductase inhibitor: S
1670	Gestational hypertension	Family A G protein-coupled receptor	Anxiety disorder, unspecified	This is one of the several different receptors for 5-	5-hydroxytryptamine recepto
1671	Gynaecomastia	Oxidoreductase	Alopecia, unspecified	Converts testosterone into 5-alpha-	3-oxo-5-alpha-steroid 4-dehy
1672	Haemolytic anaemia	Lyase			Carbonic anhydrase 3 inhibito
+ 1673 (1 of 4)	Hepatitis	Nuclear receptor	Anxiety disorder, unspecified	Orphan receptor; Its natural ligand is probably	
+ 1677 (1 of 2)	Hyperhidrosis	Family A G protein-coupled receptor	Analgesics	Receptor for the neuropeptide nocipeptin/orphanin	Kappa-type opioid receptor a
1679	Hypokalaemia	Electrochemical transporter	Crohn's disease, unspecified	Plays a critical role in the sodium-dependent	Ileal sodium/bile acid cotrans
1680	Mania	Oxidoreductase	Major depressive disorder	Catalyzes the oxidative deamination of biogenic and	Amine oxidase [flavin-contain
+ 1681 (1 of 6)	Myalgia	Nuclear receptor	Hyperglycemia	Receptor that bind peroxisome proliferators such as	Peroxisome proliferator-activ
1687	Pain	Cytochrome P450	Breast cancer	Catalyzes the formation of aromatic c18 estrogens	Aromatase inducer: Chlorphe
+ 1688 (1 of 2)	Palpitations	Family A G protein-coupled receptor	Asthma	Receptor for adenosine. The activity of this receptor	Adenosine A2b receptor anta
1690	Purine metabolism disorder	Hydrolase	Atrial fibrillation and flutter	This is the catalytic component of the active	
+ 1691 (1 of 2)	Respiratory depression	Ligand-gated ion channel	Convulsions	The glycine receptor is a neurotransmitter-gated ion	Glycine receptor subunit alph

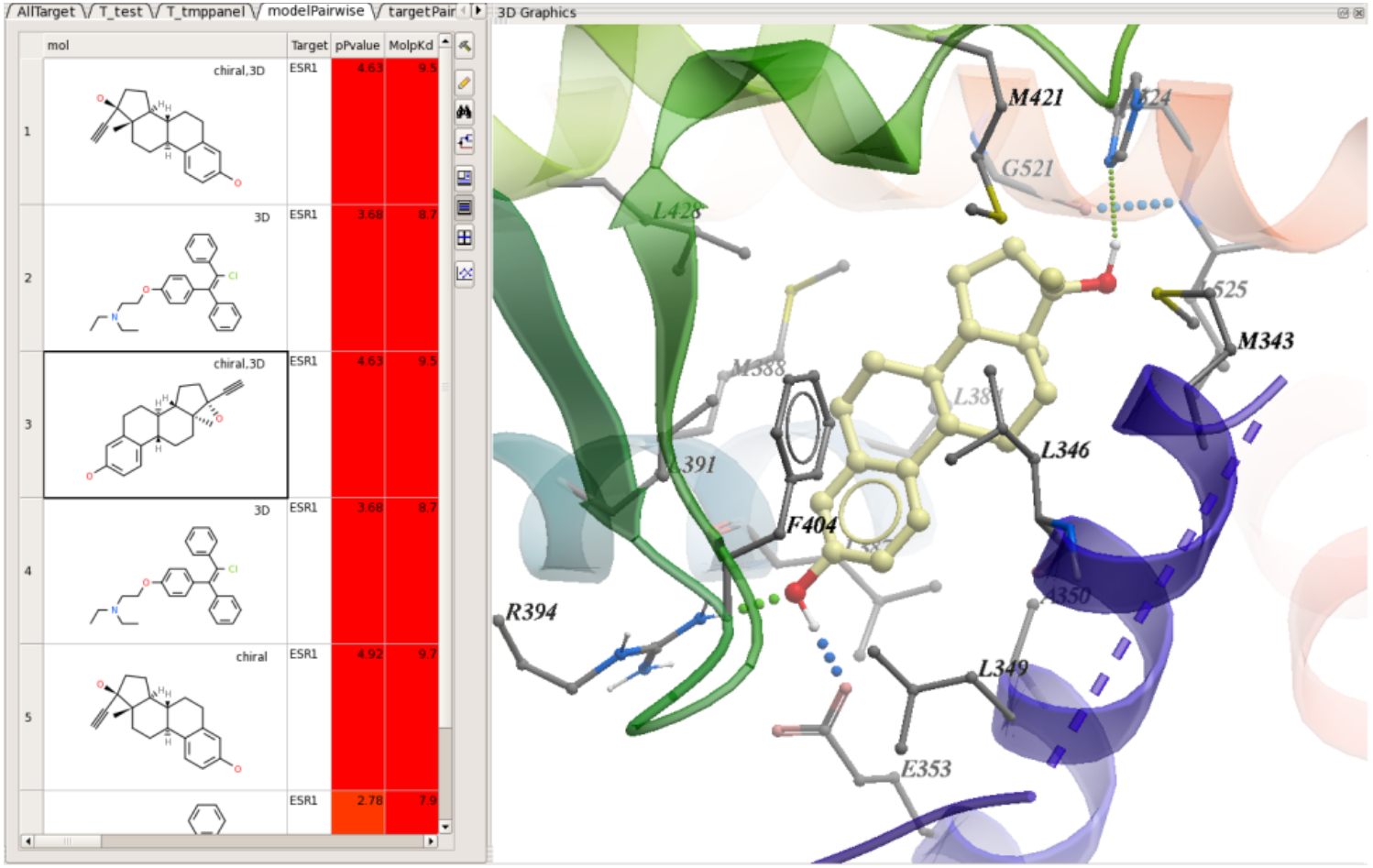
MolScreen

Individual model's detail

- Load Panel -> Check Model Performance:



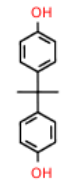
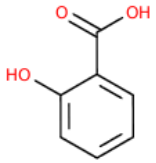
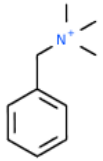
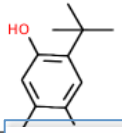
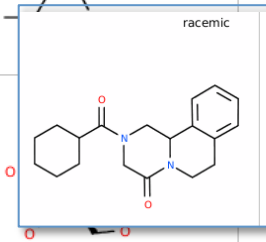
Predicted Pose (docking type)



- Live results for the dfa, dpc, dfz models.

Endocrine Disruption: Testing TiPED

The Tiered Protocol for Endocrine Disruption (TiPED)

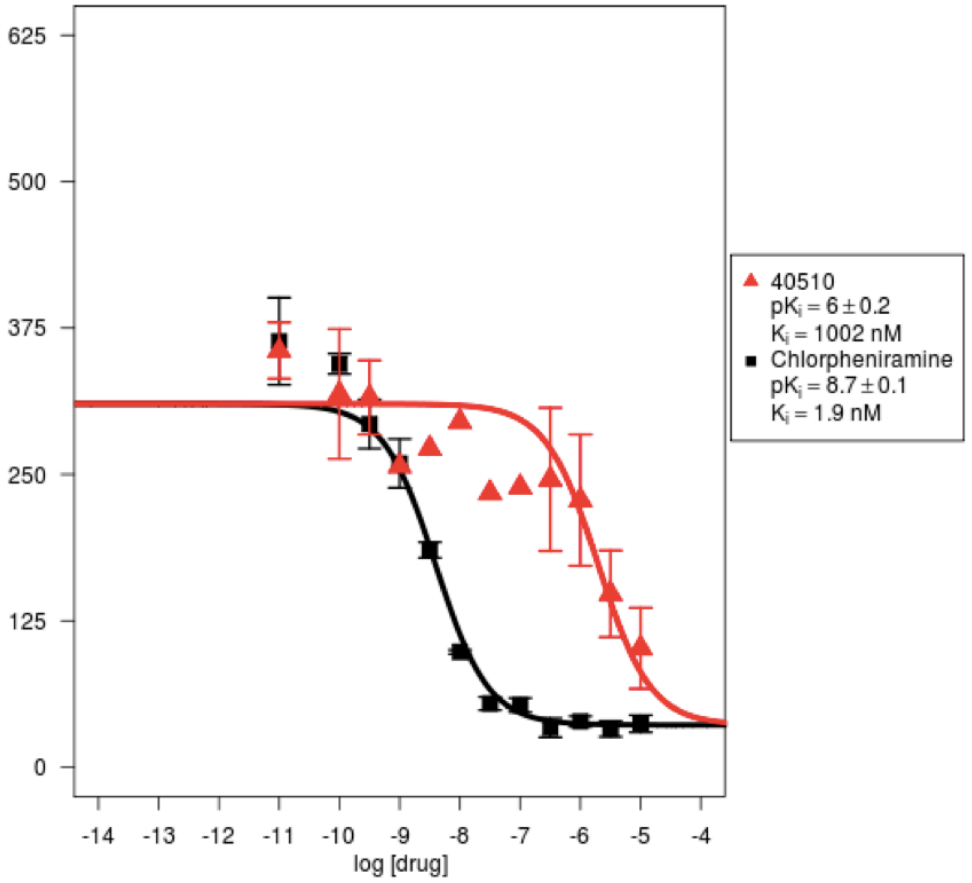
mol	Molecule Name	dfaPXR	dfaANDR	dfaESR1	dfaESR2	dfaGCR	dfaMCR	dfaPRGR	dfaTHA	dfaTHB	dfaKCNH2	dfaNR1H2	dfaNR1H3	dfaPPARA	dfaPPARD	dfaPPARG
	Bisphenol A	4.645	ND	6.737	6.512	6.169	4.882	5.836	4.939	5.557	3.586	ND	ND	ND	ND	2.425
	Salicylic acid	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	Benzyltrimethylammonium	ND	ND	ND	ND	ND	ND									
	2,5-tert-butylhydroquinone	5.082	5.233	ND	ND	ND	ND									
	racemic	1	7.375	7.366	7.269	6.983	6.863	6.748	6.714	6.689	6.686	6.639	6.247	6.175	6.156	6.092	6.005	5.9	5.88	5.872	5.857	5.854	5.787

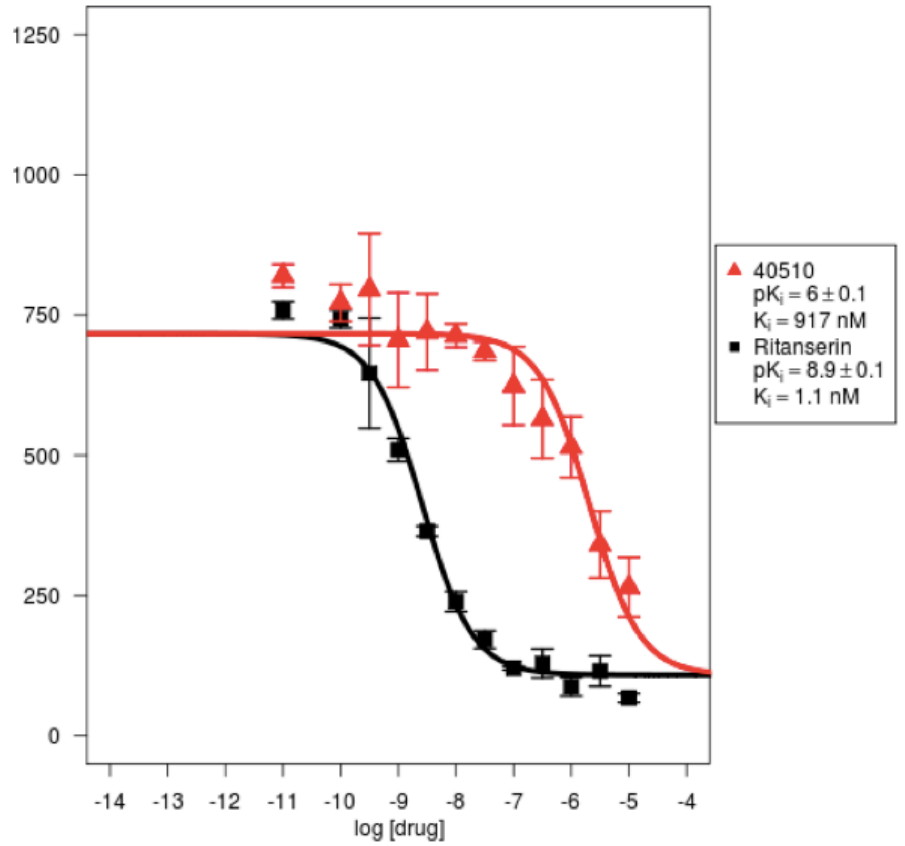
Schug, Abagyan, Blumberg, et al. **Designing endocrine disruption out of the next generation of chemicals** *Green Chem*istry, 2013, 15, 181-198

Maricel Maffini, Bruce Blumberg et al. (2015, in preparation)

Praziquantel target screen

Two activities confirmed!



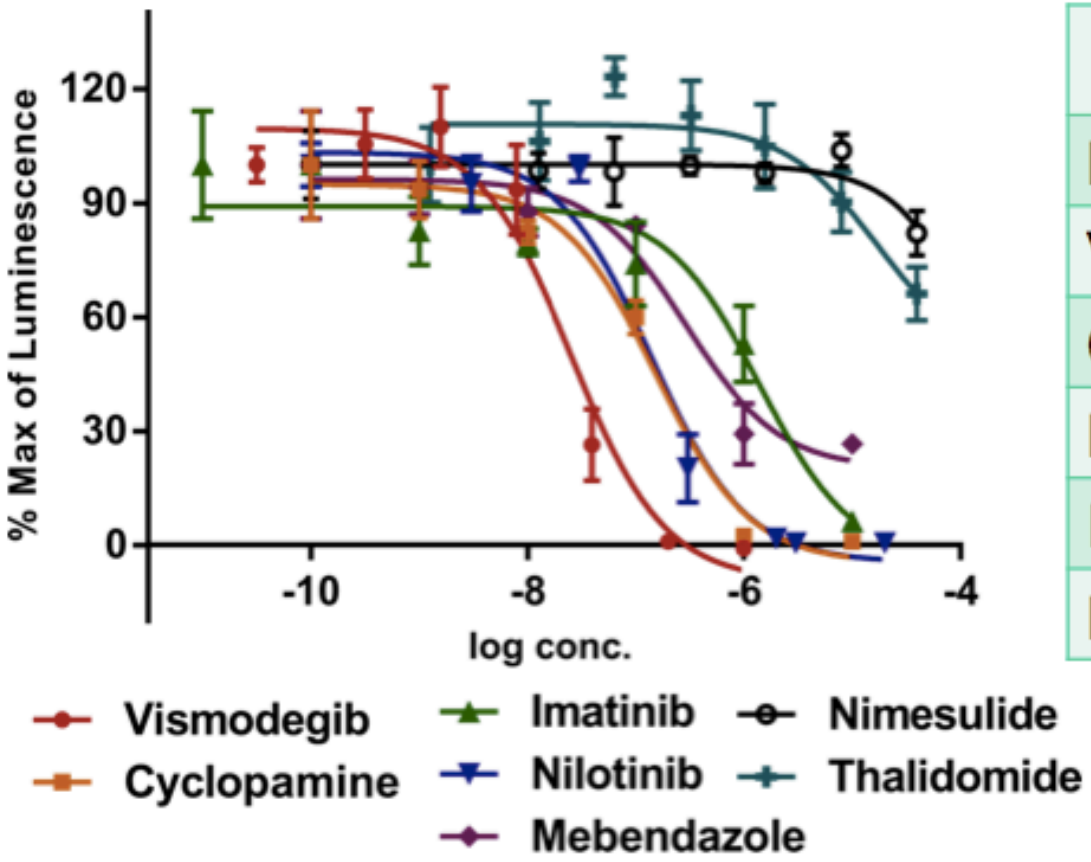


~500 nM activities for a single stereoisomer

Service by Bryan Roth Center, UNC

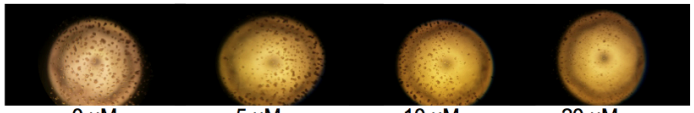
Three drugs identified as SMO modulators

Kirti Kandhwal



Effect of Drugs on Hedgehog (Hh) Pathway Activity (Gli- assay)

Table 1	
Drug	**IC50 (nM)**
Vismodegib	25
Cyclopamine	153
Nilotinib	140
Imatinib	1402
Mebendazole	287



Effect of Nilotinib on Neurosphere Formation (PDX-05 cells)