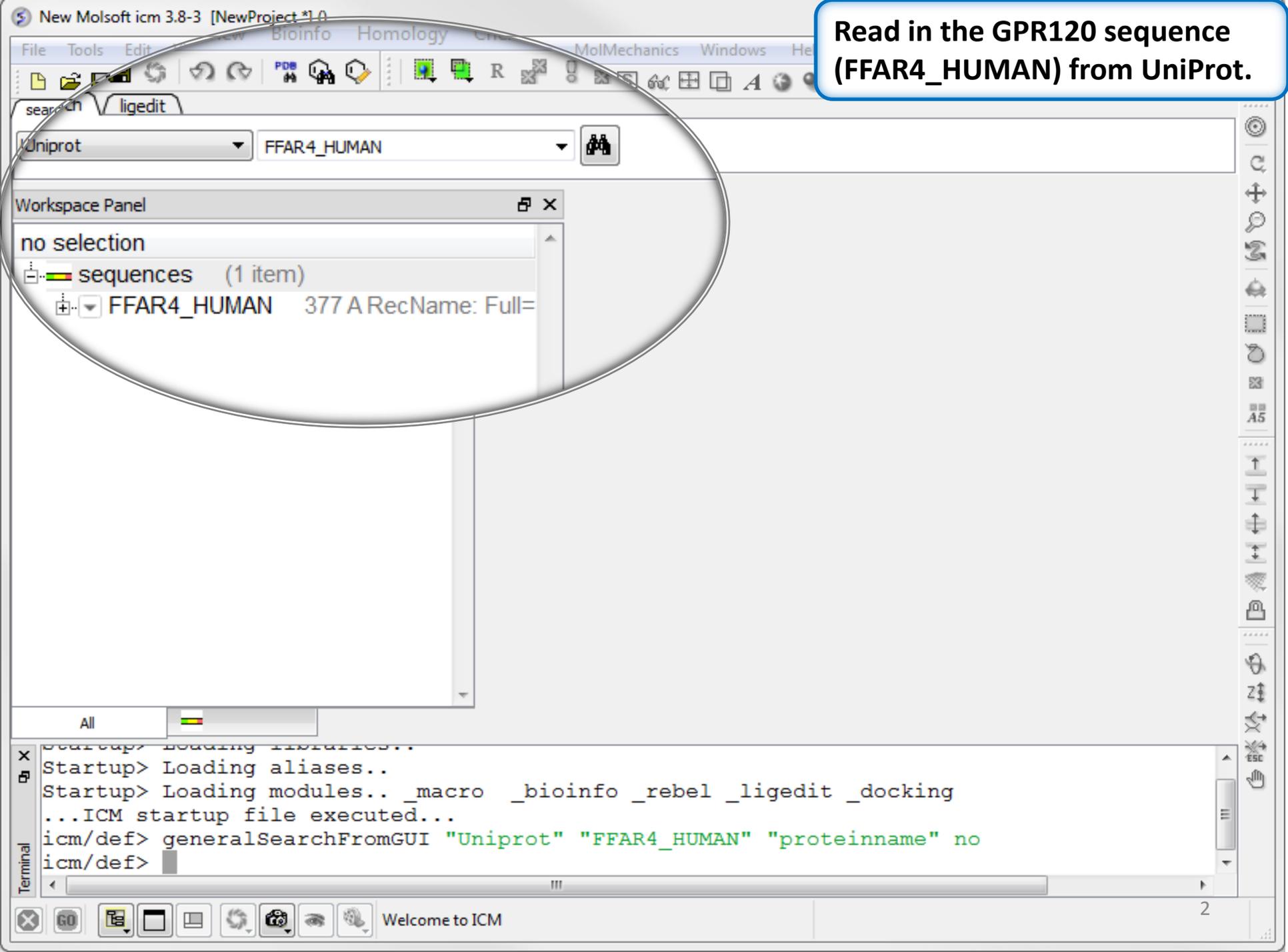


# Homology Modeling

## GPR120 Model – GPCR Modeling

GPR120 is a member of the Class A GPCRs. It 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.

Read in the GPR120 sequence  
(FFAR4\_HUMAN) from UniProt.



Read in the the template structure 3P0G.

3p0g Molsoft icm 3.8-3 [NewProject \*] (1 object)

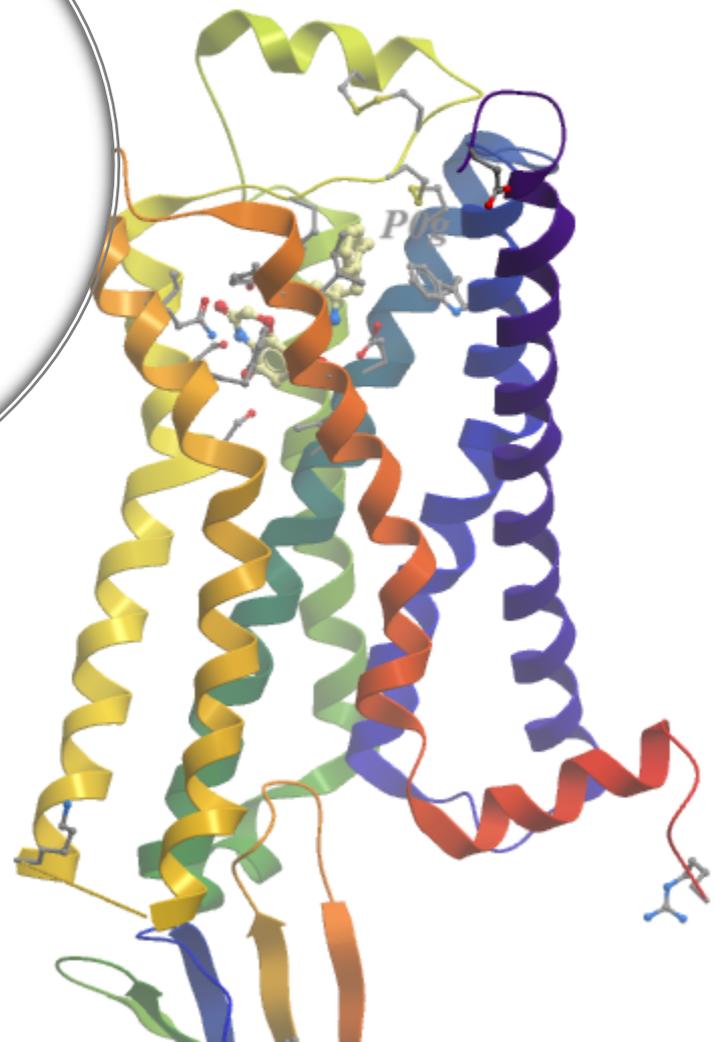
File Tools Edit View Bioinfo **Chemistry** Docking MolMechanics

display layout labels meshes search ligedit

PDB Search 3p0g Append

Workspace Panel

- no selection
- objects (1 item)
  - 3p0g [1\*] XR; 3.5Å
    - a 284 A 4 sites ADRB2\_HUMAN
    - b 121 A 2 sites
    - ap0g H 8-[(1r)-2-[[1,1-dimethyl-2-phenylethyl]amino]ethoxy]propane
    - beta-2 adrenergic receptor, lysozyme
  - sequences (1 item)
    - FFAR4\_HUMAN 377 A RecName: Full=



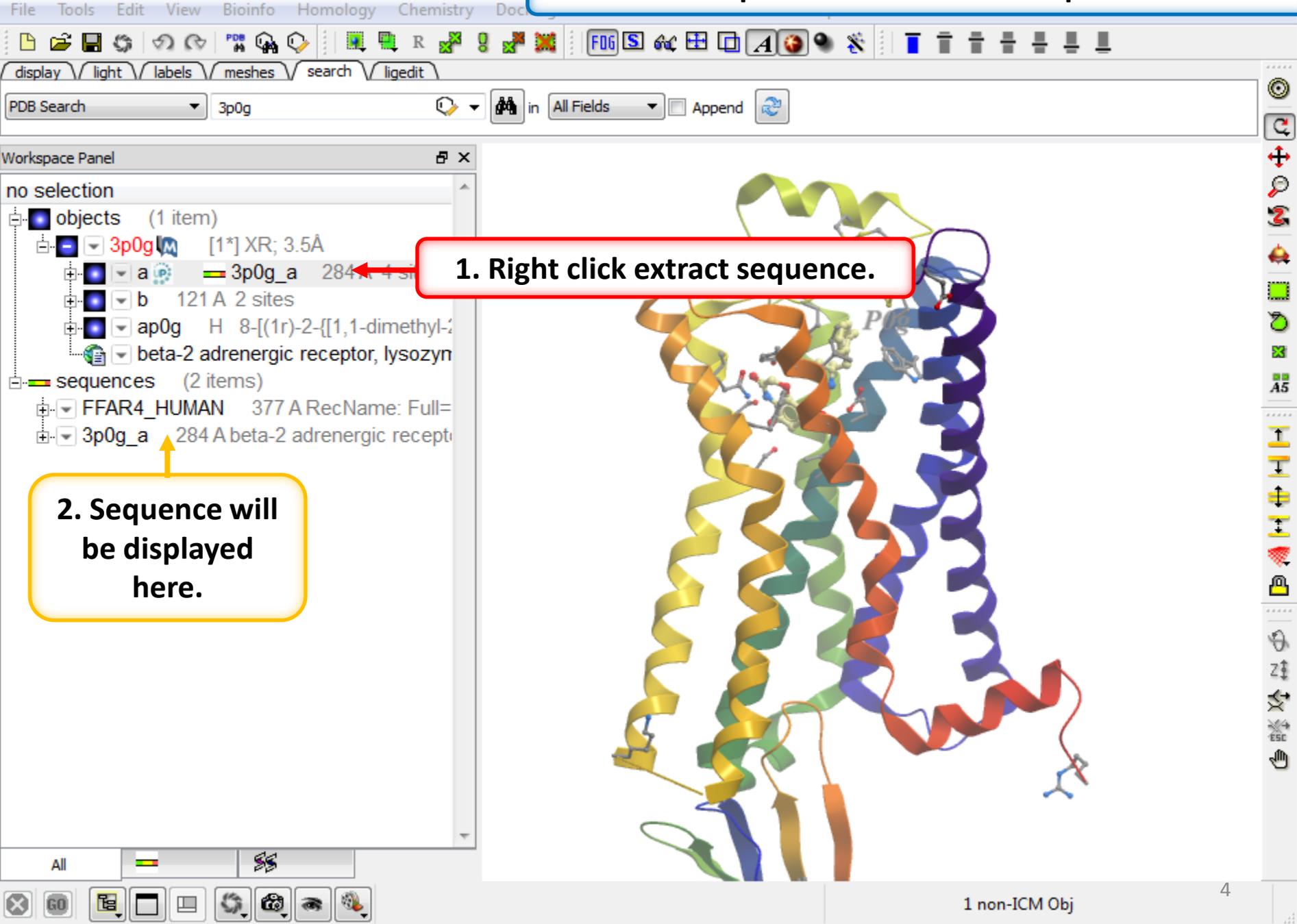
All

GO

Welcome to ICM

1 non-ICM Obj

Extract the sequence from the 3D template structure.



1. Right click extract sequence.

2. Sequence will be displayed here.

Align the template and query sequence.

File Tools Edit View Bioinfo Homology Chemistry Docking MolMechanics

display light labels meshes search ligedit

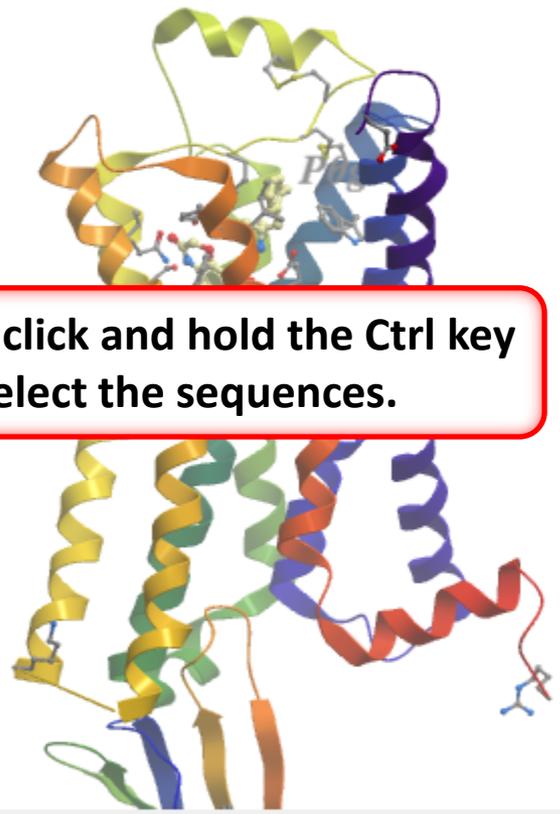
PDB Search 3p0g in All Fields Append

Workspace Panel

- 1 3p0g a,1 FFAR4 HUMAN
  - objects (1 item)
    - 3p0g [1\*] XR; 3.5Å
      - a 3p0g\_a aln 284 A
      - b 121 A 2 sites
      - ap0g H 8-[(1r)-2-[[1,1-dimethyl-2-(2,4,6-trimethylphenyl)pyridin-5-yl]ethoxy]propanoic acid
      - beta-2 adrenergic receptor, lysozyme
  - sequences (2 items)
    - FFAR4\_HUMAN 377 A RecName: Full
    - 3p0g\_a 284 A beta-2 adrenergic receptor
  - alignments (1 item)
    - aln ID=23% pP=9.1

1. Double click and hold the Ctrl key to select the sequences.

2. Alignment will be displayed here.



Alignments

FFAR4_HUMAN	45	.V#.LI###.##GNV###.##A+..R...#T'###..L#CADL#
3p0g_a	38	TVLVLI FAVSLLGNVCALV LVAR-RRRRGATA CLV LNLFCADLL
		IVMSLIVLAI VFGNVLVITAI AKFERLQTVTNYFITS LACADLV

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#  
 SLEQANRTRFPFFSDVKGDHRLVLAAVETTVLVLI FAVSLLGNVC  
 -----DVTQQRDEVWVVGMI VMSLIVLAI VFGNVL

FFAR4\_HUMAN  
 3p0g\_a  
 3p0g\_a

61 ALVLVAR-RRRRGATA CLVLNLF CADLLFISAIPLVLA VRWTE-AWLLG P VACHLLFYVM  
 54 VITAI AKFERLQ TVTNYFITSLACADLVMGLAVV PFGAAHILMKMWT FGNFWCFWTSID

FFAR4\_HUMAN  
 3p0g\_a  
 3p0g\_a

119 TLG SVTILTLAAVSLERMVCIVHLQ RGV RGPGRRARAVLLALIWGYS AVAALPLCVF FR  
 114 VLCVTAS IETLCVIAVD RYFAITSPFKYQ SLLTKNKARV IILMVWIVSGLT SFLPIQMHW

FFAR4\_HUMAN  
 3p0g\_a  
 3p0g\_a

179 --VVPQRLPGADQEISICTLIWPTIPGEISWDV SFVTLNFLV PGLVIVISYSKILQ TSEH  
 174 YRATHQEAINCYAEETCCDFF-----TNQAYAIASSIVSFYVPLVIMVFVYSRVFQ EAK-

FFAR4\_HUMAN  
 3p0g\_a  
 3p0g\_a

237 LLDARAVVTHSEITKASRKRLTVSLAYSESHQIRV SQQDFRLFR TFLFLMV SFFIMWSP I  
 266 -----LKEHKALKTLGIIMGTFTLCWLPF

FFAR4\_HUMAN  
 3p0g\_a  
 3p0g\_a

297 IITILLILIQNFKQDLVIWPSLFFWVVAFTFANSALNPILYNM T LCRNEWKKIFCCFWFP  
 290 FIVNIVHVIQD----NLIRKEVYILLNWIGYVNSGFNPLIYCRSPDFRIAFQELLCLRR-

# Homology/Batch Build and Refine. Use the Quick Test method first.

The screenshot shows the Molsoft icm 3.8-2 software interface. The main window displays a protein structure (beta-2 adrenergic receptor) in a ribbon representation. A dialog box titled "Launch background model building and refinement" is open, showing the following settings:

- Sources
- Sequence: FFAR4\_HUMAN
- 3D template: a\_3p0g.a
- Alignment: aln
- Quick Test
- First use quick test mode to adjust alignment, then run full build/refine job

Below the dialog box, the "Alignments" panel shows the following sequence alignment:

Sequence	Length	Alignment
FFAR4_HUMAN	45	.V#.LI###.##GNV###.##A+..R...#T.###..L#CADL#
3p0g_a	38	TVLVLIFAVSLGNCALVLVAR-RRRRGATACLVLNLFCADLL
3p0g_a	38	IVMSLIVLAIIVFGNVLVITAIKFERLQTVTNYFITSLACADLV

The alignment shows a red box around the sequence "TTLVLIFAVSLGNCALVLVAR-RRRR" and a green box around "GATACLVLNLFCADLL". Below the alignment, there is a red bar with a purple segment, likely representing a quality or confidence score.

You will be notified when the modeling job has completed.

File Tools Edit View Bioinfo Homology Chemistry



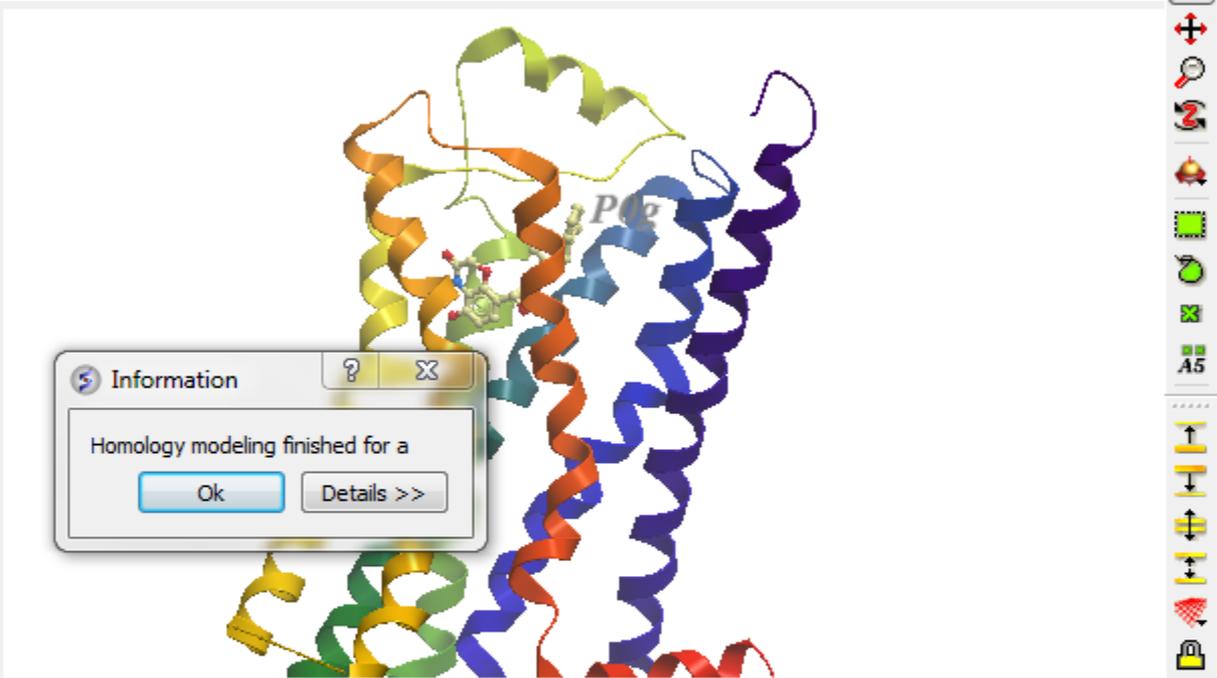
display light labels meshes search ligedit



Workspace Panel

no selection

- objects (1 item)
  - 3p0g [1\*] XR; 3.5Å
    - a 3p0g\_a aln 284
    - b 121 A 2 sites
    - ap0g H 8-[(1r)-2-[[1,1-dimethy
    - beta-2 adrenergic receptor, lysoz
- sequences (2 items)
  - FFAR4\_HUMAN 377 A RecName: Ful
  - 3p0g\_a 284 A beta-2 adrenergic rece
- alignments (1 item)
  - aln ID=20% pP=5.8



Information

Homology modeling finished for a

Ok Details >>

Alignments

			.V#.LI###.##GNV###.##A+..R...#T.###..L#CADL#
FFAR4_HUMAN	45		TVLVLIFAVSLGNCALVLR-RRRRGATACLVLNLFCADLL
3p0g_a	38		IVMSLIVLAIIVFGNVLVITAIKFERLQTVTNYFITSLACADLV
3p0g_a			



2 Mol 1 Obj





# Loop Modeling

- In this example we will use the high precision loop modeling method Arnautova *et al* Proteins. 2010, 79:477.
- We will 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.

Read in two Anthrax Protective Antigen structures PDB codes 1acc and 1tzo.

1. Use the search tab to load 1acc and 1tzo

The screenshot shows the ICM software interface. At the top, the menu bar includes File, Edit, View, Bioinfo, Tools, Homology, Chemistry, Docking, and MolMechanics. Below the menu is a toolbar with various icons. The main workspace is divided into a left panel and a right 3D view.

**Workspace Panel:** The top of the panel shows "no selection". Below that, the "objects" list contains 3 items:

- 1acc** [1] XR; 2.1Å; anthrax protecti...
  - a 1acc\_a 665 A 2 sites PAG\_...
  - aca M calcium +2
  - aca2 M calcium +2
  - W (388 water molecules)
- 1tzo** [2\*] XR; 3.6Å; protective antig...
  - a 1tzo\_a 553 A 2 sites PAG\_...
  - b 1tzo\_a 553 A 2 sites PAG\_...
  - c 1tzo\_a 553 A 2 sites PAG\_...
  - d 1tzo\_a 553 A 2 sites PAG\_...
  - e 1tzo\_a 553 A 2 sites PAG\_...
  - f 1tzo\_a 553 A 2 sites PAG\_E...
  - g 1tzo\_a 553 A 2 sites PAG\_...
  - h 1tzo\_a 553 A 2 sites PAG\_...
  - i 1tzo\_a 553 A 2 sites PAG\_E...
  - j 1tzo\_a 553 A 2 sites PAG\_E...
  - k 1tzo\_a 553 A 2 sites PAG\_I...
  - l 1tzo\_a 553 A 2 sites PAG\_E...
  - m 1tzo\_a 553 A 2 sites PAG\_...

**3D View:** The right side of the workspace shows a 3D ribbon representation of the protein structures. One structure is colored yellow, and another is colored green. They are shown in a complex, intertwined arrangement.

**Terminal:** The bottom of the interface features a terminal window with the following commands:

```
icm/1tzo> color Res(a_*.//DD) object ribbon
icm/1tzo> color Res(a_*.//DD) & a_*.//c* object xstick
icm/1tzo> undisplay store a_1tzo.o
icm/1tzo> cool a_1tzo.a
icm/1tzo> delete newAli
icm/1tzo> center static as_graph
icm/1tzo>
```

Delete everything except the first molecule a in each object. The easiest way to do this is to type `delete a_*.!a` in the terminal window.

The screenshot displays the MolScribe software interface. The main window shows a ribbon representation of a protein complex, with one subunit colored yellow and the other green. The Workspace Panel on the left lists the objects: '1acc' (anthrax protective antigen) and '1tzo' (protective antigen), each with a corresponding 'a' molecule. The terminal window at the bottom shows the following commands and their outputs:

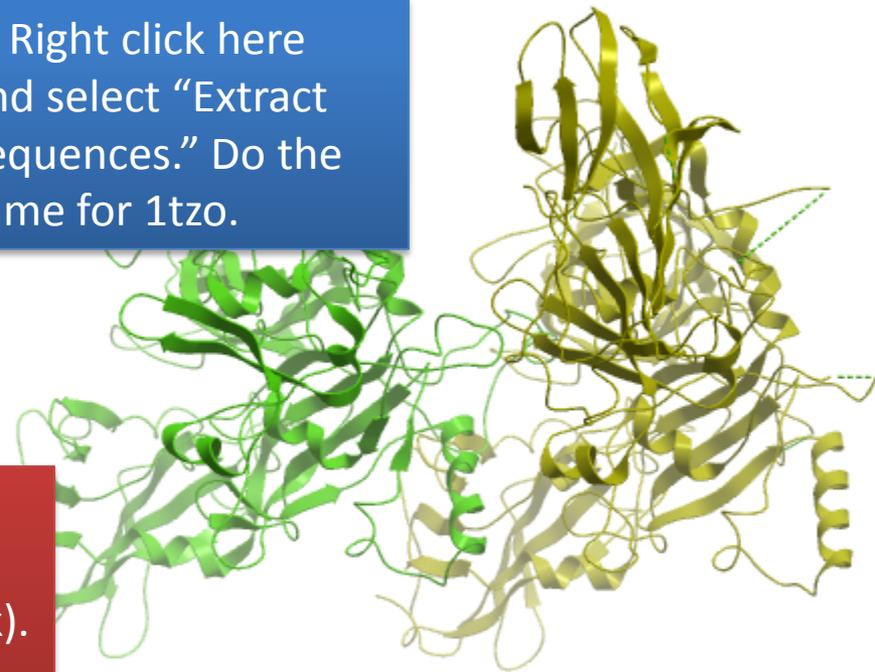
```
icm/1tzo> display ribbon a_1acc.a  
icm/1tzo> display ribbon a_1tzo.a  
icm/1tzo> delete newAli  
icm/1tzo> delete gui sequence  
Info> 2 sequences deleted  
icm/1tzo> delete a_*.!a  
Info> 431 Molecules deleted  
icm/1tzo>
```

The command `delete a_*.!a` is circled in red in the terminal window. The status bar at the bottom right indicates '2 non-ICM Obj' and the page number '12'.

Extract sequences from both structures and align.

1. Right click here and select "Extract Sequences." Do the same for 1tzo.

2. Select the sequences (hold control key and click). Then right click and align sequences.



Workspace Panel

no selection

- objects (3 items)
  - 1acc [1] XR; 2.1Å; anthrax protecti
  - a 1acc\_a newAli 665 A
  - 1tzo [2\*] XR; 3.6Å; protective antig
  - a 1tzo\_a newAli 553 A 2
- sites (1 item)
- sequences (2 items)
  - 1acc\_a 665 Amino 2 sites
  - 1tzo\_a 553 Amino 2 sites
- alignments (1 item)
  - newAli ID=94% pP=65

```
Info> alignment newAli
icm/1tzo> align $$_ou
Info> alignment 'new
icm/1tzo>
```

ID= 93.9 [r\_2out] pP=65.1 ) created

Alignments

newAli

1tzo_a	250	GVNISTGKRIITSEVYIGNAEVIRASFPDITGGVYSGAGFQNSGNSQTVADIDGSDERGERIWAETFGSDNTADITKRNDRNIVNIGIQA
1acc_a	374	IYNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYYP SKNLAPIALNAQD...SSTPITMNYNQFLELEKTKQLRLDTDQVYG

Select the loop region residues 410-414 and superimpose.

4. Click on the display tab

5. Click on the superimpose button

1. Click and drag over residues 410-414 to select

2. Click here to display alignment panel

3. Click here to Propagate selection to All Sequences

```
271  LSKNETISKN
293  TTSRTHTSE
303  VVSAGFSNSN
329  SSTVAIDHSL
339  SLAGGLNTAD
357  TARLNANIRY
367  VNTGTAPIYN
377  VLPTTSLVLG
387  KNQTLATIKA
397  KENQLSQILA
407  PNNYYP SKNL
417  APIALNAQDD
427  FSSTPITMNY
437  NQFLELEKTK
447  QLRRLTDQVY
457  GNIATYNFEN
467  GRVRVDTGSN
```

```
Info> 5 50%Dev=0.4 wRmsd=0.3 Rmsd=0.3 for
Info> 6 50%Dev=0.4 wRmsd=0.2 Rmsd=0.3 for
Info> 15 of 20 atoms superimposed (i_2out,
icm/ltzo>
```

```
newAli
1cc20_a 170 NDNDGIEEDDEVEGTVDYNNKRIEDSEWNIHNRGDIKNSPEEN
STADPYSDFEKVTGRIDKNVSPPEARHPLVAAAYPIVHVDMENIILSKNE
1acc a 227 STADPYSDFEKVTGRIDKNVSPPEARHPLVAAAYPIVHVDMENIILSKNE
```

Select

Propagate to ALL sequences

Rv Consensus x

Observe that Proline 412 is in the cis conformation in the high resolution structure 1acc and the trans conformation in the low resolution structure 1tzo.

1. Click and hold stick button to color by object.

loopModelExample.icb Molsoft icm 3.7-2c [H:\icmd\man\loopModelExample.\*] (2 objects)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics

display light labels meshes search ligedit movie

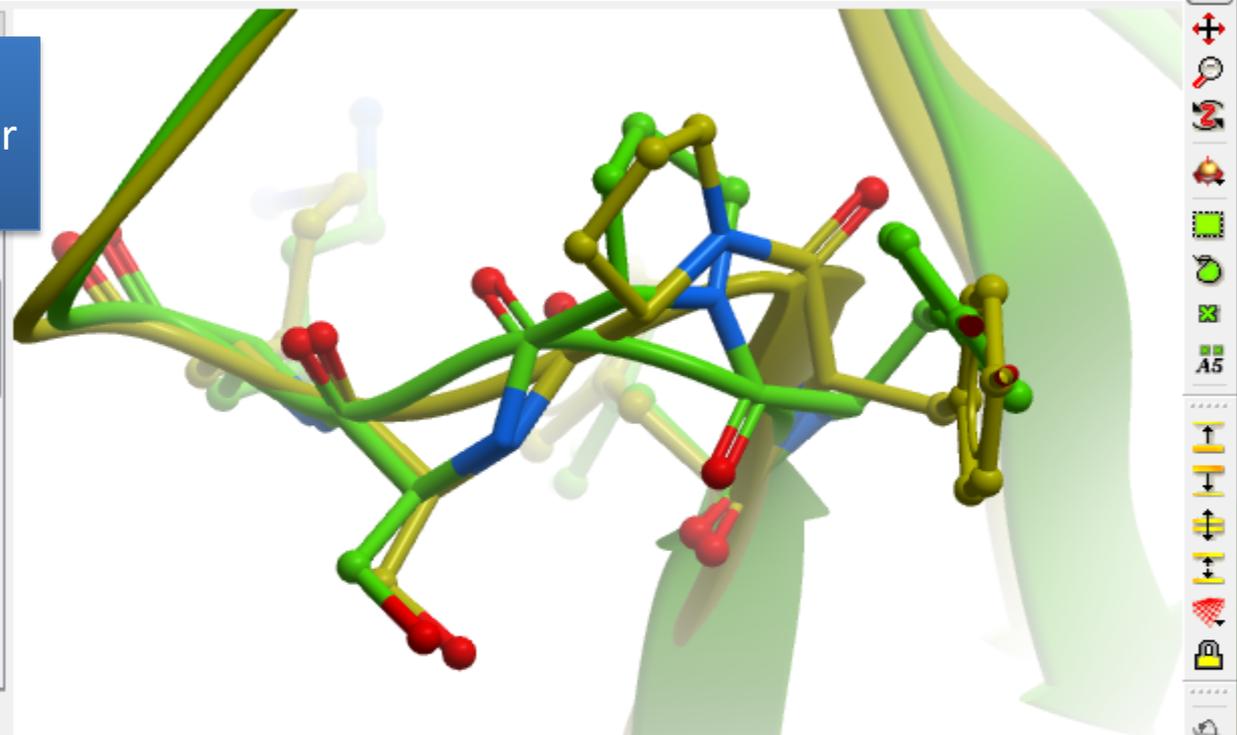
RES AS ATOM hbt VAR SITE

Workspace Panel

no selection

261	I
271	I
293	T
303	V
329	<u>SSTVAIDHSL</u>
339	SLAGGLNTAD
357	TARLNANIRY
367	VNTGTAPIYN
377	VLPTTSLVLG
387	KNQTLATIKA
397	KENQLSQILA
407	PNNYYPSKNL
417	APIALNAQDD
427	FSSTPITMNY
437	<u>NQFLELEKTK</u>
447	QLRLDTDQVY
457	GNIATYNFEN
467	GRVRVDTGSN

All



```
icm/1tzo> as_graph = Atom( Res( as_graph ) )
icm/1tzo> display xstick as_graph
icm/1tzo> color as_graph & a_*.//c* object xstick
icm/1tzo>
```

Alignments

1tzo_a	170	<u>STADSPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNE</u>
1acc_a	227	<u>STADSPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNE</u>

Selection

Propagate to ALL sequences

Rv.Consensus x

So let's re-model (1tzo) the low conformation structure and see if the cis conformation is more favorable... First we need to convert the structure to an ICM object.



Workspace Panel

no selection

- objects (3 items)
  - 1acc [1] XR; 2.1Å; anthrax protecti
    - 1acc\_a newAli 665 A
  - 1tzo [2\*] XR; 3.6Å; protective antig
    - 1tzo\_a newAli 553 A
- sites (1 item)
- sequenc
- 1acc\_a
- 1tzo\_a
- alignmer
- newAli

1. Right click here and select convert object.

Convert protein object to ICM

- delete water
- optimize hydrogens  optimize HisProAsnGlnCys
- replace the original  display the result

Ok Cancel Help

2. Choose the options shown here.

```
icm/1tzo> as_graph = Atom( Res( as_graph ) )
icm/1tzo> display xstick as_graph
icm/1tzo> color as_graph & a_*./c* object xstick
icm/1tzo>
```

Alignments

newAli	1tzo_a	170	NDNDGIEDEEVEGTVDYNNKRIEESWIKIENKQDINRQSEEM
			STASDPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNE
	1acc_a	227	STASDPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNE

Selection

- Propagate to ALL sequences

Rv.Consensus x

Select the loop region residues 410-414 in 1tzo. Make sure the loop in 1acc is not selected.

loopModelExample.icb Molsoft icm 3.7-2c [H:\icmd\man\loopModelExample.icb] (2 objects)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics

display light labels meshes search ligedit movie

RES AS ATOM hbt VAR 2 SITE

Workspace Panel

4 Res 1 Mol, 1 Obj

381	<u>TSLVVLGKNQT</u>
391	LATIKAKENQ
401	LSQILAPNNY
411	<u>YPSKNLAPIA</u>
421	LNAQDSSTPI
433	<u>TMNYNQFLEL</u>
443	<u>EKTQQLRLDT</u>
453	DQVYGNIAATY
463	NFENGRVRVD
473	<u>TGSNWSEVLP</u>
483	<u>QIQETTARII</u>
493	FNGKDLNLVE
503	RRIAAVNPSD
513	<u>PLETTRPDMT</u>
523	<u>LKEALKIAFG</u>
533	FNEPNGNLQY
543	<u>QGKDITEFDF</u>
553	NEDQOQSONT

All

Warning> amide group of a\_1tzo.a/^Q670 was flipped  
Warning> amide group of a\_1tzo.a/^N705 was flipped  
icm/1tzo> if yes cool a\_no  
icm/1tzo>

newAli

1tzo_a	170	<u>NDNDGIEEDDEVEGIVDYNNKRIEDSEWIKHNRGDIKNSPEEN</u>
1acc_a	227	<u>STADPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNE</u>

Selection

Propagate to ALL sequences  
Rv.Consensus x

2 Obj [1 non-ICM]

17

Start loop modeling simulation.

1. MolMechanics/Sample loop

2. Choose the "High Precision Sampling in Batch" tab.

Generate Stack of Conformations

Interactive | High Precision Sampling in Batch

Loop Residues:  all Selection (4 res)

thoroughness: 1.00

Sample surrounding sidechains

Truncate remote parts

Ok Cancel

3. Keep the options as shown.

Workspace Panel

4 Res 1 Mol, 1 Obj

objects (3 items)

- 1acc [1] XR; 2.1Å
- a 1acc\_a
- 1tzo [2\*] ICM; 3.6Å
- a 1tzo a newAli 553 A 2

```

174 TVPDRDNDGI
184 PDSLEVEGYT
194 VDVKNKRTFL
204 SPWISNIHEK
214 KGLTKYKSSP
224 EKWSTASDPY
234 SDFEKVTGRI
244 DKNVSPEARH
254 PLVAAYPIVH
264 VDMENIILSK
274 NEDSQTRTIS
291 KNTSTSRHT

```

```

Warning> amide group of a
Warning> amide group of a
icm/1tzo> if yes cool a_no
icm/1tzo>

```

Alignments

newAli

```

1tzo_a 170 STASDPYSDFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNE
1acc a 227 STASDPYSDFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNE

```

Selection

Propagate to ALL sequences

Rv.Consensus: x

Background job will run.

loopModelExample.icb Molsoft icm 3.7-2c [H:\icmd\man\loopModelExample.icb] (2 objects 1 alignment 1 bgnd job)

File Edit View Bioinfo Tools Homology Chemistry Docking MolMechanics Windows Help

display light labels meshes search ligedit movie

RES A5 ATOM hbt VAR 2 SITE S

3.4 1.3 A

Workspace Panel

4 Res 1 Mol, 1 Obj

objects (3 items)

- 1acc [1] XR; 2.1Å; anthrax protecti
- a 1acc\_a newAli 665 A
- 1tzo [2\*] ICM; 3.6Å; protective anti
- a 1tzo a newAli 553 A

174 TVPDRDNDGI

184 PDSLEVEGYT

194 VDVKNKRTFL

204 SPWISNIHEK

214 KGLTKYKSSP

224 EKWSTASDPY

234 SDFEKVTGRI

244 DKNVSPEARH

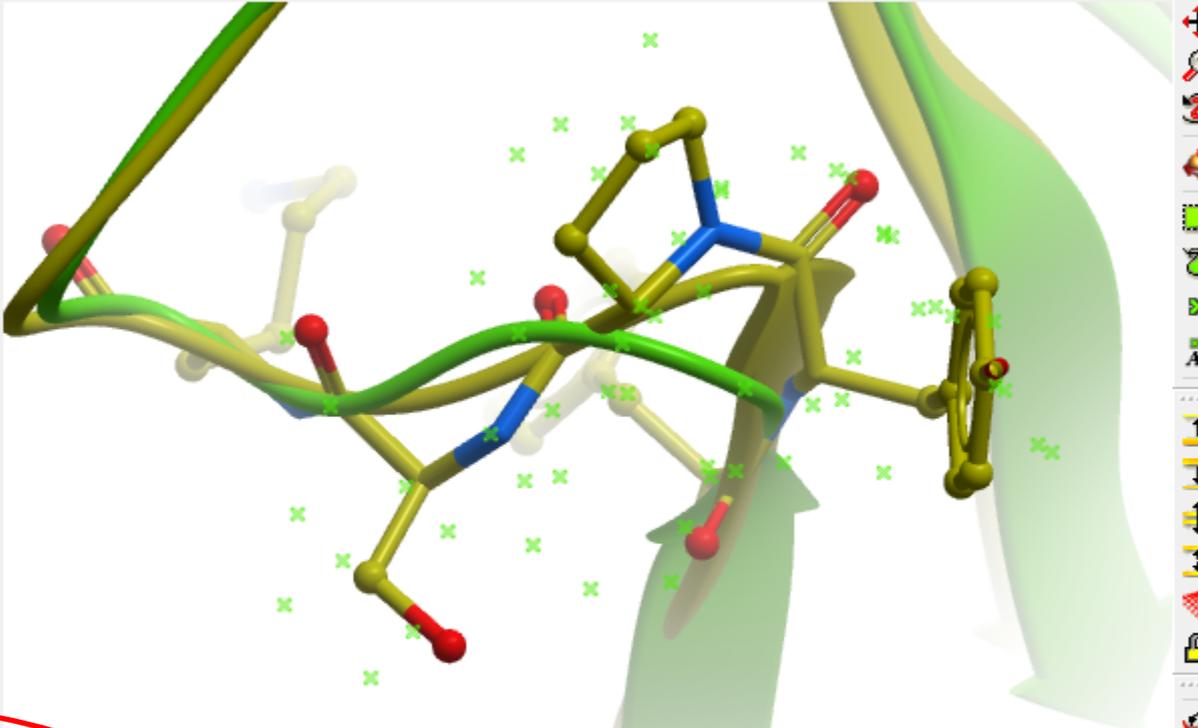
254 PLVAAYPIVH

264 VDMENIILSK

274 NEDSQTRTIS

291 KNTSTSRHT

All



```
icm/1tzo> loopmodelBG as_graph 1. no no
Loop: a/410:413
Info> Loop simulation started in the background. You will be notified when the job is complete.
icm/1tzo>
```

Alignments

newAli

1tzo_a	170	<u>NDNDGIEPDSLEVEGYTVDVKNKRTFLSPWISNIHEKDKNSPEARH</u>
1acc_a	227	<u>STASDPYSDFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNE</u>

Selection

Propagate to ALL sequences

Rv Consensus Y

When simulation has finished – press OK to load the results.

Workspace Panel

4 Res 1 Mol, 1 Obj

objects (3 items)

- 1acc [1] XR; 2.1Å; anthrax protecti
- a 1acc\_a newAli 665 A
- 1tzo [2\*] ICM; 3.6Å; protective anti
- a 1tzo a newAli 553 A

174 TVPDRDNDGI

184 PDSLEVEGYT

194 VDVKNKRTFL

204 SPWISNIHEK

214 KGLTKYKSSP

224 EKWSTADPY

234 SDFEKVTGRI

244 DKNVSPPEARH

254 PLVAAYPIVH

264 VDMENIILSK

274 NEDSQTRTIS

291 KNTSTSRHT

Information

background job 'icmjob000' completed.  
Press OK to load the results

Ok Details >>

icm/1tzo> loopmodelBG as\_graph 1. no no

Loop: a/410:413

Info> Loop simulation started in the background. You will be notified when the job is complete.

icm/1tzo>

newAli

1tzo\_a 170 NDNDGIEPDSLEVEGYTVDYVKNKRTFLSPWISNIHEKNDINRSEEN

STADSPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNE

1acc a 227 STADSPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNE

Selection

Propagate to ALL sequences

Rv Consensus x

2 Obj [1 non-ICM] 100%



