

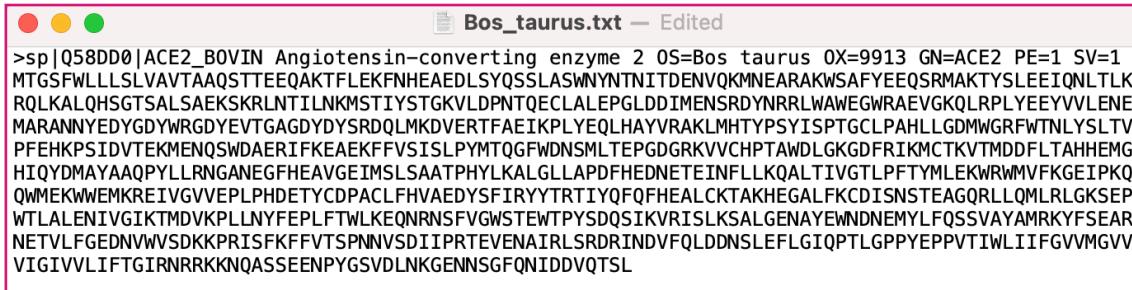
Homology Modelling and Interaction between ligand and receptor

(Here we will be using ACE2 of *Bos taurus* and Spike protein of COVID)

Manas Ranjan Praharaj and Ravi Kumar Gandham

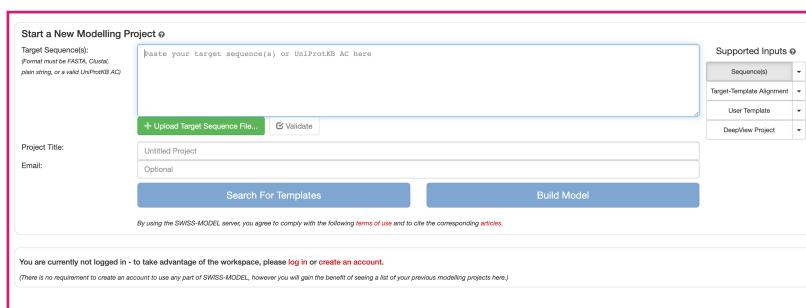
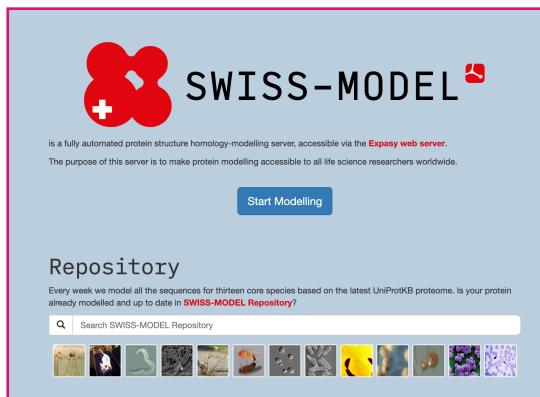
Step 1. Homology modelling of ACE2

- Download the fasta sequence of the ACE2 protein from NCBI

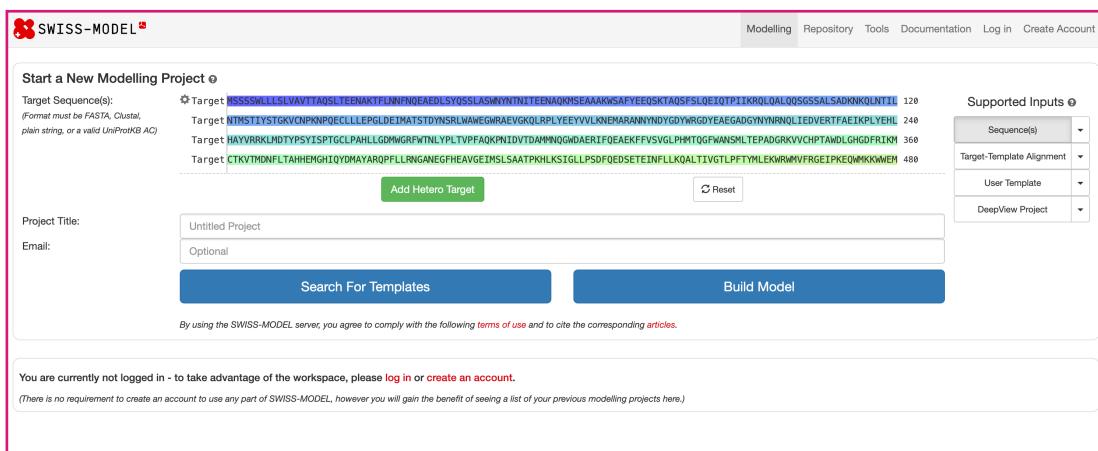


```
>sp|Q58DD0|ACE2_BOVIN Angiotensin-converting enzyme 2 OS=Bos taurus OX=9913 GN=ACE2 PE=1 SV=1 MTGSFWLLLVLAVTAQSTTEEAKTFLKEFNHEADELSYQSSLASWNYNTNTIDENVKMNEARAKWSAFYEEQSRAKTYSLLEIQNLTLLK RQLKALQHSGTSALSAEKS KRLNTILNKMSTIYSTGKVLDPTQEC LAEPGLDDIMENS RDYNRRLWAEGWRAEVGKQLRPLYEEYV VLENE MARANNNYEDYGDYWRGDYEV TGAGDYDPSRDLQMLKDVERTFAEIKPLYEQLHAYVRAKL MHTPSYISPTGCLPAHLLGDMWGRFWTNLYSLTV PFEHKPSIDVTEKMEN QSWDAERIFKEAEKFV SISLPYMTQGFWDNSMLTEPGDGRKV/CHPTAWDLGKGDRIKMC TKVTMDDFLTAHH EHG HIQYD MAYAAQPYLLRNGANEGRHEA VGEIMS LSAATPHYLK ALGL LAPDFHEDN ETEINFL LKO QALTIV GTLPFTYMLEKWRM VFKGEIPKQ QWMEKWWEMKREIVGVVVEPLPHDETYCDPACLFHV AEDYSFIR YTT RIYQFQFHEALCKTAKHEGA LFKCDISN STEAGQRL LQMLRLGKSEP WTLALENI VGIKTMDV KPLN YFEPLFTWLKE QNRNS FVGW STEWP YSDQSI KV RISL KSALGENAYE WNDN EMYLFQSSVAYAMRKYFSEAR NETVLFGEDNVWVSDKKPRISFKFFF TSPNNV S DII PRTEVENAIRLSRDRINDV FQLDDNS LEFL G I QPTLGPPYEP PV TIWI LIIFGVMGVV VIGIVV LIFTGIRNRRKKNQASSEENPYGSV DLNKGENNSGFQ NIDDV QTS L
```

- Open SWISS-MODEL and paste the sequence to build model to get the homology model by comparing with the available structures in the database



This screenshot shows the "Start a New Modelling Project" page. It has a "Target Sequence(s)" input field where the ACE2 sequence from NCBI was pasted. Below it are fields for "Project Title" (Untitled Project) and "Email" (Optional). To the right, there is a "Supported Inputs" dropdown menu with options like "Sequence(s)", "Target-Template Alignment", "User Template", and "DeepView Project". At the bottom, there is a "Build Model" button and a note about agreeing to terms of use.



This screenshot shows the same "Start a New Modelling Project" page as above, but with a "Hetero Target" added. The "Target Sequence(s)" field now contains the ACE2 sequence and a second sequence: "Target: CTKVTMONFLTAHH EHG H1QYD MAYAAQPYLLRNGANEGRHEA VGEIMS LSAATPHYLK ALGL LAPDFHEDN ETEINFL LKO QALTIV GTLPFTYMLEKWRM VFKGEIPKQ QWMEKWWEMKREIVGVVVEPLPHDETYCDPACLFHV AEDYSFIR YTT RIYQFQFHEALCKTAKHEGA LFKCDISN STEAGQRL LQMLRLGKSEP WTLALENI VGIKTMDV KPLN YFEPLFTWLKE QNRNS FVGW STEWP YSDQSI KV RISL KSALGENAYE WNDN EMYLFQSSVAYAMRKYFSEAR NETVLFGEDNVWVSDKKPRISFKFFF TSPNNV S DII PRTEVENAIRLSRDRINDV FQLDDNS LEFL G I QPTLGPPYEP PV TIWI LIIFGVMGVV VIGIVV LIFTGIRNRRKKNQASSEENPYGSV DLNKGENNSGFQ NIDDV QTS L". The "Add Hetero Target" button is visible. The rest of the page is identical to the previous screenshot.

- The model built is ranked base on GMQE (Global Model Quality Estimation), which is expressed as a number between 0 and 1 with 1 being the highest accuracy and 0 the lowest.
- QMEANDisCO compares interatomic distances in the model with ensemble information extracted from experimentally determined protein structures of target sequence homologues. The score shows similarity of the residues to the experimental structure and if it drops below 0.6, modelled residues are in general of low quality.

Model Results  Order by: GMQE 

Rank	Score
1	804

Model 02



Ligands
1 x ZN 

QMEANDisCo Local

QMEAN Z-Scores

Template
7e3j.1.A ACE2
Crystal structure of SARS-CoV-2 RBD binding to dog ACE2

Structure Assessment

Compare

Download files 

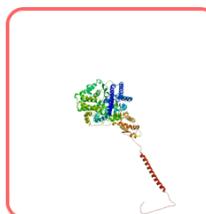
QMEANDisCo Global:
0.81 ± 0.05

Seq Identity
82.94%

Coverage 

Model-Template Alignment 

Model 01



Oligo-State
Monomer

GMQE
0.91

We predict the structure to have a transmembrane segment.

Show / Hide

Structure Assessment

Compare

Download files 

Template 

- The Model 02 is considered looking at GMQE of 0.77 and QMEANDisCo of 0.81 - the pdb of the model is downloaded

Step 2. Validating the model

- For validating the model feed the pdb @ <https://saves.mbi.ucla.edu/>

UCLA-DOE LAB – SAVES v6.0

**To run any or all programs:
upload your structure, in PDB format only**

Choose file

Customize job name:

model_02.pdb

Run programs

- The following results are obtained

UCLA-DOE LAB – SAVES v6.0

UCLA

Job 1688238 has been created

[New Job](#)

job #1688238: model_02.pdb [job link] [3D Viewer]

ERRAT Analyzes the statistics of non-bonded interactions between different atom types and plots the value of the error function versus position of a 9-residue sliding window, calculated by a comparison with statistics from highly refined structures. Start	Verify3D Determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar etc) and comparing the results to good structures. Start	PROVE <small>Temporarily down at the moment</small>
WHATCHECK <small>Derived from a subset of protein verification tools from the WHATIF program (Vriend, 1990), this does extensive checking of many stereochemical parameters of the residues in the model.</small> Start	PROCHECK <small>Checks the stereochemical quality of a protein structure by analyzing residue-by-residue geometry and overall structure geometry.</small> Start	OPEN <small>We are open to suggestions for a 6th program to operate in this window. If you know of a program that we could run locally on our server that would be most useful, please let us know: email holton at mbi dot ucla dot edu with your suggestion</small>

UCLA-DOE LAB — SAVES v6.0

UCLA

Job 1688238 has been created

New Job

job #1688238: model_02.pdb [job link] [3D Viewer]

ERRAT Complete

Overall Quality Factor

98.4012

[Results](#)

VERIFY Complete

89.25% of the residues have averaged 3D-1D score ≥ 0.1

Pass

At least 80% of the amino acids have scored ≥ 0.1 in the 3D/1D profile.

[Results](#)

PROVE

Temporarily down at the moment

WHATCHECK Complete



[Results](#)

PROCHECK Complete

Out of 9 evaluations

- Errors: 1
- Warnings: 6
- Pass: 2

[Results](#)

OPEN

We are open to suggestions for a 6th program to operate in this window. If you know of a program that we could run locally on our server that would be most useful, please let us know: email holton at mbi dot ucla dot edu with your suggestion

Overall Quality Factor

98.4012

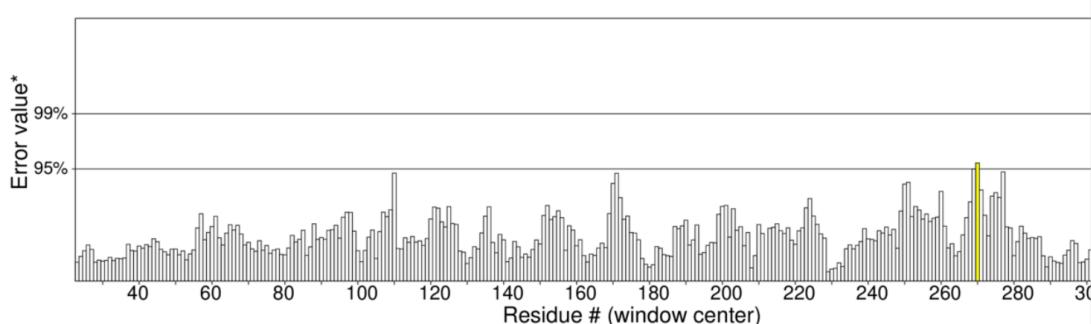
[Log](#) [PostScript](#) [PDF](#)

Chain A

Chain A

Chain A

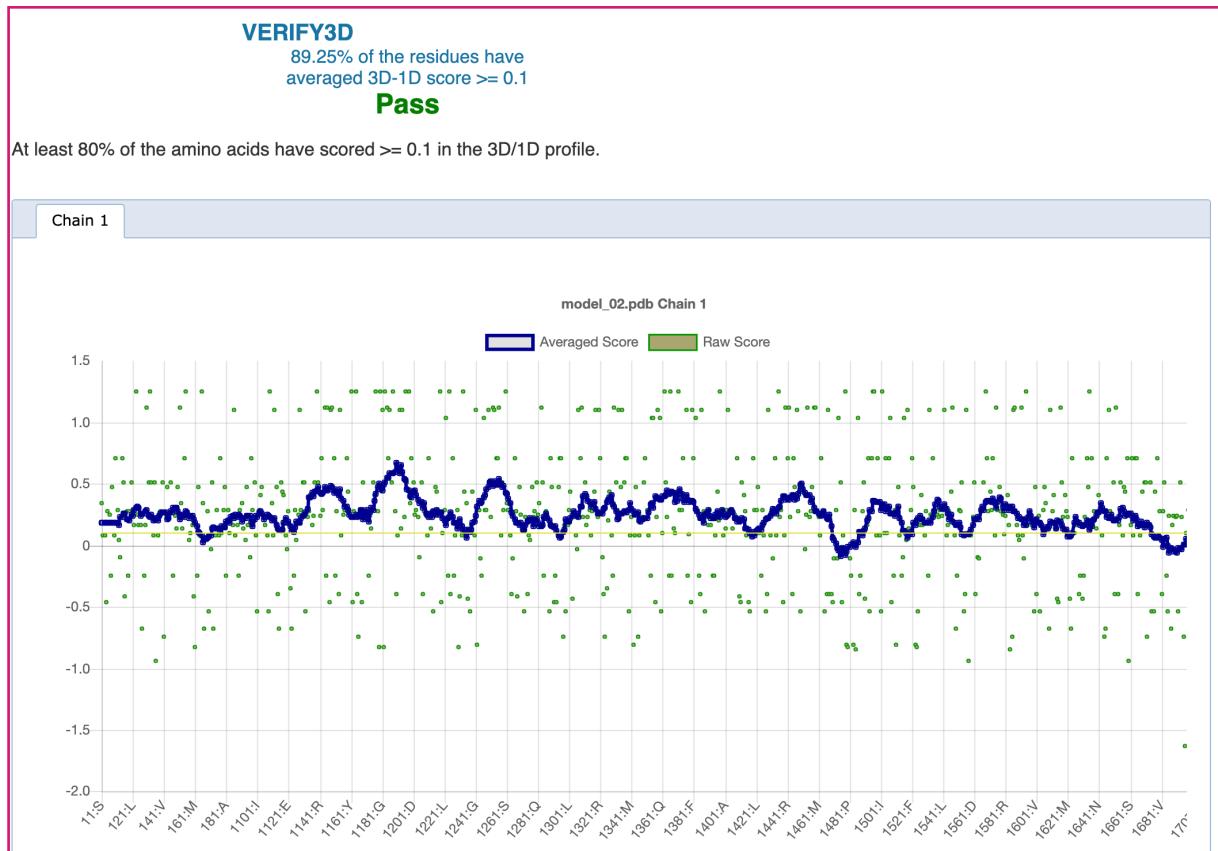
Program: ERRAT2
File: model_02.pdb
Chain#:A
Overall quality factor**: 98.401



*On the error axis, two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value.

**Expressed as the percentage of the protein for which the calculated error value falls below the 95% rejection limit. Good high resolution structures generally produce values around 95% or higher. For lower resolutions (2.5 to 3 Å) the average overall quality factor is around 91%.

- ERRAT is a program for verifying protein structures determined by crystallography. Error values are plotted as a function of the position of a sliding 9-residue window. The error function is based on the statistics of non-bonded atom-atom interactions in the reported structure (compared to a database of reliable high-resolution structures).



UCLA-DOE LAB — SAVES v6.0

[← Control panel](#) [↑ New Job](#) UCLA

job #1691070: model_02.pdb

ERRAT **VERIFY3D** **WHATCHECK** **PROCHECK**

WHATCHECK

Headings

1	2	3	4	5
6	7	8	9	10
11	12	13	14	
15	16	17	18	
19	20	21	22	
23	24	25	26	
27	28	29	30	
31	32	33	34	
35	36	37	38	
39	40	41	42	
43	44	45	46	
47				

#47. Note: Overall summary report

Note: Overall summary report
This is an attempt to create an overall summary of the quality of the structure. We do not recommend anyone to look at these numbers, please look at the complete report instead.

Structure Z-scores, positive is better than average:
1st generation packing quality : -0.897
Ramachandran plot appearance : 0.398
chi-1/chi-2 rotamer normality : -0.861
Backbone conformation : -33.473 (bad)

RMS Z-scores, should be close to 1.0:
Bond lengths : 0.659 (tight)
Bond angles : 0.968 (tight)
Omega angle restraints : 1.089
Side chain planarity : 1.134

REFERENCES
=====

PROCHECK

Out of 9 evaluations

- Errors: 1
 - Warning: 6
 - Pass: 2

The evaluations are the '+' (Warning) and '*' (Error) in the summary. The categories on the left do not always correspond in number due to PROCHECK output documents.

Summary	
Ramachandran plot	Warning
All Ramachandrans	Warning
Chi1-chi2 plots	Pass
Main-chain params	
Side-chain params	Warning
Residue properties	Warning
Bond len/angle	Warning
M/c bond lengths	
M/c bond angles	
Planar groups	Warning
Program output	

```
+-----<<< P R O C H E C K   S U M M A R Y >>>-----+
| /var/www/SAVES/Jobs/1688238/saves.pdb    1.5          708 residues |
* Ramachandran plot:  92.0% core      7.6% allow      0.2% gener      0.3% disall
+ All Ramachandrans:  9 labelled residues (out of 708)
+ Chi1-chi2 plots:   2 labelled residues (out of 475)
  Side-chain params: 5 better       0 inside       0 worse
+ Residue properties: Max.deviation:      6.6          Bad contacts:     0
+                               Bond len/angle:   4.8          Morris et al class: 1 1 2
+                               1 cis-peptides
+ G-factors           Dihedrals:   -0.07        Covalent:     0.06        Overall:  -0.01
+ Planar groups:      90.9% within limits  9.1% highlighted  7 off graph
+
+ May be worth investigating further. * Worth investigating further.
```

Summary file

Step 3. Selecting the domain of Spike (COVID-19 virus) that interacts with ACE2

- Download the desired structure from pdb and then select the domain that interacts with ACE2 in chimera

Structure Summary

Biological Assembly 1

6VW1

Structure of SARS-CoV-2 chimeric receptor-biotin human ACE2

PDB DOI: <https://doi.org/10.2210/pdb6VW1/pdb>

Classification: **CELL INVASION**

Organism(s): **Homo sapiens, Severe acute respiratory syndrome coronavirus 2**

Expression System: **Spodoptera frugiperda**

Mutation(s): **No**

Membrane Protein: **Yes**

Deposited: 2020-02-18 Released: 2020-03-04

Author(s): Shang, J., Ye, G., Shi, K., Wan, Y.S., Ali

Funding Organization(s): National Institutes of Health/Nation

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 2.68 Å

R-Value Free: 0.229

R-Value Work: 0.197

R-Value Observed: 0.199

Global Symmetry: Asymmetric - C1

Global Stoichiometry: Hetero 2-mer - A1B1

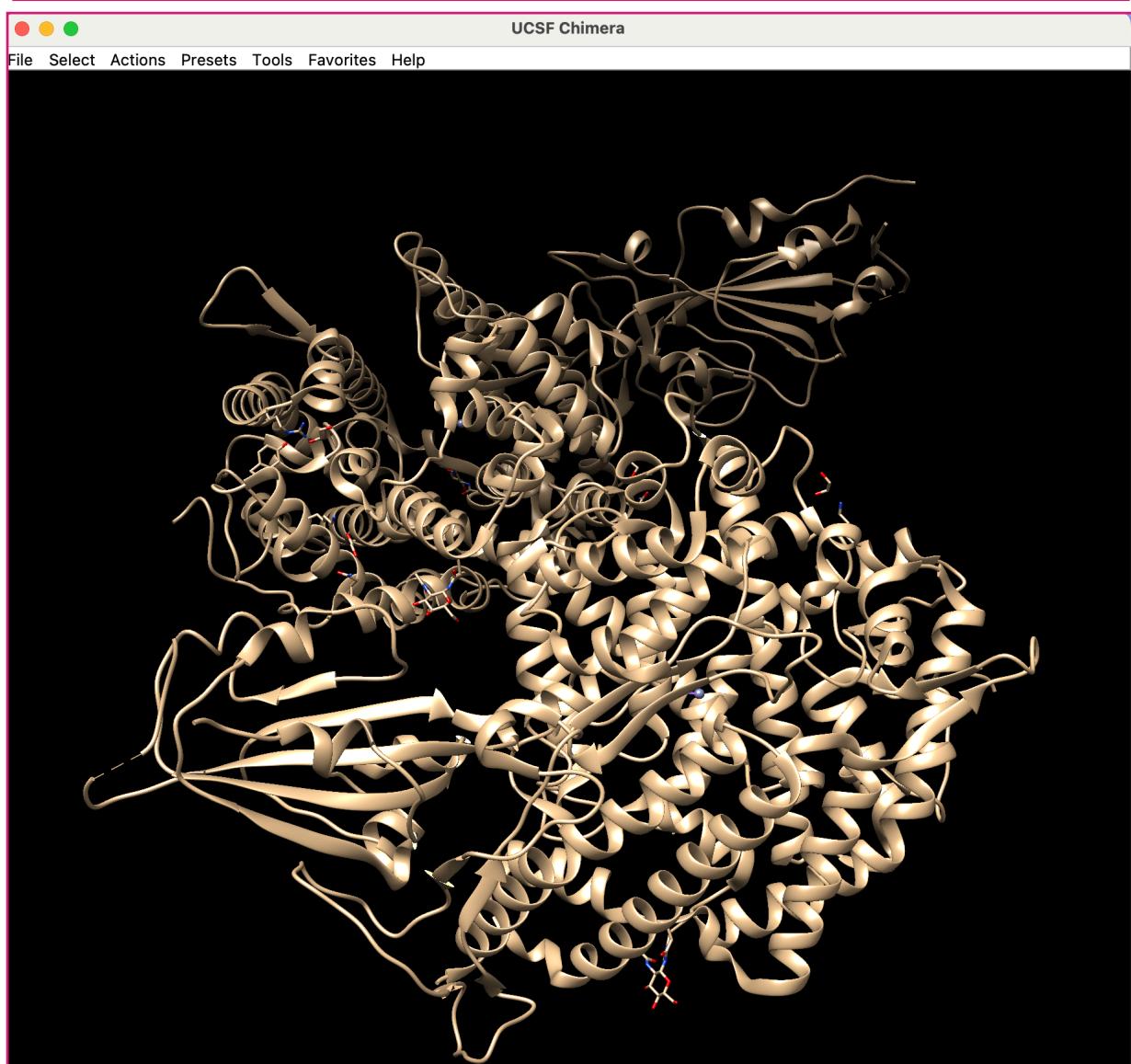
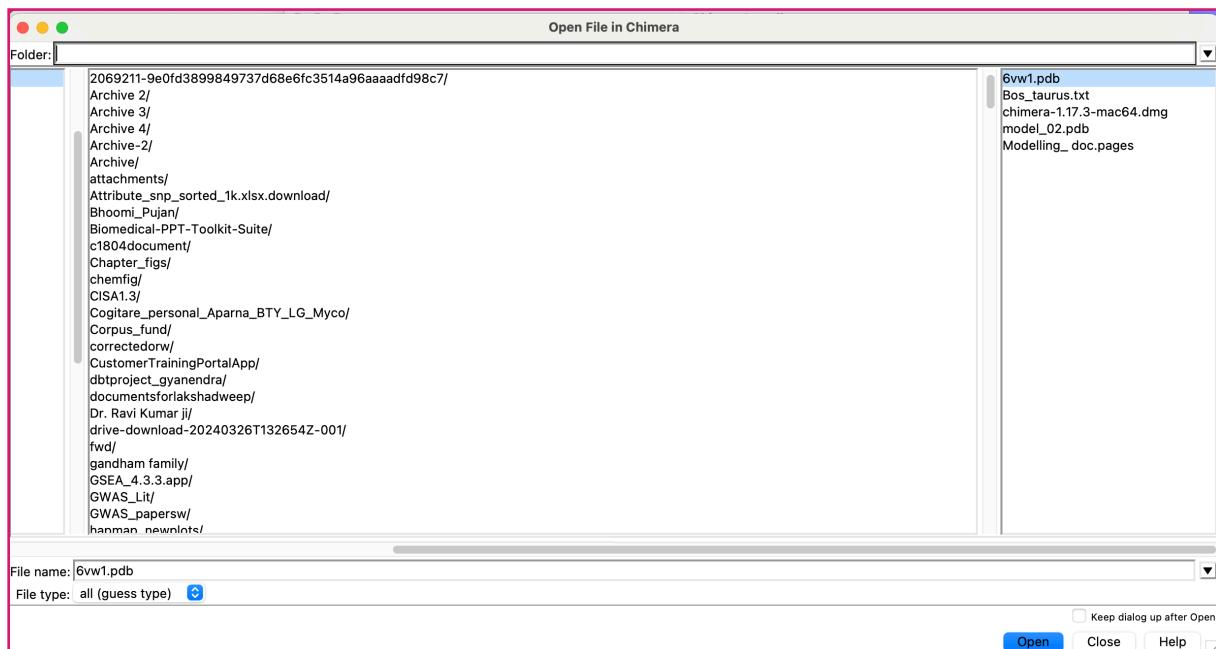
Explore in 3D: Structure | Sequence Annotations | Electron Density | Validation Report | Ligand Interaction (NAG) | Predict Membrane

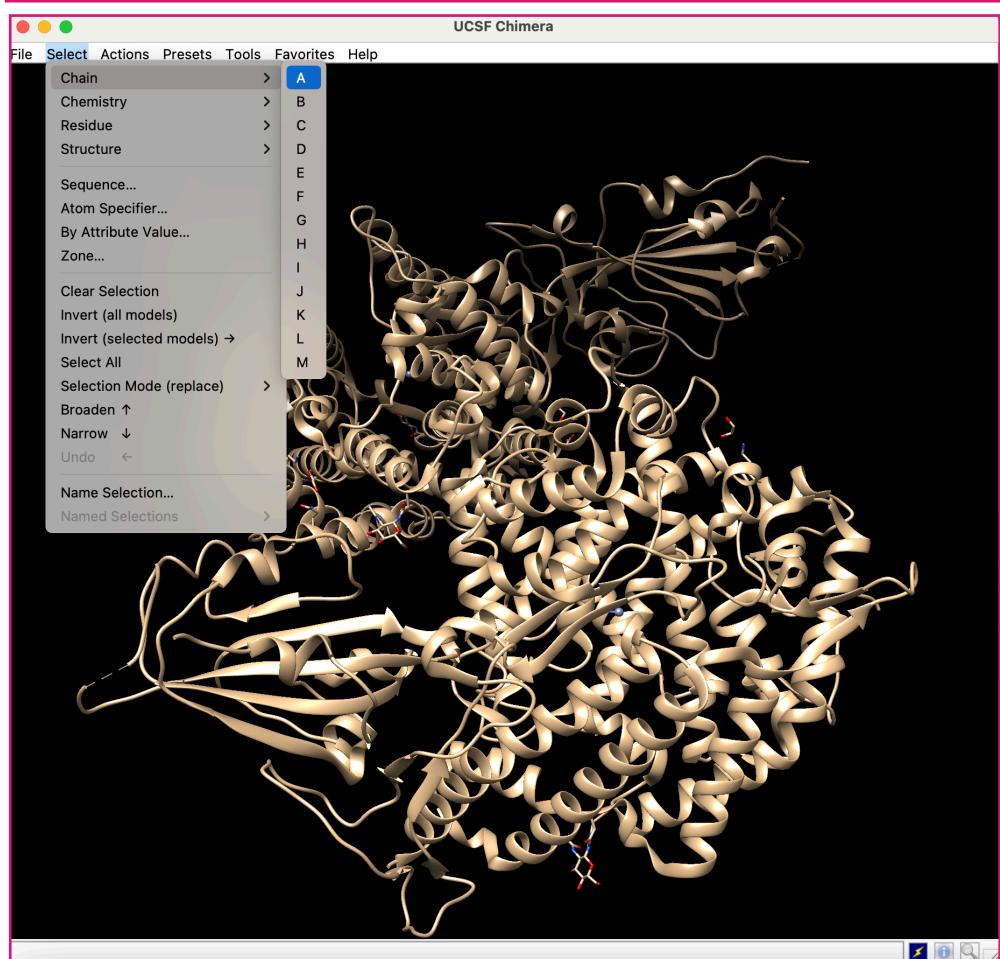
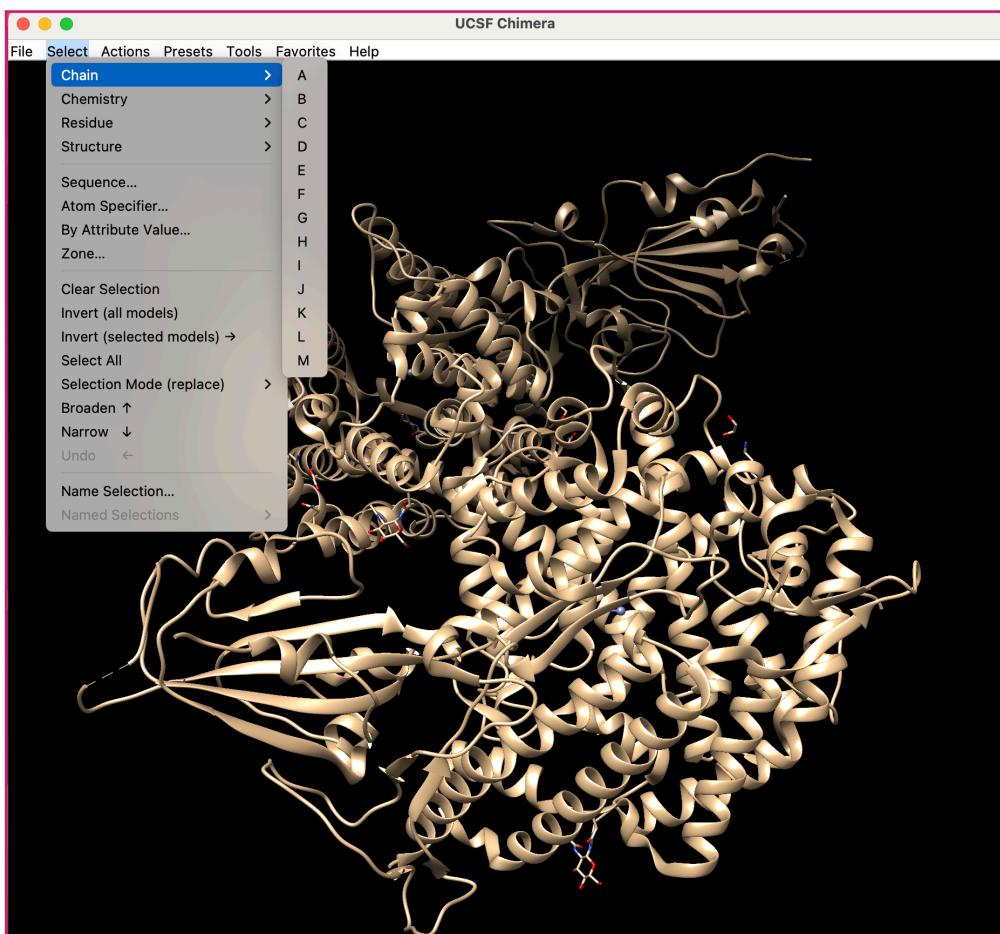
Global Assembly: Biological Assembly 1 (CIF - gz) | Biological Assembly 2 (CIF - gz)

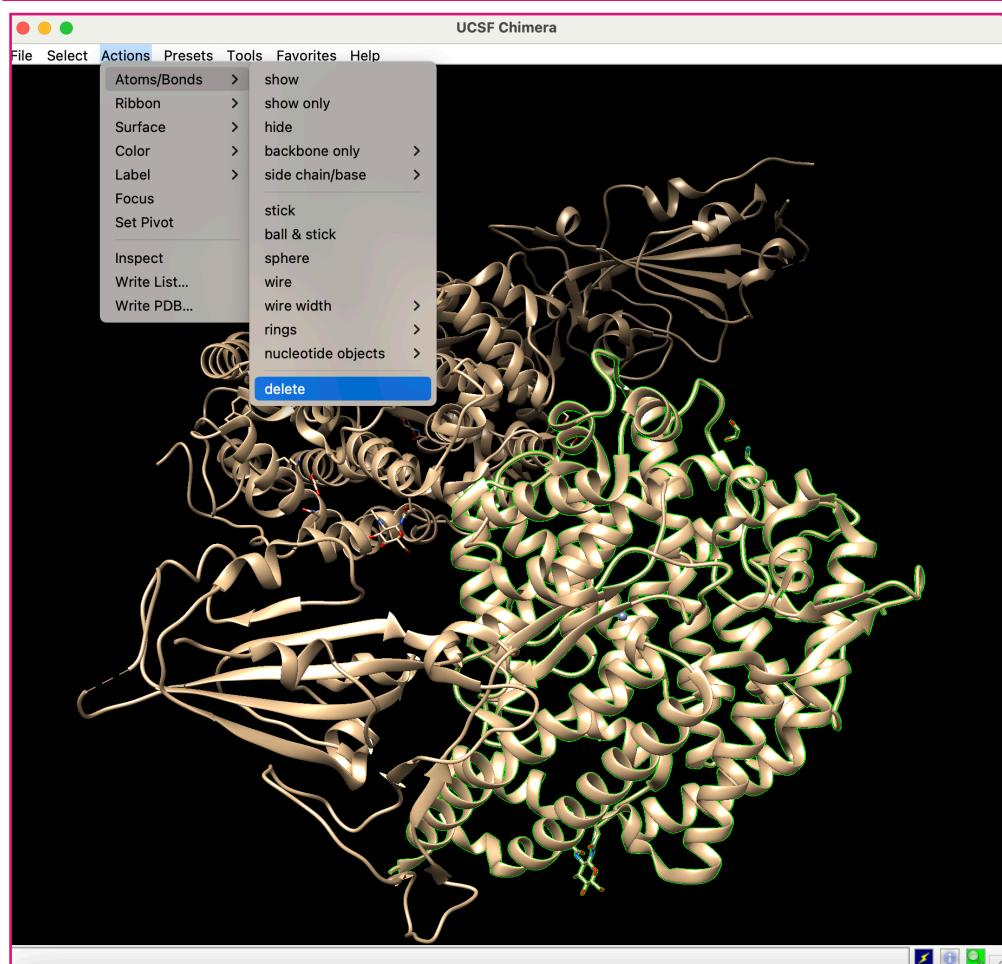
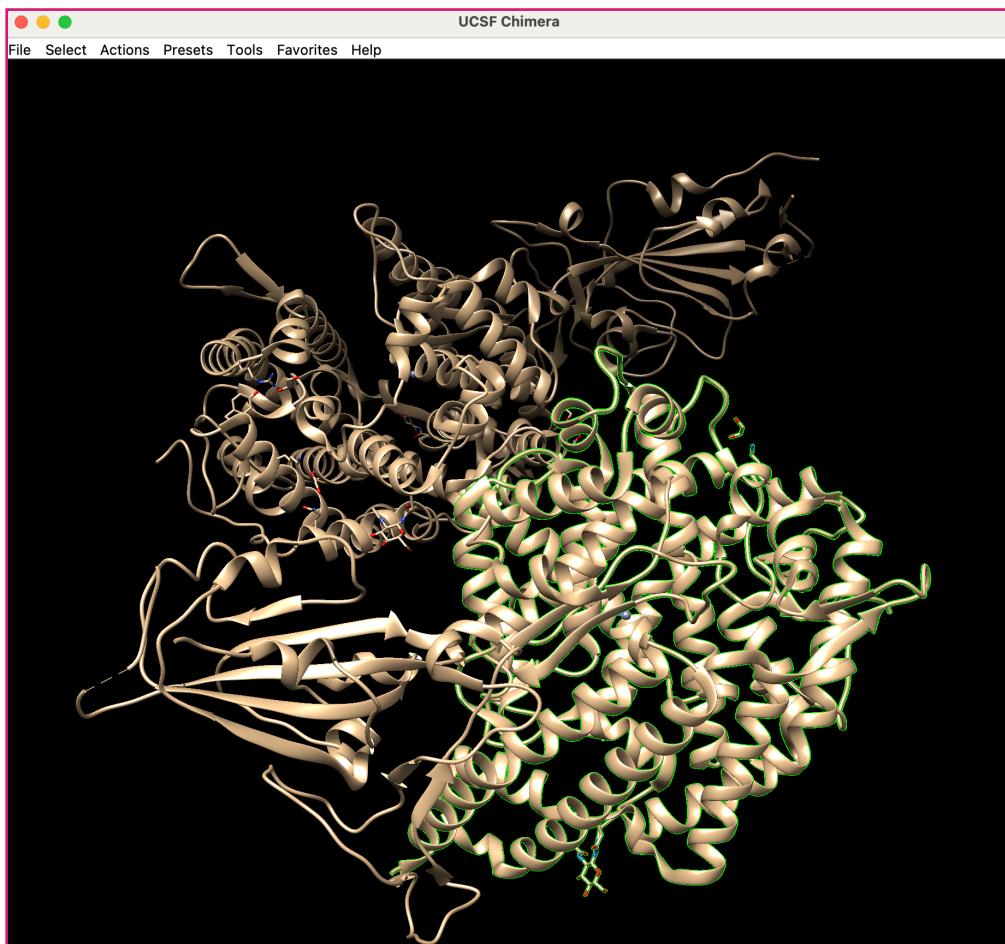
Validation: Validation Full PDF | Validation (XML - gz) | Validation (CIF - gz)

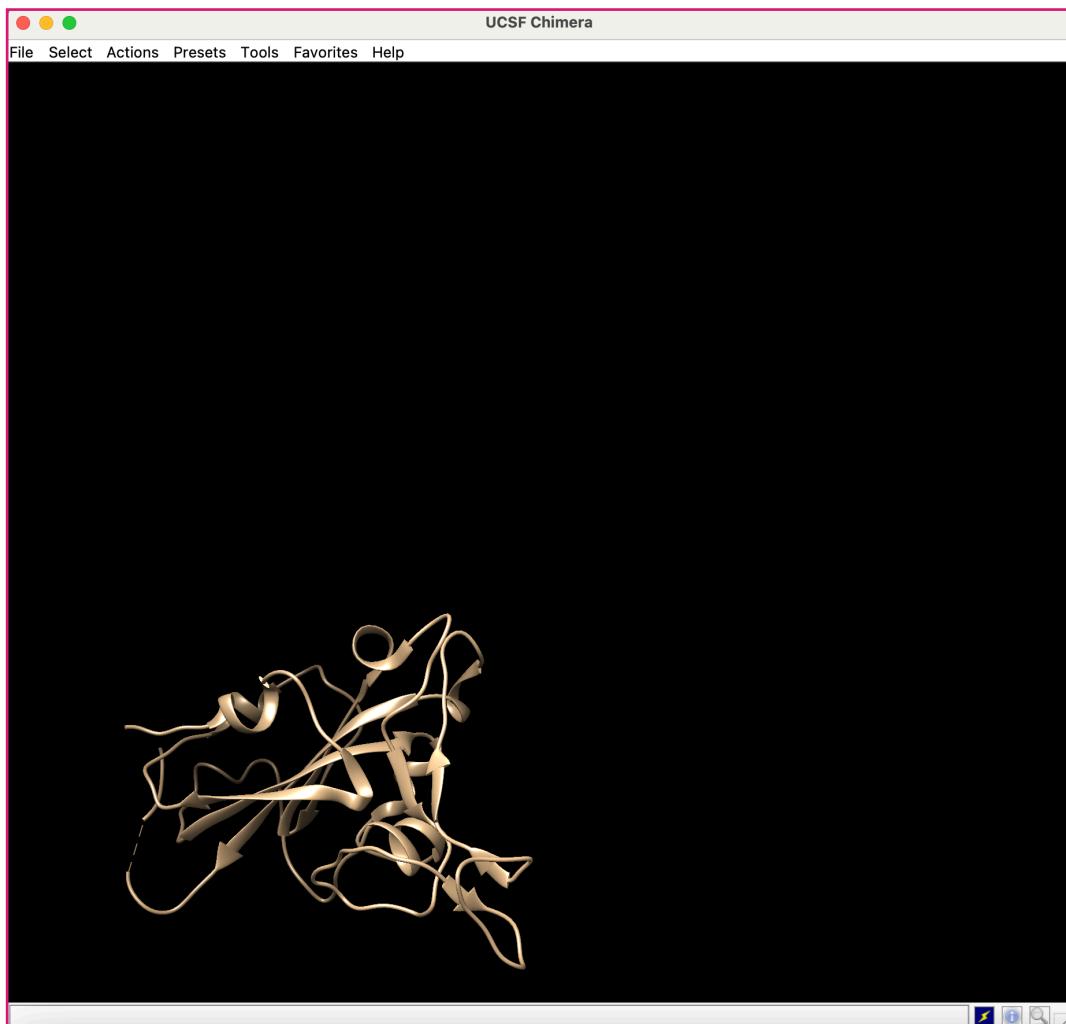
Format: Display Files | Download Files | Data API

- In Chimera open 6VW1.pdb , select the domains and delete all domains except the E domain as shown below









- Save the domain as 6VW1_E.pdb - load this along with the model2.pdb for Bos taurus in GRAMM web - <https://gramm.compbio.ku.edu/process-request>

Step 4. Interaction model between 6VW1_E.pdb and model2.pdb

GRAMM: Docking Web Server

This is the Web interface to our protein docking software **GRAMM** (Global RAnge Molecular Matching). GRAMM systematically maps the intermolecular energy landscape by predicting a spectrum of docking poses corresponding to stable (deep energy minima) and transient (shallow minima) protein interactions. More information on the GRAMM methodology is on our [laboratory website](#).

[Start GRAMM docking](#)

Conditions of Use
This computationally intensive service is provided to the research community free of charge utilizing our computational resources, which are limited.

To ensure the fair use of this server, we ask that no person submits more than ten simulation requests per day. We reserve the right to block access from those not honoring this request or involved in any other misuse of this service.

The simulation requests will be queued and the wait time will vary depending on the load of our cluster.

Disclaimer:
This is research software under active development, and we make no guarantees.

References
GRAMM web-server should be cited as:
[Singh, A., Copeland, M.M., Kundrotas, P.J., Vakser, I.A., 2024, GRAMM Web Server for Protein Docking, Methods in Mol. Biol., 2714:101-112.](#)
Other GRAMM docking references:
[Singh A., Dauzhenka T., Kundrotas P.J., Sternberg M.J.E., Vakser I.A., 2020, Application of docking methodologies to modeled proteins, Proteins, 88:1180–88.](#)
[Tsvchigrechko, A., Vakser, I.A., 2006, GRAMM-X public web server for protein-protein docking, Nucleic Acids Research, 34:W310-W314.](#)

[Home](#) [Start GRAMM docking](#) [Vakser Lab](#)

Start GRAMM docking

We are looking for Postdocs and PhD students, preferably with physics/math background, to work on modeling of protein interactions and whole cell modeling. Letters of interest and CVs can be sent to vakser@ku.edu.

Email Address:

PDB file of the receptor: model_02.pdb

PDB file of the ligand: 6VW1_E.pdb

Docking methodology:

Number of top matches to output as PDB files:

► Advanced options (optional):

Center for Computational Biology, The University of Kansas, USA © Vakser Lab - 2024

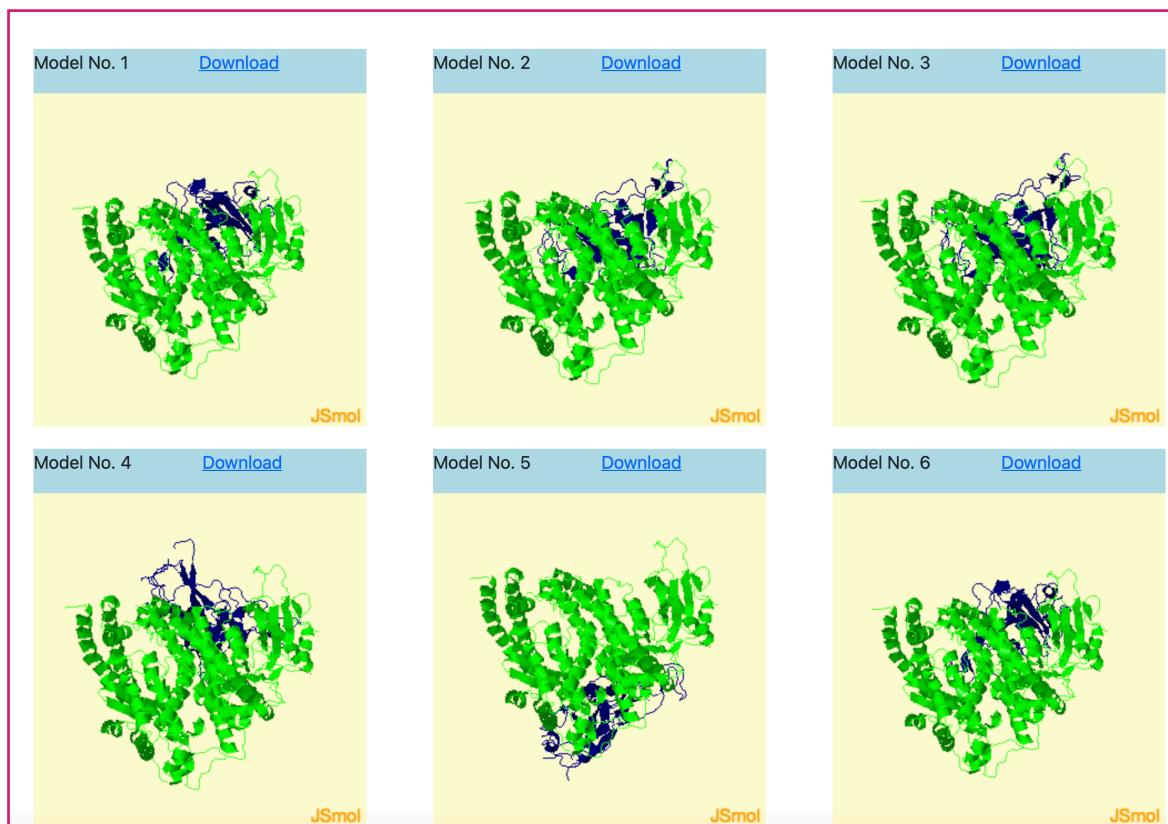
[Home](#) [Start GRAMM docking](#) [Vakser Lab](#)

Request Received

The submitted Job id is: 49785. An email will be sent to gandham71@gmail.com once the job is processed

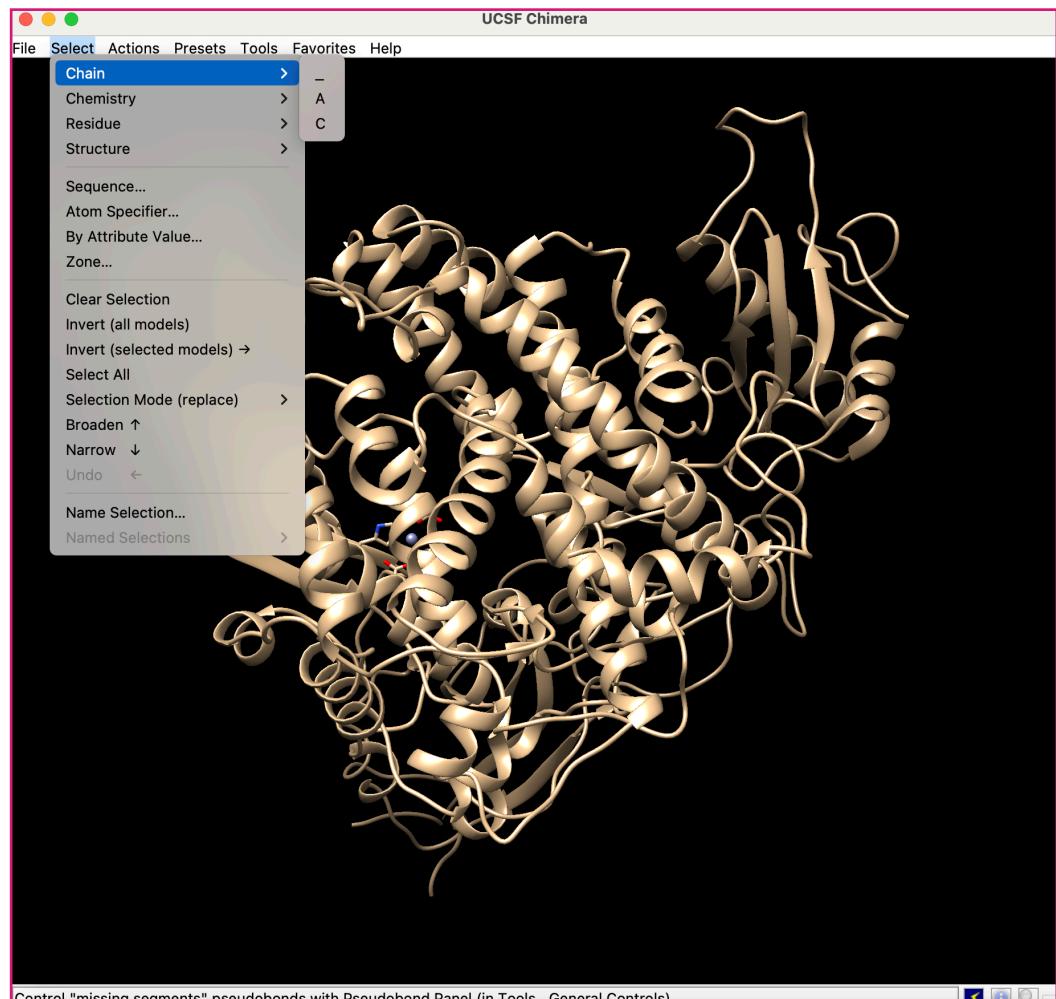
Please [click here](#) to check the the job status.

Center for Computational Biology, The University of Kansas, USA © Vakser Lab - 2024



- The top 10 models given by GRAMM are taken into PRODIGY to find the most stable model

Step 5.Taking the models into PRODIGY - to get the best model



PRODIGY [Home](#) [Manual](#) [Method](#) [Dataset](#) [Example](#) [Reference](#) [Login](#)

If you would like to try it out, please click [click here.](#)

PRODIGY (protein-protein) **PRODIGY-LIGAND** (protein-small molecule) **PRODIGY-CRYSTAL** (biological or crystallographic)

Please provide the PDB ID of the target complex or submit a file in [PDB](#) or [mmcCIF](#) format.

An archive of multiple PDB/mmcCIF files can also be provided (as a multi-model PDB), together with the interacting chains.

Information about the predictive approach can be found at the online [method page](#).

Structure(s)* OR receptor..I10.pdb

Interactor 1* A

Interactor 2* C

Temperature (in °C)* 25.0

Job ID Custom tag - Optional
Add a custom Job ID to identify your run.

Email gandham71@gmail.com
Optional email address to which the results will be send.

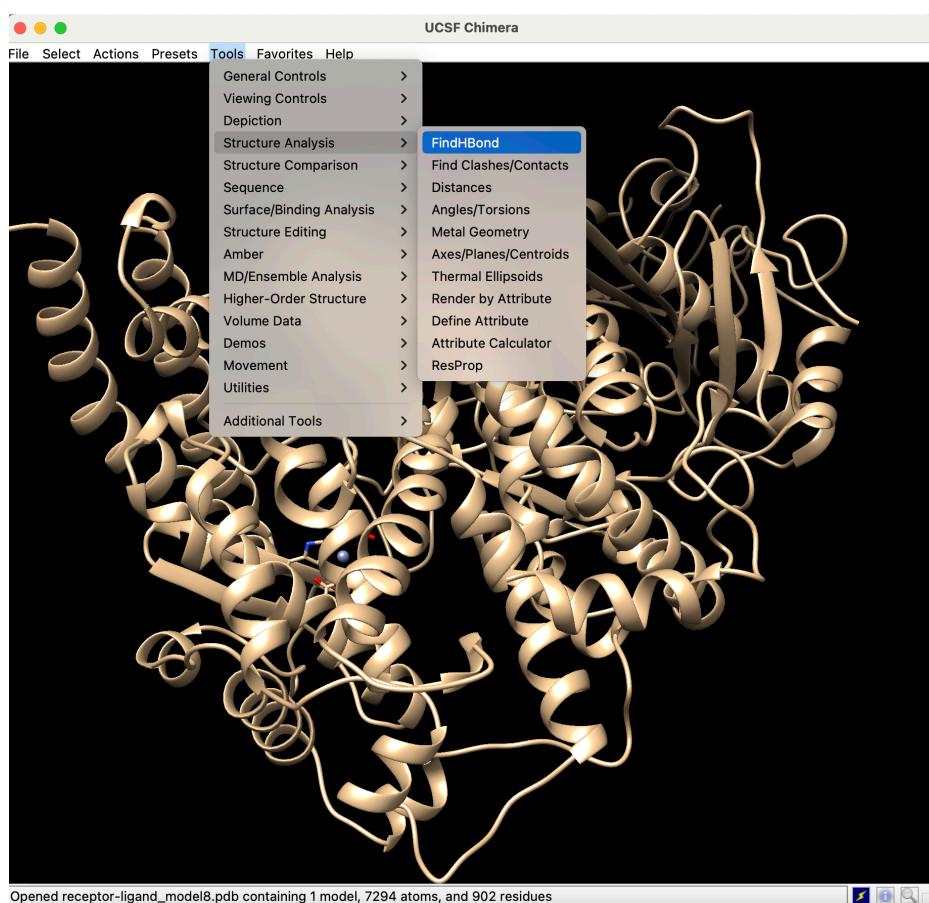
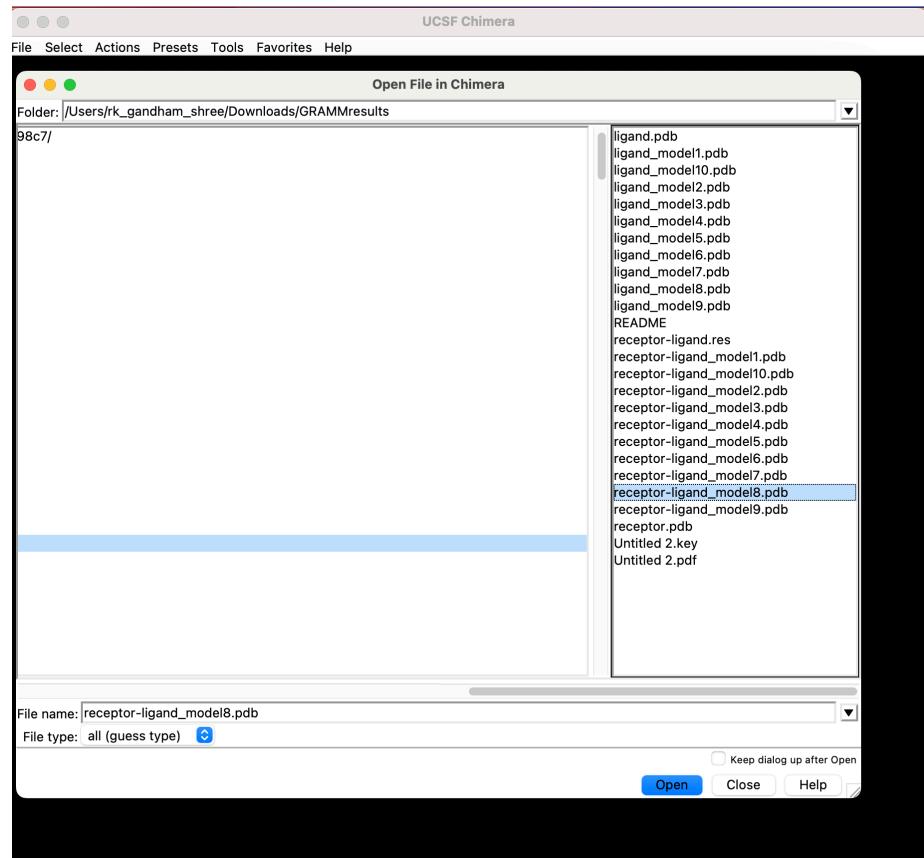
I'm not a robot 
reCAPTCHA
Privacy - Terms

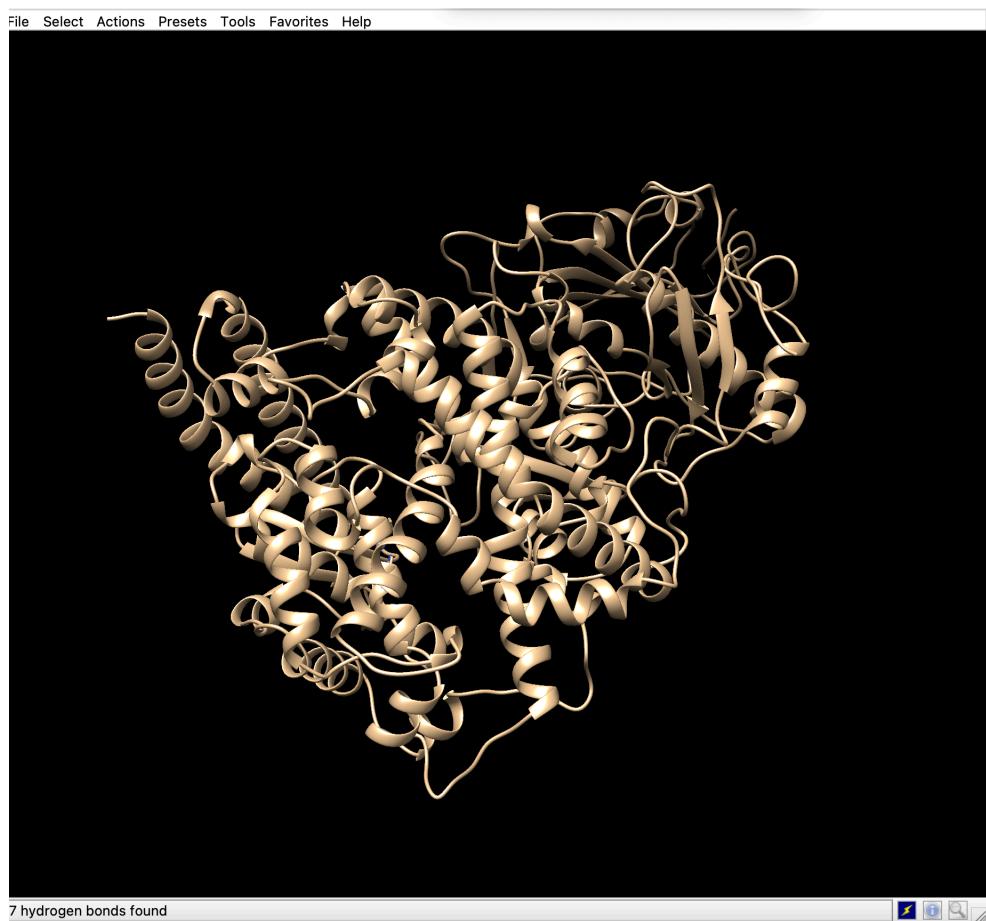
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model1_1	-16.9	4.3e-13	12	22	37	8	34	30	26.54	34.26	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model3_3	-19.2	7.7e-15	12	19	28	6	47	37	26.78	34.37	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model2_2	-18.2	4.8e-14	14	20	26	15	50	33	26.96	34.17	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model4_4	-19.1	9.7e-15	17	31	53	27	51	32	26.58	34.02	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model5_5	-17.9	8e-14	12	29	44	31	56	33	26.69	35.16	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model6_6	-19.1	9.3e-15	13	27	34	8	45	33	26.67	34.26	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model7_7	-20.5	9.2e-16	20	32	45	15	50	40	26.97	34.86	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model8_8	-21.7	1.1e-16	16	31	56	13	50	48	26.45	34.59	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model9_9	-20.3	1.3e-15	12	28	43	16	54	25	26.74	34.78	
BINDING AFFINITY AND K_D PREDICTION											
Protein-protein complex	ΔG (kcal mol ⁻¹)	K_d (M) at °C	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged	NIS apolar	
receptor_ligand_model9_9	-20.3	1.3e-15	12	28	43	16	54	25	26.74	34.78	

Model 8 is considered to be the best and will further be evaluated for RMSD, hydrogen bond using Chimera and other parameters using FoldX

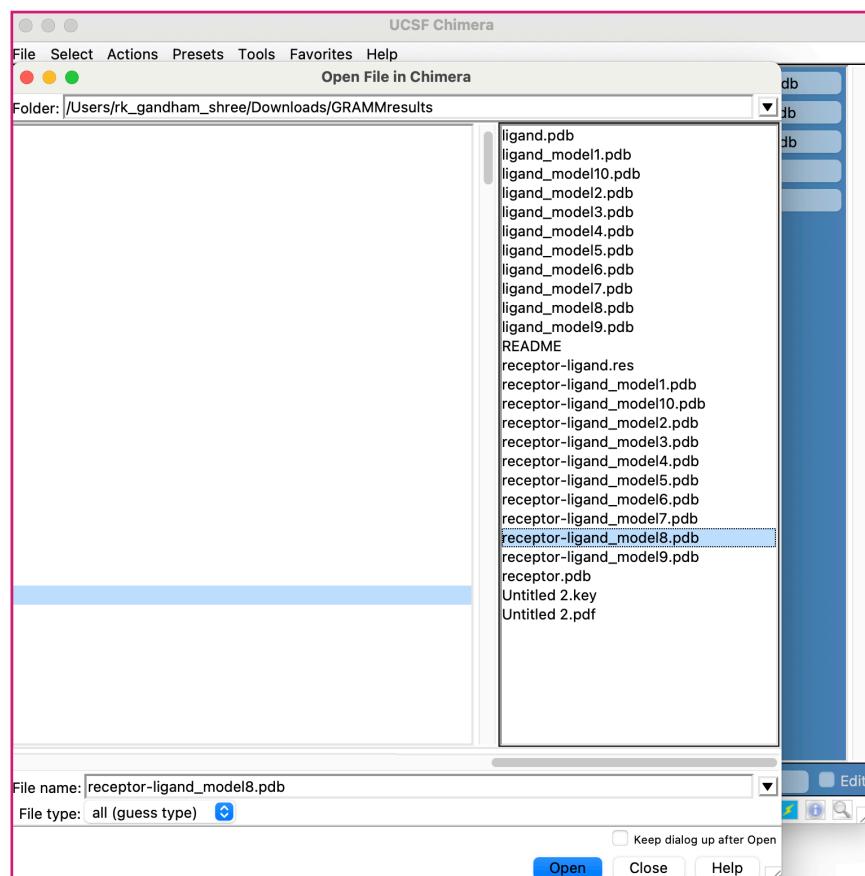
Step 5. Evaluating the model using Chimera

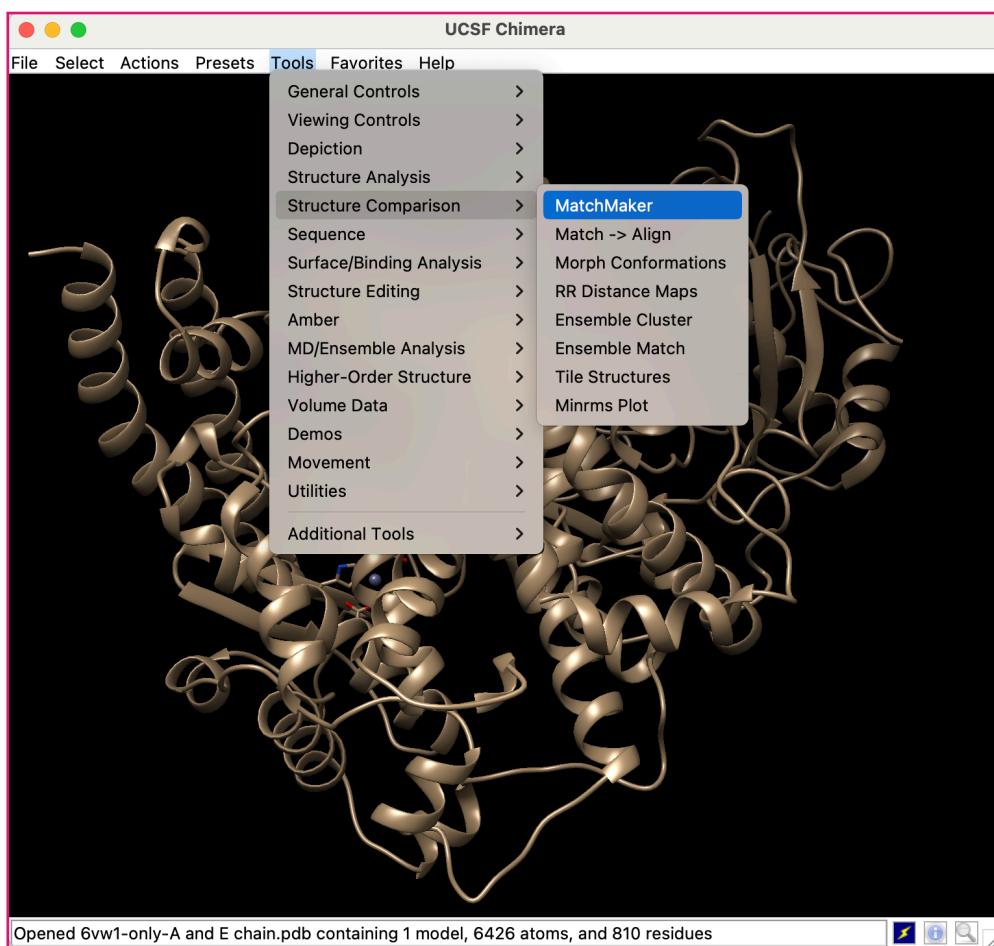
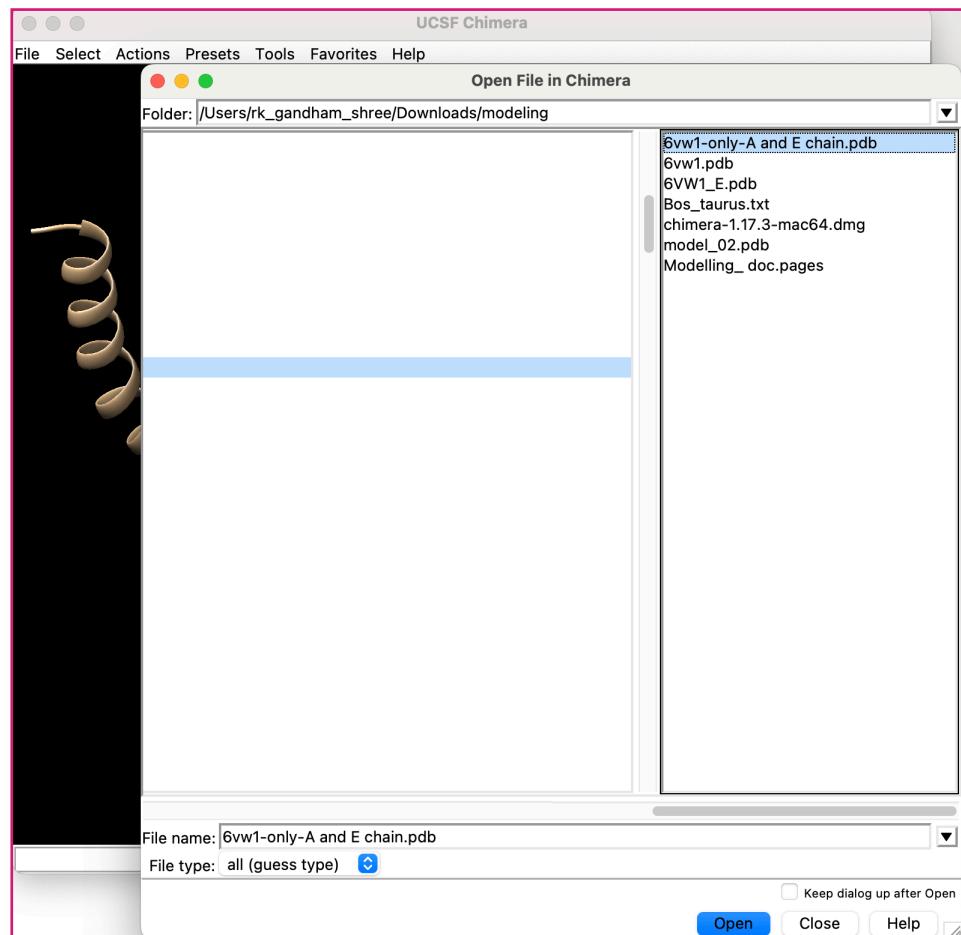
- Finding the number of intermolecular Hydrogen bonds



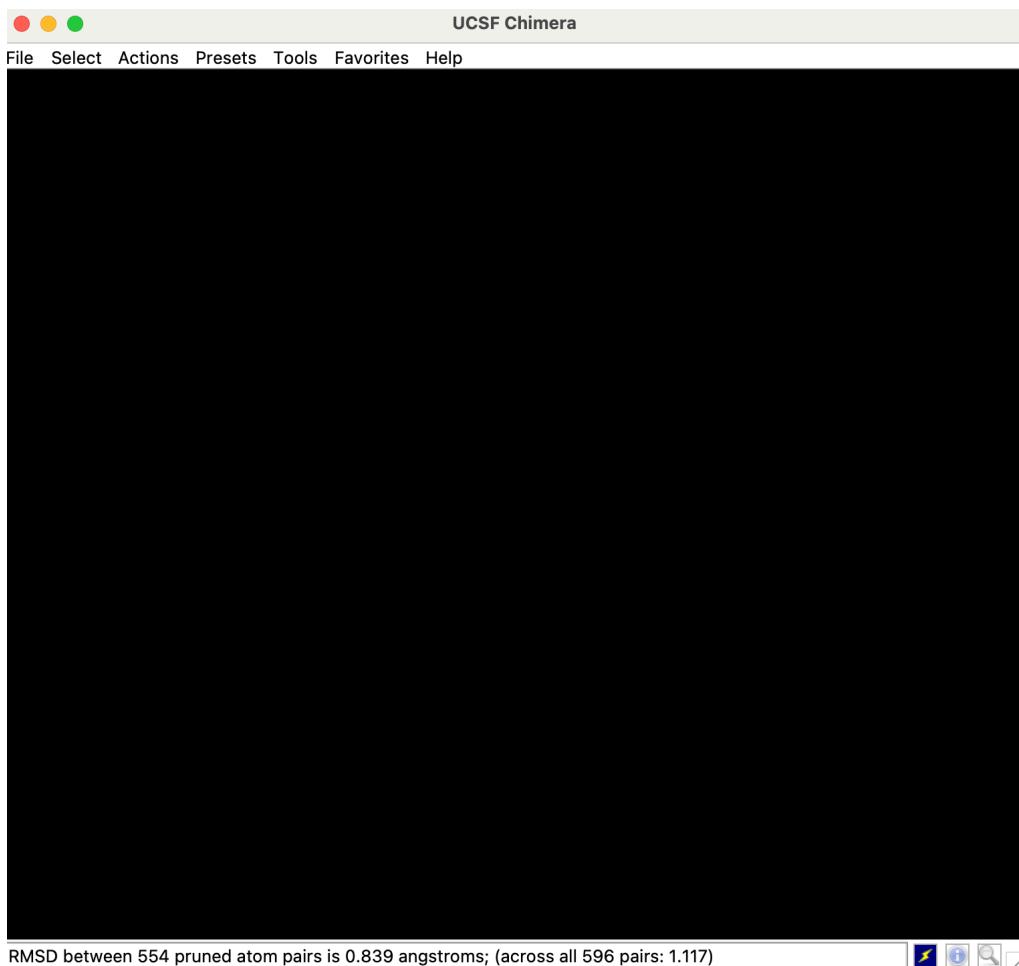
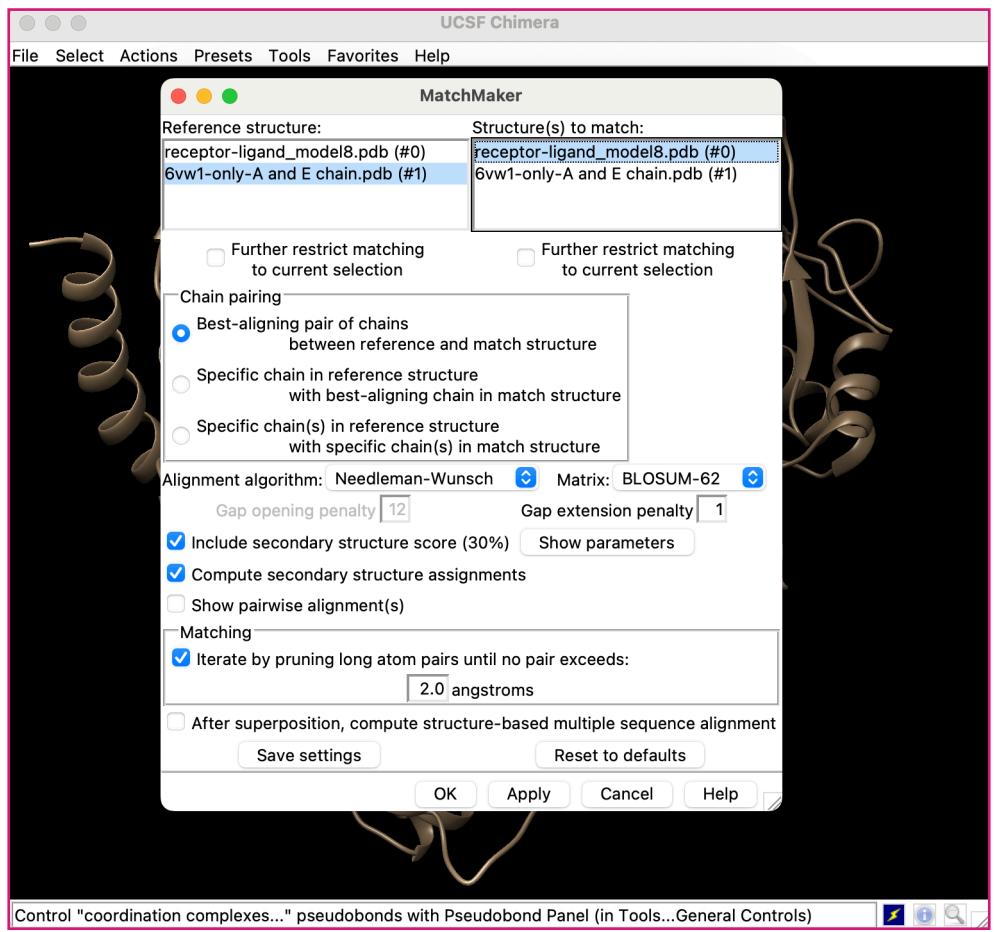


- Calculating RMSD - call in both the best model (receptor_ligand_model8.pdb) and proven Crystallography model (6vw1-only-A and E chain.pdb)





Opened 6vw1-only-A and E chain.pdb containing 1 model, 6426 atoms, and 810 residues



RMSD between 554 pruned atom pairs is 0.839 angstroms

Step 5. FoldX

```
foldx_20241231 -c AnalyseComplex --pdb=receptor-ligand_model8.pdb
```

interaction between A and C

BackHbond	=	-1.88
SideHbond	=	-3.47
Energy_VdW	=	-37.25
Electro	=	-2.20
Energy_SolvP	=	73.56
Energy_SolvH	=	-45.26
Energy_vdwclash	=	312.38
energy_torsion	=	1.08
backbone_vdwclash=		15.47
Entropy_sidec	=	17.73
Entropy_mainc	=	9.07
water bonds	=	0.00
helix dipole	=	-0.56
loop_entropy	=	0.00
cis_bond	=	0.00
disulfide	=	-0.00
kn electrostatic=		0.17
partial covalent interactions	=	0.00
Energy_Ionisation	=	0.00
Entropy Complex	=	2.38

Total	=	323.37
-------	---	--------

Complex Analysis of ./receptor-ligand_model8.pdb went fine

Your file run OK

End time of FoldX: Tue Jun 4 16:17:05 2024

Total time spend: 2.89 seconds.