Homology Modeling and GPCR Drug Discovery with BioHPC

Peng Lian
Computer Scientist, BioHPC
April 27, 2022

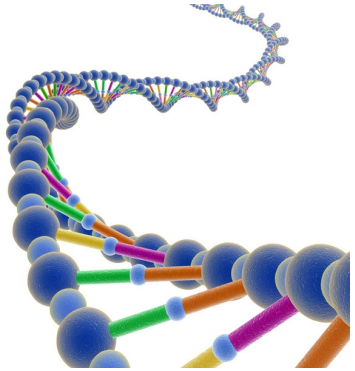
Outline

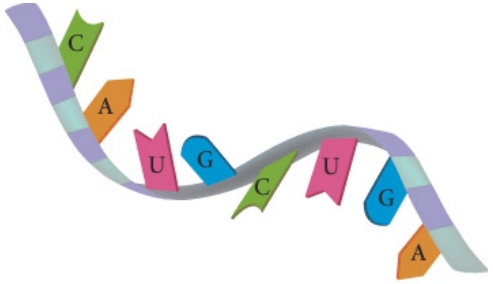
- Background on Biology, sequences, and structures

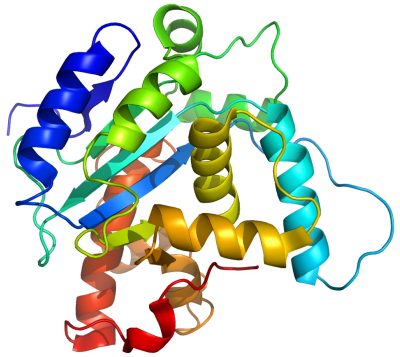
- Homology modeling
- Molecular docking

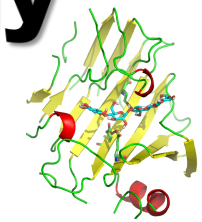
- GPCR as drug targets
- Hands-on Homology modeling of 5-HT_{1A} receptor
- Hands-on Molecular docking with Autodock on BioHPC
- Discussion on Virtual Screening

The Central Dogma of Biology

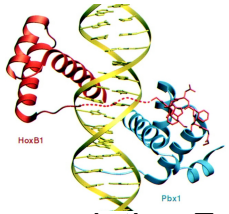




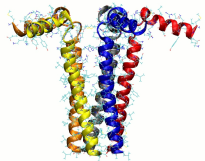




Enzymes



Transcription Factors



Signaling & Binding

GPCR receptors

Kinases

Structural Proteins

……

Functions

DNA

Transcription

Reverse Transcription

RNA

Translation

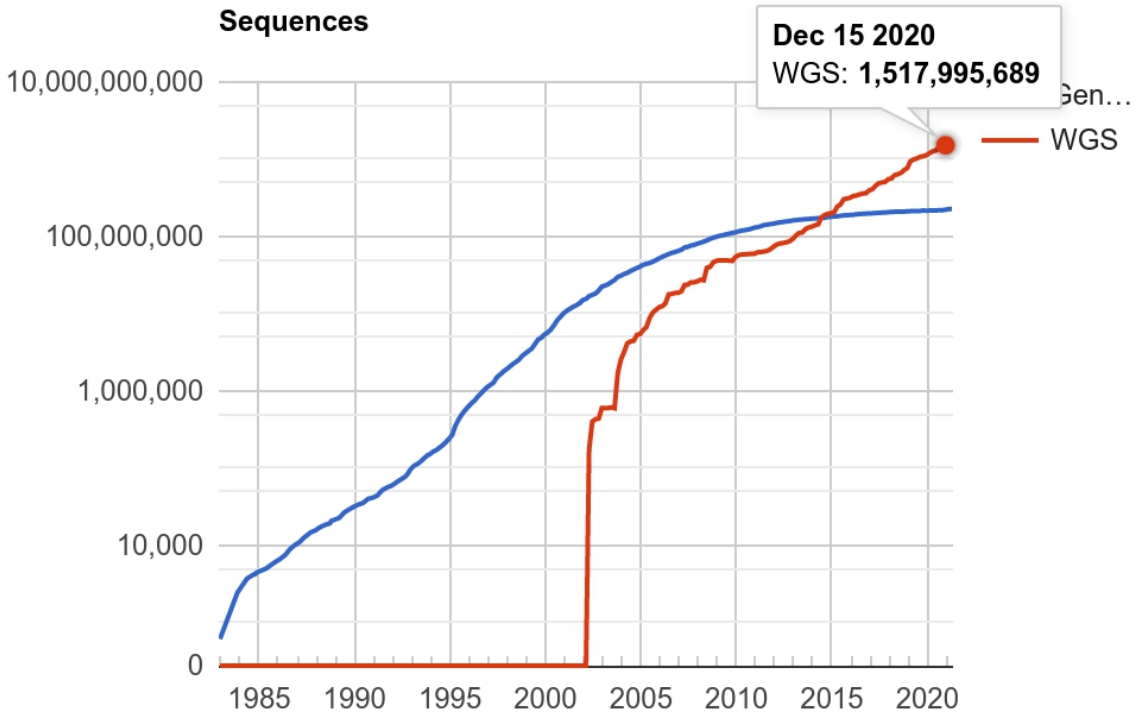
Protein

Replication

Replication

From Sequence to Structure

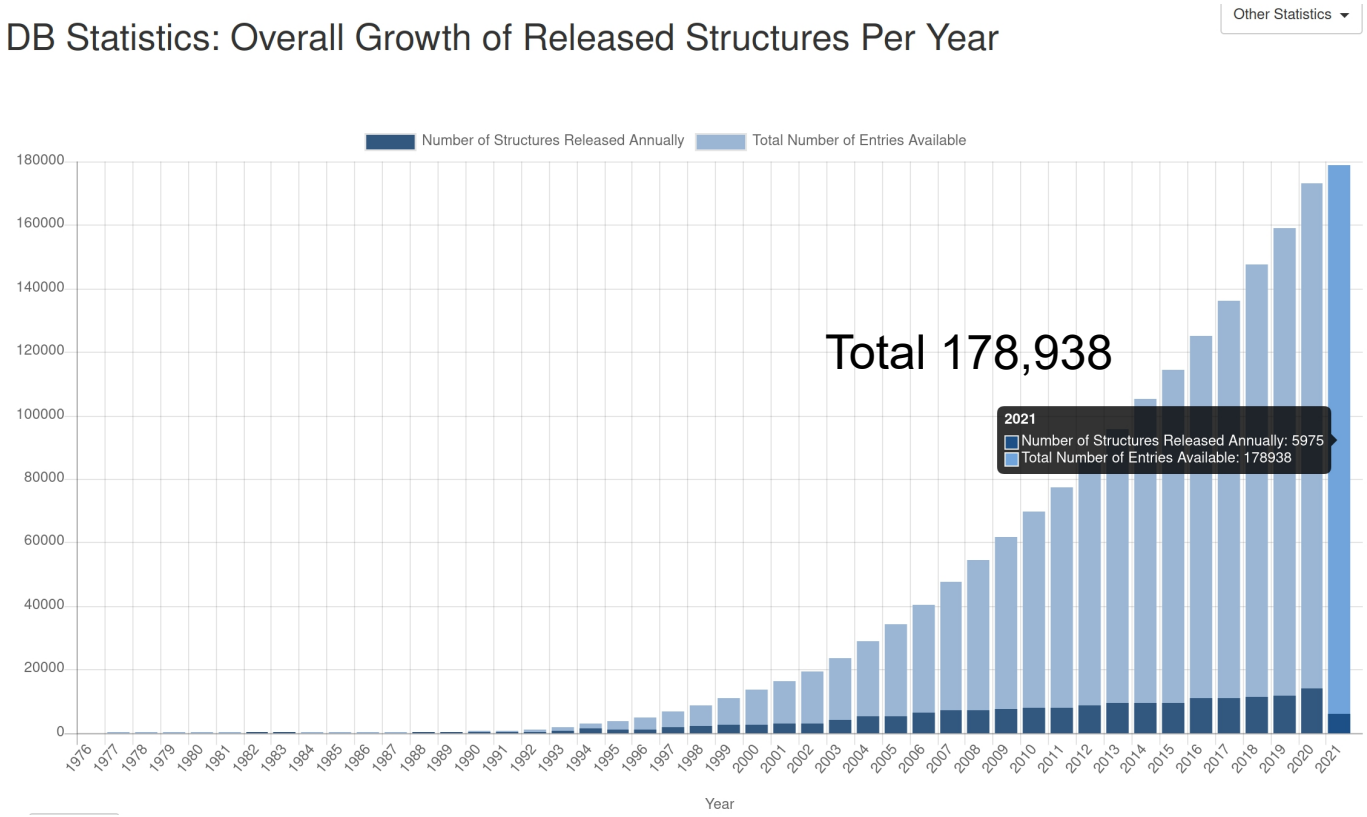
Sequence



GenBank

https://www.ncbi.nlm.nih.gov/genbank/statistics/

Structure

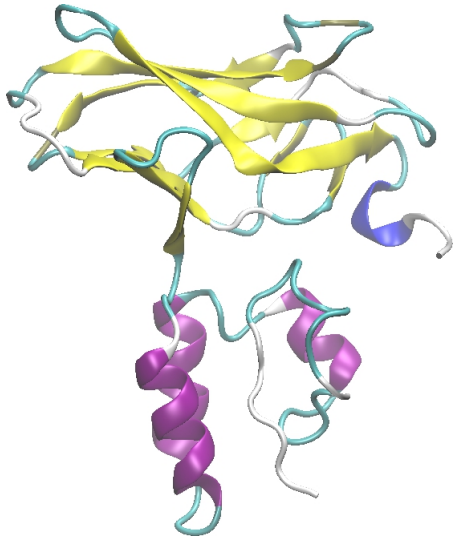


Protein Data Bank

https://www.rcsb.org/stats/growth/growth-released-structures

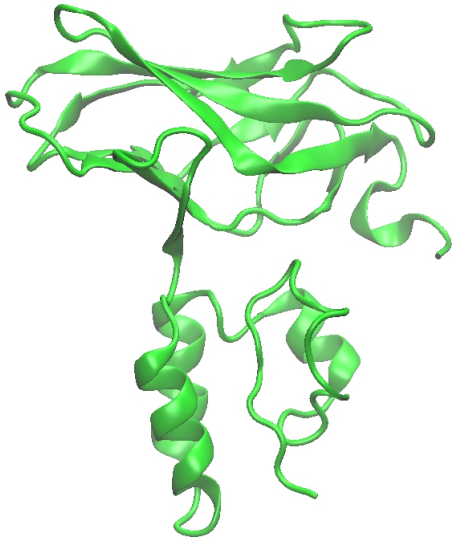
Homology Modeling

Template



Target

New Structure



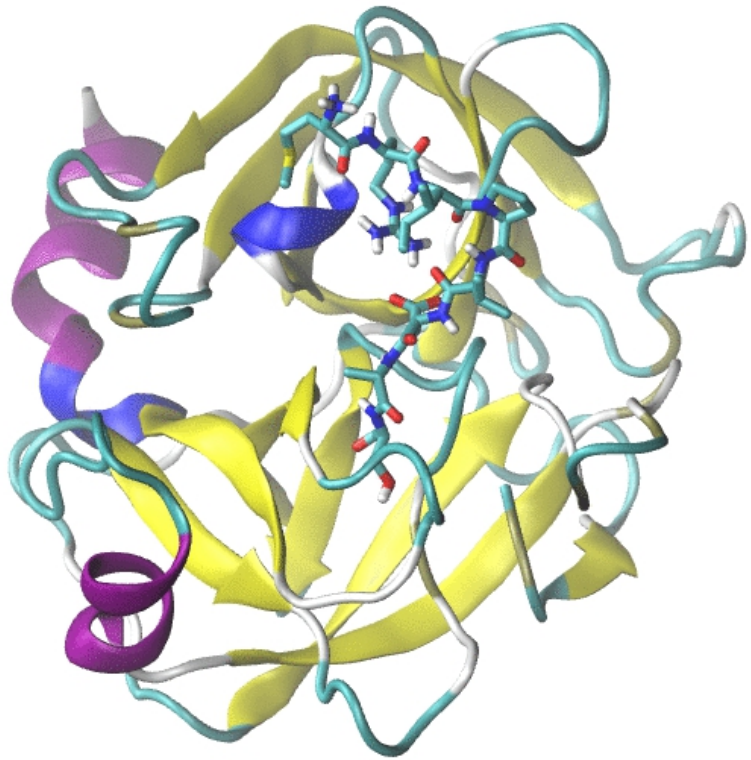
ELAGIILTVSYIPSAEKIA ELAIGILTVSYIPSAEKIR ELAIGILTVSYIPSAEKIR

Sequence Alignment

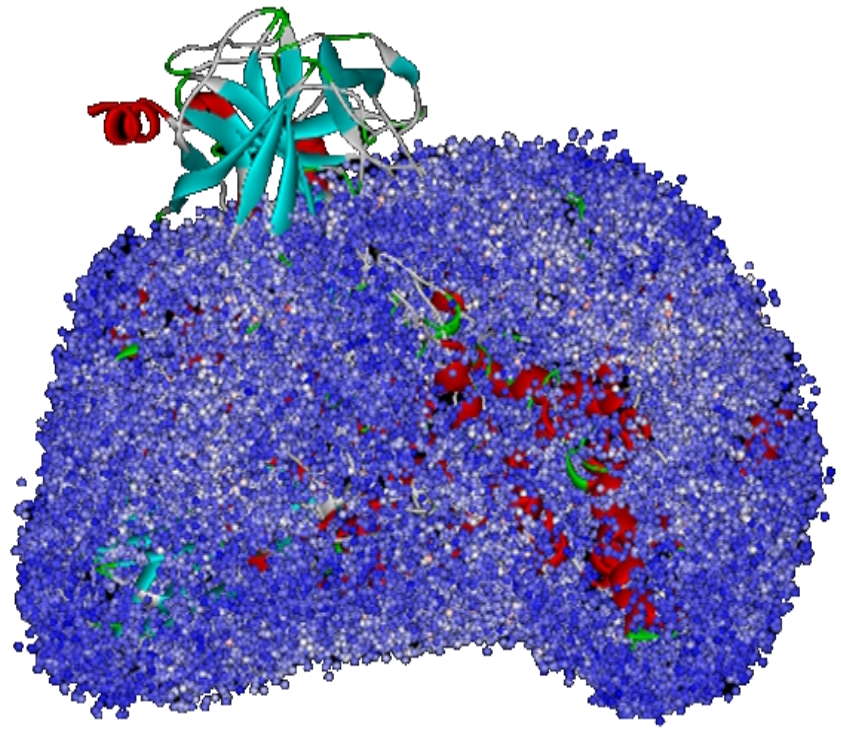
ELAGI-ILGVSYIPSAEKI-ARACELTI

ELA-IGILTVSYIPSAEKIRAR--ELTI

Molecular Docking

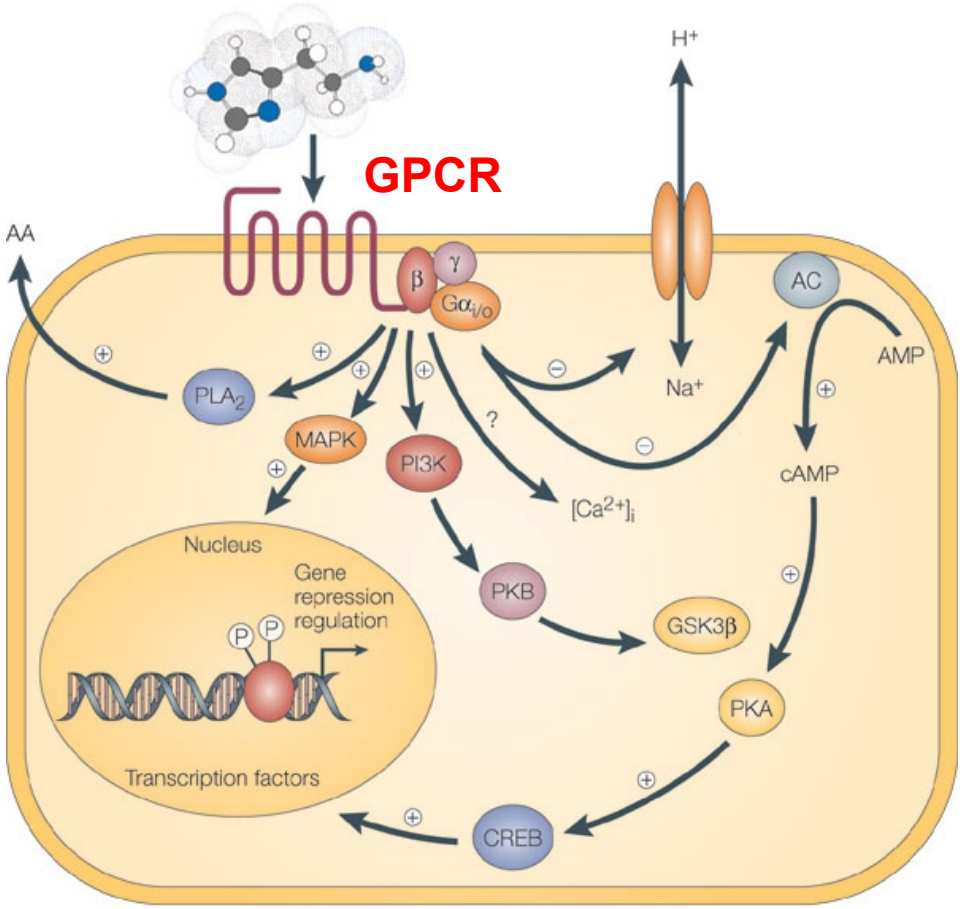


Small molecule Docking



Protein-Protein Docking

GPCR and Signaling Cascades

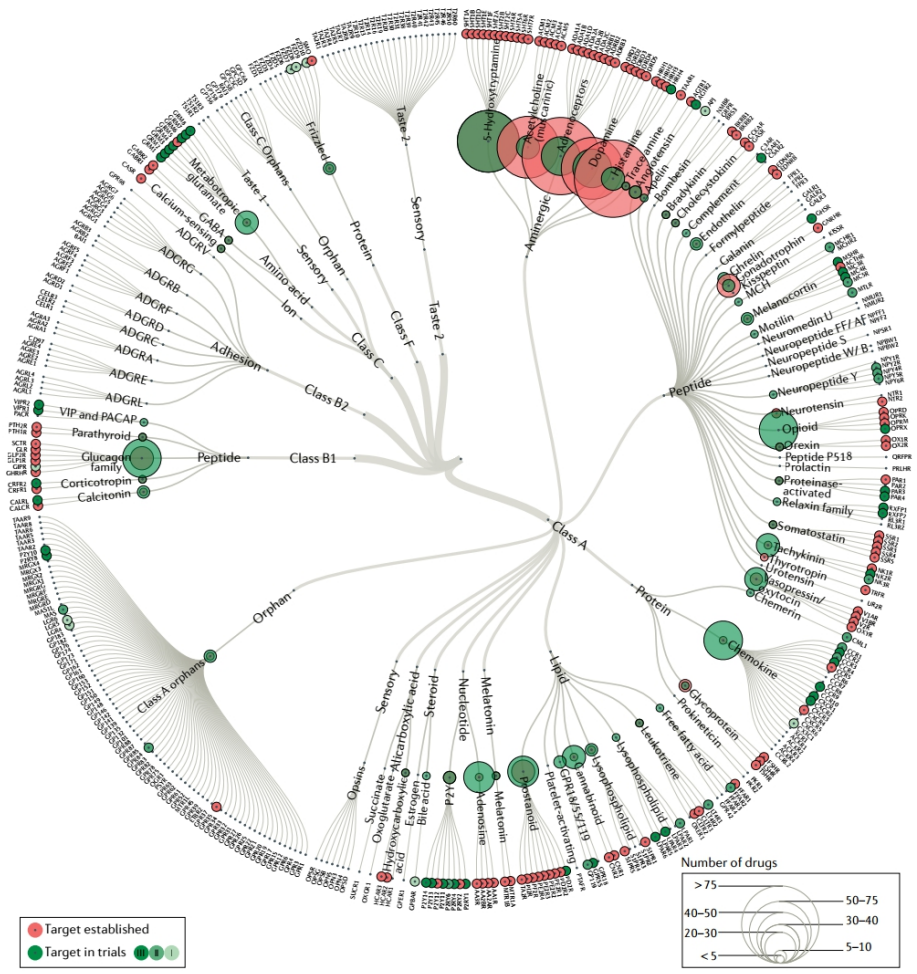


Nature Reviews | Drug Discovery

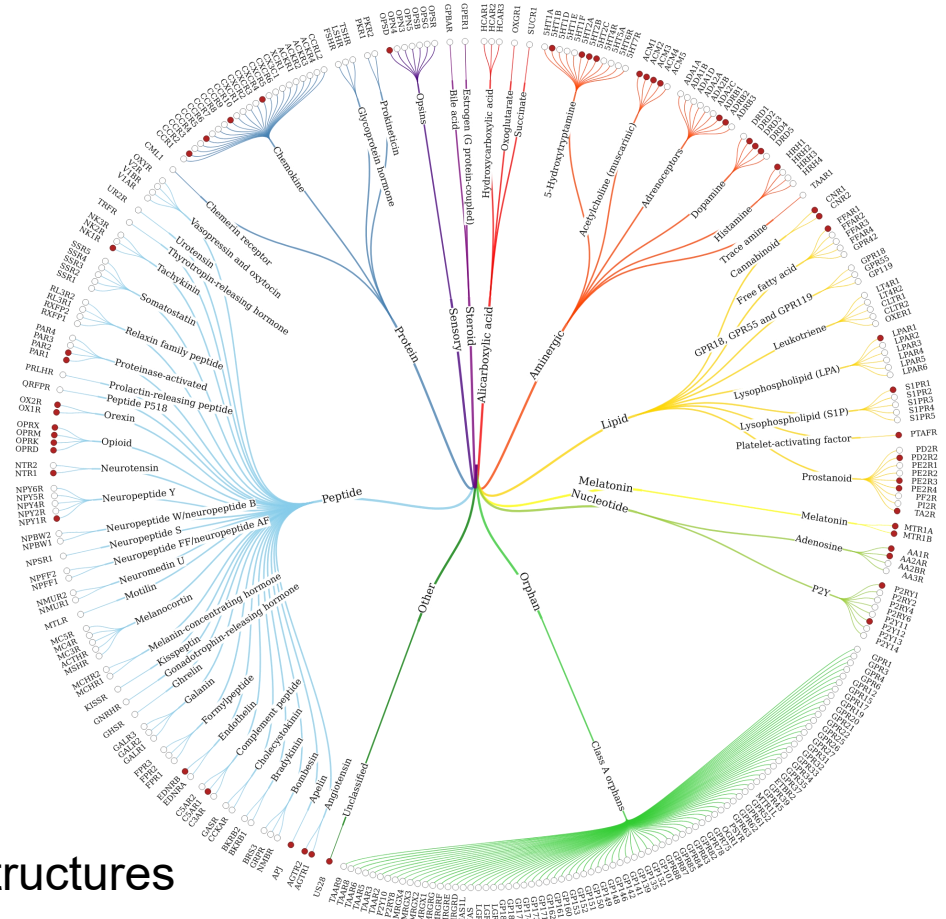
GPCR	Ligands
Active state	Agonists
Inactive state	Antagonists
Rest state	Inhibitors

https://www.nature.com/scitable/content/signaling-cascades-within-a-cell-can-interact-14673527/

GPCR as Drug Targets



475 FDA approved drugs (~34%) act on 108 unique GPCR targets. 321 are in clinical trail.



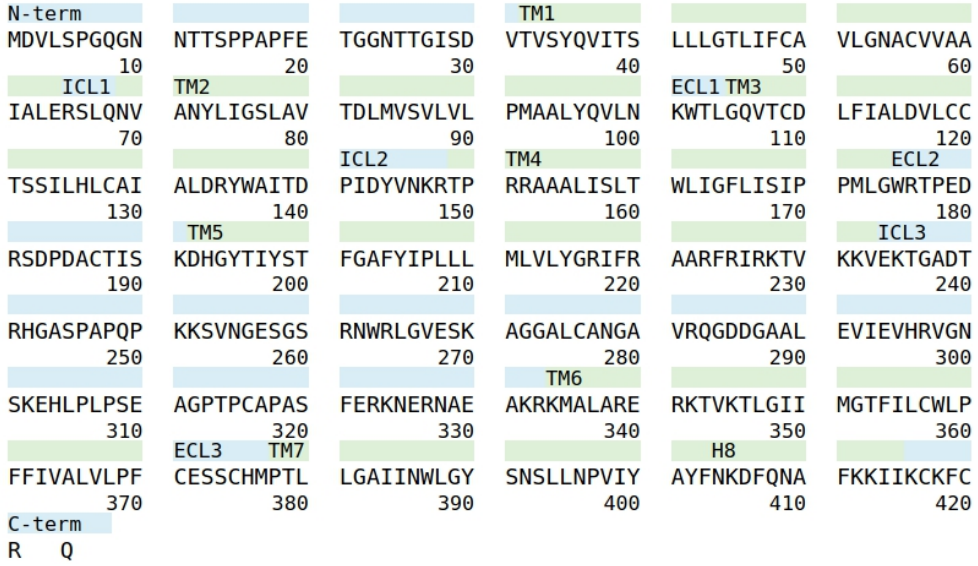
54 structures
vs
800+ class A
receptors

A. S. Hause, et al. *Nature Review*, 2017, vol 16, 829
https://gpcrdb.org/structure/statistics

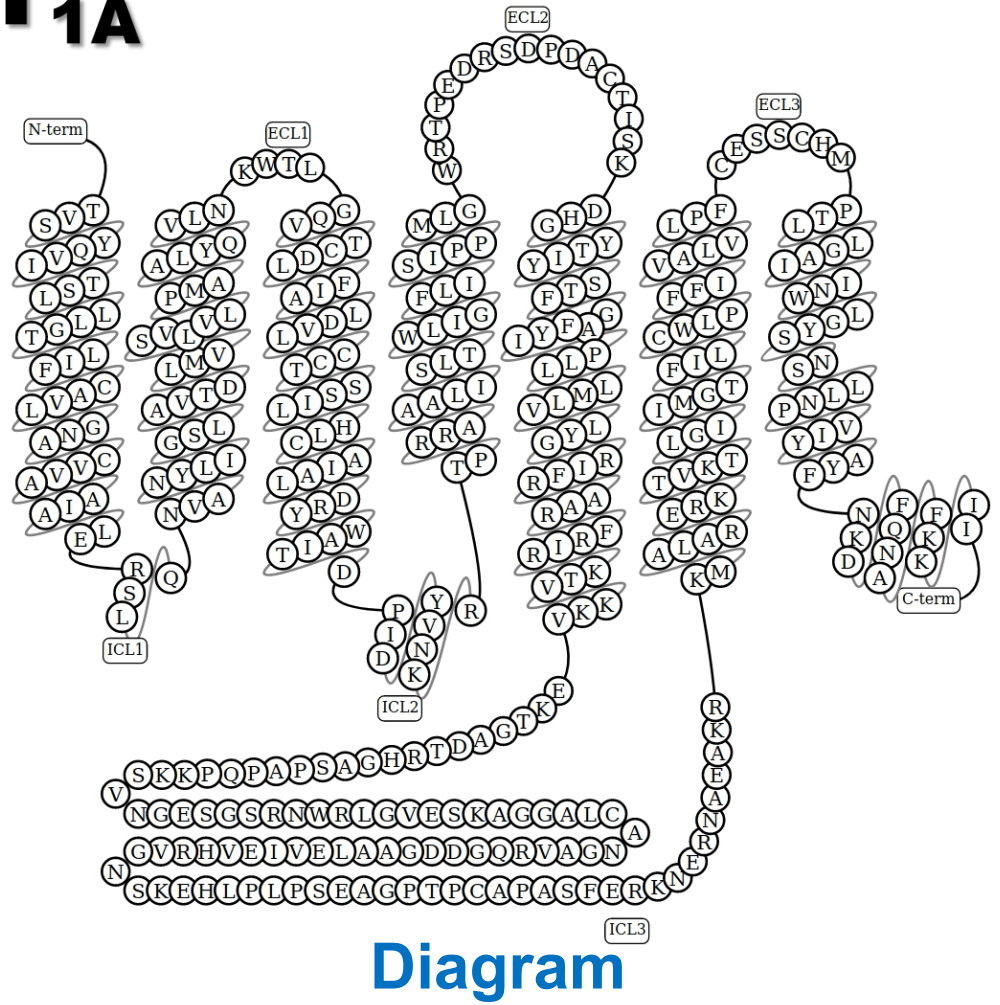
Serotonin Receptor 5-HT$_{1A}$

```
>sp|P08908|5HT1A_HUMAN 5-hydroxytryptamine receptor 1A
OS=Homo sapiens OX=9606 GN=HTR1A PE=1 SV=3
MDVLSPGQGNNTTSPPAPFETGGNTTGISDVTVSYQVITSLLLGTLIFCAVLGNACVVAA
IALERSLQNVANYLIGSLAVTDLMVSVLVLPMAALYQVLNKWTLGQVTCDLFIALDVLCC
TSSILHLCAIALDRYWAITDPIDYVNKRTPRRAAALISLTWLIGFLISIPPMLGWRTPED
RSDPDACTISKDHGYTIYSTFGAFYIPLLLMLVLYGRIFRAARFRIRKTVKKVEKTGADT
RHGASPAPQPKKSVNGESGSRNWRLGVESKAGGALCANGAVRQGDDGAALEVIEVHRVGN
SKEHLPLPSEAGPTPCAPASFERKNERNAEAKRKMALARERKTVKTLGIIMGTFILCWLP
FFIVALVLPFCESSCHMPTLLGAIINWLGYSNSLLNPVIYAYFNKDFQNAFKKIIKCKFC
RQ
```

Sequence



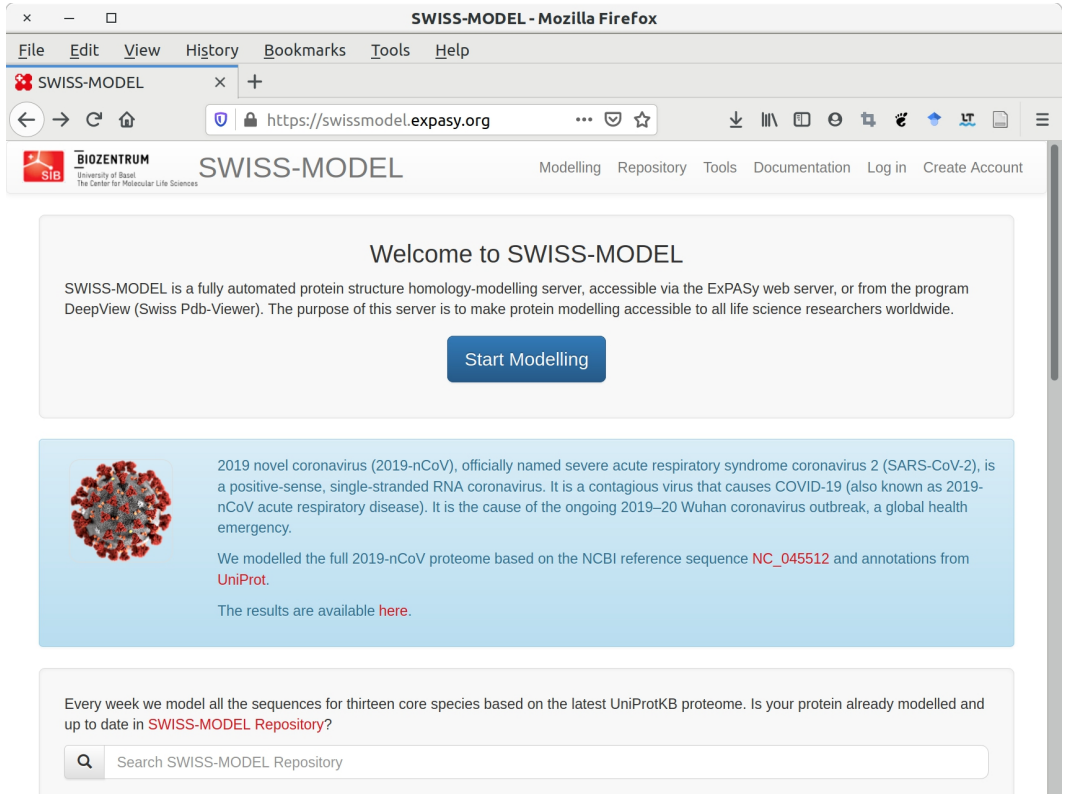
Domain Information

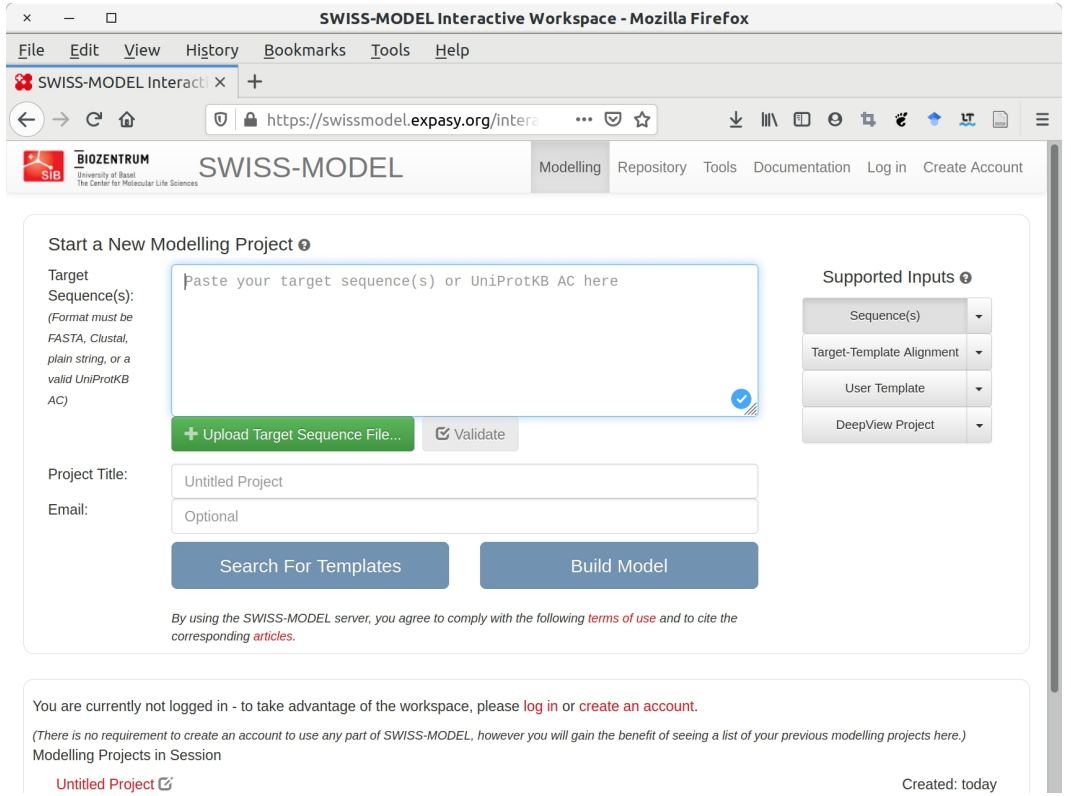


Diagram

But no 3D Structures yet!

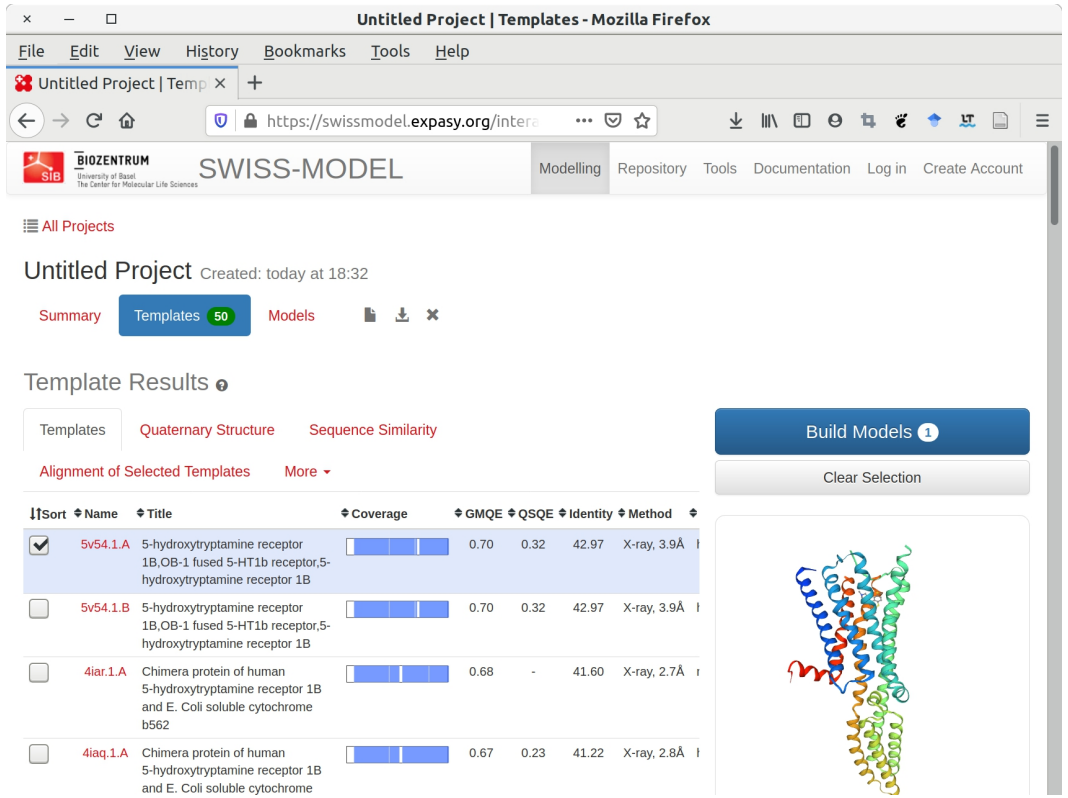
Build 3D Structure of 5-HT$_{1A}$ Receptor with Swiss-model

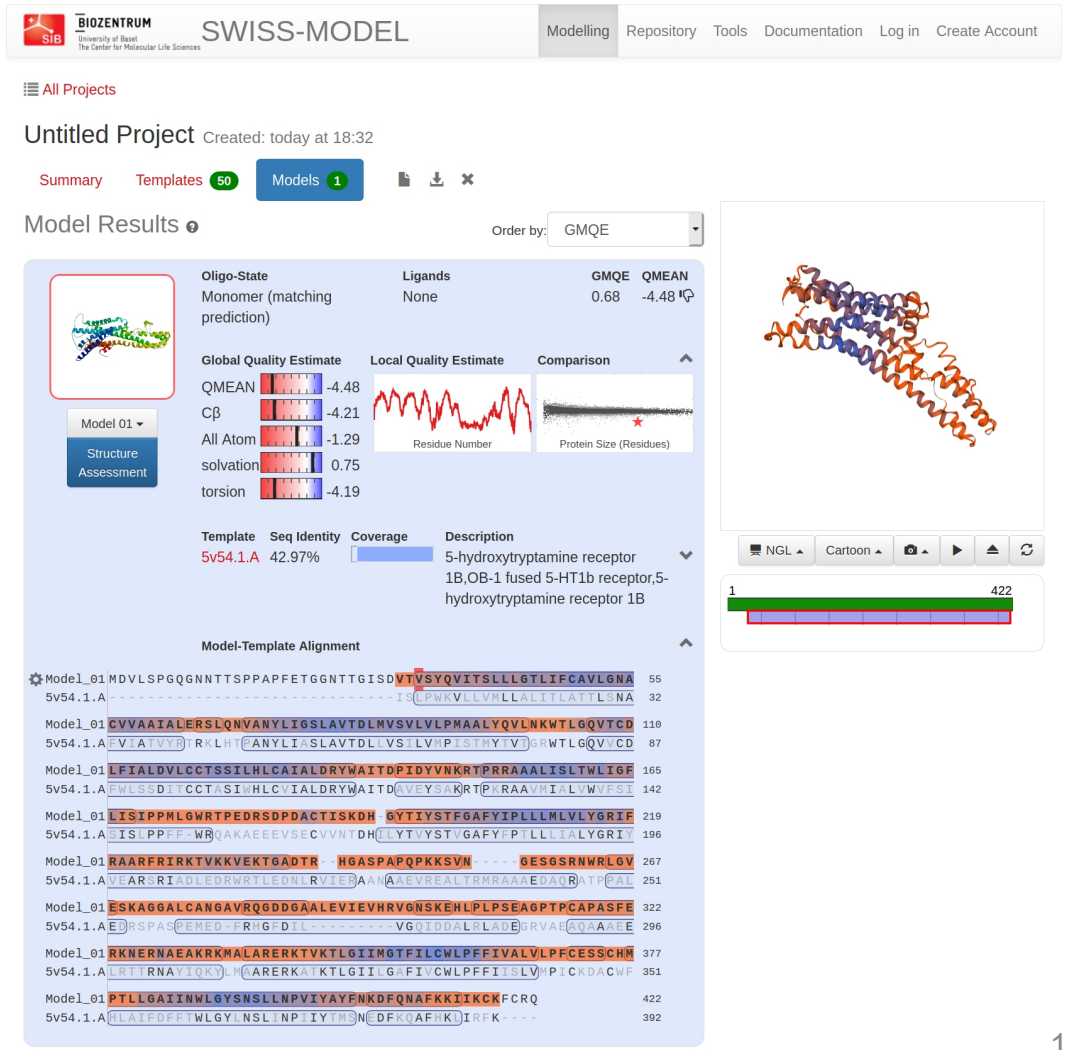




https://swissmodel.expasy.org/

Build 3D Structure of $5HT_{1A}$ Receptor with Swiss-model





Build 3D Structure of $5HT_{1A}$ Receptor with Astrocyte Alphafold

UTSouthwestern Medical Center | BioHPC — Lyda Hill Department of Bioinformatics

Astrocyte 1.0.1

Logged in as: [redacted]

Astrocyte Home | My Projects | Browse Workflows | Documentation

Workflow: biohpc/astrocyte_alphafold.git

Latest Published Version:	0.0.3
Git Repository	git@git.biohpc.swmed.edu:biohpc/astrocyte_alphafold.git (Link)
Last Synchronized	March 24, 2022, 6:08 p.m.
Created	Oct. 3, 2021, 10:26 p.m.
Updated	March 24, 2022, 6:08 p.m.

Check for Updates — Synchronization status: SUCCESS Version **publish_0.0.3** status: READY

UTSouthwestern Medical Center | BioHPC — Lyda Hill Department of Bioinformatics

This workflow provided by: UTSW BioHPC

Published Versions

Version	Git Tag	
astrocyte_alphafold - 0.0.3	**Astrocyte AlphaFold Workflow** This workflow is based on AlphaFold 2.1.1 that supports multimer **Author:** Peng Lian, Xiaochu Lou, Yingfei Chen **Contact:** biohpc-help@utsouthwestern.edu	Run this Version Documentation Developer Information
astrocyte_alphafold - 0.0.2	**Astrocyte AlphaFold Workflow** This is a workflow based on AlphaFold 2.0 **Author:** Peng Lian, Erand Smakaj, Xiaochu Lou, Yingfei Chen **Contact:** biohpc-help@utsouthwestern.edu	Run this Version Documentation Developer Information

https://astrocyte.biohpc.swmed.edu/workflow/48/view

Parameters

Project: Project 1905: alphafold

Name for this run:

A file contains one fasta sequence for monomer OR multiple sequences for multimer prediction (required): 5HT1A.fasta

Choose preset model configuration - the monomer model, the monomer model with extra ensembling, monomer model with pTM head, or multimer model (default: 'monomer') (required): monomer

This is the path to the database folder on BioHPC (required): /project/apps_database/alphafold_2.1.1/database_full

Choose preset MSA database configuration - smaller genetic database config (reduced_dbs) or full genetic database config (full_dbs) (default: 'full_dbs') (required): Full Databases

Maximum template release date to consider (ISO-8601 format - i.e. YYYY-MM-DD). Important if folding historical test sets (required): 2021-12-06

Optional for multimer system, not used by the single chain system. A boolean specifying true where the target complex is from a prokaryote, and false where it is not, or where the origin is unknown. This value determine the pairing method for the MSA (default: 'None'): Unknown

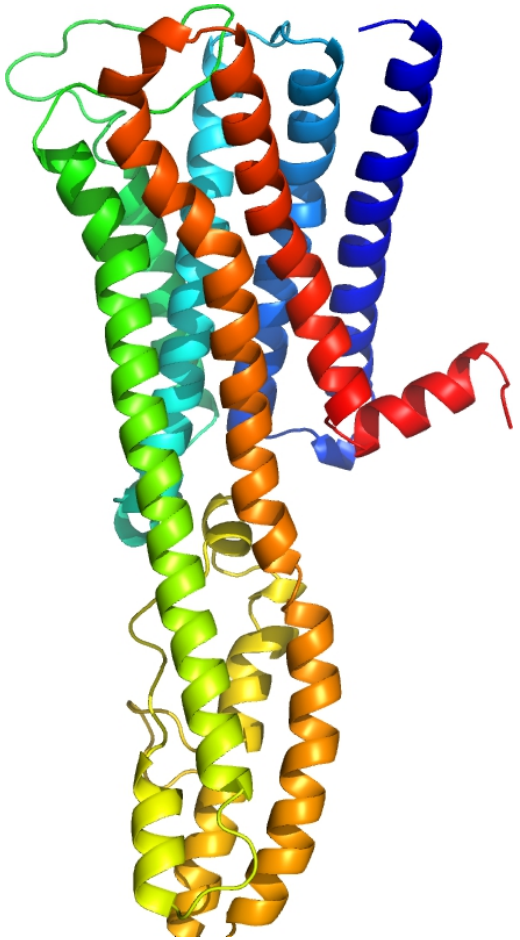
Run multiple JAX model evaluations to obtain a timing that excludes the compilation time, which should be more indicative of the time required for inferencing many proteins (default: 'False') (required): Do NOT use benchmark

Enable NVIDIA runtime to run with GPUs (default: True) (required): Use GPU

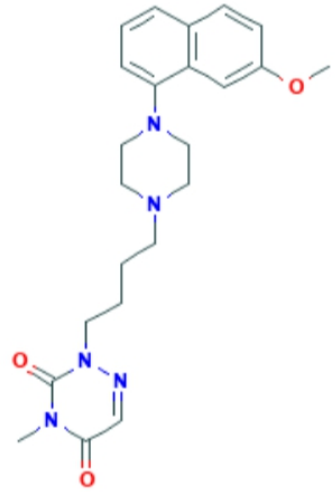
Comma separated list of devices to pass to 'CUDA_VISIBLE_DEVICES' (default: 0) (required): Use GPU 0

Run Workflow

Molecular Docking with Autodock on BioHPC



Homology model of 5-HT_{1A} Receptor



PubChem ID: 10324985
EC50 with 5-HT_{1A}: 0.47 nM

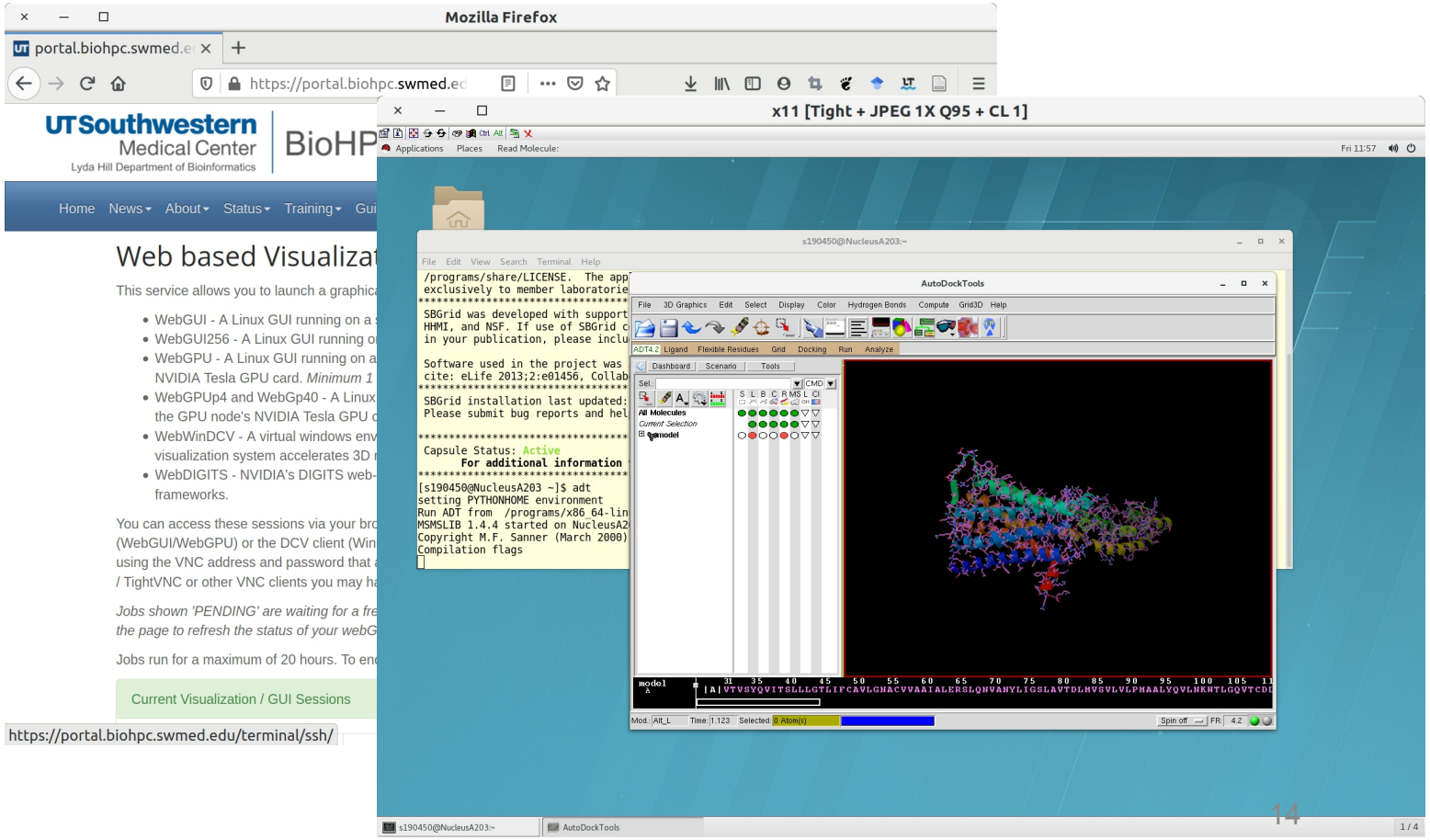
Molecular Docking with Autodock on BioHPC

https://portal.biohpc.swmed.edu/

-> Cloud Services

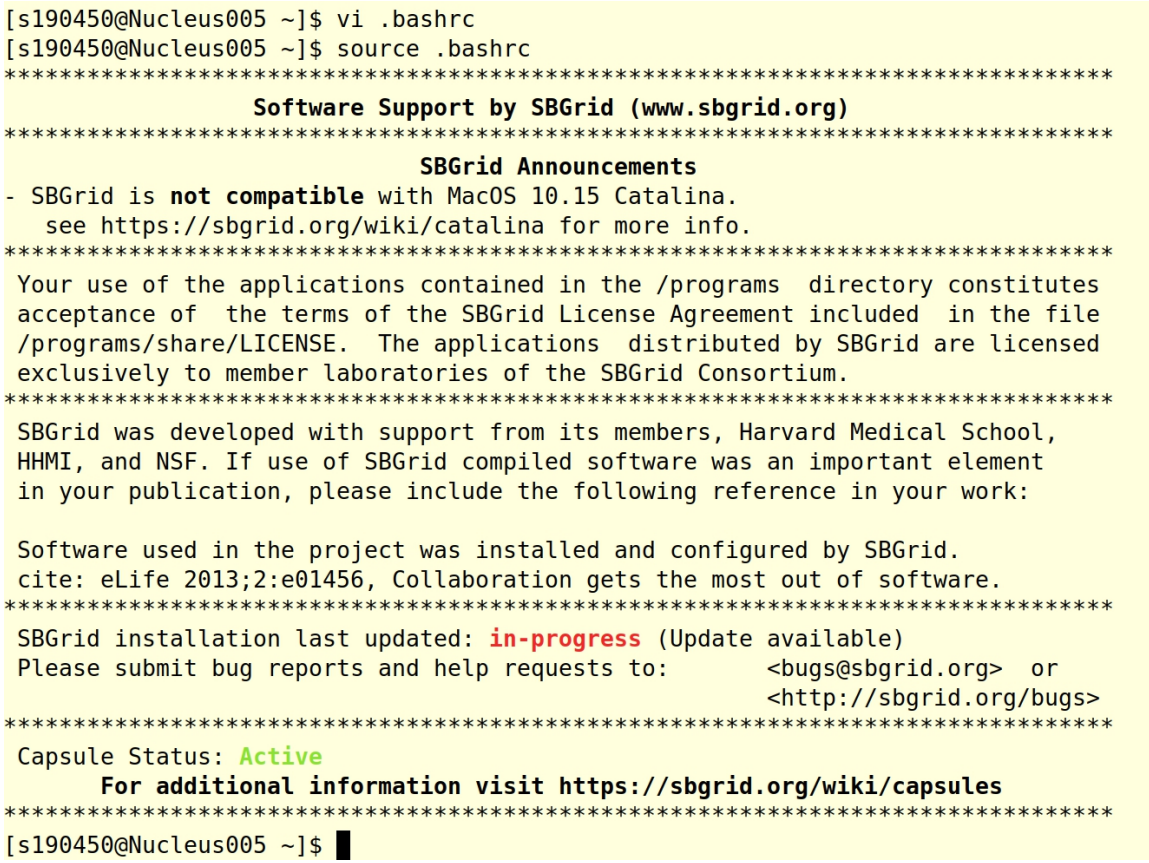
-> Web Visualization

-> VNC Viewer



Molecular Docking with Autodock on BioHPC

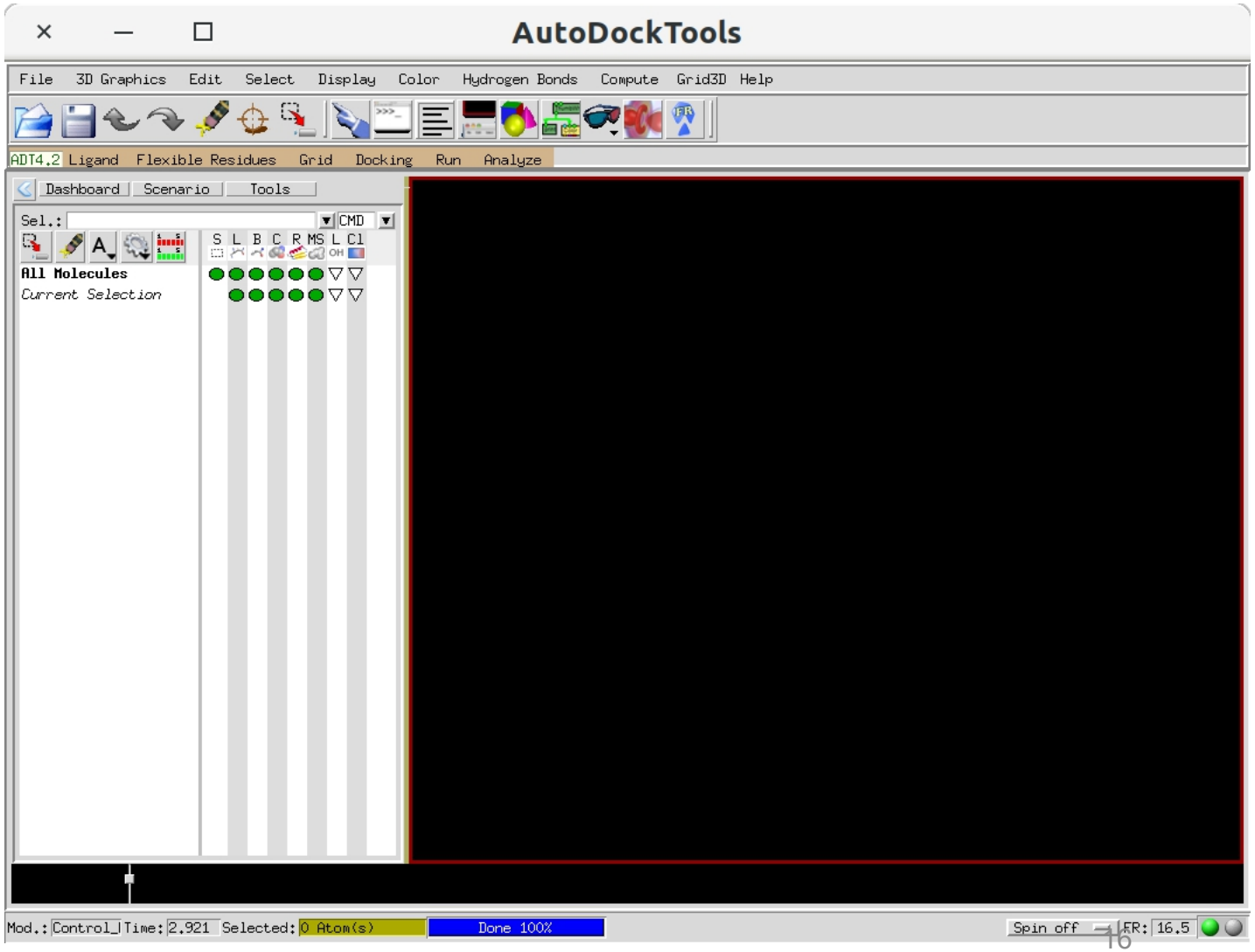
```
# Add sbgrid to your environment

vi .bashrc

. /programs/sbgrid.shrc

source .bashrc
```



Molecular Docking with Autodock on BioHPC

```
# start Autodock Tools
```

```
adt
```



Molecular Docking with Autodock on BioHPC

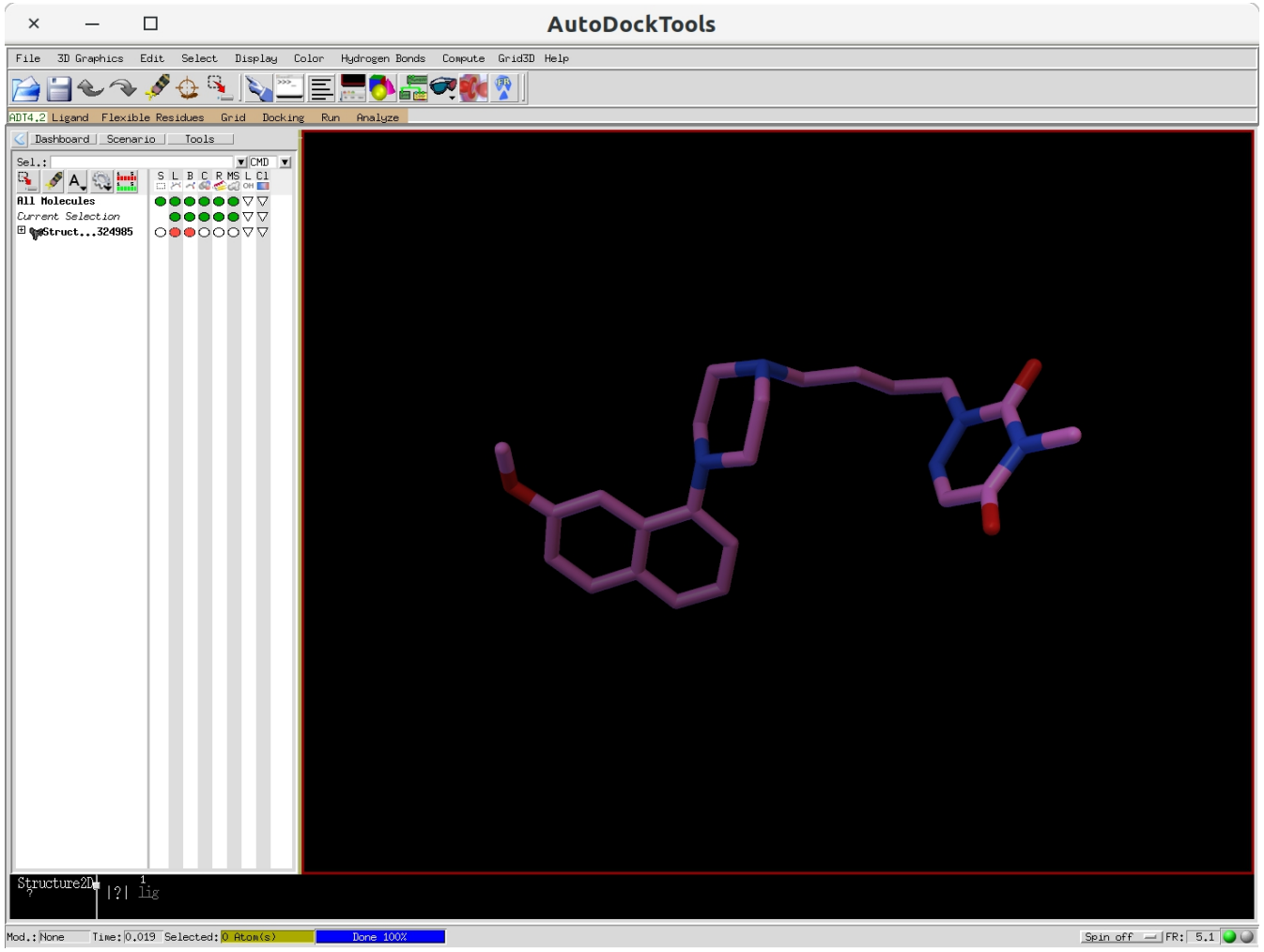
 # Prepare the ligand

Ligand -> Input -> Open -> lig.pdb

Ligand -> Torsion Tree -> Choose Torsions

Ligand -> Output -> Save as lig.pdbqt

Edit -> Delete -> Delete all Molecules



Molecular Docking with Autodock on BioHPC

```
# Prepare the receptor
```

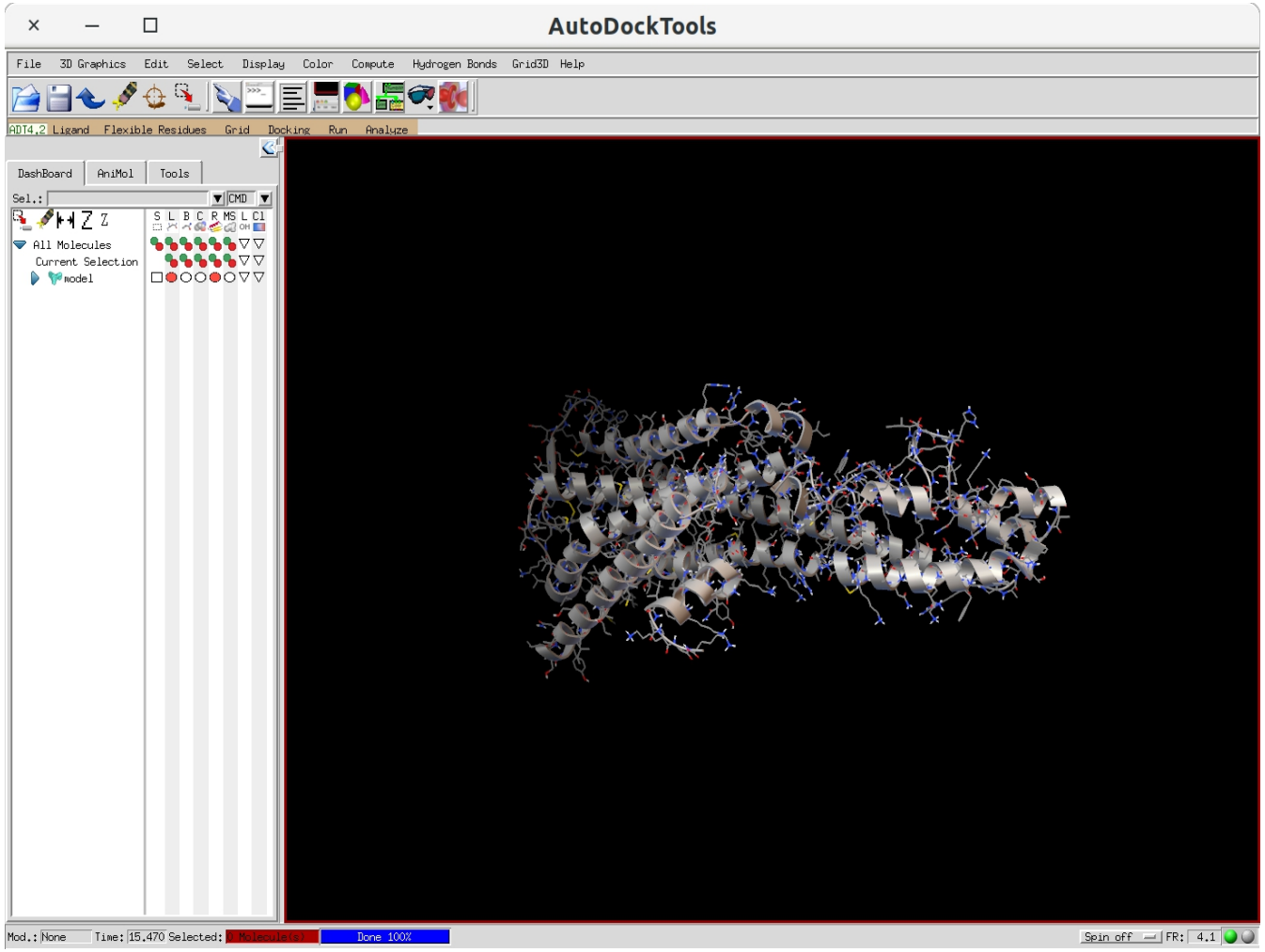
File-> Read Molecule -> model.pdb

Edit-> Hydrogens -> Add -> Polar only

Grid-> Macromolecule -> Choose ->

Save as model.pdbqt

Edit -> Delete -> Delete all Molecules



Molecular Docking with Autodock on BioHPC

 # Prepare the grids

Grid -> Macromolecule -> Open -> model.pdbqt

Grid -> Set Map Types -> Open ligand -> lig.pdbqt

Grid -> Grid Box -> (File -> Close saving current)

Grid -> Output -> box.gpf

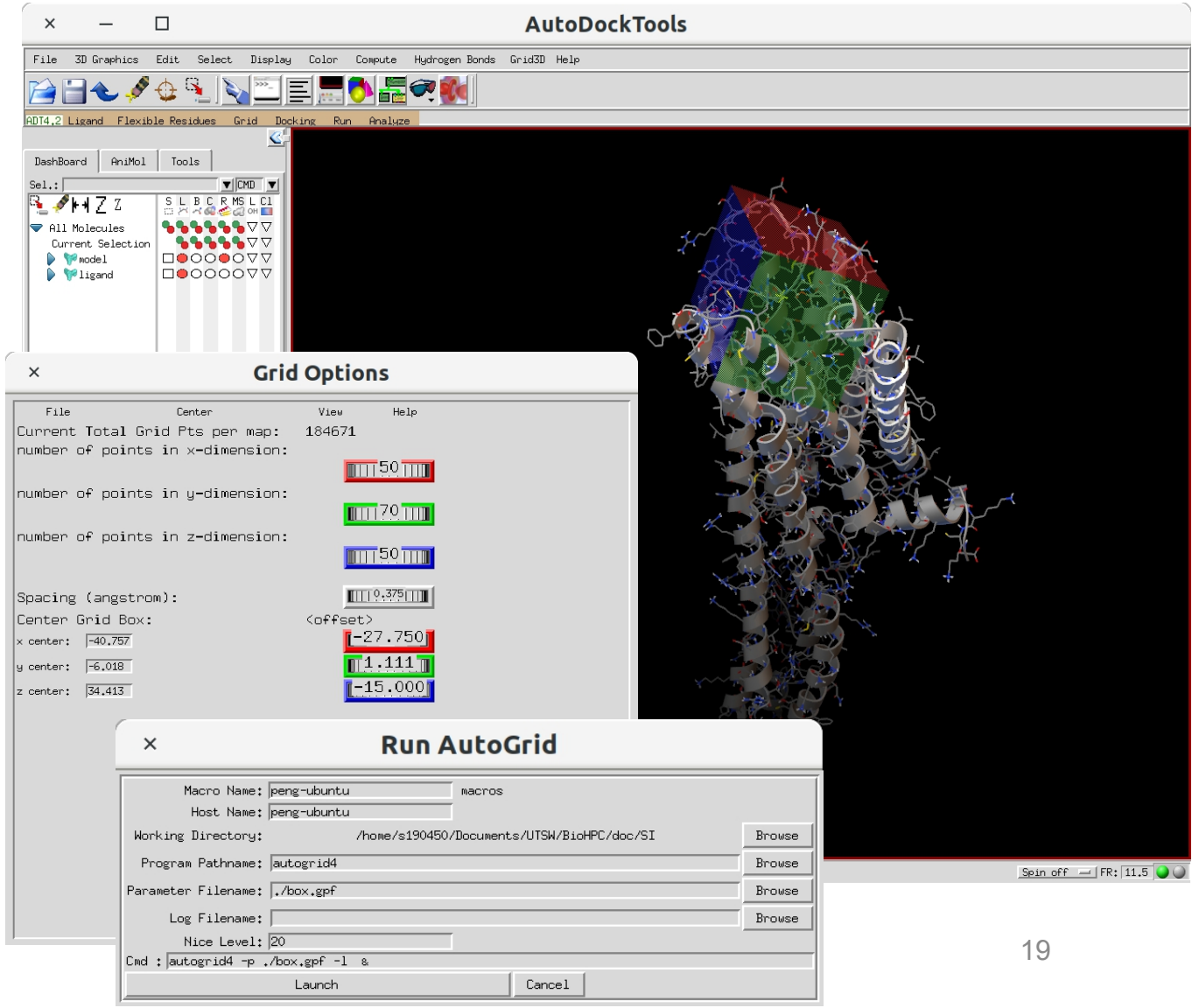
 # run autogrid4

autogrid4 -p box.gpf -l autogrid.log

(Run -> Run AutoGrid)

 # output files

model.A.map model.d.map
model.maps.fld model.NA.map
model.OA.map model.C.map
model.e.map model.maps.xyz
model.N.map



Molecular Docking with Autodock on BioHPC

 # Docking

Docking -> Macromolecule ->

Set Rigid Filename, rigid.pdbqt

Docking -> Ligand -> Open -> lig.pdbqt

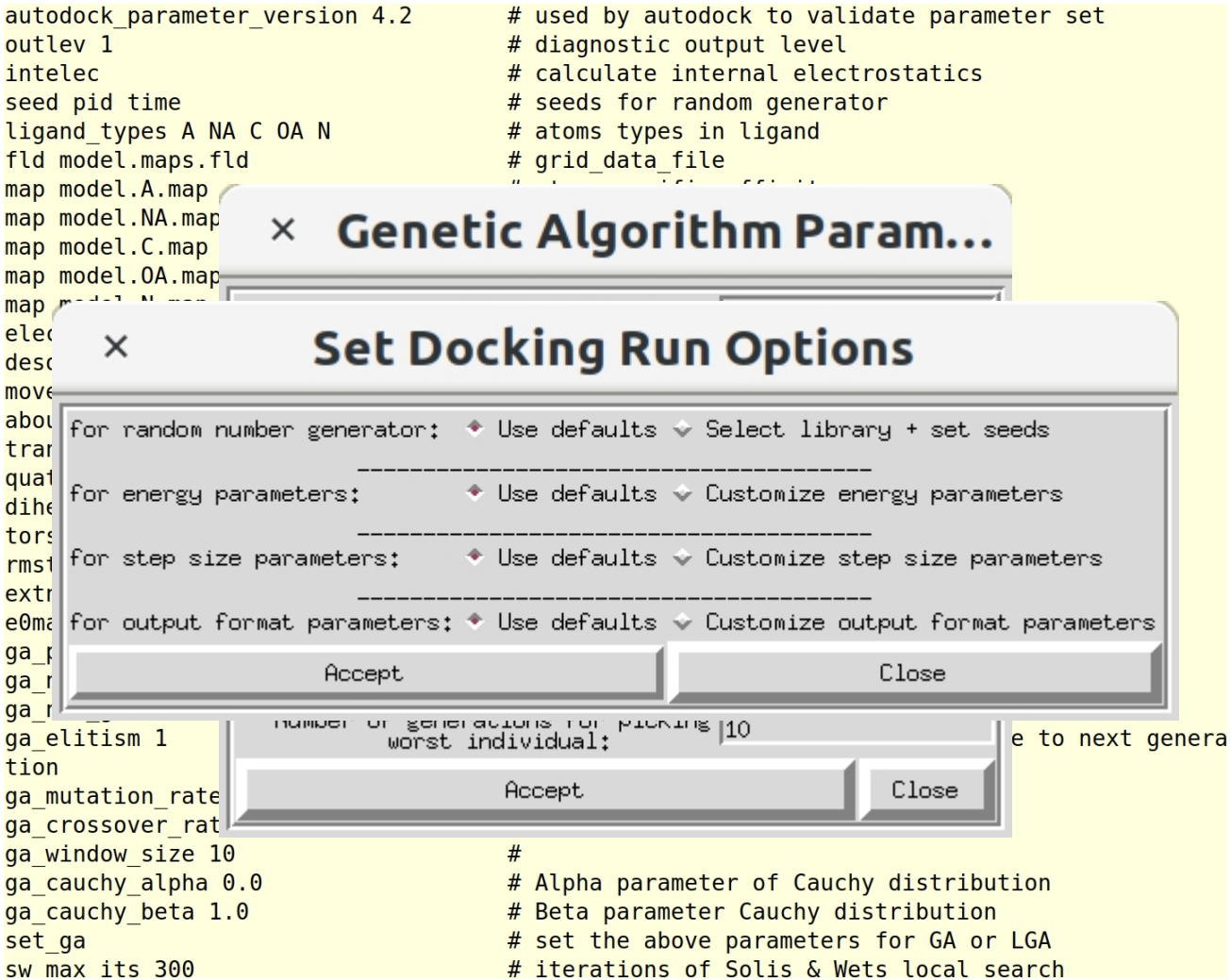
Docking -> Search Parameters -> Genetic Algorithm

Docking -> Docking Parameters

Docking -> Output -> Lamarckian GA(4.2) -> dock.dpf

autodock4 -p dock.dpf -l dock.dlg

(Run -> Run AutoDock)

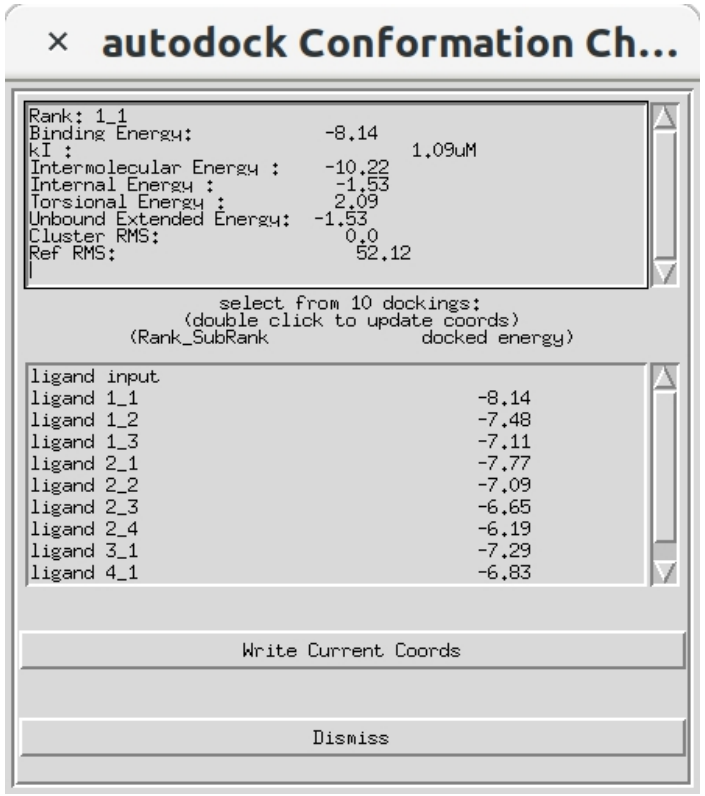


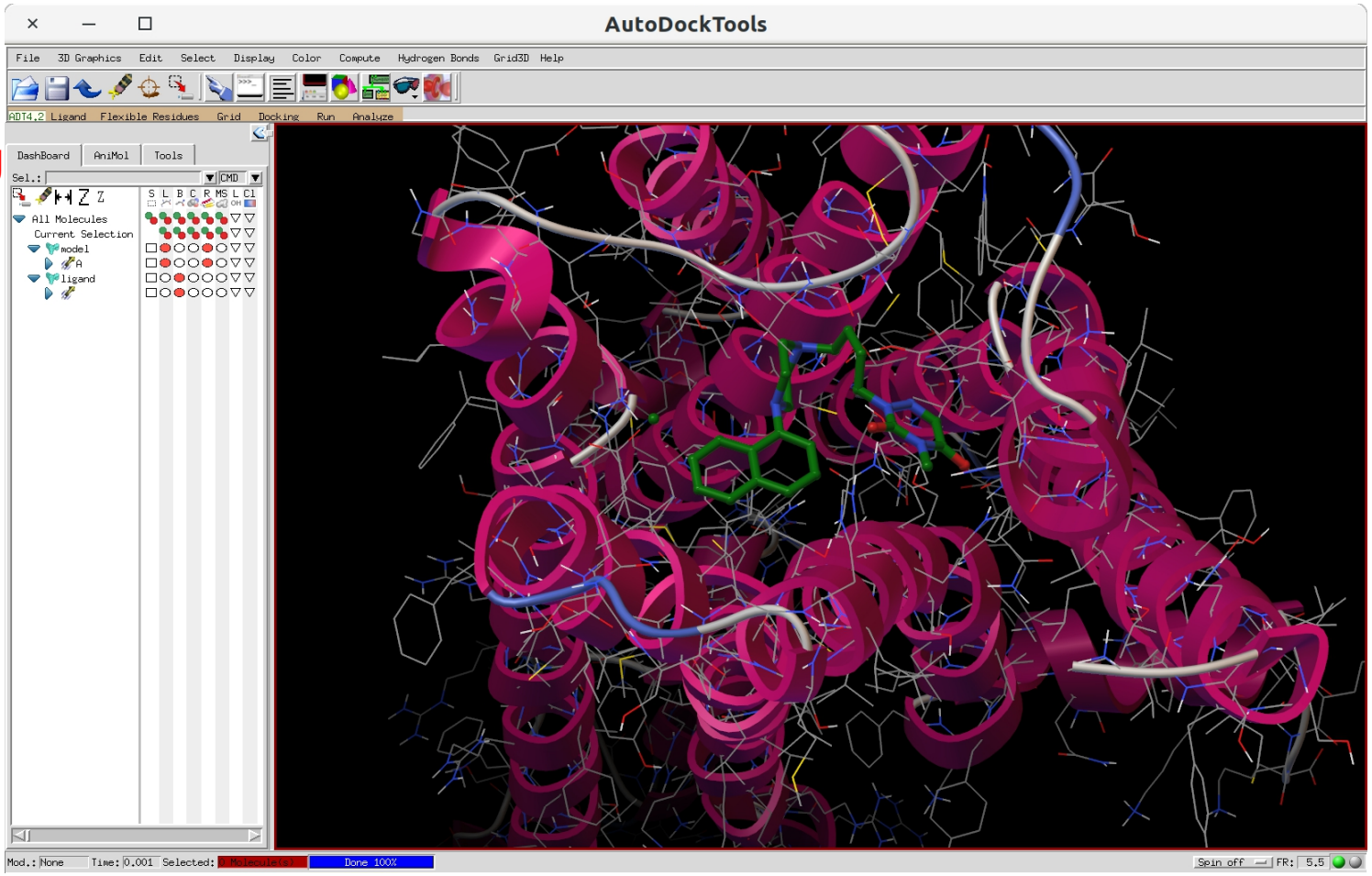
Molecular Docking with Autodock on BioHPC

 # Analysis

Analyze -> Docking -> Open -> dock.dlg

Analyze -> Conformations -> Load -> dock.dlg





Run on Your PC

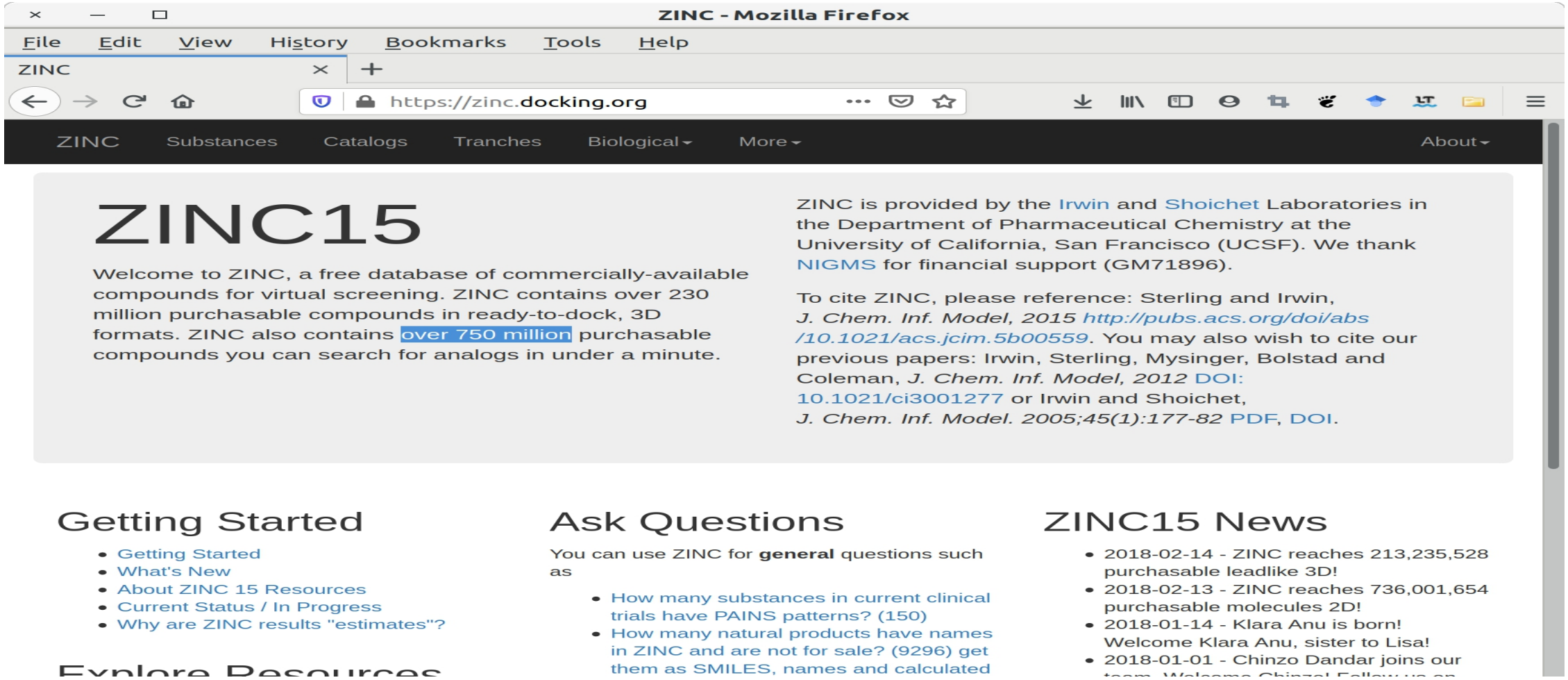
AutoDockTools

http://autodock.scripps.edu/downloads/resources/adt/index_html

AutoDock

http://autodock.scripps.edu/downloads/autodock-registration/autodock-4-2-download-page/

Virtual Screening



Thanks for your attention!

biohpc-help@utsouthwestern.edu