In silico DRUG DESIGN

Presenter: Afolabi-Owoloye Ojochenemi and Afolabi Owoloye

Background

The discovery and development of a new drug is not only expensive, but one that takes as long as 20 years from starting idea to approval and marketing of the new medication. Computer-aided drug discovery (CADD) tools have provided an effective way in the drug discovery pipeline for expediting of this long process and economizing the cost of research and development. These tools are applied in rational drug design to minimize the time for identification, characterization and structure-optimization for novel drug candidates. CADD are categorized into ligand-based drug design and structure-based drug design. The scope of this training will cover the structure-based drug design strategy. Structure-based drug design relies on the knowledge of the three-dimensional structure of the biomolecular target. Here, it is assumed that a drug molecule exerts its biological activity through specific binding to a macromolecular target receptor. So, we attempt to use the structure of proteins as a basis for designing new ligands by applying the principles of molecular recognition

Protein Modeling and Molecular Docking

Protein Modeling

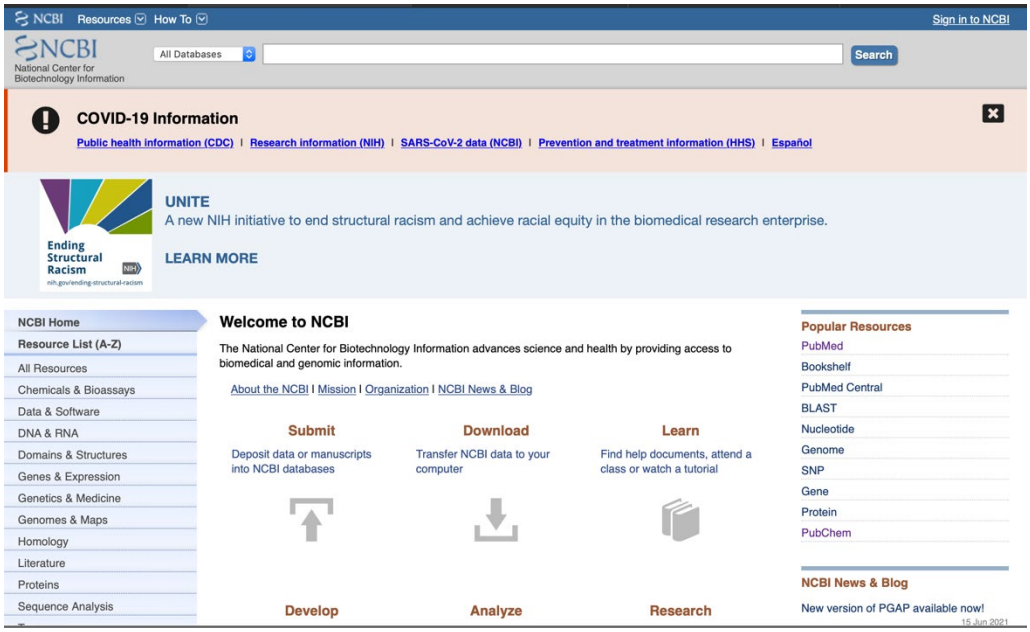
Proteins are modelled when their 3D structures are not available in the protein database. Modeling predicts the 3-dimentional structure of a protein. Homology modeling is the most common type of protein modeling. Tools for homology modeling include

Software: Schrodinger, advanced modeling tool kit, Chimera, MOE.

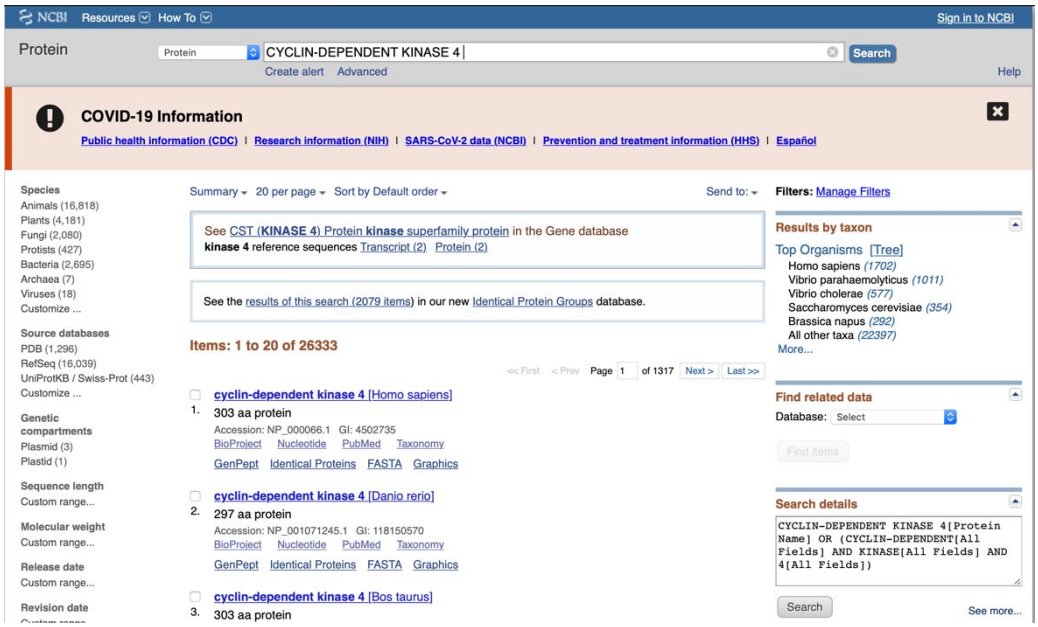
Online Servers: SWISS-MODEL

Homology Modeling Steps:

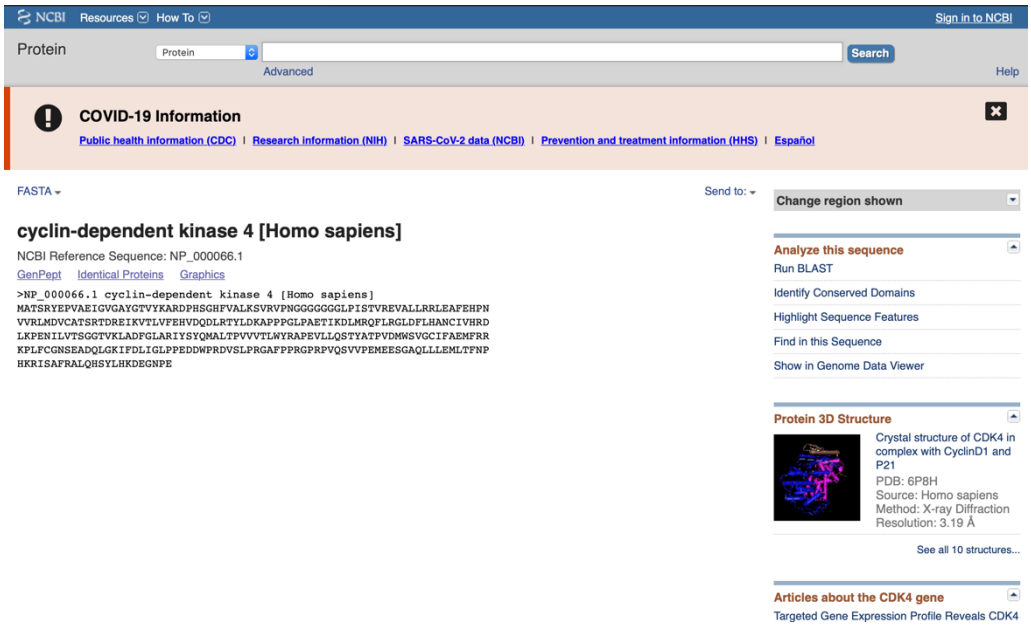
- Retrieve nucleotide sequence of protein of interest from National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov)



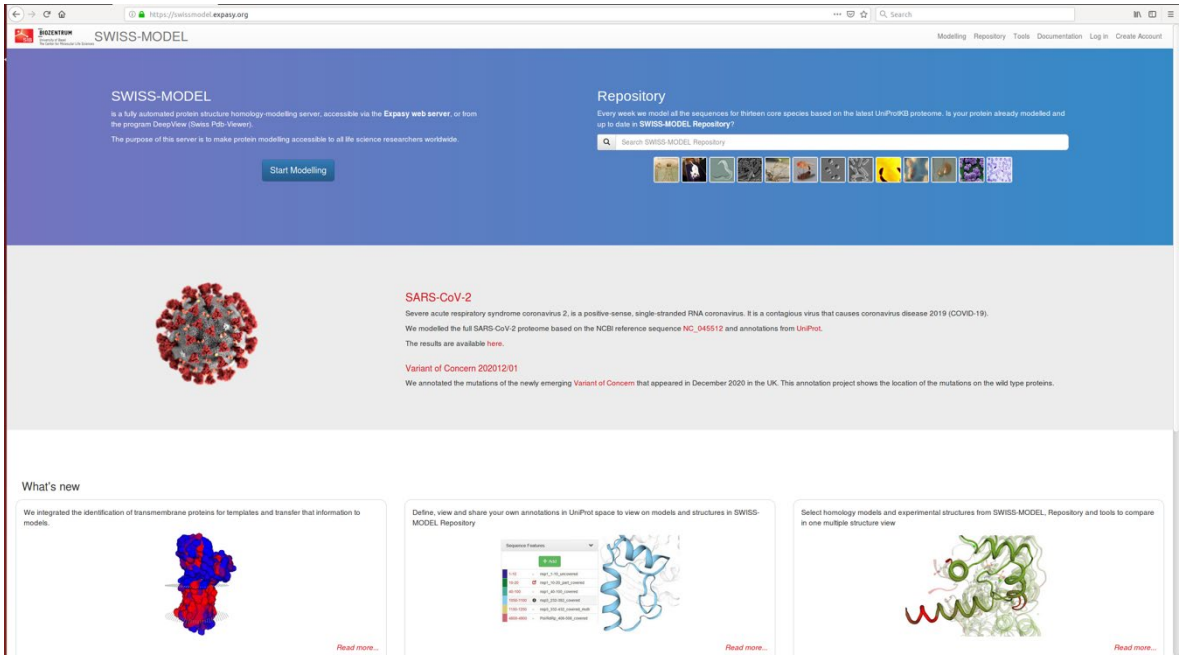
- Change the 'Database Type" to Protein. Type the name of the protein in the box beside it.



- Copy the FASTA format of the nucleotide sequence.



- Go to SWISS-MODEL (https://swissmodel.expasy.org)
- Click "Start Modelling" and paste the nucleotide sequence


https://swissmodel.expasy.org
SWISS-MODEL
Modelling Repository Tools Documentation Log in Create Account
SWISS-MODEL
is a fully automated protein structure homology-modelling server, accessible via the Expasy web server, or from the program DeepView (Swiss Pdb-Viewer).
The purpose of this server is to make protein modelling accessible to all life science researchers worldwide.
Start Modelling
Repository
Every week we model all the sequences for thirteen core species based on the latest UniProtKB proteome. Is your protein already modelled and up to date in SWISS-MODEL Repository?
Search SWISS-MODEL Repository
SARS-CoV-2
Severe acute respiratory syndrome coronavirus 2, is a positive-sense, single-stranded RNA coronavirus. It is a contagious virus that causes coronavirus disease 2019 (COVID-19).
We modelled the full SARS-CoV-2 proteome based on the NCBI reference sequence NC_045512 and annotations from UniProt.
The results are available here.
Variant of Concern 202012/01
We annotated the mutations of the newly emerging Variant of Concern that appeared in December 2020 in the UK. This annotation project shows the location of the mutations on the wild type proteins.
What's new
We integrated the identification of transmembrane proteins for templates and transfer that information to models.
Read more...
Define, view and share your own annotations in UniProt space to view on models and structures in SWISS-MODEL Repository
Read more...
Select homology models and experimental structures from SWISS-MODEL, Repository and tools to compare in one multiple structure view
Read more...

Molecular Docking

Molecular docking has become a very important tool in the drug discovery pipeline. Essentially, the aim of molecular docking is to give a prediction of the ligand-receptor complex structure using computation methods. Docking can be achieved through two interrelated steps: first by sampling conformations of the ligand in the active site of the protein; then ranking these conformations via a scoring function. Ideally, sampling algorithms should be able to reproduce the experimental binding mode and the scoring function should also rank it highest among all generated conformations

Small molecule compounds are easily obtained from databases. Compounds not available in databases can be drawn. The compounds are usually prepared before docking procedure is carried out.

Software for drawing structures: Chemaxon suite, Schrödinger, MOE
Online database: Pubchem database, Zinc database

We start by creating a library of small molecule compounds

- Go to www.pubchem.ncbi.nim.nih.gov
- Change the resource box option to "Pubchem Compound"
- Type the name of the phytocompound
- Click "search"
- The search returns a list of options, from which you select your compound (Be careful to choose the compound of interest and not its derivative)
- The next page displays information about the compound
- Click download 2D save in 'SDF' format

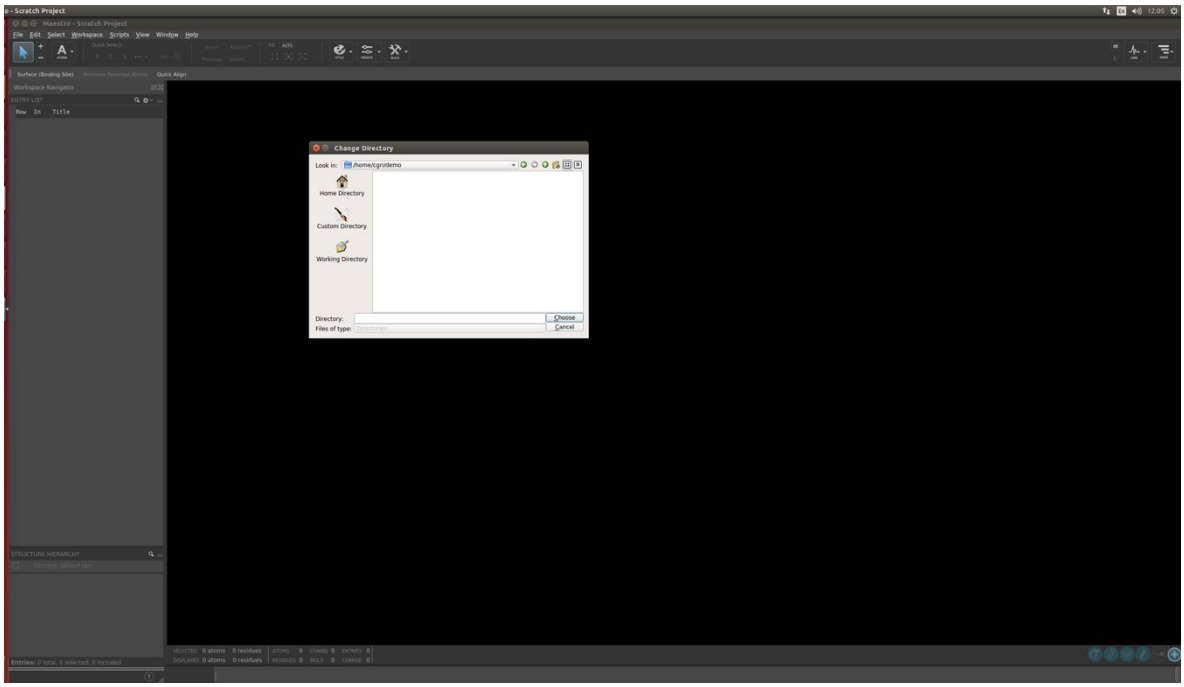
Your protein target must also be prepared before docking procedure.
To retrieve your protein of interest

- Go to https://www.rcsb.org/
- Type the name of the protein in the search box
- Select the protein you want to use. Note: it is good practice to choose a protein with a co-crystallized ligand. Another factor to consider in selecting which protein to download from PDB is the resolution of the protein; the lower, the better.

Molecular docking tools: small molecular weight compounds can be docked into the binding pocket of a protein molecule in order to understand their binding interactions. Scoring functions are used to determine how well a ligand binds to the protein molecule. Tools used: Schrödinger, Autodock vina (docking), open babel, MGL tools, PyRx, etc

Molecular Docking Using Schrödinger Software

The first thing you do after loading Schrodinger is to change your working directory. This helps to keep all your files tidy and in one place.

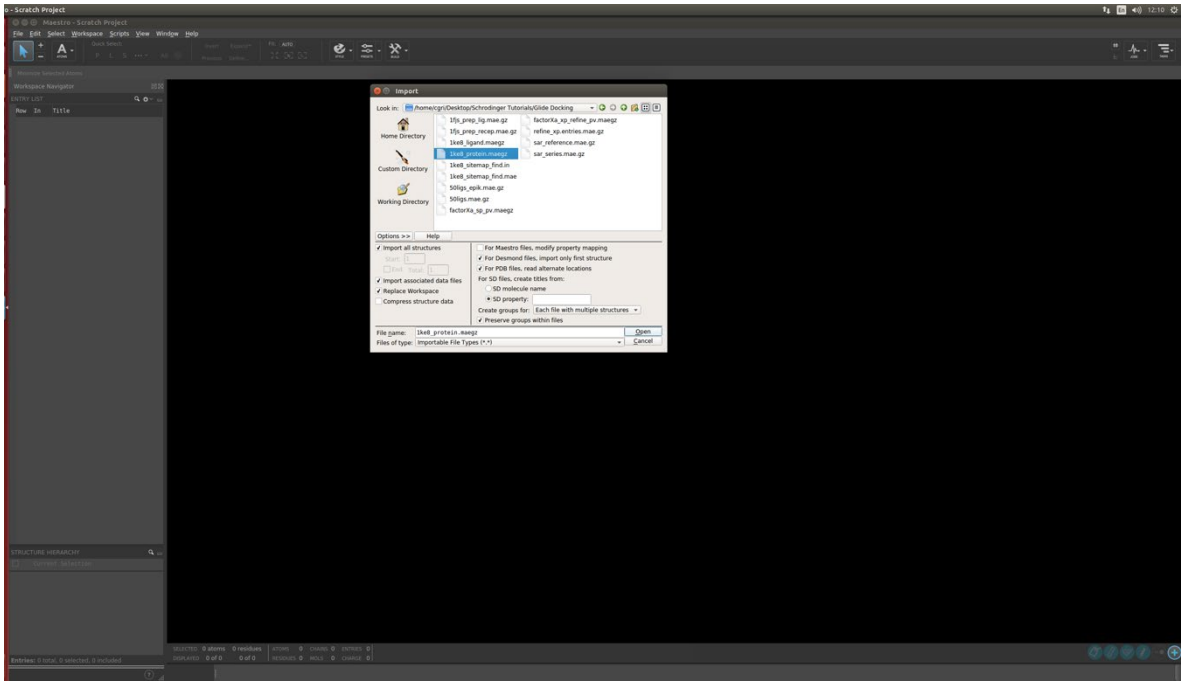


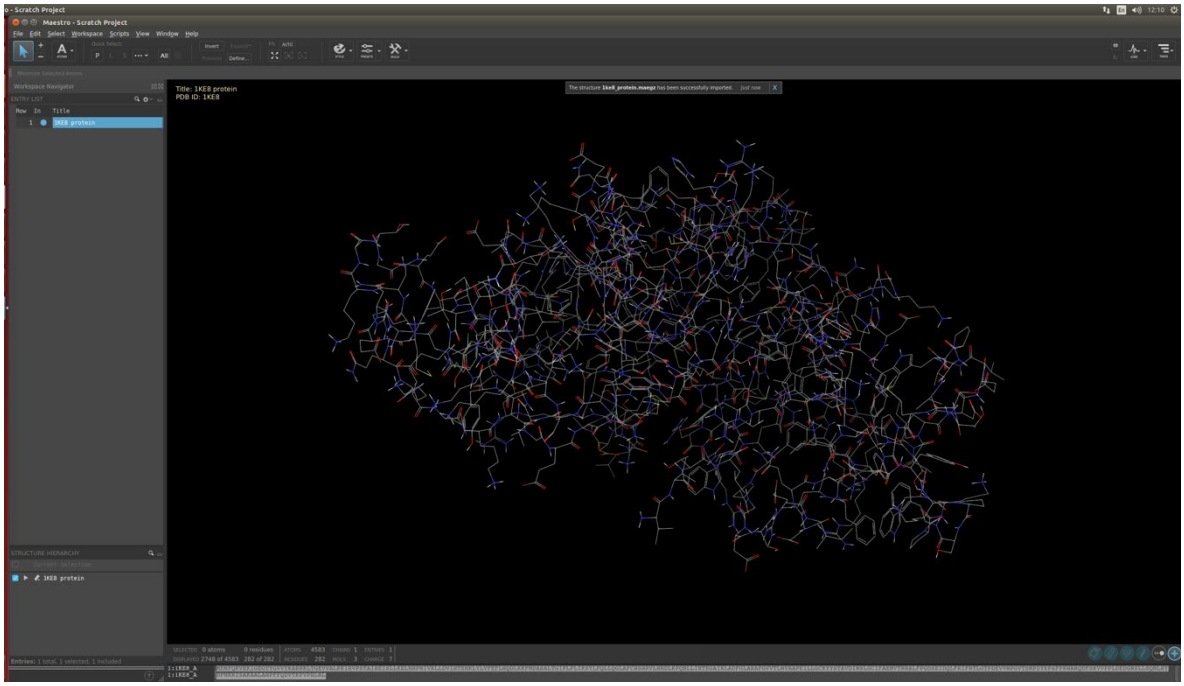
Next, you import the protein structure and prepare it.
To import the protein structure…
Go to “File” and select “import Structures”
Select the file and click “open”

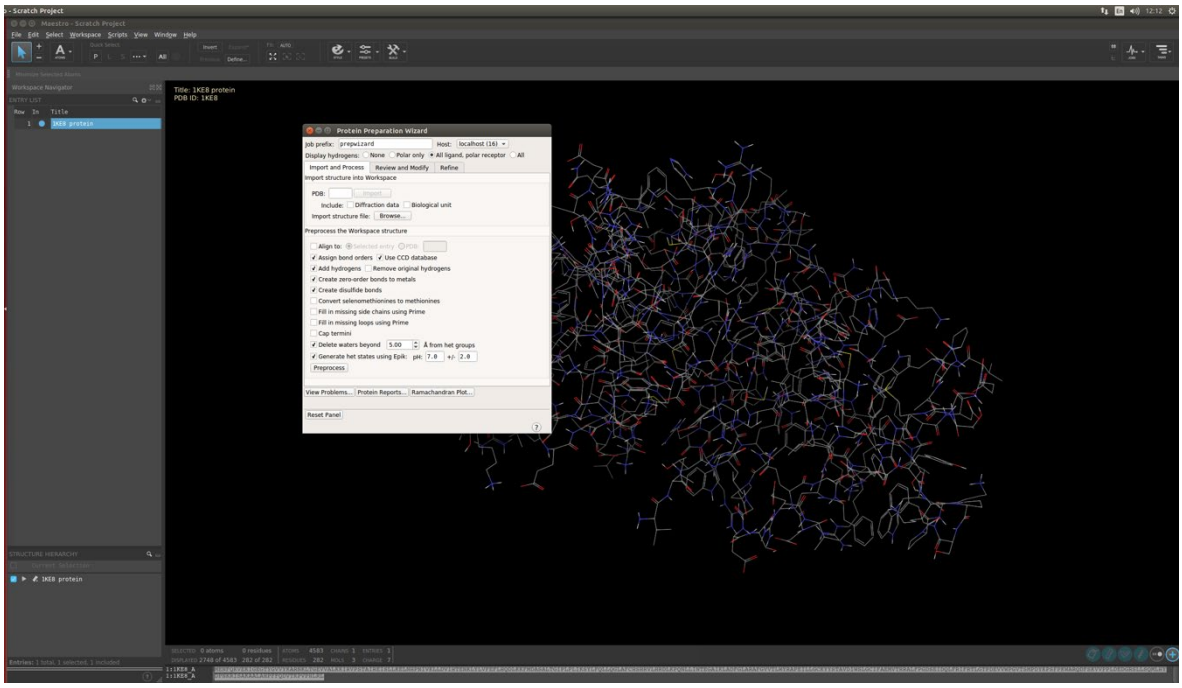
To prepare the protein…
Go to “TASKS” type “Protein Preparation Wizard”
When the option pops up, click on it, a dialog box pops up
In the “Import and Process” tab, select ‘Fill in missing side chains using Prime” and “Fill in missing loops using Prime”
Leave every other thing as default
Click “Preprocess”

When that is done, go to “Review and Modify” tab delete unnecessary chains and water

Go to “Refine” tab click “optimize” after that process is complete, click “Minimize”
Your protein has now been successfully prepared.

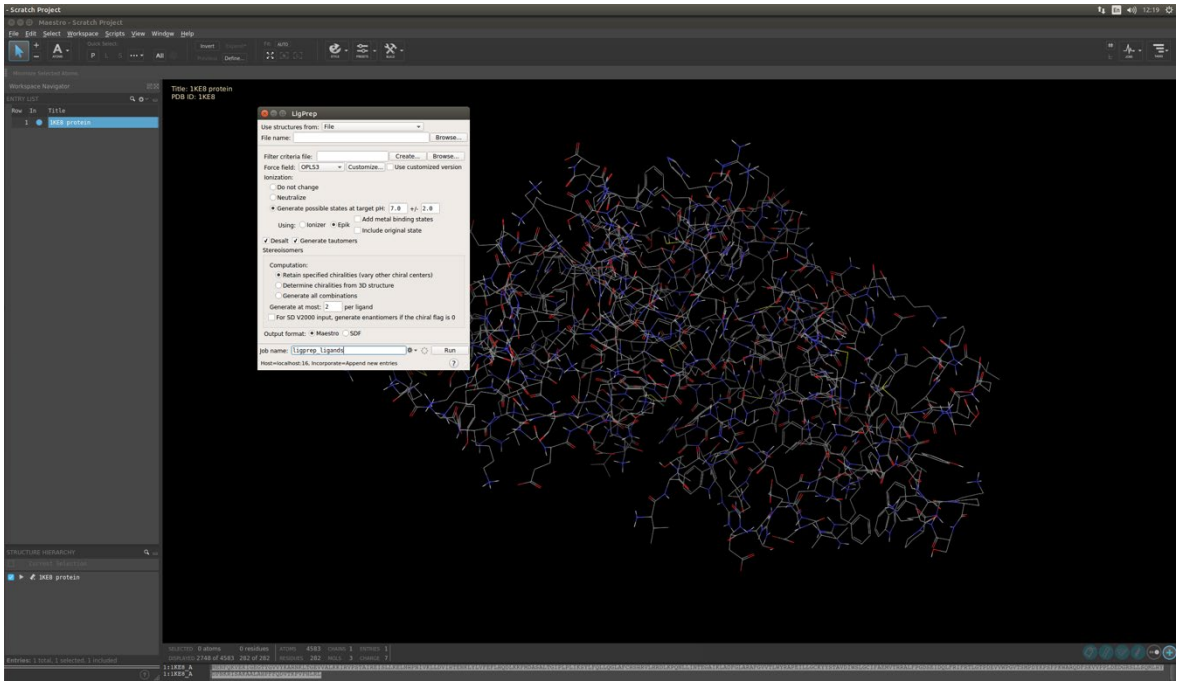

Import
Look in: /home/cgri/Desktop/Schrodinger Tutorials/Glide Docking
Home Directory
Custom Directory
Working Directory
1fjs_prep_lig.mae.gz
1fjs_prep_recep.mae.gz
1ke8_ligand.maegz
1ke8_protein.maegz
1ke8_sitemap_find.in
1ke8_sitemap_find.mae
50ligs_epik.mae.gz
50ligs.mae.gz
factorXa_sp_pv.maegz
factorXa_xp_refine_pv.maegz
refine_xp.entries.mae.gz
sar_reference.mae.gz
sar_series.mae.gz
Options >>
Help
Import all structures
Import associated data files
Replace Workspace
Compress structure data
For Maestro files, modify property mapping
For Desmond files, import only first structure
For PDB files, read alternate locations
For SD files, create titles from:
SD molecule name
SD property
Create groups for: Each file with multiple structures
Preserve groups within files
File name: 1ke8_protein.maegz
Files of type: Importable File Types (*.*)
Open
Cancel


Title: 1KE8 protein
PDB ID: 1KE8
The structure 1ke8_protein.maegz has been successfully imported. Just now



After the protein has been prepared, the next step is to locate the protein binding site.
Go to “TASKS” type “Receptor Grid Generation”
When the option pops up, click on it, a dialog box pops up
In the Receptor tab, leave everything as default
In the Site tab, ensure “Centroid of workspace ligand (selected in the Receptor tab)” is selected.
Give your job a name
Click Run

For proteins without co-crystallized ligands, run ‘SiteMap’ to predict the binding site of the protein.
Go to “TASKS” type “SiteMap”
When the option pops up, click on it
A new dialog box pops up
Leave everything as default
Give your job a name
Click Run



When the binding site of the protein has been generated, go to “Tasks” type “Ligand Docking”
In the Receptor grid box, choose the zip file that has details of the grid generated.
Under the Ligands tab, go to “Use ligands from:” selected the location of your previously prepared ligands.
Under Settings tab, go to Precision: select “XP (extra precision)”; Ligand sampling: select Flexible.
Give your job a name and click Run.

When the job is done, click on the “Show Project Table” icon to view your results