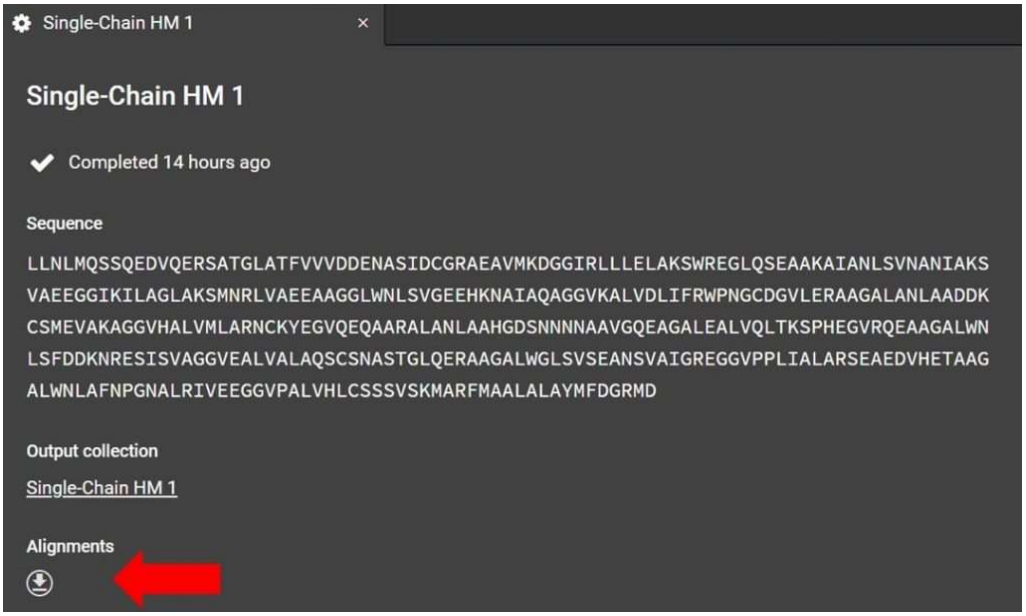
Quick Start Guide for Cyrus Bench

Detailed below are instructions for advanced users of Cyrus Bench. If you are not familiar with these processes please refer to the support.cyrusbio.com site for our in-depth tutorials and analysis.

1.) Running HM to get a structure prediction

a. Go to https://cad.cyrusbio.com/projects
b. Click Projects ⊕ to create a new project. Name it and click Create project.
c. Click on the project name in the list to open the project
d. Click Single-Chain HM to create a new HM run.
e. Enter your sequence into the box or click ⬆ to upload a sequence. Click Run 1 sequence.
f. Wait patiently for 1 to 10 hours for it to finish (depending on sequence length).
g. The results of your completed HM run will appear in a folder to the left in CAD Single-Chain HM 1. Click on the folder to view your results.
h. Click 👁 on each model to bring its structure into the viewer on the right.
    i. Superpose these structures; if these models are very similar to each other, it indicates your HM run reached the energetically minimum structure.
    ii. Regions that have variation could mean it is a dynamic region or that there is some inadequacy of the modeling (lack of good template for this region, missing component for modeling, etc).
i. Open the original HM run tab to access the alignment file, as shown below.



j. Alignment files(.json) can be viewed with any Text Editor, such as Atom.

Alignments file will list each homolog in this format:

{
"targetAlignment":
"TQIAYRSDDWRDLKEAWKKGADILIVDATDKDEAWKQVEQLRREGATQI--------
----AYRSDDWRDLKEAWKKGADILIVDATD---K-DEAWKQVEQLRRE--
GATQIAYRSDDWRDLKEAWKKGADILIVDA---TD-------KDEAWKQVEQLRREGATQI-A-
YRSDDWRDLKEAWKKGADILICDATGLE",
"targetStart": 19,
"templatePdbId": "1y0e_A",
"templateAlignment": "EPLHSSFIMSKMALAAYEGGAVGIRANT------
KEDILAIKETVDLPVIGIVKRDYDHSDVFITATSKEVDELIESQCEVIALDATLQQRPKETL
DELVSYIRTHAP NV-
EIMADIATVEEAKNAARLGFDYIGTTLHGYTSYTQGQLLYQNDFQFLKDVLQSVDAKVIAEG
NVITPDMY KRVMDLGVHCSVVGGAITR",
"templateStart": 17,
"weight": 0.6070577751720255
}

k. You may find it easier to visualize the alignment files by running each alignment in Clustal Omega. The two structures with a weight of 0.6 are the most important to look at. If these structures don't have complete sequence coverage look at the two with a weight at 0.3, and so on.

2) Load and optimize structures into CAD

a. Go to https://cad.cyrusbio.com/
b. Create a new Project.
c. Click 

, and either enter your PDB ID in 

to load the structure directly from the PDB or click 

to load a structure from your computer. The structure will be loaded into the collections section in a new folder.
d. Open the folder and click the box for the structure in the center window 

. This allows you to perform an action on that structure.
e. Click 

in the list of actions on the right window. Click 

in the center window.
    i. This will do a model optimization that will relieve any clashes or non-preferred orientations of side chains.
f. Select the collection for the finished prepared structure and click the box for that structure in the center window. Click 

, then in the center window increase 

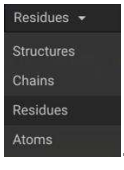
and click 

.
    i. This will run 1,000 Relaxes. These will optimize the structure in terms of Rosetta energy, but keep the structure close to the initial position. This is necessary to sample more conformations to find the global minima.
g. Once the Relaxes are done, look at the results. Click the Relax folder on the left to bring the structures into the center window. Click 

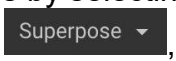
to rank them by score.

   The lowest energy model can be viewed by clicking the eye 

. Click 

to select the original model in order to see what kind of changes were made by Relax.
h. The structure viewer is in the bottom center window.
    i. Slide the top sections out of the way by click and drag of 
.
    ii. Clicking 
will bring the sequences into view.
    iii. By default, you can click on individual residues, but this can be changed by clicking 

.
    iv. Select or deselect all residues by clicking the structure name left of the sequence. Selecting several in a row is done with the shift-click. Selecting noncontiguous residues is done by command-click (Mac) or control-click (PC).
    v. Once you have selected residues, you can change their appearance by clicking 
to show backbone cartoon, 
to show sticks, and 
to show spheres.

vi. Change color by clicking  and you can color by element, secondary structure, polarity, or choose a color.
vii. If you highlight a section or all of the protein, you can choose a portion of that selection based on features such as selecting polar, + or – charge, hydrophobic. Or you can keep just cysteines, backbone or sidechain by clicking  .
viii. You can make an inverse selection by clicking  .
ix. Hide your selection from view by clicking  or bring something into view by clicking  .
x. Align multiple structures by selecting residues you want to base the alignment on and click  , which lets you align my selected atoms or backbone. This can be challenging when structures aren't identical in terms of sequence so use the most homologous region for the alignment.
xi. Click  to bring selection into the center of structure viewer.
xii. Enter full screen mode by clicking  . Press Esc to exit.
xiii. By selecting multiple sequences and clicking  , you can highlight the differences.
xiv. Clicking  will remind you what the key shortcuts do.
xv. Once you have created Selectors, you can click one to cause them to be selected in the structure viewer.

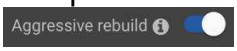
i. Select most energetically favorable structure that has the desired structural features for design. Potentially need to run iterative Relaxes using the best structure from each Relax run as a starting structure for more Relaxes.
j. Choose the most energetically favorable structure and click  . This will find the local minimum for this structure. There is no need for repeats because it is deterministic.

3) Loop Rebuild

a. Select a structure you would like to run loop rebuild on by clicking the  next to the structure name.
Then click 

.

b. Create a selector that includes a continuous loop region. We do not recommend that it be larger than 12 residues. The smaller the section, the better the prediction.

c. Click 

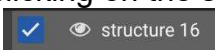
if you would like the original structure to be removed so that Ab Initio modeling will be done. This will be able to sample more search space. If you would like to limit the search space, leave aggressive build off and the modeling will start from the conformation of your model.

d. Enter a number into 

. The number will depend on the size of the loop. If doing 12 residues, bring it up to 500. Next click 

.

4) Run Design

a. Create a Selector by clicking on the structure, then clicking the eye to the left of the structure ID 

to bring it into the structure viewer.

b. Select all the residues that you want to allow to be mutated during design, then click 

to the left. By default, it will be named Selector X (X = the number, which increases with each Selector created).

c. Click  to change the name. Click  to change the selection residues. This will open a window in the center so you can add or change selections. Change the Min or Max to change the residue range that you want.

d. Select the starting structure you want for design by clicking on its collection folder on the left, then clicking the box by its ID in the center 
.

e. Click on the design tool in the right window 

.
This will bring the action into the center window.
   i. Design will keep the backbone rigid.
   ii. Flex Design will allow backbone movement only near the region of design.
   iii. Relax Design allows full protein backbone movement.

f. Choose 

from the drop down window. This will open the mutation table.

Pos.	Original Residues	A	R	N	D	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V
53	N	A	R	N	D	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V

f. For each row, select which residues you will allow design to sample.

g. For conservative mutating, this chart will help you select residues that have the same features.

5) Running ddG.

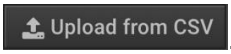
a. Select a folder on the right and the structure in the center window. Click the  .

b. Then you can click the 

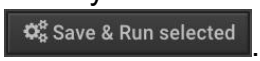
button on the right lower panel under Metrics to bring the run into the center window. By default, the selector will say 

, but you can switch to one of the selectors that you have created. The chart showing all original sequences and all potential point mutations is shown in the bottom center.

c. You can select (or deselect) all mutations at all location by clicking 

. You can also just click an entire row or column, as well as any individual box. Original sequences cannot be selected because the G is 0. You can upload your desired mutations by clicking 

.

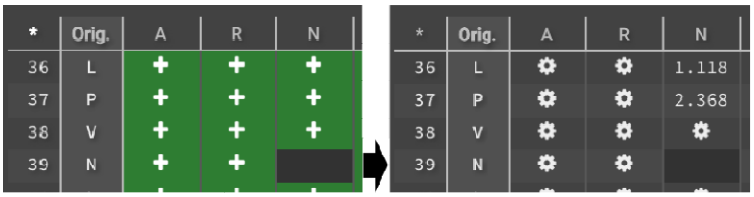
d. Once you have selected all the mutations that you want to measure, click 

.

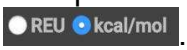
e. Before beginning the run, your selections show 

, will show 

while the calculation is running, and once it is done will show the ΔG



f. Negative ΔG indicates that the mutation will improve stability while positive ΔG will disrupt stability. You can download the entire chart by clicking 

. The default output in kcal/mol, you can switch to Rosetta Energy Units by clicking 

.

g. Use this to guide future design jobs

6) **Running design with PSSM.**
   a. PSSM = Position Specific Scoring Matrix is created by downloading the Blast database search tool along with the NCBI protein database, installing them locally, and then running Blast locally with specific options to generate the PSSM.
   b. Then you can use your sequence to BLAST NCBI. When you run Blast with the appropriate options, the resulting PSSM reflects the types of residue type variation seen at each position of your query among the identified homologs. This can be used to guide design in CAD.
   c. Run a design like described above, but after you select a Selector, click on 

to load the PSSM file. Then click OK to overwrite your mutation table. By default, it will select all the mutations that exist in the PSSM for each position because the threshold is set to 0 

Increasing the threshold will bias the mutations towards the most common mutations.

7) **Creating mutations that add or remove residues.**
   a. You cannot change the length of your sequence in CAD, but you can manually add or remove residues to a sequence and run it with HM, then load that into CAD to run design.
   b. Run HM as described above using your altered sequence. If adding residues, it is best to add a small residue that matches its environment. For an addition to the core, a small hydrophobic residue like Alanine is usually well tolerated. Adding something to the middle of a helix or a beta sheet is more challenging because it is likely to break the helix or beta sheet. Loops are easiest to add to and usually tolerant of Glycine.
   c. Run HM then run Design on the structures as suggested above.

8) **Predict regions that are T Cell epitope**
   a. Select a folder on the right and the structure in the center window. Click the  .
   b. Then you can click the Epitope Scan button on the right lower panel under Metrics to bring the run into the center window.
   c. The metric will immediately begin the calculation and could take a few minutes.
   d. Click structure name to bring the results table into view.