

Homology Modeling with YASARA

1) Search for desired protein on Uniprot.org

UniProtKB results

Filter by

- Reviewed (21)
- Unreviewed (187)

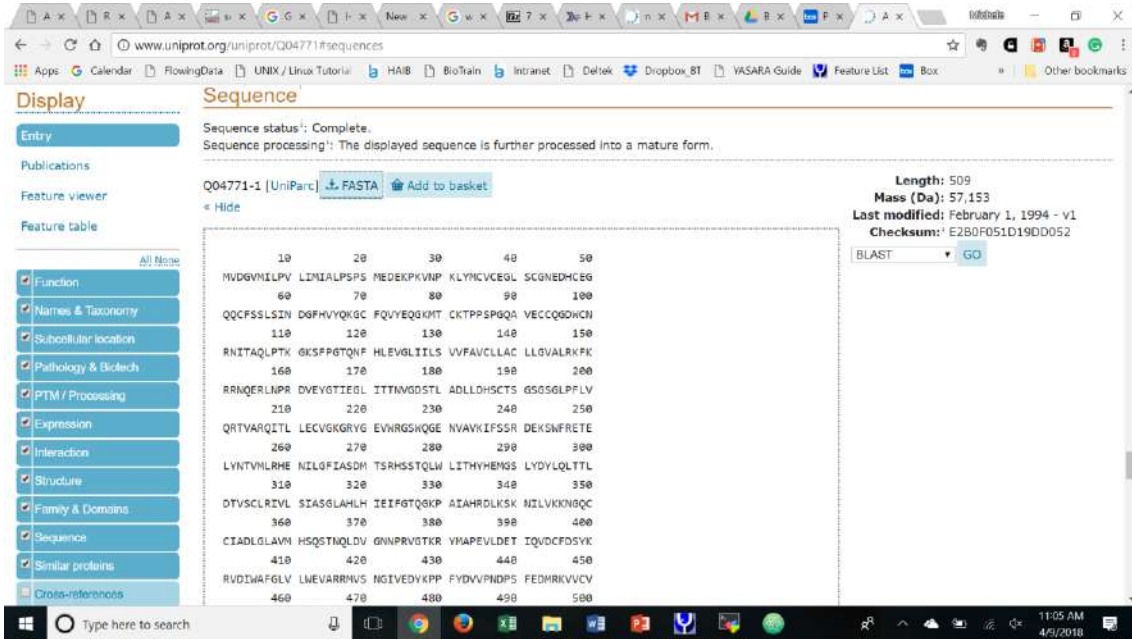
Popular organisms

- Human (18)
- Mouse (8)
- Rat (5)
- Bovine (3)
- Zebrafish (3)
- Other organisms

Entry	Entry name	Protein names	Gene names	Organism	Length
Q04771	ACVR1_HUMAN	Activin receptor type-1	ACVR1 ACVRLK2	Homo sapiens (Human)	509
P37172	ACVR1_MOUSE	Activin receptor type-1	Acvr1 Acvrik2, Tgfb1	Mus musculus (Mouse)	509
P80201	ACVR1_RAT	Activin receptor type-1	Acvr1 Acvrik2	Rattus norvegicus (Rat)	509
Q90ZK6	ACVR1_CHICK	Activin receptor type-1	ACVR1	Gallus gallus (Chicken)	504
Q28041	ACVR1_BOVIN	Activin receptor type-1	ACVR1 ACVRLK2	Bos taurus (Bovine)	509
A2VDM5	A2VDM5_BOVIN	Serine/threonine-protein kinase rec...	ACVR1	Bos taurus (Bovine)	509
E2RKJ3	E2RKJ3_CANLF	Serine/threonine-protein kinase rec...	ACVR1	Canis lupus familiaris (Dog) (Canis familiaris)	469

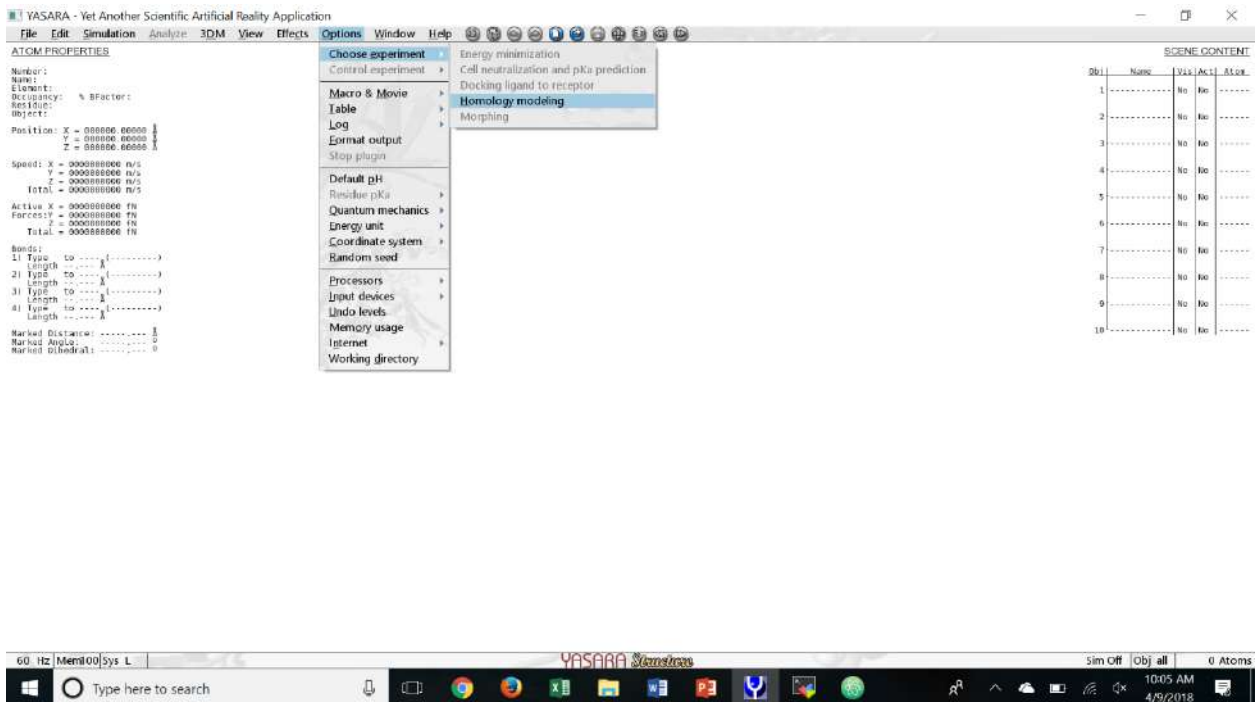
- a) Select species of choice
- b) Scroll down to sequence

c) Right click on Fasta, save to a folder named after your protein

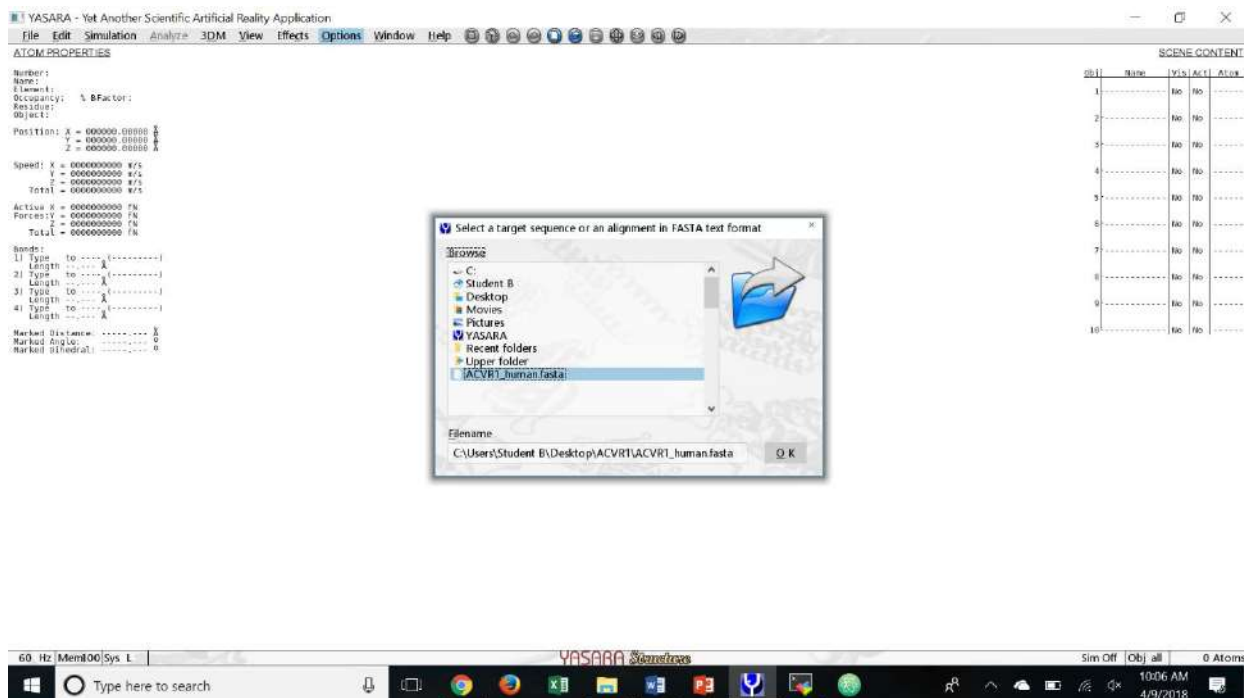


2) Open YASARA

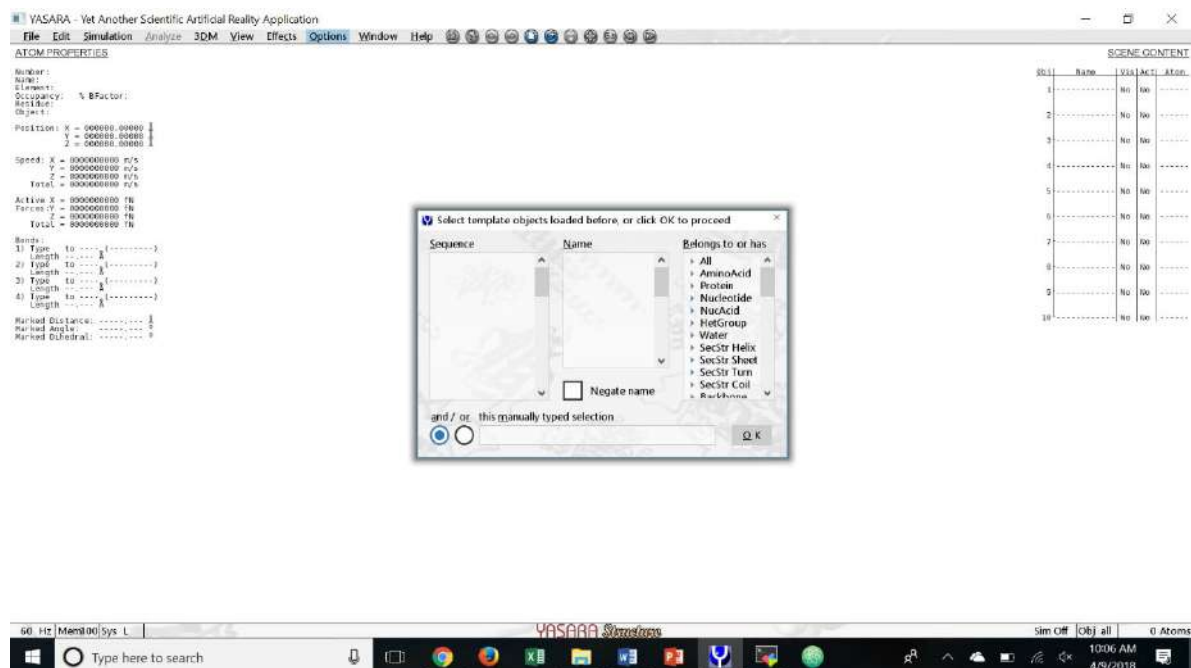
a) Click Options->Choose Experiment->Homology Modeling



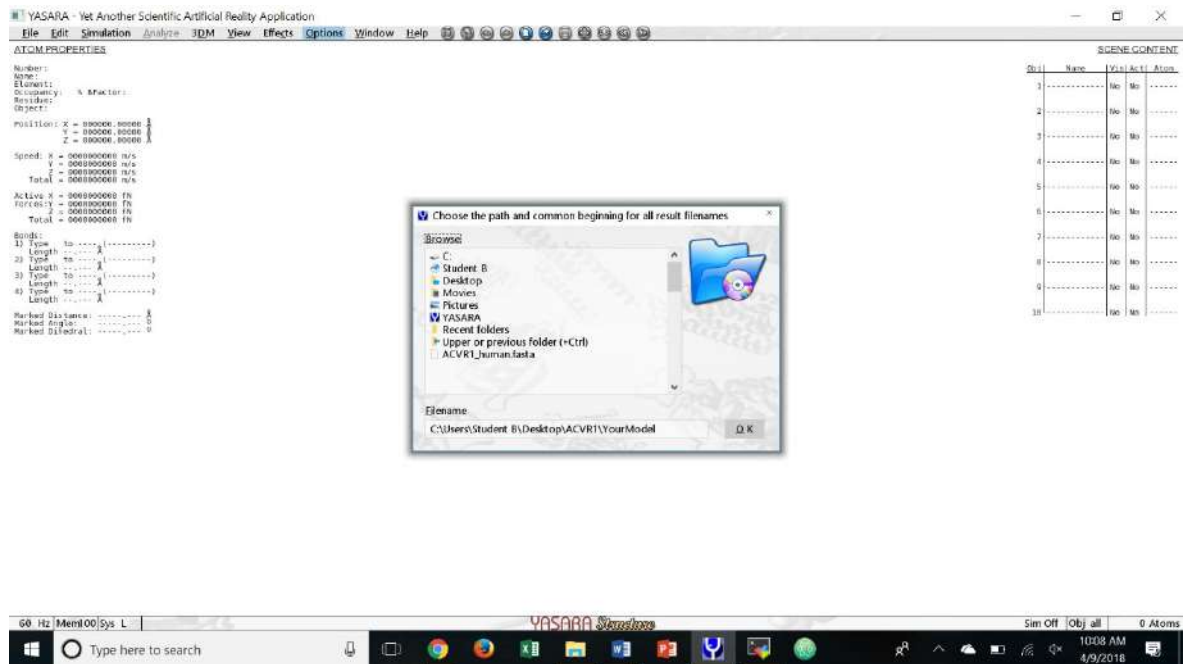
b) Find target sequence in File Browser, then click OK



c) If using a pre-modeled template select sequence in menu now, otherwise select OK



d) Designate the location that you want the output files to be saved to, this is typically the folder that you saved your.fasta sequence in.

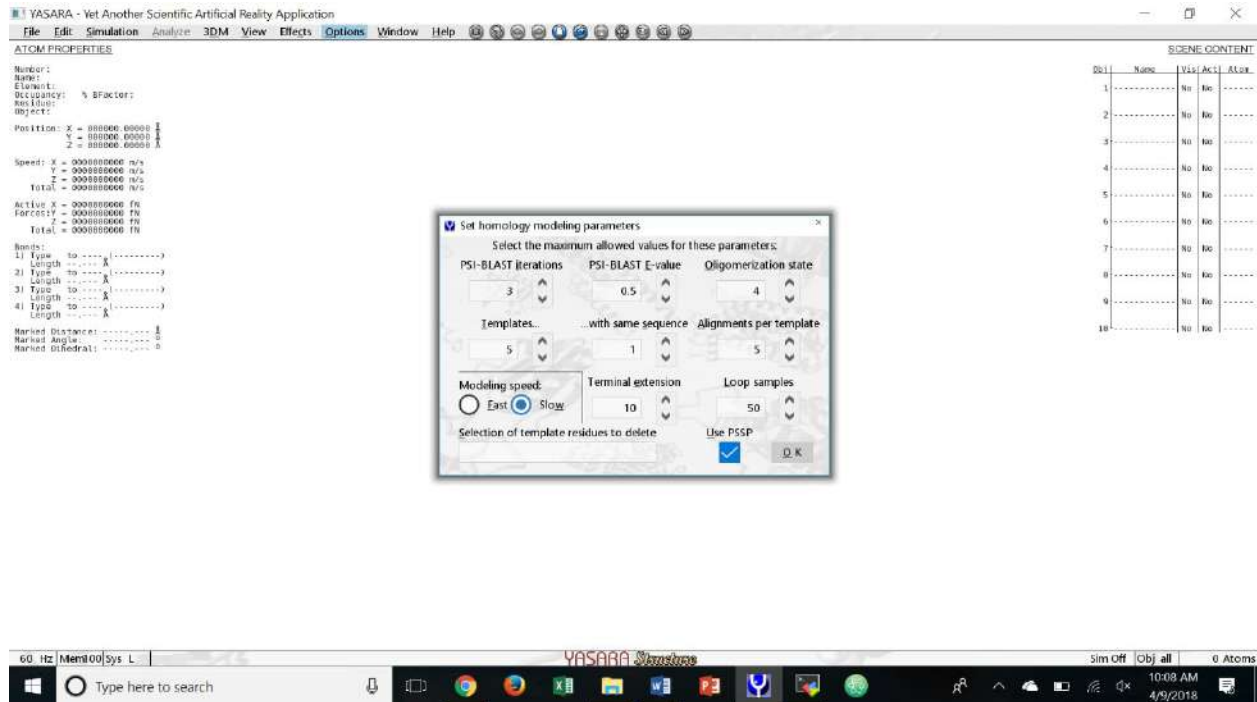


3) Modeling Parameters

a) For Standard Running Parameters:

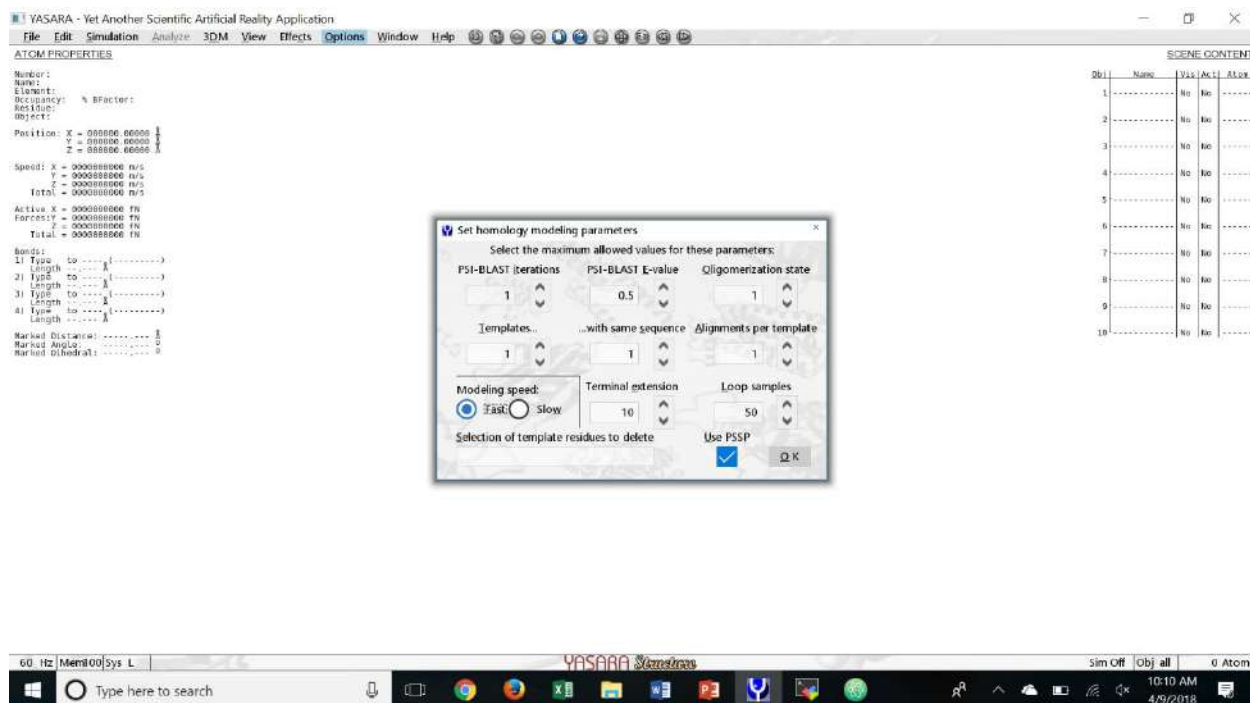
- i) Leave all options as is inside of the parameter box

- ii) A typical Standard (slow) Running Model can take around a full day to run. Occasionally longer or shorter depending on protein size and complexity.



b) For Fast Running Parameters:

- i) Change PSI-Blast iterations to 1
- ii) Change Oligomerization State to 1
- iii) Change Templates to 1
- iv) Change Alignments Per Template to 1
- v) Change Modeling Speed to Fast
- vi) Select OK and modeling will begin
- vii) A typical Fast Running Model will take anywhere from 10 minutes to a few hours depending on protein size and complexity



4) What to do when modeling is finished?

- a) If ran on **Standard Run Parameters**, a varying number of models will appear superposed upon each other. It may not let you toggle through each one.
 - i) First, go to the Chrome HTML Document that will be placed in the chosen output folder
 - ii) Upon opening this in your browser you will find an abundance of information regarding potential templates gathered, scoring parameters for models, as well as

7. The model ranking

The following table lists the 17 models sorted by their overall quality Z-scores. The models have been superposed and saved together as [YourModel.sce](#)

Rank	Z-score	Structure	State	Model ID	Filename	Original number	Residues	Comment
1	0.172		monomer	6EIX-A01	YourModel_6eix-a01.yob	9	192-503	Optimal
2	0.065		monomer	5E8S-A01	YourModel_5e8s-a01.yob	11	207-501	Optimal
3	0.029		monomer	5E8X-A01	YourModel_5e8x-a01.yob	14	203-503	Optimal
4	-0.007		monomer	6EIX-A02	YourModel_6eix-a02.yob	10	192-503	Good
5	-0.014		monomer	5E8S-A03	YourModel_5e8s-a03.yob	13	204-501	Good
6	-0.032		homodimer	3Q4U--01	YourModel_3q4u--01.yob	1	200-499	Good
7	-0.160		monomer	5E8X-A03	YourModel_5e8x-a03.yob	16	203-507	Good
8	-0.181		monomer	5E8X-A02	YourModel_5e8x-a02.yob	15	203-503	Good
9	-0.184		monomer	5E8X-A04	YourModel_5e8x-a04.yob	17	207-503	Good

visual representations of the top models. Scroll down and find the model with the highest overall Z-score, you can then rename to TopModel or your name of choice to indicate that it is the file that should be used on any future projects.

- b) What is a Z-Score? A Z-score describes how many standard deviations the model quality is away from the average high-resolution X-ray structure. Higher values are better, negative values indicate that the homology model looks worse than a high-resolution X-ray structure. The overall Z-scores for all models have been calculated as the weighted averages of the individual Z-scores using the formula $\text{Overall} = 0.145 \times \text{Dihedrals} + 0.390 \times \text{Packing1D} + 0.465 \times \text{Packing3D}$. The overall score thus captures the correctness of backbone-(Ramachandran plot) and side-chain dihedrals, as well as packing interactions. It applies to globular proteins only and can be misled by artificial structures like long single alpha helices (which have perfect dihedrals and are free of packing errors, since there is no packing).
- 5) If ran on Fast Run Parameters, you will see the exact number of models as you set for templates.
- a) Typically, templates are set to 1 so you would only see the top model
 - b) This saves you from having to analyze the Z-scores of multiple models.
 - i) However, it is a good idea to document the Z-score of your model to assure accuracy.
 - c) You will still need to save this model to a name of your choice to avoid confusion later on.