MODELING AND SIMULATION LAB REPORT

[Target Protein: Human 4 aminobutyrate]

SUBMITTED BY: ONEZA HASSAN ALVI

REGISTRATION NUMBER: 04281913032

SUMBITTED TO: SIR FAISAL AHMED

MODELING AND SIMULATION (BIF-213)

BSBIF NCB QAU

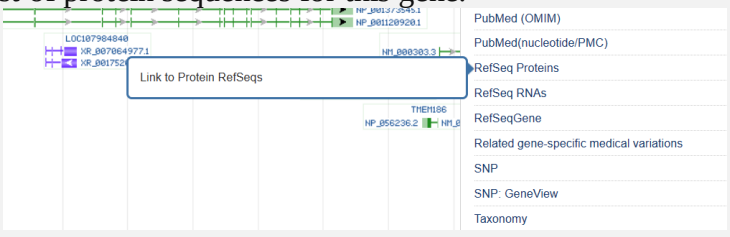
LAB REPORT:

- Download sequence file for Human 4 aminobutyrate
- Generate a homology model using SWISS-model server http://www.expasy.ch/swissmod/
- Download the template structure from www.rcsb.org
- Compare the model and template structures
- Validate the Structure

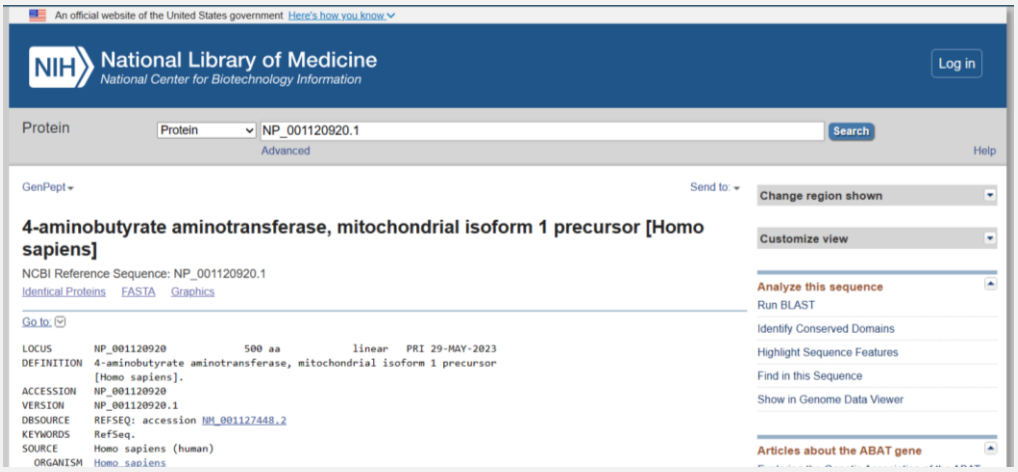
Download the protein sequence in FASTA format:

- Go to the NCBI Gene website and enter the gene ID 18 in the search box.
- Click on the gene name "ABAT 4-aminobutyrate aminotransferase [Homo sapiens (human)]" to go to the gene summary page.

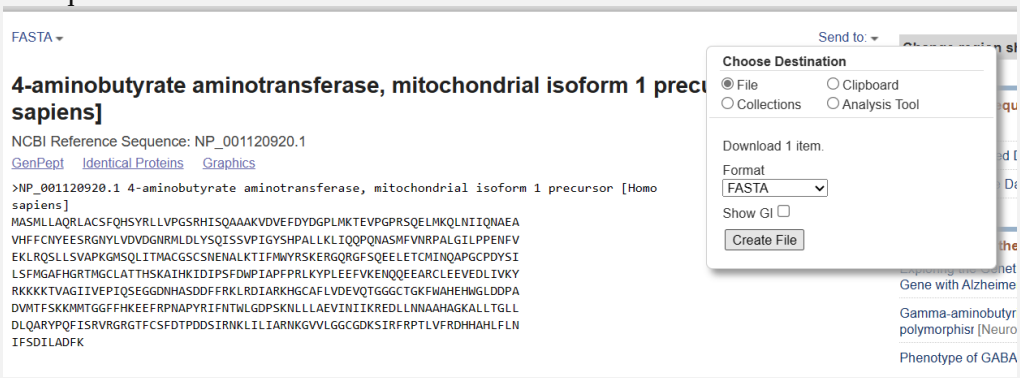
- Scroll down to the "Related information" section and click on the link "Protein (RefSeq)" under the "Gene" category. This will take you to the NCBI Protein website with a list of protein sequences for this gene.



- Click on the protein accession number "NP_001120920.1" to go to the protein summary page. This is the protein sequence that corresponds to the longest transcript variant of this gene.

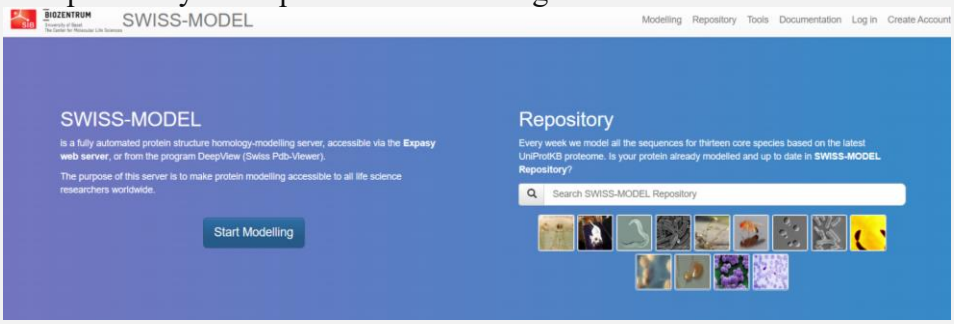


- Click on the "FASTA" link under the "Display Settings" section to download the protein sequence in FASTA format.

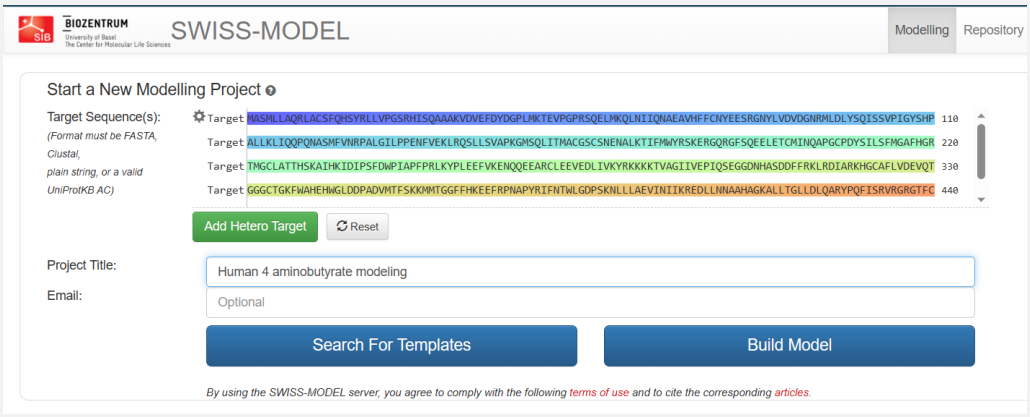


Generate a homology model using SWISS-model server:

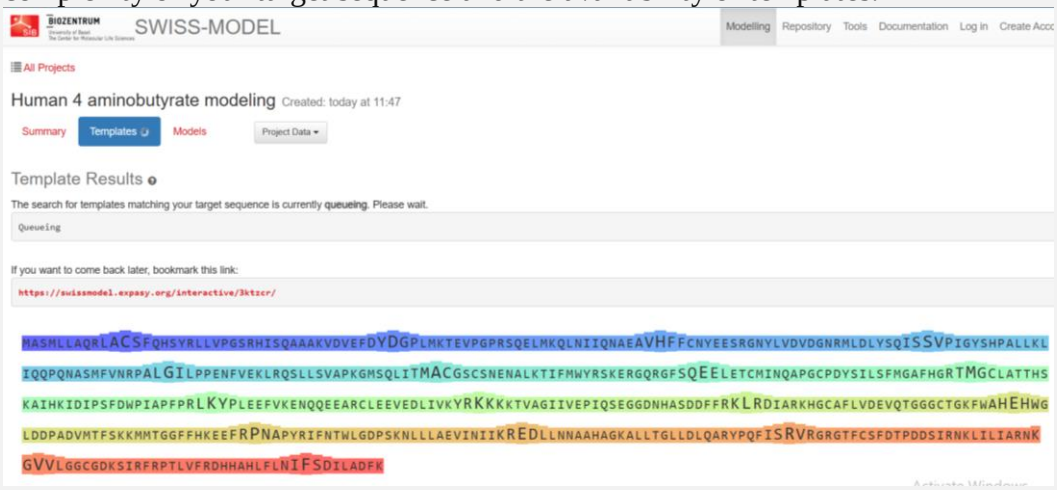
- Go to the SWISS-model website and click on the "Start Modelling" button or select the option "myWorkspace" from the navigation bar.



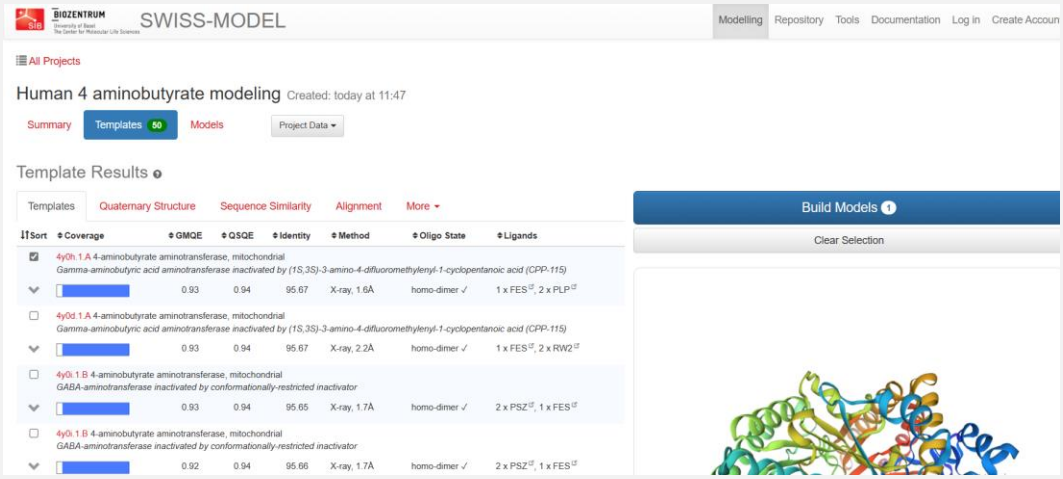
- Paste the protein sequence or provide the UniprotKB AC (P61851) of your target sequence in the input form. A project title is automatically suggested from the information retrieved from SWISS-MODEL's own copy of the UniprotKB annotation, if a UniProtKB AC is used.



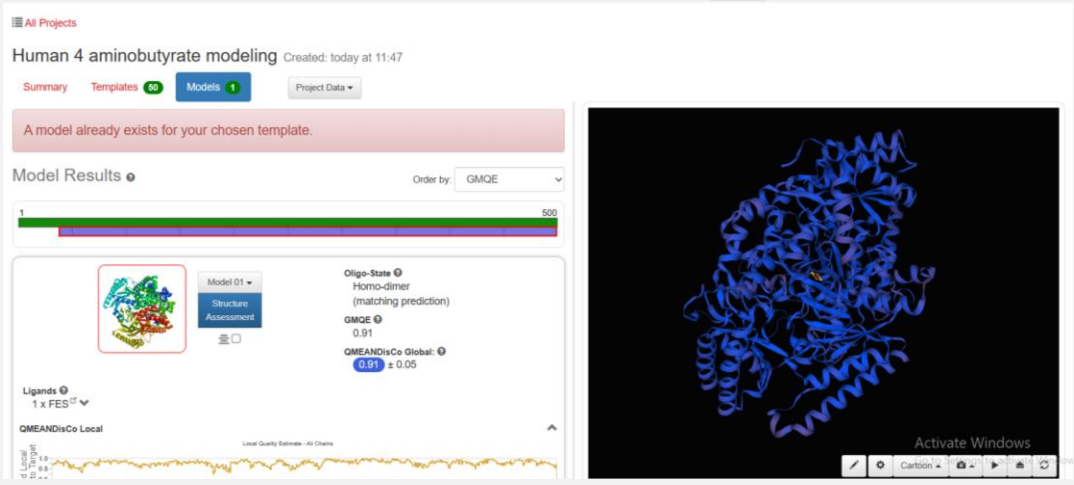
- Click on the “Search for Templates” button to find available template structures for your target sequence. The template search may take some time, depending on the complexity of your target sequence and the availability of templates.



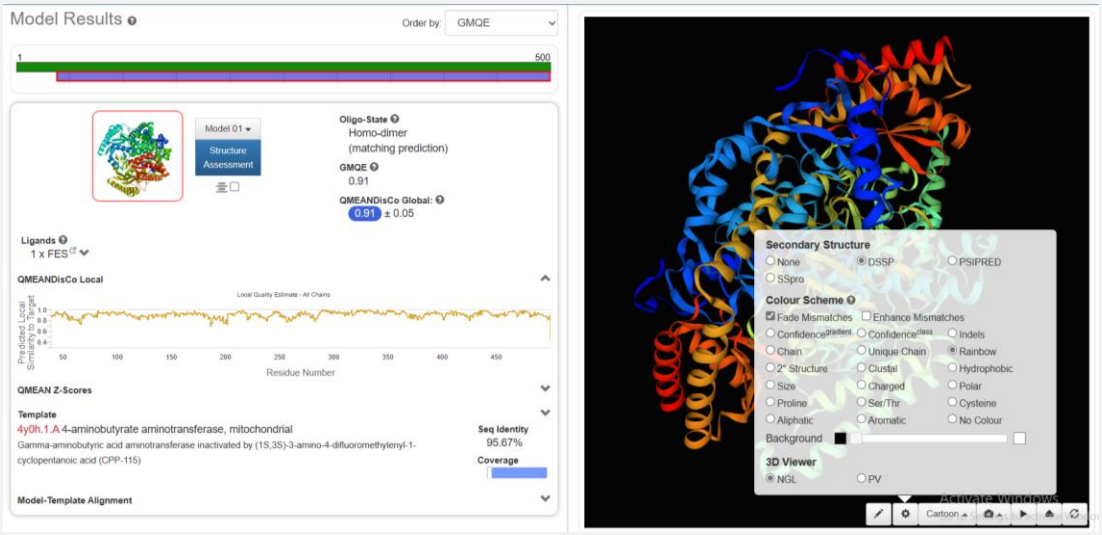
- When the template search is finished, you will see a table showing the list of available templates ranked according to the expected quality of the resulting models. You can select a template based on the GMQE (Global Model Quality Estimation), QSQE (Quaternary Structure Quality Estimation), sequence identity, and the presence of ligands or metal-binding sites. You can also visualize the template in 3D by clicking on the checkbox next to the template ID.

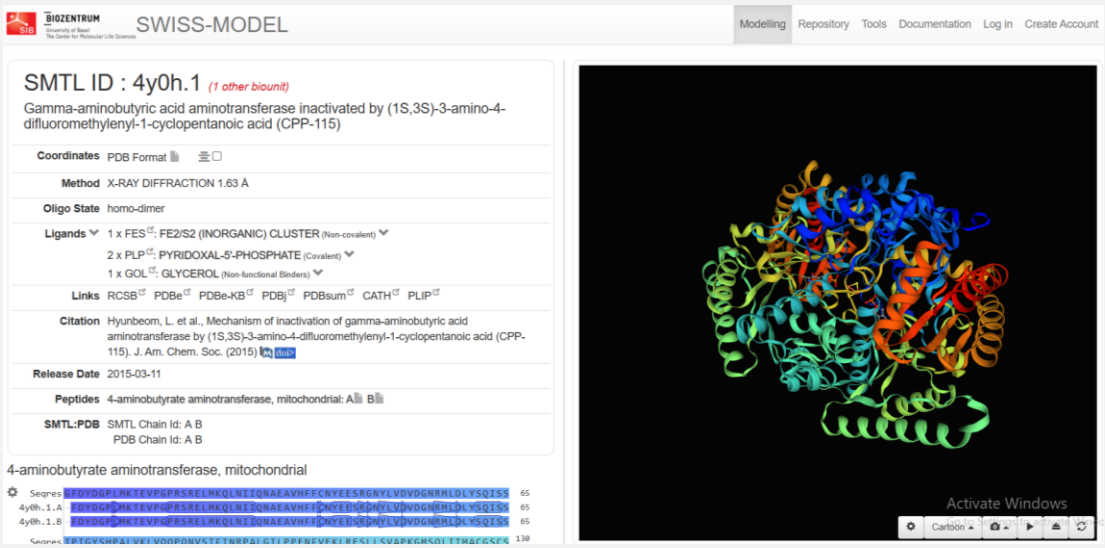


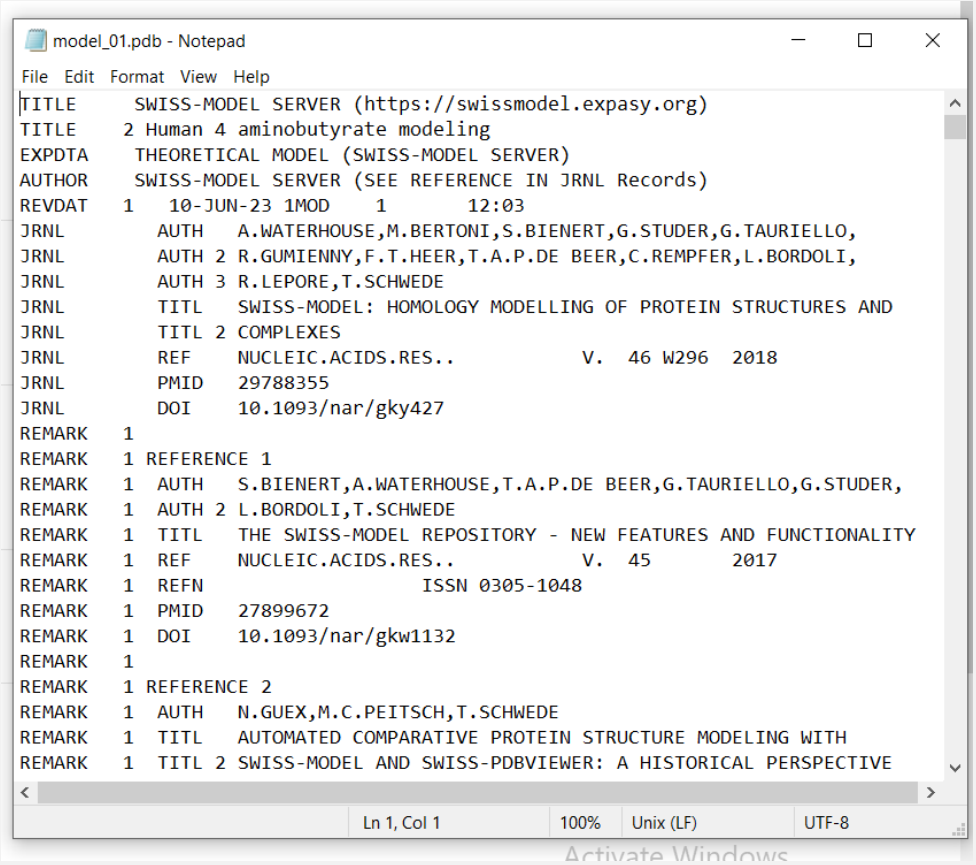
- After selecting a template, you can click on the "Build Model" button to start the modelling process. You can also adjust the target-template alignment if needed by clicking on the "Edit Alignment" button. The model building may take some time, depending on the size and complexity of your target sequence and template structure.



- When the model is built, you will see a summary of the modelling results, including a 3D visualization of your model, a quality assessment score (QMEAN), and a report of potential errors or issues in your model. You can download your model in PDB format by clicking on the "Download Model" button.


Model Results
Order by: GMQE
1
500
Model 01
Structure Assessment
Oligo-State
Homo-dimer
(matching prediction)
GMQE
0.91
QMEANDisCo Global:
0.91 ± 0.05
Ligands
1 x FES
QMEANDisCo Local
Local Quality Estimate - All Chains
Predicted Local Similarity to Target
Residue Number
QMEAN Z-Scores
Template
4y0h.1.A 4-aminobutyrate aminotransferase, mitochondrial
Gamma-aminobutyric acid aminotransferase inactivated by (1S,3S)-3-amino-4-difluoromethylenyl-1-cyclopentanoic acid (CPP-115)
Seq Identity
95.67%
Coverage
Model-Template Alignment
Secondary Structure
None
DSSP
PSIPRED
SSpro
Colour Scheme
Fade Mismatches
Enhance Mismatches
Confidence gradient
Confidence class
Indels
Chain
Unique Chain
Rainbow
2° Structure
Clustal
Hydrophobic
Size
Charged
Polar
Proline
Ser/Thr
Cysteine
Aliphatic
Aromatic
No Colour
Background
3D Viewer
NGL
PV
Cartoon

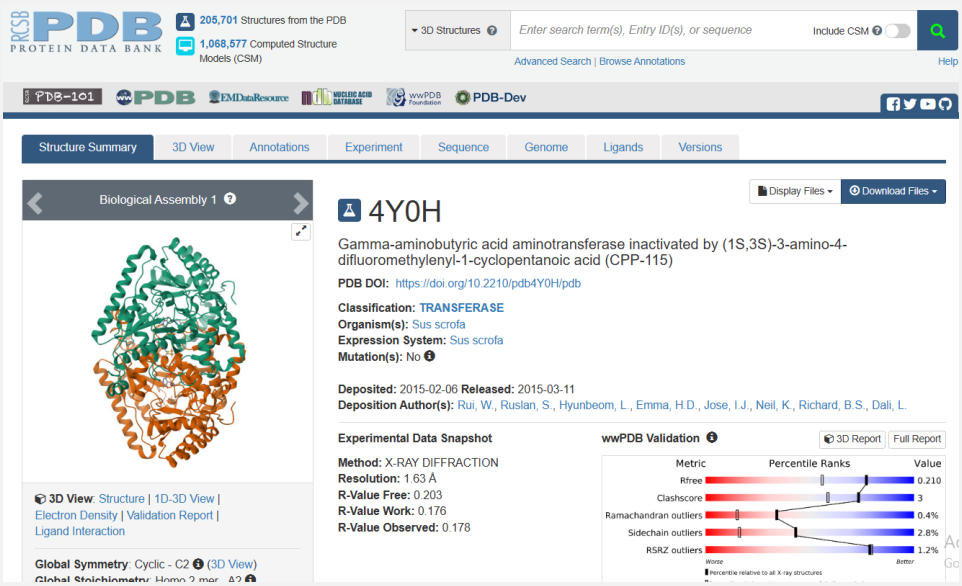

BIOZENTRUM
SWISS-MODEL
Modelling
Repository
Tools
Documentation
Log in
Create Account
SMTL ID : 4y0h.1 (1 other biounit)
Gamma-aminobutyric acid aminotransferase inactivated by (1S,3S)-3-amino-4-difluoromethylenyl-1-cyclopentanoic acid (CPP-115)
Coordinates PDB Format
Method X-RAY DIFFRACTION 1.63 Å
Oligo State homo-dimer
Ligands 1 x FES: FE2/S2 (INORGANIC) CLUSTER (Non-covalent)
2 x PLP: PYRIDOXAL-5'-PHOSPHATE (Covalent)
1 x GOL: GLYCEROL (Non-functional Binders)
Links RCSB PDBe PDBe-KB PDBj PDBsum CATH PLIP
Citation Hyunbeom, L. et al., Mechanism of inactivation of gamma-aminobutyric acid aminotransferase by (1S,3S)-3-amino-4-difluoromethylenyl-1-cyclopentanoic acid (CPP-115). J. Am. Chem. Soc. (2015)
Release Date 2015-03-11
Peptides 4-aminobutyrate aminotransferase, mitochondrial: A B
SMTL:PDB SMTL Chain Id: A B
PDB Chain Id: A B
4-aminobutyrate aminotransferase, mitochondrial
Activate Windows
Cartoon



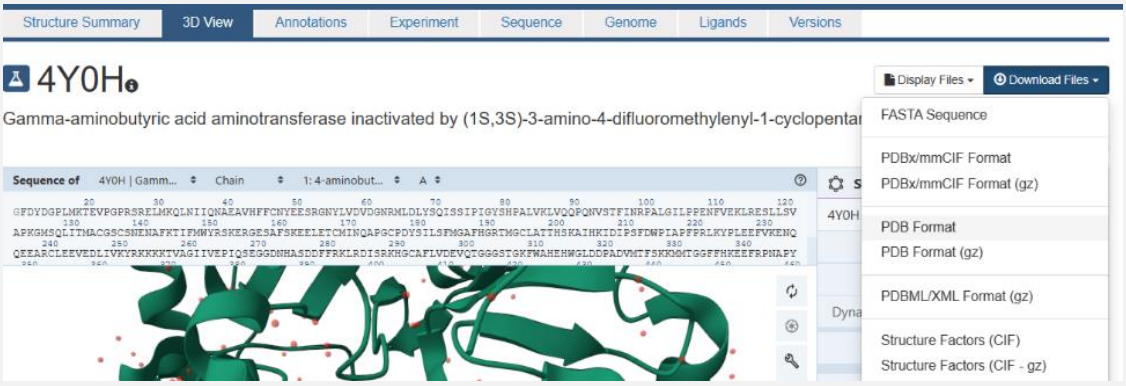
Download the template structure from www.rcsb.org:

- Go to www.rcsb.org website and enter PDG ID of template.

- You will see a summary page of the template structure, including its name, description, resolution, and 3D visualization. You can click on the “View in 3D” button to see the structure in more detail.

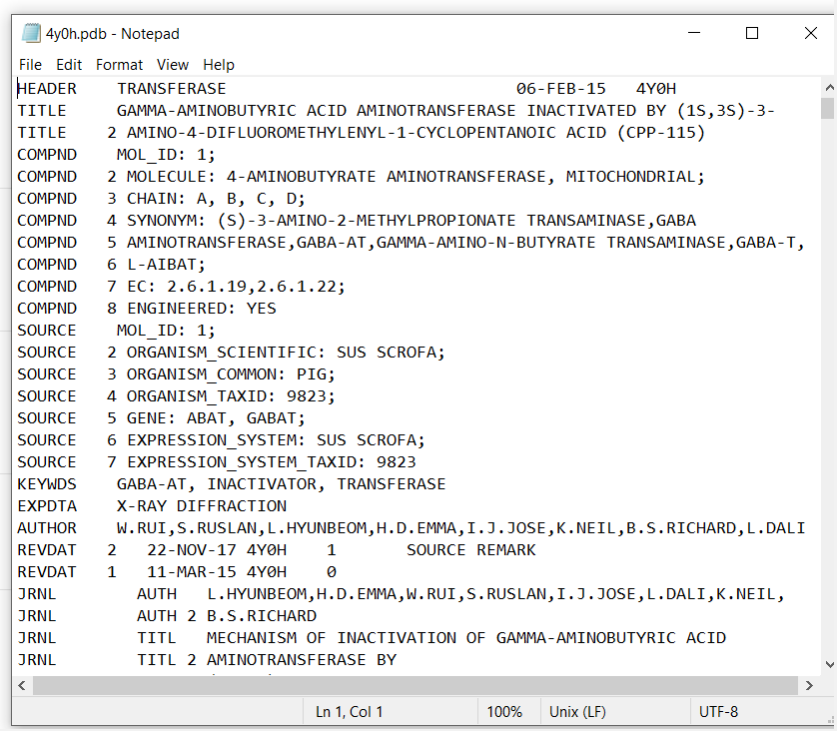


- To download the template structure in PDB format, click on the “Download Files” button on the top right corner of the page. A drop-down menu will appear with different



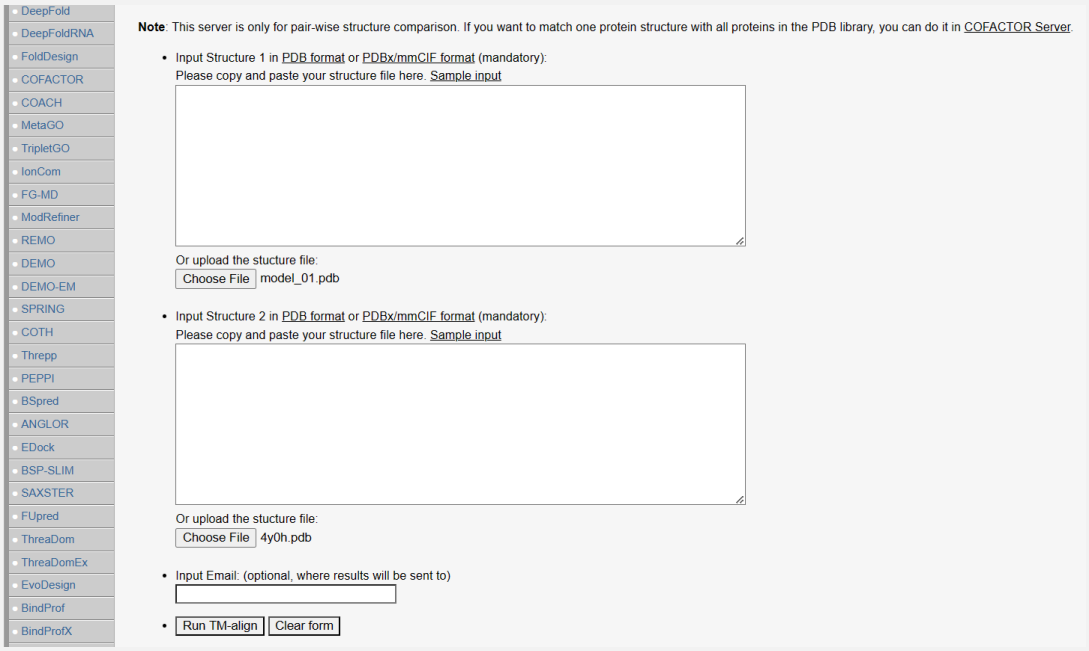
file formats and options. Select “PDB Format” to download the structure in PDB format. You can also choose other formats or options depending on your needs.

- A pop-up window will appear asking you to save the file to your computer. Choose a location and a name for your file and click save. You have now downloaded the template structure in PDB format.



Compare the model and template structures using TM-align or DALI:

- Go to the TM-align website or the DALI website (I chose TM-align).
- Click on the “Choose file” button and select your model file in PDB format from your computer. Repeat the same for your template file in PDB format.
- Click on the “RUN TM-align” button to start the alignment. The alignment may take a few seconds or minutes, depending on the size and complexity of your structures.

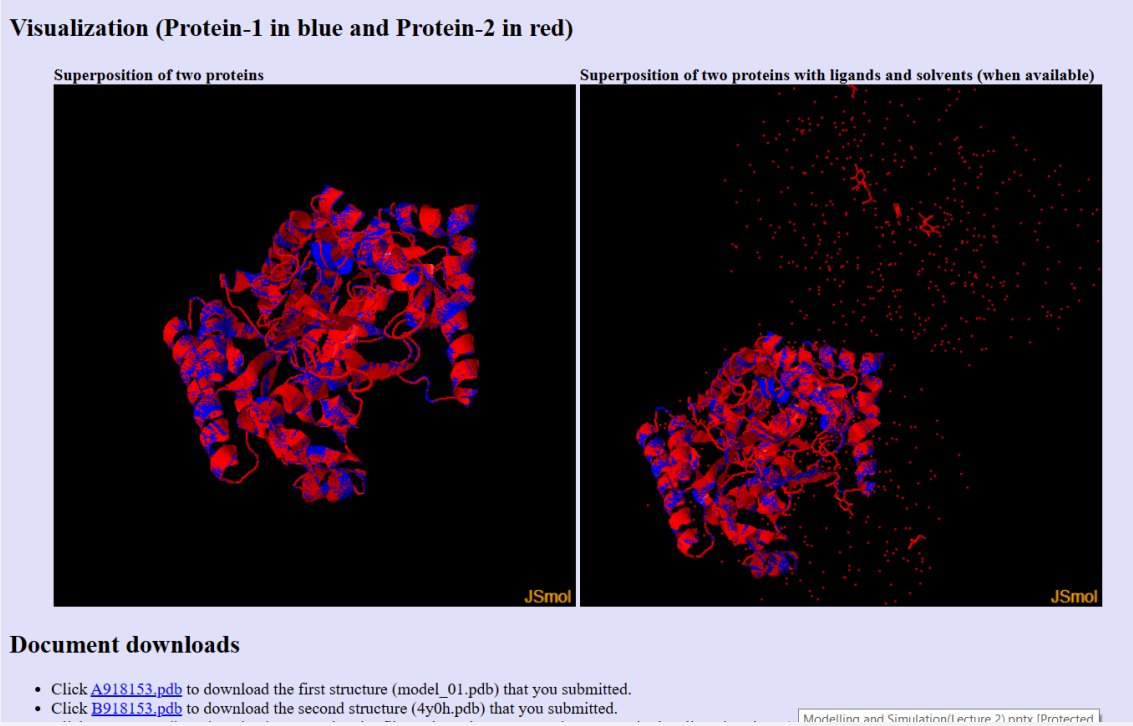


- When the alignment is finished, you will see a summary of the results, including a score and a visualization of the structural similarity between your model and template. The score is a measure of how well your structures match each other, based on their shape and orientation. A higher score means a better alignment and a more similar structure. The visualization shows your structures in different colors, superimposed on each other. You can rotate, zoom, or change the display options of the visualization using the buttons or mouse controls.

TM-align Results

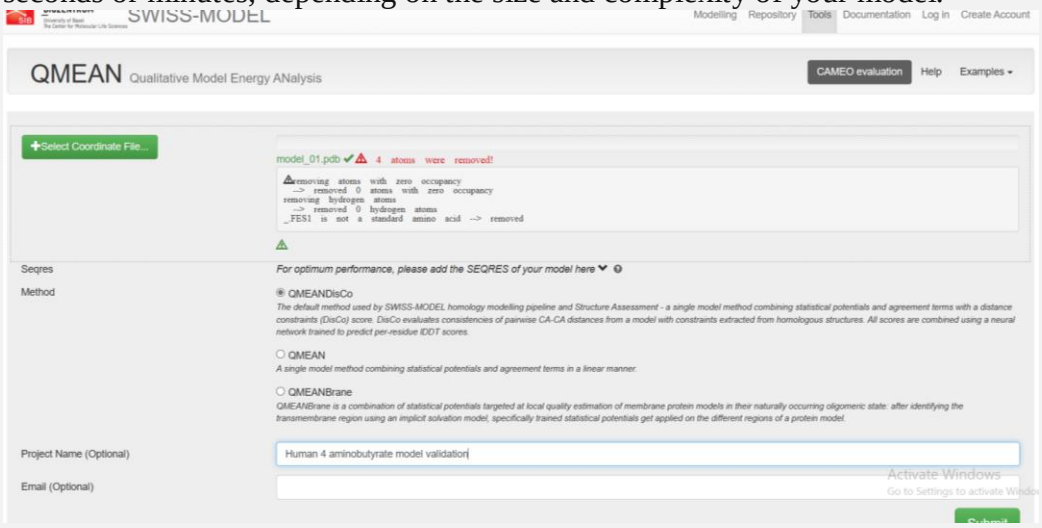
```
 **************************************************************************
 *                      TM-align (Version 20190822)                       *
 * An algorithm for protein structure alignment and comparison            *
 * Based on statistics:                                                   *
 *       0.0 < TM-score < 0.30, random structural similarity              *
 *       0.5 < TM-score < 1.00, in about the same fold                    *
 * Reference: Y Zhang and J Skolnick, Nucl Acids Res 33, 2302-9 (2005)    *
 * Please email your comments and suggestions to: zhng@umich.edu          *
 **************************************************************************

Name of Chain_1: A918153
Name of Chain_2: B918153
Length of Chain_1:  461 residues
Length of Chain_2:  461 residues

Aligned length=  461, RMSD=   0.07, Seq_ID=n_identical/n_aligned= 0.957
TM-score= 0.99992 (if normalized by length of Chain_1)
TM-score= 0.99992 (if normalized by length of Chain_2)
(You should use TM-score normalized by length of the reference protein)

(":" denotes aligned residue pairs of d < 5.0 A, "." denotes other aligned residues)
FDYDGPLMKTEVPGPRSQELMKQLNIIQNAEAVHFFCNYEESRGNYLVDVDGNRMLDLYSQISSVPIGYSHPALLKLIQQPQNASMFVNRPALG
::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
FDYDGPLMKTEVPGPRSRELMKQLNIIQNAEAVHFFCNYEESRGNYLVDVDGNRMLDLYSQISSIPIGYSHPALVKLVQQPQNVSTFINRPALG
```

- You can also download the aligned structures or the alignment file for further analysis.

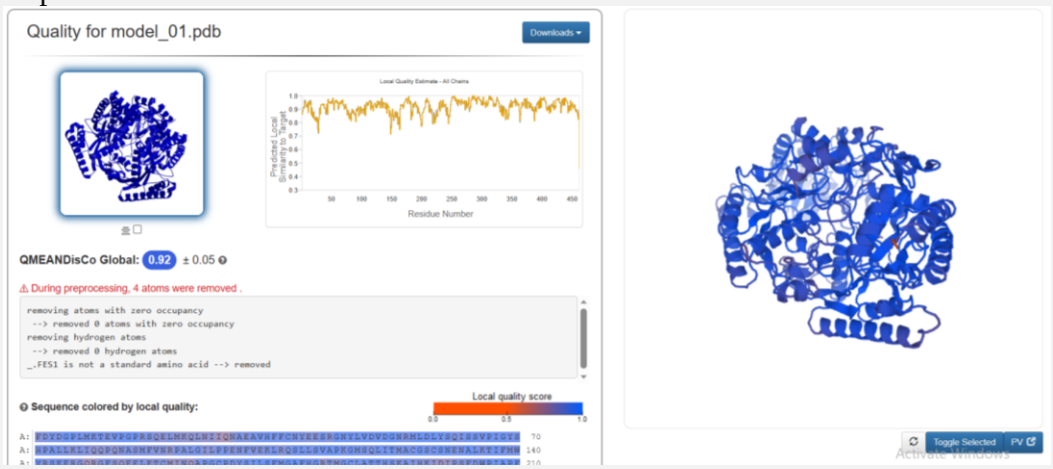


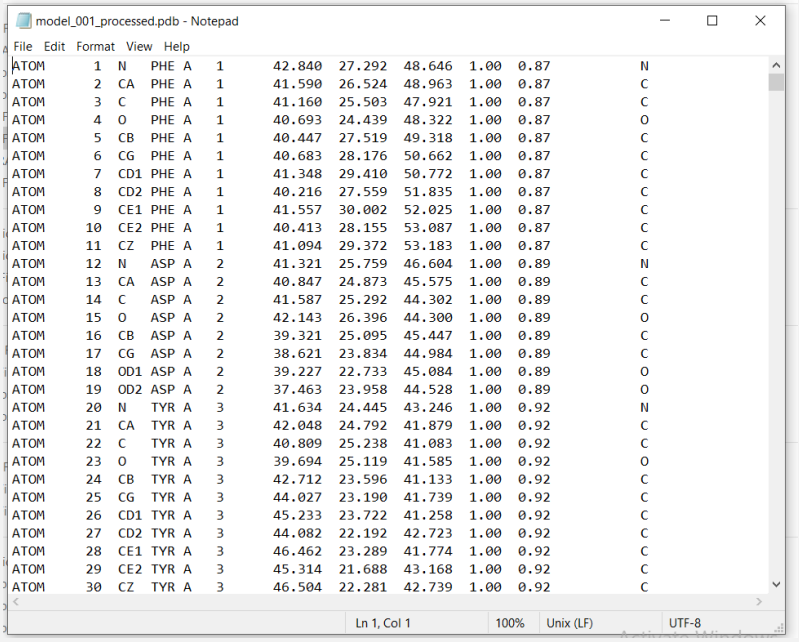
Validate the structure of your model:

- Go to the QMEAN website or the MolProbity website.
- Click on the "Browse" button and select your model file in PDB format from your computer.
- Click on the "Submit" button to start the analysis. The analysis may take a few seconds or minutes, depending on the size and complexity of your model.



- When the analysis is finished, you will see a summary of the results, including a score and a report of potential errors or issues in your model. The score is a measure of how well your model conforms to the expected stereochemistry and geometry of a protein structure, based on various criteria such as bond lengths, bond angles, torsion angles, clashes, and packing. A higher score means a better quality and a more reliable model. The report lists the specific errors or issues that were detected in your model, such as missing atoms, outliers, bad contacts, or Ramachandran violations. You can also download the annotated model or the report file for further inspection.



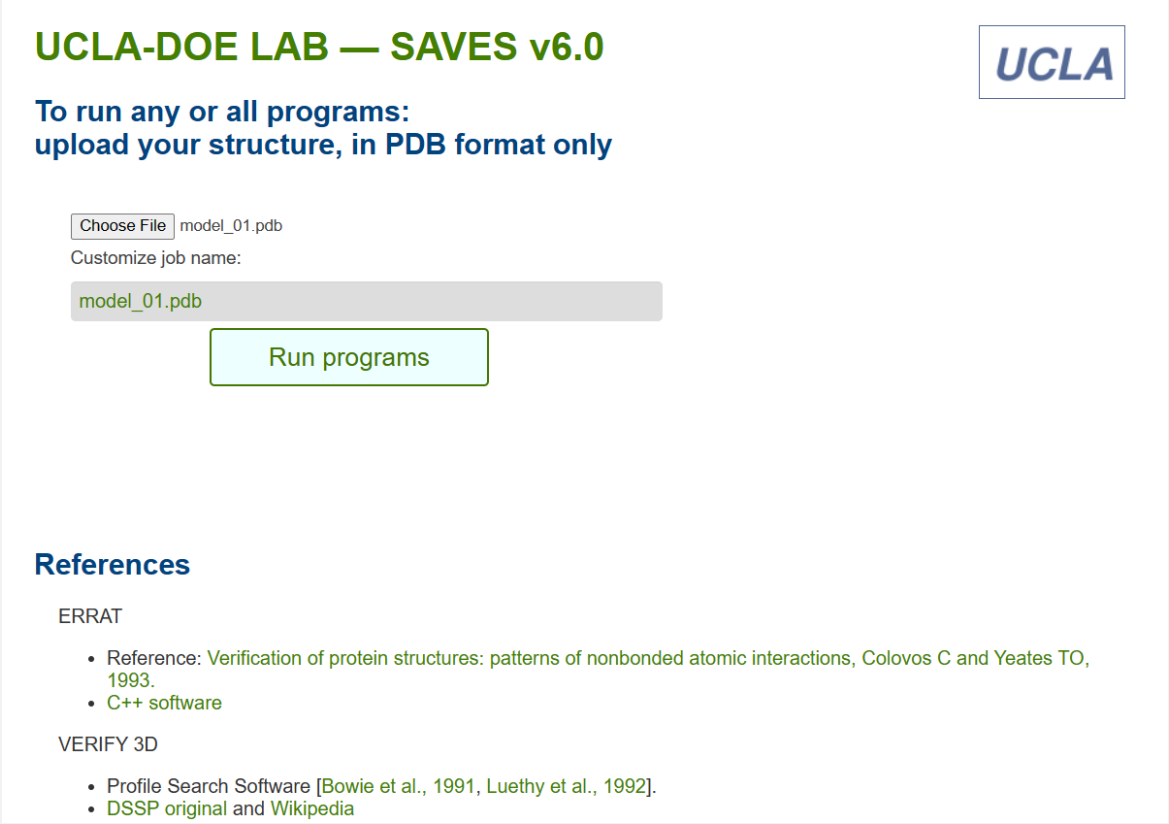


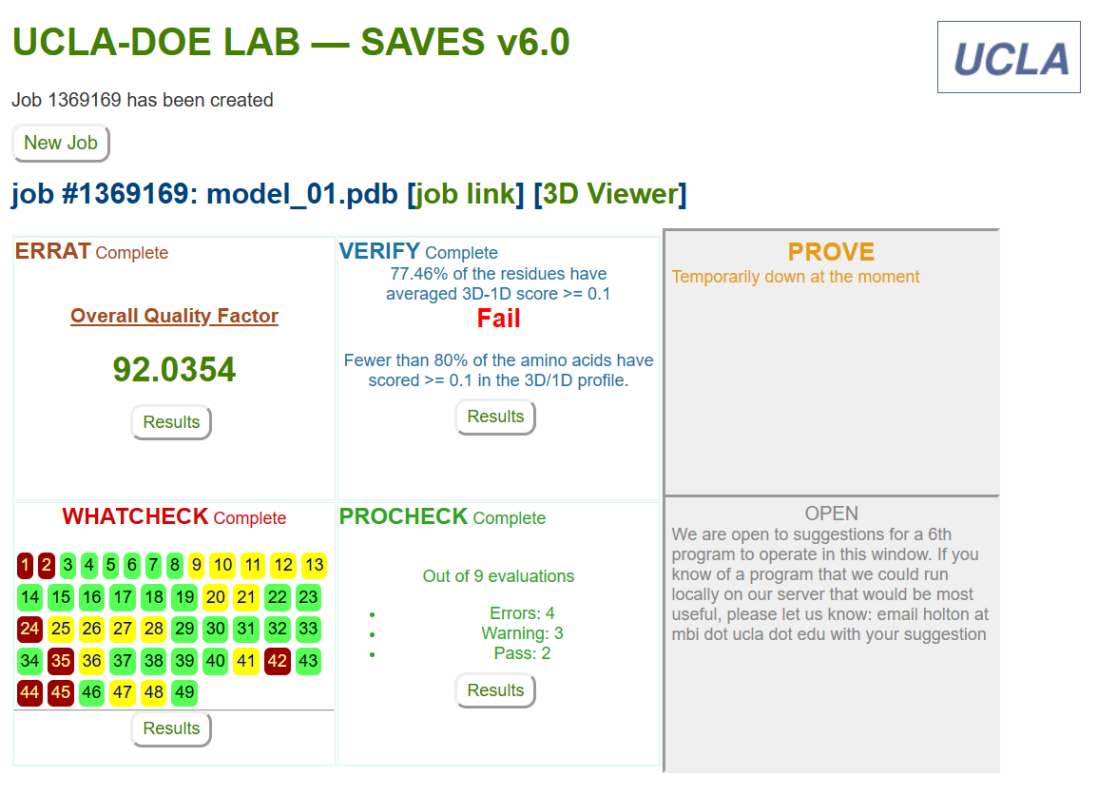
- QMEANDisCo Global is a score that measures the quality of a protein model based on various criteria, such as bond lengths, bond angles, torsion angles, clashes, and packing. It also takes into account the distance constraints derived from homologous structures in the Protein Data Bank. The score ranges from 0 to 1, where a higher score means a better quality and a more reliable model. The score is also normalized by the length of the model, to account for the different sizes of proteins. The value 0.92 ±

0.05 means that your model has a high quality and is very similar to the experimental structures of proteins of the same size. The value ± 0.05 means that there is a small uncertainty or variation in the score, which is normal for any prediction method. You can also look at the QMEANDisCo Local Quality Estimate and QMEAN Z-Scores to see how the quality of your model varies along the sequence and how it compares to other models.

Another method to validate the model: UCLA-DOE LAB --- SAVES v6.0

To validate your homology model structure using UCLA, you can use the SAVESv6.0 - Structure Validation Server[1]. This server allows you to run various programs to check the quality of your model, such as ERRAT, VERIFY 3D, PROVE, PROCHECK and WHATCHECK. To use this server, you need to upload your structure in PDB format and select the programs you want to run.





- **ERRAT**: The Overall Quality Factor is the percentage of the protein that falls below the 95% rejection limit for the error values. A higher value indicates a better quality model. This model has an Overall Quality Factor of 92.0354, which is quite good.
- **PROCHECK**: The program gives a summary of the results with Errors, Warnings, and Pass for each evaluation. Errors indicate serious problems, Warnings indicate potential problems, and Pass indicate acceptable results. This model has 4 Errors, 3 Warnings, and 2 Pass out of 9 evaluations, which suggests that your model has some issues that need to be fixed.
- **WHATCHECK**: The program gives a list of headings for each test result and highlights the problematic regions in red. You can select a test result box to see more details about the issues and how to fix them.
- **VERIFY3D**: The program also gives a percentage of residues that have a 3D-1D score above 0.1, which is considered a threshold for acceptable models. This model has 77.46% of residues above this threshold, which is slightly below the recommended value of 80%. This means that your model has some regions that are not well matched with their sequence.
