



Aug 10, 2021

Homology modeling with trRosetta

Chris Berndsen¹¹James Madison University

In Development

Share

dx.doi.org/10.17504/protocols.io.bw9nph5e

Berndsen Lab

Chris Berndsen
James Madison University

ABSTRACT

Protocol for homology modeling using trRosetta written for students in Biochemistry I at James Madison University.

DOI

dx.doi.org/10.17504/protocols.io.bw9nph5e

PROTOCOL CITATION

Chris Berndsen 2021. Homology modeling with trRosetta. **protocols.io**
<https://dx.doi.org/10.17504/protocols.io.bw9nph5e>

MANUSCRIPT CITATION please remember to cite the following publication along with this protocol

J Yang, I Anishchenko, H Park, Z Peng, S Ovchinnikov, D Baker, Improved protein structure prediction using predicted interresidue orientations, PNAS, 117: 1496-1503 (2020)

KEYWORDS

null, homology modeling, trRosetta, protein, prediction

LICENSE

————— This is an open access protocol distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

CREATED

Aug 10, 2021

LAST MODIFIED

Aug 10, 2021

PROTOCOL INTEGER ID

52238

MATERIALS TEXT

Internet access
[trRosetta server](#)

BEFORE STARTING

Have a sequence in FASTA format.

Ex.

>protein_seq

Setting up modeling

- 1 Navigate to the [trRosetta homepage](#).
- 2 Input your sequence in FASTA format.
The first list should start with > as shown in the example below.

• Provide the protein data (mandatory)

Input a protein sequence (Click for an example input) or a multiple sequence alignment (MSA) below:

```
>UFL1
HADAMEEIRRLAADFQRAQFAEATQRLSERNCIEIVNKLIAQKQLEVHTLDGKEYITPA
QTSKEHRELDHVRGGRVNIVDLQQVIHVDLIHENTIGDIKSEKHQVLGQLIDENYL
DRLAEENVIDLQESGQVTISELCKTYDLPGNFLTQALTQRLGRIISGHIDLNRGVIFTE
AFVARIKARIGLFSAITRPTAVNSISKYGFQEQLLYSVLEELVNSGLRGTVVGGRRQD
KAVFVPDIYRTQSTWVDSFFRNGYLEFDALSRGLPDVASYIKRYKTTQLFLKAAAC
VGGQLVDQVEASVEEATISSGTWVDIAPLPTSLSVEDAAILQQVWRAFQKASTVVFSD
TVVVEKFINDCITELFRELHMQKAEKKNPNVHLITEEDLKQISTLESVSTSKKKKDE
RRRKATEGSGSMRGGGGNAREYIKKVKKGRKDDSDDESQSHTGKKKPEISFHFQD
EIEDFLRKHIQDAPEEFISELAEYLKPLNKTYLEVRSVFSSTTSASGTGRKRTIKDL
QEEVSNLYNIRLFKGMKFFADDTQAALTKHLKSVCTDTNLTFNFLASDLMMVDDP
AATTSETRKKILSKLSEETKVALTKLHNSLNEKSIEDFISCLDSAAEACDIWVRGDKKR
ERQILFQHRQALAEQLKVTEDPALTLHLTSVLLFQFSTHSHLHAPGRCPQITAFNLNKI
PEDQHALLVKYQGLVVKQLVQSKKTKGQGYPLNNELDKEQEDVASTTRKELQELSSSIK
DLVLKSRKSSVTEE
```

Sequence format for trRosetta

- 3 Enter your email address and give a target name.

It is always a good idea to include target name or identifier in case you have multiple sequences that you are predicting.

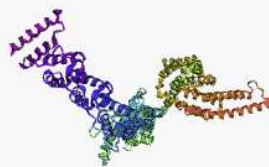
- 4 Adjust any options that you want and record these changes as a note on this step.
- 5 Press submit and wait for the job finish email. Typically takes 2 to 24 hours to get a result.

Modeling results

- 6 When you receive the job finished email, navigate to the results link.

Help guide for trRosetta: <https://yanglab.nankai.edu.cn/trRosetta/help/>

- 7 The top model will be shown in a window like below with an estimated TM score.



Download [Model 1](#) (colored by rainbow from N to C terminus)
Estimated TM-score: 0.472 ?

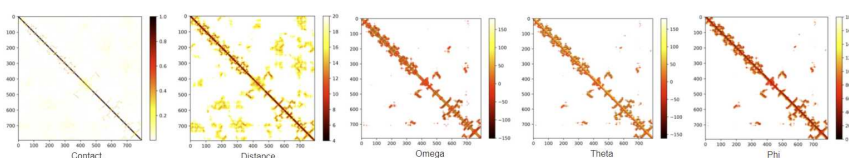
The TM-score is a template model score and measures the similarity between two proteins. 0 is no match, while 1 is a good match. Models with scores >0.5 are likely to be reliable.

- 8 The summary indicates how the model was built during the modeling along with links to the data that was used during modeling and the 4 "next best" models.

Summary of predicted models	
	• The confidence of the predicted model on the left is medium (with estimated TM-score=0.472).
	• The model was built by trRosetta based on de novo folding , guided by deep learning restraints.
	• You can download other lower-ranked models: model2 , model3 , model4 , model5 .
	• Download the multiple sequence alignment used by trRosetta (with 345 homologous sequences from the uniclust30_2018_08 database).
	• Download the predicted inter-residue distance and orientations .

Summary of modeling. Blue text indicates links to data.

- 9 The predicted 2D information shows the predicted long range interactions within the structure along with information on the dihedral angles.



2-D prediction heatmaps

This information is good for quickly identifying interesting inter-domain interactions or potential sites of regulation.

- 10 The predicted 1-D information shows the secondary structure prediction and predicted regions of disorder.



Secondary structure prediction

Saving data

- 11 At the top of the window you can download the entire file in .tar.bz2 format. To open this file it must be decompressed.

Alternatively, you can download the relevant models and data individually from the summary table.

Save files as

[date]_[proteinname]_[teamname]_trRosetta

Replace [proteinname] with the target protein name, [teamname] with your name or your team's name, [date] with the date.

Ex.: 20210810_UFL1_Berndsen_trRosetta

- 12 Indicate WHERE you saved the file as a note on this step.