[Home] [Server] [Queue] [About] [Remove] [Statistics]

I-TASSER results for job id S536199

(Click on S536199_results.tar.bz2 to download the tarball file including all modeling results listed on this page. Click on Annotation of I-TASSER Output to read the instructions for how to interpret the results on this page. Model results are kept on the server for 60 days, there is no way to retrieve the modeling data older than 2 months)

Submitted Sequence in FASTA format

```
>protein
MSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQY
GSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS
FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSA
LLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQD
SLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRL
ITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSA
PHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIIT
TDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASV
VNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC
MTSCCSCLKGCC
```

Predicted Secondary Structure

```
                20        40        60        80        100       120
                |         |         |         |         |         |
Sequence   MSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIED
Prediction CCCCCCCCCCSSSCCSSCCCCSSSSSSSSSSSSCCCCSSSSCCSSSCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCSSSSSCCCCCCCCCCCHHHHHH
Conf.Score 977778886532434287158628877778998327856873440657998899999999887799889999988777653889999985031210244245687035540489988786027887
           H:Helix; S:Strand; C:Coil
```

Predicted Solvent Accessibility

```
                20        40        60        80        100       120
                |         |         |         |         |         |
Sequence   MSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIED
Prediction 6523444221122230301441102013110214153130201200033252034003400310330240033013322330240132135214133143133020110013464352100000
           Values range from 0 (buried residue) to 9 (highly exposed residue)
```

Predicted normalized B-factor

(B-factor is a value to indicate the extent of the inherent thermal mobility of residues/atoms in proteins. In I-TASSER, this value is deduced from threading template proteins from the PDB in combination with the sequence profiles derived from sequence databases. The reported B-factor profile in the figure below corresponds to the normalized B-factor of the target protein, defined by B= (B'-u)/s, where B' is the raw B-factor value, u and s are respectively the mean and standard deviation of the raw B-factors along the sequence. Click here to read more about predicted normalized B-factor)

Top 10 threading templates used by I-TASSER
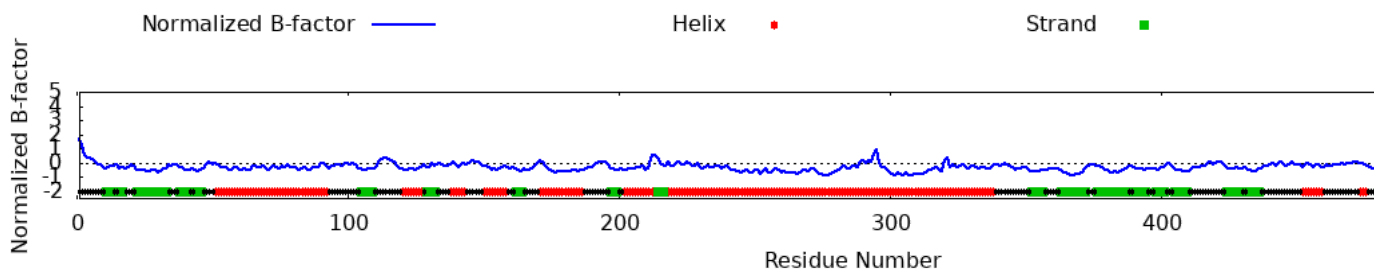
(I-TASSER modeling starts from the structure templates identified by LOMETS from the PDB library. LOMETS is a meta-server threading approach containing multiple threading programs, where each threading program can generate tens of thousands of template alignments. I-TASSER only uses the templates of the highest significance in the threading alignments, the significance of which are measured by the Z-score, i.e. the difference between the raw and average scores in the unit of standard deviation. The templates in this section are the 10 best templates selected from the LOMETS threading programs. Usually, one template of the highest Z-score is selected from each threading program, where the threading programs are sorted by the average performance in the large-scale benchmark test experiments.)

Rank	PDB Hit	Iden1	Iden2	Cov	Norm. Z-score	Download Align.					
								20	40	60	80
								|	|	|	|
						Sec.Str	CCCCCCCCCCSSSCCSSCCCCSSSSSSSSSSSSCCCCSSSSCCSSSCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCC				
						Seq	MSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPP				
1	6nb6A	0.88	0.69	0.79	2.68	Download	MSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPT-				
2	5szsA	0.37	0.31	0.81	5.64	Download	---SDNGISAIITANLSIPSNWTTSVQVEYLQITSTPIVVDCATYVCNGNPRCKNLLKQYTSACKTIEDALRLSAHLETNDVSSMLTFDSNAFSLAN				
3	6nzkA	0.42	0.35	0.79	3.58	Download	VNDSLEPVGGL--YEIQIPSEFTIGNMVEFIQTSSPKVTIDCAAFVCGDYAACKSQLVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLSG-				
4	6jx7	0.35	0.34	0.83	2.71	Download	VEIVDDSVKPVSTGNITIPKNFTVAVQAEYVQIQVKPVAVDCAKYVCNGNRHCLNLLTQYTSACQTIENSLNLGARLESLMLNDMITVSDRSLEFAT				
5	6jx7	0.35	0.34	0.83	2.00	Download	VDDSVGVIKPVSTGNITIPKNFTVAVQAEYVQIQVKPVAVDCAKYVCNGNRHCLNLLTQYTSACQTIENSLNLGARLESLMLNDMITVSDRSLEFAT				
6	6nzkA	0.41	0.35	0.80	4.00	Download	NSVNDSLEPVGGLYEIQIPSEFTIGNMVEFIQTSSPKVTIDCAAFVCGDYAACKSQLVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLGV				

7	6jx7	0.36	0.34	0.82	3.00	Download	`------VIKPVSTGNITIPKNFTVAVQAEYVQIQVKPVAVDCAKYVCNGNRHCLNLLTQYTSACQTIENSLNLGARLESLMLNDMITVSDRSLEFAT`
8	6b7nA	0.40	0.33	0.77	9.39	Download	`LQNTRPSIVSLYDGEVEIPSAFSLSVQTEYLQVQAEQVIVDCPQYVCNGNSRCLQLLAQYTSACSNIEAALHSSAQLDSREIINMFKTSTQSLQLAN`
9	5x58A	0.89	0.64	0.72	3.84	Download	`-------SIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPT`
10	6nzkA	0.41	0.36	0.79	19.72	Download	`---NDSLEPVGGLYEIQIPSEFTIGNMVEFIQTSSPKVTIDCAAFVCGDYAACKSQLVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLST`

(a) All the residues are colored in black; however, those residues in template which are identical to the residue in the query sequence are highlighted in color. Coloring scheme is based on the property of amino acids, where polar are brightly coloured while non-polar residues are colored in dark shade. (more about the colors used)

(b) Rank of templates represents the top ten threading templates used by I-TASSER.

(c) Ident1 is the percentage sequence identity of the templates in the threading aligned region with the query sequence.

(d) Ident2 is the percentage sequence identity of the whole template chains with query sequence.

(e) Cov represents the coverage of the threading alignment and is equal to the number of aligned residues divided by the length of query protein.

(f) Norm. Z-score is the normalized Z-score of the threading alignments. Alignment with a Normalized Z-score >1 mean a good alignment and vice versa.

(g) Download Align. provides the 3D structure of the aligned regions of the threading templates.

(h) The top 10 alignments reported above (in order of their ranking) are from the following threading programs:

    1: MUSTER   2: FFAS-3D   3: SPARKS-X   4: HHSEARCH2   5: HHSEARCH I   6: Neff-PPAS   7: HHSEARCH   8: pGenTHREADER   9: PROSPECT2   10: PRC

Top 5 final models predicted by I-TASSER

(For each target, I-TASSER simulations generate a large ensemble of structural conformations, called decoys. To select the final models, I-TASSER uses the SPICKER program to cluster all the decoys based on the pair-wise structure similarity, and reports up to five models which corresponds to the five largest structure clusters. The confidence of each model is quantitatively measured by C-score that is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. C-score is typically in the range of [-5, 2], where a C-score of a higher value signifies a model with a higher confidence and vice-versa. TM-score and RMSD are estimated based on C-score and protein length following the correlation observed between these qualities. Since the top 5 models are ranked by the cluster size, it is possible that the lower-rank models have a higher C-score in rare cases. Although the first model has a better quality in most cases, it is also possible that the lower-rank models have a better quality than the higher-rank models as seen in our benchmark tests. If the I-TASSER simulations converge, it is possible to have less than 5 clusters generated; this is usually an indication that the models have a good quality because of the converged simulations.)

- More about C-score
- Local structure accuracy profile of the top five models

(By right-click on the images, you can export image file or change the configurations, e.g. modifying the background color or stopping the spin of your models)

Reset to initial orientation   ☐ Spin On/Off   Reset to initial orientation   ☐ Spin On/Off   Reset to initial orientation   ☐ Spin On/Off
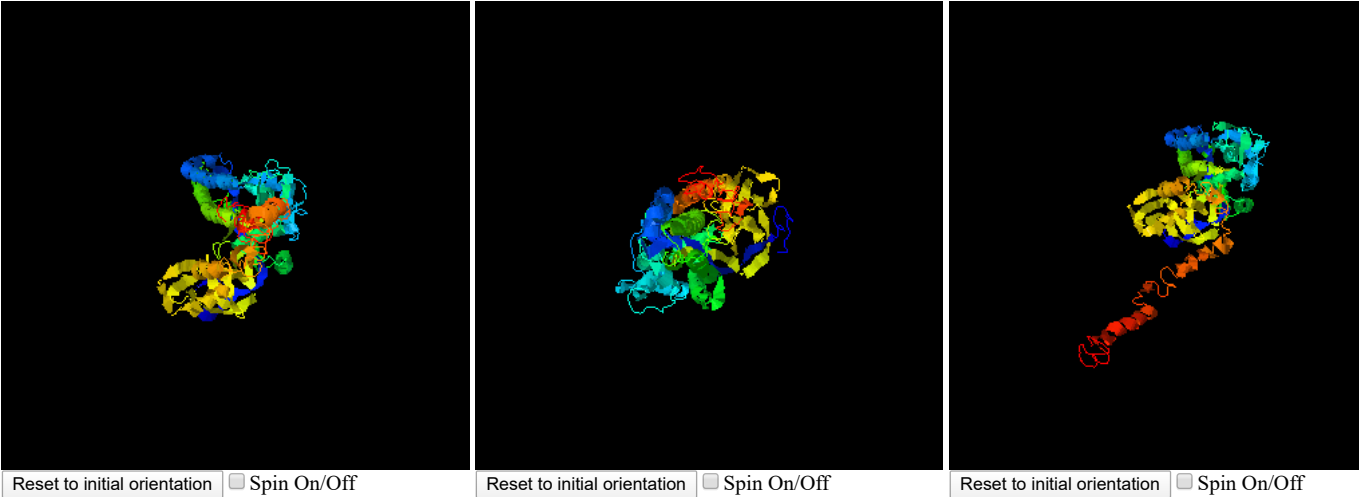
- Download Model 1
- C-score=-1.25 (Read more about C-score)
- Estimated TM-score = 0.56±0.15
- Estimated RMSD = 10.5±4.6Å

- Download Model 2
- C-score = -1.25

- Download Model 3
- C-score = -1.86

Proteins structurally close to the target in the PDB (as identified by TM-align)

(After the structure assembly simulation, I-TASSER uses the TM-align structural alignment program to match the first I-TASSER model to all structures in the PDB library. This section reports the top 10 proteins from the PDB that have the closest structural similarity, i.e. the highest TM-score, to the predicted I-TASSER model. Due to the structural similarity, these proteins often have similar function to the target. However, users are encouraged to use the data in the next section 'Predicted function using COACH' to infer the function of the target protein, since COACH has been extensively trained to derive biological functions from multi-source of sequence and structure features which has on average a higher accuracy than the function annotations derived only from the global structure comparison.)

Top 10 Identified stuctural analogs in PDB

Click to view	Rank	PDB Hit	TM-score[a]	RMSD[a]	IDEN[a]	Cov	Alignment
○	1	6jx7A	0.816	1.16	0.349	0.832	Download
○	2	5szsA	0.779	1.83	0.359	0.813	Download
○	3	6b7nA	0.775	1.66	0.379	0.802	Download
○	4	6u7kA	0.772	1.54	0.367	0.795	Download
○	5	6cv0A	0.710	3.05	0.359	0.783	Download
○	6	6nzkA	0.701	3.50	0.409	0.788	Download
○	7	6u7hA	0.697	3.31	0.341	0.775	Download
○	8	6nb3A2	0.694	3.41	0.393	0.783	Download
○	9	6nb6A	0.688	3.46	0.780	0.774	Download
○	10	5x59A	0.653	3.95	0.322	0.759	Download

(a) Query structure is shown in cartoon, while the structural analog is displayed using backbone trace.
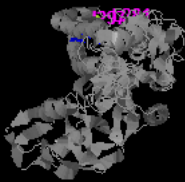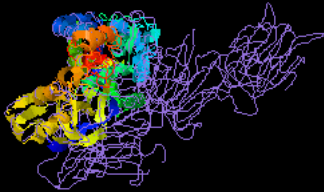
(b) Ranking of proteins is based on TM-score of the structural alignment between the query structure and known stru...

(c) RMSD^a is the RMSD between residues that are structurally aligned by TM-align.

(d) IDEN^a is the percentage sequence identity in the structurally aligned region.

(e) Cov represents the coverage of the alignment by TM-align and is equal to the number of structurally aligned res... query protein.

Pre...

(This s... d COACH based on the I-TASSER structure prediction. While COFACTOR deduces protein functions (ligand-binding... COACH is a meta-server approach that combines multiple function annotation results (on ligand-binding sites) from the CO...

| Click to view	Rank	C-score	Cluster size	PDB Hit	Lig Name	Download Complex	Ligand Binding Site Residues
◉	1	0.05	2	1vd5A	GLY	Rep, Mult	281,297
○	2	0.05	2	2eatA	CZA	Rep, Mult	40,71,74,75,259,262,311,312,313,318
○	3	0.05	2	4zgmA	32M	Rep, Mult	308,312
○	4	0.05	2	2h7hA	Nuc.Acid	Rep, Mult	298,299,302,306
○	5	0.02	1	4s1yA	CPT	Rep, Mult	78,362

Download the residue-specific ligand binding probability, which is estimated by SVM.

Download the all possible binding ligands and detailed prediction summary.

Download the templates clustering results.

(a) **C-score** is the confidence score of the prediction. C-score ranges [0-1], where a higher score indicates a mo...

(b) **Cluster size** is the total number of templates in a cluster.

(c) **Lig Name** is name of possible binding ligand. Click the name to view its information in the BioLiP database.

(d) **Rep** is a single complex structure with the most representative ligand in the cluster, i.e., the one listed in the l... **Mult** is the complex structures with all potential binding ligands in the cluster.

Enzyme Commission (EC) numbers and active sites
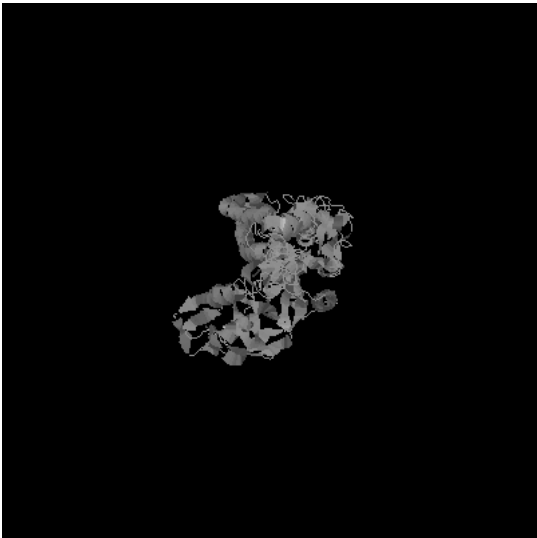
Click to view	Rank	Cscore^{EC}	PDB Hit	TM-score	RMSD^a	IDEN^a	Cov	EC Number	Active Site Residues
○	1	0.114	2vz9B	0.312	7.90	0.053	0.520	2.3.1.85	NA
○	2	0.114	2vz8B	0.310	7.93	0.041	0.518	2.3.1.85	NA
○	3	0.110	2vz8A	0.325	8.35	0.037	0.574	2.3.1.85	NA
○	4	0.109	2hg4C	0.307	8.18	0.028	0.534	2.3.1.94	NA
○	5	0.108	3b8eA	0.296	7.37	0.029	0.466	3.6.3.9	41

Click on the radio buttons to visualize predicted active site residues.

(a) Cscore^{EC} is the confidence score for the EC number prediction. Cscore^{EC} values range in between [0-1]; where a higher score indicates a more reliable EC number prediction.

(b) TM-score is a measure of global structural similarity between query and template protein.

(c) RMSD^a is the RMSD between residues that are structurally aligned by TM-align.

(d) IDEN^a is the percentage sequence identity in the structurally aligned region.

(e) Cov represents the coverage of global structural alignment and is equal to the number of structurally aligne... by length of the query protein.

Gene Ontology (GO) terms

Top 10 homologous GO templates in PDB

Rank	Cscore^{GO}	TM-score	RMSD^a	IDEN^a	Cov	PDB Hit	Associated GO Terms
1	0.17	0.2961	7.37	0.03	0.47	3b8eA	GO:0016323 GO:0002026 GO:0005886 GO:0006814 GO:0006813 GO:0016021 GO:0003869 GO:0006811 GO:0042383 G(GO:0008217 GO:0042470 GO:0045989 GO:0045822 GO:0000166 GO:0042493 GO:0016020 GO:0031947 GO:0006810 G(GO:0046872 GO:0005391 GO:0003824 GO:0006754 GO:0006812 GO:0008152 GO:0015077 GO:0015662 GO:0015672 G(
2	0.12	0.3416	6.60	0.06	0.51	2b9bB	GO:0006948
3	0.11	0.3116	7.90	0.05	0.52	2vz9B	GO:0016740 GO:0000166 GO:0000036 GO:0003824 GO:0004312 GO:0004313 GO:0004314 GO:0004315 GO:0004316 G(GO:0004320 GO:0005488 GO:0008152 GO:0008270 GO:0009058 GO:0016295 GO:0016296 GO:0016297 GO:0016491 G(GO:0031177 GO:0048037 GO:0055114

4	0.11	0.3097	7.93	0.04	0.52	2vz8B	GO:0016740 GO:0000166 GO:0000036 GO:0003824 GO:0004312 GO:0004313 GO:0004314 GO:0004315 GO:0004316 GO GO:0004320 GO:0005488 GO:0008152 GO:0008270 GO:0009058 GO:0016295 GO:0016296 GO:0016297 GO:0016491 GO GO:0031177 GO:0048037 GO:0055114
5	0.11	0.3149	8.29	0.04	0.55	3hhdC	GO:0003824 GO:0005488 GO:0008152 GO:0009058 GO:0016740
6	0.11	0.3128	7.89	0.04	0.53	2uvaG	GO:0003824 GO:0004312 GO:0005835 GO:0006633 GO:0008152 GO:0016491 GO:0016740 GO:0055114
7	0.11	0.3254	7.53	0.04	0.53	3feyA	GO:0034660 GO:0006379 GO:0006810 GO:0031047 GO:0051028 GO:0016032 GO:0006406 GO:0006417 GO:0006370 GO GO:0008334 GO:0006369 GO:0005488 GO:0005634 GO:0005737 GO:0006446 GO:0031124 GO:0003723 GO:0006366 GO GO:0030529 GO:0050434 GO:0005515 GO:0006368 GO:0005845 GO:0000387 GO:0000398 GO:0010467 GO:0000339 GO GO:0000184
8	0.11	0.3142	8.09	0.03	0.54	2qo3B	GO:0003824 GO:0005488 GO:0008152 GO:0009058 GO:0016740
9	0.11	0.3253	8.35	0.04	0.57	2vz8A	GO:0016740 GO:0000166 GO:0000036 GO:0003824 GO:0004312 GO:0004313 GO:0004314 GO:0004315 GO:0004316 GO GO:0004320 GO:0005488 GO:0008152 GO:0008270 GO:0009058 GO:0016295 GO:0016296 GO:0016297 GO:0016491 GO GO:0031177 GO:0048037 GO:0055114
10	0.11	0.3083	8.14	0.03	0.53	2hg4D	GO:0003824 GO:0005488 GO:0008152 GO:0009058 GO:0016740 GO:0048037

Consensus prediction of GO terms

Molecular Function	GO:0016418	GO:0016419	GO:0072341	GO:0016628	GO:0016297	GO:0019842	GO:0019171	GO:0032559	GO:0015081	GO:0035639
GO-Score	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.33	0.33	0.33
Biological Process	GO:0045933	GO:0009206	GO:0031946	GO:0090032	GO:0051241	GO:0051234	GO:0006942	GO:0042221	GO:0008016	GO:0031944
GO-Score	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
Cellular Component	GO:0031224	GO:0048770	GO:0042598	GO:0044459						
GO-Score	0.33	0.33	0.33	0.33						

(a) Cscore^{GO} is a combined measure for evaluating global and local similarity between query and template protein. It's range is [0-1] and higher values indicate more confident predictions.

(b) TM-score is a measure of global structural similarity between query and template protein.

(c) RMSD^a is the RMSD between residues that are structurally aligned by TM-align.

(d) IDEN^a is the percentage sequence identity in the structurally aligned region.

(e) Cov represents the coverage of global structural alignment and is equal to the number of structurally aligned residues divided by length of the query protein.

(f) The second table shows a consensus GO terms amongst the top scoring templates. The GO-Score associated with each prediction is defined as the average weight of the GO term, where the weights are as template.

[Click on S536199_results.tar.bz2 to download the tarball file including all modeling results listed on this page]

Please cite the following articles when you use the I-TASSER server:

1. Yang Zhang. I-TASSER: Fully automated protein structure prediction in CASP8. Proteins, 77 (Suppl 9): 100-113, 2009.

2. Ambrish Roy, Jianyi Yang, Yang Zhang. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research, 40: W471-W477, 2012.

3. Jianyi Yang, Yang Zhang. I-TASSER server: new development for protein structure and function predictions, Nucleic Acids Research, 43: W174-W181, 2015.