Exp 11: Structure modeling

BY YUEJIAN MO

11510511

May 24, 2018

1 Intruduction

Alwaysly, it easy to get the sequence of protein by the DNA suqence. But it is difficult to know the struture of protein. Here we will some famous server,such as Robeta and SWISS-MODEL to do predict unkonw protein KANK3 structure.

2 Methods

1. Find the amino acid sequence of human KANK3 from uniprot.org.
2. Predict the ANK repeat region in KANK3, using SMART.
3. Predict the 3D structure of the ANK repeat region of KANK3, using SWISS-MODL
4. Analyze the predictiokn results.
5. Predict the KN motif in KANK3
6. Predict the 3D structure of the KN motif of KANK3.

3 Results

1. Sequence Info

Protein name	KN motif and ankyrin repeat domain-containing protein3
Gene name	KANK3
Organism	Homo sapiens
Uniprot ID	UniProtKB-Q6NY19(KANK3_HUMAN)
Source	http://www.uniprot.org/uniprot/Q6NY19

Table 1. Gene KANK3 Information from uniprot.org

2. SMART confidently predicted 4 ANK domains

Name	Start	End	E-value
ANK	622	652	0.0000143
ANK	656	690	503
ANK	695	724	0.000161
ANK	728	758	0.0136

Table 2. Confidently predicted ANK domains of KANK3 by SMART

3. We choose 617-764th amino acid as target sequence for SWISS-MODEL. When 5ybj.1.A was as template, we get best global and local quality estimate and comparison.

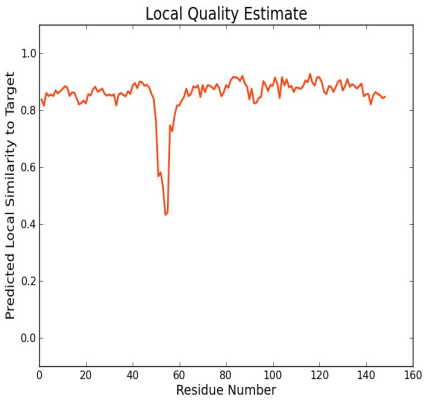

Figure 1. The local quality estimate

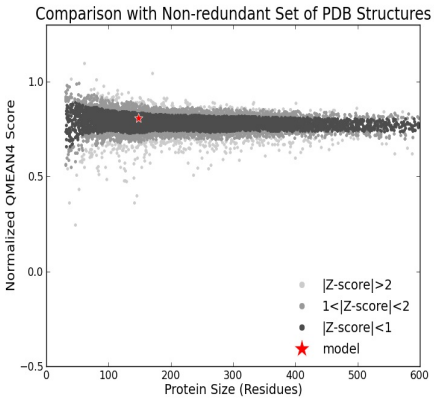

Figure 2. Comparison with Non-redundant Set of PDB Structure

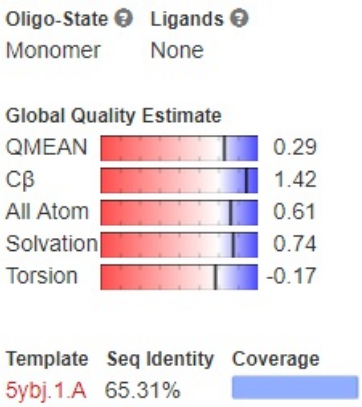

Figure 3. The Global Quality Estimate

Figure 4. The predicted 3D structure of ANK domain

4. As show in figure, most structure show high confience, execpt about 4 amino acid show low confidence(red color amino acid in Figure 4).

5. By the help of Pfam, we found out KN motif start from 33 to 74 in KANK3 gene. We choose gene from 34 to 74 as target for SWISS-MODEL, but SWISS-MODEL failed to found the template to model. So we use profile-based threading server Hhpred and *De novo* modeling server Robetta. From the results of Hhpred and Robetta, we find that KN motif has less features to analysis.(Figure 5, 6)

Figure 5. The 3D structure of KN motif predicted by Hhpred

Features and Secondary Structure

```
                        1   .   10    .   20    .   30    .   40
                        SSPYSVETPYGFHLDLDFLKYIEELERGPAARRAPGPPTSRRPR
tmhmm (0)               ----------------------------------------
low complexity (41%)    -------------------------XXXXXXXXXXXXXXXXX
coiled-coils (0%)       ----------------------------------------
disordered (45%)        XX-----------------------XXXXXXXXXXXXXXXXX
psipred                 ----EE-----EEE--HHHHHHHHHH----------------
```

```
# SignalP-4.0 gram- predictions
# Measure  Position  Value  Cutoff  signal peptide?
  max. C     21      0.100
  max. Y      2      0.112
  max. S     30      0.129
  mean S      1-1    0.129
       D      1-1    0.120    0.570  NO
```

Figure 6.

4 Conclusions

If we can find suitable templates from blast, then we can use the homology modeling server. When the gene has less homology, profile-based threading and *De novo* do better.

5 Reference

- http://pfam.xfam.org/protein/Q6NY
- Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. J Mol Biol. 2017 Dec 16. S0022-2836(17)30587-9.
- Comparative Protein Structure Modeling Using MODELLER.
  Webb B, Sali A. Curr Protoc Protein Sci. 2016 Nov 1;86:2.9.1-2.9.37.