Protein disorder and disordered binding regions

by Zsuzsanna Dosztányi

Exercise 1

DISPROT database and analyzing Nucleoprotein from Nipah virus.

Nipah virus (NiV) is a highly pathogenic emergent paramyxovirus causing deadly encephalitis in humans. Its replication requires a constant supply of unassembled nucleoprotein (N(0)) in complex with its viral chaperone, the phosphoprotein (P).

Find the Uniprot entry: **Q9IK92**

Find the DisProt link on the webpage.

- What is the Disprot ID corresponding to this protein?

You could have search for this protein directly through the DisProt database (https://www.disprot.org/) using the search box.

Explore the information on this page!

- Which region is annotated as disordered?

Check out one the one of the evidence for *Structural state.*

- Name at least one method that was used to show the disordered nature of this region.
- What is the corresponding publication? (PMID?)
- Are there more than one method that supports that this region is disordered?
- Based on the annotation for this protein, can you confidently say that this region is disordered?

Check out the structural transition section:

- What type of transition is this region involved in?

Check out the functional annotation/ interaction section:

- What type of transition is this region involved in?
- What is the function of the region? What types of molecule does it interaction with?

Check out domain annotation for this protein.

- Which regions has a corresponding PFAM annotation? Does it overlap with a disorder annotation? Is there a contradiction?

(You could also check Uniprot's Structure annotation.)

Exercise 2

Predicted disordered and disordered binding regions for Nucleoprotein from Nipah virus.

Collect prediction output for Q9IK92 using az IUPred2A method. IUPred2A is available at https://iupred2a.elte.hu/

For disorder prediction methods the input can be:

- the amino acid sequence in FASTA format
- amino acid sequence in raw format (without header)
- UNIPROT ID or accession number

First, ignore ANCHOR predictions for the moment.
(once you get the results, click on ANCHOR, it will turn it off, or ignore the blue line).

- Which regions are predicted mostly disordered?
- Does the prediction agree with the annotation?

Try the different flavors of IUPred (long/short/structured domain). These options are based on the same approach, but the parameters are tuned slightly differently.

- Is there a big difference in the predictions?
- Where are the structure domains according to IUPred?

Now turn the ANCHOR prediction back on.

- Which regions are predicted as disordered binding sites?
- Are these sites located within the experimentally verified disordered regions? Are they located within the predicted disordered regions?

(Optional)

- Can you find N proteins in Phasepro database? From which viruses?
- Does it have any implications for Nucleoprotein from Nipah virus?

Exercise 3

MobiDB Database – Using direct and indirect data to characterize Cyclin-dependent kinase inhibitor 1B (p27).

Find the p27 in the MobiDB database: It is crosslinked from the DisProt database, or from Uniprot, or use the link:
http://mobidb.bio.unipd.it/P46527.

Go to the ***Curated*** section.

- Which region is annotated as disordered according to DisProt?
- Are there other databases that support the same region?
- In which other databases can you find evidence for the disorder status of this protein?
- Which regions are disordered based on the consensus? Is there a region in conflict?
- Which evidence points to that this regions is ordered, and which points to order?

Go the ***Predictions*** section. Here you can find the results of additional prediction methods.

- How well do the prediction method agree? Which regions has the most disagreement?

Check the results of MobiDB-lite (top method). It is a conservative consensus prediction method (tries to avoid overprediction of disorder) .

- How is the ambigous region predicted by this method?

Go to the ***Indirect Evidence*** section.
MobiDB automatically collects complexes that contain so-called Linear Interacting Peptides (LIPS-- regions that form more contacts with their partner than within themselves).
Click on LIPs

- Which regions are involved in linear peptide interactions?
- Which are the corresponding PDB IDs?

Go to the ***Interactions*** section.
Check the interactions sites with known partners.

Exercise 4

DIBS database - Disordered binding sites involved in the CDK/cyclin complex.

The DIBS database (http://dibs.enzim.ttk.mta.hu/) contains interactions that are formed with ordered protein partners:

Find the DIBS entry corresponding to the complex with the PDB code 1jsu. (Corresponding to p27).

- What is the DIBS id?

Explore the entry page.

- What does the disordered region bind to?
- What is the evidence for the disordered status and for the ordered status?

Click on the Domain Type (on the left), and check if you can find other protein regions binding to this complex.

- How many entries did you find?
- Do they have the same Kds?
- Which interaction is the strongest?

(Optional)

Different interactions that involve the same disordered segment of a protein but different partners tend to compete with one another. In the case of simple competition, the interaction with the smaller K_d is favored. If the difference between K_ds is large, the higher K_d interaction would normally never occur. What other factors can be at play in this specific case, so that both interactions can form inside the cell?

DATABASES

DisProt	https://www.disprot.org/
IDEAL	http://idp1.force.cs.is.nagoya-u.ac.jp/IDEAL/
DIBS	http://dibs.enzim.ttk.mta.hu/
MFIB	http://mfib.enzim.ttk.mta.hu/
FuzDB	http://protdyn-database.org/
MobiDB	http://mobidb.bio.unipd.it/
PhaSePro	https://phasepro.elte.hu/

PREDICTION METHODS

https://en.wikipedia.org/wiki/List_of_disorder_prediction_software