Master Module Proteinbiochemistry and Bioinformatics December 2023

Session: Protein interaction networks

3. Resources for protein interactions

How can I use protein interaction data in biological research?



Interactome mapping





Prediction

Literature curation



Which information for a published interaction should be curated?

IMEx consortium to standardize curation efforts



Orchard et al Nature Methods 2012

Literature curation

Are human genes/proteins equally well studied?

What does this mean for availability of protein interactions?



Key facts

- quite comprehensive
- mix of different interaction types
- biased towards well-studied genes



Systematic protein interactome mapping

HuRI (binary)

BioPlex (co-complex)



Systematic protein interaction mapping

Key facts

- systematic -> not biased towards highly studied genes
- highly controlled experiments
- well documented
- not as comprehensive as curated protein interaction resources

Literature curation



Systematic mapping





Prediction of protein interactions

- Identification of interologs

- Structure-based modeling





Co-evolution,
 co-regulation,
 co-occurrence

Conditions

The XPA (•) protein functions together with the single-stranded DNA (ssDNA) binding protein RPA as the central scaffold to ensure proper positioning of repair factors in multi-protein nucleotide excision repair (NER) machinery. We previously determined the structure of a short motif in the disordered XPA (•) N-terminus bound to the RPA32C domain. However, a second contact between the XPA (•) DNA-binding domain (XPA (•) DBD) and the RPA70AB tandem ssDNA-binding domains, which is likely to influence the orientation of XPA (•) and RPA on the damaged DNA substrate, remains poorly characterized. NMR was used to map the binding interfaces of XPA (•) DBD and RPA70AB. Combining NMR and X-ray scattering data with comprehensive docking and refinement revealed how XPA (•) DBD and RPA70AB orient on model NER DNA substrates. The structural model enabled design of XPA (•) DBD-RPA70AB interaction. Our results inform ongoing controversy about where XPA (•) is bound within the NER bubble, provide structural insights into the molecular basis for malfunction of disease-associated XPA (•) missense mutations, and contribute to understanding of the structure and mechanical action of the NER machinery. **Excerpts from full text:**

... the globular core and basic residues in the C-terminal extension. RPA is a heterotrimer of RPA70, RPA32 (I) and RPA14 subunits (Figure 1C). The tandem high affinity DNA-binding domains [...] intensity at q = 0 Aa-1 (I(0)), relative to the recorded frame. Uniform RG values across an XPA (I) DBD-DNA-RPA70AB elution peak represented a homogenous assembly (Supplementary Figure ...

Prediction of protein interactions

Key facts

- high false positive rate
- highly biased (orthologs, structures available)
- for some species only way to get protein interaction data



Exercise: Explore STRING DB

1st part:

- Explore the STRING DB (<u>string-db.org</u>) with the help of the questions (STRING_questions.txt) and input list of proteins (STRING_input_28_genes.txt) provided
- Take notes and/or screenshots of your observations

2nd part:



30 mir

Discussion of results with everyone



I found 28 genes in my screen that are likely associated with Neurodevelopmental disorders:

- do these genes work in the same biological process?
- are these genes part of the same protein complex?
- -> do these proteins (tend to) interact with each other?

Exercise: Explore protein interaction databases

Take home messages

- STRING contains predicted and experimentally based protein associations -> only a small fraction corresponds to actual protein interactions
- You can filter your search results based on your question/interest
 -> make use of it to get a meaningful output
- STRING provides many tools to analyse and explore your network
- Make sure you understand the content of a bioinformatic resource before using it