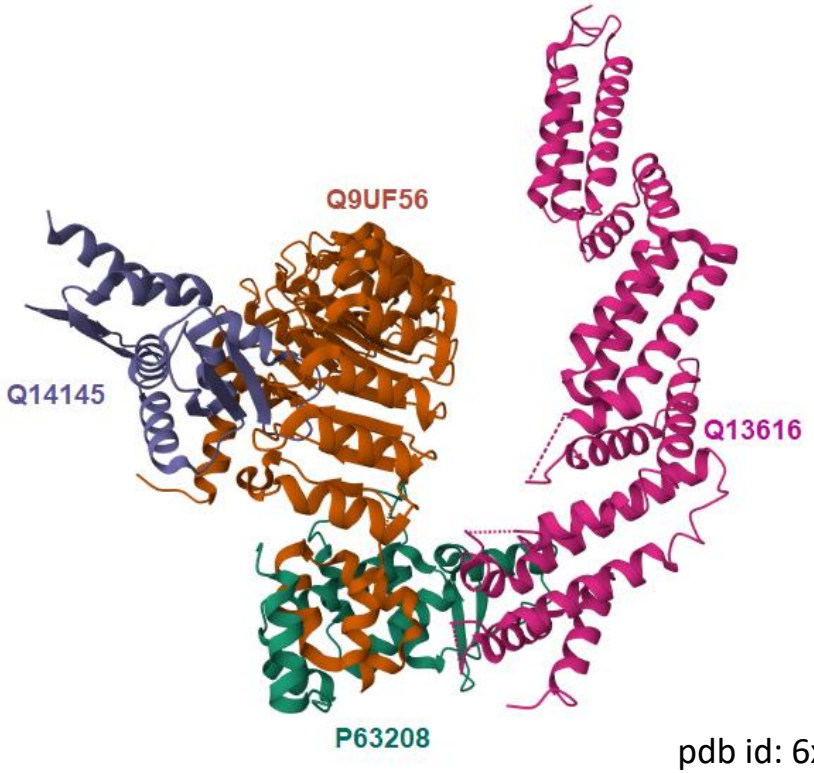
Unraveling the Function of Protein Interactions: Insights and Future Predictions

2025 M.Sc. Biology & Biomedicine
Module: Proteinbiochemie und Bioinformatik

Aimilia Christina Vagiona
PhD student
avagiona@uni-mainz.de
Supervisor: Miguel Andrade

1. Why protein interactions are important?



pdb id: 6xcq

They help us to:

- Study **biological processes** and **molecular functions**
- Understand **disease mechanisms**
- Develop **new therapies**
- **Evolutionary and functional** annotations

2. Classification of protein interactions

a. Composition

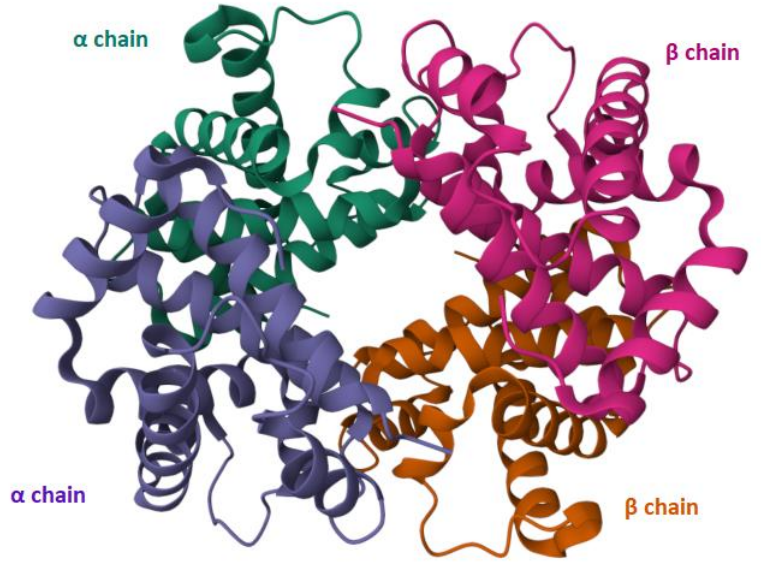
Homo-oligomers:

- If interacting partners are identical

Hetero-oligomers:

- If interacting partners are non-identical

Heat Shock Protein 27
(Homo-oligomer)



Hemoglobin
(Hetero-oligomer)

2. Classification of protein interactions

b. Affinity

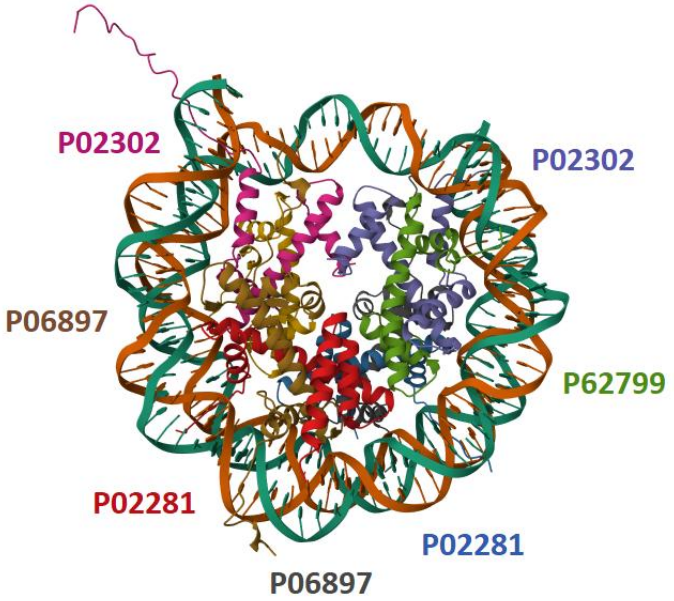
Obligate:

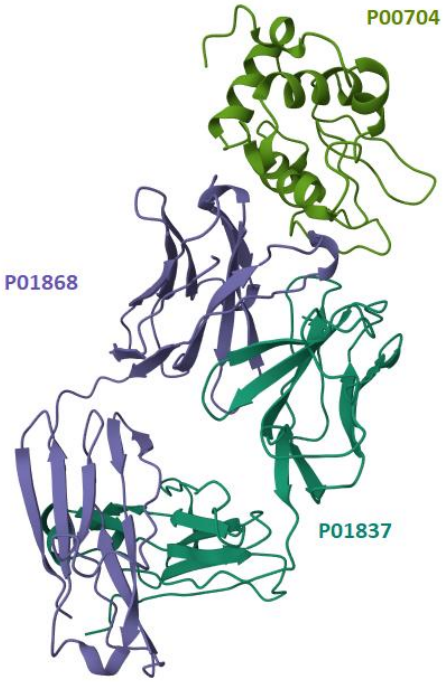
- Constituents of a complex are unstable on their own in vivo

Non obligate:

- The components of non-obligate interactions can exist independently
- Ofter are regulated by environmental or cellular conditions

Histone octamer
(Obligate)





RNAse-Antibody
(Non- obligate)

2. Classification of protein interactions

c. Lifetime

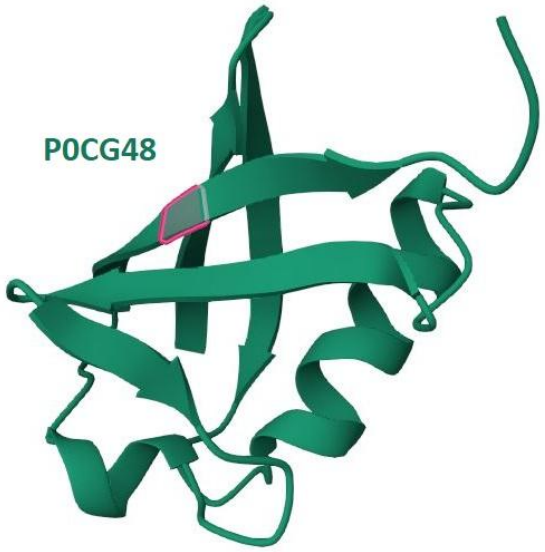
Transient:

- The components of transient interaction associate and dissociate temporarily in vivo

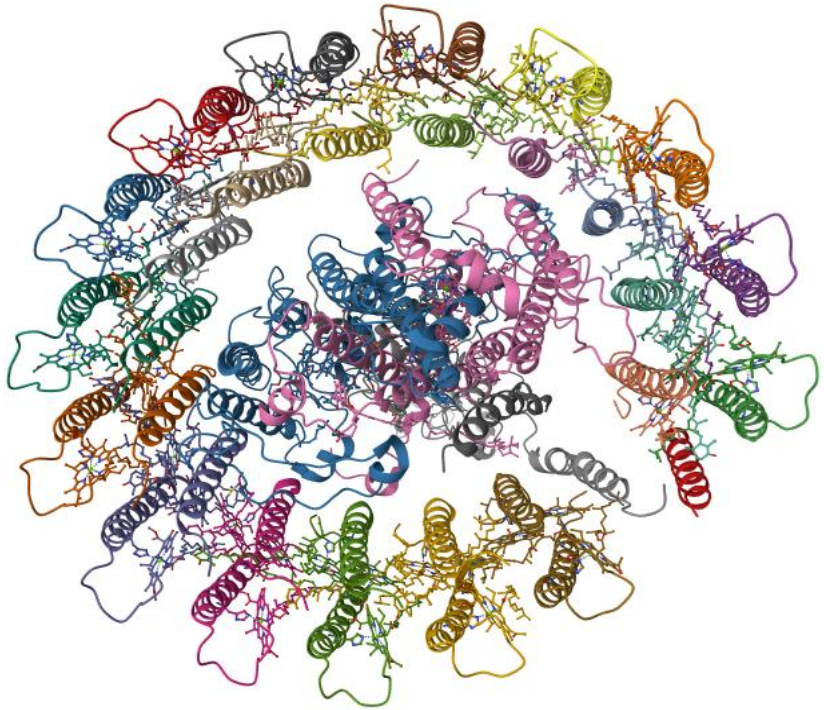
Permanent:

- Permanent interactions are usually very stable and irreversible

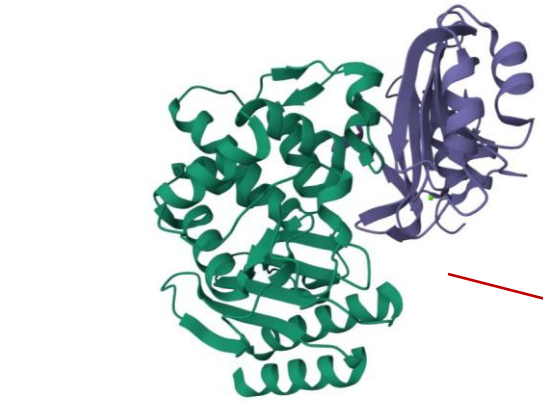
1UBQ
(transient)

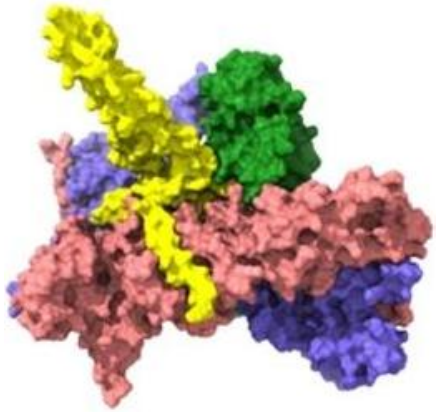


ATP Synthase Complex
(permanent)



3. Studying Protein-Protein Interaction Networks (PPINs)





It’s a very complex system

3. How do we study protein interactions on a large scale computationally?

Graph theory!

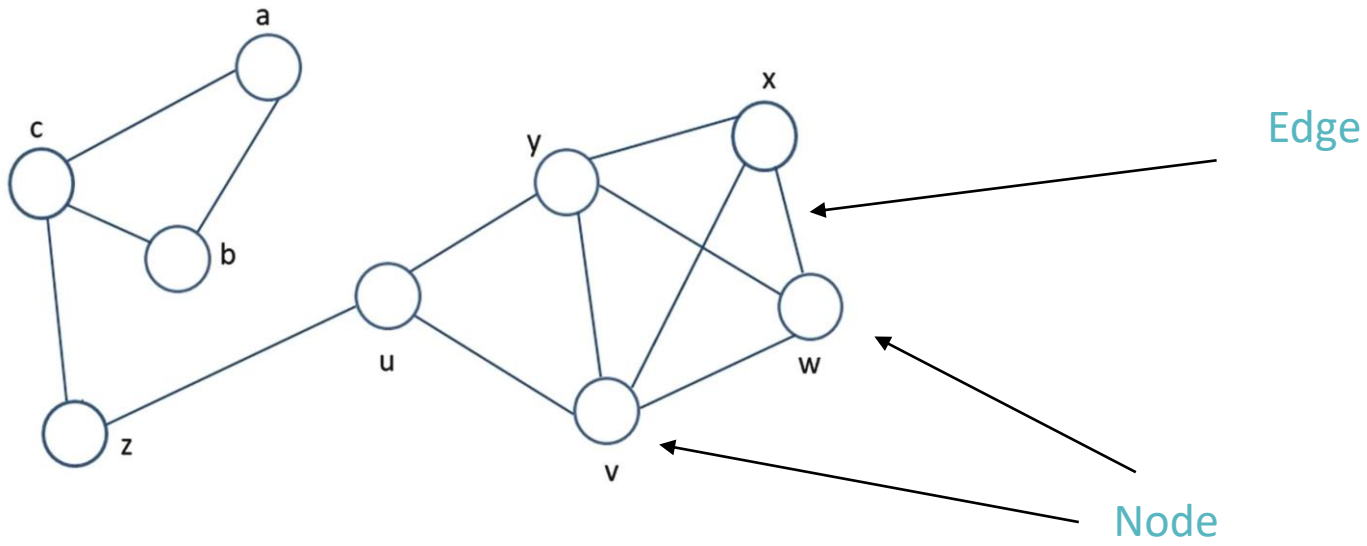
"[…] the study of graphs, mathematical structures used to model pairwise relations between objects. A graph in this context is made up of vertices, nodes, or points which are connected by edges, arcs, or lines".

Wikipedia

Graph

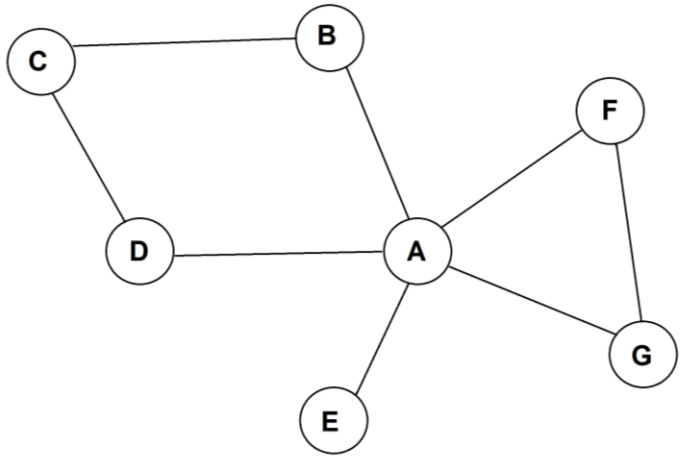
V = {v1,v2,v3,v4,…}

E = {(v1,v2),(v2,v3),(v2,v4),…}



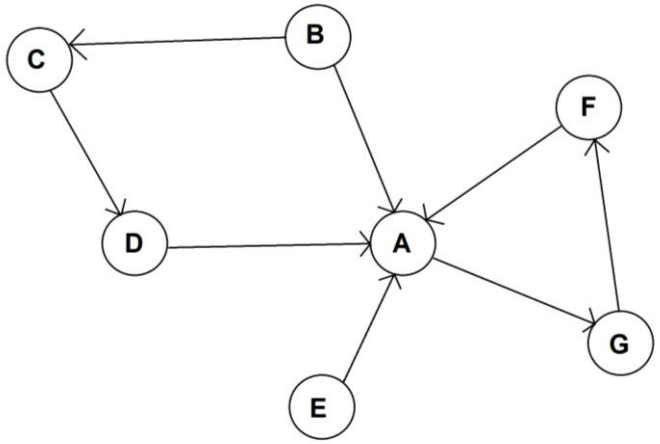
3. Graph Theory: types of graphs

undirected



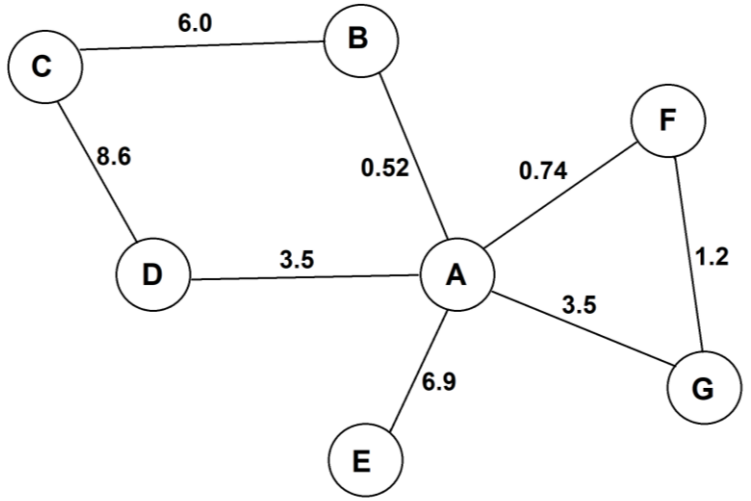
- protein protein interaction networks

directed



- metabolic networks
- regulatory networks

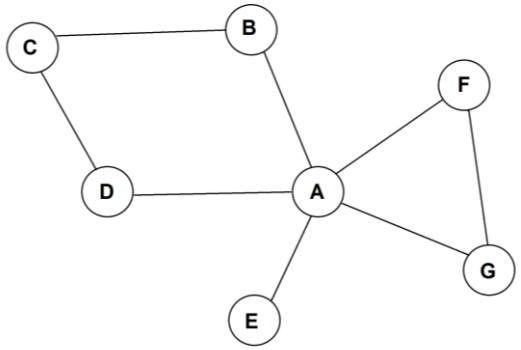
weighted



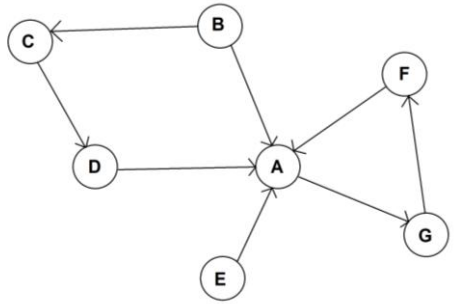
- gene co-expression networks

3. Graph Theory: adjacency matrices

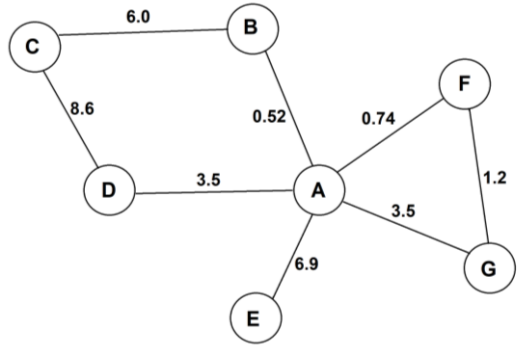
undirected



directed



weighted



	A	B	C	D	E	F	G
A	0	1	0	1	1	1	1
B	1	0	1	0	0	0	0
C	0	1	0	1	0	0	0
D	1	0	1	0	0	0	0
E	1	0	0	0	0	0	0
F	1	0	0	0	0	0	1
G	1	0	0	0	0	1	0

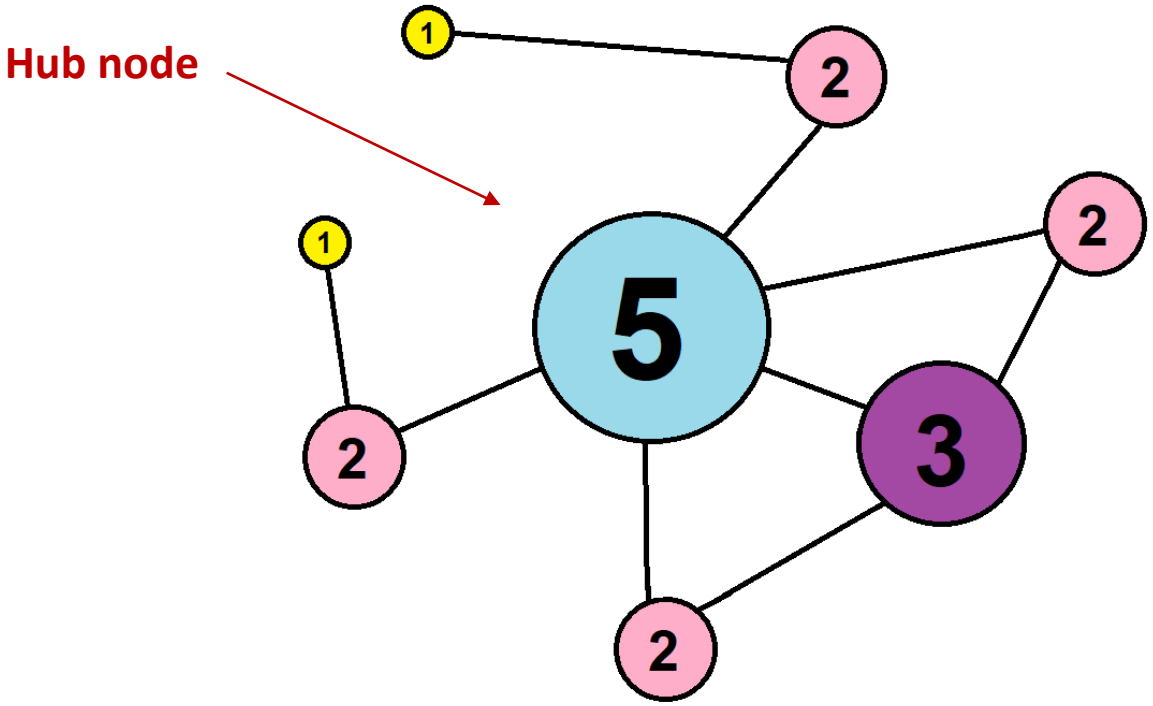
Ending points

Starting points

	A	B	C	D	E	F	G
A	0	0	0	0	0	0	1
B	1	0	1	0	0	0	0
C	0	0	0	1	0	0	0
D	1	0	0	0	0	0	0
E	1	0	0	0	0	0	0
F	1	0	0	0	0	0	0
G	0	0	0	0	0	1	0

	A	B	C	D	E	F	G
A	0	0.52	0	3.5	6.9	0.74	3.5
B	0.52	0	6.0	0	0	0	0
C	0	6.0	0	8.6	0	0	0
D	3.5	0	8.6	0	0	0	0
E	6.9	0	0	0	0	0	0
F	0.74	0	0	0	0	0	1.2
G	3.5	0	0	0	0	1.2	0

3. Graph Theory: network topological properties



Degree Centrality

- node property
- number of edges of a vertex (node)

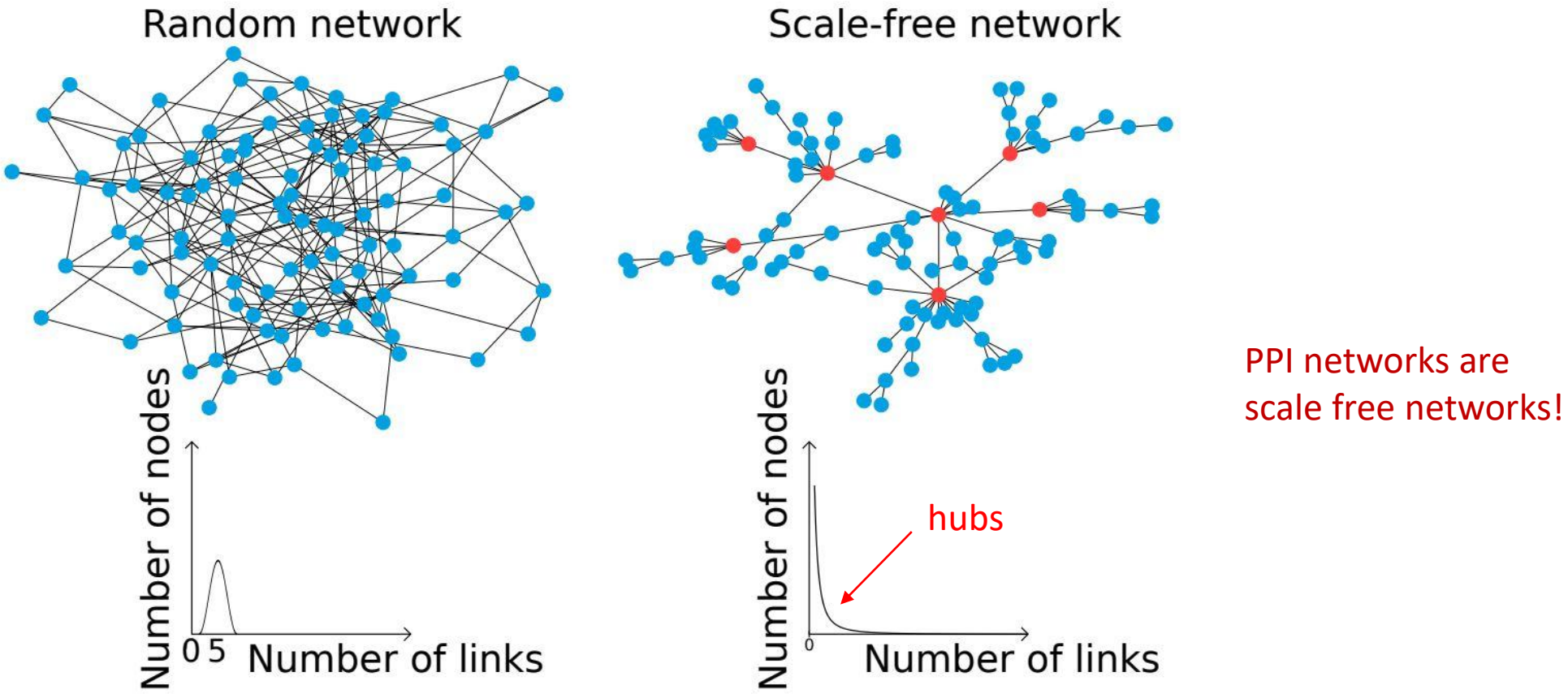
Average degree

- network property
- mean over all degrees in the network

Degree distribution

- network property
- informs about the topology of the network

3. Graph Theory: degree distribution

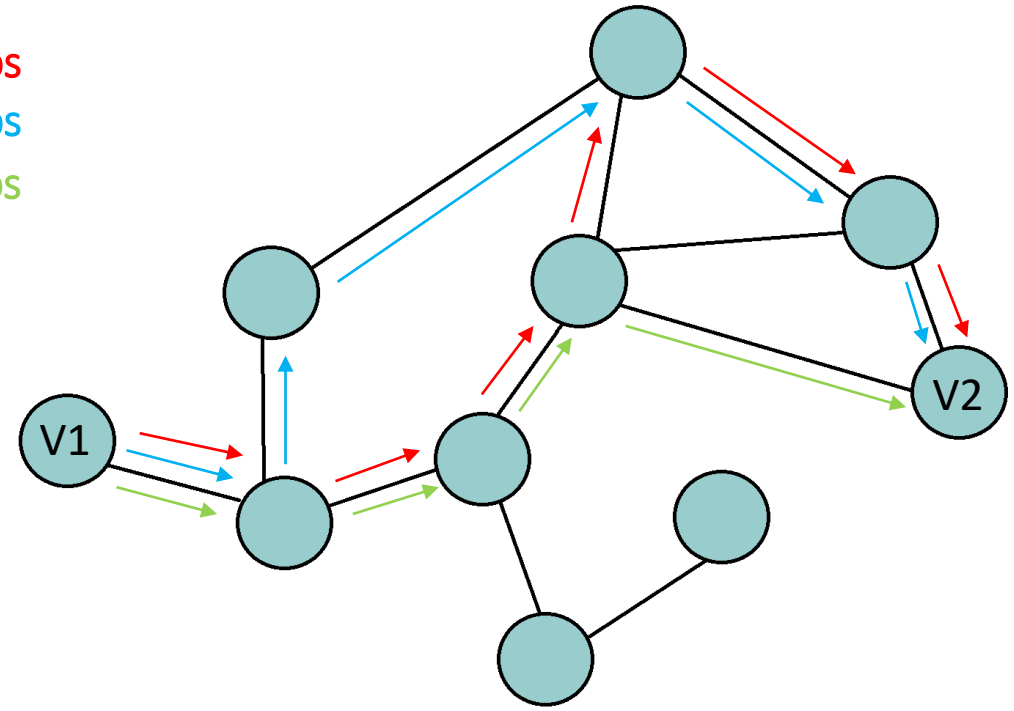


PPI networks are scale free networks!

3. Graph Theory: shortest path

Paths lengths:

6 steps
5 steps
4 steps



Path

- between two vertices is formed by the edges that lead from one vertex to another

Shortest path

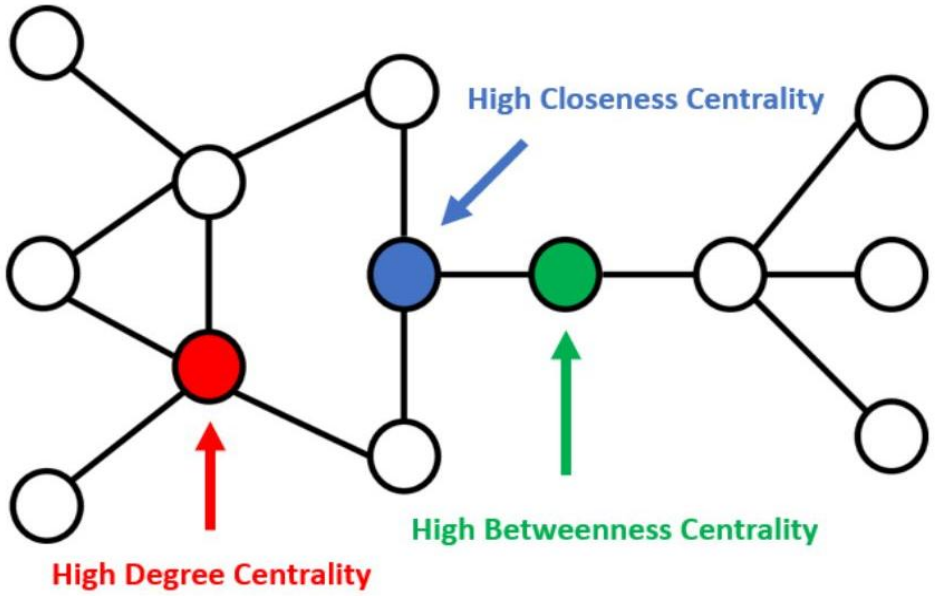
- shortest path between the two vertices
- used to model how information flows

The shortest path between two proteins in a PPI network often represent **the most efficient routes** for signaling

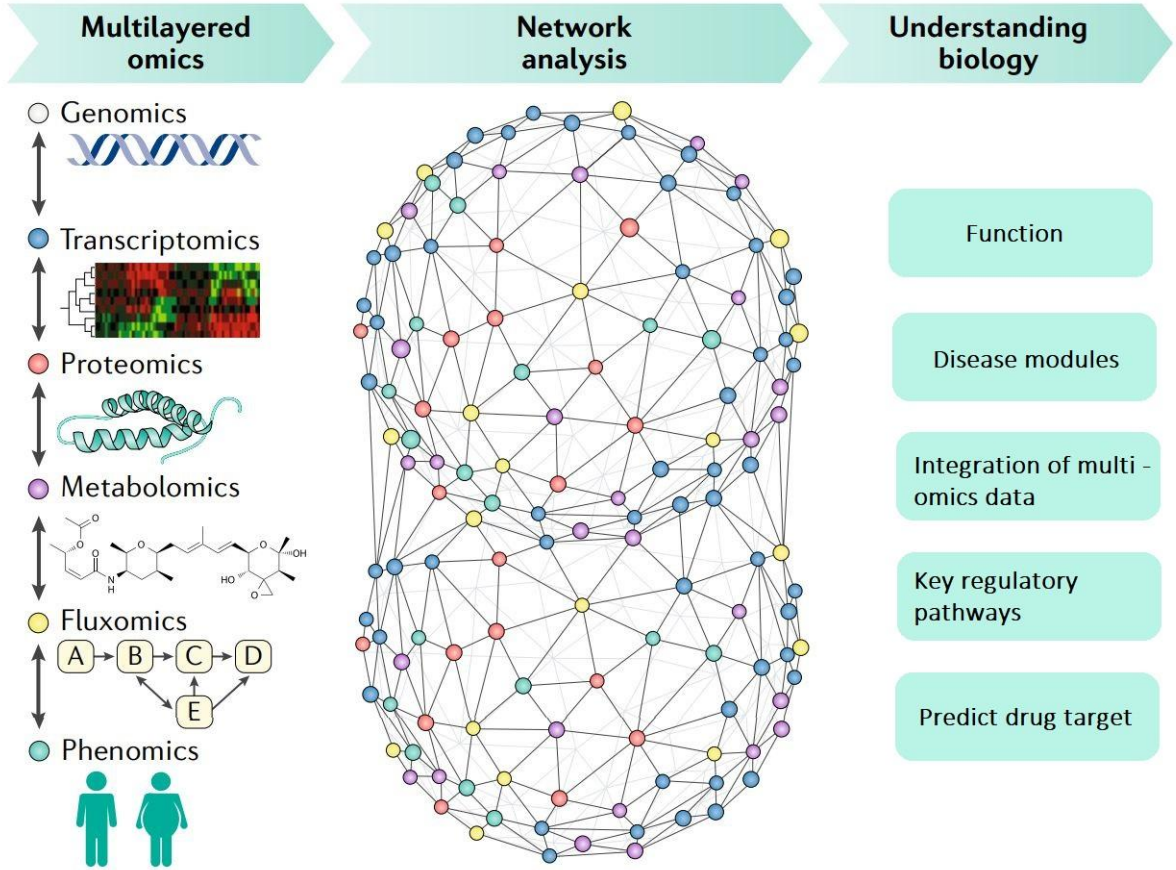
3. Graph Theory: betweenness and closeness centrality

Many shortest paths pass through a node → High **betweenness** centrality → Flow of information in the graph

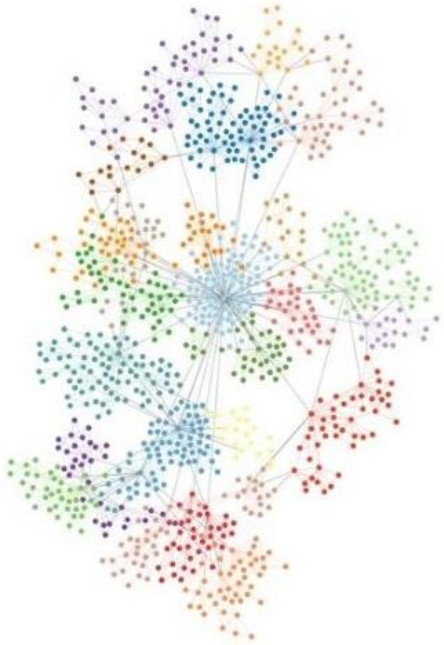
The length of shortest paths passing through the node is **low** → High **closeness** centrality → Measure of centrality in the graph



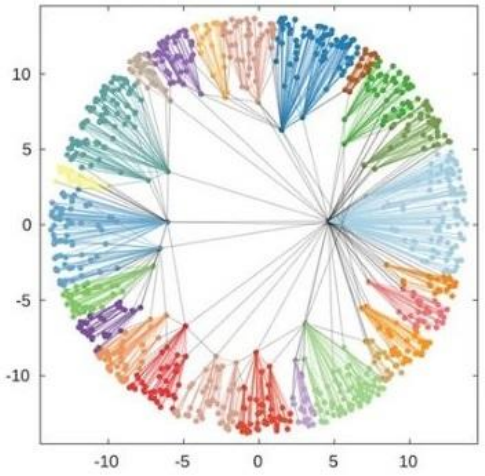
Networks in Biology



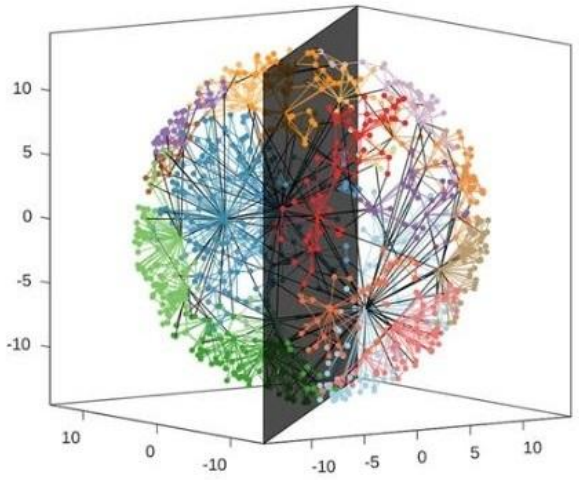
Hyperbolic network model



Euclidean



Hyperbolic model in 2D

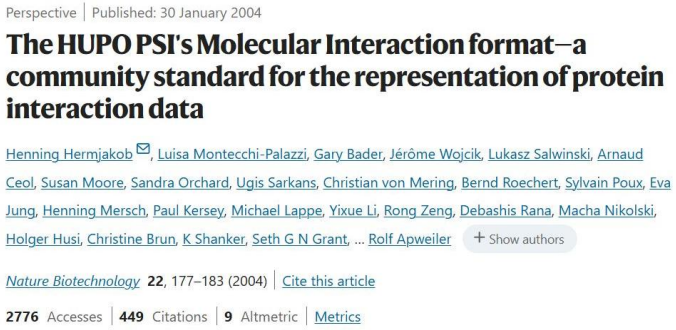


Hyperbolic model in 3D

4. How to link ontologies and PPIs

The Molecular Interactions (MI) ontology forms a structured controlled vocabulary for the annotation of experiments concerned with protein-protein interactions.

HUPO HUMAN PROTEOME ORGANIZATION | translating the code of life



Perspective | Published: 30 January 2004

The HUPO PSI's Molecular Interaction format—a community standard for the representation of protein interaction data

Henning Hermjakob, Luisa Montecchi-Palazzi, Gary Bader, Jérôme Wojcik, Lukasz Salwinski, Arnaud Ceol, Susan Moore, Sandra Orchard, Ugis Sarkans, Christian von Mering, Bernd Roechert, Sylvain Poux, Eva Jung, Henning Mersch, Paul Kersey, Michael Lappe, Yixue Li, Rong Zeng, Debashis Rana, Macha Nikolski, Holger Husi, Christine Brun, K Shanker, Seth G N Grant, ... Rolf Apweiler + Show authors

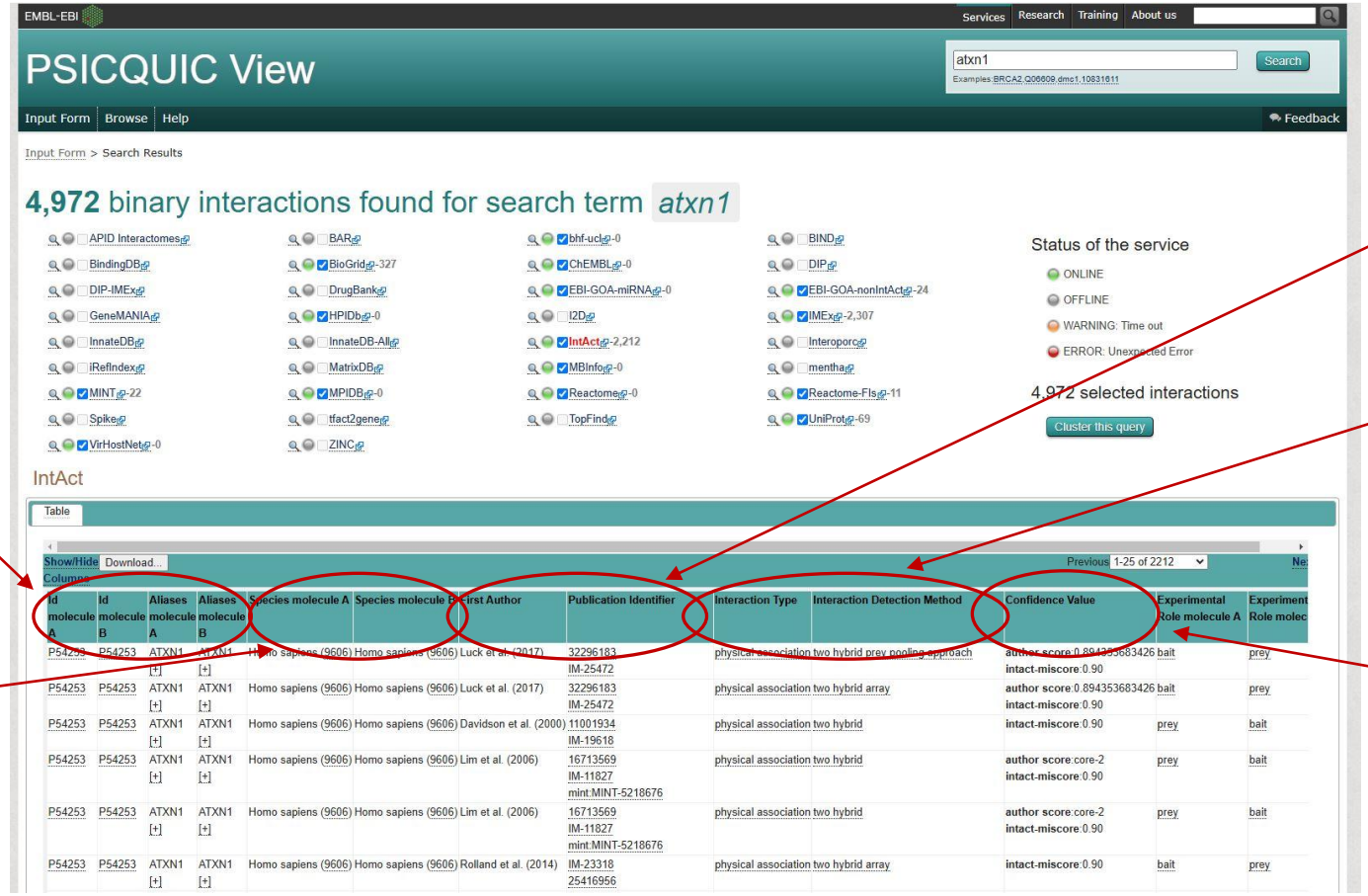
Nature Biotechnology **22**, 177–183 (2004) | Cite this article

2776 Accesses | **449** Citations | **9** Altmetric | Metrics

PSI-MI TAB Format (MITAB):

1. Interactor A & B: Unique identifiers (e.g., UniProt IDs)
2. Interaction Type: Nature of the interaction (e.g., physical association, enxymatic activity)
3. Detection Method: Experimental approach used (e.g., yeast two-hybrid, co-IP)
4. Confidence Score: Quantifies the reliability of the interaction
5. Source Database: Where the data originated (e.g., IntAct, DIP)

4. How to link ontologies and PPIs



Publication info

Which proteins

Experimental info

Which species

Confidence score

5. Databases for protein interactions

Primary Databases
(experimental interaction data)

BioGRID 4.4

MINT

IntAct

Secondary Databases
(consensus-based)

Agile Protein Interactomes DataServer

HIPPIE » Human Integrated Protein-Protein Interaction rEference

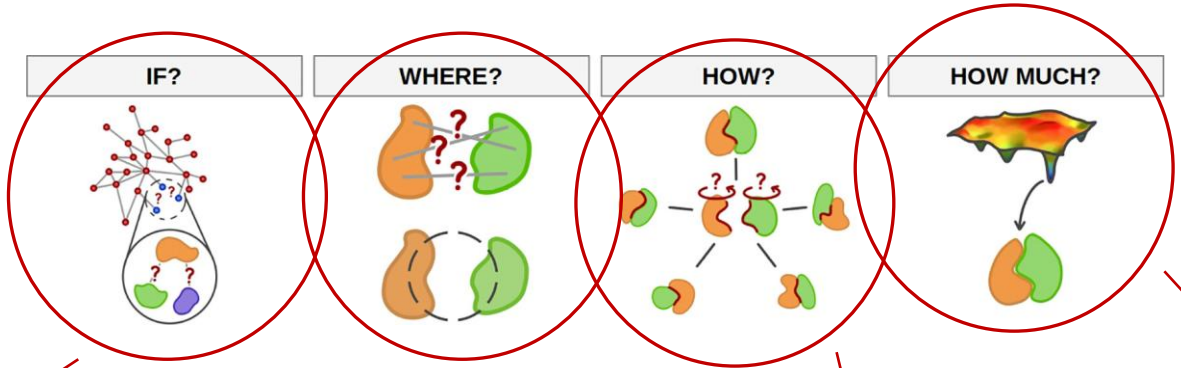
Predictive Databases
(experimental data with computational predictions)

STRING

PIPs Human Protein-Protein Interaction Prediction

UniHI -- the Unified Human Interactome database

6. Computational predictors of PPIs



if?	
PPI-Detect	sequence-based ML
DeepFE-PPI	sequence-based ML

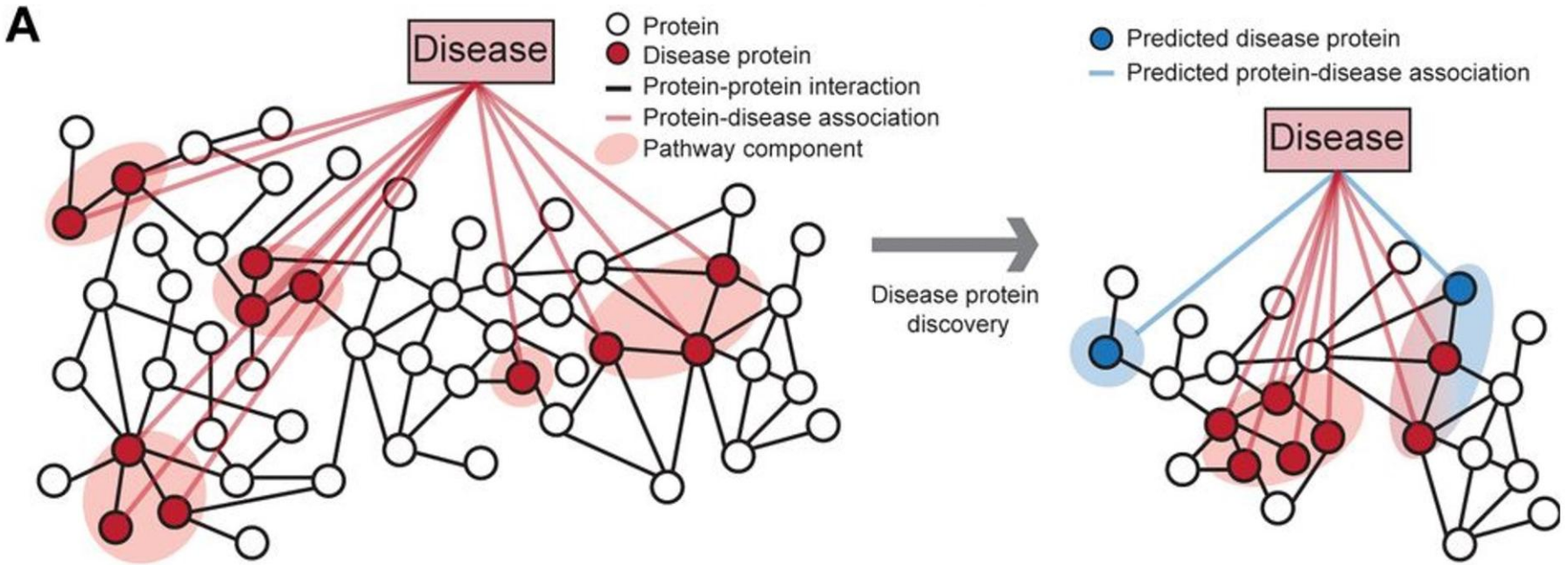
where?	
HADDOCK2	simulated annealing docking
HDOCK	homology + FFT docking
ClusPro	FFT docking
MEGADOCK	FFT docking
SWISS-MODEL	homology modeling
AlphaFold2	sequence-based ML

how?	
GROMACS	**MD simulations**
AMBER	**MD simulations**
CHARMM	**MD simulations**
PyEMMA 2	**MSM construction**
OpenMM	**MD simulations**
GENESIS	**MD simulations**

how much?	
PIPR	sequence-based ML
ISLAND	sequence-based ML
mmCSM-PPI	structure-based ML
AffPred	knowledge-based energy function
PerSpect-EL	structure-based ML
PPI-Affinity	structure-based ML

Grassmann et al. (2024), *Chemical Reviews*

7. PPINs and diseases



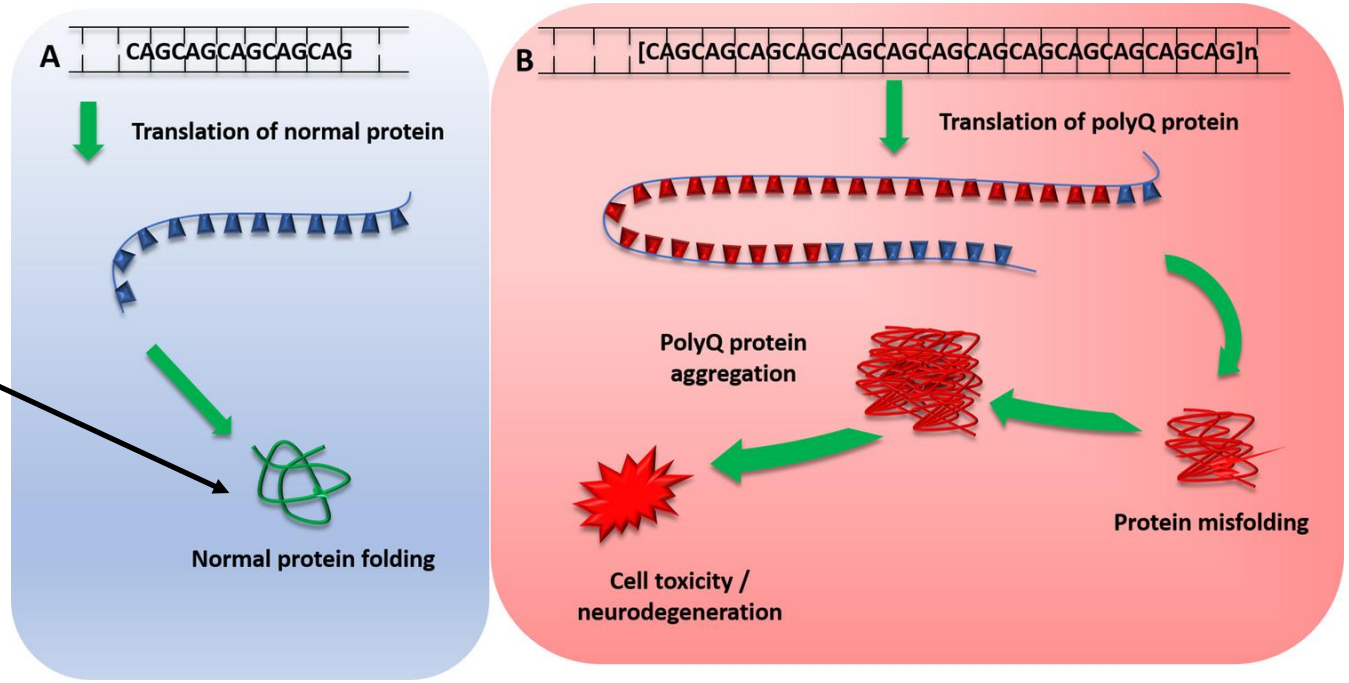
- Investigate disease pathogenesis
- Identification of critical nodes
- Drug discovery
- Protein networks can model how a mutation affects cellular signaling over time, offering predictions about disease onset and progression.

7. PPINs and diseases

A survey of SCA1

Spinocerebellar Ataxia Type 1
→ Neurodegenerative disease

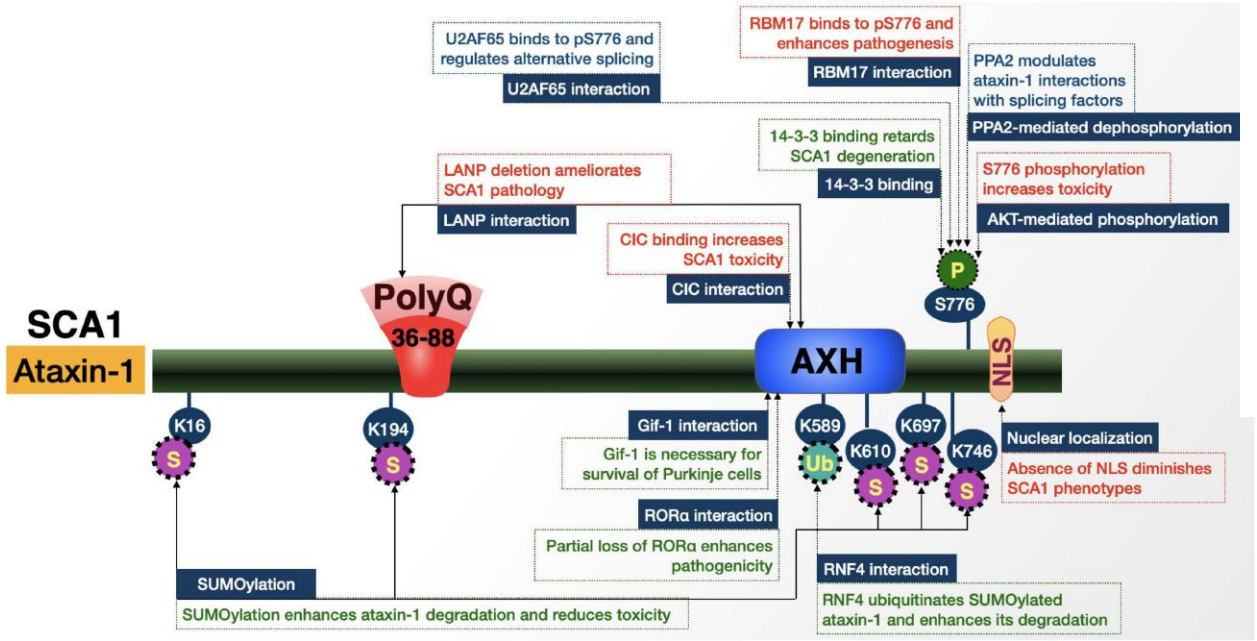
Ataxin-1 (RNA binding protein)



Addition of PolyQ track in N-terminal region causes protein aggregation!

7. PPIs and diseases

A survey of SCA1



Graphic representation of the SCA1 protein, its domains, interactions, and modifications.

› JCI Insight. 2021 Feb 8;6(3):e144955. doi: 10.1172/jci.insight.144955.

Modulation of ATXN1 S776 phosphorylation reveals the importance of allele-specific targeting in SCA1

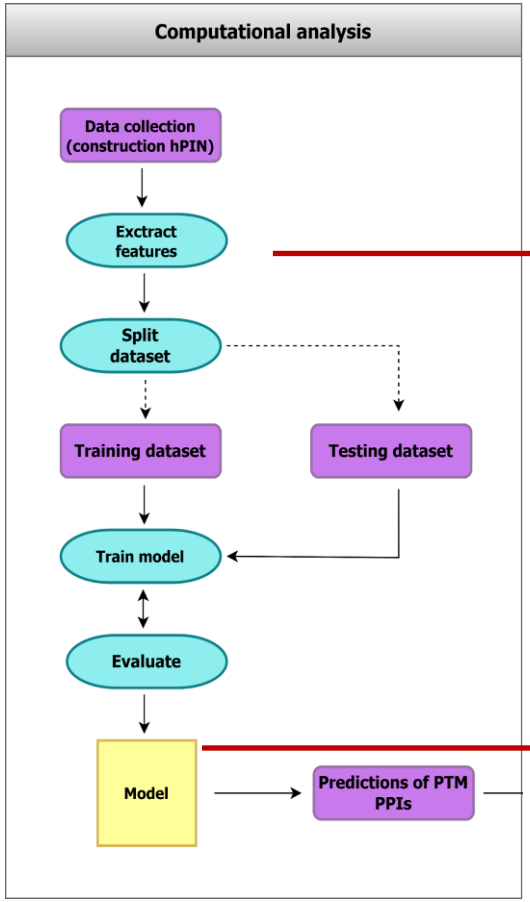
Larissa Nitschke [1 2 3], Stephanie L Coffin [2 3 4], Eder Xhako [2 3 4], Dany B El-Najjar [2 3], James P Orengo [3 5], Elizabeth Alcala [2 3], Yanwan Dai [3 6], Ying-Wooi Wan [2 3], Zhandong Liu [3 6], Harry T Orr [7], Huda Y Zoghbi [1 2 3 4 5 6 8]

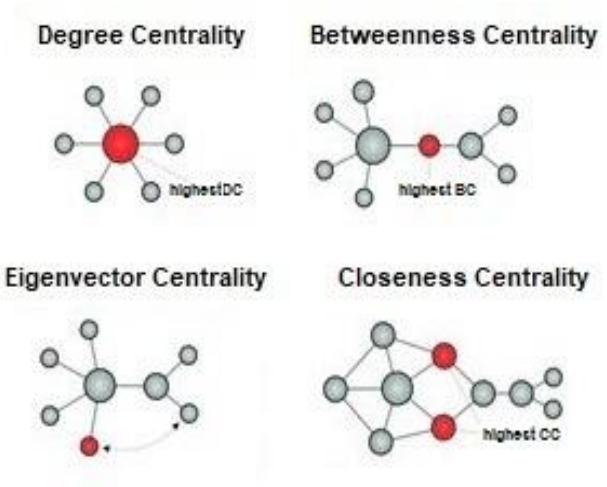
Affiliations + expand

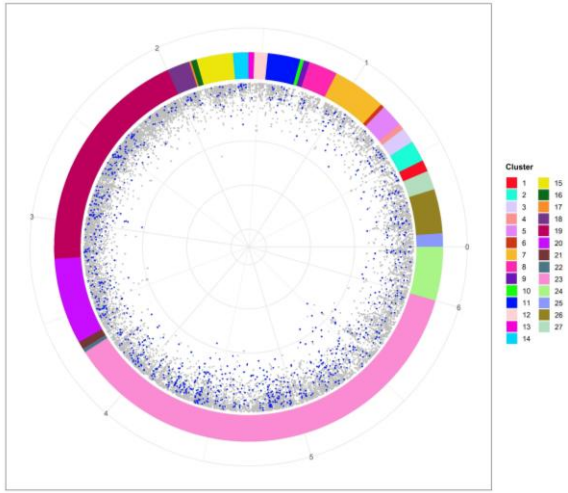
PMID: 33554954 PMCID: PMC7934855 DOI: 10.1172/jci.insight.144955

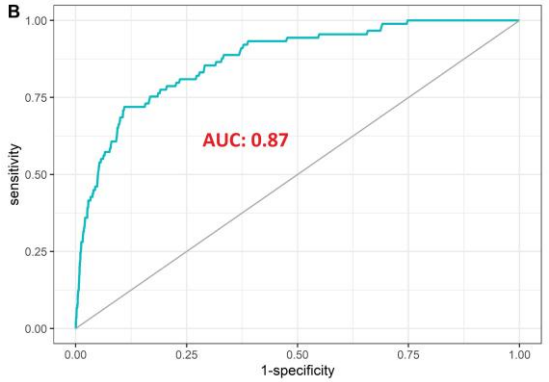
Disruption of S776 phosphorylation on the polyQ-expanded ATXN1 results in an **improvement** of SCA1 pathogenesis!

7. PPIs and diseases

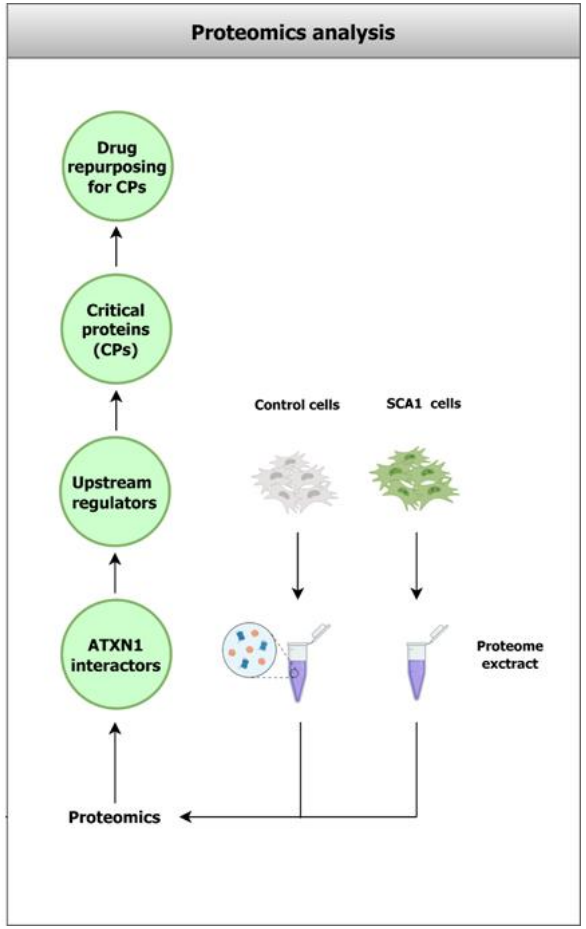


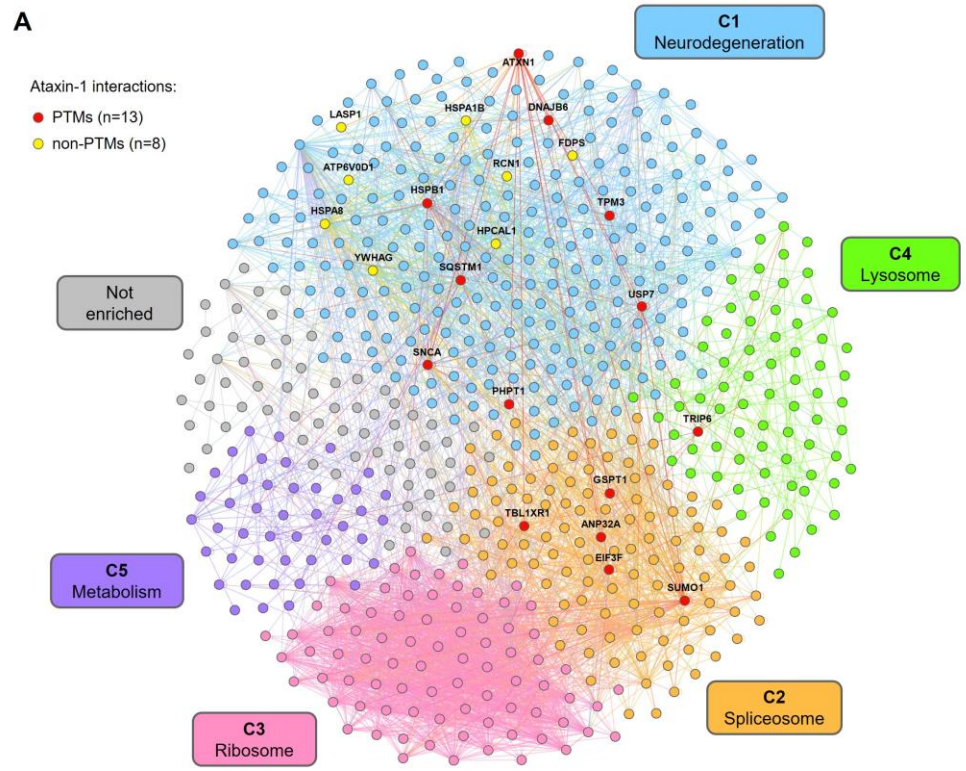






7. PPIs and diseases





PPI network of significantly dysregulated proteins in SCA1 cells

Takeaways for therapeutic usage of PPINs

PPI networks reveal disease mechanisms: From hijacked host pathways in infections to disrupted molecular interactions in neurodegeneration.

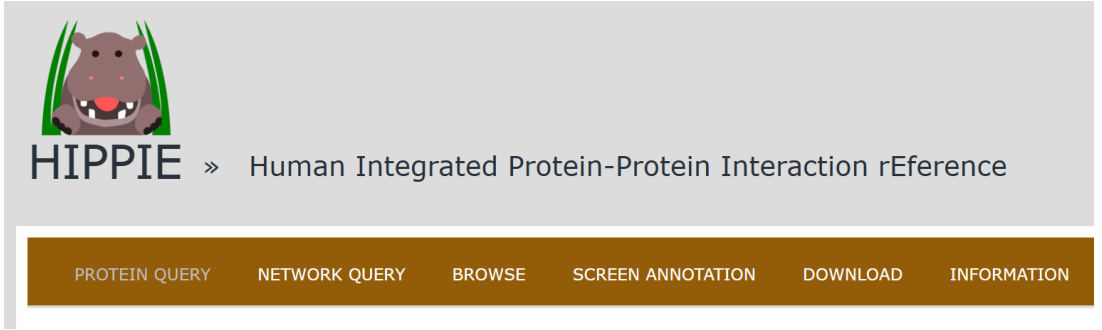
They guide drug discovery: Many successful therapies were developed by targeting interactions within these networks.

They facilitate multi-omics integration: Combining genetics, transcriptomics, and proteomics with PPI networks provides a comprehensive disease understanding.

Lovely to meet you all !!!!!

Thank you!
Any questions?

8. Time to play…



1) Google “hippie database” and go to https://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/
2) Click on “NETWORK QUERY” and type on the box: *ATXN1*
3) Scroll down on the website and set the output type as: show in browser-text, set the HIPPIE confidence score = 0,7 and select on the tissue filter the “brain-cerebellum”
4) Click on search

QUESTIONS

1) How many interactor ATAXIN 1 has?
2) Is ATAXIN 1 interacting with CIC protein?
3) If yes, how many publications validate this interaction?