Introduction to Protein-Protein Interaction

Master-Module Biological Networks

July 18, 2016

Emidio Capriotti

http://biofold.org/

Biomolecules Folding and Disease

Institute for Mathematical Modeling
of Biological Systems
Department of Biology

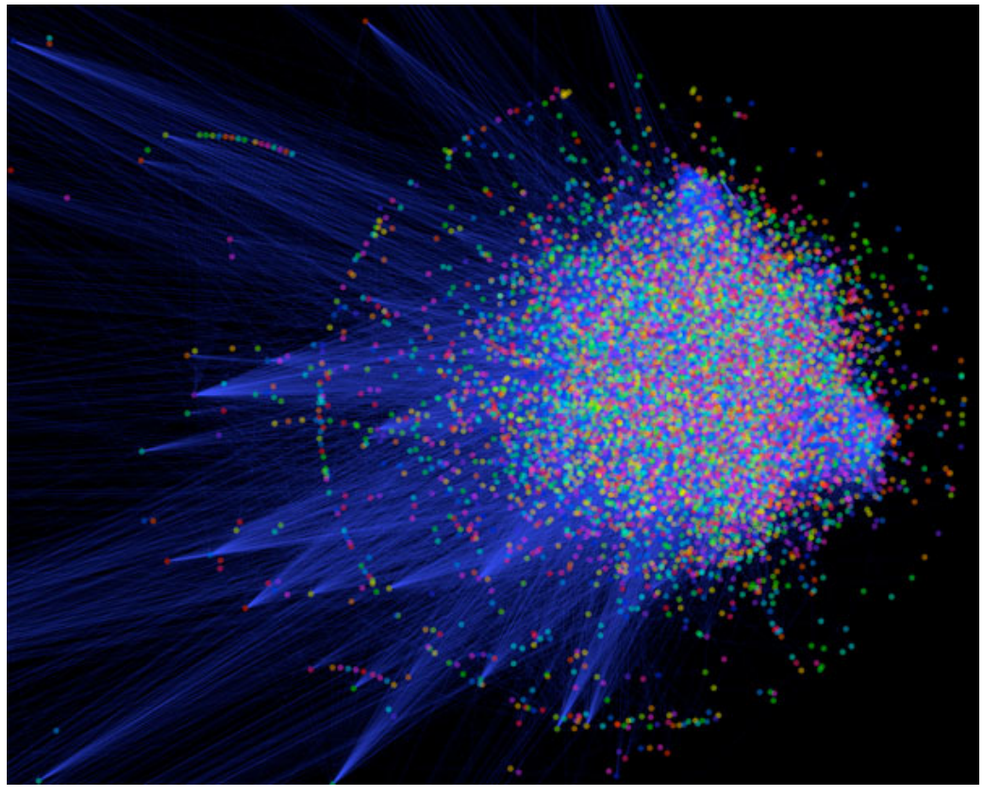
HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Why Protein Interaction?

Protein interactions underlie the assembly of macromolecular machines, mediate signaling pathways in cellular networks, and control cell-to-cell communication.

In an organism, PPIs form a huge complex network known as an “interactome”.

Nearly 650,000 interactions regulate human life.



Any deregulation leads to a disease state or death.

Levels of Complexity

Reductionist biology – molecular viewpoint

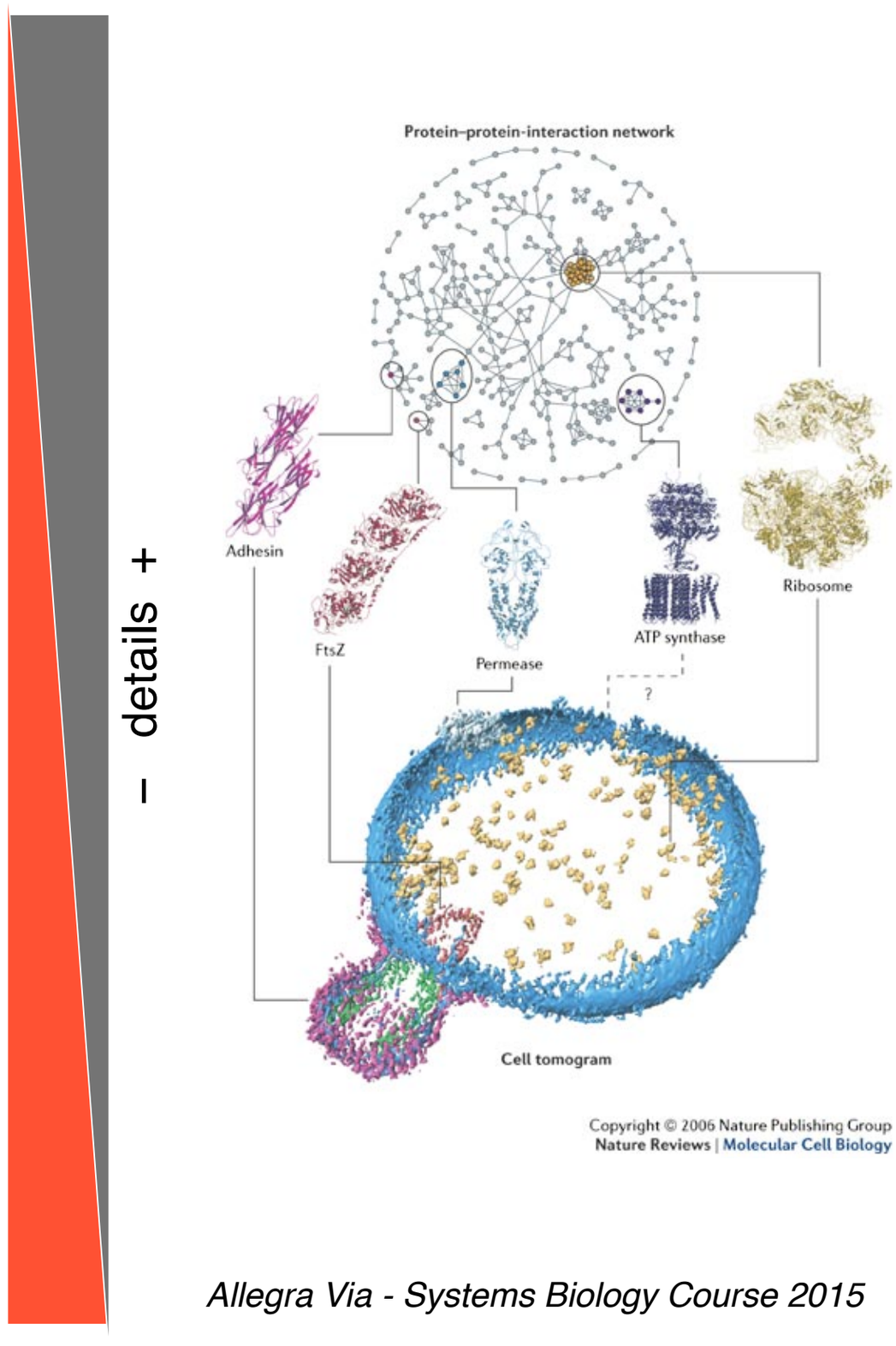
- **Specific molecule(s)** of interest
- Experiments to determine interaction partners and modes of interaction
- Prediction of interaction partners and modes of interaction
- Analysis of specific interaction

Protein networks

- Identification of **functional modules** (set of proteins highly connected to each other)
- Hubs, singletons
- Networks' analysis -> biological hypotheses
- Prediction of new interactions

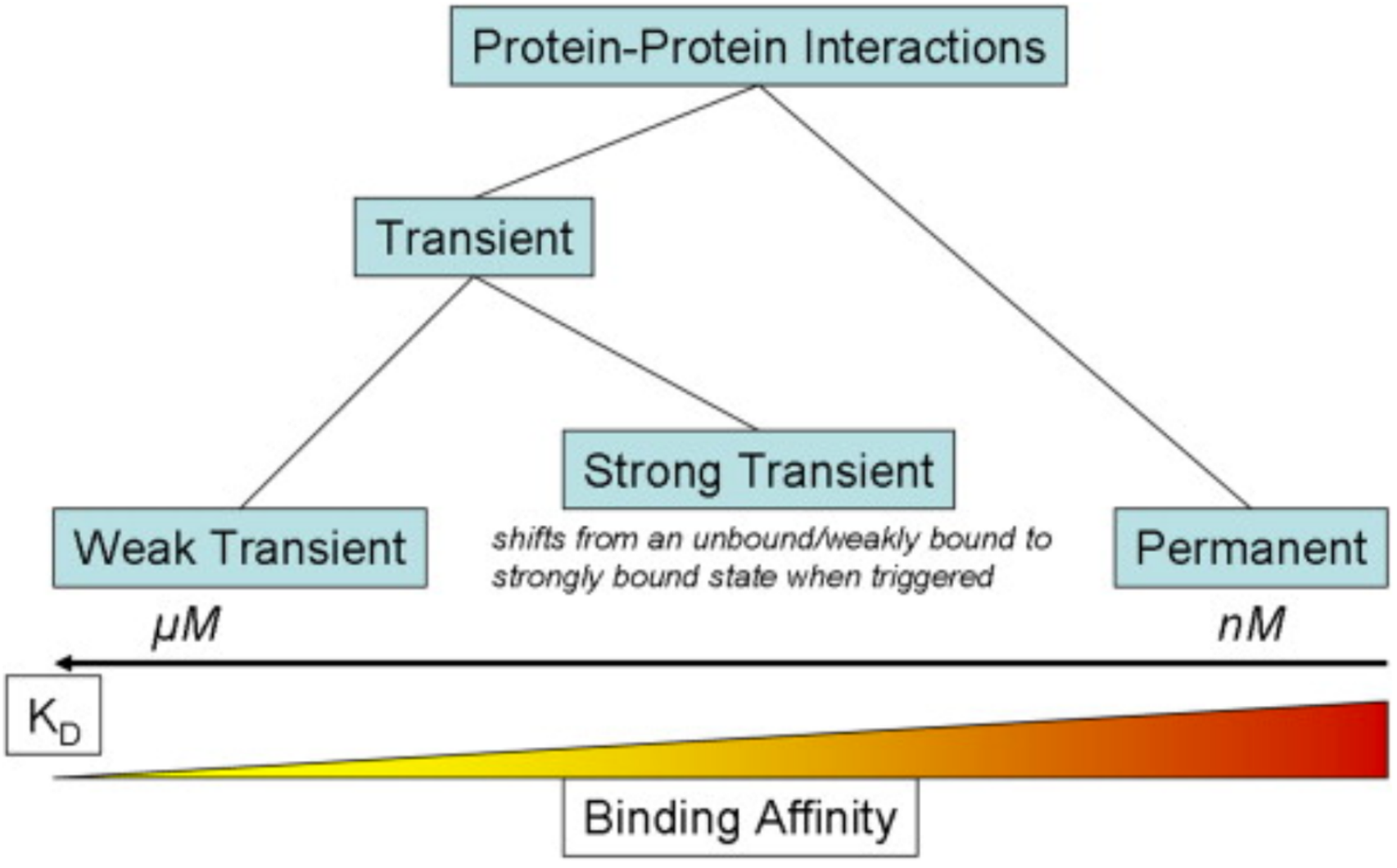
Systems biology

- **Networks, pathways** implicated in a condition
- Identify perturbed or deregulated systems
- Modelling of the system to infer to signals and/or deregulation events
- Experiments to determine responses of the system



Allegra Via - Systems Biology Course 2015

Protein-Protein Interactions



Strong transient: This category includes interactions that are triggered/stabilised by an effector molecule or conformational change. An example is given by the Ras proteins, which form tight complexes with their partners when GTP-bound and only weak complexes when GDP-bound.

Perkins et al, Structure, 2010

Transient vs Stable

- Transient (relatively weak)

  Brief and reversible interactions occurring in specific cellular contexts
  - Interactions mediated by short linear motifs
  - Interactions mediated by PTMs
  - Disorder-to-order transitions
  - Proteins involved in signalling cascades

- Stable (for a longer period of time)

  Proteins take part of permanent complexes as subunits, in order to carry out structural or functional roles

  Permanent, obligate, oligomeric, tight, more stable
  - Homo-oligomeric or hetero-oligomeric complexes
  - Interactions mediated by PTMs
  - enzyme-inhibitor
  - antibody-antigen
  - domain-domain
  - domain-peptide

PPI Identification Methods

Experimental (*in vivo*)	• Yeast two-hybrid system • Split ubiquitin system • Split lactamase/galactosidase system • Split yellow fluorescent protein system
Experimental (*in vitro*)	• Co-immunoprecipitation • Tagged Fusion Proteins • X-Ray Diffraction • Biacore • Phage Display
Computational (*in silico*)	• BIND • DIP • MINT • IntAct

The Molecular Viewpoint

- The affinity of PPI varies from millimolar to picomolar, depending on the type of interaction and signaling needed (Chen et al. Protein Sci. 2013)
- Despite affinity varies over a wide range, proteins maintain a high degree of specificity for their partners
- Many proteins exhibit specificity for multiple partners (Reichmann et al. Curr. Opin. Struct. Biol. 2007).
- The nature of the interaction surface determines how proteins interact
- A detailed knowledge of the interaction surfaces of proteins and their energetics is necessary to understand the regulatory mechanisms of biochemical pathways (especially to modulate or block these pathways for therapeutic purposes)

Surface of Interaction (I)

- The area of PPI interfaces is large (1000 to 4000 Å^2)
- Standard-sized interfaces are 1200 to 2000 Å^2
- Short-lived and low-stability complexes -> smaller interfaces (1150–1200 Å^2)

- large surfaces (2000 to 4600 Å^2) ->
  - proteases and particular inhibitors
  - G-proteins and other components of the signal transduction system

- Protein-small molecule interaction surfaces have an area of 300 to 1000 Å^2.

Surface of Interaction (II)

- Surfaces of PPIs are generally flat and lack the grooves and pockets that are present at the surfaces of proteins that bind to small molecules.

- PPI surfaces are generally hydrophobic in nature.

- Only certain hydrophobic spots contribute to the free energy of binding and help to hold the two proteins together.

- Such regions are called hot spots.

Hot Spots

- Hot spots account for less than 50% of the contact area of PPI

- A region of protein surface is called a hot spot when replacement of an amino acid residue by alanine in that spot lowers the free energy of binding by at least 2 kcal/mol

- Analysis of the amino acid composition of hot spots shows that some residues are found more frequently in hot spots (Tyr, Trp, and Arg)

- The hot spots are surrounded by energetically less important residues that separate/prevent bulk water from hot spots

Analysis of Protein Complex

- identification of interface residues/hot spots

- details about the interface
  solvent accessible surface area, shape, complementarity between surfaces, residue interface propensities, hydrophobicity, segmentation and secondary structure, and conformational changes on complex formation

- assignment of protein function

- recognition of specific residue motifs

Structure PPI Data

- The most significant contribution to understanding the PPI surface comes from structural biology via X-ray crystallography or NMR as well as mutational studies

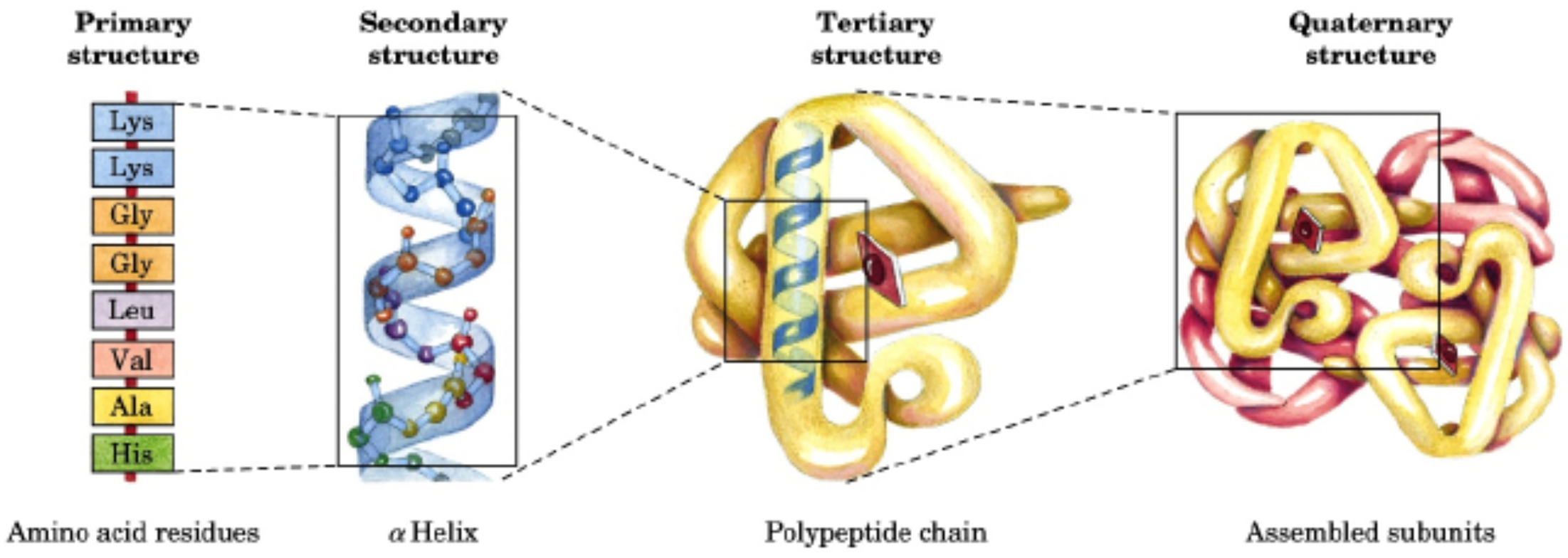
- Prediction of interaction/binding sites

- Prediction of protein-protein complexes

Introduction to Protein Structure Analysis

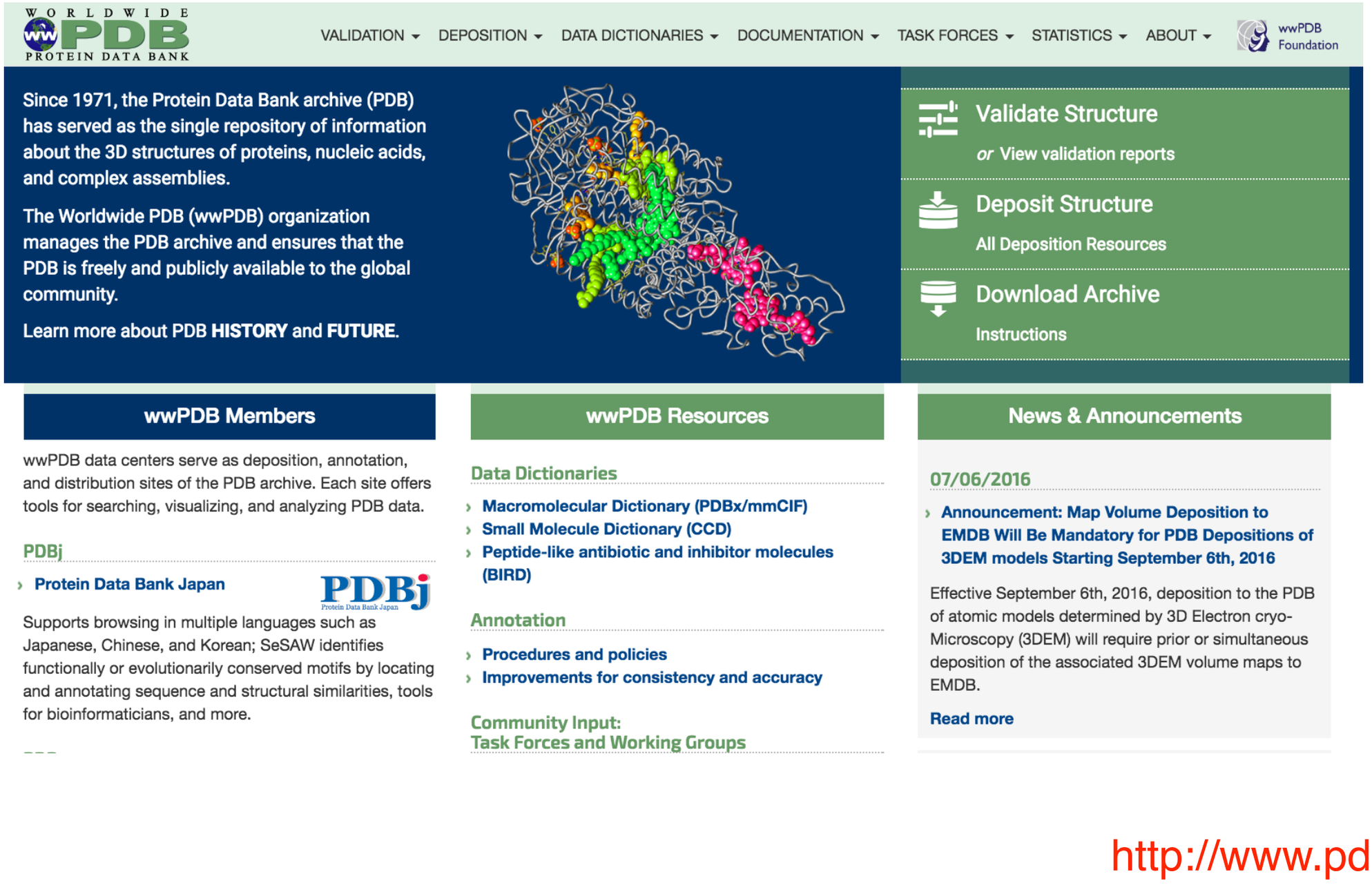
Hierarchical organization of protein structure

Protein structure is defined by four levels of hierarchical organization.



The Protein Data Bank

The largest repository of macromolecular structures obtained mainly by X-ray crystallography and NMR



http://www.pdb.org

Data Summary

Currently more than 120,000 structures have been deposited. Most of them (~100,000) are obtained by X-ray Crystallography.

Exp.Method	Proteins	Nucleic Acids	Protein/NA Complex	Other	Total
X-RAY	100628	1753	5173	4	107558
NMR	10064	1148	235	8	11455
ELECTRON MICROSCOPY	784	30	268	0	1082
HYBRID	90	3	2	1	96
other	174	4	6	13	197
Total	111740	2938	5684	26	120388

PDB data

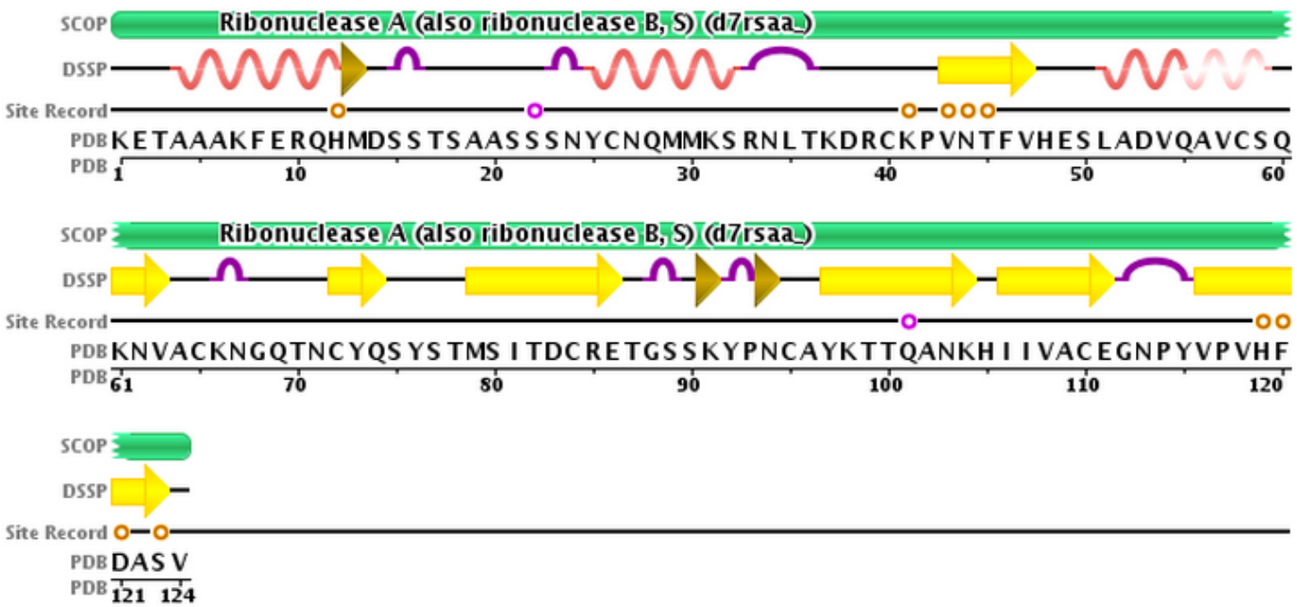
The most important information are the atomic coordinates.

		AT	RES	CH	POS	X	Y	Z			
ATOM	2145	N	GLU	B	10	150.341	72.309	103.145	1.00	99.90	N
ATOM	2146	CA	GLU	B	10	150.096	71.519	101.907	1.00	99.90	C
ATOM	2147	C	GLU	B	10	150.425	70.046	102.190	1.00	99.90	C
ATOM	2148	O	GLU	B	10	151.326	69.770	102.983	1.00	99.90	O
ATOM	2149	CB	GLU	B	10	150.963	72.057	100.790	1.00	99.90	C
ATOM	2150	N	PRO	B	11	149.661	69.092	101.595	1.00	99.90	N
ATOM	2151	CA	PRO	B	11	149.856	67.644	101.778	1.00	99.90	C
ATOM	2152	C	PRO	B	11	150.783	66.845	100.844	1.00	99.90	C
ATOM	2153	O	PRO	B	11	151.938	66.593	101.185	1.00	99.90	O
ATOM	2154	CB	PRO	B	11	148.425	67.108	101.722	1.00	99.90	C
ATOM	2155	CG	PRO	B	11	147.816	67.948	100.672	1.00	99.90	C
ATOM	2156	CD	PRO	B	11	148.333	69.350	101.000	1.00	99.90	C
ATOM	2157	N	SER	B	12	150.258	66.422	99.691	1.00	99.90	N
ATOM	2158	CA	SER	B	12	150.965	65.585	98.710	1.00	99.90	C
ATOM	2159	C	SER	B	12	150.922	64.167	99.292	1.00	99.90	C
ATOM	2160	O	SER	B	12	150.493	63.222	98.632	1.00	99.90	O
ATOM	2161	CB	SER	B	12	152.410	66.042	98.440	1.00	99.90	C
ATOM	2162	OG	SER	B	12	152.907	65.499	97.219	1.00	99.90	O

http://deposit.rcsb.org/adit/docs/pdb_atom_format.html

The Bovine Ribonuclease A

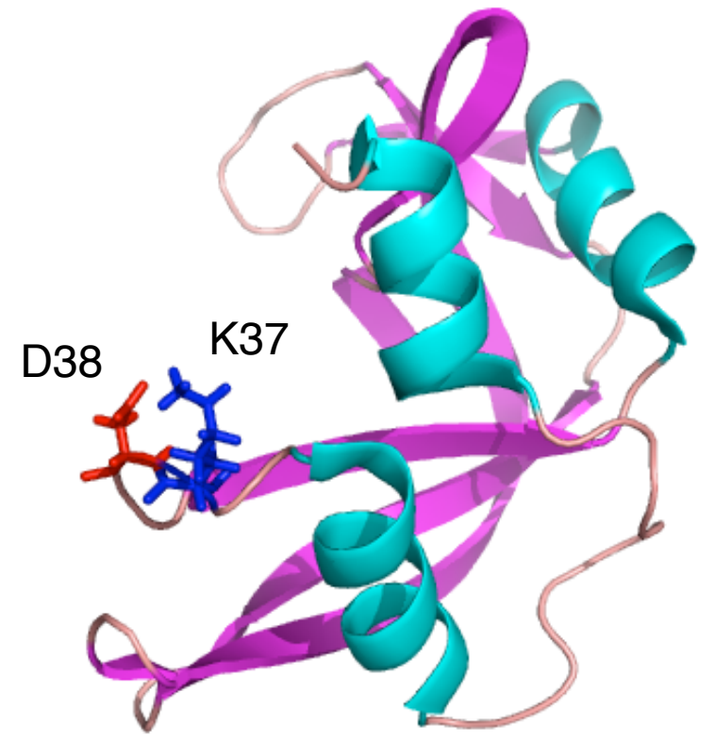
Ribonuclease A (RNase A) is a pancreatic ribonuclease that cleaves single-stranded RNA.

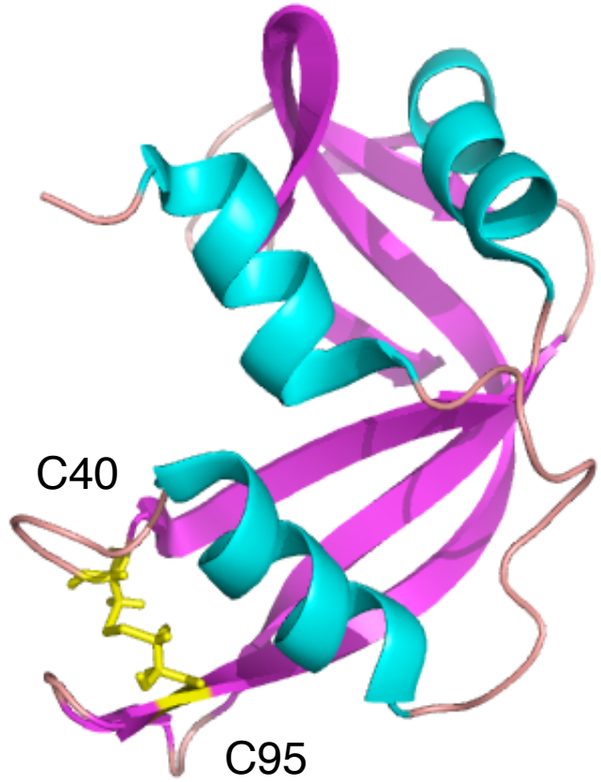


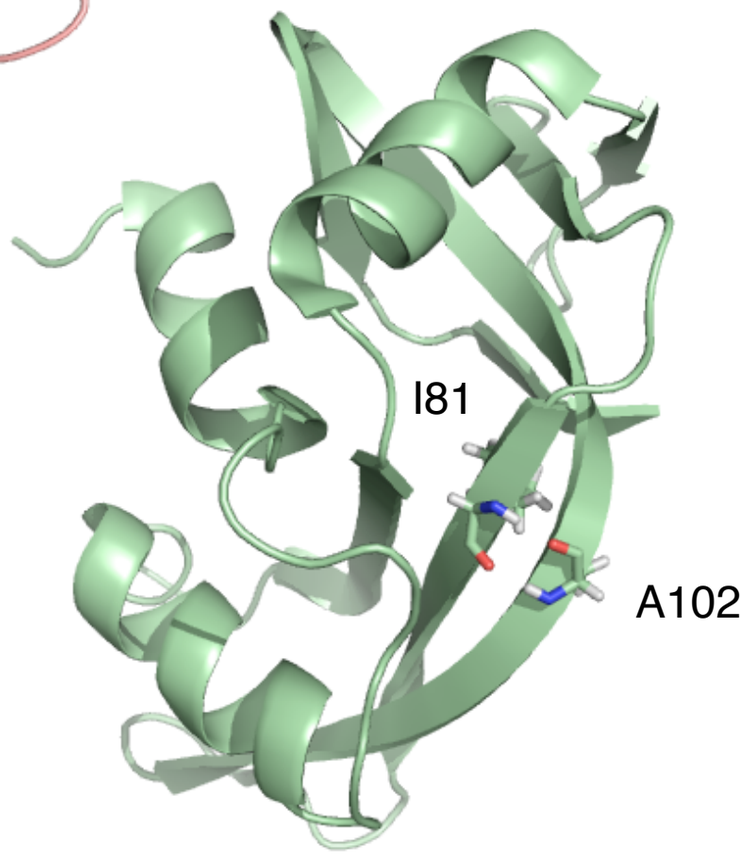
Wlodawer, A. et al. (1988) Biochemistry 27: 2705-2717

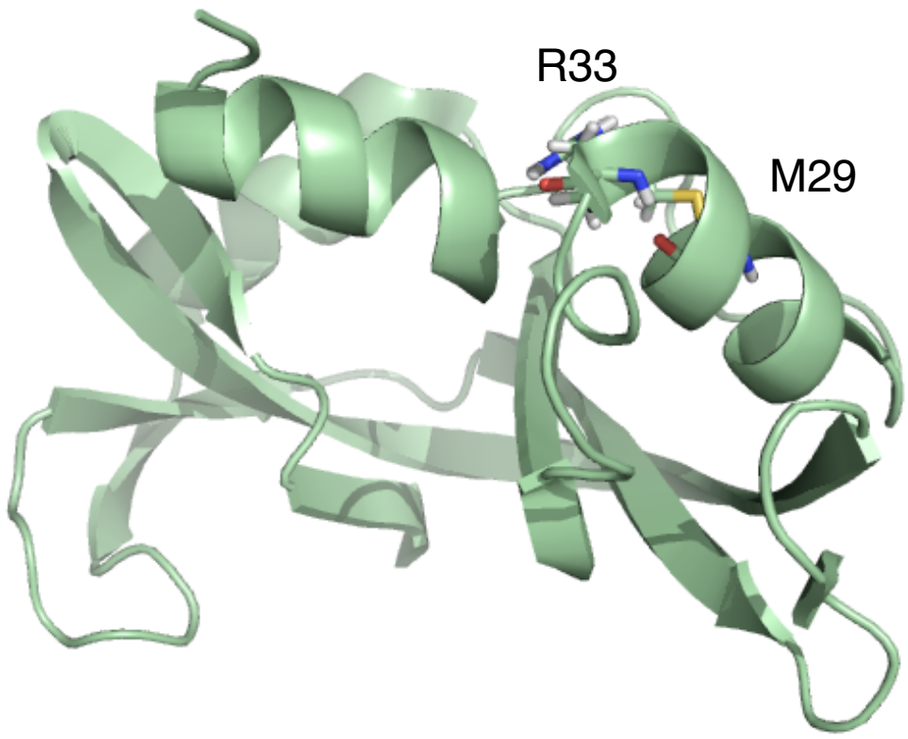
Bonds and interactions

Examples of salt bridge, disulfide bond and hydrogen bonds in ribonuclease A





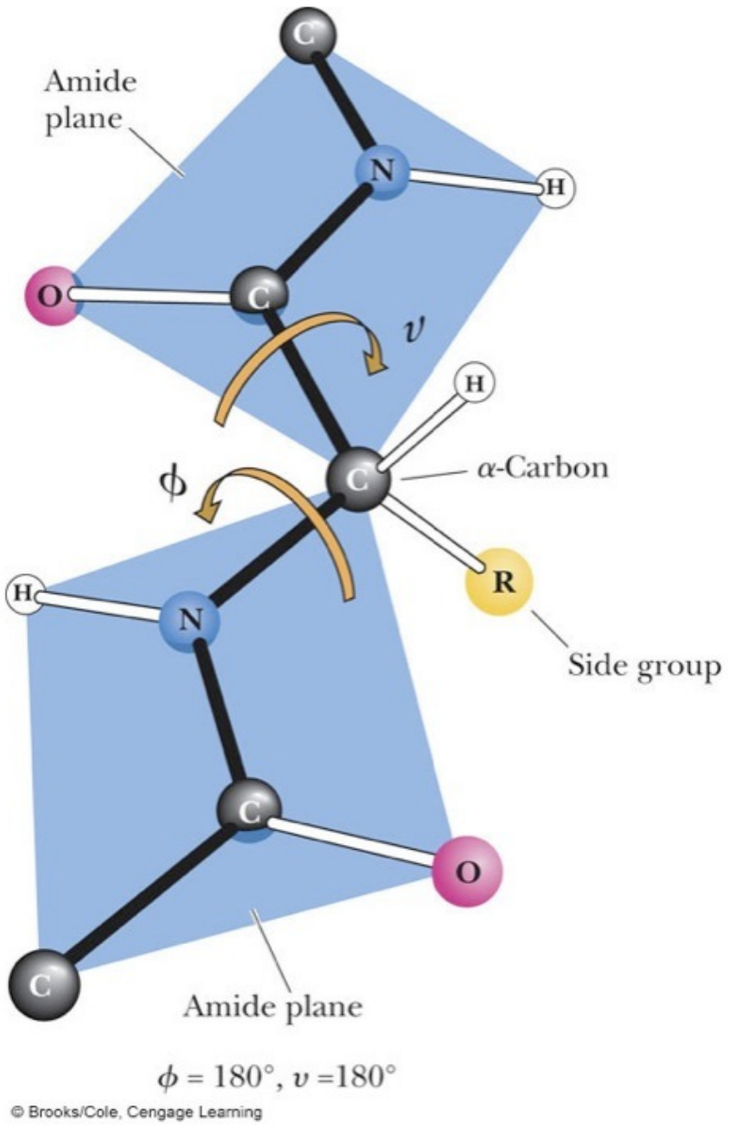




Defining protein structure

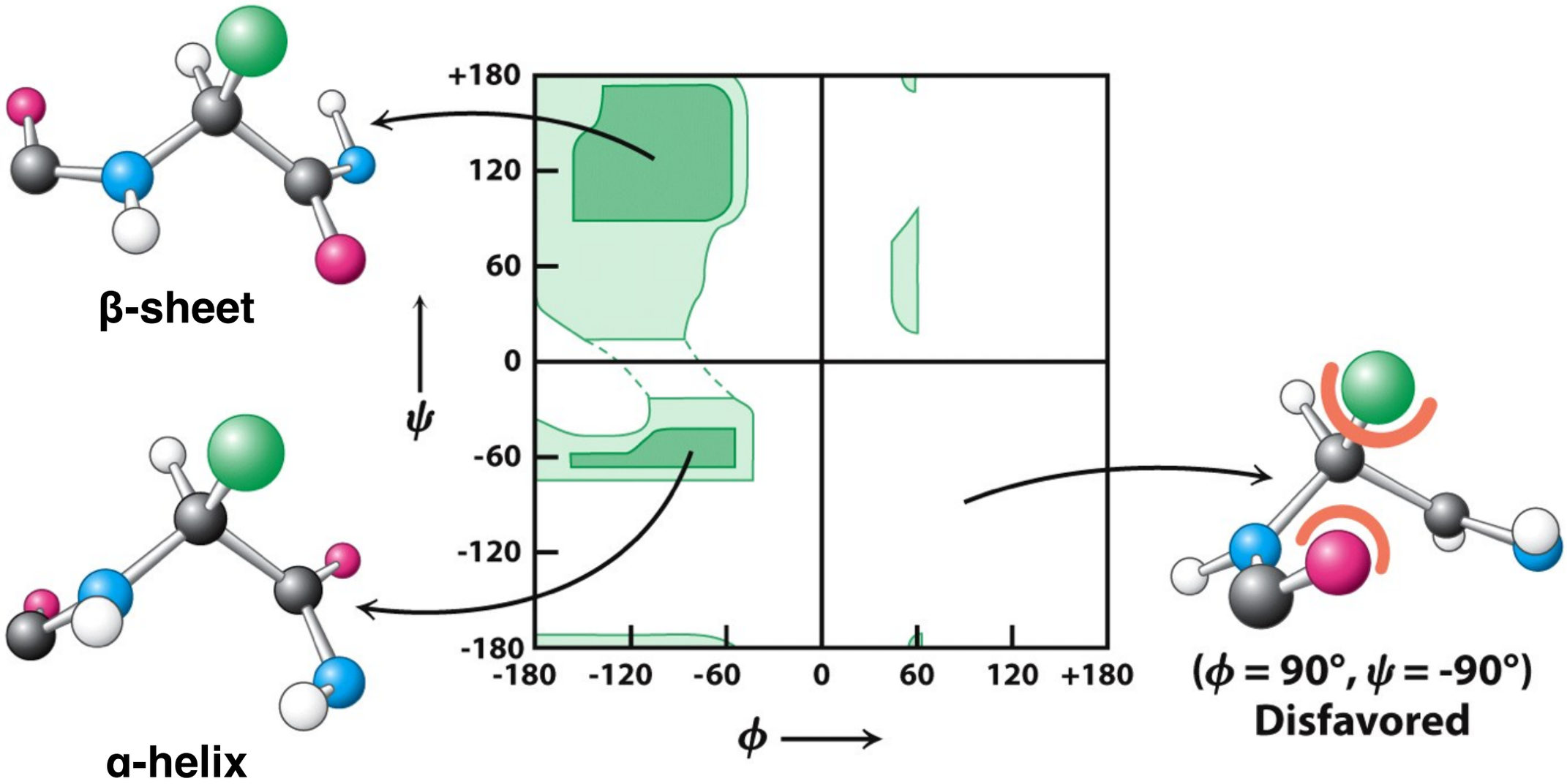
Basic information for the characterization of the protein three-dimensional structures are:

- ϕ, ψ values for each residue in the protein chain
- secondary structure
- solvent accessible area



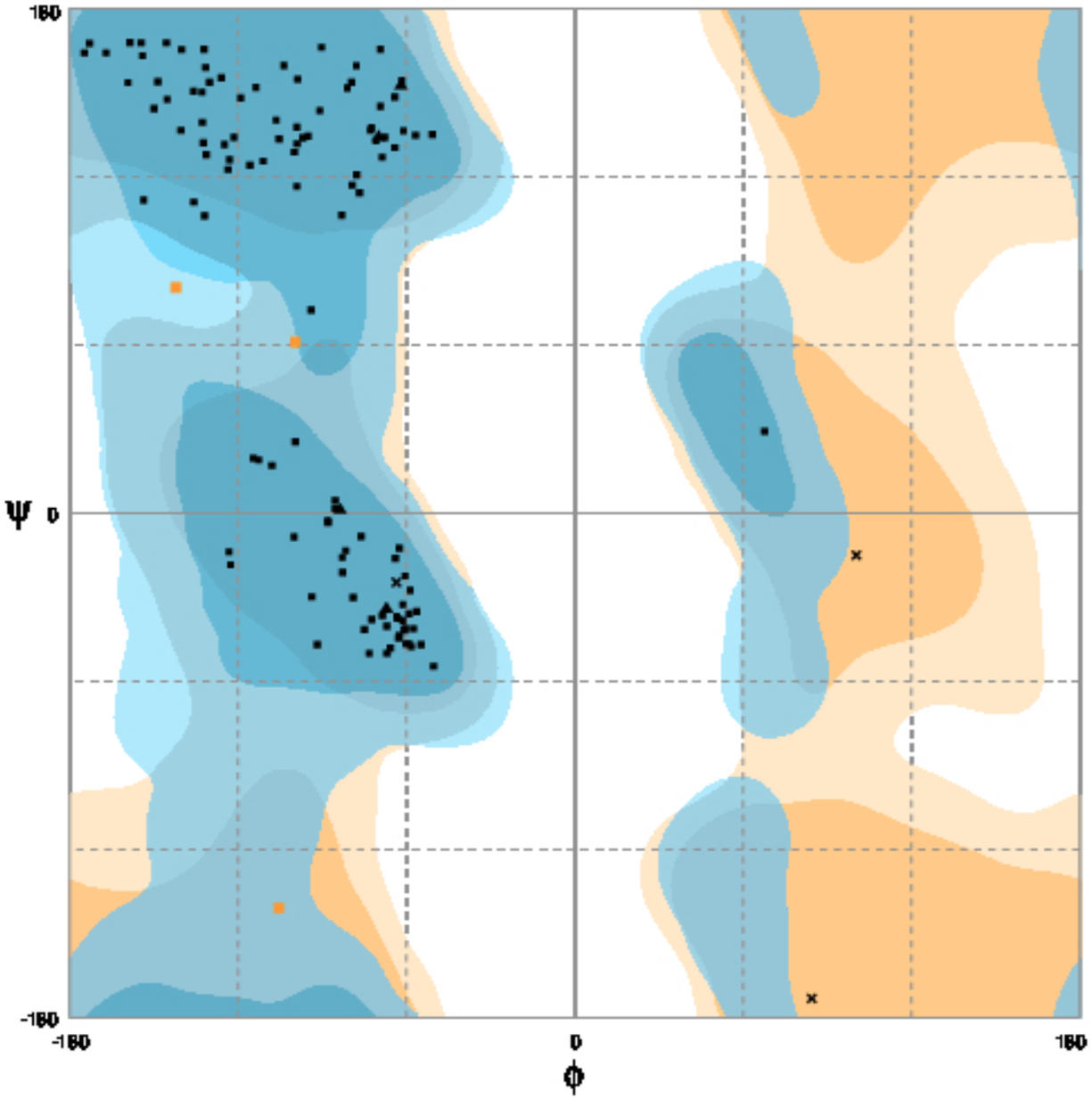
Ramachandran Plot

The backbone of the protein structure can be defined providing the list of ϕ, ψ angles for each residue in the chain.



Berg JM et al. (2012). Biochemistry VII Ed.

Ramachandran plot



http://mordred.bioc.cam.ac.uk/~rapper/rampage2.php

DSSP program

Program that implements the algorithm “Define Secondary Structure of Proteins”.

The method calculates different features of the protein structure such as the ϕ, ψ angles for each residue, its secondary structure and the solvent accessible area.

#	RESIDUE	AA	STRUCTURE	BP1	BP2	ACC	...	PHI	PSI	X-CA	Y-CA	Z-CA
1	10	B	E	0	0	153	...	360.0	144.2	150.1	71.5	101.9
2	11	B	P +	0	0	83	...	-90.2	-84.0	149.9	67.6	101.8
3	12	B	S S >> S+	0	0	60	...	77.6	-51.1	151.0	65.6	98.7
4	13	B	A T 34 S+	0	0	6	...	-82.3	73.7	151.3	62.7	101.2
5	14	B	D T 3> S+	0	0	39	...	-154.6	-41.3	147.5	62.2	100.9
6	15	B	W H <> S+	0	0	170	...	-60.8	-41.6	148.0	61.1	97.3
7	16	B	L H X S+	0	0	0	...	-62.9	-38.5	150.2	58.6	98.9
8	17	B	A H > S+	0	0	3	...	-62.0	-58.1	147.4	57.5	101.3
9	18	B	T H X S+	0	0	72	...	-56.4	-34.0	144.9	56.8	98.6

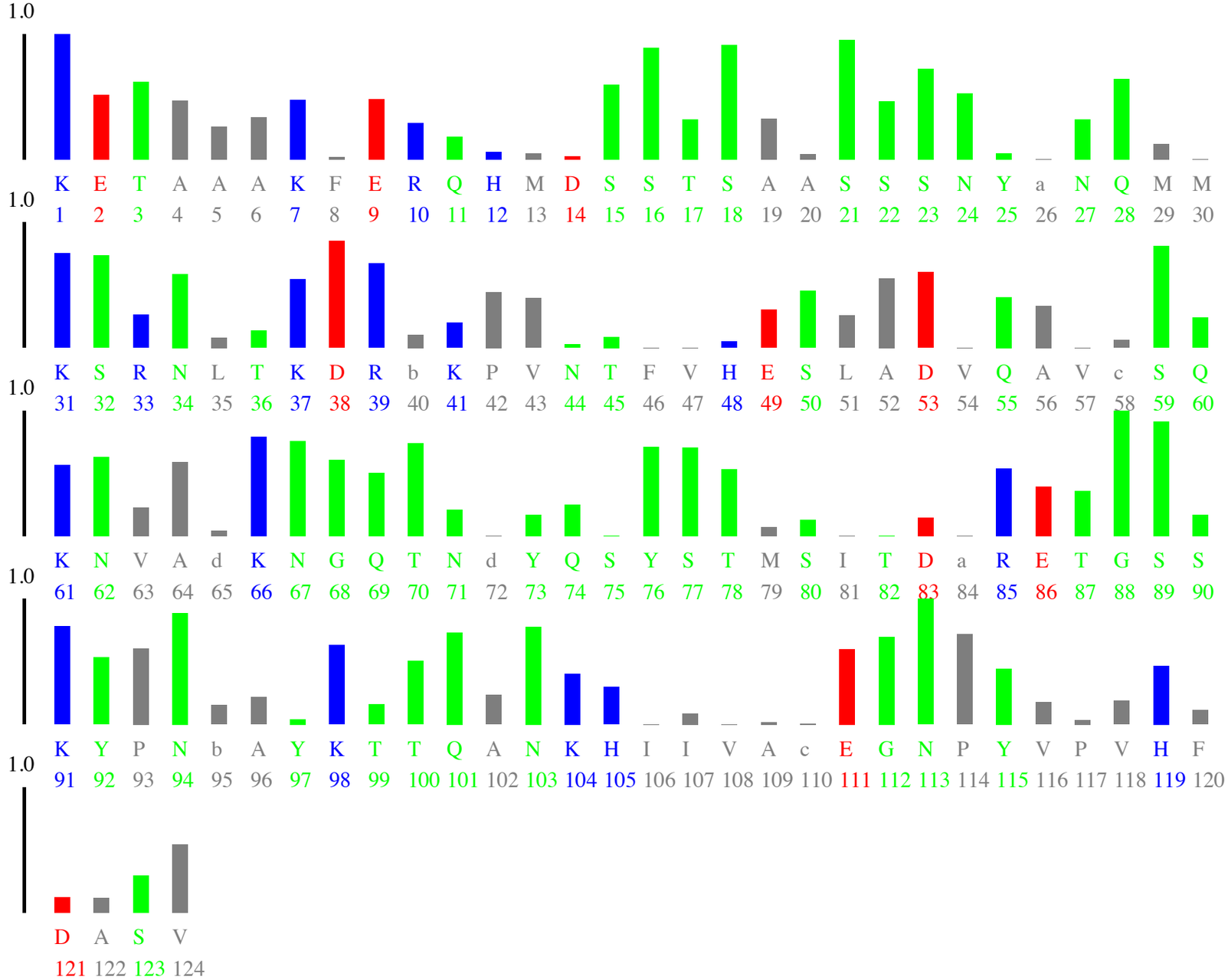
SS **SAA** **PHI** **PSI**

DSSP: ftp://ftp.cmbi.ru.nl/pub/software/dssp
more details at http://www.cmbi.ru.nl/dssp.html

Kabsch W, and Sander C, (1983). Biopolymers. 22 2577-2637.

Relative solvent accessibility

The relative solvent accessible area is obtained dividing the accessible area of the residue by an estimation of the its maximum accessible surface.



http://www.abren.net/asaview/

Exercise

Download the DSSP file of the Ribonuclease A (PDB: 7RSA) from the web (ftp://ftp.cmbi.ru.nl//pub/molbio/data/dssp/7rsa.dssp) and answer the following questions

- What is the total number of residues in helical and extended conformations?
- What is the average value of the ϕ and ψ angles for the residues in helical and extended conformations?
- Are the average values falling the the correct region of the Ramachandran plot?
- Considering the solvent accessibility values reported in the DSSP file, calculate the average relative solvent accessible area for Lysine (205), Valine (142) and Glutamine (198).
- Are this value compatible with the physico-chemical properties of the residues?

SS = col 17, Acc: cols 36-38, Phi: cols 104-109, Psi: cols 110-115