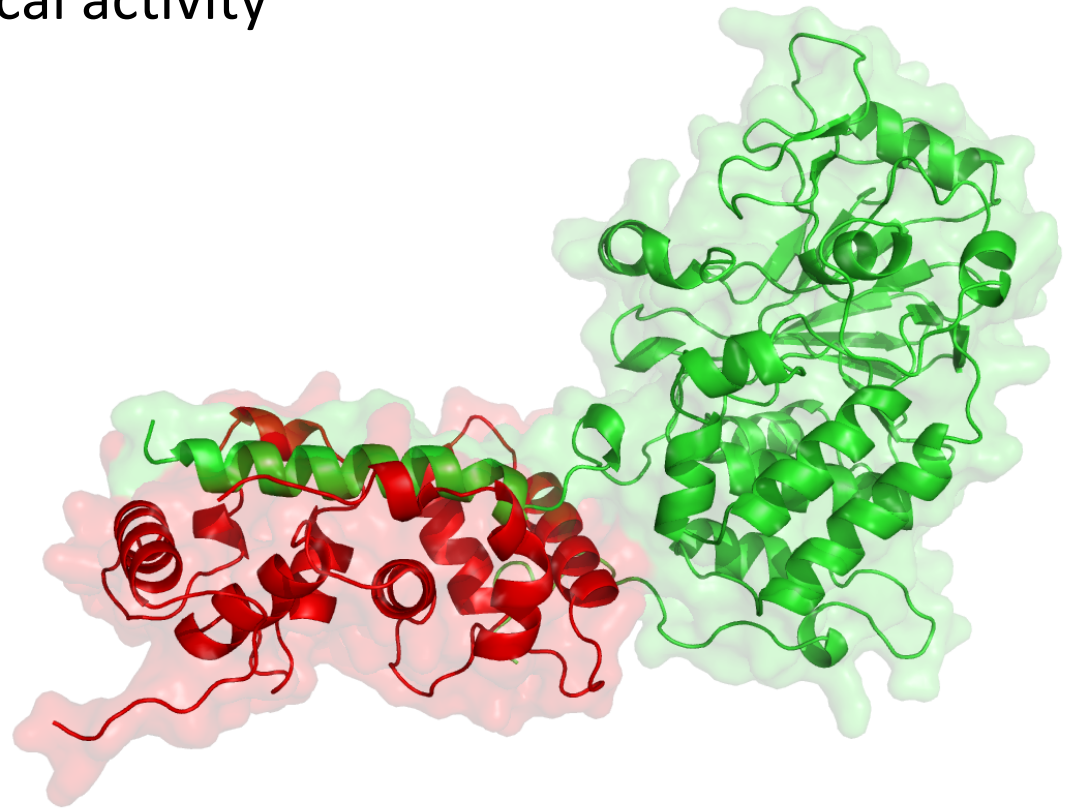


Základy molekulární biofyziky (in English)

Part5: Protein non-folding

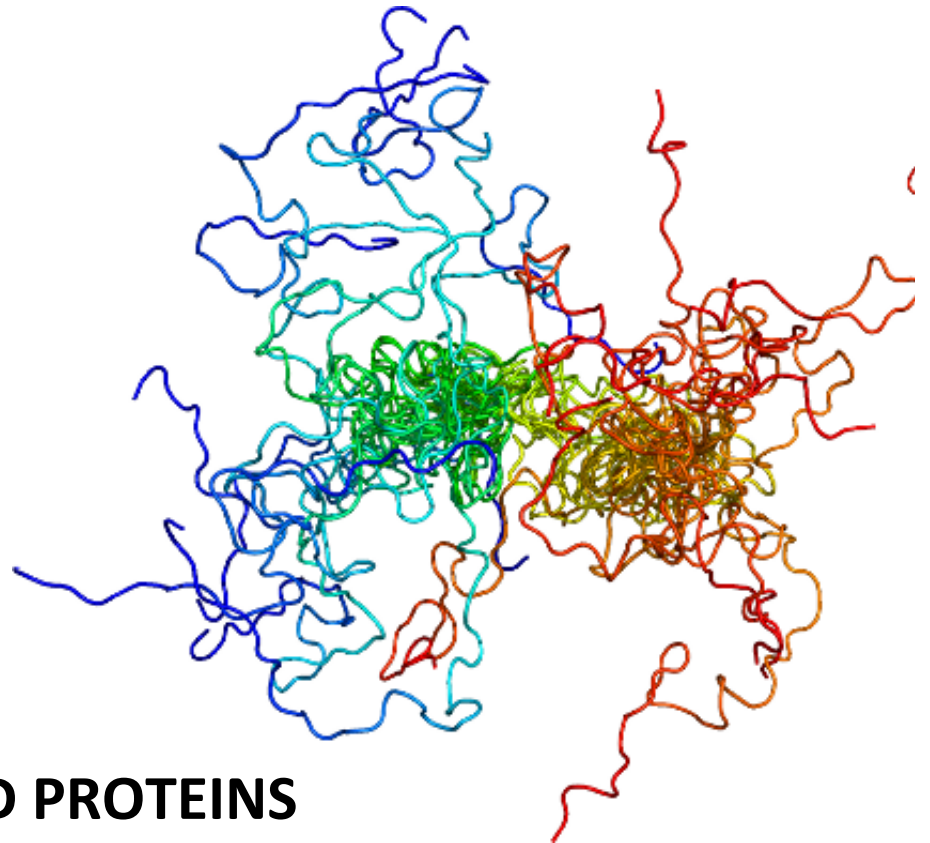
Structure-function paradigm in biochemistry

Proteins need to fold into their 3D structure to have biological activity



... the picture is not so simple

NOT ALL proteins need to fold into their 3D structure
to have biological activity



INTRINSICALLY DISORDERED PROTEINS

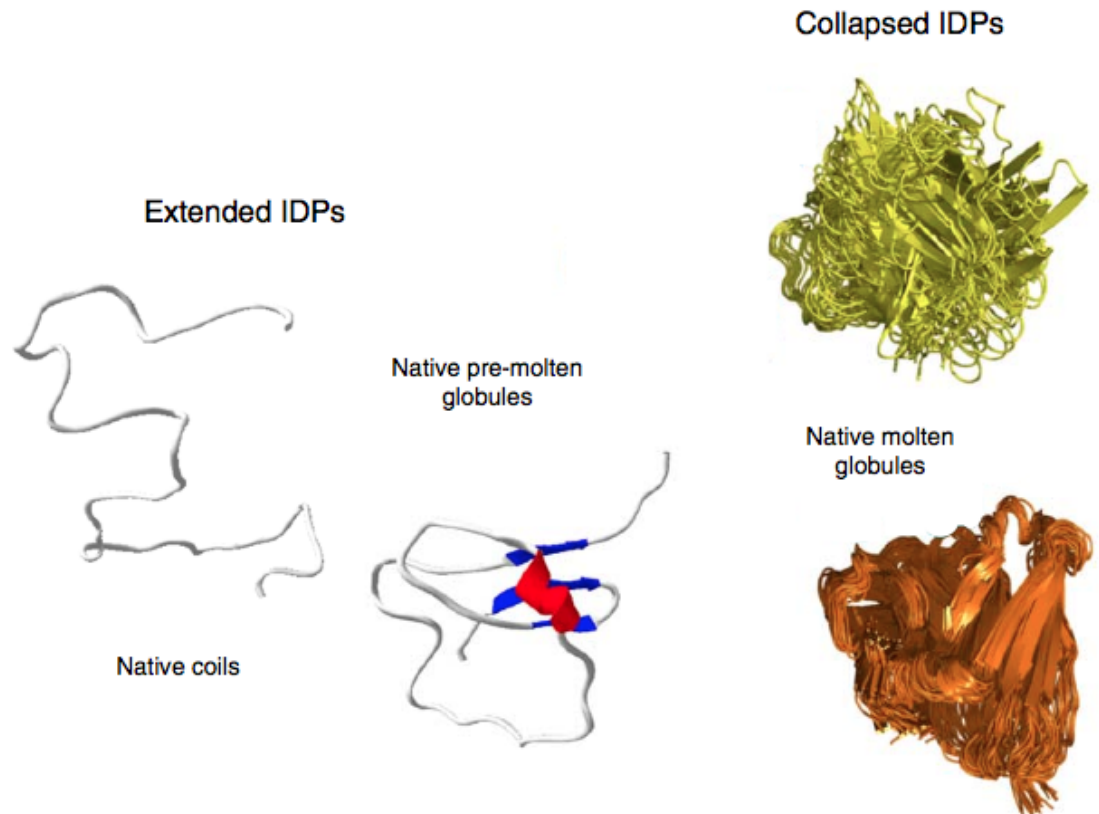
INTRINSICALLY DISORDERED PROTEINS (IDP)

Definition:

IDP proteins characterized by lack of stable **tertiary** structure.

Synonyms:

Intrinsically unfolded proteins,
natively unfolded proteins,
intrinsically unstructured proteins

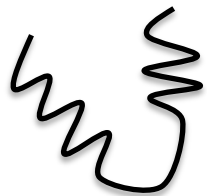


DISORDERED PROTEIN REGIONS

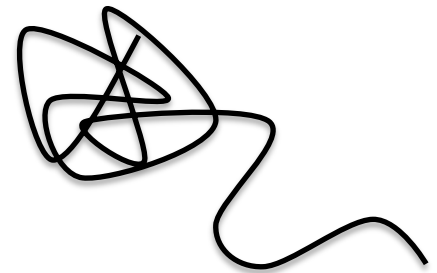
Definition:

Functional protein regions longer than 30 AA characterized by lack of stable **secondary & tertiary** structure.

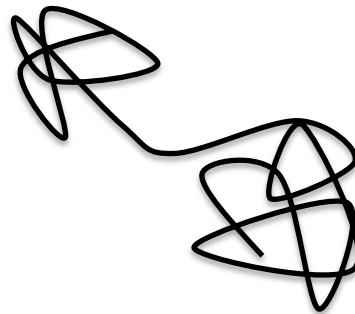
Loops between 2^o elements



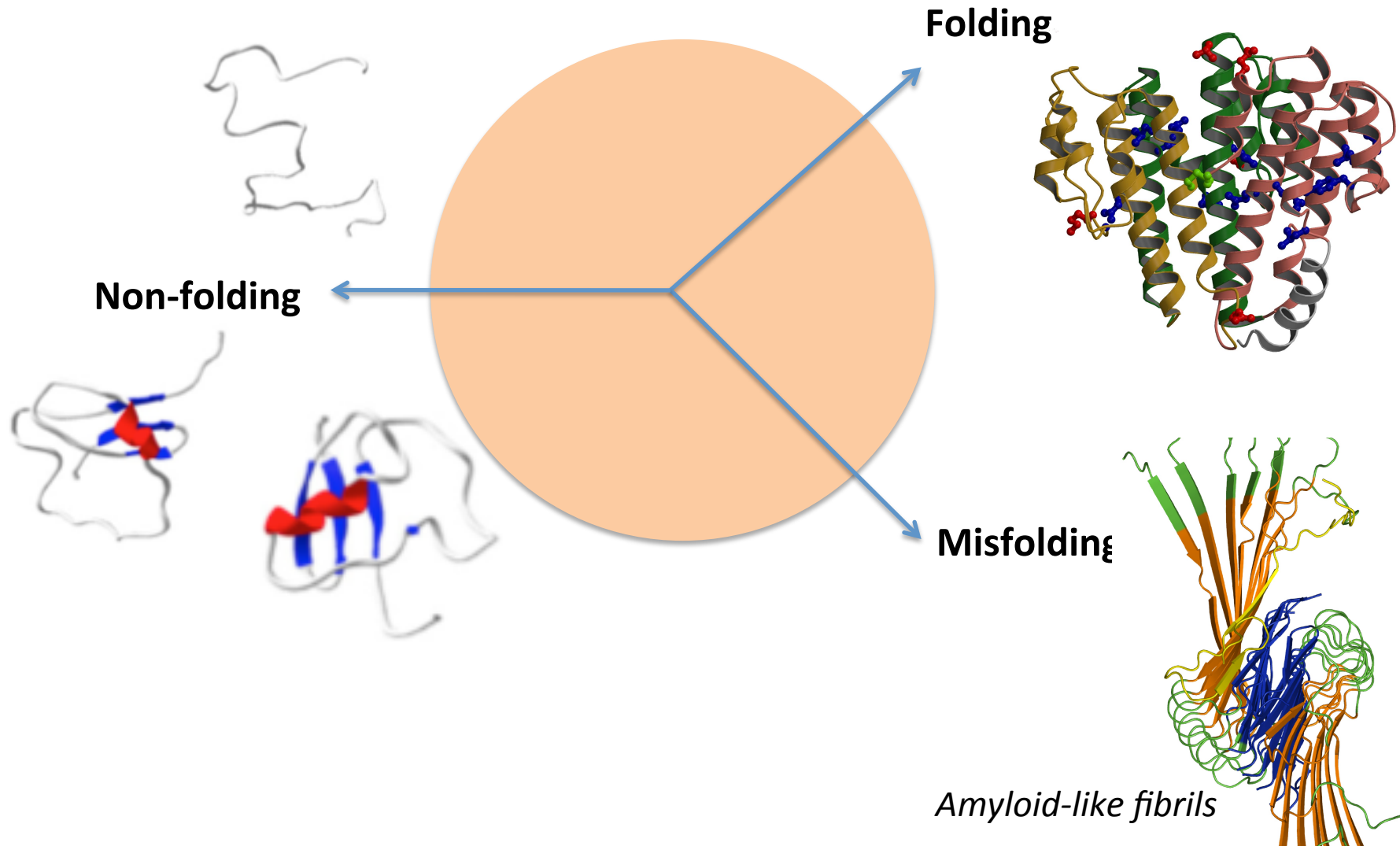
Tails of the folded proteins



Linkers between folded domains

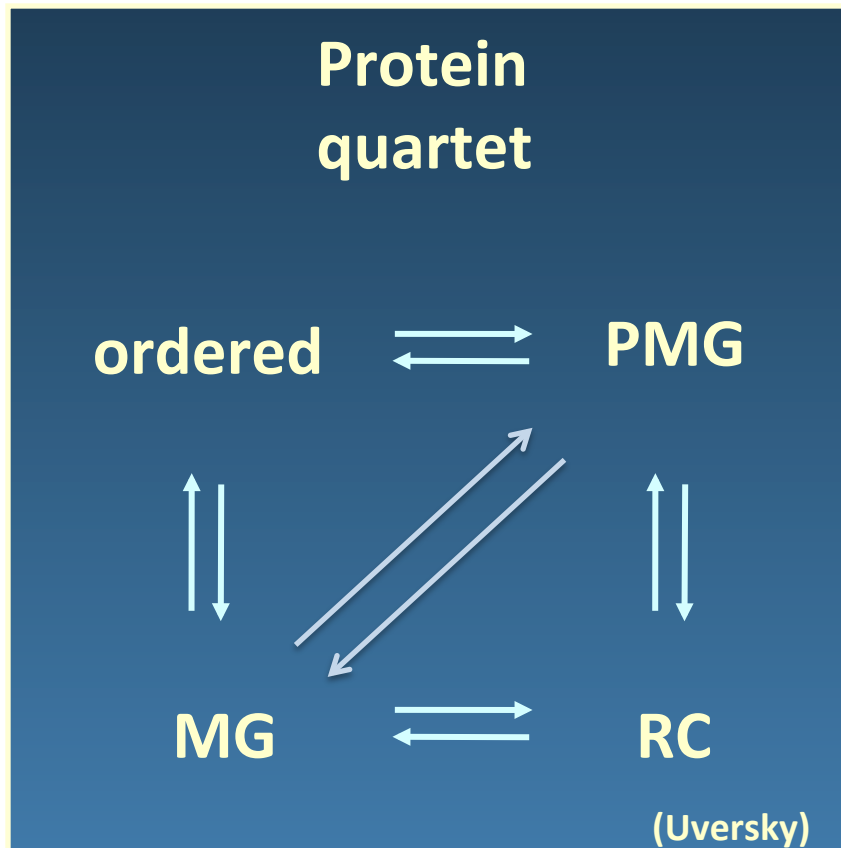


The modern understanding of the fate of a polypeptide chain inside a cell.



New model of structure-function relationships

Function can arise from any of these conformations or transitions between them.



PMG pre-molten globule
MG molten globule
RC random coil

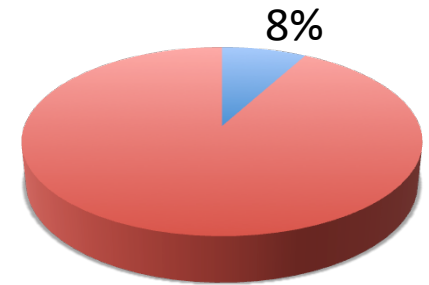
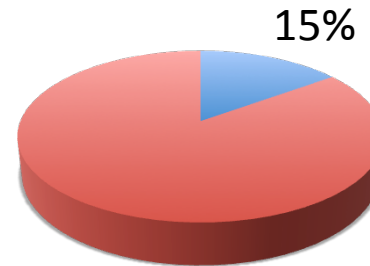
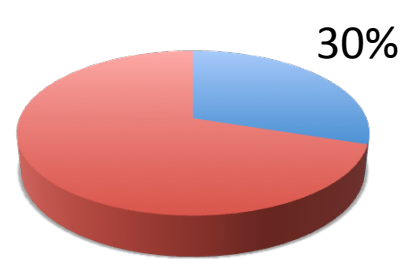
Disorder in complete genomes [%]

L > 30

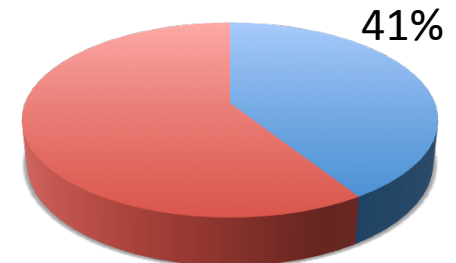
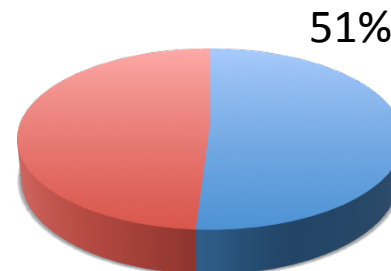
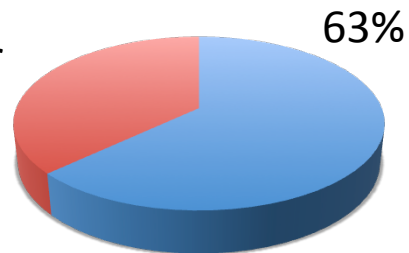
L > 40

L > 50

Bacillus subtilis
(Bacteria)



Drosophila melanogaster
(Eukaryota)

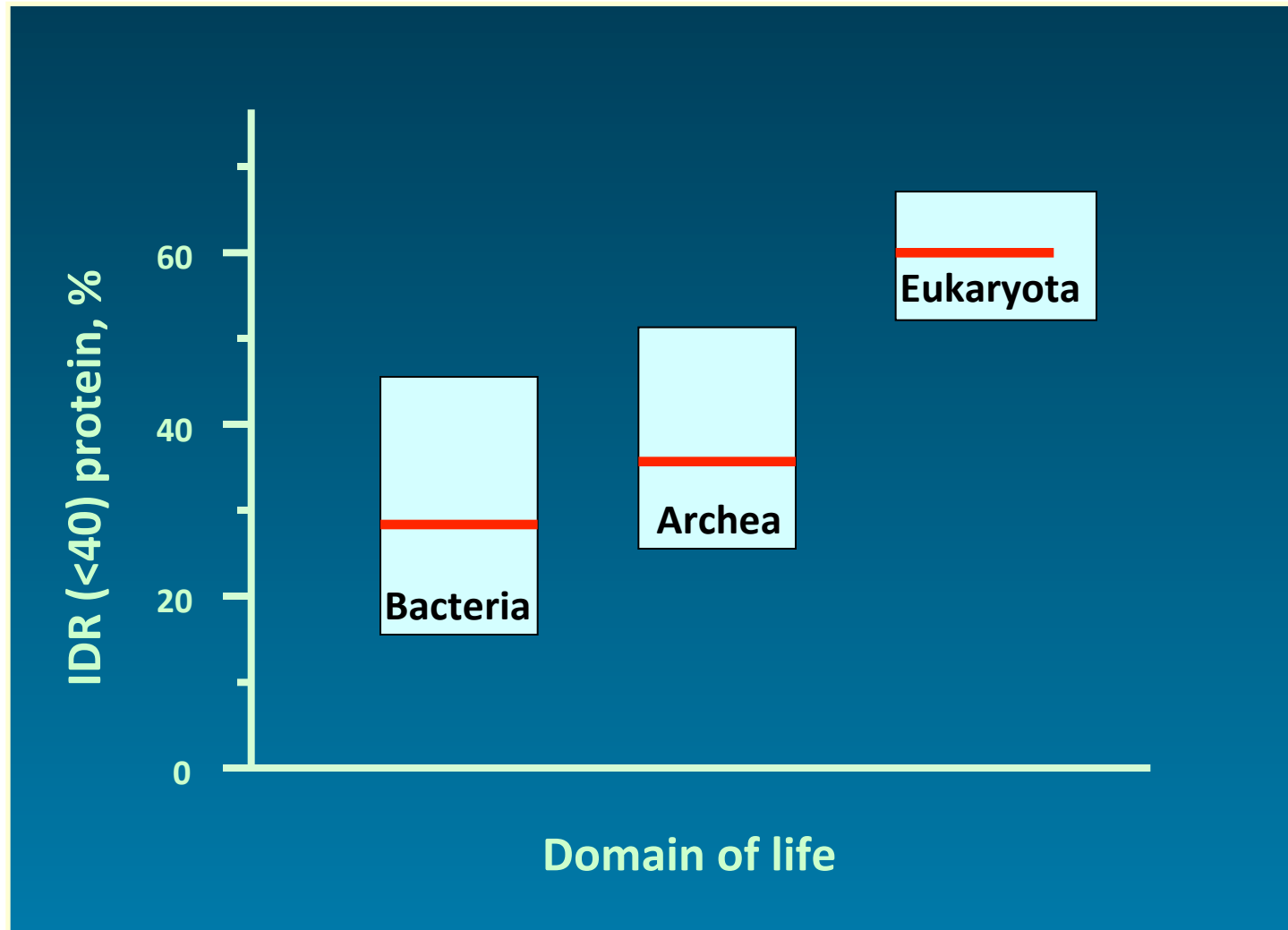


disordered

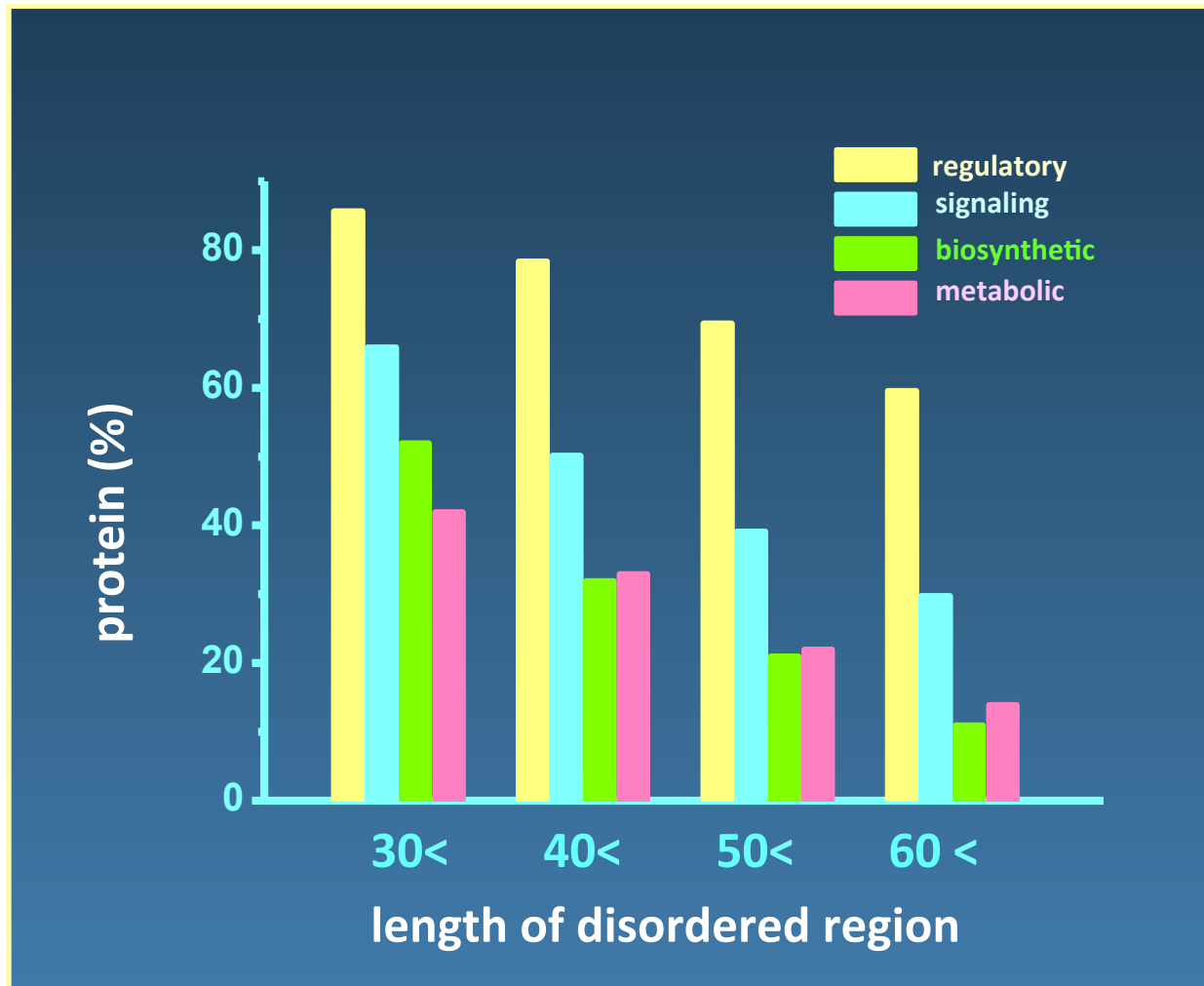


ordered

Structural disorder: evolutionary success

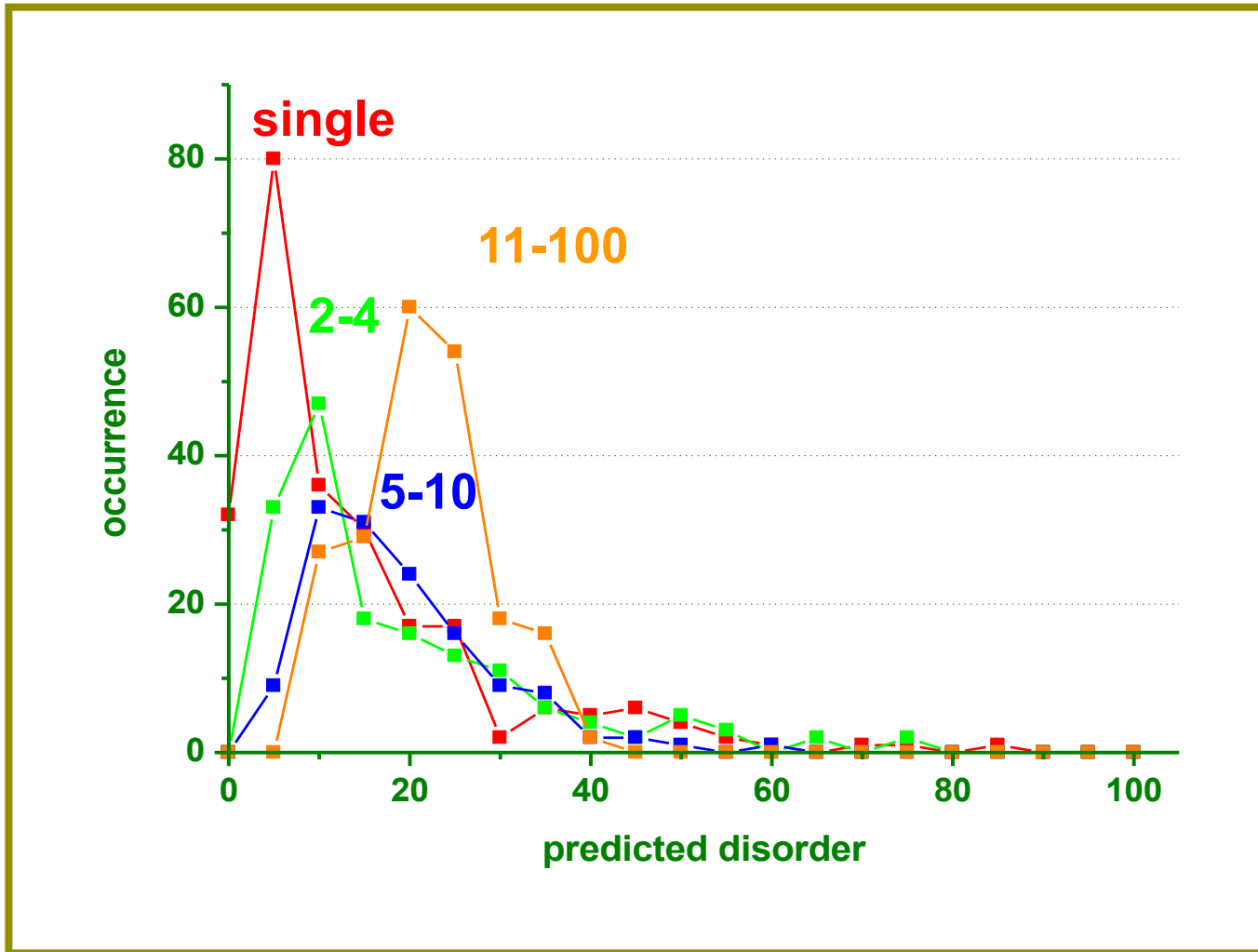


Disorder prevails in regulatory proteins

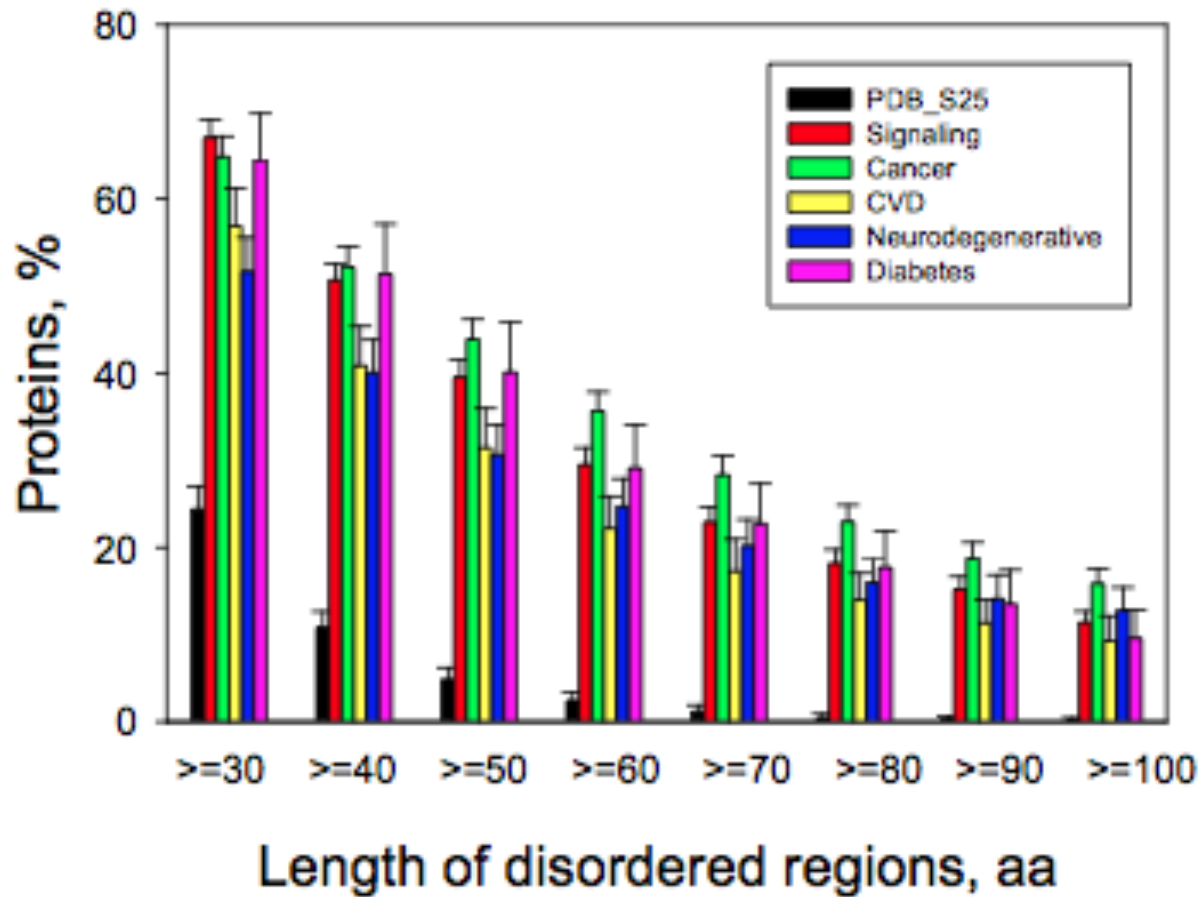


Disorder correlates with complex size

Larger protein complexes have more disorder



IDP regions are abundant in disease related proteins.

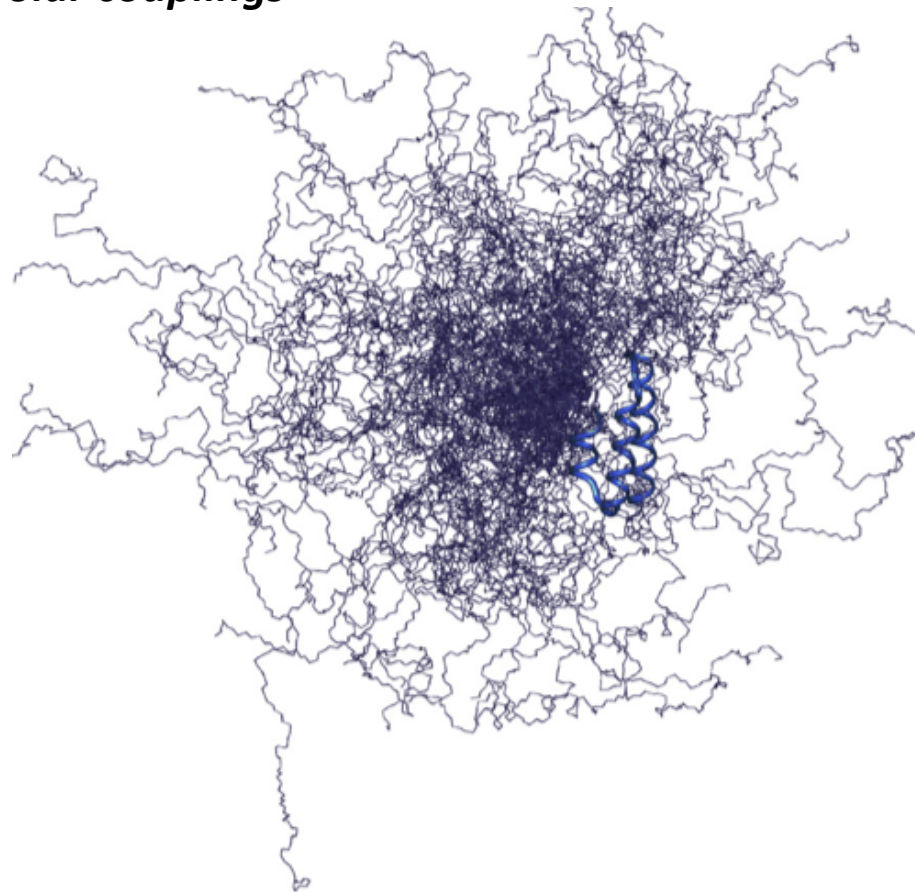
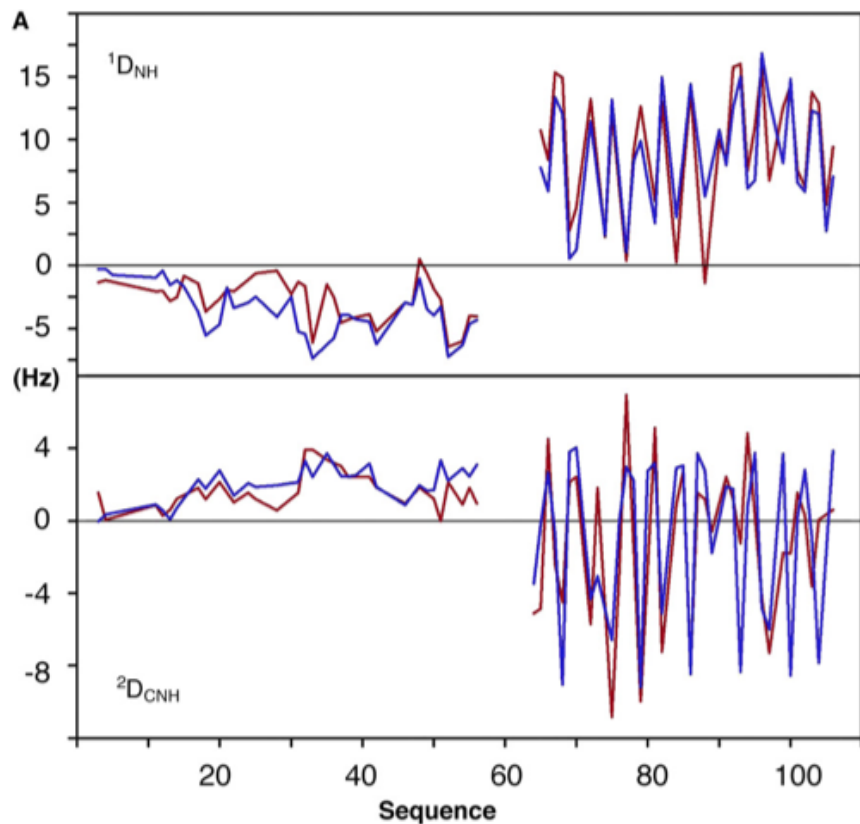


Physico-chemical properties of intrinsically unfolded proteins

Stereo-chemical properties of IUP

IUPs are dynamic, but their structures are NOT random

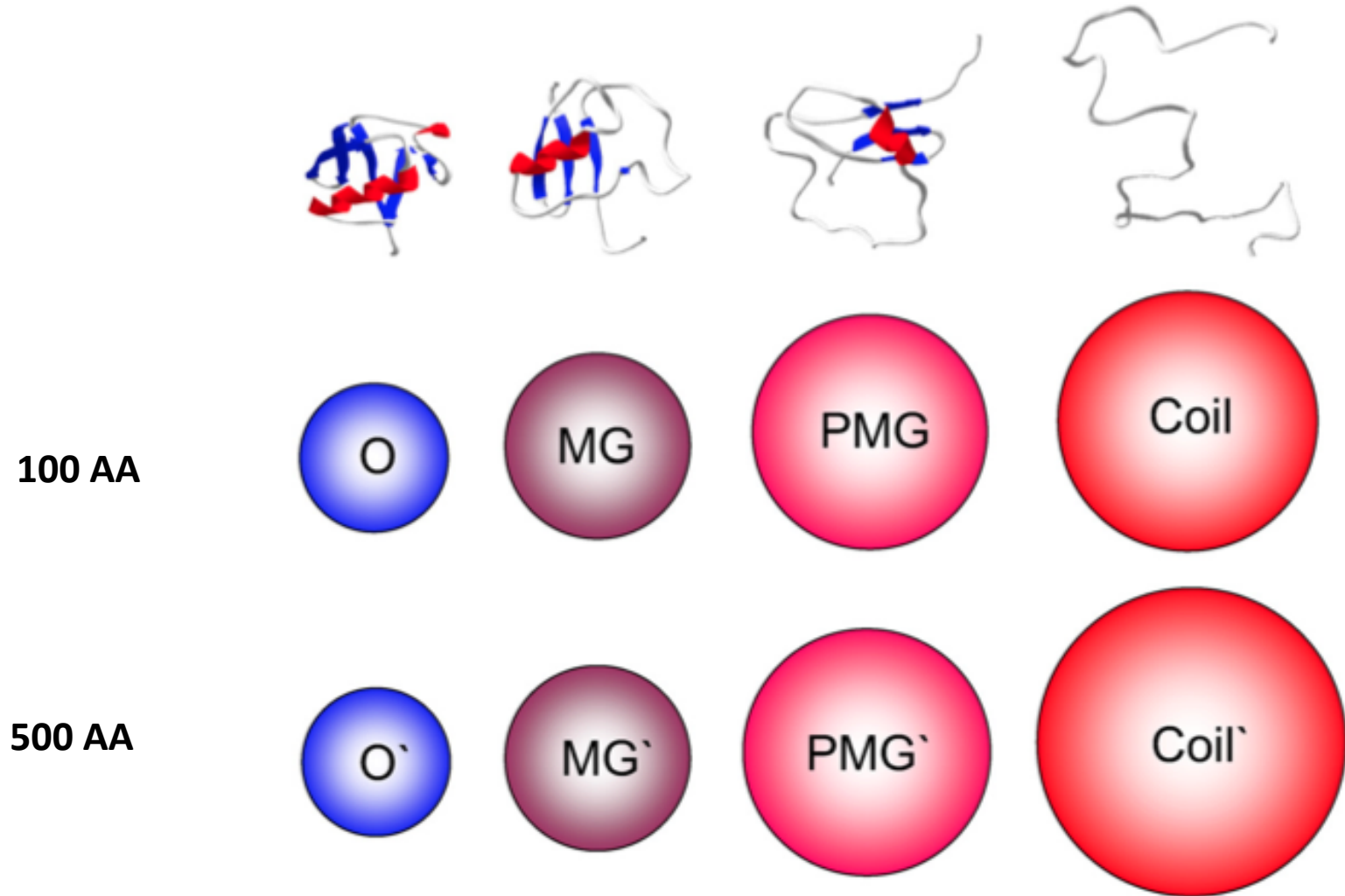
Experimental evidence: non-zero residual dipolar couplings



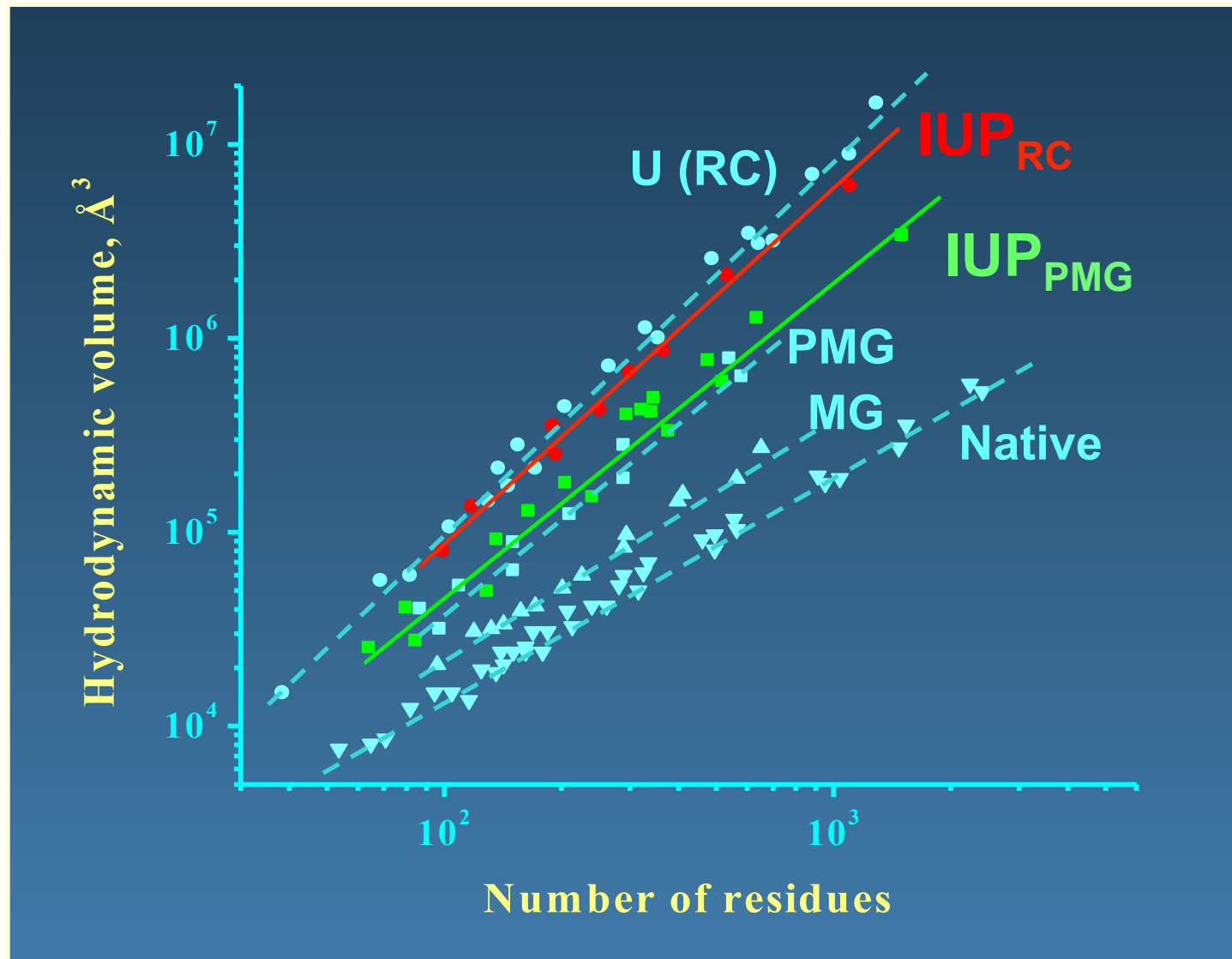
$$D_{ij} = -\frac{\gamma_i \gamma_j \hbar \mu_0}{4\pi^2 r^3} \left\langle \frac{(3\cos^2 \theta - 1)}{2} \right\rangle$$

Physico-chemical properties of IUP

IUPs have a low “density” connected with large radius of gyration.



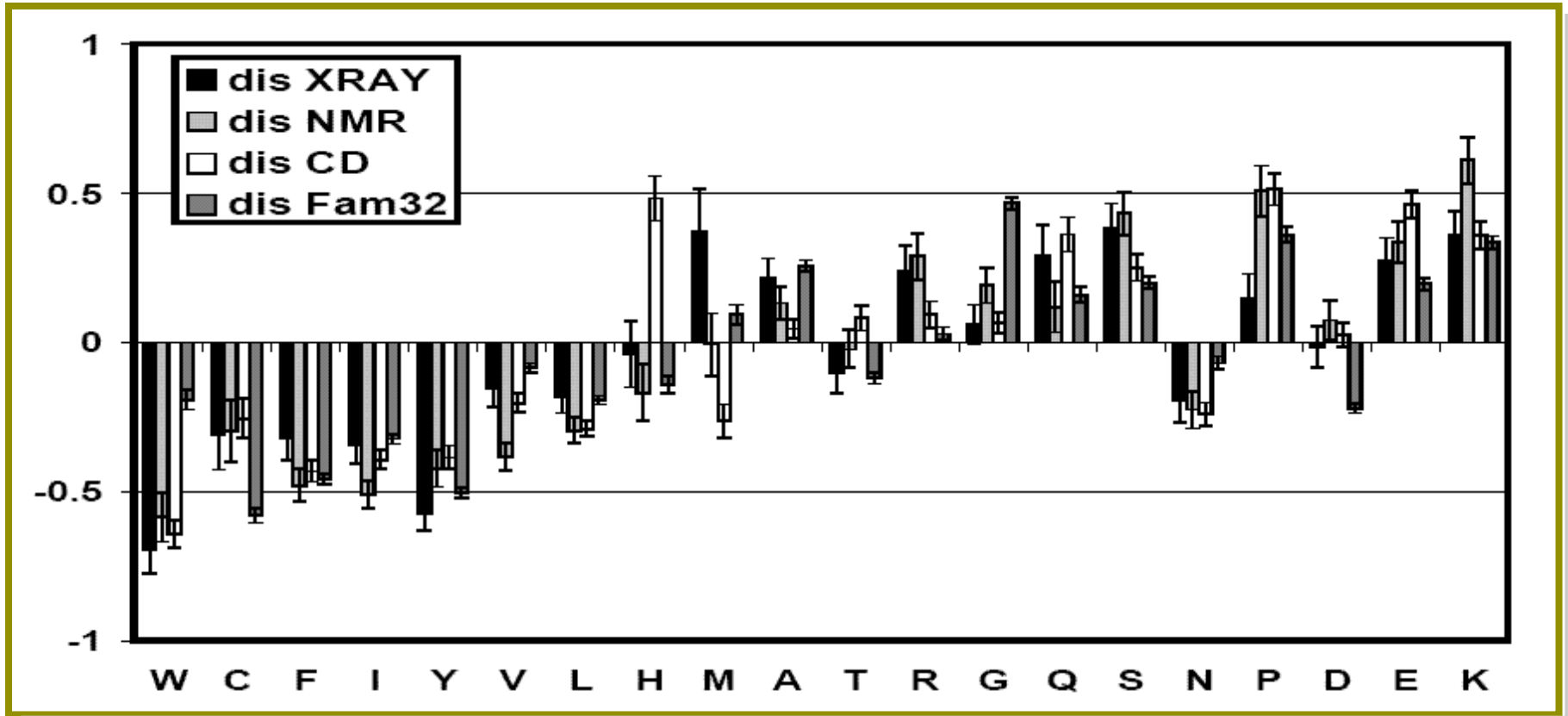
Overall structure of IUPs



Sequence signatures of disorder

- low content of bulky, hydrophobic amino acids
- high proportion of polar and charged amino acids.
- often low complexity sequences, i.e. overrepresentation of a few residues.

Sequence signatures of disorder

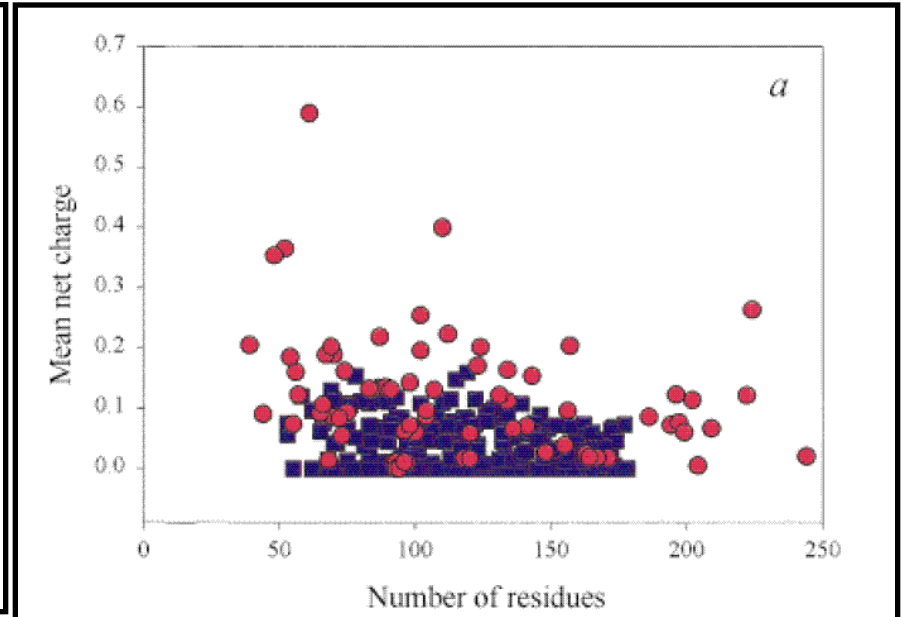
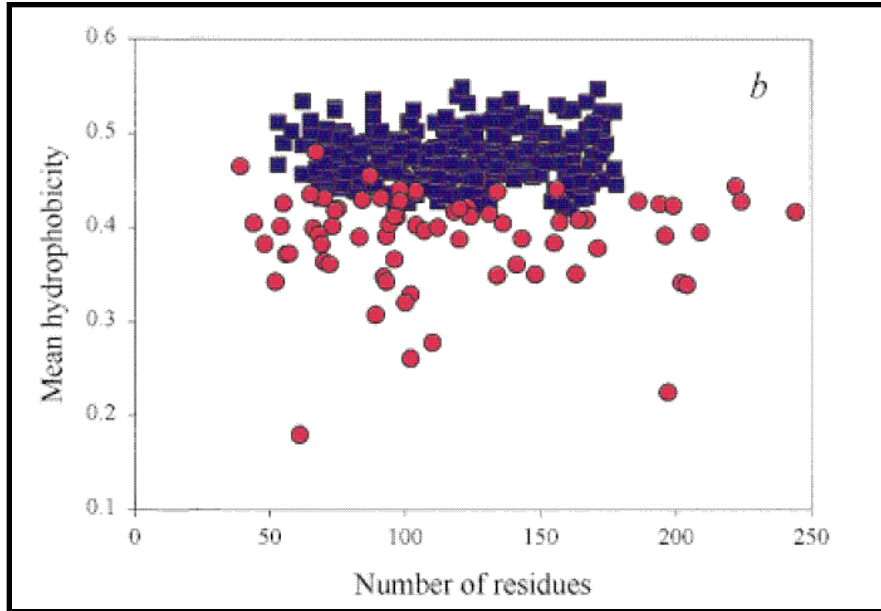


order-promoting

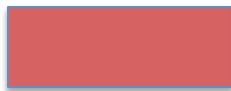


disorder-promoting

IUPs: AA charge & hydrophobicity

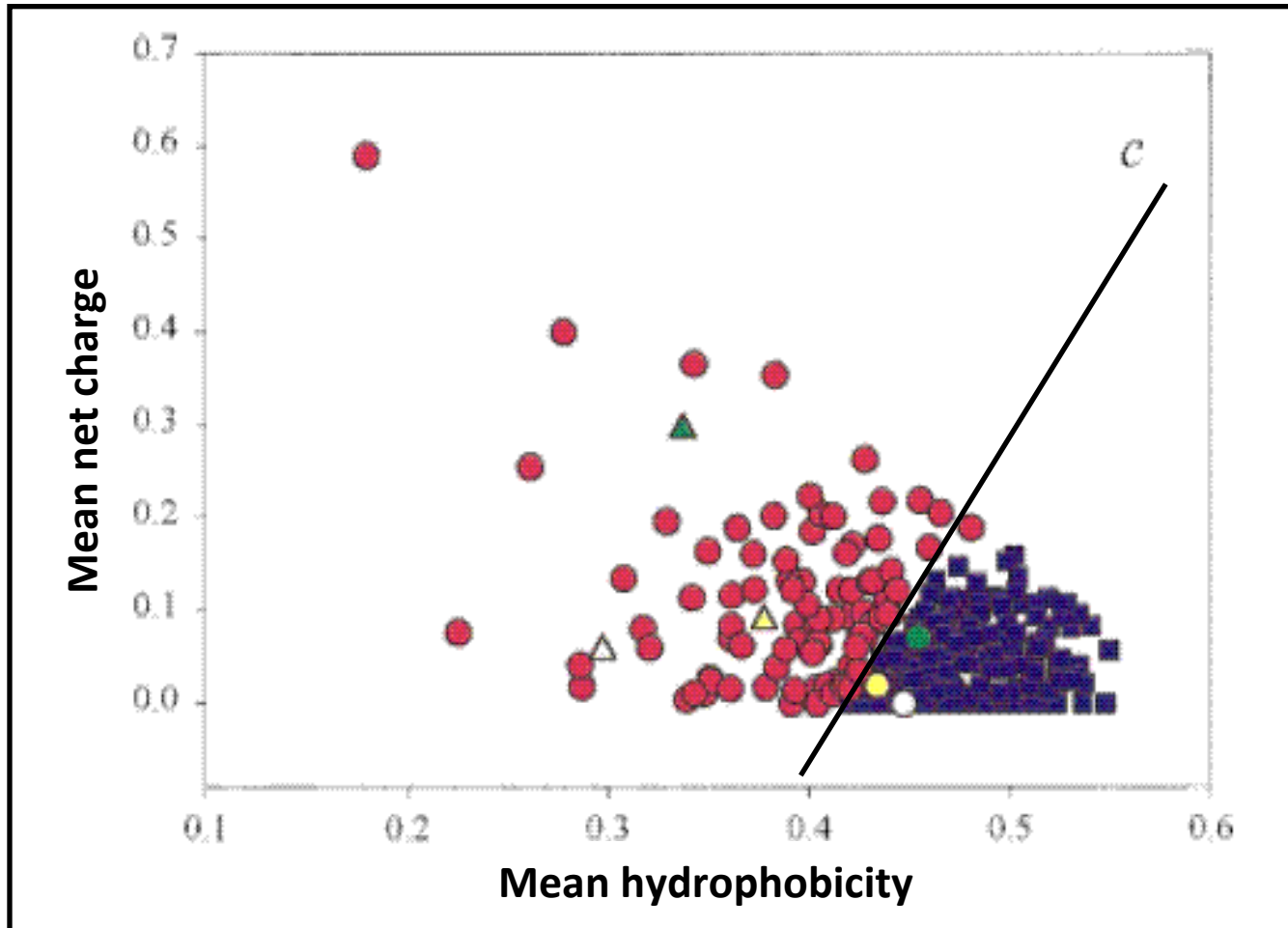


Folded



Unfolded

AA charge & hydrophobicity (Uversky plot)

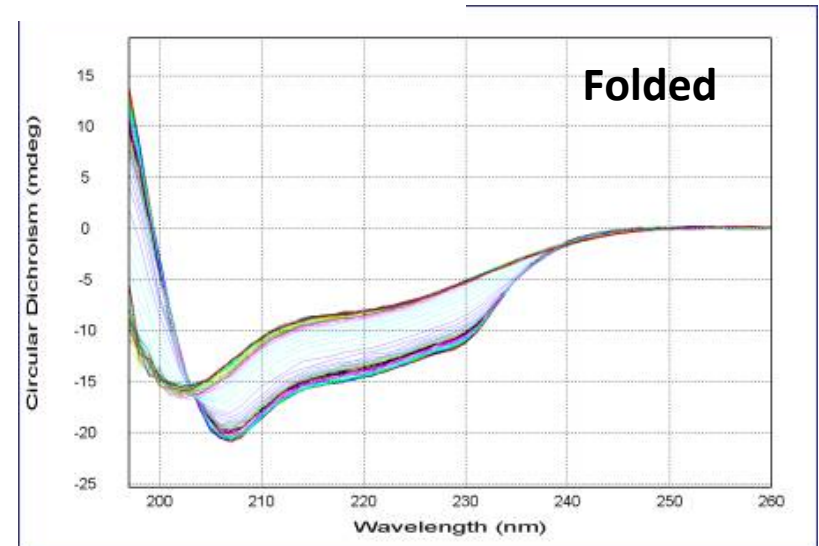
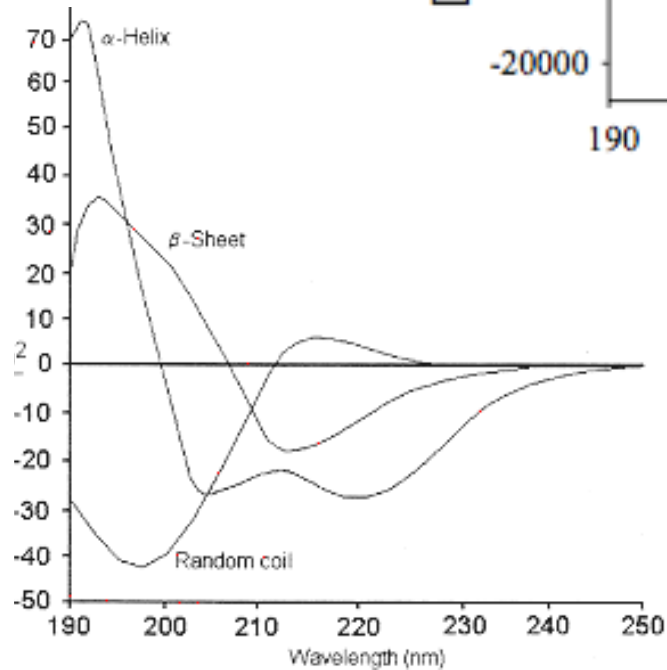
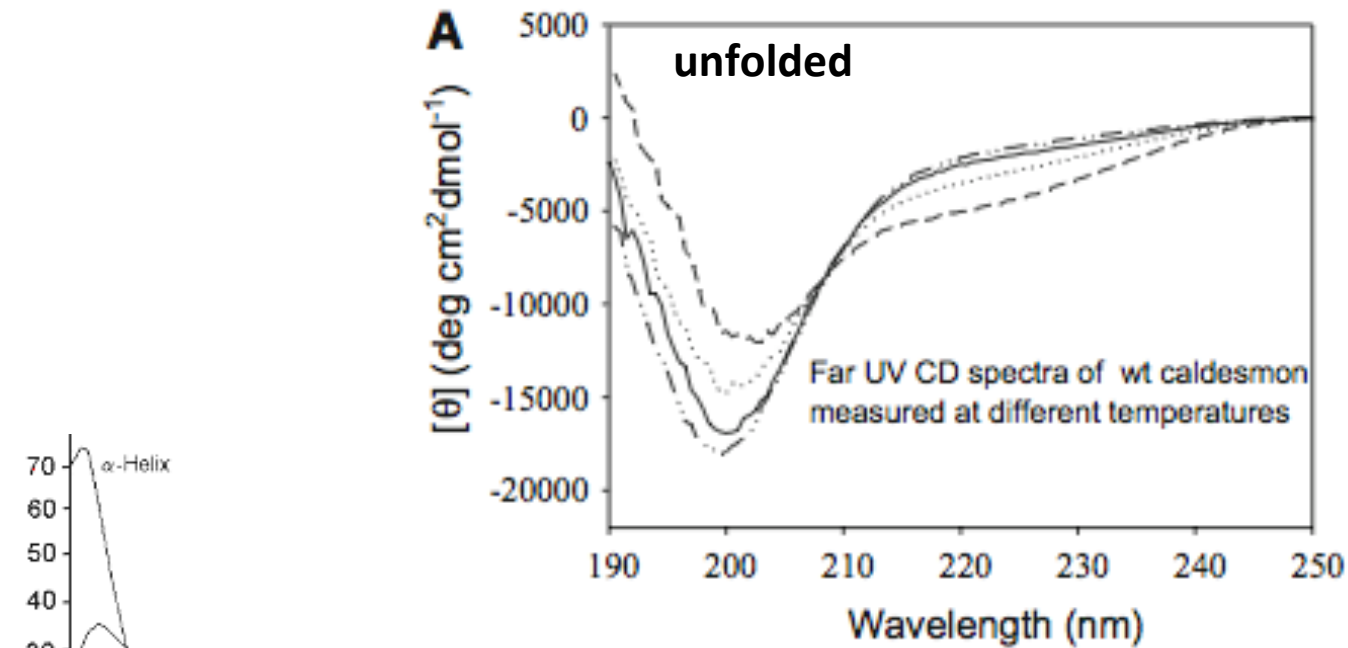


Folded

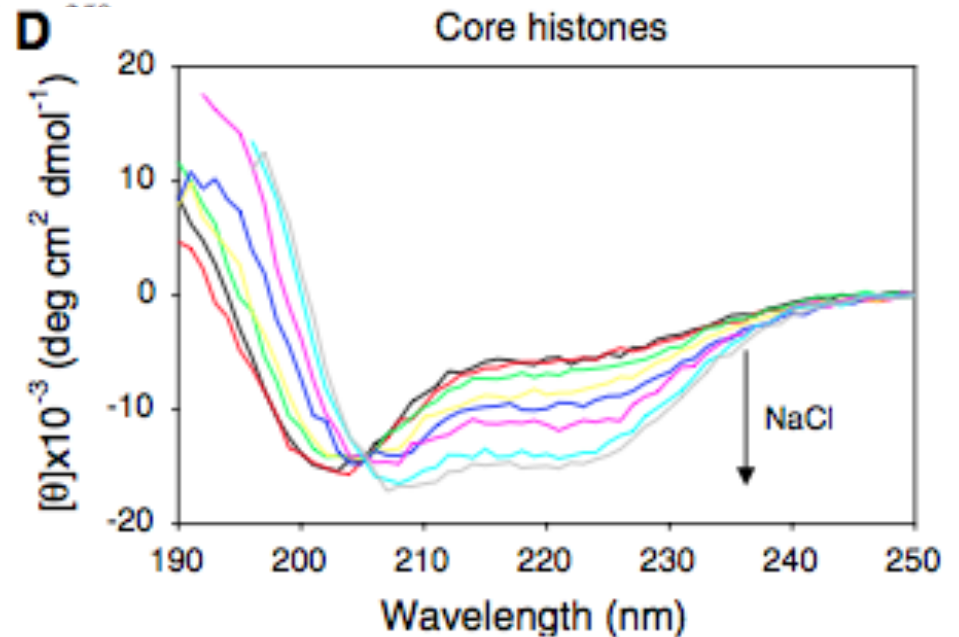
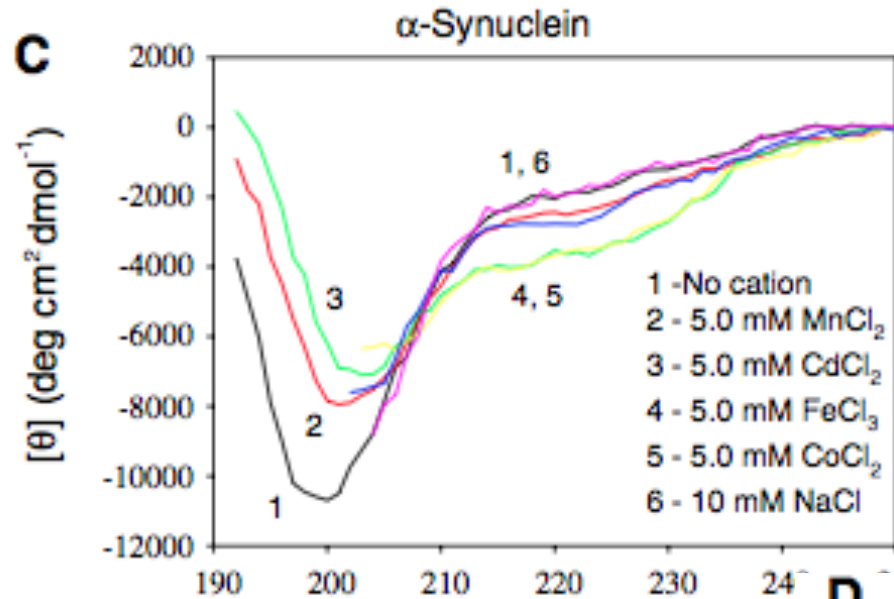


Unfolded

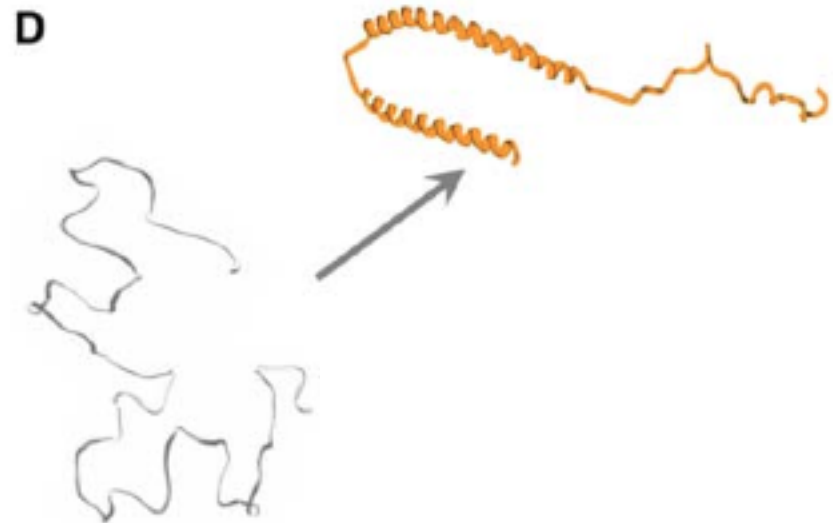
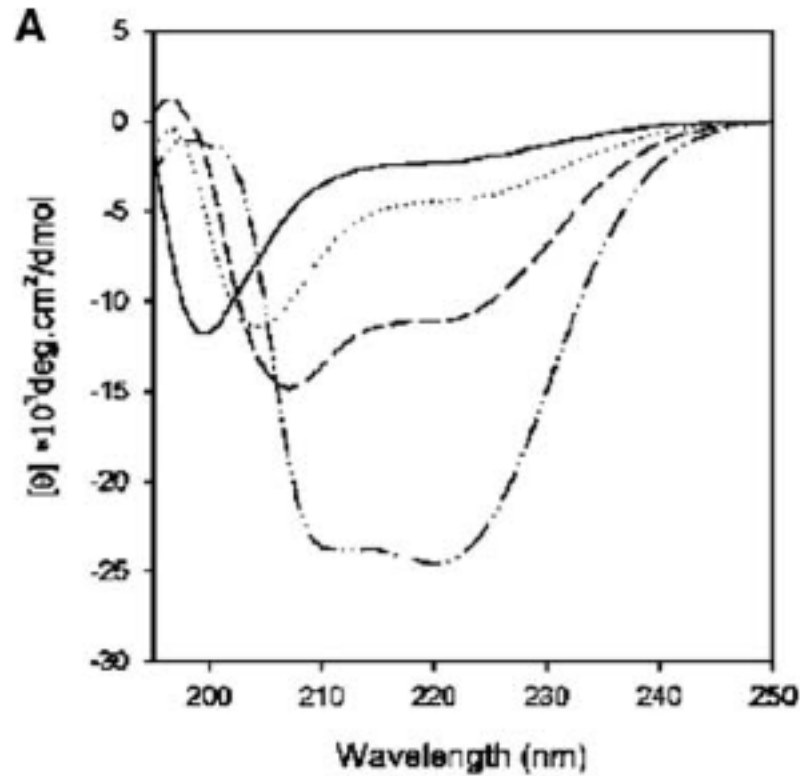
IUPs: “turned out” response to heat



IUPs: counter ions might promote folding



IUPs: membrane field can promote 2° structure



IUPs properties overview:

- IUP stereochemistry is sensitive to environmental factors
- Large hydrodynamic radius
- High content of polar AA
- High solvent accessibility
- Low content of hydrophobic, bulky AA
- Resistant to heat
- Dynamic

Biophysical tools for identification of IUPs

methods that are sensitive to molecular size, density or hydrodynamic drag:

size exclusion chromatography, analytical ultracentrifugation, small angle X-ray scattering (SAXS), or NMR measurements of the diffusion constant.

methods able to detect a lack of secondary structure:

far-UV (170-250 nm) circular dichroism, infrared spectroscopy, NMR spectroscopy

methods able to probe solvent accessibility:

Limited proteolysis proteases, hydrogen-deuterium exchange (MS and NMR)

The primary method

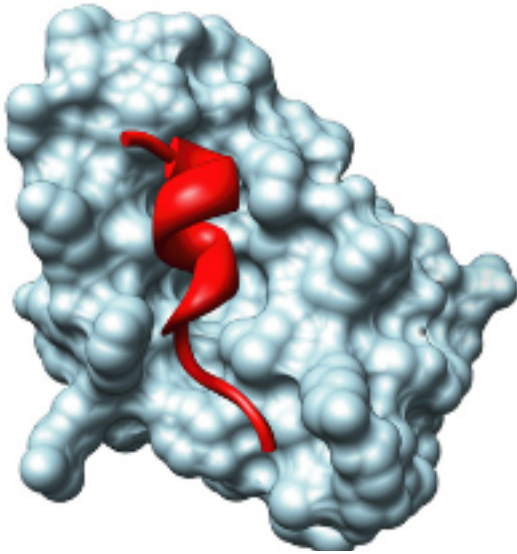
to obtain information on disordered regions of a protein is **NMR spectroscopy**.

(Quantitative description: residual structure & dynamics)

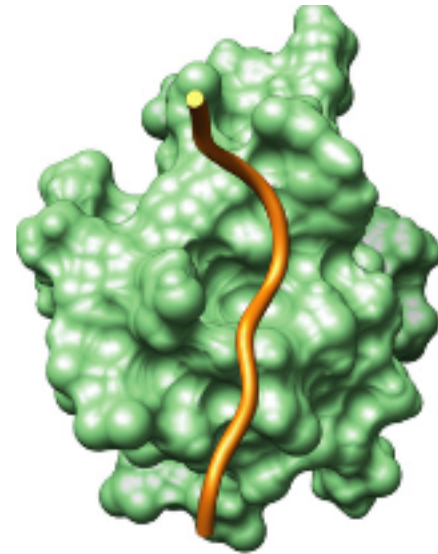
...others (SDS PAGE, AFM, DLS, ...)

Coupled folding and binding

Folding upon binding

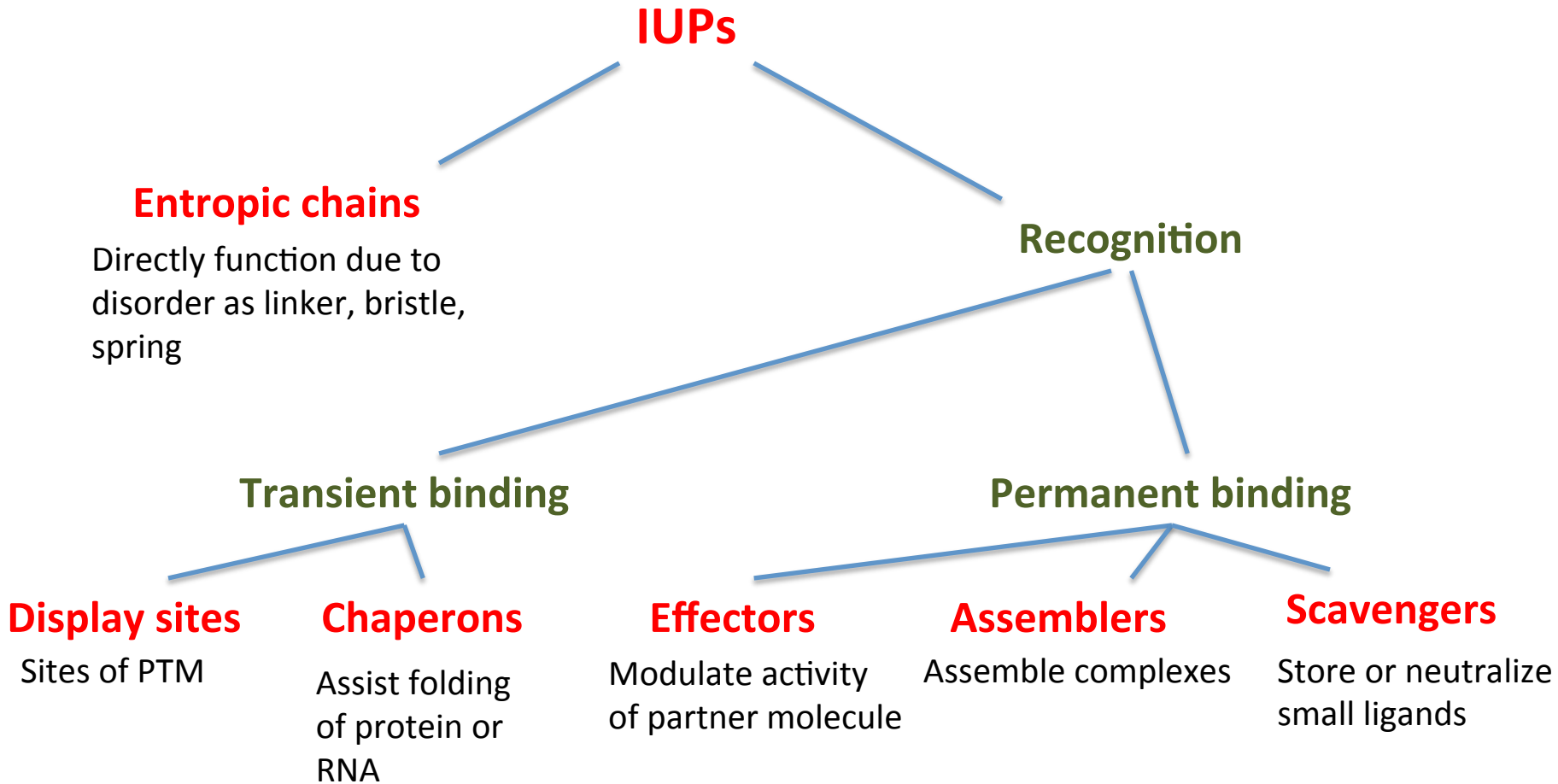


Restricting mobility upon binding



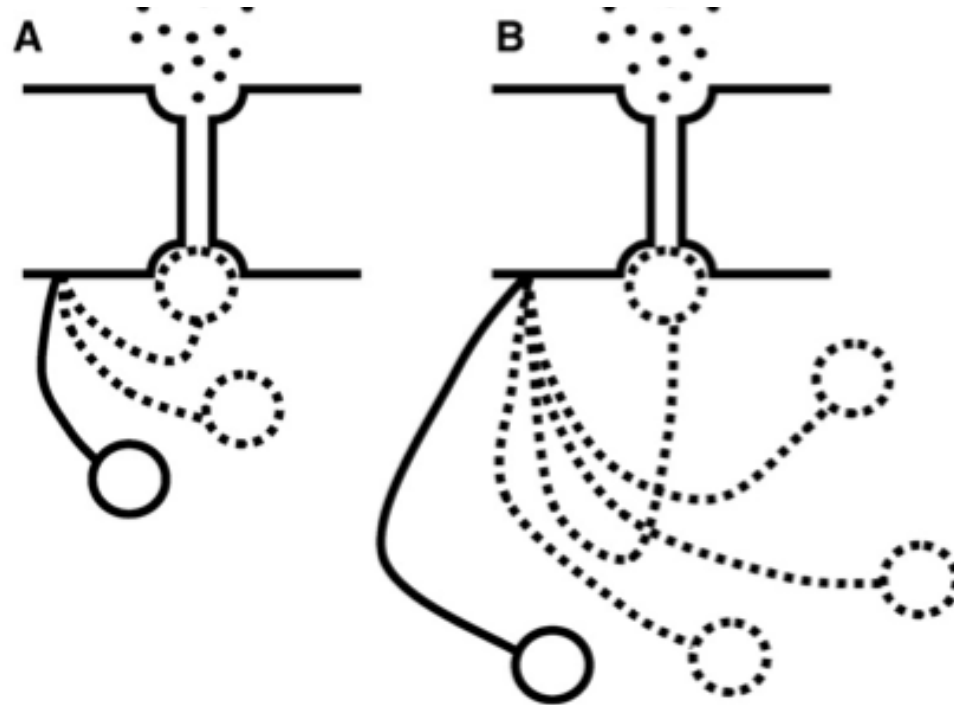
The ability of disordered proteins to bind, and thus to exert a function, shows that stability is not a required condition.

Functions of intrinsically unfolded proteins:



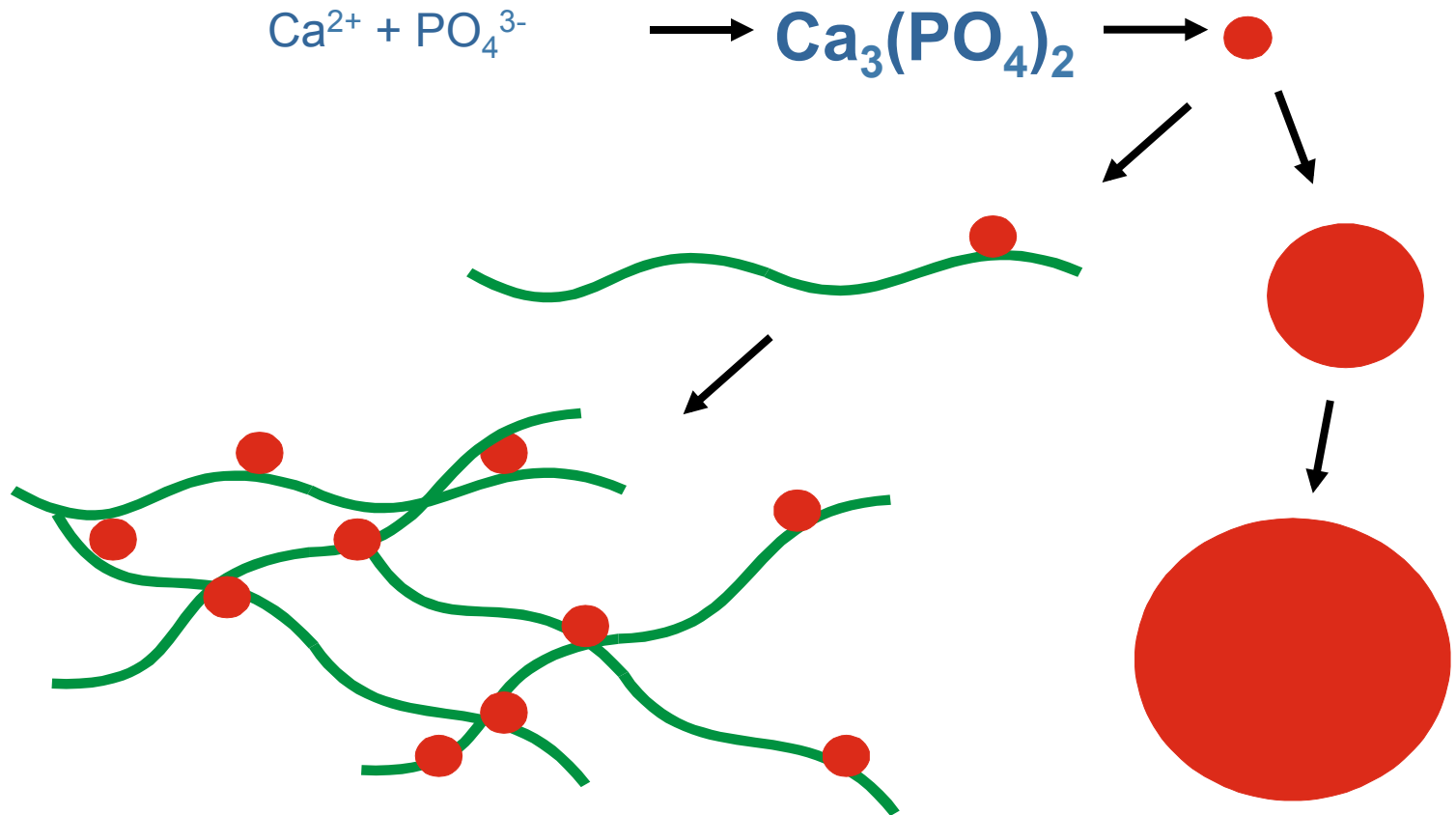
IUPs: Entropic chains

e.g. entropic clock



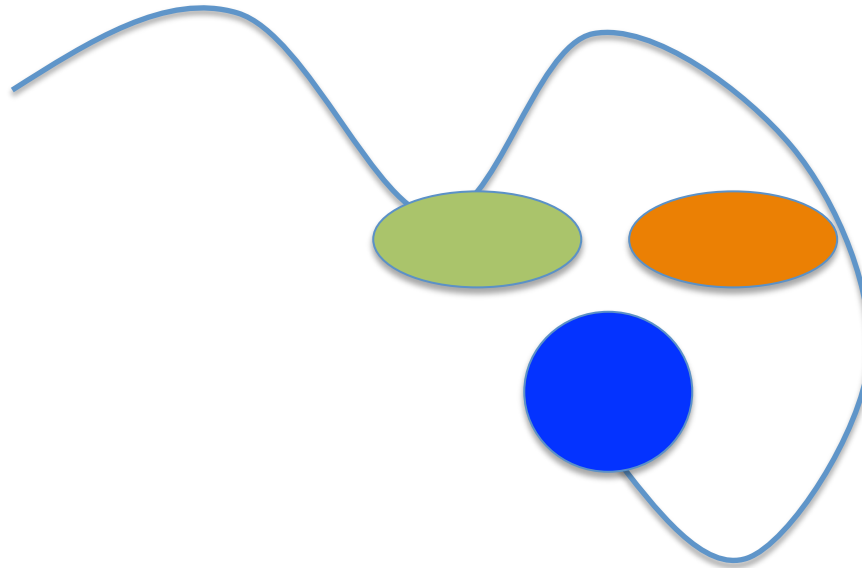
IUPs: Scavengers

Casein – preventing Ca^{2+} precipitation



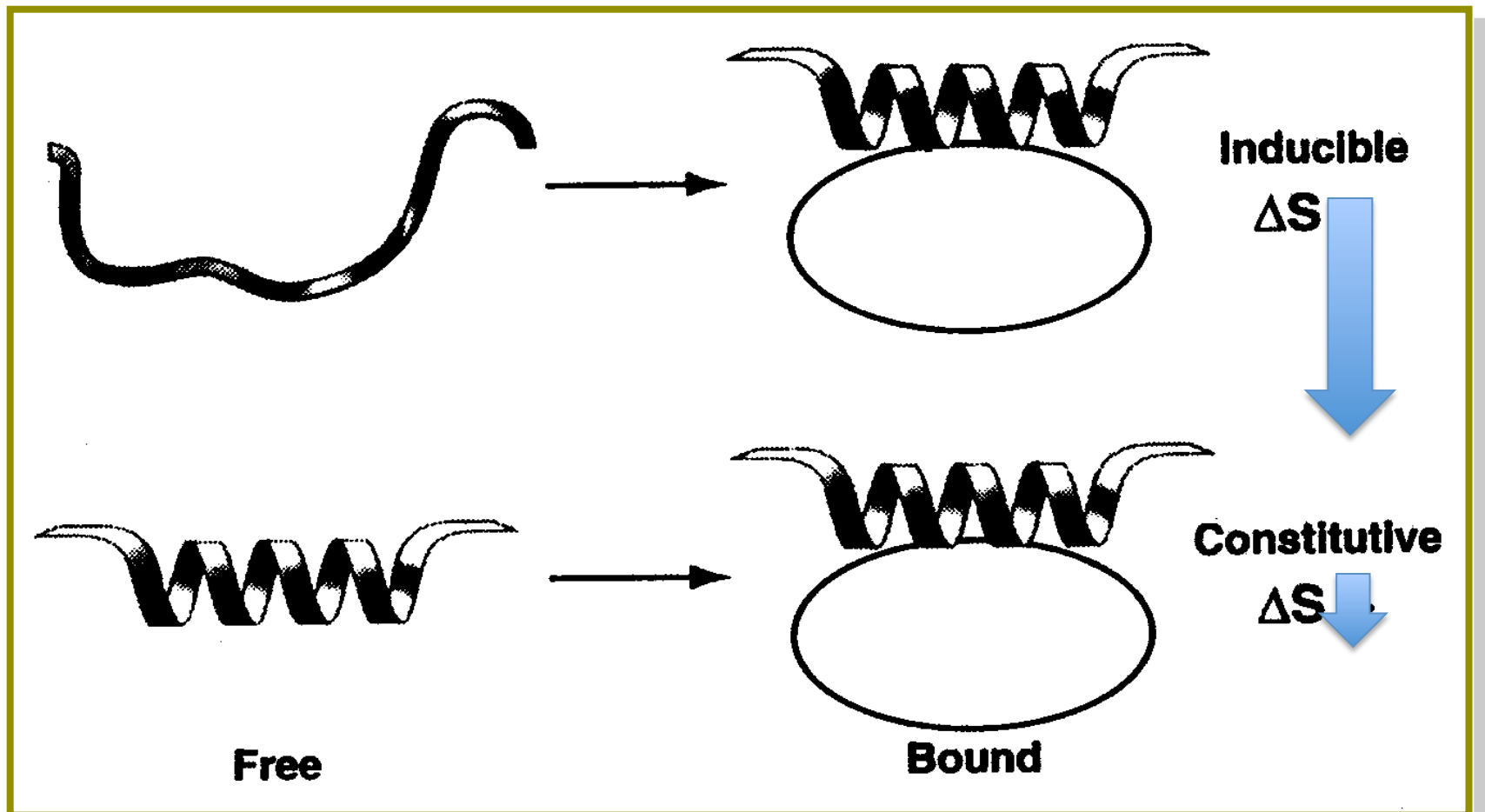
IUPs: Assemblers/Scaffolds

Assemble complexes – IUP brings binding partners together



Functional advantages of IUPs

Specificity without strong binding
(binding promiscuity, increased speed of interaction)

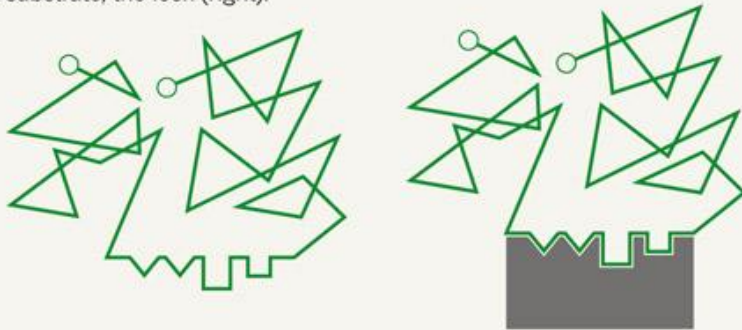


Possible mechanisms of IDP function in signal transduction

Ordered proteins

LOCK AND KEY

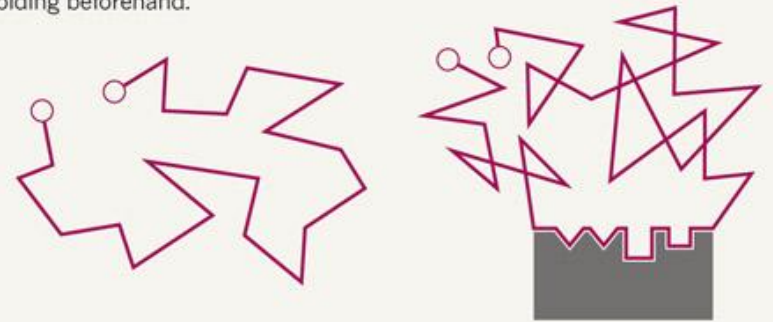
In the conventional view, an enzyme folds up immediately into a unique and stable 3D shape, the key (left). Its shape perfectly matches and allows it to bind its substrate, the lock (right).



IDP

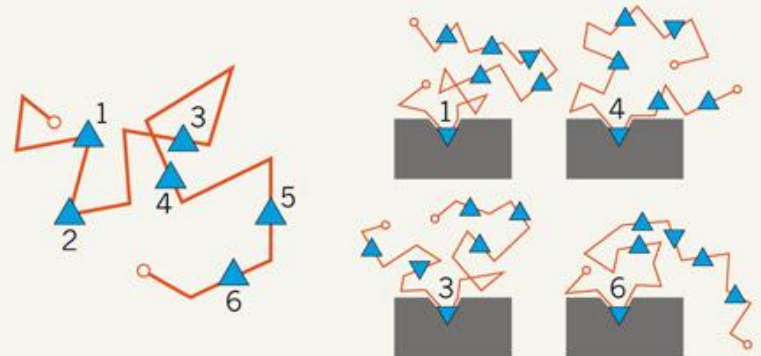
FOLD AS YOU BIND

A disordered part of the gene-regulatory protein CREB (left) uses the lock to mould itself into the shape of the key when the two meet (right), rather than folding beforehand.



SHAPE SHIFTING

The signalling protein Sic1 remains disordered in its bound state, and each of six phosphate groups occupies the binding site in turn. The protein is a mix of different conformations shifting around in constant dynamic equilibrium.



Shape shifting

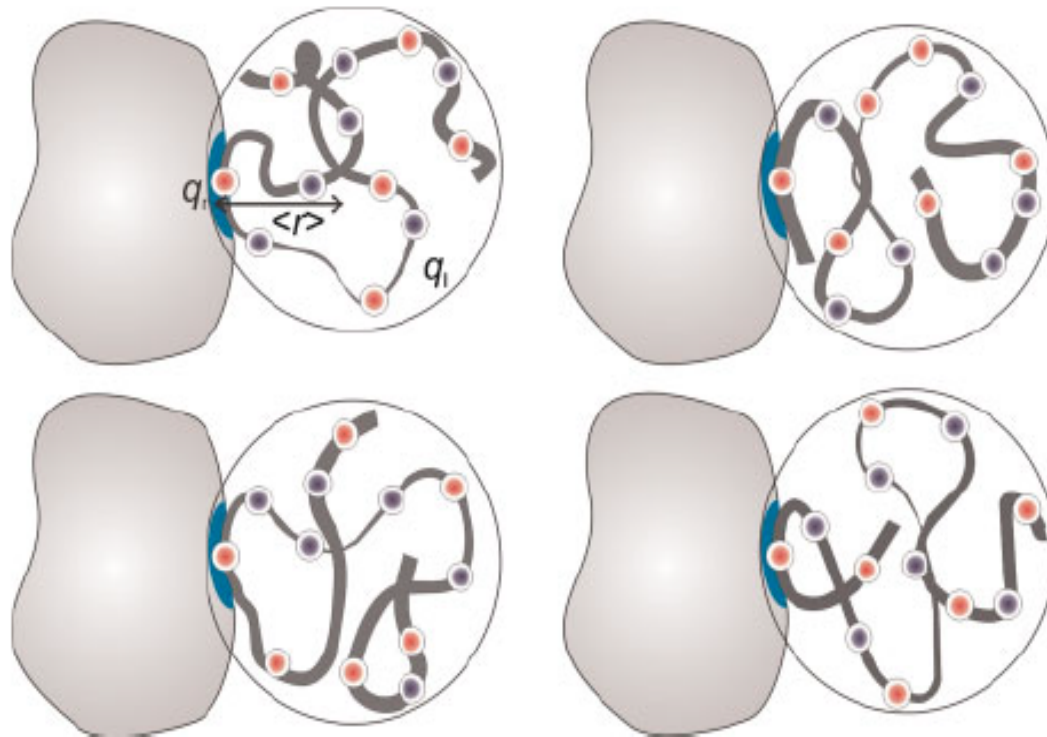


Figure 1. "Polelectrostatic" model of interaction of intrinsically disordered proteins. Schematic of an intrinsically disordered protein (ribbon) interacting with a folded receptor (gray shape) through several distinct binding motifs and an ensemble of conformations (indicated by four representations of the interaction). The intrinsically disordered protein possesses positive and negative charges (depicted as blue and red circles, respectively) giving rise to a net charge q_i , while the binding site in the receptor (light blue) has a charge q_r . The effective distance $\langle r \rangle$ is between the binding site and the centre of mass of the intrinsically disordered protein.

Database of Protein Disorder & IDP predictors

<http://www.disprot.org/>