

Macromolecules

Proteins

**Csaba Magyar, Institute of Enzymology,
Research Centre for Natural Sciences
2019 November**

Structure of proteins

Bioinformatical methods, protein modelling

Transmembrane proteins

Disordered proteins

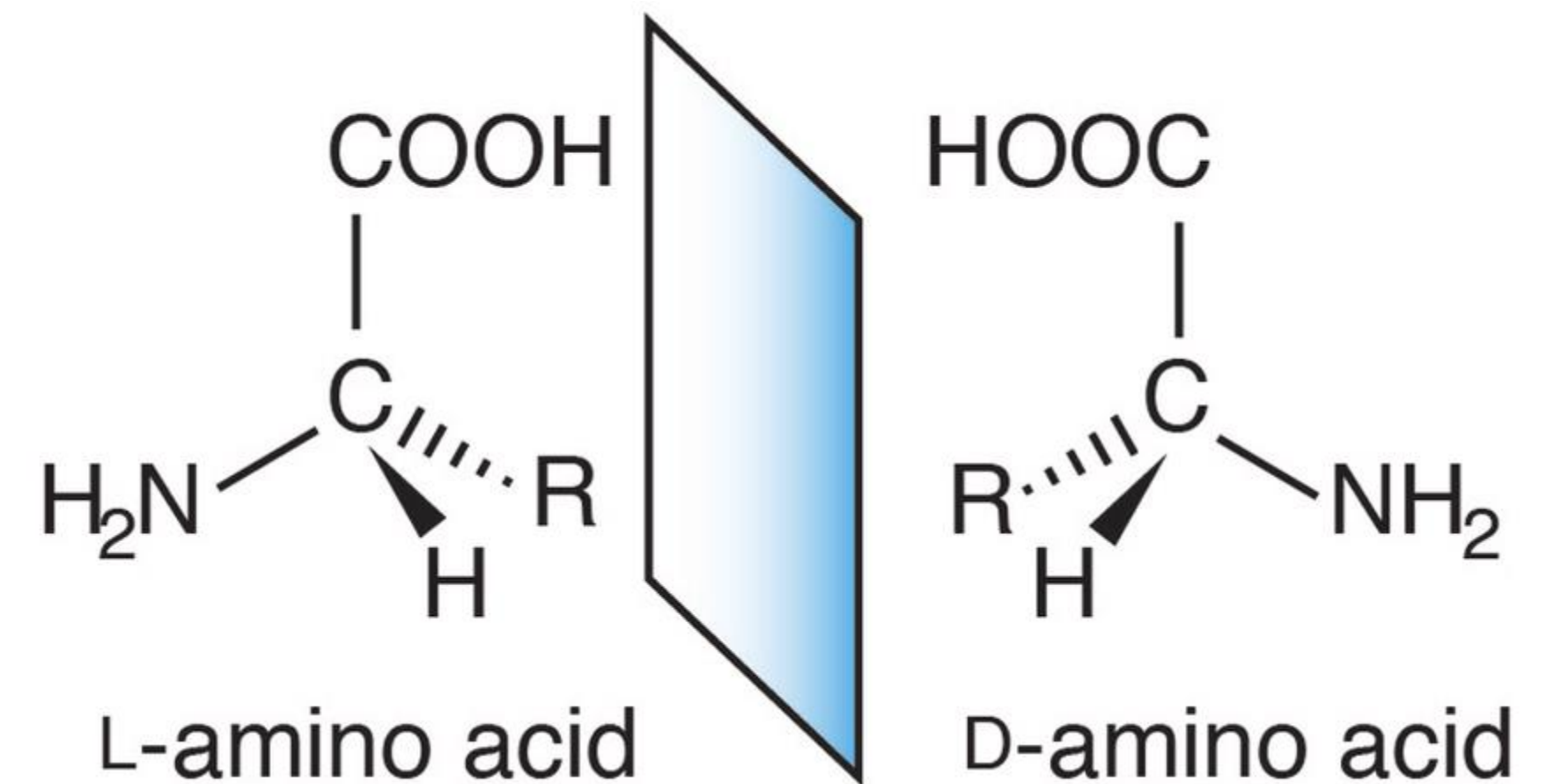
Proteins are macromolecules built from amino acids

Almost exclusively L-enantiomers of amino acids are present in proteins. For sugars the opposite is true (D-glucose).

Meaning of the word *proteos*: primary or „first place”
proteins represent ~55% of the dry cell weight

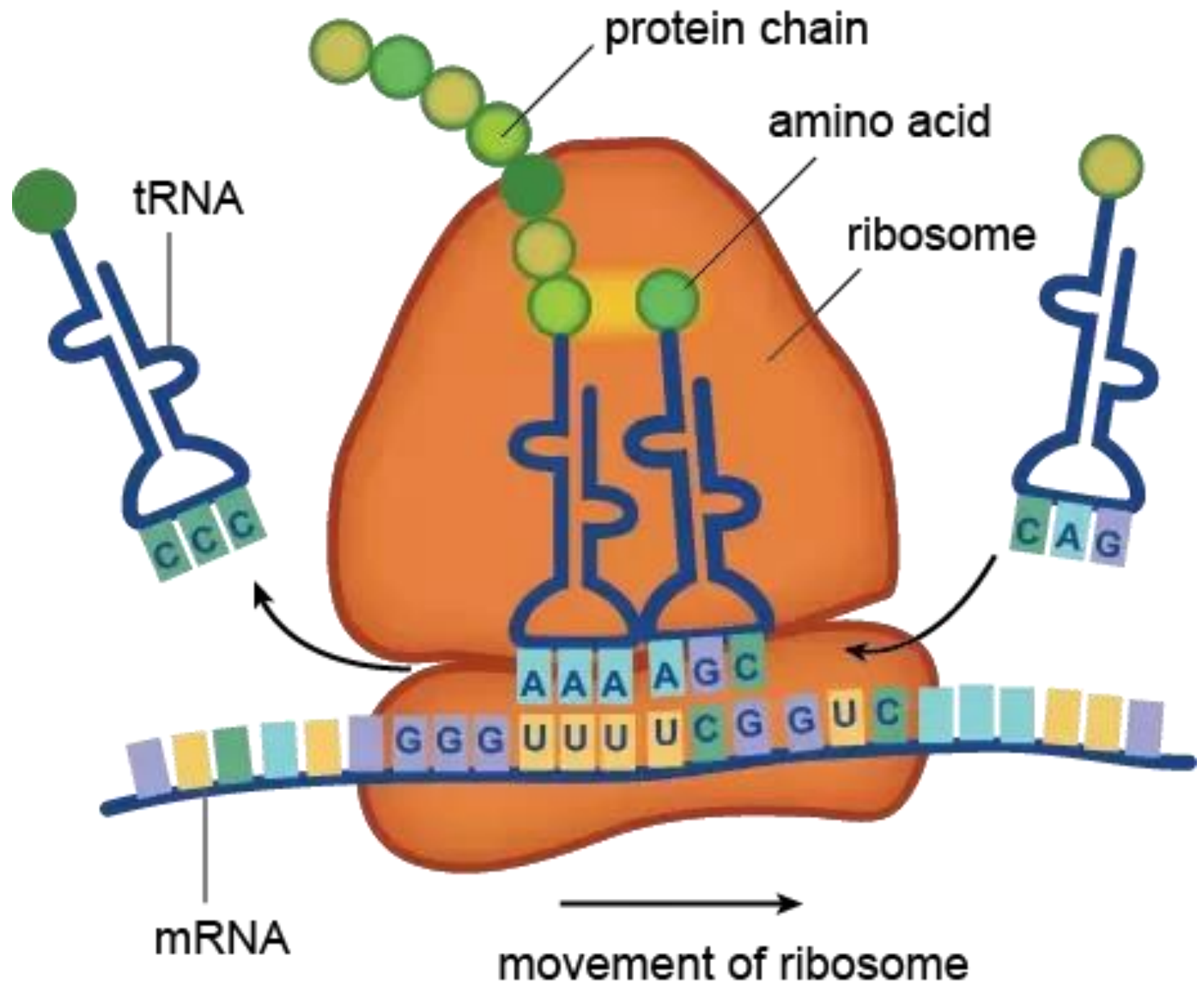
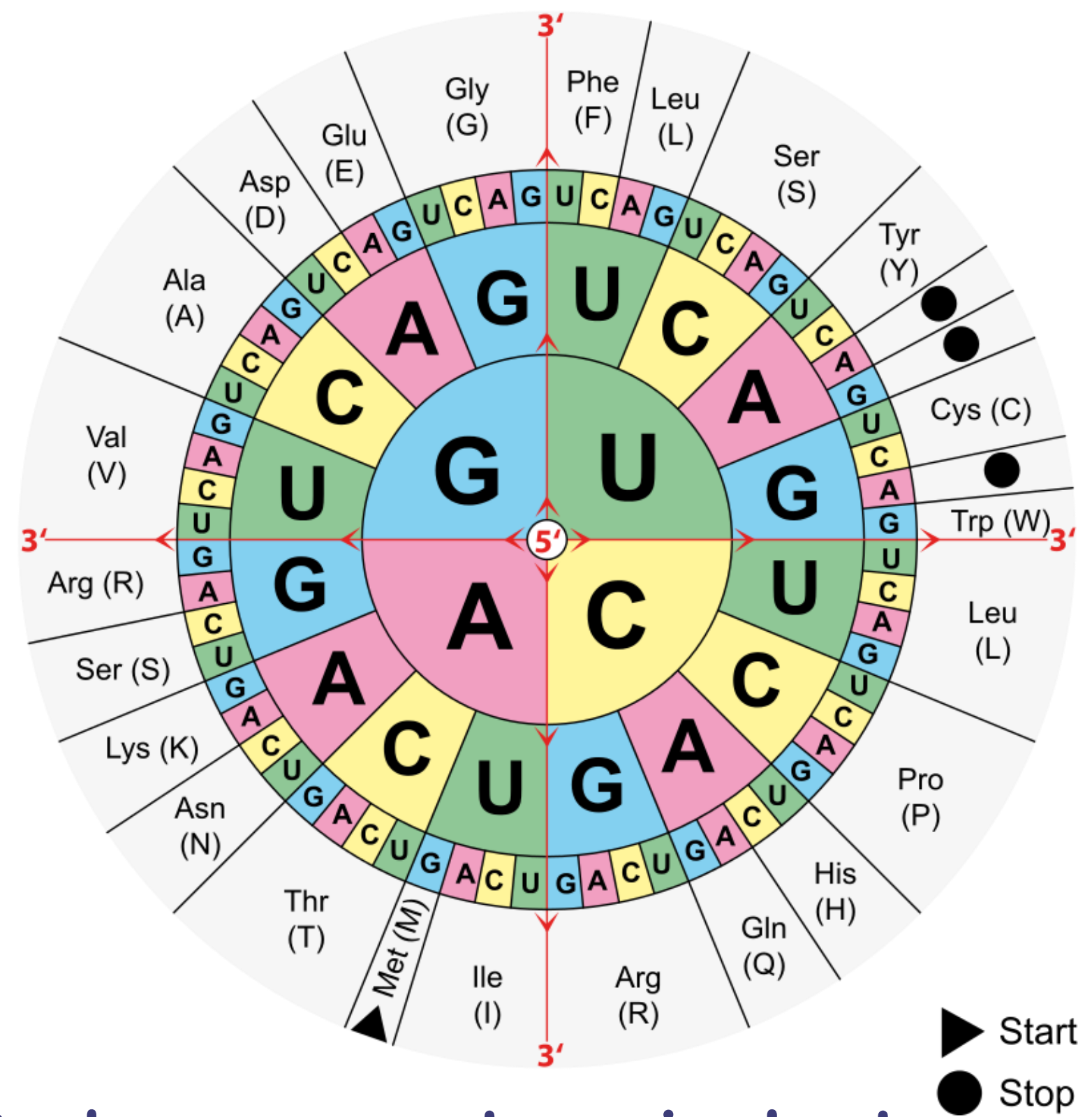
Multiples roles:

- structural proteins
- enzymes
- antibodies
- communication
- transport
- receptors
- muscles



Protein translation

Amino acids are coded by the nucleic acids triplets



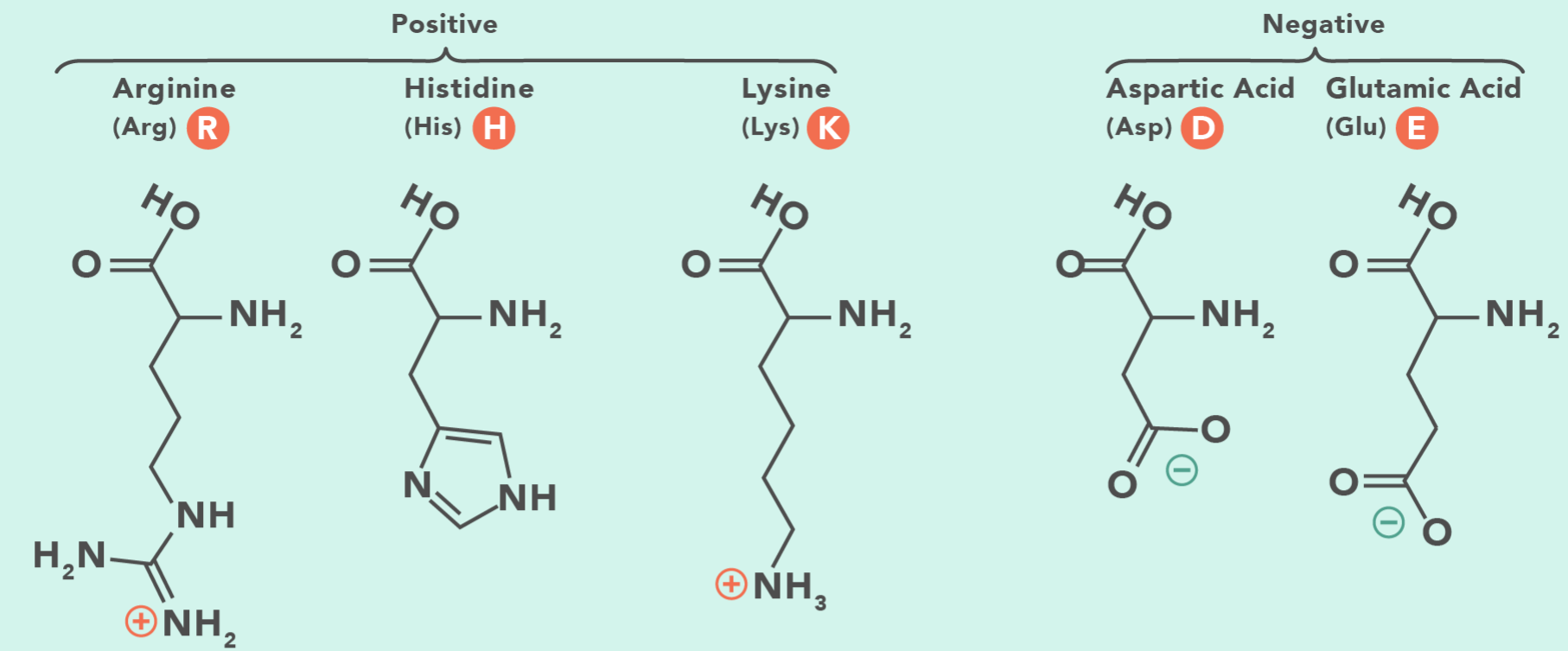
Codons are not equivalent

Translation speed depends on the codon, effect on protein folding

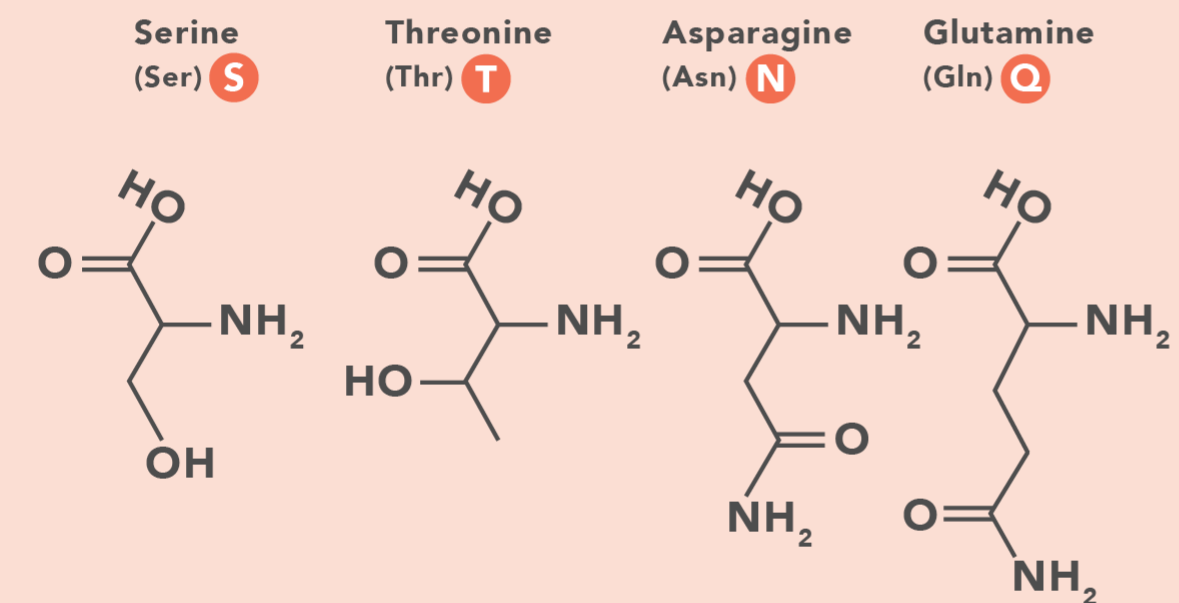
The 20 standard amino acids

3-letter code
1-letter code

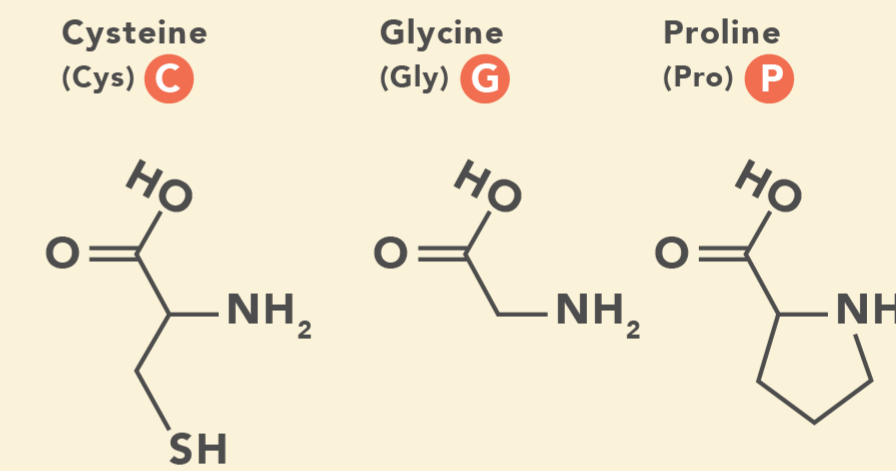
A. Amino Acids with Electrically Charged Side Chains



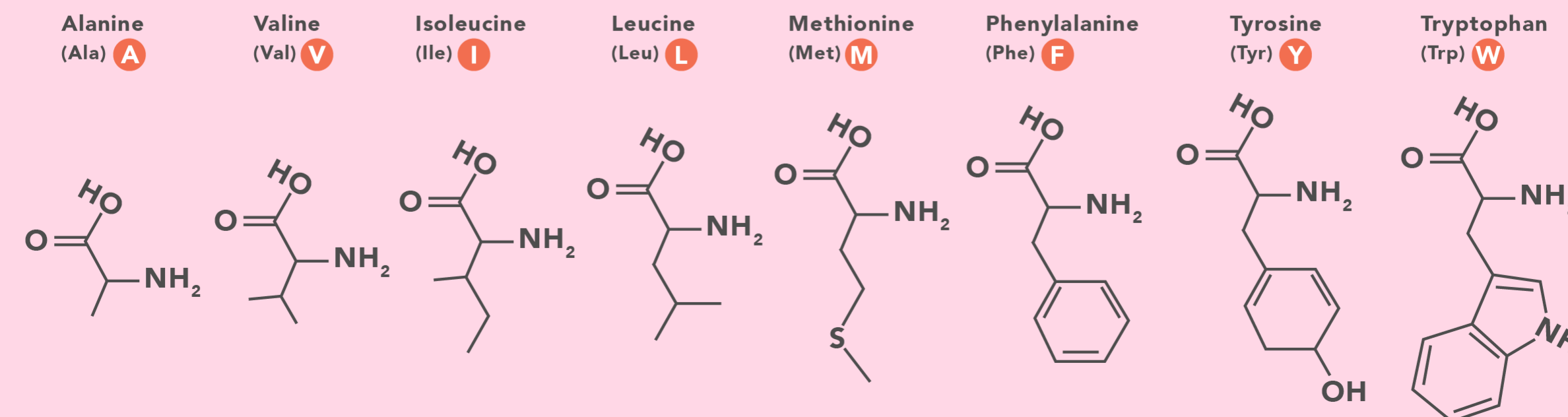
B. Amino Acids with Polar Uncharged Side Chains



C. Special Cases

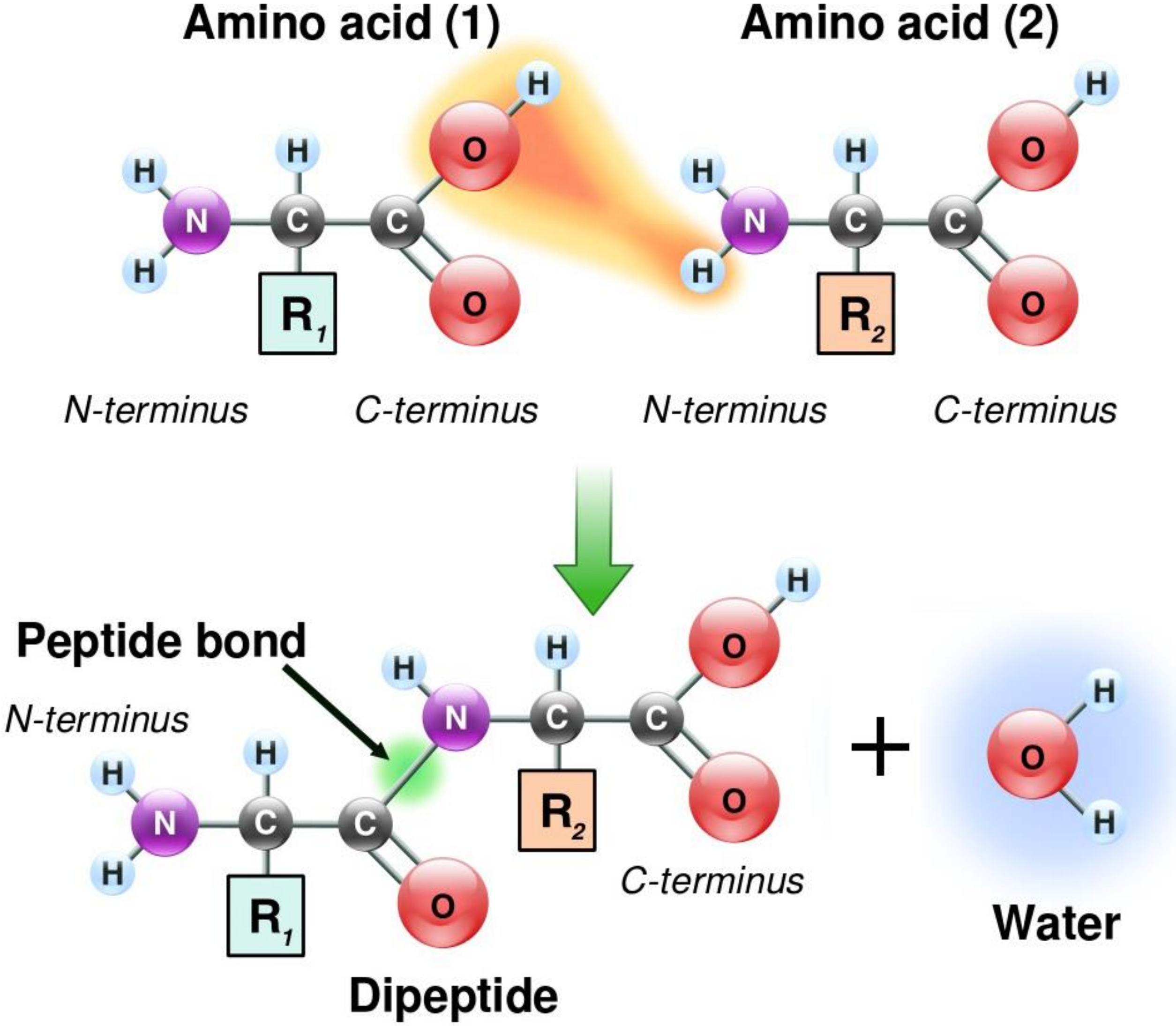


D. Amino Acids with Hydrophobic Side Chains



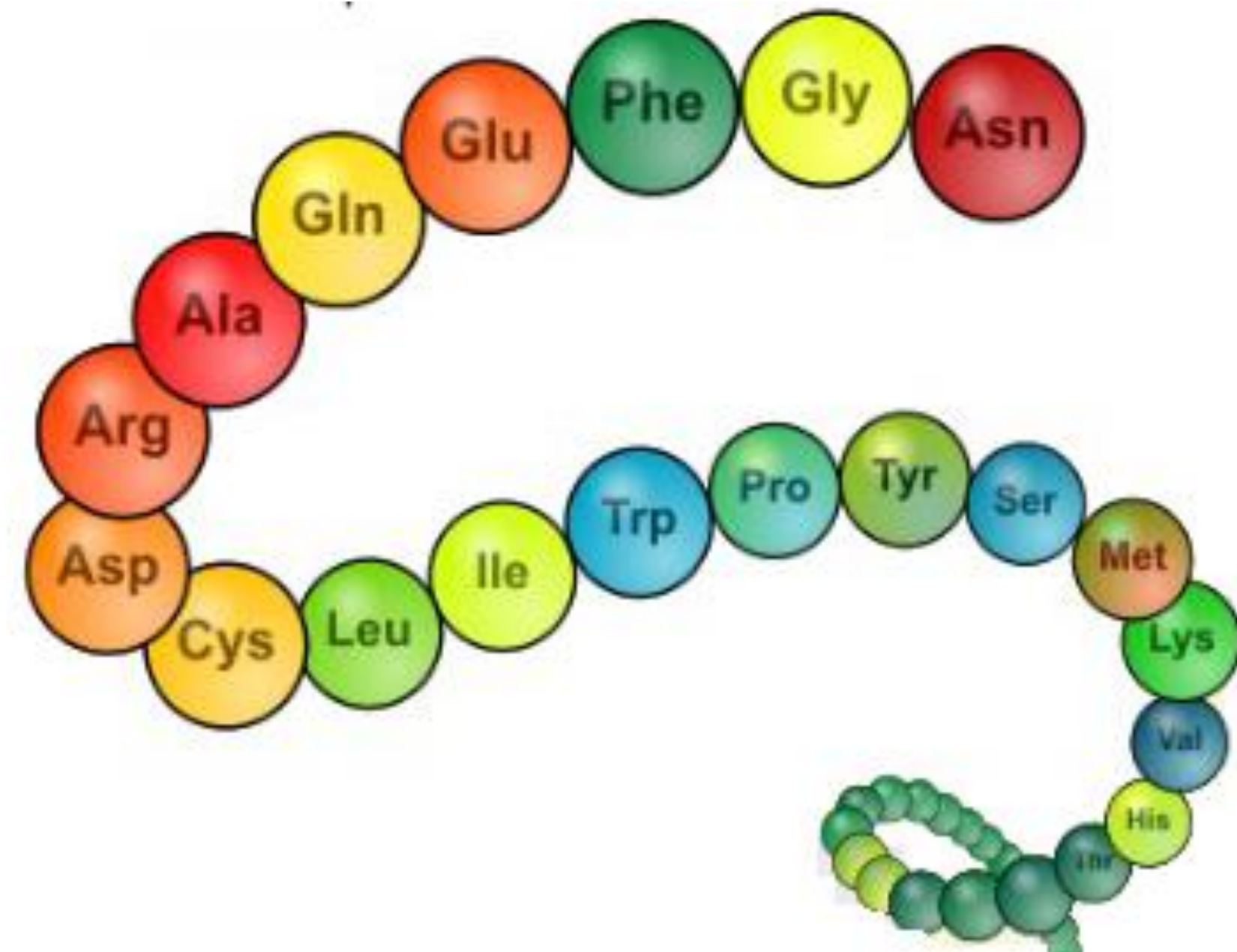
From amino acids to proteins

covalent peptide bond formation

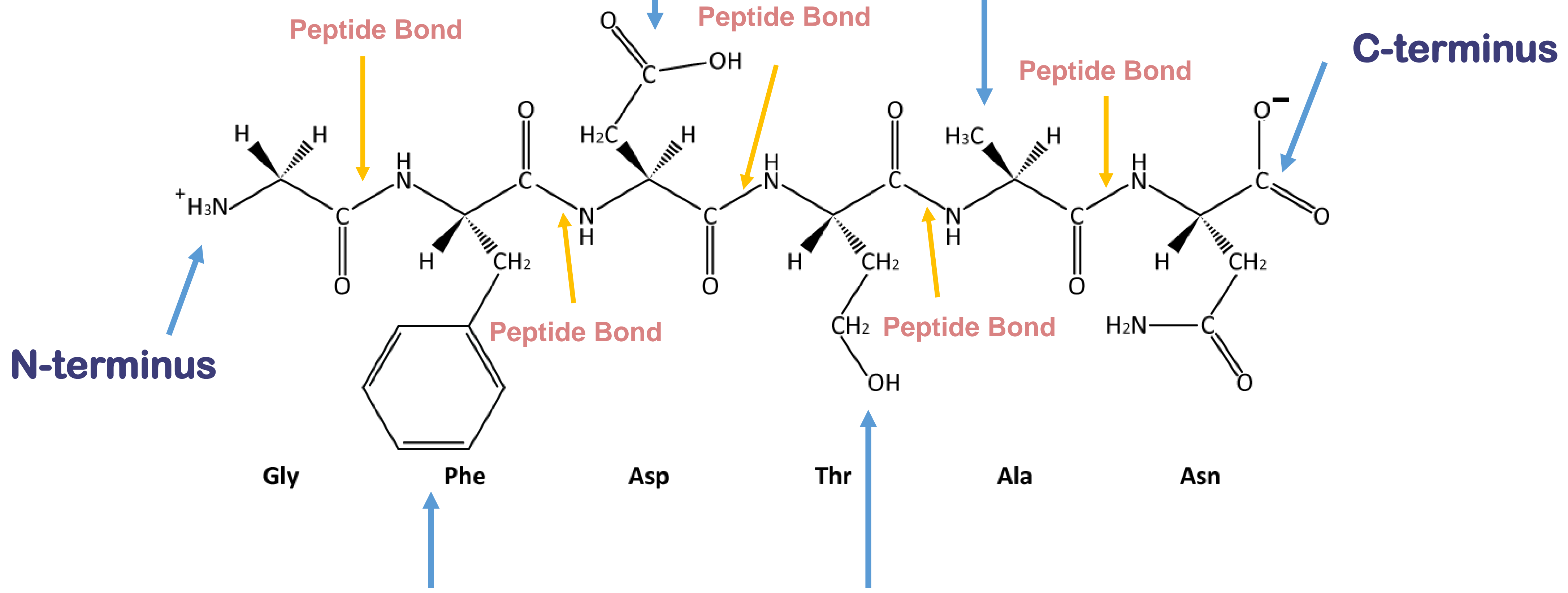


Primary structure of proteins

sequence of amino acids connected by peptide bonds
translated from mRNA using Genetic Code
synthesis begins at N-terminal and terminates at C-terminal
determines all properties of a protein

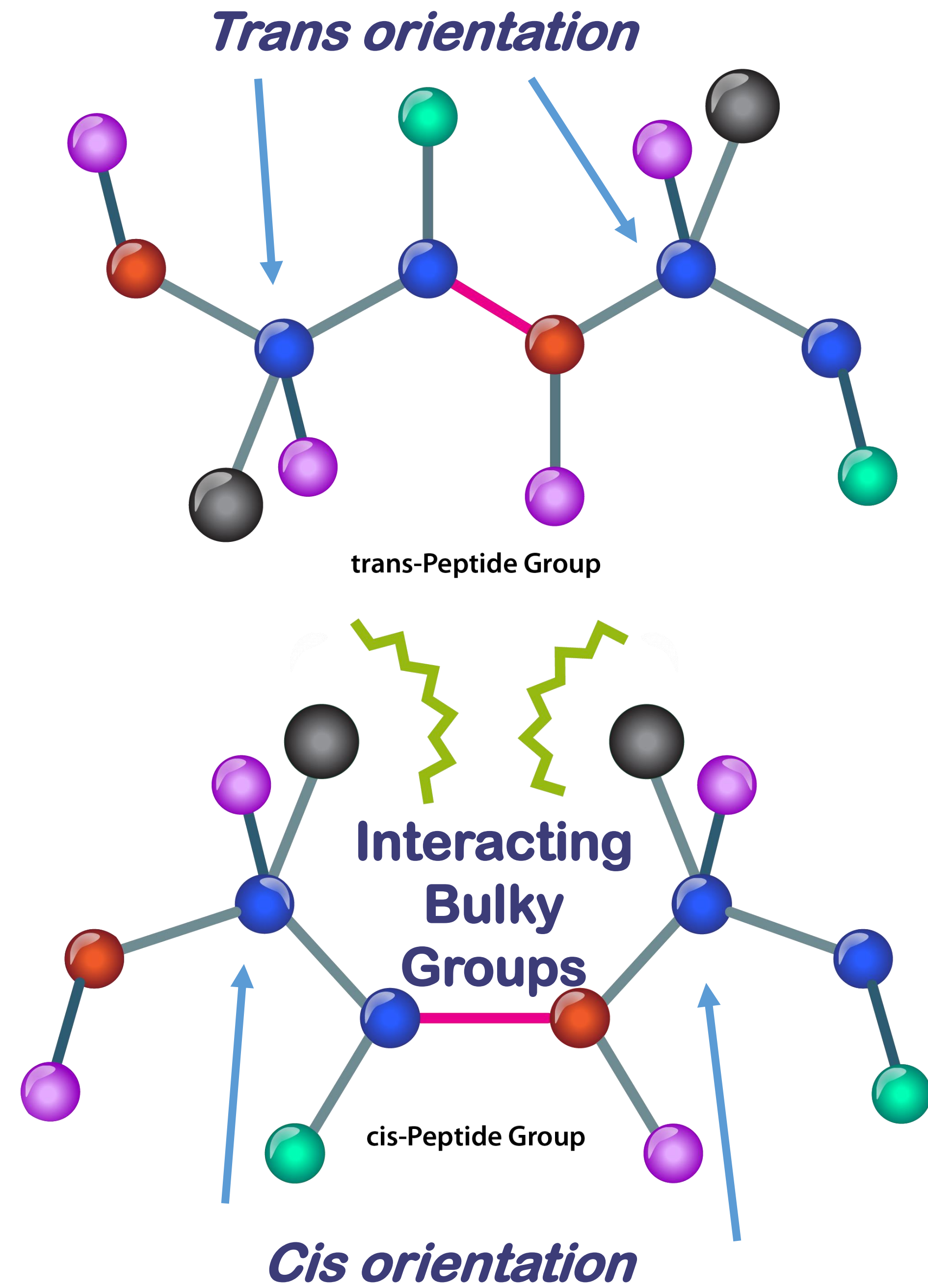


Alternating Orientations of side chains



Alternating Orientations of side chains

Orientation



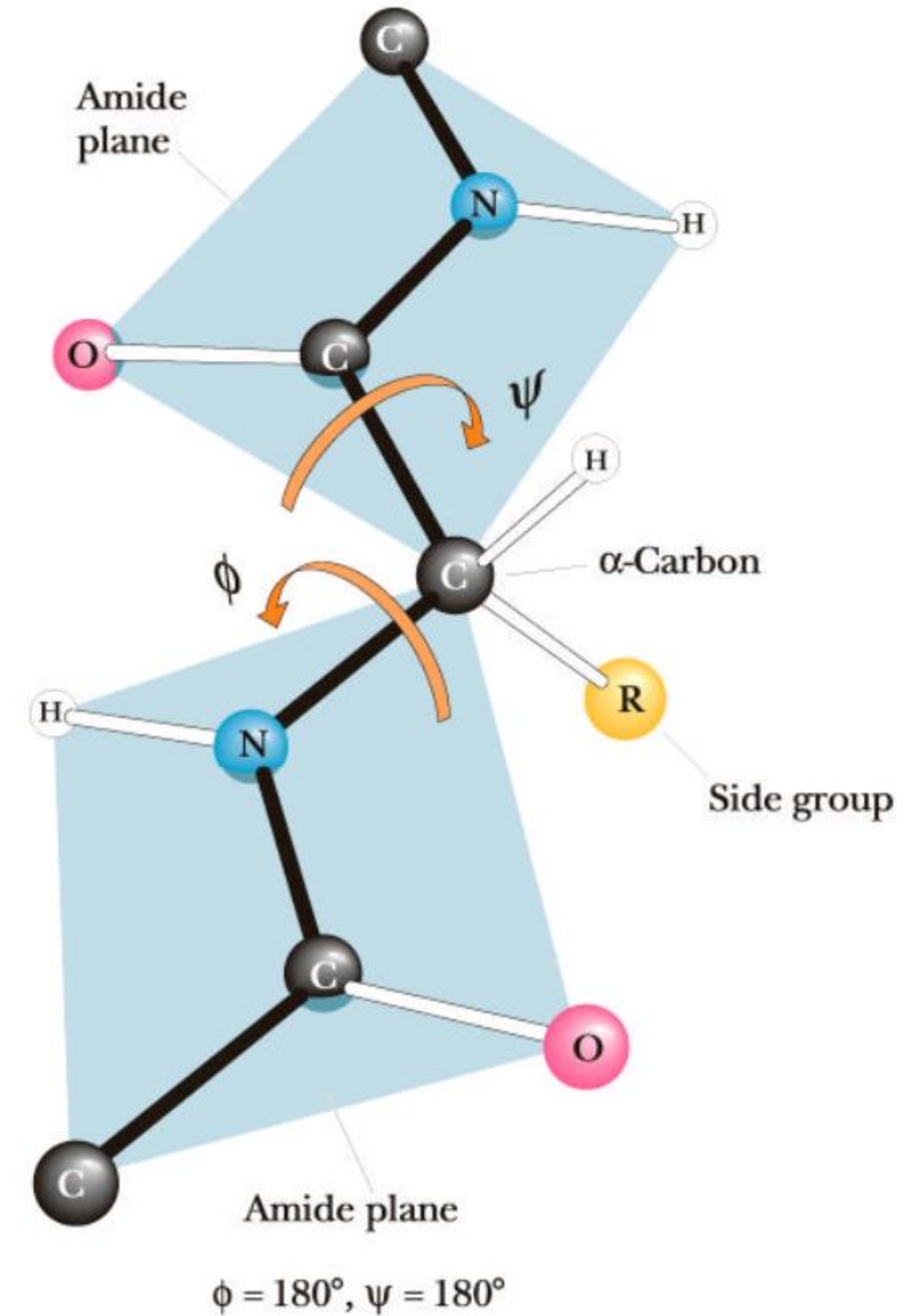
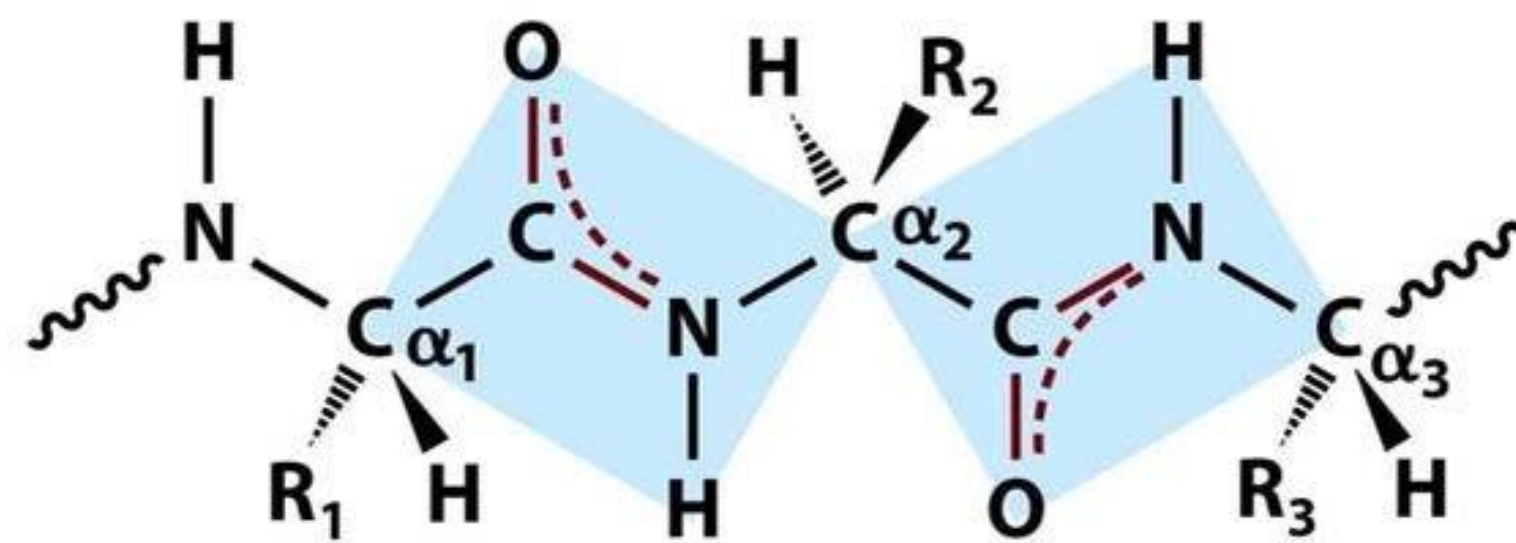
Torsional angles

partial double bond nature

restricted rotation

bond lengths are almost constant

two dihedral angles describe
the structure of the polypeptide backbone

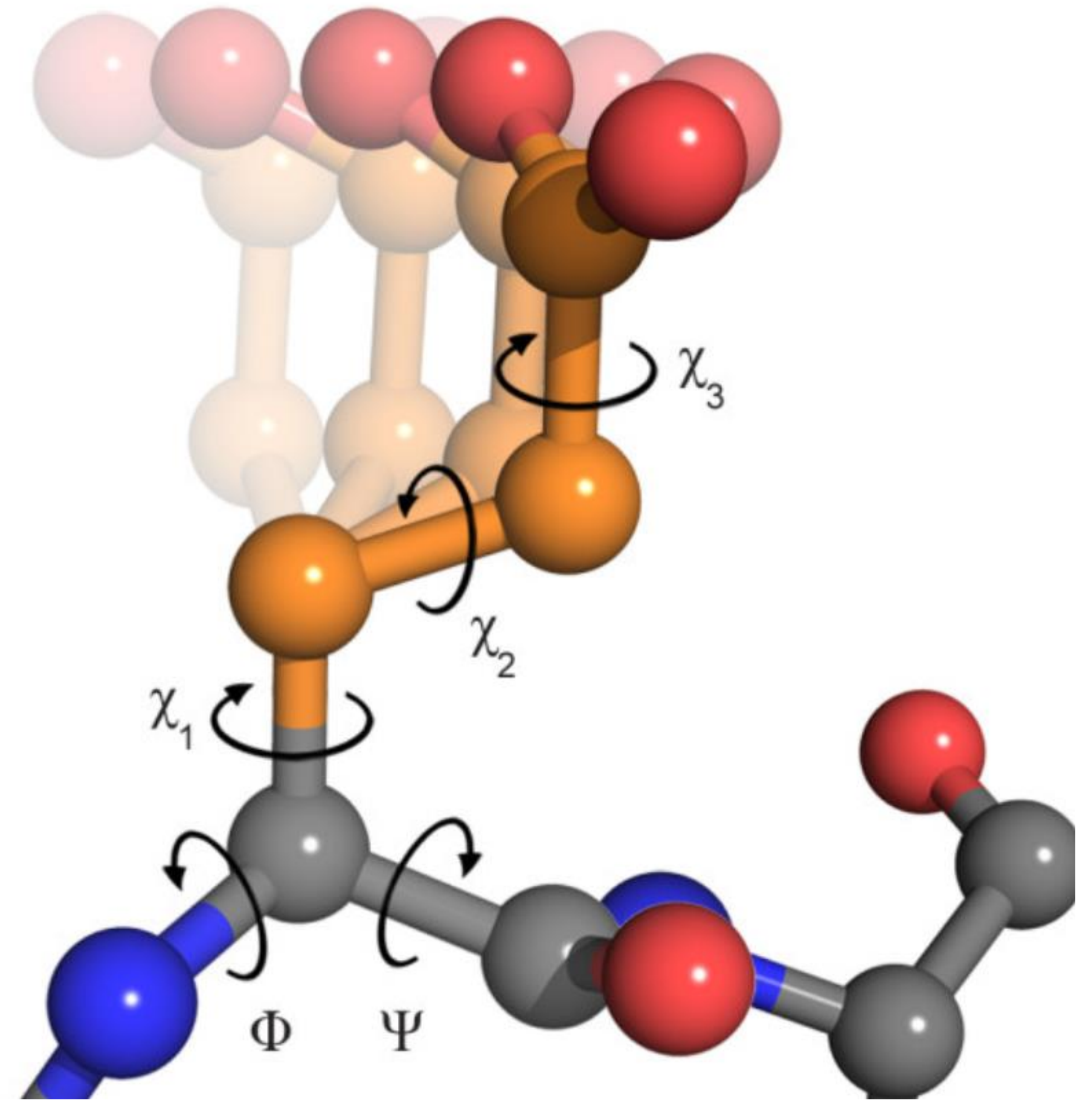


Torsional angles

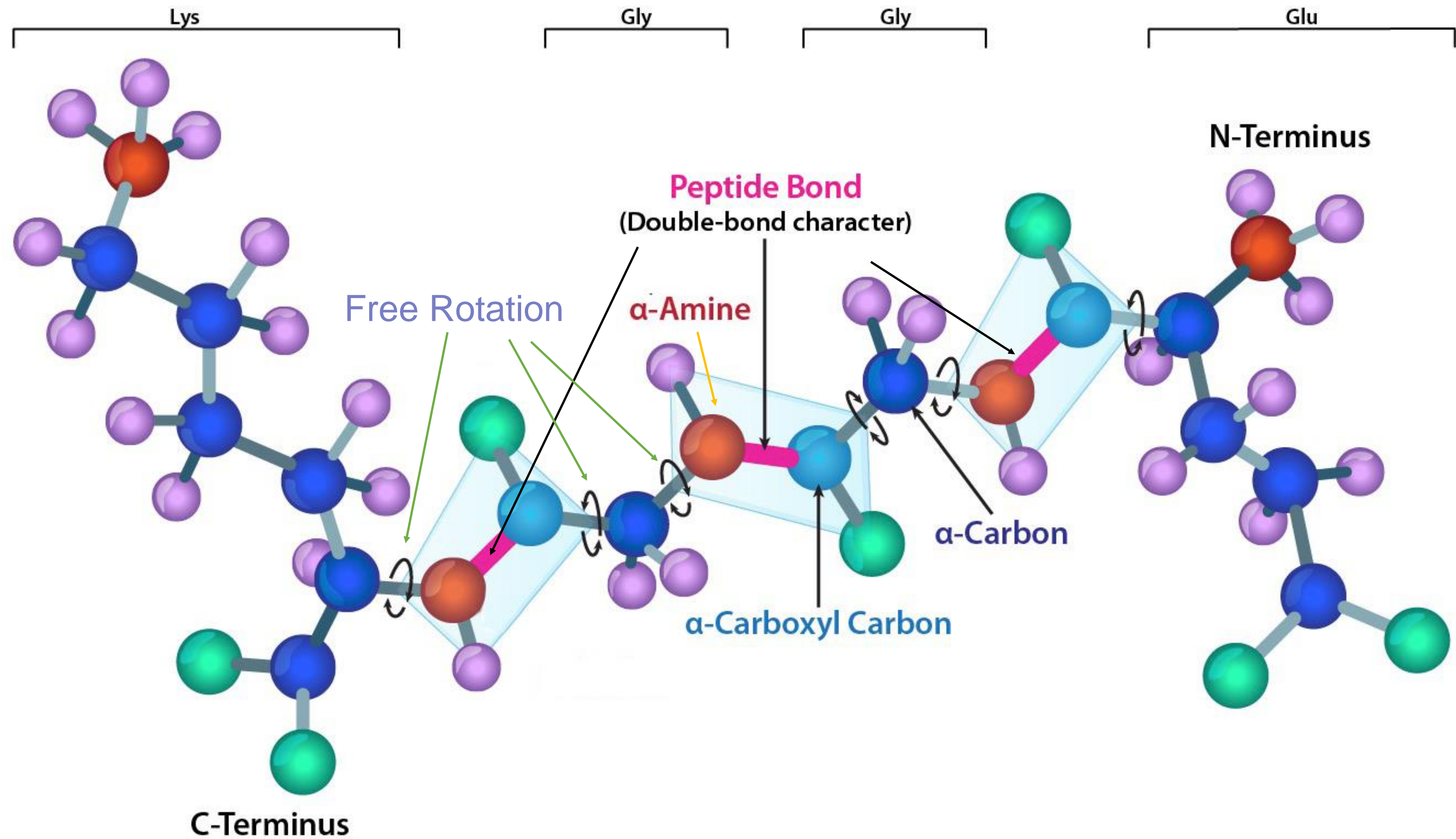
bond lengths are almost constant

two dihedral angles describe
the structure of the polypeptide backbone

the structure of the side-chain can be given by
additional dihedral angles



Structure of polypeptides

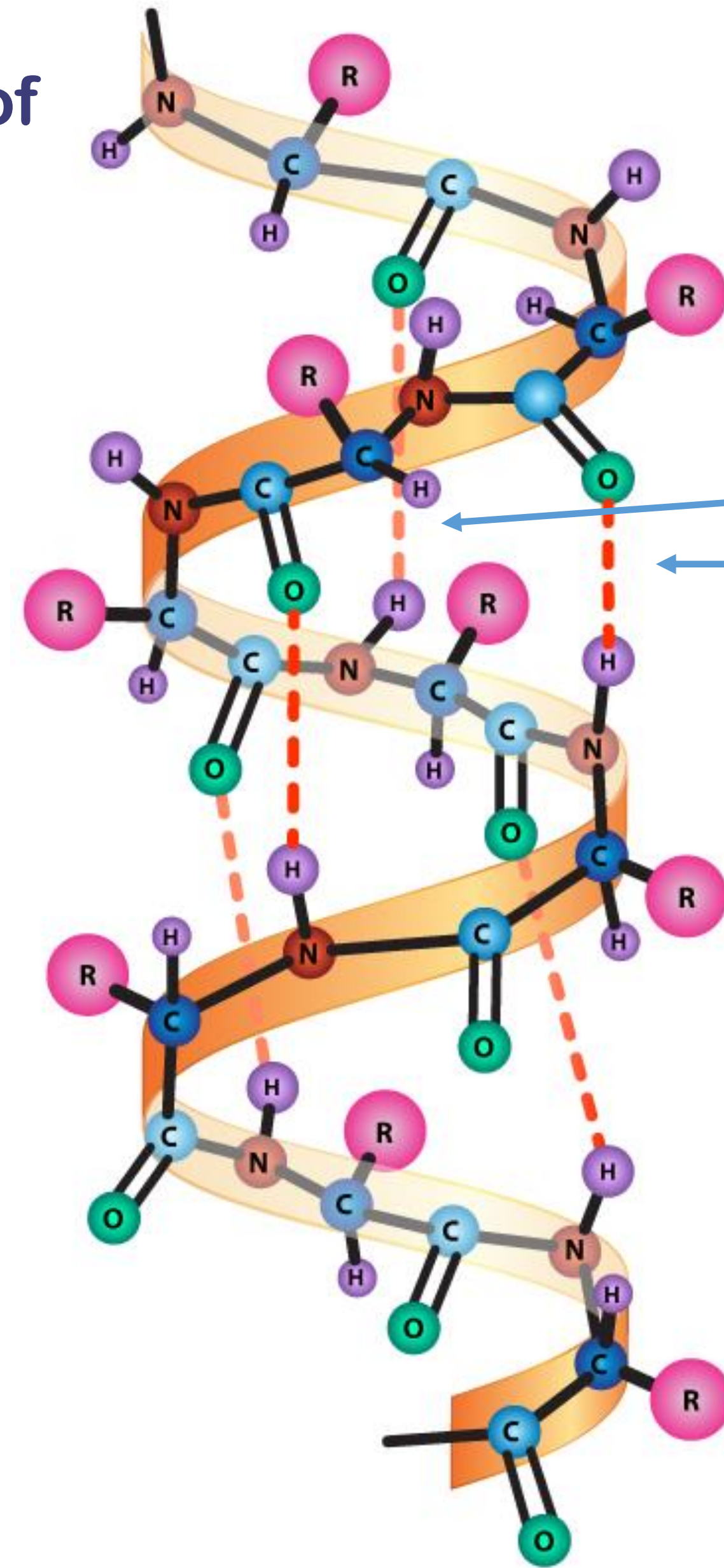


Secondary structure

Secondary structure is defined by the pattern of hydrogen bonds between the amino hydrogen and carboxyl oxygen atoms in the peptide backbone.

Secondary structural elements

alpha helix



Hydrogen bonds stabilize structure

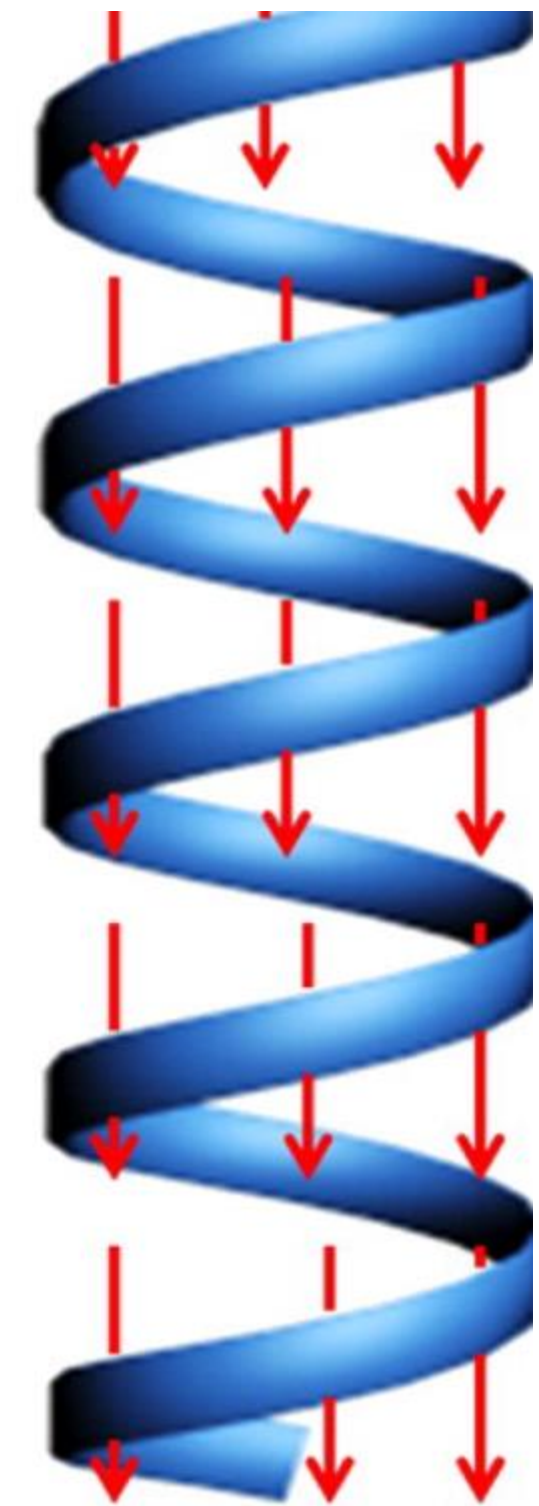
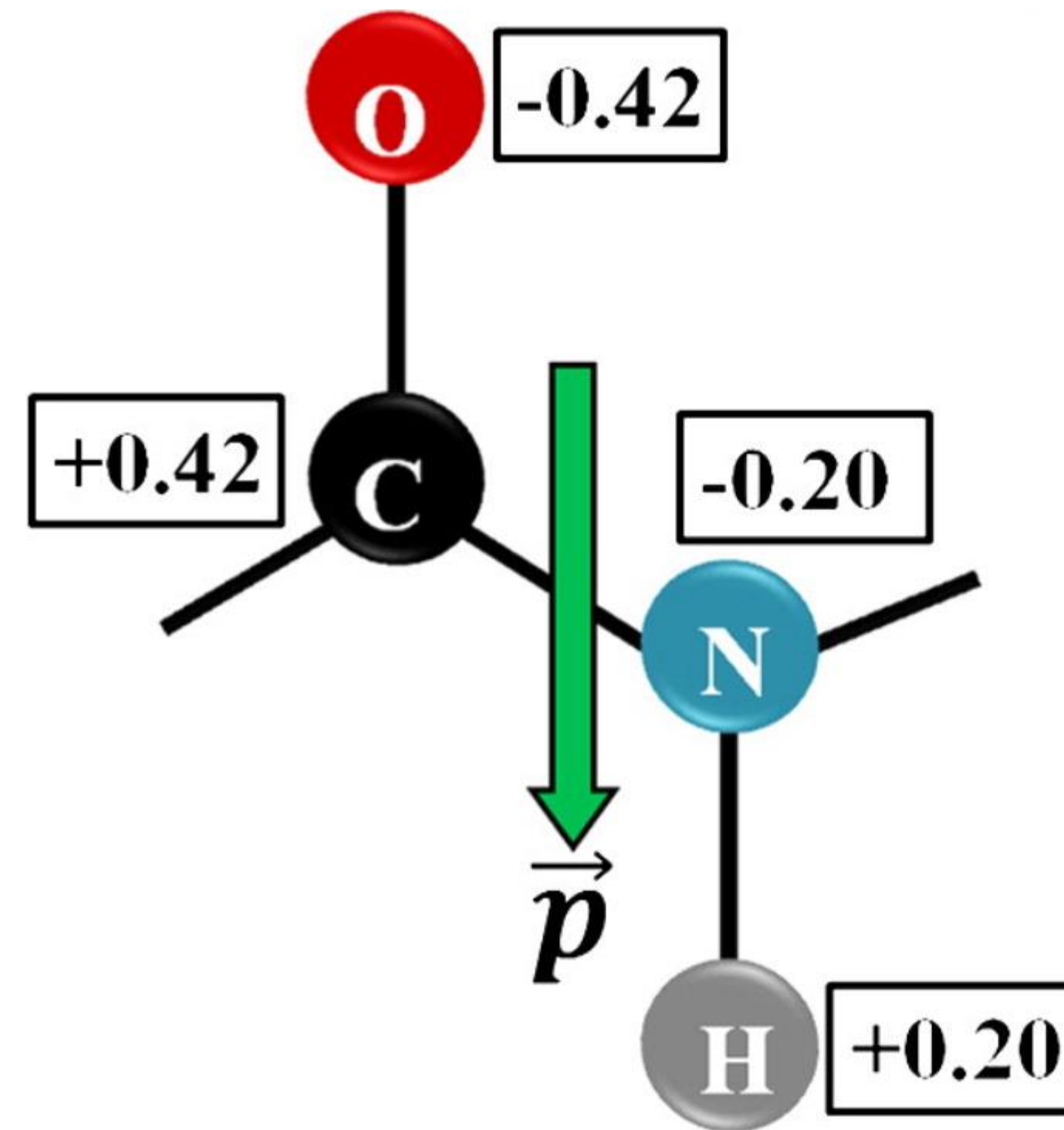
Secondary structure

Secondary structure is defined by the pattern of hydrogen bonds between the amino hydrogen and carboxyl oxygen atoms in the peptide backbone.

Secondary structural elements

alpha helix

dipole moment



Secondary structure

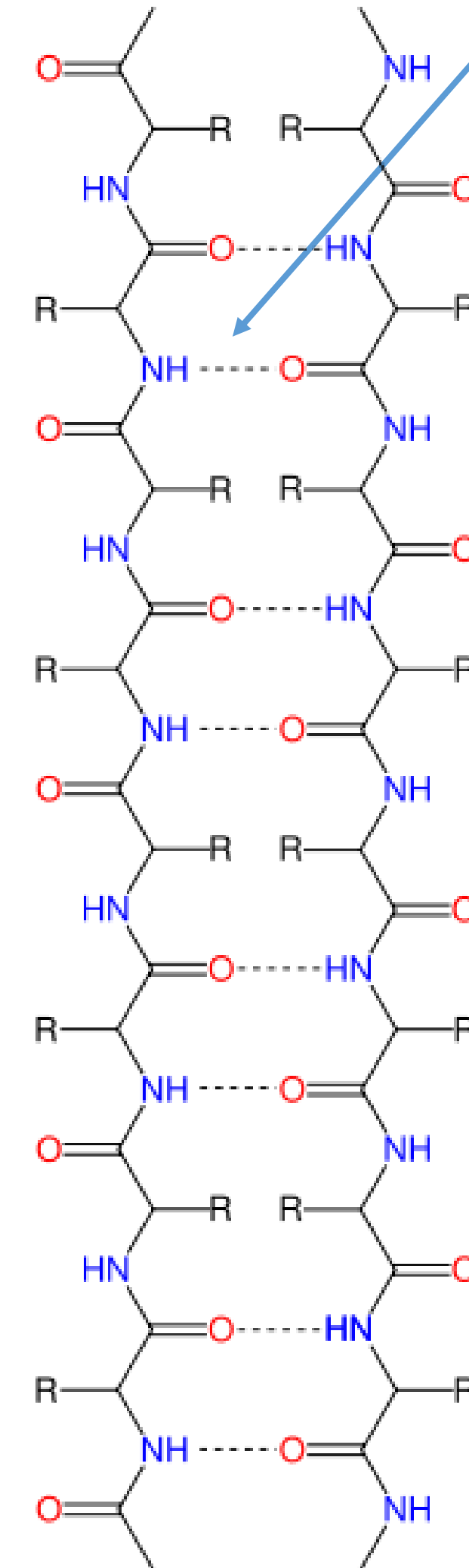
beta strand



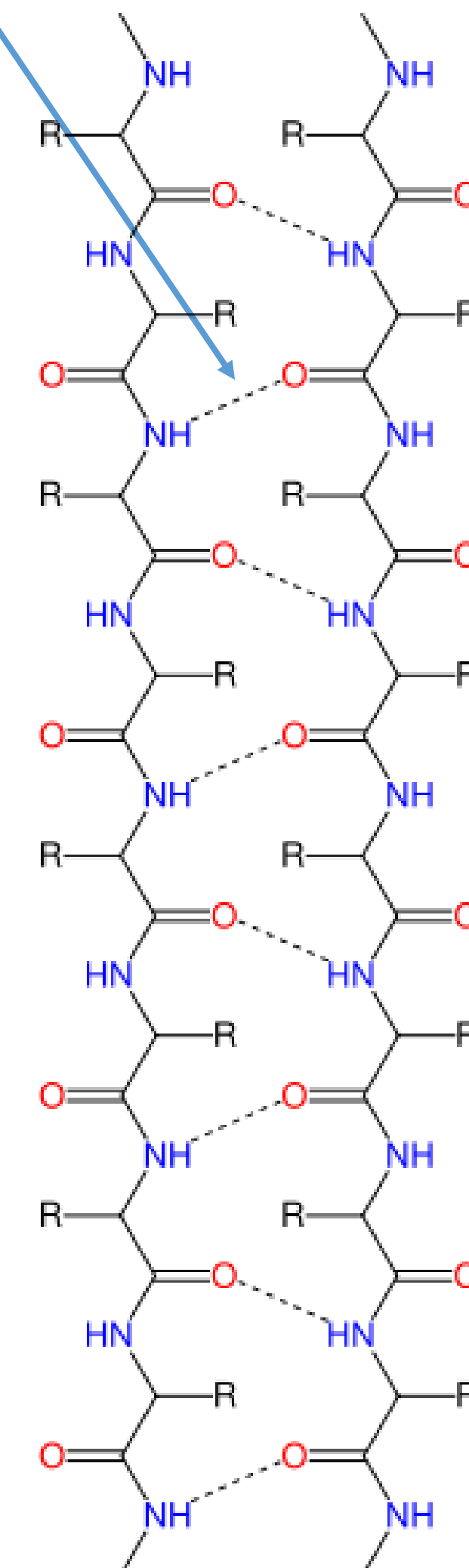
beta sheet



Hydrogen Bonds



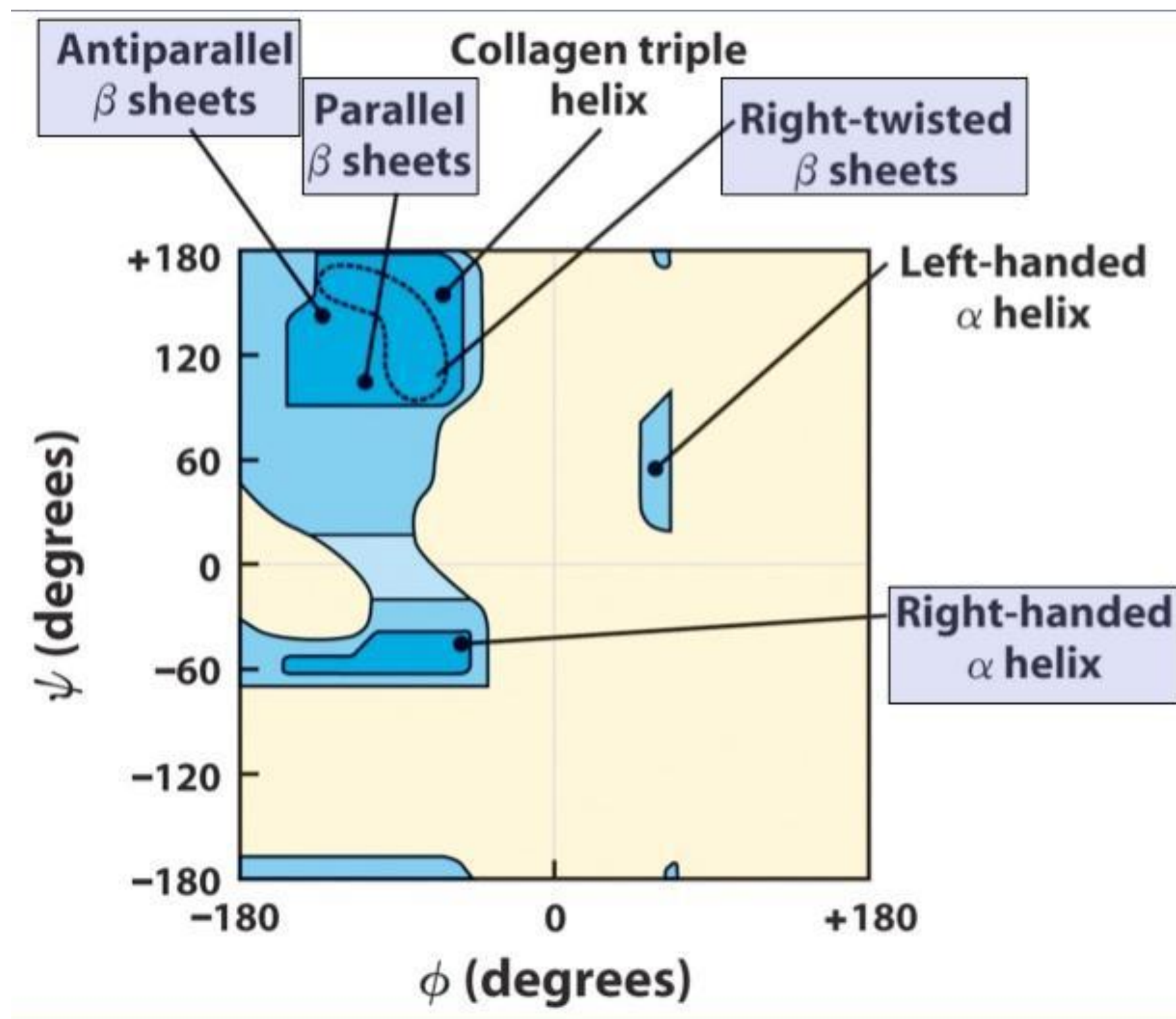
anti-parallel



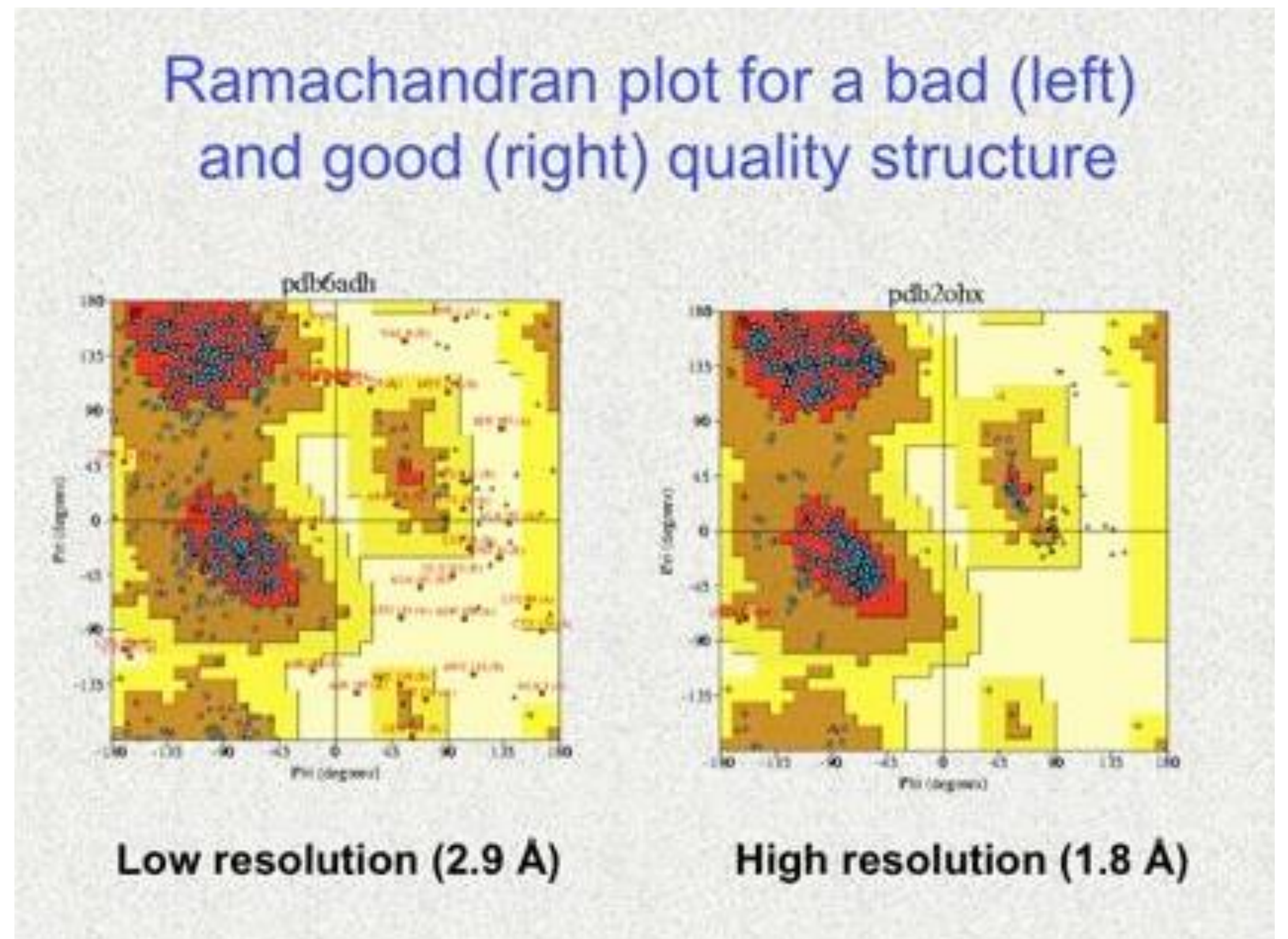
parallel

Ramachandran Plot

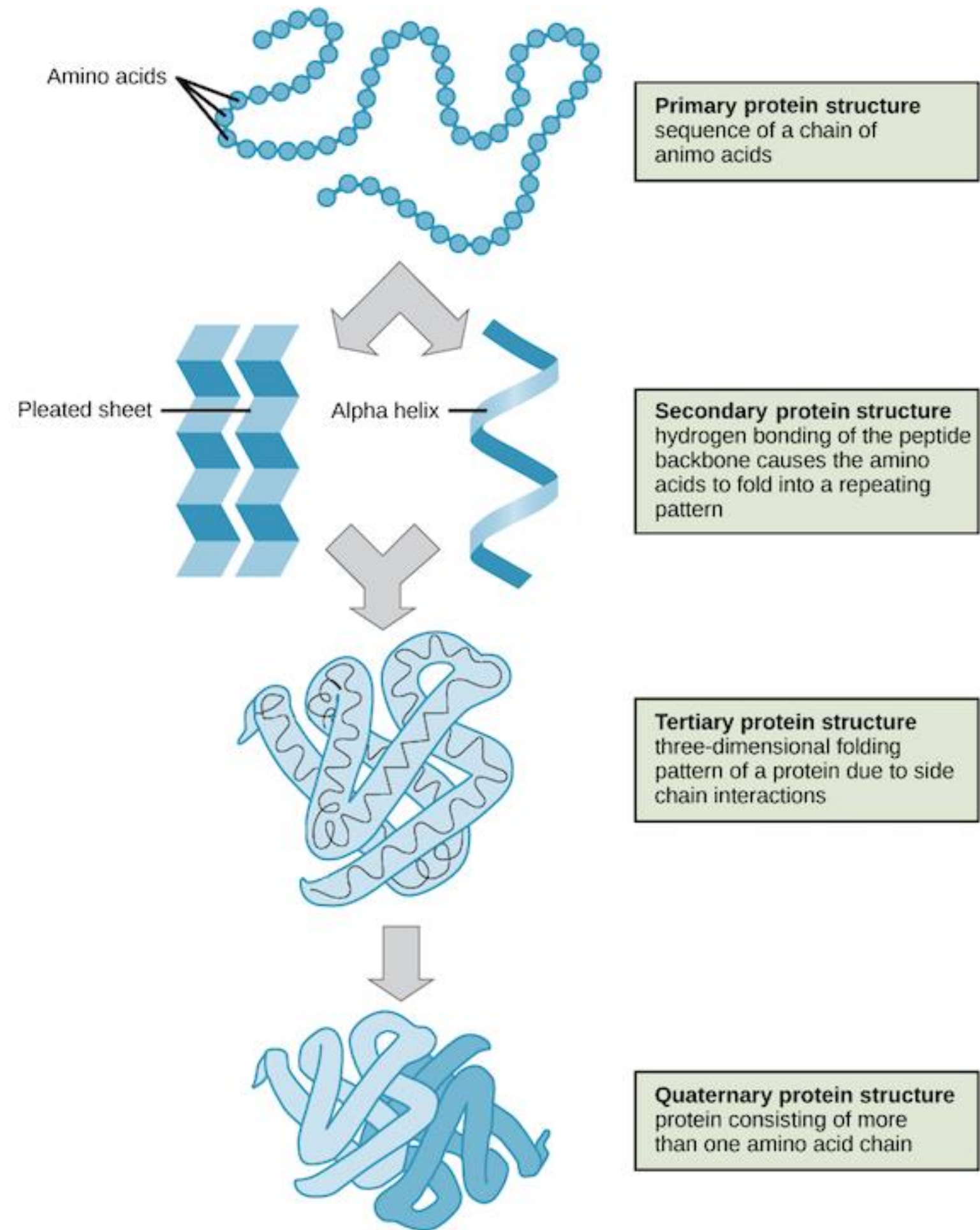
not all phi-psi values are allowed



Ramachandran plot for a bad (left) and good (right) quality structure



Tertiary and quaternary structure



Primary protein structure
sequence of a chain of amino acids

Secondary protein structure
hydrogen bonding of the peptide backbone causes the amino acids to fold into a repeating pattern

Tertiary protein structure
three-dimensional folding pattern of a protein due to side chain interactions

Quaternary protein structure
protein consisting of more than one amino acid chain

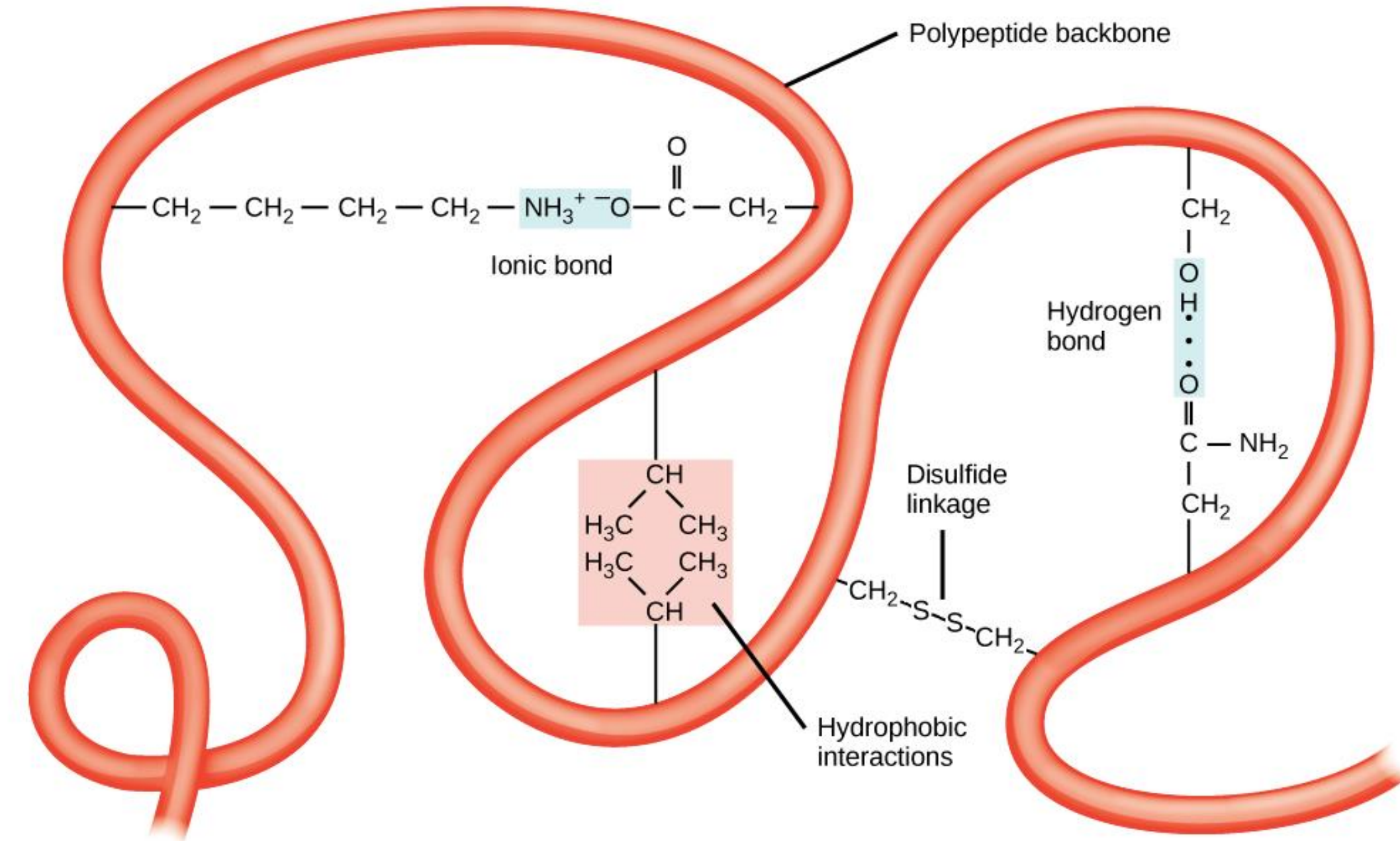
interactions are formed between amino acids

disulphide bonds

hydrogen bonds

ionic interactions

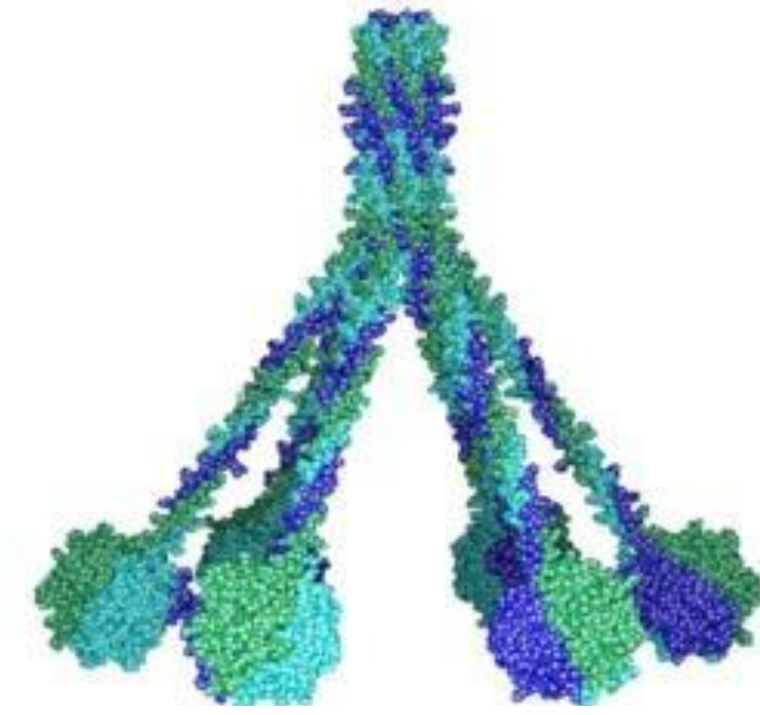
hydrophobic interactions



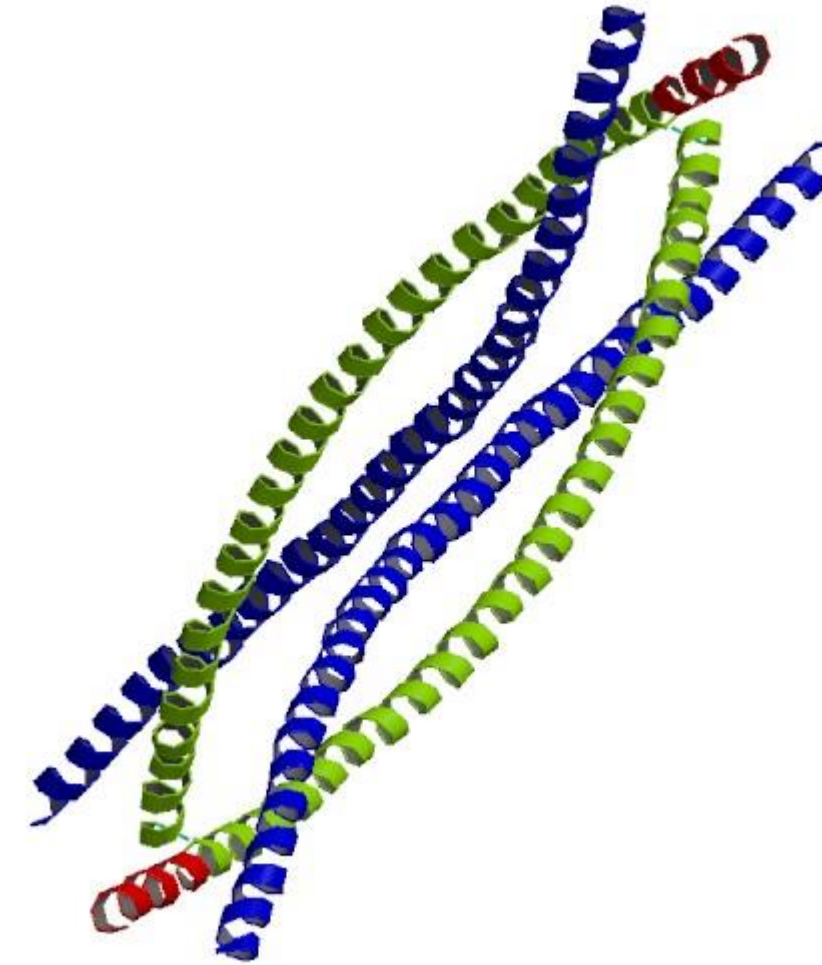
Primary structure amino acid sequence – covalent peptide bonds
 Secondary / supersecondary structures – local interactions
 Tertiary / quaternary structures – longer range interactions

Structural protein with super secondary structures

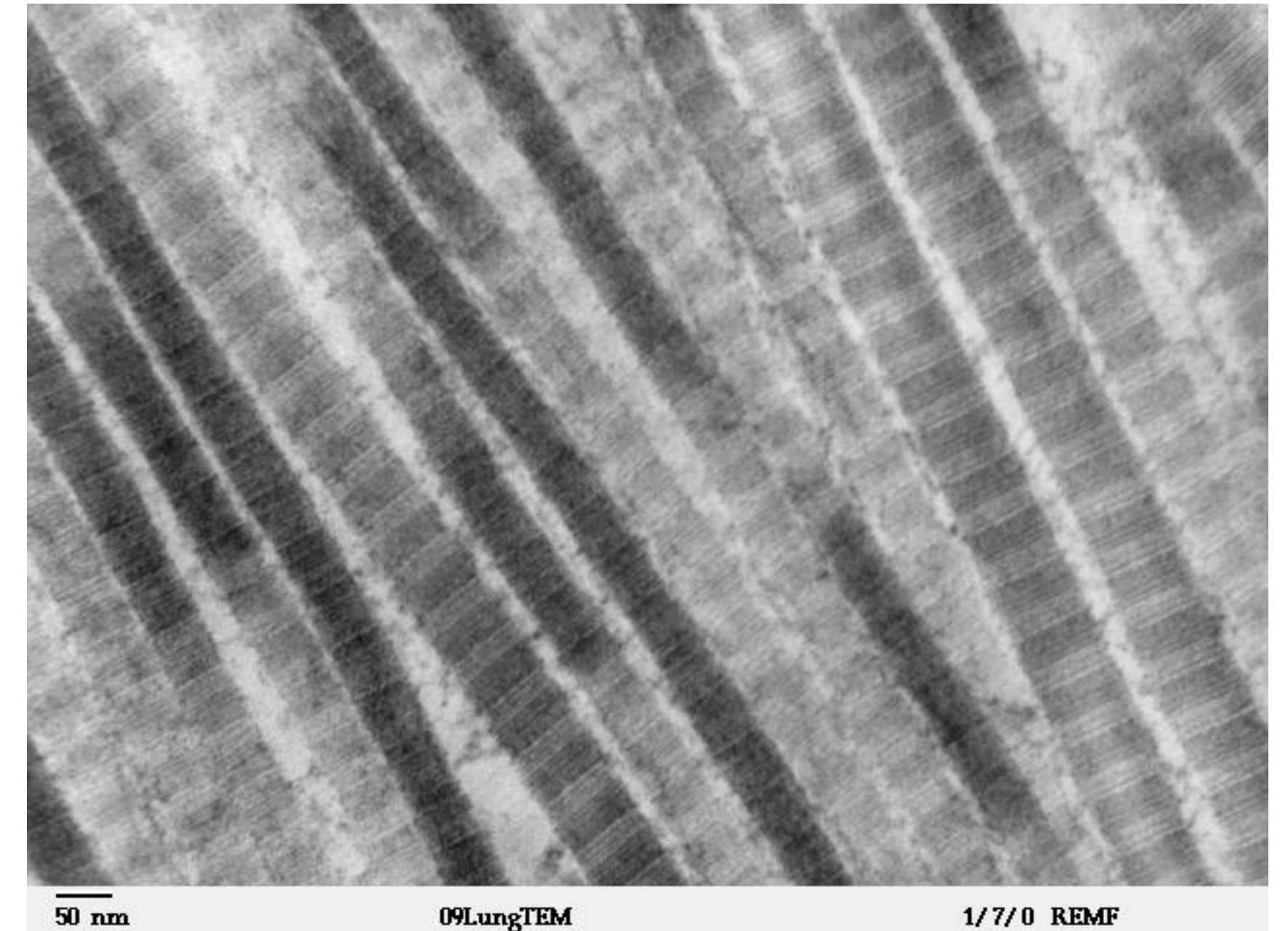
Collagen
Connective tissue,
C1q



Keratin
Hair, nails



Fibroin
Silk



Actin filaments

scaffold for cell motion and division, muscle contraction

globular monomer



filamentous



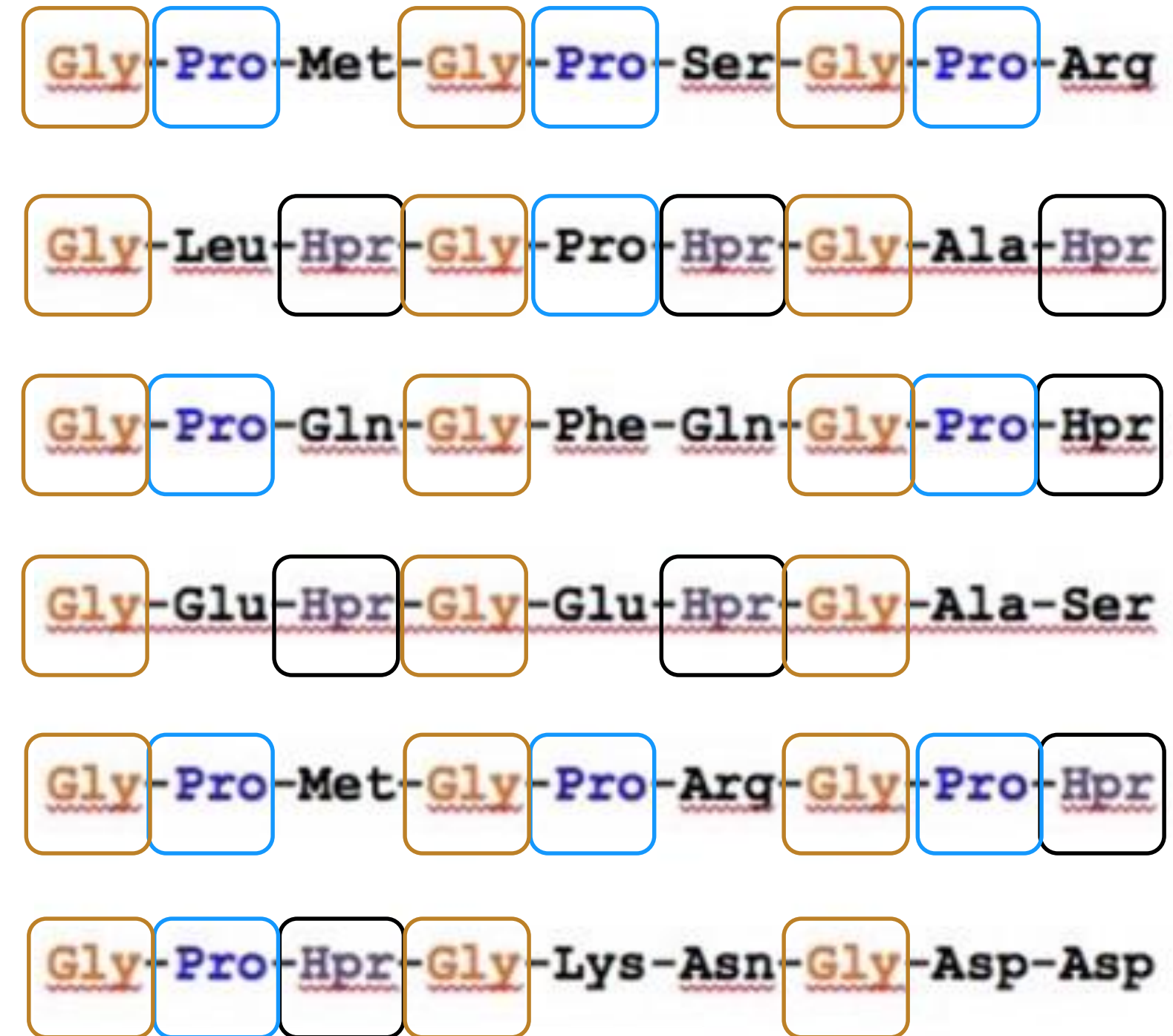
Collagen

Partial Sequence

Hydroxyproline

Proline

Glycine



Structural Proteins

Keratins

Fibrous

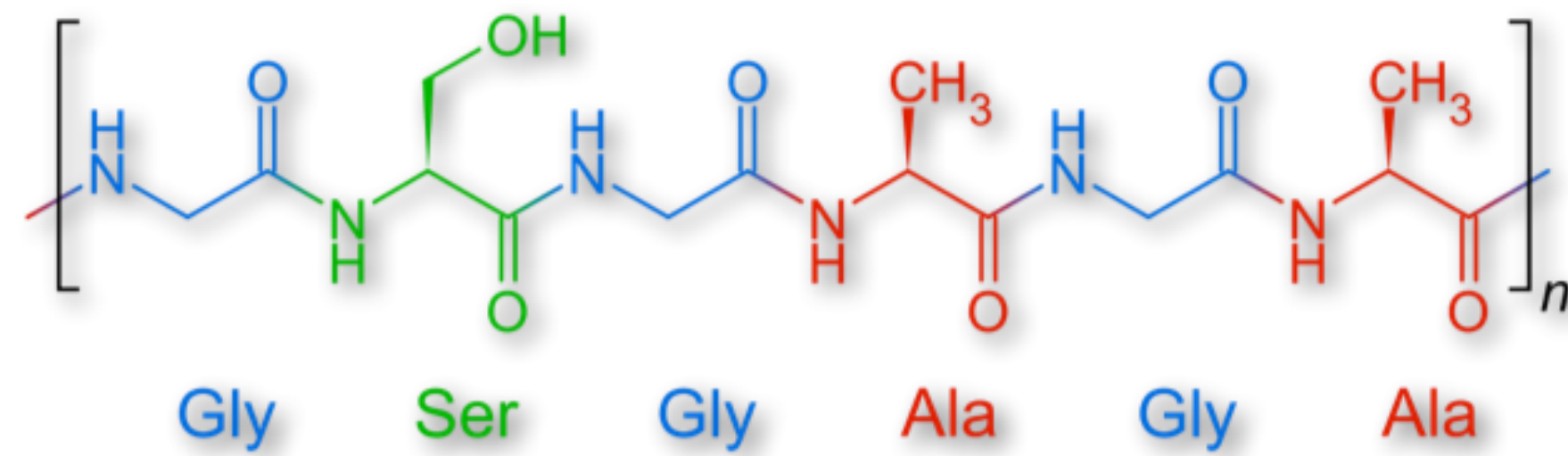
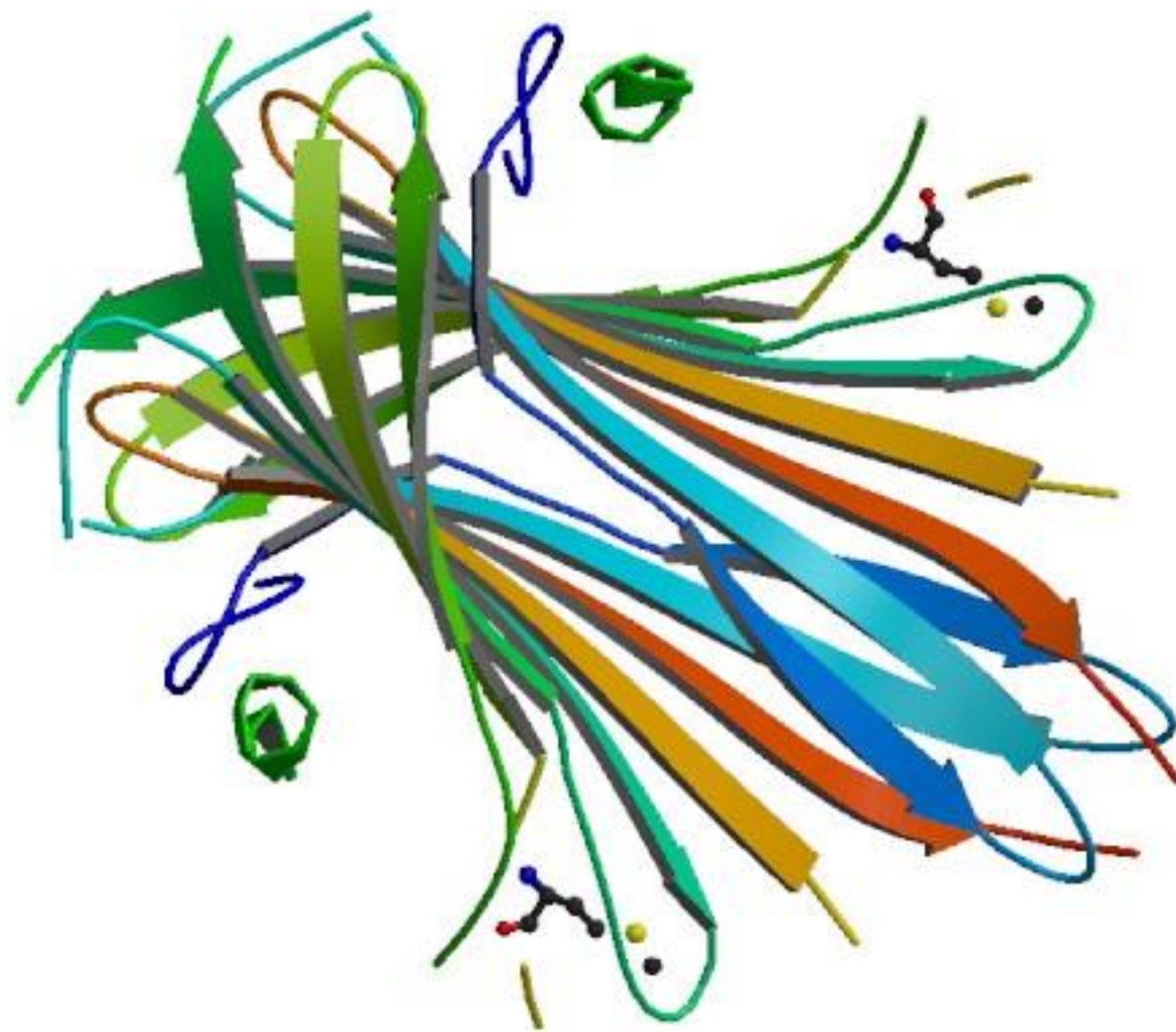
Intermediate Filaments of Cytoskeleton

Hair, nails, horns



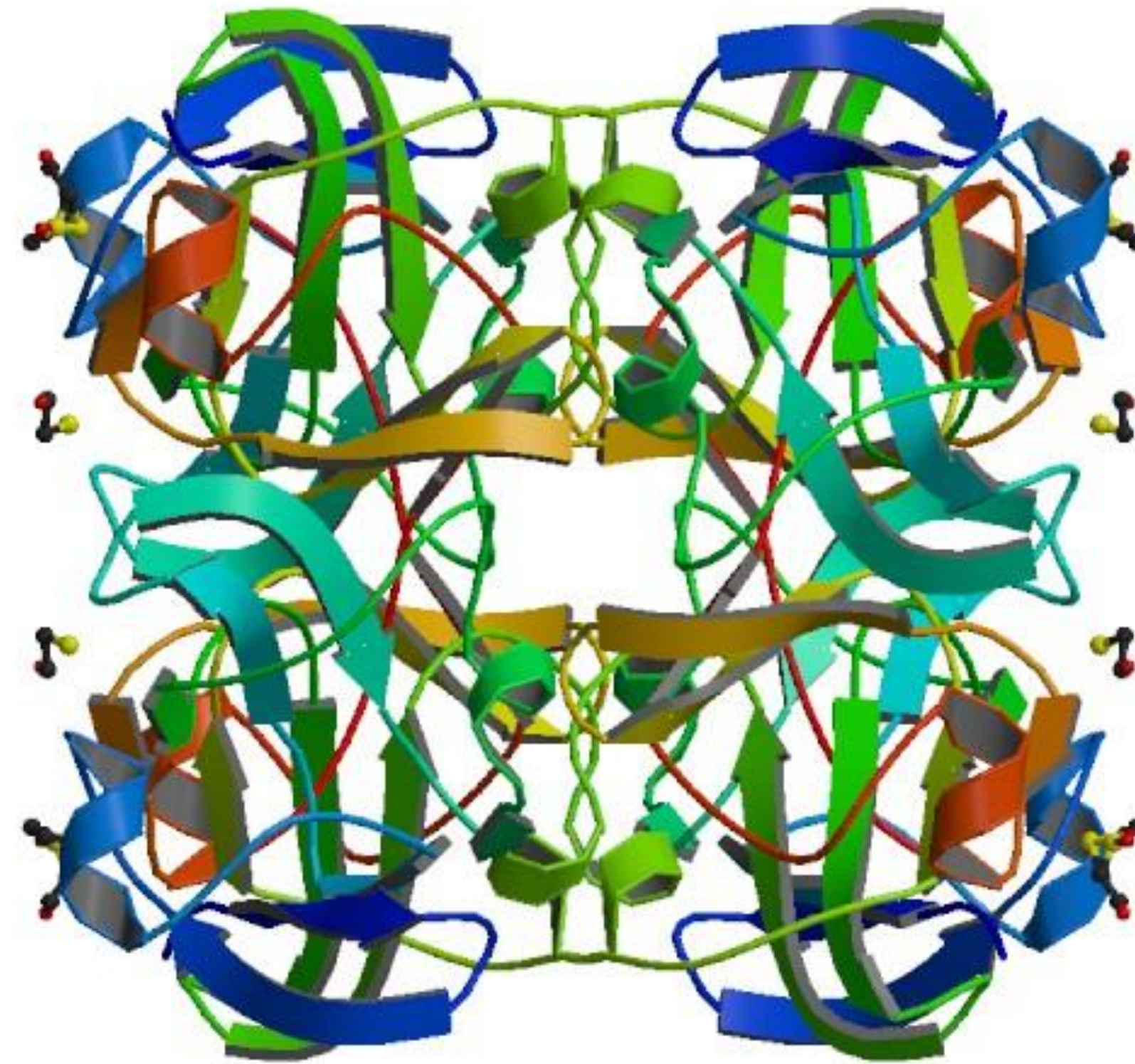
Fibroin

Silk
Beta sheets
Repeating glycines



A curiosity: crystallins

eye lens protein

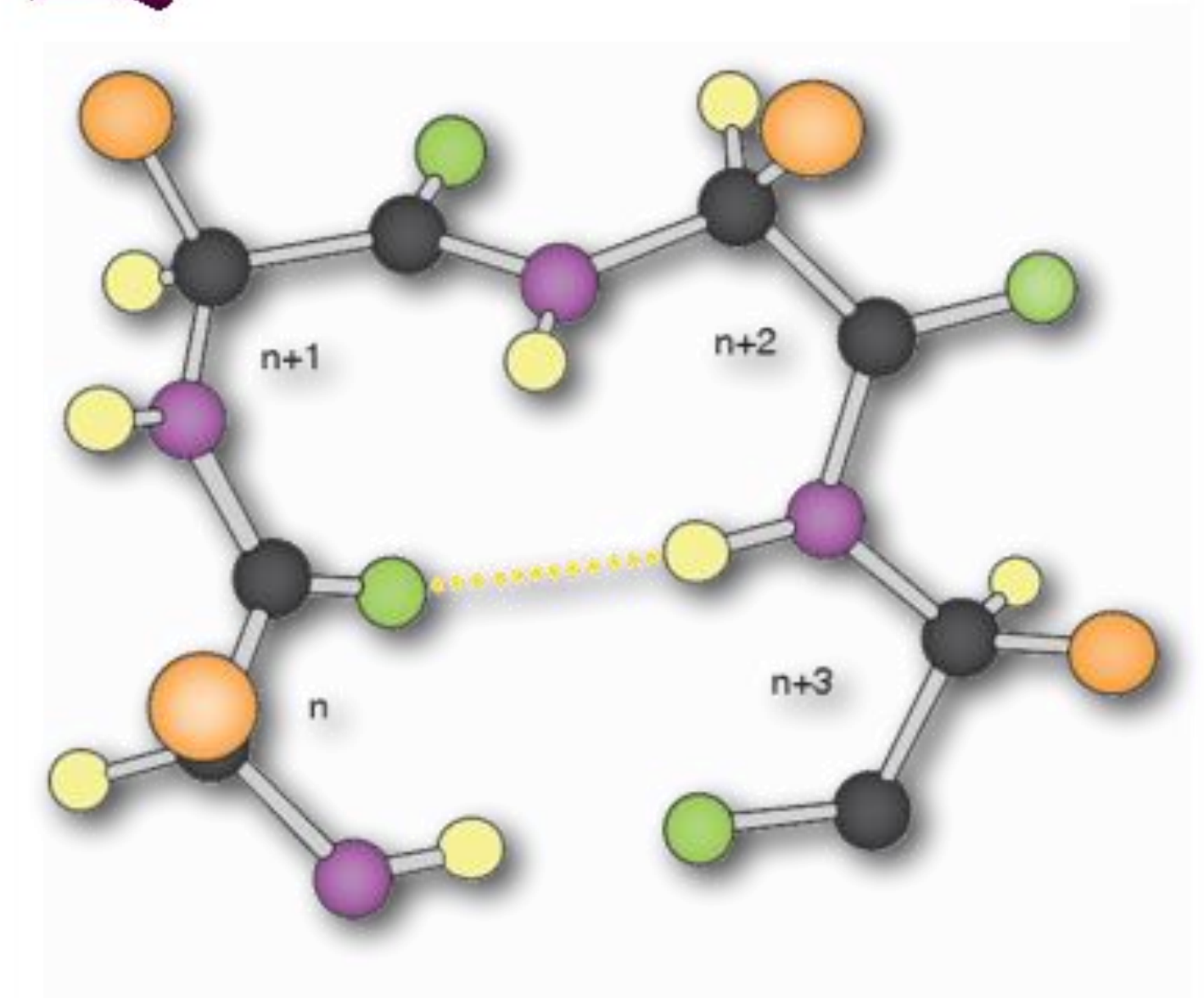
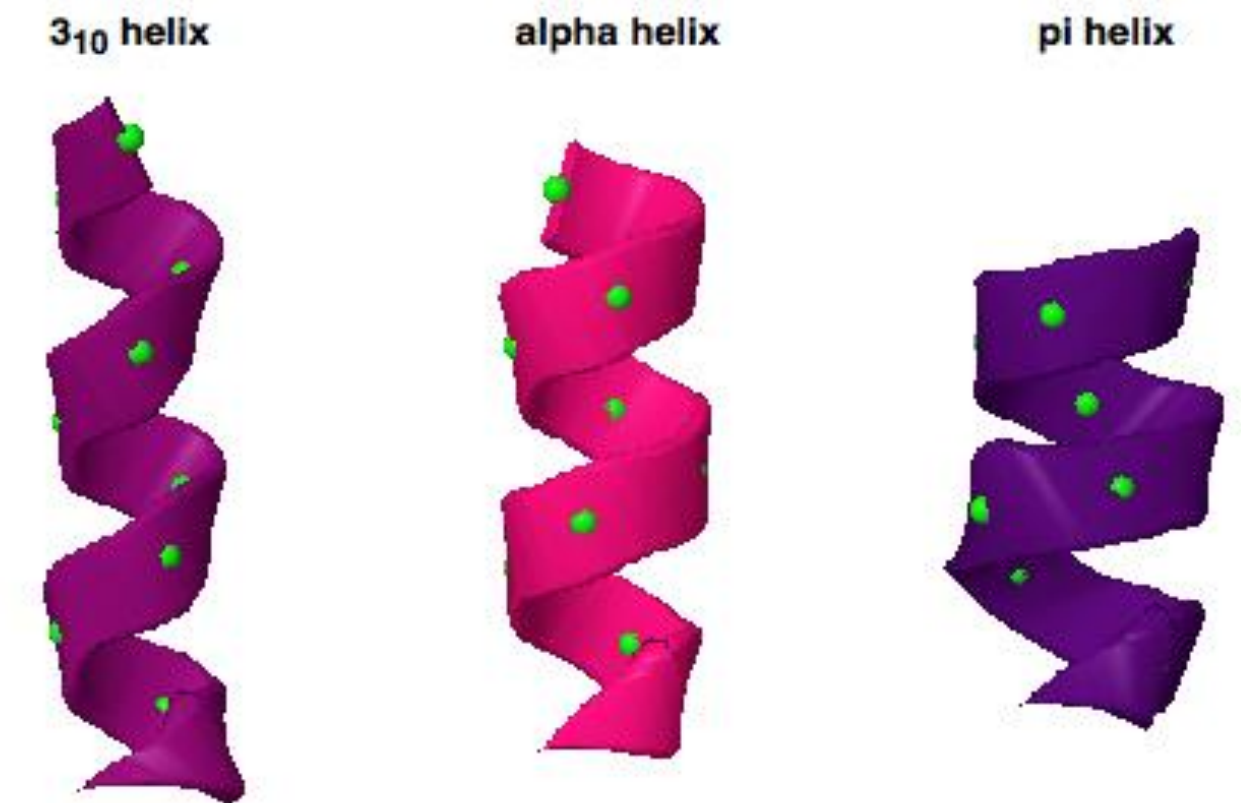


Additional secondary structural elements

DSSP *Define Secondary Structure of Proteins*

Alpha Helix – 3/10 helix

Turns and loops



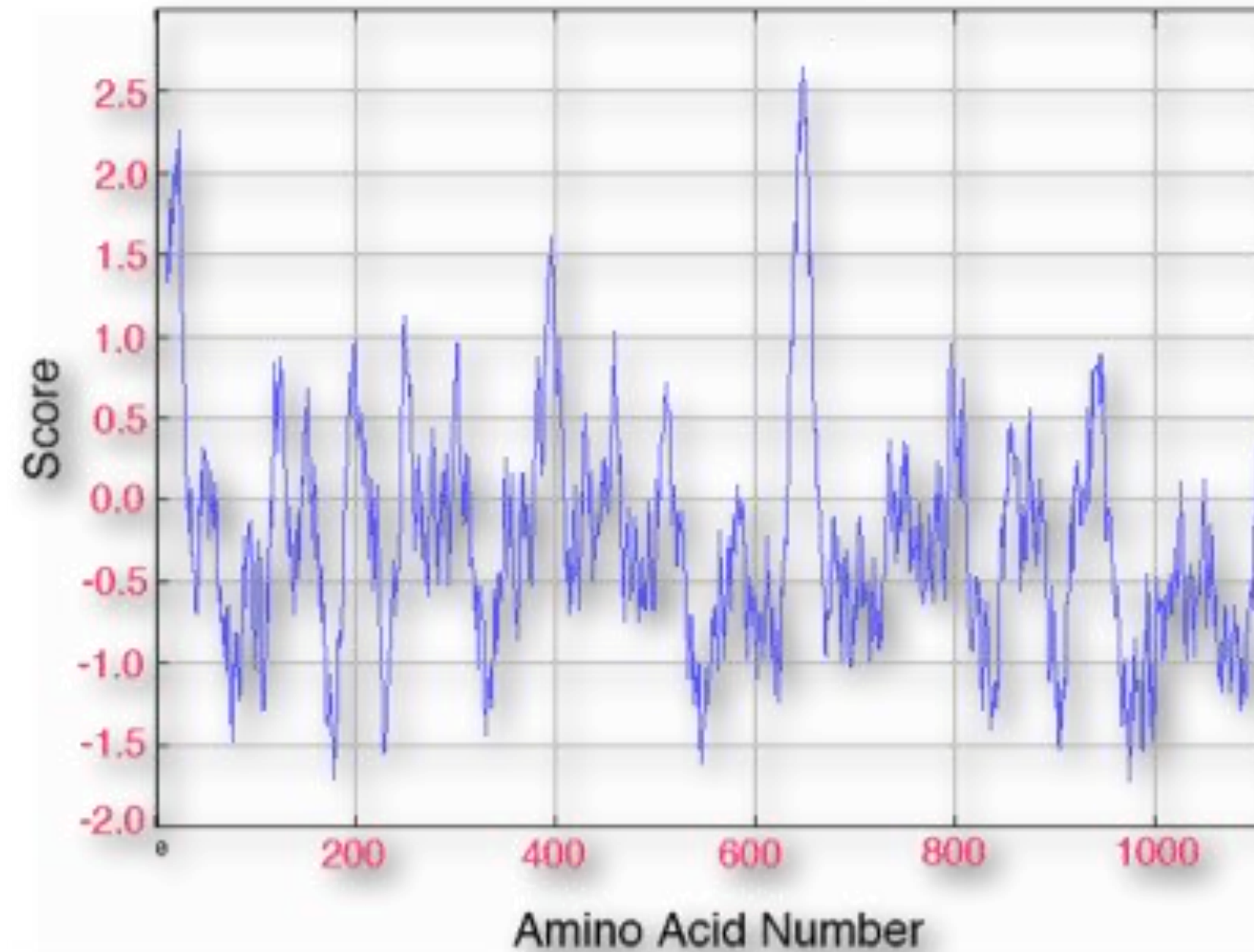
Secondary structural propensities

	Amino Acid	Alpha-H.	Beta-St.	Reverse T.	
	Ala	1.41	0.72	0.82	
	Arg	1.21	0.84	0.9	
	Asn	0.76	0.48	1.34	
High Propensity for Alpha Helices	Asp	0.99	0.39	1.24	
	Cys	0.66	1.4	0.54	
	Gln	1.27	0.98	0.84	
	Glu	1.59	0.52	1.01	
	Gly	0.43	0.58	1.77	
	His	1.05	0.8	0.81	
	Ile	1.09	1.67	0.47	
High Propensity for Beta Strands	Leu	1.34	1.22	0.57	High Propensity for Reverse Turns
	Lys	1.23	0.69	1.07	
	Met	1.3	1.14	0.52	
	Phe	1.16	1.33	0.59	
	Pro	0.34	0.31	1.32	Proline SS breaker
	Ser	0.57	0.96	1.22	
	Thr	0.76	1.17	0.96	
	Trp	1.02	1.35	0.65	
	Tyr	0.74	1.45	0.76	
	Val	0.9	1.87	0.41	

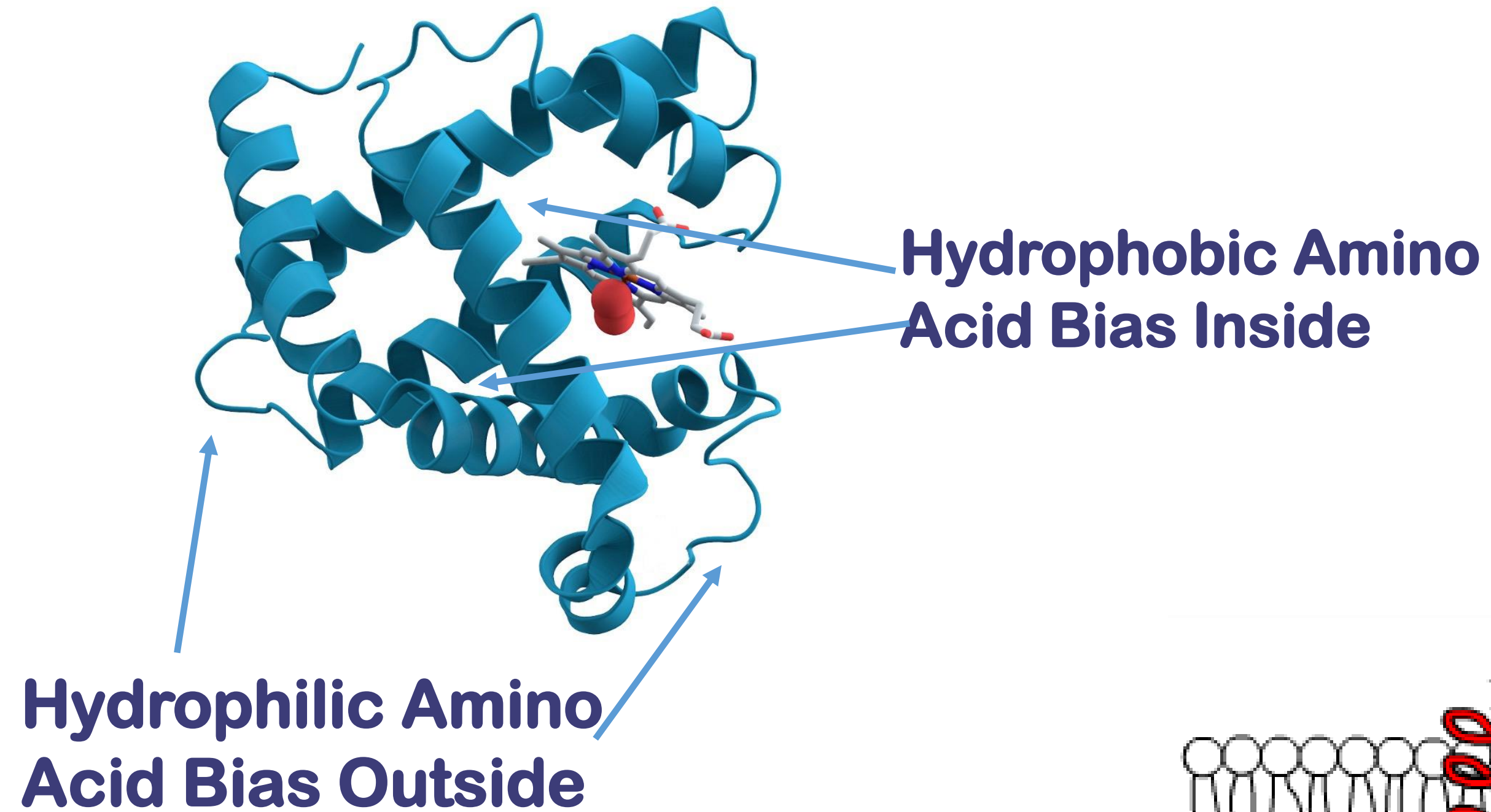
Amino Acid Hydropathy

Amino Acid Hydropathy Scores

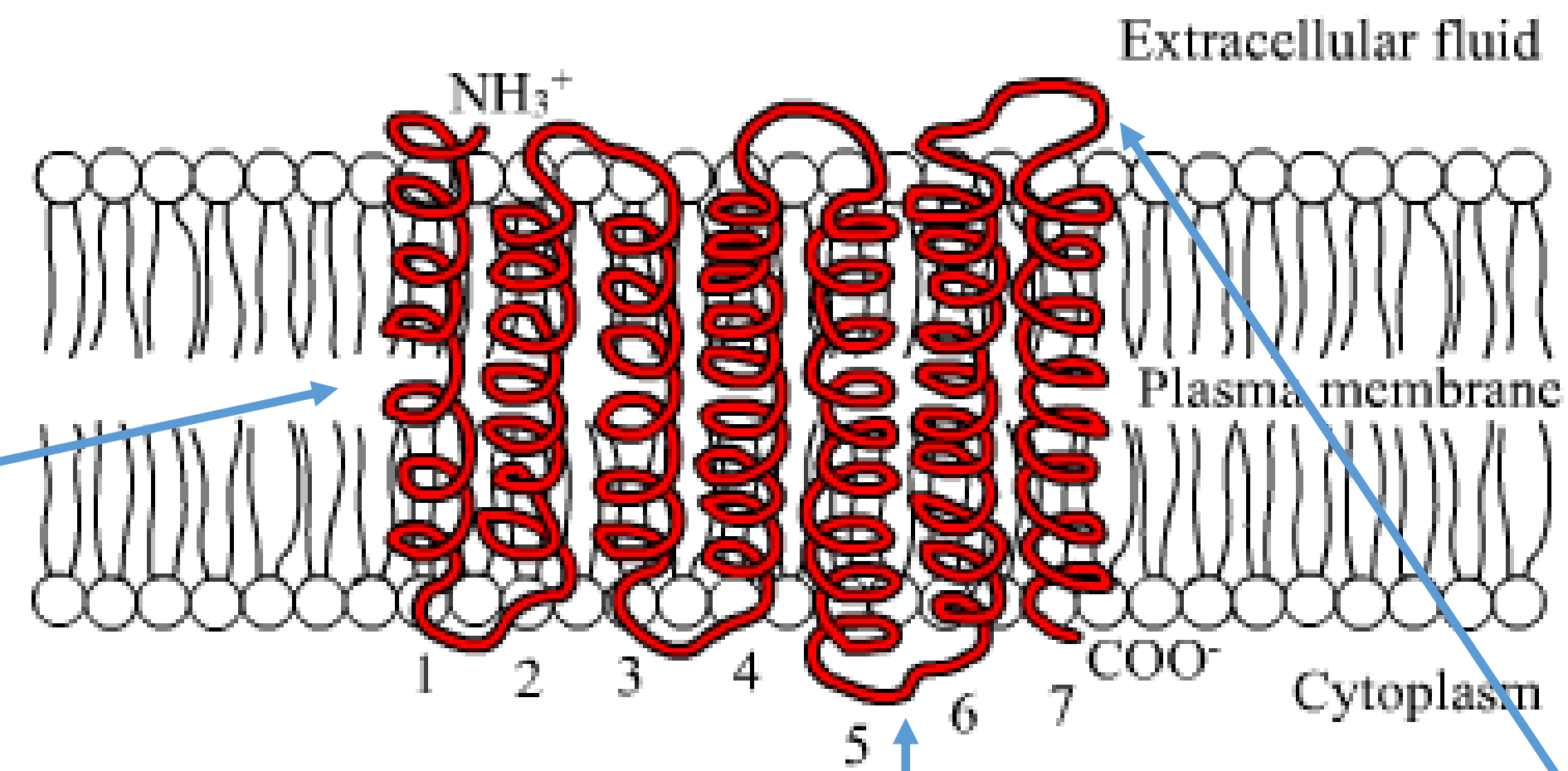
Amino Acid	One Letter Code	Hydropathy Score
Isoleucine	I	4.5
Valine	V	4.2
Leucine	L	3.8
Phenylalanine	F	2.8
Cysteine	C	2.5
Methionine	M	1.9
Alanine	A	1.8
Glycine	G	-0.4
Threonine	T	-0.7
Tryptophan	W	-0.9
Serine	S	-0.8
Tyrosine	Y	-1.3
Proline	P	-1.6
Histidine	H	-3.2
Glutamic acid	E	-3.5
Glutamine	Q	-3.5
Aspartic acid	D	-3.5
Asparagine	N	-3.5
Lysine	K	-3.9
Arginine	R	-4.5



Water soluble and membrane proteins

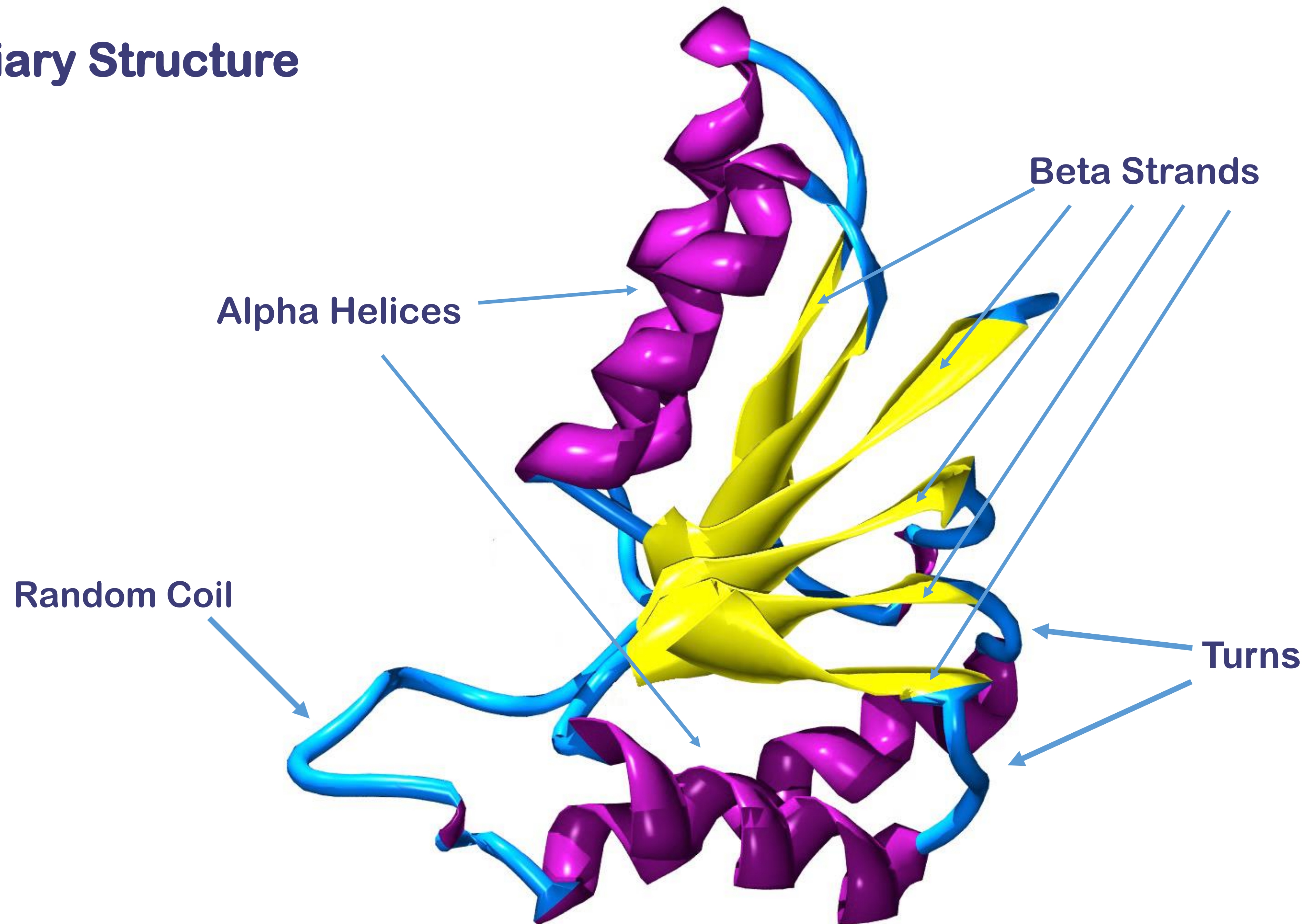


Hydrophobic Amino Acid Bias In Bilayer

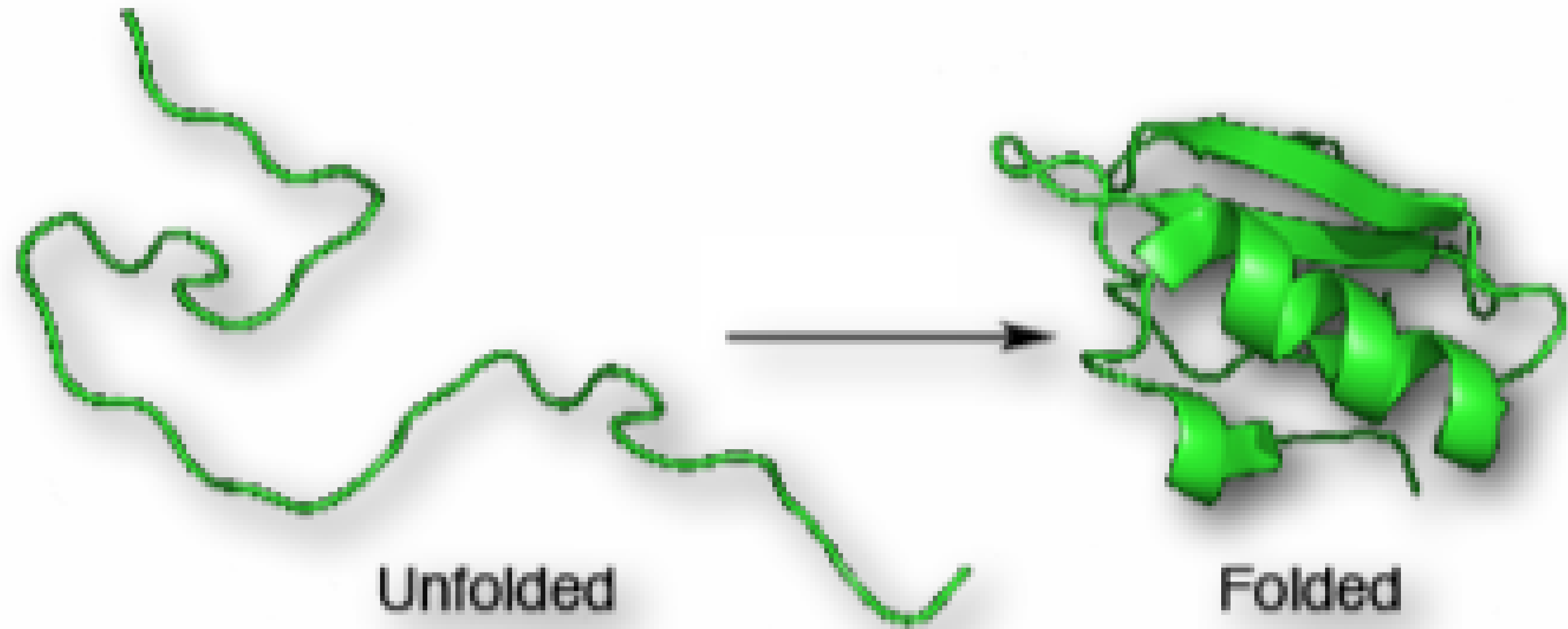


Hydrophilic Amino Acid Bias Outside of Bilayer

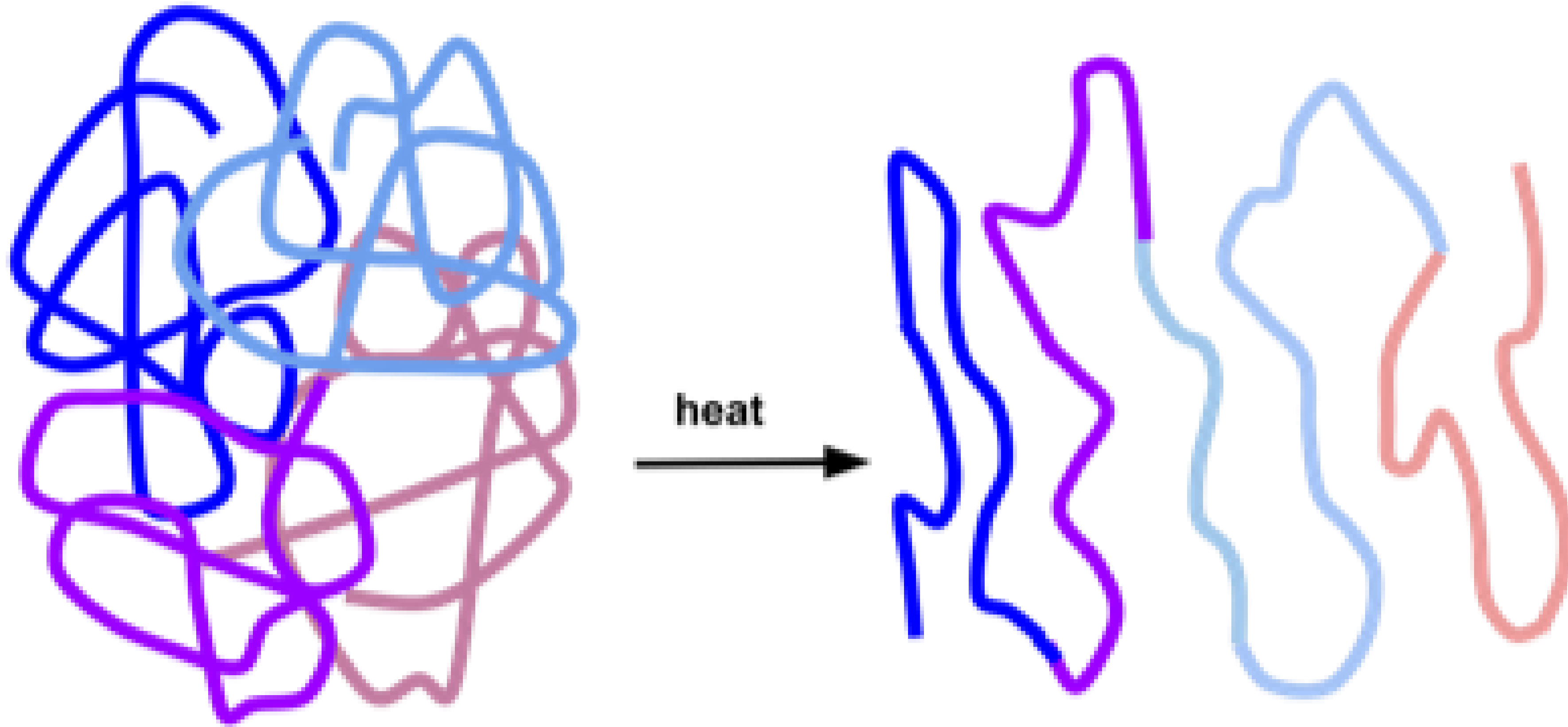
Tertiary Structure



Folding of a Globular Protein



Unfolding of a Globular Protein

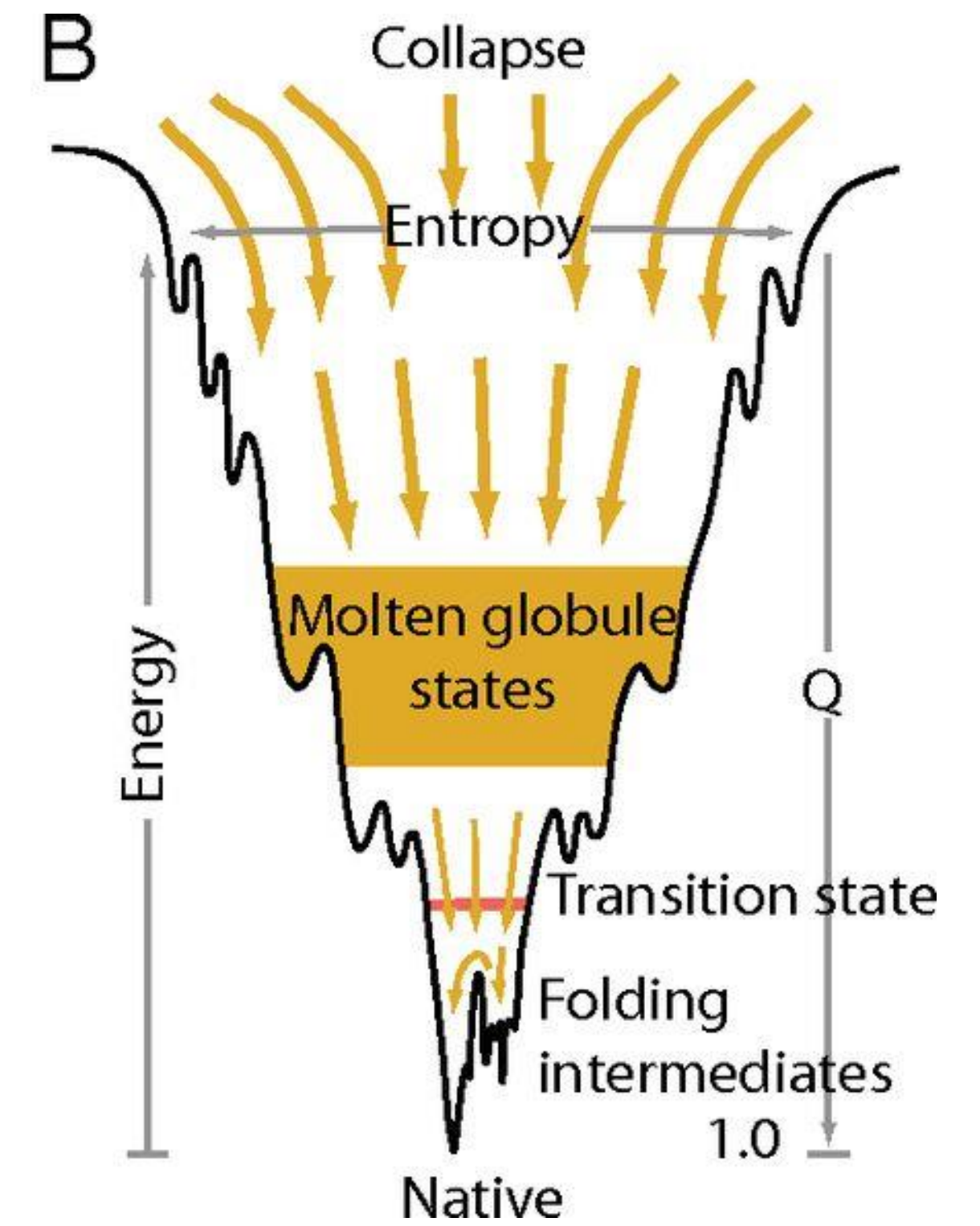
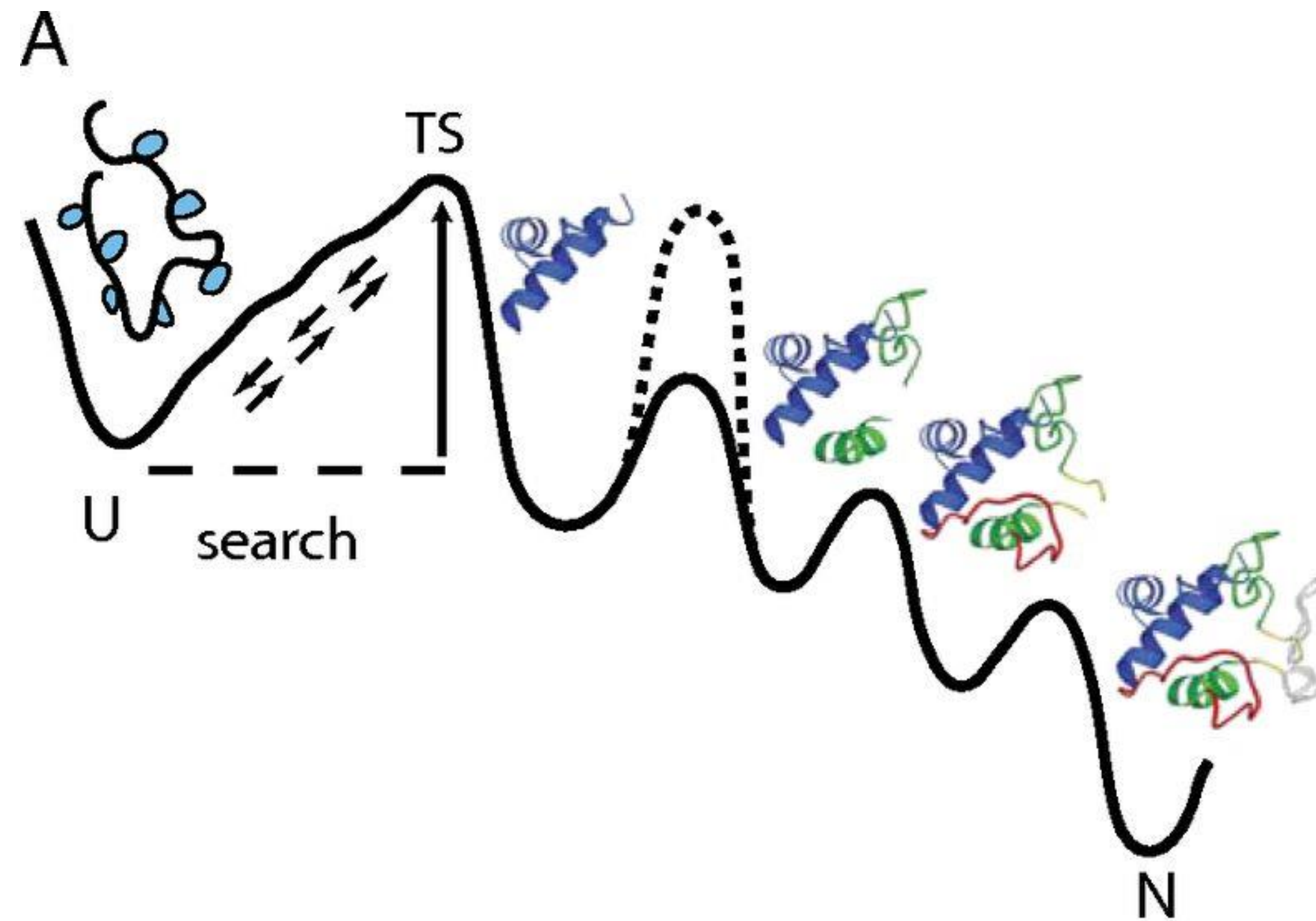


Energetics of folding

Free enthalpy (Gibbs free energy)

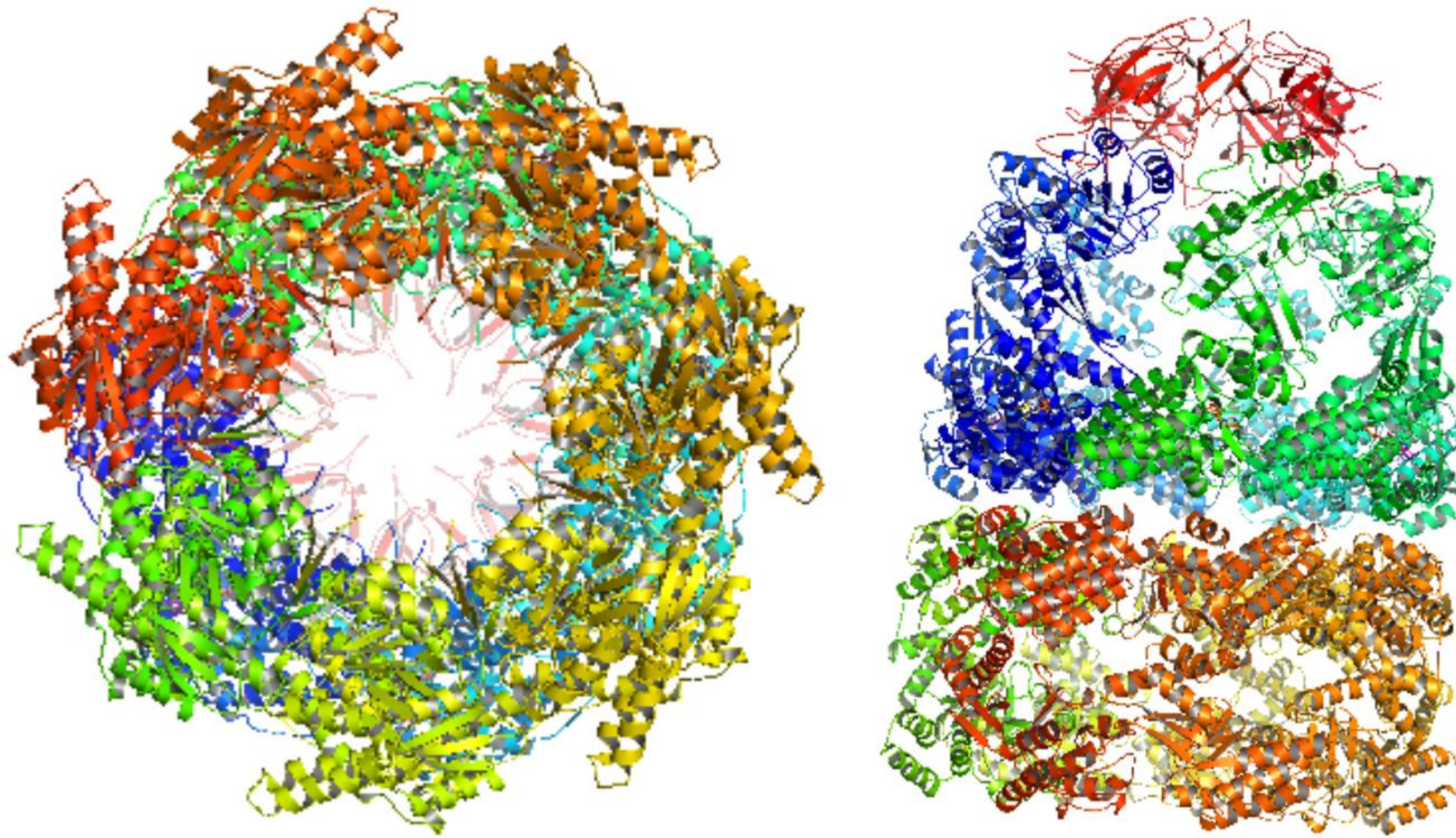
thermodynamic potential describing the system at constant T and p

$$\Delta G = \Delta H - T\Delta S$$



Chaperons / Heat Shock Proteins

provide proper environment for trapped folding / help in renaturation



Levinthal's paradoxon

even if we represent a protein structure by only the dihedral angles the number of the degrees of freedom is huge

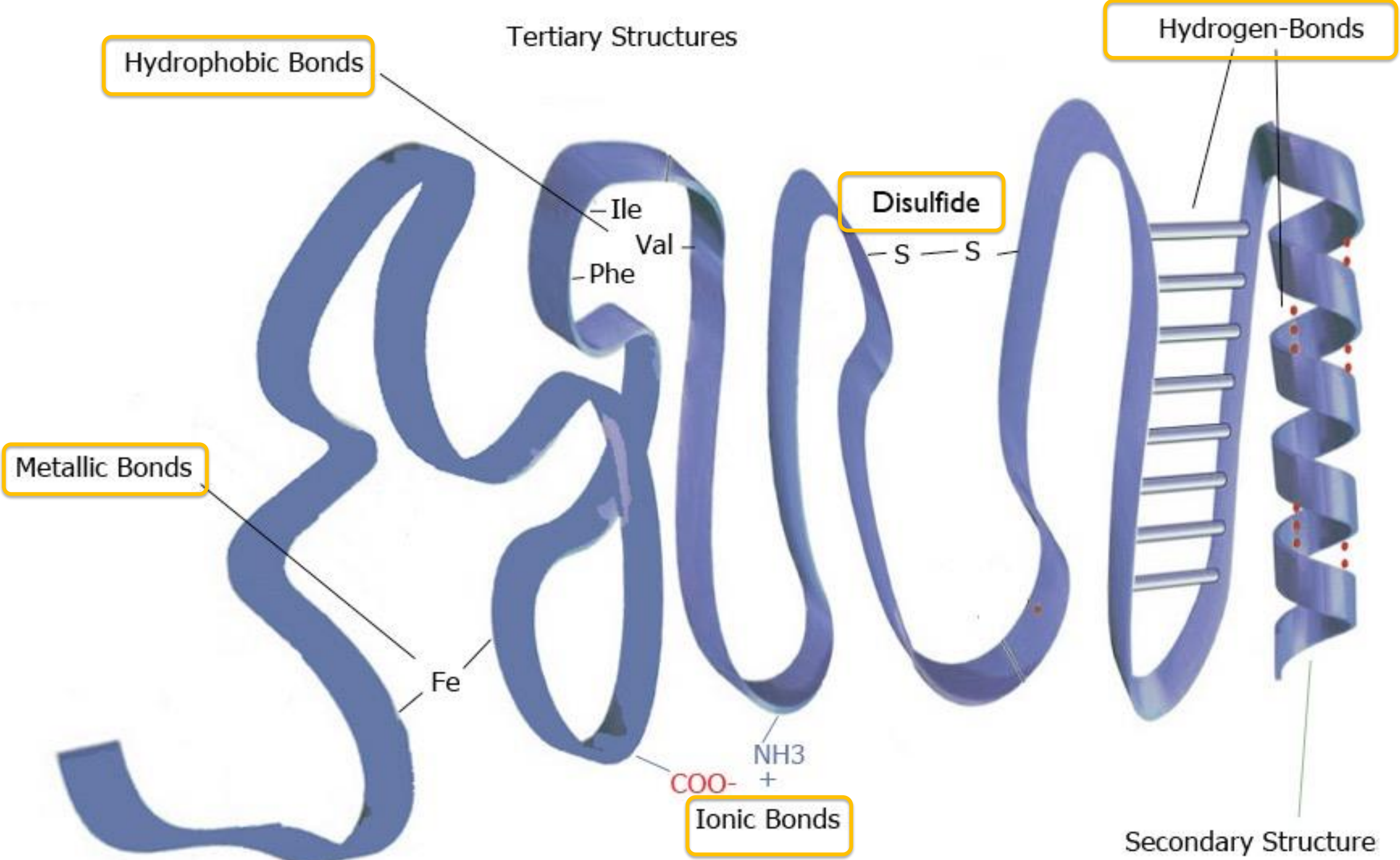
a protein can not go through all possible conformations to find the lowest energy

at 1 conformation/ps rate folding would take longer than the age of the universe

folding usually occurs at the ms – s timescale

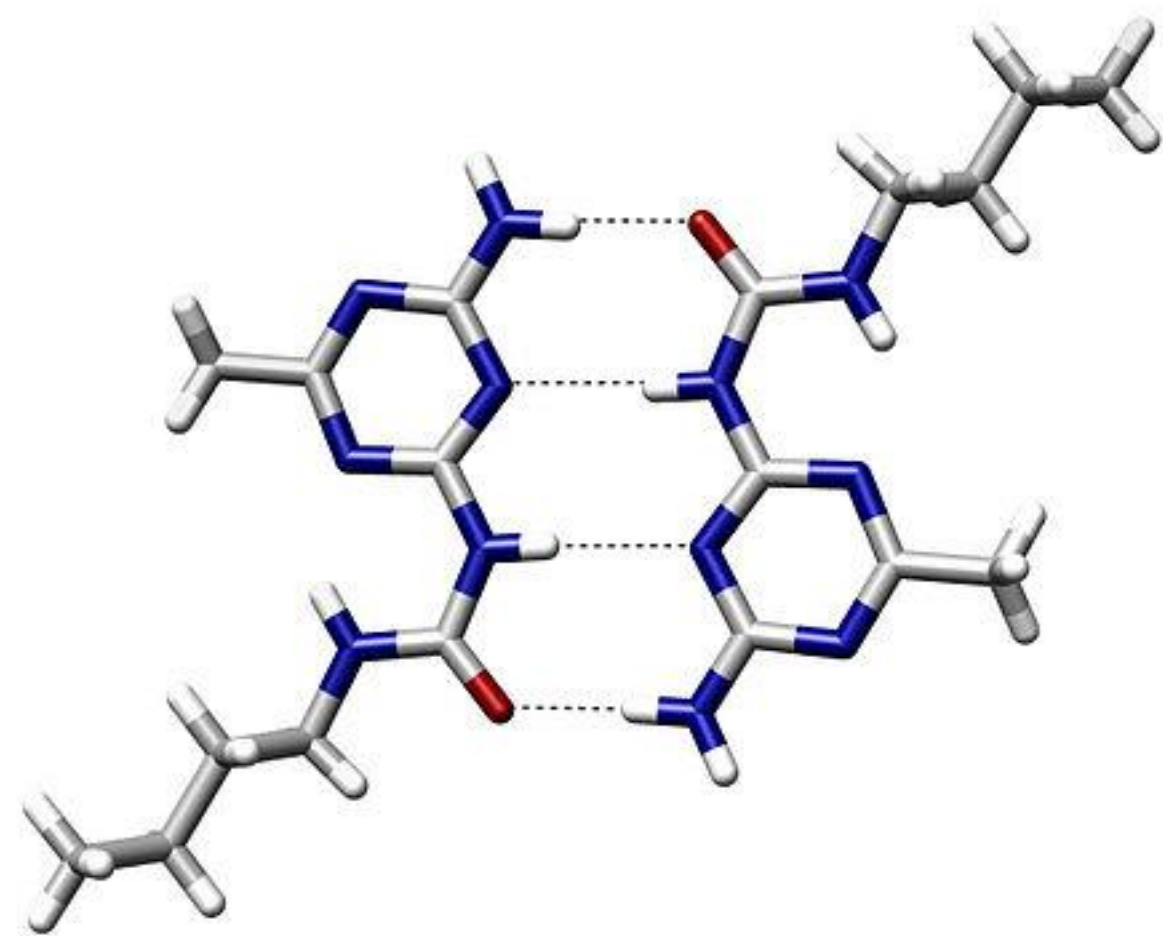
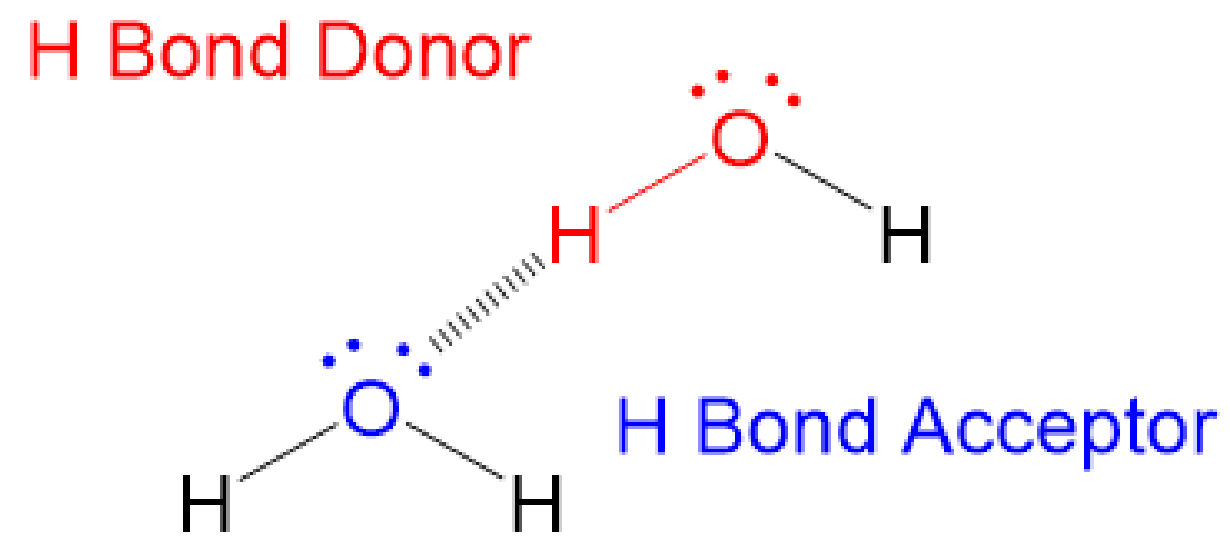
resolution: guided folding by local interactions, subunit folding

Forces Stabilizing Tertiary Structure

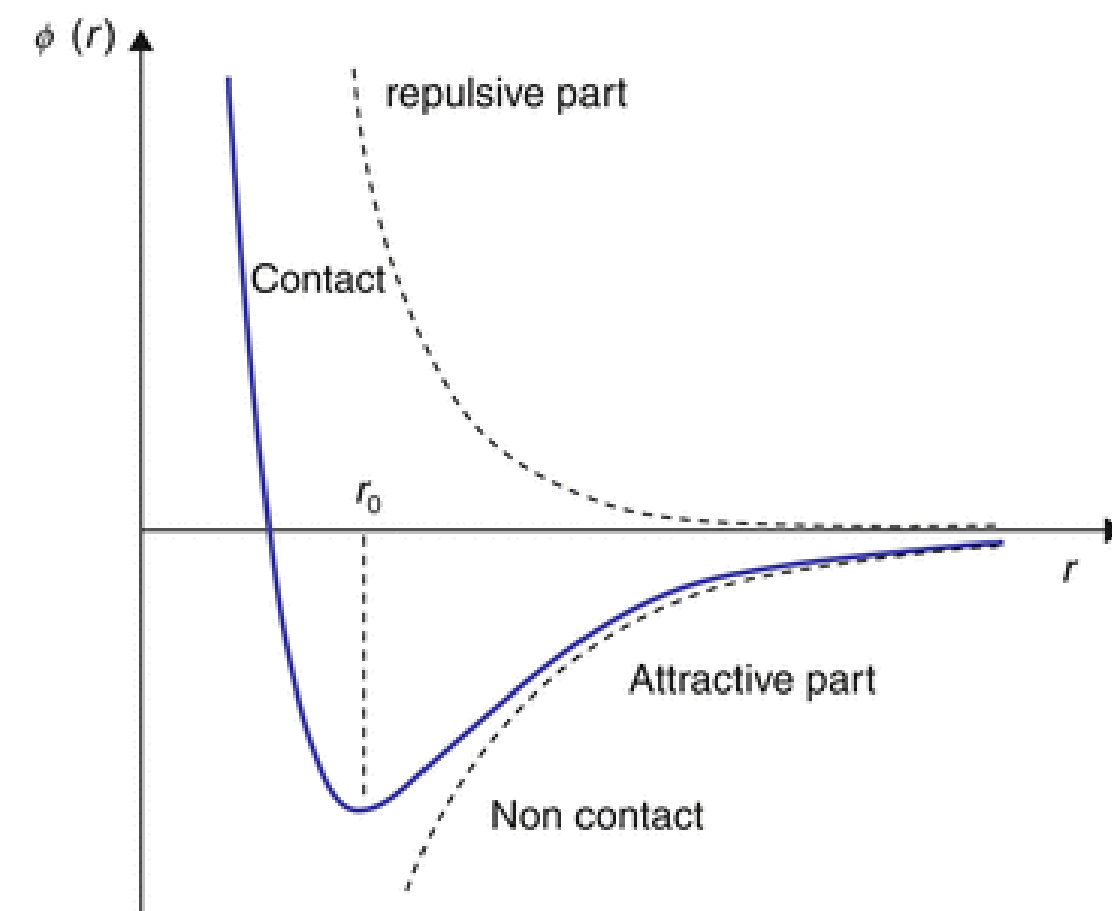
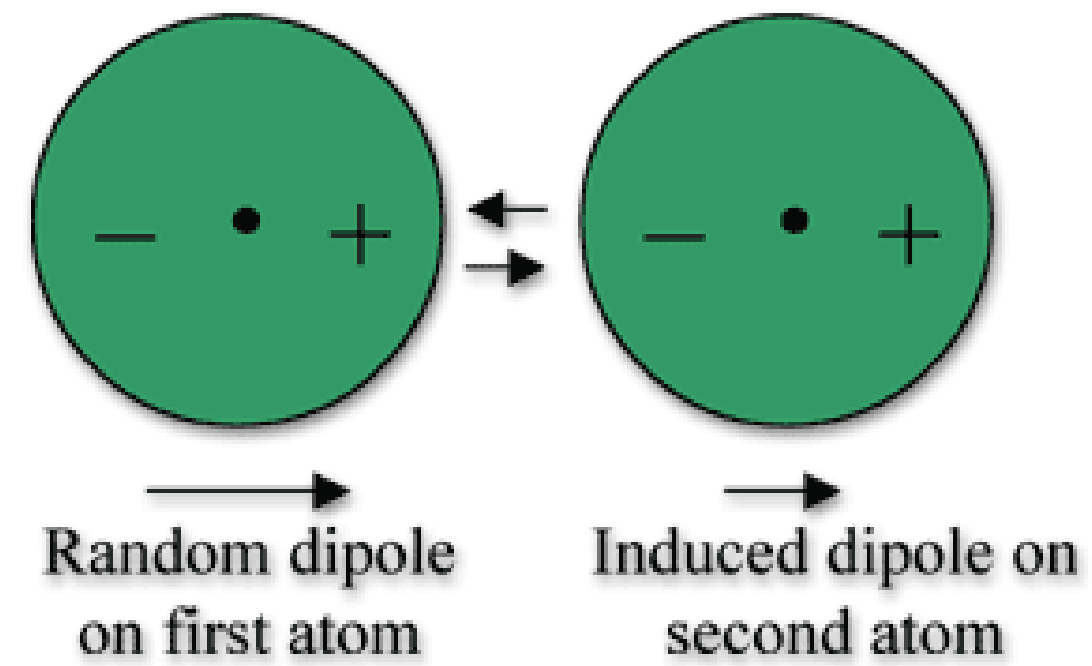


Forces Stabilizing Tertiary Structure

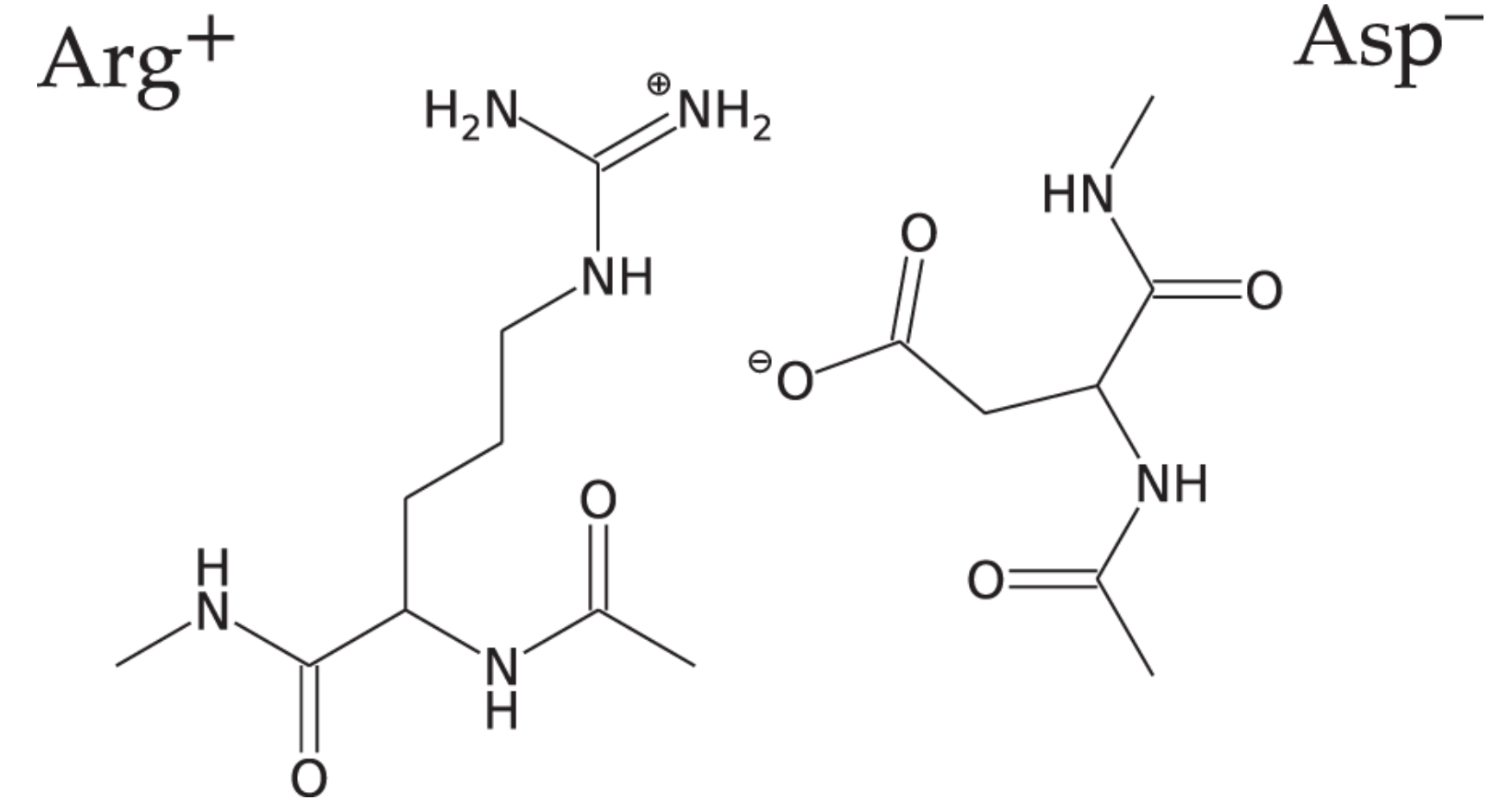
Hydrogen Bonds



Van der Waals forces

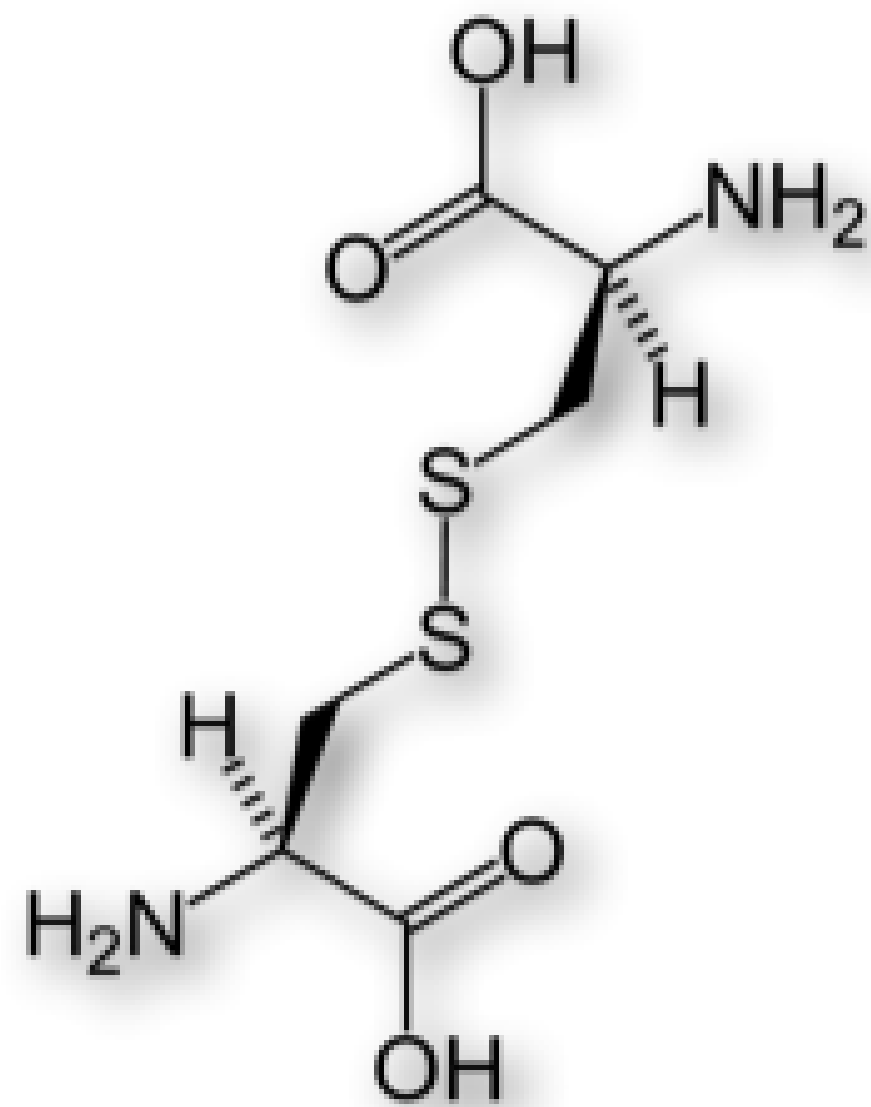
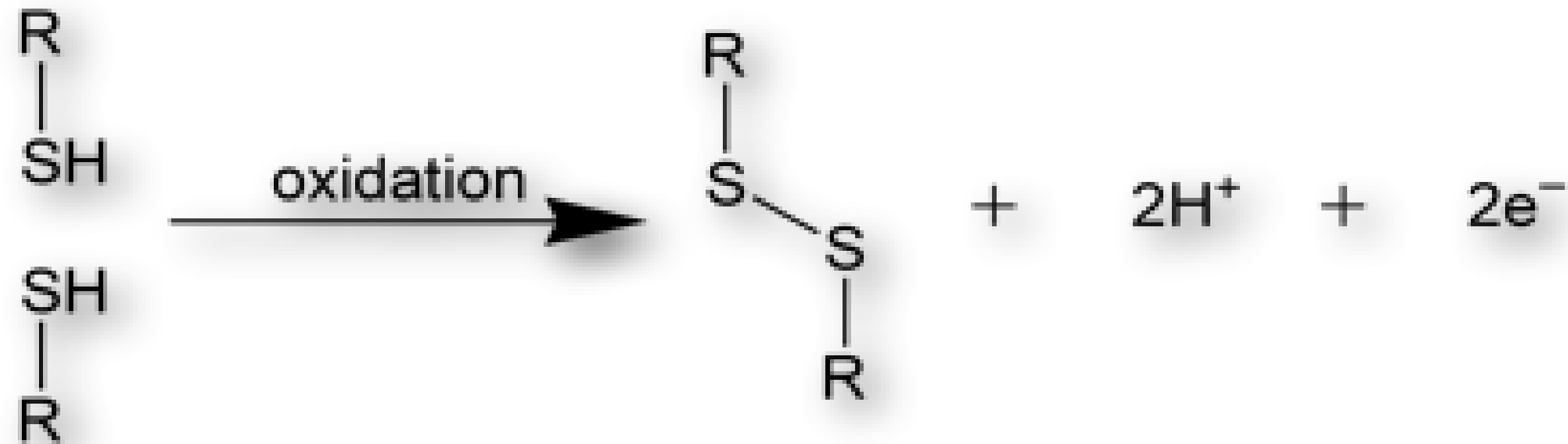


Salt bridges

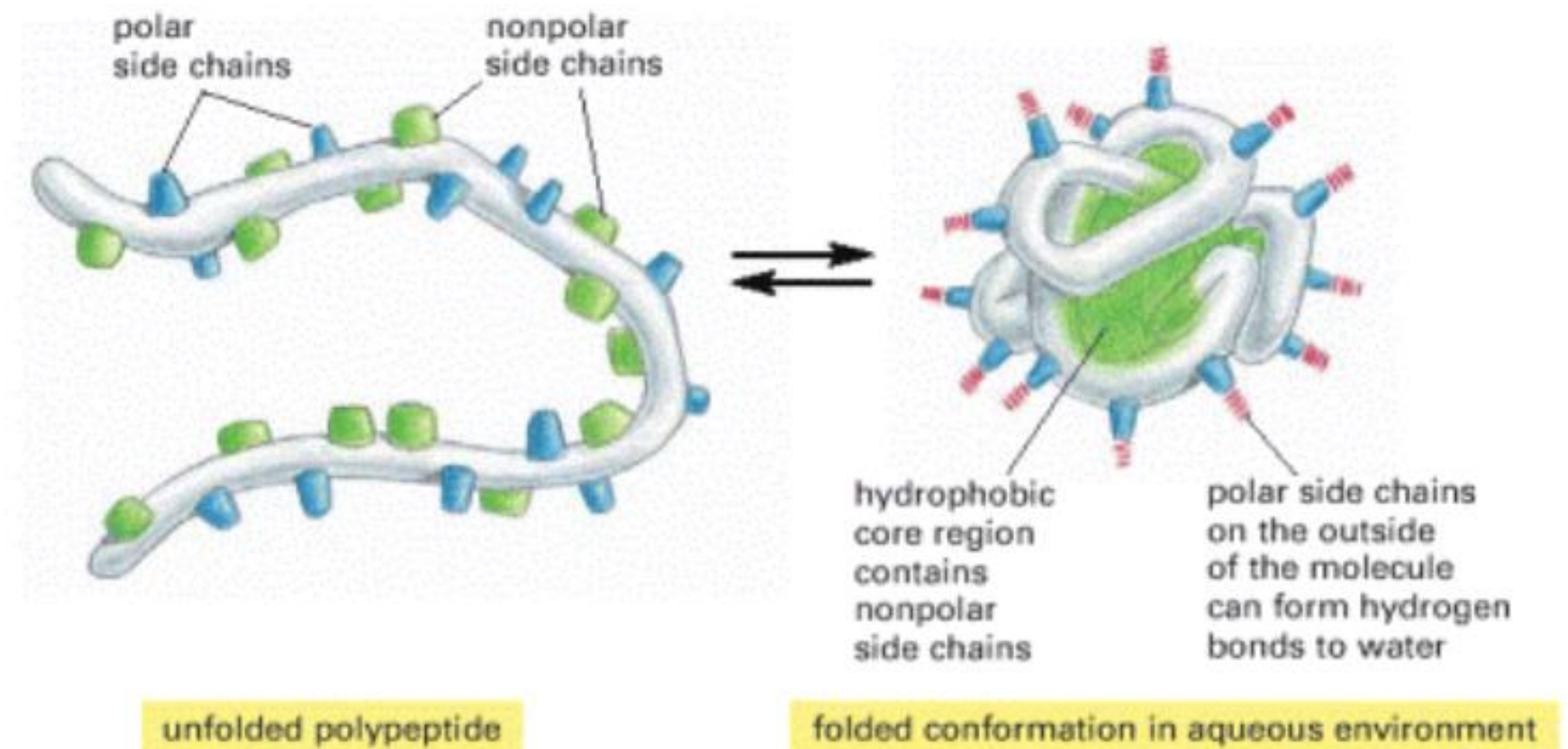


Forces Stabilizing Tertiary Structure

Disulphide bonds: the only covalent bond



Hydrophobic interaction
disruption of the solvent H-bond network



Denaturing/Unfolding Proteins

Break forces stabilizing them

Mercaptoethanol breaks disulfide bonds

Detergents disrupt hydrophobic interactions

Heat breaks hydrogen bonds

pH change alters ionic interactions

Chelators bind metal ions

Methods to detect denaturation

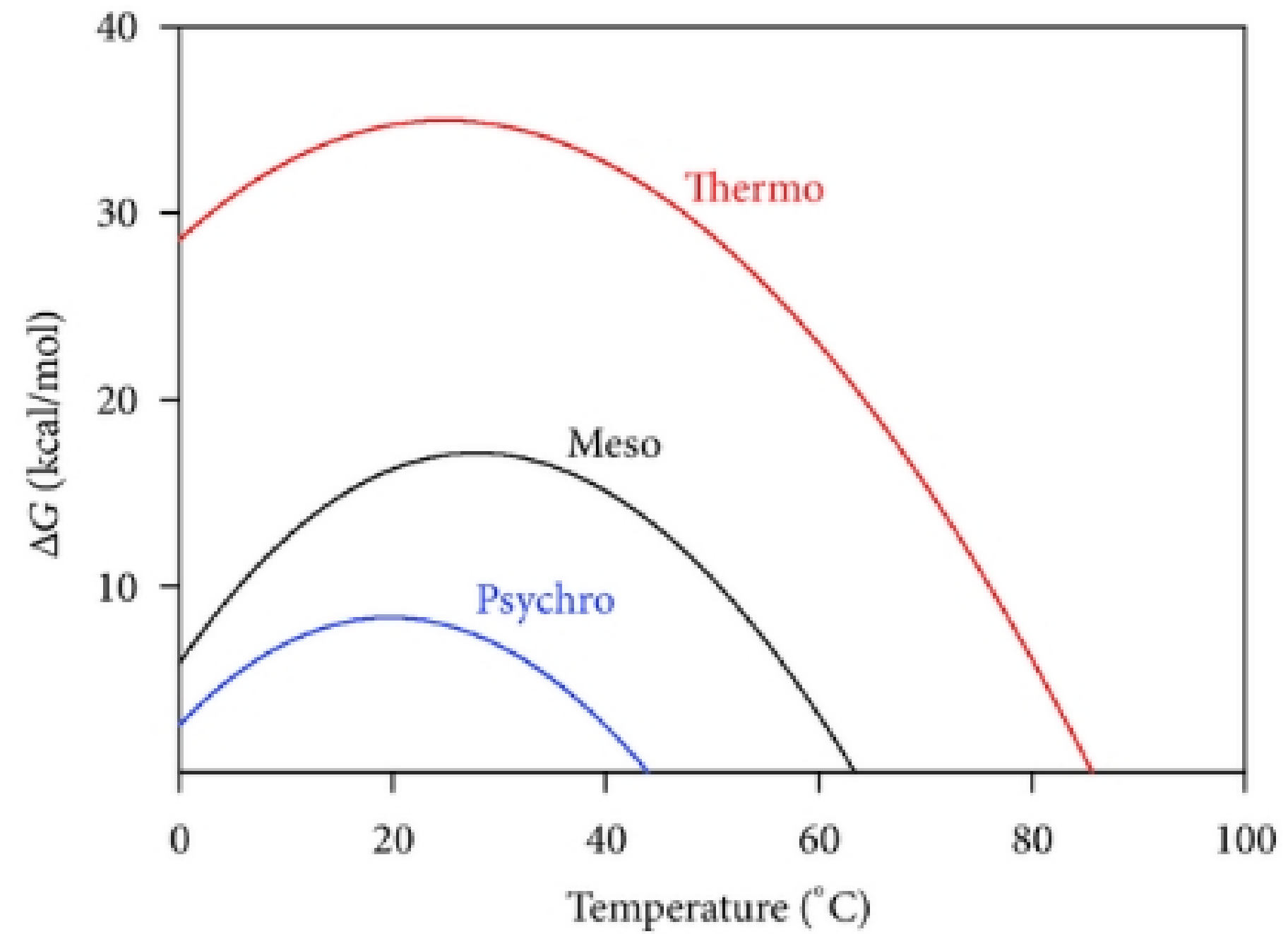
Fluorescence, UV, CD spectroscopy

DSC: Differential Scanning Calorimetry

Protein stability

marginal stability, tuned flexibility / stability

thermal adaptation

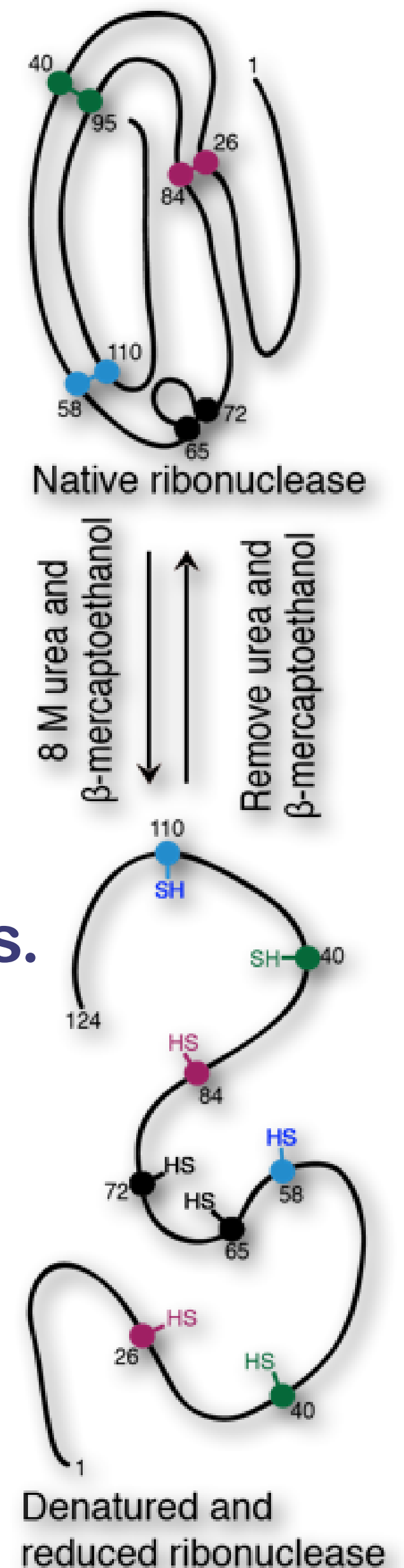


Anfinsen's dogma

For a small globular protein in its standard physiological environment, the native structure is determined only by the protein's amino acid sequence.

At the environmental conditions (temperature, solvent concentration and composition, ...) at which folding occurs, the native structure is a unique, stable and kinetically accessible minimum of the free energy.

Proteins requiring chaperones for folding, prions and amyloids are exceptions.



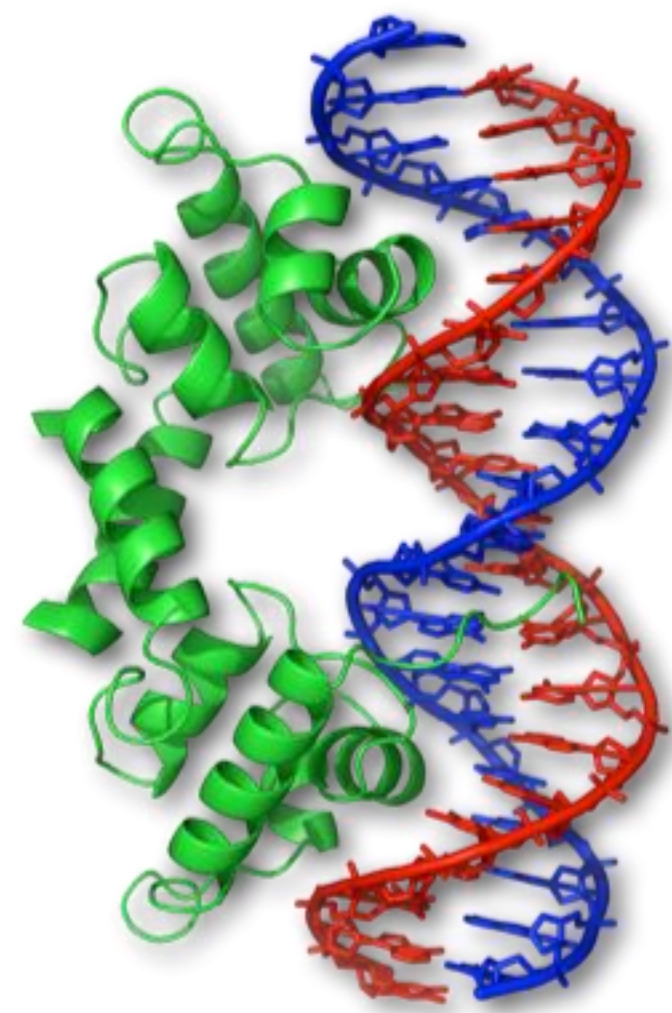
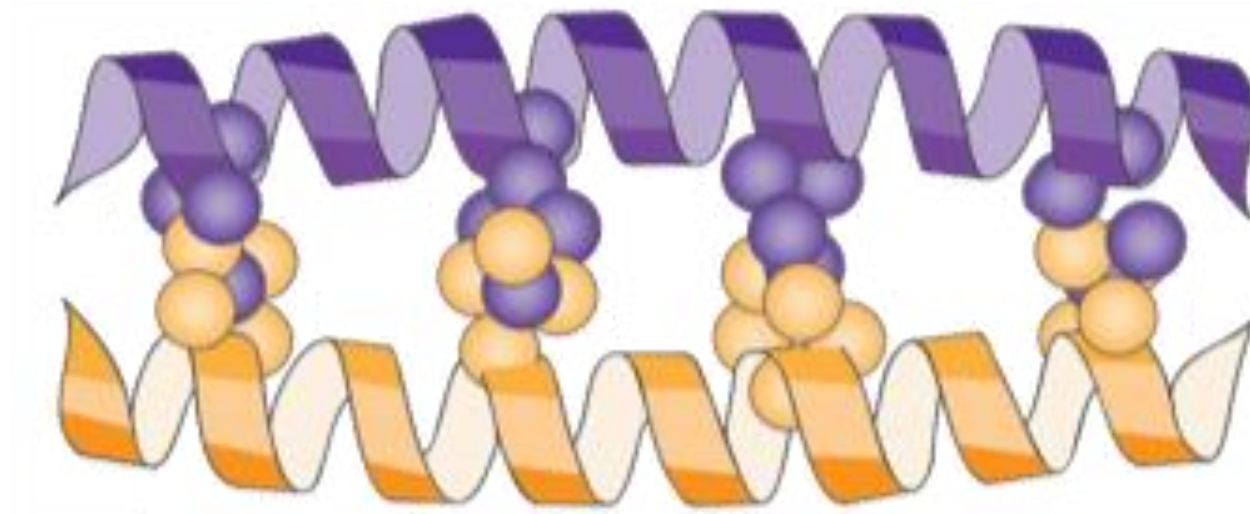
Protein Structural Domains

Leucine Zipper - Prot.-Prot. and Prot.-DNA

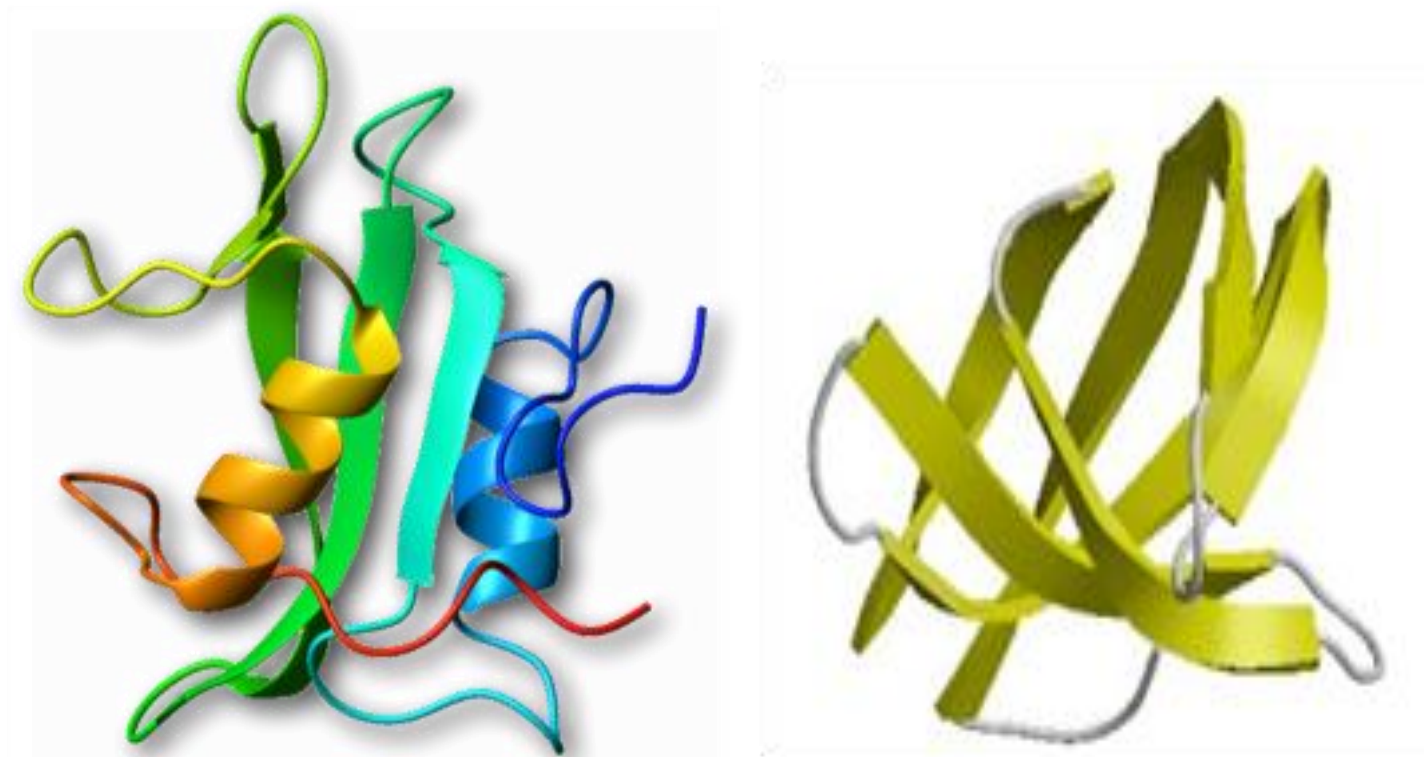
Helix Turn Helix - Protein-DNA

Zinc Fingers

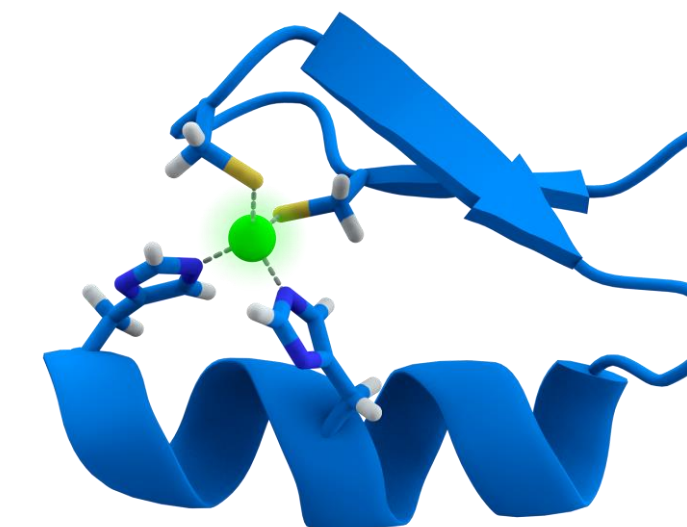
SH Domains - Protein-Protein



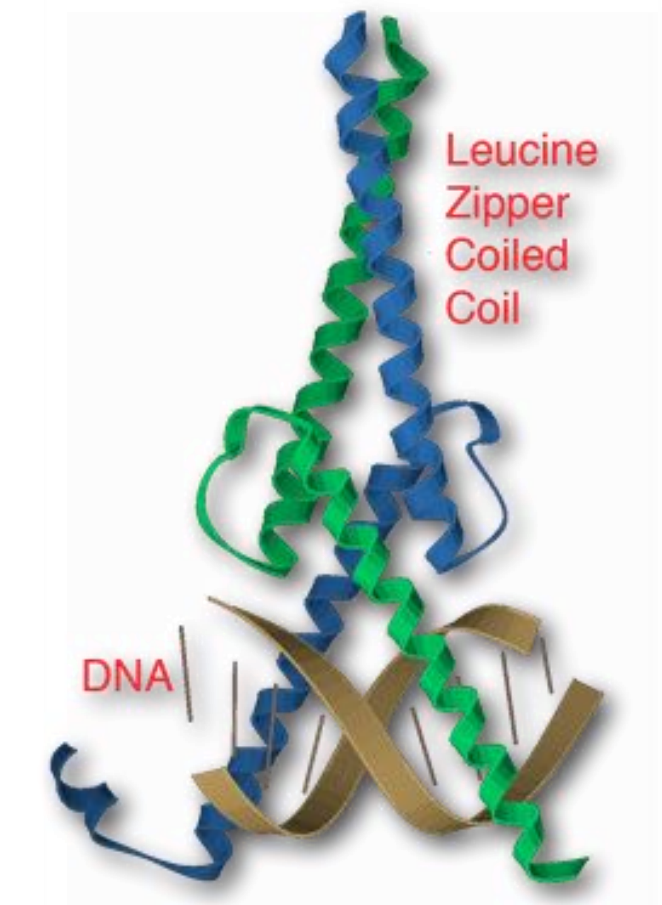
Helix-Turn-Helix



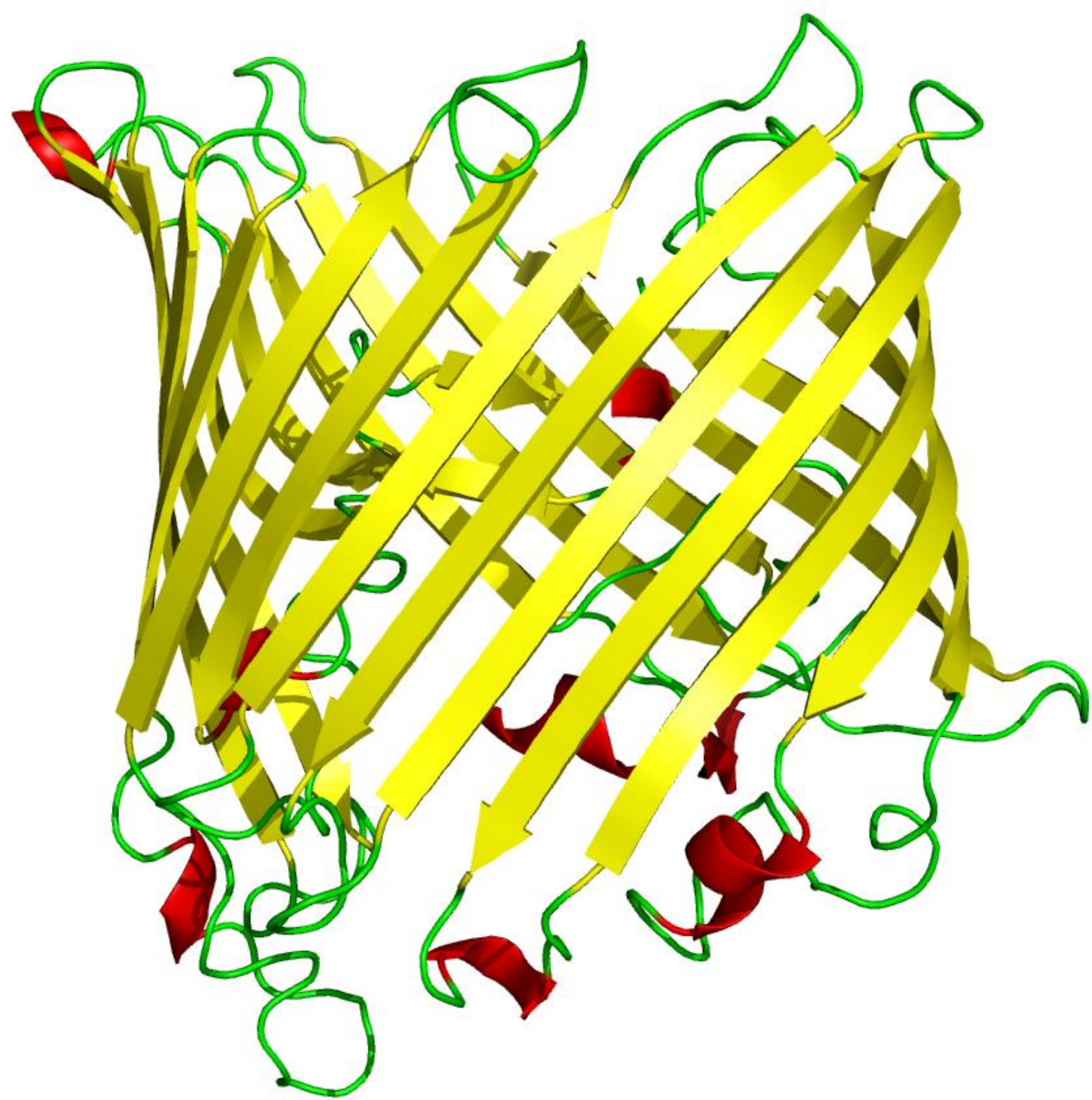
SH2 and SH3 domains



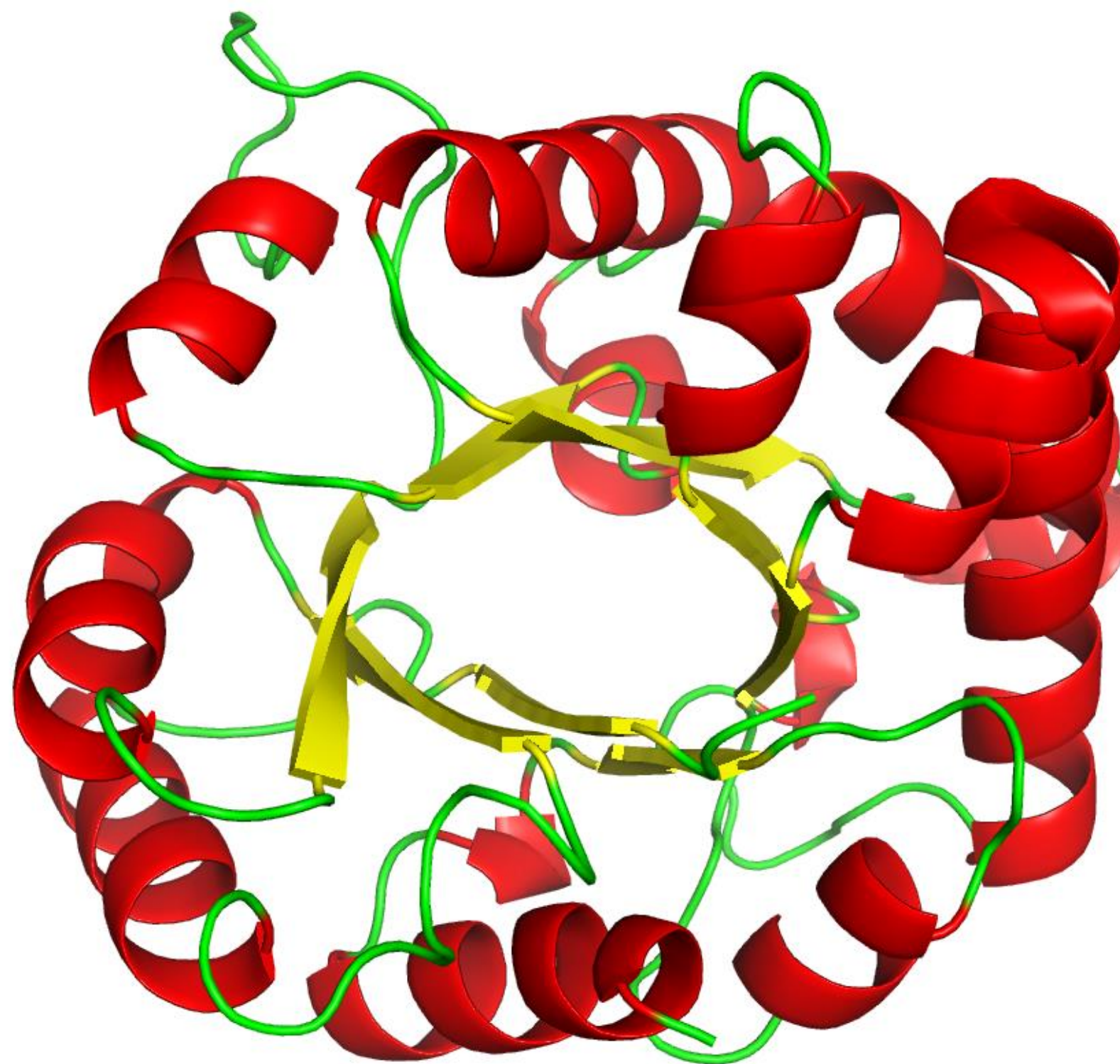
Zinc finger



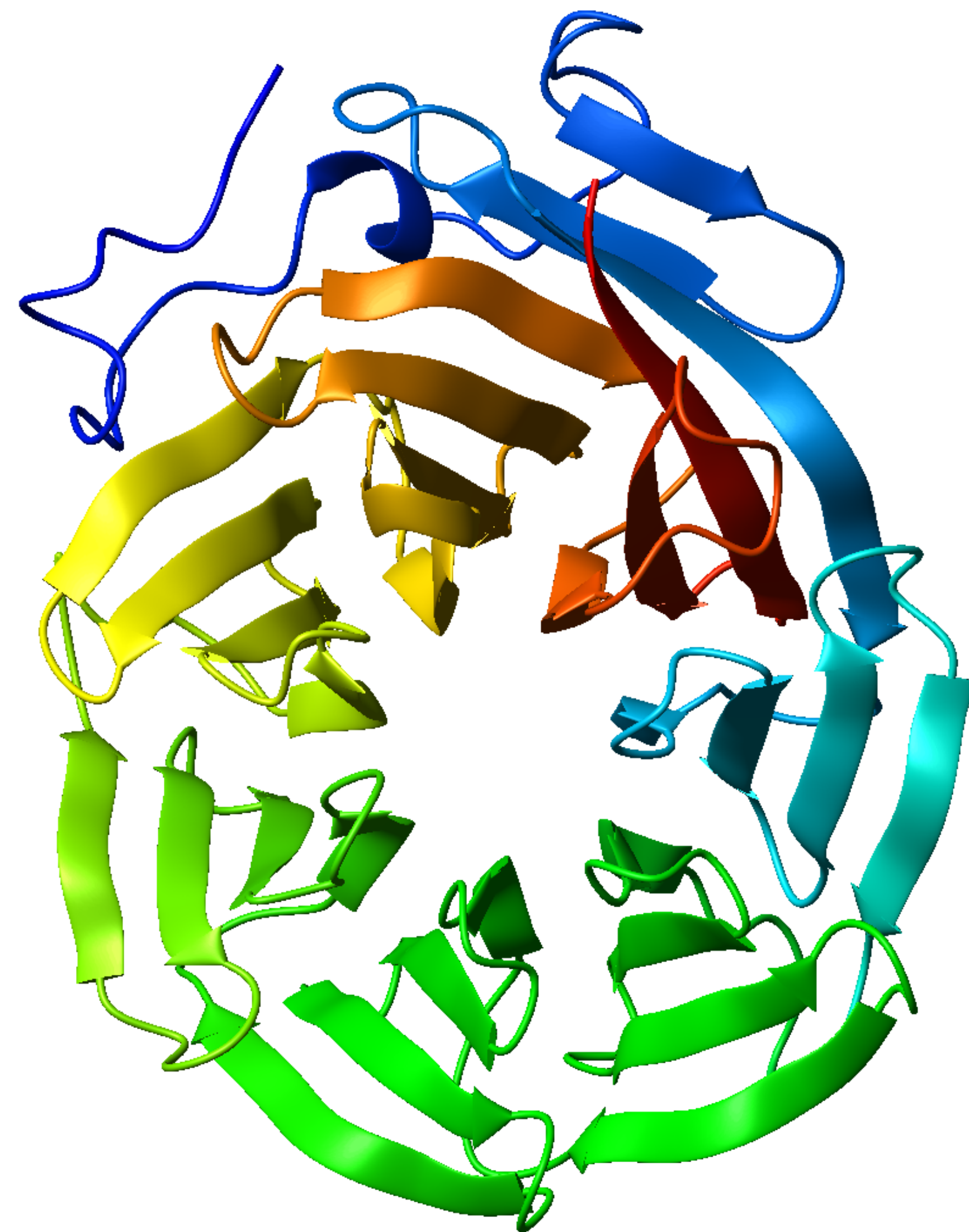
Leucine zipper



Beta barrel



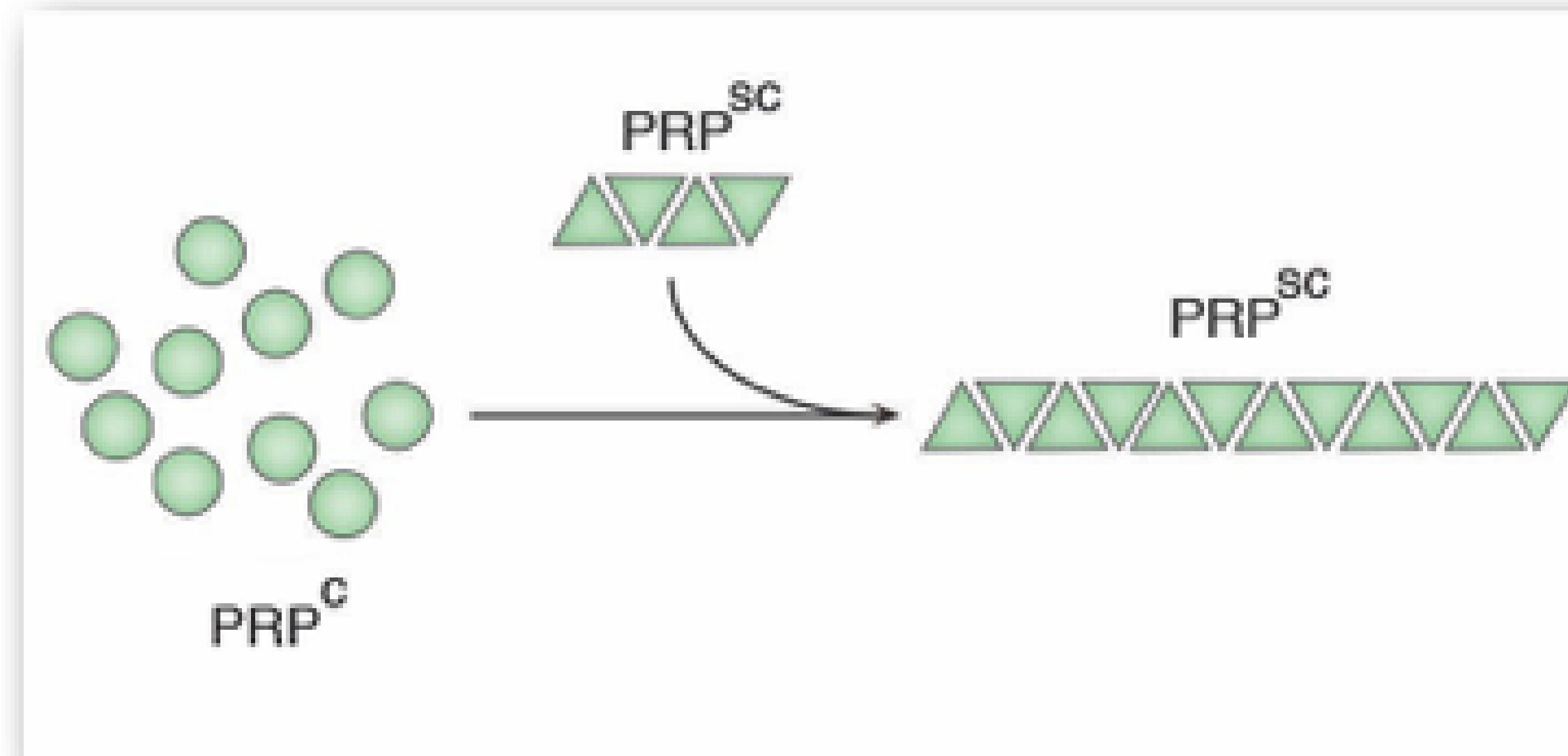
TIM barrel



Beta propeller

Folding Errors

prion protein



Amyloids and Disease

Amyloids - a collection of improperly folded **protein** aggregates found in the human body.

When misfolded, they are insoluble and contribute to some twenty human diseases including important neurological ones involving prions.

Amyloid diseases include (affected protein in parentheses) -

Alzheimer's disease (**Amyloid β**)

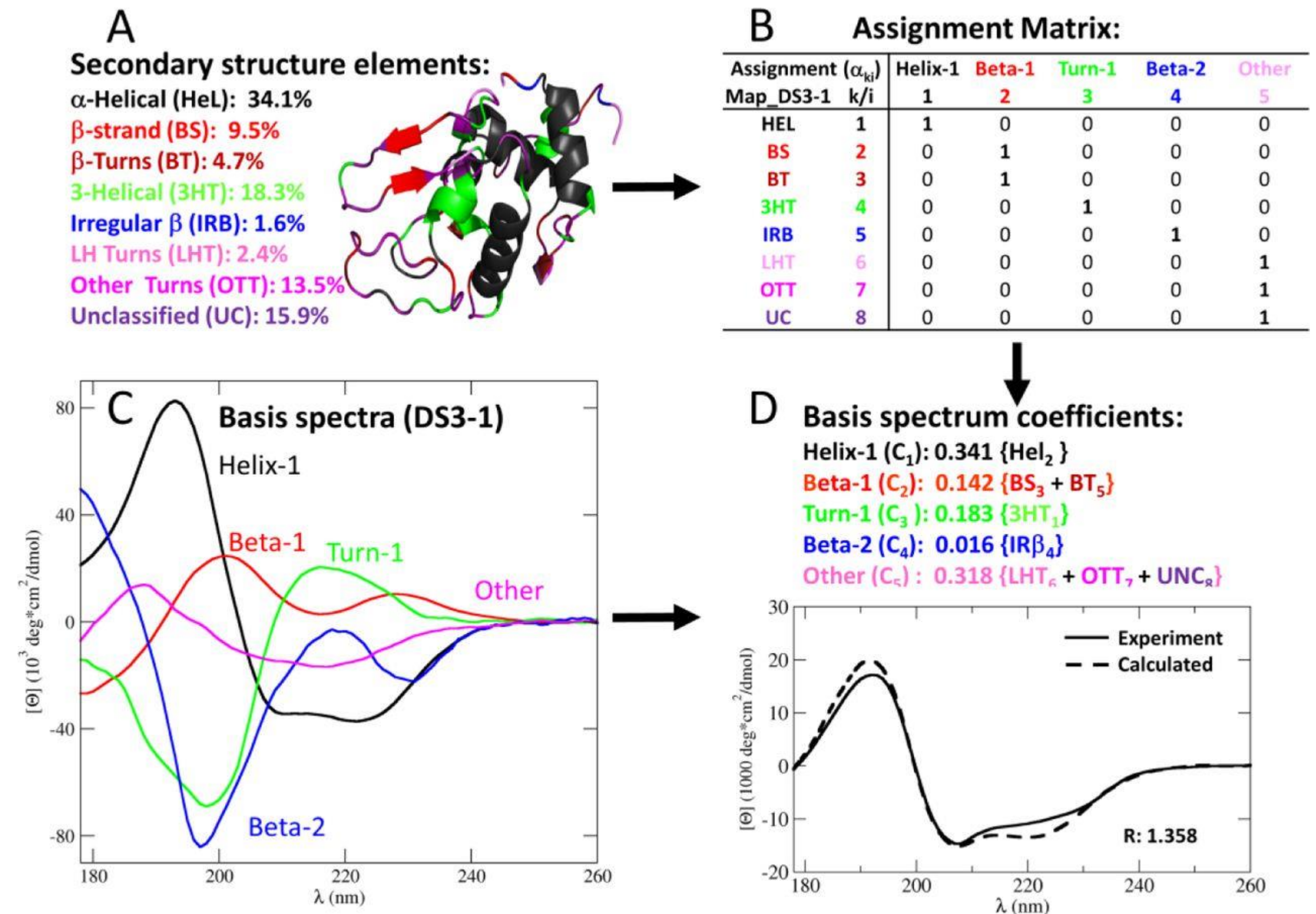
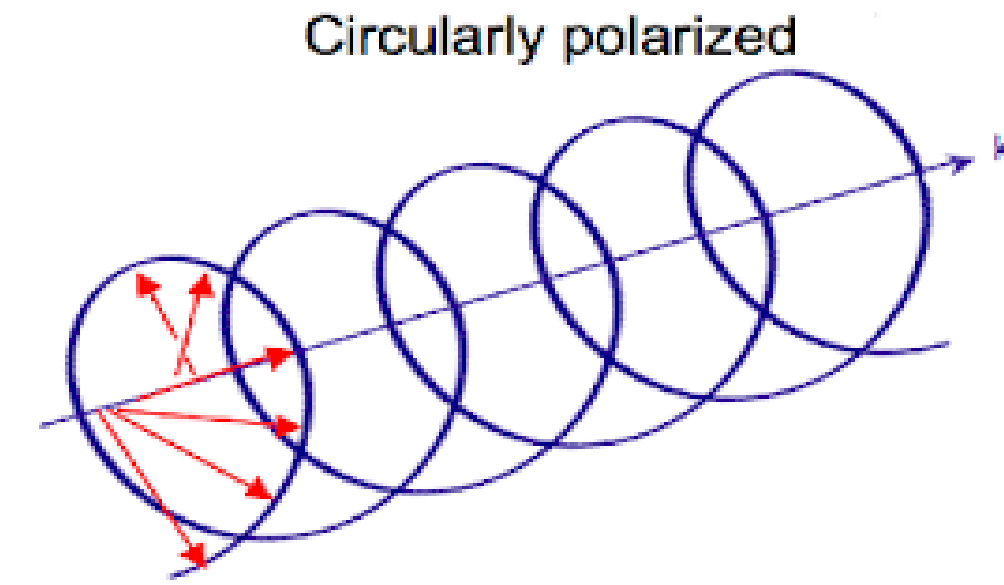
Parkinson's disease (**α -synuclein**)

Huntington's disease (**huntingtin**),

Rheumatoid arthritis (**serum amyloid A**),

Structure determination methods

circular dichroism spectroscopy
far UV near UV

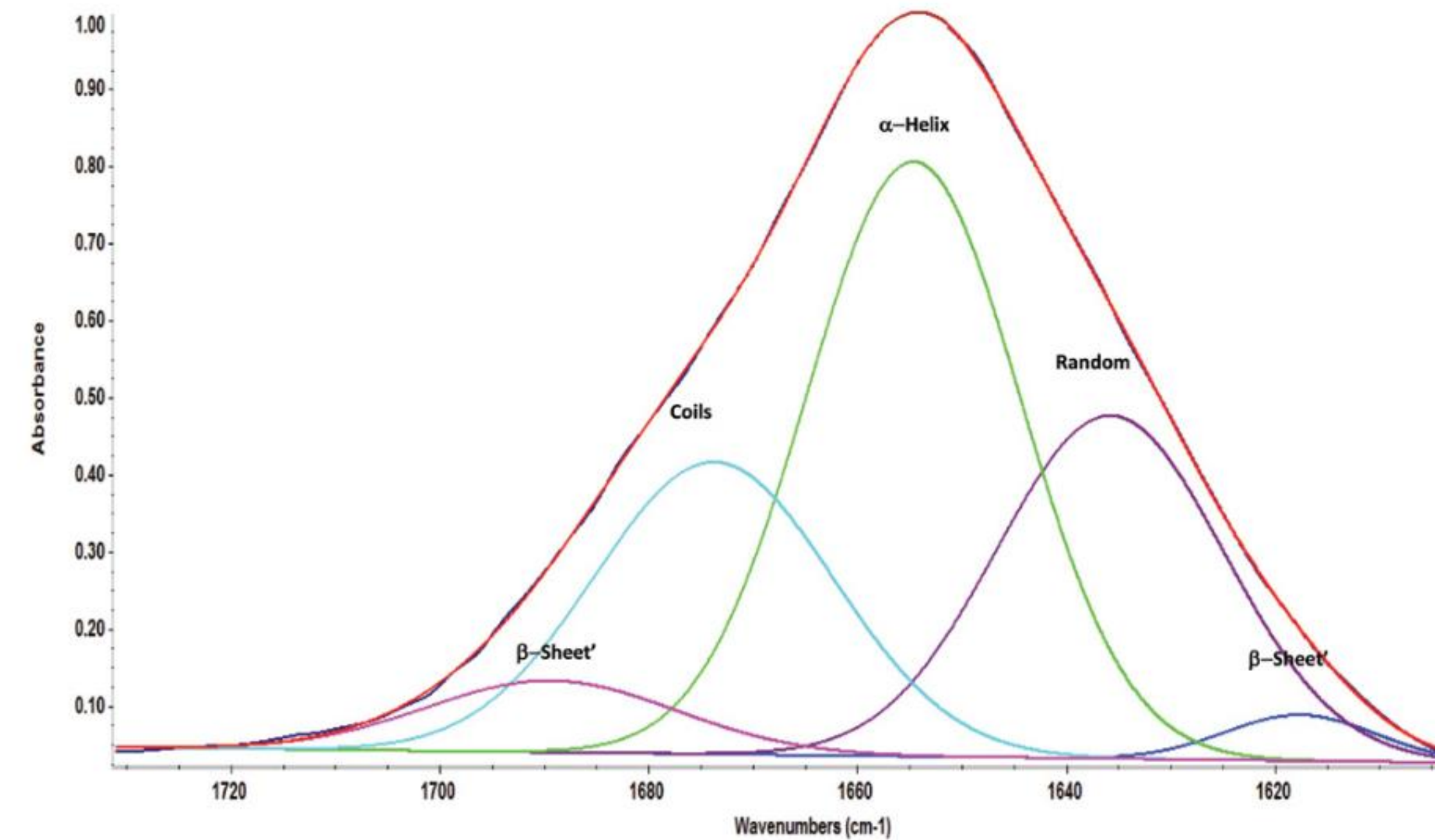
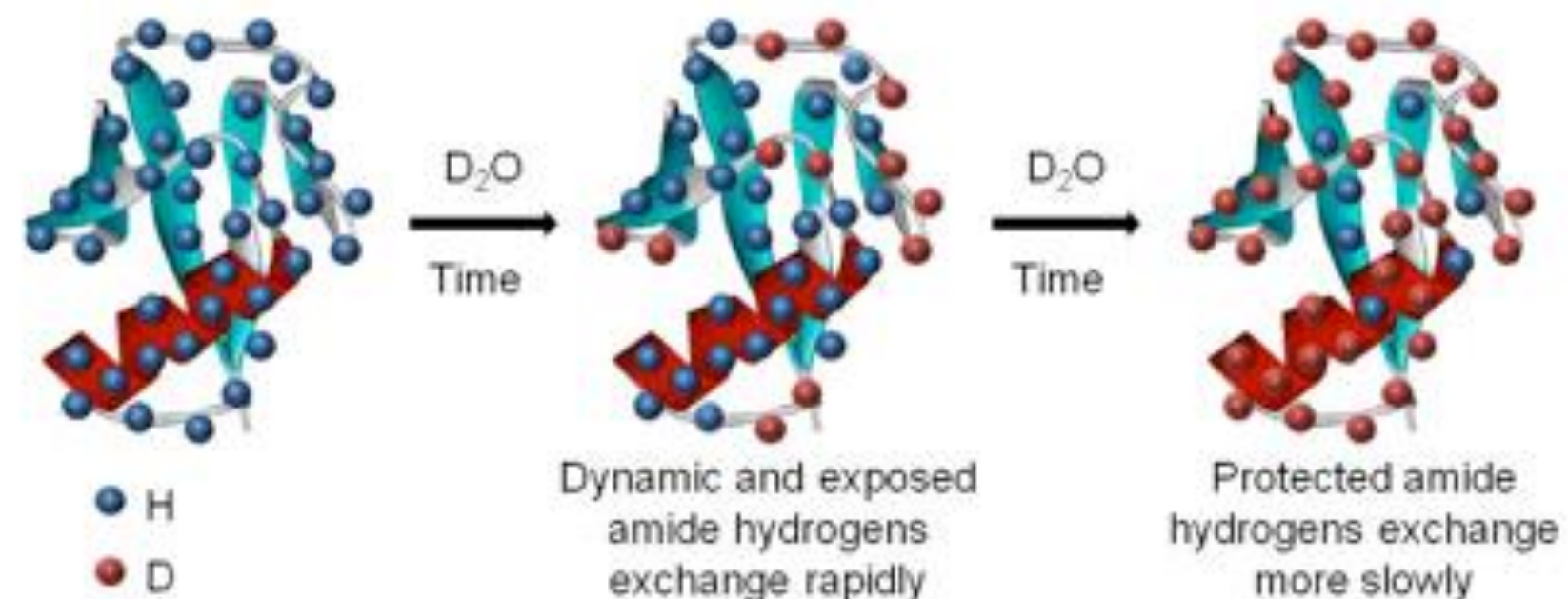


FTIR

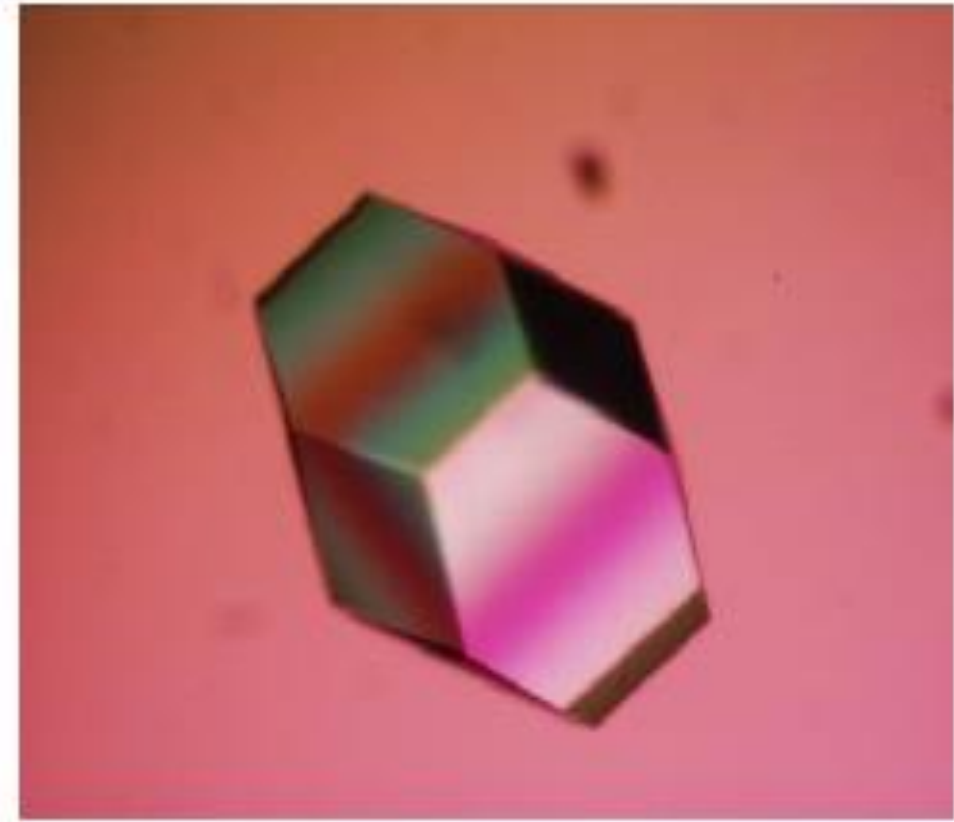
Infrared radiation absorption – molecular vibrations
amide bands are influenced by secondary structure
deconvolution of the spectra

hydrogen-deuterium exchange

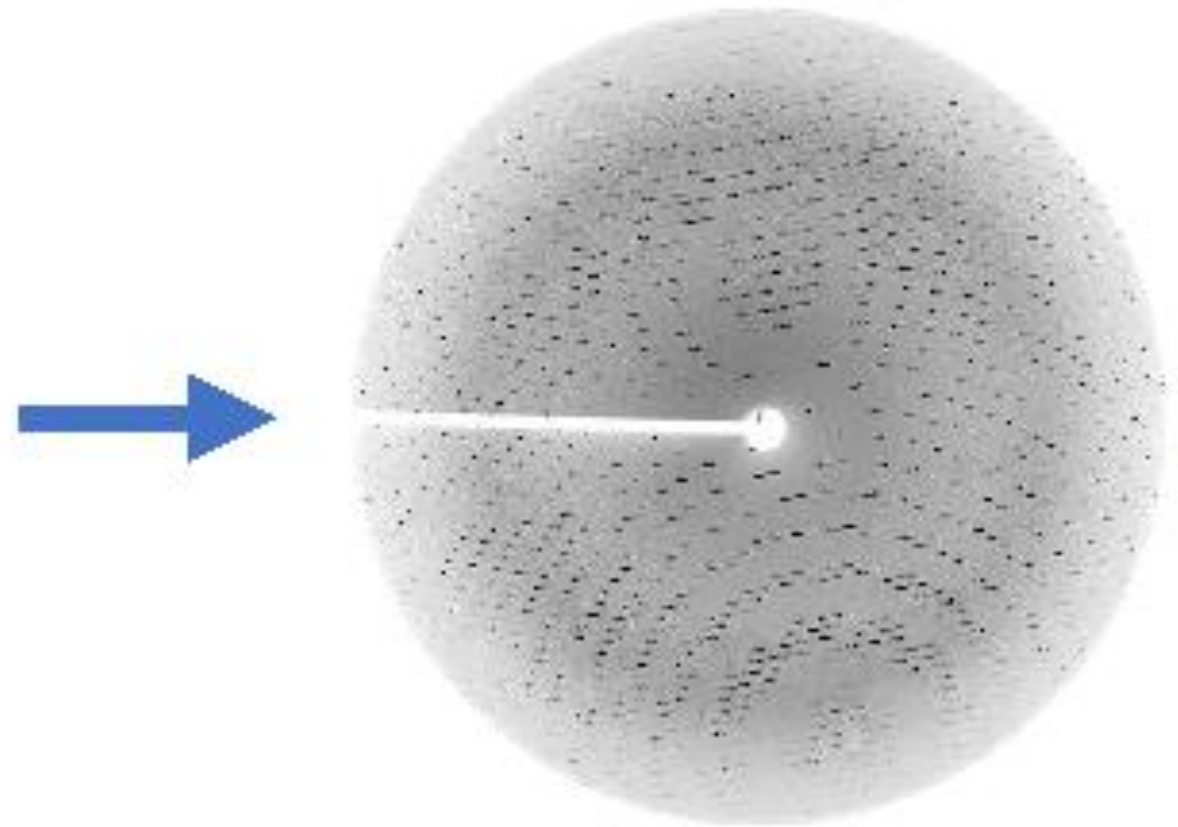
H/D Exchange



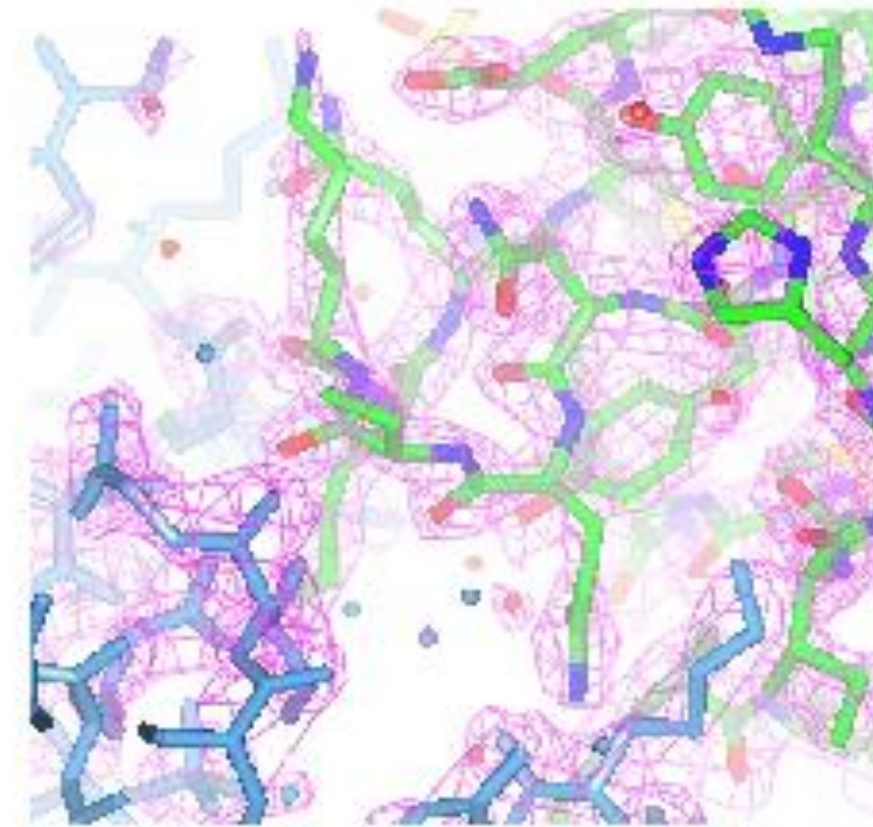
X-ray crystallography



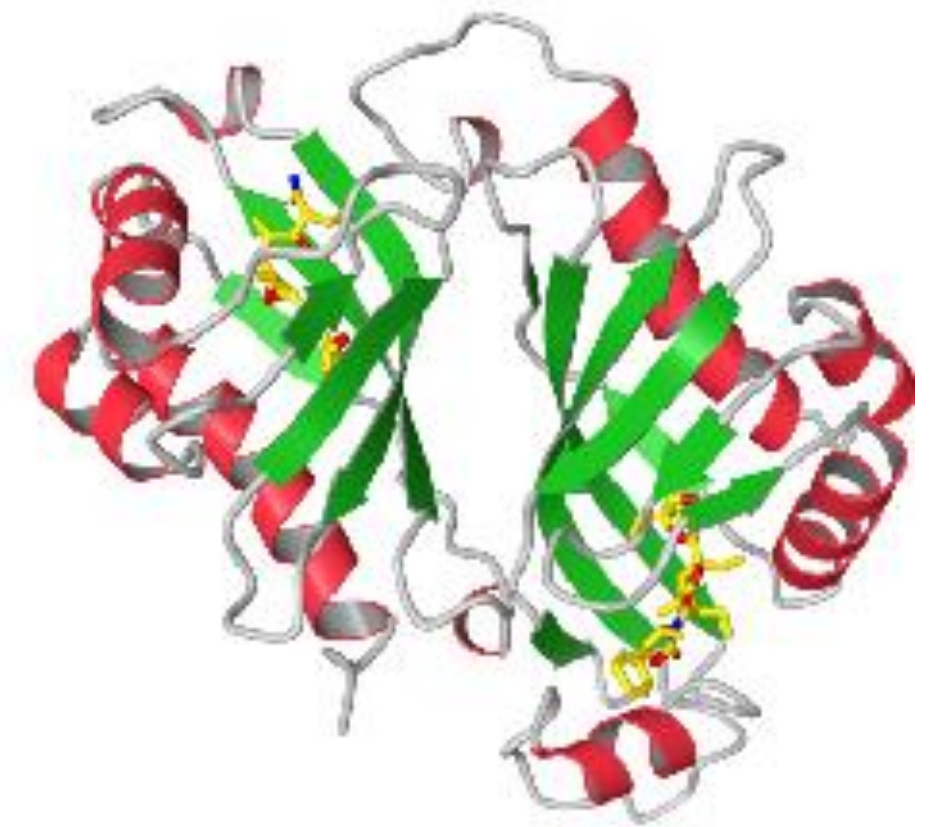
Crystal



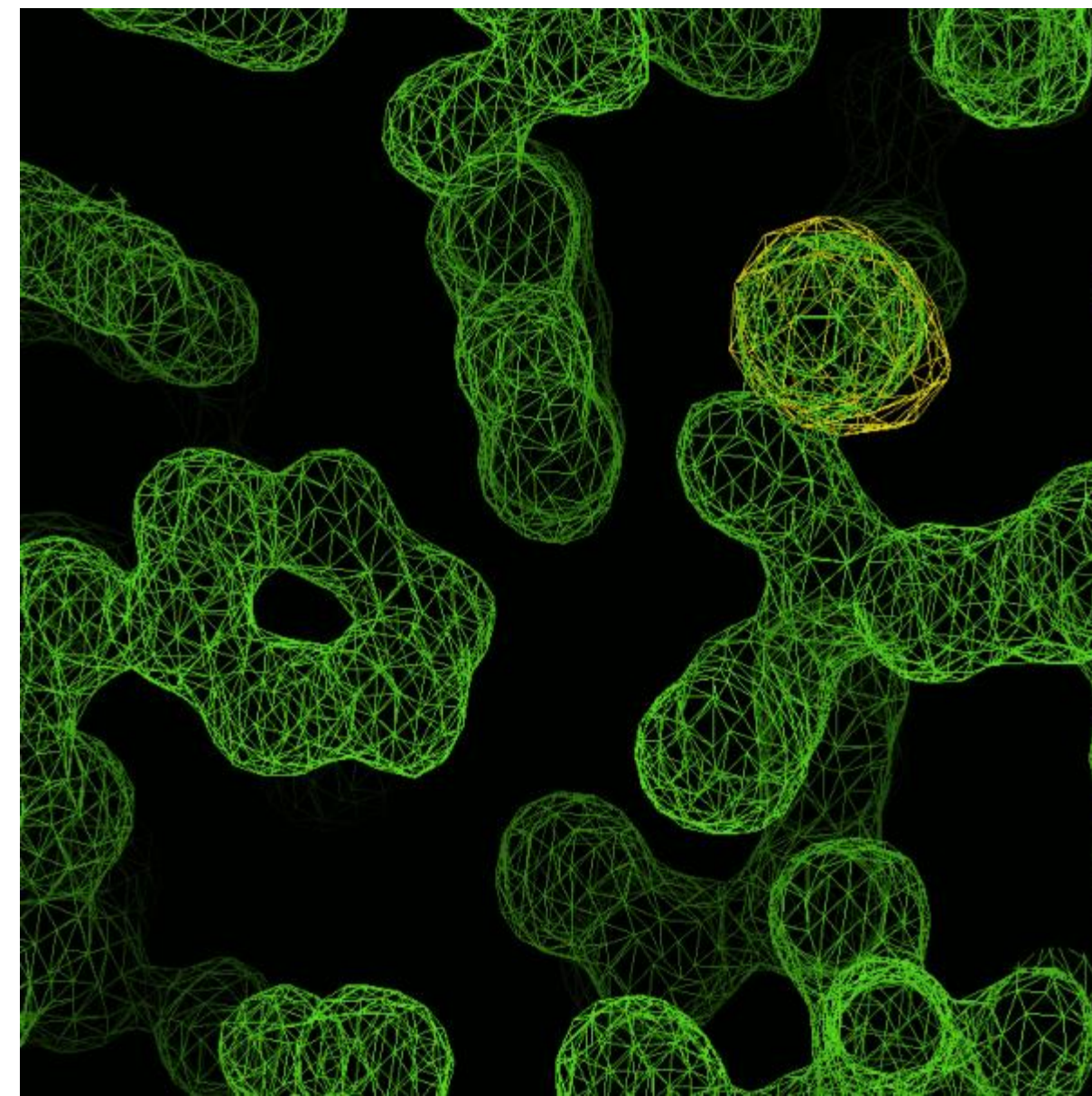
Diffraction pattern



Electron density map

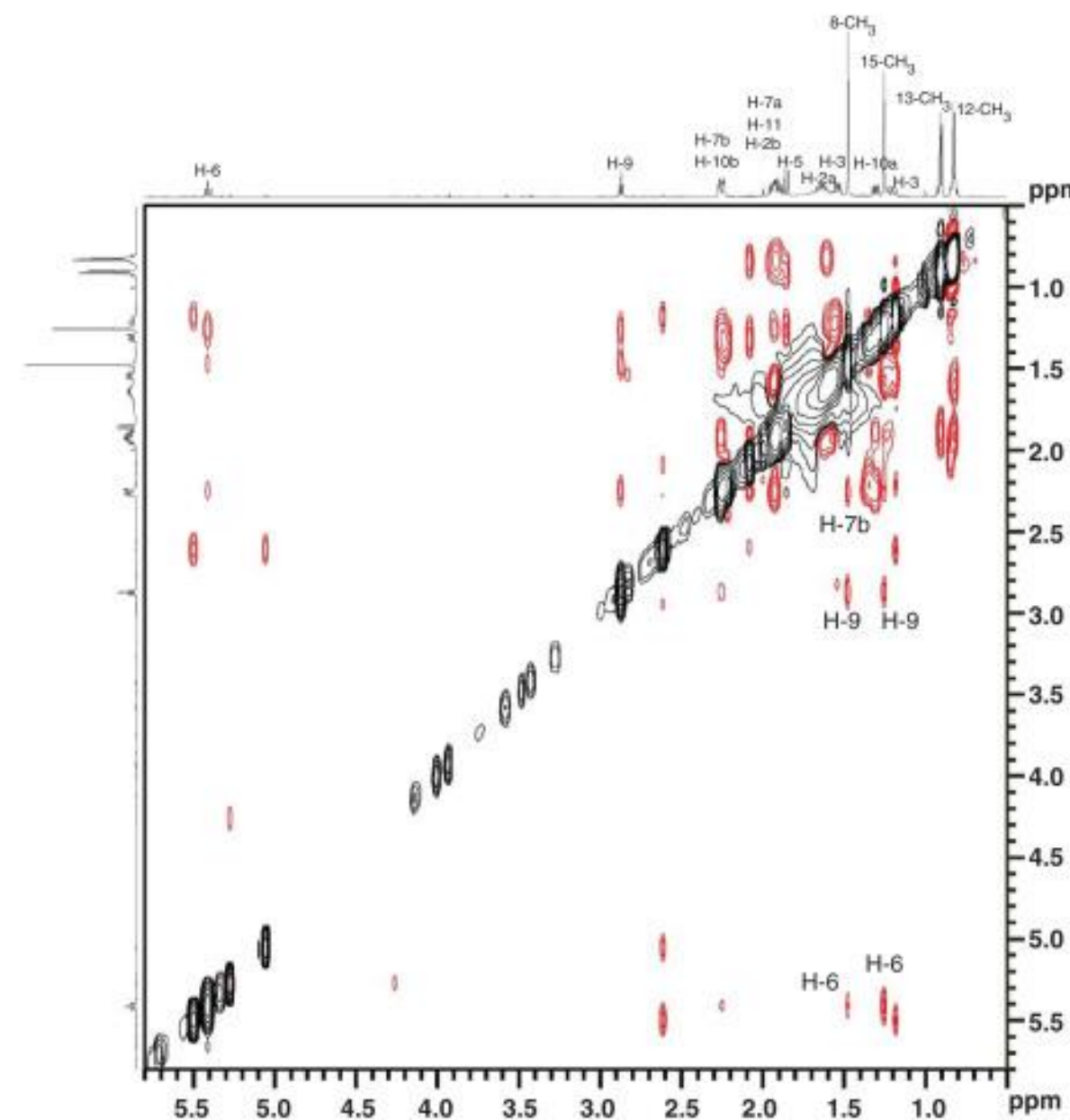
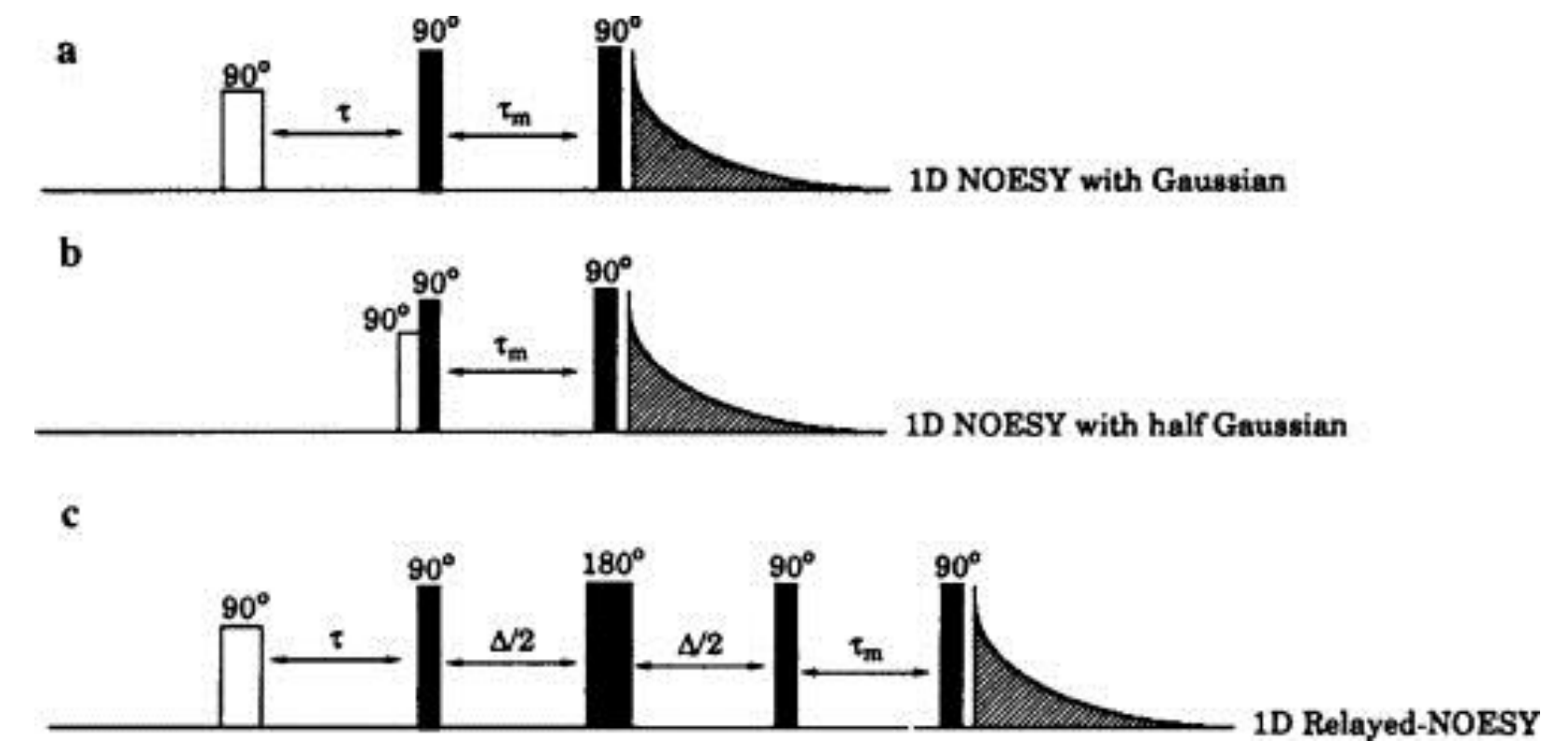
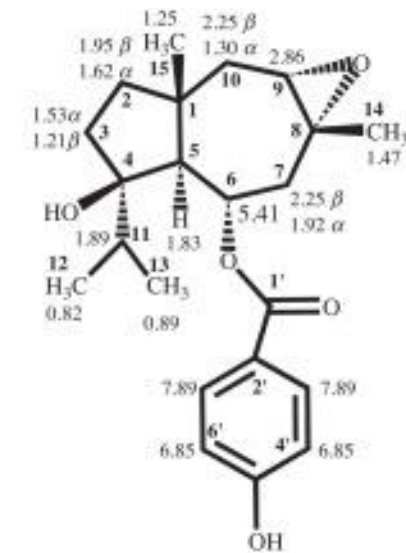
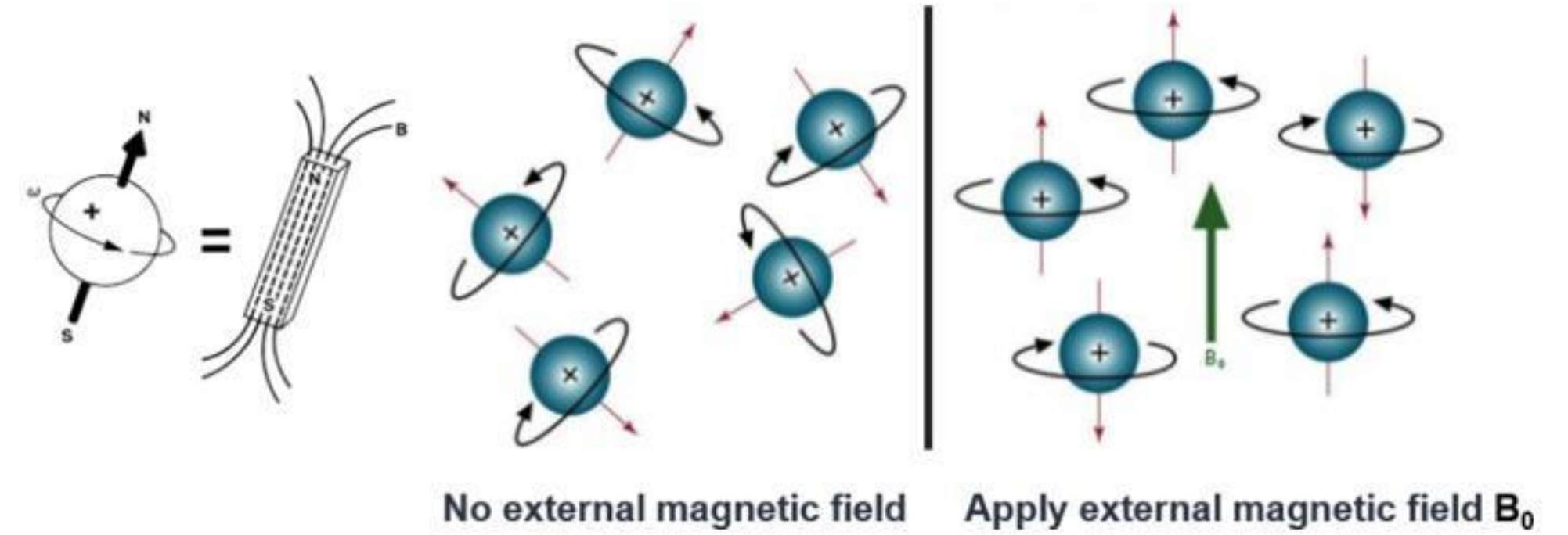


Protein model



NMR

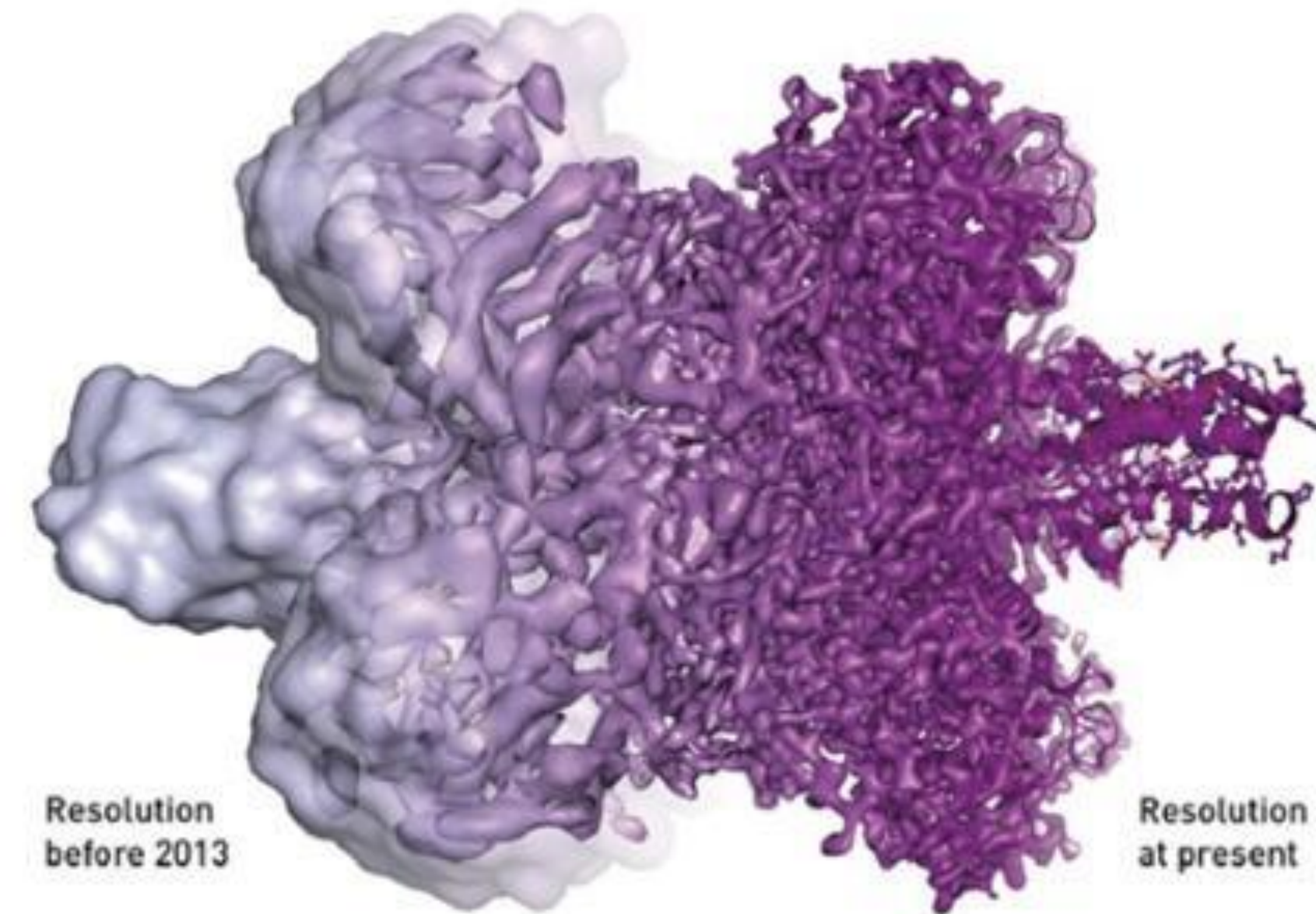
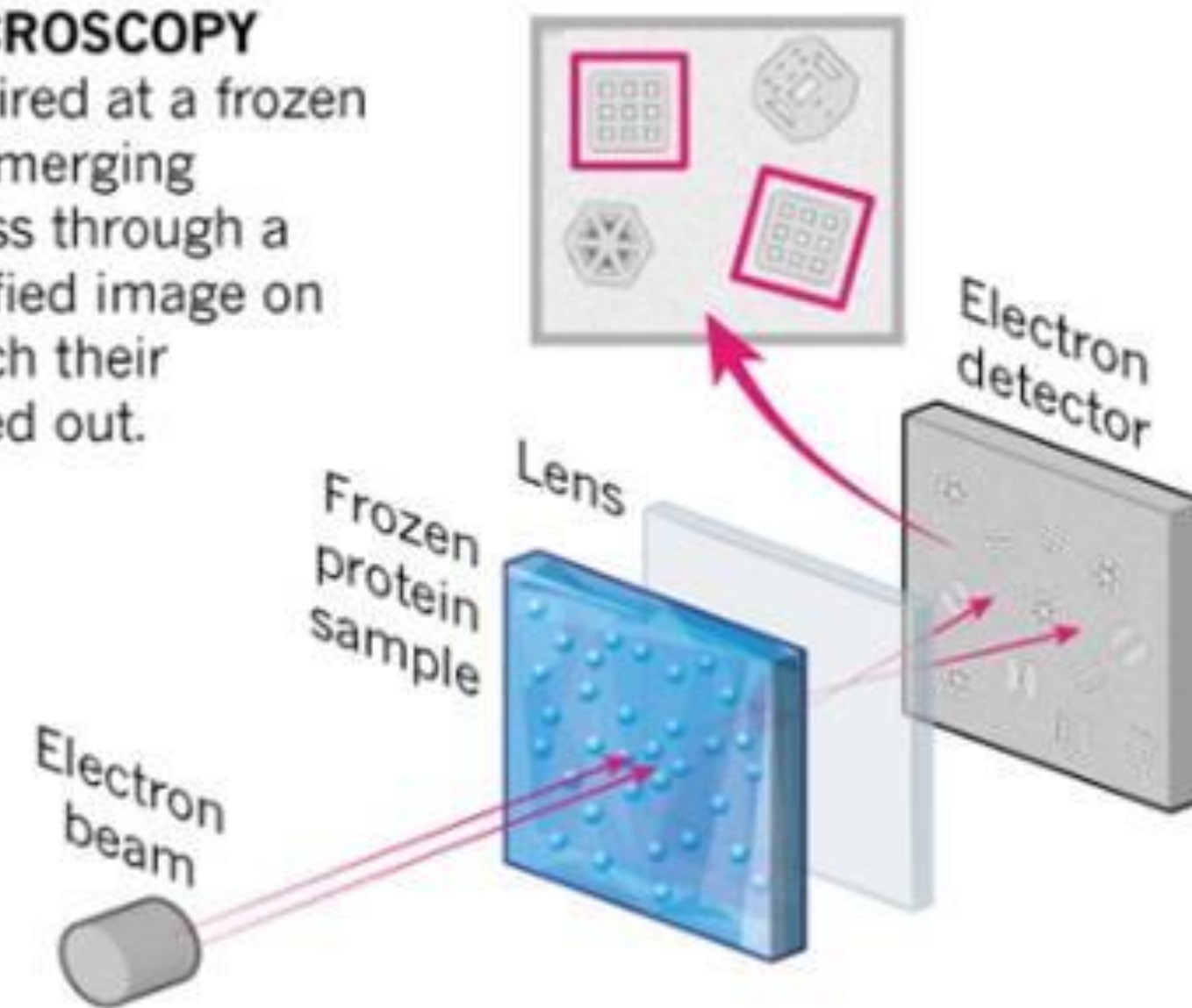
chemical shift depends on the environment
pulse sequences – selection
NMR spectra
distance constraints
model building



Cryo-EM

CRYO-ELECTRON MICROSCOPY

A beam of electron is fired at a frozen protein solution. The emerging scattered electrons pass through a lens to create a magnified image on the detector, from which their structure can be worked out.



Comparison of methods

	Advantages	Disadvantages	Objects	Resolution
X-ray Crystallography	<ul style="list-style-type: none"> • Well developed • High resolution • Broad molecular weight range • Easy for model building 	<ul style="list-style-type: none"> • Difficult for crystallization • Difficult for diffraction • Solid structure preferred • Static crystalline state structure 	<ul style="list-style-type: none"> • Crystallizable samples • Soluble proteins, membrane proteins, ribosomes, DNA/RNA and protein complexes 	High
NMR	<ul style="list-style-type: none"> • High resolution • 3D structure in solution • Good for dynamic study 	<ul style="list-style-type: none"> • Need for high sample purity • Difficult for sample preparation • Difficult for computational simulation 	<ul style="list-style-type: none"> • MWs below 40–50 kDa • Water soluble samples 	High
Cryo-EM	<ul style="list-style-type: none"> • Easy sample preparation • Structure in native state • Small sample size 	<ul style="list-style-type: none"> • Relatively low resolution • Applicable to samples of high molecular weights only • Highly dependent on EM techniques • Costly EM equipment 	<ul style="list-style-type: none"> • >150 kDa • Virions, membrane proteins, large proteins, ribosomes, complex compounds 	Relatively Low (<3.5 Å)

- Welcome
- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

A Structural View of Biology

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

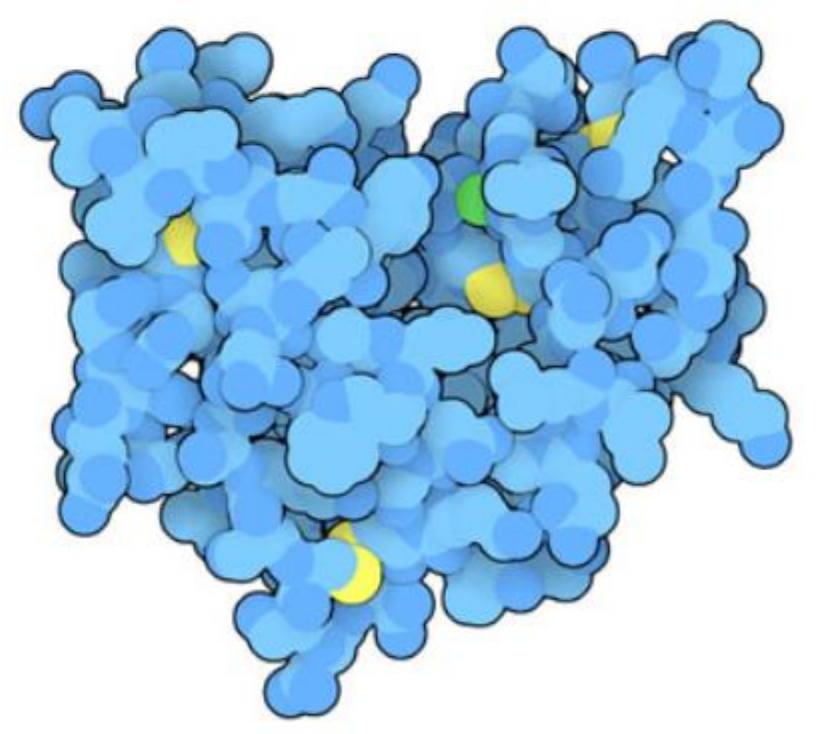
As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

Video: How Enzymes Work



November Molecule of the Month



Phospholipase A2



- Single crystal X-ray diffraction (SC-XRD)
- Nuclear magnetic resonance (NMR)
- Cryo-electron microscopy (Cryo-EM)

Thank you for your attention!