Molecular structure of proteins
Relationship between protein structure and function
Physical and chemical properties of proteins and their applications

Ph.D., D.Sc. Svetlana Demyanenko

2026

General lecture plan:

Characteristics of the bonds involved in the formation of protein structures

Covalent bonds

1. Structure and properties of the peptide bond

Φ (phi) and y (psi) angles, Ramachandran conformational plot.

2. Disulfide bond

Non-covalent bonds

1. Hydrogen bond
2. Salt bridges
3. Hydrophobic interaction

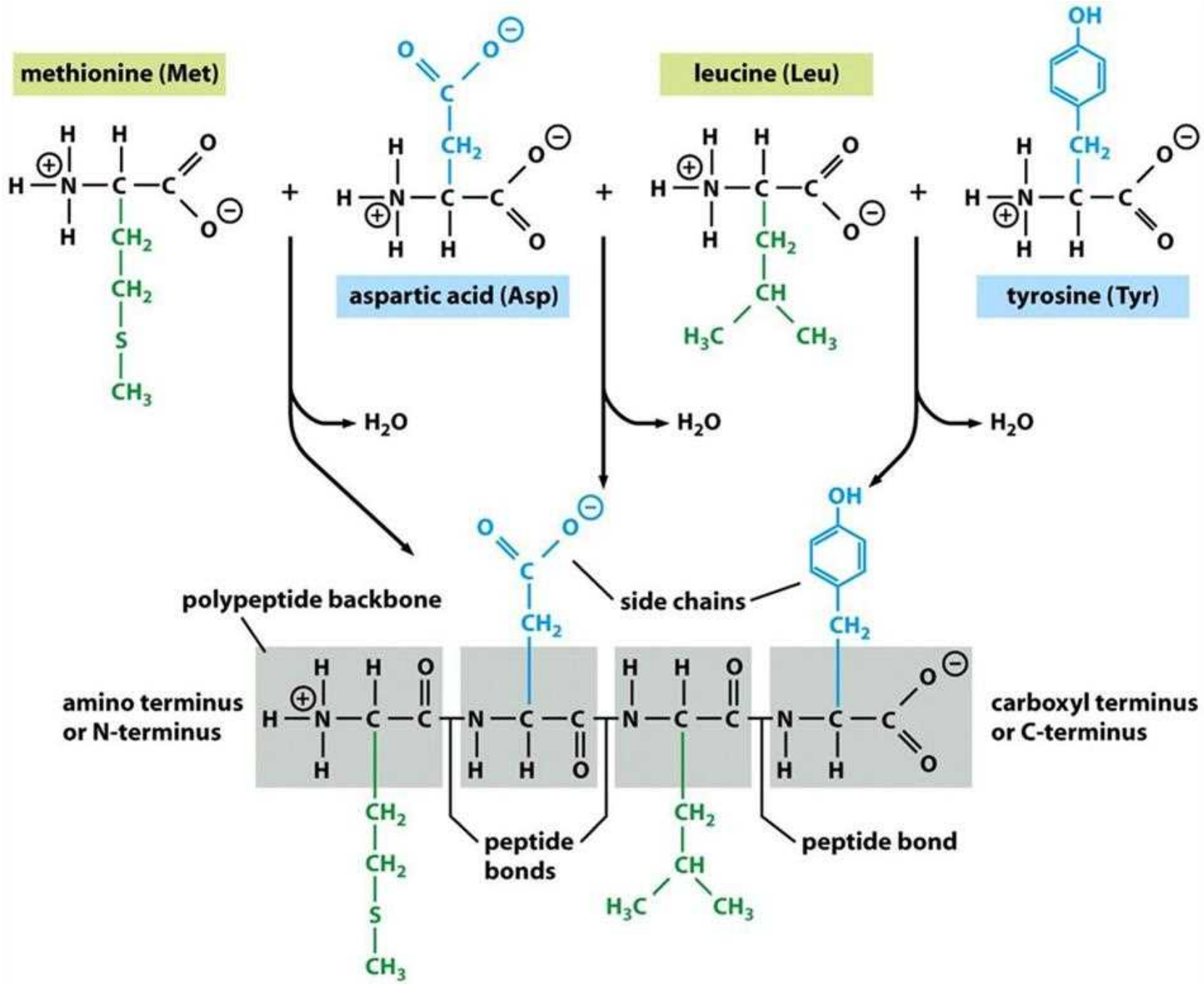
Spatial organization of proteins

1. a-helix. Secondary structure of collagen
2. β-sheet. Hydrogen bond models in β-sheet: antiparallel, parallel, mixed b-sheet. Jsmol (Jmol) program. Participation of b-sheet in the fibril's formation.
3. Tertiary structure of protein. Bonds. Domains.
4. Quaternary Structure of Protein. Fibrillary and globular proteins. Structure of collagen and globular proteins (myoglobin and hemoglobin).

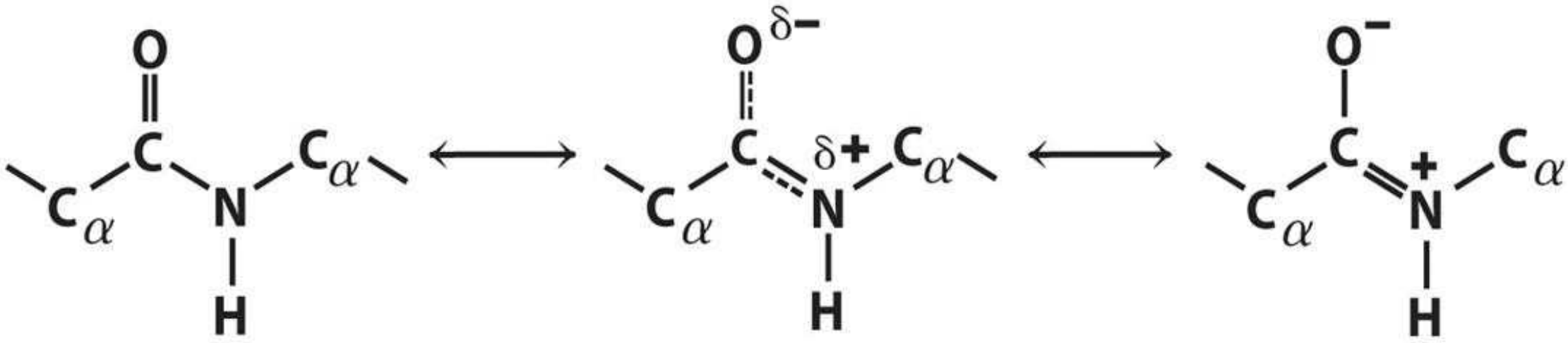
Physico-chemical properties of proteins

Hydration and solubility of proteins

Protein denaturation


methionine (Met)
aspartic acid (Asp)
leucine (Leu)
tyrosine (Tyr)
H_2O
H_2O
H_2O
polypeptide backbone
side chains
amino terminus
or N-terminus
carboxyl terminus
or C-terminus
peptide
bonds
peptide bond

The amide bond that links two amino acids is called a peptide bond

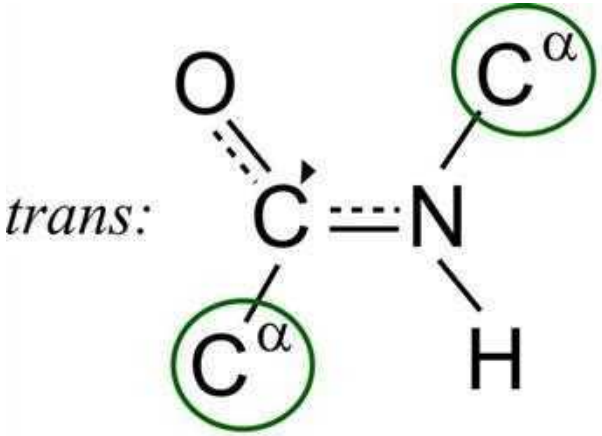


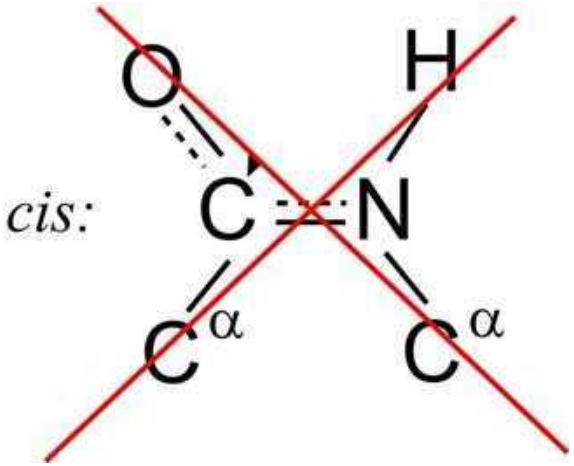
The carbonyl oxygen has a partial negative charge and the amide nitrogen has a partial positive charge, setting up a small electric dipole.

- The C-N bond is partially multiple due to the interaction of the lone pair of electrons of the N atom with the π-electrons of the C=O carbonyl group (p-π conjugation).
- This leads to hindered free rotation around the C-N bond (rotational barrier 63-84 kJ/mol)

The peptide bond in proteins is in trans form!!!

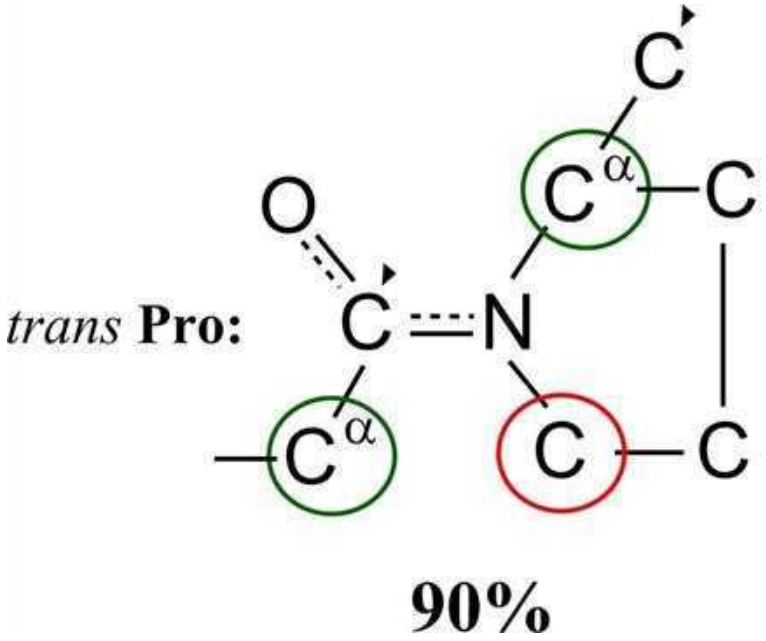
All except proline:



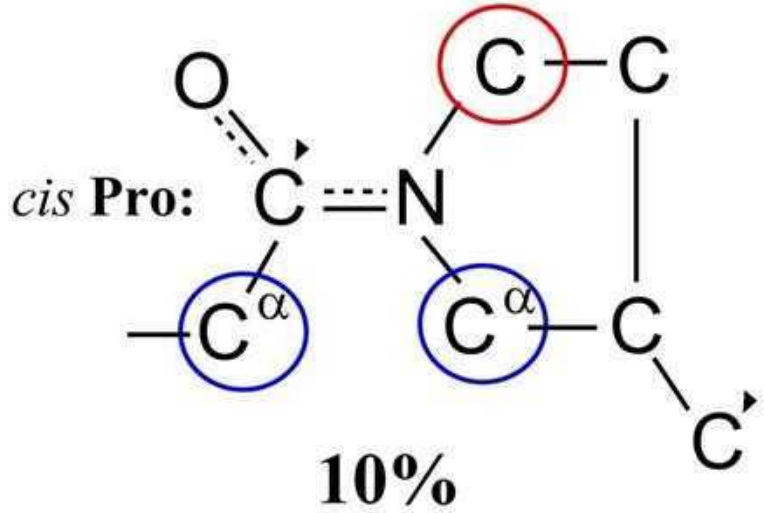


Exceptions: the peptide bond can be in a planar cis-form:

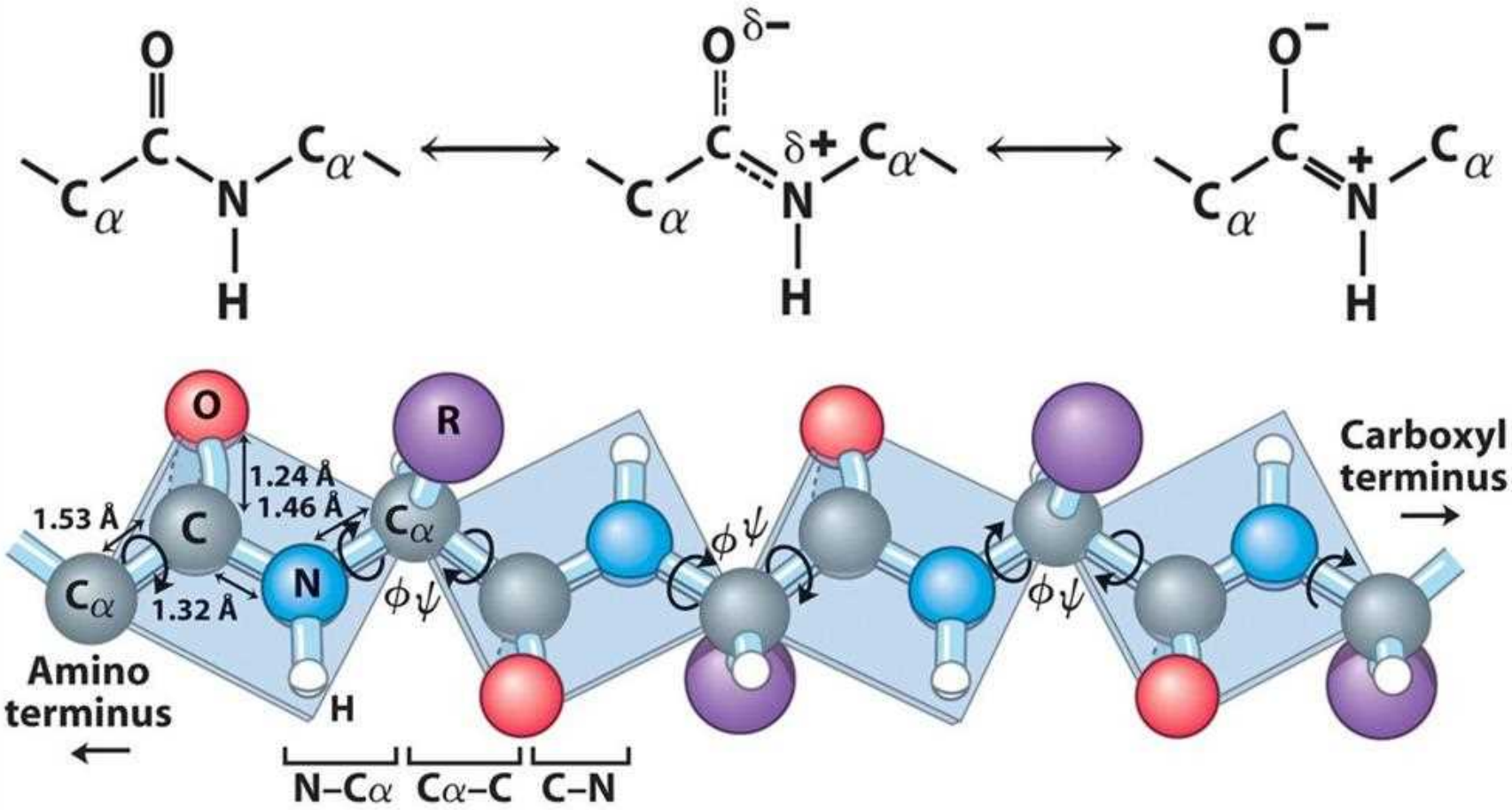
- In strained cyclic systems (cyclopeptides, proline derivatives).
- With large substituents at the N atom (alkylated derivatives of amino acids)



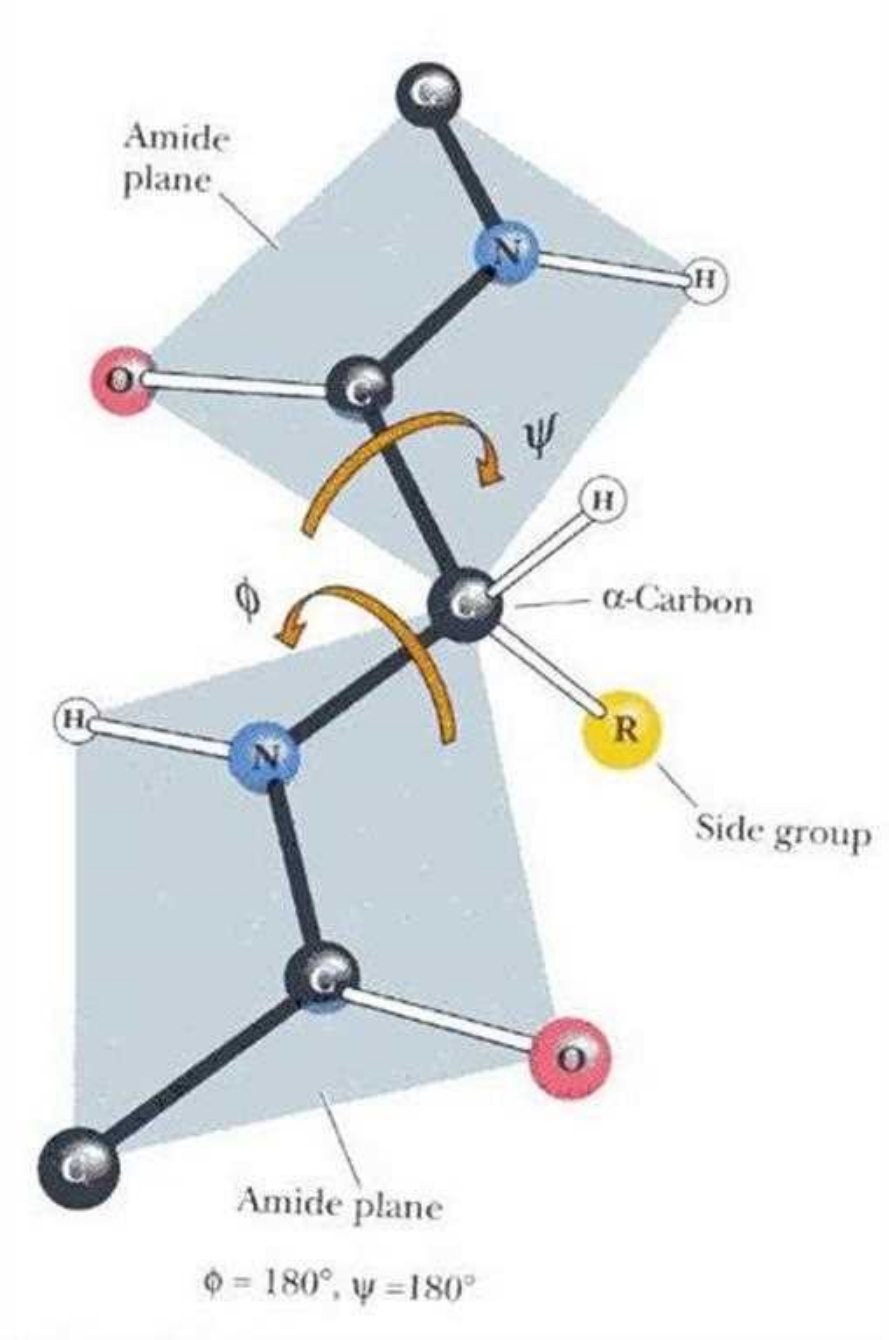
90%



10%



1. The peptide bond is intermediate between a single and a double bond. Its length is 1.32 A.
2. Rotation around the peptide bond is impossible.
3. 6 atoms near the peptide bond (Ca, O, N, H) are in the same plane.
4. The peptide bond can be considered a dipole because the carbonyl oxygen has a partially negative charge (d-) and the nitrogen has a partially positive charge ($\delta+$).



Dihedral angles between bonds:

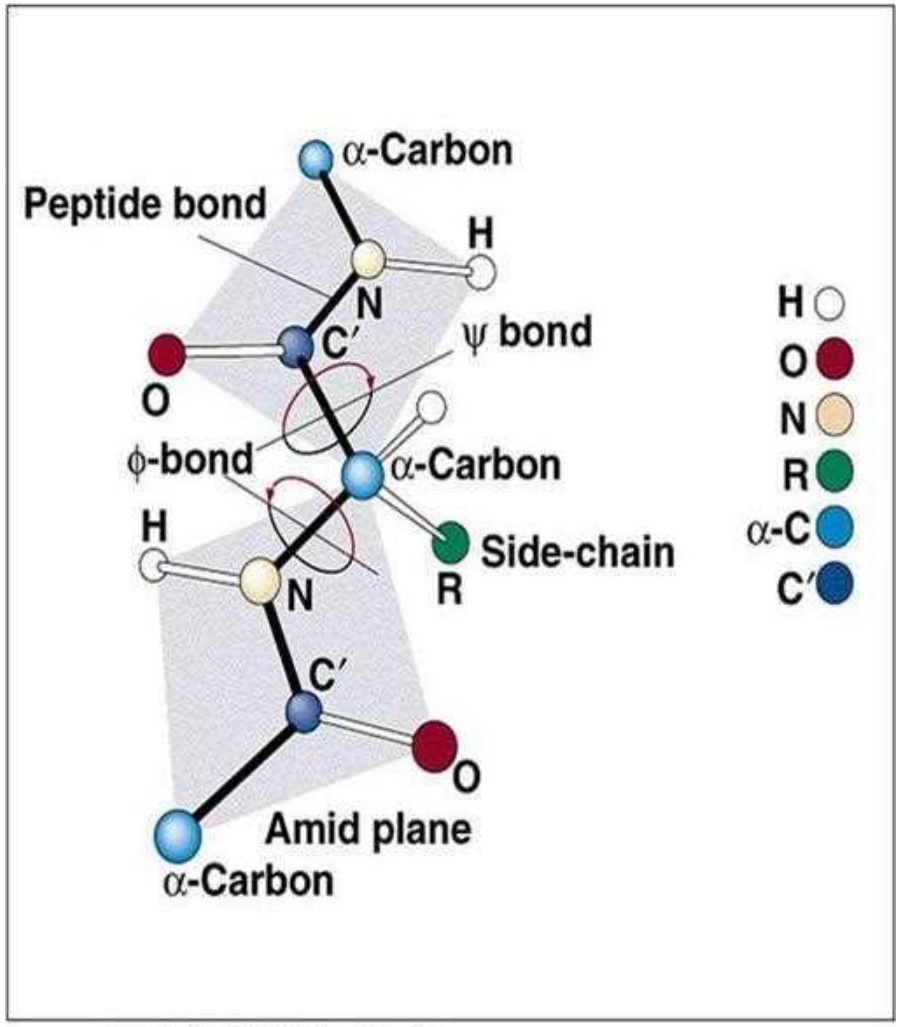
φ (phi) –angle of rotation around the bond N-Cα

ψ (psi) –angle of rotation around the bond Cα-C

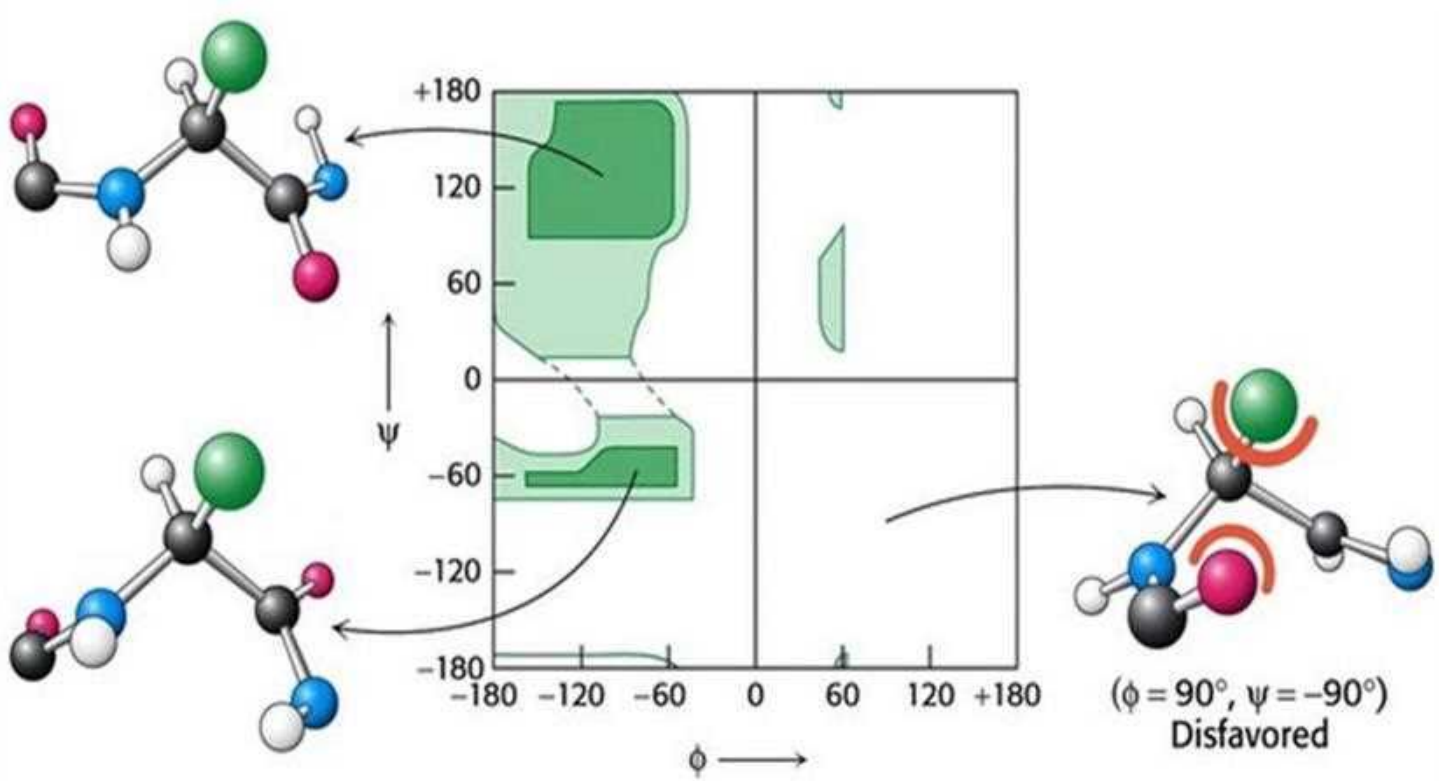
ω (omega): the angle ω determines the rotation around the peptide bond, which is usually fixed, so φ and ψ are the main angles describing the bends of the polypeptide chain.

The values of the dihedral angles between the N-Cα and Cα-C bonds cannot be arbitrary!!!
Some pairs of values (φ, ψ) are prohibited due to steric constraints.

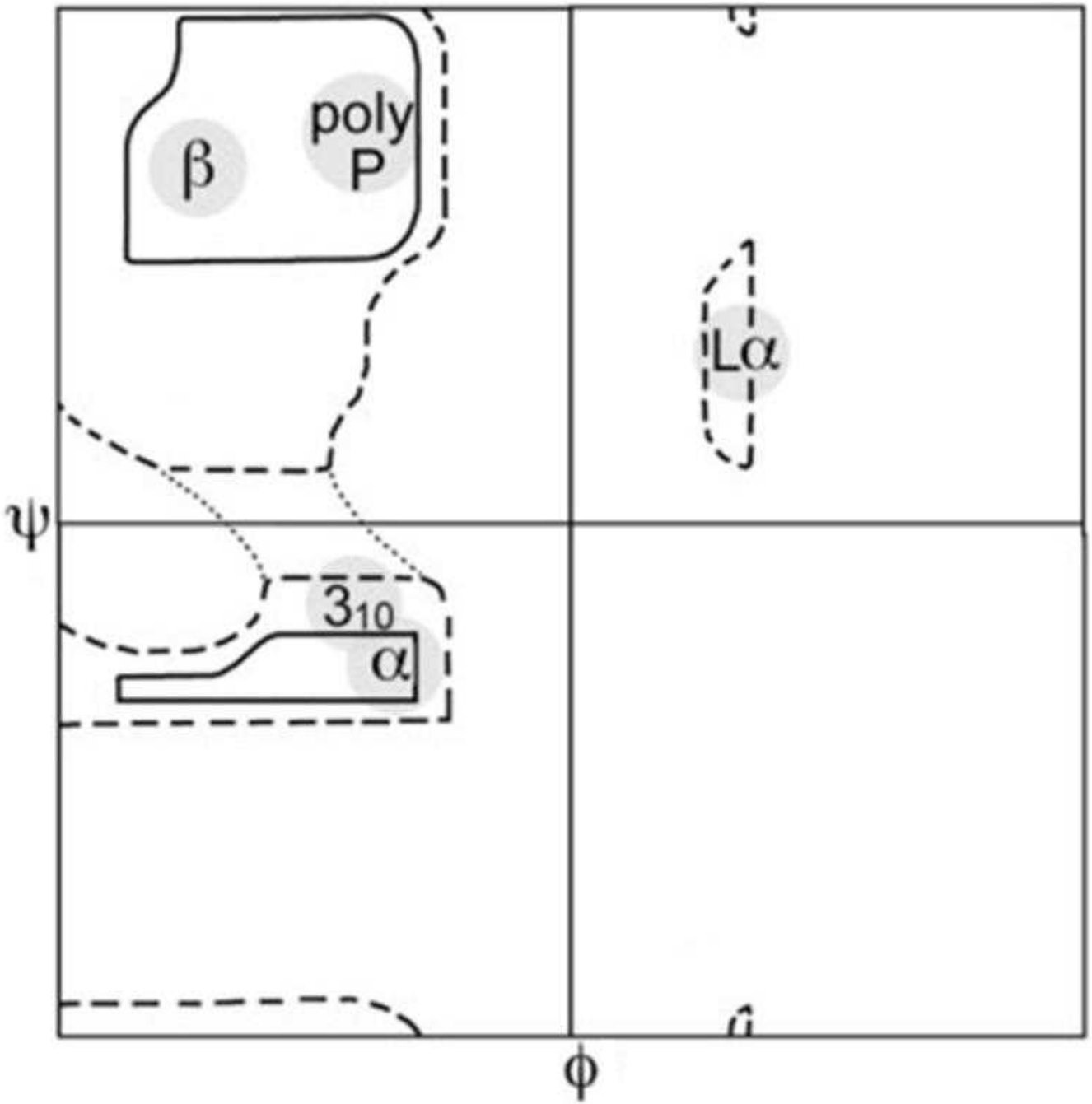
Ramachandran conformational plot



The (φ, ψ)-map shows which pairs of angle values φ and ψ are allowed for a given residue.

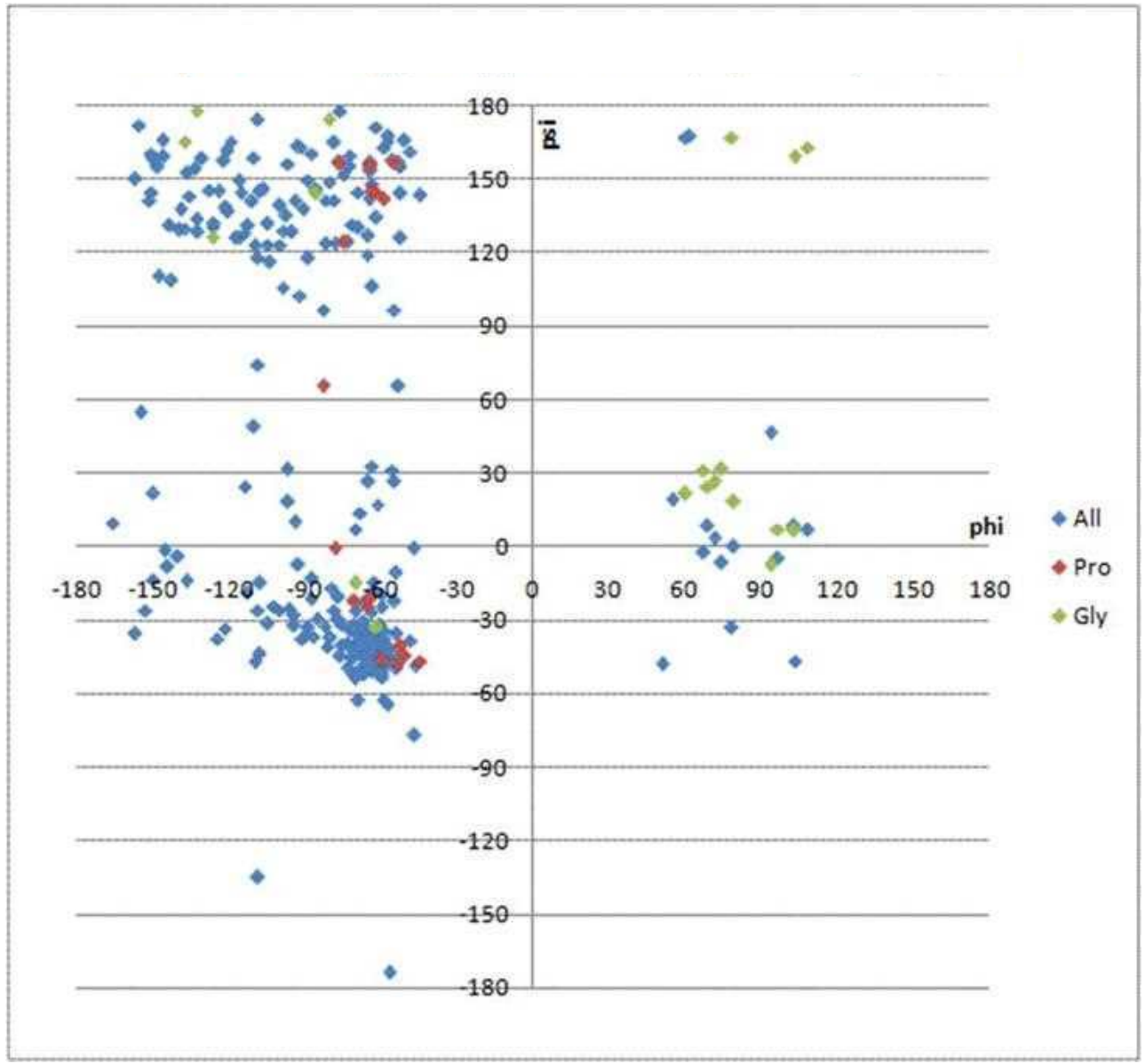


All three angles are 180° in the configuration shown.

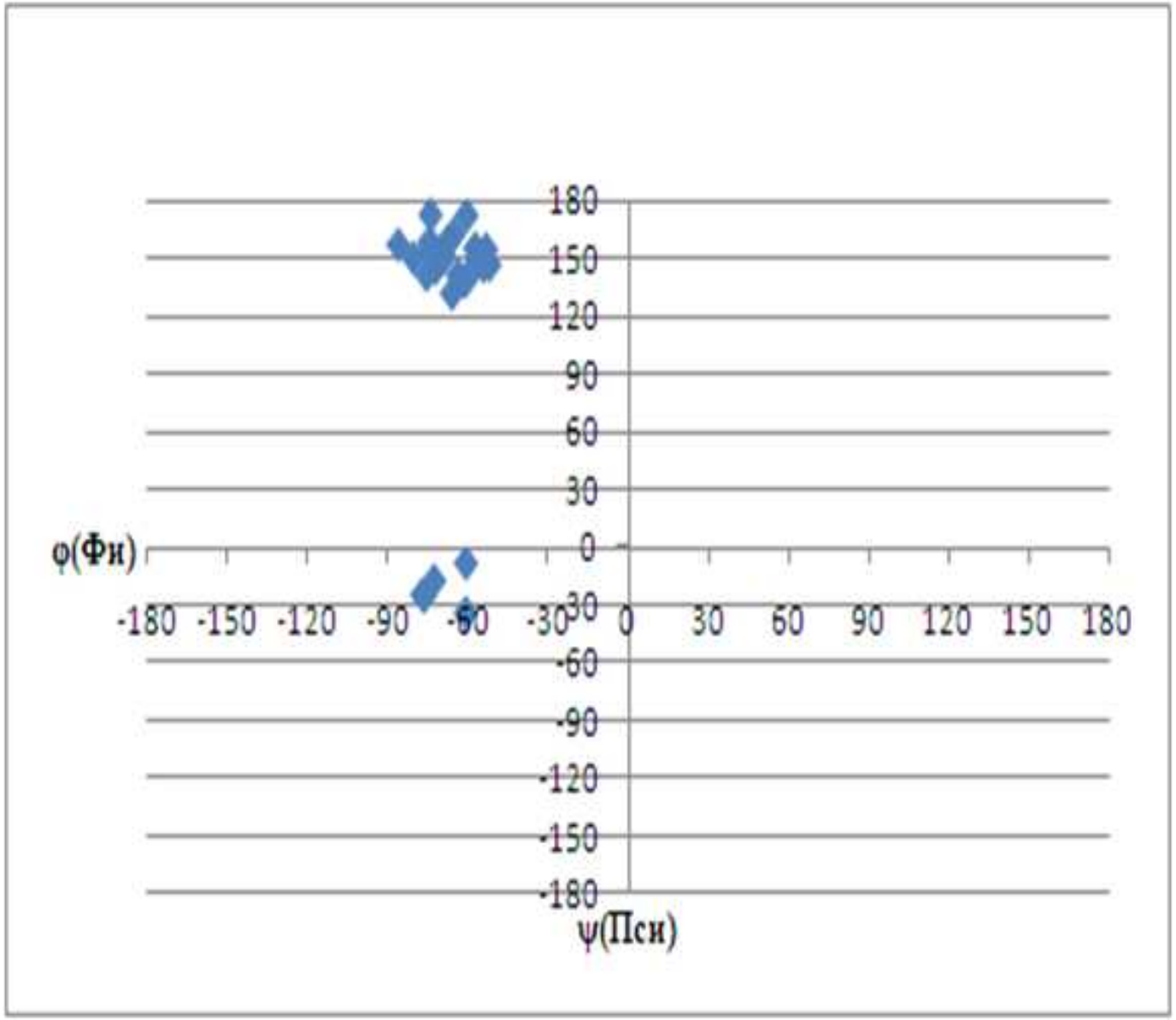


The figure shows the allowed φ,ψ conformational regions of the backbone from Ramachandran et al.'s 1963 and 1968 hard sphere calculations: the full radius is indicated by the solid outline, the reduced radius by the dashed line, and the weakened tau angle (N-Cα-C) by the dotted lines [doi:10.1016/S0065-3233(08)60402-7]. Because the dihedral angle values are circular, and 0° corresponds to 360°, the edges of the Ramachandran plot "wrap" from right to left and bottom to top.

Ramachandran’s plot for all amino acids, glycine and proline



Ramachandran’s plot for proline

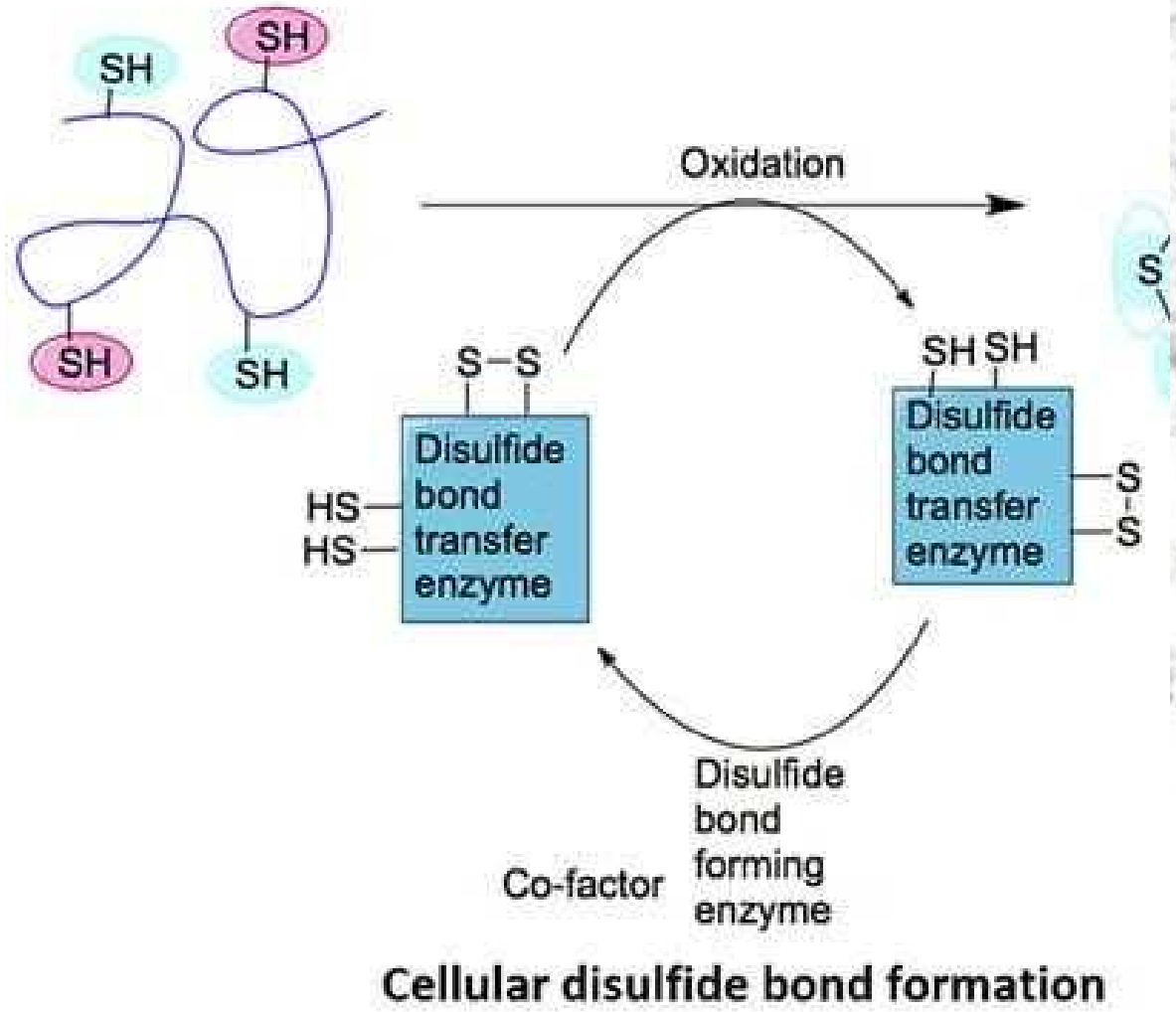


Proline is all about rigidity. It disrupts regular structures like protein’s helices but stabilizes turns, defining the protein’s structure.

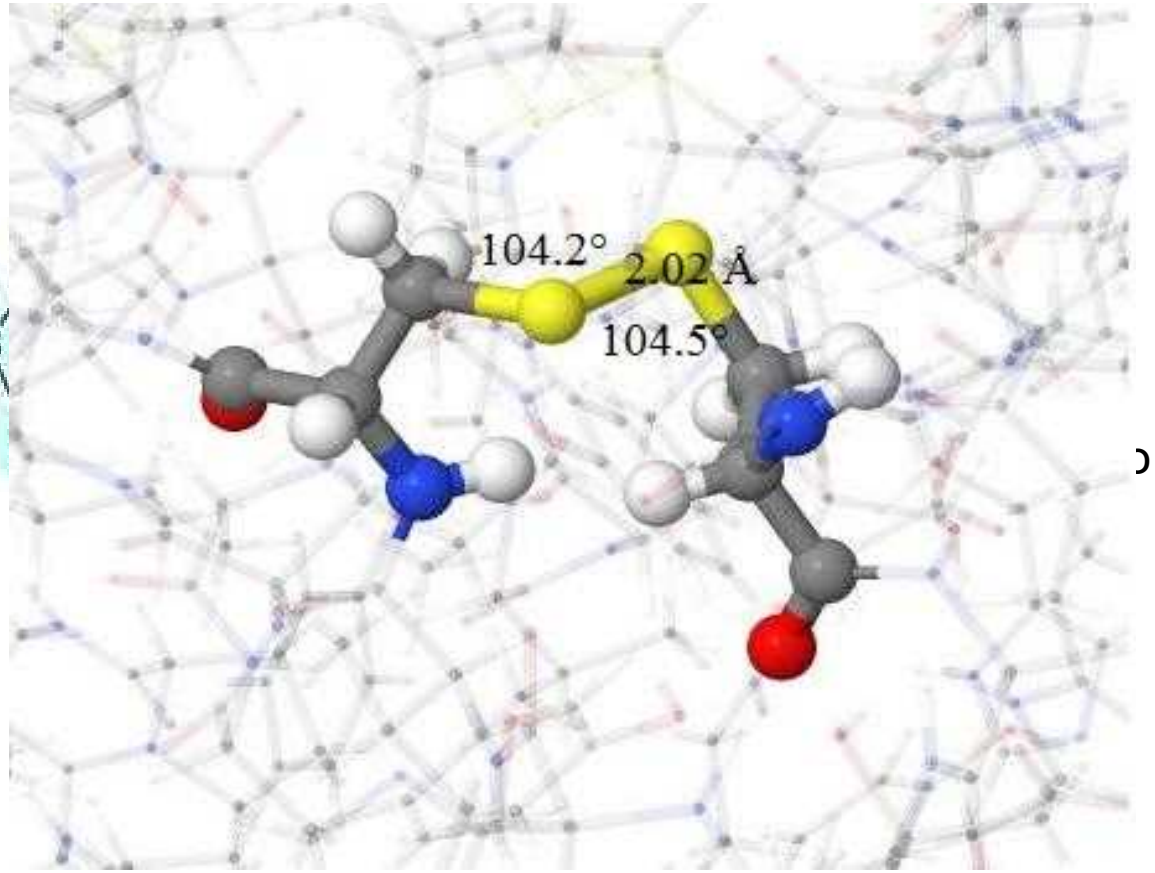
Glycine is all about flexibility. Its small size allows the protein chain to make sharp turns and adopt conformations that are sterically hindered for any other amino acid.

Disulfide bond

Disulfide bonds form when two cysteine residues, with their sulfhydryl groups, are oxidized to create a covalent (S-S) bond.



Cellular disulfide bond formation

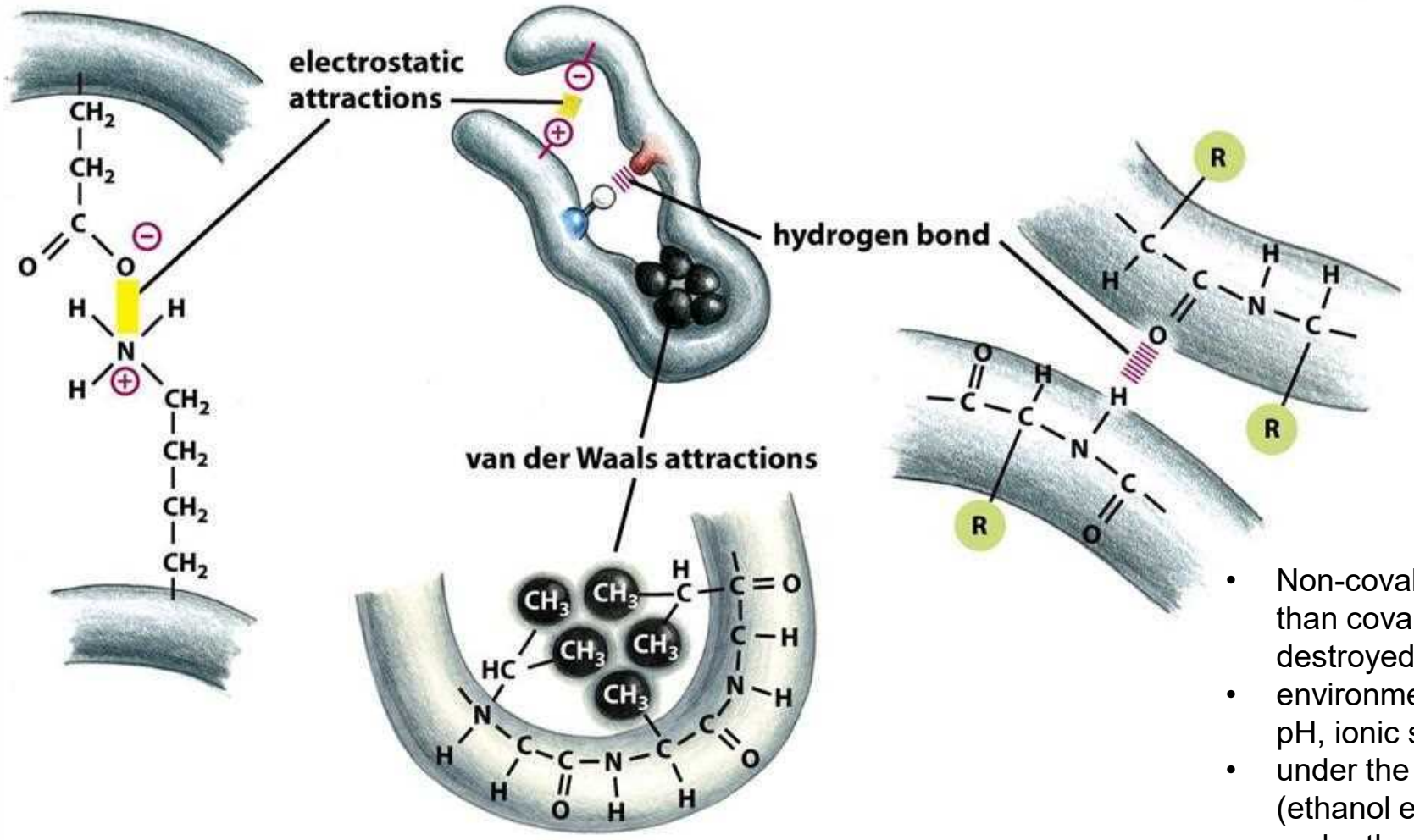


A disulfide bridge in the insulin protein. Atoms are colored: gray are carbon atoms, blue are nitrogen atoms, and red are oxygen atoms. The image was generated using the Jmol.

Non-covalent bonds -

These are physicochemical interactions that occur between different chemical groups that have a high affinity for each other (affinity is the ability to interact).

Non-covalent bonds determine the spatial organization of the protein molecule



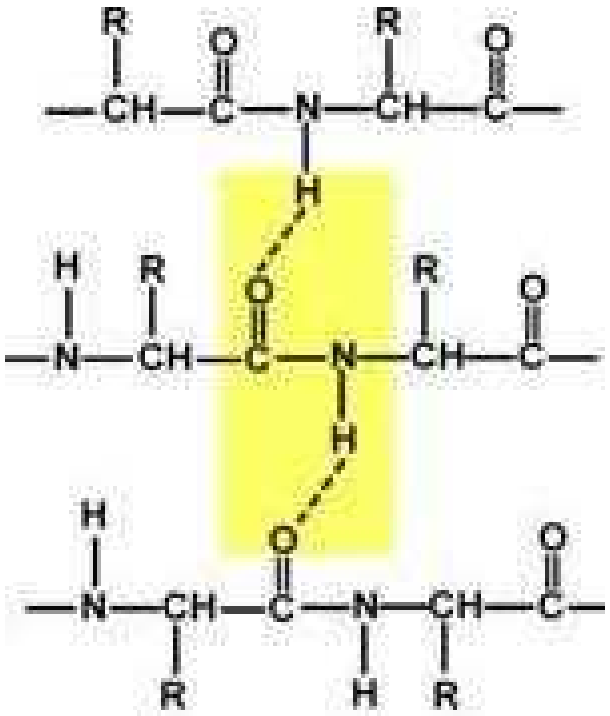
Non-covalent bonds include:

- hydrogen bonds
- salt bridges (ionic bonds)
- hydrophobic interactions

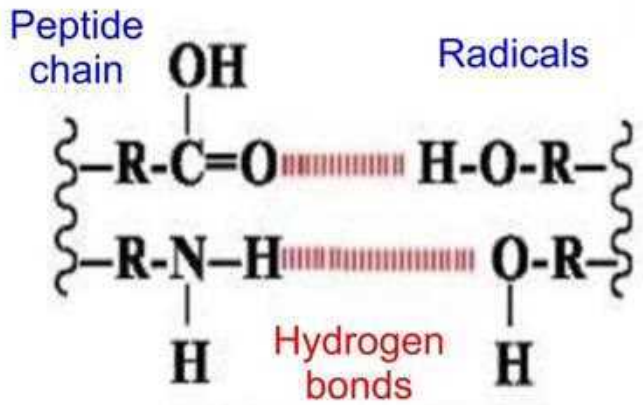
- Non-covalent bonds are tens of times weaker than covalent bonds, and therefore are destroyed when:
- environmental conditions change (temperature, pH, ionic strength of the solution)
- under the influence of organic solvents (ethanol et al.)
- under the influence of detergents

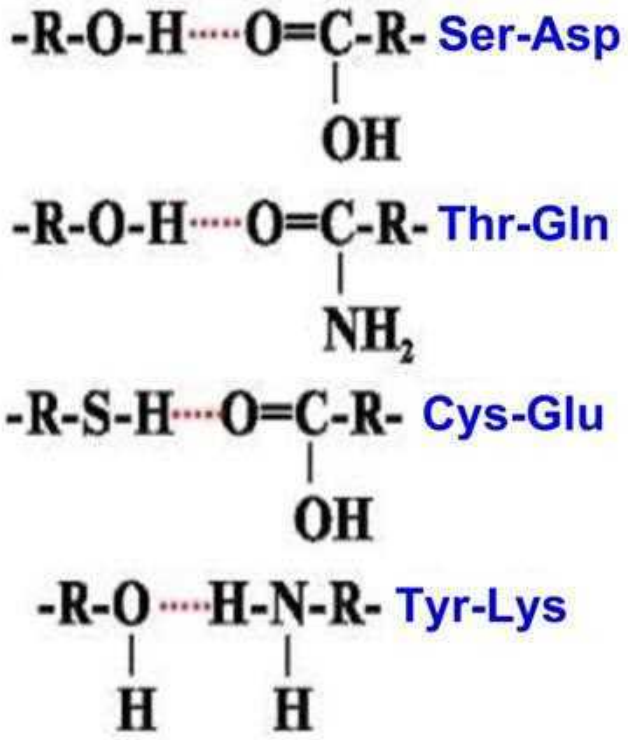
A hydrogen bond is a bond between two electronegative atoms through a hydrogen atom that is covalently bonded to one of them

Between the carbonyl (>C=O) and imino (>N-H) groups of peptide bonds

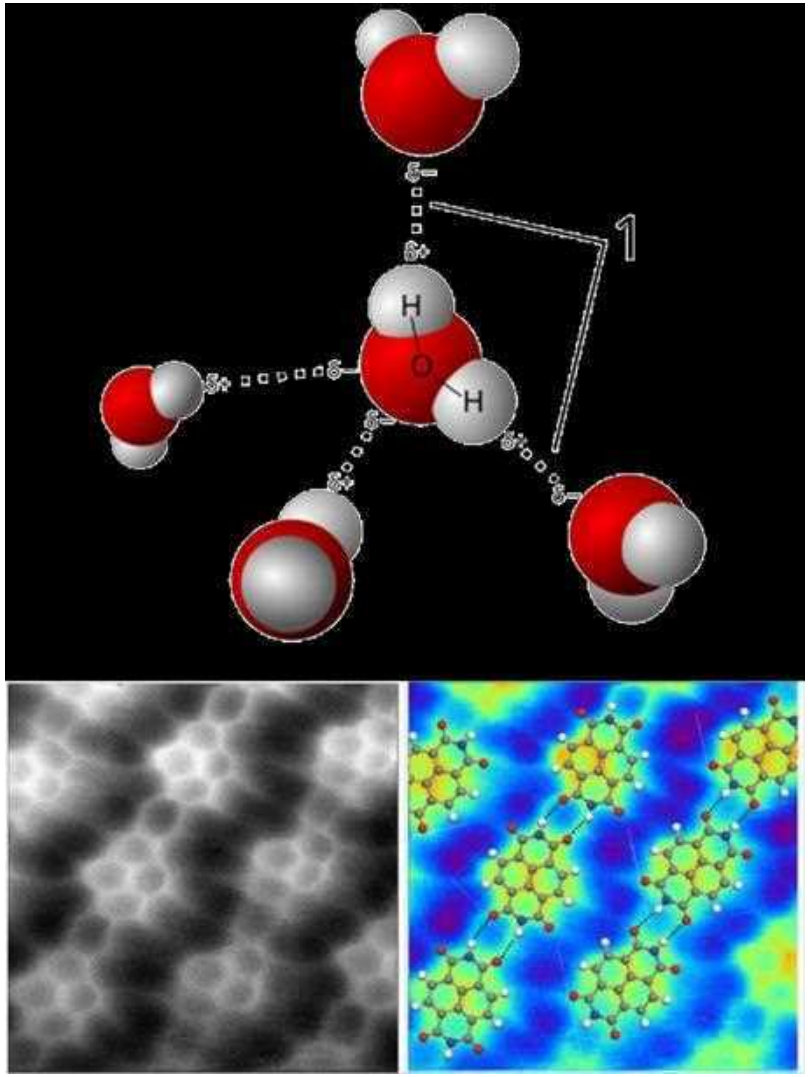


Between the side chains of polar amino acids

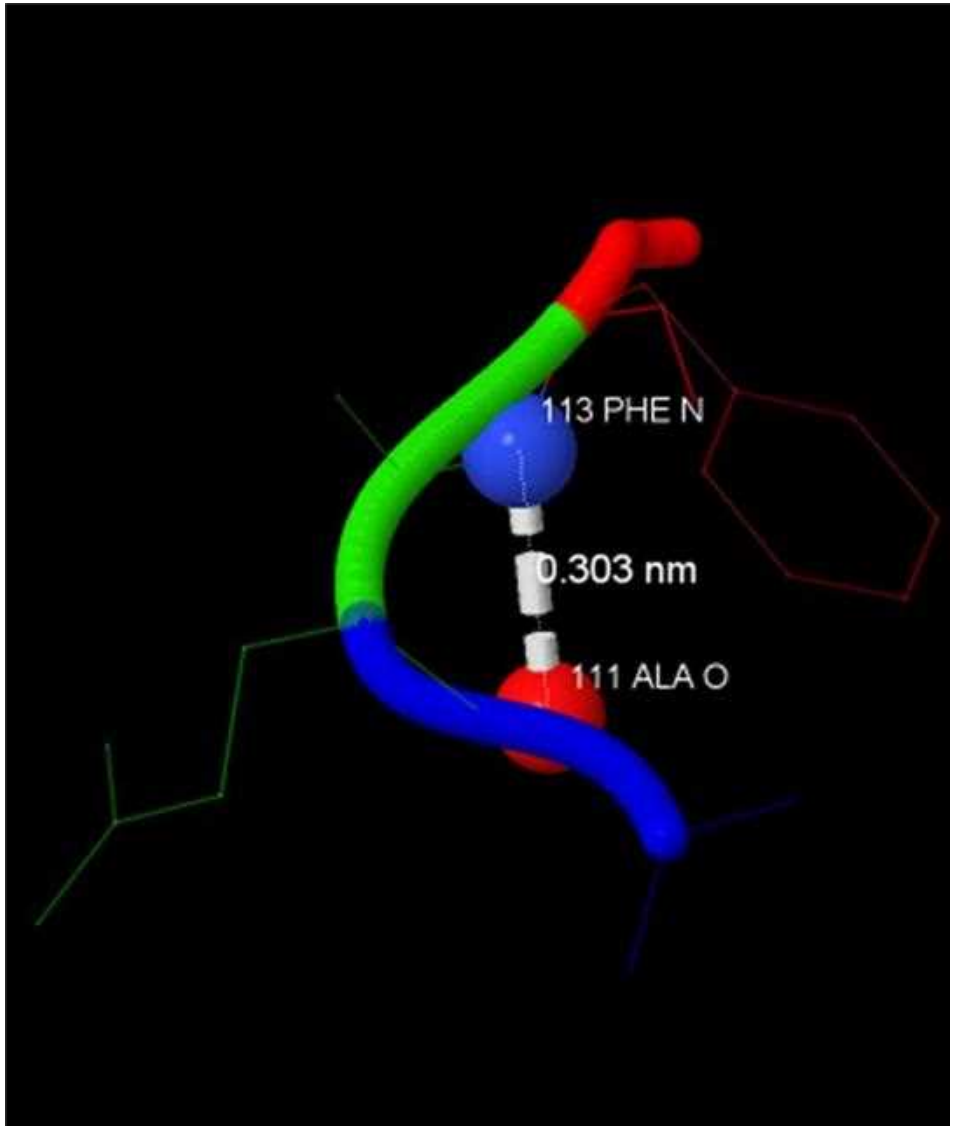




Hydrogen bond between ALA and PHE residues. The donor, acceptor, and bond length are shown.



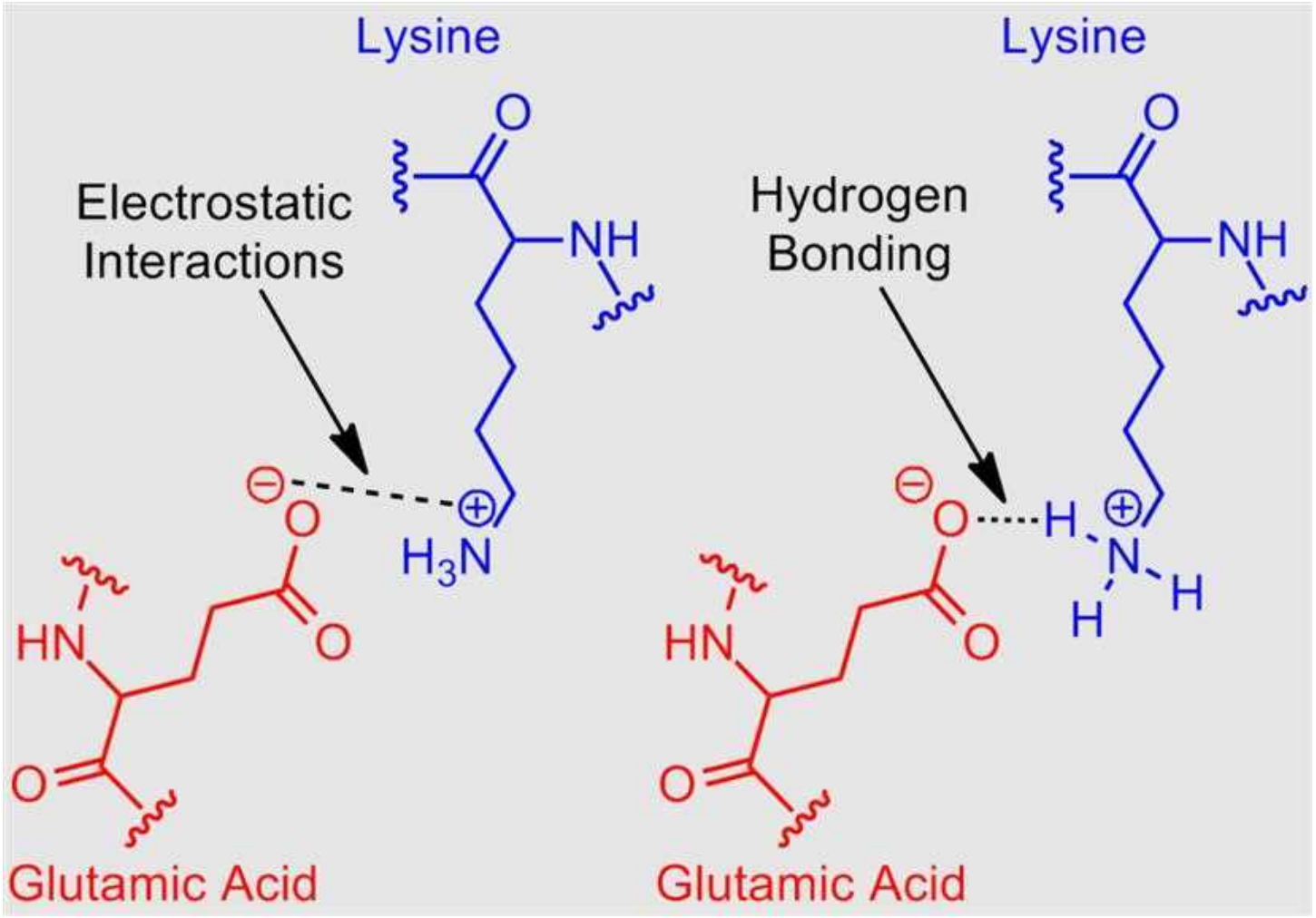
AFM image of naphthalene tetracarbonic diimide molecules on silver-tipped silicon interacting via hydrogen bonds



Hydrogen bond between ALA and PHE residues. The donor, acceptor, and bond length are shown

Salt bridges (ionic bonds)

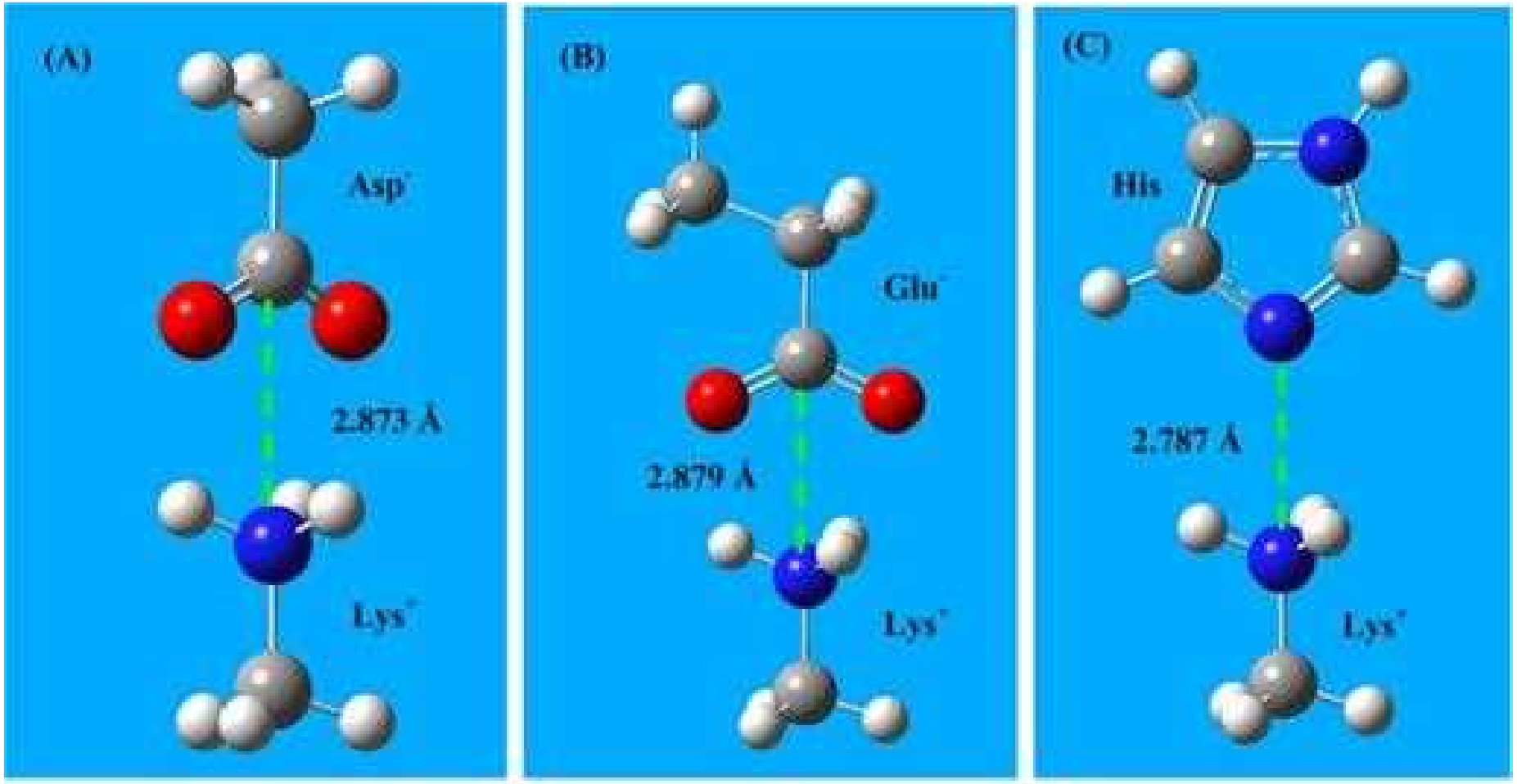
A salt bridge is a non-covalent interaction that combines an electrostatic attraction between oppositely charged groups with a hydrogen bond and are formed by the interaction of positively charged amino acid side chains (like lysine or arginine) with negatively charged ones (like aspartate or glutamate)



Example of salt bridge between amino acids glutamic acid and lysine demonstrating electrostatic interaction and hydrogen bonding

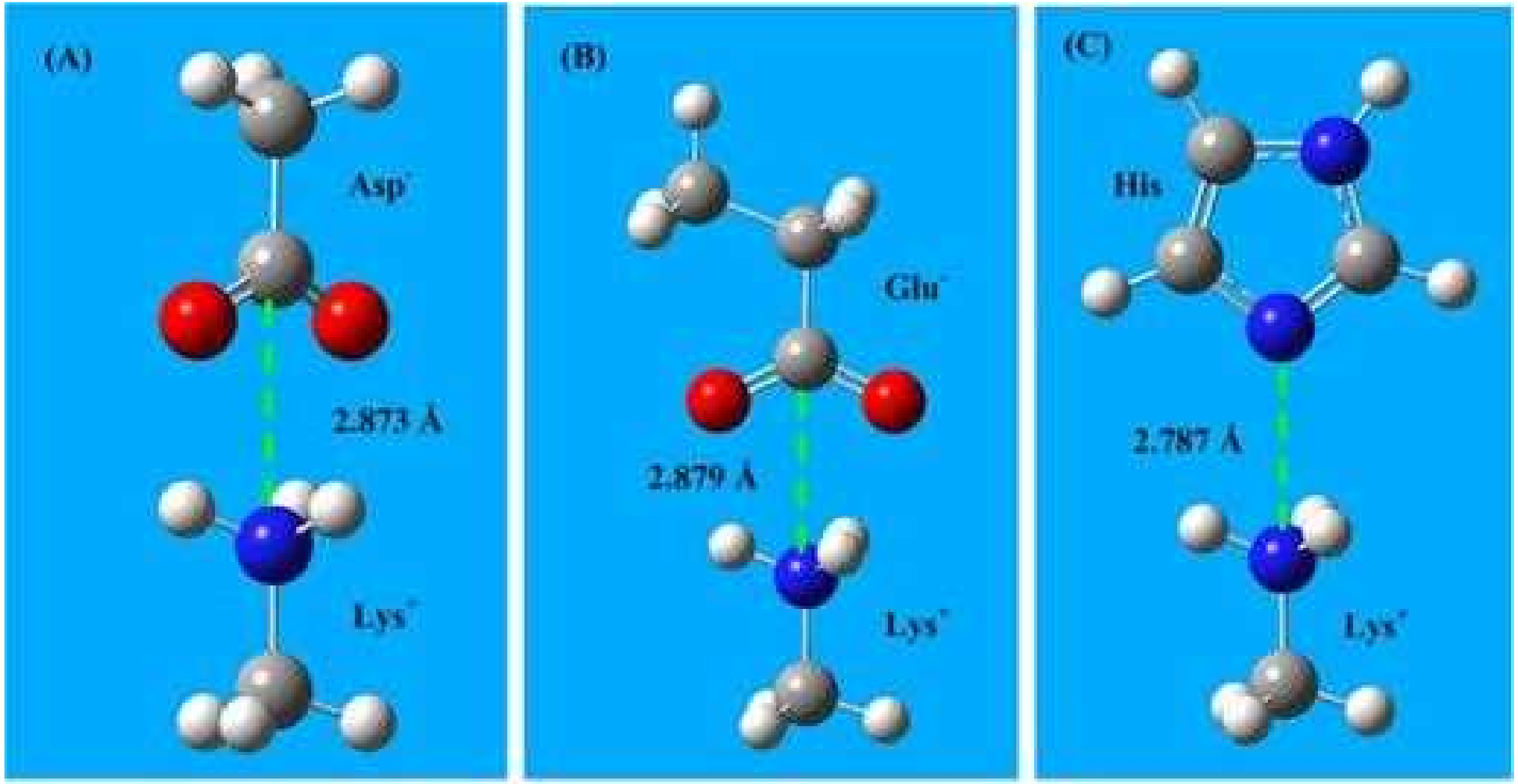
Salt bridges (ionic bonds)

The salt-bridge interaction structures between three amino acid cations (Arg+ Lys+, and His+) and three acidic amino acids (Asp-, Glu-, and His)



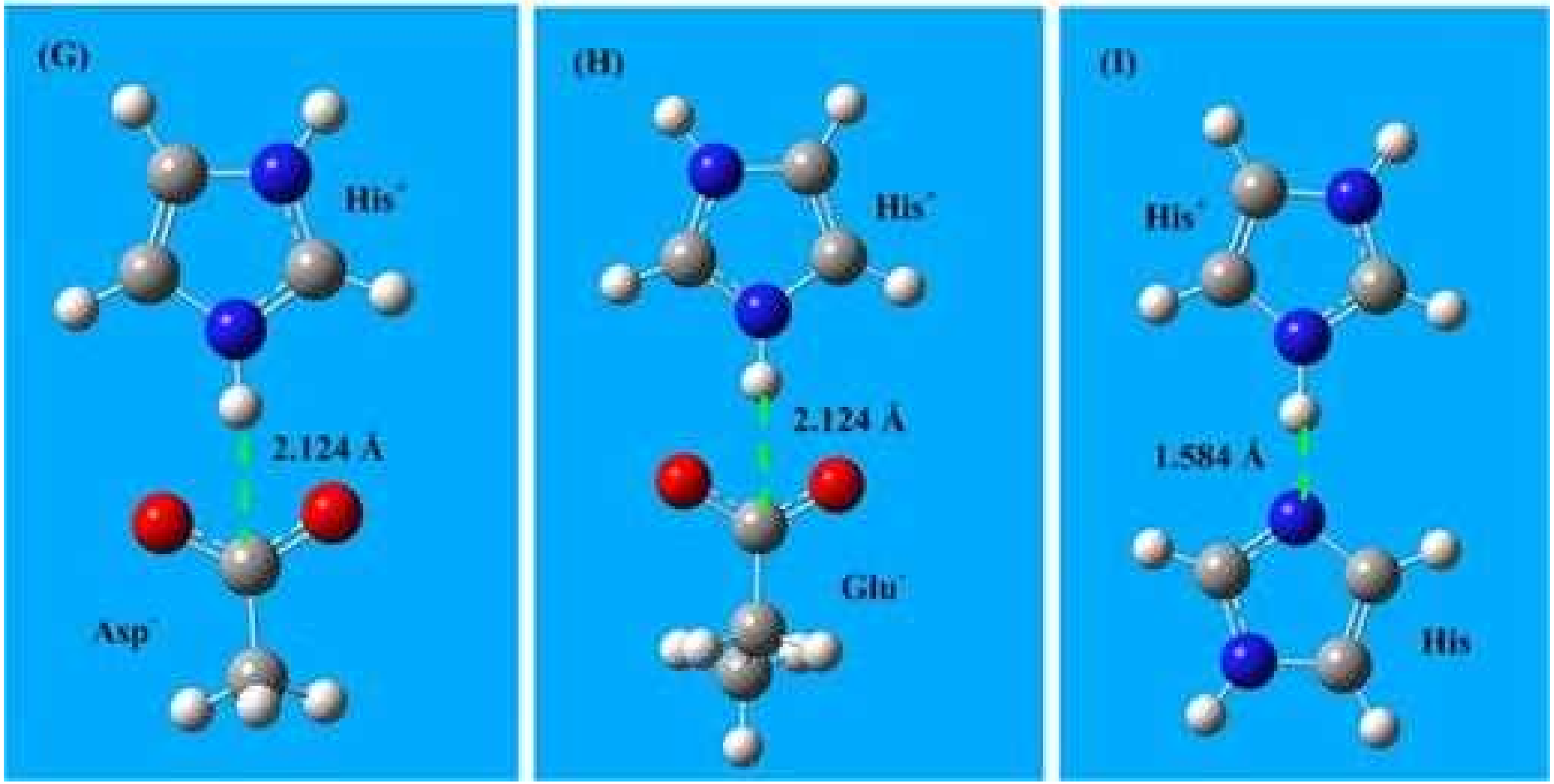
Salt bridges (ionic bonds)

The salt-bridge interaction structures between three amino acid cations (Arg+ Lys+, and His+) and three acidic amino acids (Asp-, Glu-, and His)



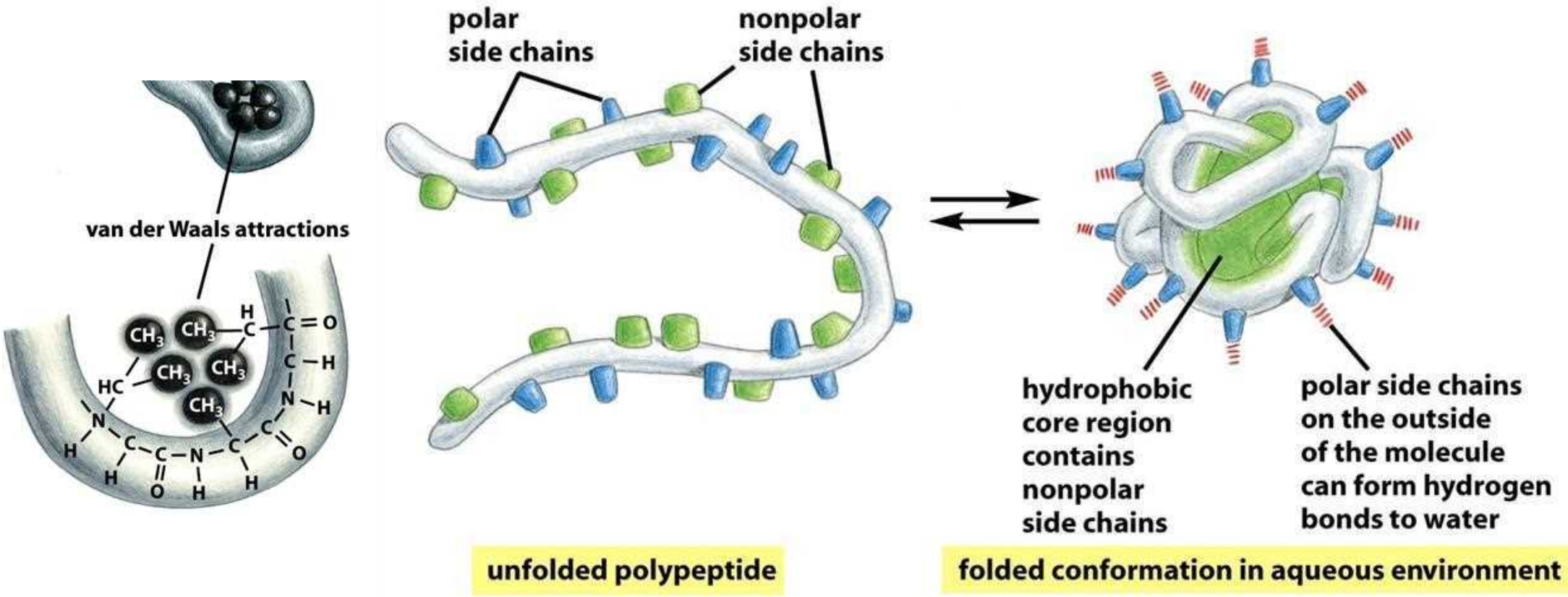
Salt bridges (ionic bonds)

The salt-bridge interaction structures between three amino acid cations (Arg+ Lys+, and His+) and three acidic amino acids (Asp-, Glu-, and His)



Hydrophobic interaction

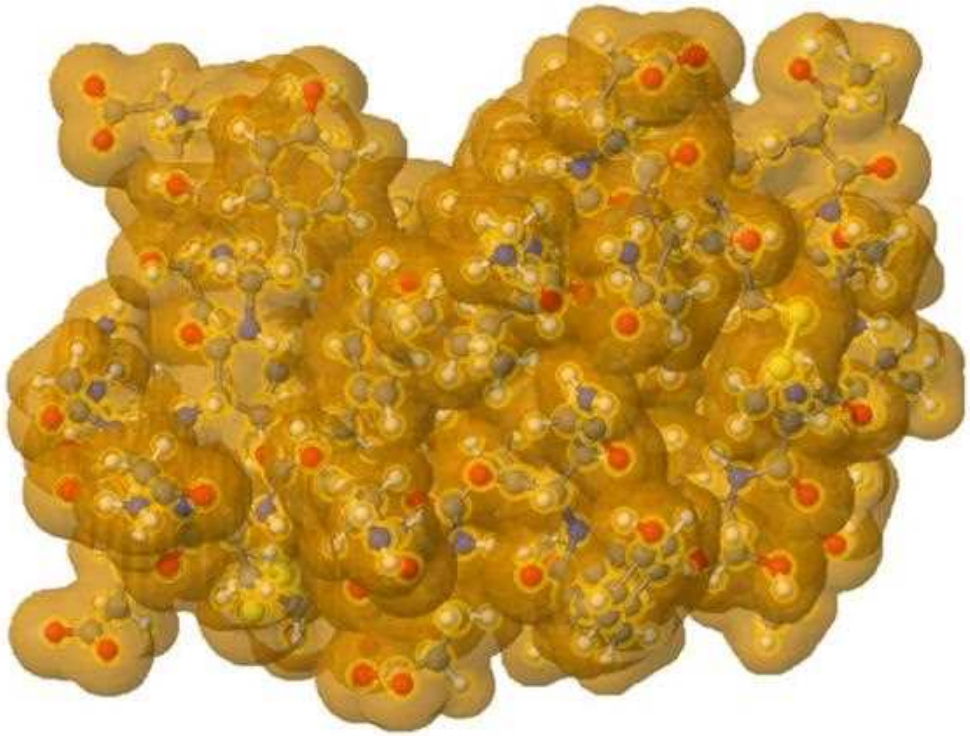
Non-specific attraction between radicals of hydrophobic amino acids due to Van der Waals forces (electrostatic interaction between opposite poles of a dipole)



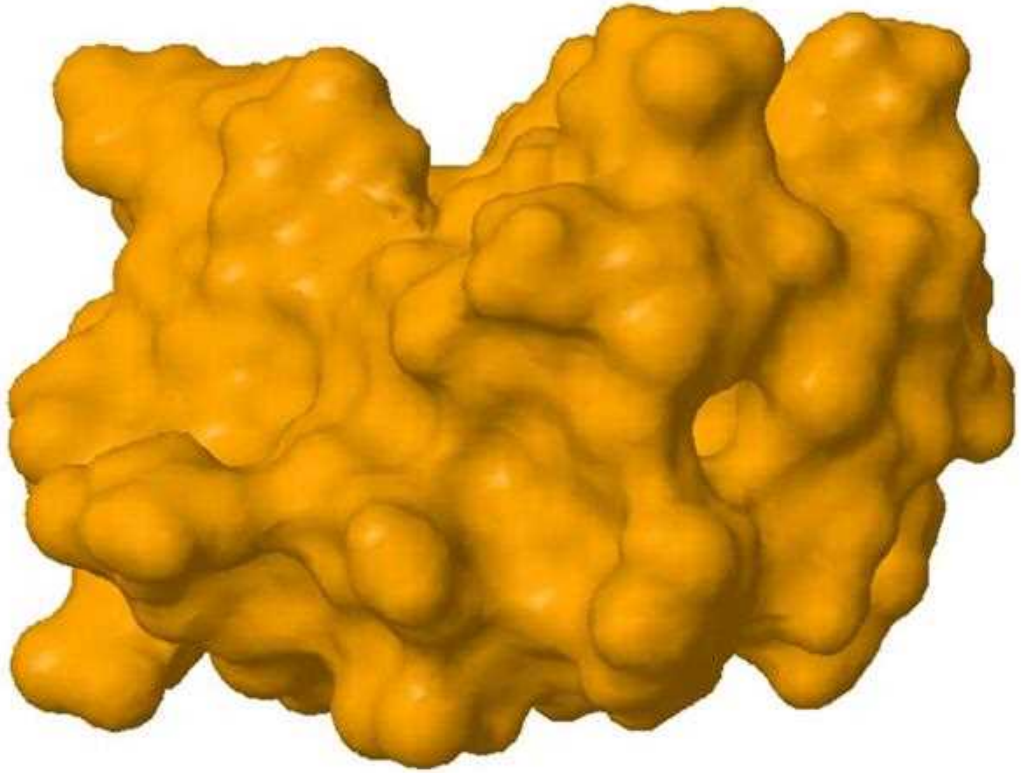
Hydrophobic radicals, avoiding contact with water, gather inside the protein molecule, forming a dense *hydrophobic core.* Polar radicals tend to move outward toward the aqueous phase.

Hydrophobic interaction

Non-specific attraction between radicals of hydrophobic amino acids due to Van der Waals forces (electrostatic interaction between opposite poles of a dipole)



Van der Waals surface of insulin



Molecular surface of insulin

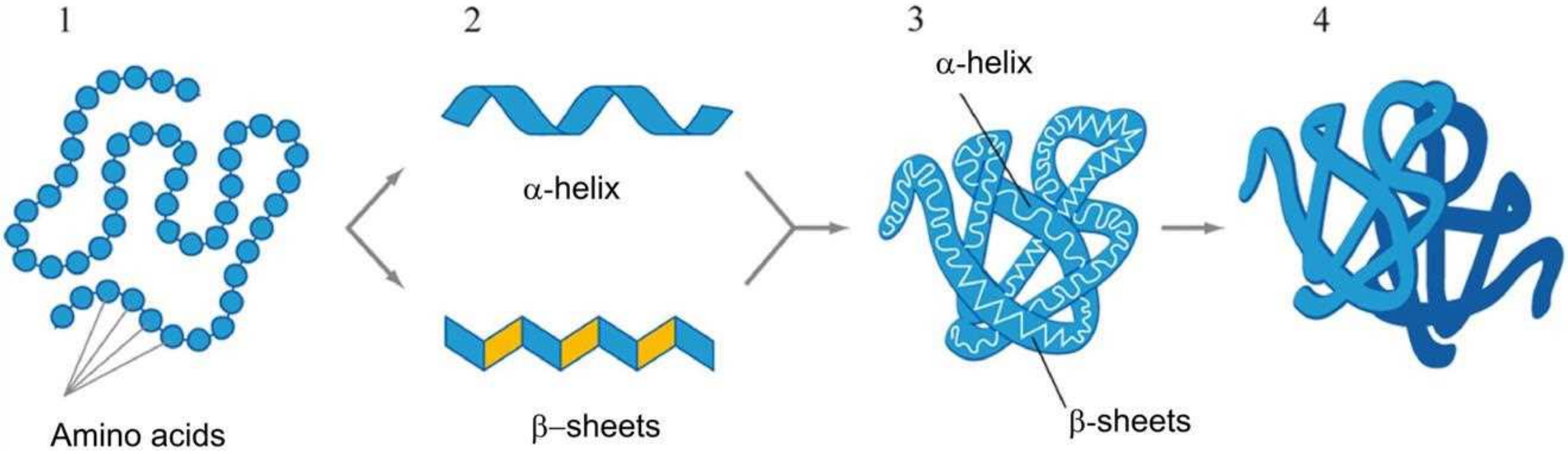
Created using the Jsmol

Spatial organization of proteins

At the core of every protein is a polypeptide chain. It's not simply elongated in space, but organized into a three-dimensional structure—α– conformation.

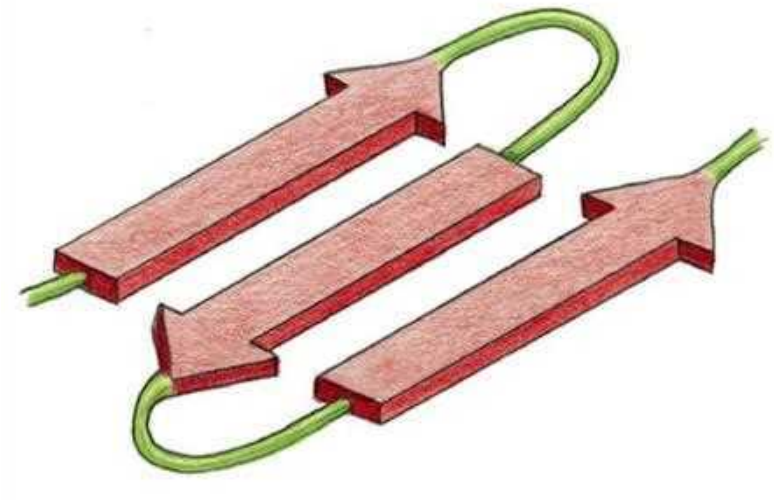
There are four levels of protein spatial organization:

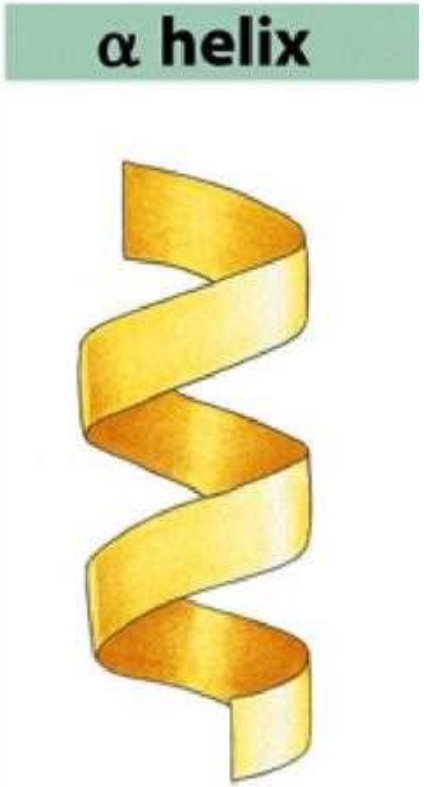
- *primary*—the polypeptide chain
- *secondary*—the α-helix and β-sheets
- *tertiary*—the polypeptide chain's spatial arrangement
- *quaternary*—a complex of several polypeptide chains with a tertiary structure



Types of protein secondary structure in order of prevalence

- α-helix 4_{13} or 3.6_{13}
- β-structure or sheet
- reverse turns
- α-helix 3_{10}



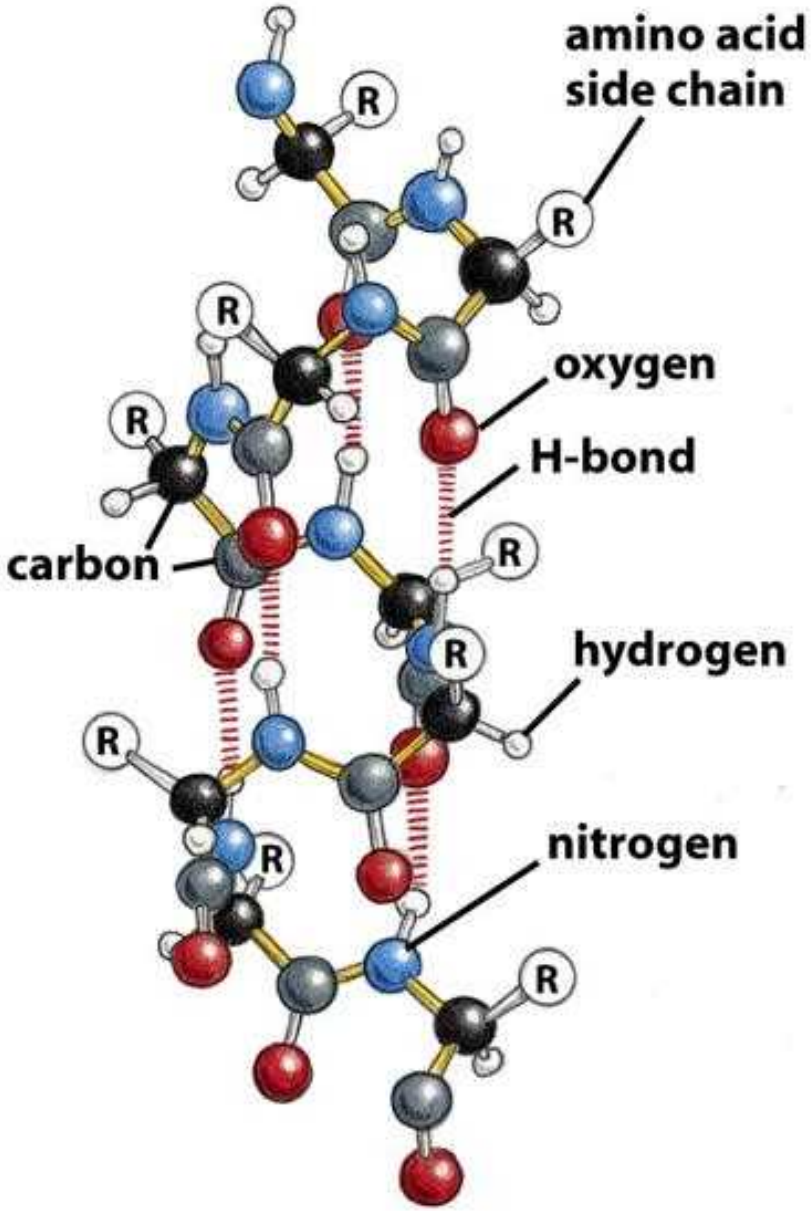


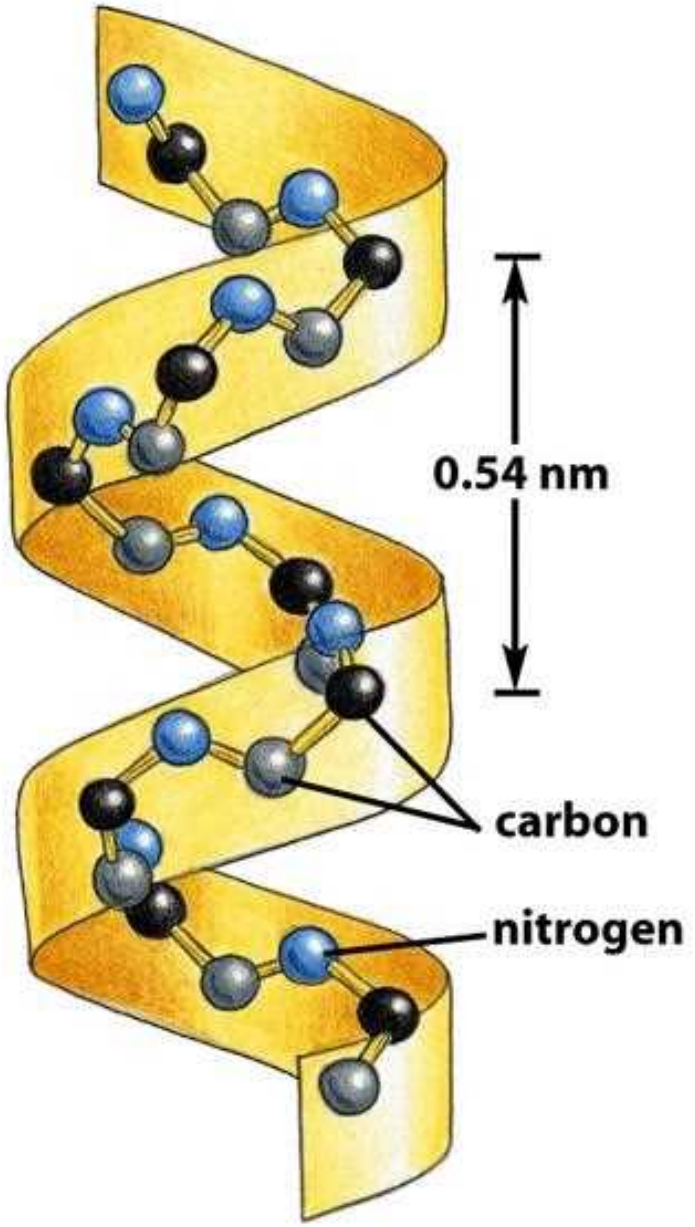
α-helix

The a-helix is the most common structural organization in protein secondary structure.

The a-helix has a right-handed helical conformation, in which each N−H group of the backbone forms *hydrogen bonds* with a C=O group of the backbone AA located four residues earlier in the protein sequence.

Hydrogen bonds are parallel to the chain axis, and the atoms involved in its formation are arranged linearly. This gives the molecule stability.



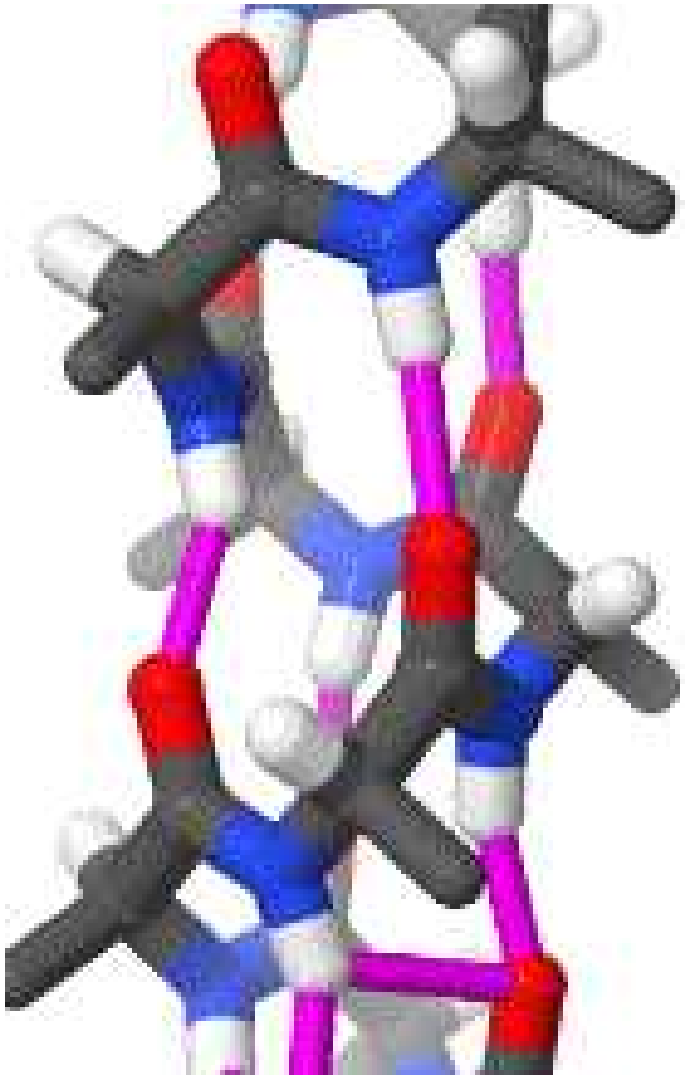


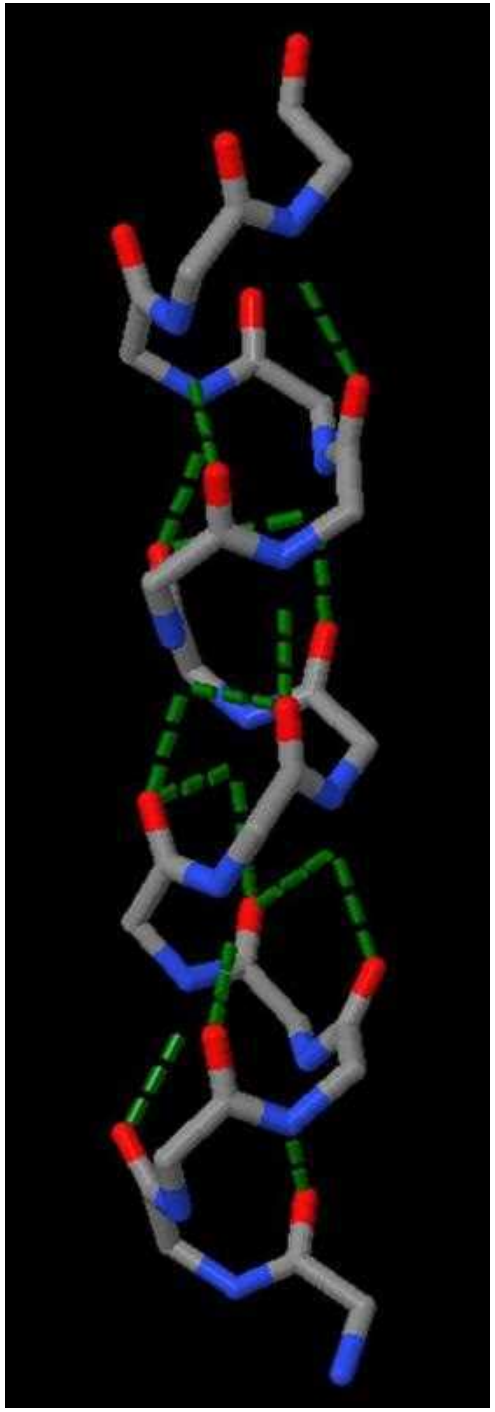
α-helix

This type of a-helix is also commonly called: Pauling-Corey-Branson α-helix (named after the three scientists who described its structure)

3.6 $_{13}$ -helix or 4 $_{13}$ -helix, since each ring contains 3.6 amino acids

Each residue in the helix participates in two hydrogen bonds (except for the residues at the ends). These hydrogen bonds are nearly parallel to the axis. However, the hydrogen bond geometry is not ideal, as the peptide NH is not aligned precisely with the lone pair of electrons on the oxygen atom.



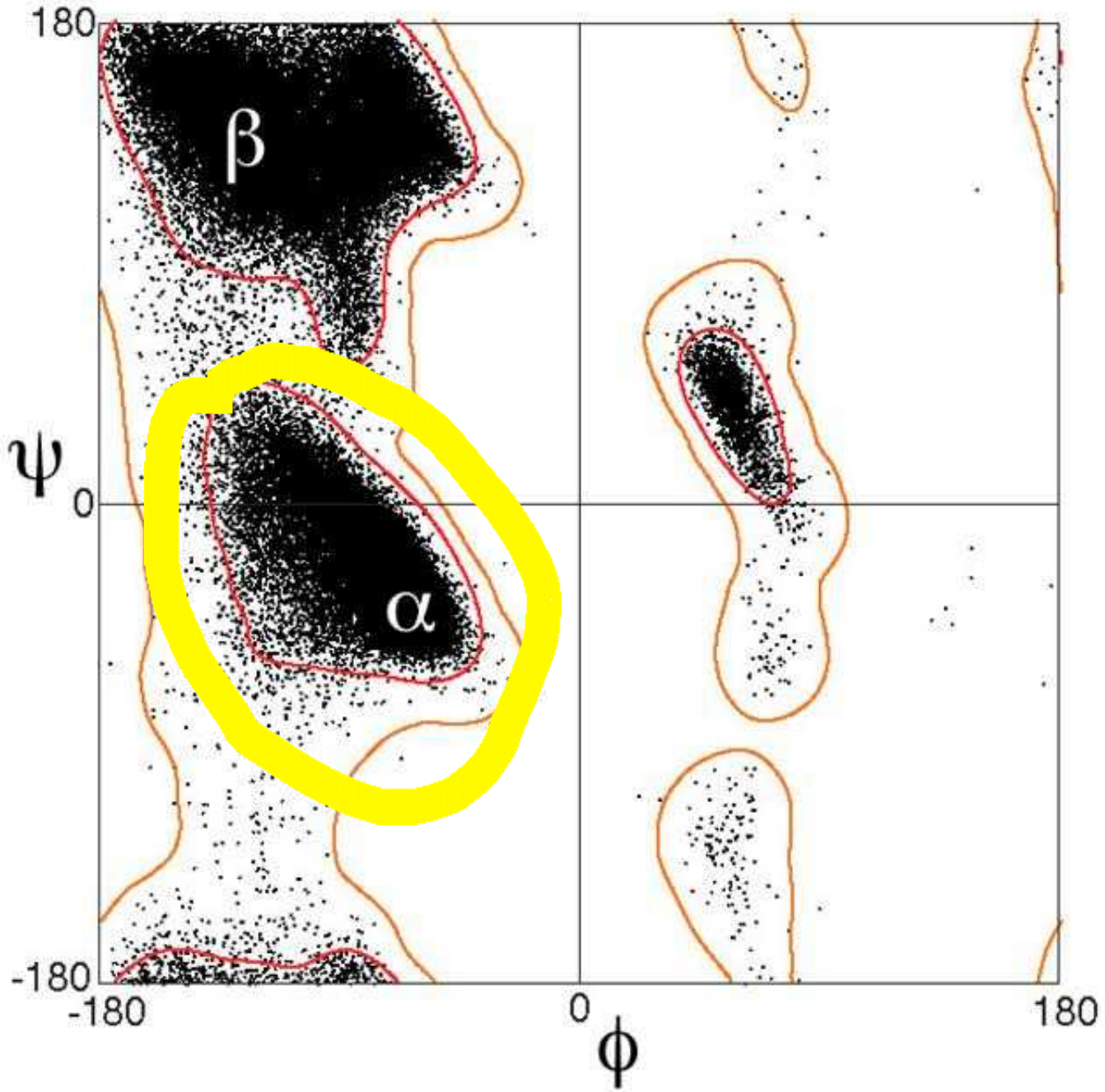


α-helix

Alpha helices can be right-handed (clockwise) or left-handed (counterclockwise). Although the Ramachandran plot shows that both right- and left-handed alpha helices are allowed conformations, the right-handed alpha helix is energetically more favorable due to fewer steric clashes between the side chains and the backbone.

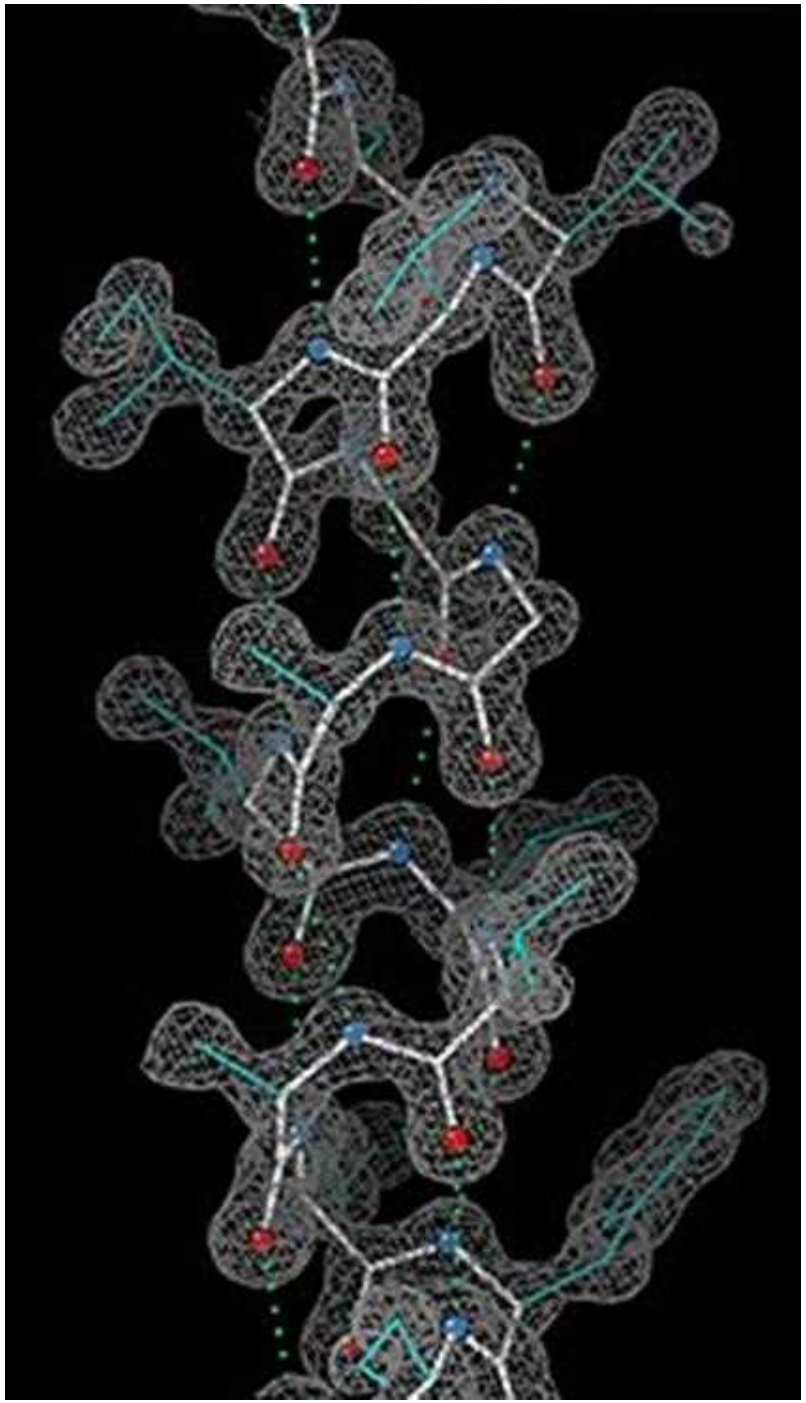
Therefore, all alpha helices in proteins are right-handed.

Ramachandran plot (φ,ψ plot) with data points for α-helix residues forming a tight diagonal cluster below and left of center, around the global energy minimum for the backbone conformation.



α-helix

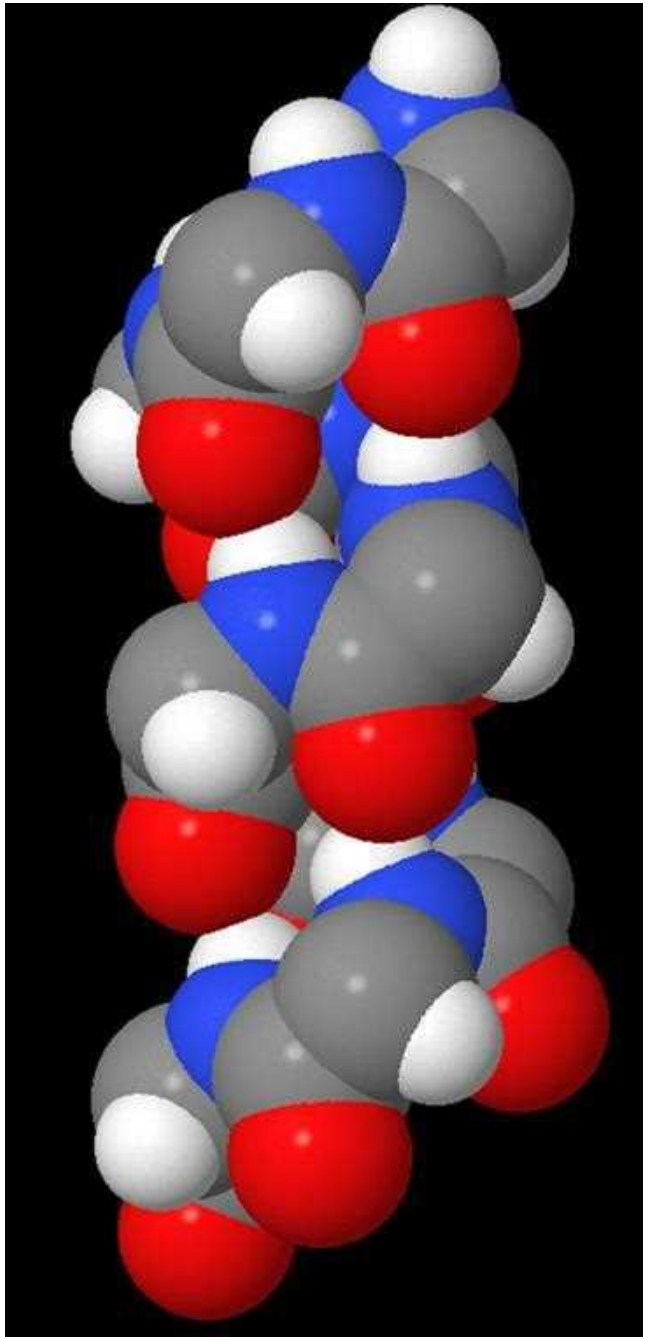
- The helices observed in proteins can range in length from four to over forty residues, but a typical helix contains about ten amino acids (about three turns).
- Short polypeptides do not exhibit significant a-helical structure in solution, as the entropic costs associated with folding the polypeptide chain are not offset by sufficient stabilizing interactions.
- In general, the backbone hydrogen bonds of a-helices are considered slightly weaker than those found in β-sheets and are easily attacked by surrounding water molecules.



The α-helix in ultra-high-resolution electron density contours, with oxygen atoms in red, nitrogen atoms in blue, and hydrogen bonds as green dashed lines. The N-terminus is at the top.

α-helix

- However, in more hydrophobic environments, such as the plasma membrane, oligopeptides readily adopt a stable a-helical structure.
- a-helices have been shown to be more stable, resistant to mutations, and amenable to engineering than β-sheet in natural proteins as well as in artificially designed proteins.



Space filling model of α-helix

Different amino acid sequences have different propensities to form an α-helical structure

Alanine, uncharged glutamate, leucine, charged arginine, methionine, and charged lysine have a particularly high propensity to form helices, while proline and glycine have a low propensity to form helices.

Higher numbers (more positive free energy changes) are less preferred

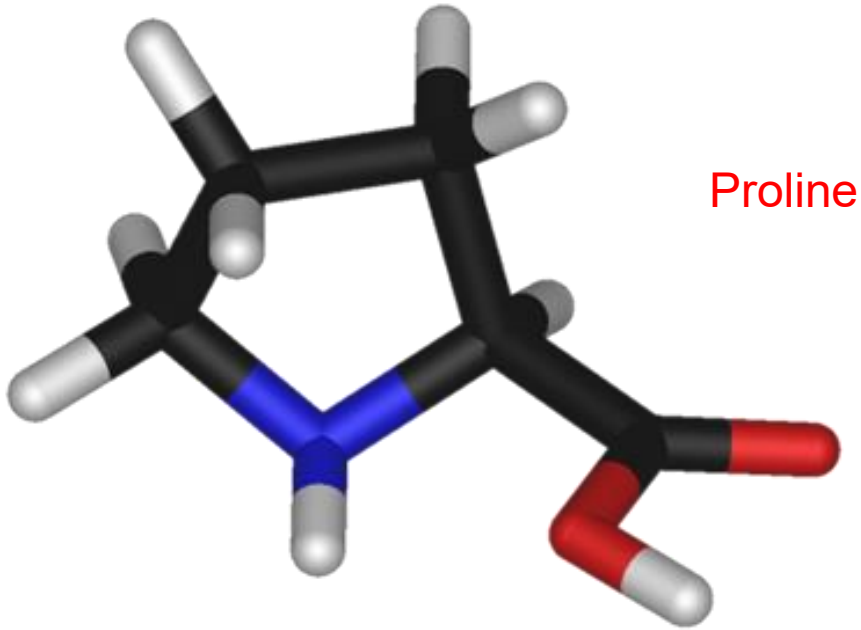
Amino acid	3-letter	1-letter	Helical penalty	
			kcal/mol	kJ/mol
Alanine	Ala	A	0.00	0.00
Arginine	Arg	R	0.21	0.88
Asparagine	Asn	N	0.65	2.72
Aspartic acid	Asp	D	0.69	2.89
Cysteine	Cys	C	0.68	2.85
Glutamic acid	Glu	E	0.40	1.67
Glutamine	Gln	Q	0.39	1.63
Glycine	Gly	G	1.00	4.18
Isoleucine	Ile	I	0.41	1.72
Leucine	Leu	L	0.21	0.88
Lysine	Lys	K	0.26	1.09
Methionine	Met	M	0.24	1.00
Phenylalanine	Phe	F	0.54	2.26
Proline	Pro	P	3.16	13.22
Serine	Ser	S	0.50	2.09
Threonine	Thr	T	0.66	2.76
Tryptophan	Trp	W	0.49	2.05
Tyrosine	Tyr	Y	0.53	2.22
Valine	Val	V	0.61	2.55

Different amino acid sequences have different propensities to form an α-helical structure

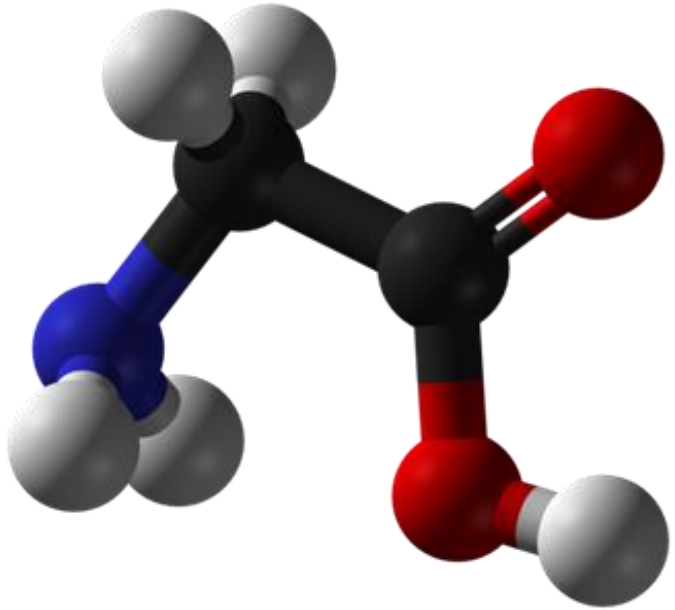
Proline either breaks or bends the helix, both because it cannot donate an amide hydrogen bond (because it lacks one) and because its side chain sterically interferes with the backbone of the previous turn—inside the helix, causing a bend of approximately 30° along the helical axis.

However, proline is often the first residue in a helix, presumably due to its structural rigidity.

On the other hand, glycine also tends to collapse helices because its high conformational flexibility makes it entropically expensive to adopt a relatively constrained α-helical structure.



Proline



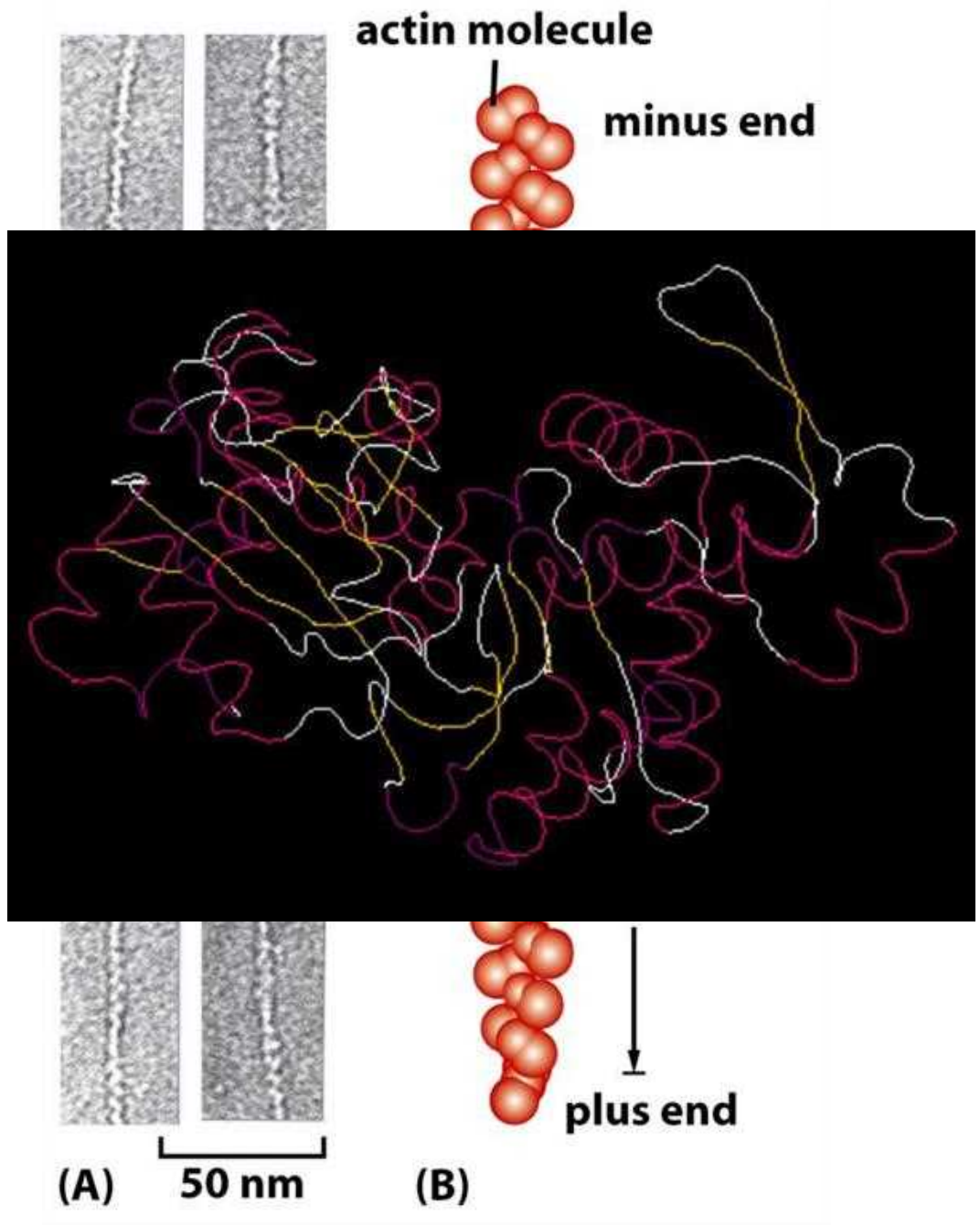
Glycine

The **actin helix** is fairly straight

- However, in many cases, proline causes a bend in the helix; this often aids in the packing of long helices in both cytosolic and transmembrane proteins.

- The proline residue lacks an amide proton. This precludes the formation of hydrogen bonds between it and hydrogen bond acceptors, and thus often limits the proline residue to the first four positions of the α-helix. Proline has "helical" dihedral angles in the backbone, which help initiate helical folding.

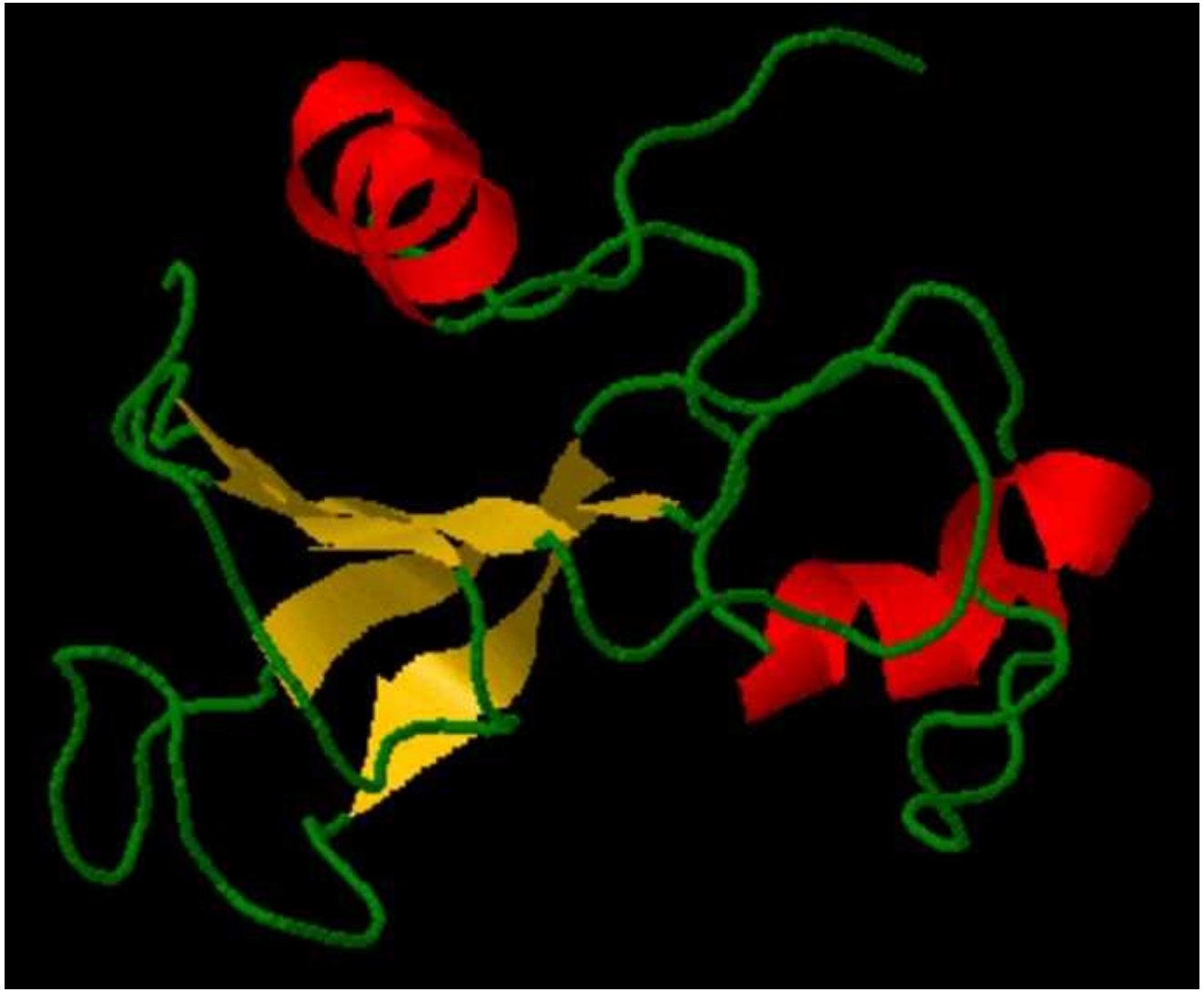
- The presence of proline within the α-helix causes a bend in the helical axis. Helices with a proline bend are almost always long (>4 turns); this often aids in the packing of long helices within the protein.



α-helix

Amphipathic α-helices (an α-helix with hydrophobic and hydrophilic amino acid residues) can be found on the surface of a water-soluble globular protein, while hydrophobic helices are found internally.

When soluble proteins fold in water, the most stable conformation should be one in which the maximum number of polar groups are on the surface and in contact with water, while the maximum number of nonpolar side chains are buried and away from the surface.



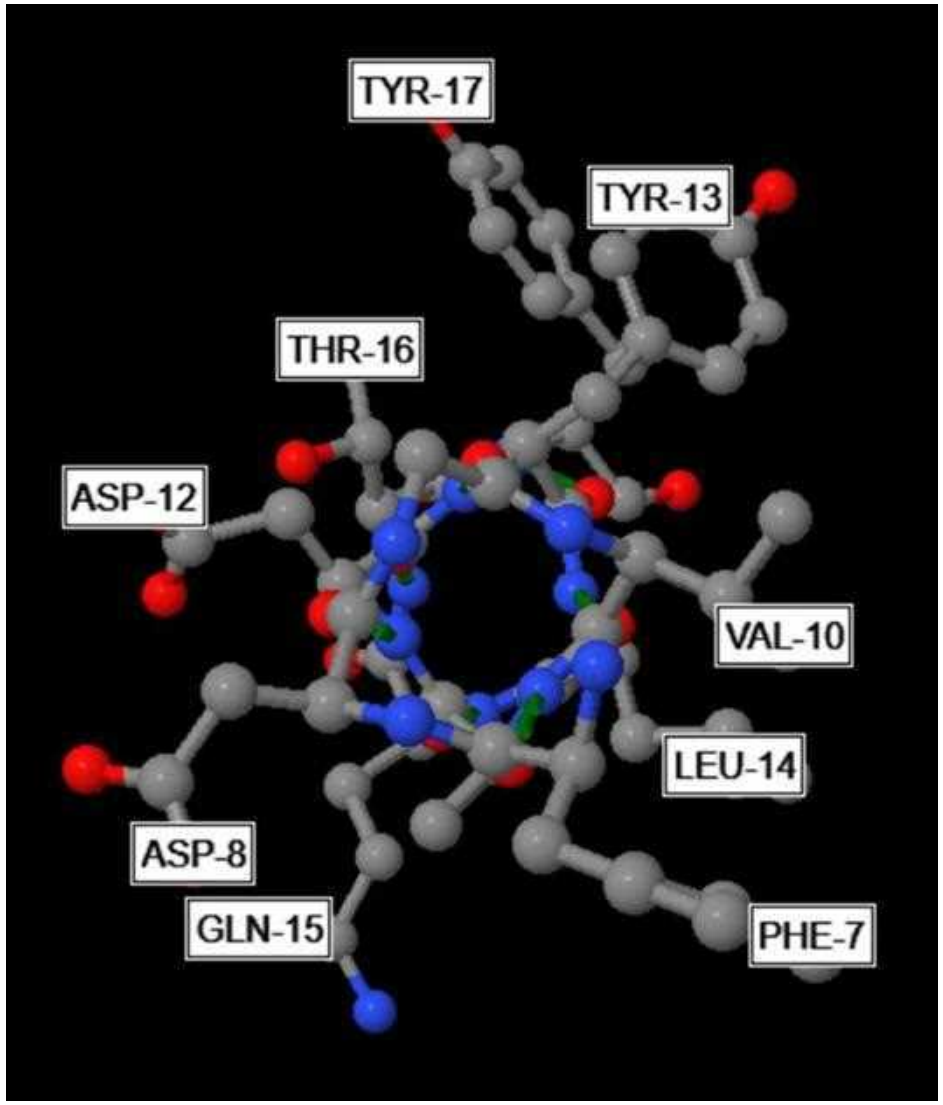
Model of barnase, a 110-residue extracellular ribonuclease from Bacillus amyloliquefaciens

Note that all three α-helices are on the surface.

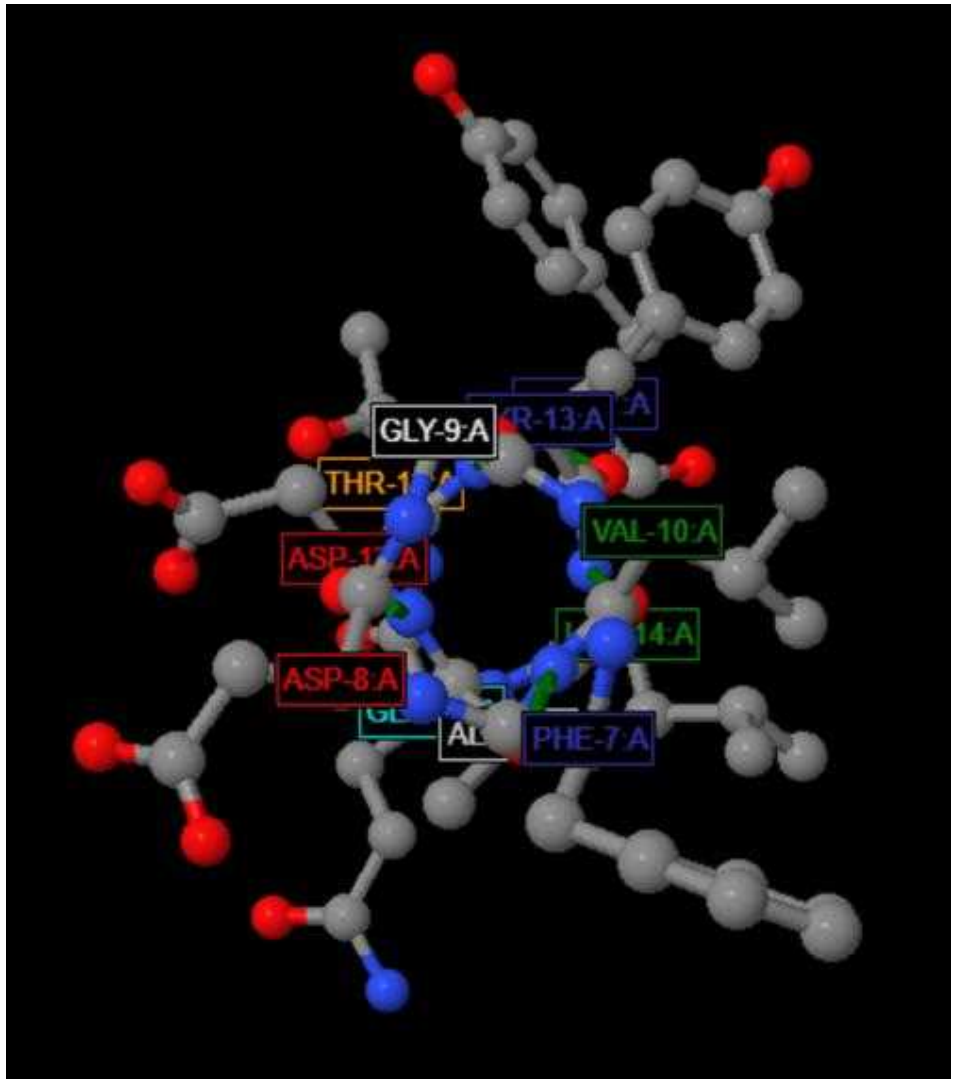
a-helix

Model of barnase

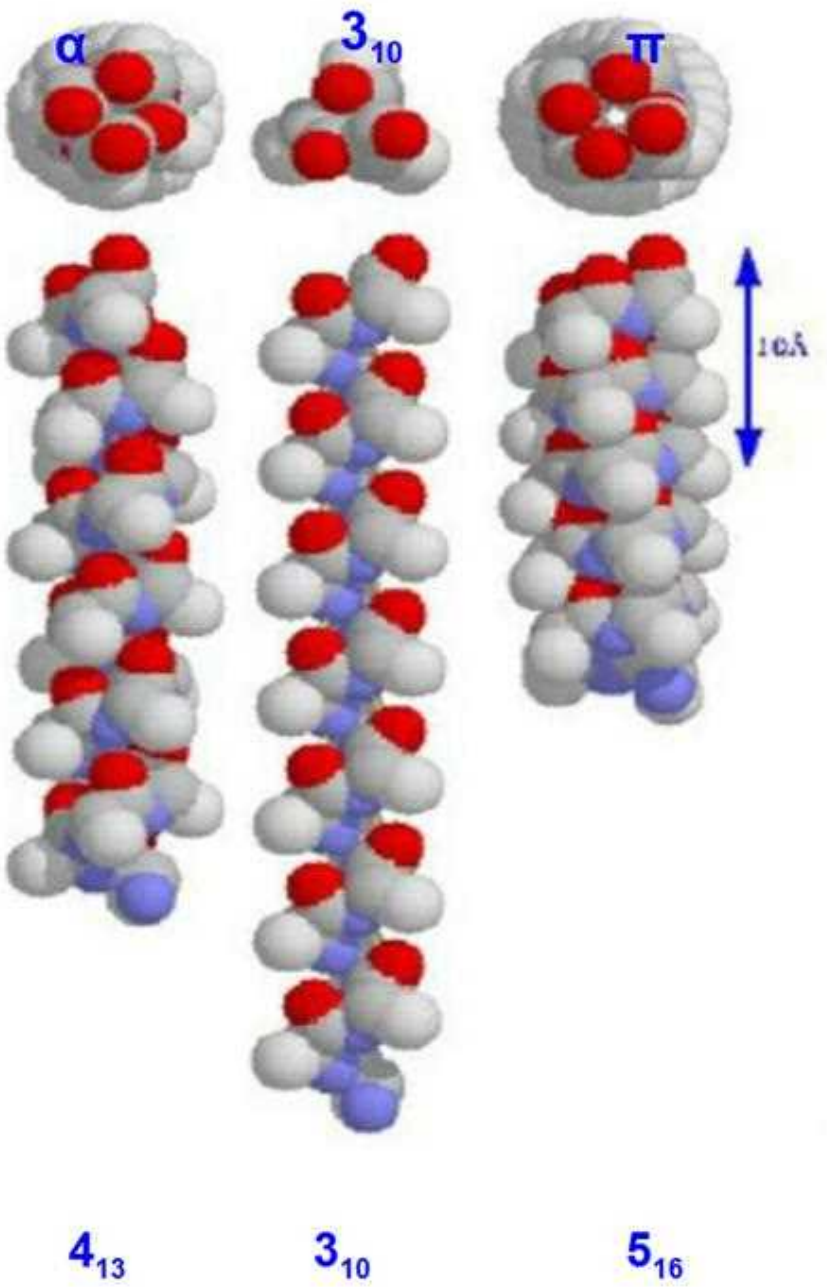
This is a typical amphipathic helix



The hydrophobic side chains of Phe-7, Val-10, and Leu-14, as well as the hydrophobic moiety of Tyr-13, are located on one side of the helix.



This side is packed opposite the hydrophobic side chains in the nonpolar interior of the protein. The other side of the helix, which is exposed to water, contains the hydrophilic side chains.



Other types of alpha-helices

The 3_{10}-helix is more extended; the axial rise per residue is 0.2 nm. The number of residues per turn is 3.0. The hydrogen bond length is 0.183 nm.

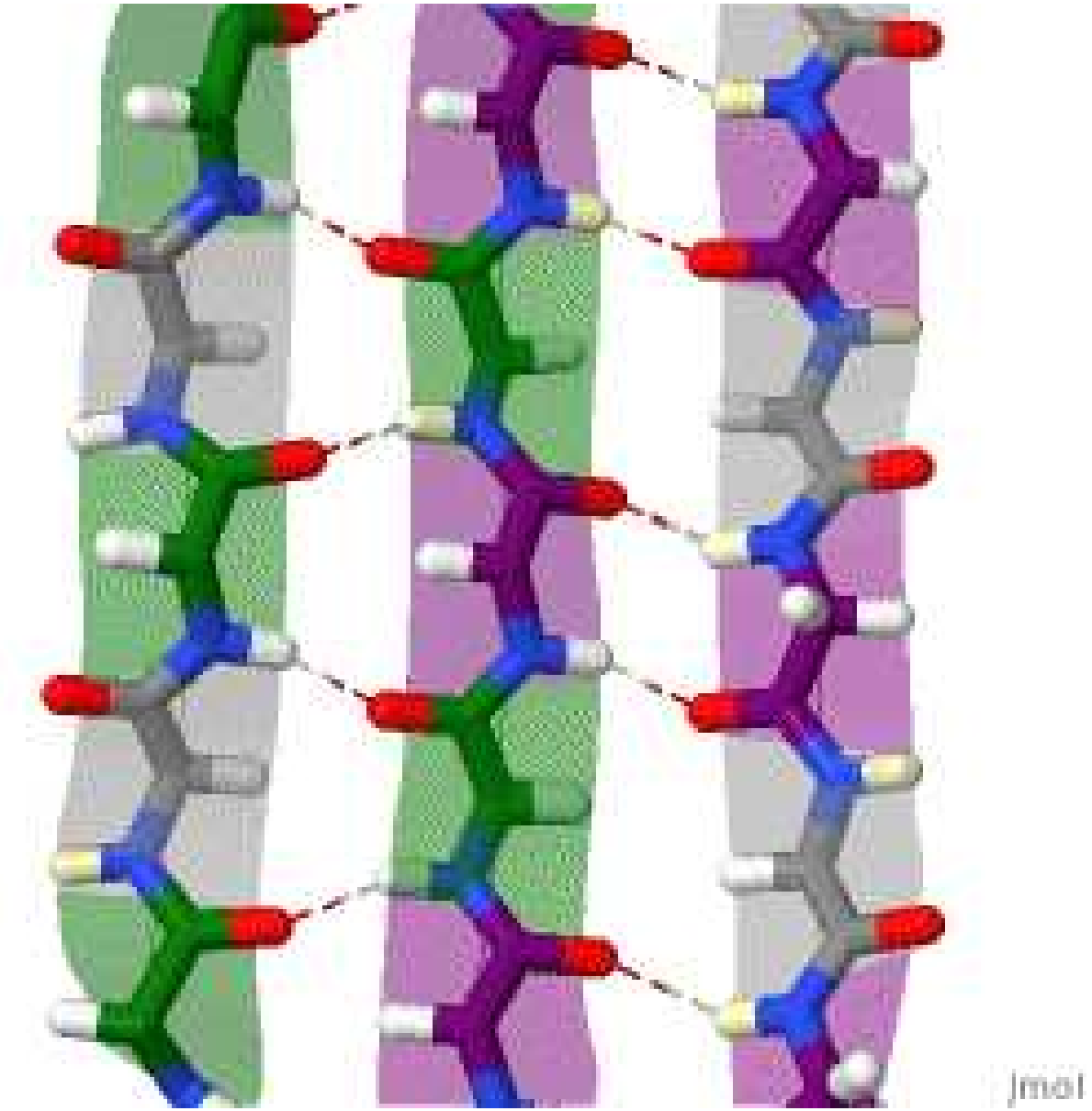
The π-helix is more compact; the axial rise per residue is 0.11 nm. The number of residues per turn is 4.3. The hydrogen bond length is 0.165 nm.

β-sheet

The β-sheet structure (also known as a β-structure or β-pleated sheet) is a common secondary structure motif in conventional proteins.

B-sheets are composed of beta strands (β-strands) linked laterally by at least two or three hydrogen bonds to the backbone, forming an overall twisted, folded sheet.

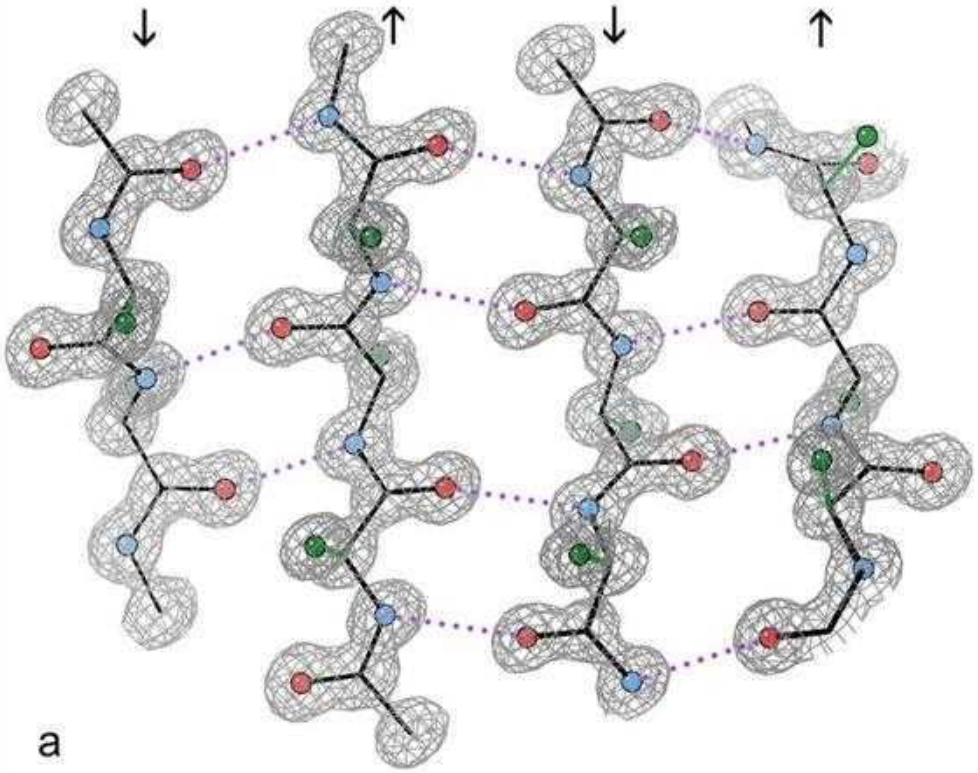
A β-strand is a section of a polypeptide chain, typically 3 to 10 amino acids long, with the backbone in an extended conformation.

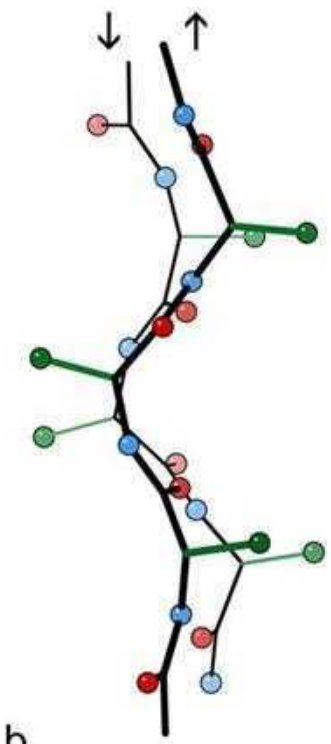


Three-dimensional structure of parts of the β–sheet

β-sheet

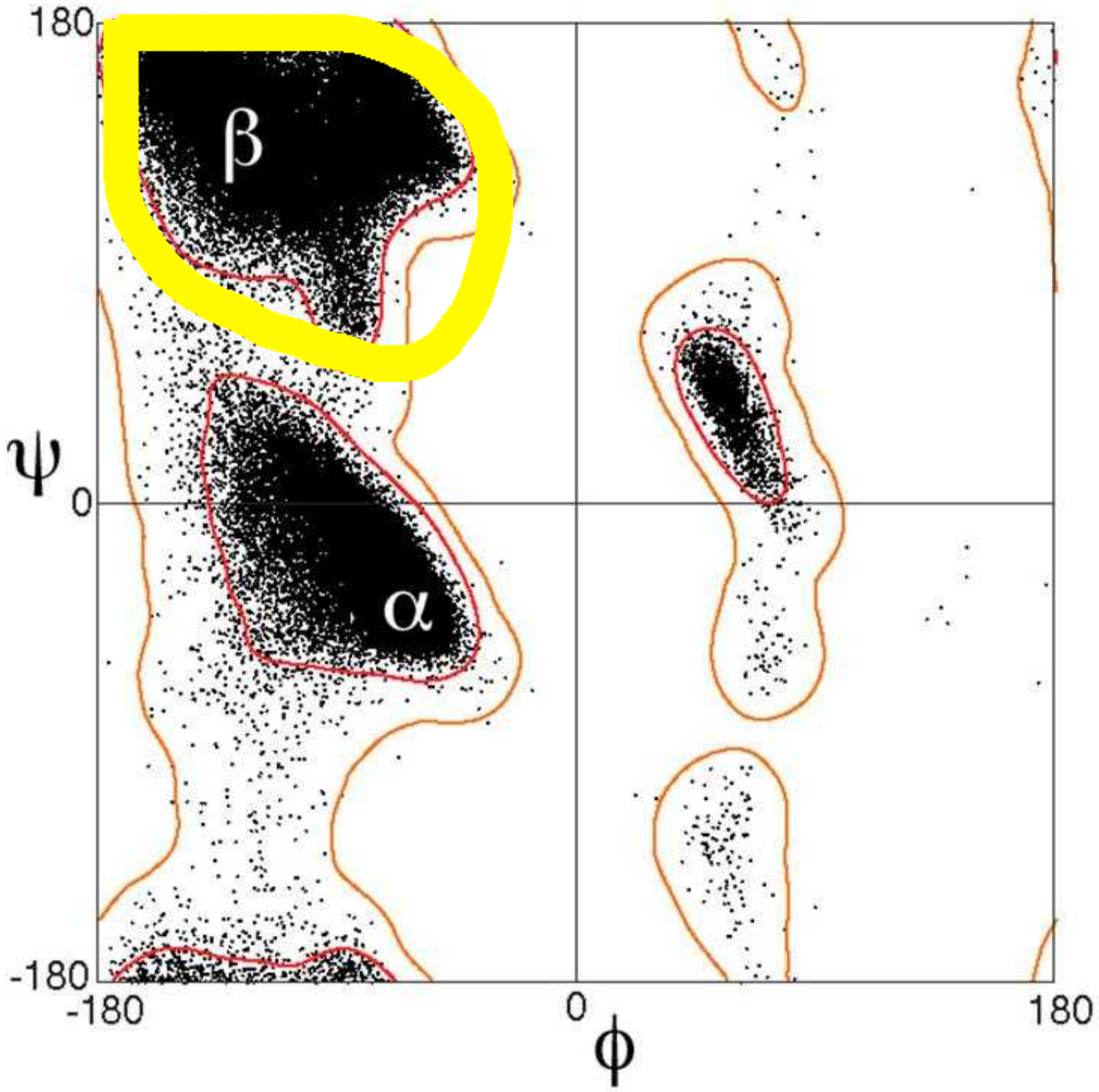
Most β-sheet are located next to other strands and form an extensive hydrogen bond network with their neighbors, in which N−H groups in the backbone of one strand establish hydrogen bonds with C=O groups in the backbone of adjacent strands.





Linus Pauling and Robert Corey

The "puckered" appearance of the β-sheet arises from the tetrahedral chemical bonding at the Cα atom; for example, if the side chain points straight up, then the bonds to C′ must point slightly downwards, since its bond angle is approximately 109.5°.



β-sheet

However, β-strands are rarely perfectly extended; rather, they exhibit twisting. Energetically preferred dihedral angles near (φ, ψ) = (–135°, 135°) (broadly speaking, the upper left region of the Ramachandran plot) deviate significantly from the fully extended conformation (φ, ψ) = (–180°, 180°).

β-sheet

Because peptide chains have a directionality imparted by their N-terminus and C-terminus, β-strands can also be called directional. They are typically represented in protein topology diagrams by an arrow pointing toward the C-terminus. Adjacent β-strands can form hydrogen bonds in:

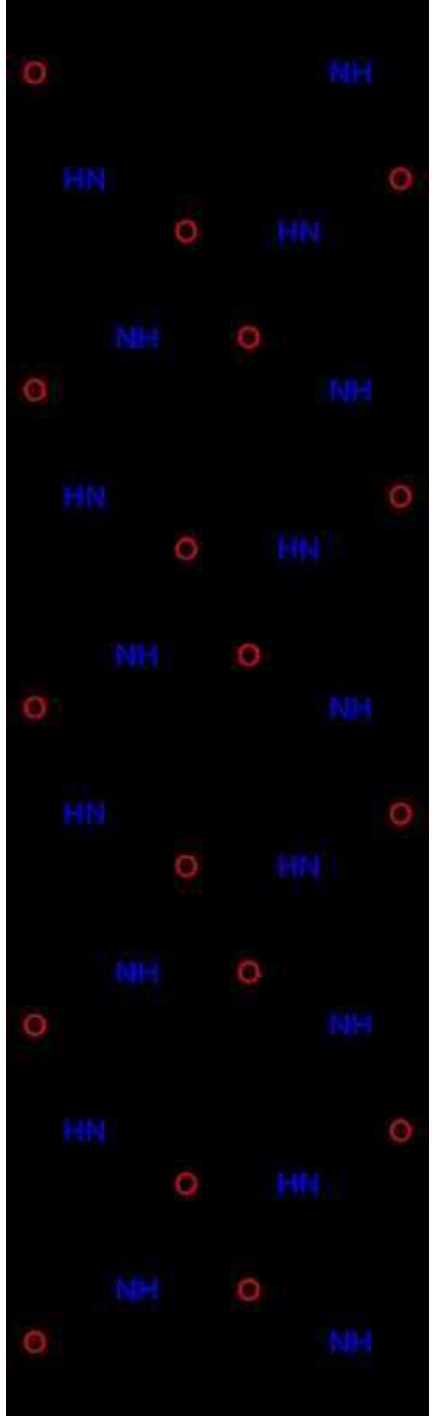
- *Antiparallel*
- *Parallel*
- *Mixed β-sheet*

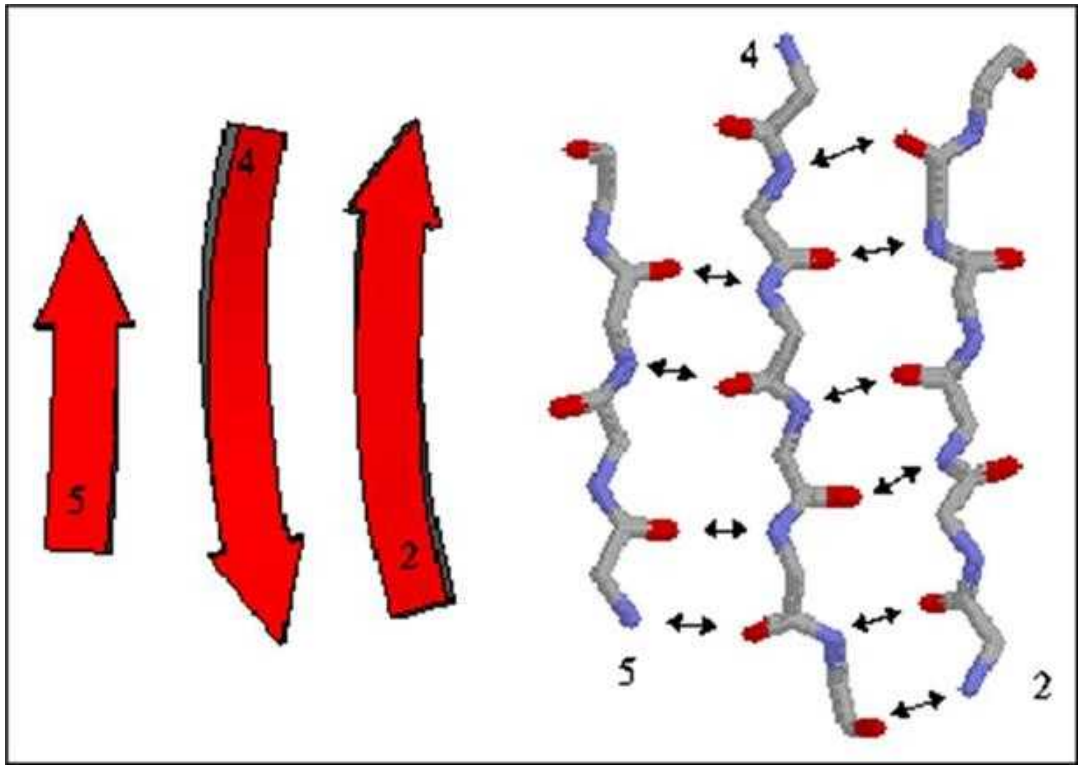
β-sheet

Antiparallel

In an antiparallel arrangement, successive β-strands alternate directions so that the N-terminus of one strand is adjacent to the C-terminus of the next.

This arrangement provides the strongest interchain stability, as it allows interchain hydrogen bonds between carbonyls and amines to be planar, which is their preferred orientation.

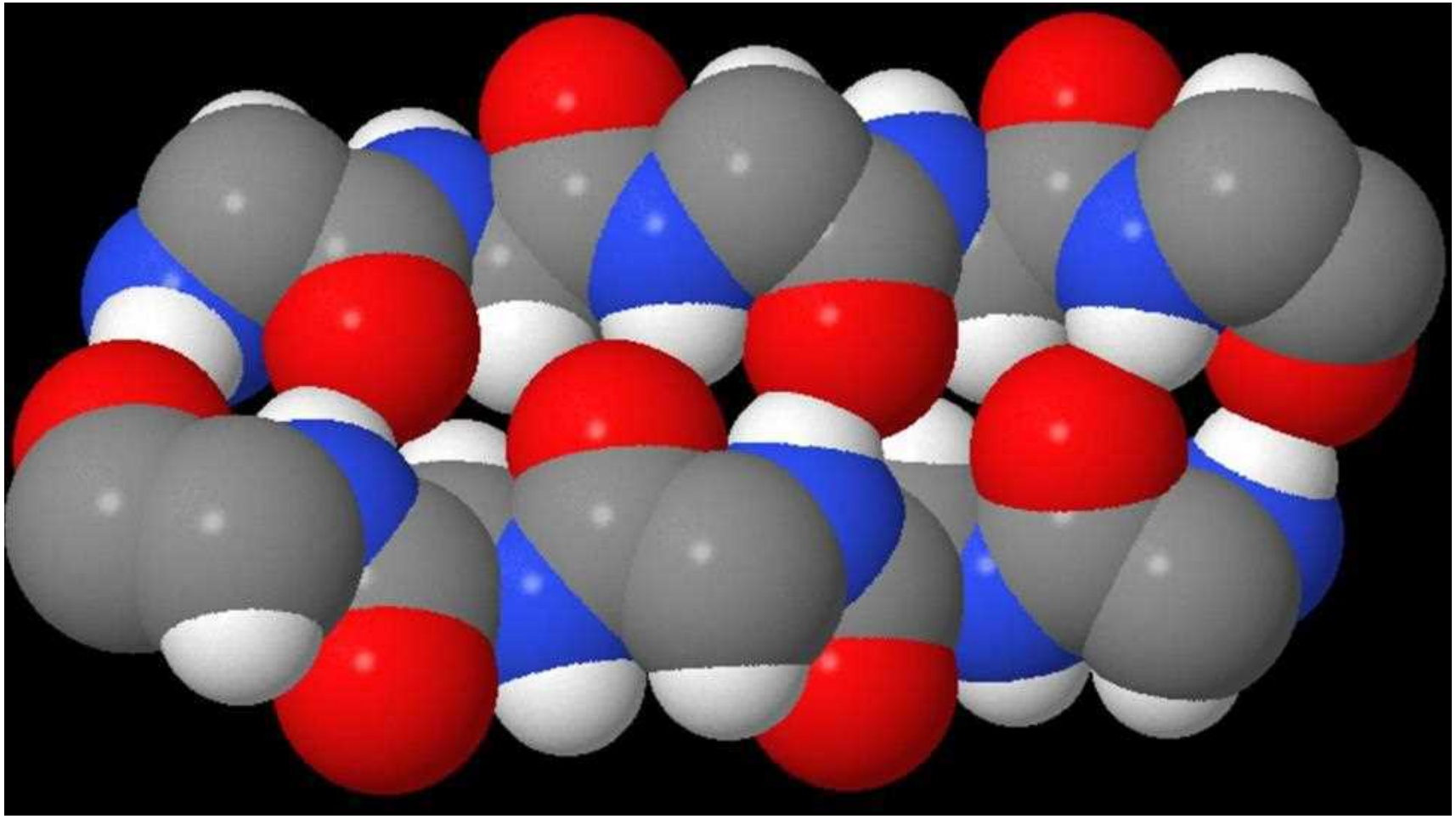




Hydrogen bonds are located perpendicular to the amino acid chain and parallel to each other

Antiparallel β-sheets

In a β-sheet, polypeptide chains are linked regularly by hydrogen bonds between the C=O and N-H groups of the backbone. Let's start with two β-sheet

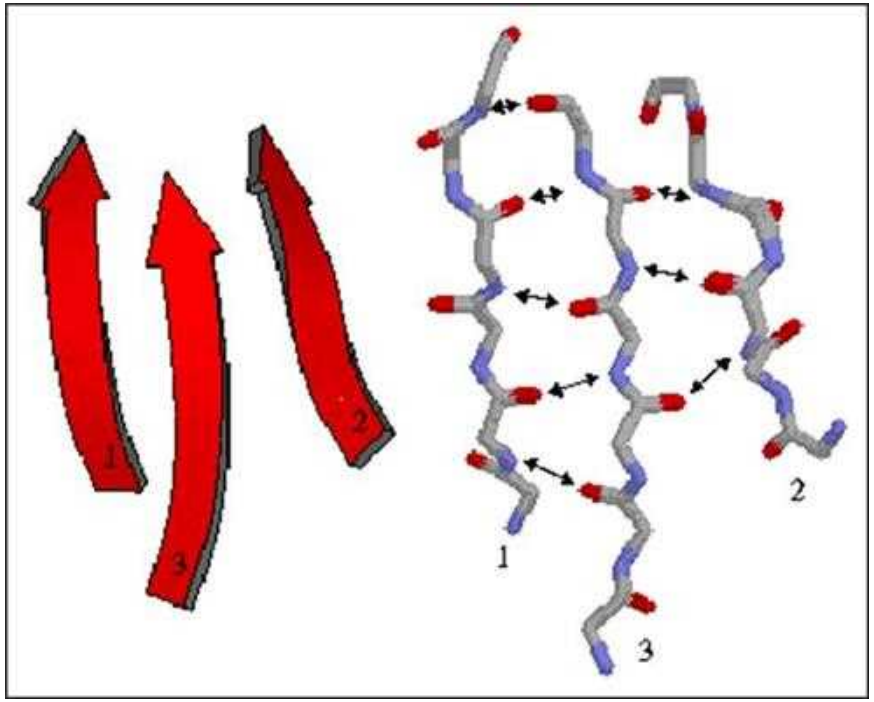


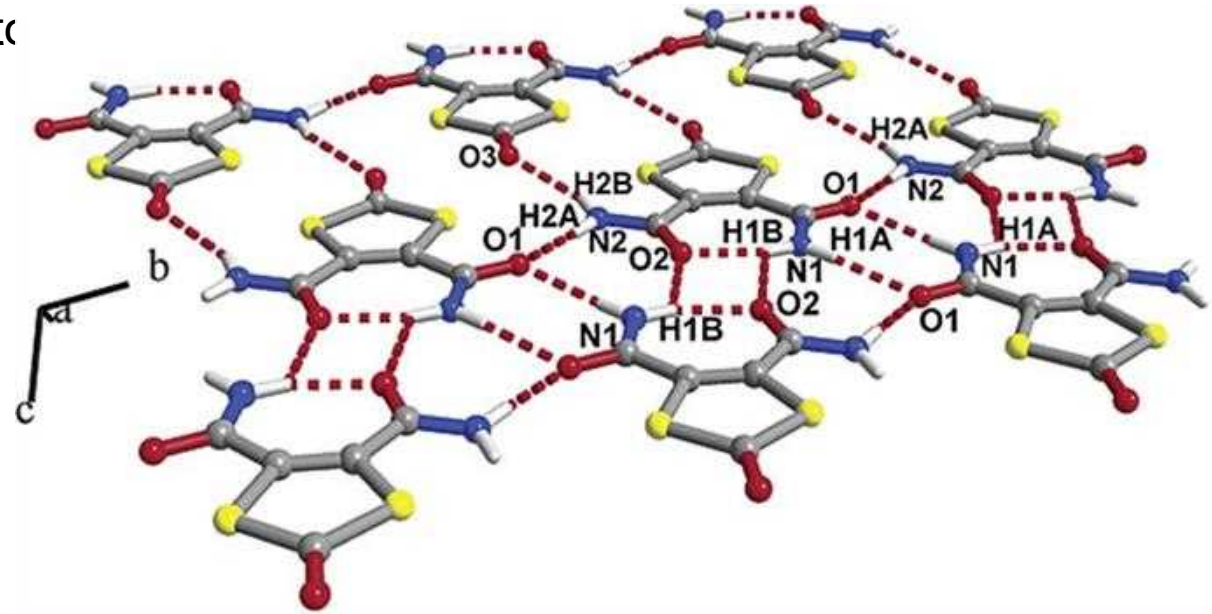
Can you identify the locations of additional hydrogen bonds between the two β-sheet? Find five more.

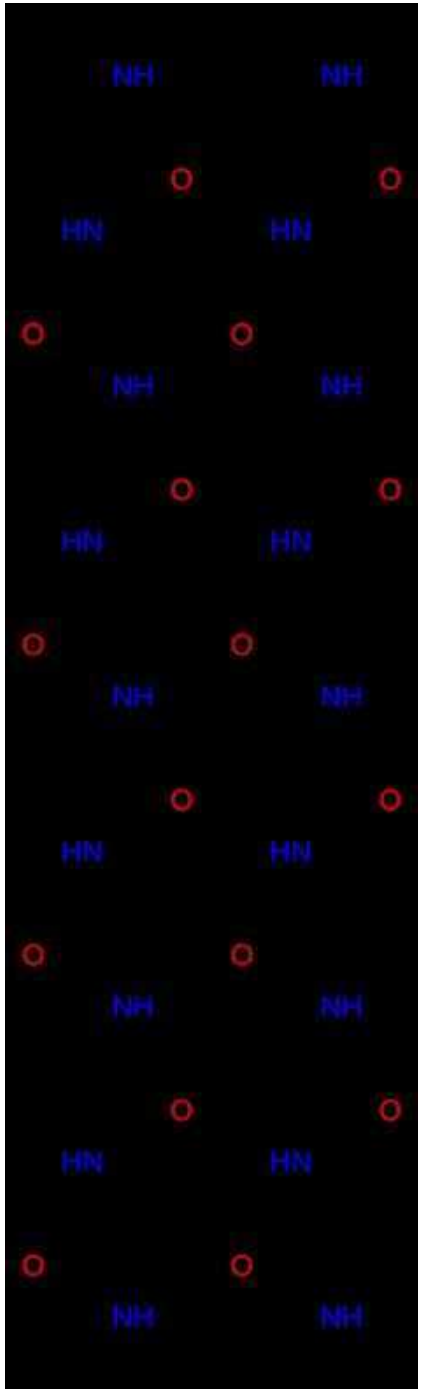
Parallel β-sheet

Parallel b-sheet – all linked strands have the same N to C direction. As a result, they must be separated by long stretches of sequence. Hydrogen bonds are equally spaced.

The arrangement of hydrogen bonds in the parallel b-sheet resembles that in the 11-atc amide ring motif.



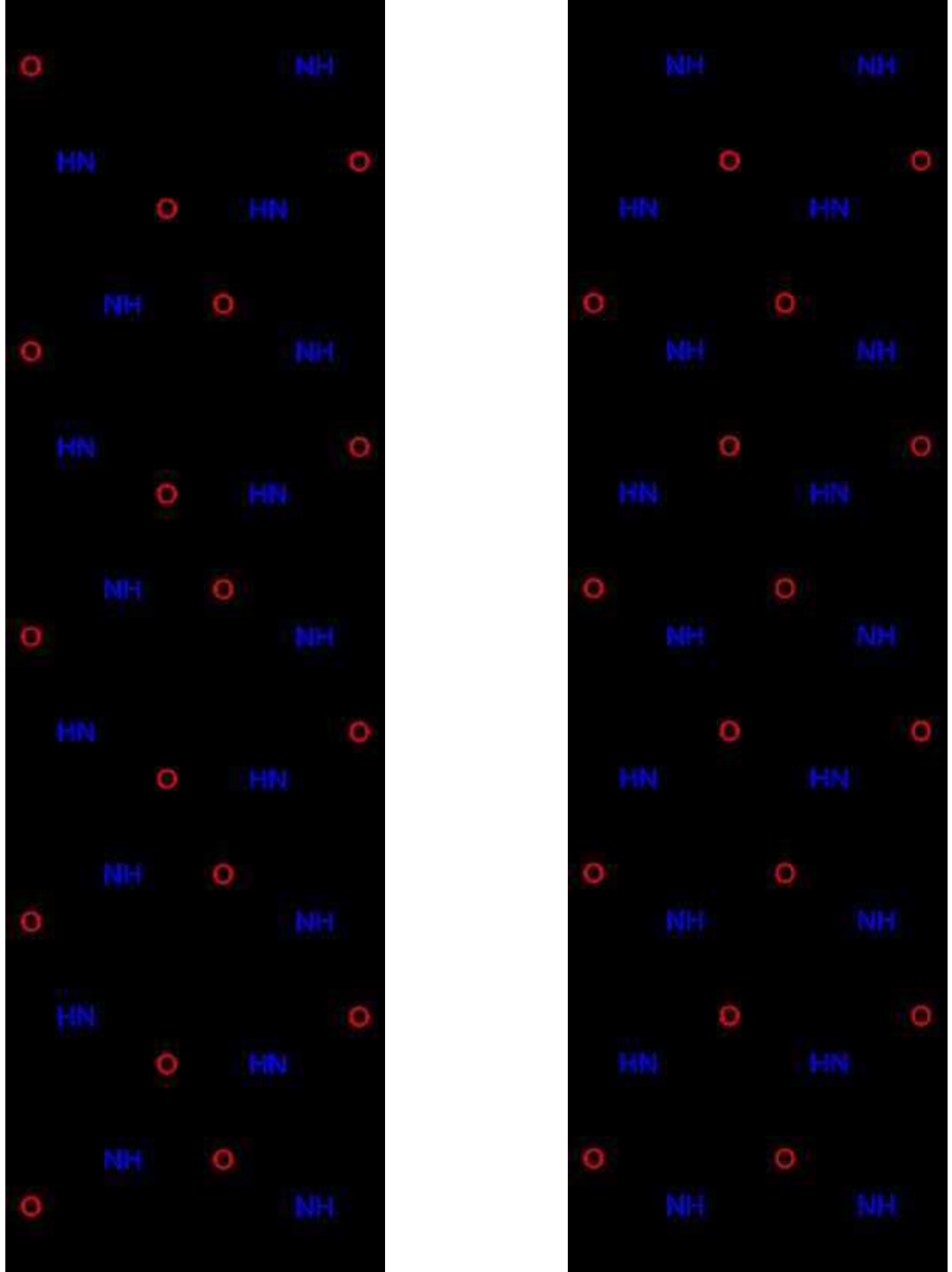




β–sheet

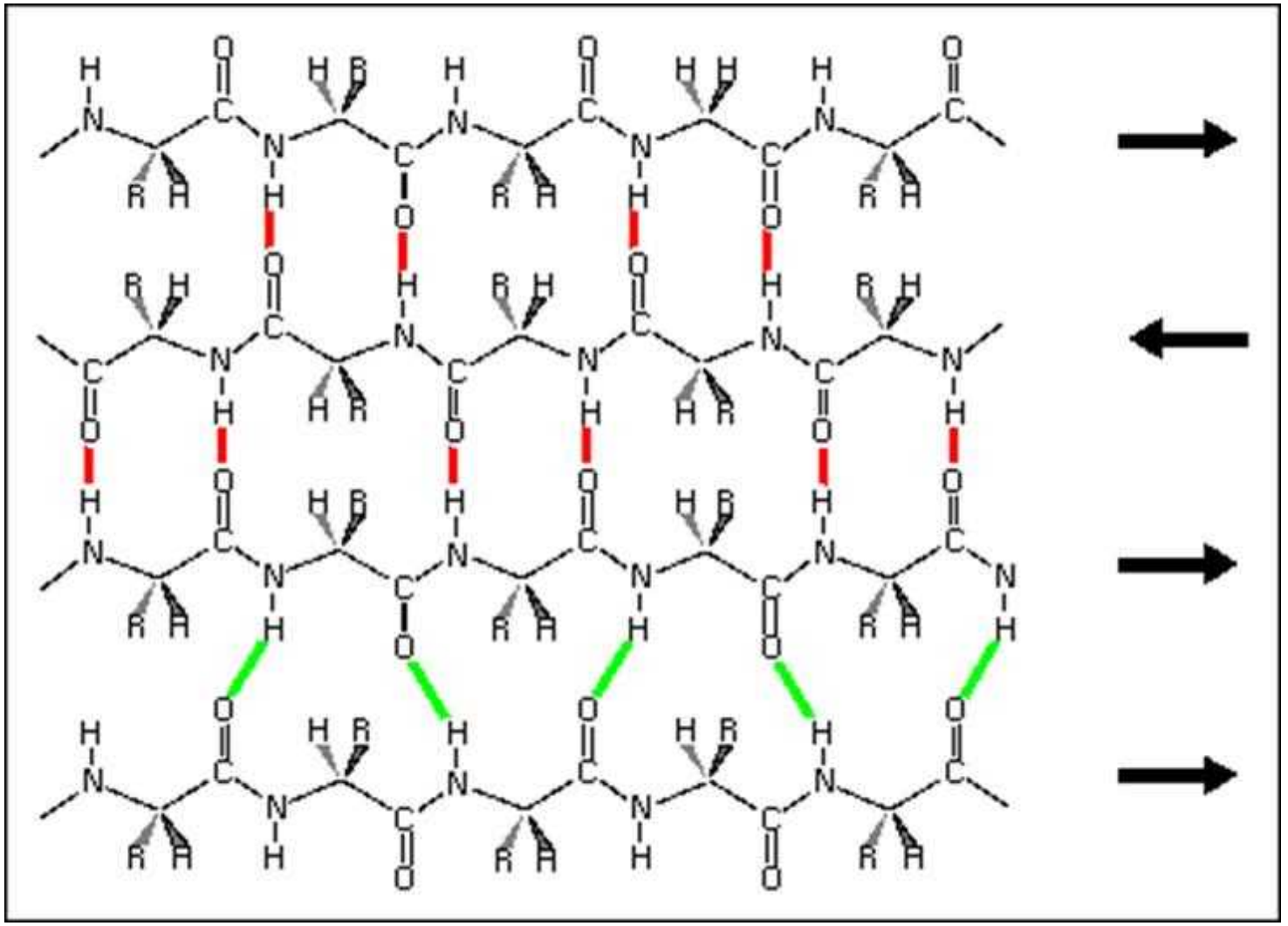
The hydrogen bonds are parallel to each other in an antiparallel β-sheet;

for a parallel β-sheet (the strands are oriented in the same direction), the hydrogen bonds are not parallel.



Mixed β-sheet

An individual strand may exhibit a mixed binding pattern, with a parallel strand on one side and an antiparallel strand on the other. Such arrangements are less common



Hydrogen bond patterns in a mixed b-sheet. This diagram shows a schematic of a four-stranded beta sheet containing three antiparallel strands and one parallel strand. Hydrogen bonds between antiparallel strands are shown as red lines, and between parallel strands as green lines.