2. Proteins have Hierarchies of Structure

Protein structure is usually described at four different levels (*Fig. II.2.1*). The first level, called the *primary structure,* describes the linear sequence of the amino acids in the chain. The different primary structures correspond to the different sequences in which the amino acids are covalently linked together. The *secondary structure* describes two common patterns of structural repetition in proteins: the coiling up into *helices* of segments of the chain, and the pairing together of *strands* of the chain into β-*sheets*. The *tertiary structure is* the next higher level of organization, the overall arrangement of secondary structural elements. The *quaternary structure* describes how different polypeptide chains are assembled into complexes.

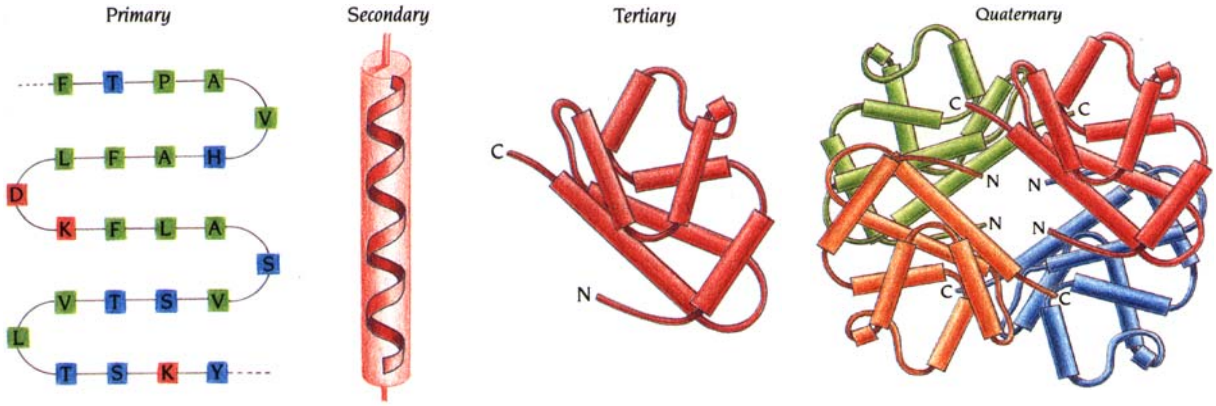

Figure II.2.1. *Different levels of protein structure. A protein chain's primary structure is its amino acid sequence. Secondary structure consists of the regular organization of helices and sheets. The example shown is a schematic representation of an α helix. Tertiary structure is a polypeptide chain's three-dimensional native conformation, which often involves compact packing of secondary structure elements. The example given in the figure is a schematic drawing of one of the four polypeptide chains (subunits) of hemoglobin, the protein that transports oxygen in the blood. α-helices along the chain are represented as cylinders. N is the amino terminus and C is the carboxyl terminus of the polypeptide chain. Quaternary structure is the arrangement of multiple polypeptide chains (subunits) to form a functional biomolecular structure. The figure shows the quaternary arrangement of four subunits to form the functional hemoglobin molecule. By definition, proteins that function as single chains and do not form multimers have no quaternary structure. (adapted from {Branden & Tooze 1991 ID: 507})*

a. Secondary Structures: Helices and sheets are common motifs of proteins

α-Helices. One natural conformation for a polymer molecule, or a piece of rope, or any other flexible one-dimensional string-like entity is a helix. If the elementary units of a chain have the same fixed angle between every sequential pair of monomers, it defines a repeating pattern; simple geometry dictates that it will be a helix (which is three-dimensional), or a planar zig-zag (which is two-dimensional), or, if the angle is zero, a straight line (which is one-dimensional). Many *homopolymers* have a single favored repeat angle (at low temperatures) and so they crystallize into helical or zig-zag structures. Of the 176 crystal structures of polymers known in 1979 {Tadokoro 1979 ID:

508}, 79 were helices of 22 different types and 49 form in-plane zigzags. The helical pitch is dictated by the physical structure of the monomer unit.

Polypeptides also form helices. The most common type of helix in proteins is called the α-helix. Discovered by Linus Pauling and his colleagues in the early l950s, the α-helix was unexpected at the time because proteins were then believed to have a very high degree of symmetry, yet the number of monomer units per helical turn predicted in the Pauling α-helix is not an integer. The defining feature of the α-helix is the backbone hydrogen bonds formed between the carbonyl oxygen of amino acid i and the amide hydrogen of amino acid i + 4 (see *Figure II.2.2*). Therefore, the α-helix has 3.6 amino acids per turn. α-helices can be formed by any of the amino acids because the hydrogen bonds are among backbone atoms, not side chains, and therefore helices are only weakly affected by side chain type, with the exception of proline that is missing a proton on its backbone N atom. Most helices in proteins are relatively short.

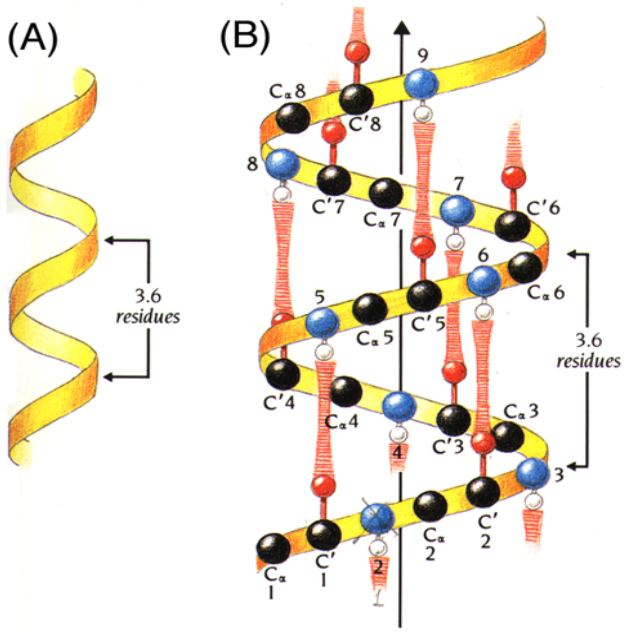

Figure II.2.2. *The α-helix conformation in proteins. (A) idealized drawing of the path of the main chain in an α-helix, with 3.6 residues per turn, which corresponds to 5.4 Å distance between successive turns (or 1.5 Å per residue). (B) includes the approximate positions of the main chain atoms. The atoms are labeled by the residue they belong to, starting from the N-terminus. Black balls are carbon atoms, blue ones along the backbone are nitrogens, and those red ones appended to the backbone carbon atoms are the oxygen atoms. Amide hydrogens are shown by white balls. Side chains are not shown. Spring-like connections between the carbonyl O of residue i and the amide H of residue i+4 indicate the (i, i+4) hydrogen bonds. (from {Branden & Tooze 1991 ID: 507})*

There are four reasons for the stability of an α-helix: the hydrogen bonds, the ready accessibility to the helical (*ϕ, ψ)* angles - in the α-helix (see Ramachandran plots in § II.1); the side chains do not interfere with the backbone; the favorable van der Waals interactions inside the helix due to the small hole down the helix axis; and the good alignment of the electrical dipole moments of the amino acids parallel to the helix axis. *Figure II.2.3* shows that the side chains are on the outside of the helix. Like the individual amino acids, helices have chirality. Almost all helices in proteins are *right-handed.* In this way, the chain avoids steric conflicts between carbonyl groups and the side chains of *L*-amino acids, as discussed in § II.1.c. But, there is an effect of the helix on side chain conformations (see *Fig. II.1.18),* where a gauche$^-$ state for χ_1 is excluded by steric conflicts with the helical backbone. If proteins were made of D-amino acids, then helices would be left-handed.

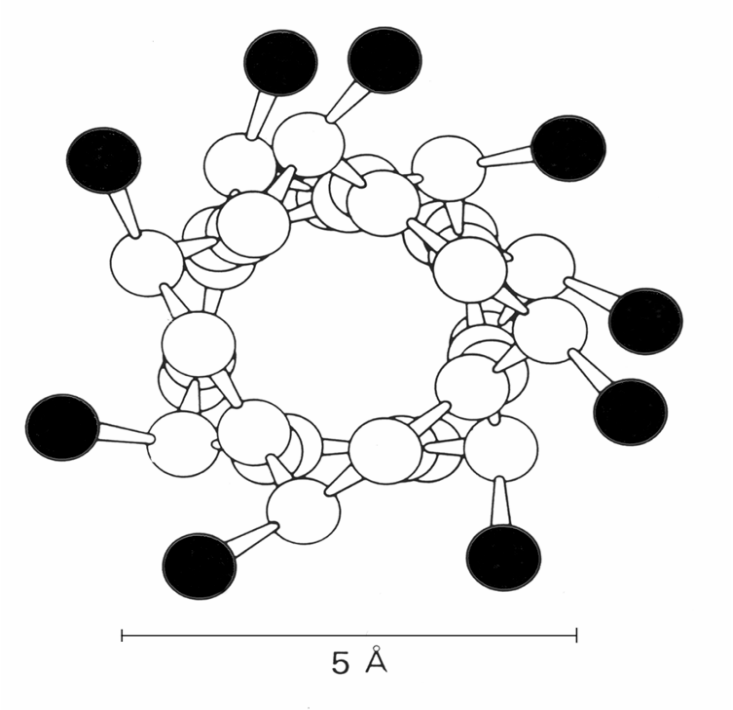

Figure II.2.3. *Cross-sectional view of an a-helix. The sidechains (shaded spheres) are on the outside. Note that the van der Waals radii of the atoms are larger than shown in the ball-and-stick model here. There is almost no free space inside of an α–helix. (from {Stryer 1988 ID: 509})*

Table II.2.1.
Approximate Geometric Parameters for Some Regular Protein Conformations

Secondary structure	Dihedral Angle ϕ	ψ	ω	Residues[a] per turn & chirality	Translation per residue (Å)
Right-handed α-helix	-57	-47	+180	+3.6	1.50
Left-handed α--helix	+57	+47	+180	-3.6	1.50
3_{10}-helix	-49	-26	+180	+3.0	2.0
α-helix	-57	-70	+180	+4.4	1.15
Antiparallel β-sheet[b]	-139	+135	-178	±2.0	3.4
Parallel β-sheet[c]	-119	+113	+180	±2.0	3.2
Twisted antiparallel or parallel β-sheet[d]	≈-110	≈+140	≈+180	≈-2.4	≈3.3
Fully extended chain[e]	-180	+180	+180	±2.0	3.6

Partly adapted from *Table 2* of {IUPAC-IUB Commission on Biochemical Nomenclature. 1970. Abbreviations and symbols for the description of the conformation of polypeptide chains. Tentative rules (1969) [published simultaneously in Biochemistry 9: 3271-3479, J. Biol. Chem. 24: 6489-6497; J. Mol. Biol. 52: 1-17]} and *Table 5.1* of {Schulz & Schirmer 1979 ID: 513}.

[a] + and – correspond, respectively to situations when successive C^{α}-C^{α} virtual bonds follow a right-handed and a left-handed helical path, ± is used when such helices become planar, i.e. when the C^{α}-C^{α} virtual bonds follow a planar zigzag form.

[b, c] Pleated sheets with Pauling-Corey idealized geometries, rare in proteins.

[d] Twisted β -sheets are abundant in proteins. There are considerable variations among the dihedral angles of twisted β-strands. The values given here only roughly indicate the location of the center of the β -sheet region on the (ϕ,ψ)-map; see *Figure II.1.9*.

[e] Included for reference only. Fully extended chains are not part of secondary structures. They do not form stable sheet-like chain organizations.

A *helical wheel diagram is* a projection of a helix onto a plane perpendicular to its axis. It shows the periodicity of amino acids around the helix (*Figure II.2.5)*. What is often observed in such diagrams is that for helices in globular proteins, the hydrophobic amino acids tend to cluster on one side of the helix (pointing toward the interior of the protein), and the polar and charged amino acids are on the outer face (*Figure II.2.6*). These are called *amphipathic* helices, and are said to have a *hydrophobic moment* {Eisenberg & McLachlan 1986 ID: 512}.

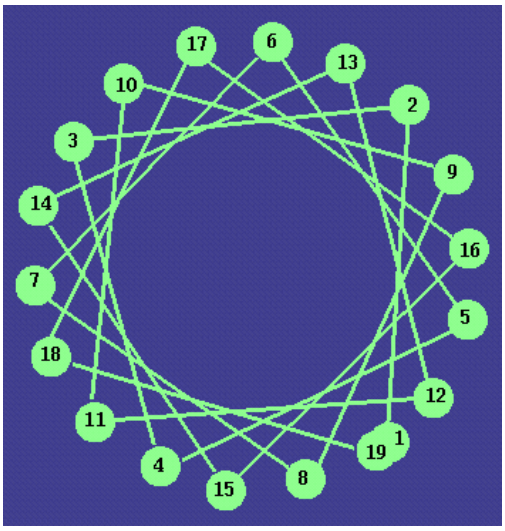

Figure II.2.5. *Helical wheel diagram for an α-helix. Residues are assigned indices conforming to their positions along the chain. In conformity with a right-handed twist, a counterclockwise rotation is observed as one proceeds from the N-terminus to the C-terminus. There are 3.6 residues per turn such that the angular separation between successive residues is 100°. A cycle is completed by five turns (or 18 residues), such that residue 19 eclipses residue 1.*

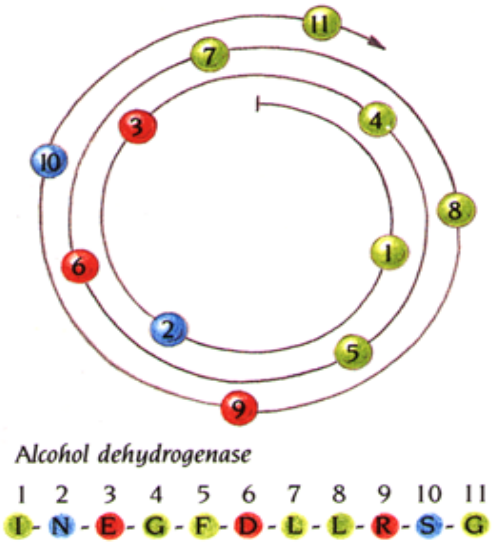

Figure II.2.6. *Helical wheel diagram of an amphipathic helix in the enzyme alcohol dehydrogenase. Hydrophobic and polar/charged sidechains are represented by green and red or blue balls, respectively. (from {Branden & Tooze 1991 ID: 507})*

β-sheets. When Pauling and his colleagues predicted the α-helix, they also predicted that hydrogen bonding would lead to parallel and anti-parallel traintrack-like structures they

called “β-sheets" (see *Figs. II.2.7* and *8*). A β-sheet is comprised of individual *strands* each of which is a nearly planar zigzag. Strands pair through amide-to-carbonyl hydrogen bond links. The side chains lie above or below the sheet, and they are well - placed to interact with neighboring side chains *(Figure II.2.9)*. β-sheets are stabilized by hydrogen bonds, by their side chains’ interactions, by favorable (ϕ, ψ) angles (in the β-region of the Ramachandran map), and by van der Waals attractions. *Figure II.2.10* shows that large β-sheets are not planar; but actually have a twist. The regularity of β-sheets can sometimes be interrupted by a ‘β-bulge’. An example is shown in *Figure II.2.11*.

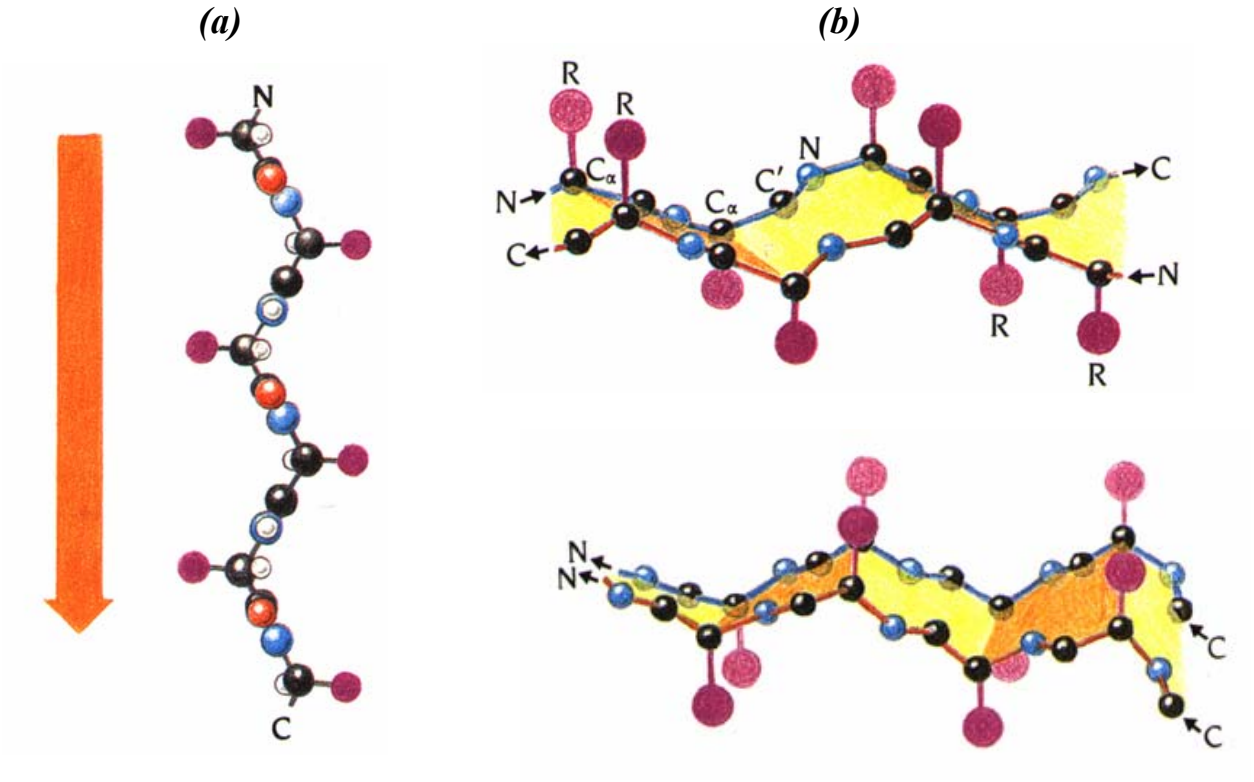

Figure II.2.7. *Schematic illustration of β-sheets.* ***(a)*** *The extended conformation of a β-strand in a ball-and-stick model. The arrow indicates the direction from the N-terminus to the C-terminus. Sidechains are shown as purple spheres, N atoms are colored blue, and carbonyl carbons, red.* ***(b)*** *Illustration of the pleat of β-sheets for two antiparallel strands (top) and two parallel strands (bottom). (adapted from Figures 2.5(a), 2.5(d) and 2.6(c) of {Branden & Tooze 1999 ID: 507}*

(a) *(b)*

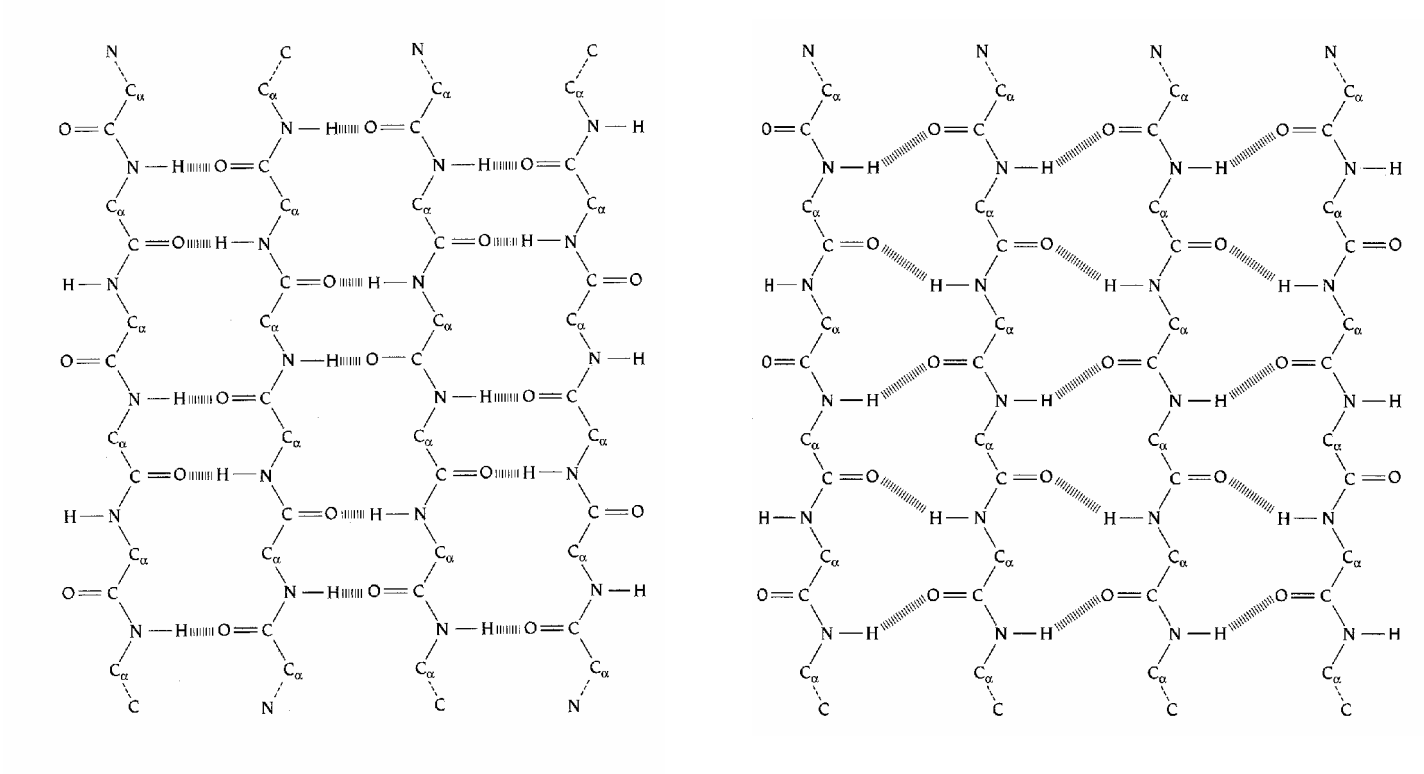

Figure II.2.8. *Hydrogen bonding pattern in (a) antiparallel and (b) parallel β-sheets. (adapted from Figures 2.5(b) and 2.6(b) of {Branden & Tooze 1999 ID: 507}.*

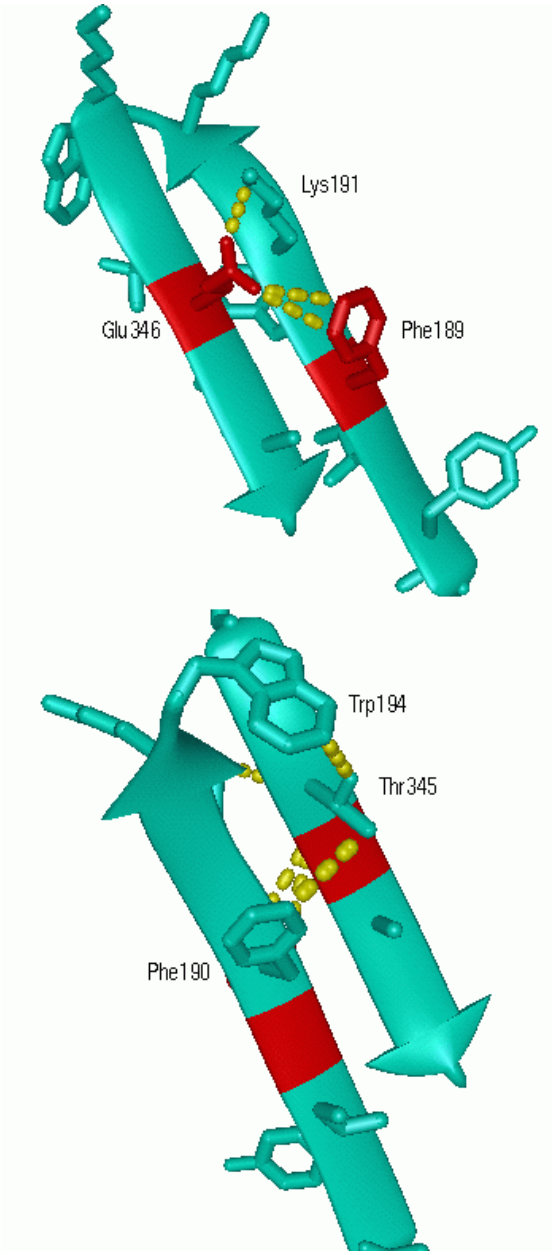

Figure II.2.9. *Interactions between sidechains belonging to adjacent β-strands contribute to the stabilization of β-sheets. The upper and lower diagrams illustrate the interactions on the opposite faces of the β-strand. Note that residue i on a given strand can simultaneously interact with residues j and j+2 (or j+4) on the adjacent strand. (diagram taken from PDB structure 9api)*

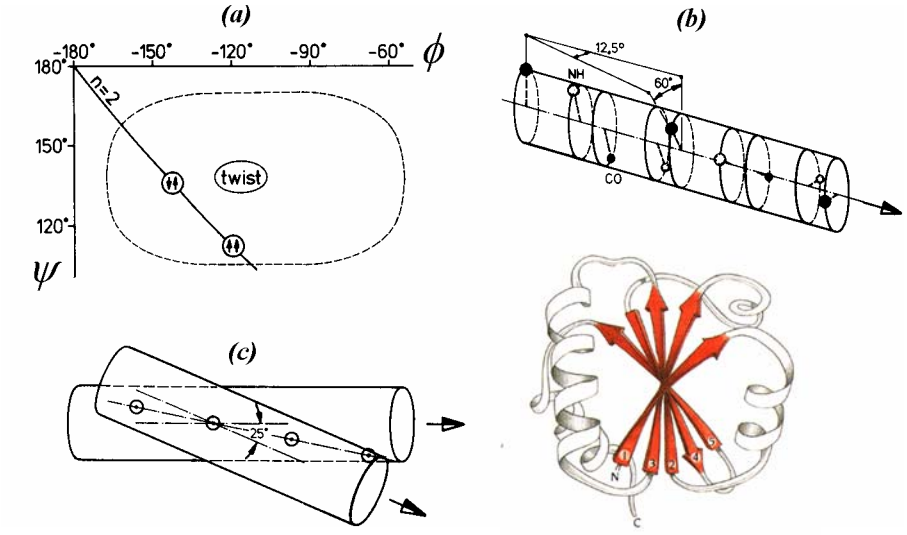

Figure II.2.10. *Twist of β-pleated sheets.* ***(a)*** *Region of (ϕ,ψ)-map corresponding to the β-sheet region (region II in Figure II.1.7(a)). The diagonal indicates the loci of dihedral angles in planar zigzag (2-fold helical) structures. The dihedral angle positions of the ideal parallel (↑↑) and antiparallel (↑↓) β-sheets are indicated. The dashed contour encloses the allowed β-region (having energy lower than 1 kcal/mol). Note that the center of this region (label "twist") is to the right of the diagonal. This corresponds to (ϕ, ψ) values of typical strands in twisted sheets.* ***(b)*** *Right-handed twist along a single strand in a twisted sheet. The rotations of the hydrogen-bonding directions of the carbonyl (filled circles) and amide (open circles) groups are indicated. Twist angles of β-strands vary considerably. Angles given here are only typical values.* ***(c)*** *Shows how two parallel strands twisted as sketched in part (b) can pair to form hydrogen bonds if the backbone directions of the two strands make an angle with each other. This leads to the left-handed twist characteristic of β-sheets observed in a number of proteins.* ***(d)*** *Ribbon diagram of the protein thioredoxin from E. coli, illustrating the left-handed twist of a β-sheet. (parts (a)-(c) adapted from Figure 5.10 in {Schulz & Schirmer 1979 ID: 513}, and (d) taken from Figure 2.7(a) of {Branden & Tooze 1999 ID: 507})*

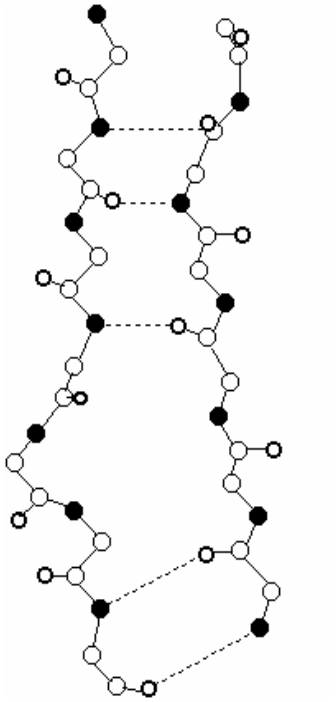

Figure II.2.11. *A β-bulge in a sheet in immunoglobulin fragment V_L. (taken from {Lesk 1991 ID: 514})*

Turns and loops. The secondary structures described above - α-helices and β-sheets - are regular and repeating and the most common types of local structures. When there is a reversal of chain direction, secondary structures are connected by *turns* (also called *reverse turns*) and *loops*. Although turn and loop conformations do not continue in a repeating fashion along the backbone, sometimes they are included as secondary structures and detailed classifications of turns have even been developed. See *Figure II.2.12*. According to the original definition of Linderstrom-Lang in 1953, secondary structure meant only helices. Soon thereafter, "secondary structure" came to include sheets. We follow that convention here: and refer to secondary structures as only the repeating regular structures - helices and sheets, but not turns and loops. Turns are usually short and tight, in the range of 3-5 monomers, typically self-hydrogen bonded; loops are longer.

Reverse turns occur almost exclusively at the protein surfaces. Reverse turns therefore contain polar residues, together with glycine and proline. As shown in § II.1, having no side chain hindrance, Gly can adopt a broad range of dihedral angles giving rise to kinks in the main chain. Pro is unique, because its sidechain curls back to the main chain and seizes it in a ring.

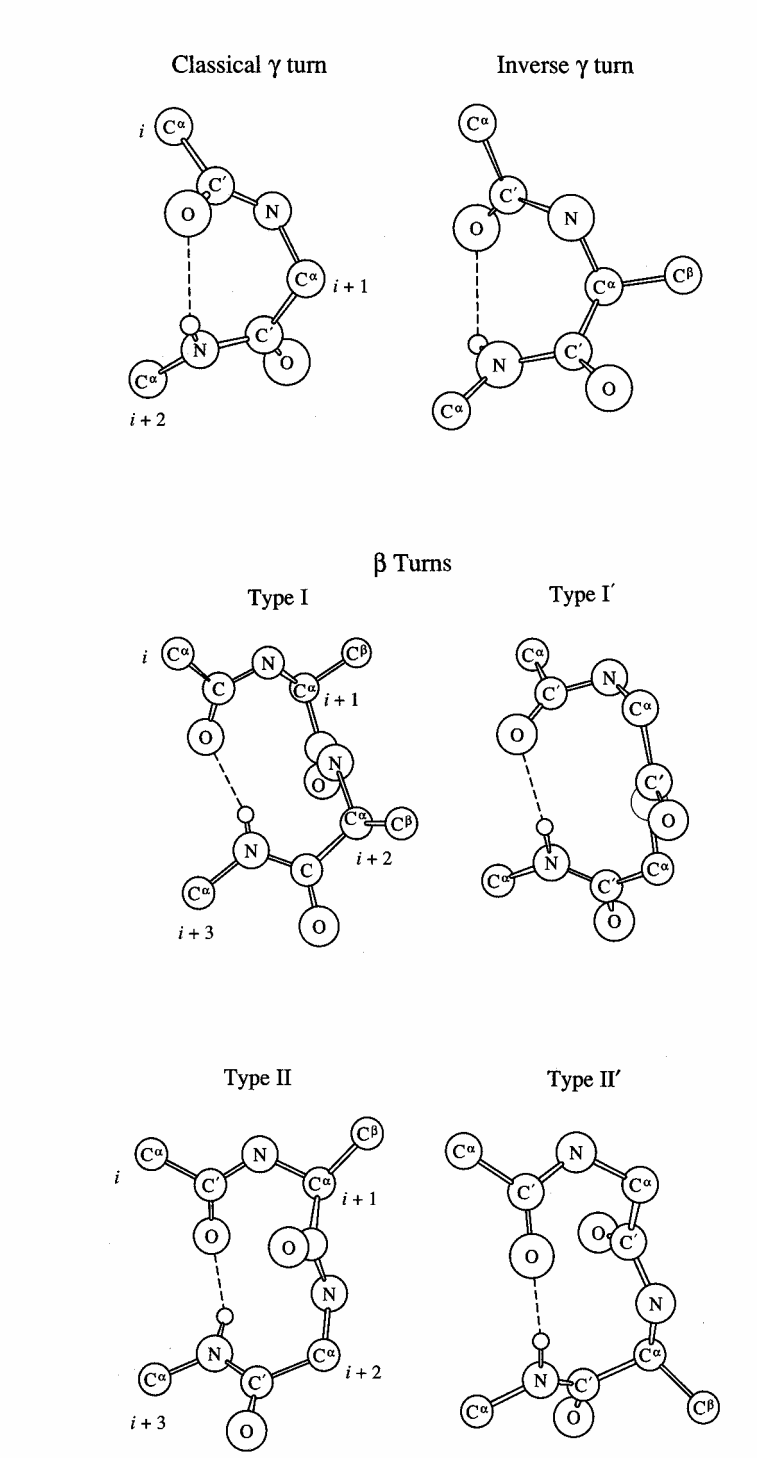

Figure II.2.12. Common reverse turns γ and β turns connect adjacent strands of an antiparallel β-sheet. The dashed lines indicate the last hydrogen bond of the β-sheet. γ turns are very tight; they are made up of three residues, one of which is not involved in hydrogen bonding. The more common β turns involve four residues, two of which do not form hydrogen bonds. Types I' and II' backbone conformations are mirror images of types I and II backbone conformations, respectively. C^{β} atoms are included only at positions where residues other than Gly occur frequently. Not shown here are Type III β turns. They may be considered as very short segments of 3_{10}-helical conformation. Many other types of turns have been defined as well. (taken from {Creighton 1993 ID: 495}