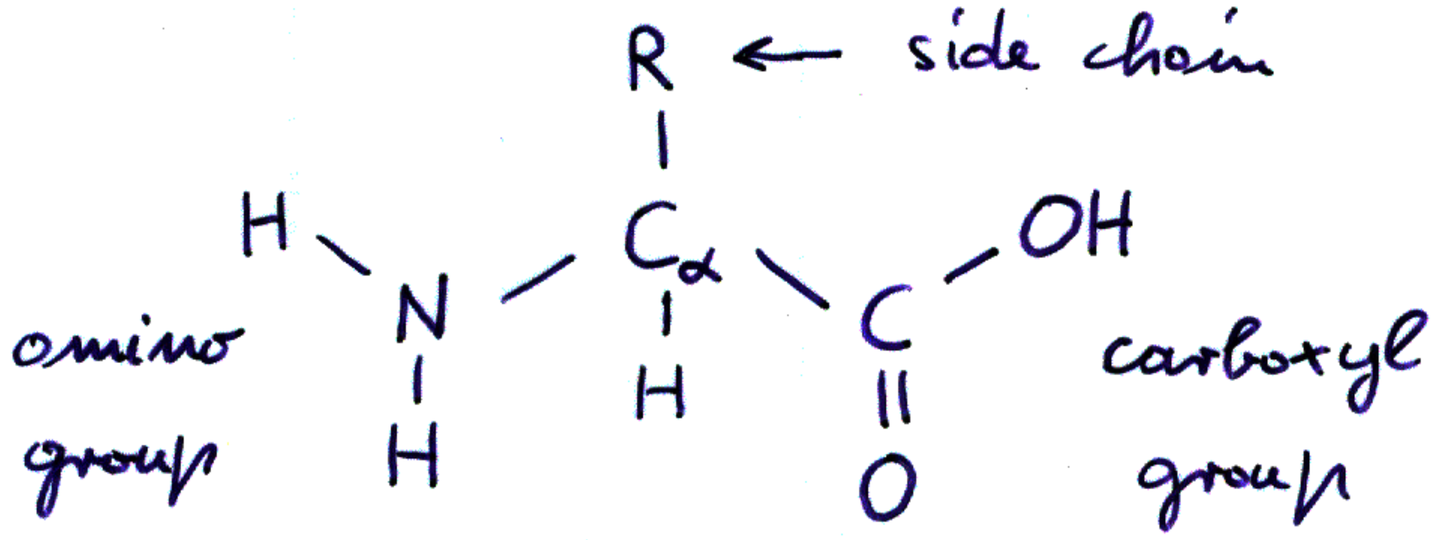
Protein Structure

A protein is a polymer made of aminoacids.

Aminoacids:



The simplest side chain is just $-H$ → Glycine

Hydrophobic side chains: e.g. $-CH_3$ → Alanine

Hydrophilic side chains:

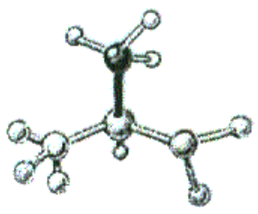
a) charged e.g. $-CH_2-COO^-$ Aspartic acid

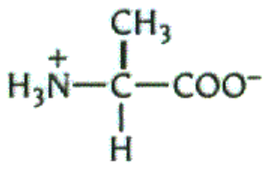
b) polar e.g. $-CH_2-OH$ Serine

Amino acids

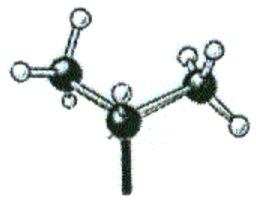
(a) Hydrophobic amino acids

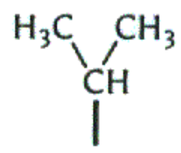
Hydrophobic



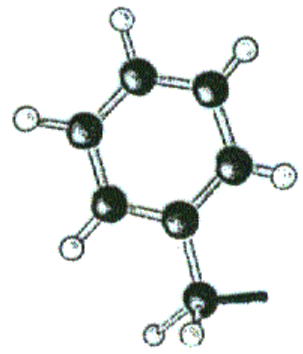


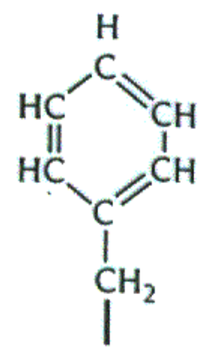
A Ala, Alanine



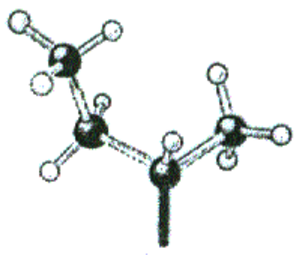


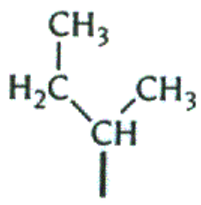
V Val, Valine



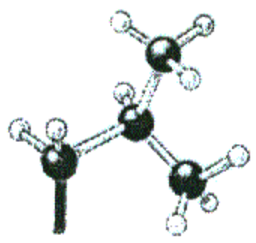


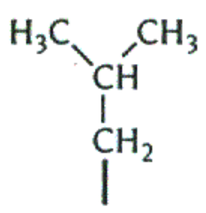
F Phe, Phenylalanine





I Ile, Isoleucine

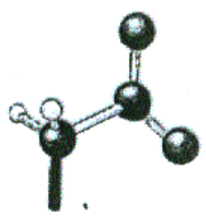


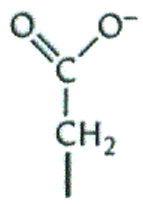


L Leu, Leucine

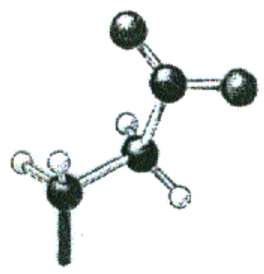
(b) Charged amino acids

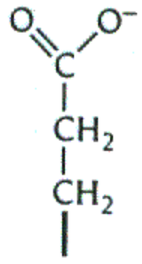
Charged



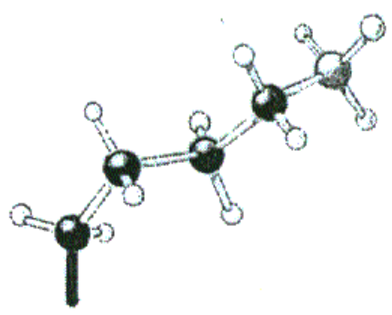


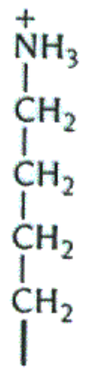
D Asp, Aspartic acid





E Glu, Glutamic acid

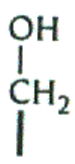




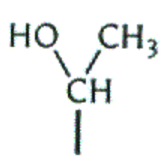
K Lys, Lysine

(c) Polar amino acids

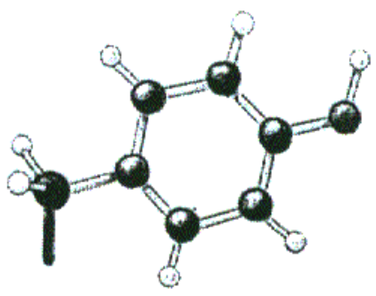
Polar

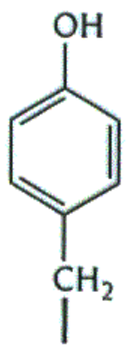


S Ser, Serine

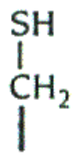


T Thr, Threonine

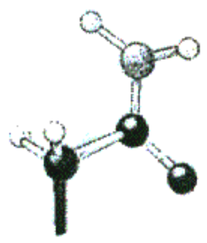


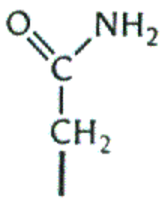


Y Tyr, Tyrosine

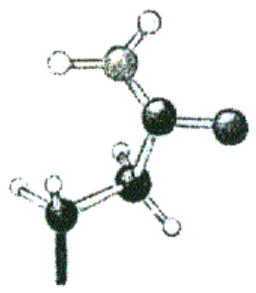


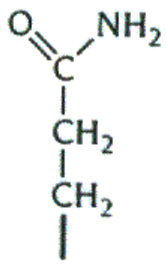
C Cys, Cysteine





N Asn, Asparagine

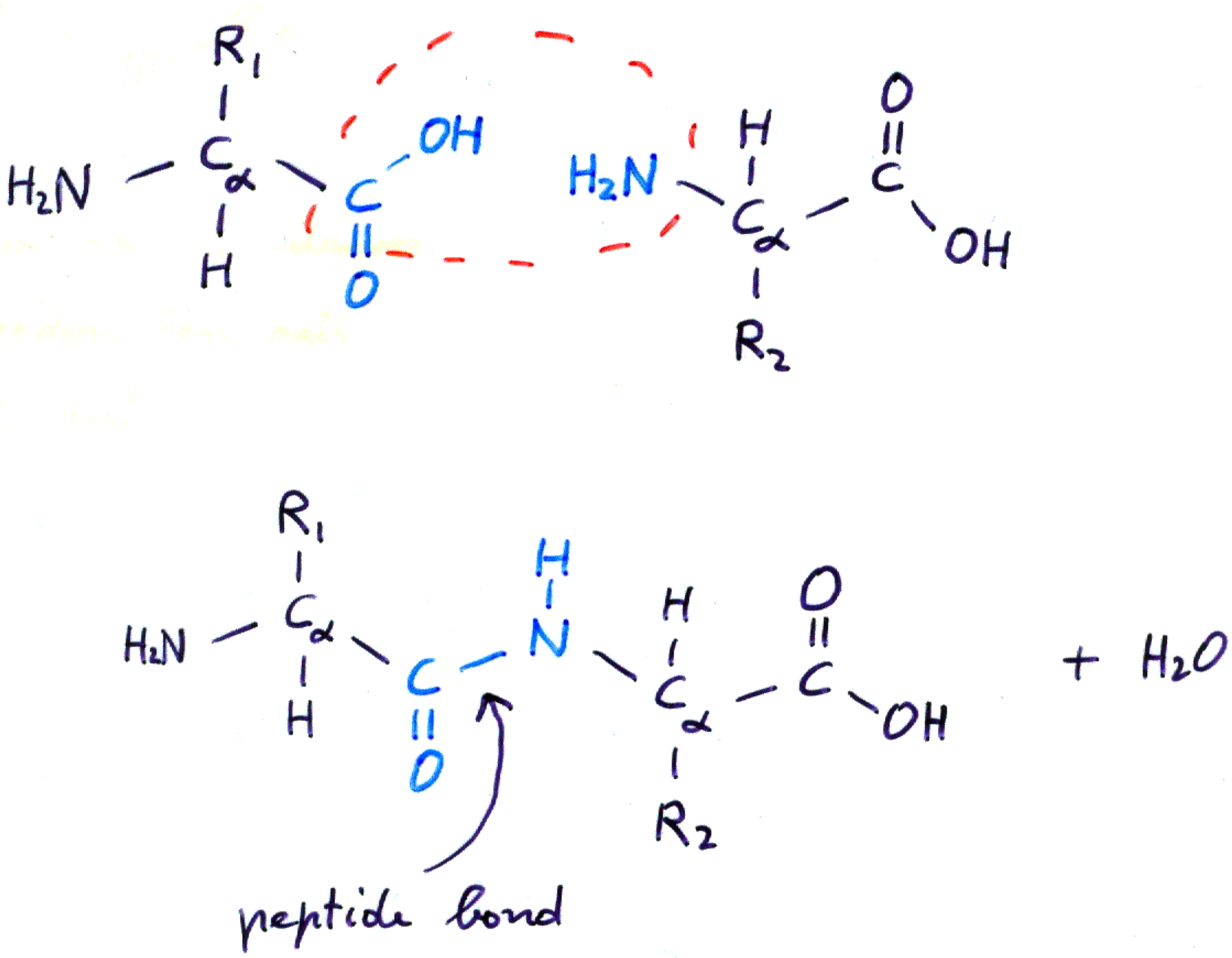




Q Gln, Glutamine

Peptide bond

The carboxyl group of one amino acid condenses with the amino group of the next → C – N bond :

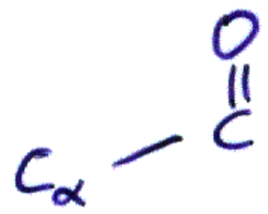


→ a polypeptide chain has an amino terminus and a carboxy terminus.

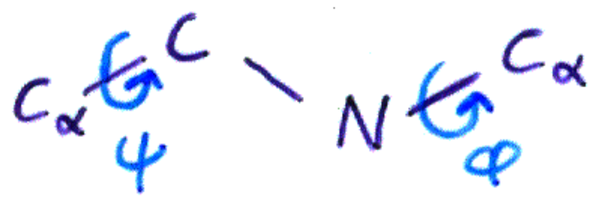
A typical protein consists of 100 – 1000 amino acids.

The sequence of amino acids is specified by the genetic code.

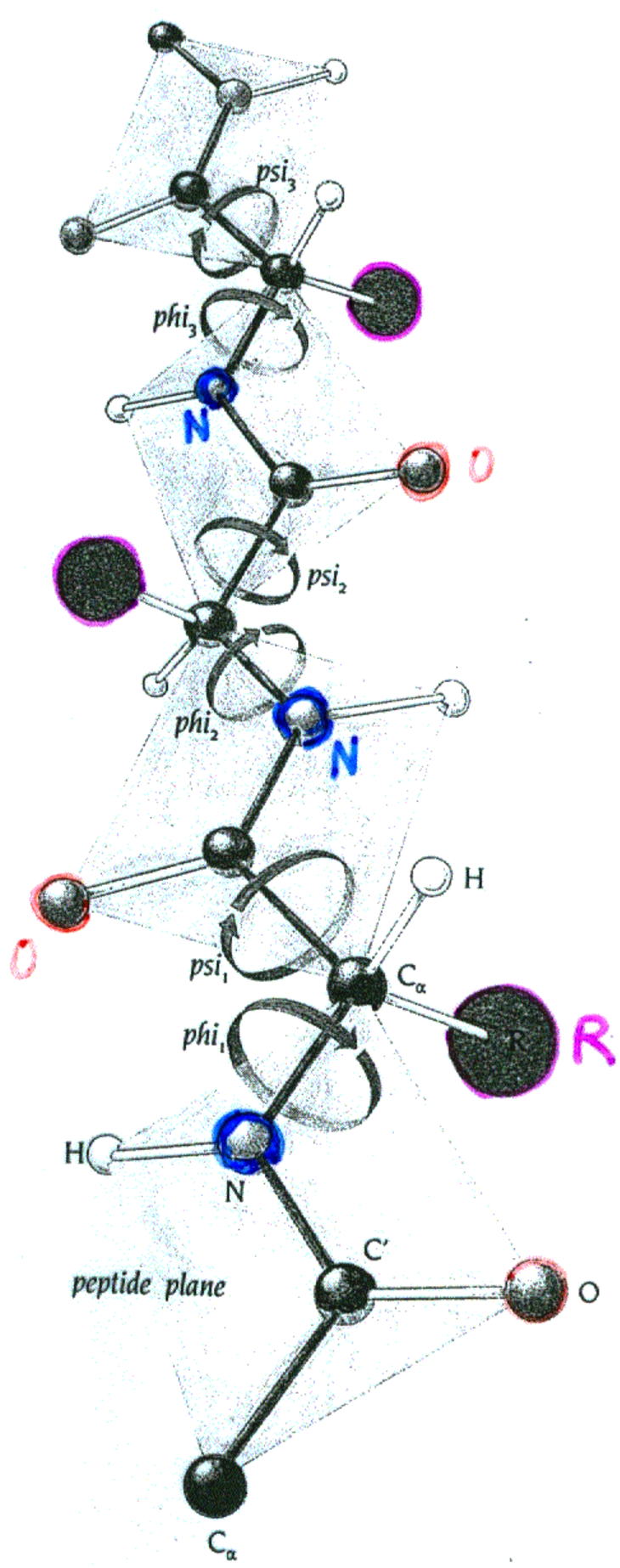
Polypeptide chains are flexible :

The  atoms lie in a plane,

but rotations are allowed around the $C_\alpha - C$ and $N - C_\alpha$ bonds :



→ there are 2 degrees of freedom for each peptide unit

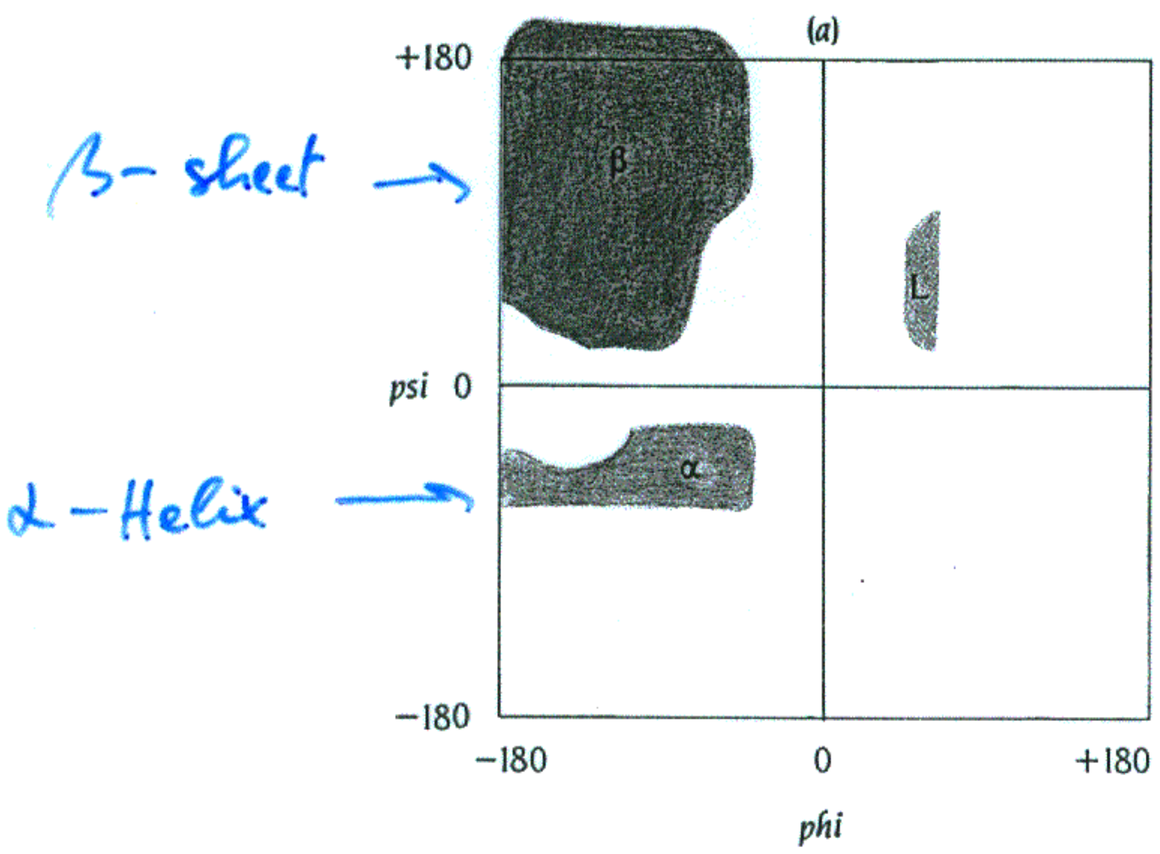


gure 1.6 Diagram showing a polypeptide .ain where the main chain atoms are presented as rigid peptide units, linked rough the C_α atoms. Each unit has two grees of freedom; it can rotate around two nds, its C_α–C′ bond and its N–C_α bond. The gle of rotation around the N–C_α bond is lled phi (ϕ) and that around the C_α–C′ bond i (ψ). The conformation of the main chain oms is therefore determined by the values of ese two angles for each amino acid.

Because of overlap between atoms, only certain combinations of ϕ, ψ can occur in the conformation of a real protein.

Ramachandran plot (allowed combinations of ψ, ϕ)

Calculation

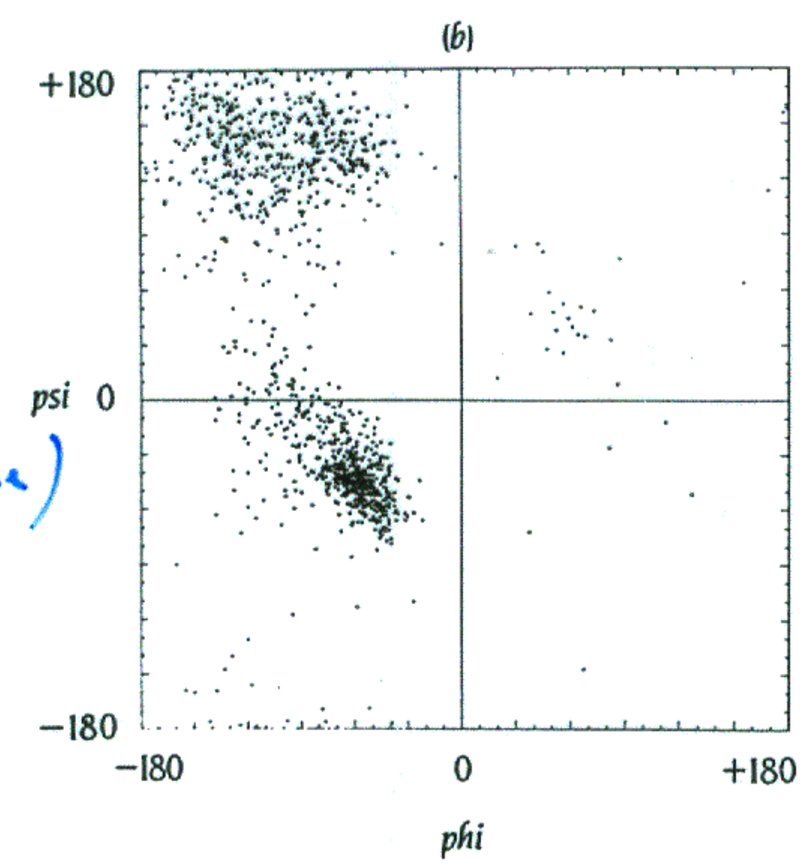


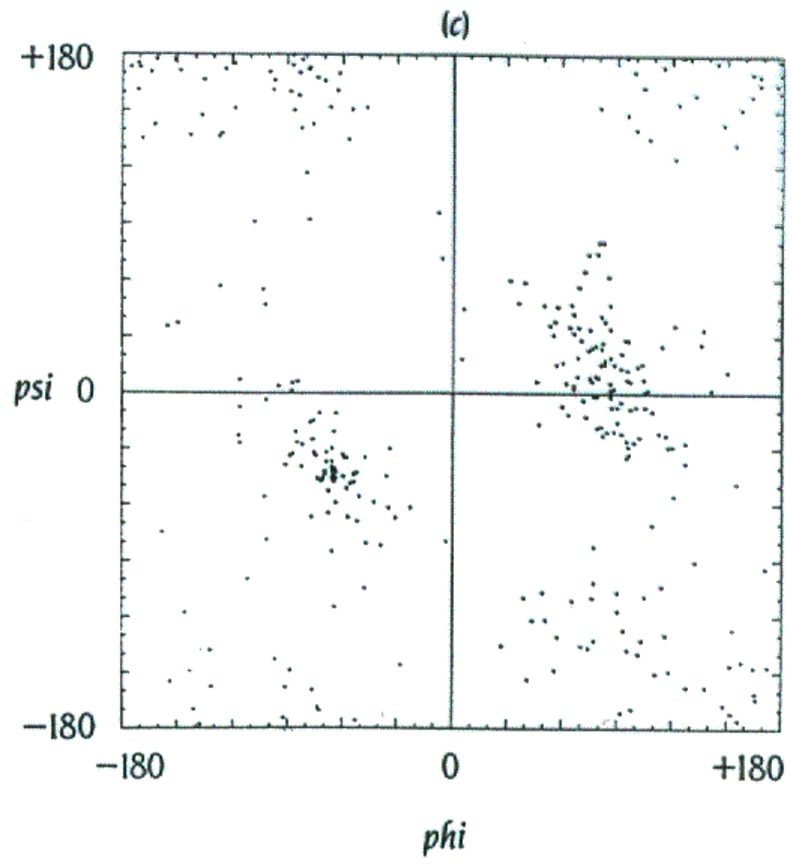
β-sheet →

α-Helix →

Figure 1.7 Ramachandran plots showing allowed combinations of the conformational angles phi and psi defined in Figure 1.6. Since phi (ϕ) and psi (ψ) refer to rotations of two rigid peptide units around the same C_α atom, most combinations produce steric collisions either between atoms in different peptide groups or between a peptide unit and the side chain attached to C_α. These combinations are therefore not allowed. (a) Colored areas show sterically allowed regions. The areas labeled α, β, and L correspond approximately to conformational angles found for the usual right-handed α helices, β strands, and left-handed α helices, respectively. (b) Observed values for all residue types except glycine. Each point represents ϕ and ψ values for an amino acid residue in a well-refined x-ray structure to high resolution. (c) Observed values for glycine. Notice that the values include combinations of ϕ and ψ that are not allowed for other amino acids. (From J. Richardson, *Adv. Prot. Chem.*34: 174–175, 1981.)

Observed values (any residue except Glycine)

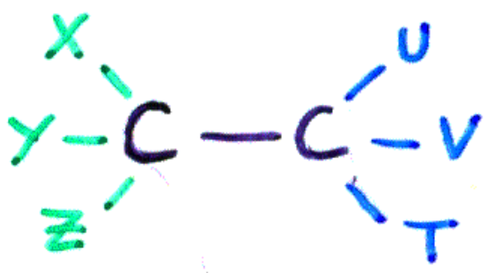


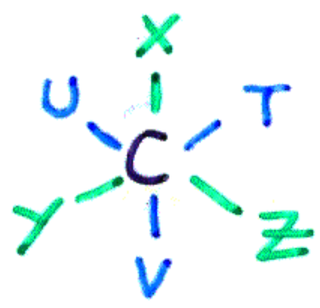


Observed values for Glycine

Energy landscape for the φ, ψ rotations

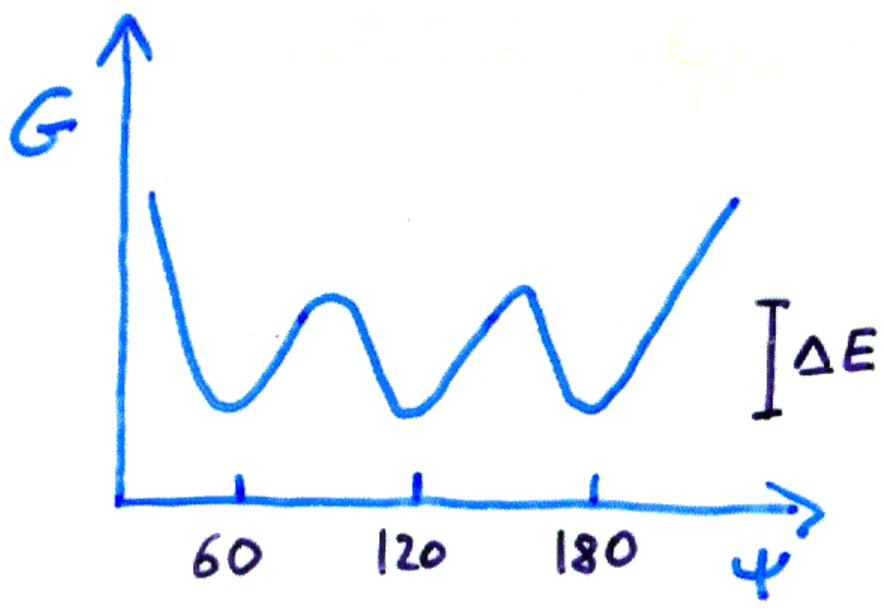
e.g. two carbon atoms (ψ):





"staggered"

→ 3 valleys in the energy landscape:



barriers are small:

$\Delta E \lesssim 5\ kT$

→ flipping is fast

With φ, ψ there are ~ 6 conformational minima per aminoacid → tot. number of minima for a 100 residue protein is $\sim (6)^{100} \simeq 10^{77}$!

One special aminoacid is Cysteine: $R = CH_2 - SH$

two Cys can form a disulfide bridge:

$$-CH_2-SH + HS-CH_2- \longrightarrow -CH_2-S-S-CH_2-$$

(requires an oxidative environment)

The peptide bond has partial double-bond character:

so rotations around the C–N bond are restricted, and the  atoms lie in a plane.

Given the bond lengths and angles, the length of a fully stretched polypeptide chain is 3.6 Å per residue.

With $\Delta E \sim 2 \div 3\ kT$

and $a = 3.6$ Å

$\rightarrow l_p = a\, e^{\Delta E/kT}$

$\sim 25 \div 70$ Å

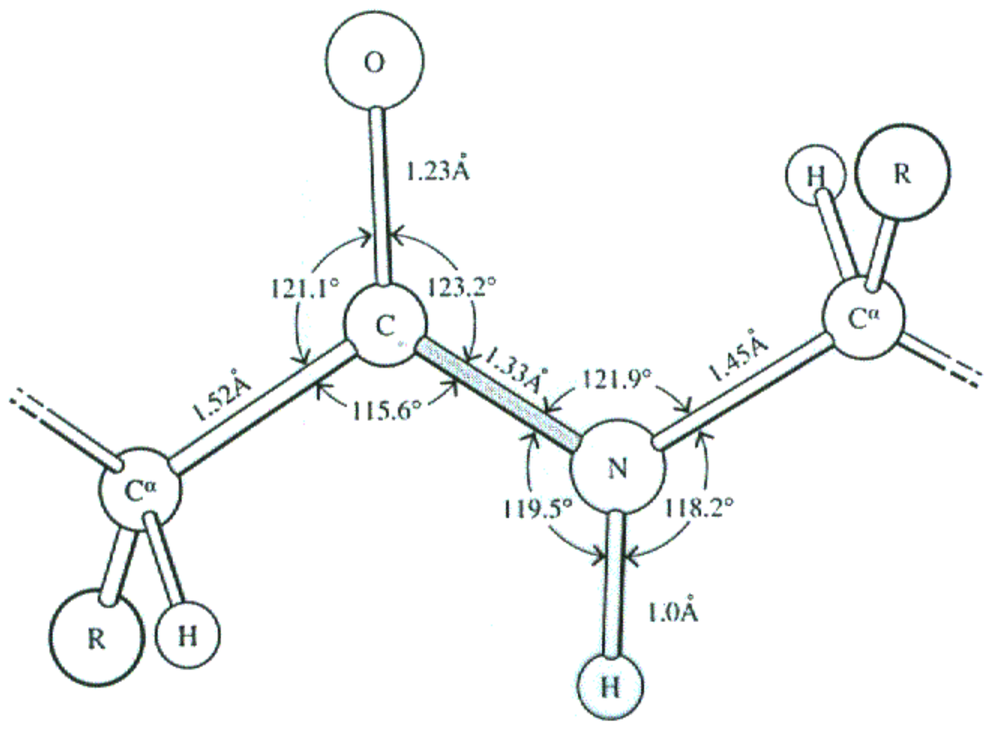
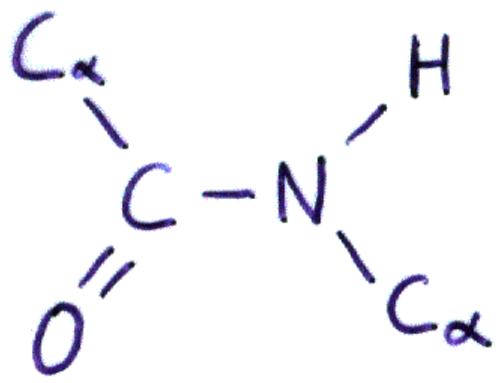
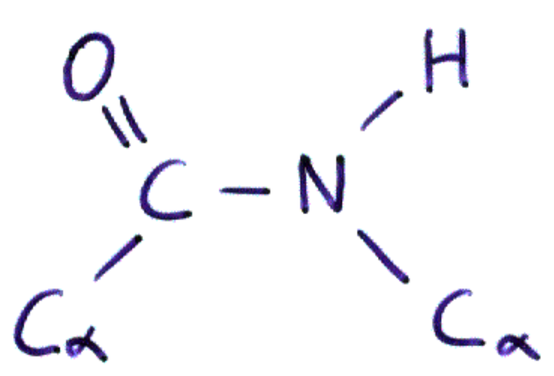


FIGURE 1.2
The geometry of the peptide backbone, with a *trans* peptide bond, showing all the atoms between two C^{α} atoms of adjacent residues. The peptide bond is stippled. The dimensions given are the averages observed crystallographically in amino acids and small peptides. (G. N. Ramachandran et al., *Biochim. Biophys. Acta* 359:298–302, 1974.)

There are two possible conformations for the peptide bond =

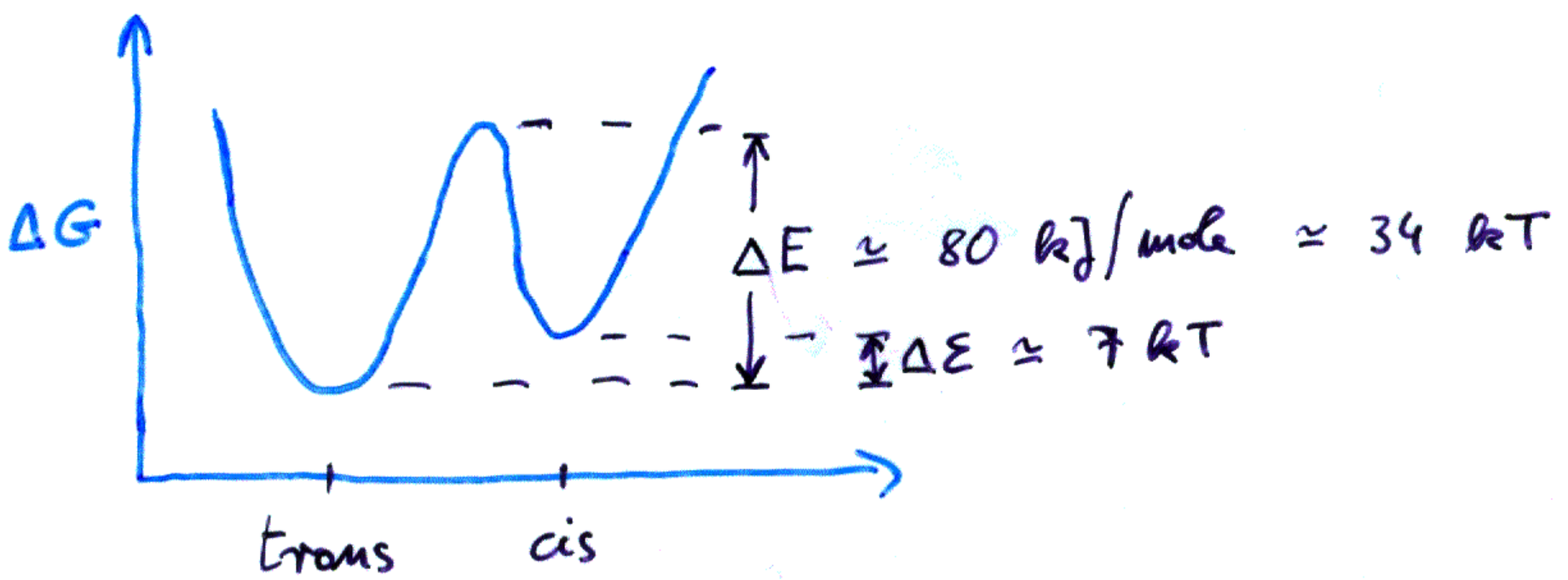


trans



cis

The barrier between cis and trans is large :



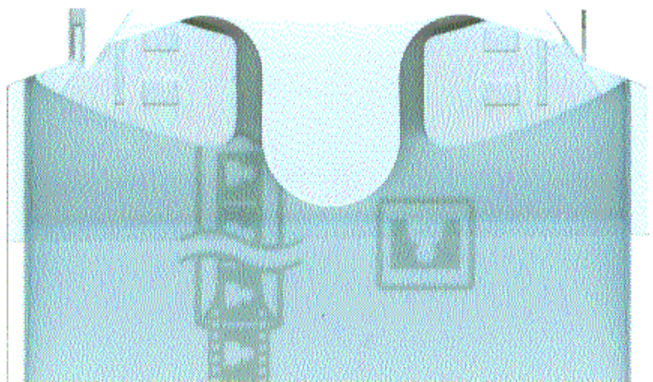
→ rate of cis-trans isomerization is slow :

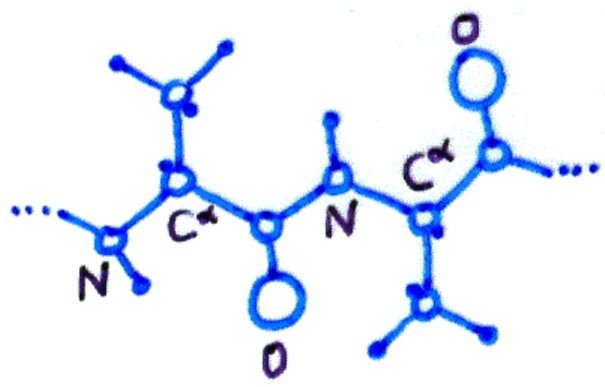
$R \sim R_0 e^{-34}$; $R_0 \simeq 10^{12}$ Hz $\Rightarrow$ $R \sim 2 \times 10^{-3}$ Hz

or $1/R \sim 500$ s

The trans conformation is lower energy : $\Delta E \sim 7\ kT$

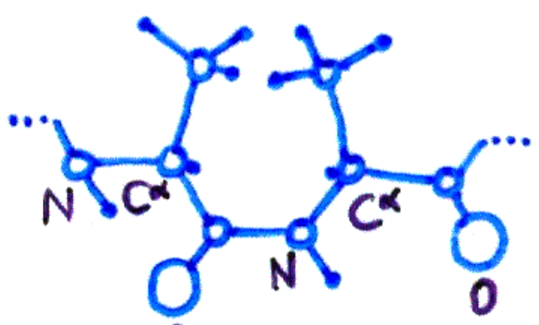
→ $N_{cis} / N_{trans} \sim 10^{-3}$





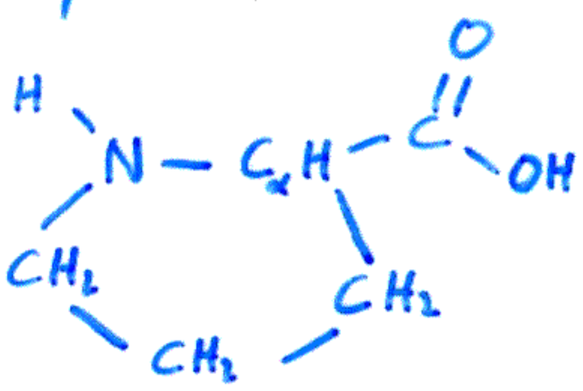
trans

steric repulsion between side chains



cis

Exception : Proline

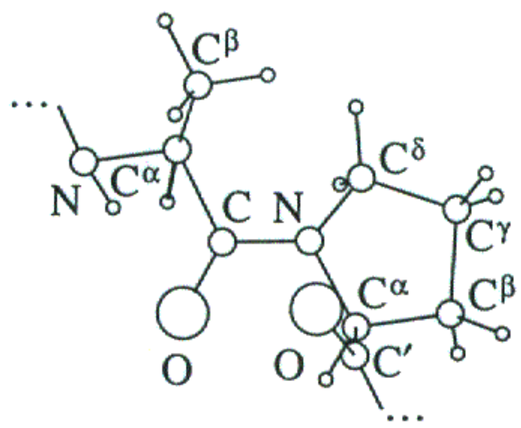


Pro :

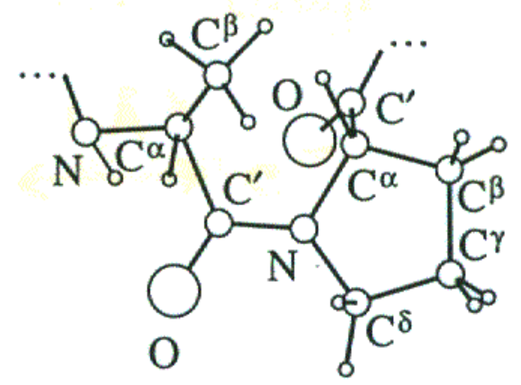
$\Delta \varepsilon$ (cis - trans) $\simeq$ 1.4 kT

$\rightarrow$ $N_{cis} / N_{trans} \simeq 1/4$

When residue $i + 1$ is Pro, however, there is very little difference between the two forms:



trans

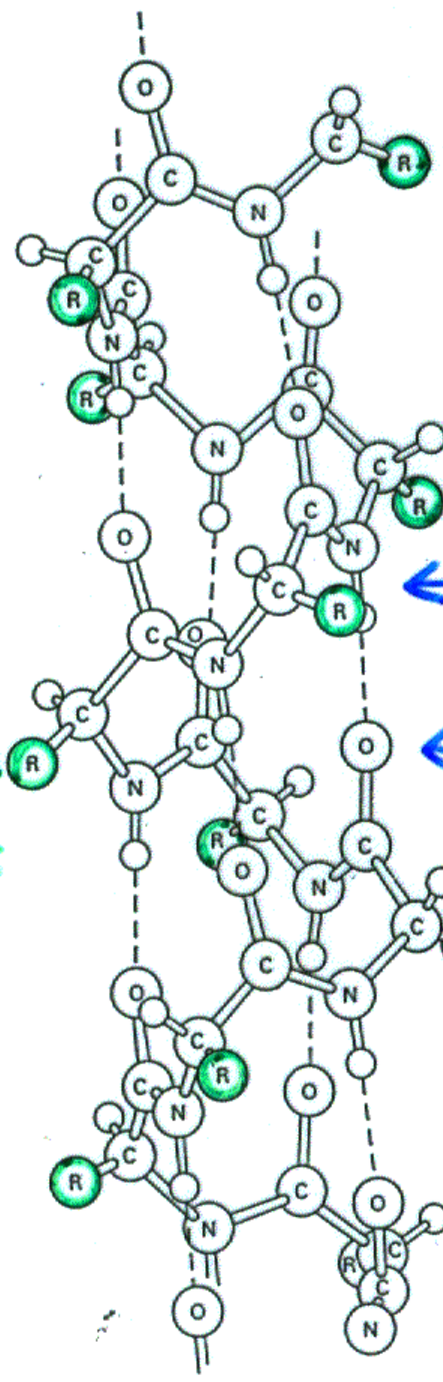


cis (5.6)

and the *trans* form is only slightly favored, generally by a ratio of about 4 : 1. The peptide bond preceding a Pro

Regular conformations of polypeptides

right handed α-Helix : 3.6 residues per turn
= 5.4 Å ~ ~

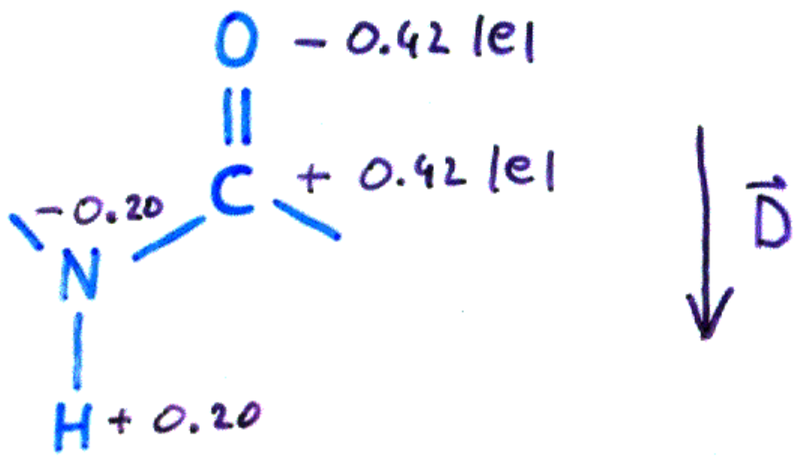


side chains → stick towards the outside

The backbone
NH ← forms a hydrogen bond with the
carbonyl oxigen ← of the backbone
four residues down the chain
(→ 2 hydrogen bonds per residue)

FIGURE 5.6
The classical right-handed α-helix.

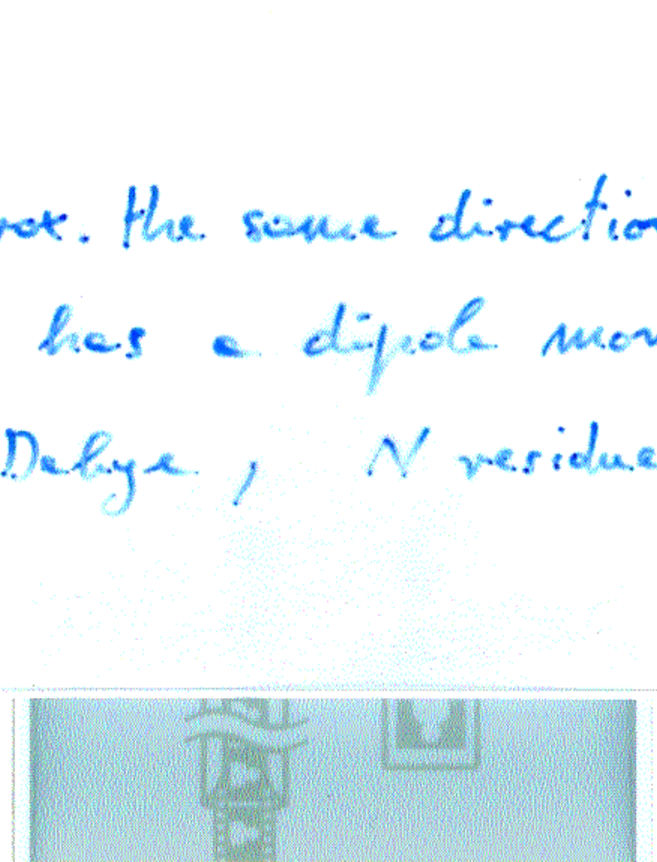
Each peptide bond has a dipole moment:



D ≃ 3.5 Debye

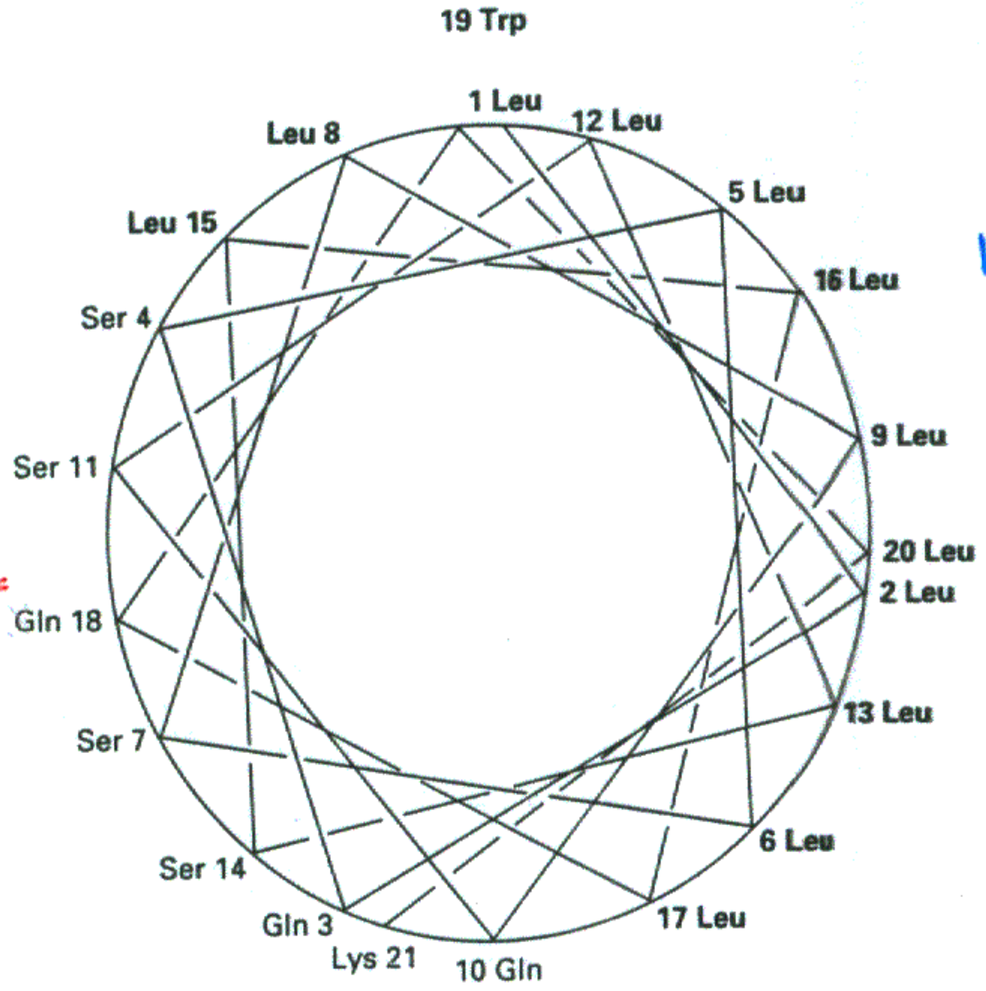
($e^+ \xleftrightarrow{1\,Å} e^-$ D = 4.8 Debye)

they all point in approx. the same direction in the α-Helix
→ the whole Helix has a dipole moment
(of ~ N × 4 Debye, N residues Helix)



Amphiphilic α-Helix: polar residues on one side and non-polar on the other → helices pack together

Example:

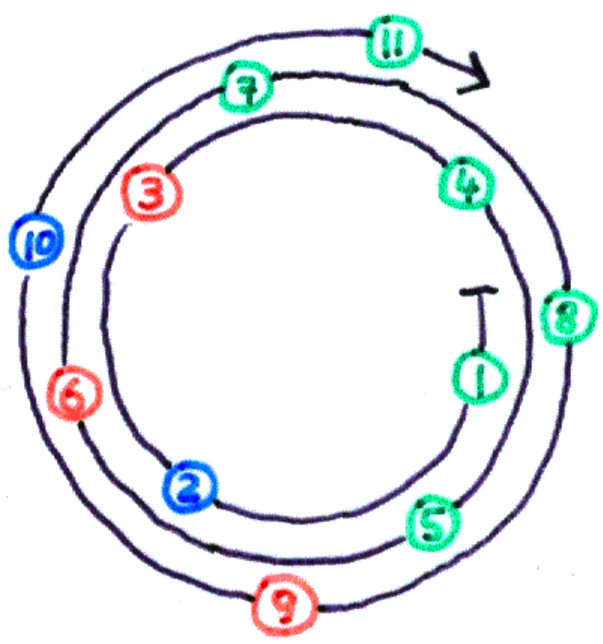


hydrophobic side

hydro-philic side

FIGURE 5.8
Helical wheel representation of an α-helix. The positions of the side chains are shown in projection down the helix axis. In an ideal α-helix, there are 3.6 residues per complete turn, or a rotation of 100° per residue. The helical wheel consequently repeats after five turns of 18 residues; residues 19–21 are offset slightly here to make them visible. In the amphipathic helix of the peptide shown, the hydrophobic residues are indicated in bold, and they can be seen to lie solely on one side of the helix; the opposite side is composed solely of polar residues. (From W. F. DeGrado et al., *J. Amer. Chem. Soc.* 103:679–681, 1981.)

α-Helix from Alcohol Dehydrogenase



- hydrophobic
- polar
- charged

] hydrophilic

β - sheet (antiparallel)

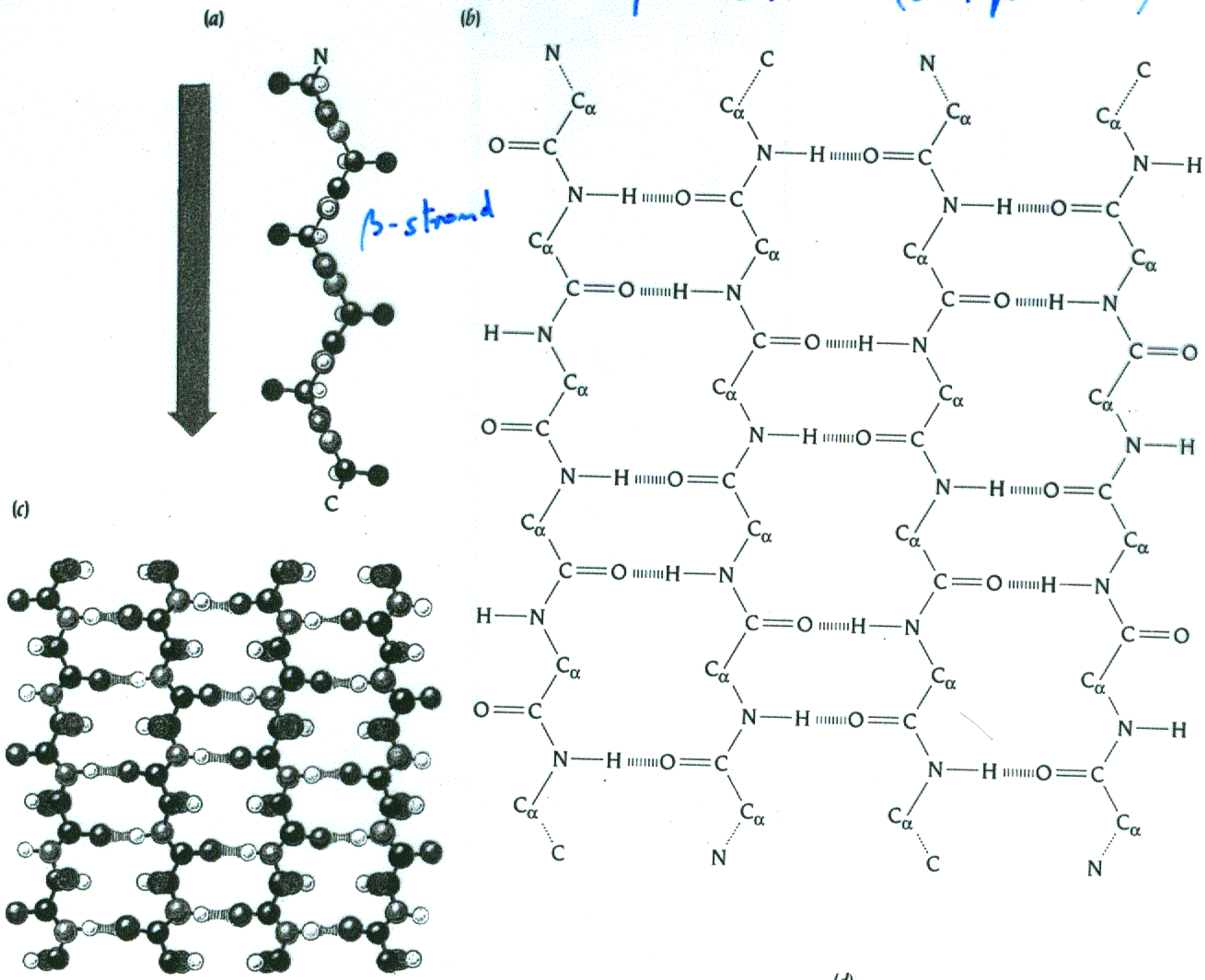
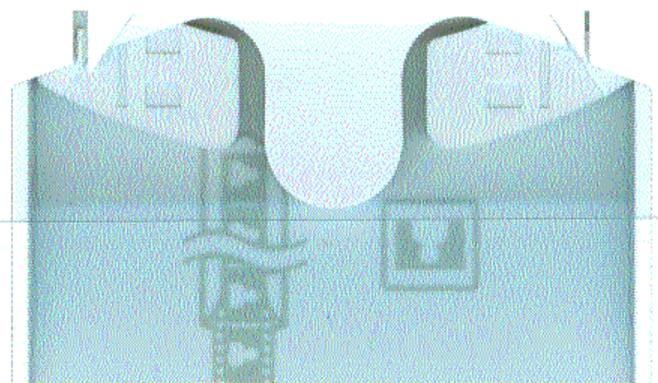


Figure 2.5 Schematic illustrations of antiparallel β sheets. β sheets are the second major element of secondary structure in proteins. The β strands are either all antiparallel as in this figure or all parallel or mixed as illustrated in following figures. (a) The extended conformation of a β strand. Side chains are shown as purple circles. The orientation of the β strand is at right angles to those of (b) and (c). A β strand is schematically illustrated as an arrow, from N to C terminus. (b) Schematic illustration of the hydrogen bond pattern in an antiparallel β sheet. Main chain NH and O atoms within a β sheet are hydrogen bonded to each other. (c) A ball-and-stick version of (b). Oxygen atoms are red; nitrogen atoms are blue. The hydrogen atom in N–H...O is white. The carbon atom in the main chain, C_α, is black. Side chains are illustrated by one purple atom. The orientation of the β strands is different from that in (a). (d) Illustration of the pleat of a β sheet. Two antiparallel β strands are viewed from the side of the β sheet. Note that the directions of the side chains, R (purple), follow the pleat, which is emphasized in yellow.

β - strands (extended conformation of the polypeptide chain) assemble into β - sheets.

Again, each oxigen of the backbone forms a hydrogen bond with an NH group of the next strand



parallel β-sheet

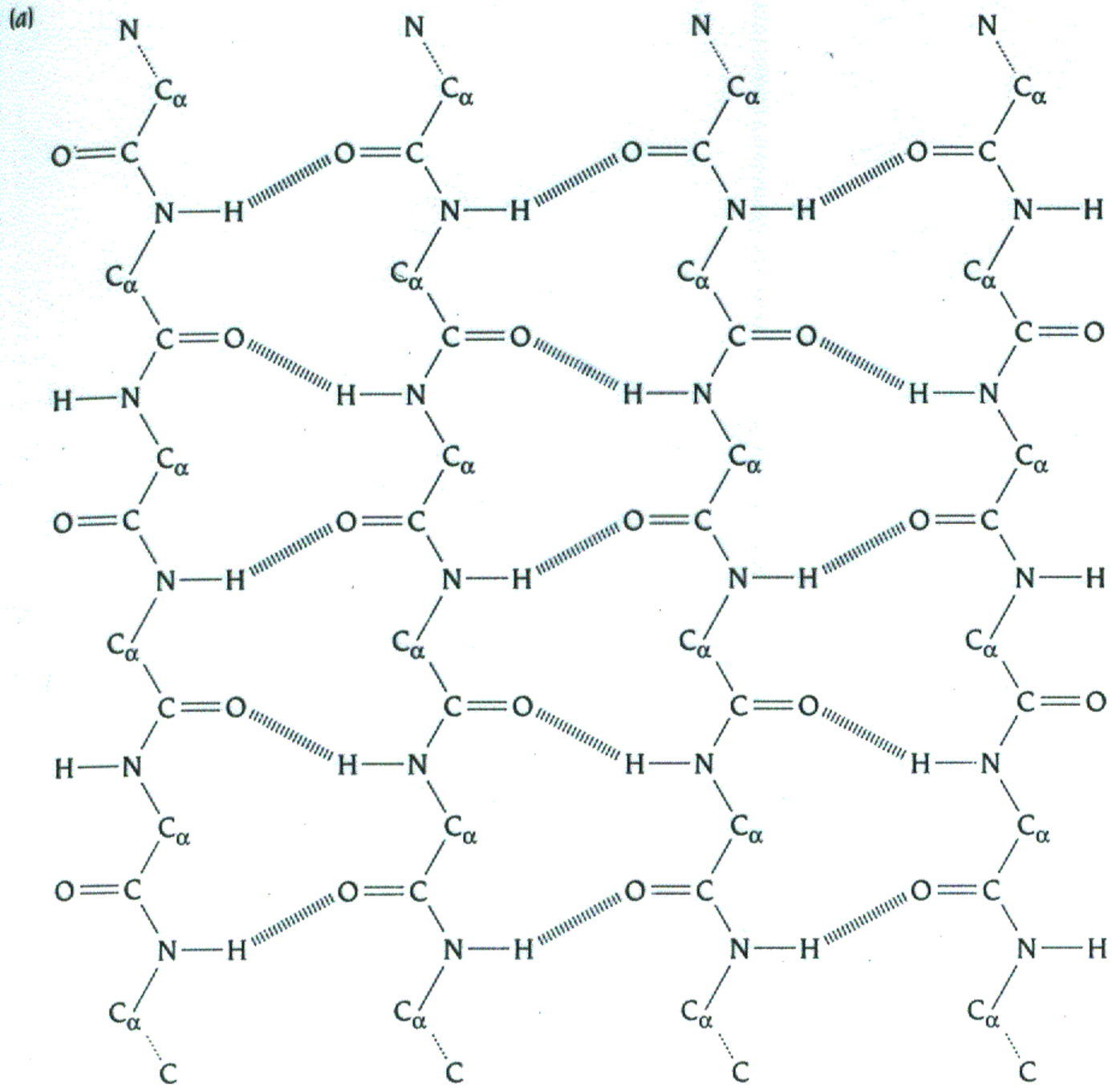
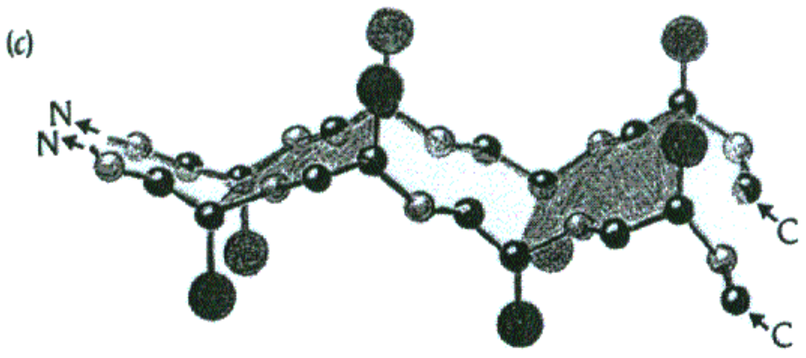


Figure 2.6 Parallel β sheet. (a) Schematic diagram showing the hydrogen bond pattern in a parallel β sheet. (b) Ball-and-stick version of (a). The same color scheme is used as in Figure 2.5c. (c) Schematic diagram illustrating the pleat of a parallel β sheet.



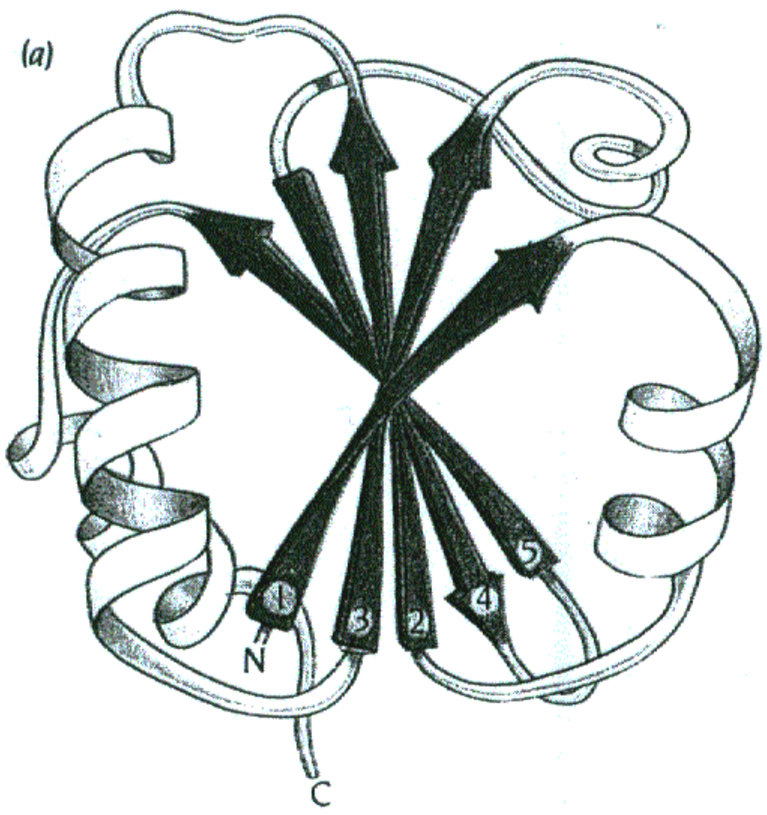


Figure 2.7 (a) Illustration of the twist of β sheets. β strands are drawn as arrows from the amino end to the carboxy end of the β strand in this schematic drawing of the protein thioredoxin from *E. coli*, the structure of which was determined in the laboratory of Carl Branden, Uppsala, Sweden, to 2.8 Å resolution. The mixed β sheet is viewed from one of its ends. (Adapted from B. Furugren.) (b) The hydrogen bonds between the β strands in the mixed β sheet of the same protein (page 18).

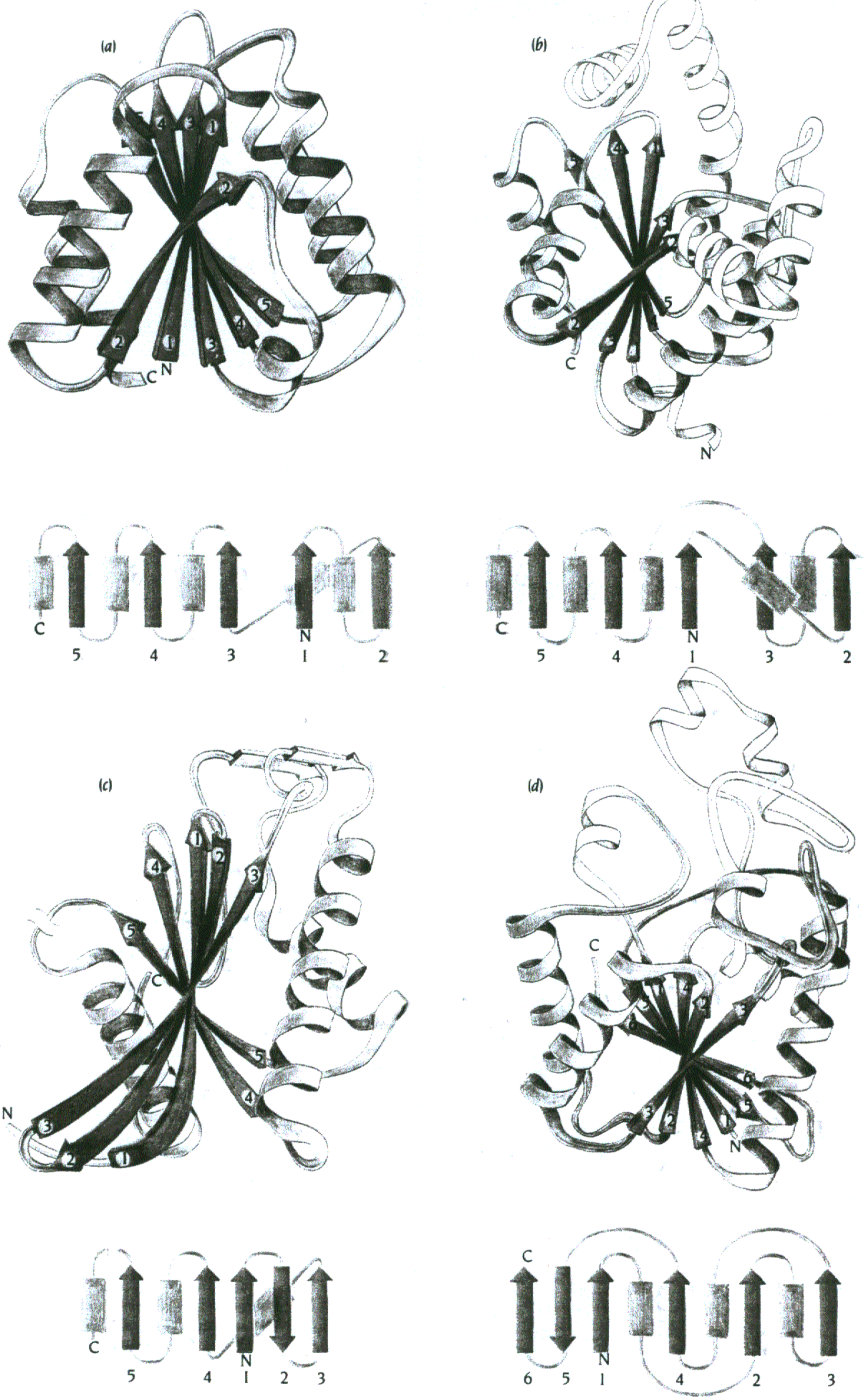


Figure 4.8
Examples of different types of open twisted α/β structures. Both schematic and topological diagrams are given. Arrows denote strands of β sheet and rectangles denote α helices. (a) The FMN-binding redox protein flavodoxin. (b) The enzyme adenylate kinase, which catalyzes the reaction AMP + ATP = 2ADP. The structure was determined to 3.0 Å resolution in the laboratory of Georg Schulz in Heidelberg, Germany. (c) The ATP-binding domain of the glycolytic enzyme hexokinase, which catalyzes the phosphorylation of glucose. The structure was determined to 2.8 Å resolution in the laboratory of Tom Steitz, Yale University. (d) The glycolytic enzyme phosphoglycerate mutase, which catalyzes transfer of a phosphoryl group from carbon 3 to carbon 2 in glycerate. The structure was determined to 2.5 Å resolution in the laboratory of Herman Watson, Bristol University, UK. (Adapted from J. Richardson.)