

## Lecture 6: Protein Primary, Secondary and Tertiary Structure

Assigned reading in Campbell: Chapter 4.1-4.3, Chapter 5.4

Key Terms:

- **Native conformation**
- **Subunit**
- **Domain**
- **Edman degradation**
- **Cyanogen bromide cleavage**
- **Chymotrypsin & trypsin cleavage**
- **$\alpha$ -Helix**
- **Parallel  $\beta$ -sheet**
- **Antiparallel  $\beta$ -sheet**

### 6.1 Protein Structure Determines Function

*Overview of the next few lectures:*

Protein structures, even of the smallest proteins, can appear at first glance to be a daunting jumble of variously colored balls and sticks. A typical (and reasonable) response from first-time viewers is, "What am I looking at?" or "What's important here?" An efficient route to the goal of understanding structure is to first organize our learning, and then our thinking into a hierarchy. For proteins, the **structural hierarchy** is:

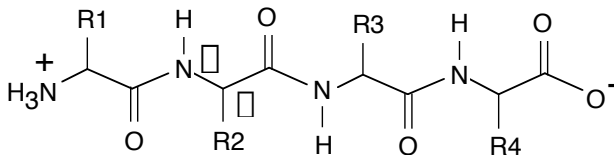
1. **Primary structure (1°): the amino acid sequence.**
2. **Secondary structure (2°): helices, sheets and turns.**
3. **Tertiary structure (3°): side chain packing in the 3-D structure.**
4. **Quaternary structure (4°): association of subunits.**

A few additional divisions in the above are also useful. For example, frequently found 2° structure patterns are termed "**super secondary structures**"; some proteins fold into one or more independent 3° structure "**domains**"; and the 4° structural association can occur between identical or dissimilar subunits. Viewing protein structures at the various hierarchical levels listed above is an essential part of understanding the overall and the detailed aspects of protein structure and function.

**Native Structure:** The native conformation of a given protein is its functionally active conformation. This is a well-defined arrangement of atoms in three-dimensional space that generates a complex surface topography. The catalytic, binding, and structural roles played by proteins are all dependent on the correct folding of the protein into a unique 3-D structure and generation of the correct surface topography. In some cases, this native structure can be reversibly unfolded and then renatured to the native state. In other cases, denaturation is irreversible, e.g. a boiled egg (extreme heat) or cottage cheese (extreme pH).

### 6.2 Primary Structure of Proteins

The primary structure of proteins (i.e. the amino acid sequence) can be determined using a number of methods. We will focus on N-terminal sequencing using Edman degradation coupled with fragmentation of the polypeptide (DNA sequence and mass spectroscopy are more common).



The sequencing of a protein is accomplished by repeated application of a procedure called Edman degradation. The Edman reagent reacts with a peptide's N-terminal residue and subsequent cleavage of the modified residue allows its identification. The actual chemical mechanism of Edman degradation is described in Campbell and will *not* be covered.

It is generally not possible to sequence an entire protein from the amino terminus (10-40 residues is reasonable). Therefore, the sequence information is extended by first fragmenting the protein into smaller peptides with overlapping sequences. Note these three examples:

- a. Cyanogen bromide (CNBr) cleaves the peptide bond after Methionine residues. You do not need to know the mechanism, just the specificity of the reaction.
- b. Chymotrypsin hydrolyzes the peptide bonds that follow large hydrophobic residues, e.g. Phenylalanine, Tyrosine, Tryptophan. You should remember this cleavage pattern.
- c. Trypsin hydrolyzes the peptide bonds that follow positively charged residues, e.g. Lysine and Arginine. You should remember this cleavage pattern as well.

After cleavage, the individual peptide fragments are separated from each other and subject to N-terminal sequencing using the Edman degradation method. What is important is the logic of ordering the overlapping sequence segments to produce a final primary structure.

**An example:** Underlined bold residues indicate sequence determined by Edman degradation. In this example we are assuming that **7 cycles** of Edman degradation are possible, but in practice 10-40 can be accomplished.

**Ala-Gly-Met-Ser-Thr-Gly-Val**-Val-Lys-Gly-Ser-Ala-Phe-Leu-Ser

CNBr:            **Ala-Gly-Met**                            **Ser-Thr-Gly-Val-Val-Lys-Gly**-Ser-Ala-Phe-Leu-Ser  
(2 fragments after treatment)

Trypsin:        **Ala-Gly-Met-Ser-Thr-Gly-Val**-Val-Lys                            **Gly-Ser-Ala-Phe-Leu-Ser**  
(2 fragments after treatment)

Strategy:

1. Find overlaps between fragments obtained with different cleavage reagents (e.g. Ser-Thr-Gly).
2. Pick a fragment from one cleavage reaction (e.g. CNBr) and look for cleavage sites for another reagent (e.g. Trypsin) within that fragment. Then identify the sequence in the fragments produced by the other reagent (e.g. Gly-Ser...)

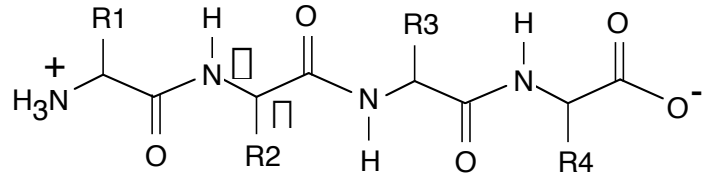
### 6.3 Secondary Structure of Proteins:

**Conformational Geometry of Peptides and Proteins:** The peptide bond is planar and trans. However, free rotation can and does occur about the two single bonds on either side of the  $\alpha$ -carbon:

$\phi$  (Phi), the bond between N and  $C_{\alpha}$

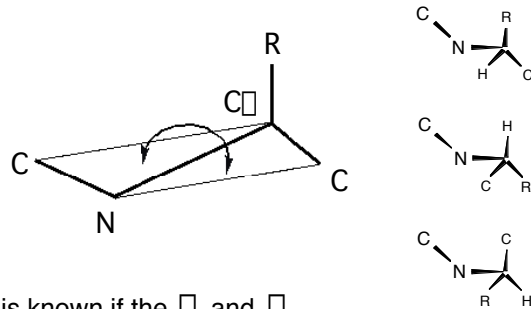
$\psi$  (Psi), the bond between  $C_{\alpha}$  and C.

In the figure of the tetra-peptide shown here, the phi  $\phi$  and psi  $\psi$  angles are labeled for residue 2. Each residue will have its own particular  $\phi$  and  $\psi$  angles.



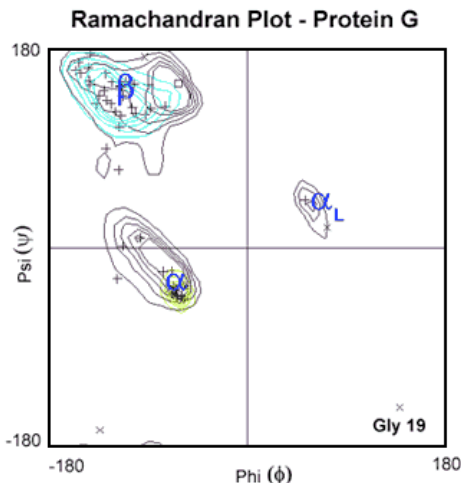
As drawn above, the figure is said to be in the “extended chain conformation”, i.e. the peptide bonds lie in the plane and both the  $\phi$  and  $\psi$  torsion angles are *defined* to equal  $180^\circ$  in this conformation.

A torsion angle defines the relative orientation of four atoms in space. In case of the  $\phi$  angle, the C, N,  $C_{\alpha}$  and C are the four atoms. Although there is free rotation about the N- $C_{\alpha}$  bond, three conformations are most stable. In these conformations the atoms are not eclipsed, avoiding steric clashes between non-bonded atoms. The same is true about the  $\psi$  angle.



The entire path of the peptide backbone in a protein is known if the  $\phi$  and  $\psi$  angles are specified (note: this is not the entire 3-D structure). Many combinations of  $\phi$  and  $\psi$  angles are not allowed due to steric clash. Those combinations that *are* sterically allowed represent the subclasses of common secondary structure found in most proteins.

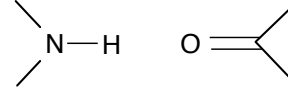
The existence of favorable combinations of  $\phi$  and  $\psi$  angles in proteins is exemplified by the Ramachandran plot. Note the clustering of  $\phi$  and  $\psi$  values in a few regions of the plot. Sparsely populated regions of the plot represent unfavorable combinations of  $\phi$  and  $\psi$  values.



## Common Secondary Structure: $\alpha$ -Helix and $\beta$ -sheet:

The  $\alpha$ -Helix and  $\beta$ -sheet conformations of the peptide chain are low energy conformations because they:

- Maximize main-chain hydrogen bonding
- Maximize van der Waals interactions of main-chain atoms
- Minimize steric clashes of main-chain and side-chain atoms



The regular repeating secondary structures of proteins have characteristic values of  $\phi$  and  $\psi$  angles that is the same for *each* residue within the element of secondary structure (see Ramachandran plot above).

### A. Helical Structures

#### $\alpha$ -Helix:

3.6 residues/turn  
rise/residue = 1.5 Å  
pitch = 5.4 Å/turn  
side-chains point outwards

### B. Beta Structures

#### $\beta$ -hairpins:

two strands connected by a sharp turn

#### $\beta$ -Sheet:

a. parallel  
b. anti-parallel  
side-chains point up and down

In both  $\alpha$  and  $\beta$  types of structures, each peptide bond is rigid and planar in the *trans* conformation. The two single bonds adjacent to the  $\alpha$ -carbon are *free* to rotate. In principle, both  $\phi$  and  $\psi$  angles can assume the three possible torsional angles that lead to no eclipsed atoms. However, the presence of the bulky atoms on the side chain as well as the necessity to form hydrogen bonds restricts possible  $\phi$  and  $\psi$  angles to 3 values:

**Right handed  $\alpha$ -helix:**  $\phi = -60^\circ$  and  $\psi = -45^\circ$

**Left handed  $\alpha$ -helix:**  $\phi = +60^\circ$  and  $\psi = +45^\circ$

**$\beta$ -sheet:**  $\phi = -120^\circ$  and  $\psi = 125^\circ$

The particular conformations listed above permit the formation of hydrogen bonds between the protein backbone in helices and sheets.

