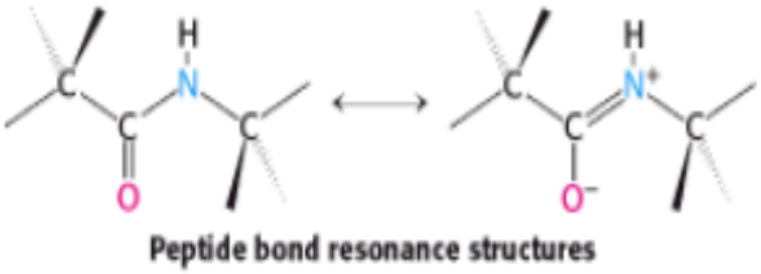
# RAMACHANDRAN MAP

## INTRODUCTION

### Polypeptide Chains Are Flexible Yet Conformationally Restricted:

Examination of the geometry of the protein backbone reveals several important features. First, the peptide bond is essentially planar (Figure). Thus, for a pair of amino acids linked by a peptide bond, six atoms lie in the same plane: the α-carbon atom and CO group from the first amino acid and the NH group and α-carbon atom from the second amino acid. The nature of the chemical bonding within a peptide explains this geometric preference. The peptide bond has considerable double-bond character, which prevents rotation about this bond.




Peptide bond resonance structures

The inability of the bond to rotate constrains the conformation of the peptide backbone and accounts for the bond's planarity. This double-bond character is also expressed in the length of the bond between the CO and NH groups. The CN distance in a peptide bond is typically 1.32 Å, which is between the values expected for a C-N single bond (1.49 Å) and a C N double bond (1.27 Å), as shown in Figure 3.24. Finally, the peptide bond is uncharged, allowing polymers of amino acids linked by peptide bonds to form tightly packed globular structures.

Two configurations are possible for a planar peptide bond. In the trans configuration, the two α-carbon atoms are on opposite sides of the peptide bond. In the cis configuration, these groups are on the same side of the peptide bond. Almost all peptide bonds in

proteins are trans. This preference for trans over cis can be explained by the fact that steric clashes between groups attached to the α-carbon atoms hinder formation of the cis form but do not occur in the trans configuration (Figure 3.25). By far the most common cis peptide bonds are X-Pro linkages. Such bonds show less preference for the trans configuration because the nitrogen of proline is bonded to two tetrahedral carbon atoms, limiting the steric differences between the trans and cis forms (Figure 3.26).

In contrast with the peptide bond, the bonds between the amino group and the a -carbon atom and between the α-carbon atom and the carbonyl group are pure single bonds. The two adjacent rigid peptide units may rotate about these bonds, taking on various orientations. This freedom of rotation about two bonds of each amino acid allows proteins to fold in many different ways. The rotations about these bonds can be specified by dihedral angles (Figure 3.27). The angle of rotation about the bond between the nitrogen and the a -carbon atoms is called phi ( fi ). The angle of rotation about the bond between the a -carbon and the carbonyl carbon atoms is called psi (psi). A clockwise rotation about either bond as viewed from the front of the back group corresponds to a positive value. The φ and ψ angles determine the path of the polypeptide chain.

**Dihedral angle A measure of the rotation about a bond, usually taken to lie between - 180° and +180°. Dihedral angles are sometimes called torsion angles.**

Are all combinations of φ and ψ possible? G. N. Ramachandran recognized that many combinations are forbidden because of steric collisions between atoms. The allowed values can be visualized on a two-dimensional plot called a Ramachandran diagram (Figure 3.28). Three-quarters of the possible (φ, ψ) combinations are excluded simply by local steric clashes. Steric exclusion, the fact that two atoms cannot be in the same place at the same time, can be a powerful organizing principle.

The ability of biological polymers such as proteins to fold into welldefined structures is remarkable thermodynamically. Consider the equilibrium between an unfolded polymer that exists as a random coil that is, as a mixture of many possible conformations and the folded form that adopts a unique conformation. The favorable entropy associated with the large number of conformations in the unfolded form opposes folding and must be overcome by interactions favoring the folded form. Thus, highly

flexible polymers with a large number of possible conformations do not fold into unique structures. The rigidity of the peptide unit and the restricted set of allowed f and y angles limits the number of structures accessible to the unfolded form sufficiently to allow protein folding to occur.

The Ramachandran plot is the 2D plot of the φ-ψ tor- sion angles of the protein backbone. It provides a simple view of the conformation of a protein. The φ-ψ angles cluster into distinct regions in the Ramachandran plot where each region corresponds to a particular secondary structure. There are four basic types of Ramachandran plots, depending on the stereo-chemistry of the amino acid: generic (which refers to the 18 non-glycine non-pro- line amino acids), glycine, proline, and pre-proline (which refers to residues preceding a proline). The generic and proline Ramachandran plots are now well understood but the glycine and pre-proline Ramach-andran plots are not.

The generic Ramachandran plot was first explained by Ramachandran and co-workers in terms of steric clashes . This has become the standard explanation for the observed regions in the Ramachandran plot . How- ever, recent studies found significant discrepancies between the classic steric map and the Ramachandran plot of high-resolution protein structures . These dis- crepancies have now been resolved. The first discrepancy is that the $N\cdots H_{i+1}$ and $O_{i-1}\cdots C$ steric clashes in the classic steric map have no effect in the observed Ramach- andran plot . By removing these steric clashes, a better steric map can be constructed. The second discrepancy is that the Ramachandran plot cluster into distinct regions within the sterically-allowed regions of the Ramachandran plot . These clusters have now been explained in terms of backbone dipole-dipole interactions . The proline Ramachandran plot has been reproduced in a calculation. The proline Ramachandran plot is severely restricted by the pyrrolidine ring, where the flexi- bility in the pyrrolidine ring couples to the backbone .

The observed glycine Ramachandran plot has a distinctive distribution (Figure 1A) quite different to the generic Ramachandran plot. An early attempt to explain the observed Ramachandran plot in terms of a steric map of glycine (Figure 2A) fails to

account for the observed distribution. It does not explain the observed clustering at $\psi$ = 180° and $\psi$ = 0°, nor the clustering into 5 distinct regions . Using a molecular-dynamics simulation of Ace-Gly-Nme , Hu and co-workers found that the gly- cine Ramachandran plot generated by standard force- fields reproduced the original steric map but not the observed Ramachandran plot. They calculated a some- what better result with a quantum-mechanics/molecular- mechanics model, which reproduced the observed cluster- ing along $\psi$, but not the partitioning into the 5 clusters. In this study, we identify the specific interactions that define the observed glycine Ramachandran plot by studying the conformations of glycine in the structural database. We test these interactions with a simple model based on elec- trostatics and Lennard-Jones potentials.

***Regions in the glycine Ramachandran plot***

Glycine is fundamentally different to the other amino acids in that it lacks a sidechain. In particular, glycine does not have the $C\beta$ atom, which induces many steric clashes in the generic Ramachandran plot. We call the hydrogen atom that is shared with the other amino acids, the $H^{\alpha 1}$ atom. We call the hydrogen atom that replaces the $C\beta$ atom, the $H^{\alpha 2}$ atom. The absence of the $C\beta$ atom allows the glycine Ramachandran plot to run over the borders at -180° and 180° (Figure 1A).

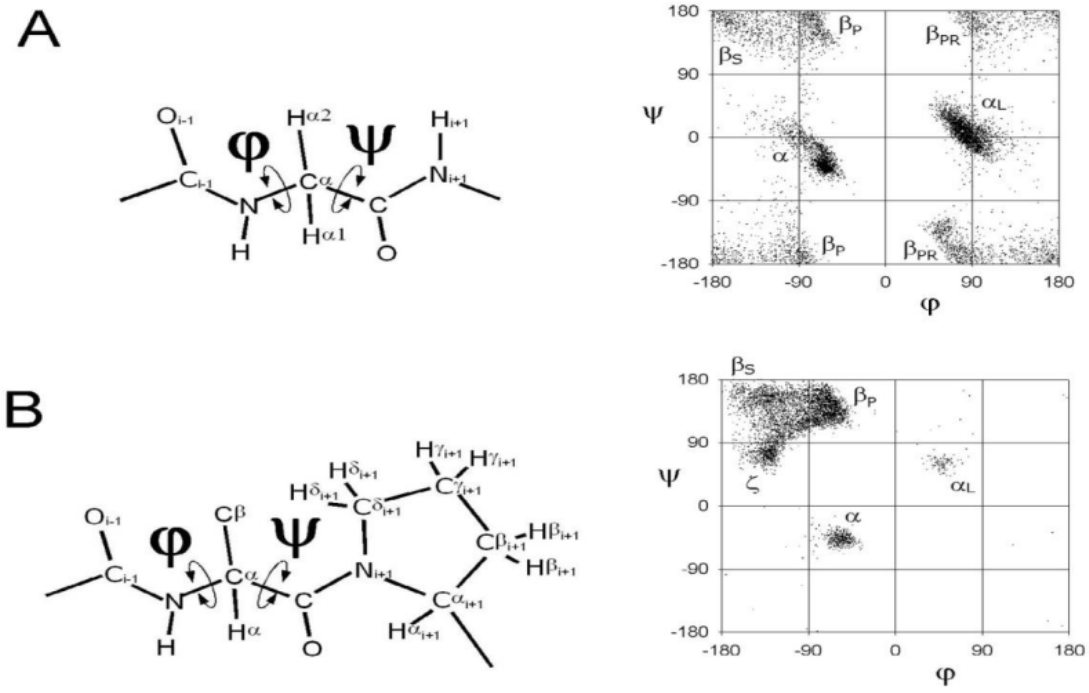
The observed glycine map has 5 regions of density . In order to display the observed density in one continuous region, we shift the coordinates from $\varphi$-$\psi$ to $\varphi'$-$\psi'$ where $\varphi'$: $0° < \varphi' < 360°$, and $\psi'$: $-90° < \psi' < 270°$.

With the shifted glycine Ramachandran plot (Figure 3A), we can clearly identify the different regions. Along the horizontal strip $\psi' \sim 180°$, there are three separate regions. One of these is an elongated version of the $\beta_P$ region of the generic Ramach- andran plot. The $\beta_P$ region corresponds to the polyproline II structure, which forms an extended left-handed helix along the protein chain . The $\beta_{PR}$ region is a reflection of the $\beta_P$ region where a sequence of glycine residues in the $\beta_{PR}$ conformation will form a right-handed helix. Finally, there is a region that corresponds to the $\beta_S$ region of the generic Ramachandran plot. This region corre- sponds to the extended conformation of

residues in β- sheets. However, the glycine $\beta_S$ region, centered on $(\varphi', \psi') = (180°, 180°)$, is slightly displaced from the $\beta_S$ region of the generic Ramachandran plot. There is also the diagonal α and $\alpha_L$ regions (Figure 3A), which are associated with helices and turns . Unlike the generic Ramachandran plot, the glycine α region is symmetric to the $\alpha_L$ region . In the generic Ramachandran plot, there is also a γ region corresponding to the hydrogen bonded γ-turn . The glycine Ramachandran plot does not have any density in the γ region.

***Steric interactions in glycine***

The original steric map of glycine (Figure 2A) fails to explain large parts of the observed glycine Ramachandran plot (Figure 1A). In the observed glycine Ramachandran (Figure 3A), there are two large excluded horizontal strips at $50° < \psi' < 120°$ and $-120° < \psi' < -50°$, which are not excluded in the glycine steric map (Figure 2A). Con- versely, the glycine steric map excludes a horizontal strip at $-30° < \psi' < 30°$ (Figure 2A), but this region is populated in the observed plot (Figure 1A). There are also diagonal steric boundaries in the observed glycine Ramachandran plot (Figure 1A), whereas the steric map predicts vertical boundaries (Figure 2A).



**Backbone conformations of glycine and pre-proline.** Backbone schematic (left) and observed Ramachandran plot (right). of (A) glycine and (B) pre-proline. Taken from the data-set of Lovell et al. (2003). The clustered regions are labeled on the Ramachandran plots.