Ramachandran Plots. Amino Acid Configuration in Proteins

Introduction

The secondary structures that polypeptides can adopt in proteins are governed by hydrogen bonding interactions between the electronegative carbonyl oxygen atoms and the electropositive amide hydrogen atoms in the backbone chain of the molecule. These hydrogen-bonding interactions can form the framework that stabilizes the secondary structure. Many secondary structures with reasonable hydrogen bonding networks could be proposed but we see only a few possibilities in polypeptides composed of L-amino acids (proteins). Most of the possible secondary structures are not possible due to limits on the configuration of the backbone of each amino acid residue. Understanding these limitations will help you to understand the secondary structures of proteins.

Most biochemistry textbooks briefly mention this important topic but I have found that many students need more examples than are provided in the textbook for a good understanding of this subject. The purpose of this tutorial is to provide further examples of structural limitations and a detailed overview of the regions in the set of possible amino acid configurations that are allowed and disallowed. This set of values is often graphically represented as a Ramachandran diagram.

The topics in this tutorial are covered in the textbook: Stryer, chapter 3 (3.1, 3.2, 3.3); Bruice, chapter 21 (21.1, 21.7, 21.13).

Assumed Knowledge

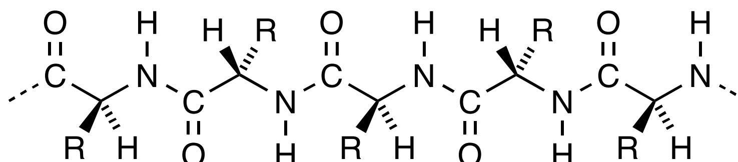
This tutorial assumes the following…

1. That you understand the function groups relevant to amino acids: carboxylic acids, amines, alcohols, thiols, nitrogen heterocycles (imidazole, indole), guanidine groups and amides. **It is especially important to be familiar with amides.** (Stryer, 3.2; Bruice, 21.7). See the Organic Chemistry Review for more on the subject.
2. That you are familiar with the structures of amino acids and proteins. It is especially important that you are familiar with the primary and secondary structures of proteins. (Stryer, chapter 3.1 to 3.3; Bruice, chapter 21.1, 21.2, 21.12, 21.13).
3. That you are familiar with the conventional representation of molecular structures of amino acids and peptides. This includes perspective Kekule structures and Fischer projections for depicting the 3-dimensional structures of molecules. (Stryer, pages 16 to 17; Bruice, chapter 2; see the Fischer Projection Tutorial and the Organic Chemistry Review).

Amide Linkages in Peptides

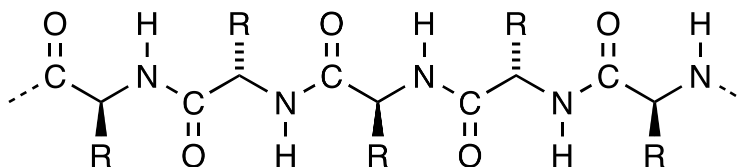
Below is a typical graphic representation of a polypeptide chain in a protein. The R groups are the side chains of the amino acids. The amide bonds are the linkages between the individual amino acids. You must be able to recognize the amide linkages in a peptide.

Figure 1. Perspective Kekulé representation of a polypeptide.



I will usually draw the above molecule is a graphically simpler manner by not writing out the hydrogen atoms on the α-carbons. It is still there and you should always remember that it must exist.

Figure 2. Perspective Kekulé representation with implied hydrogen atoms.



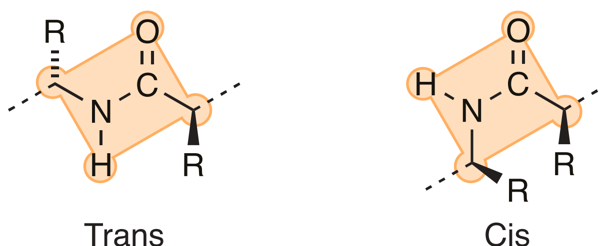
There are many bonds in the amide backbone of this peptide, and many, many more in a protein. With all those degrees of freedom (each bond that can rotate is a degree of freedom) we would expect a near infinite variety of secondary structures in proteins. Yet we see only 2 major classes of defined secondary structure: the α-helix and the β-sheet. What's up with that?

Limits on Rotation in Amides

We do not have as many degrees of freedom as we have bonds. One type of bond in the backbone does not rotate freely. The amide bonds are held in a planar orientation due to orbital overlap between the π-orbitals of the SP_2 hybridized atoms of the carbonyl group and the π-orbital of the SP_2 hybridized amine group (Stryer, chapter 3.2.2). This distributed π-orbital system resists rotation, as the orbital overlap would have to be broken. The distributed nature of the system also makes the amide bond resistant to hydrolysis.

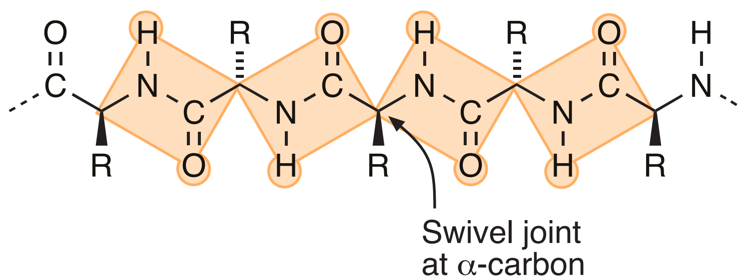
The amide group will prefer to have all atoms connected to it in the same plane as the amide bond (no rotation). This can result in a 'cis' or 'trans' configuration. Due to steric and electronic reasons the 'trans' configuration is the predominant form for the amide bond. Figure 3 shows the two possible planar configurations. Note the steric clash between the side chain groups in the 'cis' form.

Figure 3. Planar amide bonds.



So we can look at a polypeptide chain as a series of planar structures connected by swivel joints at the α-carbons of each amino acid residue. The two atoms of each amide bond and the four atoms connected to them are coplanar for each individual amide bond. Any two planar amide groups share a common atom – the α-carbon of an amino acid residue.

Figure 4. Polyamide with planar amide bonds indicated.



These swivel joints have two connections that can rotate freely. Like a universal joint in the drive train of a car, these two circular (cylindrical) rotations can allow for a fully spherical range of motion. We will consider each of these two rotating bonds separately.

The Ramachandran Angles

Each residue will have two bonds that can rotate freely. These two angles define the conformation of that residue in a protein and are called the Ramachandran angles, ψ (psi) and ϕ (phi).

The ψ (psi) Angle

The bond from the α-carbon to the carbonyl group (at the C-terminus) of the amino acid residue can rotate and turn the whole plane of the amide group, which includes the carbonyl carbon, in a 360-degree range. This angle is measured by looking along that bond with the carbon of the carbonyl group in the rear and the α-carbon to the front. We measure the apparent angle between the two bonds to nitrogen that you can see coming out of the axis of the $C_\alpha \rightarrow C_{(C=O)}$ bond. This angle is labeled ψ (psi) and is measured from –180° to +180° with the positive direction being when you turn the rear group clockwise so that the rear nitrogen bond is clockwise of the front nitrogen bond (or when you turn the front group counterclockwise so that the rear nitrogen bond is clockwise of the front.)

Figure 5. The ψ (psi) angle in an amino acid residue. (In these examples ϕ = 180˚)

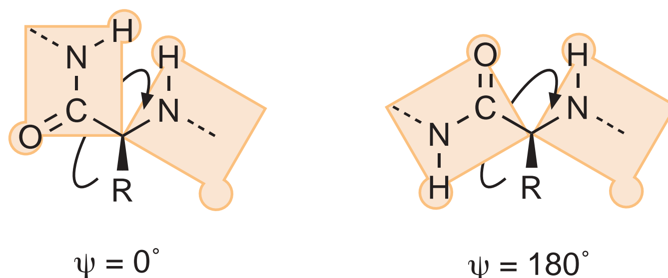
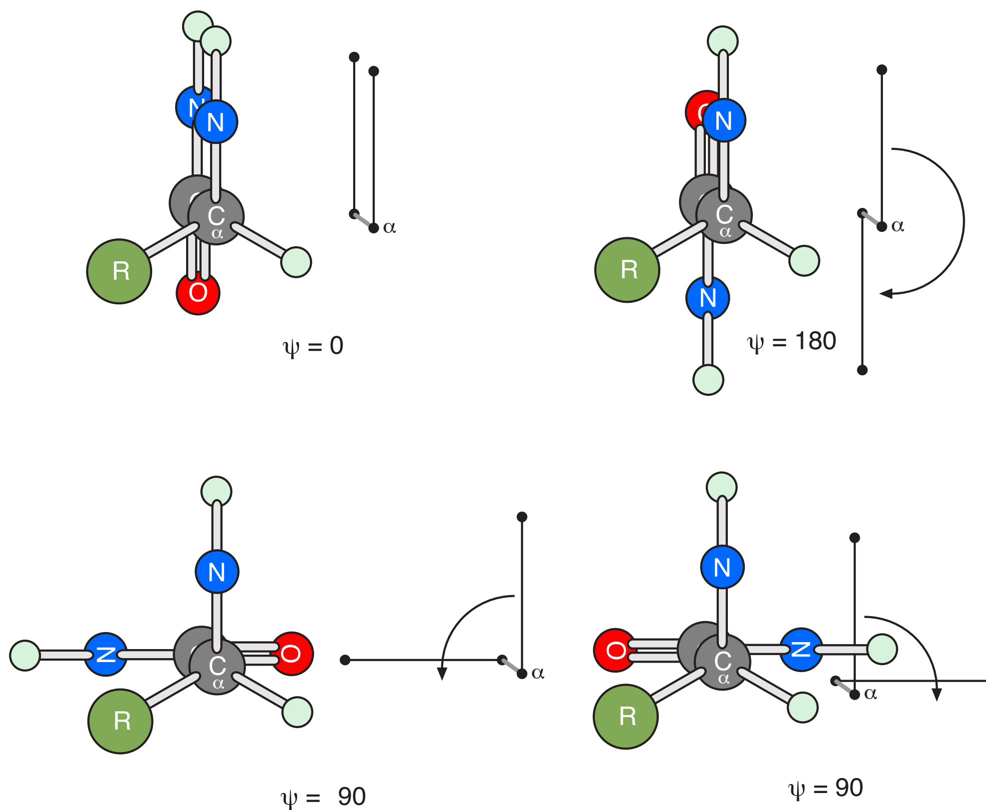


Figure 6. The ψ (psi) angle seen along the $C_\alpha \rightarrow C_{(C=O)}$ axis. (In these examples ϕ = 180˚)



The ϕ (phi) Angle

The bond from the nitrogen (at the N-terminus) to the α-carbon of the amino acid residue can rotate and turn the whole plane of the other amide group, which includes the nitrogen, in a 360-degree range. This angle is measured by looking along that bond with the nitrogen atom in front and the α-carbon to the rear. We measure the apparent angle between the two bonds to the carbonyl carbons that you can see coming out of the axis of the $N \rightarrow C_\alpha$ bond. This angle is labeled ϕ (phi) and is measured from –180˚ to +180˚ with the positive direction being when you turn the rear group clockwise so that the rear

carbonyl bond is clockwise of the front carbonyl bond (or when you turn the front group counterclockwise so that the rear carbonyl bond is clockwise of the front.)

Figure 7. The ϕ (phi) angle in an amino acid residue. (In these examples $\psi = 180^\circ$)

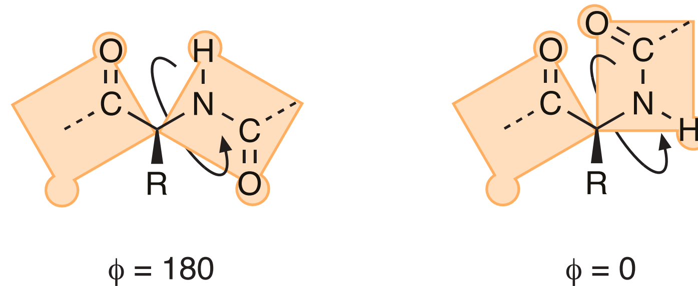
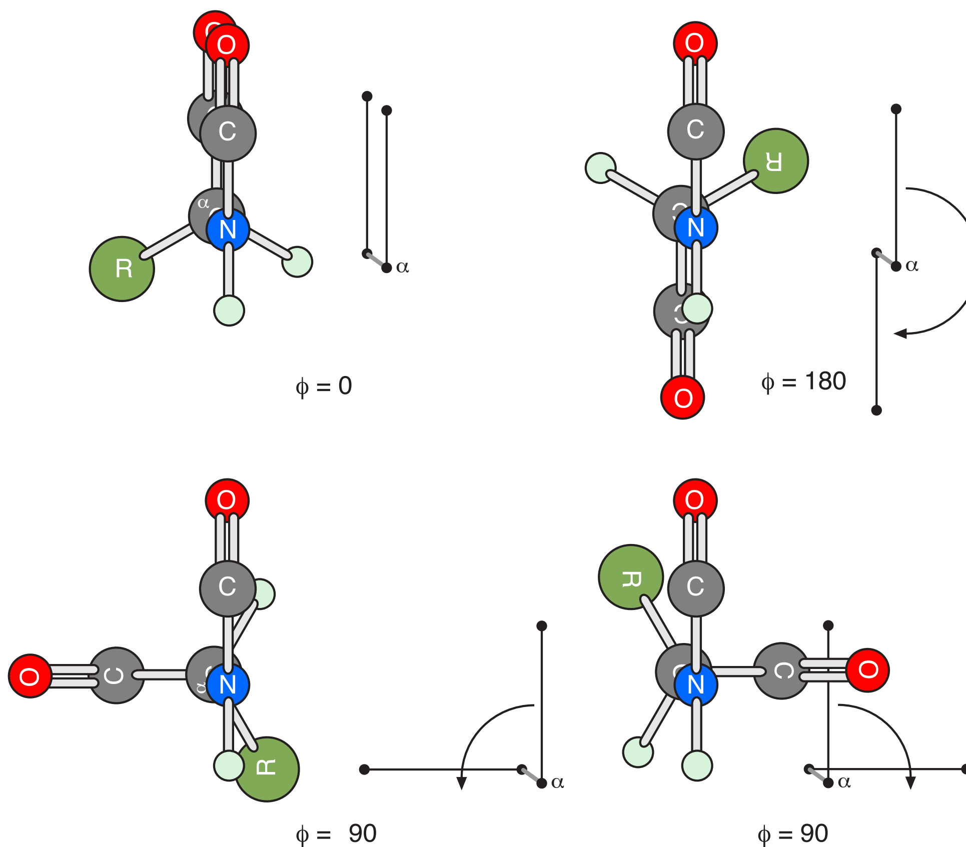


Figure 8. The ϕ (phi) angle seen along the $N \rightarrow C_\alpha$ axis. (In these examples $\psi = 180^\circ$)



The Ramachandran Plot

We can vary ψ from –180˚ to 180˚ and we can vary ϕ from –180˚ to 180˚ (that is 360˚ of rotation for each). But many combinations of these angles are almost never seen and others are very, very common in proteins.

Let us plot the values of ψ vs. the values of ϕ for an example globular protein. We will obtain a data set for the positions of each atom in space. We can get such data from one

of the several depositories of protein structural data. We will use X-ray crystallography data, as it is the most precise. (Although it may not be completely accurate, as crystal-packing forces usually distort proteins slightly. NMR data for proteins in solution are not very precise but are considered to be more accurate as the protein is in its native environment.)

We can take this data and detect the amino acid residues (this is usually done for us as the X-ray crystallography ".PDB" data format labels all atoms and assigns them to residues). Next we use a computer to compute the dihedral angles that define ψ and ϕ. A university course in linear algebra will show you many easy ways to do this but we are biochemists so lets just buy a software program rather than write one.

So a computer program can take a ".PDB" file and report the of ψ and ϕ angles for each and every residue. Lets plot these values for the protein hexokinase from yeast. A plot of ψ vs. ϕ is called a Ramachandran plot.

Figure 9. Ramachandran plot for hexokinase from yeast.

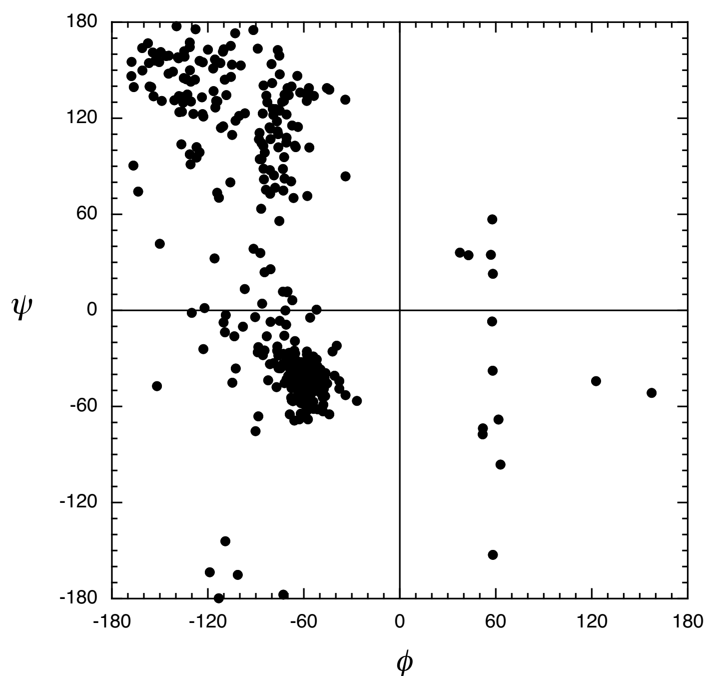
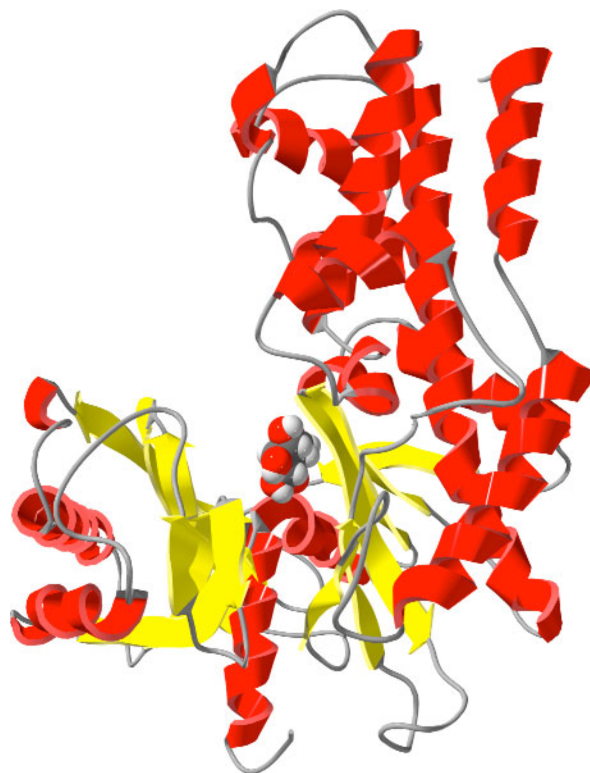


Figure 10. Ribbon diagram of hexokinase from yeast with glucose substrate in the active site.

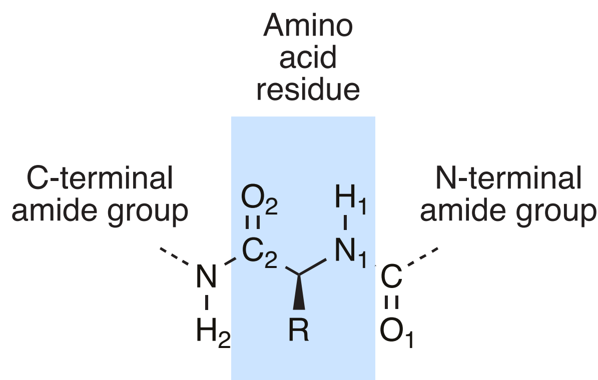


Observe the Ramachandran plot in figure 7. There are large areas where there is no examples of certain combinations of ψ and ϕ. Why is this so?

Steric Limits of ψ and ϕ

Atoms take up space and cannot occupy the same space at the same time. Covalent bonds connect them and these bonds cannot be broken. So the movements that we are concerned with involve rotations only. And we have seen previously that there are only two angles that rotate in a given residue, ψ and ϕ. But not all combinations are possible due to physical clashes of atoms in 3-dimensional space. These physical clashes are called steric interactions and the limit the available values for ψ and ϕ.

Figure 11. Labeling of atoms for declaring steric clashes for a given amino acid residue.



Case 1.

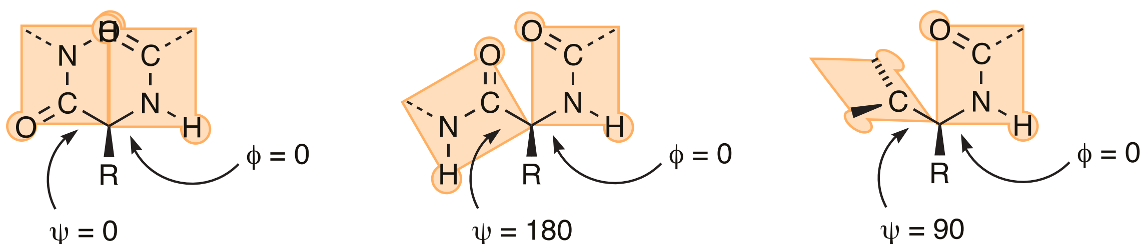
Let us **set ϕ to zero and vary ψ**. Figure 12 shows some examples of ψ angles with $\phi = 0^\circ$.

Observe that when **ψ is near 0°** we have very serious steric clashes between the N-H group of the amino acid residue attached to the C-terminus of the amino acid residue in question (its position in space is controlled by ψ) and the carbonyl group attached to the N-terminus (its position in space is controlled by ϕ). This clash is labeled an **O_1-H_2 clash** (for the backbone O on one residue clashing with the backbone amide H in the next residue due to the ψ and ϕ angles).

Observe that when **ψ is near 180°** we have very serious steric clashes between the oxygen of the carbonyl group at the C-terminus of the residue and the carbonyl group attached to the N-terminus. This clash is labeled an **O_1-O_2 clash** (for the backbone O on one residue clashing with the backbone O in the next residue due to the ψ and ϕ angles).

Observe that for **any value of ψ** we will have significant steric clashes with the carbon atom of the carbonyl group at the C-terminus and the carbonyl group attached to the N-terminus. This clash is labeled an **O_1-C_2 clash** (for the backbone O on one residue clashing with the backbone C of the carbonyl group in the next residue due to the ψ and ϕ angles).

Figure 12. An amino acid residue with various ψ angles with $\phi = 0^\circ$. From left to right: O_1-H_2 clash; O_1-O_2 clash; O_1-C_2 clash.



But we can adjust the ϕ angle to avoid these steric clashes. If we remove the oxygen and the amine of the amide group, we will have just a methyl group left. There will be no possibility of a O_1-H_2 clash or a O_1-O_2 clash. We can now explore the range of ϕ over which the O_1-C_2 clash is present. We can see from Figure 13 that the will be an O_1-C_2 clash when ϕ = 0˚ but not when ϕ = 180˚. In figure 14 we can see the approximate range over which the O_1-C_2 clash is significant.

Figure 13. A model compound for exploring forbidden ϕ angles due to O_1-C_2 clash.

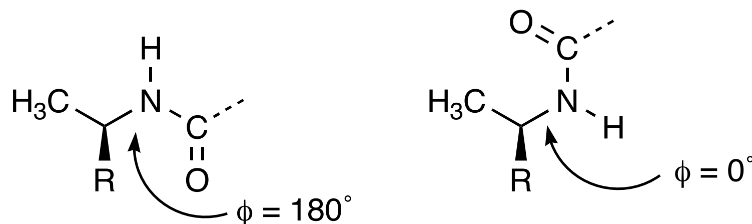


Figure 14. Examining the range of ϕ angles that give a significant O_1-C_2 clash.

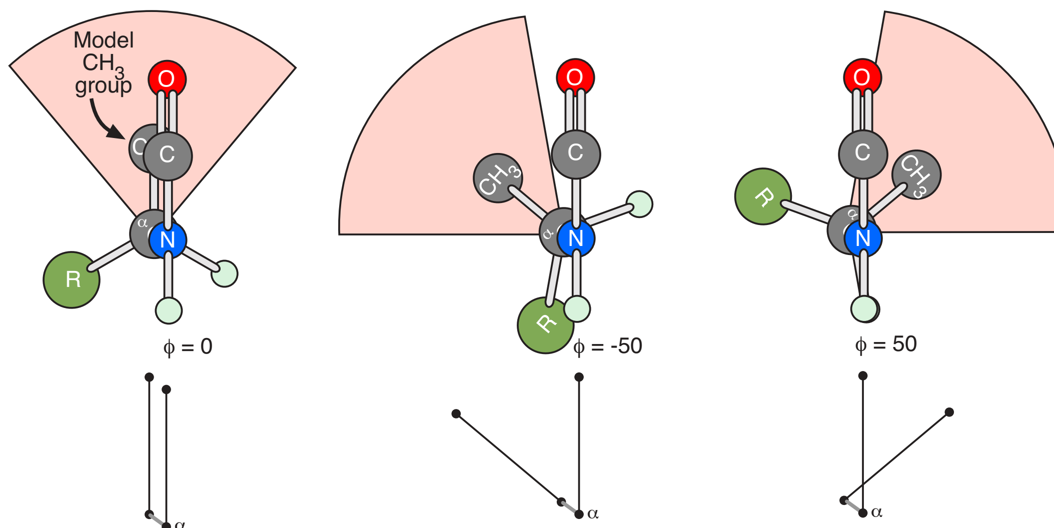
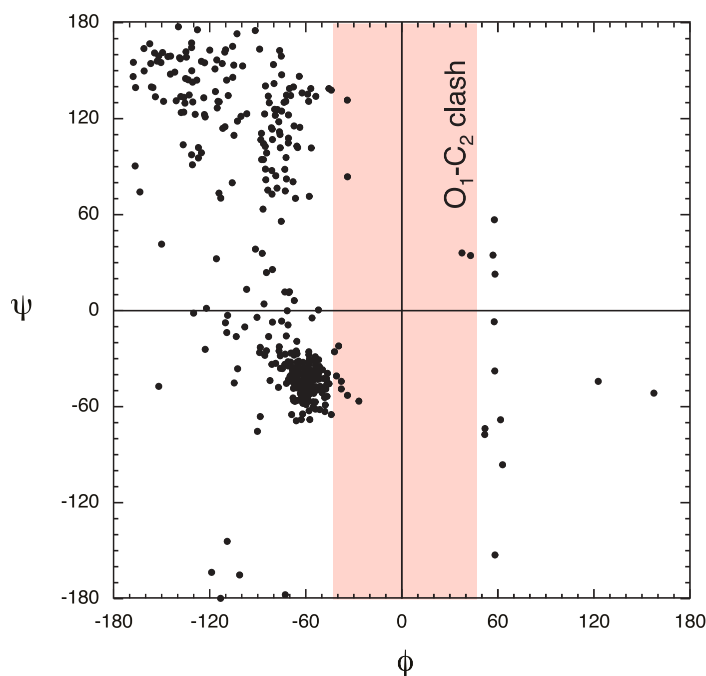


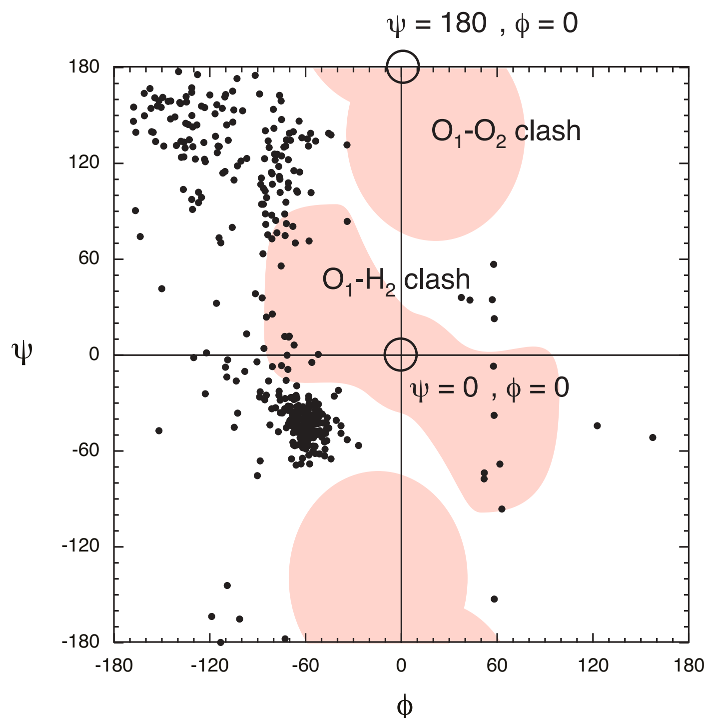
Figure 15. Ramachandran plot with O_1-C_2 clash region indicated.



Due to the size of the carbon atom, which is bigger than pictured relative to the bond lengths (Van der Waals radius, see Stryer, pages 16 and 17; Bruice, 2.12, 2.13), we have an **O_1-C_2 clash between about ϕ = -45˚ and 45˚**. Let us apply this sterically forbidden region to our Ramachandran plot from figure 9. The result is figure 15.

For the **O_1-H_2 clash** and the **O_1-O_2 clash** there will be regions of the plot where a combination of ψ and ϕ values will result in steric collisions. As seen in figure 16 the O_1-H_2 clash region includes $\psi = 0˚$ and $\phi = 0˚$ and the O_1-O_2 clash region includes $\psi = 180˚$ and $\phi = 0˚$. Figure 12 shows the structures that correspond to these points.

Figure 16. Ramachandran plot with the O_1-H_2 clash and the O_1-O_2 clash regions indicated.



So we have explored the region along the $\phi = 0^\circ$ axis and we have determined sterically forbidden regions associated with that axis. Now let us explore the $\psi = 0^\circ$ axis and vary the value of ϕ to find forbidden regions.

Case 2.

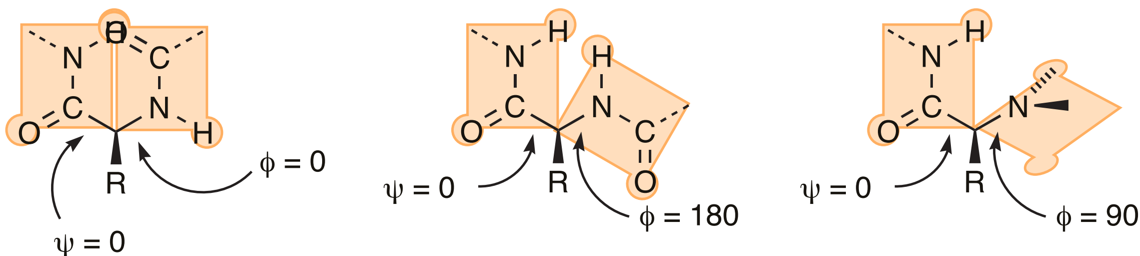
Let us **set ψ to zero and vary ϕ**. Figure 17 shows some examples of ϕ angles with $\psi = 0^\circ$.

Observe that when **ϕ is near 0°** we the exact same O_1-H_2 clash observed in case 1. The ϕ and ψ angles are both 0°, the same as in the example in case 1.

Observe that when **ϕ is near 180°** we have very serious steric clashes between the hydrogen on the amide nitrogen of the C-terminus of the residue and the hydrogen on the amide nitrogen on the N-terminus of the residue. This clash is labeled an **H_1-H_2 clash** (for the backbone H on one residue clashing with the backbone H in the next residue due to the ψ and ϕ angles.)

Observe that for **any value of ψ** we will have significant steric clashes with the hydrogen on the nitrogen atom of the amide group at the C-terminus and the nitrogen attached to the N-terminus. This clash is labeled an **N_1-H_2 clash** (for the backbone N on one residue clashing with the backbone H of the amide group in the next residue due to the ψ and ϕ angles).

Figure 17. An amino acid residue with various ϕ angles with $\psi = 0^\circ$. From left to right: O_1-H_2 clash; H_1-H_2 clash; N_1-H_2 clash.



But we can adjust the ψ angle to avoid these steric clashes. If we remove the carbonyl group of the amide group, we will have just an amine left. There will be no possibility of a O_1-H_2 clash or a H_1-H_2 clash. We can now explore the range of ψ over which the N_1-H_2 clash is present. We can see from Figure 17 that the will be an N_1-H_2 clash when $\psi = 0^\circ$ but not when $\psi = 180^\circ$. In figure 19 we can see the approximate range over which the O_1-C_2 clash is significant.

Figure 18. A model compound for exploring forbidden ψ angles due to N_1-H_2 clash.

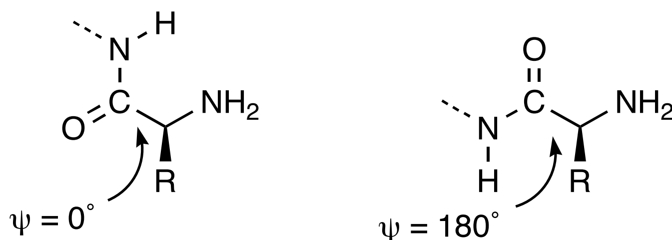
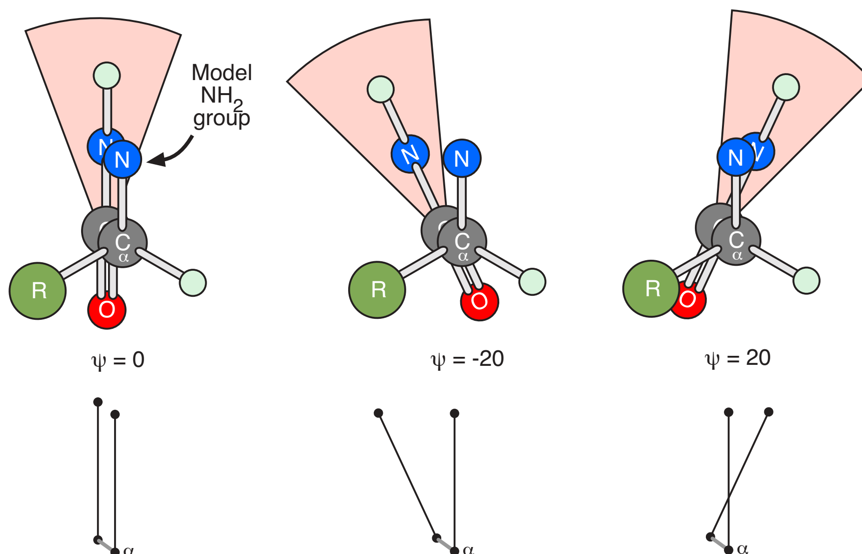
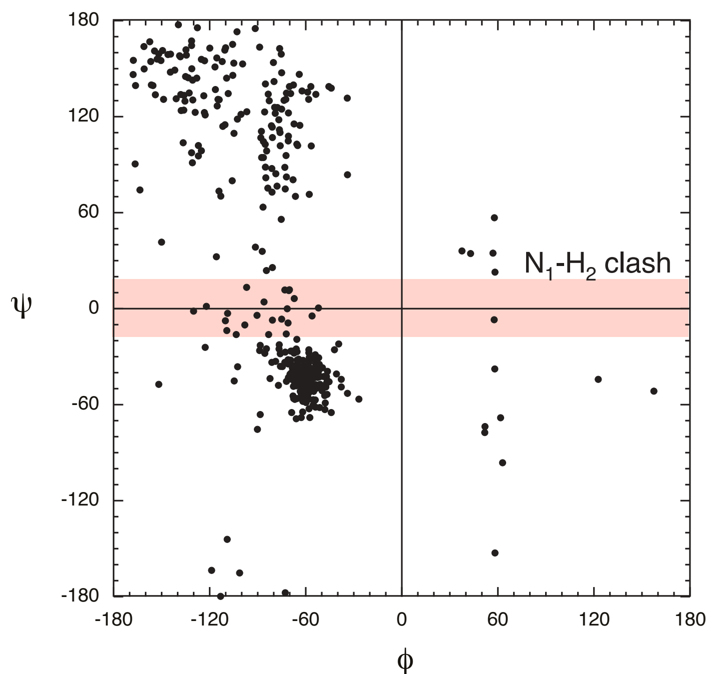


Figure 19. Examining the range of ψ angles that give a significant N_1-H_2 clash.



The hydrogen atom is smaller than a carbon atom so there is less steric hindrance involved. We have an **N_1-H_2 clash between about ψ = -20˚ and 20˚**. Let us apply this sterically forbidden region to our Ramachandran plot from figure 20.

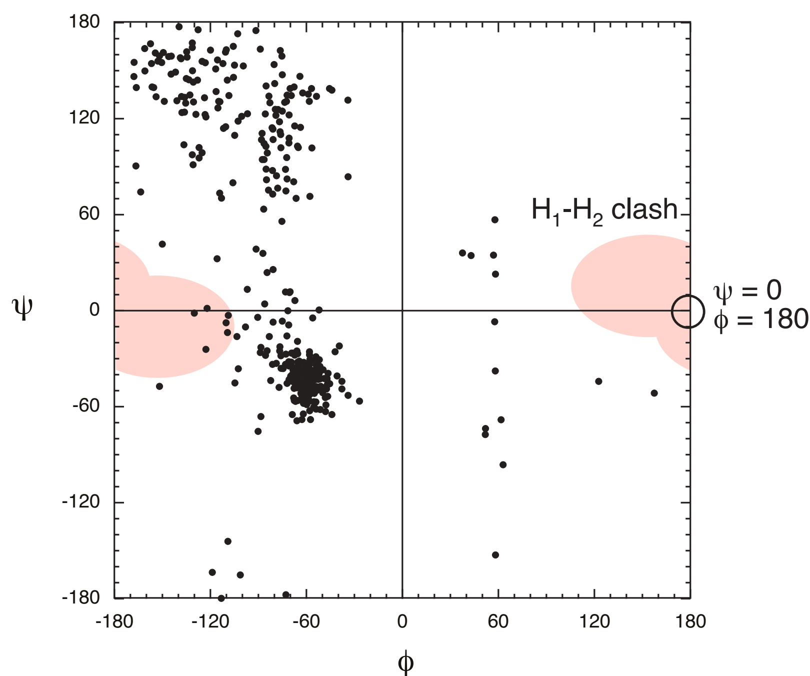
Figure 20. Ramachandran plot with N_1-H_2 clash region indicated.



For the H_1-H_2 clash there will be regions of the plot where a combination of ψ and ϕ values will result in steric collisions. As seen in figure 17 and 18, the **H_1-H_2 clash** region includes ψ = 0˚ and ϕ = 180˚.

The O_1-O_2 clash region includes ψ = 180˚ and ϕ = 0˚ (identical to case 1, see figure 16).

Figure 21. Ramachandran plot with the H_1-H_2 clash region indicated.



Peptide Backbone-Backbone Interactions

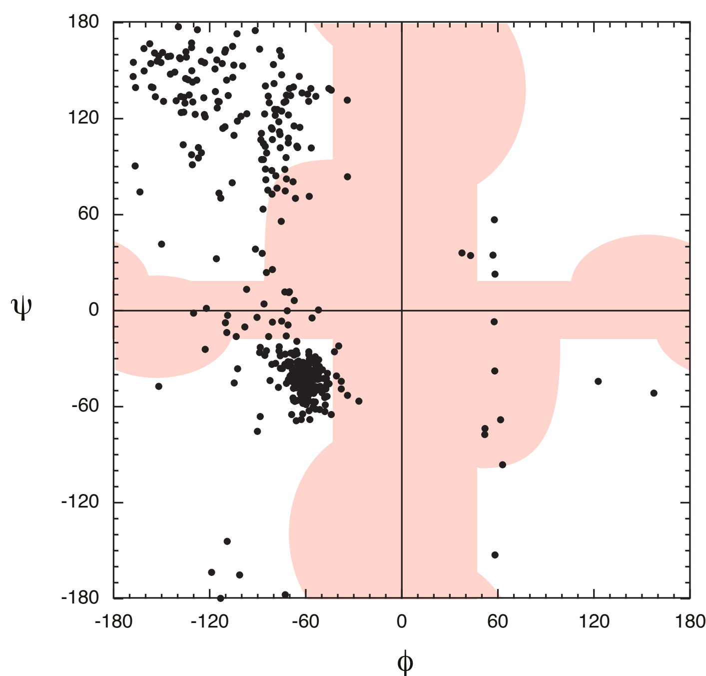
So we have determined the regions of the Ramachandran plot where the values of the ψ and ϕ angles result in steric clashes. If we combine the forbidden regions from figures 12, 13, 17 and 18 we will see the region of the plot that is forbidden **due to interactions between atoms in the peptide backbone**.

Figure 22 shows the forbidden regions for combinations of ψ and ϕ angles for backbone-backbone steric clashes. If we had a polypeptide composed entirely of glycine residues we would expect this description of conformational restrictions to define the regions where we would observe data points on the plot. However, we observe that the right-hand side and the bottom third of the Ramachandran plot for hexokinase (figure 9 and 19) do not contain very many examples of ψ and ϕ angle combinations in these regions. What's up with that?

Glycine is the only achiral (not chiral) amino acid. There is no side chain group. Since glycine is symmetrical, we would expect to observe a symmetrical set of steric limitations in a Ramachandran plot. Examine the plot in figure 22. Are the forbidden regions symmetrical?

Every other amino acid is not symmetrical. There is a side chain group on the α-carbon. This side chain group has the potential to clash with the hydrogen atom on the amide group at the C-terminus of a residue (R-H_2 clash) or the oxygen of the amide group of the N-terminal of a residue (O_1-R clash).

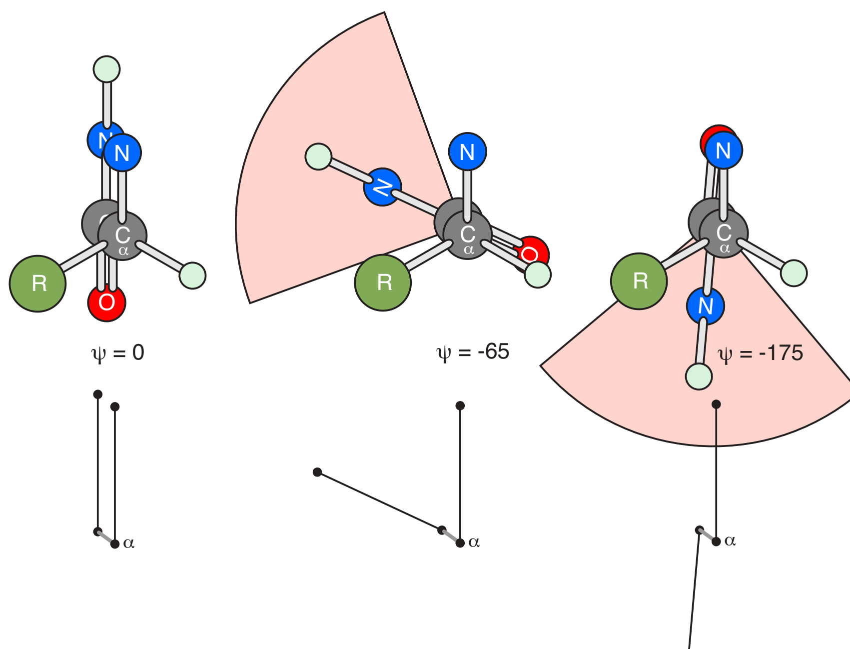
Figure 22. Ramachandran plot with backbone-backbone clash region indicated.



Case 3.

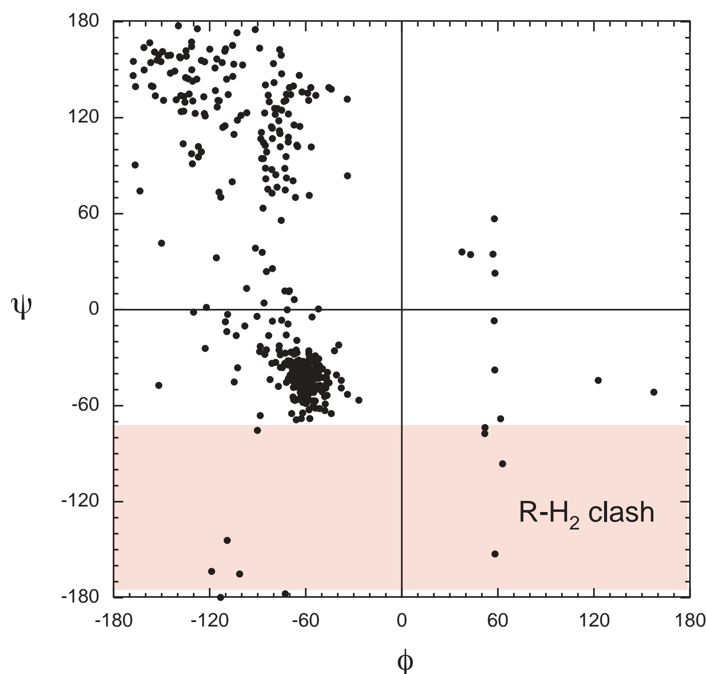
Let us vary ψ and determine what values result in a **R-H_2 steric clash**. R is the side chain and H_2 is the hydrogen on the amide group of the C-terminus. Figure 23 shows some examples of ψ angles using the model compound from figures 15 and 16.

Figure 23. R-H_2 steric clashes as the ψ angle is varied.



The R group can be bulky or very bulky depending on the amino acid. For an average we consider that it excludes an arc of 90˚ due to steric interactions with the amide hydrogen. We observe from figure 23 that the ψ angles that result in significant $R\text{-}H_2$ steric clash interactions are the range from about ψ = -65˚ to ψ = -175˚.

Figure 24. Ramachandran plot with the $R\text{-}H_2$ clash region indicated.



Case 4.

Let us vary ϕ and determine what values result in a **O_1-R steric clash**. R is the side chain and O_1 is the oxygen on the amide group of the N-terminus. Figure 25 shows some examples of ϕ angles using the model compound from figures 10 and 11.

As in case 3, the R group can be bulky or very bulky depending on the amino acid. For an average we consider that it excludes an arc of 90˚ due to steric interactions with the amide oxygen. We observe from figure 25 that the ϕ angles that result in significant O_1-R steric clash interactions are in the range from about ϕ = 65˚ to ϕ = 175˚.

Figure 25. O_1-R steric clashes as the ϕ angle is varied.

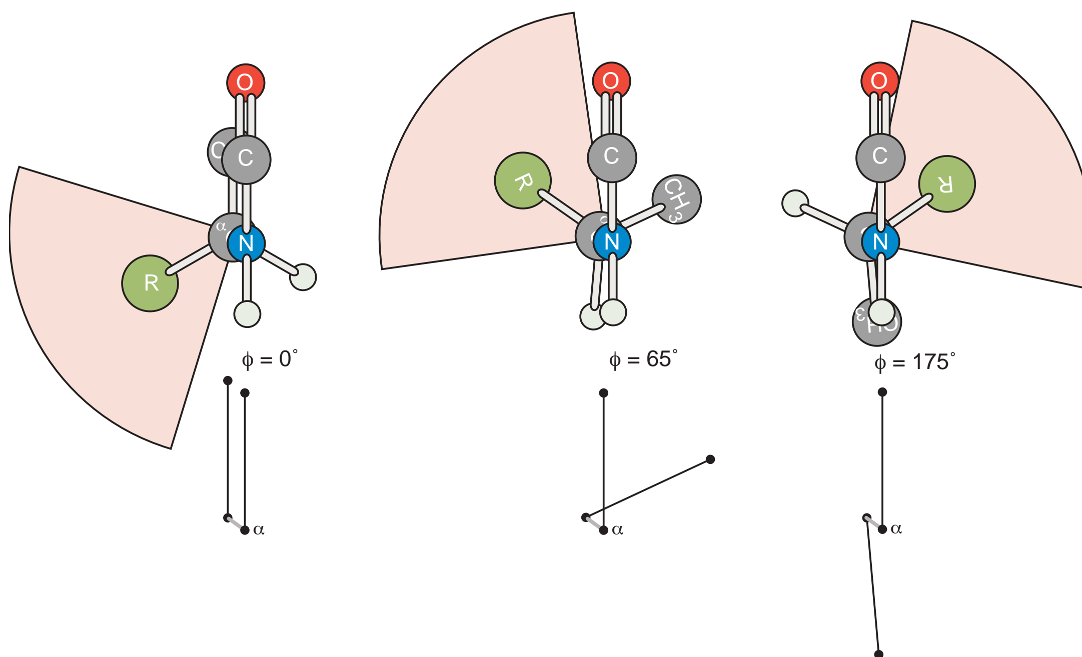
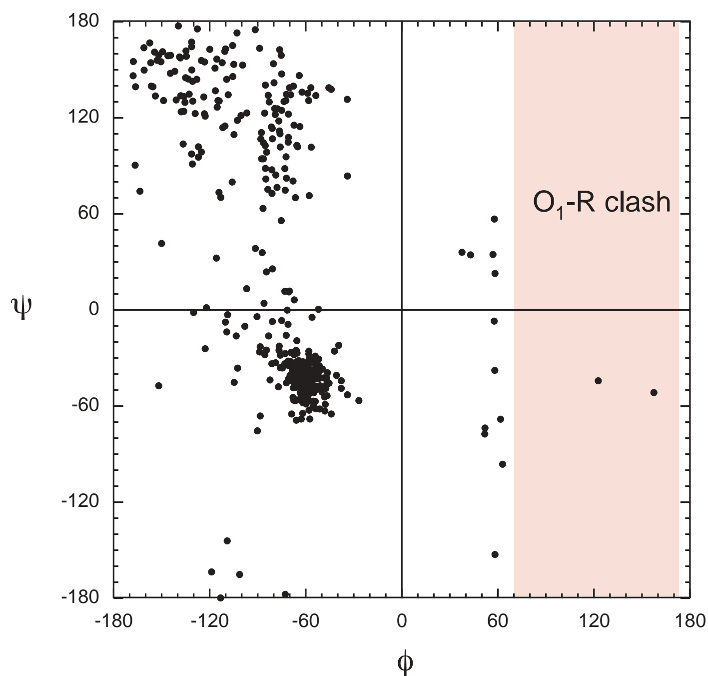


Figure 26. Ramachandran plot with the O_1-R clash region indicated.

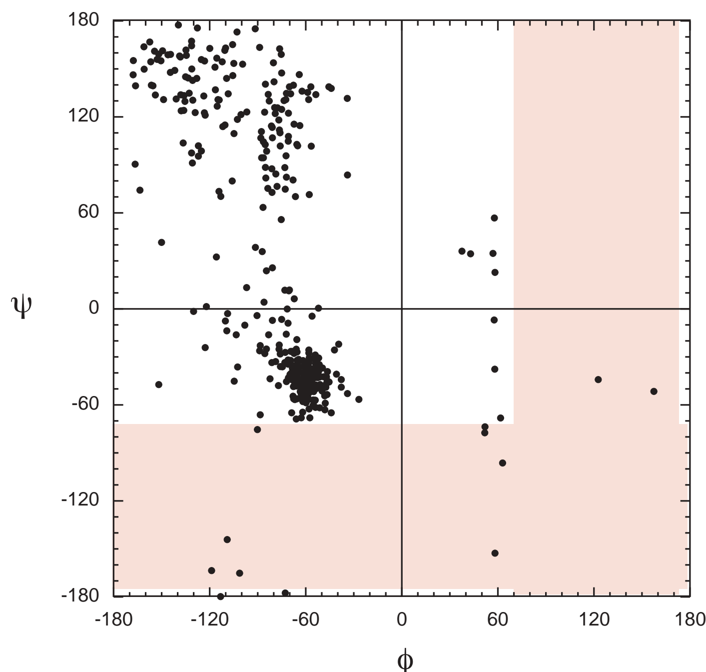


Peptide Sidechain-Backbone Interactions

Let us combine the forbidden regions for backbone-sidechain interactions (R-H_2 clashes and O_1-R clashes) in a single plot. The result is shown in figure 27. The shaded region is the set of ψ and ϕ angle combinations that result in steric clashes. This region is not symmetrical with respect to the axes of the plot.

The plot is not symmetrical because the chiral amino acid residue is not symmetrical. Can you imagine what the forbidden O_1-R and R-H_2 clash regions would look like for a polypeptide composed of R-amino acids?

Figure 27. Ramachandran plot with the O_1-R and R-H_2 clash regions indicated.



The Allowed Regions in a Ramachandran Plot

Now let us combine all the regions in a Ramachandran plot where the set of ψ and ϕ angle combinations result in steric clashes. The results are shown in figure 28.

Why are there data points in the forbidden regions? These regions represent high-energy combinations of ψ and ϕ angle but that doesn't mean that the energies are impossibly high. Slight twists in the planar amide bond can go a long way to help reduce steric clashes for certain forbidden ψ and ϕ angle combinations. Highly favourable intra-residue interactions can pull a residue's conformation into high-energy territory if the combined energy is at a minimum. Some amino acid side chains are less bulky and more forgiving. The majority of the of ψ and ϕ angle combinations observed will be in the allowed regions but some data points will appear in the forbidden regions under special local conditions in a polypeptide chain.

Figure 28. Finally! The combined forbidden regions in a Ramachandran plot.

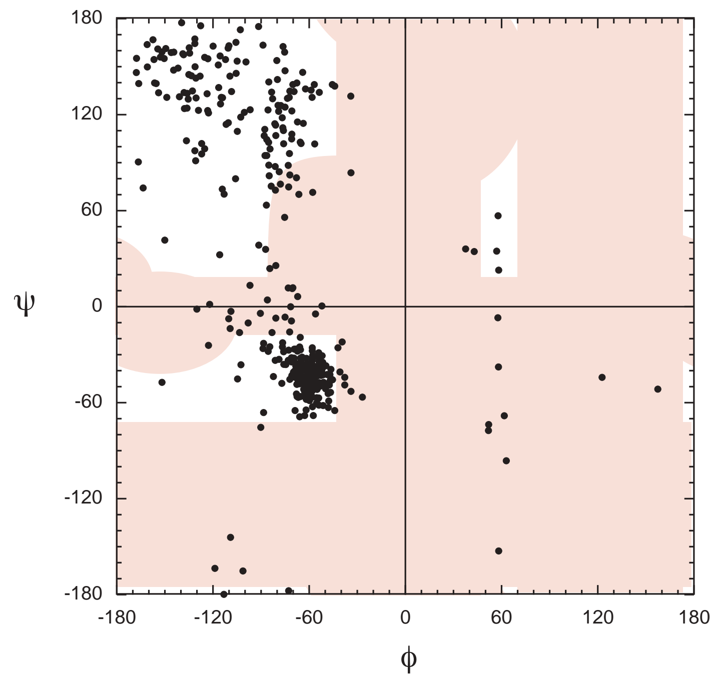
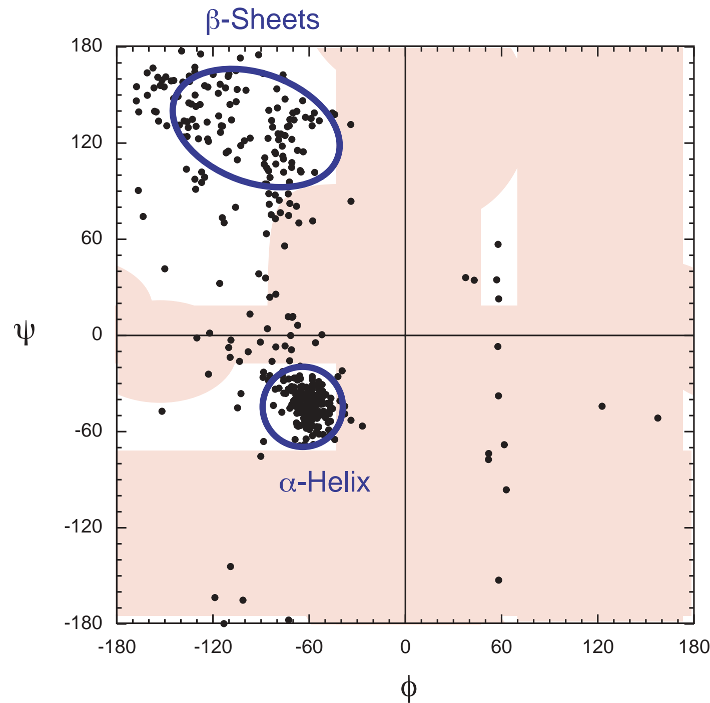


Figure 29. Ramachandran plot for hexokinase in yeast showing ψ and ϕ angle combinations for β-sheet and α-helix secondary structures.



Glycine residues do not have the steric constraints of the other 19 residues. In general practice, glycine residue data is not plotted in Ramachandran plots. **Proline** residues are

conformationally restricted due to the ring being part of the backbone. In general practice, proline residue data is not plotted in Ramachandran plots.

So let us examine the allowed regions in the Ramachandran plot for hexokinase in yeast. The β-sheet structures feature ψ and ϕ angle combinations that appear in the upper left corner of the plot and the α-helix regions have ψ and ϕ angle combinations that appear in the lower left quadrant as shown in figure 29.

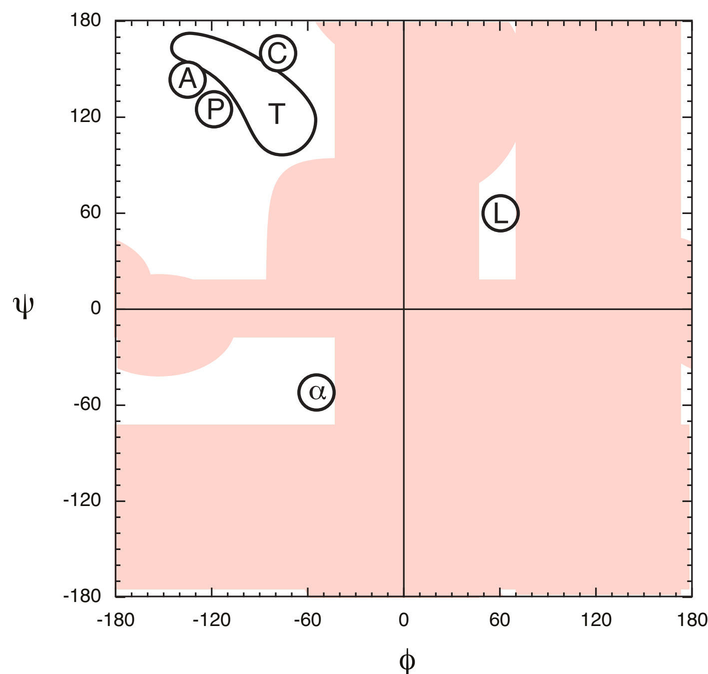
We can see from figure 29 that hexokinase has both α-helices and β-sheet regions. Why are the most common secondary structures observed in proteins α-helices and β-sheets? Because those structures have ψ and ϕ angle combinations that avoid serious steric clashes between the atoms of the amide backbone and side chain groups.

Conclusion

Understanding the steric limitations in individual amino acid residues reveals how these limitations result in the observed types of secondary structures found in nature. The complex folding of proteins is controlled, in a large part, by the limitations on the ψ and ϕ angles available to each amino acid residue.

If one inspects the ψ and ϕ angle combinations that exist in idealized secondary structures we can place the sites in a Ramachandran plot where we would expect to see data related to these structures. In a real protein we would expect data points from these defined secondary structures to appear near the ideal location.

Figure 50. Idealized secondary structures located on a Ramachandran plot.



In figure 50 we see the ψ and ϕ angle combinations located for the following secondary structures: **A**, antiparallel β-sheet; **P**, parallel β-sheet; **T**, right-hand twisted parallel and

antiparallel β-sheets (the most common type, see Stryer, pages 59, 60); **α**, the right-handed α-helix; **L**, the left-handed α-helix (very rare); C, collagen triple helix. Compare the Ramachandran plot that we have created with those in Stryer, chapter 3.3.

A fun observation is to recognize that a right-handed α-helix of polyglycine and a left-handed α-helix of polyglycine are mirror images of each other. They would be enantiomers and be identical in energy. We see from figure 50 that both are allowed. Look at figure 50 closely, what is the relationship between the α locations and the L location on the diagram? What symmetry operation would interchange the two? The answer is that they are symmetrical about a centre of inversion at the origin.

A right-handed α-helix and a left-handed α-helix in a natural protein are NOT mirror images of each other. They are not enantiomers and are not identical in energy. The right-handed α-helix is much more stable and is much more common in proteins. Why is this so? Consider the differences between glycine residues and the other amino acids in your answer.

Problems and Exercises

1. What is a condensation reaction? Give an example using two amino acids. Give an example of the reaction that is the net reverse of a condensation reaction. What kind of reaction is it?
2. Why is the rotation of the amide bond between two amino acid residues restricted?
3. Why is the pattern of forbidden regions in a Ramachandran plot for a polyglycine protein symmetrical? Why is the pattern of forbidden regions in a Ramachandran plot for a naturally occurring protein unsymmetrical?
4. Why is the variety of secondary structures in a protein limited?
5. A right-handed α-helix and a left-handed α-helix in a natural protein are NOT mirror images of each other. They are not enantiomers and are not identical in energy. Why are they not enantiomers? The right-handed α-helix is much more stable and is much more common in proteins. Why is this so?