

Protein Statistics

Scott Clark
University of California: Davis
Oregon State University

August 24, 2006

Abstract

Proteins are vital for the functionality of any organic system. The folding of proteins, based on physical laws, governs whether or not a protein will become useful or detrimental to an organism. Missfolded, disease causing proteins, or *prions*, are responsible for diseases such as alzheimers and mad cow disease. We will examine the statistics and properties of various proteins in order to determine characteristics that correspond to proper and low energy folds, as well as general information about the inherent structure of a stable protein.

1 Introduction

Biophysics (also *biological physics*) is an interdisciplinary science that applies the theories and methods of physical sciences, especially those of physics, to questions of biology.

This introduction verses the reader in the terminology of biophysics and outlines some key topics about proteins.

1.1 Protein Basics

Proteins are what makes an organism function. They carry out any number of tasks, from transporting oxygen through the blood stream to helping the immune system fight off infectious diseases. When a DNA is the code that

build the proteins, if DNA is a blueprint then the proteins are the house. Proteins are made up of amino acids. Each amino acid has a set of 3 DNA codons that correspond to it. As the DNA is being processed the amino acids are formed into chains, which are the proteins.

1.1.1 Amino Acids

Amino acids are the building blocks of proteins. There are many thousands of proteins, each made up of a unique string of amino acids, yet there are only 20 unique amino acids that occur in nature. Each amino acid has a basic form that allows it to bond with other amino acids, and a side chain, which allows it to add a special function or property to the protein. The basic form of an amino acid is shown in figure 1.

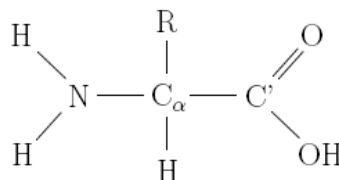


Figure 1: A single amino acid, where R is where the side chain is that determines which amino acid it is and what its properties are.

The side chain can range from a single hydrogen atom in Glycine to a complex aromatic ring in Tryptophan. A complete list of the naturally occurring amino acids can be found in the table in Appendix A. Amino acids are rarely in their solitary form, however, because they are found in proteins which are made up of many amino acids bound together.

1.1.2 Amino Acid Bonding

Two amino acids form a peptide bond when the amino group of one amino acid reacts with the carboxyl group of the other, producing water and the start of an amino acid chain (see figure 2).

These bonds occur many times over creating polypeptide chains with an average of 50 to 300 residues (amino acids). These polypeptide chains then fold into more complex structures; proteins.

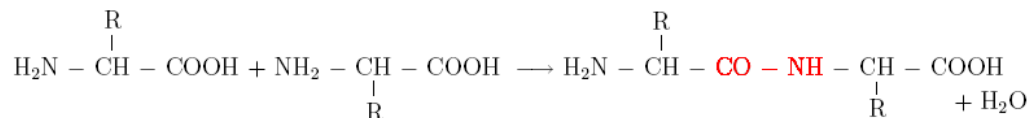


Figure 2: Two amino acids bonding, creating water

1.2 Protein Structure

Proteins fold into complex shapes that help determine their properties and function. The aspects of the proteins rather complex structure can be divided into three main: primary, secondary and tertiary structure.

1.2.1 Primary Structure

Primary structure is merely the linear sequence of amino acids that make up the polypeptide chain. This primary structure, or sequence, is unique to each protein and defines what the protein is.

1.2.2 Secondary Structure

Secondary structure is how the amino acids close to each other in the sequence interact with each other. This can be as simple as a linear chain of amino acids twisting into an α -helix so that the side chains can hydrogen bond with each other or as complex as a massive β -helix. A β -helix is a structure in which the polypeptide chain wraps back onto itself, tracing out a helix in 3-space (See Figure 3). This is an energetically preferred state because the backbones of the residues hydrogen bond with each other. This paper will be primarily concerned with the structure and properties of β -helices.

1.2.3 Tertiary Structure

Tertiary structure is a compilation of secondary structures. This is the form of the protein as a whole, a combination of α and β helices as well as other secondary structures.

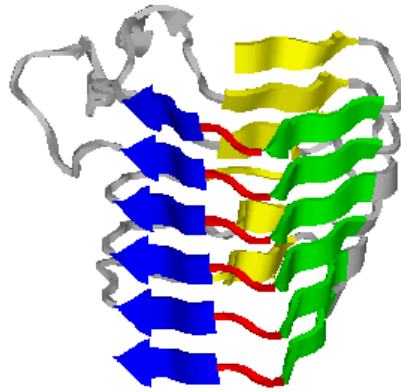


Figure 3: The β -helix 1PCL. The image only shows a cartoon of the backbone of the chain, we see that it traces out a helix. The arrows denote regions where the sequence is Beta, meaning there are backbone hydrogen bonds keeping the helix stable in these regions.

2 Techniques

The program that I wrote compiles various statistics about β -helical proteins.

2.1 Protein Database

Thousands of proteins are catalogued on an online protein database that is run by many different universities. Each protein database file (*.pdb file) contains information about the protein including name, discoverer, publication and other information. The most important information that we will use is the coordinate information. In this section of the file every atom of the protein (except hydrogen) has its coordinates explicitly defined. This allows us to rebuild a protein one piece at a time and compute statistics about it.

2.2 Statistics

My program calculates many statistics about both right and left handed helices including information about hydrophobicity (frustration), packing, Ramachandran statistics, stacking, hydrogen bonding, and general statistics.

2.2.1 Hydrophobicity

Each amino acid has identical backbone structure, only differing in the side chain that each carries. These side chains determine how each amino acid affects the protein as a whole. The side chains can be big or small, unreactive or highly reactive, susceptible to stabilizing bonds or detrimental to them. The aspect we will look at in this part is their hydrophobicity; whether it is more energetically acceptable for them to be near water (a polar molecule) which is abundant in the surroundings of the protein or buried within the helix where water cannot reach it. The amino acids can be divided up into a "periodic table" of polar (hydrophilic), non-polar (hydrophobic), acidic (hydrophilic), and basic (hydrophilic) side chains (See figure 1 in appendix A).

The protein, in its natural state, is emersed in an aqueous solution. This means that water surrounds it, and the hydrophobic residues will want to point inward, while the hydrophilic residues will want to point outward. This is the idea of minimum frustration: the residues want to be oriented in such a way that their individual energies are lower, but not at the expense of the energetics of the system as a whole.

My program runs through a *.pdb file for a protein and determines whether the residue is oriented inward or outward and then where it lies on the periodic table and calculates the overall statistics for the hydrophobicity of the protein. It calculates how many hydrophobic, hydrophilic, basic and acidic residues point in or out.

To determine whether an individual amino acid is oriented inwards or outwards a simple algorithm was devised. We use a clever combination of cross and dot products to determine whether or not the residue points inward. For all amino acids besides Glycine the side chain starts with a carbon denoted C_β . We make a vector that points from the C_α to the C_β , $\hat{v}_{\alpha\beta}$ and say that this is the direction that the side chain points. Now we need to determine if this is in or out of the protein. For this we make vectors connecting the C_α in question and the C_α atoms of the residues directly above ($\hat{v}_{\alpha+n}$), below ($\hat{v}_{\alpha-n}$), to the left ($\hat{v}_{\alpha+1}$) and to the right ($\hat{v}_{\alpha-1}$) of the residue in question (less if it is on the top or bottom rung of the helix) (See figure 4). We note that for a left handed β -helix that the following statements imply that $\hat{v}_{\alpha\beta}$ and thus the side chain points inwards.

$$(\hat{v}_{\alpha+n} \times \hat{v}_{\alpha+1}) \cdot \hat{v}_{\alpha\beta} > 0$$

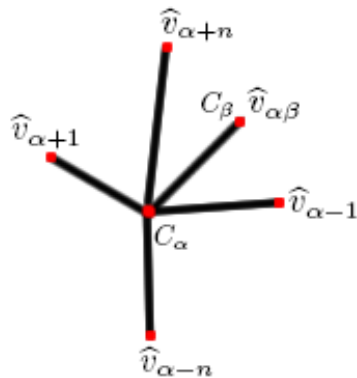


Figure 4: Vectors pointing from the C_α .

$$\begin{aligned}
 (\hat{v}_{\alpha-1} \times \hat{v}_{\alpha+n}) \cdot \hat{v}_{\alpha\beta} &> 0 \\
 (\hat{v}_{\alpha-n} \times \hat{v}_{\alpha-1}) \cdot \hat{v}_{\alpha\beta} &> 0 \\
 (\hat{v}_{\alpha+1} \times \hat{v}_{\alpha-n}) \cdot \hat{v}_{\alpha\beta} &> 0
 \end{aligned}$$

This is the exact opposite ($\neq 0$ and switch $\hat{v}_{\alpha-1} \leftrightarrow \hat{v}_{\alpha+1}$) for right handed helices. Using this technique we can determine whether or not a single amino acid is oriented inward or outward. Using this information we can see how many of the hydrophobic and hydrophilic side chains are pointing in or out respectively.

2.2.2 Packing

Packing determines how much of the interior volume of the protein is actually filled by the side chains. As we have seen the side chains come in all shapes and sizes. For a protein to be stable it must have sufficient packing so as to stand up to collisions with other proteins and molecules around it. A large protein made entirely of glycine (side chain is only a hydrogen) would be unstable in this regard.

To determine the volume of the amino acids we simply run through the sequence and determine whether the side chain points in or out (using the algorithm described above) and add up the respective volumes.

To determine the volume of the protein we use two techniques, a cylindrical approximation and a Monte Carlo integration.

The cylindrical approximation assumes that the protein conforms to the common triangular cylinder shape with an equilateral triangle as its base (each side consisting of 6 residues) and a pitch of 4.8 angstroms. The program finds the length of the longest side of the triangle (d_m) and then applies the formula,

$$V = \frac{\sqrt{3}}{4} d_m^2 \frac{\text{residues}}{18} 4.8A$$

To determine the approximated volume.

The second approximation is a Monte Carlo integration. The program first finds the range in 3-space that the helix lies (x_{min} , x_{max} , etc.). Then it applies a simple algorithm:

1. Choose a random point in the range.
2. Determine whether it is "in or out" using same algorithm described before.
3. If it is in, add one to in_{total} .
4. Repeat 500,000 times.

When this completes we apply the formula,

$$V = \frac{tot_{in}}{500000} (x_{max} - x_{min}) (y_{max} - y_{min}) (z_{max} - z_{min})$$

to find a very accurate approximation to the volume of the helix.

We then compute the packing ratio by using the total amino acid volume (V_{aa}) and the triangle approximation volume (V_{ta}) and the Monte Carlo approximated volume (V_{MC}),

$$P_{ta} = \frac{V_{aa}}{V_{ta}}, \quad P_{MC} = \frac{V_{aa}}{V_{MC}}$$

2.2.3 Ramachandran Statistics

Ramachandran angles are defined by the dihedral angles of the atomic lattice that is made by the backbones of the amino acids when a β -helix is formed. We define the ϕ angle using the angle formed by the two planes defined by C'-N-C $_{\alpha}$ -C' and the ψ angle as the angle formed by the planes N-C $_{\alpha}$ -C'-N (See figure 5).

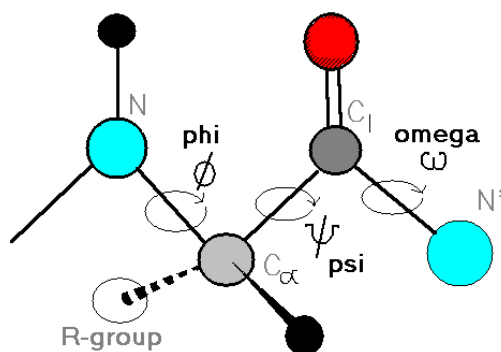


Figure 5: Graphical representation of $\phi - \psi$ definition

Using a *Ramachandran Plot* (a plot of ϕ vs. ψ) we can determine what secondary structure any single residue is conforming to within any larger structure. These angles are strictly limited by energetics and only certain combinations are allowed. Each area of Ramachandran space has an associated secondary structure (See figure 2, Appendix A).

2.2.4 Stacking

Stacking occurs when two identical amino acids lie one on top of the other spatially (One is directly above the other). This was initially thought to increase stability through bonding although it was later discovered that only side chains with aromatic rings or sulfuric side chains actually bond when stacked. It was initially thought to be energetically favorable because of the high abundance of stacking, but another program I wrote, that just relied on placement statistics showed that the frequency of stacks was just a random side effect of hydrophobicity (which determines where the residues want to be on the helix). The stacking statistics are still included because it can be illuminating for certain amino acids (namely the aromatic or sulfuric ones).

2.2.5 Hydrogen Bonding

Side chain hydrogen bonding can occur when a residue has a nitrogen or oxygen at the end of it. Many different hydrogen bonds can form, usually including the amino acids GLN, ASN, SER, HIS, THR and TYR.

The program does not perform electron density calculations or formulate complex potentials (like AMBER, a molecular dynamics simulator), but

rather just checks to see if the individual atoms involved in the hydrogen bond are close enough to each other spatially, and within a certain critical angle with each other (40 degrees). Because of this some hydrogen bonds are missed or thrown out, but it still gives a good first approximation at extremely fast speeds.

2.2.6 General Statistics

Some general statistics are calculated about the helix at the end including how many turns (layers in the helix) there were and how many residues the helix was composed of. These are just for statistical and reference purposes.

3 Results

I ran my code on 9 different β helices (7 Left handed, 2 Right handed). Proteins found in the Govaerts article were used although not all mentioned in the article were tested because some were too ill behaved (too many side loops, non helical structure etc.). It was found that most proteins behaved similarly with hydrophobic residues pointing in, about a 0.70-0.90 packing density, mostly beta structure with around 1 stack for every 10 residues. It was seen that the most stable configurations exhibited the behaviors that were discussed earlier as having good energetics, which was to be expected.

3.1 Individual Results

Individual statistics pages were made for each protein that include all information described in the preceding section. An example of one of these output pages can be seen at the end of Appendix B. These pages can be readily generated for these or other proteins using the users manual to the program found in Appendix B.

3.2 Total Results

The total results; important statistics and averages can be found in Appendix A, table 2.

3.3 Discussion

The program performed the way it was supposed to, calculating accurate statistics about files in the Protein Database. The biophysics group was interested in the structure of Scrapie, a prion disease found in sheep. The program was able to confirm the statistical stability (similarity to stable samples) of the structure that the group had proposed for one end of the prion. Using techniques from my program a model for the other end of the prion (the N-terminus) was able to be developed as well. This model, although not performing as well as the C-terminus, outperforms (in stability calculations) the current models for it. So the program was not only able to start a statistical base of the PDB, but it also can test theoretical models for yet unknown structures. The program is important because it gives a good quick overview of many areas of a β -helix, without having to run time consuming molecular dynamics tests.

4 Conclusion and Future Work

Proteins are vital to every living organism. Prions, or disease causing proteins are currently incurable and little is known about them. Knowing the structure of a protein allows for developments to be made towards fighting it (in the prion case). This program gives useful statistical data about any β -helical protein that can then be used to determine the stability of theoretical models for proteins or prions based off of already stable structures. The program has compiled a small statistical base from proteins in the PDB that can be used to check these new protein models. It was implemented in confirming the statistical stability of the C-terminus to Scrapie, a prion found in sheep.

Future work involving this program could include a unified scoring model that would rank proteins based off of performance in each statistical category, giving one final "score" that can be judged as stable or not. If this scoring model were made then a Monte Carlo Metropolis search could be used to find more stable configurations of proteins without having to result to very time intensive molecular dynamics.

References

- [1] Berger, Bonnie, Lenore Cowen, and Jonathan King. "BetaWrapPro." MIT. MIT Biology Dept. 24 Aug. 2006 <<http://betawrappro.csail.mit.edu/>>.
- [2] Douzery, Emmanuel J. "Molecular Evolution." Université Montpellier II. 24 Aug. 2006 <<http://www.isem.univ-montp2.fr/PPP/PM/TEACH/MBPE/MOLEVOL/@M1-MolEvol.php>>.
- [3] Govaerts, Cedric, Holger Wille, Stanley B. Prusiner, and Fred E. Cohen. "Evidence for Assembly of Prions with Left-Handed β -Helices." PNAS 101 (2004): 8342-8347.
- [4] Nelson, David L., and Michael M. Cox. "Amino Acids." Wikipèida. 2000. 24 Aug. 2006 <<http://en.wikipedia.org/wiki/Aminoacid>>.
- [5] Wampler, J E. "Peptide and Protein Structure." Secondary Structure. 1996. University of Georgia Biology Dept. 24 Aug. 2006 <<http://www.bmb.uga.edu/wampler/tutorial/prot2.html>>.
- [6] Whitford, David. Proteins: Structure and Function. New York: Wiley, 2005.

1 Appendix A

Table 1.1: A list of the amino acids and their names and volumes.

Name			Volume (\AA^3)	Side Chain
Alanine	Ala	A	67	$\text{CH}_3\text{-CH(NH}_2\text{)-COOH}$
Arginine	Arg	R	167	$\text{HN=C(NH}_2\text{)-NH-(CH}_2\text{)}_3\text{-CH(NH}_2\text{)-COOH}$
Asparagine	Asn	N	148	$\text{H}_2\text{N-CO-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Aspartate	Asp	D	67	$\text{HOOC-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Cysteine	Cys	C	86	$\text{HS-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Glutamine	Gln	Q	114	$\text{H}_2\text{N-CO-(CH}_2\text{)}_2\text{-CH(NH}_2\text{)-COOH}$
Glutamate	Glu	E	109	$\text{HOOC-(CH}_2\text{)}_2\text{-CH(NH}_2\text{)-COOH}$
Glycine	Gly	G	48	$\text{NH}_2\text{-CH}_2\text{-COOH}$
Histidine	His	H	118	$\text{NH-CH=N-CH=C-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Isoleucine	Ile	I	124	$\text{CH}_3\text{-CH}_2\text{-CH(CH}_3\text{)-CH(NH}_2\text{)-COOH}$
Leucine	Leu	L	124	$\text{(CH}_3\text{)}_2\text{-CH-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Lysine	Lys	K	135	$\text{H}_2\text{N-(CH}_2\text{)}_4\text{-CH(NH}_2\text{)-COOH}$
Methionine	Met	M	124	$\text{CH}_3\text{-S-(CH}_2\text{)}_2\text{-CH(NH}_2\text{)-COOH}$
Phenylalanine	Phe	F	135	$\text{Ph-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Proline	Pro	P	90	$\text{NH-(CH}_2\text{)}_3\text{-CH-COOH}$
Serine	Ser	S	73	$\text{HO-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Threonine	Thr	T	93	$\text{CH}_3\text{-CH(OH)-CH(NH}_2\text{)-COOH}$
Tryptophan	Trp	W	163	$\text{Ph-NH-CH=C-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Tyrosine	Tyr	Y	230	$\text{HO-p-Ph-CH}_2\text{-CH(NH}_2\text{)-COOH}$
Valine	Val	V	105	$\text{(CH}_3\text{)}_2\text{-CH-CH(NH}_2\text{)-COOH}$

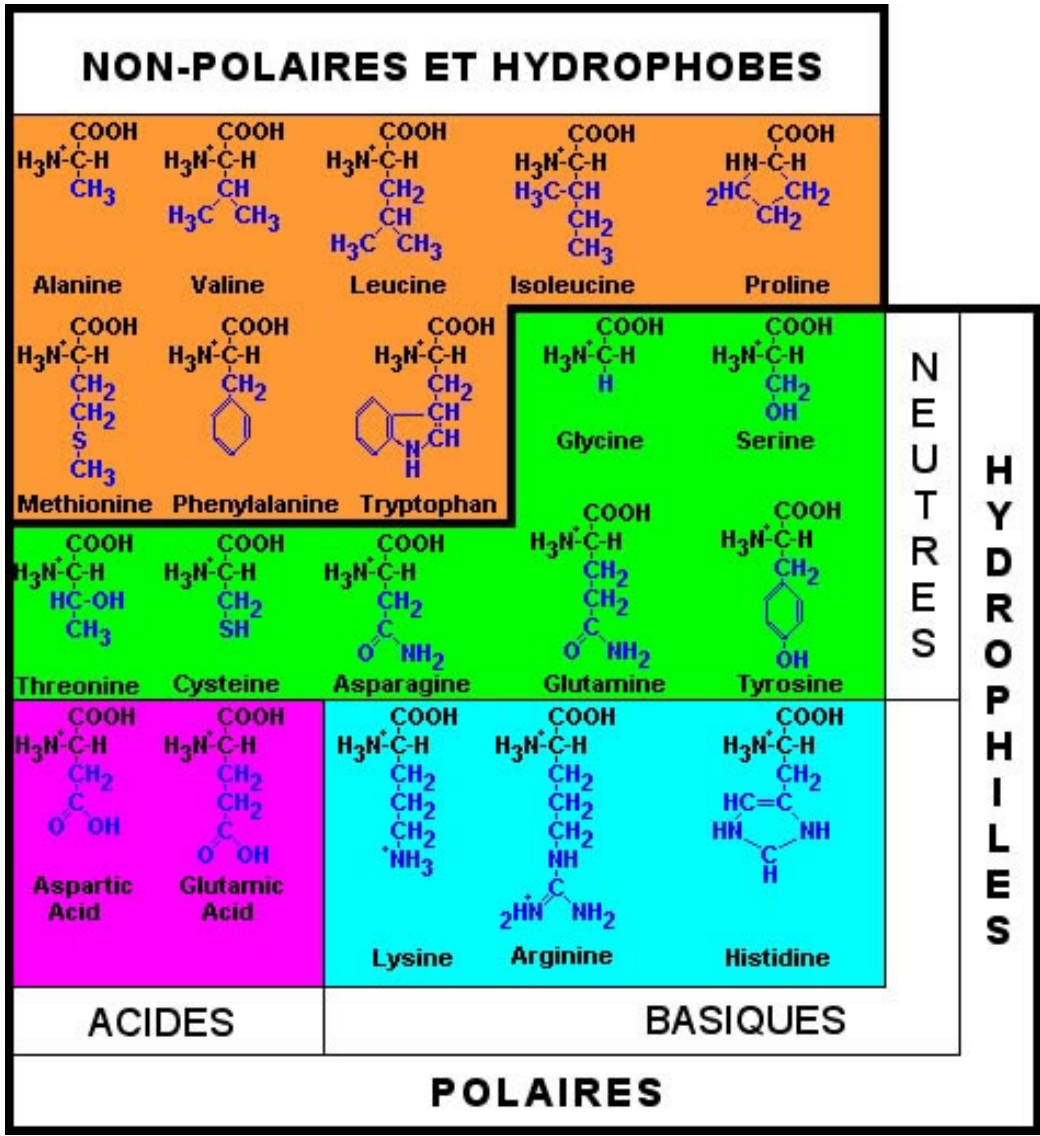


Figure 1: The periodic table of amino acids.

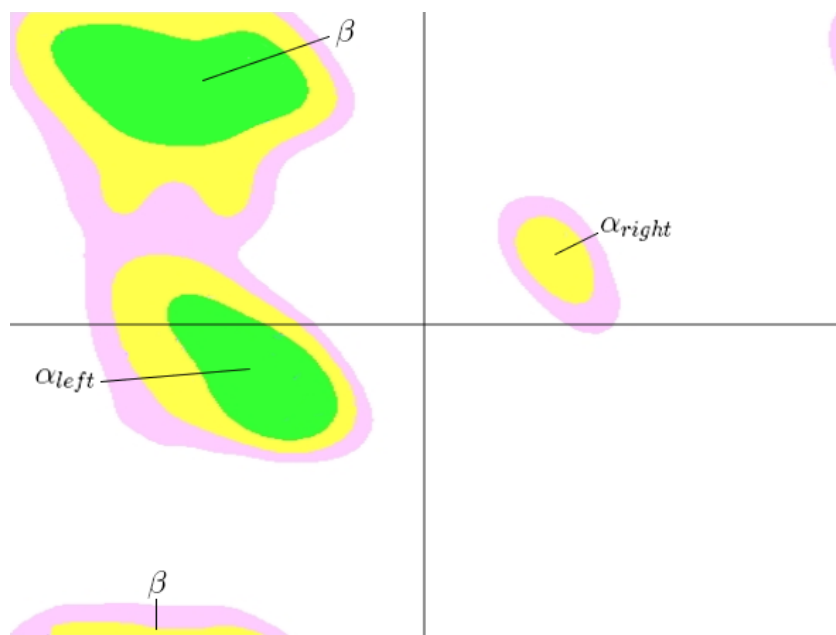


Figure 2: The secondary structures associated with different areas of Ramachandran space.

Protien	(L/R)	Volume	P_{MC}	P_{Ap}	Hydro In	Out	β	α_{RH}	α_{LH}	Runs	Stacks	H-b	Residues
1G97	L	6859	0.867	0.852	9	-45	130 (0.833)	8 (0.052)	18 (0.115)	1(2)	21	1	178
1HV9	L	6791	-	0.736	20	-56	131 (0.856)	5 (0.033)	17 (0.111)	-	16	2	178
1KK6	L	1429	-	0.693	7	1	31 (0.838)	2 (0.054)	4 (0.108)	-	3	0	42
1LOS	L	2308	-	0.837	-14	-25	45 (0.726)	5 (0.081)	12 (0.194)	3(2)	4	9	68
1LXA	L	6650	0.899	0.914	15	-42	126 (0.783)	15 (0.093)	20 (0.124)	5(2) 1(3)	12	0	185
1XAT	L	1633	0.747	0.692	7	-6	30 (0.833)	2 (0.055)	4 (0.111)	-	4	0	43
2TDT	L	3718	0.848	0.772	14	-20	81 (0.880)	4 (0.043)	7 (0.076)	1(2)	7	2	108
L_{avg}	-	4198	.840	.785	8.3	-27.6	82 (0.821)	5.9 (0.059)	11.7 (0.120)	1(2)	9.6	1	115
1EE6	R	7350	-	-	29	-83	144 (0.760)	30 (0.171)	12 (0.069)	3(2) 1(3) 1(4)	13	0	196
1DAB	R	6971	0.730	-	23	-57	129 (0.754)	29 (0.170)	13 (0.076)	6(2) 1(3)	8	1	196

1 Appendix B

1.1 Users Manual to Reader3.java

This manual will guide the user through using the program Reader3 and show how to generate nice output files like the one at the end of this appendix. The program is written in Java, although no knowledge of java is needed to run the program.

1.1.1 Setting up the program

To run the program on a PDB file first open the file Reader3.java in your favorite text editor. You will need to modify specific lines for it to run correctly through the protein.

Line 36: Modify what **START** is equal to. This is what residue the β -helix starts on in the *.pdb file.

Line 37: Modify what **END** is equal to. This is the residue that the β -helix ends at in the *.pdb file.

Line 38: Modify what **isLH** is equal to. True for left handed helices and false for right handed ones.

Line 41: Modify what **ins** is equal to. This is the name of the *.pdb that you want to run the program on, it must be in the same directory as Reader3.java.

Line 51-60: Modify the true/false values of the switches. This turns specific tests on or off (see comments).

1.1.2 Running the program

To run the program open a terminal to the directory with Reader3.java and the PDB file in it and run the commands:

```
javac Reader3.java
java Reader3
```

1.1.3 Generating an output file

All of the information that was sent to the terminal is saved in the file *stats.dat* that is now in the folder. The Ramachandran plot is stored (as a coordinate file) in the file *RAM.dat*. To make the nice output file we will need to open up a graphical text editor like Open Office Writer or LaTeX.

First we copy and paste all of the information in stats.dat into a file. We will then need to make the Ramachandran plot, which is what we will discuss in the next section.

1.1.4 Making the Ramachandran Plot

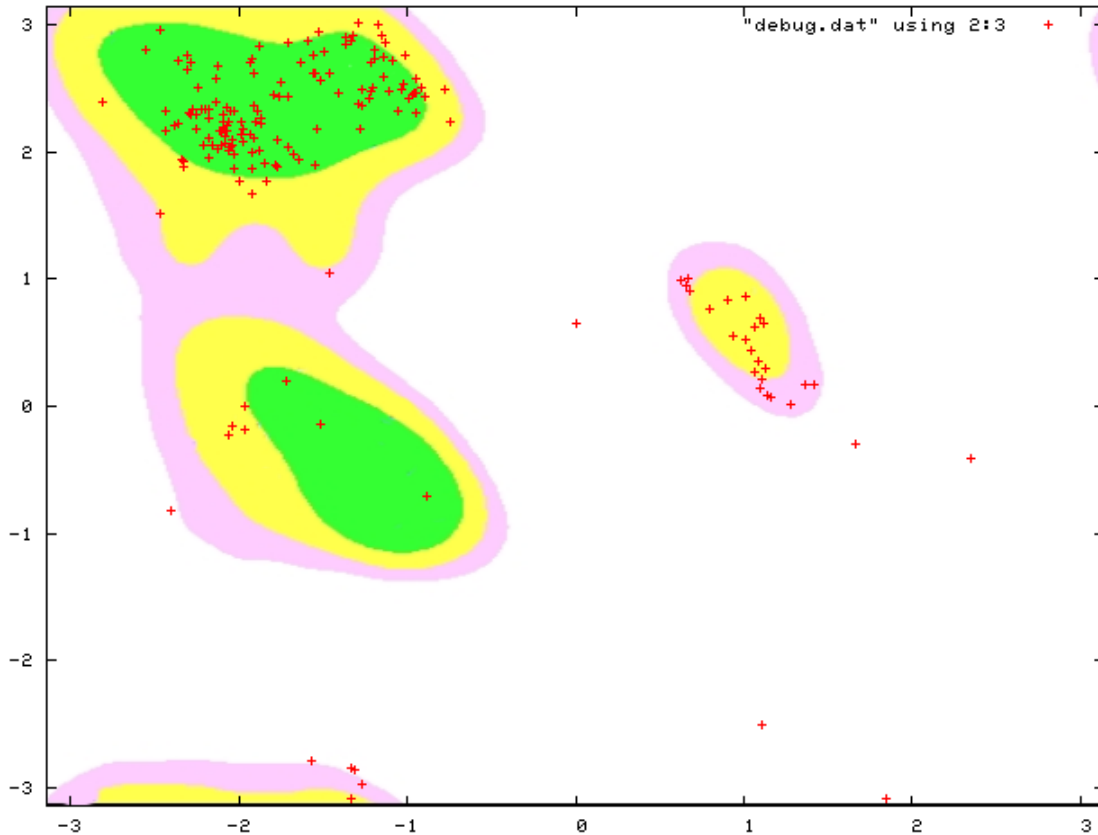
Making the Ramachandran plot can be quite involved, depending on how nice you would like it to look. This guide will walk the user through making the nicest plot, but the user can stop at any point along the way.

1. Open a terminal to the folder where RAM.dat is.
2. Run the following commands:
3. gnuplot
4. set xrange[-3.14159:3.14159]
5. set yrange[-3.14159:3.14159]
6. plot "RAM.dat" using 2:3
7. set terminal png
8. set output "RAMout.png"
9. plot "RAM.dat" using 2:3
10. exit
11. Now open the file RAMout.png using GIMP
12. Choose the button: Select Regions by Color
13. Select a red point on the graph,
14. While holding shift, select a black point on the graph
15. Copy (Ctrl+C)
16. Click on File→New
17. Change the size to 640x480

18. Paste (Ctrl+V)
19. Right click on "pasted layer" in the layers panel
20. Select new layer
21. Open the file Ram4.png (should be distributed with Reader3.java, mirrored at www.oregonstate.edu/clarksc/Ram4.png) in GIMP
22. Select the whole image
23. Copy (Ctrl+C)
24. Select the new file that you made a second ago
25. Paste (Ctrl+V)
26. Go to the layers panel, right click on pasted layer, and create a new layer
27. Right click on this new layer in the layers panel and lower the layer
28. Select the move tool from the main panel
29. Move the background colors to fit the boundaries
30. Save the file, and you're done.

Now we insert this new picture file into our document, save the document and it should look like the following page.

1G97.pdb (LH)



Volume: 6859.0

Volume total (of helix) (MC): 7909.561344482157

Volume ratio V/Vt(MC): 0.867178304999803

Volume total (of helix) (Ap): 8047.239679050651

Volume ratio V/Vt: 0.8523419549508398

Hydrophobicity:

IN: 41 Hydrophobic residue(s) 32 Hydrophilic residue(s) including 12 GLY residue(s)

IN: 0 basic residue(s) 0 acidic residue(s)

IN: Total hydrophobicity: 9.0

OUT: 30 Hydrophobic residue(s) 45 Hydrophilic residue(s) including 10 GLY residue(s)

OUT: 12 basic residue(s) 18 acidic residue(s)

OUT: Total hydrophobicity: -45

Burried Hydrophobicity:

IN: 33 Hydrophobic residue(s) 30 Hydrophilic residue(s) including 12 GLY residue(s)

IN: 0 basic residue(s) 0 acidic residue(s)

IN: Total hydrophobicity: 12.0

OUT: 20 Hydrophobic residue(s) 35 Hydrophilic residue(s) including 7 GLY residue(s)

OUT: 9 basic residue(s) 11 acidic residue(s)

OUT: Total hydrophobicity: -35

There were 130 (0.8333333333333334) Beta residues, 8 (0.05128205128205128) RH alpha and 18 (0.11538461538461539) LH alpha of 156 non-GLY residues.

1G97.pdb (LH)

On the turns: 45 (0.6617647058823529) were Beta, 5 (0.07352941176470588) RH alpha and 18 (0.2647058823529412) LH alpha of 68 non-GLY residues

There was a sequence of 2 alpha pieces in a row (265-266).

Inward Stack at 268 and 286 with I

Outward Stack at 269 and 288 with A

Outward Stack at 271 and 289 with E

Outward Stack at 277 and 294 with N

Inward Stack at 286 and 303 with I

Outward Stack at 288 and 305 with A

Outward Stack at 291 and 308 with V

Outward Stack at 293 and 310 with T

Outward Stack at 294 and 311 with N

Inward Stack at 314 and 331 with I

Inward Stack at 317 and 335 with S

Outward Stack at 318 and 336 with S

Outward Stack at 320 and 339 with A

Outward Stack at 336 and 353 with S

Outward Stack at 357 and 374 with N

Inward Stack at 364 and 381 with T

Outward Stack at 374 and 399 with N

Outward Stack at 399 and 417 with N

Inward Stack at 402 and 420 with V

Outward Stack at 407 and 425 with T

Inward Stack at 408 and 426 with I

Inward double stack of I at 268-286-303

Outward double stack of A at 269-288-305

Outward double stack of N at 277-294-311

Outward double stack of S at 318-336-353

Outward double stack of N at 357-374-399

Outward double stack of N at 374-399-417

A stacked 3 time(s)

N stacked 5 time(s)

E stacked 1 time(s)

I stacked 4 time(s)

S stacked 3 time(s)

T stacked 3 time(s)

V stacked 2 time(s)

Side chain hydrogen bonds

Hydroxyl - carbonyl H bond at 296 T and backbone O 309

There were 178 residues (259-437), and approximately 9.88888888888889 turns.