

The study of protein secondary structure and stability at equilibrium

Michelle Planicka

Dept. of Physics, North Georgia College and State University, Dahlonega, GA

REU, Dept. of Physics, University of Florida, Gainesville, FL

30 July 2003

ABSTRACT

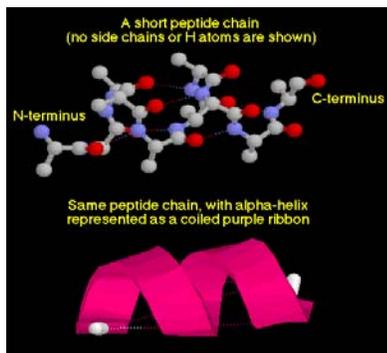
In order to study the process by which a protein folds, it is first necessary to characterize the protein, its structure, and its stability at equilibrium. Circular dichroism (CD) spectroscopy is used to characterize three novel proteins recently isolated from E.coli (EC298 and EC308) and from the thermophilic bacterium methanobacterium thermoautotrophicum (MT865). As a probe of protein secondary structure, measurement of UV-CD spectra under varying conditions of temperature and solvent allowed the determination of the stability of the properties of these proteins.

1. Background Information

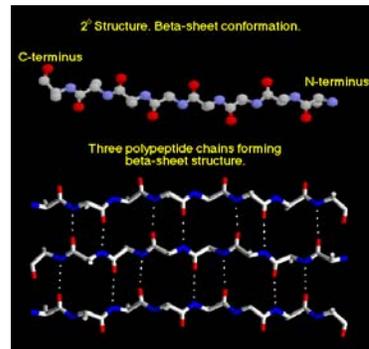
A protein is made up of a string of amino acids. There are 20 different amino acids naturally occurring in proteins. They are listed below with their three-letter and one-letter symbols [1]:

- Aspartic acid (Asp, D)
- Glutamic acid (Glu, E)
- Arginine (Arg, R)
- Lysine (Lys, K)
- Histidine (His, H)
- Asparagine (Asn, N)
- Glutamine (Gln, Q)
- Serine (Ser, S)
- Theronine (Thr, T)
- Tyrosine (Tyr, Y)
- Alanine (Ala, A)
- Glycine (Gly, G)
- Valine (Val, V)
- Leucine (Leu, L)
- Isoleucine (Ile, I)
- Proline (Pro, P)
- Phenylalanine (phen, F)
- Methionine (Met, M)
- Tryptophan (Trp, W)
- Cysteine (Cys, C)

Each amino acid consists of an identical carbon backbone but different side chains. Side-chains are the molecules attached to the carbon backbone. The amino acid sequence along with the side chains is known as the primary structure of a protein. The primary structure influences the secondary structure of the protein. The secondary structure is the shape that the amino acid chain conforms to. Two main secondary structures are found in a protein: Alpha-helix and Beta-Sheet (see Fig. 1).



Alpha-Helix



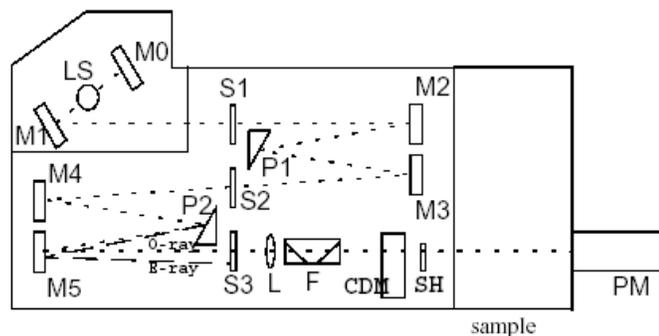
Beta-Sheet

FIGURE 1. The two main secondary structures of a protein. Source: (Alpha helix/Beta sheet) <http://www.sp.uconn.edu/~terry/images/mols/alphahelix.gif>

The amino acid sequence along with the side-chains in its secondary structure determines the biochemical function of the protein. In order for a protein to carry out its biochemical function, it must ‘fold’. After folding, the result is a three-dimensional structure that is predetermined by the amino acid sequence to minimize the free energy of the protein structure. This three-dimensional folded structure is known as the protein’s tertiary structure [1].

Protein can also be forced to denature, or unfold, in two ways. One way is to use a chemical called denaturant. The other way is by heating a protein to high temperatures (above 50°C). When denaturant is added to a protein solution, it causes the protein to unfold at a lower temperature. By adding the denaturant and unfolding the protein at a lower temperature, it allows us to easily look at the folding/unfolding process of the protein. Once the denaturant is removed, the protein spontaneously refolds to its original conformation. Using denaturant has become a standard for observing protein folding. How a protein folds has been a mystery to scientists for over half a century.

To study the stability and folding of proteins, a Circular Dichroism Spectrometer is used (see Fig. 2).

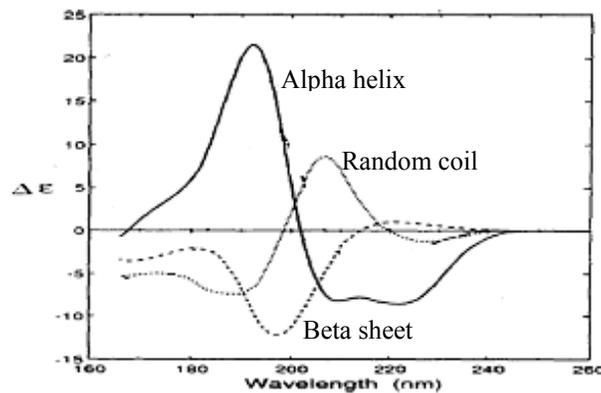


- | | |
|---|--|
| <i>LS</i> | <i>light source (DC xenon lamp, 150W)</i> |
| <i>M₁</i> | <i>elliptical mirror</i> |
| <i>M₀</i> | <i>spherical mirror</i> |
| <i>S₁, S₂, S₃</i> | <i>entrance, intermediate and exit slits</i> |
| <i>M₁, M₂, M₃, M₄</i> | <i>collimating spherical mirrors</i> |
| <i>P₁, P₂</i> | <i>quartz prisms</i> |
| <i>L</i> | <i>quartz lens</i> |
| <i>F</i> | <i>filter (quartz plates to clean the linear polarization)</i> |
| <i>CDM</i> | <i>CD modulator (piezoelectric element operated as an achromatic quarter wave plate)</i> |
| <i>SH</i> | <i>shutter</i> |
| <i>PM</i> | <i>photomultiplier tube</i> |

FIGURE 2. The layout of the inside of the Circular Dichroism Spectrometer.

Source: <http://www.jasco-europe.com>

The CD spectrometer provides a way of determining whether or not a protein is folded under a given set of conditions. The circular dichroism of a material is a measure of the difference between the material's absorbance of right-circularly polarized light versus left-circularly polarized light. The CD of a material normally varies as a function of wavelength. The CD spectrometer produces a plot of the CD versus wavelength over a range of wavelengths. For proteins, the CD spectrum in the UV range (190 nm-250 nm), known as far UV, provides information about the secondary structure of the protein.



α helix: positive band around 192nm, negative ones at 209 and 222nm
β sheet: negative band at 216 and positive one at 197nm
random coil: positive band at 212 and negative one at 198nm

FIGURE 3. There are three standard plots produced by the CD spectrometer. They are alpha helix, beta sheet, and random coil. Based on these three standard plots, one can determine the secondary structure of their protein. Source: <http://www.jasco-europe.com>

If the resultant plot shows that the protein structure does not unfold under a given set of conditions, then the protein structure is stable over those conditions. Being stable means that the protein structure, in its folded state, has the lowest free energy under the given conditions. When the conditions are changed rather severely and the protein structure remains folded over the severe change, it is highly stable. Protein structures can also be unstable. This condition occurs when the resultant plot shows a protein structure unfolding under a given set of conditions. If the protein structure can unfold under a given set of conditions, its free energy is not at a minimum relative to the folded state of

the structure. The given conditions can be dependent on wavelength, temperature, or denaturant.

By looking at the amino acid sequence of our three proteins (See Table 1), we can see that both EC298 and EC308 contain tryptophan. Since EC298 and EC308 contain tryptophan, we can measure a protein's concentration using a spectrophotometer (see Fig. 4).



FIGURE 4. Spectrophotometer used to measure protein concentration if the protein contains tryptophan. Source: <http://www.ssi.shimadzu.com/products/spectro/newindex.cfm?product=uv1601>

The spectrophotometer measures the ability of a sample to absorb light at a given wavelength. In measurements on typical materials, this absorbance varies as a function of wavelength. The spectrophotometer will produce a plot of the absorbance versus wavelength. The spectrophotometer is similar to the CD spectrometer via hitting a protein sample with a beam of Ultraviolet light. Once the UV light hits the sample, it is absorbed by the tryptophan present in the protein. By looking at the amount of UV light that gets absorbed by the tryptophan, we can determine the protein concentration.

Recently, three new proteins, EC298, EC308, and MT865, have been engineered in Toronto, Ontario (see Fig. 5).

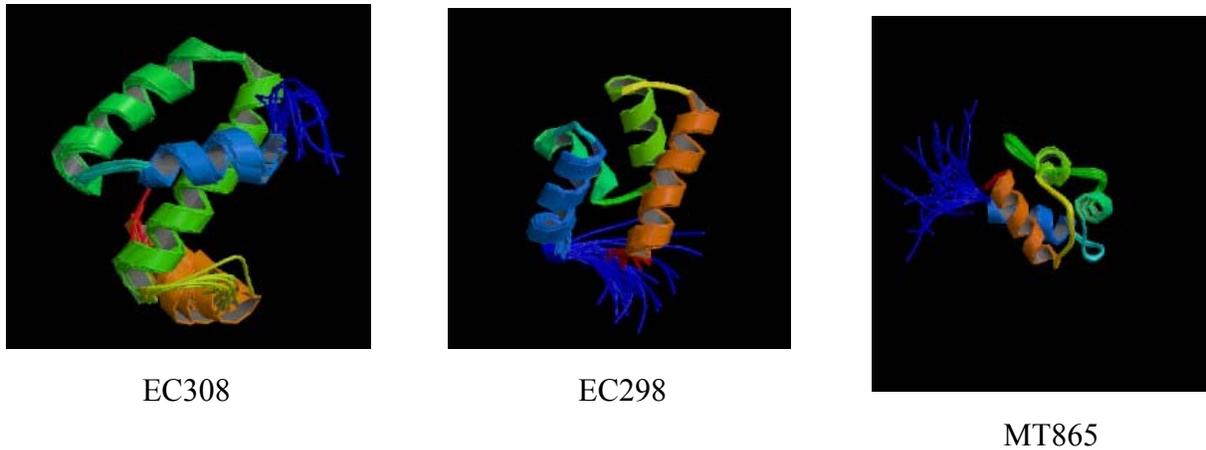


FIGURE 5. Secondary structures of EC308, EC298, and MT865. Source: www.rcsb.org

Due to the lack of knowledge known about these proteins, Karen Maxwell (CGC Proteomics at the University of Toronto) and Professor Stephen J. Hagen (University of Florida) have teamed up in a scientific collaboration to discover the basic characteristics of EC298, EC308, and MT865. EC298 and EC308 both originated from the bacterium *E.coli*, while MT865 originated from the thermophilic bacterium *methanobacterium thermoautotrophicum*. EC308 functions as a gene regulator, and the functions of EC298 and MT865 are unknown. In particular, we are looking at the stability and folding process of each protein.

2. Methods

In order to observe the stability and folding of EC298, EC308, and MT865, we had to remove approximately all of the salt present. The measurement of folding properties is done under a set of standard conditions in which the presence of salt is undesirable. Starting with EC298 and EC308, we placed 2 mL of solution into a filter

and used a centrifuge to spin the two samples multiple times at 4500 rpm for 30 minutes. By using this process, we were able to get the salt concentrations of EC298 and EC308 down to 4.17 mM and 3.94 mM from 200 mM.

Since we were using a filtration process, we were concerned with the amount of protein that was being lost (leaking through the filter). The spectrophotometer helped me to determine the amount of protein that was lost in the filtration. I was able to use the Beer Lambert Law, $A=C\epsilon L$, to calculate the protein concentration of the filtered sample. For example, my 0M guanidine reading on the spectrophotometer at 280 nm (maximum absorbance) was .11. The protein concentration is calculated by: $C = A/\epsilon L$, where A is the absorbance (0.11), ϵ is the extinction coefficient for tryptophan(5580 (M-cm)^{-1}), and L is the cell path (.1 cm). After completing this simple calculation, I divided my resultant answer by the number of tryptophan amino acids in EC298. The protein concentration of EC298 was 49.3 micro Molar (μM), and the concentration of EC308 was 18 μM . From the initial protein concentration of 50 μM for both protein solutions, it seemed that there was a defective filter used for EC308.

Because the filtration didn't work as well as we wanted it to with EC308, we chose to use the method of dialysis for the removal of salt from the original protein solution (see Fig. 6). We placed approximately 2mL of EC308 solution into a special membrane cell made for dialysis. The membrane cell was then submerged into a beaker filled with 50mM pH7 Phosphate Buffer where it was then placed on top of a magnetic stirrer for the next 6 hours. The protein molecules are trapped inside the membrane due to their size, whereas the salt molecules are small enough to slip through the membrane out into the buffer.

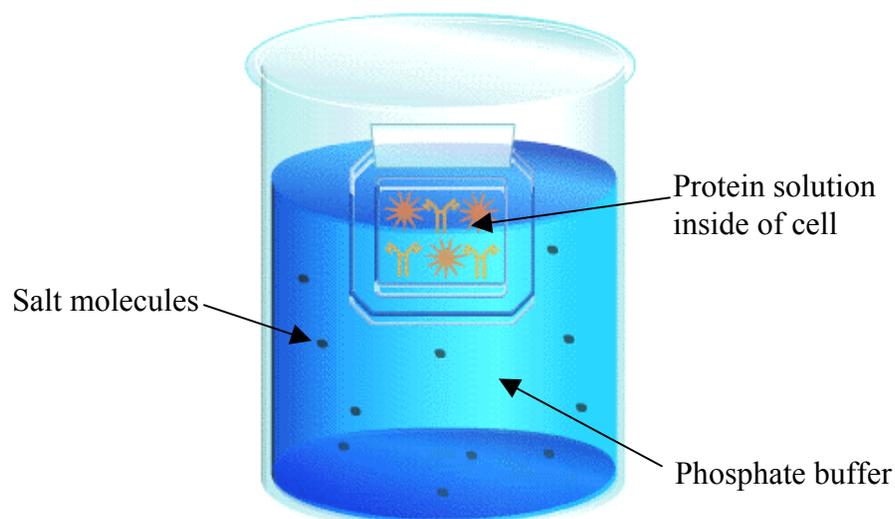


FIGURE 6. The process of dialysis. Source: <http://www.piercenet.com/dialysis/imgs/slidefront.gif>

After we replaced the buffer solution once and used the spectrophotometer, the protein concentration of EC308 was $98.56 \mu\text{M}$.

While EC308 was being dialyzed, we prepared EC298 for the CD Spectrometer. We ran several scans to test the limits of the protein. The first scan that we took was a wavelength scan at 25°C . This scan strikes our sample of EC298 with wavelengths of light ranging from 250 nm to 190 nm while the temperature is being held constant. This process enables us to see what the protein structure looks like folded/unfolded and observe its secondary structure by measuring the difference in the amount of right and left circularly polarized light that is being absorbed by EC298.

Following the initial wavelength scan, we ran two more wavelength scans, one at 5°C and one at 95°C . Looking at a wide variety of temperatures for the wavelength scans gives us a clear picture of the protein's structural stability. We also took a temperature scan. A temperature scan holds the wavelength constant while the temperature varies. The wavelength that we observed was 222nm because this is where the greatest difference of left/right circularly polarized light is absorbed between the folded and unfolded state. The temperature varies from 5°C to 95°C . This scan allowed us to see at what temperature the protein structure unfolds. We repeated this process for EC308 after

it was dialyzed. While looking at the wavelength scan at 95°C, we noticed that EC308 comes out of solution. This means that the protein actually sticks to the side of the CD spectrometer cell and is not dissolved in the solution. We left EC308 in the salt solution and proceeded to take the three wavelength scans and the temperature scan.

After looking at the amino acid sequence of MT865, we chose to use a similar method as before. The method of dialysis seemed to be the best way to remove the salt concentration based on our previous experience with EC308. We followed the same procedure while dialyzing the MT865 as we did previously with the EC308 solution. Because MT865 lacked tryptophan, we had no way to check the protein concentration after dialysis. The CD spectrometer was then used to run the same scans for MT865 as we did for EC298 and EC308.

To take folding measurements and check the stability of our three proteins, we needed to add denaturant. Guanidine is the denaturant we chose to use (see Fig. 7).

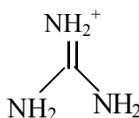


FIGURE 7. Chemical structure of guanidine. Source: Maryadele J. O'Neil...[et al], The Merck Index: An encyclopedia of chemicals, drugs, and biologicals, 13th ed. (Merck & CO., Inc., New Jersey, 2001), p .4580

Crystals of guanidine were measured out to make a five molar solution when mixed with water. After the samples were made, the guanidine molarities were checked using an index of refraction measurement. EC298 was the protein we chose to start with. We then decided to make ten sample solutions from zero molar guanidine up to five molar guanidine. Each of the ten samples contained 100 μL of EC298 filtered protein. We randomly chose different molarities to look at using the CD spectrometer to decide the greatest molarity needed. After doing several different molarities, 1.5 molar proved to be as high as we needed to go. Ten new samples were made from the range of zero molar guanidine to 1.5 molar guanidine, each still containing 100 μL of EC298 filtered protein.

Because the guanidine makes the protein unfold at a lower temperature than normal, we chose not to take wavelength scans at such high temperatures as we had done earlier. Each guanidine sample (0M-1.5M) had wavelength measurements taken at 5°C, 25°C, and 55°C from 250 nm to 190 nm. A temperature scan at 222 nm was also taken using the appropriate parameters.

3. Results

The plots produced from our first set of protein samples (without guanidine) confirmed that EC298, EC308, and MT865 are alpha-helical structures. Comparison of the EC298 wavelength scan at 25°C (see Fig. 8) with the standard CD spectrometer plot shows that at 222 nm and at 208 nm the plot has a minimum, which is characteristic of an alpha-helical protein.

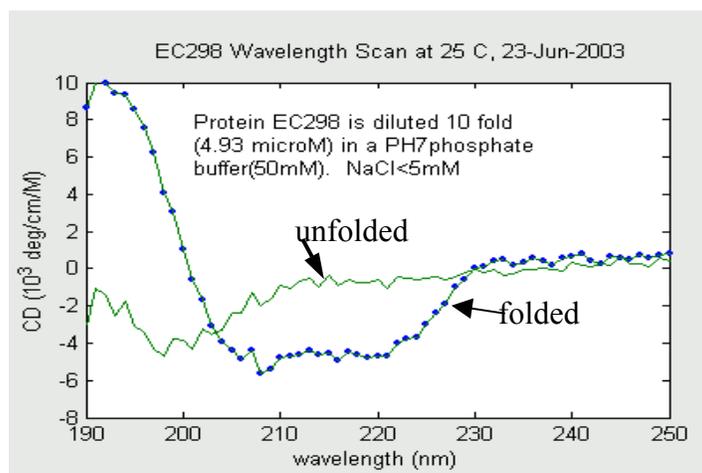


FIGURE 8. EC298 Wavelength scan at 25°C.

The temperature scans confirmed that EC298, EC308, and MT865 are also capable of unfolding. A temperature scan compares the state of folded to that of unfolded at 222 nm due to the maximum change in UV absorbance at that wavelength. Using the EC298 data as before, the unfolding of the protein is represented by tracking the data of the folded state to the unfolded state from 5°C to 95°C. Referring to the EC298 wavelength plot at

222 nm, we see that the folded state is at negative four, where the unfolded state is approaching zero. From this information we know that our temperature plot should start at negative four and end at approximately zero, as it does (see Fig. 9).

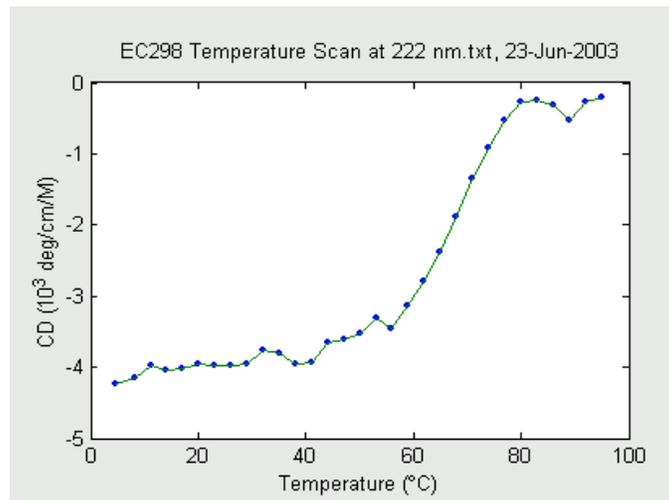


FIGURE 9. EC298 Temperature scan at 222 nm.

It is the transition between the two states that gives the plot its steep climb. This steep climb is where the protein is unfolding. We see here that EC298's midpoint temperature of unfolding at 0M guanidine is roughly 70°C. EC308 and MT865 behaved similarly at 0M guanidine.

Reviewing the resultant plots from the ten-guanidine concentrations of EC298 showed that EC298 is less stable than other proteins due to its unfolding at only 0.84M guanidine (see Fig. 10).

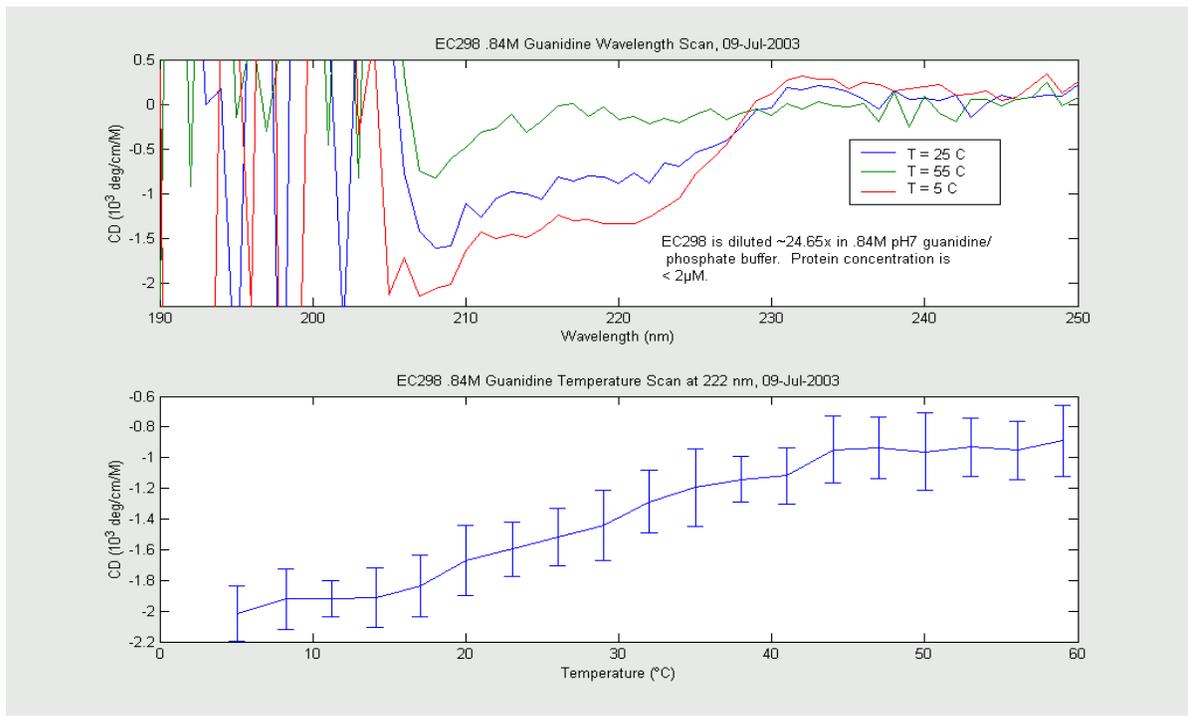


FIGURE 10. EC298 0.84M guanidine wavelength and temperature scans.

The wavelength scan of EC298 shows that a scan at 25°C is different from one at 5°C. If EC298 were more stable, the red and blue lines would be the same, showing that it is folded. The temperature scan at 222 nm of EC298 supports the idea that EC298 is beginning to unfold before 25°C. There is a steep increase in the plot just before 20°C suggesting that it is starting to unfold. The guanidine has shifted the temperature at which EC298 unfolds, and it has also caused EC298 to have a broader temperature scale for unfolding.

4. Conclusion

By using the CD spectrometer, we were able to characterize our proteins, study their structure, and observe their stability at equilibrium. Based on our CD data, we confirmed that EC298, EC308, and MT865 were all folded at room temperature and would unfold when heated to 95°C. The CD plots also confirmed the belief that all three structures were alpha helical. From our testing, we have discovered that EC298 and

MT865 have a high solubility in buffer solution without the presence of salt. We were very surprised to find that EC308 has a low solubility in buffer solution. At high temperatures, salt needs to be present to make EC308 more soluble; otherwise, it comes out of solution. The reason for this occurrence is yet to be determined. All three proteins in the presence of 0M guanidine unfold at approximately 70°C. Based on our guanidine samples experiment, EC298, EC308, and MT865 are less stable relative to other proteins. Each protein unfolded at room temperature (25°C) before 1M guanidine. A summary of our results can be found in Table 1.

	EC298	EC308	MT865
PDB Code	1JYG	1JW2	1IIO
Origin	E.coli	E.coli	methanobacterium thermoautotrophicum bacterium
# amino acids	69	72	84
Sequence	MNKDEAGGNW QFKGKVKEQ WGKLTDDDMT IIEGKRDQLV GKIQERYGYQ KDQAEKEVVD WETRNEYRW	MSEKPLTKTD YLMRLRRCQT IDTLERVIEK NKYELSDNEL AVFYSAADHR LAELTMNKLY DKIPSSVWKF IR	GSHMKMGVKE DIRGQIIGAL AGADFPINSP EELMAALPNG PDTTCKSGDV ELKASDAGQV LTADDFPFKS AEEVADTIVN KAGL
Secondary structure	Alpha helix	Alpha helix	Alpha helix
Temp. of unfolding (0M guanidine)	70°C	70°C	70°C
Temp. of unfolding (1M guanidine)	5°C		
Temp. of unfolding (1.5M guanidine)	5°C		5°C

TABLE 1. Summary of EC298, EC308, and MT865. Source: www.rcsb.org

REFERENCES

[1] Bruce Alberts...[et al], *Essential Cell Biology: An Introduction to the Molecular Biology of the Cell*, 1st ed. (Garland, New York, 1998), p. 133-156.

ACKNOWLEDGEMENTS

I would like to thank the University of Florida Physics Department and the NSF for funding the REU program. I would also like to thank Dr. Kevin Ingersent and Dr. Alan Dorsey for coordinating such a wonderful summer program. Special thanks to Dr. Stephen Hagen for allowing me to work and learn in his lab and for his unbiased opinion.