Georgia Journal of Science

Volume 64 *No. 2 Scholarly Contributions from the Membership and Others*

Article 8

2006

Constructing Globular, Sheeted, and Helical Polyalanine Structures using Nanotubes as Templates for Computational Studies

Caley Allen

Travis Bailey

Opie Bowen

Virginia Bryant

Nickolas Kitchens

See next page for additional authors

Constructing Globular, Sheeted, and Helical Polyalanine Structures using Nanotubes as Templates for Computational Studies

Authors
Caley Allen, Travis Bailey, Opie Bowen, Virginia Bryant, Nickolas Kitchens, Jason Lackey, Inez Mills, Om Patel, Nicholas Richardson, Shanda Sermons, Thomas Shaw, Calandra Stafford, Tiffany Stoney, Chad Ward, Samuel Zipperer, and Thomas J. Manning

CONSTRUCTING GLOBULAR, SHEETED AND HELICAL POLYALANINE STRUCTURES USING NANOTUBES AS TEMPLATES FOR COMPUTATIONAL STUDIES

Caley Allen, Travis Bailey, Opie Bowen, Virginia Bryant, Nickolas Kitchens, Jason Lackey, Inez Mills, Om Patel, Nicholas Richardson, Shanda Sermons, Thomas Shaw, Calandra Stafford, Tiffany Stoney, Chad Ward, Samuel Zipperer, and Thomas Manning*
Department of Chemistry
Valdosta State University
Valdosta, GA 31698

*Corresponding author: tmanning@valdosta.edu

ABSTRACT

In this experiment polyalanine was folded into globular, sheeted and helical structures through the use of carbon nanotubes. The rigidity of the nanotubes allowed for molding the polyalanine into the various structures. Nanotubes of different diameters and volumes were used in this experiment. Once the three dimensional peptide structures were formed and detached from the carbon nanotubes, a number of thermodynamic calculations were performed. Computational methods were used to calculate parameters such as Gibbs free energy, enthalpy, entropy, and molecular volume. By attaining the measurements of the Φ and Ψ angles, Ramachandran plots were constructed using linear, globular, sheeted and helical protein structures.

INTRODUCTION

i. Protein Structure

Proteins serve many biological functions in the body such as O_2 transporters and building up immunity against infection or disease. The functions of proteins are dictated by their structure or conformation. There are twenty well known amino acids that are classified into five categories: nonpolar, aliphatic R groups: (i.e. glycine, alanine); aromatic R groups: (i.e phenylalanine); polar, uncharged R groups: (i.e. serine, asparagines); positively charged R groups: (i.e. lysine); and negatively charged R groups: (i.e. aspartate). Alanine (see Figure 1) and peptides of polyalanine are the focus of this study.

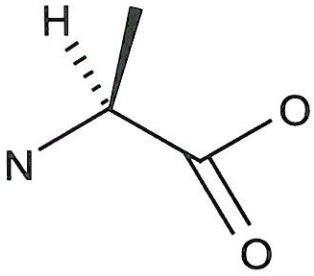

Figure 1. The amino acid alanine ($C_3H_7NO_2$)

Proteins are made of amino acids joined covalently by peptide bonds. The peptide bonds are formed through condensation reactions in which a water molecule is lost. Proteins are long polypeptide chains ranging from one hundred to several thousand amino acid residues and can adapt to different geometries. The backbone of the polypeptide chain is composed of the alpha carbon (C_α) which is linked to the carbonyl carbon. The carbonyl carbon is bonded to the nitrogen forming the backbone (see Figure 2).

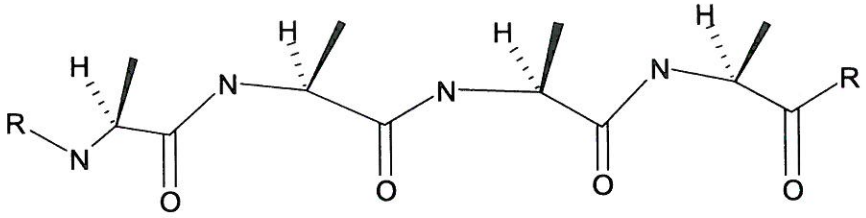

Figure 2. Polyalanine $(C_3H_7NO_2)x$ is the structure used in this study.

If the carbon-nitrogen bond has double bond characteristics, it cannot rotate freely. The bonds between the carbonyl carbon and the alpha carbon and between the nitrogen and the alpha carbon are the only bonds in the backbone that can move freely. If a large side chain is attached then the movement of the backbone will be stericly inhibited. Since the carbonyl carbon and nitrogen bond are fixed, it is possible to calculate the angles made between two nitrogen atoms (Φ) or two carbonyl carbons (Ψ) relative to the plane of the carbonyl carbon-nitrogen bond. Since the bond angles can range from -180 to +180, there are a number of conformations a protein can undergo.

There are four levels of protein structures. These structures. These include primary, secondary, tertiary, and quaternary structures. The primary structure is the order or sequence of linear chains of amino acids present. The secondary structure is a local region of organized structure and is heavily impacted by hydrogen bonds. There are two types of arrangement for secondary structures: the alpha-helix and beta sheets. In the alpha helix there is hydrogen bonding between the hydrogen on the nitrogen and the oxygen

on the carbonyl carbon within the same sheet. In the beta sheet conformation, the polypeptides are parallel and the hydrogen bonds link the adjacent polypeptides (1).

Proteins can be divided into two major categories: fibrous proteins and globular proteins. A fibrous protein has polypeptide chains arranged in long strands and/or sheets. These proteins only have one type of secondary structure. Generally fibrous proteins provide strength, flexibility, shape, and support to vertebrates. A globular protein is formed when the polypeptides fold into a ball or a globular shape. Different types of secondary structures, generally enzymes and regulatory proteins, are globular proteins. Both groups are explored in this work. If a protein's three-dimensional conformation shifts, or if there is a loss of structure, then the protein level of functioning will deteriorate. Denaturation is an example of this process. Three common causes of denaturation are increased heat, shifts in pH, and the introduction of organic solvents and/or detergents. The effects of which can be difficult or impossible to model in their totality.

ii. Thermodynamics of Protein Folding

Protein folding is influenced by the change in enthalpy (ΔH) of the structure. The amount of energy a protein absorbs is related to the temperature of the surroundings and the entropy (ΔS) associated with the organization of the protein structure. Entropy cannot be determined as a singular entity, but the change in ΔS can be estimated. The free energy change (ΔG) of a protein folding reaction can be calculated via,

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

Where ΔG is the Gibbs free energy (J/mol), ΔH is enthalpy (J/mol), T is the temperature in Kelvin (K), and ΔS is the entropy (J/mol K). Protein folding is typically an endothermic process. The transition of a protein from its unfolded state to its more ordered, folded state results in a decrease in entropy or disorder. Considering the thermodynamic parameters and neglecting all outside interactions, proteins favor the unfolded conformation.

Since the energy associated with the protein structure favors the unfolded form, another source of energy must be utilized for protein folding to occur. The solvent in which the protein is suspended provides this energy. The solvent will exchange some of its energy (- ΔH) with the protein. The ΔS of the environment value will increase due to the disruption of ordered solvent molecules positioned in between the amino acid side chain groups (1). The negative ΔG for the solvent outweighs the positive ΔG for protein, resulting in a negative ΔG for the system, which permits protein folding.

Figure 3 illustrates a folding reaction of polyalanine in transition from a helical form to a more globular structure. The absorbed energy allows the protein to form intramolecular bonds, resulting in a more stable conformation. This will allow for greater structural organization and a lower energy

state than that of its unfolded form. Figure 4 outlines the construction of a nanotube used as a template.

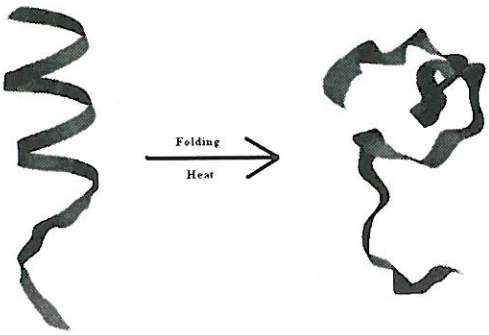

Figure 3. This illustration generalizes the thermodynamic reaction that occurs during protein folding/unfolding. The reactant protein is a helical form of polyalanine. As the protein folds intermolecular bonds form and aggregate the peptide chain forming a more ordered structure. Energy can be absorbed by the protein ($+\Delta H_{protein}$) from the environment ($-\Delta H_{solvent}$) allowing the protein to become more organized in structure as its ΔS decreases.

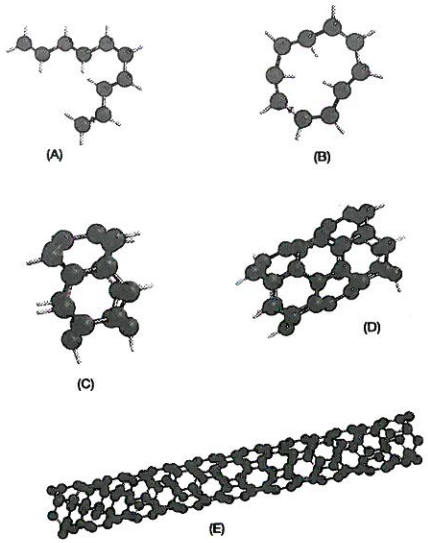

Figure 4. (A) A carbon chain composed of ten carbon atoms (B) the C_{10} chain is formed into a loop (C) 2 loops are connected using every other available bond forming a series of 6-membered rings (D) the two ring subunit is connected (E) and this unit is multiplied until the desired nanotube is constructed. The notation "10,0" is derived from the ten membered ring.

The Gibbs free energy for this process can be estimated by,

$$\Delta G = -RT \ln(K) \text{ where } K = (\text{folded})/(\text{unfolded}) \quad (2)$$

Where ΔG = standard Gibbs free energy (J/mol), R = constant (8.314 J/mol K), and K is the equilibrium constant for the folding/unfolding reaction. The equilibrium constant in this case is quite simplistic because of the number of structural variations, folded and unfolded, it can assume. This (equilibrium constant) assumes that the starting and ending structures are rigid in their geometry. Due to the small amount of denaturation of the protein under physiological conditions, the ratio K is measured at temperatures greater than 37°C and is corrected to reflect physiological conditions (2). The folding reaction will proceed spontaneously when K>1 and can be estimated by,

$$\Delta G = -RT \ln K \quad (3).$$

The interactions that occur between a protein and the solvent in sub-unit association are entropy driven reactions (1, 2). The reaction involves the breaking of protein-solvent bonds and the formation of protein-protein and solvent-solvent bonds in equal number as illustrated in figures 5 and 6. Entropy driven reactions require a sufficient number of weak product bonds that are supplied by apolar contacts within the proteins. For the entropy driven reactions to occur, the enthalpy value must not contribute heavily towards the effects on the overall free energy. As this happens the ΔG will vary tremendously while the temperature only changes slightly. Weber (4) postulated that small changes in the energy of the apolar bonds of proteins play an important part in optimizing the affinities of enzymes for substrates and coenzymes within a certain temperature range.

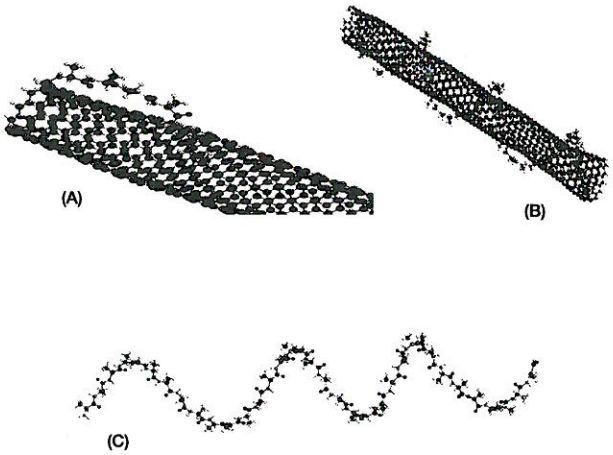

Figure 5. (A) First stage of wrapping polyalanine around (24,0) nanotubes in Spartan; Attaching the individual alanines along the nanotube allowed us to manipulate the peptide into a helical structure. (B) Next stage with polyalanine completely wrapped around the (24,0) nanotube. This allows you to see the benefits of using carbon nanotubes to mold polypeptides. (C) Final Stage of polyalanine helical with nanotube removed. As you can see the helical structure is uniform and intact.

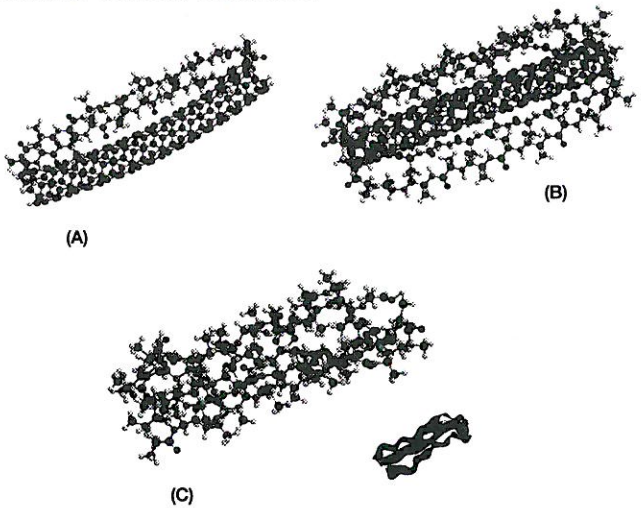

Figure 6. (A) Two strands of polyalanine stretched along a (10,0) carbon nanotube. Furthermore, SWNT's are very beneficial for molding the polyalanines into different chains. (B) Five individual strands of polyalanine stretched along the (10,0) nanotube. The polyalanine will be disconnected from the tube and formed into a sheet structure. The ends disconnected from the tube will be connected to the adjacent polyalanine chain. (C) Completed Polyalanine sheeted protein structure developed in Spartan. The structure is also displayed in ribbon form.

The unfolded states have a high level of entropy as well as high ΔG. The small dentations along the side of the free energy diagram represent semi-stable intermediates that are capable of slowing the protein folding process. To reach the protein's native conformation all the folding intermediates must be reduced to their most stable form. As the degree of folding increases both ΔG and ΔS decrease. The number of native conformations achieved within the protein increases during the folding process (4, 5).

The thermodynamics of protein folding is also affected by the bond energies and the angles associated with the protein's conformational structures. The bond energies vary due to the strain placed on Φ and Ψ angles of the amino acids in polypeptide chains and the strain of intramolecular hydrogen bond locations when overlapping occurs. Folding is a molecular response to minimize these geometric strains (5, 6, 7, 8).

The thermodynamics of protein folding are affected by a myriad of factors, which makes modeling thermodynamics of folding a difficult endeavor, because experimental results can be obtained under conditions that vary significantly from physiological conditions (9, 10, 11). However, while the methods and quantization discussed may yield inexact approximations they remain invaluable tools of investigation. The folding of a protein defines its structure which in turn dictates its functioning. In this study computational methods were incorporated to model the geometric factors and thermodynamic parameters of polyalanine. The computational experiments are divided into 4 sections and are aimed at analyzing small changes in geometric factors that may correlate with the calculated thermodynamic values;

1. Calculation of molecular volume, ΔG, ΔH, and ΔS for polyalanine with 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, and 50 residues (called a 2-50 sequence from here forward) resulting in the linear type structure.
2. Calculation of volume, ΔG, ΔH, and ΔS for 2-50 sequence for which the terminals of the peptide chain are connected and the energy reduced resulting in a globular type structure.
3. Carbon nanotubes, which are a stiff, inert compound, are used as a template to form sheeted, globular, and helical structures of polyalanine. The ΔG, ΔH, ΔS, and molecular volume are then calculated for these peptide structures.
4. Both alpha and beta secondary structures of polyalanine are constructed and evaluated for ΔG, ΔH, ΔS, and molecular volume parameters.

An important parameter in conducting this work was to be able to systematically get the geometry of the polyalanine in the same configuration each time. Specifically, a method needed to be developed whereby the exact three dimensional helical structure could be easily reproduced by different groups, a sheeted structure could be outlined in a systematic fashion or a globular structure can be confined to a specific volume. Carbon nanotubes, specifically single walled carbon nanotubes (SWNT), were selected as the template for these structures. Many materials (salts, graphite, pure metals,

etc.) do not have an independent structure that can be isolated and used as a template for nanometer sized structures. While the length of a SWNT is a variable, its diameter can be exactly reproduced.

EXPERIMENTAL

In this research the Spartan Molecular Modeling software was incorporated to visualize, construct and calculate polyalanine into sheeted, globular, and helical structures. Spartan was also used to construct carbon nanotubes, which were used to shape polyalanines. Bond angles, volume, area, phi (Φ), psi (Ψ), and the empirical formula were also calculated. Spartan '02 on desktop computers was used for constructing structures and low end calculations, and Spartan '04 (Linux, cluster) version was used for higher level work. A SUN Microsystem cluster was used in the high end calculations. The software incorporates molecular modeling, semi-empirical and Hartree-Fock methods. Much of this work was conducted with semi-empirical (PM3) for the systems with less than 600 atoms; however, molecular mechanics was used for larger systems.

RESULTS AND DISCUSSION

This study incorporated carbon nanotubes of various diameters and lengths as templates and frames for the different polyalanine structures. Both experimentally and theoretically carbon nanotubes are inert and rigid structures that provide tubular geometries for helical, sheeted or globular structures. Figure 4 A-E illustrates the construction strategy for a carbon nanotube (10,0) using sp^2 hybridized carbons. Figure 5 A-C illustrates the construction of helical polyalanine structure using a (24,0) nanotube. The diameter of the nanotube dictates the diameter of the helical structure. In Figure 6A-C a nanotube is used to construct a sheeted polyalanine structure (50 residues). Alpha and beta structures were constructed using this method. In Figure 7 A-C a (50,0) nanotube is used to mimic a cell pore (diameter $\approx$ 2 nm) and a globular polyalanine is constructed within the nanotube, forcing the globular structure to assume the nanotube's dimensions. Ramachandran plots can be easily generated for different polyalanine structures constructed here. Plots generated in this study clearly indicate that there are geometric shifts in the polypeptides as a function of the form in which they are minimized, and compares polyalanines with 50 residues for a sheeted structure (five parallel sheets), an alpha helix and a beta helix, with the alpha structure showing the widest range of angles. Figures 8, 9, 10 provide some of the computational results including entropy, enthalpy, free energy, and molecular volume for the various polyalanine structures and conformations.

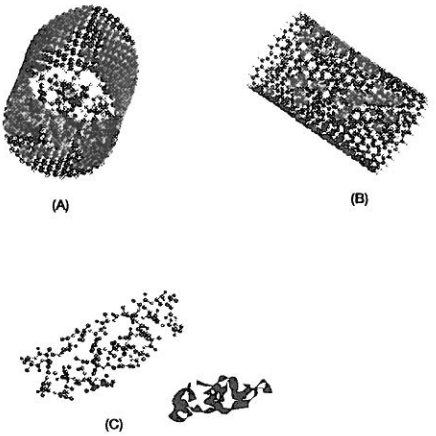

Figure 7. (A) Initial Stage of globular polyalanine being constructed inside (50,0) carbon nanotube. The first alanine is connected to a broken bond within the nanotube structure. In this experiment working inside the nanotube allowed us to keep a globular protein form that could easily pass through the small diameter of a cell pore. (B) The completed globular 50 Polyalanine protein still attached to the inner walls of the nanotube. (C) Final globular protein after being removed from inside the nanotube. The structure is also displayed in ribbon form.

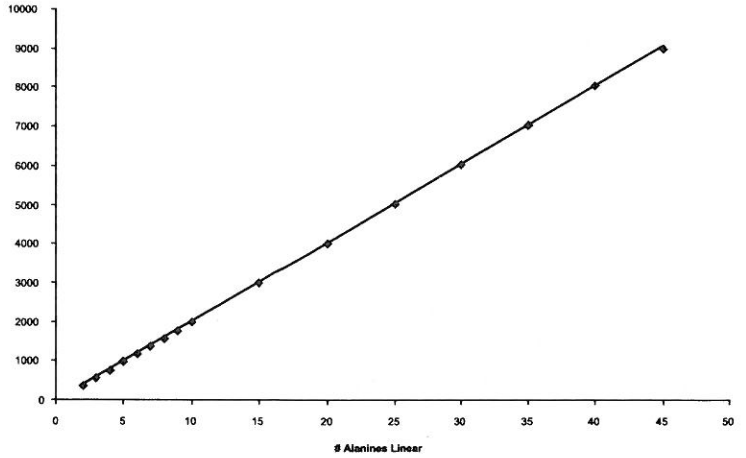

Figure 8. Plot of the calculated Gibbs free energy (kJ/mol) versus the number of alanines in the linear structure. There is a calculated ΔG per residue of 201.45 kJ for this linear structure. The ΔG_{rxn} for the transformation of $\Delta G_{Globular}$ to ΔG_{Linear} is approximately 8kJ, but this value can change significantly if different parameters are varied (i.e. density of the globular structure, the degree of hydrogen bonding in either structure, etc...).

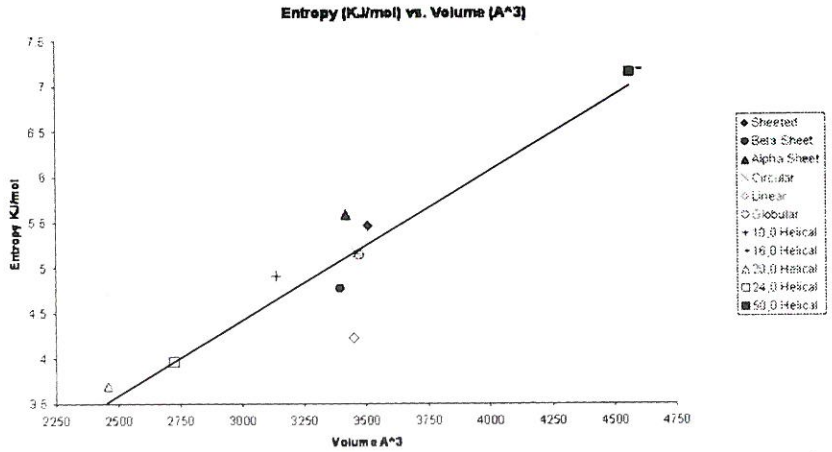

Figure 9. This figure displays the entropy (KJ/mol) versus the volume (Å^3) of multiple polyalanines (50 residues). The entropy increases as the volume of the structures increases with the exception of the linear polyalanine. The linear polyalanine is a highly order molecule, therefore the entropy will be lower in comparison with structures of a similar volume.

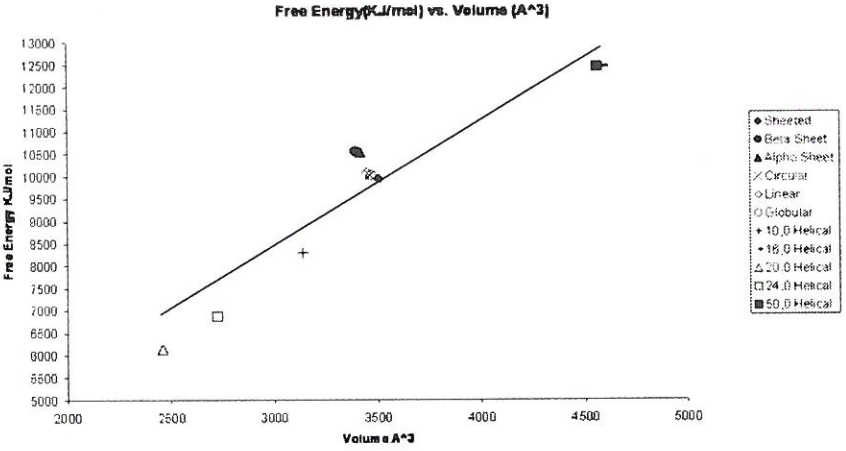

Figure 10. This figure displays the free energy (KJ/mol) versus the volume (Å^3) of multiple polyalanines (50 residues). The free energy increases with an increase in volume.

The nanotubes are removed from the polypeptide structure once the desired structure is complete, leaving the free standing polyanaline. Depending on the desired accuracy of the calculated values, the number of atoms in the structure, and the computational time constraints, molecular mechanics, semi empirical (PM3) or Hartree Fock method is chosen (12-14).

CONCLUSIONS

This work shows that carbon nanotubes can be used as templates and frames for constructing a range of protein structures. This approach is developed to build specific and reproducible peptide structures for geometric and thermodynamic studies. Their highly ordered inert structure can be used to form sheeted, helical and globular structures. Second, the computational data illustrates a number of interesting geometric and thermodynamic trends such as the variation of angles between the linear, helical, and globular structures and the difference in free energy per alanine for linear structures versus that of globular structures. Finally, this work suggests that thermodynamic values of protein folding can be estimated for the transition between one conformation and another. For example, using the equality

$$\Delta H_{rxn} = \Delta H_{products} - \Delta H_{reactants} = \Delta H_{globular} - \Delta H_{helical} \quad (4)$$

the enthalpy of the transition from a helical structure to a globular or a sheeted structure can be estimated, assuming no interactions with the solvent It should be pointed out that small differences in geometries such as the diameter of the helical structure, the degree and location of hydrogen bonding in alpha and beta sheeted structures, or the density of the globular structure can impact the thermodynamic parameters so the exact conformation/geometry needs to be known if estimates are going to be made between folded and unfolded states (eq. 2). A simple solution does not exist to the question, "Can we calculate the thermodynamics of protein or peptide folding." Two primary reasons might be cited; first because of the almost infinite number of geometric variations possible for the both the reactant and the product; second because the computational power needed for highly accurate work with a large number of atoms. On the other hand, this work does illustrate that highly ordered, reproducible structures can be constructed and estimates of the energies between these states calculated with computational methods.

REFERENCES

1. Garrett, R. H., Grisham, C. M., Biochemistry second edition. 192. (1999).
2. Lazaridis, T., Karplus, M., Thermodynamics of Protein Folding. Biophysical Chemistry. 100, 367-395 (2003).
3. Nelson, D. L., Cox, M.M., Lehninger Principles of Biochemistry fourth edition. 149-150 (2005).
4. Weber, G. Comments Mol. Cel. Biophys. 9, 201-218 (1999).
5. Dyson, H. J; Wright, PE.; Scheraga, HA. Proceedings of the National Academy of Sciences of the United States of America (2006), 103(35), 13057-13061.
6. Razvi, Abbas; Scholtz, J. Martin. Lessons in stability from thermophilic proteins. Protein Science (2006), 15(7), 1569-1578.

7. Mikulecky, P J.; Feig, AL. Heat capacity changes associated with nucleic acid folding. Biopolymers (2006), 82(1), 38-58.
8. JWH. Schymkowitz, F. Rousseau, L. Serrano, PNAS, 2002, vol. 99(25), 15846-15848
9. Benjwal, S; Verma, S; Rohm, KH; Gursky, O. Protein Science (2006), 15(3), 635-639.
10. Thompson, LC.; Walters, J; Burke, J; Parsons, JF.; Armstrong, RN.; Dirr, HW. Biochemistry (2006), 45(7), 2267-2273.
11. Topping, Traci B.; Hoch, Duane A.; Gloss, Lisa M. Folding Mechanism of FIS, the Intertwined, Dimeric Factor for Inversion Stimulation. Journal of Molecular Biology (2004), 335(4), 1065-1081.
12. Allen, C; Bailey, T; Bowen, O; Bryant, V; Lackey, J; Mills, I; Kitchens, N.,Patel, O; Richardson, N; Sermons, S; Shaw, T; Stafford, C; Stoney, T; Ward, C; Zipperer, S; Barton, I; Manning, T. Abstracts, 58th Southeast Regional Meeting of the American Chemical Society, Augusta, GA, United States, November 1-4 (2006).
13. Hayer-Hartl, M; Minton, AP. Biochemistry (2006), 45(44), 13356-13360.
14. Komori, H; Inai, Y. J. Phys. Chem. A (2006), 110(29), 9099-9107.

ACKNOWLEDGEMENTS

We'd like to thank SUN Microsystems for funding the cluster used in this work. We'd like to thank Jonathan Bryant, Ike Barton and Joe Newton of the VSU IT division for supporting this work with their expertise and resources. We'd like to thank and acknowledge Dr. John Barbas and Dr. Jim Baxter of the VSU chemistry department for their initial and ongoing use of Spartan software which laid the groundwork for this project. We'd like to thank Ms. Giso Abadi of Sunderland University who has worked with a number of students helping them improve their computational skills.