Inquiry: The University of Arkansas Undergraduate Research Journal

Volume 5 Article 11

Fall 2004

The Effects of Multiple Mutations in the Hydrophobic Core Upon the Stability of Staphylococcal Nuclease

Rebecca L. Danforth
University of Arkansas, Fayetteville

THE EFFECTS OF MULTIPLE MUTATIONS IN THE HYDROPHOBIC CORE UPON THE STABILITY OF STAPHYLOCOCCAL NUCLEASE

By Rebecca L. Danforth
Department of Chemistry/Biochemistry

Faculty Mentor: Dr. Wesley Stites
Department of Chemistry/Biochemistry

Abstract:

Previous work in the laboratory of my research advisor, Dr. Wesley Stites, has investigated the core packing of the protein staphylococcal nuclease. The core of a protein is critical in determining a protein's structure and stability. The hydrophobicity of the core has long been thought to be the principal driving force for folding, but recent work in the Stites lab has shown that optimization of van der Waals contacts and minimization of cavities, in our shorthand term, packing, is at least as energetically important. We are building upon this information in our attempt to better pack the protein core. If we can do this, we predict that the improvement in packing will make the protein more stable overall.

This project takes a closer look at the thermodynamically unfavorable left-handed alpha helix region in the core of staphylococcal nuclease. It has already been shown that there are angle strains on some of the residues that can be repaired by replacing the leucine at residue 38 with glycine. However, this created empty space within the core that greatly destabilized the protein. Our hypothesis was that by filling this space with larger amino acids at nearby locations, we would be able to correct this problem. Several mutations were made at residues 38, 39, and 125. The residues are all within close contact with each other and in the vicinity of the left-handed alpha helix. The following procedures were used: Kunkel DNA mutagenesis, transformation and preparation of M13 single stranded DNA, transfer of nuclease mutant gene from M13 to plasmid, Laemmli discontinuous protein SDS-Page gel, protein preparation and purification, and fluorometric titration.

The hypothesis that relieving angle strain near the left-handed alpha helix with a glycine, and then filling the space caused by that mutation with larger amino acids near it would increase the protein's stability was confirmed. However, new packing problems were generated so most mutations resulted in an overall decrease in stability.

Introduction:

Compared to many scientific disciplines, protein folding is a fairly new area of study (it has only been around for about 40 years) and is very highly researched. It is known that a protein's function is dictated by its shape after folding—its natural or wild-type state. That folding is a natural process that is not yet understood. It has been shown that the information needed for a protein to spontaneously fold into its biologically active three-dimensional state can be found within the protein's amino acid sequence.[1] This would imply that scientists should be able to determine a protein's three-dimensional structure as well as its function by analyzing the protein's amino acid sequence. However, this is not yet possible and there are still many gaps in the known information that must be filled.

Useful techniques for the engineering of protein stability should make it possible to create a wide array of potential medical and medicinal improvements, as well as advancements in biochemical and industrial research.[2] Proteins are commonly used in pharmaceuticals, and increasing their stability would be a great advancement for that industry because it would allow these medicines to have longer shelf lives, and, potentially, to no longer require refrigeration.

The core of a protein is critical in determining a protein's structure and stability. The hydrophobicity of the core has long been thought to be the principle driving force for folding.[2] However, recent research by the Stites group has shown that optimization of van der Waals contacts and minimization of cavities, in a commonly used shorthand term, packing, is at least as energetically important. Van der Waals interactions occur when two dipoles interact with one another; the closer these two dipoles are to each other, the stronger the interaction will be.[1]

Proteins are formed by chains of amino acids connected together in polypeptides, and then folded in various ways, forming multiple subunits that may also interact covalently.[3] During folding, the non-polar residues almost always end up in the interior of the protein, and the polar ones on the exterior. This is what makes the core of the protein hydrophobic. Non-polar residues being buried happens as a result of the aqueous environments in which proteins often reside. In order for the protein to function properly, it must be able to do so within that water, instead of being forced away by its own non-polar residues. The clustering of non-polar residues into the interior of the protein, instead of out into the aqueous environment, also

makes the protein more thermodynamically stable.[3] Other forces that contribute to protein stability are hydrogen bonding and electrostatic interactions between charged residues.[4] If the amino acids are arranged neatly, and closely within the protein, then the van der Waals interactions will contribute to the overall stability.[1] Also, it has been shown that within the core of a protein the number of van der Waals interactions is much higher than the number of van der Waals interactions that occur between an unfolded protein and the solution. This indicates that the favorability of those interactions are much more important to the protein's overall stability.[1]

Many studies have shown that in a well-packed protein core, mutations can make substantial changes.[5] Proteins accommodate for these mutations by making small shifts in the backbone and to the side torsion angles.[5] These shifts can cause the protein to become unfavorably packed and therefore more unstable. There have been attempts to correct for this computationally, but that usually requires repairing the backbone and/or correcting the torsion angles of core side chains.[6] Since the strain on these angles seems to have such an affect on the way the protein behaves, we have begun to wonder if it would be possible, by eliminating some of the angle strain that is naturally in a protein, to increase the packing, and therefore the overall stability of the protein.

Inside the core of staphylococcal nuclease, there is a specific amino acid sequence that forms a left-handed alpha helix (see Figure 1). The left-handed helix is an unusual structure and is said to have an energetically unfavorable structure.[14] This is shown in the protein's Ramachandran plot (Figure 2). The plot shows four residues in a left-handed alpha helix conformation within the core of the protein.[14] It has been theorized that by packing this helix more tightly, it is possible to increase the van der Waals interactions, and therefore, the stability of the protein. Previously, Stites *et al.* published an article in the Journal of Molecular Biology[2] in which they attempted to decrease the number of unfavorable interactions by replacing strained residues with glycine. Glycine is a common mutation found at these sites. Glycine did relieve the angle strain in this area, however, this resulted in an overall decrease in stability due to the formation of empty space near the helix, specifically the space between residues 125 and 38 (both are leucine in the wild-type state). The hole that was created here disturbs packing. From this data it is now hypothesized that, in addition to the glycine mutation, if the surrounding residues are replaced with residues that will have better steric interactions, without leaving the space between these residues empty, then it will be possible to increase the protein's stability.[14]

Staphylococcal nuclease has 42 known homologues[2] that have variations in side chains of the hydrophobic core from the wild-type sequence. The most common variations involve isoleucine, leucine, and valine. Because these are the most frequently occurring residues, they define the consensus sequence of nuclease.[2] This consensus is not in agreement with the wild-type nuclease sequence (see Table 1), showing that it is possible for more than one stable packing arrangement to exist.[2] The next thing to consider is how each of the residues affect each other. If one residue is a glycine, does that mean that another is always isoleucine? This is just one of many questions that our studies of nuclease seek to answer.

A series of papers published by Stites et al in *Biochemistry*[7-10] in 2001 established that packing is highly critical to protein stability. This paper will confirm the hypothesis that correcting the unfavorable interactions generated when replacing the leucines at residues 38 and 125 with glycines, as outlined in the previously discussed Stites paper[14], will increase the protein's stability. Table 2 provides all of the mutations that are to be made at these two residues, as well as at residue 39, to correct for the empty space generated with the glycine mutations. By filling the space, but still finding a way to correct for the unfavorable van der Waals interaction that is caused by the left-handed alpha helix in this location, it is thought that the overall stability of the core will be improved. The mutation at residue 39 is used because there seems to be a relationship between residues 38 and 39. Also, by looking at the relationship between the most common amino acids at residues 38, 39, and 125, we will be able to have a better idea of how each of those residues affect each other.

Materials and Methods:

Mutagenesis, Protein Expression, and Purification. The method used to generate the specific mutations desired in the protein was Kunkel mutagenesis.[15] Everything was done according to that previously published procedure.

Fluorometric Titration. In order to determine the free energy difference between the native and denatured states of the protein, fluorometric titration was used.[1] This free energy difference is commonly referred to as protein stability. It can be measured using an automated titrating fluorometer that was developed by the Stites lab and made by Aviv Associates.[16, 17]

The computer software on the fluorometer generated a plot of fluorescence intensity versus guanidine hydrochloride concentration. From this plot it was possible to determine the stability of the protein. A good plot had many data points at the beginning and end of the curve (the flat portions). The curve should also have been smooth, indicating that the protein had been adequately equilibrated (see Figure 3).[17] The data set and the titration curve were saved and taken to a computer that has Microsoft Excel. Once in this program, the data was put into a linear extrapolation plot of free energy versus GuHCl concentration (see Figure 4). The free energy was determined from the following equation:

$$\hat{I}G_{app} = -RT\ln K_{app}$$

$\hat{I}G_{app}$ is the apparent free energy when no denaturant is present[17] and K_{app} is found by subtracting the fluorescence intensity measured at a specific denaturant concentration (I) from the fluorescence intensity of the native state (I_n) and divided that by the fluorescence intensity of the denatured state (I_d) subtracted from I:

$$K_{app} = (I_n - I) / (I - I_d).$$

The program also provides the values for the slope of the plot of the change in free energy with respect to the change in GuHCl concentration, and the midpoint concentration (the concentration of GuHCl at which half of the protein is denatured).

For each mutation, $\hat{I}\hat{I}G$ was then calculated by the subtraction of the free energy of wild-type nuclease (5.4 kcal/mol) from the apparent free energy of the mutant. For the double and triple mutants, $£\hat{I}\hat{I}G_{single}$ was calculated by adding together the free energies of each of the single mutations that make up the double or triple mutant in question. Lastly, $\hat{I}^2G_{int}$ was calculated for the double mutants, and $\hat{I}^3G_{int}$ for the triple mutants. This was done by subtracting $£\hat{I}\hat{I}G_{single}$ from $\hat{I}\hat{I}G_{double}$ (or $\hat{I}\hat{I}G_{triple}$ as the case may be).

Results and Discussion:

All of the guanidine hydrochloride denaturation data from the single mutants can be found in Table 3. Double mutant data is in Table 4 and triple mutant data is in Table 5. As predicted, none of the single mutations by themselves showed any improvement in the stability of the protein. Most had fairly small effects, but L125Y greatly destabilizes the protein.

$\hat{I}G$ is the change in free energy as the protein is denatured. The $\hat{I}G$ for wild-type nuclease is 5.4 kcal/mol; a $\hat{I}G$ value higher than this indicates that the protein has a higher stability that the wild-type, whereas a lower $\hat{I}G$ value indicates that the protein has been destabilized. $\hat{I}\hat{I}G$ is the value of the wild-type free-energy subtracted from the mutant free-energy. $£\hat{I}\hat{I}G_{single}$ is the predicted $\hat{I}\hat{I}G$ of each mutation based on the sum of the free-energies of its single mutations. The energy of interaction ($\hat{I}^2G_{int}$) is the difference between the $\hat{I}\hat{I}G_{double}$ and the $£\hat{I}\hat{I}G_{single}$. This is the primary calculation that is used to determine the effects of the double mutations on the protein. A positive value of $\hat{I}^2G_{int}$ indicates that the combination of mutations is more favorable than expected. A negative value indicates that the mutations introduce more strain to the protein. The error for this calculation was estimated to be ±0.2 kcal/mol. Within error, L38G/V39L showed no difference from the predicted value. All other double mutations showed an increase in stability. The most significant mutation is L38G/L125Y with a $\hat{I}^2G_{int}$ of 1.2 kcal/mol. This is interesting because the single mutation of L125Y has a $\hat{I}\hat{I}G$ of -4.3 kcal/mol. This indicates that by combining 125Y with 38G, the interaction is favorable and both relieves angle strain and fills empty space. The 38G/39I mutation had a $\hat{I}G$ value of 5.5 kcal/mol, which is, within error, equivalent to wild-type; its $\hat{I}^2G_{int}$ value was 0.5 kcal/mol. This data indicates that the mutation is more favorable because it corrects for the angle strain in the area without affecting the overall stability of the protein.

The energy of interaction for the triple mutants ($\hat{I}^3G_{int}$) is the difference between the $\hat{I}\hat{I}G_{triple}$ and the $£\hat{I}\hat{I}G_{single}$, just as it is for the double mutations. All of the triple mutants had positive values for the energy of interaction, showing that the strain has been reduced. The most striking of these is 38G/39I/125Y with a $\hat{I}^3G_{int}$ of 1.8. This mutation is the most favorable of all of the mutations prepared in this study, further showing that the glycine-isoleucine combination is highly favorable, and even more so when the tyrosine is added to fill empty space, although the stability does decrease overall due to newly generated packing problems. The mutation at 38G/39L/125Y has a $\hat{I}^3G_{int}$ of 1.2, which is the same value as the $\hat{I}^2G_{int}$ for 38G/125Y. This once again shows that the leucine at residue 39 makes little difference when combined with a glycine at residue 38. Within error, the same thing is true for 38G/39L/125F and 38G/125F which have $\hat{I}^{3,2}G_{int}$ values of 0.5 and 0.6, respectively.

None of the mutations increased the overall stability of the protein; this is a problem caused by the protein's packing, and will require more investigation. However, the method of combining single mutations to fill space, and correct for strained angles and unfavorable interactions has proven to be useful. The multiple mutations investigated here are much more stable than the effects of the single mutations at each position would lead one to predict.

Conclusions:

Although none of the mutations resulted in an overall increase in stability, the energies of interaction were almost always positive. This confirms the hypothesis that correction of the angle-strain in the left-handed alpha helical region of the protein increases the stability in that region by relieving angle strain. However, the data also shows that the overall stability of the protein is decreased with almost all of the mutations. This is the result of further interactions that are generated in the original mutations. Further study would reveal what these interactions are. Correcting for these interactions would likely create more unfavorable interactions, starting a chain reaction that could quickly spread across the protein (although would probably only require correction at five or six more residues), but eventually lead to a more stable version of the entire protein. This is the packing problem. We have solved the issue of angle strain, and small unfavorable interactions. We can correct for these things, but they do not correct for the overall packing problem, and more research must be done to find a way to solve this bigger protein stability issue.

Appendix A (Figures):

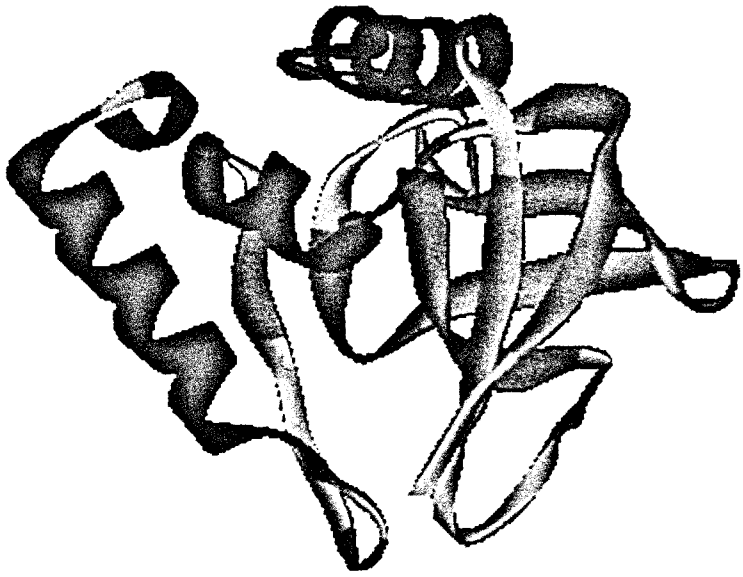
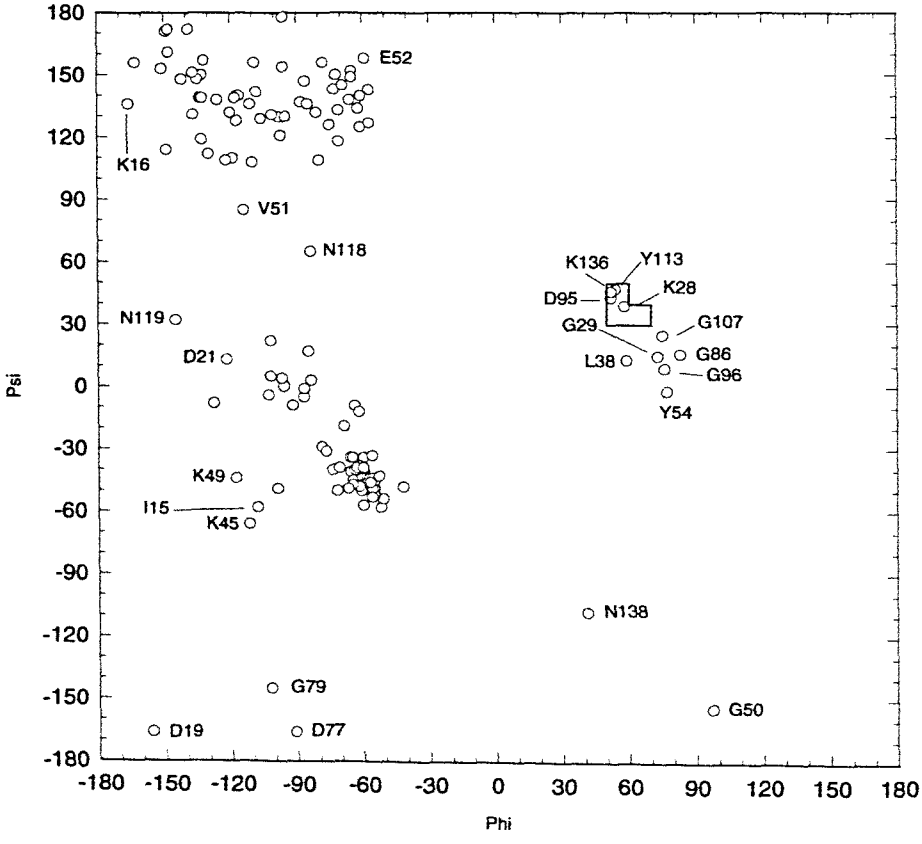


Figure 1: A ribbon diagram of staphylococcal nuclease. The left-handed alpha helix is shown in red and is displayed on the left of the molecule.



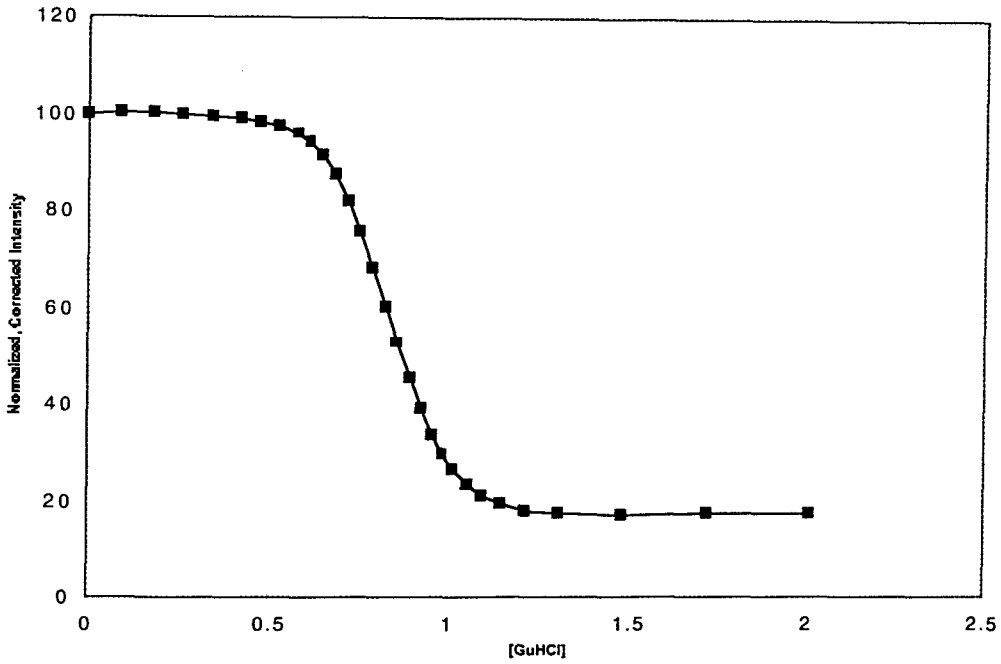


Figure 3: **A fluorescence intensity plot of wild-type staph. nuclease..**

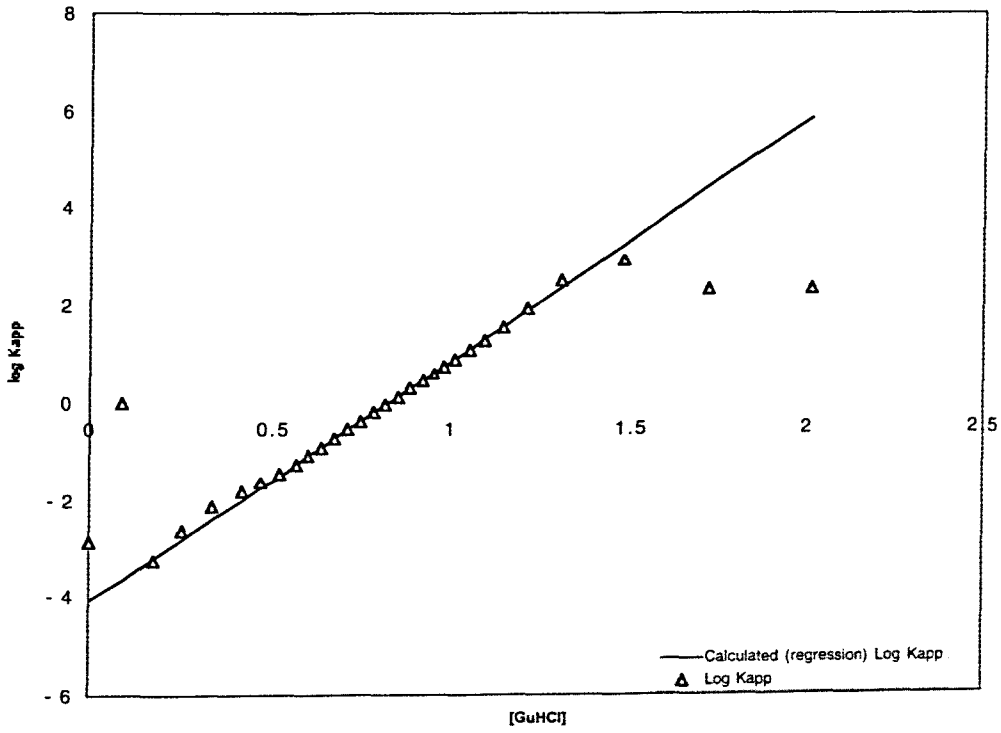


Figure 4: **The log of the denatured intensities versus guanidine hydrochloride concentration plot with a linear extrapolation to zero guanidine hyrdochloride.**

Appendix B (Tables):

Table 1: The occurrence of each amino acid at each of the three residues analyzed in this study, with the consensus and nuclease sequences.

Residue Number	*38*	*39*	*125*
Amino Acid			
Isoleucine	0	42	2
Leucine	4	1	15
Valine	3	14	2
Phenylalanine	0	0	10
Tyrosine	1	0	16
Methionine	0	1	1
Cysteine	4	0	1
Alanine	3	0	1
Aspartate	3	5	0
Glutamate	2	0	1
Glycine	35	0	0
Histidine	0	0	1
Asparagine	4	1	0
Proline	0	0	1
Glutamine	1	0	2
Serine	1	0	1
Threonine	0	0	1
All others	0	0	0
Consensus	G	I	Y
Wild-type	L	V	L

References:

1. Schwehm, J.M. Doctoral Dissertation, University of Arkansas, May 1999.
2. Stites, W.E. Protein Core Packing: Structure Function Relationships (unpublished grant proposal)
3. Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002) Biochemistry (5th ed.), W.H. Freeman and Company; New York
4. Anderson, D.E., Becktel, W.J., and Dahlquist, F.W. (1990) *Biochemistry.* **29**, 2403
5. Chen, J. Doctoral Dissertation, University of Arkansas, December 2000.
6. Baldwin, E.P. and Matthews, B.W. (1994) *Curr. Opin. Biotechnol.* **5**, 396
7. Holder, J.B., Bennett, A.F., Chen, J., Spencer, D.S., Byrne, M.P., and Stites, W.E. (2001) *Biochemistry.* **40**, 13998
8. Chen, J. and Stites, W.E. (2001) *Biochemistry.* **40**, 14004
9. Chen, J. and Stites, W.E. (2001) *Biochemistry.* **40**, 14012
10. Chen, J. and Stites, W.E. (2001) *Biochemistry.* **40**, 15280
11. Byrne, J.M. Doctoral Dissertation, University of Arkansas, August 1995.
12. Shortle, D. and Lin, B. (1985) *Genetics.* **110**, 539
13. Tucker, P.W., Hazen, E.E., and Cotton, F.A. (1979) *Mol. Cell. Biochem.* **23**, 13
14. Stites, W.E., Meeker, A.K., and Shortle, D. (1994) *J. Mol. Biol.* **235**, 27-32
15. Kunkel, T.A., Roberts, J.D., and Zakour, R.A. (1987) *Meth. In Enzymol.* **154**, 367
16. Stites, W.E., Byrne, M.P., Aviv, J., Kaplan, M. and Curtis, P. (1995) *Anal. Biochem.* **227**, 112
17. Schwehm, J.M., and Stites, W.E. (1998) *Meth. In Enzymol.* **295**, 150

Table 2: A list of all of the mutations to be made and their shorthand codes, and mutation numbers.

Position(s)	Mutation	Code	Assigned Number
Leu38	Gly	L38G	32
Leu38, Val39	Gly, Ile	L38G/V39I	33
Leu38, Val39	Gyl, Leu	L38G/V39L	34
Leu38, Val39, Leu125	Gly, Leu, Phe	L38G/V39L/L125F	141
Val39	Ile	V39I	23
Leu125	Phe	L125F	35
Leu125	Tyr	L125Y	36
Leu38, Leu125	Gly, Phe	L38G/L125F	83
Leu38, Leu125	Gly, Tyr	L38G/L125Y	84
Leu38, Val39, Leu125	Gly, Ile, Phe	L36G/V39I/L125F	139
Leu38, Val39, Leu125	Gly, Ile, Tyr	L38G/V39I/L125Y	140
Leu38, Val39, Leu125	Gly, Leu, Tyr	L38G/V39L/L125Y	142
Val39, Leu125	Ile, Phe	V39I/L125F	8
Val39, Leu125	Ile, Tyr	V39I/L125Y	
Val 39, Leu 125	Leu, Phe	V39L/L125F	
Val 39, Leu 125	Leu, Tyr	V39L/L125Y	

Table 3: **Single Mutant Data**

Mutant	m_{GuHCl} [a]	C_m [b]	$\hat{i}G_{H_2O}$ [c]	$\hat{i}\hat{i}G_{single}$ [d]
L38G	6.26	0.81	5.1	-0.3
V39I	5.3	0.83	5.3	-0.1
V39L*	6.68	0.68	4.5	-0.9
L125Y	5.85	0.20	1.1	-4.3
L125F	6.33	0.68	4.3	-1.1
WT	6.53	0.82	5.4	-

a Slope value (change in free energy with respect to the change in GuHCl concentration), units of kcal $mol^{-1}M^{-1}$. Error estimated to be ±0.09.
b Midpoint concentration (concentration of GuHCl at which half of the protein is denatured), units of mol/L. Error estimated to be ±0.01 M.
c Free energy difference between native and denatured states in the absence of denaturant, units of kcal/mol. Error estimated to be ±0.1 kcal/mol.
d Difference between the stability of wild-type protein, and the apparent stability of the mutant. Error estimated to be ±0.17 kcal/mol.
* data previously published in
Biochemistry, **40**, 46, 2001, 13999***Table 4**: *Solvent denaturation data for double packing mutants.*

Table 5: **Triple mutant data.**

Mutant	ΔG_{H_2O} [a]	C_m [b]	m_{GuHCl} [c]	$\Delta\Delta G_{triple}$ [d]	$\Sigma\Delta\Delta G_{single}$ [e]	$\Delta^3 G_{int}$ [f]
L38G/V39L/L125Y	1.1	0.21	*5.31*	-4.3	-5.5	1.2
L38G/V39I/L125Y	2.5	0.41	*6.07*	-2.9	-4.7	1.8
L38G/V39L/L125F	3.6	0.65	*5.55*	-1.8	-2.3	0.5
WT	5.4	0.82	*6.53*	-	-	-

[a] Free energy difference between native and denatured states in the absence of denaturant in units of kcal/mol. Error is estimated to be ± 0.1 kcal/mol.
[b] Midpoint concentration (concentration of guanidine hydrochloride at which half of the protein is denatured) in units of M. Error is estimated to be ± 0.01 M.
[c] Slope value (change in free energy with respect to change in guanidine hydrochloride concentration) expressed relative to wild-type value of 6.53 kcal/(molΣM). Error is estimated to be ± 0.02.
[d] Difference in free energy between the free energy of the protein with triple substitutions and the free energy of wild-type protein. DDG = DG_{H2O} (triple mutant) - 5.4 (WT). Error is estimated to be ± 0.2 kcal/mol.
[e] The sum of the DDG_{single} values of corresponding single substitutions.
[f] D^3G_{int} = DDG_{triple} - $SDDG_{single}$. Error is estimated to be ± 0.2 kcal/mol
*Data is original, table and footnotes are from reference 9.