# Stability Studies of Amino Acid Substitutions at Tyrosine 27 of the Staphylococcal Nuclease $\beta$-Barrel[†]

Mulki G. Bhat,[‡,§] Lisa M. Ganley,[‡,||] David W. Ledman,[⊥,@] Margaret A. Goodman,[⊥,#] and Robert O. Fox*[,‡,@]

*The Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511, and The Howard Hughes Medical Institute at Yale University, New Haven, Connecticut 06511*

*Received April 15, 1997; Revised Manuscript Received July 22, 1997*[⊗]

ABSTRACT: In order to help determine the extent to which side chain interactions within the staphylococcal nuclease $\beta$-barrel affect its global stability, a full set of point mutants was generated for residue 27. Intrinsic tryptophan fluorescence was monitored during solvent denaturation with guanidine hydrochloride (GuHCl) and was used to calculate $\Delta G_{H_2O\ unfolding}$ and $m$ values for each mutant. In the wild type protein, residue 27 is a tyrosine which is at the first position of a type I′ $\beta$-turn, and which participates in both hydrophobic interactions and side chain to side chain hydrogen bonding. The hydrophobicity of the mutant residue was found to be the dominant factor in determining global protein stability within this series of nuclease mutants.

Staphylococcal nuclease is a small, calcium-dependent, globular protein of 149 residues. It contains no disulfide bonds and unfolds reversibly. Extensive work has been done in measuring the stabilities of nuclease mutants using solvent denaturation to extrapolate the Gibbs free energies of unfolding of these mutants (Green & Shortle, 1993; Green et al., 1992, and references therein). Many of these studies have focused on changes in the relative stabilities of single and multiple mutants and on the effects of amino acid substitutions on the structure and stability of the mutant proteins.

The site chosen for examination was Tyr 27. It is the first position of a type I′ $\beta$-turn, comprising residues 27−30 and linking antiparallel strands 2 and 3 (Loll & Lattman, 1989; Hynes & Fox, 1991), also known as strands $\beta$2 and $\beta$3 (Wang et al., 1995; Wang & Shortle, 1995; see Figure 1A) which form part of the early-forming $\beta$-meander. The reasons for examining this site are threefold. As the first position in a $\beta$-turn, it is one of the proposed sites for the initiation of hydrogen bonding and thus of native structure formation (Jacobs & Fox, 1994). Studies of a set of nuclease mutants having multiple random substitutions in this turn have shown that only 15% of the examined mutants exhibited detectable enzymatic activity, indicating that the conformation of this $\beta$-turn is important in determining overall protein stability (M. A. Goodman and R. O. Fox, unpublished data). Second, the tyrosine side chain is involved in both hydrophobic and hydrogen bonding interactions with the surrounding residues in the native state. The tyrosine's aromatic ring lies on the protein's surface above, and roughly perpendicular to, the aromatic rings of Phe 34 and Phe 76, and is covered by Lys 28. The side chain has a solvent accessibility of 23.24 Å$^2$ calculated with X-PLOR (Brünger, 1992) using a 1.6 Å probe. The hydroxyl group of the tyrosine lies within 3 Å of the carboxyl group of Glu 10, indicating the presence of a hydrogen bond (Hynes & Fox, 1991; see Figure 1B). The tyrosine's multiple interactions thus seem to allow it to act as an anchor for this particular $\beta$-turn, not only because of the regular amide hydrogen−carbonyl oxygen hydrogen bond between the $n$ and the $n + 3$ residues of a $\beta$-turn but also because of the side chain's interactions. Since this residue is making both hydrophobic and hydrogen bonding interactions, the stability data reported herein should provide a clearer picture of the relative influences of these two factors in the $\beta$-barrel core and their affect on the overall stability of nuclease. Third, this residue, as part of a $\beta$-turn, forms one of the basic secondary structural units of the nuclease $\beta$-barrel, which recent data suggest may be a distinct subdomain (Ermácora et al., 1996; Griko et al., 1994).

## EXPERIMENTAL PROCEDURES

*Mutagenesis*. These single-position mutants were created using a cassette mutagenesis method developed by Goodman (1989). Briefly, a M13 phage vector was created which contains the wild type (WT) nuclease gene from *Staphylococcus aureus* (Foggi strain) with a cassette introduced between codons 23 and 35. Partially double-stranded oligonucleotides were cloned into the cassette and identified by DNA sequencing. The mutated nuclease gene thus created was subcloned into the expression plasmid pAS1 under $\lambda$pL promoter control, introduced into *Escherichia coli* strain AR120, and screened for nuclease activity as described previously (Goodman, 1989).

*Protein Growth and Purification*. AR120 cells containing the pAS1 plasmid were grown in 2xYT media at 37 °C. Nuclease expression was induced with nalidixic acid at an $A_{600}$ of 0.6−0.8 (Fox et al., 1986; Hynes, 1989). Four hours after induction, cells were harvested by centrifugation

[†] This work was supported by NIH Grants RO1 AI 23923 and GM 51332.

* To whom correspondence should be addressed.

[‡] The Howard Hughes Medical Institute at Yale University.

[§] Present address: The University of Pennsylvania School of Medicine, Philadelphia, PA.

[||] Present address: The Department of Pathobiology, University of Connecticut, Storrs, CT.

[⊥] Yale University.

[@] Present address: The Department of Human Biological Chemistry and Genetics, The University of Texas Medical Branch at Galveston, Rm. 634 B5B, Galveston, TX 77555-0647.

[#] Present address: The Department of Biology, Wittenberg University, Springfield, OH.

[⊗] Abstract published in *Advance ACS Abstracts*, September 15, 1997.

S0006-2960(97)00876-3 CCC: $14.00 

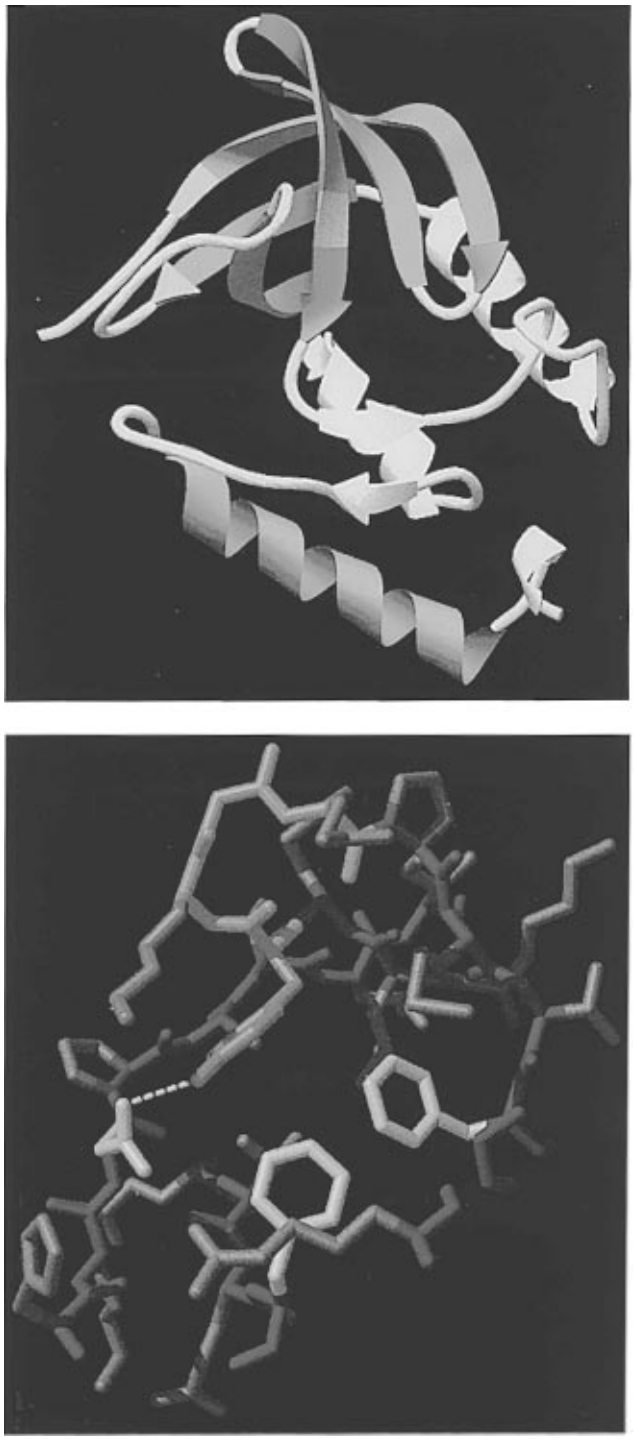

FIGURE 1: (A, top) Ribbon drawing of wild type staphylococcal nuclease (residues 6−141). The $\beta$-barrel is in green with the backbone of Tyr 27 in red and the backbones of Glu 10, Phe 34, and Phe 76 in yellow. α-Helix 3 is in cyan. The full 149 residues are not shown as the N and C termini are poorly defined in crystal structures and presumably adopt a random coil conformation. Coordinates are from 1STN, from the Brookhaven Protein Data Bank (Hynes & Fox, 1991). The figure was made using the program Ribbons (Carson, 1991). (B, bottom) Residues of wild type staphylococcal nuclease that are within 8 Å of Tyr 27. The Tyr 27 side chain is in red and is shown with the ring nearly edge-on. In yellow are the side chains of Glu 10 (shown with a dotted yellow line to represent hydrogen bonding to Tyr 27), Phe 34 (lower left of Tyr 27), and Phe 76 (left of Tyr 27). The Lys 28 side chain is in magenta. The $\beta$-turn backbone (residues 27−30) is in green, with the other backbone atoms in blue. All other side chains are in gray. Coordinates are from 1STN, from the Brookhaven Protein Data Bank (Hynes & Fox, 1991). The figure was made using the program MidasPlus from the Computer Graphics Laboratory of the University of California, San Francisco (Ferrin et al., 1988).

and frozen at −20 °C for later use. Cell pellets were thawed in 100 mL of 100 mM sodium phosphate (pH 7.5) with 100 $\mu$M $CaCl_2$ (buffer A) and sonicated on ice using a Branson Sonifier for five intervals of 30 s. Cell debris was removed by centrifugation at 20 000 rpm for 30 min at 4 °C, and the supernatant was stored on ice. The resulting pellet was resuspended in 20 mL of 100 mM sodium phosphate at pH 7.5 (buffer B) and spun at 20 000 rpm at 4 °C for 15 min, pooled, and spun at 20 000 rpm at 4 °C for 15 min. The supernatant was applied to a 100 mL S-Sepharose (Pharmacia) column and washed with buffer B, and nuclease was eluted with 0.6 M NaCl in 100 mM sodium phosphate (pH 7.5) at 30 mL/min. The protein was dialyzed overnight in 4 L of buffer A, and the precipitate was removed by centrifugation at 20 000 rpm at 4 °C for 60 min. The supernatant was run through an 80 mL Q-Sepharose (Pharmacia) column equilibrated with 50 mM HEPES at pH 7.8, and the effluent was dialyzed overnight in 4 L of buffer A. Approximately 50 mg of protein was loaded onto a Millipore SP MemSep 1500 cartridge equilibrated with buffer B. The nuclease was eluted in a linear NaCl gradient from 0 to 200 mM. The main peak was collected and consecutively dialyzed for a minimum of 8 h each against 4 L of 2 M NaCl and 4 L of 0.2 M NaCl and twice against 4 L of distilled water. The protein was frozen at −70 °C, lyophilized, and stored at −20 °C.

*Mass Spectrometry*. Mass spectra for confirmation of the identity and purity of the protein samples were acquired on a VG Platform single quadrupole mass spectrometer. The ion source was at 3.2 kV. Typically, the instrument was scanned from *m*/*z* 600 to 1600 in 10 s; 10 scans were acquired for each sample. Horse heart myoglobin, with an average mass of 16 951.50 Da (Sigma), was used as the standard for instrument calibration. Samples of protein at approximately 0.01 mg/mL in 50% 2-propanol and 0.1% formic acid were introduced into the spectrometer at 5 $\mu$L/min. Nitrogen at 40 mL/min was used to pneumatically assist the electrospray process. Y27C was reduced with 2 mM DTT (dithiothreitol) for 4 h at room temperature prior to the addition of acidic propanol for mass spectral analysis.

*Protein Identity and Purity Confirmation*. All pAS1 plasmid constructs were sequenced in the $\beta$-turn region to ensure the identity of each mutation at codon 27. Purified proteins were estimated to be at least 95% pure by SDS−PAGE and had an $Abs_{280/260}$ of 2.5 or higher, indicating an absence of significant amounts of nucleic acid. Purified proteins were examined by mass spectrometry to confirm mutant identity; each protein showed a mass within 2 Da of the calculated value (Table 1). In addition, a randomly selected group of protein samples was also analyzed in the negative ion mode with 50% 2-propanol and 0.1% ammonium hydroxide as the solvent. No species with a mass of greater than 800 Da were detected with an ion current of greater than 0.1% of that obtained for the protein in the positive ion mode. This suggests that any DNA present was negligible.

*Fluorescence Procedure*. To measure the Gibbs free energy of folding, the intrinsic fluorescence of the tryptophan residue at position 140 was monitored as each protein was denatured by addition of guanidine hydrochloride in order to observe the apparent equilibrium between the native and the unfolded states. All fluorescence measurements were performed in a temperature-controlled cuvette holder at 20.0 °C with excitation at 295 nm and emission intensity recorded

Table 1: Mass Analysis of Substituted Nuclease Proteins ($MH^+$)

| residue 27 | calculated mass (Da) | measured mass (Da) | error (Da) | percent error |
|---|---|---|---|---|
| Ala | 16 720.24 | 16 721.24 | 1.00 | 0.006 |
| Arg | 16 805.35 | 15 803.83 | −1.52 | −0.009 |
| Asn | 16 763.26 | 16 762.58 | −0.68 | −0.004 |
| Asp | 16 764.25 | 16 762.84 | −1.41 | −0.008 |
| Cys | 16 752.30 | 16 751.54 | −0.76 | −0.005 |
| Gln | 16 777.29 | 16 778.19 | 0.90 | 0.005 |
| Glu | 16 778.27 | 16 779.01 | 0.74 | 0.004 |
| Gly | 16 706.21 | 16 706.58 | 0.37 | 0.002 |
| His | 16 786.30 | 16 787.25 | 0.95 | 0.006 |
| Ile | 16 762.32 | 16 764.12 | 1.80 | 0.011 |
| Lue | 16 762.32 | 16 761.14 | −1.18 | −0.007 |
| Lys | 16 777.33 | 16 777.64 | 0.31 | 0.002 |
| Met | 16 780.36 | 16 780.29 | 0.07 | 0.000 |
| Phe | 16 796.34 | 16 795.50 | −0.84 | −0.005 |
| Pro | 16 746.28 | 16 745.69 | −0.59 | −0.004 |
| Ser | 16 736.24 | 16 736.04 | −0.20 | −0.001 |
| Thr | 16 750.27 | 16 751.05 | 0.78 | 0.005 |
| Trp | 16 835.37 | 16 834.21 | −1.16 | −0.007 |
| Tyr | 16 812.34 | 16 813.05 | 0.71 | 0.004 |
| Val | 16 748.29 | 16 749.12 | 0.83 | 0.005 |

at 325 nm. These wavelengths were chosen as they were used in the original Shortle (1986) paper and in subsequent studies of nuclease. Lyophilized proteins were freshly suspended before each experiment in 100 mM NaCl and 25 mM potassium phosphate (pH 7.0) to a concentration of 40−50 $\mu$g/mL as measured by UV spectroscopy with $A_{280} = 0.93$/mg of protein/(mL of solution) (Goodman, 1989). Before denaturation, the protein solutions were stirred rapidly at 37 °C for 1 h to ensure the breakup of any protein aggregation remaining from the lyophilization process (D. Shortle, personal communication). A 6.00 M solution of guanidine hydrochloride [determined by refractive index (Nozaki, 1972)] was prepared with a small amount of Tris base to give a pH of 7.0 upon dilution to 1 M (Shortle, 1986). A single preparation of guanidine hydrochloride solution was used.

A calibrated Harvard Apparatus model 55-1111 syringe pump was used to transfer the 6.00 M guanidine hydrochloride solution from a Hamilton gastight syringe into 2.00 mL of the above protein solution held in a 1.0 × 1.0 cm path Suprasil quartz (Hellma Cells) fluorescence cuvette. The denaturant was added at a constant rate of 0.25 mL/h for at least 4 h, and the sample was vigorously mixed with a round Scientific Products Cell Stirring Bar magnetic stirrer to ensure that the sample remained homogeneous and at equilibrium at all times. The accuracy of the syringe pump was verified periodically during the series of experiments by confirming that, when loaded with distilled water and set to the above delivery rate, it delivered 1.00 g of water to the cuvette. The guanidine hydrochloride solution was added in close proximity to the upper surface of the stirring bar to ensure thorough mixing. The emission intensity at 325 nm was collected and averaged over 1 s every 60 s during the entire experiment using a FluoroMax (Spex Industries) fluorimeter. Unfolding data for some unstable proteins were also obtained in the presence of stabilizing ammonium sulfate (0.20−0.75 M in both solutions), with all other buffer conditions maintained.

Wild type and Y27K mutant protein were test denatured for 1, 2, 3, 4, 5, and 10 h and analyzed as described below. For both proteins, stability and slope values were the same when calculated from the data collected during either the 3, 4, 5, or 10 h denaturations.

*Data Analysis.* Data for all proteins other than the four unstable mutants (Y27D, Y27E, Y27G, and Y27P) were analyzed by simultaneously fitting all of the data points to equations that describe a protein undergoing denaturation as a two-state process. The program Excel (Microsoft) was used to fit the data to eqs 1 and 2 (Bolen & Santoro, 1988) using a nonlinear least-squares analysis:

$$I = [(a_N[\text{GuHCl}] + b_N) + (a_D[\text{GuHCl}] + b_D) \times \exp(-\Delta G'/RT)]/[1 + \exp(-\Delta G'/RT)] \quad (1)$$

$$\Delta G' = \Delta G_{H_2O\ \text{unfolding}} - m_{\text{GuHCl}}[\text{GuHCl}] \quad (2)$$

where $I$ is the monitored fluorescence intensity, $a_N$ and $a_D$ are the slopes of the native and denatured baselines, respectively, $b_N$ is the intensity of the native protein at 0 M GuHCl, $b_D$ is the extrapolated intensity at 0 M GuHCl, and $\Delta G'$ is the apparent change in free energy of unfolding at a given GuHCl concentration. $\Delta G'$ is assumed to vary linearly according to eq 2, in which $\Delta G_{H_2O\ \text{unfolding}}$ is the extrapolated free energy of unfolding at 0 M GuHCl. The fit was performed with parameters that allowed for the largest radius of convergence from all possible starting values to a single set of solution values for the six variables in eqs 1 and 2 ($a_N$, $a_D$, $b_N$, $b_D$, $m_{\text{GuHCl}}$, and $\Delta G_{H_2O\ \text{unfolding}}$). This ensured that the same solution values for these variables would be found for the same set of data points independent of the starting values of the variables. All six variables were solved as independent values in the equation-fitting analysis. $\Delta G_{H_2O\ \text{unfolding}}$ and $m_{\text{GuHCl}}$ values were obtained directly from the fits of the data to eqs 1 and 2; $C_m$ was calculated by comparison of decreasing fluorescence to the $b_N$ and $b_D$ values obtained from the equation-fitting procedure. The equation-fitting method of stability analysis was selected because the simultaneous solving of all of the equation's six parameters suggested less error in the results than the more commonly used method of fitting the pre- and post-transition phases to baselines, calculating the apparent $K_{eq}$, and extrapolating apparent $\Delta G_{H_2O\ \text{unfolding}}$ values back to 0 M GuHCl as described in detail in Shortle (1986). Using this previously described method, slight changes in the slopes of the fitted lines can lead to significant differences in calculated $\Delta G_{H_2O\ \text{unfolding}}$ values. In addition, many of the mutants studied here were significantly destabilized, causing a shortened pretransition phase and correspondingly greater error in baseline fitting.

The method of two-state equation fitting was used for all proteins, except in cases where the protein was unstable ($\Delta G_{H_2O\ \text{unfolding}} \leq 0$ kcal/mol), causing a significant fraction of the protein to be denatured at 0 M GuHCl and thus leaving only a fraction of the transition phase observable during denaturation. In these cases, a protein solution was made and a denaturation curve obtained as described above. The post-transition baseline of the denaturation curve was extrapolated back to 0 M GuHCl to obtain the fluorescence of fully unfolded protein ($I_u$). Ammonium sulfate was added to another sample of the same protein solution to stabilize the protein. The rise in intrinsic fluorescence was monitored until it reached a plateau which was taken as the fluorescence of the protein in the fully folded native state ($I_n$), and the $K_{app\ eq}$ was taken as the ratio $(I_n - I_i)/(I_i - I_u)$, where $I_i$ was the initial fluorescence intensity of the solution at 0 M GuHCl (Shortle et al., 1990).

Because of the apparent three-state denaturation of Y27G and Y27P, these mutants were also analyzed by simultaneous

equation fitting to equations that described a three-state transition. The equations used were eqs 3−6, based on those of Barrick and Baldwin (1993):

$$I = [(I_N + I_I K_{N-I} - I_U + I_U K_{N-I})/(1 + K_{N-I} + K_{N-I}K_{I-U})] + I_U \quad (3)$$

in which

$$K_{N-I} = \exp[-(\Delta G^\circ_{N-I\ H_2O\ unfolding} - m_{N-I}[GuHCl])/RT] \quad (4)$$

$$K_{I-U} = \exp[-(\Delta G^\circ_{I-U\ H_2O\ unfolding} - m_{I-U}[GuHCl])/RT] \quad (5)$$

$$I_S = I_S^\circ - m_S[GuHCl] \quad (6)$$

where S can represent the native N, the intermediate I, or the unfolded U state. In these equations, $K_{N-I}$ and $K_{I-U}$ are the equilibrium constants between the native and intermediate and the intermediate and unfolded states, $\Delta G^\circ_{N-I\ H_2O\ unfolding}$ and $\Delta G^\circ_{I-U\ H_2O\ unfolding}$ are the changes in free energy in these unfolding transitions extrapolated to water (i.e. no denaturant), $I$ is the overall fluorescence intensity (the data collected), $I_N$, $I_I$, and $I_U$ are the intensities of the native, intermediate, and unfolded states at a given denaturant concentration, respectively, and $I_N^\circ$, $I_I^\circ$, and $I_U^\circ$ are the intensities of fluorescence at these states extrapolated to 0 M GuHCl. The free energies and the fluorescence intensities were assumed to vary linearly with denaturant concentration by the $m$ (slope) values, where $m_{N-I}$ and $m_{I-U}$ correspond to the $m_{GuHCl}$ values in Table 2. During the equation fitting, $I_N^\circ$ and $I_U^\circ$ were proportioned to each other on the basis of ratios obtained from denaturation in high ammonium sulfate in order to compensate for the "missing" native state in the data set. However, $\Delta G^\circ_{N-I\ H_2O\ unfolding}$, $\Delta G^\circ_{I-U\ H_2O\ unfolding}$, $I_I^\circ$, and all of the $m$ values were solved as independent variables in the equation fitting.

## RESULTS

Figure 2 illustrates results typical of the fluorescence denaturation and data collection procedures described above. The curves trace the fraction of initial fluorescence ($F_0$) of Trp 140 excited at 295 nm and monitored at 325 nm. The slight differences in the fraction of fluorescence in the denatured state are likely due to differences in the baseline count between different experiments and to variations in the protein concentration. The higher fluorescence level of unfolded Y27W is due to the presence of a second tryptophan in the protein. The wide range of slopes and stabilities obtained for this set of mutants is visible in this figure.

Apparent overall $\Delta G_{H_2O\ unfolding}$, $C_m$, and $m_{GuHCl}$ values calculated for these 19 mutants are summarized in Table 2. Relative to the wild type protein, some mutants display increased or decreased $m_{GuHCl}$ values (hereafter referred to as $m_{GuHCl}+$ and as $m_{GuHCl}-$ mutants, respectively). All but one of the mutant proteins were significantly less stable than that of the wild type, the exception being the Y27F mutant, which displayed equal apparent stability. The $C_m$ and $\Delta G_U$ values are correlated with the exception of Y27I and Y27V which show higher $C_m$ values than mutants of comparable stability due to their longer transition range (smaller slope). In general, the mutants exhibited slopes equal to or less than that of the wild type.

Table 2: Guanidine Hydrochloride Denaturation Data for Point Mutants at Position Y27 of Staphylococcal Nuclease[a]

| protein | apparent $\Delta G_{H_2O}$ unf[b] (kcal/mol) | $C_m$[h] (M) | slope $m_{GuHCl}$[e] | amino acid volumes[g] (Å$^3$) |
|---|---|---|---|---|
| WT | 5.0 | 0.84 | 1.00 | 203.6 |
| WT | 5.4[d] | 0.84[d] | 1.10[d] | |
| Y27A | 2.2 | 0.29 | 1.07 | 91.5 |
| -C | 2.5,[d] 2.1[f] | 0.41[d] | 1.02[d] | 105.6 |
| -D | −0.4[c] | | | 124.5 |
| -E | 0.0[c] | | | 155.1 |
| -F | 5.0 | 0.77 | 1.10 | 203.4 |
| -G | −0.1[c] | | | 66.4 |
| -H | 3.4 | 0.48 | 1.19 | 167.3 |
| -I | 2.5 | 0.49 | 0.81 | 168.8 |
| -K | 1.1, 1.2[c] | 0.22 | 1.07 | 171.3 |
| -L | 3.5 | 0.57 | 1.07 | 167.9 |
| -M | 3.0 | 0.52 | 0.98 | 170.8 |
| -N | 1.0 | 0.22 | 1.15 | 135.2 |
| -P | −1.1[c] | | | 129.3 |
| -Q | 1.8 | 0.26 | 1.27 | 161.1 |
| -R | 2.1 | 0.30 | 1.29 | 173.4 |
| -S | 2.0 | 0.26 | 1.32 | 99.1 |
| -T | 1.7 | 0.32 | 1.07 | 122.1 |
| -V | 2.0 | 0.45 | 0.76 | 141.7 |
| -W | 4.4 | 0.70 | 1.08 | 237.6 |

[a] Data collected in 100 mM NaCl/25 mM sodium phosphate buffer at pH 7.0 and 20.0 °C. Data obtained by two-state equation fitting or by sulfate renaturation. [b] Average estimated error of less than ±0.1 kcal/mol. [c] Values obtained by sulfate renaturation as described in Shortle et al. (1990); $C_m$ and $m_{GuHCl}$ values are not given as Y27D, -E, -G, and -P are significantly unfolded at 0 M GuHCl and much of the transition phase cannot be observed. [d] Buffer contained 1 mM DTT. [e] Values normalized to the WT value of 5.9 kcal mol$^{-1}$ M$^{-1}$; estimated average error of less than ±0.07. [f] Assumed true value for Y27C in the absence of DTT as calculated from the WT stability values with and without DTT. [g] Values from Chothia (1975) except for that of Arg from Zamyatnin (1972). [h] M refers to the molarity of guanidine hydrochloride.

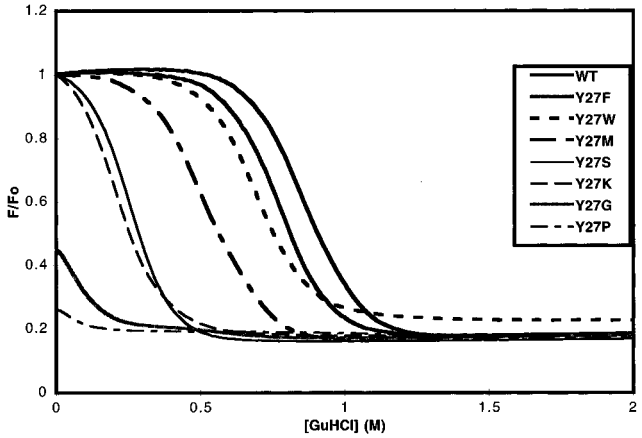


FIGURE 2: Various denaturation curves. All curves are normalized to the initial level of fluorescence except for those of Y27G and Y27P, which are scaled in proportion to the other curves using their post-transition (fully denatured) baselines. The slight differences in the post-transition baselines are due to differences in blank counts and to slight variations in protein concentration except for that of Y27W where the high unfolded baseline is due to the presence of a second fluorescent tryptophan.

Four mutants were destabilized enough to exhibit $\Delta G_{H_2O\ unfolding}$ values equal to or less than 0.0 kcal/mol, so $C_m$ values could not be determined. As the $\Delta G_{H_2O\ unfolding}$ values for these proteins in Table 2 were measured by using ammonium sulfate renaturation to find $K_{eq}$ at 0 M GuHCl and not by using equation fitting, $m_{GuHCl}$ values could not be determined reliably. As a simple cross-check, the ammonium sulfate renaturation method was also used to measure the stability of the Y27K protein, and the result with

Table 3: Guanidine Hydrochloride Denaturation Data for Y27G and Y27P[a]

| | apparent $\Delta G_{H_2O\ unfolding}$ for the N to I transition | $m_{GuHCl}$ for the N to I transition[c] | parent $\Delta G_{H_2O\ unfolding}$ for the I to U transition[b] | $m_{GuHCl}$ for the I to U transition[c] |
|---|---|---|---|---|
| Y27G | −0.1 | 1.69 | 3.3 | 0.98 |
| Y27P | −0.8 | 1.31 | 5.1 | 0.69 |

[a] Data collected in 100 mM NaCl/25 mM sodium phosphate buffer at pH 7.0 and 20.0 °C. Data obtained by three-state equation fitting. [b] Average estimated error of less than ±0.1 kcal/mol. [c] Values normalized to the WT value of 5.9 kcal mol$^{-1}$ M$^{-1}$; estimated average error of less than ±0.07.

this method, 1.2 kcal/mol, agreed with the value obtained from the standard equation fit of 1.1 kcal/mol.

The value for the estimated average error in the stability values in Table 2 (±0.1 kcal/mol) is based on repeated analysis of different sets of data of the same proteins under identical conditions; the average estimated error for the $m_{GuHCl}$ values normalized to the wild type slope is less than ±0.07. The range of $\Delta G_{H_2O\ unfolding}$ for the wild type was from 4.9 to 5.2 kcal/mol with a standard deviation of 0.09 kcal/mol. Ranges of $\Delta G_{N-I\ H_2O\ unfolding}$ values for Y27G and Y27P respectively were from −0.06 to −0.1 kcal/mol and from −0.8 to −0.9 kcal/mol; ranges of $\Delta G_{I-U\ H_2O\ unfolding}$ values for Y27G and Y27P respectively were from 3.2 to 3.4 kcal/mol and from 5.1 to 5.3 kcal/mol. Correlation coefficients were also calculated for those proteins that were stable enough to be analyzed by the equation-fitting procedure described above. Equations 1 and 2 fit the 240 data points for each protein denaturation curve extremely well with correlation coefficients of at least 0.9998, with many being 0.9999 or better.

The results for the three-state equation fitting for Y27G and Y27P are shown in Table 3. Both mutants show a great deal of instability of the native state under these experimental conditions, as can be seen from the favorable difference for unfolding free energies between the native and the intermediate state, as well as the steep slope of this transition. In contrast, the intermediate state is relatively stable, and as indicated by the fluorescence levels in panels A and B of Figure 3 seems to be structurally close to the fully unfolded state. The uncertainties for the values in Table 3 were determined in the same manner as above; correlation coefficients between the data and the equation-fitted curve were all 0.9989 or better. The series of denaturations of Y27G in varying concentrations of ammonium sulfate were also examined (Figure 4).

The hydrophobic mutations resulted in the most stable proteins. As might be expected due to its similarity to the wild type tyrosine residue in volume and in shape, the phenylalanine substitution gives the most stable mutant, although with a significant 10% increase in the *m* value compared to that of the wild type. The hydrophobic isoleucine, methionine, and leucine mutants all have similar side chain volumes (see Table 2) which are smaller than that of phenylalanine. This similarity in size and charge thus makes these mutations especially useful in observing the effects of shape on substituents. We would expect these proteins to have roughly equal stabilities of less than 5.0 kcal/mol (the stability of Y27F) if side chain hydrophobicity and volume were the only significant factors in creating stable substitutions at this position. However, this is not the case, as Y27I is the least stable of the three with a $\Delta G_{H_2O\ unfolding}$ value of 2.5 kcal/mol. Y27L has a stability of 3.5 kcal/mol, while Y27M has an intermediate stability of 3.0 kcal/mol. The $m_{GuHCl}$ values also change significantly, but with the least stable (Y27I) denaturing with the most reduced slope.

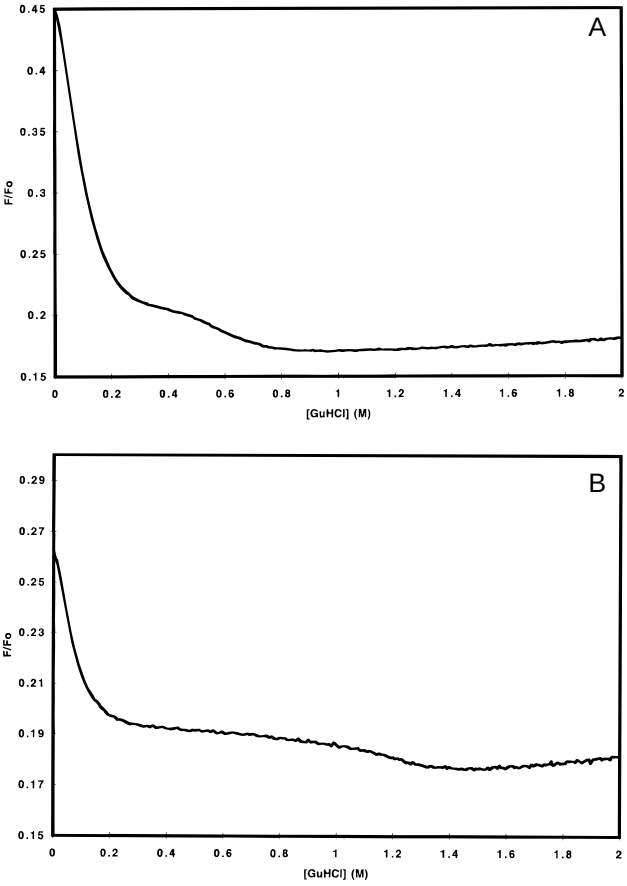


FIGURE 3: (A) Exaggerated scale plot of the Y27G denaturation curve shown in Figure 2. (B) Exaggerated scale plot of the Y27P denaturation curve shown in Figure 2.

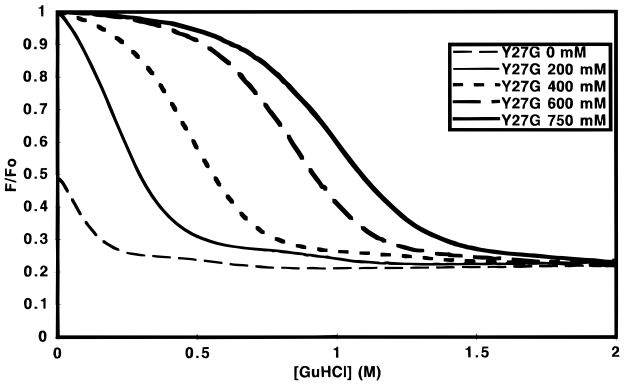


FIGURE 4: Y27G denatured in various concentrations of ammonium sulfate, given in millimolar. All curves are normalized to their initial fluorescence level except for that of Y27G with no ammonium sulfate, which was scaled in proportion to the post-transition baselines of the other curves.

The two smaller hydrophobic substitutions at position 27, alanine and valine, produce mutants with similar stabilities despite having substitutions of different shapes and sizes. The overall stability of Y27V (2.0 kcal/mol) is comparable to the 2.2 kcal/mol stability of the alanine mutant, even

though valine has a side chain that is approximately 3 times larger in volume than alanine (Table 2).

The tryptophan mutant at this position is also a stable mutant with a $\Delta G_{H_2O\ unfolding}$ value of 4.4 kcal/mol. Since this is the only substitution of a residue larger than the wild type, this data can be used as an indication of how well the local structure can accommodate expansion. The His 27 mutant is relatively stable (3.4 kcal/mol); however, its charge state through this unfolding transition is unknown.

The five polar substitutions (asparagine, glutamine, threonine, serine, and cysteine) display a set of characteristics different from those of the hydrophobic mutants. Serine and threonine both have a hydroxyl group, as does the wild type tyrosine, but are substantially less hydrophobic. The amide- and the sulfhydryl-containing substituents, on the other hand, provide examples of how polar groups in general interact with this pocket. As a group, the polar mutants span a lower range of stabilities than do the purely hydrophobic mutants, and have slopes equal to or greater than that of the wild type. The amides show a high range of variation; while the Y27N mutant is almost 20% unfolded at 0 M GuHCl, the other amide (Y27Q) which is only slightly bigger and longer is less than 5% unfolded in buffer. The hydroxyls show less variation, with a difference in the fraction of unfolded in buffer of only 2−3% (2.0 kcal/mol for Y27S vs 1.7 kcal/mol for Y27T). The cysteine is the most stable of the polar substitutions, but as it is the only one with a sulfhydryl group, conclusions that can be made about its interactions are limited. In addition, since a cysteine at this site is accessible to the surface, 1 mM DTT (dithiothreitol) was added to prevent oxidation of the sulfur and the formation of dimers through disulfide links with other Y27C nuclease molecules. In order to assess the effects of the DTT (if any), control denaturations of wild type protein were performed with the same 1 mM DTT concentration as in the Y27C data. Surprisingly, the DTT seemed to have a significant stabilizing effect on the wild type, although the nature of this effect is unclear. Since the addition of 1 mM DTT increased the wild type stability by 0.4 kcal/mol, we have assumed that the same is true of the Y27C data, and that the true $\Delta G_{H_2O\ unfolding}$ value of fully reduced Y27C protein lacking disulfide links in buffer alone is 2.1 kcal/mol.

The longer of the two positively charged substituents, arginine, gives the more stable mutant with a $\Delta G_{H_2O\ unfolding}$ value of 2.1 kcal/mol compared to the 1.1 kcal/mol value for Y27K, and is thus over 5 times more stable on the basis of relative fractions unfolded. Although its overall stability is greater, the Y27R mutant also exhibits a 29% increase in slope compared to that of the wild type protein.

The mutants with negatively charged side chains are quite destabilized. The shorter aspartate (Y27D) is the more unstable mutant with a free energy value that slightly favors the denatured state ($\Delta G_{H_2O\ unfolding} = -0.4$ kcal/mol). Since both have the same carboxylate group with the same type of delocalized charge, these provide excellent material for comparisons on the effect of length and hydrophobicity.

## DISCUSSION

One of the goals of protein study has been to predict the effects that a proposed mutation would have in a given tertiary context of a protein. Developing rules with regard to the relative contributions to stability of mutant amino acid properties, such as hydrophobicity, size, degree of branching, polarity, charges, etc., would serve as a good beginning. Using the free energy of unfolding values obtained for this set of mutants, we have qualitatively determined that the degree of hydrophobicity of the substituents is the most important criterion for overall stability at this site, with branching pattern and interaction with the Glu 10 residue acting as modifying factors that either enhance or decrease the overall stability. Implicit in the following discussion is the assumption that greater stability corresponds to greater similarity to the wild type structure, and thus better incorporation of the mutant amino acid into the tertiary nuclease structure; a correlation between mutant thermodynamic stability and overall enzymatic activity has previously been demonstrated for a set of mutants at this turn (Goodman, 1989).

Our results indicate that side chain composition (hydrophobicity or hydrophilicity) is the main factor in determining how acceptable (how stable) a particular mutant is at this site. Specifically, we have found that hydrophobic substitutions make for stable mutants, and that polar or charged residues are less stable at this site despite the potential for making a salt link or a hydrogen bond (as occurs in the wild type) with the Glu 10 residue. If the $\Delta G_{H_2O\ unfolding}$ values in Table 2 are divided into three groups, "stable" ($\Delta G_{H_2O\ unfolding} \geq 2.0$ kcal/mol or as defined by $\Delta G_{H_2O\ unfolding} = -RT \ln K_{eq}$ over 95% native in the buffer alone), "partially stable" ($\Delta G_{H_2O\ unfolding}$ of between 0.0 and 2.0 kcal/mol), and "unstable" ($\Delta G_{H_2O\ unfolding} < 0.0$ kcal/mol, or 50% or more denatured in buffer alone) mutants, the hydrophobic substitutions all produce stable mutants: valine, alanine, isoleucine, methionine, leucine, tryptophan, and phenylalanine. The exception to this general rule is proline which produces a highly unstable protein. This is expected since the conformational restriction of the proline immino ring ($\phi \sim 65°$) is inconsistent with this type of $\beta$-turn geometry.

The Y27F mutant gives the best system with which to test the interaction of the residue at position 27 with the Glu 10 residue. Table 2 shows that it is an $m_{GuHCl}+$ mutant, although its stability is equivalent to that of the wild type. Although mutational studies at a surface site of bacteriophage T4 lysozyme have indicated that side chain hydrogen bonds increase overall protein stability (Alber et al., 1987; Grutter et al., 1987), investigations into the effects of hydrogen bonding as a whole have generally given varied results with regard to the magnitude and influence of hydrogen bonding in protein stability [reviewed extensively in Dill (1990)]. The similarities in stability suggest that the Glu 10 residue in the Y27F mutant may stabilize the protein through interactions with another side chain. Stabilization of Glu 10 could result from a close proximity to the charged Lys 28 side chain, which can be modeled as a charged hydrogen bond. The involvement of Glu 10 in protein stabilization could also be present in the other mutants, but may only be enough to overcome the destabilization of the position 27 substitution in the Y27F mutant. The 10% (0.6 kcal mol$^{-1}$ M$^{-1}$ GuHCl) increase in the slope of Y27F shows that the hydrogen bond between Tyr27 and Glu 10 may be responsible for the maintenance of a compact denatured state with persistent $\beta$-sheet structure (Wynn et al., 1995).

In general, the hydrophobic substitutions result in proteins with the greatest stability ($\Delta G_{H_2O\ unfolding} \geq 2.0$ kcal/mol), indicating that hydrophobicity is the significant factor affecting stability at this site. There is a definite hierarchy in the stabilities of these mutants (Table 2), indicating that

a hydrophobic composition is the most important, but not the only, factor affecting protein stability. The significant range of stabilities for the isoleucine, methionine, and leucine mutants indicates that the structure of each substitution acts to modify the effect of hydrophobicity. The Y27I mutant contains the only $\beta$-branching substituent, in the stable hydrophobic class, making it the most unlike the wild type tyrosine of these four mutants. The leucine substituent is $\gamma$-branching, making it similar to the wild type tyrosine. The methionine can be thought of as an intermediate, having no branching at all. Since the leucine mutant is the most stable of the three and the isoleucine mutant is the least stable, similarity to the wild type residue shape implies greater stability, indicating that the side chain branching pattern is a modifier of stability, secondary to the influence of substituent hydrophobicity.

While a member of the stable class, the stability and *m* value of the His 27 mutant are difficult to interpret because of potential imidazole side chain titration within the unfolding transition. In its uncharged state, this $\gamma$-branched side chain should make interactions similar to those seen in the Tyr 27 wild type structure. The charged state could make similar packing interactions but could additionally make favorable electrostatic interactions with Glu 10.

The Y27V and the Y27A mutants provide an example of a case where the branching pattern may modify the primary effect of hydrophobicity. Valine is bulkier with a hydrophobicity more similar to that of the wild type tyrosine than that of alanine. Despite its greater hydrophobicity, Y27V is less stable than Y27A. This may be because the branching pattern of valine at the $\beta$-carbon is destabilizing enough to overcome the seeming advantage that its greater volume would have in occupying the Tyr 27 hydrophobic cavity. Therefore, the shape and accompanying steric complementarity to the local environment may in some cases modify the importance of hydrophobicity enough to make a less hydrophobic substituent at this site result in a more globally stable mutant protein.

The Y27W protein is very stable, perhaps due to tryptophan having no $\beta$- and two $\gamma$-branches just as the wild type residue at this position does. The substantial difference between tryptophan and tyrosine or phenylalanine is its significantly greater size. Modeling using the program MidasPlus (Ferrin et al., 1988) and the nuclease crystal structure (Hynes & Fox, 1991) shows that, in its most common sets of rotomer angles (Ponder & Richards, 1987), the tryptophan substitution either clashes significantly with the residues lining the cavity (particularly with Phe 34 and Val 74) or projects out of the cavity into the water, a very unlikely conformation for a large hydrophobic group. Thus, the tryptophan may insert itself into the cavity, but only with some disruption of the surrounding structure, or a deviation from ideal rotomer geometry.

The Y27W mutant contains two fluorescent probes, one at position 27 and the naturally occurring tryptophan at 140. If the two probes were reporting different local structural changes during the denaturation, we might expect to observe two separate transitions, each reflecting local denaturation. Despite the presence of two probes that are spatially far apart, only one transition is visible (Figure 2). This may be due to the greater exposure of Trp 27 to solvent resulting in modest changes in fluorescence upon denaturation.

In general, the mutants with polar substitutions at this position have lower stabilities. A possible reason for this is the fact that these polar residues are located at the protein surface with accessibility to the polar water environment, but because of the backbone geometry of the $\beta$-turn, they are directed into a pocket that is highly hydrophobic. A comparison of the two amide substitutions shows that the extra methylene group in glutamine gives almost 1 kcal/mol of greater stability than the asparagine mutant. For these polar residues, side chain length may be an important factor not only because a longer methylene chain confers greater hydrophobicity (hydrophobicity presumed from the above discussion to be the primary source of stability at this site) but also because increased length might facilitate interaction with Glu 10. This may be why the cysteine mutant is slightly more stable than the serine mutant. The volumes and geometries of the side chains are comparable, but the C−S and S−H bonds are both longer than the comparable bonds in serine. The sulfhydryl is also less polar than the hydroxyl; however, if length (and thus proximity to Glu 10) were the primary determinant of stability of these polar mutants, then the amide substitutions would be expected to be the most stable of the polar mutants. Since they are not the most stable, the larger polar group of the amides, by being forced into this pocket, would appear to detract from any energetic advantage asparagine and glutamine might have by being longer and better able to hydrogen bond.

Similarly, the advantage in being less polar is tempered by the disadvantage of being $\beta$-branched at this position, rather than being $\gamma$-branched as is the wild type tyrosine. Just as we have seen above with Y27V compared to Y27A, threonine, which with its extra methyl group is more hydrophobic than serine, would be expected to produce a more stable mutant than serine if hydrophobicity were the sole determinant of stability. The lower stability of Y27T compared with that of Y27S may be due to an unfavorable $\beta$-branch structure for the Thr 27. This difference in branching pattern may, at this site, destabilize threonine enough to overcome any advantage that threonine has in being more hydrophobic than serine, with the result that Y27T is less stable than Y27S.

The stability of the two mutants with positively charged substitutions (Y27K and Y27R) would be expected to be the result of their long length, making favorable hydrophobic interactions combined with the potential to form a stabilizing salt link with Glu 10. Modeling using the program MidasPlus (Ferrin et al., 1988) shows that such linking is sterically possible.

The greater stability of Y27E compared to Y27D is likely due to its longer aliphatic region. The only other mutant pair where such a direct comparison can be made is the two amide side chains, Y27N and Y27Q. The extra methylene in Y27Q adds 0.8 kcal/mol of stability, while in the carboxylate side chain pair above, only 0.4 kcal/mol of stability is added, indicating that, although the extra carbons are making both Y27Q and Y27E more stable than Y27N and Y27D, factors other than just the extra methylene group (and the added hydrophobicity that accompanies it) are affecting stability. The extra methylene in glutamine brings the amide group closer to Glu 10, which could enhance the protein's stability; similarly, the extra methylene brings Glu 27 closer to Glu 10, which would likely offset some of the added stability that the extra methylene alone (as a hydrophobic group filling up space in a hydrophobic pocket) would contribute to overall stability. Thus, if for instance we were interested in quantifying the stability contribution of an

"extra" methylene group at this site, we might ideally want to compare stabilities of two purely aliphatic substitutions to nullify the effect of the charge on Glu 10.

These results show that this site exhibits similarities with respect to the positions identified as core sites by Lim and Sauer (1989, 1991) as reviewed in Richards and Lim (1994). Lim and Sauer found that hydrophobicity was critical for maintaining protein activity and stability at closely packing sites of the $\lambda$-repressor core, which indicated that side chain composition, rather than shape and packing, was the key to acceptability at these positions. However, latter studies showed that close packing does play a secondary role in which it provides the detailed information for the activity and the stability of their $\lambda$-repressor mutants. This finding of two levels of structural information encoded in core residues (Lim & Sauer, 1989) seems to mirror what we have found in the stabilities of mutants at this particular site, with both shape/branching pattern and polarity/charge interaction with glutamate 10 modifying the main factor in stability, the hydrophobicity of the substituent. Lim and Sauer (1989, (1991) observed that, in order to maintain mutant activity comparable to that of the wild type $\lambda$-repressor, both volume and steric complementarity were constraints on what mutations were allowed in the core.

Overall, nearly all of the mutants show similar or larger slopes than the wild type, which corresponds to this site being part of the main hydrophobic core as reported in Shortle et al. (1990). It has been noted previously (Shortle & Meeker, 1986; Shortle et al., 1990; Green et al., 1992; Sondek & Shortle, 1992; Green & Shortle, 1993) that $m_{GuHCl}+$ mutants always destabilize the native structure in staphylococcal nuclease.

This project is part of a larger ongoing study examining the effects of single and multiple mutations in the $\beta$-turn from residues 27 to 30. The ultimate goal of this study is to obtain enough data on these turn mutants to fully characterize and quantify the influence of residues in this region on overall protein stability. We have derived general rules concerning the factors influencing the acceptability of amino acid substitutions at this site based on overall free energies of unfolding. Our data indicate that the degree of hydrophobicity of a substituent is the primary criterion for acceptability, with side chains of higher hydrophobicity in general producing more stable mutants. This primary stability-affecting factor has been found to be modified by two secondary factors which provide more detailed information about the protein's final stability. One of these is the shape or branching pattern of the substituent, with $\beta$-branching being disfavored and $\gamma$-branching being favored; evidence has been found that the favorability or unfavorability of shape can supersede the effect of hydrophobicity. The other secondary influence is the type and degree of interaction with the Glu 10 residue. This has again been shown to modify the influence of hydrophobicity, although no examples have been seen among these mutants where the influence of Glu 10 is clearly stronger than that of substituent hydrophobicity. The evident importance of hydrophobicity in the composition of substituents at this site for overall protein stability appears to support the concept of hydrophobic collapse as the initial step in formation of this $\beta$-barrel and therefore of the native structure (Jacobs & Fox, 1994).

## ACKNOWLEDGMENT

We thank Drs. Cort Anderson and Marc Jacobs for helpful comments and suggestions and Dr. Patrick Flemming for help with the computer modeling. The double-stranded DNA sequencing was carried out at the W. M. Keck Foundation's Biotechnology Resources Laboratory at Yale University.

## REFERENCES

Alber, T., Dao-pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., & Matthews, B. W. (1987) *Nature 330*, 41−46.
Barrick, D., & Baldwin, R. L. (1993) *Biochemistry 32*, 3790−3796.
Bolen, D. W., & Santoro, M. M. (1988) *Biochemistry 27*, 8069−8074.
Brünger, A. T. (1992) *X-Plor V 3.1*, Yale University Press, New Haven, CT.
Carson, M. (1991) *J. Appl. Crystallogr. 24*, 958−961.
Chothia, C. (1975) *J. Mol. Biol. 105*, 1−14.
Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
Ermácora, M. R., Ledman, D. W., & Fox, R. O. (1996) *Nat. Struct. Biol. 3*, 59−66.
Ferrin, T. E., Huang, C. C., Jarvis, L. E., & Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.
Fox, R. O., Evans, P. A., & Dobson, C. M. (1986) *Nature 320*, 192−194.
Goodman, M. A. (1989) Ph.D. Thesis, Stanford University, Stanford, CA.
Green, S. M., & Shortle, D. (1993) *Biochemistry 32*, 10131−10139.
Green, S. M., Meeker, A. K., & Shortle, D. (1992) *Biochemistry 31*, 5717−5728.
Griko, Y. V., Gittis, A., Lattman, E. E., & Privalov, P. L. (1994) *J. Mol. Biol. 243*, 93−99.
Grutter, M. G., Gray, T. M., Weaver, L. H., Alber, T., Wilson, K., & Matthews, B. W. (1987) *J. Mol. Biol. 197*, 315−329.
Hynes, T. R. (1989) Ph.D. Thesis, Stanford University, Stanford, CA.
Hynes, T. R., & Fox, R. O. (1991) *Proteins: Struct., Funct., Genet. 10*, 92−105.
Jacobs, M. D., & Fox, R. O. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 449−453.
Lim, W. A., & Sauer, R. T. (1989) *Nature 339*, 31−36.
Lim, W. A., & Sauer, R. T. (1991) *J. Mol. Biol. 219*, 359−376.
Loll, P. J., & Lattman, E. E. (1989) *Proteins: Struct., Funct., Genet. 5*, 183−201.
Nozaki, Y. (1972) *Methods Enzymol. 26*, 43−50.
Ponder, J. W., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775−791.
Richards, F. M., & Lim, W. A. (1994) *Q. Rev. Biophys. 26*, 423−498.
Shortle, D. (1986) *J. Cell. Biochem. 30*, 281−289.
Shortle, D., & Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81−89.
Shortle, D., Stites, W. E., & Meeker, A. K. (1990) *Biochemistry 29*, 8033−8041.
Sondek, J., & Shortle, D. (1992) *Proteins: Struct., Funct., Genet. 13*, 132−140.
Wang, Y., & Shortle, D. (1995) *Biochemistry 34*, 15895−15905.
Wang, Y., Alexandrescu, A. T., & Shortle, D. (1995) *Philos. Trans. R. Soc. London, Ser. B 348*, 27−34.
Wynn, R., & Zamyatnin, A. A. (1972) *Prog. Biophys. Mol. Biol. 24*, 107−123.

BI970876R