Articles

Contributions of the Large Hydrophobic Amino Acids to the Stability of Staphylococcal Nuclease†

David Shortle,* Wesley E. Stites, and Alan K. Meeker
Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205
Received March 5, 1990; Revised Manuscript Received May 7, 1990

ABSTRACT: To quantitate the contributions of the large hydrophobic residues in staphylococcal nuclease to the stability of its native state, single alanine and glycine substitutions were constructed by site-directed mutagenesis for each of the 11 leucine, 9 valine, 7 tyrosine, 5 isoleucine, 4 methionine, and 3 phenylalanine residues. In addition, each isoleucine was also mutated to valine. The resulting collection of 83 mutant nucleases was submitted to guanidine hydrochloride denaturation using intrinsic tryptophan fluorescence to monitor the equilibrium constant between the native and denatured states. From analysis of these data, each mutant protein's stability to reversible denaturation (ΔG_{H_2O}) and sensitivity to guanidine hydrochloride (m_{GuHCl} or $d(\Delta G)/d[GuHCl]$) were obtained. Four unexpected trends were observed. (1) A striking bipartite distribution was found for sites of mutations that altered m_{GuHCl}: mutations that increased this parameter only involved residues that contribute side chains to the major hydrophobic core centered around a five-strand β-barrel, whereas mutations that caused m_{GuHCl} to decrease clustered around a second, smaller and less well-defined hydrophobic core. (2) The average stability loss for mutants in each of the six residue classes was 2–3 times greater than that estimated on the basis of the free energy of transfer of the hydrophobic side chain from water to *n*-octanol. (3) The magnitude of the stability loss on substituting Ala or Gly for a particular type of amino acid varied extensively among the different sites of its occurrence in nuclease, indicating that the environment surrounding a specific residue determines how large a stability contribution its side chain will make. On the basis of statistical analyses, the parameter that provided the best estimate of this environmental effect on $\Delta\Delta G$ is the number of C_α carbons within a sphere of 10-Å radius. (4) A significant correlation was found between the absolute value of the change in m_{GuHCl} and the loss of stability. This correlation strongly supports the conclusion that amino acid substitutions can destabilize a protein indirectly via their effects on the structure and free energy of the denatured state.

In efforts to understand the physical–chemical processes that drive a polypeptide chain in aqueous solution into its highly ordered native state, biochemists have focused primarily on the structural patterns observed among proteins whose three-dimensional structures have been defined by X-ray crystallography. Among the most general patterns found in virtually all globular proteins are (1) tight packing of backbone and side-chain atoms to densities comparable to those of crystals of amino acids (Richards, 1977); (2) extensive formation of intramolecular hydrogen bonds (Baker & Hubbard, 1984); (3) a majority of hydrophobic side chains buried within the protein interior (Chothia, 1976); and (4) a large majority of charged side chains located at the surface in contact with water (Chothia, 1976).

On the basis of these patterns, the conclusion seems reasonable that all of the major classes of noncovalent forces—van der Waals interactions, hydrogen bonds, hydrophobic interactions, and electrostatic forces—play major roles in specifying the native structure of proteins. But when the question is phrased somewhat differently, "What net contribution do each of these forces make to the stability of the native state?", the issue becomes one of the magnitude by which the free energy of the native state is lowered relative to the free energy of the denatured state. Certainly, strong hydrogen bonding and van der Waals interactions occur between the polypeptide chain and solvent in the denatured state, and charged groups are likely to be as well solvated, if not better solvated, in the denatured state. Therefore, some uncertainty surrounds the contribution of these three forces to the net stability of protein structures.

However, since the publication of Kauzmann's (1959) classic paper entitled "Some factors in the interpretation of protein denaturation", a major role of hydrophobic interactions in specifying and stabilizing protein structure has been asserted with confidence. When a relatively structureless polypeptide chain folds in water, the hydrophobic amino acid residues are removed from an aqueous environment (high free energy) and transferred into the relatively nonpolar interior of the protein (low free energy). In other words, hydrophobic interactions can be viewed as acting primarily to raise the free energy of the denatured state relative to the free energy of the native state. A large literature has been published of estimates of the net contributions of the different hydrophobic amino acids to the stability of proteins [for reviews, see Rose et al. (1985) and Nakai et al. (1988)], and several authors have pointed out that most of these estimates are in relatively good agreement with one another (Rose et al., 1985; Guy, 1985).

Recent studies of the stability consequences of altering the amino acid sequence of small proteins strongly support the conclusion that hydrophobic interactions play a very important role. For example, substitutions of 12 different side chains for isoleucine 3 in T4 lysozyme, which is 80% solvent inac-

† This work was supported by a grant (GM34171) from the National Institutes of Health and an NIH postdoctoral fellowship to W.E.S.
* Author to whom correspondence should be addressed.

0006-2960/90/0429-8033$02.50/0

cessible, exhibited a good correlation between the reduction in protein stability relative to wild type, $\Delta\Delta G$, and the reduction in the transfer free energy from water to ethanol of the mutant side chain relative to an isoleucyl side chain (Matsumura et al., 1988). However, in a mutational study of two isoleucine residues in barnase, each mutated to both valine and alanine, Kellis et al. (1989) concluded that the stability contributions of hydrophobic side chains were twice as great as predicted by transfer free energies alone.

To systematically address the issues of how and by how much do hydrophobic amino acids stabilize the structure of proteins, we have mutated every large hydrophobic residue of staphylococcal nuclease, except the single tryptophan at position 140, to both alanine and glycine. Since these two types of mutations shorten a hydrophobic side chain instead of modifying its shape, we assume the results can be interpreted simply in terms of loss of side-chain atoms, without raising the issue of steric hindrance problems around the mutant side chain. By mutating all of the wild-type residues to alanine and glycine, we can compare the results to a common denominator, namely, the same mutant amino acid residue. Most importantly, a difference in the consequences of mutating the same residue type at two different sites in a protein must originate in the different environments surrounding the side chain at the two sites. Thus, this approach may permit identification of the most important parameters that determine how much a residue at a specific position will contribute to the stability of the native state.

In this report, we present the results of quantitating the change in stability ($\Delta\Delta G$) and the rate of change of ΔG with respect to GuHCl[1] concentration (m_{GuHCl}) for alanine and glycine substitutions at the positions of all of the leucine, valine, tyrosine, isoleucine, methionine, and phenylalanine residues, along with valine substitutions for the isoleucines. When the data from these mutant forms of nuclease were analyzed by statistical methods, several unexpected trends were found.

Experimental Procedures

Recombinant DNA. All mutants of staphylococcal nuclease were constructed in the wild-type gene, cloned into the phage vector M13mp9, according to the Kunkel modification (1985) of oligonucleotide-directed mutagenesis. Each oligonucleotide used for mutagenesis was 2- or 3-fold degenerate, specifying both alanine (GCN) and glycine (GGN) codons [plus a valine codon (GTN) for mutagenesis of isoleucine sites]. After transformation of mutagenized phage DNA into competent DH5α F′ cells, nuclease mutants were identified by sequencing phage DNA from a small number of well-separated plaques. Since occasional secondary mutations were found at other sites in the gene where a mutagenic oligonucleotide appeared to have annealed by partial homology and primed DNA synthesis, the entire gene (450 nucleotides) was sequenced for each recovered mutant. Each mutant gene was then transferred to the expression plasmid pL9 (Shortle & Meeker, 1989) by using T4 DNA polymerase to synthesize the second strand from the phage template, cleaving out the nuclease sequence from codon 3 (a unique *Spe*I site) to a unique *Sph*I site (Shortle, 1983) downstream of the gene, and then ligating the cleavage products to the corresponding fragment derived from pL9. To eliminate any possibility of accidentally recovering wild-type isolates, this fragment was purified from a pL9 derivative with a frame-shifting insertion of a *Xho*I linker at codon 101 in the nuclease gene.

[1] Abbreviation: GuHCl, guanidine hydrochloride.

Protein Purification. Mutant proteins were purified from *Escherichia coli* strain AR120 carrying the appropriate mutant gene on plasmid pL9 by a previously published protocol (Shortle & Meeker, 1989), except that the final chromatographic step was omitted. Protein concentration was determined by using an extinction coefficient at 280 nm of 0.93 (Fuchs et al., 1967), except for the 14 mutants that removed a single tyrosine; in these cases, a calculated extinction coefficient of 0.85 was used. Protein purity was monitored by SDS–polyacrylamide gel electrophoresis followed by staining with Coomassie Brilliant Blue and, for all proteins, was estimated to be between 90 and 98%.

Guanidine Hydrochloride Denaturation and Data Analysis. To quantitate the stability changes resulting from each amino acid substitution, intrinsic fluorescence of the single tryptophan residue at position 140 was monitored as a function of guanidine hydrochloride concentration at 20 °C, pH 7.0, by using conditions that have been described previously (Shortle & Meeker, 1986). The apparent equilibrium constant K_{app} for the reaction *native state* $\rightleftharpoons$ *denatured state* was calculated at a series of GuHCl concentrations, varied in 0.05 M increments, by using the equation

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

where I is the fluorescence intensity of the sample, I_n is the extrapolated value of the native state fluorescence, and I_d is the extrapolated value for the denatured state. As discussed previously (Shortle & Meeker, 1986), I_d can be readily estimated, since the fluorescence of the denatured state of wild type and all mutant forms increases very slightly with GuHCl concentration in a manner essentially identical with that for *N*-acetyltryptophanamide. For wild-type and stable mutant proteins, I_n does not appear to change with increasing concentrations of GuHCl. However, for unstable proteins where a significant fraction of molecules may be denatured in the absence of denaturant, I_n cannot be measured directly. Therefore, for mutants with a ΔG_{H_2O} less than +1.4 kcal/mol, I_n was estimated by three different methods. (1) Fluorescence intensity was measured as a function of ammonium sulfate concentration, until it reached a plateau. Since this salt is a strong renaturant and a 75% saturated solution alters the fluorescence of wild-type nuclease by less than 0.5%, the plateau value is assumed to be the fluorescence from a pure population of native molecules, and the gain in relative fluorescence is assumed to be proportional to the fraction of molecules that have been renatured (Shortle et al., 1989; Loll et al., 1988). (2) For wild-type nuclease and virtually all stable mutants, I_n is equal to 8.3 times the value of I_d. (3) I_n was adjusted to a value that gives the best linear fit of ΔG_{app} versus [GuHCl]. For 22 of the 33 unstable mutants, the three values of I_n agreed to within ±4%. For the remaining 11, which exhibited discrepancies no larger than 7%, the value of I_n obtained by method 1 was used for calculating K_{app}.

By use of the equation

$$\Delta G_{app} = -RT \ln K_{app}$$

the free energy change on denaturation ΔG_{app} was calculated for all values of K_{app} between 0.1 and 10. To obtain ΔG_{H_2O} and m_{GuHCl}, a straight line was fit to ΔG_{app} versus [GuHCl] by using the linear least-squares routine of Lotus 1-2-3. Typically, an average of 6–8 data points were fit, giving a value of r^2 between 0.9990 and 0.9999.

Statistical Analysis. The statistical analysis program SPSS/PC+ Version 3.0 was used to search for correlations between $\Delta\Delta G$ or m_{GuHCl} for alanine and glycine substitutions at a specific residue position and a number of physical pa-

Table I: Stability Parameters of Mutant Nucleases

mutant	C_m[a]	m_{GuHCl}[b]	$\Delta\Delta G$[c]	mutant	C_m[a]	m_{GuHCl}[b]	$\Delta\Delta G$[c]	mutant	C_m[a]	m_{GuHCl}[b]	$\Delta\Delta G$[c]
L7A	0.64	0.89	−1.6	V66A	0.44	1.10	−2.2	M26A	0.55	1.07	−1.5
L7G	0.65	0.89	−1.5	V66G	0.13	1.17	−4.4	M26G	0.43	1.11	−2.2
L14A	0.44	1.05	−2.3	V74A	0.31	1.13	−3.1	M32A	0.52	1.06	−1.7
L14G	0.22	1.18	−3.7	V74G		1.07	−6.6	M32G	0.41	1.09	−2.4
L25A	0.36	1.12	−2.7	V99A	0.29	1.13	−3.2	M65A	0.49	1.05	−2.0
L25G	0.12	1.18	−4.5	V99G		1.18	−5.0	M65G	0.11	1.18	−4.6
L36A	0.26	1.13	−3.5	V104A	0.40	0.93	−2.9	M98A	0.18	0.75	−4.6
L36G	0.02	1.23	−5.3	V104G		0.82	−6.5	M98G	0.16	0.88	−4.5
L37A	0.63	0.89	−1.7	V111A	0.19	0.64	−4.7	Y27A	0.18	1.10	−4.2
L37G	0.26	0.94	−3.8	V111G	0.11	0.75	−4.9	Y27G		1.26	−5.8
L38A	0.60	0.94	−1.7	V114A	0.84	0.96	0	Y54A	0.54	0.89	−2.2
L38G	0.80	0.90	−0.6	V114G	0.82	0.94	−0.2	Y54G	0.57	0.93	−1.9
L89A	0.42	1.00	−2.6	I15V	0.66	1.05	−0.8	Y85A	0.74	1.01	−0.4
L89G	0.34	0.96	−3.2	I15A	0.39	1.05	−2.7	Y85G	0.64	1.03	−1.0
L103A	0.13	0.96	−4.6	I15G	0.28	1.12	−3.3	Y91A	0.02	1.01	−5.3
L103G		0.88	−6.6	I18V	0.68	0.94	−1.1	Y91G		1.10	−6.7
L108A		0.77	−5.8	I18A	0.39	1.12	−2.5	Y93A		1.05	−6.5
L108G			−7.2	I18G	0.41	1.05	−2.5	Y93G			−7.5
L125A	0.12	0.66	−4.9	I72V	0.54	1.02	−1.8	Y113A	0.81	0.99	0
L125G			−7.0	I72A	0.05	1.29	−5.1	Y113G	0.78	0.97	−0.3
L137A	0.50	0.95	−2.3	I72G		1.07	−6.5	Y115A	0.77	0.99	−0.3
L137G	0.18	0.74	−4.6	I92V	0.73	1.01	−0.5	Y115G	0.74	0.96	−0.7
V23A	0.33	1.18	−2.9	I92A	0.18	1.18	−4.0	F34A	0.23	1.12	−3.7
V23G		1.34	−5.6	I92G		1.13	−6.6	F34G		1.30	−6.2
V39A	0.49	0.99	−2.2	I139V	0.62	0.94	−1.5	F61A	0.47	0.99	−2.3
V39G	0.14	0.79	−4.7	I139A	0.35	0.85	−3.5	F61G	0.10	0.94	−4.8
V51A	0.82	0.93	−0.3	I139G	0.21	0.73	−4.4	F76A	0.20	1.11	−4.0
V51G	0.80	0.94	−0.4					F76G	0.11	1.05	−4.7

[a] Midpoint concentration of GuHCl in M. [b] Units are relative to the wild-type value of 6.85 kcal/mol M. Estimated average error is less than ±0.02. [c] ΔG_{H_2O} (mutant) − 5.48 (wt) in kcal/mol. Estimated average error is less than ±0.1.

rameters that describe that position in the X-ray crystal structure of wild-type nuclease, both with (Loll & Lattman, 1989) and without Ca^{2+} and thymidine 3′,5′-bisphosphate bound (kindly provided by Dr. Robert O. Fox).

Results

Data and Estimate of Errors. In Figure 1, the positions of the 39 large hydrophobic amino acids where mutants have been constructed are shown in the context of the secondary structural features of wild-type staphylococcal nuclease. Using oligonucleotide-directed mutagenesis, both alanine and glycine were substituted for each of the 11 leucines, 9 valines, 7 tyrosines, 5 isoleucines, 4 methionines, and 3 phenylalanines. (The single tryptophan residue at position 140 was the only large hydrophobic residue that was not altered; it served as the fluorescence probe for monitoring the reversible denaturation reaction.) With each of the isoleucines also mutated to valine, a total of 83 different mutant nucleases were generated and purified.

The results of single guanidine hydrochloride denaturation analyses on each member of this collection of mutant proteins are listed in Table I. (The methods used for calculating ΔG_{H_2O}, the free energy change for reversible denaturation in the absence of denaturant, and m_{GuHCl}, the rate of change of ΔG with respect to changes in the GuHCl concentration, are described under Experimental Procedures.) For 15 mutant proteins, no value of C_m, the concentration of GuHCl at which the protein is 50% denatured, is given because ΔG_{H_2O} is negative. In other words, the protein is already beyond the midpoint of the denaturation transition in the absence of denaturant. For the three most unstable proteins, m_{GuHCl} could not be determined, since fewer than three values of K_{app} at different GuHCl concentrations could be accurately measured. By use of ammonium sulfate to renature the more unstable mutant proteins, values of ΔG_{H_2O} were obtained for all 83 proteins, including several that were more than 90% denatured in the absence of denaturant.

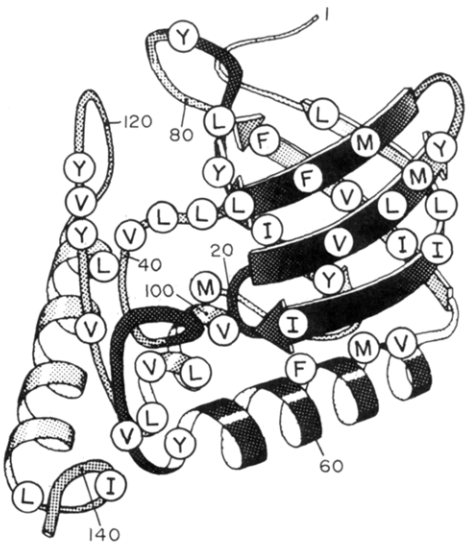

FIGURE 1: Ribbon diagram showing the positions of the large hydrophobic residues altered in mutant forms of staphylococcal nuclease. The approximate position of the C_α carbon is labeled by using the one-letter code. Numbers designate every 20th residue. This figure was drawn and copyrighted by Jane Richardson (1981) and is used with her permission.

To estimate the precision of the data obtained from this set of single denaturation experiments, wild-type nuclease was analyzed six separate times during the course of characterizing the mutant proteins. In addition, after completing one experiment on all of the mutant proteins, 16 mutants were picked at random and repeated a second time. For wild-type nuclease, the standard deviation from the mean value of C_m was ±0.5% and the standard deviations for both m_{GuHCl} and ΔG_{H_2O} were ±1.8%, in good agreement with previous estimates of the error (Shortle & Meeker, 1986). On repeating the 16 mutant nucleases, the average variation in ΔG_{H_2O} was ±0.04 kcal/mol, and the average variation of m_{GuHCl} was ±1.3%. Therefore,

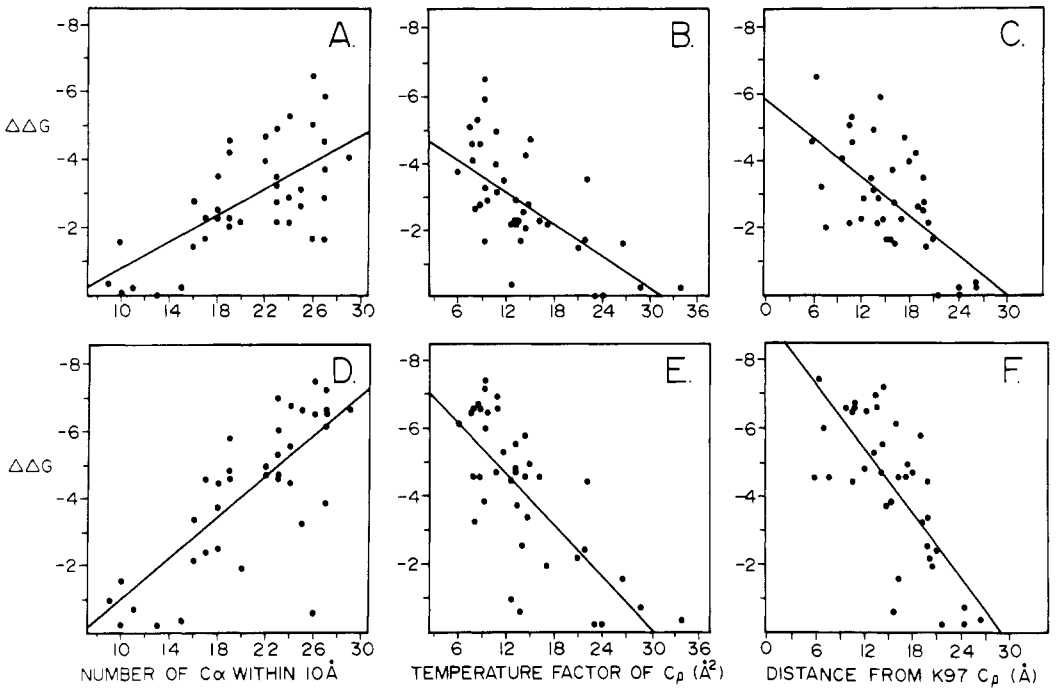

FIGURE 2: Scatterplots of $\Delta\Delta G$ for substitutions at a unique residue position versus three parameters that describe the environment surrounding that position in the wild-type native structure. Alanine substitutions are shown in panels A–C; glycine substitutions, in panels D–F. $\Delta\Delta G$ is expressed in kilocalories/mole. (A and D) The number of C_α carbons within a sphere of 10 Å centered on the C_α of the residue. (B and E) The crystallographic temperature factor of the C_β of the residue. (C and F) The distance from the C_β of lysine 97 to the C_β of the residue.

it seems reasonable to conclude that the average experimental error in Table I is less than ±0.1 kcal/mol for $\Delta\Delta G$ and less than 0.02 in relative m_{GuHCl}.

As mentioned under Experimental Procedures, for 11 mutant unstable nucleases for reasons that are not understood, the three methods for estimating the base-line value of fluorescence for the native state, I_n, disagreed by more than 4%. The value obtained by ammonium sulfate renaturation was assumed to be the most reliable and was used here to analyze all data. If instead I_n is determined by constraining m_{GuHCl} to a constant that best fits the experimental data and extrapolating to zero GuHCl concentration, the calculated ΔG_{H_2O} values for these 11 mutants differ by an average of 0.18 kcal/mol, with 0.24 kcal/mol being the largest discrepancy. Differences in m_{GuHCl} were somewhat larger—an average difference of 8.5%, with a maximal discrepancy of 11%. The satistical analyses reported below were carried out with both sets of data, and the effects of these uncertainties on the reported correlation coefficients were very small.

$\Delta\Delta G$ and Statistical Correlations with Parameters of Residue Environment. Within a class of one residue type, the loss in stability ($\Delta\Delta G = \Delta G_{mutant} - \Delta G_{WT}$) produced by alanine or glycine substitutions varied over a large range (Table I). In the case of leucine mutating to alanine, $\Delta\Delta G$ varied from −1.6 kcal/mol (L7A) to −5.8 kcal/mol (L108A), whereas for valine mutating to glycine, the range was from −0.2 (V114G) to −6.6 kcal/mol (V74G). In fact, for all mutant classes except those involving phenylalanine, there was more than a 2-fold difference between the smallest and largest change in stability. As mentioned in the introduction, these variations must be a consequence of the variations in the environment surrounding the different residue positions. For instance, a leucine residue would be expected to make a greater contribution to protein stability if it were fully buried than if it were fully exposed to solvent.

To gain more insight into the origins of the large variations in $\Delta\Delta G$ within each class of substitution mutants, statistical correlations were sought between the value of $\Delta\Delta G$ and the parameters listed in Table II. These parameters were obtained directly from, or calculated from, an X-ray diffraction model of wild-type staphyococcal nuclease and in some way describe the environment, or the effects of the environment, surrounding each wild-type residue in the native structure. The alanine and glycine substitution mutants were each analyzed in three separate groups: all 39 positions together, only the 11 leucine positions, and only the 9 valine positions. For each grouping, the Pearson correlation coefficient r was calculated by assuming a linear relationship between the dependent variable $\Delta\Delta G$ and each independent parameter. In addition, the two-tailed significance level p was calculated; this quantity is a measure of the probability that the calculated correlation coefficient between two independent, normally distributed variables could occur by chance. For most analyses, scatter plots were also plotted to search for nonlinear relationships between variables.

For the 39 mutant sites, the four parameters that gave the best linear correlations with $\Delta\Delta G$ for an alanine substitution at position N were the following: (1) the number of α-carbons within 10 Å of the α-carbon of residue N ($r = 0.6768$); (2) the crystallographic thermal B factor of the β-carbon of residue N ($r = -0.6759$); (3) the distance from the β-carbon of N to the β-carbon of residue 97 ($r = -0.6577$); and (4) the fraction of residue N's side-chain surface area that is exposed to solvent ($r = -0.6268$). The values of p for all of these correlations were less than 0.001, and scatter plots displaying $\Delta\Delta G$ as a function of each of the first three parameters are shown in Figure 2. All four of these parameters were significantly correlated with each of the other three. No noteworthy differences were found in any of these correlations on using either the X-ray coordinates or thermal B factors from wild-type nuclease complexed with the tight-binding ligands Ca^{2+} and thymidine 3′,5′-bisphosphate.

For the 11 leucine to alanine substitutions, correlations with these four parameters were not as strong: number of C_α, $r = 0.4390$, $p = 0.177$; C_β thermal factor, $r = -0.5199$, $p = 0.123$; distance from C_β at 97, $r = -0.5716$, $p = 0.066$; and

Table II: Parameters Used To Characterize the Environment of an Amino Acid Residue

residue no.
ϕ, ψ, χ, and ω angles
solvent exposure
fractional exposure of side chain
absolute exposure, backbone and side chain
temperature factors, backbone and side chain
hydrophobicity/hydrophilicity profiles in windows of 3, 5, 7, 9, and 11 residues
Kyte–Doolittle (1982)
Hopp–Woods (1981)
no. of residues within R angstroms, with R varied from 3 to 15
C_α relative to C_α of test residue
C_β relative to C_β of test residue
as above, but classified by hydrophobicity
as above, but classified by conformational flexibility of side chain (Ponders & Richards, 1987)
distance from each of the 148 other residues
various functions of number of neighbors and their distance to the nth power, $n = -1, -2, -1/2$

fractional exposure, $r = -0.2696$, $p = 0.423$. However, for the 9 valine to alanine substitutions, correlations were similar to those for the entire group of 39 sites: number of C_α, $r = 0.7769$, $p = 0.014$; C_β thermal factors, $r = -0.7483$, $p = 0.020$; distance from C_β at 97, $r = -0.6267$, $p = 0.071$; and fractional exposure, $r = -0.5496$, $p = 0.125$.

For the 39 glycine substitutions, these same four parameters also displayed the strongest correlations with $\Delta\Delta G$, with consistently higher correlation coefficients of 0.7571, −0.7458, −0.7471, and −0.6568, respectively. Again, as with the alanine substitutions, the correlation coefficients were lower for substitutions at the leucine sites and higher at the value sites.

With regard to correlations involving other parameters based on the number of residues within a sphere of radius R angstroms centered on the residue of interest, several interesting patterns were noted. The correlation coefficient for $\Delta\Delta G$ relative to the number of β-carbons was almost as large as for the number of α-carbons. As R was increased from 5 Å, both correlations improved rapidly, peaked at 10 Å, and slowly declined at greater distances. If only the number of hydrophobic residues were counted inside the sphere, the correlation coefficient was always significantly lower. For example, on using the number of α-carbons of hydrophobic residues within a 10-Å sphere, r was only 0.5035 for the 39 alanine substitutions and 0.6265 for the 39 glycine substitutions. Likewise, if the number of backbone nitrogen and oxygen atoms within a 10-Å sphere centered on the C_α carbon were counted, the correlation coefficient was reduced to $r = 0.5695$ for the alanines but was essentially the same at 0.7608 for the glycines.

m_{GuHCl} and Statistical Correlations with Residue Environment Parameters. As noted previously (Shortle & Meeker, 1986; Shortle et al., 1989; Sondek & Shortle, 1990), the value of m_{GuHCl} for mutant forms of staphylococcal nuclease can vary considerably from the wild-type value of 6.85 kcal/mol M. Among the mutant nucleases analyzed in this report, m_{GuHCl} ranges between 0.66 (L125A) and 1.34 (V23G), when normalized to wild type. Unfortunately, the chemical interactions by which GuHCl and other denaturants destabilize protein structure have not been established, so that the cause of these variations cannot be stated with confidence. To gain further insights into the origins and significance of this phenomenon, statistical correlations were sought between the list of parameters describing the environment around each residue (Table II) and both m_{GuHCl} and $|\Delta m_{GuHCl}|$, the absolute value by which the mutant m_{GuHCl} differs from that of wild type.

Surprisingly, no highly significant correlations were found with any of the parameters that correlated with $\Delta\Delta G$. The parameters that did show a significant correlation for the complete group of 39 sites—residue number ($r = -0.4301$, $p = 0.006$ for Ala substitutions, $r = -0.5366$, $p = 0.001$ for Gly substitutions), distance from residue 24 ($r = -0.6683$ and −0.8178, respectively), and distance from residue 133 ($r = 0.6116$ and 0.7064, respectively)—are basically measures of the spacial position of a residue within the native structure. Thus, it would appear that the value of m_{GuHCl} is not determined by the local chemical environment of a residue but rather by some sort of higher order, more global structural information.

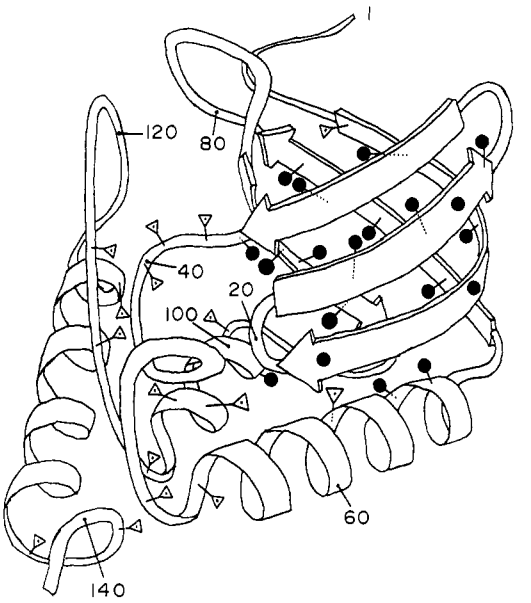

FIGURE 3: Ribbon diagram showing the approximate position of the wild-type side chain for two classes of staphylococcal nuclease mutants. (●) Residues where substitution with an alanine and/or glycine increases m_{GuHCl} to 1.05 or more of the wild-type value. (▲) Residues where substitution with an alanine and/or glycine decreases m_{GuHCl} to 0.95 or less of the wild-type value.

As shown in Figure 3, when the residue positions of the two classes of mutants, those that increase and those that decrease m_{GuHCl}, are plotted with respect to the three-dimensional structure of wild-type staphylococcal nuclease, a markedly asymmetric distribution is revealed. *All* of the hydrophobic side chains that, when shortened to alanine and/or glycine, increase m_{GuHCl} contribute to the major hydrophobic core centered on the five-strand β-barrel structure. All of the hydrophobic side chains that have the opposite effect of decreasing m_{GuHCl} are located outside of this hydrophobic core region, with the possible exceptions of L7, F61, and L103, which cannot be unambiguously classified with regard to their participation in this structure.

Of the nineteen positions for which either the alanine or the glycine substitution increased m_{GuHCl}, thirteen displayed a higher value for glycine relative to alanine, whereas five displayed a lower m_{GuHCl} for glycine, and one was indeterminant because the value for glycine could not be measured. For the sixteen positions where one or both substitutions reduced m_{GuHCl}, four exhibited a higher value for the glycine versus the alanine mutant, eight exhibited a lower value with glycine, two displayed no significant differences, and the remaining two were indeterminant. At only four positions out of the 39 analyzed did neither alanine nor glycine substitutions alter m_{GuHCl} significantly from the wild-type value.

To address the question as to whether the structural changes that are reflected in variations of m_{GuHCl} play a role in reducing the stability of a mutant protein, correlation coefficients be-

Table III: Average Values of Stability Losses

wt residue	no. of positions	$\Delta\Delta G$ (→Ala)[a]	$\Delta\Delta G$ (→Gly)[a]	$\Delta G_{transfer}$ (side chain)[a,b] $H_2O \rightarrow$ *n*-octanol
leucine	11	−3.1 ± 1.5	−4.4 ± 2.1	−2.3
valine	9	−2.4 ± 1.5	−4.3 ± 2.4	−1.6
isoleucine	5	−3.6 ± 1.0	−4.7 ± 1.8	−2.4
methionine	4	−2.5 ± 1.5	−3.4 ± 1.3	−1.7
tyrosine	7	−2.7 ± 2.7	−3.4 ± 3.1	−1.3
phenylalanine	3	−3.3 ± 0.9	−5.2 ± 0.8	−2.4

[a] Free energy changes in kcal/mol. [b] Fauchere & Pliska, 1983.

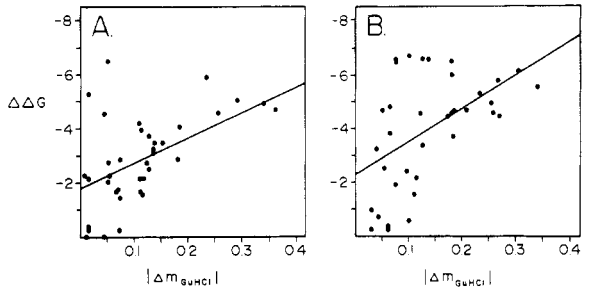

FIGURE 4: Scatterplots of $\Delta\Delta G$ for a mutant nuclease versus $|\Delta m_{GuHCl}|$, the absolute difference in m_{GuHCl} between mutant and wild type. Alanine substitutions are shown in panel A and glycine substitutions in panel B. $\Delta\Delta G$ is expressed in units of kilocalories/mole; $|\Delta m_{GuHCl}|$ is expressed relative to the wild-type value of m_{GuHCl}.

tween $\Delta\Delta G$ versus both m_{GuHCl} and $|\Delta m_{GuHCl}|$ were calculated. Although the correlations of $\Delta\Delta G$ for the alanine and glycine substitutions with m_{GuHCl} were not significant, the correlations were $|\Delta m_{GuHCl}|$ with highly significant: for the 39 alanine substitutions, $r = 0.5476$, $p < 0.001$; for the 36 glycine substitutions, $r = 0.5210$, $p = 0.001$. Scatterplots displaying these correlations are shown in Figure 4. Since the correlation is positive, large deviations of m_{GuHCl} from the wild-type value for both sets of mutants correlate with large reductions in stability.

Discussion

By changing each of the large hydrophobic amino acid residues in staphylococcal nuclease to both alanine and glycine, a data base has been accumulated that can be used to analyze the contribution of each of these residues to the stability of nuclease to reversible denaturation. When data from all 39 residue positions are grouped together, a sufficiently large data set is formed to permit the use of statistical methods to identify some of the mechanisms by which hydrophobic amino acids confer stability. The most obvious such mechanism involves the free energy lost on removing their side chains from water and transferring them into the relatively nonpolar interior of the native protein (Kauzmann, 1959). A large number of empirical measures of residue hydrophobicity have been developed over the past 25 years. Several authors have shown that most of these scales, even though based on widely different techniques for measuring this free energy change or for calculating it on the basis of how side chains are distributed relative to the surface of proteins of known structure, are in relatively good agreement in their estimates, especially for the aliphatic and aromatic side chains (Rose et al., 1985; Guy, 1985). The hydrophobicity scale of Fauchere and Pliska (1983), determined by measurements of the partition coefficients of *N*-acetyl amino acid amides from water to *n*-octanol, is fairly representative. Free energies of side-chain transfer of each of the six large hydrophobic amino acids plus alanine are listed in Table III.

When the average stability loss caused by glycine substitutions for each of the six types of residue is compared to the free energy of transfer of that residue's side chain, the value of $\Delta\Delta G$ is 2–3 times larger than expected if hydrophobicity were the sole source of stability. For example, the average $\Delta\Delta G$ for leucine to glycine substitutions is -4.3 kcal/mol versus the Fauchere–Pliska value of -2.3 kcal/mol, the predicted stability loss due to the lack of a leucyl side chain for transfer from water to the protein interior on folding. For valine to glycine mutants, the average $\Delta\Delta G$ is -4.4 kcal/mol, as opposed to -1.6 kcal/mol predicted from the transfer free energy lost in this type of mutant.

As with all measures of hydrophobicity based on side-chain partitioning, the values of Fauchere and Pliska reflect the free energy of removing a fully solvated side chain from dilute solution and inserting it into an organic solvent. Since the denaturation reaction being studied for the mutant nucleases involves transfer of a hydrophobic side chain from its environment in the denatured state into the protein interior, the question should be raised as to how good a model a dilute solution is for the denatured state. To the extent that residual structure persists in the denatured state, hydrophobic side chains may not be fully solvated, and the actual stability loss on folding of a mutant protein may not be as great as the Fauchere–Pliska values.

From Tanford's model of solvent denaturation, a value for the fractional change in solvent exposure for buried residues on protein denaturation can be experimentally determined (Tanford, 1970). For all proteins analyzed (Pace, 1975), an average buried residue undergoes a fractional increase in solvent exposure $\Delta\alpha$, where $\Delta\alpha$ lies between 0 (no increase in exposure) and 1 (transition from fully buried to fully exposed). Pace et al. (1990) have determined the value of $\Delta\alpha$ for staphylococcal nuclease by using both urea ($\Delta\alpha = 0.44$) and GuHCl ($\Delta\alpha = 0.43$) denaturation. In light of these relatively small changes in solvent exposure for the average buried residue, it may be that the hydrophobicity values of Fauchere–Pliska overestimate the free energy for side-chain burial by a factor of 2 or more. Consequently, the 2- to 3-fold discrepancy discussed above between $\Delta\Delta G$ for mutant nucleases and the transfer free energy may, in reality, be significantly larger.

A second mechanism by which hydrophobic side chains stabilize protein structure is through van der Waals interactions. The presence of an empty cavity in the mutant native state where the wild-type side chain normally resides would destabilize the native state relative to the denatured state because of lost interactions. Kellis et al. (1989) suggested this explanation to account for similar large stability losses in mutants of the small protein barnase. Since in some instances the protein structure may be able to relax around the cavity and partially fill it, one might expect that the component of $\Delta\Delta G$ arising from cavity formation would vary considerably from one residue position to another.

The free energy cost of cavity formation in the native state should depend on the size of the cavity. For cavities formed by shortening aliphatic side chains, both the volume and the

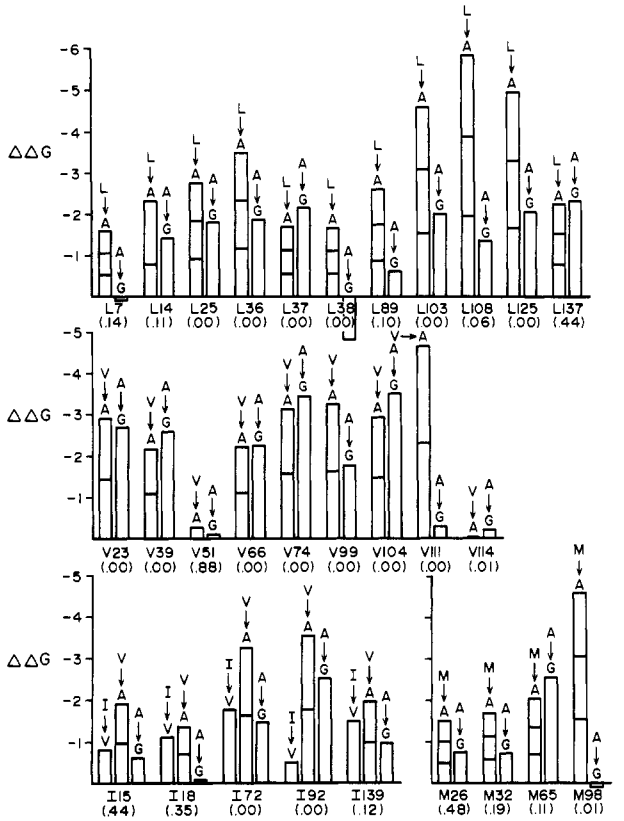

FIGURE 5: Schematic diagram showing the stability difference ($\Delta\Delta G$) between two forms of staphylococcal nuclease which differ at one amino acid residue by the number of side-chain methylene equivalents (methyl, methylene, or thioether groups). Each open box corresponds to one methylene equivalent. The height of each bar (or stack of boxes) represents the experimentally determined difference in free energy between the two forms, which are indicated above the bar. When the mutant side chains differ by two or three methylene equivalents, the bar is divided into two or three equal segments to represent the average free energy loss per methylene equivalent, in units of kilocalories/mole. The wild-type residue types and their positions are labeled below; the number in parentheses is the fraction of the side-chain surface area exposed to solvent.

surface area, and therefore the increase in free energy, should increase in an approximately linear way with the number of methyl carbons, methylene carbons, and thioether sulfurs removed. Since hydrophobicity, as measured by transfer free energies, also changes linearly with the number of methylene carbons (Tanford, 1980), the loss of stability due to cavity formation plus hydrophobicity should exhibit such a linear dependence. The value of $\Delta\Delta G$ per methylene equivalent may vary from site to site within the protein because of different local factors, such as hydrophobicity, packing density, or polarizability. These local factors, however, are unlikely to vary extensively across the volume of the protein interior occupied by a single side chain. Thus, it seems reasonable to assume that, for a given residue site, the free energy cost per deleted methylene euqivalent would be a constant. In fact, such a linear dependence of $\Delta\Delta G$ on methylenes removed was found by Kellis et al. (1989) for mutants at two isoleucine sites in barnase.

Figure 5 shows diagrammatically the change in $\Delta\Delta G$ for each methylene equivalent removed, or the average for two or three removed as a group, at each of the leucine, valine, isoleucine, and methionine positions in staphylococcal nuclease. In addition, the fractional exposure of the wild-type side chain is given in parentheses below that position. At some sites where the side chain is fully buried, such as I72, $\Delta\Delta G$ per methylene is relatively constant: −1.8 kcal/mol for the first (isoleucine to valine), -1.7 kcal/mol average for the next two (valine to alanine), and -1.4 kcal/mol for the fourth (alanine to glycine). A similar constant value of approximately -1.6 kcal/mol is found at a few other sites, such as L103, L125, and V99.

However, at other buried residue positions, no such rough linear relationship holds. Most striking of these is V104, where $\Delta\Delta G$ averages -1.4 kcal/mol per methylene carbon for the first two (valine to alanine) but is -3.6 kcal/mol for the third (alanine to glycine). At V74, the discrepancy is almost as large: an average of -1.6 kcal/mol for each of the first two but -3.5 kcal/mol for the third. Inspection of Figure 5 reveals that, for 15 of the 29 residue positions, the free energy loss varies by more than a factor of 2 among the three or four different methylene carbons in the side chain.

Thus, it appears that, for some positions in staphylococcal nuclease, the combination of hydrophobicity and van der Waals interactions in the native state may not be sufficient to completely explain the hydrophobic side chain's contribution to stability. For example, the contributions of each of the two γ-carbons of V104 (-1.3 kcal/mol) fall within the range of average values found at a number of sites, namely, -1.2 to -1.8 kcal/mol. If the assumption is made that the free energy cost per methylene equivalent for a given side chain should be a constant, these two γ methyl groups provide a "local calibration" for the sum of hydrophobic transfer free energy plus van der Waals interaction in the unique environment surrounding this valyl side chain. Yet the β-carbon contributes $2^1/_2$ times as much in free energy as each γ-carbon! The ΔG upon its removal, −3.6 kcal/mol for the alanine to glycine substitution, is higher than measurements and theoretical calculations of the enthalpy of melting plus vaporization per CH_2 moiety in straight-chain hydrocarbons, which fall in the range from 1.4 to 1.8 kcal/mol (Israelachvili, 1985). Since the entropy change on forming a cavity is unlikely to be negative, these enthalpy values represent a reasonable upper limit on estimates of the free energy cost of cavity formation. And since any relaxation of the native structure to partially fill the cavity or otherwise rearrange the packing of residues will only occur if it lowers the free energy of the native state, these enthalpy values, when added to the free energy of side transfer, define a reasonable upper limit of 2.0–2.5 kcal/mol for the increase in native-state free energy produced by removing a single methylene group.

Since the largest stability losses per methylene equivalent involve β-carbons (alanine to glycine substitutions), the much greater range of allowed ϕ,ψ angles for glycine, and the accompanying increase in chain entropy on denaturation, must be considered. Nemethy et al. (1966) have estimated the entropy gain of the polypeptide backbone as -2.4 cal/(deg·mol) on substituting glycine for alanine and assuming that in the denatured state all allowed ϕ,ψ angles for both residue types are equally populated. This estimate for the increase in chain entropy translates into a contribution to $\Delta\Delta G$ of -0.7 kcal/mol at 20 °C, the temperature under which all studies were done. Obviously, any increase in entropy of the native state caused by the substitution, or any residual local structure around this residue in the denatured state, would lead to a smaller effect on $\Delta\Delta G$.

In an attempt to identify additional sources of stabilizing interactions involving hydrophobic side chains, statistical methods were employed to search for correlations between $\Delta\Delta G$ and m_{GuHCl} and a variety of parameters that describe some feature of the residue environment (see Table II). For both the alanine and glycine substitutions, good correlations were found between $\Delta\Delta G$ and the fraction of a side chain's

surface area that is not exposed to solvent. Better correlations were found with parameters that reflect how buried, or how centrally located, the residue is within the protein interior. For example, the number of α-carbons within a sphere of 10 Å centered on the residue's β-carbon, the thermal B factor of the β-carbon, and the distance of a residue's β-carbon from the β-carbon of lysine 97 were correlated with $\Delta\Delta G$ to the same extent. [Alber et al. (1987) have previously reported a correlation between positions where temperature-sensitive mutations have been identified in T4 lysozyme and lower than average thermal B factors for both side chain and backbone atoms plus lower than average fractional accessible side-chain surface area.] However, all four of these parameters are significantly correlated to each of the other three. The only clear inference from these findings is that, on average, the more surrounded a position is by segments of the polypeptide chain the more critical the role played by the side chain at that position in determining the stability of the native state. Apparently, it is not essential that the surrounding segments involve hydrophobic residues, since the number of hydrophobic residues (either C_α or C_β) within a sphere of 10-Å radius correlated consistently less well with $\Delta\Delta G$ than did the number for all residues.

This statistical analysis of structural parameters that correlate with mutant values of $\Delta\Delta G$ has been confined to a small number of variables that contain relatively little detailed information about the native structure. Considering the high resolution to which staphylococcal nuclease has been solved (1.6 Å, R value of 0.161; Loll & Lattman, 1989), a more sophisticated analysis, one dealing with multiple variables and with more precise descriptions of the interactions of each side chain with neighboring groups, might reveal stronger empirical relationships with $\Delta\Delta G$. Such relationships might prove useful for estimating the stability effects of substitutions in other proteins. However, because of the relatively small size of the mutant collection, our plan is to conduct a more extensive analysis after alanine and glycine substitutions have been examined at all 149 residue positions.

Several noteworthy results were obtained on screening for statistical correlations between the same set of parameters (Table II) and the values of m_{GuHCl} for the alanine and glycine substitutions. No significant correlations were found between mutant values of m_{GuHCl} and any parmeter that describes some feature of the local environment around the mutant residue positions, such as hydrophobicity/hydrophilicity profiles, solvent exposure, temperature factors, number of neighboring residues of various types, etc. The only parameters that did correlate with m_{GuHCl} reflect a residue's position in the structure. When viewed from the orientation used in Figure 3, this positional effect can be seen as two distributions separated left to right. Alanine and/or glycine substitutions that increased m_{GuHCl} relative to the wild-type value are distributed throughout a major hydrophobic core centered on the five-strand β-barrel structure, whereas the substitutions that decrease m_{GuHCl} are concentrated on the opposite side of the molecule. When a statistical technique known as cluster analysis was applied to the distances between C atoms of the hydrophobic residues in the native structure, clear evidence was found for considering this second grouping of hydrophobic side chains a second, major hydrophobic core.

Previous studies of the physical–chemical basis of changes in values of m_{GuHCl} for mutants of staphylococcal nuclease (Shortle & Meeker, 1986, 1989; Shortle et al., 1989) have supported the conclusion that an increase in this quantity reflects an increase in the solvent-exposed surface area of the denatured state, whereas a reduced value appears to reflect a decrease in the solvent exposure of the denatured state.[2] If one assumes that this is in fact the case, then the observed bipartite distribution of the two mutant classes suggests that there are two different roles for hydrophobic residues in the denatured state. Residues that are buried in the β-barrel structure in the native state also cluster in the denatured state, and reducing the hydrophobicity of any one of them diminishes this clustering so that more surface is exposed in the denatured state than for the wild-type case. On the other hand, the hydrophobic residues in the second, "minor" hydrophobic core, by some mechanism that is not at all apparent, act together to effectively increase the solvent-exposed surface in the denatured state. Consequently, when the hydrophobicity of their side chains is reduced, the denatured-state structure presumably relaxes to a form that exposes less surface area to solvent or exposes surface area that is easier to solvate. Preliminary results from small-angle X-ray scattering studies of the denatured state of wild-type and several mutant nucleases reveal a bilobed structure that undergoes changes in its size and asymmetry with changes in amino acid sequence (J. Flanagan, M. Kataoka, D. Shortle, and D. Engelman, unpublished observations).

Although the detailed structural basis of the large changes in m_{GuHCl} are unlikely to be elucidated soon, an important question that has not been addressed previously is their energetic significance. Are these alleged modifications in the residual structure of the denatured state affecting the stability of the native state indirectly by lowering the free energy of the denatured state? As shown in Figure 4, a correlation can be demonstrated between the absolute value of the change in m_{GuHCl} away from the wild-type value and $\Delta\Delta G$. Therefore, the conclusion can be drawn that, for some mutants at least, the factor(s) responsible for change in m_{GuHCl} is(are) responsible for some of the stability loss manifest in the value of $\Delta\Delta G$. It is tempting to speculate that the wild-type denatured state has evolved to a unique ensemble of structures in which an optimal balance has been achieved between the competing objectives of maximal structure (to minimize the entropy increase on denaturation) and maximal exposure of hydrophobic residues (to minimize the stability of the denatured state). Small alterations in this optimal "statistical structure" either increase or decrease m_{GuHCl}, and the denatured state is lowered in free energy in proportion to the extent that m_{GuHCl} deviates from the wild-type value.

If this conjecture is not too far from the truth, then some of the unusually large changes in $\Delta\Delta G$ observed with certain mutants, such as V104G and V74G, may be a consequence of major structural changes in the denatured state, which are accompanied by large entropy increases that stabilize this state. It may be no accident that such pronounced effects are seen with substitutions that remove large hydrophobic amino acids, since interactions among hydrophobic side chains in the denatured state, especially when it is relatively compact, are probably the major structuring force acting on the polypeptide

[2] The magnitude of the changes in m_{GuHCl} argue against mechanisms based exclusively on structural changes in the native state. For example, consider tyrosines 85, 113, and 115, all three of which are on the surface. On substitution with glycine, no significant change in m_{GuHCl} is observed, even though the hydrophobic surface area of the native state has been reduced. However, on substituting the buried tyrosine at 27 with glycine, the m_{GuHCl} increases by 26%. It is extremely difficult to explain this increase in terms of a decrease in the solvent-accessible surface of the native state, since the surface area of the native state of the Y27G mutant would have to decrease by an amount that is several times greater than occurs for the Y85G, Y113G, and Y115G mutants.

chain (Dill, 1985). Indeed, the correlation noted above between $\Delta\Delta G$ for mutant nucleases and parameters that measure how centrally located the corresponding site is within the native structure may arise, in part, from a hierarchy of hydrophobic interactions in a highly compact denatured state that has many of the structural features of the native state. For it is only in the denatured state that hydrophobic interactions, in the conventional sense of this term, can occur. Perhaps the packing of hydrophobic residues in the native state should be considered more of a vestige of dominant interactions that occurred in the denatured state than an ongoing source of stability that acts exclusively of the native state.

Acknowledgments

We thank Robert Fox for allowing us to use his unpublished coordinates of the apo form of staphylococcal nuclease, Ed Lattman for assistance with the structure analysis, William Gabella for assistance with the statistical analysis, and Susan Green and John Sondek for helpful discussions and their careful reading of the manuscript.

Supplementary Material Available

Two tables giving fractional side-chain solvent-accessible surface area for wild-type nuclease, Kabsch–Sander secondary structural assignments, temperature factors for β-carbons, the number of neighboring C_α atoms within a 10-Å radius sphere, distances of C_β atoms to C_β at positions 24, 97, and 133, and Pearson correlation coefficients determined by using data sets in which values of 11 unstable mutants are calculated in different ways (4 pages). Ordering information is given on any current masthead page.

References

Alber, T., Dao-pin, S., Nye, J. A., Muchmore, D. C., Matthews, B. W. (1987) *Biochemistry 26*, 3754–3758.
Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97–141.
Chothia, C. (1976) *J. Mol. Biol. 105*, 1–14.
Dill, K. A. (1985) *Biochemistry 24*, 1501–1509.
Fauchere, J. L., & Pliska, V. (1983) *Eur. J. Med. Chem. 18*, 369–375.
Fuchs, S., Cuatrecasas, P., & Anfinsen, C. B. (1967) *J. Biol. Chem. 242*, 4768–4770.
Guy, H. R. (1985) *Biophys. J. 47*, 61–70.
Hopp, T. P., & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3824–3828.
Israelachvili, J. N. (1985) *Intermolecular and Surface Forces*, Academic Press, New York.
Kauzmann, W. (1959) *Adv. Protein. Chem. 14*, 1–63.
Kellis, J. T., Jr., Nyberg, K., Sali, D., & Fersht, A. R. (1988) *Nature 333*, 784–786.
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488–492.
Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.
Loll, P. J., & Lattman, E. E. (1989) *Proteins: Struct., Funct., Genet. 5*, 183–201.
Loll, P. J., Meeker, A. M., Shortle, D., Pease, M., & Lattman, E. E. (1988) *J. Biol. Chem. 263*, 18190–18192.
Matsumura, M., Becktel, W. J., & Matthews, B. W. (1988) *Nature 334*, 406–410.
Nakai, K., Kidera, A., & Kanehisa, M. (1988) *Protein Eng. 2*, 93–100.
Nemethy, G., Leach, S. J., & Scheraga, H. A. (1966) *J. Phys. Chem. 70*, 998–1004.
Pace, C. N. (1975) *Crit. Rev. Biochem. 3*, 1–43.
Pace, C. N., Laurents, D. V., & Thomson, J. A. (1990) *Biochemistry 29*, 2564–2572.
Ponder, J. W., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775–791.
Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.
Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., & Zehfus, M. (1985) *Science 229*, 834–838.
Schellman, J. A. (1978) *Biopolymers 17*, 1305–1322.
Shortle, D. (1983) *Gene 22*, 181–189.
Shortle, D., & Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81–89.
Shortle, D., & Meeker, A. K. (1989) *Biochemistry 28*, 936–944.
Shortle, D., Meeker, A. K., & Gerring, S. L. (1989) *Arch. Biochem. Biophys. 272*, 103–113.
Sondek, J. E., & Shortle, D. (1990) *Proteins: Struct., Funct., Genet. 7*, 299–305.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1–95.
Tanford, C. (1980) *The Hydrophobic Effect*, Wiley & Sons, New York.