Energetics of Side Chain Packing in Staphylococcal Nuclease Assessed by Systematic Double Mutant Cycles†

Junmei Chen‡ and Wesley E. Stites*

Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville Arkansas 72701-1201

Received June 19, 2001; Revised Manuscript Received September 11, 2001

ABSTRACT: All 44 possible double mutant permutations of isoleucine, leucine, and valine were constructed in 11 pairings of six sites in the core of staphylococcal nuclease. The stabilities of these mutants were determined by guanidine hydrochloride denaturation. Comparison of the stabilities of all double mutants with those expected from addition of the corresponding single mutants showed that the effects of the two single mutations are energetically independent of each other in 30 of the double mutants. However, a substantial minority, 14, of the double mutants have stability effects that are not additive. In these cases, it appears that direct van der Waals contacts between the two side chains are present. The requirement of direct van der Waals contact for the interdependence of mutational stability effects is somewhat surprising in light of results previously reported by others. In addition, it was found that double mutants that did not alter or lower the overall number of atoms in the core and that showed nonadditive behavior were more stable than expected from addition of the effects of the corresponding single mutants. A net increase in the number of atoms in the core usually, but not always, resulted in a mutant that was less stable than expected. In contrast to previous staphylococcal nuclease double mutants, energetically significant changes to the denatured state do not appear to be occurring in these packing mutants. These conclusions imply that attempts to engineer protein stability based on single mutant data will be generally successful if overall core size is preserved and if residues are not in van der Waals contact.

Hydrophobic interactions have been considered the major driving force for protein folding (*1*). However, recently, the van der Waals interactions afforded by optimal close packing of the protein have been proposed to be just as or possibly even more energetically important (*2*−*9*). Although perturbing the packing of the hydrophobic core has small effects on overall protein structure, it often results in dramatic energetic effects, although precise effects vary considerably from site to site. Understanding how protein packing affects protein stability will be critical for structure prediction and rational protein design.

The relationship between a protein's packing and its stability has been studied by mutagenesis. Packing mutants with single mutations have been constructed in various proteins (*2*, *10*−*20*). Although considerable progress has been made, the energetic effects caused by any single mutation in a protein core are context-dependent and remain difficult to predict. In part, this is because mutation of a single residue disrupts multiple interactions between that side chain and its environment. It is difficult to completely understand how a residue interacts with the surrounding environment without carrying out further experiments to alter that environment. Therefore, the simultaneous introduction of multiple mutations is necessary to explore this aspect of packing in proteins.

The stability effects of multiple packing mutations have been studied previously in surprisingly few cases and proteins. These include T4 lysozyme (*21*−*23*), bacteriophage f1 gene V protein (*24*−*26*), *Escherichia coli* RNase HI (*27*), tryptophan synthase (*28*), *E. coli* dihydrofolate reductase (*29*), ubiquitin (*30*), and staphylococcal nuclease (*31*). Virtually all of these examples are double mutants. Usually relatively few positions or possible permutations of side chains have been examined. This relative lack of data may be in part why attempts to predict the stability effects of multiple packing mutations have met with little success (*32*).

To fully understand the interactions among residues in a protein core, we felt a large database of stability effects from sets of systematically constructed single, double, and higher-order mutations would be very useful. In the preceding paper (*2*), we report the effects of single mutants which exchange isoleucine, leucine, and valine in staphylococcal nuclease.

In this work, 44 double mutants have been made at six selected positions in the hydrophobic core of staphylococcal nuclease. This is an exhaustive examination of every possible permutation of isoleucine, leucine, and valine at 11 different pairs of positions. The stabilities of these mutants have been determined by solvent denaturation, and the interactions between the mutated side chains are analyzed by double mutant cycles (*33*−*35*). The effects of triple and quadruple mutations at these same sites are reported in the paper that immediately follows.

EXPERIMENTAL PROCEDURES

Mutagenesis. Twelve double mutants at positions 23 and 25, 66 and 72, and 92 and 99 were constructed by one cycle

† Funding for this work was provided by NIH Grants GM52714 and 1P20RR15569-1.

* To whom correspondence should be addressed. Phone: (501) 575-7478. Fax: (501) 575-4049. E-mail: wstites@uark.edu.

‡ Current address: N1302 Alkek Research Building, One Baylor Plaza, Baylor College of Medicine, Houston, TX 77030.

10.1021/bi011268l CCC: $20.00
Published on Web 10/26/2001

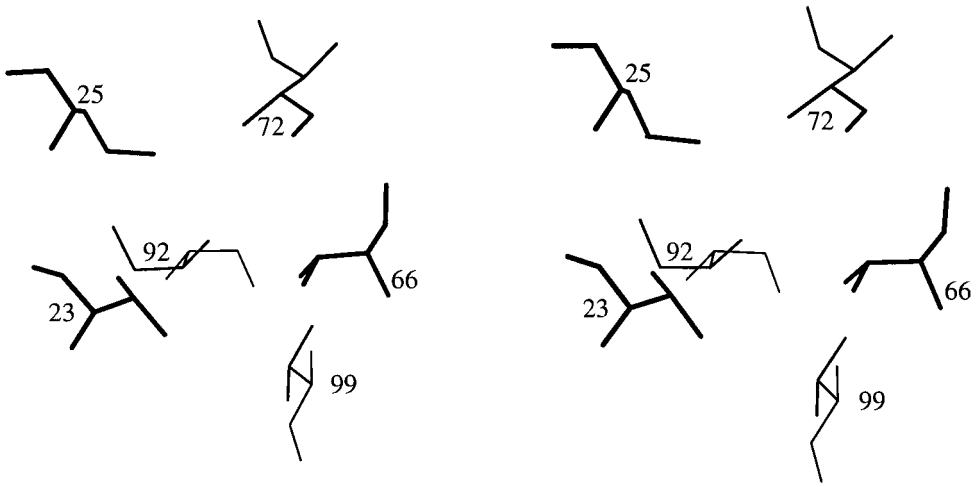

FIGURE 1: Stereoview of six core residues at positions 23, 25, 66, 72, 92, and 99 in the 1.6 Å structure of wild-type staphylococcal nuclease (PDB entry 1EY0).

of oligonucleotide-directed Kunkel mutagenesis (*36*, *37*) as previously described (*38*). Because the two residues at the pairs of positions 23 and 25, 66 and 72, and 92 and 99 are relatively close to each other in the primary sequence, the oligonucleotides were designed to create the desired mutations at these pairs of positions simultaneously. The 32 remaining double mutants in which the two mutations are far apart in the primary sequence were constructed by double cycles of Kunkel mutagenesis. The first cycle of Kunkel mutagenesis was performed by using a Kunkel template containing the wild-type gene and an oligonucleotide encoding a single mutation. The DNA containing the desired single mutations was passaged through the CJ236 cell line and served as a Kunkel template for the second round of mutagenesis, adding the other desired mutation.

Protein Expression and Purification. Protein expression and purification followed the procedures previously described (*38*, *39*).

Protein Stability Determination. The stabilities of each of the mutants in this study were characterized by guanidine hydrochloride denaturation at 20 °C using fluorescence as a probe of structure as previously described (*2*, *39*−*41*).

Calculation of Interaction Energies. The difference between the stability change in the double mutant and the summed stability effects of the component single mutants is the energy of the interaction between the two mutated residues. A general theory for examining the energetics of side chain interactions between multiple residues has been laid out in detail by Horovitz and Fersht (*35*). This interaction energy, $\Delta^n G_{int}$ in the nomenclature of Horovitz and Fersht, where n indicates the total number of side chains being mutated, is defined as follows:

$$\Delta^2 G_{int} = \Delta\Delta G_{double} - (\Delta\Delta G_{single1} + \Delta\Delta G_{single2})$$

In other double mutants in staphylococcal nuclease, this quantity has been called $\Delta\Delta\Delta G$ (*42*).

Error Estimation. The experimental error for determination of ΔG_{H_2O} based on the reproducibility of wild-type denaturation is ±0.1 kcal/mol. When interaction energies, $\Delta^2 G_{int}$, are calculated, the errors accumulate. The error in $\Delta^2 G_{int}$ is of course equal to the square root of the sum of the square of the error from each component. Since the general definition of each stability change is $\Delta\Delta G = \Delta G_{H_2O,mutant} - \Delta G_{H_2O,WT}$, each of the $\Delta\Delta G_{double}$, $\Delta\Delta G_{single1}$, and $\Delta\Delta G_{single2}$ values in the calculation contains a common constituent, $\Delta G_{H_2O,WT}$. Substituting into the equation for $\Delta^2 G_{int}$ above, we have

$$\Delta^2 G_{int} = \Delta G_{H_2O,double} - \Delta G_{H_2O,WT} - (\Delta G_{H_2O,single1} - \Delta G_{H_2O,WT} + \Delta G_{H_2O,single2} - \Delta G_{H_2O,WT})$$

After $\Delta G_{H_2O,WT}$ terms are canceled, this equation can be rearranged to Therefore, the error in $\Delta^2 G_{int}$ for double

$$\Delta^2 G_{int} = \Delta G_{H_2O,double} - \Delta G_{H_2O,single1} - \Delta G_{H_2O,single2} + \Delta G_{H_2O,WT}$$

mutants is given by the following equation:

$$\sigma_{\Delta^2 G_{int}} = \sqrt{(\sigma_{\Delta G_{H2O,double}})^2 + (\sigma_{\Delta G_{H2O,single1}})^2 + (\sigma_{\Delta G_{H2O,single2}})^2 + (\sigma_{\Delta G_{H2O,WT}})^2}$$

$$= \sqrt{4 \times (\pm 0.1)^2}$$

$$= \pm 0.2 \text{ kcal/mol}$$

Thus, if $\Delta^2 G_{int}$ falls in the range of 0 ± 0.2 kcal/mol, the stability of the double mutant is experimentally indistinguishable from that predicted by the effects of the single mutants. In other words, the effects are perfectly additive. While recognizing that the actual error is slightly lower, we have chosen to designate any double mutant with a value of $\Delta^2 G_{int}$ falling within the slightly less stringent range of 0 ± 0.3 kcal/mol as having additive stability effects. This standard is adopted for two reasons. First, it is the standard previously applied by other workers (*42*) to staphylococcal nuclease double mutants. Second, it facilitates comparison with triple and quadruple mutants in the following paper, which have slightly greater errors associated with them and require the slightly greater range.

Analysis of van der Waals Contacts. The analysis of van der Waals interactions was performed on the structure of

Table 1: Distance Matrix among the Six Selected Sites[a]

	V23	L25	V66	I72	I92	V99
V23	−	6.1	6.0	8.8	6.9	8.2
L25	6.1	−	8.7	6.9	7.8	11.5
V66	6.0	8.7	−	6.1	7.4	6.6
I72	8.8	6.9	6.1	−	6.7	8.9
I92	6.9	7.8	7.4	6.7	−	4.8
V99	8.2	11.5	6.6	8.9	4.8	−

[a] $C_\beta - C_\beta$ center to center distance (in angstroms).

wild-type staphylococcal nuclease [PDB entry 1EY0 (*43*)] using the Whatif (*44*) software package.

RESULTS

In this study, six residues at positions 23, 25, 66, 72, 92, and 99 in the major hydrophobic core of staphylococcal nuclease were selected and double mutants were made. These six residues are close to each other in the tertiary structure (Figure 1 and Table 1).

The relationship between the three pairs of residues (23 and 25, 66 and 72, and 92 and 99) is roughly triangular. Because the pair of residues 66 and 72 is close to the other two pairs of positions (23 and 25 and 92 and 99) and positions 23 and 25 are relatively far away from positions 92 and 99, we decided to limit the number of mutants by just constructing double mutants at two subsets of four positions: 23, 25, 66, and 72 or 66, 72, 92, and 99. Conservative side chain size and shape substitutions were made in which possible side chains were limited to just isoleucine, leucine, and valine. These positions are highly conserved as isoleucine, leucine, or valine among staphylococcal nuclease homologues (data not shown). A total of 44 double mutants are possible, and all were constructed.

Twelve of 44 double mutants are made at three pairs of positions: 23 and 25, 66 and 72, and 92 and 99. These pairs are very close to each other in both primary and tertiary structures. The remaining 32 double mutants are constructed at pairs of sites further separated in primary structure but near one another in the tertiary structure by combining one of the single mutants at position 23, 25, 92, or 99 with a single mutant at position 66 or 72.

Guanidine Hydrochloride Denaturation Data. A summary of guanidine hydrochloride (GuHCl) denaturation data of the 44 double mutants is shown in Table 2. Stabilities (ΔG_{H_2O}) of double mutants varied from 2.4 to 5.5 kcal/mol. Only two double mutants (V66L/I92V and V66L/V99L) have stabilities similar to that of the wild type, and both mutants contain the V66L mutation. The midpoint concentrations, C_m, of the double mutant denaturations ranged from 0.32 to 0.97 M. The slope values (m_{GuHCl}) of the double mutant denaturations expressed relative to the wild-type value of 6.53 kcal mol^{-1} M^{-1} vary from 0.74 to 1.09.

Double Mutants at Positions 23 and 66, 23 and 72, 25 and 66, and 25 and 72. As seen in Table 2, the stabilities of double mutants at the pairs of positions 23 and 66, 23 and 72, 25 and 66, and 25 and 72 seem decided by the identity of the side chain substitution at position 66 or 72 regardless of what substitution is at the second position. As an example, although V23I/V66I and V23L/V66I have different side chains at position 23, they have the same stability. It appears that the side chain substitutions at position 66 or 72 determine the energetic difference among the four double mutants at each pair of the above positions. Such behavior is not seen for any other pair of positions.

Additivity. The stability changes of double mutants relative to the wild type ($\Delta\Delta G_{double}$) are compared to the sum of stability changes of component single mutants ($\sum\Delta\Delta G_{single}$). If the difference between $\Delta\Delta G_{double}$ and $\sum\Delta\Delta G_{single}$ (i.e., $\Delta^2 G_{int}$) is 0 ± 0.3 kcal/mol (a range similar to the experimental error), the double mutant is said to have additive stability effects. Otherwise, the double mutant is said to have nonadditive stability effects.

Additive effects are generally found when the changes in stability (Figure 2, $r = 0.8572$), midpoint concentration (Figure 3, $r = 0.9274$), and slope values (Figure 4, $r = 0.9045$) of double mutants relative to that of the wild type are compared to the sum of the changes of their component single mutants. Of the 44 double mutants, 30 have additive stabilities and 14 nonadditive.

In Figure 2, the sum of stability changes of component single mutants ($\sum\Delta\Delta G_{single}$) is plotted against the stability changes of double mutants ($\Delta\Delta G_{double}$). The dotted line is the theoretical line with a slope of unity on which all points would lie if additivity were perfect. The solid line is a regression ($r = 0.8572$) through the data. Most of the data points are near the theoretical line, which indicates that additive behaviors are general among the double mutants. However, some data points plot far from the theoretical line, indicating that two mutated side chains in these double mutants interact strongly with each other.

The energy of the interaction between two mutated side chains, $\Delta^2 G_{int}$, is shown in Table 2 and is equal to the difference between the stability changes of double mutants relative to the wild type ($\Delta\Delta G_{double}$) and the sum of the stability changes of corresponding single mutants ($\sum\Delta\Delta G_{single}$). If $\Delta^2 G_{int}$ is equal to or less than ± 0.3 kcal/mol, the estimated error, the double mutant has additive stability effects and no experimentally significant interaction between the corresponding single mutants. If $\Delta^2 G_{int}$ falls outside the range of 0 ± 0.3 kcal/mol, the double mutant has nonadditive stability effects. When $\Delta^2 G_{int}$ of a double mutant with nonadditive stability effects is positive, the double mutants are more stable than expected from the sum of the stability changes from single mutants. Thus, a positive value of $\Delta^2 G_{int}$ implies that there are favorable interactions among the mutated side chains. If $\Delta^2 G_{int}$ is negative, it indicates that strains are introduced between the mutated side chains. This causes the double mutant to be less stable than the sum of the stability effects from constituent single mutants.

All double mutants at the pairs of positions 25 and 66, 23 and 72, 72 and 92, and 72 and 99 have additive stability effects. Some double mutants at the pairs of positions 23 and 25, 23 and 66, 66 and 92, 66 and 99, 25 and 72, and 92 and 99 have additive stability effects, while others do not. Nonadditive stability effects appear in all four double mutants only at positions 66 and 72. Further, V66L/I72V and V66I/I72V show the largest nonadditive stability effects ($\Delta^2 G_{int}$). These two mutants are 1.0 kcal/mol more stable than expected from the component single mutants.

Correlation between Slope and Energetic Changes. Previous work on double mutants of staphylococcal nuclease (*42*) found significant correlation between the slope change and

Table 2: Solvent Denaturation Data for Double Packing Mutants

	ΔG_{H_2O}[a]	C_m[b]	m_{GuHCl}[c]	$\Delta\Delta G_{double}$[d]	$\sum\Delta\Delta G_{single}$[e]	$\Delta^2 G_{int}$[f]	ΔCH_2[g]
V23I/L25I	3.7	0.60	0.93	−1.7	−1.9	0.2	1
V23I/L25V	4.0	0.63	0.95	−1.4	−2.0	0.6	0
V23L/L25I	3.9	0.60	0.97	−1.5	−1.8	0.3	1
V23L/L25V	3.6	0.53	1.01	−1.8	−1.9	0.1	0
V66L/I72L	4.5	0.75	0.89	−0.9	−0.5	−0.4	1
V66L/I72V	4.9	0.80	0.92	−0.5	−1.5	1.0	0
V66I/I72V	4.2	0.63	0.99	−1.2	−2.2	1.0	0
V66I/I72L	3.7	0.64	0.87	−1.7	−1.2	−0.5	1
I92L/V99I	4.5	0.67	1.01	−0.9	−0.8	−0.1	1
I92L/V99L	4.1	0.63	0.96	−1.3	−0.9	−0.4	1
I92V/V99I	4.7	0.73	0.96	−0.7	−0.6	−0.1	0
I92V/V99L	4.5	0.67	1.00	−0.9	−0.7	−0.2	0
V23I/V66I	3.9	0.75	0.78	−1.5	−1.2	−0.3	2
V23I/V66L	4.4	0.89	0.74	−1.0	−0.5	−0.5	2
V23L/V66I	3.9	0.67	0.86	−1.5	−1.1	−0.4	2
V23L/V66L	4.3	0.78	0.83	−1.1	−0.4	−0.7	2
L25I/V66I	2.4	0.38	0.92	−3.0	−2.7	−0.3	1
L25I/V66L	3.5	0.60	0.88	−1.9	−2.0	0.1	1
L25V/V66I	2.5	0.40	0.94	−2.9	−2.8	−0.1	0
L25V/V66L	3.6	0.60	0.89	−1.8	−2.1	0.3	0
V23I/I72L	4.9	0.80	0.91	−0.5	−0.4	−0.1	1
V23I/I72V	4.3	0.64	1.01	−1.1	−1.4	0.3	0
V23L/I72L	5.0	0.79	0.95	−0.4	−0.3	−0.1	1
V23L/I72V	4.3	0.62	1.04	−1.1	−1.3	0.2	0
L25I/I72L	3.7	0.53	1.06	−1.7	−1.9	0.2	0
L25I/I72V	2.4	0.32	1.09	−3.0	−2.9	−0.1	−1
L25V/I72L	3.8	0.56	1.03	−1.6	−2.0	0.4	−1
L25V/I72V	2.5	0.35	1.09	−2.9	−3.0	0.1	−2
V66I/I92L	3.5	0.59	0.88	−1.9	−1.6	−0.3	1
V66I/I92V	4.7	0.69	1.00	−0.7	−1.4	0.7	0
V66L/I92L	4.4	0.79	0.83	−1.0	−0.9	−0.1	1
V66L/I92V	5.5	0.89	0.92	0.1	−0.7	0.8	0
V66I/V99I	4.7	0.80	0.88	−0.7	−1.2	0.5	2
V66I/V99L	4.1	0.70	0.86	−1.3	−1.3	0.0	2
V66L/V99I	4.8	0.89	0.81	−0.6	−0.5	−0.1	2
V66L/V99L	5.4	0.97	0.84	0.0	−0.6	0.6	2
I72L/I92L	4.7	0.74	0.94	−0.7	−0.8	0.1	0
I72L/I92V	4.7	0.69	1.01	−0.7	−0.6	−0.1	−1
I72V/I92L	3.6	0.51	1.05	−1.8	−1.8	0.0	−1
I72V/I92V	3.7	0.54	1.03	−1.7	−1.6	−0.1	−2
I72L/V99I	4.9	0.76	0.97	−0.5	−0.4	−0.1	1
I72L/V99L	4.6	0.71	0.99	−0.8	−0.5	−0.3	1
I72V/V99I	4.2	0.61	1.05	−1.2	−1.4	0.2	0
I72V/V99L	3.7	0.53	1.03	−1.7	−1.5	−0.2	0
wild type	5.4	0.82	1.00	−	−	−	−

[a] Free energy difference between the native and denatured states in the absence of denaturant in units of kilocalories per mole. The error is estimated to be ±0.1 kcal/mol. [b] Midpoint concentration (concentration of guanidine hydrochloride at which half of the protein is denatured) in units of molar. The error is estimated to be ±0.01 M. [c] Slope value (change in free energy with respect to the change in guanidine hydrochloride concentration) expressed relative to the wild-type value of 6.53 kcal mol^{-1} M^{-1}. The error is estimated to be ±0.02. [d] Difference in free energy between the free energy of the protein with double substitutions and the free energy of the wild-type protein. $\Delta\Delta G = \Delta G_{H_2O}$ (double mutant) − 5.4 (WT). The error is estimated to be ±0.2 kcal/mol. [e] The sum of the $\Delta\Delta G_{single}$ values of corresponding single substitutions. [f] $\Delta^2 G_{int} = \Delta\Delta G_{double} - \sum\Delta\Delta G_{single}$. The error is estimated to be ±0.2 kcal/mol. [g] Difference in the number of methylene groups between the mutant and wild type.

stability change. The quantity $\Delta\Delta m_{double}$, equal to the slope value change of double mutant relative to that of the wild type minus the sum of the slope value changes of constituent single mutants, was found to be correlated with $\Delta^2 G_{int}$ [$\Delta\Delta\Delta G$ in the nomenclature of Green and Shortle (*42*)]. In contrast, we find that changes in slope values in our double mutants are largely additive from single mutants (Figure 4); i.e., values of $\Delta\Delta m_{double}$ are generally low where Green and Shortle found large values. Even when limited to double mutants with nonadditive stability effects, the changes in slope values are additive from the slope value changes of the corresponding single mutant. There is only poor correlation ($r = 0.3265$) between $\Delta\Delta m_{double}$ and $\Delta^2 G_{int}$ ($\Delta\Delta\Delta G$). This is in marked contrast to the results of Green and Shortle.

DISCUSSION

Side chain packing in the core of proteins, already acknowledged as being important, may be on par with hydrophobicity as a factor in stabilizing the native state (*2*). The degree of additivity in packing stability effects is of great importance for protein engineering (*25*). If multiple mutants generally have additive stability effects, protein engineering will be much more predictable given the structural and energetic information of single mutants. However, if residues have energetically significant interactions with one another at great distances, it will be very difficult to correctly predict stability effects in multiple mutants. Even simple rules, such as the effect of distance between the substituted side chains upon the additivity of their stability effects, would be very

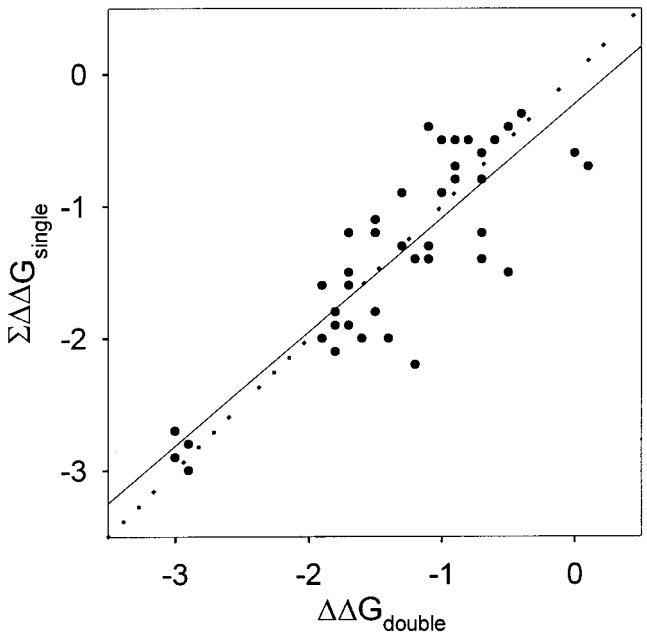

FIGURE 2: Plot of $\sum\Delta\Delta G_{single}$ (in kilocalories per mole) calculated from the corresponding values for the single mutants vs $\Delta\Delta G_{double}$ for the double mutants (in kilocalories per mole). The dotted line is the theoretical line with a unit slope. Points below this line are double mutants that are more stable than predicted from the single mutants. Those above the line are less stable than predicted. The solid line is the regression line through the data points. The correlation coefficient (r) equals 0.8572.

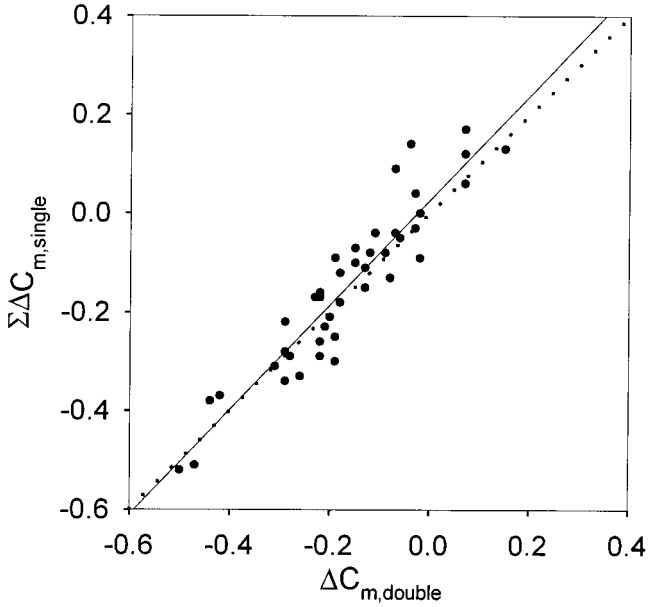

FIGURE 3: Plot of $\sum\Delta C_{m,single}$ (in units of moles per liter) as calculated from single mutants vs $\Delta C_{m,double}$ of double mutants (in units of moles per liter). The dotted line is the theoretical line with a unit slope. Points below this line are double mutants that have higher values of C_m than predicted from the single mutants. Those above the line have values lower than predicted. The solid line is the regression line through the data points. The correlation coefficient (r) equals 0.9274.

useful in rationally modifying protein structure. We have therefore created the largest systematically sampled set of double mutant proteins yet made to investigate the additivity of packing effects.

The energetic effects of interactions between two side chains can be analyzed by a double mutant cycle (*33*, *35*, *45*), in which double mutants and corresponding single mutants are constructed. The energetic difference between the stability change of a double mutant and the sum of the stability change of the component single mutants is the energy of the interaction between two mutated side chains.

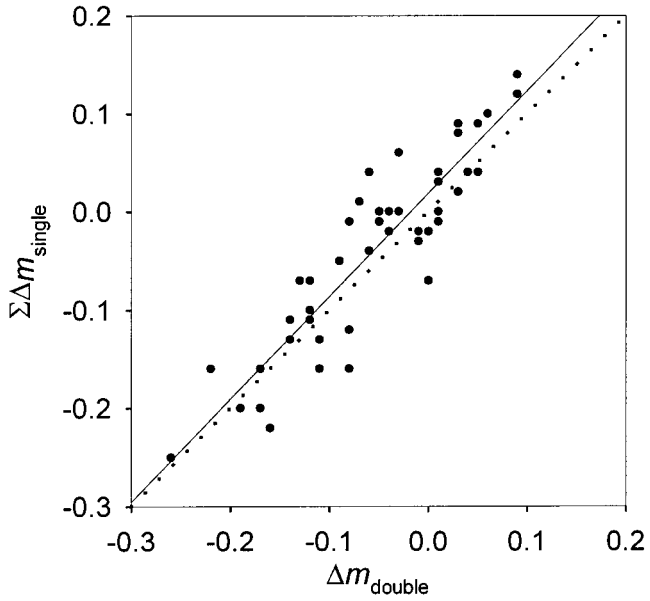

FIGURE 4: Plot of $\sum\Delta m_{single}$ as calculated from the values for single mutants vs Δm_{double} values for the double mutants. The dotted line is the theoretical line with a unit slope. Points below this line are double mutants that have higher values of m_{GuCHl} than predicted from the single mutants. Those above have values lower than predicted. The solid line is the regression line through the data points. The correlation coefficient (r) equals 0.9045.

As detailed in the Results, most of the double mutants generated here have largely additive behaviors in changes in stability, slope value, and midpoint concentration. This is the ideal situation from a protein design standpoint. However, not all changes in protein structure have additive effects, and it is important to identify the mutations that will fall into this category. Two factors that are obviously likely to be related to additivity are the distance between the two mutated residues and the degree to which the mutations alter the size of the residues. We consider each of these factors in turn.

It seems obvious that if two residues are close to one another there will be some interaction. However, when the interaction energy (Δ^2G_{int}) is plotted versus the distance between the β-carbons of the mutated side chains in the wild-type structure (not shown), there is no simple correlation between that nonadditive stability effects and the distances between two mutated side chains in the wild-type structure. On the other hand, when the $C_\beta-C_\beta$ distance between two residues is >8.0 Å, Δ^2G_{int} is zero within experimental error, consistent with the additive behavior of the single mutant effects.

The $C_\beta-C_\beta$ distances are not a very precise indicator of how close two residues are and their degree of interaction. We therefore determined all instances where any carbon atom in the side chain of these six residues approaches within ≤6 Å of another carbon atom in another side chain of the six positions. (A table of these atom−atom distances is available as Supporting Information.) If the hydrogens bonded to these carbons are taken into consideration, most of these 6 Å carbon−carbon interactions could be considered van der

Waals contacts. Of the 11 pairs of positions examined here, eight were found to have such interactions in the wild type: 23 and 25, 23 and 66, 25 and 72, 66 and 72, 66 and 92, 66 and 99, 72 and 92, and 92 and 99.

Recall that at least some double mutants at the pairs of positions 23 and 25, 23 and 66, 25 and 72, 66 and 72, 66 and 92, 66 and 99, and 92 and 99 have nonadditive stability effects. This list of pairs with nonadditive effects is nearly identical to the list of pairs with close contacts.

Therefore, we propose that nonadditive stability effects in the core of staphylococcal nuclease in the main occur only when the side chains are in van der Waals contact in the wild-type and/or mutant protein. It is the alteration of the van der Waals contacts that presumably leads to nonadditive effects.

Note that the presence of close contact between two residues appears to be necessary but not sufficient for nonadditive behavior. Not all the double mutants in contact had nonadditive behavior, and the number of contacts does not correlate well with the magnitude of the interaction energy. This does not invalidate the idea that van der Waals interactions are at the heart of nonadditivity. While the wild type, single, and double mutants would all have van der Waals interactions that were different in a geometrical or structural sense, at least some of these sets of interactions would be similar energetically. The double mutant would consequently be energetically additive. We also note that pairs in which at least one of the two is valine in the wild type seem likely to exhibit nonadditive behavior. This makes sense as either isoleucine or leucine increases the size of the side chain and the opportunity for interaction with other nearby residues.

If we test this hypothesis by examining the opposite extreme, mutants with additive effects, it seems to hold. Recall that all double mutants at the pairs of positions 25 and 66, 23 and 72, 72 and 92, and 72 and 99 have additive stability effects. Of these four pairs, only the pair of positions 72 and 92 appears on the list of pairs with close contacts. However, it is interesting to note that most of the interactions between positions 72 and 92 are near the edge of our somewhat arbitrary 6 Å cutoff. Further, both positions 72 and 92 are isoleucines in wild-type nuclease. This means that neither the valine nor leucine substitution at either site adds overall size or length to the side chain. Thus, the mutants are unlikely to have a greater level of interaction than the wild type. Also, an isoleucine to valine substitution is fairly nondisruptive to packing, retaining as it does the same branching point on the side chain. Thus, the presence of additive stability effects at positions 72 and 92 still appears to be largely consistent with the concept that altered van der Waals interactions lead to nonadditive energetic behavior.

We find it somewhat surprising that if there is no van der Waals contact between two residues they have additive behavior. It is intuitive that the degree of nonadditive behavior should be reduced as distance and intervening side chains are introduced between two mutations in the core of a protein. Yet a gradual attenuation of nonadditive effects is not found. The situation is instead simpler. As soon as van der Waals contact is lost, the two single mutations are largely energetically independent. Again, there is no obvious connection at this still rather simple level of analysis between the degree of nonadditivity and the number of van der Waals contacts, but the idea that contact is necessary for nonadditive behavior seems to hold.

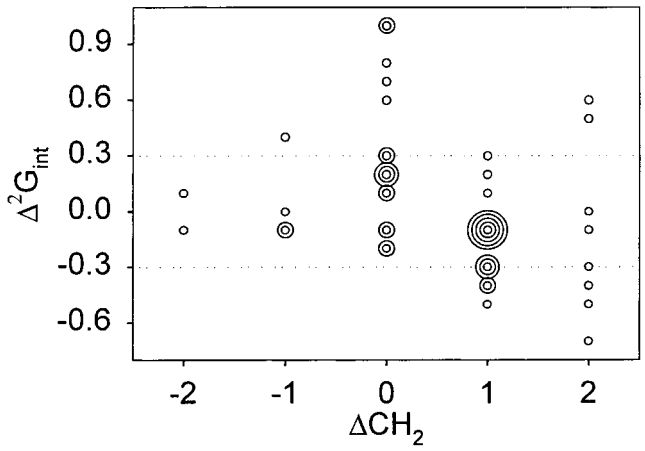

FIGURE 5: Bullseye-bubble plot of double mutant interaction energies, Δ^2G_{int} (in kilocalories per mole), organized by the value of ΔCH_2 (the change in the number of methylene groups upon mutation). In the case of coincident data points, the size of the bubble and the number of rings indicate the number of mutants with that particular interaction energy. The dashed lines at -0.3 and 0.3 kcal/mol show the range of Δ^2G_{int} values within which the effects of the single mutants are considered additive.

While the presence of nonadditivity is related to contact between two sites, the mathematical sign of any nonadditivity effect appears to be controlled primarily by the change in the size of the side chains. We approximate the change in side chain volume by the change in the number of methylene groups (ΔCH_2). For example, a valine to leucine mutant adds one methylene group, but if a double mutant also had an isoleucine to valine mutation, the value of ΔCH_2 for that double mutant would be zero. The values of Δ^2G_{int} for the double mutants are shown in Figure 5 grouped by ΔCH_2. As seen in Figure 5, the double mutants with one or two methylene group reductions almost all have additive stability effects, although the limited number of such mutations does not give much statistical weight to this observation. More telling is the large number of double mutants with no change in the number of methylene groups. Those with nonadditive stabilities have positive values of Δ^2G_{int}. In other words, these nonadditive double mutants are always more stable than expected from the stabilities of the component single mutants. On the other hand, when the sizes of the two side chains are increased by a net one or two methylene groups relative to the wild type, the nonadditive double mutants usually, but not always, have negative values of Δ^2G_{int}. In other words, these nonadditive double mutants are usually less stable than expected from the stabilities of the component single mutants.

Considered from the standpoint of side chain packing, this trend makes perfect sense. Most single mutations are destabilizing whether they increase or decrease side chain size, presumably because they lose favorable van der Waals interactions and introduce strained contacts or unfavorable bond angles into the structure. A second mutation allows the possibility of further rearrangement of side chains to find lower-energy conformations. This is especially true if the mutation causes a diminution in overall chain size, as the new void space allows still more side chain flexibility. On the other hand, increasing the size of the side chain reduces or may entirely negate the opportunity the second mutation

presents for side chain rearrangement and may cause unfavorable steric interactions. Generally speaking, the absence of voids in a structure is a good thing. Each added methylene represents more potential van der Waals interactions and hydrophobic transfer free energy in some optimal arrangement. However, the more atoms that are forced into a space, the fewer possible ways there are to arrange them and the less likely that such an optimal arrangement may be found. Indeed, it may not be possible to pack them all into the previous space, and larger-scale rearrangements of the structure may be required. There are two double mutants here, V66L/V99L and V66I/V99I, which have Δ^2G_{int} values of 0.6 and 0.5 kcal/mol, respectively, showing that such optimized arrangements can indeed be found occasionally even when two methylene equivalents are added. But these positive values are the exception to the rule, and increasing side chain size overall usually causes a greater loss of stability than expected from single mutant effects. We view this experimental trend as further evidence that the energetic importance of packing is much greater than previously appreciated.

Double mutants have been previously made in staphylococcal nuclease (*31*), and comparison to this earlier work is appropriate. In particular, Green and Shortle (*42*) previously studied the stability effects of 71 double mutants in staphylococcal nuclease. Nonadditive stability effects were common among these double mutants, with 39 of 71 having nonadditive effects. Not only is this percentage much higher than we found, it is even more striking because most of the mutants in this earlier work were much more widely separated in the structure than those discussed here. This appears at first to be in direct contradiction to our findings. However, two points should be considered. First, we have found, as detailed in the preceding paper (*2*), that the energetic value used by Green and Shortle (*42*) for the I72V single mutant was incorrect by 0.6 kcal/mol. Correction of this error shows that five double mutants, previously thought to be nonadditive, are in fact additive. The percentage of nonadditive mutants is therefore reduced to a significant degree. Nevertheless, the percentage of nonadditive mutants is higher than we find, and most are not in van der Waals contact, still in apparent contradiction to our findings.

The second point to consider is that the main focus of this earlier work was exploration of the additivity of changes in the denatured state as reflected in the slope value. The slope value, m_{GuHCl}, reflects the sensitivity of the protein unfolding reaction to the changes in denaturant concentration. The change in the slope value has been suggested to be correlated with the change in the amount of residual structure in the denatured state based on both theory (*46*) and experimental results (*31*, *38*, *47*−*51*). Therefore, the component single mutants in Green and Shortle (*42*) were selected in large part because of their large slope changes and presumed large changes in the denatured state. These single mutants were generally not in close contact in the native state, and many were solvent-exposed since packing effects in the protein core were not of interest to them.

Consistent with the idea that energetically significant changes can occur in the denatured state, Green and Shortle (*42*) found a good correlation between Δ^2G_{int} ($\Delta\Delta\Delta G$ in their nomenclature) and $\Delta\Delta m_{double}$. They therefore proposed that nonadditive stability effects of many of their double mutants, perhaps especially those widely separated in the native state, were related to the changes in the residual structure in the denatured state. We, however, find poor correlation between $\Delta\Delta m_{double}$ and Δ^2G_{int}, again in apparent contradiction to Green and Shortle's results. In addition, our values of $\Delta\Delta m_{double}$ are generally low where Green and Shortle (*42*) found large values.

This limited range of $\Delta\Delta m_{double}$ values in our data obscures the possible correlation with Δ^2G_{int}. Indeed, if our data are combined with those of Green and Shortle, the constant of the correlation between $\Delta\Delta m_{double}$ and Δ^2G_{int} for the global data set is comparable to that found in Green and Shortle's data with its more extreme range of $\Delta\Delta m_{double}$ alone (not shown). Therefore, our data do not contradict those of Green and Shortle. When single mutants with large slope effects likely related to denatured state perturbations are combined, it is quite reasonable to find nonadditive effects related to changes in the denatured state in the double mutant.

In the case of the data presented here, where single mutants were chosen on the basis of their proximity in the core of the protein, the conclusions are, not surprisingly, different. The clear relationship between Δ^2G_{int} and van der Waals contacts as well as the influence of the change of the size of the side chain both indicate that the nonadditive stability effects among our double mutants are dominated by interactions between the two mutated side chains in the protein native state. Poor correlation between Δ^2G_{int} and $\Delta\Delta m_{double}$ was found, even if comparison was restricted to the double mutants with nonadditive stability effects. This does not rule out denatured state effects, but it does indicate that changes in the denatured state are much less important. Therefore, we conclude that when packing of the core of staphylococcal nuclease is being considered, the denatured state effects are negligible compared to the native state effects.

If this is the case in staphylococcal nuclease, arguably the protein with the best documented case for energetically important denatured state changes, it seems likely that denatured state effects are likely to be of relatively minor importance compared to native state effects when considering the packing of most protein cores.

Since predicting changes in the native state alone is difficult enough without the additional complication of attempting to take into account denatured state changes, this is excellent news for those who wish to rationally alter the packing of a protein core. Coupled with our other major conclusion that complex, difficult-to-predict, nonadditive behaviors occur in only residues in van der Waals contact, the problem of predicting core packing is likely to be amenable even to simplified computational approaches.

ACKNOWLEDGMENT

We thank Zhiqiang Lu for his help with the figures. We also thank him and Joshua Sakon for helpful discussion.

SUPPORTING INFORMATION AVAILABLE

A table of carbon−carbon distances of ≤6 Å between pairs of selected side chains. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.

2. Holder, J., Bennett, A. F., Chen, J., Spencer, D. S., Byrne, M. P., and Stites, W. E. (2001) *Biochemistry 40*, 13998−14003.
3. Vlassi, M., Cesareni, G., and Kokkinidis, M. (1999) *J. Mol. Biol. 285*, 817−827.
4. Ratnaparkhi, G. S., and Varadarajan, R. (2000) *Biochemistry 39*, 12365−12374.
5. Richards, F. M., and Lim, W. A. (1993) *Q. Rev. Biophys. 26*, 423−498.
6. Hurley, J. H. (1994) in *The Protein Folding Problem and Tertiary Structure Prediction*, pp 549−578, Birkhäuser, Boston.
7. Hubbard, S. J., and Argos, P. (1995) *Curr. Opin. Biotechnol. 6*, 375−381.
8. Lee, B., and Vasmatzis, G. (1997) *Curr. Opin. Biotechnol. 8*, 423−428.
9. Lazar, G. A., and Handel, T. M. (1998) *Curr. Opin. Chem. Biol. 2*, 675−679.
10. Buckle, A. M., Cramer, P., and Fersht, A. R. (1996) *Biochemistry 35*, 4298−4305.
11. Takano, K., Ogasahara, K., Kaneda, H., Yamagata, Y., Fujii, S., Kanaya, E., Kikuchi, M., Oobatake, M., and Yutani, K. (1995) *J. Mol. Biol. 254*, 62−76.
12. Takano, K., Yamagata, Y., Fujii, S., and Yutani, K. (1997) *Biochemistry 36*, 688−698.
13. Matthews, B. W. (1996) *FASEB J. 10*, 35−41.
14. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) *Science 255*, 178−183.
15. Rashin, A. A. (1993) *Prog. Biophys. Mol. Biol. 60*, 73−200.
16. Lee, B. (1993) *Protein Sci. 2*, 733−738.
17. Wynn, R., Harkins, P. C., Richards, F. M., and Fox, R. O. (1996) *Protein Sci. 5*, 1026−1031.
18. Wynn, R., Harkins, P. C., Richards, F. M., and Fox, R. O. (1997) *Protein Sci. 6*, 1621−1626.
19. Khorasanizadeh, S., Peters, I. D., Butt, T. R., and Roder, H. (1993) *Biochemistry 32*, 7054−7063.
20. Wynn, R., and Richards, F. M. (1993) *Protein Sci. 2*, 395−403.
21. Hurley, J. H., Baase, W. A., and Matthews, B. W. (1992) *J. Mol. Biol. 224*, 1143−1159.
22. Baldwin, E., Xu, J., Hajiseyedjavadi, O., Baase, W. A., and Matthews, B. W. (1996) *J. Mol. Biol. 259*, 542−559.
23. Daopin, S., Alber, T., Baase, W. A., Wozniak, J. A., and Matthews, B. W. (1991) *J. Mol. Biol. 221*, 647−667.
24. Sandberg, W. S., and Terwilliger, T. C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1706−1710.
25. Skinner, M. M., and Terwilliger, T. C. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10753−10757.
26. Zhang, H., Skinner, M. M., Sandberg, W. S., Wang, A. H., and Terwilliger, T. C. (1996) *J. Mol. Biol. 259*, 148−159.
27. Akasako, A., Haruki, M., Oobatake, M., and Kanaya, S. (1997) *J. Biol. Chem. 272*, 18686−18693.
28. Tsuji, T., Chrunyk, B. A., Chen, X., and Matthews, C. R. (1993) *Biochemistry 32*, 5566−5575.
29. Ohmae, E., Iriyama, K., Ichihara, S., and Gekko, K. (1998) *J. Biochem. 123*, 33−41.
30. Finucane, M. D., and Woolfson, D. N. (1999) *Biochemistry 38*, 11613−11623.
31. Shortle, D., and Meeker, A. K. (1986) *Proteins 1*, 81−89.
32. Kono, H., Nishiyama, M., Tanokura, M., and Doi, J. (1998) *Protein Eng. 11*, 47−52.
33. Horovitz, A. (1996) *Folding Des. 1*, R121−R126.
34. Horovitz, A., and Fersht, A. R. (1992) *J. Mol. Biol. 224*, 733−740.
35. Horovitz, A., and Fersht, A. R. (1990) *J. Mol. Biol. 214*, 613−617.
36. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
37. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
38. Shortle, D., Stites, W. E., and Meeker, A. K. (1990) *Biochemistry 29*, 8033−8041.
39. Byrne, M. P., Manuel, R. L., Lowe, L. G., and Stites, W. E. (1995) *Biochemistry 34*, 13949−13960.
40. Stites, W. E., Byrne, M. P., Aviv, J., Kaplan, M., and Curtis, P. M. (1995) *Anal. Biochem. 227*, 112−122.
41. Schwehm, J. M., and Stites, W. E. (1998) *Methods Enzymol. 295*, 150−170.
42. Green, S. M., and Shortle, D. (1993) *Biochemistry 32*, 10131−10139.
43. Chen, J., Lu, Z., Sakon, J., and Stites, W. E. (2000) *J. Mol. Biol. 303*, 125−130.
44. Vriend, G. (1990) *J. Mol. Graphics 8*, 52−59.
45. Tuffery, P., Etchebest, C., Hazout, S., and Lavery, R. (1991) *J. Biomol. Struct. Dyn. 8*, 1267−1289.
46. Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.
47. Shortle, D., and Meeker, A. K. (1989) *Biochemistry 28*, 936−944.
48. Green, S. M., Meeker, A. K., and Shortle, D. (1992) *Biochemistry 31*, 5717−5728.
49. Wrabl, J., and Shortle, D. (1999) *Nat. Struct. Biol. 6*, 876−883.
50. Baskakov, I. V., and Bolen, D. W. (1998) *Biochemistry 37*, 18010−18017.
51. Bowler, B. E., May, K., Zaragoza, T., York, P., Dong, A., and Caughey, W. S. (1993) *Biochemistry 32*, 183−190.

BI011268L