Protein Science (1995), *4*:1815–1823. Cambridge University Press. Printed in the USA.
Copyright © 1995 The Protein Society

Interactions in nonnative and truncated forms of staphylococcal nuclease as indicated by mutational free energy changes

RICHARD WYNN,[1] CORT L. ANDERSON,[1,2,3] FREDERIC M. RICHARDS,[1] AND ROBERT O. FOX[1,2,3]

[1] Department of Biochemistry and Biophysics and [2] Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06511

(RECEIVED March 23, 1995; ACCEPTED June 14, 1995)

Abstract

Several mixed disulfide variants of staphylococcal nuclease have been produced by disulfide bond formation between nuclease V23C and methane, ethane, 1-propane, 1-*n*-butane, and 1-*n*-pentane thiols. Although CD spectroscopy shows that the native state is largely unperturbed, the stability toward urea-induced unfolding is highly dependent on the nature of the group at this position, with the methyl disulfide protein being the most stable. The variant produced by modification with iodoacetic acid, however, gives a CD spectrum indicative of an unfolded polypeptide. Thiol-disulfide exchange equilibrium constants between nuclease V23C and 2-hydroxyethyl disulfide have been measured as a function of urea concentration. Because thiol-disulfide exchange and unfolding are thermodynamically linked, the effects of a mutation (disulfide exchange) can be partitioned between various conformational states. In the case of unmodified V23C and the 2-hydroxyethyl protein mixed disulfide, significant effects in the nonnative states of nuclease are observed.

Truncated forms of staphylococcal nuclease are thought to be partially folded and may be good models for early folding intermediates. We have characterized a truncated form of nuclease comprised of residues 1–135 with a V23C mutation after chemical modification of the cysteine residue. High-resolution size-exclusion chromatography indicates that modification brings about significant changes in the Stokes radius of the protein, and CD spectroscopy indicates considerable differences in the amount of secondary structure present. Measurement of the disulfide exchange equilibrium constant between this truncated protein and 2-hydroxyethyl disulfide indicate significant interactions between position 23 and the rest of the protein when the urea concentration is lower than 1.5 M. Comparison of disulfide exchange equilibrium constants for the truncated and nonnative full-length forms indicates that some of the interactions in unfolded nuclease are dependent upon residues far removed in sequence and not in contact in the native structure.

Keywords: cysteine modification; thermodynamic cycle; truncated nuclease; unfolded state

The stability of a protein is determined by the difference in free energy between the folded and unfolded forms. Although most researchers agree that mutations can cause large effects on the free energy of the native state, it has been difficult to determine the thermodynamic consequence of mutations on the nonnative state(s). Despite early assumptions that the unfolded state of a protein was nearly a random coil, recent NMR studies have shown that this is not always the case (Evans et al., 1991; Neri et al., 1992). This view has been supported by thermodynamic arguments (Shortle & Meeker, 1986; Green et al., 1992) and also by theoretical studies (Shortle et al., 1992). Presently, the linkage between the nonrandom character of nonnative proteins and the thermodynamic consequence of amino acid changes has not been established. Here we investigate the effects of amino acid changes on the native and various nonnative states of staphylococcal nuclease (Fig. 1).

Previously we have introduced a combined mutagenesis/chemical modification method that facilitates the incorporation

Reprint requests to: Richard Wynn, Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208114, 260 Whitney Avenue, New Haven, Connecticut 06520-8114; e-mail: wynn@csbmet.csb.yale.edu.

[3] Present address: Department of Human Biological Chemistry & Genetics, University of Texas Medical Branch at Galveston, Galveston, Texas 77555.

Abbreviations: BME, β-mercaptoethanol; BMEDS, 2-hydroxyethyl disulfide; ACM, *N*-acetylcysteine methylamide; Nuc, staphylococcal nuclease; ΔNuc, truncated nuclease consisting of residues 1–135; Gdn-HCl, guanidine hydrochloride; ΔG_u^0, standard-state free energy of unfolding; IAC, iodoacetic acid.

Fig. 1. Ribbon diagram of staphylococcal nuclease. Coordinates were from Hynes and Fox (1991). The valine side chain at position 23 is shown in white. Also shown in black at the α carbon is position 135, the last residue in the truncated forms used in this study.

of unnatural amino acids into the core of a protein (Wynn & Richards, 1993a). In this method, a cysteine residue is placed at a position that is normally buried in the wild-type structure. After denaturant-induced unfolding of the protein, the cysteine residue is modified with one of a variety of available thiol modification reagents. After sufficient time, folding is driven by removal of the denaturant and the new protein can be characterized. Using this technique, it is possible to introduce a large array of unnatural side chains into a protein core. We are able to change a side chain by a single carbon atom at a time in a prespecified manner. Here we focus on a series of modifications of nuclease V23C. Although the CD spectra for most of these proteins do not detect a change in the folded structures, there are large differences in the thermodynamic stabilities. There are changes in the dependence of the free energy of unfolding on the urea concentration, the m_{den} value, which has been interpreted as indicative of either significant interactions in the unfolded state or formation of an equilibrium folding intermediate.

A problem with directly determining the effect of a mutation on a particular conformational state, the mutational free energy, is that the conversion from one amino acid to another at equilibrium is not an experimentally accessible process. This problem can be overcome if normal mutagenesis is replaced by reversible chemical modification. Then protein unfolding and protein modification can be incorporated into a thermodynamic cycle (see the Materials and methods). Using such a scheme allows one to partition the effects of a change in a protein between the folded and unfolded forms. In previous studies with T4 lysozyme and *Escherichia coli* thioredoxin using this approach, no effects were observed in the unfolded states (Lu et al., 1992; Wynn & Richards, 1993b). In contrast, here we report significant effects in the nonnative form(s) of nuclease. Extrapolation of our data suggests that a large portion of the change in stability for two nuclease variants is due to interactions in the nonnative form(s). Assuming a three-state unfolding equilibrium, we estimate values of 1.05, 3.32, and 2.00 for the disulfide exchange equilibrium constant in the native, intermediate, and unfolded forms, respectively.

Truncated forms of staphylococcal nuclease have been used as models of nonnative states (Shortle & Meeker, 1989). Residual structure in these forms has been detected by CD spectroscopy, size-exclusion chromatography (Shortle & Meeker, 1989), small angle X-ray scattering (Flanagan et al., 1992), and NMR spectroscopy (Alexandrescu et al., 1994). Here we characterize a nuclease fragment comprised of residues 1–135 with a V23C mutation. Modification of the thiol groups gives a series of proteins that show changes in the Stokes radius as measured by high-performance size-exclusion chromatography and corresponding differing amounts of secondary structure as judged by CD spectroscopy. By measuring the disulfide exchange equilibrium constant between the truncated protein and BMEDS, we have been able to quantitate the thermodynamic effect of an amino acid change on these proteins. The results presented here show that despite the fact that these truncated proteins do not

adopt the native folded state, there are significant intramolecular interactions and the strength of these interactions is dependent upon the nature of the side chain at position 23. The value for the disulfide exchange equilibrium constant in the absence of denaturant for the truncated form is essentially equal to that estimated for the intermediate formed during the unfolding of full-length nuclease. This suggests that the truncated form of nuclease studied here is a good model for the equilibrium folding intermediate.

Results

Full-length nuclease

Several variants of nuclease V23C have been produced using the chemical modification scheme described below (see the Materials and methods). In this report, the protein variants will be named according to the thiol that forms a mixed disulfide to the protein. The unmodified protein will be referred to as simply the unmodified protein or PSH. The protein that has been reacted with IAC will be referred to as the IAC protein. Figure 2A shows the CD spectra for these variants.

There is little difference between the spectra for all of the full-length variants in this study, except the IAC protein, indicating little change in the secondary structures of these proteins. Additionally, these spectra are indistinguishable from that of wild-type nuclease. The spectrum of the IAC variant is highly perturbed. The large negative peak at approximately 197 nm is indicative of high amounts of random coil. In fact, size-exclusion chromatography indicates that the average structure for this variant is highly expanded relative to both the full-length and truncated variants in this study (see below).

The data from urea unfolding curves using fluorescence and CD (data in brackets) are presented in Table 1. The signal strength was typically lower when the unfolding process was followed by CD and, as a consequence, the estimated error is larger for these measurements. The maximum difference in the means is 9% or less for all m_{den} values and 5% or less for the $[\text{urea}]_{1/2}$ values. Combined with the estimated errors there is no significant difference between the values obtained by the two methods. Although this does not prove rigorously a two-state scheme, it does prove that any intermediate that is populated is spectroscopically indistinguishable from the unfolded form, a conclusion reached by others (Carra et al., 1994). Because the fluorescence data are more precise, we have used the numbers derived from these measurements in all calculations in this work.

The stability, as represented by ΔG^0_{H2O}, increases dramatically as the unmodified protein is converted to the protein-methyl mixed disulfide (the methyl protein), which is the most stable variant in this study. ΔG^0_{H2O} drops off consistently as the modification group becomes larger (methyl to pentyl) or more polar (BME). The m_{den} value, the dependence of $\Delta G^0_{\text{unfolding}}$ on urea concentration, decreases asymptotically as the side chain at position 23 becomes more hydrophobic. The value for the BME protein is slightly higher than the value for the methyl protein. Estimations of the hydrophobicity of these two side chains based on transfer free energies of thiols and disulfides (Saunders et al., 1993) and group transfer free energies (Eisenberg & McLachlan, 1986) indicate that the BME side chain is slightly more polar than the methyl side chain.

The data for the disulfide exchange reaction between BMEDS and protein are presented in Figure 3. Data were collected down to 1.0 M urea for full-length nuclease. Experiments at lower concentrations were not possible due to extremely long reaction

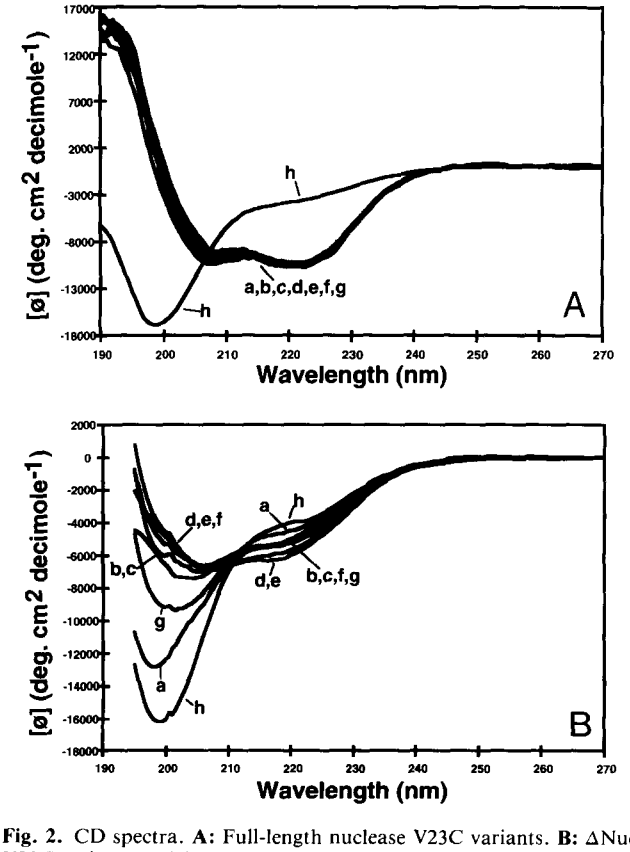

Fig. 2. CD spectra. **A:** Full-length nuclease V23C variants. **B:** ΔNuc V23C variants. a, PSH; b, methyl; c, ethyl; d, propyl; e, butyl; f, pentyl; g, BME; h, IAC.

Table 1. *Urea unfolding data for nuclease V23C variants*[a]

Protein	ΔG_{H2O}[b]	m_{den}[c]	$[\text{Urea}]_{1/2}$[d]
PSH	4.08 (4.14)	2.50 (2.41)	1.63 (1.72)
Methyl	5.87 (5.69)	2.20 (2.09)	2.67 (2.72)
Ethyl	5.48 (5.47)	2.06 (2.13)	2.66 (2.57)
Propyl	4.09 (4.14)	1.95 (2.00)	2.10 (2.07)
Butyl	3.37 (3.65)	1.91 (2.10)	1.76 (1.74)
Pentyl	2.77 (2.71)	1.95 (1.83)	1.42 (1.48)
BME	3.20 (3.52)	2.24 (2.42)	1.43 (1.46)

[a] Proteins are named according to the thiol that forms the mixed disulfide with the protein. The first value listed is that determined using fluorescence to monitor the unfolding reaction. The value in brackets was derived from CD data.

[b] ΔG_{H2O} is in kcal mol^{-1}. Estimated error is ±0.14 (fluorescence values) or ±0.30 (CD values).

[c] m_{den} is in kcal/mol^{-1} M^{-1}. Estimated error is ±0.15 (fluorescence values) or ±0.28 (CD values).

[d] $[\text{Urea}]_{1/2}$ is in M. Estimated error is 0.05 (fluorescence values) or 0.10 (CD values).

times, which are incompatible with urea due to modification of the protein by cyanate ions formed by urea decomposition (Stark, 1965; Wynn & Richards, 1993b). Guanidine hydrochloride was not used because it has been shown to affect disulfide exchange reactions (Creighton, 1977). At high concentrations of urea, the value for $K_{red,app}$ is 2.01. This is within experimental error of the value determined for *N*-acetylcysteine methylamide, a cysteine analog that we use to account for the free energy change associated with the covalent bond change (see the Materials and methods and Wynn & Richards [1993b]). As the urea concentration is lowered to within the folding-unfolding transition region for the two proteins, the value of $K_{red,app}$ drops as predicted by Equation 1 and the relative stabilities of the two proteins. The value of $K_{red,N}$ is constant at approximately 1.1 in the folding-unfolding transition region. The value of $K_{red,\Sigma Ui}$, on the other hand, is constant at 2.0 from 3.0 M to approximately 2.0 M urea and then increases as the urea concentration is lowered further. The increasing value for $K_{red,\Sigma Ui}$ indicates an increasingly favorable modification reaction.

ΔNuc

The CD spectra of the various ΔNuc variants are shown in Figure 2B. A negative peak just under 200 nm is indicative of an unstructured polypeptide chain (Brahms & Brahms, 1980). By this criterion, the IAC protein is clearly the least structured variant studied here followed by the unmodified protein. The methyl, ethyl, propyl, butyl, and pentyl variants cluster with a lower percentage of unstructured polypeptide chain, whereas the unmodified and BME proteins lie intermediate between these and the IAC protein. The signal at 220 nm is indicative of helix and sheet content. The signal at this wavelength also indicates the least amount of structure for the IAC protein. The spectra for the remainder of the variants indicate more secondary structure than the IAC protein with a roughly inverse relationship between the signal at 198 nm and 220 nm. Additionally, the spectrum for the full-length IAC protein is virtually superposable with that of the truncated IAC protein.

The results from the SEC experiments are listed in Table 2. Elution time from a size-exclusion column is dependent upon the average Stokes radius of the solute molecule of interest. The Stokes radius for a protein is dependent upon the chain length and the conformation of the protein. Because all of the ΔNuc variants in this study are comprised of residues 1–135, any change in the Stokes radius, as manifested by a change in elution time, must reflect a change in the average conformation. The IAC protein has the smallest elution time for any of the truncated forms (43.87 min), which indicates the largest Stokes radius. This correlates with the fact that this form has the largest fraction of unstructured peptide chain as judged by CD spectroscopy (discussed above). This correlation holds with the unmodified and BME proteins having elution times (44.18 and 45.08 min, respectively) between that of the IAC protein and the values for the methyl, ethyl, propyl, butyl, and pentyl forms, which again cluster at approximately 45.5 min. As expected, the full-length methyl protein has an elution volume higher than any of the truncated forms despite having an additional 14 residues. Clearly, this is due to the highly compact nature of protein native states (Richards, 1977). The full-length IAC protein has the smallest elution time measured. This presumably reflects the combined effects of having a low percentage of structured peptide chains (see above) and an extra 15 residues relative to any of the truncated forms.

Table 2. *SEC results for ΔNuc V23C variants*

Protein	Elution time (min)
PSH	44.18
Methyl	45.58
Ethyl	45.51
Propyl	45.58
Butyl	45.51
Pentyl	45.49
BME	45.08
IAC	43.87
IAC-F.L.[a]	43.07
Methyl-F.L.[a]	46.40

[a] -F.L. designates the full-length V23C protein.

The results from the disulfide exchange measurements between ΔNuc (1–135) V23C and BMEDS are shown in Figure 3. A value for $K_{red,app}$ of 2.0 represents the free energy change associated with the covalent bond change in the absence of any intramolecular interactions (Wynn & Richards, 1993b). The value for the truncated protein is equal to 2.03 at 2.0 M urea but significantly increases as the urea concentration is lowered (see Fig. 3). In the absence of denaturants, the difference in the free energy for the intramolecular interactions between the unmodified and BME forms of the truncated protein is approximately 0.3 kcal/mol ($\Delta\Delta G^0 = -RT \ln[3.41/2.0]$).

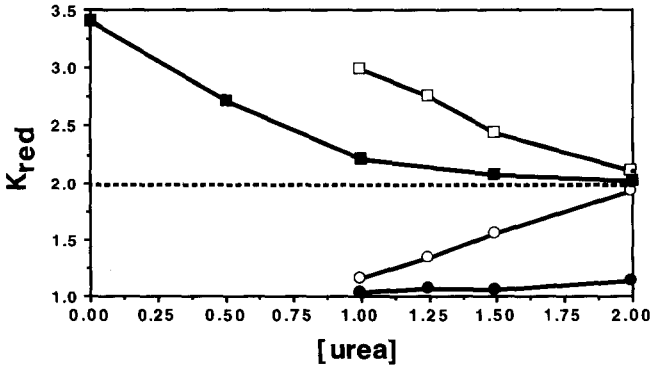

Fig. 3. Plot of k_{red} for BMEDS and full-length unfolded nuclease (open squares), native nuclease (filled circles), truncated nuclease (filled squares), and $K_{red,app}$ (open circles). Data for ACM are represented by the dashed line.

Discussion

The folded state of staphylococcal nuclease

Figure 1 shows a ribbon diagram of nuclease (Hynes & Fox, 1991). The structure is composed of a five-stranded β-barrel and three α-helices. Residues 1–5 and 142–149 are not observable in electron density maps. The wild-type valine side chain at position 23 is included and positions 23 and 135 are labeled. Posi-

tion 23 lies at the beginning of the second strand in the β-barrel. The valine side chain contacts residues in the first, second, third, and fifth strands of the β-barrel and additionally by residues 62 and 66 of helix 1.

Several unnatural amino acids have been incorporated into nuclease at position 23. The side chains produced here are mostly disulfide structures, whereas the IAC protein contains a thioether moiety. The methyl cysteine disulfide side chain is isosteric with methionine. Each increase of a carbon atom increases the size, hydrophobicity, and number of rotatable torsional angles. All of these side chains, with the exception of unmodified cysteine, are larger than the wild-type valine side chain, yet the native structure can apparently accommodate all of these changes without much change in the secondary structure, as indicated by the CD spectra, and presumably the tertiary structure. However, reaction with IAC introduces a negative charge into the protein, which apparently cannot be accommodated into the core of the protein. Additionally, we have obtained crystals for each variant in this study, except the IAC protein, which are nearly isomorphous with wild-type nuclease crystals, indicating only small changes in the folded structures of these proteins. It is now well known that protein structures can adjust to accommodate many amino acid changes (reviewed by Richards & Lim, 1993). Examination of the wild-type crystal structure shows a large cavity in the vicinity of position 23. X-ray crystallographic analysis of these structures is in progress and should help delineate the response of the proteins to these large changes in residue volume.

CD spectroscopy and size-exclusion chromatography clearly indicate that the IAC protein is unfolded. Although the native structure may be able to adjust to large changes in size at position 23, the introduction of a negative charge is apparently too destabilizing under the conditions of this study (pH 8.0). A survey of high-resolution structures indicates that accommodation of charge in a protein interior occurs relatively rarely and only when compensating hydrogen bonding/salt bridging partners are available (Rashin & Honig, 1986). However, mutational analysis of T4 lysozyme (Daopin et al., 1991), staphylococcal nuclease (Stites et al., 1991), and *E. coli* thioredoxin (Hellinga et al., 1992) have identified buried positions that can tolerate potentially charged side chains. In the case of V66K nuclease, the pH dependence of stability suggests that the lysine residue is unprotonated in the folded state; i.e., there is a marked perturbation of the pK_a of this residue in the folded state. If the IAC side chain must be protonated in the folded state ($pK_a \gg 8.0$), then the cost of protonation would be approximately 5 kcal/mol ($\Delta\Delta G^0 = -RT \ln\{(1 + K_{ionization,unfolded}/[H^+])/(1 + K_{ionization,folded}/[H^+])\} \approx 5$ kcal/mol at pH 8.0, if $pK_{unfolded} = 4.0$ and $pK_{folded} \gg 8.0$). The protonation penalty would be greater than the stability of any variant in this study. This estimate does not even take into account that burying a polar, yet uncharged, residue in the hydrophobic core most likely also carries an additional penalty. The fact that the BME protein is 0.9 kcal/mol less stable than the propyl protein suggests that this is true for position 23 of nuclease (the propyl and BME side chains have approximately the same volume but different polarities).

The stabilities of these proteins are markedly affected by chemical modification. The transfer of a hydrophobic group from solvent into the hydrophobic interior of a protein is thought to be one of the main driving forces for protein folding (Dill, 1990). If this were the only effect to be considered for stability, then the larger hydrophobic chains should produce the most stable proteins. However, the volume of an interior residue must be constrained by the tight packing of the protein core (Ponder & Richards, 1987). The wild-type valine residue is intermediate in volume between the unmodified cysteine and the methyl cysteine disulfide. The pattern in ΔG^0_{H2O} that is observed is likely a balance between overpacking of the native structure and favorable transfer free energy, although other factors such as steric packing and entropic effects are possible. In addition, changes in the nonnative states of these proteins could have dramatic effects on the stabilities of these proteins.

The nonnative states of staphylococcal nuclease

The unfolded state of proteins has recently come under intense scrutiny (Dill & Shortle, 1991). The characteristics of this state are important in order to understand the thermodynamic stability of proteins because the stability is defined by the difference in free energy between the folded and unfolded forms. Schellman (1978) has proposed that the dependence of the free energy of unfolding on the denaturant concentration, the m_{den} value, is proportional to the surface area exposed during the unfolding process ($m \propto [A_u - A_f]$; A_u and A_f are the surface areas of the unfolded and folded states, respectively). It has been experimentally observed in a number of systems that the m_{den} value is constant in the folding–unfolding transition region within experimental error (ΔG^0_u is linear with denaturant concentration). Under this theory, it is likely that a change in m_{den} reflects a change in the nonnative state(s) because it is hard to imagine a significant change in the surface area of the highly compact native state. Changes in m_{den} have been observed for a large number of variants of staphylococcal nuclease (Shortle & Meeker, 1986; Green et al., 1992). Based on the correlation between changes in m_{den} values and changes in stability, Shortle and colleagues have concluded that amino acid changes can affect the structure and free energy of the denatured state (Shortle et al., 1990). It is also possible to account for changes in m_{den} by postulating a three-state mechanism ($N \leftrightarrow I \leftrightarrow U$). Changes in m_{den} would then be brought about by changes in the relative populations and/or stabilities of I and U (Carra et al., 1994).

The m_{den} values for the variants in this study are listed in Table 1. The largest value is for the unmodified protein, which has the most polar amino acid side chain. The m_{den} value asymptotically decreases with increased hydrophobicity. Changes in m_{den} have previously been observed in conventional mutagenesis experiments at position 23 in nuclease (Shortle et al., 1990). The change in m_{den} values, under the Schellman theory discussed above, suggests decreased surface area in the nonnative state(s) of nuclease as the hydrophobicity is increased. Surprisingly, this effect levels off as the volume increases to that of a pentyl group. If an intermediate is populated, the decreasing m_{den} value would imply increasing amounts of intermediate being formed and therefore stabilization of the intermediate relative to the unfolded form with increasing hydrophobicity at position 23. The leveling off of the m_{den} value may reflect 100% intermediate (I) being formed to the exclusion of U. In a two-state scheme, the leveling off of m_{den} might reflect a critical density that cannot be exceeded while the protein remains "unfolded" with the critical density determined by the balanc-

ing of hydrophobic clustering forces and conformational entropy.

$K_{red,\Sigma Ui}$ for the V23C nuclease-BMEDS equilibrium is equal to approximately 2.0 at urea concentrations of 2.0 M and higher. This value is within experimental error of the value determined for the small molecule cysteine analog, ACM, which indicates that there are no significant interactions in the nonnative state(s) at these concentrations and also that the cysteine side chain is solvated equally well in the unfolded protein and ACM. As the urea concentration is lowered, $K_{red,\Sigma Ui}$ continuously increases. This may reflect interactions in the nonnative state(s) that are more favorable in the BME protein. The nonpolar nature of the BME side chain, relative to the unmodified cysteine side chain, could conceivably increase hydrophobic clustering.

The fact that $K_{red,\Sigma Ui}$ changes with urea concentration suggests that the nature and presumably the surface area of the nonnative state(s) change concomitantly. This could reflect a change in the average properties of a single ensemble (the unfolded state) or a change in the populations of an intermediate and the unfolded state. Using the three-state unfolding scenario and curve fitting to Equation 17 as described in the Materials and methods gives values of 2.81 kcal/mol, 3.11 kcal/mol, and 2.13 kcal/mol M for $\Delta G^0_{2,PSH,H2O}$, $\Delta G^0_{2,PSSR,H2O}$, and m_2. Use of these values and Equation 5 yields an estimate of $K_{red,I}$ of 3.32. Thus, in the intermediate the unmodified and BME forms would differ by approximately 0.3 kcal/mol in intramolecular interactions ($-RT\ln[3.3/2.0]$). This value is close to that obtained for the truncated form in the absence of urea ($K_{red,app} = 3.41$, Fig. 3).

ΔNuc V23C variants

CD spectroscopy indicates changes in the amount of secondary structure and size-exclusion chromatography indicates changes in the Stokes radius when chemical modification is carried out at position 23 of the truncated nuclease. The disulfide exchange measurements indicate significant interactions between position 23 and the rest of the protein ($\Delta\Delta G^0 \approx 0.3$ kcal/mol). This difference is most likely to be smaller than the differences that would be measured between the unmodified form and any of the straight aliphatic chains studied on the basis of the differences in the CD spectra and size-exclusion retention times. Unfortunately, disulfide exchange measurements for these variants are precluded by the limited solubilities of the thiol and disulfide molecules involved in these equilibria.

The interactions that drive the observed behavior are likely to be hydrophobic in nature but less rigidly defined than native state interactions. Consequently, the most compact and structured forms are observed for the aliphatic variants. The relatively small differences between the methyl to pentyl forms suggests that there are no severe volume constraints. This is in contrast to the native forms, where the drop off in ΔG^0_{H2O} for the larger side chains presumably reflects overpacking of the folded structure. Additionally, the tight packing in the native state is unfavorable for the change to the BME side chain from the unmodified protein as indicated by the value of $K_{red,N}$ (less than 2, Fig. 3), whereas $K_{red,app}$ (Fig. 3) for this change in the truncated form is favorable at low urea concentrations (greater than 2). Similarly, this change is favorable for the full-length nonnative forms. $K_{red,\Sigma Ui}$ for the full-length nonnative form (Fig. 3) is larger than $K_{red,app}$ for truncated nuclease at any given urea concentration. This reflects interactions, whether direct or indirect, between position 23 and parts of the peptide chain that have been deleted in the truncated form. Conspicuous among deleted residues are leucine 137, isoleucine 139, and tryptophan 140, which could participate in hydrophobic interactions.

Conclusion

The data we have presented indicate significant thermodynamic consequences for amino acid changes in the nonnative and truncated forms of nuclease. Although the measured mutational free energies are modest (≈0.3 kcal/mol), the actual contributions to the free energies of these states should be quite significant because: (1) The measured free energies are for a change to a relatively polar side chain (BME). The magnitude of these effects should be larger for a more typical hydrophobic side chain. (2) These interactions should be present in many residues throughout the protein sequence and could add up to a much larger contribution. Knowledge of the link between structure and free energy for nonnative states is crucial to an understanding of protein stability. Future work should help elucidate the nature of this relationship.

Materials and methods

Chemicals

Methyl methanethiosulfonate was purchased from Aldrich. The synthesis and purification of ethyl ethanethiosulfonate, 1-*n*-propyl 1-*n*-propanethiosulfonate, 1-*n*-butyl 1-*n*-butanethiosulfonate, 1-*n*-pentyl pentanethiosulfonate, and 2-hydroxyethyl methanethiosulfonate has been described elsewhere (Wynn & Richards, 1993a, 1993b). Ultrapure urea was purchased from Boehringer Mannheim. Sequanal grade Gdn-HCl was purchased from Pierce. BME was purchased from Aldrich and purified by distillation under reduced pressure. BMEDS was synthesized by Cu^{+2}-catalyzed O_2 oxidation of BME as described (Hopton et al., 1968). IAC was purchased from Sigma.

Proteins

Nuclease V23C was produced using the T Gen In Vitro Mutagenesis kit from USB according to the manufacturer's protocol. Resultant nuclease genes were sequenced in their entirety in order to confirm the desired mutation and check for mutations at other sites. The gene was subcloned into pAS1 and the protein was overexpressed in *E. coli* strain AR120 and purified as described previously (Evans et al., 1989). ΔNuc (1-135) V23C was produced by PCR mutagenesis and confirmed by DNA sequencing. Protein was overexpressed and purified as described by Alexandrescu et al. (1994).

Chemical modification of nuclease V23C and ΔNuc (1-135) V23C

Protein was modified with thiosulfonate reagents and checked for protein-protein disulfides and unreacted thiol as described previously except that a G-50 size-exclusion column was used for solvent exchange instead of centriprep concentrators (Wynn & Richards, 1993a). Protein modification with IAC was performed in 6 M Gdn-HCl as described in Crestfield et al. (1963).

CD spectra

Measurements were made in 0.01 M sodium phosphate, pH 8.0, at 25 °C on an Aviv model 62DS CD spectrometer. Protein concentration was 0.05 mg/mL. Measurements were taken every 0.5 nm with a 1.5-nm bandwidth using a 1.0-mm-pathlength cuvette. A total of five scans were averaged and a blank was subtracted. In the case of the unmodified protein, the buffer also contained 0.001 M BME. For the ΔNuc variants, the buffer also contained 0.15 M NaCl.

Size-exclusion chromatography

Size-exclusion chromatography was performed using a Superose 6 column (1.0 cm × 30.0 cm) from Pharmacia with a flow rate of 0.4 mL/min and the effluent was monitored at 219 nm. The buffer used was 0.05 M Tris, 0.15 M NaCl, pH 8.0. An average of at least five runs is reported. The average standard deviation for the variants in this study was ±0.05 min. The elution time was not dependent upon the amount or concentration of protein injected onto the column. Thus, dimerization or higher order aggregation is not likely.

Protein unfolding measurements

All measurements were carried out in 0.05 M Tris, 0.001 M EDTA, pH 8.0, plus the indicated amount of urea at 25 °C. Urea concentrations were determined by index of refraction measurements (Nozaki, 1972) and all solutions were prepared fresh and purged with argon just prior to use. Urea unfolding curves were determined by following the fluorescence at 320 nm with excitation at 295 nm on a SPEX DM3000 spectrofluorometer or the ellipticity at 222 nm on an Aviv model 62DS CD spectrometer of a 10 μM protein solution. The data are fit to the following equation, which assumes two-state behavior:

$$\begin{aligned} Y = \{&(Y_n + m_n[\text{urea}]) \\ &+ (Y_u + m_u[\text{urea}]) \\ &\times [\exp(-\Delta G^0_{\text{H2O}}/RT + m_{den}[\text{urea}]/RT)] \\ &\div [1 + \exp(-\Delta G^0_{\text{H2O}}/RT + m_{den}[\text{urea}]/RT)]\}, \end{aligned}$$

where Y is the signal at a given urea concentration. Y_n, Y_u, m_n, and m_u represent the intercepts and slopes of the pre- and post-transition baselines. ΔG^0_{H2O} is a fitting parameter normally considered to be the free energy of unfolding in the absence of urea, and m_{den} represents the dependence of the free energy of unfolding on the urea concentration. This scheme assumes that the free energy of unfolding varies linearly with denaturant concentration ($\Delta G^0 = \Delta G^0_{\text{H2O}} - m[\text{urea}]$). This is a common assumption in reversible protein unfolding experiments (Pace, 1986) and is well tested in the transition region for a large number of systems.

The optical data can be fit within experimental error by the two-state model. The native state is presumed to be unique and characterized by specific CD and fluorescence parameters. However, the unfolded states may be either a single conformation or more likely an ensemble of conformations with characteristic average values for the optical parameters.

Disulfide exchange measurements

Measurement of disulfide exchange between protein and BMEDS was as described previously (Wynn & Richards, 1993b) except that protein was separated from small molecules on a Sep-pak C18 cartridge and rinsed with 20% ACN/0.01% TFA instead of 35% ACN/0.01% TFA. The thermodynamic cycle involves protein folding and protein modification (disulfide exchange). For a two-state conformational equilibrium, the cycle would be:

$$\begin{array}{ccc} RSSR + P_N SH & \xleftrightarrow{K_{PSH}} & P_U SH + RSSR \\ K_{red,N} \updownarrow & & \updownarrow K_{red,U} \\ RSH + P_N SSR & \xleftrightarrow[K_{PSSR}]{} & P_U SSR + RSH \end{array}$$

$P_N SH$ and $P_U SH$ designate the unmodified protein in the folded and unfolded forms, respectively. $P_N SSR$ and $P_U SSR$ designate the modified protein in the folded and unfolded forms, respectively. The small molecule thiol and disulfide, *RSH* and *RSSR*, are spectator molecules in the horizontal unfolding reactions that are designated by K_{PSH} and K_{PSSR}. $K_{red,N}$ and $K_{red,U}$ designate the disulfide exchange equilibrium constant in the folded and unfolded forms, respectively. Because the above scheme is a true thermodynamic cycle, the following relationship holds:

$$K_{PSH}/K_{PSSR} = K_{red,N}/K_{red,U}. \quad (1)$$

It has been shown elsewhere (Wynn & Richards, 1993b) that, for the above scheme, the experimentally determined $K_{red,app}$, the apparent K_{red} that entails all conformational states, is equal to

$$K_{red,app} = K_{red,N}(1 + K_{PSSR})/(1 + K_{PSH}). \quad (2)$$

$K_{red,N}$ may then be calculated according to Equation 2 and knowledge of $K_{red,app}$, K_{PSH}, and K_{PSSR}, which are all determined experimentally independently. By Equation 1

$$K_{red,U} = K_{red,N}(K_{PSSR}/K_{PSH}). \quad (3)$$

In order to separate contributions to k_{red} from intramolecular interactions and covalent bond changes, we have compared the protein-BMEDS k_{red} values to those for ACM and BMEDS (Wynn & Richards, 1993b). The value for the exchange reaction of these small molecules under the conditions used in these experiments is 2.04 and is independent of the urea concentration.

There is evidence that, under some conditions, nuclease will unfold through an intermediate that is populated at equilibrium (Carra et al., 1994). The thermodynamic cycle above would then be expanded to include three conformational states as follows:

$$\begin{array}{ccc} RSSR + P_N SH & \xleftrightarrow{K_{red,N}} & P_N SSR + RSH \\ K_{1,PSH} \updownarrow & & \updownarrow K_{1,PSSR} \\ RSSR + P_I SH & \xleftrightarrow{K_{red,I}} & P_I SSR + RSH \\ K_{2,PSH} \updownarrow & & \updownarrow K_{2,PSSR} \\ RSSR + P_U SH & \xleftrightarrow[K_{red,U}]{} & P_U SSR + RSH \end{array}$$

where $K_{1,x}$ are conformational equilibrium constants for the transition from the native state to the intermediate. $K_{2,x}$ describes the transition from the intermediate to the unfolded form. $K_{red,N}$, $k_{red,I}$, and $K_{red,U}$ describe the disulfide exchange equilibria for the native, intermediate, and unfolded forms, respectively. This scheme involves two cycles with the following relationships:

$$K_{red,N}/K_{red,I} = K_{1,PSH}/K_{1,PSSR}, \quad (4)$$

$$k_{red,I}/K_{red,U} = K_{2,PSH}/K_{2,PSSR}. \quad (5)$$

The intermediate and unfolded states have been shown to be indistinguishable using fluorescence or CD spectroscopy (Carra et al., 1994). Thus, if an intermediate is populated under the conditions of our experiments, the measured conformational equilibrium constants (K_{PSH} and K_{PSSR}) would describe unfolding of the native state to both intermediate and unfolded states.

$$K_{PSH} = ([P_USH] + [P_{ISH}])/[P_NSH] \text{ or}$$

$$K_{PSH} = \sum_{i=1-n} [P_iSH]/[P_NSH], \quad (6)$$

$$K_{PSSR} = ([P_USSR] + [P_ISSR])/[P_NSSR] \text{ or}$$

$$K_{PSSR} = \sum_{i=1-n} [P_iSSR]/[P_NSSR], \quad (7)$$

where the extended equations on the right refer to any number (n) of unfolded conformers from a single fully unfolded state, a single or multiple intermediates, or a large ensemble on nonnative states. Note that under this scheme:

$$K_{PSH} = (K_{1,PSH} \times K_{2,PSH}) + K_{1,PSH} \text{ or}$$

$$K_{PSH} = \sum_{J=1-n} \left(\prod_{i=1-j} K_{i,PSH} \right), \quad (8)$$

$$K_{PSSR} = (K_{1,PSSR} \times K_{2,PSSR}) + K_{1,PSSR} \text{ or}$$

$$K_{PSSR} = \sum_{J=1-n} \left(\prod_{i=1-j} K_{i,PSSR} \right). \quad (9)$$

The apparent disulfide exchange equilibrium constant would then be equal to:

$$K_{red,app} = ([P_NSSR] + [P_ISSR] + [P_USSR]) \times R/([P_NSH] + [P_ISH] + [P_USH]), \quad (10)$$

or in general

$$K_{red,app} = \left(P_NSSR + \sum P_iSSR\right)/\left(P_NSH + \sum P_iSH\right) \times R, \quad (11)$$

where R is equal to [BME]/[BMEDS]. Combining Equation 10 with Equations 6 and 7 gives

$$K_{red,app} = K_{red,N}(1 + K_{PSSR})/(1 + K_{PSH}). \quad (12)$$

The above equation is identical to that derived for the two-state cycle above. Thus, the value of $K_{red,N}$ is not dependent on whether a two- or three-state unfolding scheme is postulated. An apparent k_{red} for all nonnative states can be described by:

$$K_{red,\Sigma Ui} = ([P_USSR] + [P_ISSR]) \times R/([P_USH] + [P_ISH]), \quad (13)$$

or for multiple nonnative states:

$$K_{red,\Sigma Ui} = \left(\sum [P_iSSR]/\sum [P_iSH]\right) \times R. \quad (14)$$

The combination of Equation 13 with Equations 6 and 7 gives

$$K_{red,\Sigma Ui} = K_{red,N} \times K_{PSSR}/K_{PSH}. \quad (15)$$

Note that the expression for $K_{red,\Sigma Ui}$ is identical to that for $K_{red,U}$ in the two-state cycle scheme.

For simplicity, the following discussion is limited to three-state denaturation. Equation 13 is equivalent to:

$$K_{red,\Sigma Ui} = F_{U,PSH} K_{red,U} + F_{I,PSH} K_{red,I}, \quad (16)$$

where $F_{U,PSH} = [P_USH]/([P_USH] + [P_ISH])$ and $F_{I,PSH} = [P_ISH]/([P_USH] + [P_ISH])$. Thus, $K_{red,\Sigma Ui}$ will be a weighted average of $K_{red,U}$ and $k_{red,I}$ in the three-state scheme. $K_{red,\Sigma Ui}$ can be expressed in terms of $K_{red,U}$, $K_{2,PSSR}$, and $K_{2,PSH}$:

$$K_{red,\Sigma Ui} = \frac{K_{red,U}(K_{2,PSSR} + 1)K_{2,PSH}}{(K_{2,PSH} + 1)K_{2,PSSR}}. \quad (17)$$

Three assumptions appear reasonable when analyzing the $K_{red,\Sigma Ui}$ data.

1. The free energy for the transition form I to U will be linearly related to the urea concentration ($\Delta G^0_{2,x} = \Delta G^0_{2,x,H2O} - m_2[\text{urea}]$; $K_{2,x} = \exp\{-\Delta G^0_{2,x,H2O}/RT + m_2[\text{urea}]/RT\}$).
2. If all of the experimentally measured change in m_{den} is due to different amounts of intermediate formation, then it is necessary that $m_{2,PSH} = m_{2,PSSR}$.
3. At all urea concentrations $K_{red,U}s = 2.0$. At high urea concentrations, where I is likely to be completely converted to U, $K_{red,\Sigma Ui}$ is equal to 2.0

$K_{red,\Sigma Ui}$ can then be described as a function of three variables, $\Delta G^0_{2,PSSR,H2O}$, $\Delta G^0_{2,PSH,H2O}$, and m_2, by substituting $K_{2,x} = \exp\{-\Delta G^0_{2,x,H2O}/RT + m[\text{urea}]/RT\}$ into the above equation for $K_{red,\Sigma Ui}$. Curve fitting of the values for $K_{red,\Sigma Ui}$ at several urea concentrations will yield estimates for the free energy differences between the intermediate and unfolded forms. A value for $k_{red,I}$ may then be determined using Equation 5.

Acknowledgments

This work was supported by grant GM-22778 from the National Institute of General Medical Sciences (R.W. and F.M.R.) and the Howard Hughes Medical Institute (C.L.A. and R.O.F.). We thank Dr. Marc Jacobs for reading the manuscript and for helpful discussions.

References

Alexandrescu AT, Abeygunawardana C, Shortle D. 1994. Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: A heteronuclear NMR study. *Biochemistry 33*:1063–1072.

Brahms S, Brahms J. 1980. Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J Mol Biol 138*: 149–178.

Carra JH, Anderson EA, Privalov PL. 1994. Three-state thermodynamic analysis of the denaturation of staphylococcal nuclease mutants. *Biochemistry 33*:10842–10850.

Creighton TE. 1977. Effects of urea and guanidine-HCl on the folding and unfolding of pancreatic trypsin inhibitor. *J Mol Biol 113*:313–328.

Crestfield AM, Moore S, Stein WH. 1963. The preparation and enzymatic hydrolysis of reduced and *S*-carboxymethylated proteins. *J Biol Chem 238*:622–628.

Daopin S, Anderson DE, Baase WA, Dahlquist FW, Matthews BW. 1991. Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry 30*:11521–11529.

Dill KA. 1990. Dominant forces in protein folding. *Biochemistry 29*:7133–7155.

Dill KA, Shortle D. 1991. Denatured states of proteins *Annu Rev Biochem 60*:795–825.

Eisenberg D, McLachlan AD. 1986. Solvation energy in protein folding and binding. *Nature 319*:199–203.

Evans PA, Kautz RA, Fox RO, Dobson CM. 1989. A magnetization-transfer nuclear magnetic transfer resonance study of the folding of staphylococcal nuclease. *Biochemistry 28*:362–370.

Evans PA, Topping KD, Woolfson DN, Dobson CM. 1991 Hydrophobic clustering in nonnative states of a protein: Interpretation of chemical shifts in NMR spectra of denatured states of lysozyme. *Proteins Struct Funct Genet 9*:248–266.

Flanagan JM, Kataoka M, Shortle D, Engelman DM. 1992. Truncated staphylococcal nuclease is compact but disordered. *Proc Natl Acad Sci USA 89*:748–752.

Gittis AG, Stites WE, Lattman EE. 1993. The phase transition between a compact denatured state and a random coil state in staphylococcal nuclease is first-order. *J Mol Biol 232*:718–724.

Green SM, Meeker AK, Shortle D. 1992. Contributions of the polar, uncharged amino acids to the stability of staphylococcal nuclease: Evidence for mutational effects on the free energy of the denatured state. *Biochemistry 31*:5717–5728.

Hellinga HW, Wynn R, Richards FM. 1992. The hydrophobic core of *E. coli* thioredoxin shows a high tolerance to non-conservative single amino acid substitutions. *Biochemistry 31*:11203–11209.

Hopton JD, Swan CJ, Trimm DL. 1968. Liquid-phase oxidation of thiols to disulfides. *Adv Chem Ser 75*:216–224.

Hynes TR, Fox RO. 1991. The crystal structure of staphylococcal nuclease refined at 1.7 Å resolution. *Proteins Struct Funct Genet 10*:92–105.

Lu J, Baase WA, Muchmore MC, Dahlquist FW. 1992. Protein folding: Assignment of the energetic changes of reversible chemical modifications to the folded or unfolded states. *Biochemistry 31*:7765–7772.

Neri D, Billeter M, Wider G, Wuthrich K. 1992. NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science 257*:1559–1563.

Nozaki Y. 1972. The preparation of guanidine hydrochloride. *Methods Enzymol 26*:43–51.

Pace CN. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol 131*:266–279.

Pakula AA, Sauer RT. 1990. Reverse hydrophobic effects relieved by amino-acid substitutions at a protein surface. *Nature 344*:363–364.

Ponder JW, Richards FM. 1987. Tertiary templates for proteins: Use of packing criterion in the enumeration of allowed sequences for different structural classes. *J Mol Biol 193*:775–791.

Rashin AI, Honig B. 1986. On the environment of ionizable groups in globular proteins. *Biochemistry 173*:515–521.

Richards FM. 1977. Areas, volumes, packing and protein structures. *Annu Rev Biophys Bioeng 6*:151–176.

Richards FM, Lim WA. 1993. An analysis of packing in the protein folding problem. *Rev Biophys 26*:423–498.

Saunders AJ, Young GB, Pielak GJ. 1993. Polarity of disulfide bonds. *Protein Sci 2*:1183–1184.

Schellman JA. 1978. Solvent denaturation. *Biopolymers 17*:1305–1322.

Shortle D, Chan HS, Dill KA. 1992. Modeling the effects of mutations on the denatured states of proteins. *Protein Sci 1*:201–215.

Shortle D, Meeker AK. 1986. Mutant forms of staphylococcal nuclease with altered patterns of guanidine hydrochloride and urea denaturation. *Proteins Struct Funct Genet 1*:81–89.

Shortle D, Meeker AK. 1989. Residual structure in large fragments of staphylococcal nuclease: Effects of amino acid substitution. *Biochemistry 28*:936–944.

Shortle D, Stites WE, Meeker AK. 1990. Contributions of the large hydrophobic amino acids to the stability of staphylococcal nuclease. *Biochemistry 29*:8033–8041.

Stark GR. 1965. Reactions of cyanate with functional groups of proteins. III. Reactions with amino and carboxyl groups. *Biochemistry 4*:1030–1036.

Stites WE, Gettis AG, Lattman EE, Shortle D. 1991. In a staphylococcal nuclease mutant the side-chain of a lysine replacing valine 66 is fully buried in the hydrophobic core. *J Mol Biol 221*:7–14.

Wynn R, Richards FM. 1993a. Unnatural amino acid packing mutants of *Escherichia coli* thioredoxin produced by combined mutagenesis/chemical modification techniques. *Protein Sci 2*:395–403.

Wynn R, Richards FM. 1993b. Partitioning the effects of changes in a protein to the folded or unfolded forms by using a thermodynamic cycle: A change in *Escherichia coli* thioredoxin does not affect the unfolded state. *Biochemistry 32*:12922–12927.