PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Initial Studies of the Equilibrium Folding Pathway of Staphylococcal Nuclease

Yi Wang, Andrei T. Alexandrescu and David Shortle

Phil. Trans. R. Soc. Lond. B 1995 **348**, 27-34
doi: 10.1098/rstb.1995.0042

Initial studies of the equilibrium folding pathway of staphylococcal nuclease

YI WANG, ANDREI T. ALEXANDRESCU AND DAVID SHORTLE

The Department of Biological Chemistry, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, U.S.A.

SUMMARY

Spectroscopic methods were used to examine the sequential build up of structure in the denatured state of staphylococcal nuclease. The 'free energy distance' between the native and denatured states was manipulated by altering conditions in solution (for example altering urea or glycerol concentration) and by changing the amino acid sequences. Initial studies employed a fragment of nuclease, referred to as Δ131Δ, which lacks six structural residues from the amino terminus and one structural residue from the carboxy-terminus. Nuclear magnetic resonance analysis of this fragment in solution revealed a modest quantity of dynamic structure which is native-like in character. With the addition of urea, 12 new H_N peaks appeared in the ^{1}H–^{15}N correlation spectrum, presumably as a result of the breakdown of residual structure involving the first three beta strands. With the addition of glycerol, there was a rapid increase in the quantity of beta sheet structure detected by circular dichroism spectroscopy. At very high glycerol concentrations, an increase in helical structure became apparent. These data in addition to previously published results suggest that: (i) a beta-meander (strands β1-β2-β3) and the second alpha helix (α2) are among the most stable local structures; (ii) the five-strand beta-barrel forms in a reaction which does not require the presence of several other native substructures; and (iii) the last step on the equilibrium folding pathway may be the formation and packing of the carboxy terminal alpha helix (α3) to give the native state.

1. THEORETICAL FRAMEWORK

Contrasting the very rapid rate of spontaneous protein folding with the huge number of possible conformations that would have to be searched by a random folding mechanism, the need to invoke some sort of hierarchical mechanism for assembling native protein structure seems unavoidable (Levinthal 1968). Perhaps the simplest such mechanism would involve a series of clustering events that bring together successive segments of the protein chain as each segment assumes its correct, native-like structure. By proceeding through a series of intermediates which contain only native-like substructures, such a mechanism could efficiently funnel the chain into the native state, avoiding the necessity for breaking down or rearranging any non-native structures. Recent studies of denatured proteins, partly folded proteins and molten globules do in fact suggest that most of their structure is native-like (Matthews 1993; Shortle 1993). If proteins do fold by such a simple hierarchical mechanism, important goals for experimentalists will be to identify the various intermediate substructures formed as folding progresses towards the native state and to elucidate the chain–chain interactions which drive their formation and subsequent merger into higher order structures. Using the hypothetical folding diagram in figure 1, this goal can be restated in more concrete terms.

First, the individual steps in the clustering diagram, which involve combination of two substructures, must be defined and placed in hierarchical order. Second, the free energy change (ΔG) for each clustering step must be quantitated by measuring the relative concentrations of intermediates under specified sets of conditions. Finally, the physical chemical interactions that determine ΔG for each step must be defined.

Data of this type collected for several simple proteins would undoubtedly lead to many quantitative insights into the mechanisms by which amino acid sequence specifies protein structure. This framework for describing the folding of proteins focuses on the structural and free energy differences between intermediates, and not on the kinetics of their formation. As a result, time is not given the central role it plays in the kinetic approach to analysing the mechanism of protein folding. Instead of describing the pathway of structure formation as a function of time, such an equilibrium-based approach could be used to define an 'equilibrium folding pathway', one in which the appearance of structure is examined as a function of the free energy distance between a polypeptide chain blocked in its efforts to fold and the native state that it is attempting to reach. To vary this free energy distance, two distinctly different types of manipulation could be used to raise the free energy of the native state by controlled amounts relative to that of the denatured state.

Phil. Trans. R. Soc. Lond. B (1995) **348**, 27–34
Printed in Great Britain

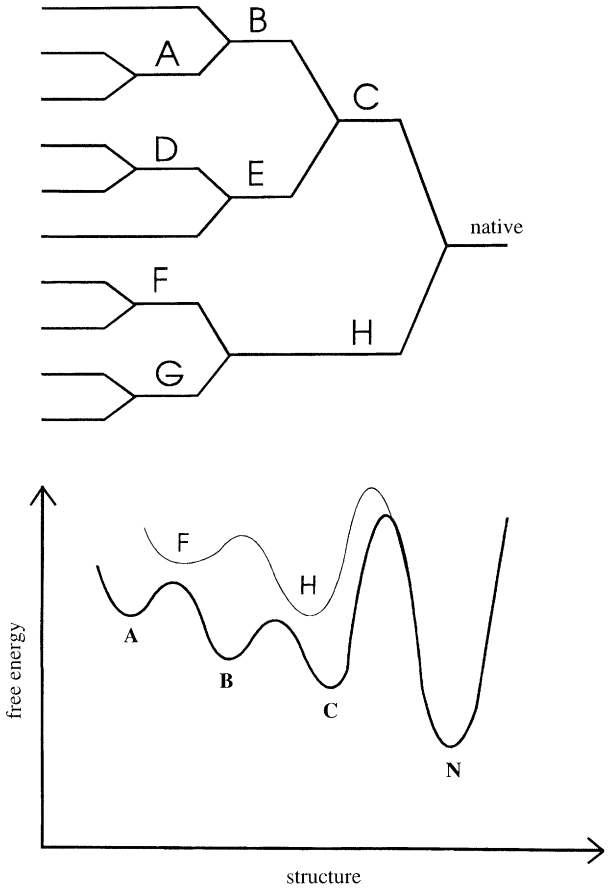


Figure 1. Top panel: hierarchical clustering diagram showing a series of hypothetical steps in the folding of a protein. A–H indicate folding intermediates, with increasing amounts of native-like structure proceeding from left to right. Bottom panel: free energy diagram of folding along two branches shown above involving several hypothetical intermediates.

First, disruption of native state interactions locally through changes in amino acid sequence increases the free energy of the native state with respect to the denatured state (Shortle 1992). Within the context of the model of structure assembly outlined in figure 1, a reasonable working assumption is that mutations will increase the free energy of only those partly folded intermediates that are stabilized by interactions which include the residue(s) at the site of the mutation (see figure 2*a*). In summary, formation of two different equilibrium intermediates will involve independent events unless they share a common substructure; intermediates that do not contain the interaction which has been disrupted by the mutation will not be affected. With this basis it should be possible to disrupt single branches on the hierarchical folding diagram, leaving other branches basically undisturbed, and by characterizing the structures of a large number of mutant denatured proteins (that have different sets of interactions disrupted) it may be possible to trace out the entire diagram. Alternatively, multiple native state interactions can be disrupted through changes in physical conditions of the protein solution, such as temperature, pH or solvent composition. When the free energy of the native state is increased by one of these variables, the effects on partly folded forms will be more universal, dependent on the physical chemistry underlying that variable (see figure 2*b*). For example, when the pH is lowered below 3–5, the high concentration of protons raises the free energy of protein structures through increasing the entropy cost of maintaining unprotonated side chains and by increasing electrostatic repulsion involving positively charged groups (Tanford 1970). This will specifically disrupt substructures in proportion to how many unprotonated sidechains they bury and/or how many charged groups are brought into positions where they repel each other. Alternatively, when denaturants such as urea and guanidine hydrochloride are added, the free energy of protein structures is increased in proportion to the amount of nonpolar surface area that becomes solvent inaccessible on formation of that structure (Tanford 1970; Timasheff 1992). Instead of attempting to follow the very rapid appearance of structure by fast methods after the initiation of protein folding, experiments within this framework can employ slower, but higher resolution physical methods such as nuclear magnetic resonance (NMR) spectroscopy. Although this strategy for analysing protein folding is based on several assumptions, it provides a framework for interpretting structural data obtained under a series of equilibrium conditions. To evaluate the overall merit of this approach, our laboratory has begun NMR characterization of the structure that persists in staphylococcal nuclease after it has been denatured (i.e. in a state where it is unable to fold completely to the native state) by mutations or by physical agents. These data will be integrated into a consistent model – an equilibrium folding pathway – based on the expectations that: (i) the structure in a majority of intermediates will be native-like; and (ii) clustering steps are independent unless they involve intermediates that share a common substructure. In this report, previous NMR studies of a large, denatured fragment of nuclease are extended by monitoring the structural consequences of changing the free energy distance with urea and the stabilizing solute glycerol.

2. DATA

(*a*) *Δ131Δ: A low density denatured state of staphylococcal nuclease*

A fragment of staphylococcal nuclease, Δ131Δ, which is 131 residues in length but missing one structural residue from the carboxy terminus, and six structural residues from the amino terminus has been estimated to be denatured more than 99% of the time at 32 °C and pH 5.3 (Alexandrescu *et al.* 1994). Yet, upon addition of Ca^{2+} and substrate DNA (or inhibitory nucleotides), this protein can readily refold to the native state with essentially normal catalytic activity. Based on the backbone and side chain assignments of 105 of the 131 residues plus additional NMR data, a low resolution model of Δ131Δ has been proposed (Alexandrescu *et al.* 1994).

Figure 3 shows the principal structural features identified, which include: (i) a significant population of Type I and I′ reverse turns at 83–86 and 94–97, respectively; (ii) approximately 30% population of the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

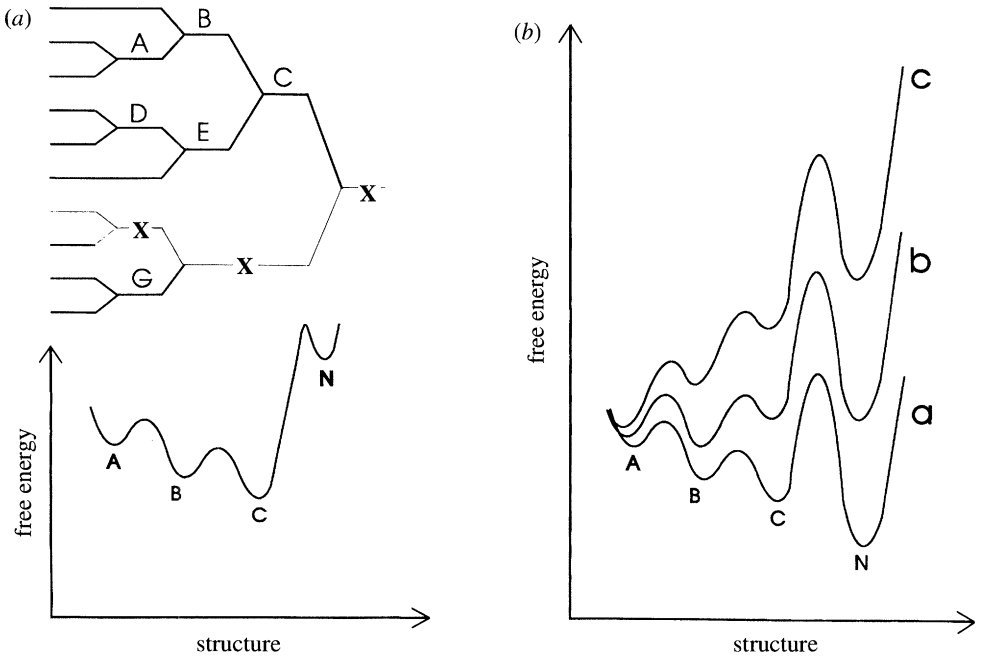


Figure 2. (*a*) Effects of mutating an amino acid residue that disrupts a chain–chain interaction present in intermediates F and H and in the native state. 'X' indicates that a clustering step in folding has been blocked by the mutation. Formation of intermediates A–E and G is unaffected by the mutation because these intermediates do not contain the disrupted interaction. (*b*) Effects on the free energy diagram of altering the solvent composition with denaturants (e.g. urea) and stabilizing solutes (e.g. glycerol). For denaturants/stabilizers, the free energy of a protein structure is raised/lowered, respectively, in proportion to the amount of solvent accessible surface area buried. Therefore the native state and intermediates with large amounts of structure will be destabilized/stabilized to a greater degree than intermediates with small amounts of structure by increasing the denaturant/stabilizer concentration in the order a–b–c/c–b–a, respectively.

second alpha helix; (iii) 10–20% population of alpha helix 1; and (iv) less than 10% population of alpha helix 3. Although no direct evidence of partial or transient beta strand formation has been found for any of the five strands that form the small Greek-key beta barrel in nuclease, 19 of the 24 H_N resonances corresponding to the first three beta strands are missing from the ^1H–^{15}N correlation spectrum (Alexandrescu *et al.* 1994). Because proton exchange rates under these solution conditions should be less than 1 s^{-1} for fully exposed amides (Molday *et al.* 1972) and as very strong resonances are observed for parts of the chain that have been shown to be solvent exposed (Alexandrescu & Shortle 1994), these resonances are unlikely to have been eliminated by proton exchange. Instead, the most probable explanation for severe broadening of a continuous segment of residues is conformational exchange on an intermediate timescale.

(*b*) *Δ131Δ*: *structure as a function of urea concentration*

To test the assumption that strands β1-β2-β3 are undergoing conformational exchange between an unfolded and a folded structure with very different values of chemical shift, the conditions of solution were made less favorable for structure formation by addition of a denaturant (see figure 2*b*). The ^1H–^{15}N spectrum of Δ131Δ was obtained at a series of urea concentrations from 0 M to 7 M in 1 M increments. Inspection of these spectra revealed several interesting features.

Those peaks present in the absence of urea exhibited only slight changes in their ^1H and ^{15}N chemical shifts between 0 M and 7 M urea. Whereas the large majority of peaks showed no significant change in intensity, 12 new peaks emerged, most of which first appeared at 3 or 4 M urea and increased in intensity at higher concentrations (see figure 4). One of these peaks falls in the range of H_N and ^{15}N chemical shifts that is diagnostic of a glycine residue in a random configuration. Because all glycines except Gly20 have been previously identified in Δ131Δ, this new peak can be assigned with confidence to Gly20, a residue which occupies the second position in a reverse turn connecting strands β1 and β2.

For five residues (Lys24, Leu25, Tyr27, Gly29 and Met32) in the peptide segment corresponding to β1-β2-β3, the H_N peaks have been assigned and in most cases these resonances are found to be significantly weaker than average. Unfortunately, except for Gly29, these peaks lie within crowded regions of the ^1H–^{15}N correlation spectrum and cannot be tracked with confidence between spectra collected at increasing urea concentrations to ascertain if they increase in intensity, as might be expected. In the case of Gly29, however, no significant increase in peak intensity was found, even at 7 M urea.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

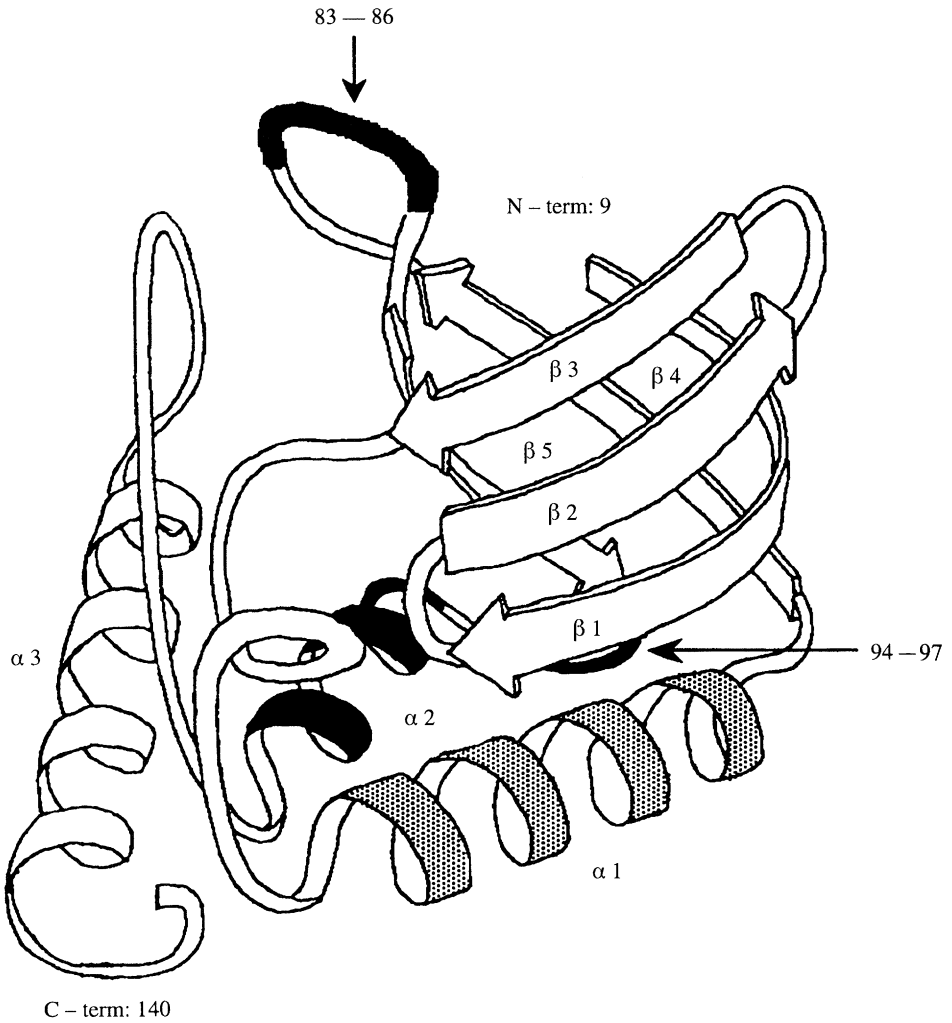


Figure 3. A ribbon diagram illustrating the substructures of native staphylococcal nuclease that have been detected in the large nuclease fragment Δ131Δ. The shaded regions indicate the elements of native-like residual structure (black, most persistent; grey, less persistent) based on secondary chemical shifts, NOEs and $^3J_{N_\alpha}$ coupling constants. Few of the residues in the first three beta strands β1-β2-β3 have been assigned. (From Alexandrescu *et al.* 1994).

(*c*) *Δ131Δ: structure as a function of glycerol concentration*

To study the response of Δ131Δ to conditions which are more favourable for folded structure (see figure 2*b*), the stabilizing solute glycerol was added to solutions of Δ131Δ at 32 °C and pH 5.3. As can be seen in figure 5*a*, large changes in the circular dichroism spectra suggestive of major increases in secondary structure occur between 0% and 70% glycerol (by volume). The fraction of molecules that have refolded to the native state can be monitored by the intrinsic fluorescence of tryptophan 140 (Shortle & Meeker 1986) with the result that less than 5% of molecules are in the native state at 30% glycerol, whereas approximately 15% have refolded in 70% glycerol (data not shown).

Further evidence that the additional structure induced at low concentrations of glycerol is not a consequence of complete folding to the native state is shown in figure 5*b*. These circular dichroism (CD) spectra of Δ131Δ were collected with increasing concentrations of the competitive inhibitor 3,5 diphosphothymidine (pdTp) in the presence of a fixed concentration of Ca^{2+}. The nucleotide plus one calcium ion bind to the active site of nuclease with very high affinity (Serpersu *et al.* 1986) and refold Δ131Δ by substrate stabilization of the native state. As can be seen, these spectra, which represent contributions from varying fractions of denatured and native states, look distinctly different from the spectra obtained in glycerol. These observations can be interpretted more quantitatively by calculating the difference spectra between Δ131Δ in the presence and absence of these two types of structure stabilizers. In figure 5*c*, the difference spectrum for 30% glycerol minus 0% glycerol displays the characteristic minimum at 215 nm typically seen for pure beta sheet structures (Johnson 1988), whereas the difference spectrum for 70% glycerol minus 30% glycerol has the minima at 222 nm and 208 nm diagnostic of alpha helix. In contrast, the difference spectra obtained as a function of pdTp are distinctly different from these two, exhibiting a strong minimum at 222 nm without a pronounced minimum at 208 nm. This suggests that alpha-helix and beta sheet are forming simultaneously rather than consecutively, as expected for one-step refolding to the native state. The 30%–0% difference spectrum sug-

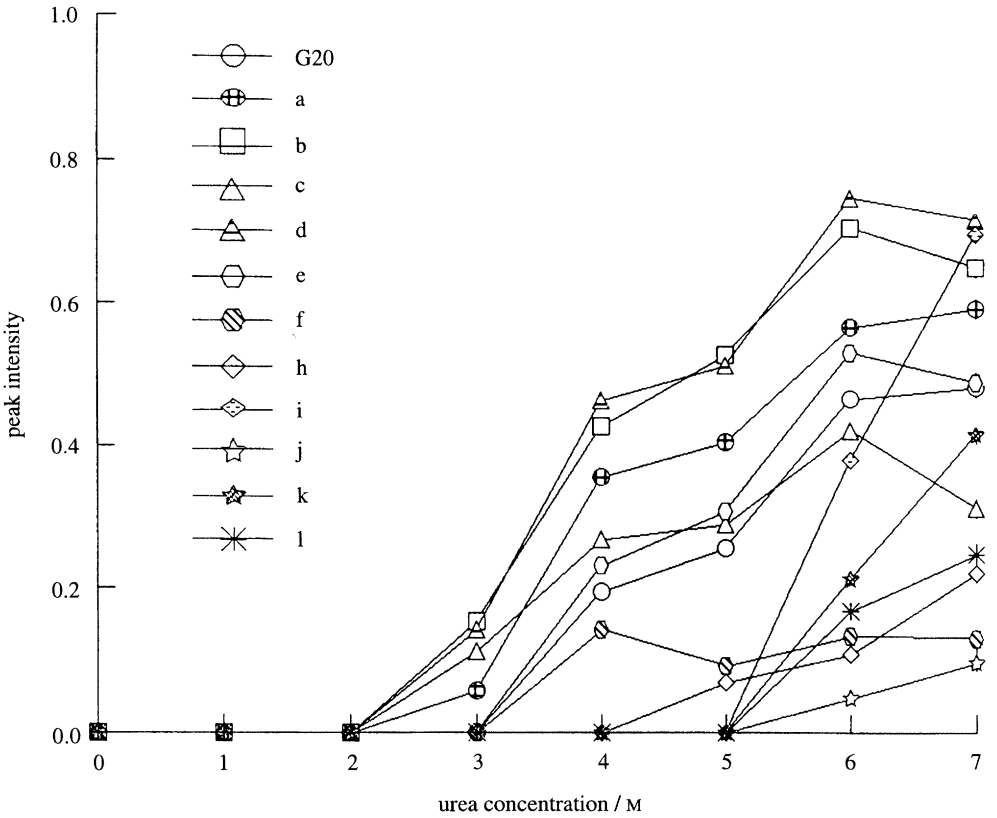


Figure 4. The relative intensity of twelve 'new' H_N resonances of Δ131Δ resolved in a series of ^1H–^{15}N–HSQC (Messerle *et al.* 1989) spectra collected at different concentrations of urea at pH 5.3 and 32 °C. Relative intensity is the measured peak height normalized to the height of the H_N of Serine 128. Peaks that were not present in 0 M urea and that were clearly resolved from all other peaks are shown. Only one of these, Glycine 20, has been assigned.

gests the transition at lower glycerol concentrations predominantly involves the formation of beta sheet. When the assigned H_N resonances are followed in the ^1H–^{15}N correlation spectrum as a function of glycerol, all but one (Ile92) of the peaks retain essentially the same position in the spectrum from 0% to 30% glycerol. However, all peaks undergo significant reductions in intensity, albeit at very different rates.

As seen in figure 6, the H_N resonances from beta strands β4 and β5 diminish in intensity most rapidly with increasing glycerol. The H_N resonances from the third alpha helix α3 and the catalytic loop (residues 43–54; data not shown) decrease by a relatively small amount. Finally, the H_N resonances from the first and second alpha helix α1 and α2 display a behaviour which is intermediate between these two extremes. It is expected that concentrations of glycerol above 5–10% significantly increase the solution viscosity, which will increase the rotational correlation time and thereby reduce T_2 for both H_N and ^{15}N directly via effects on dipolar–dipolar and chemical shift anisotropy relaxation mechanisms (Abragam 1961; Wagner 1993).

However, when such direct broadening effects are measured on wild-type nuclease, they are found to be relatively small. For example, quantitation of the peak intensities of 23 well resolved H_N peaks of native nuclease (molecular mass 17000) in the presence and absence of 30% glycerol (at 32 °C, pH 5.3) revealed an average reduction of peak intensity to 0.64 of the initial value (data not shown). Because the apparent rotational correlation times of residues in Δ131Δ have been demonstrated to be significantly shorter than that of wild-type nuclease (Alexandrescu & Shortle 1994), a 36% decrease in peak intensity would appear to represent an upper estimate of the direct effect of glycerol on T_2 and the observed line widths.

As mentioned above, only slight changes in ^1H and ^{15}N chemical shifts were observed for the H_N resonances in Δ131Δ (except Ile92) as the glycerol was increased. From this observation alone, it can be concluded that: (i) conformational exchange is unlikely to be fast on the chemical shift time scale; and (ii) the observed chemical shifts at all glycerol concentrations are those of an unstructured state which is more highly populated than the state with which it exchanges. Given the extensive broadening of so many residues, the H_N chemical shifts in the less populated, more structured state state must be significantly different for essentially all residues in Δ131Δ.

3. CONCLUSIONS

The most interesting observation reported here is that, upon addition of the stabilizing solute glycerol, the polypeptide chain of Δ131Δ undergoes a major transition to a structure with less alpha helical content than the native state. A prominent feature of this transition appears to be very extensive exchange broadening of the H_N resonances, suggesting that the final structure is native-like. This conclusion is based on the fact that the chemical shift environments of many resonances are distinctly different from their

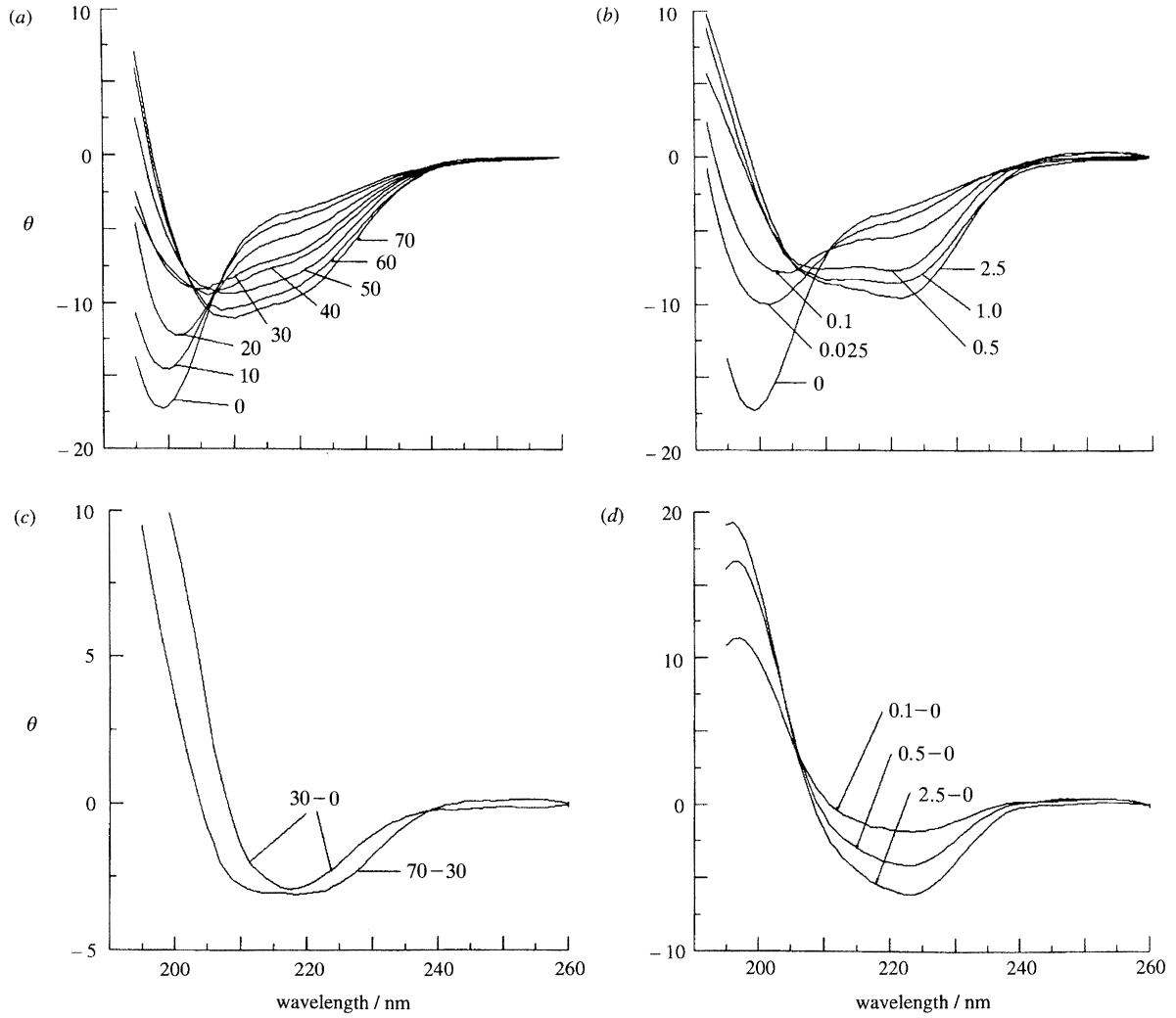


Figure 5. Far ultraviolet circular dichroism spectra of Δ131Δ. The *x* axis is wavelength of ultraviolet light in nm, the *y* axis is the mean residue molar ellipticity ($\times 10^{-3}$) (θ). (*a*) As a function of glycerol concentration, each curve is labelled with the % glycerol. (*b*) As a function of 3,5 diphosphothymidine concentration (pdTp) at pH 6.5 and 5 mM $CaCl_2$, after subtraction of a blank spectrum collected with the same concentrations of pdTp and Ca^{2+} but without protein, each curve is labelled with the pdTp concentration in mM. (*c*) Calculated difference spectra from part (*a*), 30–0 is the difference between 30% and 0% glycerol; 70–30 is the difference between 70% and 30% glycerol. (*d*) Calculated difference spectra from panel (*b*) 0.1–0 (0.5–0, 2.5–0) represent the difference between 0.1 (0.5, 2.5) mM pdTp and 0 mM pdTp.

random, highly averaged values. Because this transition is rapidly promoted by increasing concentration of glycerol and because glycerol affects conformational transitions in proportion to the amount of nonpolar surface area that undergoes burial (Gekko *et al.* 1981), the structure formed in this transition must involve the burial of a large number of hydrophobic residues. At concentrations of glycerol less than 35%, this state lacks much of the alpha helical structure found in the native state. As the glycerol concentration is raised even further, there is a specific increase in alpha helical structure. From these observations, the simplest conclusion is that the five-strand beta barrel structure of nuclease forms initially in the absence of some of the alpha helices. At higher glycerol concentrations, these helices become increasingly populated and pack against the beta barrel to give the native or a near-native structure. Previous NMR analysis of Δ131Δ in the absence of glycerol has revealed no evidence for beta structure formation by chain segments corresponding to beta strands β4 and β5 (Alexandrescu *et al.* 1994). However, the H_N resonances of all of the residues in β1 and over half of the residues in β2 and β3 are missing from the ^1H–^{15}N correlation spectrum. In view of the relatively slow (approximately 1 s^{-1}) proton exchange rate predicted at the temperature and pH used, the most likely explanation for this extensive broadening is intermediate exchange between an unstructured state and a structured state, with the beta meander formed by β1-β2-β3 as the obvious candidate for the structured state. Consistent with such a broadening mechanism is the appearance of 12 new peaks upon adding the denaturant urea to concentrations of 3 M and above. One of these peaks can be assigned to Gly20, a residue which plays a central role in this beta meander. Broadening of the H_N resonances corresponding to the third alpha helix α3 only becomes signficant at high glycerol concentrations. Because the decrease in peak intensity of these residues is approximately the same as is observed for native WT nuclease

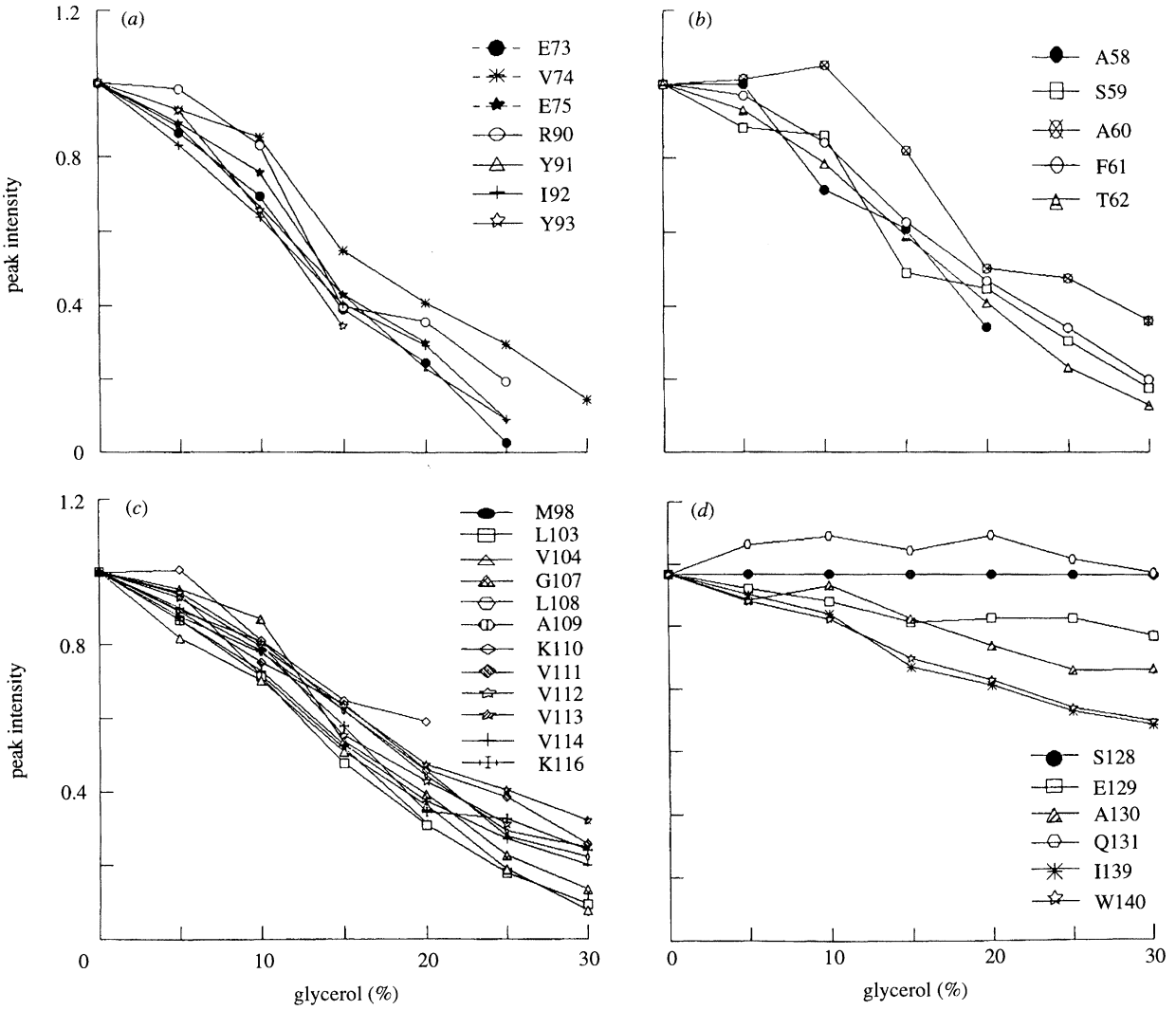
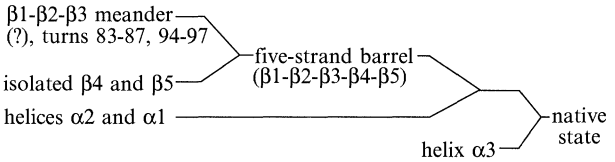


Figure 6. The relative intensity of the H_N resonances of Δ131Δ resolved in a series of ^1H–^{15}N–HSQC spectra (Messerle *et al.* 1989) collected a different glycerol concentrations at pH 5.3 and 32 °C. Relative intensity is the measured peak height normalized to the height of the H_N of Serine 128. (*a*) Residues corresponding to beta strands β4–β5. (*b*) Residues corresponding to helix α1. (*c*) Residues corresponding to helix α2 plus an extended segment. (*d*) Residues corresponding to helix α3.

in glycerol, the observed broadening could be a consequence of increasing solution viscosity rather than formation of an alpha helix. Previous ^{15}N relaxation studies clearly indicated the residues that comprise this chain segment are among the most mobile and unconstrained in Δ131Δ (Alexandrescu & Shortle 1994). The disinclination of this structure to form was also noted in a structural analysis of a quite different denatured form of staphylococcal nuclease, involving a large fragment missing five structural residues from the carboxy terminus and none from the amino terminus (Shortle & Abeygunawardana 1993). In this case, the five strand beta barrel was demonstrated to be intact, yet the α3 segment displayed extremely sharp backbone resonances with random-coil chemical shifts, suggesting a very high mobility. Thus, several lines of evidence place the packing of helix α3 against the remainder of the molecule as one of the last steps, if not the final step, in the assembly of the wild-type native structure. The results presented here, coupled with previous solution studies of Δ131Δ (Alexandrescu *et al.* 1994; Alexandrescu & Shortle 1994) allow a tentative, partial equilibrium folding pathway to be defined for staphylococcal nuclease, as shown below.



Future studies of additional mutant forms of staphylococcal nuclease of normal length should allow a more complete dissection of this hierarchy of clustering steps plus an estimation of the various values of ΔG involved.

We thank Joel Gillespie for help with the CD spectroscopy. This work was supported by NIH Grant GM34171 to D.S.

REFERENCES

Abragam. A. 1961 *The principles of nuclear magnetism.* London: Oxford University Press.

Alexandrescu, A.T. & Shortle, D. 1994 Backbone dynamics

of a highly disordered 131 residue fragment of staphylococcal nuclease. *J. molec. Biol.* **242**, 527–546.

Alexandrescu, A.T., 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.

Gekko, K. & Timasheff, S.N. 1981 Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry* **20**, 4667–4676.

Johnson, W.C. Jr 1988 Secondary structure of proteins through circular dichroism spectroscopy. *A. Rev. biophys. Biophys. Chem.* **17**, 145–166.

Levinthal, C. 1968 Are there pathways for protein folding? *J. chim. Phys.* **65**, 44–45.

Matthews, C.R. 1993 Pathways of protein folding. *A. Rev. Biochem.* **62**, 653–683.

Messerle, B.A., Wider, G., Otting, G., Weber, C. & Wuthrich, K. 1989 Solvent suppression using a spin-lock in 2D and 3D NMR spectroscopy with H_2O solutions. *J. magn. Reson.* **85**, 608–613.

Molday, R.S., Englander, S.W. & Kallen, R.G. 1972 Primary structure effects on peptide group hydrogen exchange. *Biochemistry* **11**, 150–159.

Serpersu, E.H., Shortle, D. & Mildvan, A.S. 1986 Kinetic and magnetic resonance studies of effects of genetic substitution of a Ca^{2+}-liganding amino acid in staphylococcal nuclease. *Biochemistry* **25**, 68–77.

Shortle, D. 1992 Mutational studies of protein structures and their stabilities. *Q. Rev. Biophys.* **25**, 205–250.

Shortle, D. 1993 Denatured states of proteins and their roles in folding and stability. *Curr. Opin. struct. Biol.* **3**, 66–74.

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

Shortle, D. & Abeygunawardana, C. 1993 NMR analysis of the residual structure in the denatured state of an unusual mutant of staphylococcal nuclease. *Structure* **1**, 121–134.

Tanford, C. 1970 Protein denaturation. Part C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.* **24**, 1–95.

Timasheff, S.N. 1992 Water as ligand: preferential binding and exclusion of denaturants in protein unfolding. *Biochemistry* **31**, 9857–9864.

Wagner, G. 1993 NMR relaxation and protein mobility. *Curr. Opin. struct. Biol.* **3**, 748–754.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES