Mutations Can Cause Large Changes in the Conformation of a Denatured Protein†

John M. Flanagan, Mikio Kataoka,‡ Tetsuro Fujisawa,§ and Donald M. Engelman*

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511

Received March 23, 1993; Revised Manuscript Received July 8, 1993⊕

ABSTRACT: Deletion of 13 amino acids from the carboxyl terminus of staphylococcal nuclease (WT SNaseΔ) results in a denatured, partially unfolded molecule that lacks significant persistent secondary structure but is relatively compact and monomeric under physiological conditions [Shortle & Meeker (1989) *Biochemistry 28*, 936–944; Flanagan *et al.* (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 748–752]. Because of these and other properties of the SNaseΔ polypeptide, it is a useful model system for investigating the conformation of the denatured state of a protein without using extreme temperature or solvent conditions. Moreover, since the modification is a carboxyl-terminal deletion, SNaseΔ may also resemble a transient state of the polypeptide chain as it emerges from a ribosome prior to its folding. In the present study, we have examined the sizes and conformations of mutated forms of SNaseΔ, using small-angle X-ray scattering and circular dichroism spectroscopy. Seven mutated forms were studied: four with single substitutions, two with double substitutions, and one triple substitution. When present in the full-length SNase, each of these mutated forms exhibited unusual behavior upon solvent or thermal denaturation. In the case of the truncated form (SNaseΔ), the small-angle scattering curves of the mutated forms fall into two classes: one resembling the scattering curve of compact native nuclease and the other having features consistent with those expected for an expanded coil-like polymer. In contrast, the scattering curve of WT SNaseΔ exhibits features intermediate between those observed for globular proteins and random polymers. The amino acid substitutions that gave rise to compact, native-like versions of SNaseΔ were all of the m– type (m– substitutions are predicted to decrease the size of the denatured state). Those which gave rise to versions of SNaseΔ that were more extended and coil-like than WT SNaseΔ were of the m+ type (m+ substitutions are predicted to increase the size of the denatured state). Estimates of the residual secondary structure present in WT SNaseΔ, as well as both the m+ and m– substituted versions of SNaseΔ, as determined by CD, suggest that the formation of secondary structure and compaction of the polypeptide chain occur concurrently. Our results show that single amino acid substitutions can radically alter the conformational distribution of a partially condensed polypeptide chain. The correspondence between changes in R_g produced by amino acid substitutions in SNaseΔ and the size of the denatured state predicted on the basis of solvent denaturation studies for these mutated forms supports the view that some amino acid substitutions can affect stability by changing the average conformation of the denatured state. In addition, our results are consistent with the idea that secondary structure formation and chain condensation occur simultaneously in this system.

Staphylococcal nuclease has been the subject of numerous folding studies since it is small (149 amino acids), lacks a prosthetic group, lacks cysteine, and exhibits highly reversible folding *in vitro*. Solvent denaturation studies, performed principally in the laboratory of David Shortle (Shortle 1983, Shortle & Meeker, 1986, 1989; Shortle *et al.*, 1988, 1990; Sondek & Shortle, 1990), have suggested that some mutations that affect the stability of the protein may do so, at least in part, by altering the distribution of denatured states. However, the use of high concentrations of denaturants, such as urea or guanidine hydrochloride (GuHCl), has complicated interpretation of these studies as well as attempts to characterize the denatured state. We have examined a truncated form of staphylococcal nuclease (SNaseΔ), which lacks the carboxyl-terminal 13 amino acid residues. The 136 amino acid protein is destabilized, and is denatured at room temperature and pH 7.0 in the absence of any added denaturant (Shortle & Meeker, 1989; Flanagan *et al.*, 1992). It has been shown to lack significant persistent secondary structure in aqueous buffer even though it is compact compared to the denatured state produced by high concentrations of urea or GuHCl denaturation (Shortle & Meeker, 1989; Flanagan *et al.*, 1992).

Here we describe solution X-ray scattering studies that indicate alterations in the conformation of SNaseΔ, at neutral pH and in the absence of denaturant, as a result of single amino acid substitutions. The mutations examined have been classified as either m+ or m– depending upon their effects upon the solvent denaturation of SNase. This nomenclature is derived from the slope of the dependence of the apparent equilibrium constant between the native and denatured states [$\ln(k_{app})$] on denaturant concentration. For SNase, it is linear: $\ln(k_{app}) = m_{den}[\text{denaturant}] + c$, where m_{den} is the slope of this line and is proportional to the rate of change of ΔG with denaturant concentration. Substitutions of the m+ type result in an increase in the value of m_{den}, while m– substitutions result in a decrease. The physical significance of m_{den} is not completely clear, though at least one model suggests that it is related to differences in the amount of denaturant that interacts with the native and denatured states of the polypeptide chain. Thus, m_{den} may reflect the difference

† This research was supported by grants from the NIH (GM22778, GM39546), the NSF (DMB8805587), and the National Foundation for Cancer Research.

* To whom correspondence should be addressed at the Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Ave., P.O. Box 6666, New Haven, CT 06511. Telephone: 203-432-560. FAX: 203-432-5175.

‡ Present address: Department of Biology, Faculty of Science, Osaka University, Osaka, Japan.

§ Present address: RIKEN, The Institute of Physical and Chemical Research, Saitama, Japan.

⊕ Abstract published in *Advance ACS Abstracts,* September 1, 1993.

0006-2960/93/0432-10359$04.00/0

between the accessible surface areas of these two states (Schellman, 1978). Using this model, Shortle and Meeker (1986) have argued that m+ substitutions increase the solvent-accessible surface area (or equivalently the size of the denatured state) while m– substitutions decrease the solvent-accessible surface area (or size of the denatured state) compared to that for the wild-type sequence. Our studies show dramatic changes in the radius of gyration (R_g) of SNaseΔ due to amino acid substitutions; m+ SNaseΔ polypeptides are larger than WT SNaseΔ, while m–SNaseΔ polypeptides are nearly as compact as native folded nuclease. The observed changes in the R_g resulting from m+/– substitutions in SNaseΔ are consistent with the changes in denatured state size predicted from their effects upon m_{den}. The various conformational states accessible to the SNaseΔ may represent the cellular state/states of the polypeptide emerging from the ribosome as it is recognized by the folding machinery (Langer *et al.*, 1992).

MATERIALS AND METHODS

Production and Purification of Various Forms of SNase and SNaseΔ. Wild-type staphylococcal nuclease from the Foggi strain of *Staphylococcus aureus* (WT SNase) and each of the substituted full length SNases were purified from *Escherichia coli* strain MGT7 (Supplied by D. LeMaster) containing the plasmid pSNWT11a or an equivalent plasmid containing substituted versions of SNase or truncated forms. These plasmids are derivatives of pET11a (Studier *et al.*, 1990) and contain the entire coding sequence of nuclease cloned into the *Nde*I and *Bam*HI sites of this vector (J. M. Flanagan, unpublished results). Large-scale growths for expression of SNase were conducted as described in Lemmon *et al.* (1992). The cells were harvested by centrifugation 3 h after isopropyl β-thiogalactoside (IPTG)[1] induction and then were resuspended in $^1/_{20}$th culture volume of 100 mM Tris·HCl, pH 8.0, 0.1 mM EDTA, and 1 mM PMSF. Cell lysis was accomplished by three rounds of probe sonication on ice; each round consisted of 1 min of sonication at the highest power setting followed by 1 min off. To promote the hydrolysis of DNA, $CaCl_2$ was added to a final concentration of 10 mM, and the solution was then incubated on ice for 15 min. The solution was clarified by centrifugation in a Beckman 45Ti type rotor at 40 000 rpm for 60 min, and then the supernatant was passed over a DEAE cellulose column (the volume of this column was at least 20 mL per liter of original culture) equilibrated with 50 mM Tris·HCl, pH 7.5, 0.1 mM EDTA, and 10% glycerol. The DEAE column was washed with 1.5 column volumes of equilibration buffer; these fractions contained the bulk of the SNase activity. The pooled DEAE flow through and wash were loaded onto a strong cation-exchange column [Fractogel SO_3(M) EM Science] equilibrated with 50 mM Tris·HCl, pH 7.5, 0.1 mM EDTA, and 100 mM NaCl (at least 10 mL of resin per liter of culture), and then the column was washed with 5 column volumes of 50 mM Tris·HCl, pH 7.5, 0.1 mM EDTA, 250 mM NaCl, and 0.1mM PMSF. SNase was eluted in a step gradient consisting of 50 mM Tris·HCl, pH 7.5, 0.1 mM EDTA, and 1.0 M NaCl. The SNase-containing fractions were pooled and concentrated to approximately 20 mg/mL by ultrafiltration in an Amicon flow cell with YM10 membranes. The concentrated SNase solution was dialyzed against two changes of 2 L each, of 50 mM MOPS, pH 6.8, 0.1 mM EDTA, and 100 mM NaCl, after which it was further purified by size-exclusion chromatography (SEC) on a 600 × 21.5 mm TSK-250 (Bio-Rad) column equilibrated with 50 mM MOPS, pH 6.8, 0.1 mM EDTA, and 100 mM NaCl. The SEC column was developed in equilibration buffer at a flow rate of 2.0 mL/min. The majority of SNase eluted in a volume consistent with a molecular mass of approximately 17 kDa. The fractions corresponding to monomeric nuclease were pooled, concentrated, and then dialyzed exhaustively against 10 mM TES, pH 7.5 and 10 mM NaCl. In some instances, the protein was subjected to an additional chromatographic step on a hydroxyapatite column (HAP). SNase could be eluted from the HAP column in a linear gradient from 10 to 400 mM $NaKPO_4$, pH 7.5. This additional chromatographic step eliminated a very minor contaminant (~0.1%), which in most cases (CD or X-ray analysis) had no effect upon subsequent experiments. Prior to long term storage at –80 °C, SNase was concentrated to 25 mg/mL and then dialyzed exhaustively against 10 mM TES, pH 7.5, and 10 mM NaCl. Purification of WT SNaseΔ and SNaseΔ containing various amino acid substitutions was accomplished as described in Flanagan *et al.* (1992).

Circular Dichroism Measurements. CD spectra were collected on an Aviv Associates (Lakewood, NJ) 60DS spectropolarimeter using a Hellma cuvette made of Suprasil quartz with a 0.1 mm path length. Spectra were collected in 10 mM TES, pH 7.5, and 10 mM NaCl. The final spectra used were the average of seven scans collected at 0.5 nm intervals (1 s signal integration time; 1.5 nm bandwidth), corrected with the appropriate base-line scan, and smoothed using a third-order polynomial with a ±2 point averaging algorithm. In all experiments, the temperature was maintained at 20 ± 0.2 °C with a water-jacketed cuvette holder. In order to calculate the molar ellipticity, the concentration of nuclease samples containing a single tryptophan was determined from the absorbance at 280 nm using 0.93 as the absorbance of a 1 mg/mL solution of wild-type nuclease. The concentration of each of the SNaseΔ samples was determined by quantitative amino acid analysis.

Sample Preparation and Small-Angle Scattering Measurements (SAXS). Concentrated stock solutions (25–30 mg/mL) of wild type, WT SNaseΔ, and the various substituted versions of SNaseΔ were dialyzed at room temperature against 2 L of 10 mM NaCl/10 mM TES, pH 7.5, overnight. Urea denatured nuclease was prepared by addition of 10 M urea to a 50 mg/mL nuclease-containing solution to a final concentration of 8 M urea. In experiments where the concentration of NaCl was varied, the concentration of NaCl in both the sample and buffer was adjusted by dilution from a 5 M stock solution immediately prior to data collection. The pH of the solution was varied as required by dialysis against the appropriate buffer. SAXS data were collected at the Photon Factory (PF) in Japan (Ueki *et al.*, 1985), with beam line 12B at the National Synchrotron Light Source (NSLS), or at Yale University on a small-angle scattering station as previously described (Kataoka *et al.*, 1989). The correction for the dependence of the scattering curves upon protein concentration in order to obtain the scattering curve at infinite dilution has been described (Kataoka *et al.*, 1989). In all cases, the final dialysate was used for buffer subtraction; for the samples containing urea or NaCl, solutions prepared by dilution into the appropriate buffer were used.

The measurement time for all SAXS experiments was 15 min. The sample holders at Yale and the Photon Factory

[1] Abbreviations: TES, *N*-[tris(hydroxymethyl)methyl]-2-aminoethanesulfonic acid; MOPS, 3-(*N*-morpolino)propanesulfonic acid; PMSF, phenylmethanesulfonyl fluoride; EDTA, ethylenediaminetetraacetic acid; IPTG, isopropyl β-thiogalactoside; HPLC, high-pressure liquid chromatography; SDS, sodium dodecyl sulfate.

were brass, while at NSLS they were plexiglass. The windows for the sample holders at Yale and NSLS were kaptan, and at the Photon Factory they were quartz. The temperature of the samples was maintained at the indicated temperatures by means of a water jacketed sample holder and a circulating water bath.

Analysis of Small-Angle X-ray Scattering. The scattering data are presented as the relative scattered intensity $I(Q)$, where the momentum transfer, or scattering vector Q, is calculated as

$$Q = 4\pi \sin\theta/\lambda$$

where λ is the wavelength of X-ray radiation and 2θ is the scattering angle. Following point-by-point extrapolation to zero protein concentration at each Q, Guinier analysis (Guinier & Fournet, 1955) was performed on the resulting scattering curve to obtain an estimate of the radius of gyration, R_g. Both the linearity of the resultant Guinier plots and the scattered intensity at zero angle, $I(0)$ [the value of $I(0)$ is proportional to the molecular weight of the protein], were used to assess the aggregation state of each sample. We have also analyzed the scattering data by use of the pair-distance distribution function, $P(r)$. This function may be calculated from the scattering data collected over a wider Q range. The $P(r)$ functions were calculated using the real space algorithms of both Glatter (1980) and Svergun (1992) or according to

$$P(r) = \left(\frac{1}{2\pi^2}\right) \int I(Q)(Qr) \sin(Qr)\, dQ$$

using the indirect Fourier inversion algorithm as described by Moore (1980). The version of the Moore algorithm employed in this study extends the data to infinite scattering angle by means of Porod's Law to minimize the effects of series truncation (Moore, 1980). In all cases, implementations of these algorithms employed slit-smearing corrections. The R_g of the particle was also estimated from the second moment of the $P(r)$ function by the equation

$$R_g^2 = \frac{\int r^2 P(r)\, dr}{2\int P(r)\, dr}$$

RESULTS AND DISCUSSION

The Effects of Amino Acid Substitution and Urea Denaturation on the X-ray Scattering Curve of Full-Length SNase. Solution scattering can provide several useful pieces of information about a protein in solution [for a review of small-angle scattering, see Feigin and Svergun (1987)]. Guinier analysis, based on the Gaussian shape of the scattering curve near zero angle, yields the radius of gyration of the particle (R_g), and extrapolation of the scattered intensity to zero angle gives a value related to the molecular weight. Further, Fourier inversion of a more extended region of the scattering curve gives the pair-distance distribution function, P(r), which represents the length distribution of all vectors relating scattering elements in the particle. In all cases, it is important that the scattering data be collected over a range of concentrations and extrapolated to an ideal scattering curve at infinite dilution.

Figure 1 shows a Guinier plot of the X-ray scattering data extrapolated to zero protein concentration for WT SNase in the presence and absence of 8 M urea. For comparison, the scattering curves have been normalized such that they have the same intensity at zero angle. Under both of these solvent conditions, it is clear that there is a linear Guinier region at

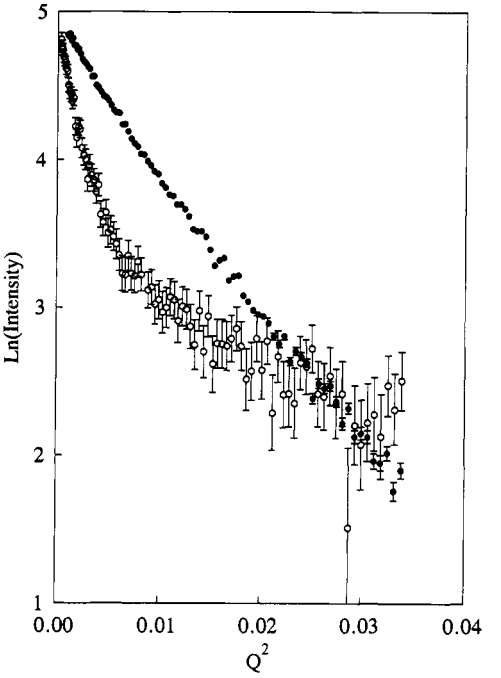

FIGURE 1: Comparison of SAXS curves of native and urea-denatured SNase. The Guinier region of the SAXS curves of WT SNase (10 mg/mL) either in 10 mM TES (pH 7.5)/10 mM NaCl (●) or in 8 M urea, 10 mM TES, pH 7.5, and 10 mM NaCl (○) buffer are shown. The data were collected at 25 °C using a synchrotron source (Photon Factory, Japan, or at NSLS).

sufficiently small angles, from which the R_g of the protein may be obtained. The slopes of the linear portion of these curves correspond to R_g values of 15.9 ± 0.1 and ~33 ± 1 Å for native and urea denatured SNase, respectively. The molecular mass of native SNase estimated from the value of $I(0)$ obtained by linear extrapolation of the Guinier region to zero scattering angle was 16.5 ± 1kDa, as expected for a monomer of SNase. The $I(0)$ value for SNase in 8 M urea was significantly smaller (not apparent in Figure 1 because of normalized intensities) than for native SNase and reflects the decrease in contrast between the solvent and protein molecule resulting from the high concentrations of urea employed in this experiment. We were therefore unable to use the value of $I(0)$ to estimate directly the molecular mass of SNase in 8 M urea; however, the linearity of the Guinier region, as well as the absence of any effect of protein concentration on the shape of the scattering curve, argues that SNase is monomeric in this solvent environment. In addition to native WT SNase, we have also measured the scattering curves of four m– (see the introduction) substituted versions of the folded full-length protein (V66L, G88V, V66L+G88V, and V66L+G79S+G88V) and three m+ substituted versions (A69T, L25A, and A69T+A90S). In none of these cases was any significant difference in the scattering profiles seen compared with those obtained for the wild-type sequence (data not shown).

The observed R_g of SNase in 8 M urea is significantly smaller (~33 Å) than that predicted for a freely jointed random coil of 149 amino acid residues (~48 Å) (Miller & Gobel, 1968), and may reflect residual interactions between amino acids along the chain that persist under these solvent conditions. This interpretation is supported by a recent NMR study of the first 63 residues of the phage 434 repressor in 8 M urea, in which the existence of a hydrophobic core that is stable in 8 M urea was demonstrated (Neri *et al.*, 1992). Further, theoretical calculations suggest that 8 M urea is far from a

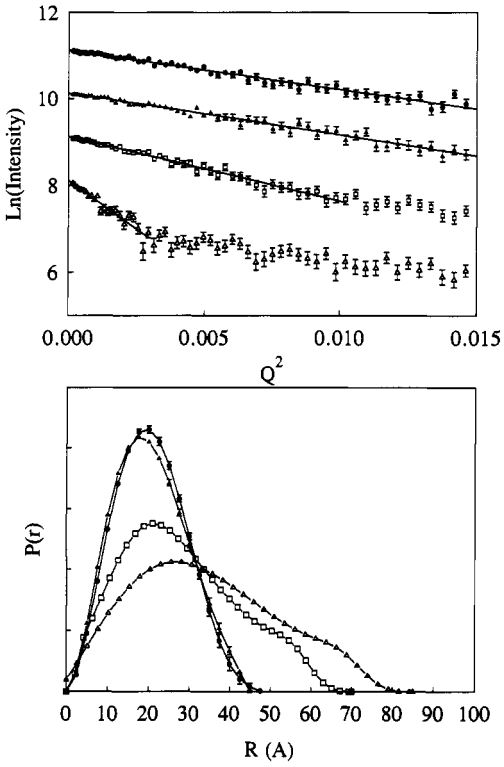

FIGURE 2: Effects of amino acid substitutions of the R_g vs $P(r)$ profiles of SNaseΔ. The Guinier regions of the SAXS curves of WT (●), WT SNaseΔ (□), V66L SNaseΔ (m–) (▲), and A69T SNaseΔ (m+) (△) in 10 mM TES/10 mM NaCl at 25 °C are shown (upper panel). The real-space equivalent of the scattering curve [$P(r)$ profile] calculated using the SAXS data from 1/500 to 1/50 Å for these four proteins is shown in the lower panel.

θ solvent, and thus the denatured state of SNase present under these conditions should be more compact than that for a random coil (Alonso & Dill, 1991). Reduced, denatured RNase A has also been shown to be more compact than a random coil (Sosnick & Trewhella, 1992). For the purposes of this study, we will use the scattering curve of SNase in 8 M urea to represent the "denatured state" of nuclease.

Changes in the Small-Angle X-ray Scattering Curve of SNaseΔ Due to Amino Acid Substitutions. Removal of the 13 carboxyl-terminal amino acids of SNase to yield SNaseΔ destabilizes the protein such that it is denatured at room temperature in the absence of denaturants; this state was previously shown to be compact and to lack persistent secondary structure (Shortle & Meeker, 1989; Flanagan *et al.*, 1992). To investigate the effects of the above mentioned amino acid substitutions on SNaseΔ we have measured the small angle X-ray scattering curves of WT SNaseΔ and various substituted versions of this polypeptide. The Guinier region for the scattering curves of WT SNase, WT SNaseΔ, A69T SNaseΔ (m+ substitution), and V66L SNaseΔ (m– substitution) are shown in Figure 2 (upper panel). As stated above, the presence of these substitutions in folded, full-length SNase had no measurable effect upon the R_g for the native state of the protein. In contrast, the presence of these substitutions in SNaseΔ produced large changes in R_g (Table I), as well as in the shape of the scattering curves at larger angles. The R_g of A69T SNaseΔ (m+ substitution) obtained from the data in the Guinier region of the scattering curve was 33 ± 0.9 Å compared with 20.2 ± 0.4 Å for WT SNaseΔ and 17.6 ± 0.2 Å for V66L SNaseΔ (m– substitution). The differences in R_g do not reflect changes in the aggregation state of the SNaseΔ polypeptide due to these substitutions, since the molecular masses obtained from $I(0)$ values for WT SNaseΔ and each of the m+ and m– substituted SNaseΔ polypeptides examined in this study were identical within experimental error (16 ± 1 kDa).

Table I: Summary of SAXS Results

protein	class	R_g(Guinier) (Å)	R_g(Moore) (Å)	D_{max} $P(r)$ (Å)
WT SNase	m^0	15.9 ± 0.1	16.2 ± 0.2	47
WT 8 M urea		33.0 ± 1.0	ND[a]	ND
WT SNaseΔ	m^0	20.2 ± 0.4	21.2 ± 0.2	63
V66L SNaseΔ	m–	17.6 ± 0.2	18.0 ± 0.2	54
G88V SNaseΔ	m–	17.4 ± 0.2	17.8 ± 0.2	53
V66L+G88V SNaseΔ	m–	17.3 ± 0.2	17.6 ± 0.2	53
V66L+G79S+ G88V SNaseΔ	m–	17.0 ± 0.2	17.3 ± 0.2	51
A90S SNaseΔ	m+	33.0 ± 0.9	35.2 ± 0.5	75
L25A SNaseΔ	m+	24.1 ± 0.7	25.0 ± 0.5	67
A69T+A90S SNaseΔ	m+	35.2 ± 1.0	37.0 ± 0.8	75

[a] ND, not determined.

To further characterize the behavior of m+ substitutions in this system, the scattering curves of two additional m+ substitutions (L25A SNaseΔ, A69T+A90S SNaseΔ) were measured. In both of these cases, as for A69T SNaseΔ, the R_g's were significantly larger than that of WT SNaseΔ (Table I). The increase in sizes, as reflected by the R_g's of these m+ substituted SNaseΔ molecules, correlates well with the greater degree of hydrophobic surface exposure in the denatured state implied by their effects on m_{den} in full-length SNase. By contrast, all four of the m– substitutions studied in SNaseΔ (V66L, G88V, V66L+G88V, and V66L+G79S+G88V) resulted in polypeptides with smaller R_g values than WT SNaseΔ, consistent with a smaller degree of hydrophobic exposure in the denatured state predicted for these substitutions in SNase. In fact, the R_g values for the m– substituted SNaseΔ were nearly as small as those of WT native SNase even though both the CD and wide-angle scattering data for these polypeptides indicate that they adopt non-native conformations (see below). It is surprising that single amino acid substitutions can produce such large effects upon the size of the polypeptide chain. The extent of this effect is best illustrated by comparing the R_g value for A69T SNaseΔ (33 Å) with that for V66L SNaseΔ (17.6 Å). The R_g for the former polypeptide is as great as that for urea-denatured SNase, while the latter more closely approximates that of native SNase (at higher NaCl concentration, >50 mM, the R_g of V66L SNaseΔ is 16.4 ± 0.2 Å; data not shown). An important feature of all of the SNaseΔ polypeptides studied is that they show catalytic activity (Shortle & Meeker, 1989; J. M. Flanagan, unpublished data). Furthermore, in some cases (WT SNaseΔ and all of the m– substituted SNaseΔ polypeptides), addition of pdTp and Ca^{2+} shifts the equilibrium very far toward the folded state (Shortle & Meeker, 1989; Flanagan *et al.*, 1992; data not shown).

For the polypeptides having R_g's significantly greater then WT SNase (WT SNaseΔ and m+ substituted SNaseΔ), the effects of interparticle interference lead to increased amounts of error in the innermost portion of the Guinier curve after extrapolation to infinite dilution. The validity of the R_g values determined for these versions of the SNaseΔ is borne out by the consistency of the values determined using various SAXS instruments (at Yale, the Photon Factory, and NSLS) under diverse collimation conditions as well as by the direct determination of the R_g from the most dilute samples (data not shown). Further, estimates of the relative size of these polypeptides using size exclusion chromatography lead to

Table II: Effects of Temperature and pH

(A) Effect of Temperature

	WT SNaseΔ		V66L SNaseΔ		A69T SNaseΔ	
temp (°C)	R_g (Å)	$I(0)$	R_g (Å)	$I(0)$	R_g (Å)	$I(0)$
4	30.0	5100	17.0 ± 0.2	2676		
10	19.1 ± 0.4	2800	17.0 ± 0.2	2702		
15	19.3 ± 0.4	2690	17.1 ± 0.2	2581		
20	19.1 ± 0.4	2830	17.2 ± 0.2	2622		
25	19.1 ± 0.4	2705	16.9 ± 0.2	2730	31 ± 1.0	2633
30	19.3 ± 0.4	3090	17.3 ± 0.2	2628		
35	33.0 ± 0.2	6080	21 ± 0.4	3300		

(B) Effect of pH

	WT SNaseΔ		V66L SNaseΔ		A69T SNaseΔ	
pH	R_g (Å)	$I(0)$	R_g (Å)	$I(0)$	R_g (Å)	$I(0)$
4.0	20.1 ± 0.5	2700	17.0 ± 0.2	2652		
4.5	19.2 ± 0.5	2900	17.2 ± 0.2	2721		
5.0	19.1 ± 0.2	2635	17.1 ± 0.2	2632		
5.5	19.1 ± 0.3	2700	16.8 ± 0.2	2708		
6.0	19.0 ± 0.3	2635	17.3 ± 0.2	2800		
6.5	19.2 ± 0.3	2526	17.5 ± 0.2	2632		
7.0	19.2 ± 0.3	2801	16.9 ± 0.2	2430	31 ± 1.0	2730
7.5	19.1 ± 0.3	2702	17.6 ± 0.2	2780	30 ± 1.0	2850
8.0	23.0 ± 0.5	3381	18.0 ± 0.5	3506		

similar conclusions for the relative sizes of the various polypeptides (Shortle *et al.*, 1989). The primary reason for the high degree of interparticle interference seen in these samples is probably the low ionic strength required to assure that all of the proteins are monomeric and the large net positive charge on SNase (pI ~10), resulting in an excluded volume from the electrostatic repulsion between individual protein molecules.

Effects of Temperature and pH on the R_g of WT and Mutated Versions of SNaseΔ. We have documented the effects of various substitutions on the SAXS curve of SNaseΔ as seen in changes in the values of R_g for the various truncated forms of SNase. One possible explanation that must be considered is that WT SNaseΔ exists as an equilibrium distribution between the native and denatured states and that amino acid substitutions affect the scattering curve by altering the relative population of these two states. Under this scenario, m– substitutions in SNaseΔ shift the distribution toward the native state, and result in a scattering curve having more native-like features. The effect of m+ substitutions in SNaseΔ would be to shift the distribution toward the denatured state; thus, the scattering curve will more closely resemble that of a denatured protein. To examine this possibility, we have measured the scattering curves of WT and V66L SNaseΔ at various temperatures (4–40 °C) and pH values (pH 4–8). Variation of these two parameters is expected to affect the relative stability of the native and denatured states or, for that matter, all thermodynamically linked compact and expanded states of the molecule.

Table IIA summarizes the results from a Guinier analysis for the scattering curves of WT and V66L SNaseΔ in the temperature range from 4–40 °C. The R_g and $I(0)$ values for both WT and V66L SNaseΔ were unaffected by changes in temperature between 7 and 35 °C within experimental error. For WT SNaseΔ at temperatures below 7 °C or above 35 °C, both the R_g and $I(0)$ values obtained from the Guinier analysis were significantly larger than those obtained at 20 °C indicating the presence of protein aggregation. In contrast, the scattering curve for V66L SNaseΔ shows no evidence of aggregation at 4 °C, though at high temperatures (above 40 °C) the effects of aggregation could also be seen. The onset of aggregation at high temperatures for V66L SNaseΔ paralleled the small but significant thermal transition observed in the DSC thermogram for this polypeptide (Flanagan, Tanaka, Engelman, and Sturtevant unpublished results). The temperature dependence of the scattering curves for m+ SNaseΔ was not examined in detail due to the high degree of aggregation present at temperatures below 15 °C and above 25 °C (data not shown). These results demonstrate that WT and V66L SNaseΔ represent thermally stable conformations of the polypeptide chain, and exclude the possibility that their scattering curves result from a combination of the curves for the native and denatured states of the molecules.

To further examine the stability of WT and V66L SNaseΔ, the scattering curves for these polypeptides were collected at 20 °C in the presence of 10 mM NaCl but varying the pH of the solution over the range from 4 to 8. Thermal denaturation studies of full length versions of these two proteins have shown that variation of the solution pH over this range results in a 20 °C range of denaturation temperatures (Shortle et al., 1989; Tanaka et al., 1993). The scattering curves for V66L SNaseΔ, as suggested by the values of R_g and $I(0)$, were independent of solution pH in this range (Table IIB). Similarly, the R_g and $I(0)$ values for WT SNaseΔ between pH 4 and 7.5 were also identical, though at pH 8.0 both R_g and $I(0)$ increased as a result of protein aggregation (Table IIB). These results provide further support for the idea that WT and the various substituted versions of SNaseΔ represent thermodynamically stable conformations of the polypeptide chain.

Comparison of the Circular Dichroism Spectra of WT and Substituted Forms of SNaseΔ. Circular Dichroism (CD), using wavelengths near the absorption band for amide bonds, provides a measure of residual structure present in SNaseΔ. For native proteins, spectra in this region are sensitive to the amount of α-helix and to a lesser extent β-sheet present in the structure. For polypeptides where a large fraction of their residues are in disordered or "random" conformations, there is considerable ambiguity in the use of CD spectra to quantitate the types and amounts of residual secondary structure (Shortle & Meeker, 1989; Flanagan et al., 1992); this said, CD spectra still provide useful qualitative indications of the amount of secondary structure. In a previous study, the CD spectra of WT and several of the versions of SNaseΔ examined in this study were measured under similar, though not identical, solvent conditions to those employed in the SAXS measurements (Shortle & Meeker, 1989). To provide a direct comparison between the SAXS and CD results, we have measured the CD spectra of WT SNase, WT SNaseΔ, and all seven of the substituted versions of SNaseΔ examined by SAXS. The protein samples were equilibrated against 10 mM TES (pH 7.5)/10 mM NaCl, and the spectra were collected at 20 °C, conditions that were identical to those used in the bulk of the SAXS measurements. For the subset of proteins (WT SNase, and WT, V66L, G88V, V66L+G88V, A69T, and A69T+A90S versions of SNaseΔ) examined in this study and by Shortle and Meeker (1989), the CD spectra were qualitatively and quantitatively identical (data not shown). Because of this fact, we have chosen to summarize our results (Table III) at three wavelengths, 199, 208, and 222 nm; these wavelengths correspond to minima in the CD spectra for native SNase (208 and 222 nm) and to the minimum present in the more expanded (m+) forms of SNaseΔ (199 nm). The magnitude of the negative ellipticity at 208 and 222 nm is primarily due to residues in an α-helical conformation (Johnson, 1991), while the minimum in the region near 199 nm is characteristic of polypeptides containing a large number

Table III: Effects of Amino Acid Substitutions on the CD of SNAseΔ

protein	class	[θ] (deg·cm²·dMOL) 199 nm	208 nm	222 nm
WT SNase	m^0	+6602	−12884	−12584
WT SNaseΔ	m^0	−10430	−9011	−6130
V66L SNaseΔ	m−	−5915	−9226	−6990
G88V SNaseΔ	m−	3065	−10430	−8560
V66L+G88V SNaseΔ	m−	−3000	−10645	−8581
V66L+G79S+G88V SNaseΔ	m−	−2950	−10802	−8620
A69T SNaseΔ	m+	−15055	−7520	−5430
L25A SNaseΔ	m+	−11028	−7922	−5642
A69T+A90S SNaseΔ	m+	−19226	−7420	−4097

of disordered residues (Johnson, 1991; Shortle & Meeker, 1989). A limit tryptic digest of nuclease gives a CD spectrum having a minimum at 199 nm as do the CD spectra of m+ substituted versions of SNaseΔ (Shortle & Meeker, 1989).

Comparing the magnitude of the ellipticity at 208 and 222 nm for WT SNase with that for the various forms of SNaseΔ suggests that these polypeptides each adopt a less structured average conformation than WT SNase. Since negative ellipticity in this region predominantly reflects helical content, these polypeptides contain reduced amounts of α-helix compared to native nuclease. These interpretations of the CD spectra are identical to those suggested in the previous study (Shortle & Meeker, 1989). Comparing the relative amount of negative ellipticity at 208 and 222 nm for eight versions of SNaseΔ shows that all four of the m− versions of SNaseΔ have a greater amount of negative ellipticity at these two wavelengths than for WT SNaseΔ, which in turn has a greater negative ellipticity than the three m+ versions of SNaseΔ at these wavelengths. Assuming that residues in a disordered or random coil conformation contribute little negative ellipticity at these wavelengths, the data suggest that m+ substitutions decrease the amount of residual structure while m− substitutions increase the amount of residual structure when present in SNaseΔ. Further support for this interpretation is obtained by an examination of the ellipticity at 199 nm for these molecules. At 199 nm, WT SNase has a positive ellipticity, while all of the forms of SNaseΔ have negative ellipticity at this wavelength. The magnitude of the negative ellipticity at 199 nm for WT SNaseΔ is intermediate between that for m− versions of SNaseΔ and m+ versions of SNaseΔ, with A69T+A90S SNaseΔ (m+) having the most negative ellipticity at this wavelength. In fact, the results of the previous study demonstrate that the CD spectra for m+ substituted SNaseΔ polypeptides most closely resemble that of a limit tryptic digest of SNase (Shortle & Meeker, 1989). Our results, and those of the previous study, support the idea that single amino acid substitutions can alter the amount of residual structure present in the SNaseΔ polypeptide. Taken together with the results from the SAXS data, they also suggest that the amount of residual structure increases as the polypeptide chain becomes more compact (decreasing R_g) in agreement with the results of several theoretical studies (Chan & Dill, 1989; 1990).

Calculation of P(r) Curves for the Various Forms of SNase and SNaseΔ. Examination of the scattering curve collected to larger angles can provide an estimate of the molecule's longest linear dimension (d_{max}) and an indication of its overall shape (Moore, 1980). A convenient means of representing this information is the radial Patterson [$P(r)$] curve, obtained by Fourier inversion of the scattering data, which is the real space analog to the scattering curve. The $P(r)$ profile is rarely calculated by direct inversion of the scattering data due, in part, to truncation artifacts or "Fourier ripples" arising from the limited range of reciprocal space encompassed by the actual measured scattering data [see Glatter and Kratky (1982) and Feigin and Svergun (1987) as well as references cited therein]. In practice, $P(r)$ curves are calculated using one of the many indirect Fourier transformation algorithms that suppress the artifacts produced by truncation of the scattering curves. However, all indirect transformation procedures require inclusion of additional information not directly specified by the data (Glatter, 1978; Moore, 1980; Taupin & Luzzati, 1985; Kube & Springer, 1987).

One widely used indirect Fourier transformation algorithm, employed in the programs ITP (Glatter, 1978) and GNOM (Svergun, 1992), synthesizes the $P(r)$ curve in real space using a sum of B-spline functions, the coefficients of which are derived by inverting each spline with the requirement that the sum of the splines adds up to the observed data. The other approach, employed in the program XCALC (Moore, 1980), fits the scattering data with a sum of $\sin(x)/x$ functions. The Fourier transforms of these are simply sine functions. Calculation of the $P(r)$ by either of these two algorithms includes an estimate of the longest linear dimension in the molecule (d_{max}). In addition, both ITP and GNOM require that two additional parameters be specified: N, the number of B-spline functions, and α, a Lagrangian multiplier that acts as a stabilizer. As has been discussed previously (Moore, 1980; Kube & Springer, 1987), the need for two additional parameters leads to some ambiguity in the calculation of the $P(r)$ curve by this method. The primary limitation of the reciprocal space algorithm of Moore (1980) is its sensitivity to "Fourier ripples". Direct application of this procedure works best when the scattering curve has been collected to large enough angles that the intensity does not differ significantly from the noise, thus minimizing the effects of truncation (Flanagan, Kataoka, and Engelman, unpublished results). To avoid this problem, when the scattering data are well determined, some implementations of XCALC artificially extend the data in the large angle region by means of Porod's Law (Moore, 1980), which asserts that for homogeneous globular proteins at sufficiently large angles the intensity of scattered radiation will be proportional to the inverse fourth power of the scattering angle [$I(Q) \propto Q^{-4}$]. Thus, for scattering curves where a Q^{-4} trend is apparent, the data can be extrapolated to infinite Q (Luzzati, 1979; Moore, 1980). At very large Q, Porod's Law no longer holds because the scattering arises from internal inhomogeneities in the molecule. Application of this procedure to scattering curves that do not conform to Porod's Law [polymer scattering $I(Q) \propto Q^{-2}$] or to portions of the scattering curve that are not in the Porod region results in distortions of the resultant $P(r)$ curve (see below).

In the course of this work, we have used XCALC, ITP, and GNOM to calculate $P(r)$ curves. For globular proteins, such as folded SNase, and using the same value of d_{max}, all three programs gave essentially identical results. Calculations of $P(r)$ curves for urea denatured SNase or for the more extended SNaseΔ (m+) polypeptides were more problematic. Figure 2 (lower panel) shows the $P(r)$ for WT SNase, WT SNaseΔ, V66L SNaseΔ, and A69T SNaseΔ calculated from the scattering data extrapolated to infinite dilution using the program ITP. Previously, we calculated the $P(r)$ curve for WT SNaseΔ using the program XCALC (Flanagan et al., 1992); the curve obtained in this manner showed a clear biphasic behavior, leading us to suggest that this protein adopted a bilobed structure in solution. In the current study,

using the programs ITP and GNOM, the $P(r)$ curves calculated for WT SNaseΔ do not exhibit as prominent a biphasic trend, so we feel that the $P(r)$ curve calculated for WT SNaseΔ with these programs more accurately reflects the molecular organization of this polypeptide (see below). We still believe that the overall structure of WT SNaseΔ contains at least two structural domains (one that is partially ordered and one that is primarily disordered); however, it is unlikely that these regions are as well-defined as we had previously thought. The $P(r)$ curves of WT SNase (d_{max} = 47.0 Å) and V66L SNaseΔ (d_{max} = 54 Å) are similar to those observed for compact, roughly spherical proteins. For most globular proteins it is possible to obtain a reasonable estimate of d_{max} using XCALC by examining the dependence of the reduced χ^2 (χ_v^2) for the fit on the input value of d_{max}. For solutions using too small a value of d_{max}, the resulting representation will fit the measured data badly, and the χ_v^2 value will be large. The χ_v^2 value of the fit will drop until it reaches a value of about $\sim$0.7 (Moore, 1980), and for larger values of d_{max}, the χ_v^2 value will remain nearly constant. The input value of d_{max} for which the χ_v^2 value first reaches a minimum is the best estimate for the "real" d_{max} value (Moore, 1980; Ramakrishnan et al., 1991). All of the d_{max} values reported in this work have been determined by this method. For trial values of d_{max} larger than the actual longest linear dimension of the molecule, there may be subsequent minima in the χ_v^2 value, some of which may correspond to better fits in terms of this criterion; however, for globular proteins, $P(r)$ profiles calculated using these values contain negative excursions at large r values. The estimated d_{max} value for WT SNase determined using this procedure was within 2 Å of that found in the crystal structure (Hynes & Fox, 1992). Further, $P(r)$ curves calculated with values larger than our estimates of d_{max} contained negative excursions at the longer values of r, while those calculated with values smaller than our estimates of d_{max} showed sharp truncations in the $P(r)$ profile. This behavior was observed for all folded SNases as well as all m– versions of SNaseΔ (data not shown).

Calculation of $P(r)$ curves of m+ versions of SNaseΔ and to a lesser extent WT SNaseΔ with XCALC was not straightforward, since the scattering curves of these molecules do not contain a pronounced Porod region (see below). Further, their scattering curves beyond the Guinier region are closer to those observed for coil molecules, and thus the decrease in the intensity of scattered radiation is more gradual than that of a compact globular protein (see below) and never approaches zero. As a result, direct application of Moore's algorithm in calculating the $P(r)$ profiles of WT SNaseΔ and the m+ versions of SNaseΔ leads to truncation distortions. In principle, the artifacts due to truncation of the high Q data do not affect the determination of d_{max} by this method. The value of d_{max} is determined primarily by the dependence of the scattered radiation in the small angle region; in fact, as long as the data have been collected to at least $Q < \pi/d_{max}$, this method should allow d_{max} to be estimated (Moore 1980). All of the $P(r)$ profiles were calculated from data collected to Q values of at least 0.0126 Å^{-1}; therefore, the data are sufficient to determine d_{max} values as large as 200 Å. Surprisingly, $P(r)$ profiles for WT SNaseΔ calculated using trial d_{max} values 20 Å larger than the optimum value determined as described above did not result in negative excursions in the profile. For A69T SNaseΔ, the situation was even less satisfactory, since we were unable to find a value of d_{max} that resulted in $P(r)$ profiles containing negative excursions (data not shown).

There are two probable explanations for this behavior: first, both of these polypeptides exist in an expanded conformation, and their scattering curves represent an ensemble average of a distribution of conformations; thus, there will be a distribution of d_{max} values for the protein molecules in the ensemble which is broad compared to that for native compact proteins. The "true" d_{max} for the ensemble is the maximum value of d_{max} represented in the distribution and is determined by the range of conformations accessible to the polypeptide chain under these conditions. For a broad distribution in d_{max} values, the d_{max} for the ensemble will be substantially larger than its mean value, and only a small fraction of the polypeptide chains in the sample adopt a conformation consistent with this value. Consequently, chains having a conformation consistent with the ensemble d_{max} contribute little to the total intensity of scattered radiation. This situation has been discussed previously with reference to the calculation of $P(r)$ profiles for polystyrene in benzene (Kube & Springer, 1987). In this work, the values of d_{max} required to fit the scattering data accurately showed a threshold behavior, where scattering curves could be accurately represented by all input values of d_{max} above some threshold. For systems of this type, the physical significance of d_{max} obtained in this way is obscure. A second, and to some extent related, problem is whether the χ_v^2 used to determine d_{max} for globular proteins is appropriate for nonglobular systems. The criterion used to obtain d_{max} for globular, roughly spherical proteins is based upon the value of d_{max} where the χ_v^2 value first reaches a minimum, and the argument for this choice is summarized as follows: trial values of d_{max} that are smaller than the optimum d_{max} fit the data poorly since contributions from a large number of intraparticle vectors are omitted; thus, the χ_v^2 will be large; as the optimum value of d_{max} is approached, the χ_v^2 value will monotonically approach a minimum value, since we are systematically including more scattering vectors. However, for particles that are very asymmetric, or that contain multiple separated domains, or where there is a distribution of particle conformations, the χ_v^2 value may not monotonically approach a minimum at the optimum d_{max}. This is best illustrated by considering a particle containing two domains of equal size separated by some interdomain distance. The number of intradomain vectors will be substantially larger than of interdomain vectors. The χ_v^2 value at a trial d_{max} corresponding to the intradomain d_{max} would be a local minimum for the reasons given above, but this value of d_{max} is clearly incorrect. In this situation, the χ_v^2 value at the local minimum will still be rather large, and the fit to the data in the small angle region will be poor. The choice of the appropriate χ_v^2 criterion for the other situations is not absolutely clear.

Given the problems in obtaining a $P(r)$ curve for the denatured forms of nuclease (WT SNaseΔ and m+ forms of SNaseΔ), it might be asked whether one should attempt to calculate them at all. There are two answers to this question. First, $P(r)$ profiles are one of the typical ways to represent the scattering data from biological samples, and thus they provide an important point for comparison with a body of previous literature. Second, and more importantly, they provide additional estimates for the values of R_g and $I(0)$ where the effects of instrumental smearing have been eliminated. Previously, it has been pointed out that the Guinier approximation for a random coil leads to an underestimate of the R_g while that determined from the best-fit $P(r)$ curve more accurately reflects its "true" value (Glatter, 1978; Kube & Springer, 1987). Table I shows a summary of the R_g values determined by XCALC. These values are all larger than the

quantities obtained from a Guinier analysis. The main reason for this difference is that Guinier analysis is most often performed on data that have not been corrected for the effects of smearing due to the geometry of the scattering apparatus, and the difficulties inherent in defining an appropriate Guinier region (Glatter, 1978).

*Comparison of the Intermediate-Angle X-ray Scattering Profiles for Various Forms of SNase and SNase*Δ. The information contained in a scattering curve may be represented in numerous ways to accentuate various features of the data; the Guinier plot and the $P(r)$ profile are only two of these (Glatter & Kratky, 1982; Feigin & Svergun, 1987). Previously, we have shown that the urea denaturation transition of WT SNase may be accurately followed by SAXS and that representing the data as a Kratky plot provides a useful means of following the globular (native) to coil (denatured) transition of this molecule (Kataoka et al., 1993) and for cytochrome *c* (Kataoka et al., 1993). The Kratky plot has also been used to analyze the neutron scattering data for human plasminogen (Ramakrishnan et al., 1991).

Figure 3 (upper panel) shows the Kratky plots for three model particles: a sphere, a Gaussian coil, and a wormlike or persistent chain (Kratky & Porod, 1949), all calculated using the same value for $I(0)$ and a R_g of 60 Å. The sphere is a reasonable model for a folded protein, while both Gaussian coils and persistent coils represent hypothetical models for the denatured state. A Gaussian coil model describes a combination of N-linked elements, each of which is rigid and has a length l. For a polypeptide chain, l is ~3.8 Å and corresponds to the distance between adjacent C_α atoms. For a chain, the distance r_{ij} between the ith and jth elements varies with time; thus, we can define a mean square interelement distance, $\langle r_{ij}^2 \rangle$, between these two elements. A Gaussian coil is defined as one where the probability density between intersegment distances is

$$p(r_{ij}^2) = \left(\frac{3}{2\pi\langle r_{ij}^2\rangle}\right)^{3/2} 4\pi r_{ij}^2 \exp\left(-\frac{3r_{ij}^2}{2\langle r_{ij}^2\rangle}\right)$$

which is a Gaussian distribution (Glatter & Kratky, 1982). The scattering function for a Gaussian coil has been calculated by Debye (1947) with the result:

$$I(Q) = \frac{2(e^{-x} + x - 1)}{x^2}$$

with $x = \langle R_g^2 \times 85 \rangle Q^2$. At very small angles, the Guinier approximation will hold and the root mean square radius of gyration, $\langle R_g^2 \rangle$, for the coil can be obtained. At larger scattering angles, the curve will have an asymptotic Q^{-2} course. This feature is clearly seen in the Kratky plot, and corresponds to a line parallel to the Q axis (upper panel of Figure 3).

The persistent coil model assumes that a polymer can be represented as an infinitely thin filament subjected to a specific bending procedure. The Gaussian coil is a limiting case; its specific bending feature is l, the intersegment length, and a rod of length Nl is the other limit. A persistent coil is characterized by its degree of flexibility, defined by the persistence length, a. For a Gaussian coil, the persistence length is $a = l/2$, while for a chain of arbitrary stiffness, $a = b/2$, where b is the length of the statistical chain element. As one might expect from this definition, the scattering function for a persistent coil will have three defined regions. First, at very small scattering angles, the Guinier approximation will describe the scattering curve. Second, at higher angles, its scattering curve will resemble that for a Gaussian

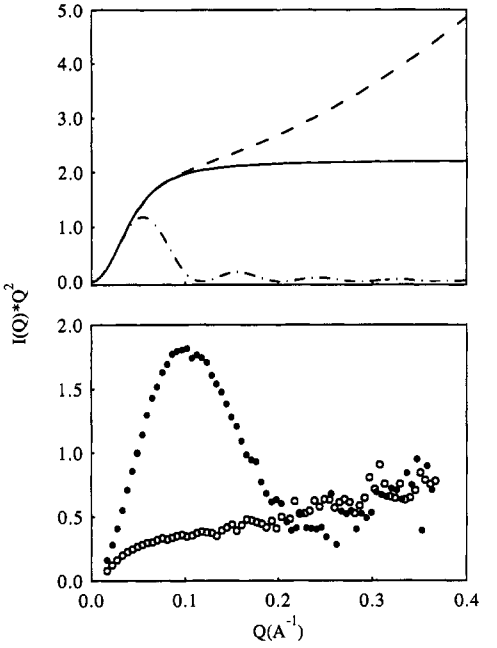

FIGURE 3: Kratky plots [$I(Q)\cdot Q^2$ vs Q] of SAXS data for model functions WT, and denatured SNase. Kratky plots for a sphere (– · –), a Gaussian coil (—), and a persistent or wormlike coil (– –), all having the same molecular weight and an R_g of 60 Å; calculated from analytically derived SAXS functions (upper panel). For comparison, the Kratky plots of SAXS data for native SNase (●) and SNase in 8 M urea (○) are shown in the lower panel.

coil having an effective segment length of $l = 2a$. Finally, at still higher angles ($Q \rightarrow \infty$), the scattering curve for a persistent coil will approach that for an infinitely thin rod (Glatter & Kratky, 1982; Feigin & Svergun, 1987). In a Kratky plot, the scattering curve for an infinitely thin rod is a straight line through the origin with a positive slope, thus, at high Q the curve for a persistent coil will have a region that asymptotically approaches the line describing an infinitely thin rod. These features are readily seen in a Kratky plot (upper panel of Figure 3), where at small to intermediate scattering angles the curves for a persistent coil and for a Gaussian coil superimpose, and at higher angles that for a persistent coil tends toward a linear dependence of $I(Q)\cdot Q^2$ on Q. The persistence length of the polymer can be estimated from the intersection between the asymptotes for the Gaussian coil and rodlike portions of the curve (Kratky & Porod, 1947).

In the Kratky plot, the scattering curve for a sphere shows a peak, as is clear from Figure 3 (upper panel). Likewise, the scattering curve for a folded, roughly spherical globular protein should give a similar peak in the Kratky plot. The steep rise in the low Q region, which is observed for all of the model functions, results from the multiplication of the Guinier (Gaussian) portion of the curve by relatively small Q^2 values. At higher angles, where the Guinier approximation no longer holds ($Q > QR_g \sim 3$–5), the scattering curve can be approximated by Porod's Law, or $I(Q) \propto Q^{-4}$. As a result, the intensity should decrease rapidly as Q^{-2} when the observed intensity is multiplied by Q^2. The position of the peak will depend upon the R_g of the particle; for increasing values of R_g, the peak position will move to smaller values of Q. From this analysis we can conclude that a peak in the Kratky plot indicates a compact globular structure while the absence of a peak is an indication of a more coil-like structure (Kataoka et al., 1993a,b). The primary advantage of the Kratky plot over the $P(r)$ profile is that the Kratky representation is based solely upon the observed intensity curve and does not require

the specification of additional parameters such as a knowledge of d_{max}.

Figure 3 (lower panel) shows Kratky plots for the scattering curves of WT SNase and WT SNase in 8 M urea. From the Kratky plot, it is clear that urea denaturation of WT SNase produces a transition from a compact, globular state to one that is coil-like in nature. This result is consistent with the bulk of spectroscopic and hydrodynamic measurements for urea denaturation of various proteins. A similar transition has been observed by small-angle neutron scattering for human Glu-plasminogen (Ramakrishnan et al., 1991). In this system, it is thought that domain interaction between individual kringles in Glu-plasminogen leads to a compact, nearly globular structure, binding of 6-aminohexanoic acid abolishes the interdomain interaction, leading to a structure that is best described as a Gaussian coil. Qualitatively, the Kratky plots for human Glu-plasminogen obtained in the absence and presence of 6-aminohexanoic acid are similar to those for SNase in the presence and absence of 8 M urea (Mangel *et al.*, 1990; Ramakrishnan *et al.*, 1991).

Examination of the Kratky plot for WT SNase in 8 M urea at slightly larger angles suggests that the molecule may be best represented as a persistent coil. In the larger angle region ($Q \geq 0.2$), the quantity $I(Q)\cdot Q^2$ increases nearly linearly with Q. For well-determined data the crossover point between coil-like and rodlike behavior defines the persistence length, a, which is given by $a = 1.91/Q^*$ where Q^* is the crossover point. For SNase in the presence of 8 M urea, the data above 0.12 Å^{-1} are very noisy and do not allow a precise value of Q^* to be estimated. However, from the general trend of the data in this region, it is likely that Q^* lies between 0.12 and 0.2 Å^{-1}. This range of estimates for Q^* corresponds to estimates of a between 9.6 and 15.9 Å. If WT SNase in 8 M urea can be thought of as a persistent coil, then the R_g of value for the polypeptide can be estimated [$R_g^2 = aL/3$, where L is the contour length of the polypeptide chain ($L = Nl$)]. The value of R_g consistent with the range of persistence length is between 42.6 and 54.8 Å. These values for R_g are somewhat larger than that obtained by the Guinier analysis. The difference between the R_g values estimated using the two different procedures may result either from an improper estimate of the persistence length or from the limitations of our assumption that the denatured state of SNase in 8 M urea can be approximated by a persistent coil model. A more detailed discussion of the urea-denatured state of SNase will be presented elsewhere (Flanagan, Kataoka, and Engelman, manuscript in preparation).

The usefulness of the Kratky plot in comparing the native state and urea-denatured states is clear, and the results for these two states provide a basis for further characterizing the various forms of SNaseΔ. Figure 4 (upper panel) shows the Kratky plot of the scattering data for WT SNaseΔ, V66L SNaseΔ (m– substitution), and A69T SNaseΔ (m+ substitution) normalized to equivalent $I(0)$ values. The profile for V66L SNaseΔ is similar in general shape to that observed for globular, compact proteins, and supports the previous conclusions resulting from its $P(r)$ profile and nativelike R_g value. In contrast, the scattering curve for A69T SNaseΔ is quite similar to that for SNase in 8 M urea, though a small broad peak can be seen in the Kratky plot (lower panel of Figure 4). The scattering curve for WT SNaseΔ is intermediate between those for the V66L and A69T versions of SNaseΔ, in that a slightly large peak is observed in the Kratky plot for WT SNaseΔ. Comparison of the Kratky plot for A69T SNaseΔ with that obtained for WT SNase in 8 M urea

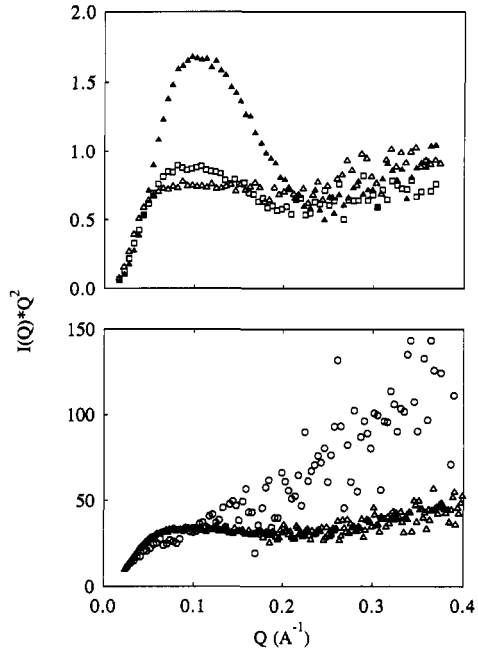

FIGURE 4: Effects of amino acid substitutions on the conformation of Δ(137–149). Kratky plots of WT SNaseΔ (□), V66L SNaseΔ (▲), and A69T SNaseΔ (△) are shown (upper panel). The lower panel shows a comparison of the Kratky plots of SNase in 8 M urea (○) and A69T SNaseΔ (△). The error bars have been omitted for clarity in the large Q region (>0.225 Å^{-1}).

demonstrates that these two polypeptides do not adopt the same conformation (lower panel of Figure 4). In the region of the scattering curve between a Q value of 0.02 and 0.1 Å^{-1} (real-space distances of 340–62.8 Å), the scattering curve of A69T SNaseΔ shows excess scattered intensity compared with that for urea-denatured SNase, though both polypeptides have nearly the same R_g. The height of the plateau region in the Kratky curve is dictated by the value of $\langle R_g^2 \rangle$ for a Gaussian coil (see scattering function for Gaussian coil given above). In the case of a persistent coil, this analysis is complicated by the effects of the persistence length on the shape of the Kratky plot, though for a persistent coil the quantity $I(Q)\cdot Q^2$ will be less than that for a Gaussian coil having the same R_g [see Figure 8 in Chapter 12 in Glatter and Kratky (1987)]. In addition, for A69T SNaseΔ, there appears to be no region of the scattering curve that exhibits rodlike scattering behavior. Unfortunately, the information content of the scattering curve is not sufficient to uniquely determine the structure of the molecule. However, assuming that in 8 M urea SNase is indeed well described as a persistent coil, the fact that the R_g values of A69T SNaseΔ and SNase in 8 M urea are nearly identical argues that A69T SNaseΔ adopts a more extended (elongated) conformation. Comparing the Kratky plots of WT and A69T SNaseΔ with that of SNase in 8 M urea suggests that the overall structure of A69T SNaseΔ is probably more similar to WT SNaseΔ than urea-denatured SNase, and thus probably retains some globular-like domains. This interpretation is consistent with the observed decrease in elution volume from a SEC column of the A69T SNaseΔ polypeptide in the presence of urea, suggesting that it too has some residual structure (Shortle & Meeker, 1989). Thus, it appears that A69T SNaseΔ adopts a conformation that differs significantly from SNase in 8 M urea and that its conformation cannot be explained by a simple persistent coil model. Further, this analysis shows that single amino acid substitutions can produce radical changes in the conformation of the polypeptide: the m– substitutions result in polypeptides whose conformations

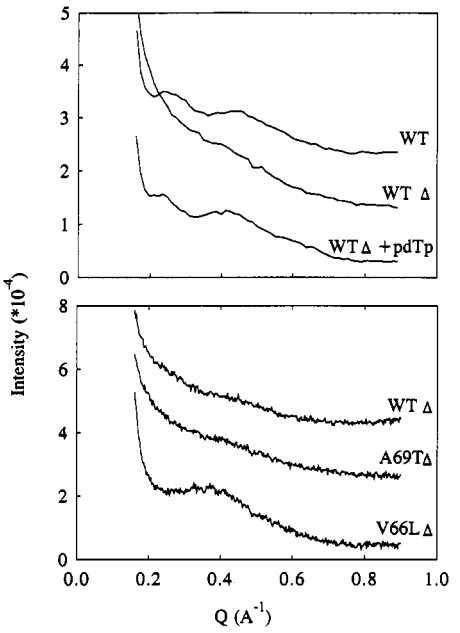

FIGURE 5: Extended scattering curves of SNase and substituted versions of SNaseΔ. The very large angle scattering curves of WT SNase, WT SNaseΔ, and WT SNaseΔ + 3mM pdTp and 10mM Ca^{2+} are shown (upper panel). The protein concentration of these samples was 30 mg/mL, and they were collected at ambient temperature (20–25 °C). The lower panel shows a comparison of the large-angle data for WT SNaseΔ, A69T SNaseΔ, and V66L SNaseΔ for 30 mg/mL samples of these proteins. The scattering curves have been displaced along the intensity axis for clarity.

are more compact and thus constrained, while m+ substitutions result in polypeptide conformations that are less globular and consequently less constrained then WT SNaseΔ.

Examination of the Scattered Intensity of X-rays at Very Large Angles for Several Forms of SNaseΔ. The information contained in the very large angle region of the scattering curve ($^1/_{40}$–$^1/_6$ Å) represents the internal packing of amino acid residues in the molecule; thus, it is related to the tertiary structure of the molecule. Unfortunately, the three-dimensional structure cannot be extracted from the information contained in this region, and secondary structural themes are hard to derive (Pickover & Engelman, 1981). Nevertheless, it does provide a fingerprint for the protein being studied. The very large angle region of the scattering curve for WT SNase contains two features: one, a small peak, corresponding to a *d*-spacing of approximately 25 Å and the other, a broader peak, to a *d*-spacing of about 15 Å (Figure 5). The scattering curve for WT SNaseΔ collected in the absence of pdTp·Ca^{2+}, a tightly binding inhibitor of SNase, shows a smooth decrease in the scattered intensity as a function of angle; however, in the presence of these two ligands, the scattering curve of WT SNaseΔ is nearly identical to that for native SNase (Figure 5, upper panel), demonstrating that this molecule is capable of adopting a near-native conformation.

A comparison of the scattered intensity of X-rays at these angles for WT, A69T, and V66L SNaseΔ is shown in the lower panel of Figure 5. The scattering curves for WT and A69T SNaseΔ are nearly identical; the dependence of the intensity of scattered radiation on angle varies rather smoothly in this range for both polypeptides. In contrast, the scattering curve for V66L SNaseΔ shows a single broad peak with a maximum at a *d*-spacing of about 17 Å. This peak in the scattering curve of V66L SNaseΔ is similar, though not identical, to the maximum that occurs at higher angles (second maximum ~15 Å) in the scattering curve of native SNase. The scattering data do not provide sufficient information to resolve the issue of whether this feature arises from similar structures present in both molecules or from entirely different structures in the two molecules. However, the presence of this peak in the curve of V66L SNaseΔ combined with the absence of the lower angle peak (25 Å) present in the scattering curve of native SNase confirms our earlier interpretation that the scattering curve of V66L SNaseΔ does not result from a combination of native and denatured molecules. The lack of any apparent maxima in this region of the scattering curves for WT and A69T SNaseΔ demonstrates that these molecules contain little persistent tertiary structure.

COMMENTS AND CONCLUSIONS

Protein stability arises from energy differences between native and denatured state conformations, so changes in the distribution of conformations in the denatured state are expected to have an important influence. Until recently, it has been assumed that the effects of amino acid substitutions on the denatured state are small (except those involving proline or glycine; Nemethy et al., 1966) compared to their effects on the native state. If this assumption is valid, the observed changes in stability (ΔG), enthalpy (ΔH), and entropy (ΔS) can be safely interpreted as resulting from alterations in the native state. Our results contribute to the small but growing body of evidence that call this assumption into question by demonstrating that the denatured states are condensed and affected by mutations.

In the present work, we show that some single amino acid substitutions produce large changes in the conformation of a denatured state of SNase produced by a carboxyl terminal deletion of 13 amino acid residues. The effects of these substitutions on the R_g (a crude measure of size) for SNaseΔ could be predicted on the basis of their effects upon m_{den} for solvent denaturation of full-length SNase. These results are consistent with the view that some amino acid substitutions affect the stability of SNase by altering its denatured-state conformation.

In addition to their effects upon the solvent denaturation of SNase, the mutations that we studied are known to produce large changes in the values of ΔH, ΔS, and ΔC_p obtained from thermal denaturation studies (Shortle *et al.*, 1990; Tanaka, *et al.*, 1993). These substitutions could be placed into two groups on the basis of their effects on ΔH, ΔS, and ΔC_p: m+ substitutions (L25A, A69T, A90S, and A69T+A90S) resulted in large increases in the values of ΔH, ΔS, and ΔC_p, while m– substitutions (V66L, G88V, V66L+G88V, and V66L+G79S+G88V) all substantially decreased the values of ΔH, ΔS, and ΔC_p [with the exception of the calorimetrically determined ΔC_p value for G88V (Tanaka *et al.*, 1993)]. The most extreme example of this behavior was V66L+G79S+G88V SNase, which is marginally more stable than WT SNase ($\Delta\Delta G$ = 0.38 kcal/mol); however, the stabilizing effects of these substitutions result from large opposing decreases in both ΔH and ΔS [$\Delta\Delta H$ = 52 kcal/mol and $\Delta\Delta S$ = 160 cal/(K·mol) at pH 7.0 and 51.4 °C). Moreover, the value of ΔC_p was not a constant, but decreased with decreasing $t_{1/2}$ (Tanaka *et al.*, 1993). These values imply either that the native state of V66L+G79S+G88V SNase is thermodynamically closer to its denatured state than to the native state of WT or that its denatured state is very compact and has an almost nativelike enthalpy and entropy (Tanaka *et al.*, 1993).

A plausible explanation for the observed thermodynamic changes resulting from these substitutions is suggested by the

results of this study as well as those obtained previously (Shortle & Meeker, 1989), provided that the conformation of the denatured state produced, by deletion of the carboxyl-terminal 13 amino acid residues, retains some properties common to the denatured state produced at high temperatures. With this as an assumption, the substantially larger values of ΔH, ΔS, and ΔC_p for SNase containing m+ substitutions result from the greater degree of unfolding and consequent loss of residual structure in the denatured state of the polypeptide compared to WT. The thermodynamic behavior of the m- versions of SNase can be explained if their denatured states are more compact and contain a larger amount of residual structure compared to that for WT SNase; a denatured state having these properties is indeed observed for the m- versions of SNaseΔ. This interpretation not only accounts for the direction of the changes in ΔH and ΔS but also accounts for the large magnitude of these changes (Shortle *et al.*, 1990; Tanaka *et al.*, 1993), since, at least in SNaseΔ, the alterations in the denatured-state conformations are quite dramatic and represent changes in the overall organization of these polypeptides.

The assumption that the genetically produced denatured state (SNaseΔ) approximates, in some form, features present in the thermally or solvent-denatured states of SNase is critical to this interpretation. One argument supporting this view is the recent observation of residual structure in the urea-denatured state of the 434 phage repressor (Neri et al., 1992). This work demonstrates that there is residual structure even in 7 M urea. A second argument is that the denatured states produced at high temperatures, in the presence of solvent denaturants (urea or guanidinium chloride) and at extremes of pH, are quite different (Aune et al., 1967; Kugiyama & Bigelow, 1973; Howarth & Lian, 1981; Dobson et al., 1984; Evans et al., 1991; Tamura et al., 1991a,b); however, the values for ΔG obtained by extrapolation from the transition region for the various types of denaturation curves to equivalent solvent conditions are nearly identical (Hu et al., 1992; Sanatoro & Bolen, 1992; Privalov et al., 1989). The reference conditions employed in these extrapolations are usually 25 °C, neutral pH, and the absence of added denaturant, conditions similar to those employed in our study.

The information provided by SAXS and CD is not sufficient to specify the conformation of SNaseΔ, however, there is a wealth of ancillary data describing the effects of amino acid substitutions on the stability of SNase (Shortle et al., 1990; Green *et al.*, 1992). In one study, all of the Leu, Ile, Val, Phe, Tyr, and Met (large hydrophobic) residues were substituted with both Ala and Gly residues, and in the other, the Ala, Gly, Thr, Pro, Gln, Asn, and Ser (polar uncharged amino acids) residues were replaced with both alanine and glycine or with one of these two amino acids plus valine. The residues examined in these two studies constitute more than half of the total amino acids in SNase. Two conclusions have emerged from these studies. First, m+ substitutions occur only in the major hydrophobic core of the molecule, composed of the strands of the highly twisted β-barrel (excluding the amino-terminal half of β-strand 1) and portions of α-helices 1 and 2 (Green et al., 1992), while m- substitutions occurred in the remainder of the molecule. Second, these authors observed a strong correlation between m_{den} and $\Delta\Delta G$, the stability of SNase, suggesting that these substitutions affect stability in part by altering the conformation of the denatured state (Green *et al.*, 1992, Shortle *et al.*, 1990). This conclusion follows from the three major models for the mechanism of urea or Gu·HCl denaturation, each of which suggests that m_{den} reflects differences in the interaction of the denaturant with the native and denatured states of the molecule (Tanford, 1970; Shellman, 1978). As has been argued previously, changes in m_{den} are most readily explained by changes in the surface area of the denatured state (Shortle & Meeker, 1987; Shortle & Meeker, 1989; Green et al., 1993). Accepting these arguments, our results and those from the previous study (Shortle & Meeker, 1989) suggest a model for the denatured state of SNaseΔ. The primary feature of this model is that the major hydrophobic core of native SNase is maintained in some form in the denatured state of WT SNaseΔ. This is consistent with our previous observation that WT SNaseΔ contains some residual structure that includes a small amount of α-helix (Flanagan *et al.*, 1992). The structure of this region is probably not fully native (Flanagan *et al.*, 1992), and there is little evidence of tertiary interactions as judged by the scattering at very large angles (Figure 5). This model is supported by two observations: first, all of the m+ substitutions occur in this region; second, given the alternating hydrophobic/hydrophilic pattern of amino acid sequence of SNase in this region, it is unlikely that the polypeptide would be soluble under these conditions unless a large fraction of the residues were buried. The collapse of this region may account for the relatively compact state of this fragment.

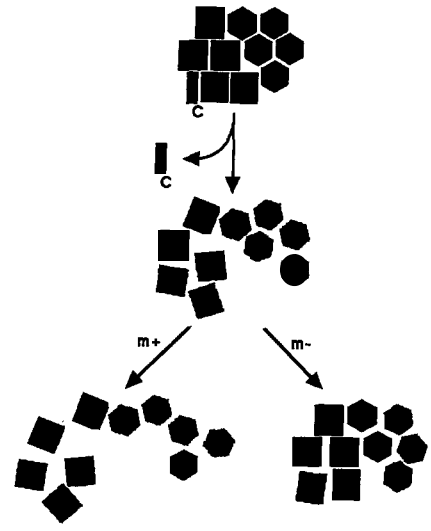

FIGURE 6: Conceptual model for the effects of amino acid substitutions on the conformation of SNaseΔ.

Given a hydrophobic core in the native sequence-denatured state, the effects of m- substitutions would then be to further organize the polypeptide chain by permitting the second, smaller hydrophobic core in SNase (carboxyl-terminal region of the molecule) to form. The increase in secondary structure and the presence of some tertiary-like structure (Figure 5, bottom panel) in the m- version of SNaseΔ suggest that increased compaction requires the formation of secondary structure to satisfy the hydrogen-bonding potential of the chain due to the exclusion of water. A schematic of our model is shown in Figure 6. This model merely summarizes the observed effects of m+/- substitutions on the R_g of SNaseΔ using the constraints provided by the extensive mutational analysis of full-length SNase.

In summary, this study provides direct support for the view that the denatured state of a protein is not a random coil, that its conformation depends upon the solvent conditions employed, and that the conformation of the denatured state of SNase nuclease produced by deletion of the carboxyl terminal portion of the chain is very sensitive to certain point mutations. Our results and those of Shortle and Meeker (1989) strongly

support the possibility that some mutations affect protein stability by altering the conformation of the denatured state.

ACKNOWLEDGMENT

We thank Mark Lemmon and Dr. Fred Richards for carefully reading the manuscript. In addition, we are indebted to Dr. Malcom Capel for the use of the X12B beam line at NSLS.

REFERENCES

Alonso, D. O. V., & Dill, K. A. (1991) *Biochemistry 30*, 5974–5985.
Chan, H. S., & Dill, K. A. (1989) *Macromolecules 22*, 4559–4573.
Chan, H. S., & Dill, K. A. (1990) *Proc. Natl. Acad. Sci. U. S. A. 87*, 6388–6392.
Debey, P. (1947) *J. Phys. Colloid Chem. 51*, 18–32.
Dobson, C. M., Evan, P. A., & Willamson, K. L. (1984) *FEBS Lett. 168*, 331–334.
Evan, P. A., Topping, K. D., Woolfson, D. N., & Dobson, C. M. (1991) *Proteins: Struct., Funct., Genet. 9*, 48–66.
Feigin, L. A., & Svergun, D. I. (1987) *Structural Analysis by Small-Angle and Neutron Scattering*, Plenum Press, New York.
Flanagan, J. M., Kataoka, M. K., Shortle, D., & Engelman, D. M. (1992) *Proc. Natl. Acad. Sci. U. S. A. 89*, 748–752.
Glatter, O. (1977) *J. Appl. Crystallogr. 10*, 415–421.
Glatter, O. (1980) *J. Appl. Crystallogr. 13*, 7–11.
Glatter, O. (1982) in *Small Angle X-ray Scattering* (Glatter, O., & Kratky, O., Eds.) pp 119–165, Academic Press, London.
Green, S. M., Meeker, A. K., & Shortle, D. (1992) *Biochemistry 31*, 5717–5728.
Guinier, A., & Fournet, B. (1955) *Small Angle X-ray Scattering*, pp 5–82, John Wiley, New York.
Howarth, O. W., & Lian, Y. (1981) *J. Chem. Soc., Chem. Commun.*, 258–259.
Hu, Q.-C., Sturtevant, J. M., Thomson, J. A., Erickson, R. E., & Pace, N. C. (1992) *Biochemistry 31*, 4876–4882.
Kataoka, M., Head, J. F., Seaton, B. A., & Engelman, D. M. (1989) *Proc. Natl. Acad. Sci. U. S. A. 86*, 6944–6948.
Kataoka, M., Flanagan, J. M., & Engelman, D. M. (1993a) in *Synchrotron Radiation in Life Science* (Sturmann, H. B., et al., Eds.) Oxford University Press, London (in press).
Kataoka, M., Hagihara, Y., Mihara, K., & Goto, Y. (1993b) *J. Mol. Biol.* (in press).
Kratky, O., & Porod, G. (1949) *Recl. Trav. Chim. Pays-Bas. 68*,1106–1122.
Kube, O., & Springer, J. (1987) *J. Appl. Crystallogr. 20*, 41–47.
Langer, T., Chi, L., Harrison, E., Flanagan, J., Hayer, M. K., & Hartl, F. U. (1992) *Nature 356*, 683–689.
Lemmon, M. A., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B-J., Dempsey, C. E., & Engelman, D. M. (1992) *J. Biol. Chem. 267*, 7683–7689.
Luzzati, V. (1979) *Imaging Processes and Coherence in Pysics*, Springer-Verlag, Heidelberg.
Mangel, W. F., Lin, B., & Ramakrishnan, V. (1990) *Science 248*, 69–73.
Miller, W. G., & Goebel, C. V. (1968) *Biochemistry 7*, 3925–3935.
Moore, P. B. (1980) *J. Appl. Crystallogr. 13*, 168–175.
Neri, D., Wider, G., & Wuthrich, K. (1992a) *Proc. Natl. Acad. Sci. U. S. A. 89*, 4397–4402.
Neri, D., Billeter, M., Wider, G., & Wuthrich K. (1992b) *Science 257*, 1559–1563.
Pickover, C., & Engelman, D. M. (1981) *Biopolymers 21*, 817–831.
Privalov, P. L., Tiktopolo, E. I., Venyaminov, S. Yu., Griko, Yu. V., Makhatadze, G. I., & Khechinashvili, N. N. (1989) *J. Mol. Biol. 205*, 737–750.
Ramakrishnan, V., Patthy, L., & Mangel, W. F. (1991) *Biochemistry 30*, 3963–3969.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Santoro, M. M., & Bowlen, D. W. (1992) *Biochemistry 31*, 4901–4907.
Shortle, D. (1983) *Gene 22*, 181–189.
Shortle, D., & Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81–89.
Shortle, D., & Meeker, A. K. (1989) *Biochemistry 28*, 936–944.
Shortle, S. Meeker, A. K., & Freire, E. (1988) *Biochemistry 27*, 4761–4768.
Shortle, D., Stites, W. E., & Meeker, A. K. (1990) *Biochemistry 29*, 8033–8041.
Sondek, J. E., & Shortle, D. (1990) *Proteins: Struct., Funct., Genet. 7*, 299–305.
Sosnick, T. R., & Trewhella, J. (1992) *Biochemistry 31*, 8329–8335.
Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60–89.
Svergun, D. I. (1992) *J. Appl. Crystallogr. 25*, 495–503.
Tamura, A., Kimura, K., Takahara, H., & Akasaka, K. (1991a) *Biochemistry 30*, 11307–11313.
Tamura, A., Kimura, K., Takahara, H., & Akasaka, K. (1991b) *Biochemistry 30*, 11313–11320.
Tanaka, A., Flanagan, J. M., & Sturtevant, J. M. (1992) *Protein Sci.* (in press).
Taupin, D., & Luzzati, V. (1982) *J. Appl. Crystallogr. 15*, 289–300.
Ueki, T., Hiragi, Y., Kataoka, M., Inoko, Y., Amemiya, Y., Izumi, Y., Tagawa, H., & Muroga, Y. (1985) *Biophys. Chem. 23*, 115–124.
Yamakawa, H., & Fujii, M. (1974) *Macromolecules 7*, 649–654.