ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/227611728

15N backbone dynamics of the S-peptide from ribonuclease A in its free and S-protein bound forms: Toward a site-specific analysis of entropy changes upon folding

ARTICLE *in* PROTEIN SCIENCE · FEBRUARY 2008

Impact Factor: 2.85 · DOI: 10.1002/pro.5560070220 · Source: PubMed

CITATIONS
25

READS
12

6 AUTHORS, INCLUDING:

Andrei T Alexandrescu
University of Connecticut
74 PUBLICATIONS **2,217** CITATIONS
SEE PROFILE

Therese Flora Schulthess
University of Basel
52 PUBLICATIONS **3,378** CITATIONS
SEE PROFILE

Available from: Andrei T Alexandrescu
Retrieved on: 26 January 2016

Protein Science (1998), 7:389–402. Cambridge University Press. Printed in the USA.
Copyright © 1998 The Protein Society

^{15}N backbone dynamics of the S-peptide from ribonuclease A in its free and S-protein bound forms: Toward a site-specific analysis of entropy changes upon folding

ANDREI T. ALEXANDRESCU,[1] KLARA RATHGEB-SZABO,[1] KLAUS RUMPEL,[2] WOLFGANG JAHNKE,[3] THERESE SCHULTHESS,[2] AND RICHARD A. KAMMERER[2]

[1]Department of Structural Biology, Biozentrum, University of Basel, Basel, CH-4056 Switzerland
[2]Department of Biophysical Chemistry, Biozentrum, University of Basel, Basel, CH-4056 Switzerland
[3]Novartis Pharma AG, CTA/PSF, Basel, CH-4002 Switzerland

Abstract

Backbone ^{15}N relaxation parameters (R1, R2, ^{1}H–^{15}N NOE) have been measured for a 22-residue recombinant variant of the S-peptide in its free and S-protein bound forms. NMR relaxation data were analyzed using the "model-free" approach (Lipari & Szabo, 1982). Order parameters obtained from "model-free" simulations were used to calculate ^{1}H–^{15}N bond vector entropies using a recently described method (Yang & Kay, 1996), in which the form of the probability density function for bond vector fluctuations is derived from a diffusion-in-a-cone motional model. The average change in ^{1}H–^{15}N bond vector entropies for residues T3–S15, which become ordered upon binding of the S-peptide to the S-protein, is -12.6 ± 1.4 J/mol·residue·K. ^{15}N relaxation data suggest a gradient of decreasing entropy values moving from the termini toward the center of the free peptide. The difference between the entropies of the terminal and central residues is about -12 J/mol·residue·K, a value comparable to that of the average entropy change per residue upon complex formation. Similar entropy gradients are evident in NMR relaxation studies of other denatured proteins. Taken together, these observations suggest denatured proteins may contain entropic contributions from non-local interactions. Consequently, calculations that model the entropy of a residue in a denatured protein as that of a residue in a di- or tri-peptide, might over-estimate the magnitude of entropy changes upon folding.

Keywords: backbone entropy; conformational entropy; end effects; NMR relaxation; order parameter; protein folding; protein stability; random coil

Biological systems are inherently dynamic. A description of biological function that neglects dynamic aspects is thus neither entirely satisfying nor complete (Williams, 1989). Although the past decade has witnessed the transformation of NMR spectroscopy into a powerful technique for structure determination, it is perhaps in the area of molecular dynamics that NMR can make the most unique and significant contributions to an improved understanding of biological function. Molecular dynamics, for example, are intimately connected to NMR models of structure. Inasmuch as any type of NMR structural restraint such as the NOE (Neuhaus & Williamson, 1989; Williams, 1989), the scalar coupling constant (Smith et al., 1996), the chemical shift (Wishart et al., 1991), or residual dipolar coupling (Bax & Tjandra, 1997; Tolman et al., 1997) represents a time average over the ensemble of conformational sub-states, the precision and accuracy of an NMR structure will depend on the molecule's dynamics. It is of considerable interest to have estimates of both the average structure in solution, and of the amplitudes of fluctuations away from this mean. In practical endeavors such as rational drug design, it is often the case that the conformational freedom of a ligand and/or receptor is changed upon complex formation (Spolar & Record, 1994; Nicholson et al., 1995). Here, information on the location and magnitude of motional disorder in complexed and free states could lead to improved strategies for drug design (Wagner, 1995). A consensus appears to be emerging that the process of protein structure formation is best described in terms of the reduction in the number of accessible conformations upon protein folding (Baldwin, 1995). In principle, NMR can provide site-specific information on the degree of conformational variability along the polypeptide chain. Full exploitation of this information could lead to a more detailed description of the ordering of a polypeptide chain than in terms of the vague concepts of "folded" and "unfolded."

Reprint requests to: Andrei T. Alexandrescu, Department of Structural Biology, Biozentrum, University of Basel, Klingelbergstr. 70, CH-4056 Basel, Switzerland; e-mail: alexandrescu@ubaclu.unibas.ch.

In this paper we consider the fastest motions accessible by NMR spectroscopy, which are in the range between ps and ns. These motions are manifested through the relaxation of spins to equilibrium following excitation by a radio frequency pulse. More specifically, experimental NMR relaxation parameters reflect motions that induce fluctuations in the magnetic environments of spins with frequencies matching those of the transitions between nuclear states. The probability distribution of these frequencies is usually expressed in terms of the spectral density function $J(\omega)$. In contrast to the case for ^{1}H spins where dipolar interactions with any other ^{1}H spin within about 5 Å must be considered (Cavanagh et al., 1996), the relaxation of backbone ^{15}N spins is dominated by dipolar interactions with the directly bonded amide proton (r = 1.02 Å) and by chemical shift anisotropy (CSA). Relaxation of an ^{15}N nucleus can thus be treated in terms of a ^{15}N–^{1}H two spin approximation, and is a function of the values of the spectral density at five characteristic frequencies: $J(0)$, $J(\omega_N)$, $J(\omega_H)$, $J(\omega_H - \omega_N)$, and $J(\omega_H + \omega_N)$. Equations describing the relationship between the spectral density and experimentally observable relaxation parameters (e.g., R1, R2, NOE) are given in a number of works (Abragam, 1961; Lipari & Szabo, 1982; Wagner, 1993). The most widely used approach to extract motional parameters from the spectral density $J(\omega)$, is the "model-free" (*MF*) formalism (Lipari & Szabo, 1982). In this approach it is assumed that the spectral density can be factored into contributions from molecular tumbling and internal motion.

$$J(\omega) = \frac{2}{5}\left[(A^2)\frac{\tau_m}{1+(\omega\tau_m)^2} + (1-A^2)\frac{\tau_e}{1+(\omega\tau_e)^2}\right],$$

$$\tau_e = \frac{\tau_i\tau_m}{\tau_i+\tau_m}. \qquad (1)$$

Molecular tumbling is described by an effective global correlation time τ_m that pertains to all sites in the molecule, while internal motion is described in terms of a site-specific effective correlation time τ_i. The unit-less order parameter A^2 gives the relative contributions of τ_m and τ_i to the spectral density, and is a measure of the restriction of internal motion. $A^2 = 1$ implies that internal motion is completely restricted, $A^2 = 0$ implies that internal motion is completely unrestricted. The Lipari and Szabo (1982) formalism is "model free" in that it describes the time scales and spatial restriction of internal motion, without alluding to a specific model for internal motion (Lipari & Szabo, 1982). Historically, motional models were incorporated into equations for the spectral density, an approach that often entailed a more cumbersome analysis of the experimental relaxation data. Lipari and Szabo (1982) proposed that the unique information content of a given set of NMR relaxation data can be specified in terms of the time scale (τ) and spatial restriction (A^2) of internal motion. Once these parameters are determined they can be interpreted in terms of any physically reasonable model of motion (Lipari & Szabo, 1982). The need to interpret the relaxation data in terms of a motional model cannot be understated. Invariably, it is necessary to assume some reference frame to describe motion, such as an overall correlation time, or a reference frame that considers the motional properties of a given site relative to those of other sites in the protein. In the absence of models one is simply left with the raw relaxation data. Indeed the *MF* approach pre-supposes a model inasmuch as it is assumed that the spectral density can be factored into components from overall and internal motion. It is from this model that the power of the approach derives, in particular, in its description of the spatial restriction of internal motion in terms of the order parameter A^2.

Recent work has attempted to link the order parameter obtained from NMR relaxation experiments to conformational entropy (Akke et al., 1993; Li et al., 1996; Yang & Kay, 1996). These approaches consider that both the order parameter A^2 (Lipari & Szabo, 1982) and the statistical entropy of an ^{1}H–^{15}N bond vector "j" (Akke et al., 1993; Yang & Kay, 1996) are functions of the equilibrium probability density distribution that gives the probability that a vector j assumes a given orientation. Under the assumption that bond vector fluctuations are independent, the probability density function for individual bond vectors in the quantum mechanical formalism is given by

$$p^i(j) = \frac{e^{-E^i(j)/kT}}{\sum_i e^{-E^i(j)/kT}} \qquad (2)$$

where $E^i(j)$ is the quantized energy of state i, k is the Boltzmann constant, and T the absolute temperature. In the classical approach,

$$p(q) = \frac{e^{-U(q)/kT}}{\int_v e^{-U(q)/kT}\,dv} \qquad (3)$$

where q are the coordinates of vector j, $U(q)$ is the potential energy associated with these coordinates, and the integration $\int_v dv$ is over the ensemble of coordinate variables. The potential energy surface for motion of an ^{1}H–^{15}N bond vector in a macromolecule is unknown. Yang and Kay (1996) examined several simple motional models used in the analysis of NMR relaxation data to evaluate the relationship between A^2 and statistical entropy, using both quantum mechanical and classical derivations. For models where classical and quantum mechanical derivations were compared, it was found that both approaches gave similar results as long as $A^2 < 0.95$. Of the models considered, a model assuming bond vector diffusion in a cone appeared to be most consistent with the motion exhibited by ^{1}H–^{15}N bond vectors in a 1.12 ns molecular dynamics trajectory of RNase HI (Philippopoulos & Lim, 1995). In the classical formalism, the probability density function for this model is given by (Yang & Kay, 1996)

$$p(q) = \frac{1}{\phi_0(1-\cos\theta_0)} \qquad \text{for } 0 \le \theta \le \theta_0 \text{ and } 0 \le \phi \le \phi_0,$$

$$p(q) = 0 \qquad \text{for } \theta > \theta_0 \text{ or } \phi > \phi_0 \qquad (4)$$

where θ and ϕ are polar angles describing the orientation of the bond vector, and θ_0 is semi-angle of the cone. The relationship between the order parameter and the conformational entropy $\delta S_p(j)$ for this model $\phi_0 = 2\pi$ is given by (Yang & Kay, 1996):

$$\delta S_p(j)/R = \ln \pi[3 - (1+8A)^{1/2}] \qquad \text{for } 0 \le A^2 \le 1$$

$$\delta S_p(j)/R = \ln \pi[3 \pm (1+8A)^{1/2}] \qquad \text{for } 0 \le A^2 \le 1/64 \qquad (5)$$

where R is the gas constant, and A is the square root of the order parameter A^2. Equation 5 was applied to calculate changes in

^{1}H–^{15}N bond vector entropies upon unfolding of the N-terminal SH3 domain of the protein drk (Yang & Kay, 1996). In the present work, we have adopted the nomenclature $\delta S_p(j)$ rather than $S_p(j)$ of Yang and Kay (1996). The symbol "δ" is meant to emphasize that the entropies calculated from NMR order parameters reflect contributions to the entropy of an ^{1}H–^{15}N bond vector from a potential $U(q)$, rather than absolute entropies. Only differences in $\delta S_p(j)$ values "$\Delta\delta S_p(j)$" between a site in different conformational states, or between different sites in one conformational state can be meaningfully interpreted. In the present work we use a modified version of Equation 5:

$$\delta S_p(j)/R = \ln(\pi[3 - (1 + 8A)^{1/2}]) - \ln(4\pi) \quad \text{for } 0 \le A^2 \le 1$$

$$\delta S_p(j)/R = \ln(\pi[3 \pm (1 + 8A)^{1/2}]) - \ln(4\pi) \quad \text{for } 0 \le A^2 \le 1/64 \qquad (6)$$

to avoid possible confusion. The difference between Equations 5 and 6 is the inclusion of the constant term $-\ln(4\pi)$ on the right-hand sides of Equation 6, which reflects $\ln\{\int_v dv\}$, the upper bound on $\delta S_p(j)$ for two-dimensional internal rotation with $U(q) = 0$. Entropy values calculated by neglect of this constant (Equation 5) can be either negative or positive, reaching a maximum of $R * \ln(4\pi)$ for $A^2 = 0$ (Yang & Kay, 1996). The entropy values calculated by use of Equation 6 are shifted from those calculated by use of Equation 5 by $-R * \ln(4\pi)$, and are always negative. Although the difference between Equations 5 and 6 is semantic [the inclusion of the constant term $-R * \ln(4\pi)$ makes no difference in the calculation of a change in entropy], we feel that use of the latter equation better emphasizes that the entropy in the presence of a potential $U(q)$ is always more negative than in the absence of a potential $U(q) = 0$ (i.e., conformational freedom is always more restricted in the presence of a potential).

A number of caveats in the interpretation of $\Delta\delta S_p(j)$ should be noted. The applicability of the *MF* analysis to flexible polypeptides remains controversial; we explore some of the issues involved in Discussion. The *MF* approach is sensitive only to motions on time scales between ps and ns. Consequently, entropic contributions from conformational variability on slower time scales will not be detected (Akke et al., 1993; Yang & Kay, 1996). An association reaction such as the one between S-peptide and S-protein contains an unfavorable entropy term (ΔS_{rt}) due to the reduction in rotational and translational degrees of freedom on complex formation. The value of ΔS_{rt} is expected to be relatively insensitive to the size of the associating species, and is estimated to be on the order of -210 J/mol K (Finkelstein & Janin, 1989; Spolar & Record, 1994). Because A^2 order parameters only pertain to the amplitudes of internal motions, however, $\Delta\delta S_p(j)$ values are not expected to be sensitive to global entropic contributions such as ΔS_{rt}. The present work only addresses the entropies of a subset of sites of the system. The total entropy change between two states, however, contains terms from all components in the system: other bonds in the backbone, side chains, and solvent (Akke et al., 1993; Yang & Kay, 1996). The issue of separability of the total entropy into components, which concerns cross-correlations between sites, remains a subject of controversy (Mark & Van Gunsteren, 1994; Boresch & Karplus, 1995; Brady et al., 1996; Brady & Sharp, 1997).

The principal aim of this work is to reach a better understanding of the factors contributing to the $\delta S_p(j)$ values of ^{1}H–^{15}N bond vectors. Our approach to this aim is to correlate the entropies obtained from NMR relaxation data with estimates of conformational entropy derived from other methods. For these comparisons we have chosen as a model system the well characterized S-peptide of ribonuclease A.

Results

In the presence of subtilisin the 124 residue protein RNAse A is cleaved between residues 20 and 21 (Fig. 1). The two chromatographically separable components of this reaction (Richards & Vithayathil, 1959) are referred to as the S-protein (residues 21–124) and the S-peptide (residues 1–20). The S-peptide and S-protein reassociate in stoichiometric amounts to form a tight non-covalent "S-complex" ($K_D < 10^{-7}$ M at pH 6.8, 25 °C) that has full enzymatic activity (Schreier & Baldwin, 1977), and a structure that is very similar to that of wild-type RNAse A (Kim et al., 1992). Studies on modified S-peptides indicate that deletion of residues 16–20 has no observable effect on the S-peptide's ability to reassociate with the S-protein into a fully active S-complex. Deletions

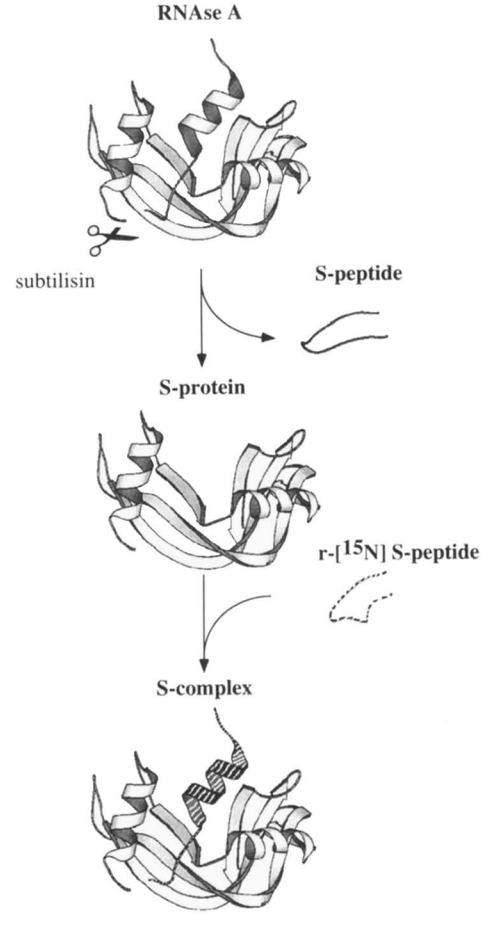

Fig. 1. Molscript (Kraulis, 1991) diagrams based on the X-ray coordinates of RNase A (Howlin et al., 1989), illustrating the relationships between S-peptide, S-protein, and S-complex. The stippled ribbons indicate the recombinant ^{15}N-labeled S-peptide.

upstream of S15, however, causes a marked reduction in binding affinity (Potts et al., 1963). A study of S-peptides in which residue R10 was replaced by ornithine, concluded that deletion of residues K1–E2 leads to a loss in activity of 20%, while deletion of the first three N-terminal residues (K1, E2, T3) leads to a drop in activity in excess of 70% (Moroder et al., 1969). The residues that are dispensable for tight complex formation (K1–E2 and S16–A20) are disordered in the X-ray structure of the S-complex (Kim et al., 1992).

For the present work we studied an ^{15}N-labeled 22-residue recombinant variant of the S-peptide. This peptide was synthesized as part of a heterologous fusion protein in *Escherichia coli*, and contains two additional residues at its N-terminus due to a cloning artifact. The two foreign residues are included in the numbering scheme of this paper, so that a −2 residue offset is introduced relative to the sequence of the wild-type S-peptide (e.g., residue F10 in our peptide corresponds to F8 in the wild-type sequence). ^{1}H-NMR assignments have been previously reported for a version of the free S-peptide without A20 (Gallego et al., 1983) and for intact RNase A (Rico et al., 1989; Robertson et al., 1989). For the present study of backbone dynamics it was necessary to obtain ^{15}N-NMR resonance assignments. Sequence-specific assignments for the free S-peptide (Fig. 2A and Electronic Appendix) were obtained from gradient 3D ^{1}H–^{15}N TOCSY-HSQC and NOESY-HSQC (150 ms NOE mixing time) spectra (Marion et al., 1989; Kay, 1995) collected with a reduced spectral width of 800 Hz, and only eight complex points in the ^{15}N dimension. The chemical shift dispersion of the bound S-peptide is considerably larger (Fig. 2B), and 2D ^{15}N-edited TOCSY and NOESY (300 ms NOE mixing time) sufficed for resonance assignments. The ^{1}H chemical shifts obtained for the free S-peptide are in good agreement with those of Gallego et al. (1983), except for residues K3, E4, S18, and A21. These differences are probably due to the different N- and C-termini of the two peptides. The chemical shifts of aliphatic protons in the bound S-peptide also show good agreement with those for intact RNAse A (Robertson et al., 1989), except for residues downstream of S16. These differences are likely to reflect the break in the polypeptide chain between residues 20 and 21 in the complex, compared to intact RNAse A. The agreement between ^{1}HN chemical shifts in the complex and intact RNAse A is somewhat poorer, with differences as large as 0.2 ppm, possibly a result of the higher sensitivity of ^{1}HN chemical shifts to solution conditions such as pH and temperature.

The S-peptide, which includes the N-terminal helix of native RNAse A (residues 3–13), is reported to have a fractional population of α-helical structure of about 30% at a temperature of 0 °C, pH 3.8, 0.1 M NaCl (Kim & Baldwin, 1984). At 25 °C the S-peptide is reported to have the characteristics of a random coil (Kim & Baldwin, 1984). Based on an analysis of the chemical shifts of the F8 δ protons in a 1–19 variant of the S-peptide, Rico et al. (1986) estimated helix molar fractions of about 24% at pD 2.5, and 31% at pD 5.4, for a temperature of 10 °C, and in the presence of 1 M NaCl. It is important to note that in the present study both the free peptide and the complex were studied in deionized (Milli-Q) water. Low ionic strength destabilizes α-helical structure in the free peptide (Brown & Klee, 1969). The NOE data for the free S-peptide in the present study show very strong $d\alpha N(i,i+1)$ and weak $dNN(i,i+1)$ NOEs. No NOEs between residues separated by more than one position in the sequence were detected, suggesting vanishingly small populations of for α-helical structure. These observations, however, could be due to the low sensitivity of the NOESY experiment to small populations of α-helical structure. Furthermore, the small size of the peptide places it in near the $\omega\tau_c \approx 1$ regime where the homonuclear ^{1}H NOE tends to zero (Neuhaus & Williamson, 1989). It is notoriously difficult to estimate fractional populations of α-helical structure from NOE or ^3J coupling constant data (Bonvin & Brünger, 1996). Chemical shift data might be more sensitive to the presence of small populations of α-helical structure. With the exception of residue K9, which immediately precedes a phenylalanine, the Hα proton chemical shifts for the free S-peptide are all within ±0.1 ppm of "random coil" values (Wishart et al., 1995). Residues T5–M15, which correspond to the α-helix between T3 and M13 in the native structure of ribonuclease A (Rico et al., 1989), however, show systematic upfield ^{1}Hα chemical shift deviations from random coil values (Fig. 3A). These deviations are characteristic of α-helical struc-

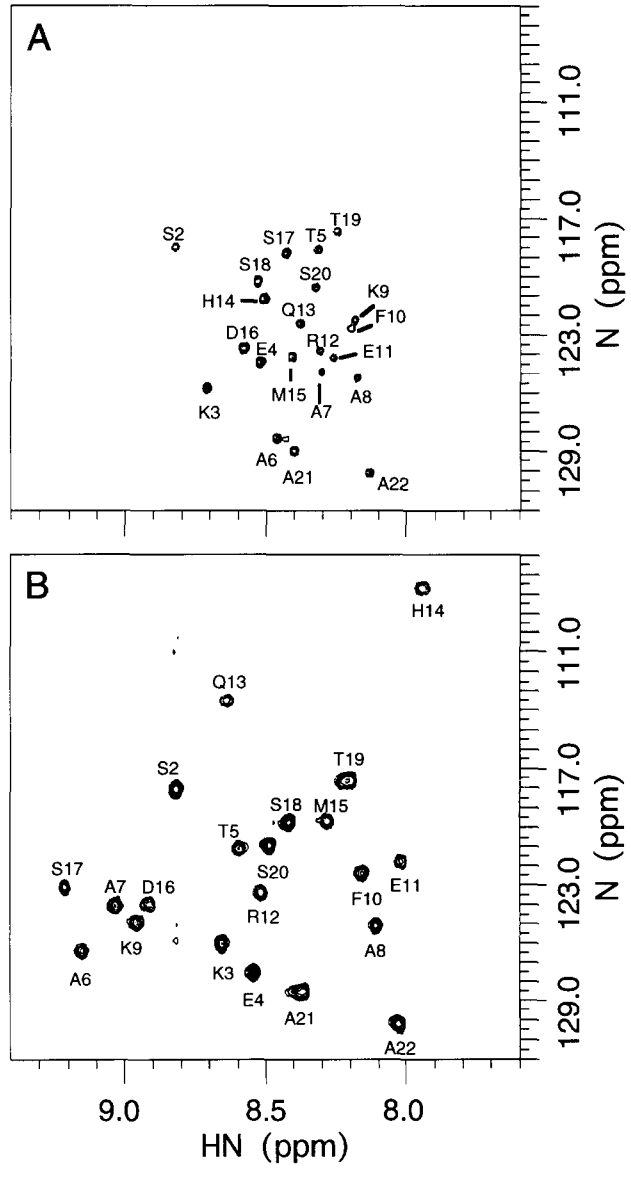

Fig. 2. ^{1}H–^{15}N gradient-HSQC spectra of (**A**) free and (**B**) S-protein bound S-peptide. For both the free and the bound S-peptide all backbone amide protons except for the N-terminal residue G1 could be detected.

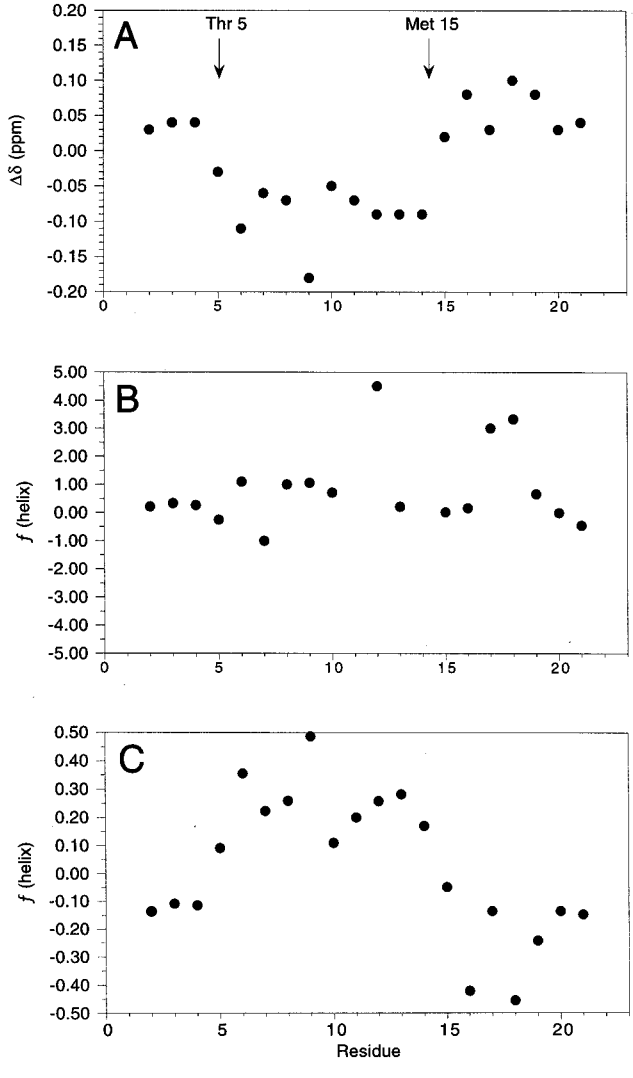

Fig. 3. ^{1}Hα conformational shifts of the free S-peptide. **A:** Difference in ^{1}Hα chemical shifts between the S-peptide and the "coil" values of (Wishart et al., 1995). **B:** Fractional populations of α-helix [f(helix)] calculated from Equation 7 by taking δ(helix) as the chemical shift values of the bound S-peptide. **C:** Fractional populations of α-helix obtained by taking δ(helix) values from a database of ^{1}Hα chemical shifts in α-helical proteins (Wishart et al., 1991; Table 6).

ture. An estimate of the fractional population of α-helical structure in the free S-peptide can be arrived at from the formula (Alexandrescu et al., 1994):

$$f(\text{helix}) = \frac{\delta(\text{observed}) - \delta(\text{coil})}{\delta(\text{helix}) - \delta(\text{coil})} \quad (7)$$

where δ(observed) are the chemical shifts of the S-peptide; δ(helix), and δ(coil) (Wishart et al., 1995) are limiting α-helix and random coil chemical shifts, respectively. If δ(helix) are taken as the chemical shifts of the bound peptide, the data are not readily interpretable in terms of the amount of residual α-helical structure (Fig. 3B). Presumably, the ^{1}Hα conformational shifts of the peptide in the complex are masked by large chemical shift contributions that are not related to α-helical structure (e.g., ring current shifts). Figure 3C shows fractional populations of α-helix calculated by taking δ(helix) values from a larger database of ^{1}Hα chemical shifts in α-helical proteins (Wishart et al., 1991). The data in the later analysis suggest a small population of α-helical structure for residues T5–M15 in the free S-peptide. The average of f(helix) over these residues is 22%, with a standard error of 4%. The ^{1}Hα chemical shift data suggest that the population of α-helix for residues T5–M15 might be as high as 20 to 25% in the free S-peptide under the conditions of this study. The uncertainty of this estimate is, however, very large. Note that residues outside of the putative helix have down-field chemical shift deviations of similar magnitude to the up-field deviations of T5 to M15. Even if the population of α-helix in the free S-peptide is as high as 30%, as suggested by experiments at 0 °C (Kim & Baldwin, 1984), the relaxation parameters in the present study should predominantly reflect contributions from the large proportion (>70%) of non α-helical conformers.

Relaxation parameters for the free and bound S-peptide are summarized in Figure 4. One factor contributing to the different re-

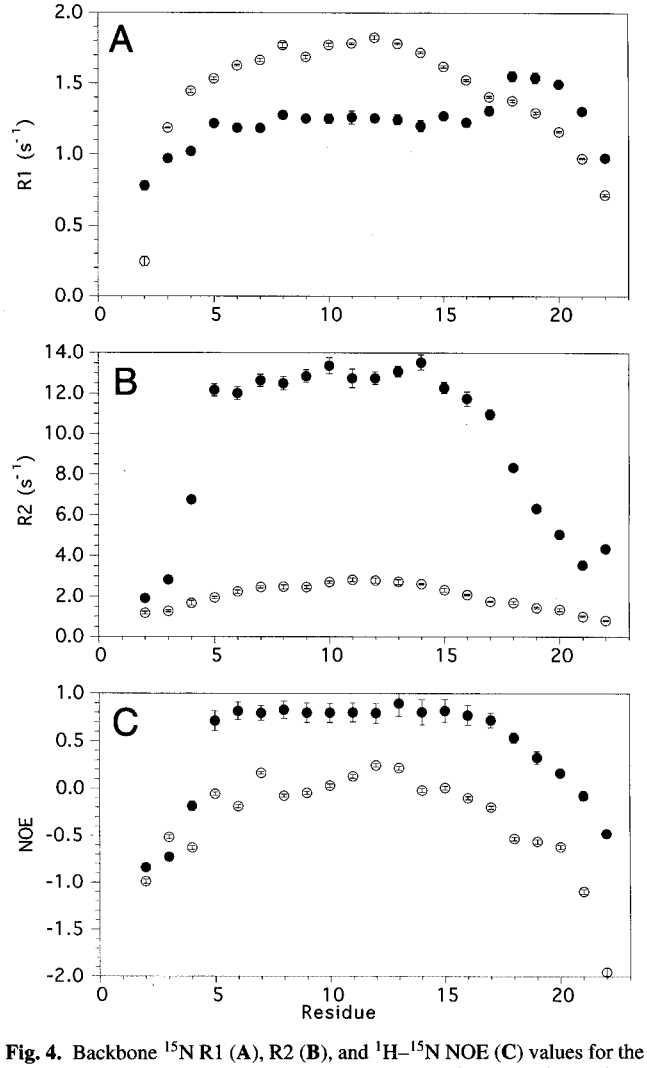

Fig. 4. Backbone ^{15}N R1 (**A**), R2 (**B**), and ^{1}H–^{15}N NOE (**C**) values for the free (open circles) and bound (closed circles) S-peptide. Experimental uncertainties (taken as the standard errors of the fits) are given for all data points; however, in some cases these are smaller than the sizes of the symbols.

laxation parameters for the two forms are the different molecular masses of the free S-peptide (2.3 kDa) and S-complex (13.7 kDa). Another factor, reflected in the variability of relaxation parameters as a function of sequence, is the different mobilities of individual sites within each species. The relaxation data were analyzed in terms of the *MF* formalism (Lipari & Szabo, 1982). The first step in this analysis is to obtain an estimate for the value of the global correlation time τ_m. This estimate can be obtained from ratios of R2/R1 values, under the assumption that R2 does not contain contributions from exchange broadening ($R2_{ex}$) and that $\omega\tau_e \ll 1$ (Kay et al., 1989). For the bound S-peptide we obtained a trimmed mean R2/R1 value of 10.0 ± 0.6 after excluding residues with NOE values below 0.7, and R2/R1 values outside one standard deviation of the mean. This R2/R1 value corresponds to a τ_m of 9.7 ± 0.1 ns for the complex. For the free S-peptide we obtained a trimmed mean R2/R1 of 1.33 ± 0.04 after excluding residues with R2/R1 ratios outside one standard deviation of the mean, corresponding to a τ_m of 1.5 ± 0.1 ns.

Alternatively, we performed a grid search in τ_m for each residue while optimizing A^2 and τ_i (Farrow et al., 1997). This procedure assumes that the spectral density for each residue is that of Equation 1 and yields an effective $\tau_{m,\text{local}}$ for each residue. Trimmed mean $\tau_{m,\text{local}}$ values were calculated by excluding values outside one standard deviation of the mean, and residues that gave Γ^2 values for the goodness of fit greater than zero (Mandel et al., 1995). Using this procedure, trimmed mean $\tau_{m,\text{local}}$ values of 1.7 ± 0.1 ns and 9.4 ± 0.2 ns were obtained for the free and bound S-peptide, respectively. These later values are highly similar to the τ_m values of 1.5 ± 0.1 ns and 9.7 ± 0.1 ns estimated from R2/R1 ratios.

Using initial guess values of 1.7 ns and 9.4 ns for the global τ_m's of the free and bound S-peptide, respectively, we calculated motional parameters for five different models of the spectral density (Mandel et al., 1995). *Model 1* assumes that $\tau_i \rightarrow 0$, and is valid if internal motions are on a timescale faster than 20 ps. For this model Equation 1 reduces to

$$J(\omega) = \frac{2}{5}\left[A^2 \frac{\tau_m}{1 + (\omega\tau_m)^2}\right]. \tag{8}$$

Model 2 is equivalent to *Model 1* with the inclusion of an $R2_{ex}$ term to correct for possible exchange broadening contributions to R2 (Clore et al., 1990a)

$$R2 = R2_{dd} + R2_{ex}. \tag{9}$$

Model 3 is the original *MF* formulation (Equation 1). *Model 4* is equivalent to *Model 3* with the inclusion of an $R2_{ex}$ term. *Model 5* corresponds to the Clore et al. (1990a) extension

$$J(\omega) = \frac{2}{5}\left[\frac{(A_f^2 A_s^2)\tau_m}{1 + (\omega\tau_m)^2} + \frac{A_f^2(1 - A_s^2)\tau_s'}{1 + (\omega\tau_s')^2} + \frac{(1 - A_f^2)\tau_f'}{1 + (\omega\tau_f')^2}\right],$$

$$\tau_s' = \frac{\tau_s\tau_m}{\tau_s + \tau_m},$$

$$\tau_f' = \frac{\tau_f\tau_m}{\tau_f + \tau_m} \tag{10}$$

of the Lipari & Szabo (1982) formalism, which assumes internal motion on fast (τ_f) and slow (τ_s) time scales, which differ by more than an order of magnitude ($\tau_m > \tau_s > \tau_f$). Assuming that $\tau_f \rightarrow 0$ gives

$$J(\omega) = \frac{2}{5}\left[\frac{(A_f^2 A_s^2)\tau_m}{1 + (\omega\tau_m)^2} + \frac{A_f^2(1 - A_s^2)\tau_s'}{1 + (\omega\tau_s')^2}\right],$$

$$\tau_s' = \frac{\tau_s\tau_m}{\tau_s + \tau_m}. \tag{11}$$

The goodness of fit for each model was assessed by the statistical criteria described by Mandel et al. (1995).

Once the optimal model was assigned for each residue, a final optimization was performed for the free parameters within each model, and for the value of the global correlation time τ_m. The motional parameters A^2 and τ_i obtained from these calculations, which assumes a global correlation time τ_m, are summarized in Figure 5. The optimized values of 1.62 ± 0.04 ns and 9.63 ± 0.05 ns for the global τ_m's of the free and bound S-peptide, respectively, were close to the initial guess values.

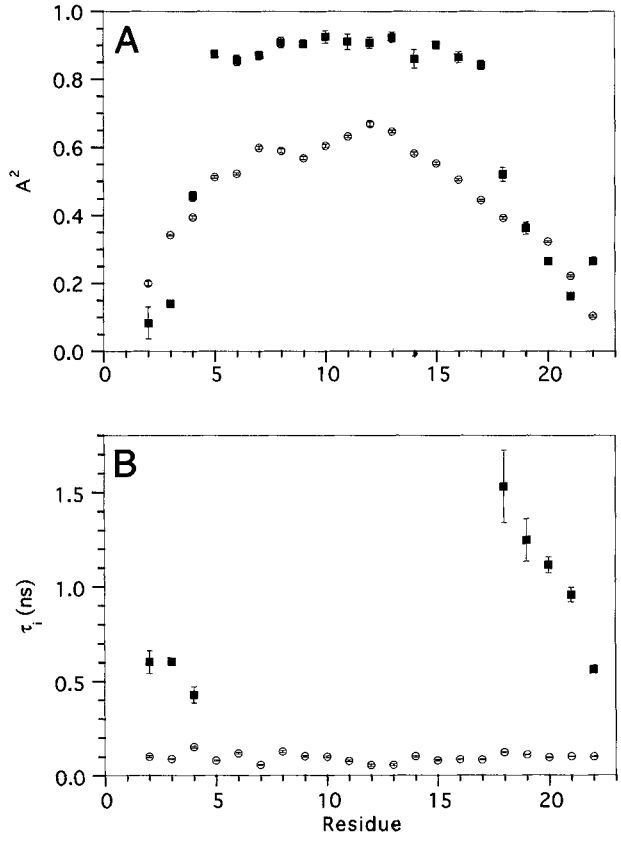

Fig. 5. Results of the MF analysis using a global correlation time, τ_m. Values for the generalized order parameters $A^2 = A_f^2 A_s^2$ (**A**), and correlation times for internal motion τ_i (**B**). Data for the free and bound S-peptide are indicated by open circles and closed squares, respectively. In the bound S-peptide all residues experience motion on a time scale faster than 20 ps with an amplitude of A_f^2. Residues S2–K4 and S18–A22 of the bound S-peptide experience an additional slower motion with an amplitude of A_s^2 on the ns time scales indicated by the filled squares in Figure 5B.

In the case of the free S-peptide, *Model 3* was the optimal model for all residues except F10 and H14. These residues were fit by *Model 4* with very small $R2_{ex}$ contributions of 0.33 ± 0.05 Hz and 0.32 ± 0.04 Hz, respectively. Residues S2, K3, and T19 gave high χ^2 values above 30, indicating that all of the models considered gave a poor fit to the experimental relaxation data. To test whether a single model could account for the relaxation data of the free S-peptide, an additional calculation was performed with all residues constrained to *Model 3*. The τ_m value obtained from this calculation was 1.70 ± 0.03 ns, and the values for A^2 and τ_i were closely similar to those obtained when residues F10 and H14 were treated according to *Model 4*. The number of residues with χ^2 values above 30, however, increased from three to six. We therefore concluded that the appropriate model for residues F10 and H14 of the free S-peptide is *Model 4*. For the bound S-peptide, residue H14 was best fit by *Model 2*, and the remainder of residues T5–S15 by *Model 1*. Residues S2–E4 and S18–A22, which are dispensable for complex formation (Potts et al., 1963; Moroder, 1969), were best-fit by *Model 5*.

Discussion

On the applicability of the "model-free" analysis to flexible polypeptides

In the *MF* formalism, NMR relaxation data are analyzed in the framework of a single correlation time for overall molecular reorientation (Lipari & Szabo, 1982). Yet the assumption of a single global correlation time presents an immediate paradox. In principle, any motional disorder should lead to a distribution of conformations, and thus to a distribution of correlation times for molecular tumbling (Miyake, 1958). The difficulties associated with describing flexible molecules, or flexible portions of molecules, in terms of a single global correlation time have long been appreciated (Torchia et al., 1975; Lipari & Szabo, 1982; vanMierlo et al., 1993; Alexandrescu & Shortle, 1994; Frank et al., 1995; Yang & Kay, 1996; Farrow et al., 1997; Kemple et al., 1997). Recently, Wagner and co-workers (Peng & Wagner, 1995; Lefevre et al., 1996) have stated that the use of the *MF* analysis is inappropriate for molecules where there is no well-defined correlation time. They proposed that in these cases it might be more profitable to discuss dynamics in terms of raw relaxation rates or spectral density functions (Lefevre et al., 1996).

Distributions of correlation times were first invoked to account for deviations of experimental relaxation data (such as dielectric relaxation) from those predicted by a model of a rigid isotropic rotor[4] (Connor, 1964). In NMR spectroscopy distributions of correlation times have been used to account for disagreements between experimental T2, T1, and NOE data, at a single field or multiple fields, from those predicted by models which incorporate a single τ_m (Schaefer, 1973; Levy et al., 1978; Wittebrot et al., 1980). Up to the early 1980s, at least nine types of correlation time distribution functions were proposed (Connor, 1964; Schaefer, 1973; Levy et al., 1978; Wittebrot et al., 1980). It seems fair to surmise, however, that in the absence of experimental methods to characterize such distributions; the shape, asymmetry, and width of a correlation time distribution for any real molecule remains unknown. Lipari & Szabo (1982b) proposed that the "double-δ-function"

$$p(\tau) = B\delta(\tau - \tau_1) + (1 - B)\delta(\tau - \tau_2), \tag{12}$$

where $0 \leq B \leq 1$ and δ is Dirac's delta function, can reproduce experimental relaxation data as well as distribution functions of a more complicated form. It is unclear, however, if this is more a reflection of the insensitivity of the correlation function to the particular form of the distribution (Lipari & Szabo, 1982) or a result of characterizing more complicated spectral density functions in terms of a limited number of relaxation parameters. The resulting spectral density for the "double-δ-function" distribution of correlation times is given by (Lipari & Szabo, 1982b)

$$J(\omega) = \frac{2}{5}\left[\frac{BA^2\tau_1}{1 + (\omega\tau_1)^2} + \frac{(1 - B)A^2\tau_2}{1 + (\omega\tau_2)^2} + \frac{B(1 - A^2)\tau'}{1 + (\omega\tau')^2} + \frac{(1 - B)(1 - A^2)\tau''}{1 + (\omega\tau'')^2}\right],$$

$$\tau' = \frac{\tau_i\tau_1}{\tau_i + \tau_1},$$

$$\tau'' = \frac{\tau_i\tau_2}{\tau_i + \tau_2}. \tag{13}$$

The spectral density of Equation 13 is isomorphous with that of the two-time scale model of Clore et al. (1990a). Setting $B = A_s^2$, $A^2 = A_f^2$, $\tau_1 = \tau_m$, $\tau_2 = \tau_s' = \tau_s\tau_m/(\tau_s + \tau_m)$, $\tau_i = \tau_f \rightarrow 0$ in Equation 13, gives Equation 11. More generally, models that assume a distribution of correlation times may be mathematically indistinguishable from models that assume motion on different time scales, because both lead to nonexponential correlation functions (Connor, 1964; Levy et al., 1978; Clore et al., 1990b). Indeed, the two-time scale model of Clore et al. (1990a, 1990b) was initially proposed to account for an under-estimation of experimental NOE values by the simpler model of Equation 1. Similar discrepancies between NOE and T1 values, led Schaefer (1973) to a model that assumes a distribution of correlation times.

Whether there are any decisive criteria that determine the types of molecules for which a *MF* analysis is appropriate (Lefevre et al., 1996) is an open question. Residues 5–17 of the "refolded" bound S-peptide can be fit to models (*Models 1* and *2*), which assume a single effective correlation time (τ_i) for internal motion (for some proteins >90% of sites show this type of behavior). One can derive some degree of comfort from the statistics that indicate that most of residues in the bound S-peptide can be adequately described in terms of an exponential correlation function. Residues 2–4 and 18–22, however, require a two-time scale model (*Model 5*). The relaxation data for these residues indicate a nonexponential correlation function, which can be interpreted either in terms of a distribution of overall correlation times, or in terms of a single overall correlation time with internal motion on multiple time scales. The key point is that the simple behavior of residues 5–17 may have little bearing on the nonexponential correlation functions of resi-

[4]It is important to distinguish between anisotropic overall motion of a rigid non-spherical molecule (e.g., a rigid ellipsoid) from anisotropic overall motion due to a distribution of correlation times in a flexible molecule (Lipari & Szabo, 1982). Operationally, both lead to a correlation function that is more complex than that for a spherical rigid rotor.

dues 2–4 and 18–22.[5] Perhaps even an apparent exponential correlation function is not conclusive evidence for a single τ_m, or for a single time scale for internal motion. In the case of the free S-peptide all of the residues are fit by either *Models 3* or *4*, which correspond to an exponential correlation function. In general, it is not possible to exclude that a finer sampling of the correlation function might indicate a nonexponential shape that is not evident in a coarser sampling. These concerns about nonexponential correlation functions may not be unique to the *MF* formalism. To interpret raw relaxation data or spectral density functions in terms of mobility, it is necessary to have a reference frame. Relaxation parameters for any given molecule will depend on factors such as solvent viscosity, temperature, or the subdivision of structure into domains. These factors need to be corrected for to analyze the degree of restriction of internal motion, for example, by resorting to a model that assumes overall tumbling (Lefevre et al., 1996). The normalization of raw relaxation data or spectral densities in terms of overall tumbling is probably no more straightforward than that used for the *MF* analysis.

A possible resolution to these issues is to consider that the analysis of experimental relaxation parameters in terms of a distribution of τ_m's might be equivalent to an analysis in terms of internal motion in the context of a single τ_m. If motion occurs on a time scale much faster than overall tumbling, it should be possible to define an average correlation time for overall tumbling. On the slower time scale of molecular tumbling this average could correspond to the effective average of a distribution of correlation times. This average τ_m may not pertain to any single conformation in the distribution, much in the same way as an average over an ensemble of NMR structures may be a poor representation of any individual conformation. The utility of an average τ_m is to assess the extent of deviations from this average, much in the same way as the RMSD from the average can reflect the breadth of conformational heterogeneity in an ensemble of NMR structures. Operationally, it is possible to define such an effective average τ_m for both the free and the bound S-peptide. *MF* simulations using global τ_m values of 1.62 and 9.63 ns for the free and bound S-peptides, respectively, reproduce the experimental ^{15}N relaxation data for most of the residues in the peptide (see Results).

A possible advantage of the *MF* approach is that in many cases A^2 order parameters evince a correlation with other indicators of conformational variability. Examples include, in particular RMSD values of NMR structures (Redfield et al., 1992; Buck et al., 1995; Alexandrescu et al., 1996; Fushman et al., 1997); and other structural parameters such as ^3J coupling constants, and chemical shift deviations from "coil" values (Buck et al., 1995; Alexandrescu et al., 1996). The question of whether a quantitative correspondence exists has not yet been examined in detail; nevertheless, a qualitative relationship is suggested. A distinct disadvantage of models that invoke a distribution of correlation times is the nearly complete lack of data on the range of conformations, or on the range of correlation times in flexible polypeptides. In the absence of such data, *Occam's Razor* would seem to dictate the use of the simplest analysis consistent with the experimental data. At the present time, we believe the *MF* approach is the simplest method to extract information on the restriction of motion in polypeptides from NMR relaxation data.

[5] A similar conclusion for residues treated by *Model 5* in ribonuclease HI appears to have been reached by Mandel et al. (1995).

Interpretation of NMR-derived H–N bond vector entropies

$\delta S_p(j)$ values calculated from Equation 6 using the order parameters from the *MF* analysis (Fig. 5A) are shown in Figure 6A and B for the free and bound S-peptide, respectively. The differences in $\delta S_p(j)$ values between the free and bound S-peptide are summarized in Figure 6C.

Thermodynamic changes associated with binding of the S-peptide to S-protein have been studied by calorimetry (Hearn et al., 1971; Varadarajan et al., 1992). In a seminal paper on entropic changes associated with induced fit binding, Spolar and Record (1994) attributed a value of −400 J/mol·K to the change in the conformational entropy ($\Delta S_{conf} = S_{folded} - S_{unfolded}$) of the S-peptide on binding to the S-protein. This estimate was based on an extrapolation of the calorimetric data of Varadarajan et al. (1992) to a temperature of −20 °C (where $T\Delta S_{total} = 0$), and correction for entropic terms other than ΔS_{conf}. For this calculation it was assumed that the structure of the S-protein is the same in its free and bound states, and that 17 residues in the S-peptide fold on binding with an average ΔS_{conf} of −23 J/mol·residue·K. The $\delta S_p(j)$ values in Figure 6, together with the information on the minimal peptide epitope required for binding (Potts et al., 1963; Moroder, 1969), however, suggest that only 13 residues in the S-peptide

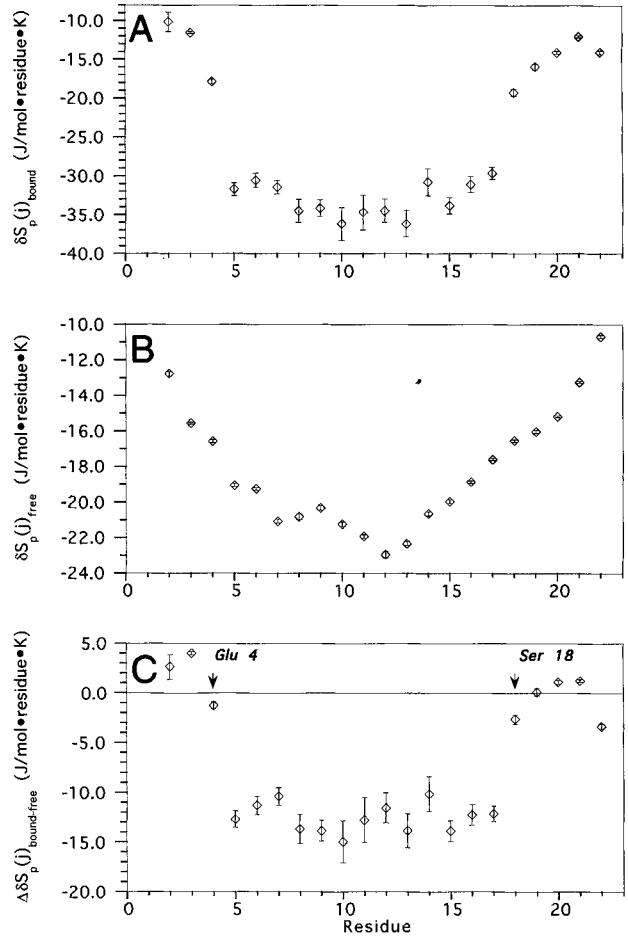

Fig. 6. ^1H–^{15}N bond vector entropies [$\delta S_p(j)$] for the bound (**A**) and free (**B**) S-peptide, and (**C**) the difference in conformational entropy [$\Delta\delta S_p(j) = \delta S_p(j)_{unfolded} - \delta S_p(j)_{folded}$].

become ordered on binding. The S-protein has a considerably smaller chemical shift dispersion than wild-type RNAse A (ATA, unpubl. obs.). This, together with the much smaller protection from solvent exchange afforded by the S-protein compared to intact RNase A (Rosa & Richards, 1981), suggest that the S-protein is more flexible than RNase A. Formation of the S-complex may involve an ordering of both the S-peptide and the S-protein components, or a mutual induced fit. In this case, the value of −400 J/mol·K attributed to the change in conformational entropy of the S-peptide on binding might represent an over-estimate.

The total change in configurational entropy upon folding contains contributions from both conformational and vibrational degrees of freedom $\Delta S_{config} = \Delta S_{conf} + \Delta S_{vib}$ (Karplus & Kushick, 1981; Karplus et al., 1987; Doig & Sternberg, 1995). The conformational entropy S_{conf} is a function of the number of conformations, or potential energy wells. For example, S_{conf} could reflect three backbone conformations with ϕ angles of −140, −120, and −100 degrees, respectively. The vibrational entropy S_{vib} describes local fluctuations in the neighborhood of a well-defined conformation; the width of a potential energy well. For example, S_{vib} could reflect fluctuations of ±5° about an energy minimum defined by a backbone ϕ angle of −120°.[6] Based on molecular dynamics simulations, S_{vib} in folded proteins is an order of magnitude larger than S_{conf} (Karplus et al., 1987). The contributions from S_{vib}, however, may be of a similar magnitude for folded and unfolded proteins (Karplus et al., 1987; Doig & Sternberg, 1995). In this case the difference in entropy between two states will only reflect changes in conformational entropy.

The change in conformational entropy upon folding contains contributions from the restriction of side chain and backbone conformations ($\Delta S_{conf} = \Delta S_{conf,sc} + \Delta S_{conf,bb}$). $\Delta S_{conf,sc}$ is believed to be about −14 J/mol·residue·K (Doig & Sternberg, 1995). Assuming separability of the two contributions, $\Delta\delta S_p(j)$ for a backbone ^1H–^{15}N bond vector should reflect $\Delta S_{conf,bb}$. The conformational entropy of an ^1H–^{15}N bond vector j is a function of the distribution of the torsional angles ϕ_j and ψ_{j-1}, because rotation about the C′–N bond is restricted by its partial double-bond character. We thus examined the possibility of a relationship between $\Delta\delta S_p(j)$ values obtained from NMR, and $\Delta S_{conf,bb}$ estimated from changes in the probability distributions of ϕ,ψ angles on folding (Brady & Sharp, 1997).

Estimates of $\Delta S_{conf,bb}$ based on changes in ϕ,ψ angle distributions upon folding range between about −19 J/mol·residue·K and −30 J/mol·residue·K (Yang & Honig, 1995; D'Aquino et al., 1996; Wang & Purisima, 1996; Zhang et al., 1997). In these calculations the distribution of ϕ,ψ angles for a residue in the unfolded state is modeled according to that of a residue in a di- or tri-peptide. Implicit to these calculations is the assumption that the conformational entropy of an unfolded protein is devoid of contributions from long-range interactions (Yang & Honig, 1995; D'Aquino et al., 1996; Wang & Purisima, 1996; Zhang et al., 1997). If we calculate $\Delta\delta S_p(j)$ as the difference $\delta S_p(j)_{bound} - \delta S_p(j)_{free}$, we obtain an average value of −12.6 ± 1.4 J/mol·residue·K, for residues T5–S17, which make up the binding epitope of the S-peptide (Fig. 6C). This value is highly similar to the average $\Delta\delta S_p(j)$ of −12 J/mol·residue·K, obtained from the difference in $\delta S_p(j)$ values between the folded and unfolded forms of the N-terminal SH3 domain of drk (Yang & Kay, 1996). The magnitude of $\Delta\delta S_p(j)$ in the present study and in the study of Yang & Kay (1996) is only about half of that estimated for $\Delta S_{conf,bb}$ from di- and tri-peptide models. One of the advantages of the S-peptide as a model system is that the bound form of the S-peptide contains both ordered and disordered segments of polypeptide chain. The studies of Potts et al. (1963) and Moroder et al. (1969) indicate that residues T5–S17 are sufficient for tight complex formation. The average of the $\delta S_p(j)$ values for the remaining disordered residues S2–E4, and S18–A22 in the bound S-peptide is −14.3 J/mol·residue·K. If we take this average as representative for $\delta S_p(j)_{unfolded}$, we obtain a higher average $\Delta\delta S_p(j)$ of −18.7 ± 2.2 J/mol·residue·K for residues T5–S17. Furthermore, if we take the $\delta S_p(j)$ value of the N-terminal residue S2 in the bound S-peptide (−10.2 J/mol·residue·K) as representative of $\delta S_p(j)_{unfolded}$, we obtain an even higher average $\Delta\delta S_p(j)$ of −22.9 ± 2.2 J/mol·residue·K for residues T5–S17. We conclude that a possible source for the discrepancy between $\Delta S_{conf,bb}$ values and $\Delta\delta S_p(j)$ values obtained by NMR is a systematic reduction in the entropy of the free S-peptide proceeding from the ends to the center of the chain. The difference in entropy between the terminal and central residues is about −10 to −12 J/mol·residue·K (Figs. 6B, 7), a value comparable to the change in $\delta S_p(j)$ between the free and bound S-peptide. The values of $\Delta\delta S_p(j)$ approach those for $\Delta S_{conf,bb}$ (estimated by taking the entropy of a residue in an unfolded protein to be equivalent to that in a di- or tri-peptide) only when the entropies of the N- and C-termini of the free or bound S-peptide are taken to represent those of the unfolded state. The assumption that long-range interactions do not contribute to the entropy of the free S-peptide does not appear to be valid.

The differences in ^1H–^{15}N bond vector entropies between individual positions and the center of the free S-peptide ($\Delta S_{i\text{-center}}$) are shown by the filled circles in Figure 7A. The crosses in Figure 7A represent values simulated for the free S-peptide according to a model in which the rate of entropy increase as a function of displacement from the center of the chain (residue 12) is given by a parameter μ, and the shape of the $\Delta S_{i\text{-center}}$ profile is given by $n_i/(N - n_i)$, with N the total number of bonds in the chain, and n_i the displacement in number of bonds from the center of the chain. We arrived at this model as described in the following.

Two types of non-local effects that are expected to contribute to the entropy of a denatured protein are "compactness" and "excluded volume" (Dill, 1990). Both of these effects can be described in terms of lattice models of polypeptide chains (Dill, 1990; Chan & Dill, 1991; Dill et al., 1995). The coordination number for a residue in a cubic lattice is 6 (up, down, left, right, front, back). The total number of conformations for a three-dimensional cubic lattice is thus proportional to 6^N (Chan & Dill, 1989), where N is the number of bonds ($N - 1$ residues).[7] Excluded volume effects (steric clashes) prohibit any two chain segments from occupying the same space. This leads to a reduction in the number of accessible conformations compared to the factor 6^N predicted for a free-flight chain. Numerical simulations of self-avoiding walks on cubic lattices suggest that for large N, the total number of conformations corrected for excluded volume is proportional to $N^\gamma \mu^N$ (Sykes, 1963; Fisher 1966; Chan & Dill, 1989). The constant μ corresponds to the limiting number of choices per step in a self-avoiding walk, and is estimated to have a value of

[6] It is unclear if the distinction between configurational and conformational entropy is blurred for a polypeptide unit in an unfolded protein because energy minima may be extremely broad.

[7] In the following we consider the 22-residue S-peptide to be made up of 21 repeating units, or bonds. The free amino terminus (residue 1) is not included in calculations.

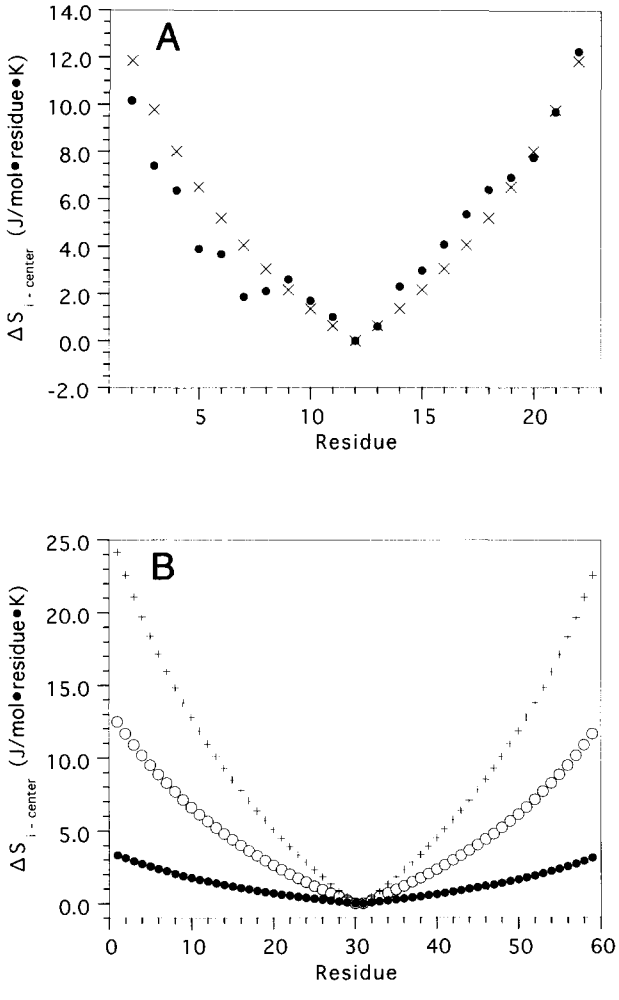

Fig. 7. A: Difference in $\delta S_p(j)$ values between individual residues i and the center of the free S-peptide (residue 12). The filled circles represent experimental ($\Delta S_{i\text{-center}}$) values, the crosses represent values simulated from Equation 18 with $\gamma = 1/6$ and $\mu = 4.68$. **B:** Simulated $\Delta S_{i\text{-center}}$ profiles for a 60-residue polypeptide as a function of the constant μ (Equation 18): filled circles $\mu = 1.5$, open circles $\mu = 4.68$, crosses $\mu = 20$.

4.68 for the simple cubic lattice (Sykes, 1963). The constant γ depends on the dimensionality but not on the specific geometry of the lattice, and is estimated to have a value of 1/6 for any three-dimensional chain (Sykes, 1963). Assuming equal weights for conformational states, the entropic correction for excluded volume per residue in a simple cubic lattice is

$$\Delta S = \frac{R}{N} \ln\left(\frac{N^{\gamma}\mu^{N}}{6^{N}}\right). \tag{14}$$

This correction changes slowly as a function of the number of residues in the sequence, reaching an asymptotic value of about -2.0 J/mol·residue·K for chains longer than 50 residues. A much greater reduction in the number of accessible conformations occurs if chains are constrained to compact conformations. Maximally compact conformations are defined as those with the maximum number of topological contacts (each of the six coordination sites is occupied for every residue), or equivalently those with the minimal surface/volume ratio (Chan & Dill, 1989). The number of maximally compact conformations varies approximately as κ^N for large N, with $\kappa = 1.74$ for a simple cubic lattice (Chan & Dill, 1991). The decrease in entropy per residue corresponding to the reduction in the number of conformations in maximally compact chains, compared to those allowed by excluded volume is

$$\Delta S = \frac{R}{N} \ln\left(\frac{\kappa^{N}}{N^{\gamma}\mu^{N}}\right). \tag{15}$$

Neglecting the smaller term N^{γ}, this corresponds approximately to $R \cdot \ln(1.74/4.68)$ or -8.3 J/mol·residue·K, irrespective of the size of the chain.

The lattice models discussed above predict that the entropic contributions from compactness and excluded volume are uniform as a function of sequence. As such, these results cannot be used to explain the entropic differences between the center and the ends of the S-peptide. NMR parameters reflect time and population-averaged properties over the ensemble of interconverting conformations. The larger entropies observed for the ends of the free S-peptide suggest that on a time average these residues have more accessible conformations than residues at the center of the chain. The principal force favoring compactness is the hydrophobic effect, which leads to the sequestering of non-polar functional groups from water. The principal force opposing compactness is the loss of chain entropy that results when a polypeptide is constrained to a compact conformation (Dill, 1990). This reduction in chain entropy should be most severe when the ends of the chain are restricted to be close in space. The larger entropies observed for the termini of the S-peptide may thus reflect that these segments are least favored to be constrained to compact conformations.

We initially assign the entropy of a maximally compact chain (Equation 15) to the center of the S-peptide (residue 12, $i = 11$). We include a second entropy term that reflects the gain in the number of conformations that would result if the chain becomes increasingly less compact towards the chain ends. For this correction, the number of compact conformations for a bond n_i positions

$$n_i = \left|\frac{(N+1)}{2} - i\right| \qquad \text{for } N \text{ odd}$$

$$n_i = \left|\frac{(N+1)}{2} - i\right| - \frac{1}{2} \qquad \text{for } N \text{ even} \tag{16}$$

away from the center of the chain, is taken to represent that of the center of a new chain that is shortened by n_i bonds. The total number of conformations $N^{\gamma}\mu^{N}$ in this second term is held constant. The resulting expression for the change in entropy relative to the center of the chain is

$$\Delta S_{i\text{-center}} = \left[\frac{R}{N} \ln\left(\frac{\kappa^{N}}{N^{\gamma}\mu^{N}}\right)\right] + \left[\frac{R}{N - n_i} \ln\left(\frac{N^{\gamma}\mu^{N}}{\kappa^{N-n_i}}\right)\right] \tag{17}$$

where the second term in Equation 17 is normalized by the factor of $N - n_i$, the number of residues remaining in the shortened chain.

By use of the relations $\ln(x^{y}) = y\ln(x)$ and $\ln(xy) = \ln(x) + \ln(y)$, Equation 17 can be considerably simplified to

$$\Delta S_{i\text{-center}} = R \cdot \frac{n_i}{N - n_i}\left[\ln(\mu) + \frac{\gamma}{N}\ln(N)\right]. \quad (18)$$

The term $(\gamma/N)\ln(N)$ decays to zero for large N. For the 22-residue S-peptide, neglect of the term $(\gamma/N)\ln(N)$ leads to a uniform 1.5% reduction in the values of $\Delta S_{i\text{-center}}$. Equation 18 can thus be approximated by

$$\Delta S_{i\text{-center}} \approx R \cdot \frac{n_i}{N - n_i}[\ln(\mu)]. \quad (19)$$

Note that $\Delta S_{i\text{-center}}$ in Equation 17 does not depend on the constant κ. The shape of the $\Delta S_{i\text{-center}}$ profile as function of sequence is determined by $n_i/(N - n_i)$. For the center of the chain, n_i and $\Delta S_{i\text{-center}}$ are zero. The maximum value of n_i occurs at the chain termini, where the factor $n_i/(N - n_i)$ tends toward unity, and $\Delta S_{i\text{-center}}$ toward a maximum given by $R\ln(\mu)$. Because the constant γ is 1/6 for all three-dimensional chains, the magnitude of entropy changes as a function of displacement from the center of chains of equal length N is uniquely specified by the value of the constant μ (Fig. 7B).

The simulated $\Delta S_{i\text{-center}}$ profile for the S-peptide (crosses in Fig. 7A) shows good agreement with experimental results (filled circles in Fig. 7A). Some fine structure and asymmetry indicative of residual entropic terms is evident in the experimental data. In particular, entropy changes are over-estimated by the model for the N-terminal half, and under-estimated for the C-terminal half. These discrepancies cannot be accounted for by the model (Equation 17), which is a homopolymer approximation. It is possible that the lower entropies (and order parameters) for the N-terminal half of the free S-peptide reflect the presence of a small population of α-helical conformations suggested by the $^1H\alpha$ chemical shift data. Alternatively, the lower entropies for the N-terminal half may be related to a slightly higher hydrophobicity for this portion of the molecule. Nevertheless, the values simulated by Equation 13 are all within 2.6 J/mol·residue·K of experimental values. The average difference between experimental and simulated values is 1.0 J/mol·residue·K, or about 8% of the total change in entropy between the center and the termini of the chain. The logarithm of the constant μ describes how quickly ^1H–^{15}N bond vector entropies rise as a function of n_i, the displacement from the center of the chain. For the simulation of the $\Delta S_{i\text{-center}}$ profile of the S-peptide, the value of μ was taken to be 4.68. This value is suggested by simulations of chains on simple cubic lattices (Sykes, 1963). If μ is treated as a free variable in a least-squares fit of $\Delta S_{i\text{-center}}$ as a function of position in the sequence (Equations 16 and 17), we obtain a slightly smaller value of 4.1 ± 0.3 for μ. The agreement between the value of $\mu = 4.68$ predicted from lattice simulations and the experimental value of 4.1 ± 0.3, however, is fortuitous. Although insights from lattice calculations are useful for developing a working model [e.g., the $n_i/(N - n_i)$ dependence of the $\Delta S_{i\text{-center}}$ profile], a chain constrained to a cubic lattice is a rather crude approximation of a real polypeptide. The value for μ in real polypeptides is unknown (Dill, 1985), and should depend on solution conditions in as much excluded volume and compactness depend on solvent (Flory, 1953). Taken together, these considerations suggest that for real polypeptides μ should be interpreted as an unknown that depends on the properties of a given polypeptide chain under a given set of solution conditions. The value of this parameter provides a useful measure of the rate of entropy change as a function of position in the polypeptide.[8]

An important question is whether the shape of the $\Delta S_{i\text{-center}}$ profile manifested by the S-peptide is observed more generally. NMR relaxation studies have been reported for a number of models of unfolded polypeptides (Torchia et al., 1975; vanMierlo et al., 1993; Alexandrescu & Shortle, 1994; Logan et al., 1994; Frank et al., 1995; Buck et al., 1996; Farrow et al., 1997; Kemple et al., 1997; Schwalbe et al., 1997). We exclude from this set of studies proteins constrained by disulfide bridges (vanMierlo et al., 1993), proteins in structure-inducing solvents (Buck et al., 1996), and proteins for which there is evidence of specific structure (Gillespie & Shortle, 1997). Of the remaining subset, the 59-residue N-terminal SH3 domain of drk in 2 M Gdn·HCl (Farrow et al., 1997) and the 26-residue monomeric melittin (Kemple et al., 1997) may show similar behavior to the free S-peptide. Although melittin shows a large drop in order parameters from the center to the N-terminus, the delta are limited as only 9 out of 26 sites were ^{13}C labeled (Kemple et al., 1997). Published A^2 order parameters for these proteins were converted to $\delta S_p(j)$ values (Equation 6), and the $\Delta S_{i\text{-center}}$ profiles were analyzed by a least-squares fit as described for the S-peptide. Experimental values for $\Delta S_{\text{center-end}}$ are on the order of −7.9, −5.7, and −4.2 J/mol·residue·K for monomeric melittin, the drk domain at a temperatures of 30 °C, and for the drk domain at a temperatures of 14 °C, respectively. As for the S-peptide, these values imply non-local contributions to the entropy of the chain. The respective values for the constant μ were 2.5 ± 0.3, 1.8 ± 0.1, and 1.5 ± 0.1, with average residuals between simulated and experimental data of 1.3, 0.8, and 2.0 J/mol·residue·K. The fit for the drk domain data is considerably poorer than that for the S-peptide, as might be expected from its larger size and complexity. The largest discrepancies occur for residues 35–38 and 50–53, which belong to two distinctly hydrophobic subdomains in the sequence of the protein (Farrow et al., 1997), and give lower entropies than predicted. Nevertheless, the overall shape of the $\Delta S_{i\text{-center}}$ profile is consistent with that predicted by Equation 19. Interestingly, the data for the drk domain at 14 °C give a slightly lower value for μ than that at 30 °C.

A number of studies point to a quite different trend in relaxation parameters as a function of sequence: an initial decrease in flexibility moving away from the chain termini, followed by a plateau of constant relaxation parameters for all residues more than a few positions away from the chain ends. Examples of this type of behavior include FK506 binding protein (Logan et al., 1994), GB1 (Frank et al., 1995),[9] and reduced lysozyme (Schwalbe et al., 1997), all studied at high concentrations of urea (6 to 8 M). This type of profile may be indicative of segmental reorientation (Allerhand & Hailstone, 1972). The relaxation parameters show an initial dependence on molecular weight but reach constant values above a critical size (Allerhand & Hailstone, 1972). This type of behavior is conceptually similar to that described for a freely rotating chain with a characteristic ratio in the C_∞ limit (Flory, 1953;

[8] A parameter such as μ might also be applicable to the characterization of the rate of entropy increase as a function of sequence position for flexible loops in an otherwise rigid protein.

[9] Relaxation parameters for this protein were analyzed in terms of the "model-free" approach. The A^2 values for the chain termini are about 0.20. For the residues more than five positions away from the ends, A^2 values reach a plateau of about 0.46. The difference between order parameters of 0.20 and 0.46 corresponds to a $\Delta\delta S_p(j)$ value of −5.1 J/mol·residue·K.

Cantor & Schimmel, 1980). Indeed, Schwalbe et al. (1997) have recently analyzed the plateau in the R2 relaxation profile of reduced lysozyme in terms of an (average) "persistence length" of the chain. It is interesting to note that all three proteins that exhibit a plateau were studied in high concentrations of urea. As noted by Frank et al. (1995), the viscosity of a concentrated urea solution (7.4 M) is 1.56-fold higher than that of water. It is possible that an increase in viscosity could change the relative weights of global and segmental reorientation. A more likely explanation, however, is that the solvent forces favoring hydrophobic collapse are considerably weaker in concentrated urea solutions than in water, or under milder denaturing conditions (e.g., lower concentrations of denaturant). Systematic studies of backbone dynamics as a function of increasing concentrations of denaturant would help clarify some of these issues. Such studies (Zhang & Forman-Kay, 1997) seem particularly important in light of the commonly held assumption that the conformational properties of unfolded proteins are independent of solvent.

At the present time NMR relaxation studies of highly dynamic polypeptides are few in number (Torchia et al., 1975; vanMierlo et al., 1993; Alexandrescu & Shortle, 1994; Logan et al., 1994; Frank et al., 1995; Alexandrescu et al., 1996; Buck et al., 1996; Farrow et al., 1997; Kemple et al., 1997; Schwalbe et al., 1997). One of the most interesting trends emerging from these studies is the extremely diverse behavior of proteins commonly lumped together under the designation "unfolded." Perhaps a more pragmatic description of these polypeptide chains is in terms of their chain entropy under a specified set of conditions (e.g., pH, temperature, solvent). The recent theoretical work aimed at establishing a link between NMR relaxation data and conformational entropy (Akke et al., 1993; Yang & Kay, 1996; Li et al., 1996) may provide the means for achieving such a description. The results of the present work suggest that backbone H–N bond vector entropies vary by more than 10 J/mol·K between the ends and the middle of the free S-peptide. The magnitude of these entropic changes may by open to question due to assumptions involved in the *MF* approach, the validity of the potentials used to calculate $\delta S_p(j)$ values (Yang & Kay, 1996), as well as the caveats described in the Introduction. Taken together, the relaxation data for the free S-peptide as well as for other proteins strongly suggest that the entropy of denatured polypeptide chains can be heterogeneous as a function of sequence. In general, the properties of two states must be considered in discussions of protein stability (Shortle, 1996):

$$\Delta G = G_D - G_N. \tag{20}$$

Although the properties of proteins under native conditions are by now fairly well understood, there is clearly much more work needed on "the other half of the folding equation."

Materials and methods

A synthetic cDNA fragment coding for residues 1–20 of bovine pancreatic RNAse A was cloned into the bacterial expression vector pPEP-T (Brandenberger et al., 1996). The resulting recombinant S-peptide was synthesized as part of a fusion protein, and contained the two additional residues G1 and S2 at its N-terminal end due to a thrombin cleavage site used to separate the S-peptide from the carrier protein. *E. coli* host strain JM109(DE3) (Promega), which has the gene for bacteriophage T7 RNA polymerase integrated into its genome, was used for all expression experiments. Single colonies of bacteria transformed with recombinant plasmid were grown in 3 mL of LB media with 100 μg/mL ampicillin. Following induction with 1 mM IPTG and further incubation for 2 h, 100-μL aliquots of bacterial cultures were analyzed by tricine-SDS-PAGE (Schägger & von Jagow, 1987) to test for the expression of the fusion protein. DNA insert sequences of plasmids expressing the desired recombinant gene were verified by Sanger dideoxy DNA sequencing. ^{15}N isotopic labeling was carried out in 1 L MOPS minimal media (Niedhardt et al., 1974; Serpersu et al., 1986) supplemented with 100 μg/mL ampicillin, and 15 mM $^{15}NH_4Cl$. His-tagged fusion protein was purified by immobilized metal affinity chromatography (Novagen) according to the manufacturer's instructions, and dialyzed against thrombin cleavage buffer (150 mM NaCl, 2.5 mM $CaCl_2$, 20 mM Tris-HCl pH 8.4). After dialysis, the S-peptide was separated from the carrier protein by proteolytic cleavage for 4 h at room temperature using human thrombin (Sigma) at a concentration of 5 units/mg recombinant protein. The S-peptide was purified by reverse-phase preparative HPLC on a 250 by 16 mm C18-100 column (Knauer). Analytical HPLC on a 250 by 4.6 mm C18 column (Vydac) was used to determine a purity of ≥97%. The identity of the S-peptide was further confirmed by mass spectrometry (experimental molecular mass 2,340 Da, calculated 2,341 Da). The yield for purified ^{15}N-labeled S-peptide was 1.5 mg/L of minimal media. S-protein (grade XII-PR) was from Sigma and was ≥92% pure by SDS-PAGE and HPLC.

All NMR experiments were performed on a 600 MHz Varian *Unity +* machine, with the probe thermostated at 10 °C. NMR data for the "free" S-peptide were collected on a 1.7 mM sample (1 mg/0.25 mL) of ^{15}N recombinant S-peptide in 90% H_2O (Milli-Q water)/10% D_2O, at a pH of 3.7. NMR data for the "bound" S-peptide were collected on a pH 3.7 sample containing 1.7 mM S-peptide and 2.9 mM S-protein (RNase A residues 21–124). The samples contained no added buffers or salts. Backbone ^{15}N R1, R2, and NOE NMR relaxation data were obtained using previously published experiments (Farrow et al., 1994; Alexandrescu et al., 1996). For the free S-peptide T1 and T2 relaxation parameters were measured from 8 data sets: T1 relaxation periods were 0.1, 0.2, 0.4, 0.5, 0.6, 0.7, 0.8, and 1.0 s; T2 relaxation periods were 16, 32, 64, 128, 200, 256, 300, and 400 ms. The recycling delay between scans was 2 s. In the ^{1}H-^{15}N NOE experiment a delay of 1 s was followed by ^{1}H saturation for 6.5 s. In the control experiment the ^{1}H saturation period was replaced by an equivalent delay. For the bound S-peptide 7 T1 (relaxation periods 0.07, 0.14, 0.28, 0.42, 0.56, 0.84, and 1.12 s) and 10 T2 (relaxation periods 16, 32, 48, 64, 86, 96, 120, 160, 200, and 256 ms) data sets were measured. The recycling delay between scans was 1.3 s. In the ^{1}H–^{15}N NOE experiment a delay of 1 s was followed by ^{1}H saturation for 3 s. Recycling delays and saturation periods were optimized as previously described (Alexandrescu & Shortle, 1994). The data sizes for the free S-peptide experiments were 1,024 × 25 complex points (spectral widths: ^{1}H—5,000 Hz, ^{15}N—530 Hz). Data sizes for the bound S-peptide experiments were 1,024 × 40 complex points (spectral widths: ^{1}H—6,000 Hz, ^{15}N—1600 Hz).

The *MF* analysis was performed using the program *Modelfree 3.1* (Mandel et al., 1995). Statistical properties of the *MF* parameters (Mandel et al., 1995) were obtained from Monte Carlo simulations using 500 randomly distributed synthetic data sets for both model selection, and for the global simulations.

Supplementary material in electronic appendix

Six tables containing ^{1}H- and ^{15}N-NMR assignments; R1, R2, and ^{1}H–^{15}N relaxation parameters; and the results of *MF* analyses for the free and bound S-peptide.

Acknowledgments

We thank Drs. A. G. Palmer III for the *Modelfree* v 3.1 computer program, H. Schwalbe for providing a preprint of his manuscript prior to publication, and D. Yang for stimulating discussions. This work was supported by Swiss National Science Foundation Grant 31-43091.95 to A.T.A.

References

Abragam A. 1961. *Principles of nuclear magnetism*. Oxford: Oxford University Press.

Akke M, Brüschweiler R, Palmer AG. 1993. NMR order parameters and free energy: An analytical approach and its application to cooperative Ca^{+2} binding by calbindin D_{9k}. *J Am Chem Soc 115*:9832–9833.

Alexandrescu AT, Shortle D. 1994. Backbone dynamics of a highly disordered 131 residue fragment of staphylococcal nuclease. *J Mol Biol 242*:527–546.

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

Alexandrescu AT, Jahnke W, Wiltscheck R, Blommers MJJ. 1996. Accretion of structure in staphylococcal nuclease: An ^{15}N NMR relaxation study. *J Mol Biol 260*:570–587.

Allerhand A, Hailstone RK. 1972. Carbon-13 Fourier transform nuclear magnetic resonance. X. Effect of molecular weight on ^{13}C spin-lattice relaxation times of polystyrene in solution. *J Chem Physics 56*:3718–3720.

Baldwin RL. 1995. The nature of protein folding pathways: The classical versus the new view. *J Biomol NMR 5*:103–109.

Bax A, Tjandra N. 1997. Are proteins even floppier than we thought? *Nat Struct Biol 4*:254–256.

Bonvin AM, Brünger AT 1996. Do NOE distances contain enough information to assess the relative populations of multi-conformer structures. *J Biomol NMR 7*:72–76.

Boresch S, Karplus M. 1995. The meaning of component analysis: Decomposition of the free energy in terms of specific analysis. *J Mol Biol 254*:801–807.

Brady GP, Szabo A, Sharp KA 1996. On the decomposition of free energies. *J Mol Biol 263*:123–125.

Brady GP, Sharp KA 1997. Entropy in protein folding and in protein–protein interactions. *Curr Opin Struct Biol 7*:215–221.

Brandenberger R, Kammerer RA, Engel J, Chiquet M. 1996. Native chick laminin-4 containing the β2 chain (s-laminin) promotes axon growth. *J Cell Biol 135*:1583–1592.

Brown JE, Klee WA. 1969. Conformational studies of a series of overlapping peptides from ribonuclease and their relationship to the protein structure. *Biochemistry 8*:2876–2879.

Buck M, Boyd J, Redfield C, MacKenzie DA, Jeenes DJ, Archer DB, Dobson CM. 1995. Structural determinants of protein dynamics: Analysis of ^{15}N NMR relaxation measurements for main-chain and side-chain nuclei of hen egg white lysozyme. *Biochemistry 34*:4041–4055.

Buck M, Schwalbe H, Dobson CM. 1996. Main-chain dynamics of a partially folded protein: ^{15}N NMR relaxation measurements of hen egg white lysozyme denatured in trifluoroethanol. *J Mol Biol 257*:669–683.

Cantor CR, Schimmel PR. 1980. *Biophysical chemistry, vol III, chap. 18*. New York: W.H. Freeman and Company.

Cavanagh J, Fairbrother WJ, Palmer AG, Skelton NJ. 1996. *Protein NMR spectroscopy principles and practice*. San Diego: Academic Press.

Chan HS, Dill KA. 1989. Compact polymers. *Macromolecules 22*:4559–4573.

Chan HS, Dill KA. 1991. Polymer principles in protein structure and stability. *Annu Rev Biophys Biophys Chem 20*:447–490.

Clore GM, Driscoll PC, Wingfield PT, Gronenborn AM. 1990a. Analysis of the backbone dynamics of interleukin 1-β using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy. *Biochemistry 29*:7378–7401.

Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM. 1990b. Deviations from simple two-parameter model-free approach to the interpretation of Nitrogen-15 nuclear magnetic relaxation in proteins. *J Am Chem Soc 112*:4989–4991.

Connor TM. 1964. Distributions of correlation times and their effects on the comparison of molecular motions derived from nuclear spin-lattice and dielectric relaxation. *Trans Faraday Soc 60*:1574–1591.

D'Aquino JA, Gómez J, Hilser V, Lee KH, Amzel LM, Freire E. 1996. The magnitude of the backbone conformational entropy change in protein folding. *Proteins Struct Funct Genet 25*:143–156.

Dill KA. 1985. Theory for the folding and stability of globular proteins. *Biochemistry 24*:1501–1509.

Dill KA. 1990. Dominant forces in protein folding. *Biochemistry 31*:7133–7155.

Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS. 1995. Principles of protein folding—A perspective from simple exact models. *Protein Sci 4*:561–602.

Doig AJ, Sternberg MJE. 1995. Side-chain conformational entropy in protein folding. *Protein Sci 4*:2247–2251.

Farrow NA, Munhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. 1994. Backbone dynamics of a free and a phosphopeptide-complexed Src Homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry 33*:5984–6003.

Farrow NA, Zhang O, Forman-Kay JD, Kay LE 1997. Characterization of the backbone dynamics of folded and denatured states of an SH3 domain. *Biochemistry 36*:2390–2402.

Finkelstein AV, Janin J. 1989. The price of lost freedom: Entropy of biomolecular complex formation. *Protein Eng 3*:1–3.

Fisher ME. 1966. Effect of excluded covolume on phase transitions in biopolymers. *J Chem Phys 45*:1469–1473.

Flory PJ. 1953. *Principles of polymer chemistry*. Ithaca, New York: Cornell University Press.

Frank MK, Clore GM, Gronenborn AM. 1995. Structural and dynamic characterization of the urea denatured state of streptococcal protein G by multidimensional heteronuclear NMR spectroscopy. *Protein Sci 4*:2605–2615.

Fushman D, Cahill S, Cowburn D 1997. The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: Analysis of ^{15}N relaxation with monomer/dimer equilibration. *J Mol Biol 266*:173–194.

Gallego E, Herranz J, Nieto JL, Rico M, Santoro J. 1983. ^{1}H NMR parameters of the N-terminal 19-residue S-peptide of ribonuclease in solution. *Int J Pept Protein Res 21*:242–253.

Gillespie JR, Shortle D. 1997. Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J Mol Biol 268*:170–184.

Hearn RP, Richards FM, Sturtevant JM, Watt GD. 1971. Thermodynamics of the binding of the S-peptide to S-protein to form Ribonuclease S'. *Biochemistry 10*:806–817.

Howlin B, Moss D, Harris GW. 1989. Segmented anisotropic refinement of bovine ribonuclease A by the application of the rigid-body TLS model. *Acta Crystallogr A45*:851–861.

Karplus M, Kushick JN. 1981. Method for estimating the configurational entropy of macromolecules. *Macromolecules 14*:325–332.

Karplus M, Ichiye T, Pettit BM. 1987. Configurational entropy of native proteins. *Biophys J 52*:1083–1085.

Kay LE. 1995. Pulsed field gradient multi-dimensional NMR methods for the study of protein structure and dynamics in solution. *Prog Biophys Mol Biol 63*:277–299.

Kay LE, Torchia DA, Bax A. 1989. Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry 28*:8972–8979.

Kemple MD, Buckley P, Yuan P, Prendergast FG. 1997. Main chain and side chain dynamics of peptides in liquid solution from ^{13}C NMR: Melittin as a model peptide. *Biochemistry 36*:1678–1688.

Kim EE, Varadarajan R, Wyckoff HW, Richards FM. 1992. Refinement of the crystal structure of ribonuclease S. Comparison with and between the various ribonuclease A structures. *Biochemistry 31*:12304–12314.

Kim PS, Baldwin RL 1984. A helix stop signal in the isolated S-peptide of ribonuclease A. *Nature 307*:329–334.

Kraulis PJ. 1991. MOLSCRIPT—A program to produce both detailed and schematic plots of protein structure. *J Appl Crystallogr 24*:946–950.

Lefevre J-F, Dayie KT, Peng JW, Wagner G. 1996. Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N–H spectral density functions. *Biochemistry 35*:2674–2686.

Levy GC, Axelson DE, Schwartz R, Hochmann J. 1978. Interpretation of complex molecular motions in solution. A variable frequency Carbon-13 relaxation study of chain segmental motions in poly(*n*-alkyl methacrylates). *J Am Chem Soc 100*:410–424.

Li Z, Raychaudhuri S, Wand AJ. 1996. Insights into the local residual entropy of proteins provided by NMR relaxation. *Protein Sci 5*:2647–2650.

Lipari G, Szabo A. 1982a. Model-free approach to the interpretation of nuclear magnetic resonance in macromolecules: 1. Theory and range of validity. *J Am Chem Soc 104*:4546–4559.

Lipari G, Szabo A. 1982b. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules: 2. Analysis of experimental results. *J Am Chem Soc 104*:4559–4570.

Logan TM, Thériault Y, Fesik SW 1994. Structural characterization of the FK506 binding protein unfolded in urea and guanidine hydrochloride. *J Mol Biol 236*:637–648.

Mandel AM, Akke M, Palmer AG III. 1995. Backbone dynamics of *Escherichia coli* ribonuclease HI: Correlations with structure and function in an active enzyme. *J Mol Biol 246*:144–163.

Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. 1989. Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H–^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: Application to interleukin 1β *Biochemistry 28*:6150–6156.

Mark AE, Van Gunsteren WF. 1994. Decomposition of the free energy of a system in terms of specific interactions. *J Mol Biol 240*:167–176.

Miyake A. 1958. On the theory of nuclear magnetic resonance in polymers. *J Polym Sci 28*:476–480.

Moroder L, Marchiori F, Rocchi R, Fontana A, Scoffone E. 1969. Synthesis of peptide analogs of the N-terminal eicosapeptide sequence of ribonuclease A. XII. Synthesis of des-Lys^1-[Orn^{10}]-, des-Lys^1,Glu^2-[Orn^{10}]-, des-Lys^1,Glu^2,Thr^3-[Orn^{10}]-S-peptide. *J Am Chem Soc 91*:3921–3926.

Neuhaus D, Williamson M 1989. *The nuclear Overhauser effect in structural and conformational analysis*. New York: VCH Publishers.

Nicholson LK, Yamazaki T, Torchia DA, Grzesiek S, Bax A, Stahl SJ, Kaufman JD, Wingfield PT, Lam PYS, Jadhav PK, Hodge CN, Domaille PJ, Chang C-H. 1995. Flexibility and function in HIV-1 protease. *Nat Struct Biol 2*:274–279.

Niedhardt FC, Bloch PL, Smith DF. 1974. Culture medium for enterobacteria. *J Bacteriol 119*:736–747.

Palmer AG III, Williams J, McDermott A. 1996. Nuclear magnetic resonance studies of biopolymer dynamics. *J Phys Chem 100*:13293–13310.

Peng JW, Wagner G. 1995. Frequency spectrum of NH bonds in Eglin c from spectral density mapping at multiple fields. *Biochemistry 34*:16733–16752.

Philippopoulos M, Lim C. 1995. Molecular dynamics simulation of *E. coli* ribonuclease HI in solution: Correlation with NMR and X-ray data and insights into biological functions. *J Mol Biol 254*:771–792.

Potts JT, Young DM, Anfinsen CB. 1963. Reconstitution of fully active RNase S by carboxypeptidase-degraded RNase S-peptide. *J Biol Chem 238*:2593–2594.

Redfield C, Boyd J, Smith LJ, Smith RA, Dobson CM. 1992. Loop mobility in a four-helix-bundle protein: ^{15}N NMR relaxation measurements on human interleukin-4. *Biochemistry 31*:10431–10437.

Richards FM, Vithayathil PJ. 1959. The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein components. *J Biol Chem 234*:1459–1464.

Rico M, Santoro J, Bermejo FJ, Herranz J, Nieto JL, Gallego E, Jiménez MA. 1986. Thermodynamic parameters for the helix-coil thermal transition of ribonuclease-S-peptide and derivatives from ^{1}H-NMR data. *Bioploymers 25*:1031–1053.

Rico M, Bruix M, Santoro J, Gonzales C, Neira JL, Nieto JL, Herranz J. 1989. Sequential ^{1}H-NMR assignment and solution structure of bovine pancreatic ribonuclease A. *Eur J Biochem 183*:623–638.

Robertson AD, Purisima EO, Eastman MA, Scheraga HA. 1989. Proton NMR assignments and regular backbone structure of bovine pancreatic ribonuclease A in solution. *Biochemistry 28*:5930–5938.

Rosa JJ, Richards FM. 1981. Hydrogen exchange from identified regions of the S-protein component of ribonuclease as a function of temperature, pH, and the binding of S-peptide. *J Mol Biol 145*:835–851.

Schaefer J. 1973. Distributions of correlation times and the Carbon-13 nuclear magnetic resonance spectra of polymers. *Macromolecules 6*:882–888.

Schägger H, von Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem 166*:368–379.

Schreier AA, Baldwin RL. 1977. Mechanism of dissociation of S-peptide from Ribonuclease A. *Biochemistry 16*:4203–4209.

Schwalbe H, Fiebig KM, Buck M, Jones JA, Grimshaw SB, Spencer A, Glaser SJ, Smith LJ, Dobson CM. 1997. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8M urea. *Biochemistry 36*:8977–8991.

Serpersu EH, Shortle D, Mildvan AS. 1986. Kinetic and magnetic resonance studies of effects of genetic substitutions of a Ca^{2+}-liganding amino acid in staphylococcal nuclease. *Biochemistry 25*:68–77.

Shortle D. 1996. The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J 10*:27–34.

Smith LJ, Bolin KA, Schwalbe H, MacArthur MW, Thornton JM, Dobson CM. 1996. Analysis of main chain torsion angles in proteins: Prediction of NMR coupling constants for native and random coil conformations. *J Mol Biol 255*:494–506.

Spolar RS, Record MT Jr. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science 263*:777–784.

Sykes MF. 1963. Self-avoiding walks on the simple cubic lattice. *J Chem Phys 39*:410–412.

Tolman JR, Flanagan, JM, Kennedy MA, Prestegard JH. 1997. NMR evidence for slow collective motions in cyanometmyoglobin. *Nat Struct Biol 4*:292–297.

Torchia DA, Lyerla JR, Quattrone AJ. 1975. Molecular dynamics and structure of the random coil and helical states of the collagen peptide α1-CB2 as determined by ^{13}C magnetic resonance. *Biochemistry 14*:887–900.

vanMierlo CPM, Darby NJ, Keeler J, Neuhaus D, Creighton TE. 1993. Partially folded conformation of the (30–51) intermediate in the disulphide folding pathway of bovine pancreatic trypsin inhibitor: ^{1}H and ^{15}N resonance assignments and determination of backbone dynamics from ^{15}N relaxation measurements. *J Mol Biol 229*:1125–1146.

Varadarajan R, Connely PR, Sturtevant JM, Richards FM. 1992. Heat capacity changes for protein-peptide interactions in the ribonuclease S system. *Biochemistry 31*:1421–1426.

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

Wagner G. 1995. The importance of being floppy. *Nat Struct Biol 2*:255–257.

Wang J, Purisima EO. 1996. Analysis of thermodynamic determinants in helix propensities of nonpolar amino acids through a novel free energy calculation. *J Am Chem Soc 118*:995–1001.

Williams RJP. 1989. NMR studies of mobility within protein structure. *Eur J Biochem 183*:479–497.

Wishart DS, Sykes BD, Richards FM. 1991. Relationship between nuclear magnetic resonance chemical shifts and protein secondary structure. *J Mol Biol 222*:311–333.

Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. 1995. ^{1}H, ^{13}C, and ^{15}N random coil chemical shifts of the common amino acids. I. Investigation of nearest-neighbor effects. *J Biomol NMR 5*:67–81.

Wittebrot RJ, Szabo A, Gurd FRN. 1980. Rotational motions of side chains of poly-L-lysine. *J Am Chem Soc 102*:5723–5728.

Yang A-S, Honig B. 1995. Free energy determinants of secondary structure formation: I. α-helices. *J Mol Biol 252*:351–365.

Yang D, Kay LE. 1996. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: Application to protein folding. *J Mol Biol 263*:369–382.

Zhang C, Cornette JL, Delisi C. 1997. Consistency in structural energetics of protein folding and peptide recognition. *Protein Sci 6*:1057–1064.

Zhang O, Forman-Kay JD. 1997. NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions. *Biochemistry 36*:3959–3970.