Sensitivity of NMR Residual Dipolar Couplings to Perturbations in Folded and Denatured Staphylococcal Nuclease†

Christine O. Sallum, David M. Martel, Robert S. Fournier, William M. Matousek, and Andrei T. Alexandrescu*

Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269

Received December 18, 2004; Revised Manuscript Received March 8, 2005

ABSTRACT: The invariance of NMR residual dipolar couplings (RDCs) in denatured forms of staphylococcal nuclease to changes in denaturant concentration or amino acid sequence has previously been attributed to the robustness of long-range structure in the denatured state. Here we compare RDCs of the wild-type nuclease with those of a fragment that retains a folded OB-fold subdomain structure despite missing the last 47 of 149 residues. The RDCs of the intact protein and of the truncation fragment are substantially different under conditions that favor folded structure. By contrast, there is a strong correlation between the RDCs of the full-length protein and the fragment under denaturing conditions (6 M urea). The RDCs of the folded and unfolded forms of the proteins are uncorrelated. Our results suggest that RDCs are more sensitive to structural changes in folded than unfolded proteins. We propose that the greater susceptibility of RDCs in folded states is a consequence of the close packing of the polypeptide chain under native conditions. By contrast, the invariance of RDCs in denatured states is more consistent with a disruption of cooperative structure than with the retention of a unique long-range folding topology.

Whether in a test tube or a cell, protein folding starts from denatured states (*1*, *2*). Structure that persists in denatured states has the potential to affect protein stability (*3*), and to facilitate or hinder protein folding (*1*, *4*). Partially formed, unfulfilled structures in denatured proteins may play roles in misfolding processes such as amyloid fibril formation (*5*, *6*). Moreover, modeling the mechanisms by which proteins become structured requires information about the initial denatured states in protein folding.

It is increasingly accepted that structure persists in denatured proteins and that the "random coil" model is a poor approximation for unfolded proteins (*3*, *7*, *8*). Nevertheless, fundamental questions remain about the type and amount of structure present. Foremost among these questions is whether long-range structure persists under denaturing conditions. Standard NMR methods typically show evidence for only short-range structure (*9*−*12*). Indirect support for long-range structure comes from relaxation data that identify segments of unfolded polypeptides connected through common dynamic properties (*13*−*17*). If the chain were a random coil, the motion of different segments would be uncorrelated. Nevertheless, a structural description of the dynamics described by NMR relaxation data remains elusive. More direct evidence for long-range structure has been obtained for the intrinsically unfolded state of an SH3 domain (*18*). Distance contacts were detected using HSQC−NOESY−HSQC experiments in conjunction with perdeuterated protein samples. Perdeuteration extends the distance sensitivity of NOE spectroscopy from ∼5 to ∼10 Å (*18*). Long-range distance contacts were identified, and an ensemble of structures was calculated for the unfolded form of the SH3 domain (*18*). Subsequent studies, however, led to a downward revision of the number of long-range distance contacts in the unfolded SH3 domain (*19*). An alternative way to probe long-range distance contacts is with paramagnetic spin-labels, which are sensitive to distance contacts as large as ∼20−25 Å (*20*). Paramagnetic probes were used to identify long-range distance contacts in the unfolded Δ131Δ fragment of staphylococcal nuclease (SN),[1] and these restraints were used to model the structure of the denatured fragment (*20*). The paramagnetic labeling approach has also shown evidence for long-range structure in the denatured states of protein L (*21*), apomyoglobin (*22*), and acyl-coenzyme A binding protein (*23*). Features that complicate the spin-labeling approach include the fact that distance contacts are transient in nature. Like the NOE, relaxation of spins by paramagnetic labels will be dominated by the distance of closest approach. As such, the observed paramagnetic effects are subject to dynamic averaging, which in principle could include averaging between the unfolded protein and a small population of the native protein.

Following the pioneering work of Shortle's group, residual dipolar couplings (RDCs) have become part of the repertoire of tools used to investigate structure in denatured proteins (*24*). RDCs for bonded pairs of NMR sensitive nuclei such as ^{1}H and ^{15}N depend on the angular orientation of the bonds relative to the spectrometer magnetic field (*25*, *26*). In solution, dipolar couplings are averaged to very small values

† This work was supported by NSF Grant MB 0236316 to A.T.A.

* To whom correspondence should be addressed: Department of Molecular and Cell Biology, University of Connecticut, 91 N. Eagleville Rd., U-3125, Storrs, CT 06269-3125. Telephone: (860) 486-4414. Fax: (860) 486-4331. E-mail: andrei@uconn.edu.

[1] Abbreviations: pdTp, thymidine 3′,5′-bisphosphate; RDC, residual dipolar coupling; rmsd, root-mean-square deviation; SN, staphylococcal nuclease; SN·T, SN−pdTp−Ca^{2+} ternary complex; SNOB, folded fragment of residues 1−103 of SN with the global suppressor mutations V66L and G88V.

10.1021/bi0473410 CCC: $30.25
Published on Web 04/05/2005

by the rapid and nearly isotropic tumbling that is typical of globular proteins. The intrinsic rotational diffusion anisotropy of proteins can be amplified by immersing the solutes in one of a number of anisotropic matrices such as bicelles, stretched or compressed gels, or phage particles (*25*). These matrices impart a weak net magnetic alignment on the solute. To retain the sharp signals typical of solution NMR, the degree of alignment is tuned so that the dipolar couplings have a magnitude that is $\sim 10^4$-fold smaller than in solid-state NMR (*26*). In contrast to the distance contact information provided by NOE spectroscopy or paramagnetic spin-labels, RDCs provide information about angles of individual bonds in the overall alignment frame of the molecule. RDCs are exquisitely sensitive to local geometry, and consequently have become a valuable tool for NMR structure refinement (*27*, *28*). They are sensitive to long-range structure in the sense that for a rigid structure the RDCs must satisfy both individual bond angles and parameters related to the overall anisotropy of the molecule, specifically, the axial and rhombic components of the molecular alignment tensor. The latter parameters determine the distribution of RDCs observed for a given solute (*29*, *30*). The situation becomes considerably more complicated, however, if a molecule is composed of multiple independent domains. The hallmarks of a protein with independently oriented domains are that (i) a single alignment tensor may fail to simultaneously satisfy all the RDCs and (ii) the magnitudes of RDCs can show systematic differences between independent domains (*31*−*33*).

RDCs, comparable to those seen in native proteins, were first observed by Shortle's group for the denatured Δ131Δ fragment of staphylococcal nuclease oriented in bicelles or gels (*24*, *34*, *35*). These findings have been confirmed by the observation of RDCs in denatured forms of eglin C (*36*), protein GB1 (*37*), acyl-CoA binding protein (*38*), apomyoglobin (*39*), and three additional proteins from our lab studied in the presence of 6 M urea: the all-β OB-fold protein CspA, the mixed $\alpha+\beta$ OB-fold protein LysN, and the GCN4-p coiled coil (A. T. Alexandrescu, C. O. Sallum, and W. M. Matousek, unpublished observations). RDCs have the potential to shed new light on the structural properties of denatured states by providing hitherto unavailable angular information. Except for eglin C (*36*), and the relatively simple α-helix formed by the S-peptide fragment of ribonuclease A (*40*), RDCs in denatured proteins do not appear to agree with the RDCs of the corresponding native states (*24*, *37*, *39*). The disagreement between RDCs of native and denatured states of the same protein could be due to (i) differences in structure, (ii) differences in overall alignment, (iii) differences in localized motions that affect the *S* order parameters and hence the magnitudes of RDCs, and (iv) segmentation of the unfolded polypeptide chain into multiple domains with noncoincident alignment tensors.

RDCs for the denatured Δ131Δ fragment of SN were attributed to the persistence of a nativelike topology in the denatured state (*24*, *34*). The evidence that a nativelike topology persists in the denatured nuclease comes from paramagnetic spin-labeling studies of the Δ131Δ fragment (*20*, *24*, *36*). The RDCs of denatured staphylococcal nuclease were shown to be nearly invariant with changes in denaturant concentration, pH, and perturbations of the amino acid sequence (*24*, *34*, *35*). The conservation of RDCs over a wide range of denaturant conditions, and when the polypeptide is truncated by as much as 20%, was interpreted as a reflection of the resilience of structure in the denatured state (*35*). The topology of the denatured state was hypothesized to be determined by local steric interactions between the main chain and the various genetically encoded side chains (*35*). This novel model of the denatured state has been considered by some of the most authoritative theoreticians in protein folding (*41*−*44*).

Because there are a number of folded forms known for staphylococcal nuclease, the protein provides a unique opportunity to examine how RDCs are affected by perturbations in different native and denatured states. Here we show that the subtle differences in structure between SN (*45*) and SN·T (*46*), the ternary complex of SN with Ca^{2+} and the inhibitor pdTp, are comparable to the differences reported between denatured forms of the protein. Still larger differences in RDCs are observed when wild-type SN is compared to the folded SNOB fragment (*47*) which is missing the last 47 of 149 residues. In contrast, the two proteins give very similar RDCs in 6 M urea, confirming previous reports about other denatured forms of staphylococcal nuclease (*24*, *35*).

Our results argue against the presence of long-range structure in SN under strongly denaturing conditions. Whereas in folded proteins changes in one part of the structure affect sites distant in the polypeptide chain, the invariance of RDCs in the denatured state suggests a lack of cooperatively stabilized structure. We propose that the RDCs observed in denatured proteins could reflect the presence of localized anisotropic structures that are independent of global structural constraints.

EXPERIMENTAL PROCEDURES

Materials. Staphylococcal nuclease (SN) was purified from *Escherichia coli* strain T7/MG containing plasmid pSNWT11a (*48*), a gift from J. Flanagan (The Pennsylvania State University, State College, PA). Expression and purification of ^{15}N-labeled SN were carried out as previously described for the Δ131Δ fragment of the protein (*10*), except that the *E. coli* culture was grown at 28 °C, and the temperature was increased to 37 °C only after induction of the nuclease gene with IPTG. The highly active wild-type SN is toxic to cells, and *E. coli* cells quickly lose the plasmid harboring the DNA and RNA-hydrolyzing enzyme at the higher temperature (D. Shortle, personal communication). Cell growth at 37 °C is not a problem with less active mutants of nuclease such as the SNOB fragment. The SNOB fragment was expressed in *E. coli* HMS174 (DE3) cells using a pET11a-derived vector designed in the Shortle lab (*47*). ^{15}N-labeled SNOB was expressed and purified as previously described (*47*).

The inhibitor pdTp was a gift from B. Garcia-Moreno and D. Karp (Johns Hopkins University, Baltimore, MD). Pf1 phage was purchased from ASLA Biotech Ltd. (Riga, Latvia). The apparatus used to prepare stretched gels and NMR tubes for gels were purchased from New Era Enterprises (Vineland, NJ). Electrophoresis-grade acrylamide and bisacrylamide were from Bio-Rad. $CaCl_2$ and "ultra-grade" urea were from Sigma-Aldrich. Concentrations of urea stock solutions were determined from refractive index measurements (*49*).

Phage Alignment. Samples aligned with Pf1 bacteriophage, strain LP11-92 (ASLA Biotech Ltd.), were prepared in a total

volume of 360 μL and had a final phage concentration of 15.5 mg/mL. We found that under these conditions SN and SNOB are insoluble, but that the proteins readily dissolve if NaCl is included at a final concentration of 0.47 M. Nuclease is highly positively charged (pI 9.9), and high salt concentrations presumably reduce the number of strong electrostatic interactions with the negatively charged phage particles (*50*). The phage-aligned and isotropic control samples in the presence of 0.47 M salt had ^{1}H$-^{15}$N HSQC chemical shifts that were nearly identical to those of the protein in the absence of salt. The inclusion of 0.47 M salt in the phage medium used for alignment thus does not appear to significantly affect the structure of SN. This is also supported by the agreement of the RDC data sets for SN and SN·T with their respective X-ray structures. The only region in SN that appears to be perturbed by salt and/or phage is the loop segment between residues 113 and 118 (see the text). For the phage alignment experiments, SN and SNOB concentrations were 0.3 mM, the pH was 7.6, and the temperature was 25 °C.

Polyacrylamide Gels. Gels for RDC measurements were prepared from a stock solution of 40% (w/v) acrylamide and 2% (w/v) *N*,*N*′-methylenebisacrylamide, to a final acrylamide:bisacrylamide ratio of ~5:1 (v/v of stock solutions). Cylindrical 7.5% gels (500 μL) were cast using the 600 MHz Gel NMR Kit (New Era Enterprises) to a diameter of 6 mm. The gels were dialyzed three times in distilled water, dried overnight at 32 °C, and rehydrated by being soaked overnight in solutions of 0.3−0.5 mM protein in a 85% H_2O/15% D_2O mixture. The slightly higher D_2O concentration of 15% was used to facilitate locking on the D_2O signal of the aligned samples. Native SN and SNOB were in 20 mM acetate buffer at pH 4.9 and 5.3, respectively. Denatured SN and SNOB samples were in 6 M urea at pH 3. Once rehydrated with protein solutions, the gels were gently pushed into open-ended 600 MHz gel NMR sample tubes from New Era Enterprises, as described in the manufacturer's instructions. The bottom end of the tube was sealed with a plug, and the top was sealed with the plunger provided in the kit. NMR data were collected at 25 °C for the folded proteins and at 32 °C for the denatured proteins.

NMR Spectroscopy. All NMR data were collected on a Varian Inova spectrometer operating at 600 MHz. NMR assignments for folded SN, SN·T, and SNOB were obtained from the BioMagResBank. Denatured SN and SNOB in 6 M urea at pH 3 and 32 °C were assigned by extending published assignments for SN at pH 3.0 and 32 °C in the absence of urea (*51*). Starting from the published assignments, we used a series of HSQC spectra in increments of ~1 M urea to follow the titration of SN peaks from 0 to 6 M urea. Assignments for denatured SNOB were obtained by comparison to the spectrum of the full-length SN, and also made use of published ^{1}H$-^{15}$N HSQC assignments for a 1−110(G88W) fragment of SN in the presence of 6 M urea at pH 4.9 and 32 °C (*52*).

$^1J_{HN}$ splittings were measured using twice the frequency difference in hertz between corresponding peaks in ^{1}H$-^{15}$N HSQC and TROSY spectra (*53*). We previously measured $^1J_{HN}$ splittings from spectra that selectively recorded the top-right and bottom-right (TROSY) components of coupled ^{1}H$-^{15}$N quartets (*40*). For large molecules, the relaxation properties of the top-right component are unfavorable, resulting in large sensitivity losses. We found that pairs of HSQC−TROSY spectra (*53*) are a more effective way to measure $^1J_{HN}$ splittings. The spectra were acquired with 1024 × 256 complex points for the folded proteins and 1024 × 512 complex points for the denatured proteins. Spectra were typically zero filled to 2048 × 2048 points to allow accurate cross-peak picking and ^{15}N frequency measurements with Felix 2000.1 (Accelrys). RDCs were calculated from the difference in $^1J_{HN}$ values between isotropic and aligned samples (RDC = $^1J_{HN,isotropic} - {}^1J_{HN,aligned}$) (*38*).

Table 1: Summary of RDC Data Statistics, Fits to the Wild-Type Nuclease Structure, and Correlations

RDC data set[a]	no. of points	range/mean error (Hz)[b]	*Q*-factor for fit to structure[c]
*f*_SN·T (pf1)	65	−15 to 27/ND	0.28
*f*_SN (pf1)	63	−29 to 40/1.5	0.19
*f*_SNOB (pf1)	74	−21 to 21/0.5	0.75
*f*_SN (gel)	58	−29 to 33/1.1	0.25
*f*_SNOB (gel)	66	−10 to 10/ND	0.81
*d*_SN (gel)	66	−16 to 3/0.5	0.95
*d*_SNOB (gel)	41	−10 to −1/0.3	0.93

correlation	no. of points	*R*-value[d]	ρ (correlation probability)
*f*_SN vs *f*_SN·T (pf1)	48	0.90	<0.0001
*f*_SN vs *f*_SNOB (pf1)	46	0.63	<0.0001
*f*_SN vs *f*_SNOB (gel)	30	0.67	<0.0001
*d*_SN vs *d*_SNOB (gel)	33	0.86	<0.0001
*f*_SN vs *d*_SN (pfl/gel)	33	−0.21	0.25
*f*_SN vs *d*_SN (gel)	32	−0.16	0.36
*f*_SNOB vs *d*_SNOB (pf1/gel)	30	0.10	0.62
*f*_SNOB vs *d*_SNOB (gel)	30	−0.10	0.61

[a] The italic characters *f* and *d* indicate proteins in their folded and denatured forms, respectively. The parentheses indicate whether the sample was aligned using polyacrylamide gels or pf1 phage particles. [b] The range of RDC values in the given data set and the mean error of the RDC estimated from replicate (duplicate or triplicate) measurements. [c] The RDC data were fit to X-ray structures of wild-type staphylococcal nuclease. The *Q*-factor [Q = rms(RDC^{exp} − RDC^{calc})/rms(RDC^{exp})] tends to zero as the fit between experimental RDCs and those calculated from a structural model improves. [d] Pearson's correlation coefficient.

To obtain estimates for the uncertainties in RDCs, we obtained duplicate or triplicate measurements of $^1J_{HN,aligned}$. Uncertainties were then calculated as the average over these replicate measurements (Table 1). In most cases, the uncertainties in $^1J_{HN}$ were on the order of ~1.0 Hz. Larger uncertainties were obtained for folded SN in pf1 phage particles or gels, but the range of RDC values for these samples was almost 3 times as large. In most cases, RDC measurements were repeated as much as 4−6 months after the original experiments, and then once again back to back. The RDCs of denatured samples in 6 M urea exhibited little if any change over a period of months. Differences for RDC measurements separated by months were comparable to the differences between RDC data sets collected back to back. For folded SN in pf1 phage, there was evidence for small changes in the sample over time. A mean error of 1.8 Hz was obtained between data sets collected 6 months apart compared to an error of 0.8 Hz for data collected 1 day apart. The uncertainty quoted in Table 1 for folded SN aligned in pf1 phage particles may thus be overestimated.

Module (*54*) and PALES (*55*) were used for the structural analysis of RDCs.

RESULTS

Different Forms of Staphylococcal Nuclease Used for RDC Studies. NMR residual dipolar couplings (RDCs) for denatured nuclease were previously reported to be nearly invariant with changes in denaturant concentration or truncation of the polypeptide chain by as much as 20% (*24*, *35*). The conservation of RDCs was attributed to the robustness of long-range structure in the denatured state (*35*). In this work, we compare the response of RDCs to perturbations of native and denatured states. In the presence of Ca^{2+} and the substrate analogue inhibitor pdTp, SN forms a ternary complex which we will designate SN·T. The X-ray structures of the SN apoprotein [PDB entry 1EY0 (*56*)] and of the ternary complex SN·T [PDB entry 1SNC (*46*)] are very similar, with an rmsd of 0.67 Å over 133 residues. Significant differences between the structures are restricted to parts of the active site loops that participate in Ca^{2+} or inhibitor binding (residues 18−22, 38−53, 83−87, and 112−116). The apoprotein shows slightly increased backbone dynamics compared to the complex (*14*).

The wild-type SN was also compared to a very unusual mutant fragment called SNOB. The SNOB fragment corresponds to residues 1−103 of the 149-residue wild-type SN, and contains two rare "global suppressor" mutations, V66L and G88V (*47*). Although SN unfolds with a cooperativity characteristic of a single protein domain (*1*), the SNOB fragment retains a well-defined tertiary structure in the absence of the C-terminal third of the chain. The NMR structure of SNOB (*47*) suggests that the V66L and G88V global suppressor mutations stabilize structure in the fragment by reinforcing the hydrophobic core of the protein's β-barrel. This five-stranded β-barrel is a conserved structural motif in the OB-fold family of proteins, and its relatively autonomous folding may reflect a role in the evolution of OB-fold proteins (*47*, *57*). OB-fold proteins combine the conserved five-stranded β-barrel structural motif with more variable elements of structure (*58*, *59*). In the three nonhomologous OB-fold proteins studied so far in our lab, the conserved β-barrel is more stable to partial unfolding than the nonconserved elements of structure (*59*). The structures of SN and SNOB have an rmsd of 4.0 Å over 95 residues. If disordered loops outside of the conserved cores of the two proteins are excluded, SN and SNOB agree to an rmsd of 1.9 Å over 60 residues (residues 10−35, 55−76, and 89−100). Backbone flexibility as measured by NMR relaxation parameters is much more pronounced in SNOB than in SN. The portions of the structure that are packed against the two C-terminal α-helices in the wild-type SN show the largest flexibility increases in SNOB (*14*). Chemical shift differences between SN and SNOB are also larger than between SN and SN·T (Figure 1A,B).

In addition to native forms of the protein, we also compared RDCs for SN and SNOB unfolded in 6 M urea. Denatured SN shows large amplitude backbone dynamics, and the degree of main chain flexibility is highly variable along the length of the polypeptide chain (*14*). The chemical shift differences between SN and SNOB in 6 M urea are small relative to those between the folded forms of the protein (Figure 1C).

Structural Analysis of RDCs in Folded Nuclease. Figure 2 shows fits of RDC data to X-ray structures of nuclease.

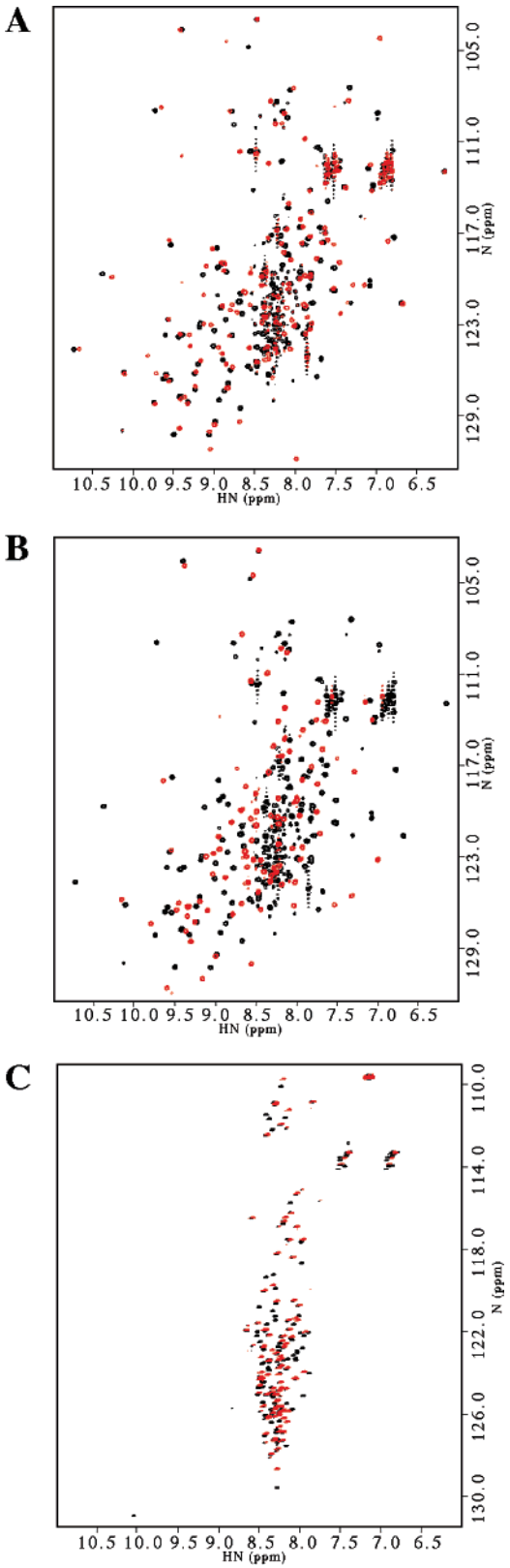

FIGURE 1: Superposition of NMR spectra for different forms of staphylococcal nuclease: (A) SN (black) and SN·T (red), (B) SN (black) and SNOB (red), and (C) denatured SN (black) and denatured SNOB (red) in the presence of 6 M urea. All spectra were collected at 25 °C. The pH values for the samples were 5.8 for SN, 6.1 for SN·T, 5.9 for SNOB, and 3.0 for denatured SN and SNOB.

Figure 3 shows the principal axes of the alignment tensors obtained from the fits, overlaid on the structures.

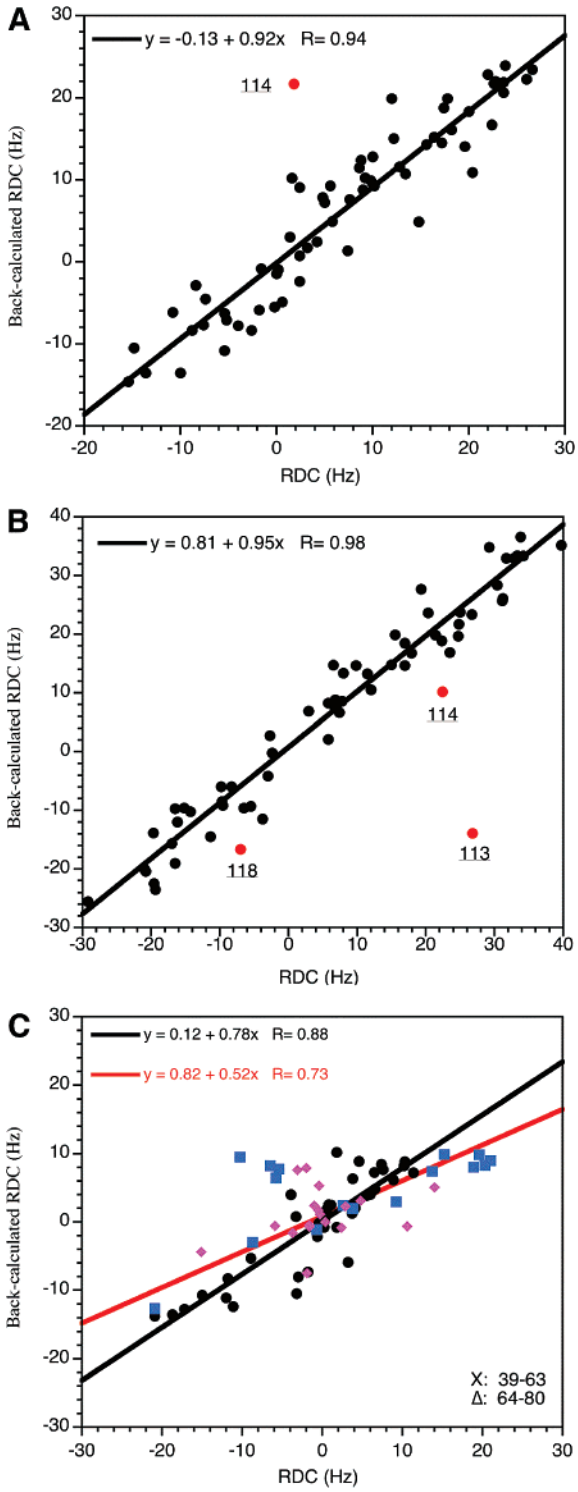

FIGURE 2: Fits of RDC data obtained from samples oriented in pf1 phage particles to native state X-ray structures. (A) RDC data for SN·T fit to the X-ray structure of the ternary complex (1SNC) (B) RDC data for SN fit to the X-ray structure of the apoprotein (1KDB). (C) RDC data for SNOB fit to residues 7−103 of the X-ray structure of the intact nuclease apoprotein (1KDB). The red points for SN and SN·T were outliers not included in the fits (see the text). The blue and purple points for SNOB are RDCs from two highly flexible segments in the protein. The lines show the fits to all the data (red) and the data excluding the two flexible loops (black). The quality of RDC fits to structures can be described by the *Q*-factor [$Q = \text{rms}(\text{RDC}^{\text{exp}} - \text{RDC}^{\text{calc}})/\text{rms}(\text{RDC}^{\text{exp}})$]. The *Q*-factor decreases toward zero as the agreement between experiment and model improves (*69*). The *Q*-factors for the fits were 0.28 for SN·T, 0.19 for SN, and 0.75 for SNOB with all points (0.46 for the partial fit indicated by the black line).

In the case of SN·T, the data were fit to the 1.65 Å resolution 1SNC structure (*46*). Even though an older structure (2SNS) of the ternary complex has a better resolution of 1.5 Å, a better fit was obtained with the 1SNC structure. The sole exception is residue V114, which is a pronounced outlier and was not included in the fit (Figure 2A). V114 gives acceptable fits to other nuclease structures, so the discrepancy appears to be an artifact of the 1SNC structure. Fits of the RDC data to other structures gave similar results, a good fit overall, except for one or two strong outliers (for example, K84 in the 2SNS structure). Because it exhibited better agreement with the RDC data, we chose to use the 1SNC structure for analysis.

The X-ray structure with the best agreement with the RDC data for unbound SN is the 1.9 Å resolution 1KDB structure (*60*). As with SN·T, there are X-ray structures with better resolution than 1KDB (e.g., 1EY0 has a resolution of 1.6 Å), but the 1KDB structure gives the best agreement with the SN RDC data. The 1KDB structure is of a K116E single-site mutant that changes the conformation of the loop between α-helices 2 and 3, which includes the site of the mutation. This region shows the strongest outliers in fits of the RDC data to the SN structure. Because of these factors, we decided to omit the three residues (Y113, V114, and N118) flanking the K116E mutation from the analysis of RDCs for SN (Figure 2B).

Figure 2C shows a fit of the RDC data for SNOB to residues 1−103 of the X-ray structure for wild-type SN (1KDB). The fit for the SNOB RDC data is much poorer than that for SN·T and SN. Similar results were obtained when the RDC data for SNOB were fitted to the NMR structure of SNOB. The largest residuals for the fit to the X-ray structure are clustered to distinct regions of the protein. The first region is between residues 39 and 63 (Figure 2C). This consists of a long disordered loop and an α-helix between β-strands 3 and 4 (*47*). Both of these elements of structure pack against the two C-terminal α-helices of SN, which are missing in the SNOB structure. The second region is between residues 64 and 80. This region encompasses the fourth strand of the β-sheet (residues 70−75) and a dynamically disordered loop between β-strands 4 and 5 (*47*). The two regions, particularly residues 39−63, increase the scatter of the fit and weaken correlations to the RDC data. Interestingly, experimentally determined RDCs for the flexible regions in SNOB span a greater range of values (from −21 to 21 Hz) than those predicted on the basis of the structure (from −13 to 10 Hz) when all the data are used in the analysis. The larger-than-expected RDCs for the unstructured loops in SNOB suggest that the flexible parts of the molecule exhibit intrinsic anisotropy that is distinct from that of the globular core of the molecule.

Modeling experiments with Module further support this hypothesis. Module makes it possible to subdivide a structure and fit the RDC data for segments separately (*54*). If the segments are rigidly linked, alignment of the principal axes of the individual segments can be used to model the change in relative orientation between the segments. Applying this procedure to residues 64−88 of SNOB results in a displacement of strand β4 from the rest of the barrel structure, which would preclude hydrogen bonding with strand β5 and completion of the native OB-fold β-barrel structure. When

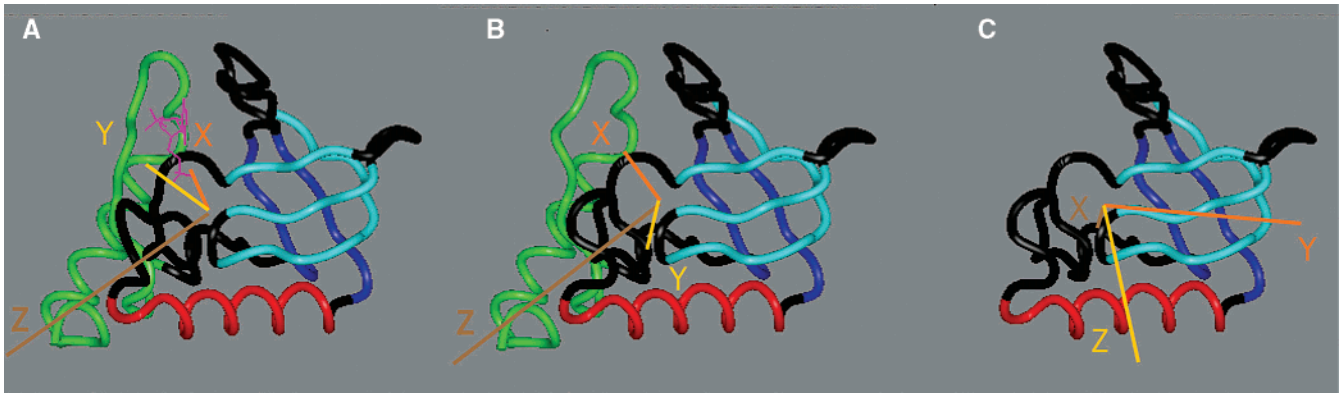

FIGURE 3: Principal axes of alignment tensors obtained from the pf1 phage RDC data depicted in Figure 2. Alignment tensors were calculated with Module (*54*). (A) SN•T: $A_a = 11 \times 10^{-4}$, $A_r = 4 \times 10^{-4}$, $\alpha = -87$, $\beta = 118$, and $\gamma = -35$. (B) SN: $A_a = 16 \times 10^{-4}$, $A_r = 5 \times 10^{-4}$, $\alpha = -55$, $\beta = 115$, and $\gamma = -30$. (C) SNOB: $A_a = -6 \times 10^{-4}$, $A_r = -3 \times 10^{-4}$, $\alpha = -87$, $\beta = 41$, $\gamma = 25$. The inhibitor pdTp is shown in purple.

this structure is stored and used as an alternative model for the experimental RDC data, the fit for residues 64−88 remains as poor as with the original native structure. This behavior is a hallmark of a molecule with multiple segmental alignment tensors (*32*, *33*). It might seem counterintuitive for disordered segments of a protein to have sufficient anisotropy to give RDCs comparable or larger than those in a folded globular protein. Collapse of a protein to a folded globule structure should minimize anisotropy as a protein comes close to achieving a spherical shape. On the other hand, partially folded incomplete structures will have a greater chance to deviate from globular structures, which is borne out by the sizable RDCs observed for a number of proteins under denaturing conditions. The tendency to adopt anisotropic structure in partially folded proteins will be balanced by an increase in internal dynamics, which will have the effect of decreasing the sizes of RDCs.

Figure 3 shows the principal axes of the alignment tensors for the three folded forms of nuclease superimposed on their structures. The alignment tensors of SN•T (Figure 3A) and SN (Figure 3B) are similar, but have small differences. The alignment tensor of the SNOB fragment (Figure 3C) shows large differences compared to the alignment tensors of the two intact proteins.

In addition to pf1 phage particles, RDC data for SN and SNOB were obtained using axially stretched 7.5% polyacrylamide gels as an alignment medium. The results from the gels are in excellent agreement with those from the pf1 phage system (Table 1). As with the phage data, RDCs for SNOB aligned in gels fall into two classes. The RDCs for the rigid β-barrel core agree with the nuclease structure, and give a *Q*-factor of 0.54, an *R*-value of 0.85, and a slope of 0.72. By contrast, the RDC data for residues 38−80 give a *Q*-factor of 0.87, an *R*-value of 0.41, and a slope of 0.18. The RDCs predicted for the flexible portion of the molecule when all the data are fitted to the structure (−4 to 5 Hz) were smaller than those observed experimentally for the flexible regions (−8 and 9 Hz).

The gel and phage data are internally consistent, and strongly suggest that the discrepancies between the RDC data for SNOB and SN reflect intrinsic differences in structure and dynamics as opposed to interactions with the alignment media. Interestingly, in the gel system, the SN loop segment between residues 114 and 118 agrees with the 1KDB structure, which is not the case with samples oriented in phage. This suggests that the aberrant RDCs for this segment with the phage system (Figure 2B) may be due to specific interactions with phage, or due to the high salt concentrations (∼0.5 M) needed to measure RDCs for nuclease in the phage medium.

In gels, solutes achieve a weak net alignment through steric hindrance. With the phage systems, solutes are oriented through a combination of electrostatic and steric effects (*25*, *34*, *55*). Nevertheless, the RDC data obtained for nuclease in the phage and gel systems are correlated. The *R*-values for the correlations are 0.87 and 0.70 for SN and SNOB, respectively. These correlations indicate that the alignment tensors and hence the alignment mechanism in the two media are not unique (*25*, *34*). At the high salt concentrations needed to dissolve the nuclease in the phage system, steric hindrance rather than electrostatics may dominate alignment with the rodlike pf1 phage particles.

Comparisons of RDCs in Folded and Unfolded Nuclease. Figure 4 compares the sequence profiles of RDCs in native and denatured SN and SNOB. Whereas the native proteins show both positive and negative RDCs with an average near zero, the RDCs for the denatured proteins in stretched gels are almost all negative. The prevalence of RDCs of the same sign has been previously observed for a number of denatured proteins (*36*, *38*, *39*, *61*) and has been attributed to extended conformations that align along the long axis of the orienting media (*38*, *39*, *61*, *62*). Thus, in a stretched gel, the polypeptide chain would align longitudinally with the H−N vectors perpendicular to the magnetic field. The couplings expected for this type of mechanism would be predominantly negative, as observed for denatured SN and SNOB (*39*). Although couplings of the same sign have been noted for the majority of denatured proteins studied thus far, we are aware of at least three cases in which both positive and negative couplings are observed: the protein eglin C in 8 M urea which shows a correlation between RDCs of the native and denatured states (*36*), the acid-denatured form of ACBP (*38*), and the protein CspA in the presence of 6 M urea (A. T. Alexandrescu, unpublished observations).

The angular dependence of RDCs predicts that H−N bond vectors at 90° to the spectrometer magnetic field should give the largest negative values, while bond vectors parallel to the magnetic field should give the largest positive values

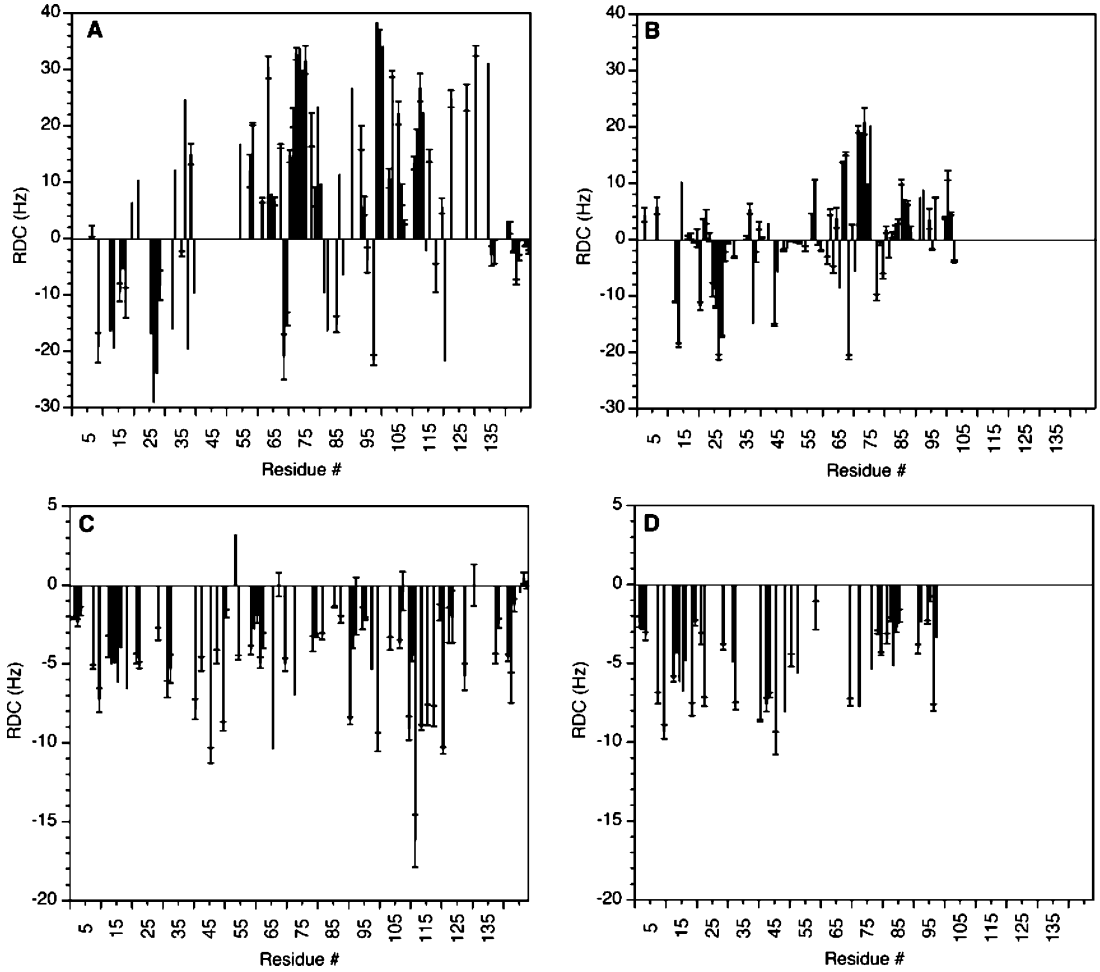

FIGURE 4: Sequence profiles of RDCs in (A) native SN, (B) native SNOB, (C) denatured SN, and (D) denatured SNOB. For the data in panels C and D, the proteins were denatured in 6 M urea. The samples in panels A and B were aligned in pf1 phage, while the samples in panels C and D were aligned in stretched gels. Uncertainty bounds (error bars in the plots) were estimated from replicate measurements.

(*29*). The presence of all negative RDCs in denatured forms of nuclease suggests that all H−N vectors are distributed along the equatorial section of a unit sphere (*39*, *61*) and precludes a nativelike protein structure. This can be tested by setting all the RDCs observed for native SN in stretched gels to negative values and fitting the new data set to the native 1KDB structure. The parameters for the fit deteriorate from an R of 0.97 and a Q-factor of 0.24 when the actual values are used to an R of 0.58 and a Q-factor of 0.73 for the synthetic data set with all RDCs set to negative values. The poor quality of the fit occurs even though in the actual data set for native SN two-thirds of the RDCs are negative.

A random flight Gaussian chain model (*62*) and a valance chain model (*63*) have been proposed to explain the uniform signs of RDCs in most denatured proteins studied to date. In these models, the denatured state behaves like an extended chain that orients along the longest dimensions of the orienting medium. The models predict RDCs of the same sign, with a flat sequence profile that decreases at the ends of the chain (*62*, *63*). The RDCs of denatured nuclease are almost all negative, and there is a leveling off of the RDCs at the chain ends as predicted by the model. On the other hand, the sequence profiles for denatured SN and SNOB are more complex than the flat profile predicted for a random flight chain (Figure 4C,D). The fine structure observed in the RDC data of both denatured SN and SNOB is consistent with the presence of nonrandom structure (*61*, *63*).

For stiff chains, the magnitudes of RDCs are expected to level off with an increase in segment length (*62*, *63*). By contrast, the RDCs of denatured SNOB are nearly coincident with those of the longer SN chain (Figure 4). This suggests that the persistence length of structure in denatured SN must be shorter than the entire chain.

RDCs for native and denatured SN both measured in stretched 7.5% polyacrylamide gels are uncorrelated (Figure 5), in agreement with previous results (*24*). There is also no correlation between the RDC data for native and denatured SNOB in 7.5% gels ($R = -0.1$, $\rho = 0.6$; data not shown). To date, a weak correlation between the RDCs of a native and denatured protein has been reported for only eglin C (*36*). The absence of a correlation need not imply a difference in structures. A change in alignment tensor caused by a substantially different orientation of the same structure could give rise to different RDCs (*64*). We thus attempted to fit the RDC data for denatured SN to the X-ray structure of native nuclease, 1KDB, with Module (*54*). The fits demonstrated a lack of correlation between the RDCs of the denatured state and the best-fitting alignment tensor obtained for the native SN structure ($R = 0.13$, $\rho = 0.30$, and Q-factor $= 0.96$). Comparable correlations were obtained with alternative structural models, including unrelated folds such as a TIM-barrel structure (PDB entry 1TIM), structures with random dihedrals, and structures with all dihedrals set to the α-helix or β-strand region of $\phi-\psi$ space (not shown).

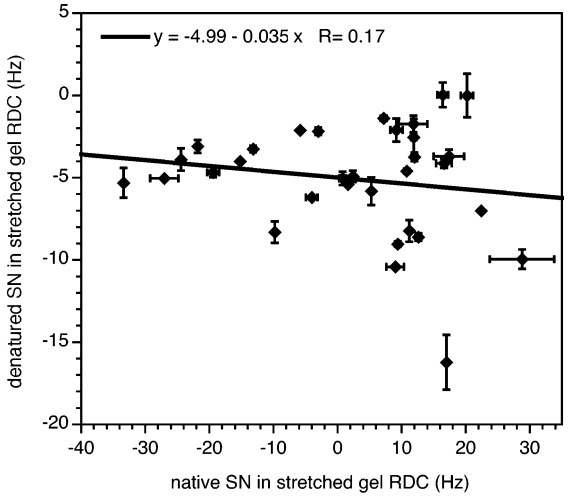

FIGURE 5: Linear correlation of RDC data for native and denatured SN, both aligned in stretched gels. At the 95% confidence level, ρ is 0.37, indicating the absence of a statistically significant correlation.

Correlations of RDCs in Folded and Unfolded Nuclease. Figure 6 shows linear correlations between different native and denatured forms of SN. Parameters for these correlations are summarized in Table 1. The RDCs of the apoenzyme, SN, and the ternary complex SN·T are strongly correlated (Figure 6A), consistent with the similar X-ray structures of the bound and unbound forms of the protein. Similarly strong correlations with *R*-values above 0.8 have been observed between RDCs of denatured forms of nuclease subjected to different urea concentrations, changes in up to 10 sequence positions, and truncation of 30 residues from either the N- or C-terminus of the chain (*35*). A much weaker correlation is seen between the folded forms of SN and the truncation fragment SNOB, which is missing the last 47 residues in the protein (Figure 6B). As described previously, the differences between the RDCs of folded SN and SNOB are likely to result from changes in both structure and dynamics. As a control, we also looked at denatured SN and SNOB in the presence of 6 M urea (Figure 6C). The two proteins in their denatured states (Figure 6C) have more strongly correlated RDCs than in their folded states (Figure 6B). We performed replicate measurements of the data sets for native and denatured SN and native and denatured SNOB to obtain estimates in the errors of RDCs (Table 1). The replicate measurements also allowed us to obtain uncertainty bounds for the correlations in panels B and C of Figure 6. From cross correlations of three data sets for native SN to two data sets for native SNOB, we obtain six correlation plots with an average *R*-value and standard deviation of 0.46 ± 0.12. From cross correlations of three data sets for denatured SN to each of three data sets for denatured SNOB, we obtain nine correlation plots with an average *R*-value of 0.85 ± 0.04. The correlations obtained for RDCs show the same trend as correlations of chemical shifts between the proteins (Figure 1). Pairwise comparisons of ^1H_N shifts between the different forms of nuclease gave the following *R*-values: 0.99 for SN and SN·T, 0.72 for folded SN and SNOB, and 0.93 for denatured SN and SNOB.

The principal axes of the alignment tensors for SN·T and SN are shown in panels A and B of Figure 3. There is a slight change in the alignment tensor between SN·T and SN,

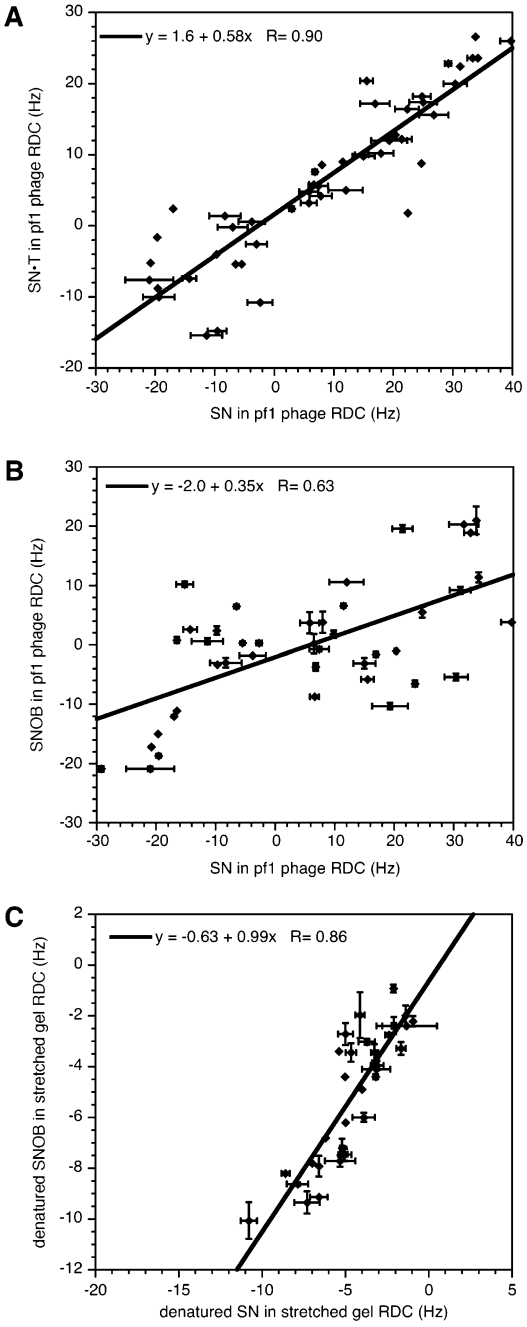

FIGURE 6: Linear correlation between RDC data sets for different forms of staphylococcal nuclease. (A) Folded SN·T and SN, both oriented in pf1 phage particles. (B) Folded SN and SNOB, in pf1 phage particles. (C) Denatured SN and SNOB (6M urea) in stretched gels. Error bars were estimated from replicate measurements, except for SN·T (*y*-axis in panel A) for which uncertainties were not determined. For data points where error bars are not given, replicate values were not determined.

and this accounts for the majority of the dispersion in the correlation between SN and SN·T (Figure 6A). When the alignment tensor obtained from the best fit of the SN·T RDC data to the SN·T X-ray structure (PDB entry 1SNC) is imposed on the apo SN structure (PDB entry 1SNC), the resulting correlation between the actual values for SN·T and those back calculated from the SN structure and the SN·T alignment tensor has an *R*-value of 0.99. By contrast, the experimental SN·T and SN RDCs correlate with an *R*-value of 0.90 (Figure 6A). This suggests that the differences between the two experimental data sets are primarily due to

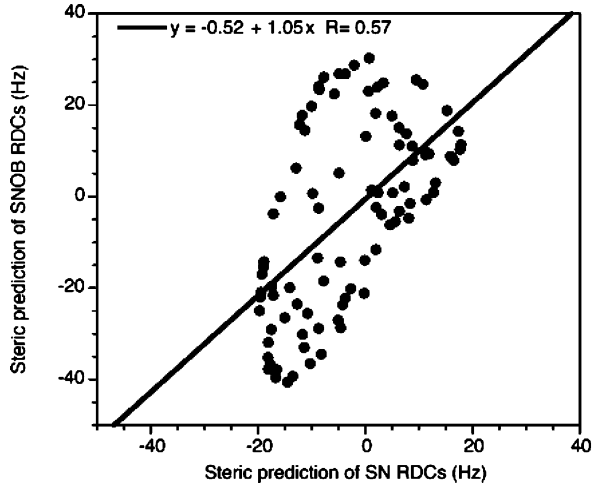

FIGURE 7: First-principles steric PALES analysis (*50*) of the effects of a hypothetical C-terminal truncation (SN residues 104−149) on RDCs. Residues 104−141 of PDB entry 1KBD were deleted with INSIGHT II (Accelrys), and residues 7−103 were used for analysis with PALES (*55*). The simulation specified an alignment in a concentration of bicelles of 5.6% and a radius of gyration for the protein of 13 Å, calculated according to the formula $R_{gyr} = 2.2(N)^{0.38}$ (*70*).

a change in alignment, rather than structure. For the very similar SN and SN·T structures, the slight difference in alignment in the phage medium may be due to a difference in the exposure of charged groups between the apoenzyme and the enzyme in the ternary complex. The analysis highlights that a correlation between two RDC data sets requires both the structure and alignment to be conserved.

There are major differences between the RDCs of SN and the SNOB fragment (Figure 6B). The increased dynamics in the SNOB fragment are partially responsible for these differences. Structural differences, as well as the difference in the overall "shape" or rhombicity of SN and SNOB, are additional factors contributing to the scatter in the RDC correlation. A first-principle steric analysis with PALES (*55*) of the orientation of SN and SNOB in bicelles predicts intrinsically different RDCs (Figure 7). The simulations were carried out for a bicelle medium rather than gels or phage, because parameters for the former are more accurate (*55*). Similar results, however, were obtained with simulations of alignment in phage. The SNOB structure was modeled as residue 1−103 of wild-type nuclease. Thus, the correlation in Figure 7 is a conservative estimate of the minimum change in RDCs that would result as a consequence of changes in the rhombicity of the protein if the last 47 residues of the wild-type nuclease were removed and the remainder of the structure stayed exactly the same. In addition to SNOB, we carried out similar simulations to examine the effects of deletions of 30 residues from the N- or C-terminus of SN. The deletions correspond to fragments that were reported (*35*) to have denatured state RDCs that were strongly correlated with those of a larger denatured fragment, Δ131Δ. *R*-values of 0.98 and 0.97 were obtained experimentally for the N- and C-terminal deletions, respectively (*35*). By contrast, our simulations show that if the deletions did not affect the remaining structure at all, the RDCs of the deletion fragments would at best correlate with those of native SN structures with *R*-values of only 0.67 and 0.80 due to the change in rhombicity when 30 residues are deleted either from the N- or C-terminus of the protein.

The relatively strong correlation of RDCs for denatured SN and SNOB (Figure 6C) is at odds with models in which the denatured state is a stiff chain, or a globular, expanded version of the native state (*35*). If the denatured state had a compact nativelike globular structure, the changes in rhombicity that would result from truncation of the SN chain to the shorter SNOB fragment should give differences in RDCs larger than those observed between the denatured states of SN and SNOB (Figures 6C and 7). If, on the other hand, the denatured protein were an extended nonglobular but stiff structure, the impact of a truncation on the anisotropy of the chain, and hence on the magnitudes of the RDCs, should be even greater (*55*). The simplest explanation for the stronger correlations observed for the RDC values of the denatured states of SN and SNOB is that long-range structural constraints no longer dominate. Structural similarity persists as borne out by the correlation of RDC values, but only along short stretches of the polypeptide chain that appear to be aligned independently of each other.

DISCUSSION

The conservation of RDCs in denatured forms of staphylococcal nuclease (SN) has previously been attributed to a conservation of folding topology (*35*). The results of this work suggest an alternative explanation. RDCs in denatured states are conserved because the structure in the denatured state is fragmented. A perturbation in one part of the protein will thus not affect the alignment of a second disconnected element of structure. By contrast, in a folded protein, structure is cooperatively stabilized. A perturbation in one part of the chain affects noncontiguous segments brought together in the three-dimensional fold of the protein. Our hypothesis is supported by analysis of RDC data sets for other proteins in the BioMagResBank. $^{1}H-^{15}N$ RDC data are available for three folded forms of troponin C aligned in pf1 phage particles (*65*): the apoprotein, the Ca^{2+}-saturated holoprotein, and the protein in complex with a 31-residue peptide from troponin I (BMRB entries 4822−4824, respectively). The *R*-values for pairwise correlations between the RDCs of the three different forms are in the range of 0.71−0.81, when $n \sim 50$. These correlations are weaker than those for various forms of denatured staphylococcal nuclease (cf. Table 1 of ref *35*). A second example, a PDZ domain of PTP-BL, has been characterized in two splice isoforms (*66*). Although the isoforms differ by only a five-residue insertion, their H−N RDCs in pf1 phage (BMRB entries 5131 and 5762) correlate with an *R*-factor of only 0.12 as a result of a change in alignment rather than structure. The RDCs of folded proteins thus appear to be very sensitive to relatively subtle changes in structure and dynamics. By contrast, RDCs are conserved in denatured proteins, as has been illustrated in this work by the direct comparison of SN and SNOB in both their folded and unfolded forms.

To obtain a good match between RDC data sets, subtle details of the structure as well as its overall alignment mode must be preserved. The conservation of alignment modes is difficult to reconcile with a model of the denatured state in which structure is fragmented and dynamic. For example, if an element of short-range structure were roughly independent of the rest of the polypetide chain, it would be improbable for the segment to randomly adopt the same orientation under different conditions. A conspicuous feature of the data for

denatured nuclease is that all the RDCs have the same sign. This constrains possible models. For example, the data do not support a model in which the topology remains native, but different segments of the polypeptide have RDCs dampened to different extents by backbone dynamics (*36*). Such a model would affect the magnitude of the RDCs but not their sign. The presence of only negative RDCs is also difficult to reconcile with a change in shape but not in structure (*36*). In a rigid molecule, RDCs of uniform sign would imply a uniform orientation of bond vectors. This could be consistent with a fibrous structure such as an extended α-helix or β-sheet but is unlikely for a globular protein. It has been proposed that the preponderance of RDCs with the same sign is consistent with extended conformations that align along the longest axes of the matrix used for alignment (*38*, *39*, *62*, *63*). Another factor that might come into play is that structures need to be anisotropic to give rise to RDCs. As such, the RDC may detect nonspherical conformation in preference to compact conformations in the denatured state ensemble. As previously mentioned, however, the RDC data of denatured SN are more complex than those predicted for an extended random coil (*62*, *63*). There is no correlation apparent between the RDC data of denatured SN and two other denatured proteins in stretched gels (*4*, *39*), as might be expected for a simple homopolymer model of the denatured state. Perhaps, as has been suggested (*38*, *39*), the complexity of the RDC data for denatured states is due to an averaged superposition of extended and folded conformations. Finally, if proteins in their denatured states aligned preferentially along the crevices of the alignment media, it would imply that the media used for alignment could affect the conformational properties of the denatured state conformational ensemble (for example, by constraining two extended segments in a parallel rather than random orientation). To examine whether the anisotropic structure necessary to observe RDCs is inherent to denatured proteins, we measured RDCs for the protein CspA in 6 M urea using the magnetic field dependence of $^1J_{HN}$ splittings (*67*). The RDCs we observed were greater than experimental error, and comparable in magnitude to those seen by the same method for native CspA (C. O. Sallum, D. Seo, E. Irimies, K. Hallenga, R. A. Kammerer, and A. T. Alexandrescu, manuscript in preparation). This result indicates that anisotropic structure is inherent to denatured proteins. However, it does not address whether the alignment medium could shift the distribution of denatured state conformations. Experiments that are sensitive to rare species in a protein's energy landscape may shed further light on this issue.

A hallmark of proteins with a well-defined tertiary structure is that atoms are densely packed. A perturbation of the structure affects multiple sites. Protein unfolding thus occurs through a very cooperative transition. Once the protein has been denatured, further changes in experimental variables are gradual, a consequence of the loss of cooperatively stabilized long-range structure. It has been hypothesized that structure in denatured proteins is primarily determined by steric interactions between side chains and the main chain (*8*, *35*). Such structure may well account for the anisotropy that gives rise to RDCs in denatured states (*35*). Nevertheless, computer simulations show that in the absence of a compacting force such as the hydrophobic effect, or specific long-range contacts, even very tight restraints of backbone dihedral angles are insufficient to specify a unique folding topology (*68*). While steric preferences for particular regions of $\phi-\psi$ space may be sufficient to ensure a locally stiff chain, even a minute variance in dihedral angles would allow enough slack in the polypeptide to result in globally heterogeneous structures (*68*). Determining the persistence length of structures in denatured proteins remains a challenging and important goal for future protein folding studies.

ACKNOWLEDGMENT

We thank John M. Flanagan (The Pennsylvania State University) for the T7/Mg plasmid encoding wild-type nuclease, Bertrand Garcia-Moreno and Daniel Karp (Johns Hopkins University) for the pdTp used in these studies, and anonymous reviewers for their suggestions. We are grateful to David Shortle (Johns Hopkins University) for the SNOB construct, and for his advice and insights into protein folding and RDCs.

SUPPORTING INFORMATION AVAILABLE

One table containing $^1H-^{15}N$ RDC values for the folded and unfolded forms of staphylococcal nuclease described in this work. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science 181*, 223−230.
2. Ghaemmaghami, S., and Oas, T. G. (2001) Quantitative protein stability measurement in vivo, *Nat. Struct. Biol. 8*, 879−882.
3. Shortle, D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability, *FASEB J. 10*, 27−34.
4. Meisner, W. K., and Sosnick, T. R. (2004) From the Cover: Fast folding of a helical protein initiated by the collision of unstructured chains, *Proc. Natl. Acad. Sci. U.S.A. 101*, 13478−13482.
5. Alexandrescu, A. T., and Rathgeb-Szabo, K. (1999) An NMR investigation of solution aggregation reactions preceding the misassembly of acid-denatured cold shock protein A into fibrils, *J. Mol. Biol. 291*, 1191−1206.
6. Dobson, C. M. (2003) Protein folding and misfolding, *Nature 426*, 884−890.
7. Baldwin, R. L., and Zimm, B. H. (2000) Are denatured proteins ever random coils? *Proc. Natl. Acad. Sci. U.S.A. 97*, 12391−12392.
8. Shortle, D. (2002) The expanded denatured state: An ensemble of conformations trapped in a locally encoded topological space, *Adv. Protein Chem. 62*, 1−23.
9. Alexandrescu, A. T., Evans, P. A., Pitkeathly, M., Baum, J., and Dobson, C. M. (1993) Structure and dynamics of the acid-denatured molten globule state of α-lactalbumin: A two-dimensional NMR study, *Biochemistry 32*, 1707−1718.
10. Alexandrescu, A. T., Abeygunawardana, C., and Shortle, D. (1994) Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: A heteronuclear NMR study, *Biochemistry 33*, 1063−1072.
11. Shortle, D. R. (1996) Structural analysis of non-native states of proteins by NMR methods, *Curr. Opin. Struct. Biol. 6*, 24−30.
12. Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR, *Chem. Rev. 104*, 3607−3622.
13. Alexandrescu, A. T., and Shortle, D. (1994) Backbone dynamics of a highly disordered 131 residue fragment of staphylococcal nuclease, *J. Mol. Biol. 242*, 527−546.
14. Alexandrescu, A. T., Jahnke, W., Wiltscheck, R., and Blommers, M. J. (1996) Accretion of structure in staphylococcal nuclease: An ^{15}N NMR relaxation study, *J. Mol. Biol. 260*, 570−587.
15. Alexandrescu, A. T., Rathgeb-Szabo, K., Rumpel, K., Jahnke, W., Schulthess, T., and Kammerer, R. A. (1998) ^{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, *Protein Sci. 7*, 389−402.

16. Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H. J. (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding, *Biochemistry 40*, 3561−3571.
17. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) Long-range interactions within a nonnative protein, *Science 295*, 1719−1722.
18. Mok, Y. K., Kay, C. M., Kay, L. E., and Forman-Kay, J. (1999) NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions, *J. Mol. Biol. 289*, 619−638.
19. Crowhurst, K. A., Choy, W.-Y., Mok, Y. K., and Forman-Kay, J. (2003) Corrigendum to the Paper by Mok et al. (1999) NOE Data Demonstrating a Compact Unfolded State for an SH3 Domain under Non-denaturing Conditions, *J. Mol. Biol. 329*, 185−187.
20. Gillespie, J. R., and 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.
21. Yi, Q., Scalley-Kim, M. L., Alm, E. J., and Baker, D. (2000) NMR characterization of residual structure in the denatured state of protein L, *J. Mol. Biol. 299*, 1341−1351.
22. Lietzow, M. A., Jamin, M., Jane Dyson, H. J., and Wright, P. E. (2002) Mapping long-range contacts in a highly unfolded protein, *J. Mol. Biol. 322*, 655−662.
23. Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling, *J. Mol. Biol. 324*, 349−357.
24. Shortle, D., and Ackerman, M. S. (2001) Persistence of native-like topology in a denatured protein in 8 M urea, *Science 293*, 487−489.
25. Prestegard, J. H., al-Hashimi, H. M., and Tolman, J. R. (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings, *Q. Rev. Biophys. 33*, 371−424.
26. Bax, A. (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics, *Protein Sci. 12*, 1−16.
27. Tjandra, N., and Bax, A. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, *Science 278*, 1111−1114.
28. Chou, J. J., Li, S., and Bax, A. (2000) Study of conformational rearrangement and refinement of structural homology models by the use of heteronuclear dipolar couplings, *J. Biomol. NMR 18*, 217−227.
29. Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude, *J. Magn. Reson. 131*, 159−162.
30. Delaglio, F., Kontaxis, G., and Bax, A. (2000) Protein structure determination using molecular fragment replacement and NMR dipolar couplings, *J. Am. Chem. Soc. 122*, 2142−2143.
31. Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1995) Nuclear magnetic dipole interactions in field-oriented proteins: Information for structure determination in solution, *Proc. Natl. Acad. Sci. U.S.A. 92*, 9279−9283.
32. Fischer, M. W., Losonczi, J. A., Weaver, J. L., and Prestegard, J. H. (1999) Domain orientation and dynamics in multidomain proteins from residual dipolar couplings, *Biochemistry 38*, 9013−9022.
33. Yu, L., Gunasekera, A. H., Mack, J., Olejniczak, E. T., Chovan, L. E., Ruan, X., Towne, D. L., Lerner, C. G., and Fesik, S. W. (2001) Solution structure and function of a conserved protein SP14.3 encoded by an essential *Streptococcus pneumoniae* gene, *J. Mol. Biol. 311*, 593−604.
34. Ackerman, M. S., and Shortle, D. (2002) Molecular alignment of denatured states of staphylococcal nuclease with strained polyacrylamide gels and surfactant liquid crystalline phases, *Biochemistry 41*, 3089−3095.
35. Ackerman, M. S., and Shortle, D. (2002) Robustness of the long-range structure in denatured staphylococcal nuclease to changes in amino acid sequence, *Biochemistry 41*, 13791−13797.
36. Ohnishi, S., Lee, A. L., Edgell, M. H., and Shortle, D. (2004) Direct demonstration of structural similarity between native and denatured eglin C, *Biochemistry 43*, 4064−4070.
37. Ding, K., Louis, J. M., and Gronenborn, A. M. (2004) Insights into conformation and dynamics of protein GB1 during folding and unfolding by NMR, *J. Mol. Biol. 335*, 1299−1307.
38. Fieber, W., Kristjansdottir, S., and Poulsen, F. M. (2004) Short-range, long-range and transition state interactions in the denatured state of ACBP from residual dipolar couplings, *J. Mol. Biol. 339*, 1191−1199.
39. Mohana-Borges, R., Goto, N. K., Kroon, G. J., Dyson, H. J., and Wright, P. E. (2004) Structural Characterization of Unfolded States of Apomyoglobin using Residual Dipolar Couplings, *J. Mol. Biol. 340*, 1131−1142.
40. Alexandrescu, A. T., and Kammerer, R. A. (2003) Structure and disorder in the ribonuclease S-peptide probed by NMR residual dipolar couplings, *Protein Sci. 12*, 2132−2140.
41. Baldwin, R. L. (2002) A new perspective on unfolded proteins, *Adv. Protein Chem. 62*, 361−367.
42. Zagrovic, B., Snow, C. D., Khaliq, S., Shirts, M. R., and Pande, V. S. (2002) Native-like mean structure in the unfolded ensemble of small proteins, *J. Mol. Biol. 323*, 153−164.
43. Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thiyagarajan, P., Sosnick, T. R., Hasan, M. Z., Pande, V. S., Ruczinski, I., Doniach, S., and Plaxco, K. W. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins, *Proc. Natl. Acad. Sci. U.S.A. 101*, 12491−12496.
44. Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins, *Proc. Natl. Acad. Sci. U.S.A. 101*, 12497−12502.
45. Hynes, T. R., and Fox, R. O. (1991) The crystal structure of staphylococcal nuclease refined at 1.7 Å resolution, *Proteins 10*, 92−105.
46. Loll, P. J., and Lattman, E. E. (1989) The crystal structure of the ternary complex of staphylococcal nuclease, Ca^{2+}, and the inhibitor pdTp, refined at 1.65 Å, *Proteins 5*, 183−201.
47. Alexandrescu, A. T., Gittis, A. G., Abeygunawardana, C., and Shortle, D. (1995) NMR structure of a stable "OB-fold" subdomain isolated from staphylococcal nuclease, *J. Mol. Biol. 250*, 134−143.
48. Flanagan, J. M., Kataoka, M., Fujisawa, T., and Engelman, D. M. (1993) Mutations can cause large changes in the conformation of a denatured protein, *Biochemistry 32*, 10359−10370.
49. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol. 131*, 266−280.
50. Zweckstetter, M., and Bax, A. (2001) Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage, *J. Biomol. NMR 20*, 365−377.
51. Wang, Y., and Shortle, D. (1995) The equilibrium folding pathway of staphylococcal nuclease: Identification of the most stable chain-chain interactions by NMR and CD spectroscopy, *Biochemistry 34*, 15895−15905.
52. Ye, K., and Wang, J. (2001) Self-association reaction of denatured staphylococcal nuclease fragments characterized by heteronuclear NMR, *J. Mol. Biol. 307*, 309−322.
53. Lukin, J. A., Kontaxis, G., Simplaceanu, V., Yuan, Y., Bax, A., and Ho, C. (2003) Quaternary structure of hemoglobin in solution, *Proc. Natl. Acad. Sci. U.S.A. 100*, 517−520.
54. Hus, J. C., Marion, D., and Blackledge, M. (2001) Determination of protein backbone structure using only residual dipolar couplings, *J. Am. Chem. Soc. 123*, 1541−1542.
55. Zweckstetter, M., and Bax, A. (2000) Predictions of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR, *J. Am. Chem. Soc. 122*, 3791−3792.
56. Chen, J., Lu, Z., Sakon, J., and Stites, W. E. (2000) Increasing the thermostability of staphylococcal nuclease: Implications for the origin of protein thermostability, *J. Mol. Biol. 303*, 125−130.
57. Riechmann, L., and Winter, G. (2000) Novel folded protein domains generated by combinatorial shuffling of polypeptide segments, *Proc. Natl. Acad. Sci. U.S.A. 97*, 10068−10073.
58. Murzin, A. G. (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences, *EMBO J. 12*, 861−867.
59. Alexandrescu, A. T., Jaravine, V. A., Dames, S. A., and Lamour, F. P. (1999) NMR hydrogen exchange of the OB-fold protein LysN as a function of denaturant: The most conserved elements of structure are the most stable to unfolding, *J. Mol. Biol. 289*, 1041−1054.
60. Hodel, A., Kautz, R. A., and Fox, R. O. (1995) Stabilization of a strained protein loop conformation through protein engineering, *Protein Sci. 4*, 484−495.

61. Meier, S., Guthe, S., Kiefhaber, T., and Grzesiek, S. (2004) Foldon, The Natural Trimerization Domain of T4 Fibritin, Dissociates into a Monomeric A-state Form containing a Stable β-Hairpin: Atomic Details of Trimer Dissociation and Local β-Hairpin Stability from Residual Dipolar Couplings, *J. Mol. Biol. 344*, 1051−1069.
62. Louhivuori, M., Paakkonen, K., Fredriksson, K., Permi, P., Lounila, J., and Annila, A. (2003) On the origin of residual dipolar couplings from denatured proteins, *J. Am. Chem. Soc. 125*, 15647−15650.
63. Louhivuori, M., Fredriksson, K., Paakkonen, K., Permi, P., and Annila, A. (2004) Alignment of chain-like molecules, *J. Biomol. NMR 29*, 517−524.
64. Barrientos, L. G., Dolan, C., and Gronenborn, A. M. (2000) Characterization of surfactant liquid crystal phases suitable for molecular alignment and measurement of dipolar couplings, *J. Biomol. NMR 16*, 329−337.
65. Paakkonen, K., Sorsa, T., Drakenberg, T., Pollesello, P., Tilgmann, C., Permi, P., Heikkinen, S., Kilpelainen, I., and Annila, A. (2000) Conformations of the regulatory domain of cardiac troponin C examined by residual dipolar couplings, *Eur. J. Biochem. 267*, 6665−6672.
66. Walma, T., Aelen, J., Nabuurs, S. B., Oostendorp, M., van den Berk, L., Hendriks, W., and Vuister, G. W. (2004) A closed binding pocket and global destabilization modify the binding properties of an alternatively spliced form of the second PDZ domain of PTP-BL, *Structure (Cambridge) 12*, 11−20.
67. Tjandra, N., Grzesiek, S., and Bax, A. (1996) Magnetic field dependence of nitrogen-proton *J* splittings in ^{15}N-enriched human ubiquitin resulting from relaxation interference and residual dipolar coupling, *J. Am. Chem. Soc. 118*, 6264−6267.
68. Alexandrescu, A. T. (2004) Strategy for supplementing structure calculations using limited data with hydrophobic distance restraints, *Proteins 56*, 117−129.
69. Cornilescu, G., Marquardt, J. L., Ottiger, M., and Bax, A. (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase, *J. Am. Chem. Soc. 120*, 6836−6837.
70. Skolnick, J., Kolinski, A., and Ortiz, A. R. (1997) MONSSTER: A method for folding globular proteins with a small number of distance restraints, *J. Mol. Biol. 265*, 217−241.

BI0473410