Characterizing the Residue Level Folding of the Intrinsically Unstructured IA_3†

Omjoy K. Ganesh,‡ Terry B. Green,§ Arthur S. Edison,‡,⊥ and Stephen J. Hagen*,‡,||

Department of Biochemistry & Molecular Biology, University of Florida, Gainesville, Florida 32610-0245, Center for Medical, Agricultural, and Veterinary Entomology, USDA/ARS, Gainesville, Florida 32608, McKnight Brain Institute and National High Magnetic Field Laboratory, University of Florida, Gainesville, Florida 32610, and Department of Physics, University of Florida, Gainesville, Florida 32611-8440

Received July 5, 2006; Revised Manuscript Received September 7, 2006

ABSTRACT: Residue level analysis of the folding of simple proteins may hold the key to understanding folding pathways and aid in structure prediction. IA_3, the endogenous inhibitor of yeast aspartic proteinase A (YPrA), is an unstructured protein in solution. Comparison of the 2D ^{15}N-HSQC spectra of IA_3 in water and in 23% 2,2,2-trifluoroethanol (TFE) shows that the individual residue cross peaks of IA_3 become more dispersed in the presence of TFE, indicating that the protein undergoes an unstructured to structured transition in the presence of TFE. This transition can be monitored by the movements of the cross peaks. Following the individual cross peaks, however, is complicated and does not establish whether a single transition occurs globally in the sequence. In this equilibrium study, we apply singular value decomposition (SVD) to elucidate both the main features of the TFE-driven transition and the residue-level deviations from the average behavior. This analysis has yielded a two-state folding description as well as specifics of NMR frequency shifts of individual residues, indicating that the N-terminus of IA_3 has a higher helical propensity than the C-terminus. Additionally, we discuss possible mechanisms for observed deviations from a two-state folding transition. When combined with a traditional biochemical understanding of interactions between individual residues, this approach leads to a better understanding of protein folding.

IA_3 is a 7.7 kDa protein that is found in the cytoplasm of *Saccharomyces cerevisiae* and is an endogenous inhibitor of the yeast aspartic proteinase YPrA. Previous biophysical studies have shown that in the absence of YPrA, IA_3[1] exists as an unstructured protein in solution (*1*, *2*). In the X-ray crystal structure of the YPrA·IA_3 complex (PDB accession number 1DPJ), only the N-terminal residues Thr 3−Lys 31 of the 68 amino acid protein are seen; this N-terminus forms an α-helix in the active site cleft of YPrA. The remaining C-terminal residues are believed to be disordered, since electron density from this region is not observed in the crystal structure (*2*). Only the 34 amino acid N-terminus of IA_3 is necessary to inhibit YPrA with the same nanomolar potency as the full-length wild-type IA_3 (*3*).

Using far-UV circular dichroism (CD) we have previously shown that 2,2,2-trifluoroethanol (TFE) facilitates a two-state transition in IA_3 from unstructured protein to α-helix with a transition midpoint around 16.3% (v/v) TFE (*1*). Due to the global nature of CD, these studies could not establish which residues in the IA_3 sequence became helical, or whether the transition is uniformly two-state for all the residues.

NMR chemical shift data from ^{15}N heteronuclear single quantum correlation (^{15}N-HSQC) experiments on proteins contain important structural information (*4*). Folded proteins show greater distribution of chemical shift dispersion, while spectral cross peaks of unfolded proteins are less dispersed with possible overlap in NMR frequencies (*5*, *6*). Here, we use chemical shift data from ^{15}N-HSQC TFE titrations in conjunction with singular value decomposition (SVD) to monitor the transition of IA_3 in the presence of TFE (*7*, *8*). Assignment of the protein through triple-resonance methods at intermediate TFE concentrations was used to aid in tracking the cross peaks as the TFE concentration varies. We also use parameters derived from a two-state fit to the TFE-driven transition to extrapolate to the high and low TFE states to extend our current understanding of how the protein folds.

The tracks taken by individual chemical shifts, in response to the addition of TFE, contain a large amount of information, and it can be difficult to extract a clear picture of either the general trends or the individual, residue-level departures from the general trends. One approach to this problem could be to use chemical shift mapping (CSM) to identify residues that are undergoing large changes. CSM is also generally used to identify interactions between proteins and nucleic

† S.J.H. acknowledges the support of the National Science Foundation, MCB #0347124. Additional support was provided by NIH research Grant NCRR P41 RR016105.

* Address correspondence to Dr. Stephen J. Hagen, Department of Physics, University of Florida, P.O. Box 118440, Gainesville, Florida 32611-8440. Phone, 352-392-4716; fax, 352-392-7709; e-mail, sjhagen@ufl.edu.

‡ Department of Biochemistry & Molecular Biology, University of Florida.

§ Center for Medical, Agricultural, and Veterinary Entomology, USDA/ARS.

⊥ McKnight Brain Institute and National High Magnetic Field Laboratory, University of Florida.

|| Department of Physics, University of Florida.

[1] Abbreviations: CD, far-UV circular dichroism; IA_3, inhibitor of the aspartic proteinase YPrA, yeast proteinase A; ^{15}N-HSQC, nitrogen-15 heteronuclear single quantum correlation spectroscopy; NMR, nuclear magnetic resonance; SVD, singular value decomposition; TFE, 2,2,2-trifluoroethanol.

10.1021/bi061358w CCC: $33.50
Published on Web 10/20/2006

acids, other proteins, or membranes (*9*−*12*). In this method, the chemical shift of each residue is assessed and mapped onto an image of the primary structure or the folded structure of the protein (*9*). In the present case, we do not have a structure for the high-TFE state, and we cannot assume that this structure is identical to the crystal structure obtained for the biologically active complex. More importantly, CSM identifies the residues involved in an interaction but does not reveal the nature of the transition or indicate how closely an individual residue does or does not follow that of other assigned residues.

As an alternative, we used singular value decomposition (SVD). SVD is a mathematical tool for identifying the most important patterns in a dataset (*13*). It breaks apart a two-dimensional matrix of data (in this case, chemical shifts recorded for an array of sequence positions and TFE concentrations) into a sum of simpler matrices. Each of those component matrices contains a single two-dimensional pattern, that is, a particular pattern of chemical shifts versus sequence position, which varies with a particular TFE dependence. The original data matrix, with all of its sequence-dependent and TFE-dependent variations, can be reconstituted exactly by adding together all of the component matrices. The primary benefit of applying SVD to a complex data set is that it essentially identifies all the major patterns that are present in the data and ranks them in order of descending strength or significance. SVD has been an invaluable tool in other areas of biophysical spectroscopy, such as transient absorbance spectroscopy (*13*).

By applying SVD to the chemical shift data, we show that the major transition is two-state and that it occurs in the same residues necessary to inhibit YPrA. Additionally, another component to the folding is shown to be sensitive to changes in the amide proton chemical shift and may be indicative of changes to the hydrogen bonding of backbone amide protons. A biochemical interpretation of the residue interactions, combined with the global description provided by SVD, results in a relatively complete residue-level picture of the structural changes induced by TFE. Our results demonstrate how the application of SVD to a two-dimensional NMR dataset can efficiently extract both a detailed description of the global structural changes in a molecule and the rather small, residue-specific deviations from that global behavior.

EXPERIMENTAL PROCEDURES

Sample Preparation. IA_3 was expressed in *Escherichia coli* and purified as previously described by Phylip et al. (*3*). IA_3 samples were prepared at a concentration of 1 mM sample in 250 μL of 50 mM PO_4 buffer, pH 5.6, as described previously (*1*), with the inclusion of increasing volumes of TFE. A TFE titration from 0% to 23% (v/v) was made by the addition of constant volumes of 5 μL of TFE to the IA_3 sample. In addition, a fresh IA_3 sample was prepared at the endpoints and middle of the TFE titration for triple resonance experiments.

NMR Data Acquisition. All NMR data were collected on Bruker Avance spectrometers in the Advanced Magnetic Resonance Imaging and Spectroscopy (AMRIS) facility at the University of Florida's McKnight Brain Institute. ^{1}H data were referenced to DSS (2,2-dimethyl-2-silapentane-5-sulfonic acid) (0.0 ppm). ^{15}N and ^{13}C 0.0 Hz frequencies were obtained through indirect referencing to DSS using the chemical shift ratios provided by the BMRB database (*1*, *14*). All NMR experiments were run at 20 °C.

Two-dimensional ^{15}N-HSQC (*15*, *16*) spectra of the TFE titrations were collected at 750 MHz on a Bruker Avance (DRX)-750 console in a 17.6 T magnet using either a 2.5 or 5 mm TXI probe. The ^{1}H carrier frequency was centered on water which was reduced using a 3-9-19 WATERGATE sequence (*15*). ^{15}N-HSQC spectra were collected with 2048 complex points in the ^{1}H dimension, and 256 complex points in the ^{15}N dimension. Quadrature detection in the indirect dimension was achieved using the States-TPPI method (*17*). All 2D spectra were collected with a 10.69 ppm (8012.82 Hz) spectral width in the proton dimension, and 30.0 ppm (2280.30 Hz) in the nitrogen dimension. The carrier frequencies for ^{1}H and ^{15}N were centered on 4.86 and 116.68 ppm, respectively.

Triple resonance data were collected at 600 MHz on a Bruker Avance (DRX)-600 console in a 14.1 T magnet equipped with a 5 mm TXI probe. Spectra were collected with a 11.02 ppm (6613.76 Hz) spectral width in the proton dimension, and 21.7 ppm (1315.79 Hz) in the nitrogen dimension. For the HNCACB (*18*) and CBCA(CO)NH (*19*), the spectral widths were 50.9 ppm (7692.31 Hz) and carrier frequencies were 42.51 ppm for the carbon nuclei. The HNCO (*20*, *21*) carbonyl spectral width was 9.5 ppm (1428.57 Hz) with a carrier frequency of 176.09 ppm. Quadrature detection in the ^{15}N dimension was achieved using the Echo-AntiEcho method (*22*−*24*) and, in the ^{13}C dimension, using the States-TPPI method. All three-dimensional spectra were collected with 1024 complex points in the $^1H^N$ dimension. HNCO spectra were collected with 32 complex points in the ^{15}N dimension and 30 complex points in the ^{13}C dimension. HNCACB spectra were collected with 32 complex points in the ^{15}N dimension and 50 complex points in the ^{13}C dimension. CBCA(CO)NH spectra were collected with 18 complex points in the ^{15}N dimension and 50 complex points in the ^{13}C dimension.

Processing and Referencing of NMR Data. Multidimensional spectra were processed using NMRPipe (*25*), and the resulting spectra were analyzed with NMRView software (One Moon Scientific, Trenton, NJ) (*26*). We applied cosine-squared apodization functions in the direct and the indirect dimensions during processing of the 2D ^{15}N-HSQC spectra. To increase the digital resolution, spectra were zero-filled to 4096 points in the proton dimension, and to 1024 points in the indirect dimension. Polynomial baseline correction was applied in both dimensions.

For the CBCA(CO)NH, HNCACB, and HNCO triple resonance spectra, the direct dimension was zero-filled to 2048 points. Cosine-squared apodization and zero-filling were applied in all dimensions. Linear prediction was used to expand the data in the indirect carbon dimension only. We used model-free baseline correction to remove noise from the resulting 3D spectra (*25*, *27*).

Assignments and Cross Peak Tracking. Resonance assignments were made manually using standard triple resonance methods. We tracked individual residue chemical shifts in the 2D ^{15}N-HSQC as a coordinate from the centroid of the cross peak. The cross peak $^1H^N$ and ^{15}N coordinates of individual residues were tracked as they varied with TFE concentration and displayed as two-dimensional tracks. We

tracked only cross peaks for which an unambiguous assignment was possible and where the cross peak could be followed through the whole TFE titration. Of the 68 residues in IA_3, 40 cross peaks were assigned and tracked in this study.

SVD Analysis. Random-coil chemical shifts and sequence-dependent correction values were obtained from methods developed by Schwaringer et al. (*28*, *29*), using a program written in PERL by the authors. The chemical shift data, after this correction, can be represented by two separate matrices containing the $^1H^N$ and ^{15}N values, respectively. Each matrix has 40 rows corresponding to the 40 different residues tracked and 16 columns corresponding to the 16 different TFE concentrations studied; the element in row *i* and column *j* then gives the value of the NMR chemical shift for residue *i* at TFE concentration *j*. We then decomposed each matrix by singular value decomposition (SVD). SVD is a method for producing a unique factorization of a given matrix as the product of three basis matrices (*30*−*33*): it extracts the essential features of a data matrix, preserving the significant information in the matrix while sharply reducing the number of parameters needed to represent that information. As SVD itself is a purely mathematical and model-independent procedure, it lacks intrinsic bias toward any particular interpretation of the data and, as such, reliably identifies interesting features in the data that rise above the noise background.

SVD decomposes a data matrix **A** into the matrix product,

$$\mathbf{A} = \mathbf{U}\cdot\mathbf{S}\cdot\mathbf{V}^{\mathbf{T}} \tag{1}$$

where, for a matrix **A** of size 40 × 16, **U** is a matrix of size 40 × 40, **V** is a matrix of size 16 × 16, and **S** is a matrix of size 40 × 16. The 16 diagonal elements of **S** are the only nonzero elements and are known as the singular values, s_n (with $n = 1-16$). The columns of **U** comprise an orthonormal set of 40-element vectors, u_n ($n = 1-40$), and the columns of **V** comprise an orthonormal set of 16-element vectors v_n ($n = 1-16$).

From eq 1, the data matrix **A** can be reconstructed exactly as the sum

$$\mathbf{A} = \sum_{n=1}^{16} u_n \cdot s_n \cdot v_n^T \tag{2}$$

In our case the vectors u_n may be interpreted as functions of sequence position (residue number), and the vectors v_n as functions of TFE concentration. The dataset can then be reconstructed from the series

$$\mathbf{A}(\text{residue, TFE}) = u_1\,(\text{residue})\cdot s_1 \cdot v_1(\text{TFE})^{\text{T}} + u_2\,(\text{residue})\cdot s_2 \cdot v_2(\text{TFE})^{\text{T}} + \ldots \tag{3}$$

Since the singular values are ordered from largest (s_1) to smallest (s_{16}), and the *u* and *v* vectors are normalized, the weight or importance of each successive term in this series is smaller than that of the preceding term. More precisely, the sum of the first *n* terms in the series provides the best (in the least-squares sense) *n*-component representation of the complete dataset **A** (*33*).

Fitting and Rotation Matrices. If the response to TFE is primarily the induction of a two-state folding transition, then this transition should be represented in the **V** (TFE) data and can be described by the fraction of molecules folded, F_f (*34*):

$$F_f = [1 + \exp(-\Delta G/RT)]^{-1} \tag{4}$$

where ΔG, the free energy of folding (or unfolding), is given by $\Delta G = \Delta G_0 + m[\text{TFE}]$ in the conventional case of a linear relationship between folding free energy and TFE concentration (*35*). This two-state model requires only two free parameters, *m* and ΔG_0, to fit the data. A two-state transition will be represented by two SVD components. Therefore, we expect the v_1 and v_2 vectors obtained from both the $^1H^N$ and ^{15}N datasets to have the TFE dependence of eq 4. We fit these two vectors in both datasets simultaneously in order to find the parameters. The residue dependence should then follow $s\cdot v_n(\text{TFE}) = a_n + b_n F_f$ for the constants a_n, and b_n, $n = 1, 2$.

Extrapolation to TFE → + ∞ or TFE → − ∞ should then result in limiting spectra. TFE → + ∞ would correspond to a high TFE state with a large population of folded molecules ($F_f \approx 1$); while TFE → − ∞ would correspond to a low TFE state with a high population of unfolded molecules ($F_f \approx 0$). The limiting values of v_1 and v_2 can be calculated as $sv_n = a_n + b_n$ for the high TFE state and $sv_n = a_n$ for the low TFE state. Combining the limiting TFE dependence with the u_1 and u_2 vectors produces the high- and low-limiting spectra:

$$\text{limiting spectra} = [u_1 \;\; u_2]\cdot\begin{bmatrix} sv_1 \\ sv_2 \end{bmatrix} \tag{5}$$

The first two **V**(TFE) components were fit extremely well by a two-state function (eq 4). The vectors of each of the basis sets **U** and **V** produced from the SVD are orthogonal; therefore, the physical phenomena of interest in the data are most likely spread across several of the SVD components. Accordingly, it was noted that rotation of the components of the $^1H^N$ and ^{15}N basis sets would minimize the deviations between the two-state fit and the first two SVD components, resulting in more accurate ΔG_0 and *m*-values. A rotation of the basis vectors can preserve the orthogonality of the set, while ensuring that separate physical phenomena are maximally represented in the minimal number of components (one or few). The rotation, however, is not a fit or approximation and does not change or reduce the information contained in the data.

Rotation of **V** requires us to find an optimal rotation matrix **R** (*36*). To find **R**, we rotated the dataset while simultaneously fitting to the two-state equation (eq 4). In doing so, we were able to find a new set of basis vectors that spanned the same vector space, yet also provided the best fit to the two-state equation for v_1 and v_2, while preserving the orthonormality of v_3. The third component would then also contain deviations which did not conform to a two-state fit.

RESULTS

Initial NMR Analysis. Using the ^{15}N-HSQC, individual residue chemical shifts were tracked as a $^1H^N$ or ^{15}N coordinate and as a function of TFE. We subtracted random coil values (*28*) from all chemical shift values and corrected the resulting values for sequence-dependent effects

(*29*). Several residues could not be tracked throughout the titration, due to overlap of cross peaks at some concentrations, or due to the difficulty of completely assigning all peaks in the TFE titration. However, 40 of the 68 residues in IA_3 were fully tracked through the titration. Initial analysis of these residue tracks indicates that there is a general tendency toward an α-helical structure according to cutoff values given by Wishart and Sykes (*4*) (Supporting Information). As shown in Figure 1, the tracks of many of the residues throughout the protein exhibit a large curvature while moving in a generally α-helical direction with the addition of TFE. $^{13}C^{\alpha}$ chemical shifts at a few TFE concentrations verify the helical transition (data not shown). However, several residues do not follow this general trend. Additionally, residues such as Asp 22, Asp 28, Phe 30, Ala 34, Ala 45, Glu 49, Tyr 53, and Asn 58 move in a manner contrary to the general motion of the other tracked residues of the protein.

Evaluation of SVD Analysis. The ^{15}N and $^{1}H^{N}$ coordinates of the chemical shift tracks were collected as separate data matrices as described in Experimental Procedures. We used SVD to decompose the chemical shift data into three matrices, **U**, **S**, and **V**. The **V** matrix describes the TFE dependence of the chemical shift, with the columns of **V** representing the TFE dependence. The columns of the **U** matrix represent the residue dependence of the chemical shift. **S** is a diagonal matrix of singular values, with diagonal elements representing the weights for the corresponding columns of the **U** and **V** matrices. Hence the *n*th column of the **V** matrix, together with the *n*th column of **U** and the *n*th diagonal element of **S**, together comprise the *n*th SVD component of the data (eq 2).

The weights of the first three components from the SVD describe more than 96% of the information in the chemical shift data matrix (Figure 2). All remaining higher order SVD components are less than 1% of the total chemical shift data. Further, the TFE dependence and residue dependence of the higher components appeared randomly distributed around zero, indicating that these components essentially reflect noise in the data (Figure 3). Therefore, in the discussion that follows, we focus only on the first three components of the SVD analysis.

TFE Dependence. The overall effect of TFE on the coil-to-helix transition of IA_3 can be evaluated from Figure 4. It shows the TFE dependence of the chemical shifts as a function of increasing TFE concentration for the first three components of the SVD analysis. Plots of the first three components of this TFE dependence revealed that both the $^{1}H^{N}$ and ^{15}N exhibited similar responses to TFE. The first two components of the TFE dependence of both the $^{1}H^{N}$ and ^{15}N chemical shifts exhibit sigmoidal trends with increasing amounts of TFE. These movements are strongly indicative of a two-state transition (*37*), where one state is progressively depopulated in favor of a second state as the TFE concentration rises. The third component in each figure exhibits a weak maximum near 15% (v/v) TFE. This third component corresponds to only a small portion of the signal, around 4.1% of the $^{1}H^{N}$ and 2.6% of the ^{15}N chemical shifts (Figure 2).

We fit the first two components of both the $^{1}H^{N}$ and ^{15}N simultaneously to a two-state folding−unfolding model. This model implicitly assumes that the free energy of folding is linear in the TFE concentration. The fit yielded a folding free energy in water, ΔG_0, of −21.2 kJ/mol with an *m*-value, which measures the helix stabilizing effect of TFE, of 1.2 $(kJ\ mol)^{-1}$ at 20 °C. Simultaneous fitting of the first and second components of the TFE dependence of the chemical shift for both $^{1}H^{N}$ and ^{15}N cross peak movements yielded a TFE refolding midpoint of 18.3%, which was in agreement with the TFE refolding midpoint obtained from far-UV circular dichroism (CD) studies of IA_3 in increasing concentrations of TFE (*1*). Fitting of $^{1}H^{N}$ and ^{15}N datasets separately yielded TFE refolding midpoints of 18.4% and 18.1%, respectively, indicating that the model provides a robust fit to the data (Table 1).

Residue Dependence of Chemical Shift. Plotting the residue dependence of the chemical shifts versus the amino acid sequence provides insight into the folding of IA_3 at the residue level (Figure 4). Bar plots of the first through third components of the residue dependence on the chemical shift proved to be very revealing. The first component of the $^{1}H^{N}$ residue dependence indicates that IA_3 exhibits a large tendency toward α-helical structure throughout the length of the sequence. This tendency is present, albeit not as clearly, in the first component of the ^{15}N residue dependence. The second component of the residue dependence for the $^{1}H^{N}$ and ^{15}N can be broken down into distinct regions. In $^{1}H^{N}$, the N-terminal half of the protein (residues 1−28) exhibits a negative trend. This trend appears to be minimized or reversed in the C-terminal portion of the protein. A central region of positive values, with flanking negative values, can also be seen around residues 28−31. The second component of the ^{15}N residue dependence also shows this trend with the N-terminus showing generally larger negative changes as compared to the C-terminus, with residues 29 and 30 being positive. The third component residue dependence indicates negative deviations in both ^{1}H and ^{15}N.

The limiting high and low TFE spectra obtained from the fit and the SVD components are shown in Figure 5. Figure 6 shows the 0% and 23% ^{15}N-HSQC TFE spectra, along with the same spectra as reconstructed from just the first three SVD components (i.e., omitting all higher order components). The original and reconstructed spectra agree well, illustrating that discarding the higher order components of the SVD has a negligible effect on the initial NMR data, and that the rotation included in our SVD treatment preserves all the information in the data.

The Third SVD Component. Panels G and H of Figure 4 illustrate the deviations seen in the third component. All the deviations seen in this component are small in both the $^{1}H^{N}$ and ^{15}N dimension. Residues such as Thr 3, Gly 40, Glu 49, and Tyr 53 show bars which are very small or nearly zero. A very small third component, or the apparent lack of a third component, indicates that the chemical shift tracks of these residues can be described primarily by the first and second components, both of which were well-fit by a two-state transition. Ser 9, Val 26, Gly 62, and other residues which are influenced by this third component are indicated by very large bars in panels G and H.

Reconstruction of the NMR chemical shift tracking spectra based only on the third SVD component and residuals allowed us to evaluate the third component (Figure 7). Residues which have no third component in $^{1}H^{N}$ and ^{15}N would be represented here as a track in which all the points

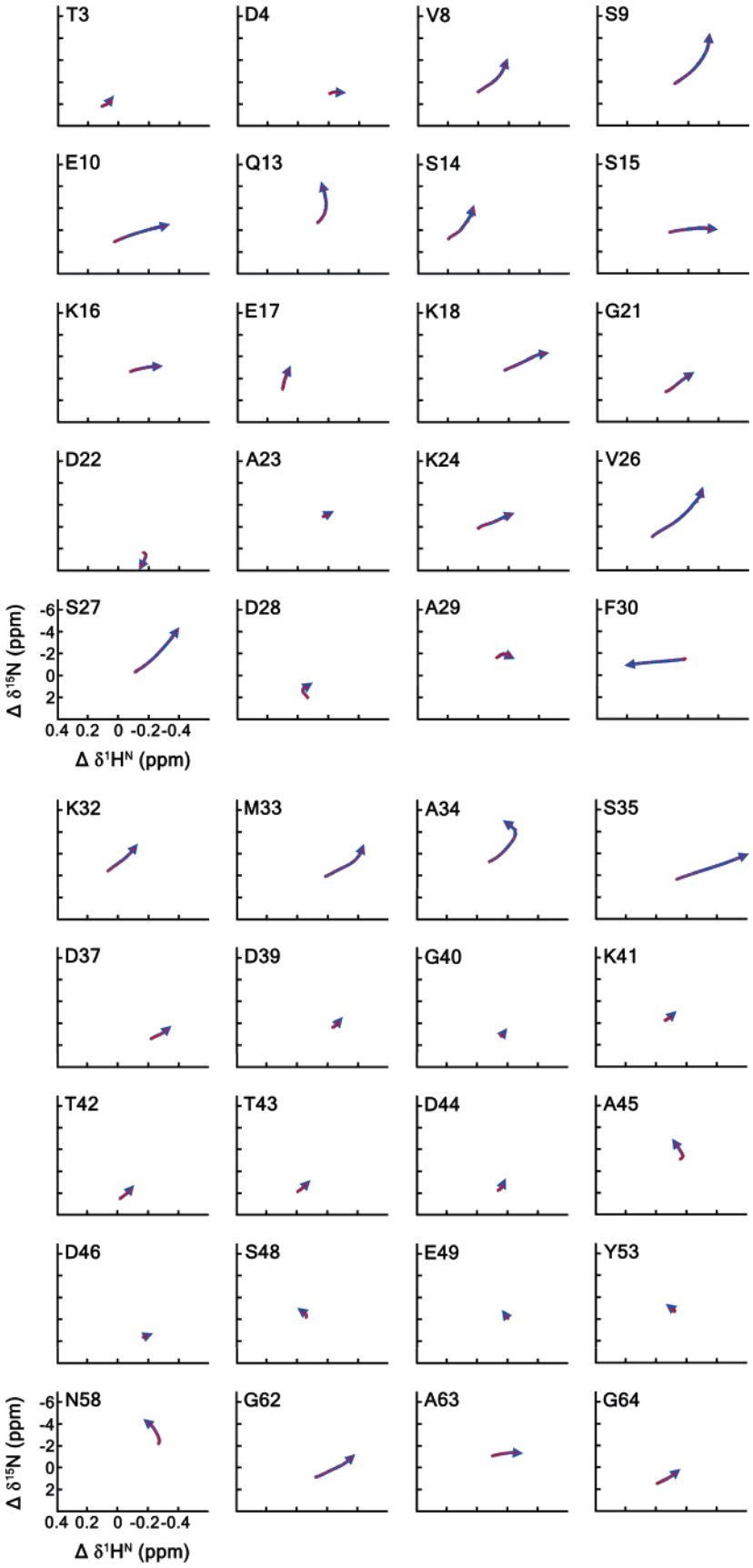

FIGURE 1: Changes in the chemical shifts ($^1H^N$, *x*-axis; ^{15}N, *y*-axis), in response to increasing TFE concentration, for the 40 residues of the IA_3 sequence followed in this study. Random coil values have been subtracted from the $^1H^N$ and ^{15}N chemical shifts (*28*, *29*). Individual residue chemical shifts were tracked as a function of TFE concentration from 0% TFE (v/v) to 23% TFE, via analysis of sequential ^{15}N-HSQC spectra. Red dots indicate data points, arrowhead indicates direction of track as TFE concentration rises.

lie on top of each other. Closer inspection of such a track would reveal that the points have small deviations from one another and are randomly distributed. Figure 7 provides several examples of this: Ala 23, Phe 30, Gly 40, Thr 43,

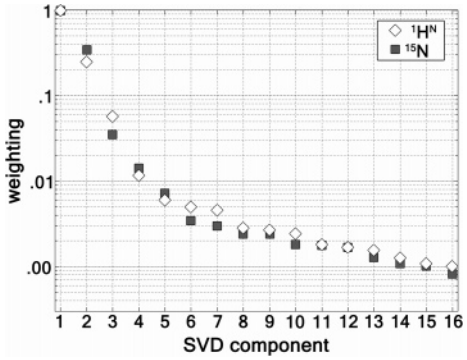

FIGURE 2: Relative weighting of the SVD components obtained from the $^1H^N$ (◇) and ^{15}N (■) chemical shifts. The relative weighting of component *i* is defined as weighting = s_i/s_1, where s_i is the *i*th element along the diagonal of the weighting matrix **S** obtained by SVD of the data matrix **A** (see eq 1). The first three components of the SVD analysis constitute 96.4% and 96.7% of the $^1H^N$ and ^{15}N matrices, respectively. Each of the remaining components of the $^1H^N$ and ^{15}N decompositions is under 1%. This analysis indicates that the $^1H^N$ and ^{15}N chemical shift matrices can be represented by the first three components of the SVD analysis.

Glu 49, and Tyr 53. Residues such as Ser 9, Val 26, Lys 32, Met 33, Ala 34, and Gly 62 in Figure 7 are examples of large deviations of the third component seen in Figure 4. From Figure 7, we can see that deviations in the third component primarily manifest themselves as chemical shift movements along the $^1H^N$ axis.

DISCUSSION

The native IA_3 provides an ideal model system to study the conformational features of intrinsically unstructured proteins. It is known to be unfolded in solution; however, the N-terminus of the protein forms an α-helix upon binding to YPrA (*2*). We have previously shown by far-UV CD that TFE can induce an α-helical structure in IA_3 (*1*), and here, we expand this work using NMR spectroscopy along with SVD analysis to understand changes in chemical shift behavior to gain an insight into the folding of IA_3.

TFE Driven Transition in the NMR Data. It is well-understood that $^1H^N$ shifts are influenced by local electronic environment (*38*) and ^{15}N shifts are sensitive to amino acid type and the effect of neighboring amino acids. Both $^1H^N$ and ^{15}N shifts are influenced by hydrogen bonding and the φ/ψ_{i-1} backbone dihedral angles (*4*, *39*, *40*). In general, cross peaks which move to lower frequencies are indicative of helical structure, while cross peaks which move to higher frequencies indicate β structure (*4*, *38*).

Subtraction of random coil and sequence-dependent effects from the chemical shift data allowed us to evaluate the change in secondary structure of a given residue due to the increase of TFE. According to the analysis of Wishart et al, a residue track starting near the origin and ending in the upper-right of each graph in Figure 1 would clearly indicate a transition from coil to helix. Titration tracks generated by following residue chemical shifts as a function of TFE (Figure 1) indicate that the general chemical shift movements of most of the residues of IA_3 appear to be consistent with the formation of an α-helical structure as the concentration of TFE in solution is increased. Most of the paths of the individual residue $^1H^N$ and ^{15}N cross peak titration tracks contain this positive slope movement in a $\Delta\delta^1H^N$ and $\Delta\delta^{15}N$ plot. The residue tracks of Lys 18, Gly 21, Lys 24, Val 26, Ser 27, Lys 32, Ser 35, Lys 41, Gly 62, and Gly 64 exemplify this behavior.

However, this type of analysis does not cover the large variation in chemical shift tracks when we consider all of the tracked residues. Many of the residue chemical shift tracks maintain some degree of a positive slope as the TFE concentration is increased. This indicates that there is a large degree of helical tendency (*4*, *41*), yet very few of these tracks have a strictly positive linear slope. In Figure 1, the residue tracks of Asp 4, Glu 10, Ser 15, Lys 16, Ala 29, and Ala 63 have slopes which are nearly horizontal, showing that the chemical shift movement occurred in the $^1H^N$ dimension only. Val 8, Ser 9, Gln 13, and Met 33 have residue tracks which bend toward the vertical axis at later points in the titration. More extreme cases of this bending can be seen in the tracks of Ala 34, Asp 45, Ser 48, and Asn 58 which bend past the vertical and turn downfield on the $^1H^N$ axis. The track of Glu 17 is nearly vertical. Also the residue tracks of Thr 3, Ser 14, Ala 23, Asp 37, Asp 39, Thr 42, Thr 43, Asp 44, and Asp 46 show positive linear slopes yet start from positions father away from the $^1H^N$ and ^{15}N origin, thereby confusing the standard coil-to-helix argument.

A few residue tracks are not consistent with the conventional coil-to-helix transition. Residues Asp 22, Asp 28, Phe 30, Ala 34, Ala 45, Ser 48, Glu 49, Tyr 53, and Asn 58 move in a manner completely contrary to what is expected for an ideal coil-to-helix transition. Asp 22, Asp 28, Ala 34, and Asn 58 make drastic hook-like turns as the TFE concentration increases. Ser 48, Glu 49, and Tyr 53 do not have long tracks; however, they move downfield in $^1H^N$ and upfield in ^{15}N. Finally, the residue track of Phe 30 is linear but moves downfield as the TFE concentration increases, and is the only residue track to have this kind of movement.

Qualitatively, many of the chemical shifts move in a manner consistent with a coil-to-helix transition. Following the chemical shift tracks of each individual residue, however, it is unclear whether these shifts constitute a uniform or global, two-state transition. This led us to use SVD on the chemical shift tracking data.

Unraveling the Transition. In our previous study, SVD analysis of the far-UV circular dichroism data yielded basis states consistent with a high TFE folded state and a low TFE unfolded state. In the presence of TFE, IA_3 transitions from an initial disordered state to a helical conformation through a two-state transition, with the midpoint of this transition observed at 16.3% TFE (*1*).

From the individual residue chemical shift titration curves in Figure 1, it is difficult to see the two-state transition. SVD allowed us to focus on the important cooperative transitions inherent in the chemical shift data without being misled by the differences between the chemical shift tracks of individual residues. This approach is fundamentally different from chemical shift mapping in that it allows us to identify and quantify the transition that IA_3 is undergoing in increasing amounts of TFE. Additionally, the use of SVD allowed us to compare and analyze the TFE and residue dependence inherent to the residue chemical shift movements separately.

SVD analysis of the 2D ^{15}N-HSQC data indicated that the first two components of the TFE dependence of both the $^1H^N$ and ^{15}N chemical shifts could be fit to a two-state

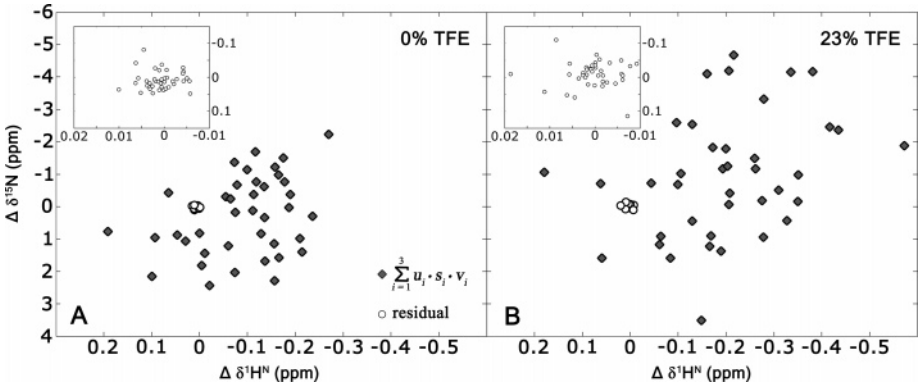

FIGURE 3: Comparison of the first three components of the SVD to the higher order components indicating that the fourth and higher order components are noise. Residual (○) is the difference between the complete dataset and the reconstruction of the first three SVD components (◆). Panel A is at 0% TFE, while panel B is at 22% TFE. The insets give an expanded view of the residual.

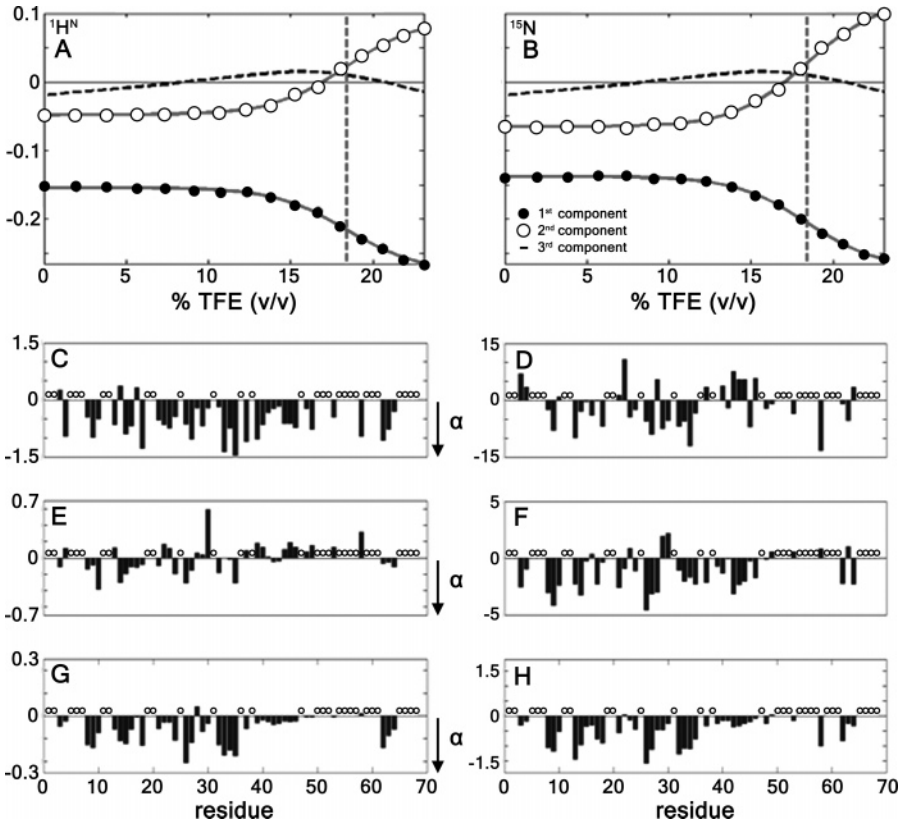

FIGURE 4: The first three components of the TFE dependence (u_1, u_2, and u_3 from eq 3) of the chemical shifts for $^1H^N$ and ^{15}N resonances are represented in panels A and B, respectively. The first and second components of the TFE dependence are shown as open (○) and filled circles (●), respectively, while the third component is shown as a dashed curve. The solid lines show a two-state model, simultaneously fit to the first two TFE components of both the $^1H^N$ and ^{15}N together. The dashed vertical line indicates the TFE refolding midpoint which was found to be 18.3%, in agreement with the TFE refolding midpoint derived from CD (*1*). Panels C−H illustrate the residue dependence of the chemical shift (the v components from eq 3). Panels C and D show the first component of the residue dependence shifts for $^1H^N$ and ^{15}N resonances associated with the first TFE-dependent curve in panels A and B; panels E and F show the second component, and panels G and H show the third component. Open circles in panels C−H indicate residues not tracked. The arrow indicates α-helical tendency as given by Wishart and Sykes (*4*).

Table 1. Computed ΔG_0, *m*-values, and $TFE_{midpoint}$ from Separate $^1H^N$, ^{15}N, Datasets and the Combined $^1H^N$ and ^{15}N Dataset

dataset	ΔG_0 (kJ/mol)	*m*-value (kJ (mol % TFE)$^{-1}$)	$TFE_{midpoint}$ (% TFE)
$^1H^N$	-21.2 ± 7.3	1.15 ± 0.21	18.4
^{15}N	-16.6 ± 6.6	0.92 ± 0.30	18.1
$^1H^N$ and ^{15}N combined	-20.4 ± 4.1	1.11 ± 0.12	18.3

thermodynamic transition, as in the earlier CD study (*1*). We simultaneously fit these four datasets to yield a TFE folding midpoint of 18.3%. This value is in close agreement with the value obtained from earlier CD studies. In Green et al. (*1*) the far-UV CD studies of IA_3 in increasing concentrations of TFE exhibited very broad transitions which have been known to influence ΔG_0 and *m*-values (*42*). This, coupled with differences in sample preparation, can explain the small difference in the TFE folding midpoint found between these two techniques.

From Figure 2, we can see that the bulk of the TFE dependence can be represented by the first two TFE components from the SVD analysis. This indicates that, to

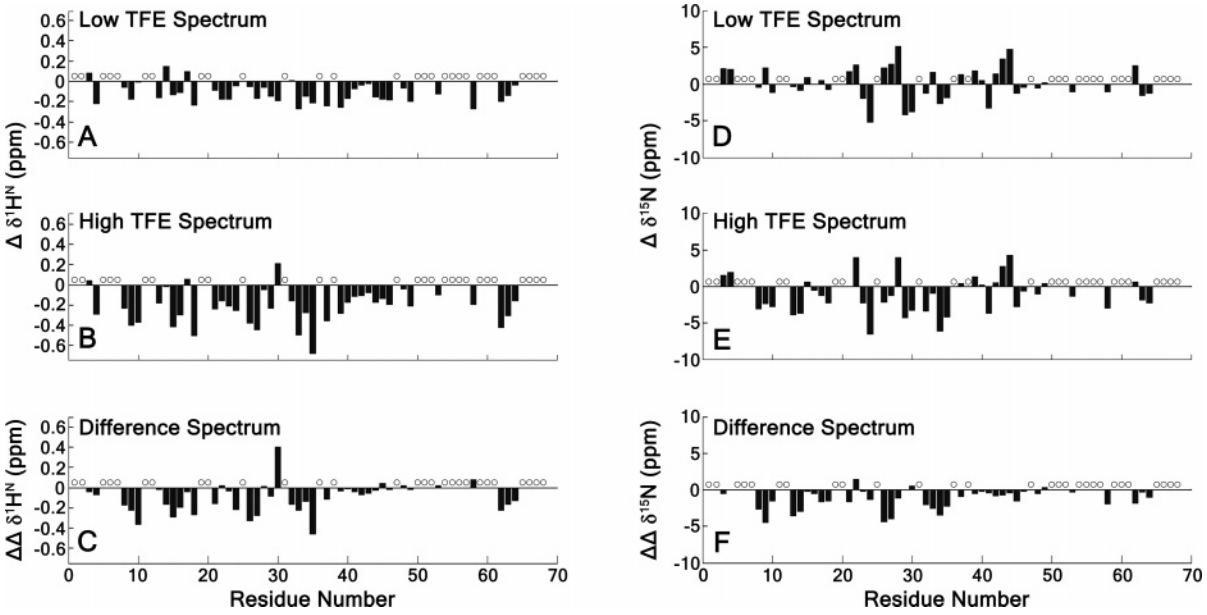

FIGURE 5: Computed high and low TFE spectra. Panels A and B indicate the computed residue dependence in the limit of low and high TFE concentrations for $^{1}H^{N}$; panel C shows the difference between the high and low TFE spectra. Panels D−F are the ^{15}N analogues to panels A−C. Residues that were not tracked are indicated by open circles.

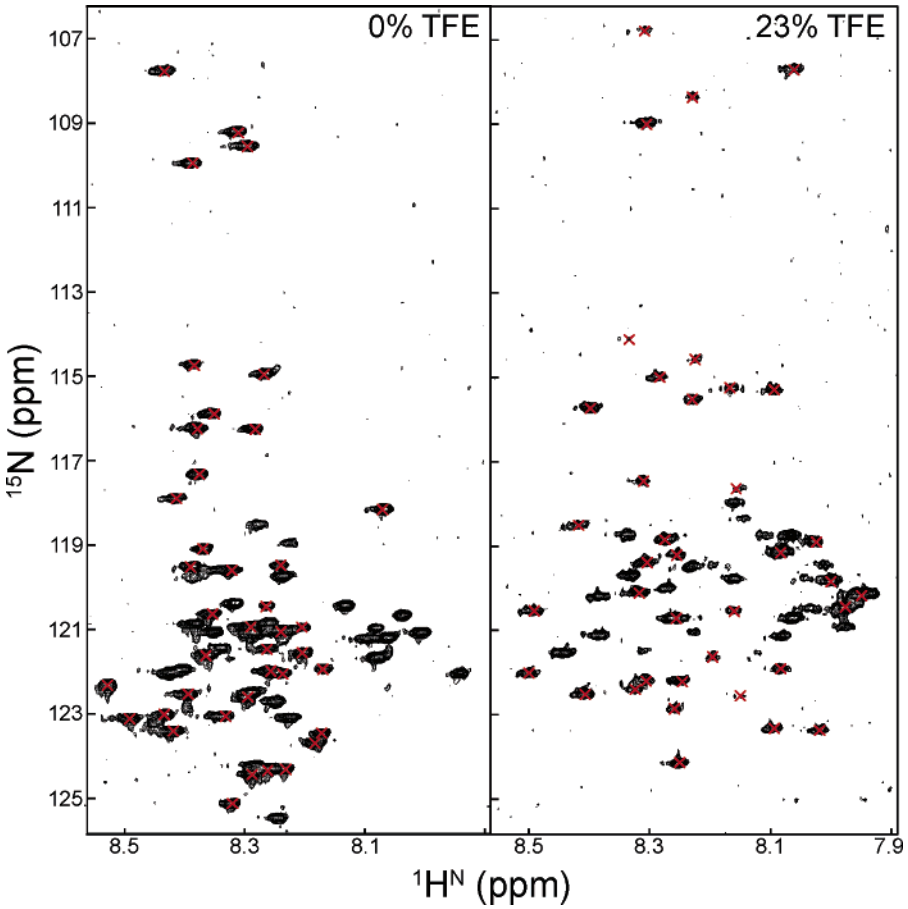

FIGURE 6: Experimental and reconstructed (red *x*) ^{15}N-HSQC spectra. Panel A shows the low TFE spectrum, and Panel B shows the high TFE spectrum.

a large extent, most residues can be described as participating in a global, two-state transition.

Different Behavior in the Two Halves of IA3. Our previous study of IA_3 had shown that at 0% TFE the bulk of the residues in IA_3 had a small tendency toward α-helical structure in ^{15}N, $^{13}C'$, $^{13}C^{\alpha}$, $^{1}H^{\alpha}$, and $^{1}H^{N}$ chemical shifts (*1*). From panels C and D in Figure 4, we see that the first component of the $^{1}H^{N}$ residue dependence of the chemical shift from the SVD analysis does indeed agree with our previous assessment (*1*). The first component of the ^{15}N residue dependence on the chemical shift qualitatively agrees as well; however, the SVD analysis indicates that several

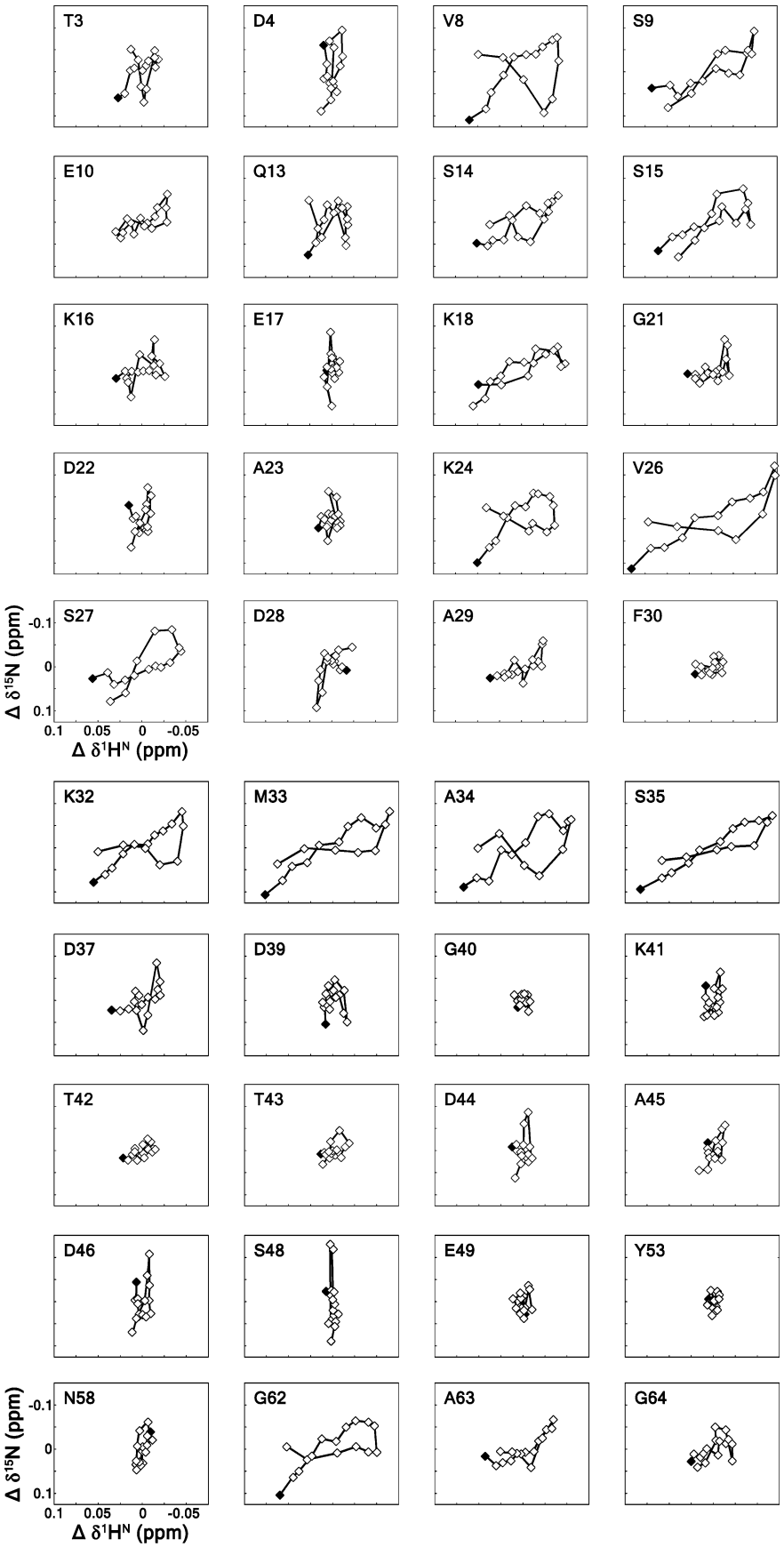

FIGURE 7: Analysis of the third SVD component. Each panel shows the reconstruction of the third component and residuals plotted for the residues tracked in this study. Each residue track illustrates the chemical shift movement of the third component from the SVD analysis for that residue as a function of TFE. Residues that do not have an appreciable third component appear to be randomly distributed around the origin with no discernible trend (e.g., Phe 30, Gly 40, and Tyr 53). Residues with large third components show an outward looping track with beginning and ending points close to each other, and are very extended along the $^{1}H^{N}$ axis (e.g., Ser 9, Val 26, and Met 33). The beginning of the TFE titration (0%) is indicated by a filled diamond.

residues around the middle of the sequence may have a propensity toward a different structure.

The data presented by Green et al. indicated that, at 0% TFE, there was a moderately populated α-helix in the C-terminus of the protein (*1*). The amplitudes of second components of both the $^{1}H^{N}$ and ^{15}N residue dependence over the whole TFE range separate the N-terminus from the C-terminus of the sequence (panels E and F in Figure 4). In both panels, the N-terminus of the sequence shows larger negative bars than the C-terminus, indicating that the N-terminus shows a larger propensity toward a helical structure, with residues around the middle of the sequence exhibiting a positive trend. As the percentage of TFE is increased, the ^{15}N-HSQC chemical shift cross peaks of these residues move downfield along the $^{1}H^{N}$ and ^{15}N frequency axes.

However, the difference spectrum of the limiting high and low TFE states (Figure 5) shows larger negative bars in the N-terminus of the protein. The first implication of this is that the N-terminal residues are undergoing a more pronounced TFE induced transition than residues in the C-terminus. Second, since the bars are predominantly negative in the N-terminus, this suggests that the N-terminus shows a larger overall helical change than the C-terminus as the TFE concentration is increased. Comparisons of the frequency shifts of these residues with limiting values from Wishart and Sykes (*4*) indicate that this region is largely α-helical.

The crystal structure of the YPrA·IA_3 complex shows only residues 3−31. Our analysis shows helical propensity throughout the sequence with the region around the middle acting as a separator between the two halves of the protein. Some difference in structure appears to exist in this region separating the N and C termini; this is exactly the boundary found in the crystal structure.

Phylip et al. had previously shown that the inhibitory action of the protein could be induced with only the first 34 residues of the N-terminus, with a K_i of less than 0.1 nM at pH 4.5 (*3*). From the crystal structure of YPrA·IA_3, it is understood that formation of an α-helix in the N-terminus of the protein is necessary for inhibition of YPrA. The folded state of IA_3 induced by TFE may not be identical with the structure found in the biologically active complex. However, this analysis shows that the TFE-induced two-state transition from a coil to a helical structure occurs in the same residues seen in the YPrA·IA_3 crystal structure and is consistent with the formation of an α-helix in the N-terminus of IA_3. The predisposition of the N-terminus of IA_3 to form an α-helix, which is indicated by our analysis, may be a direct consequence of its sequence and is necessary for its function (*43*). The results of this study indicate that TFE is driving the same kind of structural transition that occurs during the inhibition of the protease.

Additional Folding Phenomena? While the main transition seen in the chemical shifts of the residues of IA_3 appears to be two-state, a third component did become apparent from the SVD analysis. This component represented less than 5% of the observable transition in the data yet stood out from the analysis, since the first three singular values were much larger than the subsequent singular values (Figure 2).

The third SVD component of the TFE dependence of the NMR chemical shift showed a weak positive peak, as if an additional state becomes populated up to 15% TFE, then depopulates as the concentration of TFE increases further (Figure 4). This component contains all deviations from perfect two-state folding, since the rotation procedure (see Experimental Procedures) optimized the first and second components for their fit to a two-state transition. Any trend in the TFE-dependence of the data that does not match up with two-state behavior will appear in this third component (or else in a higher-order component). Several factors could contribute to this third component including differential line broadening (*44*), direct interactions with carboxylate moieties on acidic side chains (*45*, *46*), and any other deviation from a linear two-state transition.

It is well-known that 2,2,2-trifluoroethanol (TFE) stabilizes secondary structures in peptides (*47*−*50*). While TFE has been shown to stabilize many different types of secondary structures, several authors have demonstrated that there exists a good correlation between helical propensity and primary sequence (*51*−*53*). Although the exact mechanism of action of TFE in the inducement of helical structure is not fully understood, several theories of how TFE stabilizes helical structures have been proposed and continue to be studied. TFE may induce secondary structure simply by offering a more favorable environment for secondary structure formation, possibly by reducing the hydrophobic effect or decreasing the affinity of an unstructured polypeptide to hydrogen bonding with the surrounding solvent (*42*). However, TFE is also known to be a better hydrogen bond donor and a poorer acceptor than water (*42*, *54*). These effects may assist or detract from folding by stabilizing transient structures.

One possibility to explain the third TFE-dependent component of the SVD analysis could be regarded as a low-populated intermediate prior to the two-state folding in which the carboxylate moieties shift their transient hydrogen bonds from the amide protons along the backbone to TFE. This frees the amide protons, allowing them to proceed with $i + 4$ hydrogen-bonding found in the α-helix. The weak peak at 15% in both the $^{1}H^{N}$ and ^{15}N could represent the limit at which the amide proton is most free of the transient hydrogen bonds, just before it participates in cooperative hydrogen bonding of the α-helix. This could be an interesting area for future research on this system.

CONCLUSION

We have presented a method to identify folding behaviors using residue level information on IA_3. In this study, we have shown that the use of 2D ^{15}N-HSQC spectroscopy combined with an SVD analysis can be used to identify folding behaviors and is an excellent method for following individual residue movements as a function of perturbant.

The addition of TFE enhances the inherent propensity of IA_3 to form a helical structure. From the above analysis, we determine that the N-terminus of IA_3 has a greater transition toward an α-helical structure than the C-terminus of the protein, as the TFE concentration is increased. Using SVD of the ^{15}N-HSQC NMR data, we have shown that a two-state coil-to-helix transition occurs in the same residues seen in the α-helical structure of IA_3 found in the crystal structure of the YPrA·IA_3 complex.

The folding of IA_3 derived from this study indicates several features that were determined by using SVD analysis of the

NMR data. We also show that to a large degree the transition of each individual residue in the protein is two-state-like. However, the presence of a weak third component in the analysis indicates that a two-state description of IA_3 folding is not entirely complete.

ACKNOWLEDGMENT

The authors thank Ben Dunn for advice and suggestions regarding the IA_3/YPrA system, Jim Rocca and Aaron Dossey for valuable conversations, as well as Cherian Zachariah, Seth McNeill, Heather Cornnell, and Katia Ganesh for their critical reading of this paper.

SUPPORTING INFORMATION AVAILABLE

Figure depicting the 40 tracked chemical shifts of the IA_3 sequence, and triple resonance data used to identify and assign individual peaks. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Green, T. B., Ganesh, O., Perry, K., Smith, L., Phylip, L. H., Logan, T. M., Hagen, S. J., Dunn, B. M., and Edison, A. S. (2004) IA3, an aspartic proteinase inhibitor from *Saccharomyces cerevisiae*, is intrinsically unstructured in solution, *Biochemistry 43*, 4071−4081.
2. Li, M., Phylip, L. H., Lees, W. E., Winther, J. R., Dunn, B. M., Wlodawer, A., Kay, J., and Gustchina, A. (2000) The aspartic proteinase from *Saccharomyces cerevisiae* folds its own inhibitor into a helix, *Nat. Struct. Biol. 7*, 113−117.
3. Phylip, L. H., Lees, W. E., Brownsey, B. G., Bur, D., Dunn, B. M., Winther, J. R., Gustchina, A., Li, M., Copeland, T., Wlodawer, A., and Kay, J. (2001) The potency and specificity of the interaction between the IA_3 inhibitor and its target aspartic proteinase from *Saccharomyces cerevisiae*, *J. Biol. Chem. 276*, 2023−2030.
4. Wishart, D. S., and Sykes, B. D. (1994) Chemical shifts as a tool for structure determination, *Methods Enzymol. 239*, 363−392.
5. Pastore, A., and Saudek, V. (1990) The relationship between chemical shift and secondary structure in proteins, *J. Magn. Reson. 90*, 165−176.
6. Williamson, M. P. (1990) Secondary-structure dependent chemical shifts in proteins, *Biopolymers 29*, 1423−1431.
7. Matsuura, H., Shimotakahara, S., Sakuma, C., Tashiro, M., Shindo, H., Mochizuki, K., Yamagishi, A., Kojima, M., and Takahashi, K. (2004) Thermal unfolding of ribonuclease T1 studied by multidimensional NMR spectroscopy, *Biol. Chem. 385*, 1157−1164.
8. Sadqi, M., Fushman, D., and Munoz, V. (2006) Atom-by-atom analysis of global downhill protein folding, *Nature 442*, 317−321.
9. Zuiderweg, E. R. (2002) Mapping protein-protein interactions in solution by NMR spectroscopy, *Biochemistry 41*, 1−7.
10. Rajagopal, P., Waygood, E. B., Reizer, J., Saier, M. H., Jr., and Klevit, R. E. (1997) Demonstration of protein-protein interaction specificity by NMR chemical shift mapping, *Protein Sci. 6*, 2624−2627.
11. Baber, J. L., Levens, D., Libutti, D., and Tjandra, N. (2000) Chemical shift mapped DNA-binding sites and 15N relaxation analysis of the C-terminal KH domain of heterogeneous nuclear ribonucleoprotein K, *Biochemistry 39*, 6022−6032.
12. Lee, A. L., Volkman, B. F., Robertson, S. A., Rudner, D. Z., Barbash, D. A., Cline, T. W., Kanaar, R., Rio, D. C., and Wemmer, D. E. (1997) Chemical shift mapping of the RNA-binding interface of the multiple-RBD protein sex-lethal, *Biochemistry 36*, 14306−14317.
13. Henry, E. M., and Hofrichter, J. (1992) Singular value decomposition: application to the analysis of experimental data, *Methods Enzymol. 210*, 129−192.
14. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects, *J. Biomol. NMR 5*, 67−81.
15. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR 2*, 661−665.
16. Bodenhausen, G., and Ruben, D. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, *Chem. Phys. Lett. 69*, 185−189.
17. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins, *J. Magn. Reson. 85*, 393−399.
18. Grzesiek, S., and Bax, A. (1992) An efficient experiment for sequential backbone assignment of medium sized isotopically enriched proteins, *J. Magn. Reson. 99*, 201−207.
19. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *J. Am. Chem. Soc. 114*, 6291−6293.
20. Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein, *J. Magn. Reson. 96*, 432−440.
21. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR-spectroscopy of isotopically enriched proteins, *J. Magn. Reson. 89*, 496−514.
22. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients, *Prog. Nucl. Magn. Reson. Spectrosc. 34*, 93−158.
23. Palmer, A. G., Cavanagh, J., Wright, P. E., and Rance, M. (1991) Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy, *J. Magn. Reson. (1969) 93*, 151−170.
24. Kay, L., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc. 114*, 10663−10665.
25. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
26. Johnson, B., and Blevins, R. NMR View: A computer program for the visualization and analysis of NMR data, (1994) *J. Biomol. NMR 4*, 603−614.
27. Friedrichs, M. A model-free algorithm for the removal of baseline artifacts, (1995) *J. Biomol. NMR 5*, 147−153.
28. Schwarzinger, S., Kroon, G. J., Foss, T. R., Wright, P. E., and Dyson, H. J. (2000) Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView, *J. Biomol. NMR* 18, 43−48.
29. Schwarzinger, S., Kroon, G. J., Foss, T. R., Chung, J., Wright, P. E., and Dyson, H. J. (2001) Sequence-dependent correction of random coil NMR chemical shifts, *J. Am. Chem. Soc. 123*, 2970−2978.
30. Golub, G. H., and Reinsch, C. (1970) Singular value decomposition and least squares solution, *Numer. Math. 14*, 403−420.
31. Golub, G. H., and Van Loan, C. F. (1996) *Matrix Computations*, 3 ed., Johns Hopkins University Press, Baltimore, MA.
32. Nash, J. C. (1990) *Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation*, 2 ed., Adam Hilger, England, New York.
33. Press, W., Flannery, B., Teukolsky, S., and Vetterling, W. (1992) *Numerical Recipes in C: The Art of Scientific Computing*, 2 ed., Cambridge University Press, New York.
34. Santoro, M. M., and Bolen, D. W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants, *Biochemistry 27*, 8063−8068.
35. Dobson, C. M., and Fersht, A. R. (1996) *Protein Folding*, Cambridge University Press, New York.
36. Goldstein, H., Poole, C., and Safko, J. (2002) *Classical Mechanics*, 3 ed., Addison Wesley, San Francisco, CA.
37. Demchenko, A. P. (2001) Concepts and misconcepts in the analysis of simple kinetics of protein folding, *Curr. Protein Pept. Sci. 2*, 73−98.
38. Asakura, T., Kazuhiro, T., Demura, M., and Williamson, M. P. (1995) The relationship between amide proton chemical shifts and secondary structure in proteins, *J. Biomol. NMR 6*, 227−236.
39. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure, *J. Mol. Biol. 222*, 311−333.

40. Wishart, D. S., and Case, D. A. (2001) Use of chemical shifts in macromolecular structure determination, *Methods Enzymol. 338*, 3−34.
41. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, *Biochemistry 31*, 1647−1651.
42. Buck, M. (1998) Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins, *Q. Rev. Biophys. 31*, 297−355.
43. Fink, A. L. (2005) Natively unfolded proteins, *Curr. Opin. Struct. Biol. 15*, 35−41.
44. Matsuo, H., Walters, K. J., Teruya, K., Tanaka, T. T. G. G., Lippard, S. J., Kyogoku, Y., and Wagner, G. (1999) Identification by NMR spectroscopy of residues at contact surfaces in large, slowly exchanging macromolecular complexes, *J. Am. Chem. Soc. 121*, 9903−9904.
45. Edison, A. S., Espinoza, E., and Zachariah, C. (1999) Conformational ensembles: the role of neuropeptide structures in receptor binding, *J. Neurosci. 19*, 6318−6326.
46. Dossey, A. T., Reale, V., Chatwin, H., Zachariah, C., Debono, M., Evans, P. D., and Edison, A. S. (2006) NMR Analysis of *Caenorhabditis elegans* FLP-18 neuropeptides: implications for NPR-1 activation, *Biochemistry 45*, 7586−7597.
47. Santiveri, C. M., Pantoja-Uceda, D., Rico, M., and Jimenez, M. A. (2005) Beta-hairpin formation in aqueous solution and in the presence of trifluoroethanol: a (1)H and (13)C nuclear magnetic resonance conformational study of designed peptides, *Biopolymers 79*, 150−162.
48. Kemmink, J., and Creighton, T. E. (1995) Effects of trifluoroethanol on the conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor, *Biochemistry 34*, 12630−12635.
49. Bolin, K. A., Pitkeathly, M., Miranker, A., Smith, L. J., and Dobson, C. M. (1996) Insight into a random coil conformation and an isolated helix: structural and dynamical characterisation of the C-helix peptide from hen lysozyme, *J. Mol. Biol. 261*, 443−453.
50. Kentsis, A., and Sosnick, T. R. (1998) Trifluoroethanol promotes helix formation by destabilizing backbone exposure: desolvation rather than native hydrogen bonding defines the kinetic pathway of dimeric coiled coil folding, *Biochemistry 37*, 14613−14622.
51. Jasanoff, A., Davis, B., and Fersht, A. R. (1994) Detection of an intermediate in the folding of the $(\beta\alpha)_8$-barrel *N*-(5′-phosphoribosyl)anthranilate isomerase from *Escherichia coli*, *Biochemistry 33*, 6350−6355.
52. Shiraki, K., Nishikawa, K., and Goto, Y. (1995) Trifluoroethanol-induced stabilization of the alpha-helical structure of beta-lactoglobulin: implication for non-hierarchical protein folding, *J. Mol. Biol. 245*, 180−194.
53. Luidens, M. K., Figge, J., Breese, K., and Vajda, S. (1996) Predicted and trifluoroethanol-induced alpha-helicity of polypeptides, *Biopolymers 39*, 367−376.
54. Iovino, M., Falconi, M., Marcellini, A., and Desideri, A. (2001) Molecular dynamics simulation of the antimicrobial salivary peptide histatin-5 in water and in trifluoroethanol: a microscopic description of the water destructuring effect, *J. Pept. Res. 58*, 45−55.

BI061358W