The interplay between transient α-helix formation and side chain rotamer distributions in disordered proteins probed by methyl chemical shifts

Magnus Kjaergaard, Vytautas Iešmantavičius, and Flemming M. Poulsen*

Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 København N, Denmark

Received 11 May 2011; Revised 11 August 2011; Accepted 13 August 2011
DOI: 10.1002/pro.726
Published online 6 September 2011 proteinscience.org

Abstract: The peptide backbones of disordered proteins are routinely characterized by NMR with respect to transient structure and dynamics. Little experimental information is, however, available about the side chain conformations and how structure in the backbone affects the side chains. Methyl chemical shifts can in principle report the conformations of aliphatic side chains in disordered proteins and in order to examine this two model systems were chosen: the acid denatured state of acyl-CoA binding protein (ACBP) and the intrinsically disordered activation domain of the activator for thyroid hormone and retinoid receptors (ACTR). We find that small differences in the methyl carbon chemical shifts due to the γ-*gauche* effect may provide information about the side chain rotamer distributions. However, the effects of neighboring residues on the methyl group chemical shifts obscure the direct observation of γ-*gauche* effect. To overcome this, we reference the chemical shifts to those in a more disordered state resulting in residue specific random coil chemical shifts. The ^{13}C secondary chemical shifts of the methyl groups of valine, leucine, and isoleucine show sequence specific effects, which allow a quantitative analysis of the ensemble of χ_2-angles of especially leucine residues in disordered proteins. The changes in the rotamer distributions upon denaturation correlate to the changes upon helix induction by the co-solvent trifluoroethanol, suggesting that the side chain conformers are directly or indirectly related to formation of transient α-helices.

Keywords: aliphatic side chains; natively unfolded proteins; intrinsically disordered proteins; conformational ensemble; denatured state; hydrophobic packing

Introduction

The conformational properties of disordered proteins are studied for two reasons: first, intrinsically disordered proteins (IDPs) are abundant and perform many important functions in nature.[1] Second, the denatured state of a folded protein may provide insight into the initial steps in protein folding.[2] Due to their heterogeneity, these two kinds of disordered proteins pose similar challenges to the traditional methods of structural biology. NMR spectroscopy is a powerful technique for characterization of transient structure in disordered proteins. By NMR spectroscopy the conformational ensembles can be probed in detail by chemical shifts,[3] paramagnetic relaxation enhancement,[4,5] and residual dipolar couplings.[6,7] Even though these methods provide a detailed picture of the peptide backbone,[7–11]

Additional Supporting Information may be found in the online version of this article

Grant sponsors: EliteForsk programme, J.C. Jacobsen memorial scholarship from the Carlsberg Foundation (V.I.); The John and Birthe Meyer Foundation; Grant sponsor: Carlsberg Foundation; Grant number: 2008-01-0368; Grant sponsor: Danish Natural Research Council; Grant number: 272-08-0500.

*Correspondence to: Flemming M. Poulsen, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 København N, Denmark. E-mail: fmpoulsen@bio.ku.dk

comparatively little experimental information is available about the side chains of disordered proteins.[12]

In disordered proteins, the NMR signals of the side chains are less dispersed than those from the backbone.[13,14] The lack of chemical shift dispersion shows that the side chains of a given residue type experience a similar time-averaged environment. The side chain conformations can be determined experimentally using coupling constants that allow determination of the populations of each of the three staggered χ_1-angles.[15] Discrepancies between experimentally determined coupling constants from disordered proteins and those predicted from coil libraries compiled from loop regions of structured proteins suggest the formation of transient hydrophobic clusters.[12,14,16] The χ_1-angle distributions of disordered proteins can also be probed using residual dipolar couplings.[12] This method shows only small differences in the side chain conformations within a side chain type and these variations are thus difficult to interpret. The interpretation of small differences in the rotamer distributions is complicated by the effects of neighboring residues, which may be larger than the effects of transient structures.

The most precisely measured NMR parameter is the chemical shift. In side chains, the ^{13}C chemical shifts depend on χ-angles of the side chains[17] at least partly due to the γ-*gauche* effect. The γ-*gauche* effect is an upfield shift of the ^{13}C resonances caused by heavy atom γ-substituents in a *gauche* conformation.[18,19] Accordingly, a ^{13}C atom that is *gauche* to a single carbon γ-substituent will have a chemical shift that is approximately 5 ppm lower than an equivalent atom in a *trans* conformation. The γ-substituents vary between the different kinds of methyl groups in proteins, and their conformational dependence thus needs to be considered separately for each kind of methyl group. The ^{13}C chemical shifts of the two methyl groups of leucine have a particularly simple conformational dependence as the two C^{δ} atoms have a single γ-substituent: the C^{α}. The contributions from the γ-*gauche* effect can be isolated by subtracting the chemical shifts of the two geminal methyl groups, and this chemical shift difference thus measures the conformation of the side chain more precisely.[17,19] Due to steric repulsion, leucine side chains primarily populate the two staggered χ_2-angles that position one methyl group *gauche* and one methyl group *trans* relative to C^{α}. As only two rotamers are significantly populated, their populations can be estimated from a linear dependence of the ^{13}C chemical shifts.[19] The chemical shift of the isoleucine C^{δ} obeys a similar conformational dependence as that of leucine and the χ_2 distribution can similarly be estimated as a linear combination of the two most populated χ_2-angles.[20] The ^{13}C chemical shifts of γ-methyl groups depend on the χ_1-angle, however, since all three rotamers are usually populated it is not possible to extract the populations from the chemical shifts alone. The methyl groups of alanine, threonine, and methionine will not be considered in the following as the $^{13}C^{\beta}$ of alanine is primarily sensitive to the secondary structure of the backbone, the $^{13}C^{\gamma}$ of threonine is perturbed by cosolvents[12] and the $^{13}C^{\varepsilon}$ of methionine is extremely flexible.

For proton chemical shifts, the conformational dependence of the methyl signals is less clear than that of the ^{13}C chemical shifts.[21] In folded proteins, the dominant source of proton chemical shift dispersion for methyl groups is the ring current effect caused by aromatic rings,[22] as the ring current effect of a nearby aromatic group can cause an upfield shift of more than 1 ppm.[23] In protein folding studies, the disappearance of upfield shifted methyl proton signals have been used to demonstrate dynamics in the hydrophobic core of partially folded states such as molten globules.[24] When the source of the ring current can be identified, the upfield shift can thus be used to probe hydrophobic interactions. Recently, a specific ring current effect on methyl groups was used to demonstrate specific hydrophobic packing in a native molten globule.[25]

In the following, the methyl chemical shifts of two disordered systems with transient structure have been analyzed: The intrinsically disordered activation domain of the activator for thyroid hormone and retinoid receptors (ACTR) and the acid denatured state of bovine acyl-CoA binding protein (ACBP). The activation domain of ACTR binds to the nuclear coactivator domain of CREB binding protein and forms three α-helices in the complex.[26] In the free state, the activation domain of ACTR is disordered,[27] but it has transient α-helices especially in the region that corresponds to the first helix formed in the complex.[28,29] ACBP is a four helix bundle[30] that has been investigated as a model system for protein folding.[31,32] At low pH, ACBP forms a disordered state that has transient, native-like α-helices that correspond to the four helices in the folded structure.[33,34] The acid denatured state of ACBP forms transient long-range native-like interactions[35,36] that can be perturbed by single-site mutations suggesting a set of specific hydrophobic and transient interactions between the helices in the denatured ensemble.[37]

Transient secondary and tertiary structure is common in disordered proteins as is amply demonstrated by numerous NMR studies.[5,35,38–40] Transient α-helices are often amphipatic, suggesting that an important factor in secondary structure formation in disordered proteins is burial of hydrophobic side chains. Similarly, long range tertiary interactions in disordered proteins can be disrupted by removal of hydrophobic side chains,[37] implying that

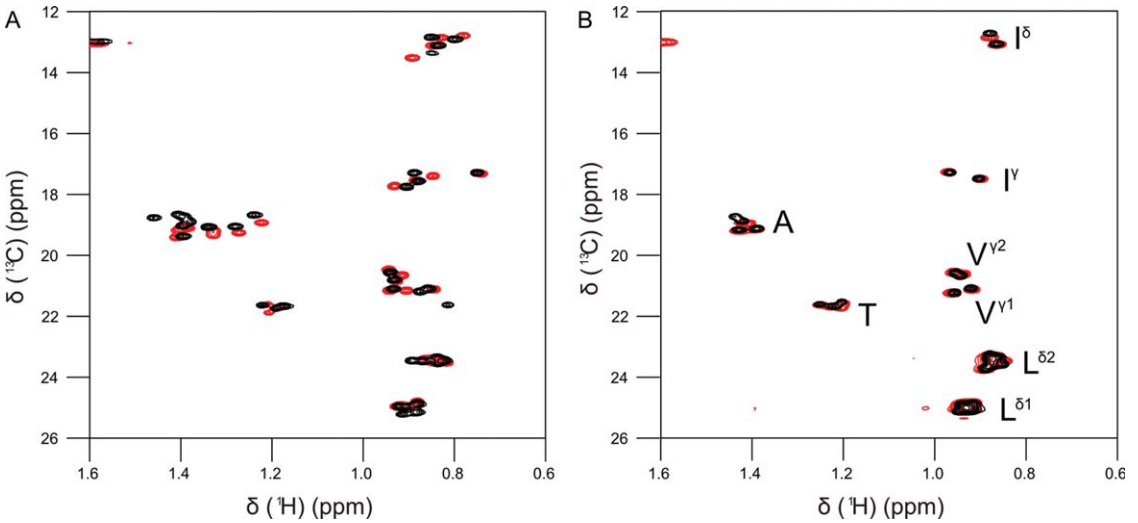

Figure 1. Constant time ^{1}H-^{13}C HSQC spectra of the methyl regions of acid denatured ACBP (A) and ACTR (B). The experiment was run without urea (black) and in the presence of 6*M* urea (red).

these interactions are in part driven by hydrophobic interactions as well. Even though hydrophobic side chains are important determinants of residual structure in disordered proteins, little is known about whether formation of secondary and tertiary interactions affect the conformations of the side chains. In the following, we explore the usefulness of methyl chemical shifts for probing the conformations of aliphatic side chains in disordered proteins. We use methyl chemical shifts to show that formation of transient helical structure leads to changes in the rotamer ensembles of hydrophobic side chains an order of magnitude smaller than the populations of the transient α-helices.

Results

Assignment of the methyl signals

The methyl signals in the ^{1}H-^{13}C HSQC are found in groups according to their amino acid types in both ACBP and ACTR (Fig. 1). As expected, the clustering shows that the hydrophobic side chains have ensemble distributions close to those of a statistical coil. The methyl group resonances can be unambiguously assigned by correlation to the backbone amide resonances of the following residue using H(CCO)NH and (H)C(CO)NH spectra[41] (Supporting Information Fig. 1). A combination of these two experiments allowed the unambiguous assignment of all methyl groups from leucine, isoleucine, and valine residues in ACBP and ACTR in the presence and absence of 6*M* urea (Supporting Information Tables 1 and 2). The chemical shifts are site-specific suggesting that the individual methyl groups experience slightly different average environments. When urea is added to the sample, the site-specific chemical shift dispersion remains approximately the same, even though the chemical shifts are perturbed for several resonances. The difference in urea sensitivity suggests that the site-specific dispersion originates from the combined effects from nearest neighbors and residual structure.

Precision of chemical shift measurements

As the chemical shift dispersion in disordered proteins is small, it is worthwhile to consider the precision of the measurements beforehand. The precision of chemical shift measurements depend on the signal-to-noise ratio and the line width[42] and can be estimated to lw/2SN.[43] For ^{13}C chemical shifts, the precision of the chemical shift measurements depend on whether the peaks are resolved in the constant time HSQC spectrum as is indicated in the Supporting Information Tables 1 and 2 for each resonance. The line widths in the constant time HSQC are on average 15 Hz (12–18 Hz), while the line widths in C(CO)NH experiment are approximately 72 Hz (64–79 Hz). With a signal-to-noise ratio of approximately 1:50, this translates into measurement uncertainties of 0.8 ppb and 3.6 ppb. For the ^{1}H chemical shifts, the line widths are similar in the H(CCO)NH experiment and the HSQC and have measurement uncertainties of approximately 0.5 ppb. In practice, however, the error associated with the measurements may be larger due to slight phase distortions or overlap, so we will not consider secondary chemical shifts below 10 ppb for ^{13}C and 1 ppb for ^{1}H to be significant. For leucines residues, this translates into an uncertainty of the χ_2-distribution of 0.2% from the chemical shift measurement alone.[19]

^{13}C methyl chemical shifts probe the individual side chain conformations in disordered proteins

In the two model proteins, the chemical shifts of the geminal methyl groups in the leucine residues

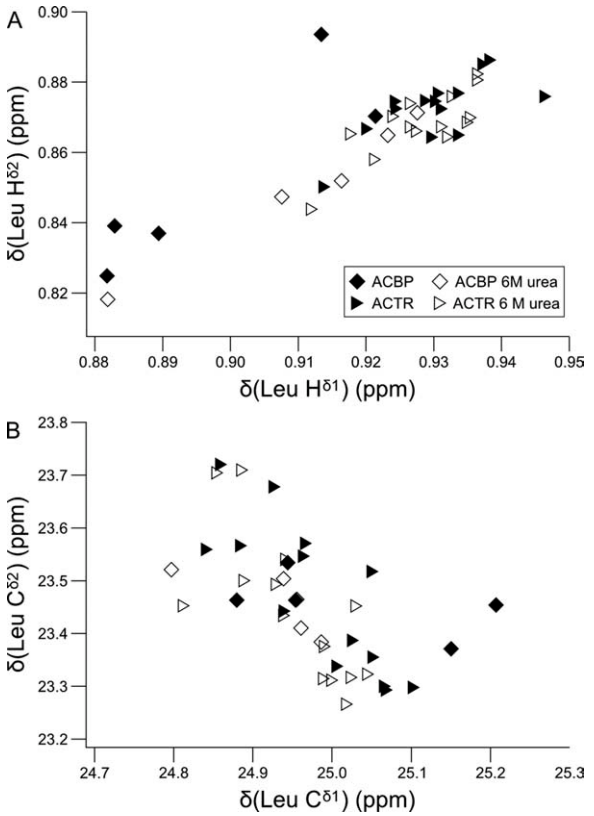

Figure 2. Correlation between chemical shifts from the geminal methyl groups of Leu side chains of ACBP and ACTR with and without urea. The $^{1}H^{\delta}$ chemical shifts are correlated ($R = 0.85$) (A) and the $^{13}C^{\delta}$ chemical shifts (B) are anticorrelated ($R = -0.62$). The anticorrelation of the ^{13}C chemical shifts is better for ACTR alone ($R = -0.74$). The anticorrelation of the ^{13}C chemical shifts suggests that variations in the χ_2-distributions affect the methyl chemical shifts.

correlate as shown in Figure 2. For the protons, this correlation suggests that the chemical environment affects the protons of the two methyl groups similarly. Correlated chemical shift changes could, for example, be caused by electrostatics or ring currents from nearby aromatic groups. In contrast, the ^{13}C chemical shifts of the geminal methyl groups are weakly anticorrelated [Fig. 2(B)] more for ACTR than for ACBP ($R = -0.74$ vs. $R = -0.54$). This difference most likely reflect the absence of aromatic residues in ACTR and hence the absence of correlated chemical shift perturbations from ring currents. The anticorrelation is not significantly perturbed by denaturation with urea (both proteins with urea: $R = -0.7$, without urea $R = -0.66$), suggesting that it is a result of factors intrinsic to the amino acid sequence rather than transient secondary or tertiary interactions. Anticorrelated changes in the ^{13}C chemical shifts is most likely due to the γ-*gauche* effect and changes in the rotamer distribution, as the two dominant χ_2-rotamers of leucine side chains have one methyl group *gauche* and one methyl group *trans* relative to C^{α}.[19] Redistribution between these two rotamers will thus lead to an increase in the γ-*gauche* effect for one methyl group and a decrease for the neighboring methyl group. The anticorrelation of the ^{13}C chemical shifts of geminal methyl groups suggests that small differences between the side chain rotamer ensembles contribute to the dispersion of methyl chemical shifts in disordered proteins. This implies that the ^{13}C methyl chemical shifts may be used to probe the individual side chain conformations in disordered proteins.

Trifluoroethanol titration of ACTR

To test whether methyl chemical shifts are affected by helical structure stabilization in the backbone, ACTR was titrated with trifluoroethanol (TFE). TFE stabilizes helices in disordered proteins through a mechanism that is not fully understood,[44] and is accordingly expected to increase the populations of the transient helices in ACTR. The helical populations were monitored using the backbone chemical shifts, C^{α} and C', that are most sensitive to the transient secondary structure. Upon addition of TFE, the ^{13}C backbone chemical shifts only increase in the regions that form transient helices in the disordered state [Fig. 3(A,B)].[28,29] TFE thus induces helicity specifically in regions with an intrinsic propensity for helix formation. Simultaneously, the side chain populations of the individual leucines were monitored using ^{13}C methyl chemical shifts, which allow determination of the population of *trans* rotamers based on the difference in ^{13}C chemical shifts between the geminal methyl groups.[19] The methyl chemical shifts suggest that the side chains ensembles change with increasing TFE concentration [Fig. 3(C)]. The chemical shift changes may be both positive and negative depending on which χ_2-rotamer is more stable in the helical conformation. The methyl chemical shift changes are largest in the regions where helical structure is induced in the backbone [Fig. 3(D)], suggesting that the side chain conformational distributions are connected to transient secondary structure in the backbone.

Secondary chemical shifts for methyl groups

Secondary chemical shifts are ubiquitously used to probe the backbone conformations of disordered proteins. Here, we examine how side chain secondary chemical shifts may be useful for defining the side chain conformations. The secondary chemical shift is calculated by subtracting a statistical coil chemical shift from the experimentally determined chemical shift. The statistical coil values are typically determined in small unstructured peptides. In this study, we have measured the methyl chemical shifts of two peptide series with the sequences Ac-GGXGG-NH_2 and Ac-QQXQQ-NH_2 (Table I), which have both

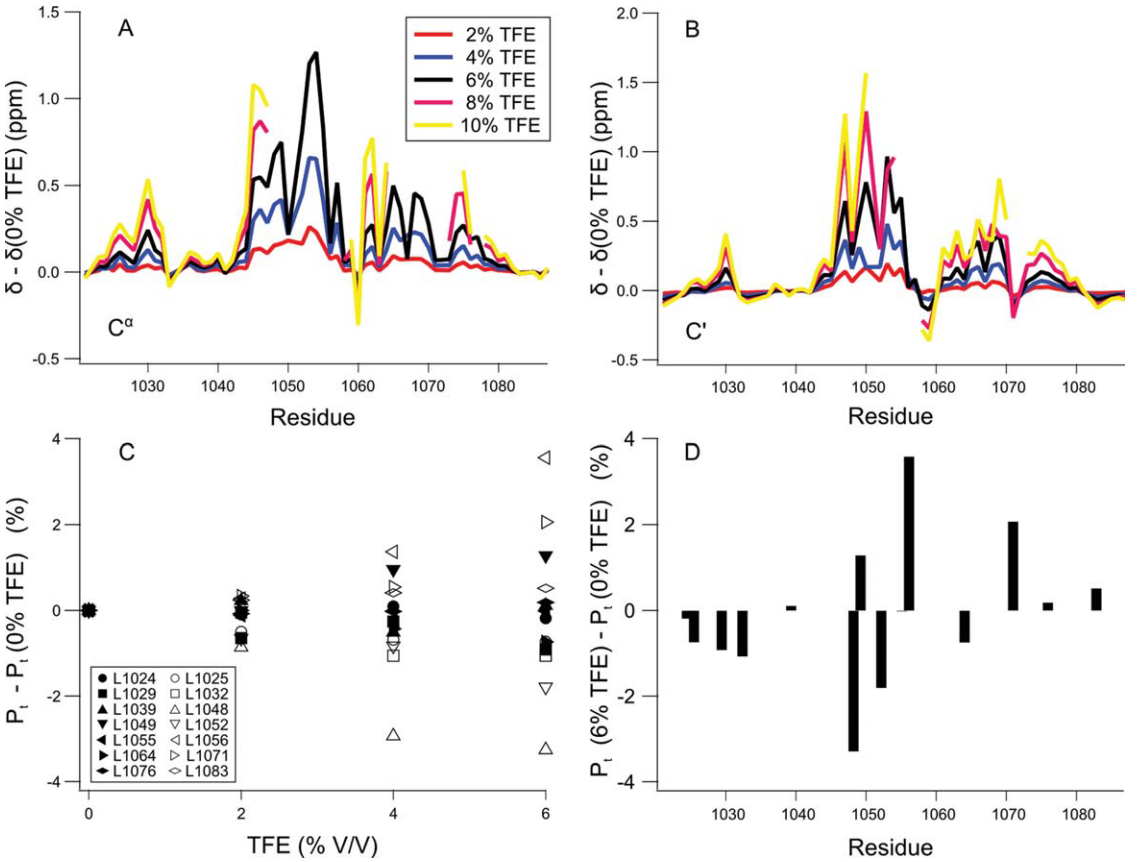

Figure 3. TFE titration of ACTR suggests interplay between backbone and side chain conformations. ACTR was titrated with the helix inducing co-solvent TFE, which results in increasing populations of the transient helices as demonstrated by C^{α} (A) and C′ (B) chemical shifts. The side chain distributions in the disordered state changes approximately linearly with increasing TFE concentration (C). The largest changes are roughly in the regions where transient helices are formed (D) suggesting that there is an interplay between the backbone and the side chain conformations. The populations of the *trans* rotamer P_t is calculated based on the difference in ^{13}C chemical shifts between the two geminal methyl groups.[19] At the highest TFE concentrations, the chemical shifts in the helical segments are missing due to exchange broadening, suggesting formation of relatively long-lived interactions stabilized by TFE.

recently been used to determine the backbone statistical coil chemical shifts.[45,46] The Ac-GGXGG-NH_2 peptide series is similar to most of the random coil peptide series in the literature and is commonly used for backbone chemical shift analysis.[47,48] It has recently been proposed that the Ac-QQXQQ-NH_2 peptides have more representative conformational distributions than the glycine containing peptides.[46] The methyl chemical shifts of these two peptide series are similar except for the ^{13}C chemical shifts of Val and Ile where differences of 0.3–0.4 ppm are seen. These differences suggest neighbor dependent changes in the conformational distributions of the side chains.

Table I. *Random Coil Chemical Shift of Methyl Groups in Two Peptide Series*

		Ac-GGXGG-NH_2		Ac-QQXQQ-NH_2	
	Atom	^1H	^{13}C	^1H	^{13}C
Ala	β	1.421	19.16	1.406	19.05
Ile	γ	0.939	17.46	0.908	17.38
	δ	0.888	13.04	0.873	12.71
Leu	δ1	0.944	24.95	0.945	24.86
	δ2	0.892	23.30	0.884	23.35
Met	ε	2.118	16.92	2.106	16.91
Thr	γ	1.226	21.52	1.222	21.66
Val	γ 1	0.963	21.06	0.938	21.07
	γ 2	0.960	20.28	0.962	20.69

The statistical coil chemical shifts in Table I were used to calculate secondary chemical shifts of the methyl groups for all leucine, isoleucine, and valine residues of ACBP and ACTR (Supporting Information Figs. 2 and 3). These secondary chemical shifts have large variations, but neither of the peptide random coil datasets result in secondary chemical shifts that correlate to regions of known structure in the backbone. This could be due to the effects of sequential neighbors that are known to have a large effect on the statistical coil chemical shifts of backbone nuclei.[48] If the neighboring residues affect the side chain chemical shifts as well,

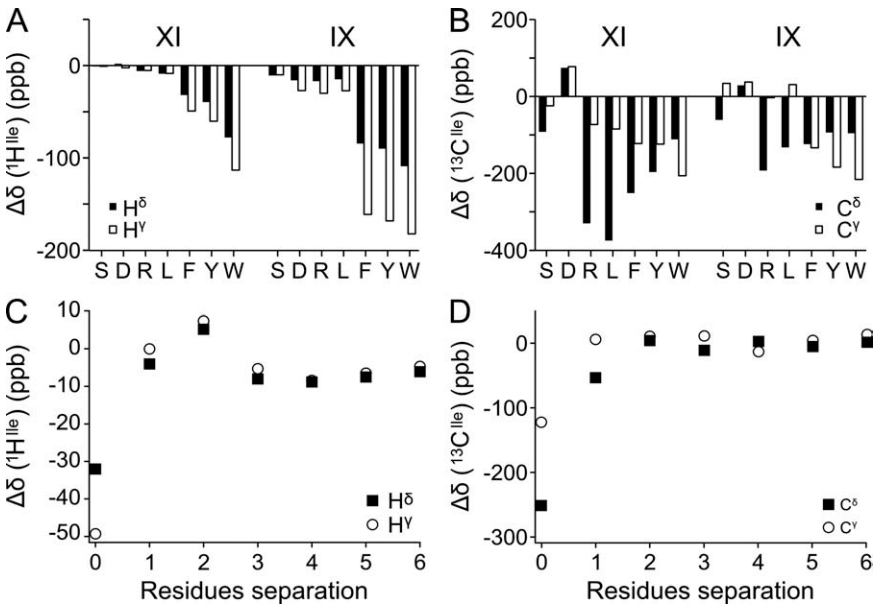

Figure 4. Methyl chemical shift perturbations from sequential neighbors. ^{1}H (A) and ^{13}C (B) methyl secondary chemical shifts of Ile were obtained for a series of Ac-GGXIGG-NH_2 and Ac-GGIXGG-NH_2 peptides. The X residue is indicated on the *X*-axis. ^{1}H (C) and ^{13}C (D) methyl secondary chemical shifts were determined in the peptide series Ac-GGFG$_n$IGG-NH_2 for separations up to 6 residues. The secondary chemical shifts were calculated using the chemical shifts from the peptide Ac-GGIGG-NH_2 in 1M urea as random coil reference.

this would obscure the detection of transient secondary or tertiary structure. To examine this effect a model system study was performed.

The effects of neighboring residues on the statistical coil chemical shifts

To investigate the effect of neighboring residues on the chemical shifts, we examined a series of peptides containing an isoleucine preceded or succeeded by either a small hydrophilic residue (S), a negatively charged residue (D), a positively charged residue (R), an aliphatic residues (L), or one of the three aromatic residues (F, Y, W). Isoleucine was chosen as the invariant residue as it contains both a γ- and a δ-methyl group. For proton chemical shifts, the neighbor effect is small for all non-aromatic residues [Fig. 4(A)]. The aromatic residues, however, have a significant impact on the proton chemical shifts of the methyl groups in the neighboring residues. This is in accordance with the neighbor effects observed for backbone nuclei[48] and is caused by a ring current from the aromatic side chain. The methyl group neighbor effects of the aromatic side chains are of the same magnitude as the upfield shifts seen for the methyl groups in acid denatured ACBP (Fig. 1). For the ^{13}C chemical shift, the neighbor effect is more complicated. The individual neighbor effects are of approximately the same magnitude as the chemical shift dispersion observed in the two disordered proteins (Fig. 1). If this neighbor effect is not taken into account, it will dominate the side chain secondary chemical shifts.

For protons, the neighbor effect was generally larger for the γ-methyl than the δ-methyl, whereas for ^{13}C the effect was largest for the δ-methyl group. The two types of nuclei report on different aspects of the local structure as the protons primarily report on the magnitude of the ring current, where the carbons report on the χ_1/χ_2 distribution. The γ-methyl is only affected by χ_1 averaging and is thus less dynamic than the δ-methyl, which is affected by both χ_1 and χ_2 averaging. The rotational freedom of the side chain is consistent with the buildup of a larger ring current effect on the less flexible γ-methyl, and a larger perturbation of the side chain rotamer distribution for the more flexible δ-methyl. To compare the effect of the preceding and the following residue, the isoleucine was placed both before and after the variable residue. For ^{1}H resonances, the upfield shift was larger when the aromatics were placed after the isoleucine, suggesting that an aromatic side chain exerts a larger ring current effect on the preceding residue. For ^{13}C, the neighbor effect varies considerably dependent on the residue type. There is not, however, an immediately apparent connection between the structure of the neighbor and the size of the neighbor effect. Determination of neighbor correction factors like those reported for the backbone nuclei[48] would require a large peptide library and is not practical, which suggests that another method has to be adopted to correct for the neighbor effects.

To probe whether more distant residues affect the methyl chemical shifts, we designed a peptide

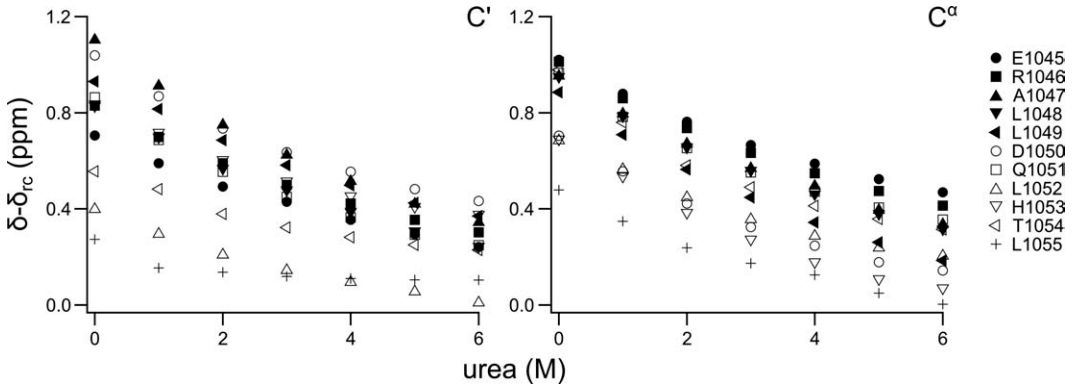

Figure 5. Urea titration of ACTR. The backbone chemical shifts of ACTR are observed with increasing urea concentration for the first transiently formed helix (residues 1045–1055). The helical population decreases rapidly at low concentrations of urea, but levels of at higher concentration. The secondary chemical shifts do not reach zero, which suggests a small helical population in the urea denatured state consistent with other studies of denatured proteins. Random coil chemical shifts are based on work from Ref. 46.

series containing an isoleucine and a phenylalanine separated by a flexible linker consisting of up to six glycine residues. The ring current from the aromatic group allows detection of even a slight interaction. Figure 4 reveals that the perturbation of the chemical shifts of the methyl groups is small when the two hydrophobic residues are separated by one or more glycines. Unless medium or long range interactions bring residues into close contact, effects beyond the nearest neighbor can thus safely be ignored. A small, but significant, distance-dependent change of secondary chemical shift can be seen for protons even at larger separations. The secondary chemical shifts have a small, positive maximum at two intervening residues, suggesting that at this separation the isoleucine methyl groups are more likely to be in the plane of the aromatic ring than above or below the ring. At longer separations, the secondary chemical shift decreases as the length of the intervening linker is increased as expected for an unspecific hydrophobic interaction. The small magnitude, however, suggests that the two hydrophobic residues only interact in a small fraction of the ensemble, implying that they are unable to bury sufficient hydrophobic surface area to compensate for the loss of conformational entropy associated with the formation of a medium range hydrophobic interaction.

Urea titration of ACTR

In backbone secondary chemical shift analyses, the neighbor effects can be removed by using a highly disordered state of the same protein as a reference, which gives clean secondary chemical shifts.[29,34] Many studies of chemically disordered proteins have demonstrated that even under highly denaturing conditions transient structure can still be observed.[6,38,49–51] Therefore, it is necessary to study the urea dependence of the secondary structure propensity. A urea titration series was recorded for ACTR from 0 to 6*M* urea, where the transient helicity was monitored by C′ and C^{α} chemical shifts. Figure 5 shows the secondary chemical shifts of residues from the helix 1 as a function of urea concentration. The secondary chemical shifts decrease most rapidly at low urea concentrations, but do not reach zero even at 6*M* urea. This is consistent with the numerous studies demonstrating transient helicity in chemically denatured states. At high urea concentrations, the chemical shift change levels off but is not constant similar to what have been observed for ACBP previously.[34] To be consistent with previous studies, we will use a urea concentration of 6*M* in the following.[29,34] The protein is not fully disordered in this state, but since intrinsic random coil referencing compares relative structural propensities it only requires that the protein is significantly more disordered than the native state.

Secondary chemical shift analysis by intrinsic statistical coil referencing

We used the urea unfolded state as the intrinsic statistical coil chemical shifts for the methyl chemical shifts of ACBP and ACTR (Fig. 6) to obtain the sequence specific random coil chemical shifts of the residue types of valine, leucine, and isoleucine. The magnitudes of the secondary chemical shifts are significantly smaller when referenced to the urea denatured state compared to those referenced to peptide derived random coil chemical shifts (Supporting Information Figs. 2 and 3) consistent with the removal of contributions from sequential neighbors.

For ACBP, the valine and leucine residue ^{13}C methyl group secondary chemical shifts range from −50 to 450 ppb and the separation between the shifts of the intraresidue methyl groups between −100 and up to 250 ppb. For the isoleucine γ2 and δ methyl groups, secondary chemical shifts between −200 and 100 ppb were observed. Most of these

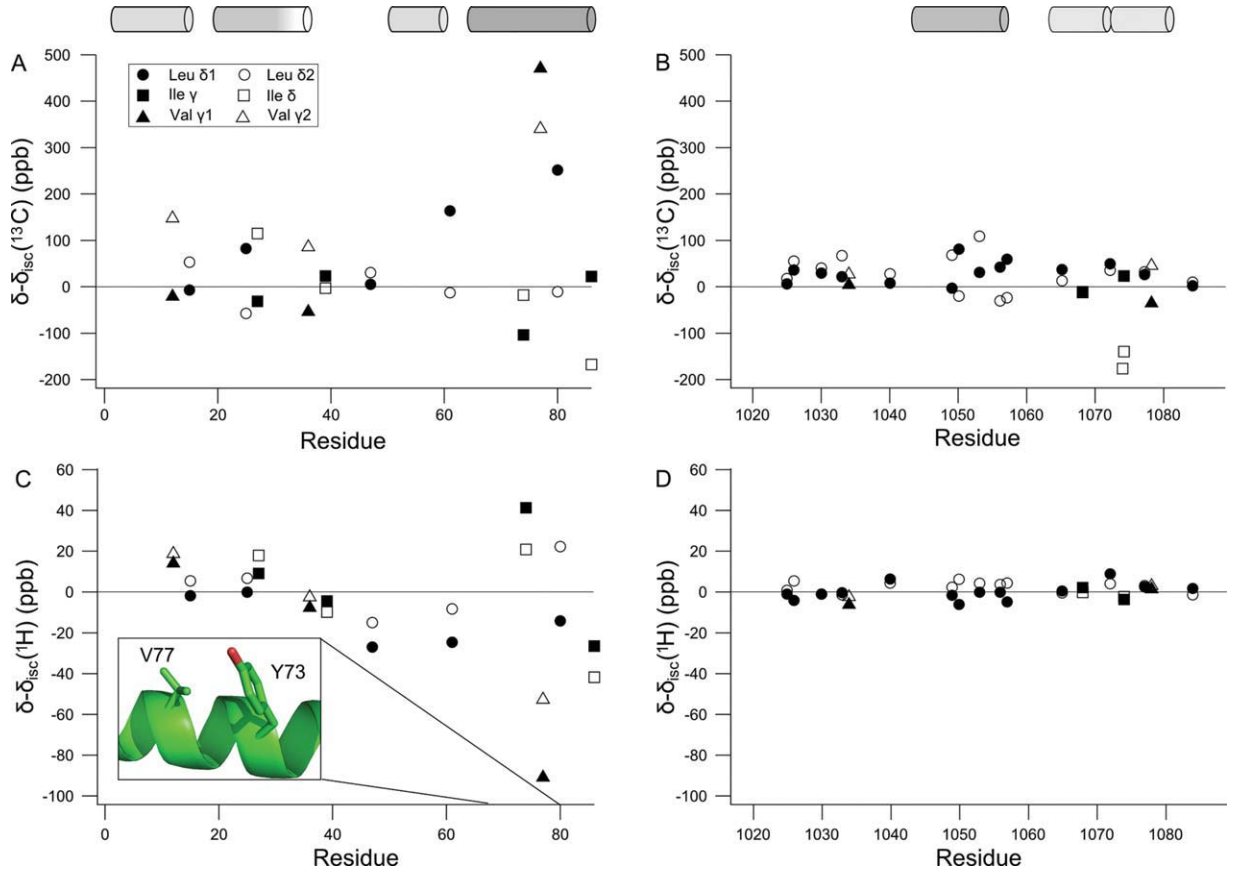

Figure 6. Methyl secondary chemical shifts were determined by intrinsic statistical coil referencing for ACBP (A,C) and ACTR (B,D). The chemical shifts determined for the highly disordered state in 6*M* urea are used as the intrinsic statistical coil chemical shifts and are subtracted from the chemical shifts of the partially folded states to give secondary chemical shifts. V77 has a particularly large ^{1}H secondary chemical shift, which suggests a ring current effect. The insert shows the conformations of V77 and Y73 the NMR structure of ACBP (PDB:1NTI). The cylinders show the locations of the α-helices in the folded state of ACBP[30] and the complex between ACTR and CBP.[26] The shading of the cylinders represents the approximate populations of the helices in the free state with fully formed helices being solid black.

residues are located in the segments of the protein that form the four helices in native ACBP. Having eliminated the contribution from neighboring residues by using intrinsic statistical chemical shift referencing, the contributions from the γ-*gauche* effect is more readily available, in particular, for leucines. These relatively large secondary chemical shifts may be due to be long- or short-range interactions favored by electrostatic or hydrophobic interactions in association with ring current shift perturbations. In the urea referenced secondary chemical shifts, a large effect is seen for all nuclei of V77 in ACBP. The magnitude of the proton secondary chemical shifts and the correlation between the two ^{13}C chemical shifts suggest that the large secondary chemical shifts are due to a ring current effect and not just due to a change in the side chain rotamer distribution. V77 does not have any aromatic neighbors, but Y73 is one turn away in a transient helix. In the folded state of ACBP, these side chains interact and a large ring current effect would be predicted for V77 [Fig. 6(C) insert]. In this case, the secondary chemical shifts report on a medium range, native-like interaction between hydrophobic side chains that forms transiently in the disordered state. For ACTR, the intrinsic statistical chemical shift referencing of the valine, leucine, and isoleucine secondary ^{13}C methyl group shifts result in a considerably smaller chemical shift range between −100 and 100 ppb, and the largest secondary shifts for the segment of residues between 1048 and 1056 coincide with the first helix in the folded form of the protein. Most of the secondary shifts from residues in other parts are very small.

Extracting rotamer population from ^{13}C methyl secondary chemical shifts

For isoleucine and leucine, the populations of χ_2-angles can be estimated from the ^{13}C chemical shifts.[19,20] Figure 7 shows the difference in the population of the *trans* rotamer between the partially folded state and the highly disordered state in urea.

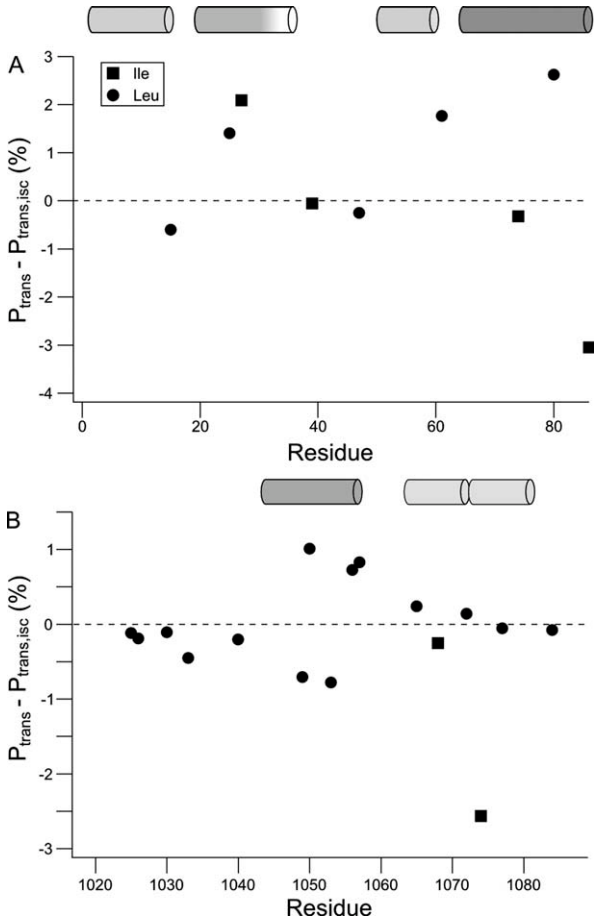

Figure 7. Perturbations of the χ_2-rotamer distributions between the partially folded states and the highly disordered state in urea for ACBP (A) and ACTR (B). The population of the trans rotamer of the χ_2-angle was estimated from the difference between $\delta(C^{\delta 1})$ and $\delta(C^{\delta 2})$ for leucine or from $\delta(C^{\delta})$ for isoleucine using the linear relationships reported previously.[19,20]

For ACTR, the population differences are close to zero. A notable exception is the region from residue 1048 to 1056 where perturbations of up to ±1% are seen. This region corresponds to the first helix in the complex between ACTR and CBP,[26] which forms transiently in the unbound state.[28,29] When the effects of sequence are removed, the ^{13}C secondary chemical shifts show that transient helix formation is accompanied by a small but significant change in the side chain rotamer distribution. Figure 6 suggests that I1073 of ACTR experience a larger change of rotamer population than the helical region. While leucine side chain populations can be estimated by the difference between two chemical shifts, the isoleucine rotamer population is based on just one chemical shift. Apart from the *γ-gauche* effect, factors affecting the chemical shifts are likely to be similar for the two geminal methyl groups as illustrated by the correlation between the proton chemical shifts of geminal methyl groups. By subtracting the two methyl chemical shifts in the leucine side chain, other contributions to the chemical shifts are removed. This is not possible for isoleucine and therefore the estimation of the χ_2-distribution is likely to be less precise for isoleucine than for leucine. It is thus not certain whether the large secondary chemical shift of I1073 represents changes in the χ_2-distribution or small chemical shift perturbations from other factors. The same consideration applies to the isoleucine side chains of ACBP. The only completely unfolded region in acid denatured ACBP is the loop between helices 2 and 3. Consistent with the low content of secondary structure, the loop region has χ_2 population perturbations that are close to zero. The residues in the helix segments have a shift in their χ_2 populations up to 3% compared to the urea unfolded form, which is significantly larger than the population shifts observed in ACTR.

If the deviations from the intrinsic statistical coil rotamer distribution observed under native conditions are due to the transient α-helices, then we would expect to see opposite changes under helix stabilizing conditions and denaturing conditions, that is, TFE and urea, respectively. In Figure 8, the change in the population of *trans* rotamers between TFE and buffer is correlated to the change between buffer and urea. The changes in the rotamer distributions are correlated with a correlation coefficient of 0.76, which suggest that it is the same kind of structural change occurring.

Discussion

We have shown that when the effect of neighbors is removed, small secondary chemical shifts can be measured for methyl groups in disordered proteins. The ^{13}C chemical shifts report primarily on the rotamer ensemble distribution, while the ^{1}H

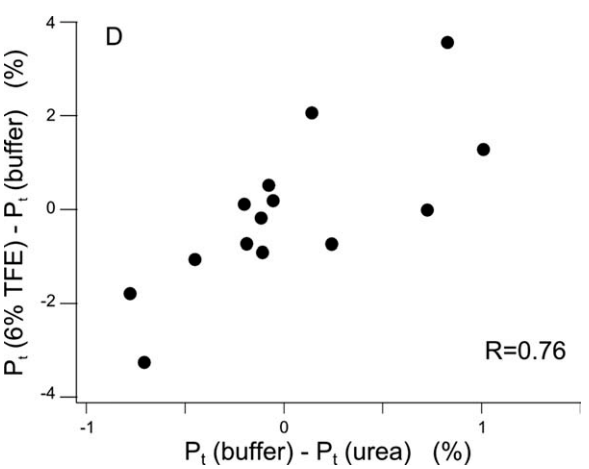

Figure 8. Correlation between changes in the rotamer distribution in solvent conditions stabilizing (TFE) and destabilizing α-helices (urea) suggests that the subtle rearrangements in the side chain rotamer distributions are mainly caused by helix formation

chemical shifts report on a mixture of other structural factors. This prompts the question whether these secondary chemical shifts probe local structural preferences or long-range interactions. In both of the two model proteins used in this study, the changes in the rotamer distribution where largest in the regions containing transient helices, and furthermore stabilization of the helices in ACTR leads to changes in the rotamer distributions. These observations suggest that formation of an α-helix leads to changes in the rotamer distribution of the side chains. These changes are, however, relatively subtle as the population changes are an order of magnitude smaller than the populations of the helices. In ACBP, the secondary chemical shifts are much larger than in ACTR, and accordingly the perturbations of the rotamer ensembles are up to threefold larger. This difference is larger than what would have been expected from the populations in the transient helices in these two proteins. The individual helix segments in acid unfolded ACBP have pairwise interactions that may contribute to the secondary chemical shifts observed[35,37] explaining the larger magnitude of the secondary chemical shifts relative to ACTR. Furthermore, ACBP has several aromatic residues that causes ring current effects. In the fourth helix segment in ACBP, the interaction between the two residues Y73 and V77 results in an up-field shift of the methyl resonances of V77. The methyl chemical shifts of this residue thus report on a medium-range side chain interaction in the transient helix that according to the random peptides in Figure 4 would not be detectable in a coil structure. This suggests that methyl secondary chemical shifts can be a useful probe of medium- to long-range structure in disordered proteins if the source of ring current effects can be traced.

The transient α-helices found in both ACTR and acid denatured ACBP are amphipathic, which means their formation is largely driven by the burial of the side chains on the hydrophobic side. Bunching all the hydrophobic side chains together on one side of a helix thus allows the side chains to interact and are thus less exposed to the solvent. Furthermore, hydrophobic clusters in disordered proteins are often involved in long-range interactions suggesting that they are also driven by burial of hydrophobic groups. Due to these hydrophobic interactions, the rotamer distributions of the side chains change relative to the random coil, but the population changes are an order smaller than populations of the structure in the backbone. This suggests that even though the side chains interact to become partially buried, their rotamer distributions are only perturbed slightly, suggesting that the interactions do not require specific hydrophobic packing. This is consistent with a model of protein folding, where the proteins collapses initially due to non-specific hydrophobic interactions, and rigid side chain packing occurs as the last step in the protein folding pathway.

Materials and Methods

Protein preparation

^{15}N and ^{13}C labeled ACBP was expressed and purified as described previously.[29,52] ACTR was coexpressed with the nuclear coactivator binding domain of CBP using a biscistronic expression vector.[26] The bacteria pellet resuspended in 10 m*M* phosphate buffer, pH 7, and heat treated at 55°C for 5 min and sonicated. Solid urea was added to the supernatant and loaded onto a HiTrap Q FF column. ACTR was eluted with a gradient from 0 to 0.5*M* NaCl and purified by RP-HPLC as described previously.[29] The NMR sample of ACTR contained 2 m*M* labeled protein, 20 m*M* phosphate buffer pH 6.5, 10% D_2O, 0.02% NaN_3, and 1 m*M* DSS. The NMR sample of ACBP contained 0.75 m*M* labeled protein, 50 m*M* phosphate buffer pH 2.4, 50 m*M* NaCl, 0.02% NaN_3, and 50 μ*M* DSS. Similar samples were prepared containing 6*M* urea and concentrations of 1 m*M* and 0.75 m*M*, respectively, for ACTR and ACBP.

NMR spectroscopy

For all four NMR samples, a constant time ^{13}C HSQC,[53] a (H)C(CO)NH,[41] and a H(CCO)NH[41] experiment were acquired at 25°C on a Varian Unity 800 MHz spectrometer equipped with either a cold probe or a room temperature probe. A FLOPSY8[54] mixing sequence was used to ensure maximal sensitivity on the methyl signals. Chemical shifts were referenced to internal DSS as described previously.[55] NMR data were processed using NMRPipe[56] and analyzed in CCPNMR Analysis 2.1.5.[57] Assignment of methyl resonances were performed using (H)C(CO)NH and H(CCO)NH experiments using previously reported backbone resonance assignments in the absence[28,33] and presence of urea.[29,34] For non-overlapped peaks, the chemical shifts were extracted from the constant time ^{13}C HSQC experiment, whereas for overlapping peaks the chemical shifts were determined from the (H)C(CO)NH and H(CCO)NH experiments. Furthermore, TFE was titrated into a sample containing 1 m*M* double labeled ACTR under similar conditions to that described above. A ^{15}N HSQC, a ^{13}C constant time HSQC of the methyl region, a (H)C(CO)NH, a 2D HNCO, and a 2D HNCOCA spectrum were acquired at 0, 2, 4, 6, 8, and 10 %(V/V) TFE. A urea titration series was acquired by recording seven datasets consisting of a ^{15}N HSQC, a 2D HNCO, and a 2D HNCOCA spectrum from 0 to 6*M* urea by addition of small quantities from a protein solution containing 10.4*M* urea.

All peptides were purchased from KJ Ross-Petersen ApS (Klampenborg, Denmark) and were

purified to more than 95% purity by reversed phase HPLC and their identities were confirmed by mass spectroscopy. The effect of the immediate neighbor was investigated using two peptide series with the sequences Ac-GGXIGG-NH_2 and Ac-GGIXGG-NH_2, where X was either of (S, D, R, L, F, Y, W). The distance dependence of the neighbor effect was investigated using a peptide series with the sequence Ac-GGFG_nIGG-NH_2, where the number of separating residues, n, was varied from 0 to 6. In these peptide series, NMR samples were prepared by dissolving 2–3 mg peptide into 500 μL 20 m*M* phosphate buffer pH 6.5 containing 5% D_2O, 3 m*M* NaN_3, and 1 m*M* DSS. pH was adjusted to the desired pH by addition of small quantities of HCl. Two series of random coil peptides with the sequence Ac-GGXGG-NH_2 and Ac-QQXQQ-NH_2 (X = A, I, L, M, T, V) were used to determine the statistical coil chemical shift values of the methyl chemical shifts. The solvent conditions were identical to previously used for determination of random coil backbone chemical shifts on the same peptides,[45,46] that is, pH 6.5 20 m*M* phosphate buffer, 5% D_2O, 3 m*M* NaN_3, and 1 m*M* DSS. The Ac-GGXGG-NH_2 peptides additionally contained 1*M* urea and was recorded at 25°C, while the Ac-QQXQQ-NH_2 peptide samples did not contain urea and were recorded at 5°C. For all peptides, the chemical shifts were determined from a ^1H-^{13}C HSQC spectrum recorded at 25°C and processed as described above.

Acknowledgments

Authors thank Kaare Teilum and Søren Brander for constructive comments to the manuscript and the Swedish NMR Centre for spectrometer time.

References

1. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Bio 6: 197–208.
2. Shortle D (1996) The denatured state (the other half of the folding equation) and its role in protein stability. FASEB J 10:27–34.
3. Jensen MR, Salmon L, Nodet G, Blackledge M (2010) Defining conformational ensembles of intrinsically disordered and partially folded proteins directly from chemical shifts. J Am Chem Soc 132:1270–1272.
4. Gillespie JR, Shortle D (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. J Mol Biol 268:158–169.
5. Lietzow MA, Jamin M, Dyson HJ, Wright PE (2002) Mapping long-range contacts in a highly unfolded protein. J Mol Biol 322:655–662.
6. Shortle D, Ackerman MS (2001). Persistence of native-like topology in a denatured protein in 8 M urea. Science 293:487–489.
7. Bernado P, Blanchard L, Timmins P, Marion D, Ruigrok RWH, Blackledge M (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle X-ray scattering. Proc Natl Acad Sci USA 102:17002–17007.
8. Jensen MR, Markwick PR, Meier S, Griesinger C, Zweckstetter M, Grzesiek S, Bernado P, Blackledge M (2009) Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. Structure 17:1169–1185.
9. Marsh JA, Forman-Kay JD (2009) Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. J Mol Biol 391: 359–374.
10. Mukrasch MD, Bibow S, Korukottu J, Jeganathan S, Biernat J, Griesinger C, Mandelkow E, Zweckstetter M (2009) Structural polymorphism of 441-residue tau at single residue resolution. PLos Biol 7:e34.
11. Huang JR, Grzesiek S (2010) Ensemble calculations of unstructured proteins constrained by RDC and PRE data: a case study of urea-denatured ubiquitin. J Am Chem Soc 132:694–705.
12. Vajpai N, Gentner M, Huang JR, Blackledge M, Grzesiek S (2010) Side-chain chi(1) conformations in urea-denatured ubiquitin and protein G from (3)J coupling constants and residual dipolar couplings. J Am Chem Soc 132:3196–3203.
13. Alexandrescu AT, Abeygunawardana C, Shortle D (1994) Structure and dynamics of a denatured 131-residue fragment of Staphylococcal nuclease—a heteronuclear NMR-study. Biochemistry 33:1063–1072.
14. Hennig M, Bermel W, Spencer A, Dobson CM, Smith LJ, Schwalbe H (1999) Side-chain conformations in an unfolded protein: chi(1) distributions in denatured hen lysozyme determined by heteronuclear C-13, N-15 NMR spectroscopy. J Mol Biol 288:705–723.
15. West NJ, Smith LJ (1998) Side-chains in native and random coil protein conformations. Analysis of NMR coupling constants and chi1 torsion angle preferences. J Mol Biol 280:867–877.
16. Mathieson SI, Penkett CJ, Smith LJ (1999) Characterisation of side-chain conformational preferences in a biologically active but unfolded protein. Pac Symp Biocomput 542–553.
17. London RE, Wingad BD, Mueller GA (2008) Dependence of amino acid side chain 13C shifts on dihedral angle: application to conformational analysis. J Am Chem Soc 130:11097–11105.
18. Tonelli AE, Schilling FC (1981) C-13 NMR chemical-shifts and the microstructure of polymers. Acc Chem Res 14:233–238.
19. Mulder FAA (2009) Leucine side-chain conformation and dynamics in proteins from C-13 NMR chemical shifts. Chembiochem 10:1477–1479.
20. Hansen DF, Neudecker P, Kay LE (2010) Determination of isoleucine side-chain conformations in ground and excited states of proteins from chemical shifts. J Am Chem Soc 132:7589–7591.
21. Sahakyan AB, Vranken WF, Cavalli A, Vendruscolo M (2011) Structure-based prediction of methyl chemical shifts in proteins. J Biomol NMR 50:331–346.
22. Perkins SJ, Wuthrich K (1979) Ring current effects in the conformation dependent NMR chemical shifts of aliphatic protons in the basic pancreatic trypsin inhibitor. Biochim Biophys Acta 576:409–423.
23. Haigh CW, Mallion RB (1979) Ring current theories in nuclear magnetic-resonance. Prog Nucl Magn Reson Spect 13:303–344.
24. Redfield C (2004) NMR studies of partially folded molten-globule states. Methods Mol Biol 278:233–254.

25. Kjaergaard M, Teilum K, Poulsen FM (2010) Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. Proc Natl Acad Sci USA 107:12535–12540.
26. Demarest SJ, Martinez-Yamout M, Chung J, Chen H, Xu W, Dyson HJ, Evans RM, Wright PE (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. Nature 415:549–553.
27. Demarest SJ, Deechongkit S, Dyson HJ, Evans RM, Wright PE (2004) Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. Protein Sci 13:203–210.
28. Ebert MO, Bae SH, Dyson HJ, Wright PE (2008) NMR relaxation study of the complex formed between CBP and the activation domain of the nuclear hormone receptor coactivator ACTR. Biochemistry 47:1299–1308.
29. Kjaergaard M, Norholm AB, Hendus-Altenburger R, Pedersen SF, Poulsen FM, Kragelund BB (2010) Temperature-dependent structural changes in intrinsically disordered proteins: formation of α-helices or loss of polyproline II? Protein Sci 19:1555–1564.
30. Andersen KV, Poulsen FM (1992) Three-dimensional structure in solution of acyl-coenzyme A binding protein from bovine liver. J Mol Biol 226:1131–1141.
31. Kragelund BB, Robinson CV, Knudsen J, Dobson CM, Poulsen FM (1995) Folding of a four-helix bundle: studies of acyl-coenzyme A binding protein. Biochemistry 34:7217–7224.
32. Kragelund BB, Osmark P, Neergaard TB, Schiodt J, Kristiansen K, Knudsen J, Poulsen FM (1999) The formation of a native-like structure containing eight conserved hydrophobic residues is rate limiting in two-state protein folding of ACBP. Nat Struct Biol 6: 594–601.
33. Thomsen JK, Kragelund BB, Teilum K, Knudsen J, Poulsen FM (2002) Transient intermediary states with high and low folding probabilities in the apparent two-state folding equilibrium of ACBP at low pH. J Mol Biol 318:805–814.
34. Modig K, Jurgensen VW, Lindorff-Larsen K, Fieber W, Bohr HG, Poulsen FM (2007) Detection of initiation sites in protein folding of the four helix bundle ACBP by chemical shift analysis. FEBS Lett 581:4965–4971.
35. Teilum K, Kragelund BB, Poulsen FM (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling. J Mol Biol 324:349–357.
36. Fieber W, Kristjansdottir S, Poulsen FM (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.
37. Bruun SW, Iesmantavicius V, Danielsson J, Poulsen FM (2010) Cooperative formation of native-like tertiary contacts in the ensemble of unfolded states of a four-helix protein. Proc Natl Acad Sci USA 107:13306–13311.
38. Neri D, Billeter M, Wider G, Wuthrich K (1992) NMR determination of residual structure in a urea-denatured protein, the 434-repressor. Science 257:1559–1563.
39. Gillespie JR, Shortle D (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. J Mol Biol 268:170–184.
40. Mohana-Borges R, Goto NK, Kroon GJA, Dyson HJ, Wright PE (2004) Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. J Mol Biol 340:1131–1142.
41. Grzesiek S, Anglister J, Bax A (1993) Correlation of backbone amide and aliphatic side-chain resonances in C-13/N-15-enriched proteins by isotropic mixing of C-13 magnetization. J Magn Reson 101:114–119.
42. Kontaxis G, Clore GM, Bax A (2000) Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. J Magn Reson 143:184–196.
43. Bax A, Kontaxis G, Tjandra N (2001) Dipolar couplings in macromolecular structure determination. Methods Enzymol 339:127–174.
44. Buck M (1998) Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. Q Rev Biophys 31:297–355.
45. Kjaergaard M, Brander S, Poulsen FM (2011) Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. J Biomol NMR 49:139–149.
46. Kjaergaard M, Poulsen FM (2011) Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. J Biomol NMR 50:157–165.
47. Schwarzinger S, Kroon GJ, Foss TR, Wright PE, Dyson HJ (2000) Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRview. J Biomol NMR 18:43–48.
48. Schwarzinger S, Kroon GJ, Foss TR, Chung J, Wright PE, Dyson HJ (2001) Sequence-dependent correction of random coil NMR chemical shifts. J Am Chem Soc 123: 2970–2978.
49. Bhavesh NS, Sinha R, Mohan PM, Hosur RV (2003) NMR elucidation of early folding hierarchy in HIV-1 protease. J Biol Chem 278:19980–19985.
50. Lindorff-Larsen K, Kristjansdottir S, Teilum K, Fieber W, Dobson CM, Poulsen FM, Vendruscolo M (2004) Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme a binding protein. J Am Chem Soc 126:3291–3299.
51. Meier S, Strohmeier M, Blackledge M, Grzesiek S (2007) Direct observation of dipolar couplings and hydrogen bonds across a beta-hairpin in 8 M urea. J Am Chem Soc 129:754–755.
52. Mandrup S, Hojrup P, Kristiansen K, Knudsen J (1991) Gene synthesis, expression in Escherichia-coli, purification and characterization of the recombinant bovine acyl-CoA-binding protein. Biochem J 276:817–823.
53. Kay LE, Keifer P, Saarinen T (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc 114:10663–10665.
54. Mohebbi A, Shaka AJ (1991) Improvements in C-13 broad-band homonuclear cross-polarization for 2d and 3d NMR. Chem Phys Lett 178:374–378.
55. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J Biomol NMR 6:135–140.
56. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) Nmrpipe—a multidimensional spectral processing system based on Unix pipes. J Biomol NMR 6:277–293.
57. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas P, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59:687–696.