JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Modulation of the Intrinsic Helix Propensity of an Intrinsically Disordered Protein Reveals Long-Range Helix−Helix Interactions

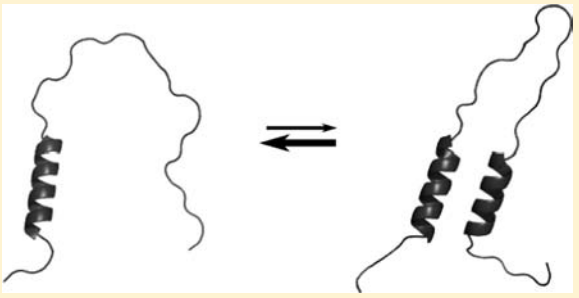
Vytautas Iešmantavičius,[†] Malene Ringkjøbing Jensen,[‡] Valéry Ozenne,[‡] Martin Blackledge,[‡] Flemming M. Poulsen,[†,∥] and Magnus Kjaergaard*[,†,§]

[†]Department of Biology, University of Copenhagen, 1017 København K, Copenhagen, Denmark

[‡]Institut de Biologie Structurale Jean-Pierre Ebel, CEA-CNRS-UJF UMR 5075, 41 rue Jules Horowitz, 38027 Grenoble, France

Supporting Information

ABSTRACT: Intrinsically disordered proteins (IDPs) are widespread and important in biology but defy the classical protein structure−function paradigm by being functional in the absence of a stable, folded conformation. Here we investigate the coupling between transient secondary and tertiary structure in the protein activator for thyroid hormone and retinoid receptors (ACTR) by rationally modulating the helical propensity of a partially formed α-helix via mutations. Eight mutations predicted to affect the population of a transient helix were produced and investigated by NMR spectroscopy. Chemical shift changes distant to the mutation site are observed in regions containing other transient helices indicating that distant helices are stabilized through long-range hydrophobic helix−helix interactions and demonstrating the coupling of transient secondary and tertiary structure. The long-range structure of ACTR is also probed using paramagnetic relaxation enhancements (PRE) and residual dipolar couplings, which reveal an additional long-range contact between the N- and C-terminal segments. Compared to residual dipolar couplings and PRE, modulation of the helical propensity by mutagenesis thus reveals a different set of long-range interactions that may be obscured by stronger interactions that dominate other NMR measurements. This approach thus offers a complementary and generally applicable strategy for probing long-range structure in disordered proteins.



INTRODUCTION

A large fraction of all genomes codes for functional proteins that lack a stable folded structure. These intrinsically disordered proteins (IDPs) play important roles in, e.g., neurodegenerative diseases and cancer,[1,2] which has promoted a surge in interest in understanding their functions. For folded proteins, the function is closely tied to 3D structure in what is usually referred to as the structure−function paradigm. IDPs challenge this paradigm by being functional in the absence of a stable fold, and accordingly a better understanding of the relationship between primary sequence and function needs to be developed for this class of proteins.[3,4]

IDPs lack the single unique fold of structured proteins, but their conformations are far from random. The conformational ensembles of IDPs can be probed at residue-specific resolution using NMR spectroscopy, and these proteins have been found to contain a wide range of structure.[5−7] Transient secondary structure elements are typically identified based on their characteristic secondary chemical shifts,[8−12] and long-range interactions can be identified by paramagnetic relaxation enhancement (PRE) and residual dipolar couplings (RDCs) and in some cases chemical shifts.[13−22] While short- and long-range interactions can routinely be identified, there are limited data on how the different levels of structure interact. Thus, it is not known how commonly observed transient α-helices affect long-range interactions or how the helices interact with each other.

The activator for thyroid hormone and retinoid receptors (ACTR) is a nuclear coactivator involved in transcriptional activation by steroid receptors.[23] ACTR activates transcription via acetylation of histones either directly by the enzymatic domain of ACTR or through recruitment of another coactivator called CREB binding protein. The first activation domain of ACTR binds to the nuclear coactivator binding domain (NCBD) of CREB binding protein, and together the proteins form an entangled complex, where both proteins contribute three α-helices.[24] Due to its small size and the structural plasticity of NCBD, this complex has become a favorite model system for both experimental and computational studies of the coupled folding and binding mechanism.[25−34] The pre-recognition state plays an important role in mechanistic studies of coupled folding and binding as one of the most popular mechanisms, conformational selection, directly predicts the existence of a fraction of a folded-like conformation in the unbound state. The residual structure of ACTR in its pre-recognition state will thus influence the binding kinetics and mechanism, and a thorough description of the disordered

Received: May 7, 2013
Published: June 13, 2013

   dx.doi.org/10.1021/ja4045532 | J. Am. Chem. Soc. 2013, 135, 10155−10163

Table 1. Data Reproduction from Ensembles with Different Numbers of Helical Conformers Based on Chemical Shift and RDC Data of Wild-Type and Mutant ACTR

no. of helical ensembles	χ^{2a}	optimized parameters[b]	helical ensembles[c]	population (%)[d]	significance[e]
		Wild-Type			
1	179	4	1045−1054	31	
2	122	7	1045−1052	19	P = 0.0207
			1045−1054	19	
3	97	10	1044−1049	12	P = 0.1608
			1045−1052	29	
			1052−1055	17	
		Triple Mutant			
1	264	4	1045−1052	70	
2	111	7	1045−1052	40	P = 0.0002
			1045−1055	30	
3	96	10	1045−1049	15	P = 0.4188
			1045−1055	31	
			1046−1052	30	

[a]Target function for the χ^2 included 32 and 29 experimental data points (^{1}H−^{15}N RDCs and Cα chemical shifts for residues 1042−1057) for the wild-type and the triple mutant, respectively. The RDCs were included in the ASTEROIDS optimization with an uncertainty of 1.0 Hz, while the chemical shifts were included with an uncertainty of 0.15 ppm. [b]One helix implies the optimization of three parameters: starting amino acid, final amino acid and the population. In addition a scaling factor is optimized to take into account the absolute level of alignment for the RDCs. [c]Range of the invoked helices. [d]Population of the invoked helices. The remaining conformers are completely unfolded. [e]Significance of the improvement of this model as compared to the simpler model calculated using a standard F-test.

ensemble may guide simulations of this, and related systems.[25−29] Of the three helices formed in complex with NCBD, only the first (helix 1) is appreciably populated in the isolated state as demonstrated by chemical shift analysis[35,36] with the populations of helices 2 and 3 remaining below 10%. The three helices are amphipathic, where the hydrophobic side is buried in the complex with NCBD.[24] In addition to the helices observed in the complex, urea denaturation revealed a small region with helical propensity in the N-terminal end (helix 0),[36] which remains flexible in the complex with NCBD.[24] Despite the presence of transient amphipathic α-helices, ACTR does not form a hydrophobic core with rigid side chain packing, as the rotamer distributions only experience small perturbations from their random coil values.[37] In general, the activation domain of ACTR is a fairly typical IDP with transient complex-like secondary structure and is thus a suitable model for studying general aspects of the coupling between secondary and tertiary interactions in disordered proteins.

This study is inspired by a recent paper where residues important for the folding of acyl-coenzyme A binding protein (ACBP) were mutated, and the effects on helices in the denatured ensemble were followed by chemical shift changes.[22] This study unambiguously demonstrated that helices can be stabilized by long-range interactions, however, the observed chemical shift changes were caused by a combination of changes in the helix propensities and removal of side chains involved in the helix−helix interactions. In this study, we avoid changing the residues involved in the helix−helix interactions, in order to isolate the effect of the helix formation. We thus use a tangential approach where we modulate the helical propensity of helix 1 of the activation domain of ACTR through site-directed mutagenesis and study the effects on the residual structure throughout the domain. We find that increases in the helical propensity lead to increases in helicity distant from the mutation site, which demonstrates the existence of long-range stabilizing interactions between transient α-helices. This indicates that secondary and tertiary structure is closely linked in disordered proteins and that rational mutagenesis coupled with NMR may be a generally applicable method for detecting interactions between helices. These long-range interactions are unlikely to be picked up by other NMR methods, principally because the populations involved are small. This suggests that modulation of the helical propensity by mutagenesis offers a complementary strategy to the existing experimental approaches.

METHODS

Protein Preparation and Mutations. The pET-22b plasmid encoding cDNA of ACTR (1018−1088) and its binding partner NCBD was a kind gift from Peter E. Wright (The Scripps Research Institute) and was described previously.[24] Mutations were introduced in the gene of ACTR using the QuikChange Lightning site-directed mutagenesis kit (Stratagene). ACTR was expressed in *E. coli* BL21 (DE3) cell lines at 37 °C in M9 minimal medium supplemented with 4 g/L ^{13}C glucose and 1.5 g/L ^{15}N ammonium sulfate as the only carbon and nitrogen sources. Protein expression was induced by addition of 1 mM IPTG to bacterial cultures at OD_{600} = 0.8, harvested after 6 h and resuspended into 20 mM piperazine buffer pH 5, sonicated, and centrifuged for 30 min at 20000 g. Urea was added to the supernatant to a final concentration of 6 M and loaded onto HiTrap Q FF column. ACTR was eluted with linear gradient from 0 to 1 M of NaCl in 6 M urea, 20 mM piperazine pH 5.7. The pH was adjusted to 2 in the fraction containing ACTR and further purified with Source 15RPC reversed phase HPLC column using a linear gradient from 30% to 60% (v/v) acetonitrile in 0.1% TFA. The purity and identity of the proteins were validated by MALDI-TOF mass spectrometry, and the concentrations of ACTR were determined using BCA assays. Spin-labeling of ACTR was achieved by overnight reaction at room

dx.doi.org/10.1021/ja4045532 | J. Am. Chem. Soc. 2013, 135, 10155−10163

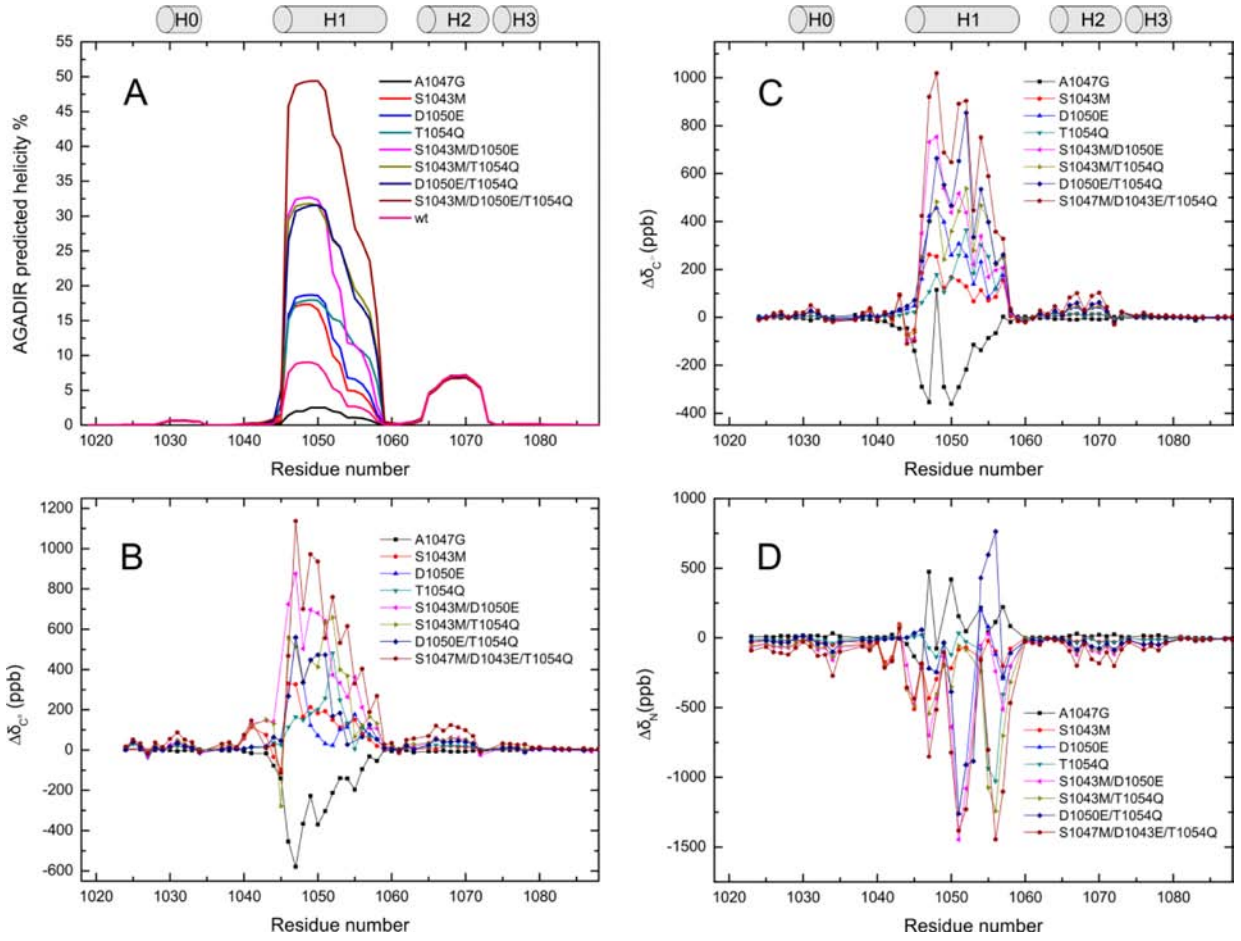


Figure 1. Modulation of the intrinsic helix propensity by site-directed mutagenesis. (A) AGADIR predictions for the helical propensities of the eight variants of ACTR used in the study. The difference between the chemical shifts of mutant and wild-type ACTR for C^{α} (B), C′ (C), and N (D). The cylinders at the top indicate the regions of transiently populated helices.

temperature in the dark with (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl) methanesulfonate (MTSL, Toronto Research Chemicals Inc., Canada) followed by removal of unlabeled protein by reversed phase HPLC. The covalent attachment of the spin labels was verified by MALDI-TOF mass spectrometry.

NMR Spectroscopy. All samples contained 10% D_2O and 0.02% NaN_3 in 20 mM phosphate buffer pH 6.5, and spectra were recorded at 37 °C. For chemical shift and PRE measurements the protein concentrations were 0.5 mM, whereas a concentration of 1 mM was used for RDC measurements. All mutants were buffer exchanged against the same buffer stock to ensure exactly the same conditions. Most experiments were acquired on a Varian Unity 800 MHz spectrometer equipped with a cold probe. The RDCs on the triple mutant were acquired on a Varian Unity 600 MHz with a room temperature probe. Standard HSQC, HNCO, and HNCOCA experiments were used for chemical shift and PRE measurements. A partially aligned sample for RDC measurements was prepared in stretched polyacrylamide gels as described previously,[38,39] and the RDCs were recorded using modified HSQC[40] experiments for spin-coupling detection in the ^{15}N dimension. Four MTSL modified mutants were used for PRE measurements: Q1020C, N1038C, N1058C, and N1078C. The PREs were measured as the ratio between peak intensities in two 2D ^{15}N HSQC spectra corresponding to the paramagnetic and diamagnetic states, respectively. Initially the paramagnetic spectrum was acquired, and subsequently the samples were reduced by addition of a 2.5-fold molar excess of ascorbate before recording the diamagnetic spectrum. Chemical shifts were referenced to internal DSS as described previously.[41] NMR data were processed using NMRPipe[42] and analyzed in CCPNMR Analysis 2.1.5.[43] To study the correlation between chemical shift changes in different helical segments, the chemical shift changes were averaged in the helical regions defined as H0: 1023−1035, H1: 1044−1055, H2: 1064−1072, and H3: 1074−1079. A non-helical region in the C-terminus (1080−1087) was included as a control.

Ensemble Description of Helix 1 in Wild-Type and Triple Mutant of ACTR Using Chemical Shifts and RDCs. Experimental RDCs and C^{α} chemical shifts were used to obtain an ensemble description of helix 1 of wild-type and triple mutant of ACTR using the previously described minimal ensemble approach.[44,45] We obtained a representative ensemble description of helix 1 (defined between residues 1044−1055) by generating ensembles of wild-type and triple mutant ACTR each consisting of 10 000 conformers using Flexible-Meccano[46,47] containing varying helix lengths and positions within the 1044−1055 region. All helices were invoked with a population of 100%. A minimum helix length of 4 residues and a maximum length of 12 residues were used giving rise to a total of 46 conformational ensembles. In addition, a random coil ensemble was created consisting of 50 000 structures. The alignment tensor of each conformer in the ensembles was calculated using PALES[48] and ensemble-averaged RDCs were obtained for each of the 47 ensembles. Ensemble-averaged chemical shifts were obtained by SPARTA[49] using 1000 randomly picked conformers. The number of helical ensembles, N, necessary to describe the experimental data was determined by incrementing N (starting at N = 1). For each step, the genetic algorithm ASTEROIDS[50] was used to select N helical ensembles and their associated populations such that the predicted population weighted RDCs and chemical shifts were in agreement with the experimental values.[45] For each step a standard F-test was used to evaluate the significance of one model over the other (Table 1).

dx.doi.org/10.1021/ja4045532 | J. Am. Chem. Soc. 2013, 135, 10155−10163

PRE Contact Maps. ASTEROIDS ensemble selections on the basis of PREs were performed as described previously[13,51] using a starting pool of 10 000 Flexible-Meccano conformers. The dynamics of the spin label for each Flexible-Meccano conformer was explicitly taken into account as described previously.[13,52] Long-range contacts within the selected ensembles were identified using the metric Δ_{ij} that compares the distance distribution of the selected ensemble with that of the reference ensemble carrying no specific long-range contacts:

$$\Delta_{ij} = \log(\langle d_{ij}\rangle / \langle d_{ij}^0\rangle) \quad \text{(Eq. 1)}$$

Here, d_{ij} is the distance between residues i and j in the ASTEROIDS ensemble, while d_{ij}^0 is the corresponding distance in the reference ensemble.

RESULTS

Mutational Strategy and Chemical Shift Assignment. Our initial goal was to modulate the intrinsic helix propensity of ACTR by site-directed mutagenesis to study the coupling between secondary and tertiary structure. Ideally, the interaction with NCBD should not be perturbed, to be consistent with future studies of the binding reaction, so all residues whose side chains are within 4 Å of NCBD in the complex[24] were ruled out as mutation sites. This in effect preserves the hydrophobic faces of the amphipathic helices and left us with nine potential mutation sites mostly in helix 1. To evaluate all possible mutants, a matrix of bulk helix propensity was generated for all possible substitutions using AGADIR (Table S1).[53−55] This approach has been successfully used to cause a maximal disruption of the helical propensity of another IDP for functional characterization.[56] To focus on the intrinsic helix propensity, we avoided modifying the charge of the molecule or introducing or removing proline residues. We chose to focus on helix 1 that is the longest and most populated helix. Four mutations were selected for experimental characterization: three mutations predicted to cause ∼2-fold increases in helicity each (S1043M, D1050E, T1054Q) and one mutation predicted to cause a ∼3-fold reduction in helicity (A1047G). The predicted helicity of the double and triple mutants and all combinations of the three stabilizing mutants suggest that the effects are additive (Figure 1A). These eight variants of ACTR are thus predicted to have a helical population in helix 1 ranging from almost 0 to ∼50%. NMR samples were prepared of each of the eight mutants and the wild-type protein. Sequence assignments were readily transferred from previous studies except for the regions surrounding the mutation sites.[35,36] Triple resonance experiments were used to confirm the assignments and to obtain the C′ and C^α chemical shifts for the mutants. The chemical shift changes for these nuclei show that the population of helix 1 changes as function of the mutations in a similar manner as predicted by AGADIR (Figure 1B−D).

Ensemble description of helix 1 in wild-type and triple mutant of ACTR from chemical shifts and RDCs. Helical regions in disordered proteins usually consist of a mixture of helical segments with slightly different termini, sometimes known as 'fraying'.[44,45] Residual dipolar couplings (RDCs) of helical segments are sensitive to the projection of the unfolded chains from the termini of the helices[57] and can therefore be used to probe the position of the termini and the populations of helical elements formed in disordered proteins.[44] Samples of wild-type and the triple mutant of ACTR were partially aligned using polyacrylamide gels,[38,39] and $^1H-^{15}N$ RDCs were collected for each residue. To determine which helical elements are populated in helix 1, the minimal ensemble approach[44,45] was used. This approach involves the selection of a minimal structural ensemble that best fits the experimental C^α chemical shifts and $^1H-^{15}N$ RDCs within helix 1 as described in the Methods section. $^1H-^{15}N$ RDCs allow the determination of the nature of the helices, whereas C^α chemical shifts are used to obtain the populations of the helical elements. Table 1 summarizes the results of this analysis for both wild-type and the triple mutant of ACTR. The analysis shows that helix 1 in both wild-type and triple mutant of ACTR can best be described by two helical elements in rapid conformational exchange with a completely disordered state. The χ^2-values from the ensemble selections demonstrate that several different helical ensembles give comparable agreement with the experimental data (Table S1), although the helical elements differ only by one residue. In the following we will only discuss the best solution for which the agreement between measured and back-calculated secondary chemical shifts, and RDCs are shown in Figure 2B,C,E,F. For both the wild-type and triple mutant, the two helical fragments start at E1045 indicating that the helical elements are stabilized by N-capping interactions via D1044.[58] For the wild-type the helices are capped at the C-terminus at L1052 and T1054 and for the triple mutant at L1052 and L1055 with roughly equal populations of the two types of helices. The change in the C-terminus in the triple mutant is most likely due to the mutation T1054Q that changes the helical capping and shifts the end of the helix by one residue. In general, it is seen that the mutations modulate the population of helix 1 from a total of 38% helix in the wild-type protein to a total of 70% in the triple mutant without changing significantly the distribution of helical elements in the pre-recognition state.

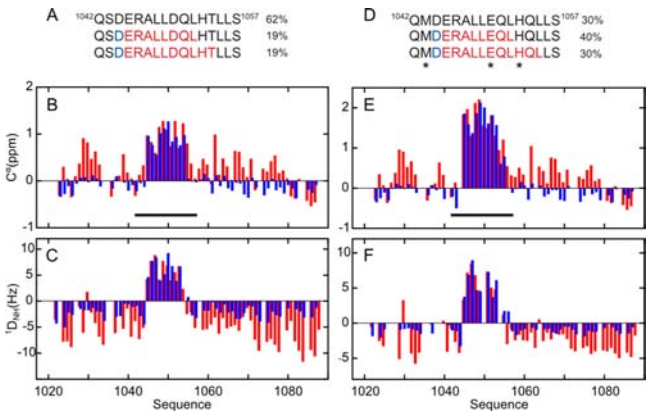


Figure 2. Helical populations in helix 1 mapped by RDCs and C^α chemical shifts. The two helical ensembles that best reproduce the data for (A) the wild-type and (D) the triple mutant of ACTR. The populations in the best fit are given as percentages and the mutation sites are marked with *. The experimental data (red) were compared to back-calculated data for C^α secondary chemical shifts (B and E) and $^1D_{NH}$ RDCs (C and F). Note that only the experimental data in the region 1044−1055 were included in the ensemble selection.

Chemical Shift Changes Distant from the Mutation Site. In addition to the chemical shift changes observed in helix 1, smaller chemical shift changes are also observed distant to the mutation site (Figure 1). The chemical shift changes are found in the three other regions that have transient helicity, i.e., helices 0, 2, and 3, and no significant changes were seen in regions without helical propensity, e.g., the C-terminal region.

dx.doi.org/10.1021/ja4045532 | J. Am. Chem. Soc. 2013, 135, 10155−10163

The chemical shift changes are observed for all nuclei and are consistent with increased helicity in helix 0, 2, and 3 for all seven mutations within helix 1. For the helix-disrupting mutant (A1047G) the opposite pattern was observed. The local chemical shift changes in helix 1 are linearly related to the chemical shift changes observed in distant helices (Figure 3). This correlation found for a range of different mutations and nuclei suggests that the chemical shift changes are caused by a change in helicity, rather than unintended side-effects of the mutations. Notably, the correlation is stronger for the ^{13}C chemical shifts that are most sensitive to helical populations. The slopes are similar for C′ and C^{α} but differ for ^{15}N (Table 2). As we are mainly interested in changes in the helical populations, we will focus on the ^{13}C chemical shifts in the following. The distant chemical shift changes are most likely explained by transient long-range interactions between helix 1 and the other helices (Figure 3). When helix 1 is stabilized, the population with the long-range helix–helix interaction is increased, which thus leads to an increase in the populations of the interacting helices. When a linear function is fitted to the correlation between chemical shift changes in helix 1 and distant helices, the slopes correspond to the ratio between the fraction of induced helicity in the distant segment and helix 1. The slopes thus depend on both the strength of the helix–helix interaction and the entropic cost of forming the interaction. The average slopes for ^{13}C are 0.014, 0.086, and 0.023 for helices 0, 2 and 3, respectively. The slopes of the correlations decrease as the distance between the helical segments increases, which is expected as the entropic cost associated with forming a long-range interaction increases with the length of the linker between the two interacting regions. The slopes correspond to an equilibrium constant between a closed conformation with a helix–helix interaction and an open helical conformation. This equilibrium constant, however, only provides a lower bound on the equilibrium constant for the interaction as non-helical states will not give rise to chemical shift changes. The larger slopes observed for ^{15}N chemical shift changes suggest that the mutations also induce non-helical interactions.

Describing long-range interactions in wild-type ACTR using paramagnetic relaxation enhancements. To describe the long-range interaction network of ACTR further, we employed PREs. We made four mutants of ACTR (Q1020C, N1038C, N1058C, and N1078C), attached MTSL spin labels to each of the four cysteine residues, and four sets of PREs were measured. In the MTSL-modified mutants, a small but significant increase in the populations of all the helical segments was observed (Figure S1) corresponding to a 3–7% increase in the helical populations. The small structural perturbations induced by the presence of the spin label are an intrinsic limitation of PRE measurements but are unlikely to significantly influence the subsequent ensemble calculations.

To describe the long-range interaction network in ACTR, we selected conformational ensembles using ASTEROIDS[50] based on the experimental PREs. Initially, we systematically investigated the number of conformers necessary to reproduce the experimental data based on the four sets of PREs obtained from the wild-type protein. Ensembles were selected using ASTEROIDS with an increasing number of conformers (from 10 to 300 structures) from a pool of 10 000 conformers generated using Flexible-Meccano.[46,47] We carried out ASTEROIDS selections on the basis of the PREs of Q1020C, N1058C, and N1078C ("active" data) and used the PREs of N1038C ("passive" data) for cross-validation. The fit of both

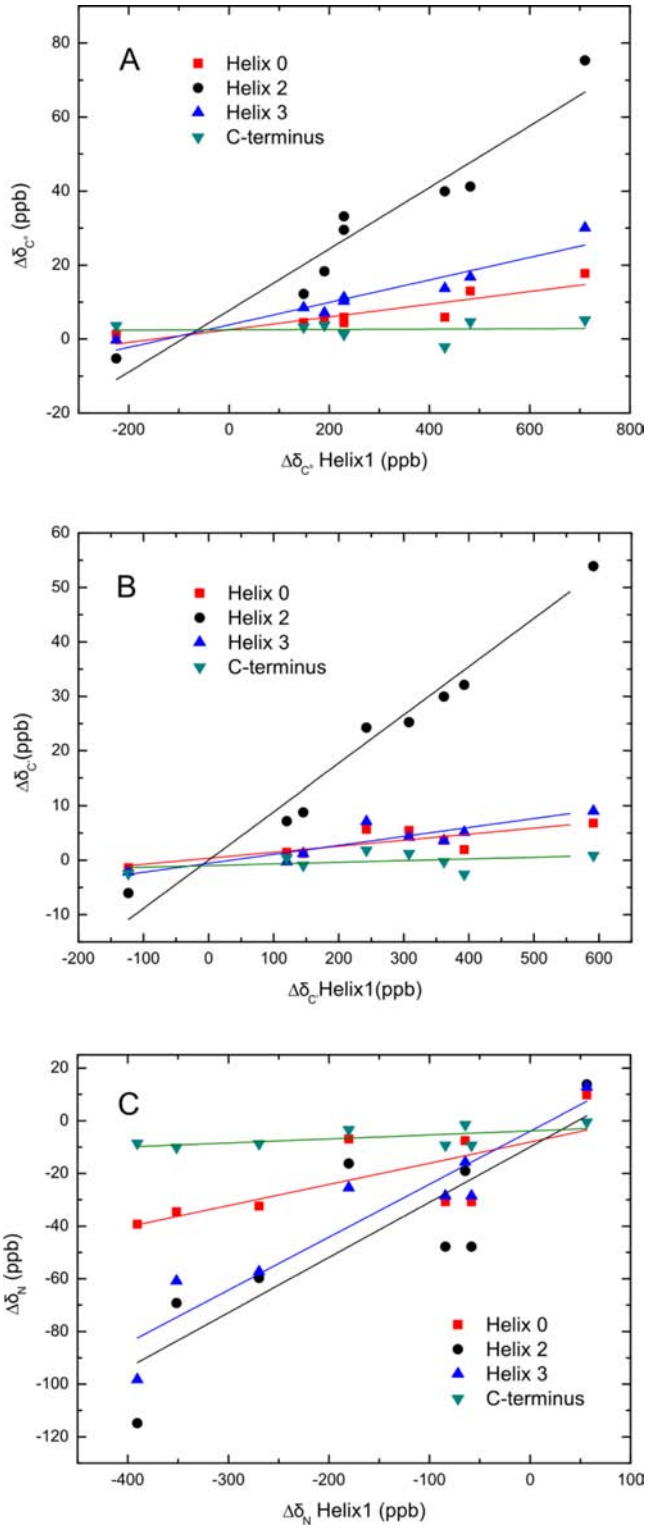


Figure 3. Correlation between local and non-local chemical shift changes caused by mutations. The average C^{α} (A), C′ (B) and N (C) chemical shift changes in helix 1 (residues 1044–1055) versus the average chemical shifts changes in helix 0 (residues 1023–1035), helix 2 (residues 1064–1072), helix 3 (residues 1074–1081), and the C-terminal region (residues 1080–1087). The linear dependence (indicated by full drawn lines) of chemical shift changes suggests the presence of long-range helix–helix interactions, where the slopes of the lines correspond to pseudoequilibrium constants.

dx.doi.org/10.1021/ja4045532 | J. Am. Chem. Soc. 2013, 135, 10155–10163

Table 2. Parameters from Linear Fits of Mutation-Induced Chemical Shift Changes

	C′		C^{α}		N	
	slope	R^2	slope	R^2	slope	R^2
helix 0	0.011 ± 0.003	0.60	0.017 ± 0.004	0.75	0.08 ± 0.03	0.44
helix 2	0.089 ± 0.007	0.95	0.083 ± 0.010	0.91	0.21 ± 0.05	0.67
helix 3	0.016 ± 0.003	0.77	0.030 ± 0.004	0.89	0.20 ± 0.03	0.87
C terminus	0.003 ± 0.003	0.00	0.000 ± 0.003	0.16	0.015 ± 0.008	0.27

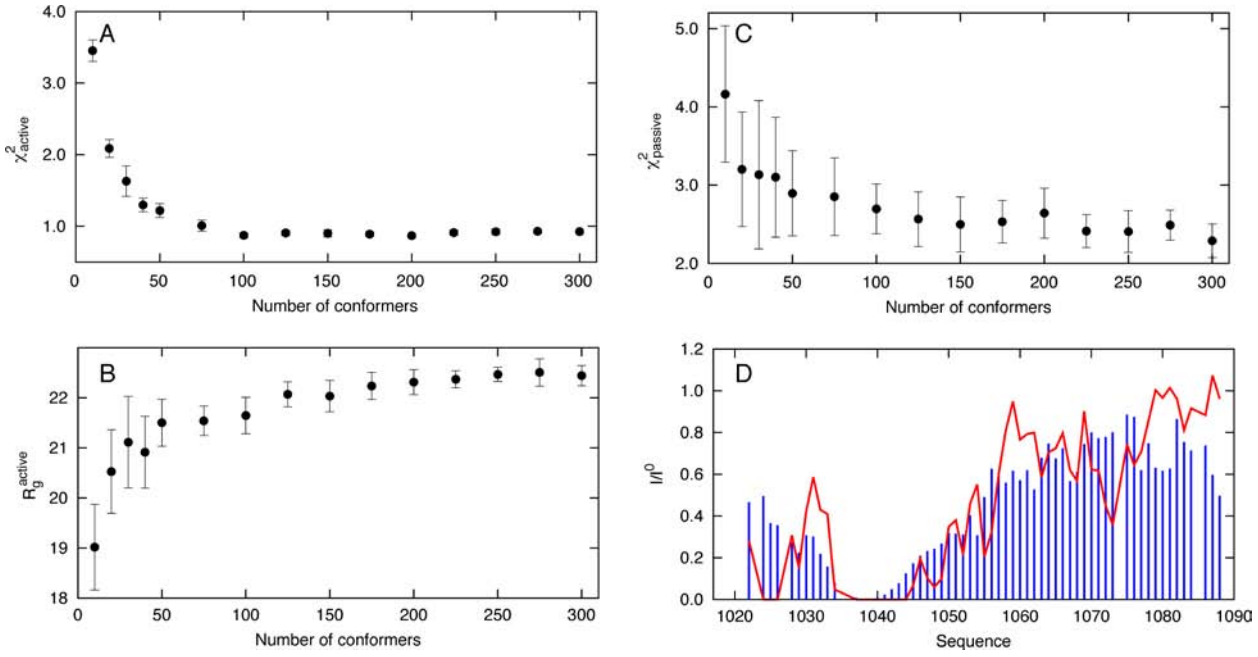


Figure 4. Selection of ensembles using ASTEROIDS on the basis of the experimental PRE data of wild-type ACTR. (A) χ^2 for the active PRE data (Q1020C, N1058C, and N1078C). (B) Average radius of gyration for the selected ensembles on the basis of the active PRE data. (C) χ^2 for the passive PRE data (N1038C). (D) Comparison (cross-validation) of PRE data from N1038C (red lines) with those back-calculated from an ASTEROIDS ensemble comprising 150 conformers selected on the basis of the PREs of Q1020C, N1058C, and N1078C (blue bars). Error bars correspond to averages over nine independent ensemble selections.

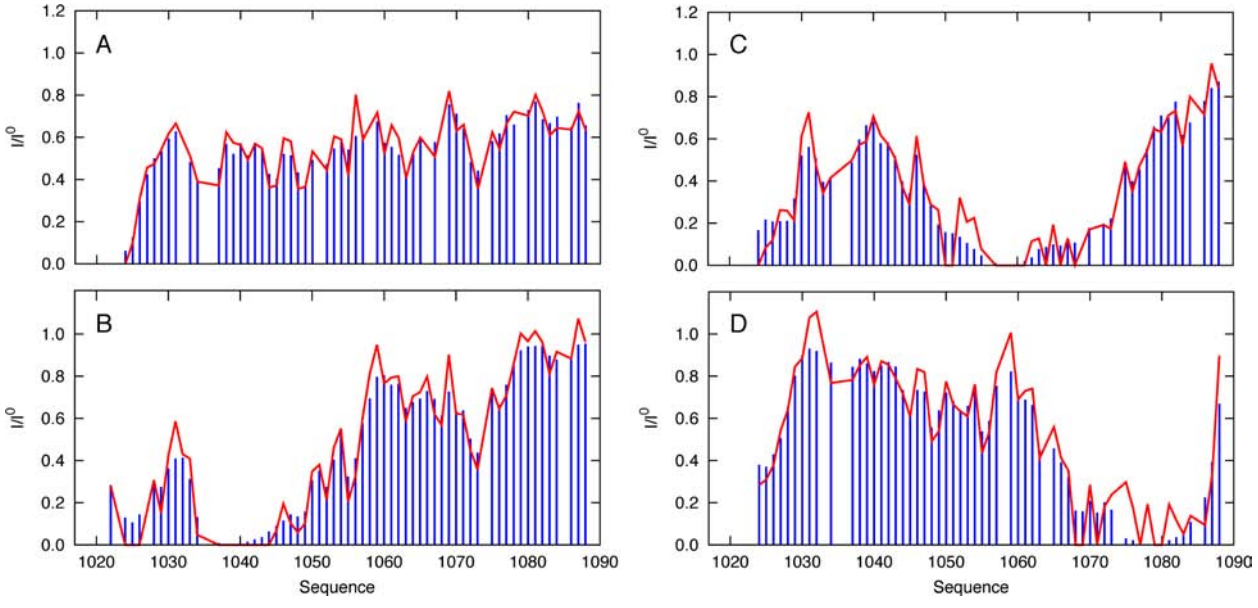


Figure 5. Reproduction of experimental PRE data from the ASTEROIDS-selected ensemble (150 conformers) of ACTR on the basis of the four sets of PRE data. Experimental values are shown as red lines, while back-calculated PREs from the ASTEROIDS ensemble are indicated by blue bars for each of the PRE data sets: Q1020C (A), N1038C (B), N1058C (C), and N1078C (D).

dx.doi.org/10.1021/ja4045532 | J. Am. Chem. Soc. 2013, 135, 10155−10163

the active and passive data as well as the convergence of the average radius of gyration show that an ensemble of 150 conformers is appropriate for mapping the long-range interactions in ACTR (Figure 4). We then selected an ensemble comprising 150 conformers on the basis of all four PRE data sets (Figure 5). The resulting contact map reveals a dominant long-range contact between the N-terminus of ACTR (residues 1018−1030) and the region comprising residues 1060−1080 (Figure 6).

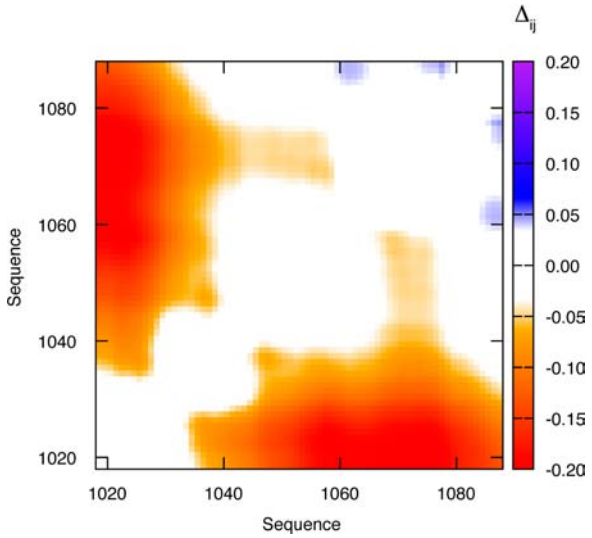


Figure 6. Contact map displaying long-range interactions in ACTR determined on the basis of the four experimental PRE data sets. A significant contact is seen between the N- and C-terminal regions.

Long-Range Interactions Probed by Residual Dipolar Couplings. Interestingly, the experimental $^{1}H-^{15}N$ RDCs disagree with those back-calculated from the best fitting structural ensembles for residues outside helix 1 (Figure 2C,D). As described previously, RDCs are sensitive to persistent long-range interactions between different parts of the disordered chain.[13,17] This dependence can be taken into account by considering experimental RDCs as a product of two contributions: a contribution from local conformational sampling only[21] and a contribution from a baseline that takes into account the chain-like nature of the disordered protein.[13,50] In the absence of specific long-range contacts, the baseline is a bell-shaped curve that becomes strongly modulated in the presence of specific long-range contacts by reinforcing the RDCs in the regions of the contact. Remarkably, the agreement between experimental and predicted RDCs outside helix 1 is improved considerably by applying an RDC baseline mimicking the long-range interaction between the N-terminus and the region comprising residues 1060−1080 derived from the experimental PREs (Figure 7). The RDC data are thus consistent with the dominating interaction detected by PREs.

The structural ensemble determined for ACTR should in principle contain the helix−helix interactions detected by mutagenesis. The ensemble selection procedure does not include an energy function that drives the hydrophobic helices together so we are unlikely to find direct contacts between the helices in the ensemble. Instead we tested proximity of the helical segments by measuring the fraction of the ensemble where the C^{α} of the central residue of each helix is within 10 Å. For all three helices the fraction of structures with the helix−

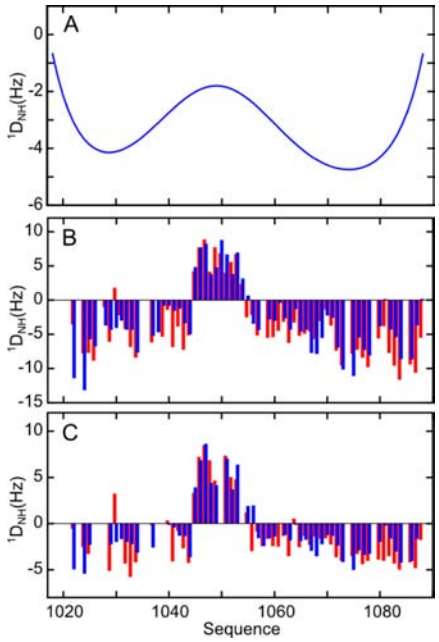


Figure 7. Introduction of a long-range contact improves the back-calculation of RDCs. (a) RDC baseline derived from the main long-range contact observed in the PRE ensemble between the N-terminus and the region 1060−1080. The agreement between experimental (red) and calculated (blue) RDCs improves in the regions surrounding helix 1 when the long-range contact is included for both the wild-type (B) and the triple mutant (C). The corresponding plots without the baseline are found in Figure 2C,F.

helix contacts are higher than in the random coil ensemble (Table 3).

Table 3. Number of Conformers with a Distance <10 Å to Helix 1 from Other Parts of the Ensemble

contact between residue 1049 and:	structural element	conformers within 10 Å in selected ensemble[a]	conformers within 10 Å in random coil ensemble[b]
1026	helix 0	8.5	2.4
1067	helix 2	3.0	1.0
1077	helix 3	4.5	1.0
1084	C-terminus	2.0	0.7

[a]Ensemble consists of 150 structures. [b]Ensemble consists of 10 000 structures.

DISCUSSION

The relationship between secondary structure and transient tertiary fold in intrinsically disordered or partially folded proteins is of fundamental importance if we are to develop a better understanding of the functional modes of these highly flexible proteins. Here, we demonstrate that modulation of the intrinsic helix propensity of a transiently formed α-helix in ACTR leads to non-local structural changes. This observation is similar to what was observed in the acid denatured state of acyl-coenzyme A binding protein (ACBP) for a range of single-site mutations.[22] The ACBP study aimed to remove side chains important for long-range interactions, whereas the present study preserves the interacting side chains and only modifies the helical propensities. The linear relationship between the

dx.doi.org/10.1021/ja4045532 | J. Am. Chem. Soc. 2013, 135, 10155−10163

induced chemical shifts in the modified helix and those in distant helices across a range of different mutations suggests that the changes in the long-range structure are almost exclusively a consequence of the change in helicity.

The transient helices in ACTR are all either negatively charged or neutral, and it is thus unlikely that long-range contacts are caused by electrostatic interactions. The helices are all amphipathic, and we have attempted not to alter the hydrophobic sides of these helices. When the helical population in helix 1 is increased, the effective concentration of the hydrophobic patch formed by the helix is thus increased even though the mutants themselves do not necessarily increase the hydrophobicity of the protein. If this helix participates in long-range hydrophobic interactions, the populations of these interactions will also be increased due to the mutations. The observation of increased helicity in distant regions further tells us that the mutated helix interacts with a subset of the helical conformations of these regions. In acid denatured ACBP, mutations throughout the four amphipathic helices lead to chemical shift changes distant in the primary sequence.[22] In contrast, mutational disruption of a hydrophilic helix only caused local chemical shift changes in the intrinsically disordered intracellular domain of the Na^+/H^+ exchanger 1 from the flounder *Pleuronectes americanus*.[56] These observations suggest that long-range chemical shift changes upon mutation are more common for amphipathic helices and that the interactions are dominated by hydrophobic interactions.

The question remains, whether the observed interactions are specific to ACTR or a manifestation of a general property of disordered proteins. Helix 1 interacts with all of the three other amphipathic helices in the domain suggesting that this interaction may not be specific. Interestingly, the chemical shift changes upon introduction of the MTSL tag support this conclusion. Modification with the hydrophobic MTSL tag slightly increases the helicity in all helical regions regardless of the modification site. This suggests that the tag generally leads to a slightly increased hydrophobic collapse that stabilizes the amphipathic helices via transient hydrophobic interactions. For an exposed hydrophobic patch in aqueous solution, it is favorable to interact with any other hydrophobic patch. Similar situations were observed in both denatured ACBP and lysozyme, where mutations of hydrophobic residues leads to less compact states.[22,59] Therefore the transient hydrophobic interactions we observe here are likely to be general to all flexible proteins with clusters of hydrophobic residues.

The PRE data and ensuing ensemble calculation are dominated by a long-range contact between the N-terminus and the region comprising residues 1060−1080. This interaction is likely to be electrostatic as the first ten residues in our construct have a net positive charge of three while the remainder of the domain is negatively charged. The long-range contact predominately observed in the ensemble is thus different from the contacts deduced by the chemical shift changes upon mutation. We can estimate the populations of the contacts observed by mutagenesis, if we consider the slope of the line relating chemical shift changes in helix 1 to changes in the other helices (maximally 0.086) as a pseudoequilibrium constant for the formation of the helix−helix interactions. In wild-type ACTR that has 38% transient helicity in helix 1, we arrive at a maximal contact of <4%. Contacts with such a low population are likely to be obscured by stronger contacts and are thus not apparent in the ensemble selected on the basis of PREs. The population of contacting residues in the ensemble is approximately the same size as the estimates from the mutations (Table 3), which suggests that the helix−helix interactions are consistent with the calculated ensemble, although they do not contribute detectably to the contact map. The mutagenesis approach thus provides complementary information about long-range contacts as it only reports on helix−helix interactions, whereas PREs are sensitive to proximity in general.

CONCLUSION

In conclusion, we have demonstrated that modulation of the population of a transient helix has long-range effects due to hydrophobic helix−helix interactions. Rational mutagenesis monitored by NMR spectroscopy thus provides an attractive and generally applicable experimental strategy for probing this type of interactions. Furthermore, as long-range contacts in the disordered state may affect binding kinetics and mechanisms, the conformational ensemble determined here may be an optimal starting point for future simulations of the coupled folding and binding reaction between NCBD and ACTR.

ASSOCIATED CONTENT

Supporting Information

Chemical shift differences between wild-type ACTR and MTSL-labeled cysteine mutants. Statistics for the top three helical fit models. This information is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author

mk710@cam.ac.uk

Present Address

[§]Department of Chemistry, Cambridge University, United Kingdom

Notes

The authors declare no competing financial interest.

[‖]Deceased

ACKNOWLEDGMENTS

This work was supported by a J. C. Jacobsen memorial scholarship from the Carlsberg Foundation (V.I.), The John and Birthe Meyer Foundation, the Carlsberg Foundation grant no. 2008-01-0368, The Danish Natural Research Council grant no. 272-08-0500 (F.M.P.) and the French Agence National pour la Recherche through ANR JCJC Protein Disorder (M.R.J.) and ANR MALZ TAUSTRUCT (M.B.). We thank Kaare Teilum and Gitte Wolfsberg Haxholm for critical comments to the manuscript.

REFERENCES

(1) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. *Annu. Rev. Biophys.* **2008**, *37*, 215−46.
(2) Uversky, V. N. *Front. Biosci.* **2009**, *14*, 5188−238.
(3) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321−31.
(4) Uversky, V. N. *Protein Sci.* **2002**, *11*, 739−56.
(5) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197−208.
(6) Mittag, T.; Forman-Kay, J. D. *Curr. Opin. Struct. Biol.* **2007**, *17*, 3−14.
(7) Eliezer, D. *Curr. Opin. Struct. Biol.* **2009**, *19*, 23−30.
(8) Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490−5492.
(9) Wishart, D.; Sykes, B. *J. Biomol. NMR* **1994**, *4*, 171−180.

(10) Kjaergaard, M.; Poulsen, F. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *60*, 42−51.
(11) Marsh, J. A.; Singh, V. K.; Jia, Z.; Forman-Kay, J. D. *Protein Sci.* **2006**, *15*, 2795−804.
(12) Tamiola, K.; Mulder, F. A. A. *Biochem. Soc. Trans.* **2012**, *40*, 1014−20.
(13) Salmon, L.; Nodet, G.; Ozenne, V.; Yin, G.; Jensen, M. R.; Zweckstetter, M.; Blackledge, M. *J. Am. Chem. Soc.* **2010**, *132*, 8407−8418.
(14) Gillespie, J. R.; Shortle, D. *J. Mol. Biol.* **1997**, *268*, 158−69.
(15) Gillespie, J. R.; Shortle, D. *J. Mol. Biol.* **1997**, *268*, 170−84.
(16) Bertoncini, C. W.; Jung, Y.-S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 1430−5.
(17) Schneider, R.; Huang, J.; Yao, M.; Communie, G.; Ozenne, V.; Mollica, L.; Salmon, L.; Jensen, M. R.; Blackledge, M. *Mol. BioSyst.* **2012**, *8*, 58−68.
(18) Jensen, M. R.; Markwick, P. R. L.; Meier, S.; Griesinger, C.; Zweckstetter, M.; Grzesiek, S.; Bernadó, P.; Blackledge, M. *Structure* **2009**, *17*, 1169−85.
(19) Bernadó, P.; Bertoncini, C. W.; Griesinger, C.; Zweckstetter, M.; Blackledge, M. *J. Am. Chem. Soc.* **2005**, *127*, 17968−9.
(20) Marsh, J. A.; Forman-Kay, J. D. *J. Mol. Biol.* **2009**, *391*, 359−74.
(21) Marsh, J. A.; Baker, J. M. R.; Tollinger, M.; Forman-Kay, J. D. *J. Am. Chem. Soc.* **2008**, *130*, 7804−5.
(22) Bruun, S. W.; Iesmantavicius, V.; Danielsson, J.; Poulsen, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 13306−11.
(23) Chen, H.; Lin, R. J.; Schiltz, R. L.; Chakravarti, D.; Nash, A.; Nagy, L.; Privalsky, M. L.; Nakatani, Y.; Evans, R. M. *Cell* **1997**, *90*, 569−80.
(24) Demarest, S. J.; Martinez-Yamout, M.; Chung, J.; Chen, H.; Xu, W.; Dyson, H. J.; Evans, R. M.; Wright, P. E. *Nature* **2002**, *415*, 549−53.
(25) Knott, M.; Best, R. B. *PLoS Comput. Biol.* **2012**, *8*, e1002605.
(26) Burger, V. M.; Ramanathan, A.; Savol, A. J.; Stanley, C. B.; Agarwal, P. K.; Chennubhotla, C. S. *Pac. Symp. Biocomput.* **2012**, *1*, 70−81.
(27) Zhang, W.; Ganguly, D.; Chen, J. *PLoS Comput. Biol.* **2012**, *8*, e1002353.
(28) Ganguly, D.; Zhang, W.; Chen, J. *Mol. BioSyst.* **2012**, *8*, 198−209.
(29) Naganathan, A. N.; Orozco, M. *J. Am. Chem. Soc.* **2011**, *133*, 12154−61.
(30) Kjaergaard, M.; Teilum, K.; Poulsen, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 12535−40.
(31) Dogan, J.; Schmidt, T.; Mu, X.; Engstrom, A.; Jemth, P. *J. Biol. Chem.* **2012**, *287*, 34316−24.
(32) Kjaergaard, M.; Poulsen, F. M.; Teilum, K. *Biophys. J.* **2012**, *102*, 1627−1635.
(33) Kjaergaard, M.; Andersen, L.; Nielsen, L. D.; Teilum, K. *Biochemistry* **2013**, *52*, 1686−1693.
(34) Japrung, D.; Dogan, J.; Freedman, K. J.; Nadzeyka, A.; Bauerdick, S.; Albrecht, T.; Kim, M. J.; Jemth, P.; Edel, J. B. *Anal. Chem.* **2013**, *85*, 2449−56.
(35) Ebert, M.-O.; Bae, S.-H.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2008**, *47*, 1299−308.
(36) Kjaergaard, M.; Nørholm, A.-B.; Hendus-Altenburger, R.; Pedersen, S. F.; Poulsen, F. M.; Kragelund, B. B. *Protein Sci.* **2010**, *19*, 1555−64.
(37) Kjaergaard, M.; Iešmantavièius, V.; Poulsen, F. M. *Protein Sci.* **2011**, *20*, 2023−34.
(38) Sass, H. J.; Musco, G.; Stahl, S. J.; Wingfield, P. T.; Grzesiek, S. *J. Biomol. NMR* **2000**, *18*, 303−9.
(39) Chou, J. J.; Gaemers, S.; Howder, B.; Louis, J. M.; Bax, A. *J. Biomol. NMR* **2001**, *21*, 377−82.
(40) Ottiger, M.; Delaglio, F.; Bax, A. *J. Magn. Reson.* **1998**, *131*, 373−8.
(41) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67−81.
(42) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−93.
(43) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins* **2005**, *59*, 687−96.
(44) Jensen, M. R.; Houben, K.; Lescop, E.; Blanchard, L.; Ruigrok, R. W. H.; Blackledge, M. *J. Am. Chem. Soc.* **2008**, *130*, 8055−61.
(45) Jensen, M. R.; Communie, G.; Ribeiro, E. A.; Martinez, N.; Desfosses, A.; Salmon, L.; Mollica, L.; Gabel, F.; Jamin, M.; Longhi, S.; Ruigrok, R. W. H.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 9839−44.
(46) Ozenne, V.; Bauer, F.; Salmon, L.; Huang, J.-R.; Jensen, M. R.; Segard, S.; Bernadó, P.; Charavay, C.; Blackledge, M. *Bioinformatics* **2012**, *28*, 1463−70.
(47) Bernadó, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R. W. H.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17002−7.
(48) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.
(49) Shen, Y.; Bax, A. *J. Biomol. NMR* **2007**, *38*, 289−302.
(50) Nodet, G.; Salmon, L.; Ozenne, V.; Meier, S.; Jensen, M. R.; Blackledge, M. *J. Am. Chem. Soc.* **2009**, *131*, 17908−18.
(51) Bibow, S.; Ozenne, V.; Biernat, J.; Blackledge, M.; Mandelkow, E.; Zweckstetter, M. *J. Am. Chem. Soc.* **2011**, *133*, 15842−5.
(52) Iwahara, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 5879−96.
(53) Munoz, V.; Serrano, L. *Nat. Struct. Biol.* **1994**, *1*, 399−409.
(54) Muñoz, V.; Serrano, L. *J. Mol. Biol.* **1995**, *245*, 275−296.
(55) Muñoz, V.; Serrano, L. *J. Mol. Biol.* **1995**, *245*, 297−308.
(56) Nørholm, A.-B.; Hendus-Altenburger, R.; Bjerre, G.; Kjaergaard, M.; Pedersen, S. F.; Kragelund, B. B. *Biochemistry* **2011**, *50*, 3469−3480.
(57) Jensen, M. R.; Blackledge, M. *J. Am. Chem. Soc.* **2008**, *130*, 11266−7.
(58) Serrano, L.; Sancho, J.; Hirshberg, M.; Fersht, A. R. *J. Mol. Biol.* **1992**, *227*, 544−59.
(59) Wirmer, J.; Schlörb, C.; Klein-Seetharaman, J.; Hirano, R.; Ueda, T.; Imoto, T.; Schwalbe, H. *Angew. Chem., Int. Ed.* **2004**, *116*, 5904−5909.