BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

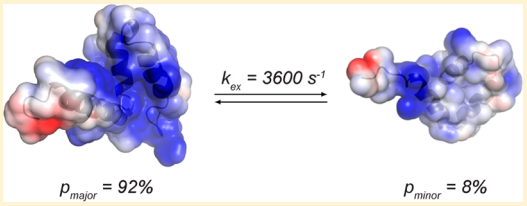
pubs.acs.org/biochemistry

A Folded Excited State of Ligand-Free Nuclear Coactivator Binding Domain (NCBD) Underlies Plasticity in Ligand Recognition

Magnus Kjaergaard,[†] Lisbeth Andersen, Lau Dalby Nielsen, and Kaare Teilum*

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

ABSTRACT: Intrinsically disordered proteins are renowned for their structural plasticity when they undergo coupled folding and binding to partner proteins. The nuclear coactivator binding domain of CBP is a remarkable example of this adaptability as it folds into two different conformations depending on the binding partner. To understand the role of the conformational ensemble for plasticity in ligand recognition, we investigated the millisecond dynamics of this domain using relaxation dispersion NMR spectroscopy. All NMR signals originating from the domain are broadened, demonstrating that the whole domain experience conformational exchange. The dispersion data can be described by a global two-state exchange process between a ground state and an excited state populated to 8%. The three helices are still folded in the excited state but have a different packing from the ground state; the contact between helices 2 and 3 found in the ground state is broken in the excited state, and a new one is formed between helices 1 and 3. This suggests that while NCBD in the ground state has a structure similar to the complex with the ligand ACTR, the conformation of NCBD in the excited state has some similarity with that of NCBD in complex with the ligand IRF-3. The energy landscape of this domain is thus proposed to resemble the fold-switching proteins that have two coexisting native states, which may serve as a starting point for binding via conformational selection.



Intrinsically disordered proteins (IDPs) are heavily associated with molecular signaling processes where their ability to bind different proteins allows multiple signaling pathways to compete for the cellular transcriptional machinery.[1] Binding to several different interaction partners often requires that the protein adapts its structure to fit that of its partners.[2] The lack of a unique stable structure in the unbound state makes IDPs especially appropriate for this task. While many of these proteins are promiscuous, they still need to recognize only a small subset of the potential binding partners. For folded proteins, discrimination between partners and nonpartners relies on complementary structures. This is not immediately possible for IDPs due to their lack of structure. Instead, recent research has focused on describing the conformational ensembles of IDPs. For some IDPs the conformational ensemble approaches a random coil while for other IDPs it contains transient structure and is molten-globule-like. The overall aim of this study is to address the role of transient structures in promiscuous molecular recognition. A key question is whether the conformations found in the complexes are populated in the ligand-free ensemble.

CREB binding protein (CBP) is a transcriptional coactivator and is a particularly apt example of promiscuous ligand recognition as a plethora of transcription factors compete for binding to its small ligand binding domains.[3] Of these ligand binding domains, the nuclear coactivator binding domain (NCBD) is an especially interesting example as it adopts two different conformations when bound to two of its interaction partners: the activator for thyroid hormone and retinoid receptors (ACTR) and interferon regulatory factor 3 (IRF-3).[4,5] These two conformations have the same three helical regions, but the helices are organized into different tertiary structures. Because of its remarkable adaptability, NCBD is studied as a model system to understand plasticity in molecular recognition.[6−10] In the ligand-free state, NCBD has the characteristics of a molten globule with a high degree of helicity,[11,12] rapid hydrogen exchange,[13] ANS binding suggesting exposed hydrophobic groups,[11] and a gradual thermal denaturation profile.[6,11] Despite the molten-globule-like behavior, a conformation that resembles the complex with ACTR dominates the free ensemble.[6] This conformation unfolds cooperatively as demonstrated by the simultaneous change of all NMR chemical shifts in a urea titration.[14] The denaturation experiments also revealed that the folded state of NCBD is only marginally stable. Consequently, a small fraction of the molecules are unfolded under native conditions.[6,14] Molten globules are generally believed to have hyperdynamic hydrophobic cores.[15] The hydrophobic core of free NCBD, however, was found by NMR relaxation experiments to be only slightly more dynamic than an average folded protein,[14] which suggests that NCBD behaves differently from the archetypical molten globules. Recently, experimental evidence for additional conformational states of ligand-free NCBD have come from single molecule translocations through solid-state nanopores. These experiments revealed three different translocation events corresponding to three different conformations.[16]

Because of its small size and dynamics, NCBD is an optimal model system for computational studies of flexible molecular recognition. The time scales of folding and large scale

Received: January 28, 2013
Published: February 1, 2013

dx.doi.org/10.1021/bi4001062 | *Biochemistry* 2013, 52, 1686−1693

conformational reconfigurations of NCBD are still beyond the reach of all-atom simulations. Therefore, the computational studies of NCBD have used various tricks ranging from coarse-grained representations, where the protein is modeled as a string of C^{α} atoms[9,10] to replica-exchange all-atom simulations in explicit water.[17] The conformational ensembles resulting from these molecular simulations are quite different, but all find multiple different folded conformations. One study suggested that NCBD can interconvert between multiple folded conformations without passing a free energy barrier in a mechanism that resembles downhill folding.[9] This suggests that the conformational ensemble was continuous rather than composed of several distinct states, although this prediction awaits experimental confirmation.

In this study, we used nuclear spin relaxation experiments, which are capable of detecting minute quantities of short-lived excited states exchanging with the dominant background of ground-state protein on the millisecond time scale.[18] The relaxation dispersion of a nuclear spin depends on the rate of exchange between the ground state and the excited state, together with their relative populations and the chemical shift difference between them. Thus, under optimal conditions the experiments determine the kinetics and thermodynamics of the conformational transition and also provide structural information on the excited state. We show that the dominant folded conformation exchanges rapidly with a low-populated minor state that represents an alternative folded state of NCBD. The two discrete states of NCBD in the free ensemble can thus provide the starting point for binding in a conformational selection type mechanism, resulting in two radically different ligand-bound conformations.

EXPERIMENTAL PROCEDURES

Protein Production. The plasmid for coexpression of NCBD of murine CBP (UniProt: P45481, residues 2059−2117 in the mature sequence) with the activation domain of ACTR (UniProt: Q9Y6Q9, residues 1023−1093) was a gift from Peter E. Wright (The Scripps Research Institute).[4] NCBD was expressed in M9 minimal medium and purified as described previously.[14] In addition to a uniform ^{15}N-labeled sample, three samples produced with different carbon isotope labeling schemes were used: (i) uniform ^{13}C labeling using [$^{13}C_6$]-glucose (Glc) as sole carbon source and 100% H_2O as growth solvent; (ii) [1-^{13}C]-Glc as sole carbon source and 100% H_2O as growth solvent; (iii) [1-^{13}C]-Glc as sole carbon source and 100% D_2O as growth solvent. All labeling schemes also contained uniform ^{15}N labeling. Growth on [1-^{13}C]-Glc ensures that most methyl groups do not have a ^{13}C neighbor and consequently avoid strong ^{13}C−^{13}C one-bond couplings.[19] [1-^{13}C]-Glc as carbon source also ensures that only the C^{δ} position of Tyr and Phe become ^{13}C labeled.[20] This sample was used for ^{15}N backbone and ^{13}C methyl Carr−Purcell−Meiboom−Gill (CPMG) relaxation dispersion experiments and for off-resonance $R_{1\rho}$ experiments on aromatic side chains. The sample grown in 100% D_2O is enriched in the $^{13}CHD_2$ isotopomer[21] and was used for methyl ^{1}H CPMG experiments. Uniformly ^{13}C-labeled NCBD was used for ^{13}C′ CPMG experiments. The ^{15}N single-labeled sample was used for $R_{1\rho}$ relaxation experiments on backbone ^{15}N. NMR samples were prepared by dissolving lyophilized NCBD in 20 mM phosphate, 10% D_2O, and adjusting the pH to 6.5. The protein concentration in the samples was 3.9 mM for uniformly labeled NCBD and 1.5 mM for the other samples.

NMR Spectroscopy. NMR experiments were recorded on Varian VnmrS 500 MHz (11.7 T), Varian Inova 750 MHz (17.6 T), and Agilent DD2 800 MHz (18.8 T) spectrometers equipped with conventional triple-resonance probes. CPMG relaxation dispersion profiles of ^{15}N, ^{13}C′, and ^{13}C in $^{13}CH_3$ groups and ^{1}H in $^{13}CHD_2$ groups were recorded using published pulse sequences.[21−24] The apparent transverse relaxation rates, $R_{2,app}$, were determined with the constant time approach.[25] ^{15}N dispersions were recorded with a relaxation delay of 40 ms by sampling 15 different CPMG field strengths ranging from 50 to 1000 Hz. The dispersions were measured at 19, 22, 25, 28, and 31 °C at 17.6 T and at 19 °C at 11.7 T. The remaining three CPMG dispersion experiment were recorded only at 19 °C. ^{13}C′ and $^{13}CH_3$ dispersions were recorded with a relaxation delay of 40 ms by sampling 13−15 different CPMG field strengths ranging from 50 to 900 Hz at both 17.6 and 11.7 T. CHD_2 dispersions were recorded with a relaxation delay of 90 ms by sampling 55 different CPMG field strengths ranging from 22 to 5700 Hz only at 17.6 T. Off-resonance $R_{1\rho}$ relaxation dispersion experiments on ^{15}N were recorded at 18.8 T and 31 °C using the pulse sequence of Mulder et al.[26] with spin-lock field strengths from 431 to 1649 Hz and offsets ranging from 0 to 10 000 Hz. Off-resonance $R_{1\rho}$ relaxation dispersion experiments on $^{13}C^{\delta}$ of Phe and Tyr were recorded at 17.6 T and 19 °C using the pulse sequence of Igumenova and Palmer[27] modified for ^{13}C, with spin-lock field strengths from 810 to 4452 Hz and offsets ranging from 500 to 20 000 Hz.

Data Analysis. All spectra were processed and analyzed with nmrPipe.[28] Peak intensities were quantified by summing the data in a 3 × 3 point window centered on the cross-peak. Relaxation dispersions were fitted both to a constant value and to the general expression for a two-state process.[22,29] For the ^{15}N dispersions measured at varying temperatures at 17.6 T, all data for a single residue were fit simultaneously with the chemical shift difference between the exchanging states as a global parameter. For the dispersions measured at 19 °C, the data for each nucleus and for each residue were initially fit individually. The *F*-test was used to determine whether each data set showed significant dispersions with $P > 0.001$ as significance criteria. Data sets with significant relaxation dispersion and with dispersion steps larger than 1 s^{-1} were included in a global fit with the exchange rate, k_{ex}, and the population of the minor state, p_{minor}, as global parameters. All fits were performed using IgorPro v6 (WaveMetrics) or OriginPro v8.6 (OriginLab). To assess the robustness and quality of the global fits, 20% of the data sets that were included in the full global fit were randomly left out in a new global fit; this procedure was repeated 1000 times. The 1000 sets of optimized parameters from these reduced data sets were analyzed and compared to the optimized global fit parameters of the full data set. The $R_{1\rho}$ rates were fitted to the expression for a two-state process in fast exchange, $R_{1\rho} = R_1 \cos^2\theta + R_{2,0} \sin^2\theta + [k_{ex}\phi_{ex}/(k_{ex}^2 + \omega_{eff}^2)]\sin^2\theta$, where θ is the tilt angle of the magnetization, k_{ex} is the exchange rate between the two states, ω_{eff} is the effective field of the spin-lock, $\phi_{ex} = \Delta\omega^2 p_{minor}(1 - p_{minor})$, and $\Delta\omega$ is the difference in chemical shifts between the exchanging states. Secondary chemical shifts were determined using backbone random coil chemical shifts calculated using the server at http://www.bio.ku.dk/randomcoil/,[30,31] and methyl random coil chemical shifts were based on previous peptide studies.[32]

dx.doi.org/10.1021/bi4001062 | *Biochemistry* 2013, 52, 1686−1693

Distribution of ^{1}H Methyl Chemical Shifts More Than 6 Å from Aromatic Groups. Nineteen structures were selected from recent NMR structures deposited by the NorthEast Structural Genomics consortium. Structures determined using sparse data sets and of proteins containing cofactors were excluded. The PDB/BMRB accession codes are 2LFI/17754, 2LK2/17971, 2LLK/18051, 2LM4/18098, 2LN3/18145, 2LND/18161, 2LNI/18166, 2LPK/18263, 2LQ3/18290, 2LR0/18337, 2LRQ/18390, 2LSE/18429, 2LSO/18438, 2LTA/18465, 2LTD/18469, 2LTE/18468, 2LTL/18487, 2LVA/18560, and 2LVB/18561. Only the first conformer in each ensemble and only methyl groups where the protons are more than 6 Å away from any atom in the aromatic ring were considered. The average proton chemical shift was calculated for each residue type and subtracted from the chemical shift. A single standard deviation was determined for all residue types.

RESULTS

NCBD Undergoes Reversible Conformational Exchange. NMR spectra of free NCBD are characterized by line-broadening, suggesting that the protein undergoes reversible conformational exchange on the millisecond time scale. From ^{1}H,^{15}N-heteronuclear single quantum coherence (HSQC) spectra measured at 19 and 31 °C, it is readily apparent that the line widths of peaks from the backbone amides in NCBD are strongly temperature dependent (Figure 1A). At 19 °C, it is still possible to reliably quantify the intensity of most peaks, but at lower temperatures the peaks are broadened to an extent that makes a qualitative analysis impossible. Even at 31 °C the peaks are much broader than expected for a 59-residue protein like NCBD. This is consistent with previous NMR relaxation studies on NCBD that were performed at 31 °C.[12,14] To quantify the exchange contribution to transverse relaxation, we recorded ^{15}N CPMG relaxation dispersion experiments at five temperatures (19, 22, 25, 28, and 31 °C). Residue specific exchange contribution varies from 0 to more than 30 s^{-1} at 19 °C (Figure 1B). Most of the residues with no exchange contribution to the transverse relaxation are localized at the flexible N- and C-termini. At increasing temperatures the exchange contribution decreases, and the dispersion profiles for all residues are almost flat at 31 °C. It is important to notice that the peaks from the amides are still broadened at 31 °C, suggesting the presence of an additional exchange process, which cannot be suppressed by the CPMG experiment. The CPMG dispersion profiles from 22 amide nitrogens showing significant dispersion were fit globally to a two-state exchange process including data from all five temperatures. Because of the decreasing amplitude of the exchange contribution, the fits of the data from 25−31 °C are not well constrained. At 19 and 22 °C, k_{ex} extracted from the fit are $(2.8 \pm 0.2) \times 10^3$ and $(4.0 \pm 0.4) \times 10^3$ s^{-1}, respectively. At such high exchange rates and with data from only one static field, the optimized values of k_{ex} and p_{minor} become highly correlated.[33]

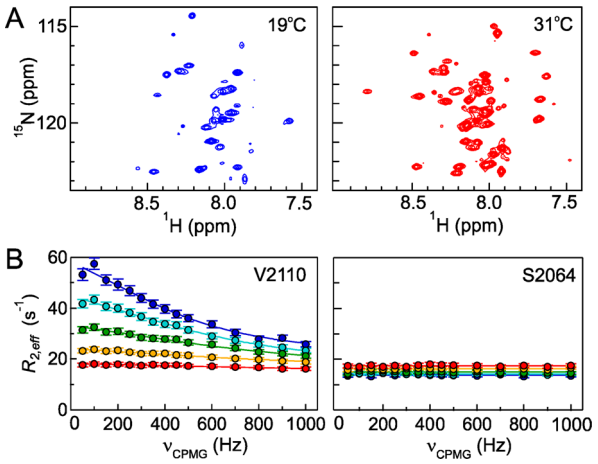


Figure 1. Temperature-dependent line-broadening of NMR signals in NCBD. (A) Extract of ^{1}H,^{15}N-HSQC spectra of NCBD recorded at 19 and 31 °C contoured at the same intensity level. (B) Representative ^{15}N CPMG relaxation dispersions ($R_{2,eff}$ vs ν_{CPMG}) recorded for V2110 and for S2064, which shows no dispersion, at five temperatures (blue, 19 °C; cyan, 22 °C; green, 25 °C; yellow, 28 °C; red, 31 °C). The solid lines for V2110 represent a global fit to a two-state process of the ^{15}N CPMG relaxation dispersion data at five temperatures for 22 residues with significant dispersion.

To characterize the conformational exchange remaining at 31 °C, we recorded off-resonance $R_{1\rho}$ relaxation dispersion of amide ^{15}N (Figure 2). The $R_{1\rho}$ relaxation dispersion method is suitable for characterizing exchange processes with rates up to 50 000 s^{-1}, whereas the CPMG relaxation method is limited to exchange processes with rates up to ~4000 s^{-1}. At 31 °C, 15 residues show significant $R_{1\rho}$ relaxation dispersion, which can be fit to a two-state process with k_{ex} as global parameter. The optimized value for k_{ex} is $(1.3 \pm 0.2) \times 10^4$ s^{-1}.

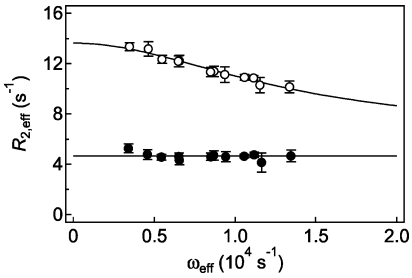


Figure 2. Relaxation dispersion profiles of R_2 as a function of the effective field strength ω_{eff} extracted from off-resonance $R_{1\rho}$ relaxation measurement of $^{15}N^{H}$ from S2064 (filled symbols) and V2110 (open symbols). The solid line through the data points for V2110 represents a global fit to a two-state process of the ^{15}N $R_{1\rho}$ relaxation dispersion data for 15 residues with significant dispersion. S2064 does not show significant exchange and fits a constant value. The data were recorded at a static magnetic field of 18.8 T and 31 °C.

NCBD Has Two Folded Conformations. To get more structural information about the minor state and at the same time get more data to constrain k_{ex} and p_{minor} in our data analysis at 19 °C, we recorded additional CPMG relaxation dispersions on backbone $^{15}N^{H}$ and $^{13}C'$ as well as on methyl ^{13}C and methyl ^{1}H (Figure 3). Except for the methyl ^{1}H, all experiments were measured at two static magnetic fields. 100 data sets with significant dispersion were fit to a global two-state process. This global fit is stable to random reduction of the data by 20%, and the correlation between k_{ex} and p_{minor} is broken. The fit satisfactorily describes the measured dispersions with $k_{ex} = (3.6 \pm 0.2) \times 10^3$ s^{-1} and $p_{minor} = 8.2 \pm 0.7\%$ (Figure 3). In addition and for each nucleus, the unsigned difference in chemical shift between the major and the minor state, $|\Delta\omega|$, is obtained from the fit.

dx.doi.org/10.1021/bi4001062 | Biochemistry 2013, 52, 1686−1693

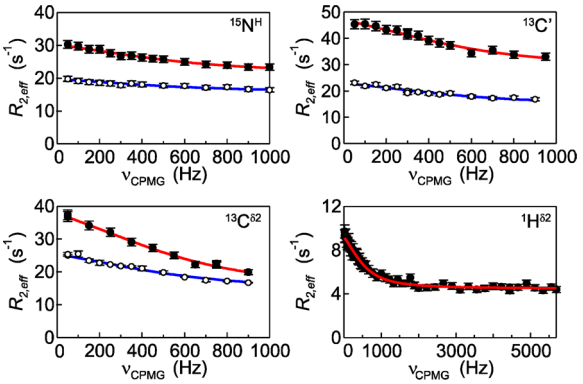


Figure 3. CPMG relaxation dispersions for $^{15}N^{H}$, $^{13}C'$, $^{13}C^{\delta2}$, and $^{1}H^{\delta2}$ of residue L2068 in NCBD at 19 °C. The dispersions were recorded at static magnetic fields of 17.6 T (filled symbols) and 11.7 T (open symbols). The solid lines represent a global fit to a two-state process of all CPMG relaxation dispersion data at 19 °C with significant dispersion.

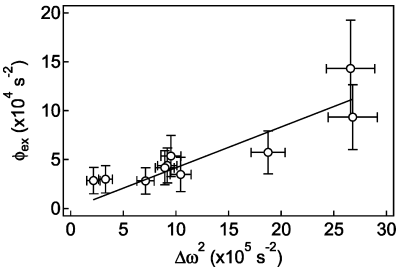


Figure 5. Structural similarity of the excited state of NCBD at 19 and 31 °C. ϕ_{ex} from ^{15}N $R_{1\rho}$ relaxation dispersions at 31 °C plotted against $\Delta\omega^2$ from ^{15}N CPMG relaxation dispersion at 19 °C. The solid line represents the best fit of a line through (0,0) with slope $p_{minor}(1 - p_{minor})$ = 0.041 ± 0.005 equiv of p_{minor} = 4.3%.

Urea denaturation experiments of NCBD have shown that the folded state is marginally stable and that the unfolded state is populated to a few percent under native conditions.[6,14] To test whether the excited state probed by the relaxation dispersion experiments corresponds to the unfolded state, $|\Delta\omega|$ extracted from the CPMG experiments were compared to the secondary chemical shift (SCS), i.e., the difference in chemical shift between the native state and the unfolded state (Figure 4A). If the excited state is fully unfolded, a strong correlation between SCS and $|\Delta\omega|$ is expected. There is, however, poor correlation for the backbone nuclei, suggesting that the excited state probed by the relaxation dispersion experiments is not unfolded.

To test if the exchange process probed by the $R_{1\rho}$ relaxation dispersion at 31 °C is similar to the process probed by the CPMG relaxation dispersion at 19 °C, we compared ϕ_{ex} from the fit of the $R_{1\rho}$ dispersions with $\Delta\omega$ from the CPMG dispersions. If the excited states have similar structures, ϕ_{ex} is expected to be directly proportional to $\Delta\omega^2$. This is indeed the situation, as evident from Figure 5 from which the population of the excited state p_{minor} can be estimated to 4.3% from the slope of best fit straight line. The population of the minor state thus decreases with increasing temperature, which infers $\Delta H < 0$ for the change from the major to the minor state. As we at the same time have that $\Delta G > 0$, the transition to the minor state must be entropically unfavorable, which could indicate exposure of hydrophobic surface area. This behavior is similar to that observed for an alternative compact state in superoxide dismutase.[34]

The Hydrophobic Core Is Rearranged in the Alternative Folded Conformation. Chemical shifts of excited states extracted from relaxation dispersion experiments have previously with success been used to calculate three-dimensional models of excited states of proteins.[35−37] This kind of structure calculation requires determination of the signs of the chemical shift differences. Structural information about excited states may also be obtained from residual dipolar couplings, but again the signs of the chemical shift differences are needed.[38] We tried unsuccessfully to determine the signs of the chemical shift differences using both the H(S/M)QC and the $R_{1\rho}$ methods.[39,40] Both methods work best for relatively slow processes, whereas the exchange rate for NCBD is at the upper end of what can be detected by CPMG methods. The kinetics thus makes a full chemical shift based structure determination unfeasible. Still, $|\Delta\omega|$ extracted from the CPMG experiments contain valuable information about both the secondary and the tertiary structure of the excited state. Carbonyl chemical shifts are especially sensitive to the secondary structure of the backbone, and the $^{13}C'$ CPMG experiment thus reports on the

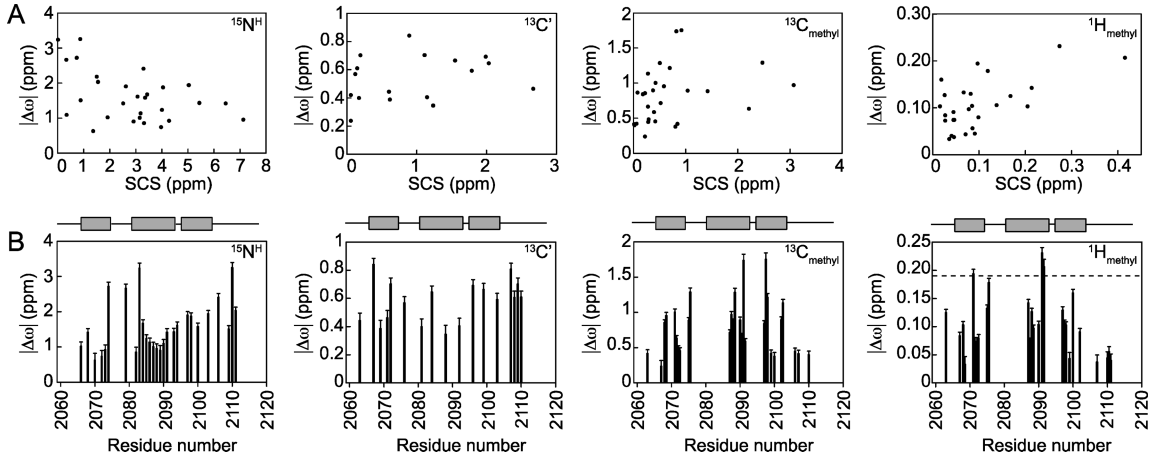


Figure 4. Chemical shift differences from relaxation dispersion measurements on NCBD at 19 °C. (A) Correlation plots of $|\Delta\omega|$ versus the secondary chemical shifts, SCS. (B) $|\Delta\omega|$ versus sequence number. The locations of the three helices are indicated above each panel. The dashed line in the panel with $^{1}H_{methyl}$ indicates 2σ = 0.19 ppm for the variance of chemical shifts from $^{1}H_{methyl}$ not influenced by aromatic groups.

dx.doi.org/10.1021/bi4001062 | *Biochemistry* 2013, 52, 1686−1693

degree of helix formation in the excited state. The $^{13}C'$ chemical shift changes between the ground state and the excited state are much smaller than the secondary chemical shifts of the ground state (Figure 4A), demonstrating that the three helices are preserved in the excited state. Methyl ^{13}C chemical shifts depend on the rotamer distribution of the side chain and thus probe the hydrophobic packing.[7,41−44] The ^{13}C methyl secondary chemical shifts of the unbound state of NCBD are of similar magnitude to those of other folded proteins, demonstrating significant hydrophobic packing in the dominant conformation. The ^{13}C chemical shift differences between the ground state and the excited state are of the same magnitude as the secondary chemical shift, but they are not highly correlated ($r = 0.30$). This suggests that the hydrophobic core of the excited state has specific packing, albeit one that is different from the ground state.

The methyl ^{1}H chemical shift differences derived from the CPMG experiments can potentially be used to probe the tertiary structure of the excited state. In the dominant conformation of free NCBD, the only methyl ^{1}H that have substantial secondary chemical shifts are from V2087 and L2091. This is due to ring current effects from the aromatic ring of F2101 consistent with the major conformation of free NCBD resembling that of the complex with ACTR.[6] The construct of NCBD used here contains only two aromatic residues, F2101 and Y2109. Aromatic ^{13}C $R_{1\rho}$ experiments show significant exchange contribution to the transverse relaxation only for F2101, most likely because Y2109 is highly flexible in both states (Figure 6). The optimized values of k_{ex}

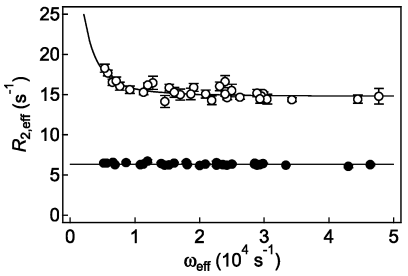


Figure 6. Relaxation dispersion profiles of R_2 as a function of the effective field strength ω_{eff} extracted from off-resonance $R_{1\rho}$ relaxation measurement of $^{1}H^{\delta}$ in the aromatic rings of F2101 (open symbols) and Y2109 (filled symbols). For F2101 the data were fitted to a two-state exchange model. Y2109 does not show significant exchange and fit a constant value. The data were recorded at a static magnetic field of 17.6 T and at 19 °C.

and ϕ_{ex} from the fit of $R_{1\rho}$ data for F2101 are highly correlated and individual values for the two parameters cannot be reliably extracted. Still the exchange contribution is significant for F2101, and we consequently attribute any putative ring current effect observed for methyl groups in the minor state to originate from F2101. Y2109 remains highly dynamic and consequently is unlikely to give strong recurrent shifts. Furthermore, Y2109 does not undergo exchange on the relevant time scale.

The largest contributions to methyl ^{1}H chemical shifts are usually ring current effects from aromatic groups.[45] The methyl chemical shifts can thus potentially be used to probe the tertiary interaction in the excited state; however, it is necessary to estimate how large a chemical shift change needs to be to indicate proximity to an aromatic group. To this end we want to estimate a cutoff for how large a chemical shift change must be to necessarily be an effect of ring currents. To estimate a reasonable threshold, we extracted the chemical shifts of all methyl groups that were more than 6 Å from the nearest aromatic side chain in 19 recent NMR structures from the NorthEast Structural Genomics consortium, resulting in 149 usable methyl chemical shifts. For each of these methyl chemical shifts, the residue average was subtracted to compare different residue types. The standard deviation for methyl ^{1}H chemical shifts not influenced by aromatic rings was estimated to be $\sigma = 0.095$ ppm, and chemical shift differences larger than 2σ most likely indicate the effects of an aromatic ring.

In our methyl ^{1}H CPMG dispersion experiments on free NCBD a few methyl groups have $|\Delta\omega|$ above 2σ, and their chemical shifts are thus likely influenced by F2101 (Figure 5B). The two largest $|\Delta\omega|$ in the methyl ^{1}H CPMG dispersion experiments of 0.23 and 0.21 ppm are observed for $H^{\delta 2}$ and $H^{\delta 1}$ in L2091. These values match the magnitude of the secondary chemical shifts of 0.27 and 0.42 ppm of the same groups. It is unlikely that these methyl groups should move this much further away from their random coil positions, suggesting that the ring current effect dominating the ground state is not present in the excited state. L2091 lies directly under the ring of F2101 in both the ACTR complex and the dominant conformation of free NCBD (Figure 7). The tertiary

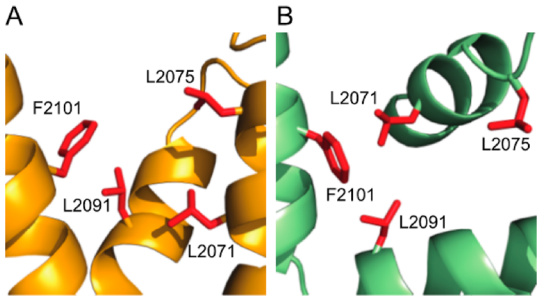


Figure 7. Structural environment of F2101 in NCBD in complex with (A) ACTR (PDB ID: 1KBH) and (B) IRF-3 (PDB ID: 1ZOQ). In both panels, L2071, L2075, and L2091 that show large $|\Delta\omega|$ in ^{1}H methyl CPMG are shown as sticks.

interaction between helices 2 and 3 is thus broken or reconfigured in the excited state. V2087 $H^{\gamma 2}$ that is also near the ring of F2101 in the ground state has the same trend as L2091 even though its $|\Delta\omega|$ is smaller. L2071 $H^{\delta 1}$ with a ground-state methyl ^{1}H chemical shift close to the random coil value has $|\Delta\omega|$ above the 2σ threshold, strongly suggesting this methyl group to be in proximity of F2101 in the excited state and that helices 1 and 3 consequently form direct interactions in the excited state. This is further supported by the large, albeit below the limit, value of $|\Delta\omega|$ for L2075 $H^{\delta 1}$.

DISCUSSION

The exact definition of an IDP has always been vague. NCBD displays clear signs of dynamic averaging in the NMR spectra, but also solid evidence of a dominant conformation with well-defined tertiary structure.[6,12,14] It is thus one of the most folded proteins commonly described as being intrinsically disordered and is frequently referred to as a molten globule. While the dominant conformation resembles that found in the complexes with ACTR,[4] steroid receptor coactivator-1 (SRC-1),[46] and p53,[47] it is unclear which other structures are contained in the

ensemble. Several options can explain the averaging observed by NMR and can conveniently be considered using an energy landscape containing one or more energy minima. The energy landscape could contain only a single energy well centered on the dominant ACTR-like conformation, allowing significant structural fluctuations within this well, but not to conformations with significantly different tertiary contacts. Naganathan and Orozco proposed an alternative energy landscape, where the single energy minimum is wide enough to allow the protein to sample states with significantly different tertiary structure without passing a kinetic barrier.[9] Finally, it is possible that the main energy well is not necessarily wider than that of other folded proteins but that two or more energy wells with folded conformations and similar stabilities exist simultaneously. The latter model resembles fold-switching proteins where the protein can switch between two well-defined conformations.[48] Numerous molecular dynamics simulations of NCBD have been reported, and all suggest multiple coexisting conformations although the specific structures vary.[7−9,17] Because of the simplified force fields used in most of these simulations, however, experimental verification is required.

The experimental approach in this paper is directly aimed at testing the presence of alternative folded conformations. The relaxation dispersion throughout NCBD demonstrates that large structural changes that fit a two-state model will take place on the microseconds time scale. A good fit to the two-state equation does not in itself guarantee an underlying two-state process. In a global cooperative process, all data from a protein needs to fit to the same global parameters for the exchange rate and the population of the excited state. Furthermore, when an external parameter such as the temperature is changed, the kinetic parameters need to change in a physically realistic way. All these conditions are fulfilled by the relaxation dispersion of NCBD, suggesting that there are at least two well-defined energy minima.

The structural similarity between the dominant conformation of free NCBD and the complex with ACTR was previously used to suggest a binding process resembling conformational selection.[6] The pre-existence of a complex-like structure does not guarantee that a process occurs through conformational selection, and the final conclusion must be based on the flux.[49] Pre-existence of the folded conformation in the ligand free ensemble is, however, a prerequisite for the conformational selection, and subsequent induced-fit adjustments after the initial encounter are still likely to occur.[50] In this study, we have demonstrated the existence of a sparsely populated folded conformation in addition to the dominant conformation described previously. It is tempting to speculate that the two conformations in the free ensemble mirror the two different conformations found in complexes, with the dominant conformation resembling the complexes with ACTR, SRC-1, and p53 and the excited state conformation resembling the structure in complex with IRF-3. While we have not been able to determine the structure of the excited state, we have extracted structural characteristics of the excited state. In summary, we found (a) that the excited state has roughly the same helical content as the major state; (b) that the contact between helices 2 and 3 found in the dominant conformation is broken, and a new contact is formed between helices 1 and 3; and (c) that the excited state may expose more hydrophobic surface than the major state. Indeed, NCBD adopts a topology with these characteristics in complex with IRF-3 (Figures 7 and 8). For this topology the methyl groups of L2071 are predicted

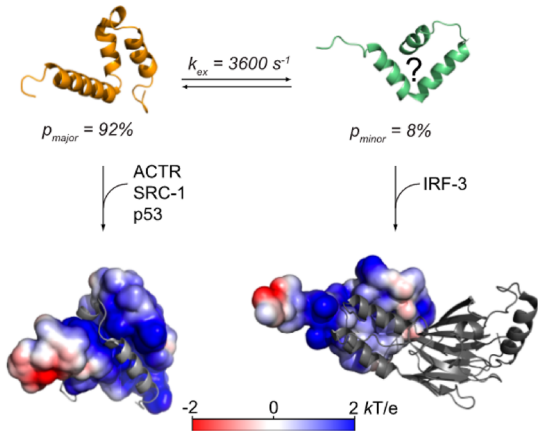


Figure 8. Hypothesis for ligand binding to NCBD. In its highly malleable free state NCBD is in exchange between two folded conformations. As previously characterized by NMR, the structure of NCBD in the major state resembles that in complex with ACTR, SRC-1, and p53. The excited state is represented by the structure of NCBD in complex with IRF-3 although the evidence for this is not conclusive. In the lower part of the figure, NCBD is displayed with the solvent accessible surface area colored according to the electrostatic potential[51] and with the ligands ACTR (in gray to the left) and IRF-3 (in gray to the right).

to experience the largest ring current effect from F2101, exactly as we observe in our methyl ^{1}H CPMG relaxation dispersion experiments. While several other structures can be envisaged with this contact and with the other characteristics of the excited state, it does demonstrate that the excited state has a tertiary arrangement that resembles the IRF-3 bound structure in this key aspect and as a consequence may be mechanistically important for the binding reaction.

The minor state of NCBD that we have characterized here exchanges with the major state at a rate of around 3600 s^{-1}. This is faster than the rate constants recently determined for the binding of NCBD to ACTR, suggesting that the process we observe is not rate limiting for the ligand binding to the major conformation. This does not, however, rule out the possibility that it could be mechanistically important. The observation of a folded, excited state thus prompts the question of whether it serves as a starting point for binding to IRF-3 and other putative ligands that bind NCBD in a similar conformation. In other words, can e.g. IRF-3 and ACTR form encounter complexes with NCBD in both its minor and major conformations, or are they restricted to binding through a conformer that resembles their final bound conformation? The structures of the complexes give us a hint to the answer. NCBD has a high positive charge, and ACTR, SRC-1, and p53 are all highly negatively charged and wrap around NCBD, thus neutralizing its charge. The binding kinetics depend on ionic strength,[50] which suggests that electrostatic interactions are important for formation of the encounter complex. In contrast, in the interaction with IRF-3 the interface is composed of the hydrophobic sides of amphipathic helices for both proteins. There are no obvious charges in IRF-3 that neutralize the charges in NCBD and the positively charged side chains are mainly located on the side facing away from IRF-3. It is thus unlikely that NCBD would form an electrostatic encounter complex with IRF-3, and consequently NCBD needs to expose a hydrophobic patch to initiate the binding reaction. The temperature dependence of the excited state shows that it is

dx.doi.org/10.1021/bi4001062 | *Biochemistry* 2013, 52, 1686−1693

entropically unfavored relative to the ground state, which is expected if it exposes additional hydrophobic groups to the solvent. The excited state may thus be a mechanistic adaptation to allow binding to a biophysically different subset of proteins that present a folded uncharged binding site. This type of complex is so far only exemplified by IRF-3, but future structural studies of NCBD's many interactions will reveal how general this type of interaction is.

In conclusion, our results show that NCBD adopts two distinct folded conformations. This is similar to what is observed for fold-switching proteins, except in this case the population is heavily skewed toward one state. The exchange between the two states is faster than the effective binding rate to ACTR recently determined by stopped flow[50] and is thus not likely to be rate limiting for the binding process. The low energy barrier to interconversion efficiently provides a structural repertoire, where different ligands can initiate the binding reaction from conformers with complementary structures. Our current hypothesis is outlined in Figure 8. The exact importance of the preformed structures for the binding process would require a detailed kinetic analysis of the binding of ACTR and IRF-3 and especially studying the kinetic effects of shifting the equilibrium between the two states. Furthermore, a high-resolution model of the minor state would provide insights into how similar the excited state is to the IRF-3 bound structure and may clarify how much of the structural change takes part before and after the initial association reaction. In other words, knowing the structure of the excited state would shed light on the relative importance of induced fit and conformational selection. The current work lay the foundation for such experiments.

AUTHOR INFORMATION

Corresponding Author

*Phone +45-35322029; e-mail kaare.teilum@bio.ku.dk.

Present Address

[†]Department of Chemistry, University of Cambridge, Cambridge, UK.

Funding

This work was supported by the EliteForsk program (to M.K.), by the Lundbeck Foundation grant no. R77-A7094 (to K.T.), and by the Carlsberg Foundation grant no. 2010-01-0317.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Peter E. Wright (The Scripps Research Institute) for sharing the coexpression plasmid for NCBD, Flemming M. Poulsen and Birthe B. Kragelund for helpful comments, and Mikael Akke for access to the Varian VnmrS 500 MHz spectrometer at Lund University.

ABBREVIATIONS

ACTR, activator for thyroid hormone and retinoid receptors; CBP, CREB binding protein; CPMG, Carr−Purcell−Meiboom−Gill; Glc, glucose; HSQC, heteronuclear single quantum coherence; IDP, intrinsically disordered protein; IRF-3, interferon regulatory factor 3; NCBD, the nuclear coactivator binding domain; SCS, secondary chemical shift; SRC-1, steroid receptor coactivator-1.

REFERENCES

(1) Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2008) Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys. 37*, 215−246.

(2) Wright, P. E., and Dyson, H. J. (2009) Linking folding and binding. *Curr. Opin. Struct. Biol. 19*, 31−38.

(3) Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol. 6*, 197−208.

(4) Demarest, S. J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H. J., Evans, R. M., and Wright, P. E. (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature 415*, 549−553.

(5) Qin, B. Y., Liu, C., Srinath, H., Lam, S. S., Correia, J. J., Derynck, R., and Lin, K. (2005) Crystal structure of IRF-3 in complex with CBP. *Structure 13*, 1269−1277.

(6) Kjaergaard, M., Teilum, K., and Poulsen, F. M. (2010) Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. U. S. A. 107*, 12535−12540.

(7) Zhang, W., Ganguly, D., and Chen, J. (2012) Residual structures, conformational fluctuations, and electrostatic interactions in the synergistic folding of two intrinsically disordered proteins. *PLoS Comput. Biol. 8*, e1002353.

(8) Burger, V. M., Ramanathan, A., Savol, A. J., Stanley, C. B., Agarwal, P. K., and Chennubhotla, C. S. (2012) Quasi-anharmonic analysis reveals intermediate States in the nuclear co-activator receptor binding domain ensemble. *Pac. Symp. Biocomput.*, 70−81.

(9) Naganathan, A. N., and Orozco, M. (2011) The native ensemble and folding of a protein molten-globule: functional consequence of downhill folding. *J. Am. Chem. Soc. 133*, 12154−12161.

(10) Ganguly, D., Zhang, W., and Chen, J. (2012) Synergistic folding of two intrinsically disordered proteins: searching for conformational selection. *Mol. BioSyst. 8*, 198−209.

(11) Demarest, S. J., Deechongkit, S., Dyson, H. J., Evans, R. M., and Wright, P. E. (2004) Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. *Protein Sci. 13*, 203−210.

(12) Ebert, M.-O., Bae, S.-H., Dyson, H. J., and Wright, P. E. (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.

(13) Keppel, T. R., Howard, B. A., and Weis, D. D. (2011) Mapping unstructured regions and synergistic folding in intrinsically disordered proteins with amide H/D exchange mass spectrometry. *Biochemistry 50*, 8722−8732.

(14) Kjaergaard, M., Poulsen, F. M., and Teilum, K. (2012) Is a malleable protein necessarily highly dynamic? The hydrophobic core of the nuclear coactivator binding domain is well ordered. *Biophys. J. 102*, 1627−1635.

(15) Dolgikh, D. A., Kolomiets, A. P., Bolotina, I. A., and Ptitsyn, O. B. (1984) 'Molten-globule' state accumulates in carbonic anhydrase folding. *FEBS Lett. 165*, 88−92.

(16) Japrung, D., Dogan, J., Freedman, K. J., Nadzeyka, A., Bauerdick, S., Albrecht, T., Kim, M. J., Jemth, P., and Edel, J. B. (2013) Single molecule studies of intrinsically disordered proteins using solid-state nanopores. *Anal. Chem.*, DOI: 10.1021/ac3035025.

(17) Knott, M., and Best, R. B. (2012) A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: evidence from molecular simulations. *PLoS Comput. Biol. 8*, e1002605.

(18) Baldwin, A. J., and Kay, L. E. (2009) NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol. 5*, 808−814.

(19) Lundström, P., Teilum, K., Carstensen, T., Bezsonova, I., Wiesner, S., Hansen, D. F., Religa, T. L., Akke, M., and Kay, L. E. (2007) Fractional 13C enrichment of isolated carbons using [1-13C]- or [2-13C]-glucose facilitates the accurate measurement of dynamics at backbone Calpha and side-chain methyl positions in proteins. *J. Biomol. NMR 38*, 199−212.

(20) Teilum, K., Brath, U., Lundström, P., and Akke, M. (2006) Biosynthetic 13C labeling of aromatic side chains in proteins for NMR relaxation measurements. *J. Am. Chem. Soc. 128*, 2506−2507.
(21) Otten, R., Villali, J., Kern, D., and Mulder, F. A. (2010) Probing microsecond time scale dynamics in proteins by methyl (1)H Carr-Purcell-Meiboom-Gill relaxation dispersion NMR measurements. Application to activation of the signaling protein NtrC(r). *J. Am. Chem. Soc. 132*, 17004−17014.
(22) Tollinger, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D., and Kay, L. E. (2001) Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc. 123*, 11341−11352.
(23) Lundström, P., Hansen, D. F., and Kay, L. E. (2008) Measurement of carbonyl chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy: comparison between uniformly and selectively (13)C labeled samples. *J. Biomol. NMR 42*, 35−47.
(24) Mulder, F. A., Hon, B., Mittermaier, A., Dahlquist, F. W., and Kay, L. E. (2002) Slow internal dynamics in proteins: application of NMR relaxation dispersion spectroscopy to methyl groups in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc. 124*, 1443−1451.
(25) Mulder, F. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Measurement of slow (micros-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc. 123*, 967−975.
(26) Mulder, F. A., De Graaf, R. A., Kaptein, R., and Boelens, R. (1998) An off-resonance rotating frame relaxation experiment for the investigation of macromolecular dynamics using adiabatic rotations. *J. Magn. Reson. 131*, 351−357.
(27) Igumenova, T. I., and Palmer, A. G. (2006) Off-resonance TROSY-selected R1rho experiment with improved sensitivity for medium- and high-molecular-weight proteins. *J. Am. Chem. Soc. 128*, 8110−8111.
(28) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.
(29) Carver, J., and Richards, R. (1972) A general two-site solution for the chemical exchange produced dependence of T2 upon the carr-Purcell pulse separation. *J. Magn. Reson. 6*, 89−105.
(30) Kjaergaard, M., Brander, S., and Poulsen, F. M. (2011) Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J. Biomol. NMR*, 139−149.
(31) Kjaergaard, M., and Poulsen, F. M. (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.
(32) Kjaergaard, M., Iešmantavičius, V., and Poulsen, F. M. (2011) The interplay between transient α-helix formation and side chain rotamer distributions in disordered proteins probed by methyl chemical shifts. *Protein Sci. 20*, 2023−2034.
(33) Kovrigin, E. L., Kempf, J. G., Grey, M. J., and Loria, J. P. (2006) Faithful estimation of dynamics parameters from CPMG relaxation dispersion measurements. *J. Magn. Reson. 180*, 93−104.
(34) Teilum, K., Smith, M. H., Schulz, E., Christensen, L. C., Solomentsev, G., Oliveberg, M., and Akke, M. (2009) Transient structural distortion of metal-free Cu/Zn superoxide dismutase triggers aberrant oligomerization. *Proc. Natl. Acad. Sci. U. S. A. 106*, 18273−18278.
(35) Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R., and Kay, L. E. (2010) A transient and low-populated protein-folding intermediate at atomic resolution. *Science 329*, 1312−1316.
(36) Bouvignies, G., Vallurupalli, P., Hansen, D. F., Correia, B. E., Lange, O., Bah, A., Vernon, R. M., Dahlquist, F. W., Baker, D., and Kay, L. E. (2011) Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature 477*, 111−114.
(37) Neudecker, P., Robustelli, P., Cavalli, A., Walsh, P., Lundström, P., Zarrine-Afsar, A., Sharpe, S., Vendruscolo, M., and Kay, L. E. (2012) Structure of an intermediate state in protein folding and aggregation. *Science 336*, 362−366.
(38) Vallurupalli, P., Hansen, D. F., Stollar, E., Meirovitch, E., and Kay, L. E. (2007) Measurement of bond vector orientations in invisible excited states of proteins. *Proc. Natl. Acad. Sci. U. S. A. 104*, 18473−18477.
(39) Baldwin, A. J., and Kay, L. E. (2012) Measurement of the signs of methyl 13C chemical shift differences between interconverting ground and excited protein states by R(1ρ): an application to αB-Crystallin. *J. Biomol. NMR 53*, 1−12.
(40) Auer, R., Hansen, D. F., Neudecker, P., Korzhnev, D. M., Muhandiram, D. R., Konrat, R., and Kay, L. E. (2010) Measurement of signs of chemical shift differences between ground and excited protein states: a comparison between H(S/M)QC and R1rho methods. *J. Biomol. NMR 46*, 205−216.
(41) Mulder, F. A. (2009) Leucine side-chain conformation and dynamics in proteins from 13C NMR chemical shifts. *ChemBioChem 10*, 1477−1479.
(42) Hansen, D. F., Neudecker, P., and Kay, L. E. (2010) Determination of isoleucine side-chain conformations in ground and excited states of proteins from chemical shifts. *J. Am. Chem. Soc. 132*, 7589−7591.
(43) Hansen, D. F., Neudecker, P., Vallurupalli, P., Mulder, F. A., and Kay, L. E. (2010) Determination of Leu side-chain conformations in excited protein states by NMR relaxation dispersion. *J. Am. Chem. Soc. 132*, 42−43.
(44) London, R. E., Wingad, B. D., and Mueller, G. A. (2008) Dependence of amino acid side chain 13C shifts on dihedral angle: application to conformational analysis. *J. Am. Chem. Soc. 130*, 11097−11105.
(45) Sahakyan, A. B., Vranken, W. F., Cavalli, A., and Vendruscolo, M. (2011) Structure-based prediction of methyl chemical shifts in proteins. *J. Biomol. NMR 50*, 331−346.
(46) Waters, L., Yue, B., Veverka, V., Renshaw, P., Bramham, J., Matsuda, S., Frenkiel, T., Kelly, G., Muskett, F., Carr, M., and Heery, D. M. (2006) Structural diversity in p160/CREB-binding protein coactivator complexes. *J. Biol. Chem. 281*, 14787−14795.
(47) Lee, C. W., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2010) Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry 49*, 9964−9971.
(48) Bryan, P. N., and Orban, J. (2010) Proteins that switch folds. *Curr. Opin. Struct. Biol. 20*, 482−488.
(49) Hammes, G. G., Chang, Y.-C., and Oas, T. G. (2009) Conformational selection or induced fit: a flux description of reaction mechanism. *Proc. Natl. Acad. Sci. U. S. A. 106*, 13737−13741.
(50) Dogan, J., Schmidt, T., Mu, X., Engström, A., and Jemth, P. (2012) Fast association and slow transitions in the interaction between two intrinsically disordered protein domains. *J. Biol. Chem. 287*, 34316−34324.
(51) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A. 98*, 10037−10041.