# Is a Malleable Protein Necessarily Highly Dynamic? The Hydrophobic Core of the Nuclear Coactivator Binding Domain Is Well Ordered

Magnus Kjaergaard, Flemming M. Poulsen, and Kaare Teilum*
Department of Biology, University of Copenhagen, Copenhagen, Denmark

ABSTRACT The nuclear coactivator binding domain of CREB binding protein folds into remarkably different structures in complex with different ligands. To understand the mechanism of the structural adaptability in the nuclear coactivator binding domain (NCBD), we have compared the dynamics of the hydrophobic core of NCBD in the ligand-free state and in a well-folded complex with the ligand activator for thyroid hormone and retinoid receptors using multiple NMR methods including methyl chemical shifts, coupling constants, and methyl order parameters. From all NMR measures, the aliphatic side chains in the hydrophobic core are slightly more dynamic in the free protein than in the complex, but have mobility comparable to the hydrophobic cores of average folded proteins. Urea titration monitored by NMR reveals that all parts of the protein, including the side-chain packing in the hydrophobic core, denatures in a single cooperative process. The molten globule characteristics of NCBD are thus restricted to a slowly fluctuating tertiary structure. Consequently, the conformational plasticity of the protein is most likely related to its low overall stability rather than an intrinsically flexible protein structure. The well-defined structure supports a model of molecular recognition dominated by conformational selection, whereas only minor structural adjustments are necessary after the association.

## INTRODUCTION

Flexibility plays an important role in many protein-protein interactions (1). An especially extreme case of flexibility is the intrinsically disordered proteins (IDPs) that lack a folded native structure (2–4). IDPs frequently interact with many other proteins, where they often fold upon binding to their partner proteins (5). The flexibility of the native state of these proteins is believed to be key to their binding promiscuity, as they can fold into different structures suited to bind structurally different ligands. An alternative mechanism to achieve binding to structurally different proteins is the fold-switching mechanism. Fold-switching proteins have at least two well-defined conformations with comparable stability that are both significantly populated in solution (6–8). Consequently, different ligands can bind to either of the folded forms in a conformational selection-like reaction. Each of these conformations may be well ordered and not extraordinary flexible as long as they have a low thermodynamic stability (8).

The nuclear coactivator binding domain (NCBD) of CREB binding protein (CBP) is a particularly interesting example of plasticity in ligand recognition. NCBD forms well-defined complexes with a number of ligands, where it folds into remarkably different tertiary structures (9–12), as illustrated in Fig. 1. There has been much interest in understanding the basis of this structural plasticity (13–16) and NCBD has emerged as a model system for studying the mechanism of coupled folding and binding (14,17). Three hypotheses have emerged so far that are not necessarily mutually exclusive. The adaptability has either been explained by the free protein being a molten globule (9,13,18), having a low overall structural stability (14), or being a downhill folder (15). NCBD has most of the characteristics typical for a molten globule: 8-anilinonaphthalene-1-sulphonic acid binding, broad urea denaturation profile, the absence of an excess heat capacity peak in DSC, and exchange broadening of the NMR spectra (18). The molten globule state is frequently observed as a partially folded intermediate of proteins under weakly denaturing conditions (19,20). More recently, a number of proteins have been discovered that behave as molten globules under native-like conditions, which makes them the most folded members of the family of IDPs (2). In the ligand-free state, secondary chemical shifts show that the three helices of NCBD are almost fully formed (13,14) and therefore the protein have a high secondary structure content. All these observations are characteristic of a molten globule. The dynamics of the ligand-free state were examined using $^{15}$N relaxation techniques, which suggest that the backbone experiences a uniform degree of dynamics that is comparable to that of the most folded regions of the apomyoglobin molten globule (13,21), although the free protein seems to experience exchange broadening. Despite the dynamic nature of NCBD, the structural ensemble is dominated by a tertiary fold that resembles the conformations in complex with the ligands activator for thyroid hormone and retinoid receptors (ACTR), steroid receptor coactivator-1, and p53 (14). The fold appears to have well-defined tertiary

*Submitted November 21, 2011, and accepted for publication February 6, 2012.*

This article is dedicated to Flemming M. Poulsen, who passed away on November 9, 2011.

*Correspondence: kaare.teilum@bio.ku.dk

Magnus Kjaergaard's present address is Department of Chemistry, University of Cambridge, Cambridge, UK.

Editor: Josh Wand.



0006-3495/12/04/1627/9 $2.00

doi: 10.1016/j.bpj.2012.02.014

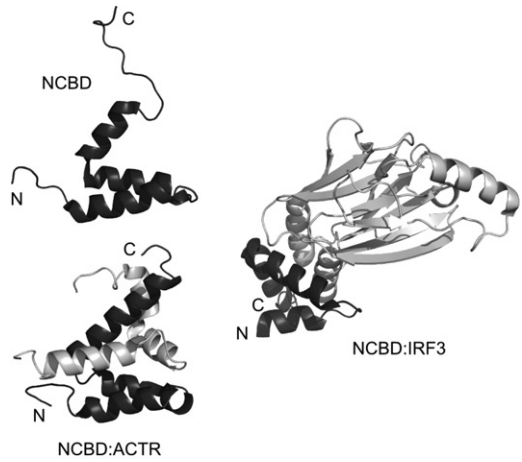


FIGURE 1 Representative structures of NCBD. The dominating structure of the ligand-free NCBD (2KKJ) (14), the complex with ACTR (1KBH) (9), and the complex with IRF-3 (1ZOQ) (10). The NCBD structures are shown in black, aligned along helix 1, and display remarkably different conformations in the two complexes.

interactions and to fold in a cooperative manner even though it is only marginally stable (14), which led to the suggestion that the low stability might be important for the binding plasticity. The third explanation of NCBDs plasticity is that it is a downhill folder. This conclusion was based on its broad denaturation profile and coarse-grained molecular simulations (15). According to these simulations, the barrierless nature of NCBD extends to the native ensemble of NCBD, where it samples highly diverse structures without passing significant energy barriers (15). The absence of energy barriers allows NCBD to sample a continuum of conformations that the protein's diverse ligands can bind to.

To clarify what role the intrinsic flexibility of NCBD plays for its ability to adopt different conformations in complex with ligands, we have characterized the dynamics of the methyl groups in NCBD by NMR relaxation, and the unfolding process of NCBD by following chemical shifts in a urea titration. Collectively, the results show that NCBD has a well-ordered core that is slightly more dynamic in the free state than in complex with a ligand, and that all of NCBD cooperatively loses its structure during unfolding. Despite having many of the characteristics of a molten globule, NCBD thus behaves much as a fully compact protein with a hydrophobic core that fluctuates slowly. This suggests that the conformation of NCBD observed both in the free state and in complex with ligand indeed represent a minimum in conformational space surrounded by energy barriers.

## MATERIALS AND METHODS

### Protein preparation

The plasmid for coexpression of the NCBD (UniProt: P45481, residues 2059–2117) of murine CBP and the activation domain of ACTR (UniProt: Q9Y6Q9, residues 1018–1088) was a gift from Peter E. Wright (The Scripps Research Institute, San Diego, CA) and was described previously (9). Isotope labeling was achieved by growing the cells in M9 medium with 4 g/L $^{13}$C glucose and 1.5 g/L $^{15}$N ammonium sulfate as the only carbon and nitrogen sources. Four different labeling schemes were used: Uniform $^{15}$N labeling, uniform $^{13}$C and $^{15}$N labeling, 10% $^{13}$C and uniform $^{15}$N labeling [10% $^{13}$C], and uniform $^{13}$C and $^{15}$N labeling combined with 50% random $^{2}$H labeling [50% $^{2}$H]. Random $^{2}$H labeling was achieved by growing the bacteria in 50% $^{2}H_2O$. Protein purification was performed as described previously for ACTR (22) except that a FastFlow SP column was used for the ion exchange step.

### NMR spectroscopy

Lyophilized NCBD was dissolved into 20 mM phosphate in 10% $D_2O$ and adjusted to pH 6.5. Relaxation experiments were recorded at 31°C on Varian VnmrS 500 MHz and Varian Inova 750 MHz spectrometers equipped with conventional triple resonance probes. Chemical shift and coupling constant measurements were recorded on a Varian Inova 800 MHz spectrometer equipped with a cold probe. NMR spectra were processed using nmrPipe (23) and analyzed in CCPNMR analysis (24). Constant time $^{13}$C-HSQC spectra were recorded on samples containing 1 mM [10% $^{13}$C] NCBD, and 1 mM [10% $^{13}$C] NCBD complexed with 1.5 mM unlabeled ACTR. LRCC (25) spectra were recorded on a sample containing 0.7 mM uniformly labeled NCBD and 0.7 mM uniformly labeled NCBD in complex with 1.2 mM unlabeled ACTR.

Deuterium relaxation experiments were recorded on a sample containing 1.3 mM [50% $^{2}$H]-NCBD and a sample containing 1.3 mM [50% $^{2}$H]-NCBD in complex with 1.8 mM ACTR. Relaxation times of 0.2, 1.3, 2.8, 4.4, 6.2, 8.4, 10.9*, and 15.1 ms were used for the $T_{1\rho}$ experiment and relaxation times of 0.05, 2.8, 4, 10, 16*, 22, 35, and 40 ms were used for measurement of the $D_z$, $3D_z^2$-2, $D_+^2$, and $D_+D_z + D_zD_+$ coherences as described (26,27). * denotes points determined in triplicate. Experiments were recorded with 64–76 increments in the indirect dimension, a recycle delay of 2 s and 16 transients collected for each increment. For both free NCBD and NCBD in complex with ACTR the five deuterium relaxation experiments were measured at 750 MHz. For free NCBD, $D_z$, $3D_z^2$-2, $D_+^2$ were also measured at 500 MHz. $^{15}$N relaxation experiments were recorded on a sample containing 1.3 mM $^{15}$N-NCBD and a sample containing 1.3 mM $^{15}$N-NCBD in complex with 1.8 mM ACTR at 750 MHz using previous published pulse sequences (28). All relaxation data were analyzed in CCPNMR analysis. The change in peak intensities were fit to monoexponential decays and errors on the optimized parameters were determined by bootstrap analyses.

From the five relaxation rates measured at 750 MHz three spectral densities, $J(0)$, $J(\omega_D)$, $J(2\omega_D)$, were extracted by optimizing their values in the expressions describing the relations between the relaxation rates and the spectral densities using a quadrupolar coupling constant of 167 kHz (26). For free NCBD $J(\omega_D)$ and $J(2\omega_D)$ were also determined at 500 MHz. For NCBD in complex with ACTR the correlation time, $\tau_r$, for the overall rotational diffusion was estimated from the $^{15}$N $R_1$ and $R_2$ relaxation rates by fitting their trimmed mean to the following expression as described by Fushman et al. (29):

$$\tau_r = \frac{\sqrt{6R_2/R_1 - 7}}{2\omega_N}.$$

For NCBD in complex with ACTR three spectral densities for each methyl group were fit to the Lipari-Szabo two-parameter model (LS2) for the spectral densities:

$$J(\omega) = \frac{S^2\tau_r}{1 + (\omega\tau_r)^2} + \frac{(1 - S^2)\tau}{1 + (\omega\tau)^2},$$

where $1/\tau = 1/\tau_r + 1/\tau_f$. $\tau_r$ is the correlation time for the overall rotational diffusion of the complex and $\tau_f$ is the correlation time for the fast (ps)

internal motion of the methyl group. $\tau_r$ was kept constant in the fit of $S$ and $\tau_f$. This model fit all data well and the current data set does not warrant the use of a more complex model. The order parameter describing the fluctuation of the average methyl axis is defined as $S^2_{\text{axis}} = 9\ S^2$.

As the backbone resonances of free NCBD are significantly broadened by exchange (13) a reliable correlation time for the rotational diffusion cannot be estimated from $^{15}$N relaxation rates. Therefore, the five spectral densities were fit to a model with an individual efficient correlation time for each methyl group, i.e., a three-parameter model (LS3):

$$J(\omega) = \frac{S^2\tau_C^{eff}}{1+\left(\omega\tau_C^{eff}\right)^2} + \frac{(1-S^2)\tau}{1+(\omega\tau)^2},$$

where $1/\tau = 1/\tau_C{}^{eff} + 1/\tau_f$ . $\tau_C{}^{eff}$ is the local efficient correlation time including both the overall rotational diffusion and internal motions on the same (ns) timescale. $\tau_f$ is the correlation time for the fast (ps) internal motion of the methyl group. Again the axial order parameter is defined as $S^2_{\text{axis}} = 9\ S^2$.

Error bars on plots of spectral densities and order parameters indicate the 95% confidence intervals determined from the distribution of the parameters in 100 Monte Carlo simulations. The average order parameters for methyl groups derived from Carbonell and del Sol (30) are $A_\beta = 0.81$, $I_{\gamma 2} = 0.78$, $I_\delta = 0.60$, $L_\delta = 0.58$, $M_\varepsilon = 0.35$, $T_\gamma = 0.74$, and $V_\gamma = 0.70$.

## Urea denaturation

The denaturation of NCBD was followed by NMR using a urea titration at 31°C. Starting from a uniformly labeled 1 mM sample of NCBD, the urea concentration was increased in 0.5 M steps by addition from a stock containing 10 M urea but similar in other respects. At each denaturant concentration the following experiments were acquired: $^{15}$N HSQC, constant time $^{13}$C HSQC, 2D HNCO, and 2D HNCOCA. The chemical shifts as a function of urea concentration was fitted to a two-state transition with sloping baselines using Igor Pro (WaveMetrics, Lake Oswego, OR). The chemical shifts were fitted to a two-state equilibrium by considering either each curve individually, all backbone chemical shifts from each residue together or three clusters corresponding to the three helices in free NCBD. Methyl chemical shifts are left out of the global fits as the posttransition baseline typically is poorly defined due to spectral overlap. In the global fits the *m*-values and denaturation midpoints are treated as global parameters and errors on the parameters are determined from the asymptotic standard errors.

# RESULTS

## Stereospecific assignment of methyl groups

Meaningful comparison of the methyl groups in the ligand-free and -bound states of the NCBD requires stereospecific assignments, which can be determined from the sign of the methyl crosspeaks in a $^{13}$C CT-HSQC spectrum recorded on a sample containing 10% $^{13}$C labeling as described previously (31). $^{13}$C CT-HSQC spectra were recorded on ligand-free [10% $^{13}$C]-NCBD and [10% $^{13}$C]-NCBD in complex with the activation domain of ACTR (Fig. 2). In cases where the methyl peaks are overlapped in the $^{13}$C CT-HSQC spectrum, the chemical shifts were determined using a C(CO)NH spectrum recorded on a uniformly $^{13}$C-labeled sample (32), which correlates the $^{13}$C chemical shifts to the well-dispersed amide protons. In these cases, the stereospecific assignments are based on the geminal neighbor in the [10% $^{13}$C]-NCBD CT-HSQC. Using this method, stereospecific assignments were obtained for all geminal methyl groups of free NCBD and in complex with ACTR.

## $\chi$-Angle dynamics from chemical shifts

Aliphatic side chains primarily occupy the three possible staggered conformations for each $\chi$-angle and may populate either a single rotamer or a distribution of staggered conformations. $^{13}$C chemical shifts report on the rotamer distribution of the protein side chain due to the $\gamma$-*gauche* effect (33,34). The $\gamma$-*gauche* effect decreases the chemical shifts

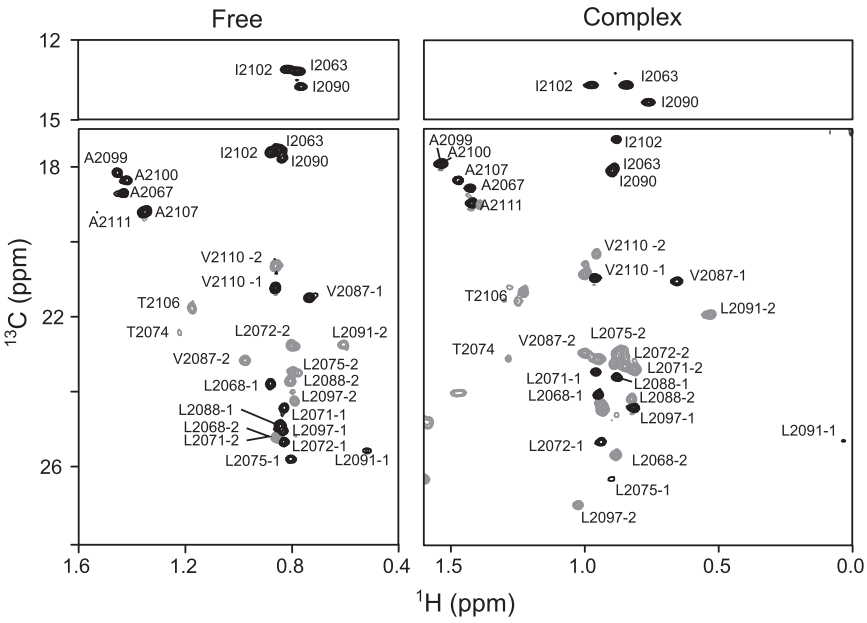


FIGURE 2 Methyl region of the $^{13}$C constant time HSQC spectra of 10%-$^{13}$C NCBD in the free form (*left*) and in complex with ACTR (*right*). Stereospecific assignments are obtained from the signs of the peaks that are shown as black (*positive*) and gray (*negative*). The Ile-$\delta$1 peaks (*top panel*) are recorded on a uniformly labeled sample.

of nuclei that are *gauche* to another heavy atom by ~5 ppm. This effect has been particularly useful for determining the side-chain conformations of leucine (22,35,36), isoleucine (37), and valine residues (38) from their chemical shifts. Leucine and isoleucine side chains mainly occupy two $\chi^2$-angles, which simplify the interpretation of the $^{13}C$ methyl chemical shifts. For leucine, the populations of the staggered conformations of the $\chi^2$-angle depend linearly on the chemical shift difference between the two methyl groups, $\delta(C^{\delta1})$ - $\delta(C^{\delta2})$ (35), and for isoleucine it depends directly on $\delta(C^{\delta1})$ (37). Fig. 3 *A* shows the populations of the *trans* rotamer of the leucine and isoleucine residues of unbound NCBD and in complex with NCBD. Valine residues populate all three $\chi^1$-angles and are thus more difficult to interpret quantitatively. Instead, we report the chemical shift difference between the two geminal methyl groups of each valine residue, which may serve as a qualitative probe of the side-chain dynamics (Fig. 3 *B*). In both the free state and in the complex, NCBD does not have any leucine and isoleucine side chains with unique rotamers. Disordered proteins that do not form hydrophobic cores have side-chain distributions close to their random coil values (22), which is signified by the horizontal lines in Fig. 3, *A* and *B*. Depending on the dynamics, the side-chain rotamer distributions in a mobile hydrophobic core would approach the values of the random coil. In both states, NCBD has methyl rotamer distributions distinct from random coil. Comparison of the free and the ACTR-bound states of NCBD reveals that 9 out of 12 methyl-bearing side chains are closer to the random coil values in the free form than in the complex. This suggests that the side chains in the hydrophobic core undergo slightly more rotameric averaging in the unbound state.

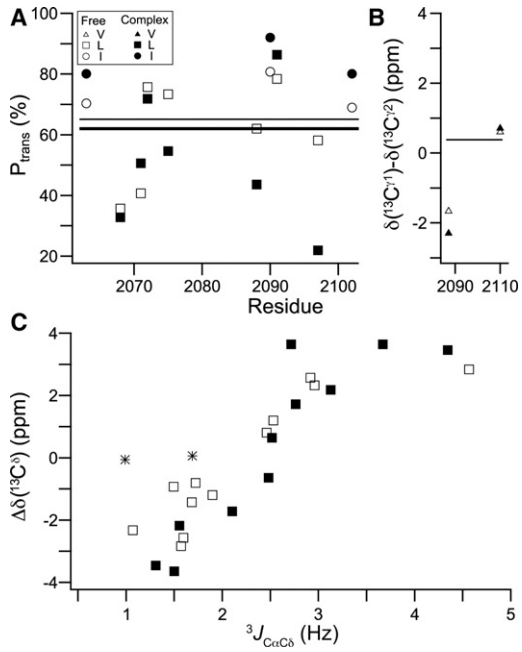


FIGURE 3 Comparison of the rotamer distribution of aliphatic side chains in free NCBD and in the complex with ACTR. (*A*) Populations of the *trans* conformer of the $\chi^2$-angle for isoleucine and leucine residues derived from $^{13}C$ methyl chemical shifts. Horizontal lines correspond to the random coil values of Leu (*thin*) and Ile (*bold*) adopted from (22). (*B*) $^{13}C$ chemical shift differences between the geminal methyl groups of valine side chains. (*C*) Correlation between rotamer distributions determined from chemical shifts and $^3J_{C\alpha C\beta}$ couplings. The * corresponds to residue L2071 in the free state that predominantly populated the rare *gauche*-rotamer and thus is an outlier.

## $\chi$-Angle dynamics from coupling constants

The rotamer distribution of the side chains can also be measured using coupling constants. For aliphatic side chains, the $^3J_{C\delta C\alpha}$ is particularly useful as it probes the $\chi^2$-angle and can thus be directly compared to the values obtained from methyl chemical shifts of leucine and isoleucine residues. $^3J_{C\delta C\alpha}$ values of ~4 Hz suggest that the $C^{\delta}$ is *trans* to the $C^{\alpha}$ and values of <1 Hz suggest that the $C^{\delta}$ is *gauche* relative to the $C^{\alpha}$. The coupling constants were determined using the long-range C-C experiment as described previously (25). The overlap particularly in the leucine methyl region makes this approach challenging. Still, the $^3J_{C\delta C\alpha}$ could be determined for 16 methyl groups in the free state and for 15 in the complex with ACTR. Because the $^3J_{C\delta C\alpha}$ coupling constant and the methyl chemical shifts report on the same dihedral angle, their internal consistency can be checked by correlating the chemical shift difference to the coupling constants as described previously (35). The correlation in Fig. 3 *C* shows that the coupling constants generally agree with the conclusion from the chemical shifts, except for residue L2071. Both methyl chemical shifts of this residue are low suggesting that they are both *gauche* to the $C_{\alpha}$ and thus populate the rare *gauche*-conformation. As the coupling constants report on the same information as the chemical shifts and the latter can be measured more precisely, we did not interpret the coupling constants beyond noting that they support the conclusions from the methyl chemical shifts.

## Deuterium relaxation

The ps-ns timescale motions of methyl groups can be quantitatively probed using the deuterium relaxation rates for the $CH_2D$ isotopomer of methyl groups (26,39). The relaxation rates of five coherences were recorded on both NCBD in the unbound state and in complex with ACTR. Relaxation rates were determined by fitting of a single exponential decay to the peak intensity as a function of relaxation time. 21 and 24 methyl groups yielded signals of sufficient quality to allow further analysis of the free and ACTR-bound states, respectively (Fig. 4 and Fig. S1 and Tables S1 and S2 in the Supporting Material). To determine the overall correlation time of NCBD, $R_1$ and $R_2$ relaxation rates were determined for $^{15}N$ nuclei in the unbound state and in complex with

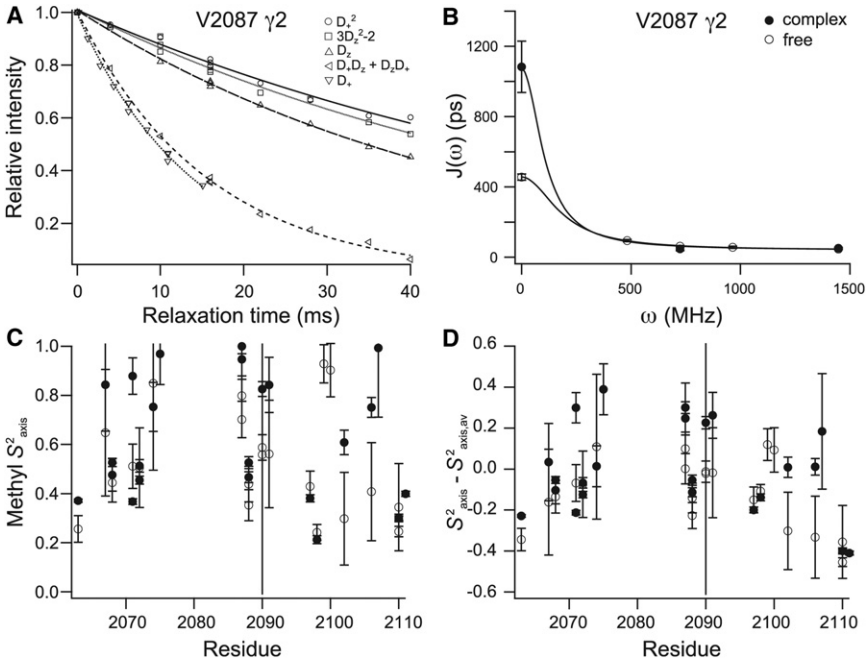


FIGURE 4 Hydrophobic core dynamics of NCBD probed by deuterium relaxation. (*A*) Representative relaxation decays from the V2087 γ2 methyl group of the five deuterium coherences recorded on free NCBD at 750 MHz. Additionally, three relaxation decays were measured for the free form at 500 MHz. (*B*) Spectral densities were extracted from the relaxation rates and two parameters (ACTR complex) and three parameters (free NCBD) model-free spectral density functions were fitted to extract methyl order parameters. (*C*) Comparison of methyl order parameters for free NCBD and the complex with ACTR. (*D*) Methyl order parameters relative to average order parameters for each residue type determined from a library of 18 well-folded proteins.

ACTR (Fig. S2). In the complex, the $^{15}$N relaxation rates are similar for most residues and accordingly a correlation time of 9.3 ns is estimated based on the ratio of the truncated averaged values of the $R_1$ and $R_2$ rates as described previously (29). This value is consistent with a complex with a molecular mass of 14.5 kDa. To interpret the relaxation rates in terms of order parameters for the complex, the spectral densities $J(0)$, $J(\omega_D)$, and $J(2\omega_D)$ were extracted from the measured relaxation rates and fitted to the two-parameter spectral density function (LS2) proposed by Lipari and Szabo (40,41) by constraining the rotational correlation time to the value obtained from $^{15}$N relaxation data (Fig. 4, *B–C*, and Table S3). All data sets are well fit by the LS2 model and the fit fall within the error bars of the spectral densities. With the current data set, fits to more complex models for the spectral densities would lead to overfitting of the data. If some local ns motions of the methyl groups are present, it would mean that $\tau_r$ would overestimate the correlation time for the ns motions, which would lead to underestimation of $S^2_{axis}$. In such a case the reported order parameters are lower limits (27). However, as $J(\omega_D) \approx J(2\omega_D)$ within error for all methyl groups the correlation times for the ns motions cannot be much shorter than the 9.3 ns determined for $\tau_r$. Free NCBD has highly variable $R_2$ rates (Fig. S2) and at low temperatures most signals are broadened beyond detection suggesting μs-ms dynamics throughout the sequence, as also reported previously (13). Consequently, the correlation time of free NCBD cannot be reliably determined from $^{15}$N relaxation rates. Therefore, it was necessary to record deuterium relaxation data at two static magnetic fields to sample the spectral density function at more values. Five spectral densities were extracted from the relaxation rates and fitted to a three-parameter spectral density function, LS3 (27), where the local ns dynamics and the overall rotation of the molecule are combined into an effective correlation time for each methyl group. All methyl groups of free NCBD fit LS3 with $\langle \tau_C^{eff} \rangle = 4.8 \pm 0.8$ ns (Table S4). The average $\tau_C^{eff}$ is in the expected range for the rotational correlation time for a protein with the size of NCBD (6.6 kDa). The longest $\tau_C^{eff}$ from the fits (up to 6.3 ns) are longer than expected for the overall tumbling of NCBD and may thus be overestimated. For these residues the fitted $S^2_{axis}$ are thus lower limits as discussed previously.

The order parameters for unbound NCBD and NCBD in complex with ACTR are compared in Fig. 4 *C*. The methyl groups near the C-terminus are more flexible in the free form than in the complex, which is in agreement with the conclusion from $^{15}$N backbone experiments (13). This region forms a part of helix 3 in the complex (9), but is largely disordered in the unbound state (14). In the core region containing the three helices, the dynamics is reduced in the complex as often observed upon ligand binding. Unlike the backbone, methyl groups have a high variation in the intrinsic dynamics between side chains due to the large difference in chain lengths, and the variation in order parameters shown Fig. 4 *C* is both a result of chain length and local side-chain packing. To isolate the contribution from side-chain packing, we subtracted the average order parameter for each type of methyl group. The average values were obtained from a database consisting of methyl relaxation studies of 18 well-folded proteins used for prediction of methyl order parameters (30). Additionally, this approach allows us to compare the dynamics of the hydrophobic core of the free state of NCBD with that of folded

proteins. In the core region of NCBD, the residue-type corrected order parameters of unbound NCBD is scattered around 0, indicating that in the ps-ns timescale the hydrophobic core of the protein is approximately as dynamic as the core of an average folded protein. Fig. 5 shows the difference from the average order parameters plotted onto the structures of NCBD in the complex with ACTR and in the conformation dominating the free state. Without a bound ligand, NCBD has a cluster of methyl groups in the interfaces between the three helices that have average or above average rigidity. These residues correspond to the small hydrophobic core that we have recently proposed to be reasonably well ordered based on ring-current effects (14). In contrast, methyl groups located on the outer surface of the protein are more dynamic than average. Exceptions from this pattern are the two methyl groups of A2099 and A2100 that are exposed and rigid. Alanine methyl groups, however, are closely related to the dynamics of the backbone (42) and thus the rigidity of these nuclei do not represent side-chain packing. In the complex with ACTR, the interface between helix 1 and 2 is more rigid than in the free state. The hinge region and outer surface of helix 1, in contrast, seemingly becomes more dynamic in the complex.

Order parameters for methyl groups are not uniformly distributed, but tend to cluster in three classes (43). The first class is centered around $S^2_{axis}$ ~0.35 and results from rotameric interconversion; the second class is centered around $S^2_{axis}$ ~0.6 and results from motions within a rotameric well and motions of connecting bonds; the third class is centered around $S^2_{axis}$ ~0.8 and results from restricted motions within a rotameric well (43,44). The distribution of order parameters in NCBD is shown in Fig. S3. The distributions of $S^2_{axis}$ for both free and complexed NCBD are multimodal and roughly fall into three classes. This analysis supports the notion that free NCBD has a well-ordered core with several nonalanine methyl groups in the most rigid class. Upon ligand binding a shift in the distribution from the group centered around 0.6 to the group around 0.8 is observed, suggesting that the motions of the methyls on average become more restricted by restricting the motions of bonds adjacent to the methyls that are themselves restricted to motions within a rotameric well even in the free state. This is consistent with the observation that the methyl chemical shifts generally move away from their random coil positions upon ligand binding.

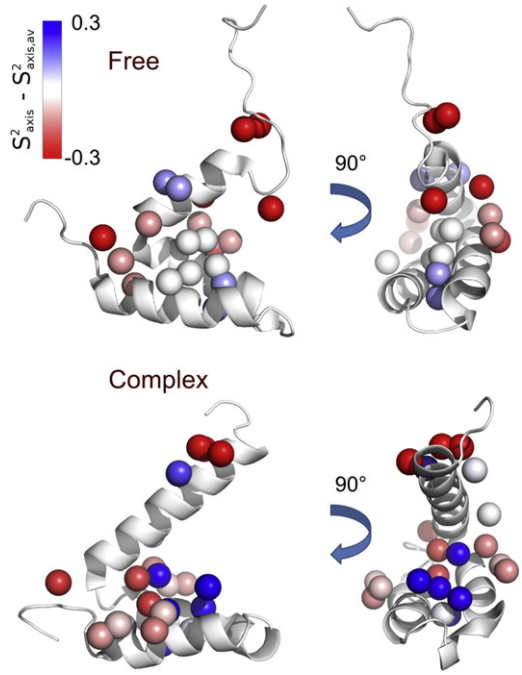


FIGURE 5 Differences from average methyl order parameters, $S^2_{axis}$ -$S^2_{axis,v}$ plotted on the structure of NCBD in the complex with ACTR and in the dominating form of the free state. Methyl groups are shown as spheres and color coded by $S^2_{axis}$-$S^2_{axis,v}$ from red (−0.3) to blue (0.3) as indicated by the scale.

## Urea denaturation

Denaturation of NCBD using urea has previously demonstrated that the protein has a folded core that unfolds cooperatively (14). The denaturation was monitored using far ultraviolet circular dichroism spectroscopy that probes the average loss of helicity with increasingly denaturing conditions, but tells little about the hydrophobic core and the behavior of individual residues. To probe the change in the side-chain packing in the hydrophobic core during unfolding, we recorded NMR spectra in increasing concentrations of urea. The peaks in all spectra are in fast exchange and can be followed throughout the titration range from native conditions to highly denaturing conditions. Six types of chemical shifts were measured for each urea concentration, resulting in between 4 and 8 chemical shifts for each residue. The backbone $^{13}C$ nuclei, $C^{\alpha}$, and $C'$, are sensitive to the secondary structure; the amide $^{15}N^{H}$ and $^{1}H^{N}$ nuclei are reporting on both the secondary and tertiary structure and the methyl $^{13}C$ and $^{1}H$ are sensitive to the hydrophobic packing. Fig. 6 and Fig. S4 show the chemical shift change with increasing urea concentration. For all nuclei, a sigmoidal chemical shift change occurs between 0 and 4 M urea, which is consistent with the data obtained by circular dichroism (14). The nuclei from each of the three helices were fit globally to a two-state transition assuming a shared *m*-value and denaturation midpoint. Despite their poor definition by the data the pretransition baselines have small slopes of the same magnitude as the posttransition baselines showing that the baselines do not absorb variations in chemical shift between each titration curve (Fig. S5) as proposed previously for BBL (45). The global model fits the data and results in *m*-values of (4.4, 4.3, and 4.1 kJ/molM) and a denaturation midpoint of (2.37, 2.36, and 2.32 M) for helix 1, 2, and 3, respectively. This corresponds

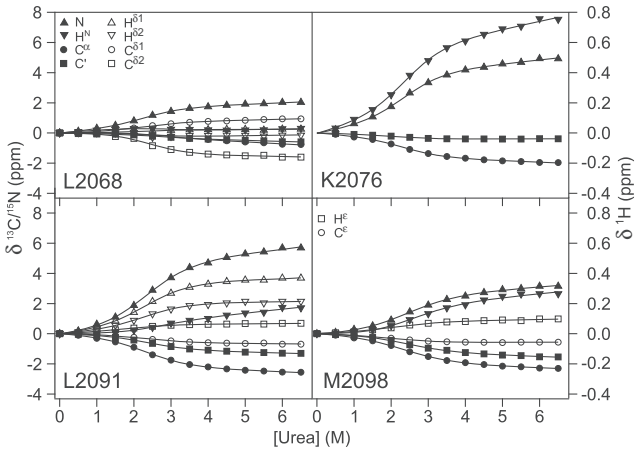


FIGURE 6 Urea denaturation of NCBD monitored by NMR for selected residues from each of the three helices. The chemical shifts were measured for NCBD from $^{15}$N-HSQC, 2D HNCOCA, 2D HNCO and constant-time $^{13}$C HSQC experiments allowing us to specifically probe helicity ($^{15}$N, C$^{\alpha}$, and C′), hydrophobic packing (methyl $^{1}$H and $^{13}$C), and tertiary structure (methyl $^{1}$H, amide $^{1}$H, and $^{15}$N).

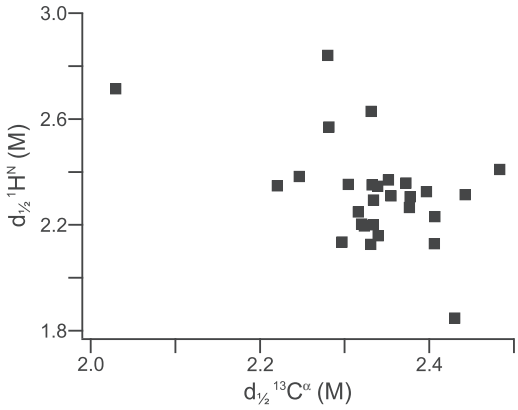


FIGURE 7 Correlation of unfolding midpoints obtained from individually fitted unfolding curves measured on H$^{N}$ and C$^{\alpha}$ of the same residue.

to a thermodynamic stability of 9.9 ± 0.3 kJ/mol, which is equivalent to ~2% of the protein being in the unfolded state. The stability is slightly higher than the one found using circular dichroism spectroscopy, most likely due to the different solvent conditions used in this study.

If each urea unfolding curve is fitted individually, we observe a large spread in thermodynamic parameters, e.g., the $^{1}$H$^{N}$ curves have unfolding midpoints between 1.8 and 2.8 M (Fig. S6). Diversity in the unfolding midpoints have previously been interpreted as evidence of downhill folding (46), and in such a case global fitting of the unfolding curves would be invalid. It has been suggested that NCBD folds downhill (15), and it is important to clarify whether this diversity represents differential stability of different parts of the molecule. The spread in unfolding midpoints, however, differ between the nuclei. The parameters from individual fits of C$^{\alpha}$ curves are thus distributed over a much narrower range than $^{1}$H$^{N}$ (Fig. S6). The uncertainties of the fits are substantial, which can be attributed to the poorly restrained pretransition baselines due to the low stability of the domain. The error bars of the data points thus overlap in the central region of the plot. To further test whether the dispersion in fitting parameters represents differential stability or fitting uncertainty, we correlated the fitting parameters from the multiple nuclei within the same residue. As is shown for $^{1}$H$^{N}$ and C$^{\alpha}$ in Fig. 7 and by the pairwise correlation coefficients for all other nuclei in Table S1, there is no correlation between these parameters. This shows that the spread in the fitting parameters is a consequence of fitting uncertainty and does not represent differential stability. Of importance, the correlation of multiple nuclei within each residue is a rigorous test for whether dispersion in the structural stability is true. From the current data set with unusually many data points per residue, we thus conclude that the unfolding of NCBD is well described by a barrier-limited two-state model.

## DISCUSSION

Here, we have demonstrated that the protein domain NCBD has a hydrophobic core as well ordered as the hydrophobic core of an average compact folded protein and almost as rigid as the ligand-bound state of the same protein. Furthermore, we have shown that the hydrophobic core unfolds cooperatively in a two-state transition. Together with our previous results that show that the dominant structure of NCBD in its ligand-free state under our experimental conditions resembles its structure in complex with the ligand ACTR (14), these results show that the behavior of NCBD is remarkably similar to a compact folded protein. This behavior may appear contradictory to the classification of NCBD as a molten globule. In several other biophysical experiments NCBD behaves as an archetypical molten globule (9,13,18). The NMR signals of NCBD are poorly dispersed and are broadened significantly, indicative of exchange between at least two conformational states (13). This conformational exchange contributes significantly to the transverse relaxation of $^{1}$H, $^{13}$C, and $^{15}$N. In contrast, the quadrupolar relaxation of deuterium that we measure here is so fast that it is only minimally sensitive to conformational exchange on the $10^{-3}$–$10^{-6}$ timescale (47). Our data are therefore consistent with previous relaxation data and suggest that NCBD forms well ordered and slowly interchanging structures. The definition of a molten globule has always been vague. It is thus unclear whether a protein that experiences slow conformational exchange, but has a well-defined hydrophobic packing at least in the major state, is a molten globule or not. Ultimately, this is a semantic discussion that depends on how inclusive the term molten globule should be. Such a discussion should not, however, derail the more important discussion of how the dynamic behavior of NCBD affects the mechanism of ligand binding.

NCBD binds several different ligands and adopts different conformations in the resulting complexes. Such behavior is also observed for other IDPs (48) and it is commonly attributed to the inherent flexibility of the disordered polypeptide chain. How important disorder is for the structural malleability of NCBD is, however, unclear. Based on coarse-grained molecular dynamics simulations and DSC, it was argued that the high flexibility of NCBD and thus its promiscuity in ligand binding is a consequence of a downhill folding mechanism for NCBD (15). It has, however, been demonstrated that downhill folding cannot be proven by DSC alone (49) calling the alleged downhill folding of NCBD into question. Accordingly, the well-ordered hydrophobic core of NCBD that we have presented here is in direct conflict with such a folding behavior that predicts rapid interconversion in a wide energy basin. In support of our results from the relaxation measurements on the methyl groups in NCBD, the urea induced unfolding of NCBD followed by chemical shift changes of nuclei in the entire protein is explained by a single cooperative transition. Our experimental data make it highly unlikely that NCBD follows a barrierless downhill folding mechanism.

Collectively, our data suggest that structural plasticity of NCBD is not due to an extraordinarily flexible structure as is usually emphasized for other intrinsically disordered proteins. In contrast, NCBD has evolved to use a different mechanism where at least the dominant fold is well defined but only marginally stable. The marginal thermodynamic stability is important because it allows a fully disordered state to be populated. Recently, this observation was supported by hydrogen exchange measurements showing that the exchange behavior of free NCBD is dominated by the fast exchange of a minor conformation in a highly disordered state (50). This highly disordered state may thus fold into alternative conformations by an induced fit mechanism. Alternatively, several proteins have recently been observed to have two coexisting well-defined conformations in solution (6–8). Fold-switching proteins invariably have low free energies of folding (8) because an exceedingly stable conformation would require alternative folds to have similar stabilities to be appreciably populated, which is likely to be evolutionarily unattainable. The ligand-free state of NCBD may thus contain alternatively folded conformers, which could initiate the binding reaction with other ligands, e.g., IRF-3 in a mechanism resembling conformational selection. Even though the dominant conformation is as well ordered as NCBD, a protein may still adapt to different ligands through both induced fit and conformational selection type mechanisms.

## SUPPORTING MATERIAL

Six figures and five tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)00216-0.

We thank Peter E. Wright (The Scripps Research Institute) for sharing the coexpression plasmid for NCBD and Mikael Akke for access to the Varian VnmrS 500 MHz spectrometer at Lund University.

This work was supported by the EliteForsk programme (to M.K.), The John and Birthe Meyer Foundation, the Carlsberg Foundation grant No. 2008-01-0368, the Danish Natural Research Council grant No. 272-08-0500 (to F.M.P), and by the Lundbeck Foundation (to K.T.).

## REFERENCES

1. Teilum, K., J. G. Olsen, and B. B. Kragelund. 2009. Functional aspects of protein flexibility. *Cell. Mol. Life Sci.* 66:2231–2247.
2. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6:197–208.
3. Wright, P. E., and H. J. Dyson. 1999. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 293:321–331.
4. Uversky, V. N. 2002. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* 11:739–756.
5. Wright, P. E., and H. J. Dyson. 2009. Linking folding and binding. *Curr. Opin. Struct. Biol.* 19:31–38.
6. Gambin, Y., A. Schug, …, A. A. Deniz. 2009. Direct single-molecule observation of a protein living in two opposed native structures. *Proc. Natl. Acad. Sci. USA.* 106:10153–10158.
7. Alexander, P. A., Y. He, …, P. N. Bryan. 2009. A minimal sequence code for switching protein structure and function. *Proc. Natl. Acad. Sci. USA.* 106:21149–21154.
8. Bryan, P. N., and J. Orban. 2010. Proteins that switch folds. *Curr. Opin. Struct. Biol.* 20:482–488.
9. Demarest, S. J., M. Martinez-Yamout, …, P. E. Wright. 2002. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature.* 415:549–553.
10. Qin, B. Y., C. Liu, …, K. Lin. 2005. Crystal structure of IRF-3 in complex with CBP. *Structure.* 13:1269–1277.
11. Waters, L., B. G. Yue, …, D. M. Heery. 2006. Structural diversity in p160/CREB-binding protein coactivator complexes. *J. Biol. Chem.* 281:14787–14795.
12. Lee, C. W., M. A. Martinez-Yamout, …, P. E. Wright. 2010. Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry.* 49:9964–9971.
13. Ebert, M. O., S. H. Bae, …, P. E. Wright. 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.
14. Kjaergaard, M., K. Teilum, and F. M. Poulsen. 2010. Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. USA.* 107:12535–12540.
15. Naganathan, A. N., and M. Orozco. 2011. The native ensemble and folding of a protein molten-globule: functional consequence of downhill folding. *J. Am. Chem. Soc.* 133:12154–12161.
16. Burger, V. M., A. Ramanathan, …, C. S. Chennubhotla. 2012. Quasi-anharmonic analysis reveals intermediate States in the nuclear co-activator receptor binding domain ensemble. *Pac. Symp. Biocomput.* 17:70–81.
17. Ganguly, D., W. Zhang, and J. Chen. 2011. Synergistic folding of two intrinsically disordered proteins: searching for conformational selection. *Mol. Biosyst.* 8:198–209.
18. Demarest, S. J., S. Deechongkit, …, P. E. Wright. 2004. Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. *Protein Sci.* 13:203–210.
19. Dobson, C. M. 1994. Protein folding. Solid evidence for molten globules. *Curr. Biol.* 4:636–640.

20. Ptitsyn, O. B. 1995. Molten globule and protein folding. *Adv. Protein Chem.* 47:83–229.

21. Eliezer, D., J. Yao, …, P. E. Wright. 1998. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* 5:148–155.

22. Kjaergaard, M., V. Iešmantavičius, and F. M. Poulsen. 2011. The interplay between transient $\alpha$-helix formation and side chain rotamer distributions in disordered proteins probed by methyl chemical shifts. *Protein Sci.* 20:2023–2034.

23. Delaglio, F., S. Grzesiek, …, A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 6:277–293.

24. Vranken, W. F., W. Boucher, …, E. D. Laue. 2005. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins.* 59:687–696.

25. Bax, A., D. Max, and D. Zax. 1992. Measurement of long-range $^{13}C$-$^{13}C$ J couplings in a 20-kDaprotein-peptide complex. *J. Am. Chem. Soc.* 114:6923–6925.

26. Millet, O., D. R. Muhandiram, …, L. E. Kay. 2002. Deuterium spin probes of side-chain dynamics in proteins. 1. Measurement of five relaxation rates per deuteron in $^{(13)}C$-labeled and fractionally $^{(2)}H$-enriched proteins in solution. *J. Am. Chem. Soc.* 124:6439–6448.

27. Skrynnikov, N. R., O. Millet, and L. E. Kay. 2002. Deuterium spin probes of side-chain dynamics in proteins. 2. Spectral density mapping and identification of nanosecond time-scale side-chain motions. *J. Am. Chem. Soc.* 124:6449–6460.

28. Farrow, N. A., R. Muhandiram, …, L. E. Kay. 1994. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry.* 33:5984–6003.

29. Fushman, D., R. Weisemann, …, H. Rüterjans. 1994. Backbone dynamics of ribonuclease T1 and its complex with 2′GMP studied by two-dimensional heteronuclear NMR spectroscopy. *J. Biomol. NMR.* 4:61–78.

30. Carbonell, P., and A. del Sol. 2009. Methyl side-chain dynamics prediction based on protein structure. *Bioinformatics.* 25:2552–2558.

31. Tugarinov, V., and L. E. Kay. 2004. Stereospecific NMR assignments of prochiral methyls, rotameric states and dynamics of valine residues in malate synthase G. *J. Am. Chem. Soc.* 126:9827–9836.

32. Grzesiek, S., J. Anglister, and A. Bax. 1993. Correlation of backbone amide and aliphatic side-chain resonances in C-13/N-15-enriched proteins by isotropic mixing of C-13 magnetization. *J. Magn. Reson. B.* 101:114–119.

33. Tonelli, A. E., and F. C. Schilling. 1981. C-13 Nmr chemical-shifts and the microstructure of polymers. *Acc. Chem. Res.* 14:233–238.

34. London, R. E., B. D. Wingad, and G. A. Mueller. 2008. Dependence of amino acid side chain 13C shifts on dihedral angle: application to conformational analysis. *J. Am. Chem. Soc.* 130:11097–11105.

35. Mulder, F. A. A. 2009. Leucine side-chain conformation and dynamics in proteins from 13C NMR chemical shifts. *Chem. Bio. Chem.* 10: 1477–1479.

36. Hansen, D. F., P. Neudecker, …, L. E. Kay. 2010. Determination of Leu side-chain conformations in excited protein states by NMR relaxation dispersion. *J. Am. Chem. Soc.* 132:42–43.

37. Hansen, D. F., P. Neudecker, and L. E. Kay. 2010. Determination of isoleucine side-chain conformations in ground and excited states of proteins from chemical shifts. *J. Am. Chem. Soc.* 132:7589–7591.

38. Hansen, D. F., and L. E. Kay. 2011. Determining valine side-chain rotamer conformations in proteins from methyl 13C chemical shifts: application to the 360 kDa half-proteasome. *J. Am. Chem. Soc.* 133:8272–8281.

39. Muhandiram, D. R., T. Yamazaki, …, L. E. Kay. 1995. Measurement of H-2 T-1 and T-1p relaxation-times in uniformly C-13-labeled and fractionally H-2-labeled proteins in solution. *J. Am. Chem. Soc.* 117:11536–11544.

40. Lipari, G., and A. Szabo. 1982. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* 104:4546–4559.

41. Lipari, G., and A. Szabo. 1982. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* 104:4559–4570.

42. Mittermaier, A., L. E. Kay, and J. D. Forman-Kay. 1999. Analysis of deuterium relaxation-derived methyl axis order parameters and correlation with local structure. *J. Biomol. NMR.* 13:181–185.

43. Igumenova, T. I., K. K. Frederick, and A. J. Wand. 2006. Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chem. Rev.* 106:1672–1699.

44. Best, R. B., J. Clarke, and M. Karplus. 2004. The origin of protein side-chain order parameter distributions. *J. Am. Chem. Soc.* 126:7734–7735.

45. Sadqi, M., D. Fushman, and V. Munoz. 2007. Structural biology: analysis of 'downhill' protein folding; analysis of protein-folding cooperativity. *Nature.* 445:E17–E18.

46. Sadqi, M., D. Fushman, and V. Muñoz. 2006. Atom-by-atom analysis of global downhill protein folding. *Nature.* 442:317–321.

47. Choy, W. Y., and L. E. Kay. 2003. Model selection for the interpretation of protein side chain methyl dynamics. *J. Biomol. NMR.* 25:325–333.

48. Oldfield, C. J., J. Meng, …, A. K. Dunker. 2008. Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics.* 9(*Suppl 1*):S1.

49. Farber, P., H. Darmawan, …, A. Mittermaier. 2010. Analyzing protein folding cooperativity by differential scanning calorimetry and NMR spectroscopy. *J. Am. Chem. Soc.* 132:6214–6222.

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