



A structurally heterogeneous transition state underlies coupled binding and folding of disordered proteins

Received for publication, September 13, 2018, and in revised form, November 30, 2018. Published, Papers in Press, December 4, 2018, DOI 10.1074/jbc.RA118.005854

✉ Elin Karlsson[‡], Eva Andersson[‡], Jakob Dogan[§], Stefano Gianni[¶], Per Jemth^{†1}, and Carlo Camilloni^{||2}

From the [‡]Department of Medical Biochemistry and Microbiology, Uppsala University, SE-75123 Uppsala, Sweden, the

[§]Department of Biochemistry and Biophysics, Stockholm University, SE-10691 Stockholm, Sweden, the [¶]Istituto Pasteur-

Fondazione Cenci Bolognetti and Istituto di Biologia e Patologia Molecolari del CNR, Dipartimento di Scienze Biochimiche "A. Rossi Fanelli," Sapienza Università di Roma, 00185 Rome, Italy, and the ^{||}Dipartimento di Bioscienze, Università degli Studi di Milano, 20133 Milano, Italy

Edited by Wolfgang Peti

Many intrinsically disordered proteins (IDPs) attain a well-defined structure in a coupled folding and binding reaction with another protein. Such reactions may involve early to late formation of different native structural regions along the reaction pathway. To obtain insights into the transition state for a coupled binding and folding reaction, we performed restrained molecular dynamics simulations using previously determined experimental binding Φ_b values of the interaction between two IDP domains: the activation domain from the p160 transcriptional co-activator for thyroid hormone and retinoid receptors (ACTR) and the nuclear co-activator binding domain (NCBD) of CREB-binding protein, each forming three well-defined α -helices upon binding. These simulations revealed that both proteins are largely disordered in the transition state for complex formation, except for two helices, one from each domain, that display a native-like structure. The overall transition state structure was extended and largely dynamic with many weakly populated contacts. To test the transition state model, we combined site-directed mutagenesis with kinetic experiments, yielding results consistent with overall diffuse interactions and formation of native intramolecular interactions in the third NCBD helix during the binding reaction. Our findings support the view that the transition state and, by inference, any encounter complex in coupled binding and folding reactions are structurally heterogeneous and largely independent of specific interactions. Furthermore, experimental Φ_b values and Brønsted plots suggested that the transition state is globally robust with respect to most mutations but can display more native-like features for some highly destabilizing mutations, possibly because of Hammond behavior or ground-state effects.

Protein–protein interactions involving intrinsically disordered proteins (IDPs)³ are pivotal in cellular signaling pathways. The flexibility of the IDP allows for extended binding surfaces for recognition and easy access for kinases and phosphatases that modulate the affinity of the protein–protein interaction by (de)phosphorylation. Another aspect of protein–protein interactions of IDPs is that the binding often involves (partial) folding of the IDP into a well-defined structure in the protein–protein complex, a process that in itself can modulate the affinity of the interaction. The details of such coupled binding and folding reactions have therefore recently been addressed to better understand how the disorder of the interacting proteins is affecting the mechanism of the interaction and how it may provide functional advantages (1, 2). The emerging picture is that coupled binding and folding reactions, much like protein folding, can employ mechanisms ranging from a largely nonspecific initial encounter (corresponding to pure nucleation–condensation in protein folding) (3–6) to substantial preformed structure (corresponding to diffusion–collision in protein folding) (7–9). Most studies suggest that the initial stages of a protein–protein interaction, in which at least one binding partner is an IDP, contain only a very small fraction of “native” interactions, *i.e.* interactions that are present in the bimolecular complex at equilibrium. However, there is one example of a highly structured α helix in the transition state (TS) for coupled binding and folding of cMyb to KIX (10).

We have previously performed extensive mechanistic studies on one of the classical examples of a coupled binding and folding reaction, that between ACTR and NCBD (11). The kinetics are complex, reflecting accumulation of various intermediates depending on the conditions. First, NCBD, which was described as a molten globule–like protein domain (12), displays ground state heterogeneity on a submillisecond time scale (13, 14) and on a time scale over tens of seconds related to a proline *cis-trans*-isomerization (15). Second, the binding between NCBD and the highly disordered ACTR involves two concentration-dependent kinetic phases related to the ini-

This work was supported by Swedish Research Council Grant 2016-04965 (to P. J.). The authors declare that they have no conflicts of interest with the contents of this article.

✂ Author's Choice—Final version open access under the terms of the Creative Commons CC-BY license.

This article contains Table S1, Figs. S1–S5, and supporting Excel file.

¹ To whom correspondence may be addressed. Tel.: 46-18-471-4557; E-mail: Per.Jemth@imbim.uu.se.

² To whom correspondence may be addressed. Tel.: 39-02-503-14918; E-mail: carlo.camilloni@unimi.it.

³ The abbreviations used are: IDP, intrinsically disordered protein; ACTR, activation domain from the p160 transcriptional co-activator for thyroid hormone and retinoid receptors; NCBD, nuclear co-activator binding domain; TS, transition state; TMAO, trimethylamine *N*-oxide; MD, molecular dynamics.

tial steps of association: one detected in stopped-flow experiments that is the major binding phase (16) and one detected in single molecule experiments involving the *cis*-proline conformation of NCBD (15). Third, in stopped-flow binding experiments, two additional phases were observed ($\tau \sim 25$ ms and $\tau \sim 0.9$ s, respectively, at 4 °C). The 25-ms phase is observed only in presence of trimethylamine *N*-oxide (TMAO) or high salt. The bases for both of these phases are likely structural rearrangements following the initial binding event. Fourth, in single molecule experiments, a submillisecond phase was detected following initial binding that might be related to a high-energy intermediate (17). Our experimental studies on this interaction involved site-directed mutagenesis and kinetic experiments to obtain Φ_b values reporting on the degree of native contacts formed in the TS for the initial binding event. The implications of the complex binding kinetics for interpretation of Φ_b values were discussed in detail (18). It was shown that Φ_b values will report on the initial binding event for the major binding pathway, and this holds true for the *cis-trans*-proline heterogeneity in NCBD as well, where the major *trans*-proline NCBD conformation will dominate the observed kinetic phase in stopped-flow experiments. The Φ_b analysis suggested that most native hydrophobic interactions form late in the reaction, downhill of the TS for binding (18), whereas helix 1 in ACTR is partially formed at the top of the barrier (19). By inference, several initial and TS interactions along the reaction pathway are non-native, *i.e.* not found in the complex at equilibrium. In the present work we use molecular dynamics (MD) simulations restrained by the previous Φ_b values, followed by additional experiments and reanalysis of previous double-mutant data to obtain a complete picture of the binding TS of this paradigmatic coupled binding and folding reaction.

Results

We have here obtained a structural model of the TS for the coupled binding and folding reaction of NCBD and ACTR by a three-step approach. First we used restrained MD simulations (20–22) using previously determined experimental Φ_b values based on single-point mutations (18, 19). Hereby we obtained a structural ensemble for the binding TS. Second, we made new single-point mutants to test the MD simulations. Third, to map the plasticity of the TS, we used linear free energy diagrams to reanalyze another previously published kinetic data set on the ACTR–NCBD interaction that was based on double mutants. Briefly, the binding kinetics of all single mutants of ACTR were measured with all single mutants of NCBD. This will probe how sensitive the second mutation is to perturbation by the first mutation.

Φ_b restrained MD simulations of the ACTR–NCBD interaction

To obtain a structural ensemble of the TS, MD simulations were run in GROMACS using the Amber03w force field as detailed under “Experimental procedures.” We included Φ_b values in the 0–1 range from single-point mutations (18, 19) to restrain the MD simulation (Table S1). The reason for using the 0–1 range is that the standard interpretation of a Φ value is in terms of fraction of native contact. Most of the included values were conservative deletion mutations. However, three substi-

tutions in helix 1 of ACTR involved polar or charged groups, which complicates interpretation of Φ values caused by solvation effects, which do not cancel out (23). However, these mutations were designed to specifically probe secondary structure propensity, and this effect was confirmed by NMR (19). Furthermore, one of the mutations in ACTR helix 1 was the preferred Ala \rightarrow Gly, which reported a Φ_b value of 0.24 (Table S1). Thus, overall we deem the input Φ_b value data set sound, but as pointed out (23), the number of any single Φ value should be treated with caution and interpreted as weak, intermediate, or strong. Furthermore, several Φ values should preferably be measured for every structural region, which is the case here for helix 1 of ACTR.

Overall, the TS structure is very extended and disordered (Fig. 1). In particular a comparison to a simulation for the ground state of the bimolecular complex highlighted how the TS ensemble is characterized by a highly heterogeneous size (measured by the radius of gyration; Fig. 1c), as well as by highly heterogeneous conformers (measured by the distribution of pairwise RMSDs; Fig. 1d). Analysis of the secondary structure populations calculated using DSSP (24) and averaged over the ensemble shows that most of the native helical structure in the TS is not formed, except for helix 1 of ACTR ($\text{A}\alpha 1$) and helix 3 of NCBD ($\text{N}\alpha 3$) that are almost native-like (Fig. 1e). The tertiary and quaternary structure arrangement as revealed by a contact analysis contains many more weakly populated contacts in the TS than in the native state (Fig. 1b). The interface between the two proteins in the TS is diffused and disordered as compared with that of the native complex (Fig. 1b), with the TS showing a larger number of less populated contacts (Fig. S1). In the TS the interdomain contacts between the native-like $\text{A}\alpha 1$ and $\text{N}\alpha 3$ are relatively well-formed with respect to other interface contacts (Fig. 1b). The overall picture suggests that the two helices are the structural core of the TS around which the remaining elements of the complex organizes to form a very heterogeneous TS structure stabilized by unspecific interactions.

Testing the binding TS model with mutation

To probe the effect of the many weakly populated contacts at the interface between the two proteins, we selected residues involved in such contacts for additional experiments. Based on the contact maps (Fig. 1b), six mutants ($\text{ACTR}_{\text{E}1065\text{A}}$, $\text{ACTR}_{\text{E}1066\text{A}}$, $\text{ACTR}_{\text{P}1074\text{A}}$, $\text{NCBD}_{\text{Q}2068\text{A}}$, $\text{NCBD}_{\text{T}2073\text{A}}$, and $\text{NCBD}_{\text{A}2098\text{G}}$) were expressed, purified, and subjected to kinetic binding experiments with pseudo-WT $\text{NCBD}_{\text{Y}2108\text{W}}$ and WT ACTR using the stopped-flow technique (Table 1, Fig. 2, and Figs. S2 and S3) as previously described (16). The same caveats with regard to interpretation of Φ_b values, as discussed in the previous section, applies here (23).

$\text{NCBD}_{\text{A}2098\text{G}}$ was used to probe native formation of $\text{N}\alpha 3$ in the TS. The shape of the CD spectrum, which reflects secondary structure content, was slightly different for $\text{NCBD}_{\text{A}2098\text{G}}$ as compared with the other NCBD variants (208 *versus* 222 nm; Fig. S2). This suggests a structural perturbation, and kinetic stopped-flow experiments were therefore conducted in the presence of 0.7 M TMAO, which restored the shape of the CD spectrum for this mutant. The association rate constant for

A heterogeneous transition state for an IDP interaction

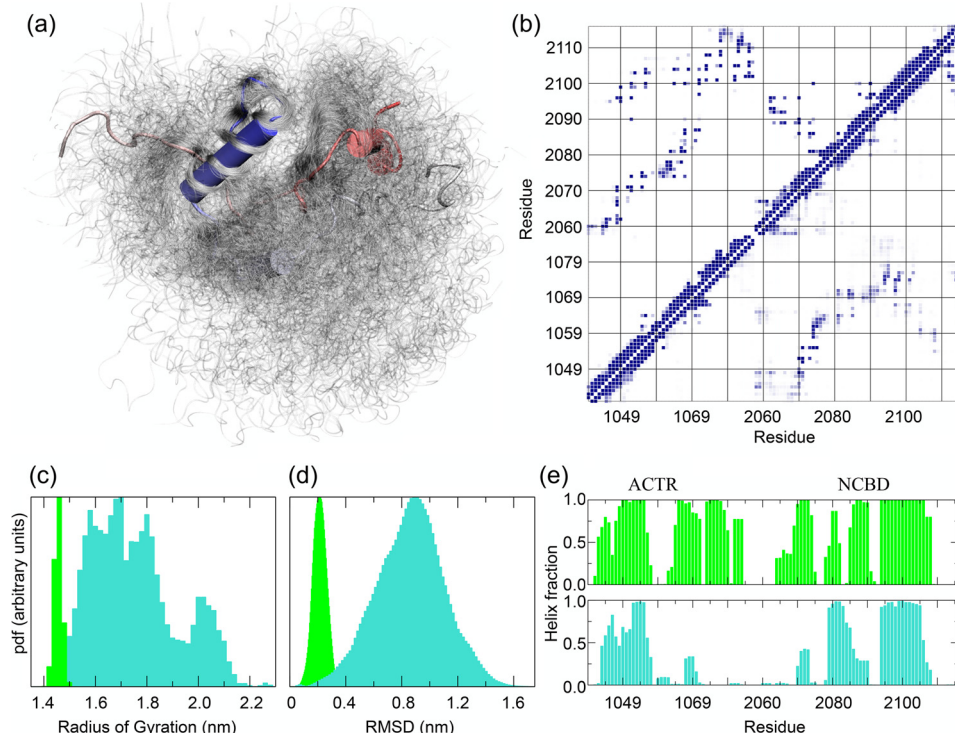


Figure 1. Model of the transition state ensemble for the binding between ACTR and NCBD. *a*, structural representation of the TS ensemble with ACTR in violet (helix 1) and NCBD in blue (helix 3). *b*, contact map analysis for the bound native state (*upper diagonal*) and the TS ensemble (*lower diagonal*). The probability of contact formation goes from white (0) to blue (1). *c* and *d*, comparison of the radius of gyration (*c*) and root mean square deviation distributions (*d*) for the native bound state of ACTR–NCBD (*green bars*) and the TS ensemble (*turquoise bars*). Probability density function (*pdf*) in arbitrary units is shown on the y axis. *e*, comparison of the helical content for a simulation of the native bound state of ACTR–NCBD (*green bars*) and the TS ensemble (*turquoise bars*).

Table 1

Experimentally determined binding rate constants, dissociation constants, and Φ_b values for the association between ACTR and NCBD

The measurements were conducted in 20 mM sodium phosphate (pH 7.4), 150 mM NaCl at $T = 277$ K unless otherwise stated. Typically, NCBD_{Y2108W} was held constant at 1 μ M, and ACTR was added in excess. Association rate constants and K_d values from experiments in which NCBD_{Y2108W} was added in excess over ACTR are shown in parentheses.

ACTR variant	NCBD variant	Interaction probed	k_{off} s^{-1}	k_{on} $\mu\text{M}^{-1} s^{-1}$	K_d nM	Φ_b
WT	NCBD _{Y2108W}		2.2 ± 0.1	21.3 ± 0.4 (17.0 ± 0.6)	103 ± 7 (129 ± 7)	
WT	NCBD _{Y2108W} ^a		0.49 ± 0.02	95 ± 10	5.2 ± 0.8	
WT	NCBD _{Q2068A}	Non-native interactions with the N-terminal of ACTR	2.7 ± 0.1	28.1 ± 0.6	96 ± 6	
WT	NCBD _{T2073A}	Non-native interaction with Leu-1052 and residual helical structure	1.08 ± 0.03	41 ± 2	26 ± 2	0.49
WT	NCBD _{N2088A}	Cluster of contacts with region 1069–1072 of ACTR	2.3 ± 0.1	22.2 ± 0.8	104 ± 8	
WT	NCBD _{A2098G} ^a	Formation of N α 3	1.68 ± 0.04	33.8 ± 0.3	50 ± 2	0.45
ACTR _{E1065A}	NCBD _{Y2108W}	Non-native interaction with backbone of NCBD Lys-2075	2.2 ± 0.1	21.3 ± 0.6 (18 ± 1)	104 ± 8 (122 ± 9)	
ACTR _{E1066A}	NCBD _{Y2108W}	Non-native interaction with backbone of NCBD Ser-2076	1.73 ± 0.02	22.4 ± 0.3 (17.9 ± 0.7)	77 ± 2 (97 ± 4)	
ACTR _{P1074A}	NCBD _{Y2108W}	Cluster of contacts with region 2094–2098 of NCBD	1.9 ± 0.1	18.5 ± 0.6 (19.1 ± 0.6)	103 ± 9 (99 ± 6)	

^a The experiments were conducted in 20 mM sodium phosphate (pH 7.4), 150 mM NaCl, 0.7 M TMAO.

NCBD_{A2098G} decreased 3-fold as compared with the pseudo-WT NCBD_{Y2108W} (in 0.7 M TMAO), whereas the dissociation rate constant increased 3-fold, yielding a Φ_b value of 0.45 and thus confirming early formation of native interactions in N α 3. (Φ_b in buffer without TMAO was very similar and also in the intermediary category, 0.38.) We note that the binding of the pseudo-WT NCBD_{Y2108W} to WT ACTR in the presence of 0.7 M TMAO exhibited biphasic kinetics, where one phase increased linearly with the concentration of ACTR (the one used to calculate k_{on}) whereas the second phase had a rather constant k_{obs} value of ~ 50 s⁻¹. The linear kinetic phase reports on the binding event as discussed previously (18) and in the introduction.

The NCBD_{T2073A} mutant, probing both native and non-native intramolecular interactions in the TS, showed a 2-fold increase in association rate constant and a 2-fold decrease in the dissociation rate constant, suggesting a deletion of a nonfavorable interaction in the TS in agreement with the simulation. The results are also consistent with deletion of a native unfavorable interaction between Ile-2089 and the hydroxyl group of Thr-2073 that imposes frustration in the ground states of both free and complexed NCBD, although analysis using the Frustratometer 2 (25) suggests little frustration in this region. In either case, NCBD_{T2073A} yielded a Φ_b value of 0.49 compatible with partial native helical population in the TS (Fig. 1*b*). In addition, this mutant exhibited biphasic kinetics with one lin-

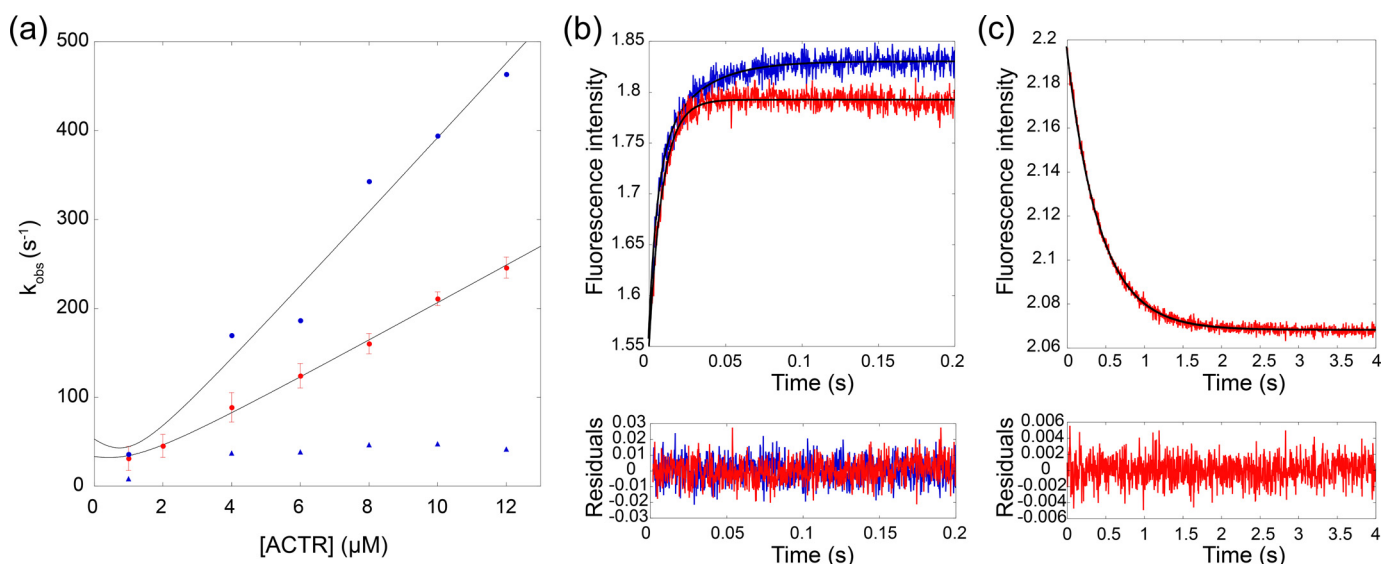


Figure 2. Binding kinetics for ACTR and NCBD. *a*, the dependences of k_{obs} on ACTR concentration for NCBD_{Y2108W} (pseudo-WT, shown in red) and NCBD_{T2073A} (blue) at 277 K in 20 mM sodium phosphate (pH 7.4), 150 mM NaCl. NCBD_{T2073A} exhibited biphasic kinetics where the second phase (blue triangles) was rather constant with ACTR concentration. The linear phase was used to calculate the Φ_b value. Each k_{obs} value of NCBD_{Y2108W} is an average of two different experiments conducted with different samples, and the error bars show the standard deviation for k_{obs} at each ACTR concentration. *b*, a typical kinetic trace for the binding of NCBD_{T2073A} (blue) and NCBD_{Y2108W} (red) to ACTR_{WT}. The concentrations were 1 μM NCBD and 6 μM ACTR. The black line represents the fit to a double exponential function for NCBD_{T2073A} and a single exponential curve for NCBD_{Y2108W}. The residuals for the fits are shown below each plot. *c*, the dissociation rate constants were determined in displacement experiments in which NCBD_{WT} was used to displace the different NCBD variants from the complex. The figure shows the displacement of NCBD_{Y2108W} from its complex with ACTR upon addition of NCBD_{WT}. The black line represents the fit to a single exponential curve, and the residuals are shown below the graph.

ear and one apparently constant phase with a k_{obs} value of $\sim 40 \text{ s}^{-1}$ in the absence of TMAO, showing that the hydroxyl of the Thr facilitates an intramolecular step on the path to the native bound conformation. The Φ_b value calculated from the linear phase reports on the TS for binding as previously discussed (18).

The remaining point mutants (NCBD_{Q2068A}, NCBD_{N2088A}, ACTR_{E1065A}, ACTR_{E1066A}, and ACTR_{P1074A}) showed no significant effects on the association and dissociation rate constants. For this reason, $\Delta\Delta G_{\text{Eq}}$ values were too low ($< 0.18 \text{ kcal mol}^{-1}$) to allow calculation of Φ_b values and validation of any specific non-native interactions. However, although not a proof, the lack of effect of the mutations is consistent with an interface between the two proteins in the TS that is malleable, stabilized by unspecific interactions and able to easily rearrange upon mutagenesis.

Brønsted plots and Φ_b values of the ACTR–NCBD interaction

To further assess the role of the respective protein domain in determining the properties of the TS in the coupled binding and folding reaction, we reanalyzed a previous comprehensive data set on the binding kinetics between ACTR and NCBD (26). This data set contains kinetic experiments with pairs of mutations, one in ACTR and one in NCBD (and also including the pseudo-WT variants). Such a double-mutant cycle gives the possibility to calculate interaction (or coupling) free energies $\Delta\Delta G_C$ between any two residues (27), which was done in the previous study (26). However, these data can also be subjected to Brønsted/Leffler type analysis (28) in which the change in free energy for the TS ($\Delta\Delta G_{\text{TS}}$) upon mutational perturbation is plotted versus the change in free energy at equilibrium ($\Delta\Delta G_{\text{Eq}}$).

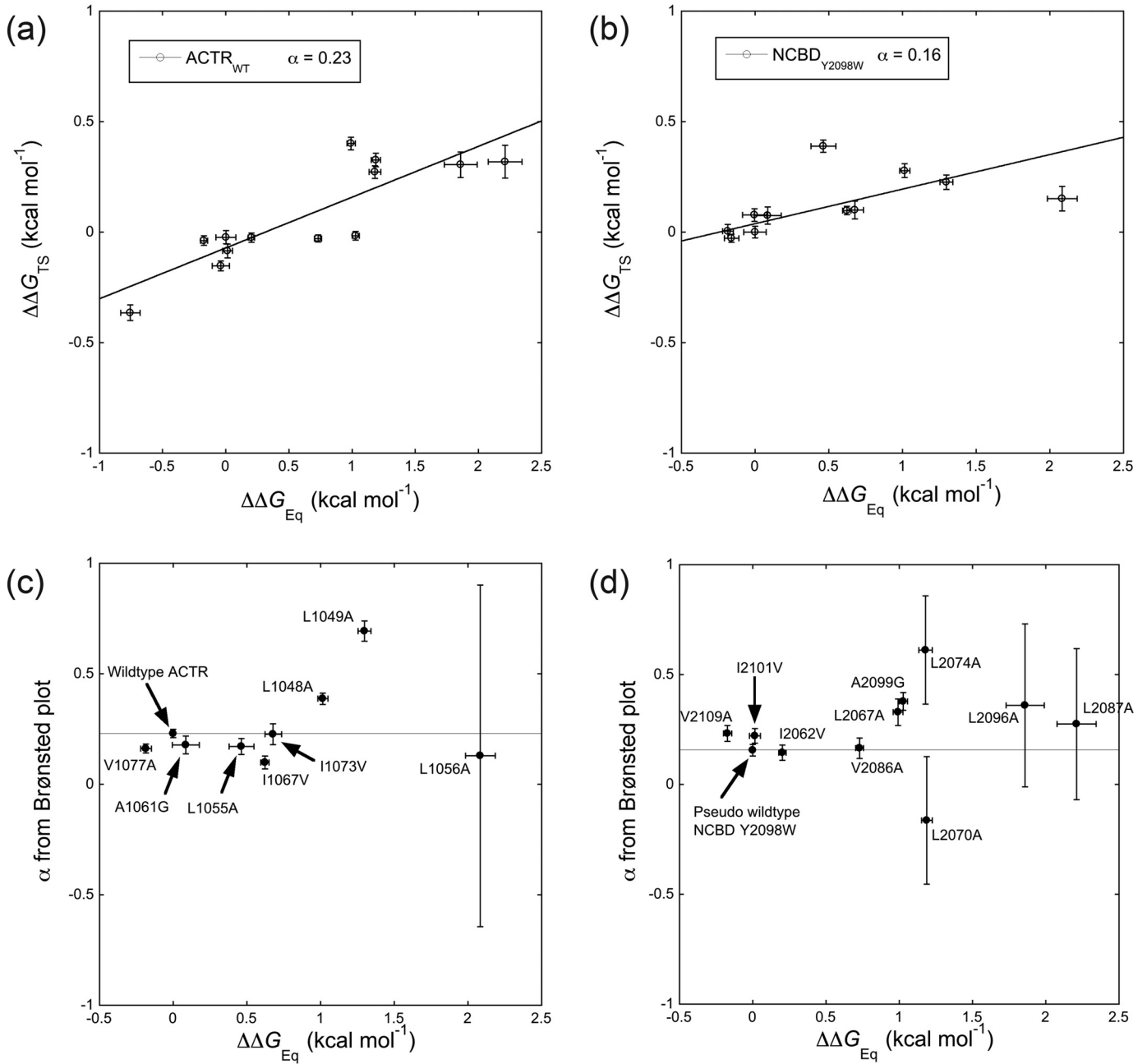
Such linear free energy relationships were originally devised for physical organic chemistry but adopted by enzymology (29), protein folding (30), and binding (31) studies. We evaluated, first, the slope α of Brønsted plots ($\Delta\Delta G_{\text{TS}}$ versus $\Delta\Delta G_{\text{Eq}}$), which is a measure of the overall resemblance of the TS to the native state ($\alpha = 1$ for full native interactions and 0 for no native interactions) and, second, Φ_b values, which report on local structure. In these data sets, binding kinetics were measured for pseudo-WT NCBD_{Y2108W} and 10 point mutants of NCBD_{Y2108W} versus WT ACTR and 8 point mutants of ACTR. For each NCBD variant, 6–8 kinetic experiments with distinct ACTR variants could be performed. Likewise, from the perspective of ACTR, for each ACTR variant 3–10 kinetic experiments with distinct NCBD variants were performed. In total, 87 binding experiments with ACTR and NCBD wildtype and point mutants of 99 possible combinations were included in the analysis. The complexes of some pairs, for example NCBD_{I2062V} and ACTR_{L1056A}, were too unstable to allow a kinetic analysis and thus excluded. This data set was previously interpreted in terms of coupling free energies between specific positions in ACTR and NCBD (26).

The slopes of Brønsted plots were overall similar among NCBD variants (median $\alpha = 0.23$, average $\alpha = 0.24$), but with large errors in the slope for NCBD mutations where $\Delta\Delta G_{\text{Eq}} > 1 \text{ kcal mol}^{-1}$ (Fig. 3). A similar pattern is observed for Brønsted plots among ACTR variants (median $\alpha = 0.18$, average $\alpha = 0.25$). However, α differs for some point mutants. For NCBD, all four α values, which deviate significantly from that of pseudo-WT NCBD_{Y2108W}, were from variants with $\Delta\Delta G_{\text{Eq}} \geq 1 \text{ kcal mol}^{-1}$ and with slightly perturbed CD spectra (suggesting ground state structural effects on NCBD). Turning to ACTR we

A heterogeneous transition state for an IDP interaction

observe a similar trend with larger deviation from WT α values when $\Delta\Delta G_{\text{Eq}} > 1$ kcal mol⁻¹. The mutated residues of two ACTR variants with higher α values (L1048A and L1049A) cluster in the complex with three NCBD residues giving destabilized NCBD variants (L2067A, L2070A, and L2074A). These

residues are part of the $\alpha 1$ helices of ACTR and NCBD, respectively, and could either perturb the folding pathway or, because they are all highly destabilizing and potentially disruptive Leu \rightarrow Ala mutations, result in structural changes in the ground state complex.



(e)

	pWT NCBD	I2062V	L2067A	L2070A	L2074A	V2086A	L2087A	L2096A	A2099G	I2101V	V2109A
WT ACTR	x	x	x	x	x	x	x	x	x	x	x
L1048A	x	x	x	x	x	x	x	x	x	x	x
L1049A	x	x	x	x	x	x			x	x	x
L1055A	x	x	x	x	x	x	x	x	x	x	x
L1056A	x					x				x	x
A1061G	x		x	x	x	x	x	x	x	x	x
I1067V	x	x	x	x	x	x	x	x	x	x	x
I1073V	x	x	x	x	x	x	x	x	x	x	x
V1077A	x	x	x			x	x	x	x	x	x

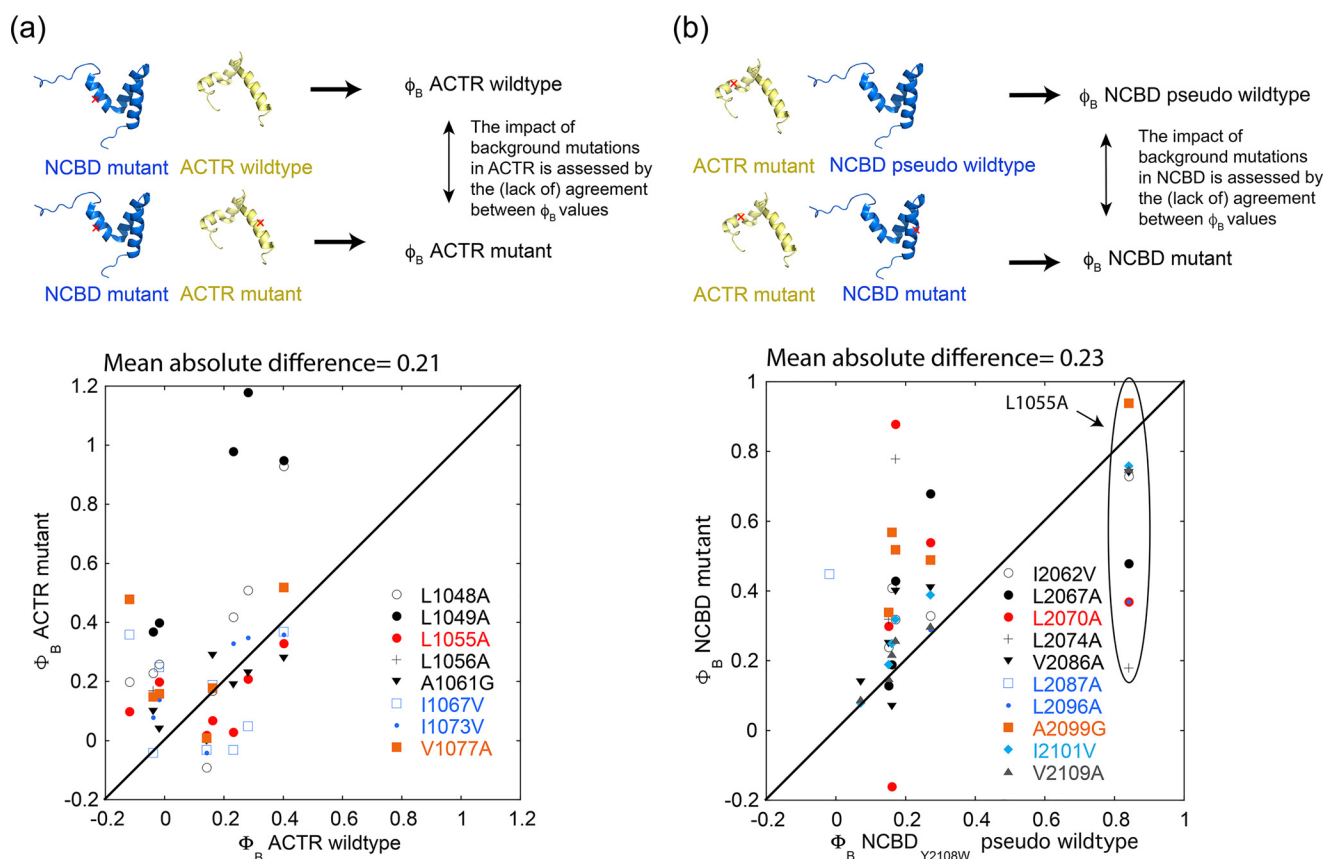


Figure 4. Φ_b values in WT versus mutants. *a*, Φ_b values for mutations in NCBD determined for different variants of ACTR. *b*, Φ_b values for mutations in ACTR determined for different variants of NCBD. The Leu-1055 mutations are in the oval. The solid line corresponds to perfect agreement (slope = 1). Some of the Φ_b values have very large propagated errors that are not shown for clarity. Φ_b values with errors of >1 are not shown. The mean absolute difference between Φ_b values in WT and mutant background is shown above each graph. The Protein Data Bank entries 1KBH (11) and 2KKJ (12) were used to draw models of ACTR (*i.e.* hypothetical) and NCBD, respectively.

Second, we calculated Φ_b values with each ACTR and NCBD variant as a WT. In this way, we could assess how point mutation in NCBD affected Φ_b values for mutations in ACTR and vice versa. Plots of Φ_b values for WT ACTR and NCBD, respectively, versus Φ_b values of their point mutants result in a large scatter. The mean absolute difference between Φ_b values in WT and mutant background is ~ 0.2 for both NCBD and ACTR variants (Fig. 4). In other words, many of the Φ_b values differ between NCBD variants, which means that the mutational perturbation in NCBD affects the measured Φ_b value obtained from mutation in ACTR. As visualized by the error bars for the slope α from Brønsted plots (Fig. 3, *c* and *d*), Φ_b values are more similar for smaller mutational perturbations in NCBD (<1 kcal mol⁻¹) in agreement with general guidelines for Φ value analysis in protein folding (23). Considering only $\Delta\Delta G_{Eq} < 1$ kcal mol⁻¹, the $\Phi_b - \Phi_b$ values show an improved agreement. This is evident from the lower mean absolute difference (0.15 and 0.10 for background mutations in ACTR and NCBD, respectively)

between the $\Phi_b - \Phi_b$ pairs (Fig. S4). Four of the NCBD variants with $\Delta\Delta G_{Eq} > 1$ have significant ground state effects based on CD. However, Φ for folding were never considered perfectly accurate because of several caveats and should be viewed as low, intermediate, or high. Indeed, the Φ_b values in the present study mainly fluctuate within the same category (0–0.3, 0.3–0.6, and 0.6–1).

One of the substitutions, L1049A in helix 1 of ACTR ($\Delta\Delta G_{Eq} = 1.3$ kcal mol⁻¹), was particularly intriguing because all five Φ_b values, which could be determined by mutation in NCBD, were significantly higher in the context of ACTR_{L1049A} as compared with ACTR WT (Fig. 4a). This is also reflected in the high $\alpha = 0.69 \pm 0.05$. Leu-1049 is situated in the more structured part of the TS (Fig. 1), and destabilization by L1049A might result in a shift of the position of the TS along the reaction coordinate toward a more native conformation in agreement with Hammond behavior as seen in protein folding studies (30). There is a trend to higher Φ_b values (Hammond behavior) also for the neighboring amino acid,

Figure 3. Linear free energy relationships for the ACTR–NCBD interaction. *a* and *b*, show two examples of Brønsted plots. Shown are the differences in free energies upon mutation for the binding between ACTR and NCBD ($\Delta\Delta G_{TS}$ plotted versus $\Delta\Delta G_{Eq}$) for WT ACTR versus a panel of pseudo-WT NCBD_{Y2108W} mutants (*a*) and for pseudo-WT NCBD_{Y2108W} versus a panel of ACTR mutants (*b*). A slope (α) close to 0 indicates a low degree of native contacts in the TS, whereas a slope of one suggests that all native contacts are formed in the TS. Moreover, a linear relationship between $\Delta\Delta G_{TS}$ and $\Delta\Delta G_{Eq}$ is indicative of a nucleation–condensation type mechanism where all native contacts form simultaneously around a “nucleus.” On the other hand, nonlinearity is more consistent with intermediates and preformed structure. Similar Brønsted plots were made for each ACTR and pseudo-WT NCBD_{Y2108W} variant. *c* and *d*, the slope α from these Brønsted plots was then plotted against $\Delta\Delta G_{Eq}$ for each variant: ACTR variants (*c*) and pseudo-WT NCBD_{Y2108W} variants (*d*). The horizontal lines correspond to the α value of WT ACTR (from *a*, shown in *c*) and pseudo-WT NCBD_{Y2108W} (from *b*, shown in *d*). *e*, all pairs of variants that were included in the analysis are marked by an X. The pairs of variants not included were too destabilized by the two mutations, resulting in binding kinetics too fast for the stopped-flow experiments.

A heterogeneous transition state for an IDP interaction

L1048A ($\alpha = 0.39 \pm 0.03$). Although Φ_b values for NCBD correlate well in the backgrounds of ACTR WT and L1055A, Φ_b values for Leu-1055 are generally lower with different NCBD variants as compared with pWT NCBD thus displaying anti-Hammond behavior (inside *oval* in Fig. 4*b*). Interestingly, this L1055A mutation is the only one giving a high Φ_b value with pWT NCBD. Finally, we note that conservative deletion mutations and in particular Leu \rightarrow Ala, which removes three methyl groups, will yield Φ_b values that report on the tertiary interaction(s) of the side chain but also on the secondary structure, which the residue is part of, because any tertiary interaction will stabilize its associated secondary structure and vice versa.

Discussion

Coupled binding and folding of IDPs is a complex reaction involving formation of a large number of noncovalent interactions (11, 32, 33). Similarly to folding of protein domains, these reactions can experimentally appear as two state (10, 34) or as in the case of ACTR–NCBD (15, 16) and HPV E7/Rb (35), can involve detectable intermediates. In either case, the fine details of the coupled binding and folding are hard to probe with currently available kinetic instruments. However, by combining experiment with simulation, it is possible to gain a deeper understanding of the molecular events along the binding and folding pathway. We have here used restrained MD simulation in combination with mutagenesis and linear free energy relationships to better understand the binding transition state for the coupled binding and folding of ACTR and NCBD.

IDP interactions studied by Φ_b analysis differ in complexity ranging from a single β strand or α helix to several helices (10, 31, 34, 36–41). The ACTR–NCBD interaction involves formation of three helices and a relatively large interface that is topologically complex, resulting in a multistep binding mechanism as reflected in multiple kinetic phases (16, 18, 15). Our present results suggest that the TS for binding is a structurally heterogeneous ensemble of complexes, with a large fraction of weakly populated native and non-native interface contacts, the most native-like regions being helix 1 of ACTR and helix 3 of NCBD. Of notice, a former computational study based on a simplified structure-based model suggested a disordered transition state for the ACTR–NCBD complex associated with an induced-fit mechanisms but could not provide a description of the TS including non-native interactions (42). Mutagenesis based on the structure of the TS indicates that many of these native and non-native interdomain interactions contribute little or nothing in energetic terms to the binding and folding reaction reinforcing a plastic TS. Brønsted plots and Φ_b analysis suggests that the total amount of structure in the TS is generally robust, but that specific contacts formed in the TS can change upon mutation of the partner protein. Furthermore, some highly destabilizing mutants suggest that the TS might display Hammond behavior, *i.e.* that the position and structure of the TS approaches the native state when the ACTR–NCBD complex is destabilized by mutation. The Hammond effect is well-established in physical–organic chemistry (43) and has been observed in protein folding (30) but challenged because an apparent TS movement could result from changes in the ground state structures upon mutation (44). In the present case,

although the ACTR–NCBD complex tolerates most mutations, some might slightly perturb both the native free and bound states and the TS. Indeed, for a coupled binding and folding reaction with plastic TS and native structures, such movement of TS position and ground state structure appears reasonable and could be a general property whereby IDPs can accommodate mutations more easily than folded proteins (45).

Overall, our data support a model in which the initial events in coupled binding and folding reactions are not following obligatory pathways but involves encounter complexes with large structural variation (46). Thus, most native interactions are formed in an induced fit type mechanism after the rate-limiting barrier as previously suggested for ACTR–NCBD (18, 42, 47), and where intermediates may be detected by kinetics under certain conditions (16). A disordered TS does not require the free states to sample differently ordered “bound-like” states interconverting by extensive conformational changes. Still, such conformational selection may occur in shorter segments like helix 1 for ACTR and helix 3 for NCBD. However, although our kinetic and computational data show that these helices are partially formed in the TS, we cannot resolve the order of events (19), and it is likely that both disordered and ordered conformations of the helices can form the initial encounter complex as suggested for other IDP interactions (32, 48). In a broader context, our data provide support for the emerging view that structural heterogeneity is essential for many protein–protein interactions, with implications for the binding pathway, as well as for the most stable complex (49, 50).

Experimental procedures

Determination of the transition state ensembles

Molecular dynamics simulations were performed using the Amber03w force field (51) with the TIP4P/05 water model (52). All the simulations were run in GROMACS (53) using PLUMED2 (54). The van der Waals and Coulomb interactions were implemented with a cutoff at 0.9 nm, and long-range electrostatic effects were treated with the particle mesh Ewald method. Simulations were carried out in the canonical ensemble by thermostating the system using a stochastic velocity rescaling (55). The starting conformation was taken from an available NMR structure (Protein Data Bank code 1KBH) and modified to account for the Y2108W mutation using SCWRL4 (56). The structure was solvated in ~ 7200 water molecules. A standard 298 K simulation, 150 ns long, was performed as a reference for the native state ensemble.

Binding TS ensembles were determined following a previously described procedure based on the interpretation of Φ value analysis in terms of fraction of native contacts (20–22). Given a set of experimental Φ values, a pseudo energy term is added to a force field as the squared difference between experimental and simulated Φ values. This restraint provides an effective potential that can in principle account for multiple effects like the ionic strength. The Φ value for a residue is calculated from the fraction of native contacts that it makes in a conformation. Given two residues that are not nearest neighbors, the native contacts between them are defined as the number of heavy side-chain atoms within 0.65 nm in the native

structure. With this approach only Φ values between 0 and 1 can be incorporated as structural restraints. The underlying hypothesis is that the average structure satisfying all Φ values can be a good representation, at least at the residue-level resolution, of the TS. This hypothesis, even if not necessarily true, has previously proved powerful by providing testable predictions (22, 57–60). The TS ensemble was generated using simulated annealing, 1340 annealing cycles, 150 ps long, in which the temperature is varied between 298 and 383 K and performed for a nominal simulation time of 200 ns. Only the structures sampled at the reference temperature for the latter 1000 cycles are retained for further analysis, resulting in a TS ensemble of 1000 structures. The sampling of the conformational space is robust given its heterogeneity. The pair-wise RMSD support the notion that a number of uncorrelated structures are sampled. Furthermore, a cluster analysis on the first 500 structures or including all the structures shows a converged number of clusters (Fig. S5).

Protein expression and purification

Site-directed mutants of ACTR and the pseudo-WT NCBD_{Y2108W} were generated by whole plasmid mutagenesis (QuikChange), using a complementary primer pair with basepair mismatches at the mutation site and flanked by 15 bp on each side. Protein expression and purification were performed as previously described (16). Protein concentrations were determined by absorbance at 280 nm for protein variants that contained Tyr or Trp residues. To determine the concentration of the ACTR variants, which are lacking aromatic residues, absorbance at 205 nm was measured, and an extinction coefficient of 250,000 cm⁻¹ M⁻¹ was used to estimate the protein concentration.

Kinetic experiments

The kinetic experiments were conducted on an SX-17MV stopped-flow spectrophotometer. The excitation wavelength was set to 280 nm and a 320-nm-long pass filter was used to detect the emitted light. All experiments were performed at 277 K in 20 mM sodium phosphate (pH 7.4), 150 mM NaCl with or without 0.7 M TMAO. It has been shown previously that NCBD remains monomeric in presence of TMAO (12). Typically, the binding kinetics were measured using 1 μ M of the NCBD_{Y2108W} pseudo-WT and varying the concentration of ACTR between 1 and 12 μ M. In addition, the binding kinetics of the ACTR mutants were measured using 1 μ M of the respective ACTR mutant and varying concentrations of the NCBD_{Y2108W} pseudo-WT. The kinetic traces (each trace typically an average of three traces) were fitted to either a single exponential or double exponential function using the Applied Photophysics software to obtain the observed rate constant (k_{obs}) at each ACTR or NCBD concentration. The resulting plots of k_{obs} values against different ACTR or NCBD concentrations were fitted using KaleidaGraph *versus* 4.5 (Synergy Software) to the solution for reversible bimolecular binding to obtain the association rate constant (k_{on}) for each variant (61). The dissociation rate constants (k_{off}) were measured in displacement experiments, where WT NCBD was used to displace the different NCBD variants from the complexes. The concentration of pro-

teins in the complex was 1 μ M NCBD and 1 μ M ACTR. At high concentrations of WT NCBD, k_{obs} reaches a plateau, and the mean k_{obs} value from the plateau region is approximated as k_{off} . The standard deviation of the average from these points is taken as the error for k_{off} . All kinetic data are provided as a [supporting Excel file](#).

CD spectroscopy

CD spectra were measured on a JASCO J-1500 CD spectrophotometer between 200 and 260 nm. The buffer conditions were 20 mM NaP_i (pH 7.4), 150 mM NaCl with or without 0.7 M TMAO, and the temperature for all measurements was 277 K. The bandwidth was set to 0.1 nm, scanning speed was 50 nm/min, and data points were sampled every 0.1 nm. Each spectrum is an average of three accumulations.

Author contributions—E. K., P. J., and C. C. data curation; E. K., J. D., S. G., P. J., and C. C. formal analysis; E. K., E. A., and C. C. investigation; E. K., P. J., and C. C. visualization; E. K., E. A., and C. C. methodology; E. K., P. J., and C. C. writing—original draft; E. K., J. D., S. G., P. J., and C. C. writing—review and editing; S. G., P. J., and C. C. conceptualization; P. J. and C. C. resources; P. J. supervision; P. J. funding acquisition; P. J. project administration.

References

- Gianni, S., Dogan, J., and Jemth, P. (2016) Coupled binding and folding of intrinsically disordered proteins: what can we learn from kinetics? *Curr. Opin. Struct. Biol.* **36**, 18–24 [CrossRef Medline](#)
- Uversky, V. N., and Dunker, A. K. (2012) Multiparametric analysis of intrinsically disordered proteins: looking at intrinsic disorder through compound eyes. *Anal. Chem.* **84**, 2096–2104 [CrossRef Medline](#)
- Abkevich, V. I., Gutin, A. M., and Shakhnovich, E. I. (1994) Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry* **33**, 10026–10036 [CrossRef Medline](#)
- Abkevich, V. I., Gutin, A. M., and Shakhnovich, E. I. (1995) Impact of local and non-local interactions on thermodynamics and kinetics of protein folding. *J. Mol. Biol.* **252**, 460–471 [CrossRef Medline](#)
- Fersht, A. R. (1995) Optimization of rates of protein folding: the nucleation–condensation mechanism and its implications. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10869–10873 [CrossRef Medline](#)
- Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation–condensation mechanism for protein folding. *J. Mol. Biol.* **254**, 260–288 [CrossRef Medline](#)
- Karplus, M., and Weaver, D. L. (1976) Protein-folding dynamics. *Nature* **260**, 404–406 [CrossRef Medline](#)
- Cohen, F. E., Sternberg, M. J., Phillips, D. C., Kuntz, I. D., and Kollman, P. A. (1980) A diffusion–collision–adhesion model for the kinetics of myoglobin refolding. *Nature* **286**, 632–634 [CrossRef Medline](#)
- Bashford, D., Weaver, D. L., and Karplus, M. (1984) Diffusion–collision model for the folding kinetics of the λ -repressor operator-binding domain. *J. Biomol. Struct. Dyn.* **1**, 1243–1255 [CrossRef Medline](#)
- Giri, R., Morrone, A., Toto, A., Brunori, M., and Gianni, S. (2013) Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14942–14947 [CrossRef Medline](#)
- 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 [CrossRef Medline](#)
- 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 [CrossRef Medline](#)

A heterogeneous transition state for an IDP interaction

13. Kjaergaard, M., Andersen, L., Nielsen, L. D., and Teilum, K. (2013) A folded excited state of ligand-free nuclear coactivator binding domain (NCBD) underlies plasticity in ligand recognition. *Biochemistry* **52**, 1686–1693 [CrossRef Medline](#)
14. Dogan, J., Toto, A., Andersson, E., Gianni, S., and Jemth, P. (2016) Activation barrier-Limited folding and conformational sampling of a dynamic protein domain. *Biochemistry* **55**, 5289–5295 [CrossRef Medline](#)
15. Zosel, F., Mercadante, D., Nettels, D., and Schuler, B. (2018) A proline switch explains kinetic heterogeneity in a coupled folding and binding reaction. *Nat. Commun.* **9**, 3332 [CrossRef Medline](#)
16. Dogan, J., Schmidt, T., Mu, X., Engström Å., Jemth, P. (2012) Fast association and slow transitions in the interaction between two intrinsically disordered protein domains. *J. Biol. Chem.* **287**, 34316–34324 [CrossRef Medline](#)
17. Sturzenegger, F., Zosel, F., Holmstrom, E. D., Buholzer, K. J., Makarov, D. E., Nettels, D., and Schuler, B. (2018) Transition path times of coupled folding and binding reveal the formation of an encounter complex. *Nat. Commun.* **9**, 4708 [CrossRef Medline](#)
18. Dogan, J., Mu, X., Engström Å., Jemth, P. (2013) The transition state structure for coupled binding and folding of disordered protein domains. *Sci. Rep.* **3**, 2076 [CrossRef Medline](#)
19. Ieșmăntăvičius, V., Dogan, J., Jemth, P., Teilum, K., and Kjaergaard, M. (2014) Helical propensity in an intrinsically disordered protein accelerates ligand binding. *Angew. Chem. Int. Ed. Engl.* **53**, 1548–1551 [CrossRef Medline](#)
20. Vendruscolo, M., Paci, E., Dobson, C. M., and Karplus, M. (2001) Three key residues form a critical contact network in a protein folding transition state. *Nature* **409**, 641–645 [CrossRef Medline](#)
21. Paci, E., Vendruscolo, M., Dobson, C. M., and Karplus, M. (2002) Determination of a transition state at atomic resolution from protein engineering data. *J. Mol. Biol.* **324**, 151–163 [CrossRef Medline](#)
22. Toto, A., Camilloni, C., Giri, R., Brunori, M., Vendruscolo, M., and Gianni, S. (2016) Molecular recognition by templated folding of an intrinsically disordered protein. *Sci. Rep.* **6**, 21994 [CrossRef Medline](#)
23. Fersht, A. R., and Sato, S. (2004) Phi-value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7976–7981 [CrossRef Medline](#)
24. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 [CrossRef Medline](#)
25. Parra, R. G., Schafer, N. P., Radusky, L. G., Tsai, M.-Y., Guzovsky, A. B., Wolynes, P. G., and Ferreiro, D. U. (2016) Protein Frustratometer 2: a tool to localize energetic frustration in protein molecules, now with electrostatics. *Nucleic Acids Res.* **44**, W356–W360 [CrossRef Medline](#)
26. Jemth, P., Mu, X., Engström Å., Dogan, J. (2014) A frustrated binding interface for intrinsically disordered proteins. *J. Biol. Chem.* **289**, 5528–5533 [CrossRef Medline](#)
27. Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M., and Smith, M. (1982) Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding. *Nature* **299**, 756–758 [CrossRef Medline](#)
28. Leffler, J. E. (1953) Parameters for the description of transition states. *Science* **117**, 340–341 [CrossRef Medline](#)
29. Toney, M. D., and Kirsch, J. F. (1989) Direct Brønsted analysis of the restoration of activity to a mutant enzyme by exogenous amines. *Science* **243**, 1485–1488 [CrossRef Medline](#)
30. Matouschek, A., and Fersht, A. R. (1993) Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7814–7818 [CrossRef Medline](#)
31. Haq, S. R., Chi, C. N., Bach, A., Dogan, J., Engström Å., Hultqvist, G., Karlsson, O. A., Lundström, P., Montemiglio, L. C., Stromgaard, K., Gianni, S., and Jemth, P. (2012) Side-chain interactions form late and cooperatively in the binding reaction between disordered peptides and PDZ domains. *J. Am. Chem. Soc.* **134**, 599–605 [CrossRef Medline](#)
32. Schneider, R., Maurin, D., Communie, G., Kragelj, J., Hansen, D. F., Ruigrok, R. W., Jensen, M. R., and Blackledge, M. (2015) Visualizing the molecular recognition trajectory of an intrinsically disordered protein using multinuclear relaxation dispersion NMR. *J. Am. Chem. Soc.* **137**, 1220–1229 [CrossRef Medline](#)
33. Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 [CrossRef Medline](#)
34. Rogers, J. M., Oleinikovas, V., Shammass, S. L., Wong, C. T., De Sancho, D., Baker, C. M., and Clarke, J. (2014) Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. *Proc. Natl. Acad. Sci.* **111**, 15420–15425 [CrossRef Medline](#)
35. Chemes, L. B., Sánchez, I. E., and de Prat-Gay, G. (2011) Kinetic recognition of the retinoblastoma tumor suppressor by a specific protein target. *J. Mol. Biol.* **412**, 267–284 [CrossRef Medline](#)
36. Toto, A., and Gianni, S. (2016) Mutational analysis of the binding-induced folding reaction of the mixed-lineage leukemia protein to the KIX domain. *Biochemistry* **55**, 3957–3962 [CrossRef Medline](#)
37. Bonetti, D., Troilo, F., Toto, A., Brunori, M., Longhi, S., and Gianni, S. (2017) Analyzing the folding and binding steps of an intrinsically disordered protein by protein engineering. *Biochemistry* **56**, 3780–3786 [CrossRef Medline](#)
38. Dahal, L., Kwan, T. O. C., Shammass, S. L., and Clarke, J. (2017) pKID Binds to KIX via an unstructured transition state with nonnative interactions. *Biophys. J.* **113**, 2713–2722 [CrossRef Medline](#)
39. Lindström, I., and Dogan, J. (2017) Native hydrophobic binding interactions at the transition state for association between the TAZ1 domain of CBP and the disordered TAD-STAT2 are not a requirement. *Biochemistry* **56**, 4145–4153 [CrossRef Medline](#)
40. Crabtree, M. D., Mendonça, C. A. T. F., Bubb, Q. R., and Clarke, J. (2018) Folding and binding pathways of BH3-only proteins are encoded within their intrinsically disordered sequence, not templated by partner proteins. *J. Biol. Chem.* **293**, 9718–9723 [CrossRef Medline](#)
41. Lindström, I., Andersson, E., and Dogan, J. (2018) The transition state structure for binding between TAZ1 of CBP and the disordered Hif-1 α CAD. *Sci. Rep.* **8**, 7872 [CrossRef Medline](#)
42. Knott, M., and Best, R. B. (2014) Discriminating binding mechanisms of an intrinsically disordered protein via a multi-state coarse-grained model. *J. Chem. Phys.* **140**, 175102 [CrossRef Medline](#)
43. Hammond, G. S. (1955) A correlation of reaction rates. *J. Am. Chem. Soc.* **77**, 334–338 [CrossRef](#)
44. Sánchez, I. E., and Kiefhaber, T. (2003) Hammond behavior versus ground state effects in protein folding: evidence for narrow free energy barriers and residual structure in unfolded states. *J. Mol. Biol.* **327**, 867–884 [CrossRef Medline](#)
45. Hultqvist, G., Åberg, E., Camilloni, C., Sundell, G. N., Andersson, E., Dogan, J., Chi, C. N., Vendruscolo, M., and Jemth, P. (2017) Emergence and evolution of an interaction between intrinsically disordered proteins. *eLife* **6**, e16059 [Medline](#)
46. Dogan, J., Jonasson, J., Andersson, E., and Jemth, P. (2015) Binding rate constants reveal distinct features of disordered protein domains. *Biochemistry* **54**, 4741–4750 [CrossRef Medline](#)
47. 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 [CrossRef Medline](#)
48. Kragelj, J., Palencia, A., Nanao, M. H., Maurin, D., Bouvignies, G., Blackledge, M., and Jensen, M. R. (2015) Structure and dynamics of the MKK7-JNK signaling complex. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3409–3414 [CrossRef Medline](#)
49. Fuxreiter, M. (2018) Fuzziness in protein interactions: a historical perspective. *J. Mol. Biol.* **430**, 2278–2287 [CrossRef Medline](#)
50. Fuxreiter, M. (2018) Fold or not to fold upon binding: does it really matter? *Curr. Opin. Struct. Biol.* **54**, 19–25 [CrossRef Medline](#)
51. Best, R. B., and Mittal, J. (2010) Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *J. Phys. Chem. B* **114**, 14916–14923 [CrossRef Medline](#)
52. Abascal, J. L., and Vega, C. (2005) A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **123**, 234505 [CrossRef Medline](#)

53. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25
54. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. (2014) PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 [CrossRef](#)
55. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 [CrossRef](#) [Medline](#)
56. Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L. (2009) Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **77**, 778–795 [CrossRef](#) [Medline](#)
57. Gianni, S., Geierhaas, C. D., Calosci, N., Jemth, P., Vuister, G. W., Travanzi-Allocatelli, C., Vendruscolo, M., and Brunori, M. (2007) A PDZ domain recapitulates a unifying mechanism for protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 128–133 [CrossRef](#) [Medline](#)
58. Paci, E., Clarke, J., Steward, A., Vendruscolo, M., and Karplus, M. (2003) Self-consistent determination of the transition state for protein folding: application to a fibronectin type III domain. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 394–399 [CrossRef](#) [Medline](#)
59. Lindorff-Larsen, K., Vendruscolo, M., Paci, E., and Dobson, C. M. (2004) Transition states for protein folding have native topologies despite high structural variability. *Nat. Struct. Mol. Biol.* **11**, 443–449 [CrossRef](#) [Medline](#)
60. Gianni, S., Camilloni, C., Giri, R., Toto, A., Bonetti, D., Morrone, A., Sormanni, P., Brunori, M., and Vendruscolo, M. (2014) Understanding the frustration arising from the competition between function, misfolding, and aggregation in a globular protein. *Proc. Natl. Acad. Sci.* **111**, 14141–14146 [CrossRef](#) [Medline](#)
61. Malatesta, F. (2005) The study of bimolecular reactions under non-pseudo-first order conditions. *Biophys. Chem.* **116**, 251–256 [CrossRef](#) [Medline](#)