



## OPEN

# The transition state structure for coupled binding and folding of disordered protein domains

Jakob Dogan, Xin Mu, Åke Engström &amp; Per Jemth

Department of Medical Biochemistry and Microbiology, Uppsala University, BMC Box 582, SE-75123 Uppsala, Sweden.

SUBJECT AREAS:  
INTRINSICALLY  
DISORDERED PROTEINS  
PROTEIN FOLDING  
PROTEINS  
BIOPHYSICAL CHEMISTRYReceived  
24 April 2013Accepted  
11 June 2013Published  
25 June 2013Correspondence and  
requests for materials  
should be addressed to  
J.D. (Jakob.Dogan@  
imbim.uu.se) or P.J.  
(Per.Jemth@imbim.uu.  
se)

Intrinsically disordered proteins are abundant in the eukaryotic proteome, and they are implicated in a range of different diseases. However, there is a paucity of experimental data on molecular details of the coupled binding and folding of such proteins. Two interacting and relatively well studied disordered protein domains are the activation domain from the p160 transcriptional co-activator ACTR and the nuclear co-activator binding domain (NCBD) of CREB binding protein. We have analyzed the transition state for their coupled binding and folding by protein engineering and kinetic experiments ( $\Phi$ -value analysis) and found that it involves weak native interactions between the N-terminal helices of ACTR and NCBD, but is otherwise "disordered-like". Most native hydrophobic interactions in the interface between the two domains form later, after the rate-limiting barrier for association. Linear free energy relationships suggest a cooperative formation of native interactions, reminiscent of the nucleation-condensation mechanism in protein folding.

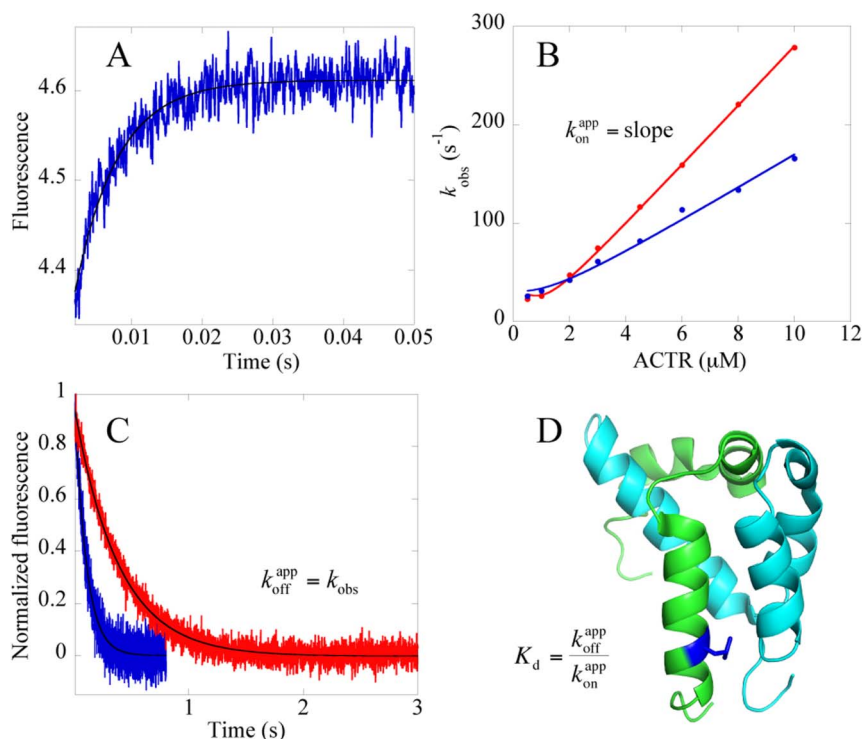
Intrinsically disordered proteins (IDPs) and disordered regions of proteins are crucial for central cellular processes such as receptor signalling, cell-cycle control, and transcription<sup>1–5</sup>. It is believed that the flexibility of IDPs gives a functional advantage, for example, it enables binding of the polypeptide in an extended conformation with a larger interface area per residue than globular proteins<sup>6</sup> and many specific contacts. The flexibility is also likely the reason that IDPs often can evolve to interact with multiple partners like, for example, p53<sup>7</sup> and NCBD<sup>8–11</sup> (see below). However, experimental data on details of the binding reaction mechanisms of IDPs are scarce and many of the hypotheses as to why disorder is beneficial remain untested by experiment.

Two disordered domains that participate in the formation of a multicomponent protein assembly that is involved in the activation and regulation of gene expression<sup>12,13</sup> are 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 (CBP). These protein domains and their interaction have been characterized in detail by NMR<sup>8,14,15</sup>, SAXS<sup>14,16</sup> and stopped-flow spectroscopy<sup>17</sup>. ACTR is highly disordered with some degree of secondary structure in the free state<sup>15,16</sup>, while NCBD has molten globule characteristics with a small hydrophobic core and low stability<sup>14,15,18,19</sup>. The bimolecular complex of the two domains displays a well-defined tertiary structure<sup>8</sup> (Fig. 1) and is a classical example of coupled binding and folding of IDPs.

To shed light on the molecular details of the coupled binding and folding reaction of ACTR and NCBD we have here used protein engineering in conjunction with kinetic experiments ( $\Phi$  value analysis<sup>20,21</sup>) to determine the structure of the transition state in terms of formation of native hydrophobic tertiary contacts in the binding interface. Our study, which is the first one reported on a system, in which one of the components is completely disordered (ACTR) and the other one has molten globule characteristics (NCBD), demonstrates that very few native interactions are in the process of forming in the transition state for binding. However, a nucleus with partial formation of native hydrophobic interactions was found to be present between the N-terminal helices of both IDP domains.

## Results

**Design of site-directed mutants.** In a previous study we characterized the kinetic binding reaction of ACTR and NCBD using engineered Trp residues as fluorescent probes<sup>17</sup>. In the current study we made ten deletion mutations (purportedly non-disruptive side chain truncations) at hydrophobic residues in ACTR<sub>WT</sub> and an additional ten in a pseudo wild type of NCBD (NCBD<sub>Y2108W</sub>). The mutations were made in the binding interface between the two domains based on a published NMR structure of the complex<sup>8</sup>. ACTR<sub>WT</sub> is highly disordered in



**Figure 1 | The experimental system.** (A) Example of an experimental binding trace for NCBD<sub>Y2108W</sub> and an ACTR mutant, L1048A. Shown here is the main (fast) phase and protein concentrations are 1  $\mu$ M NCBD<sub>Y2108W</sub> and 10  $\mu$ M L1048A-ACTR. (B)  $k_{\text{obs}}$  values for the main phase obtained from binding traces (such as in panel A) plotted versus [ACTR] to obtain association rate constants,  $k_{\text{on}}^{\text{app}}$  by fitting the data to the general equation for the reversible association of two molecules under second order conditions<sup>46</sup> (Eq. 4). At high concentrations of ACTR the equation approaches a linear function with a slope equal to  $k_{\text{on}}$ . Red, NCBD<sub>Y2108W</sub> and ACTR<sub>WT</sub>; blue, NCBD<sub>Y2108W</sub> and ACTR<sub>L1048A</sub>. (C) Dissociation rate constants,  $k_{\text{off}}^{\text{app}}$ , were determined in displacement experiments, in which NCBD<sub>WT</sub> was used to compete out NCBD<sub>Y2108W</sub> from the complex. The solid line is a fit to a single exponential equation. At high concentration of NCBD<sub>WT</sub> the dissociation of NCBD<sub>Y2108W</sub>/ACTR is virtually irreversible and  $k_{\text{obs}}$  from the curve fitting is equal to  $k_{\text{off}}^{\text{app}}$ . Red, NCBD<sub>Y2108W</sub>/ACTR<sub>WT</sub>; blue, NCBD<sub>Y2108W</sub>/ACTR<sub>L1048A</sub>. (D) Structure of the complex between ACTR (green) and NCBD (cyan), with L1048 highlighted in dark blue. The figure was generated using PyMol (The PyMol Molecular Graphics System, Version 1.3 Schrödinger, LLC).

solution<sup>8,15</sup> and mutations are not expected to disrupt any hydrophobic core. NCBD, on the other hand, displays significant residual structure as shown by NMR and circular dichroism (CD)<sup>11,14,15,19</sup>. Far-UV CD was used to assess the effect of the mutations on the secondary structure of NCBD (Supplementary Fig. S1). Six of the mutants displayed a CD spectrum identical to that of NCBD<sub>Y2108W</sub>, whereas four appeared to have lost some  $\alpha$ -helical structure. Trimethylamine *N*-oxide (TMAO) is known to shift the equilibrium towards the folded state<sup>22</sup> and 0.7 M partially or fully restored the CD spectra for these four mutants (Supplementary Fig. S1).

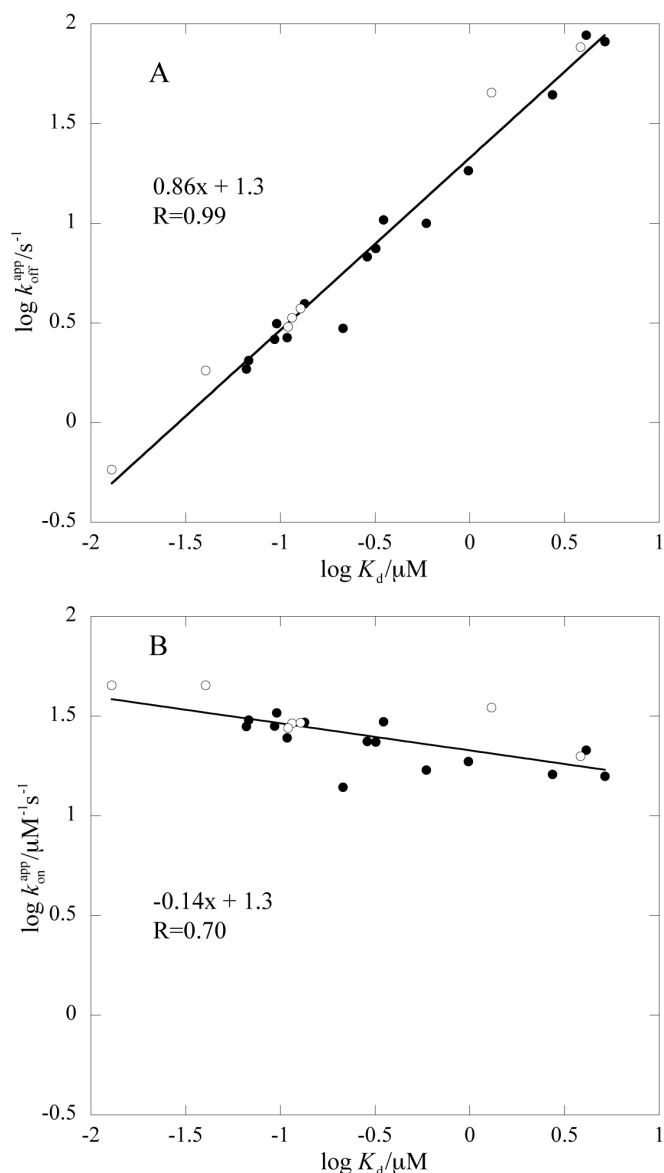
**Binding kinetics of mutants of ACTR and NCBD.** The binding kinetics of the ten ACTR mutants was measured with NCBD<sub>Y2108W</sub> and the kinetics of the ten NCBD<sub>Y2108W</sub> mutants measured with ACTR<sub>WT</sub>, using the stopped-flow technique (Fig. 1 and Supplementary Table S1, Fig. S2 and Fig. S3). Two of the mutations (L1064A and L1071A in ACTR) were highly destabilizing for the bimolecular complex and their kinetics were measured in the presence of 0.7 M TMAO to reduce the observed rate constants ( $k_{\text{obs}}$ ) to a magnitude accessible by the stopped-flow instrument. Likewise, the four mutants with altered CD spectra were measured in the presence of 0.7 M TMAO to avoid ground state effects on the observed kinetics. Two more mutants, L2090A NCBD and L1052A ACTR, were also purified but did not yield reliable kinetics data, due to elevated  $k_{\text{obs}}$  values, even in the presence of TMAO.

The binding kinetics of almost all mutants was biphasic, with a fast phase, which was linear with increasing concentration of ACTR and a slow phase, which appeared rather constant throughout the measured concentration interval. Similar kinetics was observed

previously for NCBD<sub>Y2108W</sub> and ACTR<sub>WT</sub>, and a detailed analysis has been published<sup>17</sup> (see also text in Supplementary Information for further discussion on the slow phase). From the fast phase we extracted the apparent association rate constant ( $k_{\text{on}}^{\text{app}}$ ) as the slope of the observed rate constant  $k_{\text{obs}}$  versus ACTR concentration (Fig. 1, Supplementary Table S1, Fig. S2). The apparent dissociation rate constant ( $k_{\text{off}}^{\text{app}}$ ) for mutant ACTR/NCBD complexes was determined in separate displacement experiments (Fig. 1, Supplementary Table S1) as detailed in the Materials section.

In the experiments with TMAO, an additional phase of intermediate magnitude was clearly visible for one mutant, L2067A ( $k_{\text{obs}} \sim 19 \text{ s}^{-1}$ ). Interestingly, a similar phase was previously detected for a double mutant involving a buried salt bridge, but also for the wild type at high ionic strength<sup>17</sup>. Careful analyses of binding traces for NCBD<sub>Y2108W</sub> and wild-type ACTR (i.e., our wild type pair) in 0.7 M TMAO revealed that a similar intermediate phase could be fitted to the data. This result may be explained by an induced-fit scenario, in which the population of an intermediate is promoted by TMAO, high salt, and certain mutations. Thus, TMAO and salt may be used to tune the shape of the energy landscape for coupled binding and folding for ACTR/NCBD. However, while the presence of TMAO may shift the distribution of populations of free ACTR and NCBD, and the height of energetic barriers, it is unlikely to affect the transition state structure and overall mechanism as shown by the linear free energy diagrams (Brønsted plot), in which data collected in presence and absence of TMAO fall on the same line (Fig. 2 and Supplementary Fig. S4B).

For three out of the four mutants, TMAO did not restore the magnitude of the CD signal completely, but they are close, having



**Figure 2 | Linear free energy relationships of rate and equilibrium binding constants upon mutation.** (A)  $\log k_{\text{off}}^{\text{app}}$  versus  $\log K_d$  and (B)  $\log k_{\text{on}}^{\text{app}}$  versus  $\log K_d$ . Solid circles represent data that were obtained in 20 mM phosphate (pH = 7.4), 150 mM NaCl, whereas constants that were obtained in the same buffer but supplemented with 0.7 M TMAO are shown as open circles.

90%, 88%, and 81% of the wild-type CD signal at 222 nm. However, as judged from Fig. 2, these small differences are not enough to result in a significant deviation from the overall linear trend in the Brønsted plot. Thus, for these mutants as well as the other ones measured in the presence of TMAO, the data report on the same transition state.

**Structural details of the interface between ACTR and NCBD.** The 3D structure of the complex between ACTR and NCBD<sup>8</sup> shows that the leucine rich binding interface of NCBD/ACTR is well packed and contains specific hydrophobic interactions. Accordingly, several of the conservative deletion mutations resulted in significantly reduced binding affinities (Table 1 and Supplementary Table S1), in particular those from Leu to Ala in LXXLL/LLXXL motifs. These results agree well with previous mutational studies of the interaction between NCBD and ACTR<sup>19</sup> or TIF2, an ACTR homolog<sup>23</sup>, as well as other NCBD binding proteins<sup>24,25</sup>. On the other end, the mutations V2109A NCBD and V1077A ACTR were not destabilizing but displayed even lower  $k_{\text{off}}^{\text{app}}$  and  $K_d$  values than the wild type (Supplementary Table S1). Both of these mutated residues are positioned at the C-terminal helices of the respective domain, and as shown in the NCBD/ACTR structure<sup>8</sup> they also interact with each other. While the thermodynamic origin of these changes is not clear, it has previously been shown that the C-terminal part of helix three in unbound NCBD displays significant fast backbone dynamics<sup>15</sup> compared to the rest of NCBD, and the region in bound ACTR that forms the C-terminal helix, has very little helical content in unbound ACTR<sup>15,16</sup>. One may speculate that the disorder that is present in the respective helix is modulated by mutation, for example through stabilization of the helix, such that a slightly higher affinity is obtained, compared to the wild type.

**Calculation of  $\Phi_{\text{binding}}$  values.** Linear free energy relationships clearly demonstrate that the effect of the mutations is largely in the dissociation rate constant (Fig. 2A). However, a few mutations affected the association rate constant (Supplementary Fig. S2), suggesting that some side-chains have begun to form native-like interactions in the rate limiting transition state for binding. In order to quantify the degree of native contact formation and thus get a picture of the structure of the transition state for the coupled binding and folding of ACTR/NCBD, we calculated  $\Phi$  values for binding,  $\Phi_{\text{binding}}^{21,26}$  (Table 1, see Supplementary Information for a detailed discussion of the  $\Phi_{\text{binding}}$  values in relation to  $k_{\text{on}}^{\text{app}}$ ,  $k_{\text{off}}^{\text{app}}$  and the slow phase for each plausible reaction scheme) and mapped the values onto the structure of the complex between the two proteins (Fig. 3).  $\Phi_{\text{binding}}$  values were calculated by relating the change in free energy for the rate limiting barrier for binding ( $\Delta\Delta G_{\text{TS}}$ , calculated from  $k_{\text{on}}$  values) with the total change in free energy for the binding reaction at equilibrium ( $\Delta\Delta G_{\text{Eq}}$ , calculated from  $K_d$  values, which in

**Table 1 |  $\Phi_{\text{binding}}$  values for the interaction between wild type and mutants of ACTR<sub>WT</sub> and NCBD<sub>Y2108W</sub>, respectively, measured in 20 mM phosphate (pH = 7.4), 150 mM NaCl, and 277 K**

NCBD <sub>Y2108W</sub> mutant	ACTR <sub>WT</sub>		ACTR mutant	NCBD <sub>Y2108W</sub>	
	$\Delta\Delta G_{\text{Eq}}$ kcal/mol	$\Phi_{\text{binding}}$		$\Delta\Delta G_{\text{Eq}}$ kcal/mol	$\Phi_{\text{binding}}$
I2062V $\alpha$ 1	0.20 $\pm$ 0.01	-0.12 $\pm$ 0.07	L1048A $\alpha$ 1	1.02 $\pm$ 0.02	0.27 $\pm$ 0.02
V2086A $\alpha$ 2	0.73 $\pm$ 0.02	-0.04 $\pm$ 0.02	L1049A $\alpha$ 1	1.30 $\pm$ 0.02	0.17 $\pm$ 0.02
L2087A $\alpha$ 2	2.2 $\pm$ 0.1	0.14 $\pm$ 0.03	L1055A $\alpha$ 1	0.46 $\pm$ 0.05	0.85 $\pm$ 0.10
L2096A $\alpha$ 3	1.9 $\pm$ 0.1	0.16 $\pm$ 0.03	L1056A $\alpha$ 1	2.09 $\pm$ 0.1	0.07 $\pm$ 0.02
I2101V <sup>a</sup> $\alpha$ 3	0.01 $\pm$ 0.02	-	A1061G <sup>a</sup> loop	0.08 $\pm$ 0.06	-
V2109A $\alpha$ 3	-0.17 $\pm$ 0.02	0.22 $\pm$ 0.09	I1067V $\alpha$ 2	0.62 $\pm$ 0.02	0.16 $\pm$ 0.02
			I1073V $\alpha$ 3	0.68 $\pm$ 0.03	0.15 $\pm$ 0.05
			V1077A $\alpha$ 3	-0.19 $\pm$ 0.02	-0.01 $\pm$ 0.13

<sup>a</sup>Too low  $\Delta\Delta G_{\text{Eq}}$  to calculate a reliable  $\Phi_{\text{binding}}$  value.



turn were calculated from  $k_{\text{off}}^{\text{app}}$  divided by  $k_{\text{on}}^{\text{app}}$  (Equations 1–3, Supplementary Information).

$$\Phi_{\text{binding}} = \Delta\Delta G_{\text{TS}} / \Delta\Delta G_{\text{Eq}} \quad (1)$$

$$\Delta\Delta G_{\text{TS}} = RT \ln(k_{\text{on}}^{\text{mutant}} / k_{\text{on}}^{\text{wild type}}) \quad (2)$$

$$\Delta\Delta G_{\text{Eq}} = RT \ln(k_{\text{d}}^{\text{wild type}} / k_{\text{d}}^{\text{mutant}}) \quad (3)$$

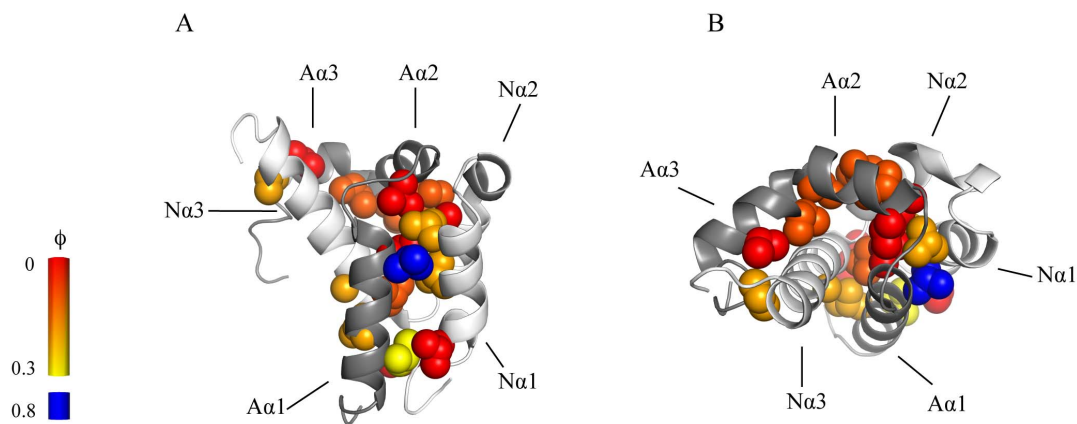
The wild type in the calculations refers to the interaction between NCBD<sub>Y2108W</sub> and ACTR<sub>WT</sub>. A  $\Phi_{\text{binding}}$  value of zero would imply that the mutated residue is not making any native interactions in the transition state of the binding reaction. A  $\Phi_{\text{binding}}$  value of 1, on the other hand, implies that the residue makes a full native interaction in the transition state. Intermediate values are subject to a number of caveats<sup>20</sup>, but are usually interpreted as partial formation of native interactions in the transition state. We calculated a  $\Phi_{\text{binding}}$  value if the absolute value of  $\Delta\Delta G_{\text{Eq}} > 0.16$  kcal/mol. This may appear as a very low cut-off value, but is based on the very accurate determinations of  $k_{\text{on}}^{\text{app}}$  and  $k_{\text{off}}^{\text{app}}$ . For most mutants, the largest source of error is the concentration determination of ACTR variants, which directly affects  $k_{\text{on}}^{\text{app}}$  and which we estimate to about 5%. The  $\Phi_{\text{binding}}$  values in Table 1 report on the transition state that is rate limiting for binding in 20 mM sodium phosphate (pH = 7.4), 150 mM NaCl. In the presence of TMAO, it is likely that the same transition state has been stabilized resulting in a higher  $k_{\text{on}}^{\text{app}}$  value. An alternative explanation is that an earlier transition state becomes (partially) rate limiting.  $k_{\text{on}}^{\text{app}}$  for NCBD<sub>Y2108W</sub> and wild-type ACTR in presence of 0.7 M TMAO, was determined to be around  $45 \mu\text{M}^{-1}\text{s}^{-1}$ , which was used to calculate  $\Phi_{\text{binding}}$  values in Table 2.

**The structure of the transition state for coupled binding and folding.** Following the guidelines of Fersht<sup>20</sup>, it is useful to categorize  $\Phi$  values as low, intermediate and high. The low  $\Phi_{\text{binding}}$  values ( $<0.2$ ) at almost all positions show that most of the native hydrophobic interactions in the interface between ACTR and NCBD have not formed in the transition state for the binding reaction, suggesting that most of the initial encounter complex does not contain native-like hydrophobic interactions. However, a few ACTR mutations situated in helix 1 make direct interactions with NCBD and displayed intermediate or even high  $\Phi_{\text{binding}}$  values. Although  $\Phi_{\text{binding}}$  values for the NCBD mutants were low, two mutations in helix 1 gave  $\Phi_{\text{binding}}$  values of 0.2, suggesting that the N-terminal helices of ACTR and NCBD form the initial native intermolecular contacts.

## Discussion

Protein engineering together with detailed kinetic analyses allowed us to determine a structural model of the rate-limiting transition state in the coupled binding and folding involving the IDP, ACTR, and the molten globule, NCBD (Fig. 3). In terms of hydrophobic interactions in the interface between the two disordered domains, the transition state for the initial association is more disordered-like than native-like. Overall, native interactions in the binding interface are formed and consolidated subsequent to the rate-limiting barrier for association. This is in agreement with previous studies on the kinetic binding mechanism of IDPs and disordered regions using NMR<sup>27</sup>, MD<sup>28</sup> or stopped-flow spectroscopy<sup>17,29–32</sup>, in which data also suggest that native interactions form late on the reaction pathway. It is usually very difficult to prove multistep binding reactions and there are a number of possible binding mechanisms for the association of ACTR and NCBD. Importantly, our  $\Phi_{\text{binding}}$  values report on the transition state for the productive binding pathway for either of the four mechanisms (Scheme 1–4 in Supplementary Fig. S5), which are overall consistent with the data set, as discussed in detail in Supplementary Information. Recent data on NCBD show that it exists in at least two different conformations in the free state<sup>33,34</sup>, which lends support to model 4 in Fig. S5. Thus, in their free states, there may be several different forms of ACTR and NCBD, with different degrees of structure and multiple pathways for binding, with external conditions determining the flux through each pathway<sup>35</sup>. However, the pathways converge after their initial association and ACTR and NCBD will experience a cooperative formation of native bonds as the barrier is crossed as shown by the linear free energy relationship (Fig. 2).

Looking at the transition state in more detail, a few native contacts have begun to form between the N-terminal helices of the respective protein in the transition state for binding. In fact, the linear free energy relationship (Fig. 2) is reminiscent of those obtained for proteins obeying the nucleation-condensation mechanism in protein folding<sup>36</sup>. Thus, following formation of a few native interactions during the crossing of the main rate-limiting barrier, the remaining native contacts form after the main barrier. Binding of IDPs often involves extended protein conformations with a binding surface containing hydrophilic as well as hydrophobic interactions. Interestingly, the region with highest  $\Phi_{\text{binding}}$  values contains the LXXLL/LLXXL recognition motifs, where X stands for any amino acid. These motifs, in which any bulky hydrophobic residue can replace Leu, are known to mediate protein-protein interactions in transcriptional regulation<sup>37</sup>.



**Figure 3 | A structural model of the transition state for coupled binding and folding.**  $\Phi_{\text{binding}}$  values from Table 1 and 2 mapped onto the structure of the complex between ACTR and NCBD. ACTR is shown in dark grey and NCBD in light grey. Residues are colour coded by the magnitude of their  $\Phi_{\text{binding}}$  values, with a gradient ranging from 0 (red) to 0.3 (yellow). ACTR<sub>L1055A</sub>, which has the highest  $\Phi_{\text{binding}}$  value, is blue colored. The negative  $\Phi_{\text{binding}}$  values were close to zero and are colored red. Panel A and B show the same structure from two different angles. Images were generated using PyMol (The PyMol Molecular Graphics System, Version 1.3 Schrödinger, LLC).



**Table 2** |  $\Phi_{\text{binding}}$  values for the interaction between wild type and mutants of ACTR<sub>WT</sub> and NCBD<sub>Y2108W</sub>, respectively, measured in 20 mM phosphate (pH = 7.4), 150 mM NaCl, 0.7 M TMAO, and 277 K

NCBD <sub>Y2108W</sub> mutant	ACTR <sub>WT</sub>		ACTR mutant	NCBD <sub>Y2108W</sub>	
	$\Delta\Delta G_{\text{Eq}}$ kcal/mol	$\Phi_{\text{binding}}$		$\Delta\Delta G_{\text{Eq}}$ kcal/mol	$\Phi_{\text{binding}}$
L2067A $\alpha$ 1	0.63 $\pm$ 0.08	0.0 $\pm$ 0.1	L1064A $\alpha$ 2	2.5 $\pm$ 0.1	0.06 $\pm$ 0.02
L2070A $\alpha$ 1	1.2 $\pm$ 0.1	0.20 $\pm$ 0.04	L1071A loop/ $\alpha$ 2	3.1 $\pm$ 0.2	0.14 $\pm$ 0.03
L2074A $\alpha$ 1	1.3 $\pm$ 0.1	0.19 $\pm$ 0.04			
A2099G $\alpha$ 3	1.2 $\pm$ 0.1	0.23 $\pm$ 0.04			

Furthermore, NMR and CD studies<sup>15,16</sup> have shown that the region that constitutes helix one in bound ACTR in fact displays transient helix formation in the free state. Previous studies<sup>38–41</sup> have emphasized the functional importance of preformed structural elements and our observation that such elements form the initial native contacts during the binding between ACTR and NCBD lends support to this hypothesis.

In our initial study on ACTR/NCBD<sup>17</sup> we characterized a buried salt bridge by mutagenesis (R2104L in NCBD and D1068L in ACTR). It is clear that the effect on the association rate constant was large (~20 fold) while the  $k_{\text{off}}^{\text{app}}$  was similar to that of the wild type proteins. This would imply a  $\Phi_{\text{binding}}$  value of one, however, the R2104L mutation has clear effects on the ground state of NCBD<sup>19</sup> and the two Leu residues that replaced the salt bridge will most probably contribute to binding by forming new hydrophobic interactions. Thus, while mutation of the salt bridge strongly affects  $k_{\text{on}}^{\text{app}}$ , suggesting that the salt bridge stabilizes the transition state for binding, we cannot estimate a  $\Phi_{\text{binding}}$  value for this interaction. Previous studies on interactions involving IDPs<sup>42</sup> or folded proteins<sup>43</sup> have demonstrated that mutation of charged residues may affect the association rate constant such that high  $\Phi_{\text{binding}}$  values are obtained. Thus, charged residues have a higher tendency than hydrophobic amino acids to form native-like interactions in the transition state for binding, due to long-range electrostatic effects<sup>21,44</sup>.

A recent molecular dynamics simulation study<sup>38</sup> concluded that the end helices  $\alpha$ 1 and  $\alpha$ 3 in NCBD, which dominate the number of contacts it makes with ACTR in the complex, are most readily formed in the free state, having a conformation similar to the bound state. Here we show that  $\alpha$ 1 in NCBD forms weak native-like interactions with ACTR in the transition state, and two out of four mutations in  $\alpha$ 3 in NCBD has an intermediate  $\Phi$  value (Table 1). It should be noted that these mutations involve hydrophobic substitutions, whereas  $\alpha$ 3 in NCBD also contains charged residues, which, as previously discussed, tend to adopt higher  $\Phi$  values than hydrophobic positions<sup>21</sup>. Thus, it is possible that this region may make long-range native-like electrostatic interactions in the rate-limiting transition state.

In summary, we propose, based on available experimental data, the following scenario for the productive interaction between ACTR and NCBD. Helix one of the highly disordered ACTR forms transiently<sup>15,16</sup>, with folding and unfolding probably on the ns- $\mu$ s time-scale<sup>45</sup>. In a helix-like state it makes initial weak native-like interactions with helix one of NCBD that are rate limiting for the binding reaction. The initial association is also promoted by R2104 in NCBD and D1068 in ACTR<sup>17</sup>, which may steer the two protein domains into the correct orientation as the native salt bridge is formed. Formation of native hydrophobic contacts in the binding interface proceeds in a cooperative fashion following crossing of the rate-limiting transition state for association. It is likely that this mechanism is common for coupled binding and folding reactions of IDPs.

## Methods

**Protein expression and purification.** Human NCBD and ACTR were expressed and purified as described previously<sup>17</sup>. Briefly, BL21(DE3) pLysS cells were used to express

NCBD or ACTR in fusion with an N-terminal His-tagged lipoyl fusion protein with a thrombin site separating the lipoyl and NCBD or ACTR sequences. The fusion protein was first subjected to a Ni-sepharose fast flow (GE Healthcare) column purification step followed by thrombin cleavage, after which a second Ni-sepharose fast flow purification step was utilized to remove the lipoyl protein and other impurities. This was followed by reversed phase chromatography using a C-8 (Grace Davison Discovery) column. Mutations were generated by inverted PCR and mutants were expressed and purified as described for the wild-type proteins<sup>17</sup>. The identity of each mutant was verified with MALDI-TOF. The concentration of NCBD was determined by measuring the absorbance at 280 nm, whereas the concentration of ACTR, which lacks Tyr and Trp residues, was determined by measuring the absorbance at 205 nm and using an extinction coefficient of 31.9 mL mg<sup>-1</sup> cm<sup>-1</sup> obtained from amino acid analysis.

**Binding kinetics.** Binding kinetics of NCBD/ACTR was measured using an upgraded SX-17MV stopped-flow spectrometer (Applied Photophysics, Leatherhead, U.K.) and monitoring Trp fluorescence<sup>17</sup>. Excitation was at 280 nm and a 320 nm long pass filter was used to monitor the emission. Measurements were done at 277 K in either 20 mM sodium phosphate (pH = 7.4), 150 mM NaCl, or the same buffer supplemented with 0.7 M TMAO. In order to obtain association rate constants ( $k_{\text{on}}^{\text{app}}$ ) for NCBD<sub>Y2108W</sub>/ACTR<sup>mutant</sup> and NCBD<sub>Y2108W</sub><sup>mutant</sup>/ACTR<sub>WT</sub>, the concentration of ACTR was varied while keeping the concentration of NCBD constant at 1  $\mu$ M. Observed rate constants were plotted versus [ACTR] and the data were fitted to the general equation for the reversible association of two molecules<sup>46</sup> to determine  $k_{\text{on}}^{\text{app}}$  (Fig. 1) (Eq. 4).

$$k_{\text{obs}} = \left( k_{\text{on}}^2 ([\text{NCBD}]_0 - [\text{ACTR}]_0)^2 + k_{\text{off}}^2 + 2k_{\text{on}}k_{\text{off}} ([\text{NCBD}]_0 + [\text{ACTR}]_0) \right)^{0.5} \quad (4)$$

[NCBD]<sub>0</sub> and [ACTR]<sub>0</sub> are the total concentrations of the respective species. Note that the equation breaks down to a linear function under pseudo-first order conditions, that is when [ACTR]<sub>0</sub>  $\gg$  [NCBD]<sub>0</sub>.

Apparent dissociation rate constants were determined through displacement experiments, by mixing a pre-formed NCBD<sub>Y2108W</sub>/ACTR<sup>mutant</sup> or NCBD<sub>Y2108W</sub><sup>mutant</sup>/ACTR<sub>WT</sub> complex (1.1–2.5  $\mu$ M NCBD mixed with 1–2  $\mu$ M ACTR) with an excess of [NCBD<sub>WT</sub>], such that ACTR is trapped in a spectroscopically silent complex. Higher concentration of mutant NCBD/ACTR complex was used for mutants with low affinity, to improve the amplitude of the kinetic trace. The observed rate constant at high [NCBD<sub>WT</sub>] is equal to  $k_{\text{off}}^{\text{app}}$ .

**Circular dichroism spectroscopy.** CD spectra were recorded using a JASCO-810 spectrometer with a peltier temperature control system. Spectra were recorded between 260 nm and 200 nm, at 298 K or 277 K, and 20 mM phosphate (pH = 7.4), 150 mM NaCl, either with or without 0.7 M TMAO.

- Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M. & Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry* **41**, 6573–82 (2002).
- Wright, P. E. & Dyson, H. J. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 (2005).
- Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**, 527–533 (2002).
- Liu, J. *et al.* Intrinsic disorder in transcription factors. *Biochemistry* **45**, 6873–6888 (2006).
- Dyson, H. J. Expanding the proteome: disordered and alternatively folded proteins. *Q. Rev. Biophys.*, 1–52 (2011).
- Meszaros, B., Tompa, P., Simon, I. & Dosztanyi, Z. Molecular principles of the interactions of disordered proteins. *J. Mol. Biol.* **372**, 549–561 (2007).
- Vousden, K. H. & Lane, D. P. p53 in health and disease. *Nat. Rev. Mol. Cell Biol.* **8**, 275–283 (2007).
- Demarest, S. J. *et al.* Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553 (2002).
- Waters, L. *et al.* Structural diversity in p160/CREB-binding protein coactivator complexes. *J. Biol. Chem.* **281**, 14787–14795 (2006).
- Qin, B. Y. *et al.* Crystal structure of IRF-3 in complex with CBP. *Structure* **13**, 1269–1277 (2005).



11. Lee, C. W., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry* **49**, 9964–9971 (2010).
12. Chen, H. *et al.* Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell* **90**, 569–580 (1997).
13. Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14**, 1553–1577 (2000).
14. Kjaergaard, M., Teilum, K. & Poulsen, F. M. Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. USA* **107**, 12535–12540 (2010).
15. Ebert, M. O., Bae, S. H., Dyson, H. J. & Wright, P. E. NMR relaxation study of the complex formed between CBP and the activation domain of the nuclear hormone receptor coactivator ACTR. *Biochemistry* **47**, 1299–1308 (2008).
16. Kjaergaard, M. *et al.* Temperature-dependent structural changes in intrinsically disordered proteins: formation of alpha-helices or loss of polyproline II? *Protein Sci.* **19**, 1555–1564 (2010).
17. Dogan, J., Schmidt, T., Mu, X., Engström, Å. & Jemth, P. Fast Association and Slow Transitions in the Interaction between Two Intrinsically Disordered Protein Domains. *J. Biol. Chem.* **287**, 34316–34324 (2012).
18. Kjaergaard, M., Poulsen, F. M. & Teilum, K. Is a malleable protein necessarily highly dynamic? The hydrophobic core of the nuclear coactivator binding domain is well ordered. *Biophys. J.* **102**, 1627–1635 (2012).
19. Demarest, S. J., Deechongkit, S., Dyson, H. J., Evans, R. M. & Wright, P. E. Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. *Protein Sci.* **13**, 203–210 (2004).
20. Fersht, A. R. & Sato, S. Phi-value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. USA* **101**, 7976–7981 (2004).
21. Schreiber, G., Haran, G. & Zhou, H. X. Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* **109**, 839–860 (2009).
22. Baskakov, I. & Bolen, D. W. Forcing thermodynamically unfolded proteins to fold. *J. Biol. Chem.* **273**, 4831–4834 (1998).
23. Voegel, J. J. *et al.* The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways. *EMBO J.* **17**, 507–519 (1998).
24. Matsuda, S. *et al.* A conserved alpha-helical motif mediates the binding of diverse nuclear proteins to the SRC1 interaction domain of CBP. *J. Biol. Chem.* **279**, 14055–14064 (2004).
25. Ryan, C. M., Harries, J. C., Kindle, K. B., Collins, H. M. & Heery, D. M. Functional interaction of CREB binding protein (CBP) with nuclear transport proteins and modulation by HDAC inhibitors. *Cell cycle* **5**, 2146–2152 (2006).
26. Taylor, M. G., Rajpal, A. & Kirsch, J. F. Kinetic epitope mapping of the chicken lysozyme. HyHEL-10 Fab complex: delineation of docking trajectories. *Protein Sci.* **7**, 1857–1867 (1998).
27. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1026 (2007).
28. Turjanski, A. G., Gutkind, J. S., Best, R. B. & Hummer, G. Binding-induced folding of a natively unstructured transcription factor. *PLoS Comp. Biol.* **4**, e1000060 (2008).
29. Haq, S. R. *et al.* Side-chain interactions form late and cooperatively in the binding reaction between disordered peptides and PDZ domains. *J. Am. Chem. Soc.* **134**, 599–605 (2012).
30. Karlsson, O. A., Chi, C. N., Engström, Å. & Jemth, P. The transition state of coupled folding and binding for a flexible beta-finger. *J. Mol. Biol.* **417**, 253–261 (2012).
31. Bachmann, A., Wildemann, D., Praetorius, F., Fischer, G. & Kiefhaber, T. Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proc. Natl. Acad. Sci. USA* **108**, 3952–3957 (2011).
32. Gianni, S., Morrone, A., Giri, R. & Brunori, M. A folding-after-binding mechanism describes the recognition between the transactivation domain of c-Myb and the KIX domain of the CREB-binding protein. *Biochem. Biophys. Res. Commun.* **428**, 205–209 (2012).
33. Kjaergaard, M., Andersen, L., Nielsen, L. D. & Teilum, K. A Folded Excited State of Ligand-Free Nuclear Coactivator Binding Domain (NCBD) Underlies Plasticity in Ligand Recognition. *Biochemistry* **In press**, doi:10.1021/bi4001062 (2013).
34. Japrun, D. *et al.* Single-molecule studies of intrinsically disordered proteins using solid-state nanopores. *Anal. Chem.* **85**, 2449–2456 (2013).
35. Hammes, G. G., Chang, Y. C. & Oas, T. G. Conformational selection or induced fit: a flux description of reaction mechanism. *Proc. Natl. Acad. Sci. U S A* **106**, 13737–13741 (2009).
36. Itzhaki, L. S., Otzen, D. E. & Fersht, A. R. 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 (1995).
37. Plevin, M. J., Mills, M. M. & Ikura, M. The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends Biochem. Sci.* **30**, 66–69 (2005).
38. Knott, M. & Best, R. B. A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: evidence from molecular simulations. *PLoS Comp. Biol.* **8**, e1002605 (2012).
39. Fuxreiter, M., Simon, I., Friedrich, P. & Tompa, P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J. Mol. Biol.* **338**, 1015–1026 (2004).
40. Sivakolundu, S. G., Bashford, D. & Kriwacki, R. W. Disordered p27Kip1 exhibits intrinsic structure resembling the Cdk2/cyclin A-bound conformation. *J. Mol. Biol.* **353**, 1118–1128 (2005).
41. Cino, E. A., Wong-ekkkabut, J., Karttunen, M. & Choy, W. Y. Microsecond molecular dynamics simulations of intrinsically disordered proteins involved in the oxidative stress response. *PLOS ONE* **6**, e27371 (2011).
42. Hemsath, L., Dvorsky, R., Fiegen, D., Carlier, M. F. & Ahmadian, M. R. An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. *Mol. Cell* **20**, 313–324 (2005).
43. Selzer, T., Albeck, S. & Schreiber, G. Rational design of faster associating and tighter binding protein complexes. *Nat. Struct. Biol.* **7**, 537–541 (2000).
44. Kiel, C., Selzer, T., Shaul, Y., Schreiber, G. & Herrmann, C. Electrostatically optimized Ras-binding Ral guanine dissociation stimulator mutants increase the rate of association by stabilizing the encounter complex. *Proc. Natl. Acad. Sci. USA* **101**, 9223–9228 (2004).
45. Neuweiler, H., Johnson, C. M. & Fersht, A. R. Direct observation of ultrafast folding and denatured state dynamics in single protein molecules. *Proc. Natl. Acad. Sci. USA* **106**, 18569–18574 (2009).
46. Malatesta, F. The study of bimolecular reactions under non-pseudo-first order conditions. *Biophys. Chem.* **116**, 251–256 (2005).

## Acknowledgements

This work was supported by the Swedish Research Council (NT) and the Human Frontiers Young Investigator Science Program (to P. J.) and the Lars Hierta Foundation (to J. D.) and the Magnus Bergvall Foundation (to J. D.).

## Author contributions

J.D. and P.J. planned and designed the experiments. J.D., X.M. and Å.E. performed the experiments. J.D. and P.J. analyzed data and wrote the paper.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Dogan, J., Mu, X., Engström, Å. & Jemth, P. The transition state structure for coupled binding and folding of disordered protein domains. *Sci. Rep.* **3**, 2076; DOI:10.1038/srep02076 (2013).



This work is licensed under a Creative Commons Attribution 3.0 Unported license. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0>