

ARTICLE

DOI: 10.1038/s41467-018-05725-0

OPEN

# A proline switch explains kinetic heterogeneity in a coupled folding and binding reaction

Franziska Zosel <sup>1,3</sup>, Davide Mercadante<sup>1</sup>, Daniel Nettels<sup>1</sup> & Benjamin Schuler <sup>1,2</sup>

The interactions of intrinsically disordered proteins (IDPs) with their molecular targets are essential for the regulation of many cellular processes. IDPs can perform their functions while disordered, and they may fold to structured conformations on binding. Here we show that the *cis/trans* isomerization of peptidyl–prolyl bonds can have a pronounced effect on the interactions of IDPs. By single-molecule spectroscopy, we identify a conserved proline residue in NCBD (the nuclear-coactivator binding domain of CBP) whose *cis/trans* isomerization in the unbound state modulates the association and dissociation rates with its binding partner, ACTR. As a result, NCBD switches on a time scale of tens of seconds between two populations that differ in their affinities to ACTR by about an order of magnitude. Molecular dynamics simulations indicate as a cause reduced packing of the complex for the *cis* isomer. Peptidyl-prolyl *cis/trans* isomerization may be an important previously unidentified mechanism for regulating IDP interactions.

<sup>1</sup>Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. <sup>2</sup>Department of Physics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. <sup>3</sup>Present address: Novo Nordisk A/S, Novo Nordisk Park 1, 2760 Måløv, Denmark. Correspondence and requests for materials should be addressed to B.S. (email: [schuler@bioc.uzh.ch](mailto:schuler@bioc.uzh.ch))

Intrinsically disordered proteins (IDPs) and disordered regions are abundant in the eukaryotic proteome. They are particularly prevalent in regulatory processes such as signal transduction, transcription, and many other complex interaction networks<sup>1–3</sup>. Several explanations have been put forward as to why disorder can be beneficial. For instance, structural flexibility enables the interaction with multiple binding partners<sup>4</sup> and facilitates interactions with both high association and dissociation rates<sup>5</sup>, rendering signaling processes more rapid. Moreover, disordered segments are easily accessible for post-translational modifications that modulate interactions and the corresponding signaling pathways<sup>6</sup>.

The conformational heterogeneity underlying the properties of IDPs results from structurally diverse ensembles that are sampled on a broad range of timescales<sup>7–9</sup>. The fastest long-range dynamics occur in the sub-microsecond range and correspond to the reconfiguration of the disordered polypeptide chain, sometimes impeded by residual non-native interactions causing internal friction<sup>10</sup>. Transient secondary structure formation typically occurs on the microsecond time scale<sup>11</sup>, and more persistent long-range interactions in IDPs can lead to dynamics in the millisecond range<sup>12</sup>. All of these dynamics are expected to be faster than the typical kinetics of binding at cellular protein concentrations<sup>13,14</sup>. As a result, it is usually reasonable to assume a separation of timescales between the internal dynamics of an IDP and its association and dissociation kinetics. In this case, the binding kinetics can be approximated by a two-state process with single-exponential relaxation, as frequently observed experimentally<sup>14–16</sup>.

There is, however, a source of slow conformational dynamics in IDPs whose influence on binding kinetics has hardly been investigated, but that is known to occur on timescales of seconds to minutes: the *cis/trans* isomerization of peptide bonds involving proline residues<sup>17,18</sup>. Proline isomerization is a classic source of slow phases in protein folding kinetics<sup>19</sup>, and given that Pro is the amino acid most highly enriched in IDPs compared to folded proteins<sup>20</sup>, it may be expected to influence the behavior of many IDPs and cause heterogeneity in their association and dissociation kinetics. Here we use single-molecule Förster resonance energy transfer (FRET) to probe the kinetic heterogeneity and its structural origin for the interaction between the nuclear-coactivator binding domain (NCBD) of the CBP/p300 transcription coactivator and the activation domain of SRC-3 (ACTR). NCBD and ACTR are a classic example of two IDPs that undergo a coupled folding and binding reaction in which both proteins gain structure and form a cooperatively folded core in the complex<sup>4,21</sup>. Unbound ACTR contains only little residual secondary structure<sup>22</sup>. In contrast, unbound NCBD is a marginally stable, molten-globule-like protein, with a large content of  $\alpha$ -helical structure<sup>23,24</sup>, and different arrangements of the helices have been observed in complexes with its different binding partners<sup>21,25–27</sup>. We find that NCBD occupies two distinct subpopulations at equilibrium that differ by the conformation of one proline residue. Both the *trans*- and the *cis*-Pro populations are able to bind ACTR, but with different affinities and kinetics, thus revealing a potentially widespread source of slow dynamics and heterogeneity in IDP interactions that may have an important role in regulation.

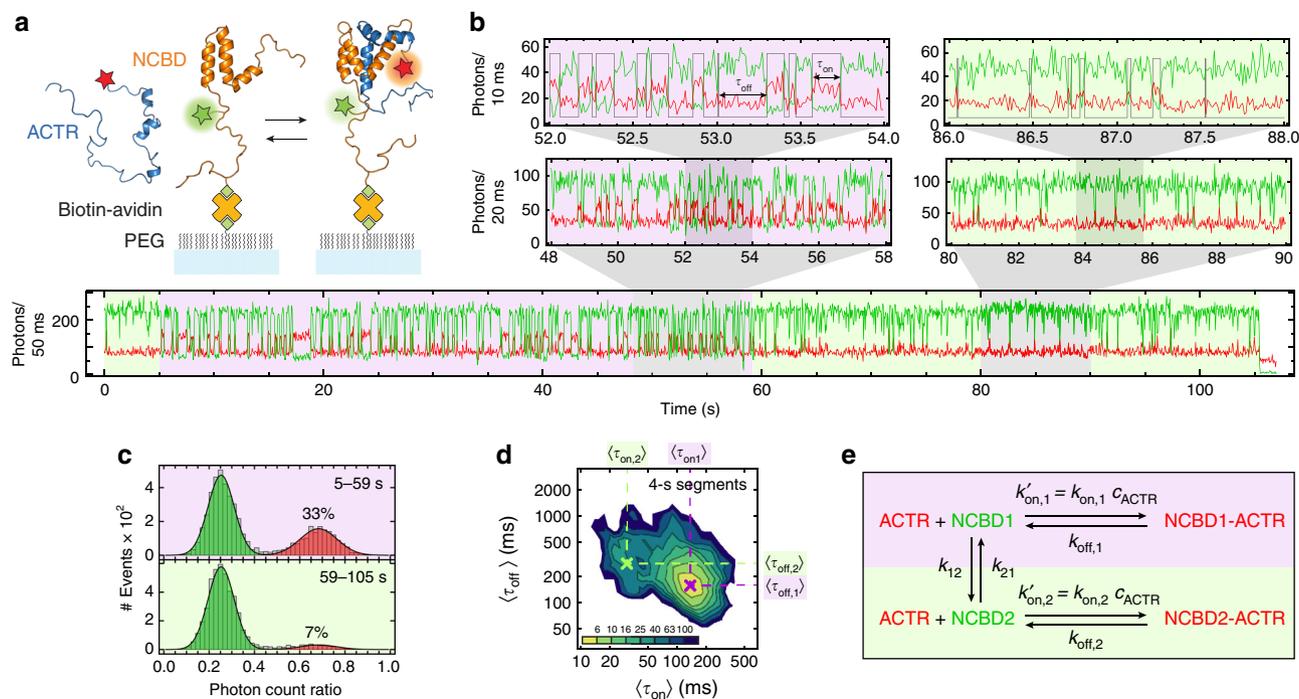
## Results

**Single-molecule FRET of immobilized molecules.** Investigating heterogeneous kinetics that span a broad range of timescales with single-molecule spectroscopy requires individual molecules to be observed for extended periods of time with high time resolution. We thus immobilized biotinylated NCBD labeled with Cy3B as a

donor fluorophore on a polyethylene glycol- (PEG-) passivated glass surface via avidin (see Methods) and recorded fluorescence emission from single molecules by confocal single-photon counting. Including low concentrations of ACTR (labeled with CF680R as an acceptor fluorophore) free in solution allowed us to monitor association and dissociation events via FRET in great detail (Fig. 1a). In the absence of bound ACTR, only donor emission is observed from a single immobilized NCBD molecule (plus a small background contribution from direct excitation of labeled ACTR molecules in solution). Upon binding of an ACTR molecule, FRET results in a decrease in donor fluorescence and an increase in acceptor fluorescence (Fig. 1b). Upon dissociation of ACTR, FRET ceases and acceptor emission is lost. At a typical ACTR concentration,  $c_{\text{ACTR}}$ , of 65 nM used here (determined from a fluorescence correlation spectroscopy (FCS) measurement in the solution above the surface; see Methods), several association and dissociation events revealed by such anticorrelated changes in donor and acceptor signal occur each second (Fig. 1). With an average duration of the single-molecule time traces of  $\sim 1$  min until the donor dye photobleaches (Supplementary Fig. 2e), we can thus record hundreds to thousands of binding and dissociation events of a single NCBD molecule, yielding excellent statistics even from individual time traces (Fig. 1, Supplementary Fig. 2). Note that acceptor photobleaching is not an issue, since the dissociation of ACTR is much faster than photobleaching ( $< 0.1 \text{ s}^{-1}$ ), and new ACTR molecules are always replenished from the solution.

**Slowly interconverting states with different ACTR affinity.** With this approach, the interaction kinetics of ACTR and NCBD can be analyzed at equilibrium from milliseconds to minutes by extracting association and dissociation rate coefficients from the distributions of dwell times in the bound and unbound states or by using maximum likelihood (MLH) analysis in combination with hidden Markov models<sup>28,29</sup>, as we will show below. However, besides the rapid transitions between the bound and the unbound state, inspection of the time traces reveals two kinetic regimes that alternate on a much slower time scale: in the time trace of Fig. 1b, e.g., the binding events between  $\sim 5$  s and  $\sim 59$  s last much longer on average than outside that period. Correspondingly, the relative population of the bound state is increased during this time (Fig. 1c). This behavior is observed in many time traces (Supplementary Fig. 1c) and reveals that a single NCBD molecule can switch between two kinetic regimes that persist for tens of seconds. Notably, the transfer efficiencies are very similar in both kinetic regimes (Fig. 1c), so only their kinetic characteristics—and not the photon count ratios or transfer efficiencies—can be used to distinguish them.

To illustrate the difference in binding kinetics between the two regimes quantitatively, we split all time traces into 4-second segments and determined for each segment the mean dwell time of NCBD in the bound state,  $\langle \tau_{\text{on}} \rangle$ , and in the unbound state,  $\langle \tau_{\text{off}} \rangle$ , using the Viterbi algorithm (see Methods). The corresponding 2D histogram of dwell times (Fig. 1d) shows two populations that reflect the two kinetic regimes. We used the resulting average dwell times as initial values for a global MLH analysis of all recorded time traces with a hidden Markov model based on the kinetic scheme in Fig. 1e. The model describes association and dissociation of ACTR and NCBD, with NCBD additionally interconverting between two states, NCBD1 and NCBD2, that bind ACTR with different kinetics. The resulting rate coefficients (Table 1) translate into equilibrium dissociation constants ( $K_{\text{d},1} = k_{\text{off},1}/k_{\text{on},1}$  and  $K_{\text{d},2} = k_{\text{off},2}/k_{\text{on},2}$ ) of ACTR to NCBD1 and NCBD2, respectively, that differ by about an order of magnitude, with  $78 \pm 6$  nM for the high-affinity state, NCBD1,



**Fig. 1** Binding of ACTR to single surface-immobilized NCBD molecules. **a** Schematic representation of acceptor-labeled ACTR (blue) binding to surface-immobilized donor-labeled NCBD (orange). **b** A single-molecule time trace showing donor (green) and acceptor fluorescence (red) until donor photobleaching, with progressive magnifications above. A period where ACTR forms relatively long-lived complexes with NCBD is shaded in purple; binding events are sparser and on average shorter in the segments shaded in light green. The most likely state trajectory (bound/unbound) identified by the Viterbi algorithm is depicted as gray lines for the largest magnification. 283 binding events were identified in this time trace. Note that the lower quantum yield of the acceptor compared to the donor dye leads to a decrease in total fluorescence intensity upon ACTR binding. **c** Histograms of photon count ratios (acceptor over total number of photons) for the two kinetic regimes in **b**, calculated for 20-ms time bins. The two populations arise from the unbound (green, low photon count ratio) and the bound state (red, high photon count ratio) of NCBD; the corresponding fractions bound and the time ranges in the trajectory are indicated. **d** 2D histogram of the mean time in the bound state,  $\langle \tau_{on} \rangle$ , and in the unbound state,  $\langle \tau_{off} \rangle$ , generated by splitting all 163 measured time traces into 4-s segments, resulting in two clusters. **e** Kinetic binding model used to analyze the time traces, including the interconversion between the two states of NCBD with different kinetics of dissociation and association with ACTR

**Table 1** Kinetic parameters determined with maximum likelihood analysis of single-molecule FRET experiments

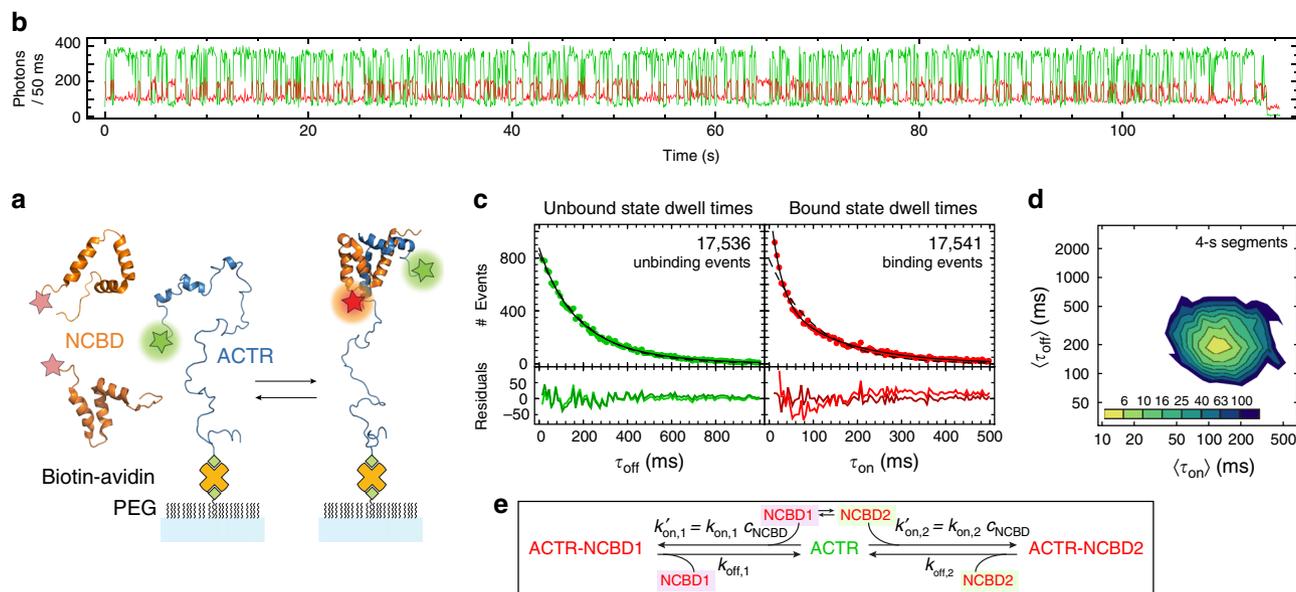
Immobilized protein	$k'_{on,1} = k_{on,1} \cdot c_{ligand}$ ( $s^{-1}$ ) <sup>a</sup>	$k'_{on,2} = k_{on,2} \cdot c_{ligand}$ ( $s^{-1}$ ) <sup>a</sup>	$k_{off,1}$ ( $s^{-1}$ )	$k_{off,2}$ ( $s^{-1}$ )	$c_{ligand}$ (nM) <sup>b</sup>	$k_{on,1}$ ( $10^8 M^{-1}s^{-1}$ ) <sup>c</sup>	$k_{on,2}$ ( $10^8 M^{-1}s^{-1}$ ) <sup>c</sup>	$k_{12}$ ( $s^{-1}$ )	$k_{21}$ ( $s^{-1}$ )
NCBD	$6.0 \pm 0.2$	$3.0 \pm 0.5$	$7.3 \pm 0.3$	$30 \pm 3$	65	$0.93 \pm 0.06$	$0.46 \pm 0.08$	$0.04 \pm 0.01$	$0.07 \pm 0.01$
ACTR	$3.3 \pm 0.3$	$1.6 \pm 0.3$	$5.6 \pm 0.5$	$30 \pm 6$	17	$3.2 \pm 0.5$	$2.5 \pm 0.6$	n.a.	n.a.
NCBD P20A	$5.1 \pm 0.2$	n.a.	$23 \pm 1$	n.a.	65	$0.78 \pm 0.05$	n.a.	n.a.	n.a.

Errors are the standard deviations of ten bootstrapping trials if not stated otherwise  
 n.a., not applicable  
<sup>a</sup>Pseudo-first-order association rate coefficient observed in the time traces. The second-order association rate coefficient is calculated based on the ligand concentration  $c_{ligand}$  measured by FCS in solution (see Methods)  
<sup>b</sup>Concentrations of the acceptor-labeled ligand were determined using FCS (see Methods). An uncertainty of  $\pm 5\%$  was estimated from two independent measurements conducted before and after recording the time traces  
<sup>c</sup>Second-order association rate coefficient calculated from the pseudo-first-order association rate coefficient based on the ligand concentration,  $c_{ligand}$ , measured by FCS in solution (see Methods). Uncertainty from propagating the error of  $k'_{on}$  and the ligand concentration. The error for immobilized ACTR is greater owing to the greater uncertainty in the relative populations of NCBD1 and NCBD2,  $p_1$  and  $p_2$ :  $p_1 = k_{21}/(k_{12} + k_{21})$  and  $p_2 = k_{12}/(k_{12} + k_{21})$

and  $650 \pm 150$  nM for the low-affinity state, NCBD2 (errors are standard deviations based on bootstrapping). The interconversion rates between NCBD1 and NCBD2 are slow ( $k_{12} = 0.04 \pm 0.01 s^{-1}$  and  $k_{21} = 0.07 \pm 0.01 s^{-1}$ ) (Table 1), as expected. If direct interconversion between the bound states, NCBD1-ACTR and NCBD2-ACTR, is included in the model, MLH analysis returns values for the corresponding rate coefficients indistinguishable from zero; we thus exclude this link.

To test the proposed kinetic model further, we take advantage of the possibility to invert the roles of the two binding partners in the single-molecule experiments and probe association of surface-immobilized ACTR with freely diffusing NCBD (Fig. 2). This

arrangement allows us to assess whether ACTR exhibits similar kinetic heterogeneity as NCBD, with long-lived interconverting states. Such behavior is absent in the 75 time traces recorded (see Fig. 2b and Supplementary Fig. 3a for examples). Consistently, a 2D dwell-time histogram for time trace segments of 4 s shows only one population (Fig. 2d), indicating the presence of a single, kinetically homogenous population of ACTR. However, if NCBD exists in a high- and a low-affinity state as proposed above, the reaction should be described by the kinetic model in Fig. 2e, where two different NCBD populations bind to ACTR. The model predicts single-exponential association kinetics decaying with the sum of the two pseudo-first-order



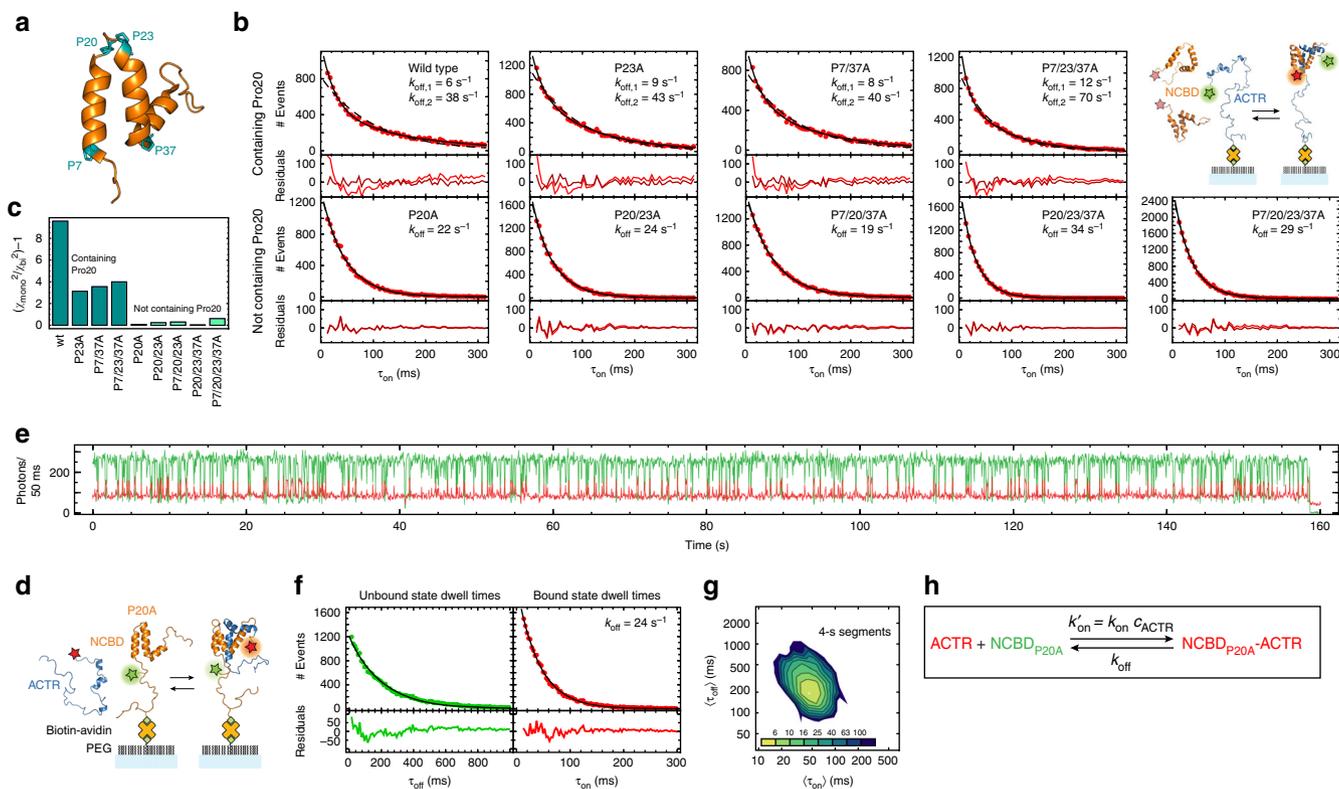
**Fig. 2** Binding of NCBD to single surface-immobilized ACTR molecules. **a** Schematic representation of acceptor-labeled NCBD (orange) binding to surface-immobilized donor-labeled ACTR (blue). **b** Representative single-molecule time trace (donor emission green, acceptor emission red). **c** Dwell-time distributions of the unbound state (green) and the bound state (red) with single-exponential (dashed lines, light residuals) and double-exponential fits (solid lines, dark residuals). **d** 2D histogram of  $\langle\tau_{\text{on}}\rangle$  and  $\langle\tau_{\text{off}}\rangle$  from 75 time traces, showing a single cluster. **e** Kinetic model used to analyze the time traces, with two different NCBD states interacting with ACTR

association rate coefficients,  $k'_{\text{on},1} + k'_{\text{on},2}$  ( $k'_{\text{on},1} = k_{\text{on},1} \cdot c_{\text{NCBD}1}$  and  $k'_{\text{on},2} = k_{\text{on},2} \cdot c_{\text{NCBD}2}$ , where  $k_{\text{on},1/2}$  are the second-order rate coefficients and  $c_{\text{NCBD}1/2}$  the concentrations of NCBD1/2 free in solution). The dissociation kinetics, however, are predicted to be bi-exponential, with the two phases (with  $k_{\text{off},1}$  and  $k_{\text{off},2}$ ) arising from the high- and low-affinity NCBD-ACTR complexes, respectively<sup>30</sup>. Indeed, this is what we observe (Fig. 2c), and a global MLH analysis of all time traces reveals dissociation rate coefficients in good agreement with those observed for immobilized NCBD (Table 1). Moreover, the observed fraction of fast transitions [given by  $k_{12}k_{\text{on},2}/(k_{21}k_{\text{on},1} + k_{12}k_{\text{on},2})$ ] for immobilized ACTR ( $32 \pm 4\%$ ) is close to the corresponding value from the measurements with immobilized NCBD ( $24 \pm 6\%$ ). This agreement supports the presence of the two different states of NCBD and furthermore suggests that surface interactions do not have a pronounced effect on the measurements.

We note, however, that the association rate coefficients,  $k_{\text{on},1}$  and  $k_{\text{on},2}$ , are three to five times greater when ACTR is immobilized rather than NCBD (cf. Table 1), in better agreement with published values<sup>31</sup>. To test whether this discrepancy is caused by NCBD immobilization, we performed stopped-flow experiments with the same constructs used for immobilization (Supplementary Fig. 4). The resulting association rate coefficients [ $(0.89 \pm 0.07) \cdot 10^8 \text{ M}^{-1} \text{ s}^{-1}$  for ACTR binding to biotinylated NCBD and  $(3.0 \pm 0.1) \cdot 10^8 \text{ M}^{-1} \text{ s}^{-1}$  for NCBD binding to biotinylated ACTR; an average from the contributions of the high- and the low-affinity population] agree well with the ones obtained from the single-molecule experiments, indicating that the dominant effect on the kinetics originates from biotinylation and not surface immobilization per se. A likely cause is the negative net charge ( $-3$ ) of the Avi-tag introduced for biotinylation in vivo, which either reduces (in the case of NCBD-Avi) or increases (in the case of ACTR-Avi) the charge-promoted interaction between ACTR and NCBD (net charges of  $-8$  and  $+6$ , respectively)<sup>31</sup>. The charges of the dyes (Cy3B: 0, CF680R:  $-1$ ) appear to have a smaller influence than the Avi-Tag.

**Pro20 isomerization induces a conformational switch in NCBD.** What is the molecular origin of the long-lived kinetic heterogeneity in NCBD? Given the marginal stability of NCBD, the presence of conformational states that persist for seconds may be surprising. However, NCBD contains seven proline residues, four of which are located in the central part of the protein, where they terminate helices (Fig. 3a). Moreover, the time scale of peptidyl-prolyl *cis/trans* isomerization ( $\sim 0.06 \text{ s}^{-1}$  at  $22^\circ \text{C}$  in unfolded peptides<sup>32,33</sup>) is in accord with the slow dynamics observed in NCBD. A first corroboration of the hypothesis that peptidyl-prolyl *cis/trans* isomerization causes the slow kinetic heterogeneity comes from the addition of the peptidyl-prolyl *cis/trans* isomerase Cyclophilin A. Cyclophilin A accelerates the switching between low- and high-affinity subpopulations of NCBD ( $k_{12}$  and  $k_{21}$ ) by almost an order of magnitude, while leaving the association rate coefficients unaffected (Supplementary Fig. 5, Supplementary Table 1). But which proline residue(s) cause(s) the observed switching?

To address this question, we prepared NCBD variants with Pro to Ala substitutions and tested their interaction in single-molecule FRET experiments with immobilized ACTR, analogous to Fig. 2. The dissociation kinetics of the nine NCBD variants with different combinations of proline substitutions fall into two classes: all variants where Pro20 is substituted by alanine (P20A) decay single-exponentially, whereas the variants containing P20 decay bi-exponentially (Fig. 3b, c; for rate coefficients, see Supplementary Table 2). Pro20 is thus the most likely cause of the slow kinetics observed in NCBD. To test the importance of Pro20 also in the inverse experiment, we incorporated the P20A substitution into surface-immobilized NCBD and followed the binding kinetics of freely diffusing ACTR, analogous to Fig. 1. As expected, both binding and dissociation kinetics are single-exponential and can be described by a simple two-state model (Fig. 3d, f, h), with dissociation rate coefficients similar to measurements where ACTR is immobilized and NCBD lacking Pro20 is free in solution. Most importantly, the time traces with immobilized NCBD P20A no longer exhibit the slow switching



**Fig. 3** Isomerization of Pro20 is the cause of kinetic heterogeneity. **a** Model of NCBD (PDB entry 2KKJ)<sup>23</sup>, with helix-terminating proline residues highlighted in turquoise. **b** Dwell-time distributions for the bound state of surface-immobilized ACTR for all NCBD variants, with single-exponential (dashed line, light red residuals) and double-exponential fits (solid line, dark red residuals) and the fitted rates indicated. Dwell times below 12 ms are too short to be identified reliably by the Viterbi algorithm (Supplementary Fig. 6) and are hence omitted. MLH analysis results are compiled in Supplementary Table 2. **c** Normalized ratio of chi-squared as an indicator for the quality of single- and double-exponential fits in **b**, showing variants containing Pro20 in dark green and those with Pro20 exchanged by Ala in light green. **d** Schematic representation of acceptor-labeled ACTR (blue) binding to surface-immobilized donor-labeled NCBD P20A (orange). **e** Corresponding representative single-molecule time trace (donor emission green, acceptor emission red). **f** Dwell-time distributions and residuals for ACTR binding to and dissociating from surface-immobilized NCBD P20A, with single-exponential fits (solid lines). **g** 2D histogram of  $\langle\tau_{on}\rangle$  and  $\langle\tau_{off}\rangle$  from 113 time traces showing a single kinetic cluster. **h** Two-state kinetic model used to analyze the time traces of NCBD P20A

between the two kinetic regimes (Supplementary Fig. 3b), and splitting the time traces into 4-s segments reveals only a single kinetic cluster (Fig. 3g). In summary, we thus conclude that *cis/trans* isomerization of Pro20 in NCBD causes the protein to switch between two conformational ensembles with different affinities for ACTR.

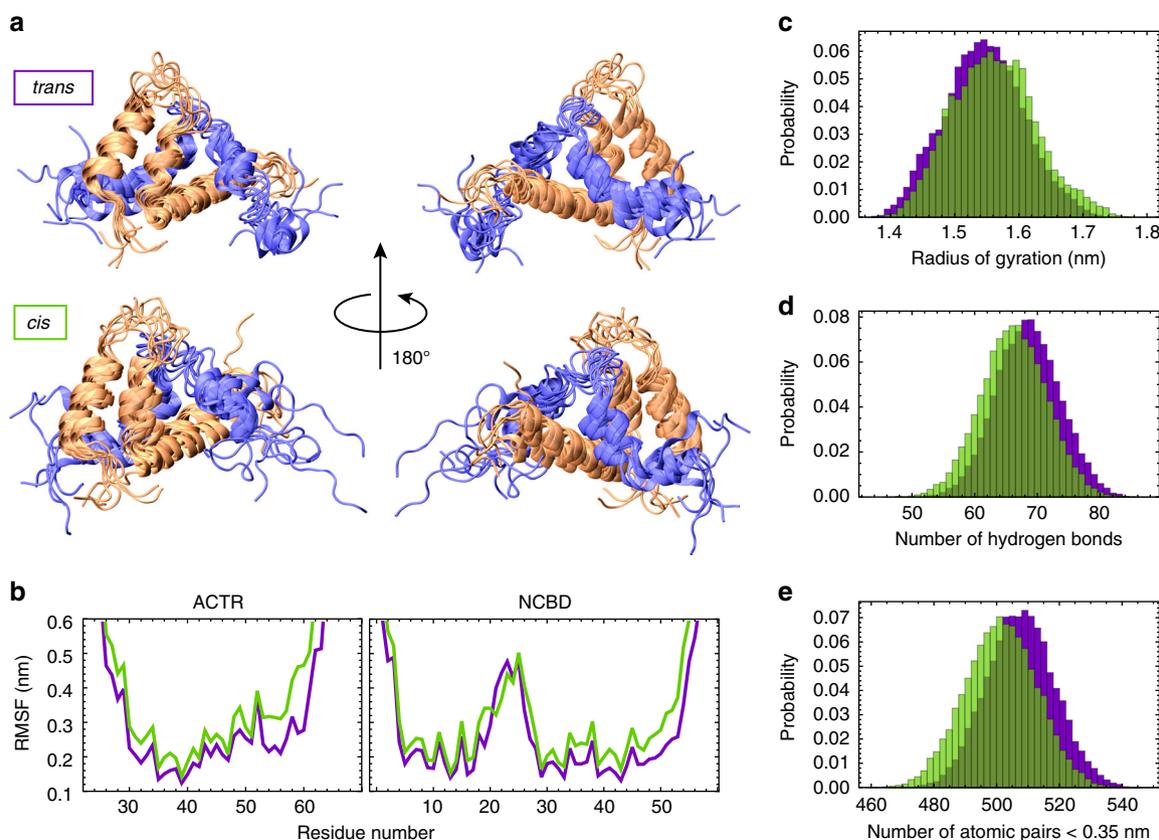
**Loss of complex stability in the *cis*-state.** To probe the structural origin of the experimentally observed differences in the affinity of the NCBD-ACTR complex depending on the isomerization state of Pro20, we employed molecular dynamics (MD) simulations to sample the dynamics of the complex in the *trans* and the *cis* state of Pro20. Modeling the *cis* state of Pro20 by changing the value of the S19-P20  $\omega$ -angle from  $180^\circ$  to  $0^\circ$  would disrupt the topology of NCBD in the complex. Since the inter-dye distances in the complexes of the *trans* and *cis* forms are very similar (Fig. 1c), the arrangement of the NCBD helices in complex with ACTR is unlikely to be altered. We thus modeled the *cis* state by perturbing the  $\omega$ -angle by well-tempered metadynamics (see Methods for details), which shows an energy barrier between the two states of  $\sim 80$  kJ mol<sup>-1</sup>, in line with experimental results<sup>34</sup> and previous calculations using the AMBER2003 force field<sup>35</sup>. We then simulated the *cis* and *trans* conformations separately by replica exchange MD to investigate the stability of the complex when P20 is either in the *trans* or the *cis* state. Within the simulated time and at the chosen temperatures, the simulations do not sample prolyl isomerization, as expected<sup>36</sup> (Supplementary Fig. 7). The

simulations suggest that both ACTR and NCBD are consistently more dynamic when Pro20 is in the *cis* state (Fig. 4a), as indicated by greater root mean square fluctuations in *cis* along the entire sequence of both NCBD and ACTR (Fig. 4b). As a result, the distribution of the radius of gyration of the complex is shifted to slightly higher values with Pro20 in *cis* (Fig. 4c).

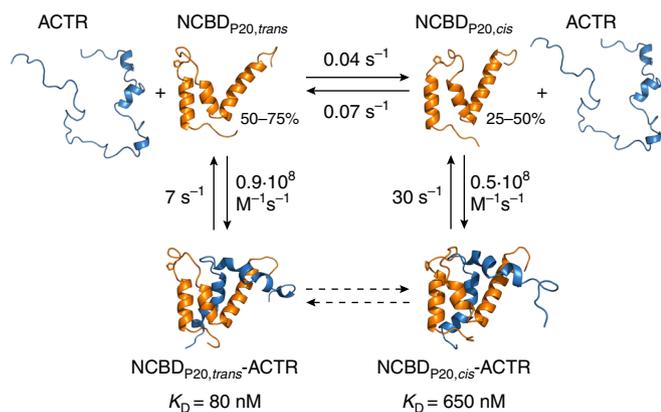
This increase in fluctuations and dimensions results from a loss of interactions between the two partners: both the number of hydrogen bonds and of all interactions between atomic pairs within 0.35 nm are on average greater in the *trans* state (Fig. 4d, e). A more detailed analysis indicates a loss of contacts in the loop of NCBD containing Pro20, which is accompanied by a reduced average contact time in the interface between the C-termini of the proteins, as well as a slight loss of packing in the hydrophobic core (Supplementary Fig. 8). Altogether, these findings indicate that a change in the isomerization state of Pro20 from *trans* to *cis* can generate long-range effects that perturb interactions in the NCBD-ACTR complex and lead to faster dissociation and reduced affinity.

## Discussion

We find that NCBD undergoes slow conformational switching between two subpopulations that differ by the conformation of a proline residue. *Cis/trans*-isomerization as the origin of the kinetic heterogeneity observed in single-molecule experiments is confirmed by proline substitutions and the effect of proline *cis/trans*-isomerase on the kinetics. The overall kinetic mechanism (Fig. 5) thus involves



**Fig. 4** Effect of Pro20 isomerization from replica exchange MD simulations. **a** Overlap of representative conformations of the NCBD (orange)-ACTR (blue) complex with Pro20 in the *trans* and *cis* state. **b** Root mean square fluctuation (RMSF) profiles along the sequence of ACTR and NCBD, with NCBD Pro20 in *trans* (purple) or *cis* (green). The RMSF was averaged over all atoms of each residue. **c** Probability distributions for the radii of gyration of the NCBD-ACTR complex, with NCBD Pro20 in *trans* (purple) or *cis* (green). **d**, **e** Probability distributions of intermolecular hydrogen bonds (**d**) or any intermolecular atomic pair falling within a 0.35-nm distance cutoff (**e**) in the MD simulations with Pro20 in *trans* (purple) or *cis* (green)



**Fig. 5** Kinetic model of the NCBD-ACTR interaction. The two subpopulations of NCBD differ by the *cis/trans* isomers of the peptide bond involving Pro20, which are interconverting slowly; both bind ACTR, but with different association rate coefficients. The two complexes can be distinguished experimentally by their dissociation rate coefficient, which is greater for the NCBD<sub>P20,cis</sub>-ACTR complex. The difference in association and dissociation rates translates into an eightfold higher affinity of the NCBD<sub>P20,trans</sub>-ACTR complex

four states: NCBD with Pro20 in *trans*, which binds ACTR with higher affinity, and NCBD with Pro20 in *cis*, which binds ACTR with lower affinity. This difference in affinity is largely caused by a change in dissociation rate coefficient; the association rate

coefficients are similar for both subpopulations. The binding interface between ACTR and NCBD has been shown to be sub-optimal and frustrated<sup>37</sup>, which might explain why the affinity loss is not even larger in response to the proline switch. We find no evidence for conformational exchange between the *cis*- and *trans*-isomers of Pro20 while NCBD is bound to ACTR. It seems likely that binding locks NCBD in the respective conformation by raising the energy barrier for *cis/trans* isomerization, but we cannot exclude a low rate of exchange that is beyond our current detection limit. Based on the parameters we obtained here (Fig. 5), the *cis* complex is only populated to 6% at equilibrium at the concentrations of 3 mM NCBD and 6 mM ACTR where the NMR structure was solved - consistent with the absence of a population with Pro20 in *cis* in the structure<sup>21</sup>. In unbound NCBD, however, evidence for the presence of *cis/trans* isomerization of Pro20 has been reported based on peak splitting of NMR resonances<sup>24,38</sup>. Peak splitting is also observed in the spectra reported by Kjaergaard et al.<sup>23</sup> if analyzed at lower contour levels (K. Teilum and M. Kjaergaard, personal communication). Moreover, multi-exponential kinetics have been observed in the association and dissociation of the NCBD-ACTR complex with stopped-flow techniques<sup>39</sup>, suggesting the presence of a complex interaction mechanism and/or multiple states of the interacting molecules<sup>40</sup>, in accordance with our findings.

The interaction between NCBD and ACTR is a prime example for intrinsic disorder of both ligands that enables the formation of transient high-affinity interactions by combining very fast, charge-promoted association rates with dissociation on the millisecond to seconds time scale<sup>2,41</sup>. NCBD has more than ten

known interaction partners<sup>24,42</sup>, among them the activation domains of the steroid receptor coactivator (SRC) family<sup>21</sup>, the p53 activation domain<sup>43</sup>, and viral factors that hijack the host transcription machinery (e.g., adenoviral E1A)<sup>44</sup>. Rapid dissociation facilitates a fast response of the coactivator machinery to this wide variety of regulators. Structures of NCBD in complex with its other interaction partners reveal large variations, with very different arrangements of the helices, e.g., in the IRF-3<sup>26</sup> and ACTR-bound states<sup>21</sup>. Peptidyl-prolyl isomerization raises the interesting possibility of modulating the interaction with different ligands and thus differentially regulating downstream signaling. Similar to ACTR, other disordered targets of NCBD may bind both subpopulations, but some might have higher affinity for the *cis* conformation. Structured ligands, such as the pointed domain of Ets-2<sup>42</sup> or IRF-3, might bind to one conformation only, as they cannot as readily adapt to different binding interfaces. The high evolutionary conservation of Pro20 in NCBD (Supplementary Fig. 9) supports such an important functional role. Post-translational modifications of the phosphorylation sites reported for NCBD<sup>45</sup> could further regulate the equilibrium and dynamics between the subpopulations and subsequent ligand binding.

Several cases where peptidyl-prolyl *cis/trans* isomerization regulates binding have been described in the context of IDPs. One example is the C-terminal disordered tail of RNA-Polymerase II, where *cis/trans* isomerization of Pro6 (regulated by phosphorylation of the preceding serine residue) allows binding to different downstream factors<sup>46</sup>. Other examples include p53, where a *cis* proline mediates the apoptosis-triggering interaction with BAX<sup>47</sup>, as well as the disordered transactivation domain of BMAL-1, where the switch is linked to the circadian rhythm<sup>48</sup>. Furthermore, in an antibody-IDP interaction, proline *cis/trans* isomerization was linked to slow kinetic phases in binding<sup>49</sup>. Considering the high relative abundance of proline residues in IDPs<sup>20</sup>, this mechanism of regulating differential binding could be widespread. Elucidating the underlying kinetic mechanisms will be essential for understanding how such regulation by peptidyl-prolyl *cis/trans* isomerization occurs and how it couples to other contributions, such as secondary structure content<sup>50</sup>. Single-molecule spectroscopy, which has previously been employed to expose molecule-to-molecule variation in the dynamics of nucleic acids<sup>51,52</sup> and the influence of proline isomerization on protein folding dynamics<sup>53</sup>, provides a versatile toolbox for addressing such questions.

## Methods

**Protein expression.** All protein sequences used in this study are compiled in Supplementary Table 3. All primers (Microsynth) used for site-directed mutagenesis are listed in Supplementary Table 4. The coding sequences of Avi-tagged single-cysteine NCBD and ACTR variants for immobilization were cloned via BamHI/HindIII into a pAT222-pD expression vector (gift of J. Schöppe and A. Plückthun)<sup>54</sup>, yielding an expression construct with an N-terminal Avi-tag and a Thrombin-cleavable C-terminal His<sub>6</sub>-tag. The NCBD-Avi P20A mutant was generated by site-directed mutagenesis of the NCBD-Avi construct. pBirAcm (Avidity) was co-transfected for *in vivo* biotinylation of Lys12 in the Avi-tag, and expression was carried out in *E. coli* C41(DE3) (Merck). Cells were grown at 37 °C in TYH medium (for 1 l: 20 g trypton, 10 g yeast extract, 11 g HEPES, 5 g NaCl, 1 g MgSO<sub>4</sub>, pH 7.3), supplied with 0.5% (w/v) glucose, until they reached an OD<sub>600</sub> of 0.8. Then, 50 μM biotin in 10 mM bicine buffer (pH 8.3) and 1 mM IPTG were added to the culture. Expression continued for 3 h at 37 °C, after which cells were harvested by centrifugation. The harvested cells were lysed by sonication and the His<sub>6</sub>-tagged protein enriched via IMAC on Ni-IDA resin (ABT). The His<sub>6</sub>-tag was then cleaved off with Thrombin (Serva Electrophoresis) and separated from the protein by another round of IMAC. Finally, biotinylated protein was separated from impurities and non-biotinylated protein via reversed-phase HPLC (RP-HPLC) on a C18 column (Reprosil Gold 200, Dr. Maisch) with a H<sub>2</sub>O/0.1% TFA—acetonitrile gradient and lyophilized.

Single-cysteine variants of NCBD and ACTR were generated by site-directed mutagenesis and co-expressed<sup>55</sup> from a pET-47b(+) vector. The expression construct contained an N-terminal His<sub>6</sub>-tag cleavable with HRV 3C protease. Transformed *E. coli* C41(DE3) cells were grown at 37 °C in TYH medium supplied

with 0.5% glucose until they reached an OD<sub>600</sub> of 0.8, after which expression was induced with 1 mM IPTG for 1 h at 37 °C. Cell harvest, lysis, and protein enrichment via IMAC were carried out as described above, followed by enzymatic cleavage of the His<sub>6</sub>-tag with HRV 3C protease and separation of the tag from the proteins via another round of IMAC. Finally, ACTR and NCBD were separated with RP-HPLC as described above.

NCBD proline variants were generated by site-directed mutagenesis, co-expressed with ACTR analogously to wild-type NCBD and purified according to the procedures described above.

**Protein labeling.** Lyophilized Avi-tagged protein (NCBD-Avi and ACTR-Avi) was dissolved under nitrogen atmosphere to a concentration of 200 μM in 100 mM potassium phosphate buffer, pH 7.0. The protein was then labeled for 3 h at room temperature with a 0.8:1 molar ratio of Cy3B maleimide dye (GE Healthcare) to protein. Labeled protein was separated from unlabeled protein with RP-HPLC on a Sunfire C18 column (Waters) as described above. The correct mass of labeled protein was confirmed by electrospray ionization mass spectrometry (ESI-MS).

Analogously, the single-cysteine variants of NCBD and ACTR were labeled with a 1.5-fold molar excess of CF680R maleimide dye (Biotium). The free dye was separated from the labeled protein by RP-HPLC on a C18 column (Reprosil Gold 200) as described above and the protein was lyophilized. The correct mass of labeled protein was confirmed by ESI-MS.

Labeling of the NCBD proline variants was carried out as outlined above, but with a 0.7:1 molar ratio of CF680R dye to protein. The constructs contain two cysteine residues, but as Cys3 is more reactive with maleimides due to its lower pK<sub>a</sub>, site-specific labeling can be achieved. The single-labeled fraction was separated from unlabeled and double-labeled protein by RP-HPLC (first on a Reprosil Gold 200 column, followed by a Sunfire C18 column). A tryptic digest in combination with ESI-MS confirmed the correct mass and labeling position.

**Surface immobilization.** Adhesive silicone hybridization chambers (Secure Seal Hybridization Chambers, SA8R-2.5, Grace Bio-Labs) were cut in half and bound to PEGylated, biotinylated quartz coverslips (Bio\_01, MicroSurfaces, Inc.) to form 150-μl reaction chambers. For surface immobilization experiments, 10 nM NCBD-Avi was pre-coupled to 1 μM Avidin D (Vector Labs) in NaP buffer (50 mM sodium phosphate pH 7.0, 0.01% Tween 20), and immobilized for 10 min in a reaction chamber at a concentration of 20 pM NCBD-Avi/2 nM Avidin D. ACTR-Avi was immobilized by incubating 3 μM Avidin D (Vector Labs) in NaP buffer for 5 min in a reaction chamber, followed by three washing steps with NaP buffer to remove unbound Avidin. Afterwards, the surface was treated with 10 pM ACTR-Avi. With these protocols, a surface coverage of 0.1–0.3 molecules/μm<sup>2</sup> was achieved.

**Single-molecule instrumentation and experiments.** All single-molecule experiments were conducted on a custom-built confocal instrument<sup>56</sup> equipped with a 532 nm cw laser (LaserBoxx LBX-532-50-COL-PP, Oxixus) and a 635 nm diode laser (LDH-D-C-635M, PicoQuant). Fluorescence photons were separated from scattered photons with a triple band mirror (zt405/530/630rpc, Chroma) and split onto two channels with a dichroic mirror (T635LPXR, Chroma) according to their wavelength. Donor photons were filtered with an ET585/65 m bandpass filter (Chroma) before detection on a τ-SPAD avalanche photodiode (PicoQuant). Acceptor photons were filtered with a LP647RU long-pass filter (Chroma) and detected with a SPCM-AQRH-14 single-photon avalanche diode (Perkin Elmer). The objective (xUplanApo 60/1.20-W, Olympus) was mounted onto a piezo stage (P-733.2 and PIFOC, Physik Instrumente GmbH) for scanning.

Binding experiments were conducted at 22 °C under argon atmosphere in NaP buffer supplied with ACTR/NCBD-CF680R. 1% (w/v) glucose, 20 μg/ml glucose oxidase (Sigma) and 20 U/ml catalase (Sigma) were included as oxygen scavenging system, as well as 1 mM methyl viologen and 1 mM ascorbic acid as triplet quenchers<sup>57</sup>. Human Cyclophilin A with an N-terminal FLAG-tag was prepared according to published procedures<sup>58</sup> and included in the measurements shown in Supplementary Fig. 5. Single immobilized NCBD/ACTR molecules were localized by scanning a 16 × 16 μm<sup>2</sup> area (130 nm/pixel) with a 532-nm cw laser at a power of 4 μW (measured at the back aperture of the objective). The objective collar and z-positioning of the focus were optimized for molecular brightness. For the acquisition of binding-unbinding time traces, the laser power was reduced tenfold. Time traces for 50–150 immobilized molecules were recorded until the Cy3B dye photobleached. Data were acquired with the SymphoTime software (PicoQuant).

**Analysis of single-molecule time traces.** Single-molecule time traces were inspected to ensure that no substantial brightness variations were occurring (e.g., caused by a drift of the molecule's position, long-lived dark states, or background fluctuations). Suitable traces were analyzed until photobleaching of the donor dye. Single-step photobleaching indicated that only one molecule was present in the confocal volume.

On first inspection, the time traces reveal jumps between two states of low and high FRET efficiency. Hence, in a first analysis step, we approximate the system

with two states, unbound and bound, whose exchange is described with the rate matrix

$$\mathbf{K}_{2\text{state}} = \begin{pmatrix} -k'_{\text{on}} & k_{\text{off}} \\ k'_{\text{on}} & -k_{\text{off}} \end{pmatrix} \quad (1)$$

$k'_{\text{on}}$  and  $k_{\text{off}}$  are the transition rate coefficients between the bound and the unbound state of the immobilized protein. Note that for analyzing the data, we use the pseudo-first-order rate coefficients,  $k'_{\text{on}} = k_{\text{on}} \cdot c_{\text{ligand}}$ , the product of the second-order rate coefficient,  $k_{\text{on}}$ , and the concentration of acceptor-labeled binding partner,  $c_{\text{ligand}}$ . We determined the donor and acceptor photon rates associated with each state by applying the MLH method on binned photon time traces rather than individual photons to achieve sufficient numerical efficiency for a global analysis of all measurements<sup>29,59</sup>. The likelihood,  $L_m$ , of a time trace  $m$  with bin size  $\Delta$ , number of bins  $T_m$  and  $(N_{D,i}, N_{A,i})$  donor and acceptor photons detected in time bin  $i$  is calculated from

$$L_m = \mathbf{1}^T \left[ \prod_{t=1}^{T_m} \mathbf{F}_t e^{\Delta \mathbf{K}} \right] \mathbf{p}_{\text{eq}} \quad (2)$$

Here, the population vector  $\mathbf{p}_{\text{eq}}$  describes the equilibrium distribution of states for which  $\mathbf{K} \mathbf{p}_{\text{eq}} = 0$ .  $\mathbf{F}_t$  is a diagonal matrix with elements

$(\mathbf{F}_t)_{ii} = \frac{(n_{D,i} \Delta)^{N_{D,i}}}{N_{D,i}!} e^{-n_{D,i} \Delta} \times \frac{(n_{A,i} \Delta)^{N_{A,i}}}{N_{A,i}!} e^{-n_{A,i} \Delta}$ , assuming Poisson statistics for the number of photons per bin.  $n_{D,i}$  and  $n_{A,i}$  are the mean photon detection rates of the  $i$ th state in the donor and acceptor channel, respectively. We found the photon detection rates and transition rate coefficients for each individual time trace (binned at 20 ms) by maximizing  $\ln(L_m)$ .

In the next step, we used the photon rates and transition rate coefficients to identify the most likely state trajectories using the Viterbi algorithm<sup>60,61</sup>. For this purpose, the photon time traces were binned at 1 ms to avoid averaging over fast events. If the rate matrix  $\mathbf{K}_{2\text{state}}$  is used as input, occasional short blinking events are misrecognized as binding events. Therefore, we modify  $\mathbf{K}_{2\text{state}}$  to include a dark state accounting for blinking in the low-FRET-efficiency unbound state, which is populated and depopulated with rate coefficients  $k_{+b}$  and  $k_{-b}$ :

$$\mathbf{K}_{2\text{state,blink}} = \begin{pmatrix} -(k'_{\text{on}} + k_{+b}) & k_{\text{off}} & k_{-b} \\ k'_{\text{on}} & -k_{\text{off}} & 0 \\ k_{+b} & 0 & -k_{-b} \end{pmatrix} \quad (3)$$

Blinking also occurs in the high-FRET-efficiency bound state but does not need to be included in the model since it is not misrecognized as a transition. We note that blinking events can be well distinguished from binding and dissociation transitions, since they have a lower photon rate both in the donor and the acceptor channel (see Supplementary Fig. 2f, l). MLH analysis yields rate coefficients of  $k_{+b} = 1.7 \text{ s}^{-1}$  and  $k_{-b} = 38 \text{ s}^{-1}$  for immobilized NCBD and  $k_{+b} = 5 \text{ s}^{-1}$  and  $k_{-b} = 200 \text{ s}^{-1}$  for immobilized ACTR. From the identified state trajectory, dwell times in the bound ( $\tau_{\text{on}}$ ) and unbound state ( $\tau_{\text{off}}$ ) are obtained to construct 2D-plots of  $\langle \tau_{\text{on}} \rangle$  vs.  $\langle \tau_{\text{off}} \rangle$  (Figs. 1d, 2d, 3g) and dwell-time histograms (Figs. 2c, 3b, f). Even though the underlying kinetics are clearly more complex than two-state, the Viterbi algorithm still reliably identifies over 92% of all transitions (as shown for simulated data in Supplementary Fig. 6).

Depending on the complexity of the underlying kinetics, the strategy to fit the individual transition rate coefficients was adjusted. NCBD constructs with the P20A mutation (Fig. 3) exhibit single-exponential binding and dissociation kinetics, so the rate matrix  $\mathbf{K}_{2\text{state,blink}}$  was used to obtain rate coefficients. We maximize  $\sum_m \ln(L_m)$ , the sum over the likelihood of all time

traces (binned at 1 ms), with respect to  $k'_{\text{on}}$  and  $k_{\text{off}}$ . The mean photon rates of the individual time traces were fixed to the values found in the previous step. NCBD constructs without the P20A mutation exchange slowly between a high-affinity *trans*-proline state and a low-affinity *cis*-proline state. Surface-immobilized ACTR is hence observed in three states: unbound, bound to NCBD1, or bound to NCBD2 (see Fig. 2e). Including donor blinking of the unbound state, we obtain the rate matrix:

$$\mathbf{K}_{3\text{state,blink}} = \begin{pmatrix} -(k'_{\text{on},1} + k'_{\text{on},2} + k_{+b}) & k_{\text{off},1} & k_{\text{off},2} & k_{-b} \\ k'_{\text{on},1} & -k_{\text{off},1} & 0 & 0 \\ k'_{\text{on},2} & 0 & -k_{\text{off},2} & 0 \\ k_{+b} & 0 & 0 & -k_{-b} \end{pmatrix} \quad (4)$$

The MLH analysis was carried out as described above, with the two bound states having identical photon rates, as suggested by the experimental data.

Finally, if wild-type NCBD is immobilized instead of ACTR (Fig. 1), we observe four states (see Fig. 1e): unbound NCBD1, unbound NCBD2, NCBD1 bound to ACTR, and NCBD2 bound to ACTR. Including donor blinking of

the unbound states, we obtain:

$$\mathbf{K}_{4\text{state,blink}} = \begin{pmatrix} -(k_{12} + k'_{\text{on},1} + k_{+b}) & k_{\text{off},1} & k_{-b} & k_{21} & 0 & 0 \\ k'_{\text{on},1} & -k_{\text{off},1} & 0 & 0 & 0 & 0 \\ k_{+b} & 0 & -(k_{-b} + k_{12}) & 0 & 0 & k_{21} \\ k_{12} & 0 & 0 & -(k_{21} + k'_{\text{on},2} + k_{+b}) & k_{\text{off},2} & k_{-b} \\ 0 & 0 & 0 & k'_{\text{on},2} & -k_{\text{off},2} & 0 \\ 0 & 0 & k_{12} & k_{+b} & 0 & -(k_{-b} + k_{21}) \end{pmatrix} \quad (5)$$

Switching events between *cis* and *trans* are rare, do not entail a significant change in transfer efficiency of the bound state, and can only be recognized by the persistent change of the binding kinetics of NCBD. A good estimate for the rate coefficients associated with binding ( $k'_{\text{on},1}$ ,  $k_{\text{off},1}$ ,  $k'_{\text{on},2}$ ,  $k_{\text{off},2}$ ) is thus needed for a reliable fit of the much lower isomerization rate coefficients,  $k_{12}$  and  $k_{21}$ . We obtain these estimates by an iterative procedure, where time traces are initially split into *cis* and *trans* segments for separate analysis of the binding kinetics. In the first step, we estimate the blinking rate coefficients,  $k_{+b}$  and  $k_{-b}$ , by maximizing the likelihood of the kinetic model based on the rate matrix  $\mathbf{K}_{2\text{state,blink}}$ , and fix them in all further steps. Second, we fix  $k'_{\text{on},1}$ ,  $k_{\text{off},1}$ ,  $k'_{\text{on},2}$  and  $k_{\text{off},2}$  to the values estimated from the  $\langle \tau_{\text{on}} \rangle$  vs.  $\langle \tau_{\text{off}} \rangle$  plot (Fig. 1d) and estimate the *cis/trans*-isomerization rate coefficients,  $k_{12}$  and  $k_{21}$ , individually. Next, we identify the most likely state trajectory with the Viterbi algorithm, fixing the rate coefficients to the previously estimated values. Now, the time traces are split according to their *cis* or *trans* state identified by the Viterbi algorithm. The *cis* and *trans* segments are analyzed with the rate matrix  $\mathbf{K}_{2\text{state,blink}}$  to obtain  $k'_{\text{on},1}$  and  $k_{\text{off},1}$  (from the *trans* segments) as well as  $k'_{\text{on},2}$  and  $k_{\text{off},2}$  (from the *cis* segments). Finally, this set of refined rate coefficients is used in a second iteration to fit the isomerization rate coefficients,  $k_{12}$  and  $k_{21}$ , with the model  $\mathbf{K}_{4\text{state,blink}}$  in a global analysis of all time traces.

**Fluorescence correlation spectroscopy.** Before and after recording time traces, the concentration of CF680R-labeled ligand in solution was estimated by measuring FCS curves with the 635-nm diode laser at a power of 10  $\mu\text{W}$  (measured at the back aperture of the objective). The laser was focused into the solution, 20  $\mu\text{m}$  above the surface. Fluorescence photons were separated according to their polarization and detected on two channels. The fluorescence signal of both channels was cross-correlated. The amplitude of the cross-correlation of the acceptor signal,  $G_{AA}(\tau)$ , was used to estimate the concentration of the CF680R-labeled binding partner in the surface experiments. The amplitude of  $G_{AA}(\tau)$  depends on the average number of molecules  $\langle N \rangle$  in the confocal volume, which is proportional to the concentration of the fluorescent species in the solution<sup>62</sup>. Occasional high-intensity fluorescence bursts caused by protein aggregates were removed from the raw acceptor signal. Then, the mean number of molecules in the confocal volume was obtained by fitting the FCS curve with a model including terms for translational diffusion and triplet dynamics:

$$G_{AA}(\tau) = 1 + \frac{1}{\langle N \rangle} \left( 1 + \frac{\tau}{\tau_D} \right)^{-1} \left( 1 + s^2 \frac{\tau}{\tau_D} \right)^{-1/2} \left( 1 + c_T \exp \left[ -\frac{\tau}{\tau_T} \right] \right) \quad (6)$$

The parameter  $s$  (describing the aspect ratio of the confocal volume) was set to 0.27; the triplet time  $\tau_T$  and the triplet amplitude  $c_T$  were averaged over all fits and used as constants in a second iteration of the fit. The remaining free parameters were the diffusion time,  $\tau_D$ , and the mean number of molecules in the confocal volume,  $\langle N \rangle$ . A calibration curve<sup>63</sup> was recorded to convert  $\langle N \rangle$  to the actual concentration of acceptor-labeled species. Two independent FCS curves (before and after the recording of time traces) were measured and averaged to obtain concentrations. The variation in  $\langle N \rangle$  between those two measurements was <5%.

**Stopped-flow experiments.** Stopped-flow experiments were conducted either with biotinylated NCBD-Avi-Cy3B and ACTR-CF680R or with biotinylated ACTR-Avi-Cy3B and NCBD-CF680R in NaP buffer at 20  $^{\circ}\text{C}$  on a PiStar-180 spectrometer (Applied Photophysics). Cy3B-labeled protein was excited with the 546-nm line of an Hg/Xe lamp and the loss of donor fluorescence after binding was monitored. Fluorescence emission was filtered with a 550-nm long-pass filter (the concomitant increase in acceptor fluorescence could not be detected, as the photomultiplier tube is not sensitive to emission wavelengths >650 nm). Concentrations of labeled protein were monitored via absorbance at 680 nm before and after the stopped-flow measurement to correct for adhesion of protein to the glass syringe or tubing. In one experiment, 105 nM of donor-labeled NCBD were mixed with 130, 260, 530, and 1540 nM of acceptor-labeled ACTR. In the other experiment, 390 nM of donor-labeled ACTR were mixed with 250, 570, and 1600 nM of acceptor-labeled NCBD. The mixing ratio was always kept at 1:2.5 (donor:acceptor), with final concentrations of 30 nM NCBD-Avi-Cy3B or 110 nM ACTR-Avi-Cy3B as the non-varied species after mixing.

**MD simulations.** MD simulations were performed using GROMACS<sup>64</sup> version 2016.4 (<http://www.gromacs.org>). In a first step, the NCBD-ACTR complex was

equilibrated. The initial structure of the bound NCBD-ACTR complex from the NMR ensemble was used (PDB accession code: 1KBH)<sup>21</sup>. The molecular topology of the complex was created according to the parameters from the AMBER12 force field<sup>65</sup>, and the complex was placed in a dodecahedral box solvated with water molecules parameterized according to the TIP4PD water model<sup>66</sup>. This combination of force field and water model has proven effective for sampling the dynamics of IDPs<sup>66</sup>. Na<sup>+</sup> and Cl<sup>-</sup> ions were added to the solution to neutralize the net charge of the complex and reach an ionic strength of 0.15 M. The system was then energy-minimized using a steep descent algorithm with an initial step size of 0.01 nm for 5000 steps or when the maximum force converged below 10 kJ mol<sup>-1</sup> nm<sup>-1</sup>. The solvent was then equilibrated in two runs, each 500 ps long. A first equilibration run was performed in the nVT ensemble, in which the temperature was coupled every 0.1 ps at a value of 300 K using the V-rescale thermostat. Random velocities drawn from a Boltzmann distribution obtained at 300 K were assigned to each particle of the system. The second equilibration run was in the nPT ensemble, in which temperature and pressure, coupled isotropically, were kept constant at values of 300 K and 1 bar using the V-rescale thermostat and the Parrinello-Rahman barostat, respectively. Temperature was coupled every 0.1 ps, whereas pressure was coupled every 2.0 ps. In both the nVT and nPT runs, protein atoms were subject to a position-restraining potential of 1000 kJ mol<sup>-1</sup> nm<sup>2</sup>. Subsequent to solvent equilibration, a 500-ns MD production run was performed in order to equilibrate the protein complex, using a time step of 2 fs. Interactions were calculated by building pair-lists using the Verlet scheme. Van-der-Waals (vdW) and Coulomb interactions were computed using a cutoff of 1.4 nm. The use of interaction cutoffs longer than 1.0 nm has been shown to improve the sampling of the conformational dynamics of IDPs<sup>67</sup>. Beyond the cutoff, electrostatics were treated using the Particle Mesh Ewald (PME) algorithm using a grid spacing of 0.16 and a cubic B-spline interpolation level of 4.

The last conformer obtained from the equilibration run was used as starting point for well-tempered metadynamics simulations<sup>68</sup>, which were performed to generate initial conformations of NCBD with Pro20 either in *trans* or in *cis*. Thus, the  $\omega$  angle formed by the atoms <sup>19</sup>C<sub>α</sub>-<sup>19</sup>C-<sup>20</sup>N-<sup>20</sup>C<sub>α</sub> was used as the reaction coordinate, on which a history-dependent bias represented by Gaussian potential functions was applied. The simulations were carried out using GROMACS version 2016.4 patched with the Plumed plugin version 2.3<sup>69</sup> by adding Gaussian potentials of height equal to 0.2 kJ mol<sup>-1</sup> and width of 0.2 radians at a temperature of 300 K, using as bias factor a value of 6 and a Gaussian deposit time of 1 ps.

To sample the conformational dynamics of the NCBD-ACTR complex at equilibrium in either the *trans* or *cis* state, two conformations with Pro20 either in *trans* or in *cis* (RMSD difference 0.25 nm) were taken from the well-tempered metadynamics run and used as starting points for temperature replica exchange MD. The systems were prepared as in the run performed to initially equilibrate the system (as described above), and the solvent was equilibrated again in the nVT and nPT ensembles at 14 different temperatures between 298.15 K and 323.11 K. The equilibrated replicas were run for 300 ns each using a time step of 2 fs, and replicas were allowed to swap every 500 steps (every ps). The number of replicas and the respective set of temperatures were chosen using a temperature predictor<sup>70</sup> with the aim to achieve an exchange rate of at least 20% between replicas, with the exchange probability calculated for each replica after the simulations being between 21 and 22%. The analysis of the collected trajectory was then carried out considering the replica at 298.15 K using either GROMACS or VMD (<https://www.s.ks.uiuc.edu/Research/vmd/>) tools, as well as in-house analysis scripts. Contact analysis was performed using the program CONAN<sup>71</sup>. Hydrogen bonds were identified by choosing a distance between polar hydrogens and hydrogen acceptors lower than or equal to 0.3 nm and a donor-hydrogen-acceptor angle of less than 20°.

**Data availability.** Data supporting the findings of this manuscript are available from the corresponding author upon reasonable request. A custom module for Mathematica (Wolfram Research) used for the analysis of single-molecule fluorescence data is available upon request.

Received: 15 May 2018 Accepted: 19 July 2018

Published online: 20 August 2018

## References

- Liu, J. et al. Intrinsic disorder in transcription factors. *Biochemistry* **45**, 6873–6888 (2006).
- Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 (2015).
- Tompa, P., Schad, E., Tantos, A. & Kalmar, L. Intrinsically disordered proteins: emerging interaction specialists. *Curr. Opin. Struct. Biol.* **35**, 49–59 (2015).
- Dyson, H. J. & Wright, P. E. Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding Protein (CBP) and p300. *J. Biol. Chem.* **291**, 6714–6722 (2016).
- Dogan, J., Gianni, S. & Jemth, P. The binding mechanisms of intrinsically disordered proteins. *Phys. Chem. Chem. Phys.* **16**, 6323–6331 (2014).
- Bah, A. & Forman-Kay, J. D. Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem.* **291**, 6696–6705 (2016).
- Best, R. B. Computational and theoretical advances in studies of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **42**, 147–154 (2017).
- Salvi, N., Abyzov, A. & Blackledge, M. Atomic resolution conformational dynamics of intrinsically disordered proteins from NMR spin relaxation. *Prog. Nucl. Magn. Reson. Spectrosc.* **102–103**, 43–60 (2017).
- Schuler, B., Hofmann, H., Soranno, A. & Nettels, D. Single-molecule FRET spectroscopy and the polymer physics of unfolded and intrinsically disordered proteins. *Annu. Rev. Biophys.* **45**, 207–231 (2016).
- Soranno, A. et al. Integrated view of internal friction in unfolded proteins from single-molecule FRET, contact quenching, theory, and simulations. *Proc. Natl Acad. Sci. USA* **114**, E1833–E1839 (2017).
- Eaton, W. A. et al. Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327–359 (2000).
- Brucale, M., Schuler, B. & Samori, B. Single-molecule studies of intrinsically disordered proteins. *Chem. Rev.* **114**, 3281–3317 (2014).
- Milo, R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *Bioessays* **35**, 1050–1055 (2013).
- Shammas, S. L., Crabtree, M. D., Dahal, L., Wicky, B. I. & Clarke, J. Insights into coupled folding and binding mechanisms from kinetic studies. *J. Biol. Chem.* **291**, 6689–6695 (2016).
- Gianni, S., Dogan, J. & Jemth, P. Coupled binding and folding of intrinsically disordered proteins: what can we learn from kinetics? *Curr. Opin. Struct. Biol.* **36**, 18–24 (2016).
- Kiefhaber, T., Bachmann, A. & Jensen, K. S. Dynamics and mechanisms of coupled protein folding and binding reactions. *Curr. Opin. Struct. Biol.* **22**, 21–29 (2012).
- Brandts, J. F., Halvorson, H. R. & Brennan, M. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **14**, 4953–4963 (1975).
- Schmidpeter, P. A., Koch, J. R. & Schmid, F. X. Control of protein function by prolyl isomerization. *Biochim. Biophys. Acta* **1850**, 1973–1982 (2015).
- Kim, P. S. & Baldwin, R. L. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **51**, 459–489 (1982).
- Theillet, F. X. et al. The alphabet of intrinsic disorder: I. Act like a Pro: on the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disord. Proteins* **1**, e24360 (2013).
- Demarest, S. J. et al. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553 (2002).
- 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).
- 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).
- Lin, C. H. et al. A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. *Mol. Cell* **8**, 581–590 (2001).
- 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).
- Qin, B. Y. et al. Crystal structure of IRF-3 in complex with CBP. *Structure* **13**, 1269–1277 (2005).
- Waters, L. et al. Structural diversity in p160/CREB-binding protein coactivator complexes. *J. Biol. Chem.* **281**, 14787–14795 (2006).
- Blanco, M. & Walter, N. G. Analysis of complex single-molecule FRET time trajectories. *Methods Enzymol.* **472**, 153–178 (2010).
- McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* **91**, 1941–1951 (2006).
- Steinfeld, J. I., Francisco, J. S. & Hase, W. L. *Chemical Kinetics and Dynamics* 2nd edn (Prentice Hall, Inc., New Jersey, 1999).
- Dogan, J., Jonasson, J., Andersson, E. & Jemth, P. Binding rate constants reveal distinct features of disordered protein domains. *Biochemistry* **54**, 4741–4750 (2015).
- Harrison, R. K. & Stein, R. L. Substrate specificities of the peptidyl prolyl cis-trans isomerase activities of cyclophilin and FK-506 binding protein: evidence for the existence of a family of distinct enzymes. *Biochemistry* **29**, 3813–3816 (1990).
- Reimer, U. et al. Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J. Mol. Biol.* **279**, 449–460 (1998).
- Kern, D., Schutkowski, M. & Drakenberg, T. Rotational barriers of cis/trans isomerization of proline analogues and their catalysis by cyclophilin. *J. Am. Chem. Soc.* **119**, 8403–8408 (1997).

35. Melis, C., Bussi, G., Lummis, S. C. & Molteni, C. Trans-cis switching mechanisms in proline analogues and their relevance for the gating of the 5-HT<sub>3</sub> receptor. *J. Phys. Chem. B* **113**, 12148–12153 (2009).
36. Neale, C., Pomes, R. & Garcia, A. E. Peptide bond isomerization in high-temperature simulations. *J. Chem. Theory Comput.* **12**, 1989–1999 (2016).
37. Jemth, P., Mu, X., Engstrom, A. & Dogan, J. A frustrated binding interface for intrinsically disordered proteins. *J. Biol. Chem.* **289**, 5528–5533 (2014).
38. Haberz, P., Arai, M., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Mapping the interactions of adenoviral E1A proteins with the p160 nuclear receptor coactivator binding domain of CBP. *Protein Sci.* **25**, 2256–2267 (2016).
39. Dogan, J., Schmidt, T., Mu, X., Engstrom, A. & Jemth, P. Fast association and slow transitions in the interaction between two intrinsically disordered protein domains. *J. Biol. Chem.* **287**, 34316–34324 (2012).
40. Dogan, J., Mu, X., Engstrom, A. & Jemth, P. The transition state structure for coupled binding and folding of disordered protein domains. *Sci. Rep.* **3**, 2076 (2013).
41. Zhou, H. X. Intrinsic disorder: signaling via highly specific but short-lived association. *Trends Biochem. Sci.* **37**, 43–48 (2012).
42. 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).
43. Livengood, J. A. et al. p53 Transcriptional activity is mediated through the SRC1-interacting domain of CBP/p300. *J. Biol. Chem.* **277**, 9054–9061 (2002).
44. Kurokawa, R. et al. Differential use of CREB binding protein-coactivator complexes. *Science* **279**, 700–703 (1998).
45. Zhou, H. et al. Toward a comprehensive characterization of a human cancer cell phosphoproteome. *J. Proteome Res.* **12**, 260–271 (2013).
46. Gibbs, E. B. et al. Phosphorylation induces sequence-specific conformational switches in the RNA polymerase II C-terminal domain. *Nat. Commun.* **8**, 15233 (2017).
47. Follis, A. V. et al. Pin1-induced proline isomerization in cytosolic p53 mediates BAX activation and apoptosis. *Mol. Cell* **59**, 677–684 (2015).
48. Gustafson, C. L. et al. A slow conformational switch in the BMAL1 transactivation domain modulates circadian rhythms. *Mol. Cell* **66**, 447–457 (2017). e447.
49. Fassolari, M. et al. Minute time scale prolyl isomerization governs antibody recognition of an intrinsically disordered immunodominant epitope. *J. Biol. Chem.* **288**, 13110–13123 (2013).
50. Crabtree, M. D. et al. Conserved helix-flanking prolines modulate intrinsically disordered protein:target affinity by altering the lifetime of the bound complex. *Biochemistry* **56**, 2379–2384 (2017).
51. Hyeon, C., Lee, J., Yoon, J., Hohng, S. & Thirumalai, D. Hidden complexity in the isomerization dynamics of holliday junctions. *Nat. Chem.* **4**, 907–914 (2012).
52. Solomatina, S. V., Greenfeld, M., Chu, S. & Herschlag, D. Multiple native states reveal persistent ruggedness of an RNA folding landscape. *Nature* **463**, 681–684 (2010).
53. Rognoni, L., Most, T., Zoldak, G. & Rief, M. Force-dependent isomerization kinetics of a highly conserved proline switch modulates the mechanosensing region of filamin. *Proc. Natl Acad. Sci. USA* **111**, 5568–5573 (2014).
54. Schilling, J., Schoppe, J. & Pluckthun, A. From DARPins to LoopDARPins: novel LoopDARPin design allows the selection of low picomolar binders in a single round of ribosome display. *J. Mol. Biol.* **426**, 691–721 (2014).
55. Soranno, A. et al. Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc. Natl Acad. Sci. USA* **111**, 4874–4879 (2014).
56. Konig, I. et al. Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat. Methods* **12**, 773–779 (2015).
57. Vogelsang, J. et al. A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes. *Angew. Chem. Int. Ed. Engl.* **47**, 5465–5469 (2008).
58. Morger, D. et al. The three-fold axis of the HIV-1 capsid lattice is the species-specific binding interface for TRIM5alpha. *J. Virol.* **92**, e01541 (2018).
59. Gopich, I. V. & Szabo, A. Decoding the pattern of photon colors in single-molecule FRET. *J. Phys. Chem. B* **113**, 10965–10973 (2009).
60. Viterbi, A. J. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE T. Inform. Theory* **13**, 260–269 (1967).
61. Chung, H. S. et al. Extracting rate coefficients from single-molecule photon trajectories and FRET efficiency histograms for a fast-folding protein. *J. Phys. Chem. A* **115**, 3642–3656 (2011).
62. Sauer, M., Hofkens, J. & Enderlein, J. *Handbook of Fluorescence Spectroscopy and Imaging: From Single Molecules to Ensembles* (Wiley-VCH, Germany, 2011).
63. Zijlstra, N. et al. Rapid microfluidic dilution for single-molecule spectroscopy of low-affinity biomolecular complexes. *Angew. Chem. Int. Ed. Engl.* **56**, 7126–7129 (2017).
64. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
65. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
66. Piana, S., Donchev, A. G., Robustelli, P. & Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **119**, 5113–5123 (2015).
67. Mercadante, D., Wagner, J. A., Aramburu, I. V., Lemke, E. A. & Gräter, F. Sampling long- versus short-range interactions defines the ability of force fields to reproduce the dynamics of intrinsically disordered proteins. *J. Chem. Theory Comput.* **13**, 3964–3974 (2017).
68. Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, 020603 (2008).
69. Bonomi, M. et al. PLUMED: a portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **180**, 1961–1972 (2009).
70. Patriksson, A. & van der Spoel, D. A temperature predictor for parallel tempering simulations. *Phys. Chem. Chem. Phys.* **10**, 2073–2077 (2008).
71. Mercadante, D., Gräter, F. & Daday, C. CONAN: a tool to decode dynamical information from molecular interaction maps. *Biophys. J.* **114**, 1267–1273 (2018).

## Acknowledgements

We thank Jane Dyson, Peter Wright, Kaare Teilum, and Magnus Kjærgaard for the discussion of peak splitting due to peptidyl-prolyl *cis/trans* isomerization in their NMR data; Attila Szabo for the discussion of single-molecule kinetics; Damien Morger, Maximilian Mittelviehhaus, and Markus Grütter for the gift of Cyclophilin A; Jendrik Schöppe and Andreas Plückthun for the gift of a pAT222-pD expression vector; Tina Schelbert and Karin Buholzer for help with protein expression; and the Functional Genomics Center Zurich for expert mass spectrometry analysis. This work was supported by the Swiss National Science Foundation (grant numbers 200021\_169741, 310030B\_173333, CRSII5\_170976). Molecular dynamics simulations were carried out on Piz Daint at the CSCS Swiss National Supercomputing Centre.

## Author contributions

F.Z. and B.S. designed the research. F.Z. prepared the protein samples, conducted experiments, and analyzed and interpreted the data. D.M. conducted and analyzed the MD simulations. D.N. developed instrumentation and data analysis tools, and assisted with data analysis. B.S. supervised the work. F.Z. and B.S. wrote the paper with the help of the other authors.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-05725-0>.

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

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

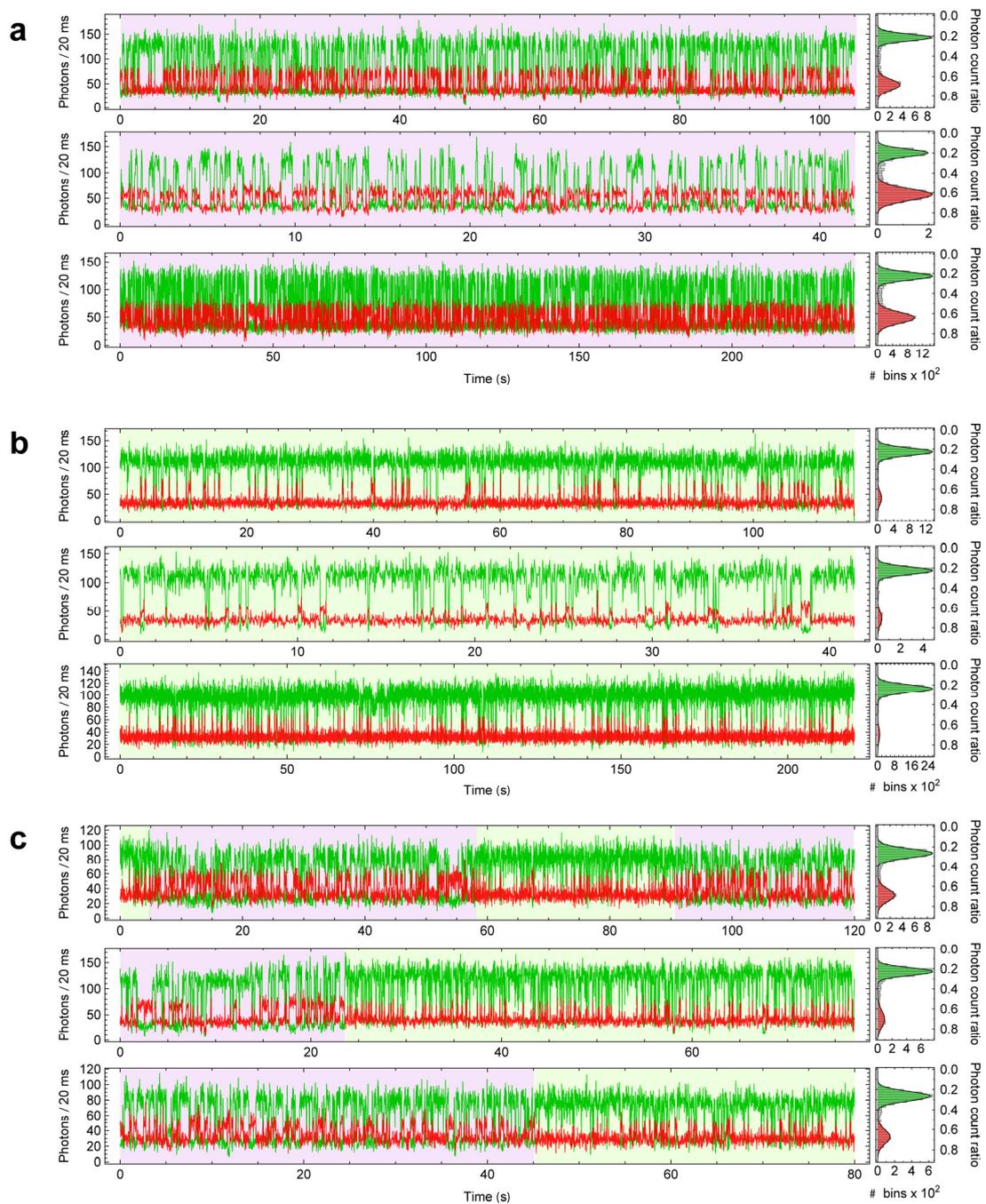
© The Author(s) 2018

## Supplementary Information for

# A proline switch explains kinetic heterogeneity in a coupled folding and binding reaction

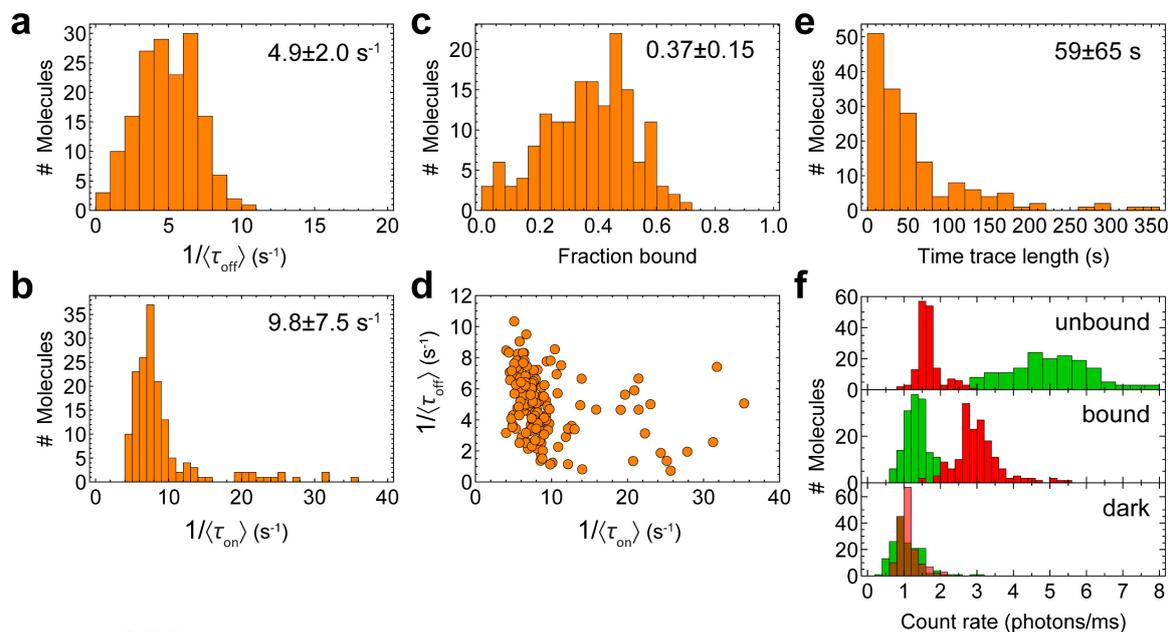
Zosel et al.

## Supplementary Figures

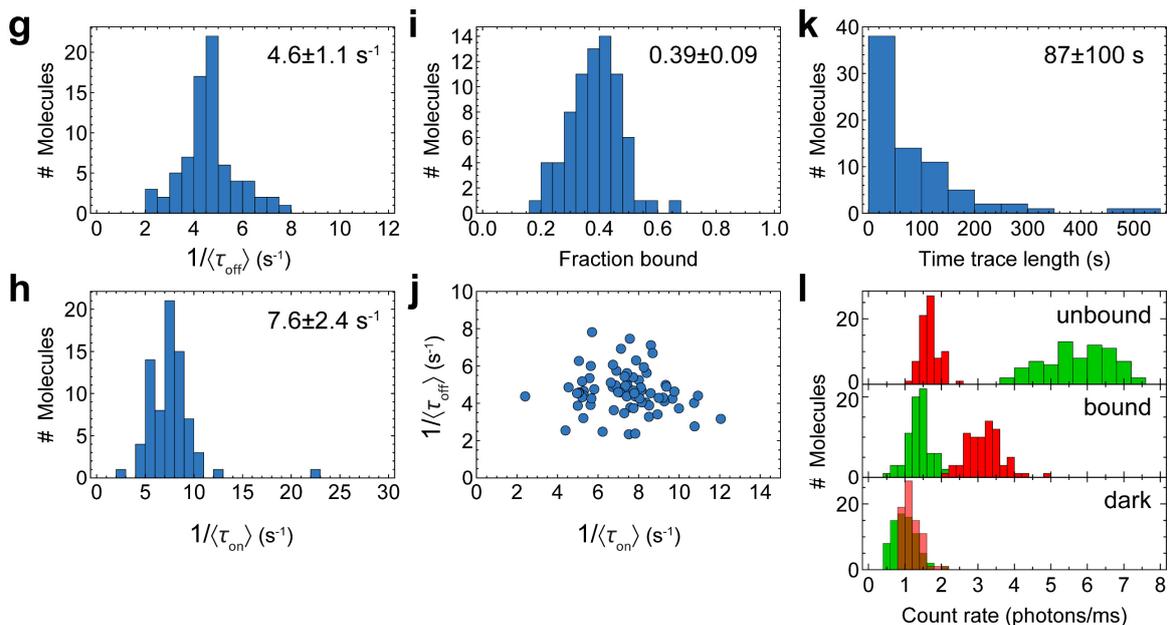


**Supplementary Figure 1. Representative time traces of acceptor-labeled ACTR binding to immobilized, donor-labeled NCBD.** Green: Donor photons; red: acceptor photons. The photon count ratio histograms (acceptor over total number of photons) of the traces are displayed to the right. Time traces were analyzed with the Viterbi algorithm using the model from Fig. 1e. **(a)** Examples of time traces where NCBD is in the high-affinity state for the whole time. **(b)** Examples of time traces where NCBD is in the low-affinity state for the whole time. **(c)** Examples of time traces where NCBD switches between high- and low-affinity state, with the high- and low-affinity segments (as identified by the Viterbi algorithm) shaded in light purple and light green, respectively.

## NCBD immobilized

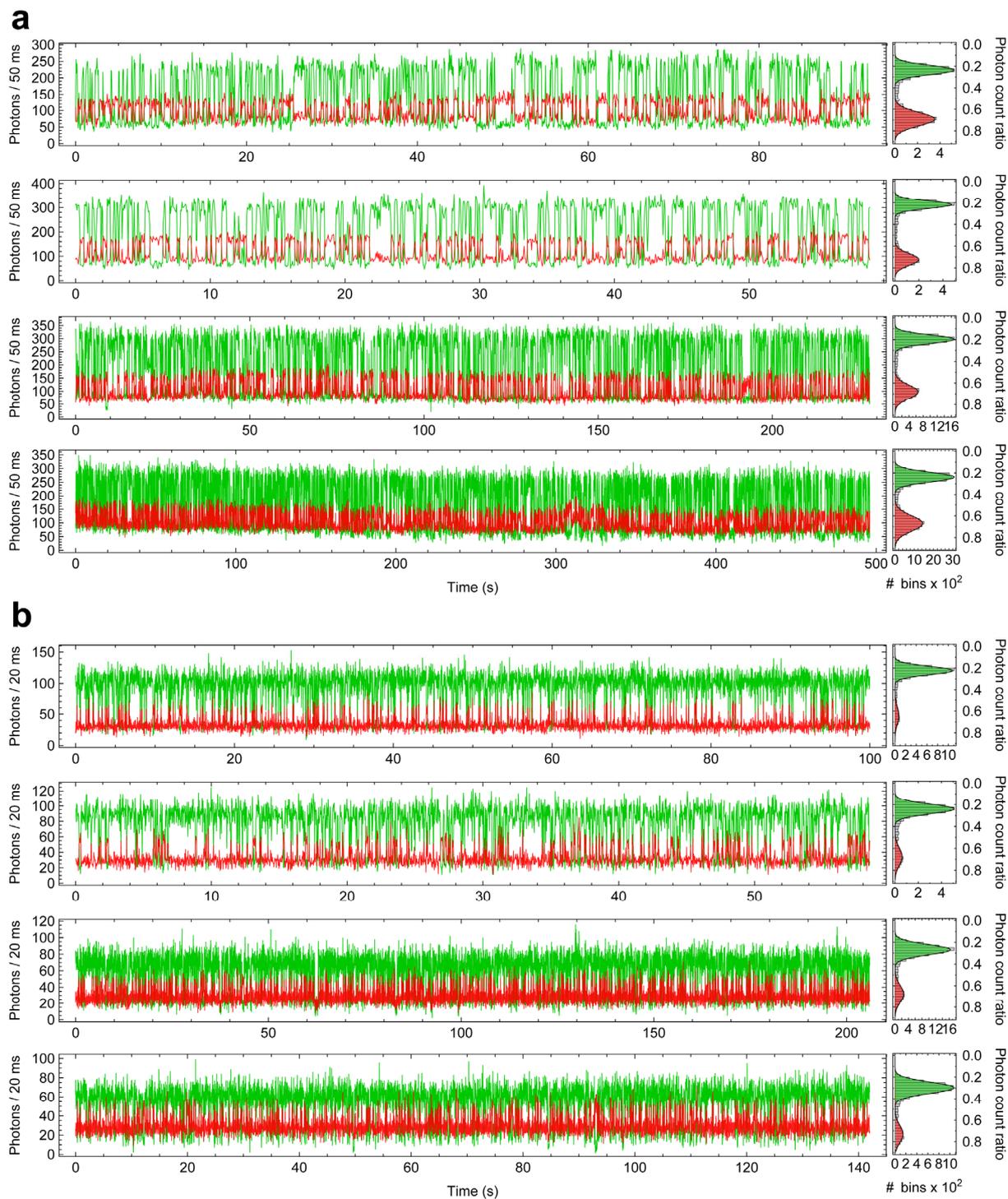


## ACTR immobilized

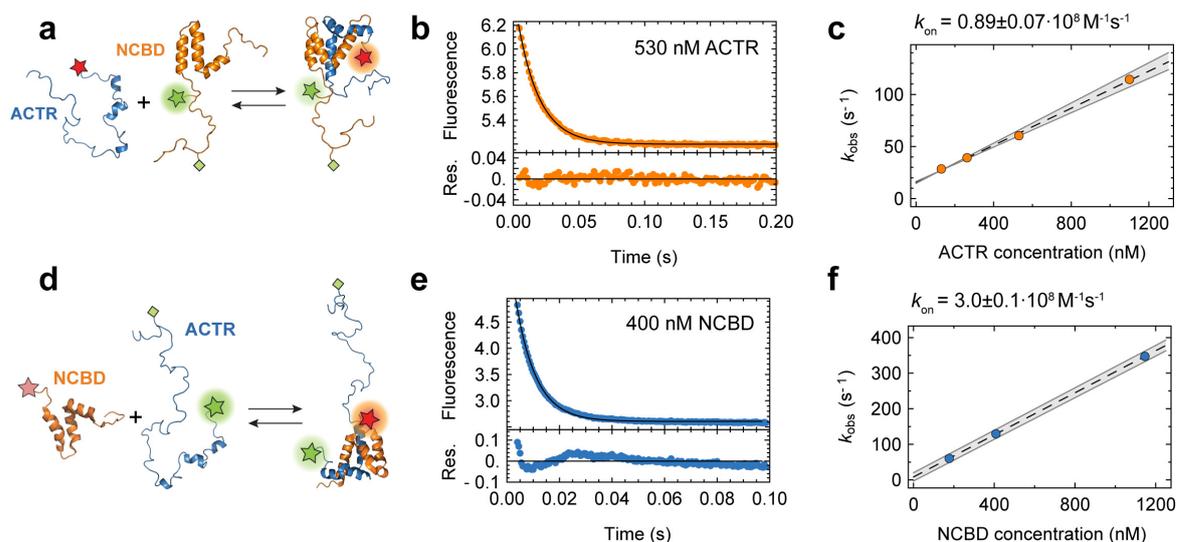


**Supplementary Figure 2. Molecule-to molecule variability of immobilized NCBD and ACTR.** Each event in the histograms corresponds to a single NCBD or ACTR molecule. Means and standard deviations of the displayed parameters are given. **(a-f)** Immobilized NCBD (cf. Fig. 1). **(a)**  $k'_{\text{on}}$  estimated from the reciprocal mean time spent in the unbound state,  $1/\langle\tau_{\text{off}}\rangle$ . **(b)**  $k_{\text{off}}$  estimated from the reciprocal mean time spent in the bound state,  $1/\langle\tau_{\text{on}}\rangle$ . The two subpopulations of NCBD (*cis/trans*) give rise to two clusters, around 7 and 30  $\text{s}^{-1}$ . Molecules that switch between the two subpopulations within a time trace appear with intermediate  $1/\langle\tau_{\text{on}}\rangle$ . **(c)** Fraction of time that molecules spend in the bound state, calculated from  $\sum \tau_{\text{on}} / \text{total time}$ . **(d)** Scatter plot of  $1/\langle\tau_{\text{on}}\rangle$  and  $1/\langle\tau_{\text{off}}\rangle$ . Molecules

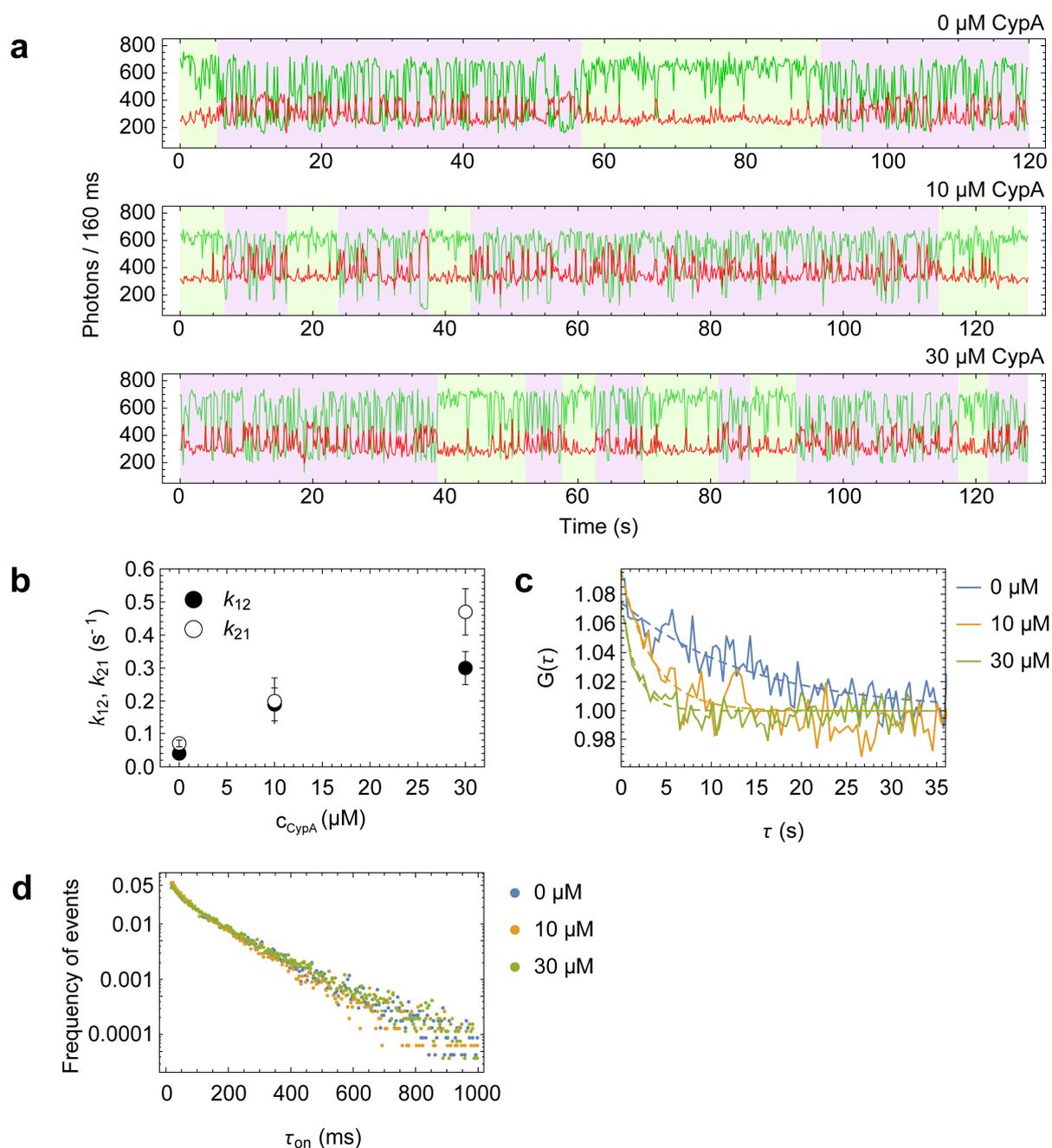
with a large  $1/\langle\tau_{\text{on}}\rangle$  tend to have a smaller  $1/\langle\tau_{\text{off}}\rangle$ , in accordance to the individual rate coefficients (Table 1). **(e)** Length of time traces until photobleaching. **(f)** Photon detection rates in the donor (green) and acceptor (red) channels, in the unbound (*upper panel*) and ACTR-bound state (*middle panel*), as well as in the dark state (*lower panel*). **(g-l)** Same as above, but for the experiment recorded with immobilized ACTR (cf. Fig. 2). The histograms of  $1/\langle\tau_{\text{off}}\rangle$ ,  $1/\langle\tau_{\text{on}}\rangle$  and the fraction bound are more homogeneous than for immobilized NCBD, since both NCBD subpopulations bind to ACTR randomly and their binding events are thus uncorrelated.



**Supplementary Figure 3. Representative time traces.** Green: Donor photons; red: acceptor photons. The photon count ratio histograms (acceptor over total number of photons) of the traces are displayed to the right. **(a)** Examples of time traces of NCBD binding to immobilized ACTR. **(b)** Examples of time traces of ACTR binding to immobilized NCBD P20A.

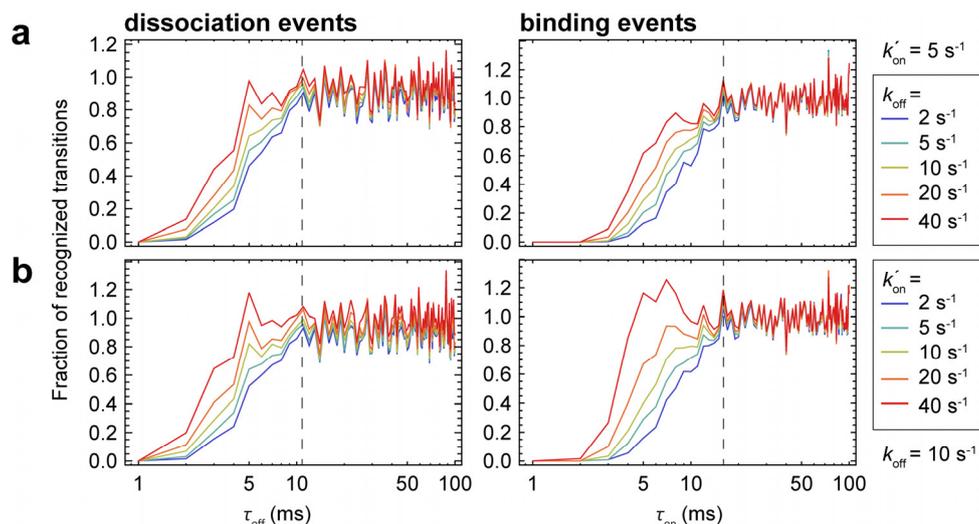


**Supplementary Figure 4. Immobilization has only a small influence on the association rate coefficient.** (a) Schematic representation of the reaction probed in the stopped-flow experiment. The same constructs as in the experiments with surface-immobilized molecules are used, but Cy3B-labeled, biotinylated NCBD remains free in solution. An excess of ACTR-CF680R is mixed with NCBD; the loss of Cy3B fluorescence reports on the formation of the NCBD-ACTR complex. (b) *Upper panel:* Average of five stopped-flow traces recorded at 530 nM ACTR, fitted with a single-exponential decay (black line). *Lower panel:* Residuals of the fit. (c) Association rate coefficients from stopped-flow experiment fitted with  $k_{\text{obs}} = k_{\text{on,2state}} \cdot c_{\text{NCBD}} + k_{\text{off}}$  (dashed line). Even though this two-state simplification does not reflect the true binding mechanism, it suffices to estimate the overall association rate coefficient,  $k_{\text{on}}$ . The largest error arises from the uncertainty in ACTR concentration, due to adsorption of the protein to the syringe and tubing (the gray-shaded bands and errors given correspond to the uncertainty estimated from the difference in protein concentrations determined before and after the stopped-flow measurement). (d-f) Same as (a-c), but for binding of CF680R-labeled NCBD to Cy3B-labeled, biotinylated ACTR.

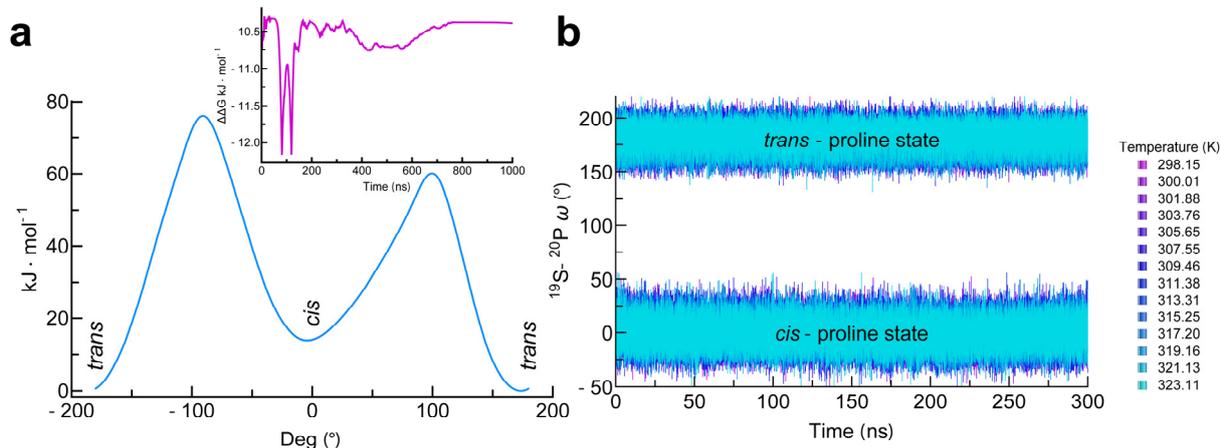


**Supplementary Figure 5. Influence of Cyclophilin A (CypA) on the kinetics of ACTR binding to NCBD.** (a) Representative time traces of CF680R-labeled ACTR binding to immobilized, Cy3B-labeled NCBD in the presence of CypA at the concentrations indicated. Green: donor emission; red: acceptor emission. The low-affinity segments (as identified by the Viterbi algorithm using the model from Fig. 1E) are shaded in light green, the high-affinity segments in light purple. (b)  $k_{12}$  and  $k_{21}$  as a function of CypA concentration determined by MLH analysis. (c) The correlation of the duration of consecutive binding events ( $G(\tau)$ , solid line) provides additional evidence that CypA accelerates switching between the two kinetic regimes.  $G(\tau)$  was calculated for all time traces containing  $> 80$  binding events and averaged. A value of  $G(\tau) > 1$  indicates clustering of binding events with similar duration (long in the

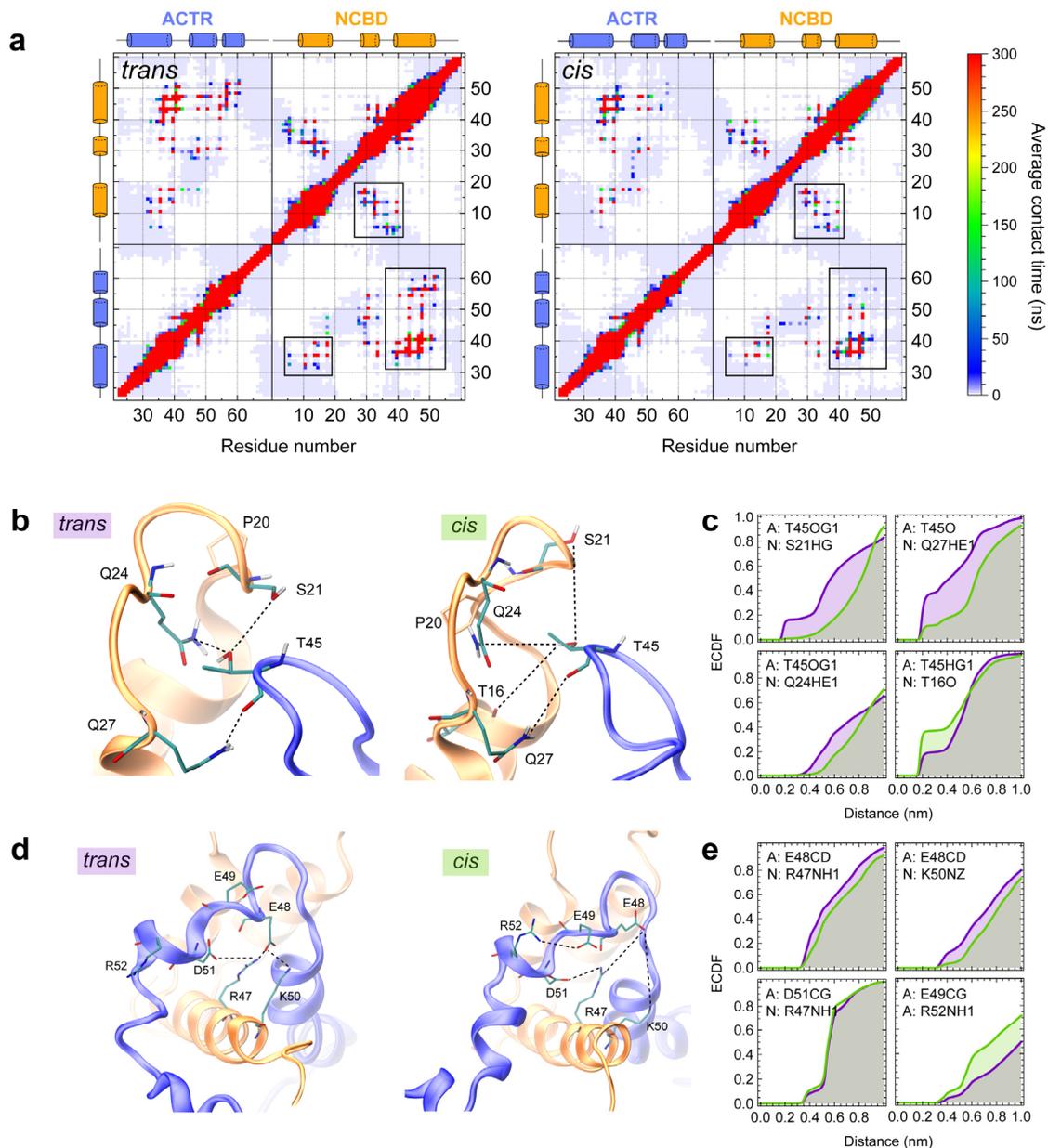
high-affinity NCBD1 state, short in the low-affinity NCBD2 state). The decay of  $G(\tau)$  reports on the exchange rate coefficient,  $k_{\text{ex}}$ , between the clusters. With increasing CypA concentration,  $G(\tau)$  decays faster. The decay rate coefficients of single-exponential fits to  $G(\tau)$  (dashed lines) are shown in Supplementary Table 1. **(d)** Overlay of the normalized bound-state dwell-time histograms in the absence and presence of CypA. The histograms are almost identical, indicating that CypA has no effect on the dissociation kinetics.



**Supplementary Figure 6. Performance of the Viterbi algorithm for detecting short events.** 162 single-molecule time traces were simulated based on the rate matrix  $\mathbf{K}_{4\text{state,blink}}$  and the best-fit parameters from Table 1 (see Methods). Photon rates and durations of the time traces were taken from the measurements (see Supplementary Fig. 2). The simulated data were analyzed with the Viterbi algorithm using the rate matrix  $\mathbf{K}_{2\text{state,blink}}$ , which distinguishes between an unbound and a bound state, as well as a dark state that arises from donor blinking. Plotted is the fraction of recognized events, calculated by dividing the dwell time histogram of the identified trajectory by the dwell-time histogram of the original simulated state trajectory. Different guess values for  $k'_{\text{on}}$  and  $k'_{\text{off}}$  were used, as indicated in the legend on the right. **(a)** guess value for  $k'_{\text{off}}$  was varied; **(b)** guess value for  $k'_{\text{on}}$  was varied. Binding events are identified reliably for  $\tau_{\text{on}} > 16$  ms; dissociation events for  $\tau_{\text{off}} > 11$  ms (dashed lines). Taking the different photon count rates in both states into account,  $\sim 70$  photons have to be detected for an event to be recognized correctly. Increasing the guess values for  $k'_{\text{on}}$  and  $k'_{\text{off}}$  facilitates the identification of shorter events but has no influence on the identification of longer events. In all cases, the Viterbi algorithm recognizes  $> 89\%$  of all events; that fraction increases to  $92\%$  if values of  $k'_{\text{on}} = 5 \text{ s}^{-1}$  and  $k'_{\text{off}} = 10 \text{ s}^{-1}$  are used, which are closest to the simulation input.



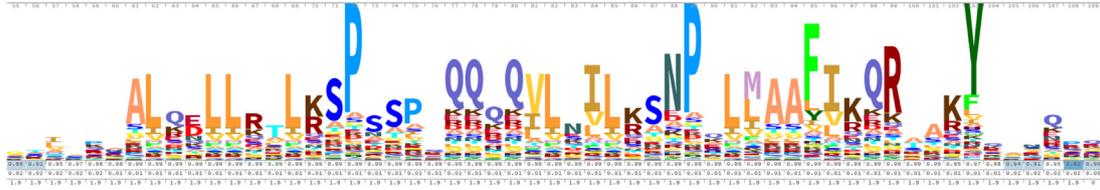
**Supplementary Figure 7. Free energy surface of *cis/trans* isomerization of Ser19-Pro20 obtained from WT-metadynamics and  $\omega$ -angle values sampled in REMD simulations. (a) Free energy profile of the *cis* and *trans* states of Pro20 integrated after 1  $\mu$ s WT-metadynamics simulations. The inset shows the free-energy difference between the wells defining the *cis* (between -50 and 50 degrees) and the *trans* states as a function of the simulated time. After 700 ns, the energy between the states is not subject to variations anymore, indicating that the simulation has converged. (b)  $\omega$ -angle as a function of time in the *trans* or *cis* states of Pro20 in REMD at the different temperatures chosen for each replica.**



**Supplementary Figure 8. Analysis of interactions within in the NCBD-ACTR complex from the MD simulations.** (a) Diagonalized contact maps colored according to the average contact time of two residues falling within a distance cutoff of 0.65 nm. The secondary structure elements of ACTR and NCBD (according to the NMR structure, PDB code 1KBH<sup>1</sup>) are indicated next to the matrix. The black boxes highlight parts of the contact map with noticeable differences between the NCBD Pro20 *trans* and *cis* complexes. The maps indicate that the reduced stability mostly arises from the loss of intermolecular contacts. (b-e) Representative conformations of ACTR (blue) and NCBD (orange), showing which interactions are perturbed when NCBD Pro20 changes its isomerization state from *trans* to *cis*. Residues are shown using a stick representation, with atoms colored by atom type (carbon – cyan, oxygen – red, nitrogen – blue and polar hydrogen atoms – white). Figures were created using VMD and rendered using the Tachyon ray tracer built into VMD<sup>2</sup>. In (b), the interactions involving the loop connecting helices  $\alpha 1$  and  $\alpha 2$  of NCBD are depicted. The loop shifts its binding register in the *cis* state upon losing some interactions that occur in the *trans* state. Particularly, ACTR residue T45, which is involved in a series of hydrogen bonds with S21, Q24 and Q27 of NCBD, loses these interactions and establishes, in *cis*, a hydrogen bond with T16. The corresponding empirical distribution functions

(*trans*: purple, *cis*: green) are depicted in (c). In (d), interactions toward the C-terminal end of the molecules are shown, with the corresponding empirical distribution functions (*trans*: purple, *cis*: green) depicted in (e). The network of salt bridges rearranges in the *cis* state, with intermolecular interactions (ACTR E48–NCBD R47 and K50) being weakened, and intramolecular salt bridges (ACTR E49–R52) strengthened instead. Notably, the interaction between ACTR D51 and NCBD R47, which falls in the same region and was indicated as a crucial determinant of the NCBD-ACTR complex stability before<sup>1</sup>, does not appear to be affected by the *cis/trans* switch.

...RSISPSALQDLLRTLKSPSSPQQQQVLNILKSNPQLMAAFIKQRTAKYVANQPG...



**Supplementary Figure 9. Sequence conservation in NCBD.** Relative amino acid probabilities at each position are shown as HMM logo<sup>3</sup>, based on the analysis of 236 Creb-binding (PF09030) sequences in the Pfam database<sup>4</sup>. The sequence of the NCBD construct used here is shown on top.

## Supplementary Tables

**Supplementary Table 1. Addition of the peptidyl-prolyl *cis/trans* isomerase Cyclophilin A (CypA) accelerates interconversion between NCBD populations.** Kinetic parameters of ACTR binding to surface-immobilized NCBD were obtained with MLH analysis, according to the kinetic model in Fig. 1E.  $k_{off,1}$  and  $k_{off,2}$  were constrained to the values in the absence of CypA, since the dwell-time histograms of the bound state are virtually identical under all conditions (cf. Supplementary Fig. 5d).  $k_{ex}$  reports on the decay of the correlation of the duration of consecutive binding events (cf. Supplementary Fig. 5c).

[CypA] ( $\mu\text{M}$ )	$k_{on,1} \cdot c_{ACTR}$ ( $\text{s}^{-1}$ )	$k_{on,2} \cdot c_{ACTR}$ ( $\text{s}^{-1}$ )	$k_{off,1}$ ( $\text{s}^{-1}$ )	$k_{off,2}$ ( $\text{s}^{-1}$ )	$c_{ACTR}$ (nM) <sup>a</sup>	$k_{on,1}$ ( $10^8 \text{ M}^{-1}\text{s}^{-1}$ ) <sup>b</sup>	$k_{on,2}$ ( $10^8 \text{ M}^{-1}\text{s}^{-1}$ ) <sup>b</sup>	$k_{12}$ ( $\text{s}^{-1}$ )	$k_{21}$ ( $\text{s}^{-1}$ )	$k_{ex}$ ( $\text{s}^{-1}$ ) <sup>c</sup>
0	6.0 $\pm$ 0.2	3.0 $\pm$ 0.5	7.3 $\pm$ 0.3	30 $\pm$ 3	65	0.93 $\pm$ 0.06	0.46 $\pm$ 0.08	0.04 $\pm$ 0.01	0.07 $\pm$ 0.01	0.07 $\pm$ 0.02
10	4.4 $\pm$ 0.1	2.2 $\pm$ 0.2	7.3 <sup>d</sup>	30 <sup>d</sup>	61	0.72 $\pm$ 0.04	0.36 $\pm$ 0.04	0.19 $\pm$ 0.05	0.20 $\pm$ 0.07	0.29 $\pm$ 0.09
30	6.2 $\pm$ 0.2	2.6 $\pm$ 0.3	7.3 <sup>d</sup>	30 <sup>d</sup>	72	0.86 $\pm$ 0.05	0.36 $\pm$ 0.05	0.30 $\pm$ 0.05	0.47 $\pm$ 0.07	0.58 $\pm$ 0.14

Errors are the standard deviation based on ten bootstrapping trials if not stated otherwise.

<sup>a</sup> Concentrations of acceptor-labeled ACTR were determined using FCS (see Methods). An uncertainty of  $\pm 5\%$  was estimated from two independent measurements conducted before and after recording the single-molecule time traces.

<sup>b</sup> Uncertainty from propagating the error of  $k'_{on}$  and the ACTR concentration.

<sup>c</sup> Determined by correlating the duration of consecutive binding events (Supplementary Fig. 5c).

<sup>d</sup>  $k_{off,1}$  and  $k_{off,2}$  were constrained to the values in the absence of CypA.

**Supplementary Table 2.** Table of all kinetic parameters of NCBD variants binding to surface-immobilized ACTR, determined with MLH analysis.

variant	$k_{on,1} \cdot C_{NCBD}$ (s <sup>-1</sup> )	$k_{on,2} \cdot C_{NCBD}$ (s <sup>-1</sup> )	$k_{off,1}$ (s <sup>-1</sup> )	$k_{off,2}$ (s <sup>-1</sup> )	$C_{NCBD}$ (nM) <sup>a</sup>	$k_{on,tot}$ (10 <sup>8</sup> M <sup>-1</sup> s <sup>-1</sup> )	# events <sup>c</sup>
wild type	3.3±0.3	1.6±0.3	5.6±0.5	30±6	17	2.9±0.2	34852
P23A	5.2±0.2	1.5±0.1	8.3±0.3	48±2	25	2.7±0.2 <sup>b</sup>	44380
P7/37A	2.6±0.1	0.89±0.10	7.4±0.2	55±5	19	1.8±0.1 <sup>b</sup>	31939
P7/23/37A	2.1±0.1	0.81±0.08	11±1	62±10	13	2.2±0.1 <sup>b</sup>	32058
P20A	4.3±0.2	n.a.	23±1	n.a.	14	3.2±0.2	24294
P20/23A	4.0±0.1	n.a.	25±1	n.a.	16	2.5±0.1	30348
P7/20/23A	5.8±0.2	n.a.	19±1	n.a.	26	2.2±0.2	31617
P20/23/37A	5.7±0.2	n.a.	33±1	n.a.	23	2.4±0.2	21841
P7/20/23/37A	6.8±0.1	n.a.	28±2	n.a.	32	2.1±0.1	35533

n.a. not applicable

Errors are the standard deviation based on ten bootstrapping trials if not stated otherwise.

<sup>a</sup> Concentrations of acceptor-labeled NCBD were determined using FCS (see Methods). An uncertainty of ±5% was estimated from two independent measurements conducted before and after recording the single-molecule time traces.

<sup>b</sup> For bi-exponential decays, the sum  $k'_{on,1} + k'_{on,2}$  was used to calculate  $k_{on,tot}$ .

<sup>c</sup> Sum of all detected binding and dissociation events.

**Supplementary Table 3.** Amino acid sequences of all protein variants used (after proteolytic cleavage of the His-Tag). The proline residue where the residue numbering of NCBD starts (position 1) is highlighted in bold face; positions where Pro residues were exchanged to Ala are underlined; the cysteine residues used for labeling are shown in red. The sequences correspond to UniprotKB entries Q92793 (CBP\_HUMAN), residues 2060-2216 (NCBD) and Q9Y6Q9 (NCOA3\_HUMAN), residues 1023-1093 (ACTR).

Variant	Sequence
NCBD-Avi	MAGLNDIFEAQKIEWHEGSMGS GSC PNRSISPSAL QDLLRTLKSP SSPQQQQQVL NILKSNPQLM AAFIKQRTAK YVANQPGMQ GGPR
NCBD-Avi P20A	MAGLNDIFEAQKIEWHEGSMGS GSC PNRSISPSAL QDLLRTLKSA SSPQQQQQVL NILKSNPQLM AAFIKQRTAK YVANQPGMQ GGPR
NCBD wt	GPC PNRSISPSAL QDLLRTLKSP SSPQQQQQVL NILKSNPQLM AAFIKQRTAK YVANQPGMQ
NCBD P20A	GPC PNRSISPSAL QDLLRTLKSA SSPQQQQQVL NILKSNPQLM AAFIKQRTAK YVANQPGMQC
NCBD P23A	GPC PNRSISPSAL QDLLRTLKSP SSAQQQQQVL NILKSNPQLM AAFIKQRTAK YVANQPGMQC
NCBD P20/23A	GPC PNRSISPSAL QDLLRTLKSA SSAQQQQQVL NILKSNPQLM AAFIKQRTAK YVANQPGMQC
NCBD P7/37A	GPC PNRSISASAL QDLLRTLKSP SSPQQQQQVL NILKSNAQLM AAFIKQRTAK YVANQPGMQC
NCBD P7/20/23A	GPC PNRSISASAL QDLLRTLKSA SSAQQQQQVL NILKSNPQLM AAFIKQRTAK YVANQPGMQC
NCBD P20/23/37A	GPC PNRSISPSAL QDLLRTLKSA SSAQQQQQVL NILKSNAQLM AAFIKQRTAK YVANQPGMQC
NCBD P7/23/37A	GPC PNRSISASAL QDLLRTLKSP SSAQQQQQVL NILKSNAQLM AAFIKQRTAK YVANQPGMQC
NCBD P7/20/23/37A	GPC PNRSISASAL QDLLRTLKSA SSAQQQQQVL NILKSNAQLM AAFIKQRTAK YVANQPGMQC
ACTR-Avi	MAGLNDIFEAQKIEWHEGSMGS GS GTQNRPLLRLN SLDDLVGPPS NLEGQSDERA LLDQLHTLLS NTDATGLEEI DRALGIPELV NQGQALEPKQ DC GGPR
ACTR	GP GTQNRPLLRLN SLDDLVGPPS NLEGQSDERA LLDQLHTLLS NTDATGLEEI DRALGIPELV NQGQALEPKQ DC

**Supplementary Table 4.** Primers used in the construction of NCDB and ACTR variants. All primers were obtained at Microsynth.

<b>Primer</b>	<b>Sequence</b>
NCBD_P20A_fw	GCT GAA ATC GGC ATC TTC ACC GCA ACA GCA ACA GC
NCBD_P20A_rev	GGT GAA GAT GCC GAT TTC AGC GTC CTC AGA AGG
NCBD_P23A_fw	CCA TCT TCA GCG CAA CAG CAA CAG CAA GTT CTT AAC
NCBD_P23A_rev	GCT GTT GCG CTG AAG ATG GCG ATT TCA GCG TCC
NCBD_P20A_P23A_fw	GCA TCT TCA GCG CAA CAG CAA CAG CAA GTT CTT AAC
NCBD_P20A_P23A_rev	GCT GTT GCG CTG AAG ATG CCG ATT TCA GCG TCC
NCBD_P7A_fw	GTA TAA GCG CGT CCG CCC TGC AGG ACC
NCBD_P7A_rev	GGC GGA CGC GCT TAT ACT ACG GTT CGG
NCBD_P37A_fw	AAA GCA ATG CTC AAC TGA TGG CGG CTT TC
NCBD_P37A_rev	CAG TTG AGC ATT GCT TTT AAG AAT GTT AAG
ACTR_Cys3rem_fw	CAG GGC CCT GGT ACC CAG AAT CGC CCG TTG C
ACTR_Cys3rem_rev	CTG GGT ACC AGG GCC CTG GAA TAA CAC CTC
ACTR_Avi_Cys3rem_fw	GGG ATC CGG CAG CGG TAC CCA GAA TCG CCC GTT GCT GCG
ACTR_Avi_Cys3rem_rev	CTG GGT ACC GCT GCC GGA TCC CAT GGA ACC TTC GTG CC

## Supplementary References

- 1 Demarest, S. J. *et al.* Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549-553, (2002).
- 2 Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33-38, (1996).
- 3 Schuster-Bockler, B., Schultz, J. & Rahmann, S. HMM Logos for visualization of protein families. *BMC Bioinformatics* **5**, 7, (2004).
- 4 Punta, M. *et al.* The Pfam protein families database. *Nucleic Acids Research* **40**, D290-301, (2012).