Diffusion of the disordered E-cadherin tail on β -catenin

Felix Wiggers¹, Samuel Wohl², Artem Dubovetskyi¹, Gabriel Rosenblum¹,
Wenwei Zheng^{3*} and Hagen Hofmann^{1*}

¹Department of Structural Biology, Weizmann Institute of Science, Herzl St.
234, 76100 Rehovot, Israel, ²Department of Physics, Arizona State University,
Tempe, AZ 85287, USA, ³College of Integrative Sciences and Arts, Arizona
State University, Mesa, AZ 85212, USA

*To whom correspondence may be addressed:

hagen.hofmann@weizmann.ac.il

wenweizheng@asu.edu

Abstract. Specific protein interactions typically require well-shaped binding interfaces. Here, we report a cunning exception. The disordered tail of the cell-adhesion protein E-cadherin dynamically samples a large surface area of the proto-oncogene β -catenin. Single-molecule experiments and molecular simulations resolve these motions with high resolution in space and time. Contacts break and form within hundreds of microseconds without dissociation of the complex. A few persistent interactions provide specificity whereas unspecific contacts boost affinity. The energy landscape of this complex is rugged with many small barriers ($3 - 4 k_B T$) and reconciles specificity, high affinity, and extreme disorder. Given the roles of β -catenin in cell-adhesion, signalling, and cancer, this Velcro-like design has the potential to tune the stability of the complex without requiring dissociation.

Specific molecular interactions orchestrate a multitude of simultaneous cellular processes. The discovery of intrinsically disordered proteins (IDPs)^{1,2} has substantially aided our understanding of such interactions. More than two decades of research revealed a plethora of functions and mechanisms^{2,3,4} that complemented the prevalent structure-based view on protein interactions. Even the idea that IDPs always ought to fold upon binding has largely been dismantled by recent discoveries of high affinity disordered complexes^{5,6}. Classical shape-complementary is indeed superfluous in the complex between prothymosin- α and histone H1 in which charge-complementary is the main driving-force for binding⁵. Yet, specificity remains limited in these complexes and it is unclear whether extreme levels of disorder, specificity, and affinity are compatible. In fact, many complexes between IDPs and folded proteins such as Sic1 and Cdc4⁷, the Na⁺/H⁺ exchanger tail and ERK2⁸, nucleoporin tails and nuclear transport receptors⁹ balance these factors. For instance, whereas nucleoporin tails are disordered in complex with Importin- β , affinities are weak and association/dissociation is fast⁹.

Here, we show that a cell adhesion complex involved in growth pathologies and cancer¹⁰ reconciles high affinity and specificity with flexibility. The intrinsically disordered cytoplasmic tail of E-cadherin retains disorder in complex with the multifunctional signalling factor β -catenin^{11,12,13}. Large segments of E-cadherin diffuse dynamically on the β -catenin surface. By integrating single-molecule FRET (smFRET) experiments with molecular simulations, we obtained a residue-resolved understanding of these motions: A small number of persistent interactions provide specificity whereas many weak multivalent contacts boost affinity. With this design, regulatory enzymes can easily access their recognition motifs on E-cadherin and β -catenin whereas the elastic properties of the complex might dampen forces exerted by neighbouring epithelial cells.

E-cadherin is a trans-membrane protein that mediates cell-cell adhesions by linking actin filaments of adjacent epithelial cells (Fig. 1a). The cytoplasmic tail of E-cadherin binds β -catenin, which establishes a connection to the actin-associated protein

Fig. 1. Complex between the cytoplasmic tail of E-cadherin (E-cad) and β -catenin (β -cat). (a) Schematics of cell-cell junctions mediated by E-cadherin and β -catenin. (b) The two X-ray structures of the complex between the tail of E-cadherin (red) and β -catenin (white) resolve different parts of E-cadherin (PDB: 1i7x). Bottom: Cartoon representation of the resolved E-cadherin parts. (c) Scheme showing the resolved parts of E-cadherin (red).

α -catenin^{14, 15, 16}. β -catenin on the other hand, is a multifunctional repeat protein^{17, 18, 19, 20} that mediates cadherin-based cell adhesions²¹ and governs cell fate decisions during embryogenesis²². It contains three domains: an N-terminal domain (130 aa), a central repeat domain (550 aa), and a C-terminal domain (100 aa). The 12 repeats of the central domain arrange in a super-helix²³ and the X-ray structure shows that the E-cadherin tail (hereafter E-cad) wraps around the central domain of β -catenin (hereafter β -cat)²³ (Fig. 1b). However, not only is half of the electron density of E-cad missing, the X-ray unit cell also comprises two structures with different resolved parts of E-cad (Fig. 1b). In fact, only 45% of all resolved E-cad residues are found in both structures (Fig. 1c). This ambiguity together with the large portion of missing residues might suggest²⁴ that E-cad is highly dynamic in complex with β -cat.

To understand how E-cad binds β -cat with high specificity and high affinity while retaining flexibility, we integrated smFRET experiments and molecular simulations. In our bottom-up strategy, we first probed intra-molecular interactions within E-cad using smFRET to parameterize a coarse-grained (CG) model. In a second step, we monitored E-cad on β -cat, integrated this information into the CG-model, and obtained a dynamic picture that reconciles specificity, affinity, and disorder.

E-cad is expanded in solution. To probe the conformation of E-cad, we labelled different regions with AlexaFluor 488 as donor and AlexaFluor 594 as acceptor. We created three constructs in which the FRET labels map the N-terminal (A), central (B), and C-terminal (C) segments of E-cad (Fig. 2a, Supplementary Table 1). In addition, we also checked the longer segments AB, BC, and ABC to characterize the global behaviour of E-cad. The labelled E-cad constructs were monitored while freely diffusing through the confocal volume of our microscope. For all six constructs, we obtained homogeneous FRET distributions (Fig. 2a) with mean positions that scaled with the sequence separation between the dyes (Fig. 2b), as expected for an IDP. The persistence lengths of all segments, except of the A-segment, significantly exceeded the value of 0.4 nm found for other disordered or unfolded proteins²⁵ (Fig. 2c). Hence, E-cad is expanded in solution, likely due to electrostatic repulsions given its high net charge (-22)^{26, 27, 28}. When we screened these repulsions with KCl, constructs ABC, AB, and BC collapsed substantially, a feature that is well described by mean-field polyampholyte theory²⁹ (Fig. 2d). Yet, compaction was not uniform across the chain. Whereas the local segments B and C also compacted, the N-terminal A-segment expanded (Fig. 2d). Such chain expansion upon charge screening is known for polyampholytes with balanced numbers of positive and negative charges^{28, 29, 30, 31, 32}. However, the A-segment is not charge-balanced. It even has a more negative net charge per residue (-0.186) than B (-0.093) and C (-0.166), thus excluding polyampholyte effects²⁹. Instead, the A-segment exhibits significant charge segregation: a positively charged N-terminus is followed by a negative C-terminal part (Fig. 2a). Such charge patterning has previously been shown to compact disordered proteins³³, which we

Fig. 2 E-cad is expanded in solution. **(a)** SmFRET histograms of six segments of E-cad (A, B, C, AB, BC, ABC) with n indicating the total number of molecules. A scheme shows the division of E-cad into the segments probed by FRET and the distribution of positively (blue) and negatively (red) charged residues. Histogram insets indicate the labelling sites as green and red spheres for donor and acceptor, respectively. **(b)** The mean FRET efficiencies scale with the sequence separation of the dyes. The colour code is identical to (a). **(c)** Persistence lengths of the individual segments. Error bars are the uncertainty introduced by the model of the distance distribution used (see Methods). **(d)** Salt-induced change in the donor-acceptor distance R_{DA} of the six E-cad segments (see Methods). A fit with the polyampholyte theory^{28, 29} (dashed line) and the best fit of the CG-model (solid line) are shown for comparison. **(e)** Length-scaling exponent map resulting from the best-fit CG-model at low (62 mM, upper panel) and high (512 mM, lower panel) ionic strength. Inset: Segment A.

confirmed using a sequence charge decoration metric ($SCD_{\text{low salt}}$)^{34, 35} (Supplementary Fig. 1a). Mean-field polyampholyte theory is therefore inappropriate to describe E-cad, despite its success for other IDPs^{28, 30} (Fig. 2d). To convert our smFRET experiments into structural ensembles, we modelled E-cad as a coarse-grained (CG) heteropolymer of beads. Each bead corresponded to an amino acid with charge (+1, 0, -1) and hydrophobicity. The strength of hydrophobic interactions was adjustable via a global parameter ϵ (see Methods). Remarkably, a single value ($\epsilon = 0.16$ kcal/mol) reproduced all experimental data, including those for the A-segment (Fig. 2d). A 2D-map of length scaling exponents (ν) for individual residue pairs³⁶ now clearly uncovers the effect of charge segregation in the A-segment (Fig. 2e). The scaling exponents for residue pairs from the positively charged N-terminus and the negatively charged C-terminus of the A-segment are small ($\nu < 0.5$), which suggested strong attractions between the termini. Swapping a pair of charged residues between the termini would therefore revert the salt-dependence of the A-segment (Supplementary Fig. 1b). Notably, high salt concentrations (>500 mM) collapsed all segments (Fig. 2d), a salting-out effect³⁰ that is irrelevant at the physiological salt concentrations relevant for our study.

E-cad forms a pliable complex with β -cat. To monitor E-cad in the complex, we added unlabelled β -cat, which gave rise to a second population of E-cad molecules with altered FRET efficiencies (Fig. 3a). In most segments, the FRET efficiency was higher than in free E-cad, indicating a compaction upon binding to β -cat. An exception was again the A-segment. The disruption of interactions between its oppositely charged termini slightly expanded this segment upon binding to β -cat. We next performed smFRET experiments at different β -cat concentrations to determine the affinity of the complex (Fig. 3b), which we found to be in the range of 4 ± 2 nM (Fig. 3c), i.e., very high for all constructs. This agreed with previous measurements of unlabelled E-cad³⁷, showing that labelling did not interfere with binding. Remarkably, whereas affinity was only moderately affected by salt (Fig. 3c), we found that increasing KCl concentrations caused a significant decrease in the FRET efficiencies of E-cad, indicative of an expansion (Fig. 3d). This solvent sensitivity suggested that E-cad in complex with β -cat retains significant flexibility.

To understand this flexibility, we extended our CG-model by intermolecular contacts between E-cad and β -cat (see Methods). We first tested a CG-model in which we rigidified the parts of E-cad that are resolved in the X-ray structures²³. Hence, structured parts of E-cad were essentially kept static while the unresolved parts were flexible but allowed to contact β -cat. Indeed, this ‘rigid’ model described the observed FRET efficiencies (Fig. 3e). The model even partially captured the experimentally observed expansion of E-cad at higher salt concentrations (Fig. 3d), suggesting that flexible parts are responsible for the effect. How flexible are the individual segments?

Fig. 3 The conformation of E-cad in complex with β -cat is pliable. **(a)** smFRET histograms of the six E-cad constructs in the presence of β -cat (A: 4 nM, B: 85 nM, C: 15 nM, AB: 7 nM, BC: 8.5 nM, ABC: 5 nM). Solid lines are fits of the histograms with a superposition of Gaussian and log-normal peaks and dashed lines represent free E-cad (see Methods). The number of molecules in each histogram is indicated by n . **(b)** Binding isotherms of selected E-cad segments (filled circles) and of the core-binding peptide (white circles). **(c)** Dissociation constants (K_D) for the six constructs at 82 mM ionic strength and a comparison of the affinity of the ABC segment at higher ionic strength. **(d)** Change in FRET efficiency of the E-cad constructs ABC and AB in complex with β -cat with increasing ion concentrations. Circles indicate the experimental data and lines are the result from the rigid CG-model when keeping the X-ray resolved E-cad parts static (dashed) and the best fit of the flexible CG-model. **(e)** Comparison of the experimental FRET efficiencies with the rigid CG-model (open circles) and the flexible CG-model (filled circles). **(f)** Donor-acceptor distance correlation functions from the rigid CG-model (dashed lines) and the flexible CG-model (solid lines).

Our model predicted extensive motions in all segments except the C-segment, which interacted rigidly with β -cat (Fig. 3f). Unfortunately, simplifications inherent to coarse-grained models prevented accurate timescale estimates. We therefore measured the dynamics of E-cad in complex with β -cat experimentally.

E-cad is highly dynamic in complex with β -cat. To probe the dynamics of the E-cad ensemble, we performed nanosecond FCS (nsFCS)^{38, 39} experiments of labelled E-cad in the absence and presence of β -cat. Disordered proteins such as E-cad exhibit large-scale distance fluctuations that cause anti-correlated donor-acceptor intensity fluctuations³⁹. Indeed, the cross-correlation functions of free E-cad reveal anti-correlated decays in the 100 ns timescale for almost all segments (Fig. 4a). Only the C-segment lacks this signal, likely due to inter-dye contacts caused by a close proximity of the dyes in this short segment⁴⁰ (Supplementary Fig. 2a).

When we formed the complex by adding β -cat, the correlation functions changed. Whereas signals in A and AB were nearly unaltered, the cross-correlation amplitude of the ABC segment decreased, indicating reduced fluctuations at the sub-microsecond timescale (Fig. 4a). Segments B, C, and BC even showed positive cross-correlation amplitudes, thus providing no information about potential distance dynamics. Dye-sticking to protein surfaces had previously been reported as cause for such positive cross-correlation amplitudes⁴¹, but we excluded this possibility with fluorescence anisotropy experiments (Supplementary Fig. 3). Instead, we found increased static quenching, likely due to dye contacts with aromatic residues of β -cat (Supplementary Fig. 2b).

To bypass this effect, we determined donor fluorescence lifetimes for all segments, a quantity that is unaffected by static quenching. For a static well-defined distance, these lifetimes are expected to correlate linearly with the FRET efficiency³⁹ (Fig. 4b). Yet, in contrast to this prediction, we found substantial deviations from the expected values, suggesting an ensemble of structures. Importantly, we found this deviation also for the C-segment (Fig. 4b). This is surprising given that the C-segment is resolved in the X-ray structure (Fig. 1b), did not show flexibility in our nsFCS experiments (Fig. 4a) and therefore was not expected to be dynamic. However, nsFCS experiments only probe dynamics on sub-microsecond timescales but cannot capture slower motions. In fact, we observed substantial broadening of the FRET peaks of all segments in complex with β -cat (Fig. 4b inset), which is symptomatic for dynamics in the millisecond regime.

We can readily identify such slow dynamics using RASP (recurrence analysis of single particles)^{42, 43, 44}. The method exploits the fact that a diffusing molecule can enter and exit the confocal volume multiple times. Once a molecule leaves the observation volume, the chance of it returning within a short time interval is greater than the chance of detecting a new molecule. This effect allowed us to monitor conformational

Fig. 4 Timescales of motion of E-cad free and bound to β -cat. **(a)** Donor-acceptor cross-correlation functions of all six FRET constructs of E-cad in the absence (darker colour) and presence (lighter colour) of saturating amounts of β -cat (100 nM). Solid lines are fits to a product of exponential terms (see Methods). **(b)** 2D correlation map between the donor fluorescence lifetime and FRET efficiency of all six E-cad constructs in complex with β -cat. The solid line shows the expected dependence for a single donor-acceptor distance. The dashed line indicates the dependence for a Gaussian chain distribution as an upper limit of distance heterogeneity. Inset: Measured FRET histogram of the ABC segment of E-cad in complex with β -cat in comparison to the expected shot-noise limit (white area). **(c)** RASP time decays of the fraction of molecules with FRET values lower than the mean value of the FRET histogram for all six E-cad constructs in complex with β -cat. Solid lines are exponential fits of the data. Inset: residuals of the exponential fit. **(d)** Comparison of the relaxation times of free E-cad (upper panel) and β -cat bound E-cad (lower panel) as a function of the donor-acceptor sequence separation. The Pearson correlation coefficient is indicated.

switching between the low-FRET (left) and high-FRET (right) side of the FRET distributions (see Methods, Supplementary Fig. 4). The resulting exchange kinetics indeed showed pronounced decays on timescales of micro- to millisecond for all six constructs (Fig. 4c), including the C-segment. The relaxation times ranged from $250 \pm 30 \mu$ s for the C-segment to $820 \pm 50 \mu$ s for the BC-segment, i.e., three orders of magnitude slower than for free E-cad. Moreover, the relaxation times did not scale with the sequence separation of the dyes as expected for polymers^{45, 46, 47} such as free E-cad (Fig. 4d). This showed that motions were determined by contacts between E-cad and β -cat rather than within E-cad. Most importantly, these experiments clearly showed that even the C-segment, which was expected to be rigidly bound to β -cat, is highly dynamic. How can we reconcile this finding with the X-ray structure? In fact, the C-segment is only resolved in one of two structures in the unit cell (Fig. 1b,c). Such ambiguity has also been observed in other complexes²⁴ and suggests that the C-segment can indeed dissociate from the β -cat surface, which explains the dynamics found experimentally. Hence, our results show that the E-cad/ β -cat complex is more dynamic than indicated by the X-ray structure and our rigid CG-model.

A structural model of E-cad in complex with β -cat. To obtain a more realistic picture of the E-cad ensemble in complex with β -cat, we revised our CG-model. Instead of fully rigidifying the resolved parts of E-cad, we introduced a parameter (ξ) that

Fig. 5 The flexible CG-model predicts a heterogeneous structural ensemble of E-cad in complex with β -cat. (a) Distribution of contact types in the complex. The classification of contacts is indicated. A Kolmogorov-Smirnov test gives a probability of $p = 0.37$ that both distributions originate from the same sample. (b) Intermolecular contact probabilities (P) based on the flexible CG-model (blue) overlaid with the contact map derived from the X-ray map (gray dots). The highest contact probabilities are found in a 20 amino acid long fragment in the B-segment (gray bar) that is well resolved in both X-ray structures (right panel). (c) Structure of β -cat indicating the directionality of the axis between (left) and ensemble of the centre of mass positions (right) of A (blue), B (gray), and C (red) segment. (d) Centre of mass angular distribution along the axis between the N-terminus (NT) and C-terminus (CT) of β -cat (schematically indicated) at low (82 mM) and high (532 mM) ionic strength. Colour code is identical to (c). (e) Centre of mass positional distribution of E-cad projected onto the NT-CT axis of β -cat for the segments A, B, and C at two ionic strengths (indicated). The colour code is identical to (d) and (c).

tunes the strength of all intermolecular X-ray contacts in a global manner. We found that $\xi = 0.6$ kcal/mol described the experimentally determined FRET-values and the salt-sensitivity of the complex equally good as the rigid model (Fig. 3d-e). Most importantly, despite the fact that all X-ray contacts share the same interaction parameter ξ , the model indeed predicted a dynamic C-segment (Fig. 3f), in accord with our experimental findings.

To understand the deviations of this model from the known X-ray structure, we compared the distribution of contact types. To this end, we classified amino acids as polar (P), charged (C), hydrophobic (H), and other amino acids (O). Interestingly, the most abundant intermolecular contacts in the CG-model were of the uncommon H-C type instead of the commonly discussed C-C type (Fig. 5a), which explains the weak salt-dependence of the binding affinity (Fig. 3c). However, this distribution of contact types agrees well with the X-ray structure (Fig. 5a) and an intermolecular contact map shows that interactions found in the X-ray structure were also preserved in the CG-model (Fig. 5b). Yet, the smFRET-based CG-model differed from the known structure in one important aspect. Contact probabilities (Fig. 5b) and lifetimes (Supplementary Fig. 5) were broadly distributed in the CG-model, reflecting many weak and unspecific interactions. The highest contact probabilities, i.e., the most persistent interactions, were found in a short 20 amino acid stretch of the B-segment (Fig. 5b), a region that is well resolved in both X-ray structures (Fig. 1b,c). It has previously been identified as

core-binding region of E-cad⁴⁸ and likely provides specificity. To disentangle contributions of specific and unspecific interactions to the overall stability of the E-cad/ β -cat complex, we determined its affinity to β -cat (Fig. 3b). We found a low affinity of the core-binding region with an upper limit of $K_D > 1 \mu\text{M}$ compared to $K_D \sim 4 \text{ nM}$ for full-length E-cad. Hence, flexible segments of E-cad boost affinity by more than $5.5 k_B T$. How are these segments distributed on the surface of β -cat?

The three-dimensional distribution of E-cad on β -cat differed strongly between segments (Fig. 5c). For instance, not only did the A-segment contact the N-terminus (NT) and the C-terminus (CT) of β -cat, but also the C-segment explored a large surface area of β -cat. The ensemble of the B-segment on the other hand was much more confined, as expected from the higher contact probabilities and the fact that 42% of it is resolved in both X-ray structures (Fig. 1c). We visualized the ensembles of the three segments by the distribution of their centre-of-mass along two coordinates: an angular coordinate describing the distribution around the long axis of β -cat (Fig. 5d), and an axial coordinate for the distribution along β -cat (Fig. 5e). These distributions showed that the observed salt-induced changes of E-cad in complex with β -cat (Fig. 3d) were predominantly due to changes in the A-segment (Fig. 5e,d) whereas changes in the B- and C-segment were much less pronounced.

In summary, the refined flexible CG-model indeed recapitulated the experimentally observed large-scale motions of E-cad on the β -cat surface. A total of 91% of the β -cat surface was explored by E-cad in a rather diffusive, i.e., continuous, manner without encountering major barriers. However, given the simplification inherent to CG-models, it remains to be determined whether a continuous diffusion of E-cad on the β -cat surface is realistic. An alternative model requires the concerted association and dissociation of E-cad segments, thus leading to dominant barriers that must be overcome to reconfigure on the β -cat surface. In fact, such a scenario would explain why E-cad motions in the complex (Fig. 4c) are orders of magnitude slower than those found for free E-cad (Fig. 4a). We therefore aimed at identifying the magnitude of these barriers.

Intra-chain diffusion of E-cad on β -cat. The experimental RASP-kinetics of the six constructs agreed with exponential decays (Fig. 4c), a hallmark of two-state kinetics, and indicative of a dominant barrier. We tested this hypothesis by monitoring the motions of E-cad ABC at different temperatures (Supplementary Fig. 6). A barrier should cause Arrhenius behaviour, i.e., a significant slow-down of the dynamics with decreasing temperature. Yet, we only found a twofold slow-down of the kinetics in the range from 37°C to 5°C (Fig. 6a). A fit with the Arrhenius equation $\tau = A \exp(\beta E_a)$ with $\beta^{-1} = k_B T$ yielded an activation energy of $E_a = 28 \pm 16 \text{ kJ/mol}$ (Fig. 6b). However, this analysis neglected that the pre-exponential factor A depends on the viscosity of the medium⁴⁹, which itself scales with temperature.

Fig. 6 E-cad dynamics in complex with β -cat are diffusive. **(a)** Normalized time decays of the E-cad ABC construct at two temperatures (indicated). Solid and dashed lines are fits with an exponential function and with the Smoluchowski equation, respectively. **(b)** Relaxation times of the E-cad ABC construct as function of temperature. The solid line is the dependence without activation energy (see Methods). The red area indicates the 90% confidence interval of this fit. The dashed line is an Arrhenius fit. Inset: Donor-acceptor distance distributions for E-cad bound (black) and free (gray) from smFRET-based CG-model. **(c)** Comparison of the diffusion coefficient of E-cad free (gray circle) and bound to β -cat (white circles). Shaded areas represent the effect of different roughness values (σ in $k_B T$ at 23°C). Black solid and dashed lines are fits with the Zwanzig expressions⁵⁰ for periodic ($\sigma = 4.3 \pm 0.1$) and Gaussian ($\sigma = 2.9 \pm 0.1$) modulated roughness, respectively. **(d)** Scheme of E-cad on the surface of β -cat. Parts of E-cad can partially detach, thus giving access to regulatory enzymes. **(e)** Same as d but illustrating the spring-like properties of the E-cad/ β -cat complex.

Accounting for this effect, we found that the slow-down was also reproduced without activation energy (Fig. 6b). We therefore checked whether the kinetics were indeed consistent with single-well, i.e., barrier-less diffusion as suggested by our CG-model. To this end, we used the donor-acceptor distance distribution $P(r)$ for the ABC construct from our flexible CG-model (Fig. 6b inset) to fit the experimental kinetics by solving the Smoluchowsky equation for diffusion in a single-well potential given by $V(r) = -\ln P(r)$. Indeed, this barrier-less model also resulted in excellent fits. Intra-chain diffusion coefficients ranged from $(4.4 \pm 0.1) \times 10^{-3} \text{ nm}^2/\mu\text{s}$ to $(9.5 \pm 0.1) \times 10^{-3} \text{ nm}^2/\mu\text{s}$ in the temperature range from 5°C to 37°C (Fig. 6c, see Methods). Notably, intra-chain diffusion was four orders of magnitude slower than that of free E-cad ABC ($42 \pm 1 \text{ nm}^2/\mu\text{s}$, see Methods) despite the absence of a dominant barrier (Fig. 6c).

Ruggedness in energy landscapes, i.e., a plethora of small barriers, is known to slow down motions as effectively as a dominant barrier^{50, 51, 52, 53}. Zwanzig⁵⁰ derived expressions for the apparent diffusion coefficient (D^*) (i) for randomly distributed barriers with average height σ , leading to $D^* = D e^{-\beta^2 \sigma^2}$ and (ii) for a periodic modulation, leading to $D^* = D e^{-2\beta \sigma}$. Here, D and D^* are diffusion coefficients in the absence and presence of ruggedness, respectively. Experimentally, it has been challenging to obtain absolute values for σ because D and D^* are typically not simultaneously known. Our experiments allowed us to overcome this problem. Since the donor-acceptor distance distribution of E-cad ABC changed only slightly upon

binding to β -cat (Fig. 6b inset), the free energy potentials obtained from the smFRET-based CG-model were very similar between free and bound E-cad. Hence, D and D^* were simply the diffusion coefficients of isolated E-cad and bound to β -cat, respectively. Using the above relationships, we found the amplitude to be 2.9 ± 0.1 and 4.3 ± 0.1 (in $k_B T$ for $T = 23^\circ\text{C}$) for randomly and periodically fluctuating ruggedness, respectively. These numbers suggest that the individual E-cad segments diffuse on a frustrated energy landscape with many local barriers of height $3 - 4 k_B T$.

Conclusion.

Whereas disordered complexes, sometimes being referred to as ‘fuzzy’, have been identified in the past⁵⁴, their dynamics were often only accessible in simulations⁵⁵. Our measurements of such dynamics in the E-cad/ β -cat complex uncovered a rugged energy landscape of E-cad with many shallow minima. The minima depth, i.e., the total strength of all contacts is in the order of a hydrogen bond⁵⁶, which is insufficient to rigidify E-cad. Nevertheless, the flexible parts of E-cad serve an important function: they boost affinity. This is necessary because specific contacts of the core-binding region fail to generate sufficient binding strength as indicated by the low affinity of the core-binding region of E-cad (Fig. 3b). Yet, the correct register of this element in the binding groove of β -cat provides specificity, which sets this complex apart from fully disordered complexes^{5,6}.

What might be the functional role of this unusual design? In-cell concentrations of β -cat are very low ($< 1 \text{ nM}$)^{57,58}, even in adherent cancer cell lines where β -cat is overexpressed. Given the nano-molar affinity of the complex, adherens junctions would be unstable. Post-translational modifications of E-cad are known to increase the stability of the E-cad/ β -cat complex⁵⁹. However, allowing kinases to access their substrate by dissociating the entire complex requires hours for re-associating it with the known rates ($10^5 - 10^6 \text{ M}^{-1}\text{s}^{-1}$)³⁷. The continuous attachment and detachment of local E-cad segments bypasses this problem by providing easy access for regulatory enzymes without dissociating and re-associating the complex (Fig. 6d). In addition, the complex can also serve mechanical functions. E-cad and β -cat link the cytoskeletons of neighbouring epithelial cells (Fig. 1a), suggesting that rupture forces can propagate from cell to cell. Like all polymers, IDPs are soft entropic springs. Many weak contacts to the β -cat surface can increase the spring constant without causing a rigid complex (see Methods). The E-cad/ β -cat complex has therefore the potential to act as a tuneable damper for mechanical forces between cells.

The Velcro-like design of many weak contacts on top of a few persistent interactions reconciles three seemingly contradictory factors: specificity, high affinity, and flexibility. Given the simplicity and advantages of this design together with the fact that partially disordered complexes are highly abundant⁵⁴, the mode of interaction described here may be common in biology.

Methods.

Constructs. The plasmid containing the cytoplasmic domain of mouse E-cad (residues 734-884, UniProt ID P09803) was a kind gift of William I. Weis from Stanford University School of Medicine. E-cad was transferred from the pGEX vector to a K151 vector system using transfer-PCR. The procedure was carried out as described previously⁶⁰. Additionally, cysteine linkers were incorporated in the E-cad-K151 vector using the restriction free cloning approach⁶¹ at different positions to generate the six labelling variants (Table S1). Full-length murine-beta-catenin (UniProt ID Q02284, purchased from GenScript) was transferred to a K151 vector system using transfer-PCR. Both proteins start with an N-terminal His₁₂-tag followed by the Sumo protein.

Protein expression. Proteins were expressed in *E. coli* BL21(DE3) cells. A 50 ml overnight culture grown at 37 °C (100 µg ml⁻¹ Kanamycin) was used to inoculate 4 l of super broth (100 µg ml⁻¹ Kanamycin). The cultures were grown at 37 °C to an optical density of 0.6 - 0.8 cm⁻¹ (at 600 nm). Then 0.5 mM isopropyl-β-D-thiogalactoside was added to induce protein expression for 3 h. After induction, the cultures were cooled to 4 °C and centrifuged at 3000 g for 30 min at 4 °C. The cell pellet was washed with 80 ml of ice-cold buffer (10 mM Tris-HCl pH 8.0, 300 mM NaCl) and centrifuged at 3000 g for 30 min at 4 °C. The supernatants were discarded and the pellets were stored at -80 °C.

Protein purification. The cell pellet from 2 l culture was re-suspended in 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 % glycerol, 1 mM DTT, 0.5 mM phenylmethanesulfonyl fluoride (PMSF), 5 µM leupeptin, 2.5 µg ml⁻¹ pepstatin (25 ml per 5 g wet weight) followed by sonication on ice using an ultrasonicator (Vibra-Cell, Sonics) at 70 % amplitude for 5 times 30 s with a 2 s on / 8 s off pulse. The soluble fraction was collected by centrifugation at 40,000 g for 30 min at 4 °C. Following sonication, DNA impurities were digested (DNase: Fisher Optizyme, 2 U/ml) and the supernatant was filtered (0.45 µm, Sartorius) and loaded at 4 °C with a flow rate of 0.5 ml min⁻¹ onto a 5 ml HisTrap HP column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 % glycerol. The column was subsequently washed with 20 ml of the equilibration buffer followed by 20 ml of 50 mM Tris-HCl pH 8.0, 500 mM NaCl and with additional 20 ml of 50 mM Tris-HCl pH 8.0, 470 mM NaCl, 75 mM imidazole at a flow rate of 1 ml min⁻¹. The protein was eluted with 20 ml of 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 500 mM imidazole at a flow rate of 1 ml min⁻¹. Protein fractions were pooled.

Removal of the His₆-Sumo tag. The affinity tags of both proteins were removed using Sumo protease, bdSENP1⁶². The plasmid was a kind gift of Professor Dirk Görlich, Max-Planck-Institute for Biophysical Chemistry (Göttingen, Germany). A

solution containing 300 nM Sumo protease was dialyzed twice against 50 mM Tris-HCl pH 8.0, 500 mM NaCl for 2h at 4°C. Afterwards, the sample was loaded at 2 ml min⁻¹ onto a 5 ml HisTrap HP column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 8.0, 470 mM NaCl, 75 mM imidazole at 4°C. The column was washed with 20 ml of the same buffer. The flow-through containing the cleaved protein was collected. The protease and non-cleaved proteins were eluted with 20 ml of buffer B at a flow rate of 2 ml min⁻¹. Protein fractions were pooled and submitted to size exclusion chromatography (SEC) using a Superdex 200 pg HiLoad 26/600 column (GE Healthcare Life Sciences) equilibrated with 20 mM Tris/HCl, 300 mM NaCl, 5 mM DTT, pH 8.0. Protein fractions were collected and pooled. Final refinement of the β -cat was conducted with a buffer exchange using a 5 ml GE, HiTrap desalting column into 25 mM Tris/HCl, 50 mM KCl, 200 mM L-Arg, pH 8, followed by concentrating the sample (VivaSpin 6 MWCO 3 kDa cut-off, GE Healthcare) to 8 μ M. The E-cad constructs were concentrated to ~ 14 μ M. Protein concentrations were determined based on the absorbance at 280 nm using the extinction coefficient at 280 nm of 18910 M⁻¹cm⁻¹ for the E-cad variants and 63830 M⁻¹cm⁻¹ for β -cat. All E-cad constructs were additionally purified using reverse-phase HPLC purification with the ZORBAX Eclipse Plus C18 (3.5 μ m) column (Agilent) (see following section). The correct protein mass was confirmed via ESI mass spectrometry analysis. The proteins were drop-frozen in liquid nitrogen and stored at -80 °C.

Fluorophore labeling with AlexaFluor 488 (donor) and AlexaFluor 594 (acceptor). The six E-cad variants were reduced before labelling by the addition of 100 mM DTT. After 15 min at room temperature, the proteins were purified via reversed phase HPLC using a ZORBAX Eclipse Plus C18 (3.5 μ m) column (Agilent) equilibrated with 0.1% trifluoroacetic acid (TFA). The E-cad constructs were eluted with gradients from 15% - 60% acetonitrile (ACN) in 13 ml. Purified proteins were lyophilized and stored at -80 °C. The lyophilized proteins were re-solubilized in 50 mM sodium phosphate pH 7.3, 6 M GdmCl at a protein concentration of 150 μ M. The precise concentrations were determined using the absorbance at 280 nm. For labelling, the proteins were incubated with 0.6 equivalents of Alexa 488 C5 maleimide (Invitrogen) for 60 min (25 °C, 300 rpm). Afterwards, a tenfold excess of β -mercaptoethanol (compared to the dye) was added to quench unreacted dye. Reversed-phase HPLC was then used to remove unreacted dye, unlabelled and double-donor labeled protein. To this end, a ZORBAX Eclipse Plus C18 (3.5 μ m) column (Agilent) equilibrated with water containing 0.1 % TFA was used. The gradient was 38% - 46% ACN in 15 ml. The pooled fractions were lyophilized and stored at -80 °C. The donor-labeled protein was re-solubilized in the labelling buffer to obtain a protein concentration of approximately 20 μ M and the correct protein concentration was determined using the absorbance at 493 nm and the extinction coefficient of the Alexa 488 dye (72000 M⁻¹cm⁻¹). A 5-fold excess of acceptor dye Alexa 594 C5 maleimide

(Invitrogen) was added and samples were incubated for 180 min at (25°C, 300 rpm). Afterwards, the reaction was quenched using β -mercaptoethanol. Again, reversed-phase HPLC was used to remove free label using the same column as described above. The doubly labeled proteins were eluted using with a 25% - 40% ACN gradient. The pooled fractions were lyophilized and stored at -80 °C. The correct labelling of all protein variants was confirmed using ESI mass spectrometry.

Single-molecule fluorescence spectroscopy. All single-molecule fluorescence experiments were performed with a MicroTime 200 confocal microscope (PicoQuant) equipped with an Olympus IX73 inverted microscope. We either used linearly polarized light from a 485 nm diode laser (LDH-D-C-485, PicoQuant) adjusted to 100 μ W (measured at the back aperture of the objective) to excite the donor fluorophore with a repetition rate of 40 MHz or two pulsed lasers controlled by a PDL 828-L "Sepia II" (PicoQuant, Germany) for Pulsed Interleaved Excitation (PIE) experiments^{63, 64}. In PIE experiments, the interleaved excitation of the acceptor dye allows us to identify molecules carrying both donor and acceptor dyes. To this end, two light sources, a 485 nm diode laser (LDH-D-C-485, PicoQuant) and a white-light laser (Solea, PicoQuant) set to an excitation wavelength of 595 nm, were used to excite the donor and the acceptor dyes alternately at a total repetition rate per period of 20 MHz. The laser intensities were adjusted to 100 μ W at 485 nm and 20 μ W at 595 nm (Pm100D, Thor Labs). The excitation beam was guided through a major dichroic mirror (ZT 470-491/594 rpc, Chroma) to a 60x, 1.2NA water objective (Olympus) that focuses the beam into the sample. The sample was placed in a homemade cuvette (quartz 25 mm diameter round cover slips, Esco Optics), borosilicate glass 6 mm diameter cloning cylinder (Hilgenberg)) with a volume of 50 μ l. Photons emitted from the sample were collected by the same objective and after passing the major dichroic mirror (ZT 470-491/594 rpc, Chroma), the residual excitation light was filtered by a long-pass filter (BLP01-488R, Semrock) and sent through a 100 μ m pinhole to remove out-of-focus light. The sample fluorescence was detected with either two channels or four channels for fluorescence anisotropy measurements. Donor and acceptor fluorescence was separated via a dichroic mirror (T585 LPXR, Chroma) and each colour was focused onto a single-photon avalanche diode (SPAD) (Excelitas) with additional bandpass filters: FF03-525/50, (Semrock) for the donor SPAD and FF02-650/100 (Semrock) for the acceptor SPAD. For fluorescence anisotropy measurements, the emission light was first separated into its parallel and perpendicular components with respect to the linearly polarized excitation light via a polarizing beam splitter and each component was separated by two dichroic mirrors into donor and acceptor photons resulting. The arrival time of every detected photon was recorded with a HydraHarp 400 time-correlated single photon counting (TCSPC) module (PicoQuant) and stored with a resolution of 8 ps (16 ps for PIE). The labelled proteins were diluted to a concentration of approximately 50 pM in 20 mM Tris-HCl pH 8.0 (12 mM ionic

strength) at the appropriate concentrations of KCl in the absence or presence of unlabelled β -cat. To prevent surface adhesion of the proteins and to maximize photon emission, 0.001% Tween 20 (Pierce) and 20 mM DTT were included in the buffer. Measurements in the presence of β -cat contained in addition 20 mM L-Arg (for a total of 32 mM ionic strength together with the Tris-HCl buffer) to prevent aggregation. All measurements were performed at 23°C.

Single-molecule identification and FRET corrections. Instrumental imperfections and differences in the brightness of donor and acceptor require a correction of the detected raw photon counts^{65, 66}. Although many of the experiments were performed with four detection channels, two donor channels and two acceptor channels with different polarizations, we describe the molecule identification and corrections for a two-channel (channel 1 and 2) setup. However, the extension to four channels is straightforward. These corrections are defined by five parameters: γ_1 and γ_2 that account for the different detection probabilities of photons from the two dyes, β_{21} and β_{12} , the leakage of donor photons into the acceptor channel 1 and the leakage of acceptor photons into the donor channel 2, respectively, and α , the probability to directly excite the acceptor dye at the wavelength specific for the donor. If n_1 and n_2 are the detected photons in the acceptor and donor channels, respectively, and b_1 and b_2 are the background rates in both channels, the corrected photon counts for acceptor and donor (n'_A and n'_D) are given by

$$\begin{pmatrix} n_A \\ n'_D \end{pmatrix} = \begin{pmatrix} \gamma_1 & -\beta_{21} \\ -\beta_{12} & \gamma_2 \end{pmatrix} \begin{pmatrix} n_1 - b_1 T \\ n_2 - b_2 T \end{pmatrix} \text{ and } n'_A = n_A - \alpha(n'_D + n_A) \quad (1)$$

Here, T is a time specifying the length of a burst. The correction parameters were determined with two separate samples of the dyes in which their concentrations were adjusted such that both samples have an absorbance of 0.1 at the excitation wavelength (485 nm)⁶⁷. Setting arbitrarily $\gamma_1 = 1$, we obtain $\gamma_2 = 1.12 \pm 0.09$, $\beta_{21} = 0.050 \pm 0.003$, and $\beta_{12} = 0.0021 \pm 0.0004$ over 5 years with 21 measurements of these correction factors. The probability of directly exciting the acceptor dye at 485 nm is given by $\alpha = \epsilon_A/(\epsilon_A + \epsilon_D)$ where ϵ_A and ϵ_D are the extinction coefficients of the dyes at the donor excitation wavelength of 485 nm. For our dye pair, $\alpha = 0.05$. Transfer efficiency histograms were computed with the fully corrected photon counts according to

$$E = \frac{n'_A}{n'_A + n'_D} \quad (2)$$

Importantly, the corrections (eq. 1) are already taken into account during burst identification. Bursts were identified from the measured photon traces following Eggeling *et al.*⁶⁵ and Hoffmann *et al.*⁶⁶. If $\Delta t_i = t_i - t_{i-1}$ is the inter-photon time

between the i^{th} photon and its predecessor, the photon i is retained if $\Delta t_i \leq \gamma_j(i) \Delta t_{\text{max}}$ ($\Delta t_{\text{max}} = 100 \mu\text{s}$) with $\gamma_j(i)$ being the correction factor of the i^{th} photon detected in channel $j = [1, 2]$ (see eq. 1). The algorithm then proceeds to the next photon $i + 1$, stops after n photons once $\Delta t_{i+n} > \gamma_j(i+n) \Delta t_{\text{max}}$, and provides the total length of the burst by $T = t_{n-1} - t_{i-1}$. The resulting string of photons is now corrected via eq. 1 using estimated background rates b_1 and b_2 . The initial guess of b_1 and b_2 is given by all detected photons in channel 1 and 2, respectively, divided by the total measurement time. A burst is then identified if $(n'_A + n'_D) > 100$. The photons belonging to this burst are removed from the photon trace and a new guess for b_1 and b_2 is computed based on the remaining photons. Subsequently, the burst search is performed again with updated background rates. This procedure converges after three iterations to constant background rates and a constant number of identified bursts.

Since the identified bursts also contain molecules for which the acceptor bleached during the transit through the confocal spot, thus masking the true transfer efficiency, we further cleaned the FRET histograms from these events⁶⁸. For a burst with n'_D donor photons with the arrival times $t_{D,1} \dots t_{D,n'_D}$ and n'_A acceptor photons with the arrival times $t_{A,1} \dots t_{A,n'_A}$, the average arrival times are given by $\langle t_D \rangle = n'^{-1}_D \sum_i t_{D,i}$ and $\langle t_A \rangle = n'^{-1}_A \sum_i t_{A,i}$. The burst asymmetry is defined by $\alpha_{DA} = \langle t_D \rangle - \langle t_A \rangle$. If the acceptor dye bleaches, we clearly find $\alpha_{DA} > 0$. Taking shot noise into account, the distribution of α_{DA} has a standard deviation given by

$$\sigma_{DA} = \frac{T}{2\sqrt{3}} \left(\frac{1}{n'_D} + \frac{1}{n'_A} \right)^{1/2}. \quad (3)$$

To eliminate molecules with a bleached acceptor, we excluded all molecules for which $|\alpha_{DA}| > \sigma_{DA}$. In addition, only molecules containing active acceptor and donor dyes were included in the analysis. To this end, we computed the donor-acceptor stoichiometry (S) for each burst according to

$$S = \frac{n'_{DD} + n'_{DA}}{n'_{DD} + n'_{DA} + \gamma(n'_{AD} + n'_{AA})}. \quad (4)$$

Here, γ is a correction factor to account for the different excitation intensities for donor and acceptor. Furthermore, the first subscript indicates the emission and the second subscript indicates the excitation. Only molecules with $S < 0.8$ were used for constructing smFRET histograms.

FRET histograms were fitted with a combination of empirical log-normal and Gaussian distributions⁶⁹. For binding experiments with unlabelled β -cat, the width and position of the FRET-peaks were fixed to minimize the number of free fitting parameters and the area under the fitted histogram curve for each subpopulation was determined using numerical integration. For distance calculations based on the mean

transfer efficiencies, the Förster radius measured in water $R_0^6(x)$ was corrected for the different refractive indices of the solutions according to:

$$R_0^6(x) = R_{0,0}^6 \left(\frac{n_0}{n(x)} \right)^4 \quad (5)$$

where $n(x)$ is the refractive index of the sample at condition x . Refractive indices were measured with an Abbe refractometer (Krüss) and were used to calculate the exact KCl concentrations^{70, 71}. The Foerster distance in water is 5.4 nm⁷².

Two-focus FCS (2fFCS). 2fFCS measurements were conducted on a MicroTime200 (PicoQuant, Berlin) confocal microscope (Olympus). The light of two orthogonally polarized, pulsed diode lasers is combined with a polarization sensitive beam splitter and afterwards coupled into a polarization maintaining single-mode fibre. Both lasers (483 nm, LDH-D-C-485, PicoQuant) are pulsed alternately with a total repetition rate of 40 MHz and a laser power of 30 μ W each. Before entering the objective, the laser beam passes through a Nomarski prism (U-DICTHC, Olympus). Afterwards, the two laser beams are focused by the objective thus causing two overlapping excitation volumes with a small lateral shift. Emission light passes through the objective, prism, and the dichroic beamsplitter (see section Single-molecule fluorescence spectroscopy) and is focused onto a pinhole (150 μ m). Behind the pinhole the light is collimated and divided by a polarizing beam-splitter and focused onto two single-photon avalanche diodes (SPADs). For each laser focus, the corresponding autocorrelation function and the cross-correlation functions between foci were computed and fitted as described previously⁷³. To determine precise diffusion coefficients, we determined the distance between the two foci using the known Stokes radius of the dye Oregon Green and the water viscosity at the known lab-temperature of $23 \pm 0.6^\circ\text{C}$. Fitting of the auto- and cross-correlation functions is performed as described by Dertinger *et al.*⁷³. We found a focal distance of 353 nm.

Determination of the affinity of the core-binding region of E-cad for β -cat.

The core-binding region of E-cad with the highest contact probabilities (Fig. 5b) consisted of a 20 amino acid stretch. Since smFRET is not suitable for monitoring conformational changes in such a short peptide, we used 2fFCS to determine the Stokes radius of the peptide at increasing concentrations of β -cat. To minimize that dye-labelling with AlexaFluor 488 massively impacted binding, we used two peptides in which the dye was either placed at the N-terminus or at the C-terminus (Supplementary Table 1). In addition, we introduced one serine-glycine repeat between the peptide and the cysteine used for labelling to increase the distance between the dye and the binding competent peptide. The free peptides had an average (average over both labelling constructs) Stokes radius of 1.5 ± 0.2 nm. At the highest β -cat concentration

measured in our experiment, the Stokes radius was found to be 1.6 ± 0.2 nm, i.e., unchanged within the error. For comparison, the expected Stokes radius⁷⁴ of β -cat with a molecular weight of 85 kDa is 3.5 nm, which agrees with the Stokes radius of 3.9 nm found for donor-labelled E-cad in complex with β -cat (1 μ M). The increased value in presence of E-cad is likely due to the flexible parts of E-cad that render the complex slightly larger than β -cat alone.

Temperature-controlled smFRET. Temperature-controlled smFRET experiments were conducted in a custom-built system. The design is very similar to a previous version^{75, 76} and includes a temperature controlled sample holder and a second device to control the temperature of the objective. To precisely determine the temperature inside the confocal spot, we measured the diffusion coefficient of the dye Oregon Green (ThermoFisher Scientific) in water using 2f-FCS⁷³ (Supplementary Fig. 7). The average of the fits from three independent measurements at each temperature in the range of 0-70 °C were used to determine the viscosity of water given the hydrodynamic radius of Oregon Green (0.6 nm)⁷⁷. The Stokes equation gives a relationship between the hydrodynamic radius R_S , the diffusion coefficient D , and the viscosity of the medium η :

$$\eta = \frac{k_B T}{6\pi D R_S} \quad (6)$$

where k_B is the Boltzmann constant and T the temperature. On the other hand, the water viscosity η has a known temperature dependence⁷⁸ given by

$$\eta = \eta_0 e^{ap + \frac{H-bp}{R(T-\theta-cp)}} \quad (7)$$

with p being the pressure in bar, T the temperature in K, and η_0, a, b, c, H and θ being constants ($E = 4.753 \text{ kJ mol}^{-1}$, $\eta_0 = 2.4055 \times 10^{-5} \text{ Pa s}$, $\theta = 139.7 \text{ K}$, $a = 4.42 \times 10^{-4} \text{ bar}^{-1}$, $b = 9.565 \times 10^{-4} \text{ kJ mol}^{-1} \text{ bar}^{-1}$, and $c = 1.24 \times 10^{-2} \text{ K bar}^{-1}$). When combining eq. 6 and 7, the temperature satisfies an implicit equation that relates the measured diffusion coefficient of Oregon Green to the temperature of the sample, which was used to compute the actual temperature in the confocal spot (Supplementary Fig. 7).

Determination of average donor-acceptor distances from FRET efficiencies. The mean FRET efficiency of all E-cad variants in the absence of β -cat were converted to distances according to⁷²

$$\langle E \rangle = \frac{\int_0^L E(r) P(r) dr}{\int_0^L P(r) dr} . \quad (8)$$

Here, $E(r)$ is the well-known Foerster equation

$$E(r) = \frac{R_0^6}{R_0^6 + r^6} \quad (9)$$

with $R_0 = 5.4 \text{ nm}$. The conversion of mean transfer efficiencies to donor-acceptor distances (R_{DA}) is clearly model-dependent and we determined R_{DA} using eq. 8 with four polymer models: the Gaussian chain model⁷², the Worm-like chain model⁷⁹, the Sanchez model^{25, 80, 81, 82}, and the self-avoiding random walk model (SAW)^{79, 83}. The resulting R_{DA} -distances are all very similar. The model-dependent variation is shown as an error band in Fig. 2d.

Two-dimensional fluorescence lifetime vs. transfer efficiency plots. To identify whether the observed broadening of the FRET histograms arises from conformational heterogeneity, we computed two-dimensional plots that correlate the donor excited-state lifetime with the measured FRET efficiency⁶⁶. For a fixed distance r , the mean donor lifetime in the presence of acceptor is given by

$$\frac{\tau_{DA}(r)}{\tau_D} = 1 - E(r) \quad (10)$$

with τ_D being the lifetime of the donor in absence of the acceptor. Yet, in presence of a distribution of distances $P(r)$, the relationship between mean fluorescence lifetime and mean FRET efficiency changes. The mean fluorescence lifetime of the donor in the presence of an acceptor is then given by

$$\langle \tau_{DA} \rangle = \frac{\int_0^\infty t I(t) dt}{\int_0^\infty I(t) dt} \text{ with } I(t) = I_0 \int_0^\infty P(r) e^{-t/\tau_{DA}(r)} dr. \quad (11)$$

Here, $I(t)$ is the time-dependent fluorescence emission intensity. Simplifying equation (11), the average fluorescence lifetime is given by

$$\langle \tau_{DA} \rangle = \frac{\int_0^\infty \tau_{DA}(r)^2 P(r) dr}{\int_0^\infty \tau_{DA}(r) P(r) dr}. \quad (12)$$

In Fig. 4b, we used the distance distribution for a Gaussian chain given by

$$P(r) = 4\pi r^2 \left(\frac{3}{2\pi \langle r^2 \rangle} \right)^{\frac{3}{2}} e^{-\frac{3r^2}{2\langle r^2 \rangle}} \quad (13)$$

with $R_{DA} = \langle r^2 \rangle^{\frac{1}{2}}$ being the average donor-acceptor distance to demonstrate the expected lifetime-FRET correlation for a fully disordered chain.

Polyampholyte theory. To describe the chain collapse with increasing concentrations of KCl, we fitted the change in donor-acceptor distance of all E-cad variants with a polyampholyte polymer theory developed by Higgs and Joanny²⁹

$$\alpha^5 - \alpha^3 = \frac{4}{3} \left(\frac{3}{2\pi} \right)^{3/2} \sqrt{N} \nu \quad \text{with} \quad \alpha^2 = \frac{\langle r^2 \rangle}{R_{ideal}^2} \quad (14)$$

Here, $R_{ideal}^2 = Nb^2$ with N being the number of peptide bonds and $b = 0.38 \text{ nm}$, being the distance between two C_α -atoms. Expression (14) measures the deviation of the end-to-end distance with respect to the end-to-end distance of an ideal Gaussian chain. The two-body interaction term (excluded volume) ν has two contributions $\nu = \nu_0 + \nu_{el}$, non-electrostatic interactions (ν_0) and electrostatic interactions (ν_{el}). The latter is given by Higgs and Joanny as

$$\nu_{el} = \frac{4\pi l_B (f-g)^2}{b^3 \kappa^2} - \frac{\pi l_B^2 (f+g)^2}{b^3 \kappa}. \quad (15)$$

Here, f and g are the fractions of positive and negative charges in the sequence of the individual E-cad variants, respectively. The Debye length is defined as $\kappa^{-1} = (8\pi l_B I)^{-\frac{1}{2}}$, where I is the ionic strength of the solution and l_B is the Bjerrum length, i.e., the distance at which the electrostatic energy of the interaction between two elementary charges equals thermal energy. We use the known relation $l_B = \frac{e^2}{4\pi\epsilon_0\epsilon_r k_B T}$, where e is the elementary charge, ϵ_0 is the permittivity of vacuum, ϵ_r is the dielectric constant, k_B is Boltzmann's constant, and T is the temperature. Hence, eq. 15 is fully determined by the protein sequence and the ionic strength of the medium. To describe the collapse at high concentrations of KCl, we assume a linear dependence between the non-electrostatic two-body interaction term $\nu_0 = c_1 + c_2[I]$ with I being the ionic strength of the solution and c_1 and c_2 are fitting constants³⁰.

Sequence charge decoration (SCD). We have used a variant of the sequence charge decoration^{34, 35}, $SCD_{\text{low salt}}$,

$$SCD_{\text{low salt}} = \frac{1}{N} \sum_{i=2}^N \sum_{j=1}^{i-1} q_i q_j (i-j) \quad (16)$$

in which q_i and q_j are the charges of the i -th and j -th amino acids in the sequence. This metric is designed to predict the salt-induced conformational change of the intrinsically disordered proteins due to charge patterning. The chain will compact with addition of salt if $SCD_{\text{low salt}}$ is positive, whereas increasing salt concentration will induce chain expansion if $SCD_{\text{low salt}}$ is negative.

Nanosecond Fluorescence Correlation Spectroscopy (nsFCS). We compute subpopulation specific correlation functions. To this end, we are using sample concentrations of around 100-500 pM that still allow us to differentiate populations of molecules that exhibit different FRET efficiencies. In a first step, we identify the photon bursts from individual molecules as described above. For these bursts, the FRET efficiencies are determined and finally the correlation functions for the selected subset of molecules were computed. By distributing the photons onto two donor and two acceptor detectors, dead time and after-pulsing of the SPADs are avoided. The cross-correlation functions were fitted using

$$g(t) = N^{-1} \left[1 - c_{ab} e^{\left(\frac{t-t_0}{\tau_{ab}} \right)} \right] \left[1 - c_c e^{\left(\frac{t-t_0}{\tau_c} \right)} \right] \left[1 + c_T e^{\left(\frac{t-t_0}{\tau_T} \right)} \right]. \quad (17)$$

The three terms in brackets describe photon antibunching (*ab*), conformational dynamics (*c*), and triplet blinking of the dyes (*T*). In addition, N is the effective number of molecules in the confocal volume and c_{ab} , c_c , and c_T are amplitudes. The correlation times τ_{ab} , τ_c , τ_T and the time origin t_0 are fitting parameters. To graphically compare different experiments, the correlation functions in Fig. 4a were normalised by N and the triplet term.

Determining the intra-chain diffusion coefficient of E-cad ABC from nsFCS.

We determined the intra-chain diffusion coefficient of free E-cad ABC by modelling its dynamics as a diffusive process in the potential of mean force given by the donor-acceptor distance distribution that resulted as a best fit from the CG-simulations (Fig. 6c). In this picture, the measured correlation time τ_c obeys a direct relationship to the intra-chain diffusion coefficient D given by⁸⁴

$$\tau_c = D^{-1} \int_0^L P(r)^{-1} \left[\int_0^r \delta n(\rho) P(\rho) d\rho \right] / \int_0^L \delta n(r)^2 P(r) dr \quad (18)$$

with

$$\delta n(r) = n(r) - \langle n \rangle, \quad (19)$$

where $P(r)$ is the donor-acceptor distance distribution determined from the CG-simulation and $n(r)$ is the donor photon rate. The latter is determined from the photophysics of the FRET system with donor (D) and acceptor (A) that includes the four states DA, D*A, D*A*, DA* where the asterisk indicates an excited state. The rate matrix for this system is given by (in the basis given above)

$$\mathbf{K} = \begin{pmatrix} -(1 + \alpha)k_{ex} & k_D & 0 & k_A \\ k_{ex} & -(k_D + \alpha k_{ex} + k_T(r)) & k_A & 0 \\ 0 & \alpha k_{ex} & -(k_D + k_A) & k_{ex} \\ \alpha k_{ex} & k_T(r) & k_D & -(k_{ex} + k_A) \end{pmatrix} \quad (20)$$

where $k_{ex} = 0.018 \text{ ns}^{-1}$ is the donor excitation rate, $k_D = 0.28 \text{ ns}^{-1}$ is the donor emission rate, $k_A = 0.25 \text{ ns}^{-1}$ is the acceptor emission rate, and $\alpha = 0.05$ is the direct excitation probability. The energy transfer rate is given by $k_T(r) = k_D(R_0/r)^6$. The steady-state vector of the four states $\mathbf{p}_{ss}$ is the solution of the algebraic system $\mathbf{0} = \mathbf{K}\mathbf{p}_{ss}$ where $\mathbf{0}$ is the null vector. The photon rate $n(r)$ of the system is then given by $n(r) = \mathbf{1}^T \mathbf{V} \mathbf{p}_{ss}$ where

$$\mathbf{V} = k_D \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (21)$$

is the donor photon detection matrix. We determined the intra-chain diffusion coefficient D of free E-cad ABC (Fig. 6c) by numerically solving eq. 18.

Recurrence analysis of single particles (RASP). The recurrence analysis of single particles (RASP) uses the fact that a freely diffusing molecule can be observed multiple times in a single-molecule experiment. Once a molecule leaves the observation volume, the chance of it returning to the confocal spot is greater than the chance of detecting a new molecule for short time intervals (Supplementary Fig. 4a-b). To extract kinetics based on this effect, we binned our data in 100 μs bins and computed the bin-time autocorrelation function $G(\tau)$ that contains the information of the likelihood to observe two molecules separated by the delay time τ

$$G(\tau) = \frac{p(\{b_1, t_1\}, \{b_2, t_1 + \tau\})}{p(\{b_1, t_1\})p(\{b_2, t_1\})} \quad (22)$$

Here, $p(\{b_1, t_1\}, \{b_2, t_1 + \tau\})$ denotes the joint probability of observing two bins b_1 and b_2 at times t_1 and $t_1 + \tau$, respectively, and $p(\{b_1, t_1\})$ and $p(\{b_2, t_1\})$ are the probabilities of detecting b_1 and b_2 , respectively, at time t_1 . This correlation function can be converted to $P_{same}(\tau)$, i.e., the probability that bins separated by the time τ originate from the same molecule⁴² (Supplementary Fig. 4b):

$$P_{same}(\tau) = 1 - \frac{1}{G(\tau)}. \quad (23)$$

During the time between the first and the second detection, the molecule can change its conformation, thus causing a different FRET efficiency in the second detection. To

quantify the conformational dynamics with RASP, we define two regions of the broad FRET histograms of the E-cad/ β -cat complex, region 1 and region 2 (Supplementary Fig. 4c). Bins originating from the FRET efficiency chosen in region 1 are selected (E_0 -bins) and recurrence histograms are calculated from the photons in bins (E_1 -bins) at a time delay τ after the initial bins (Supplementary Fig. 4d). The recurrence FRET-histograms were globally fitted with a sum of a log-normal function for molecules that lack an active acceptor dye ($E = -0.05$) and two Gaussian functions for the populations from regions 1 and 2. The positions, widths, and asymmetries were determined from the histograms of both regions at the initial time window, i.e., at $\tau = 0$. The relative populations of the molecules in region 1 and region 2 were obtained by integrating the respective sub-populations. Supplementary Fig. 4e depicts the time course of the fraction of molecules in region 1 after originally selecting molecules in region 2 (red) or molecules in region 1 (blue), respectively. By increasing the time delay τ , the fraction of molecules in region 1 when initially selecting region 1 decrease and those in region 1 when selecting region 2 increase. The observed kinetics $P_m(t)$ is a combination of two contributions: conformational switching between the two selected regions, characterized by $P_{conf}(t)$, and the time-dependent likelihood that two observed bins or “snapshot” arise from the same molecule, $P_{same}(t)$. Since we can compute $P_{same}(t)$ directly from the burst time autocorrelation function of our data, we can determine $P_{conf}(t)$, the actual kinetics of conformational switching. Thus, the observed increase in the population of molecules in region 1 in the recurrence histograms after initially selecting molecules in region 2 is given by

$$P_m(t) = P_{same}(t)P_{conf}(t) + (1 - P_{same}(t))\rho_{eq}. \quad (24)$$

Here, $P_m(t)$ is the measured time decay, $P_{same}(t)$ is the probability that the same molecule recurs, $P_{conf}(t)$ is the kinetics of switching to region 1 due to conformational dynamics, $1 - P_{same}(t)$ is the probability that a new molecule enters the confocal volume, and ρ_{eq} is the probability that a new molecule is in region 1, which is given by the equilibrium probability that a molecule is in region 1. For a better visualization and easier analysis, we directly determine the ‘true’ kinetics of switching $P_{conf}(t)$ that we obtain by re-arranging equation (24) for $P_{conf}(t)$. In addition, we average the two kinetic traces, i.e., the increase of molecules in region 1 after initial selection of molecules from region 2 and the depletion of molecules in region 1 after initial selection of molecules from region 1. The resulting averaged traces are shown in Fig. 4c.

Fitting RASP kinetics by numerically solving the Smoluchowski equation.

As an alternative to exponential fitting, we also fitted the RASP kinetics of the variant E-cad ABC bound to β -cat by solving the Smoluchowski equation

$$\partial u(t, r) / \partial t = D \partial / \partial r [V'(r) + \partial / \partial r] u(t, r) \quad (25)$$

with $u(t, r)$ being the time-dependent donor-acceptor distance distribution and $V(r) = -\ln P(r)$ being the potential of mean force with $P(r)$ being the normalized equilibrium donor-acceptor distance distribution obtained as a best fit from the CG-model (Fig. 6b inset). Equation 25 was solved with the boundary conditions

$$u(0, r) = \begin{cases} 0 & 0 \leq r < r_b \\ Z(r_b)^{-1}P(r) & r_b \leq r \leq r_{max} \end{cases} \text{ with } u(t, 0) = 0 \text{ and } u(t, r_{max}) = 0. \quad (26)$$

Here, $Z(r_b)$ is a normalization factor, r_b defines the distance that separates high-FRET (region 2) ($r < r_b$) from low-FRET (region 1) ($r \geq r_b$) molecules, and r_{max} is the upper integration limit (20 nm). An analytical expression for $P(r)$ suitable to efficiently solve eq. 25 was obtained by fitting the distance distribution for E-cad ABC from the CG-model with the superposition of three Gaussian functions. The fraction of low-FRET molecules $f(t)$, i.e., molecules from region 1, was obtained from

$$f(t) = \int_{r_b}^{r_{max}} u(t, r) dr / \int_0^{r_{max}} u(t, r) dr. \quad (27)$$

The model contains two free parameters, the diffusion coefficient D and the separation distance r_b . The latter quantity was determined prior to fitting by numerically solving

$$\langle f \rangle = \int_{r_b}^{r_{max}} P(r) dr / \int_0^{r_{max}} P(r) dr \quad (28)$$

for r_b where $\langle f \rangle$ is the average fraction of the last 2 ms of the kinetic time traces, i.e., from 8 – 10 ms. The resulting value for r_b was then used for fitting. We performed an error estimation of the diffusion coefficient using bootstrapping the kinetic time traces with $n = 3$.

Molecular simulations with the CG-model. We parameterized a coarse-grained (CG) model to characterize the interactions between E-cad and β -cat. The model was based on our recently developed HPS model⁸⁵ in which each amino acid was represented as a bead. The potential energy can be written as,

$$E = E_{\text{bonded}} + E_{\text{electrostatic}} + E_{\text{IDP}}(\epsilon) + E_{\text{complex}}(\zeta), \quad (29)$$

in which there are three types of interactions for disordered E-cad: bonded, electrostatic and short-range pairwise interactions, and one additional term for the interactions between E-cad and β -cat. The bonded interactions were modelled by a harmonic potential with a bond length of 3.8 Å and a spring constant of 10 kJ/Å². The electrostatic interactions were modelled with a Coulombic term using Debye-Hückel electrostatic screening⁸⁶. The short-range pairwise interactions were modelled

using Ashbaugh-Hatch functional form⁸⁷ and the depths of the potential minimums are proportional to the amino acid hydrophobicity values (λ)⁸⁸. The free parameter ϵ , which was originally tuned according to the size of the IDPs, has been adjusted to 0.16 kcal/mol using the experimental FRET efficiencies of the six E-cad segments (Supplementary Fig. 8). To simulate β -cat in the CG model, we represented the structured part of β -cat in the experimentally solved structure (PDB: 1i7x)²³ as a rigid body in the simulation. The interactions between E-cad and β -cat were also modelled based on the native interactions in the same structure. A native interaction was assumed if the C α distance between two amino acids is smaller than 1.2 times the sum of the amino acid radii used in the CG model. For every pair of amino acids that form native interactions in the complex, we have introduced one rigid and one flexible CG-models, with different types of potential energy functions. First for the rigid model, a harmonic potential energy function with the native interaction distance as the minimum position and a large spring constant of 20 kcal/mol were introduced. This effectively rigidified all the native contacts as seen in both X-ray structures. Second for the flexible model, a 12-6 Leonard-Jones (LJ) potential energy function with the native interaction distance as the minimum position was introduced. The depth of the LJ-well, i.e., the interaction strength ξ was adjusted to 0.6 kcal/mol according to the FRET efficiencies of the six segments of E-cad in the complex at an ionic strength of 82 mM. The deviations to the experimental FRET efficiencies for both models were shown in Supplementary Fig. 8. The CG simulation of E-cad alone was run for 10 μ s and the complex simulation was run for 20 μ s, using LAMMPS⁸⁹ and HOOMD-Blue⁹⁰ to benefit from both CPU and GPU resources. The first 1 μ s of the simulation was always excluded from the analysis. To obtain the FRET efficiencies from the simulations, the linker of both dyes were taken to be equivalent 9 amino acids⁹¹, i.e., 4.5 amino acids per linker.

Rouse model with surface interactions. Here we show that for a simple bead-and-spring model, contacts with a surface increase the effective spring constant. In one dimension, the Langevin equation for the coordinate of the k^{th} bead ($x_k = x_k(t)$) that contacts its neighbour beads is well known^{46, 92}. All beads also contact a specific location on a surface ($y = 0$), which, for simplicity, is identical for all beads and does not fluctuate in time

$$\xi \dot{x}_k(t) + k_1(2x_k(t) - x_{k-1}(t) - x_{k+1}(t)) + k_2 x_k(t) = X_k(t) \quad (30)$$

Here, ξ is a friction coefficient and k_1 and k_2 are spring constants for bonds in the chain and with the surface, respectively. X_k is Gaussian distributed white noise. Assuming a circular chain⁴⁶, taking the Fourier transform in bead number ($k \rightarrow q$) and time ($t \rightarrow \omega$), and rearranging, we obtain

$$x''_q(\omega) = X''_q(\omega)/[-i\omega\xi + 2k_1(1 - \cos q) + k_2]. \quad (31)$$

The relaxation time for the largest modes of such a chain is then given by $\tau_q = \xi/k_{eff}$ with the effective spring constant $k_{eff} = k_1q^2 + k_2$, showing that the spring constant of a chain is affected by surface interactions.

Data availability.

Data supporting the findings of this study are available from the corresponding authors upon reasonable request. A reporting summary of this article is available as a Supplementary information file. Source data are provided with this paper.

Code availability.

A custom WSTP add-on for Mathematica (Wolfram Research) used for the analysis of single-molecule fluorescence data is available at <https://schuler.bioc.uzh.ch/programs/>. The implementation of the coarse-grained HPS model in HOOMD-Blue can be downloaded at https://bitbucket.org/jeetain/hoomd_slab_builder.

References.

1. Schuler B, *et al.* Binding without folding - the biomolecular function of disordered polyelectrolyte complexes. *Curr Opin Struct Biol* **60**, 66-76 (2020).
2. van der Lee R, *et al.* Classification of Intrinsically Disordered Regions and Proteins. *Chem Rev* **114**, 6589-6631 (2014).
3. Csizmok V, Follis AV, Kriwacki RW, Forman-Kay JD. Dynamic Protein Interaction Networks and New Structural Paradigms in Signaling. *Chem Rev* **116**, 6424-6462 (2016).
4. Levy Y, Onuchic JN, Wolynes PG. Fly-casting in protein-DNA binding: frustration between protein folding and electrostatics facilitates target recognition. *J Am Chem Soc* **129**, 738-739 (2007).
5. Borgia A, *et al.* Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61-66 (2018).
6. Holmstrom ED, Liu Z, Nettels D, Best RB, Schuler B. Disordered RNA chaperones can enhance nucleic acid folding via local charge screening. *Nat Comms* **10**, 2453-2411 (2019).
7. Mittag T, *et al.* Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc Natl Acad Sci USA* **105**, 17772-17777 (2008).
8. Hendus-Altenburger R, *et al.* The human Na⁺/H⁺ exchanger 1 is a membrane scaffold protein for extracellular signal-regulated kinase 2. *BMC Biol* **14**, 1-17 (2016).

9. Milles S, *et al.* Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* **163**, 734-745 (2015).
10. Nusse R, Clevers H. Wnt/ β -Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985-999 (2017).
11. McCreas PD, Turck CW, Gumbiner B. A homolog of the armadillo protein in *Drosophila* (plakoglobin) associated with E-cadherin. *Science* **254**, 1359-1361 (1991).
12. Kemler R. From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion. *Trends Genet* **9**, 317-321 (1993).
13. Behrens J, *et al.* Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature* **382**, 638-642 (1996).
14. Aberle H, Butz S, Stappert J, Weissig H, Kemler R, Hoschuetzky H. Assembly of the cadherin-catenin complex in vitro with recombinant proteins. *J Cell Sci* **107**, 3655-3663 (1994).
15. Ozawa M, Baribault H, Kemler R. The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. *EMBO J* **8**, 1711-1717 (1989).
16. Rimm DL, Koslov ER, Kebriaei P, Cianci CD, Morrow JS. Alpha 1(E)-catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex. *Proc Natl Acad Sci USA* **92**, 8813-8817 (1995).
17. Huber AH, Nelson WJ, Weis WI. Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* **90**, 871-882 (1997).
18. Behrens J, *et al.* Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature* **382**, 638-642 (1996).
19. Hoschuetzky H, Aberle H, Kemler R. Beta-catenin mediates the interaction of the cadherin-catenin complex with epidermal growth factor receptor. *J Cell Biol* **127**, 1375-1380 (1994).
20. Hülsken J, Birchmeier W, Behrens J. E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton. *J Cell Biol* **127**, 2061-2069 (1994).
21. Aberle H, Schwartz H, Hoschuetzky H, Kemler R. Single amino acid substitutions in proteins of the armadillo gene family abolish their binding to alpha-catenin. *J Biol Chem* **271**, 1520-1526 (1996).
22. MacDonald BT, Tamai K, He X. Wnt/ β -Catenin Signaling: Components, Mechanisms, and Diseases. *Dev Cell* **17**, 9-26 (2009).

23. Huber AH, Weis WI. The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell* **105**, 391-402 (2001).
24. DeForte S, Uversky VN. Resolving the ambiguity: Making sense of intrinsic disorder when PDB structures disagree. *Protein Sci* **25**, 676-688 (2016).
25. Hofmann H, Soranno A, Borgia A, Gast K, Nettels D, Schuler B. Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc Natl Acad Sci USA* **109**, 16155-16160 (2012).
26. Ha, Thirumalai. Conformations of a polyelectrolyte chain. *Phys Rev A* **46**, R3012-R3015 (1992).
27. Hofmann H, Golbik R, Ott M, Hübner C, Ulbrich-Hofmann R. Coulomb forces control the density of the collapsed unfolded state of barstar. *J Mol Biol* **376**, 597-605 (2008).
28. Müller-Späth S, *et al.* Charge Interactions can Dominate the Dimensions of Intrinsically Disordered Proteins. *Proc Natl Acad Sci USA* **107**, 14609-14614 (2010).
29. Higgs PG, Joanny J-F. Theory of polyampholyte solutions. *J Chem Phys* **94**, 1543-1554 (1991).
30. Vancraenenbroeck R, Harel YS, Zheng W, Hofmann H. Polymer effects modulate binding affinities in disordered proteins. *Proc Natl Acad Sci USA* **116**, 19506-19512 (2019).
31. Bhattacharjee A, Kundu P, Dua A. Self-consistent theory of structures and transitions in weak polyampholytes. *Macromol Theory Simul* **20**, 75-84 (2011).
32. Samanta HS, Chakraborty D, Thirumalai D. Charge fluctuation effects on the shape of flexible polyampholytes with applications to intrinsically disordered proteins. *J Chem Phys* **149**, 163323 (2018).
33. Das RK, Pappu RV. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc Natl Acad Sci U S A* **110**, 13392-13397 (2013).
34. Sawle L, Ghosh K. A theoretical method to compute sequence dependent configurational properties in charged polymers and proteins. *J Chem Phys* **143**, 085101 (2015).
35. Huihui J, Firman T, Ghosh K. Modulating charge patterning and ionic strength as a strategy to induce conformational changes in intrinsically disordered proteins. *J Chem Phys* **149**, 085101 (2018).
36. Gomes G-NW, *et al.* Conformational Ensembles of an Intrinsically Disordered Protein Consistent with NMR, SAXS, and Single-Molecule FRET. *J Am Chem Soc* **142**, (2020).

- 1028 37. Choi H-J, Huber AH, Weis WI. Thermodynamics of β -Catenin-Ligand Interactions
1029 THE ROLES OF THE N- AND C-TERMINAL TAILS IN MODULATING BINDING
1030 AFFINITY. *J Biol Chem* **281**, 1027-1038 (2006).
1031
- 1032 38. Nettels D, Gopich I, Hoffmann A, Schuler B. Ultrafast dynamics of protein collapse
1033 from single-molecule photon statistics. *Proc Natl Acad Sci USA* **104**, 2655-2660 (2007).
1034
- 1035 39. Schuler B, Soranno A, Hofmann H, Nettels D. Single-Molecule FRET Spectroscopy and
1036 the Polymer Physics of Unfolded and Intrinsically Disordered Proteins. *Annu Rev*
1037 *Biophys* **45**, 207-231 (2016).
1038
- 1039 40. Haenni D, Zosel F, Reymond L, Nettels D, Schuler B. Intramolecular Distances and
1040 Dynamics from the Combined Photon Statistics of Single-Molecule FRET and
1041 Photoinduced Electron Transfer. *J Phys Chem B* **117**, 13015-13028 (2013).
1042
- 1043 41. Hillger F, *et al.* Probing protein-chaperone interactions with single-molecule
1044 fluorescence spectroscopy. *Angew Chem Int Ed Engl* **47**, 6184-6188 (2008).
1045
- 1046 42. Hoffmann A, *et al.* Quantifying heterogeneity and conformational dynamics from single
1047 molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP).
1048 *Phys Chem Chem Phys* **13**, 1857-1871 (2011).
1049
- 1050 43. Grossman-Haham I, Rosenblum G, Namani T, Hofmann H. Slow domain
1051 reconfiguration causes power-law kinetics in a two-state enzyme. *Proc Natl Acad Sci*
1052 *USA* **115**, 513-518 (2018).
1053
- 1054 44. Lasitza-Male T, *et al.* Membrane Chemistry Tunes the Structure of a Peptide
1055 Transporter. *Angew Chem Int Ed Engl* **28**, 51 (2020).
1056
- 1057 45. Soranno A, *et al.* Quantifying internal friction in unfolded and intrinsically disordered
1058 proteins with single molecule spectroscopy. *Proc Natl Acad Sci USA* **109**, 17800-17806
1059 (2012).
1060
- 1061 46. Soranno A, Zosel F, Hofmann H. Internal friction in an intrinsically disordered protein-
1062 Comparing Rouse-like models with experiments. *J Chem Phys* **148**, 123326 (2018).
1063
- 1064 47. Rouse PE. A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling
1065 Polymers. *J Chem Phys* **21**, 1272-1280 (1953).
1066
- 1067 48. Stappert J, Kemler R. A short core region of E-cadherin is essential for catenin binding
1068 and is highly phosphorylated. *Cell Adhes Commun* **2**, 319-327 (1994).
1069
- 1070 49. Kramers H. Brownian motion in a field of force and the diffusion model of chemical
1071 reactions. *Physica* **7**, 284-304 (1940).
1072
- 1073 50. Zwanzig R. Diffusion in a rough potential. *Proc Natl Acad Sci USA* **85**, 2029-2030
1074 (1988).
1075

- 1076 51. Frauenfelder H, Sligar SG, Wolynes PG. The energy landscapes and motions of
1077 proteins. *Science* **254**, 1598-1603 (1991).
1078
- 1079 52. Ansari A, *et al.* Protein states and proteinquakes. *Proc Natl Acad Sci USA* **82**, 5000-
1080 5004 (1985).
1081
- 1082 53. Wensley B, *et al.* Experimental evidence for a frustrated energy landscape in a three-
1083 helix-bundle protein family. *Nature* **463**, 685-U122 (2010).
1084
- 1085 54. Fuxreiter M, Tompa P. *Fuzzy Complexes: A More Stochastic View of Protein*
1086 *Function*. Springer US (2012).
1087
- 1088 55. Bigman LS, Levy Y. Tubulin tails and their modifications regulate protein diffusion on
1089 microtubules. *Proc Natl Acad Sci USA* **117**, 8876-8883 (2020).
1090
- 1091 56. Sheu S-Y, Yang D-Y, Selzle HL, Schlag EW. Energetics of hydrogen bonds in peptides.
1092 *Proc Natl Acad Sci USA* **100**, 12683-12687 (2003).
1093
- 1094 57. Sievers S, Fritzsche C, Grzegorzczak M, Kuhnen C, Müller O. Absolute beta-catenin
1095 concentrations in Wnt pathway-stimulated and non-stimulated cells. *Biomarkers* **11**,
1096 270-278 (2006).
1097
- 1098 58. Theillet F-X, *et al.* Physicochemical properties of cells and their effects on intrinsically
1099 disordered proteins (IDPs). *Chem Rev* **114**, 6661-6714 (2014).
1100
- 1101 59. Choi H-J, Loveless T, Lynch AM, Bang I, Hardin J, Weis WI. A conserved
1102 phosphorylation switch controls the interaction between cadherin and β -catenin in vitro
1103 and in vivo. *Dev Cell* **33**, 82-93 (2015).
1104
- 1105 60. Erijman A, Dantes A, Bernheim R, Shifman JM, Peleg Y. Transfer-PCR (TPCR): a
1106 highway for DNA cloning and protein engineering. *J Struct Biol* **175**, 171-177 (2011).
1107
- 1108 61. Unger T, Jacobovitch Y, Dantes A, Bernheim R, Peleg Y. Applications of the
1109 Restriction Free (RF) cloning procedure for molecular manipulations and protein
1110 expression. *J Struct Biol* **172**, 34-44 (2010).
1111
- 1112 62. Frey S, Goerlich D. Purification of protein complexes of defined subunit stoichiometry
1113 using a set of orthogonal, tag-cleaving proteases. *J Chromatogr A* **1337**, 106-115
1114 (2014).
1115
- 1116 63. Müller BK, Zaychikov E, Bräuchle C, Lamb DC. Pulsed interleaved excitation. *Biophys*
1117 *J* **89**, 3508-3522 (2005).
1118
- 1119 64. Kapanidis AN, Laurence TA, Lee NK, Margeat E, Kong X, Weiss S. Alternating-laser
1120 excitation of single molecules. *Acc Chem Res* **38**, 523-533 (2005).
1121
- 1122 65. Eggeling C, *et al.* Data registration and selective single-molecule analysis using multi-
1123 parameter fluorescence detection. *J Biotechnol* **86**, 163-180 (2001).
1124

- 1125 66. Hoffmann A, *et al.* Mapping protein collapse with single-molecule fluorescence and
1126 kinetic synchrotron radiation circular dichroism spectroscopy. *Proc Natl Acad Sci USA*
1127 **104**, 105-110 (2007).
1128
- 1129 67. Schuler B. Application of single molecule Förster resonance energy transfer to protein
1130 folding. *Methods Mol Biol* **350**, 115-138 (2007).
1131
- 1132 68. Hillger F, Nettels D, Dorsch S, Schuler B. Detection and analysis of protein aggregation
1133 with confocal single molecule fluorescence spectroscopy. *J Fluoresc* **17**, 759-765 (2007).
1134
- 1135 69. Benke S, Nettels D, Hofmann H, Schuler B. Quantifying kinetics from time series of
1136 single-molecule Förster resonance energy transfer efficiency histograms. *Nanotechnology*
1137 **28**, 114002 (2017).
1138
- 1139 70. Haynes WM. *Handbook of Chemistry and Physics*, 95 edn. CRC Press (2014).
1140
- 1141 71. Pace CN. Determination and analysis of urea and guanidine hydrochloride denaturation
1142 curves. *METHOD ENZYMOL* **131**, 266-280 (1986).
1143
- 1144 72. Schuler B, Lipman E, Eaton W. Probing the free-energy surface for protein folding with
1145 single-molecule fluorescence spectroscopy. *Nature* **419**, 743-747 (2002).
1146
- 1147 73. Dertinger T, Pacheco V, von der Hocht I, Hartmann R, Gregor I, Enderlein J. Two-
1148 focus fluorescence correlation spectroscopy: a new tool for accurate and absolute
1149 diffusion measurements. *Chemphyschem* **8**, 433-443 (2007).
1150
- 1151 74. Wilkins D, Grimshaw S, Receveur V, Dobson C, Jones J, Smith L. Hydrodynamic radii
1152 of native and denatured proteins measured by pulse field gradient NMR techniques.
1153 *Biochemistry* **38**, 16424-16431 (1999).
1154
- 1155 75. Aznauryan M, Nettels D, Holla A, Hofmann H, Schuler B. Single-molecule spectroscopy
1156 of cold denaturation and the temperature-induced collapse of unfolded proteins. *J Am*
1157 *Chem Soc* **135**, 14040-14043 (2013).
1158
- 1159 76. Nettels D, *et al.* Single-molecule spectroscopy of the temperature-induced collapse of
1160 unfolded proteins. *Proc Natl Acad Sci USA* **106**, 20740-20745 (2009).
1161
- 1162 77. Russinova E, Tretyachenko-Ladokhina V, Vele OE, Senear DF, Alexander Ross JB.
1163 Alexa and Oregon Green dyes as fluorescence anisotropy probes for measuring protein-
1164 protein and protein-nucleic acid interactions. *Anal Biochem* **308**, 18-25 (2002).
1165
- 1166 78. Likhachev ER. Dependence of water viscosity on temperature and pressure. *Tech Phys*
1167 **48**, 514-515 (2003).
1168
- 1169 79. O'Brien EP, Morrison G, Brooks BR, Thirumalai D. How accurate are polymer models
1170 in the analysis of Förster resonance energy transfer experiments on proteins? *J Chem*
1171 *Phys* **130**, 124903 (2009).
1172

80. Sanchez I. Phase Transition Behavior of the Isolated Polymer Chain. *Macromolecules* **12**, 980-988 (1979).
81. Ziv G, Haran G. Protein folding, protein collapse, and tanford's transfer model: lessons from single-molecule FRET. *J Am Chem Soc* **131**, 2942-2947 (2009).
82. Sherman E, Haran G. Coil-globule transition in the denatured state of a small protein. *Proc Natl Acad Sci U S A* **103**, 11539-11543 (2006).
83. Zheng W, Zerze GH, Borgia A, Mittal J, Schuler B, Best RB. Inferring properties of disordered chains from FRET transfer efficiencies. *J Chem Phys* **148**, 123329 (2018).
84. Gopich IV, Nettels D, Schuler B, Szabo A. Protein dynamics from single-molecule fluorescence intensity correlation functions. *J Chem Phys* **131**, 095102 (2009).
85. Dignon GL, Zheng W, Kim YC, Best RB, Mittal J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput Biol* **14**, e1005941 (2018).
86. Debye P, Hückel E. Zur Theorie der Elektrolyte: I. Gefrierpunktserniedrigung und verwandte Erscheinungen. *Phys Z* **24**, 185-206 (1923).
87. Ashbaugh HS, Hatch HW. Natively unfolded protein stability as a coil-to-globule transition in charge/hydrophathy space. *J Am Chem Soc* **130**, 9536-9542 (2008).
88. Kapcha LH, Rossky PJ. A simple atomic-level hydrophobicity scale reveals protein interfacial structure. *J Mol Biol* **426**, 484-498 (2014).
89. Plimpton S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J Comput Phys* **117**, 1-19 (1995).
90. Anderson JA, Glaser J, Glotzer SC. HOOMD-blue: A Python package for high-performance molecular dynamics and hard particle Monte Carlo simulations. *Comput Mater Sci* **173**, 109363 (2020).
91. McCarney ER, *et al.* Site-specific dimensions across a highly denatured protein; a single molecule study. *J Mol Biol* **352**, 672-682 (2005).
92. Makarov DE. Spatiotemporal correlations in denatured proteins: The dependence of fluorescence resonance energy transfer (FRET)-derived protein reconfiguration times on the location of the FRET probes. *The Journal of chemical physics* **132**, 035104 (2010).

Acknowledgements.

We thank William I. Weis for providing the E-cad plasmid and Dirk Görlich for the plasmid containing the SUMO protease. We are also grateful to Nir London and Christian Dubiella for their help with mass spectrometry. In addition, we enjoyed critical discussions with Gilad Haran, Amnon Horovitz, Koby Levy, Benjamin Schuler, Robert Best, and Philipp Selenko. This work was supported by the Israel Science Foundation (grant no. 1549/15), the Benozio Fund for the Advancement of Science, the Carolito Foundation, the Leir Charitable Foundation, and the Koshland family. W.Z. acknowledges the support from the National Science Foundation (MCB-2015030) and the research computing at Arizona State University for providing HPC.

Supplementary Figures

Supplementary Fig. 1: Sequence charge decoration metric (SCD_{lowsalt}, see Methods) for all E-cad segments computed based on the amino acid sequence (**a**) and salt-dependent FRET efficiencies of wild type E-cad and two charge-swapped variants of the A-segment (**b**).

Supplementary Fig. 2. nsFCS of all E-cad variants in the absence (left) and presence (right) of β -cat after exciting the acceptor directly. Quantitative comparison of the quenching amplitudes show that the C-segment has the highest amplitude in free E-cad. All quenching amplitudes are elevated in complex with β -cat.

Supplementary Fig. 3. Steady-state donor anisotropy of the six E-cad variants (indicated) in complex with β -cat. The population at a FRET efficiency close to zero arises from molecules that lack an active acceptor (gray shaded area). The second population at higher FRET efficiency (box) results from E-cad molecules with donor and acceptor in complex with β -cat. For all variants, the steady-state anisotropies are between 0.1 and 0.2, indicating no substantial restriction in the rotational freedom of the donor dye. The average donor anisotropies of the FRET-population is indicated.

Supplementary Fig. 4. Recurrence analysis of single particles (RASP) of the E-cad/β-cat complex. **(a)** Scheme of single molecules that are detected while they diffuse through the confocal volume. After the initial molecule exits the confocal volume, it can either return or a new molecule can enter. **(b)** Time-dependent probability to detect the same molecule several time decays. **(c)** Equilibrium FRET histogram of the E-cad/β-cat complex. For RASP, molecules were selected either from region 1 (low FRET) or from region 2 (high FRET). The dashed line indicates the mean FRET position (E_m). **(d)** Time-dependent recurrence histograms of molecules selected from region 1 (left) and region 2 (right). Solid lines are fits with a superposition of Gaussian peaks. The dashed line indicates the mean FRET position (E_m) of the equilibrium FRET histogram shown in c. **(e)** Time decays of the fraction of molecules in region 1 (low-FRET) after initially selecting molecules from region 1 (blue) or region 2 (red). The dashed line indicates the expected kinetics in the absence of conformation dynamics as computed from the p_{same} (see panel b).

Supplementary Fig. 5. Contact lifetimes in the E-cad/β-cat complex from the flexible CG-model. The contact lifetime map shows a significant distribution of contacts across the whole β-cat surface.

Supplementary Fig. 6. Thermal stability of β -cat monitored with nano-differential scanning fluorimetry (nanoDSF). The melting curve shows a denaturation at 57°C, indicating that β -cat is stable at the highest temperature used in the temperature-dependent RASP experiments (main text Fig. 6). Inset: SmFRET histograms of E-cad ABC at 37°C in the absence (blue) and presence (red) of β -cat (200 nM). The significant broadening of the FRET-distribution at 37°C and the shift to higher FRET-values indicate binding of E-cad at 37°C. Solid lines are fits with a superposition of two Gaussian peaks.

Supplementary Fig. 7. Calibration of smFRET experiments at different temperatures. **(a)** Peltier-controlled sample holder (left) and device to control the temperature of the objective (right). **(b)** Dual-focus FCS autocorrelation functions (top) and cross-correlation functions (bottom) of the dye Oregon Green at set temperatures 0°C (blue circles) and 70°C (red circles). Solid lines are fits as described in the Methods section. The diffusion coefficients at the two temperatures are indicated. **(c)** Temperature calibration curve obtained by determining the diffusion coefficient of Oregon Green using 2fFCS. The x-axis shows the temperature set on the sample controller. The objective controller is set to the same temperature up to a set-temperature of 40°C. For higher temperatures, the objective is kept at a set-temperature of 40°C. The y-axis shows the temperature determined based on the diffusion coefficient of Oregon Green and the known temperature dependence of the water viscosity (see Methods). The solid line is a polynomial fit of fourth order to the data and the blue shaded are indicates the 90% confidence band.

Supplementary Fig. 8. Parameterization of the coarse-grained (CG) model. Left: Deviation of the calculated FRET-values from the experimental data of the six constructs when varying the interaction strength (ϵ) of E-cad only. Right: The deviation of the model from the FRET experiment of the six constructs at the experimental ionic strength of 82 mM when varying the interaction strength (ζ) between E-cad and β -cat for the flexible CG-model (see Methods). The blue dashed lines indicate the minimum and the black dashed lines indicate the deviation for the rigid CG-model.

Supplementary Table 1. E-cad variants and sequence of β -cat used in the smFRET experiments.	
ABC	A ^C SGLRRRTVVKEPLLPPDDDDTRDNVYYYDEEGGGEEDQDFDLSQLHRGLDARPEVTRNDVAPTLMSVPQYRPRPANPDEIGNFIDENLKAADSPTAPPYDSLLVFDYEGSGSEAA ^C LSLNSSESDDQDQDYDLNEWGNRFKKLADMYGGGE DDGS ^C
AB	A ^C SGLRRRTVVKEPLLPPDDDDTRDNVYYYDEEGGGEEDQDFDLSQLHRGLDARPEVTRNDVAPTLMSVPQYRPRPANPDEIGNFIDENLKAADSPTAPPYDSLLVFDYEGSGSEAA ^C LSLNSSESDDQDQDYDLNEWGNRFKKLADMYGGGE DD
BC	ALRRRTVVKEPLLPPDDDDTRDNVYYYDEEGGGEEDQDFDL ^C QLHRGLDARPEVTRNDVAPTLMSVPQYRPRPANPDEIGNFIDENLKAADSPTAPPYDSLLVFDYEGSGSEAA ^C LSLNSSESDDQDQDYDLNEWGNRFKKLADMYGGGEDDGS ^C
A	A ^C SGLRRRTVVKEPLLPPDDDDTRDNVYYYDEEGGGEEDQDFDL ^C QLHRGLDARPEVTRNDVAPTLMSVPQYRPRPANPDEIGNFIDENLKAADSPTAPPYDSLLVFDYEGSGSEAA ^C LSLNSSESDDQDQDYDLNEWGNRFKKLADMYGGGE DD
B	ALRRRTVVKEPLLPPDDDDTRDNVYYYDEEGGGEEDQDFDL ^C QLHRGLDARPEVTRNDVAPTLMSVPQYRPRPANPDEIGNFIDENLKAADSPTAPPYDSLLVFDYEGSGSEAA ^C LSLNSSESDDQDQDYDLNEWGNRFKKLADMYGGGEDD
C	ALRRRTVVKEPLLPPDDDDTRDNVYYYDEEGGGEEDQDFDLSQLHRGLDARPEVTRNDVAPTLMSVPQYRPRPANPDEIGNFIDENLKAADSPTAPPYDSLLVFDYEGSGSEAA ^C LSLNSSESDDQDQDYDLNEWGNRFKKLADMYGGGEDDGS ^C
Core 1	^C SGAADSPTAPPYDSLLVFDYE
Core 2	AADSPTAPPYDSLLVFDYEGS ^C
β -cat	ATQADLMELDMAMEPDRKAAVSHWQQQSYLDSGIHSGATTAPSLSGKGNPEEEDVDTSQVLYEWEQGSQSFTQEQVADIDGQYAMTRAQRVRAAMFPETLDEGMQIPSTQFDDAAHPTNVQRLAEPQMLKHAVVNLINYQDDAELATRAIPE LTKLLNDEQVVVNKAAMVHQLSKKEASRHAIMRSPQMVS ^A IVRTMQNTNDVETARCTAGTLHNLSSHREGLLAIF KSGGIPALVKMLGSPVDSVLFYAITTLHNLLHQEGAKMAVRLAGGLQKMVALLNKTNVKFLAITTDCLQILAYGNQ ESKLIILASGGPQALVNIMRTYTYEKLLWTTSRVLKVL ^S VCSSNKPAIVEAGGMQALGLHLTDP ^S QRLVQNC ^L WLTLR NLSDAATKQEGMEGLLGT ^L VQLLGSD ^D INVVTC ^A AAGILSNLTCNNYKNKMMVCQVGGIEALVRTVLRAGDREDITEP AICALRHLTSRHQEAEMAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLALCPANHAPLREQGAIPRLVQLL VRAHQDTQRRTS ^M GGTQQQFVEGV ^R MEEIVEGCTGALHILARDVHNRI ^V IRGLNTIPLFVQLLYSP ^I ENIQ ^R VAAAGV LCELAQDK ^E AAEAIEAEGATAPL ^T ELLHSRNEG ^V ATYAAV ^L FRMS ^E DKPQDYKK ^R LSVELTSS ^L FRTEPMAWNETA DLGLDIGAQGEALGYRQDDPSYRSFHS ^G GYQDALGMDPMMEHEMGGHHPGADYPVDGLPDLGHAQDLMDGLPPGDS NQLAWFD ^T DL