ARTICLE

Extreme disorder in an ultrahigh-affinity protein complex

Alessandro Borgia[1]*, Madeleine B. Borgia[1]*, Katrine Bugge[2]*, Vera M. Kissling[1], Pétur O. Heidarsson[1], Catarina B. Fernandes[2], Andrea Sottini[1], Andrea Soranno[1,3], Karin J. Buholzer[1], Daniel Nettels[1], Birthe B. Kragelund[2], Robert B. Best[4] & Benjamin Schuler[1,5]

Molecular communication in biology is mediated by protein interactions. According to the current paradigm, the specificity and affinity required for these interactions are encoded in the precise complementarity of binding interfaces. Even proteins that are disordered under physiological conditions or that contain large unstructured regions commonly interact with well-structured binding sites on other biomolecules. Here we demonstrate the existence of an unexpected interaction mechanism: the two intrinsically disordered human proteins histone H1 and its nuclear chaperone prothymosin-α associate in a complex with picomolar affinity, but fully retain their structural disorder, long-range flexibility and highly dynamic character. On the basis of closely integrated experiments and molecular simulations, we show that the interaction can be explained by the large opposite net charge of the two proteins, without requiring defined binding sites or interactions between specific individual residues. Proteome-wide sequence analysis suggests that this interaction mechanism may be abundant in eukaryotes.

In the traditional paradigm of structural biology, intermolecular interactions are thought to be encoded in complementary shapes and noncovalent forces between folded biomolecules. However, it has become increasingly clear that many proteins involved in cellular interactions are fully or partially unstructured under physiological conditions[1,2]. Some of these intrinsically disordered proteins (IDPs) form well-defined 3D structures on binding to their targets[1]; in other complexes, parts of the IDP remain disordered. A broad spectrum of these protein complexes with differing degrees of disorder is known[3]. In some cases, a well-defined and structured binding interface is formed in the bound state, and only some loops or the chain termini remain disordered. In other cases, one of the binding partners remains almost completely unstructured in the complex, and its multiple binding motifs dynamically interact with the folded partner. Examples of the latter complexes include interdomain interactions in the cystic fibrosis transmembrane regulator[4]; the cyclin-dependent kinase inhibitor Sic1 in yeast binding to the substrate recognition subunit of its ubiquitin ligase subunit Cdc4[5]; the tail of the human sodium/proton exchanger 1 with the extracellular signal-regulated kinase ERK2[6]; and nuclear transport receptors interacting with nucleoporins[7]. The underlying multivalent binding enables unique regulatory mechanisms[8] and can mediate the formation of liquid–liquid phase separation[9], indicating the emergence of new modes of biomolecular interactions.

We have discovered a pair of proteins that constitute an extreme case of a highly unstructured protein complex with physiological function. One of the binding partners is the linker histone H1.0 (H1), which is involved in chromatin condensation by binding to nucleosomes[10,11]; this protein is largely unstructured[12] and highly positively charged, with two disordered regions that flank a small folded globular domain (Fig. 1, Extended Data Table 1). The other partner is the abundant nuclear protein prothymosin-α (ProTα), which is a fully unstructured, highly negatively charged IDP[13,14] involved in chromatin remodelling[15],

transcription, cellular proliferation and apoptosis[16]. ProTα acts as a linker histone chaperone by interacting with H1 and increasing its mobility in the nucleus[17]. Here we show that ProTα and H1 bind to one another with very high affinity, but that both proteins fully retain their structural disorder. By integrating experimental techniques and molecular simulations, we obtain a detailed model of this highly disordered and dynamic protein complex, which presents a previously undescribed paradigm of biomolecular binding.
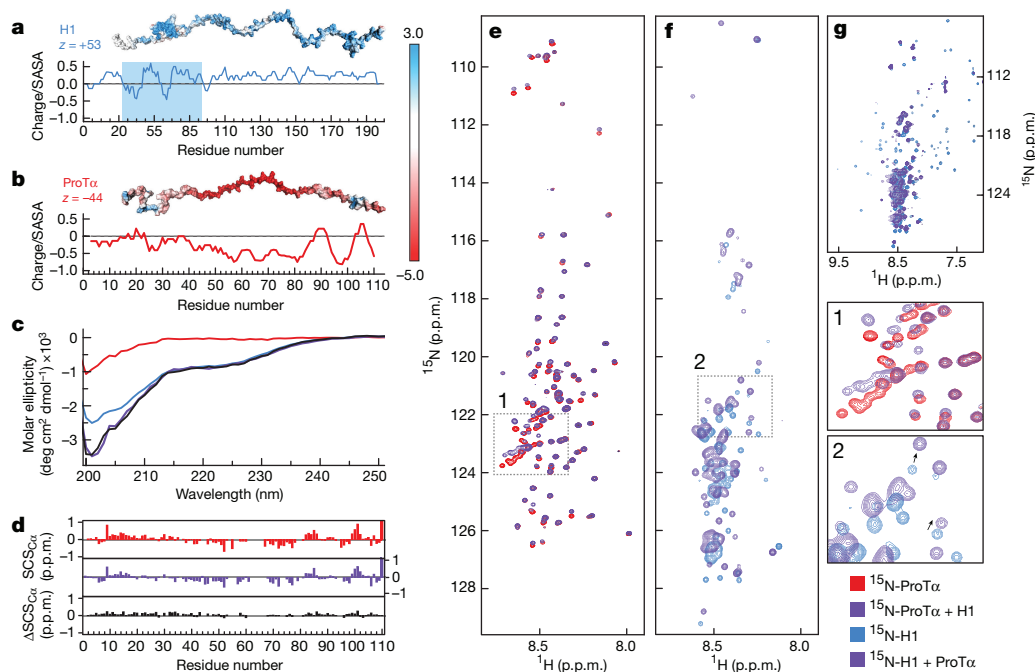
A highly unstructured protein complex

The binding of H1 to ProTα has been demonstrated both *in vitro*[18] and *in vivo*[17]. However, the high net charge, low hydrophobicity and pronounced disorder in the free proteins raise the question of how much structure is formed when they interact. We used circular dichroism and nuclear magnetic resonance (NMR) spectroscopy to investigate the formation of secondary and tertiary structure in ProTα and H1, separately and in complex with one another. The circular dichroism spectra of unbound ProTα and H1 reflect the low secondary structure content of each individual IDP, except for the small helix-turn-helix domain of H1[13,19,20] (Fig. 1c). Notably, the circular dichroism spectrum of an equimolar mixture of the two proteins can be explained by the simple sum of the individual spectra, indicating that complex formation entails minimal changes in average secondary structure content.

To obtain residue-specific information, we employed NMR spectroscopy. ^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) spectra of the individual proteins exhibit low dispersion of the ^{1}H chemical shifts, as expected for IDPs[14,21–23] (Fig. 1e, f). Only the globular domain of H1, which is stably folded even in isolation (Extended Data Fig. 1), shows the large dispersion of resonances characteristic of tertiary structure[23,24] (Fig. 1g). Remarkably, the overall peak dispersion remains unchanged on complex formation, confirming that no pronounced tertiary structure is formed on binding. Nevertheless, small but clearly

Figure 1 | ProTα and H1 remain unstructured upon binding.
a, b, Extended configurations of H1 (**a**) and ProTα (**b**), net charges (*z*), and surface electrostatic potentials with colour scale (units in $k_B T$ per e⁻). For the globular domain of H1, only residues with a solvent-accessible surface area (SASA) larger than $0.5\,nm^2$ are included, and are indicated by blue shading (see Extended Data Table 1). **c,** Far-UV circular dichroism spectra of ProTα (red), H1 (blue), the H1–ProTα mixture (purple) and their calculated sum (black) at 5 μM for each protein; curves are the mean of $n = 60$ individual spectra; $n = 2$ repeats of this measurement yielded consistent results. **d,** Cα secondary chemical shifts ($SCS_{C\alpha}$) of ProTα free (red), in complex with H1 (purple) and their differences ($\Delta SCS_{C\alpha}$, black). **e,** ^1H–^{15}N HSQC spectra of ^{15}N-ProTα in the absence (red) and presence (purple) of unlabelled H1; $n = 5$ repeats of this measurement yielded consistent results. **f,** ^1H–^{15}N HSQC spectra of ^{15}N-H1 in the absence (blue) and presence (purple) of unlabelled ProTα ($n = 2$), with zoomed-in regions corresponding to boxes in **e** (1) and **f** (2); arrows in 2 indicate the shift for selected resonances (see also Extended Data Fig. 3b). **g,** H1 spectra from **f** shown at a lower contour level.

detectable peak shifts observed for ProTα and H1 indicate changes in the average chemical environment of the corresponding residues, as expected on interaction with the large opposite charge of the other IDP. For ProTα, 95% of the amide backbone nuclei could be assigned (Extended Data Fig. 2), which enabled a residue-specific analysis: the Cα secondary chemical shifts[25] of ProTα show no evidence for the induction of persistent or transiently populated secondary structure on complex formation (Fig. 1d). The pronounced overlap in the NMR spectra of the unstructured parts of H1 precluded residue-specific assignments, but the clusters of Hα–Cα peaks in the ^1H–^{13}C HSQC spectrum from the lysine-rich disordered regions of H1 do not exhibit detectable chemical shift perturbations on titration with ProTα, and additional resonances do not emerge (Extended Data Fig. 3e, f). We therefore have no indications of changes in secondary structure content in H1 on ProTα binding.

The lower intensity of the resonances corresponding to the H1 globular domain (Fig. 1f, g, Extended Data Fig. 3) is likely to originate from the faster transverse (T_2) relaxation of structured, compared to unstructured, regions; additionally, tumbling of the globular domain is decelerated by the drag of the unstructured regions in which it is embedded[26]. On complex formation, the intensity of many H1 and ProTα resonances decreases, and those of the globular domain drop below the noise (Fig. 1f, g, Extended Data Fig. 3b). The large hydrodynamic radii of H1 and the complex (Extended Data Fig. 4a, b) support a large effective rotational correlation time as the origin of peak broadening, but a contribution from chemical exchange cannot be excluded. However, the globular domain is dispensable for complex formation (Fig. 2b).

High-affinity binding in spite of disorder

To quantify the strength of the interaction between H1 and ProTα, we used single-molecule Förster resonance energy transfer (FRET), which

enables measurements over a very broad range of affinities down to the picomolar regime. By labelling two positions with a donor and an acceptor dye, distances and distance changes between or within the polypeptides can be determined by confocal fluorescence detection of molecules freely diffusing in solution[27,28]. ProTα labelled at positions 56 and 110 (ProTα$_{56/110}$; all labelled residues are cysteines) exhibits a mean transfer efficiency, $\langle E \rangle$, of 0.33 at near-physiological ionic strength (Fig. 2a, Extended Data Table 2), as expected for this IDP, which is highly expanded owing to its large negative net charge[13,29,30]. On addition of unlabelled H1, a population with higher $\langle E \rangle$ of 0.58 (that is, shorter average distance within the ProTα chain) emerges: binding to the positively charged H1 evidently leads to a compaction of ProTα by charge screening, analogous to the compaction on addition of salt[29]. The same behaviour is observed for doubly labelled H1 (Extended Data Table 2), which demonstrates mutual adaptation of the conformational ensembles. The resulting dissociation constant in the low picomolar range reveals an extremely strong interaction (Fig. 2b, Extended Data Table 2), consistent with the physiological role of ProTα as a linker histone chaperone[17] that competes with the tight binding of H1 to chromatin[31]. Measurements with other FRET dyes and label positions resulted in similar affinities (Extended Data Table 2), indicating that labelling has only a small effect on binding. The dominant contribution to the interaction with ProTα stems from the unstructured C-terminal part of H1, which in isolation still binds with picomolar affinity. The N-terminal half and the isolated globular domain of H1 also bind to ProTα, but with much lower affinity (Fig. 2b). At least four isolated globular domains can bind to one ProTα molecule at the same time, with modest chemical shift changes (Extended Data Fig. 1), suggesting the absence of a specific binding interface.

The large and opposite net charges of ProTα (−44) and H1 (+53) imply a strong electrostatic contribution to binding. Indeed, a mere doubling of the ionic strength from the physiological 165 mM to
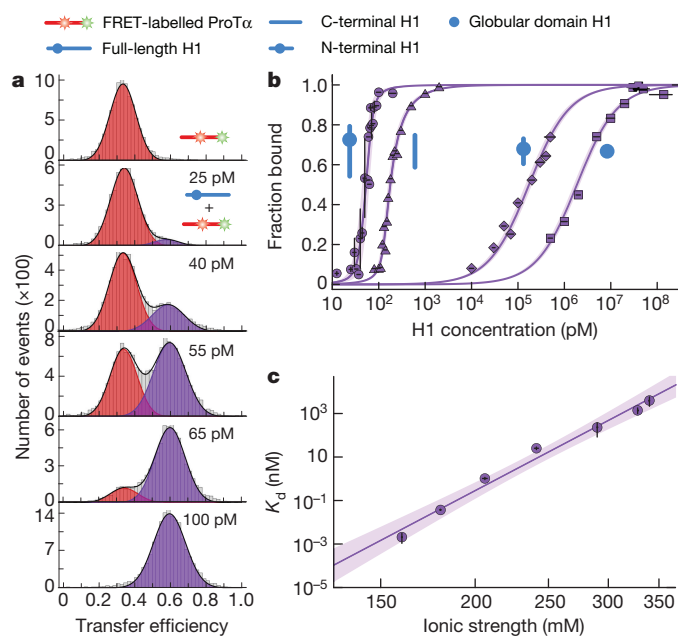
Figure 2 | ProTα and H1 form an electrostatically driven high-affinity complex. a, Single-molecule transfer efficiency histograms of ProTα$_{56/110}$ (FRET-labelled at positions 56 and 110) with varying concentrations (0 pM, 25 pM, 40 pM, 55 pM, 65 pM or 100 pM) of unlabelled H1 as indicated in the panels, fitted with an unbound peak (red) and a bound peak (purple). **b**, Binding isotherms based on transfer efficiency histograms for full-length H1 (circles, $K_d = 2.1^{+1.1}_{-0.8}$ pM), H1 N-terminal region (diamonds, $K_d = 173^{+29}_{-28}$ nM), H1 C-terminal region (triangles, $K_d = 40^{+6}_{-4}$ pM) and the globular domain of H1 (squares, $K_d = 1.9^{+0.3}_{-0.3}$ μM) at 165 mM ionic strength (see Extended Data Table 1 for details). **c**, K_d of H1–ProTα complex as a function of ionic strength with fit in terms of counter-ion release[50] (purple line) and 95% confidence interval (shaded). See Methods for details of statistics and data analysis.

340 mM reduces the affinity by six orders of magnitude (Fig. 2c). By extrapolation, a reduction of ionic strength to approximately 140 mM would take this interaction into the femtomolar range. From low picomolar to 100 μM protein concentrations, the stoichiometry from intermolecular FRET (Extended Data Fig. 4c), NMR chemical shift titrations (Extended Data Figs 2, 3) as well as the hydrodynamic radii measured with pulsed-field gradient NMR and two-focus fluorescence correlation spectroscopy (Extended Data Fig. 4a, b) all indicate the predominant formation of one-to-one dimers and the absence of large oligomers or coacervates[32]. However, in the presence of a large excess of one of the binding partners, we observe a decrease in FRET efficiencies that is indicative of the weak association of additional molecules with a dissociation constant (K_d) in the 10 to 100 μM range (Extended Data Fig. 4d, e), a propensity that is also observed in the simulations described later.

A highly dynamic complex

The lack of structure formation in the H1–ProTα complex implies great flexibility and a highly dynamic interconversion within a large ensemble of configurations and relative arrangements of the two IDPs. The presence of a broad and rapidly sampled distance distribution is supported by the analysis of fluorescence lifetimes[28,33,34] (Extended Data Fig. 5). Because fluctuations in distance cause fluctuations in the fluorescence intensity of donor and acceptor, the timescale of these long-range distance dynamics can be measured by single-molecule FRET combined with nanosecond fluorescence correlation spectroscopy[34,35]. For unfolded or disordered proteins, reconfiguration times (the relaxation times for distances within the chain) of approximately 20–200 ns have previously been observed[27]. ProTα is a particularly dynamic IDP, because of its highly expanded chain[13,29] and corresponding lack of

impeding intramolecular interactions[36]; in isolation it yields reconfiguration times (τ_r) between 29^{+2}_{-2} ns and 78^{+15}_{-9} ns, depending on the chain segment probed[34,36] (Extended Data Table 2). H1 labelled at positions 113 and 194 reconfigures more slowly than ProTα ($\tau_r = 118^{+24}_{-14}$ ns), but is within the range previously observed for unfolded and disordered proteins[27,34].
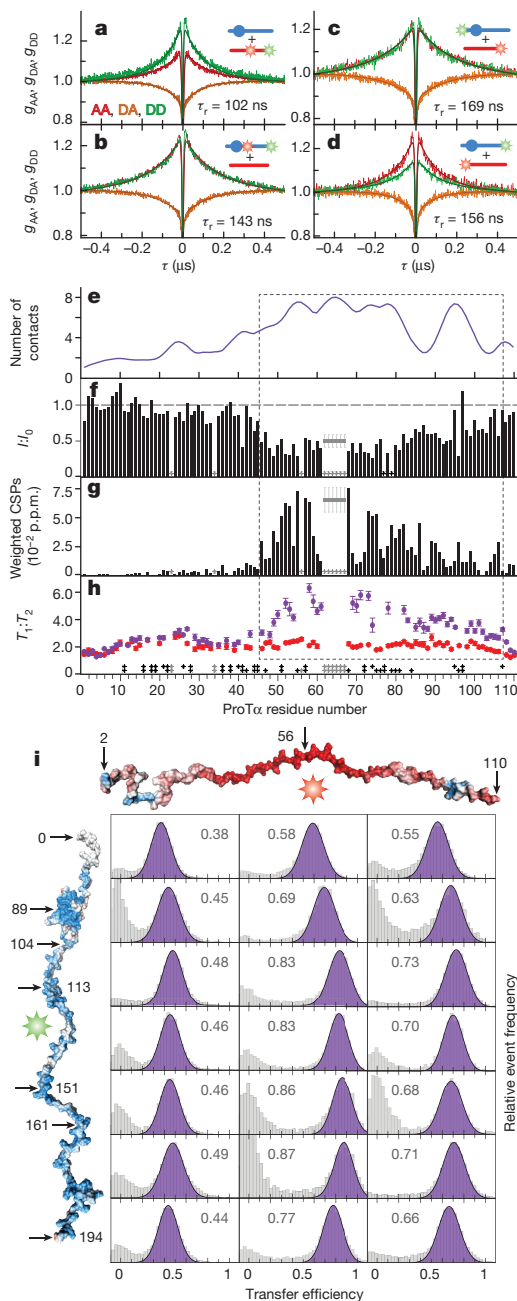
Notably, these pronounced and rapid long-range dynamics are retained in the H1–ProTα complex; we observed values for τ_r of between 66^{+2}_{-2} ns and 191^{+22}_{-19} ns for 13 different labelling pairs throughout the dimer (Fig. 3a–d, Extended Data Table 2). The similarity between the τ_r values of H1 and ProTα when bound in the complex suggests a coupling of the dynamics of the two intertwining chains. The highly dynamic nature of the complex is further supported by NMR: the longitudinal (T_1) and transverse (T_2) ^{15}N relaxation times reflect rapid backbone dynamics in the pico- to nanosecond range for ProTα in both the free and the bound state (Fig. 3h, Extended Data Fig. 2). The increase in T_1/T_2 (Fig. 3h) and in the hydrodynamic radius (Extended Data Fig. 4), as well as the reduced peak intensities (Fig. 3f) are consistent with the increase in τ_r for ProTα observed by nanosecond fluorescence correlation spectroscopy in the complex (Fig. 3a), in which chain–chain interactions are expected to slow down both local and long-range dynamics.

Architecture of an unstructured protein complex

To develop a structural representation of the conformational ensemble of the H1–ProTα complex, we combined single-molecule FRET, NMR and molecular simulations. We first mapped the complex with single-molecule FRET by probing a total of 28 intra- and intermolecular distances with donor and acceptor dyes in specific positions (Figs 3i, 4a). The resulting intermolecular transfer efficiencies lack the pronounced patterns that would be expected if persistent site-specific interactions or chain alignment in a preferred register were present. The intermolecular transfer efficiencies are most sensitive to the labelling position on ProTα, with the highest efficiencies (that is, shortest average distances) for the central position at residue 56 (ProTα$_{56}$), intermediate efficiencies for ProTα$_{110}$ and lowest efficiencies (that is, longest average distances) for ProTα$_2$. These results indicate that the region of highest charge density of ProTα (Fig. 1b) most strongly attracts H1. The charge density along H1 is more uniform (Fig. 1a), as are the transfer efficiencies to ProTα, albeit with some decrease towards the termini (Fig. 3i).

On the basis of this information we sought to establish a molecular model of the H1–ProTα complex. Given the lack of structure formation and residue-specific interactions, the dominance of electrostatics and the size of the system, we used a simplified coarse-grained model in which each residue is represented by a single bead. Coulombic interactions between all charged residues are included explicitly, with a screening factor to account for an ionic strength of 165 mM. Other attractive interactions and excluded volume repulsion are captured using a short-range potential, with the radius of the residues determined from their volumes[37]. A structure-based potential[38] is used to describe the folded globular domain of H1. The transfer efficiencies computed from Langevin dynamics simulations can be matched to the measured values (Fig. 4a) via the single adjustable parameter in our model—namely, the contact energy of the short-range potential—which is the same for all residues (see Methods); explicitly including a representation of the chromophores in the simulations yielded very similar results (Fig. 4a). The resulting intra- and intermolecular distance distributions (Extended Data Fig. 6d) are smooth and unimodal, which is consistent with the absence of site-specific interactions and structure formation observed experimentally and attests to the convergence of the simulations. The good agreement between the transfer efficiencies observed in our experiments and those obtained from the simulation indicates that this simple model captures the essential properties of the structural ensemble. Considering its simplicity, the femtomolar affinity estimated from the model

Figure 3 | Dynamics, interactions, and distances in the complex.
a–d, Examples of nanosecond fluorescence correlation spectroscopy probing long-range dynamics, based on intra- and intermolecular FRET (see Extended Data Table 2 for details); curves are the averages of $n = 3$ independent measurements of acceptor–acceptor (g_{AA}), donor–acceptor (g_{DA}) and donor–donor (g_{DD}) correlation, performed with doubly labelled ProTα with unlabelled H1 (**a**), doubly labelled H1 with unlabelled ProTα (**b**) and singly labelled H1 and singly labelled ProTα (**c**, **d**). Pictograms as in key in Fig. 2. **e**, Average number of contacts of each ProTα residue with H1 based on the simulations (Fig. 4b). **f**, Ratios of NMR resonance intensities of ProTα in the presence (I) and absence (I_0) of H1. **g**, Weighted backbone amide chemical shift perturbations (CSPs) of ProTα induced by equimolar H1 binding (see Extended Data Fig. 2 for other stoichiometries); $n = 5$ repeats of this measurement yielded consistent results. In **f** and **g**, the grey horizontal lines represent the average of three unassigned but traceable Glu residues in the range of residues 62–67 with error bars from their s.d. (see Methods). **h**, Ratios of T_1 and T_2 ^{15}N relaxation times of ProTα in free (red) and bound (purple) states (see Extended Data Fig. 2 for details). Light grey stars, prolines and unassigned residues; black stars, resonance overlap and/or insufficient data quality. Circles are mean values from $n = 3$ consecutive measurements, errors are s.d. The dashed box in **e–h** indicates the sequence range with the largest changes. **i**, Single-molecule transfer efficiency histograms from intermolecular FRET experiments between different positions in acceptor-labelled (red star) ProTα and donor-labelled (green star) H1, fitted with a single peak (purple, mean transfer efficiency shown). The signal at $E \approx 0$ originates from molecules without active FRET acceptor. For further information on statistics, see Methods.
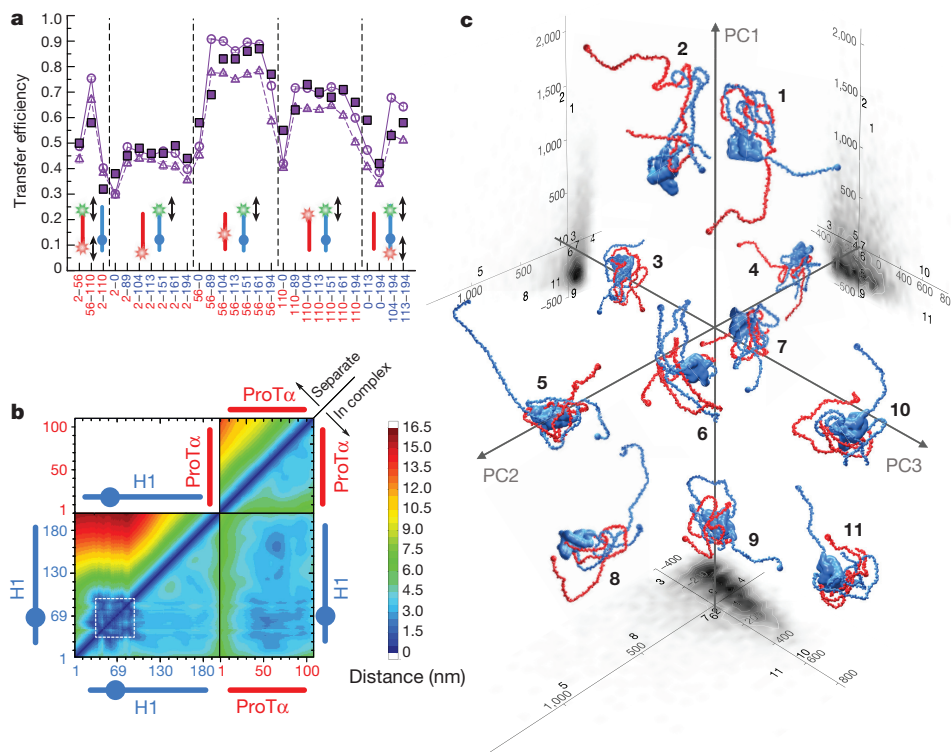
peaks, but the chemical shift and intensity perturbations in this cluster are similar to those observed in the rest of the region from residues 46 to 106 (Fig. 3f, g).

Further analysis of the simulated structural ensemble (see Supplementary Video 1) shows a lack of distinct conformational clusters (Extended Data Fig. 6a), which implies a continuous distribution of configurations. A projection of the simulation onto the first three principal components of the inter-residue distances (Extended Data Fig. 6c) reveals a highly heterogeneous ensemble of arrangements of the two entwining flexible chains (Fig. 4c). Given the rapid intramolecular dynamics and lack of structure in the complex, the activation barrier for binding is likely to be close to zero. Indeed, association of H1 and ProTα occurs at the diffusion limit, with a binding rate coefficient of $3.1 \pm 0.1 \times 10^9$ M^{-1} s^{-1} (Extended Data Fig. 7). The simulations support this mechanism, with a downhill free-energy surface for binding and attractive fly-casting[39] interactions enhanced by electrostatics[40] already emerging at a distance of approximately 22 nm, which is much greater than the sum of the hydrodynamic radii (Extended Data Fig. 6b).

Conclusions

Our results suggest that high-affinity complex formation between two oppositely charged IDPs is possible without the formation of structure or the need for folded domains. In contrast to the current paradigm for molecular recognition in biomolecular interactions, this type of highly dynamic complex does not require structurally defined binding sites or specific persistent interactions between individual residues. Instead, our findings are well-described as being the result of long-range electrostatic attraction between the two interpenetrating polypeptide chains, especially between their charge-rich regions. The exceedingly rapid interconversion of many different arrangements and configurations on the 100-ns timescale results in efficient averaging and essentially corresponds to a mean-field-type interaction[41,42] between all charges. This type of complex expands the known spectrum of protein–protein interactions. Although the H1–ProTα complex is extreme in the extent of its disorder for both binding partners, the possibility of this interaction mechanism is not entirely unexpected, given the prevalence of charged amino acids in many IDPs[2], the previous observation of disorder in IDPs interacting with folded proteins[3–7]

(Extended Data Fig. 5b) is remarkably consistent with the affinities that were observed experimentally near this ionic strength (165 mM). The affinity for a second molecule of H1 or ProTα to the complex is predicted to be orders of magnitude weaker than for the first molecule, consistent with the experimental results (Extended Data Figs 4d, e, 6b).

The intra- and intermolecular distance maps from the simulation (Fig. 4b) indicate that the interactions between ProTα and H1 are broadly distributed along their sequences, but also reflect the asymmetry in electrostatic attraction owing to the higher charge density of ProTα in its central and C-terminal regions (Figs 1b, 4a). The NMR results provide an independent experimental test of the model: the distribution of the average number of contacts made by the residues of ProTα based on the simulation (Fig. 3e) is notably similar to the distribution of changes in chemical shifts, peak intensities and T_1/T_2 ratios observed on binding (Fig. 3f–h). These changes occur across the same broad region between residues 46 and 106, encompassing the most acidic tracts of ProTα. Overlap within the Glu cluster of the NMR spectra prevents the quantitative analysis of some of the corresponding

Figure 4 | Architecture of the complex from simulations. a, Comparison of experimental (filled squares) and simulated transfer efficiencies (empty symbols) in the H1–ProTα complex for the pairs of dye positions indicated on the *x* axis (red, position on ProTα; blue, position on H1). Triangles and circles indicate simulations with and without explicit chromophores, respectively. Pictograms as in key in Fig. 2. **b**, Intra- and intermolecular average distance maps of H1 and ProTα from the simulations, separately and in complex. The white dashed square indicates the globular domain (only surface-exposed residues shown, see Extended Data Table 1). **c**, Examples of configurations of H1 (blue) and ProTα (red) in the complex; N termini are indicated by small spheres. The structures are projected onto the first three principal components (PC1, PC2 and PC3) of the distance map, with projections of the full ensemble shown as grey scatter plots (units of Å, see also Extended Data Fig. 6 and Supplementary Video 1). Bold numbers indicate the positions of the structures in the principal component projections.

and the role of electrostatics in the formation of dynamic binding interfaces between folded proteins[43]. Moreover, the H1–ProTα interaction resembles polyelectrolyte complexes formed by charged synthetic polymers[42], even though the latter usually phase-separate into coacervates[42]. The absence of coacervation[32,42] or liquid–liquid phase separation[9] for ProTα and H1 at concentrations ranging from the picomolar to high micromolar regime may be due to the complementarity[44] of the two proteins in terms of effective length and opposite net charge, leading to optimal and mutually saturating electrostatic interactions, or to the lack of hydrophobic and aromatic side chains and cation–π interactions, which have previously been suggested to favour phase separation mediated by proteins (refs 32, 45 and R. M. Vernon *et al.*, unpublished).

There are several functional implications of this mechanism for forming a high-affinity yet unstructured dynamic complex between two IDPs. Histone H1 is a key factor in chromatin condensation and transcriptional regulation[11], and ProTα acts as a chaperone of H1 that facilitates its displacement from and deposition onto chromatin[17]. ProTα thus needs to be able to compete with the very high affinity of the histone to chromatin[31]. However, high affinities between structured biomolecules are usually linked to exceedingly slow dissociation[40], which is incompatible with fast regulation. By contrast, the high affinity of the H1–ProTα complex is facilitated by its ultra-fast association, which allows dissociation on a biologically relevant timescale in spite of the high affinity required for function. Another consequence of polyelectrolyte interactions is the possibility of ternary complex formation[46], signs of which are detected here with a large excess of ProTα or H1 (Extended Data Figs 4d, e, 6b), resulting in mostly unexplored kinetic mechanisms that cannot be explained by competition via simple dissociation and re-association[47]. Finally, the flexibility within such unstructured complexes may facilitate access for enzymes that add posttranslational modifications; these modifications have key roles in the regulation of cellular processes, including those of H1. One example of this mechanism may be the interaction of the acidic domain of the oncogene SET with the lysine-rich C-terminal tail of p53, which is regulated by acetylation[48].

The behaviour we observe for ProTα and H1 might be surprisingly widespread, as highly charged protein sequences that could form such complexes are abundant in eukaryotes. In the human proteome alone, several hundred proteins that are predicted to be intrinsically disordered[49] contain contiguous stretches of at least 50 residues with a fractional net charge similar to that of H1 or ProTα. As the interaction of highly oppositely charged IDPs is unlikely to be very sequence-specific[18], achieving binding selectivity may be linked to other regulatory mechanisms such as cellular localization or synchronized expression during relevant stages of development or the cell cycle.

1. Wright, P. E. & Dyson, H. J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **19,** 31–38 (2009).
2. Habchi, J., Tompa, P., Longhi, S. & Uversky, V. N. Introducing protein intrinsic disorder. *Chem. Rev.* **114,** 6561–6588 (2014).
3. Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.* **33,** 2–8 (2008).
4. Baker, J. M. *et al.* CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat. Struct. Mol. Biol.* **14,** 738–745 (2007).
5. 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).
6. 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,** 31 (2016).
7. Milles, S. *et al.* Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* **163,** 734–745 (2015).
8. Csizmok, V., Follis, A. V., Kriwacki, R. W. & Forman-Kay, J. D. Dynamic protein interaction networks and new structural paradigms in signaling. *Chem. Rev.* **116,** 6424–6462 (2016).
9. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18,** 285–298 (2017).
10. Robinson, P. J. & Rhodes, D. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr. Opin. Struct. Biol.* **16,** 336–343 (2006).
11. Hergeth, S. P. & Schneider, R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* **16,** 1439–1453 (2015).
12. Hansen, J. C., Lu, X., Ross, E. D. & Woody, R. W. Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J. Biol. Chem.* **281,** 1853–1856 (2006).
13. Gast, K. *et al.* Prothymosin α: a biologically active protein with random coil conformation. *Biochemistry* **34,** 13211–13218 (1995).

14. Uversky, V. N. *et al.* Natively unfolded human prothymosin α adopts partially folded collapsed conformation at acidic pH. *Biochemistry* **38,** 15009–15016 (1999).

15. Gómez-Márquez, J. & Rodríguez, P. Prothymosin α is a chromatin-remodelling protein in mammalian cells. *Biochem. J.* **333,** 1–3 (1998).

16. Mosoian, A. Intracellular and extracellular cytokine-like functions of prothymosin α: implications for the development of immunotherapies. *Future Med. Chem.* **3,** 1199–1208 (2011).

17. George, E. M. & Brown, D. T. Prothymosin α is a component of a linker histone chaperone. *FEBS Lett.* **584,** 2833–2836 (2010).

18. Papamarcaki, T. & Tsolas, O. Prothymosin α binds to histone H1 *in vitro. FEBS Lett.* **345,** 71–75 (1994).

19. Barbero, J. L., Franco, L., Montero, F. & Morán, F. Structural studies on histones H1. Circular dichroism and difference spectroscopy of the histones H1 and their trypsin-resistant cores from calf thymus and from the fruit fly *Ceratitis capitata. Biochemistry* **19,** 4080–4087 (1980).

20. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. & Sweet, R. M. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* **362,** 219–223 (1993).

21. Yi, S., Brickenden, A. & Choy, W. Y. A new protocol for high-yield purification of recombinant human prothymosin α expressed in *Escherichia coli* for NMR studies. *Protein Expr. Purif.* **57,** 1–8 (2008).

22. Khan, H. *et al.* Fuzzy complex formation between the intrinsically disordered prothymosin α and the Kelch domain of Keap1 involved in the oxidative stress response. *J. Mol. Biol.* **425,** 1011–1027 (2013).

23. Zhou, B. R. *et al.* Structural insights into the histone H1–nucleosome complex. *Proc. Natl Acad. Sci. USA* **110,** 19390–19395 (2013).

24. Zarbock, J., Clore, G. M. & Gronenborn, A. M. Nuclear magnetic resonance study of the globular domain of chicken histone H5: resonance assignment and secondary structure. *Proc. Natl Acad. Sci. USA* **83,** 7628–7632 (1986).

25. Kjaergaard, M., Brander, S. & Poulsen, F. M. Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J. Biomol. NMR* **49,** 139–149 (2011).

26. Bae, S. H., Dyson, H. J. & Wright, P. E. Prediction of the rotational tumbling time for proteins with disordered segments. *J. Am. Chem. Soc.* **131,** 6814–6821 (2009).

27. Schuler, B., Soranno, A., Hofmann, H. & Nettels, D. Single-molecule FRET spectroscopy and the polymer physics of unfolded and intrinsically disordered proteins. *Annu. Rev. Biophys.* **45,** 207–231 (2016).

28. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods Enzymol.* **475,** 455–514 (2010).

29. Müller-Späth, S. *et al.* Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA* **107,** 14609–14614 (2010).

30. Hofmann, H. *et al.* Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl Acad. Sci. USA* **109,** 16155–16160 (2012).

31. White, A. E., Hieb, A. R. & Luger, K. A quantitative investigation of linker histone interactions with nucleosomes and chromatin. *Sci. Rep.* **6,** 19122 (2016).

32. Pak, C. W. *et al.* Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Mol. Cell* **63,** 72–85 (2016).

33. Chung, H. S., Louis, J. M. & Gopich, I. V. Analysis of fluorescence lifetime and energy transfer efficiency in single-molecule photon trajectories of fast-folding proteins. *J. Phys. Chem. B* **120,** 680–699 (2016).

34. Soranno, A. *et al.* Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proc. Natl Acad. Sci. USA* **109,** 17800–17806 (2012).

35. Nettels, D., Gopich, I. V., Hoffmann, A. & Schuler, B. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl Acad. Sci. USA* **104,** 2655–2660 (2007).

36. Soranno, A. *et al.* Integrated view of internal friction in unfolded proteins from single-molecule FRET, contact quenching, theory, and simulations. *Proc. Natl Acad. Sci. USA* **114,** E1833–E1839 (2017).

37. Creigthon, T. E. *Proteins: Structures and Molecular Properties* 2nd edn (W. H. Freeman and Co., 1993).

38. Karanicolas, J. & Brooks, C. L. III. The origins of asymmetry in the folding transition states of protein L and protein G. *Protein Sci.* **11,** 2351–2361 (2002).

39. Shoemaker, B. A., Portman, J. J. & Wolynes, P. G. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl Acad. Sci. USA* **97,** 8868–8873 (2000).

40. Schreiber, G., Haran, G. & Zhou, H. X. Fundamental aspects of protein–protein association kinetics. *Chem. Rev.* **109,** 839–860 (2009).

41. Borg, M. *et al.* Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc. Natl Acad. Sci. USA* **104,** 9650–9655 (2007).

42. Srivastava, S. & Tirrell, M. V. in *Advances in Chemical Physics* (eds Rice, S. A. & Dinner, A. R.) Ch. 7, 499–544 (John Wiley & Sons, 2016).

43. Ahmad, A.*et al.* Heat shock protein 70 kDa chaperone/DnaJ cochaperone complex employs an unusual dynamic interface. *Proc. Natl Acad. Sci. USA* **108,** 18966–18971 (2011).

44. Freeman Rosenzweig, E. S. *et al.* The eukaryotic CO_2-concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171,** 148–162.e119 (2017).

45. Nott, T. J. *et al.* Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57,** 936–947 (2015).

46. Peng, B. & Muthukumar, M. Modeling competitive substitution in a polyelectrolyte complex. *J. Chem. Phys.* **143,** 243133 (2015).

47. Berlow, R. B., Dyson, H. J. & Wright, P. E. Hypersensitive termination of the hypoxic response by a disordered protein switch. *Nature* **543,** 447–451 (2017).

48. Wang, D.*et al.* Acetylation-regulated interaction between p53 and SET reveals a widespread regulatory mode. *Nature* **538,** 118–122 (2016).

49. Dosztányi, Z., Csizmok, V., Tompa, P. & Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21,** 3433–3434 (2005).

50. Record, M. T. Jr, Anderson, C. F. & Lohman, T. M. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* **11,** 103–178 (1978).

Author Contributions A.B., M.B.B., K.B., B.B.K., R.B.B. and B.S. designed research; M.B.B., A.B., V.M.K. and A.Sot. produced and labelled fluorescent protein variants; A.B. and M.B.B. performed single-molecule experiments; A.B., M.B.B., A.Sor. and D.N. analysed single-molecule data; D.N. developed single-molecule instrumentation and data analysis tools; A.Sot. and A.B. carried out stopped-flow measurements, A.B., M.B.B., K.J.B. and A.Sot. established experimental conditions for single-molecule measurements; C.B.F. and P.O.H. produced protein samples for NMR; K.B. and C.B.F. performed and analysed NMR measurements; A.Sor. carried out the bioinformatics analysis; R.B.B. conducted and analysed simulations; A.B., B.B.K. and C.B.F. carried out circular dichroism experiments; B.B.K., R.B.B. and B.S. supervised research; B.S., A.B., R.B.B., B.B.K. and K.B. wrote the paper with help from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to B.S. (schuler@bioc.uzh.ch), A.B. (a.borgia@bioc.uzh.ch), R.B.B. (robert.best2@nih.gov) or B.B.K. (bbk@bio.ku.dk).

Reviewer Information *Nature* thanks E. Zuiderweg and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample sizes. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Protein preparation. ProTα was prepared and labelled as previously described[29]. For a complete list of all protein variants, labelling positions and fluorophores used for single-molecule FRET, see Extended Data Table 1. The correct molecular mass of all protein variants and labelled constructs was confirmed by mass spectrometry.

For experiments using wild-type human linker histone H1.0 (H1), recombinant protein was used (New England Biolabs, cat.# M2501S). For the production of labelled H1 variants and wild-type H1 for NMR, the cDNA of the human *H1F0* gene (UniProt P07305) was cloned into a modified version of the pRSET vector[51]. In this plasmid, the N-terminal hexahistidine tag and thrombin cleavage site were removed and replaced by a C-terminal hexahistidine tag and thrombin cleavage site (with sequence GGPRGSRGSHHHHHH) to enable purification of H1 free of degradation products. Cysteine mutations were introduced using site-directed mutagenesis, to enable labelling with fluorescent dyes using maleimide coupling (see Extended Data Table 1 for a complete list of variants). All H1 variants were expressed in *Escherichia coli* C41 cells and terrific broth medium at 37 °C, induced with 0.5 mM isopropylthiogalactopyranoside (IPTG) at an OD_{600} of ~0.6, and grown for 3 further hours. Cell pellets were collected and resuspended in denaturing buffer (6 M guanidinium chloride (GdmCl) in phosphate-buffered saline (PBS, 10 mM sodium phosphate pH 7.4, 137 mM NaCl, 2.7 mM KCl), the soluble fraction was collected and applied to a Ni-IDA resin (ABT Beads) in batch. The resin was washed twice with 5 resin volumes of denaturing buffer including 25 mM imidazole, three times with 5 resin volumes of PBS including 25 mM imidazole, and the protein was eluted with PBS including 250 or 500 mM imidazole. The protein was dialysed against PBS, filtered and its hexahistidine tag cleaved off with 5 U of thrombin (Serva) per milligram of H1, for 2 h at room temperature. To remove uncleaved protein and the tag, the mixture was run through a HisTrap HP 5-ml column (GE Healthcare) in PBS including 25 mM imidazole. H1 was further purified using a Mono S ion exchange chromatography column (GE Healthcare), washed with 20 mM Tris (pH 8.0) including 200 mM NaCl, and eluted in 20 mM Tris (pH 8.0) buffer with a gradient from 200 mM to 1M NaCl. Finally, samples for labelling were reduced with 20 mM dithiothreitol and purified by reversed-phase high-performance liquid chromatography (HPLC) with a Reprosil Gold C4 column with a gradient from 5% acetonitrile and 0.1% trifluoroacetic acid in aqueous solution to 100% acetonitrile. H1-containing fractions were lyophilized and resuspended in degassed 6 M GdmCl, 50 mM sodium phosphate buffer (pH 7.0). For double labelling, both dyes (dissolved in dimethylsulfoxide) were added to the protein in a 1:1:1 molar ratio; for single labelling, dye was added at a 0.7:1 molar ratio of dye to protein. Reactions were incubated at room temperature for 2 h, and stopped by adding 20 mM dithiothreitol. Products were purified by reversed-phase HPLC, and the correct mass of all labelled proteins confirmed by mass spectrometry (see Extended Data Fig. 8 for an example). Lyophilized labelled protein was dissolved in 8 M GdmCl and stored at −80 °C.

For NMR experiments, H1, ProTα and the H1 globular domain (Extended Data Table 1) were either produced in unlabelled form by growing cells in LB medium, uniformly labelled with ^{15}N by growing cells in M9 minimal medium containing $^{15}NH_4Cl$ as the sole source of nitrogen, or uniformly labelled with ^{15}N and ^{13}C by growing cells in M9 minimal medium containing $^{15}NH_4Cl$ as the sole source of nitrogen and $^{13}C_6$-glucose as the sole source of carbon, as previously described[6] and then purified essentially as explained earlier. The H1 globular domain was expressed as a GST-fusion protein with a TEV protease site, and purified on a glutathione sepharose 4 fast-flow column (GE Healthcare). The column was washed with 10 column volumes (CV) of PBS and the tagged protein eluted with 5 CV of elution buffer (50 mM Tris-HCl pH 8.0, 10 mM reduced glutathione). All fractions containing GST–H1 globular domain were pooled and cleaved with TEV protease (100 μl of 0.5 mg ml$^{−1}$ stock solution) overnight, and subsequently applied to a HiTrap SP FF 5 mL (GE Healthcare) with 50 mM sodium phosphate pH 9.0 and eluted with 50 mM sodium phosphate pH 9.0, 1 M NaCl over 25 CV. The protein-containing fractions were applied to a Superdex 75 10/300 (GE Healthcare) in TBS buffer (10 mM Tris, 157 mM KCl, 0.1 mM EDTA, pH 7.4) and further concentrated using an Amicon Ultra-15 centrifugal filter device (Millipore) with a molecular weight cutoff of 3 kDa. Protein concentrations of H1 and H1 globular domain were determined by UV absorbance, and the concentration of ProTα was determined by BCA assay (Thermo Fisher Scientific).

NMR spectroscopy. To minimize amide exchange, all NMR spectra were acquired at 283 K, unless otherwise specified, on a Varian INOVA 800-MHz (^1H) spectrometer with a room temperature probe or Bruker AVANCE III 600- or

750-MHz (^1H) spectrometers equipped with cryogenic probes. Free induction decays were transformed and visualized in NMRPipe[52] or Topspin (Bruker Biospin) and analysed using CcpNmr Analysis software[53]. Assignments of backbone nuclei of ^{13}C–^{15}N-labelled ProTα in the unbound state (0.1 mM ^{13}C–^{15}N-labelled ProTα, TBS buffer, 10% D_2O (v/v), 0.7 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS)) and at sub-saturating concentration (1:0.8 molar ratio) of H1 (0.1 mM ^{13}C–^{15}N-labelled ProTα, 0.08 mM H1, TBS buffer, 10% D_2O (v/v), 0.7 mM DSS) were performed manually from the analysis of ^1H–^{15}N HSQC, HNCACB, CBCA(CO)NH, HN(CO)CA, HNCO and HN(CA)NNH spectra acquired with non-uniform sampling[54] using standard pulse sequences. At saturating concentrations of H1, backbone resonances of ProTα became too weak for successful assignments. Proton chemical shifts were referenced internally to DSS at 0.00 p.p.m., with heteronuclei referenced by relative gyromagnetic ratios. The content of transient structure in ProTα was evaluated for each state from secondary Cα-chemical shifts assigned in the free form and at 80% saturation of H1 using a random coil reference set for IDPs[25]. In both states, three transient marginally populated α-helices were identified: residues Ser9–Glu19 (~10% populated), Ala82–Thr86 (~13% populated) and Val99–Lys102 (~18% populated). The populations of the transient α-helices were estimated from the average SCS value of the residues of the transient helices divided by 2.8 p.p.m. ($SCS_{Cα}$ value expected for a fully populated α-helix[55]) and were very similar in the free and bound states. ^1H–^{15}N HSQC spectra of ^{15}N-labelled H1 (40 μM) were recorded in the absence and presence of ProTα (40 μM). ^1H–^{13}C HSQC and/or ^1H–^{15}N HSQC spectra were acquired on four different sequential titrations: addition of up to 44 μM H1 to 11 μM ^{15}N-labelled ProTα; addition of up to 140 μM H1 globular domain to 20 μM ^{15}N-labelled ProTα; addition of up to 400 μM ProTα to 100 μM ^{13}C–^{15}N-labelled H1–Gly-Ser-6 × His; and addition of up to 700 μM ProTα to 100 μM ^{13}C–^{15}N-labelled H1 globular domain. Before each titration, the proteins were concentrated and dialysed in the same beaker. Subsequently, the solution of labelled protein was split equally into two samples, to one of which the unlabelled titrant was added at the maximum concentration, and to the other the same volume of dialysis buffer. After acquisition of NMR spectra on the two samples, they were used to obtain titration points between the end points by sequentially mixing the sample of the complex into the free protein. All NMR titration data were recorded in TBS buffer, 10% D_2O (v/v), 0.7 mM DSS. Binding-induced weighted CSPs were calculated as the weighted Euclidean distance between the peaks using $|\gamma_N|/|\gamma_H| = 0.154$. Owing to extensive resonance overlap of H^N, N, Cα and Cβ resonances in the 2D and 3D NMR spectra, assignments of backbone nuclei were not possible for the Glu repeat region from Glu62–Glu67. Nonetheless, spin systems displaying resonances consistent with Glu residues with Glu neighbours could be identified, and by exclusion were assigned to be part of the Glu62–Glu67 Glu repeat. For three of these systems, amide backbone peaks could be confidently tracked in the titration of ^{15}N-labelled ProTα with H1. The intensity ratios and weighted CSPs of the three Glu amide backbone peaks on addition of equimolar H1 were calculated and the average value used to represent the Glu repeat region in Fig. 3f, g. Peaks from the remaining Glu residues were present in the spectra of both free and bound states of ProTα, but could not be followed unambiguously during the titrations.

The hydrodynamic radii (R_H) of ProTα alone and at saturating concentrations of H1 or H1 globular domain were determined from a series of ^1H–^{15}N HSQC spectra with preceding pulse-field gradient stimulated-echo longitudinal encode–decode diffusion filter[56] and with the gradient strength increasing linearly from 0.963 to 47.2 G cm$^{−1}$. To determine the diffusion coefficients (D) the decay curves of the amide peaks were plotted against the gradient strength and fitted in Dynamics Center (Bruker) using $I = I_0\exp(−Dx^2\gamma^2\delta^2(\Delta − \delta/3) \times 10^4)$, in which I is the intensity of the NMR signal at the respective gradient strength, I_0 the intensity without applied gradient, x the gradient strength in G cm$^{−1}$, $\gamma = 26752$ rad Gs$^{−1}$, $\delta = 3$ ms, $\Delta = 250$ ms. R_H was calculated from the diffusion coefficient using the Stokes–Einstein relation, $R_H = k_BT/(6\pi\eta D)$, with η being the viscosity of water at 283 K.

T_1 and T_2 ^{15}N relaxation times were determined from 2× two series of ^1H–^{15}N HSQC spectra with varying relaxation delays using the pulse sequence of reference[57], and using pulsed-field gradients to suppress solvent resonances. The series were recorded on free ^{15}N-ProTα and on ^{15}N-ProTα with saturating concentrations of H1 at 800 MHz (^1H), using eight (10 ms, 100 ms, 300 ms, 500 ms, 700 ms, 1,100 ms, 1,300 ms and 1,500 ms) and seven (50 ms, 90 ms, 130 ms, 190 ms, 230 ms, 390 ms and 490 ms) different relaxation delays for T_1 and T_2, respectively, plus triplicate measurements. The relaxation decays were fitted to single exponentials and relaxation times determined using CcpNmr Analysis software[53].

Single-molecule fluorescence spectroscopy. Single-molecule measurements were performed using either a custom-built confocal instrument[58] or a MicroTime 200,

both equipped with a HydraHarp 400 counting module (PicoQuant). The donor dye was excited with light from a 485-nm diode laser (LDH-D-C-485, PicoQuant) at an average power of $100\,\mu W$ at the sample. The laser was operated in continuous-wave mode or in pulsed mode with alternating excitation of the dyes, achieved using pulsed interleaved excitation[59]. The wavelength range used for acceptor excitation was selected with a z582/15 band pass filter (Chroma) from the emission of a supercontinuum laser (EXW-12 SuperK Extreme, NKT Photonics) driven at 20 MHz, which triggers interleaved pulses from the 485-nm diode laser used for donor excitation. Emitted photons were collected by the microscope objective (Olympus UplanApo $60\times/1.20\,W$), focused onto a 100-μm pinhole, and then separated into four channels with a polarizing beam splitter and two dichroic mirrors (585DCXR, Chroma). Emission was additionally filtered by bandpass filters (ET525/50M and HQ650/100, Chroma) before being focused onto one of four single-photon avalanche detectors (Optoelectronics SPCM AQR-15, PerkinElmer or τ-SPADs, PicoQuant).

FRET efficiency histograms of doubly labelled ProTα and H1 were acquired on samples with concentrations of labelled protein between 10 and 100 pM. For intermolecular measurements, up to 500 pM of acceptor-labelled protein were used to ensure saturation of binding. Measurements were performed in TBS buffer (165 mM ionic strength) or in an analogous buffer with higher ionic strength (adjusted by increasing the KCl concentration, as noted), in the presence of 140 mM β-mercaptoethanol (Sigma-Aldrich) for photoprotection[60] and 0.01% Tween 20 (Pierce) to minimize surface adhesion[61]. To avoid the pronounced interaction of H1 with glass surfaces, more-inert polymer sample chambers (μ-Slide, ibidi) were used throughout. Transfer efficiencies were obtained from $E = n_A/(n_A + n_D)$, in which n_D and n_A are the numbers of donor and acceptor photons, respectively, in each burst, corrected for background, channel crosstalk, acceptor direct excitation, differences in quantum yields of the dyes and detection efficiencies[61]. Even in cases in which pulsed interleaved excitation was insufficient to completely eliminate the donor-only contribution to the signal (Fig. 3i), the population at zero transfer efficiency was sufficiently well separated from the FRET population that the reliability of the transfer efficiencies was not affected. Fluorescence anisotropy values were determined for all labelling positions using polarization-sensitive detection in the single-molecule instrument[28,62], and were between 0.04 and 0.16 for the monomeric proteins, and between 0.08 and 0.22 in the complex, indicating sufficiently rapid orientational averaging of the fluorophores to justify the approximation $\kappa^2 \approx 2/3$ used in Förster theory[63].

The low fluorescence anisotropy values, the consistency of the FRET and NMR results, and the self-consistency of a large number of labelling positions suggest that the fluorophores do not entail a severe perturbation of the interaction between ProTα and H1. However, to assess the effect of fluorophore labelling in more detail, we tested how different dye pairs and labelling positions influence the affinity between ProTα and H1 and the inferred inter-dye distances (Extended Data Table 2). In view of the high net charge of the proteins, alternative fluorophores with a net charge different from Alexa 488 and 594 (both net charge −2) were chosen: Cy3B (GE Healthcare Life Sciences; zwitterionic with zero net charge), Abberior STAR 635 (Abberior GmbH; zwitterionic with zero net charge), and Atto550 and Atto647N (ATTO-TEC; both net charge +1). The K_d values for the respective binding partner were between 1.0 nM and 3.5 nM (at 205 mM ionic strength to simplify quantification) for all labelling positions and dye pairs, corresponding to an energetic perturbation of binding by at most ~1 k_BT. To test for the effect of the fluorophores on the inferred distances, we recorded single-molecule transfer efficiency histograms of ProTα labelled at positions 56 and 110 (ProTα$_{56/110}$) with Cy3B/Abberior STAR 635 and Atto550/Atto647N, and of H1 labelled at positions 104 and 194 (H1$_{104/194}$) with Cy3B/Abberior STAR 635, both with and without the respective unlabelled binding partner present. The resulting transfer efficiency values yielded root mean square interdye distances consistent with those inferred from measurements with Alexa 488/594 (assuming a Gaussian chain distribution of inter-dye distances[29] and an experimental uncertainty of ± 0.05 for the transfer efficiency due to instrument calibration for the different dye pairs).

Fluorescence lifetime analysis. The comparison of ratiometric transfer efficiencies with the mean fluorescence lifetimes of donor and acceptor provides a further diagnostic for the presence of a broad distribution rapidly sampled during the time of a fluorescence burst[28,33,34]. Average lifetimes were estimated by using the mean donor ($\langle t_D \rangle$) and acceptor ($\langle t_A \rangle$) arrival times of the respective photons in a burst relative to the exciting laser pulse, and were combined with transfer efficiencies in a two-dimensional plot (Extended Data Fig. 5), in which $\tau_D^D/\tau_D^0 = \langle t_D \rangle/\tau_D^0$ and $\tau_D^A/\tau_D^0 = (\langle t_A \rangle - \tau_A^0)/\tau_D^0$ were calculated for each burst. In these equations, τ_D^0 is the intrinsic donor lifetime in the absence of the acceptor, and τ_A^0 is the intrinsic acceptor lifetime. For a single, fixed inter-dye distance (and thus transfer efficiency, E), one finds $\tau_D^D/\tau_D^0 = \tau_D^A/\tau_D^0 = 1 - E$, as illustrated by the diagonal line in Extended Data Fig. 5.

Nanosecond fluorescence correlation spectroscopy. Data for nanosecond fluorescence correlation spectroscopy were acquired at a concentration of ~100 pM of the protein carrying the donor (or both donor and acceptor) and an excess of the partner (either unlabelled or acceptor-labelled) to saturate binding. Donor and acceptor fluorescence emission (on continuous-wave excitation at 485 nm) from the subpopulation corresponding to the H1–ProTα complex in a transfer efficiency histogram was correlated with a binning time of 1 ns. To avoid the effects of detector dead times and after-pulsing on the correlation functions, the signal was recorded with two detectors each for donor and acceptor and cross-correlated between detectors[34,35]. Autocorrelation curves of acceptor and donor channels and cross-correlation curves between acceptor and donor channels were computed from the measurements and analysed as previously described[34,64]. In brief, auto- and cross-correlation curves were fitted over a time window of 2.5 μs with

$$g_{ij}(\tau) = 1 + \frac{1}{N}(1 - c_{ab}e^{-|\tau|/\tau_{ab}})(1 + c_{cd}e^{-|\tau|/\tau_{cd}})(1 + c_T e^{-|\tau|/\tau_T}) \text{ and } i,j = A, D$$

in which i and j correspond to donor or acceptor fluorescence emission; N is the effective mean number of molecules in the confocal volume; c_{ab}, τ_{ab}, c_{cd} and τ_{cd} are the amplitudes and time constants of photon antibunching and chain dynamics, respectively; and c_T and τ_T refer to the triplet blinking component on the microsecond timescale. Distance dynamics result in a characteristic pattern of the correlation functions based on donor and acceptor emission, with a positive amplitude in the autocorrelations ($c_{cd} > 0$) and a negative amplitude in the cross-correlation ($c_{cd} < 0$), but with identical decay times. All three correlation curves were thus fitted globally with the same values of τ_{cd}. Independent values of c_{cd}, c_{ab}, τ_{ab}, τ_T and c_T were used as free-fit parameters for each correlation curve. τ_{cd} was converted to the reconfiguration time of the chain, τ_r, as previously described[64], by assuming that chain dynamics can be modelled as a diffusive process in the potential of mean force derived from the sampled inter-dye distance distribution $P(r)$[35,64]. In light of the good agreement between the transfer efficiencies observed experimentally and in the simulations, we employed the $P(r)$ distributions obtained from the simulations for the respective pairs of labelling sites (intra- or intermolecular). This conversion does not entail a large change in timescale, and τ_{cd} and τ_r differ by less than 20% in all cases investigated here, depending on the average distance relative to the Förster radius[64]. The correlation functions shown in Fig. 3a–d were normalized to 1 at their respective values at 0.5 μs to facilitate direct comparison.

Two-focus fluorescence correlation spectroscopy. Two-focus fluorescence correlation spectroscopy measurements of Alexa 594-labelled ProTα were performed at 295 K on a MicroTime 200 confocal microscope equipped with a differential interference contrast prism. Alexa 594 was excited alternately with two orthogonally polarized laser beams: one beam with a wavelength range centred at 582 nm, selected with a z582/15 band pass filter (Chroma) from the emission of a supercontinuum fibre laser (EXW-12 SuperK Extreme, NKT Photonics) driven at 20 MHz, triggers (interleaved) pulses from a second supercontinuum laser with wavelength-selected output at 585 ± 3 nm (Solea, PicoQuant), with a combined repetition rate of 40 MHz and a power of 15 μW per laser at the sample. The distance between the two foci, δ, was calibrated as previously described on the basis of sample standards quantified under identical conditions using dynamic light scattering[30,65], yielding a δ of 490 ± 15 nm at $\lambda_{ex} = 585$ nm, corresponding to a systematic error of 3% of the calculated value of the hydrodynamic radius R_H. The concentration of labelled protein used in these experiments was ~4 nM in TBS buffer in the presence of 140 mM β-mercaptoethanol and 0.01% Tween 20. Translational diffusion coefficients were obtained from fits of the correlation functions[66] and converted to R_H using the Stokes–Einstein equation.

Analysis of binding isotherms. At ionic strengths of 200 mM and above, binding titrations of ProTα and H1 were hyperbolic and could be described well with a Langmuir-type isotherm, valid when the ligand concentration is sufficiently large compared to the analyte concentration. For example, with H1 as the ligand and ProTα as the analyte,

$$\frac{c_{H1-ProT\alpha}}{c_{ProT\alpha}^{tot}} = \frac{c_{H1}^{tot}}{K_d + c_{H1}^{tot}},$$

in which the subscripts to c indicate the concentration of the species (that is, c_x for species x), and c_x^{tot} the total concentration of x. However, below an ionic strength of about 200 mM the affinity of H1 for the surface of the sample chambers in which the measurements were performed was so high that the surface of the chamber noticeably competed with H1 binding by ProTα; the polymeric sample chambers we used already exhibit much lower affinity for H1 than glass surfaces, which are negatively charged. This results in a decrease in the effective H1 bulk concentration available for binding to ProTα that leads to a shift of the apparent midpoint of the titration to higher H1 concentrations and to a distortion of the curve to a

non-hyperbolic shape. To account for this effect, we need to take into account two coupled equilibria, one for the adsorption of H1 to surface binding sites, S, and one for H1 binding to ProTα:

$$H1 + S \rightleftarrows H1S \text{ and } H1 + ProT\alpha \rightleftarrows H1\text{–}ProT\alpha,$$
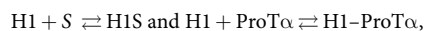
with the dissociation constants

$$K_d^{H1S} = \frac{c_{H1} \times \Gamma_S}{\Gamma_{SH1}} \quad (1)$$

and

$$K_d^{H1\text{–}ProT\alpha} = \frac{c_{H1} \times c_{ProT\alpha}}{c_{H1\text{–}ProT\alpha}} \quad (2)$$

in which c_{H1}, $c_{ProT\alpha}$, and $c_{H1\text{–}ProT\alpha}$ are the bulk concentrations of free H1, free ProTα and complex, respectively. Γ_S and Γ_{SH1} are the surface concentrations (that is, binding sites per area) of free binding sites and of binding sites occupied by H1, respectively. The resulting corresponding total concentrations are given by three equations:

$$c_{H1}^{tot} = c_{H1} + c_{H1\text{–}ProT\alpha} + \alpha\, \Gamma_{SH1} \quad (3)$$

$$c_{ProT\alpha}^{tot} = c_{ProT\alpha} + c_{H1\text{–}ProT\alpha} \quad (4)$$

and

$$\Gamma_{S\,tot} = \Gamma_S + \Gamma_{SH1} \quad (5)$$

In these equations, α is the surface-to-volume ratio of the sample well. Equations (1) to (5) were solved for the fraction of H1-bound ProTα using Mathematica (Wolfram Research) and the solution used to fit the titrations with full-length H1 and the H1 C-terminal fragment at 165 mM ionic strength (Fig. 2b) and full-length H1 at 185 mM ionic strength (Fig. 2c), with the adjustable parameters $K_d^{H1\text{–}ProT\alpha}$ and K_d^{H1S}, and with the product $\alpha \times \Gamma_{SH1}$; $c_{ProT\alpha}^{tot}$ was fixed to the known value. The vertical error bars in Fig. 2b were estimated from five independent measurements. The horizontal error bars represent the pipetting errors estimated for the applied sequences of dilution steps. We obtained upper and lower bounds for the binding isotherms by taking into account the uncertainty of the ProTα concentration; these bounds are displayed as shaded bands in Fig. 2b. The resulting K_d for the full-length proteins at 165 mM ionic strength follows the trend expected from the measurements at higher ionic strength (Fig. 2c), validating the analysis. The weak association of additional monomers at high micromolar excess of binding partner was ignored in this analysis because it occurs in a different concentration regime. The dependence of the K_d on ion activity, a, (Fig. 2c) was analysed using a previously developed formalism[50], according to the approximation: $d\ln(K_d)/d\ln(a) \approx -\Delta n = 18 \pm 1$ (standard error of the fit), in which $-\Delta n$ corresponds to the number of anionic and cationic counter ions released upon association of the two proteins, and the ion activity was approximated by the ionic strength.

Circular dichroism spectroscopy. Far-UV circular dichroism measurements were carried out on a Jasco J-810 spectropolarimeter, using a 1-mm path length quartz cuvette. Wild-type H1 and ProTα$_{56}$ samples were measured at a concentration of 5 μM in TBS and 5 mM β-mercaptoethanol at 20 °C. A total of 20–60 spectra per sample were recorded between 250 and 195 nm with 1-nm step size, averaged and a buffer spectrum was subtracted. The far-UV circular dichroism spectrum of the H1 globular domain was recorded at 283 K from 260 to 198 nm with a scan speed of 20 nm/min, 10 accumulations and a response time of 2 s at a protein concentration of 10 μM in TBS, and the buffer spectrum was subtracted. To assess the thermal stability of the H1 globular domain, thermal unfolding was monitored at 222 nm from 283 to 378 K in increments of 1 K per minute. The ellipticity as a function of temperature was fitted with $\theta(T) = f_U(T)\theta_U(T) + (1 - f_U(T))\theta_N(T)$. In this equation,

$$f_U(T) = \left(1 + \exp\left(-\frac{\Delta H_m}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right)\right)\right)^{-1}$$

which represents the fraction of unfolded H1 globular domain; ΔH_m represents the enthalpy change of folding at the transition midpoint, and R the gas constant. $\theta_N(T)$ and $\theta_U(T)$ are linear baselines from the folded and unfolded states, respectively, as a function of absolute temperature, T.

Binding kinetics of H1 and ProTα. Mixing experiments were carried out with an Applied Photophysics Pi Star-180 stopped-flow spectrometer. A solution of ProTα doubly labelled with Alexa 488/594 (at positions 56 and 110) at a concentration of

2.2 nM was mixed with a solution of unlabelled H1 at variable concentrations using a 1:10 mixing ratio. The increase in acceptor fluorescence emission resulting from the compaction of ProTα on H1 binding (see Figure 2a) was used to monitor the binding reaction by exciting at 436 nm with a 10-nm bandwidth using a HgXe lamp and recording fluorescence emission using a 580-nm long-pass filter. The buffer used was TBS in the presence of 0.01% Tween 20 to minimize surface adhesion of the proteins. For each final H1 concentration between 5 nM and 100 nM, at least 80 measurements were recorded and averaged.

Simulation methods. A coarse-grained model was used for both proteins, in which each residue is represented by a single bead centred on the Cα atom. The potential energy had the functional form:

$$\begin{aligned} V = &\frac{1}{2}\sum_{i<N} k_b(d_i - d_i^0)^2 + \frac{1}{2}\sum_{i<N-1} k_\theta(\theta_i - \theta_i^0)^2 \\ &+ \sum_{i<N-2}\sum_{n=1}^{4} k_{i,n}(1 + \cos(n\phi_i - \delta_{i,n})) + \sum_{i<j}\frac{q_i q_j}{4\pi\varepsilon_d\varepsilon_0}\exp\left[-\frac{d_{ij}}{\lambda_D}\right] \\ &+ \sum_{(i,j)\notin nat} 4\varepsilon_{pp}\left(\left(\frac{\sigma_{ij}}{d_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{d_{ij}}\right)^6\right) \\ &+ \sum_{(i,j)\in nat} \varepsilon_{ij}\left(13\left(\frac{\sigma_{ij}}{d_{ij}}\right)^{12} - 18\left(\frac{\sigma_{ij}}{d_{ij}}\right)^{10} + 4\left(\frac{\sigma_{ij}}{d_{ij}}\right)^6\right) \end{aligned}$$

The first two terms describe harmonic bond and angle energies, respectively, with force constants $k_b = 3.16 \times 10^5$ kJ mol^{-1} nm^{-2} and $k_\theta = 6.33 \times 10^2$ kJ mol^{-1} rad^{-2}, and reference values d_i^0 and θ_i^0 taken from an extended backbone structure. The third term is a sequence-based statistical torsion potential taken from the Go model[38], and is applied to all residues, and the fourth term is a screened coulomb potential, with Debye screening length λ_D that is applied to all residues with non-zero charges q_i. ε_0 is the permittivity of free space and ε_d the dielectric constant, set here to 80. The fifth term is a generic short-range attractive potential applied to all residue pairs not identified as being part of the natively folded globular domain of H1. This interaction is characterized by a contact distance $\sigma_{ij} = (\sigma_i + \sigma_j)/2$, in which σ_{ij} represents the residue diameters (all ~6 Å) determined from residue volumes[37], and by a contact energy ε_{pp}, which is the same for all such non-native residue pairs. The final term is an attractive potential applied only to the residues identified as native in the folded histone domain. The Go model[38] gives the residues that are considered native as well as the values of the parameters σ_{ij} and ε_{ij} for native pairs. For the electrostatic term, the charges are +1 for lysine and arginine, −1 for glutamate and aspartate, and +0.5 for histidine (to account for its pK_a near 6). The screening length, or Debye length, λ_D is given by

$$\lambda_D = \left(\frac{\varepsilon_d\varepsilon_0 k_B T}{2 N_A e^2 I}\right)^{1/2}$$

in which k_B is the Boltzmann constant, T is the temperature, N_A is the Avogadro constant, e is the elementary charge and I is the ionic strength in molar units. The variation of ionic strength only enters the model through the screening length. Although this treatment of electrostatics is very simplified, it is consistent with the coarse-grained level of the rest of the model.

There was therefore only one free parameter to be determined (ε_{pp}); the same value was used for all inter- and intramolecular interactions. We varied ε_{pp} in order to obtain an optimal agreement with all the FRET data. This optimal value was found to be 0.16 $k_B T$, or ~0.40 kJ mol^{-1}. Langevin dynamics simulations were run at a temperature of 300 K, with a friction coefficient of 0.1 ps^{-1} and a time step of 10 fs for 20 μs for each run; the mass of each bead was that of the corresponding residue. Simulations of the bound complex were started either with the molecules separated, or in an initially contacting configuration. Results from either simulation were the same, neglecting the equilibration part of the simulation. We also tested the effect of variations of the model. Using a residue-independent value of 6 Å for σ_{ij} for all residue pairs did not appreciably change the results. Similarly, using a residue-specific short-range potential similar to that devised in a previous protein interaction model[67] did not improve the agreement with experiment. However, a model with randomized or uniform charges (equal to the average) for the two proteins was unable to capture the important qualitative features of the data, in particular the difference in FRET efficiencies between the N and C termini of ProTα, and H1. This result emphasizes the dominant role of electrostatics in determining the properties of the complex.

We also considered whether the results may have been influenced by the presence of the FRET chromophores and the linkers used to covalently attach them to the protein. We therefore ran an additional set of simulations, one for each labelling

combination, in which we included an explicit, coarse-grained representation of the linkers. The linkers and dyes were approximated by 5 beads (for each dye plus linker), in an unbranched chain, and with similar properties to the protein (bond lengths 3.8 Å, all bond angles 110°; the dihedral angle term was omitted). One end of the chain was bonded to the bead for the labelled residue. The motivation for the choice of 5 beads and protein-like geometry was the earlier finding that the effect of linkers on unfolded proteins can be accounted for by adding an extra 9–10 residues to the true number of residues separating the labelling positions[30,65]. The short-range interaction of the dyes with themselves and the protein was given by a Lennard–Jones term similar to that used for the other non-native interactions in the model, but the parameters were set so as to give only a short-range repulsion, with $\varepsilon = 0.001$ kJ mol^{-1}, and $\sigma = 6$ Å. Each chromophore carries a net charge of -2, which was included by adding a charge of -1 to each of the two beads furthest from the attachment point to the protein. Explicit simulations were run for each labelling combination considered in the paper.

Dissociation constants were estimated by umbrella sampling using the centre-of-mass distance between the proteins as coordinate, with harmonic umbrellas spaced between 0 and 25 nm and a force constant of 10 kJ mol^{-1} nm^{-2}. The potential of mean force $F_{WHAM}(r)$ along the distance r between the centres of mass of the proteins was reconstructed using weighted histogram analysis (WHAM)[68], and the effective pair potential $F_{eff}(r)$ (Extended Data Fig. 6b) was obtained from $F_{eff}(r) = F_{WHAM}(r) + 2k_BT\log(r)$, in which k_B is the Boltzmann constant and T the temperature. F_{eff} was shifted by a constant energy so that the interaction energy at large separations was zero. The dissociation constant K_d was calculated from

$$K_d^{-1} = 4\pi N_A \int_0^{r_b} \exp[-\beta F_{eff}(r)]r^2 dr$$

in which r_b is the radius defining the maximum extent of the bound state (in which $F_{eff}(r)$ becomes non-zero), and $\beta = 1/k_BT$.

Conformations were initially analysed using a previously devised clustering algorithm[69], which was applied to the Hamming distances between the binary contact maps of different conformations (using a distance cut-off of 8 Å to define a contact). This algorithm identifies cluster centres as structures, i, with a high density of neighbours, ρ_i, (many structures at a short distance), but which have a large distance to the nearest structure with higher neighbour density, δ_i. The 'decision graph' consists of plotting δ_i versus ρ_i for all structures. Cluster centres should appear as points at the top right of the graph. The decision graph in this case (Extended Data Fig. 6a) shows only a single cluster. Other clustering algorithms also provided little evidence for distinct clusters, suggesting that all structures fall into a single, very broad state. We therefore used principal component analysis as a way of projecting out the structural variations. We used a set of coarse-grained inter-residue distances as the space in which to perform principal component analysis, in which only every fifth residue in the sequence was considered, and all pair distances between such residues were computed. We obtained the principal components by diagonalization of the variance–covariance matrix of this set of distances. The first three components are represented as matrices in Extended Data Fig. 6c.

Statistics and sample sizes of single molecule experiments and simulations. Minimum and average numbers of single molecules for which fluorescence has been recorded and used to build transfer efficiency histograms are indicated below every figure.

In Fig. 1c, 2 independent measurements: curves shown are the average of 60 spectra each.

In Fig. 2a, 5 independent titrations, 19 different protein concentrations, minimum number of molecules > 1,000 each (~4,000 molecules on average). Each transfer efficiency histogram constitutes an independent measurement of the affinity, because the relative populations can be determined directly from the peak integrals.

In Fig. 2b, for full-length H1: 5 independent titrations, 19 different protein concentrations, minimum number of molecules > 1,000 each (~4,000 molecules on average); for H1 C-terminal disordered region: 1 titration, 19 different protein concentrations, minimum number of molecules > 2,500 each (~4,600 molecules on average); for H1 N-terminal disordered region: 1 titration, 10 different protein concentrations, minimum number of molecules > 1,100 each (~4,200 molecules on average); for H1 globular domain: 1 titration, 12 different protein concentrations, minimum number of molecules > 8,200, (~1.6 × 10^4 molecules on average).

In Fig. 2c, for 'Ionic strength (IS) = 180 mM': 1 titration with 8 different protein concentrations, minimum number of molecules > 2,000 each (~4,000 molecules on average); for 'IS = 205 mM': 2 titrations with an average of 14 different protein concentrations, minimum number of molecules > 2,900 each (~5,300 molecules on average); for 'IS = 240 mM': 2 titrations with an average of 10 different protein

concentrations, minimum number of molecules > 3,000 each (~4,800 molecules on average); for 'IS = 290 mM': 3 titrations with an average of 12 different protein concentrations, minimum number of molecules > 2,700 each (~6,000 molecules on average); for 'IS = 330 mM': 2 titrations with an average of 8 different protein concentrations, minimum number of molecules > 3,100 each (~5,700 molecules on average); for 'IS = 340 mM': 2 titrations with an average of 6 different protein concentrations, minimum number of molecules > 950 each (~4,000 molecules on average).

In Fig. 3, ProT$\alpha_{56/110}$ + H1: 3 independent measurements, sample size: ~5 × 10^6 molecules (Fig. 3a); ProTα + H1$_{0/194}$: 3 independent measurements, sample size: ~5 × 10^5 molecules (Fig. 3b); ProT$\alpha_{110(Alexa 594)}$ + H1$_{0(Alexa 488)}$: 3 independent measurements, sample size: ~5 × 10^5 molecules (Fig. 3c); ProT$\alpha_{2(Alexa 594)}$ + H1$_{194(Alexa 488)}$: 3 independent measurements, sample size: ~9 × 10^5 molecules (Fig. 3d). In Fig. 3i, 1 measurement each, minimum number of molecules > 2,900 each, ~6,800 molecules on average.

In Extended Data Fig. 4c, 1 measurement each; ProT$\alpha_{2(Alexa 594)}$ + H1$_{194(Alexa 488)}$: ~7,300 molecules; ProT$\alpha_{56(Alexa 594)}$ + H1$_{194(Alexa 488)}$: ~6,900 molecules; ProT$\alpha_{110(Alexa 594)}$ + H1$_{194(Alexa 488)}$: ~8,700 molecules. In Extended Data Fig. 4d, 1 measurement, minimum number of molecules > 3,600 (~4,500 on average). In Extended Data Fig. 4e, 1 measurement, minimum number of molecules > 800 (~1,400 on average).

In Extended Data Fig. 5, 1 experiment each; ProT$\alpha_{56/110}$: ~7,500 molecules; ProT$\alpha_{56/110}$ + H1: ~3,000 molecules; H1$_{0/113}$: ~2,000 molecules; H1$_{0/113}$ + ProTα: ~3,300 molecules; ProT$\alpha_{2(Alexa 594)}$ + H1$_{194(Alexa 488)}$: ~7,400; ProT$\alpha_{110(Alexa 594)}$ + H1$_{194(Alexa 488)}$: ~8,700.

The uncertainty of the FRET efficiencies estimated from simulations was determined by block analysis in a similar fashion to a previously described method[70], in which the error around the mean is estimated from statistically independent blocks of simulation data.

Code availability. A custom module for Mathematica (Wolfram Research) used for the analysis of single-molecule fluorescence data is available upon request.

Data availability. All data supporting the findings of this study are available within the paper and its Supplementary Information. The raw data are available from the corresponding authors upon reasonable request. Source Data for Fig. 2 and Extended Data Table 2 are provided with the paper. The chemical shift assignments of ProTα alone and in complex with H1 have been deposited to the Biological Magnetic Resonance Bank (http://www.bmrb.wisc.edu/) under accession numbers 27215 and 27216, respectively.

51. Scott, K. A., Steward, A., Fowler, S. B. & Clarke, J. Titin: a multidomain protein that behaves as the sum of its parts. *J. Mol. Biol.* **315**, 819–829 (2002).
52. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
53. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
54. Orekhov, V. Y. & Jaravine, V. A. Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Prog. Nucl. Magn. Reson. Spectrosc.* **59**, 271–292 (2011).
55. Fedyukina, D. V. *et al.* Contribution of long-range interactions to the secondary structure of an unfolded globin. *Biophys. J.* **99**, L37–L39 (2010).
56. Gibbs, S. J. & Johnson, C. S. Jr. A PFG NMR experiment for accurate diffusion and flow studies in the presence of eddy currents. *J. Magn. Reson.* **93**, 395–402 (1991).
57. Farrow, N. A. *et al.* Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* **33**, 5984–6003 (1994).
58. Benke, S. *et al.* The assembly dynamics of the cytolytic pore toxin ClyA. *Nat. Commun.* **6**, 6198 (2015).
59. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed interleaved excitation. *Biophys. J.* **89**, 3508–3522 (2005).
60. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3**, 891–893 (2006).
61. Schuler, B. Application of single molecule Förster resonance energy transfer to protein folding. *Methods Mol. Biol.* **350**, 115–138 (2007).
62. Kellner, R. *et al.* Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. *Proc. Natl Acad. Sci. USA* **111**, 13355–13360 (2014).
63. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **437**, 55–75 (1948).
64. Gopich, I. V., Nettels, D., Schuler, B. & Szabo, A. Protein dynamics from single-molecule fluorescence intensity correlation functions. *J. Chem. Phys.* **131**, 095102 (2009).
65. Borgia, A. *et al.* Consistent view of polypeptide chain expansion in chemical denaturants from multiple experimental methods. *J. Am. Chem. Soc.* **138**, 11714–11726 (2016).
66. Dertinger, T. *et al.* Two-focus fluorescence correlation spectroscopy: a new tool for accurate and absolute diffusion measurements. *ChemPhysChem* **8**, 433–443 (2007).

67. Kim, Y. C. & Hummer, G. Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding. *J. Mol. Biol.* **375,** 1416–1433 (2008).

68. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A. & Rosenberg, J. M. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13,** 1011–1021 (1992).

69. Rodriguez, A. & Laio, A. Machine learning. Clustering by fast search and find of density peaks. Science **344,** 1492–1496 (2014).

70. Flyvbjerg, H. & Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **91,** 461–466 (1989).

71. Kapanidis, A. N. *et al.* Alternating-laser excitation of single molecules. *Acc. Chem. Res.* **38,** 523–533 (2005).

72. Hoffmann, A. *et al.* Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc. Natl Acad. Sci. USA* **104,** 105–110 (2007).
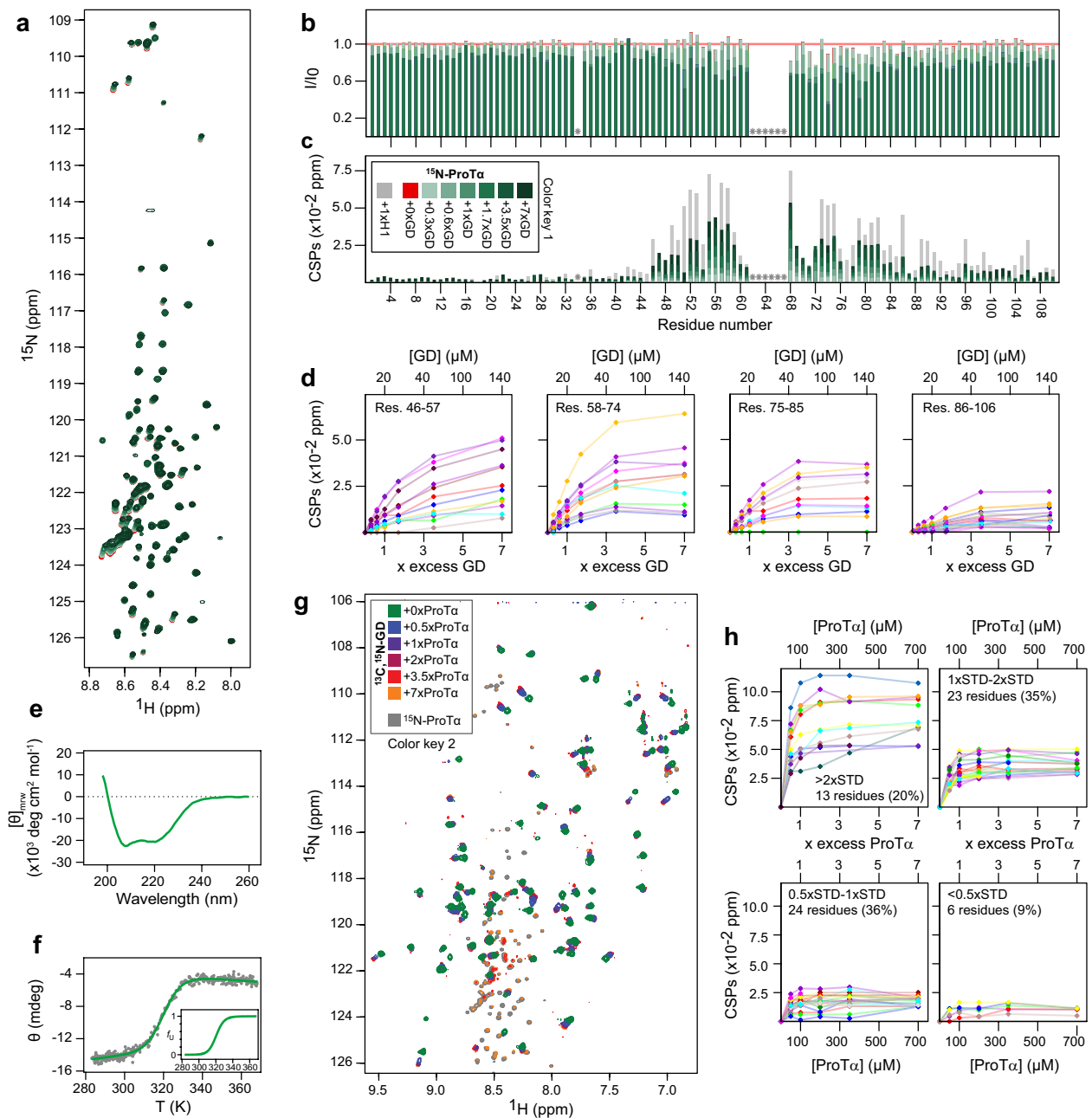
Extended Data Figure 1 | Titrations of ProTα and H1 globular domain.
a, Titration of ^{15}N-ProTα with zero to sevenfold molar addition of the H1 globular domain followed by ^{1}H–^{15}N HSQC spectra; $n = 2$ repeats of this measurement yielded consistent results. **b**, Peak intensity ratios for assigned residues of ProTα relative to the free state induced by zero to 1.7-fold molar addition of the H1 globular domain ($n = 2$). **c**, Weighted backbone CSPs per residue of ProTα induced by zero to sevenfold molar addition of the H1 globular domain ($n = 2$). For comparison, CSPs of ProTα with equimolar addition of H1 are shown in grey ($n = 5$). In **a**–**c**, 'colour key 1' applies; grey stars, prolines and unassigned residues. **d**, ProTα CSPs plotted against concentration and times excess of the H1 globular domain relative to the free state for residues 46–106 upon zero to sevenfold molar addition of the H1 globular domain. Curves corresponding to individual residues are shown in different colours for clarity. **e**, Far-UV circular dichroism spectrum of the H1 globular domain.
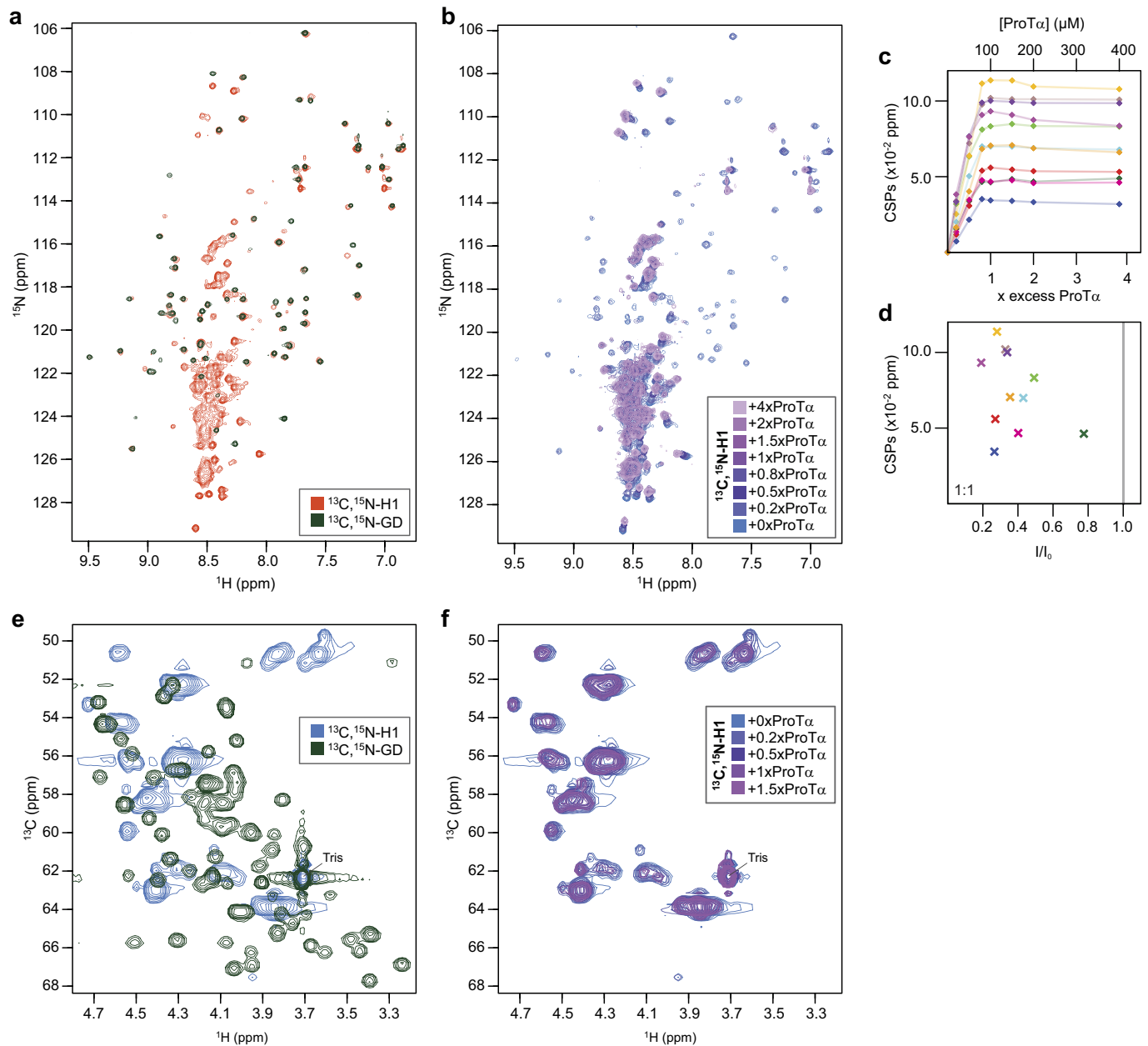
f, Thermal denaturation of the H1 globular domain followed by the change in ellipticity at 222 nm ($T_m = 320.5 \pm 0.3$ K, $\Delta H_m = -44 \pm 2$ kcal mol^{-1}). Inset in **f** shows fraction of unfolded H1 globular domain (f_u) as a function of temperature. **g**, Titration of 100 μM ^{13}C–^{15}N- H1 globular domain with zero to sevenfold molar addition of ProTα followed by ^{1}H–^{15}N HSQC spectra. Peak intensities gradually decrease during the titration. At 3.5- and 7-fold molar excess ProTα, natural abundance peaks of free ProTα appear. ^{1}H–^{15}N HSQC spectrum of ^{15}N-ProTα is shown in grey for comparison. **h**, Weighted backbone CSPs of the H1 globular domain plotted against concentration and times excess of ProTα relative to the free state on zero to sevenfold molar addition of ProTα. A total of 66 (unassigned) amide backbone peaks were followed and grouped according to the standard deviation (STD) of the CSPs (1 s.d. = 0.0254 p.p.m.). Of these, 55% had CSPs larger than 1 STD.
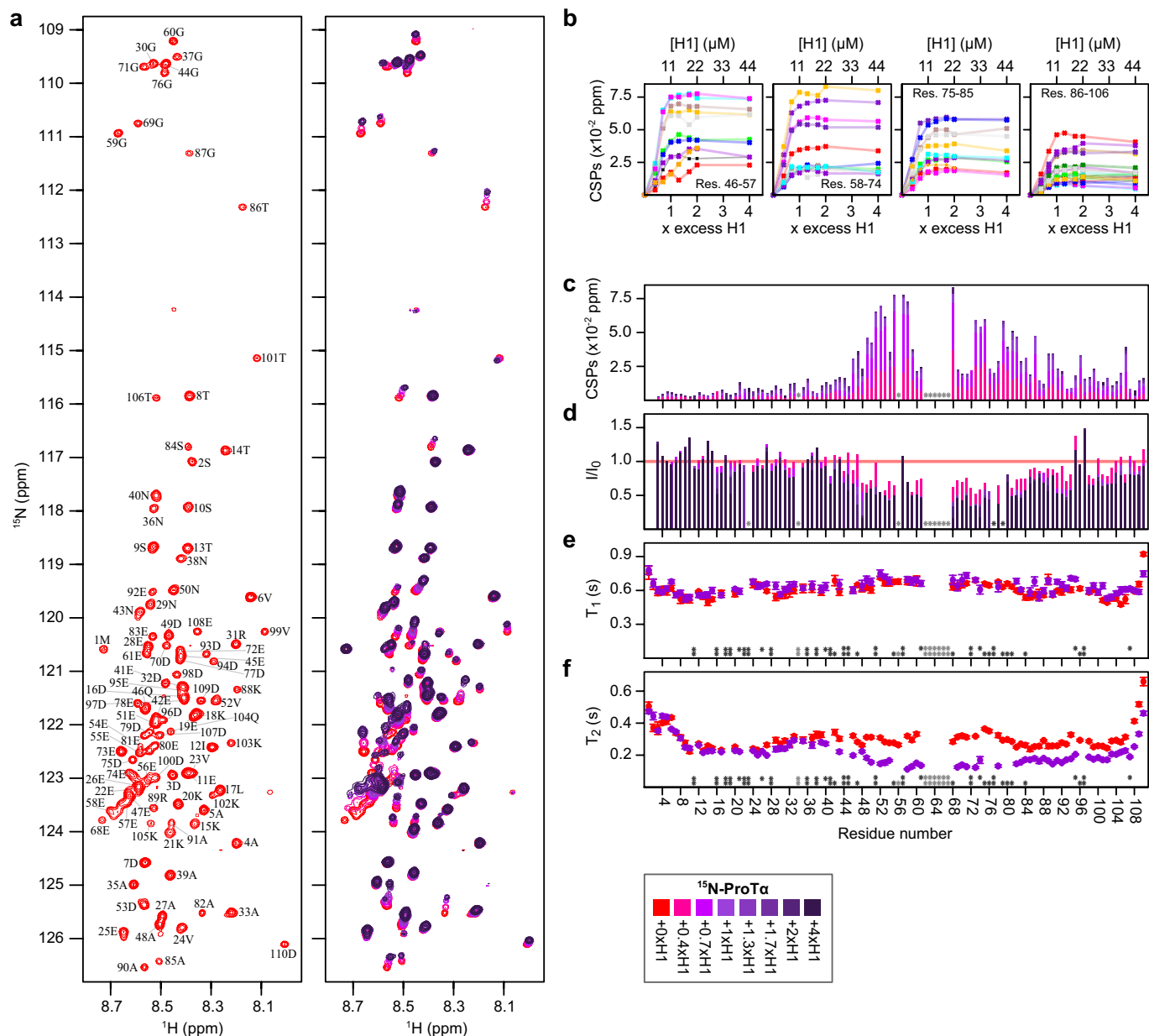
Extended Data Figure 2 | Titration of ^{15}N-ProTα with H1. a, ^{1}H–^{15}N HSQC spectrum of 11 μM free ^{15}N-ProTα with assigned residues labelled (left) and titrated with zero to fourfold molar addition of H1 (right); $n = 5$ individual repeats of this measurement yielded consistent results. **b**, Weighted backbone CSPs of ProTα (residues 46–106) relative to the free state on zero to fourfold molar addition of H1, plotted against concentration and times excess of H1. Curves corresponding to individual residues are shown in different colours for clarity. **c, d**, CSPs (**c**) and peak intensity ratios (**d**) for assigned residues of ProTα induced by

zero to fourfold molar addition of H1 (bar colours correspond to key); $n = 5$ for both. **e**, T_1 ^{15}N relaxation times of free (red) and H1-bound (purple) ^{15}N-ProTα. $\langle T_1 \rangle = 610$ ms (free) and 636 ms (complex); $n = 2$ individual repeats of this measurement yielded consistent results. **f**, T_2 ^{15}N relaxation times of free (red) and H1-bound (purple) ^{15}N-ProTα. $\langle T_2 \rangle = 302$ ms (free) and 217 ms (complex). In **c–f**, light grey stars, prolines and unassigned residues; dark grey stars, overlap and/or insufficient data quality. Circles in **e** and **f** are mean values from $n = 3$ consecutive data acquisitions on the same samples, errors are s.d.

**Extended Data Figure 3 | Titration of $^{13}C-^{15}N$-H1 with ProTα.**
a, $^1H-^{15}N$ HSQC spectra of the free $^{13}C-^{15}N$-H1 globular domain (dark green) and free $^{13}C-^{15}N$-H1 (orange). The majority of the amide peaks of the H1 globular domain overlap with the more dispersed peaks from H1, indicating the similarity in structure of the H1 globular domain in isolation and within H1. **b**, Titration followed by $^1H-^{15}N$ HSQC spectra of $^{13}C-^{15}N$-H1 with zero to fourfold molar addition of ProTα. Data acquired on His_6-tagged H1; $n = 2$ individual repeats of this measurement yielded consistent results. **c**, CSPs relative to free H1 of 11 traceable H1 amide backbone peaks from the intrinsically disordered region (based on overlay with $^1H-^{15}N$ HSQC spectra of the H1 globular domain (in **a**)) on zero to fourfold molar addition of ProTα plotted against concentration and times

excess. Curves corresponding to individual residues are shown in different colours for clarity. **d**, CSPs plotted against peak intensity ratios relative to the free state of H1 of the 11 H1 amides at $1\times$ excess of ProTα. Colours correspond to those in **c. e**, Overlay of the Cα–Hα region from $^1H-^{13}C$ HSQC spectra of free $^{13}C-^{15}N$-H1 (blue) and the $^{13}C-^{15}N$-H1 globular domain (green). The H1 $^1H-^{13}C$ HSQC spectrum is dominated by intense clusters of peaks not present in the H1 globular domain spectrum, consistent with the large fraction of residue repeats in the H1 disordered regions. **f**, Cα–Hα region of $^{13}C-^{15}N$-H1 on titration with ProTα. The lack of detectable changes in Cα–Hα resonances is consistent with the absence of secondary structure induction in the disordered regions of H1 on binding.
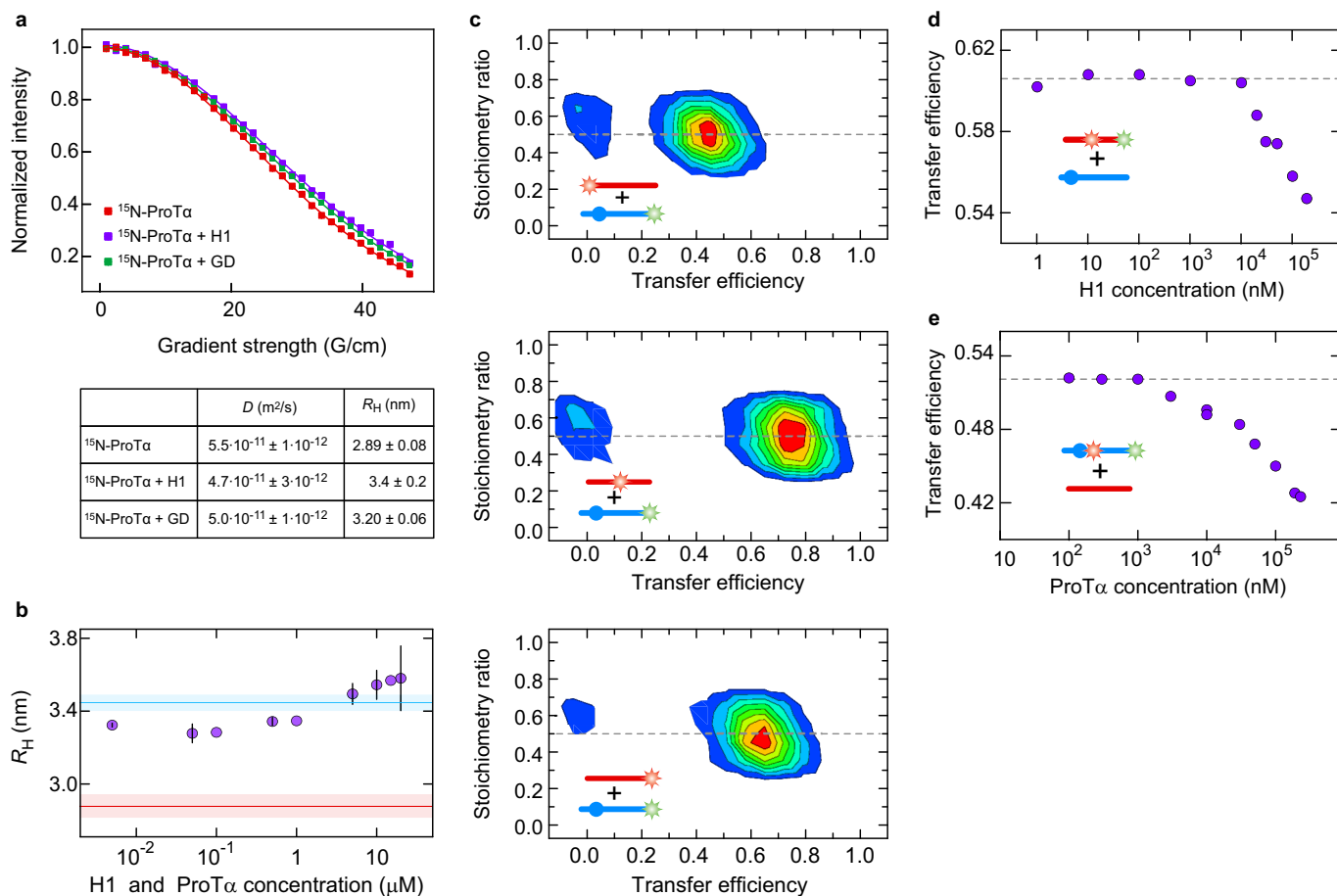
Extended Data Figure 4 | Hydrodynamic radii and stoichiometry of the H1–ProTα complex. a, R_H of free and bound ^{15}N-ProTα (100 μM) determined with pulsed-field gradient NMR at 283 K. The signal decays of free ^{15}N-ProTα (red), with H1 at a 1:1 molar ratio (purple) and with the H1 globular domain at a 1:7 molar ratio (green) as a function of gradient strength, together with corresponding fits and a table of the diffusion coefficients and resulting R_H values. **b**, R_H measured by two-focus fluorescence correlation spectroscopy at 295 K. Lines show the mean R_H from $n = 2$ independent measurements of $H1_0$ (blue) and $ProTα_2$ (red) labelled with Alexa 594 in the absence of the binding partner. Symbols represent the mean R_H from $n = 2$ independent measurements of labelled ProTα (5 nM) in the presence of equimolar concentrations

of unlabelled ProTα and unlabelled H1. Error bars or shaded bands, s.d. **c**, Stoichiometry ratio[71] versus transfer efficiency plots from intermolecular single-molecule FRET measurements of $ProTα_2$ + $H1_{194}$ (top), $ProTα_{56}$ + $H1_{194}$ (middle), and $ProTα_{110}$ + $H1_{194}$ (bottom); pictograms in panels indicate labelling locations. A stoichiometry ratio of 0.5 indicates a 1:1 complex. The peaks at $E \approx 0$ originate from molecules or complexes that lack an acceptor dye and remain after filtering for donor-only fluorescence bursts based on pulsed-interleaved excitation. **d**, **e**, Transfer efficiency changes at a large excess of unlabelled binding partner for FRET-labelled $ProTα_{56/110}$ (**d**) and $H1_{104/194}$ (**e**). See Methods for further information on statistics.
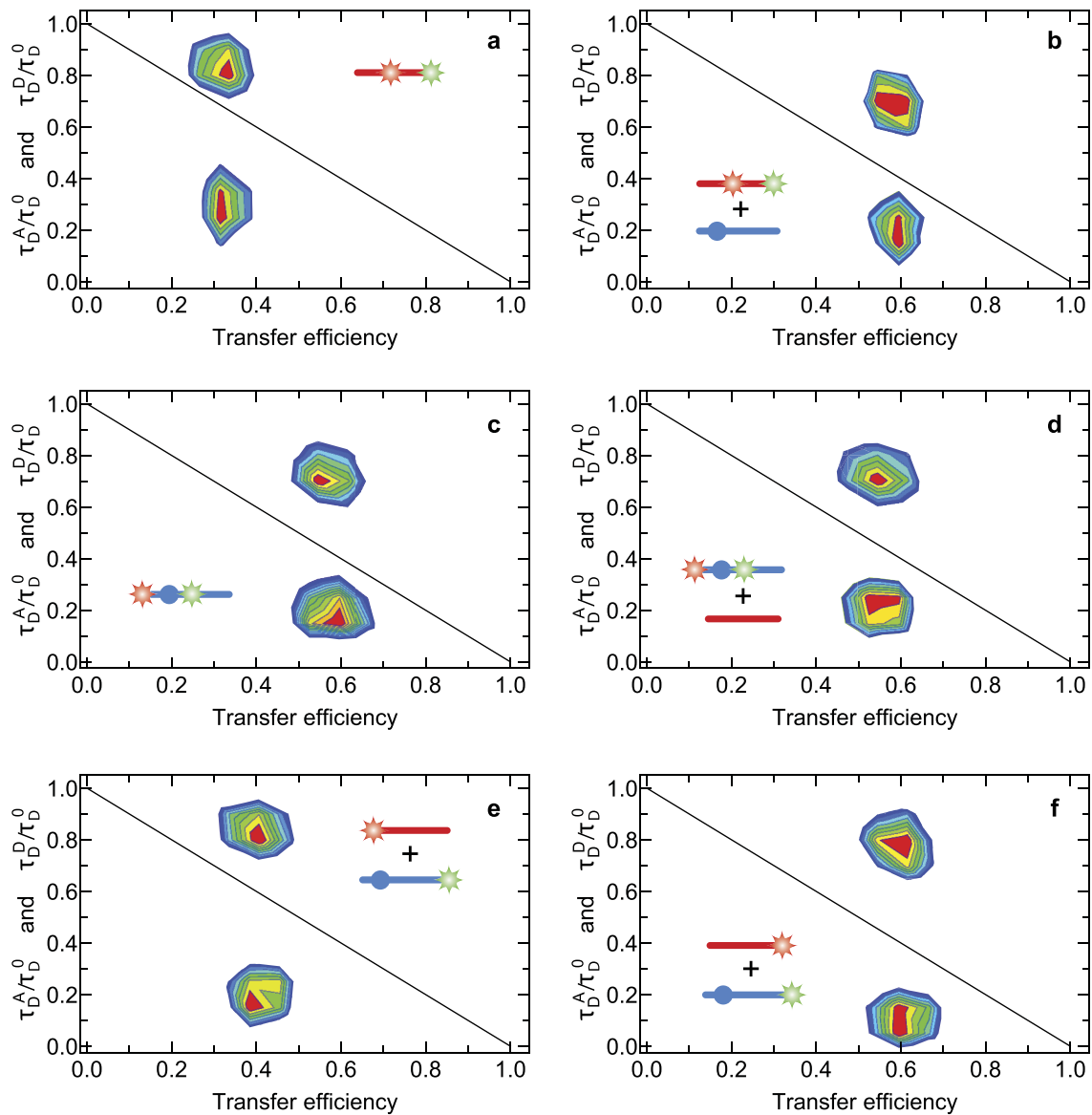
Extended Data Figure 5 | Fluorescence lifetime analysis. a–f, Plots of the fluorescence lifetimes of Alexa 488 donor (τ_D^D) and Alexa 594 acceptor (τ_D^A) normalized by the intrinsic donor lifetime (τ_D^0) versus the ratiometric transfer efficiency E (calculated from the number of donor and acceptor photon counts), as a diagnostic for the presence of a broad distance distribution rapidly sampled during the time of a fluorescence burst[28,33,34]. If fluctuations in transfer efficiency occur on a timescale between the donor fluorescence lifetime (∼4 ns) and the burst duration (∼1 ms), the normalized donor lifetimes cluster above—and the acceptor lifetimes

below—the solid diagonal line expected for a single fixed distance, as previously observed for intrinsically disordered proteins[34,72]. The large deviation from the diagonal observed for both unbound and bound ProTα and H1 supports the presence of broad and rapidly sampled distance distributions. **a,** ProTα$_{56/110}$; **b,** ProTα$_{56/110}$ + unlabelled H1; **c,** H1$_{0/113}$; **d,** H1$_{0/113}$ + unlabelled ProTα; **e,** ProTα$_2$ + H1$_{194}$; and **f,** ProTα$_{110}$ + H1$_{194}$. All variants labelled with Alexa 488 (green) and/or Alexa 594 (red) as indicated by the pictograms in the figure panels.
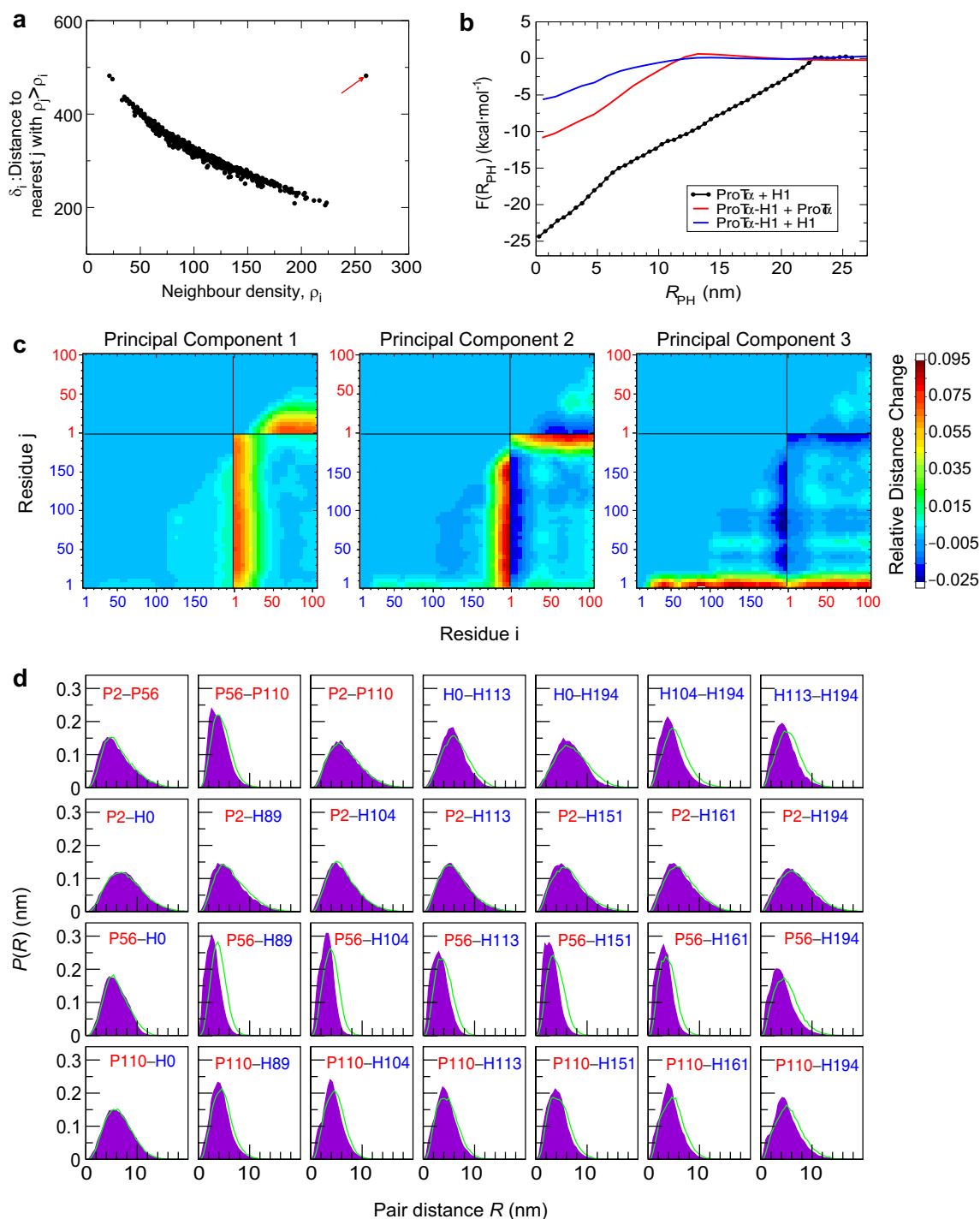
Extended Data Figure 6 | Simulation results. a, Decision graph using the Rodriguez–Laio clustering algorithm[69], showing only a single density maximum distant from other density maxima (that is, a single distinct cluster). **b**, Free energy of association from simulation for ProTα and H1 along the distance between their centres of mass, R_{PH}, yielding a K_d of 7 fM (black curve). Blue and red curves are the free energies for addition of a second H1 or a second ProTα, respectively, to an existing H1–ProTα complex. **c**, Principal component vectors shown as contact maps. Colours indicate the increase or decrease in each pair distance for that principal component, relative to the other distances. ProTα and H1 residue numbers are indicated in red and blue, respectively. Each principal component describes a feature of the chain arrangement: principal component 1, for example, captures the presence or absence of interactions between the ProTα N terminus and H1. **d**, Intramolecular (top row) and intermolecular (bottom three rows) distributions of distances corresponding to FRET labelling sites, within the H1–ProTα complex. P and H numbers refer to ProTα and H1 residues, respectively. Filled distributions, simulations without explicit chromophores; green lines, simulations with explicit chromophores.
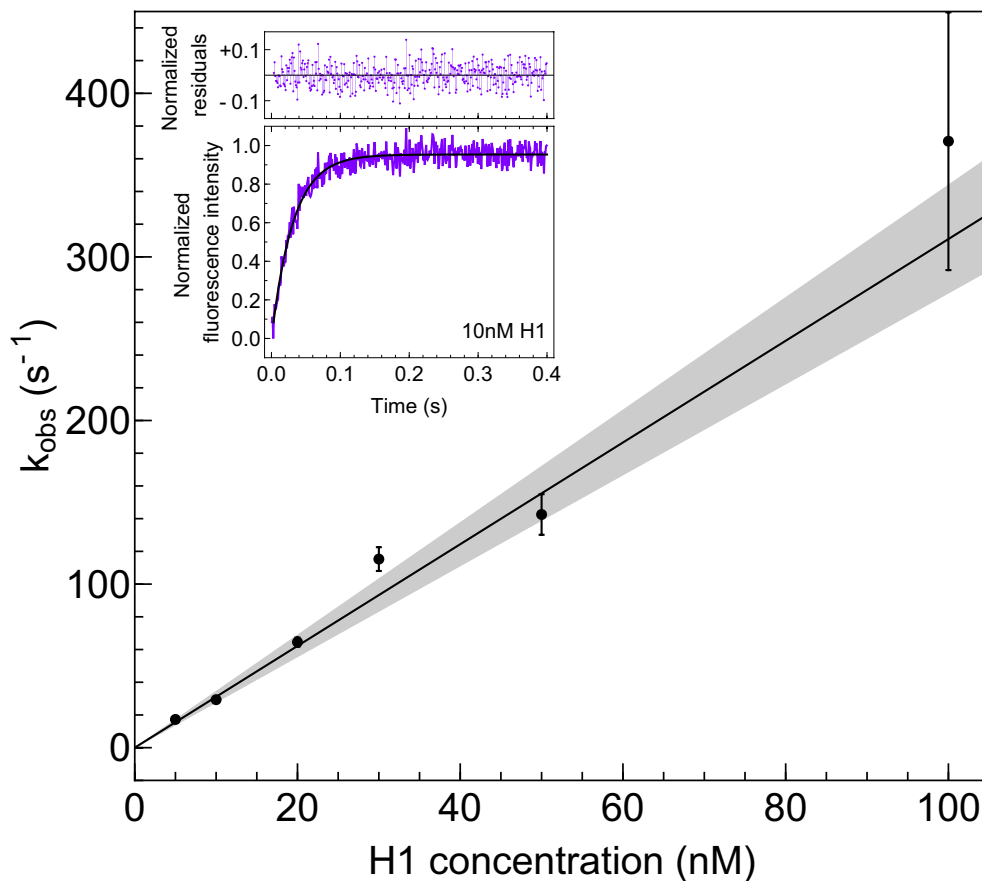
Extended Data Figure 7 | Kinetics of H1–ProTα binding measured by stopped flow. FRET-labelled ProTα$_{56/110}$ was mixed rapidly with unlabelled H1 in TBS buffer, and the resulting increase in acceptor fluorescence was monitored. Inset, example at 10 nM H1 with single-exponential fit and residuals shown above (see Methods for details). Decay rates were obtained from single-exponential fits, with an instrument dead time of 3 ms. Standard errors for each H1 concentration were obtained using bootstrapping. The observed rate, k_{obs}, is shown as a function of H1 concentration (c_{H1}); for H1 concentrations between 10 and 100 nM—for which pseudo-first order conditions apply (ProTα concentration after mixing was 2 nM)—the observed rates were fit with $k_{obs} = k_{on}c_{H1} + k_{off} = k_{on}c_{H1} + k_{on}K_d$, using the independently determined K_d of 2.1 pM (Extended Data Table 2). The fit yields a bimolecular association rate coefficient of $k_{on} = 3.1 \pm 0.1 \times 10^9 \, \text{M}^{-1} \, \text{s}^{-1}$ and an apparent dissociation rate coefficient of $k_{off} = 6.5 \pm 3.1 \times 10^{-3} \, \text{s}^{-1}$. The grey area represents the 95% confidence band.
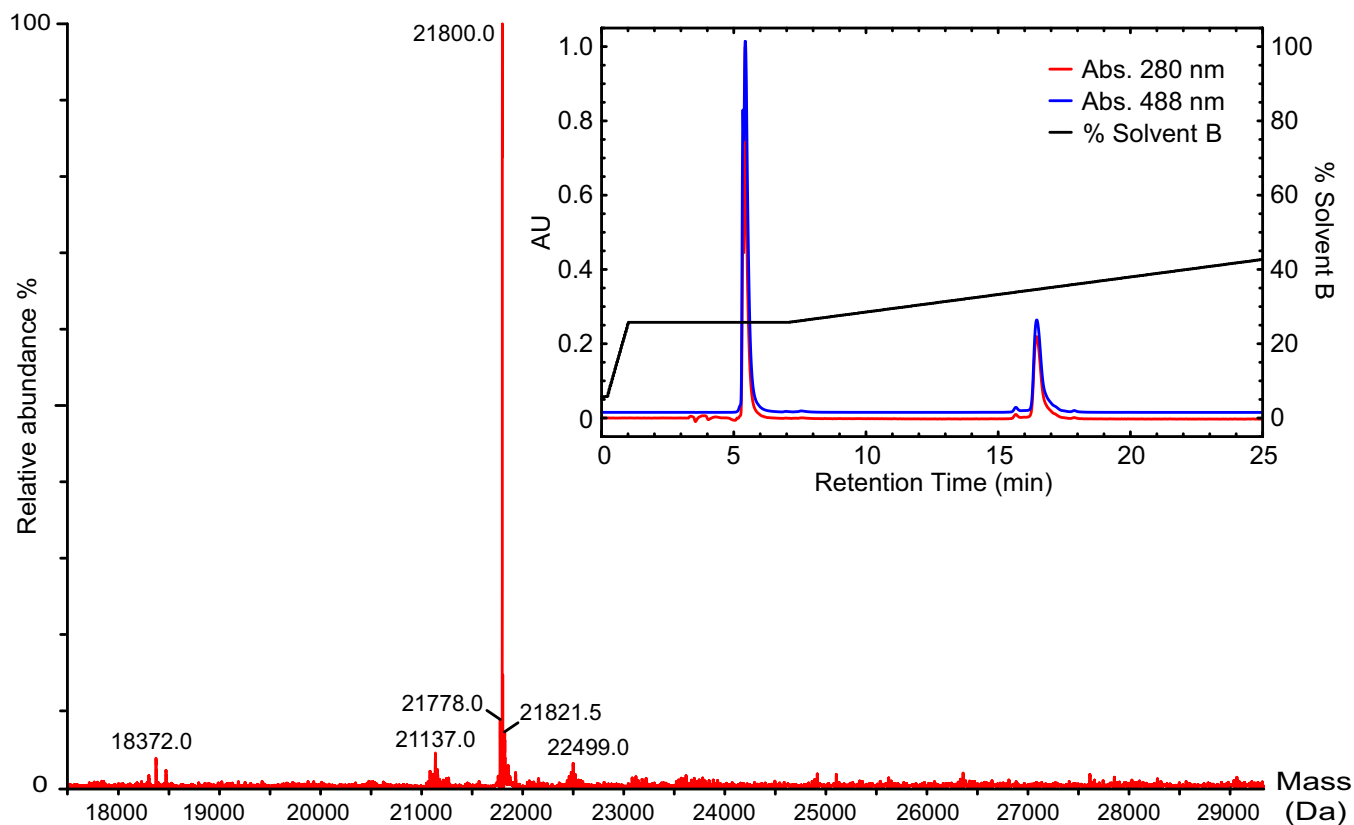
Extended Data Figure 8 | Example of the quality of the H1 preparation. Electrospray ionization mass spectrum of H1(T161C) labelled with Alexa 488 (calculated mass 21,800 Da) and preparative reversed-phase HPLC (Vydac C4) chromatogram (inset) showing absorption at 280 nm (red) and 488 nm (blue) and the elution gradient from solvent A (H_2O + 0.1% TFA) to solvent B (100% acetonitrile) (black), illustrating the high purity of the sample. The peak at approximately 5.5 min corresponds to free Alexa 488, and the peak at approximately 16.8 min to H1(T161C) labelled with Alexa 488.

Extended Data Table 1 | Sequences of protein constructs and fluorescently labelled variants of H1 and ProTα

H1 (+53)	0 23
	CTENSTSAPAAKPKRAKASKKST DHPKYSDMIVAAIQAEKNRAGSSRQSIQKYIKSHYKVGENADSQI
	          89 96 104 113
	KLSIKRLVTTGVLKQTKGVGA**S**GSFRLAKSDEPKKS**V**AFKKTKKE**I**KKVATPKKASKPKKAASKAPTK
	      151 161 194
	KPKATPVKKAKKKLA**A**TPKKAKKPK**T**VKAKPVKASKPKKAKPVKPKAKSSAKRAGKKK**G**GPR

H1 †CTR (+39)	103 SVAFKKTKKEIKKVATPKKASKPKKAASKAPTKKPKATPVKKAKKKLAATPKKAKKPKTVKAKPVKASK 193 PKKAKPVKPKAKSSAKRAGKKKGGPR
H1 ‡NTR (+18)	0 G**C**TENSTSAPAAKPKRAKASKKSTDHPKYSDMIVAAIQAEKNRAGSSRQSIQKYIKSHYKVGENADSQI 113 KLSIKRLVTTGVLKQTKGVGASGSFRLAKSDEPKKSVAFKKTKKE**I**
GD (+9)	23 DHPKYSDMIVAAIQAEKNRAGSSRQSIQKYIKSHYKVGENADSQIKLSIKRLVTTGVLKQTKGVGASGS 96 FRLAK
ProTα (−44)	2 56 GPS**D**AAVDTSSEITTKDLKEKKEVVEEAENGRDAPANGNAENEENGEQEADNEVDEE**E**EGGEEE 110 EEEEEGDGEEEDGDEDEEAESATGKRAAEDDEDDDVDTKKQKTDED**D**

H1				ProTα				
Singly labeled		**Doubly labeled**		**Singly labeled**		**Doubly labeled**		
Alexa 488	Alexa 594	Alexa 488/ Alexa 594 (R₀=5.4 nm) §	Cy3B/ Abberior* 635 (R₀ = 5.9 nm) §	Alexa 488	Alexa 594	Alexa 488/ Alexa 594 (R₀ = 5.4 nm) §	Cy3B/ Abberior* 635 (R₀ = 5.9 nm) §	Atto 550/ Atto 647N (R₀ = 6.6 nm) §
C0, S89C, V104C, I113C, A151C, T161C, G194C	C0, S89C, G194C	C0/I113C, C0/G194C, V104C/G194C, I113C/G194C	V104C/G194C	D110C	D2C, E56C, D110C	D2C/E56C, E56C/D110C, D2C/D110C	E56C/D110C	E56C/D110C

Top, sequences of wild-type H1, H1 fragments used, and wild-type ProTα. Residues in bold yellow are positions mutated to Cys for fluorophore conjugation. Residues in red are remain after proteolytic cleavage of the HisTag with thrombin (Gly-Gly-Pro-Arg or Gly-Cys) or HRV-3C (Gly-Pro). Note that the wild-type sequence of H1 starts with Thr1 and ends with Lys193; the preceding Cys residue (0) was added for labelling. The underlined part of the H1 sequence indicates the globular domain (GD), identified based on a sequence alignment with the *Gallus gallus* homologue[20] (RCSB Protein Data Bank access code: 1HST, 82% sequence identity). Surface-exposed residues in GD (as shown in Fig. 1a and 4b) are shaded in light blue. The net charge of each variant is indicated in parentheses
Bottom, labelled variants of H1 and ProTα.
†CTR, C-terminal disordered region.
‡NTR, N-terminal disordered region including H1 globular domain.
§Förster radius of the corresponding dye pair.

Extended Data Table 2 | Binding affinities, molecular dimensions and reconfiguration times of fluorescently labelled H1 and ProTα

Affinity of ProTα 56/110 Alexa 488/594 for full-length H1 and fragments in TBS				Affinities in TBS 205 mM		Intramolecular transfer efficiencies and distances in TBS 205 mM				
Ionic strength (mM)	K_d (nM)	H1 fragm.	K_d (TBS 165 mM)	ProTα	$\|K_d$ (nM)	ProTα	E unbound	E bound	$R_{unbound}$ (nm)	R_{bound} (nm)
†165	$(2.1^{+1.1}_{-0.8})\cdot 10^{-3}$	‡CTR	40^{+6}_{-4} pM	D2C/D110C Alexa 488/594	2.0 ± 0.13	E56C/D110C Alexa 488/594	0.36	0.54	$7.5^{+1.0}_{-0.3}$	$5.8^{+0.7}_{-0.1}$
180	$(37 ± 5)\cdot 10^{-3}$					E56C/D110C Cy3B/Abb.*635	0.41	0.56	$7.6^{+0.6}_{-0.5}$	6.2± 0.4
205	1.0 ± 0.1	‡NTR	173^{+29}_{-28} nM	E56C/D110C Cy3B/Abb.*635	1.0 ± 0.10	E56C/D110C Atto 550/647N	0.45	0.59	$8.1^{+0.6}_{-0.5}$	$6.7^{+0.5}_{-0.4}$
240	25 ± 3			E56C/D110C Atto 550/647N	3.1 ± 0.20	D2C/D110C Alexa 488/594	0.18	0.33	$10.6^{+1.6}_{-1.1}$	$7.9^{+0.7}_{-0.6}$
290	$(2.3 ± 1.5)\cdot 10^2$	‡GD	$1.9^{+0.3}_{-0.3}$ μM§	H1		H1	In TBS 165 mM			
						V104C/G194C Alexa 488/594	0.15	0.53	$11.4^{+2.1}_{-1.4}$	5.9± 0.4
330	$(1.4 ± 0.4)\cdot 10^3$			V104C/G194C Alexa 488/594	3.5 ± 0.23	V104C/G194C Cy3B/Abb.*635	0.18	0.52	$11.6^{+1.8}_{-1.2}$	$6.5^{+0.5}_{-0.4}$
340	$(4.0 ± 1.8)\cdot 10^3$					I113C/G194C Alexa 488/594	0.23	0.58	$9.5^{+1.1}_{-0.9}$	5.6± 0.4

ProTα A-594 / H1 A-488	τ_r (ns)			Labeled Protein (Alexa 488/ Alexa 594)	τ_r (ns)	
	D2C	E56C	D110C		unbound	bound
H1 C0	180^{+19}_{-16}	191^{+22}_{-19}	169^{+19}_{-16}	ProTα E56C/D110C	29±2	102^{+3}_{-2}
H1 I113C	121^{+13}_{-11}					
H1 A151C	124^{+13}_{-12}	98^{+16}_{-2}		ProTα D2C/E56C	33±2	66±2
H1 G194C	156^{+16}_{-14}		142^{+19}_{-15}			
A-488 / A-594				ProTα D2C/D110C	78^{+15}_{-9}	133^{+10}_{-7}
H1 G194C			120^{+13}_{-12}	H1 I113C/G194C	118^{+24}_{-14}	143^{+5}_{-4}

Top left, affinities of labelled ProTα for unlabelled H1 at different ionic strengths (IS), and for H1 fragments at 165 mM IS. Uncertainties for the IS dependence are standard errors estimated from independent titrations (see Statistics in Methods). Top centre, binding affinities of ProTα and H1 labelled with different dye pairs for the respective unlabelled partner. Top right, transfer efficiencies and average distances of ProTα and H1 labelled with different dye pairs in the unbound ($R_{unbound}$) and bound state (R_{bound}) (with the respective unlabelled partner; Abb.*635 = Abberior Star 635). Uncertainties in distance are based on an estimated systematic error of ± 0.05 in the transfer efficiency from instrument calibration for the different dye pairs. Bottom left, intermolecular reconfiguration times for the complex of donor-labelled H1 (Alexa 488) and acceptor-labelled ProTα (Alexa 594) and vice versa. Bottom right, reconfiguration times of doubly labelled ProTα and H1 (unbound and bound). Uncertainties estimated by propagating the systematic error on the transfer efficiency (± 0.05).
†Uncertainty at 165 mM (see Methods for details).
‡For H1 fragment sequences, see Table 1).
§Apparent K_d from fraction of all bound species.
||Uncertainties based on dilution errors.

nature research

Corresponding author(s): Benjamin Schuler, Alessandro Borgia, Robert B. Best, Birthe B. Kragelund

☐ Initial submission   ☐ Revised version   ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

▶ Experimental design

1. **Sample size**

   Describe how sample size was determined.

   > Sample sizes of 800-5,000,000 molecules per single-molecule measurement were used and found to satisfactorily represent the conformational distribution in the ensemble.

2. **Data exclusions**

   Describe any data exclusions.

   > No data were excluded from the analysis

3. **Replication**

   Describe whether the experimental findings were reliably reproduced.

   > All key experimental findings were successfully reproduced in multiple measurements performed on different days, with freshly prepared buffer and from protein aliquots from the same or different protein preparations.

4. **Randomization**

   Describe how samples/organisms/participants were allocated into experimental groups.

   > Samples were not randomized.

5. **Blinding**

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   > Experiments were not blinded.

   Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. **Statistical parameters**

   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

   See the web collection on statistics for biologists for further resources and guidance.

▶ Software

7. Software

Describe the software used to analyze the data in this study.

Wolfram Mathematica 11; Modified Gromacs 4.0.5; Python; CcpNmr Analysis; NMRPipe, Topspin, Dynamics Center (Bruker); SymphoTime (PicoQuant).

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

▶ Materials and reagents

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

Histone H1.0 wt was purchased from New England Biolabs. Chemicals, enzymes and other consumables were standard and sources from common suppliers.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used.

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No eukaryotic cell lines were used.

▶ Animals and human research participants

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

No animals were used.

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

No human participants were involved.

nature research

Corresponding author(s): Benjamin Schuler, Alessandro Borgia, Robert B. Best, Birthe B. Kragelund

☐ Initial submission ☒ Revised version ☐ Final submission

Editorial Policy Checklist

This form is used to ensure compliance with Nature Research editorial policies related to research ethics and reproducibility in the life sciences. For further information, please see our Authors & Referees site. All questions on the form must be answered.

▸ Data availability

Policy information about availability of data

Data availability statement

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated source data
- A description of any restrictions on data availability

☒ A full data availability statement is included in the manuscript.

Required accession codes

Data deposition is mandated for certain types of data.

Confirm that all relevant data have been deposited into a public repository and that all accession codes are provided.

☐ Accession codes will be available before publication ☐ No data with mandated deposition ☒ All relevant accession codes are provided

▸ Data presentation

Image integrity

☒ Confirm that all images comply with our image integrity policy.

Unprocessed data must be provided upon request. Please double-check figure assembly to ensure that all panels are accurate (e.g. all labels are correct, no inadvertent duplications have occurred during preparation, etc.).

Data distribution

Data should be presented in a format that shows data distribution (dot-plots or box-and-whisker plots), with all box-plot elements (e.g. center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers) defined. If bar graphs are used, the corresponding dot plots must be overlaid.

☒ Confirm that all data presentation meets these requirements.

☒ Confirm that in all cases where the number of data points is <10, individual data points are shown.

▸ Structural data

Policy information about special considerations for specific types of data

☒ If this study did not involve data of these types, check here and skip the rest of this section.

Electron microscopy

☐ For all electron microscopy work, confirm that you have deposited any density maps and coordinate data in EMDB.

Macromolecular structures

☐ For all macromolecular structures studied, confirm that you have provided an official validation report from wwPDB.

▸ Code availability

Policy information about availability of computer code

Code availability statement

For all studies using custom code, the Methods section must include a statement under the heading "Code availability" describing how readers can access the code, including any access restrictions.

☒ A full code availability statement is included in the manuscript ☐ No custom code used

▶ Research animals

Policy information about studies involving animals; follow the ARRIVE guidelines for reporting animal research

☒ If this study did not use animals and/or animal-derived materials for which ethical approval is required, check here and skip the rest of this section.

Ethical compliance

☐ Confirm that you have complied with all relevant ethical regulations and that a statement affirming this is included in the manuscript.

Ethics committee

☐ Confirm that you have stated the name(s) of the board and institution that approved the study protocol in the manuscript.

▶ Human research participants

Policy information about studies involving human research participants

☒ If this study did not involve any human research participants, check here and skip the rest of this section.

Ethical compliance

☐ Confirm that you have complied with all relevant ethical regulations and that a statement affirming this is included in the manuscript.

Ethics committee

☐ Confirm that you have stated the name(s) of the board and institution that approved the study protocol in the manuscript.

Informed consent

☐ Confirm that informed consent was obtained from all participants.

Identifiable images

For publication of identifiable images of research participants, confirm that consent to publish was obtained and is noted in the Methods.

Authors must ensure that consent meets the conditions set out in the Nature Research participant release form.

☐ Yes   ☐ No identifiable images of human research participants

▶ Clinical studies

Policy information about clinical studies

☒ If this study was not a clinical trial, check here and skip the rest of this section.

Clinical trial registration

☐ Confirm that you have provided the trial registration number from ClinicalTrials.gov or an equivalent agency in the manuscript.

Phase 2 and 3 randomized controlled trials

Confirm that you have provided the CONSORT checklist with your submission.

☐ Yes   ☐ No   ☐ Not a phase 2/3 randomized controlled trial

Tumor marker prognostic studies

Did you follow the REMARK reporting guidelines?

☐ Yes   ☐ No   ☐ Not a tumor marker prognostic study

▶ Methods reporting

Nature Research wishes to improve the reproducibility of the work we publish. As part of this effort, all life science manuscripts require a reporting summary; certain types of research require specialized modules in addition to this form.

☒ Confirm that you have provided a complete and accurate reporting summary.

n/a | Confirmed

☒ ☐ For MRI studies, confirm that you have completed the additional MRI module.

☒ ☐ For flow cytometry studies, confirm that you have completed the additional flow cytometry module.

☒ ☐ For ChIP-seq studies, confirm that you have completed the additional ChIP-seq module.

I certify that all the above information is complete and correct.

Typed signature Benjamin Schuler

Date 20 December , 2017