Article

Disordered regions of human eIF4B orchestrate a dynamic self-association landscape

Bikash Chandra Swain [1,5], Pascale Sarkis [1,5], Vanessa Ung [2,5], Sabrina Rousseau [1], Laurent Fernandez [1], Ani Meltonyan[1], V. Esperance Aho [1,4], Davide Mercadante [2], Cameron D. Mackereth [3] ✉ & Mikayel Aznauryan [1] ✉

Eukaryotic translation initiation factor eIF4B is required for efficient cap-dependent translation, it is overexpressed in cancer cells, and may influence stress granule formation. Due to the high degree of intrinsic disorder, eIF4B is rarely observed in cryo-EM structures of translation complexes and only ever by its single structured RNA recognition motif domain, leaving the molecular details of its large intrinsically disordered region (IDR) unknown. By integrating experiments and simulations we demonstrate that eIF4B IDR orchestrates and fine-tunes an intricate transition from monomers to a condensed phase, in which large-size dynamic oligomers form before mesoscopic phase separation. Single-molecule spectroscopy combined with molecular simulations enabled us to characterize the conformational ensembles and underlying intra- and intermolecular dynamics across the oligomerization transition. The observed sensitivity to ionic strength and molecular crowding in the self-association landscape suggests potential regulation of eIF4B nanoscopic and mesoscopic behaviors such as driven by protein modifications, binding partners or changes to the cellular environment.

As their names suggest, eukaryotic translation initiation factors (or eIFs) play an important role in the first stage of protein translation. Their functions span from facilitating 5' mRNA 7-methylguanosine cap recognition to orchestrating mRNA scanning and start codon recognition in concert with the 40S small ribosomal subunit, to ensure the effective start of protein synthesis[1–8]. A critical contribution to the mechanistic understanding of eukaryotic translation initiation resulted from a burst of high-resolution structural characterizations of various translation initiation complexes and sub-complexes[6,8–12]. However, due to the nature of cryo-EM and crystallography approaches, these experimental advances nevertheless omit the characterization of the constituent proteins and protein regions that are intrinsically disordered.

Specifically within this latter category of proteins is eIF4B, the major portion of which is predicted to be intrinsically disordered (Fig. 1a, b). eIF4B is essential for optimal unwinding of 5'-proximal mRNA secondary structures as it significantly enhances the activity of RNA helicase eIF4A and stimulates its processivity[3,13–18]. In addition to the specific activation of eIF4A, eIF4B is generally required for efficient control of cell survival and proliferation[19–21] given that it selectively regulates a set of proliferative and prosurvival mRNAs that encompass long structured 5' untranslated regions[19,22]. In this context, eIF4B was shown to be overexpressed in certain cancer cells and the reduction of eIF4B expression is sufficient to decrease synthesis of proteins associated with enhanced tumor cell survival[23]. Thus, eIF4B is considered as

[1]University of Bordeaux, Inserm, CNRS, ARNA Laboratory, U1212, UMR 5320, Institut Européen de Chimie et Biologie, F-33600 Pessac, France. [2]School of Chemical Sciences, University of Auckland, Auckland, New Zealand. [3]University of Bordeaux, Inserm, CNRS, ARNA Laboratory, U1212, UMR 5320, F-33000 Bordeaux, France. [4]Present address: Institut de Biologie Structurale (IBS), UMR 5075, F-38044 Grenoble, France. [5]These authors contributed equally: Bikash Chandra Swain, Pascale Sarkis, Vanessa Ung. ✉e-mail: cameron.mackereth@inserm.fr; mikayel.aznauryan@inserm.fr
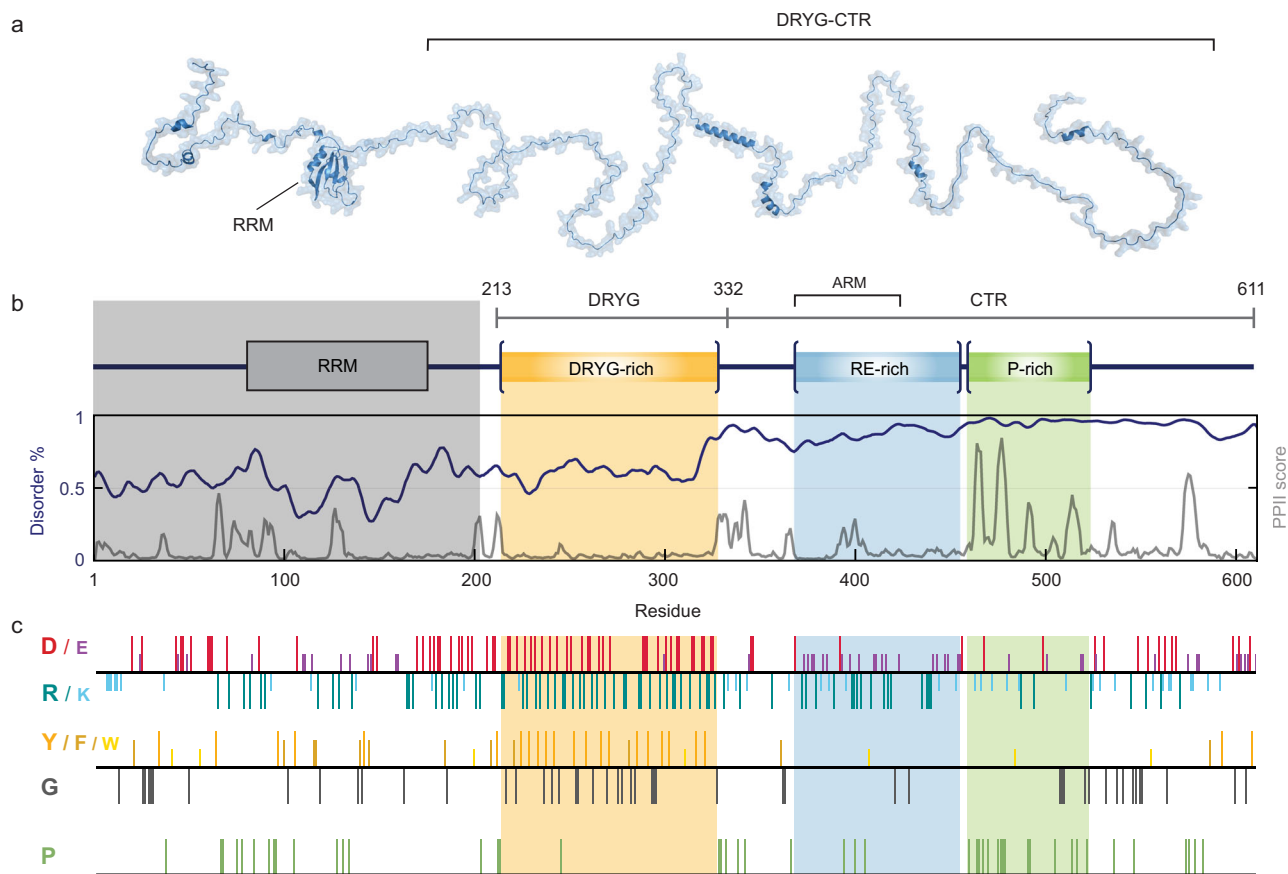
Fig. 1 | An overview of sequence and structural characteristics of eIF4B.
a Cartoon representation of human eIF4B based on the RRM domain structure (PDB: 2J76) and predicted IDR structural propensity from AlphaFold2. **b** Schematic domain representation of human eIF4B, together with a disorder prediction plot based on IUPred3[91] (dark blue) and the prediction of polyproline II secondary structure propensity (gray)[92]. The structured RRM domain and short disordered N-terminal tail are shadowed in gray. The boundaries of protein constructs used in this work are also indicated. **c** Amino acid distribution plots for negative (D and E, shown in red and purple, respectively), positive (R and K, shown in teal and light blue, respectively), aromatic (Y, F and W, shown in orange, brown and yellow, respectively), glycine (G, shown in black) and proline (P, shown in green) residues, particularly highlighting clustering of these residues within several subregions, i.e. DRYG-rich region (residues 214–327; highlighted in light orange), RE-rich region (residues 367–455; highlighted in light blue) and the P-rich region (residues 460–522; highlighted in light green). Source data are provided as a Source Data file.

a potential anti-cancer target[21]. eIF4B has also been reported to localize into stress granules[24,25], and the reduction of cellular eIF4B levels was shown to affect stress granule formation[26].

Human eIF4B consists of 611 residues and is composed of a folded RNA recognition motif (RRM) domain preceded by a short unstructured tail[27], and an approximately 400 residue intrinsically disordered region (IDR) downstream of the RRM (Fig. 1a, b). As the single folded domain present in eIF4B, the RRM domain has been previously studied at the atomic level[27]. Recently, we have also characterized the atomic details of the C-terminal region (CTR) of eIF4B, encompassing residues 333 to 611, and experimentally confirmed that this region is indeed highly intrinsically disordered[28]. Nonetheless, we could also identify a short helical segment that coincides with the previously mapped arginine-rich motif (ARM) implicated in RNA binding[17,29] (Fig. 1b).

Another notable segment of the IDR encompasses a defined low-complexity region enriched in aspartate, arginine, tyrosine and glycine residues (the DRYG-rich region, Fig. 1b, c)[30,31]. Initially this region consisted of residues 164–356[30], but later it was refined and mapped to residues 214–327[17,31], to precisely bracket the cluster of tyrosine and glycine residues (Fig. 1c). In some deletion variants, the DRYG-rich region was alternatively defined to extend further towards the C-terminus, covering the residues 213–340[29]. This DRYG-rich region was shown to facilitate the direct interaction between eIF4B and eIF3a, which is believed to mediate the association of the mRNA with 40S[31]. At

the same time, the deletion of the DRYG-rich region from the full-length eIF4B reduces the RNA binding affinity[29], either because this region has intrinsic RNA affinity or it stabilizes the RNA interaction with the RNA-binding region ARM[29]. The key feature of the DRYG-rich region is its ability to promote the self-association of eIF4B, both in vitro and in vivo[31]. Moreover, the N-terminal ~210 residue portion of eIF4B that encompasses the RRM domain and its N-terminal tail is dispensable for self-association[31]. Notably, the removal of the DRYG-rich region leads to a loss of global translation inhibition associated with the overexpression of eIF4B in certain mammalian cells[29,32], suggesting that DRYG-driven self-association of eIF4B might be crucial for its function in translation regulation. The exact links between these observations however remain unexplained, partly due to the lack of description of molecular principles underlying the eIF4B self-association.

Here, with a goal of understanding the mechanistic bases of eIF4B self-association we focus on the C-terminal IDR of eIF4B, starting from the DRYG-rich region up to the complete C-terminus and perform a detailed characterization of structural dynamics and interactions using a combination of single-molecule FRET spectroscopy (smFRET) and nuclear magnetic resonance (NMR) spectroscopy experiments, coarse-grained Langevin dynamics simulations and phase separation assays. By integrating the atomic and molecular information at different scales, we obtain a comprehensive picture of factors defining the

molecular behavior of eIF4B IDR across several regimes of the self-association landscape, ranging from monomers to oligomers and condensates.

Results

DRYG-CTR undergoes homotypic phase separation

As the DRYG-rich segment in the IDR has a reported role in eIF4B self-association, we first focused on this fragment of eIF4B for detailed investigation. The minimal DRYG protein fragment (residues 214–332) was purified from inclusion bodies, where it exclusively localized during recombinant expression (see Supplementary Methods for details). In line with this observation, the DRYG-rich fragment was only soluble in the presence of guanidine hydrochloride (GdmCl) and only within a limited protein concentration range (up to ~10 μM in 0.5 M GdmCl). At higher concentrations the purified construct again underwent phase separation, effectively prohibiting characterization of the minimal DRYG-rich region under native conditions. From literature, the start of the DRYG-rich region is well defined (i.e. residue 214)[17,31] whereas the end residue is more vaguely specified and had been suggested to span beyond residue 327[29,30]. The easiest option therefore was to extend the protein construct from the DRYG-rich region towards the final C-terminal end of the protein (residues 213–611), thus encompassing the almost complete C-terminal IDR of eIF4B in a construct we term DRYG-CTR (Fig. 1a, b).

Already at the stage of protein purification, we noticed that DRYG-CTR also undergoes phase separation, albeit to much lesser degree as compared to the minimal DRYG-rich fragment and thus was readily amenable to further characterization. Differential interference contrast (DIC) and fluorescence imaging confirmed the formation of microscopic phase-separated droplets, which undergo the fast fusion characteristic of liquid-like behavior (Fig. 2a, b). To map the condensation properties of DRYG-CTR, we first quantified the protein concentration in dilute (c_{sat}) and dense (c_{dense}) phases at different temperatures. For this assay, protein solutions were prepared under conditions favoring phase separation and were incubated at different temperatures, followed by centrifugation to separate light and dense
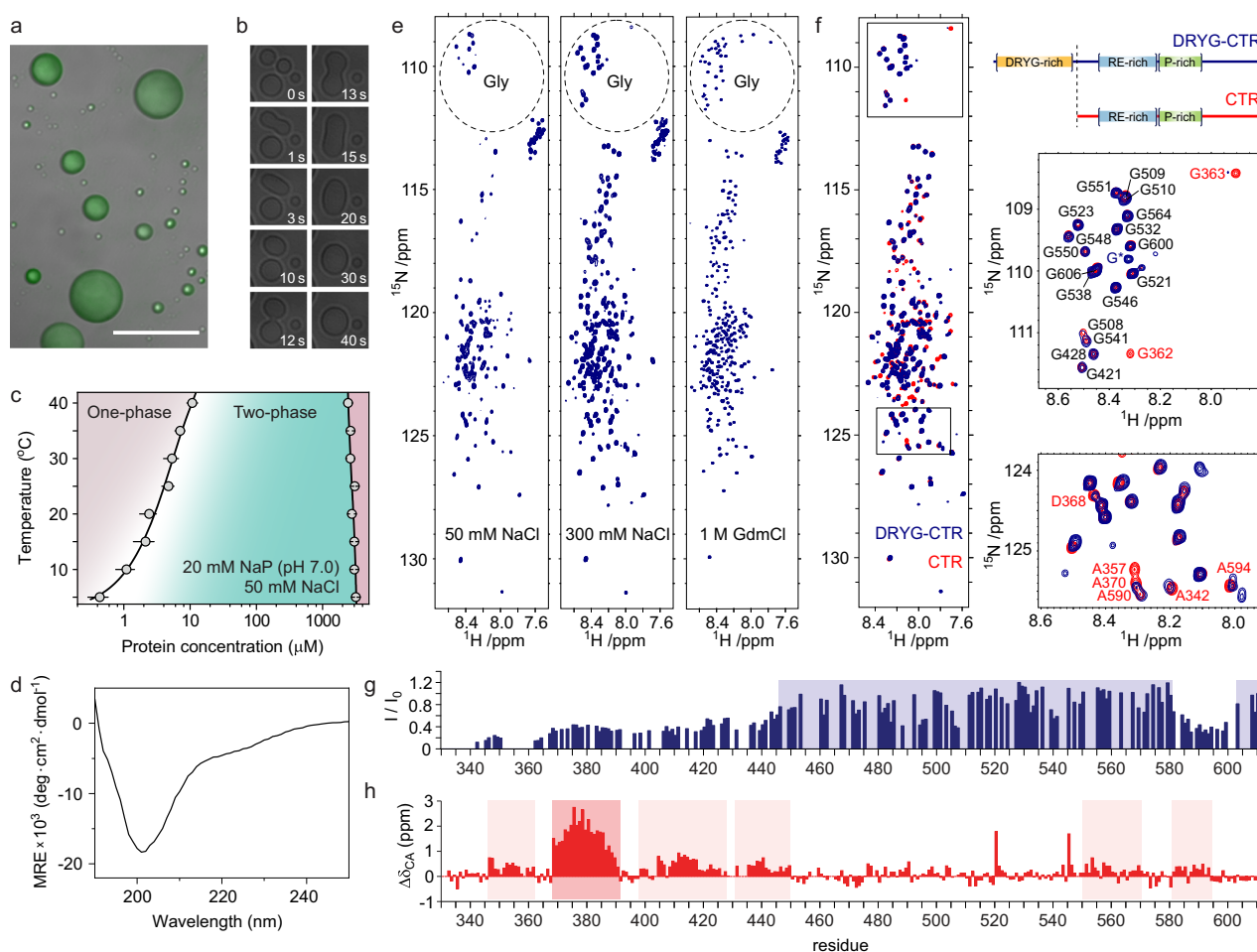

Fig. 2 | The structural and phase separation behavior of eIF4B IDR. a An overlay of DIC and fluorescence images of liquid-like droplets forming upon phase separation of DRYG-CTR. The image is representative of results from 3 independent experiments. The scale bar indicates 20 μm. **b** A time course of DIC images showing rapid fusion of protein droplets. **c** Phase diagram of DRYG-CTR in 20 mM NaP, 50 mM NaCl (pH 7.0), showing the temperature dependence of c_{sat} (left branch) and c_{dense} (right branch). Each data point was measured in 3 independent replicates; the data are presented as mean values +/− SD. The black solid lines are included as visual guides to emphasize the boundaries between the one-phase (highlighted in purple) and two-phase (highlighted in cyan) regimes. **d** Far-UV CD spectrum of DRYG-CTR in 20 mM NaP, 75 mM NaCl (pH 7.0), highlighting the low secondary structure content in DRYG-CTR. **e** ^{1}H,^{15}N-HSQC spectra of DRYG-CTR in 20 mM NaP (pH 7.0) buffer with 50, 300 mM NaCl and 1 M GdmCl. The dashed circles highlight resonance peaks for glycine residues. **f** The overlay of HSQC spectra for DRYG-CTR (blue) and CTR (red) constructs in 20 mM NaP, 150 mM NaCl (pH 7.0). The right panel shows the schematics of DRYG-CTR and CTR constructs, together with closeup segments of the overlaid HSQC spectra. **g** The ratio of peak intensities for DRYG-CTR and CTR constructs, the light blue shaded areas indicate the protein regions with comparable peak heights between the two constructs. **h** Secondary chemical shift values for $^{13}C_\alpha$ in CTR, based on Biological Magnetic Resonance Bank (BMRB) entry 51957[28] and random coil values calculated by the online ncIDP server[93]. The regions of increased secondary structure propensity are highlighted in shades of light red. Source data are provided as a Source Data file.

phases[33]. The protein concentrations in the supernatant (light phase) and pellet (dense phase) were then measured, corresponding to c_{sat} and c_{dense}, respectively (see Methods for details). This concentration window defines the DRYG-CTR phase separation regime, outside of which the protein is easily soluble and miscible in solution (Fig. 2c). The resulting temperature-dependent phase diagram reveals an increase of phase separation propensity at lower temperatures, following a characteristic upper critical solution temperature (UCST) behavior. At the same time, the phase separation of DRYG-CTR is favored in low ionic strength buffers, e.g. 20 mM sodium phosphate (NaP), 50 mM NaCl, pH 7.0 (Fig. 2c), and is largely suppressed ($c_{sat} > 1.5$ mM) at NaCl concentrations above 150 mM.

eIF4B IDR is largely devoid of stable secondary structure

To proceed with structural characterization of DRYG-CTR, we next assessed the degree of protein secondary structure by circular dichroism (CD), under conditions suppressing protein phase separation (Fig. 2d). The CD spectrum shows a large minimum at 202 nm and small ellipticity signal at 220 nm, which is consistent with low secondary structure content as expected for disordered proteins. The small contribution of ellipticity at 220 nm is likely originating from transient short helical motifs at the N-terminal half of the CTR as suggested by our previous NMR data[28] and CD data of truncated CTR constructs (Supplementary Fig. 1).

The ^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) spectrum of ^{15}N-labeled DRYG-CTR is useful to probe the nature of protein behavior at the residue level. Under a variety of conditions, the only backbone amides that are observed appear as crosspeaks in the narrow ^{1}H chemical shift range from 8.0–8.5 ppm, which is characteristic of disordered peptides (Fig. 2e). Nevertheless, at low ionic strength (50 mM NaCl), the ^{15}N-HSQC spectrum of DRYG-CTR displays only a few clear resonance peaks, which account for only a small portion of residues from the full construct. In particular, the glycine region in the spectrum is easy to analyze (circled region in Fig. 2e), and only 10 out of the total 36 expected glycine backbone amide crosspeaks are observed. Line-broadening (that is, loss of crosspeak intensity) can be caused by several factors including dynamics, heterogeneous folded states or formation of large oligomers, and in this case, it is most likely consistent with the condensation of the protein observed in low ionic strength (Fig. 2a–c). With an increase in ionic strength (300 mM NaCl) the number of observed glycine amide crosspeaks approximately doubles (Fig. 2e, middle spectrum), consistent with observed reduction in phase separation at higher salt concentration. As further shown in more detail later in the results, this behavior is also consistent with the data indicating the formation of smaller protein assemblies under these conditions. As a control experiment, the construct was fully destabilized upon addition of 1 M GdmCl which should prevent essentially all inter- and intramolecular interactions. Indeed, this ^{15}N-HSQC spectrum has equal intensity sharp amide peaks for what appears to be the entire protein sequence, as specifically evidenced by observation of all 36 glycine backbone amides (Fig. 2e, right spectrum).

To return to the DRYG-CTR construct at moderate ionic strength, the number and position of observed amide crosspeaks is remarkably similar to the pattern observed for the isolated CTR[28]. An overlay of the DRYG-CTR and CTR spectra at 150 mM NaCl highlights this similarity (Fig. 2f) and suggests that the visible residue crosspeaks all derive from the CTR region, and that the DRYG residue amides are in contrast all line-broadened. Our previous chemical shift assignment of human eIF4B CTR allows for a higher-resolution analysis, and it appears that not only the DRYG region, but most of the N-terminal third of the CTR displays markedly different behavior in the DRYG-CTR sample (residues 332–450, such as seen for G362, G363) (Fig. 2g). There also appears to be a small segment in the C-terminus from residue 580–600 (such as A590, A594) that also has more peak-broadening in

the DRYG-CTR sample. Prediction of secondary structure elements in the CTR based on the chemical shift assignments[28] (Fig. 2h) also indicates that regions of stable and transient α-helices mainly overlap with CTR residues that experience line-broadening when the DRYG region is present. Taken together, it appears that the DRYG-CTR construct spectra are consistent with the formation of a large possibly heterogenous assembly, centered on the DRYG region but also encompassing to a lesser degree the neighboring region of CTR and a small C-terminal segment. Unfortunately, the NMR data also demonstrate that increased insight in the molecular nature of DRYG-CTR will require complementary biophysical techniques.

Sequence bias defines conformational subregions of distinct compactness and dynamics

Although the NMR analysis confirms the disordered nature of the CTR, as well as general segments involved in self-association, the high-resolution data was restricted to residues outside of the DRYG region. To overcome this limitation in the NMR analysis we opted to characterize the conformational behavior and structural dynamics of DRYG-CTR using smFRET, as this technique is amenable to observation of defined subregions in a wide range of protein states, even in conformationally heterogenous samples. For this purpose, we created four different fluorescently labeled DRYG-CTR constructs: DRYG-CTR$_{P213C-P332C}$, DRYG-CTR$_{P332C-C457}$, DRYG-CTR$_{C457-G523C}$ and DRYG-CTR$_{G523C-Y609C}$ (Fig. 3a). We selected the labeling positions to ensure that each labeled construct probes distinct segments of specific sequence biases in the context of complete DRYG-CTR. In this context, the residues 213 to 332 bracket the low-complexity DRYG-rich region. Additionally, we defined the segment of residues 367 to 455, which partially overlaps with previously identified ARM region, and termed it the RE-rich region, due to its enrichment in alternating arginine and aspartate residues. Downstream to this region we defined a segment of residues 460 to 522, which we termed P-rich region, as it contains a large cluster of prolines. Additionally, we note the characteristic large negative net charge of the far C-terminal region (residues 523–611), particularly due to the cluster of glutamate and aspartate residues at the end of the protein. Based on these constructs, we then performed a large set of smFRET and nanosecond fluorescence correlation spectroscopy (nsFCS) experiments to characterize the corresponding protein regions of DRYG-CTR.

Starting with smFRET analysis of the DRYG subregion, the transfer efficiency (E) histograms of DRYG-CTR$_{P213C-P332C}$ (Fig. 3b and Supplementary Fig. 2) show single relatively narrow peaks consistent with a fast-interconverting ensemble of disordered conformations. At the same time, the analysis of fluorescence lifetimes reveals that the normalized donor lifetime versus transfer efficiency histograms systematically deviate from the static diagonal line, and instead are centered along the curved line. This behavior is expected for a broad distance distribution of disordered proteins. Considering the large sequence separation between the two fluorophores (120 amino acids), the transfer efficiency histogram of DRYG-CTR$_{P213C-P332C}$ is centered at a surprisingly high mean transfer efficiency, which indicates that this labeled segment displays an ensemble of very compacted disordered conformations. The analysis of fluorescence lifetimes of both dyes confirms that observed transfer efficiencies are not affected by dye quenching artifacts (Supplementary Table 1). Furthermore, measurements of the same construct labeled with different dye pairs yielded consistent results (Supplementary Fig. 3), additionally confirming the robustness of the observations regardless of the specific dyes used. The conformation of the probed region is only slightly affected by ionic strength, as evidenced by a small increase of mean distance (R_{DA}) between the fluorophores upon increasing concentration of NaCl (Fig. 3c). In contrast, the transfer efficiency peak monotonically shifts toward lower transfer efficiencies with increasing GdmCl concentration (Fig. 3b and Supplementary Fig. 4), indicating a continuous
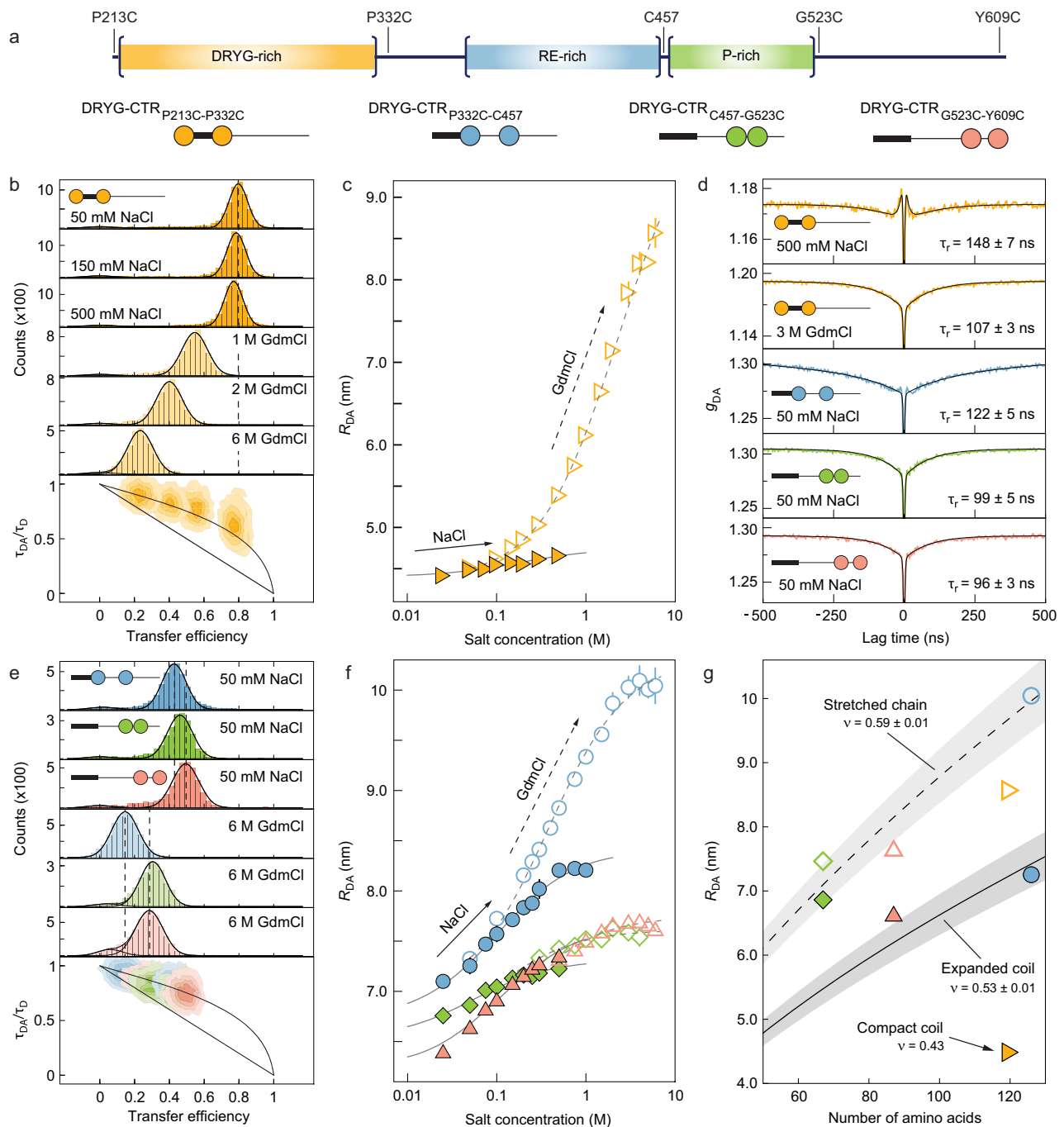
Fig. 3 | Different regions of eIF4B IDR have distinct compactness and dynamics. a A schematic representation of the DRYG-CTR construct, highlighting the cysteine residues used for fluorescent labeling. The bottom panel shows icons for four fluorescently labeled constructs used in smFRET experiments. **b** Transfer efficiency, E, histograms of DRYG-CTR$_{P213C-P332C}$ at different concentrations of NaCl and GdmCl. The dashed line indicates mean E at 50 mM NaCl. The bottom panel shows donor lifetime-E histograms, under the same conditions as the E histograms. The diagonal and curved lines indicate the dependence expected for fixed distance and broad distance distribution, respectively. **c** The mean inter-residue distance, R_{DA}, for DRYG-CTR$_{P213C-P332C}$ depending on NaCl (orange filled symbols) and GdmCl (orange open symbols) concentration. The solid and dashed lines show the fits to an empirical binding isotherm. **d** nsFCS donor-acceptor correlation curves for DRYG-CTR$_{P213C-P332C}$ at 500 mM NaCl and 3 M GdmCl (orange), and for DRYG-CTR$_{P332C-C457}$ (blue), DRYG-CTR$_{C457-G523C}$ (green) and DRYG-CTR$_{G523C-Y609C}$ (pink) at 50 mM NaCl. The solid black lines are fits used to determine the τ_r reconfiguration times, as indicated (see "Methods" for details). **e** Transfer efficiency

histograms of DRYG-CTR$_{P332C-C457}$, DRYG-CTR$_{C457-G523C}$ and DRYG-CTR$_{G523C-Y609C}$ at 50 mM NaCl and 6 M GdmCl. The dashed lines indicate mean E for DRYG-CTR$_{P332C-C457}$ and DRYG-CTR$_{G523C-Y609C}$ in the presence of NaCl and GdmCl. The bottom panel shows donor lifetime-E histograms, under the same conditions as the E histograms. **f** The R_{DA} for DRYG-CTR$_{P332C-C457}$ (blue), DRYG-CTR$_{C457-G523C}$ (green) and DRYG-CTR$_{G523C-Y609C}$ (pink) depending on NaCl (filled symbols) and GdmCl (open symbols) concentration. The solid and dashed lines show the fits to an empirical binding isotherm. Each data point in (**c**, **f**) was acquired in 3 independent replicates and are presented as mean values +/− SD. **g** The dependence of R_{DA} for DRYG-CTR$_{P213C-P332C}$ (orange), DRYG-CTR$_{P332C-C457}$ (blue), DRYG-CTR$_{C457-G523C}$ (green) and DRYG-CTR$_{G523C-Y609C}$ (pink) on sequence separation (N_{aa}), at 50 mM NaCl (filled symbols) and 6 M GdmCl (open symbols). The solid and dashed lines indicate the R_{DA} - N_{aa} dependence with ν = 0.53 and ν = 0.59, respectively. The gray bands show the expected range considering ±0.01 uncertainty of ν. Source data are provided as a Source Data file.

expansion of this segment (Fig. 3c), similar to other unfolded proteins and IDPs[34–36].

The nsFCS measurements with DRYG-CTR$_{P213C-P332C}$ enabled us to quantify the long-range distance dynamics between the two fluorophores bracketing this protein segment (Fig. 3d and Supplementary Fig. 5). In the absence of GdmCl, we observed chain reconfiguration times (τ_r) in the range of 150 ns that is within a typical timescale for IDPs[37], albeit reflecting overall reduced fluctuations of the polypeptide chain. The relatively small R_{DA} and larger τ_r values for DRYG-CTR$_{P213C-P332C}$ indicate a large contribution of transient intramolecular contacts within this segment that compact the conformational ensemble and slow down the chain dynamics of the corresponding DRYG-rich region. Both the R_{DA} and τ_r values are only slightly affected by the increase of the ionic strength, which is in contrast to what would be expected for purely electrostatic intramolecular contacts. Thus, despite the large fraction of alternating charged residues, additional intramolecular interactions also contribute to the overall chain dimensions and dynamics. These interactions can include the cation-π contacts between the arginine and tyrosine residues and π-π stacking interactions between two tyrosine residues or two arginine residues (via stacking of guanidinium groups), with the latter possibly being stabilized by charge-neutralization through neighboring negatively charged aspartate sidechains. As previously shown, the strength of π-π and cation−π interactions increases with increasing ionic strength[38–40] and thus may compensate for the weakening of electrostatic interactions at higher salt concentrations, therefore leading to practically unaltered chain dimensions and dynamics. Increasing the GdmCl concentration results in decrease of the τ_r values, consistent with GdmCl-induced disruption of transient intramolecular interactions within the DRYG-rich region and thus faster disordered chain fluctuations upon its expansion (Fig. 3c, d). Altogether, our smFRET data confirms that the DRYG-rich region adopts relatively compact but nevertheless disordered and flexible conformations, with overall reduced chain dynamics.

Similar to DRYG-CTR$_{P213C-P332C}$, the transfer efficiency histograms of DRYG-CTR$_{P332C-C457}$, DRYG-CTR$_{C457-G523C}$ and DRYG-CTR$_{G523C-Y609C}$ also show single narrow peaks centered at intermediate E values, meanwhile the fluorescence lifetimes cluster above the diagonal line, which collectively confirm the more expanded and flexible nature of these labeled regions (Fig. 3e and Supplementary Fig. 2). The chain reconfiguration times, τ_r, from nsFCS measurements are in the range of ~100 ns (Fig. 3d and Supplementary Fig. 5), which is typical for disordered proteins of comparable lengths[36,37]. In contrast to the DRYG-rich region, these labeled segments are sensitive to the ionic strength, expanding to various degrees upon the increase of NaCl concentration (Fig. 3f). This behavior indicates the contribution of intramolecular electrostatic contacts, due to a large fraction of oppositely charged residues (Fig. 1c), which weaken upon electrostatic screening at higher ionic strengths. A similar expansion is also observed for these segments upon increasing GdmCl concentrations (Fig. 3f and Supplementary Fig. 4).

It is remarkable that the dimensions of probed regions do not scale with their amino acid sequence length, as it would be expected for an ideal disordered chain in the absence of any intra-chain interactions (Fig. 3g). As already mentioned, the main outlier is DRYG-CTR$_{P213C-P332C}$, which has a scaling exponent ($\nu = 0.43$) characteristic of compact coils. The high degree of compactness observed in this region is somewhat unexpected, considering the significant presence of well-interspersed alternating charged residues ($\kappa = 0.033$) and is likely due to the substantial fraction of aromatic amino acids and particularly tyrosine residues that over-compacts the disordered chain through a network of cation-π and π-π contacts, as observed for other IDRs[41]. To investigate this possibility, we created a mutant construct where all tyrosine residues within the DRYG region were replaced with serines i.e. DRYG-

CTR$_{P213C-P332C}$-YtoS construct (see Supplementary Methods for details) and performed similar smFRET experiments (Supplementary Fig. 6). As expected, in the absence of tyrosines, this region is more expanded (R_{DA} ~ 6.4 nm at 50 mM NaCl) with scaling exponents $\nu = 0.50–0.52$ in the range of 0–500 mM NaCl (Supplementary Fig. 6a, b). At the same time, this mutated segment exhibits faster intrachain dynamics ($\tau_r = 114 \pm 1$ ns; Supplementary Fig. 6c) compared to the wild-type construct, which altogether supports the significant role of tyrosine residues in defining the particular conformational behavior of the DRYG-rich region. For the other three constructs we observe larger scaling exponents ($\nu > 0.5$) consistent with more expanded conformations of corresponding probed regions. It is worth mentioning that DRYG-CTR$_{C457-G523C}$ possesses a larger scaling exponent ($\nu = 0.58$) compared to DRYG-CTR$_{P332C-C457}$ and DRYG-CTR$_{G523C-Y609C}$, which can be attributed to the presence of a large cluster of proline residues (Fig. 1b). The proline-rich region is expected to increase the stiffness of the polypeptide chain and promote transient polyproline-II helix formation (Fig. 1b), thus resulting in a more expanded conformation of this region[42,43]. The stiffness of DRYG-CTR$_{C457-G523C}$ is also evident from the relatively larger values of τ_r, (e.g. compared to DRYG-CTR$_{G523C-Y609C}$), as it would also be expected to scale with the chain length in case of a similar degree of internal friction[35,44]. Scaling exponents of these three regions systematically increase upon addition of GdmCl and cluster around $\nu = 0.59 \pm 0.01$, as expected for largely expanded disordered chains in the presence of denaturants. Here as well, DRYG-CTR$_{P213C-P332C}$ possesses a smaller scaling exponent, indicating that the corresponding protein region is yet more compact compared to the ideal chain, likely due to a fraction of intrachain cation-π and π-π contacts that persist even at high GdmCl concentrations.

Despite being largely unstructured, the eIF4B IDR therefore does not exhibit a uniform conformational behavior and plasticity. Different regions instead possess a variable degree of compactness and dynamics, primarily dictated by the amino acid composition of these segments and underlying intrachain contacts. These two features also define the protein response to changes of ionic strength, with the DRYG-rich region remaining essentially unaffected at increasing salt concentrations, whereas the CTR expands upon screening of intrachain charge-charge interactions.

eIF4B IDR self-associates into large-size oligomers driven by the DRYG-rich region

As eIF4B was previously reported to form simple homodimers mediated by the DRYG-rich region[31], we wondered if DRYG-CTR may be able to self-associate at sub-saturation concentrations independently of phase separation (Fig. 2a–c), and in particular at conditions in which phase separation is not observed. We therefore used single-molecule fluorescence correlation spectroscopy (FCS) to probe the sub-nanomolar concentrations of fluorescently labeled monomeric DRYG-CTR upon titration with unlabeled protein (Fig. 4a). Upon increasing unlabeled protein concentrations, we observed a shift of FCS curves towards longer times, consistent with DRYG-CTR self-association and molecular size increase for diffusing protein species (Fig. 4b). Similar experiments with the CTR construct show no change in the protein diffusion rate, confirming the essential role of the DRYG-rich region in oligomerization (Supplementary Fig. 7). Fitting FCS data allowed us to extract the average diffusion rates of diffusing protein species. The relative diffusion rates gradually decrease upon self-association at increasing protein concentrations (Fig. 4c). Surprisingly, the observed drop of diffusion rates extends beyond what would be anticipated for simple dimers, reaching a plateau at values corresponding to oligomers with an apparent average size of >10-mers. The formation of large oligomers is further confirmed by crosslinking-coupled SDS-PAGE and electron microscopy (EM) analyses. At high salt (300 mM NaCl) we primarily observe relatively small oligomers
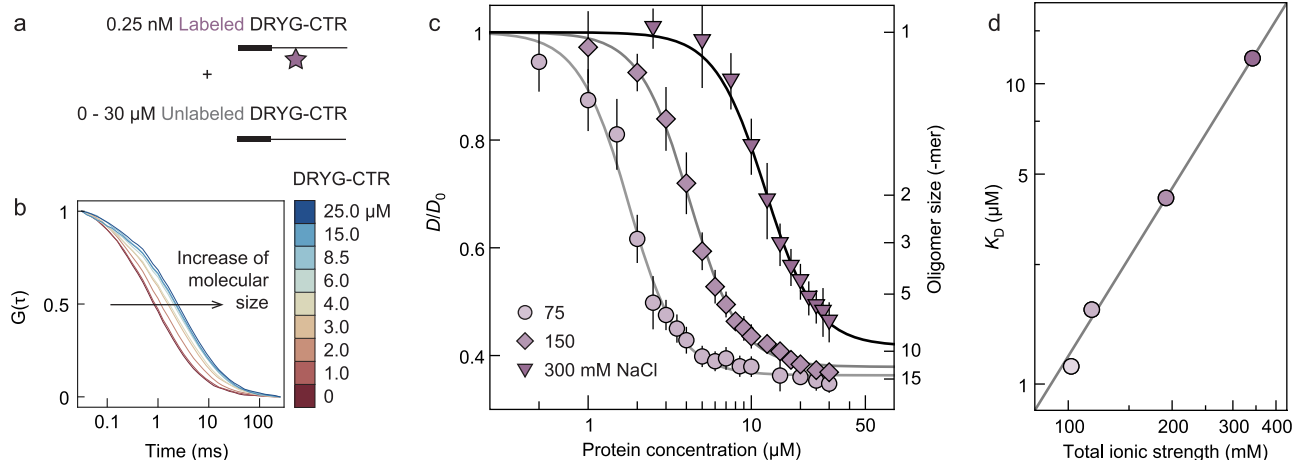
Fig. 4 | eIF4B IDR undergoes self-association and forms large-size oligomers, with weak dependence on ionic strength. a A schematic of the FCS oligomerization assay, based on a titration of sub-nanomolar labeled DRYG-CTR with micromolar concentrations of unlabeled protein. **b** A set of FCS curves at 75 mM NaCl obtained in the presence of increasing protein concentration (from dark red to dark blue). The arrow highlights a shift of FCS curves towards higher times, indicating a growing size of diffusing species. **c** The dependence of relative diffusion rates (normalized to the diffusion rate of monomeric protein) on protein concentration, acquired in buffers containing 75 (light purple circles), 150 (purple rhombi) and 300 (violet triangles) mM NaCl, indicates a reduction of protein association affinity with increasing buffer ionic strength. Each data point was measured in 3 independent replicates; the data are presented as mean values +/− SD. The solid lines show global fits to the Hill equation. The right axis shows the size of oligomers, considering the $D \sim 1/N^{1/3}$ dependence (see "Methods" for details). **d** The dependence of apparent dissociation constant on buffer ionic strength. Error bars are represented by standard errors of the fits in (**c**) and are smaller than plot symbols, thus may not be visible. The black solid line is a fit with the Lohman–Record model[94]. Source data are provided as a Source Data file.

(ranging from 2- to 6-mers), whereas at lower ionic strength the majority of oligomeric species are unable to enter the gel due to their larger size (Supplementary Fig. 8a). These larger oligomers could be observed by negative stain EM (Supplementary Fig. 8b–e), as round-shaped particles in the range of 25–35 nm. The formation of such heterogeneous size oligomers is in line with several observations from FCS experiments, such as the monotonous shift of FCS curves with increasing oligomer fraction (Fig. 4b), as opposed to a simple two-state behavior expected in case of monomer to fixed-size oligomer transition, as well as the apparent cooperativity in monomer-oligomer transition curves (that is binding isotherm with $n > 1$; Fig. 4c and Supplementary Table 2). Thus, the eIF4B IDR does not undergo simple dimerization, as previously proposed[31], but forms a range of variable-size oligomers, including large oligomer assemblies.

Interestingly, the observed average size of the oligomers is slightly dependent on ionic strength, plateauing at average oligomer size of ~15-mers in the presence of 75–150 mM NaCl, but dropping to <10-mers at higher ionic strength, which is in line with crosslinking experiments that indicate a reduction of the large oligomer fraction at higher salt concentrations. The midpoint of monomer to oligomer transition curves shown in Fig. 4c reflects the apparent dissociation constant, K_D, of oligomers, which decreases upon lowering the NaCl concentration (Fig. 4d), indicating an increased DRYG-CTR association affinity. Interestingly, the effect of ionic strength on DRYG-CTR association is more pronounced compared to the salt dependence of DRYG-rich region dimensions and dynamics. Nevertheless, the change of dissociation constants with ionic strength is modest and the analysis of the NaCl dependence of K_D reveals a relatively small number of counterions ($\Delta n = 1.86 \pm 0.04$) released upon protein self-association (Fig. 4d). This observation points towards an overall weak electrostatic nature of DRYG-DRYG mediated protein association and implies that DRYG-CTR intermolecular interactions also contain contributions from cation-π and π-π contacts, similar to the intramolecular interactions observed within the DRYG-rich region. As in case of intramolecular interactions, tyrosine residues also have a strong impact on DRYG-DRYG association, exemplified by the absence of any detectable oligomerization for DRYG-CTR$_{213-332}$-YtoS construct (Supplementary Fig. 6d).

DRYG-CTR oligomers are disordered and dynamic complexes

Although FCS experiments allowed us to extract substantial details on DRYG-CTR self-association behavior, they did not provide information on the structural and dynamic aspects underlying the oligomerization process. To address this aspect, we took advantage of the double-labeled DRYG-CTR constructs, and performed smFRET oligomerization assays probing transfer efficiency histograms of labeled DRYG-CTR in their monomeric and oligomeric states (Fig. 5a–c). For DRYG-CTR$_{P332C-C457}$, increasing the concentration of the unlabeled protein resulted in a second peak appearing at a lower transfer efficiency. This new low transfer efficiency peak originates from labeled DRYG-CTR molecules that associate to unlabeled oligomers (one labeled molecule per oligomer). The fraction of this peak increases at the expense of the higher transfer efficiency peak, with increasing concentration of unlabeled protein (Fig. 5a). This behavior indicates an increase of the fraction of oligomer-bound labeled protein (Fig. 5d), in parallel with increasing fraction of oligomers (Fig. 4c). For DRYG-CTR$_{C457-G523C}$ and DRYG-CTR$_{G523C-Y609C}$, we instead observed an apparent gradual shift of the single peak towards lower transfer efficiencies (together with increasing peak width at intermediate protein concentrations) (Fig. 5b, c, e). Thus, all three protein regions display a consistent reduction of mean transfer efficiency of the protein when bound to oligomers, in keeping with an oligomerization-driven increase of inter-residue distance throughout the entire IDR. Unfortunately, for the DRYG-CTR$_{P213C-P332C}$ construct, the accurate quantitative interpretation of transfer efficiencies was not possible due to pronounced reduction of dye mobility (i.e. sticking) and quenching upon oligomer formation, which further hindered the assessment of dimensions of this segment in the oligomer-bound state. We wondered if the observed distance increase originates from the overall expansion of the disordered ensemble, or if it can be explained by the formation of local structure. An answer is provided by the analysis of the fluorescence lifetimes of three DRYG-CTR constructs, which systematically cluster at the dynamic line, throughout the course of oligomerization, indicating that the protein remains disordered also in the context of the oligomers (Fig. 5a–c). The same conclusion is also supported by essentially identical CD spectra at low (monomeric) and high (oligomeric) protein concentrations (Supplementary Fig. 9). Interestingly, the three regions
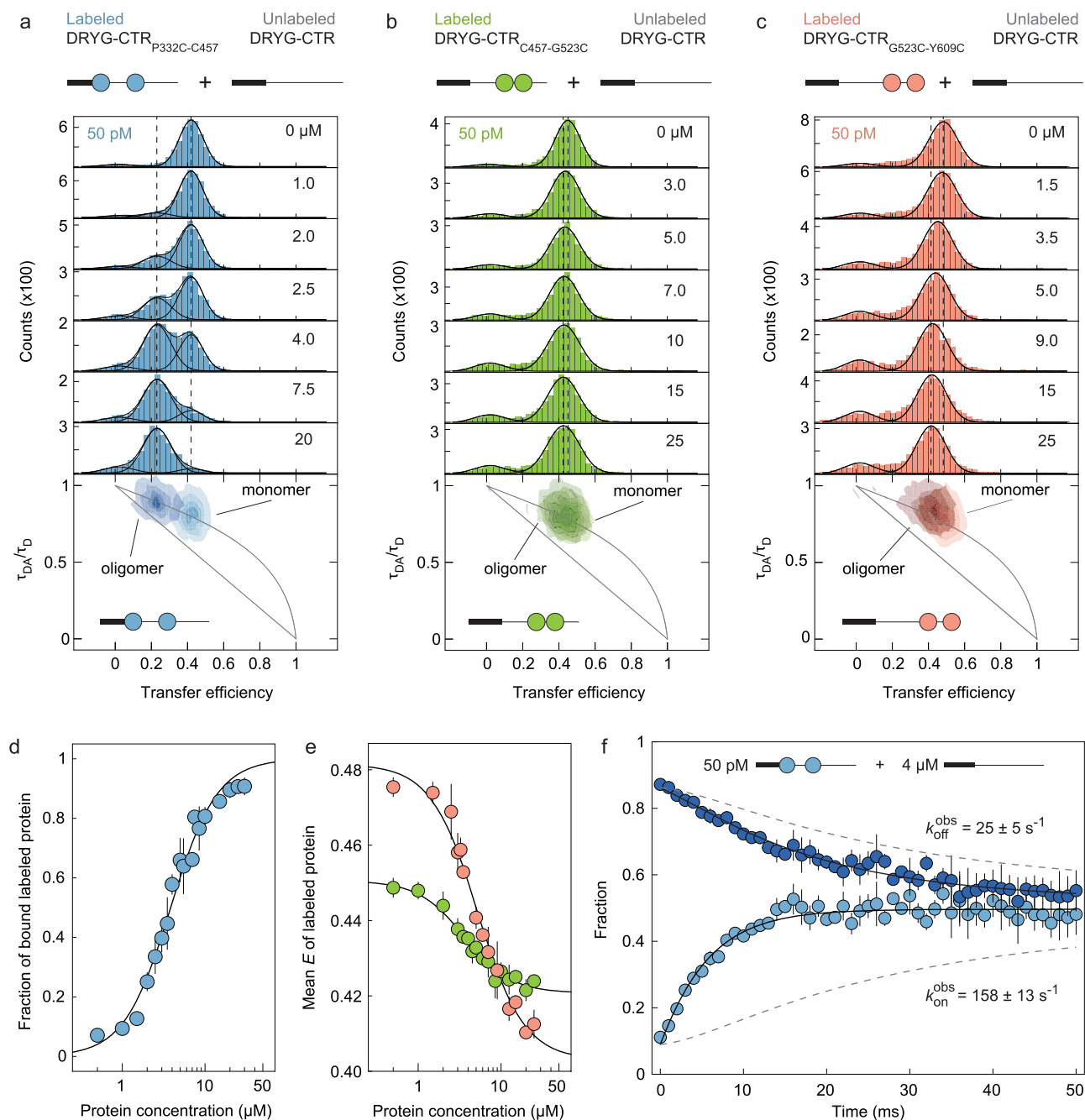
Fig. 5 | Oligomerization of eIF4B IDR and underlying conformational dynamics.
a–c Transfer efficiency histograms of DRYG-CTR$_{P332C-C457}$ (blue), DRYG-CTR$_{C457-G523C}$ (green) and DRYG-CTR$_{G523C-Y609C}$ (pink) constructs at increasing concentrations of unlabeled protein. The dashed lines indicate mean E of unbound DRYG-CTR (monomeric) and bound DRYG-CTR (oligomeric), for each construct, respectively. The bottom panels show donor lifetime - E histograms for each construct in the absence (light colored) and presence (dark colored) of unlabeled protein, corresponding to monomeric and oligomeric states, as respectively indicated. The diagonal and curved lines indicate the static and dynamic lines, as in Fig. 3. **d** The fraction of oligomer-bound DRYG-CTR$_{P332C-C457}$ (blue) as a function of overall protein concentration. Each data point was measured in 3 independent replicates; the data are presented as mean values +/− SD. The black line is a fit to

the Hill equation. **e** The mean transfer efficiencies of DRYG-CTR$_{C457-G523C}$ (green) and DRYG-CTR$_{G523C-Y609C}$ (pink) as a function of overall protein concentration. Each data point was measured in 3 independent replicates; the data are presented as mean values +/− SD. The black lines are fits to the Hill equation. **f** Plot of DRYG-CTR association kinetics in the presence of 4 μM unlabeled protein, showing the relaxation of bound (dark blue) and unbound (light blue) states over time. Each data point was measured in three independent replicates; the data are presented as mean values +/− SD. The solid lines show global fits with single relaxation rates, as indicated alongside the corresponding plots. Dashed lines show the kinetic behavior in the absence of protein association (i.e. originating from the occurrence of new molecules only). Source data are provided as a Source Data file.

nevertheless show variable response to oligomerization, as evidenced by different extent of mean transfer efficiency change for monomeric and oligomeric protein states (Fig. 5a–c). The DRYG-CTR$_{P332C-C457}$ construct undergoes larger expansion ($\Delta R_{DA} \sim 2$ nm), whereas DRYG-CTR$_{C457-G523C}$ and DRYG-CTR$_{G523C-Y609C}$ show more modest expansion

($\Delta R_{DA} < 1$ nm) (Supplementary Fig. 10). Thus, the protein regions at immediate vicinity to the oligomerization core (such as the DRYG-rich region) undergo larger conformational changes upon protein self-assembly, compared to the rest of the chain that senses the oligomerization to a lesser degree.

To probe the self-association kinetics we focused on the DRYG-CTR$_{P332C-C457}$ construct, which has sufficient separation of FRET states corresponding to the oligomer bound and unbound protein, and is thus amenable to recurrence analysis of single particles (RASP)[45]. Indeed, RASP experiments confirm the relatively fast kinetics at a ms timescale (Fig. 5f). Considering the observed on/off rates, we calculated the exchange rate between the bound and unbound proteins, $k_{ex} \approx 184 \pm 18\,s^{-1}$, which indicates that a single DRYG-CTR molecule on average spends ~5 ms as part of oligomers. Thus, combining the characterization of structural behavior and dynamics we can describe the DRYG-CTR oligomers as large disordered complexes that dynamically form on the basis of several protein molecules.

Molecular simulations confirm DRYG-mediated oligomerization and associated IDR expansion
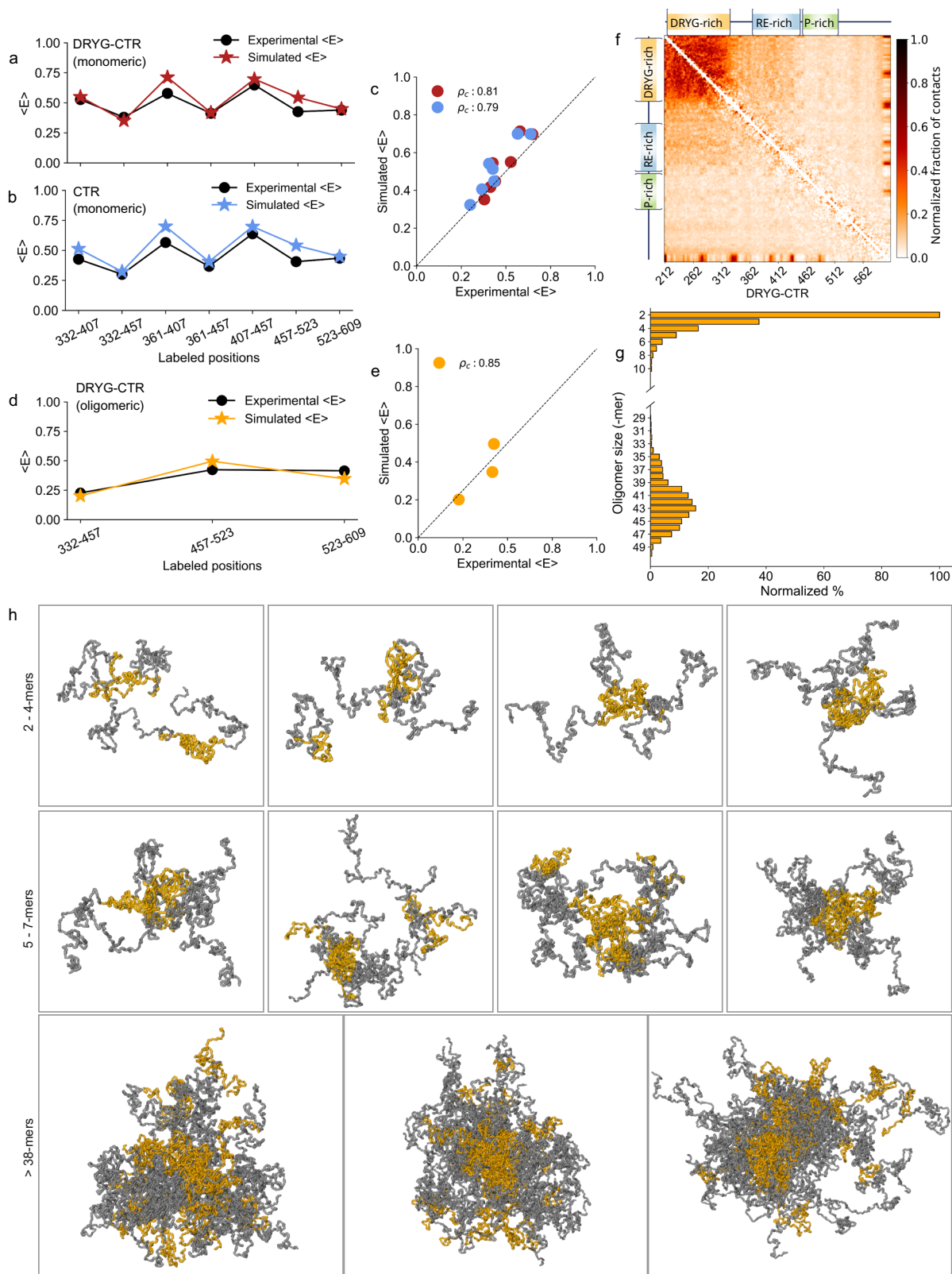
Coarse-grained (CG) simulations have been consistently leveraged against smFRET data, and have yielded experimentally sound ensembles for a series of IDPs[46–49]. We employed CG Langevin dynamics simulations to reconstruct the ensemble of eIF4B IDR constructs (DRYG-CTR and CTR; Supplementary Movies 1, 2), by quantitatively matching experimental and computed mean transfer efficiencies simultaneously, across multiple eIF4B positions. To further enhance the mechanistic understanding of the oligomerization behavior of eIF4B we simulated the eIF4B IDR both as a single chain and oligomers (see Supplementary Movies 3–7 and "Methods" for details). To simulate eIF4B under conditions that favor oligomerization, we modelled 50 molecules in an elongated box ($25 \times 25 \times 200$ nm) and under periodic boundary conditions as per a slab geometry approach[50]. By restricting the box dimensions along one axis, this approach has shown to be effective in sampling differently concentrated regimes at equilibrium, without the need to restrict the conformational space of some of the simulated chains or via random insertions/deletions of chains. At each frame, we monitor oligomeric size by counting the number of molecules within a radius of 1 nm from the center of mass of each other molecules. This approach allowed us to pinpoint the role of the DRYG region in oligomerization of eIF4B, to assess the protein conformational ensemble in the oligomeric state, and to understand how the size of eIF4B oligomers responds to changes in ionic strength, as investigated by experiments. As shown in Fig. 6a–c, the computed mean transfer efficiencies show an excellent agreement (ρ_c of ~0.8) with experimental transfer efficiencies of single chains for both the DRYG-CTR and CTR constructs (Supplementary Fig. 2, 11 and 12). Simulations also recapitulate the drop in the mean transfer efficiencies observed for entirety of the CTR construct when the DRYG-rich region is truncated (Fig. 6b), with the largest effect seen for the 332-457 region. We note that for two regions (361-407 and 457-523) simulations overestimate mean FRET efficiencies by approximately 10–15%. Notably, the 361-407 region brackets the transient helical motif identified in NMR experiments (Fig. 2h), and the proline-rich 457-523 region has a high prediction of PPII helical propensity (Fig. 1b), providing a rationale of more expanded conformations. These regions most prominently contribute to the inability of simulations to account for approximately 20% ($1-\rho_c$) of the variance between experimental and computed FRET efficiencies. Generally, CG potentials still lack the ability to simulate secondary structure elements in disordered chains. The model we employ only contains bonded terms for two subsequently linked beads, without taking into consideration angle and dihedral torsional terms, which contribute to the local and likely transient secondary structure propensity of IDRs.

Given the good agreement with experiments for the mean transfer efficiencies of single chains, we attempted the sampling of the eIF4B ensemble in an oligomeric state. As before, the simulated DRYG-CTR constructs within oligomers show an excellent agreement with experimental transfer efficiencies (Fig. 6d). The simulations capture the reduction of mean transfer efficiencies upon oligomerization,

indicative of expansion of the conformational ensemble upon monomer-oligomer transition as observed in experiments (Fig. 5). In this case, the ρ_c is higher than in the case of single chain simulations ($\rho_c = 0.85$) (Fig. 6e). This result is somewhat unsurprising considering the higher level of constraints in the oligomeric state, which possibly accounts for the rigidity otherwise mediated by proline-rich regions (Fig. 1c) in the single chains. Additionally, we extracted the end-to-end distances (R_{ee}) for the different regions labeled in smFRET experiments both for single chains and oligomers. The comparison reports on an increase of average R_{ee} to various degrees for three constructs, in line with the expansion of the CTR upon oligomerization, as observed in smFRET experiments (Supplementary Fig. 13). Notably, in parallel with the increase of average R_{ee}s, the R_{ee} distributions get systematically narrower. Simulations, together with smFRET experiments (Fig. 5), suggest that the conformational ensemble of eIF4B CTR undergoes pronounced reshaping upon oligomerization. We then further analyzed the molecular components of the ensemble retrieved in the oligomeric state, to understand if specific regions within eIF4B IDR drive association. We quantified the map of inter-chain contacts within oligomers (see "Methods" for details), which shows a significantly higher fraction of contacts concentrating in the DRYG-rich region and gradually fading towards the C-terminal end of the chain (Fig. 6f). Notably, both these observations are in line with the NMR data (Fig. 2f, g) identifying the DRYG-rich region as a key mediator of the oligomerization process. Moreover, the combination of experiments (NMR and smFRET) and simulations provide a highly consistent picture rationalizing the molecular principles of eIF4B IDR oligomer formation.

Heterogenous size distribution of DRYG-CTR oligomers is sensitive to the ionic strength

As demonstrated above, DRYG-CTR forms oligomers of varying size, with pronounced dependence on ionic strength (Fig. 4c and Supplementary Fig. 8). Nevertheless, these experiments provide limited quantitative information about the heterogeneity of the oligomer population or their size distribution. As a goal to expand this observation, we analyzed the populations of simulated DRYG-CTR oligomer ensembles at $I = 117$ mM, corresponding to buffer ionic strength of 20 mM NaP and 75 mM NaCl. In line with experiments (Supplementary Fig. 8), this analysis confirms the formation of oligomers with a diverse range of sizes. The oligomer population displays a pronounced bimodal size distribution, consisting of minimal dimers and relatively small oligomers (3-mers to 8-mers) alongside larger multimolecular oligomeric clusters comprising 30 and more protein molecules (Fig. 6g, h). These larger oligomeric assemblies closely resemble the round-shape particles observed with EM (Supplementary Fig. 8b–e). To compare with the experiments, we also looked at the sensitivity of oligomerization to changes in ionic strength and analyzed the oligomer size distributions at various salt conditions (Supplementary Fig. 14). Similar to simulations at $I = 117$ mM, bimodal oligomer size distribution is also observed at even lower ionic strength $I = 92$ mM, corresponding to buffer ionic strengths of 20 mM NaP and 50 mM NaCl (Supplementary Fig. 14a). Comparing these two conditions, a lowering of ionic strength favors the formation of larger oligomeric clusters, as evidenced by an increase of their population at the expense of the relatively smaller oligomers and an overall increase in the size of the larger clusters (Supplementary Fig. 14a, b). Interestingly, increasing the ionic strength to $I = 192$ and 342 mM (corresponding to buffer ionic strengths of 20 mM NaP with 150 and 300 mM NaCl, respectively), results in an appearance of tailed distributions. Here again, in line with experiments (Supplementary Fig. 8a), the majority of oligomers are observed in the range of 2–6-mers, followed by gradually decreasing smaller fraction of larger oligomers and a notable absence of a defined population of large oligomeric clusters, such as seen at lower ionic strengths (Supplementary Fig. 14c, d). Based on the observed oligomer size distributions, the average oligomer size across the entire

population systematically decreases from approximately 19-mers to 7-mers as ionic strength increases from 117 mM to 342 mM. This trend and the range of average oligomer sizes observed in simulations is consistent with the estimates from FCS experiments (Fig. 4c). Thus, in strong coherence between experiments (Fig. 4c and Supplementary Fig. 8) and simulations (Fig. 6g, h and Supplementary Fig. 14), the eIF4B IDR forms a heterogenous ensemble of oligomeric species whose size

distribution changes with salt concentration, along with the ionic strength dependence of protein association (Fig. 4d).

The interplay between DRYG and CTR determines the eIF4B IDR self-association behavior

As we noted earlier, upon saturation of DRYG-CTR, the protein self-association culminates in mesoscopic phase separation and formation

Fig. 6 | Coarse-grained ensembles of eIF4B IDR suggest a direct role of the DRYG-rich region in driving oligomer formation. a, b Experimental (black) and simulated (red and blue) mean transfer efficiencies, $\langle E \rangle$ as a function of labeled positions for monomeric DRYG-CTR (**a**) and CTR (**b**) at an ionic strength of $I = 192$ mM (corresponding to buffer ionic strength of 20 mM NaP with 150 mM NaCl). **c** Correlations between the experimental and simulated $\langle E \rangle$. The concordance correlation coefficients (ρ_c) are reported in the legend for DRYG-CTR and CTR, respectively. **d** Experimental (black) and simulated (orange) $\langle E \rangle$ as a function of labeled positions for DRYG-CTR simulated at an ionic strength of $I = 117$ mM (corresponding to buffer ionic strength of 20 mM NaP with 75 mM NaCl) and conditions favoring formation of oligomers (see "Methods" for details). **e** Correlation between the experimental and simulated $\langle E \rangle$, with the concordance

correlation coefficients (ρ_c) reported in the legend. **f** Difference in the fraction of contacts between monomeric and oligomeric DRYG-CTR, at an ionic strength of $I = 117$ mM, as observed in experiments. Darker regions indicate the intermolecular contacts when DRYG-CTR is in its oligomeric state, with an increase in inter-chain contacts particularly seen within the DRYG region (residues 213–331). **g** Bar plot showing the size of oligomers obtained from the simulations of DRYG-CTR in the oligomeric state at $I = 117$ mM. Each oligomer size is expressed as a percentage of the simulated trajectory and normalized considering the size of the smallest oligomer (2-mer). **h** Representative conformations of oligomeric clusters of eIF4B IDR. The DRYG-rich region is represented as gold, whilst the CTR is shown in gray. Source data are provided as a Source Data file.

of condensed protein-rich droplets (Fig. 2a–c). Already through our initial qualitative characterization we noted that the DRYG-rich region has extremely high phase separation propensity, which however decreases in the presence of the CTR. To rationalize this observation, we quantified the phase separation propensity of the isolated DRYG or CTR segments of the protein (Fig. 7a) and under the same conditions as the DRYG-CTR (Fig. 2c). The comparison of saturation concentrations for these variants (Fig. 7b), indicates that the DRYG-rich region is indeed essential for the phase separation of the eIF4B IDR. In the absence of this region the CTR construct does not show any detectable phase separation under physiological protein concentration and ionic strength, whereas DRYG alone largely phase separates at very low protein concentrations even in solutions of high ionic strength ($c_{sat} \sim 0.2\,\mu M$ in 1 M NaCl). Interestingly, the extremely low phase separation threshold of DRYG is increased by an order of magnitude in the presence of CTR, as evidenced by 2-3 µM saturation concentration of DRYG-CTR. This indicates that the CTR acts as a phase separation dampener, lowering the extremely strong intrinsic phase separation propensity of the DRYG-rich region. To dissect the role of CTR in inhibition of DRYG-driven phase separation we designed two additional deletion constructs (DRYG-CTR-N and DRYG-CTR-C), in which the C-terminal or N-terminal halves of the CTR were removed, respectively (Fig. 7a). The CTR-N segment encompasses the entire RE-rich region, almost evenly interspersed by positively and negatively charged residues (Fig. 1c), and is slightly acidic at neutral pH (net charge is −1). The CTR-C segment on the other hand encompasses the entire P-rich region and the heavily charged far C-terminal region and has a net charge of −5, particularly due to a cluster of negatively charged residues at the very end of the protein (Fig. 1c). Interestingly, the CTR-N has relatively modest inhibitory effect on DRYG-driven phase separation and increases the saturation concentration of DRYG by less than a factor of 5 (Fig. 7b). The effect of CTR-C is much more dramatic as it raises the phase separation threshold by several orders of magnitude. This could be correlated with its acidic nature, suggesting that the overall net charge of the segments downstream of DRYG-rich region defines the saturation concentration of the protein, as also observed with other IDRs[51–53].

To rationalize the mechanisms of DRYG-CTR self-association and identify the link between the phase separation and oligomerization we quantified the phase separation saturation concentrations at different ionic strengths (Fig. 7c) and compared them with oligomerization affinities also acquired at varying salt concentrations. Interestingly, the ionic strength has a pronounced effect on DRYG-CTR phase separation propensity, as seen from the large increase of the saturation concentrations at high salt concentrations. This behavior is in striking contrast to the weaker ionic strength dependence of DRYG-CTR oligomerization (Figs. 4c, d and 7c), thus indicating that additional molecular mechanisms are behind the phase separation process. Notably, the ionic strength dependence of oligomerization affinity (Supplementary Fig. 15) and phase separation propensity for DRYG-CTR-N and DRYG-CTR-C constructs follow similar trends as for DRYG-CTR (Fig. 7c). Depending on the actual CTR sequence, the ionic

strength dependent phase diagram (i.e. c_{sat} values) and oligomerization affinity (K_D values) laterally shift along the diagonal of protein - NaCl concentration plane (Fig. 7c). The comparison of these three constructs thus points towards a pronounced coupling between the oligomerization and phase separation processes. Thus, the DRYG-DRYG mediated interaction is an essential step for self-association of eIF4B IDR, whereas the CTR acts as a regulator of self-association efficacy. The balance between these two features defines both the self-association affinity and the condensation threshold at the two extremes of eIF4B IDR self-association landscape, such as oligomerization and phase separation.

Discussion

Human translation initiation factor eIF4B remains one of the least structurally and functionally characterized translation factors. In contrast to the human protein, the yeast ortholog of eIF4B was studied more extensively. Rationalizing the structural and functional behavior of human eIF4B based on these data is however challenging, as the two orthologues share minimal sequence conservation (~20%) and are not fully equivalent in function[54]. One of the unique features of the human homolog is the presence of the DRYG-rich region within the long IDR, which was shown to promote eIF4B self-association[31]. Notably the yeast ortholog lacks a DRYG-like region (Supplementary Fig. 16) and its self-association was not reported. Nevertheless, both human and yeast orthologues of eIF4B have been shown to localize to stress granules upon cellular stress[24,25,55–57], however, the actual mechanisms of their recruitment to stress granules remain elusive. Yeast eIF4B has been shown to interact with eIF4G[58], which is likely how it becomes sequestered into stress granules[57], in the absence of inherent condensation ability. In contrast to the yeast orthologue, human eIF4B does not bind eIF4G. Instead, it likely depends on its intrinsic self-association ability, driven by the DRYG-rich region, to localize within stress-induced cellular condensates. The DRYG-rich region is largely present in mammals and extends down to some insect species, with varying sequence conservation that appear to diminish with decreasing organism complexity (Supplementary Fig. 16). The DRYG-rich region may therefore have been adapted through evolution as an intrinsic feature ensuring a stimulus-responsive condensation of eIF4B, independent on other co-factors. Furthermore, the high proportion of tyrosine residues and their phosphorylation can potentially serve as an additional fine-tuning mechanism, enabling a delicate regulation of condensation behavior of eIF4B in vivo.

With a goal of obtaining a comprehensive mechanistic picture of eIF4B self-association we integrated a large set of experiments and simulations to structurally characterize the human eIF4B IDR in its various molecular states across different regimes of the self-association landscape. We have described the conformational ensemble of complete eIF4B IDR in its monomeric form and were able to detect and describe the distinct population of large and dynamically forming oligomers. Importantly, our integrative approach enabled us not only to describe the intricate details of oligomerization, but also follow the conformational reshaping of the disordered ensemble
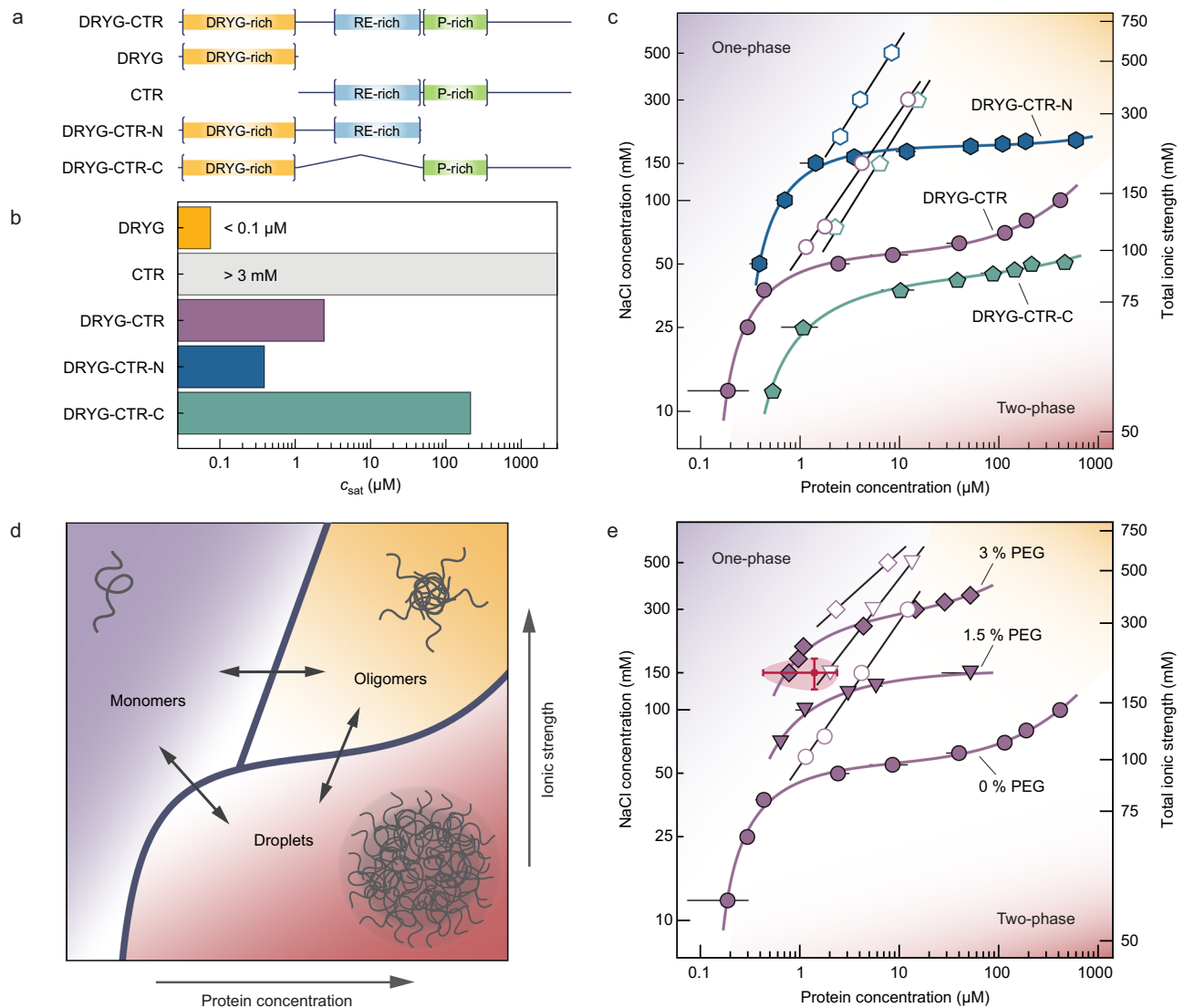
Fig. 7 | Disordered segments of eIF4B define its phase separation behavior.
a Schematics of DRYG-CTR and truncated constructs used in phase separation experiments. **b** Bar chart showing the c_{sat} values for DRYG-CTR and truncated constructs at 20 mM NaP, 50 mM NaCl (pH 7.0), 20 °C. **c** Phase diagrams of DRYG-CTR (purple circles), DRYG-CTR-N (blue hexagons) and DRYG-CTR-C (green pentagons) constructs based on c_{sat} values (filled symbols) and K_D values (open symbols, as in Fig. 4c and Supplementary Fig. 15) as a function of NaCl concentration and total ionic strength (right axis). Each c_{sat} data point was measured in 3 independent replicates; the data are presented as mean values +/− SD. The solid colored lines are fits to an empirical dose-response function. The solid black lines are fits with the Lohman–Record model[94]. The one-phase and two-phase regimes of the phase diagram are indicated. **d** Schematic phase diagram defined by protein concentration and ionic strength indicative of the self-association landscape of eIF4B

IDR and underlying transitions across monomeric, oligomeric and condensed droplet states. **e** Phase diagrams of the DRYG-CTR construct based on c_{sat} values (filled circles) and K_D values (open symbols) as a function of NaCl concentration and total ionic strength (right axis), in the absence (purple cicles) and presence (purple rhombi and triangles) of crowding agent PEG. Each c_{sat} data point was measured in 2 independent replicates; the data are presented as mean values +/− SD. The solid purple lines are fits to an empirical dose-response function. The solid black lines are fits with the Lohman–Record model[94]. The dark red point indicates the average eIF4B concentration in HeLa cells[95–97] and a range of physiological ionic strength. The shaded area shows the mean deviation along both axes. The regions of the phase diagrams in (**c**−**e**) corresponding to monomers, oligomers, and droplets are highlighted in shades of purple, orange, and red, respectively. Source data are provided as a Source Data file.

across the monomer - oligomer transition and quantify the underlying dynamics at intra- and intermolecular scales.

The growing evidence indicates that for many condensation-prone proteins, the propensity of in vitro phase separation, as well as formation of biological condensates, is often correlated with initial formation of large-size oligomers/clusters at concentrations far below the condensation thresholds[51,59–63]. Here we show that the in vitro self-association landscape of eIF4B follows a similar scenario, as exemplified by the behavior of its key self-associating long disordered region. The phase diagram of eIF4B IDR shows three distinct regions across the two-dimensional space of protein concentration versus ionic strength, corresponding to monomeric, oligomeric and condensed

(droplets) states (Fig. 7d). Although a narrow range of low ionic strengths strongly favors a direct phase transition from dispersed monomeric phase, at higher and physiologically relevant ionic strengths the formation of large-size oligomers precedes the phase separation and appear to be a critical and mandatory step along the condensation pathway (Fig. 7). Notably, the DRYG-rich region is at the core of both oligomerization and phase separation, as neither the N-terminal RRM domain (and its tail)[31] nor the C-terminal IDR (i.e. the CTR; Fig. 7b) that flank the DRYG-rich region, exhibit any detectable self-association properties.

Although DRYG-DRYG association is essential for self-association at nanoscale (i.e. oligomer formation) and mesoscale (i.e. droplet

formation), the two processes demonstrate differing responses to ionic strength (Fig. 7c). As discussed earlier, the weak ionic strength-dependent oligomerization (Fig. 4) we attribute to the intermolecular cation-π and π-π interactions, dictated by the presence of a large fraction of tyrosines, similar to the intramolecular contacts that define the compact conformational ensemble for the monomeric DRYG-rich region (Fig. 3). In contrast, the pronounced ionic strength dependence suggests an additional electrostatic driving force behind the phase separation, likely orchestrated through interactions between aspartate and arginine residues of the DRYG-rich region. These interactions are favored at a low ionic strength regime with poor electrostatic screening, leading to rapid coalescence of monomers into phase separated droplets. Higher ionic strengths weaken these contacts and raise the phase separation threshold. In this case, oligomer formation facilitates phase separation by amplifying the effective interaction strength between oligomers in proportion to their size. The phase separation propensity is also correlated with the fraction of larger size oligomers (Fig. 4c and Supplementary Fig. 8, 14). In this context, the distinct population of larger oligomeric assemblies observed in simulations and experiments at low ionic strength conditions (Fig. 6g, h and Supplementary Fig. 8, 14) is similar to the micelles formed by simple diblock copolymers[64,65]. The formation of such micelle-like oligomeric clusters by DRYG-CTR is likely to be driven by the 'blockiness' of its chain, with the DRYG-rich region favoring protein-protein interactions over protein-solvent interactions and forming the cluster core, whereas the CTR forming the surrounding corona due to favorable solvent interactions (Supplementary Fig. 6h, bottom panel). The situation where the micelle-like large oligomeric clusters form along with smaller size oligomers may resemble the so called microphase separation[66–68]. This bimodal regime is favored at low ionic strengths, whereas at higher ionic strength conditions a single continuous regime is favored, populated by continuum of increasing size oligomers with the majority of oligomers sampling 2–6-mers (Supplementary Fig. 14). The shift from a bimodal to unimodal regime is likely due to the salt-induced reduction in valency per monomer (e.g. by screening of the electrostatic charges). As a consequence, this would lead to formation of predominantly small oligomers, due to exhaustion of available valencies as previously proposed[67]. To promote the assembly of larger-size oligomers in this regime, much higher protein concentrations would therefore be required. The salt dependence of oligomer size distribution is correlated well with the effect of ionic strength on phase separation, which is favored at conditions under which large protein assemblies are observed and is disfavored when the system enters a single continuous regime. In other words, the formation of large oligomeric clusters or microphase separation might be a mandatory step for nucleating eIF4B IDR condensation and observing a mesoscopic phase separation, as a result of coalescence of large oligomeric assemblies.

While the DRYG-rich region orchestrates the self-association of eIF4B, the CTR defines and fine-tunes the precise boundaries across the self-association landscape. Comparing DRYG-CTR constructs with varying CTR lengths demonstrates that the effect of CTR, likely mediated by charged interactions, is strongly correlated between the protein oligomerization and phase separation (Fig. 7). This property thus systematically couples the eIF4B oligomerization affinity with phase separation propensity. This is an interesting observation, as it identifies the CTR as a point of control over eIF4B self-association at a global scale. The sequence conservation analysis assigns high scores to several polypeptide segments within the CTR (Supplementary Fig. 17), which coincide well with the known phosphorylation and previously-mapped interaction sites[15,17,29,31]. Phosphorylation is a well-known modulator of protein phase separation[69–71], thus it is reasonable to expect that eIF4B phosphorylation at previously identified functionally relevant Ser406 and Ser422 sites[21] located in the CTR, may have a strong impact on protein self-association behavior. Similarly, self-association regulation could also be achieved by means of reentrant phase transition, for example upon RNA binding to CTR. The potential of such regulation is further supported by the observation that phase separation behavior of the eIF4B IDR is extremely sensitive in vitro. In addition to the pronounced effect of ionic strength on eIF4B self-association (Fig. 7c), the phase behavior is also very sensitive to molecular crowding (Fig. 7e). Even minimal crowding introduced by PEG strongly influences the phase diagram, counteracting the effect of ionic strength, and shifting the phase diagram towards a more physiologically relevant ionic strength regime. Notably, under these conditions the triple point of the phase diagram is positioned at close proximity to the estimated cellular eIF4B concentration and physiological ionic strength (Fig. 7e). Although beyond the scope of this work, such behavior supports our expectation that the monomer-oligomer-condensate transitions that we observe in vitro might also be relevant at the cellular level and be used as a response and coping mechanism to rapid changes of cellular microenvironment under stress conditions, such as variation of the intracellular ionic strength or crowding upon osmotic stress[72–75]. In this context, in-cell smFRET experiments[76] might open up future possibilities for monitoring self-association transitions (e.g. monomer-oligomer transitions) directly in live cells.

It remains to be determined how oligomerization of eIF4B is utilized functionally during translation initiation and stress granule formation. Beyond previously proposed hypotheses, such as eIF4B-dimerization dependent RNA annealing activity[54] or inhibition of interaction with eIF3[31], the formation of 'fuzzy' oligomers may be a sensitive and dynamic mechanism to enhance the effective binding affinity of the protein towards its targets, such as mRNA. Moreover, the pronounced reshaping and systematic expansion of the conformational ensemble of eIF4B IDR upon oligomerization (Fig. 5) may further enable oligomerization-driven tuning of the accessibility and affinity towards its binding partners. Additionally, the formation of large micelle-like protein clusters, occurring independently of assembly of cellular condensates, and their disassembly upon post-translational modifications could serve as a regulatory mechanism to maintain the functionally optimal concentration of 'active' eIF4B (i.e. monomers or small oligomers) within the cell. Our work thus provides detailed structural insights beyond simple dimerization of eIF4B, which can serve as a basis for future validation of these important hypotheses.

Methods

Protein expression and purification

A synthetic gene based on the protein sequence of the human eIF4B protein (Uniprot entry P23588) was purchased as codon-optimized DNA for *E. coli* expression (Supplementary Table 3) (Integrated DNA Technologies). All constructs were generated based on PCR amplification of specific regions, and inserted into the *SspI* site of the 2B-T modified pET vector (Addgene plasmid #29666) using the ligation-independent cloning protocol. The vector also adds an N-terminal hexahistidine (His$_6$) tag followed by a tobacco etch virus (TEV) protease cleavage site on the expressed proteins. PCR-based site-directed mutagenesis was performed to create double site-specific cysteine constructs and truncated variants (see Supplementary Table 4 for the list of primers used). The full list of constructs is presented in Supplementary Table 5.

For protein expression, colonies of freshly transformed *E. coli* T7 Express *lysY/I^q* (New England Biolabs) were first grown overnight in 10 mL of lysogeny broth (LB) with 50 µg/ml ampicillin, then transferred to 1 L cultures at 37 °C in TB or 500 ml in M9 minimal media containing 1 g/L [^{15}N]NH$_4$Cl. At an OD$_{600nm}$ = 1.8 for TB or OD$_{600nm}$ = 0.6 for M9, protein expression was induced with a final concentration of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) followed by overnight growth at 20 °C. Bacteria were harvested by centrifugation and then resuspended in a lysis buffer consisting of 20 mM Tris-HCl (pH 7.8),

250 mM KCl, 20 mM imidazole, 100 mg/L lysozyme, 1 mM PMSF, 2 mM β-mercaptoethanol, 0.5–2 M urea (depending on the construct) and one protease inhibitor cocktail tablet (Roche). Following sonication on ice (10 min at 40% power, alternating 50 s of sonication and 60 s of pause), the sample was centrifuged at 40,000 × g for 40 min at 4 °C to remove cellular debris. Specifically, for the DRYG-CTR-N variant, the bacteria were harvested by centrifugation and then resuspended in a Triton-containing lysis buffer consisting of 50 mM Tris-HCl (pH 7.8), 250 mM KCl, 1% Triton-X100, 2 mM β-mercaptoethanol, 1 M urea. Following sonication on ice (1 min at 40% power, alternating 20 s of sonication and 20 s of pause), the sample was centrifuged at 20,000 × g for 20 min at 4 °C to remove cellular debris. The supernatant was filtered with a 0.7 μm glass Microfilter GF/F (GE Healthcare Life Sciences Whatman) and added to NUVIA Ni^{2+} affinity chromatography resin (Bio-Rad) in a plastic column, washed first with five column volumes of binding buffer (20 mM Tris-HCl (pH 7.8), 250 mM KCl, 20 mM imidazole, 2 mM β-mercaptoethanol), then three column volumes with a higher salt concentration (800 mM of KCl), and finally five column volumes with increased imidazole concentration (30 mM). The protein was eluted with a buffer containing 20 mM Tris-HCl (pH 7.8), 250 mM KCl, 500 mM imidazole, 2 mM β-mercaptoethanol. The cleavage of the N-terminal His-tag for DRYG-CTR constructs was performed by TEV protease (15 μg/mg protein) during a dialysis step against 20 mM Tris-HCl (pH 7.8), 250 mM KCl, 5 mM β-mercaptoethanol, 0.5 M urea and 1% glycerol. The same procedure for CTR constructs was performed in 20 mM Tris-HCl (pH 7.8), 250 mM KCl, 5 mM β-mercaptoethanol buffer. The protease, His-tag and remaining uncleaved protein were removed by a second Ni^{2+} affinity chromatography step, followed by a purification with anion exchange chromatography with EnrichQ (Bio-rad), MonoQ, HiTrap Q or HiTrap DEAE (Cytiva) columns. In addition, prior to fluorescent labeling, protein samples were reduced by TCEP and were further purified with reverse-phase HPLC on a C18 column (ReproSil Gold 200, Dr. Maisch), followed by lyophilization, yielding samples of 2–5 mg dried protein depending on the size of the column used. Protein samples for NMR were used after anion exchange purification and were concentrated and buffer exchanged into corresponding NMR buffers.

Fluorescent labeling of proteins

For smFRET experiments, double cysteine constructs of proteins were fluorescently labeled with Cy3B (Cytiva) and CF660R (Biotium), as donor and acceptor dyes, respectively. The labeling was performed under denaturing conditions (100 mM NaP (pH 7.1), 6 M GdmCl), in two steps. First, the protein was incubated with the Cy3B maleimide at a protein to dye ratio 1:0.7, for 1 h at room temperature, followed by purification of the labeling mixture with reverse-phase HPLC, on a C18 column (ReproSil Gold 200, Dr. Maisch). This procedure enabled separation of unlabeled, single-labeled (in most of the cases as two separate single site-specifically labeled protein permutants) and double-labeled protein fractions, as well as the small fraction of unreacted/inactive dye. All corresponding protein fractions were collected individually and were lyophilized overnight. In a second step, one of the single labeled protein fractions (corresponding to a particular site-specifically donor-labeled protein) was further labeled with CF660R maleimide at a protein to dye ratio 1:3, for 2–3 h at room temperature, followed by purification of the labeling mixture with reverse-phase HPLC, on a C18 column (ReproSil Gold 200, Dr. Maisch). This procedure enabled the separation of the single-labeled and double-labeled protein fractions, as well as the excess of the unreacted acceptor dye. After these rigorous labeling and purification steps, it was possible to obtain predominantly double labeled protein, in most cases with site-specific attachment of donor and acceptor dyes. For labeling of DRYG-CTR$_{P213C-P332C}$ construct, alternatively, ATTO 532 maleimide (ATTO-TEC) and LD655 maleimide (Lumidyne Technologies) were used as donor and acceptor, respectively. For FCS

experiments, the proteins were single-labeled with ATTO 655 maleimide (ATTO-TEC) dye and purified using a similar procedure as described above.

Phase separation assays

For phase separation experiments the lyophilized protein was dissolved in 20 mM NaP (pH 7.0), 1 mM TCEP buffer containing 100–300 mM NaCl. The dissolved protein solution was then repeatedly buffer exchanged with a centrifugal filter (Amicon, Merck) against the same buffer to eliminate the residual TFA, remaining after reverse-phase HPLC and lyophilization and was concentrated up to ~1 mM, to yield concentrated one-phase protein solution. The DRYG construct was dissolved and concentrated in double-distilled water and was used as it is.

The light phase or saturation concentrations, c_{sat}, as well as the dense phase concentrations, c_{dense}, were obtained by the centrifugation method[33]. In brief, at respective NaCl concentrations, a protein solution was prepared under conditions favoring phase separation, which was confirmed by solution turbidity and observation of droplets under the microscope. The phase separated solution was then incubated at respective temperatures in a thermocycler (Mastercycler nexus, Eppendorf) for 20 min, followed by centrifugation at 20,500 × g for 5 min at respective temperatures, using a temperature-controlled centrifuge (Eppendorf 5430 R or 5425 R). After 5 min, the dense phase is settled as a pellet and the light phase remains as supernatant. A portion of supernatant was collected without disturbing the pellet and was diluted in 6 M GdmCl. Similarly, 1 μl of dense phase was collected using a positive displacement pipette (Microman E, Gilson) and was also diluted in 6 M GdmCl. The protein concentrations were measured spectrophotometrically (Cary 60, Agilent Technologies) with a 1 mm pathlength tray cell (Hellma) or 1 cm quartz cuvettes (Hellma). The c_{sat} and c_{dense} values were then estimated with consideration of the dilution factors. Each data point was repeated at least three times. The c_{sat} and c_{dense} values were found to be independent of the initial volumes of the two-phase solution as well as the centrifugation time and speed. For phase separation assays in the presence of a crowding agent, PEG 8000 (P1458, Sigma-Aldrich) was used.

Differential interference contrast (DIC) and fluorescence imaging

DIC images were acquired using an IX83 (Olympus) brightfield microscope equipped with CMOS camera (Thorlabs), polarizer (IX-LWPO, Olympus), analyzer (IX3-AN, Olympus), Nomarski prism (IX2-DIC60, Olympus) and a water immersion objective (UplanSApo 60x/1.20 W, Olympus). The fluorescence images were acquired using a MicroTime 200 confocal microscope (PicoQuant), employing a 520 nm laser (LDH-D-C-520; PicoQuant) operating in a continuous-wave mode and a XY piezo scanning stage (PI). The solutions with phase separated DRYG-CTR droplets were prepared at 5 μM protein concentration in 20 mM NaP (pH 7.0), 50 mM NaCl, 1 mM TCEP and were imaged in ibidi multiwell chambers (μ-Slide 15 Well 3D, ibidi). For fluorescence imaging the solution was doped with 20 nM of Cy3B-labeled DRYG-CTR.

CD spectroscopy

The CD spectra were recorded with a Jasco J-1500 CD spectrometer (using Jasco Spectra Manager Suite software) equipped with a temperature-controlled cuvette holder, using 0.1 or 1 mm path-length cuvettes at 20 °C. The CD spectra of DRYG-CTR were measured in a buffer containing 20 mM NaP (pH 7.0), 75 mM NaCl or 300 mM NaF, 1 mM TCEP. The CD spectra of CTR, CTR-N, CTR-C were measured in a buffer containing 20 mM NaP (pH 7.0), 150 mM NaF, 1 mM TCEP. All CD spectra were recorded between 190 and 250 nm, with data pitch 1 nm, scan speed 50 nm/min and accumulation of 10 scans. Before recording the spectra, baseline correction was performed with the respective buffers. CD data was analyzed using Origin 2021 software.

NMR spectroscopy

NMR data were collected on a Bruker Avance 700 or 800 MHz spectrometer (Bruker TopSpin version 4.0 software), equipped with triple-resonance gradient room-temperature or cryogenic probe, respectively. NMR data were processed by using NMRPipe/NMRDraw software[77] and NMR spectra were analyzed using Sparky (T. D. Goddard & D. G. Kneller, University of California, San Francisco, USA). Samples contained 170 μL in a 3 mm NMR tube at 293 K, in a buffer containing 20 mM NaP (pH 7.0), 2 mM DTT, and varying amounts of NaCl and GdmCl. The spectra correspond to sensitivity-enhanced ^{15}N-HSQC with water flip-back suppression using selective 90° gauss pulses. The direct dimension (^{1}H) contained 1024 complex points and a sweep width of 9615 Hz centred on the water frequency (4.814 ppm). The indirect dimension (^{15}N) used 512 complex points and a sweep width of 2500 Hz centred on 117.623 ppm. Chemical shift assignments for the CTR construct of human eIF4B were taken from BMRB entry 51957 that was determined for the same buffer conditions and measurement temperature as for the DRYG-CTR spectrum[28], and this allowed for straightforward transfer of the chemical shift assignments for peaks without significant line-broadening.

Single-molecule fluorescence spectroscopy

Single-molecule fluorescence measurements were performed using a MicroTime 200 confocal microscope (PicoQuant) using a Sympho-Time 64 software. Alternating excitation of the dyes was achieved using pulsed interleaved excitation (PIE)[78], by synchronizing 520 nm (LDH-D-C-520; PicoQuant) and 638 nm (LDH-D-C-640; PicoQuant) diode lasers at 20 MHz pulse rate with a delay time of 25 ns between the two laser pulses. The power of 520 nm laser excitation was adjusted to 100 μW (measured at the back aperture of the microscope objective), whereas the direct acceptor 638 nm excitation was adjusted to ~50 μW (such that the donor and acceptor emission intensities were matching after donor and direct acceptor excitations, respectively). The lasers were focused on the sample through a water immersion objective (UplanSApo 60x/1.20 W, Olympus), 50 μm in the solution above the cover slide surface. The emitted photons from the sample were collected by the same objective, and were separated from the excitation photons and scattered light by a major dichroic mirror (ZT532/640rpc, Chroma) and 100 μm pinhole (Thorlabs), respectively. The remaining photons were then detected onto four detection channels (two donor and two acceptor) based on their polarization (using a polarizing beam splitter) and wavelength by two dichroic mirrors (T635lpxr, Chroma). The donor and acceptor photons were then further filtered by bandpass filters (ET585/65 and H690/70, Chroma) before being focused onto the SPAD detectors (SPCM AQRH-14-TR, Excelitas Technologies). The arrival time of each photon was recorded with a HydraHarp 400 TCSPC module (PicoQuant) with 16 ps time resolution.

All smFRET measurements were performed with 50–100 pM of double labeled protein solution in 20 mM NaP (pH 7.0) buffer at corresponding concentrations of NaCl or GdmCl, supplemented with 0.01% Tween 20 and 143 mM β-mercaptoethanol. The smFRET self-association assays were performed under the same solution conditions, but with increasing concentrations of unlabeled proteins. Each sample was measured at room temperature (20 ± 1 °C) with acquisition time of 15–30 mins, each dataset was repeated at least once. All measurements of DRYG-CTR constructs were performed in ibidi multiwell chambers with uncoated polymer coverslip (μ-Slide 15 Well 3D, ibidi). For measurements of DRYG-CTR at low NaCl (<100 mM), the ibidi chambers were additionally coated with PLL-PEG (SUSOS, Switzerland). PLL-PEG coating was achieved by incubation of wells with 0.5 g/L solution of PLL-PEG for 10 min, followed by a wash with distilled water and respective measurement buffer. All measurements for CTR constructs were performed in custom-made PEG coated glass cuvettes, which were prepared according to previously reported protocol[79].

smFRET data analysis was performed using the Fretica package for Mathematica 13.0/14.0 (Wolfram Research Inc, USA) developed by Daniel Nettels and Benjamin Schuler (https://schuler.bioc.uzh.ch/programs/). Briefly, fluorescence bursts were identified by the ΔT method, with interphoton time of 150 μs and a minimum number of 60–80 photons per burst. Transfer efficiencies for each burst were calculated according to:

$$E = \frac{n'_A}{n'_D + n'_A} \tag{1}$$

where n'_D and n'_A are the donor and acceptor counts (i.e. numbers of photons), respectively, corrected for background, channel crosstalk, acceptor direct excitation, differences in detector efficiencies, and quantum yields of the dyes, as previously described[80,81]. The stoichiometry ratio was calculated according to:

$$S = \frac{n_{D-D} + n_{A-D}}{n_{D-D} + n_{A-D} + \gamma_{PIE} \cdot n_{A-A}} \tag{2}$$

where n_{D-D} and n_{A-D} are the background corrected total donor and acceptor counts, respectively, after donor excitation (520 nm laser) and n_{A-A} is the total acceptor counts after acceptor direct excitation (638 nm laser). γ_{PIE} is a correction factor used for weighing the n_{A-A} to account for differences in 520 and 638 nm laser intensities. The final transfer efficiency histograms were built based on bursts originating from molecules containing both donor-acceptor fluorophores (thus eliminating molecules with photobleached acceptor or donor, i.e. donor-only and acceptor-only sub-populations). These bursts were selected based on stoichiometry ratio values $0.2 < S < 0.8$.

To determine the mean transfer efficiencies for different sub-populations, the transfer efficiency histograms were fitted by a sum of 2–3 Gaussian peak functions. When analyzing histogram series, a global fitting was used with some fitting variables (e.g. width or position) being shared within the same histogram series. The small peak at $E \approx 0$, sometimes appearing at transfer efficiency histograms, originates from the small fraction of photobleaching and blinking molecules still remaining after the PIE-cut as well as possible residual fluorescent impurities of the buffer.

In the analysis for lifetime - transfer efficiency histograms the donor lifetimes were estimated from mean donor detection times, following the corresponding excitation pulse. Alternatively, the mean fluorescence lifetimes of donor and acceptor dyes (shown in Supplementary Table 1) were obtained by fitting of time-correlated single-photon counting histograms corresponding to donor-only and acceptor-only molecules, respectively. In all smFRET measurements the donor and acceptor anisotropies are low ($r < 0.2$), indicating the free rotation of dyes.

Conversion of mean transfer efficiency to mean distances

For conversion of mean transfer efficiency into mean inter-residue (donor-acceptor) distances, the following equation was numerically solved:

$$\langle E \rangle = \int_0^{l_c} P(r)E(r)dr \tag{3}$$

where the l_c is the contour length of the probed polypeptide chain segment, and

$$E(r) = \frac{R_0^6}{R_0^6 + r^6} \tag{4}$$

with the R_0 as Förster radius at particular buffer conditions. $R_0 = 5.9$ nm was calculated for Cy3B/CF660R dye pair, with the solution refractive

index equal to 1.333, which was further scaled to account for differences in the refractive index values of particular buffers. Solution refractive index values were measured using temperature-controlled digital Abbe refractometer (DR-A1-Plus, Atago) at 20 °C.

For the $P(r)$ distance distribution we used the SAW-ν model[82], which adopts the following form:

$$P(r) = A_1 \frac{4\pi}{R_{DA}} \left(\frac{r}{R_{DA}}\right)^{2+g} \exp\left[-A_2\left(\frac{r}{R_{DA}}\right)^{\delta}\right] \quad (5)$$

where, $A_1 = \frac{\delta}{4\pi} \frac{\Gamma[5+g/\delta]^{\frac{3+g}{2}}}{\Gamma[3+g/\delta]^{\frac{5+g}{2}}}$, $A_2 = \left(\frac{\Gamma[5+g/\delta]}{\Gamma[3+g/\delta]}\right)^{\frac{\delta}{2}}$, $g = \frac{\gamma-1}{\nu}$, with $\gamma = 1.1615$, $\delta = \frac{1}{1-\nu}$, Γ is the gamma function and ν is the scaling exponent. Equation 5 has two adjustable parameters, the scaling exponent ν and the mean inter-residue distance, R_{DA}, (i.e. distance between the two dyes), which equates to the root mean squared end-to-end distance, $\langle r^2 \rangle^{1/2}$. The mean end-to-end distance for unfolded proteins approximately follows the scaling law:

$$R_{DA} = bN^{\nu} \quad (6)$$

where the $b = \sqrt{2l_p b_0}$ is an empirical prefactor, l_p is the persistence length, $b_0 = 0.38$ nm is the peptide bond length, and N is the number of amino acids between the two dyes (including the contribution of the dye linkers, with each approximated as $l = 4.5$ peptide bond). The prefactor b has been previously estimated to be equal to approximately 0.55 nm for proteins (with $l_p = 0.4$ nm). With this approximation, it is possible to solve Eq. 5 for both R_{DA} and ν.

Nanosecond fluorescence correlation spectroscopy (nsFCS) measurements

The nsFCS measurements were carried out with the MicroTime 200 confocal microscope (PicoQuant) with the 520 nm diode laser (LDH-D-C-520; PicoQuant) operating in continuous-wave mode at a laser power of 100 μW. The emitted photons (after the pinhole) were split by a 50/50 beam splitter before being separated and filtered based on wavelength and detected at respective donor and acceptor detectors. The typical data acquisition times for nsFCS measurements were 16–20 h. All measurements were performed at protein concentrations of ~1–2 nM, in 20 mM NaP (pH 7.0), 0.01% Tween-20, 143 mM β-mercaptoethanol at respective NaCl/GdmCl concentrations. The measurements at each condition were repeated 1–2 times. The auto-correlation curves of donor and acceptor channels and cross-correlation curves between donor and acceptor channels were calculated from the measurements following the previously described methods[37,83]. The calculated correlation curves (for delay times up to 4 μs) were fitted with the following equation:

$$g_{ij}(\tau) = 1 + \frac{1}{\langle N \rangle}\left(1 - c_{AB}e^{-\frac{\tau-t_0}{\tau_{AB}}}\right)\left(1 + c_{CD}e^{-\frac{\tau-t_0}{\tau_{CD}}}\right)\left(1 + c_T e^{-\frac{\tau-t_0}{\tau_T}}\right) \quad (7)$$

where i and j correspond to the type of signal (e.g. D or A, corresponding to the signal either from the donor or acceptor channels, respectively) and $\langle N \rangle$ is the mean number of molecules in the confocal volume. The three multiplicative terms describe the contributions of photon antibunching (AB), chain dynamics (CD), and triplet blinking (T) of the dyes, with corresponding amplitudes, c, and relaxation times, τ, respectively. τ_{CD} was then converted in the reconfiguration time τ_r of the interdye segment as previously described[84]. An additional multiplicative term $\left(1 + c'e^{-\frac{\tau-t_0}{\tau'}}\right)$ was used for D-D and A-A auto-correlations to describe the fast decay observed on a short timescale (<10 ns). The timescale, τ', for this decay is 6–7 ns and it can be associated with the

rotational motion/tumbling of the overall protein. The same multiplicative term was also included for D-A cross-correlation for fitting the data of DRYG-CTR$_{P213C-P332C}$ construct in the presence of NaCl. The three correlation curves from each measurement were fit globally with the time origin, t_0, and τ' as shared fit parameters.

Fluorescence correlation spectroscopy (FCS) measurements

The FCS measurements were performed with the MicroTime 200 confocal microscope with the 638 nm diode laser (LDH-D-C-640; PicoQuant or LBX-638; Oxxius) operating in continuous-wave mode at a laser power of 50 μW. The emitted photons (after the 100 μm pinhole) were split by a 50/50 beam splitter and detected on two detectors. The typical data acquisition times for FCS measurements were 120 s with each sample measured 5 times and each experiment was repeated at least once. All measurements were performed at a protein concentration of ~250 pM, in 20 mM NaP (pH 7.0), 0.01% Tween-20, 143 mM β-mercaptoethanol at respective NaCl concentrations, in the absence and presence of unlabeled protein. The correlation curves were fitted with using a three-dimensional free diffusion model (Eq. 8) with one diffusional component and an additional component for triplet blinking:

$$G(\tau) = 1 + \frac{1}{\langle N \rangle}\left(1 + \frac{\tau-t_0}{\tau_D}\right)^{-1}\left(1 + s^2\frac{\tau-t_0}{\tau_D}\right)^{-1/2}\left(1 + c_T e^{-\frac{\tau-t_0}{\tau_T}}\right) \quad (8)$$

where $\langle N \rangle$ is the mean number of molecules in the confocal volume, c_T is the amplitude of triplet component, τ_D and τ_T are the diffusion time and triplet blinking relaxation time, respectively, and s is the axial ratio of the confocal volume. The relative diffusion rates (D/D_0) were calculated based on the $\tau_D \propto D^{-1}$ relationship.

The curves corresponding to the dependence of relative diffusion on protein concentration were fitted globally with the Hill equation, with n being shared among all datasets:

$$y = 1 - \Delta D \frac{x^n}{k^n + x^n} \quad (9)$$

where $k \equiv K_D$ is the apparent dissociation constant and n is the cooperativity coefficient.

The oligomer size, N, shown on the right axis of Fig. 4c and Supplementary Fig. 15a, b was estimated based on the $\frac{D}{D_0} = \frac{1}{(N \cdot c)^{1/3}}$ relationship, where $c = 1.4725$ is a factor accounting for ~15% increase of the CTR radius of gyration upon oligomerization (Supplementary Fig. 18). We do not account for possible changes of dimensions of DRYG-rich region associated with oligomerization, considering that this region is packed within the cores of oligomers and will likely have smaller contribution to overall oligomer size (as compared to the contribution of CTR).

Recurrence analysis of single particles (RASP) measurements

The RASP[45] experiments were performed with the MicroTime 200 confocal microscope (PicoQuant) with the 520 nm (LDH-D-C-520; PicoQuant) and 640 nm (LDH-D-C-640; PicoQuant) lasers operating in PIE mode with laser powers adjusted to 100 and 50 μW, respectively. Samples consisted of 25–50 pM double-labeled DRYG-CTR with respective μM concentration of unlabeled protein, prepared in 20 mM NaP (pH 7.0), 75 mM NaCl, 0.01% Tween-20, 143 mM β-mercaptoethanol buffer. Each sample was typically acquired for ~20 h, to ensure sufficient burst statistics. Photon bursts were first identified by combining successive photons that were separated by less than 150 μs and only bursts containing >60 photons were retained for the further analysis. These bursts were then time-binned with a binning time of 1 ms and only the time intervals containing more than 40 photons were used for the recurrence analysis. To generate recurrence histograms first the initial transfer efficiency ranges of $E \in$

[0.15–0.25] and $E \in$ [0.4–0.55] were selected to define the bound and unbound populations, respectively. The burst list retained for recurrence analysis containing at least 200,000 bursts per experiment, was then scanned for burst pairs for which the first burst matched the given E range (either bound or unbound E range), followed by a second burst detected within a given time interval (i.e. recurrence interval). From the collection of second bursts, 2 sets of recurrence transfer efficiencies histograms were built corresponding to first bursts originating from either unbound or bound population, for recurrence time interval of 1–50 ms with a step of 1 ms. The resulting recurrence histograms were then fit with a sum of 2 Gaussian peak functions to describe the unbound and bound FRET populations. The fraction of each population was calculated ratiometrically based on the area under each peak and was plotted as a function of the recurrence time. The resulting curves were fitted with the following equation:

$$f(t) = (1 - p_{same}(t))f_{eq} + p_{same}(t)\left(f_{eq.} + \left(f_0 - f_{eq}\right)e^{-k^{obs}\cdot t}\right) \quad (10)$$

where f_{eq} is the transfer efficiency at equilibrium (i.e. $t = \infty$), f_0 is the fraction of either bound or unbound population at $t = 0$ ms, k^{obs} is the observed binding or unbinding rate (i.e. $k_{on}^{obs.}$ or $k_{off}^{obs.}$, respectively) and $p_{same}(t)$ is the probability that the second burst originates from the same molecule as the first burst, which can be independently quantified from the data[45]. Each experiment was repeated three times, the resulting kinetic curves were individually fitted and the resulting rate coefficients were averaged to estimate the mean rates and associated standard deviation.

Coarse-grained Langevin dynamics simulations

The CG Langevin dynamics simulations of the DRYG-CTR and CTR constructs were carried out using OpenMM version 7.5[85]. Each amino acid was represented as a single bead mapped to the C_α carbon of each residue. Molecular interactions between beads were accounted for as bonded or non-bonded. Bonded interactions were considered for two beads directly linked by a harmonic potential, with an equilibrium bond length of 0.38 nm and a force constant of 8033 kJ mol^{-1} nm^{-2}. Nonbonded interactions were modeled for all beads not directly bonded using the interaction potentials implemented in CALVADOS2[86,87]. Particularly, dispersion forces were accounted for by a truncated and shifted Ashbaugh-Hatch potential, in which a residue-specific hydrophobicity parameter λ, scales interaction strengths between beads. Electrostatic forces acting between charged particles were accounted for, as in implicit solvent, by a salt-screening Debye-Hückel based potential, truncated at a cut-off distance of 2 nm. In this model, beads representing aspartate and glutamate carry a -1 charge, while lysine and arginine beads carry a $+1$ charge.

Single chains of DRYG-CTR and CTR were each modeled as linear chains and placed in a box with dimensions 100 nm^3. To simulate the oligomeric systems, we followed the slab method suggested to well simulate IDPs in the condensed phase[50]. 50 chains of the DRYG-CTR construct were modeled as linear chains, equally spaced and placed in a triclinic box measuring $25 \times 25 \times 200$ nm along x, y and z, respectively. This corresponded to a total protein concentration of 664 μM. Full simulation details are available in Supplementary Table 6. All systems were simulated considering periodic boundary conditions and at two ionic strengths: 117 and 192 mM. After energy minimization, each system was simulated at a temperature of 293 K, with a time step of 0.01 ps and a friction coefficient of 0.01 ps^{-1}, saving coordinates every 1 ns. Each system was simulated for a total of 10 μs in triplicates (30 μs total sampling) with the first μs of each replicate discarded as considered equilibration time.

Using the Förster equation (Eq. 4), transfer efficiencies were calculated from the simulated distances between the experimentally labeled residues along the protein, implementing a Förster radius of 5.9 nm, reflective of the R_0 of the Cy3B/CF660R dye-pair used in the smFRET experiments. All the simulations were performed without explicitly modeling FRET dyes, as those were not parameterized within the used coarse-grained model. Nevertheless, the effect of dyes on the retrieved ensembles has been shown in multiple studies[46,47,49] to be negligible. The agreement between experimental and computed FRET efficiencies was obtained by calculating the concordance correlation coefficient of the two datasets as follows:

$$\rho_c = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\mu_x - \mu_y)^2} \quad (11)$$

where μ_x and μ_y are the means of the experimental and computational FRET efficiencies, respectively, while σ_x and σ_y are the standard deviations of each dataset.

The collected trajectories were analyzed using a combination of *in-house* scripts, MDAnalysis libraries[88] and tools available in GROMACS version 2021.6[89]. Clustering was performed using gmx clustsize, with a distance cutoff of 1 nm. For each simulation step, two or more chains were considered bound if the distance between two or more beads fell within 1 nm. Intramolecular and intermolecular contacts were calculated from the single-chain and oligomeric systems, respectively. Intermolecular contacts were obtained by firstly calculating the average distance between all beads of a bound cluster: distance arrays were then binarized to create a Boolean contact matrix (1 if bound, 0 if not bound) considering a cutoff of 1 nm and the final fraction of residue-wide contacts was calculated as a time-average over the simulated trajectory. The map of the fraction of contacts gained for oligomers (shown in Fig. 6f) was quantified by subtracting the intrachain contacts sampled in the single chain simulations from the interchain contacts sampled in the oligomeric state. Representative snapshots of the simulated ensembles were obtained using VMD version 1.9.3[90]. Unit root statistical tests to assess timeseries stationarity have been performed through an augmented Dickey-Fuller test. Test statistics are reported in Supplementary Table 7.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data supporting the findings of this work are available within the main text and its Supplementary Information. The gene sequence of eIF4B used in this work is provided in Supplementary Table 3. The list of oligonucleotide primers used in this work is provided in Supplementary Table 4. The list of sequences for protein constructs used in this work is provided in Supplementary Table 5. The raw smFRET data are available from the corresponding authors upon request. The chemical shift assignments for eIF4B CTR were previously deposited in the BMRB under accession code 51957. The CG simulation trajectories are available from: https://doi.org/10.5281/zenodo.13370824. Source data are provided with this paper.

Code availability

Fretica package for Mathematica v.14 (Wolfram Research) used for the analysis of single-molecule data is available at https://schuler.bioc.uzh.ch/fretica/.

References

1. Jackson, R. J., Hellen, C. U. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 113–127 (2010).
2. Hinnebusch, A. G. The scanning mechanism of eukaryotic translation initiation. *Annu Rev. Biochem* **83**, 779–812 (2014).
3. Merrick, W. C. eIF4F: a retrospective. *J. Biol. Chem.* **290**, 24091–24099 (2015).

4. Aitken, C. E. & Lorsch, J. R. A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol*. **19**, 568–576 (2012).

5. Sokabe, M. & Fraser, C. S. Toward a kinetic understanding of eukaryotic translation. *Cold Spring Harb Perspect Biol*. **11**, a032706 (2019).

6. Brito Querido, J. et al. Structure of a human 48S translational initiation complex. *Science* **369**, 1220–1227 (2020).

7. Brito Querido, J., Diaz-Lopez, I. & Ramakrishnan, V. The molecular basis of translation initiation and its regulation in eukaryotes. *Nat. Rev. Mol. Cell Biol*. **25**, 168–186 (2024).

8. Brito Querido, J. et al. The structure of a human translation initiation complex reveals two independent roles for the helicase eIF4A. *Nat. Struct. Mol. Biol*. **31**, 455–464 (2024).

9. Hashem, Y. & Frank, J. The jigsaw puzzle of mRNA translation initiation in Eukaryotes: a decade of structures unraveling the mechanics of the process. *Annu Rev. Biophys*. **47**, 125–151 (2018).

10. Simonetti, A., Guca, E., Bochler, A., Kuhn, L. & Hashem, Y. Structural insights into the mammalian late-stage initiation complexes. *Cell Rep*. **31**, 107497 (2020).

11. des Georges, A. et al. Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* **525**, 491–495 (2015).

12. Eliseev, B. et al. Structure of a human cap-dependent 48S translation pre-initiation complex. *Nucleic Acids Res*. **46**, 2678–2689 (2018).

13. Ozes, A. R., Feoktistova, K., Avanzino, B. C. & Fraser, C. S. Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B. *J. Mol. Biol*. **412**, 674–687 (2011).

14. Nielsen, K. H. et al. Synergistic activation of eIF4A by eIF4B and eIF4G. *Nucleic Acids Res*. **39**, 2678–2689 (2011).

15. Rozovsky, N., Butterworth, A. C. & Moore, M. J. Interactions between eIF4AI and its accessory factors eIF4B and eIF4H. *RNA* **14**, 2136–2148 (2008).

16. Garcia-Garcia, C., Frieda, K. L., Feoktistova, K., Fraser, C. S. & Block, S. M. RNA BIOCHEMISTRY. Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science* **348**, 1486–1488 (2015).

17. Methot, N., Pause, A., Hershey, J. W. & Sonenberg, N. The translation initiation factor eIF-4B contains an RNA-binding region that is distinct and independent from its ribonucleoprotein consensus sequence. *Mol. Cell Biol*. **14**, 2307–2316 (1994).

18. Rogers, G. W. Jr., Richter, N. J., Lima, W. F. & Merrick, W. C. Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J. Biol. Chem*. **276**, 30914–30922 (2001).

19. Shahbazian, D. et al. Control of cell survival and proliferation by mammalian eukaryotic initiation factor 4B. *Mol. Cell Biol*. **30**, 1478–1485 (2010).

20. Wang, Y. et al. Mitotic MELK-eIF4B signaling controls protein synthesis and tumor cell survival. *Proc. Natl Acad. Sci. USA* **113**, 9810–9815 (2016).

21. Shahbazian, D., Parsyan, A., Petroulakis, E., Hershey, J. & Sonenberg, N. eIF4B controls survival and proliferation and is regulated by proto-oncogenic signaling pathways. *Cell Cycle* **9**, 4106–4109 (2010).

22. Sen, N. D., Zhou, F., Harris, M. S., Ingolia, N. T. & Hinnebusch, A. G. eIF4B stimulates translation of long mRNAs with structured 5' UTRs and low closed-loop potential but weak dependence on eIF4G. *Proc. Natl Acad. Sci. USA* **113**, 10464–10472 (2016).

23. Horvilleur, E. et al. A role for eukaryotic initiation factor 4B over-expression in the pathogenesis of diffuse large B-cell lymphoma. *Leukemia* **28**, 1092–1102 (2014).

24. Buchan, J. R. & Parker, R. Eukaryotic stress granules: the ins and outs of translation. *Mol. Cell* **36**, 932–941 (2009).

25. Low, W. K. et al. Inhibition of eukaryotic translation initiation by the marine natural product pateamine A. *Mol. Cell* **20**, 709–722 (2005).

26. Mokas, S. et al. Uncoupling stress granule assembly and translation initiation inhibition. *Mol. Biol. Cell* **20**, 2673–2683 (2009).

27. Fleming, K. et al. Solution structure and RNA interactions of the RNA recognition motif from eukaryotic translation initiation factor 4B. *Biochemistry* **42**, 8966–8975 (2003).

28. Mondal, S. et al. Backbone resonance assignments of the C-terminal region of human translation initiation factor eIF4B. *Biomol. NMR Assign*. **17**, 199–203 (2023).

29. Naranda, T., Strong, W. B., Menaya, J., Fabbri, B. J. & Hershey, J. W. Two structural domains of initiation factor eIF-4B are involved in binding to RNA. *J. Biol. Chem*. **269**, 14465–14472 (1994).

30. Milburn, S. C., Hershey, J. W., Davies, M. V., Kelleher, K. & Kaufman, R. J. Cloning and expression of eukaryotic initiation factor 4B cDNA: sequence determination identifies a common RNA recognition motif. *EMBO J*. **9**, 2783–2790 (1990).

31. Methot, N., Song, M. S. & Sonenberg, N. A region rich in aspartic acid, arginine, tyrosine, and glycine (DRYG) mediates eukaryotic initiation factor 4B (eIF4B) self-association and interaction with eIF3. *Mol. Cell Biol*. **16**, 5328–5334 (1996).

32. Raught, B. et al. Phosphorylation of eucaryotic translation initiation factor 4B Ser422 is modulated by S6 kinases. *EMBO J*. **23**, 1761–1769 (2004).

33. Milkovic, N. M. & Mittag, T. Determination of protein phase diagrams by centrifugation. *Methods Mol. Biol*. **2141**, 685–702 (2020).

34. 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).

35. Aznauryan, M. et al. Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS. *Proc. Natl Acad. Sci. USA* **113**, E5389–E5398 (2016).

36. 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).

37. Schuler, B. Perspective: chain dynamics of unfolded and intrinsically disordered proteins from nanosecond fluorescence correlation spectroscopy combined with single-molecule FRET. *J. Chem. Phys*. **149**, 010901 (2018).

38. Krainer, G. et al. Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun*. **12**, 1085 (2021).

39. Hazra, M. K. & Levy, Y. Cross-talk of cation-π interactions with electrostatic and aromatic interactions: a salt-dependent trade-off in biomolecular condensates. *J. Phys. Chem. Lett*. **14**, 8460–8469 (2023).

40. MacAinsh, M., Dey, S. & Zhou, H.-X. Direct and indirect salt effects on homotypic phase separation. *bioRxiv*, https://doi.org/10.1101/2024.05.26.596000 (2024).

41. Martin, E. W. et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).

42. Martin, E. W. et al. Sequence determinants of the conformational properties of an intrinsically disordered protein prior to and upon multisite phosphorylation. *J. Am. Chem. Soc*. **138**, 15323–15335 (2016).

43. Reiersen, H. & Rees, A. R. The hunchback and its neighbours: proline as an environmental modulator. *Trends Biochem Sci*. **26**, 679–684 (2001).

44. 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).

45. Hoffmann, A. et al. Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Phys. Chem. Chem. Phys*. **13**, 1857–1871 (2011).

46. Borgia, A. et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61–66 (2018).

47. Bjarnason, S. et al. DNA binding redistributes activation domain ensemble and accessibility in pioneer factor Sox2. *Nat. Commun*. **15**, 1445 (2024).

48. Holmstrom, E. D., Liu, Z., Nettels, D., Best, R. B. & Schuler, B. Disordered RNA chaperones can enhance nucleic acid folding via local charge screening. *Nat. Commun.* **10**, 2453 (2019).

49. Heidarsson, P. O. et al. Release of linker histone from the nucleosome driven by polyelectrolyte competition with a disordered protein. *Nat. Chem.* **14**, 224–231 (2022).

50. Dignon, G. L., Zheng, W., Best, R. B., Kim, Y. C. & Mittal, J. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA* **115**, 9929–9934 (2018).

51. Guillen-Boixet, J. et al. RNA-induced conformational switching and clustering of G3BP drive stress granule assembly by condensation. *Cell* **181**, 346–361 e317 (2020).

52. Bremer, A. et al. Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. *Nat. Chem.* **14**, 196–207 (2022).

53. Dao, T. P., Rajendran, A., Galagedera, S. K. K., Haws, W. & Castaneda, C. A. Short disordered termini and proline-rich domain are major regulators of UBQLN1/2/4 phase separation. *Biophys. J.* **123**, 1449–1457 (2024).

54. Altmann, M., Wittmer, B., Methot, N., Sonenberg, N. & Trachsel, H. The Saccharomyces cerevisiae translation initiation factor Tif3 and its mammalian homologue, eIF-4B, have RNA annealing activity. *EMBO J.* **14**, 3820–3827 (1995).

55. Cherkasov, V. et al. Systemic control of protein synthesis through sequestration of translation and ribosome biogenesis factors during severe heat stress. *FEBS Lett.* **589**, 3654–3664 (2015).

56. Wallace, EdwardW. J. et al. Reversible, specific, active aggregates of endogenous proteins assemble upon heat stress. *Cell* **162**, 1286–1298 (2015).

57. Desroches Altamirano, C. et al. eIF4F is a thermo-sensing regulatory node in the translational heat shock response. *Mol. Cell* **84**, 1727–1741.e1712 (2024).

58. Park, E.-H. et al. Yeast Eukaryotic initiation factor 4B (eIF4B) Enhances Complex Assembly between eIF4A and eIF4G in Vivo*. *J. Biol. Chem.* **288**, 2340–2354 (2013).

59. Marzahn, M. R. et al. Higher-order oligomerization promotes localization of SPOP to liquid nuclear speckles. *EMBO J.* **35**, 1254–1275 (2016).

60. Dao, T. P. et al. Ubiquitin modulates liquid-liquid phase separation of UBQLN2 via disruption of multivalent interactions. *Mol. Cell* **69**, 965–978 e966 (2018).

61. Kar, M. et al. Phase-separating RNA-binding proteins form heterogeneous distributions of clusters in subsaturated solutions. *Proc. Natl Acad. Sci. USA* **119**, e2202222119 (2022).

62. Zhao, H. et al. Energetic and structural features of SARS-CoV-2 N-protein co-assemblies with nucleic acids. *iScience* **24**, 102523 (2021).

63. Ray, S. et al. Mass photometric detection and quantification of nanoscale alpha-synuclein phase separation. *Nat. Chem.* **15**, 1306–1316 (2023).

64. Kotaka, T., Tanaka, T., Hattori, M. & Inagaki, H. Block copolymer micelles in dilute solution. *Macromolecules* **11**, 138–145 (1978).

65. Riess, G. Micellization of block copolymers. *Prog. Polym. Sci.* **28**, 1107–1170 (2003).

66. Marko, J. F. & Rabin, Y. Microphase separation of charged diblock copolymers: melts and solutions. *Macromolecules* **25**, 1503–1509 (1992).

67. Ranganathan, S. & Shakhnovich, E. I. Dynamic metastable long-living droplets formed by sticker-spacer proteins. *eLife* **9**, e56159 (2020).

68. Yamazaki, T. et al. Paraspeckles are constructed as block copolymer micelles. *EMBO J.* **40**, e107270 (2021).

69. Nosella, M. L. & Forman-Kay, J. D. Phosphorylation-dependent regulation of messenger RNA transcription, processing and translation within biomolecular condensates. *Curr. Opin. Cell Biol.* **69**, 30–40 (2021).

70. Hofweber, M. & Dormann, D. Friend or foe-Post-translational modifications as regulators of phase separation and RNP granule dynamics. *J. Biol. Chem.* **294**, 7137–7150 (2019).

71. Snead, W. T. & Gladfelter, A. S. The control centers of biomolecular phase separation: how membrane surfaces, PTMs, and active processes regulate condensation. *Mol. Cell* **76**, 295–305 (2019).

72. Watanabe, K. et al. Cells recognize osmotic stress through liquid-liquid phase separation lubricated with poly(ADP-ribose). *Nat. Commun.* **12**, 1353 (2021).

73. Gao, C. et al. Hyperosmotic-stress-induced liquid-liquid phase separation of ALS-related proteins in the nucleus. *Cell Rep.* **40**, 111086 (2022).

74. Franzmann, T. M. & Alberti, S. Protein phase separation as a stress survival strategy. *Cold Spring Harb Perspect Biol* **11**, a034058 (2019).

75. Jalihal, A. P. et al. Multivalent proteins rapidly and reversibly phase-separate upon osmotic cell volume change. *Mol. Cell* **79**, 978–990 e975 (2020).

76. Konig, I. et al. Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat. Methods* **12**, 773–779 (2015).

77. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

78. Muller, B. K., Zaychikov, E., Brauchle, C. & Lamb, D. C. Pulsed interleaved excitation. *Biophys. J.* **89**, 3508–3522 (2005).

79. Cubuk, J. et al. The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nat. Commun.* **12**, 1936 (2021).

80. Schuler, B., Muller-Spath, S., Soranno, A. & Nettels, D. Application of confocal single-molecule FRET to intrinsically disordered proteins. *Methods Mol. Biol.* **896**, 21–45 (2012).

81. Holmstrom, E. D. et al. Accurate transfer efficiencies, distance distributions, and ensembles of unfolded and intrinsically disordered proteins from single-molecule FRET. *Methods Enzymol.* **611**, 287–325 (2018).

82. Zheng, W. et al. Inferring properties of disordered chains from FRET transfer efficiencies. *J. Chem. Phys.* **148**, 123329 (2018).

83. Nettels, D., Hoffmann, A. & Schuler, B. Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds. *J. Phys. Chem. B* **112**, 6137–6146 (2008).

84. Gopich, I. V., Nettels, D., Schuler, B. & Szabo, A. Protein dynamics from single-molecule fluorescence intensity correlation functions. *J. Chem. Phys.* **131**, 095102 (2009).

85. Eastman, P. et al. OpenMM 7: rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol.* **13**, e1005659 (2017).

86. Tesei, G., Schulze, T. K., Crehuet, R. & Lindorff-Larsen, K. Accurate model of liquid-liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. *Proc. Natl Acad. Sci. USA* **118**, e2111696118 (2021).

87. Tesei, G. & Lindorff-Larsen, K. Improved predictions of phase behaviour of intrinsically disordered proteins by tuning the interaction range [version 2; peer review: 2 approved]. *Open Res. Euro.* **2**, 94 (2023).

88. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput Chem.* **32**, 2319–2327 (2011).

89. Lindahl, E., Abraham, T., Hess, M. & van der Spoel, D. GROMACS 2021 Source code. *Zenodo*, https://doi.org/10.5281/zenodo.4457626 (2021).

90. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph* **14**, 27–38 (1996). 33-38.

91. Erdos, G., Pajkos, M. & Dosztanyi, Z. IUPred3: prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. *Nucleic Acids Res.* **49**, W297–W303 (2021).

92. O'Brien, K. T., Mooney, C., Lopez, C., Pollastri, G. & Shields, D. C. Prediction of polyproline II secondary structure propensity in proteins. *R. Soc. Open Sci.* **7**, 191239 (2020).

93. Tamiola, K., Acar, B. & Mulder, F. A. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.* **132**, 18000–18003 (2010).

94. 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).

95. Itzhak, D. N., Tyanova, S., Cox, J. & Borner, G. H. Global, quantitative and dynamic mapping of protein subcellular localization. *Elife* **5**, e16950 (2016).

96. Hein, M. Y. et al. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712–723 (2015).

97. Duncan, R. & Hershey, J. W. Identification and quantitation of levels of protein synthesis initiation factors in crude HeLa cell lysates by two-dimensional polyacrylamide gel electrophoresis. *J. Biol. Chem.* **258**, 7228–7235 (1983).

Acknowledgements

Author contributions

M.A. and C.D.M. conceived and designed the experiments. B.C.S., P.S., S.R., L.F. A.M. V.E.A, C.D.M and M.A. performed the experiments. D.M. conceived and designed the simulations. V.U. performed the simulations. B.C.S., P.S., V.U., L.F., D.M., C.D.M and M.A. analyzed the data. M.A. wrote the paper with the contributions from B.C.S., P.S., V.U., S.R., D.M. and C.D.M.

Competing interests

The authors declare no competing interests.

Additional information