Title

A chaperone and a heat-resistant obscure (Hero) protein mediates conformational extension and aggregation suppression of TDP-43

Authors and affiliations

Andy Y.W. Lam[1,2], Kotaro Tsuboyama[1,3,*], Hisashi Tadakuma[1,4,*], Yukihide Tomari[1,2,*]

1. Institute for Quantitative Biosciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan.
2. Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan
3. Institute of Industrial Science, the University of Tokyo, Meguro-ku, Tokyo, 153-8505, Japan
4. School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, People's Republic of China

To whom correspondence should be addressed. Tel: +81-3-5841-7839; Email: tomari@iqb.u-tokyo.ac.jp

Correspondence may also be addressed to Kotaro Tsuboyama. Tel: +81-3-5452-6333; Email: kotaro.tsuboyama@gmail.com or Hisashi Tadakuma. Tel: 86-21-20685017; Email: tadakumahisashi@shanghaitech.edu.cn

Abstract

Misfolding and aggregation of proteins are characteristic features in the progression of neurodegenerative diseases. While molecular chaperones are well-known suppressors of these aberrant events, we recently reported that a widespread family of heat-resistant obscure (Hero) proteins may also play a similar role. However, it is unclear how such electrostatically charged intrinsically disordered proteins protect client proteins from denaturation and aggregation. Here we utilize single-molecule FRET to monitor the conformations of an aggregation-prone protein, TAR DNA-binding protein 43 (TDP-43). While we observed high conformational heterogeneity of wild-type TDP-43, the ALS-associated mutation A315T promoted collapsed conformations. In contrast, a passive HSP40 family chaperone, DNAJA2, and a Hero protein, Hero11, both of which prevent TDP-43 aggregation in cells, stabilized extended states. Our results link single-molecule effects on the TDP-43 conformation to macro effects on bulk aggregation propensity, where a Hero protein, like a chaperone, maintains the conformational integrity of a client protein to prevent its aggregation.

Introduction

Most protein functions are defined by their native state, usually well-folded tertiary structures. However, intrinsically disordered proteins (IDPs) do not follow this principle. Particularly when isolated, IDPs exhibit a high degree of conformational flexibility, adopting a range of loosely defined structures due to the high hydrophilicity of their amino acid sequences and the absence of a hydrophobic 'core' necessary for a well-defined tertiary structure. This flexibility enables cooperative, multivalent interactions with various binding partners and contexts. Consequently, IDPs can play diverse roles, including gene regulation, signal transduction, trafficking, and assembly (1, 2).

Despite these important functional roles, several IDPs have become the focus of intense study for their aberrant behavior in neurodegenerative diseases. One representative of such a misbehaving IDP is transactive response DNA binding protein 43 (TDP-43), a highly abundant and essential RNA-binding protein (3, 4), that functions in various manners of RNA processing, including mRNA transcription, splicing, transport and translation, and the processing of microRNAs and long non-coding RNAs (3, 5). TDP-43 is predominantly localized to the nucleus, but can move into the cytoplasm upon cellular stress, including oxidative, osmotic, heat proteasome, and ER stresses, where it not only regulates but also forms stress granules (6, 7). During prolonged stress, TDP-43 even forms aberrant aggregates in the cytosol (8–11). Such aggregates have become known as a pathological hallmark of amyotrophic lateral sclerosis (ALS), appearing in ~97% of cases (12, 13) and in ~43% of frontotemporal lobar degeneration (FTLD) (3). In these cases, the intrinsically disordered ~141 amino acid C-terminal region of TDP-43, known as the low complexity domain (LCD), is the core to form typical steric-zipper β-sheet structures characteristic for the aggregates (14–17).

To counteract the disasters caused by protein aggregates, cells have developed mechanisms to suppress misfolding, for example molecular chaperones that mainly assist in the proper folding of proteins during or after translation (18). Chaperones usually work in an ATP-dependent manner, promoting the proper folding of unfolded client proteins or the disassembly of already aggregated proteins, for example HSP70/40/90 and the Hsp70/40/110 systems, respectively (18–20). Additionally, chaperones can also work in an ATP-independent manner, via binding to unfolded client proteins to avoid their aggregation. An example of this is HSP40, also known as DnaJ or J-domain protein, which holds client proteins until they are handed off to HSP70 (20–22). In both ATP-dependent and -independent cases, chaperones bind and stabilize the folding intermediates of their client proteins through the interfaces on the tertiary structure of the chaperones (23–25). Chaperones collaborate with each other and other processes such as the degradation pathways, forming an integral part of the protein quality control network (18, 20).

In addition to the structure-based conventional chaperones, it is increasingly evident that intrinsically disordered proteins or regions also play roles in the stabilization of client proteins (2, 26–31). For example, we previously reported that heat-resistant obscure (Hero) proteins, abundant in the boiled supernatant of human and fly cell lysates, stabilize activity of client proteins, and prevent them from aggregation (32, 33). However, hero proteins appear vastly different from structured chaperones in that they are highly hydrophilic, electrostatically charged, and intrinsically disordered. Remarkably, Hero protein activity seems to rely on their biophysical properties as highly charged and sufficiently long polymers, rather than specific sequence-encoded structure (32). However, it is still unclear how such disordered proteins can suppress aggregation, especially with these obvious structural differences and unique biophysical properties.

Here we report unexpectedly similar effects of a hero protein, Hero11, and a HSP40-family chaperone, DNAJA2, on the conformational dynamics and aggregation of TDP-43. Using genetic code expansion in HEK293T cells, we synthesized dye-labeled full-length TDP-43 for monitoring its LCD via total internal reflection fluorescence (TIRF) single-molecule fluorescence energy transfer (smFRET). While we observed high heterogeneity in the conformation of the WT LCD, the disease-associated aggregation-enhancing A315T mutation promoted collapsed conformations. In contrast, both DNAJA2 and Hero11 stabilized extended conformations, while a charge-less Hero11 mutant did not, mirroring their effects on aggregation. This suggests that Hero11, despite lacking clear structural features like conventional HSP40, can still stabilize extended conformations in a similar effect.

Materials and Methods

HEK293T Cell culture and lysate preparation.

HEK293T cells were maintained in Dulbecco's Modified Eagle Medium (DMEM, Sigma, QJ-D6429) supplemented with 10% FBS in a 5% CO_2 atmosphere at 37°C. Naïve HEK293T cell lysates were produced as previously described (32).

Plasmids for recombinant protein production in E. coli

Plasmids for recombinant expression of Hero proteins and control, pColdI_ Hero11, Hero11KR→G, and GST -FLAG-6×His were previously described (32). Recombinant plasmids for chaperones pColdI_6×His- DNAJA2 were previously described (34).

Plasmids for HEK293T cell-based assays and protein production

For transfection into HEK293T cells, pCAGEN_6×His- DNAJA2-HA was created from PCR of the ORF from pColdI_6×His- DNAJA2 and insertion into EcoRI-digested pCAGEN plasmids using NEBuilder HiFi DNA Assembly Master Mix (NEB). This plasmid was then used to produce pCAGEN_3xFLAG-DNAJA2 in a similar manner. Plasmids for Hero protein and controls, pCAGEN_3xFLAG- Hero11, Hero11KR→G, and GST were previously described (32). Plasmids for chaperones pCAGEN_3xFLAG-DNAJB8 and DNAJC8 were created as previously described (35). Plasmids for genetic code expansion, pCAGEN_U6tRNAUCmut×4_NESPylRS-Afmut and pCAGEN_U6tRNAUCmut×4_eRF1EDmut were previously described (36). Plasmid for expression of GFP-tagged TDP-43 in HEK293T cells, TDP43ΔNLS was previously described (32). The plasmid for incorporating unnatural amino acids in TDP-43, pCAGEN_3xFLAG-TEV-HALO-TDP43_R268Amb-F420Amb-2×HA was created first by PCR of the ORF from pCAGEN_eGFP- TDP43 and insertion into NotI-digested pCAGEN_FLAG-tev-HALO-Ago2. Then an ocher mutation of the final stop codon after 2×HA, and amber mutations around the TDP-43 LCD at R268 and F420 were introduced by PCR-based site-directed mutagenesis, as previously described (36). For testing the A315T mutation, pCAGEN_eGFP-TDP43_A315TΔNLS, and pCAGEN_FLAG-TEV-HALO-TDP43_R268Amb-A315T-F420Amb-2×HA were created from A315T mutation of pCAGEN_eGFP-TDP43_ΔNLS, and pCAGEN_FLAG-TEV-HALO-TDP43_R268Amb-F420Amb-2×HA, respectively, via PCR-based site-directed mutagenesis.

Protein Purification of His-tagged proteins from pColdI plasmids.

Recombinant Hero proteins and charge mutants and DNAJA2 were purified as previously described (32, 34). Briefly, pColdI plasmids for C-terminally FLAG-6×His-tagged proteins were transfected in Rosetta2 (DE3) and grown at 37°C to O.D. 0.4−0.6. Proteins were induced with 1 mM isopropyl-β-D-thiogalactoside (IPTG) overnight at 15°C and snap-frozen with liquid nitrogen. Cells were lysed by probe sonication in His A buffer (30 mM HEPES-KOH pH 7.4, 200 mM KOAc, 2 mM $Mg(OAc)_2$, 5% glycerol) supplemented with cOmplete™, EDTA-free Protease Inhibitor Cocktail (Roche), cleared by centrifugation, and purified over Ni-NTA resin (Roche) and eluted with His B buffer (HisA with 400mM imidazole). Proteins were buffer exchanged into 1×PBS using PD-10 columns (Cytiva), concentration determined by BCA assay (Pierce) before supplementation with 1 mM DTT, snap frozen in aliquots and stored at −80°C.

Protein purification via genetic code expansion

Genetic code expansion was performed as described previously (36–38). HEK293T cells were seeded and transfected after 24 h with pCAGEN plasmids in medium, EndoBCN-K (SiChem, SC-8014) supplemented, and harvested 48 h after transfection. Cells were lysed via BIORUPTOR II (BM Bio), snap-frozen with liquid nitrogen and stored at −80°C. Proteins were purified with FLAG-functionalized Dynabeads Protein G (Invitrogen), dye-labeled on beads with tetrazine functionalized- Cy3 and ATTO647N (Jena Bioscience, CLK-014-05 and CLK-012-02), functionalized with HaloTag Biotin ligand (Promega, G8281), eluted in naïve HEK293T cell lysate with TurboTEV Protease (Accelagen, T0102S), and aliquots snap-frozen in liquid nitrogen and stored at −80°C. Fluorescent dye-labeling was verified on SDS-PAGE via detection of fluorescence at indicated wavelengths with an Imager 600 (Amersham).

Cell-based aggregation assays

Cell-based aggregation assays were performed as previously described (32). HEK293T cells were co-transfected with 200 ng pCAGEN plasmid for TDP-43 and 1000 ng pCAGEN plasmid for the other protein using Lipofectamine 3000 Transfection Reagent (Invitrogen, L3000001) in 12-well plates. After 48 h, cells were imaged using FLUOVIEW FV3000 confocal laser scanning microscope (Evident) and then lysed with BIORUPTOR II (BM Bio) in 1xPBS supplemented 1x cOmplete™ Protease Inhibitor Cocktail (Roche) in Protein LoBind Tubes (Eppendorf). Protein levels were normalized via Bradford assay (Bio-Rad) and aggregates were captured on a 0.2 μm Whatman OE66 cellulose acetate membrane filter (GE Healthcare, 10404180) using 1× PBS with 1% SDS to solubilize. TDP-43 aggregation levels were detected by GFP fluorescence signal and immunoblotting for GFP using α-GFP (B-2) primary antibody (Santa Cruz Biotechnology, sc-9996) with α-mouse IgG secondary antibody (MBL).

Microscopy of GFP-TDP43ΔNLS in cells

HEK293T cells were seeded onto 35 mm glass bottom dishes (MatTek, P35G-1.5-14-C), pre-treated with Poly-L-lysine (Sigma-Aldrich, P4707), and transiently transfected with 1000 ng pCAGEN_GFP-TDP43ΔNLS and 1500 ng pCAGEN-3xFLAG- DNAJA2, Hero11, Hero11KR→G, or GST using Lipofectamine™ 3000 Transfection Reagent (Invitrogen, L3000001), and after 48h, Hoechst 33342 stained (Invitrogen, H3570) for imaging on a FLUOVIEW FV3000 confocal laser scanning microscope (Evident) using 100x oil-immersion lens, controlled via FV31S-SW Viewer software (Evident).

For FRAP experiments, cells were transfected using 1000 ng pCAGEN_GFP-TDP43ΔNLS and 1500 ng pCAGEN-3xFLAG-GST, without Hoechst staining, and using the LSM stimulation mode. Videos were taken zoomed into a single cell, at 1 second per frame. After five pre-bleaching frames, 100% power laser stimulation was applied to a small circular area within the puncta for 500 ms, and an image taken as the bleach frame, followed by 150 additional recovery frames.

Single-molecule FRET microscopy

Single-molecule assays were conducted as previously described (36, 39), with modifications to minimize aggregation. Aliquots of dye-labelled HALO-TDP43 were pre-diluted in 1×Lysis buffer (30 mM HEPES-KOH pH 7.4, 100 mM KOAc, 2 mM $Mg(OAc)_2$), and then diluted in naïve HEK293T cell lysate with 1×ATP regeneration system (1 mM ATP, 25 mM creatine monophosphate (Sigma), 0.03 U/μL creatine kinase (Calbiochem), 0.1 U/μL RNasin Plus RNase Inhibitor (Promega), in Lysis buffer) and incubated at 37°C for 30 min, to enable aggregation-clearing chaperone activities. Note the use of high-quality HEK293T cell lysate via hypotonic lysis method was crucial, as lysate prepared via sonication failed to sufficiently remove aggregates. This reaction was then centrifuged at 18,000 × G for 15 min at 25°C and supernatant collected to remove aggregated proteins in the pellet. Observation chambers were prepared as previously described (36, 39), infused with neutravidin for 2 min, and washed with Lysis buffer. Supernatant was flowed into chamber and incubated for 2 min, and then washed twice with Wash buffer (1×Lysis buffer containing 1% Triton X-100 and 800 mM NaCl), four times with Wash buffer supplemented with 10 mM ATP, and then four times with Lysis buffer. These stringent washes are necessary to remove nonspecific interactions and oligomerization of TDP-43 with the tethered TDP-43 molecules. The final Observation buffer containing 0.3 g/L of BSA or recombinant protein in 1× PBS with 1 mM DTT, O_2 scavenger system (7.1 mM PCA (3,4-Dihydroxybenzoic acid, 37580, Sigma) with 29 nM PCD (Protocatechuate 3,4-dioxygenase, P8279, Sigma) (Aitken *et al*, 2008)), and 2 mM Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid, 238813, Sigma) was flowed into chamber and then mounted onto microscope stage for smFRET observations. Three five-minute videos were taken of each slide, using alternating laser illumination of 1s each. We emphasize that these procedures are uniquely practical for TIRF-smFRET experiments, which by tethering dye-labeled molecules, allows for buffer exchange and minimization of aggregation for an extremely small sample.

Data analysis and visualization software

Image analyses of cells were performed using the Fiji distribution of ImageJ (40). Data analysis was performed in Python 3.7.10, using pandas, NumPy, SciPy, Matplotlib, seaborn in addition to the standard library, via Jupyter notebooks implemented with Spyder-notebook in Spyder IDE version 5.

Analysis of aggregation assays

Images from bulk aggregation filter trap assays were taken using Amersham ImageQuant 800 Imager (Amersham), in TIFF image format. For quantification, images were analyzed using ImageJ Fiji , using a simple macro script, in a manner similar to dot blots, first with background subtraction via the "Subtract Background" tool, with a rolling ball radius of 50 pixels. Circular areas of 125 pixels diameter were then selected around the highest concentration wells to obtain integrated densities for each sample and saved to CSV file format. These were imported into python for analysis, where integrated densities were divided by the GST co-transfection well in each blot, to quantify aggregation relative to GST, and visualized.

Microscope image analysis

HEK293T cell images in figures were prepared via the QuickFigs plugin for Fiji (41). For puncta counting and measurement analyses, images were analyzed through an ImageJ macro that, using base macro functions, thresholded, cleaned up, and identified areas of interest. These created masks for cell areas, nuclei areas, and puncta areas, and image calculator operations performed to identify overlapping areas. These defined cytoplasmic puncta were used to count and measure puncta in the original image and exported to CSV for analysis and visualization in python.

FRAP analysis

For FRAP analysis, a macro used to allow automatic selection of regions of interests: the bleach spot, the puncta area, the cell area and a background area for FRAP analysis, which were interpolated over the video to account for movement of the puncta and cell, then measured and exported to csv for analysis in python. Measurements for each recovery curve were background corrected and double normalized to average max intensity pre-bleach and min intensity post-bleach (42). A single exponential, exponential

equation (Equation 1) was fitted to each recovery curve using non-linear least squares via SciPy, for determination of the mobile fraction, $(1 - I_f)$, and visualized.

$I(t) = I_f\ (1 - e^{-kt})$ (Equation 1)

Where the intensity at time t, $I(t)$, is described by the rate parameter, k, and I_f is the plateau recovery intensity.

Analysis of smFRET assays

The smFRET videos were imported to iSMS implemented in MATLAB for image channel alignment, peak localization, FRET-pair identification, photobleaching and blinking identification, drift correction, and correction factor determination and application as previously described (36, 43, 44). Note the background correction, applied uniformly by experiment, sometimes results in negative values for FRET values measured close to 0. FRET efficiency traces were selected based on single-step photobleaching, stoichiometry ~0.4 via ALEX and anticorrelation. FRET distribution analysis data were exported as TXT files from iSMS. Visualization was conducted through python to load TXT files, and density histograms were plotted using the combined data for each condition with 25 bins. Average density plots were created by estimating the probability density functions for each independent experiment via gaussian kernel density estimation with 0.05 bandwidth, and values of each curve evaluated for determination of means and standard errors along the entire range. Mean density and standard error of the mean for each condition were plotted over the histograms.

Analysis of smFRET dynamics

For dynamics analysis a hidden Markov model (HMM) with variational Bayesian expectation maximization (VBEM) implemented in iSMS was used to extract ideal traces, as previously described (36, 44), as TXT and imported into python for analysis. Traces were denoised for counting of transition events and dwell times, and visualized. Single exponential distributions were fitted to the dwell time distributions using Bayesian modeling implemented in PyMC 5.9. A normally distributed (μ=0, σ=10) prior was used for the log of the lambda parameter (λ, equivalent to rate parameter, k) for the exponential model, and combined with the observed dwell time distribution in the likelihood function to run Markov chain Monte Carlo (MCMC) stochastic simulations with the No U-Turn Sampler (NUTS) algorithm, taking 100 samples after 100 tuning samples from one chain. We reported median and 95% credible

interval estimates of the posterior distribution of the fitted parameters. These were visualized as cumulative distribution functions (Equation 2).

$f(t) = 1 - e^{-kt}$ (Equation 2)

Where the cumulative distribution function, $f(t)$ is described by the rate parameter, k.

Wormlike chain model

To approximate end-to-end distance, we assumed that the LCD polymer behaves according to the Worm-Like Chain (WLC) model (45–47).

$P(r) = \frac{4\pi C r^2}{L^2[1-(r/L)^2]^{9/2}} e^{-\frac{3L}{4l_p[1-(r/L)^2]}}$ (Equation 3)

Where r is distance between polymer ends, L is the polymer contour length, $P(r)$ the probability of polymer with end-to-end distance of r, l_p the persistence length of the polymer (describing bending), and C the normalization factor. For simplicity, we considered a one-dimensional distribution, of contour length 47.94 nm (141 amino acids × 0.34 nm per amino acid). Persistence length of 0.8 nm was chosen to describe random coils such as experimentally observed for polyglycine (or glycine-serine repeats) (45).

Interdye distance estimations from FRET efficiencies

To approximate theoretical end-to-end distances from FRET efficiencies, we used community provided values of the Förster Radius, R_0, for FRET between Cy3 and ATTO 647N (48), following Förster's equation for relationship between FRET efficiency and distance (Equation 4).

$E = \frac{1}{1 + (R/R_0)^6}$ (Equation 4)

Where the FRET efficiency, E, is related to the distance between fluorophores, R.

Results

DNAJA2 and Hero11 suppress TDP-43 aggregation in HEK293T cells

In order to understand how Hero proteins act upon aggregation, we first sought to have a well-established chaperone protein for reference. Since we do not expect Hero proteins to use ATP for this activity, we

reasoned that the ATP-independent chaperoning ability of HSP40 proteins would be a suitable reference for studying Hero proteins. Thus, we first performed cell-based aggregation assays to identify an effective HSP40 chaperone for the aggregation of TDP-43 (32). As in the previous study, we used GFP-tagged TDP-43 lacking its nuclear localization signal (GFP-TDP43ΔNLS WT) to increase cytoplasmic mislocalization and aggregation (12, 13). GFP-TDP43ΔNLS was co-transfected with Hero proteins, HSP40 proteins, or the GST (Glutathione S-transferase) control into HEK293T cells, and aggregation was assessed by the filter trap method (Figure 1A). In agreement with our previous results, a 11 kDa highly positively charged Hero protein, Hero11 (also known as L10K protein or C19orf53) showed a strong suppression activity for TDP-43 aggregation, while its KR→G charge mutant, in which all 21 lysine and 6 arginine residues have been mutated to neutral glycine, fails to suppress aggregation (Figure 1B−D), suggesting the importance of net charge for this anti-aggregation activity shown previously (32). Of the tested HSP40s, DNAJA2 most effectively suppressed aggregation of co-transfected GFP-TDP43ΔNLS relative to the GST control in our experimental conditions (Figure 1D). Given that DNAJA2 was recently demonstrated to selectively bind to and reduce the phase separation of the TDP-43 LCD (49), our results support DNAJA2 as a representative suppressor of TDP-43 aggregation.

Many disease-associated mutations of TDP-43 have been found in the LCD, for example A315T, which stabilizes amyloid fibrils (16, 17) and promotes its aggregation and toxicity (50) Thus, we next introduced the A315T mutant into GFP-TDP43ΔNLS (GFP-TDP43ΔNLS A315T) and co-transfected it with Hero11 or DNAJA2. Similar to the WT GFP-TDP43ΔNLS, DNAJA2 and Hero11, but not Hero11KR→G, suppressed aggregation of A315T (Figure 1E), suggesting both remain effective suppressors of the pathogenic A315T mutant.

We next performed detailed microscopy analysis to investigate how Hero and DNAJA2 affect the morphology and dynamics of GFP-TDP-43ΔNLS aggregates in HEK293T cells. While large, irregularly-shaped puncta of GFP-TDP43ΔNLS appeared in the cytoplasm of cells co-transfected with GST or Hero11KR→G, these were absent in the DNAJA2 and Hero11 co-transfections (Figure 2A). Our image analysis pipeline also revealed that cells in GST and Hero11KR→G co-transfections contained ~10 μm^2 total puncta signal area on average, but DNAJA2 and Hero11 co-transfections reduced this to ~0.3 and ~0.8 μm^2 respectively (Figure 2B).

To understand how DNAJA2 and Hero11 reduce these cytoplasmic puncta, we also quantified the size of individual puncta observed in the images (Figure 2C). The sizes were unevenly distributed in all co-transfections, consisting of ~50–75% small areas ≤1μm^2, for all of the four conditions we tested here. However, in the GST and Hero11KR→G co-transfections, these were more numerous and also tended to be larger than those in DNAJA2 and Hero11. To understand the biophysical properties of these large

puncta, we performed fluorescence recovery after photobleaching (FRAP) of GFP-TDP-43ΔNLS bodies in the GST control. In our experimental conditions, we observed a discontinuous relationship between the size of puncta and internal mobility, where small puncta with areas <20 μm² tended to have recoveries >0.8 while larger bodies with areas ≥20 μm² had <0.1 recovery (Figure 2D and 2E). This suggests that while smaller puncta tend to be highly fluid, bodies larger than 20 μm² are practically immobile. Notably, while ~3% of puncta in the GST (35 out of 1156) and Hero11KR→G (38 out of 1107) co-transfections were larger than 20 μm², we observed no puncta above this size threshold out of 79 and 281 total puncta in the DNAJA2 and Hero11 transfections, respectively (Figure 2C). Altogether these suggest that both Hero11 and DNAJA2 efficiently suppress the transition of small liquid-like TDP-43 bodies into larger solid cytoplasmic aggregates, which may be toxic and cause disease (11, 51).

TDP-43 LCD takes heterogenous and compact conformation when alone

Given the importance of the intrinsically disordered LCD for TDP-43 aggregation (14–16), we next sought to observe the conformational behavior of this region at the single-molecule level. Total internal reflection fluorescence (TIRF)-based smFRET is a powerful tool for isolating and collecting long-range conformational data of individual IDPs that are difficult to obtain by other methods (52–54), resolving heterogeneity between molecules. Using genetic code expansion in HEK293T cells, we synthesized full-length TDP-43 with the unnatural amino acid, endo-BCN-L-lysine (endoBCN-K) incorporated around the LCD, via mutation of R268 in the linker six amino acids before, and in a linker six amino acids after the LCD (Supplementary Figure 1). We then performed site-specific dye-labeling at these two positions with tetrazine-coupled fluorescent dyes, Cy3 (donor) and ATTO647N (acceptor) (36–38). This dual-labeled TDP-43 was immobilized onto PEG-biotinylated glass using an N-terminally encoded HaloTag, isolating individual molecules for illumination via TIRF to observe the conformational behavior of the TDP-43 LCD (Figure 3A, B). With an alternating laser excitation (ALEX) scheme, where two lasers were used to alternately excite the donor and acceptor dyes, we selected molecules labeled with one Cy3 and one ATTO647 dye, as assessed by dye stoichiometry and one-step photobleaching (Figure 3C) (36, 55). In our experimental scheme, higher FRET efficiency occurs when fluorophores are in close proximity, corresponding to relative collapse of the LCD, whereas lower FRET efficiency indicates relative extension. Under incubation with the BSA control, we observed a broad range of FRET signals, with a slight peak at ~0.8, suggesting high heterogeneity in the distribution of LCD conformations (Figure 3D), in agreement with its disordered nature and tendency to aggregate (56, 57). We also introduced the A315T mutation into the genetic code expanded TDP-43 for single-molecule analysis. Unlike WT LCD (Figure 3D), we observed a strong high FRET peak (~0.9) for the A315T mutant LCD when incubated in

the control BSA-containing buffer (Figure 3H), indicating that A315T mutation induces more collapsed conformations than WT LCD.

Hero11 and DNAJA2 promote extension of the TDP-43 LCD

Since previous TIRF-smFRET studies of bacterial HSP40s show that they stabilize conformationally expanded states of natively folded client proteins such as luciferase and rhodanese (58, 59), we investigated if eukaryotic DNAJA2 was capable of eliciting this effect on the intrinsically disordered client TDP-43 LCD. In accordance with these studies, incubation of the LCD with DNAJA2 resulted in a very low FRET peak at ~0.0 (Figure 3F), which indicates extension of the LCD conformation. Strikingly, the addition of Hero11 also caused a similar very low FRET peak (Figure 3G). In contrast, Hero11KR→G mutant did not cause apparent changes in the FRET distribution compared with the BSA control (Figure 3E), consistent with its inability to suppress TDP-43 aggregation in cells (Figure 1D).

Similarly, consistent with their effectiveness in suppressing aggregation of the A315T mutant in cells (Figure 1E), Hero11 and DNAJA2 both promoted very low FRET peaks (~0.0) in the FRET distribution of the A315T LCD (Figure 3J, K). Unlike the WT LCD where its influence on conformational distribution was indistinguishable from BSA control (Figure 3E), incubation with Hero11KR→G resulted in loss of the very high FRET peak (~0.9), resulting in a flat conformational distribution of the A315T LCD (Figure 3I). This suggests that the KR→G mutant can reduce the fraction of LCD molecules trapped in very high FRET states, but unlike DNAJA2 and Hero11 WT, did not cause a shift to a low FRET peak. Altogether this suggests that DNAJA2 promotes extended conformations of the aggregation-prone intrinsically disordered WT and A315T TDP-43 LCD, and Hero11 despite lacking specific tertiary structures such as those of DNAJA2, is capable of eliciting a similar effect, in a manner dependent on its net charge.

Hero11 and DNAJA2 stabilize conformationally dynamic extended states of the LCD

We next investigated detailed conformational dynamics of TDP-43 LCD in the presence of the aggregation suppressors. To this end, we idealized the dynamic FRET trajectories into discrete FRET states via a hidden Markov model (HMM), allowing for quantitative analysis of conformational dynamics (Figure 3C) (36, 44). To understand how DNAJA2 and Hero11 change transition dynamics, FRET states were classified into 4 bins based on FRET efficiency: very low (<0.2); low (0.2–0.5); high (0.5–0.8); and very high (>0.8), then, we counted transition events between FRET states for all possible combinations,

summarizing the frequency of transition events occurring between these FRET state bins (Figure 4A). While there were some weak transition patterns in BSA (between low and high bins and vice versa) and Hero11KR→G (from very low to high), there was no clear overall pattern in the effects of these proteins on transition events. In contrast, we observed most transition events (~20%) occur within the very low FRET bin for the DNAJA2 and Hero11 incubations, with little pattern elsewhere. This suggests that both DNAJA2 and Hero11 can keep the dynamic transitions of the LCD within very low FRET state (i.e. the extended state) (Figure 3F, G).

For the A315T mutant (Figure 4B), the transition density histograms in BSA control displays most transition events occurring between the high and very high FRET bins. Incubation with Hero11KR→G results in evenly distributed FRET state transitions from very low to very high bins, consistent with its flat conformational distribution (Figure 3I). By contrast, Hero11 or DNAJA2 show that transition events occur more frequently within <0.2 FRET bins, suggestive of dynamic behaviors of molecules in the very low FRET peak promoted by these proteins (Figure 3F, G). Taken together, the effects of BSA, Hero11, Hero11KR→G, and DNAJA2 on the transition dynamics of the A315T mutant LCD are generally consistent with their effects on the WT LCD.

We next measured the dwell times of the FRET states (duration spent in a certain FRET state before transitioning towards another state) for the four bins (very low to very high as in Fig. 4A and B), summarizing as empirical cumulative distribution functions (Figure 4C, D). Because it is expected that transitions between the four FRET bins are governed by single-step processes, we used a Bayesian approach to fit single exponential distribution models to each of the dwell time distributions in order to quantitatively sample/estimate the kinetic rate constant (reported as dwell time half-life, $t_{½}$, see Supplementary Figure 2 for fitting accuracy).

For the WT LCD (Figure 4C), Hero11 and DNAJA2 clearly extended the dwell times of the very low FRET states of the LCD compared to BSA and Hero11KR→G (median $t_{1/2}$ ~33–36 s vs ~24–25 s, for Hero11 and DNAJA vs BSA and Hero11KR→G, respectively). Similarly, Hero11 and DNAJA2 also extended the dwell times of the very low FRET states of A315T compared to BSA and Hero11KR→G (Figure 4D; median $t_{1/2}$ ~33 s vs ~23–25 s, respectively). Additionally, this extension of dwell time also occurs for the low FRET states of A315T (median $t_{1/2}$ ~34–38 s vs ~19–26 s, for Hero11 and DNAJA2 vs BSA and Hero11KR→G, respectively). This suggests that Hero11 and DNAJA2 stabilize extended conformations of both the WT and A315T LCD.

Taken together, our detailed dynamics analyses indicate that Hero11 and DNAJA2 both stabilized dwell times of the lowest FRET states (Figure 4C and D) while biasing for transition events between these lowest FRET states (Figure 4A and B) of WT and A315T LCDs. For DNAJA2, this effect is consistent with stabilization of unfolded client proteins by HSP40s (24, 58, 59). We also found that Hero11 achieves similar effects despite lacking the structural features necessary for HSP40 interaction with TDP-43 LCD (24, 49). Thus, despite differences in structural features, these results indicate that Hero11 and DNAJA2 have surprisingly similar effects on the dynamics of the TDP-43 LCD.

Discussion

Overall, our results orchestrally support that conformational modulation of the LCD is a key to TDP-43 pathological aggregation process. Although it is currently almost impossible to directly observe conformation within aggregates via smFRET, we envision that these compact conformations, which remain liable to opening, may represent a precursor to stably aggregated states. In support of this, we estimate the very high FRET ~0.9 peak corresponds to ~3.3 nm mean end-to-end distance for our dye pair (labeled at residues 268 and 420 six amino acids before and after the LCD), which is roughly compatible with distances in collapsed LCD conformations seen in amyloid fibril structures of the in vitro-formed full LCD (7KWZ [PDB code]: F276–M414, 1.716 nm) (14). This suggests the A315T mutation induces more compact conformations in addition to its role in stabilizing amyloid fibril structure through intra- and inter-strand interactions (16, 17).

In contrast, we observed that DNAJA2 stabilized extended conformations of the TDP-43 LCD and suppressed its aggregation. This is in line with the reports that HSP40 stabilizes non-native unfolded conformations of structured client proteins leading to the native state folding (24, 58–60), and reports that DNAJA2 binds to the LCD (49). We also found that Hero11 stabilized the extended conformations of LCD of TDP-43 similarly to DNAJA2. In support of this, small but increasing numbers of Hero11 molecules leads to increased radius of gyration for LCD molecules in molecular dynamics simulations (61). The very low FRET ~0 peak observed in the presence of Hero11 and DNAJA2 should contain conformations from 6.80 nm (0.1 FRET) to wider than 10.2 nm (0.01 FRET) mean end-to-end distance (~0.01 FRET), which is comparable or more extended than the ~7.34 nm mean distance expected for a “native” state of the LCD, estimated by a wormlike chain model. Together, these support conformational modulation of the LCD as determinant of TDP-43 aggregation, where DNAJA2- or Hero11-stabilized extended conformations are excluded from progressing further towards aggregation, as expected for HSP40 chaperoning (21, 22, 62).

Albeit the similar effect on TDP-43 LCD conformation, a key difference between DNAJA2 and Hero11 is that DNAJA2 makes extensive use of its structured features, including multivalent binding on multiple sites across the surfaces of its C-terminal domains across the dimer, in its chaperoning of client proteins (24). By contrast, Hero11 is unlikely to have a solid tertiary structure or bind to TDP-43 LCD in a defined conformation. Nonetheless, it is possible that Hero11 interacts with the TDP-43 LCD via partial alpha-helical secondary structure on both IDPs, as demonstrated in molecular dynamics simulations (61). Notably, such interactions between Hero11 and the LCD made up a large fraction of intermolecular contacts in these simulations (61). This transient alpha-helical structure in the LCD is also selectively bound by DNAJA2 (49), which may serve as an anchoring interaction for stabilizing LCD conformation. Alternatively, but not mutually exclusively, this stabilization effect may occur based on composition of the biophysical properties of the Hero11 chain regardless of the alpha-helical structure, because exclusion of this alpha-helical structure did not substantially affect overall results of simulations (61). As demonstrated in designed random polymers mimicking IDPs (63), conformational flexibility and composition of Hero11 may maximize stabilizing interactions with client proteins to prevent aggregation. Notably, the highly positively charged property of Hero11 is important for its interaction with the LCD, as the charge composition mutant Hero11KR→G, failed to elicit stabilization of the extended conformations nor suppress aggregation. Thus, we speculate that the combination of transient interactions and composition properties of Hero11 enable it to influence client protein conformation.

Moreover, these biophysical properties of Hero11 enable it to suppress phase separation of the LCD (and presumably later aggregation) in molecular dynamics simulations (61). This curiously mirrors how DNAJA2 suppresses LCD phase separation, where cationic residues of DNAJA2 are thought to supply similar repulsive effects (49). However, one key difference is that DNAJA2 is a co-chaperone of Hsp70 (20, 49), allowing for recruitment of TDP-43 towards the greater chaperoning network, whereas such activity is still unknown for Hero11. Nonetheless, we expect that hero proteins may play a supporting role in stabilizing unfolded client proteins, beneficial in that there are less immediately harmful misfolded proteins burdening the chaperoning system. Future work is required to decipher the downstream mechanistic particularities of hero proteins, and examine why they even exist alongside chaperones.

Funding

This work was supported in part by JSPS KAKENHI Grant-in-Aid for Transformative Research Areas (A) 21H05278 (to Y. T.), Grant-in-Aid for Scientific Research (S) 18H05271 (to Y. T.) and Grant-in-Aid for JSPS Fellows 22J11678 (to A. Y. W. L.).

Conflict of Interest Statement

Y. T. and K. T. have patent applications related to this work. H. T. and A. Y. W. L. declare no competing interests.

Acknowledgments

We thank all members of the Tomari laboratory for critical comments on the manuscript.

References

1. Dunker,A.K., Silman,I., Uversky,V.N. and Sussman,J.L. (2008) Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology*, **18**, 756–764.
2. van der Lee,R., Buljan,M., Lang,B., Weatheritt,R.J., Daughdrill,G.W., Dunker,A.K., Fuxreiter,M., Gough,J., Gsponer,J., Jones,D.T., *et al.* (2014) Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.*, **114**, 6589–6631.
3. Ling,S.-C., Polymenidou,M. and Cleveland,D.W. (2013) Converging Mechanisms in ALS and FTD: Disrupted RNA and Protein Homeostasis. *Neuron*, **79**, 416–438.
4. Ling,S.-C., Albuquerque,C.P., Han,J.S., Lagier-Tourenne,C., Tokunaga,S., Zhou,H. and Cleveland,D.W. (2010) ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *PNAS*, **107**, 13318–13323.
5. Prasad,A., Bharathi,V., Sivalingam,V., Girdhar,A. and Patel,B.K. (2019) Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis. *Front. Mol. Neurosci.*, **12**.
6. Aulas,A., Stabile,S. and Vande Velde,C. (2012) Endogenous TDP-43, but not FUS, contributes to stress granule assembly via G3BP. *Molecular Neurodegeneration*, **7**, 1.
7. Dewey,C.M., Cenik,B., Sephton,C.F., Johnson,B.A., Herz,J. and Yu,G. (2012) TDP-43 aggregation in neurodegeneration: Are stress granules the key? *Brain Research*, **1462**, 16–25.
8. Mann,J.R., Gleixner,A.M., Mauna,J.C., Gomes,E., DeChellis-Marks,M.R., Needham,P.G., Copley,K.E., Hurtle,B., Portz,B., Pyles,N.J., *et al.* (2019) RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. *Neuron*, **102**, 321-338.e8.
9. McGurk,L., Gomes,E., Guo,L., Mojsilovic-Petrovic,J., Tran,V., Kalb,R.G., Shorter,J. and Bonini,N.M. (2018) Poly(ADP-Ribose) Prevents Pathological Phase Separation of TDP-43 by Promoting Liquid Demixing and Stress Granule Localization. *Molecular Cell*, **71**, 703-717.e9.
10. Streit,L., Kuhn,T., Vomhof,T., Bopp,V., Ludolph,A.C., Weishaupt,J.H., Gebhardt,J.C.M., Michaelis,J. and Danzer,K.M. (2022) Stress induced TDP-43 mobility loss independent of stress granules. *Nat Commun*, **13**, 5480.
11. Gasset-Rosa,F., Lu,S., Yu,H., Chen,C., Melamed,Z., Guo,L., Shorter,J., Da Cruz,S. and Cleveland,D.W. (2019) Cytoplasmic TDP-43 De-mixing Independent of Stress Granules Drives Inhibition of Nuclear Import, Loss of Nuclear TDP-43, and Cell Death. *Neuron*, **102**, 339-357.e7.
12. Neumann,M., Sampathu,D.M., Kwong,L.K., Truax,A.C., Micsenyi,M.C., Chou,T.T., Bruce,J., Schuck,T., Grossman,M., Clark,C.M., *et al.* (2006) Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science*, **314**, 130–133.
13. Arai,T., Hasegawa,M., Akiyama,H., Ikeda,K., Nonaka,T., Mori,H., Mann,D., Tsuchiya,K.,

Yoshida,M., Hashizume,Y., *et al.* (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochemical and Biophysical Research Communications*, **351**, 602–611.
14. Li,Q., Babinchak,W.M. and Surewicz,W.K. (2021) Cryo-EM structure of amyloid fibrils formed by the entire low complexity domain of TDP-43. *Nat Commun*, **12**, 1620.
15. Arseni,D., Hasegawa,M., Murzin,A.G., Kametani,F., Arai,M., Yoshida,M., Ryskeldi-Falcon,B., Hasegawa,M., Murzin,A.G., Kametani,F., *et al.* (2021) Structure of pathological TDP-43 filaments from ALS with FTLD. *Nature*, **10**.1038/s41586-021-04199-3.
16. Cao,Q., Boyer,D.R., Sawaya,M.R., Ge,P. and Eisenberg,D.S. (2019) Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat Struct Mol Biol*, **26**, 619–627.
17. Guenther,E.L., Cao,Q., Trinh,H., Lu,J., Sawaya,M.R., Cascio,D., Boyer,D.R., Rodriguez,J.A., Hughes,M.P. and Eisenberg,D.S. (2018) Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. *Nat Struct Mol Biol*, **25**, 463–471.
18. Balchin,D., Hayer-Hartl,M. and Hartl,F.U. (2016) In vivo aspects of protein folding and quality control. *Science*, **353**, aac4354.
19. Nillegoda,N.B., Kirstein,J., Szlachcic,A., Berynskyy,M., Stank,A., Stengel,F., Arnsburg,K., Gao,X., Scior,A., Aebersold,R., *et al.* (2015) Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. *Nature*, **524**, 247–251.
20. Rosenzweig,R., Nillegoda,N.B., Mayer,M.P., Bukau,B., Nillegoda,N.B., Mayer,M.P. and Bukau,B. (2019) The Hsp70 chaperone network. *Nature Reviews Molecular Cell Biology*, **20**, 665–680.
21. Mitra,R., Wu,K., Lee,C. and Bardwell,J.C.A. (2022) ATP-Independent Chaperones. *Annual Review of Biophysics*, **51**, 409–429.
22. Graff,A.M. de, Mosedale,D.E., Sharp,T., Dill,K.A. and Grainger,D.J. (2020) Proteostasis is adaptive: Balancing chaperone holdases against foldases. *PLOS Computational Biology*, **16**, e1008460.
23. Karagöz,G.E., Duarte,A.M.S., Akoury,E., Ippel,H., Biernat,J., Morán Luengo,T., Radli,M., Didenko,T., Nordhues,B.A., Veprintsev,D.B., *et al.* (2014) Hsp90-Tau Complex Reveals Molecular Basis for Specificity in Chaperone Action. *Cell*, **156**, 963–974.
24. Jiang,Y., Rossi,P. and Kalodimos,C.G. (2019) Structural basis for client recognition and activity of Hsp40 chaperones. *Science*, **365**, 1313–1319.
25. Wang,R.Y.-R., Noddings,C.M., Kirschke,E., Myasnikov,A.G., Johnson,J.L., Agard,D.A., Noddings,C.M., Kirschke,E., Myasnikov,A.G., Johnson,J.L., *et al.* (2022) Structure of Hsp90–Hsp70–Hop–GR reveals the Hsp90 client-loading mechanism. *Nature*, **601**, 460–464.
26. Yakubu,U.M., Morano,K.A. and Morano,K.A. (2021) Suppression of aggregate and amyloid formation by a novel intrinsically disordered region in metazoan Hsp110 chaperones. *Journal of Biological Chemistry*, **296**. 100567
27. Chakrabortee,S., Meersman,F., Kaminski Schierle,G.S., Bertoncini,C.W., McGee,B., Kaminski,C.F. and Tunnacliffe,A. (2010) Catalytic and chaperone-like functions in an intrinsically disordered protein associated with desiccation tolerance. *Proceedings of the National Academy of Sciences*, **107**, 16084–16089.
28. Ren,C., Zheng,Y., Liu,C., Mencius,J., Wu,Z. and Quan,S. (2022) Molecular Characterization of an Intrinsically Disordered Chaperone Reveals Net-Charge Regulation in Chaperone Action. *Journal of Molecular Biology*, **434**, 167405.
29. Huang,L., Agrawal,T., Zhu,G., Yu,S., Tao,L., Lin,J., Marmorstein,R., Shorter,J., Yang,X., Agrawal,T., *et al.* (2021) DAXX represents a new type of protein-folding enabler. *Nature*, **10**.1038/s41586-021-03824-5.
30. Ikeda,K., Suzuki,S., Shigemitsu,Y., Tenno,T., Goda,N., Oshima,A. and Hiroaki,H. (2020) Presence of intrinsically disordered proteins can inhibit the nucleation phase of amyloid fibril formation of Aβ(1–42) in amino acid sequence independent manner. *Scientific Reports*, **10**, 12334.
31. Kovacs,D., Kalmar,E., Torok,Z. and Tompa,P. (2008) Chaperone Activity of ERD10 and ERD14, Two Disordered Stress-Related Plant Proteins. *Plant Physiology*, **147**, 381–390.
32. Tsuboyama,K., Osaki,T., Matsuura-Suzuki,E., Kozuka-Hata,H., Okada,Y., Oyama,M., Ikeuchi,Y.,

Iwasaki,S. and Tomari,Y. (2020) A widespread family of heat-resistant obscure (Hero) proteins protect against protein instability and aggregation. *PLOS Biology*, **18**, e3000632.
33. Morimoto,E., Tsuboyama,K. and Tomari,Y. (2022) Fusion with heat-resistant obscure (Hero) proteins have the potential to improve the molecular property of recombinant proteins. *PLOS ONE*, **17**, e0270097.
34. Naruse,K., Matsuura-Suzuki,E., Watanabe,M., Iwasaki,S. and Tomari,Y. (2018) In vitro reconstitution of chaperone-mediated human RISC assembly. *RNA*, **24**, 6–11.
35. Iwasaki,S., Sasaki,H.M., Sakaguchi,Y., Suzuki,T., Tadakuma,H. and Tomari,Y. (2015) Defining fundamental steps in the assembly of the Drosophila RNAi enzyme complex. *Nature*, **521**, 533–536.
36. Tsuboyama,K., Tadakuma,H. and Tomari,Y. (2018) Conformational Activation of Argonaute by Distinct yet Coordinated Actions of the Hsp70 and Hsp90 Chaperone Systems. *Molecular Cell*, **70**, 722-729.e4.
37. Nikić,I., Kang,J.H., Girona,G.E., Aramburu,I.V. and Lemke,E.A. (2015) Labeling proteins on live mammalian cells using click chemistry. *Nature Protocols*, **10**, 780–791.
38. Nikić,I., Estrada Girona,G., Kang,J.H., Paci,G., Mikhaleva,S., Koehler,C., Shymanska,N.V., Ventura Santos,C., Spitz,D. and Lemke,E.A. (2016) Debugging Eukaryotic Genetic Code Expansion for Site-Specific Click-PAINT Super-Resolution Microscopy. *Angewandte Chemie International Edition*, **55**, 16172–16176.
39. Sasaki,H.M., Tadakuma,H. and Tomari,Y. (2018) Single-Molecule Analysis for RISC Assembly and Target Cleavage. In Okamura,K., Nakanishi,K. (eds), *Argonaute Proteins: Methods and Protocols*, Methods in Molecular Biology. Springer, New York, NY, pp. 145–164.
40. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods*, **9**, 676–682.
41. Mazo,G. (2021) QuickFigures: A toolkit and ImageJ PlugIn to quickly transform microscope images into scientific figures. *PLOS ONE*, **16**, e0240280.
42. Ishikawa-Ankerhold,H., Ankerhold,R. and Drummen,G. (2014) Fluorescence Recovery After Photobleaching (FRAP). In *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd.
43. Hwang,L.C., Hohlbein,J., Holden,S.J. and Kapanidis,A.N. (2009) Single-Molecule FRET: Methods and Biological Applications. In Hinterdorfer,P., Oijen,A. (eds), *Handbook of Single-Molecule Biophysics*. Springer US, New York, NY, pp. 129–163.
44. Preus,S., Noer,S.L., Hildebrandt,L.L., Gudnason,D. and Birkedal,V. (2015) iSMS: single-molecule FRET microscopy software. *Nature Methods*, **12**, 593–594.
45. Sahoo,H., Roccatano,D., Zacharias,M. and Nau,W.M. (2006) Distance Distributions of Short Polypeptides Recovered by Fluorescence Resonance Energy Transfer in the 10 Å Domain. *J. Am. Chem. Soc.*, **128**, 8118–8119.
46. Schuler,B., Lipman,E.A., Steinbach,P.J., Kumke,M. and Eaton,W.A. (2005) Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence. *PNAS*, **102**, 2754–2759.
47. Miyazono,Y., Hayashi,M., Karagiannis,P., Harada,Y. and Tadakuma,H. (2010) Strain through the neck linker ensures processive runs: a DNA□kinesin hybrid nanomachine study. *The EMBO Journal*, **29**, 93–106.
48. Lambert,T.J. (2019) FPbase: a community-editable fluorescent protein database. *Nat Methods*, **16**, 277–278.
49. Carrasco,J., Antón,R., Valbuena,A., Pantoja-Uceda,D., Mukhi,M., Hervás,R., Laurents,D.V., Gasset,M. and Oroz,J. (2023) Metamorphism in TDP-43 prion-like domain determines chaperone recognition. *Nat Commun*, **14**, 466.
50. Guo,W., Chen,Y., Zhou,X., Kar,A., Ray,P., Chen,X., Rao,E.J., Yang,M., Ye,H., Zhu,L., *et al.* (2011) An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nature Structural & Molecular Biology*, **18**, 822–830.
51. Maharana,S., Wang,J., Papadopoulos,D.K., Richter,D., Pozniakovsky,A., Poser,I., Bickle,M., Rizk,S.,

Guillén-Boixet,J., Franzmann,T.M., *et al.* (2018) RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science*, **360**, 918–921.
52. Borgia,A., Borgia,M.B., Bugge,K., Kissling,V.M., Heidarsson,P.O., Fernandes,C.B., Sottini,A., Soranno,A., Buholzer,K.J., Nettels,D., *et al.* (2018) Extreme disorder in an ultrahigh-affinity protein complex. *Nature*, **555**, 61–66.
53. Meng,F., Bellaiche,M.M.J., Kim,J.-Y., Zerze,G.H., Best,R.B. and Chung,H.S. (2018) Highly Disordered Amyloid-β Monomer Probed by Single-Molecule FRET and MD Simulation. *Biophysical Journal*, **114**, 870–884.
54. Metskas,L.A. and Rhoades,E. (2020) Single-Molecule FRET of Intrinsically Disordered Proteins. *Annual Review of Physical Chemistry*, **71**, 391-414.
55. Kapanidis,A.N., Lee,N.K., Laurence,T.A., Doose,S., Margeat,E. and Weiss,S. (2004) Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *PNAS*, **101**, 8936–8941.
56. Johnson,B.S., Snead,D., Lee,J.J., McCaffery,J.M., Shorter,J. and Gitler,A.D. (2009) TDP-43 Is Intrinsically Aggregation-prone, and Amyotrophic Lateral Sclerosis-linked Mutations Accelerate Aggregation and Increase Toxicity. *J. Biol. Chem.*, **284**, 20329–20339.
57. French,R.L., Grese,Z.R., Aligireddy,H., Dhavale,D.D., Reeb,A.N., Kedia,N., Kotzbauer,P.T., Bieschke,J. and Ayala,Y.M. (2019) Detection of TAR DNA-binding protein 43 (TDP-43) oligomers as initial intermediate species during aggregate formation. *J. Biol. Chem.*, **294**, 6696–6709.
58. Marzano,N.R., Paudel,B.P., van Oijen,A.M. and Ecroyd,H. (2022) Real-time single-molecule observation of chaperone-assisted protein folding. *Science Advances*, **8**, eadd0922.
59. Kellner,R., Hofmann,H., Barducci,A., Wunderlich,B., Nettels,D. and Schuler,B. (2014) Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. *PNAS*, **111**, 13355–13360.
60. Rüdiger,S., Schneider-Mergener,J. and Bukau,B. (2001) Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *The EMBO Journal*, **20**, 1042–1050.
61. Tan,C., Niitsu,A. and Sugita,Y. (2023) Highly Charged Proteins and Their Repulsive Interactions Antagonize Biomolecular Condensation. *JACS Au*, **3**, 834–848.
62. Hall,D. (2019) On the nature of the optimal form of the holdase-type chaperone stress response. *FEBS Letters*, **594**, 43-66.
63. Panganiban,B., Qiao,B., Jiang,T., DelRe,C., Obadia,M.M., Nguyen,T.D., Smith,A.A.A., Hall,A., Sit,I., Crosby,M.G., *et al.* (2018) Random heteropolymers preserve protein function in foreign environments. *Science*, **359**, 1239–1243.

Table and Figure Legends

Figure 1. DNAJA2 and Hero11 suppress aggregation of TDP-43 in HEK293T cells

(A) Schematic of the HEK293T cell-based filter trap assay for TDP-43 aggregation. Plasmids for GFP-TDP43ΔNLS and the indicated protein were co-transfected into HEK293T cells and incubated for 48 h. Cells were lysed, and aggregates captured on 0.2 μm filter membrane for detection via immunoblotting for GFP.

(B) Representative microscope images of GFP-TDP43ΔNLS signals in HEK293T cells with the indicated co-transfections, at 20x optical magnification.

(C) Immunoblots for aggregated GFP-TDP43ΔNLS in HEK293T cells detected with α-GFP antibody. For each filter trap membrane (three independent experiments), each sample is loaded three times at two-fold serial dilution.

(D) Quantification of GFP-TDP43ΔNLS WT aggregation amount relative to the density in the GST co-transfection in (C). Bar plots represent the mean and standard deviations as the bar heights and error bars, respectively. Dots represent the values of individual experiments.

(E) Quantification of GFP-TDP43ΔNLS A315T aggregation, relative to the density in the GST co-transfection. Bar plots represent the mean and standard deviations as the bar heights and error bars, respectively. Dots represent the values of individual experiments.

Figure 2. DNAJA2 and Hero11 suppress large immobile aggregates of TDP-43

(A) Representative microscope images of GFP-TDP43ΔNLS signals in HEK293T cells, with the indicated co-transfections, at 100x optical magnification. Top row, GFP-TDP43ΔNLS signals. Bottom row, brightfield.

(B) Total of puncta area (μm²) of GFP-TDP43ΔNLS in cytoplasm per cell. Each dot represents one image. Error bars represent the mean and standard deviations calculated from thirty images.

(C) Cytoplasmic punctum size in each of the co-transfections. Each dot represents a punctum. Error lines represent the median and 95% confidence intervals, chosen due to the asymmetrical distribution, calculated from all puncta. The dotted line represents 20 μm², red text indicates the percentage of aggregates in each co-transfection above this line. There are 1156, 1107, 79, and 281 puncta for the GST, Hero11KR→G, DNAJA2 and Hero11 co-transfections, respectively.

(D) FRAP recovery curves of puncta in GST co-transfection. Each trace represents the recovery curve for an aggregate, from -5 s pre-bleach to 150 s post-bleach (where an area within the punctum was bleached at 0 s), and colored according to aggregate size.

(E) Mobile fractions binned according to aggregate size. Mobile fractions were extracted from fits of exponential recovery functions to each curve. Each dot represents an aggregate. Boxplots represent the median and interquartile range, whiskers represent the 1.5x interquartile range.

Figure 3. Hero11 and DNAJA2 promote extended conformations of the TDP-43 LCD

(A) Schematic of TIRF-smFRET observation of the LCD (orange) on full-length TDP-43. Note that donor (green) and acceptor dyes (red) are randomly incorporated at the two sites but only

donor/acceptor-labeled molecules were selected for analysis (see methods), of which the case where the donor is before and acceptor after the LCD is displayed here.

(B) Representative smFRET microscope image (from DNAJA2 incubation). Donor (green) and acceptor (red) emissions under donor excitation are overlaid here, where candidate single molecules are highlighted in the yellow circles.

(C) Example FRET trace selected for analysis (from Hero11 incubation). The top plot is the corrected fluorescence intensity of the Cy3 donor (green) and ATTO647N acceptor (red) signals under excitation of the donor dye. The fluorescence intensity of the ATTO647N acceptor signal under excitation of the acceptor dye is in magenta. The middle plot is the dye stoichiometry (purple), where signals ~0.4 come from dual labeled single molecules. The bottom plot is the FRET efficiency (cyan), with the ideal path from a hidden markov model (blue) overlaid. Shaded background in this plot corresponds to the binning used in dynamics analyses in Figure 4. The vertical dotted line shows when the first photobleaching event occurs (in this case, of the acceptor dye).

(D – K) FRET efficiency distribution of the TDP-43 LCD WT (D – G) or A315T (H – K), incubated with BSA (D, H), Hero11KR→G (E, I), DNAJA2 (F, J) or Hero11 (G, K). Density plots represent mean ± SEM, as the solid lines and shaded areas, respectively for three repeat experiments. Histograms contain all observations. The total number of molecules in each histogram are (D), 150; (F), 160; (G), 178; (H), 168; (H), 101; (I), 255; (J), 357; and (K), 216. Note values below 0 and above 1 arise from background subtraction uniformly across molecules of an experiment, and from the kernel in the density plot, and still correspond to valid observations (see methods).

Figure 4. DNAJA2 and Hero11 stabilize dynamic and extended states of the LCD

(A, B) Transition density plots or histograms of the WT (A) or A315T (B) LCD, under incubation in the indicated protein from left to right: BSA, Hero11KR→G, DNAJA2, and Hero11. FRET states were extracted from idealized FRET traces obtained from HMM, and binned into four ranges shown in Figure 3C. The ratio of transition events between the FRET state bins are indicated by the shading.

(C, D) Dwell times of FRET states as represented by empirical cumulative distribution functions for the WT (C) or A315T (D) LCD, divided by FRET state bin. Trace colors indicated the incubated protein. Each trace was fit to a single exponential model, and estimated median dwell

time half lives, $t_{1/2}$, along with 95% credible intervals in parentheses for the fitted model are displayed. The fitting is shown in Supplementary Figure 2.

Supplementary Figure 1. Genetic code expansion in HEK293T cells for dye-labeling of TDP-43

(A) Schematic of HALO-TDP43 with unnatural amino acids incorporated around the LCD after elution via TEV protease cleavage. Boxes indicate TDP-43 regions and tags: N-terminal domain (NTD), two RNA recognition motifs (RRM1 and RRM2), low-complexity domain (LCD), HaloTag, and 2xHA tag. The last amino acid of the HaloTag is the 23rd amino acid from the start of the TDP-43 sequence. Sites of unnatural amino acid (endo BCN-L-Lysine) incorporation are indicated by the violet triangles, where arginine 268 was replaced with endo BCN-L-Lysine (R268X), while the other site was incorporated at residue 420 (420X; the sixth residue after the end of the TDP-43 sequence, residue 414), in the linker before a 2xHA-tag. The yellow triangle marks the site of the A315T mutation.

(B) Amino acid sequence of (A). Highlight color corresponds to the regions in (A). X represents endo BCN-L-Lysine, the A highlighted in yellow is alanine 315 which is mutated to threonine in A315T mutant. The N-terminal glycine is left over from the TEV protease cleavage, the asterisk marks the end of the sequence.

(C) Dye-labeling of HALO-TDP43. HEK293T cells expressing FLAG-tev-HALO-TDP43 incorporated with unnatural amino acids were lysed and incubated on anti-FLAG Dynabeads. Proteins were dye-labeled on bead and eluted via cleavage with TEV protease in naïve HEK293T cell lysate (note the multiple bands in the elution fraction of the CBB stain). Copper-free click chemistry-based, dye-labeling of endo BCN-L-Lysine with tetrazine-Cy3 (Donor) and tetrazine-ATTO-647N (Acceptor) was confirmed by fluorescence imaging in gel with excitation at 520 nm and 620 nm, respectively, before staining with Coomassie Brilliant Blue (CBB). The two bands in the Elution fraction correspond to HALO-TDP43 with only one or two dyes incorporated (lower and upper bands, respectively). The Beads fraction indicates proteins that did not elute from the Dynabeads, and the higher molecular weights likely correspond to dye-labeled but uncleaved FLAG-tev-HALO-TDP43.

Supplementary Figure 2. Fitting of exponential distribution models to dwell time distributions

(A, B) Model fits for the empirical cumulative distribution functions for WT (A) and A315T (B) LCDs. Columns are FRET state bins and rows are the incubated proteins. Each dot represents an observed dwell time, to which exponential distribution models were fitted to obtain median and

95% credible interval estimates of the rate parameter, and are plotted as the red dashed lines and shaded area.

Figure 1

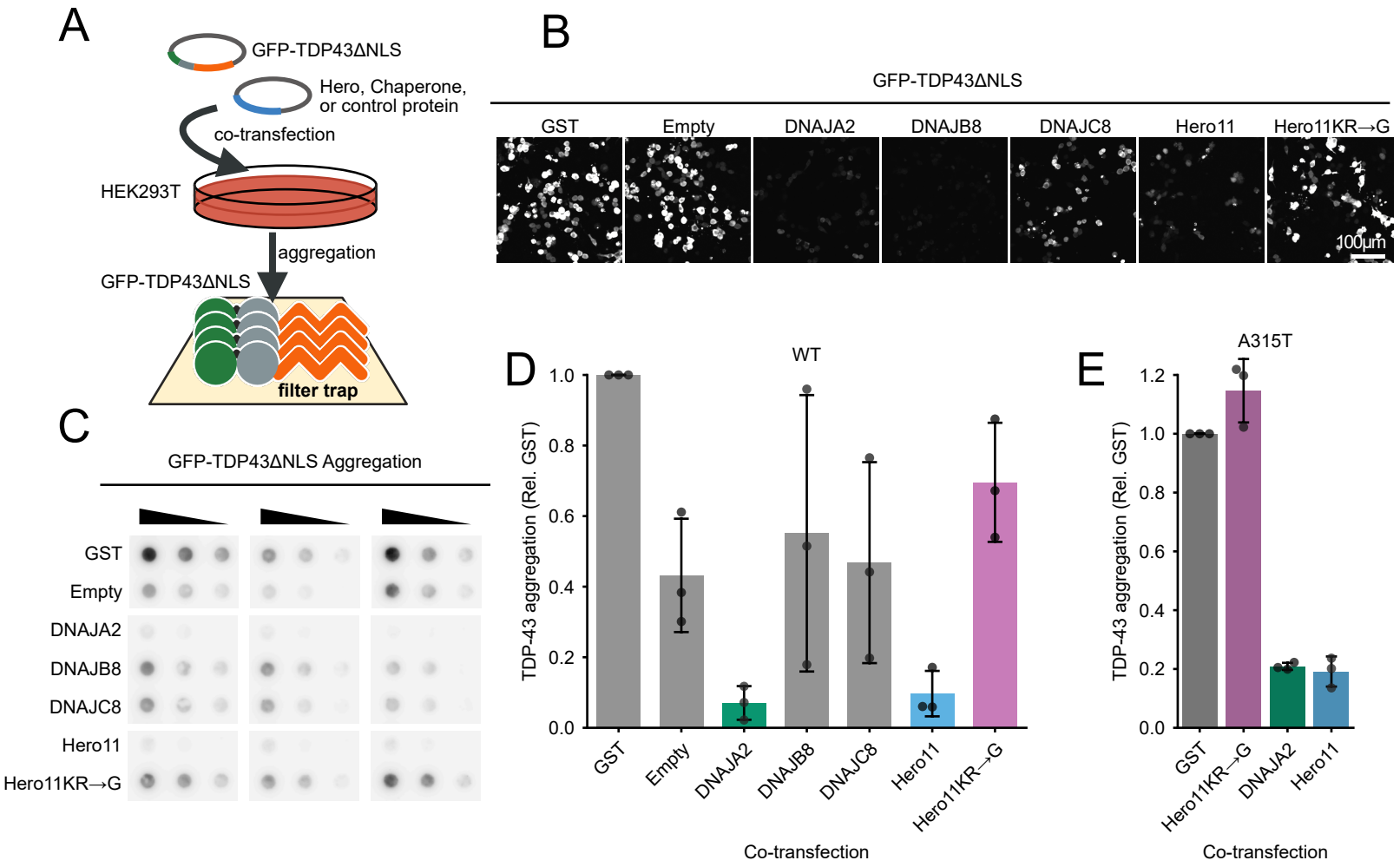

A
GFP-TDP43ΔNLS
Hero, Chaperone, or control protein
co-transfection
HEK293T
aggregation
GFP-TDP43ΔNLS
filter trap
B
GFP-TDP43ΔNLS
GST
Empty
DNAJA2
DNAJB8
DNAJC8
Hero11
Hero11KR→G
100μm
C
GFP-TDP43ΔNLS Aggregation
GST
Empty
DNAJA2
DNAJB8
DNAJC8
Hero11
Hero11KR→G
D
WT
TDP-43 aggregation (Rel. GST)
1.0
0.8
0.6
0.4
0.2
0.0
GST
Empty
DNAJA2
DNAJB8
DNAJC8
Hero11
Hero11KR→G
Co-transfection
E
A315T
TDP-43 aggregation (Rel. GST)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
GST
Hero11KR→G
DNAJA2
Hero11
Co-transfection

Figure 2

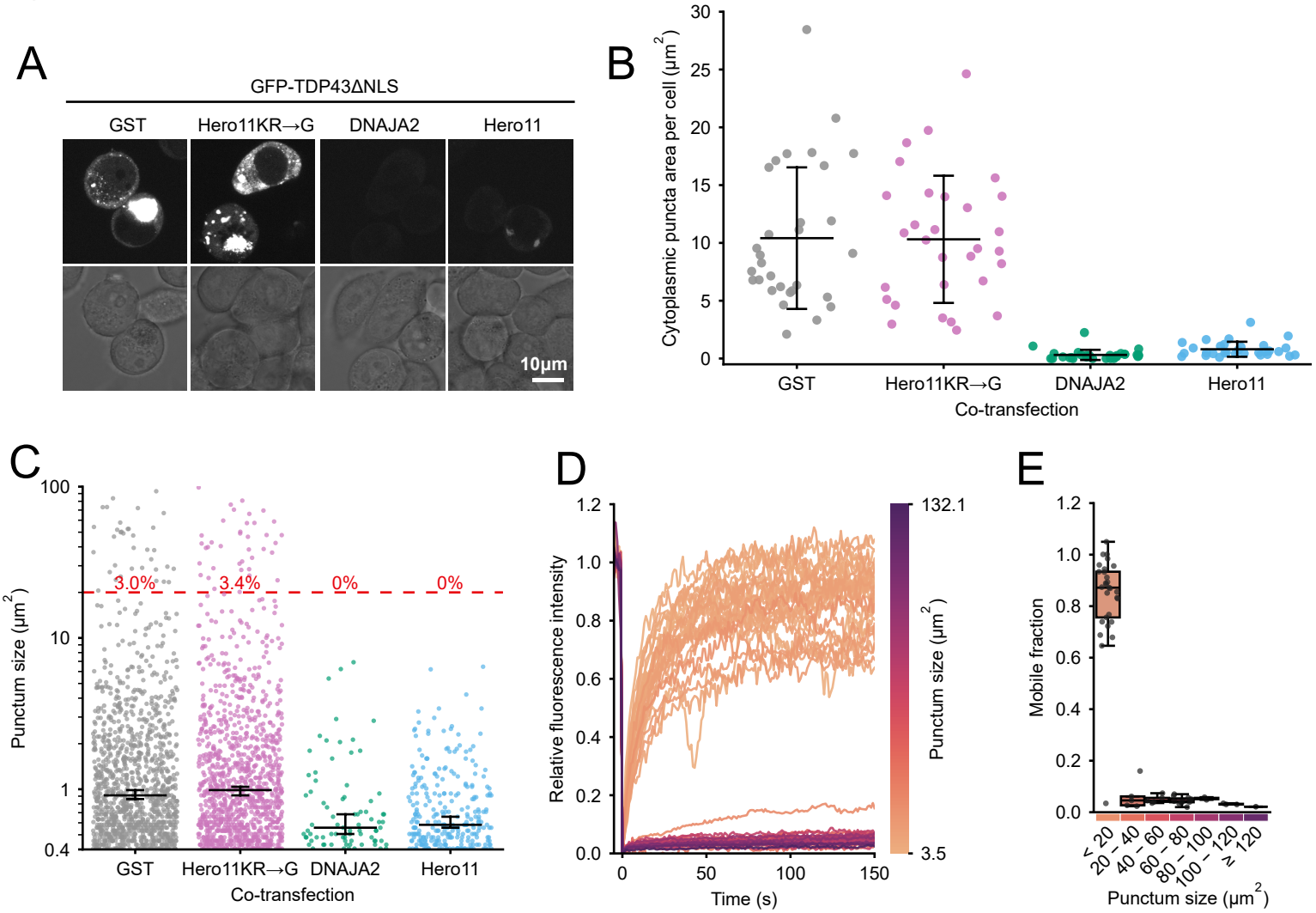

A
GFP-TDP43ΔNLS
GST
Hero11KR→G
DNAJA2
Hero11
10μm
B
Cytoplasmic puncta area per cell (μm²)
30
25
20
15
10
5
0
GST
Hero11KR→G
DNAJA2
Hero11
Co-transfection
C
Punctum size (μm²)
100
10
1
0.4
3.0%
3.4%
0%
0%
GST
Hero11KR→G
DNAJA2
Hero11
Co-transfection
D
Relative fluorescence intensity
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Time (s)
132.1
3.5
Punctum size (μm²)
E
Mobile fraction
1.2
1.0
0.8
0.6
0.4
0.2
0.0
< 20
20 – 40
40 – 60
60 – 80
80 – 100
100 – 120
≥ 120
Punctum size (μm²)

Figure 3

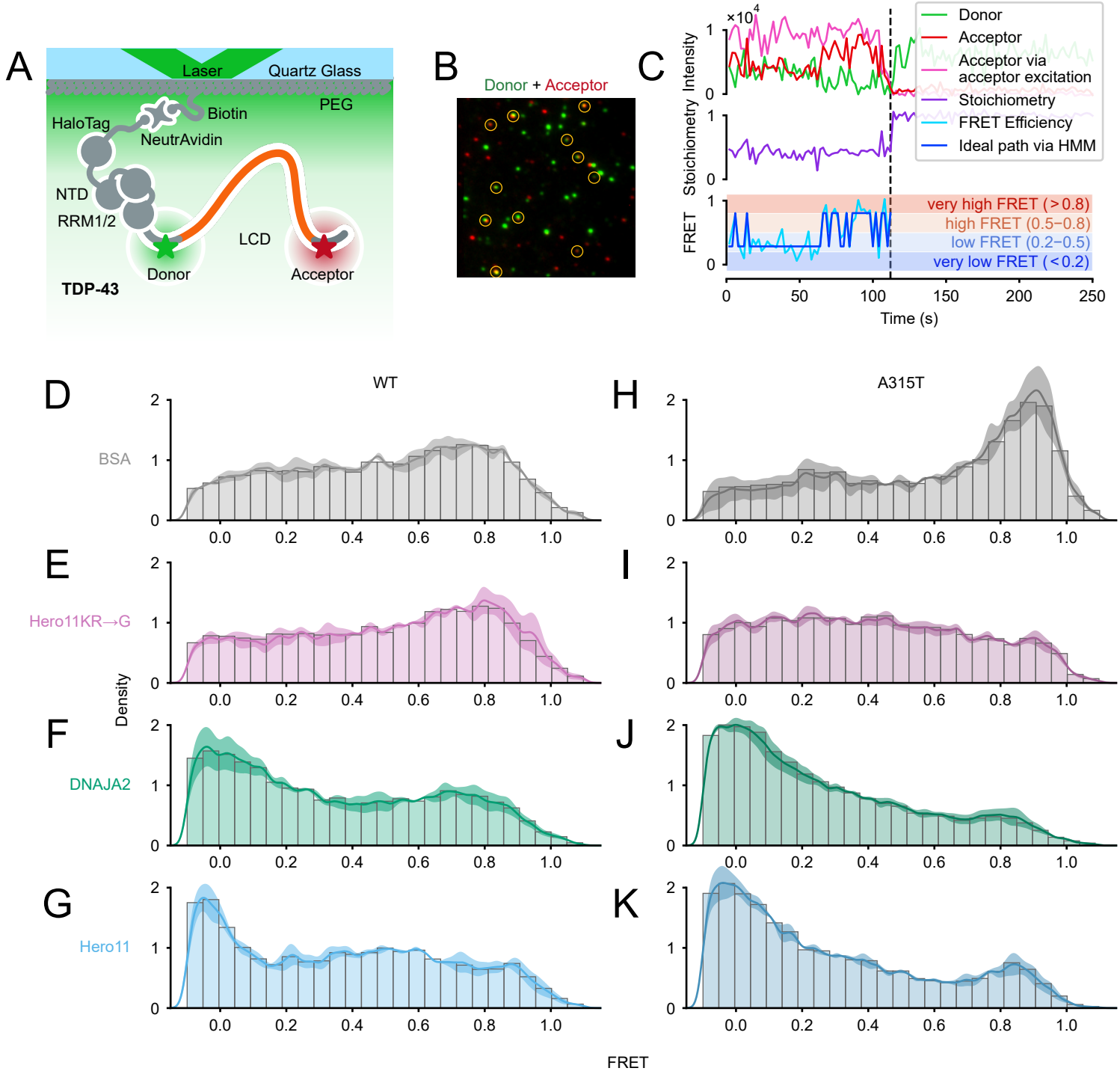
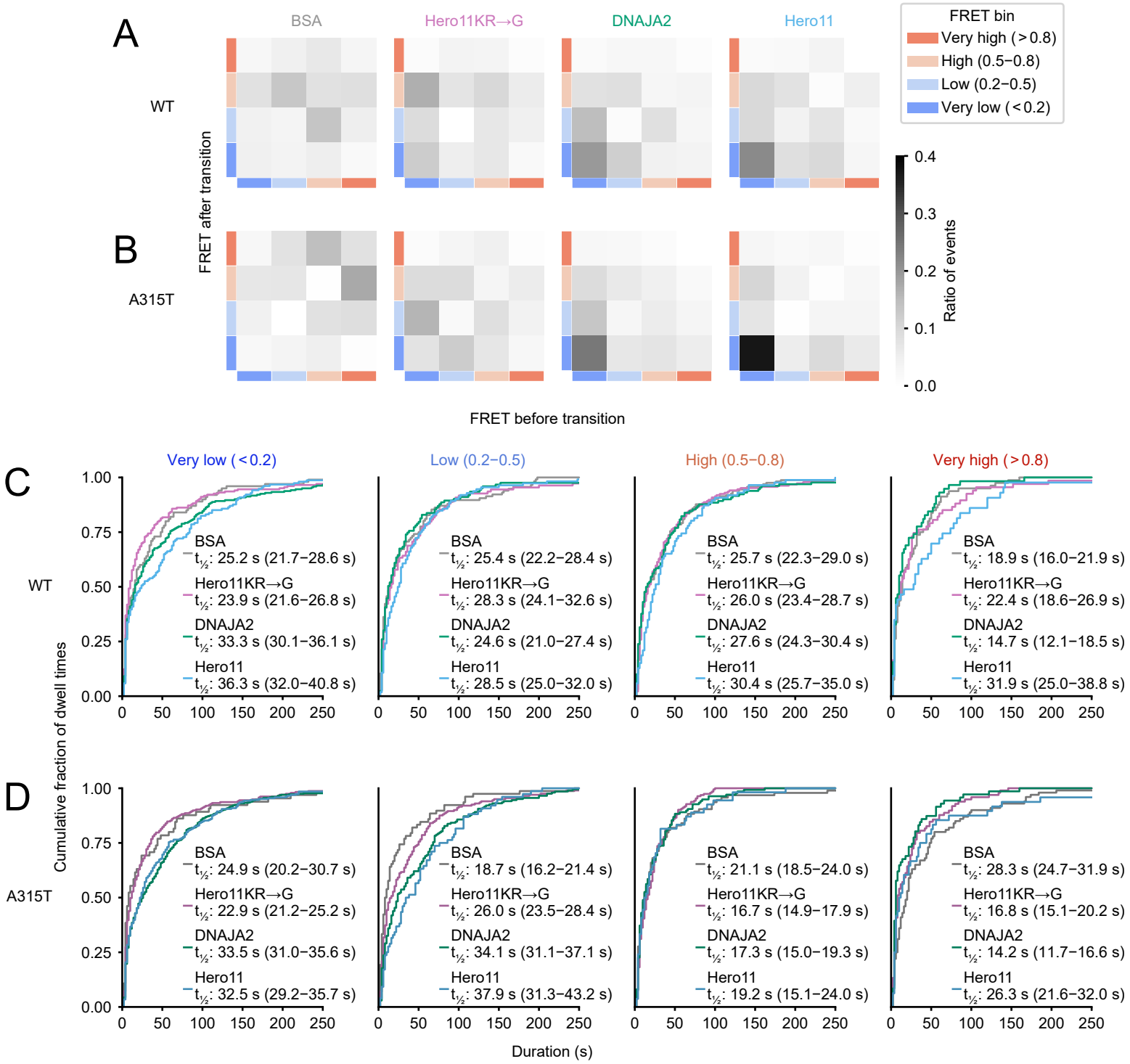

A
Laser
Quartz Glass
PEG
HaloTag
Biotin
NeutrAvidin
NTD
RRM1/2
LCD
Donor
Acceptor
TDP-43
B
Donor + Acceptor
C
×10^4
Intensity
Stoichiometry
FRET
Donor
Acceptor
Acceptor via acceptor excitation
Stoichiometry
FRET Efficiency
Ideal path via HMM
very high FRET (>0.8)
high FRET (0.5–0.8)
low FRET (0.2–0.5)
very low FRET (<0.2)
Time (s)
D
WT
BSA
E
Hero11KR→G
F
DNAJA2
G
Hero11
H
A315T
I
J
K
Density
FRET

Figure 4

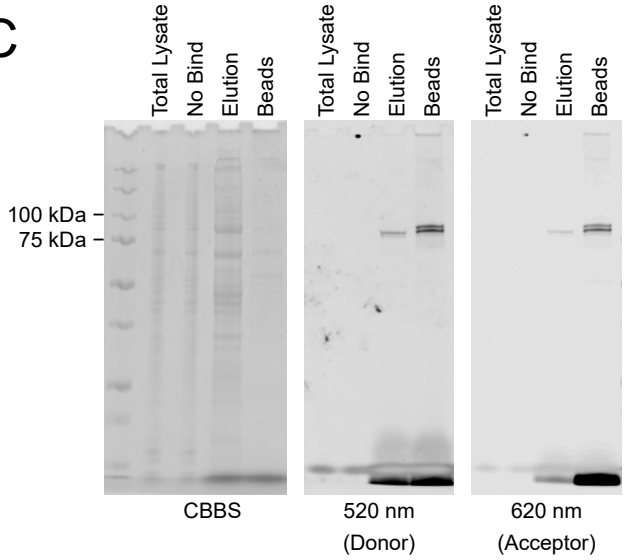
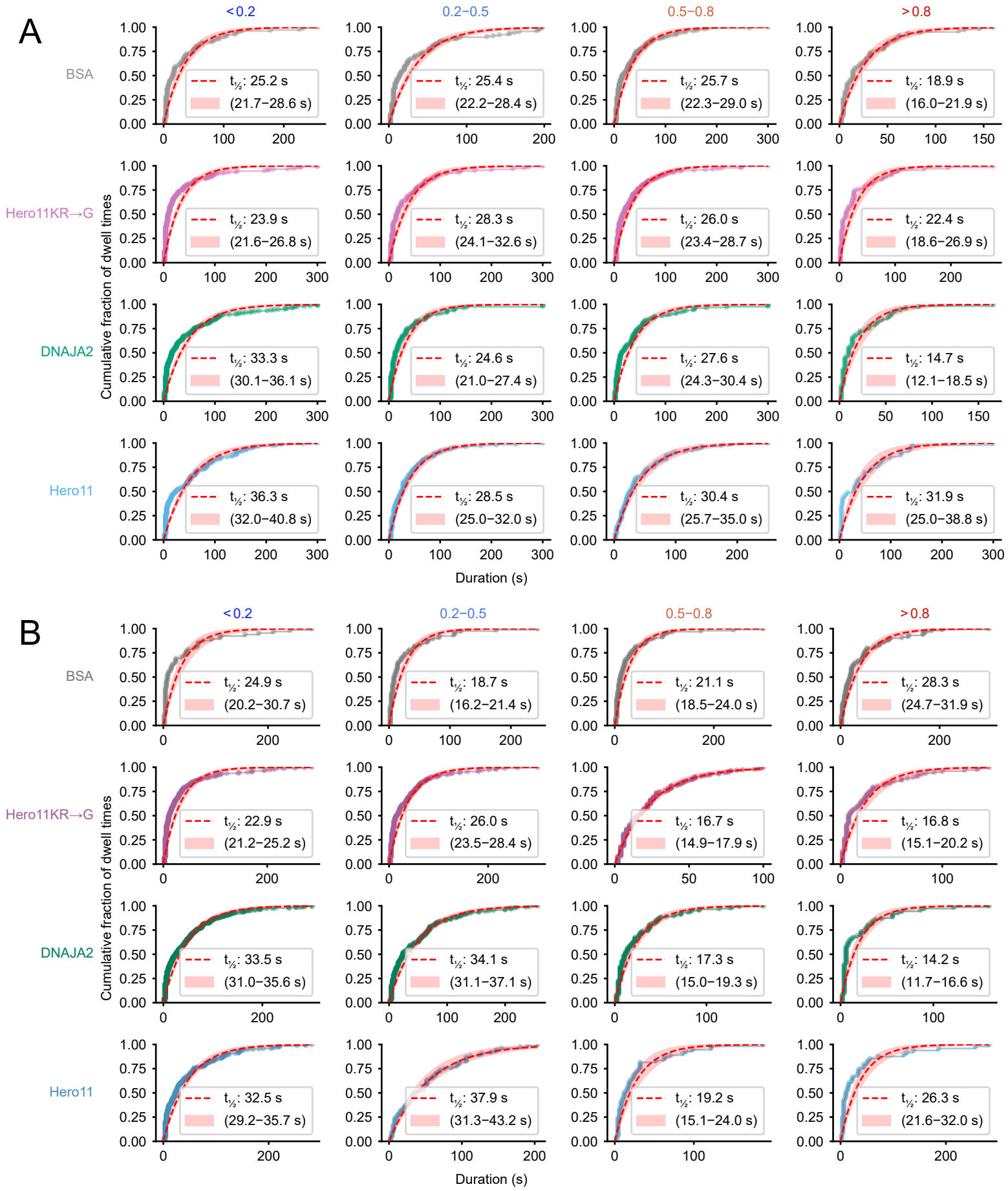

A
B
BSA
Hero11KR→G
DNAJA2
Hero11
WT
A315T
FRET after transition
FRET before transition
FRET bin
Very high (>0.8)
High (0.5–0.8)
Low (0.2–0.5)
Very low (<0.2)
Ratio of events
0.4
0.3
0.2
0.1
0.0
C
D
Very low (<0.2)
Low (0.2–0.5)
High (0.5–0.8)
Very high (>0.8)
Cumulative fraction of dwell times
Duration (s)
BSA
t½: 25.2 s (21.7–28.6 s)
Hero11KR→G
t½: 23.9 s (21.6–26.8 s)
DNAJA2
t½: 33.3 s (30.1–36.1 s)
Hero11
t½: 36.3 s (32.0–40.8 s)
BSA
t½: 25.4 s (22.2–28.4 s)
Hero11KR→G
t½: 28.3 s (24.1–32.6 s)
DNAJA2
t½: 24.6 s (21.0–27.4 s)
Hero11
t½: 28.5 s (25.0–32.0 s)
BSA
t½: 25.7 s (22.3–29.0 s)
Hero11KR→G
t½: 26.0 s (23.4–28.7 s)
DNAJA2
t½: 27.6 s (24.3–30.4 s)
Hero11
t½: 30.4 s (25.7–35.0 s)
BSA
t½: 18.9 s (16.0–21.9 s)
Hero11KR→G
t½: 22.4 s (18.6–26.9 s)
DNAJA2
t½: 14.7 s (12.1–18.5 s)
Hero11
t½: 31.9 s (25.0–38.8 s)
BSA
t½: 24.9 s (20.2–30.7 s)
Hero11KR→G
t½: 22.9 s (21.2–25.2 s)
DNAJA2
t½: 33.5 s (31.0–35.6 s)
Hero11
t½: 32.5 s (29.2–35.7 s)
BSA
t½: 18.7 s (16.2–21.4 s)
Hero11KR→G
t½: 26.0 s (23.5–28.4 s)
DNAJA2
t½: 34.1 s (31.1–37.1 s)
Hero11
t½: 37.9 s (31.3–43.2 s)
BSA
t½: 21.1 s (18.5–24.0 s)
Hero11KR→G
t½: 16.7 s (14.9–17.9 s)
DNAJA2
t½: 17.3 s (15.0–19.3 s)
Hero11
t½: 19.2 s (15.1–24.0 s)
BSA
t½: 28.3 s (24.7–31.9 s)
Hero11KR→G
t½: 16.8 s (15.1–20.2 s)
DNAJA2
t½: 14.2 s (11.7–16.6 s)
Hero11
t½: 26.3 s (21.6–32.0 s)

Supplementary Figure 1

A

B

C



Supplementary Figure 2