Highly Charged Proteins and Their Repulsive Interactions Antagonize Biomolecular Condensation

Cheng Tan, Ai Niitsu, and Yuji Sugita*

Cite This: *JACS Au* 2023, 3, 834−848

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Biomolecular condensation is involved in various cellular processes; therefore, regulation of condensation is crucial to prevent deleterious protein aggregation and maintain a stable cellular environment. Recently, a class of highly charged proteins, known as heat-resistant obscure (Hero) proteins, was shown to protect other client proteins from pathological aggregation. However, the molecular mechanisms by which Hero proteins protect other proteins from aggregation remain unknown. In this study, we performed multiscale molecular dynamics (MD) simulations of Hero11, a Hero protein, and the C-terminal low-complexity domain (LCD) of the transactive response DNA-binding protein 43 (TDP-43), a client protein of Hero11, under various conditions to examine their interactions with each other.

We found that Hero11 permeates into the condensate formed by the LCD of TDP-43 (TDP-43-LCD) and induces changes in conformation, intermolecular interactions, and dynamics of TDP-43-LCD. We also examined possible Hero11 structures in atomistic and coarse-grained MD simulations and found that Hero11 with a higher fraction of disordered region tends to assemble on the surface of the condensates. Based on the simulation results, we have proposed three possible mechanisms for Hero11's regulatory function: (i) In the dense phase, TDP-43-LCD reduces contact with each other and shows faster diffusion and decondensation due to the repulsive Hero11−Hero11 interactions. (ii) In the dilute phase, the saturation concentration of TDP-43-LCD is increased, and its conformation is relatively more extended and variant, induced by the attractive Hero11−TDP-43-LCD interactions. (iii) Hero11 on the surface of small TDP-43-LCD condensates can contribute to avoiding their fusion due to repulsive interactions. The proposed mechanisms provide new insights into the regulation of biomolecular condensation in cells under various conditions.

KEYWORDS: *biomolecular condensation, liquid−liquid phase separation, highly charged protein, molecular dynamics, multiscale MD simulation, coarse-grained model*
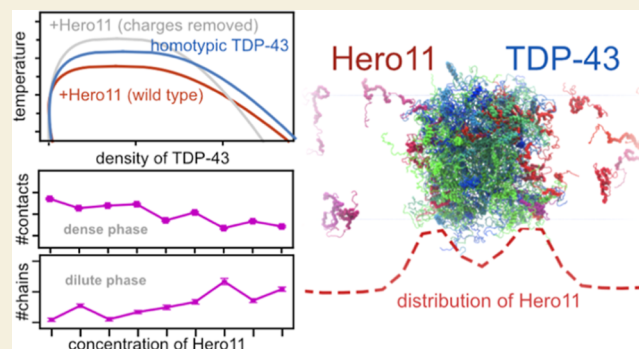
INTRODUCTION

Proteins and nucleic acids can form biomolecular condensates via liquid−liquid phase separation (LLPS).[1] LLPS is observed in many cellular processes and is considered a general mechanism for the compartmentalization of biomolecules.[2] Biomolecular condensation can be either functional or dysfunctional. The functional condensation facilitates genome organization,[3,4] assembly of transcription machinery at super-enhancers,[5] cellular signaling,[6] and response to environmental stress,[7] whereas the dysfunctional condensation is involved in the assembly of pathogenetic ribonucleoprotein granules that leads to irreversible liquid-to-solid transitions.[8] For example, the transactive response DNA-binding protein 43 (TDP-43)[9] can form a dysfunctional condensate via LLPS, whose consequent aggregation is associated with several neuro-degenerative diseases,[9−11] including limbic-predominant age-related TDP-43 encephalopathy,[12] amyotrophic lateral sclerosis,[13] and frontotemporal dementia.[14] Thus, understanding the molecular mechanisms underlying the regulation of biomolecular condensation in cells is crucial in both basic cellular biology and medical science fields.

Experimental studies have revealed several general molecular features of biological LLPS. Multivalency,[15] usually driven by the low complexity intrinsically disordered regions (IDRs) of proteins, is one of the central principles underlying LLPS.[16−19] The C-terminal prion-like domain of TDP-43 is a typical IDR that plays an essential role in TDP-43 phase transitions.[20,21] Changes in environmental conditions such as pH,[22] mutations,[23,24] and post-translational modifications, such as
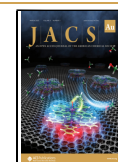
acetylation[25,26] and phosphorylation,[27] can regulate biological LLPS. Liquid droplets formed by RNA-binding proteins and RNAs can be modulated by the RNA−protein ratio and the length of RNA.[28,29] Interestingly, short-bait RNAs composed of TDP-43 target sequences have been used to prevent neurotoxic aggregation of TDP-43.[21] Similarly, a recent experimental study identified a class of highly charged intrinsically disordered proteins (IDPs), known as heat-resistant obscure (Hero) proteins, which can protect proteins from pathological aggregation.[30] TDP-43 aggregation can be suppressed by coexpressed Hero proteins in cells.[30] Interestingly, shuffling of the amino acid sequence in Hero proteins does not change their activity if the amino acid composition is maintained.[30] The experimental results suggest the importance of charged residues in the antiaggregation functions of Hero proteins; however, their detailed molecular mechanisms remain unknown.

Theoretical and computational studies have been conducted to examine the phase behavior of IDPs to understand the formation of LLPS in atomic detail.[31,32] The stickers-and-spacers model provides a framework for describing the composition and interactions among multivalent molecules.[31,33,34] Random phase approximation theory is used to determine the connections between the charge patterning and phase transition of IDPs.[35,36] However, most studies of charged IDPs or polyampholytes focus on the promotion of phase separation due to the attractive interactions between proteins or between proteins and RNAs.[35−39] Improvements in all-atom force fields[40−42] and residue-level coarse-grained (CG) models[43−46] have proven to be effective tools for investigating biomolecular interactions in dilute and dense phases. Notably, the hydrophobicity scale (HPS) CG model[43] has shown great potential for capturing sequence-specific thermodynamic properties.[47−51] Multiscale molecular dynamics (MD) simulations have been used to study the effects of a transient α-helical secondary structure[24,52] in the TDP-43 C-terminal IDR. The α-helical structure promotes the formation of a dense phase and induces a higher critical temperature for TDP-43.[52] Despite these results of TDP-43 phase separation, how other proteins regulate the TDP-43 condensate still needs to be fully understood.

In this study, we performed multiscale MD simulations of biomolecular condensates while targeting the antiaggregation functions of Hero proteins. We selected the C-terminal IDR of TDP-43 as a protein that forms high-density condensates and examined the mechanism by which Hero11 regulates the physical properties of condensations. We found that the electrostatic repulsion between Hero proteins could be a key factor, which prompted us to carry out simulations of a mixture of TDP-43 C-terminal IDR and a Hero11 mutant with all positive residues mutated to glycine. The repulsive interactions between Hero11 proteins and the attractive interactions between TDP-43 and Hero11 regulate the formation of TDP-43 droplets and kinetics of the protein. The secondary structure dependence of the antiaggregation function was examined using atomistic and CG MD simulations. Based on our results, we propose three possible molecular mechanisms for the antiaggregation function of Hero11 and discuss their general applicability to other Hero proteins as well as highly charged biomolecules, such as RNAs.

RESULTS

Single-Chain Simulations of TDP-43 and Hero11

To understand the phase behavior of TDP-43 and the antiaggregation regulation function of Hero11, we performed atomistic and CG simulations of the low-complexity domain (LCD) of TDP-43 (residues 261−414, hereafter referred to as TDP-43-LCD for convenience) and the full-length Hero11 of wild-type (WT) and mutants (sequences listed in Table S1). Previous studies have identified a short α-helical piece in TDP-43-LCD and shown that this secondary structure facilitates the phase separation of TDP-43-LCD.[24,52] Interestingly, Alpha-Fold2[53] predicts the structure of TDP-43 containing an α-helix in the same region. For Hero11, however, there are discrepancies in its secondary structure predictions between AlphaFold2 and the IDP-predicting tools IUPred3,[54] flDPnn,[55] and PONDR.[56] AlphaFold2 suggests two α-helices located in regions 17−25 (confidence score < 70) and 38−72 (confidence score > 90) of Hero11, whereas IUPred3, flDPnn, and PONDR anticipate the whole chain with IDR propensities > 0.5 (Figure S1).

To examine the secondary structure of Hero11, we employed the CHARMM36m force field[41] and performed atomistic simulations of full-length Hero11 in an explicit solvent (see the Computational Methods section for detailed simulation conditions). We started with two different initial structures: one with two α-helices obtained from the AlphaFold Protein Structure Database[53] (entry Q9UNZ5), and the other as a random coil reconstructed from a short CG simulation using the HPS model (Figure S2A,B). After energy minimization and equilibration, we performed ~1 μs simulations for each system. In the simulations starting from the AlphaFold2-predicted structure, the first α-helix (residues 17−25) unfolded after ~0.6 μs, while the second α-helix was only partly sustained (residues 38−55; Figure S2A). In the simulation starting from a random coil, a short piece of α-helical structure at residues 44−50 emerged in the early stage and was preserved until the end (Figure S2B). The structures sampled from these two simulations were quite different, suggesting that the 1 μs simulation is insufficient to achieve equilibrium. To obtain statistically more reliable results, it may be necessary to perform much longer MD simulations (>100 μs) or use an enhanced conformational sampling algorithm to explore the protein conformational landscapes.[57−59] Regardless of the initial structure, a short region (residues 44−50) was predicted as a helical structure in all MD simulations, suggesting that Hero11 is neither completely random nor stably folded in solution.

Next, we used CG simulations to explore the conformations of Hero11 and TDP-43-LCD. Because our atomistic simulations with a classic force field did not provide a concrete conclusion about the secondary structure of Hero11, we decided to model two extreme cases: one harboring the two α-helices predicted by AlphaFold2, and the other with the entire peptide as an IDP. Note that our all-atom simulations suggest that only part of the helical structures are formed in a single chain of Hero11. Particularly, residues 17−25 seem to be totally unfolded. However, these results do not guarantee the exact conformation of Hero11 to be observed in more crowded environments such as biomolecular condensates. Therefore, in CG simulations, we considered the two extreme cases of secondary structures of Hero11 to cover more possibilities. Here, we employed the HPS potentials[43] and the HPS-Urry

A Hero11-wild type (Hero11-WT)

MAQGQ**RK**FQÂH**K**PA**K**S**K**TAÂAÂAS**E**K**N**RGP**RK**GG**R**VIAP**KK**A**R**VVQQQ**KL**K
KNL**E**VGI**RKK**I**E**H**D**VVM**K**ASSSLP**KK**LALLKAPA**KKK**GAÂAÂATSS**K**TPS

e^-: 4
e^+: 27

○ Large particles: α-helix
● Red: Asp, Glu
● Blue: Arg, Lys

B TDP-43 low-complexity domain (TDP-43-LCD)

EPKHNSN**R**QL**E**RSG**R**FGGNPGGFGNQGGFGNS**R**GGGAGLGNNQGSNMGGG
MNFGAFSIN**PAMMAAAQAALQSS**WGMMGMLASQQNQSGPSGNNQNQGNMQ
REPNQAFGSGNNSYSGSNSGAAIGWGSASNAGSGSGFNGGFGSSM**D**S**K**SS
GWGM

e^-: 4
e^+: 7

○ Large particles: α-helix
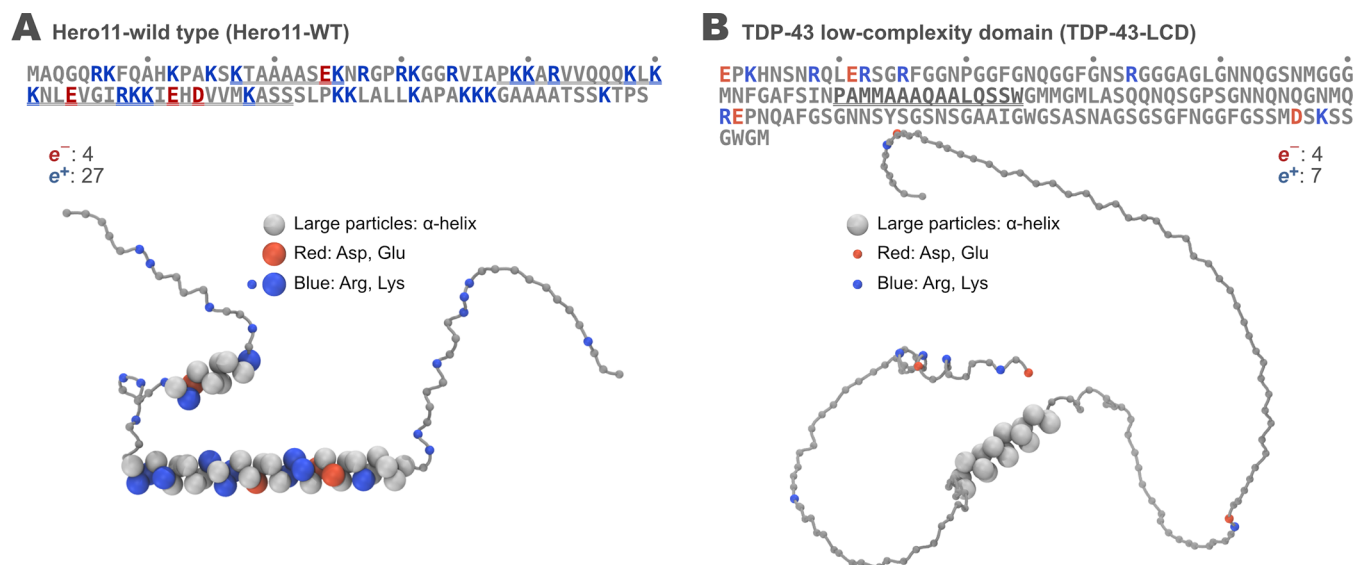● Red: Asp, Glu
● Blue: Arg, Lys

Figure 1. Sequence and CG modeling of the heat-resistant obscure protein 11-wild type (Hero11-WT) and transactive response DNA-binding protein 43 LCD (TDP-43-LCD). (A) Sequence and CG modeling of Hero11-WT. (B) is same as (A) but for TDP-43-LCD. Each amino acid residue is represented by one bead in the simulations. Particles are shown in different sizes to represent different treatments: small particles are modeled with the HPS, and large particles are modeled with AICG2+ to maintain the α-helical structures based on AlphaFold2 predictions.

parameter set[47] to define the interactions in the IDR. For the α-helical regions, we used the AICG2+ model[60] to restrain the folded secondary structures (Figure 1).

We first conducted simulations of single-chain TDP-43-LCD at temperatures ranging from 250 to 350 K. At all temperatures, the α-helices were well-maintained during the simulations (Figure S3B). The IDR tails of TDP-43-LCD were flexible, and the whole structure showed an increase in the radius-of-gyration (R_g) as the temperature increased (Figure S3C). These results show that the combination of AICG2+ and HPS models provides a way to simulate proteins consisting of both folded domains and IDRs, although its accuracy should be validated by comparison with experimental data such as small-angle X-ray scattering (SAXS) profiles. Using the same method, we simulated a single-chain Hero11-WT with α-helical structures (named Hero11-WT-α) or without any α-helical structure (named Hero11-WT-noα). Contact analysis demonstrated that when applied, the AICG2+ potentials preserved the folded structures of the α-helices (Figure S4B). Interestingly, we found that Hero11-WT-α had larger R_g values than Hero11-WT-noα at all simulated temperatures (Figure S4C). This result showed that the α-helical secondary structures resulted in relatively extended conformations of Hero11-WT.

We further validated our method using a multiple-chain system of TDP-43-LCD. Following the slab simulation strategy,[43] we generated a system composed of 100 chains of TDP-43-LCD in a 180 Å × 180 Å × 3000 Å box with periodic boundary conditions. The system was simulated at different temperatures from 260 to 350 K. During the simulations, we monitored the protein density distribution along the z-axis (the longest dimension). We divided the z-axis into small bins of size Δz = 30 Å and computed the local density of TDP-43-LCD molecules in each bin (see Computational Methods section for details). Based on the density analysis, we determined the dense and dilute phases at temperatures lower than the critical temperature (T_c). The time evolution of TDP-43-LCD density shows that TDP-43-LCD forms a stable

condensate at 295 K (Figure 2A). In the dense phase, TDP-43-LCD forms extensive intra- and intermolecular interactions (Figure 2B). In the dilute phase, TDP-43-LCD chains are primarily free and seldom form contact with each other. The density data at different temperatures were then used to plot the phase diagram of TDP-43-LCD, as shown in Figure 2C (solid line and dots). As a control, we built a model of TDP-43-LCD without holding the α-helical structure and carried out the same simulation and analysis. Our results showed that the α-helical secondary structures raised the critical temperature of TDP-43-LCD by ~5 K (Figure 2C). Detailed contact analysis revealed that the α-helical region contributed more to interchain interactions than the IDRs (Figure S5). These results are consistent with those of a previous study,[52] even though the modeling methods and parameters differ. Because the existence of the α-helix had already been proven,[52] we kept the TDP-43-LCD model with the α-helical structure for the rest of the simulations in this study.

Previous simulation studies have investigated the R_g of IDR in the condensate and its relationship with conformational entropy gain during phase transition.[61,62] Inspired by these studies, we also calculated the R_g values of TDP-43-LCD from our simulations (Figure S6). Consistent with earlier findings,[61,62] TDP-43-LCD had larger R_g values in the condensate than in the dilute phase. While dilute phase R_g grows monotonically with temperature, dense phase R_g is almost independent of temperature (Figure S6A). We also computed the standard deviation of the R_g values (s_{R_g}) and found that the dense phase R_g has larger standard deviation ($s_{R_g,dense}$ = 7.479 Å) than the dilute phase R_g ($s_{R_g,dilute}$ = 7.043 Å) (Figure S6B). These results show that the structure of TDP-43-LCD in the dense phase is not an elongated and trapped conformation due to restraining intermolecular interactions. Instead, larger s_{R_g} value suggests higher conformational fluctuations of TDP-43-LCD in the condensate.
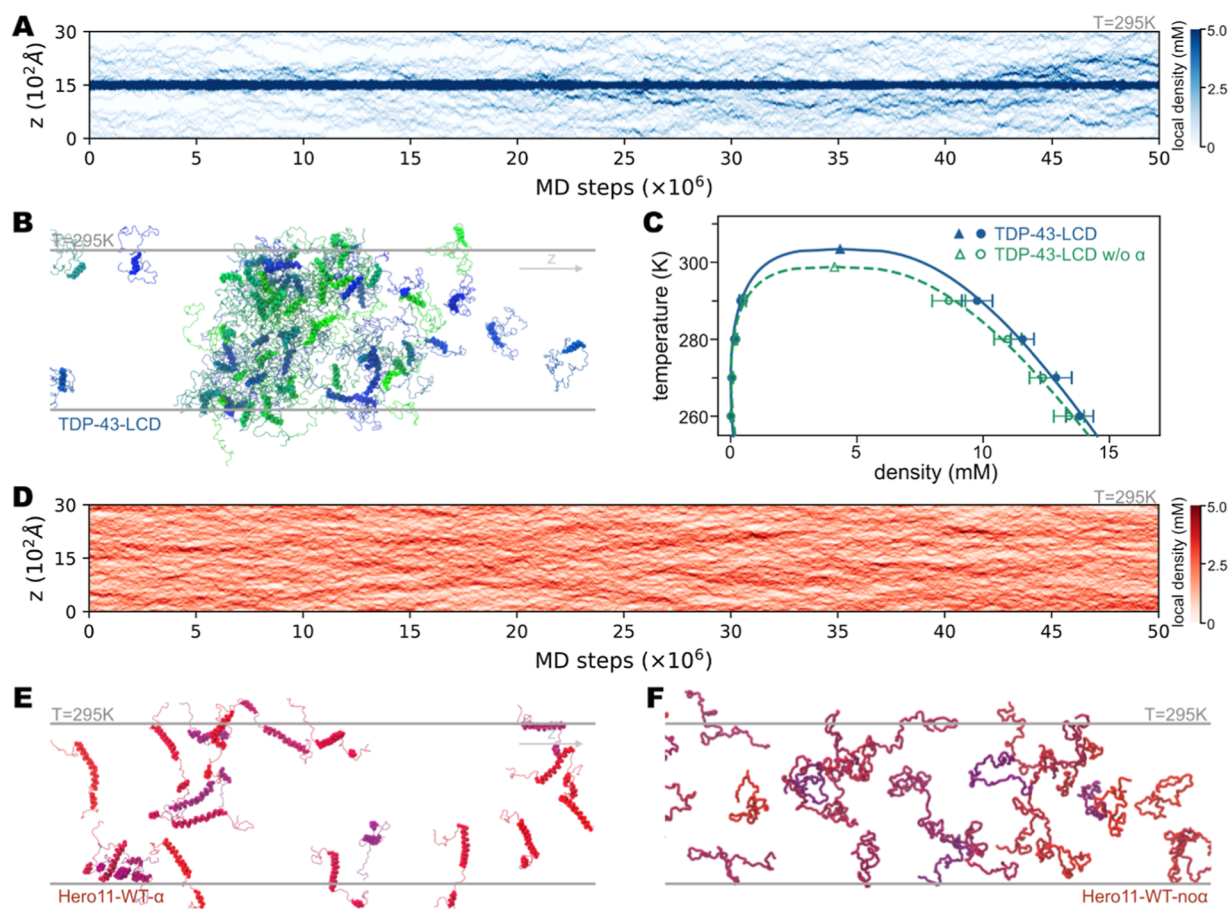
Figure 2. Multitemperature slab simulations of homotypic TDP-43-LCD and Hero11-WT. (A) Time evolution of the local density of TDP-43-LCD along the z-axis (longest dimension of the simulation box). In each simulation, 100 chains of TDP-43-LCD were put in a box of $180 \times 180 \times 3000$ Å^3. The z dimension is divided into 100 bins with length $\Delta z = 30$ Å, and local density is calculated in each $180 \times 180 \times 30$ Å^3 bin. Intensity of the blue color represents local density, as indicated by the color bar. (B) Last structure of a 5×10^7-step slab simulation of TDP-43-LCD at 295 K. For clarity, we only show the region around the condensate of TDP-43-LCD. Chains are shown in different colors to distinguish them. (C) Phase diagram of TDP-43-LCD with (solid line and dots) and without (dashed line and empty dots) the α-helical structure modeled by AICG2+. Circles represent the simulated densities, and triangles represent the fitted values (see Computational Methods for details). Error bars represent standard deviations. (D) is same as (A) but for Hero11-WT-α. Density is represented by the intensity of red, as indicated by the color bar. (E) Part of the last structure of a 5×10^7-step slab simulation of Hero11-WT-α at 295 K. Chains are shown in different colors. (F) is same as (E) but for Hero11-WT-noα. In (E,F), Hero11-WT-α and Hero11-WT-noα means Hero11-WT modeled with and without α-helix, respectively.

Phase Behavior of WT Hero11

Next, we conducted CG simulations for the multiple chains of Hero11-WT-α and Hero11-WT-noα. Hero11-WT contains many positively charged residues (Figure 1A), and the net charge of a single chain is +23. Experiments have shown that Hero11-WT has a lower tendency to form condensates than typical LLPS or aggregation-prone proteins, such as TDP-43-LCD.[30] Consistently, our simulation results demonstrated the same property of Hero11-WT and connected the Hero11-WT sequence with its phase behavior. The density profile of Hero11-WT-α simulated at 295 K showed that Hero11-WT-α did not assemble into a stable concentrated phase (Figure 2D) as TDP-43-LCD did at the same temperature (Figure 2A). Similarly, Hero11-WT-noα did not form condensates at 295 K. A snapshot of the Hero11-WT-α (Hero11-WT-noα) structure at 295 K is shown in Figure 2E,F, demonstrating an unconsolidated structure with rare interchain contacts.

The simulation temperature was then tentatively decreased to determine the T_c of Hero11-WT. We found that the Hero11-WT proteins showed no sign of condensation in simulations at temperatures higher than 200 K (Figure S7). We

were able to determine an exceptionally low critical temperature using the current model; however, because the applicability of the HPS and AICG2+ models are yet to be validated at low temperatures, we want to avoid discussing those possibly unphysical results here. Nevertheless, our simulations indicate that Hero11-WT bears a much lower T_c than LLPS-prone IDPs, such as TDP-43.

Phase Behavior of Hero11 Mutants

Mutation experiments have suggested that the positive charges on Hero11-WT dominate its phase behavior.[30] Following the experimental design, we also simulated several Hero11 mutants, including the "KRless" mutant, where all lysine and arginine residues are substituted to glycine, and the "scrambles," in which the sequence of Hero11 is randomly shuffled.[30]

Similar to Hero11-WT, we first carried out all-atom simulations for Hero11-KRless using the CHARMM36m force field.[41] Two initial structures were used: either from the AlphaFold2 predicted structure or a random coil. During the 1 μs simulations of the predicted structure, we found that the α-helix was partially unfolded, and only a short part
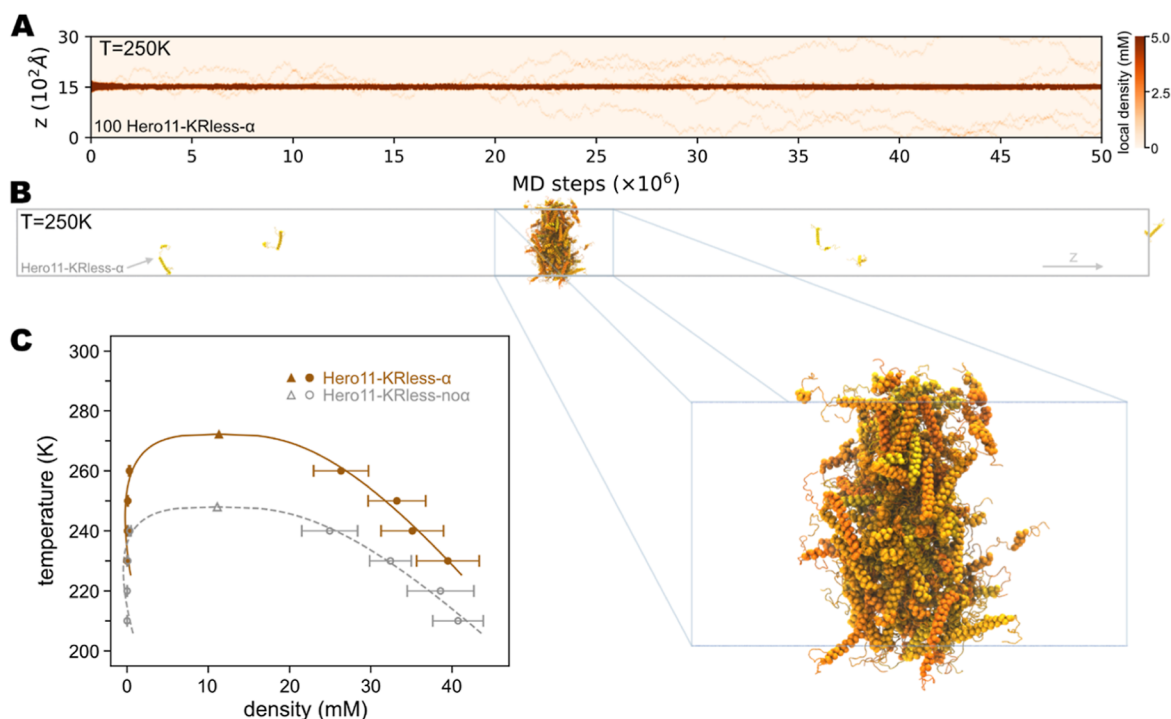
Figure 3. Phase behavior of the Hero11-KRless mutants. (A) Representative graph of the local density of Hero11-KRless-α over time. Intensity of the orange color represents local density, as indicated by the color bar. Local density has the same definition as Figure 2A. (B) Last structure of a 5 × 10^7-step simulation of Hero11-KRless-α at 250 K. Colors from yellow to orange shade are used to distinguish the different chains. (C) Phase diagram of Hero11-KRless mutants. Circles represent the simulated values, and triangles represent the fitted values (see Computational Methods for details). Error bars represent standard deviations.

(residue 43–55) was retained (Figure S2C). In the simulation from a random coil, some short α-helices and β-strands were formed transiently but were not stable (Figure S2D). Similar to Hero11-WT, the Hero11-KRless simulations also did not present convergent results. Therefore, we consider two extreme conformations (Hero11-KRless-α and Hero11-KRless-noα) in the downstream CG simulations.

We performed CG simulations for systems containing 100 chains of Hero11-KRless-α or Hero11-KRless-noα at temperatures from 210 to 300 K. Using the same analysis method as for Hero11-WT, we plotted the time evolution of density along the z-axis (Figure 3A). Unlike Hero11-WT, Hero11-KRless-α formed a stable cluster at 250 K, with dynamic dissociation and association of single chains (see a representative snapshot at 250 K in Figure 3B). We found that α-helical structures still contributed many interactions in the dense phase of Hero11-KRless-α, as also indicated by the contact analysis (Figure S8). We then used the density and temperature data to plot the phase diagram for KRless mutants (Figure 3C). Removing the positive charges from the sequence significantly raises the critical temperature of Hero11 (Figures 3C and S7). In other words, electrostatic repulsion due to the high content of positively charged residues in Hero11-WT dominates its phase behavior. To further verify this conclusion, we mutated all charged residues in Hero11 (including 27 positive and 4 negative residues; Figure 1A) to glycine and modeled them as Hero11-chargeless-α and Hero11-chargeless-noα, respectively. Compared to the Hero11-KRless mutants, which have a net charge of −4e, the Hero11-chargeless mutants were neutralized and showed even higher T_c than the Hero11-KRless mutants (Figure S9). These results supported the hypothesis that interchain electrostatic interactions between the large amount

of positively charged residues suppress self-condensate formation at physiological temperatures.[30]

In addition to the KRless and chargeless mutants, in which the fraction of charges was altered, we also simulated the mutants in which the sequence was randomly shuffled but keeping the same amount of net charge. We modeled these mutants as pure IDPs (named as Hero11-scramble n-noα, n = 1, 2, 3) because there is no reason to assume that the α-helices are preserved in the randomly shuffled sequences. Simulation results of these mutants (Figure S9) showed that all three scrambles tested had similar phase behaviors to those of Hero11-WT-noα that they do not form condensates at physiological temperatures. These results agree with previous experiments showing that the charge content is more crucial than the amino acid sequence in determining the phase behavior of Hero proteins.[30]

Regulation of TDP-43-LCD Condensates by Hero11-WT

Next, we explored the mechanism for the antiaggregation effect of Hero11 on TDP-43-LCD, which was found in previous experiments.[30] Using the slab method, we simulated a system consisting of 100 Hero11-WT-α and 100 TDP-43-LCD. We prepared a consolidated mixture of Hero11-WT-α and TDP-43-LCD as the initial structure (see Computational Methods section for details). Simulations were performed at different temperatures ranging from 250 to 350 K. The time evolution of the densities of TDP-43-LCD and Hero11-WT-α showed that the dense region of the two proteins was highly coupled (Figure 4A,B). At 250 K, TDP-43-LCD and Hero11-WT-α chains jointly formed high-density condensates, as indicated by the darkest blue and red regions (Figure 4A). The final structure of this simulation showed that some Hero11-WT-α chains joined the well-defined condensate formed by TDP-43-
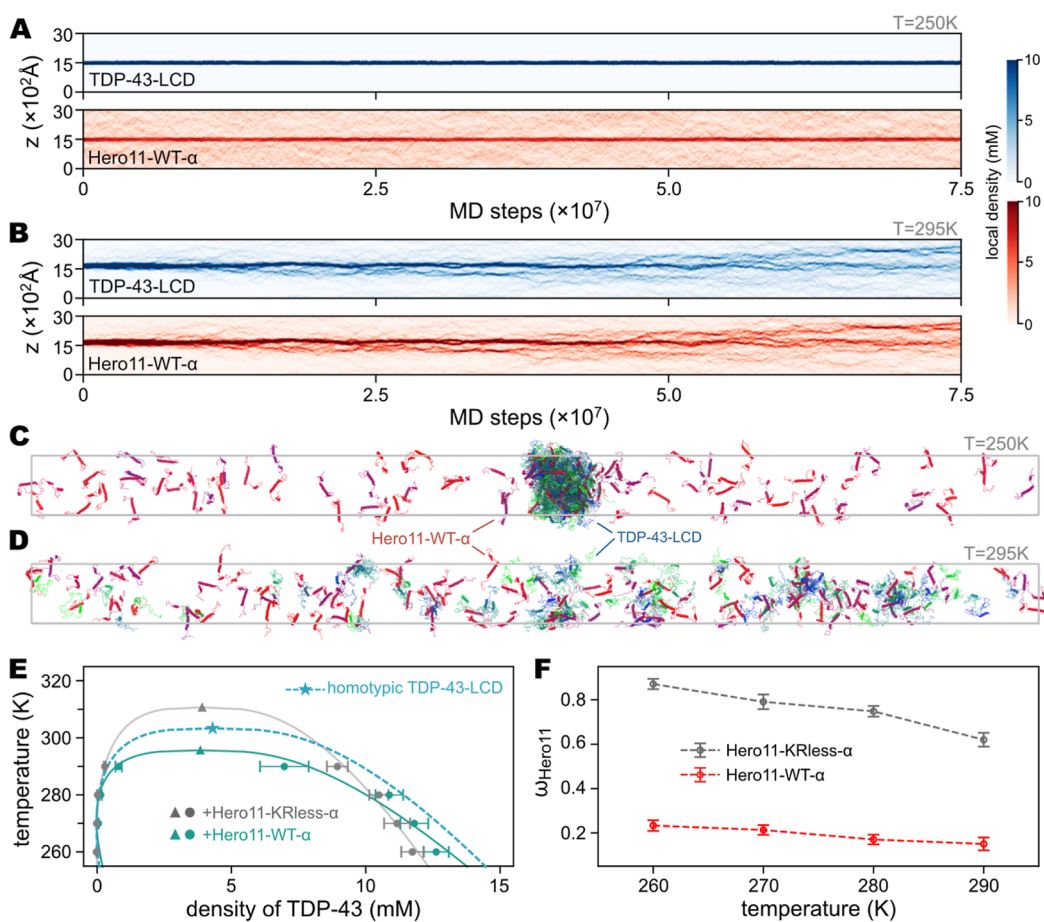
Figure 4. Slab simulations of heterotypic systems consisting of 100 TDP-43-LCD and 100 Hero11. (A,B) show representative local density profiles of TDP-43-LCD and Hero11-WT-α simulated at 250 and 295 K, respectively. Intensity of blue (red) represents the local density of TDP-43-LCD (Hero11-WT-α), as indicated by the color bars. (C,D) show the last structures of the 250 and 295 K trajectories, respectively. (E) Phase diagram of TDP-43-LCD in the presence of Hero11-WT-α (cyan) or Hero11-KRless-α (gray). Circles represent the quantities calculated from the simulations, with error bars representing standard deviations of the densities. Triangles are the fitted critical points (see Computational Methods for details). Star and dashed line represent the fitted critical point and phase diagram of the homotypic TDP-43-LCD, which is the same as Figure 2C. (F) Fraction of Hero11 located in the condensate of TDP-43-LCD (ω_{Hero11}) as a function of the simulation temperature. Error bars represent standard deviations. In (E,F), only the last one-third of each simulation was used for analysis.

LCD (Figure 4C). In contrast, at 295 K, the dense phase formed by TDP-43-LCD and Hero11-WT-α began to dissolve after the beginning of the simulation. The largest cluster diverged into several smaller clusters in the last ~2.5 × 10⁷ steps (Figure 4B). Hero11-WT-α still formed contacts with TDP-43-LCD, but a clear boundary between the two phases disappeared (Figure 4B,D). Compared to the homotypic TDP-43-LCD system at 295 K (Figure 2A), these results illustrated how Hero11-WT-α changes the behavior of TDP-43-LCD. Because no stable condensation was observed at 295 K, we did not use simulation data at this temperature for the following analysis. Instead, we used results obtained at 260, 270, 280, and 290 K, at which two phases can be clearly determined in the heterotypic systems containing TDP-43-LCD and Hero11-WT. The phase diagram of TDP-43-LCD in the presence of Hero11-WT-α (Figure 4E, cyan solid line) showed that the addition of Hero11-WT-α decreased the T_c of TDP-43-LCD.

The mutations from Lys/Arg to Gly drastically changed the phase behavior of homotypic Hero11 (Figure 3). We then investigated how these mutations affected the cocondensation of Hero11 and TDP-43-LCD. From the simulations of a mixture of 100 Hero11-KRless-α and 100 TDP-43-LCD, we found that at 295 K, the dense phase formed by both proteins

was stably maintained (Figure S10). The phase diagram of TDP-43-LCD (Figure 4E, gray line) showed that without the high content of positively charged residues, Hero11-KRless-α promotes T_c of TDP-43-LCD to ~310 K. Removing all charged residues, Hero11-chargeless-α raises T_c of TDP-43-LCD to a higher value of ~320 K (Figure S11A). Interestingly, we found that both Hero11-WT-α and Hero11-KRless-α entered the condensate of TDP-43-LCD, consistent with a previously proposed "scaffold and ligand" model.[63−65] However, the fraction of "ligand"-Hero11 (ω_{Hero11}) was smaller in the case of Hero11-WT-α than in Hero11-KRless-α, owing to the strong interchain repulsion (Figure 4F). These results show that the electrostatic interactions between positively charged residues in Hero11-WT-α determine its regulatory functions in the phase behavior of other proteins.

Mechanism Underlying the Antiaggregation Function of Hero11

Next, we investigated the detailed mechanisms of the antiaggregation function of Hero11 by performing simulations of heterotypic systems composed of 100 TDP-43-LCDs and variant numbers (10n, n = 1, 2, ..., 9) of Hero11-WT-α. We first plotted the binary phase diagram (Figure 5A) based on
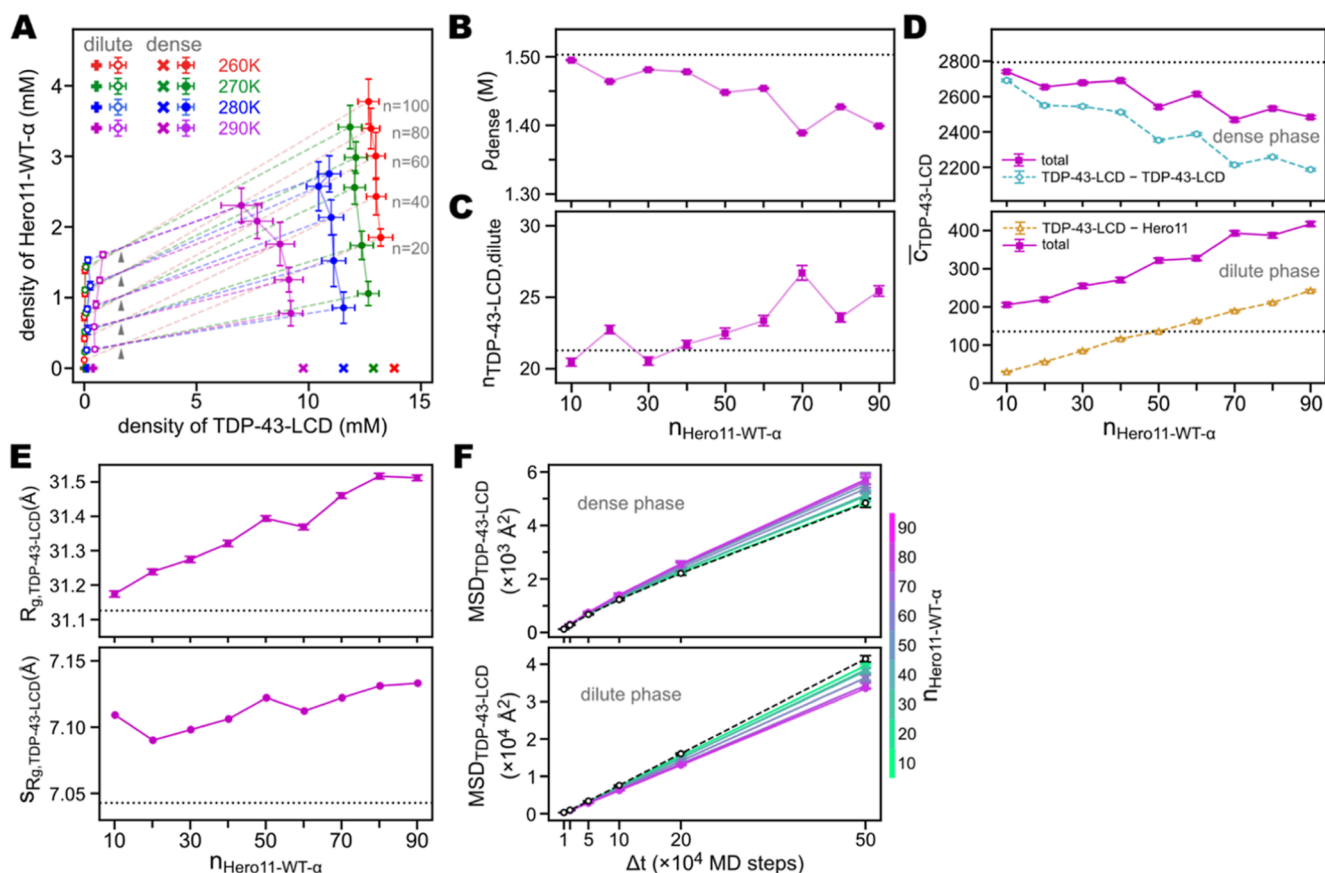
Figure 5. Effects of Hero11-WT-α on TDP-43-LCD condensation. (A) Binary phase diagram of TDP-43-LCD and Hero11-WT-α. Data dots calculated from the same simulations are connected by dashed lines. Note that simulations with the same number of Hero11-WT-α intersect at the same point (indicated by gray triangles) corresponding to the averaged densities of proteins in the entire simulation box. The number of Hero11-WT-α corresponding to each triangle is n = 20, 40, 60, 80, and 100 from the lower to upper panel, as indicated by the numbers next to the high-density dots for 260 K (red). Colors of lines and dots represent the simulation temperatures, as indicated by the legends. "×" and "+" symbols represent the densities in the dense and dilute phase of the homotypic TDP-43-LCD system, respectively. Error bars represent standard deviations. (B) Total particle density in the dense phase (ρ_{dense}) as a function of the number of Hero11-WT-α in the simulation ($n_{\mathrm{Hero11\text{-}WT\text{-}\alpha}}$). (C) Number of TDP-43-LCD in the dilute phase ($n_{\mathrm{TDP\text{-}43\text{-}LCD,dilute}}$) as a function of $n_{\mathrm{Hero11\text{-}WT\text{-}\alpha}}$. (D) Average number of contacts formed by any TDP-43-LCD chain ($\bar{c}_{\mathrm{TDP\text{-}43\text{-}LCD}}$) with other chains of TDP-43-LCD or Hero11-WT-α. Upper and lower panels are for TDP-43-LCDs in the dense and dilute phases, respectively. The total number of $\bar{c}_{\mathrm{TDP\text{-}43\text{-}LCD}}$ (purple) is the sum of TDP-43-LCD−TDP-43-LCD (cyan) and TDP-43-LCD−Hero11 (orange) contacts. For clarity, the data are not shown for TDP-43-LCD−Hero11 in the dense phase, while in the dilute phase, TDP-43-LCD−TDP-43-LCD is hidden. (E) Radius of gyration of TDP-43-LCD ($R_{\mathrm{g,TDP\text{-}43\text{-}LCD}}$, upper) and its standard deviation (($s_{R_{\mathrm{g,TDP\text{-}43\text{-}LCD}}}$, lower) in the dilute phase as a function of $n_{\mathrm{Hero11\text{-}WT\text{-}\alpha}}$. Horizontal dotted lines in (B−E) are corresponding values calculated from homotypic TDP-43-LCD simulations. Error bars in (B−E) represent the standard errors of the mean. (F) MSD of TDP-43-LCD (MSD$_{\mathrm{TDP\text{-}43\text{-}LCD}}$) in the dense (upper) and dilute (lower) phases as a function of time interval Δt, respectively. Results with different $n_{\mathrm{Hero11\text{-}WT\text{-}\alpha}}$ are shown in different colors, as indicated by the color bar. Black dashed lines and empty circles show the results of homotypic TDP-43-LCD simulations. Error bars show standard deviation of the MSD values calculated from five independent simulations. Results in (B−F) are based on simulations at temperature 290 K.

the densities of TDP-43-LCD and Hero11-WT-α in different phases and at different temperatures. In the phase diagram, the low-density branch (empty circles) shows the saturation concentration of TDP-43-LCD, whereas the high-density branch (solid circles) represents the density in the condensate. As a reference, we also show the density of TDP-43-LCD analyzed from the simulations of homotypic TDP-43-LCD ("+" and "×" dots). As the temperature increased, the biphasic region shrank, suggesting an upper consolute temperature above which no phase separation occurs (Figure 5A).

We then focused on the density changes caused by adding Hero11-WT-α to the TDP-43-LCD condensate. Compared with the densities of the homotypic TDP-43-LCD system, the addition of Hero11-WT-α decreased the density of TDP-43-LCD in the dense phase, as indicated by the high-density branches in Figure 5A. This effect was more prominent at

higher temperatures (290 K) than at lower temperatures (280−260 K). On the other side, our results also showed that Hero11-WT-α increased the saturation concentration of TDP-43-LCD, as indicated by the higher TDP-43-LCD densities in the dilute phase (Figure 5A). These results show that Hero11-WT-α not only results in a lower TDP-43-LCD density of the concentrated phase but also increases the population of TDP-43-LCD in the dilute phase. The changes of the dense phase induced by Hero11-WT-α suggest an essential role of electrostatic repulsion between the Hero11-WT-α chains, whereas the effects of Heto11-WT-α in the dilute phase imply contributions from attractive interactions between Hero11-WT-α and TDP-43-LCD.

Given that the coexistence of both repulsive and attractive interactions complicates the net effect of Hero11 on TDP-43-LCD's condensation, a more detailed analysis was conducted.

Because the density changes of TDP-43-LCD induced by Hero11-WT-α are the most significant at 290 K (Figure 5A purple dots and lines), we focused on this temperature in the following discussion. We performed five more independent simulations for every 100 TDP-43-LCD + 10n Hero11-WT-α (n = 1, 2, ..., 9) system to obtain more samples. Each simulation was run for 9×10^7 steps, and the last 3×10^7 steps were used for the data analysis. At 290 K, dense and dilute phases were well defined. Thus, we could separately analyze the physical quantities of the proteins in the two phases. We first calculated the total (including all TDP-43-LCD and Hero11-WT-α) particle density in the dense phase (ρ_{dense}) as a function of the number of Hero11-WT-α included in the simulations ($n_{\text{Hero11-WT-}\alpha}$). We found that ρ_{dense} was decreased as $n_{\text{Hero11-WT-}\alpha}$ increased (Figure 5B). Importantly, compared with the homotypic TDP-43-LCD system, ρ_{dense} is lower when Hero11-WT-α is involved, showing that Hero11-WT-α loosens the condensate. We consider this phenomenon to be nontrivial because the "ligand" proteins do not necessarily decrease the density of condensates formed by the "scaffold" proteins.[34,50,65] On the other side, number of TDP-43-LCD in the dilute phase increased as $n_{\text{Hero11-WT-}\alpha}$ increased (Figure 5C), consistent with the increased dilute phase concentration of TDP-43-LCD (Figure 5A).

To decipher the regulatory function of Hero11-WT in TDP-43-LCD condensation at the molecular level, we analyzed the interchain contacts formed by the proteins. We found that the total number of all interchain contacts in the dense phase was decreased when $n_{\text{Hero11-WT-}\alpha}$ was increased (Figure S12), confirming the loosening effect of Hero11-WT-α on TDP-43-LCD condensation. Then, more specifically, we counted the number of contacts formed by every TDP-43-LCD chain ($c_{\text{TDP-43}}$) with other TDP-43-LCDs (TDP-43-LCD−TDP-43-LCD) or Hero11s (TDP-43-LCD−Hero11). We considered the average (over all structures simulated at the same condition) interchain contact number of TDP-43-LCD ($\overline{c}_{\text{TDP-43-LCD}}$) as a function of $n_{\text{Hero11-WT-}\alpha}$. Compared with the $\overline{c}_{\text{TDP-43-LCD}}$ value in the homotypic TDP-43-LCD condensate (~2800, Figure 5D, horizontal dotted line), adding Hero11-WT-α to the system reduced the number of TDP-43-LCD−TDP-43-LCD contacts (Figure 5D). As $n_{\text{Hero11-WT-}\alpha}$ increased to 90, the TDP-43-LCD−TDP-43-LCD contact number drops to ~2200 (cyan dashed line and empty dots in Figure 5D). This phenomenon can be explained by the replacement of some TDP-43-LCD−TDP-43-LCD contacts by TDP-43-LCD−Hero11-WT-α interactions in the condensate. Indeed, as the number of Hero11-WT-α increased, the TDP-43-LCD−Hero11 contact number increased, which can be interpreted from the difference between the total and the TDP-43-LCD−TDP-43-LCD contact numbers (Figure 5D, upper panel). However, our results also show an evident decrease in the total number of contacts formed by each TDP-43-LCD, which cannot be solely explained by substituting TDP-43-LCD−TDP-43-LCD contacts with TDP-43-LCD−Hero11 contacts. Considering that the repulsive electrostatic interactions result in an extraordinarily low T_c for Hero11-WT-α, it is straightforward to speculate that the repulsion between Hero11-WT-α also induces a decrease in the contact number of TDP-43-LCD. To test this assumption, we analyzed the contact numbers of TDP-43-LCD from simulations containing 100 TDP-43-LCD and 100 Hero11-KRless-α (Figure S13). We found that when all positive charges were removed, instead of reducing the total contact number of TDP-43-LCD,

Hero11-KRless-α increased $\overline{c}_{\text{TDP-43-LCD}}$ (Figure S13), consistent with the higher T_c yielded by Hero11-KRless-α (Figure 4E). These results illustrate the function of Hero11-WT-α in the concentrated phase of TDP-43-LCD that Hero11-WT-α replaces some part of the interchain contacts of TDP-43-LCD, and the strong electrostatic repulsion between Hero11-WT-α chains "loosen" the whole condensate, resulting in fewer intermolecular interactions on TDP-43-LCD.

We also analyzed the contact number of TDP-43-LCD in the dilute phase, as shown in the lower panel of Figure 5D. We found that both the total contact number and the number of TDP-43-LCD−Hero11 contacts increased as $n_{\text{Hero11-WT-}\alpha}$ increased. This result suggests that TDP-43-LCD has a higher probability of remaining in the dilute phase due to interactions with Hero11-WT-α and explains the increased saturation concentration of TDP-43-LCD in the presence of Hero11-WT-α (Figure 5A the low-density branches, and Figure 5C). We then calculated the radius of gyration of TDP-43-LCD ($R_{\text{g,TDP-43-LCD}}$) in the dilute phase to evaluate the effect of increased intermolecular contacts on the conformation of TDP-43-LCD. We found that $R_{\text{g,TDP-43-LCD}}$ expanded from ~31.2 to ~31.5 Å when $n_{\text{Hero11-WT-}\alpha}$ increased from 10 to 90 (Figure 5E, upper panel). Notably, all these $R_{\text{g,TDP-43-LCD}}$ values were larger than those calculated in the single-chain simulations (Figure 5E, upper panel). In contrast, $R_{\text{g,TDP-43-LCD}}$ in the dense phase was independent of the concentration of Hero11-WT-α and was close to the value in the homotypic TDP-43-LCD condensation (Figure S14). As discussed earlier in this work, TDP-43-LCD in the homotypic condensate has a larger mean value and variance of R_g than in the dilute phase, indicating an entropy gain during phase separation (Figure S6).[61,62] Interestingly, we found that Hero11-WT-α induced a larger standard deviation of TDP-43-LCD's R_g ($s_{R_g,\text{TDP-43-LCD}}$) in the dilute phase (Figure 5E, lower panel) and smaller $s_{R_g,\text{TDP-43-LCD}}$ in the dense phase (Figure S14, lower panel). Importantly, the difference in $s_{R_g,\text{TDP-43-LCD}}$ in the two phases was smaller in the heterotypic system than in the homotypic system, suggesting less conformational entropy gain during phase separation. Collectively, our results showed that Hero11-WT-α regulates the thermodynamic behavior of TDP-43-LCD with both energetic (Figure 5D) and entropic (Figure 5E) contributions.

To quantitatively evaluate the effect of Hero11-WT-α on the kinetics of TDP-43-LCD, we calculated the mean square displacement (MSD) of TDP-43-LCD ($\text{MSD}_{\text{TDP-43-LCD}}$) in dense and dilute phases as a function of $n_{\text{Hero11-WT-}\alpha}$ to estimate the diffusion of TDP-43-LCD (Figure 5F). Importantly, the effects of Hero11-WT-α on TDP-43-LCD MSD in the two phases were the opposite. In the dilute phase, owing to the interactions with Hero11-WT-α, TDP-43-LCD is slowed down compared with its behavior in the homotypic condensate (Figure 5F, lower panel). In contrast, in the dense phase, TDP-43-LCD diffused faster when more Hero11-WT-α was involved (Figure 5F, upper panel). We also decomposed the three-dimensional MSD into three one-dimensional MSDs (MSD_{1D}) in the x-, y-, and z-dimensions. Due to the limited size of the condensate, the z-axis dense phase MSD_{1D} was smaller than the other two dimensions (Figure S15). However, all three dimensions showed a consistent feature that the dense phase MSD_{1D} was increased as $n_{\text{Hero11-WT-}\alpha}$ increased (Figure S15). Importantly, the higher diffusion speed of TDP-43-LCD induced by Hero11-WT-α implies that the condensation of
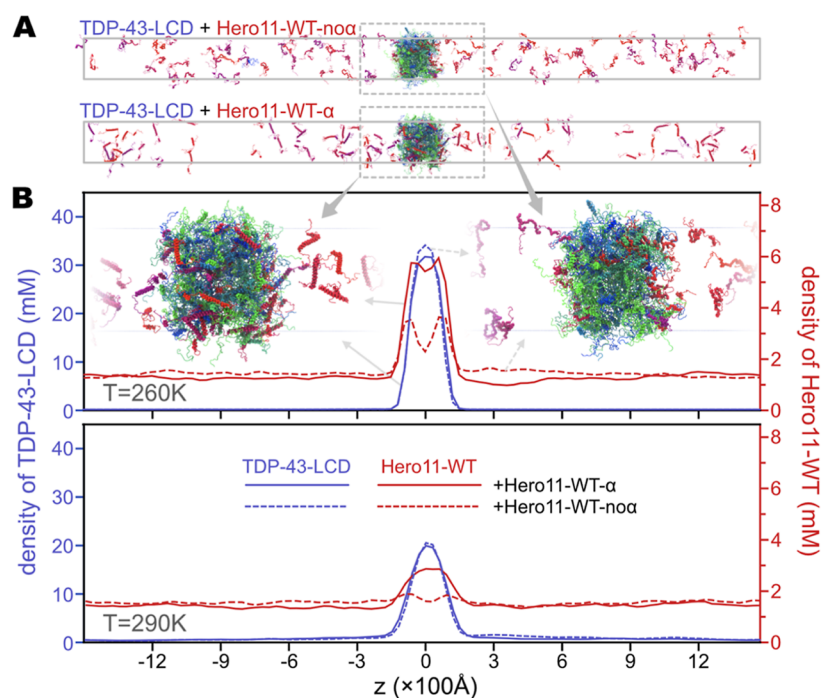
Figure 6. Effect of α-helical secondary structures on the distribution of Hero11-WT in the condensate. (A) Final structures of two simulations at T = 260 K, 100 TDP-43-LCD +90 Hero11-WT-noα (upper) or 90 Hero11-WT-α (lower). (B) Densities of TDP-43-LCD (blue) and Hero11-WT (red) along the z-axis in simulations at T = 260 K (upper) and 290 K (lower), respectively. Solid and dashed lines represent results analyzed from simulations with Hero11-WT-α and Hero11-WT-noα, respectively.

TDP-43-LCD and Hero11-WT-α is more fluidic than the homotypic TDP-43-LCD assembly. This result supports the antiaggregation function of Hero11[30] with a molecule-level kinetic explanation.

Effect of the Secondary Structure of Hero11 on Its Regulatory Function

As already shown in this study and previous studies, α-helical secondary structures affect the homotypic phase behaviors of proteins.[24,52] We then ask how heterotypic condensation is affected by secondary structural properties. To answer this question, we performed MD simulations for the TDP-43-LCD−Hero11-WT-noα mixture. We find that the "noα" systems' phase diagrams are similar to those with α-helices (Figure S11B). Notably, Hero11-WT-noα also decreased T_c of TDP-43-LCD, while Hero11-KRless-noα and Hero11-charge-less-noα did not have this effect (Figure S11B). We then calculated the partition coefficient (defined as the number of Hero11 in the dense phase divided by the number of Hero11 in the dilute phase) of Hero11-WT-α and Hero11-WT-noα to determine the effect of the α-helical secondary structure on the distribution of Hero11 proteins. Interestingly, we found that there was more Hero11-WT-α populated in the dense phase than Hero11-WT-noα under the same conditions (Figure S16). This phenomenon is more evident at relatively lower temperatures or lower Hero11 concentrations. Contact map analysis showed that in the mixture of TDP-43-LCD and Hero11-WT-α, a large fraction of intermolecular contacts between TDP-43-LCD and Hero11-WT-α was contributed by the α-helices of both proteins (Figure S17). In the TDP-43-LCD−Hero11-WT-noα system, these secondary structure-facilitated interactions were missing (Figure S18). Therefore, intermolecular interactions energetically favor Hero11-WT-α in the condensate of TDP-43-LCD.

We further analyzed the distribution of the Hero11 molecules in the simulation boxes. Representative structures simulated at 260 K for 100 TDP-43-LCD in a mixture with 90 Hero11-WT-noα and 90 Hero11-WT-α show that there seems to be more Hero11-WT-α than Hero11-WT-noα inside the concentrated phase (Figure 6A). To quantitatively assess the distribution of the Hero11-WT chains in the simulation boxes, we calculated the time-averaged local density of Hero11-WT-α and Hero11-WT-noα along the z-axis (Figure 6B). At both 260 and 290 K, the local concentration of Hero11-WT-α chains located in the dense phase was higher than that of Hero11-WT-noα. Additionally, consistent with the structures shown in Figure 6A, we found that Hero11-WT-noα had a high population at the boundary of the TDP-43-LCD condensate, whereas Hero11-WT-α entered deeper into the dense phase (Figure 6B). These results can be explained by the difference in the secondary structure-dependent intermolecular contacts. Compared with Hero11-WT-noα, the α-helical regions in Hero11-WT-α contribute more interactions with the α-helical region of TDP-43-LCD (Figures S17 and S18). These contacts make Hero11-WT-α easier than Hero11-WT-noα to permeate into the interior of TDP-43-LCD condensation. We found that the difference between Hero11-WT-α and Hero11-WT-noα weakened at relatively higher temperatures (290 K, Figure 6B, lower panel). We then calculated the Hero11 concentration-dependent thermodynamic and kinetic quantities of TDP-43-LCD (interchain contacts, particle density, R_g, and MSD) in the presence of Hero11-WT-noα at 290 K and found no substantial difference compared to the results obtained with Hero11-WT-α (Figure S19).
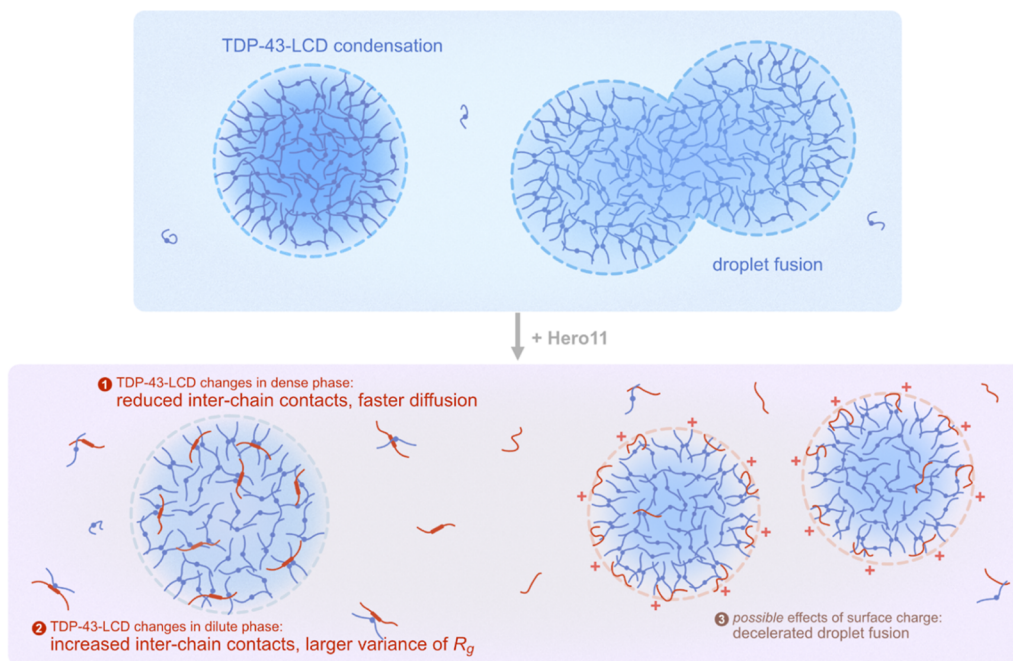
Figure 7. Molecular mechanism for the antiaggregation function of Hero11. Top: TDP-43-LCD can form homotypic condensation via intermolecular interactions formed by IDR (lines) and facilitated by α-helical structures (dots).[52] Bottom: Hero11 regulates the phase behavior of TDP-43-LCD using three different mechanisms: (1) Hero11 joins the condensate of TDP-43-LCD as a ligand and decreases the intermolecular contacts of TDP-43-LCD. Consequently, the diffusion of TDP-43-LCD is promoted. (2) Hero11 raises the probability for TDP-43-LCD to be in the dilute phase and induces larger variance in R_g of TDP-43-LCD outside of the condensate. (3) In the case of less helical structures, Hero11 tends to stay at the surface of the droplets, which may have some surface effects such as preventing droplet fusion.[66]

DISCUSSION

In the current work, we used Hero11 as an example to study the mechanisms of biomolecular condensation regulation. We found that Hero11 binds to the condensate formed by TDP-43-LCD (Figure 4A−D), which conforms to the "scaffold and ligand" paradigm.[34,63,64] The stickers-and-spacers model has been used to show that the formation or dissolution of condensate can be regulated by the valency and binding specificity of the ligand.[34] However, charges and repulsive interactions were not considered in the model.[34] Our results demonstrate that Hero11 destabilizes the TDP-43-LCD condensate by introducing strong ligand−ligand electrostatic repulsion in addition to the normally attractive ligand−scaffold interactions. By analyzing the contact number of TDP-43-LCD, we showed that Hero11 binds to TDP-43-LCD in both dense and dilute phases (Figures 5D; S17, S18, and S19B), but the effects of Hero11 are different in the two phases, as explained below.

Specifically, we considered three possible mechanisms by which Hero11 regulates TDP-43-LCD condensation (Figure 7). First, in the concentrated phase, Hero11 forms contact with the scaffolding TDP-43-LCD mainly via short-range attractive interactions (modeled by the HPS potential). Simultaneously, the positively charged residues in Hero11 exert long-range repulsive forces on Hero11. As a result, the Hero11 chains tend to repel each other and loosen the whole condensate, including the scaffold on which they bind. Accordingly, we observed lower particle densities (Figure 5B), smaller numbers of contacts (Figure 5D), and faster displacement of TDP-43-LCD (Figure 5F) in the dense phase when Hero11 was present compared to those in the homotypic condensate. Second, in the dilute phase, Hero11 increased the total contact number of TDP-43-LCD (Figure 5D, lower panel; Figures S17−S19) and

energetically raised the probability for TDP-43-LCD to stay in the dilute phase. Hero11 also increased both the mean value and standard deviation of R_g of the TDP-43-LCD chains in the dilute phase (Figure 5E). As discussed earlier, Hero11 results in a smaller difference in the conformational fluctuation of TDP-43-LCD in the two phases and thus may contribute to reducing the entropy gain during the phase separation. Therefore, these results explain the ability of Hero11 to enhance the probability of TDP-43-LCD locating in the dilute phase.

We also found that as the structural helicity decreases, Hero11 distributes more on the surface layer of the coassembly with TDP-43-LCD. Similar results were reported in a previous MD simulation study of the heterotypic condensate formed by LAF-1 and RNA.[50] Notably, amphiphilic proteins have been found to form a surface layer that can regulate the structure and size of condensates.[67] Another combined computational and experimental study revealed that surface charges stabilize and protect biomolecular condensates against fusion.[66] These observations, in combination with our findings on the surface distribution of Hero11-WT-noα (Figure 6), suggest a third possible mechanism for the regulatory function of Hero11, in which Hero11 may contribute to sustaining the condensate of TDP-43-LCD and prevent droplet coalescence (Figure 7). Note that this antifusion effect of the surface charge distribution is still a speculative model. Our current slab simulations can only provide the concentration profiles of Hero11-WT proteins in the condensate. A direct validation of this mechanism may need larger-scale simulations of the whole droplets. We also propose that these possible mechanisms can be verified and distinguished by comparing detectable quantities, such as the radius of gyration and diffusion coefficients, using in vitro and in vivo experiments.[68−70]

Our results revealed that the net charge is crucial in determining the thermodynamic properties and regulatory functions of Hero11. Similar to Hero11, a large absolute value of net charge, either positive or negative, is a common feature of other Hero proteins.[30] Therefore, we speculate that the driving force for the antiaggregation function of the other Hero proteins may also be strong electrostatic repulsion. Interestingly, RNA, which is also highly charged, has been shown to induce concentration-dependent reentrant phase behaviors in many RNA-binding proteins.[29,50,71,72] In particular, a recent computational study showed that a high concentration of RNA can slow down the aging of condensates formed by proteins composed of β-sheet secondary structures.[72] Although the computational models and target systems were different, the results of the RNA deceleration effect were very similar to the antiaggregation mechanism of Hero11 proposed here. One difference is that Tejedor et al. introduced a time-dependent model to simulate the maturation procedure of protein aggregation regulated by RNA,[72] whereas in our work, we focused on the preaged stage of TDP-43 condensation. Our conclusions can be considered as complementary to Tejedor et al. in providing a whole picture of the passive regulation of biomolecular condensation in different maturation stages. In summary, our results suggest that adding repulsive interactions to the framework of sticker−spacer or scaffold−ligand models can extend their applicability to describe the formation, regulation, and elimination of biomolecular condensation.

We note that our method can be extended to study similar systems in the future. For example, one question that remains unanswered is the specificity of the regulatory functions of the Hero proteins as different Hero proteins have distinct clients. Hero proteins have different sequences and secondary structural features whose relationships with their phase behaviors have not been thoroughly studied. We expect that our models and simulation results of Hero11-WT-α and TDP-43-LCD in the current study have provided mechanisms that can be qualitatively applied to explain other similar systems. For example, we studied two extreme cases of secondary structures and found differences in their distribution in condensates, which may help to understand the properties of some other Hero proteins that have similar structural features. We also note that the accuracy of our CG model can be improved by parameter recalibration against experimental results if they are available, for example, critical points of the scaffold proteins (such as TDP-43) and circular dichroism spectra or SAXS profiles of ligand proteins (such as Hero proteins). Alternatively, all-atom simulations using different force fields with improved parameters[73−76] and advanced sampling methods[57,58,77] can provide more direct information such as atomic-level residue−residue interactions and motions of ions and water molecules.

■ COMPUTATIONAL METHODS

All-Atom Simulations of WT and KRLess Hero11

We used AlphaFold2[53] to predict the structure of Hero11 and obtained two α-helical regions at residues 17−25 and 38−72. To test the stability of these secondary structures, we performed all-atom MD simulations of the Hero11-WT and the Hero11-KRless mutant with CHARMM36m force field.[41] More detailed information on the atomistic simulations can be found in the Supporting Information Methods section.

CG Modeling of TDP-43-LCD and Hero11

In the CG simulations, we studied full-length Hero11 (99 amino acid residues) and the C-terminal fragment of TDP-43 (TDP-43-LCD, index 261−414). To understand the effect of the charged residues, we also simulate several mutants of Hero11, including the KRless (all arginine and lysine mutated to glycine), the charge-less (all charged residues mutated to glycine), and the "scrambles" (sequence randomly permuted). The sequences of all simulated proteins are shown in Table S1.

The HPS CG model was employed for IDRs.[43] The HPS model represents each amino acid residue as a single particle. The potential energy function is given by[43]

$$V_{\mathrm{HPS}}(\Gamma) = \sum_{\mathrm{bonds}} E_{\mathrm{b}}(b_i) + \sum_{(i,j)\in \mathrm{non\text{-}bonded\ pairs}} E_{\mathrm{HPS}}(r_{ij}) + \sum_{(i,j)\in \mathrm{charged\ pairs}} E_{\mathrm{ele}}(r_{ij}) \tag{1}$$

where Γ is the conformation of an IDP and $E_{\mathrm{b}}(b_i)$ is the harmonic potential for every two neighboring CG particles with a bond length b_i

$$E_{\mathrm{b}}(b_i) = k_{\mathrm{b}}(b_i - b_{i,0})^2 \tag{2}$$

where the reference value of the bond length is $b_{i,0}$ = 3.8 Å and the force constant is k_{b} = 2.39 kcal Å^{-2}. $E_{\mathrm{HPS}}(r_{ij})$ is the interaction between two nonbonded particles and is defined as follows[43,78]

$$E_{\mathrm{HPS}}(r_{ij}) = \begin{cases} E_{\mathrm{LJ}}(r_{ij}) + (1 - \lambda_{ij})\epsilon, & r_{ij} \leq 2^{1/6}\sigma_{ij} \\ \lambda_{ij}E_{\mathrm{LJ}}(r_{ij}), & r_{ij} > 2^{1/6}\sigma_{ij} \end{cases} \tag{3}$$

where λ_{ij} is the hydropathy and $E_{\mathrm{LJ}}(r_{ij})$ is the Lennard−Jones potential

$$E_{\mathrm{LJ}}(r_{ij}) = 4\epsilon\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{4}$$

In eqs 3 and 4, ϵ = 0.2 kcal/mol. We used σ_{ij} values in the original HPS model and the λ_{ij} values in the recently proposed HPS-Urry optimization.[47] Both σ_{ij} and λ_{ij} use an arithmetic combinational rule for an interacting pair of particles (i,j).

For the electrostatic interaction $E_{\mathrm{ele}}(r_{ij})$ in eq 1, we use the Debye−Hückel term

$$E_{\mathrm{ele}}(r_{ij}) = \frac{q_i q_j \ \mathrm{e}^{-r_{ij}/\lambda_{\mathrm{D}}}}{4\pi\varepsilon_0\varepsilon_{\mathrm{r}}r_{ij}} \tag{5}$$

where r_{ij} is the distance between the nonbonded charged particles, λ_{D} is the Debye screening length, and ε_0 is the dielectric permittivity of vacuum. ε_{r} is the relative permittivity of the solution and is a function of the solution temperature T and salt molarity C: $\varepsilon_{\mathrm{r}} = e(T)a(C)$, where $e(T) = 249.4 - 0.788T + 7.20 \times 10^{-4}T^2$,[79] and $a(C) = 1 - 0.2551C + 5.151 \times 10^{-2}C^2 - 6.889 \times 10^{-3}C^3$.[80] The Debye length is given by $\lambda_{\mathrm{D}} = \sqrt{\frac{k_{\mathrm{B}}T\varepsilon_0\varepsilon_{\mathrm{r}}}{2N_{\mathrm{A}}\mathrm{e}_{\mathrm{c}}^2 I}}$, where N_{A} is the Avogadro's number, e_{c} is the elementary charge, and I is the ionic strength of the solution.

As described above, AlphaFold2 predicted the α-helical secondary structural pieces in the conformation of Hero11. However, our atomistic simulations did not provide convergent results regarding the stability of these secondary structures (Figure S2). Therefore, in the CG simulations, we modeled two extreme cases of Hero11, namely, one with a well-folded α-helical model (Hero11-WT-α) and the other with pure IDR (Hero11-WT-noα). We used the AICG2+ model[60] to maintain folded parts of proteins. The AICG2+ potential energy function has the following form[60]

$$V_{\text{AICG2+}}(\Gamma) = V_{\text{local}} + \sum_{(i,j) \in \text{native contacts}} E_{\text{G}\bar{\text{o}}}(r_{ij})$$
$$+ \sum_{(i,j) \in \text{nonnative contacts}} E_{\text{exv}}(r_{ij}) \tag{6}$$

where V_{local} includes all bonded terms, $E_{\text{G}\bar{\text{o}}}(r_{ij})$ is the G$\bar{\text{o}}$-type potential biasing toward the native structure, and $E_{\text{exv}}(r_{ij})$ is the excluded volume interaction. More details on the potential functions can be found in the Supporting Information Methods section. Specifically, residues 17−25 and 38−72 in Hero11 and residues 60−74 in TDP-43-LCD (320−334 in full-length TDP-43) were modeled with AICG2+. We used the AlphaFold2-predicted structures as reference structures for structure-based potentials. We also turned off the HPS potentials for nonlocal particle pairs that already had AICG2+ interactions.

CG MD

All CG simulations were performed using the MD package GENESIS v1.7.1.[81−83] Structure and topology files were prepared with the GENESIS-CG-tool.[83] The time-integration step size of the CG simulations was 10 fs. Nonlocal interactions E_{HPS} and E_{ele} have cutoffs of 20 and 35 Å, respectively. The potential function forms and cutoffs are the same as those used in the original HPS model[43] and the HPS-Urry optimization.[47]

For both TDP-43-LCD and Hero11, we first simulated single chains in 2×10^7 steps at different temperatures ranging from 110 to 350 K. For each protein, we simulated two conformations, one with an α-helical region modeled by AICG2+ and the other with pure IDR. The final structures of the single-chain simulations were then duplicated to construct multiple-chain systems, in which we set a minimum distance of 10 Å between neighboring chains. We then simulated homotypic systems, including 100 chains of TDP-43-LCD and Hero11. For the heterotypic systems consisting of TDP-43-LCD and Hero11, we fixed the number of TDP-43-LCD to 100 and added different numbers of Hero11 to the system. From all multiple-chain structures, we perform "shrinking" simulations to gradually squash the simulation boxes to 18 nm × 18 nm × z (z is system dependent) in 10^6 steps. We then extended z to 200 nm for the homotypic Hero11 systems and 300 nm for all others. These systems were equilibrated in the NVT ensemble for 2×10^7 steps at different temperatures. Production runs were conducted using Langevin dynamics with a friction coefficient of 0.01 ps^{-1}.

Data Analysis

To determine the density of protein particles along the longest dimension (z), we divided the z axis into $n_{\text{bin}} = 100$ bins. The z coordinates of the particles were then used to calculate the "local particle density" in each bin as $\rho_{\text{particle}}(i) = N_{\text{particle}}(i)/V_{\text{bin}}(i)$, where $N_{\text{particle}}(i)$ and $V_{\text{bin}}(i)$ are the number of particles and volume of the i-th bin, respectively. The protein density in the i-th bin was then computed as $\rho_{\text{protein}}(i) = \rho_{\text{particle}}(i)/n_{\text{aa,protein}}$, where $n_{\text{aa,protein}}$ is the number of amino acid residues in the considered protein ($n_{\text{aa,TDP-43}} = 154$ and $n_{\text{aa,Hero11}} = 99$). The slab, which corresponded to the largest cluster of proteins, was determined in two steps. We first determined the largest $\rho_{\text{protein}}(j)$ and its bin index j. Then, from this local density peak, we scanned in the $+z$ and $-z$ directions and determined the boundaries of the slab (b_{+z} and b_{-z}) by finding the first bins satisfying $\rho_{\text{protein}}(b_{\pm z}) < \rho_{\text{threshold}}$. Here, $\rho_{\text{threshold}} = 3.4$ mM, which corresponds to roughly two proteins in each bin. Practically, we shifted the center of the slab (calculated as $(b_{-z} + b_{+z})/2$) to the center of the simulation box. The dense phase is defined as the region in (b_{-z}, b_{+z}), and the dilute phase is defined as the regions $[-n_{\text{bin}}/2, b_{-z} - b_{\text{buff}}]$ and $[b_{+z} + b_{\text{buff}}, n_{\text{bin}}/2]$, where $b_{\text{buff}} = 2$ is the buffering distance from the boundaries of the dense phase. The densities of the dense (ρ_{h}, the "high" density) and dilute (ρ_{l}, the "low" density) phases were then defined as the average densities in the dense and dilute regions, respectively. The current method works only for structures containing single slabs. Therefore, we only calculated ρ_{h} and ρ_{l} for relatively low

temperatures, where both the dense and dilute phases could be well defined. The critical point (ρ_{c}, T_{c}) was computed by fitting the following equations[84]

$$\rho_{\text{h}} - \rho_{\text{l}} = \Delta\rho_0 \left(1 - \frac{T}{T_{\text{c}}}\right)^{\beta} \tag{7}$$

$$\frac{\rho_{\text{h}} + \rho_{\text{l}}}{2} = \rho_{\text{c}} + A(T_{\text{c}} - T) \tag{8}$$

where ρ_{c} and T_{c} are the critical density and temperature, respectively; $\Delta\rho_0$ and A are system-specific fitting parameters; and $\beta = 0.325$.[85]

To evaluate the interaction between two protein chains, we calculated the number of interchain contacts formed by each chain as follows

$$c_m = \sum_{i \in \text{chain } m} \sum_{j \notin \text{chain } m} \frac{1}{1 + e^{r_{ij} - \sigma}} \tag{9}$$

where c_m is the number of contacts of the m-th chain, i is the index of amino acid residues in the m-th chain, j is the index of residues from all the other chains, r_{ij} is the distance between the i-th and j-th residues, and $\sigma = 15$ Å. For TDP-43-LCD, we calculated the average number of contacts formed by each chain as $\overline{c}_{\text{TDP-43}} = \sum_{m \in \text{TDP-43}} c_m/n_{\text{TDP-43}}$, where $n_{\text{TDP-43}} = 100$.

To quantitatively describe the diffusion of proteins, we calculated the MSD of center of mass (COM) of each chain as a function of time interval (Δt): $\text{MSD} = \langle (\vec{r}(t + \Delta t) - \vec{r}(t))^2 \rangle$. For each chain, we first extracted its COM coordinates and divided CG MD trajectories into short pieces, in which the chain of interest stayed in the same phase. Then, for each phase, we gathered all short trajectory pieces and calculated MSDs with different Δt. We also calculated 1D MSDs (MSD_{1D}) on the x, y, and z dimensions. For example, for the x dimension, we define $\text{MSD}_{\text{1D}} = \langle (x(t + \Delta t) - x(t))^2 \rangle$.

■ ASSOCIATED CONTENT

Data Availability Statement

All MD simulation files and data to produce the results are available from https://github.com/RikenSugitaLab/LLPS-Regulated-by-Highly-Charged-Proteins.

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.2c00646.

     Details for all-atom and CG simulations, list of protein sequences simulated, and additional details of all-atom and CG simulation results and data analysis (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Yuji Sugita − *Computational Biophysics Research Team, RIKEN Center for Computational Science, Kobe, Hyogo 650-0047, Japan; Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Wako, Saitama 351-0198, Japan; Laboratory for Biomolecular Function Simulation, RIKEN Center for Biosystems Dynamics Research, Kobe, Hyogo 650-0047, Japan;* ⊙ orcid.org/0000-0001-9738-9216; Email: sugita@riken.jp

Authors

Cheng Tan − *Computational Biophysics Research Team, RIKEN Center for Computational Science, Kobe, Hyogo 650-0047, Japan;* ⊙ orcid.org/0000-0002-8957-4267

Ai Niitsu − *Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Wako, Saitama 351-0198, Japan;* ◎ orcid.org/0000-0002-7823-9041

Complete contact information is available at:
https://pubs.acs.org/10.1021/jacsau.2c00646

Author Contributions

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Dignon, G. L.; Best, R. B.; Mittal, J. Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. *Annu. Rev. Phys. Chem.* **2020**, *71*, 53−75.

(2) Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019**, *176*, 419−434.

(3) Heidarsson, P. O.; Mercadante, D.; Sottini, A.; Nettels, D.; Borgia, M. B.; Borgia, A.; et al. Release of linker histone from the nucleosome driven by polyelectrolyte competition with a disordered protein. *Nat. Chem.* **2022**, *14*, 224−231.

(4) Wagh, K.; Garcia, D. A.; Upadhyaya, A. Phase separation in transcription factor dynamics and chromatin organization. *Curr. Opin. Struct. Biol.* **2021**, *71*, 148−155.

(5) Cho, W.-K.; Spille, J.-H.; Hecht, M.; Lee, C.; Li, C.; Grube, V.; et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **2018**, *361*, 412−415.

(6) Su, X.; Ditlev, J. A.; Hui, E.; Xing, W.; Banjade, S.; Okrut, J.; et al. Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **2016**, *352*, 595−599.

(7) Riback, J. A.; Katanski, C. D.; Kear-Scott, J. L.; Pilipenko, E. V.; Rojek, A. E.; Sosnick, T. R.; et al. Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* **2017**, *168*, 1028−1040.

(8) Ramaswami, M.; Taylor, J. P.; Parker, R. Altered Ribostasis: RNA-Protein Granules in Degenerative Disorders. *Cell* **2013**, *154*, 727−736.

(9) Portz, B.; Lee, B. L.; Shorter, J. FUS and TDP-43 Phases in Health and Disease. *Trends Biochem. Sci.* **2021**, *46*, 550−563.

(10) Riemenschneider, H.; Guo, Q.; Bader, J.; Frottin, F.; Farny, D.; Kleinberger, G.; et al. Gel-like inclusions of C-terminal fragments of TDP-43 sequester stalled proteasomes in neurons. *EMBO Rep.* **2022**, *23*, No. e53890.

(11) Arseni, D.; Hasegawa, M.; Murzin, A. G.; Kametani, F.; Arai, M.; Yoshida, M.; et al. Structure of pathological TDP-43 filaments from ALS with FTLD. *Nature* **2022**, *601*, 139−143.

(12) Nelson, P. T.; Dickson, D. W.; Trojanowski, J. Q.; Jack, C. R.; Boyle, P. A.; Arfanakis, K.; et al. Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. *Brain* **2019**, *142*, 1503−1527.

(13) Neumann, M.; Sampathu, D. M.; Kwong, L. K.; Truax, A. C.; Micsenyi, M. C.; Chou, T. T.; et al. Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science* **2006**, *314*, 130−133.

(14) Arai, T.; Hasegawa, M.; Akiyama, H.; Ikeda, K.; Nonaka, T.; Mori, H.; et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **2006**, *351*, 602−611.

(15) Li, P.; Banjade, S.; Cheng, H.-C.; Kim, S.; Chen, B.; Guo, L.; et al. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **2012**, *483*, 336−340.

(16) Abyzov, A.; Blackledge, M.; Zweckstetter, M. Conformational Dynamics of Intrinsically Disordered Proteins Regulate Biomolecular Condensate Chemistry. *Chem. Rev.* **2022**, *122*, 6719−6748.

(17) Borcherds, W.; Bremer, A.; Borgia, M. B.; Mittag, T. How do intrinsically disordered protein regions encode a driving force for liquid-liquid phase separation? *Curr. Opin. Struct. Biol.* **2021**, *67*, 41−50.

(18) Weng, J.; Wang, W. Dynamic multivalent interactions of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2020**, *62*, 9−13.

(19) Wang, W.; Wang, D. Extreme Fuzziness: Direct Interactions between Two IDPs. *Biomolecules* **2019**, *9*, 81.

(20) McGurk, L.; Gomes, E.; Guo, L.; Mojsilovic-Petrovic, J.; Tran, V.; Kalb, R. G.; et al. Poly(ADP-Ribose) Prevents Pathological Phase Separation of TDP-43 by Promoting Liquid Demixing and Stress Granule Localization. *Mol. Cell* **2018**, *71*, 703−717.

(21) Mann, J. R.; Gleixner, A. M.; Mauna, J. C.; Gomes, E.; DeChellis-Marks, M. R.; Needham, P. G.; et al. RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. *Neuron* **2019**, *102*, 321−338.

(22) Dogra, P.; Joshi, A.; Majumdar, A.; Mukhopadhyay, S. Intermolecular Charge-Transfer Modulates Liquid-Liquid Phase Separation and Liquid-to-Solid Maturation of an Intrinsically Disordered pH-Responsive Domain. *J. Am. Chem. Soc.* **2019**, *141*, 20380−20389.

(23) Wang, A.; Conicella, A. E.; Schmidt, H. B.; Martin, E. W.; Rhoads, S. N.; Reeb, A. N.; et al. A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J.* **2018**, *37*, No. e97452.

(24) Conicella, A. E.; Zerze, G. H.; Mittal, J.; Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **2016**, *24*, 1537−1549.

(25) Yu, H.; Lu, S.; Gasior, K.; Singh, D.; Vazquez-Sanchez, S.; Tapia, O.; et al. HSP70 chaperones RNA-free TDP-43 into anisotropic intranuclear liquid spherical shells. *Science* **2021**, *371*, No. eabb4309.

(26) Cohen, T. J.; Hwang, A. W.; Restrepo, C. R.; Yuan, C.-X.; Trojanowski, J. Q.; Lee, V. M. Y. An acetylation switch controls TDP-43 function and aggregation propensity. *Nat. Commun.* **2015**, *6*, 5845.

(27) Gruijs da Silva, L. A.; Simonetti, F.; Hutten, S.; Riemenschneider, H.; Sternburg, E. L.; Pietrek, L. M.; et al. Disease-linked TDP-43 hyperphosphorylation suppresses TDP-43 condensation and aggregation. *EMBO J.* **2022**, *41*, No. e108443.

(28) Langdon, E. M.; Qiu, Y.; Ghanbari Niaki, A.; McLaughlin, G. A.; Weidmann, C. A.; Gerbich, T. M.; et al. mRNA structure determines specificity of a polyQ-driven phase separation. *Science* **2018**, *360*, 922−927.

(29) Maharana, S.; Wang, J.; Papadopoulos, D. K.; Richter, D.; Pozniakovsky, A.; Poser, I.; et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **2018**, *360*, 918−921.

(30) Tsuboyama, K.; Osaki, T.; Matsuura-Suzuki, E.; Kozuka-Hata, H.; Okada, Y.; Oyama, M.; et al. A widespread family of heat-resistant obscure (Hero) proteins protect against protein instability and aggregation. *PLoS Biol.* **2020**, *18*, No. e3000632.

(31) Choi, J.-M.; Holehouse, A. S.; Pappu, R. V. Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annu. Rev. Biophys.* **2020**, *49*, 107−133.

(32) Brangwynne, C. P.; Tompa, P.; Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **2015**, *11*, 899−904.

(33) Martin, E. W.; Holehouse, A. S.; Peran, I.; Farag, M.; Incicco, J. J.; Bremer, A.; et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **2020**, *367*, 694−699.

(34) Ruff, K. M.; Dar, F.; Pappu, R. V. Ligand effects on phase separation of multivalent macromolecules. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2017184118.

(35) Lin, Y.-H.; Chan, H. S. Phase Separation and Single-Chain Compactness of Charged Disordered Proteins Are Strongly Correlated. *Biophys. J.* **2017**, *112*, 2043−2046.

(36) Lin, Y.-H.; Forman-Kay, J. D.; Chan, H. S. Sequence-Specific Polyampholyte Phase Separation in Membraneless Organelles. *Phys. Rev. Lett.* **2016**, *117*, 178101.

(37) Dinic, J.; Marciel, A. B.; Tirrell, M. V. Polyampholyte physics: Liquid-liquid phase separation and biological condensates. *Curr. Opin. Colloid Interface Sci.* **2021**, *54*, 101457.

(38) Zeng, X.; Ruff, K. M.; Pappu, R. V. Competing interactions give rise to two-state behavior and switch-like transitions in charge-rich intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, No. e2200559119.

(39) Borgia, A.; Borgia, M. B.; Bugge, K.; Kissling, V. M.; Heidarsson, P. O.; Fernandes, C. B.; et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **2018**, *555*, 61−66.

(40) Dignon, G. L.; Zheng, W.; Mittal, J. Simulation methods for liquid-liquid phase separation of disordered proteins. *Curr. Opin. Chem. Eng.* **2019**, *23*, 92−98.

(41) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71−73.

(42) Zerze, G. H.; Zheng, W.; Best, R. B.; Mittal, J. Evolution of All-Atom Protein Force Fields to Improve Local and Global Properties. *J. Phys. Chem. Lett.* **2019**, *10*, 2227−2234.

(43) Dignon, G. L.; Zheng, W.; Kim, Y. C.; Best, R. B.; Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.* **2018**, *14*, No. e1005941.

(44) Joseph, J. A.; Reinhardt, A.; Aguirre, A.; Chew, P. Y.; Russell, K. O.; Espinosa, J. R.; et al. Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nat. Comput. Sci.* **2021**, *1*, 732−743.

(45) Garaizar, A.; Espinosa, J. R. Salt dependent phase behavior of intrinsically disordered proteins from a coarse-grained model with explicit water and ions. *J. Chem. Phys.* **2021**, *155*, 125103.

(46) Latham, A. P.; Zhang, B. Maximum Entropy Optimized Force Field for Intrinsically Disordered Proteins. *J. Chem. Theory Comput.* **2020**, *16*, 773−781.

(47) Regy, R. M.; Thompson, J.; Kim, Y. C.; Mittal, J. Improved coarse-grained model for studying sequence dependent phase separation of disordered proteins. *Protein Sci.* **2021**, *30*, 1371−1379.

(48) 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. U.S.A.* **2021**, *118*, No. e2111696118.

(49) Dignon, G. L.; Zheng, W.; Kim, Y. C.; Mittal, J. Temperature-Controlled Liquid-Liquid Phase Separation of Disordered Proteins. *ACS Cent. Sci.* **2019**, *5*, 821−830.

(50) Regy, R. M.; Dignon, G. L.; Zheng, W.; Kim, Y. C.; Mittal, J. Sequence dependent phase separation of protein-polynucleotide mixtures elucidated using molecular simulations. *Nucleic Acids Res.* **2020**, *48*, 12593−12603.

(51) Perdikari, T. M.; Jovic, N.; Dignon, G. L.; Kim, Y. C.; Fawzi, N. L.; Mittal, J. A predictive coarse-grained model for position-specific effects of post-translational modifications. *Biophys. J.* **2021**, *120*, 1187−1197.

(52) Conicella, A. E.; Dignon, G. L.; Zerze, G. H.; Schmidt, H. B.; D'Ordine, A. M.; Kim, Y. C.; et al. TDP-43 α-helical structure tunes liquid-liquid phase separation and function. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 5883−5894.

(53) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.

(54) Erdős, G.; Pajkos, M.; Dosztányi, Z. IUPred3: prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. *Nucleic Acids Res.* **2021**, *49*, W297−W303.

(55) Hu, G.; Katuwawala, A.; Wang, K.; Wu, Z.; Ghadermarzi, S.; Gao, J.; et al. flDPnn: Accurate intrinsic disorder prediction with putative propensities of disorder functions. *Nat. Commun.* **2021**, *12*, 4438.

(56) Peng, K.; Radivojac, P.; Vucetic, S.; Dunker, A. K.; Obradovic, Z. Length-dependent prediction of protein intrinsic disorder. *BMC Bioinf.* **2006**, *7*, 208.

(57) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(58) Kamiya, M.; Sugita, Y. Flexible selection of the solute region in replica exchange with solute tempering: Application to protein-folding simulations. *J. Chem. Phys.* **2018**, *149*, 072304.

(59) Stelzl, L. S.; Pietrek, L. M.; Holla, A.; Oroz, J.; Sikora, M.; Köfinger, J.; et al. Global Structure of the Intrinsically Disordered Protein Tau Emerges from Its Local Structure. *JACS Au* **2022**, *2*, 673−686.

(60) Li, W.; Wang, W.; Takada, S. Energy landscape views for interplays among folding, binding, and allostery of calmodulin domains. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 10550−10555.

(61) Chou, H.-Y.; Aksimentiev, A. Single-Protein Collapse Determines Phase Equilibria of a Biological Condensate. *J. Phys. Chem. Lett.* **2020**, *11*, 4923−4929.

(62) Hazra, M. K.; Levy, Y. Charge pattern affects the structure and dynamics of polyampholyte condensates. *Phys. Chem. Chem. Phys.* **2020**, *22*, 19368−19375.

(63) Banani, S. F.; Rice, A. M.; Peeples, W. B.; Lin, Y.; Jain, S.; Parker, R.; et al. Compositional Control of Phase-Separated Cellular Bodies. *Cell* **2016**, *166*, 651−663.

(64) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 285−298.

(65) Greig, J. A.; Nguyen, T. A.; Lee, M.; Holehouse, A. S.; Posey, A. E.; Pappu, R. V.; et al. Arginine-Enriched Mixed-Charge Domains Provide Cohesion for Nuclear Speckle Condensation. *Mol. Cell* **2020**, *77*, 1237−1250.

(66) Welsh, T. J.; Krainer, G.; Espinosa, J. R.; Joseph, J. A.; Sridhar, A.; Jahnel, M.; et al. Surface Electrostatics Govern the Emulsion Stability of Biomolecular Condensates. *Nano Lett.* **2022**, *22*, 612−621.

(67) Kelley, F. M.; Favetta, B.; Regy, R. M.; Mittal, J.; Schuster, B. S. Amphiphilic proteins coassemble into multiphasic condensates and act as biomolecular surfactants. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2109967118.

(68) Murthy, A. C.; Fawzi, N. L. The (un)structural biology of biomolecular liquid-liquid phase separation using NMR spectroscopy. *J. Biol. Chem.* **2020**, *295*, 2375−2384.

(69) Wang, Z.; Zhang, G.; Zhang, H. Protocol for analyzing protein liquid-liquid phase separation. *Biophys. Rep.* **2019**, *5*, 1−9.

(70) Sottini, A.; Borgia, A.; Borgia, M. B.; Bugge, K.; Nettels, D.; Chowdhury, A.; et al. Polyelectrolyte interactions enable rapid

association and dissociation in high-affinity disordered protein complexes. *Nat. Commun.* **2020**, *11*, 5736.

(71) Roden, C.; Gladfelter, A. S. RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 183−195.

(72) Tejedor, A. R.; Sanchez-Burgos, I.; Estevez-Espinosa, M.; Garaizar, A.; Collepardo-Guevara, R.; Ramirez, J.; et al. Protein structural transitions critically transform the network connectivity and viscoelasticity of RNA-binding protein condensates but RNA can prevent it. *Nat. Commun.* **2022**, *13*, 5717.

(73) Tang, W. S.; Fawzi, N. L.; Mittal, J. Refining All-Atom Protein Force Fields for Polar-Rich, Prion-like, Low-Complexity Intrinsically Disordered Proteins. *J. Phys. Chem. B* **2020**, *124*, 9505−9512.

(74) Nawrocki, G.; Wang, P.; Yu, I.; Sugita, Y.; Feig, M. Slow-Down in Diffusion in Crowded Protein Solutions Correlates with Transient Cluster Formation. *J. Phys. Chem. B* **2017**, *121*, 11072−11084.

(75) Matsubara, D.; Kasahara, K.; Dokainish, H. M.; Oshima, H.; Sugita, Y. Modified Protein-Water Interactions in CHARMM36m for Thermodynamics and Kinetics of Proteins in Dilute and Crowded Solutions. *Molecules* **2022**, *27*, 5726.

(76) Robustelli, P.; Piana, S.; Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E4758−E4766.

(77) Pietrek, L. M.; Stelzl, L. S.; Hummer, G. Hierarchical Ensembles of Intrinsically Disordered Proteins at Atomic Resolution in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2020**, *16*, 725−737.

(78) Ashbaugh, H. S.; Hatch, H. W. Natively Unfolded Protein Stability as a Coil-to-Globule Transition in Charge/Hydropathy Space. *J. Am. Chem. Soc.* **2008**, *130*, 9536−9542.

(79) Catenaccio, A.; Daruich, Y.; Magallanes, C. Temperature dependence of the permittivity of water. *Chem. Phys. Lett.* **2003**, *367*, 669−671.

(80) Stogryn, A. Equations for Calculating the Dielectric Constant of Saline Water (Correspondence). *IEEE Trans. Microwave Theory Tech.* **1971**, *19*, 733−736.

(81) Jung, J.; Mori, T.; Kobayashi, C.; Matsunaga, Y.; Yoda, T.; Feig, M.; et al. GENESIS: a hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2015**, *5*, 310−323.

(82) Kobayashi, C.; Jung, J.; Matsunaga, Y.; Mori, T.; Ando, T.; Tamura, K.; et al. GENESIS 1.1: A hybrid-parallel molecular dynamics simulator with enhanced sampling algorithms on multiple computational platforms. *J. Comput. Chem.* **2017**, *38*, 2193−2206.

(83) Tan, C.; Jung, J.; Kobayashi, C.; La Torre, D. U.; Takada, S.; Sugita, Y. Implementation of residue-level coarse-grained models in GENESIS for large-scale molecular dynamics simulations. *PLoS Comput. Biol.* **2022**, *18*, No. e1009578.

(84) Guggenheim, E. A. The Principle of Corresponding States. *J. Chem. Phys.* **1945**, *13*, 253−261.

(85) Rowlinson, J. S.; Widom, B. *Molecular Theory of Capillarity*; Dover Publication, 1982; pp 261−270.