Biophysical Journal
Article

Biophysical Society

A predictive coarse-grained model for position-specific effects of post-translational modifications

Theodora Myrto Perdikari,[1] Nina Jovic,[2] Gregory L. Dignon,[2,3] Young C. Kim,[4] Nicolas L. Fawzi,[5,6,*] and Jeetain Mittal[2,*]

[1]Center for Biomedical Engineering, Brown University, Providence, Rhode Island; [2]Department of Chemical and Biomolecular Engineering, Lehigh University, Bethlehem, Pennsylvania; [3]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York; [4]Center for Materials Physics and Technology, U.S. Naval Research Laboratory, Washington D.C; [5]Department of Molecular Pharmacology, Physiology, and Biotechnology; and [6]Robert J. and Nancy D. Carney Institute for Brain Science, Brown University, Providence, Rhode Island

ABSTRACT Biomolecules undergo liquid-liquid phase separation (LLPS), resulting in the formation of multicomponent protein-RNA membraneless organelles in cells. However, the physiological and pathological role of post-translational modifications (PTMs) on the biophysics of phase behavior is only beginning to be probed. To study the effect of PTMs on LLPS in silico, we extend our transferable coarse-grained model of intrinsically disordered proteins to include phosphorylated and acetylated amino acids. Using the parameters for modified amino acids available for fixed-charge atomistic force fields, we parameterize the size and atomistic hydropathy of the coarse-grained-modified amino acid beads and, hence, the interactions between the modified and natural amino acids. We then elucidate how the number and position of phosphorylated and acetylated residues alter the protein's single-chain compactness and its propensity to phase separate. We show that both the number and the position of phosphorylated threonines/serines or acetylated lysines can serve as a molecular on/off switch for phase separation in the well-studied disordered regions of Fused in Sarcoma (FUS) and DDX3X, respectively. We also compare modified residues to their commonly used PTM mimics for their impact on chain properties. Importantly, we show that the model can predict and capture experimentally measured differences in the phase behavior for position-specific modifications, showing that the position of modifications can dictate phase separation. In sum, this model will be useful for studying LLPS of post-translationally modified intrinsically disordered proteins and predicting how modifications control phase behavior with position-specific resolution.

SIGNIFICANCE Post-translational modifications are important regulators of liquid-liquid phase separation, (LLPS) which drives the formation of biomolecular condensates. Theoretical methods can be used to characterize the biophysical properties of intrinsically disordered proteins (IDPs). Our recent framework for molecular simulations using a Cα-centered coarse-grained model can predict the effect of various perturbations such as mutations (Dignon et al. *PLoS Comput. Biol.*, 2018) and temperature (Dignon et al., *ACS Cent. Sci.*, 2019) on LLPS. Here, we expand this framework to incorporate modified residues like phosphothreonine, phosphoserine, and acetyllysine. This model will prove useful for simulating the phase separation of post-translationally modified IDPs and predicting how position-specific modifications can control phase behavior across the large family of proteins known to be phosphorylated and acetylated.

INTRODUCTION

Many cellular processes depend on the formation of membraneless nuclear and cytoplasmic assemblies known as membraneless organelles or biomolecular condensates (1–3). Lacking a phospholipid membrane, these biomolecular condensates (4) can respond rapidly to environmental changes, forming cellular compartments concentrating specific proteins (5) and nucleic acids (6). The formation of many membraneless organelles appears to be driven by interactions between proteins containing intrinsically disordered regions (IDRs) (7). The detailed molecular interactions mediating these membraneless organelles may include cation-π (8), sp^2/π (9), hydrogen bonding, salt bridges, and hydrophobic interactions (10–12). Importantly, recent efforts have demonstrated that IDR liquid-liquid

Submitted June 12, 2020, and accepted for publication January 19, 2021.

*Correspondence: ; nicolas_fawzi@brown.edu or jeetain@lehigh.edu

Theodora Myrto Perdikari, Nina Jovic, and Gregory L. Dignon contributed equally to this work.

Editor: Rohit Pappu.

https://doi.org/10.1016/j.bpj.2021.01.034

phase separation (LLPS) and, hence, the formation of membranelles organelles can be regulated by a plethora of factors such as salt concentration, pH, RNA (13), ATP (14), temperature (15), and post-translational modifications (PTMs) (16,17).

PTMs enrich the repertoire of the 20 natural amino acids and have been shown to be an important potential means for modulating phase separation (17). Two of the most frequently occurring eukaryotic PTMs (18) are phosphorylation (19) and acetylation (20), which covalently attach a phosphoryl group or an acetyl group, respectively, to select amino acids (21). These modifications can reduce the binding affinity of RNA (22), impair enzyme activity (23), decrease aggregation propensity of fibril-forming segments (24), or trigger the self-assembly of RNA granules (25). Importantly, phosphorylation and acetylation both change the charge of the associated amino acid. Phosphorylated serine and threonine have a net charge between -1 and -2 at neutral pH (the actual value depends on the surrounding environment, though a pK_a of $\sim$5.9 to $\sim$6.0 for the -1 to -2 protonation state of phosphoserine/threonine (26) suggests the -2 charge state will be predominant at neutral pH) compared with no net charge for unmodified serine and threonine. Conversely, acetylation neutralizes the $+1$ charge that lysine has at physiological pH. Because IDRs generally contain many PTM sites that may be simultaneously modified (27), this can dramatically alter the net charge and charge patterning in ways that impact chain properties, phase separation, and aggregation. Hyperphosphorylation and hyperacetylation can lead to large changes in protein charge distribution and, hence, conformation and are associated with aggregation of amyloidogenic proteins in neurodegenerative diseases (20,28,29).

Although in vitro studies have demonstrated the effect of different PTMs on LLPS, studying the effect of each possible combination of PTMs is laborious and difficult. Furthermore, it is challenging to have site-specific control to reproduce modification patterns that are deposited on IDRs in vivo. Therefore, computational studies can be an effective approach to screen a large number of PTM patterns and predict their role on regulating the molecular properties and phase separation of proteins (30,31). All-atom explicit solvent simulations readily incorporate post-translationally modified amino acids (32–36) and have been used to probe the contacts leading to phase separation (37,38). However, the ability of atomistic force fields to accurately capture the properties of IDRs depends strongly on parameterization (39,40). Additionally, simulations of sufficiently long length and timescales to predict phase-separation behavior (e.g., saturation concentration) are beyond current computational capability for fully atomistic simulations with explicit water. In contrast, coarse-grained (CG) models (41–45), in particular those in which peptides are represented as chains of single beads and solvation effects are represented implicitly, have been used to efficiently investigate the phase behavior of large macromolecules (46,47). Most current CG modeling approaches capable of studying IDRs are currently parameterized only for the 20 natural amino acids (48), whereas others are suited specifically to probe phosphoregulation of folded domains (49,50). Therefore, there is a need for transferable approach to elucidate the effect of PTMs on the biophysical properties and phase behavior of IDRs associated with the assembly of biomolecular condensates.

Here, we expand our previous amino acid resolution CG hydropathy scale model (HPS) model (48) to incorporate PTMs compatible with the procedure used to obtain the parameters for the 20 natural unmodified amino acids. We use the HPS model as a starting point due to its transferable physics-based philosophy as compared with other commonly available CG paradigms. Many other CG models rely on explicit input from experiment data on a system of interest or a closely related system. Of course, further improvements in the underlying parameters may be needed (51), while preserving its transferable nature, to further enhance its applicability. We refer the readers to recent reviews for extensive discussions on CG models and their applications (41,43,44,47,52). To investigate how the position and the number of modified sites can alter the phase behavior of IDRs, we perform in silico studies of two IDRs: the N-terminal low-complexity (LC) domain of RNA-binding protein Fused in Sarcoma (FUS) LC (which is rich in serine, tyrosine, glutamine, and glycine residues and contains 12 known phosphorylation sites modified by DNA-dependent protein kinase (53)) and the first disordered region (IDR1) of the DEAD box RNA helicase 3, X-linked (DDX3X) with 10 lysine residues known to be acetylated (54). We test how the post-translational state of each molecule modulates IDR compactness and LLPS in a sequence-specific manner. Using CG molecular simulations, we examine the effect of increasing number of modified residues and test the hypothesis that position-specific patterning of PTMs can alter the collapse and phase behavior of IDR sequences. Finally, we test the ability of the model to reveal the origin of experimentally known position-specific effects on phase separation.

METHODS

Hydropathy scale model for PTMs

The functional form of the total energy of the system is as follows,

$$\Phi(r) = \sum_{bonds} k_{spring}(r - r_0)^2 + \sum_{i<j}\left[\Phi_{elec}(r_{ij}) + \Phi_{nb}(r_{ij})\right]$$

where Φ_{bond} is a standard harmonic spring with $k_{spring} = 10\frac{\text{kcal}}{\text{mol Å}^2}$ and $r_0 = 3.8$ Å. The screened electrostatic term is represented using a Debye-Hückel type form:

$$\Phi_{elec}(r) = \frac{q_i q_j}{4\pi D r} e^{\frac{-r}{\kappa}},$$

where q_i and q_j are the net charges of formally charged amino acids (D,E = −1; K,R = 1; H = 0.5), D is the dielectric constant (which is set to 80 for water), and κ is the screening length, which is set to 10 Å to represent a salt concentration of ∼100 mM. The nonbonded interactions are described by the Ashbaugh-Hatch functional form, which has been previously applied to the study of disordered proteins (55). In this Lennard-Jones-like functional form, the attractiveness of the interactions is scaled by the arithmetic average of hydropathy for the two interacting amino acid types, $\lambda = 0.5(\lambda_i + \lambda_j)$:

$$\Phi_{LJ} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right],$$

$$\Phi_{nb}(r) = \begin{cases} \Phi_{LJ} + (1-\lambda)\varepsilon, \ if\ r \leq 2^{1/6}\sigma \\ \lambda\Phi_{LJ}, \quad otherwise \end{cases},$$

where Φ_{LJ} is the standard Lennard-Jones potential and σ is the arithmetic average of the vdW radius,

$$\sigma_{ij} = \frac{(\sigma_i + \sigma_j)}{2},$$

where $\sigma = 2r_{vdW}$ because the atoms are modeled as spheres, and their impenetrable volumes are defined by the vdW radius (56) r_{vdW}.

Determining hydropathy values from atomic charges

The original HPS model is based on the hydrophobicity scale put forth by Kapcha and Rossky (57) using the atomic charges from the atomistic optimized potentials or liquid simulations (OPLS) force field for proteins. The hydropathy scale was then scaled to the range of 0–1, with 0 being the least hydrophobic and 1 being the most hydrophobic. We further extend this procedure to obtain hydropathy values for post-translationally modified amino acid residues using partial atomic charges from a set of AMBER forcefield parameters for 32 common PTMs (FF-PTM) (35). Atoms with a partial charge magnitude ranging from 0 to 0.25 are considered nonpolar and are assigned a hydropathy value of −1. In the binary atomic-level hydropathy scale, atoms with a partial charge magnitude greater than 0.25 are considered polar and are assigned a hydropathy value of +1. The residue-level hydropathy values are calculated as a weighted sum of the atomic hydropathy values according to the following rules: hydrophobic (nonpolar) atoms contribute −0.5, hydrophilic (polar) atoms contribute +1, and charged atoms contribute +2. An atom in charged residues is classified as charged if it belongs to the terminal polar group in which the magnitude of the sum of all atomic partial charges in that group is greater than 0.5 (57).

Simulations framework

To obtain the temperature under which a protein chain behaves as an ideal solvent (T_θ), single-chain simulations were conducted for 1 μs at a range of temperatures using replica exchange molecular dynamics (REMD) (58) with a temperature list of 150.0, 170.1, 193.0, 218.9, 248.3, 281.7, 300, 362.4, 411.1, 466.3, 529.0, and 600.0 K to enhance ergodic sampling of the IDR conformational ensemble (59). Simulations were conducted in cubic boxes with periodic boundaries large enough so that a protein chain will not encounter its periodic image, and temperature was maintained using a Langevin thermostat. All single-chain simulations were conducted using LAMMPS (60). To get better sampling for the Rg for proteins in Fig. 1, we simulated DDX3X IDR1, DDX3X IDR1 10acK, FUS LC, FUS LC 12 pS/pT, FUS LC 12 E (−2) using REMD (58) for 5 μs with a temperature list of 150.0, 154.49, 159.11, 163.88, 168.78, 173.84, 179.04, 184.4, 189.92, 195.6, 201.46, 207.49, 213.7, 220.1, 226.69, 233.47, 240.46, 247.66, 255.07, 262.71, 270.57, 278.67, 287.02, 295.61, 304.46, 313.57, 322.96, 332.63, 342.58, 352.84, 363.4, 374.28, 385.48, 397.02, 408.91, 421.15, 433.75, 446.74, 460.11, 473.88, 488.07, 502.68, 517.73, 533.23, 549.19, 565.63, 582.56, and 600.0 K.

Coexistence simulations were conducted using slab geometry (61,62) like in our previous work (15,37,48,63–65). Simulations were conducted on 100 chains of FUS or DDX3 using the HPS model and the new parameters for PTMs. Simulations were conducted using HOOMD-Blue v2.1.5 (66) for 5 μs each in serial with a Langevin thermostat, during which the first μs was discarded as equilibration.

Single-molecule properties

Using the radius of gyration (R_g) and the polymer-scaling exponent (ν) in different temperatures as descriptors of the phase behavior, we calculate the temperature at which a protein behaves as an ideal polymer: the point where protein-protein interactions and protein-solvent interactions become energetically equally favorable (T_θ) (67). For each temperature, we estimated the polymer-scaling exponent (ν) by fitting to the following equation:

$$R_{ij} = b\,|\,i-j|^{\nu},$$

where b is the Kuhn length, which is set to 0.55 nm for disordered proteins (68), and R_{ij} is the intrachain separation defined as the distance between residue i and residue j. Next, we interpolated the results to find the temperature at which v = 0.5, which is correlated with the critical temperature for phase separation (63).

Sequence-charge decoration

We use the sequence-charge decoration (SCD) parameter, which is a metric first presented in the analytical model of Sawle and Ghosh, to capture the effect of charge patterning on the values of radius of gyration of synthetic polyampholyte sequences (69). They showed that electrostatic interactions are enhanced by local patches of like charges and attenuated when charged residues are more mixed within the sequence. For a sequence with equal number of positive and negative residues (of net charges −1 and +1), highly negative SCD-values correspond to stretches of like charges clustered in the sequence, whereas values near zero result from well-mixed charge residues, with a sequence of strictly alternating positive and negative charges approaching SCD of 0. Positive values of SCD result from sequences that have a large net charge (see Fig. S2 for schematic of SCD). SCD is calculated (67) using the following equation:

$$SCD = \frac{1}{N}\left[\sum_{i=2}^{N}\sum_{j=1}^{i-1} q_i q_j (i-j)^{\frac{1}{2}}\right],$$

where $i-j$ is the separation of residues with charges q_i and q_j along a chain of length N.

Phase diagrams

Phase diagrams were extracted from coexistence simulations by measuring the density of protein in the dense phase in the center, and in the low-density phase outside. The critical temperature and critical density were obtained using the top seven highest temperatures for the maximal SCD case and eight for the minimal SCD case in which we observed phase coexistence. We fit the coexistence densities at these temperatures to a fitting function:

$$\rho_H - \rho_L = A(T_c - T)^{\beta},$$

where β is set to 0.325 according to the universality class of the three-dimensional Ising model (66).

RESULTS

Expanding the HPS model to incorporate phosphorylation and acetylation

To extend our primary sequence-specific CG model to study PTMs, we employed the original philosophy of computing CG interaction parameters from the size, net charge, and atomic hydropathy of the post-translationally modified amino acids. Our previous CG representation of the protein represents every natural amino acid by a single bead of unique mass, charge, size, and hydropathy parameter that were computed from a fully atomistic fixed-charge force field parameter for each residue type (35,48,57). Using previously published AMBER force field parameters for modified amino acids (34), we computed CG parameters for serine, threonine, and tyrosine phosphorylation and lysine acetylation (Table S1). We have selected to represent phosphorylated amino acids as bearing a -2 charge, given that at neutral pH the net charge on phosphorylated residues is likely to be close to -2 (see above) (26,70). We note that the proposed framework is not meant to capture dynamic pK_a shifts because of the local environment, which can change the pK_a. To simulate lower pH conditions or if the pK_a is experimentally known or likely to be shifted as may be the case in charge-dense polyelectrolyte chains, it may be useful in these situations to simulate phosphorylated residues with a different net charge (i.e., -1), as previously suggested (71). In principle, these parameterization of CG parameters for modified residues can be applied to any library of noncanonical amino acids with known structure or charge parameters to aid the high-throughput screening of proteins with unknown PTM-regulated phase separation.

To test the effects of these PTMs on CG models of IDRs, we used the HPS-PTM model to characterize the properties of two modified IDRs: phosphorylated FUS LC (FUS LC 12pS/pT) (Fig. 1 *A*) and acetylated human DDX3X IDR1 protein (DDX3X IDR1 10acK) (Fig. 1 *C*). Single-chain CG simulations were conducted for 5 μs at a range of temperatures (150–600 K) using REMD to enhance sampling (58). Unmodified FUS LC is nearly uncharged (containing only two charged residues) but can be phosphorylated by DNA-dependent protein kinase at 12 unique sites, displaying either SQ or TQ motif (53). Here, we begin by testing the effect of hyperphosphorylation at all 12 sites (FUS LC 12pS/pT). In the hyperphosphorylated state, FUS LC 12pS/pT is more extended than the unmodified state, having a much larger radius of gyration (R_g) at the reference temperature (295 K) (Fig. 1 *B*). Hyperphosphorylation also reduces T_θ (Fig. S1, *A and C*), the temperature at which attractive intramolecular interactions are balanced by repulsive excluded volume interactions (67). Both of these reflect experimental observations that phosphomimetic mutants of FUS are more extended in solution and are less prone to phase separate (24,37).

The observed difference in phase separation is likely caused by the large change in net charge of FUS LC, but phosphoserine and phosphothreonine also differ from the unmodified residues in the other model parameters. In our previous work, we showed that the electrostatic contribution of phosphomimetic mutations was larger than the impact of the change in λ (24). Here, we have explored the contribution of the CG bead size and hydropathy parameters on the chain dimensions. We simulated the FUS LC 12E variant in which a Glu residue has been substituted instead of 12pS/pT residues and has been assigned a charge of -2 to match the charge of pS and pT (instead of its usual -1 charge). This FUS LC 12E (-2) is, indeed, more extended than FUS LC but not to the same extent as FUS LC 12pS/pT. These data indicate that the role of nonelectrostatic interactions is not negligible and, thus, show additional value of the CG model that incorporates explicit modified residues (Fig. 1 *B*).

Next, we tested the effect of acetylation using single-chain REMD simulations of DDX3X IDR1 and DDX3X IDR1 10acK, which is the acetylated variant of DDX3RX IDR1 by replacing 10 lysines with acetyllysines in positions identified to be acetylated in vivo (54). At the reference 295 K temperature, we found that the modified state is also more expanded than the unmodified state, but the difference is considerably less than that of FUS LC 12pS/pT from unmodified FUS LC (Fig. 1 *D*). Similarly, T_θ is only slightly decreased (Fig. S1, *B and D)*, consistent with the small change in R_g. This finding is also consistent with experimental results showing that acetylmimetic DDX3X is less prone to phase separate (54).

We suspect that the larger changes in single-chain properties for FUS LC hyperphosphorylation compared with DDX3X IDR1 hyperacetylation may be caused by the nearly uncharged nature of unmodified FUS LC as opposed to the highly charged polyampholytic character of DDX3X IDR1 in which other positively and negatively charged residues are extensively distributed throughout the sequence. Therefore, acetylation of lysine residues across the entire DDX3X IDR1 does not dramatically alter the intramolecular interactions, which we probe further in the following sections. Taken together, these data demonstrate that we can model changes in biophysical properties of IDRs using this CG approach.

Effect of increasing number of PTMs on chain properties of FUS LC and DDX3X IDR1

We next wanted to test the effect of increasing the number of modified sites on the expansion or collapse of the IDRs. The number of possible modification combinations is very large, given by the following equation:

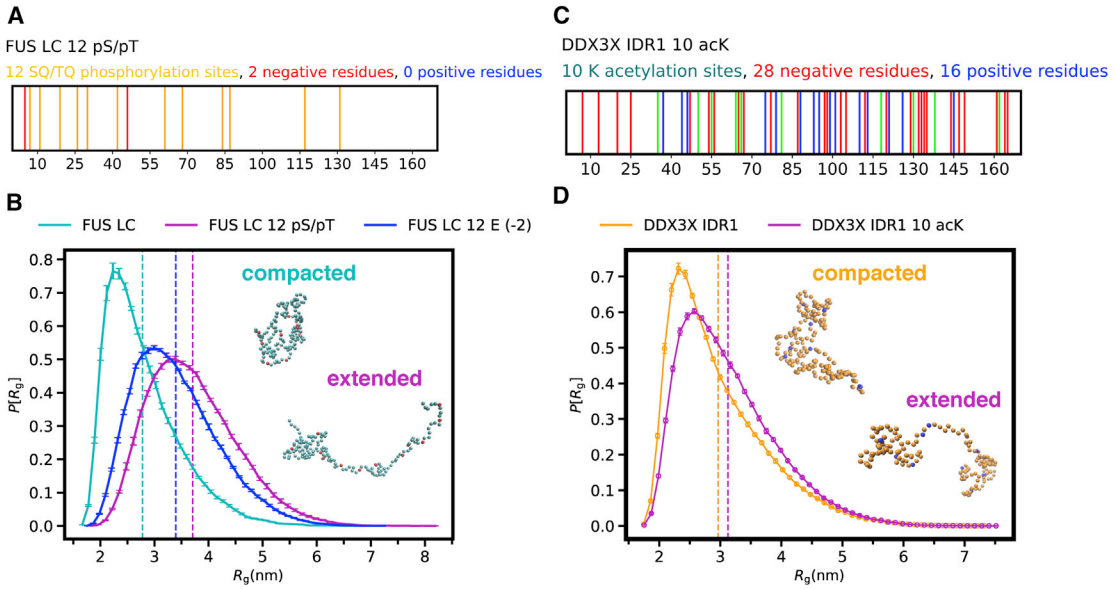

FIGURE 1 Expanding the HPS model to incorporate phosphorylation and acetylation. (*A*) Domain representation of FUS LC 12 pS/pT. Negative residues D5 and D46 are highlighted in red. Phosphorylation sites at 7, 11, 19, and 68 (phosphothreonine) and 26, 30, 42, 61, 84, 87, 117, and 131 (phosphoserine) are highlighted in orange. (*B*) Distribution of radius of gyration of FUS LC (*cyan*), FUS LC 12 pS/pT (*purple*), and FUS 12 E (−2) with −2 charge for Glu (*blue*). (*C*) Domain representation of DDX3X IDR1. Negative residues are shown in red, positive residues (other than lysine, i.e., arginine) in blue, and lysines in green. (*D*) Distribution of radius of gyration of DDX3X IDR1 (*orange*) and DDX3X IDR1 10 acK (*purple*). In representative snapshots, the unmodified beads are shown in cyan (FUS LC) or orange (DDX3X IDR1), the negatively and positively charged residues in red and blue, respectively, and the modified sites in purple. Dashed lines represent the ensemble average R_g. Error bars represent standard error of the mean (SEM) from 5 equal divisions. To see this figure in color, go online.

$$\sum_{m=0}^{M} \frac{M!}{m!(M-m)!} = 2^M,$$

where M is the total number of possible PTM sites, and m is the number of sites that are modified. Although there is only one possible state with 0, or M modifications, there are 924 possible states of FUS with six modified residues, and 252 possible states of DDX3 with five modified residues. We selected a subset of 10 random modification patterns for each number of modifications sites ($1 \leq m \leq 11$ for FUS LC and $1 \leq m \leq 9$ for DDX3X IDR), and performed simulations on these sequences. Because the position of charged residues can play an important role in chain dimensions (72), we also computed the SCD parameter (69) for all possible sequences (including both the PTMs and the other charged residues in the sequence) to determine the SCD extremes where like-charged residues are clustered together (SCD far from 0) or charged residues are well-mixed (SCD close to 0). See Materials and methods and Fig. S2 for more information on SCD. We included both the SCD minimum and maximum in our simulated sequences to complement our randomly chosen sequences.

For FUS LC, R_g increases with increasing number of phosphorylation sites (Fig. 2 *A*; Fig. S3, *A and B*), going from slightly below 3.0 nm for unmodified to 3.8 nm for FUS LC 12pS/pT. This large expansion is consistent with increasing polyelectrolytic character because of phosphorylation, adding increasing numbers of negatively charged sites to FUS LC, which is otherwise nearly uncharged.

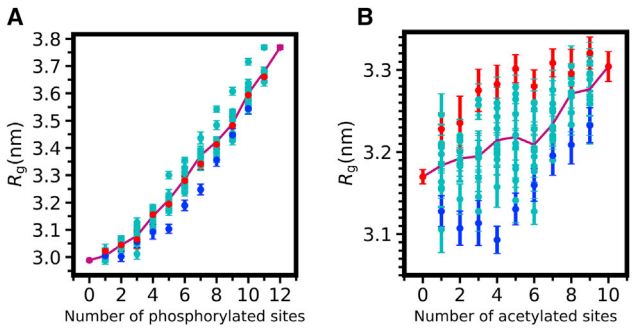

FIGURE 2 Effect of increasing number of PTMs on chain properties of FUS LC and DDX3X IDR1. (*A*) Random phosphorylation patterns expand FUS LC monotonically. Cyan data points show the ensemble average R_g for each variant of 10 randomly selected phosphorylation patterns for a given number of phosphorylation sites in FUS LC. Violet line represents the mean R_g-value from 10 simulations. Red and blue data points show the R_g of SCD-maximal and SCD-minimal sequences, respectively. (*B*) Random acetylation patterns in DDX3X yield to a moderate expansion. Cyan data points show the ensemble average R_g for each variant of 10 randomly selected acetylation patterns for a given number of acetylation sites in DDX3X IDR1. Violet line represents the mean R_g-value from 10 simulations. Red and blue data points show the ensemble average R_g of SCD-maximal and SCD-minimal sequences, respectively. Error bars represent standard error of the mean (SEM) from 10 equal divisions of the equilibrated ensembles. To see this figure in color, go online.

Interestingly, sequences with the same number of modifications but in different patterns showed a range of R_g, which we explore further (see below).

We then probed the chain properties of the DDX3X IDR1 as a function of lysine acetylation. To quantify the difference in chain properties as a function of increasing number of acetylation modifications, we performed single-chain simulations selecting 10 random acetylation patterns for each number of acetylated sites, analogous to the procedure we used for FUS LC phosphorylation above. We found that with increasing number of modifications, the chain dimension increases only slightly, from 3.2 to 3.3 nm (Fig. 2 *B*; Fig. S3, *C and D)*. Importantly, sequences carrying the same number of acetylated lysines but different arrangements of the acetylation sites also exhibit the variation of chain dimensions. As for FUS LC, we further explore the connection between the arrangement of the PTMs and the distribution of R_g below.

Comparing post-translationally modified residues to common modification mimetics

A widely used approach in experimental studies of PTMs is to mimic the change in amino acid character caused by the modification with a natural amino acid substitution. This allows for precise control of the location of modified residues. In the case of phosphorylation, serine and threonine are typically mutated to aspartic acid (73) or glutamic acid (54,74,75), which mimic the negative charge and larger size of the phosphorylated residue. Here, we evaluate aspartic acid or glutamic acid substitutions as phosphomimics using our model. We found that aspartic acid and glutamic acid do not fully recapitulate the chain dimensions observed with phosphorylation (Fig. 3 *A*), likely because phosphorylated residues are represented as having a −2 charge compared with the −1 charge of the phosphomimics. We tested this hypothesis that the dominant contribution to the difference in chain expansion for phosphorylation versus phosphomimetic mutation is caused by the difference in added charge by replotting R_g as a function of net charge (Fig. 3 *B*). Indeed, we find that the effect of phosphorylation on FUS LC chain dimensions in the CG model appears to be largely, though not entirely (see Fig. 1 *A*), driven by the charge of the modifications. Similarly, for the glutamine, which is commonly used to mimic acetylation of lysine (removing the positive charge of the amino acid) (76,77), we find that glutamine substitutions result in similar R_g to lysine acetylation in DDX3X IDR1 (Fig. 3 *C*). Small differences in the glutamine and acetyllysine chain dimensions again point to the contribution of parameters (i.e., size and hydropathy, Table S1) other than the residue net charge in determining the contacts formed. These observation in turn demonstrate that capturing such effects require the parameterization of CG models in a transferable way, which is an important feature of the approach that we follow in our modeling framework.

Different PTM patterns with the same number of modifications have differing effects on LLPS

Although adding more PTMs or mimics resulted on average in a change in the chain dimensions, different patterns of PTMs with the same number of modifications in FUS LC and DDX3X IDR1 show significantly different R_g (Fig. 2). This is an interesting observation for FUS LC and DDX3X IDR1 because it implies that the placement of the PTMs can impact the chain properties. To attempt to explain the variance in R_g, we computed and plotted the SCD for our sequences. Although the impact of charge patterning as measured by SCD is largely predictive of the R_g for FUS LC when looking across all number of modifications simultaneously (Fig. S4 *A*), the SCD does not fully capture the behavior for DDX3X IDR1 (Fig. S4 *C*) as is also apparent when examining the SCD only for a given number of modifications (Fig. S4 *D*). These deviations may be cause by contributions from hydropathy patterning (78) as well as inherent limitations in SCD (44). In summary, the impact of other physical parameters on the chain dimensions are not trivial and underscore the potential for CG models to complement theoretical predictions, which only consider charged residue patterning.

Given that IDR chain dimensions correlate with phase behavior (63,79), we next tested the effect of the placement

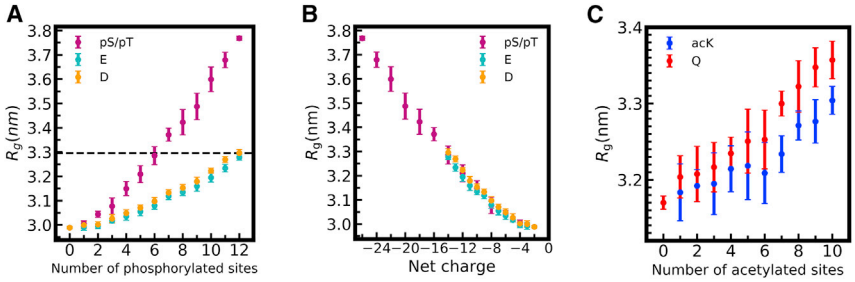

FIGURE 3 Comparing post-translationally modified residues to common modification mimetics. (*A*) Effect of FUS LC phosphomimetic substitutions glutamic acid (E, *cyan*) or aspartic acid (D, *orange*) on R_g compared with phosphorylation (pS/pT, *violet*). (*B*) The same data as in (*A*) but with the *x* axis shown as the net charge of FUS LC, demonstrating that the difference between phosphomimetic and phosphorylated R_g can be largely explained by difference in net charge of aspartic acid and glutamine acid (−1) and phosphoserine or phosphothreonine (−2). (*C*) Effect of DDX3X IDR1 acetylmimetic glutamine (Q, *red*) substitutions on R_g compared with acetylated lysine (acK, *blue*). Data represent mean and standard deviation from 10 individual simulations with random phosphorylation or substitution patterns per site except for 0 or n = maximal number of substitutions (for which there is only one pattern: all or none) where the block average and SEM is shown. To see this figure in color, go online.

of PTMs on LLPS. We noted earlier that there may be dramatic differences in the single-chain properties for sequences with the same number of PTMs. To understand the differences between the sequences, we decided to focus on DDX3X IDR1 where, even when the number of modifications is held constant, the arrangement of the acetylation modifications may have a large impact on R_g (Fig. 2 *B*). For the two sequences with highest or lowest values of R_g and two with extreme values of SCD, we computed T_θ and found that it is significantly different for sequences with the same modification number (and hence same net charge) but different sequence (Fig. 4 *A*). To demonstrate the position-specific effect of PTMs on the phase diagram of acetylated DDX3X IDR1, we simulated the two sequences with four acetylated lysines that exhibited the greatest difference in single-chain properties (Table S2). We find that positioning four acetylated lysines within the predominantly negatively charged C-terminus of DDX3X IDR1 enhances LLPS compared with the sequence that places the four acetylated lysines in the polyampholytic N-terminal/central region (Fig. 4 *B*). For reference, we also show the phase diagrams of the acetylmimetic variants in Fig. S5.

Together, these data demonstrate that the arrangement as well as the quantity of PTMs can have a large impact on the phase separation behavior of IDRs. This also highlights the necessity of methods that allow for control over which PTMs are "activated" to fully characterize a disordered protein and its range of responses to regulation by PTMs.

Predicting effect of different modification patterns on phase separation: comparison with experiment

To test the predictive ability of the model, we decided to further explore the differences in predicted phase behavior for different modification patterns at a set number of modifications. To this end, we examined a series of single-position acetylmimetic (lysine to glutamine substitutions) of DDX3X IDR1 whose phase separation has been previously examined by experiment (54). We again show that in the CG model, single-glutamine mimetics and acetylated lysine residues at the same position yield similar though not identical results (Fig. 5 *A*). We do notice that the replacement of lysines with acetylated lysines in DDX3X IDR1 has a slightly bigger effect on R_g than substitution of acetylated lysines with glutamines, most prominently near the C-terminus of the sequence (residues 118–162). However, between different single modifications, we observe widely varying effects on chain properties, highlighting the importance of not only the number of PTMs, but the location (Fig. 5 *A*).

It is worth noting that single-acetylated lysine substitutions in the C-terminus (K118, K130, K138, K162) result in DDX3X IDR1 slight collapse with respect to wild-type (WT), whereas changes in other positions (K35, K50, K55, K64, K66, K81) lead to slight chain expansion. This observation is also consistent with the result in the previous section that four simultaneously acetylated lysines can enhance LLPS or decrease LLPS depending on their location in the C- or N-terminus (because of changes in the SCD), respectively. As discussed throughout the literature (44,78–82), changes in charge patterning with fixed net charge can be an important parameter, leading to position-specific effects. Based on this, we can say that the observed position dependence in our case mainly comes from the well-known effects in the field. Furthermore, we find that even though SCD can capture changes with an increasing number of PTMs (Fig. S4), it cannot offer precise guidance for a fixed number of PTM sites located in different regions of peptide (Fig. S4, *B and D*). As we mentioned above, this can either be because of the limitations of SCD in capturing the subtle differences or the role of additional factors related to nonelectrostatic interactions and secondary structures.

Given that R_g is an excellent predictor of IDR phase separation (78), we then quantified the correlation between computationally predicted R_g and previously published data on phase separation measured experimentally by

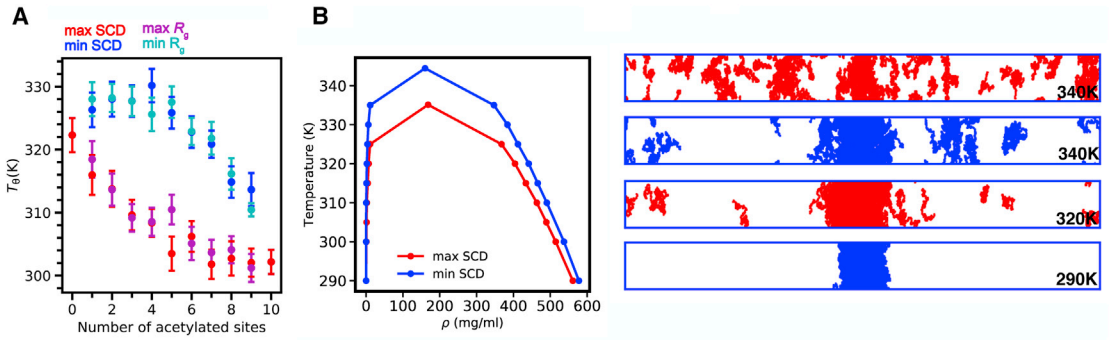

FIGURE 4 Different PTM patterns with the same number of modifications have differing effects on chain expansion and LLPS. (*A*) T_θ temperature of R_g-minimal (*cyan*) and R_g maximal (*magenta*), SCD minimal (*blue*), and SCD maximal (*red*) DDX3X IDR1 sequences acetylated at n = 1 to n = 9 sites. Non-acetylated (0) and fully acetylated (10) variants are shown in red. Error bars calculated based on the posterior uncertainty of the probabilistic surrogate used to fit (T_θ, ν) pairs (see Fig. S1) (*B*) Phase diagram of SCD minimum and SCD maximum of DDX3X IDR1 acetylated at four lysines are shown along with the simulation snapshots at multiple temperatures. To see this figure in color, go online.

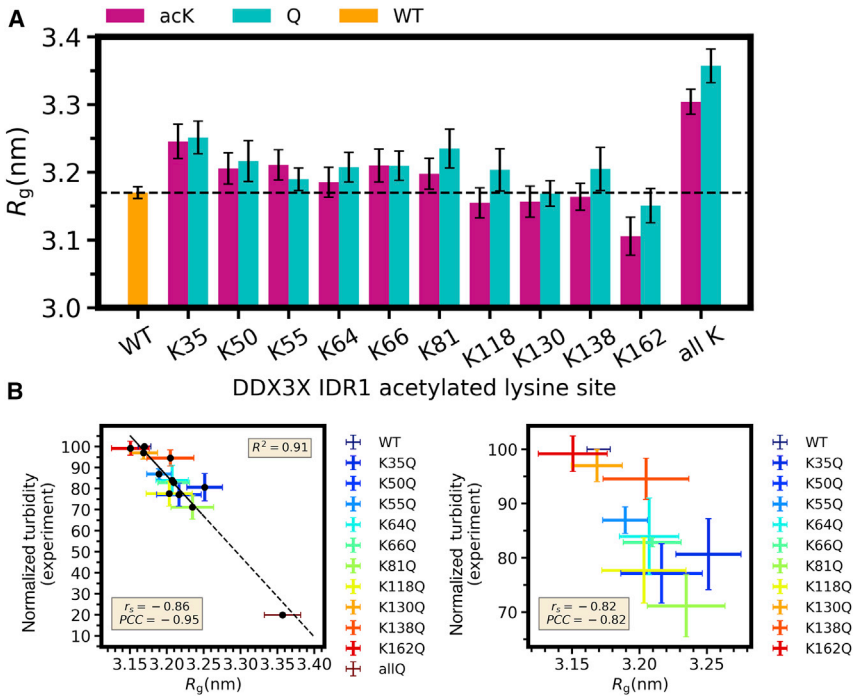

FIGURE 5 Predicting effect of different modification patterns on phase separation: comparison with experiment. (*A*) Lysine acetylation (acK, *violet*) and lysine-to-glutamine-acetylmimetic mutants (*cyan*) show a range of R_g-values compared with WT (*orange*) and all fully acetylated (all K) or fully acetylmimetic (allQ). Error bars represent SEM from 10 equal divisions of the equilibrated ensemble. (*B*) Correlation of normalized turbidity measured by experiment and R_g predicted by simulations with allQ (10 acetylation-targeted lysines converted to glutamines) (*left panel*) and without (*right panel*) to clearly illustrate the behavior of each individual mutant. The values of Spearman (rank order, r_s) and Pearson coefficient (PCC) show strong negative monotonic and linear correlation, respectively. Turbidity values (OD_{600}) and corresponding error bars represent mean value and SD from 3 independent experiments (previously published data by Saito et al. (54)). The dashed line is meant to be a guide to the eye and is not intended to support a linear correlation between the two variables. To see this figure in color, go online.

turbidity (Fig. 5 *B*), where turbidity (optical density at 600 nm (OD_{600})) of the DDX3X IDR1 K to Q mutants are normalized to the OD_{600} of the WT that phase separates avidly generating liquid droplets which scatter light (54). As expected, none of these single-position modifications has as much impact as simultaneous acetylmimetic substitutions at all 10 lysine positions (allQ) (Fig. 5 *B*). Importantly, R_g is negatively correlated with the extent of phase separation as measured by experimental turbidity values (Fig. 5 *B*; (54)). As the "allQ" mutant (Fig. 5 *B*, *left panel*) significantly differs from the singe-Q variants, we also present the same data zoomed in and the correlation coefficients without this "allQ" variant (Fig. 5 *B*, *right panel*). Still, the linear and rank order correlation coefficients show strong correlation between the experimental turbidity and the simulated R_g. It is remarkable that the model can capture site-specific changes in the phase separation of DDX3X mutants so well, highlighting the accuracy of the model with regard to its intended goal of capturing such changes predictively without experimental input.

DISCUSSION

Understanding the role of PTMs on phase separation is a critical element to controlling membraneless organelle formation in cells and engineering LLPS in vitro.

In this study, we found that phosphorylation increases chain dimensions of FUS LC steeply and monotonically but that acetylation of DDX3X IDR1 has only a small average effect on chain expansion. CG single-chain simulations revealed that phosphorylation expands FUS LC more than phosphomimetic mutations largely but not entirely because of the higher net charge of phosphorylation compared with the phosphomimetic substitutions. We found that particular patterns of PTMs with the same number of modified sites result in a large diversity in chain expansion, which is strongly correlated with the propensity to phase separate. Importantly, we show modification patterns with the same number of modified sites can have dramatically different phase separation depending on the position of the modified residues. Furthermore, the model can predict the effect of position-specific modification on experimentally observed DDX3X IDR1 phase separation. Together our data suggest that the CG model can represent how the response of disordered regions to multisite modification depends on the modification patterning. In the future, it will be interesting to probe the effect on phase separation of IDR ensembles containing heterogenous combinations of IDRs with a varied distribution of both number and positions of modifications as may be present in natural systems.

Our findings highlight the importance of combinatorial effects exerted by phosphorylation and acetylation on disordered proteins. Such effects often remain unexplored by in vitro biochemistry approaches because protein modification reactions targeting only certain modification sites are difficult to achieve. However, we have noted that DNA-dependent protein kinase targets FUS LC at 12 SQ or TQ sites for phosphorylation with high specificity, modifying only these 12 out of a total of 52 S/T sites in FUS LC (24). With our CG model we can selectively modify amino acids of disordered regions and predict their effect on single-chain properties and phase behavior. These simulations can then

enable future engineering of site-specific phosphorylation by introducing, for example, only SQ/TQ sites (and not other S/T sites) as specific targets for DNA-PK modification. Such computational studies can also aid predictions of phase separation properties of with IDRs with multisite modifications.

It should be noted that threonine/serine phosphorylation and lysine acetylation are not the only PTMs with a known or potential effect on LLPS. Methylation of arginine residues in proteins leads to stimulation (83) or suppression (8,37,84) of phase separation depending on the sequence content and the underlying cellular function. Importantly, methylation of arginine (and lysine) residues does not alter the charge of the residue (unlike lysine acetylation and serine/threonine/tyrosine phosphorylation). Hence, interactions other than charge are likely responsible for the changes to phase behavior by methylation and may be challenging to capture in CG model frameworks. Investigation of the role of methylated arginines and lysines on LLPS using transferable CG models remains important future work.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2021.01.034.

AUTHOR CONTRIBUTIONS

T.M.P. and J.M. designed and performed the single-chain CG simulations and analyzed the resulting data. N.J. and J.M. designed, performed, and analyzed slab simulations. G.L.D., Y.C.K., and J.M. developed the methodology, provided the software, and supervised the formal analysis. N.L.F. and J.M. contributed to project administration and funding acquisition. T.M.P., N.L.F., and J.M. wrote the manuscript with comments from all authors.

ACKNOWLEDGMENTS

Use of the high-performance computing capabilities of the Extreme Science and Engineering Discovery Environment, which is supported by the National Science Foundation grant TG-MCB-120014, is gratefully acknowledged. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. We thank Anastasia Murthy and Veronica Ryan for helpful discussions. We thank Patrick Matthias for providing previously published tabulated data on DDX3X phase separation.

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM136917 (to J.M.). Research was also supported in part by National Institute of General Medical Services R01GM118530 (to N.L.F.), National Institute of Neurological Disorders and Stroke and National Institute on Aging R01NS116176 (to N.L.F. and J.M.), National Science Foundation 2004796 (to J.M.), and National Science Foundation 1845734 (to N.L.F.). Y.C.K. is supported by the Office of Naval Research via the U.S. Naval Research Laboratory Base Program.

REFERENCES

1. Shin, Y., and C. P. Brangwynne. 2017. Liquid phase condensation in cell physiology and disease. *Science.* 357:eaaf4382.
2. Ditlev, J. A., L. B. Case, and M. K. Rosen. 2018. Who's in and who's out-compositional control of biomolecular condensates. *J. Mol. Biol.* 430:4666–4684.
3. Ryan, V. H., and N. L. Fawzi. 2019. Physiological, pathological, and targetable membraneless organelles in neurons. *Trends Neurosci.* 42:693–708.
4. Banani, S. F., H. O. Lee, …, M. K. Rosen. 2017. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18:285–298.
5. Molliex, A., J. Temirov, …, J. P. Taylor. 2015. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell.* 163:123–133.
6. Feric, M., N. Vaidya, …, C. P. Brangwynne. 2016. Coexisting liquid phases underlie nucleolar subcompartments. *Cell.* 165:1686–1697.
7. Darling, A. L., Y. Liu, …, V. N. Uversky. 2018. Intrinsically disordered proteome of human membrane-less organelles. *Proteomics.* 18:e1700193.
8. Qamar, S., G. Wang, …, P. St George-Hyslop. 2018. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interactions. *Cell.* 173:720–734.e15.
9. Vernon, R. M., P. A. Chong, …, J. D. Forman-Kay. 2018. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife.* 7:1–48.
10. Burke, K. A., A. M. Janke, …, N. L. Fawzi. 2015. Residue-by-residue view of in vitro FUS granules that bind the C-terminal domain of RNA polymerase II. *Mol. Cell.* 60:231–241.
11. Dignon, G. L., R. B. Best, and J. Mittal. 2020. Biomolecular phase separation: from molecular driving forces to macroscopic properties. *Annu. Rev. Phys. Chem.* 71:53–75.
12. Murthy, A. C., G. L. Dignon, …, N. L. Fawzi. 2019. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* 26:637–648.
13. Schwartz, J. C., X. Wang, …, T. R. Cech. 2013. RNA seeds higher-order assembly of FUS protein. *Cell Rep.* 5:918–925.
14. Hayes, M. H., E. H. Peuchen, …, D. L. Weeks. 2018. Dual roles for ATP in the regulation of phase separated protein aggregates in *Xenopus* oocyte nucleoli. *eLife.* 7:1–24.
15. Dignon, G. L., W. Zheng, …, J. Mittal. 2019. Temperature-controlled liquid-liquid phase separation of disordered proteins. *ACS Cent. Sci.* 5:821–830.
16. Hofweber, M., and D. Dormann. 2019. Friend or foe-post-translational modifications as regulators of phase separation and RNP granule dynamics. *J. Biol. Chem.* 294:7137–7150.
17. Bah, A., and J. D. Forman-Kay. 2016. Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem.* 291:6696–6705.
18. Khoury, G. A., R. C. Baliban, and C. A. Floudas. 2011. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* 1:1–5.
19. Andrusiak, M. G., P. Sharifnia, …, Y. Jin. 2019. Inhibition of axon regeneration by liquid-like TIAR-2 granules. *Neuron.* 104:290–304.e8.
20. Ferreon, J. C., A. Jain, …, A. C. Ferreon. 2018. Acetylation disfavors Tau phase separation. *Int. J. Mol. Sci.* 19:1360.
21. Audagnotto, M., and M. Dal Peraro. 2017. Protein post-translational modifications: *in silico* prediction tools and molecular modeling. *Comput. Struct. Biotechnol. J.* 15:307–319.
22. Kim, T. H., B. Tsang, …, J. D. Forman-Kay. 2019. Phospho-dependent phase separation of FMRP and CAPRIN1 recapitulates regulation of translation and deadenylation. *Science.* 365:825–829.
23. Brunk, E., R. L. Chang, …, N. E. Lewis. 2018. Characterizing post-translational modifications in prokaryotic metabolism using a multi-scale workflow. *Proc. Natl. Acad. Sci. USA.* 115:11096–11101.
24. Monahan, Z., V. H. Ryan, …, N. L. Fawzi. 2017. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* 36:2951–2967.

25. Tsang, B., J. Arsenault, …, J. D. Forman-Kay. 2019. Phosphoregulated FMRP phase separation models activity-dependent translation through bidirectional control of mRNA granule formation. *Proc. Natl. Acad. Sci. USA.* 116:4218–4227.

26. Xie, Y., Y. Jiang, and D. Ben-Amotz. 2005. Detection of amino acid and peptide phosphate protonation using Raman spectroscopy. *Anal. Biochem.* 343:223–230.

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

28. Kamah, A., I. Huvent, …, C. Smet-Nocca. 2014. Nuclear magnetic resonance analysis of the acetylation pattern of the neuronal Tau protein. *Biochemistry.* 53:3020–3032.

29. Wegmann, S., B. Eftekharzadeh, …, B. T. Hyman. 2018. Tau protein liquid-liquid phase separation can initiate tau aggregation. *EMBO J.* 37:e98049.

30. Martin, E. W., A. S. Holehouse, …, T. Mittag. 2016. Sequence determinants of the conformational properties of an intrinsically disordered protein prior to and upon multisite phosphorylation. *J. Am. Chem. Soc.* 138:15323–15335.

31. Huihui, J., T. Firman, and K. Ghosh. 2018. Modulating charge patterning and ionic strength as a strategy to induce conformational changes in intrinsically disordered proteins. *J. Chem. Phys.* 149:085101.

32. Margreitter, C., D. Petrov, and B. Zagrovic. 2013. Vienna-PTM web server: a toolkit for MD simulations of protein post-translational modifications. *Nucleic Acids Res.* 41:W422–W426.

33. Rani, L., J. Mittal, and S. S. Mallajosyula. 2020. Effect of phosphorylation and O-GlcNAcylation on proline-rich domains of Tau. *J. Phys. Chem. B.* 124:1909–1918.

34. Khoury, G. A., J. Smadbeck, …, C. A. Floudas. 2014. Forcefield_NCAA: ab initio charge parameters to aid in the discovery and design of therapeutic proteins and peptides with unnatural amino acids and their application to complement inhibitors of the compstatin family. *ACS Synth. Biol.* 3:855–869.

35. Khoury, G. A., J. P. Thompson, …, C. A. Floudas. 2013. Forcefield_PTM: Ab initio charge and AMBER forcefield parameters for frequently occurring post-translational modifications. *J. Chem. Theory Comput.* 9:5653–5674.

36. Zerze, G. H., and J. Mittal. 2015. Effect of O-linked glycosylation on the equilibrium structural ensemble of intrinsically disordered polypeptides. *J. Phys. Chem. B.* 119:15583–15592.

37. Ryan, V. H., G. L. Dignon, …, N. L. Fawzi. 2018. Mechanistic view of hnRNPA2 low-complexity domain structure, interactions, and phase separation altered by mutation and arginine methylation. *Mol. Cell.* 69:465–479.e7.

38. Ryan, V. H., T. M. Perdikari, …, N. L. Fawzi. 2021. Tyrosine phosphorylation regulates hnRNPA2 granule protein partitioning and reduces neurodegeneration. *EMBO J.* 40:e105001.

39. Vymětal, J., V. Jurásková, and J. Vondrášek. 2019. AMBER and CHARMM force fields inconsistently portray the microscopic details of phosphorylation. *J. Chem. Theory Comput.* 15:665–679.

40. Zerze, G. H., W. Zheng, …, J. Mittal. 2019. Evolution of all-atom protein force fields to improve local and global properties. *J. Phys. Chem. Lett.* 10:2227–2234.

41. Vitalis, A., and R. V. Pappu. 2009. ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem.* 30:673–699.

42. Amin, A. N., Y. H. Lin, …, H. S. Chan. 2020. Analytical theory for sequence-specific binary fuzzy complexes of charged intrinsically disordered proteins. *J. Phys. Chem. B.* 124:6709–6720.

43. Saunders, M. G., and G. A. Voth. 2013. Coarse-graining methods for computational biology. *Annu. Rev. Biophys.* 42:73–93.

44. Das, S., A. N. Amin, …, H. S. Chan. 2018. Coarse-grained residue-based models of disordered protein condensates: utility and limitations of simple charge pattern parameters. *Phys. Chem. Chem. Phys.* 20:28558–28574.

45. Choi, J. M., F. Dar, and R. V. Pappu. 2019. LASSI: a lattice model for simulating phase transitions of multivalent proteins. *PLoS Comput. Biol.* 15:e1007028.

46. Ruff, K. M., R. V. Pappu, and A. S. Holehouse. 2019. Conformational preferences and phase behavior of intrinsically disordered low complexity sequences: insights from multiscale simulations. *Curr. Opin. Struct. Biol.* 56:1–10.

47. Dignon, G. L., W. Zheng, and J. Mittal. 2019. Simulation methods for liquid-liquid phase separation of disordered proteins. *Curr. Opin. Chem. Eng.* 23:92–98.

48. Dignon, G. L., W. Zheng, …, J. Mittal. 2018. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.* 14:e1005941.

49. Sieradzan, A. K., M. Bogunia, …, M. Makowski. 2019. Introduction of phosphorylated residues into the UNRES coarse-grained model: toward modeling of signaling processes. *J. Phys. Chem. B.* 123:5721–5729.

50. Shen, T., C. Zong, …, P. G. Wolynes. 2005. The folding energy landscape and phosphorylation: modeling the conformational switch of the NFAT regulatory domain. *FASEB J.* 19:1389–1395.

51. Das, S., Y.-H. Lin, …, H. S. Chan. 2020. Comparative roles of charge, π, and hydrophobic interactions in sequence-dependent phase separation of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA.* 117:28795–28805.

52. Choi, K. J., P. S. Tsoi, …, A. C. M. Ferreon. 2018. A chemical chaperone decouples TDP-43 disordered domain phase separation from fibrillation. *Biochemistry.* 57:6822–6826.

53. Rhoads, S. N., Z. T. Monahan, …, F. P. Shewmaker. 2018. The role of post-translational modifications on prion-like aggregation and liquid-phase separation of FUS. *Int. J. Mol. Sci.* 19:886.

54. Saito, M., D. Hess, …, P. Matthias. 2019. Acetylation of intrinsically disordered regions regulates phase separation. *Nat. Chem. Biol.* 15:51–61.

55. Ashbaugh, H. S., and H. W. Hatch. 2008. Natively unfolded protein stability as a coil-to-globule transition in charge/hydropathy space. *J. Am. Chem. Soc.* 130:9536–9542.

56. Creighton, T. E. 1993. Proteins: Structures and Molecular Properties. Helvetian Press.

57. Kapcha, L. H., and P. J. Rossky. 2014. A simple atomic-level hydrophobicity scale reveals protein interfacial structure. *J. Mol. Biol.* 426:484–498.

58. Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314:141–151.

59. Hansmann, U. H. E. 1997. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* 281:140–150.

60. Plimpton, S. 1995. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* 117:1–19.

61. Blas, F. J., L. G. MacDowell, …, G. Jackson. 2008. Vapor-liquid interfacial properties of fully flexible Lennard-Jones chains. *J. Chem. Phys.* 129:144703.

62. Silmore, K. S., M. P. Howard, and A. Z. Panagiotopoulos. 2017. Vapour-liquid phase equilibrium and surface tension of fully flexible Lennard-Jones chains. *Mol. Phys.* 115:320–327.

63. Dignon, G. L., W. Zheng, …, J. Mittal. 2018. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA.* 115:9929–9934.

64. Schuster, B. S., E. H. Reed, …, D. A. Hammer. 2018. Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. *Nat. Commun.* 9:2985.

65. Schuster, B. S., G. L. Dignon, …, J. Mittal. 2020. Identifying sequence perturbations to an intrinsically disordered protein that determine its phase-separation behavior. *Proc. Natl. Acad. Sci. USA.* 117:11421–11431.

66. Anderson, J. A., C. D. Lorenz, and A. Travesset. 2008. General purpose molecular dynamics simulations fully implemented on graphics processing units. *J. Comput. Phys.* 227:5342–5359.

67. Flory, P. J. 1949. Statistical mechanics of dilute polymer solutions. *J. Chem. Phys.* 17:1347.

68. Hofmann, H., A. Soranno, …, B. Schuler. 2012. Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. USA.* 109:16155–16160.

69. Sawle, L., and K. Ghosh. 2015. A theoretical method to compute sequence dependent configurational properties in charged polymers and proteins. *J. Chem. Phys.* 143:085101.

70. Singer, A. U., and J. D. Forman-Kay. 1997. pH titration studies of an SH2 domain-phosphopeptide complex: unusual histidine and phosphate pKa values. *Protein Sci.* 6:1910–1919.

71. Cragnell, C., E. Rieloff, and M. Skepö. 2018. Utilizing coarse-grained modeling and Monte Carlo simulations to evaluate the conformational ensemble of intrinsically disordered proteins and regions. *J. Mol. Biol.* 430:2478–2492.

72. Das, R. K., and R. V. Pappu. 2013. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. USA.* 110:13392–13397.

73. Brady, O. A., P. Meng, …, F. Hu. 2011. Regulation of TDP-43 aggregation by phosphorylation and p62/SQSTM1. *J. Neurochem.* 116:248–259.

74. Dephoure, N., K. L. Gould, …, D. R. Kellogg. 2013. Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. *Mol. Biol. Cell.* 24:535–542.

75. Oulhen, N., S. Boulben, …, B. Cosson. 2009. A variant mimicking hyperphosphorylated 4E-BP inhibits protein synthesis in a sea urchin cell-free, cap-dependent translation system. *PLoS One.* 4:e5070.

76. Li, M., J. Luo, …, W. Gu. 2002. Acetylation of p53 inhibits its ubiquitination by Mdm2. *J. Biol. Chem.* 277:50607–50611.

77. Alaei, S. R., C. K. Abrams, …, M. M. Freidin. 2018. Acetylation of C-terminal lysines modulates protein turnover and stability of Connexin-32. *BMC Cell Biol.* 19:22.

78. Zheng, W., G. Dignon, …, J. Mittal. 2020. Hydropathy patterning complements charge patterning to describe conformational preferences of disordered proteins. *J. Phys. Chem. Lett.* 11:3408–3415.

79. Lin, Y.-H., and H. S. Chan. 2017. Phase separation and single-chain compactness of charged disordered proteins are strongly correlated. *Biophys. J.* 112:2043–2046.

80. Firman, T., and K. Ghosh. 2018. Sequence charge decoration dictates coil-globule transition in intrinsically disordered proteins. *J. Chem. Phys.* 148:123305.

81. Lin, Y. H., J. P. Brady, …, H. S. Chan. 2017. Charge pattern matching as a "fuzzy" mode of molecular recognition for the functional phase separations of intrinsically disordered proteins. *New J. Phys.* 19:115003.

82. Statt, A., H. Casademunt, …, A. Z. Panagiotopoulos. 2020. Model for disordered proteins with strongly sequence-dependent liquid phase behavior. *J. Chem. Phys.* 152:075101.

83. Arribas-Layton, M., J. Dennis, …, J. Lykke-Andersen. 2016. The C-terminal RGG domain of human Lsm4 promotes processing body formation stimulated by arginine dimethylation. *Mol. Cell. Biol.* 36:2226–2235.

84. Matsumoto, K., H. Nakayama, …, M. Tsujimoto. 2012. PRMT1 is required for RAP55 to localize to processing bodies. *RNA Biol.* 9:610–623.