Biophysical Journal, Volume *119*

Supplemental Information

Phase Separation of Toxic Dipeptide Repeat Proteins Related to C9orf72 ALS/FTD

Hamidreza Jafarinia, Erik van der Giessen, and Patrick R. Onck

Table of Contents

Coarse-grained force field

Update of the 1BPA coarse-grained force field. The 1BPA implicit solvent force field has been used before to study intrinsically disordered FG-Nups and nucleocytoplasmic transport (1-3). The bonded interactions, i.e. bending and torsion potentials, in this force field are residue and sequence specific. This force field, interestingly, differentiates between the bending and torsion potentials of Glycine, Proline, and other residues. This feature is highly important since DPRs are rich in Proline and Glycine and it has been shown that these two residues contribute to the rigidity and flexibility of an IDP (4-7). Therefore, using the 1BPA model enables us to distinguish between the properties of poly-PR and poly-GR and to obtain more accurate results for poly-GA. The 1BPA also accounts for the hydrophobic/hydrophilic and electrostatic interactions between different amino acids, polarity of the solvent and screening of free ions. This force field uses the average of several residue-based hydrophobicity scales to describe the effective interactions between the amino acids. The hydrophobic interactions in this force field have been calibrated against the experimentally known R_{h} values of FG-Nup segments (8). For Proline and Glycine the interaction parameters have also been fine-tuned using the end-to-end distance and radius of gyration of poly-Proline (9) and poly-Glycine segments (10) which makes the 1BPA a proper choice for investigating the properties of DPRs.

The majority of FG-Nup segments, however, contain less than 0.6% of Arginine (R) (see the pink shaded band in Fig. S1). For the ones with more than 0.6% of R a correlation can be observed between the R content and the R_{h} error (see black dashed line in Fig. S1a) showing that there is still room for improving the R interaction with other residues in the 1BPA force field. To achieve this we further fine-tune the relative hydrophobic strength ε_i of R. The ε_i value is a residue-specific parameter that ranges between 0 and 1 and is close to 0 for hydrophilic polar residues. Since we also study the interaction of R-DPRs with acidic molecules, we recalibrate the ε_i values of all charged residues, i.e. RDEK. The aim is to obtain an updated 1BPA force field that is more accurate for studying the properties of R-DPRs and their interaction with negatively-charged molecules.

To update the 1BPA force field we slightly increase the ε_i values of charged residues to 0.005 (see Table S1), thus reducing the R_{h} error for all the six FG-Nup segments with R $\mathrm{content} > 0.6\%$ (see Fig. S1a). This choice of parameter for the relative hydrophobic strength of charged residues gives the best results in terms of the total average error and the minimum largest error in our calibration simulations with 16 FG-Nup segments presented here and originally used for the calibration of 1BPA (2). The total average and the largest errors are found to be 8.3% and 21.1% in the 1BPA force field, and 7.5% and 17.1% in the updated 1BPA force field. The correlation mentioned earlier for an R $\mathrm{content} > 0.6\%$ still exists in the updated 1BPA (red dashed line in Fig. S1a) which might be due to the absence of cation-π interactions between R and residues with aromatic rings in our force field. However, this has no effect on our simulation results since the DPRs and acidic molecules studied in this work contain no aromatic residues. A direct comparison between the two force fields is presented in Fig. S1b. At physiological

intracellular pH between 7 and 7.4, i.e more than three pH units away from the pK_a values of Arginine and Aspartic acid, we assume R, D, E, and K to be fully charged (11).

Complex coacervation of R-DPRs. The complex coacervation of polyelectrolytes is driven by a combination of enthalpic and entropic effects (12). Coulombic energy change and counterion release entropy are the main contributors to the free energy of complexation (13). In our single-molecule and phase separation simulations of R-DPRs with stretches of acidic amino acids, we account for the screening effect of ions, but similar to previous theoretical (14, 15) and coarse-grained models (16, 17) used to study complex coacervation, the effect of counterion condensation has not been considered in our modeling. Despite this limitation, our simulations capture the experimentally observed length-dependence of ρ_{L} (concentration of the dilute phase) and ρ_{H} (concentration of the condensed phase) for polyelectrolytes (18). Here we compare the effect of Coulomb energy change and counterion release entropy for the complexation of R-DPRs with acidic molecules by calculating the Coulomb strength parameter as suggested by Ou and Muthukumar (13). In their study the Coulomb strength parameter Γ has been defined as

$$\Gamma = \frac{l_{\mathrm{B}}}{l_0}.$$

Here $l_{\mathrm{B}} = e^2/4\pi\epsilon_0\epsilon_r k_{\mathrm{B}}T$ is the Bjerrum length, where e is the elementary charge, ϵ_0 is the vacuum permitivity, ϵ_r is the relative permitivity, k_{B} is the Boltzmann constant, T is the absolute temperature and l_0 is the charge separation distance along a polymer chain. For $\Gamma < 1$ the entropic term is negligible and the complexation is driven by the change in the Coulomb energy, but for $\Gamma > 1$ the entropic term starts to play a more important role, for $\Gamma < 1.5$ the electrostatic attraction between the oppositely-charged polyelectrolytes still predominantly drives the complexation, while the counterion release entropy only plays a subsidiary role. For $\Gamma > 1.5$, the contribution of the entropic term becomes more significant, until at $\Gamma = 2.5$ the complexation is completely driven by counterion release entropy (13).

The charge separation distance l_0 is 0.76 nm (two times the bond length) for R-DPRs, and 0.38 nm (the bond length) for acidic molecules. Using the average $l_0 = (0.76 + 0.38)/2 = 0.57$ and $l_{\mathrm{B}} \simeq 0.7$ nm for water at 300 K, the Coulomb interaction parameter is found to be $\Gamma = 1.23$ that lies in the range where the Coulomb energy change plays a more significant role than the counterion release entropy, supporting our assumption to neglect the effect of counterion condensation.

Simulations

Single-molecule simulations. Langevin dynamics simulations are performed at 300 K and physiological salt concentration of 150 mM in NVT ensembles with a time-step of 0.02 ps and a Langevin friction coefficient of 0.02 ps^{-1} using GROMACS version 2016. Each simulation is performed for 3 μs and the last 1 μs is used to obtain the ρ_{L} and R_{g}. The Hydro++ program (19) is used to obtain

the R_{h} values from the trajectories. The fits for R_{h} in Fig. 1b are according to bN^{ν}. For R_{g} we use the following equation for the fits in Fig. 1c (20):

$$R_{\mathrm{g}} = \sqrt{\frac{2l_{\mathrm{p}}^{*}b}{(2\nu+1)(2\nu+2)}}N^{\nu},$$

Where $b = 0.38\,\mathrm{nm}$ and $l_{\mathrm{p}}^{*} = 0.40 \pm 0.07\,\mathrm{nm}$. The errors represent the changes in the scaling exponents for $0.33 < l_{\mathrm{p}}^{*} < 0.47$.

The asphericity in Fig. 1d is calculated using the following equation(21):

$$\mathrm{Asphericity} = 1 - 3\left\langle\frac{\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1}{(\lambda_1+\lambda_2+\lambda_3)^2}\right\rangle$$

Where $\lambda_{1,2,3}$ are the eigenvalues of the gyration tensor. Asphericity is zero for a perfect sphere and is 1 for a perfect rod.

Droplet simulations and cluster size distribution analysis. Although the slab method can minimize the finite size effects and produce reliable values for the concentrations of the two phases, this method is not suitable for investigating the properties of the droplets such as the nucleation process, droplet size and interaction between droplets. Due to these limitations we also study the 3D droplet formation of DPRs. To perform droplet simulations the DPRs are placed in a cubic box of size 80 nm and then simulated for $\simeq 3\ \mu s$ which is sufficient to reach the equilibrium state (Figs. 2a left panels, 3b, S4-S8). At equilibrium the number of the residues inside the condensed phase (droplet) and the radial density profile measured from the center of the droplet (e.g. Fig. S6) are well-converged. In our test simulations we observed no significant effect of the initial distribution of the molecules on the properties of the resulting droplets at equilibrium. The radial density profiles (Figs. 3c, S6) and cluster size distribution plots (Figs. S4a, S5) are time-averaged for at least 1 μs at equilibrium. Discrete cells of thickness 1 nm are used to obtain the radial density profiles. From the radial density profiles, one can obtain values of ρ_{H} from the flat region close to the center of the droplet, and ρ_{L} from the region far from the droplet. However, to minimize the finite-size effects (22), we use slab method (see the next subsection 'Slab simulations') to obtain ρ_{H} and ρ_{L}. The droplet radius is obtained from the radial density profile and is the average position of the first point close to the low-density region with $|d\rho/dr| \leq\ 2-5\ \mathrm{mg/ml/nm}$ (see dashed lines in Figs. 3c, S6). This term is the absolute value of the slope of the density profile which is close to zero where the density profile reaches the dilute phase region. Note that using $|d\rho/dr| \simeq 0\ \mathrm{mg/ml/nm}$ gives unrealistically large values for the droplet radius. We performed a sensitivity analysis for the critical value of $|d\rho/dr|$ and observed that selecting slope limits between 2-5 mg/ml/nm results in a maximal change of 4% for all computed values for the droplet radius.

To generate the cluster size distribution plots in Figs. S4 and S5, two chains are considered to be in the same cluster if at least two residues of those chains come closer than 0.7 nm (23). In our cluster size distribution plots (Fig. S4a and S5) the horizontal axis is the logarithm of the number of residues inside a cluster (S) and the vertical axis is the logarithm of the time-averaged number of the clusters (N_c). When phase separation occurs the curves are divided into two regions, a dilute phase containing free molecules and small clusters, and a condensed phase that exchanges molecules with the dilute phase.

Slab simulations. The initial simulations are performed in cubic boxes of size 20 nm for poly-GA and 25 nm for more extended R-DPRs. These box sizes ensure no interaction between DPRs and their periodic images for repeat lengths $n \leq 100$. For longer DPRs the box size is increased accordingly. For the initial equilibration simulations we followed the steps suggested in (24). After equilibration, the box is enlarged to a 10 times larger size than its initial value in the z direction which is sufficient to reduce the finite-size effects and to obtain reliable values for ρ_{H} and ρ_{L}. The system is then simulated for $\simeq$ 3 μs in an NVT ensemble to achieve convergence for the density profiles in the z direction (Figs. 2b, S11, S12). The density profiles are calculated using discrete cells of thickness 1 nm and time-averaged for at least 1 μs at equilibrium. When the system undergoes phase separation, the averaged concentrations in $|z| < 4$ nm region is used to obtain ρ_{H}. To obtain ρ_{L} we use the average concentrations in $|z| > 40$ nm for poly-GA (Fig. 2b) and $|z| > 65$ nm for R-DPRs (Figs. S11, S12). The simulation parameters are similar to the ones we used in the single-molecule and droplet simulations. From the slab simulations we also obtain the time-averaged exchange rates based on the fluctuations of the number of polymer units inside the condensed phases at equilibrium for around 1 μs (Table S1). To obtain the number of the molecules inside the condensed phases, we use cut-offs of 0.7 (23) and 0.9 nm for poly-GA and poly-PR+poly-D. We use a larger cut-off value for the second case since the equilibrium distances between the oppositely-charged residues, determined by both electrostatic and hydrophilic interactions, are almost 0.2 nm larger than the one for non-charged residues in our coarse-grained force field (2). To obtain the exchange rates per unit area, we divide the exchange rate by 2 times the xy-plane area of the slab box.

Phase diagrams. The vertical axis in our phase diagram is the DPR repeat length and the horizontal axis is the concentration (Figs. 2a, 4a, S12). The phase diagram is obtained by connecting values of ρ_{L} and ρ_{H} computed from the slab simulations. To find the critical point, we begin with the smallest repeat length n_1 that has produced converged ρ_{L} and ρ_{H} values. Then we perform slab simulation for repeat length of $n_{-1} = n_1 - 10$ to calculate the time-averaged density profile for 1 μs after 3 μs of simulation time. If the calculated density profile for repeat length n_{-1} is almost flat, with small fluctuations in concentration $\Delta\rho = |\rho_{\max} - \rho_{\min}| < 20$ mg/ml, we report $(n_1 + n_{-1})/2$ as the critical repeat length. With this method we estimate the critical repeat length with an error of less than 5 repeats. Choosing any value larger than 20 mg/ml for $\Delta\rho$ does not change the critical repeat length. In each phase diagram,

for the region close to the critical point, a dashed spline that reaches its minimum at $(n_1 + n_{-1})/2$, $(\rho_{\mathrm{H}}(n_1) + \rho_{\mathrm{L}}(n_1))/2$ is shown as a guide to the eye (Figs. 2a, 4a, S12).

Potential of mean force (PMF) calculation. We use umbrella sampling simulations and the weighted histogram analysis method (WHAM) via the gmx wham utility of GROMACS to calculate the PMF associated with the binding of R-DPRs to the acidic molecules in Fig. S9. The distance r between the center of masses of two molecules is considered as the reaction coordinate. For each window the simulation is conducted for 1 μs. We use $\Delta r = 0.1$ nm for the distance between two windows. For details about the umbrella sampling method the reader is referred to (3, 25).

Protein sequences

Sequences of the disordered parts of NPM1 and NCL used in Fig. S10 are listed below. The negatively charged residues are shown in blue.

$NPM1_{120\text{-}240}$	EEDAESEDEEEEDVKLLSISGKRSAPGGGSKVP QKKVKLAADEDDDDDDEEDDDEDDDDDDFD DEEAEEKAPVKKSIRDTPAKNAQKSNQNGKDS KPSSTPRSKGQESFKKQEKTPKTPKG
$NCL_{1\text{-}300}$	MVKLAKAGKNQGDPKKMAPPPKEVEEDSEDE EMSEDEEDDSSGEEVVIPQKKGKKAAATSAKK VVVSPTKKVAVATPAKKAAVTPGKKAAATPA KKTVTPAKAVTTPGKKGATPGKALVATPGKK GAAIPAKGAKNGKNAKKEDSDEEEDDDSEEDE EDDEDEDEDEDEIEPAAMKAAAAAPASEDEDD EDDEDDEDDDDDEEDDSEEEAMETTPAKGKK AAKVVPVKAKNVAEDEDEEEDDEDEDDDDDE DDEDDDDEDDEEEEEEEEEEPVKEAPGKRKKE MAKQKAAPEAKKQKVEG

Supplementary figures

Figure S1: Coarse-grained force field

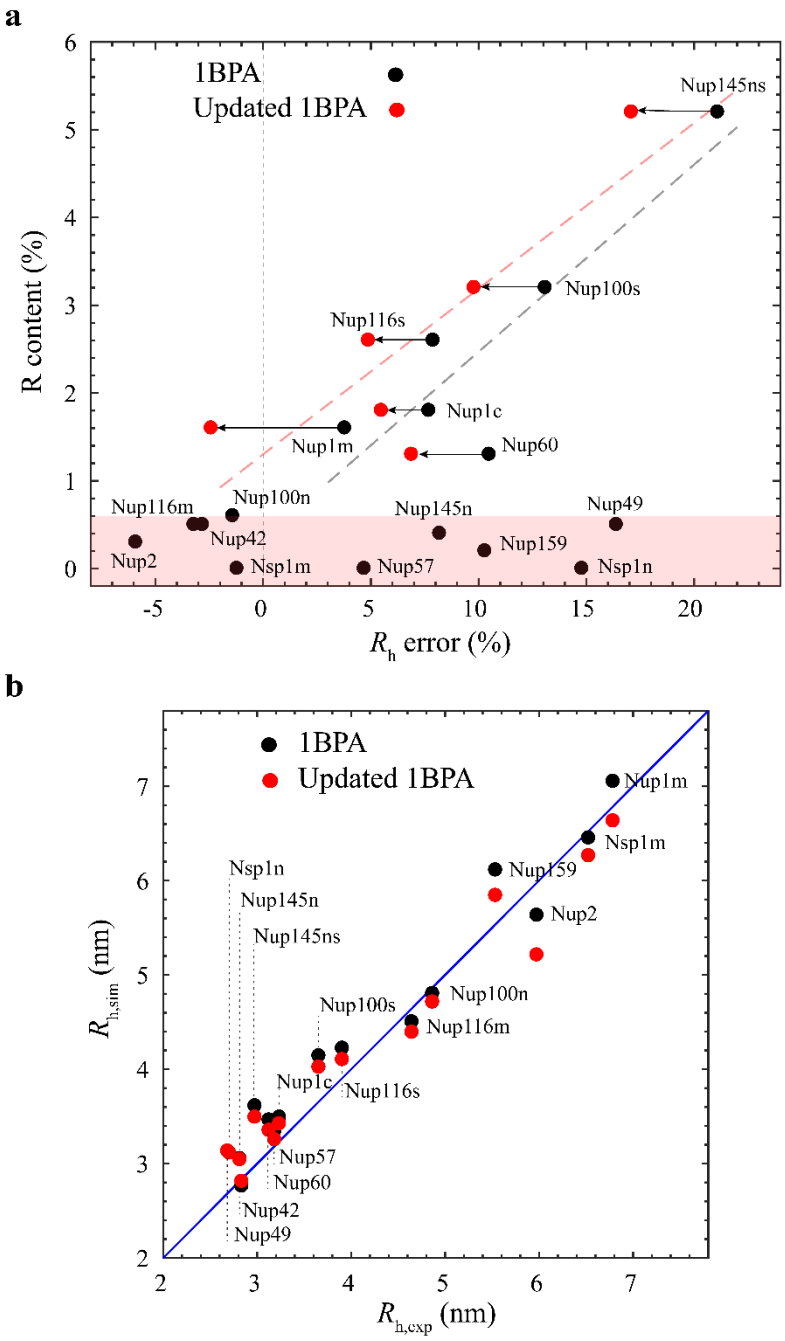

Figure S1: (a) Content of Arginine in FG-Nup segments plotted against their corresponding hydrodynamics radius error: $(R_{\mathrm{h,sim}} - R_{\mathrm{h,exp}})/R_{\mathrm{h,exp}}$ for 1BPA and the updated 1BPA force fields. The pink shaded band contains FG-Nups with R content $< 0.6\%$. The black dashed line shows the correlation between the R content $> 0.6\%$ and the R_{h} error in 1BPA. In the updated 1BPA force field the R_{h} error is reduced for all FG-Nups with R content $> 0.6\%$. The red dashed line shows the correlation between the R content $> 0.6\%$ and the R_{h} error in updated 1BPA. **(b)** A direct comparison

of the two force fields in predicting the hydrodynamic radius of FG-Nups. The total average and the largest errors are found to be 8.3% and 21.1% in the 1BPA force field, and 7.5% and 17.1% in the updated 1BPA force field.

Figure S2: Contribution of the backbone potentials to the size of the R-DPRs

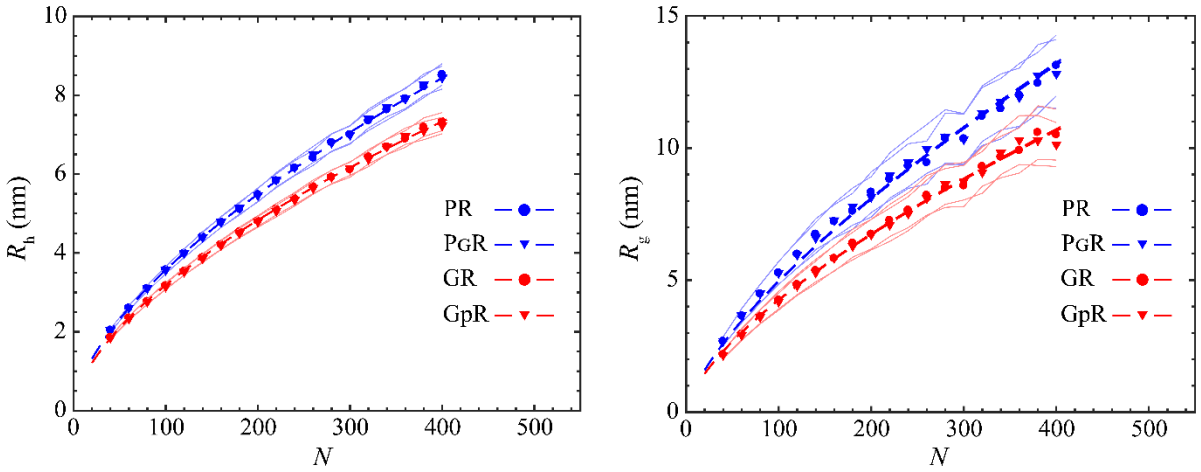

Figure S2: A comparison between the R_h and R_g of poly-PR and -GR (presented in Fig. 1b and 1c) and poly-P_GR and -G_PR. P_G has the backbone rigidity of P and the hydrophobicity of G, G_P has the backbone flexibility of G and the hydrophobicity of P. The error regions between the thin lines indicate half of the standard deviation.

Figure S3: Comparison of the hydrodynamic radius of Proline-Arginine chains

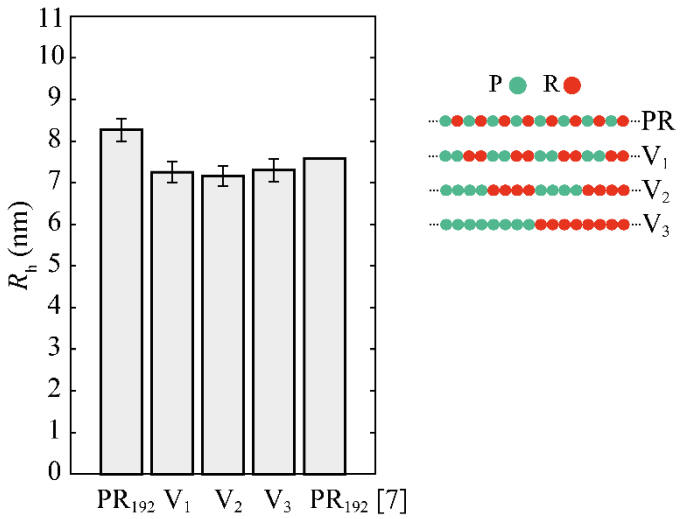

Figure S3: The R_h value of poly-PR and three variants with different patterning of the Proline and Arginine residues. The bar chart at the right has been obtained using the equation reported in (7).

Figure S4: Phase separation of poly-GA

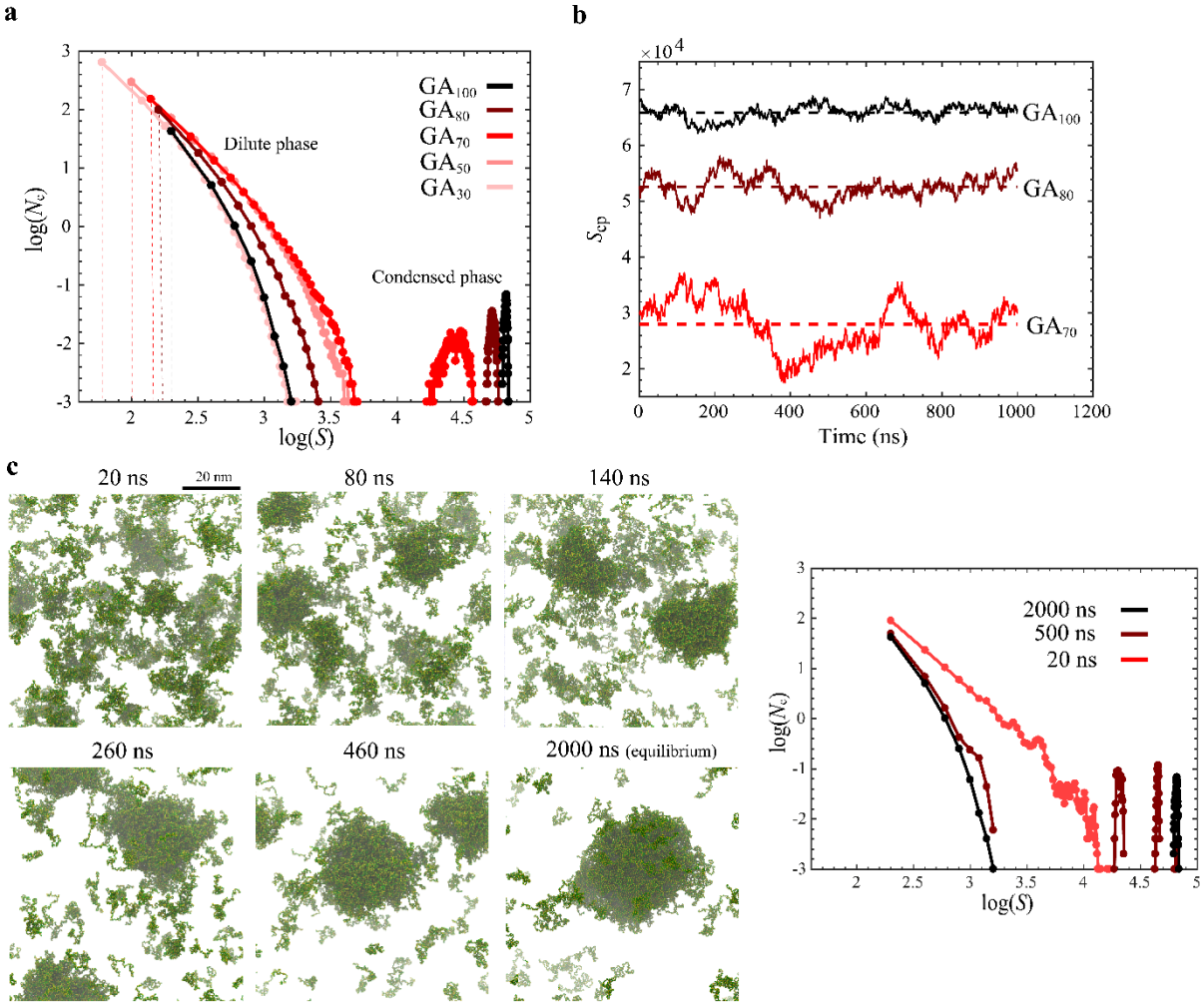

Figure S4: Phase separation of poly-GA for a total concentration of 32.2 mg/ml. **(a)** Cluster size distribution of poly-GA $30 \leq n \leq 100$ at equilibrium. S is the number of residues inside a cluster and N_c is the time-averaged number of the clusters. For comparison, the time-averaged number of free molecules is also included in this plot, indicated with a dashed line for each case. **(b)** The number of residues in the condensed phase (S_{cp}) plotted against time at equilibrium for GA_{100}, GA_{80}, and GA_{70}. Horizontal dashed lines indicate the average values. Longer dipeptides form clusters containing more residues with lower exchange with the surrounding. **(c)** Time evolution for the phase separation of GA_{100} starting from randomly distributed molecules in a cubic box of size 80 nm (left). A similar cluster size distribution analysis as presented in (a) at three different simulation times for GA_{100}. For each data set the average of the last 20 ns is used.

Figure S5: Length- and concentration-dependent phase separation of poly-GA

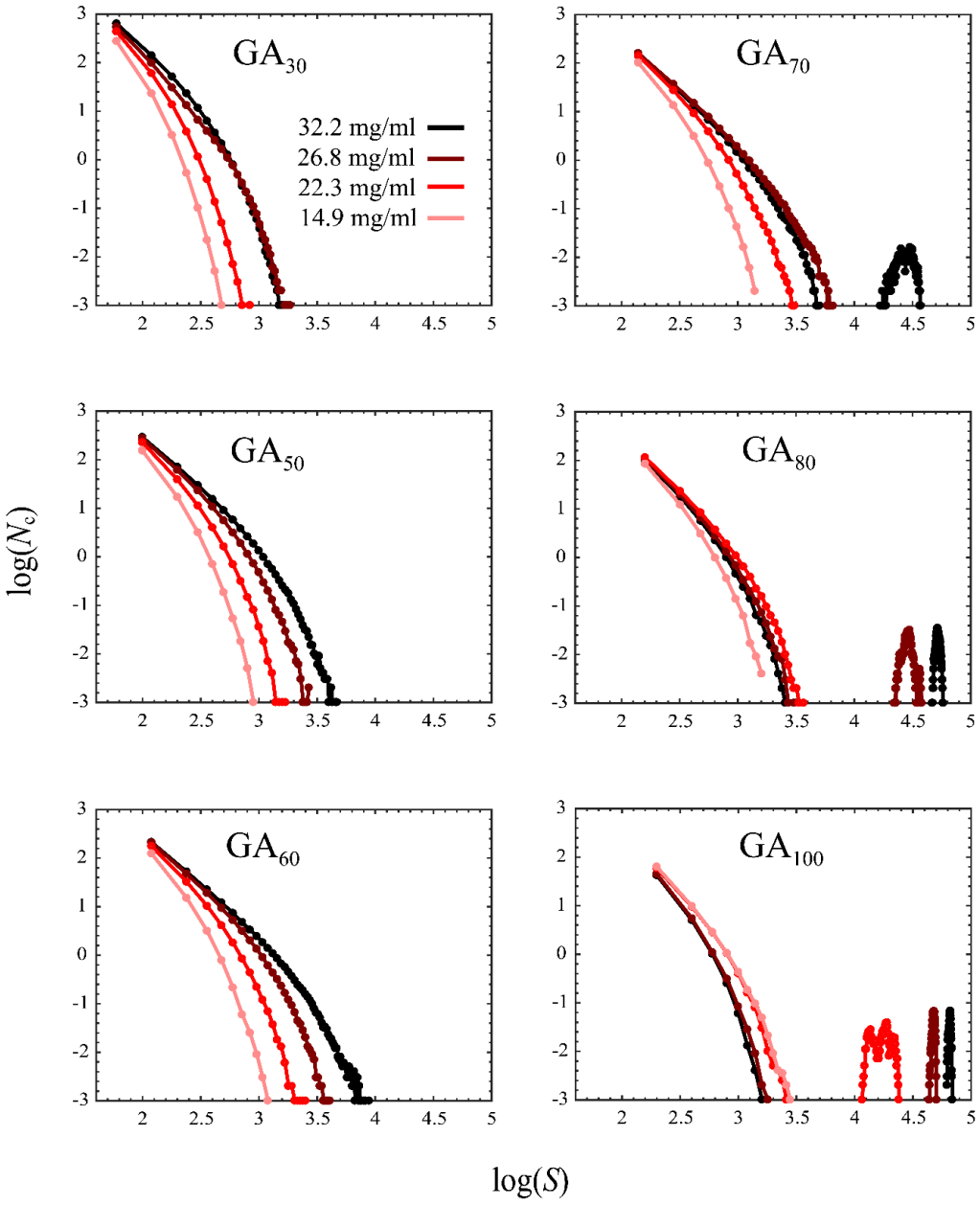

Figure S5: Cluster size distribution analysis of poly-GA at equilibrium for four different total mass concentrations of 32.2, 26.8, 22.3, and 14.9 mg/ml shows a length- and concentration-dependent phase separation. S is the number of residues inside a cluster and N_c is the time-averaged number of the clusters. For a fixed repeat length, increasing the concentration increases the average number of the molecules inside the condensed phase (see the results for GA_{80} and GA_{100}).

Figure S6: Radial density profiles for GA_{100} droplets

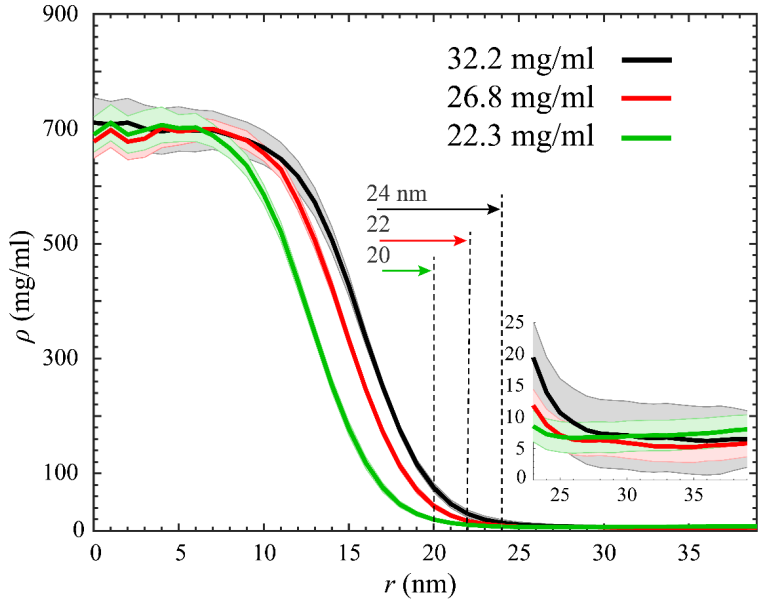

Figure S6: Radial density profiles for GA_{100} droplets for three different total mass concentrations. The shading indicates half of the standard deviations as error bars. The radius of the droplet is shown with dashed lines for each case. The inset figure shows the zoomed density profiles for $r \geq 24$ nm. The size of the droplet increases with increasing the total mass concentration. However, ρ_H (the average concentration for $r < 4$ nm) and ρ_L (the average concentration for $r > 30$ nm) remain unchanged.

Figure S7: Number of free molecules of poly-GA at equilibrium

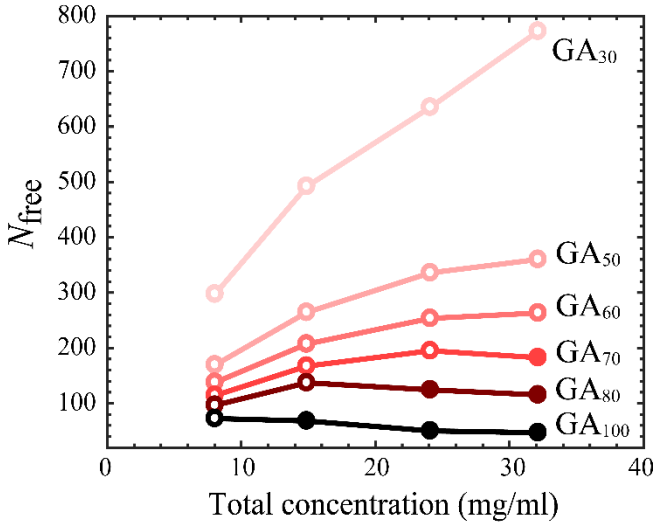

Figure S7: Number of free molecules N_{free} in the dilute phase plotted against the total concentration for different lengths of poly-GA. Filled markers are used when poly-GA molecules undergo phase separation. When phase separation occurs the N_{free} drops. At a fixed concentration, N_{free} is higher for shorter dipeptides. The data in the figure is obtained from the cluster size distribution curves of Fig. S5.

Figure S8: R-DPRs phase separation in the presence of poly-D

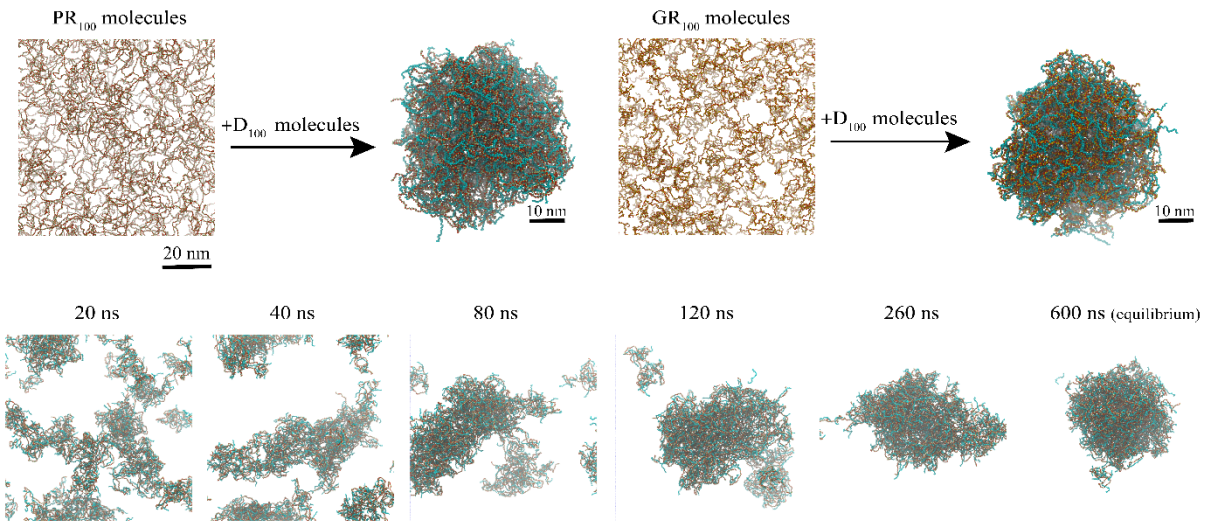

Figure S8: Long chains of poly-PR and poly-GR are not capable of forming clusters. Adding acidic molecules (poly-D) induces the phase separation of R-DPRs (top). Time evolution for the phase separation of $PR_{100} + D_{40}$ starting from randomly distributed molecules (bottom). The concentration ratios and the total concentration are $r_{PR} = r_{GR} = 0.57$ and 14.8 mg/ml for all cases.

Figure S9: PMF calculation for binding of R-DPRs to poly-D

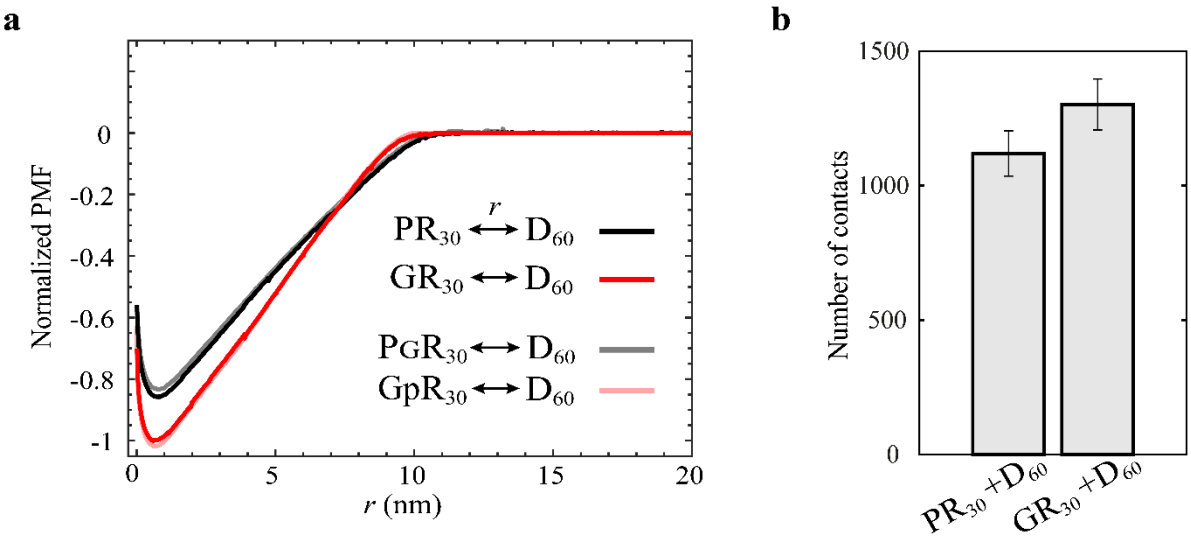

Figure S9: (a) PMF curves for binding of PR_{30} and GR_{30} to D_{60} indicates a larger free energy of binding of poly-GR to acidic molecules. The distance between the center of masses of two molecules is indicated with r. Due to the more compact conformation of poly-GR, its PMF curve vanishes at slightly shorter inter-molecule distances. Results for P_GR_{30} and G_PR_{30} binding to D_{60} show no significant change in the PMF curves. Note that P_G has the backbone rigidity of P and the hydrophobicity of G and G_P has the backbone flexibility of G and the hydrophobicity of P. All curves are normalized with the depth of the GR_{30}–D_{60} binding free energy. **(b)** Time-averaged total number of contacts after binding of PR_{30} and GR_{30} to D_{60} at equilibrium using a cut-off of 2.5 nm. The contact of one residue in the R-DPRs with one residue in the acidic molecule is counted as one contact, see Movie S4 for short trajectories.

Figure S10: NCPR plots and poly-PR binding to the disordered regions of NCL and NPM1

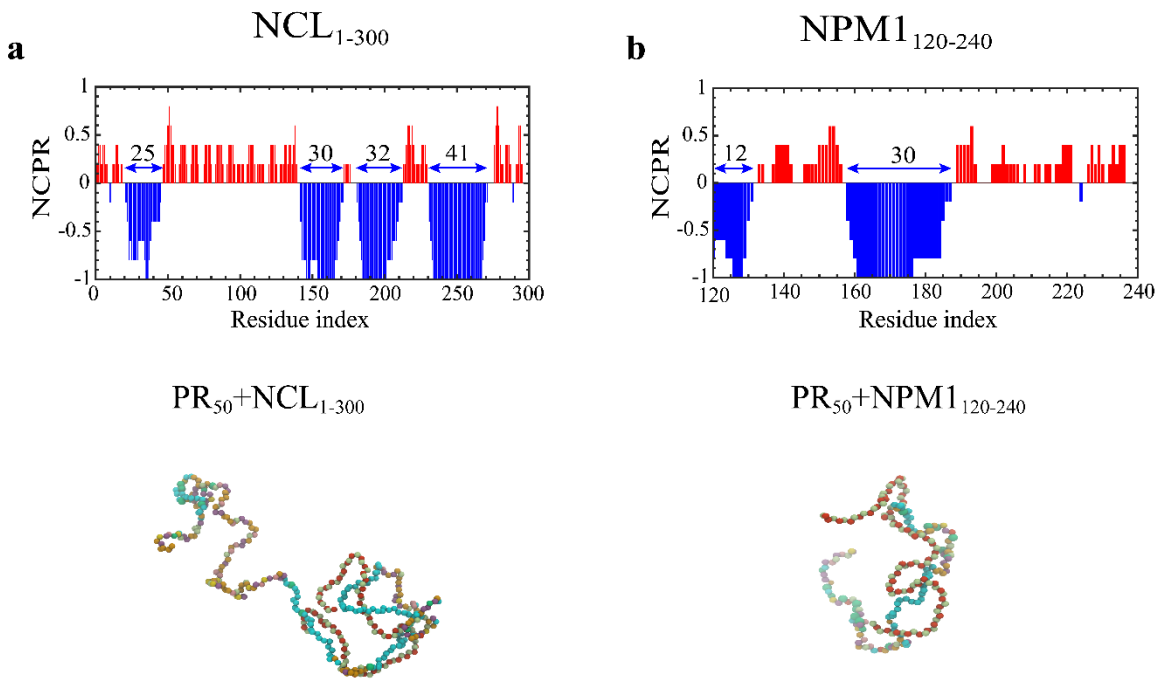

Figure S10: Top: Net charge per residue (NCPR) histograms for disordered parts of two nucleolar proteins: **(a)** nucleolin ($NCL_{1\text{-}300}$) and **(b)** nucleophasmin ($NPM1_{120\text{-}240}$). Blue arrows show acidic tracts with lengths ranging from 12 to 41. To find NCPR we use a sliding window containing 5 residues. Bottom: snapshots of binding of PR50 to $NCL_{1\text{-}300}$ and $NPM1_{120\text{-}240}$. Acidic tracts are indicated in cyan; PR chains comprise of red-green colored beads (as in Fig. 1a). The other aminoacids are given a range of colors according to their aminoacid type.

Figure S11: Slab density profiles for phase separation of poly-PR with poly-D

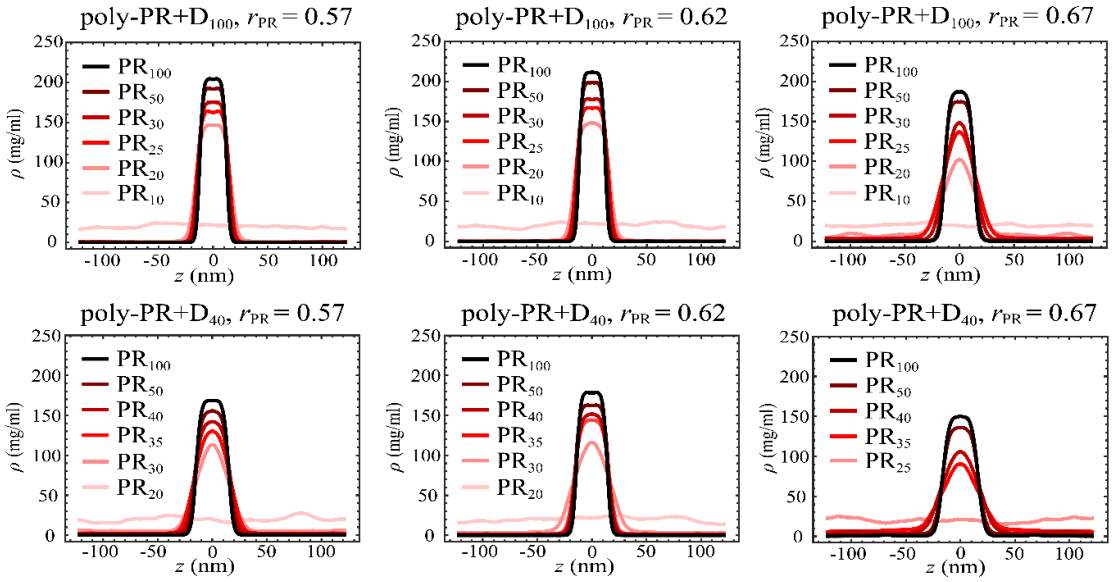

Figure S11: Slab density profiles for phase separation of poly-PR with acidic molecules of lengths 40 and 100 for three different concentration ratios of $r_{\mathrm{PR}} = 0.57, 0.62$ and 0.67. These density profiles are used to obtain the coexistence phase diagrams in Fig. 4a.

Figure S12: Comparison of the phase diagrams of poly-PR and poly-GR

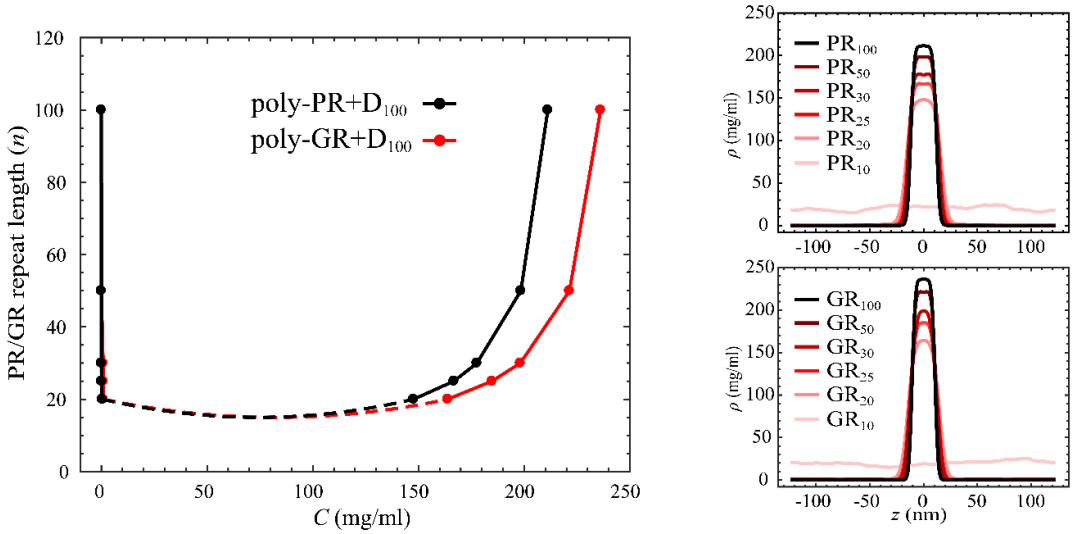

Figure S12: Coexistence phase diagram (left) and the corresponding slab density profiles (right) for phase separation of poly-PR and poly-GR with D_{100} for $r_{PR} = r_{GR} = 0.62$. With the same repeat lengths, poly-GR forms condensed phases with higher concentrations.

Supplementary tables

Table S1: The relative hydrophobic strength values of charged residues

Amino acid	R	D	E	K
ε_i (1BPA)	0	0.0005	0.0005	0.0005
ε_i (updated 1BPA)	0.005	0.005	0.005	0.005

Table S1: The relative hydrophobic strength ε_i values of charged residues in the 1BPA force field (1, 2) and updated 1BPA force field (the current paper).

Table S2: Time-averaged exchange rates

Poly-GA repeat length	Exchange rate (ns^{-1})	Exchange rate per unit area $\times 10^4$ ($ns^{-1}nm^{-2}$)	Poly-PR repeat length	Exchange rate (ns^{-1})	Exchange rate per unit area $\times 10^4$ ($ns^{-1}nm^{-2}$)
55	4.01	50.19	30	4.71	37.72
60	3.33	41.69	35	2.38	19.04
70	1.63	20.32	40	1.56	12.47
80	0.94	11.76	50	0.43	3.41
100	0.37	4.63	100	0.02	0.14

Table S2: The exchange rates between the condensed and dilute phases of poly-GA and poly-PR with different repeat lengths using the slab simulations presented in Fig. 2 and Fig. 4. The details are provided in the section 'Simulations' of the supplementary information. For poly-PR the exchange rates are reported for poly-PR phase separation with D_{40} and $r_{\mathrm{PR}} = 0.62$.

Supplementary movies

Movie S1: Phase separation of GA_{70} (left), GA_{100} (middle) and GA_{140} (right). For better visualization, in each simulation, 10 random molecules of poly-GA are indicated in red.

Movie S2: Single-molecule simulation of PR_{20} (left), GR_{20} (middle), and GA_{20} (right).

Movie S3: Fusion of two liquid droplets formed by PR_{100} and D_{40} with a poly-PR concentration ratio of $r_{\mathrm{PR}} = 0.57$.

Movie S4: Binding of PR_{30} (left) and GR_{30} (right) to D_{60}.

Supporting references

1. Ghavami, A., E. van der Giessen, and P. R. Onck. 2013. Coarse-Grained Potentials for Local Interactions in Unfolded Proteins. J. Chem. Theory Comput. 9(1):432-440.

2. Ghavami, A., L. M. Veenhoff, E. Van Der Giessen, and P. R. Onck. 2014. Probing the disordered domain of the nuclear pore complex through coarse-grained molecular dynamics simulations. Biophys. J. 107(6):1393-1402.

3. Ananth, A. N., A. Mishra, S. Frey, A. Dwarkasing, R. Versloot, E. van der Giessen, D. Görlich, P. Onck, and C. Dekker. 2018. Spatial structure of disordered proteins dictates conductance and selectivity in nuclear pore complex mimics. eLife 7:1-24.

4. Cheng, S., M. Cetinkaya, and F. Gräter. 2010. How sequence determines elasticity of disordered proteins. Biophys. J. 99(12):3863-3869.

5. Martin, E. W., A. S. Holehouse, C. R. Grace, A. Hughes, R. V. Pappu, and 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(47):15323-15335.

6. Wang, J., J. M. Choi, A. S. Holehouse, H. O. Lee, X. Zhang, M. Jahnel, S. Maharana, R. Lemaitre, A. Pozniakovsky, D. Drechsel, I. Poser, R. V. Pappu, S. Alberti, and A. A. Hyman. 2018. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell 174(3):688-699.e616.

7. Marsh, J. A., and J. D. Forman-Kay. 2010. Sequence Determinants of Compaction in Intrinsically Disordered Proteins. Biophys. J. 98(10):2383-2390.

8. Yamada, J., J. L. Phillips, S. Patel, G. Goldfien, A. Calestagne-Morelli, H. Huang, R. Reza, J. Acheson, V. V. Krishnan, S. Newsam, A. Gopinathan, E. Y. Lau, M. E. Colvin, V. N. Uversky, and M. F. Rexacha. 2010. A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. Mol. Cell. Proteomics 9(10):2205-2224.

9. Schuler, B., E. A. Lipman, P. J. Steinbach, M. Kumkell, and W. A. Eaton. 2005. Polyproline and the "spectroscopic ruler" revisited with single-molecule flourescence. Proc. Natl. Acad. Sci. U.S.A. 102(8):2754-2759.

10. Ohnishi, S., H. Kamikubo, M. Onitsuka, M. Kataoka, and D. Shortle. 2006. Conformational Preference of Polyglycine in Solution to Elongated Structure. J. Am. Chem. Soc. 128(50):16338-16344.

11. Perry, S. L., L. Leon, K. Q. Hoffmann, M. J. Kade, D. Priftis, K. A. Black, D. Wong, R. A. Klein, C. F. Pierce, K. O. Margossian, J. K. Whitmer, J. Qin, J. J. De Pablo, and M. Tirrell. 2015. Chirality-selected phase behaviour in ionic polypeptide complexes. Nat. Commun. 6(1):1-8.

12. Overbeek, J. T. G., and M. J. Voorn. 1957. Phase separation in polyelectrolyte solutions. Theory of complex coacervation. J. Cell. Comp. Physiol. 49(S1):7-26.

13. Ou, Z., and M. Muthukumar. 2006. Entropy and enthalpy of polyelectrolyte complexation: Langevin dynamics simulations. J. Chem. Phys. 124(15):154902-154902.

14. Lin, Y. H., J. P. Brady, J. D. Forman-Kay, and H. S. Chan. 2017. Charge pattern matching as a 'fuzzy' mode of molecular recognition for the functional phase separations of intrinsically disordered proteins. New Journal of Physics 19(11):115003-115003.

15. Lin, Y. H., and H. S. Chan. 2017. Phase Separation and Single-Chain Compactness of Charged Disordered Proteins Are Strongly Correlated. Biophys. J. 112(10):2043-2046.

16. Das, S., A. N. Amin, Y. H. Lin, and H. S. Chan. 2018. Coarse-grained residue-based models of disordered protein condensates: Utility and limitations of simple charge pattern parameters. PCCP 20(45):28558-28574.

17. Boeynaems, S., A. S. Holehouse, V. Weinhardt, D. Kovacs, J. Van Lindt, C. Larabell, L. V. D. Bosch, R. Das, P. S. Tompa, R. V. Pappu, and A. D. Gitler. 2019. Spontaneous driving forces give rise to protein−RNA condensates with coexisting phases and complex material properties. Proc. Natl. Acad. Sci. U.S.A. 116(16):7889-7898.

18. Spruijt, E., A. H. Westphal, J. W. Borst, M. A. Cohen Stuart, and J. van der Gucht. 2010. Binodal Compositions of Polyelectrolyte Complexes. Macromolecules 43(15):6476-6484.

19. Garcia de la Torre, J., S. Navarro, M. C. Lopez Martinez, F. G. Diaz, and J. J. Lopez Cascales. 1994. HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules. Biophys. J. 67(2):530-531.

20. Hofmann, H., A. Soranno, A. Borgia, K. Gast, D. Nettels, and B. Schuler. 2012. Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. Proc. Natl. Acad. Sci. U.S.A. 109(40):16155-16160.

21. Mao, A. H., S. L. Crick, A. Vitalis, C. L. Chicoine, and R. V. Pappu. 2010. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. Proc. Natl. Acad. Sci. U.S.A. 107(18):8183-8188.

22. Dignon, G. L., W. Zheng, R. B. Best, Y. C. Kim, and J. Mittal. 2018. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. Proceedings of the National Academy of Sciences 115(40):9929.

23. Ghavami, A., E. Van der Giessen, and P. R. Onck. 2018. Sol–gel transition in solutions of FG-Nups of the nuclear pore complex. Extreme Mechanics Letters 22:36-41.

24. Dignon, G. L., W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal. 2018. Sequence determinants of protein phase behavior from a coarse-grained model. PLoS Comp. Biol. 14(1):e1005941-e1005941.

593 25. Ghavami, A., E. van der Giessen, and P. R. Onck. 2016. Energetics of Transport through the
594 Nuclear Pore Complex. PLoS One 11(2):e0148876-e0148876.

595