An Intermediate Resolution Model of RNA Dynamics and Phase Separation with Explicit Mg^{2+}

Shanlong Li and Jianhan Chen*

Department of Chemistry

University of Massachusetts, Amherst, MA 01003, USA

* To whom correspondence should be addressed. Tel: 413-545-3386; Email: jianhanc@umass.edu

ABSTRACT

RNAs are major drivers of phase separation in the formation of biomolecular condensates. Recent studies suggest that RNAs can also undergo protein-free phase separation in the presence of divalent ions or crowding agents. Much remains to be understood regarding how the complex interplay of base stacking, base pairing, electrostatics, ion interactions, and structural propensities governs the phase behaviour of RNAs. Here we develop an **i**ntermediate resolution model for **con**densates of **RNA**s (iConRNA) that can capture key local and long-range structure features of dynamic RNAs and simulate their spontaneous phase transitions in the presence of Mg^{2+}. Representing each nucleotide using 6 or 7 beads, iConRNA considers specific RNA base stacking and pairing interactions and includes explicit Mg^{2+} ions to study Mg^{2+}-induced phase separation. Parametrized using theoretical and experimental data, the model can correctly reproduce the chain properties of A-form helical poly(rA) and coil poly(rU), and essential structures of several RNA hairpins. With an effective Mg^{2+} ion model, iConRNA simulations successfully recapitulate the experimentally observed lower critical solution temperature (LCST)-type phase separation of poly(rA) and the dissolution of poly(rU). Furthermore, the phase diagrams of CAG/CUG/CUU-repeat RNAs correctly reproduce the experimentally observed sequence- and length-dependence of phase separation propensity. These results suggest that iConRNA can be a viable tool for studying homotypic RNA and potentially heterotypic RNA-protein phase separations.

INTRODUCTION

Biomolecular condensates have attracted intensive interest in recent years and are believed to play important roles in myriad cellular functions ranging from RNA storage and processing, stress responses, metabolism, and immune response, to cellular signalling (1-5). They have also been implicated in a variety of human diseases including cancers and neurodegenerative diseases (6-11). Formation of condensates involves spontaneous phase separations, mediated by dynamic and multivalent biological macromolecules such as proteins and nucleic acids (12-14). Interestingly, even though the first biomolecular condensate discovered was germline P granules containing RNA and RNA-binding proteins (1), phase separation of proteins has drawn more intensive research interests in the field. Extensive efforts from experiment, theory, and simulation have provided many insights into the molecular driving forces and sequence-specific phase behaviours as well as the thermodynamic, dynamic, and material properties of the condensates (15-25). In contrast, less attention has been paid to RNA-mediated phase separation (26).

The importance of RNAs in biomolecular condensates such as Ribonucleoprotein (RNP) granules is well recognized (27). On one hand, RNAs can drive the formation and stability of phase-separated condensates. RNA binding proteins (RBPs) interact with RNAs through both specific bindings, medicated by well-structured RNA binding domains such as RNA recognition motifs (RRMs) (28-30), and nonspecific interactions, through positively charged arginine-glycine-glycine (RGG) domains (31-33). These interactions can promote the phase separation through seeding and accelerating condensation of RBPs (34-37), where RNA, especially long noncoding RNA (lncRNA), serves as a scaffold to bind multi-RBPs or directly participates in the phase separation. On the other hand, RNAs have potent effects on controlling material properties of phase-separated droplets, such as viscosity and fluidity (30,31,38,39). For example, the addition of RNA can increase the observed fluidity of LAF-1 condensates (31). Importantly, recent experiments have reported that RNA can undergo protein-free phase separation in the presence of multivalent cations or crowding agents (39-45). In the presence of Mg^{2+} ions, RNA homopolymers, including poly(adenylic acid) (poly(rA)), poly(uridylic acid) (poly(rU)), and poly(cytidylic acid) (poly(rC)), display thermo-responsive phase separations with lower critical solution temperature (LCST) (44). RNA triplet repeats, like CAG or CUG repeats, are able to lead to their coalescence into nuclear foci (40), which is involved in several

neurological and neuromuscular disorders (46-49). These RNA condensates show strong [Mg^{2+}], temperature, sequence, and length dependence, giving rise to the formidable complexity of RNA phase transitions. At present, much remains to be understood regarding the molecular mechanisms and driving forces of RNA-mediated phase separation. In particular, there is a critical need to dissect how the interplay of base stacking, base pairing, electrostatics, ion interactions, and dynamic structural propensities governs the phase behaviors of RNAs.

Molecular dynamics (MD) simulations have played an important role in dissecting the molecular mechanisms and sequence determinant of protein phase separation (50,51), particularly with various coarse-grained (CG) models designed to reach the timescales and length scales required for simulation of condensates (52-56). There have also been numerous CG models of RNAs, which have been summarized in a comprehensive recent review (57). Notably, most existing CG RNA models have been designed for the modelling and prediction of 2D and 3D folded structures and thermodynamics, with different levels of resolutions and consideration of key interactions such as stacking, pairing, and various electrostatic interactions (58-64). At present, only RNA models with one bead per nucleotide have been designed and applied to study phase separation, including the HPS-based model, Mpipi, and COCOMO (65-68). These models primarily focus on supplementing existing protein models for simulating heterotypic RNA-protein phase separation. The pioneering work for simulating homotypic RNA phase separation was based on the single interaction site (SIS) model developed by the Thirumalai group (69). SIS simulations were able to reproduce key experimental findings in the phase separation of RNA triplet repeats (40) and provided a qualitative description of the morphologies and dynamics of RNA chains in condensates. However, the single-bead models do not provide sufficient spatial resolution to describe and resolve the complexity of dynamic RNA structure and interactions, severely limiting their utility in studying RNA phase separation.

Here, we developed a new **i**ntermediate resolution model for **con**densates of **RNA**s (iConRNA) that represents each nucleotide using 6 or 7 beads and explicitly considers a range of backbone and base-mediated interactions. The model is carefully balanced to capture a range of local structures and long-range dynamics of flexible RNAs and thermodynamic properties of folded RNAs. Furthermore, iConRNA includes explicit Mg^{2+} ions and provides a protocol to calibrate the temperature and concentration dependence of Mg^{2+}/phosphate interactions

based on experiment, simulation, and/or theory. The final model does not only capture the salt and Mg^{2+} dependence of model flexible RNAs but also successfully folds small RNAs. We further demonstrate that iConRNA is suitable for direct simulation of RNA phase separation. It recapitulates a wide range of sequence, length, Mg^{2+} and temperature dependence of the phase separation of RNA homopolymers and triplet repeats. Taken together, iConRNA may provide a powerful platform for the study of homotypic RNA phase separation and potentially the heterotypic protein-RNA phase separation.

MATERIAL AND METHODS

All-Atom to CG Mapping

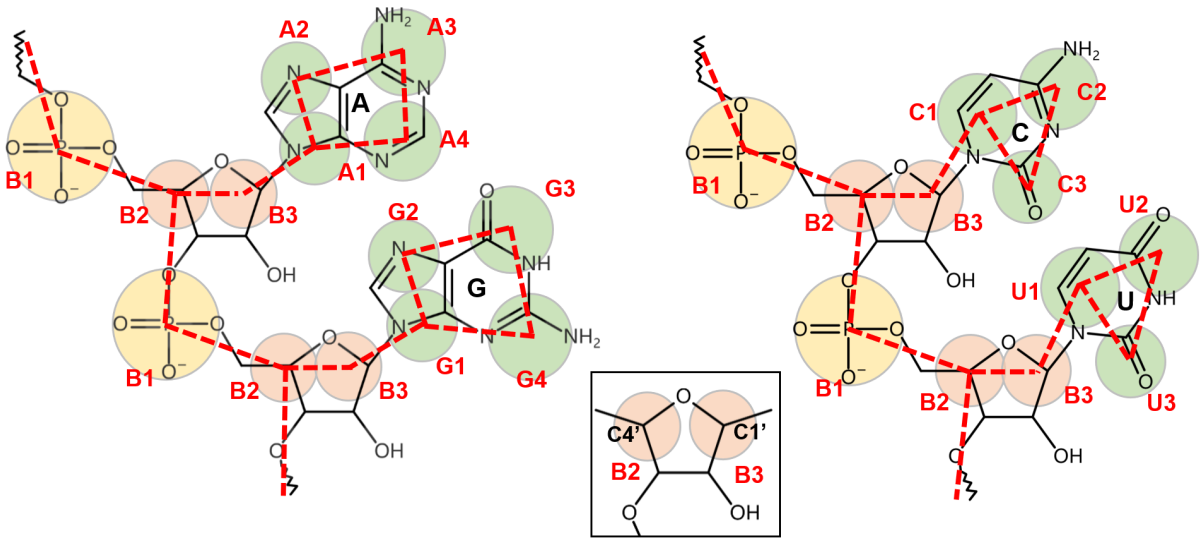

Figure 1**.** All-atom to CG mapping of RNA nucleotides in iConRNA. The CG beads are shown by coloured circles, which roughly contain their corresponding atom groups shown in the chemical structures. The bead names are given in red labels, and red dashed lines are the pseudo-bonds. B2 and B3 are located at the positions of C1′ and C4′ of ribose sugar, respectively, as shown in the inset.

Closely following Martini-RNA (59), each ribonucleotide is represented by six or seven CG beads in iConRNA. As illustrated in **Figure 1**, the phosphate group is represented as one negatively charged bead (B1), while the ribose sugar (including the backbone methylene) is mapped to two beads (B2 and B3). For bases, we use four-bead and three-bead rings to model purines in adenine (A) and guanine (G) and pyrimidines in cytosine (C) and uracil (U), respectively. Detailed all-atom (AA) to CG mapping scheme is given in **Table S1.** Note that,

when converting the atomistic model to the CG model, B2 and B2 are fixed at the positions of C4′ and C1′ of ribose, while the centres of mass are used for other CG beads.

iConRNA Potential Energy Function and Parametrization

Given the mapping scheme summarized in Figure 1, the total potential energy contains contributions from four major types of interactions,

$$U_{total} = U_{bonded} + U_{nonbonded} + U_{stacking} + U_{pairing}. \quad (1)$$

Bonded and nonbonded interactions. U_{bonded} contains contributions from the standard bond, angle, dihedral angle, and improper dihedral angle terms, as used in our previous HyRes protein model (70,71). The only exception is that a restricted bending (ReB) potential (72) is used for the angle term to prevent the bending angle from reaching the value of 180° and avoid numerical instability of related dihedral terms,

$$U_{angle} = \sum_{angles} k_\theta \frac{(\theta - \theta_0)^2}{\sin^2\theta}. \quad (2)$$

Nonbonded interactions $U_{nonbonded}$ includes both electrostatic and Van de Waals terms,

$$U_{nonbonded} = \sum_{ij} \varepsilon_{ij} \left[\left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{6} \right] + \sum_{ij} \frac{q_i q_j}{4\pi\varepsilon_r \varepsilon_0 r_{ij}} e^{-\frac{r_{ij}}{\kappa}}, \quad (3)$$

where $r_{ij}^{min} = \frac{1}{2}\left(r_i^{min} + r_j^{min}\right)$, and $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$. $r_i^{min}/2$ and ε_i are the vdW radius and interaction strength of the bead i. q_i is the charge (-1 for B1 and 0 otherwise), and r_{ij} is the distance between beads i and j. ε_0 is the permittivity of vacuum. κ is the Debye screening length, determined as $\sqrt{9.48/I}$ (in Å) at 300 K, where *I* is the ionic strength in molar. ε_r is the effective dielectric constant, which is set as 20 at 300 K following the HyRes protein model.

Parameters for bonded terms, including all force constants and equilibrium values, were assigned by reproducing the corresponding distributions derived from all-atom (AA) explicit solvent simulations of tetramer single-strand RNAs (ssRNAs), to provide an implicit description of the structural flexibility at the CG level. A set of ssRNA tetramers as used in the Martini-RNA model (59), including GNRA, CUUG, and UNCG tetramers (where N = any nucleotide and R = A or G), was selected to generate the atomistic reference distributions. For nonbonded parameters, the vdW radius ($r_i^{min}/2$) of each CG bead was assigned to reproduce the average total vdW volume of the corresponding groups of atoms derived from the tetramer simulations. The initial values of vdW interaction strengths (ε_i) were taken from the similar atom groups in

the Martini force field (59,73) and then scaled down to the same level of ε_i values in the HyRes protein model (70,71) for future compatibility. The overall strengths of vdW interactions were further optimized by a single scaling factor during the optimization of base stacking and pair interactions to produce the structure and chain flexibilities of RNA homopolymers, including the radius of gyration (R_g) of poly(adenylic acid)$_{30}$ (rA$_{30}$) and poly(uridylic acid)$_{30}$ (rU$_{30}$), and the end-to-end distance (R_e) and persistence length (l_p) of (rU$_{40}$).

Base stacking and pairing interactions. Besides the standard nonbonded interactions, iConRNA also contains explicit base stacking interactions between neighbouring bases,

$$U_{stacking} = \sum_{ij} \varepsilon_{ij}^{stack} \left[5\left(\frac{r_0}{r_{ij}}\right)^{10} - 6\left(\frac{r_0}{r_{ij}}\right)^{6} \right], \tag{4}$$

where $|i - j| = 1$, ε_{ij}^{stack} is the stacking strength between bases i and j. In this work, ε_{ij}^{stack} only depends on the types of stacking bases (denoted as $i//j$), which are taken from Li and Chen (2023) (74) and summarized in **Table S2** (relative to A//A stacking interaction). To mimic the A-form-like helical structure, this stacking interaction is imposed between two virtual sites on bases. As shown in **Figure S1**, each base i has two virtual sites, named S_{i1} and S_{i2}, located at the mass centres of the two bonded beads. For example, for base A, S_{i1} is the mass center of beads A1 and A2 and S_{i2} is the mass center of beads A3 and A4. In the current version of this model, only two neighbouring bases along the RNA chain have stacking interactions. Thus, from the 5′ to the 3′ end, the stacking force between base i and j is assigned between S_{i2} and S_{j1}, with a distance denoted as r_{ij}. The equilibrium distance r_0 between different types of base stacking was set as a uniform value of 0.34 nm.

For base-pairing interactions, only canonical WC A-U and G-C pairs are explicitly included in iConRNA. As shown in **Figure S2**, pairwise interactions $U_{pairing}$ are applied to pairs of hydrogen-bonding CG beads, namely, A3-U2 and A4-U3 for the A-U pair and G3-C2 and G4-C3 for the G-C pair, as

$$U_{pairing} = \begin{cases} \sum_{ij} \varepsilon_{ij}^{pair} \left[5\left(\frac{r_0}{r_{ij}}\right)^{10} - 6\left(\frac{r_0}{r_{ij}}\right)^{6} \right] cos^3\,\varphi\,, & if\ \varphi = \left[\frac{\pi}{2}, \pi\right); \\ 0, & if\ \varphi = [0, \frac{\pi}{2}), \end{cases} \tag{5}$$

where ε_{ij}^{pair} is the strength of hydrogen-bonding pairing, r_{ij} is the distance between corresponding CG beads. φ is the angle of C2-G3-G2 for the G-C pair and U2-A3-A2 for the

A-U pair, which is used to roughly mimic the planar structure of two pairing bases. r_0 is the equilibrium distance for each pair of hydrogen bonding CG beads, which is determined from atomistic structures and summarized in **Table S3**.

To determine the strength of A//A stacking (and thus all stacking interactions), we examine both the radius of gyration (R_g) and the structural propensity of rA_{30} in comparison to the experimental results (75). The latter is quantified by the orientation correlation function (OCF),

$$OCF(|i-j|) = \langle cos\,\theta_{ij} \rangle = \langle \hat{r}_i \cdot \hat{r}_j \rangle, \quad (6)$$

where $\hat{r}_i$ is the normalized bond vector between the ith and (i + 1)th phosphate groups along the RNA chain. After the determination of base stacking strength, the base pairing strength was optimized to reproduce the melting temperature of referenced double-stranded RNA (dsRNA), including $(CAG)_{20}$ and human telomerase RNA (hTR) hairpin. The strengths of canonical A-U and G-C pairs are fixed at the ratio of 2:3 (74). The melting curve was measured through the constant volume heat capacity (C_V). The final parameters for various pairing and stacking interactions are given in **Tables S2** and **S3**.

Treatment of explicit Mg^{2+} ions: concentration and temperature dependence

Screening of electrostatic interactions due to monovalent ions can be treated effectively using the Debye-Hückel-type electrostatic potential (Eqn. 3). as described above. However, such a mean-field treatment is problematic for interactions involving divalent ions such as Mg^{2+} due to ion-ion correlations and complex coordination with the phosphate backbone of RNA (76-78). To more accurately describe specific Mg^{2+} interactions in RNA structural dynamics and phase separation, we include explicit Mg^{2+} ions in iConRNA. A complication is that Mg^{2+} can interact with the phosphate group either as a fully hydrated ion, mostly through electrostatics, or through first-shell coordination, which has significant non-electrostatic components such as polarization and charge transfer (77,79,80). In this work, we include an effective Mg^{2+} ion, denoted as Mg, to roughly describe the essential charge screening and coordination effects. This effective Mg has a comparable volume as the fully hydrated $[Mg(H_2O)_6]^{2+}$ cluster with a radius of 0.3 nm and +2 charge (80,81).

Phosphate-Mg interaction. In the mixed salt solution, magnesium binding with RNA depends strongly on the monovalent salt concentration (82). Previous experimental, theoretical and

simulation studies have shown that the ion competition between Na^+ and Mg^{2+} can be described by the excess Mg^{2+} ions per phosphate, $\Delta n_{Mg} = \Delta N_{Mg}/N_P$, where ΔN_{Mg} is the number of excess Mg^{2+} in the RNA-containing sample compared to the bulk buffer and N_P is the number of phosphates (81-84). Even though Δn_{Mg} depends on the sequence and structure of RNA as well as the concentration of competing monovalent cations (85), it has been shown previously that Δn_{Mg} can be fit to an effective Hill equation (75),

$$\Delta n_{Mg} = \frac{F_{Mg}\left({[Mg]}/{M_{\frac{1}{2}}}\right)^n}{1+\left({[Mg]}/{M_{\frac{1}{2}}}\right)^n}, \tag{7}$$

where F_{Mg} is the number of excess Mg per phosphate in the limit [Mg]→∞, n is the Hill coefficient, and $M_{1/2}$ is the competition coefficient (with respect to Na^+). F_{Mg} is generally ~0.5. In iConRNA, the number of explicit Mg beads in the simulation box is fixed as the number of phosphate groups, which can be considered as an Mg buffer. Increased RNA binding at higher [Mg^{2+}] can be effectively described by strengthening the electrostatic interaction between phosphate groups and Mg using a scaling factor $\lambda_{P\text{-}Mg}$. As will be shown later in Results, there is an apparent linear relationship between $\lambda_{P\text{-}Mg}$ and Δn_{Mg}, which will be empirically determined and applied to simulating RNA structure and phase separation under different Mg^{2+} concentrations.

Temperature dependence of phosphate-Mg interaction. To capture the temperature dependence of electrostatic screening, we adopt an empirical relation previously derived from experimentally measured dielectric constants (86),

$$\varepsilon_r(T) = 87.74 - 0.40008T + 9.398 \times 10^{-4}T^2 - 1.41 \times 10^{-6}T^3, \tag{8}$$

where T is the temperature in Celsius. Because the value of $\varepsilon_r = 20.0$ is used in iConRNA at 303 K, this relation was then scaled down as,

$$\varepsilon_r(T) = 22.94 - 0.10466T + 2.458 \times 10^{-4}T^2 - 0.37 \times 10^{-6}T^3. \tag{9}$$

In addition, entropic contributions due to the release of counterions and coordination water molecules strengthen the interaction between reversely charged groups at higher temperatures, which drives the lower critical solution temperature (LCST)-type phase separation of polyelectrolytes (including RNAs) (44,87-92). The increased interaction between Mg and phosphate can be effectively captured by introducing a temperature dependence in $\lambda_{P\text{-}Mg}$, which can be derived from Δn_{Mg}, from simulations at various temperatures. As will be

shown in Results, a simple linear relation between $\lambda_{P\text{-}Mg}$ and temperature also apparently exists and can be empirically derived from atomistic simulations.

MD simulations and analysis

Atomistic simulations. All atomistic simulations were performed using the GROMACS 2022 package (93-95) in the NPT ensemble. The initial structures of RNAs were generated using the Nucleic Acids Builder module available in the AmberTools package (96). The AMBER99 force field (97) with TIP3P waters (98) was used in combination with the parmbsc0 (99) and the parmchiOL3 (100) corrections for RNA. The 'microMg' parameters from Grotz et al (101) were used for Mg^{2+} with the parameters of Cl^- from Mamatkulov-Schwierz (102). Packmol (103) and CHARMM-GUI (104-106) were used to build the initial simulation boxes. Periodic boundary conditions (PBC) and the particle-mesh Ewald (PME) algorithm (107,108) were applied to calculate the long-range electrostatic interactions, while the cut-off of both the short-range electrostatic and vdW potential were set to 1.1 nm. The V-rescale thermostat (109) was used to maintain the temperature with a coupling time constant of 1.0 ps, while the Parrinello-Rahman method (110) was used to maintain the pressure at 1 bar with a coupling time constant of 5.0 ps.

To derive the reference distributions of bonded terms, 500 ns production simulations were performed for ssRNA tetramers after appropriate equilibration. For each bonded term, the distribution was obtained from any tetramer that contain this term. To calculate the concentration and temperature dependence of excess Mg^{2+}, a single copy of $(CAG)_{31}$ or $(CAG)_{10}$ in the native hairpin conformation (see **Figure S3**) was placed in a simulation box with different concentrations of $MgCl_2$ and NaCl, where $MgCl_2$ was used to do neutralization. The box size is 20×20×20 nm^3 for $(CAG)_{31}$ or 10×10×10 nm^3 for $(CAG)_{10}$. For $(CAG)_{10}$, six 100-ns simulations were performed at every 10 K increment from 303 K to 353 K with 25 mM NaCl and 50 mM $MgCl_2$, to derive the temperature dependence of excess Mg^{2+}. For $(CAG)_{31}$, four 100-ns simulations were performed to predict the concentration dependence of excess Mg^{2+}, which was then used to derive F_{Mg} and $M_{1/2}$ by fitting to the Hill equation (Eqn. 7). The number of Mg^{2+} (N_{Mg}) in excess to the bulk around the RNA was calculated as (111-113)

$$N_{Mg} = 2\pi h \int_0^\infty \left(c_{Mg}(r) - c_{Mg}^0\right) r dr \,, \tag{10}$$

where h is the dimension of RNA along the axis direction, and c_{Mg}^0 is the bulk concentration.

CG simulation protocols. All CG simulations were performed using OpenMM 8.1.0 (114) in the NVT ensemble. Langevin dynamics was used with a friction coefficient of 0.1 ps^{-1}. In-house Python packages were used to create initial CG RNA structures or convert atomistic models into iConRNA representation. For all single-chain simulations, RNA was centred in the box sized from 20 nm to 50 nm for different RNAs and then simulated for at least 2 μs. For the simulations of structured RNAs, two replica simulations were performed starting from either disordered chains or structured chains. At least 2-μs simulations were run to obtain the equilibrium structures. For phase separation simulations, multi-copies of RNAs were randomly packed into cubic boxes to build the initial solution with specific RNA concentrations. The detailed numbers of copies and box sizes are summarized in **Table S4**. To build the phase diagrams, 2 replicas of at least 2-μs simulations were performed at every 10 K increment from 303 K to 353 K for each RNA and/or Mg^{2+} concentrations.

Heat capacity calculation. Following the framework presented by Walker et al. (115), we first performed replica-exchange MD simulations with 15 replicas, where the temperatures were distributed exponentially from 273 K to 393 K. Langevin integrator with a collision frequency of 5 ps^{-1} and time step of 8 fs was used. Replica exchange moves were attempted every 200 MD time steps and a total of 300 ns per replica was run. The constant volume heat capacity was computed from the mean square fluctuation of the enthalpy using the multistate Bennett acceptance ratio (MBAR) method as implemented in the pymbar package (116),

$$C_V = \left(\frac{\left\langle \delta H^2 \right\rangle}{kT^2} \right)_V . \tag{11}$$

Sequences of all RNAs simulated in this work are provided in **Table S5**.

RESULTS

Model parametrization and validation

The force field parameters for the iConRNA model were determined in a sequential matter, starting from bonded interactions, then nonbonded and base stacking interactions, and finally, base-pairing interactions. Parameters involved in bonded interactions were first iteratively optimized to match the corresponding CG and AA distributions, which are summarized in **Figures S4-S7**. Details of all bonded parameters are provided in **Tables S6-S9**. Overall, the CG

distributions match the reference AA distributions very well, but show fewer features in several cases, such as the B2-B3-A1/G1-A2/G2 dihedral angle distribution (**Figure S6**). At the atomistic level, the relative orientation of the base and ribose ring is highly flexible and the B2-B3-A1/G1-A2/G2 dihedral angle could sample both cis- and trans-like configurations for ssRNAs, especially at the termini. Thus a soft term is used to cover the main peak of the reference distribution (**Table S8**, k_ψ = 1.0 kcal/mol).

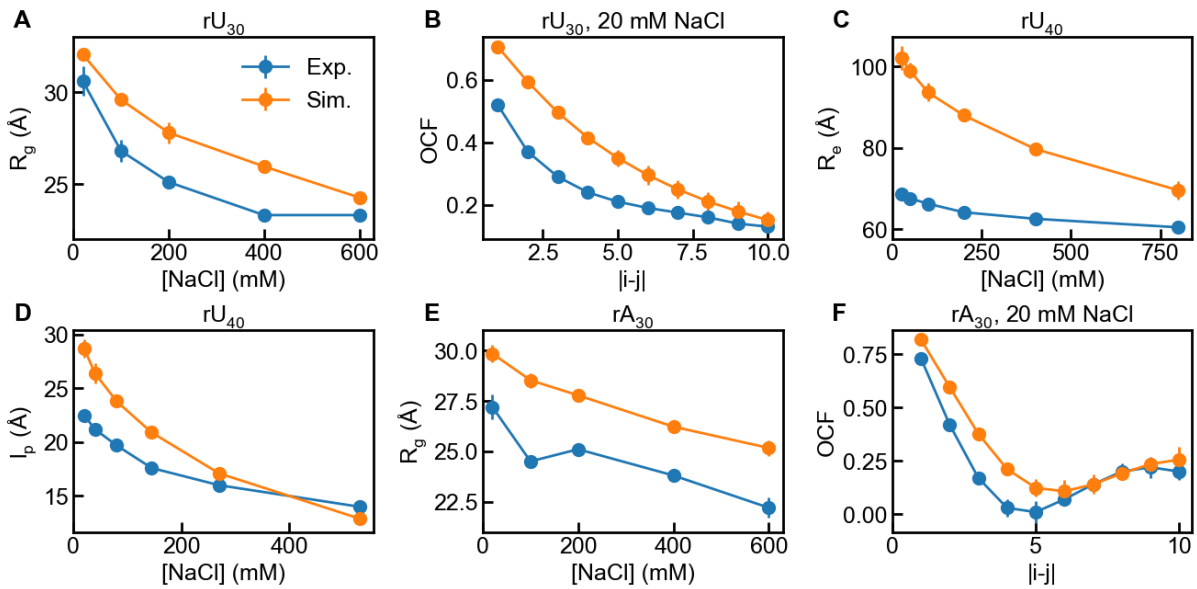

Figure 2. Salt-dependent chain properties of homopolymer RNAs. (**A**) The radius of gyration (R_g) of rU_{30} as a function of NaCl concentration. (**B**) Orientation correlation function (OCF), as defined in Eqn. 6, of rU_{30} with 20 mM NaCl. (**C, D**) The end-to-end distance (R_e) and persistence length (l_p) of rU_{40} as a function of NaCl concentration. (**E**) R_g of rA_{30} as a function of NaCl concentration. (**F**) OCF of rA_{30} with 20 mM NaCl. The experimental results (65,67) are shown in blue whereas the iConRNA simulation results are in orange. The error bars of simulation results were estimated from block analysis.

As discussed in Materials and Methods, the nonbonded and base stacking interactions were tuned by examining the salt-dependent conformational properties of homopolymer RNAs in comparison to the available experimental results (75,117). The key parameters include the overall scaling of vdW interactions and the A//A stacking strength; the latter is used to set all stacking strengths using the relative scales in **Table S2**. The optimized vdW parameters are given in **Tables S10** and **S11**. The final iConRNA model shows a strong ability to describe the structure and dynamics of model ssRNAs and their salt dependence. **Figures 2** and **S8** summarize simulated R_g and OCF of rU_{30} and R_e and l_p of rU_{40} under different salt

concentrations in comparison to available experimental data. Without any base stacking between uracils, rU_{30} is a random coil (**Figure S9A**) and its OCF profiles are featureless (**Figure 2B**), also consistent with other experiments (117-119). Even though R_g (and l_p) from iConRNA simulations is slightly over-estimated compared to the experiment, strong agreement on their salt dependence suggests that the model captures well the balance between electrostatic and vdW interactions. Curiously, it proved difficult to further reduce the overestimation of R_e of rU_{40} (**Figure 2C**). Similar R_e deviations were reported in some other CG RNA models (65,67). Our analysis suggested that this was mainly due to the strong repulsive interactions along the negatively charged backbone, leading to a longer persistence length (**Figure 2D**). Weakening of electrostatic interactions could rescue R_e of rU_{40}, but this would severely reduce the ability of the model to capture salt dependence of structural properties. With the A//A stacking strength set to 2.05 kcal/mol, iConRNA correctly predicts the helical structure of rA_{30} (**Figure S9B**), and both the OCF profiles and their salt dependences show high consistency with experimental results, as illustrated in **Figures 2F** and **S8D-F**. Nonetheless, rA_{30} appears to be slightly more dynamic in iConRNA, leading to slightly shallower features in the OCF profiles. This is likely due to the smooth potentials of the CG model and the higher freedom of the bases. Thus, the strengths of the dihedral angles B1-B2-B3-A1(G1) and B2-B3-A1(G1)-A2(G2) were increased from 8.0 and 5.0 kcal/mol to 35.0 kcal/mol to prevent the deformation.

The strength of base pairing was lastly determined. In iConRNA, the ratio between A-U and G-C pairing strength was fixed at 2:3 as frequently done in CG RNA models (74,120). Both pairs form two (pseudo-)hydrogen bonds to help retain the planar geometry (**Figure S2**). For the A-U pair, hydrogen bonds A3-U2 and A4-U3 take equilibrium distances of 0.33 nm and 0.40 nm, respectively. For the G-C pair, the equilibrium distances of G3-C2 and G4-C3 are 0.33 nm and 0.38 nm. In the final model, the strength of each hydrogen bond of the A-U pair was set as 1.44 kcal/mol, while it was 2.15 kcal/mol for the G-C pair (**Table S3**). With these parameters, iConRNA reproduces the melting temperatures of both $(CAG)_{20}$ (121) and hTR hairpin (122) (**Figure 3A**) as well as the Beet Western Yellow Virus pseudoknot (**Figure S10**). Note that experimental melting profiles frequently contain minor peaks, suggesting the existence of alternative structures and/or multiple unfolding intermediates. iConRNA cannot capture all these details, which is not surprising due to the CG nature and absence of non-canonical base pairing interactions.

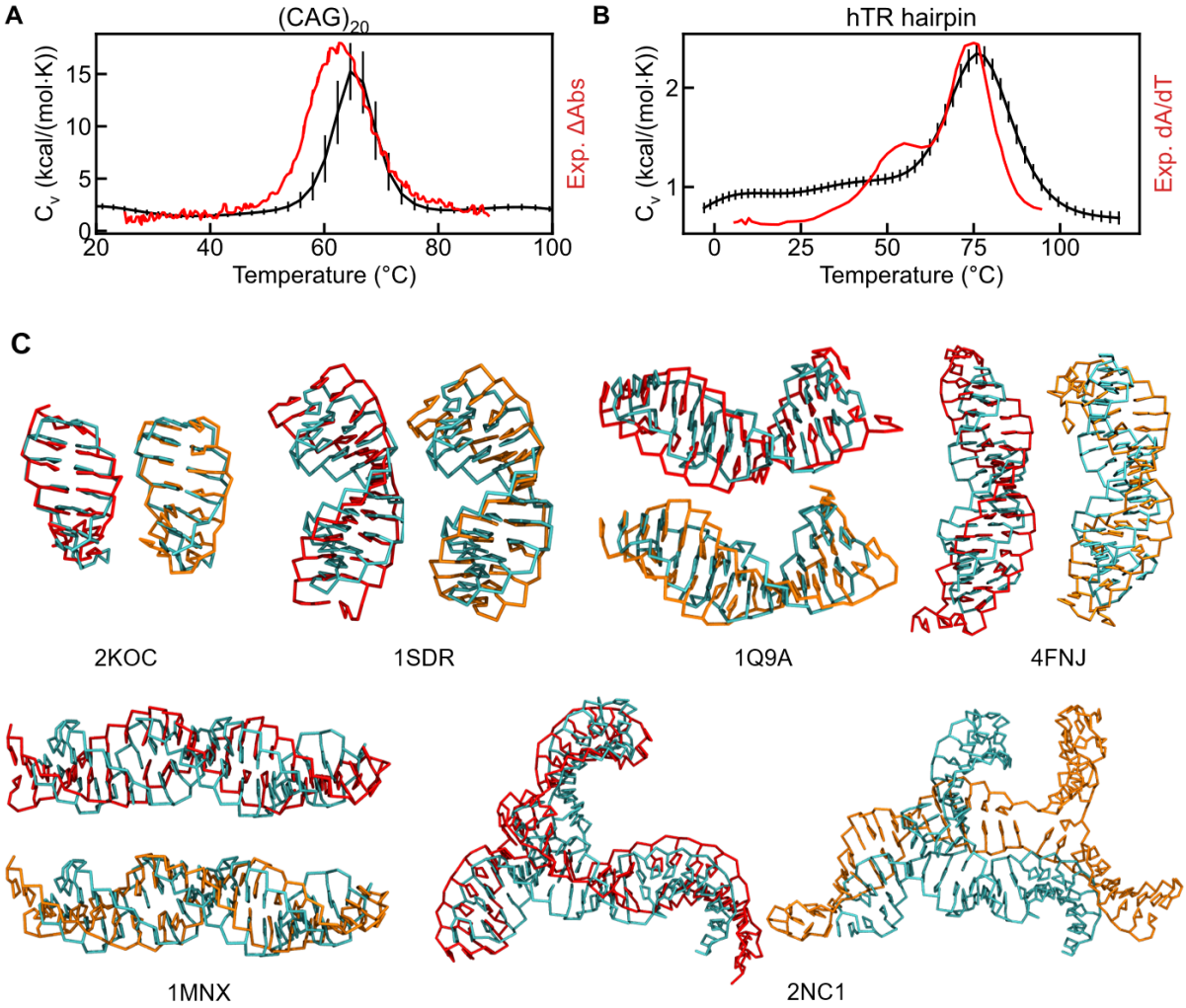

Figure 3. Melting temperature and folding of dsRNAs. (**A, B**) Heat capacity (C_v) of $(CAG)_{20}$ (**A**) and hTR hairpin (**B**). The uncertainties of simulation results were calculated using bootstrap resampling (115). Experimental data are red lines, where the change of UV-absorbance (ΔAbs) at 260 nm and the first derivative of UV-absorbance with respect to temperature (dA/dT) at 280 nm are used as the melting profiles for $(CAG)_{20}$ (121) and hTR hairpin (122), respectively. (**C**) Comparison between experimental crystal structures (cyan) and predicted structures of dsRNA (PDB id: 2KOC, 1SDR, 1Q9A, 4FNJ, 1MNX, 2NC1) (see **Table S5** for sequences). The predicted structures were obtained from single chain simulation starting from crystal structures (red) or linear chains (orange), respectively.

We further examine the ability of iConRNA to fold small dsRNA structures. Starting from crystal structures and coil single strands, we simulated a series of simple dsRNA, including 2KOC, 1SDR, 1Q9A, 4FNJ, 1MNX, and 2NC1 (in PDB id). As illustrated in **Figure 3C**, the model can basically predict the hairpin structures with A-form-like helices, although the detailed helical characteristics are not always consistent with those in crystal structures. For complex RNA like 2NC1, some special interactions including intra-molecular cross-base stacking (not the

stacking between neighbouring bases) and non-canonical base pairing are necessary to retain the structure, which is not included in iConRNA. As such, the folded structure sampled by iConRNA displays large deviations. The ability of iConRNA to correctly fold small RNAs is notable, as the model has been mainly designed to capture the structural properties of dynamic RNAs. The apparent balance between structure and dynamics provided by iConRNA is an important strength for capturing the complex interplay of various interactions and structural propensities of RNA in phase separation.

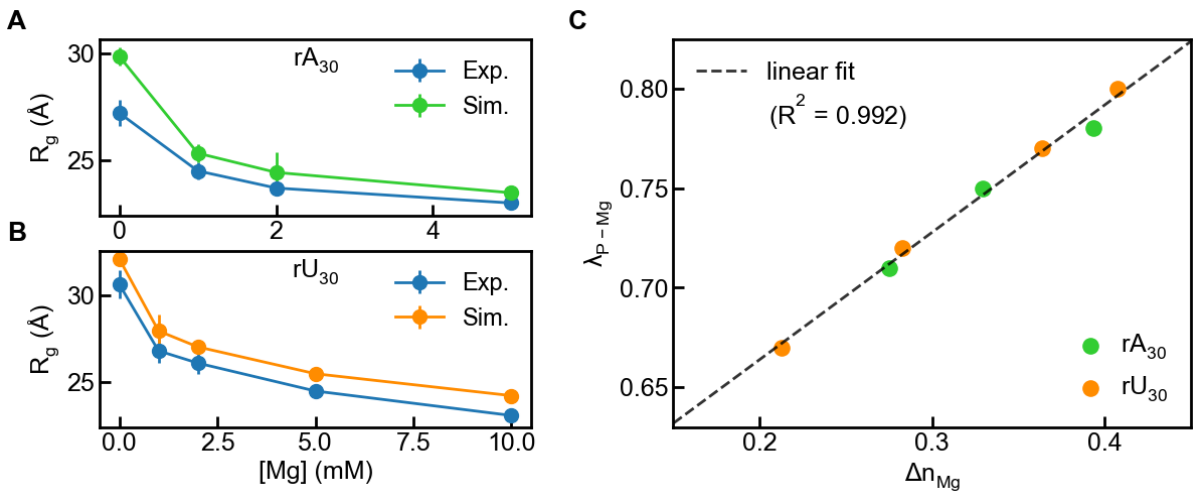

Figure 4. Calibration of phosphate-Mg interaction. (**A, B**) [Mg] dependence of R_g of rA_{30} (**A**) and rU_{30} (**B**). Experimental data (blue lines) are taken from Ref (75). The error bars of simulation results are the standard deviations of block averages. (**C**) Linear relation between optimal $\lambda_{P\text{-}Mg}$ and Δn_{Mg} at various Mg^{2+} concentrations. The black dashed line shows the linear fit with R^2 = 0.992.

Parameterization of Explicit Mg^{2+} interactions

As described in Method section, with a fixed number of Mg in the simulation box, the actual effect of Mg^{2+} concentration on binding to RNA backbone is effectively modelled through scaling the phosphate-Mg interaction strength with $\lambda_{P\text{-}Mg}$, which is related to the excess number of Mg^{2+} per phosphate group (Δn_{Mg}). We first determined the optimal $\lambda_{P\text{-}Mg}$ by reproducing the [Mg^{2+}] dependence of R_g of rA_{30} and rU_{30} (**Figures 4A** and **4B**) (75). The final optimized $\lambda_{P\text{-}Mg}$ at various [Mg^{2+}] are summarized in **Table S12**. The OCF profiles derived from the simulations at 1 and 2 mM [Mg^{2+}] are in reasonable agreement with the experimental references (**Figure S11**), but display larger discrepancies compared to Mg^{2+}-free conditions (e.g. **Figure 2**). This is likely due to that the effective Mg bead in iConRNA is as big as the fully

hydrated Mg^{2+}, which does not account for the dehydration of chelated Mg^{2+} ions and resulting structural differences.

Interestingly, there is a strong linear relation between Δn_{Mg} and the optimal $\lambda_{P\text{-}Mg}$ determined from reproducing R_g of rA_{30} and rU_{30} (**Figure 4C**),

$$\lambda_{P-Mg} = 0.536 + 0.64\,\Delta n_{Mg}, \tag{12}$$

with R^2 of 0.992. This suggests that, for any RNA system in the mixed Na^+/Mg^{2+} solution, the effective phosphate-Mg interaction can be determined through Δn_{Mg} without the need for reparameterization based on measured or predicted R_g of the RNA of interest. Note that Δn_{Mg} itself can be determined through experiment or predicted using simulation or theory (75,81-84,112,123-125). For rA_{30} and rU_{30}, Δn_{Mg} and its fit to the Hill equation have been determined experimentally (75). For RNA triplet repeats to be studied in this work, we performed additional all-atom simulations to determine Δn_{Mg}, using $(CAG)_{31}$ as a representative system In the simulations. The calculated Δn_{Mg} converges well within the 100 ns simulation timeframe (**Figure 5A**), and can be fitted well to the non-cooperative Hill-equation (Eqn. 7) (**Figure 5B**). Clearly, $(CAG)_{31}$ has a lower level of Mg^{2+} binding compared to RNA homopolymers. The fit, given in **Table S13**, will be used for the simulation of all RNA triplet repeats in this work.

The temperature dependence of Δn_{Mg} was further investigated through all-atom simulations of short RNA triplet repeats $(CAG)_{10}$. The distribution of Mg^{2+} around phosphate groups was quantified within the temperature range of 303 K to 353 K under 25 mM NaCl and 50 mM $MgCl_2$. Similar to previous simulations of polyphosphate (44), the resulting radial distribution functions (RDFs) show that the binding of Mg^{2+} ions with phosphate groups strengthens with the increase in temperature (**Figure 5C**). The increase is particularly pronounced in the inner-sphere binding within 4 Å (the first peak), where the Mg^{2+} ion was directly coordinated with the oxygen of the phosphate group (**Figure S12**) (126,127). Interestingly, a linear dependence exists between Δn_{Mg} and the temperature (**Figure 5D**),

$$\Delta n_{Mg}(T) = 0.3716 + 0.0012T, \tag{14}$$

where T is in Celsius, or

$$\Delta n_{Mg}(T) = \Delta n_{Mg}(303\,K) + 0.0012(T - 303), \tag{15}$$

where T is in Kelvin, and Δn_{Mg} (303 K) is the excess binding Mg^{2+} per phosphate at 303 K. The linear relationship suggests that the entropic effects do dominate the temperature dependence of Mg^{2+} binding to RNA backbone phosphates as expected. As such, the above

relationship should roughly hold for all RNAs, and it will be used to model the temperature dependence of Mg^{2+}-phosphate interaction for all subsequent simulations in the work.

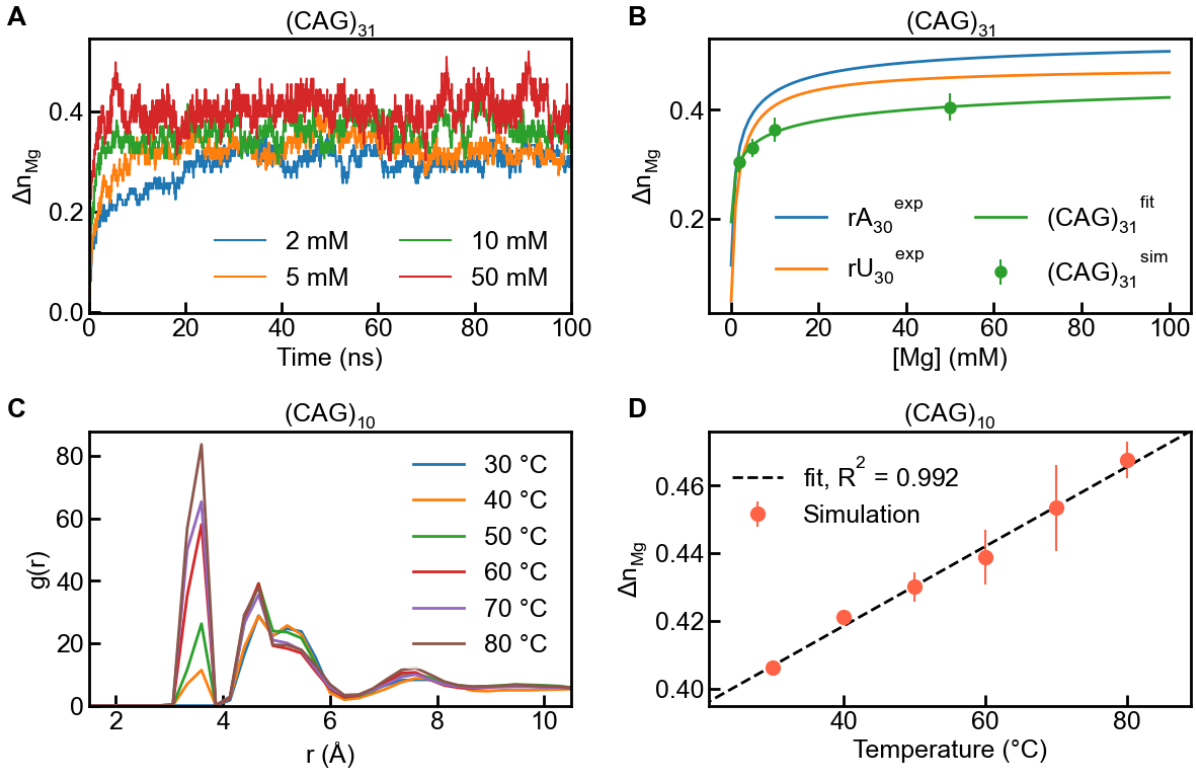

Figure 5. Determination of Δn_{Mg} and its temperature dependence for CAG repeat. (**A**) Δn_{Mg} of $(CAG)_{31}$ as a function of simulation time under different Mg^{2+} concentrations. (**B**) Hill fits of Δn_{Mg} to [Mg] for rA_{30}, rU_{30}, and $(CAG)_{31}$. Labels for legend entries: "exp" is for experimental data, "sim" is for simulation data estimated from the last 50 ns, and 'fit' is for the fit curve. Error bars for simulation data were evaluated as block deviation. (**C**) RDFs of Mg^{2+} ions around phosphorus atoms at different temperatures. (**D**) Linear fit of Δn_{Mg} to temperature (in Celsius). Simulation results are given in orange scatters with error bars estimated from multiple replicas. The black dashed line shows the linear fit with R^2 = 0.992.

Notably, the number of Mg is fixed to be the same as the number of phosphate groups, which does not change with the box size. In principle, $\lambda_{P\text{-}Mg}$ should be re-optimized for different box sizes to reproduce Δn_{Mg}. Therefore, we further evaluated the effect of box size with a fixed number or concentration of Mg on RNA structure. The results show that, without adjusting $\lambda_{P\text{-}Mg}$, only up to 4% change was observed in R_g even with box sizes varying from 15 nm to 50 nm in both setups (**Figure S13**). Therefore, the final model will not correct for the box size dependence of $\lambda_{P\text{-}Mg}$ with a fixed number of explicit Mg beads. In summary, we developed a protocol for setting up explicit magnesium ions in iConRNA simulation. The number of Mg

was first set as the number of phosphate groups. Next, Δn_{Mg} was determined based on reported experiments, theories or atomistic simulations. If temperature varied in the system, Δn_{Mg} at different temperatures could be predicted by **Eqn. 14** or **15**. Finally, Δn_{Mg} was converted into $\lambda_{P\text{-}Mg}$ following **equation 12** as an input parameter of CG simulation.

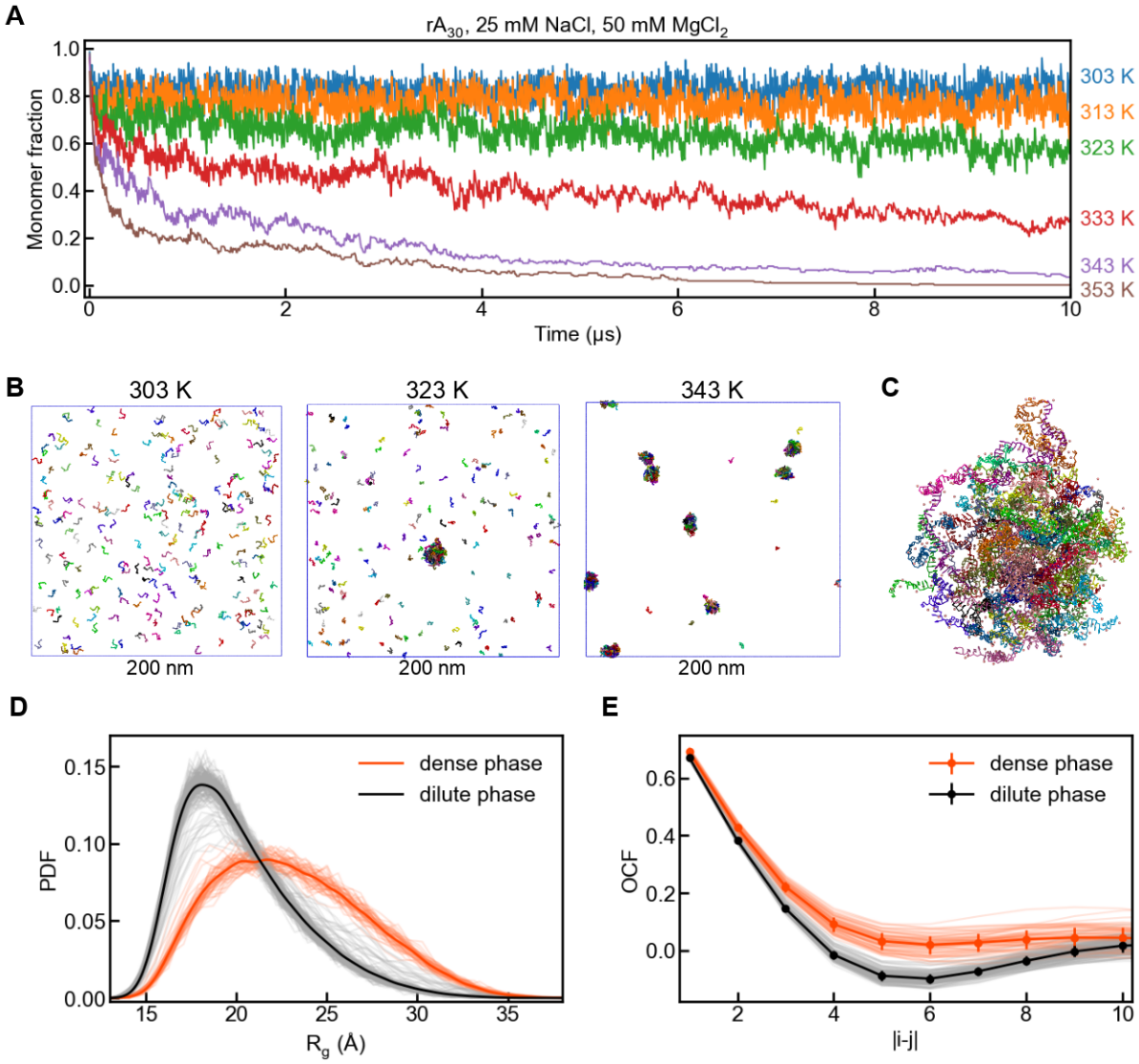

Figure 6. Phase separation of rA_{30}. (**A**) Monomer fractions at different temperatures under 25 mM NaCl and 50 mM $MgCl_2$. The total RNA concentration is 50 μM. (**B**) Final snapshots at 303, 323, and 343 K. Mg were omitted for clarity. The box size is 200 nm. (**C**) Snapshot of a representative rA_{30} droplet. Each RNA chain is coloured differently, while Mg ions are shown as pink beads. (**D, E**) R_g distributions (as probability density function, PDF) and OCFs of RNA chains in the dense (orange-red lines) and dilute phases (black and grey lines) at 323 K. Light-coloured lines plot distributions of individual chains, and their averages are in bold lines. Error bars are the standard deviation of related chain groups.

Phase separation of poly(rA) and poly(rU)

Recent works have demonstrated that RNA homopolymers can undergo homotypic phase separation in the presence of multivalent cations, and exhibit very different phase behaviours when composed of different nucleotides (43-45,91,128). For example, rA_{30} readily undergoes phase separation under 50 mM $MgCl_2$, while poly(rU) remains soluble under the same conditions (45). We first examined the ability of iConRNA to recapitulate the phase separation of RNA homopolymers in the presence of Mg^{2+}. Starting from dispersed initial states with 240 copies in a cubic box of 200 nm in dimension (50 μM), rU_{30} remained dispersed and can only form small oligomers (up to pentamers) with temperatures ranging from 303 to 353 K even with 50 mM $MgCl_2$ (and 25 mM NaCl) (**Figure S14**). In contrast, rA_{30} clearly shows LCST-type phase separation behaviours as observed in experiments (44) (**Figure 6A** and **6B**). At lower temperatures (300 K and 313 K), only dimers and trimers were obtained. Starting from 323 K, stable droplets can be observed and the number of monomers in the dilute phase decreases rapidly with the increase in temperature. We note that experimentally rA_{30} undergoes phase separation as low as 303 K with 50 mM $MgCl_2$ (45). Therefore, iConRNA underpredicts the phase separation propensity of rA_{30}. A possible reason is that only the stacking between neighbouring bases is included in our model, such that rA_{30} is unable to form intermolecular base stacking interactions potentially important in the phase separation of rA_{30}.

Within the condensate, rA_{30} forms extensive intermolecular interactions through phosphate-Mg coordination and base-base vdW interactions among base As. These base-base interactions are similar to A-A stacking and non-canonical A-A pairing interactions, although iConRNA doesn't have these terms explicitly. (**Figure 6C**). We further characterized the conformations of rA_{30} inside the dense phase. The results show that RNA chains within the condensates are more extended, exhibiting a larger average value and broader distribution of R_g (**Figure 6D**), Furthermore, each individual R_g distribution (the light orange lines) is very close to the average one (bold orange line), which indicates a more homogeneous conformation distribution of rA_{30} inside the condensates. Additionally, we found condensates are less favourable for the single-strand helical structure of rA_{30} compared to the dilute phase. OCFs of RNA chains within the condensates exhibit flatter trends and fewer structural details (**Figure 6E**), indicating weakened base stacking. In the crowded environment of condensates, intermolecular interactions among rA_{30} chains are highly enhanced, which results in the partial disruption of local intramolecular base stacking.

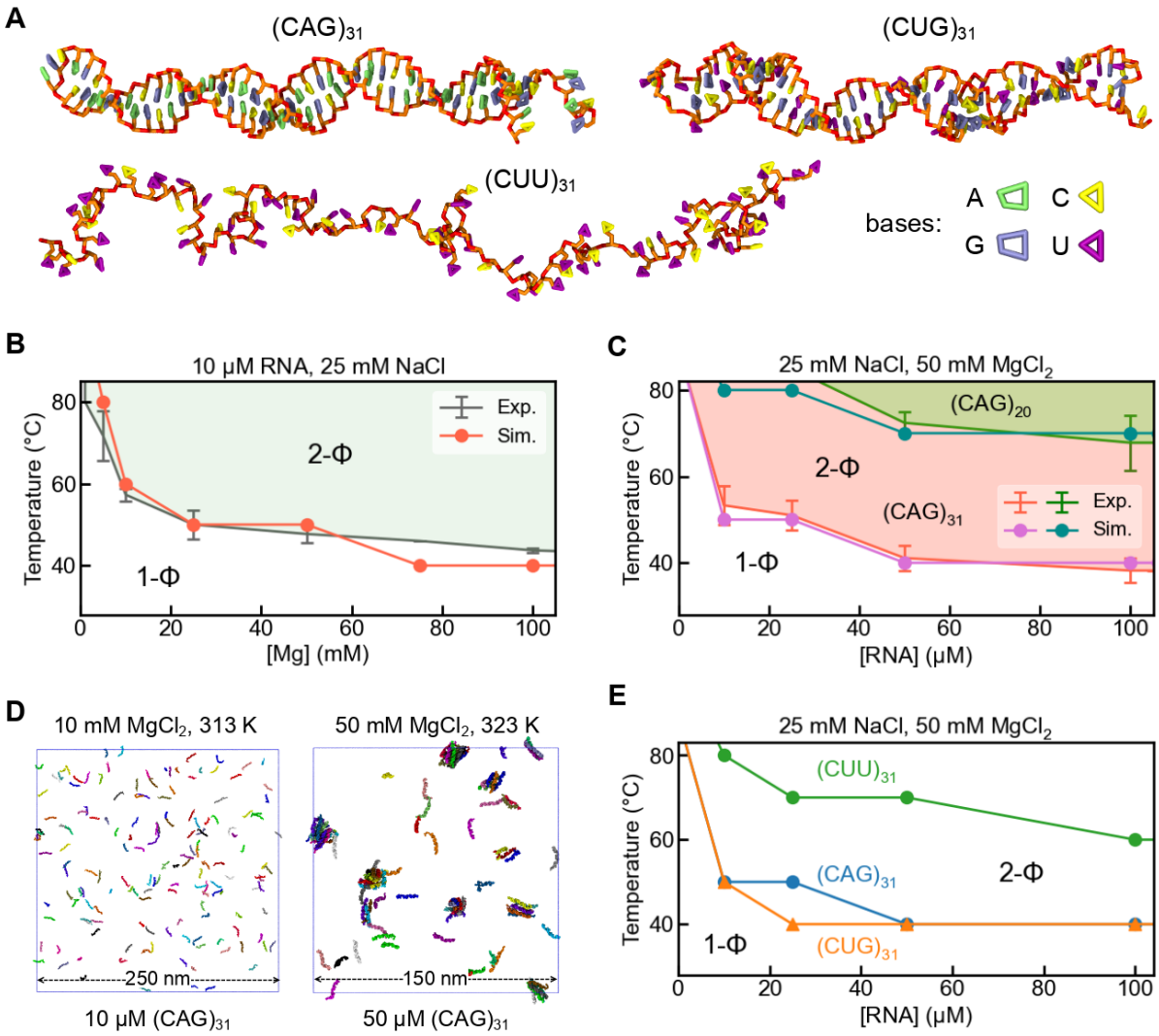

Figure 7. Phase separation of RNA triplet repeats. (**A**) Snapshots of representative structures of monomeric $(CAG)_{31}$, $(CUG)_{31}$, and $(CUU)_{31}$ under 25 mM NaCl at 303 K. Phosphate bead B1 and ribose beads B2/B3 are in red and orange, while bases A, G, C, and U are in lime, ice blue, yellow, and purple, respectively. (**B**) Phase diagram of 10 μM $(CAG)_{31}$ under 25 mM NaCl, where transition temperatures are plotted as a function of [Mg]. (**C**) Phase diagrams of $(CAG)_{31}$ and $(CAG)_{20}$ under 25 mM NaCl and 50 mM $MgCl_2$, where transition temperatures are plotted as a function of [RNA]. Experimental 2-phase regions of $(CAG)_{31}$ and $(CAG)_{20}$ are shown in pink and lime areas. In B and C, experimental phase boundaries are plotted as lines with error bars, while simulation ones are lines with filled circles. (**D**) Snapshots of representative points in 1-phase (1-Φ, left panel) and 2-phase (2-Φ, right panel) regions, which were obtained from simulations of 10 μM $(CAG)_{31}$ at 313 K and 50 μM $(CAG)_{31}$ at 323 K, respectively (with 25 mM NaCl and 10 mM $MgCl_2$). RNAs are coloured in chains. Mg were omitted for clarity. (**E**) Phase diagrams of $(CAG)_{31}$ (blue), $(CUG)_{31}$ (orange), and $(CUU)_{31}$ (green) under 25 mM NaCl and 50 mM $MgCl_2$, where transition temperatures are plotted as a function of [RNA]. For each RNA, the lower left area of the boundary line is the 1-Φ region, while the upper right one is the 2-Φ region.

Phase separation of RNA triplet repeats

Nucleotide triplet repeats have been implicated in several neurological and neuromuscular disorders (46,48,49). Recently, Jain and Vale found CAG/CUG repeat-containing RNAs could undergo phase separation and gelation, which could be a contributing factor to neurological disease (40). The temperature, sequence and Mg^{2+} dependence of the phase separation of CAG/CUG repeats and additional percolation transition within condensates have been further characterized in detail (44). The formidable complexity of trinucleotide repeat phase separation provides a challenging test case for evaluating the strengths and weaknesses of the new RNA CG model.

We first examine the ability of iConRNA to capture the structures of these triplet repeats in the monomer state. It has been shown that CAG/CUG repeats can form stable hairpins, while the CUU repeat remains a single strand without any higher-order structure (49,129,130). As shown in **Figure 7A**, starting from single strands, iConRNA successfully predicts the hairpin structures of $(CAG)_{31}$ and $(CUG)_{31}$ and the disordered single strand of $(CUU)_{31}$. Consistent with experimental observations (130), CAG hairpins were observed to be more stable than CUG ones. The backbone RMSD is 11.8 ± 3.5 Å for CAG, but 34.1 ± 12.4 Å for CUG. Although only canonical A-U and G-C pairing interactions are included in our model, some semi-stable A-A pairs can be found in CAG hairpins due to neighbouring A-U and G-C pairs and base stacking interactions, while U-U pairs are very dynamic in CUG hairpins.

We then calculated the phase diagrams of $(CAG)_{31}$ RNA as a function of magnesium concentration (**Figure 7B**) and RNA concentration (**Figure 7C**). As detailed above, for CAG repeats, Δn_{Mg} and its temperature dependence were first determined through all-atom simulations. Similar to rA_{30}, $(CAG)_{31}$ undergoes LCST-type phase separation, where the transition temperature decreases with the increase of [Mg], indicating the entropic contribution to phase separation. Notably, an excellent agreement was between simulations and experiments for both [Mg]-dependence and [RNA]-dependence (44). For the [Mg]-dependence, the critical temperature is slightly overestimated at the low-Mg^{2+} region (below 25 mM) but underestimated at the high-Mg^{2+} region above 50 mM. It might result from the corresponding under- or over-estimation of excess binding Mg in the all-atom simulations. Notably, in the 1-Φ region, at the temperatures close to the phase boundary, there always exist

small RNA oligomers (from dimer to pentamer, left panel in **Figure 7D**), which are dynamic and undergo fast exchange with the monomers in the buffer.

We further evaluated the ability of iConRNA to capture the length dependence of LCST-type phase transition of CAG repeats observed experimentally (44). As summarized in **Figure 7C**, decreasing the RNA length to 20 repeats dramatically weakens the propensity of CAG repeats phase separation. In particular, the critical temperature increases throughout the concentration range examined. For example, it increases from 40 to 70 °C at 100 μM RNAs in simulation, while in experiments, it increases from 38.0 °C for 100 μM $(CAG)_{31}$ RNA to 59.9 °C for 155 μM $(CAG)_{20}$. In addition, under all the conditions tested for $(CAG)_{20}$ and $(CAG)_{31}$ repeats in **Figure 7C**, $(CAG)_{10}$ repeats didn't undergo any phase separation. For example, under the most favourable conditions, 80 °C and 100 μM RNAs, only a few dimers and trimers were found throughout the simulation box (**Figure S15**).

Experimentally, it was reported the purine-to-uracil substitution could suppress LCST phase transitions (44). Compared to $(CAG)_{31}$, $(CUG)_{31}$ exhibited slightly lower phase separation propensity, and $(CUU)_{31}$ did not undergo phase separation under the same conditions. In our simulations, $(CUG)_{31}$ shows a very similar phase boundary with $(CAG)_{31}$ (**Figure 7E**), indicating the weakening effect of A-to-U substitution for $(CAG)_{31}$ was not well captured by this model. However, compared to $(CUG)_{31}$, $(CUU)_{31}$ does show a significantly weaker phase separation propensity, requiring much higher temperature (and thus stronger Mg^{2+}/phosphate interactions) to phase separate in simulation. The limited ability of iConRNA to capture the negative effect of purine-to-uracil substitution in phase separation trinucleotide repeats suggests that there is still substantial room for improvement in balancing various interactions within iConRNA. In particular, it is likely that vdW interactions between uracils are over-estimated. Another possible reason could be that the excess binding Mg^{2+} of $(CUG)_{31}$ and $(CUU)_{31}$ may be weaker than $(CAG)_{31}$ used to drive $\lambda_{P\text{-}Mg}$ for all trinucleotide repeat simulations.

DISCUSSION

We have developed a new intermediate resolution coarse-grained RNA model with explicit magnesium ions, named iConRNA, for the simulation of dynamic RNAs and their phase separation. Representing each nucleotide using 6 or 7 CG beads, the model is designed to describe a diverse range of interactions that RNA can make through the backbone and bases,

specifically electrostatics, vdW interactions, base stacking and base pairing. Even with only WC canonical base pairing and neighbouring base stacking, iConRNA can be parameterized to capture the essential conformational features of flexible RNAs and fold small RNAs including tetraloop, duplex and hairpins. One of the major challenges in RNA modelling is describing the interactions with Mg^{2+} ions, which are important in RNA structure and phase separation and display complex structure, ion concentration and temperature dependence. To achieve a balance between accuracy and computational efficiency, we developed a framework for including explicit effective CG Mg ions in iConRNA. The sequence, concentration and temperature dependence are captured by parametrizing a scaling factor for Mg^{2+}/phosphate interaction strengths, based on excess Mg^{2+} binding to RNA derived from experiment, simulation and/or theory. The resulting model can successfully capture the Mg^{2+} dependence of conformational properties of flexible RNAs. We further demonstrated that iConRNA was able to reproduce a wide range of nontrivial phase separation properties of homopolymer and trinucleotide repeat RNAs, including length, sequence, Mg^{2+}, and temperature dependence.

Extensive benchmarks also revealed several limitations of iConRNA for further improvement. For example, base stacking occurs only between two neighbouring bases in iConRNA. The stacking between intramolecular non-neighbouring bases and intermolecular bases could be important for RNA phase separation, which was suggested as the primary contributing factor to the phase separation of poly(rA) under high salt concentration (90). In our tests, the experimentally observed temperature-dependent conformation and phase separation of poly(rA) could not be reproduced accurately. Furthermore, the phase separation propensity of rA_{30} in the presence of magnesium was underestimated. Another limitation is that iConRNA only considers the canonical WC base pairs and neglects the formation of non-canonical base pairs (131-133). It's a challenge to mimic each kind of base pairing by specific interactions. A simple and practical strategy may be to introduce a general hydrogen bond formation ability into each potential donor and acceptor that could contribute to the formation of non-canonical base pairing. However, it has proven challenging to balance the strength and propensity of different types of base pairing at the CG level. Finally, although the effective magnesium did perform well in the benchmarks summarized in this work, it's still limited in describing chelated Mg^{2+} binding through the inner sphere. Using the same effective concentration and temperature dependence for a given RNA also neglects the difference in

Mg^{2+} binding depending on the local structural environment. Despite these limitations, the successes demonstrated in this work support the great potential of using intermediate resolution CG models for simulating RNA dynamic conformation and phase separation. The current iConRNA model, albeit with limitations, already provides a powerful tool for studying RNA condensates with much greater details compared to the existing single-bead models.

DATA AVAILABILITY

The data underlying this article including all iConRNA force field files are available freely from GitHub at: https://github.com/lslumass/iConRNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

All simulations were performed on the Pikes GPU cluster housed in the Massachusetts Green High-Performance Computing Cluster (MGHPCC).

FUNDING

This work was supported by the National Institutes of Health [Grant R35 GM144045 to J.C.].

REFERENCES

1. Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F. and Hyman, A.A. (2009) Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, **324**, 1729-1732.
2. Banani, S.F., Lee, H.O., Hyman, A.A. and Rosen, M.K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.*, **18**, 285-298.
3. Brangwynne, C.P., Tompa, P. and Pappu, R.V. (2015) Polymer physics of intracellular phase transitions. *Nat. Phys.*, **11**, 899-904.
4. Holehouse, A.S. and Pappu, R.V. (2018) Functional Implications of Intracellular Phase Transitions. *Biochemistry*, **57**, 2415-2423.
5. Alberti, S., Gladfelter, A. and Mittag, T. (2019) Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell*, **176**, 419-434.
6. Alberti, S. and Dormann, D. (2019) Liquid–Liquid Phase Separation in Disease. *Annu. Rev. Genet.*, **53**, 171-194.
7. Zbinden, A., Pérez-Berlanga, M., De Rossi, P. and Polymenidou, M. (2020) Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force. *Dev. Cell*, **55**, 45-68.
8. Wegmann, S., Eftekharzadeh, B., Tepper, K., Zoltowska, K.M., Bennett, R.E., Dujardin, S., Laskowski, P.R., MacKenzie, D., Kamath, T., Commins, C. *et al.* (2018) Tau protein liquid–liquid phase separation can initiate tau aggregation. *EMBO J.*, **37**, e98049.
9. Taylor, J.P., Brown, R.H. and Cleveland, D.W. (2016) Decoding ALS: from genes to mechanism. *Nature*, **539**, 197-206.
10. Davis, R.B., Moosa, M.M. and Banerjee, P.R. (2022) Ectopic biomolecular phase transitions: fusion proteins in cancer pathologies. *Trends Cell Biol.*, **32**, 681-695.
11. Wiedner, H.J. and Giudice, J. (2021) It's not just a phase: function and characteristics of RNA-binding proteins in phase separation. *Nat. Struct. Mol. Biol.*, **28**, 465-473.
12. Hyman, A.A., Weber, C.A. and Jülicher, F. (2014) Liquid-Liquid Phase Separation in Biology. *Annu. Rev. Cell. Dev. Biol.*, **30**, 39-58.
13. Pappu, R.V., Cohen, S.R., Dar, F., Farag, M. and Kar, M. (2023) Phase Transitions of Associative Biomacromolecules. *Chem. Rev.*, **123**, 8945-8987.
14. Guo, Q., Shi, X. and Wang, X. (2021) RNA and liquid-liquid phase separation. *Non-Coding RNA Res.*, **6**, 92-99.
15. Mittag, T. and Pappu, R.V. (2022) A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell*, **82**, 2201-2214.
16. Schmit, J.D., Bouchard, J.J., Martin, E.W. and Mittag, T. (2020) Protein Network Structure Enables Switching between Liquid and Gel States. *J. Am. Chem. Soc.*, **142**, 874-883.
17. Dignon, G.L., Best, R.B. and Mittal, J. (2020) Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. *Annu. Rev. Phys. Chem.*, **71**, 53-75.
18. Villegas, J.A., Heidenreich, M. and Levy, E.D. (2022) Molecular and environmental determinants of biomolecular condensate formation. *Nat. Chem. Biol.*, **18**, 1319-1329.
19. Li, S., Zhang, Y. and Chen, J. (2024) Backbone interactions and secondary structures in phase separation of disordered proteins. *Biochem. Soc. Trans.*, **52**, 319-329.
20. Pappu, R.V., Cohen, S.R., Dar, F., Farag, M. and Kar, M. (2023) Phase Transitions of Associative Biomacromolecules. *Chem. Rev.*, **123**, 8945-8987.
21. McCarty, J., Delaney, K.T., Danielsen, S.P.O., Fredrickson, G.H. and Shea, J.E. (2019) Complete Phase Diagram for Liquid-Liquid Phase Separation of Intrinsically Disordered Proteins. *J Phys Chem Lett*, **10**, 1644-1652.
22. Li, S.-F. and Muthukumar, M. (2022) Theory of Microphase Separation in Concentrated Solutions of Sequence-Specific Charged Heteropolymers. *Macromolecules*, **55**, 5535-5549.
23. Wessén, J., Das, S., Pal, T. and Chan, H.S. (2022) Analytical Formulation and Field-Theoretic Simulation of Sequence-Specific Phase Separation of Protein-Like Heteropolymers with

Short- and Long-Spatial-Range Interactions. *The Journal of Physical Chemistry B*, **126**, 9222-9245.
24. Biswas, S. and Potoyan, D.A. (2024) Molecular Drivers of Aging in Biomolecular Condensates: Desolvation, Rigidification, and Sticker Lifetimes. *PRX Life*, **2**, 023011.
25. Valdes-Garcia, G., Heo, L., Lapidus, L.J. and Feig, M. (2023) Modeling Concentration-dependent Phase Separation Processes Involving Peptides and RNA via Residue-Based Coarse-Graining. *J. Chem. Theory Comput.*
26. Wadsworth, G.M., Srinivasan, S., Lai, L.B., Datta, M., Gopalan, V. and Banerjee, P.R. (2024) RNA-driven phase transitions in biomolecular condensates. *Mol. Cell*, **84**, 3692-3705.
27. Ripin, N. and Parker, R. (2023) Formation, function, and pathology of RNP granules. *Cell*, **186**, 4737-4756.
28. Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M. *et al.* (2012) Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, **149**, 1393-1406.
29. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, Anderson P., Kim, Hong J., Mittag, T. and Taylor, J.P. (2015) Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell*, **163**, 123-133.
30. Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillén-Boixet, J., Franzmann, T.M. *et al.* (2018) RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science*, **360**, 918-921.
31. Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C.C.-H., Eckmann, C.R., Myong, S. and Brangwynne, C.P. (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 7189-7194.
32. Protter, D.S.W., Rao, B.S., Van Treeck, B., Lin, Y., Mizoue, L., Rosen, M.K. and Parker, R. (2018) Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Reports*, **22**, 1401-1412.
33. Yoshizawa, T., Ali, R., Jiou, J., Fung, H.Y.J., Burke, K.A., Kim, S.J., Lin, Y., Peeples, W.B., Saltzberg, D., Soniat, M. *et al.* (2018) Nuclear Import Receptor Inhibits Phase Separation of FUS through Binding to Multiple Sites. *Cell*, **173**, 693-705.e622.
34. Lin, Y., Protter, D.S.W., Rosen, M.K. and Parker, R. (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol. Cell*, **60**, 208-219.
35. Pessina, F., Giavazzi, F., Yin, Y., Gioia, U., Vitelli, V., Galbiati, A., Barozzi, S., Garre, M., Oldani, A., Flaus, A. *et al.* (2019) Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. *Nat. Cell Biol.*, **21**, 1286-1299.
36. Langdon, E.M., Qiu, Y., Ghanbari Niaki, A., McLaughlin, G.A., Weidmann, C.A., Gerbich, T.M., Smith, J.A., Crutchley, J.M., Termini, C.M., Weeks, K.M. *et al.* (2018) mRNA structure determines specificity of a polyQ-driven phase separation. *Science*, **360**, 922-927.
37. Zhang, H., Elbaum-Garfinkle, S., Langdon, E.M., Taylor, N., Occhipinti, P., Bridges, A.A., Brangwynne, C.P. and Gladfelter, A.S. (2015) RNA Controls PolyQ Protein Phase Transitions. *Mol. Cell*, **60**, 220-230.
38. Van Treeck, B. and Parker, R. (2018) Emerging Roles for Intermolecular RNA-RNA Interactions in RNP Assemblies. *Cell*, **174**, 791-802.
39. Boeynaems, S., Holehouse, A.S., Weinhardt, V., Kovacs, D., Van Lindt, J., Larabell, C., Van Den Bosch, L., Das, R., Tompa, P.S., Pappu, R.V. *et al.* (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**, 7889-7898.
40. Jain, A. and Vale, R.D. (2017) RNA phase transitions in repeat expansion disorders. *Nature*, **546**, 243-247.

41. Van Treeck, B., Protter, D.S.W., Matheny, T., Khong, A., Link, C.D. and Parker, R. (2018) RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, 2734-2739.
42. Poudyal, R.R., Sieg, J.P., Portz, B., Keating, C.D. and Bevilacqua, P.C. (2021) RNA sequence and structure control assembly and function of RNA condensates. *RNA*, **27**, 1589-1601.
43. Aumiller, W.M., Jr., Pir Cakmak, F., Davis, B.W. and Keating, C.D. (2016) RNA-Based Coacervates as a Model for Membraneless Organelles: Formation, Properties, and Interfacial Liposome Assembly. *Langmuir*, **32**, 10042-10053.
44. Wadsworth, G.M., Zahurancik, W.J., Zeng, X., Pullara, P., Lai, L.B., Sidharthan, V., Pappu, R.V., Gopalan, V. and Banerjee, P.R. (2023) RNAs undergo phase transitions with lower critical solution temperatures. *Nat. Chem.*, **15**, 1693-1704.
45. Tom, J.K.A., Onuchic, P.L. and Deniz, A.A. (2022) Short PolyA RNA Homopolymers Undergo Mg2+-Mediated Kinetically Arrested Condensation. *J. Phys. Chem. B*, **126**, 9715-9725.
46. Gatchel, J.R. and Zoghbi, H.Y. (2005) Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. *Nat. Rev. Genet.*, **6**, 743-755.
47. Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L. *et al.* (2011) A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron*, **72**, 257-268.
48. La Spada, A.R. and Taylor, J.P. (2010) Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nat. Rev. Genet.*, **11**, 247-258.
49. Krzyzosiak, W.J., Sobczak, K., Wojciechowska, M., Fiszer, A., Mykowska, A. and Kozlowski, P. (2012) Triplet repeat RNA structure and its role as pathogenic agent and therapeutic target. *Nucleic Acids Res.*, **40**, 11-26.
50. Shea, J.E., Best, R.B. and Mittal, J. (2021) Physics-based computational and theoretical approaches to intrinsically disordered proteins. *Curr. Opin. Struct. Biol.*, **67**, 219-225.
51. Choi, J.M., Holehouse, A.S. and Pappu, R.V. (2020) Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annu Rev Biophys*, **49**, 107-133.
52. Bari, K.J. and Prakashchand, D.D. (2021) Fundamental Challenges and Outlook in Simulating Liquid–Liquid Phase Separation of Intrinsically Disordered Proteins. *J. Phys. Chem. Lett.*, **12**, 1644-1656.
53. Shea, J.-E., Best, R.B. and Mittal, J. (2021) Physics-based computational and theoretical approaches to intrinsically disordered proteins. *Curr. Opin. Struct. Biol.*, **67**, 219-225.
54. Dignon, G.L., Zheng, W. and Mittal, J. (2019) Simulation methods for liquid–liquid phase separation of disordered proteins. *Curr. Opin. Chem. Eng.*, **23**, 92-98.
55. Zhang, Y., Li, S., Gong, X. and Chen, J. (2024) Toward Accurate Simulation of Coupling between Protein Secondary Structure and Phase Separation. *J. Am. Chem. Soc.*, **146**, 342-357.
56. Choi, J.M., Dar, F. and Pappu, R.V. (2019) LASSI: A lattice model for simulating phase transitions of multivalent proteins. *PLoS Comput. Biol.*, **15**, e1007028.
57. Šponer, J., Bussi, G., Krepl, M., Banáš, P., Bottaro, S., Cunha, R.A., Gil-Ley, A., Pinamonti, G., Poblete, S., Jurečka, P. *et al.* (2018) RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.*, **118**, 4177-4338.
58. Leontis, N.B. and Westhof, E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499-512.
59. Uusitalo, J.J., Ingolfsson, H.I., Marrink, S.J. and Faustino, I. (2017) Martini Coarse-Grained Force Field: Extension to RNA. *Biophys. J.*, **113**, 246-256.
60. Pasquali, S. and Derreumaux, P. (2010) HiRE-RNA: a high resolution coarse-grained energy model for RNA. *J Phys Chem B*, **114**, 11957-11966.

61. Boniecki, M.J., Lach, G., Dawson, W.K., Tomala, K., Lukasz, P., Soltysinski, T., Rother, K.M. and Bujnicki, J.M. (2016) SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res.*, **44**, e63.
62. Xia, Z., Gardner, D.P., Gutell, R.R. and Ren, P. (2010) Coarse-grained model for simulation of RNA three-dimensional structures. *J Phys Chem B*, **114**, 13497-13506.
63. Denesyuk, N.A. and Thirumalai, D. (2013) Coarse-grained model for predicting RNA folding thermodynamics. *J. Phys. Chem. B*, **117**, 4901-4911.
64. Bernauer, J., Huang, X., Sim, A.Y. and Levitt, M. (2011) Fully differentiable coarse-grained and all-atom knowledge-based potentials for RNA structure evaluation. *RNA*, **17**, 1066-1075.
65. Valdes-Garcia, G., Heo, L., Lapidus, L.J. and Feig, M. (2023) Modeling Concentration-dependent Phase Separation Processes Involving Peptides and RNA via Residue-Based Coarse-Graining. *J. Chem. Theory Comput.*, **19**, 669-678.
66. Joseph, J.A., Reinhardt, A., Aguirre, A., Chew, P.Y., Russell, K.O., Espinosa, J.R., Garaizar, A. and Collepardo-Guevara, R. (2021) Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nature Computational Science*, **1**, 732-743.
67. Regy, R.M., Dignon, G.L., Zheng, W., Kim, Y.C. and Mittal, J. (2020) Sequence dependent phase separation of protein-polynucleotide mixtures elucidated using molecular simulations. *Nucleic Acids Res.*, **48**, 12593-12603.
68. Joseph, J.A., Espinosa, J.R., Sanchez-Burgos, I., Garaizar, A., Frenkel, D. and Collepardo-Guevara, R. (2021) Thermodynamics and kinetics of phase separation of protein-RNA mixtures by a minimal model. *Biophys. J.*, **120**, 1219-1230.
69. Nguyen, H.T., Hori, N. and Thirumalai, D. (2022) Condensates in RNA repeat sequences are heterogeneously organized and exhibit reptation dynamics. *Nat. Chem.*, **14**, 775-785.
70. Liu, X.R. and Chen, J.H. (2017) HyRes: a coarse-grained model for multi-scale enhanced sampling of disordered protein conformations. *Phys. Chem. Chem. Phys.*, **19**, 32421-32432.
71. Zhang, Y., Liu, X. and Chen, J. (2022) Toward Accurate Coarse-Grained Simulations of Disordered Proteins and Their Dynamic Interactions. *J. Chem. Inf. Model.*, **62**, 4523-4536.
72. Bulacu, M., Goga, N., Zhao, W., Rossi, G., Monticelli, L., Periole, X., Tieleman, D.P. and Marrink, S.J. (2013) Improved Angle Potentials for Coarse-Grained Molecular Dynamics Simulations. *J. Chem. Theory Comput.*, **9**, 3282-3292.
73. Uusitalo, J.J., Ingólfsson, H.I., Akhshi, P., Tieleman, D.P. and Marrink, S.J. (2015) Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory Comput.*, **11**, 3932-3945.
74. Li, J. and Chen, S.-J. (2023) RNAJP: enhanced RNA 3D structure predictions with non-canonical interactions and global topology sampling. *Nucleic Acids Res.*, **51**, 3341-3356.
75. Plumridge, A., Andresen, K. and Pollack, L. (2020) Visualizing Disordered Single-Stranded RNA: Connecting Sequence, Structure, and Electrostatics. *J. Am. Chem. Soc.*, **142**, 109-119.
76. Draper, D.E. (2004) A guide to ions and RNA structure. *RNA*, **10**, 335-343.
77. Bowman, J.C., Lenz, T.K., Hud, N.V. and Williams, L.D. (2012) Cations in charge: magnesium ions in RNA folding and catalysis. *Curr. Opin. Struct. Biol.*, **22**, 262-272.
78. Draper, D.E., Grilley, D. and Soto, A.M. (2005) Ions and RNA Folding. *Annu. Rev. Biophys. Biomol. Struct.*, **34**, 221-243.
79. Yu, T. and Chen, S.-J. (2018) Hexahydrated Mg2+ Binding and Outer-Shell Dehydration on RNA Surface. *Biophys. J.*, **114**, 1274-1284.
80. Nguyen, H.T., Hori, N. and Thirumalai, D. (2019) Theory and simulations for RNA folding in mixtures of monovalent and divalent cations. *Proc. Natl. Acad. Sci. U. S. A.*, **116**, 21022-21030.
81. Xi, K., Wang, F.-H., Xiong, G., Zhang, Z.-L. and Tan, Z.-J. (2018) Competitive Binding of Mg2+ and Na+ Ions to Nucleic Acids: From Helices to Tertiary Structures. *Biophys. J.*, **114**, 1776-1790.

82. Misra, V.K. and Draper, D.E. (1999) The interpretation of Mg2+ binding isotherms for nucleic acids using Poisson-Boltzmann theory1 1Edited by B. Honig. *J. Mol. Biol.*, **294**, 1135-1147.
83. Jacobson, D.R. and Saleh, O.A. (2017) Counting the ions surrounding nucleic acids. *Nucleic Acids Res.*, **45**, 1596-1605.
84. Bai, Y., Greenfeld, M., Travers, K.J., Chu, V.B., Lipfert, J., Doniach, S. and Herschlag, D. (2007) Quantitative and Comprehensive Decomposition of the Ion Atmosphere around Nucleic Acids. *J. Am. Chem. Soc.*, **129**, 14981-14988.
85. Grilley, D., Soto, A.M. and Draper, D.E. (2006) Mg2+–RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 14003-14008.
86. Malmberg, C.G. and Maryott, A.A. (1956) Dielectric constant of water from 0 to 100 C. *J. Res. Natl. Bur. Stand.*, **56**, 1.
87. Ylitalo, A.S., Balzer, C., Zhang, P. and Wang, Z.-G. (2021) Electrostatic Correlations and Temperature-Dependent Dielectric Constant Can Model LCST in Polyelectrolyte Complex Coacervation. *Macromolecules*, **54**, 11326-11337.
88. Adhikari, S., Prabhu, V.M. and Muthukumar, M. (2019) Lower Critical Solution Temperature Behavior in Polyelectrolyte Complex Coacervates. *Macromolecules*, **52**, 6998-7004.
89. Singh, A.N. and Yethiraj, A. (2020) Driving Force for the Complexation of Charged Polypeptides. *J. Phys. Chem. B*, **124**, 1285-1292.
90. Eisenberg, H. and Felsenfeld, G. (1967) Studies of the temperature-dependent conformation and phase separation of polyriboadenylic acid solutions at neutral pH. *J. Mol. Biol.*, **30**, 17-37.
91. Pullara, P., Alshareedah, I. and Banerjee, P.R. (2022) Temperature-dependent reentrant phase transition of RNA–polycation mixtures. *Soft Matter*, **18**, 1342-1349.
92. Zeng, X., Liu, C., Fossat, M.J., Ren, P., Chilkoti, A. and Pappu, R.V. (2021) Design of intrinsically disordered proteins that undergo phase transitions with lower critical solution temperatures. *APL Materials*, **9**.
93. Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B. and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1-2**, 19-25.
94. Bauer, P., Hess, B. and Lindahl, E. (2022) GROMACS 2022 Manual. *Zenodo*.
95. Paul Bauer, Berk Hess and Lindahl, E. (2022) GROMACS 2022 Source code (Version 2022). *Zenodo*.
96. Case, D.A., Aktulga, H.M., Belfon, K., Cerutti, D.S., Cisneros, G.A., Cruzeiro, V.W.D., Forouzesh, N., Giese, T.J., Götz, A.W., Gohlke, H. *et al.* (2023) AmberTools. *J. Chem. Inf. Model.*, **63**, 6183-6191.
97. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.*, **117**, 5179-5197.
98. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926-935.
99. Zgarbová, M., Otyepka, M., Šponer, J., Mládek, A., Banáš, P., Cheatham, T.E., III and Jurečka, P. (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.*, **7**, 2886-2902.
100. Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham, T.E., Laughton, C.A. and Orozco, M. (2007) Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.*, **92**, 3817-3829.

101. Grotz, K.K., Cruz-León, S. and Schwierz, N. (2021) Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water-Exchange Properties. *J. Chem. Theory Comput.*, **17**, 2530-2540.
102. Mamatkulov, S. and Schwierz, N. (2018) Force fields for monovalent and divalent metal cations in TIP3P water based on thermodynamic and kinetic properties. *J. Chem. Phys.*, **148**.
103. Martínez, L., Andrade, R., Birgin, E.G. and Martínez, J.M. (2009) PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.*, **30**, 2157-2164.
104. Jo, S., Kim, T., Iyer, V.G. and Im, W. (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.*, **29**, 1859-1865.
105. Lee, J., Cheng, X., Swails, J.M., Yeom, M.S., Eastman, P.K., Lemkul, J.A., Wei, S., Buckner, J., Jeong, J.C., Qi, Y. *et al.* (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.*, **12**, 405-413.
106. Lee, J., Hitzenberger, M., Rieger, M., Kern, N.R., Zacharias, M. and Im, W. (2020) CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.*, **153**.
107. Darden, T., York, D. and Pedersen, L. (1993) Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089-10092.
108. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. and Pedersen, L.G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577-8593.
109. Bussi, G., Donadio, D. and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.*, **126**.
110. Nosé, S. and Klein, M.L. (1983) Constant pressure molecular dynamics for molecular systems. *Mol. Phys.*, **50**, 1055-1076.
111. Yoo, J. and Aksimentiev, A. (2012) Competitive Binding of Cations to Duplex DNA Revealed through Molecular Dynamics Simulations. *J. Phys. Chem. B*, **116**, 12946-12954.
112. Kirmizialtin, S., Silalahi, A.R.J., Elber, R. and Fenley, M.O. (2012) The Ionic Atmosphere around A-RNA: Poisson-Boltzmann and Molecular Dynamics Simulations. *Biophys. J.*, **102**, 829-838.
113. Kirmizialtin, S. and Elber, R. (2010) Computational Exploration of Mobile Ion Distributions around RNA Duplex. *J. Phys. Chem. B*, **114**, 8207-8220.
114. Eastman, P., Galvelis, R., Peláez, R.P., Abreu, C.R.A., Farr, S.E., Gallicchio, E., Gorenko, A., Henry, M.M., Hu, F., Huang, J. *et al.* (2024) OpenMM 8: Molecular Dynamics Simulation with Machine Learning Potentials. *J. Phys. Chem. B*, **128**, 109-116.
115. Walker, C.C., Meek, G.A., Fobe, T.L. and Shirts, M.R. (2021) Using a Coarse-Grained Modeling Framework to Identify Oligomeric Motifs with Tunable Secondary Structure. *J. Chem. Theory Comput.*, **17**, 6018-6035.
116. Shirts, M.R. and Chodera, J.D. (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.*, **129**.
117. Chen, H., Meisburger, S.P., Pabit, S.A., Sutton, J.L., Webb, W.W. and Pollack, L. (2012) Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 799-804.
118. Richards, E.G., Flessel, C.P. and Fresco, J.R. (1963) Polynucleotides. VI. Molecular properties and conformation of polyribouridylic acid. *Biopolymers*, **1**, 431-446.
119. Inners, L.D. and Felsenfeld, G. (1970) Conformation of polyribouridylic acid in solution. *J. Mol. Biol.*, **50**, 373-389.
120. Kapoor, U., Kim, Y.C. and Mittal, J. (2024) Coarse-Grained Models to Study Protein–DNA Interactions and Liquid–Liquid Phase Separation. *J. Chem. Theory Comput.*, **20**, 1717-1731.
121. de Mezer, M., Wojciechowska, M., Napierala, M., Sobczak, K. and Krzyzosiak, W.J. (2011) Mutant CAG repeats of Huntingtin transcript fold into hairpins, form nuclear foci and are targets for RNA interference. *Nucleic Acids Res.*, **39**, 3852-3863.

122. Theimer, C.A., Finger, L.D., Trantirek, L. and Feigon, J. (2003) Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 449-454.
123. Kankia, B.I. (2003) Binding of Mg2+ to single-stranded polynucleotides: hydration and optical studies. *Biophys. Chem.*, **104**, 643-654.
124. Tan, Z.-J. and Chen, S.-J. (2005) Electrostatic correlations and fluctuations for ion binding to a finite length polyelectrolyte. *J. Chem. Phys.*, **122**, 044903.
125. Tan, Z.-J. and Chen, S.-J. (2010) Predicting Ion Binding Properties for RNA Tertiary Structures. *Biophys. J.*, **99**, 1565-1576.
126. Pörschke, D. (1979) The mode of Mg ++ binding to oligonucleotides. Inner sphere complexes as markers for recognition? *Nucleic Acids Res.*, **6**, 883-898.
127. Nguyen, H.T. and Thirumalai, D. (2020) Charge Density of Cation Determines Inner versus Outer Shell Coordination to Phosphate in RNA. *J. Phys. Chem. B*, **124**, 4114-4122.
128. Onuchic, P.L., Milin, A.N., Alshareedah, I., Deniz, A.A. and Banerjee, P.R. (2019) Divalent cations can control a switch-like behavior in heterotypic and homotypic RNA coacervates. *Sci. Rep.*, **9**, 12161.
129. Kiliszek, A., Kierzek, R., Krzyzosiak, W.J. and Rypniewski, W. (2010) Atomic resolution structure of CAG RNA repeats: structural insights and implications for the trinucleotide repeat expansion diseases. *Nucleic Acids Res.*, **38**, 8370-8376.
130. Sobczak, K., Michlewski, G., de Mezer, M., Kierzek, E., Krol, J., Olejniczak, M., Kierzek, R. and Krzyzosiak, W.J. (2010) Structural Diversity of Triplet Repeat RNAs. *J. Biol. Chem.*, **285**, 12755-12764.
131. Lemieux, S. and Major, F. (2002) RNA canonical and non‐canonical base pairing types: a recognition method and complete repertoire. *Nucleic Acids Res.*, **30**, 4250-4263.
132. Mokdad, A., Krasovska, M.V., Sponer, J. and Leontis, N.B. (2006) Structural and evolutionary classification of G/U wobble basepairs in the ribosome. *Nucleic Acids Res.*, **34**, 1326-1341.
133. Abu Almakarem, A.S., Petrov, A.I., Stombaugh, J., Zirbel, C.L. and Leontis, N.B. (2012) Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res.*, **40**, 1407-1423.