Conformational compaction as a mechanism for ATP resolubilization of protein condensates

Zhou Gong (gongzhou@wipm.ac.cn)
State Key Laboratory of Magnetic Resonance and Atomic Molecular Physics, Innovation Academy for Precision Measurement Science and Technology Chinese Academy of Sciences https://orcid.org/0000-0001-6883-4026

Yueling Zhu
State Key Laboratory of Magnetic Resonance and Atomic Molecular Physics, Innovation Academy for Precision Measurement Science and Technology Chinese Academy of Sciences

Shiyan Lin
Zhejiang Province Key Laboratory of Quantum Technology and Device, School of Physics, Zhejiang University

Ling-Shen Meng
Peking University

Min Sun
Innovation Academy for Precision Measurement Science and Technology Chinese Academy of Sciences

Maili Liu
Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences https://orcid.org/0000-0002-9359-915X

Jingyuan Li
Zhejiang University https://orcid.org/0000-0003-2926-1864

Chun Tang
Peking University https://orcid.org/0000-0001-6477-6500

Article

Keywords: intrinsically disordered protein, phase separation, ATP, conformational compaction, HNRNPG

Posted Date: June 26th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3106412/v1

Abstract

ATP has been shown to regulate the phase separation behavior of intrinsically disordered proteins (IDPs), but a detailed mechanism remains to be fully established. Using the RG/RGG-rich motif from the HNRNPG protein as our model system, we show that the condensation of the IDP follows a biphasic relationship with the concentration of ATP. At a relatively low ATP concentration, ATP dynamically interacts with the protein and neutralizes surface charges, which promotes intermolecular interactions and favors phase separation. At the same time, ATP binding makes the protein more compact while enhancing local dynamics. As the ATP concentration increases, further compaction of the IDP hinders intermolecular interactions, and consequently prevents the protein from phase separation. We have thus identified IDP conformational compaction a mechanism for the ATP regulation of phase separation.

Introduction

Intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs) make up ~ 40% of human proteome. Unlike folded protein, IDPs are generally characterized with low-complexity sequences and have large proportion of charged residues, which prevents them from adopting compact conformations[1]. Unable to collapse and fold properly on its own, the IDPs are highly dynamic and conformationally heterogeneous, but can be stabilized when binding to folded proteins[2]. In the past ten years, it has been shown that IDPs are also important participants of macromolecular condensates through the processes of liquid-liquid phase separation (LLPS) [3, 4, 5]. Notable IDPs include those containing RG/RGG, Q/N, or polyQ motifs[6, 7, 8], which can readily afford multivalent intermolecular interactions.

A number of physicochemical factors, including temperature[9], salt[10], and macromolecular crowding[11], are known to modulate LLPS. Recently, it has been shown that ATP can reverse protein phase transition and forestalls the formation of protein fibrils and aggregates[12, 13], therefore enhancing protein solubility and functioning as a biological hydrotrope[14]. However, the role of ATP in regulating protein phase separation is likely more nuanced. At a relatively low concentration, the addition of ATP can promote the phase separation of the RG/RGG-rich IDPs derived from FUS and CAPRIN1 proteins[15, 16]. ATP can interact with protein residues through a range of interactions[16, 17, 18, 19], and a recent NMR study measuring near-surface electrostatic potential showed that ATP binding neutralizes the positive charges of CAPRIN1 and allows attractive the intermolecular interactions to take place[16]. It has also been shown that, at a relatively high concentration, the addition of ATP can reverse the electrostatic potential of CAPRIN1, introduce new electrostatic repulsions, and cause the dissipation of the protein droplets[16]. However, the binding of ATP towards the protein is very weak[18], let alone a charge-neutralized protein. In addition, from promoting to dissipating LLPS, the ATP concentration only varies by several fold, e.g., from 2 mM to 8 mM[12]. Thus, it is possible that additional factors may contribute to the solubilization of the phase-separated protein.

Heterogeneous nuclear ribonucleoprotein G (HNRNPG), also known as RBMX, is associated with nascent mRNA transcription, involved in alternative splicing of pre-mRNAs, and implicated in UV-damage response[20]. Importantly, HNRNPG is located in nuclear speckles[21, 22], a membraneless organelle formed

through phase separation[23]. An RG/RGG-rich motif is found near the C-terminus of HNRNPG, and is predicted to phase separate (Fig. 1a). In the current study, we show that, through an integrative use of solution NMR, small angle X-ray scattering (SAXS), and molecular dynamics (MD) simulations, ATP binding modulates the conformation of the RGG domain derived from HNRNPG. Importantly, we show that an extended-to-compact conformational transition hinders intermolecular interactions and disfavors protein phase separation.

Results

ATP modulates the phase-separating behavior of an RGG-rich IDP

We found that the RGG domain (residues 334–391) derived from HNRNPG protein does not phase separate on its own. The addition of 5 mM ATP lowered the threshold concentration and resulted in protein phase separation of the protein at a concentration above 250 μM. Based on GFP fluorescence, we estimated that the protein was enriched by about 130-fold in the droplets. However, increasing the concentration ATP eventually dissolved the proteins droplets (Fig. 1b). Thus, the modulatory effect of ATP on protein phase separation is not monotonic.

We doped the ATP with its fluorescent analog 2,4,6-trinitrophenol (TNP), and found the dye is co-localized with the Cy3-labeled RGG domain in the droplets (Fig. 1c). This has also been shown previously for the full length of FUS protein[12]. Thus, ATP co-phase separates with the protein, and likely directly interacts with the protein to exerts its modulatory effect. The liquid droplets are highly fluidic with the fluorescence of the labeled proteins in the droplets rapidly recovers to nearly 100% after photobleaching (Fig. 1d).

ATP makes the IDP more conformationally compact

To investigate the effect of ATP on the ensemble conformation of the RGG-rich IDP, we collected small angle X-ray scattering (SAXS) data. The experimental protein concentration was kept at ~ 200 μM, below the critical concentration of phase separation. The SAXS data indicated that the protein conformation becomes more compact with the addition of ATP, as the radius of gyration (R_g) decreases from ~ 20 Å to ~ 17 Å and end-to-end distance D_{max} decreases from ~ 70 Å to ~ 60 Å (Fig. 2a). Kratky plot for the SAXS data further showed that the protein changed from an extended conformation to a partially folded conformation with the addition of ATP (Fig. 2b). The translational diffusion coefficient obtained from NMR DOSY measurements confirmed a large increase of the diffusion rate with the addition of ATP. The viscosity may only change slightly with the addition of ATP, and therefore, the hydrodynamic radius of the protein decreases by more than 30% (Fig. 2c). The collapse of the protein conformation was further supported by all-atom MD simulations, which shows a narrower distribution of the R_g values in the presence of ATP (Fig. 2d).

We then performed NMR measurements for the HNRNPG-RGG. Resonance assignment for protein backbone was obtained using the standard triple resonance NMR spectroscopy (Supplementary Fig. 1).

The addition of 5- or 10-mM ATP to the protein causes little chemical shift perturbations (CSPs) of the backbone amide protons (Fig. 3a). Nevertheless, the intensities for many peaks increase in the presence of ATP (Fig. 3b, and Supplementary Fig. 2). Interestingly, most of these peaks are from Ser and Gly residues that have small or no side chains.

To understand which factors contribute to the increased NMR peak intensities, we measured solvent-exchange rates of amide protons using CLEANEX-PM pulse sequence [24]. The exchange rates were found 10 and 20 s^{-1} for the protein alone, but decreased by about two-fold with the addition of ATP. For several residues, the solvent exchange rates are already slowed with the addition of 5 mM ATP, while no further reduction was observed with the addition of 10 mM ATP (Fig. 3c, and Supplementary Fig. 3, Supplementary Fig. 4). As the pH of the protein solution is well maintained, the decrease of the exchange rates can only be resulted from the reduction of solvent accessibility, which can be attributed to the direct protection from ATP binding and/or conformational collapse of the protein.

HNRNPG-RGG remains highly dynamic upon ATP association

ATP interacts with the RGG domain of HNRNPG protein electrostatically. Indeed, our all-atom MD simulations showed that the phosphate groups of ATP molecules are preferentially associated with Arg residues (Fig. 4a, b). Moreover, the nucleobase of ATP transiently interacts with the backbone and sidechain atoms of neighboring Ser and Gly residues through hydrogen bonds (Fig. 4b). Thus, both local protection by the ATP molecules and global compaction of the protein would account for the decrease of solvent-exchange rates for these residues.

MD simulation also showed that the number of the protein-associated ATP molecules, i.e., within 6 Å vicinity of the protein, fluctuates. Nevertheless, in a simulated system containing five ATP molecules, the proportion of all five ATP molecules simultaneously associated with the protein is only 10% (Fig. 4c). This means that the interaction between ATP and the RGG domain is weak and dynamic, which explains why a 20-fold excess of the ATP (Fig. 1a) is required in order to effectively neutralize protein charges and promote protein phase separation. As such, despite of ATP coating, the compaction of the IDP plays a dominant role and accounts for the decreased hydrodynamic radius (Fig. 2c).

Despite the global compaction of the IDP, ATP binding also causes changes to the *ps-ns* dynamics of the protein at the same time. Increased local dynamics is consistent with transient ATP associations, and would make the overall conformational change more energetically favorable. We measured the longitudinal R_1 and transverse relaxation R_2 rates for the backbone amide nitrogen atoms of HNRNPG RGG domain (Fig. 5). The R_2 rates decrease in the presence of 10 mM ATP, from 3.0 ± 0.2 s^{-1} to 2.6 ± 0.2 s^{-1}. On the other hand, the R_1 rates increase by a larger relative amount, from 1.5 ± 0.2 to 2.1 ± 0.2 s^{-1}, in the presence of 10 mM ATP. R_1 is a function of spectral density function $J(\omega)$, while R_2 mainly depends on the $J(0)$ term of spectral density function. With only a global compaction of the protein, R_2 should exhibit a larger relative change in comparison to R_1. R_1 is more sensitive to fast dynamics at *ps*-timescale than R_2,

whereas R_2 becomes almost as small as R_1, approaching the extreme narrowing regimen of ^{15}N nuclear relaxation. Thus, an enhanced local dynamic of the protein needs to be invoked to explain the larger relative increase of R_1 rates in the presence of ATP, akin to what usually observed at protein N- or C-terminal residues.

The RGG motif is different from other low-complexity IDPs that phase separate

The addition of ATP lowers the threshold of the IDP from HNRNPG RGG to phase separate. For the liquid droplets prepared with 300 μM protein and 10 mM ATP, we estimated the concentration of HNRNPG RGG in condensed phase is 18.6 mM (~ 117 mg/mL), based on the relative peak intensities in ^1H NMR spectra. This amounts to over a hundred-fold enrichment (Fig. 6), close to estimation based on GFP fluorescence. However, the value is several-fold lower than the concentration expected for entangled random-coil polymers in the condensed phase[25], and is also lower than the values reported for other proteins. For example, the low-complexity domains from FUS, DDX4, hnRNPA2 and ELP3, were estimated at 477 mg/L, 380 mg/mL, 440 mg/mL, and 500 mg/mL, in their respective condensed phase[6, 26, 27, 28]. Moreover, from our estimation, water molecules still make up ~ 88% volume of the liquid droplets formed by HNRNPG-RGG. As such, the IDP adopts compact conformations in the condensed phase, which prevents further compaction and also accounts for the high mobility of the IDP (Fig. 1d).

To further our comparison, we prepared a protein construct from the N-terminal prion-like domain (PLD) of FUS (Fig. 7a), and found that that addition of ATP caused little NMR chemical shift perturbations or signal enhancement (Fig. 7b, 7c and Supplementary Fig. 5). Moreover, DOSY measurements showed that the addition of ATP actually slows the translational diffusion of FUS-PLD (Fig. 7d). The FUS-PLD readily phase-separates, and can eventually assembles into fibrillar structure. It has been reported that ATP functions as a hydrotrope to dissipate the gel-like, solid, and fibrillar protein condensates in a monotonic fashion[29]. The FUS-PLD is rich in Gly and Ser residues, but has no charged residues. The results of MD simulations indicate that ATP tends to form hydrogen bonds with the side chain of Gln, Ser and Thr on FUS-PLD (Supplementary Fig. 6). Therefore, the interactions between ATP and FUS-PLD should be inherently different from those for HNRNPG-RGG [15, 29].

Discussion

ATP and ATP-Mg^{2+} has been shown as hydrotrope that helps to stabilize aggregation-prone proteins in the diluted phase[12]. However, how ATP regulates protein phase-separation behavior remains to be fully understood. At an atomic level, the hydrophobic base of ATP can interact with protein residues through hydrogen bonding, π-π stacking, and NH-π interactions[30, 31, 32], while the charged phosphate group can interact with charged or polar residues through hydrogen bonding and salt bridges[16, 33]. We have shown that the ATP promotes the phase separation of an RG/RGG-rich protein at a relatively low ATP concentration, but dissolves the liquid droplets and homogenizes the protein solution at a relatively high

ATP concentration. We have further shown that such biphasic modulatory effect of the ATP is resulted from protein charge neutralization and conformational compaction.

The RGG-rich IDP is positively charged, and the protein tends to adopt extended and dynamic conformations. Electrostatic repulsion also prevents intermolecular interactions (stage I in Fig. 8). The transient association of the ATP molecules effectively neutralize the electrostatic potential, as recently shown for another RGG-rich protein[16]. As a result, two or more copies of the IDP molecules can come close to each other and interact when shieled by ATP molecules (stages II and III in Fig. 8). The situation is likely different for the FUS and other prion-like domains, which contains more hydrophobic residues and fewer charged residues. ATP likely uses its aromatic moiety to interacts with FUS, which explains the increased hydrodynamic radius (Fig. 7d), and consequently, the additional negative charges from ATP coating prevents FUS molecules interact with each other and phase separate.

The addition of ATP also modulates the ensemble conformation of the IDP, making the protein more compact, akin to a molten globule. The conformational compaction was evidenced from SAXS, NMR, and MD analyses. Such compaction is resulted from the electrostatic interaction between ATP and the RGG/RG-rich protein, which are unique for the positively charged IDPs. With phosphate moiety transiently shielded, the hydrophobic moieties of the ATP also afford additional long-range contacts. The transient association of ATP molecules and conformational compaction of protein accounts for the increased protection of backbone amides from the solvent. With an increasing amount of ATP molecules (stage IV in Fig. 8), the IDP would become so compact that precludes intermolecular interactions. As a result, the protein become once again homogenized.

Materials and methods

Sample preparation.

The HNRNPG RGG domain (L334-Y391) was cloned to pET11a vector (GenBank:CAG33028.1). The N-terminal of the RGG domain connects the GB1 protein through the cleavage site of tobacco etch virus (TEV) protease. The fusion protein was expressed in BL21 Star (DE3) cells. The *E. coli* bacteria were cultured in minimal M9 medium or LB medium to prepare isotope-enriched or unlabeled proteins. For preparing the isotope-labeled protein, 1g/L U-^{15}N-labeled NH_4Cl (Cambridge Isotope Laboratories) and/or 2 g/L U-^{13}C-labeled glucose (Sigma Aldrich) were added to the minimal M9 medium. The protein was induced at OD_{600} of 0.8 with IPTG at a final concentration of 0.25 mM for 16 h at 23°C. The protein was purified with Sepharose SP (GE Healthcare), Sephacryl S100 (GE Healthcare), and Source-S columns (GE Healthcare). The protein was buffer exchanged to TEV protease restriction buffer containing 50 mM Tris, 1 mM EDTA, and 1 mM DTT with pH 7.6. The GB1 protein was removed with TEV protease at 25°C for an hour. The RGG domain was purified with the second round of Source-S column (GE Healthcare) and then concentrated and buffer exchanged in Amicon Ultra (Millipore). The final sample was prepared in 20 mM NaH_2PO_4 and 20 mM NaCl at pH 6.8.

For the ATP-Mg solution preparation, ATP (ATP·Na_2, BBI, cat No.987-65-5) was dissolved in water and add NaOH to adjust pH to 6.8 at final concentration of 500 mM. $MgCl_2$ (Magnesium chloride hexahydrate, Diamond, cat No.7792-28-6) was dissolved in water at final concentration of 500 mM. The two solutions were mixed in the ratio of 1:1.

The FUS prion-like domain (PLD, M1-S165) was cloned to a pET11a vector (GenBank: CAG33028.1). The protein connects the His tag and GB1 protein through the cleavage site of tobacco etch virus (TEV) protease. The protein was induced at OD_{600} of 0.8 with IPTG at a final concentration of 0.25 mM for 16 h at 23℃ using BL21 Star (DE3). The protein was purified with HisTrap™ FF (GE Healthcare). The protein was buffer exchanged to TEV protease restriction buffer containing 50 mM Tris, 1 mM EDTA, and 1 mM DTT with pH 7.6. The His tag and GB1 protein was removed with TEV protease at 25℃ for an hour. The PLD was purified with of Source-Q column (GE Healthcare) and then concentrated and buffer exchanged in Amicon Ultra (Millipore). The final sample was prepared in 50 mM MES and 250 mM NaCl at pH 5.5.

The GFP-tag HNRNPG RGG domain was cloned to a pET11a vector. The N-terminal of protein connects the His tag through the flexible linker GSGSGS. The protein was expressed at OD_{600} of 0.8 with IPTG at a final concentration of 0.25 mM for 16 h at 23℃ using BL21 Star (DE3). The protein was purified with HisTrap™ FF (GE Healthcare) and change the buffer in 20 mM NaH_2PO_4 and 20 mM NaCl at pH 6.8.

Fluorescent and paramagnetic labeling

Cyanine 3 maleimide (Cy3, AAT Bioquest, cat No. 142)-labeled sample was prepared for fluorescence experiments. The Cy3 was connected to S386C on the RGG domain. The mutant protein was prepared and purified according to the method described above. A total of 1.5 times Cy3 was added into the protein under the buffer containing 50 mM HEPES and 150 mM NaCl at pH 7.4. The excess probe was then removed by a desalting column (GE Healthcare). TNP-ATP triethylammonium salt (TNP-ATP triethylammonium salt, APExBIO Technology, cat No. B7066) was dissolved in water and add NaOH to adjust pH to 6.8 at final concentration of 500 mM, and then mixed with 500 mM $MgCl_2$ in the ratio of 1:1.

Phase-separated NMR samples

The phase-separated was induced by 300 μM ^{15}N-labeled RGG domain mixed with 10 mM ATP in a centrifuge glass tube and centrifuged at 7000 rpm for 10 min in 7 ℃. The phase separation will form tiny droplets that coalesce into the large droplet. The large droplets were transferred to the NMR sample tubes (NORELL®, NI5CCI-B). The above steps will be repeated several times until enough NMR samples are obtained (40 ml 300 μM ^{15}N-labeled RGG domain sample was consumed).

Differential interference contrast (DIC) microscopy

The formation of protein droplets was observed at 25℃ with different RGG domain concentrations at 100, 200, 250, and 300 μM in 20 mM NaH_2PO_4 buffer at pH 6.8 containing 20 mM NaCl in the absence and presence of ATP-Mg complex at 0, 1, 5, 10, 15, 20 and 30 mM. After the sample is mixed in advance, it will be kept at room temperature for 30 minutes, and then the sample will be transferred to the self-made

chamber. The chamber will then be observed under Nikon A1 confocal laser-scanning microscope under a 60× oil objective (Nikon, Tokyo, Japan).

Fluorescence Recovery After Photobleaching (FRAP).

FRAP experiments were performed using the Nikon A1 confocal laser-scanning microscope under a 60× oil objective (Nikon, Tokyo, Japan). Samples were prepared using 300 μM RGG protein and 10 mM ATP-Mg with 5% of RGG protein with GFP tag. The droplet is observed as above. A circular region of interest (ROI) was drawn at the position of droplet and bleached with 100% laser power for 60 s at 488 nm. The mean fluorescence intensity from the ROI was collected until recovery was complete.

Since the mean fluorescence intensity of droplet from ROI can be recorded. The dilute of RGG protein with GFP tag was prepared at certain concentration (100 μM) under the same laser parameter setting, and the mean fluorescence intensity was also recorded. The concentration of protein in condensed phase can be calculated as the following equations:

$$C_{cond} = C_{dilute} \times \frac{I_{cond}/5\%}{I_{dilute}} \# (1)$$

C_{cond} and C_{dilute} corresponding to the protein concentration in droplet and in dilute. I was the mean fluorescence intensity, and proportion 5% indicate that only 5% of protein involved in phase separation was fused of GFP.

NMR experiments

The protein sample for the NMR experiment was in the same buffer condition described above with a 10% D_2O addition. All NMR experiments were performed at 298K on Bruker 600 MHz and 700 MHz spectrometers equipped with cryogenic probes. The NMR data were further processed using TopSpin 3.5 (Bruker), NMRPipe 2016[34], and CCPNmr Analysis V2.4 [35], respectively. A series of NMR experiments were acquired for backbone assignment, including ^{1}H-^{15}N HSQC, HNCACB, HNCA, and HNCO.

NMR titration experiment. The initial ^{15}N-labeled protein sample was prepared as 200 μM, and ATP and $MgCl_2$ were dissolved in the same buffer described above (adjust pH using NaOH). The ^{1}H-^{15}N HSQC spectrums were recorded with the addition of ATP-Mg at 0, 5 mM, and 10 mM at the final concentration. The amide CSP were calculated by the equation:

$$CSP = \sqrt{0.5 \times \Delta\delta H^2 + 0.1 \times \Delta\delta N^2} \# (2)$$

$\Delta\delta H$ and $\Delta\delta N$ are the perturbation with ATP added in proton and nitrogen dimensions in ppm unit, respectively.

Water-exchange experiment. Water-selective experiments were performed to study exchange processes using CLEANEX-PM[24]. The experiments were recorded using a standard pulse sequence in Topspin 3.5 (fhsqccxf3gpph). The mixing time τ_m was set at 0 ms, 10 ms, 15 ms, 20 ms, and 30 ms. The ATP was

added to the protein sample at 0, 5 mM, and 10 mM at the final concentration. The peak volumes with different mixing times and concentrations of ATP were processed, and exchange rates were fitted using the same algorithm described before[36].

Diffusion coefficient measurement. The pulsed field gradient NMR diffusion experiments on RGG and PLD with or without ATP were performed as pseudo-2D experiments using the standard Bruker pulse sequence with a gradient strength from 5–95%. The data were analyzed by integrating resonance corresponding to the side chains in RGG and PLD. The diffusion coefficient was fitted as the following equation:

$$I_G = I_0 \exp\left[-(2\pi\gamma\sigma G)^2 D\left(\triangle - \frac{\sigma}{3}\right) \times 10^4\right] \# (3)$$

Due to the more complex dynamic variations of FUS PLD, a single exponential function cannot accurately fit the diffusion coefficient. Therefore, we employed a dual exponential function to fit the diffusion coefficient as the following equation:

$$I_G = I_0 \exp\left[-(2\pi\gamma\sigma G)^2 D1\left(\triangle - \frac{\sigma}{3}\right) \times 10^4\right] + I_0 \exp\left[-(2\pi\gamma\sigma G)^2 D2\left(\triangle - \frac{\sigma}{3}\right) \times 10^4\right] \# (4)$$

I_0 is the signal intensity at a gradient strength of zero, G is the gradient strength, D is the diffusion coefficient, σ is the gradient pulse duration, and $\triangle$ is the diffusion time.

NMR relaxation experiment. The ^{15}N labeled protein was prepared as 200 μM and 10 mM ATP was added further. Motions of the backbone of RGG were measured at 600 MHz using standard pulse sequences (hsqct1etf3gpsi.2, hsqct2etf3gpsi). Delay for gradient recovery was set to 16.96, 169.6, 339.2, 508.8ms for R2 experiments, and 20, 200, 400, 600 and 700 ms for R1 experiments, respectively.

SAXS experiment

All SAXS experiments of the RGG domain and ATP were collected at the BL19U2 beamline in National Center for Protein Science Shanghai. We used a 1-second exposure time for each time at 25 ºC, and 20 frames were recorded and averaged for further analysis. The scattering data for the corresponding buffer and the substrate from the sample data were also recorded. The theoretical scattering curve was calculated from the related PDB file using CRYSOL modules[37] in the ATSAS 2.8 software package[38]. The paired distance distribution function (PDDF) was calculated from the scattering curve using the PRIMSQT module in ATSAS 2.8

The overlap concentration calculation

The overlap concentration C* of the RGG domain from dilute to semi-dilute regimes was calculated by the Eq. 3[9]:

$$C^* \cong M / \left[N_A (h_0/2)^3\right] \# (5)$$

where M and h_0 are the protein molecular weight and the root-mean-square end-to-end distance of the macromolecular coil, respectively. N_A is Avogadro constant. h_0 was calculated by the equation:

$$h_0{}^2 = 6R_g{}^2 \# (6)$$

Rg is the radius of gyration from the SAXS experiment (average about 18 Å). The protein molecular weight (M = 6299.79) was used in the calculation. According to these equations, the overlap concentration C^* of the RGG domain from dilute to semi-dilute regimes is 972 mg/ml.

MD simulation and structure analysis

The RGG-rich C-RBD domain of HNRNPG and the N-terminal prion-like domain (PLD) of FUS were constructed by PyMOL (The PyMOL Molecular Graphics System, Version 2.3 Schrödinger, LLC). The Nϵ group in Arg was set to the protonated state, and the carboxyl deprotonated state was employed for each Asp and Glu. The constructed peptide chain was relaxed in a vacuum for 100 *ps*, then solvated by water molecules, and subjected to 100 ns relaxation simulation. The resulting conformation was used as the initial structure for the following simulations.

The single protein was initially placed in a cubic box and solvated by water molecules; the box side length was chosen according to the protein size, namely 10.0 and 12.0 nm for RGG domain and PLD, respectively. Sodium and chloride ions were added to neutralize the systems and mimic physiological conditions (150 mM NaCl). The systems underwent a series of equilibrium procedures, including 50000-step energy minimization, 10-*ns* NVT simulation, and 10-*ns* NPT simulation for temperature relaxation. 1000-*ns* production runs were performed as control systems for the following analysis.

To study the interactions between ATP and the proteins, we added ATP and Mg^{2+} molecules to the control systems mentioned above. The number of the molecules corresponded to the concentration of 8 mM (5 and 8 for RGG domain and PLD, respectively). A similar equilibrium procedure was conducted for each system, and a 1000-*ns* production run was performed and used for the following analyses.

All simulations were run using the package GROMACS 2018 t. Amber99sb-ildn[40] force field was used to describe the interactions of the proteins and ATPs. TIP4P-D water model was adopted. The boxes were set with periodic boundary conditions. The equations of motion were numerically integrated using the Verlet leapfrog algorithm with a time step of 2 *fs*. The Particle Mesh Ewald (PME) [41] method was used to handle the long-range electrostatic interactions with cubic interpolation and a grid spacing of 0.16 nm. The short-range interactions were calculated using a cutoff of 1.2 nm. The systems were coupled to temperature baths at 300 K, with the V-rescaling thermostat and a relaxation time of 0.1 *ps*. The Parrinello-Rahman pressure coupling was used at 1.0 bar, with a relaxation time of 2 *ps*. The protein structures were rendered using PyMOL and the visual molecular dynamics (VMD) 1.9.3 program[42].

The criterion for the binding of ATP to the protein is the distance between the heavy atoms of ATP and the heavy atoms of protein is less than 6 Å.

Declarations

Acknowledgments

We thank the staff of BL19U2 beamlines at the National Facility for Protein Science Shanghai (NFPS) and Shanghai Synchrotron Radiation Facility, Shanghai, People's Republic of China, for assistance during SAXS data collection. The work has been supported by the National Key R&D Program of China (2018YFA0507700), the National Natural Science Foundation 31971155 and 21991081), and the Youth Innovation Promotion Association of the Chinese Academy of Sciences under grant No. 2020329.

Author contributions

Z.G, C.T, J-Y.L. and M-L.L. designed and guided the study, analyzed data and wrote the manuscript. Y-L.Z. purified proteins, performed NMR experiments, analyzed the data. S-Y.L. performed MD simulations, analyzed the structure data. M.S and L-S.M helped with the NMR experiments and data analysis.

Competing interests

The authors declare no competing interests.

Data Availability

The data supporting the findings of this study are available from the corresponding authors upon reasonable request.

References

1. Habchi J, Tompa P, Longhi S, Uversky VN. Introducing Protein Intrinsic Disorder. *Chemical Reviews*, **114**, 6561-6588 (2014)..
2. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* **16**, 18-29 (2015).
3. Sanders DW*, et al.* Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. *Cell* **181**, 306-324.e328 (2020).
4. Dignon GL, Best RB, Mittal J. Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. *Annu Rev Phys Chem* **71**, 53-75 (2020).
5. Wiedner HJ, Giudice J. It's not just a phase: function and characteristics of RNA-binding proteins in phase separation. *Nat Struct Mol Biol* **28**, 465-473 (2021).
6. Ryan VH*, et al.* Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol Cell* **69**, 465-479 (2018).
7. Alshareedah I*, et al.* Interplay between Short-Range Attraction and Long-Range Repulsion Controls Reentrant Liquid Condensation of Ribonucleoprotein-RNA Complexes. *J Am Chem Soc* **141**, 14593-14602 (2019).

8. Chong PA, Vernon RM, Forman-Kay JD. RGG/RG Motif Regions in RNA Binding and Phase Separation. *J Mol Biol* **430**, 4650-4665 (2018).
9. Dignon GL, Zheng W, Kim YC, Mittal J. Temperature-Controlled Liquid-Liquid Phase Separation of Disordered Proteins. *ACS Cent Sci* **5**, 821-830 (2019).
10. Krainer G*, et al.* Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat Commun* **12**, 1085 (2021).
11. Ghosh A, Mazarakos K, Zhou HX. Three archetypical classes of macromolecular regulators of protein liquid-liquid phase separation. *Proc Natl Acad Sci U S A* **116**, 19474-19483 (2019).
12. Patel A*, et al.* ATP as a biological hydrotrope. *Science* **356**, 753-756 (2017).
13. Mehringer J*, et al.* Hofmeister versus Neuberg: is ATP really a biological hydrotrope? *Cell Reports Physical Science* **2**, 100343 (2021).
14. Sridharan S*, et al.* Proteome-wide solubility and thermal stability profiling reveals distinct regulatory roles for ATP. *Nat Commun* **10**, 1155 (2019).
15. Kang J, Lim LZ, Song JX. ATP enhances at low concentrations but dissolves at high concentrations liquid-liquid phase separation (LLPS) of ALS/FTD-causing FUS. *Biochem Biophys Res Commun* **504**, 545-551 (2018).
16. Toyama Y, Rangadurai AK, Forman-Kay JD, Kay LE. Mapping the per-residue surface electrostatic potential of CAPRIN1 along its phase-separation trajectory. *Proc Natl Acad Sci U S A* **119**, e2210492119 (2022).
17. Ou X*, et al.* ATP Can Efficiently Stabilize Protein through a Unique Mechanism. *JACS Au* **1**, 1766-1777 (2021).
18. Nishizawa M*, et al.* Effects of Weak Nonspecific Interactions with ATP on Proteins. *J Am Chem Soc* **143**, 11982–11993 (2021).
19. Ren CL, Shan Y, Zhang P, Ding HM, Ma YQ. Uncovering the molecular mechanism for dual effect of ATP on phase separation in FUS solution. *Sci Adv* **8**, eabo7885 (2022).
20. Xiang Y*, et al.* RNA m6A methylation regulates the ultraviolet-induced DNA damage response. *Nature* **543**, 573-576 (2017).
21. Adamson B, Smogorzewska A, Sigoillot FD, King RW, Elledge SJ. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat Cell Biol* **14**, 318-328 (2012).
22. Heinrich B*, et al.* Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA. *J Biol Chem* **284**, 14303-14315 (2009).
23. Galganski L, Urbanek MO, Krzyzosiak WJ. Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res* **45**, 10350-10368 (2017).
24. Hwang TL, Mori S, Shaka AJ, vanZijl PCM. Application of phase-modulated CLEAN chemical EXchange spectroscopy (CLEANEX-PM) to detect water-protein proton exchange and intermolecular NOEs. *J Am Chem Soc* **119**, 6203-6204 (1997).

25. Ying QC, Chu B. Overlap Concentration of Macromolecules in Solution. *Macromolecules* **20**, 362-366 (1987).
26. Murthy AC*, et al.* Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat Struct Mol Biol* **26**, 637-648 (2019).
27. Brady JP*, et al.* Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc Natl Acad Sci U S A* **114**, E8194-E8203 (2017).
28. Reichheld SE, Muiznieks LD, Keeley FW, Sharpe S. Direct observation of structure and dynamics during phase separation of an elastomeric protein. *Proc Natl Acad Sci U S A* **114**, E4408-E4415 (2017).
29. Kang J, Lim L, Lu Y, Song J. A unified mechanism for LLPS of ALS/FTLD-causing FUS as well as its modulation by ATP and oligonucleic acids. *PLoS Biol* **17**, e3000327 (2019).
30. Shetty AS, Zhang JS, Moore JS. Aromatic pi-stacking in solution as revealed through the aggregation of phenylacetylene macrocycles. *J Am Chem Soc* **118**, 1019-1027 (1996).
31. Mao LS, Wang YL, Liu YM, Hu XC. Molecular determinants for ATP-binding in proteins: A data mining and quantum chemical analysis. *J Mol Biol* **336**, 787-807 (2004).
32. Meyer EA, Castellano RK, Diederich F. Interactions with aromatic rings in chemical and biological recognition. *Angew Chem Int Ed Engl* **42**, 1210-1250 (2003).
33. Heo CE*, et al.* ATP Kinetically Modulates Pathogenic Tau Fibrillations. *ACS Chem Neurosci* **11**, 3144-3152 (2020).
34. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).
35. Vranken WF*, et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696 (2005).
36. Hwanga T-L, C.M P, Zijl v, Moria S. Accurate quantitation of water–amide proton exchange rates using the Phase-Modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J Bioml NMR* **11**, 221-226 (1998).
37. al DSe. CRYSOL-a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Cryst* **28**, 768-773 (1995).
38. Franke D*, et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Cryst* **50**, 1212-1225 (2017).
39. Ying Q, Chu B. Overlap concentration of macromolecules in solution. *Macromolecules* **20**, 362-366 (1987).
40. Robustelli P, Piana S, Shaw DE. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A* **115**, E4758-E4766 (2018).
41. Darden T, York D, Pedersen L. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **98**, 10089-10092 (1993).
42. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, (1996).

Figures

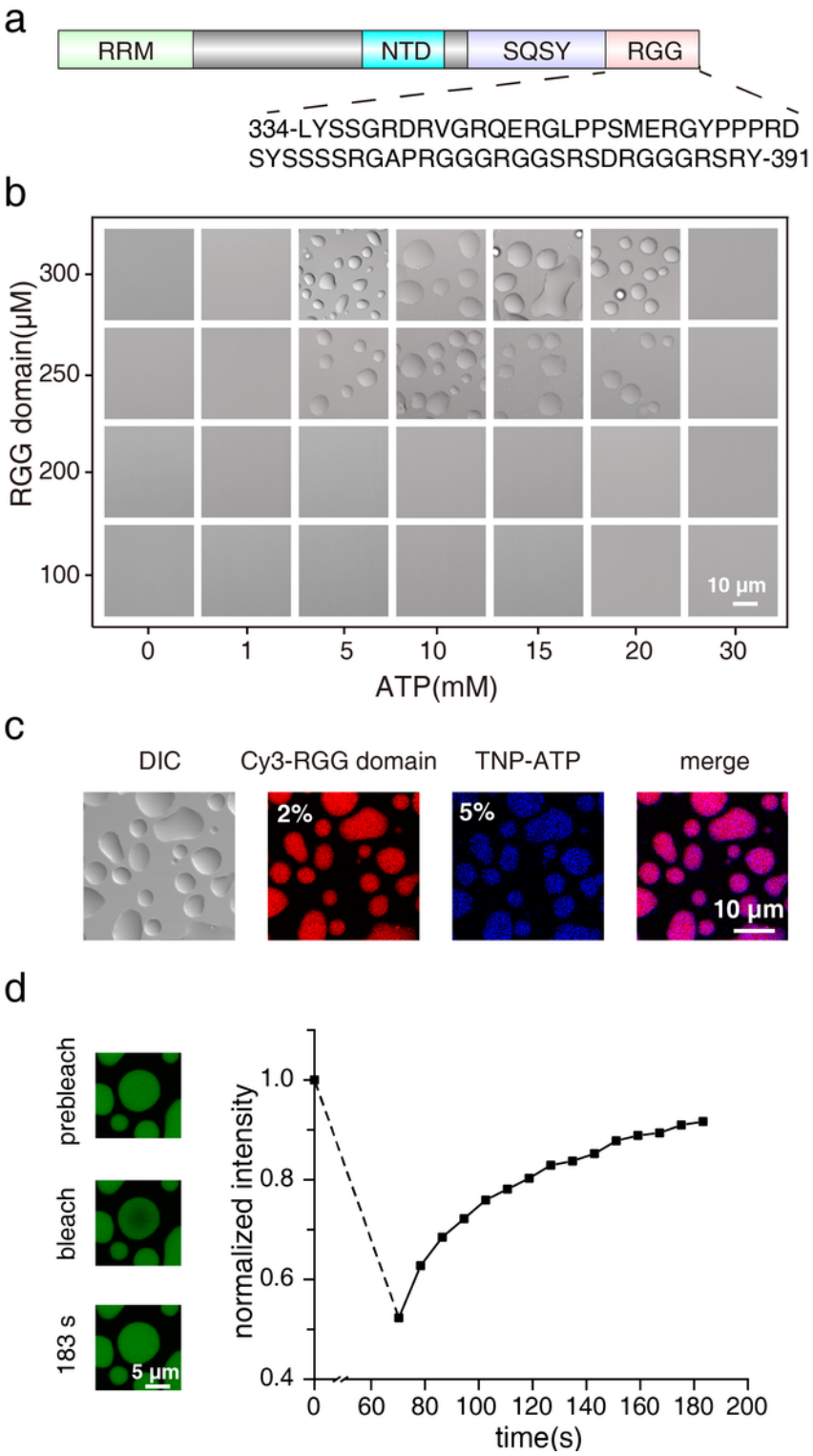

Figure 1

ATP modulates the phase separation of the RGG domain derived from HNRNPG.

a The construct of HNRNPG used in the current study, with the sequence of C-terminal RGG domain shown. **b** The DIC micrographs showed that the IDP could form liquid droplets only when the ATP concentration is in the range of 5-20 mM. **c** ATP co-localizes with the IDP in the droplets. Here Cy3-labeled protein (red) and ATP analog TNP-ATP (blue) were doped into the unlabeled protein at 2% and 5% proportions, respectively. **d**

Fluorescence recovery after photobleaching (FRAP) shows that the IDP is highly mobile in the liquid droplet. A GFP protein was fused to the N-terminus of the RGG domain.

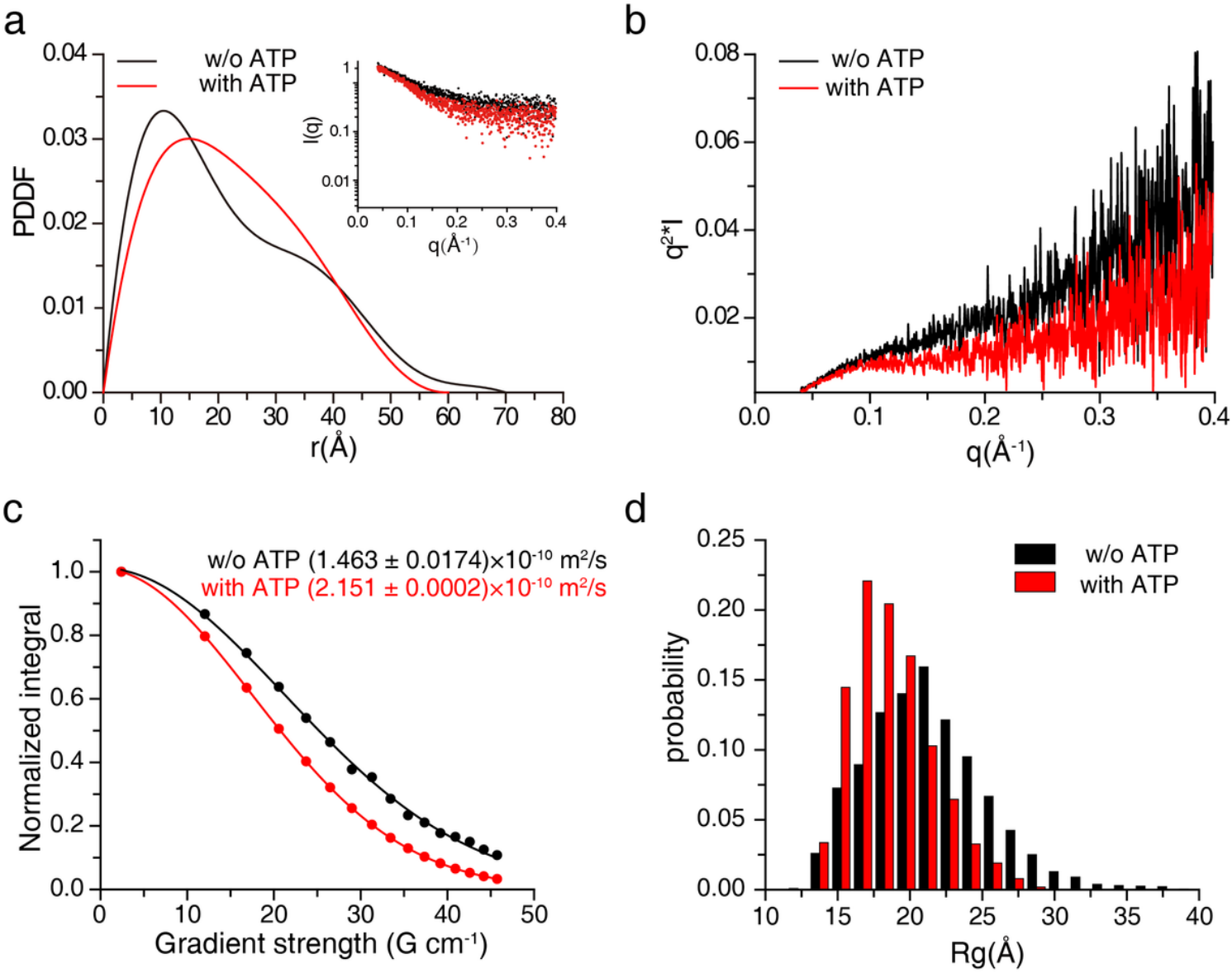

Figure 2

The addition of ATP collapses the ensemble conformation of HNRNPG-RGG.

a The paired distance distribution function (PDDF) in the absence (black) and presence of 5 mM ATP (red), experimentally derived from small angle X-ray scattering data (inset). **b** The Kratky plot was calculated from the scattering curve of the RGG domain in the absence (black) and presence of 5 mM ATP binding (red). **c** NMR DOSY measurement affords the translation diffusional coefficients. **d** The distribution of calculated radius of gyration (R_g) of the protein, obtained from MD simulations.

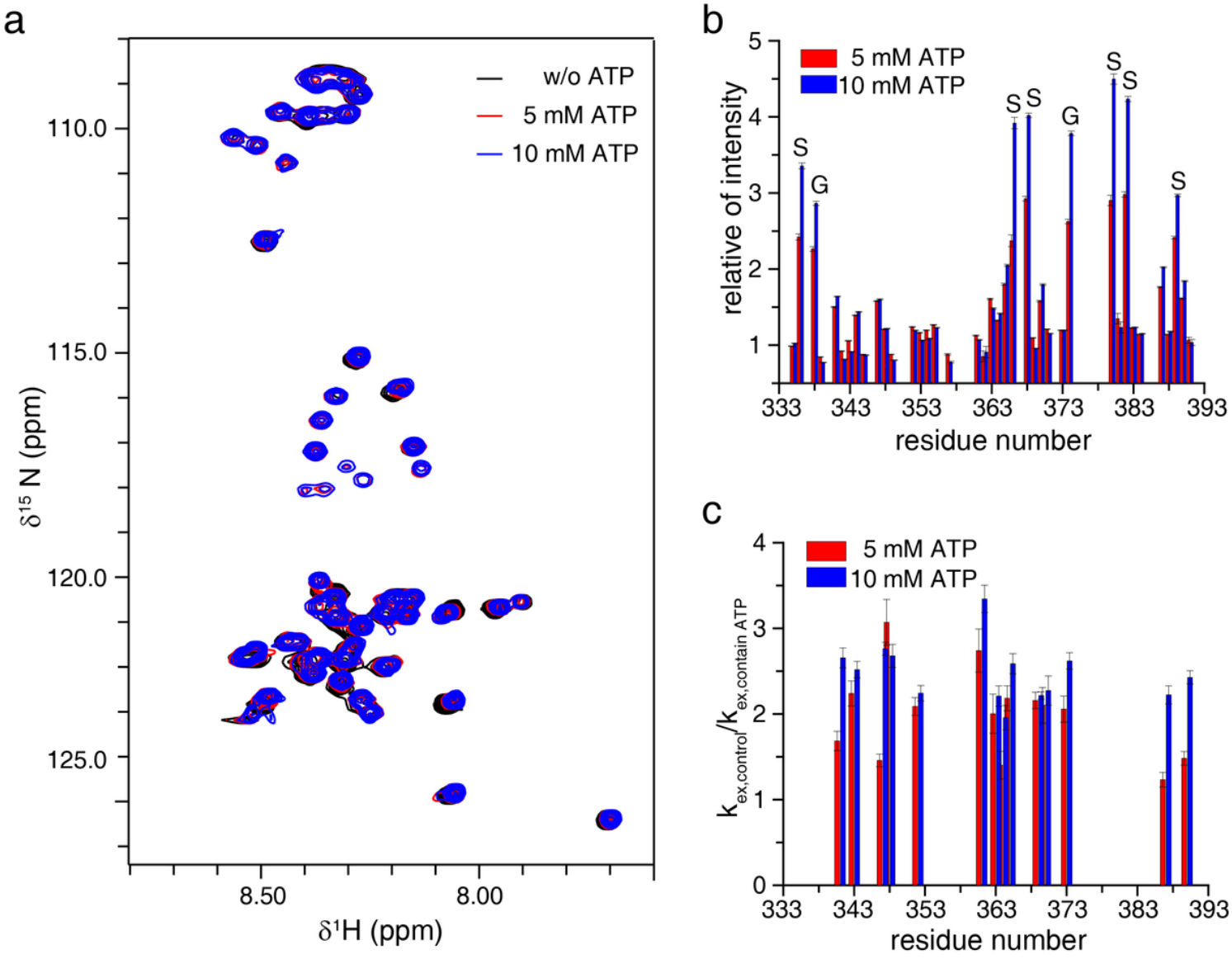

Figure 3

Solvent exchange of the IDP slows in the presence of ATP.

a Overlay of ^{1}H-^{15}N HSQC spectra of HNRNPG-RGG in the absence (black) and presence of 5 mM ATP (red) or 10 mM ATP (blue). **b** The ratios of peak intensities were calculated for the backbone amide protons in the presence and absence of ATP. **c** The ratios of solvent exchange rates k_{ex}, were measured by CLEANEX-PM scheme, were calculated for the backbone amide protons in the presence and absence of ATP.

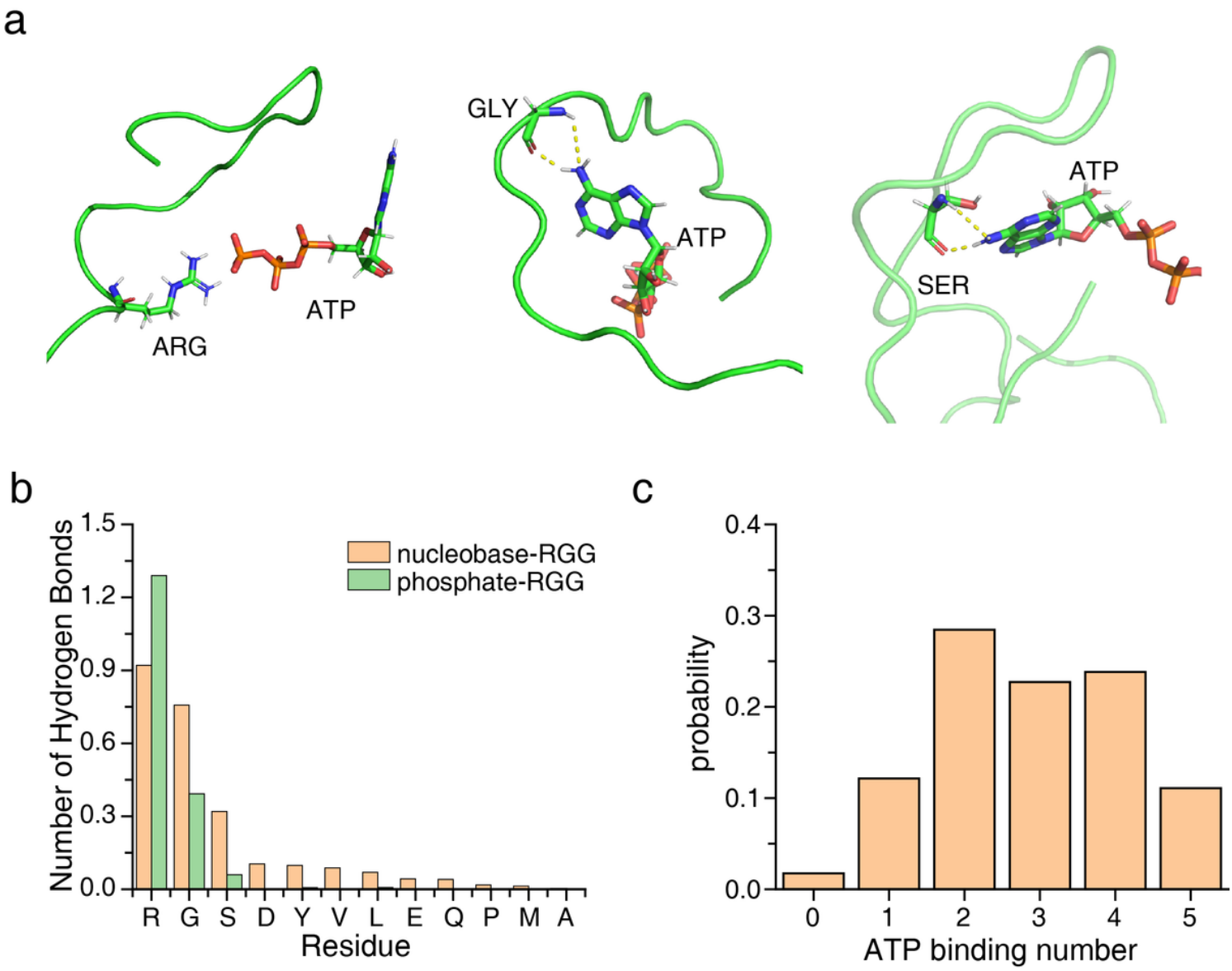

Figure 4

Structural models for the ATP-associated HNRNPG-RGG

a Representative snapshots from MD simulation indicate that ATP phosphate group interacts with Arg sidechain, while at the same time, ATP nucleobase forms hydrogen bonds with adjacent Gly and Ser residues. **b** Statistics of the number of the hydrogen bonds formed between ATP and different protein residues, with the phosphate and nucleobase groups tabulated separately. **c** The probability of the number of ATP molecules associated with the theoretical protein and ATP concentrations of 1.6 mM and 8 mM, respectively. The probability was calculated as the number of RGG domain associated with the number of ATP molecules divided by the total number of MD snapshots.

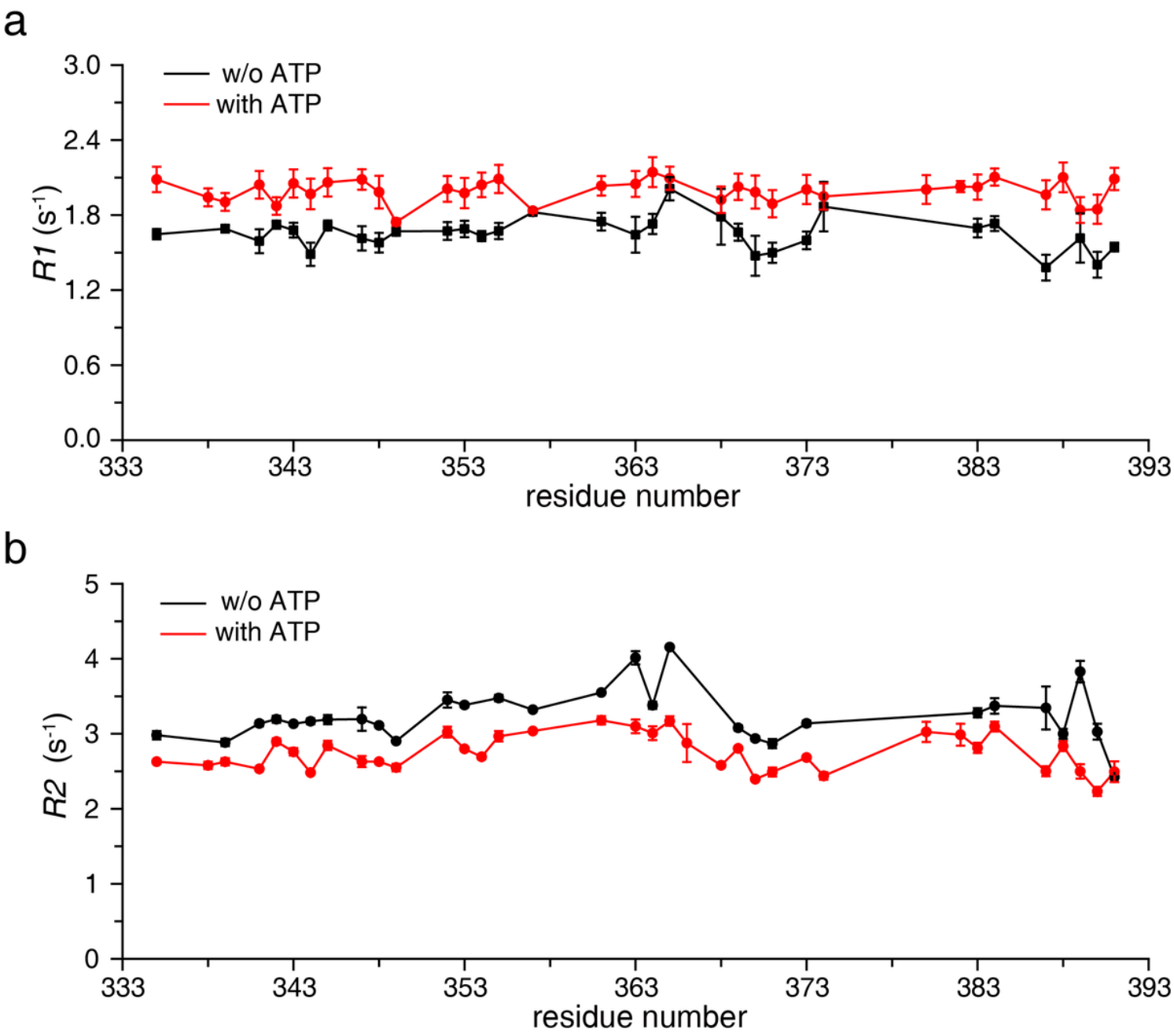

Figure 5

NMR relaxation analysis of the HNRNPG-RGG without or with 10 mM ATP.

The residue-specific R_1 (a) and R_2 (b) rates were collected at 25 °C on a 600 MHz NMR instrument.

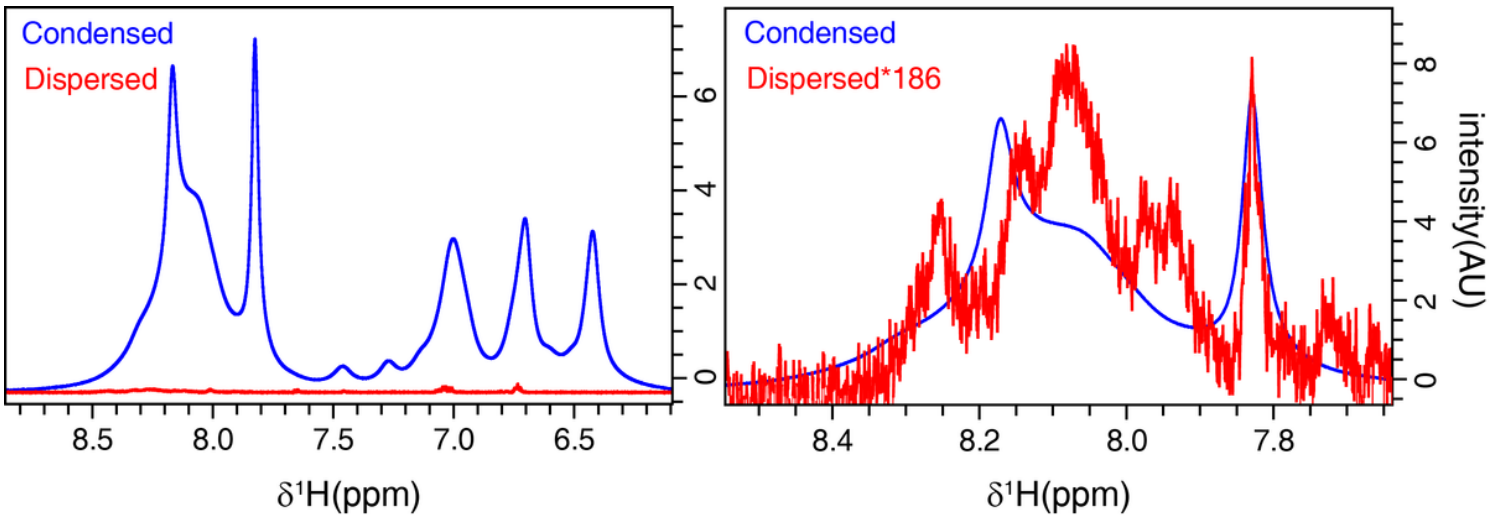

Figure 6

Overlay of the 1D NMR spectra of RGG domain in 100 μM dispersed and phase-separated states.

Based on the peak intensities, protein in the condensed phase is ~186 times more concentrated than in the diluted state. The original spectra without scaling were shown on the left.

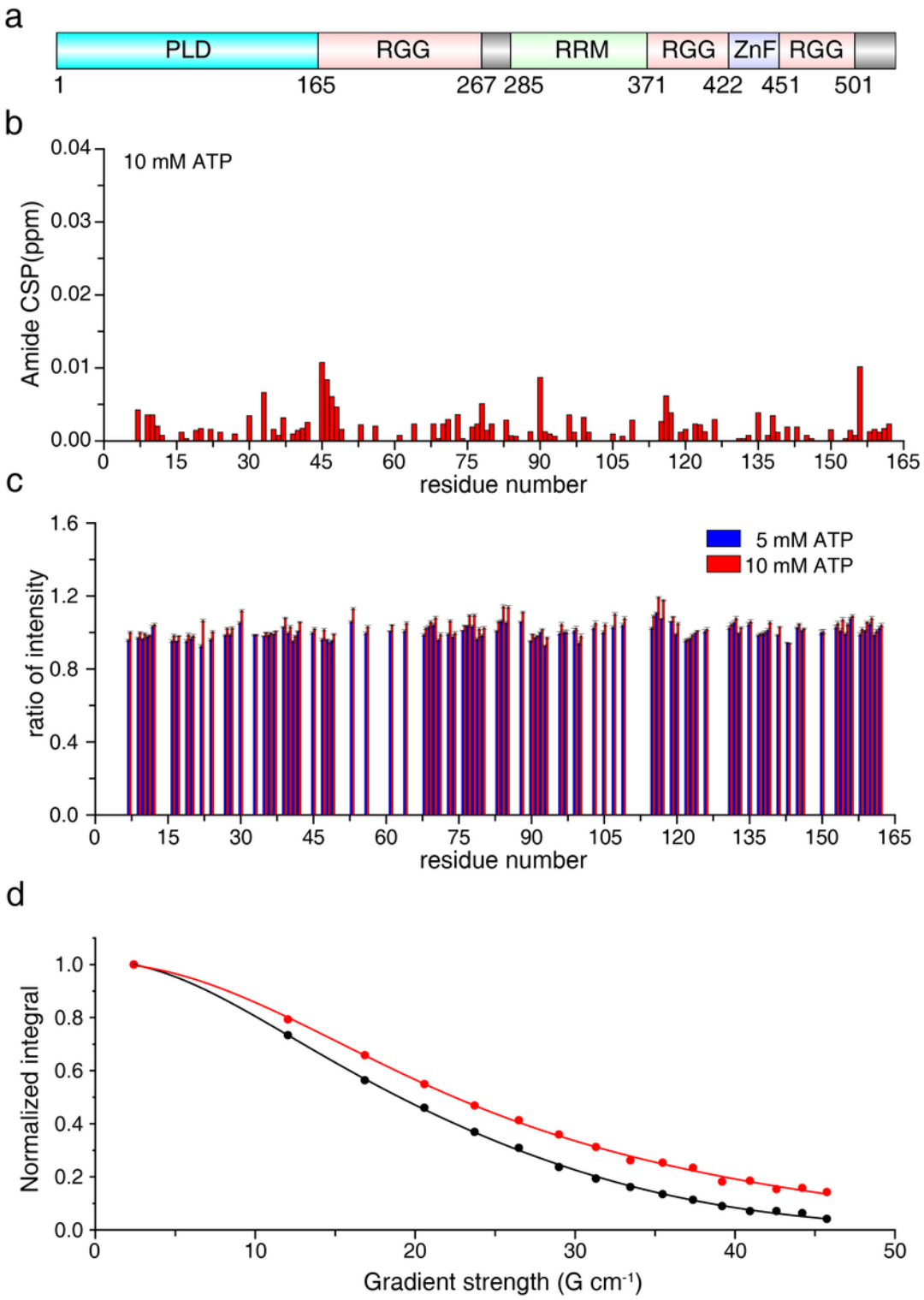

Figure 7

NMR titration of ATP to FUS PLD. a The construction of FUS, with the prion-like domain (PLD) at its N-terminus. **b** The chemical shift perturbations for the ^{1}H-^{15}N of FUS PLD in the presence of 10 mM ATP. **c** The relative peak intensities with the addition of 5 mM (red) and 10 mM ATP (blue), reference to those without ATP. **d** NMR DOSY measurement affords the translation diffusional coefficients. The NMR experiments were performed with the concentration of FUS PLD at 100 μM in the presence of 5 mM ATP.

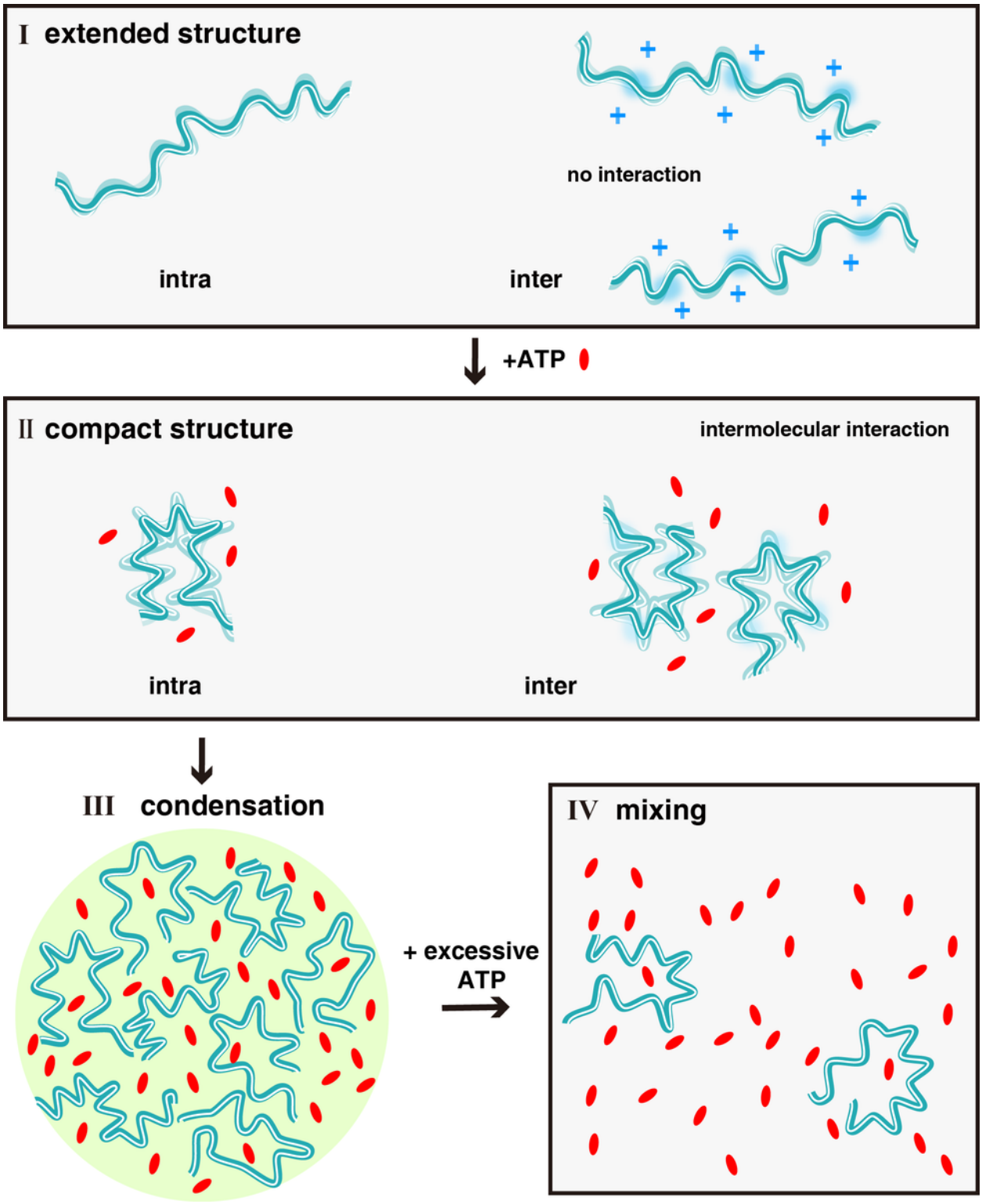

Figure 8

A scheme illustrating how ATP binding modulates the phase separation behavior of an RGG-rich IDP.

The RGG domain has a large density of positive charges that causes both intramolecular expansion and inter-molecular repulsion (I). The addition of ATP neutralizes the positive charges, allowing intermolecular interactions to take place, and also makes the protein more compact (II). Multiple copies of proteins come

together and de-mix when a critical concentration is reached (III). Nevertheless, excessive ATP further compacts the protein and dissipates the liquid droplets (IV).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementalmaterial.docx