BIOPHYSICS

Uncovering the molecular mechanism for dual effect of ATP on phase separation in FUS solution

Chun-Lai Ren[1]*†, Yue Shan[1]†, Pengfei Zhang[2], Hong-Ming Ding[3]*, Yu-Qiang Ma[1]*

Recent studies reported that adenosine triphosphate (ATP) could inhibit and enhance the phase separation in prion-like proteins. The molecular mechanism underlying such a puzzling phenomenon remains elusive. Here, taking the fused in sarcoma (FUS) solution as an example, we comprehensively reveal the underlying mechanism by which ATP regulates phase separation by combining the semiempirical quantum mechanical method, mean-field theory, and molecular simulation. At the microscopic level, ATP acts as a bivalent or trivalent binder; at the macroscopic level, the reentrant phase separation occurs in dilute FUS solutions, resulting from the ATP concentration–dependent binding ability under different conditions. The ATP concentration for dissolving the protein condensates is about 10 mM, agreeing with experimental results. Furthermore, from a dynamic point of view, the effect of ATP on phase separation is also nonmonotonic. This work provides a clear physical description of the microscopic interaction and macroscopic phase diagram of the ATP-modulated phase separation.

INTRODUCTION

Recent discoveries have shown that phase separation plays an important role in many intracellular biological processes (*1–6*), e.g., the formation of membrane-less compartments like stress granules in the cytoplasm (*7*), Cajal bodies (*8*), and nucleoli (*9*) in the nucleus. Phase separation results in the compartmentalization of biomolecules into concentrated liquid droplets (with a typical size in micrometers) via weak specific interactions at the molecular level (*10–12*). Although their organizations are typically highly dynamic and accompanied by the rapid exchange of components with their surroundings, their overall size and shape are usually stable for minutes or hours (*13, 14*).

The development of both experiment and theory (*15–22*) has revealed the possible mechanisms of phase separation. One central principle is multivalence, which can drive the formation of condensate via phase separation coupled to percolation (*23–27*). Specifically, multivalent interaction is facilitated by the intrinsically disordered regions (IDRs) in the proteins (*28–30*). As one specific class of IDRs, the prion-like domain (PLD) is frequently found in RNA binding proteins, e.g., the fused in sarcoma (FUS) protein. It was found that the molecular driving forces for phase separation in FUS were cooperatively weak interactions among tyrosine (Tyr) residues from the PLD and arginine (Arg) residues from the RNA binding domain (*31*); the valence of the multivalent interactions was a key factor determining the saturation concentration in FUS solutions.

Adenosine triphosphate (ATP) is known as an energy fuel for biological reactions typically at the concentration of micromolar (*32–34*). In cells, unexpectedly, the ATP concentration is in the range of 2 to 12 mM, which is much higher than the amount required for the energy fuel. Previous studies showed that the high concentration of ATP was related to the dissolution of protein condensates (*35*), in which ATP was explained as a biological hydrotrope (*36–38*). In chemistry, hydrotropes are small-molecule amphiphiles that contain distinct hydrophilic and hydrophobic parts (the aromatic ring/rings in most cases) and can help solubilize the hydrophobic substances in aqueous systems (*39*). Likewise, ATP consists of hydrophobic aromatic pyrimidine base ring and hydrophilic triphosphate moiety, fitting well into the concept of a hydrotrope.

However, later works found that like RNA (*40*), single-stranded DNA (*41*), and other molecules (*42*), ATP had two-stage effects on phase separation of prion-like proteins: an enhancement of phase separation at low concentrations and an inhibition of phase separation at high concentrations (*41, 43–45*). The mechanism of these dual effects of ATP is still largely unknown. It is worth noting that such reentrant phase separations are quite general in biomacromolecules, which were observed in many multicomponent biosystems (*46–50*). There is emerging evidence that abnormal protein aggregation is associated with many human diseases, including cancer, neurodegeneration, and infectious disease (*14, 51*). Therefore, it is of fundamental and physiological significance to uncover the influence of ATP on protein condensation or solubilization.

In this work, we combine the semiempirical quantum mechanical (SQM) method, mean-field theory, and coarse-grained molecular dynamics (CGMD) simulation to reveal the mechanism underlying the dual effect of ATP on phase separation. We first use the SQM method to uncover the Tyr-Arg and ATP-Arg bindings at the atomic level, which are taken as the input for the mean-field theory. We then develop a mean-field theory by explicitly taking molecular interactions into account based on a sticker-and-spacer model and provide phase diagrams with full binodals at the macroscopic level, which can be directly compared with the experimental phenomena. Furthermore, we use the CGMD simulation to investigate the phase separation from the dynamic perspective. These three methods complement each other and can build a direct relationship between the microscopic driving force and macroscopic phase separation, offering a better understanding of ATP-modulated phase separation.

[1]National Laboratory of Solid State Microstructures and Department of Physics, Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China. [2]State Key Laboratory for Modification of Chemical Fibers and Polymer Materials, Center for Advanced Low-Dimension Materials, College of Material Science and Engineering, Donghua University, Shanghai 201620, China. [3]Center for Soft Condensed Matter Physics and Interdisciplinary Research, School of Physical Science and Technology, Soochow University, Suzhou 215006, China.
*Corresponding author. Email: chunlair@nju.edu.cn (C.-L.R.); dinghm@suda.edu.cn (H.-M.D.); myqiang@nju.edu.cn (Y.-Q.M.)
†These authors contributed equally to this work.

RESULTS

The relationship between Tyr-Arg binding and phase separation in FUS solution

Multivalent interactions among Tyr and Arg residues are considered as driving forces for the phase separation in FUS solution (*31*). However, the direct relationship between the multivalent interactions and phase diagrams is still lacking. To achieve a quantitative description, we first used the SQM method to calculate the binding free energy of the Tyr-Arg interaction (more details on the SQM method are given in Materials and Methods). Figure 1 (A and B) shows two typical binding modes of the Tyr-Arg complex, where the ending charged group in Arg may be either perpendicular or parallel to the benzene plane in Tyr. Notably, the binding free energy is about $-2.0\ k_BT$ in both cases, indicating that the Tyr-Arg binding is a weak interaction. This raises the question of whether Tyr-Arg binding could be sufficient to cause the phase separation in FUS solutions.

To answer this question, we developed a mean-field theory for FUS solutions based on a sticker-and-spacer model. This model was first proposed by Semenov and Rubinstein (*52*) to treat the gelation and phase separation of thermo-responsive associative polymer solutions, following the pioneering works by Tanaka (*53*, *54*) for similar systems. Recently, this model has been successfully used to reflect the phase separation in multivalent proteins (*20*, *21*, *55*). It should be mentioned that two order parameters, i.e., the volume fraction of protein and the fraction of Tyr-Arg bonds, self-consistently appear in the free energy, which is widely used to reflect the concentration and structure in the molecular liquid (*52*, *56*). To better compare the theoretical and experimental results, we convert the calculated volume fraction into the concentration (see details of the theoretical modeling in Materials and Methods) in the following discussion. Figure 1C shows the phase diagram (i.e., the binodal curve) of phase separation in the FUS solution, depending on the FUS concentration (ρ) and the Tyr-Arg binding free energy (ϵ_1). The details of the

calculation for the phase diagram are given in the Supplementary Materials. When ϵ_1 is lower than $-1.4\ k_BT$, the phase separation could be observed in the FUS solution, meaning that the Tyr-Arg binding is strong enough to drive phase separation due to the cooperation of multiple Tyr-Arg bindings. At $\epsilon_1 = -2.0\ k_BT$, the concentration of the condensed phase is two orders of magnitude higher than that of the dilute phase, which is consistent with the characteristics of protein condensates (*12*). Because the binding free energy of Tyr-Arg is weak, there may exist a reversible breaking and formation of the Tyr-Arg binding under the thermal fluctuation. The extent of Tyr-Arg binding can be reflected by the fraction of Tyr-Arg bonds p, namely, the ratio of the number of Tyr-Arg bonds to the number of Tyr residues in the solution. As shown in Fig. 1D, the value of p in the dense phase is larger than that in the diluted phase, indicating that the protein condensates are denser liquids with network structures in which Tyr-Arg bonds act as cross-linkers (Fig. 1E). Collectively, the above results demonstrate that different FUS concentrations, as well as distinct microscopic structures, occur in the two coexisting phases.

SQM methods on the microscopic interactions between ATP and Arg

We then investigated the effect of ATP on the phase separation in FUS solutions. With the addition of ATP, there may exist an interaction between ATP and FUS, in which the ATP-Arg binding is believed to play an important role (*41*, *43*–*45*). However, the binding modes, possible binding numbers of Arg residues (to ATP), and binding strength are still lacking. To reveal the detailed microscopic mechanism of their interactions, the SQM method was again used. One study (*41*) assumed that ATP may act as a bivalent binder due to its amphiphilic nature, i.e., two Arg residues could bind to the triphosphate part and the adenine part, respectively. As shown in Fig. 2 (A and B), the Arg can bind to either the triphosphate part or
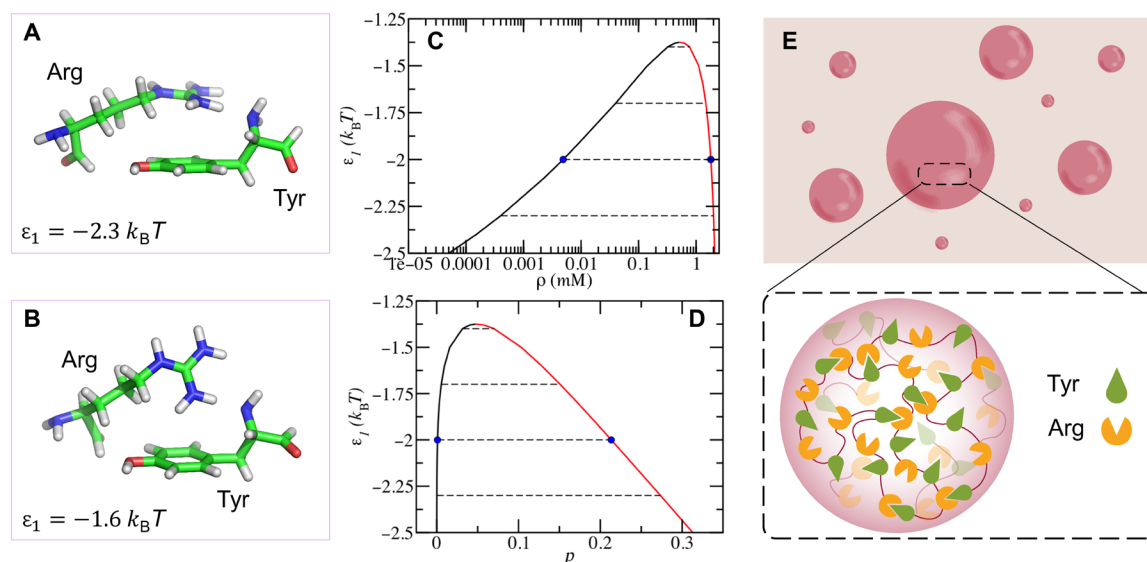

Fig. 1. Tyr-Arg binding resulting in the phase separation in FUS solution. Snapshots for the two typical binding modes between one Tyr and one Arg obtained by the SQM method (**A** and **B**), where ϵ_1 denotes the binding free energy in each case. Phase diagram and the fraction of Tyr-Arg binding calculated from the mean-field theory (**C** and **D**). The black curve represents the diluted phase, and the red one represents the dense phase. Dash lines are tie-lines relating two coexisting phases. The blue dots at $\epsilon_1 = -2\ k_BT$ are taken as the average binding free energy between Tyr and Arg for the following theoretical calculations. (**E**) Schematic representation of the formation of FUS protein droplets caused by Tyr-Arg binding.
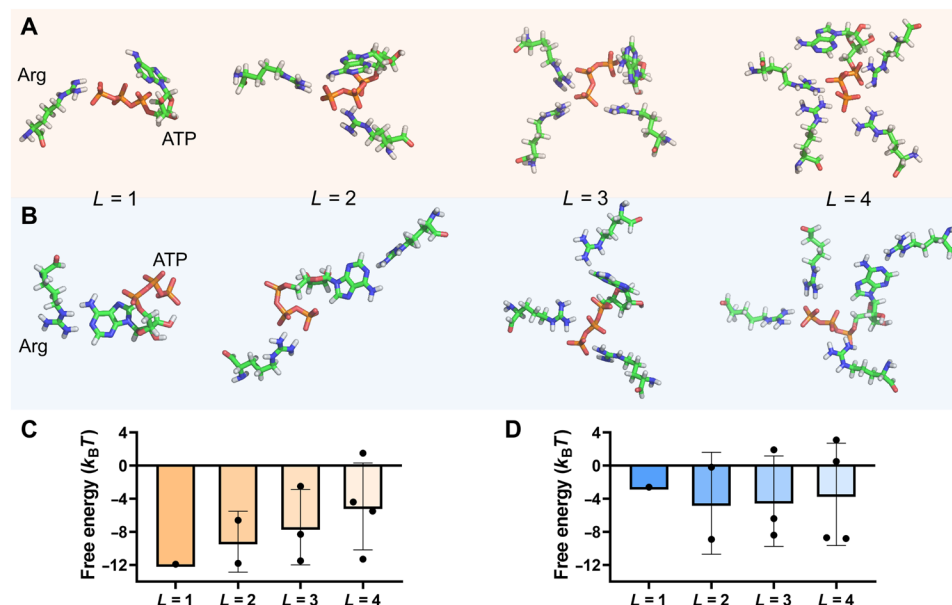
Fig. 2. The SQM result for the ATP-Arg binding. Snapshots for the typical binding modes between one ATP and one/several Arg residues: (**A**) The Arg residue(s) all bind to the triphosphate part of the ATP; (**B**) one of the Arg residue(s) binds to the adenine part of the ATP. The number of Arg residues is denoted as the valence of ATP, represented by L. (**C** and **D**) shows the corresponding binding free energy of the ATP-Arg interaction as a function of L in the case of (A) and (B), respectively.

the adenine part of ATP. Nevertheless, the binding strength is quite different in the two cases. Because the Arg carries one net positive charge and the triphosphate part of ATP is highly negatively charged, the binding free energy in the former case is much lower than that in the latter case (Fig. 2, C and D). Thus, the Arg prefers binding to the triphosphate part of the ATP. We also compared the binding preference in the case of multiple Arg residues. Similarly, these Arg residues are more likely to accumulate around the triphosphate part of the ATP because the average binding free energy is much lower. Notably, because of the steric effect, the average binding free energy increases with the increase of the Arg number. The average binding energy increases significantly, and one of the binding free energies becomes positive in the case of four Arg residues (Fig. 2C), indicating that the binding of Arg to ATP is no longer energy-favorable. Thus, the maximum binding number of Arg residues (to ATP) should be three; namely, ATP may act as a trivalent binder.

We only considered a single amino acid (i.e., Arg) in the above calculation, which may be different from the real case when the Arg is in the protein chain. To this end, we further considered a tripeptide model [i.e., glycine (Gly)–Arg-Gly] and calculated the binding free energy between the tripeptide and the ATP. The use of Gly as the terminal of the tripeptide is due to the fact that Gly is the most frequently neighboring amino acid of the Arg in FUS (there are two Arg-Gly rich regions in FUS). Gly is a very small amino acid; thus, it could have little impact on the interaction between the Arg and the ATP. As shown in fig. S1, the binding free energy of the ATP-tripeptide interaction is similar to that of ATP-Arg interaction at $L = 1$, 2, and 3 (Fig. 2C), while some increase in the binding free energy is observed at $L = 4$, probably due to the steric effect. Generally, the neighboring amino acids may not obviously affect the binding strength and the binding valance of the Arg-ATP interaction in FUS.

Mean-field theory on the relationship between microscopic interactions and macroscopic phase behaviors

According to the above SQM results, the strength of ATP-Arg binding is stronger than that of Tyr-Arg binding, implying that the phase separation of FUS driven by Tyr-Arg binding could be affected by the addition of ATP. As such, to investigate the effect of ATP on the phase separation in FUS solutions, we further developed the mean-field theory by taking one-to-multiple binding between the ATP and Arg residues into account; this was inspired by the idea of dealing with specific interactions in polymer systems (see details of the theoretical modeling in Materials and Methods) (52, 57, 58).

Figure 3A shows the phase diagram with critical points in the presence of ATP. The saturation concentration (i.e., the left part of the binodal curve separated by the critical point) of FUS shows a nonmonotonic manner with the addition of ATP when ATP serves as a multivalent binder (i.e., $L = 2$ or 3). When the amount of ATP is small, the saturation concentration decreases with the increase of ATP. However, the saturation concentration increases when the ATP concentration reaches the order of millimolar. As illustrated in Fig. 3B, ATP can bind to more than one Arg residues when the ATP concentration is low, which could be taken as an additional driving force for promoting the aggregation of proteins and causing phase separation in very dilute solutions. However, at a high ATP concentration, each ATP can only bind to one Arg because of the limited number of Arg residues, which leads to the breakage of Tyr-Arg binding and the inhibition of phase separation. This two-stage phenomenon is in good agreement with previous experiments on the effect of ATP on phase separation (43, 44): first, the enhancement and the subsequent inhibition with the increase of additives. On the contrary, if ATP is just a monovalent binder (i.e., $L = 1$), then the saturation concentration of FUS monotonically increases with the increase of ATP concentration. In other words, a higher
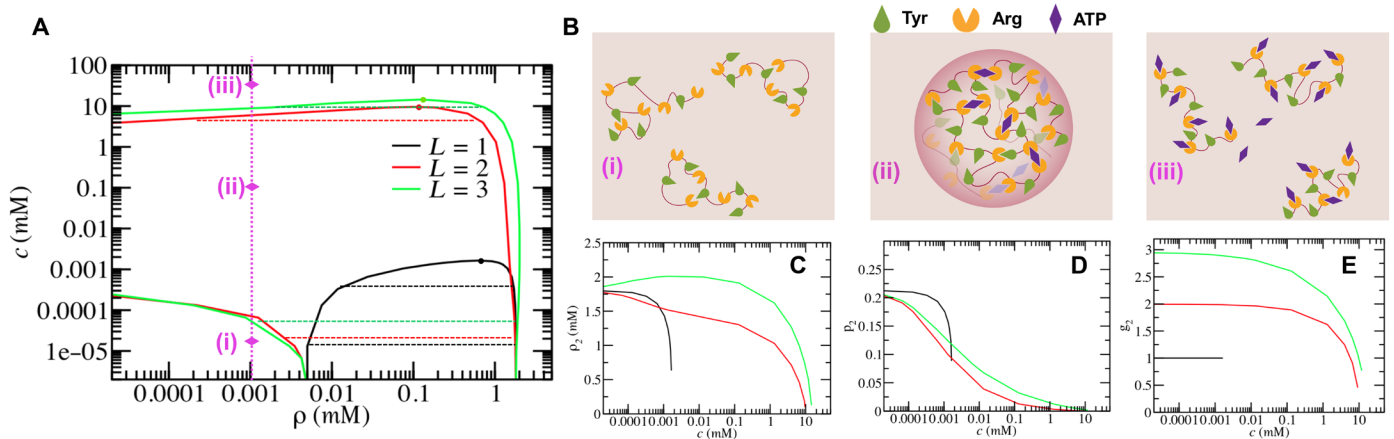
Fig. 3. The reentrant phase separation and protein dissolution in mean-field theory. (**A**) Phase diagram of FUS solutions as a function of ATP concentration (c) for three different ATP valences L. The parameter $\beta\epsilon_2$ at $L = 1$, 2, and 3 is set based on the average binding free energy in Fig. 2C. Dashed lines are tie-lines relating two coexisting phases. Dots are critical points of each cases. The dotted line with diamond symbols (i) to (iii) indicates the reentrant phase transition with the increase of ATP concentration. (**B**) Schematic representations for (i) to (iii) indicate the reentrant phase transition occurring in dilute solutions for $L = 2$. (**C**) The concentration of protein condensates (ρ_2), (**D**) the extent of Tyr-Arg binding (p_2), and (**E**) the average number of bound Arg residues per ATP (g_2) in protein condensates as functions of the ATP concentration.

concentration of protein solution is required for the phase separation in the presence of ATP, indicating that the ATP only inhibited phase separation. Therefore, a multivalent binder acted on by ATP is a prerequisite of the two-stage phenomenon.

It was reported that the dissolution of protein droplets can be observed when the ATP concentration reached its physiological concentration (*35*). As shown in Fig. 3C, the protein condensates can completely disappear when the ATP concentration is about 10 mM for $L = 2$ and 15 mM for $L = 3$, which agrees with experimental observations (∼8 to 12 mM) by Patel *et al.* (*35*). In comparison, the protein condensates dissolve when ATP concentration reached 0.001 mM under the condition of $L = 1$. These results indicate that the multivalence of ATP may be an important reason why the physiological concentration of ATP is in the order of millimolar. With the increase in ATP, the concentration of the protein condensate ρ_2 first increases and then decreases in the case of $L = 3$. This means that the addition of a small amount of ATP not only can stabilize the protein droplet but also can make the droplet denser when the ATP has a high valence ($L = 3$). Moreover, as shown in Fig. 3 (D and E), both the extent of Tyr-Arg binding and the average number of bound Arg residues per ATP show a similar way of changing with the addition of ATP in the case of $L = 2$ or 3, which is quite different from that in the case of $L = 1$. In the latter case, one molecule of ATP always links to one Arg, which strongly decreases the Tyr-Arg binding and shows the inhibition of phase separation with ATP concentration much lower than the physiological concentration. This indicates that the changeable binding ability of ATP (to Arg) is crucial for dissolving the protein droplets at a millimolar ATP concentration.

CGMD simulation on the effect of ATP on the phase separation from the dynamic view
We also used the CGMD simulation (see details of the CG modeling in Materials and Methods) to investigate the effect of ATP on the phase separation. As shown in fig. S2 (A and B), the reentrant phase separation in the diluted solution is also observed once the ATP can

bind to more than one Arg residues. Moreover, a large number of ATP molecules can dissolve the condensates at a high concentration of protein solution (fig. S2, C and D). Protein condensates are dynamic, having the characteristic of the exchange of protein molecules inside and outside the condensate. As shown in movies S1 to S3, the phase separation is in a dynamic equilibrium and the protein chains in the surrounding continuously exchange with protein chains in the droplet. In more detail, there are many exchanged chains in the absence of ATP (Fig. 4A). With the addition of a small amount of ATP, because the protein condensate becomes denser, the number of chains participating in the exchange decreases (Fig. 4B). With the further addition of ATP, the protein condensate becomes looser and the molecular exchange is promoted again (Fig. 4C). The quantitative data of the total number of exchanged chains in a given time can be found in fig. S3. Generally, from the dynamic point of view, the influence of ATP on phase behavior is also nonmonotonic.

Predictions on phase separation induced by ATP for prion-like proteins with asymmetric numbers of Tyr and Arg residues
After revealing the nonmonotonic effect of the ATP on the phase separation in the FUS solution, we further explored a more general case in prion-like proteins, in which the number of Tyr and Arg residues is not necessarily equal (*31*). To denote such a difference, here, for the sake of simplicity, we only considered two cases: the number of Tyr residues being much larger than ($m_1 = 34$, $m_2 = 11$) or smaller than ($m_1 = 11$, $m_2 = 34$) that of Arg residues. Because of the asymmetric number of Tyr and Arg residues, the phase separation is hardly observed in the pure protein solutions. With the addition of ATP, the phase separation could be observed as two looped phase diagrams (see Fig. 5A). However, the area of the loop is quite different; namely, the area in the case of $m_1 < m_2$ is much larger than that in the case of $m_1 > m_2$. This difference mainly originated from the ability of ATP binding to Arg; in other words, the average number of bound Arg residues per ATP in the case of $m_1 < m_2$ is much larger than that in the case of $m_1 > m_2$ (Fig. 5B). Furthermore, with a small amount of ATP, the Tyr-Arg binding is enhanced in the case
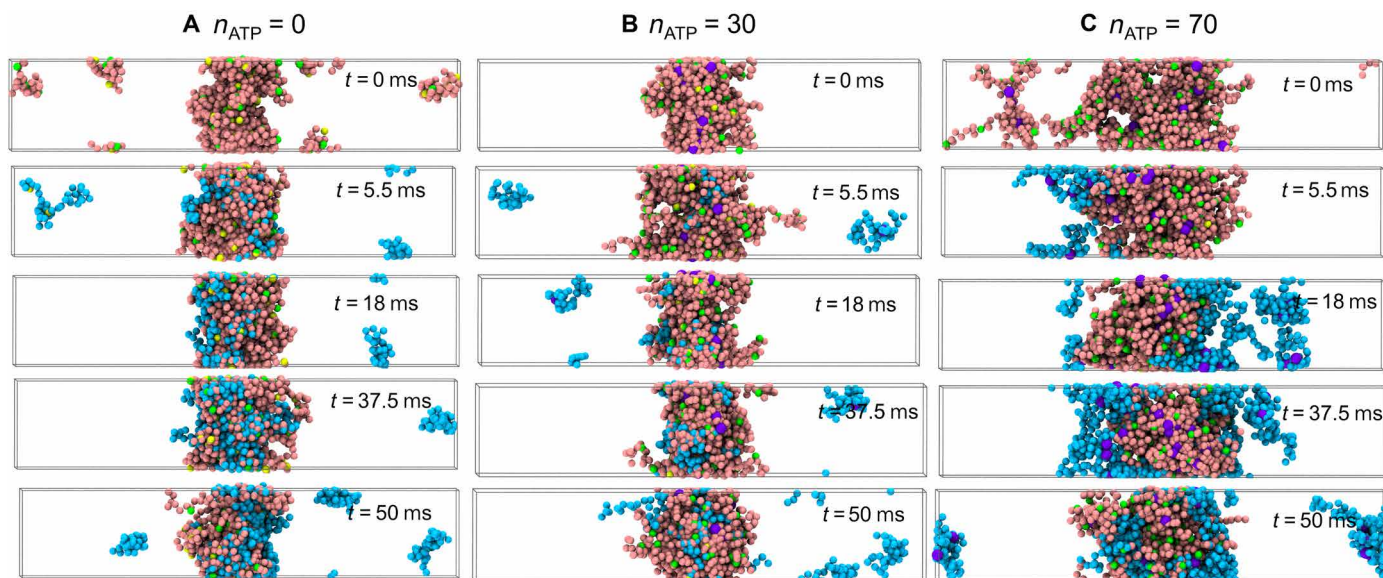
Fig. 4. Molecular exchange of protein condensates. Time sequence of snapshots illustrating the exchange of protein chains inside and outside the protein condensates in CGMD simulations: (**A**) $n_{ATP} = 0$, (**B**) $n_{ATP} = 30$, and (**C**) $n_{ATP} = 70$. The protein chains participating in the molecular exchange are shown in cyan, and others are described as pink chains. The yellow and green beads represent Arg and Tyr residues, respectively, and the purple beads stands for the ATP molecules. The total number of protein chains is $n_{pro} = 50$ in the simulations.

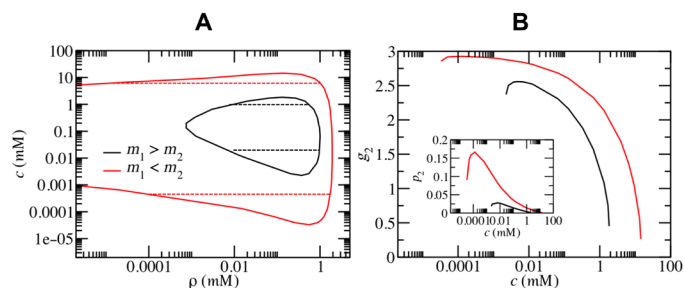**Fig. 5. The effect of ATP on other prion-like proteins.** (**A**) Phase diagram of prion-like proteins with asymmetric numbers of Tyr and Arg residues ($m_1 > m_2$ and $m_1 < m_2$) with the addition of ATP for $L = 3$. (**B**) The average number of bound Arg residues per ATP (g_2) and the extent of Tyr-Arg binding (p_2) (inset) in protein condensates as functions of ATP concentration.

of $m_1 < m_2$, which promotes phase separation (see the inset of Fig. 5B). The above theoretical results are also verified by the CGMD simulations (fig. S4). Generally, the phase separation could also be modulated by the ATP in the protein systems with asymmetrical numbers of Tyr and Arg residues, but the sensitivity of the phase separation (to the ATP concentration) is highly related to the number of Arg residues.

DISCUSSION
Why the ATP concentration in cells is so high has long been a puzzling problem. When Patel *et al.* (*35*) found that ATP can work as a biological hydrotrope in 2017, researchers began to pay special attention to the effect of ATP on the dissolution of protein condensates (*36*). Later, the dual roles of ATP on phase separation were reported (*41*, *43–45*), which again put forward the question of how ATP works.

In this work, combining the effort of the SQM method, mean-field theory, and molecular simulation, we perform a systematic study on the modulation of FUS condensates aroused by ATP. The SQM method shows that Arg prefers binding to the triphosphate part of the ATP, with the binding free energy in the order of k_BT and a maximum binding number of three. Taking these microscopic features as the input, we develop a mean-field theory to provide macroscopic phase diagrams based on the thermodynamic criterion; this demonstrates that the nonmonotonic modulation largely originates from the distinct number of binding Arg residues (to ATP) at different ATP concentrations. Furthermore, we use the CGMD simulation to demonstrate that the dynamic molecular exchange between the condensate and the surrounding also shows a nonmonotonic behavior as the ATP concentration increases.

It should be mentioned that the effects of ligands on the phase separation of proteins may be also the opposite (i.e., enhancement or inhibition), which can be well described within the framework of polyphasic linkage first developed by Wyman and Gill (*59*). Recently, Ruff *et al.* (*60*, *61*) reported that the strength, valence, structure, and concentration of the ligand can generate important and subtle effects on the phase behaviors of multivalent macromolecules with sticker-and-spacer architectures. Notably, they were the first to show that monovalent ligands destabilize phase separation driven by homotypic interactions, and the nonmonotonic behavior of ligands as a function of the ligand concentration was also anticipated and demonstrated (*60*, *61*). Very recently, Dao *et al.* (*62*) found that the type of the polyubiquitin chain could also play distinct roles in the phase separation of UBQLN2, which further bolsters the general conceptual foundations of the findings from Ruff *et al.* These results share some similarity with the dual effect of ATP (on the phase separation) reported here, which provides a key conceptual lynchpin for our current work.

In general, this work well explains the nonmonotonic effect of ATP on the phase separation. This sheds some lights on fine-tuning the phase behaviors of prion-like proteins with additional molecules and may have implications for devastating neurodegenerative disorders.

MATERIALS AND METHODS

The SQM method
Here, the latest SQM method proposed by Bannwarth *et al.* (*63*), the GFN2-xTB method, is used to calculate the free energy of the Arg binding to the Tyr and/or the ATP. The accuracy of this SQM method in binding energy calculation is believed to be close to that of the density functional theory, but it has a much lower computing cost (*64*). The binding free energy is calculated as $\Delta G = G_{com} - G_A - G_B$, where G_{com} is the free energy of the complex, G_A is the free energy of the Arg, and G_B is the free energy of the Tyr or the ATP. The free energy G usually comes from three sources: the total molecular gas-phase energy, the corresponding free solvation energy, and the thermostatistical contribution to the free energy including the translation, rotation, vibration, and conformation degrees of freedom at a given temperature (*63, 64*). To obtain the representative configuration, we first used the Molclus program (*65*) to search the possible binding structures in each case. Then, we used the GFN2-xTB method to make a preliminary optimization of the searched structures. Next, we divided the optimized structure into a complex and two monomers and performed the Hessian calculation. All the calculations were performed under the implicit solvent model, namely, using the generalized Born (GB) model with surface area (SA) contributions [i.e., GBSA(H$_2$O)] (*66*). The SQM calculations were performed using the xtb program (*67*).

Theoretical modeling of FUS solutions in the absence of ATP
In the theory, FUS proteins were modeled as linear polymers with a length of $N_1 = 526$. Because FUS belongs to a class of biomolecules with limited aqueous solubility (*20*), solvents were modeled as chains with lengths of N_2, which is a simple way to reflect the limited solubility of biomolecules (*15, 16*). Each FUS protein contains some Tyr and Arg residues (*31*). The number of Tyr or Arg residues is named as the valence of the multivalent interaction. According to experimental data from Wang *et al.* (*31*), FUS has m_1 Tyr residues and m_2 Arg residues, with $m_1 = m_2 = 34$. The incompressible solution consists of n_p FUS proteins and n_s solvent molecules. The dimensionless free energy of the reference state βF_{ref} was written as

$$\beta F_{ref} = \frac{\phi}{N_1}\ln\phi + \frac{(1-\phi)}{N_2}\ln(1-\phi) \quad (1)$$

where $\phi = \frac{n_p N_1 v}{V}$ is the volume fraction of FUS. v is the volume of one residue of the protein, which is taken as the unit volume. $V = n_p N_1 v + n_s N_2 v$ is the volume of system. Thus, the reference system is always homogeneous, which is also the system without including Tyr-Arg binding. Next, we took the specific binding between Tyr and Arg residues into account, assuming that each Tyr/Arg has only one binding site. According to previous works on associative polymers (*52, 55*), the free energy of Tyr-Arg binding can be written as

$$\beta F_{binding} = -\frac{v}{V}\ln Z_{binding} \quad (2)$$

where $Z_{binding}$ is a partition function and is equal to

$$Z_{binding} = P_{binding}W\exp(-k\beta\epsilon_1) \quad (3)$$

Here, $P_{binding}$ is a combinatorial factor that describes the number of ways to form k Tyr-Arg bonds, which can be written as

$$P_{binding} = C_{n_p m_1}^k C_{n_p m_2}^k k! \quad (4)$$

The formation of each bond requires the Tyr and the Arg to be in close proximity within some volume scale v_b. For simplicity, we take v_b to simply be the residue volume v. The overall probability of forming k bonds is then

$$W = \left(\frac{v}{V}\right)^k \quad (5)$$

In Eq. 3, $\beta\epsilon_1$ is the free energy associated with the binding between Tyr and Arg, which includes the affinity between Tyr and Arg residues, as well as the entropic loss during the formation of the Tyr-Arg bond (*57*). More specifically, the specific chemical structures of Tyr and Arg side chains appear to be important determinants of the formation of the Tyr-Arg bond (*31*). Such a reduction of conformational entropy is a common character for specific interactions, such as hydrogen bonds and cation-π interactions. Substituting Eqs. 3 to 5 into Eq. 2, the dimensionless binding free energy can be written as

$$\beta F_{binding} = \frac{m_1\phi}{N_1}[p\ln p + (1-p)\ln(1-p)] +$$
$$\frac{m_2\phi}{N_1}\left(1-p\frac{m_1}{m_2}\right)\ln\left(1-p\frac{m_1}{m_2}\right) - p\frac{m_1\phi}{N_1}\left(\ln\frac{m_2\phi}{N_1 e} - \beta\epsilon_1\right) \quad (6)$$

where $p = \frac{k}{n_p m_1}$ is the fraction of Tyr in the formation of Tyr-Arg bonds and reflects the extent of protein network cross-linked by the Tyr-Arg bonds. Thus, the total free energy of the FUS solution is $F_{tot} = F_{ref} + F_{binding}$, which includes the two order parameters ϕ and p. According to the thermodynamic criterion, phase diagrams can be obtained from the equality of chemical potential and the osmotic pressure of two coexisting phases (see the details in the Supplementary Materials). The two coexisting phases consist of a diluted phase and a condensed phase, with distinct volume fractions of ϕ_1 and ϕ_2, respectively. It should be mentioned that the volume fraction of the diluted phase (ϕ_1) can be converted to the saturation concentration (ρ_1) by considering that the hydration size of the amino acid is about 0.65 nm (*68*). This can be then mapped with the experimental saturation concentration (i.e., 5 μM) (*31*) to determine the fitting parameter $N_2 = 65$ in the theory.

Theoretical modeling of FUS solutions in the presence of ATP
The ATP is modeled as a small molecule with a volume of v (the unit volume), which is the same as one residue of FUS. The reference system is chosen as a FUS solution with freely moving ATP molecules. The dimensionless free energy of the reference state can be written as

$$\beta\mathcal{F}_{ref} = \frac{\phi}{N_1}\ln\phi + \frac{(1-\lambda\phi-\phi)}{N_2}\ln(1-\lambda\phi-\phi) + \lambda\phi\ln\lambda\phi +$$
$$\chi\lambda\phi(1-\phi-\lambda\phi) - \lambda\phi\mu_{ATP} \quad (7)$$

where $\lambda = \frac{n_{\text{ATP}}}{n_p N_1}$ is the ratio of the number of ATP molecules and the number of FUS residues. n_{ATP} is the number of ATP molecules in the solution. The first three terms are mixing the entropy for the three components of FUS, solvent, and ATP. The fourth term reflects the effective interaction between ATP and the solvent with $\chi = -1.2$, because ATP can act as a hydration mediator due to its triphosphate with a unique hydration property (44). The last term is introduced to change the system into a semiclosed system (58), which is more convenient for calculating the phase diagram. This means that ATP molecules in protein solutions are free to exchange with a reservoir at a fixed chemical potential (μ_{ATP}) determined by ATP volume fraction (φ), i.e., $\mu_{\text{ATP}} = \ln\varphi - \frac{1}{N_2}\ln(1-\varphi) - \frac{1}{N_2} + 1 + \chi(1-2\varphi)$. The ATP volume fraction (φ) can be converted into the ATP concentration (c) by $c = \frac{\varphi}{v}$. Next, we take both the Tyr-Arg binding and ATP-Arg binding into account. The former is a one-to-one binding, and the latter may be one-to-multiple binding. L is used to represent the valence of ATP. According to previous works on one-to-multiple binding (57, 58), the free energy of the formation of different kinds of bonds is

$$\beta\mathcal{F}_{\text{binding}} = -\frac{v}{V}\ln\mathcal{Z}_{\text{binding}} \tag{8}$$

$\mathcal{Z}_{\text{binding}}$ is the partition function and is given by

$$\mathcal{Z}_{\text{binding}} = \mathcal{P}_{\text{binding}}\mathcal{W}\exp(-k\beta\epsilon_1 - b\beta\epsilon_2) \tag{9}$$

Here, $\mathcal{P}_{\text{binding}}$ is a combinatorial factor that describes the number of ways to form k Tyr-Arg bonds and b ATP-Arg bonds in the system with n_p FUS proteins and n_{ATP} ATP molecules. It can be written as

$$\mathcal{P}_{\text{binding}} = C_{n_p m_1}^k C_{n_{\text{ATP}} L}^b C_{n_p m_2}^{k+b}(k+b)! \tag{10}$$

$\mathcal{W}$ in Eq. 9 is a probability of finding $(k+b)$ bonds in the vicinity of each other, which is

$$\mathcal{W} = \left(\frac{v}{V}\right)^{(k+b)} \tag{11}$$

In Eq. 9, $\beta\epsilon_2$ is the binding free energy aroused by the formation of one ATP-Arg bond. Substituting Eqs. 9 to 11 into Eq. 8, the associated dimensionless free energy is given by

$$\begin{aligned}\beta\mathcal{F}_{\text{binding}} = {} & \frac{m_1\phi}{N_1}[p\ln p + (1-p)\ln(1-p)] + L\lambda\phi[q\ln q + (1-q)\ln(1-q)] + \\ & \frac{m_2\phi}{N_1}\left(1 - \frac{pm_1}{m_2} - \frac{qL\lambda N_1}{m_2}\right)\ln\left(1 - \frac{pm_1}{m_2} - \frac{qL\lambda N_1}{m_2}\right) - \\ & p\phi\frac{m_1}{N_1}\left(\ln\frac{m_2\phi}{N_1 e} - \beta\epsilon_1\right) - qL\lambda\phi\left(\ln\frac{m_2\phi}{N_1 e} - \beta\epsilon_2\right)\end{aligned} \tag{12}$$

where $q = \frac{b}{L\lambda n_p N_1}$ is the fraction of ATP participating in ATP-Arg binding. Therefore, the total free energy becomes $\mathcal{F}_{\text{tot}} = \mathcal{F}_{\text{ref}} + \mathcal{F}_{\text{binding}}$. It can be seen that p and q are coupled in Eq. 12, indicating the competition between ATP-FUS and FUS-FUS interactions. After minimizing $\mathcal{F}_{\text{tot}}$ with respect to p, q, and μ, the phase diagram can be obtained according to the thermodynamic criteria of the same chemical potential and osmotic pressure of two coexisting phases (see the details in the Supplementary Materials).

The CGMD simulation

A CGMD simulation was used to investigate the effect of ATP on the phase separation of the protein solution. The protein was modeled as a chain with connective beads in the implicit solvent (22).

These beads were connected by the harmonic spring with the potential $U_b(r) = k(r - r_b)^2$, where $r_b = 4.5$ nm is the mean bond length and $k = 20\epsilon_0/r_b^2$ is the bond stiffness, where ϵ_0 is the energy unit. There were three types of beads in the protein chain, namely, the Tyr beads, the Arg beads, and the noninteracting beads. Here, the percentage of the Tyr or Arg beads was 0.06, similar to that in the mean-field theory. To reflect the effective binding between the Tyr bead and the Arg bead, the attractive potential $U_a(r) = U_0\left(1 + \cos\frac{\pi r}{r_0}\right)$ for $r < r_0$ was used. We chose $r_0 = 2.0$ nm and $U_0 = -20\epsilon_0$ (22). Moreover, a softened and truncated Lennard-Jones potential was adopted to avoid the aggregation of the same type of beads, which was given by $U_r(r) = 4\epsilon\lambda\left\{\left[(1-\lambda)^2 + \left(\frac{r}{\sigma}\right)^6\right]^{-2} - \left[(1-\lambda)^2 + \left(\frac{r}{\sigma}\right)^6\right]^{-1}\right\}$ when $r < r_c$. $\epsilon = 0.621\epsilon_0$ and $\epsilon = 0.15\epsilon_0$ were chosen for the Tyr bead and the Arg bead, respectively. Other parameters are fixed as $\lambda = 0.68$, $\sigma = 3.5$ nm, and $r_c = 5$ nm, which are similar to those used in the previous article (22). The ATP was modeled as one bead in the simulations. The effective binding between the ATP bead and Arg bead was also modeled as the attraction potential $U_a(r) = U_0\left(1 + \cos\frac{\pi r}{r_0}\right)$. Here, $r_0 = 2.0$ nm and $U_0 = -75\epsilon_0$ were chosen because the binding free energy of ATP-Arg was nearly four times as strong as that of Tyr-Arg according to the SQM result.

The coarse-grained simulations were performed at the NVT ensemble with the Langevin thermostat at room temperature. The size of the simulation box was 50 nm by 50 nm by 250 nm, and the periodic boundary conditions were adopted in all three directions. To promote the system to equilibrium, the slab method was adopted in the simulation (69). The time step in the equilibrium simulation was $dt = 0.5$ ns. The data were collected every 50 μs, with a total time of about 50 ms. All the simulations were performed using the LAMMPS software package (29 October 2020) (70).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abo7885

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. E. Dolgin, What lava lamps and vinaigrette can teach us about cell biology. *Nature* **555**, 300–302 (2018).
2. C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Juelicher, A. A. Hyman, Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).
3. C. P. Brangwynne, T. J. Mitchison, A. A. Hyman, Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4334–4339 (2011).
4. P. Li, S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani, P. S. Russo, Q.-X. Jiang, B. T. Nixon, M. K. Rosen, Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).
5. T. J. Nott, E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowietz, T. D. Craggs, D. P. Bazett-Jones, T. Pawson, J. D. Forman-Kay, A. J. Baldwin, Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**, 936–947 (2015).
6. M. Zeng, Y. Shang, Y. Araki, T. Guo, R. L. Huganir, M. Zhang, Phase transition in postsynaptic densities underlies formation of synaptic complexes and synaptic plasticity. *Cell* **166**, 1163–1175.e12 (2016).
7. R. Parker, U. Sheth, P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**, 635–646 (2007).
8. J. G. Gall, Cajal bodies: The first 100 years. *Annu. Rev. Cell Dev. Biol.* **16**, 273–300 (2000).
9. F.-M. Boisvert, S. van Koningsbruggen, J. Navascues, A. I. Lamond, The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* **8**, 574–585 (2007).
10. K.-H. Lee, P. Zhang, H. J. Kim, D. M. Mitrea, M. Sarkar, B. D. Freibaum, J. Cika, M. Coughlin, J. Messing, A. Molliex, B. A. Maxwell, N. C. Kim, J. Temirov, J. Moore, R.-M. Kolaitis,

T. I. Shaw, B. Bai, J. Peng, R. W. Kriwacki, J. P. Taylor, C9orf72 dipeptide repeats impair the assembly, dynamics, and function of membrane-less organelles. *Cell* **167**, 774–788.e17 (2016).

11. S. Alberti, A. Gladfelter, T. Mittag, Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 (2019).

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

13. R. D. Phair, T. Misteli, High mobility of proteins in the mammalian cell nucleus. *Nature* **404**, 604–609 (2000).

14. Y. Shin, C. P. Brangwynne, Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).

15. C. F. Lee, C. P. Brangwynne, J. Gharakhani, A. A. Hyman, F. Jüelicher, Spatial organization of the cell cytoplasm by position-dependent phase separation. *Phys. Rev. Lett.* **111**, 088101 (2013).

16. J. Berry, S. C. Weber, N. Vaidya, M. Haataja, C. P. Brangwynne, RNA transcription modulates phase transition-driven nuclear body assembly. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5237–E5245 (2015).

17. S. C. Weber, C. P. Brangwynne, Inverse size scaling of the nucleolus by a concentration dependent phase transition. *Curr. Biol.* **25**, 641–646 (2015).

18. J. Berry, C. P. Brangwynne, M. Haataja, Physical principles of intracellular organization via active and passive phase transitions. *Rep. Prog. Phys.* **81**, 046601 (2018).

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

20. E. W. Martin, A. S. Holehouse, I. Peran, M. Farag, J. J. Incicco, A. Bremer, C. R. Grace, A. Soranno, R. V. Pappu, T. Mittag, Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).

21. J.-M. Choi, A. S. Holehouse, R. V. Pappu, Physical principles underlying the complex biology of intracellular phase transitions. *Annu. Rev. Biophys.* **49**, 107–133 (2020).

22. Y. Zhang, B. Xu, B. G. Weiner, Y. Meir, N. S. Wingreen, Decoding the physical principles of two-component biomolecular phase separation. *eLife* **10**, e62403 (2021).

23. T. S. Harmon, A. S. Holehouse, M. K. Rosen, R. V. Pappu, Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *eLife* **6**, e30294 (2017).

24. J.-M. Choi, F. Dar, R. V. Pappu, LASSI: A lattice model for simulating phase transitions of multivalent proteins. *PLOS Comput. Biol.* **15**, e1007028 (2019).

25. M. C. Cohan, A. M. P. Eddelbuettel, P. A. Levin, R. V. Pappu, Dissecting the functional contributions of the intrinsically disordered C-terminal tail of *Bacillus subtilis* FtsZ. *J. Mol. Biol.* **432**, 3205–3221 (2020).

26. X. Zeng, A. S. Holehouse, A. Chilkoti, T. Mittag, R. V. Pappu, Connecting coil-to-globule transitions to full phase diagrams for intrinsically disordered proteins. *Biophys. J.* **119**, 402–418 (2020).

27. I. Seim, A. E. Posey, W. T. Snead, B. M. Stormo, D. Klotsa, R. V. Pappu, A. S. Gladfelter, Dilute phase oligomerization can oppose phase separation and modulate material properties of a ribonucleoprotein condensate. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120799119 (2022).

28. A. Crozat, P. Aman, N. Mandahl, D. Ron, Fusion of chop to a novel RNA-binding protein in human myxoid liposarcoma. *Nature* **363**, 640–644 (1993).

29. L. Malinovska, S. Kroschwald, S. Alberti, Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochim. Biophys. Acta* **1834**, 918–931 (2013).

30. L. Malinovska, S. Palm, K. Gibson, J.-M. Verbavatz, S. Alberti, Dictyostelium discoideum has a highly Q/N-rich proteome and shows an unusual resilience to protein aggregation. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E2620–E2629 (2015).

31. J. Wang, J.-M. Choi, A. S. Holehouse, H. O. Lee, X. Zhang, M. Jahnel, S. Maharana, R. Lemaitre, A. Pozniakovsky, D. Drechsel, I. Poser, R. V. Pappu, S. Alberti, A. A. Hyman, A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699.e16 (2018).

32. S. Hofmann, V. Cherkasova, P. Bankhead, B. Bukau, G. Stoecklin, Translation suppression promotes stress granule formation and cell survival in response to cold shock. *Mol. Biol. Cell* **23**, 3786–3800 (2012).

33. N. Kedersha, S. Chen, N. Gilks, W. Li, I. J. Miller, J. Stahl, P. Anderson, Evidence that ternary complex (eIF2-GTP-tRNAiMet)–Deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol. Biol. Cell* **13**, 195–210 (2002).

34. M. B. Lilly, T. C. Ng, W. T. Evanochko, C. R. Katholi, N. G. Kumar, G. A. Elgavish, J. R. Durant, R. Hiramoto, V. Ghanta, J. D. Glickson, Loss of high-energy phosphate following hyperthermia demonstrated by in vivo 31P-nuclear magnetic-resonance spectroscopy. *Cancer Res.* **44**, 633–638 (1984).

35. A. Patel, L. Malinovska, S. Saha, J. Wang, S. Alberti, Y. Krishnan, A. A. Hyman, ATP as a biological hydrotrope. *Science* **356**, 753–756 (2017).

36. H. Aida, Y. Shigeta, R. Harada, The role of ATP in solubilizing RNA-binding protein fused in sarcoma. *Proteins* **90**, 1606–1612 (2022).

37. S. Sarkar, J. Mondal, Mechanistic insights on ATP's role as a hydrotrope. *J. Phys. Chem. B* **125**, 7717–7731 (2021).

38. S. Sridharan, N. Kurzawa, T. Werner, I. Günthner, D. Helm, W. Huber, M. Bantscheff, M. M. Savitski, Proteome-wide solubility and thermal stability profiling reveals distinct regulatory roles for ATP. *Nat. Commun.* **10**, 1155 (2019).

39. J. Eastoe, M. H. Hatzopoulos, P. J. Dowding, Action of hydrotropes and alkyl-hydrotropes. *Soft Matter* **7**, 5917–5925 (2011).

40. S. Maharana, J. Wang, D. K. Papadopoulos, D. Richter, A. Pozniakovsky, I. Poser, M. Bickle, S. Rizk, J. Guillen-Boixet, T. M. Franzmann, M. Jahnel, L. Marrone, Y.-T. Chang, J. Sterneckert, P. Tomancak, A. A. Hyman, S. Alberti, RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360**, 918–921 (2018).

41. J. Kang, L. Lim, Y. Lu, J. Song, 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).

42. W. M. Babinchak, B. K. Dumm, S. Venus, S. Boyko, A. A. Putnam, E. Jankowsky, W. K. Surewicz, Small molecules as potent biphasic modulators of protein liquid-liquid phase separation. *Nat. Commun.* **11**, 5574 (2020).

43. J. Kang, L. Lim, J. Song, 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).

44. J. Song, Adenosine triphosphate energy-independently controls protein homeostasis with unique structure and diverse mechanisms. *Protein Sci.* **30**, 1277–1293 (2021).

45. M. H. Hayes, E. H. Peuchen, N. J. Dovichi, D. L. Weeks, Dual roles for ATP in the regulation of phase separated protein aggregates in *Xenopus* oocyte nucleoli. *eLife* **7**, e35224 (2018).

46. A. N. Milin, A. A. Deniz, Reentrant phase transitions and non-equilibrium dynamics in membraneless organelles. *Biochemistry* **57**, 2470–2477 (2018).

47. E. S. F. Rosenzweig, B. Xu, L. K. Cuellar, A. Martinez-Sanchez, M. Schaer, M. Strauss, H. N. Cartwright, P. Ronceray, J. M. Plitzko, F. Förster, N. S. Wingreen, B. D. Engel, L. C. M. Mackinder, M. C. Jonikas, The eukaryotic CO_2-concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**, 148–162.e19 (2017).

48. J. D. Schmit, J. J. Bouchard, E. W. Martin, T. Mittag, Protein network structure enables switching between liquid and gel states. *J. Am. Chem. Soc.* **142**, 874–883 (2020).

49. G. Krainer, T. J. Welsh, J. A. Joseph, J. R. Espinosa, S. Wittmann, E. de Csilléry, A. Sridhar, Z. Toprakcioglu, G. Gudiškytè, M. A. Czekalska, W. E. Arter, J. Guillén-Boixet, T. M. Franzmann, S. Qamar, P. St George-Hyslop, A. A. Hyman, R. Collepardo-Guevara, S. Alberti, T. P. J. Knowles, Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun.* **12**, 1085 (2021).

50. P. R. Banerjee, A. N. Milin, M. M. Moosa, P. L. Onuchic, A. A. Deniz, Reentrant phase transition drives dynamic substructure formation in ribonucleoprotein droplets. *Angew. Chem. Int. Ed.* **56**, 11354–11359 (2017).

51. C. Mathieu, R. V. Pappu, J. P. Taylor, Beyond aggregation: Pathological phase transitions in neurodegenerative disease. *Science* **370**, 56–60 (2020).

52. A. N. Semenov, M. Rubinstein, Thermoreversible gelation in solutions of associative polymers. 1. Statics. *Macromolecules* **31**, 1373–1385 (1998).

53. F. Tanaka, Theory of thermoreversible gelation. *Macromolecules* **22**, 1988–1994 (1989).

54. F. Tanaka, Thermodynamic theory of network-forming polymer solutions. 1. *Macromolecules* **23**, 3784–3789 (1990).

55. J.-M. Choi, A. A. Hyman, R. V. Pappu, Generalized models for bond percolation transitions of associative polymers. *Phys. Rev. E* **102**, 042403 (2020).

56. H. Tanaka, R. Kurita, H. Mataki, Liquid-liquid transition in the molecular liquid triphenyl phosphite. *Phys. Rev. Lett.* **92**, 025701 (2004).

57. E. E. Dormidontova, Role of competitive PEO-water and water-water hydrogen bonding in aqueous solution PEO behavior. *Macromolecules* **35**, 987–1001 (2002).

58. C.-L. Ren, I. Nakamura, Z.-G. Wang, Effects of ion-induced cross-linking on the phase behavior in salt-doped polymer blends. *Macromolecules* **49**, 425–431 (2016).

59. J. Wyman, S. J. Gill, Ligand-linked phase changes in a biological system: Applications to sickle cell hemoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5239–5242 (1980).

60. K. M. Ruff, F. Dar, R. V. Pappu, Ligand effects on phase separation of multivalent macromolecules. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2017184118 (2021).

61. K. M. Ru, F. Dar, R. V. Pappu, Polyphasic linkage and the impact of ligand binding on the regulation of biomolecular condensates. *Biophysics Rev.* **2**, 021302 (2021).

62. T. P. Dao, Y. Yang, M. F. Presti, M. S. Cosgrove, J. B. Hopkins, W. Ma, S. N. Loh, C. A. Castañeda, Mechanistic insights into the enhancement or inhibition of phase separation by polyubiquitin chains of different lengths or linkages (2022); https://doi.org/10.1101/2021.11.12.467822.

63. C. Bannwarth, S. Ehlert, S. Grimme, GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).

64. C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, Extended tight-binding quantum chemistry methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **11**, e1493 (2021).

65. T. Lu, Molclus program, version 1.9.9, www.Keinsci.Com/research/molclus.html [accessed 30 October 2020].

66. W. Im, M. S. Lee, C. L. Brooks III, Generalized born model with a simple smoothing function. *J. Comput. Chem.* **24**, 1691–1702 (2003).

67. Semiempirical extended tight-binding program package xtb (version 6.3.3); https://github.com/grimme-lab/xtb [accessed 20 December 2020].

68. T. Hajari, S. Bandyopadhyay, Water structure around hydrophobic amino acid side chain analogs using different water models. *J. Chem. Phys.* **146**, 225104 (2017).

69. G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, J. Mittal, Sequence determinants of protein phase behavior from a coarse-grained model. *PLOS Comput. Biol.* **14**, e1005941 (2018).

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