[1]Jinchen Zhong
[2]Xiaoshan Xu
[3] Chenchen Wang

Computer-Aided Design of Peptides for Suppressing Liquid-Liquid Phase Separation

Journal of Electrical Systems

Abstract: Liquid-liquid phase separation (LLPS) is a fundamental process in cellular biology, critical for the formation of membraneless organelles. Dysregulation of LLPS is linked to various diseases, highlighting the need for targeted therapeutics. Traditional computer-aided drug design has struggled to address LLPS due to the involvement of intrinsically disordered regions (IDRs). In this study, we identified a peptide derived from the low-complexity (LC) region of FUS, with serine-to-glutamic acid mutations (S2E), which significantly suppresses LC phase separation. GPU-based molecular dynamics simulations reveal that the interaction between S2E and LC is essential for their suppressive effect. Inspired by this finding, we proposed a strategy that employs multivalent interactions to design peptides capable of inhibiting phase separation. As a proof of concept, we engineered peptides that combine the LARKS domain from FUS-LC with an E-rich sequence from HSP90, yielding two potent peptides. These peptides effectively inhibit phase separation, as validated by both computer simulation and experimental data. Our findings provide a new avenue for the primary design of peptides aimed at controlling phase separation and protein aggregation, with potential applications in therapeutic development for LLPS-related diseases.

Keywords: Computer simulation, peptide design, phase separation

I. INTRODUCTION

Liquid-liquid phase separation (LLPS) is a critical process that drives the formation of membraneless organelles [1]. Proteins such as FUS, hnRNP A1, and DDX4 fulfill their physiological roles by forming droplets through LLPS [2-4]. These proteins typically possess large intrinsic disordered regions (IDRs), within which multivalent interactions among amino acids occur, thereby facilitating phase separation [5, 6]. Abnormal phase separation is associated with various diseases, including cancer and neurodegenerative disorders, making the regulation of phase separation a burgeoning field in drug discovery [7, 8]. However, due to the predominant role of IDRs in mediating phase separation, traditional protein and small molecule drug design approaches, such docking, have proven less effective. In recent years, with the development of GPU computing power, it has become possible to simulate large-scale and highly dynamic molecules. However, the development of a new computer-aided design method targeting IDRs poses a significant challenge.

Peptides hold great promise as candidates for artificial regulators of LLPS due to their potential to interact with IDRs. Earlier research has shown that polymers of specific amino acids can enhance protein phase separation, such as poly-Arg for FUS [9]. Moreover, the same research group has successfully developed a contact-energy based design strategy with computer simulation to obtain peptides that enhance p53 LLPS [10]. However, until now, a direct approach to designing peptides that inhibit phase separation has not been reported.

In this study, we demonstrated that a peptide derived from the FUS low-complexity (LC) region, with serine residues mutated to glutamic acid (S2E), had the capacity to inhibit LC phase separation. Our findings indicate that multivalent interactions between LC and S2E are the key to S2E's inhibitory effect on LC phase separation. Capitalizing on these insights, we proposed a straightforward and feasible design strategy aimed at effectively generating peptides with the ability to suppress phase separation. We designed and validated these peptides to test the efficacy of this strategy.

II. METHODS

A. Optical Microscopy

The optical microscopy was used to analyze the morphology of droplets formed after phase separation. Images were captured using the E200 optical microscope (Nikon) and analyzed with ImageJ software.

B. GPU accelerated Coarse-grain Simulation

The implicit coarse-grain simulation was conducted using the OpenABC package, employing parameters previously reported by Gregory L. Dignon [11, 12]. The initial equilibration was performed for 100 ns in the NVT ensemble, commencing from a dispersed phase of protein chains in a 15x15x250 box with periodic boundary conditions at 150 K. A time step of 10 fs was utilized for all simulations. Subsequent simulations were conducted

[1,3] Shanghai Advanced Research Institute, Chinese Academy of Sciences;
[1]University of Chinese Academy of Sciences
[2]University of Science and Technology of China
*Corresponding author: Chenchen Wang, Email: morning.365@163.com

for 6 μs under constant temperature and volume conditions, employing a Langevin thermostat with a friction coefficient of 0.01 ps−1. The temperature was incrementally increased from 150 K to the target temperature over the initial 100 ns. The subsequent 2 μs of simulation was discarded, and the remaining data was used for further analysis.

The scripts used for constructing the initial simulation systems and for conducting the product simulations are available in the supplementary files.

C. Simulation Analysis

For cluster analysis, chains with center-of-mass distances of less than 5 nm were classified as belonging to the same cluster. The final 1000 frames from the simulation were utilized to calculate the ratio of released to clustered protein. For contact analysis, we calculated the average contact number over every 1000 frames, repeating this process three times to ensure accuracy. A contact was defined by any two amino acids being within a distance of 1.5 nm. The total contact number for a pair of chains was ascertained by summing the individual contact numbers of each amino acid pair. All analyses were conducted using the MDAnalysis package.

D. Turbidity Analysis

Phase-separated protein samples were incubated for 5 minutes at room temperature in the presence of increasing mass ratios of S2E and SY2E to ensure full mixing. Turbidity was measured by monitoring the optical density (OD) at 600 nm using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific).

E. Purification of Proteins

The vector encoding FUS-LC was kindly provided by Liu's laboratory. DNA fragments encoding peptides were synthesized by Genewiz Inc. These fragments were subsequently inserted into the pGEX-GST vector using the ClonExpress II One Step Cloning Kit (Vazyme; Cat No. C112-01), in accordance with the provided manual. The vectors were then transformed into BL21(DE3) competent cells (Tsingke Biology; Cat No.TSC-E01). Upon reaching an OD 600nm of 1.0 during culturing, the expression of the target proteins were induced by adding 0.1mM IPTG to the medium.

FUS-LC was purified as previously described [13]. FUS LC variants werepurifed similarly to the wild type.

LARKS peptides pellets were resuspended in a buffer containing 20 mM HEPES, 0.5 M sodium chloride at pH 7.5. The resuspended pellets were lysed and cleared by centrifugation. The resulting soluble ERD and LARKS-ERD cell lysates were loaded onto a 5 ml GST-tag purification resin. Proteins were eluted using a gradient of 10 to 20 mM GSH. The GST tag was removed using 3C protease (Yeasen Biology; Cat No. 20409ES60) and the sample was then loaded onto a Superdex 75 sizing column equilibrated with a buffer containing 20 mM HEPES and 0.5 M sodium chloride at pH 7.5. Target protein fractions were collected and their purity assessed by gel electrophoresis. The purified proteins were stored in 20mM HEPES (pH 7.5) at 4°C until they were buffer exchanged into phase separation buffer.

F. Statistical Analysis

Student's t-test was used for a comparison of two independent treatments. The significance threshold was set at $P < 0.05$; indicated as * for $P<0.05$, ** for $P < 0.01$, and *** for $P < 0.001$.

III. RESULTS

A. A Serine-to-Glutamic Acid Mutation in an LC-Derived Peptide Inhibits LC Liquid-Liquid Phase Separation

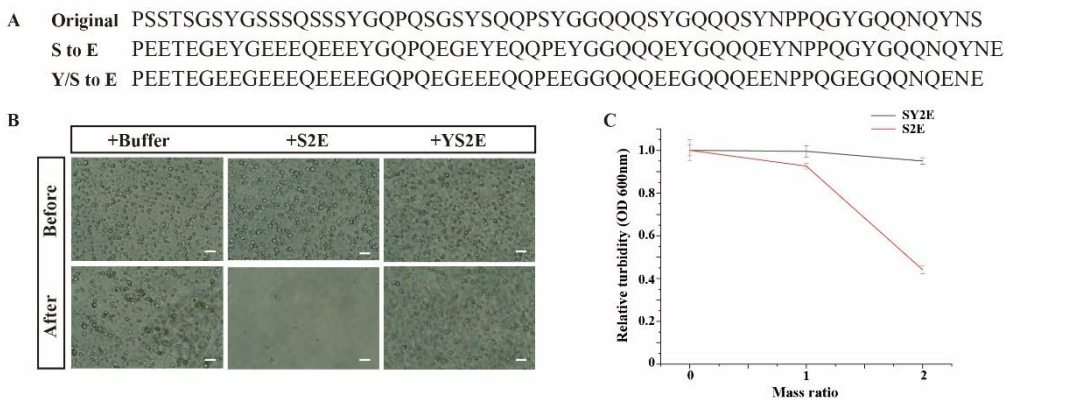

Figure 1: (A) Microscopic images depicting phase-separated droplets in the presence of either buffer or mutated LCs. The mass ratio of mutated LC to LC was maintained at 2:1. Scale bar = 10μm. (B) Turbidity assays showing the effect of mutated LCs on the droplets formation of phase-separated proteins. Turbidity was evaluated at 600 nm with increasing mutated LC mass ratios. A value of 0 denotes the addition of buffer. Data are presented as mean ± SD (n = 3).

Phosphorylation of the low-complexity (LC) region has been shown to inhibit the phase separation of LC proteins [13]. During our investigation into the phosphorylation of LC, we serendipitously discovered a peptide, which

features a serine/threonine to glutamate substitution (S2E), not only does not undergo phase separation itself but also prevents the phase separation of wild-type LC (Figure 1A and 1B). Interestingly, when additional tyrosine mutations were introduced to increase the glutamate ratio within the peptide (SY2E), the suppressive effect on phase separation was abolished. This suggests that the net charge of the peptide is not the determinant of this phenomenon. The suppress of phase separation of S2E peptide was also proven by turbidity assay (Figure 1C). With the increase in S2E concentration, the turbidity of the solution gradually decreased, indicating the dissolve of the droplets. Conversely, in the SY2E group, the turbidity remained almost unchanged.

B. Computer Simulation Reveals the Mechanism of Inhibition Effect of the S2E Peptide

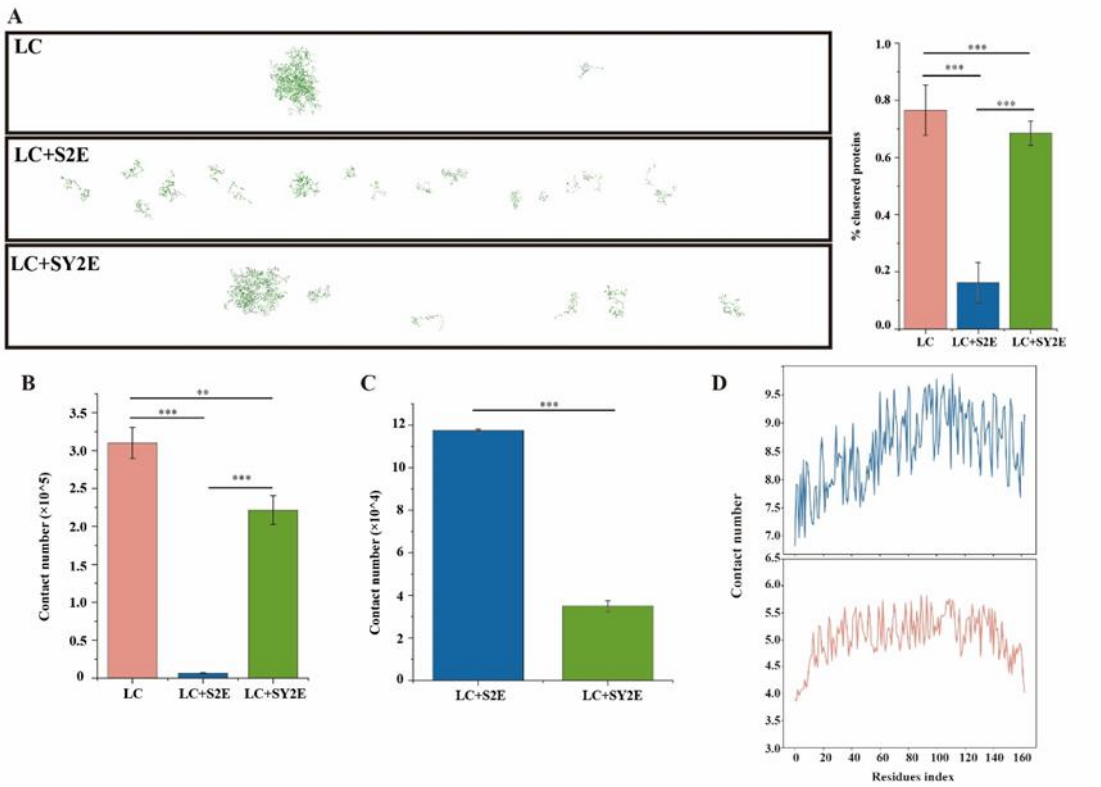

Figure 2: (A) Slab simulation snapshots depict LC and mixtures of LC with S2E and LC with SY2E. Within the mixture systems, only the trajectories of LC molecules are visualized. The histogram on the right shows the percentage of clustered LC molecules across different groups. Data are presented as mean ± SD (n = 1000). (B) The number of intermolecular contacts per frame for LC-LC interactions is presented for both the pure LC simulation system and the LC mixed with S2E or LC mixed with SY2E systems. Data are presented as mean ± SD (n = 3). (C) Comparison of the number of intermolecular contacts per frame between LC-S2E and LC-SY2E groups. Data are presented as mean ± SD (n = 3). (D) The number of intermolecular contacts per residue per frame for LC-S2E (top) and LC-LC (bottom). Abscissa represents the residue index of LC. The contact numbers were averaged across different chains.

To elucidate the underlying mechanism, we conducted molecular dynamics simulations of the LC and S2E peptides using the HPS force fields. In simulations of LC only or LC/SY2E, LC molecules reached an equilibrium state and aggregated into clusters. In contrast, LC molecules evenly dispersed within the simulation box when S2E was present, indicating that the simulations accurately recapitulate the experimental observations (Figure 2A and Movie1-2). Given that multivalent interactions between LC molecules are the primary drivers of their phase separation, we investigated the impact of S2E on these interactions. We observed a significant decrease in the contact number between LC-LC molecules upon mixing with S2E, indicating a weakening of the interaction. In contrast, the contact number exhibited only a slight change upon mixture with SY2E. These suggest that S2E has a notable impact on the interactions between LC molecules (Figure 2B). The contact number between SY2E-LC was significantly less than that between S2E-LC, indicating that LC molecules tend to engage more readily in interactions with S2E (Figure 2C). The interaction sites between LC-S2E were found to be distributed across the entire LC molecule and exhibited a similar pattern to the LC-LC interactions, confirming that S2E can replace the multivalent interactions between LC-LC molecules (Figures 2D). Collectively, these results indicate that S2E inhibits the phase separation of LC through interactions with LC.

C. Design of Peptides for Suppressing Liquid-Liquid Phase Separation

Based on our findings, we propose a hypothesis: The integration of multivalent interaction sites into monodispersed peptides enables the creation of a peptide with the ability to suppress phase separation. To test our hypothesis, we engineered a peptide that combines the LARKS domain of FUS-LC with an E-rich sequence from HSP90 [14]. The LARKS sequence was utilized to facilitate interactions with LC, while the E-rich sequence was designed to make the peptide monodispersed (Figure 3A). Previous studies have shown that the interaction strength between hnRNP A1 and LC is higher than that between LC-LC [15]. Consequently, we also constructed a similar peptide using the LARKS sequence from hnRNP A1, anticipating that it would exhibit a more potent inhibitory effect on phase separation. We employed MD simulations to evaluate the suppressive capability of these peptides. As expected, both peptides demonstrated the ability to suppress phase separation (Figure 3B). The E-rich sequence alone did not show suppression, indicating that this inhibitory effect is not derived from the HSP protein. These peptides both show interaction with LC, and diminish the LC-LC interactions (Figure 3C and D). LARKS2::ERD exhibits a stronger interference with LC-LC interactions, suggesting that it has a more potent inhibitory effect. Both peptides demonstrated a significant increase in the contact number near the LARKS motifs, which was also in line with our expectations (Figure 3E). When the peptide was mixed with LC in vitro, it suppressed LC phase separation (Figure 3F), thereby verifying the accuracy of our simulation and confirming the feasibility of our design strategy.

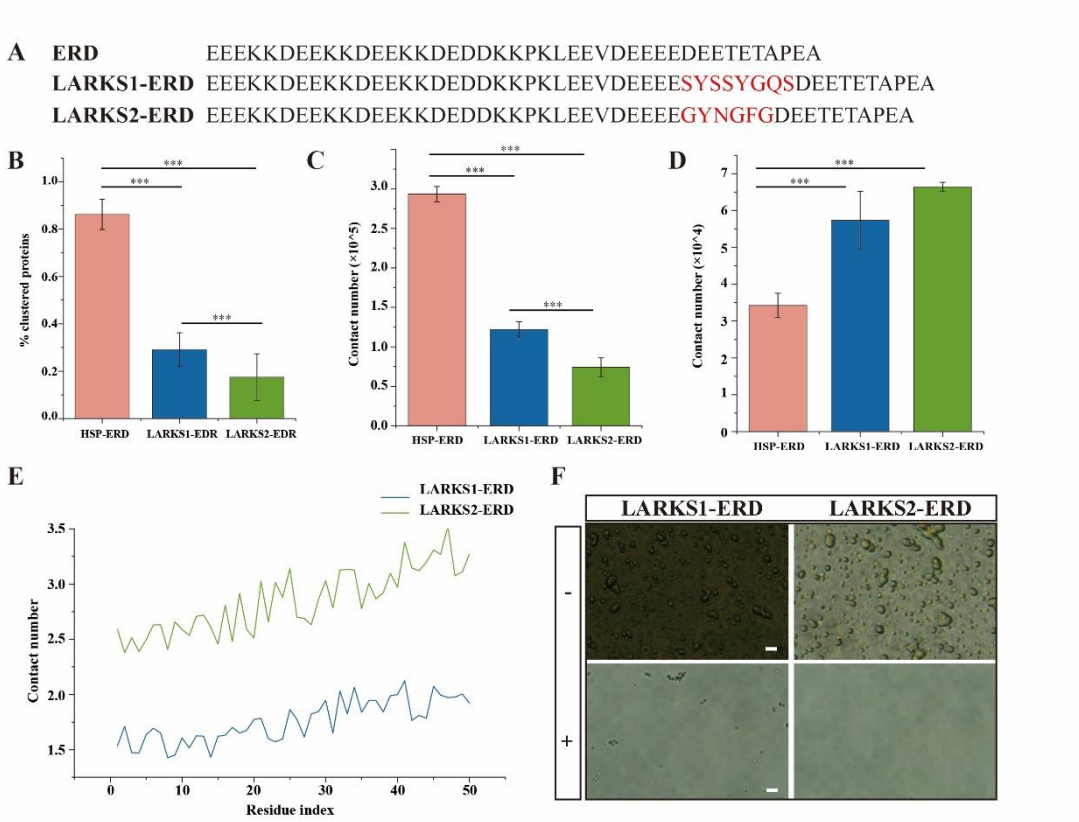

Figure 3: (A) The amino acid sequence of HSP-ERD, LARKS1::ERD and LARKS2::ERD. Amino acids of LARKS were highlighted in red. (B) Histogram shows the percentage of clustered LC molecules across different groups. Data are presented as mean ± SD (n = 500). (C)The contact number per frame for LC-LC interactions is presented for LC mixed with HSP-ERD, LARKS1::ERD and LARKS2::ERD. Data are presented as mean ± SD (n = 3). (D) Histogram illustrates the contact number per frame for both LC-ERD, LC-LARKS1::ERD and LC-LARKS2::ERD interactions. Data are presented as mean ± SD (n = 3). (E) The number of intermolecular contacts per residue per frame for LC-LARKS2::ERD (top) and LC-LARKS1::ERD (bottom). Abscissa represents the residue index of peptides. The contact numbers were averaged across different chains. (F) Microscopic images depicting phase-separated droplets in the presence of LARKS1::ERD and LARKS2::ERD. The mass ratio of peptides to FUS was maintained at 2:1.

IV. DISCUSSION

In China, there is an ancient proverb: 'He who tied the bell to the tiger must take it off.' This adage suggests that the individual responsible for creating a trouble should also be the one to resolve it. We have embraced this philosophical principle in our study of protein phase separation. Our research presents a straightforward yet potent approach for obtaining proteins with the capacity to modulate phase separation. We illustrate that proteins with the ability to inhibit phase separation can be derived from linking the multivalent interaction site of phase-separated

proteins to charged peptides. The resulting peptides maintain their multivalent interactions while simultaneously keep monodispersed in solution. Using this strategy we successfully obtain two peptides that suppress LLPS of FUS. In principle, our design can be applied to any protein with phase separation.

In addition to the glutamate-rich sequences exemplified in our study, there are numerous aspartate-rich or lysine-rich regions present in the proteome. These regions also have the potential to serve as scaffolds for the design of peptides that suppress phase separation. However, due to arginine's inherent role as a sticker site, its effectiveness in this context is uncertain [5]. Furthermore, proteins that undergo phase separation typically exhibit multiple multivalent interaction sites, and each of these sites can serve as a building block for constructing regulatory peptides. This feature renders our strategy highly scalable and versatile.

Similar conclusions have been reached in previous research. A study has identified a peptide derived from the SOD1 protein that can effectively inhibit SOD1 aggregation [16]. This finding suggests that the pattern we have identified could represent a generalizable strategy for discovering sequences capable of not only suppressing liquid-liquid phase separation (LLPS) but also inhibiting protein aggregation.

In parallel, it has been found that acidic amino acids rich sequences typically flank the most potent aggregation-prone regions to avoid protein aggregation [17]. Our strategy for designing peptides to suppress FUS-LC LLPS closely resembles this natural mechanism. This resemblance implies that nature have already utilized a similar strategy throughout evolutionary processes to protect cells from the adverse effects of abnormal protein aggregation. Furthermore, this similarity also substantiates the effectiveness of our strategy from a different perspective.

In summary, our study introduces a novel paradigm for the design of proteins with the capacity to regulate phase separation. Given the established link between aberrant phase separation and a spectrum of diseases, our findings offer innovative avenues for the screening and development of novel polypeptide therapeutics.

ACKNOWLEDGMENT

We extend our sincere thanks to Dr. Liu Cong at the Shanghai Institute of Organic Chemistry (SIOC) for generously providing the plasmid used in our study. We are also grateful to the High-Performance Computing Platform at the School of Life Science and Technology, Shanghai University of Science and Technology, for their invaluable support in supplying the necessary computational resources that facilitated our research.

REFERENCES

1. Brangwynne, C.P., et al., *Germline P granules are liquid droplets that localize by controlled dissolution/condensation.* Science, 2009. **324**(5935): p. 1729-32.
2. Murakami, T., et al., *ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function.* Neuron, 2015. **88**(4): p. 678-90.
3. Kim, H.J., et al., *Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS.* Nature, 2013. **495**(7442): p. 467-73.
4. Nott, T.J., et al., *Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles.* Mol Cell, 2015. **57**(5): p. 936-947.
5. Wang, J., et al., *A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins.* Cell, 2018. **174**(3): p. 688-699.e16.
6. Martin, E.W. and A.S. Holehouse, *Intrinsically disordered protein regions and phase separation: sequence determinants of assembly or lack thereof.* Emerg Top Life Sci, 2020. **4**(3): p. 307-329.
7. Bouchard, J.J., et al., *Cancer Mutations of the Tumor Suppressor SPOP Disrupt the Formation of Active, Phase-Separated Compartments.* Mol Cell, 2018. **72**(1): p. 19-36.e8.
8. Patel, A., et al., *A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation.* Cell, 2015. **162**(5): p. 1066-77.
9. Kamagata, K., et al., *Characterization of design grammar of peptides for regulating liquid droplets and aggregates of FUS.* Sci Rep, 2021. **11**(1): p. 6643.
10. Kamagata, K., et al., *Rational peptide design for regulating liquid-liquid phase separation on the basis of residue-residue contact energy.* Sci Rep, 2022. **12**(1): p. 13718.
11. Dignon, G.L., et al., *Sequence determinants of protein phase behavior from a coarse-grained model.* PLoS Comput Biol, 2018. **14**(1): p. e1005941.
12. Liu, S., et al., *OpenABC enables flexible, simplified, and efficient GPU accelerated simulations of biomolecular condensates.* PLoS Comput Biol, 2023. **19**(9): p. e1011442.
13. Liu, Z., et al., *Hsp27 chaperones FUS phase separation under the modulation of stress-induced phosphorylation.* Nat Struct Mol Biol, 2020. **27**(4): p. 363-372.
14. Hughes, M.P., et al., *Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks.* Science, 2018. **359**(6376): p. 698-701.
15. Farag, M., et al., *Phase separation of protein mixtures is driven by the interplay of homotypic and heterotypic*

interactions. Nat Commun, 2023. **14**(1): p. 5527.
16. Banerjee, V., et al., *Superoxide Dismutase 1 (SOD1)-Derived Peptide Inhibits Amyloid Aggregation of Familial Amyotrophic Lateral Sclerosis SOD1 Mutants.* ACS Chem Neurosci, 2016. **7**(11): p. 1595-1606.
17. Houben, B., et al., *Autonomous aggregation suppression by acidic residues explains why chaperones favour basic residues.* EMBO J, 2020. **39**(11): p. e102864.