Multi-scale Simulations Reveal Molecular Drivers in MUT-16 Scaffold Protein Phase Separations and Client Recognition

Kumar Gaurav,[†,‡,¶] Virginia Busetto,[§,∥] Diego Javier Paez Moscoso,[‡] Arya Changiarath,[¶] Sonya M. Hanson,[⊥] Sebastian Falk,[§,∥] Rene F. Ketting,[‡] and Lukas S. Stelzl[*,†,‡,¶]

†*Institute of Molecular Physiology, Johannes Gutenberg University Mainz, Germany*

‡*Institute of Molecular Biology (IMB), Mainz*

¶*KOMET1, Institute of Physics, Johannes Gutenberg University Mainz, Germany*

§*Max Perutz Lab, Vienna Biocenter Campus (VBC), Austria*

∥*University of Vienna, Center for Molecular Biology, Department of Structural and Computational Biology, Vienna, Austria*

⊥*Flatiron Institute, Center for Computational Biology & Center for Computational Mathematics, New York City, USA*

E-mail: lstelzl@uni-mainz.de

Abstract

Recruitment of biomolecules by phase-separated condensates has emerged as a fundamental organising principle of biological processes. One such process is the RNA silencing pathway, which regulates gene expression and genomic defence against foreign nucleic acids. In *C. elegans*, this pathway involves siRNA amplification at perinuclear germ granules named *Mutator foci*. The formation of *Mutator foci* depends on the phase separation of MUT-16, acting as a scaffolding protein to recruit other components of the Mutator complex. Earlier studies have indicated a crucial role for an exoribonuclease, MUT-7, in RNA silencing. The recruitment of MUT-7 to *Mutator foci* is facilitated by a bridging protein, MUT-8. However, how MUT-8 binds to MUT-16 remains elusive. We resolved the molecular drivers of MUT-16 phase separation and the recruitment of MUT-8 using multi-scale molecular dynamics simulations and *in vitro* experiments. Residue-level coarse-grained simulations predicted the relative phase separation propensities of MUT-16 disordered regions, which we validated by experiments. Near-atomic coarse-grained simulations also capture the relative tendencies of different sequences to phase-separate. Together, coarse-grained simulations at the residue level and near-atomic resolution indicated the essential role of aromatic amino acids (Tyr and Phe) in MUT-16 phase separation. Furthermore, coarse-grained and atomistic simulations of MUT-8 N-terminal prion-like domain with phase-separated MUT-16 revealed the importance of cation-π interaction between Tyr residues of MUT-8 and Arg/Lys residues of MUT-16. By re-introducing atomistic detail into condensates derived from coarse-grained models and conducting 350 µs all-atom simulations on Folding@Home, we demonstrate Arg-Tyr interaction surpasses the strength of Lys-Tyr interactions in the recruitment of MUT-8. The atomistic simulations show that the planar guanidinium group of Arg also engages in sp^2-π interaction, and hydrogen bonds with the Tyr residues and these additional favourable contacts are missing in the Lys-Tyr interactions. In agreement with simulations, the mutation of seven Arg residues in MUT-16 to Lys and Ala weakens MUT-8 binding *in vitro*.

Introduction

The formation of phase-separated condensates is a key mechanism for regulating cellular processes in time and space.[1] The assembly of such condensates is frequently underpinned by intrinsically disordered regions (IDRs) in multi-domain proteins and intrinsically disordered proteins (IDPs). The intrinsic property of IDRs and IDPs is the abundance of possible conformations and the scarcity of secondary structure. Thanks to these properties, IDRs engage in distributed, multivalent, and transient interactions.[1] Further, these characteristics enable IDRs and IDPs to undergo phase separation and form biomolecular condensates.[1,2] An unresolved question regarding biomolecular condensates is how specific regulation is achieved. The key to understanding specific regulation by phase-separated condensates is to elucidate how disordered proteins in phase-separated condensates recognise their binding partners. Previous studies have suggested that protein phase separation is facilitated by electrostatic, hydrophobic, sp^2-π, and cation-π interactions between amino acids.[3–9] The sp^2-π interactions (also called π-π interactions) are facilitated by induced quadrupolar electrostatic attraction between the π orbitals of sp^2 hybridised atoms. These interactions are notably prevalent in the aromatic, amide, carboxyl, and guanidinium groups of amino acid side chains.[6,10] Like sp^2-π, cation-π interactions are also electrostatic in nature. However, they are formed between a positively charged cation of amino or guanidinium groups and π systems of aromatic groups within amino acid side chains.[11–14] The interaction network facilitating the phase separation may also be at play in recruiting additional molecules to condensates. Important progress has been made in understanding interaction patterns and molecular grammar of phase separation.[8,15,16] However, understanding how such interactions underpin molecular recognition on a molecular scale remains elusive.[17]

Molecular recognition by disordered domains of proteins is critical for RNA silencing, also recognised as RNA interference (RNAi), in the model organism *C. elegans*. RNA silencing is an evolutionarily conserved pathway ubiquitous in most eukaryotes and regulates gene expression. It is a primaeval defence mechanism to protect the genome against viruses

and transposons. RNA silencing also plays an essential role in the production of gametes, segregation of chromosomes, and development.[18,19] Key to this pathway is the interaction of small (∼18-30 nt) RNAs with the proteins of the Argonaute family, which enables the regulation of complementary mRNAs at the levels of transcription, translation, and stability.[19,20] There are three types of small RNAs: microRNAs (miRNAs), piwi-interacting RNAs (piRNAs), and short interfering RNA (siRNA) that are synthesised through a variety of mechanisms. One key mechanism of siRNA synthesis is the cleavage of endogenous or exogenous double-stranded RNAs (dsRNAs) by RNase III-like enzyme Dicer.[21,22] The siRNAs produced through this mechanism are called primary siRNAs.

In *C. elegans*, RNA silencing involves a siRNA amplification step facilitated by RNA-dependent RNA polymerases (RdRp).[23–26] This amplification produces secondary siRNAs complementary to the target mRNAs. In *C. elegans* germ cells, siRNA amplification guided by RdRp is choreographed in perinuclear membrane-less organelles called *Mutator foci*.[27] *Mutator foci* are juxtaposed to P granules that accommodate the proteins responsible for RNA silencing and mRNA degradation (Fig. 1A).[27] Biogenesis of *Mutator foci* is underpinned by phase separation of a protein named MUT-16.[27,28]

MUT-16 contains a substantial intrinsically disordered region (IDR) and is a scaffolding protein. The C-terminal segment of MUT-16 (773-1054aa) is required for MUT-16 clustering *in vivo*, which has been suggested to be driven by MUT-16 phase separation. MUT-16 initiates the nucleation of *Mutator foci*. It recruits other proteins essential for RNA amplification, including RNA-dependent RNA polymerase RRF-1, the nucleotidyltransferase MUT-2, the DEAD-box RNA helicase MUT-14, the 3'-5' exoribonuclease MUT-7 and MUT-8 (also known as RDE-2)(Fig. 1B).[27,29] MUT-16 requires MUT-8 as a linker to recruit MUT-7 (Fig. 1B).[28] Further, it was shown that MUT-16 contains a MUT-8 binding region (M8BR, 633-772aa) and that the N-terminus of MUT-8 may interact with MUT-16.[28,30] Finally, the C-terminal domain of MUT-8 was shown to interact with the C-terminal domain of MUT-7.[30,31] Interestingly, a prion-like domain is present in the MUT-8 N-terminus (Fig. S1B). Prion-like

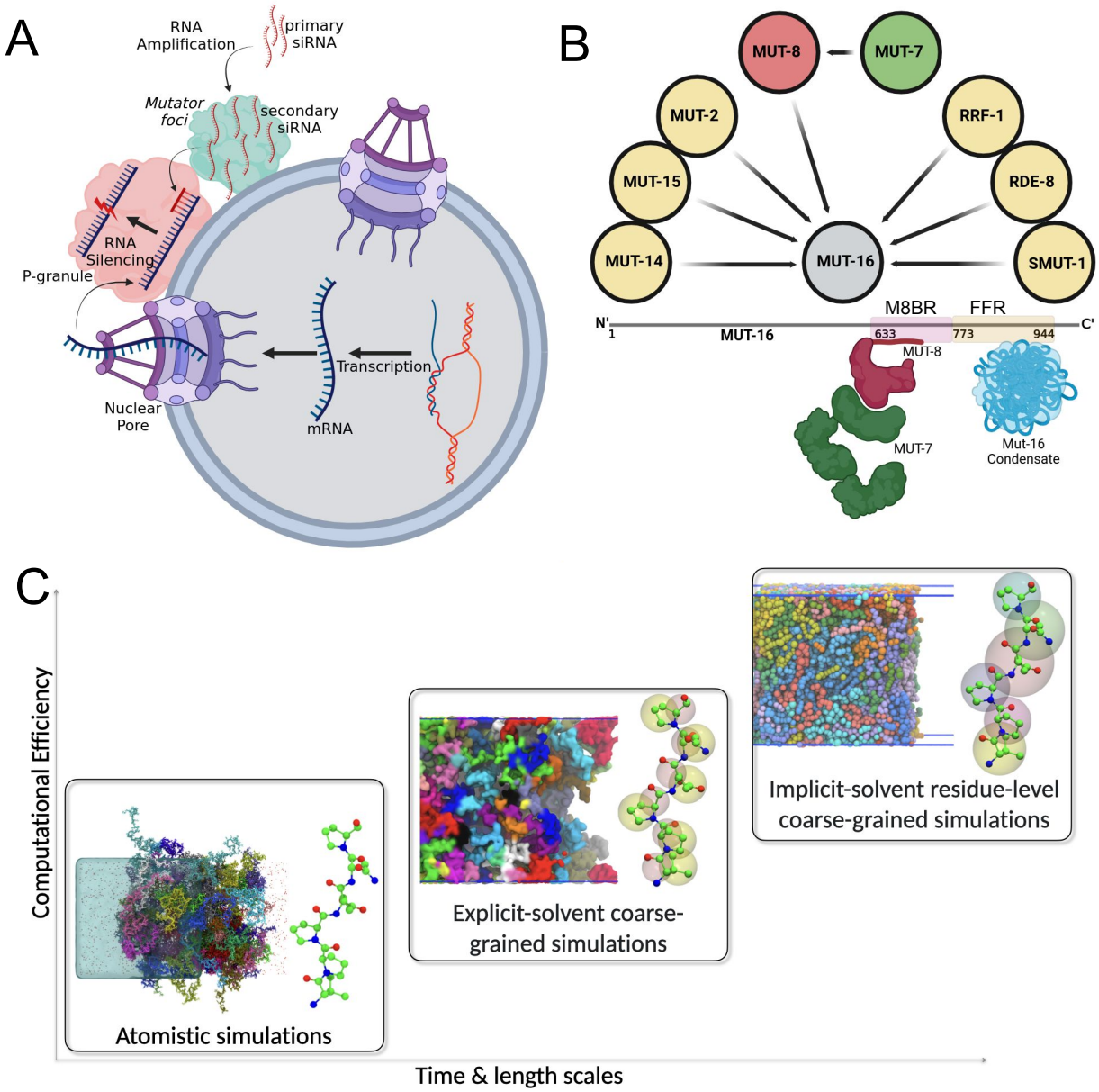
A
RNA Amplification
primary siRNA
Mutator foci
secondary siRNA
RNA Silencing
P-granule
Nuclear Pore
Transcription
mRNA
B
MUT-8
MUT-7
MUT-2
RRF-1
MUT-15
RDE-8
MUT-14
MUT-16
SMUT-1
M8BR
FFR
N'
1
MUT-16
633
773
944
C'
MUT-8
MUT-7
Mut-16 Condensate
C
Computational Efficiency
Atomistic simulations
Explicit-solvent coarse-grained simulations
Implicit-solvent residue-level coarse-grained simulations
Time & length scales

Figure 1: **A.** Illustration of RNA silencing mechanism in *C. elegans*. This process involves the amplification of small interfering RNAs (siRNAs) within a membrane-less organelle known as the *Mutator foci*. The generated siRNAs are then transported to P-granules, another type of membrane-less organelle. Within the P-granules, siRNAs guide the RNA-induced silencing complex (RISC) to target and degrade specific mRNA molecules, thereby silencing gene expression. **B.** The MUT-16 protein (grey) is essential for the formation of *Mutator foci*, acting as a scaffolding protein that recruits other proteins involved in siRNA amplification. MUT-16 recruits proteins (yellow), including MUT-8 (red), which further recruits the exoribonuclease MUT-7 (green). The regions responsible for MUT-16's phase separation and MUT-8 recruitment are highlighted as the foci-forming region (FFR, yellow) and the MUT-8 binding region (M8BR, pink). The intrinsically disordered N-terminal domain of MUT-8 (red) interacts with MUT-16 M8BR, while the C-terminal domain of MUT-8 recruits the C-terminal domain of MUT-7 (green). Chains of MUT-16 FFR (blue) are depicted forming phase-separated condensates. **C.** Multi-scale simulation strategy to identify essential drivers of phase behaviour and molecular recognition of biomolecular condensates combining Atomistic molecular dynamics simulations with explicit solvent, near-atomic coarse-grained (Martini3) simulation with explicit solvent, residue-level coarse-grained (CALVADOS2) simulations with implicit solvent.

domains (PLDs) are often associated with protein phase separation, and the MUT-8 PLD could potentially facilitate its co-condensation with MUT-16, but how the MUT-16 M8BR and MUT-8 N-terminal PLD interact with each other remains elusive.

Molecular dynamics can serve as a "computational microscope" to resolve drivers of phase behaviour and molecular recognition by phase-separated condensates at the molecular scale.[32–35] Atomistic molecular dynamics[36] can resolve interactions of individual amino acids in phase-separated condensates of proteins[34,35] albeit at a sizeable computational cost. In these simulations, the dynamics of each atom in the system are tracked, and interactions with water and ions are considered explicitly. To simulate the phase separation, it is necessary to simplify the protein representations and thus capture the relevant long-time and large-length scales.[32,37] One method of simplifying protein representations involves grouping atoms into distinct representative beads, a process known as coarse-graining. The Martini3 model[38] is a chemically-detailed coarse-grained model where near-atomic resolution and an explicit representation of the water and ions are retained. The Martini3 model holds promise

for simulations of protein, including disordered regions, which are essential drivers for phase separation but have yet to be validated in detail. Further, conducting coarse-grained simulations using CALVADOS2 model,[39,40] with one bead representing each amino acid, facilitates the exploration of protein phase separation behaviours,[41] enabling simulations at the proteomic scale.[42]

In this study, we elucidate how the MUT-16 phase separates and recruits MUT-8 through a comprehensive multi-scale simulation approach (Fig. 1C), complemented by *in vitro* experiments that report on phase behaviour and molecular recognition. We systematically elevate the resolution of our simulation models, progressively transitioning from amino acid bead-string representations to more intricate coarse-grained models that approach near-atomic resolution, eventually culminating in fully atomistic molecular dynamics simulations. Our coarse-grained simulations, employing residue-level coarse-graining and near-atomistic models, underscore the significance of aromatic interactions in stabilising MUT-16 condensates and align with the relative tendencies of protein sequences to phase separation in our experiments. Furthermore, by initiating numerous atomistic trajectories of MUT-16 condensates in complex with MUT-8 fragments using Folding@home,[43] we illustrate that Arg-Tyr interactions, as opposed to Lys-Tyr interactions, play a pivotal role in facilitating the recruitment of MUT-8 N-terminal PLD to MUT-16 condensate, which we validate by binding experiments.

Methods

We performed molecular dynamics simulations in three resolutions (Fig. 1C): implicit solvent residue-level coarse-grained simulations (CALVADOS2),[39,40] near-atomic explicit-solvent coarse-grained simulations (Martini3),[38,44] and atomistic simulations. Combining the simulations at three resolutions enables us to identify residues engaging in key interactions that favour MUT-16 condensate formation and MUT-8 recruitment to MUT-16 condensates and then zoom in on these interactions in atomistic molecular dynamics. Further, we performed *in*

vitro experiments to investigate and validate the MUT-16 phase separation and the recruitment of the MUT-8 N-terminal domain.

Residue-level coarse-grained simulations

Residue-level coarse-grained simulations were executed using the CALVADOS2 model with GPUs using the HOOMD-blue software package (v. 2.9.6)[45] extended with azplugins (v. 0.10.2). The CALVADOS2 model is an improvement of the previously established HPS model,[32] proposed by Tesei et al.[39,40] Within the framework of the CALVADOS2 model, each amino acid is treated as a single bead suspended in an implicit solvent environment. This model enables the investigation of sequence-specific interactions among biomolecules and effectively circumvents the temporal limitations associated with higher-resolution models. The potential energy function employed for the simulations encompasses bonded and non-bonded components, incorporating electrostatic and short-range pairwise interactions. A harmonic potential describes the bonded interactions,

$$U_{\text{bonded}}(r) = \frac{k}{2}(r - l)^2 \tag{1}$$

where the spring constant k = 8368 kJ mol^{-1}nm^{-2} and equilibrium bond length l = 0.38 nm. The non-bonded interaction between the monomer beads is described by Ashbaugh-Hatch potential,[46]

$$U_{\text{AH}}(r) = \begin{cases} U_{\text{LJ}} + \epsilon(1 - \lambda), & r \leq 2^{1/6}\sigma, \\ \lambda U_{\text{LJ}}, & \text{else}, \end{cases} \tag{2}$$

where $\epsilon = 0.8368$ kJ mol^{-1} and U_{LJ} is the Lennard-Jones potential:

$$U_{\text{LJ}}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right], & r \leq r_c^{\text{LJ}}, \\ 0, & \text{else}, \end{cases} \tag{3}$$

where $r_c^{\mathrm{LJ}} = 2.0$ nm.[40] σ and λ are determined by computing the arithmetic average of amino acid specific parameters denoting size and hydrophobicity, respectively. The residue-specific λ were previously optimised by Tesei et al., with a Bayesian method that uses a comprehensive experimental data set.[39,40] The non-bonded interaction also includes an electrostatic component modeled via Debye-Hückel potential:

$$U_{\mathrm{DH}}(r) = \frac{q_{\mathrm{i}} q_{\mathrm{j}}}{4\pi\epsilon r} exp(-r/\lambda_{\mathrm{D}}) \tag{4}$$

where q_{i} and q_{j} are charges. The Debye screening length (λ_{D}) and the dielectric constant (ϵ) are set to 1 nm and 80, respectively to reproduce the physiological conditions. The electrostatic potential is truncated at $r_c = 4.0$ nm.

The simulations were performed with 100 protein chains in the slab geometry of size $20 \times 20 \times 280$nm under periodic boundary conditions (See Table S1).[32] The simulations were initialised by placing the protein chains randomly in the slab. The Langevin thermostat regulated the temperature. Further, the equations of motions were integrated with a timestep (Δt) of 10 fs. The simulations were typically performed for 10 µs (Table S1).

Phase diagram from residue-level coarse-grained simulations

The residue-level coarse-grained (CALVADOS2)[39] simulations were performed at different temperatures (260 K, 265 K, 270 K, 275 K, 280 K, 285 K, 290 K, 291 K, 292 K, 293 K, 295 K, 300 K) for 10 µs each (Table S1). To determine the densities of dense and dilute phases, we followed the approach by Tesei et al.[39] The protein concentration of dilute (ρ_{l}) and dense (ρ_{h}) phases are calculated along the z-axis. The critical temperature T_c was obtained by fitting the densities obtained from the simulation to:

$$\Delta\rho = B(1 - T/T_c)^{\beta} \tag{5}$$

where $\Delta\rho = \rho_{\mathrm{h}} - \rho_{\mathrm{l}}$, B is critical amplitude and $\beta = 0.325$ is critical exponent. Further

we obtained the critical density ρ_c by fitting the simulations data to:

$$(\rho_h + \rho_l)/2 = \rho_c + C(T_c - T) \tag{6}$$

where C is a positive fitting parameter.[32]

Near-Atomic Coarse-Grained Simulations with Explicit Solvent Using the Martini3 Model

The near-atomic coarse-grained (Martini3) simulations were executed using GROMACS,[47] employing either the latest version of the Martini force field (Martini3),[38,44] or an improved variant of Martini3 featuring rescaled protein-water interactions.[48] In the Martini3 force-field framework, multiple atoms are grouped into a single particle, often representing four to five heavy atoms with a single coarse-grained bead.[38] Water and ions are represented explicitly. The initial structure of the protein IDR chains was obtained by taking in desired regions from the AlphaFold[49] structures of MUT-16 (AF-O62011-F1) and MUT-8 (AF-Q58AA4-F1). The protein chains were coarse-grained using the martinize2 python script.[50] The coarse-grained protein chains were inserted in a cubical box or a slab geometry and further solvated with the Insane.py Python script.[51] Furthermore, 0.15 M NaCl was added to the system on top of balancing the charge. Simulations were performed with periodic boundary conditions. The simulation box was subjected to minimisation using the steepest descent algorithm. Firstly, the water and salt were equilibrated in the NVT ensemble by applying the position restraint on the protein. Secondly, the system was further equilibrated by removing the position restraint on the protein in the NVT ensemble. Thirdly, the system was subjected to equilibration in an NPT ensemble for 700000 steps with a 20 fs time-step. Finally, production simulations were performed in the NPT ensemble. The final temperature was maintained at 300 K using the Bussi-Donadio-Parrinello velocity-rescaling thermostat,[52] and the pressure was maintained at 1 bar using Parrinello-Rehman barostat.[53]

All the simulations were performed for 20 µs with a 20 fs time step. Last 10 µs of the trajectory was used for the analysis.

Density calculations from Martini3 simulations

Radial density profiles were calculated to investigate the spatial distribution of the MUT-16 M8BR+FFR, MUT-16 FFR and MUT-16 M8BR chains relative to the center-of-mass (COM) of the system. The methodology was adapted from Zakarya et al.[37] Each bin had a width of 10 Å.

Contact-map analysis

Two-dimensional (2-D) contact maps were generated by computing contact probabilities between residues. Initially, distances between all pairs of residues were averaged across the protein chains. Subsequently, the distance matrix was transformed into a contact matrix by assigning values of 0 and 1 for elements above or below the defined cutoff, respectively. The pair cutoff was set to $2^{1/6}\sigma_{ij}$,[32] 6 Å, and 4.5 Å for CALVADOS2, Martini3, and atomistic simulations, respectively. The resulting contact matrix was then averaged over the frames within the simulation trajectory to obtain the final contact matrix. For atomistic and Martini3 simulations, the contact analysis was based on the procedure described in `https://github.com/dwhswenson/contact_map` and for CALVDAOS simulations MDAnalysis was used.[54,55] One-dimensional contact maps were computed by averaging the values of the final contact matrix along the respective axis. The region-wise contact map was acquired by averaging the values within the specific region of interest from the final contact matrix. On the other hand, the residue-wise contact map was generated by summing the values of each residue pair (e.g., Arg-Tyr, Phe-Met) obtained from the final contact matrix and normalising them with the corresponding number of residue pairs.

Atomistic molecular dynamics simulations

The atomistic simulations were performed for the MUT-16/MUT-8 interaction system obtained from the last frame of the Martini3 simulation trajectory. The system was backmapped using backmap.py python script.[56] For computational efficiency, we reduced the slab dimensions from 15 nm× 15 nm × 60 nm to 15 nm×15 nm×60 nm by selectively removing water molecules from the z-direction. The atomistic simulation system features 884,932 atoms. Simulation was performed using the Amber99sb-star-ildn-q force field[57–60] with TIP4P-D water[61] in GROMACS.[47] The simulation box contained 100 MUT-16 M8BR and 10 MUT-8 N-terminal chains. After backmapping, the atomistic system was minimised using the steepest-descent algorithm to remove clashes. Further, the system was equilibrated in three steps similar to Martini3 simulations. Temperature (300 K) and pressure (1 bar) were maintained using the Bussi-Donadio-Parrinello velocity rescaling thermostat[52] and Parrinello-Rahman barostat.[53] The production run was performed for 1 µs in the NPT ensemble in in-house supercomputer Mogon-II and Further, 100 parallel simulations totalling 350 µs in Folding@home[43] were launched from 100 different conformations taken from the 1 µs trajectory. For each conformation, 100 simulations with different random velocities were launched.

Cation-π and sp^2-π interactions

Cation-π interactions between Arg/Lys and Tyr residues were calculated based on previously established procedures by Zheng et al. and Vernon et al.[6,35] The Arg/Lys and Tyr interaction was filtered for the distance and angle cutoff. The distance cutoff guarantees the magnitude of the vector joining the charged nitrogen of Arg/Lys and the centre-of-mass of the π group in Tyr to be less than 6 Å. Furthermore, the angle cutoff ensures that the absolute cosine of the angle between the previously mentioned vector and the normal vector of the π group in Tyr should be greater than 0.8.

The quantifications of sp^2-π interactions between Arg and Tyr are calculated based on the established methodology reported in the recent literature.[6,35] Firstly, the distance between

the centre-of-mass of the Arg guanidinium group and the Tyr benzene ring is constrained within a cutoff of 8 Å. Secondly, the cosine of the angle between the normal vector of the plane defined by the Arg guanidinium group and the Tyr benzene group should be more than 0.8. Finally, both planes defined by the Arg guanidinium and Tyr benzene groups are elevated by 1.5 Å, and the distance between the centre-of-mass of the two new planes is computed. Pairs exhibiting center of mass distances less than 4 Å are identified as forming sp^2-π interactions.

MUT-16 constructs purification

Three different MUT-16 constructs were purified to test their condensate-formation properties in vitro: MUT-16 633-772 (M8BR), MUT-16 773-944 (FFR), MUT-16 633-944 (M8BR+FFR). The gene fragments coding for MUT-16 (O62011) were cloned into modified pET vectors using ligation-independent cloning. MUT-16 constructs were produced as fusion proteins with an N-terminal MBP-tag and a C-terminal 10xHis-tag in the *E. coli* BL21(DE3) derivatives strain in Terrific Broth (TB) medium. Protein production was induced at 18°C by adding 0.2 mM IPTG for 12-16 hours. Cell pellets expressing the MUT-16 constructs were resuspended in Lysis buffer (20 mM Tris/HCl, 50 mM NaPO4, 500 mM NaCl, 10% (v/v) glycerol, 5 mM 2-mercaptoethanol, pH 7.5) and lysed by sonication. Insoluble material was removed by centrifugation. MUT-16 constructs were purified by immobilised metal affinity chromatography using Ni^{2+}-chelating beads (HisTrap FF; GE Healthcare) and eluted with 20 mM Tris/HCl pH 7.5, 150 mM NaCl, 500 mM imidazole, 10% (v/v) glycerol, 5 mM 2-mercaptoethanol. MUT-16 was dialysed against 20 mM Tris/HCl pH 7.5, 150 mM NaCl, 10% (v/v) glycerol, and 5 mM 2-mercaptoethanol. MUT-16 was subsequently purified using anion exchange chromatography (HiTrap Q, GE Healthcare) and size-exclusion chromatography using an S200 increase 16/600 column (GE Healthcare) in a buffer containing 20 mM Tris/HCl pH 7.5, 150 mM NaCl, 10% (v/v) glycerol, 2 mM DTT. All steps were performed on ice or at 4°C. Proteins were stored at -70°C.

In vitro phase separation assays

Purified MUT-16 constructs, MUT-16 633-772 aa (M8BR), MUT-16 773-944 aa (FFR), and MUT-16 633-944 aa (M8BR+FFR), were diluted to a final concentration of 50 µM in the storage buffer, 20 mM Tris/HCl (pH 7.5), 150 mM NaCl, 10% (v/v) glycerol, and 2 mM DTT. A 1:1 serial dilution was performed in 8-strip PCR tubes (Multiply-μstrip Pro 8-strip, Sarstedt, REF 72.991.002) for droplet formation. For each construct, the reaction was done in two parallel serial dilutions. The first one, without adding 3C protease, was used as a control. The second serial dilution included the addition of 3C protease (1 mg/mL; in-house produced) at a 1:100 (w/w) ratio of 3C protease to the MUT-16 fragment. 3C protease removed the N-terminal MBP tag from the constructs and induced droplet formation. After incubating the reaction for 60 minutes at room temperature to allow cleaving, the reaction mixture was added to a slide that had been previously attached with 1 cm x 1 cm frames (Thermo Scientific, AB-0576) and then covered with a cover slip 20x20 mm (Roth Karlsruhe). Finally, the slides were imaged using Thunder (Leica), an inverted widefield microscope in bright field mode, with a 100x/1.44 oil lens and a 310 ms exposure time. Images were analysed, and the contrast was adjusted using Fiji/ImageJ (2.14.0/1.54f).

Co-expression pulldown assays

The gene fragments coding for MUT-8 (Q19672) 1-235 and MUT-16 (O62011) 584-724 were cloned into modified pET vectors using ligation-independent cloning. The MUT-16 (584-724) gene fragments carrying mutations of seven Arginine residues (Arg 642, Arg 657, Arg 658, Arg 682, Arg 689, Arg 698, Arg 699) into either Alanines or Lysines were ordered from Integrated DNA Technologies (IDT). MUT-16 (584-724) constructs were produced as fusion proteins with an N-terminal GST-tag, while MUT-8 (1-235) was produced as fusion proteins with an N-terminal MBP-tag. Vectors carry different antibiotic resistance markers to allow co-expression of MUT-8 and MUT-16 in E. coli. Plasmids containing the genes coding for MUT-16 (584-724) and MUT-8 (1-235) were co-transformed into BL21(DE3)

derivative strains to allow co-expression. Cells were grown in 50 mL TB medium shaking at 37°C, and protein production was induced at 18°C by adding 0.2 mM IPTG for 12-16 hours. Cell pellets were resuspended in 4 mL of Lysis buffer (50 mM NaH2PO4, 20 mM Tris/HCl, 250 mM NaCl, 10 mM Imidazole, 10% (v/v) glycerol, 0.05% (v/v) IGEPAL, 5 mM 2-mercaptoethanol pH 7.5). Cells were lysed by sonication, and insoluble material was removed by centrifugation at 21,000xg for 10 minutes at 4°C. 500 µL of the supernatant was applied to 35 µL amylose resin (New England Biolabs) or glutathione resin (Cytiva) and incubated for 1-2 hours at 4°C. Subsequently, the resin was washed thrice with 500 µL of Lysis buffer. The proteins were eluted in 50 µL of Lysis buffer supplemented with 10 mM maltose (amylose resin) or 20 mM of reduced glutathione (glutathione resin), respectively. Input and eluate fractions were analysed using SDS–PAGE and Coomassie staining.

Results and discussion

Residue-level coarse-grained simulations predict the tendency of MUT-16 IDRs to phase separate *in vitro*

Simulations with the residue-level coarse-grained (CALVADOS2) model[39,40] captured the relative phase-separation propensities of MUT-16 IDRs in line with *in vitro* experiments and highlighted the factors determining MUT-16 phase behaviour. Previous studies have observed that at ambient temperature, MUT-16 protein foci *in vivo* and that MUT-16 nucleates the formation of *Mutator foci*.[27,28] Uebel et al.,[28] observed that the C-terminal region of MUT-16 (773-1050 aa) is essential for forming these foci, attributed to protein phase separation. Furthermore, an intermediate region spanning amino acid 633 to 772, designated as M8BR, is required to recruit the MUT-8 protein but is dispensable for the phase separation of MUT-16. For this study, we used the MUT-16 isoform with Uniprot ID: O62011 to understand the role of disordered regions. We excluded the region from amino acids 945 to 1054 from the C-terminal domain due to the tendency of this region to adopt a three-dimensional

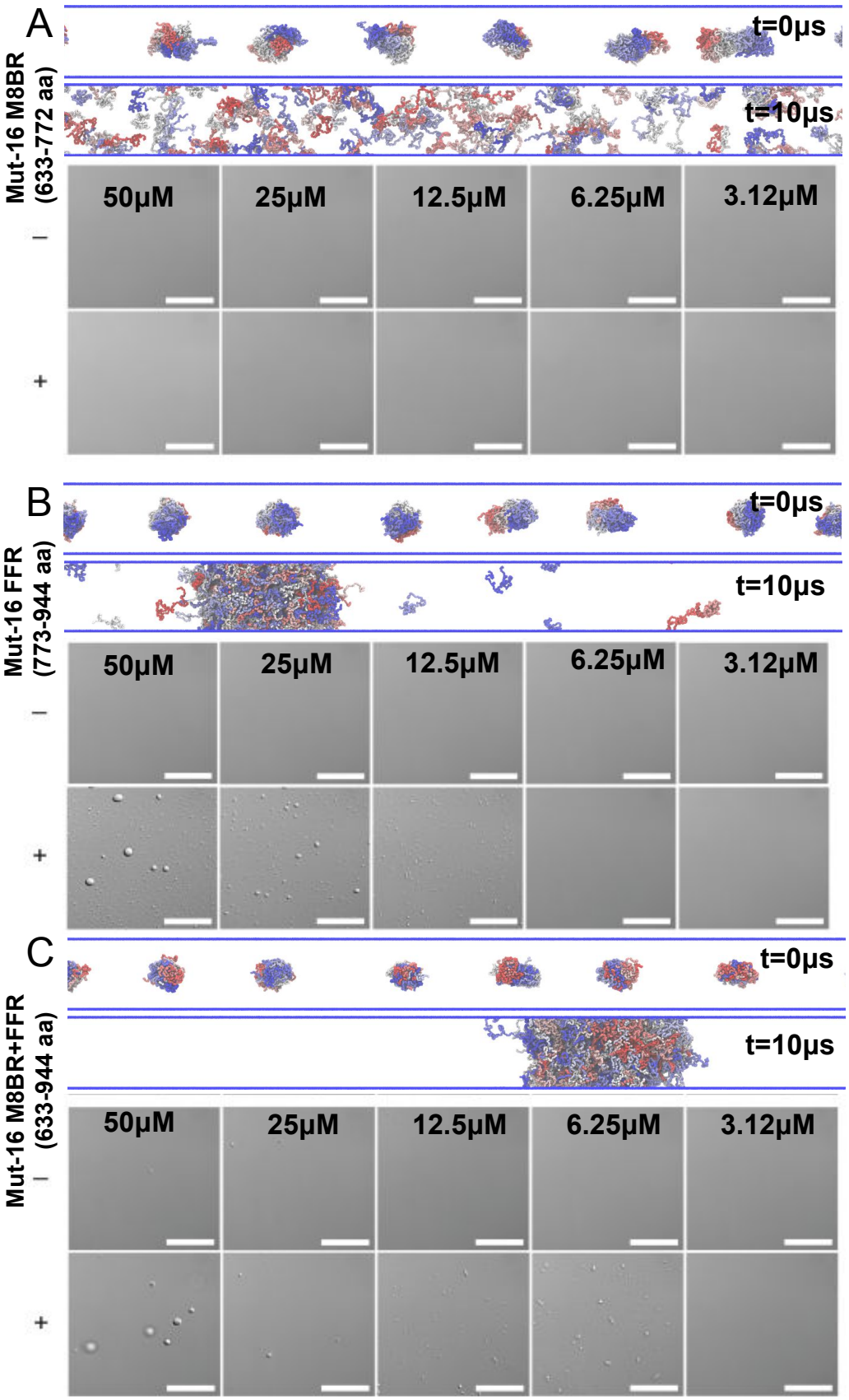
A
t=0μs
t=10μs
Mut-16 M8BR
(633-772 aa)
50μM
25μM
12.5μM
6.25μM
3.12μM
−
+
B
t=0μs
t=10μs
Mut-16 FFR
(773-944 aa)
50μM
25μM
12.5μM
6.25μM
3.12μM
−
+
C
t=0μs
t=10μs
Mut-16 M8BR+FFR
(633-944 aa)
50μM
25μM
12.5μM
6.25μM
3.12μM
−
+

Figure 2: MUT-16 FFR is sufficient to drive the phase separation. **A.** Panel above shows the initial (t=0 µs) and final (t=10 µs) frame of MUT-16 M8BR simulation in a slab geometry. In the panel below, representative bright-field images show the phase separation assays of the MUT-16 M8BR (633–772 aa) construct as a function of protein concentration. The concentration is indicated at the top of each image. Phase separation was induced by cleaving the MBP N-terminal tag by adding 1:100 (w/w) 3C protease to the MUT-16 fragment. The presence of 3C is represented as +, and its absence is indicated by – on the left part of the panel **B.** Simulation and *in vitro* observations of MUT-16 FFR construct obtained through the method described in **A**. **C.** Simulation and *in vitro* observations of MUT-16 M8BR+FFR construct obtained through the method described in **A**. Scale bar correspond to 20 µm.

structure (Fig. S1A). Hence, we first simulated a construct designated as FFR (773-944 aa) (Fig. 2). Subsequently we simulated the putative MUT8 binding region independently (M8BR, 633-772 aa) and in conjunction with the FFR (M8BR+FFR, 633-944 aa). Mut-16 M8BR chains were not observed to phase separate spontaneously (Fig. 2A). However, FFR and M8BR+FFR chains phase separate (Fig. 2B,C), with dilute phase concentration measured at around 0.28 ± 0.04 mM and 0.15 ± 0.06 mM, respectively. Experimental observation also indicated the lack of phase separation of MUT-16 M8BR at all the concentrations up to 50 µM, in line with the simulations (Fig. 2A). MUT-16 FFR was observed to phase separate *in vitro* at concentrations 12.5 µM and higher. In MUT-16 FFR, the number and size of droplets increase with the concentration (Fig. 2B). Similarly, the MUT-16 M8BR+FFR was found to also phase separate at 50 µM, 25 µM, and 12.5 µM concentrations but also at lower concentration (6.25 µM). In MUT-16 M8BR+FFR, the size of droplets also increases upon increasing the concentration, but unlike MUT-16 FFR (Fig. 2B), the number of droplets decreases (Fig. 2C). Furthermore, the timelapse of MUT-16 FFR droplets obtained *in vitro* were observed to fuse and wet the coverslip (Fig. S2), suggesting the liquid-like behaviour of the droplet. The simulations and *in vitro* experiments demonstrate that the FFR segment of MUT-16 is sufficient to drive phase separation.

Further, we performed simulations of MUT-16 M8BR+FFR at several temperatures between 260 K to 300 K to study the temperature dependence of phase separation. Based on

experimental findings from Uebel et al.,,[28] we expect that MUT-16 M8BR+FFR will cluster together, possibly via phase separation at lower temperatures but not high temperatures. Indeed, at lower temperatures, the MUT-16 M8BR+FFR protein chains phase separates and forms a dense protein-rich phase surrounded by a dilute phase depleted in proteins (Fig. 3A). As the temperature increases, MUT-16 M8BR+FFR chains are progressively observed in the dilute phase (Fig. 3A). This is in qualitative agreement with experiments by Uebel et al.,[28] where MUT-16 foci disappeared as the worms were exposed to elevated temperature (∼303 K). The condensates reappeared after lowering the ambient temperature (∼294 K). We further established the phase diagram by calculating the concentration of MUT-16 chains in the dense and the dilute phase (Fig. 3B). The phase diagram revealed the critical temperature (T_c) to be ∼296 K. The agreement of simulations with a transferable physics-based model[32,39,41] with *in vivo* behaviour adds theoretical support to the interpretation of the temperature-dependent loss of MUT-16 foci in *C. elegans* as upper critical solution temperature (UCST) phases separation.[5]

The simulations further reveal the regions and amino acids responsible for driving the phase separation. We calculated the inter-molecular contact probabilities of MUT-16 M8BR+FFR residues and observed that the FFR exhibits a stronger propensity to interact (Fig. 3C) compared to the M8BR. While residues throughout the sequences engage in interactions in the inter-molecular contact map (Fig. 3C), residues from around 663 to 713 in M8BR make relatively few interactions with residues in this region compared to residues in the remainder of the protein. The region-wise contact map demonstrates a higher FFR:FFR interaction probability followed by FFR:M8BR and M8BR:M8BR (Fig. S3A,B). The simulations indicate an essential role of FFR in driving the phase separation and validate the experimental observations by Uebel et al.,.[28] Although the M8BR has a lower probability of interaction, it maintains considerable interactions with the FFR. The 1-D contact map highlights a heightened interaction probability of aromatic residues like Tyr and Phe (Fig. 3D). Notably, the Tyr residues display the most prominent and abundant peaks, followed by Phe. The aro-

matic amino acids are relatively plentiful in the FFR (14%) compared to the M8BR (5%), possibly enabling the FFR to drive the phase separation. Additionally, the 1-D contact map shows (Fig. 3D) elevated interaction propensities for positively charged amino acids such as Arg and Lys. The amino acids F695, R698, and R699 are among the residues forming the most vital interactions in the condensate. Lys residues generally show a somewhat lower tendency to interact. A stretch of Lys in M8BR (671,674, and 678) shows only small spikes in contact probability (3D). Apart from relatively strong Tyr interaction probabilities in FFR, we found significant Tyr peaks within amino 700-772aa (724, 733, 740 and 754) of the MUT-16 M8BR. This observation aligns with the observations by Uebel et al.[28] Their experiments revealed that the foci formation of GFP-tagged MUT-16 construct spanning amino acid 704 to 1050 forms larger condensates compared to a shorter MUT-16 construct spanning amino acid 773 to 1050.[28] We computed the relative frequency of the interaction of amino acid pairs normalised (Fig. S4A) and unnormalised (Fig. S4B). The normalised contact plot indicates the relative strengths of amino acid interactions. Tyr-Tyr, Tyr-Arg, Tyr-Phe, Phe-Arg, Phe-Phe interactions are very common, suggesting that these residues engage in attractive interactions (Fig. S4A). Tyr and Phe generally form secondary interactions with Leu, Ile, and Met. Oppositely charged Arg-Asp also forms frequent contacts. Looking at the normalised interactions provides a better view of what drives phase separation, as many weak interactions can be equally or more important than a more limited number of relatively strong interactions.[16] The unnormalised plot indicates that Pro-Pro, Glu-Glu, Pro-Glu, Pro-Tyr,[62] Glu-Tyr and Asn-Tyr are the most abundant interactions (Fig. S4B). The relative abundance of interactions involving Glu, in particular Glu-Glu interactions, likely stems from the fact that Glu residues are frequently flanked by positively charged residues in MUT-16 M8BR+FFR (See Table S2). The unnormalised plot reflects the role of the abundance of Pro in MUT-16 M8BR+FFR phase separation, as these proline-based interactions are not prominent in the normalised plot. While some residue pairs are more critical, it is clear that other residues also interact.[16] Overall, our simulations rationalised

why the C-terminal region, including the FFR region with its aromatic residues like Tyr and Phe, plays an essential role in the phase separation of MUT-16.

MUT-16 phase separation in explicit solvent coarse-grained Martini3 simulations

To better characterise the molecular drivers of MUT-16 phase behaviour, we sought to employ the near-atomic resolution Martini3 coarse-grained model.[38] Martini3 is a powerful general-purpose coarse-grained model of biomolecules, which has also been used to simulate multi-domain and disordered proteins.[44,48,63,64] Unlike CALVADOS2, Martini3 has not been developed or extensively tested for the phase separation of disordered proteins.[44] Thus, comparing simulations with this model to other simulations and experiments is important. In our simulations, we observed the spontaneous phase separation of the MUT-16 M8BR+FFR (Fig. 4A), MUT-16 FFR and MUT-16 M8BR chains suspended in explicit solvent. The MUT-16 M8BR+FFR chains forms a droplet within 2 µs of simulations and the droplet sustains until 20 µs (Fig. 4A). Since the Martini3 force-field is known to overestimate protein-protein interactions, we have re-scaled the protein-water interactions[48,65] by a rescaling factor λ. We reasoned that if the propensity of MUT-16 M8BR+FFR, MUT-16 FFR and MUT-16 M8BR phase separation, captured by Martini3, results from favourable interactions in the simulations, then increasing protein-water interactions somewhat should not abrogate condensation. We observed that at lower λ values of 1.00 and 1.01, the MUT-16 M8BR+FFR, MUT-16 FFR and MUT-16 M8BR condensates persisted (Fig. S5A,B,C) as shown by the density profiles. At larger λ values of 1.05, all three condensates dissolved(Fig. S5A,B,C). However, at an intermediate lambda value of 1.03, the MUT-16 M8BR+FFR and MUT-16 FFR condensates persist, but MUT-16 M8BR condensate dissolves (Fig. S5D,E,F). Rescaling the relative strength of protein-water interactions compared to protein-protein interactions can be analogous to increasing the effective temperature.[37] Thus, the simulations with the Martini3 model are also in qualitative agreement with the dissolution of the condensates at

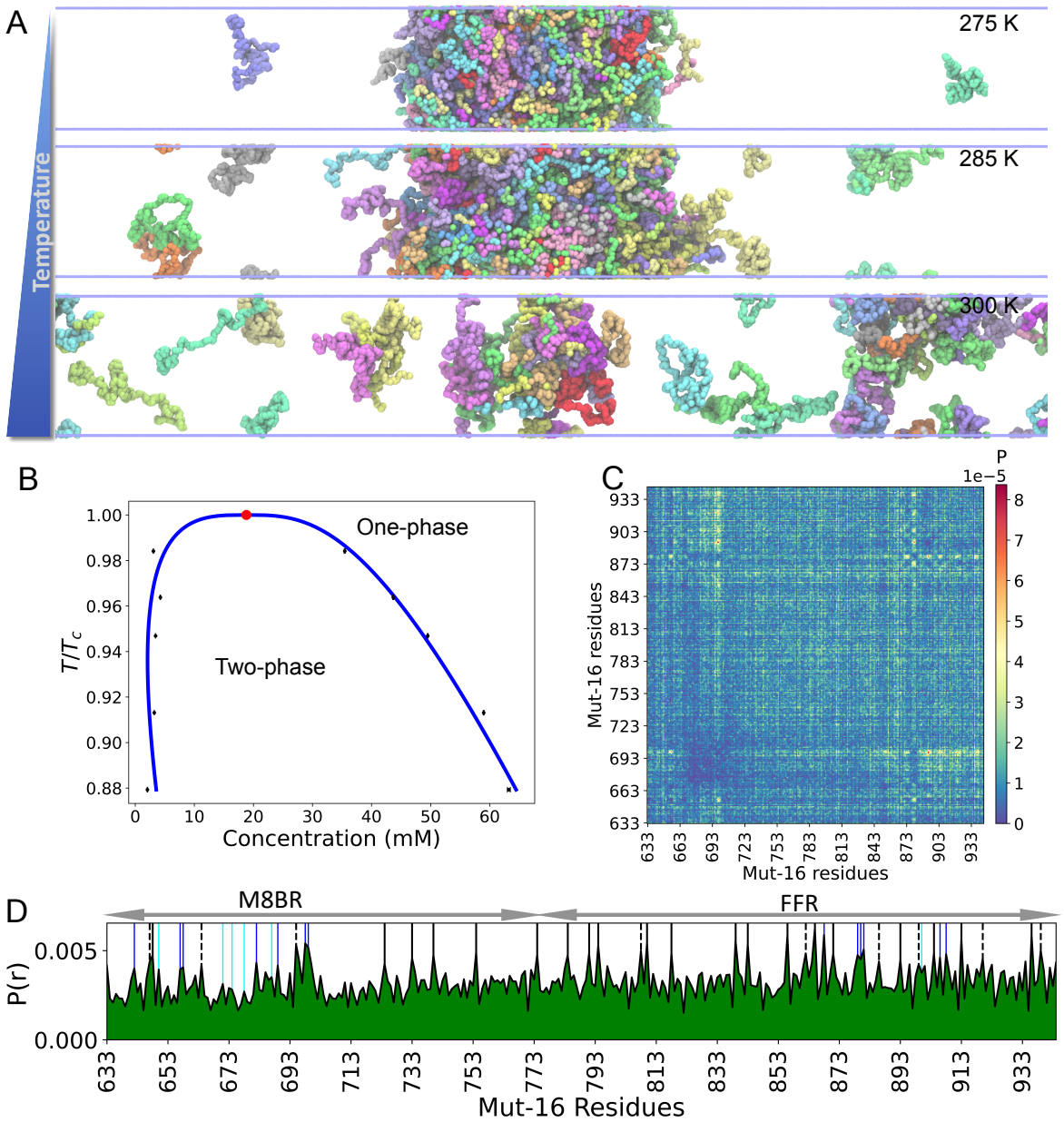

Figure 3: Phase behaviour of MUT-16 condensate in residue-level coarse-grained (CALVADOS2) simulations in implicit solvent. **A.** Simulations of 100 MUT-16 M8BR+FFR chains at different temperatures (275 K, 285 K, and 300 K). The colour indicates the different chains. **B.** Phase diagram of MUT-16 obtained by CALVADOS2 simulation at different temperatures. Black dots represent the concentration of dilute and dense phases. The critical point is depicted by a red dot at $T/T_c = 1$. Error bars are smaller than the marker size. **C.** 2-D inter-chain contact-map obtained from the simulation of MUT-16 M8BR+FFR chains at 275 K. **D.** Relative probability $P(r)$ of an amino acid to form a contact in the MUT-16 (M8BR+FFR) protein sequence. Peaks shown in the 1-D plot are relative to aromatic amino acids like Tyr (black, solid) and Phe (black, dashed) and positively charged amino acids like Arg (blue) and Lys (cyan).

higher temperatures in our simulations with the CALVADOS2 model and the temperature dependence Uebel et al., observed *in vivo*.

Further, we characterised the regions of MUT-16, which form many contacts and might stabilise the condensate. We computed the average pairwise contacts between the residues in each chain per frame, represented in a 2-D contact plot (Fig. 4B). The contact plot suggests that many contacts are congregated in the FFR (773-944aa) of MUT-16 (Fig. 4B). The observations align with our CALVADOS2 simulations and the experimental results obtained by Uebel et al.,.[28] Our results further suggest that a sub-region of MUT-16 (850-940aa) is particularly important for phase separation. Further, we computed the 1-D contact map (Fig. 4D). The 1-D map indicates significant peaks associated with aromatic amino acids such as Tyr and Phe. Additionally, supplementary peaks of positively charged amino acids like Arg and Lys are evident. We also investigated the region-wise contact map to elucidate the competition between M8BR and FFR in orchestrating the phase separation of MUT-16. The contact map reveals that the FFR:FFR interactions are not significantly more frequent than M8BR:M8BR interactions(Fig. S3B). This observation made from Martini3 simulations is in contrast to the CALVADOS2 simulations in which the M8BR:M8BR interactions are significantly weaker than FFR:FFR interaction (Fig. S3A). The 2-D contact map indicates a more substantial interaction frequency between the N-terminal region (640-700aa) and the C-terminal region (701-773aa) of MUT-16 M8BR (Fig. 4B). These interactions are missing in the 2-D contact map obtained from CALVADOS2 simulations (Fig. 3C). Furthermore, the comparison of the 1-D contact map of Martini3 and CALVADOS2 simulations indicates the higher contribution of MUT-16 M8BR N-terminal region (640-700aa) in Martini3 simulations compared to the CALVADOS2 (Fig. S3C). Since the sequence of MUT-16 M8BR is populated with the positively charged amino acids (Arg and Lys) in the N-terminal domain (640-700aa) and aromatic amino acid in the C-terminal domain (700-772aa), we hypothesise that M8BR:M8BR interaction could be facilitated by cation-π or sp^2-π interactions. We posit that cation-π or sp^2-π interactions are probably overestimated in the Martini3 model resulting

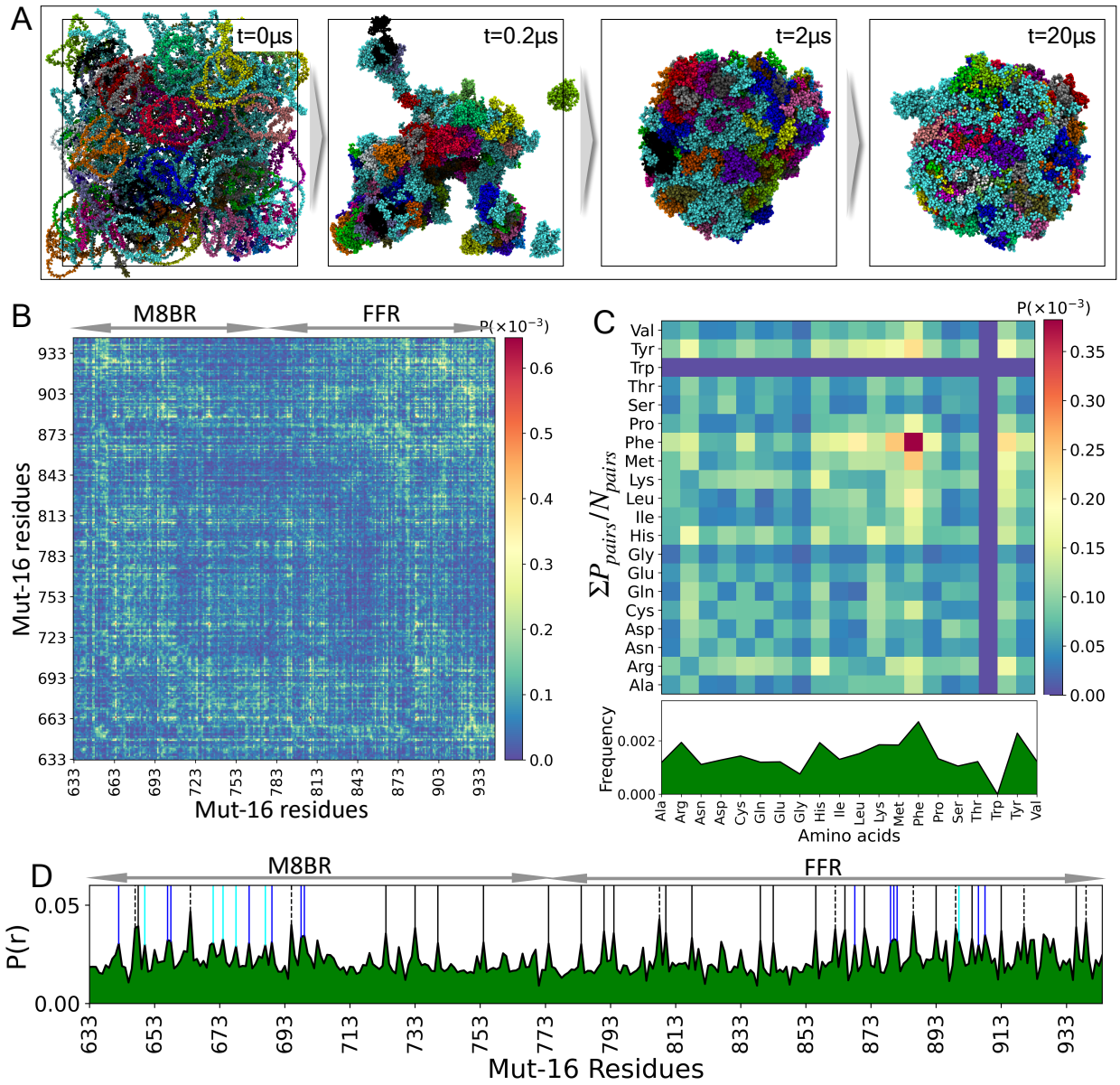

Figure 4: MUT-16 phase separation in explicit-solvent coarse-grained (Martini3) simulations. **A.** Time evolution of MUT-16 M8BR+FFR phase separation. **B.** Inter-molecular contact probabilities in the condensed phase of MUT-16 M8BR+FFR as a function of the amino acid sequence. **C.** Inter-molecular contact probability of MUT-16 M8BR+FFR as a function of residue type normalised by the relative abundance of amino acid. **D.** One-dimensional contact map representing the peaks associated with the amino acid with higher relative contact probabilities (P(r)). The peaks shown by Tyr (black), Phe (dashed), Arg (blue), and Lys (cyan) are represented by the vertical lines.

in the difference in the M8BR:M8BR interactions compared to CALVADOS2. Unlike in the CALVADAOS2 simulations, Lys 671, Lys 674, and Lys 678 have a similar relative probability of forming contacts as Arg 657 and Arg 658 (Fig. S3C), which suggests that Lys is more sticky in Martini3 than in CALVADAOS2. We also computed an amino acid-based contact map to shed light on the contributions of individual amino acids to the phase separation process. Notably, Tyr, Phe, Arg, and His were found to play crucial roles in the phase separation of MUT-16 (Fig. 4C). Interestingly, we observed that the strength of Phe-Phe interactions was greater than Tyr-Tyr interactions despite the higher abundance of Tyr residues in Mut16 (M8BR+FFR). Furthermore, the unnormalised amino-acid based contact map indicates a higher propensity of Pro-Tyr[62] and Pro-Pro interactions, followed by secondary interactions by Glu, Asn and His. Nonetheless, the relative propensities for residues to engage in inter-chain interactions in the Martini3 simulations are similar to the propensities from simulations with the CALVADOS2 model[40] (Fig. S3C).

Recruitment of MUT-8 N-terminal PLD by MUT-16 condensate

Our coarse-grained simulations revealed the molecular drivers of MUT-8 recruitment to the MUT-16 condensate (Fig. 5). Experiments performed by Uebel et al.,[28] suggest that a particular region of MUT-16 (633-772aa, M8BR) is responsible for recruiting MUT-8. MUT-8 was not observed to localise to the phase-separated MUT-16 condensate upon deleting the M8BR region.[28] However, in previous studies, the localisation of MUT-8 has not been determined based on deleting the MUT-16 C-terminus. Therefore, the experiments do not exclude the auxiliary role of the C-terminus in MUT -8 recruitment. Busetto and Falk have recently shown that the MUT-8 C-terminal domain is critical for recruitment of the MUT-7 exoribonuclease.[31] Still, a specific functional role of the N-terminal domain of MUT-8 remains elusive.[30] Interestingly, we found that the first 51aa of MUT-8 is predicted as a prion-like domain (PLD) (Fig. S1).[66] PLDs are known to be enriched in uncharged polar amino acids (Tyr, Asn, Gln) and Gly[67] and are present in proteins with propensity of phase separation like

FUS[68] and TDP-43.[69,70] Aromatic amino acids of PLDs are known to drive phase separation and recruitment to existing biomolecular condensate.[8,9,71,72] The presence of a PLD in the MUT-8 N-terminal domain makes it a promising candidate for interacting with the MUT-16 condensate. We employed residue-level CALVADOS2 (Fig. S7A) model and chemically more detailed near-atomic resolution coarse-grained Martini3 (Fig. 4A) model to elucidate the interaction network of MUT-16 M8BR+FFR and MUT-8 N-terminal domain(Uniprot ID: Q58AA4).

In simulations with both models, the MUT-16 M8BR+FFR condensate maintained its stability throughout its interaction with the MUT-8 N-terminal as we expected based on the simulations of the MUT-16 M8BR+FFR without MUT-8 (Fig. 4A). By analysing both CALVADOS2 (Fig. S7C) and Martini3 (Fig. 5B) simulations, we extracted the contact probabilities of amino acids in both MUT-16 and MUT-8, revealing critical region and residues crucial for MUT-8 recruitment. Comparing the contact map obtained from CALVADOS2 and Martini3, we observed: (i) Recruitment of MUT-8 appears to be facilitated by the MUT-16 FFR in CALVADOS2 simulations (Fig. S7A) as these regions form very frequent contacts in our simulations. (ii) In the case of Martini3 simulations, the recruitment is driven by the MUT-16 M8BR (Fig. S7B). Both simulations indicate that the recruitment of MUT-8 could be facilitated by both the M8BR and FFR of MUT-16 (Fig. 5B,S7D,E). Furthermore, we observed that the positively charged amino acids (Arg and Lys) and the aromatic amino acids (Tyr and Phe) in MUT-16 M8BR+FFR facilitate the recruitment of MUT-8 N-terminal domain (Fig. S7D). The positively charged residues (Arg and Lys), segregated in a particular region of MUT-16 M8BR (640-700aa), display significant interactions with the MUT-8 N-terminal domain in Martini3 simulations. The aromatic residues (Tyr and Phe) are enriched in the C-terminal of MUT-16 FFR (850-944aa). The observations indicate that the MUT-16 M8BR recruits MUT-8 majorly through cation-π or sp^2-π interactions between positively charged amino acids (Arg and Lys) of MUT-16 M8BR and Tyr residues of the MUT-8 N-terminal. Conversely, MUT-16 FFR recruits the MUT-8 N-terminal domain pri-

marily through the π-π interaction between aromatic residues (Tyr and Phe) of the MUT-16 FFR and Tyr of the MUT-8 N-terminal domain.

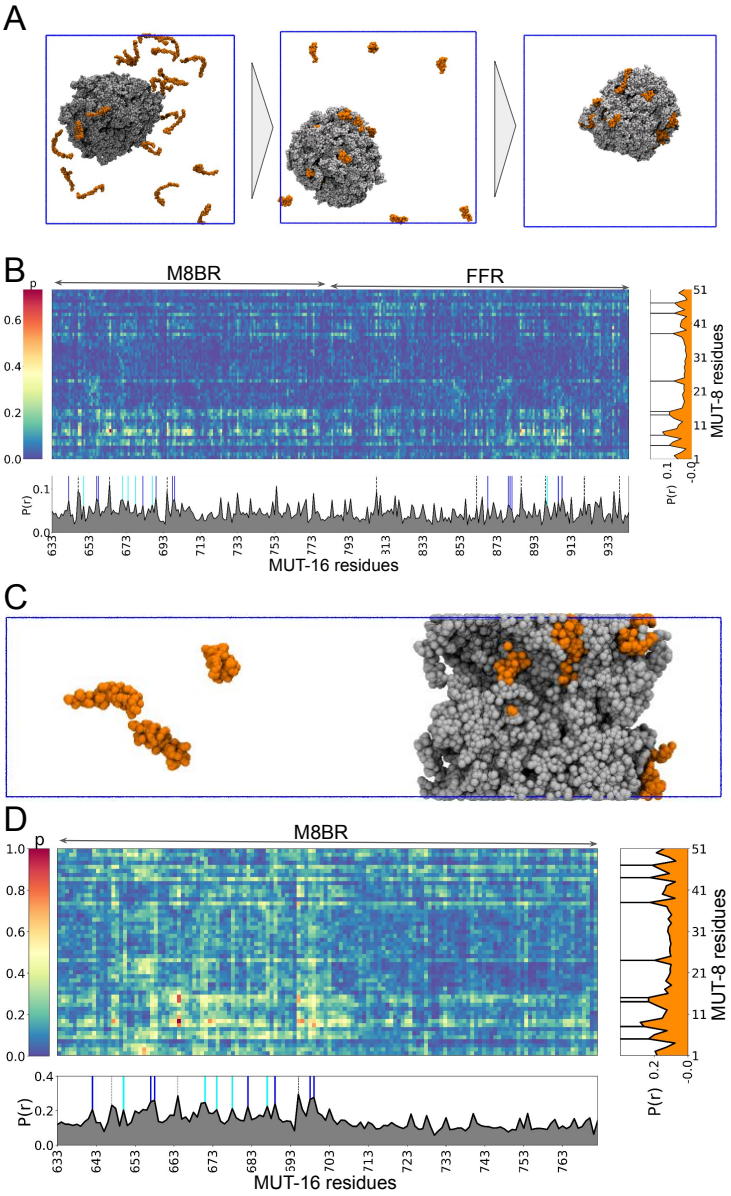

Figure 5: Recruitment of MUT-8 N-terminal PLD driven by phase-separated MUT-16. **A** Initial, intermediate, and final snapshots of a 20 µs near-atomic coarse-grained (Martini3) simulation trajectory of the interaction between MUT-8 N-terminal PLD (orange) and MUT-16 M8BR+FFR (grey) phase-separated condensate. **B** Inter-molecular contact probabilities of interaction between amino acids of MUT-8 N-terminal and MUT-16 M8BR+FFR are shown in a 2-D contact map. 1-D relative contact probability (P(r)) of the MUT-16 (grey) and MUT-8 (orange) amino acids is plotted to indicate the higher peaks by positively charged (Arg (blue), Lys (cyan)) and aromatic (Tyr (black), Phe (dashed)). **C** Intermediate snapshot of the slab simulation of MUT-8 N-terminal (orange) recruitment by phase-separated MUT-16 M8BR (grey). **D** Contact map of the interaction between MUT-8 N-terminal and MUT-16 (M8BR). Amino acids Arg (blue), Lys (cyan), Tyr (black), and Phe (dashed) are represented by vertical lines.

The comparison of the 1-D contact map of CALVADOS2 and Martini3 (Fig. S7D) revealed that in MUT-16 M8BR, the contact probability of the positively charged region is lower in the CALVADOS2 simulation compared to the Martini3 simulation. These observations indicate that the interactions of residues, which can engage in cation-π interactions, are more substantial and possibly overestimated in the Martini3 model than in the CALVADOS2 model. The Martini3 model is expected to capture better the influence of specific non-covalent interactions on protein-protein interactions since it can capture, e.g., side-chain interactions, as opposed to CALVADOS2, which captures interactions between disordered sequences based on their chemical compatibility. We note, however, that residue-level models have been validated much more extensively[39,41] and have been shown to capture sequence determinants of protein phase behaviour.[40]

Since experiments by Uebel et al.,[28] suggest a loss of MUT-8 recruitment upon deletion of the MUT-16 M8BR, we sought to further characterise the recruitment mechanism of MUT-8 by the MUT-16 M8BR by simulations of M8BR condensates and the N-terminus of MUT-8. Since we observed that the MUT-16 M8BR does not phase separate in CALVADOS2 simulations (Fig. 2A), we only conducted Martini3 simulations at $\lambda = 1$. At $\lambda = 1$, protein-protein interactions are overestimated,[48] enabling us to simulate the interactions of phase-separated MUT-16 M8BR with the MUT-8 N-terminal region (1-51aa). We note that this scaling can be understood as a change in conditions such as the solvent conditions and can be considered akin to performing a phase separation experiment with crowding agents or at a lower temperature.[37] To speed up the simulations, we also reduced the box size and are simulating the condensate in so-called slab geometry (15×15×60nm), where the condensate can also more easily equilibrate across two interfaces.[73] We observed phase separation of the MUT-16 M8BR and subsequent recruitment of the MUT-8 N-terminal domain (Fig. 5C). To compare the observations from our previous simulations of the MUT-8 N-terminal domain and MUT-16 (M8BR+FFR, Fig. 5A,B), we calculated contact probabilities (Fig. 5D). Interestingly, our contact plot indicated an essential role of the very N-terminal residues of

MUT-8 (1-15aa) and MUT-16 M8BR (640-700aa) in the recruitment mechanism. Subsequently, we calculated the one-dimensional contact map (Fig. 5D). This further validated the significant role of Tyr residues in the MUT-8 N-terminal domain recruitment process. Additionally, the 1-D contact map of MUT-16 M8BR further emphasises the critical involvement of positively charged amino acids like Arg and Lys. Interestingly, the positively charged amino acids are only present in the initial residues of MUT-16 M8BR (640-700aa), thereby enabling this region to recruit the MUT-8 N-terminal domain. Like the simulations with MUT-16 M8BR+FFR, these simulations indicate the potential involvement of cation-π or sp^2-π interaction between positively charged Arg and Lys of MUT-16 and aromatic Tyr of MUT-8 N-terminal domain in the recruitment process of MUT-8 to the *Mutator foci*. We calculated the cation-π between ARG/LYS of MUT-16 M8BR and TYR of MUT-8 N-terminal domain[35] and observed that Arg exhibits more robust cation-π interactions with Tyr compared to Lys residues (Fig. S8) in line with previous experiments.[9] However, it is essential to highlight that Martini3 simulations may not capture enough chemical detail to characterise the intricate balance of non-covalent interactions, including cation-π, sp^2-π and, e.g., hydrogen bonding interactions between Tyr and Arg/Lys and other competing interactions accurately enough.

Atomistic Simulation reveals the role of cation-π, sp^2-π, and hydrogen bond in recruitment of MUT-8 N-terminal PLD

To resolve, with atomistic detail, how Arg and Lys residues underpin the interaction of the MUT-8 N-terminal domain with MUT-16 M8BR, we conducted atomistic molecular dynamics simulations in an explicit solvent with the MUT-16 M8BR condensates bound to MUT-8 N-terminal fragments. Atomistic simulations can only partially capture the local structure and dynamics involved in the co-condensation of MUT-8 peptides with MUT-16. These simulations cannot encompass the entire association process, which is too slow to adequately sample using computationally intensive all-atom simulations. However, they can capture

molecular interactions more accurately and thus elucidate the molecular drivers[34,35,74] of peptide interaction with phase-separated condensate. The atomistic simulations (Fig. 6A) were started by backmapping the final frame of a 20 µs Martini3 coarse-grained slab simulation.[56] This backmapping enabled us to convert the coarse-grained representation of our system to an atomistically detailed system consisting of roughly 900,000 atoms, which we simulated for 1 µs. We observed that the contact map obtained from this 1 µs long trajectory is roughly analogous to the contact map of explicit solvent coarse-grained (Martini3) simulations (Fig. S9). Notably, the interaction frequency between the MUT-8 N-terminal domain and the MUT-16 M8BR increases in the atomistic simulations compared to Martini3 simulations (Fig. S9). Specifically, we have identified an elevated interaction frequency within the region spanning amino acid residues 640 to 680 and 730 to 772, predominantly facilitated by Arg, Lys, Gln, and His amino acids (Fig. S9). This aligns with the expectation that condensates do not change much on the timescale of 1 µs and that most contacts from the Martini3 simulation will be maintained. Lys687 makes fewer contacts in the atomistic contact map than the Martini3 (Fig. S9). Overall, the atomistic contact map is consistent with the essential role of Tyr residues within the MUT-8 N-terminal domain and Arg/Lys residues within the MUT-16 M8BR for MUT-8 recruitment.

To analyse the interactions in more detail we sought to create more extensive data set. We extracted 100 equidistant frames from this 1 µs trajectory. We launched 100 simulations each, utilising these frames as the initial structures in Folding@home,[43] for a total simulation duration of 350 µs.

With the atomistic simulations we assessed the relative importance of cation-π or sp^2-π interactions in Arg-Tyr and Lys-Tyr interactions (Fig. 6B). We tracked the dynamics of cation-π and sp^2-π interactions between MUT-8 and MUT-16 throughout the 1 µs in-house trajectory and 350 µs simulations on Folding@Home.[43] We calculated the abundance of cation-π interactions between Arg/Lys and Tyr[35] (see Methods). Interestingly, Arg (MUT-16) contributes to more abundant cation-π interactions with Tyr (MUT-8) compared to

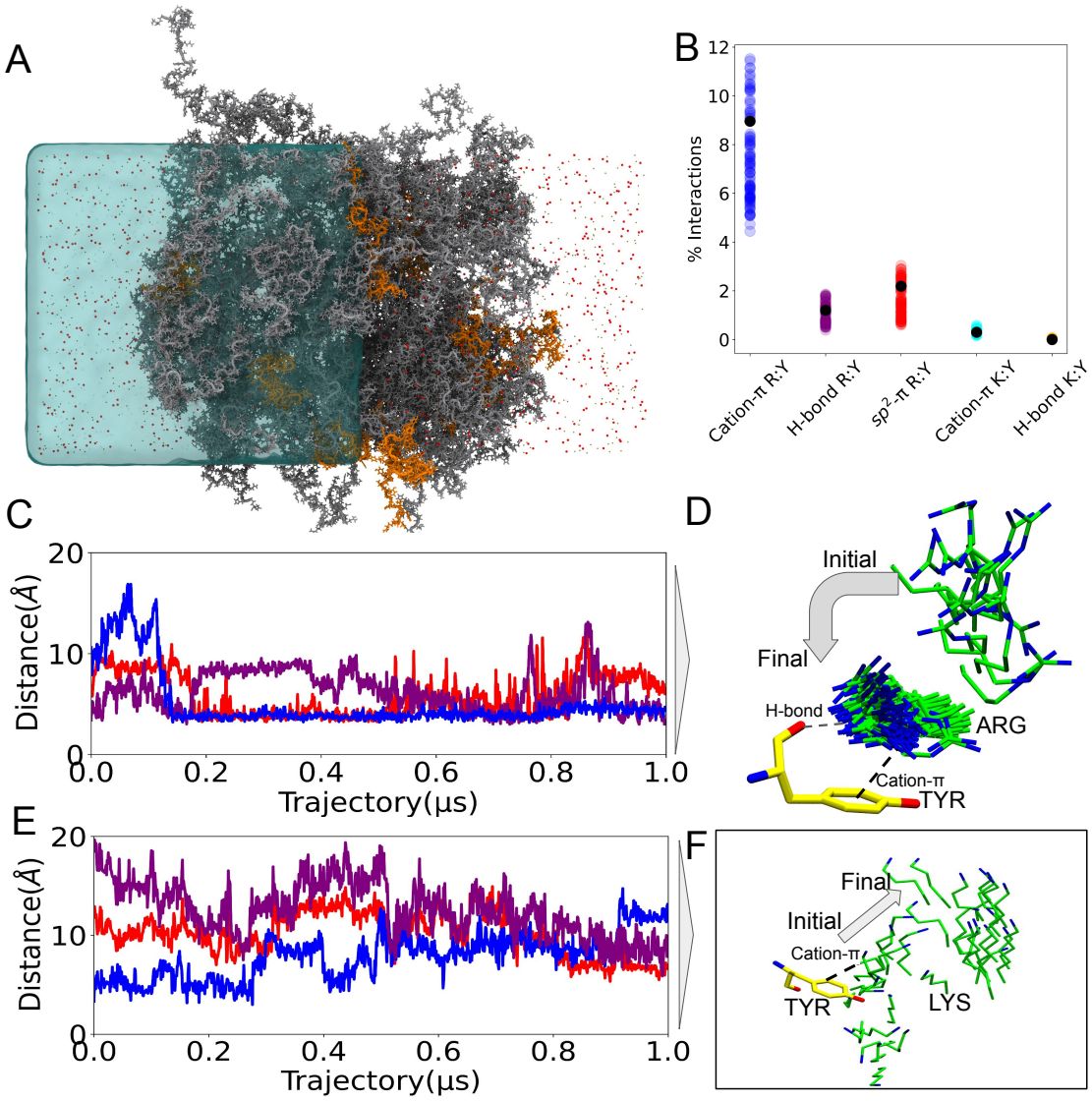

Figure 6: **A** Atomistic simulation depicting 100 chains of MUT-16 M8BR (grey) and the N-terminal domain of MUT-8 (orange). Sodium ions (Na+) and chloride ions (Cl-) are shown as green and red points, respectively, suspended in water (blue). For clarity, the water on the right half of the simulation box is deleted. **B** Comparison of the percentage of cation-π, sp^2-π and H-bond interactions made by Arg/Lys (MUT-16 M8BR) and Tyr (MUT-8 N-terminal). The percentage of H-bonds was calculated given the pair was already forming the cation-π interaction. Multiple data points were calculated from the 100 atomistic runs obtained from Folding@home (coloured, translucent) and in-house 1 µs trajectory (black, opaque). **C** Time evolution of the distance between the aromatic ring of Tyr and guanidinium group of Arg for selected pairs of residues. **D** The pictorial representation of one of the residue pairs whose time evolution (blue) is plotted in **C**. **E** Time evolution of the distance between the aromatic ring of Tyr and the amino group of Lys for selected amino acid pairs. **F** The pictorial representation of one of the residue pairs whose time evolution (blue) is plotted in **E**.

Lys(MUT-16) (Fig. 6B). We also saw that the Arg residues form many sp^2-π interactions (Fig. 6B). Further, to identify if the Arg/Lyr and Tyr form transient or persistent interactions, we calculated the distance between the guanidinium group of Arg (or the ammonium group of Lys) and the centre-of-mass of the benzene ring in Tyr for several pairs (Fig. 6C,E). Fig. 6C indicates the presence of persistent (blue line) and transient (red line, purple line) pairs of Arg and Tyr. The persistent pair depicted by the blue line in Fig. 6C and visually represented in Fig. 6D establishes contact within 150 ns and maintains it throughout the remaining trajectory. The pair tracked with the red line engages in close contact after about 200 ns of simulations, with a Tyr-Arg distance close to 5 Å similar to the persistent pair. This close contact appears stable for approximately 200 ns before the distance fluctuates more strongly, shifting to larger values. For the pair tracked by the purple line, close contacts are observed only for short periods, shorter than 100 ns. These observations align with the recent study, which demonstrated the overall rapid exchange (nanosecond scale) of contacts in a condensate.[74] Some longer-lived contacts could also be important in condensates and the recognition dynamics. Visualisation of simulations illustrated that the guanidinium group of Arg may be able to form cation-π and sp^2-π interaction with the aromatic ring due to the presence of sp^2 hybridised positively charged nitrogens and planar geometry. Arg can also form a hydrogen bond with the backbone carbonyl of Tyr (Fig. 6D). Furthermore, we also observed the coexistence of cation-π, sp^2-π and hydrogen bond in Arg and Tyr pairs (Fig. S10), with the three interactions predominately coinciding for exemplary Arg-Tyr pair, except for short transients, e.g., shortly after 0.4 µs when a sp^2-π interaction is detected but no cation-π interaction or hydrogen bond. Lys and Tyr pairs appeared less stable in the atomistic simulations (Fig. 6E). The Lys and Tyr pair indicated by the blue line in Fig. 6E and pictorially shown in Fig. 6F initially remained in contact for approximately 250 ns but eventually dissociates for the remainder of the trajectory. By contrast, other Lys-Tyr pairs start to form close contacts towards the end of the 1 µs trajectory (red line), but remaining further apart than the pair tracked by the blue line, which nonetheless shows that Lys-Tyr

contacts can form in the simulations even though they are less favoured compared to Arg-Tyr contacts. The purple line shows an example of the Lys-Tyr pair, which mostly stays separate and only starts to get closer towards the end of the trajectory. Unlike for Arg and Tyr, no side hydrogen bonds between the side chain of Lys and the Tyr backbone are found (Fig. 6B), and thus, this additional interaction cannot stabilise already formed cation-π interactions. While Arg can in principle form simultaneous hydrogen bonding, cation-π and sp^2-π interaction with Tyr, Lys can only form cation-π. This "triple" interaction could contribute to Arg's more robust interactions with Tyr than Lys. This finding is in line with previous quantum chemical calculations[14,75,76] as well as by a recent study involving an artificial IDP,[16] where atomistic molecular dynamics simulations of condensates also resolved interactions involving hydrogen bonds with the backbone.

Pull-down experiments validate the importance of Arg residues in M8BR of MUT-16 for MUT-8 recuitment

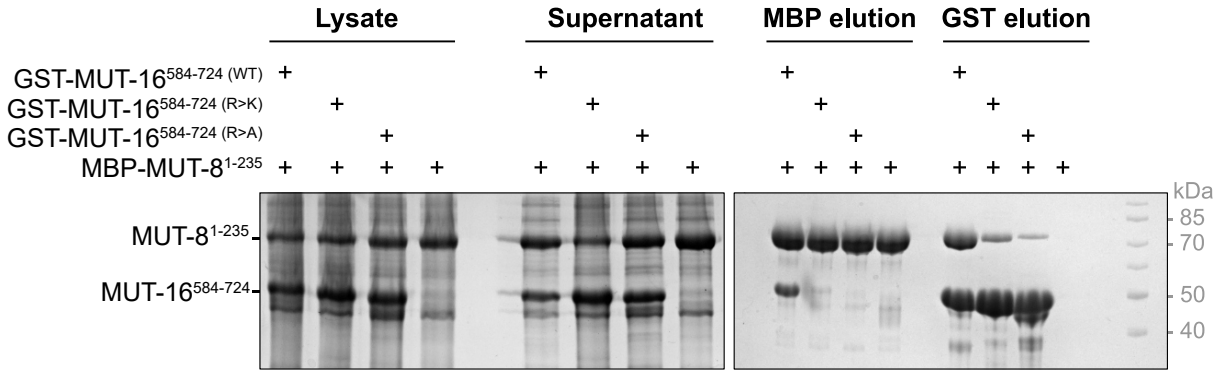

Figure 7: Contribution of Arg residues of MUT-16 M8BR in MUT-8 recruitment. Analysis of the interaction between an MBP-tagged MUT-8 N-terminal fragment (1-235) and a GST-tagged MUT-16 fragment (584-724) by pulldown assays. Three different MUT-16 constructs are tested: WT, a construct where seven Arg residues (642, 657, 658, 682, 689, 698, 699) were mutated to Lys (R > K), and a construct where the identical seven Arg residues were mutated to Ala (R > A). MUT-16 constructs were co-expressed with MUT-8 in *E. coli*. Both MBP and GST pulldowns were performed. SDS-PAGE, followed by Coomassie staining, analysed the total Lysate, supernatant, and elution.

We performed co-expression pulldown experiments to investigate the importance of the Arg/Tyr interaction for the binding between MUT-16 and MUT-8, as predicted by the simulation approaches. When seven MUT-16 Arg (642, 657, 658, 682, 689, 698, 699) were mutated to Lysines (R > K), binding between MUT-16 and MUT-8 was reduced compared to pull-downs with wt MUT-16. Mutating these seven Arg to Ala (R > A) weakened the interaction between MUT-16 and MUT-8 further (Fig. 7). The same trend is seen whether an MBP-tagged MUT-8 N-terminal fragment or a GST-tagged Mut-16 fragment was used for the pulldown. The experiments indicate the essential role of Arg residues of the MUT-16 M8BR in the recruitment of MUT-8, as highlighted by the simulations.

Conclusions

We elucidated how the MUT-16 phase separates to form *Mutator foci* and recruits MUT-8 N-terminal PLD. MUT-16 is the scaffolding protein of *Mutator foci* in *C. elegans* germ cells, which facilitates RNA silencing. Our simulations and *in vitro* experiments demonstrate that MUT-16 condensates form via phase separation, thereby extending the *in vivo* observations of MUT-16 foci reversibly dissolving at elevated temperatures. This evidence strengthens the conclusion by Uebel et al. that MUT-16 phase separation is fundamental to the function of *Mutator foci*. At the same time, we show that MUT-16 FFR, enriched in aromatic residues, is essential for MUT-16 phase separation.

MUT-16 recruits many cognate binding partners to *Mutator foci*, including MUT-8, which further recruits the exoribonuclease MUT-7. Previous studies have shown that the N-terminus of MUT-8 interact with MUT-16,[30] while its C-terminus recruits MUT-7.[31] However, the molecular mechanism by which MUT-16 specifically recruits the N-terminus of MUT-8 remains elusive. Interestingly, the MUT-8 N-terminus features an IDR (1-51 aa), which is predicted to be a prion-like domain enriched in residues that are frequently associated with protein phase separation. Our coarse-grained and atomistic simulations suggest

that Tyr residues within the MUT-8 N-terminal domain predominantly interact with Arg and, to a lesser extent, Lys residues in MUT-16 M8BR. Further, the atomistic simulations on the μs time-scale emphasise the role of cation-π and sp^2-π interactions in the interaction of MUT-16 M8BR with MUT-8 N-terminal domain. In particular, the atomistic simulations emphasise the relative stability of Arg-Tyr interactions over Lys-Tyr interactions, attributed in part to the ability of Arg, unlike Lys, to additionally stabilise the cation-π or sp^2-π interactions by hydrogen bonding with the backbone of Tyr. The *in vitro* experiments involving mutations of seven Arg in MUT-16 M8BR to Lys/Ala demonstrate the reduction in MUT-8 recruitment, indicating the pivotal role of Arg residues within MUT-16 M8BR in the recruitment of MUT-8 N-terminal domain.

Our study used a multi-scale approach by combining and comparing the coarse-grained simulations with different resolutions (residue-level CALVADOS2 and near-atomic Martini3) and atomistic simulations. The comparison establishes that the CALVADOS2 simulations effectively capture disordered regions' relative phase separation propensities. This is demonstrated by the observation that MUT-16 M8BR does not phase separate spontaneously. In contrast, MUT -16 M8BR+FFR and MUT-16 FFR phase-separates at 275 K, aligning with findings from *in vitro* phase separation assays. The chemically more detailed Martini3 simulations demonstrate the phase separation propensity of MUT-16 M8BR, MUT-16 FFR and MUT-16 M8BR+FFR. However, Martini3 simulations are known to overemphasise the protein-protein interactions, and a rebalancing of protein-protein[37,64] or protein-water[48,65] interactions can demonstrate the phase-behaviour.[48] Remarkably, applying such a rebalancing to protein-water interactions,[48] the simulations reproduce the experimental trends in the relative phase separation propensities of the MUT-16 IDRs. The chemically detailed representation offered by Martini3 simulations also lets us elucidate the molecular drivers of phase separation, including the specific involvement of amino acids. Despite this advantage, Martini3 simulations remain limited in their representation of amino acid side chains compared to the atomistic models. Nonetheless, utilising Martini3 simulations for back-mapping to

atomistic resolution facilitates investigating the atomic-level interactions between MUT-16 M8BR and MUT-8 N-terminal domain.

This ability to go back and forth between resolutions will become essential to tackle increasingly large and complex systems with substantial chemical detail and sufficiently accurate description of molecular interactions.[16] In the context of the *Mutator foci*, the same "molecular grammar" and non-covalent interactions like cation-π and sp^2-π[9,35] that regulates the dynamic interactions of disordered proteins appears to govern phase separation of MUT-16 and recruitment of the client MUT-8.

Acknowledgement

This project was funded by SFB 1551 Project No. 464588647 of the DFG (Deutsche Forschungsgemeinschaft). L.S.S. acknowledges support by ReALity (Resilience, Adaptation and Longevity) and Forschungsinitiative des Landes Rheinland-Pfalz. A.C.S and L.S.S. thank M^3ODEL for support. S.F. received funding from the Austrian Science Fund (FWF) programs I6110-B and the doc.funds DOC 177-B: RNA@core: "Molecular mechanisms in RNA biology". V.B. received funding from the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Curie Skłodowska Grant Agreement Nr. 847548. (Vienna International PostDoc Program (VIP-2). We gratefully acknowledge the advisory services offered and the computing time granted on the supercomputers Mogon II at Johannes Gutenberg University Mainz, which is a member of the AHRP (Alliance for High-Performance Computing in Rhineland Palatinate) and the Gauss Alliance e.V. K.G. thanks Yashraj Wani and Rodrique Badr for insightful discussions. L.S.S. Thanks, Dr R. Sprangers and Dr J. Mittal, for inspiring discussions.

References

(1) Hyman, A. A.; Weber, C. A.; Jülicher, F. Liquid-liquid phase separation in biology. *Annual review of cell and developmental biology* **2014**, *30*, 39–58.

(2) Li, P.; Banjade, S.; Cheng, H.-C.; Kim, S.; Chen, B.; Guo, L.; Llaguno, M.; Hollingsworth, J. V.; King, D. S.; Banani, S. F., et al. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **2012**, *483*, 336–340.

(3) Brangwynne, C. P.; Tompa, P.; Pappu, R. V. Polymer physics of intracellular phase transitions. *Nature Physics* **2015**, *11*, 899–904.

(4) Gomes, E.; Shorter, J. The molecular language of membraneless organelles. *Journal of Biological Chemistry* **2019**, *294*, 7115–7127.

(5) Martin, E. W.; Mittag, T. Relationship of sequence and phase separation in protein low-complexity regions. *Biochemistry* **2018**, *57*, 2478–2487.

(6) Vernon, R. M.; Chong, P. A.; Tsang, B.; Kim, T. H.; Bah, A.; Farber, P.; Lin, H.; Forman-Kay, J. D. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *elife* **2018**, *7*, e31486.

(7) Dignon, G. L.; Best, R. B.; Mittal, J. Biomolecular phase separation: from molecular driving forces to macroscopic properties. *Annual review of physical chemistry* **2020**, *71*, 53–75.

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

(9) Wang, J.; Choi, J.-M.; Holehouse, A. S.; Lee, H. O.; Zhang, X.; Jahnel, M.; Maharana, S.; Lemaitre, R.; Pozniakovsky, A.; Drechsel, D., et al. A molecular grammar

governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **2018**, *174*, 688–699.

(10) Sherrill, C. D. Energy component analysis of π interactions. *Accounts of chemical research* **2013**, *46*, 1020–1028.

(11) Reddy, A. S.; Sastry, G. N. Cation [M= H+, Li+, Na+, K+, Ca2+, Mg2+, NH4+, and NMe4+] interactions with the aromatic motifs of naturally occurring amino acids: a theoretical study. *The Journal of Physical Chemistry A* **2005**, *109*, 8893–8903.

(12) Gokel, G. W.; De Wall, S. L.; Meadows, E. S. Experimental evidence for alkali metal cation- π interactions. *European journal of organic chemistry* **2000**, *2000*, 2967–2978.

(13) Gokel, G. W.; Barbour, L. J.; Stephen, L.; Meadows, E. S. Macrocyclic polyethers as probes to assess and understand alkali metal cation-π interactions. *Coordination Chemistry Reviews* **2001**, *222*, 127–154.

(14) Gallivan, J. P.; Dougherty, D. A. Cation-π interactions in structural biology. *Proceedings of the National Academy of Sciences* **1999**, *96*, 9459–9464.

(15) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nature reviews Molecular cell biology* **2017**, *18*, 285–298.

(16) Rekhi, S.; Garcia, C. G.; Barai, M.; Rizuan, A.; Schuster, B. S.; Kiick, K. L.; Mittal, J. Expanding the molecular language of protein liquid–liquid phase separation. *Nature Chemistry* **2024**, 1–12.

(17) Pritišanac, I.; Alderson, T. R.; Kolarić, D̈.; Zarin, T.; Xie, S.; Lu, A. X.; Alam, A.; Maqsood, A.; Youn, J.-Y.; Forman-Kay, J. D., et al. A Functional Map of the Human Intrinsically Disordered Proteome. *bioRxiv* **2024**, 2024–03.

(18) Claycomb, J. M. Ancient endo-siRNA pathways reveal new tricks. *Current Biology* **2014**, *24*, R703–R715.

(19) Ketting, R. F. The many faces of RNAi. *Developmental cell* **2011**, *20*, 148–161.

(20) Hutvagner, G.; Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nature reviews Molecular cell biology* **2008**, *9*, 22–32.

(21) Bernstein, E.; Caudy, A. A.; Hammond, S. M.; Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **2001**, *409*, 363–366.

(22) Ketting, R. F.; Fischer, S. E.; Bernstein, E.; Sijen, T.; Hannon, G. J.; Plasterk, R. H. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes & development* **2001**, *15*, 2654–2659.

(23) Pak, J.; Fire, A. Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science* **2007**, *315*, 241–244.

(24) Sijen, T.; Steiner, F. A.; Thijssen, K. L.; Plasterk, R. H. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* **2007**, *315*, 244–247.

(25) Lee, S. R.; Collins, K. Physical and functional coupling of RNA-dependent RNA polymerase and Dicer in the biogenesis of endogenous siRNAs. *Nature structural & molecular biology* **2007**, *14*, 604–610.

(26) Colmenares, S. U.; Buker, S. M.; Buhler, M.; Dlakić, M.; Moazed, D. Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. *Molecular cell* **2007**, *27*, 449–461.

(27) Phillips, C. M.; Montgomery, T. A.; Breen, P. C.; Ruvkun, G. MUT-16 promotes formation of perinuclear mutator foci required for RNA silencing in the C. elegans germline. *Genes & development* **2012**, *26*, 1433–1444.

(28) Uebel, C. J.; Anderson, D. C.; Mandarino, L. M.; Manage, K. I.; Aynaszyan, S.; Phillips, C. M. Distinct regions of the intrinsically disordered protein MUT-16 mediate assembly of a small RNA amplification complex and promote phase separation of Mutator foci. *PLoS genetics* **2018**, *14*, e1007542.

(29) Phillips, C. M.; Montgomery, B. E.; Breen, P. C.; Roovers, E. F.; Rim, Y.-S.; Ohsumi, T. K.; Newman, M. A.; van Wolfswinkel, J. C.; Ketting, R. F.; Ruvkun, G., et al. MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Current Biology* **2014**, *24*, 839–844.

(30) Tops, B. B. J.; Tabara, H.; Sijen, T.; Simmer, F.; Mello, C. C.; Plasterk, R. H. A.; Ketting, R. F. RDE-2 interacts with MUT-7 to mediate RNA interference in Caenorhabditis elegans. *Nucleic Acids Research* **2005**, *33*, 347–355.

(31) Busetto, V.; Pshanichnaya, L.; Lichtenberger, R.; Hann, S.; Ketting, R. F.; Falk, S. MUT-7 exoribonuclease activity and localisation are mediated by an ancient domain. *bioRxiv* **2023**, 2023–12.

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

(33) Mukherjee, S.; Schäfer, L. V. Thermodynamic forces from protein and water govern condensate formation of an intrinsically disordered protein domain. *Nature Communications* **2023**, *14*, 5892.

(34) Gruijs da Silva, L. A.; Simonetti, F.; Hutten, S.; Riemenschneider, H.; Sternburg, E. L.; Pietrek, L. M.; Gebel, J.; Dötsch, V.; Edbauer, D.; Hummer, G., et al. Disease-linked TDP-43 hyperphosphorylation suppresses TDP-43 condensation and aggregation. *The EMBO journal* **2022**, *41*, e108443.

(35) Zheng, W.; Dignon, G. L.; Jovic, N.; Xu, X.; Regy, R. M.; Fawzi, N. L.; Kim, Y. C.; Best, R. B.; Mittal, J. Molecular details of protein condensates probed by microsecond long atomistic simulations. *The Journal of Physical Chemistry B* **2020**, *124*, 11671–11679.

(36) Sarthak, K.; Winogradoff, D.; Ge, Y.; Myong, S.; Aksimentiev, A. Benchmarking molecular dynamics force fields for all-atom simulations of biological condensates. *Journal of Chemical Theory and Computation* **2023**, *19*, 3721–3740.

(37) Benayad, Z.; von Bülow, S.; Stelzl, L. S.; Hummer, G. Simulation of FUS Protein Condensates with an Adapted Coarse-Grained Model. *J. Chem. Theory Comput.* **2021**, *17*, 525–537.

(38) Souza, P. C. T. et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature Methods* **2021**, *18*, 382–388.

(39) Tesei, G.; Schulze, T. K.; Crehuet, R.; Lindorff-Larsen, K. Accurate model of liquid–liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. *Proceedings of the National Academy of Sciences* **2021**, *118*, e2111696118.

(40) Tesei, G.; Lindorff-Larsen, K. Improved predictions of phase behaviour of intrinsically disordered proteins by tuning the interaction range [version 2; peer review: 2 approved]. *Open Research Europe* **2023**, *2*.

(41) Joseph, J. A.; Reinhardt, A.; Aguirre, A.; Chew, P. Y.; Russell, K. O.; Espinosa, J. R.; Garaizar, A.; Collepardo-Guevara, R. Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nature Computational Science* **2021**, *1*, 732–743.

(42) Tesei, G.; Trolle, A. I.; Jonsson, N.; Betz, J.; Knudsen, F. E.; Pesce, F.; Johans-

son, K. E.; Lindorff-Larsen, K. Conformational ensembles of the human intrinsically disordered proteome. *Nature* **2024**, *626*, 897–904.

(43) Shirts, M.; Pande, V. S. Screen savers of the world unite! *Science* **2000**, *290*, 1903–1904.

(44) Tsanai, M.; Frederix, P. W. J. M.; Schroer, C. F. E.; Souza, P. C. T.; Marrink, S. J. Coacervate formation studied by explicit solvent coarse-grain molecular dynamics with the Martini model. *Chem. Sci.* **2021**, *12*, 8521–8530.

(45) Anderson, J. A.; Glaser, J.; Glotzer, S. C. HOOMD-blue: A Python Package for High-Performance Molecular Dynamics and Hard Particle Monte Carlo Simulations. *Comput. Mater. Sci.* **2020**, *173*, 109363.

(46) Ashbaugh, H. S.; Hatch, H. W. Natively unfolded protein stability as a coil-to-globule transition in charge/hydropathy space. *Journal of the American Chemical Society* **2008**, *130*, 9536–9542.

(47) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism From Laptops to Supercomputers. *SoftwareX* **2015**, *1-2*, 19 – 25.

(48) Thomasen, F. E.; Pesce, F.; Roesgaard, M. A.; Tesei, G.; Lindorff-Larsen, K. Improving Martini 3 for Disordered and Multidomain Proteins. *Journal of Chemical Theory and Computation* **2022**, *18*, 2033–2041.

(49) Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589.

(50) Kroon, P. C.; Grünewald, F.; Barnoud, J.; van Tilburg, M.; Souza, P. C.; Wassenaar, T. A.; Marrink, S.-J. Martinize2 and vermouth: Unified framework for topology generation. *arXiv preprint arXiv:2212.01191* **2022**,

(51) Wassenaar, T. A.; Ingólfsson, H. I.; Bockmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations. *Journal of chemical theory and computation* **2015**, *11*, 2144–2155.

(52) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of chemical physics* **2007**, *126*.

(53) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics* **1981**, *52*, 7182–7190.

(54) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.

(55) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Dotson, D. L.; Domański, J.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. Proceedings of the 15th Python in Science Conference. 2016; pp 98–105.

(56) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going backward: a flexible geometric approach to reverse transformation from coarse grained to atomistic models. *Journal of chemical theory and computation* **2014**, *10*, 676–690.

(57) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712–725.

(58) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.

(59) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950–1958.

(60) Best, R. B.; De Sancho, D.; Mittal, J. Residue-Specific α-Helix Propensities From Molecular Simulation. *Biophys. J.* **2012**, *102*, 1462–1467.

(61) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. *J. Phys. Chem. B* **2015**, *119*, 5113–5123.

(62) Flores-Solis, D.; Lushpinskaia, I. P.; Polyansky, A. A.; Changiarath, A.; Boehning, M.; Mirkovic, M.; Walshe, J.; Pietrek, L. M.; Cramer, P.; Stelzl, L. S., et al. Driving forces behind phase separation of the carboxy-terminal domain of RNA polymerase II. *Nature Communications* **2023**, *14*, 5979.

(63) Zerze, G. H. Optimizing the martini 3 force field reveals the effects of the intricate balance between protein–water interaction strength and salt concentration on biomolecular condensate formation. *Journal of Chemical Theory and Computation* **2023**, *20*, 1646–1655.

(64) Thomasen, F. E.; Skaalum, T.; Kumar, A.; Srinivasan, S.; Vanni, S.; Lindorff-Larsen, K. Recalibration of protein interactions in Martini 3. *bioRxiv* **2023**, 2023–05.

(65) Larsen, A. H.; Wang, Y.; Bottaro, S.; Grudinin, S.; Arleth, L.; Lindorff-Larsen, K. Combining molecular dynamics simulations with small-angle X-ray and neutron scattering data to study multi-domain proteins in solution. *PLoS computational biology* **2020**, *16*, e1007870.

(66) Lancaster, A. K.; Nutter-Upham, A.; Lindquist, S.; King, O. D. PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* **2014**, *30*, 2501–2502.

(67) King, O. D.; Gitler, A. D.; Shorter, J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain research* **2012**, *1462*, 61–80.

(68) Gitler, A. D.; Shorter, J. RNA-binding proteins with prion-like domains in ALS and FTLD-U. *Prion* **2011**, *5*, 179–187.

(69) Babinchak, W. M.; Haider, R.; Dumm, B. K.; Sarkar, P.; Surewicz, K.; Choi, J.-K.; Surewicz, W. K. The role of liquid–liquid phase separation in aggregation of the TDP-43 low-complexity domain. *Journal of Biological Chemistry* **2019**, *294*, 6306–6317.

(70) Udan-Johns, M.; Bengoechea, R.; Bell, S.; Shao, J.; Diamond, M. I.; True, H. L.; Weihl, C. C.; Baloh, R. H. Prion-like nuclear aggregation of TDP-43 during heat shock is regulated by HSP40/70 chaperones. *Human molecular genetics* **2014**, *23*, 157–170.

(71) Qamar, S.; Wang, G.; Randle, S. J.; Ruggeri, F. S.; Varela, J. A.; Lin, J. Q.; Phillips, E. C.; Miyashita, A.; Williams, D.; Ströhl, F., et al. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interactions. *Cell* **2018**, *173*, 720–734.

(72) Holehouse, A. S.; Ginell, G. M.; Griffith, D.; Boke, E. Clustering of aromatic residues in prion-like domains can tune the formation, state, and organization of biomolecular condensates: published as part of the biochemistry virtual special issue "protein condensates". *Biochemistry* **2021**, *60*, 3566–3581.

(73) Dignon, G. L.; Zheng, W.; Mittal, J. Simulation methods for liquid–liquid phase separation of disordered proteins. *Current opinion in chemical engineering* **2019**, *23*, 92–98.

(74) Galvanetto, N.; Ivanović, M. T.; Chowdhury, A.; Sottini, A.; Nüesch, M. F.; Nettels, D.; Best, R. B.; Schuler, B. Extreme dynamics in a biomolecular condensate. *Nature* **2023**, *619*, 876–883.

(75) Dougherty, D. A. Cation-π interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* **1996**, *271*, 163–168.

(76) Kumar, K.; Woo, S. M.; Siu, T.; Cortopassi, W. A.; Duarte, F.; Paton, R. S. Cation–π interactions in protein–ligand binding: Theory and data-mining reveal different roles for lysine and arginine. *Chemical science* **2018**, *9*, 2655–2665.