Multiscale simulations of molecular recognition by phase separated MUT-16: A scaffolding protein of *Mutator foci*

Kumar Gaurav[1,2,3], Virginia Busetto[4,5], Diego Javier Paez Moscoso[2], Arya Changiarath[3], Sonya M. Hanson[6], Sebastian Falk[4,5], Rene F. Ketting[2], Lukas S. Stelzl[1,2,3*]

1 Institute of Molecular Physiology, Johannes Gutenberg University Mainz, Germany
2 Institute of Molecular Biology (IMB), Mainz
3 KOMET1, Institute of Physics, Johannes Gutenberg University Mainz, Germany
4 Max Perutz Lab, Vienna Biocenter Campus (VBC), Austria
5 University of Vienna, Center for Molecular Biology, Department of Structural and Computational Biology, Vienna, Austria
6 Flatiron Institute, Center for Computational Biology & Center for Computational Mathematics, New York City, USA

* lstelzl@uni-mainz.de

Abstract

Biomolecular recruitment by phase separated condensates has emerged as a key organising principle of biological processes. One such process is the RNA silencing pathway, which regulates gene expression and genomic defense 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. Coarse-grained simulations at residue-level and near atomic-resolution also 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 condensate 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 to condensates from coarse-grained and 350 µs all-atom simulations in explicit solvent 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 favorable 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 formation 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. These properties allow IDRs and IDPs to form distributed, multivalent, and transient interactions [1]. Further, it enables IDRs and IDPs to undergo phase separation in which they coalesce to form biomolecular condensates [1, 2]. An unresolved question in the field of 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 recognize 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 hybridized 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 [15], but an understanding of the molecular scale of how such interactions underpin molecular recognition remains elusive [16].

Molecular recognition by disordered domains of proteins is critical for RNA silencing, also recognized 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 primeval defense 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 [17, 18]. The principal mechanism of this pathway is the interaction of small (∼18-30 nt) RNA with the proteins of the Argonaute family which enables the regulation of complementary mRNAs at the levels of transcription, translation, and stability [18, 19]. There are three types of small RNAs: microRNAs (miRNAs), piwi-interacting RNAs (piRNAs), and short interfering RNA (siRNA) that are synthesized 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 [20, 21]. The siRNAs produced through this mechanism are called primary siRNAs.

In *C. elegans*, RNA silencing involves a siRNA amplification step. This amplification produces secondary siRNAs complementary to the target mRNAs, and is facilitated by RNA-dependent RNA polymerases (RdRp) [22–25]. In *C. elegans* germ cells siRNA amplification guided by RdRp is choreographed in perinuclear membrane-less organelles called *Mutator foci* [26]. *Mutator foci* are juxtaposed to P granules that accommodate the proteins responsible for RNA silencing and mRNA degradation (Fig. 1A) [26]. Biogenesis of *Mutator foci* is underpinned by phase separation of a protein named MUT-16 [26, 27].

MUT-16 contains a substantial intrinsically disordered region (IDR) and functions as 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* and 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) [26, 28]. MUT-16 requires MUT-8 as a linker to recruit MUT-7 (Fig. 1B) [27]. Further, it was shown that MUT-16 contains a MUT-8 binding region (RBR, 633-772aa) and that the N-terminus of MUT-8 may interact with MUT-16 [27, 29]. Finally, the C-terminal domain of MUT-8 was shown to interact with the C-terminal domain of MUT-7 [29, 30]. Interestingly, a prion-like domain is present in the MUT-8 N-terminus (Fig. S1B). Prion-like 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 RBR 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 behavior and molecular recognition by phase separated condensates at the molecular scale [31–33]. Atomistic molecular dynamics [34] can resolve interactions of individual amino acids in phase separated condensates of proteins [33, 35] albeit at large computational cost. In these simulations the dynamics of each atom in the system is tracked and interactions with water and ions are considered explicitly. In order to simulate the phase separation, it is necessary to simplify the protein representations, allowing the simulations to capture pertinent time and length scales [31, 36]. One method of simplifying protein representations involves grouping atoms into distinct representative beads, a process known as coarse graining. The Martini3 model [37] 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 important drivers for phase separation, but has not been validated in detail. Further, conducting coarse grained simulations using CALVADOS2 model [38], with one bead representing each amino acid, facilitates the exploration of protein phase separation behaviors [39], enabling simulations at the proteomic scale. [40].

In this study, we elcudidate how 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 behavior 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 stabilizing MUT-16 condensates. Furthermore, by initiating numerous brief atomistic trajectories of MUT-16 condensates in complex with MUT-8 fragments using Folding@home, [41] 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.

Materials and methods

We performed molecular dynamics simulations in three different resolutions: the residue-level coarse-grained model (CALVADOS2), the near-atomic coarse-grained model (Martini3), and the atomistic model. The combination of the three models allows us to establish a linear connection between the dynamics and the atomistic interaction profile of phase separation and biomolecular recognition. Further, we performed *in vitro* experiments to further investigate and validate the MUT-16 phase separation and the recruitment of the MUT-8 N-terminal domain.

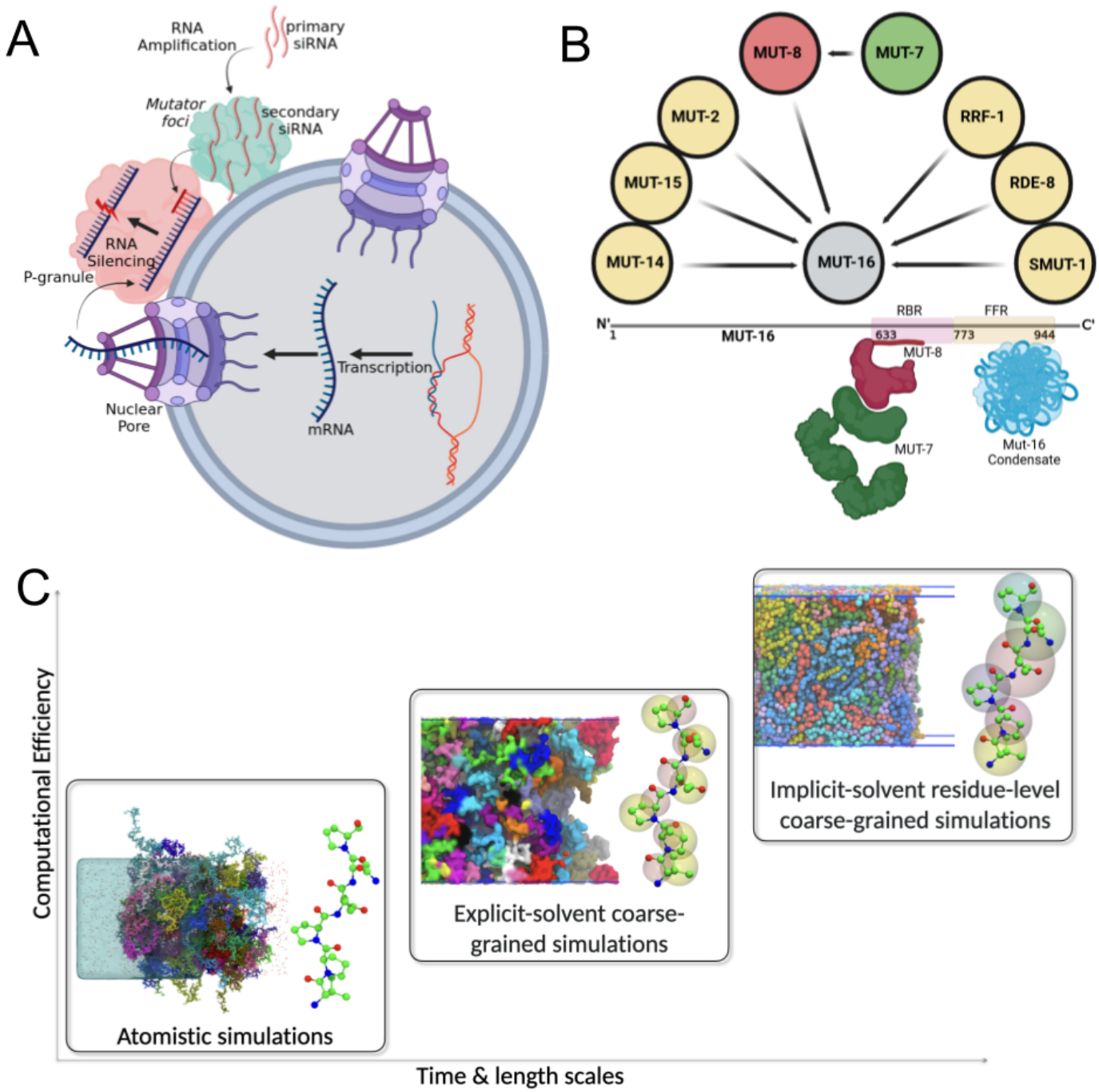

Fig 1. A. RNA silencing mechanism in *C. elegans* which constitutes RNA amplification in *Mutator foci*. **B.** MUT-16 scaffolding protein (grey) recruits other proteins (yellow) associated with the *Mutator foci*, including MUT-8 (red) which further recruits MUT-7 (green). The MUT-8 binding region (RBR, pink) and Foci forming region (FFR, yellow) are illustrated based on their position on MUT-16 protein. The intrinsically disordered N-terminal domain of MUT-8 (red) is shown to be recruited by MUT-16 RBR. MUT-8 C-terminal domain further recruits MUT-7 (green) C-terminal domain. Chains of MUT-16 FFR protein (blue) are shown to form phase separated condensate. **C.** Multi-scale simulation strategy to identify essential drivers of phase behavior and molecular recognition of 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.

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) [42] extended with azplugins (v. 0.10.2). The CALVADOS2 model is an improvement of the previously established HPS model [31], proposed by Tesei et al. [38, 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 both bonded and non-bonded components, incorporating electrostatic and short-range pairwise interactions. The bonded interactions are described by a harmonic potential,

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

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 [43],

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

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 $r_c^{\text{LJ}} = 2.0$ 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 optimized by Tesei et. al., with a Bayesian method which makes use of a comprehensive experimental data set [38, 40]. The non-bonded interaction also includes an electrostatic component modeled via Debye-Hückel potential:

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

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). The simulations were initialized by placing the protein chains randomly in the slab. The slab geometry allows the equilibration of semi-infinite condensate in two dimensions (along the x and y-axis) and provides a phase separation interface along the z-axis. The temperature was regulated by the Langevin thermostat. Further, the equations of motions were integrated by the velocity Verlet algorithm with a timestep (Δt) of 10 fs. The simulations were performed for 10 µs.

MUT-16 constructs purification

Three different MUT-16 constructs were purified to test their condensate-formation properties in vitro: MUT-16 633-772 (RBR), MUT-16 773-944 (FFR), MUT-16 633-944 (RBR+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 immobilized 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 dialyzed 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 (RBR), MUT-16 773-944 aa (FFR), and MUT-16 633-944 aa (RBR+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. For droplet formation, a 1:1 serial dilution was performed in 8-strip PCR tubes (Multiply-µstrip Pro 8-strip, Sarstedt, REF 72.991.002). For each construct, the reaction was done in two parallel serial dilutions. The first one, without the addition of 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 cleaved off the N-terminal MBP tag from the constructs, and it was used to induce 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 analyzed, 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 three times 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 analyzed by SDS–PAGE and Coomassie staining.

Phase diagram

The residue-level coarse-grained (CALVADOS2) [38] 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. To determine the densities of dense and dilute phases we followed the approach by Tesei et al [38]: (1) we positioned the dense phase at the center of the slab. (2) Clustered the proteins based on the distance between the center-of-mass (COM) of the proteins. The protein chains within 5 nm of COM were considered to be in one cluster. (3) Post clustering, the protein concentration of dilute (ρ_{l}) and dense (ρ_{h}) phase 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:

$$(\rho_{\mathrm{h}} + \rho_{\mathrm{l}})/2 = \rho_{\mathrm{c}} + C(T_{\mathrm{c}} - T) \tag{6}$$

where C is a positive fitting parameter.

Near-atomic coarse-grained simulations with explicit solvent with the Martini3 model

The near-atomic coarse-grained (Martini3) simulations were executed using GROMACS [44], employing either the latest version of the Martini force field (Martini3) [37, 45], or an improved variant of Martini3 featuring rescaled protein-water interactions [46]. In the Martini3 force-field framework, multiple atoms are grouped together into a single particle, often representing four to five heavy atoms with a single coarse-grain bead. Water and ions are represented explicitly. The initial structure of the protein IDR chains was obtained by truncating the desired region from the AlphaFold [47] structures of MUT-16 (AF-O62011-F1) and MUT-8 (AF-Q58AA4-F1). The protein chains were coarse-grained using the martinize2 python script. The coarse-grained protein chains were inserted in a cubical box or a slab geometry and further solvated

with the Insane.py Python script. Furthermore, 0.15 M salt(NaCl) was added to the system on top of balancing the charge in the system. The simulation box was subjected to minimization using the steepest descent algorithm. Simulations were performed with periodic boundary conditions. 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 [48], and the pressure was maintained at 1 bar using Parrinello-Rehman barostat [49]. All the simulations were performed for 20 µs with 20 fs time step. Last 10 µs of the trajectory was used for the analysis.

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 using MDAnalysis. 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 $2^{1/6}\sigma_{ij}$ [31], 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`. 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 normalizing 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 [50]. For computational efficiency, we reduced the slab dimensions from $15 \times 15 \times 60$nm to $15 \times 15 \times 30$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 [51–54] with TIP4P-D water [55] in GROMACS [44]. The simulation box contained 100 MUT-16 RBR and 10 MUT-8 N-terminal chains. After backmapping, the atomistic system was minimized 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 [48] and Parrinello-Rahman barostat [49]. The production run was performed for 1 µs in the NPT ensemble in in-house supercomputer Mogon-II and Further, 100 parallel simulations totaling 350 µs in Folding@home [41] 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

The calculation of cation-π interactions between Arg/Lys and Tyr residues were calculated based on the recent literature [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 center of mass 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 center 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 centers 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.

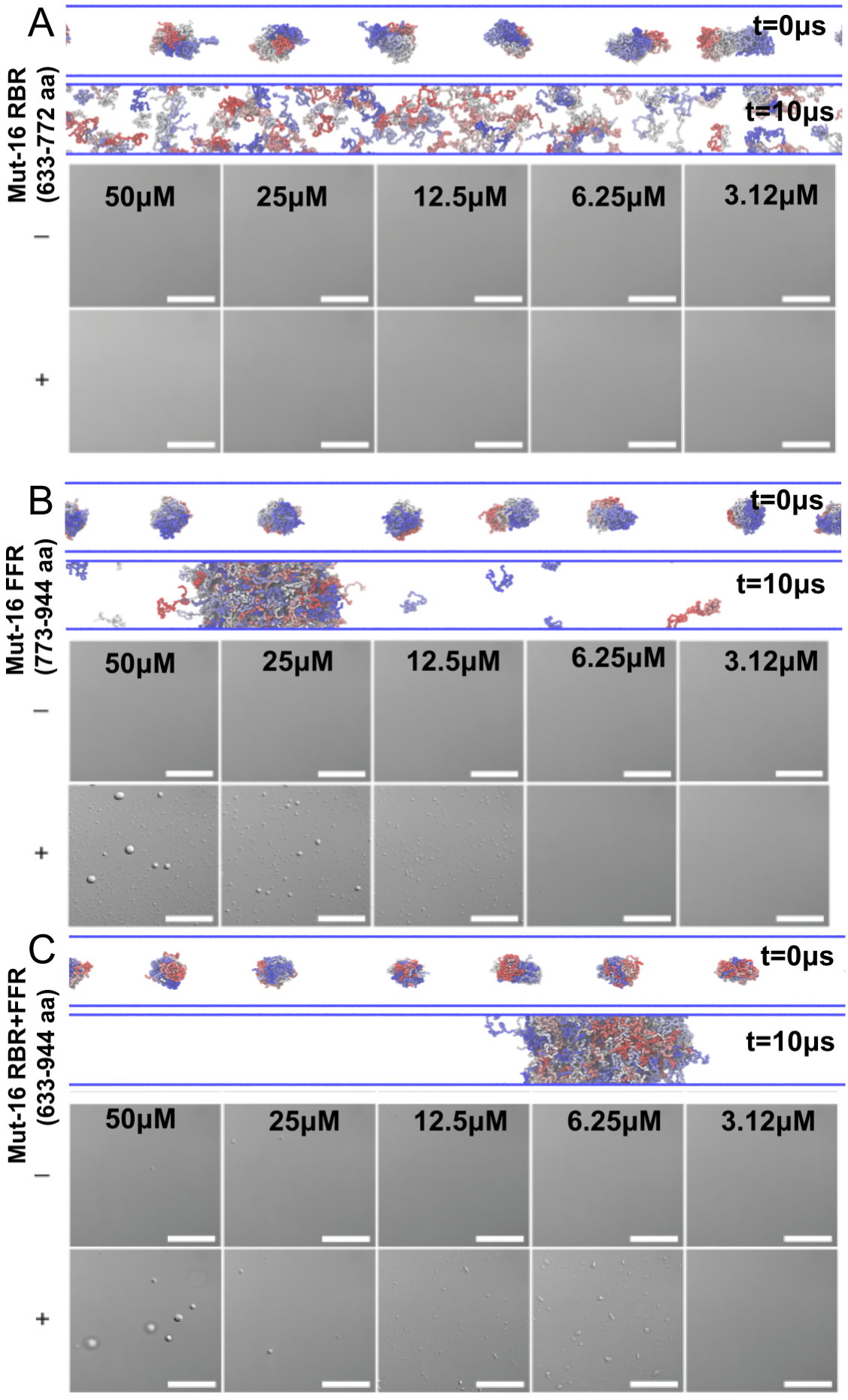

Fig 2. MUT-16 FFR is sufficient to drive the phase separation. **A.** Panel above shows the initial (t=0 µs) and final (t=10 µs) frame of MUT-16 RBR simulation in a slab geometry. In the panel below, representative bright-field images show the phase separation assays of the MUT-16 RBR (633–772 aa) construct 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 with the addition of 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 RBR+FFR construct obtained through the method described in **A**. Scale bar correspond to 20 µm.

Results and Discussion

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

We studied the phase separation of MUT-16 experimentally and through molecular dynamics simulations. The simulations were performed using the residue-level coarse-grained (CALVADOS2) model [38, 40], which captured the phase separation, and temperature dependence and highlighted the factors determining the phase behavior. Previous studies have observed that MUT-16 clusters into foci in *in vivo* at ambient temperature and nucleates the formation of *Mutator foci* [26,27]. Uebel et al., [27] observed that the C-terminal region of MUT-16 (773-1050 aa) is essential for its phase separation. Furthermore, an intermediate region spanning amino acid 633 to 772, designated as RBR, 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 945 to 1054 from the C-terminal domain, due to the tendency of this region to adopt a three-dimensional structure (Fig. S1A). Thus, we first simulated a construct ranging from amino acids 773 to 944, designated as FFR. Subsequently, as a methodological control and analytical reference, we included the RBR region to the FFR. Ultimately, the MUT-16 simulation and experimental protocol was focused on the regions RBR (633-772 aa), FFR (773-944 aa) and RBR+FFR (633-944) of MUT-16 (Fig. 2). Simulations indicate that chains of RBR do not phase separate spontaneously (Fig. 2A). However FFR and RBR+FFR chains phase separate (Fig. 2B,C), with dilute phase concentration measured at around 0.28 ± 0.042 mM and 0.15 ± 0.057 mM, respectively. Experimental observation also indicated the lack of phase separation of MUT-16 RBR 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 RBR+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 RBR+FFR, the size of droplets also increases upon increasing the concentration, but unlike MUT-16 FFR, the number of droplets decreases (Fig. 2C). Experimentally, the differences between the FFR and RBR+FFR compared to RBR seem even larger than in simulations, considering that FRR and RBR+FFR phase separate at even lower concentrations than in simulations.

Further we performed simulations of MUT-16 RBR+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., [27], we expect that MUT-16 RBR+FFR will cluster together possibly via phase separation at lower temperatures but not at high temperatures. Indeed, at lower temperatures, the MUT-16 RBR+FFR protein chains have the propensity to phase separate and forming a dense protein rich phase surrounded by a dilute phase depleted in proteins (Fig. 3A). As the temperature increases, MUT-16 RBR+FFR chains are progressively observed in the dilute phase (Fig. 3A). This is in qualitative agreement with experiments by Uebel et al., [27] where MUT-16 foci disappeared as the worms were exposed to elevated temperature (∼303 K). Upon lowering the ambient temperature (∼294 K), the condensates reappeared. 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 [31, 38, 39] with *in vivo* behavior, on the one hand, encourages further analysis of drivers of phase separation and also 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 RBR+FFR residues and observed that the FFR exhibits a stronger propensity to interact (Fig. 3C) compared to the RBR. While residues throughout the sequences engage in interactions in the inter-molecular contact map (Fig. 3C), residues from around 663 to 713 in RBR 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:RBR and RBR:RBR (Fig. S2A,B). The simulations indicate an essential role of FFR in driving the phase separation and validate the experimental observations by Uebel et al., [27]. Although the RBR 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 aromatic amino acids are relatively plentiful in the FFR (14%) compared to the RBR (5%), possibly enabling the FFR to drive the phase separation. Additionally, the 1-D contact map (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 somewhat lower tendency to interact. A stretch of Lys in RBR (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 RBR. This observation is in line with the observations by Uebel et. al.. The 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 [27]. We computed the relative frequency of the interaction of amino acid pairs normalized (Fig. S3A) and unnormalized (Fig. S3B). The normalized 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. S3A). Tyr and Phe generally form many secondary interactions with Leu, Ile, and Met. Arg-Asp also forms frequent contacts. Looking at the unormalized interactions provides a better view of what interactions are actually driving phase separation as many weak interactions can be equally or more important than a more limited number of relatively strong interactions [56]. The unnormalized plot indicates that Pro-Pro, Glu-Glu, Pro-Glu, Pro-Tyr [57], Glu-Tyr and Asn-Tyr are the most abundant interactions (Fig. S3B). The unnormalized plot reflects the role of the abundance of Pro in MUT-16 RBR+FFR phase separation as these proline based interactions are not prominent in the normalized plot. The relative abundance of interactions involving Glu, in particular Glu-Glu interactions, likely stem from the fact that Glu residues are frequently flanked by positively charged residues in MUT-16 RBR+FFR (See tableS2). While some pairs are more important, it is clear that other residues also interact [56]. Overall, our simulations rationalized 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 characterize the molecular drivers of MUT-16 phase behavior we sought to employ the near-atomic resolution Martini3 coarse-grained model [37]. Martini3 is a powerful general-purpose coarse-grained model of biomolecules, which has also been used to simulate multi-domain and disordered proteins [45, 46, 58, 59] . Unlike CALVADOS2, Martini3 has not been developed or extensively tested for the phase separation of disordered proteins. 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 (RBR+FFR) IDR chains (Fig. 4A). The MUT-16 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 [46, 60] by the rescaling factor λ. We reasoned that if the propensity of MUT-16 RBR+FFR phase separation, captured by Martini3 results from favorable interactions in the simulations, then increasing protein-water interactions somewhat should not abrogate condensation. We observed that at lower λ values of 1.01 and 1.03, the MUT-16 RBR+FFR condensates persisted (Fig. S4A,B). At larger λ values of 1.06 and 1.1 the condensates dissolved(Fig. S4C,D). The density of the MUT-16 RBR+FFR chains with respect to the distance from the center of the box also indicated the dissolution of condensate upon increasing the λ value ((Fig. S4E). Rescaling the relative strength of protein-water interactions compared to protein-protein interactions can be considered as analogous to increasing the effective temperature [36]. Thus, the simulations with the Martini3 model are also in qualitative agreement with the dissolution of the condensates at higher temperatures in our simulations with the CALVADOS2 model and the temperature dependence Uebel et. al., observed *in vivo*.

Further, we characterized the regions of MUT-16, which form many contacts and might stabilize the condensate. We computed the average pairwise contacts between the residues in each chain per frame, represented in a 2-D(Fig. 4B). The contact plot suggests that 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., [27]. 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 RBR and FFR in orchestrating the phase separation of MUT-16. The contact map reveals that the FFR:FFR interactions are not signficantly more frequent than RBR:RBR interactions(Fig. S2C,D). This observation made from Martini3 simulations is in contrast to the CALVADOS2 simulations in which the RBR:RBR interactions are significantly weaker than FFR:FFR interaction (Fig. S2A,B). 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 RBR (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 RBR N-terminal region (640-700aa) in Martini3 simulations compared to the CALVADOS2 (Fig. S4F). Since the sequence of MUT-16 RBR is populated with the positively charged amino acids

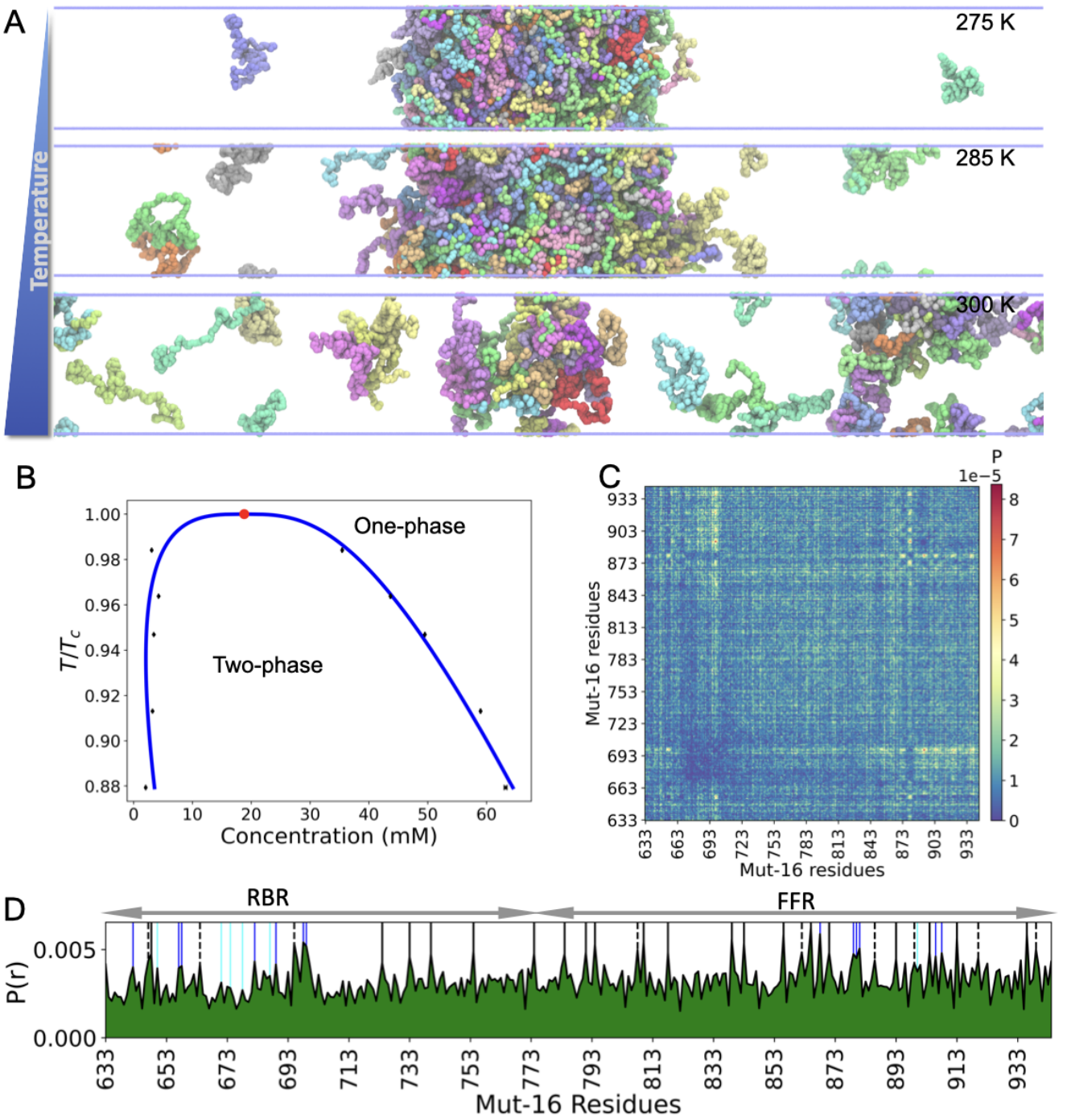

Fig 3. Phase behavior of MUT-16 condensate in residue-level coarse-grained (CALVADOS2) simulations. **A.** Simulations of 100 MUT-16 RBR+FFR chains at 275 K, 285 K, and 300 K. The color indicates the different chains. **B.** Phase diagram of MUT-16 obtained by CALVADOS2 simulation at different temperatures. 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 interchain contact-map obtained from the simulation of MUT-16 RBR+FFR chains at 275 K. **D.** Relative probability P(r) of an amino acid to form a contact in the MUT-16 (RBR+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).

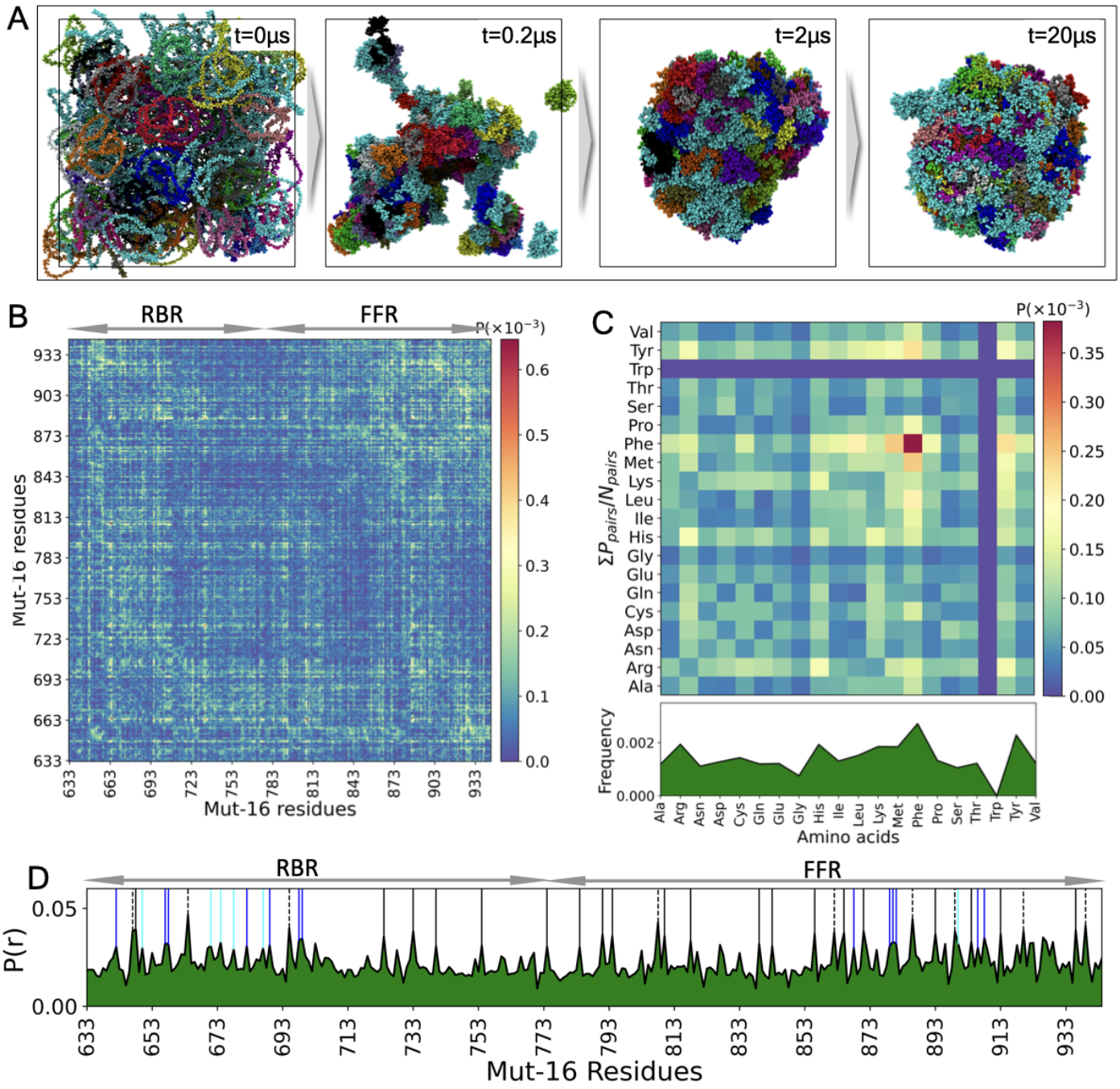

Fig 4. MUT-16 phase separation in explicit-solvent coarse-grained (Martini3) simulations. **A.** Time evolution of MUT-16 RBR+FFR phase separation. **B.** Inter-molecular contact probabilities in the condensed phase of MUT-16 RBR+FFR as a function of the amino acid sequence. **C.** Intermolecular contact probability as a function of residue type normalized 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.

(Arg and Lys) in the N-terminal domain (640-700aa) and aromatic amino acid in the C-terminal domain (700-772aa), we hypothesize that RBR:RBR interaction could be facilitated by cation-π or sp^2-π interactions. We posit that cation-π or sp^2-π interactions are either overestimated in Martini3 model resulting in the difference in the RBR:RBR 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. S4F), 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 (RBR+FFR). Nonetheless, the relative propensities for residues to engage in inter-chain interactions in the Martini3 simulations is similar to the propensities from simulations with the CALVADOS2 model [40] (Fig. S4F).

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., [27] suggest that a particular region of MUT-16 (633-772aa, RBR) is responsible for recruiting MUT-8. MUT-8 was not observed to localize to the phase separated MUT-16 condensate upon the deletion of the RBR region [27]. However the localization of MUT-8 has not been determined upon deletions of MUT-16 C-terminal in previous studies. Recent literature has elucidated the role of MUT-8 C-terminal domain in further recruitment of the MUT-7 exoribonuclease [30], but the specific role of its N-terminal domain remains elusive [29]. Interestingly, we found that the first 51aa of MUT-8 is predicted as a prion-like domain (PLD) (Fig. S1) [61]. PLDs are known to be enriched in uncharged polar amino acids (Tyr, Asn, Gln) and Gly [62] and are present in proteins with propensity of phase separation like FUS [63] and TDP-43 [64, 65]. Aromatic amino acids of PLDs are known to drive phase separation and recruitment to existing biomolecular condensate [8, 9, 66, 67]. 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. S5A) model and chemically more detailed near-atomic resolution coarse-grained Martini3 (Fig. 4A) model to elucidate the interaction network of MUT-16 RBR+FFR and MUT-8 N-terminal domain(Uniprot ID: Q58AA4). In simulations with both models, the MUT-16 RBR+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 RBR+FFR without MUT-8 (Fig. 4A). By analyzing both CALVADOS2 (Fig. S5C) 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 is majorly facilitated by the MUT-16 FFR in CALVADOS2 simulations (Fig. S5A). (ii) In the case of Martini3 simulations, the recruitment is driven by the MUT-16 RBR (Fig. S5B). Both simulations indicate that the recruitment of MUT-8 could be facilitated by both the RBR and FFR of MUT-16 (Fig. 5B,S5D,E). Furthermore, we observed that the positively charged amino acids (Arg and Lys) and the aromatic amino acids (Tyr and Phe) in MUT-16 RBR+FFR facilitate the recruitment of MUT-8 N-terminal domain (Fig. S5D). The positively charged residues (Arg and Lys), segregated in a particular region of MUT-16 RBR (640-700aa), display significant interactions with the MUT-8 N-terminal domain in Martini3 simulations. The aromatic residues (Tyr and Phe) are majorly segregated in the C-terminal of MUT-16 FFR (850-944aa). The observations indicate that the MUT-16 RBR recruits MUT-8 majorly through cation-π or sp^2-π interactions between positively charged amino acids (Arg and Lys) of MUT-16 RBR and Tyr residues of the MUT-8 N-terminal. Conversely, MUT-16 FFR recruits the MUT-8 N-terminal domain primarily 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. The comparison of the 1-D contact map of CALVADOS2 and Martini3 (Fig. S5D) revealed that in MUT-16 RBR, 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 stronger and possibly overestimated in the Martini3 model compared to the CALVADOS2 model. The Martini3 model is expected to better capture 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 [38] and have been shown to capture sequence determinants of protein phase behavior.

Since experiments by Uebel et al., [27] suggest a loss of MUT-8 recruitment upon deletion of the MUT-16 RBR, we investigated to establish the recruitment mechanism of MUT-8 by the MUT-16 RBR. Since we observed that the MUT-16 RBR 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 [46], enabling us to simulate the interactions of phase separated MUT-16 RBR 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 [36]. To accelerate the sampling of the interactions, we also reduced the box size and are simulating the condensate in so-called slab geometry ($15 \times 15 \times 60$nm), where the condensate can also more easily

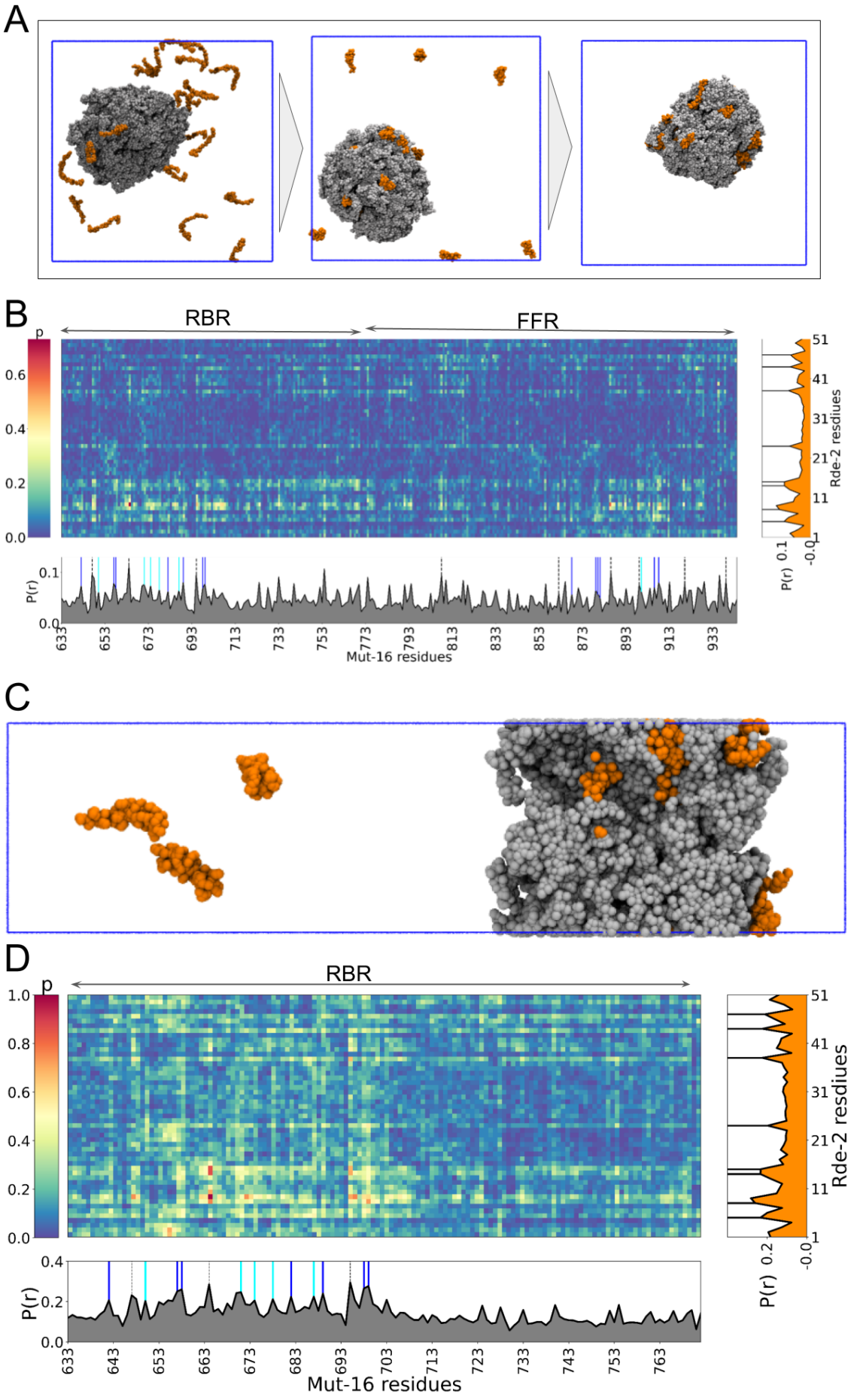

Fig 5. Recruitment of MUT-8 N-terminal PLD driven by phase separated MUT-16. A) Initial, intermediate, and final snapshots of a 20 µs near-atomic coarse-grained (Martini3) simulation trajectory of the interaction between MUT-8 N-terminal PLD (orange) and MUT-16 RBR+FFR (grey) phase separated condensate. B) Intermolecular contact probabilities of interaction between amino acids of MUT-8 N-terminal and MUT-16 RBR+FFR shown in 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 RBR (grey). D) Contact-map of the interaction between MUT-8 N-terminal and MUT-16 (RBR), Amino acids Arg (blue), Lys (cyan), Tyr (black), and Phe (dashed) are represented by vertical lines.

equilibrate across two interfaces [68]. We observed phase separation of the MUT-16 RBR and subsequent recruitment of the MUT-8 N-terminal domain (Fig. 5C). To compare the observations from our previous simulations of MUT-8 N-terminal domain and MUT-16 (RBR+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 RBR (640-700aa) in the recruitment mechanism. Subsequently, we calculated the one-dimensional contact map (Fig. 5D). This further validated the Tyr residues' significant role in the MUT-8 N-terminal domain recruitment process. Additionally, the 1-D contact map of MUT-16 RBR further emphasizes 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 RBR (640-700aa), thereby allowing this region to recruit MUT-8 N-terminal domain efficiently. Our 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 RBR 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. S6). However, it is important to highlight that Martini3 simulations may not posses the requisite fidelity to capture the necessary chemical details needed to characterize 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 RBR, we conducted atomistic molecular dynamics simulations in explicit solvent with the MUT-16 RBR condensates bound to MUT-8 N-terminal fragments. The atomistic simulations cannot capture the entire association process of MUT-8 peptides to MUT-16 as the process is far too slow to be sampled in computationally expensive all-atom simulations. However, they can capture molecular interactions more accurately and can thus elucidate the molecular drivers [33,35] 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 [50]. This backmapping allowed us to convert the coarse-grained representation of our system to an atomistically detailed system consisting of roughly 900,000 atoms. We ran one 1 µs trajectory in in-house supercomputer MOGON-II. Subsequently, we extracted 100 equidistant frames from this 1 µs trajectory and launched 100 simulations each, utilizing these frames as initial structure in Folding@home [41], for a total simulation duration of 350 µs. We observed that the contact map obtained from the 1 µs long in-house atomistic simulations is roughly analogous to the contact map of explicit solvent coarse-grained (Martini3) simulations (Fig. S7). Notably, the interaction frequency between the MUT-8 N-terminal domain and the MUT-16 RBR increases in the atomistic simulations compared to Martini3 simulations (Fig. S7). 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. S7). 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 on this timescale. Notably, Lys687 makes fewer contacts in the atomistic contact map than the Martini3 contact map (Fig. S7). 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 RBR for MUT-8 recruitment.

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. 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 Lys(MUT-16) (Fig. 6B). We also validated that the Arg significantly form sp^2-π interactions (Fig. 6B). Further, To identify if the ARG/LYS and TYR forms transient or persistent interactions, we calculated the distance between the guanidinium group of Arg (or the ammonium group of Lys) and the center-of-mass

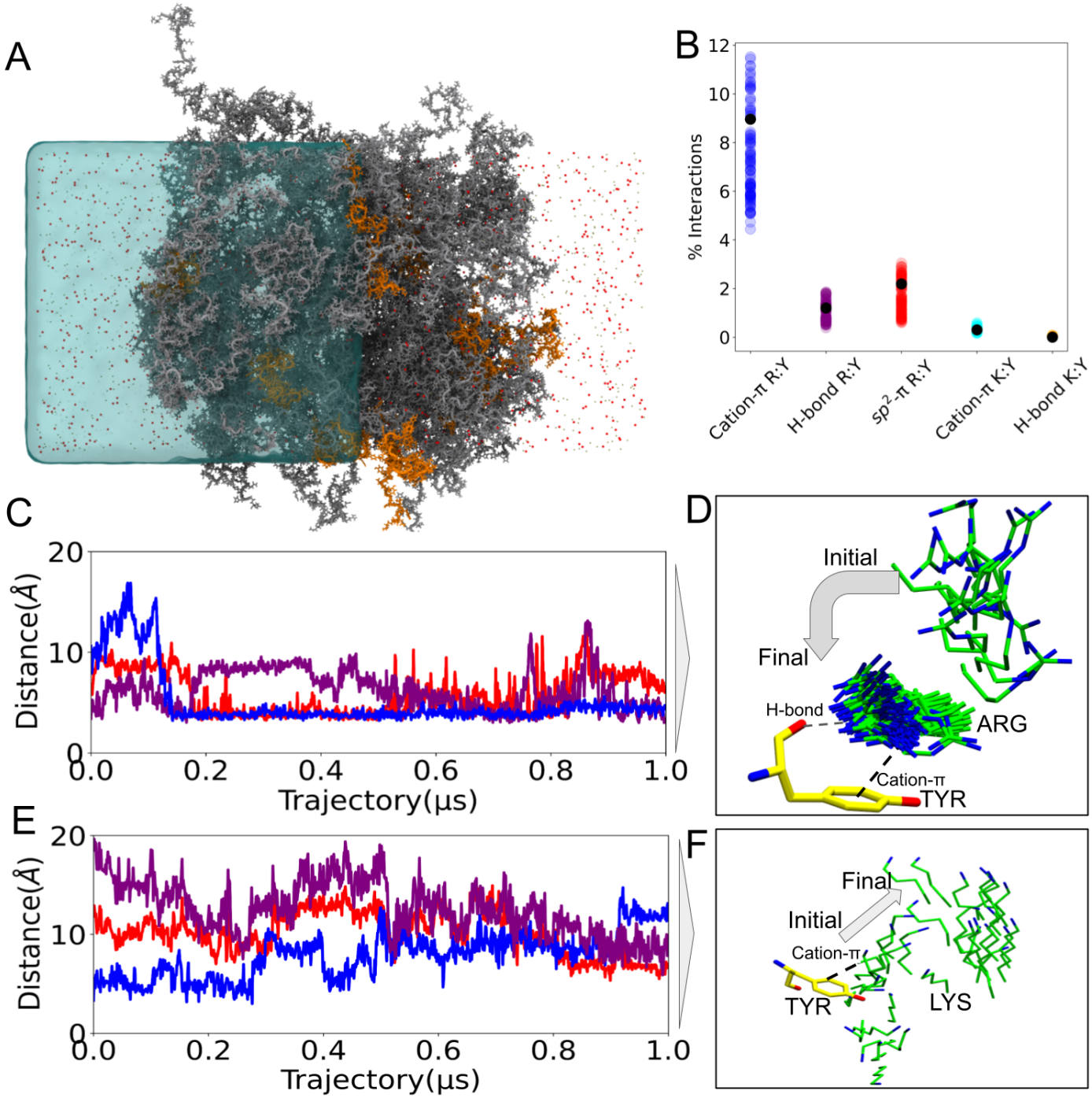

Fig 6. A) Atomistic simulation of 100 chains of MUT-16 RBR (grey) and MUT-8 N-terminal (orange). B) Comparison of the percentage of cation-π, sp^2-π and H-bond interactions made by Arg/Lys (MUT-16 RBR) and Tyr (MUT-8 N-terminal). The percentage of H-bonds was calculated given the pair was already forming the cation-π interaction. Multiple data points were calculated from the 100 atomistic runs obtained from Folding@home (colored, translucent) and initial 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.

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. Visualization of simulations illustrated that the guanidinium group of Arg may exhibit the ability to form the cation-π and/or sp^2-π interaction with the aromatic ring due to the presence of sp^2 hybridized 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. S8) suggesting the tridentate interactions. 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. The percentage of hydrogen bond between Arg side-chain and Tyr backbone, given that the pair is already forming cation-π interactions, is higher in coparison to Lys and Tyr pair (Fig. 6B). The tridentate interaction of Arg and Tyr could be attributed to the distinct chemistry of the guanidinium group. This trivalency of Arg, notably absent in Lys, could allow it to form more robust interactions with Tyr compared to Lys. This finding is in line with previous quantum chemical calculations [14, 69, 70] as well as by a recent study involving an artificial IDP [56], where interactions involving hydrogen bonds with the backbone were also resolved.

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

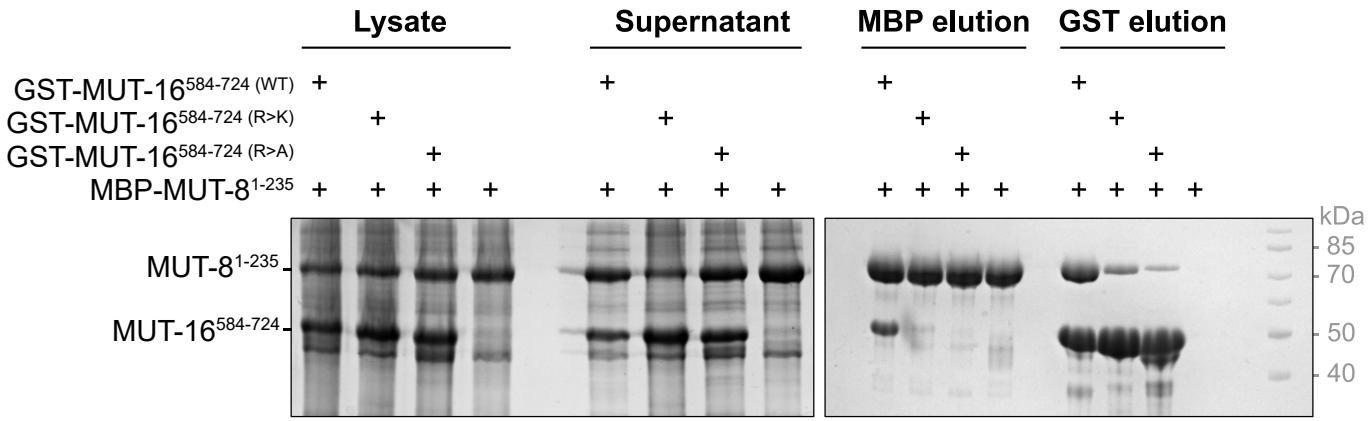

Fig 7. Contribution of Arg residues of MUT-16 RBR in MUT-8 recruitment. Analysis of the interaction between an MBP-tagged MUT-8 N-terminal fragment (1-235) and a GST-tagged MUT-16 fragment (584-724) by pulldown assays. Three different MUT-16 constructs are tested: WT, a construct where seven Arg residues (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 analyzed 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), weakend the interaction between MUT-16 and MUT-8 further (Fig. 7). The same trend is seen whether a 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 RBR in the recruitment of MUT-8, as highlighted by the simulations.

Conclusions

We elucidated how 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 the RNA silencing. Our simulations and *in vitro* experiments demonstrate that MUT-16 FFR, enriched in aromatic residues, is essential for MUT-16 phase separation. Further, the simulations also validated the temperature dependence of MUT-16 phase separation, previously observed *in vivo*.

MUT-16 is known to recruit many different 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 [29],

while its C-terminus recruits MUT-7 [30]. However, the molecular mechanism by which MUT-16 specifically recruits the N-terminus of MUT-8 remains elusive. Interestingly, 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 MUT-8 N-terminal domain predominantly interact with Arg and, to a lesser extent, Lys residues in MUT-16 RBR. Further, the atomistic simulations on the μs time-scale emphasize the role of cation-π and sp^2-π interactions in the interaction of MUT-16 RBR with MUT-8 N-terminal domain. In particular, the atomistic simulations emphasize the relative stability of Arg-Tyr interactions over Lys-Tyr interactions, attributed in part to the ability of Arg, unlike Lys, to additionally stabilize 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 RBR to Lys/Ala demonstrate the reduction in MUT-8 recruitment, indicating the pivotal role of Arg residues within MUT-16 RBR in the recruitment of MUT-8 N-terminal domain.

For our study we 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. Notably, the comparison establishes that the CALVADOS2 simulations effectively capture the relative phase separation propensities of disordered regions. This is demonstated by the observation that MUT-16 RBR does not phase separate spontaneously, whereas MUT-16 RBR+FFR does, aligning with findings from *in vitro* phase separation assays. However, the chemically more detailed Martini3 simulations does demonstrate the phase separation propensity of both RBR and RBR+FFR. Martini3 simulations are known to overemphasize the protein-protein interactions and a rebalancing of protein-protein [36, 60] and protein-water [46, 59] interactions may be required to study the phase behaviour of disordered proteins [46]. Nevertheless, the chemically detailed representation offered by Martini3 simulations allows us to elucidate the molecular drivers of phase separation, including 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, utilizing Martini3 simulations for back-mapping to atomistic resolution facilitates the investigation of the atomic-level interactions between MUT-16 RBR and MUT-8 N-terminal domain.

This ability to go back and forth between resolutions will increasingly become important to tackle increasingly large and complex systems with substantial chemical detail and sufficiently accurate description of molecular interactions [56]. In the context of the *Mutator focus*, the same "molecular grammar" and non-covalent interactions like cation-π and sp^2-π that regulates the dynamic interactions of disordered proteins appears to govern phase separation of MUT-16 and recruitment of the client MUT-8.

Supporting information

Acknowledgments

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 AA, Weber CA, 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 HC, Kim S, Chen B, Guo L, et al. Phase transitions in the assembly of multivalent signalling proteins. Nature. 2012;483(7389):336–340.

3. Brangwynne CP, Tompa P, Pappu RV. Polymer physics of intracellular phase transitions. Nature Physics. 2015;11(11):899–904.

4. Gomes E, Shorter J. The molecular language of membraneless organelles. Journal of Biological Chemistry. 2019;294(18):7115–7127.

5. Martin EW, Mittag T. Relationship of sequence and phase separation in protein low-complexity regions. Biochemistry. 2018;57(17):2478–2487.

6. Vernon RM, Chong PA, Tsang B, Kim TH, Bah A, Farber P, et al. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. elife. 2018;7:e31486.

7. Dignon GL, Best RB, Mittal J. Biomolecular phase separation: from molecular driving forces to macroscopic properties. Annual review of physical chemistry. 2020;71:53–75.

8. Martin EW, Holehouse AS, Peran I, Farag M, Incicco JJ, Bremer A, et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. Science. 2020;367(6478):694–699.

9. Wang J, Choi JM, Holehouse AS, Lee HO, Zhang X, Jahnel M, et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. Cell. 2018;174(3):688–699.

10. Sherrill CD. Energy component analysis of π interactions. Accounts of chemical research. 2013;46(4):1020–1028.

11. Reddy AS, Sastry GN. 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(39):8893–8903.

12. Gokel GW, De Wall SL, Meadows ES. Experimental evidence for alkali metal cation- π interactions. European journal of organic chemistry. 2000;2000(17):2967–2978.

13. Gokel GW, Barbour LJ, Stephen L, Meadows ES. Macrocyclic polyethers as probes to assess and understand alkali metal cation-π interactions. Coordination Chemistry Reviews. 2001;222(1):127–154.

14. Gallivan JP, Dougherty DA. Cation-π interactions in structural biology. Proceedings of the National Academy of Sciences. 1999;96(17):9459–9464.

15. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nature reviews Molecular cell biology. 2017;18(5):285–298.

16. Pritišanac I, Alderson TR, Kolarić , Zarin T, Xie S, Lu AX, et al. A Functional Map of the Human Intrinsically Disordered Proteome. bioRxiv. 2024; p. 2024–03.

17. Claycomb JM. Ancient endo-siRNA pathways reveal new tricks. Current Biology. 2014;24(15):R703–R715.

18. Ketting RF. The many faces of RNAi. Developmental cell. 2011;20(2):148–161.

19. Hutvagner G, Simard MJ. Argonaute proteins: key players in RNA silencing. Nature reviews Molecular cell biology. 2008;9(1):22–32.

20. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001;409(6818):363–366.

21. Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes & development. 2001;15(20):2654–2659.

22. Pak J, Fire A. Distinct populations of primary and secondary effectors during RNAi in C. elegans. Science. 2007;315(5809):241–244.

23. Sijen T, Steiner FA, Thijssen KL, Plasterk RH. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science. 2007;315(5809):244–247.

24. Lee SR, 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(7):604–610.

25. Colmenares SU, Buker SM, Buhler M, Dlakić M, Moazed D. Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Molecular cell. 2007;27(3):449–461.

26. Phillips CM, Montgomery TA, Breen PC, Ruvkun G. MUT-16 promotes formation of perinuclear mutator foci required for RNA silencing in the C. elegans germline. Genes & development. 2012;26(13):1433–1444.

27. Uebel CJ, Anderson DC, Mandarino LM, Manage KI, Aynaszyan S, Phillips CM. 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(7):e1007542.

28. Phillips CM, Montgomery BE, Breen PC, Roovers EF, Rim YS, Ohsumi TK, 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(8):839–844.

29. Tops BBJ, Tabara H, Sijen T, Simmer F, Mello CC, Plasterk RHA, et al. RDE-2 interacts with MUT-7 to mediate RNA interference in Caenorhabditis elegans. Nucleic Acids Research. 2005;33(1):347–355. doi:10.1093/nar/gki183.

30. Busetto V, Pshanichnaya L, Lichtenberger R, Hann S, Ketting RF, Falk S. MUT-7 exoribonuclease activity and localisation are mediated by an ancient domain. bioRxiv. 2023; p. 2023–12.

31. Dignon GL, Zheng W, Kim YC, Best RB, Mittal J. Sequence determinants of protein phase behavior from a coarse-grained model. PLoS Comput Biol. 2018;14:e1005941.

32. Mukherjee S, Schäfer LV. Thermodynamic forces from protein and water govern condensate formation of an intrinsically disordered protein domain. Nature Communications. 2023;14(1):5892. doi:10.1038/s41467-023-41586-y.

33. Gruijs da Silva LA, Simonetti F, Hutten S, Riemenschneider H, Sternburg EL, Pietrek LM, et al. Disease-linked TDP-43 hyperphosphorylation suppresses TDP-43 condensation and aggregation. The EMBO journal. 2022;41(8):e108443.

34. 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(12):3721–3740.

35. Zheng W, Dignon GL, Jovic N, Xu X, Regy RM, Fawzi NL, et al. Molecular details of protein condensates probed by microsecond long atomistic simulations. The Journal of Physical Chemistry B. 2020;124(51):11671–11679.

36. Benayad Z, von Bülow S, Stelzl LS, Hummer G. Simulation of FUS Protein Condensates with an Adapted Coarse-Grained Model. J Chem Theory Comput. 2021;17(1):525–537. doi:10.1021/acs.jctc.0c01064.

37. Souza PCT, Alessandri R, Barnoud J, Thallmair S, Faustino I, Grünewald F, et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. Nature Methods. 2021;18(4):382–388. doi:10.1038/s41592-021-01098-3.

38. Tesei G, Schulze TK, 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(44):e2111696118.

39. Joseph JA, Reinhardt A, Aguirre A, Chew PY, Russell KO, Espinosa JR, et al. Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. Nature Computational Science. 2021;1(11):732–743.

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(94). doi:10.12688/openreseurope.14967.2.

41. Shirts M, Pande VS. Screen savers of the world unite! Science. 2000;290(5498):1903–1904.

42. Anderson JA, Glaser J, Glotzer SC. HOOMD-blue: A Python Package for High-Performance Molecular Dynamics and Hard Particle Monte Carlo Simulations. Comput Mater Sci. 2020;173:109363.

43. Ashbaugh HS, Hatch HW. Natively unfolded protein stability as a coil-to-globule transition in charge/hydropathy space. Journal of the American Chemical Society. 2008;130(29):9536–9542.

44. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism From Laptops to Supercomputers. SoftwareX. 2015;1-2:19 – 25. doi:https://doi.org/10.1016/j.softx.2015.06.001.

45. Tsanai M, Frederix PWJM, Schroer CFE, Souza PCT, Marrink SJ. Coacervate formation studied by explicit solvent coarse-grain molecular dynamics with the Martini model. Chem Sci. 2021;12:8521–8530. doi:10.1039/D1SC00374G.

46. Thomasen FE, Pesce F, Roesgaard MA, Tesei G, Lindorff-Larsen K. Improving Martini 3 for Disordered and Multidomain Proteins. Journal of Chemical Theory and Computation. 2022;18(4):2033–2041.

47. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–589. doi:10.1038/s41586-021-03819-2.

48. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. The Journal of chemical physics. 2007;126(1).

49. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied physics. 1981;52(12):7182–7190.

50. Wassenaar TA, Pluhackova K, Bo¨ckmann RA, Marrink SJ, Tieleman DP. Going backward: a flexible geometric approach to reverse transformation from coarse grained to atomistic models. Journal of chemical theory and computation. 2014;10(2):676–690.

51. 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(3):712–725. doi:10.1002/prot.21123.

52. Best RB, Hummer G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. J Phys Chem B. 2009;113(26):9004–9015.

53. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. Proteins: Struct, Funct, Bioinf. 2010;78(8):1950–1958. doi:10.1002/prot.22711.

54. Best RB, De Sancho D, Mittal J. Residue-Specific α-Helix Propensities From Molecular Simulation. Biophys J. 2012;102(6):1462–1467. doi:10.1016/j.bpj.2012.02.024.

55. Piana S, Donchev AG, Robustelli P, Shaw DE. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. J Phys Chem B. 2015;119(16):5113–5123. doi:10.1021/jp508971m.

56. Rekhi S, Garcia CG, Barai M, Rizuan A, Schuster BS, Kiick KL, et al. Expanding the molecular language of protein liquid-liquid phase separation. bioRxiv. 2023; p. 2023–03.

57. Flores-Solis D, Lushpinskaia IP, Polyansky AA, Changiarath A, Boehning M, Mirkovic M, et al. Driving forces behind phase separation of the carboxy-terminal domain of RNA polymerase II. Nat Commun. 2023;14(1):5979. doi:10.1038/s41467-023-41633-8.

58. Zerze GH. 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;.

59. Thomasen FE, Skaalum T, Kumar A, Srinivasan S, Vanni S, Lindorff-Larsen K. Recalibration of protein interactions in Martini 3. bioRxiv. 2023; p. 2023–05.

60. Larsen AH, 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(4):e1007870.

61. Lancaster AK, Nutter-Upham A, Lindquist S, King OD. PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. Bioinformatics. 2014;30(17):2501–2502.

62. King OD, Gitler AD, Shorter J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain research. 2012;1462:61–80.

63. Gitler AD, Shorter J. RNA-binding proteins with prion-like domains in ALS and FTLD-U. Prion. 2011;5(3):179–187.

64. Babinchak WM, Haider R, Dumm BK, Sarkar P, Surewicz K, Choi JK, et al. The role of liquid–liquid phase separation in aggregation of the TDP-43 low-complexity domain. Journal of Biological Chemistry. 2019;294(16):6306–6317.

65. Udan-Johns M, Bengoechea R, Bell S, Shao J, Diamond MI, True HL, et al. Prion-like nuclear aggregation of TDP-43 during heat shock is regulated by HSP40/70 chaperones. Human molecular genetics. 2014;23(1):157–170.

66. Qamar S, Wang G, Randle SJ, Ruggeri FS, Varela JA, Lin JQ, et al. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interactions. Cell. 2018;173(3):720–734.

67. Holehouse AS, Ginell GM, 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(47):3566–3581.

68. Dignon GL, Zheng W, Mittal J. Simulation methods for liquid–liquid phase separation of disordered proteins. Current opinion in chemical engineering. 2019;23:92–98.

69. Dougherty DA. Cation-π interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science. 1996;271(5246):163–168.

70. Kumar K, Woo SM, Siu T, Cortopassi WA, Duarte F, Paton RS. Cation–π interactions in protein–ligand binding: Theory and data-mining reveal different roles for lysine and arginine. Chemical science. 2018;9(10):2655–2665.