Supplementary Materials for

Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions

Georg Krainer,[†,1] Timothy J. Welsh,[†,1] Jerelle A. Joseph,[†,2,3,4] Jorge R. Espinosa,[2,3,4] Ella de Csilléry,[1] Akshay Sridhar,[2,3,4] Zenon Toprakcioglu,[1] Giedre Gudiškytė,[1] Magdalena A. Czekalska,[1] William E. Arter,[1] Peter St George-Hyslop,[5,*] Rosana Collepardo-Guevara,[2,3,4*] Simon Alberti,[6,*] Tuomas P.J. Knowles[1,2,*]

Affiliations

[1] Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

[2] Cavendish Laboratory, Department of Physics, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, UK

[3] Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK

[4] Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

[5] Cambridge Institute for Medical Research, Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0XY, UK; Department of Medicine (Division of Neurology), University of Toronto and University Health Network, Toronto, Ontario M5S 3H2, Canada

[6] Biotechnology Center (BIOTEC), Center for Molecular and Cellular Bioengineering (CMCB), Technische Universität Dresden, Tatzberg 47/49, 01307 Dresden, Germany

* To whom correspondence should be addressed: Tuomas P.J. Knowles (tpjk2@cam.ac.uk), Rosana Collepardo-Guevara (rc597@cam.ac.uk), Simon Alberti (simon.alberti@tu-dresden.de), Peter St George-Hyslop (phs22@cam.ac.uk)

† These authors contributed equally to this work

This PDF file includes:

Materials and Methods

Sample preparation

All reagents and chemicals were purchased with the highest purity available. FUS and FUS G156E were produced as C-terminal EGFP fusion proteins as previously described (*1*) and stored at a concentration of 75 μM in 50 mM Tris-HCl (pH 7.4), 500 mM KCl, 1 mM DTT, 5% Glycerol. TDP-43 was similarly produced as a C-terminal GFP fusion protein as previously described (*2*) and stored at 120 μM in the same buffer as the FUS proteins. Annexin A11 was expressed and purified from insect Sf9 cells using standard procedures as previously described, labelled with Alexa647 and stored at a concentration of 150 μM in 20 mM HEPES (pH 7.0), 225 mM NaCl (*3*). The PR_{25} peptide, containing 25 proline–arginine repeats, was obtained from GenScript. N-terminally labelled PR_{25} was obtained by reacting the peptide with amine-reactive AlexaFluor546 (Sigma Aldrich). PolyU RNA with a molecular weight range from 800–1,000 kDa was purchased from Sigma Aldrich as lyophilized power and dissolved into a stock of 5 mg/mL in 50 mM Tris-HCl (pH 7.2) before use. ATP was obtained from Fisher Scientific and a 50 mM stock solution was prepared in 50 mM Tris-HCl (pH 7.2). 1,6-hexanediol was purchased from Santa Cruz Biotechnology Inc. and a 40% (*w*/*v*) stock solution was prepared in 50 mM Tris-HCl (pH 7.2). KCl was from Fisher Scientific and all the chloride salts used in the Hofmeister series experiments were from Sigma Aldrich. Phase separation of protein was induced by mixing the protein stock with the respective salts and additives as indicated. In all cases the buffer contained 50 mM Tris-HCl (pH 7.2). Phase separated samples were prepared in tubes and were always imaged within 1–5 minutes to limit any ageing effects.

Optical detection

Imaging was performed on an inverted fluorescence microscope (Zeiss AxioObserver D1) equipped with a high-sensitivity camera (Evolve 512 EMCCD, Photometrics). In experiments with FUS and TDP-43 an appropriate filter set for GFP detection was used (49002, Chroma Technology). Similarly, labelled PR_{25} was detected with the appropriate filter set for AlexaFluor546 fluorescence detection (49004, Chroma Technology). Annexin A11 was imaged with a filter set for AlexaFluor647 fluorescence detection (49009, Chroma Technology). All imaging was done with an aliquot of the sample (3 μL) placed on a microscope slide mounted on the microscope stage. For experiments involving dissolution induced by additives, components were mixed 3:1 with 3 μL of phase separated protein at the specified salt concentration with 1 μL of the additional component (*i.e.*, 1,6-hexandiol, PolyU RNA, ATP). The final concentrations of protein and additional components are stated in each figure.

Simulation methods

PMF Calculations. PMF calculations were carried out using the GROMACS simulation package (version 2019.3) (*4*). Amino acids were modelled using the AMBERff03ws force field (*5*). Since, we are interested in probing interaction potentials at very high salt concentrations (up to 3 M NaCl), it is very important that the solvent and ion model parameters employed are well-fitted to reproduce correct ion solubilities in water at 298 K (*i.e.*, in the absence of unphysical salt crystallization). The JC-SPC/E-ion/TIP4P/2005 force field has been optimized for that purpose, and so it was used in this work (*6*).The N- and C-terminal ends of each amino acid were capped with acetyl and an N-methyl capping groups, respectively. Pairs of amino acids were oriented with their sidechains facing each other, based on the most common arrangements observed in protein structures. Dimers were immersed in a cubic box containing TIP4P/2005 water molecules with a

minimum distance of 1.0 nm between the dimer and the edge of the box. Where necessary, some water molecules were replaced by Na^+ and/or Cl^- ions to produce neutral systems. Energy minimizations (force tolerance = 500 kJ mol^{-1} nm^{-1}) were performed for the neutralized systems, with positional restraints of 20,000 kJ mol^{-1} nm^{-2} applied in each dimension to all amino acid heavy-atoms. Na^+ and Cl^- ions were then added to yield the desired salt concentrations (0.5 M, 1.5 M, 3 M). For each concentration, approximately 30 windows, spaced at 0.05 nm from 0.1 to 1.6 nm, were used per amino acid pair. For production runs, positional restraints of 1000 kJ mol^{-1} nm^{-2} (in the directions perpendicular to the pulling direction) were used to constrain amino acid heavy-atoms. The COM distance between amino acid pairs was restrained with a harmonic umbrella potential (pulling force constant = 6000 kJ mol^{-1} nm^{-2}). Each window was simulated for 10 ns. Three independent simulations were conducted for each umbrella sampling window (*i.e.*, an aggregate simulation time of 30 ns per window). Umbrella sampling simulations were analyzed using WHAM, as implemented in GROMACS. The first 1000 ps of simulations were used for equilibration and were not included in the WHAM analysis. Error analysis was performed using the Bayesian bootstrap method, as described by Hub and coworkers (*7*).

Coarse-grained protein model. We used the sequence-dependent coarse-grained model of the Mittal group developed for proteins that undergo LLPS using a top-down parameterization that matches the experimental radius of gyration of a wide-range of intrinsically disordered proteins (*8*). This model treats intrinsically disordered protein regions as flexible polymers and globular regions as rigid bodies with each amino acid residue considered as a single bead. Inter-residue bonds within the disordered domains are described using harmonic springs. Long-range electrostatics are modelled using a Coulombic term with Debye–Hückel electrostatic screening. Nonbonded pairwise interactions are modelled using a knowledge-based potential termed HPS that is based on a hydrophobicity scale for amino acids (*9*). For the globular protein domains, a 30% scaled down set of the HPS parameters was used to account for 'buried' amino acids. Because the model distinguishes between disordered and globular protein regions and maintains the secondary structure of globular regions, it requires an initial atomistic model for the proteins; these are described below.

Initial atomistic models for coarse-grained simulations. We simulated the phase behavior of the full length FUS protein (Uniprot code: K7DPS7, 526 residues, 24 proteins) and a reduced version of the PR_{25} protein (13 Arg and 12 Pro residues alternately positioned, 400 proteins). Since the structure of full length FUS has not been resolved, we developed an atomistic model by attaching the intrinsically disordered regions to the resolved structural domains (residues from 285–371 (PDB code: 2LCW) and from 422-453 (PDB code: 6G99)). An initial intrinsically disordered model for PR_{25} was developed in VMD.

Coarse-grained simulation methods. To evaluate the formation of liquid condensates in the different systems, we performed direct coexistence simulations at constant volume and temperature. The direct coexistence method simulates the condensate and diluted phases in the same box separated by an interphase. The initial simulation box was prepared by running simulations at constant temperature and a pressure of 1 bar, using the Berendsen barostat, and then enlarging the simulation box in one direction ~3.5 times. The simulation temperatures were chosen to be just below the correspondent critical temperatures for each system: 280 K for full length FUS and 200 K for PR_{25}. For the production runs, each system was simulated for ~2 μs, using a Langevin thermostat with relaxation time of 5 ps and a time step of 10 fs (*10*). The LAMMPS software MD package was used to carry out all the coarse-grained simulations (*11*).

Supplementary Material for atomistic PMF calculations

To study the impact of salt on the electrostatic cation–π interactions, which are not well captured by standard atomistic force fields (*12*) as they involve the polarization of the π electron cloud of an aromatic side chain (Tyr, Phe, Trp) due to the cationic side chain (Lys, Arg), we developed a molecular model of Tyr that mimics its polarized state. The issue is that, non-polarizable atomistic force fields are unable to capture cation–π interactions; such interactions rely heavily on non-additive effects (*13*). In order to study these interactions in our atomistic simulations, we modified the charge distributions in the sidechains of arginine and tyrosine. For arginine, the terminal part of the sidechain was redefined to be purely positive, while in tyrosine all carbon atoms of the aromatic ring were assigned a negative charge. The overall charge of arginine and tyrosine was maintained at +1 and 0, respectively. The original and modified force field parameters are summarized below:

Atom in Arg	Original Charge	Modified Charge	Atom in Tyr	Original Charge	Modified Charge
CB	0.036707	-0.46489	CB	-0.051853	0.040185
CG	0.012454	-0.685774	HB1	0.019145	0.315207
CD	0.126329	-0.685774	HB2	0.019145	0.315207
NE	-0.46489	0.036707	CG	0.112601	-0.183461
NH1	-0.685774	0.012454	CZ	0.206277	-0.181823
NH2	-0.685774	0.126329			

Supplementary Figure

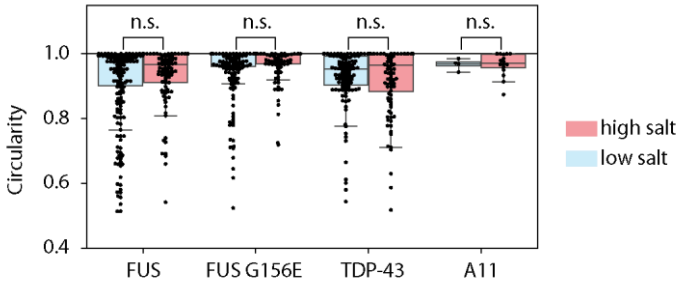

Figure S1. Analysis of condensate circularity in the low and high salt regime for FUS, FUS G156E, TDP-43, and A11. Boxes extend from the 25th to 75th percentiles, with a line at the median. Whiskers span 1.5× the interquartile range. In all cases, median circularities were >0.95. For each protein, at high salt and low salt, respectively, the median circularities were: FUS: 0.97, 0.99; FUS G156E: 1.00, 1.00; TDP-43: 0.96, 0.95; A11: 0.97, 0.97. Statistical analysis was performed using a two-sided *t*-test. No significant difference in circularity (n.s., not significant) between the low and high salt regime was found (FUS: $p > 0.215$; FUS G156E: $p > 0.508$; TDP-43: $p > 0.869$; A11: $p > 0.9973$). Circularity analysis was performed in Fiji/ImageJ and statistical analysis in Python.

Supplementary References

1. A. Patel, H. O. Lee, L. Jawerth, S. Maharana, M. Jahnel, M. Y. Hein, S. Stoynov, J. Mahamid, S. Saha, T. M. Franzmann, A. Pozniakovski, I. Poser, N. Maghelli, L. A. Royer, M. Weigert, E. W. Myers, S. Grill, D. Drechsel, A. A. Hyman, S. Alberti, A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*. **162**, 1066–1077 (2015).
2. S. Maharana, J. Wang, D. K. Papadopoulos, D. Richter, A. Pozniakovsky, I. Poser, M. Bickle, S. Rizk, J. Guillén-Boixet, T. M. Franzmann, M. Jahnel, L. Marrone, Y. Chang, J. Sterneckert, P. Tomancak, A. A. Hyman, S. Alberti, RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science*. **360**, 918–921 (2018).
3. Y. C. Liao, M. S. Fernandopulle, G. Wang, H. Choi, L. Hao, C. M. Drerup, R. Patel, S. Qamar, J. Nixon-Abell, Y. Shen, W. Meadows, M. Vendruscolo, T. P. J. Knowles, M. Nelson, M. A. Czekalska, G. Musteikyte, M. A. Gachechiladze, C. A. Stephens, H. A. Pasolli, L. R. Forrest, P. St George-Hyslop, J. Lippincott-Schwartz, M. E. Ward, RNA Granules Hitchhike on Lysosomes for Long-Distance Transport, Using Annexin A11 as a Molecular Tether. *Cell*. **179**, 147-164.e20 (2019).
4. S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. **29**, 845–854 (2013).
5. R. B. Best, W. Zheng, J. Mittal, Balanced Protein–Water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. *J. Chem. Theory Comput.* **10**, 5113–5124 (2014).
6. L. H. Kapcha, P. J. Rossky, A simple atomic-level hydrophobicity scale reveals protein interfacial structure. *J. Mol. Biol.* **426**, 484–98 (2014).
7. J. S. Hub, B. L. De Groot, D. Van Der Spoel, G-whams-a free Weighted Histogram Analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* **6**, 3713–3720 (2010).
8. G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, J. Mittal, Sequence determinants of protein phase behavior from a coarse-grained model. *PLOS Comput. Biol.* **14**, e1005941 (2018).
9. L. H. Kapcha, P. J. Rossky, A Simple Atomic-Level Hydrophobicity Scale Reveals Protein Interfacial Structure. *J. Mol. Biol.* **426**, 484–498 (2014).
10. J. S. Rowlinson and B. Widom, *Molecular Theory of Capillarity* (Clarendon Press, Oxford, 1984).
11. S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
12. H. M. Khan, C. Grauffel, R. Broer, A. D. MacKerell, R. W. A. Havenith, N. Reuter, Improving the Force Field Description of Tyrosine–Choline Cation−π Interactions: QM Investigation of Phenol–N(Me)4+ Interactions. *J. Chem. Theory Comput.* **12**, 5585–5595 (2016).
13. J. W. Caldwell, P. A. Kollman, Cation-.pi. Interactions: Nonadditive Effects Are Critical in Their Accurate Representation. *J. Am. Chem. Soc.* **117**, 4177–4178 (1995).