Liquid-liquid phase separation in amino acid mixtures is governed by composition

David De Sancho*

Polimero eta Material Aurreratuak: Fisika, Kimika eta Teknologia, Kimika Fakultatea, UPV/EHU & Donostia International Physics Center (DIPC), PK 1072, 20018 Donostia-San Sebastian, Euskadi, Spain

E-mail: david.desancho@ehu.eus

Phone: +123 (0)123 4445556. Fax: +123 (0)123 4445557

Abstract

Liquid-liquid phase separation (LLPS) has recently come to immense prominence as it is central to the formation of membraneless organelles, leading to a new paradigm of cellular organization. This type of phase transition is mediated by molecular interactions between biomolecules, including nucleic acids and both ordered and disordered proteins. In the latter case, the separation between protein-dense and dilute phases is often interpreted using models adapted from polymer theory. Specifically, the "stickers and spacers" model proposes that LLPS originates from the interplay between two classes of residues and that the main determinants for phase separation are multivalency and sequence patterning. The duality of roles of stickers (aromatics like Phe and Tyr) and spacers (Gly and polar residues) may apply more broadly in protein-like mixtures, and the presence of these two types of components alone may suffice for LLPS to take place. In order to explore this hypothesis, we use atomistic molecular dynamics simulations of capped amino-acid residues as a minimal model system. We study the behaviour of pure amino acids in water for three types of residues that are thought to play critical roles in LLPS, corresponding to the spacer and sticker categories, and their multicomponent

mixtures. In agreement with previous observations, we find that the spacer-type amino acids fail to phase-separate on their own, while the sticker is prone to aggregation. However, ternary amino acid mixtures involving both types of amino acids phase-separate into two phases that retain intermediate degrees of compaction and considerable fluidity. Our results suggest that LLPS is an emergent property of amino acid mixtures determined by composition.

Cellular organisms are well known to use lipid membranes to compartmentalise their many functions into organelles. In the last decade, it has become apparent that organelles lacking membranes are also ubiquitous in cells.[1] These membraneless organelles include the nucleolus, stress granules, nuclear speckles and Cajal bodies, among many others. They are rich in proteins and also nucleic acids,[2] which form molecular condensates as they undergo a type of phase transition called liquid-liquid phase separation (LLPS). This term is used because, as a result, a dilute and a concentrated phase are formed, both of which have liquid-like properties such as fusion into larger droplets, flow in response to stress, or wetting and dropping behaviours typical of liquids.[3] The arrival to prominence of membraneless organelles has sparked considerable interest in the types of molecules undergoing LLPS and the physical principles governing this process.[4] Although LLPS has been reported for modular folded domains,[5] many proteins that phase separate are intrinsically disordered or have intrinsically disordered regions. These IDPs or IDRs have strong biases in sequence composition,[6–8] as in the case of the "low complexity regions" (LCRs) that are abundant in biomolecular condensates.[9]

Because polypeptide chains are often the drivers of LLPS, concepts from polymer physics have been borrowed and adapted to interpret experimental results.[3] One such model is the "stickers and spacers model", where different types of modules come to play different roles in the polymer solution.[4] In the case of IDPs or IDRs, aromatic amino acid residues like Tyr or Phe behave as stickers that can form intra and intermolecular contacts, while glycine and polar amino acid residues are spacers that do not have strong interaction patterns. This simple model has been used successfully to interpret experiments on prion-like low complexity domains, which are frequently involved in the formation of condensates.[10] Both experiments and simulations show that specific

patterns of stickers and spacers in polypeptide sequences are important determinants of phase separation.[7,8,10–12] Specifically, both randomly and regularly distributed stickers result in LLPS while patchy regions with multiple stickers clustered together result in aggregation, or liquid-solid (as opposed to liquid-liquid) phase separation (i.e. LSPS).[10]

The duality from the stickers and spacers model for different amino acid types may apply more broadly, even in the absence of an ordered polypeptide sequence where multivalency and patterning effects become important. In fact, composition alone may be sufficient to drive phase separation in the absence of additional information. Here we explore this possibility using atomistic molecular dynamics (MD) simulations. Molecular simulations with atomistic resolution of condensate formation have been attempted by others before.[13–15] Unfortunately, these simulations often involve multiple protein chains and large simulation boxes, which cannot be run for sufficiently long timescales for observing LLPS. Alternatively, much more computationally cheap coarse-grained simulation models can recapitulate general principles of phase separation of disordered proteins,[6,16–18] although details of chemical interactions are inevitably lost. In this work we run atomistic MD simulations of amino acid mixtures as a minimal, albeit detailed, model for LLPS in protein-like systems. This type of simulations has been carried out in the past to study crowding and peptide solubility,[19–21] although often in single component boxes. Here we attempt to capture the behaviour that emerges from mixing amino acids of the sticker and spacer types, using compositions inspired by those observed in LCRs. Specifically, we choose Ser and Gly, which make over 50% of the sequence in some LCRs,[9] as representative spacers, and the aromatic Tyr, which has been shown to be key for phase separation,[11] as a sticker.

We have prepared elongated simulation boxes using the Packmol software,[22] adding multiple copies of acetylated and amidated amino acid residues (i.e. Ace-X-NH_2, usually termed dipeptides) into 10 nm$\times$3.5 nm$\times$3.5 nm simulation boxes in order to reach concentrations between 1 and 2.5 M (see Supporting Information, Table S1). The amino acid boxes were then solvated, energy minimized and equilibrated in the NVT and NPT ensembles at 300 K. Long production simulations of 500 ns were then run for all systems at 300 K in the NVT ensemble. In all cases,

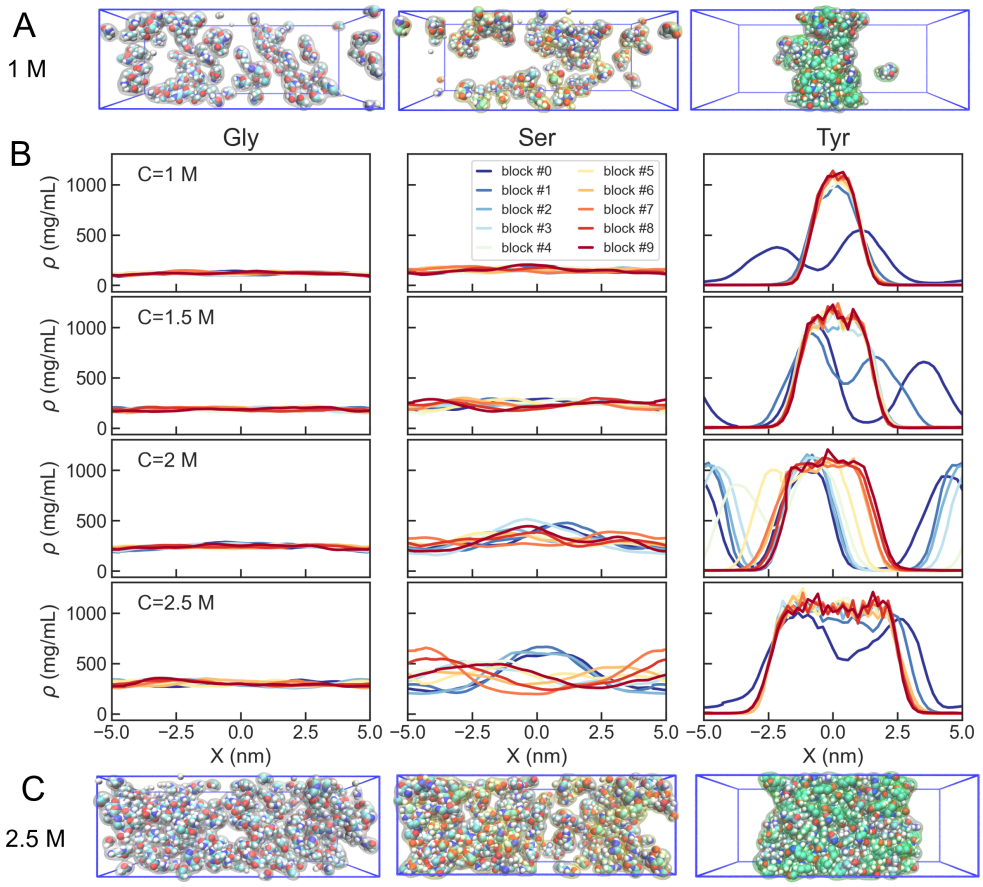

Figure 1: (A, C) Snapshots from simulations at 1 M and 2.5 M for Gly (white), Ser (yellow) and Tyr (green) dipeptide simulations (left, center and right, respectively). (B) Density profiles along the longest axis of the simulation box for the pure dipeptide solutions at different concentrations for glycine (left column), serine (center) and tyrosine (right). Different colours correspond to averages over 50 ns blocks in the 500 ns simulations.

a stochastic velocity-rescaling thermostat was used[23] and, when needing a barostat, we used the method developed by Berendsen et al.[24] The Amber 99SB-ILDN force field[25] together with the TIP3P water model[26] were used, a combination that has been shown to appropriately reproduce experimental solution densities in simulations of crowded amino acids.[19] We use the GROMACS program[27] (version 2021) to run all the simulations. Additional details of our simulations are shown in the Supplementary Methods.

First we focus on the simulations of pure amino acid solutions (see Fig. 1). In the case of the Gly boxes, we find that the peptide remains soluble even at the largest concentration for the full duration of the MD simulations, resulting in flat density profiles. For Ser, the dipeptide solution is

mostly homogeneous up to 2 M, while at 2.5 M some regions become enriched in amino acid (Fig. 1, center column). However the difference between high density and low density regions is less than two-fold, suggesting that phase separation has not occurred at least in our simulation time. On the contrary, in the case of Tyr, we find phase separation across all the concentrations explored, resulting in the formation of a slab (Fig. 1, right column). At all concentrations, condensate formation takes place within the first 50 ns of the simulation run. The differences observed between Gly/Ser and Tyr are consistent with the very low solubility reported for experiments on Tyr in its zwitterionic state[28] and suggests that very small amounts of this amino acid may have drastic effects in properties of mixtures, as indeed has been found experimentally.[11]

We then repeat the same type of calculations but in this case for amino acid mixtures of spacers and stickers (see Fig. 2). We first look at equimolecular combinations of Gly and Ser for total amino acid concentrations between 1 and 2.5 M, as before (hence ranging between 0.5 and 1.25 M for each component). We find that these mixtures retain the properties of the single amino acid solutions of their components. At all concentrations probed in our calculations, the density profiles remain flat (Fig. 2, left column), which is consistent with the mixture exhibiting properties intermediate between their individual components. However, simulations for the same range of total concentration with stoichiometric amounts of Tyr result in the separation of the amino acid solution into dilute and dense phases (see Fig. 2, right column). The densities of these protein-dilute and protein-rich phases are higher and lower, respectively, than those observed for proteins in experiments and previous atomistic simulations.[14] This may be due to excluded volume effects of the polypeptide chain, but also to the combination of force field and water model used in this work. In fact, the results are consistent with the strong tendency of the Amber 99SB-ILDN with TIP3P water combination to induce compact states in proteins.[29] We have run simulations with two other optimized force fields of the Amber family (see Figure S1 in the Supporting Information). Using the Amber03* force field[30] with TIP3P water, the density profiles for the Gly/Ser/Tyr mixture are comparable to those obtained using Amber99SB-ILDN. With the more recently derived a99SB-disp force field and water,[25] the amino acid mixture fails to phase-separate. The latter result can

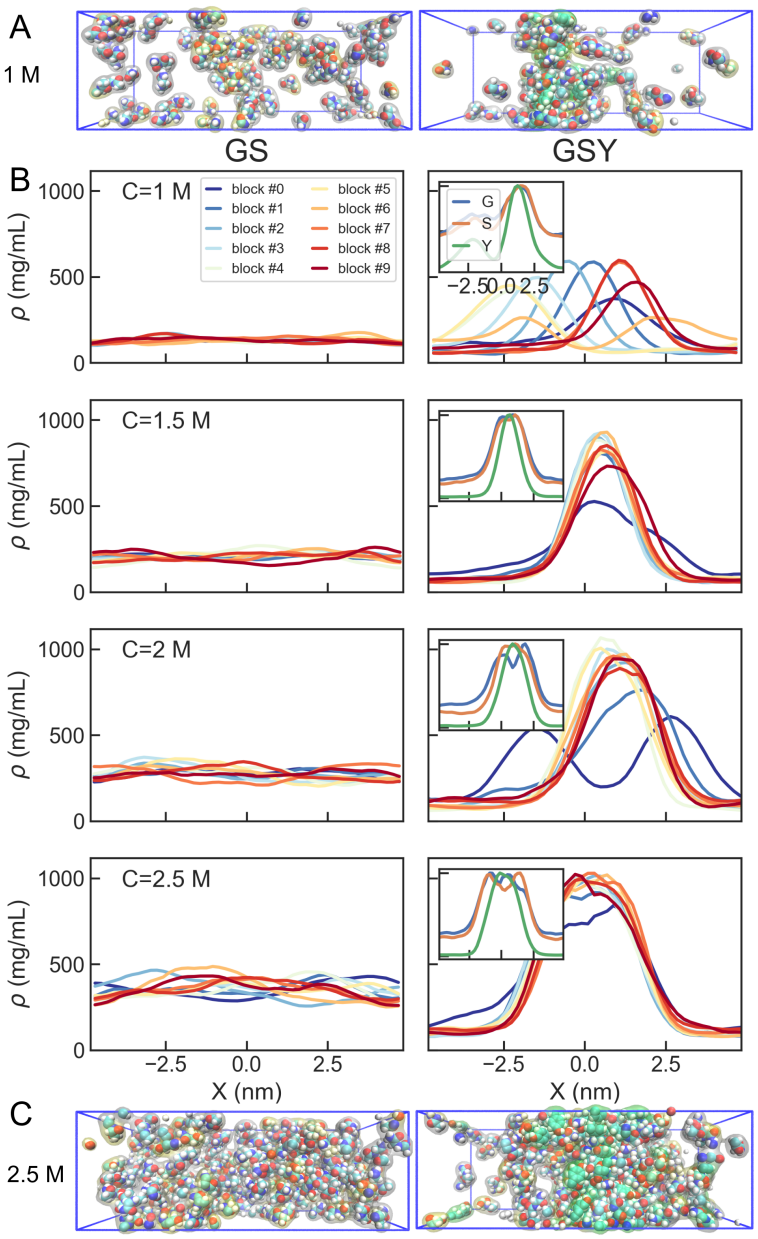

Figure 2: (A, C) Snapshots from simulations at 1 M and 2.5 M for Gly/Ser (left) and Gly/Ser/Tyr (right) mixtures. (B) Density profiles along the longest axis for the mixtures at different concentrations.

be attributed to the same reasons that prevent this specific force field and water model combination to adequately capture protein-protein binding, which led to its reparametrization.[31] In any case, the properties of the simulated amino acid condensates will depend on the choice of the force field.

We have seen that our results from simulations of very simple mixtures of capped "sticker and spacer" amino acids recapitulate properties observed in many disordered protein segments,

which experience a phase transition with an upper critical solution temperature (UCST).[9] In order to analyze the temperature dependence of condensate formation, we have run an additional set of simulations at multiple temperatures ranging between 270 and 400 K for the 2 M Gly/Ser/Tyr box. As can be seen in Figure 3A, there is coexistence between a protein-dense and a dilute phase for a broad range of conditions, and as the temperature increases the difference between the densities of the dense and dilute phases (ρ_H and ρ_L, respectively) narrows down. The resulting phase diagram can be fitted to extract a critical temperature of $T_c = 402$ K (see Fig. 3B).

So far we have treated the amino acid mixtures as homogeneous protein solutions. However, lacking the connectivity of covalently bonded polypeptide chains, amino acids could be partitioning independently of each other. We have used the density profiles for Gly, Ser and Tyr to calculate their normalized compositions along the simulation boxes (i.e. $\rho(x)/\rho_{max}$, see Fig. 2B, insets). Although all three amino acids are enriched in the slab, the condensed phase is richer in Tyr than in Gly/Ser, which is consistent with the the greater hydrophobicity of this amino acid. In order to map the composition dependence of the phase behaviour of our mixtures, we have run simulations varying the ratio between stickers and spacers while retaining the total number of amino acid residues in the mixtures constant. The resulting phase diagram is shown in Figure 3C. The emerging picture is that of a ”scaffold-client” system,[32] where the spacer-type residues, which cannot phase-separate on their own, are recruited by the stickers. This is due to the greater strength of the homotypic interactions between tyrosine residues, relative to every other pairwise interactions. However, heterotypic interactions between tyrosine and glycine/serine also contribute to the stability of the condensate. This is supported by the statistics of pairwise contacts among residues (which we show for the equimolar mixture in Fig. S2 of the Supporting Information). While Tyr-Tyr contacts dominate, Tyr-Ser and Tyr-Gly contacts are also relevant. We note that the relative weights of different types of contacts is comparable to that observed by Zheng and his co-workers in atomistic simulations of FUS LC and LAF1 RGG (i.e. with Tyrosine contributing twice as much as other residue types).[14] Our results support that spacers are not inert but relevant actors in the phase-behaviour of protein condensates.

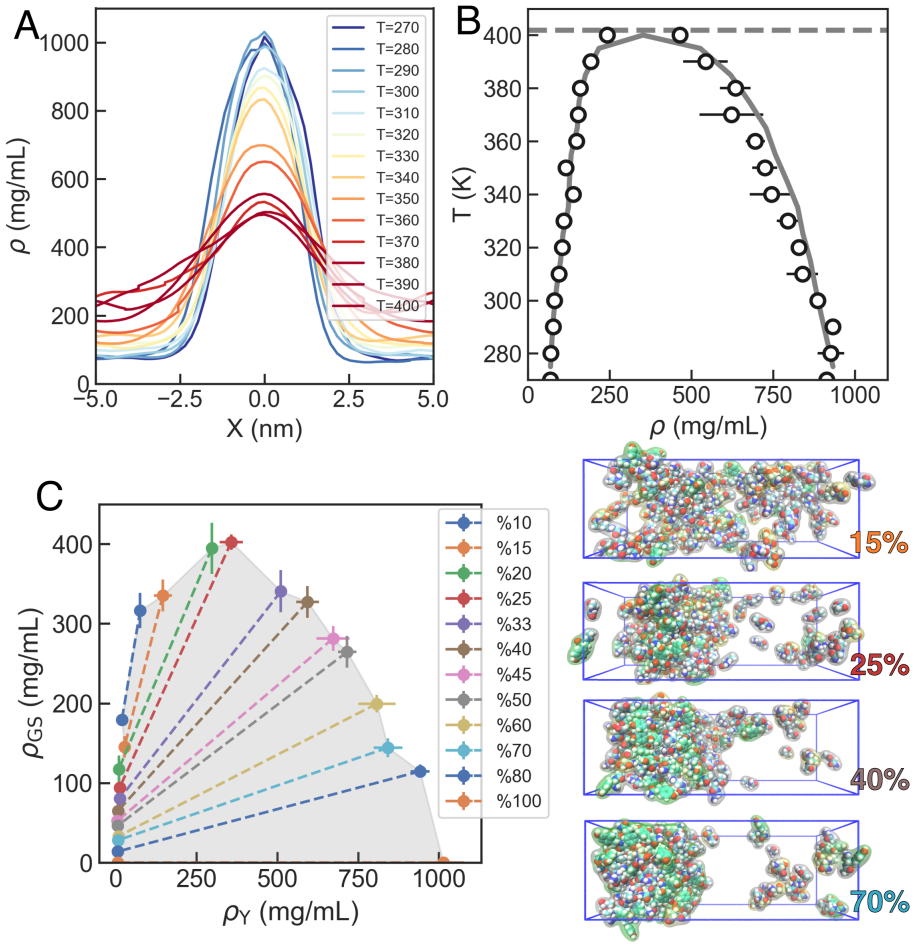

Figure 3: (A) Density profiles for simulations of Gly/Ser/Tyr at 2M concentration at multiple temperatures. (B) Phase diagram obtained from density profiles, where ρ is the total protein density and T is the simulation temperature. The dashed line indicates the fitted critical temperature, T_c. (C) Two-component phase diagram from simulations of Gly/Ser/Tyr mixtures for different compositions, represented in terms of sticker and spacer densities (ρ_Y and ρ_{GS}, respectively) for the dilute and dense phases. Different colours indicate simulated fraction percentage of Tyr within the mixture. On the right we show representative snapshots of the simulation for different compositions.

An important consideration is whether the condensates formed are aggregates (hence formed by liquid-solid phase separation, LSPS) or, instead, have liquid-like properties. Even if a neat distinction may be difficult to establish, we can inspect both equilibrium and dynamical properties to clarify this point. First we look at the accessible surface area for the pure solutions and equimolar mixtures. In Figure 4A, we show that at 2 M the solvent accessible surface area (*SASA*) for the Gly/Ser/Tyr mixes decreases to a value intermediate between that of the Gly/Ser or Gly/Ser solu-

tions and that of the Tyr simulation box. This result occurs consistently across the full range of concentrations explored in our simulations (see Supplementary Fig. S3). This type of intermediate behaviour for the Gly/Ser/Tyr mixture is also observed for the maximum cluster size (Fig. 4A, bottom, and Fig. S4 in the Supporting Information), which has also been shown to be a robust descriptor of the degree of clustering[33] (see also the average cluster size and the distribution of cluster sizes in Figs. S5 and S6 in the Supporting Information). Two useful metrics that summarize the information from these time-dependent observables are the "aggregation propensity", defined as the ratio between the initial and instantaneous solvent accessible surface areas,[34] and the "clustering degree", which we obtain normalizing the maximum cluster size by the total number of residues in the box.[33] In the projection of our datasets on these axes (Fig. 4B), the Gly/Ser/Tyr system stands out as having intermediate values of the aggregation propensity and clustering degree between the systems that do not phase separate (Gly, Ser and Gly/Ser mixture) and the Tyr boxes, which are best described as aggregates.

The results from clustering and collapse metrics are confirmed when we investigate the dynamical properties in our model systems. First, we use Einstein's relation to calculate diffusion constants for the longest box dimension, D_x, for both the pure amino acid solutions and their mixtures at the different concentrations (see Fig. 4C). In all cases, we observe a decrease of the diffusion coefficient with increasing amino acid concentration, which is consistent with expectations from crowding.[35] However, this decrease is sharper in the case of the Tyr solutions. More importantly, we find that there is a timescale gap of between 1 and 2 orders of magnitude in the diffusion constant of Gly, Ser or Gly/Ser solutions and that of Tyr, suggesting that the dynamics of Tyr are arrested relative to what we find in the homogeneous solutions (Fig. 4D). For the Gly/Ser/Tyr mixture we obtain a diffusion coefficient intermediate between those of the Gly, Ser and Gly/Ser solutions and the Tyr simulations. Differences in the dynamics can also be found in the intramolecular dynamics, which we measure using autocorrelation functions of the Ramachandran angles (Φ and Ψ) and the χ_1 sidechain torsion. The dynamics of Gly and Ser remain unchanged in the binary and ternary mixtures relative to the pure solutions (see Supplementary Information Figs. S7 and

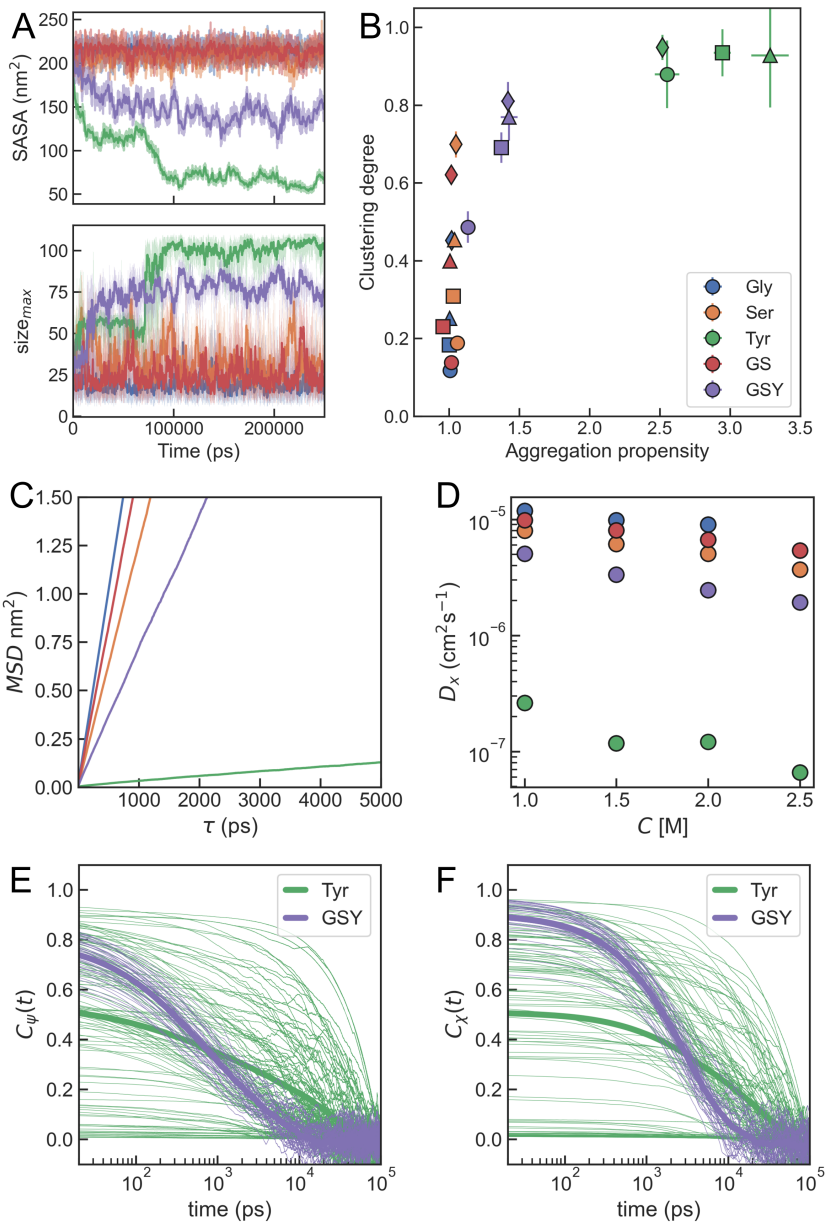

Figure 4: (A) Time series data for the solvent accessible surface area (*SASA*, top) and the clustering degree (bottom) for all systems in the 1.5 M solutions. (B) Projection of all the simulation datasets on the aggregation propensity and clustering degree. Different symbols total peptide concentrations (circles: 1 M; squares: 1.5 M; triangles: 2 M; diamonds: 2.5 M). (C) Mean squared displacement as a function of time for amino acid solutions and mixtures. (D) Concentration dependence of diffusion constants. (E) Dihedral autocorrelation functions for the individual tyrosine residues in the pure solution (green) and in the Gly/Ser/Tyr mixture (purple) for the ψ backbone torsion. Thick lines are averages over all residues. (F) Same, for the χ_1 sidechain torsion.

S8). On the other hand, for Tyr both the backbone and sidechain dynamics are considerably faster in the multicomponent mixture (see Fig. 4E-F). All of this, together with substantially greater

fluctuations in the time series data for both the *SASA* (Fig. 4A), is consistent with the Gly/Ser/Tyr condensates having greater fluidity than those formed by Tyr.

In summary, in this work we have showed with a minimalistic model system that LLPS is an emergent property of amino-acid mixtures that is governed by composition. Considering a mixture of "spacer" residues (Gly/Ser) as a reference system –which is warranted by the abundance of these amino acids in phase-separating LCRs– we show that small amounts of a "sticker" (in this case the rather insoluble Tyr amino acid) make their solutions separate into two phases with liquid-like properties. This suggests that, the duality of amino acid types in the "sticker and spacers" model holds more generally than in the context of IDPs and that the stoichiometry of the mixture alone may be predictive of the phase behaviour. Provided a small amount of aromatics or "stickers" is present, amino-acids, peptides or proteins may be able to spontaneously phase separate. The critical role of Tyr in our simulations is consistent with previous experiments on folded modules and intrinsically disordered regions.[8,10,11] The formation of aggregates as opposed to liquid-like condensates in the case of patchy sequences with clusters of stickers[10] may be understood as being closer to the results from our simulations of Tyr-only boxes, which result in the formation of a solid agregate (i.e. LSPS). Our work also confirms in full atomistic detail recent results by Tang et al,[33] who explored dipeptide solutions in simulations using the intermediate-resolution MARTINI model.[36] Further evidence for the generality of LLPS is accumulating in the experimental front in a broad range of length-scales, from short peptides[37,38] to globular proteins.[39] The occurrence of LLPS may hence turn out to be a general property of both peptides and proteins, determined by composition and further fine-tuned by multivalency and patterning,[10] the presence of charged residues,[8] nucleic acids,[40,41] ionic strength[18] or crowding agents.[39]

Acknowledgement

Financial support to DDS comes from Eusko Jaurlaritza (Basque Government) through the project IT1254-19, Grants RYC-2016-19590 and PGC2018-099321-B-I00 from the Spanish Ministry of

Science and Universities through the Office of Science Research (MINECO/FEDER) and the Donostia International Physics Center (DIPC). The author thanks Xabier López for useful discussions and Athi N. Naganathan and Robert B. Best for their comments on the manuscript. The author also acknowledges the staff at the DIPC Supercomputing Center for technical support.

References

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

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

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

(4) Choi, J.-M.; Holehouse, A. S.; Pappu, R. V. Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annu. Rev. Biophys.* **2020**, *49*, 107–133.

(5) Li, P.; Banjade, S.; Cheng, H.-C.; Kim, S.; Chen, B.; Guo, L.; Llaguno, M.; Hollingsworth, J. V.; King, D. S.; Banani, S. F.; Russo, P. S.; Jiang, Q.-X.; Nixon, B. T.; Rosen, M. K. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **2012**, *483*, 336–340.

(6) Dignon, G. L.; Zheng, W.; Kim, Y. C.; Best, R. B.; Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.* **2018**, *14*, 1–23.

(7) Lin, Y.-H.; Forman-Kay, J. D.; Chan, H. S. Theories for Sequence-Dependent Phase Behaviors of Biomolecular Condensates. *Biochemistry* **2018**, *57*, 2499–2508.

(8) Bremer, A.; Farag, M.; Borcherds, W. M.; Peran, I.; Martin, E. W.; Pappu, R. V.; Mittag, T.

Deciphering how naturally occurring sequence features impact the phase behaviors of disordered prion-like domains. *bioRxiv* **2021**,

(9) Martin, E. W.; Mittag, T. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry* **2018**, *57*, 2478–2487.

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

(11) Lin, Y.; Currie, S. L.; Rosen, M. K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J. Biol. Chem.* **2017**, *292*, 19110–19120.

(12) Holehouse, A. S.; Ginell, G. M.; Griffith, D.; Böke, E. Clustering of Aromatic Residues in Prion-like Domains Can Tune the Formation, State, and Organization of Biomolecular Condensates. *Biochemistry* **2021**, *60*, 3566–3581.

(13) Rauscher, S.; Pomès, R. The liquid structure of elastin. *eLife* **2017**, *6*, e26526.

(14) Zheng, W.; Dignon, G. L.; Jovic, N.; Xu, X.; Regy, R. M.; Fawzi, N. L.; Kim, Y. C.; Best, R. B.; Mittal, J. Molecular Details of Protein Condensates Probed by Microsecond Long Atomistic Simulations. *J. Phys. Chem. B* **2020**, *124*, 11671–11679.

(15) Paloni, M.; Bailly, R.; Ciandrini, L.; Barducci, A. Unraveling Molecular Interactions in Liquid–Liquid Phase Separation of Disordered Proteins by Atomistic Simulations. *J. Phys. Chem. B* **2020**, *124*, 9009–9016.

(16) Dignon, G. L.; Zheng, W.; Best, R. B.; Kim, Y. C.; Mittal, J. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 9929–9934.

(17) Regy, R. M.; Thompson, J.; Kim, Y. C.; Mittal, J. Improved coarse-grained model for studying sequence dependent phase separation of disordered proteins. *Protein Sci.* **2021**, *30*, 1371–1379.

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

(19) Andrews, C. T.; Elcock, A. H. Molecular Dynamics Simulations of Highly Crowded Amino Acid Solutions: Comparisons of Eight Different Force Field Combinations with Experiment and with Each Other. *J. Chem. Theory Comput.* **2013**, *9*, 4585–4602.

(20) Karandur, D.; Wong, K.-Y.; Pettitt, B. M. Solubility and Aggregation of Gly5 in Water. *J. Phys. Chem. B* **2014**, *118*, 9565–9572.

(21) Sarma, R.; Wong, K.-Y.; Lynch, G. C.; Pettitt, B. M. Peptide Solubility Limits: Backbone and Side-Chain Interactions. *J. Phys. Chem. B* **2018**, *122*, 3528–3539, PMID: 29384681.

(22) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164.

(23) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(24) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(25) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958.

(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(27) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19 – 25.

(28) Nozaki, Y.; Tanford, C. The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions: establishment of a hydrophobicity scale. *J. Biol. Chem.* **1971**, *246*, 2211–2217.

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

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

(31) Piana, S.; Robustelli, P.; Tan, D.; Chen, S.; Shaw, D. E. Development of a Force Field for the Simulation of Single-Chain Proteins and Protein–Protein Complexes. *J. Chem. Theory Comput.* **2020**, *16*, 2494–2507.

(32) Dignon, G. L.; Best, R. B.; Mittal, J. Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. *Annu. Rev. Phys. Chem.* **2020**, *71*, 53–75.

(33) Tang, Y.; Bera, S.; Yao, Y.; Zeng, J.; Lao, Z.; Dong, X.; Gazit, E.; Wei, G. Prediction and characterization of liquid-liquid phase separation of minimalistic peptides. *Cell Rep. Phys. Sci.* **2021**, *2*, 100579.

(34) Frederix, P. W. J. M.; Ulijn, R. V.; Hunt, N. T.; Tuttle, T. Virtual Screening for Dipeptide

Aggregation: Toward Predictive Tools for Peptide Self-Assembly. *J. Phys. Chem. Lett.* **2011**, *2*, 2380–2384.

(35) Nawrocki, G.; Wang, P.-h.; Yu, I.; Sugita, Y.; Feig, M. Slow-Down in Diffusion in Crowded Protein Solutions Correlates with Transient Cluster Formation. *J. Phys. Chem. B* **2017**, *121*, 11072–11084.

(36) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.

(37) Wang, Y.; Lomakin, A.; Kanai, S.; Alex, R.; Benedek, G. B. Liquid–Liquid Phase Separation in Oligomeric Peptide Solutions. *Langmuir* **2017**, *33*, 7715–7721.

(38) Yuan, C.; Levin, A.; Chen, W.; Xing, R.; Zou, Q.; Herling, T. W.; Challa, P. K.; Knowles, T. P. J.; Yan, X. Nucleation and Growth of Amino Acid and Peptide Supramolecular Polymers through Liquid–Liquid Phase Separation. *Angew. Chem. Int. Ed.* **2019**, *58*, 18116–18123.

(39) Poudyal, M.; Patel, K.; Sawner, A. S.; Gadhe, L.; Kadu, P.; Datta, D.; Mukherjee, S.; Ray, S.; Navalkar, A.; Maiti, S.; Chatterjee, D.; Bera, R.; Gahlot, N.; Padinhateeri, R.; Maji, S. K. Liquid condensate is a common state of proteins and polypeptides at the regime of high intermolecular interactions. *bioRxiv* **2022**,

(40) Mann, J. R.; Donnelly, C. J. RNA modulates physiological and neuropathological protein phase transitions. *Neuron* **2021**, *109*, 2663–2681.

(41) Dutagaci, B.; Nawrocki, G.; Goodluck, J.; Ashkarran, A. A.; Hoogstraten, C. G.; Lapidus, L. J.; Feig, M. Charge-driven condensation of RNA and proteins suggests broad role of phase separation in cytoplasmic environments. *Elife* **2021**, *10*, e64004.