Tandem-repeat proteins introduce tuneable properties to engineered biomolecular condensates

Tin Long Chris Ng*[1], Mateo P. Hoare*[1], M. Julia Maristany*[2,3], Ellis J. Wilde[1], Tomas Sneideris[2], Jan Huertas[2,4], Belinda K. Agbetiameh[1], Mona Furukawa[1], Jerelle A. Joseph[5,6], Tuomas P.J. Knowles[2], Rosana Collepardo-Guevara[2,3,4], Laura S. Itzhaki[1,§] and Janet R. Kumita[1,§]

[1]Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge, CB2 1PD, UK.

[2]Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK.

[3]Department of Physics, University of Cambridge, Cavendish Laboratory, Cambridge, CB3 0FZ, UK.

[4]Department of Genetics, University of Cambridge, Downing Pl, Cambridge CB2 3EH, UK.

[5]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ, 08544, USA.

[6]Omenn–Darling Bioengineering Institute, Princeton University, Princeton, NJ, 08544, USA.

*these authors contributed equally

[§]To whom correspondence should be addressed. E-mail: jrk38@cam.ac.uk; lsi10@cam.ac.uk.

Abstract:

The ability of the cell to rapidly partition biomolecules into biomolecular condensates has been linked to a diverse range of cellular functions. To understand how the structural and dynamic attributes of these biomolecular condensates are linked with their biological roles, it is important to develop synthetic systems that allow systematic tuning of their physicochemical properties. Here, we describe the design and characterisation of a phase-separating consensus-designed tetratricopeptide repeat (CTPR) protein platform to make precise and predictable changes that both tune the condensate propensity and endow the condensates with new functions. We show that by incorporating short binding motifs into the CTPR domain we can specifically recruit target proteins to the dense phase of the condensate. The extent of target recruitment correlates with the binding affinity of the motifs. The relationship between CTPR domain modulation and condensate propensity observed experimentally is recapitulated by *in silico* modelling, giving us the ability to apply *in silico* methods to rationally design novel functional biomolecular condensates.

Significance:

This study designs tunable synthetic condensates by fusing modular consensus-designed tetratricopeptide repeat (CTPR) proteins to intrinsically disordered domains. Notably, changes to the CTPR scaffold allow for manipulation of the range of stability and dynamicity of the condensates, as well as recruitment of specific client proteins, *in vitro*. Moreover, these attributes are supported by *in silico* analysis performed at physiologically relevant conditions demonstrating the amenability of the system to rational design.

Introduction

Biomolecular condensates, also known as membraneless cellular compartments, assemble specific proteins and other biomolecules in a high local concentration through phase separation (1). The stability and material properties of biomolecular condensates can be modulated by intrinsic parameters, such as the amino acid sequence, post-translational modifications, and the stoichiometry and composition, as well as through the environmental conditions, including heat, pH stress and salt (2-5). Biomolecular condensates exhibit a wide range of material properties, comprising low-to-high viscosity liquids, semi-fluid gels, glasses, and solid aggregates. There is growing evidence indicating that the diverse material properties of condensates are significant to their functionality (6).

Many naturally occurring RNA-binding proteins that form biomolecular condensates combine well-folded globular domains and intrinsically disordered regions (IDRs) with sequences of low complexity (LC). *In vitro* experiments have revealed that LC IDRs are strong positive and negative modulators of the stability of biomolecular condensates, due to their ability to form weak, multivalent interactions (7-11). Nonetheless, folded domains, which can establish specific interactions with other biomolecules, are also important regulators of biomolecular condensate properties (8, 10).

Recent work has focused on understanding how condensate material properties impact biological functions, and several engineered condensate systems have been developed to control the material properties in a tuneable way. These studies have predominantly focused on using IDR regions to drive phase-separation by appending these molecular adhesives to enzymes to characterise their function within the dense phase of a condensate. Such systems have included the use of the hydrophobic protein elastin (1, 12), short cationic tags (13, 14) and IDRs derived from naturally occurring phase-separating proteins (15-17). Along with using

IDRs to drive phase separation, the role of multivalent protein scaffolds and oligomerisation domains to create engineered condensates have also been explored (5, 18-20).

In this work, we fuse LC domains (LCDs) to folded consensus-designed tetratricopeptide repeat (CTPR) proteins to create a library of chimeric molecules in order to test how changes to the folded domain affect condensate properties. Two types of LCDs were added to the CTPRs, namely LCD1 and LCD2 (16) derived from the sequences of DEAD-box ATPase proteins, Dhh1 and Dpb1, that have been shown to undergo phase separation associated with the formation of processing bodies in yeast (21, 22). These sequences have been used as molecular adhesives able to drive *in vitro* phase-separation of different globular proteins to create novel microreactors (16, 23). Condensate formation for these systems can be controlled through protein concentration, pH and salt concentrations, with the LCD2 molecular adhesives showing a broader range of LLPS-inducing conditions (16). For folded domains, we used CTPRs based on tetratricopeptide repeats (TPRs) that unlike globular proteins, adopt an elongated molecular shape comprised of tandem arrays of between 3 and 16 repeats, each composed of a 34 residue helix-loop-helix motif (24). Naturally occurring TPRs have the ability to mediate protein–protein interactions via a range of target recognition modes making them useful for a number of processes, including transcriptional regulation, protein folding, cell-cycle control and neurogenesis (24-26). Main and co-workers designed CTPRs through statistical analysis of amino acid propensity at each position of the TPR, and they showed that these new proteins were ultra-stable despite the absence of disulphide bonds (27). This repetitive modular architecture and high thermodynamic stability make CTPRs a useful tool for protein engineering, where a rational design approach can be used to endow them with additional functionality including specific client recruitment via insertion of single and multiple copies of short linear binding motifs (SLiMs), and arrangement in diverse, precise, and predictable geometries for multivalent and multi-functional display (28-30). Using computational simulation and *in vitro* experiments, we demonstrate that introducing different characteristics to the CTPR domains can alter the condensate propensity in a rational way whilst also enabling specific protein recruitment to the condensates. Strikingly, the *in silico* analysis of the LCD-CTPR proteins correlates with our *in vitro* findings demonstrating the utility of using this approach for further rational design of engineered biomolecular condensates.

Results:

Design of the LCD-CTPR system:

Based on the molecular adhesive strategy developed by Arosio and co-workers (16), LCD1 and LCD2 (N- and C-terminal sequences, Table S3) were grafted onto our CTPR domains. Functionalisation of the CTPRs is achieved by incorporating SLiMs into one or more of the short (4-residue) inter-repeat loops, where even 58 residue-long inserts can be incorporated without causing major changes to the overall stability of the CTPR (31). Using this ability to modify CTPR inter-repeat loops, we incorporated a short tetracysteine motif (CCGPCC) between repeats 1 and 2. This motif binds biarsenical dyes like fluorescein arsenical hairpin binder (FlAsH) (32). When unbound, the reagent is not fluorescent, but upon interaction with the tetracysteine (TC) motif, a significant increase in fluorescence quantum yield is observed (33). The TC-tag/biarsenical dye system is advantageous due to the small size reducing the chance of tag interference with the properties of the protein of interest, and it lends itself to the visualisation of target proteins inside live cells (14, 34-36). To test whether changes to the CTPR properties can affect the propensity of condensate formation, we created the following CTPR variants: 1) Comprised of either two, three, or four repeats (LCD2-CTPR2, LCD2-CTPR3, LCD2-CTPR4), 2) presenting altered CTPR surface residues, changing all lysine residues to arginine residues (LCD2-R-CTPR3), and 3) adding a SLiM that binds specifically to the protein LC3, which is involved in sequestering substrates to the autophagy pathway

(37). The LC3-interacting region (LIR) was inserted in one inter-repeat loop to allow us to determine if LCD-CTPRs can maintain the specific binding capabilities of the original CTPRs (LCD2-CTPR4-FUNDC1 and LCD2-CTPR4-ATG13) (Fig. 1).

LCD-CTPR proteins have high thermodynamic stability and form biomolecular condensates:

Initially, a three-repeat CTPR protein flanked by N- and C-terminal LCD1 sequences (LCD1-CTPR3) was expressed and purified. The resulting protein could be labelled with FlAsH and, based on mass spectrometry, was of high purity (Fig. S1*A*, Table S1). However, unlike the LCD1-GFP protein, where high salt concentrations maintained the solubility of the protein (16), the LCD1-CTPR3 system required the addition of 1M urea and 1M NaCl to avoid phase-separation during purification. Using fluorescence microscopy, spherical droplets were observed and formed more readily for LCD1-CTPR3 compared to LCD1-GFP under identical solution conditions (Fig. S2). Condensate induction for LCD1-CTPR3 required dialysis into appropriate buffers, therefore the CTPR variants flanked by N- and C-terminal LCD2 sequences were made, as these were reported to be more condensate-forming than the LCD1 system (16) (Fig. S1*B*, Table S1). The LCD2-CTPR3 readily formed spherical droplets even at high salt concentrations and 1 M urea (data not shown). Having confirmed that all the CTPR variants under investigation maintained high stability towards chemical denaturation (Fig. S3, Table S2), with even the smallest, CTPR2 variant having a $[GdnHCl]_{50}$ = 2.58 ± 0.22 M and remaining fully folded in 3M urea (Fig. S4), we introduced 3M urea into the purification buffer to maintain the solubility of the LCD2-CTPRs. With this optimised purification method, induced condensate formation using standard dilution methods was achieved (38) and used to systematically monitor the condensate-forming properties of the LCD2-CTPR variants. In all LCD2-CTPR variants, we could induce the formation of spherical droplets with different salt concentrations, and these droplets display fusion when imaged in real-time (Fig. 2*A-D*, Fig. S5). Under moderate-to-high salt concentrations, we observe that these condensates can be dissolved with 1,6-hexanediol (Fig. 2*C*, Fig. S6). Using this library of variants, we further compared how changes to the CTPR domain can tune condensate propensity and their physico-chemical properties using *in vitro* techniques and *in silico* simulations.

Phase diagrams of LCD2-CTPRs show that condensate propensity can be systematically modulated:

To obtain the phase behaviour across variants under a broad range of solution conditions, we employed PhaseScan, a high-throughput imaging analysis technique. PhaseScan is a combinatorial droplet microfluidic platform designed for rapid, high-resolution, automated analysis of phase separation (39). We titrated the FlAsH-labelled protein solution across a range of NaCl and urea concentrations (Fig. 3). The dotted line in the phase diagram delineates the phase boundary: above this boundary the protein solution forms a single, dilute phase, and below this boundary the protein solution forms two coexisting phases (dilute and dense). The LCD2-CTPR variants show clear differences in phase behaviour. First, the propensity to form condensates increases with increasing number of CTPR modules (Fig. 3, *A, B, D*), as demonstrated by LCD2-CTPR4 forming condensates in higher urea and salt conditions than LCD2-CTPR3 and LCD2-CTPR2. Second, substituting the surface lysine residues for arginine residues greatly increases the condensate propensity (Fig. 3, *B, C*). Consistent with the *in vitro* and *in silico* findings of other research groups, our observations underline the positive role arginine residues can play in phase separation propensity of proteins (4, 11, 40-43). Lastly, PhaseScan analysis of LCD2-CTPR4 proteins containing two different LC3-binding loops shows that introducing these loops between repeats 3 and 4 decreases the condensate-forming propensity. Additionally, LCD2-CTPR4-FUNDC1 presents

a slightly higher condensate propensity than LCD2-CTPR4-ATG13 (Fig. 3, *E, F*). Overall, it is clear that the CTPR folded domain is a critical regulator of the stability of LCD2-CTPR condensates.

In silico simulations reveal molecular principles regulating LCD-CTPR condensate formation:

To complement our experimental findings, *in silico* simulations of different LCD2-CTPR variants were performed (sequences in Table S3). These simulations allow us to probe the underlying molecular mechanisms that explain the differential phase separation propensity of the LCD-CTPR variants. Initial simulations to compare LCD1-CTPR3 with LCD1-adenylate kinase, a system extensively studied *in vitro* by Arosio and co-workers (16), found that CTPR3 clearly enhanced condensate propensity (Fig. S7) and this was similar to the *in vitro* comparison with LCD1-GFP (Fig. S2).

The simulations enable contextualisation of the *in vitro* analysis by testing the same constructs in physiological salt and pH conditions. We used our Mpipi sequence-dependent residue-resolution coarse-grained model for phase-separating proteins, which has been shown to recapitulate experimental temperature-vs-density phase diagrams with near-quantitative accuracy (Fig. 4*A*-*B*) (44). During the simulation, the secondary structure of each separate CTPR module is rigidified. In contrast, the IDRs, such as the flanking LCDs and the short loops between CTPR modules are simulated as fully flexible chains. To compare the stability of the various LCD2-CTPR variants, we perform Direct Coexistence simulations at various temperatures. In a Direct Coexistence simulation, the condensed phase and diluted phase are simulated in the same elongated box, separated by an interface (Fig. 4*C*) (45-49). For each temperature probed in the simulations if, after equilibration, we detect two phases in coexistence, we measure the density of each phase and plot them versus the simulation temperature to construct coexistence curves (or binodals) in the temperature–density plane (Fig. 4*D*). Our simulations allow us to predict the Upper Critical Solution Temperatures (UCST) of each of our different constructs. That is, the maximum temperature threshold beyond which phase separation is no longer observed is the UCST or critical temperature, T_c, (Fig. 5 *A*–*F*). In our simulations, LCD2-CTPR2 presented the lowest critical temperature of the set, indicating its general lower propensity for droplet formation (Fig. 5*A*). LCD2-CTPR3 is positioned in the intermediate range, while the highest critical temperature was observed for LCD2-CTPR4, suggesting the condensates with the strongest thermodynamic stability (Fig. 5*A*). These results are in line with the PhaseScan analysis (Fig. 3*A, B, D*).

One key hypothesis is that increasing the number of globular repeats within the LCD2-CTPR variants enhances the valency of intermolecular interactions (50-52). This heightened multivalency is likely to contribute to the formation of a more stable and dense phase. To test this hypothesis, the frequency of intermolecular contacts for each LCD2-CTPR construct was computed and analyzed. The increased repeat number correlates with increased molecular valency (Fig. 5*B*). Increased valency gives rise to a more densely connected liquid and, hence, enhanced thermodynamic stability of the condensate (53). An analysis of the most frequent inter-molecular interactions inside the condensate reveals that LCD–LCD interactions contribute most strongly to driving phase-separation of the LCD–CTPR solutions, followed by contacts between the LCDs and the CTPRs (Fig. 5). The latter confirms a synergistic relationship between CTPR repeats and LCDs in promoting phase separation. In addition to the role of multivalency, we hypothesise that the globular domains in the CTPR variants facilitate a more efficient packing density of the proteins within the condensed phase, enhancing its stability. This hypothesis is supported by the large increase in the critical solution

temperature of the LCD2-CTPR3 when the surface lysine residues are substituted for arginine residues to yield the LCD2-R-CTPR3 variant (Fig. 5*F*).

Finally, the differences in condensate critical solution temperature when introducing two different LIR motifs (FUNDC1 and ATG13) between repeats 3 and 4 were analysed. In accordance with the Phasescan results, FUNDC1 exhibits an increased propensity for condensate formation when compared to ATG13. Analysis of the protein sequence (Fig. 5*D*), shows that FUNDC1 is a highly negatively charged linker. Given that the construct LCD2-CTPR4 has an excess of positive charges, a negatively charged linker can drive the resulting condensate into electroneutrality and thus, make the dense phase more energetically favourable (54, 55). Although these linker regions are relatively short (constituting only 5% of the full sequence), contact analysis of these constructs revealed that the FUNDC1 linker can facilitate intermolecular interactions with the globular domains, both with itself and between them.

This comparative analysis between *in vitro* phase diagrams and *in silico* condensate propensities confirms that the structure and chemical makeup of the CTPR domains highly influences condensate propensity in diverse solution conditions. This alignment of experimental and computational results confirms that both techniques can be used complementary to predict and influence rational designs into the LCD-CTPR systems.

CTPR attributes can alter salt-dependent condensate formation:

To futher characterise the tuneability of the LCD2-CTPR condensates, we looked at salt dependency, critical saturation concentrations (C_{sat}) and reversibility measured by turbidity assays. First, to quantify the effect of salt concentration, a sedimentation assay was used to determine the dilute-to-dense phase ratios. The FlAsH-labelled LCD2-CTPR variants were diluted 10-fold into appropriate buffers to induce condensate formation at varying salt concentrations. After visualising the appearance of condensates, the samples were centrifuged and the protein in the dilute phase was quantified using SDS-PAGE analysis. In all cases, condensates can be induced even with high NaCl concentrations (Fig. 6); however, there are significant differences in propensity of dense-phase formation between the variants. In line with our previous observations, when comparing different repeat lengths, LCD2-CTPR2 shows the lowest propensity, followed by LCD2-CTPR3, and LCD2-CTPR4 being the most condensate-prone. Interestingly, an increase in dense phase with increasing salt concentration is observed for LCD2-CTPR2, whereas for LCD2-CTPR4 the amount of dense phase decreases with increasing salt concentration (Fig. 6, *A, D*). The moderate phase-separation observed with LCD2-CTPR3 increases dramatically when lysine-to-arginine substitutions were made (LCD2-R-CTPR3) (Fig. 6*C*). Consistent with the *in silico* findings that suggest that the overall negatively charge of the FUNDC1 loop can balance the positive charge of the LCD2-CTPR4 to favour phase separation, increased salt concentration dramatically decreases LCD2-CTPR4-FUNDC1 condensate formation (Fig. 6*E*), whereas LCD2-CTPR4-ATG13 has similar properties to LCD2-CTPR4 (Fig. 6 *D, F*).

Using sedimentation assays, the critical saturation concentration for each of the LCD2-CTPR proteins was determined. The critical saturation concentration marks the threshold above which the system demixes into a dilute and a dense phase at a given set of conditions (15). Multivalency plays a pivotal role in determining C_{sat}, as increasing valency notably lowers the C_{sat} threshold, initiating phase separation at lower total protein concentrations (38). The dilute phase concentration across a range of protein concentrations with a fixed buffer condition (50 mM Tris-HCl, pH 7, 300 mM urea, 1 M NaCl) was plotted against total protein concentration and shown to plateau once it surpassed the C_{sat} value (Fig. S8). The C_{sat} values calculated from this are shown in Table 1.

When comparing the different repeat lengths, the results match well with our other results: LCD2-CTPR2 and LCD2-CTPR3 have statistically equivalent C_{sat} values, whereas LCD2-CTPR4 has a much lower C_{sat} value and is therefore more condensate-prone. Similarly, LCD2-R-CTPR3, which appears to have the highest condensate-propensity presents the lowest C_{sat} value. Interestingly, the introduction of the LIR sequences into the inter-repeat loops appears to decrease condensate formation at high salt concentrations, which agrees with the sedimentation assay; the ATG13 sequence results in a small increase in the C_{sat} value, whereas the FUNDC1 sequence increases the C_{sat} value by an order of magnitude.

CTPRs impart differences to thermal reversibility of condensates:

The physicochemical attributes of condensates likely influence their functionality due to their impact on molecular mobility and reactivity. The LCD2-proteins exhibit UCST behaviour as shown previously (16) and by our simulations. Building upon this knowledge, we evaluated the reversibility of these condensates under varying salt concentrations using a turbidity assay (56). To compare this systematically, the condensates were induced in either the 100 mM NaCl or 550 mM NaCl conditions at 20°C and heated to 70°C, the incremental change from 70 to 20°C (D1) was recorded at $A_{500\ nm}$ followed by a 20 to 70°C increase (I2) and a second incremental decrease (D2).

For all of our LCD2-CTPR variants, the condensates induced in low salt concentrations (100 mM NaCl) result in spherical condensates and turbid solutions at 20°C; however, this turbidity only decreases partially with increased temperatures to 70°C. For LCD2-CTPR2, LCD2-CTPR3, and LCD2-CTPR4-FUNDC1, we observe only small changes in turbidity throughout the temperature cycles (Fig. 7 *A, B, E,* black traces), suggesting a dynamically arrested state (57). Two exceptions to this trend are LCD2-CTPR4 and LCD2-R-CTPR3; both of these samples show temperature-dependent changes in turbidity in low salt conditions; moreover, upon repeating the increasing temperature ramp (I2) and decreasing temperature ramp (D2) of these samples, some reversibility of the turbidity is observed. LCD2-CTPR4-ATG13 has an intermediate behaviour between these two trends. With higher NaCl concentrations (550 mM), reversibility in the turbidity traces for all variants is observed.

LCD2-CTPR can be readily functionalised for target recruitment into condensates:

The CTPRs allow us to easily functionalise the condensates to recruit specific client proteins. Target-binding motifs can be grafted onto the inter-repeat loops, and here we used LC3 interacting regions (LIRs) that can bind to LC3, a member of the ATG8 family of proteins involved in target recognition for the selective autophagy pathway (37). Two different LIRs were inserted into the LCD2-CTPR4 protein: the FUNDC1 loop was derived from a protein present in the outer mitochondrial membrane that acts as a receptor for mitophagy in mammals (58) and the ATG13 loop was derived from an autophagy-related protein that regulates phagosome formation (59). These LIRs have different LC3-binding affinities, with the LIR from FUNDC1 possessing higher affinity than that from ATG13 (60, 61). This result was confirmed to still be the case when these loops were grafted into the CTPR4 scaffold without LCD regions, in a comparative binding study monitored by western blotting (Fig. S9).

To study client protein recruitment, condensate formation with the FlAsH-labelled LCD2-CTPR4 variants was induced. Once the condensates were formed, ROX-labelled LC3 protein was added and specific recruitment was monitored by fluorescence microscopy. No co-localisation of the FlAsH-labelled condensates with ROX-labelled LC3 was observed when the LIR-motif is absent (LCD2-CTPR4). The presence of either the FUNDC1 or ATG13

sequences resulted in co-localisation of LC3 within the condensates (Fig. 8*A*). To quantify the extent of this co-localisation, the partition coefficient of ROX-labelled LC3 recruitment to unlabelled LCD-CTPR4 variant condensates was measured (Fig. 8*B*). This suggests that the FUNDC1 motif resulted in more LC3 co-localisation than the ATG13 motif, corresponding to the variation in reported binding affinities to LC3. This relationship was recapitulated using a sedimentation assay to determine the proportions of LCD2-CTPRs and LC3 in the dilute and dense phases (Fig. 8*C*). These results further demonstrate that not only can we recruit client proteins to our engineered condensates using target-binding loops, but also that the extent of this recruitment can be tuned depending on the binding affinity of the loop used.

Discussion:

To understand how condensates impact biological functions it is useful to develop synthetic condensates whose properties can be tuned in a controllable and predictable way. Additionally, genetically encodable systems are advantageous so that they can be introduced into the cellular environment. LCD sequences have been used to generate synthetic condensates in bacterial, yeast and mammalian cell systems (14, 17, 57, 62, 63). The LCDs can be altered to tune the critical concentrations for phase separation by changes in valency (17). In addition, the LCDs can be rationally designed to give rise to more or less viscous liquid condensates or sustain gel-like attributes (1, 18, 62). Systems like PopZ and multivalent *de novo* coiled coils drive phase separation through oligomerisation and can be tuned by altering weak protein-protein interactions, demonstrating applicability in mammalian and bacteria cell systems, respectively (19, 20). Recently, Arosio and co-workers have described the use of LCD-fusion to promote the phase-separation of globular proteins whilst maintaining the enzymatic activity of such globular proteins (16). Here we show that fusing LCDs to both termini of a CTPR, leads to a phase-separating system. Notably, the stability of the LCD-CTPR condensates is demonstrated to be strongly influenced by the unique architecture of the CTPR and tuneable by engineering of this folded region. In addition, the modular, stable CTPR scaffold opens up the ability to incorporate target-binding peptides for specific and controlled recruitment of client molecules into condensates.

Since we have conserved the LCDs in the LCD2-CTPR systems, we can deduce the impact of changes in the CTPR on the phase behaviour of the LCD2-CTPR solutions in a rational manner. Using a complementary set of experimental and *in silico* analyses, we have demonstrated that the CTPR component strongly influences condensate formation and physicochemical attributes. Computational analysis indicates that synergistic interactions between CTPR modules and the LCD2 regions are pivotal in promoting phase separation, and with a repeat length of four, there is a strong preference for condensate formation. Specifically, our simulations reveal that the the stability of LCD2-CTPR condensates is sustained by numerous associative interactions between the LCDs and also contributed by some contacts between the LCDs and CTPRs. This prediction is corroborated by our *in vitro* analyses. Interestingly, when we replace lysine residues within the CTPRs modules with arginine residues, we observe a large increase in condensate propensity both *in silico* and *in vitro*. Recent efforts to define the impact of specific amino acids within IDRs that promote or reduce phase separation has demonstrated that saturation of phase separation is inversely proportional to the product of the numbers of arginine and tyrosine residues (11); despite being in the folded CTPR rather than LCD region, it appears that arginine residues can promote phase separation in our system also. Finally, we find that incorporating short functional peptides into the loop regions of the CTPRs alters condensate propensity tuned by their amino acid composition.

At high salt concentrations, all LCD2-CTPR variants show the ability to fuse, but in low-salt conditions, most of our LCD2-CTPRs exhibit minimal decreases in turbidity both in temperature-mediated and 1,6-hexanediol assays, suggesting that they are in a dynamically arrested state. The exceptions to this are LCD2-R-CTPR3 and LCD2-CTPR4 which have the highest propensity to form biomolecular condensates as shown by PhaseScan and *in silico* analysis. They show a significant degree of reversibility even in low-salt conditions. These findings indicate that by changing the chemical-make up of the CTPRs, we can design an array of LCD-CTPR condensates that exhibit liquid-like or gel-like material properties.

The ability to functionalise engineered biomolecular condensates in a facile and versatile manner to enable specific recruitment of target proteins into the dense phase to enhance enzymatic reactions or modulate protein-protein interactions is highly desirable. A variety of strategies to enrich target proteins into engineered condensates have been explored, including short coil-coiled tags and nanobody fusions for GFP or appending directly to a protein of interest to target its partitioning to the condensate phase (6, 14, 16-20). A unique feature of the modular CTPR scaffold is that target-binding specificity, valency and affinity can be readily manipulated, enabling recruitment of one or multiple proteins through engrafting of loops into the inter-repeat loops (28, 29). As proof-of-concept, we show that incorporating two LIR-motifs, FUNDC1 and ATG13, results in specific recruitment of LC3 to the dense phase of the condensates. Importantly also, the binding affinity determines the extent of target recruitment.

The relationship between CTPR modulation and condensate-propensity can be rationally understood by comparing *in silico* modelling with *in vitro* experiments. Not only have we demonstrated that there is a tuneable synergy within the LCD-CTPRs, but the *in silico* analysis of the phase behaviour of these systems closely mirrors our *in vitro* findings. Such a multidisciplinary approach gives us the ability to explore future designs *in silico* in a rational and predictive manner to create novel biomolecular condensates with the ability to specifically recruit target proteins of interest.

Materials and Methods:

LCD-CTPR plasmid preparation and protein purification:

LCD-CTPR constructs were cloned using gBlock oligos (Integrated DNA technologies) into the pRSET-B vector with FastDigest BamHI and HindIII restriction enzymes (ThermoFisher Scientific (UK) Ltd.) and Anza T4 DNA ligase master mix (Invitrogen). The amino acid sequences of the proteins are listed in Table S3. Recombinant proteins were expressed in chemically competent C41(DE3) *E. coli* cells (Lucigen). Upon reaching OD_{600nm} of 0.6-0.8, bacteria cultures (1L) were induced with isopropyl D-thiogalactopyranoside (IPTG, 0.5 mM, PanReac AppliChem) and incubated (18 h, 20°C, 200 rpm). Cell pellets were collected by centrifugation and resuspended in lysis buffer (50 mM Tris (pH 8.5), 1 M NaCl, 0.5 mM TCEP with 1 M urea (LCD1-CTPRs) or 3 M urea (LCD2-CTPRs), 1 μg/mL DNase I, 1 cOmplete EDTA-free protease inhibitor tablet (Roche Diagnostics GmbH) for each 50 mL of lysis buffer. Cells were lysed using an EmulsiFlex C5 homogeniser (Avestin) (3-4X, 10000-15000 psi) and centrifuged. The supernatants were purified using HisTrap HP columns (1 mL, Cytiva Ltd.) with Buffer A (50 mM Tris (pH 8.5), 1 M NaCl, 0.5 mM TCEP with 1 M urea (LCD1-CTPRs) or 3 M urea (LCD2-CTPRs)) and a linear gradient (8-60% Buffer A + 0.5 M imidazole, 15 column volumes (CV)). LC3 was purified with Buffer B (50 mM sodium phosphate buffer (pH 8.0), 150 mM NaCl (0.5 mM TCEP included for cysteine-variant)) with a linear gradient (8-80% Buffer B + 0.5 M imidazole, 15 CV). LC3 was dialysed (18h, room temperature (RT)) against 50 mM sodium phosphate (pH 8), 150 mM NaCl. LCD-CTPRs were flash frozen in purification buffer and stored at -80°C. Protein purity was confirmed using SDS-PAGE (Fig. S1) and electrospray

ionisation mass spectrometry (Table S1) performed on a Xevo G2 mass spectrometer with data analyzed using MassLynx software (Waters UK) (Yusuf Hamied Department of Chemistry, University of Cambridge, UK).

FlAsH-EDT$_2$ Labelling:

LCD-CTPRs (50 μM) were thawed, centrifuged (2 min, 15000 xg) to remove aggregates and protein was labelled with FlAsH-EDT$_2$ (100 μM, Cayman Chemical). Excess dye was removed using a Zeba 7kDa MWCO Spin Desalting Column (Thermofisher) in Buffer A. To determine the concentration of FlAsH-labelled proteins, a correction factor was calculated by measuring the A_{280nm} of the dye divided by the A_{max} of the dye. For FlAsH-EDT$_2$ this was determined to be 0.271. Protein concentrations were determined using Eq. 1 where ε_{280nm} was calculated using Expasy ProtParam Tool based on the primary sequence of the protein.

Eq. 1: Protein concentration (M) = (A280nm – (A511nm x CF))/ε_{280nm}

ROX-maleimide labelling of LC3:

LC3 with an N-terminal cysteine was labelled by reacting with a 10x molar excess of TCEP, then adding 2–5x molar excess of ROX-maleimide (Biotium) and incubated (ON, 4°C). 1 mM DTT was added, and excess dye was removed with a 7kDa MWCO Zeba Spin Desalting Column. Protein concentration and labelling efficiency were calculated using Equations 2 and 3:

Eq. 2: Protein Concentration (M) = $(A_{280} - (A_{591} \times 0.49) / \varepsilon_{280}$

Eq. 3: Labelling Efficiency (%) = $(100 \times A_{591}) / (93{,}000 \times \text{Protein Concentration (M)})$

Fluorescence Microscopy:

Imaging of condensate formation was done on a Zeiss AxioVert A1 microscope with a Colibri 7 light source (Carl Zeiss Ltd.), using a LD PN 40x/0.6 Corr objective or an EC PN 63x/1.25 oil objective. Images were captured with a Axiocam 305 mono camera and processed using ImageJ.

Real-time Microscopy of condensate fusion:

Imaging of condensate formation was done on a Leica TCS SP5 DMI6000 B inverted confocal microscope (Cambridge Advanced Imaging Centre, University of Cambridge, UK) with a 100 x 1.4 numerical aperture (NA) HCX PL APO CS total internal reflection fluorescence oil immersion lens with a 512-by-512-pixel photon multiplier tubes (PMTs) as detectors. 2000 frames on each field of view were collected for 30 min.

PhaseScan:

Polydimethylsiloxane (PDMS) devices (Sylgard 184 kit; Dow Corning) were fabricated by casting PDMS on a master wafer made from SU-8 3050 photoresist (A-Gas Electronic Materials Limited) via conventional soft-photolithography methods. Phase diagrams were generated using PhaseScan, a high-throughput phase scanning technology (39).

Briefly, stock solutions of different components were mixed at various mass ratios and encapsulated in water-in-oil droplets, individual microenvironments, using a microfluidic device. Fluorescence images of droplets in the observation chamber of the microfluidic device were taken continuously using an epifluorescence microscope (Cairn Research) equipped with 10 × objective (Nikon CFI Plan Fluor 10×, NA 0.3) and high-sensitivity camera (Kinetix, Teledyne Photometrics). Images were analysed via an automated costom Python-based image analysis script (39) to detect condensates and to determine the concentrations of different components within individual microfluidic droplets. The data was plotted as a colour-

coded scatter plot, where each point represents individual microfluidic droplets and colour represents the phase separation probability (1 or dark red – phase separated, 0 or dark blue – mixed) assigned to the droplet. This probability was calculated by averaging the phase state assignment or neighbouring droplets. The radius of the neighbourhood was defined as a percentrage of data rangewith an overlayed colour-coded heat map showing the estimated phase separation probability.

Direct Coexistence MD Simulations:

Direct coexistence simulations were performed in a slab geometry, characterised by a simulation box elongated in one of the cartesian coordinates. To coarse-grain the protein system we employed a one-bead-per-aminoacid model, that maps each amino acid of the protein into a single distinct bead. Each amino acid has its particular force field parameters, mass, and charge, to ensure single amino acid resolution. The force field chosen was the Mpipi model, a model optimized for the study of phase-separating systems (44). Each CTPR structure was computed through AlphaFold, optimized by hand via atomistic modeling on Pymol and Gromacs, and finally, systematically cross-validated with published structures and predicted minimized structures. A coarse-graining protocol was then applied, taking the atomistic structure and mapping each bead into the alpha carbons of each protein residue. Globular regions were maintained rigid during the simulation to ensure the conservation of the secondary structure, while intrinsically disordered regions were modeled in a flexible, polymer-like manner.

The simulations were performed using the LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) (64). Each phase diagram was computed at several temperature settings, with each set subjected to a thermal equilibrium period of up to 500 ns to guarantee convergence of our results. Each simulation was performed in an NVT ensemble, with a constant number of proteins (N=140) and constant temperature. The specific number of protein copies was carefully chosen to balance computational efficiency and mitigate the impact of finite-size effects.

Phase Diagram Fitting:

To estimate the critical temperature, we use the law of coexisting densities (Equation 4) (65):

Eq. 4 $$(\rho_h(T) - \rho_l(T))^{3.06} = d(1 - T/T_C),$$

where ρ_h and ρ_l are the densities of the protein-rich and protein-depleted phases respectively, 3.06 is the Ising critical exponent, T_C is the critical temperature, and d is a fitting parameter. The law of rectilinear diameter also holds for the estimation of the critical density (Equation 5):

Eq. 5 $$\boldsymbol{\rho_h(T) + \rho_l(T) = 2\rho_c + 2A(T - T_c)}$$

where ρ_c is the critical density, and A is a fitting parameter.

Contact Analysis:

The number of contacts between the domains of the CTPRs within the slab was analysed using the Visual Molecular Dynamics (VMD) software (66). Each protein construct was analyzed firstly by considering each of its distinct domains separately, i.e., the N and C terminal LCD domains, the CTPR repeats, and the designable linkers between CTPRs. For each simulation time frame, we then computed the number of amino acids of a different domain that were at a distance smaller than 7Å from the domain considered. This number, which we refer to as raw contact number, was averaged across frames. Finally, the raw contacts were then normalised by the total number of contacts, giving a resulting percentage. For the specific amino acid sequence of each protein domain, see Table S3. The LCD2

regions were as follows: LCD2-CTPR2: 1-173, 267-334; LCD2-CTPR3 & LCD2-R-CTPR3: 1-173, 301-369; LCD2-CTPR4: 1-173, 320-388; LCD2-CTPR4-ATG13: 1-173, 339-411; LCD2-CTPR4-FUNDC1: 1-173, 337-409. The linker regions were: LCD2-CTPR4-ATG13: 282-308; LCD2-CTPR4-FUNDC1: 282-306.

Sedimentation and C_{sat} Assays:

Post FlAsH-labelling, protein samples were diluted in Buffer A to 20 μM. A 10-fold dilution into 50 mM Tris (pH 7.0), varied NaCl, 0.5 mM TCEP buffers, or Buffer A was made. Samples were centrifuged (5 min, 20,000 xg), and the supernatant was mixed (1:1) with 50 mM Tris (pH 8.5), 1 M NaCl, 10 M urea, and separated by SDS-PAGE. Band densities were analysed using an Odyssey Fc Imager (LI-COR) and quantified in Image Studio Lite 5.2. Phase proportions were derived from band densities relative to a 100% dilute-phase control. To obtain saturation concentration, proteins were diluted to different concentrations in 50 mM Tris, pH 7, 300 mM urea, 1 M NaCl, 0.5 mM TCEP then centrifuged (5 min, 20,000 xg) to sediment the dense-phase. The supernatant was analysed as above. Test condition protein concentrations were deduced using linear regression on control conditions, with GraphPad Prism 10.0 employed for plotting and fitting to the model: $y = if (mx < c, mx, c)$, where x is total protein concentration, m an arbitrary constant gradient, and c the C_{sat} value.

Turbidity Assays:

Protein samples were thawed, centrifuged to remove aggregates, buffer exchanged into Buffer A using a Zeba 7kDa MWCO Spin Desalting Column. The samples (20 μM) were diluted 10-fold with 50 mM Tris (pH 7.0), varied NaCl, 0.5 mM TCEP buffer and placed in quartz cuvettes (400 μL) and analysed in a FL 6500 Fluorescence Spectrometer (PerkinElmer LAS (UK) Ltd.). Samples were heated to 70°C then turbidity (excitation and emission set to 500 nm) was recorded as temperature decreased 70°C-to-25°C (D1), then increased 25°C-to-70°C (I2) and a final decrease from 70°C-to-25°C (D2). Readings were taken every 2°C, at 5°C/min, 1.5°C error and 400V PMT. Three independent experimental repeats were performed. For 1,6-hexanediol studies, concentrations of 0%, 4%, or 8% were used. Samples were placed in a 96-well plate (Corning 3881) and turbidity was measured on a FLUOstar Omega Microplate Reader (BMG Labtech). Data were plotted and analysed using GraphPad Prism 10.0.

LC3 co-localisation Assays:

FlAsH-labelled LCD2-CTPRs were diluted to 2 μM in 100 mM Tris (pH 7.0), 150 mM NaCl, 300 mM urea, 0.5 mM TCEP and ROX-LC3 was added (1:1 ratio). Solutions were imaged using a 63x oil objective and a variable exposure (4–18 ms) for the green channel and 50 ms exposure for the red channel. For partition coefficient assays, unlabelled LCD2-CTPRs (2 μM in 50 mM Tris (pH 7.0), 150 mM NaCl, 100 mM urea, 0.5 mM TCEP) and ROX-LC3 (1:1 ratio) were combined. To determine partition coefficients, fluorescence intensities were extracted using ImageJ, with an intensity threshold to distinguish between high-intensity condensates and low-intensity dilute-phase. The sedimentation assay was carried out as above, with the addition of ROX-LC3 (1:1 ratio) to the FlAsH-labelled LCD2-CTPRs prior to centrifugation. A control of 2 μM ROX-LC3 in 50 mM Tris (pH 7.0), 100 mM NaCl, 100 mM urea, 0.5 mM TCEP was used to determine phase proportions for LC3.

Acknowledgements:

The authors would like to thank Dr. Timothy Welsh for his help with initial PhaseScan experiments, Hao Chen, William Murphy and Chloe Burgess for their help with initial condensate experiments, and Dr. Luisa Capalbo, Dr. Lenka Faltova and Prof. Paolo Arosio for insightful discussions. T.L.C.N is supported by a Gates Cambridge International Scholarship, M.J.M. is supported by the Winton Programme for Physics of Sustainability for

doctoral funding and B.K.A. is supported by the Cambridge BBSRC-DTP programme. T.S. is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant MicroREvolution (agreement no. 101023060) and the Frances and Augustus Newman Foundation. This project made use of time on HPC granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim (http://hecbiosim.ac.uk), supported by EPSRC (grant no. EP/R029407/1). J.A.J. acknowledges research support from departmental start-up funds provided by the Department of Chemical and Biological Engineering and the Omenn–Darling Bioengineering Institute at Princeton University. This work was supported by a MRC Career Development Award (MR/W01632X/1, J.R.K.), the Royal Society (RGS\R1\231207, J.R.K.), the Rosetrees Trust (CF2\100013, J.R.K and L.S.I.), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 803326 to R.C.G.), the Engineering and Physical Sciences Research Council (EPSRC) (EP/X02332X/1 to J.H.M.), and an ERC grant DiProPhys (agreement ID 10100161, T.P.J.K.).

References:

1. Y. Dai, L. You, A. Chilkoti, Engineering synthetic biomolecular condensates. *Nat Rev Bioeng* 10.1038/s44222-023-00052-6, 1-15 (2023).
2. S. Alberti, A. Gladfelter, T. Mittag, Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419-434 (2019).
3. S. Alberti, A. A. Hyman, Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat Rev Mol Cell Biol* **22**, 196-213 (2021).
4. G. Krainer *et al.*, Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat Commun* **12**, 1085 (2021).
5. W. Peeples, M. K. Rosen, Mechanistic dissection of increased enzymatic rate in a phase-separated compartment. *Nat Chem Biol* **17**, 693-702 (2021).
6. Y. Dai *et al.*, Programmable synthetic biomolecular condensates for cellular control. *Nat Chem Biol* **19**, 518-528 (2023).
7. Z. Feng, X. Chen, X. Wu, M. Zhang, Formation of biological condensates via phase separation: Characteristics, analytical methods, and physiological implications. *J Biol Chem* **294**, 14823-14835 (2019).
8. Y. Lin, S. L. Currie, M. K. Rosen, Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem* **292**, 19110-19120 (2017).
9. L. Malinovska, S. Kroschwald, S. Alberti, Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochim Biophys Acta* **1834**, 918-931 (2013).
10. D. S. W. Protter *et al.*, Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep* **22**, 1401-1412 (2018).
11. J. Wang *et al.*, A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **174**, 688-699 e616 (2018).
12. F. G. Quiroz, A. Chilkoti, Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nat Mater* **14**, 1164-1171 (2015).
13. R. A. Kapelner, A. C. Obermeyer, Ionic polypeptide tags for protein phase separation. *Chem Sci* **10**, 2700-2707 (2019).
14. V. Yeong, J. W. Wang, J. M. Horn, A. C. Obermeyer, Intracellular phase separation of globular proteins facilitated by short cationic peptides. *Nat Commun* **13**, 7882 (2022).
15. S. Elbaum-Garfinkle *et al.*, The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A* **112**, 7189-7194 (2015).
16. L. Faltova, A. M. Kuffner, M. Hondele, K. Weis, P. Arosio, Multifunctional Protein Materials and Microreactors using Low Complexity Domains as Molecular Adhesives. *ACS Nano* **12**, 9991-9999 (2018).

17. B. S. Schuster *et al.*, Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. *Nat Commun* **9**, 2985 (2018).
18. M. V. Garabedian *et al.*, Protein Condensate Formation via Controlled Multimerization of Intrinsically Disordered Sequences. *Biochemistry* **61**, 2470-2481 (2022).
19. A. T. Hilditch *et al.*, Assembling membraneless organelles from de novo designed proteins. *Nat Chem* **16**, 89-97 (2024).
20. K. Lasker *et al.*, The material properties of a bacterial-derived biomolecular condensate tune biological function in natural and synthetic systems. *Nat Commun* **13**, 5643 (2022).
21. M. Hondele *et al.*, DEAD-box ATPases are global regulators of phase-separated organelles. *Nature* **573**, 144-148 (2019).
22. C. F. Mugler *et al.*, ATPase activity of the DEAD-box protein Dhh1 controls processing body formation. *Elife* **5** (2016).
23. M. Linsenmeier *et al.*, Dynamics of Synthetic Membraneless Organelles in Microfluidic Droplets. *Angew Chem Int Ed Engl* **58**, 14489-14494 (2019).
24. E. R. Main, K. Stott, S. E. Jackson, L. Regan, Local and long-range stability in tandemly arrayed tetratricopeptide repeats. *Proc Natl Acad Sci U S A* **102**, 5721-5726 (2005).
25. R. K. Allan, T. Ratajczak, Versatile TPR domains accommodate different modes of target protein recognition and function. *Cell Stress Chaperones* **16**, 353-367 (2011).
26. A. Perez-Riba, E. Komives, E. R. G. Main, L. S. Itzhaki, Decoupling a tandem-repeat protein: Impact of multiple loop insertions on a modular scaffold. *Sci Rep* **9**, 15439 (2019).
27. E. R. Main, Y. Xiong, M. J. Cocco, L. D'Andrea, L. Regan, Design of stable alpha-helical arrays from an idealized TPR motif. *Structure* **11**, 497-508 (2003).
28. A. Diamante *et al.*, Engineering mono- and multi-valent inhibitors on a modular scaffold. *Chem Sci* **12**, 880-895 (2021).
29. S. K. Madden, A. Perez-Riba, L. S. Itzhaki, Exploring new strategies for grafting binding peptides onto protein loops using a consensus-designed tetratricopeptide repeat scaffold. *Protein Sci* **28**, 738-745 (2019).
30. A. Perez-Riba, L. S. Itzhaki, The tetratricopeptide-repeat motif is a versatile platform that enables diverse modes of molecular recognition. *Curr Opin Struct Biol* **54**, 43-49 (2019).
31. J. F. Ripka, A. Perez-Riba, P. K. Chaturbedy, L. S. Itzhaki, Testing the length limit of loop grafting in a helical repeat protein. *Curr Res Struct Biol* **3**, 30-40 (2021).
32. B. R. Martin, B. N. Giepmans, S. R. Adams, R. Y. Tsien, Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat Biotechnol* **23**, 1308-1314 (2005).
33. A. Pomorski, A. Krezel, Biarsenical fluorescent probes for multifunctional site-specific modification of proteins applicable in life sciences: an overview and future outlook. *Metallomics* **12**, 1179-1207 (2020).
34. J. S. W. Ng *et al.*, Using Tetracysteine-Tagged TDP-43 with a Biarsenical Dye To Monitor Real-Time Trafficking in a Cell Model of Amyotrophic Lateral Sclerosis. *Biochemistry* **58**, 4086-4095 (2019).
35. S. Ray *et al.*, alpha-Synuclein aggregation nucleates through liquid-liquid phase separation. *Nat Chem* **12**, 705-716 (2020).
36. M. J. Roberti, C. W. Bertoncini, R. Klement, E. A. Jares-Erijman, T. M. Jovin, Fluorescence imaging of amyloid formation in living cells by a functional, tetracysteine-tagged alpha-synuclein. *Nat Methods* **4**, 345-351 (2007).
37. T. Johansen, T. Lamark, Selective Autophagy: ATG8 Family Proteins, LIR Motifs and Cargo Receptors. *J Mol Biol* **432**, 80-103 (2020).
38. S. Alberti *et al.*, A User's Guide for Phase Separation Assays with Purified Proteins. *J Mol Biol* **430**, 4806-4820 (2018).
39. W. E. Arter *et al.*, Biomolecular condensate phase diagrams with a combinatorial microdroplet platform. *Nat Commun* **13**, 7845 (2022).

40. A. Bremer *et al.*, Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. *Nat Chem* **14**, 196-207 (2022).
41. J. M. Choi, A. S. Holehouse, R. V. Pappu, Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annu Rev Biophys* **49**, 107-133 (2020).
42. M. J. G. Maristany, A. A.;Espinosa, J.R.; Huertas, J.; Collepardo-Guevara, R.; Joseph, J.A. , Decoding phase separation of prion-like domains through data-driven scaling laws. *bioRxiv* https://doi.org/10.1101/2023.06.14.543914 (2024).
43. E. W. Martin *et al.*, Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694-699 (2020).
44. J. A. Joseph *et al.*, Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nat Comput Sci* **1**, 732-743 (2021).
45. G. A. S. Chapela, G.; Thompson, S.M.; Rowlinson, J.S., Computer simulation of a gas-liquid surface. Part 1. *J.Chem.Soc., Faraday Trans. 2* **73**, 1133-1144 (1977).
46. J. R. Espinosa, E. Sanz, C. Valeriani, C. Vega, On fluid-solid direct coexistence simulations: the pseudo-hard sphere model. *J Chem Phys* **139**, 144502 (2013).
47. R. Garcia Fernandez, J. L. Abascal, C. Vega, The melting point of ice Ih for common water models calculated from direct coexistence of the solid-liquid interface. *J Chem Phys* **124**, 144506 (2006).
48. A. J. C. W. Ladd, L.V., Triple-point coexistence properties of the lennard-jones system. *Chem Phys Lett* **51**, 155-159 (1977).
49. A. C. L. Opitz, Molecular dynamics investigation of a free surface of liquid argon. *Phys Lett A* **47**, 439-440 (1974).
50. D. Bracha, M. T. Walls, C. P. Brangwynne, Probing and engineering liquid-phase organelles. *Nat Biotechnol* **37**, 1435-1445 (2019).
51. P. Li *et al.*, Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340 (2012).
52. R. V. Pappu, S. R. Cohen, F. Dar, M. Farag, M. Kar, Phase Transitions of Associative Biomacromolecules. *Chem Rev* **123**, 8945-8987 (2023).
53. J. R. Espinosa *et al.*, Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components. *Proc Natl Acad Sci U S A* **117**, 13238-13247 (2020).
54. S. Chen, Z. G. Wang, Driving force and pathway in polyelectrolyte complex coacervation. *Proc Natl Acad Sci U S A* **119**, e2209975119 (2022).
55. Y. Dai *et al.*, Global control of cellular physiology by biomolecular condensates through modulation of electrochemical equilibria. *bioRxiv* 10.1101/2023.10.19.563018 (2023).
56. J. H. Jung *et al.*, A prion-like domain in ELF3 functions as a thermosensor in Arabidopsis. *Nature* **585**, 256-260 (2020).
57. Y. Shin *et al.*, Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168**, 159-171 e114 (2017).
58. L. Liu *et al.*, Mitochondrial outer-membrane protein FUNDC1 mediates hypoxia-induced mitophagy in mammalian cells. *Nat Cell Biol* **14**, 177-185 (2012).
59. N. Hosokawa *et al.*, Nutrient-dependent mTORC1 association with the ULK1-Atg13-FIP200 complex required for autophagy. *Mol Biol Cell* **20**, 1981-1991 (2009).
60. Y. Kuang *et al.*, Structural basis for the phosphorylation of FUNDC1 LIR as a molecular switch of mitophagy. *Autophagy* **12**, 2363-2373 (2016).
61. H. Suzuki *et al.*, Structural basis of the autophagy-related LC3/Atg13 LIR complex: recognition and interaction mechanism. *Structure* **22**, 47-58 (2014).
62. M. V. Garabedian *et al.*, Designer membraneless organelles sequester native factors for control of cell behavior. *Nat Chem Biol* **17**, 998-1007 (2021).
63. S. P. Wei *et al.*, Formation and functionalization of membraneless compartments in Escherichia coli. *Nat Chem Biol* **16**, 1143-1148 (2020).
64. C. Grindon *et al.*, Large-scale molecular dynamics simulation of DNA: implementation and validation of the AMBER98 force field in LAMMPS. *Philos Trans A Math Phys Eng Sci* **362**, 1373-1386 (2004).
65. K. G. Weil, *Molecular Theory of Capillarity* (Clarendon Press, 1984).

66. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).

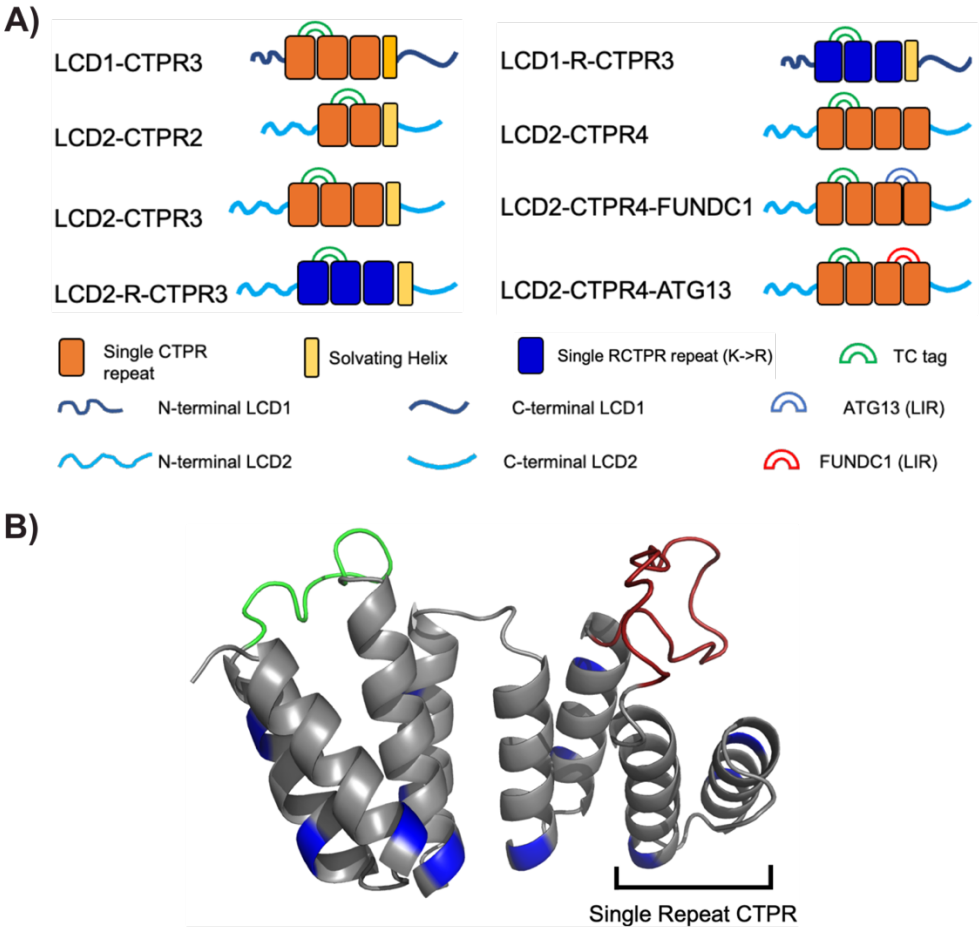

Figure 1: A) Composition of different LCD-CTPR variants. B) AlphaFold 2.0 predicted structure of the 4-repeat CTPR containing a tetracysteine motif (TC, green) between repeat 1-2 and a LC3 interacting region (LIR, red) between repeat 3-4. The lysine residues mutated to arginine in LCD-R-CTPR3 variant are highlighted in blue.

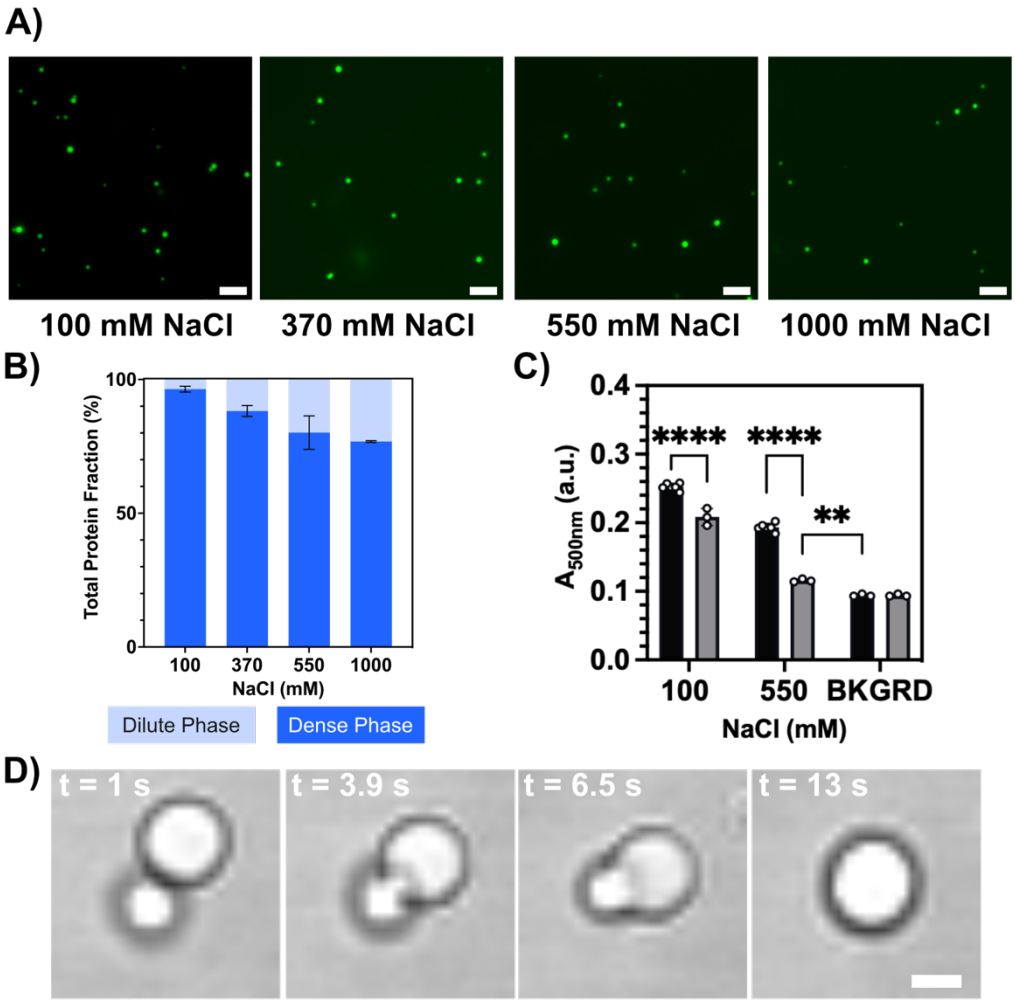

Figure 2: A) Fluorescent microscopy images of LCD2-CTPR4 after dilution into 50 mM Tris-HCl pH 7, 300 mM urea and containing varied salt concentrations. Scale = 5 µm. **B)** Sedimentation assay showing the proportion of dilute phase (light blue) to dense phase (dark blue). Error is standard deviation (n = 3). **C)** Turbidity assay of condensates formed with 50 mM Tris-HCl pH 7, 300 mM urea and either 100 mM or 550 mM NaCl (black) followed by treatment with 1,6-hexanediol (light grey). BKGRD represents buffer only absorbance in both conditions. **D)** Real-time microscopy images showing fusion events of LCD2-CTPR4 in 50 mM Tris-HCl pH 7, 300 mM urea, 1000 mM NaCl. Scale = 2 µm.

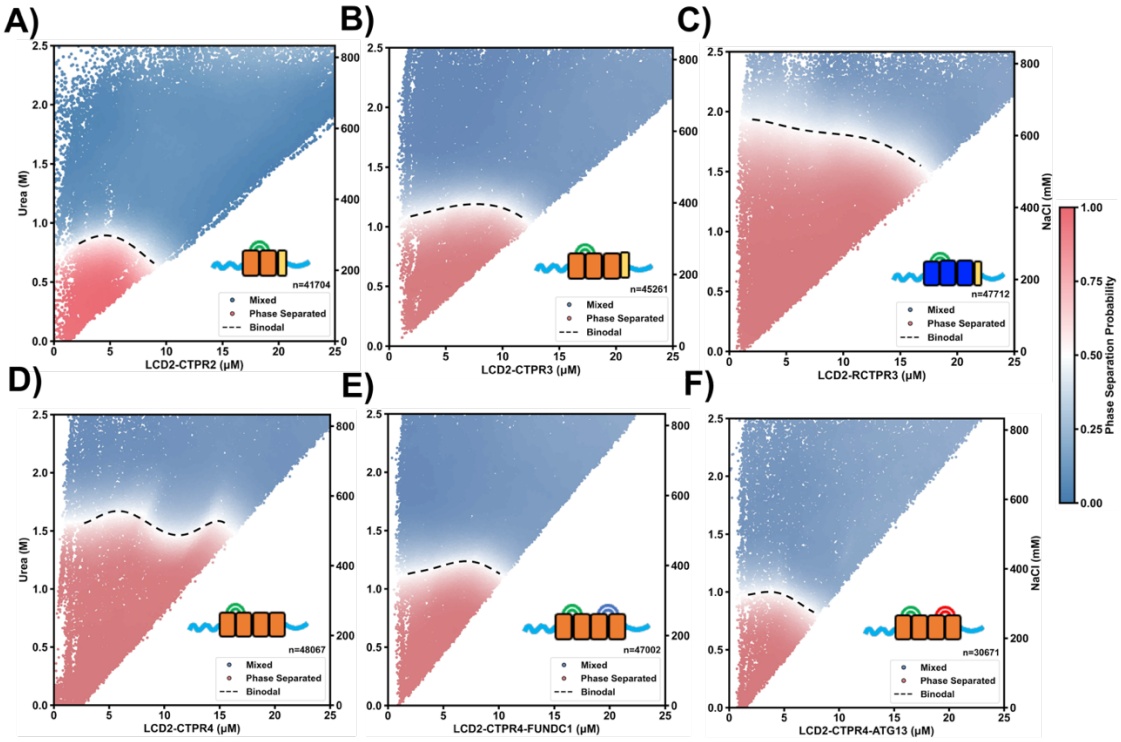

Figure 3: A-F) Phase diagrams of LCD2-CTPR variants vs urea and NaCl concentration, 50 mM Tris pH 8.5. Red and blue data points in the scatter plot correspond to individual microdroplets classified as phase-separated or homogeneous, respectively. The colour-coded heat map shows the estimate phase separation probability over the range of protein vs. urea/salt concentrations. The boundary (dashed line) was acquired by finding a region where the phase separation probability is equal to 0.5 and fitting a line to these coordinates. Data are representative images (n=3 experimental repeats).

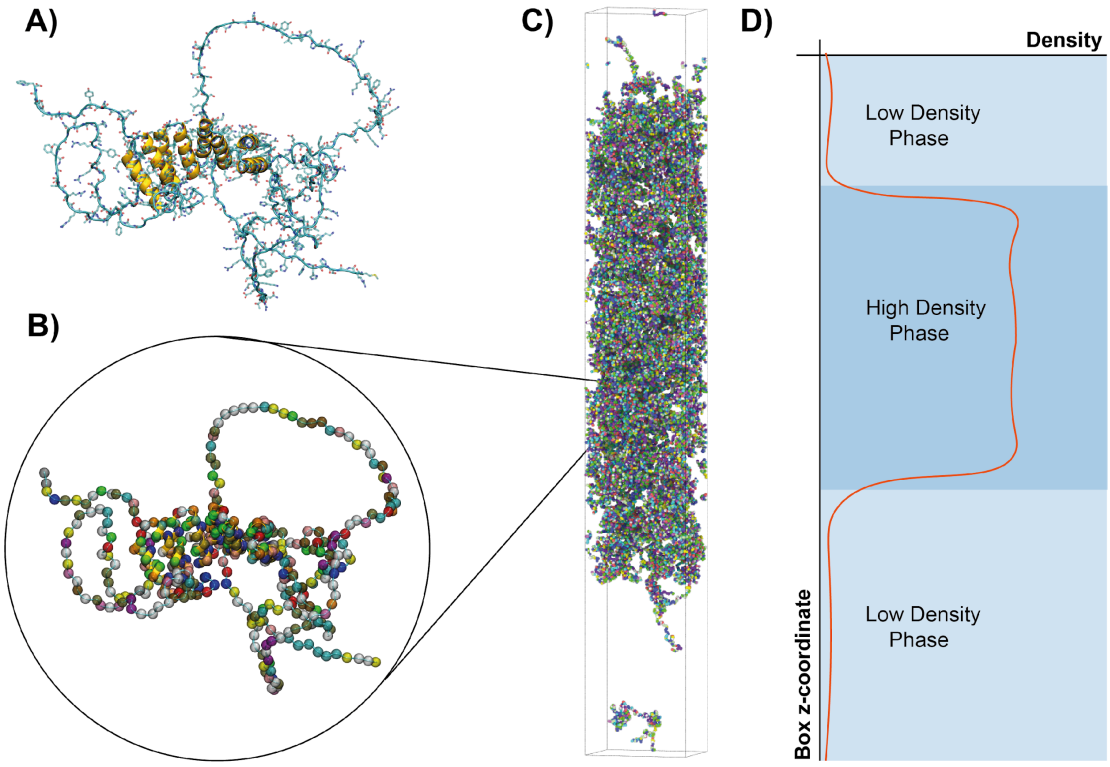

Figure 4: A) Predicted structure of a CTPR construct (LCD2-CTPR4) in atomistic resolution, determined using AlphaFold [Jumper 2021]. (42) **B)** Coarse-grained representation of the CTPR-protein. **C)** Simulation box for a Direct Coexistence simulation. **D)** Profile of the concentration density of the simulated.

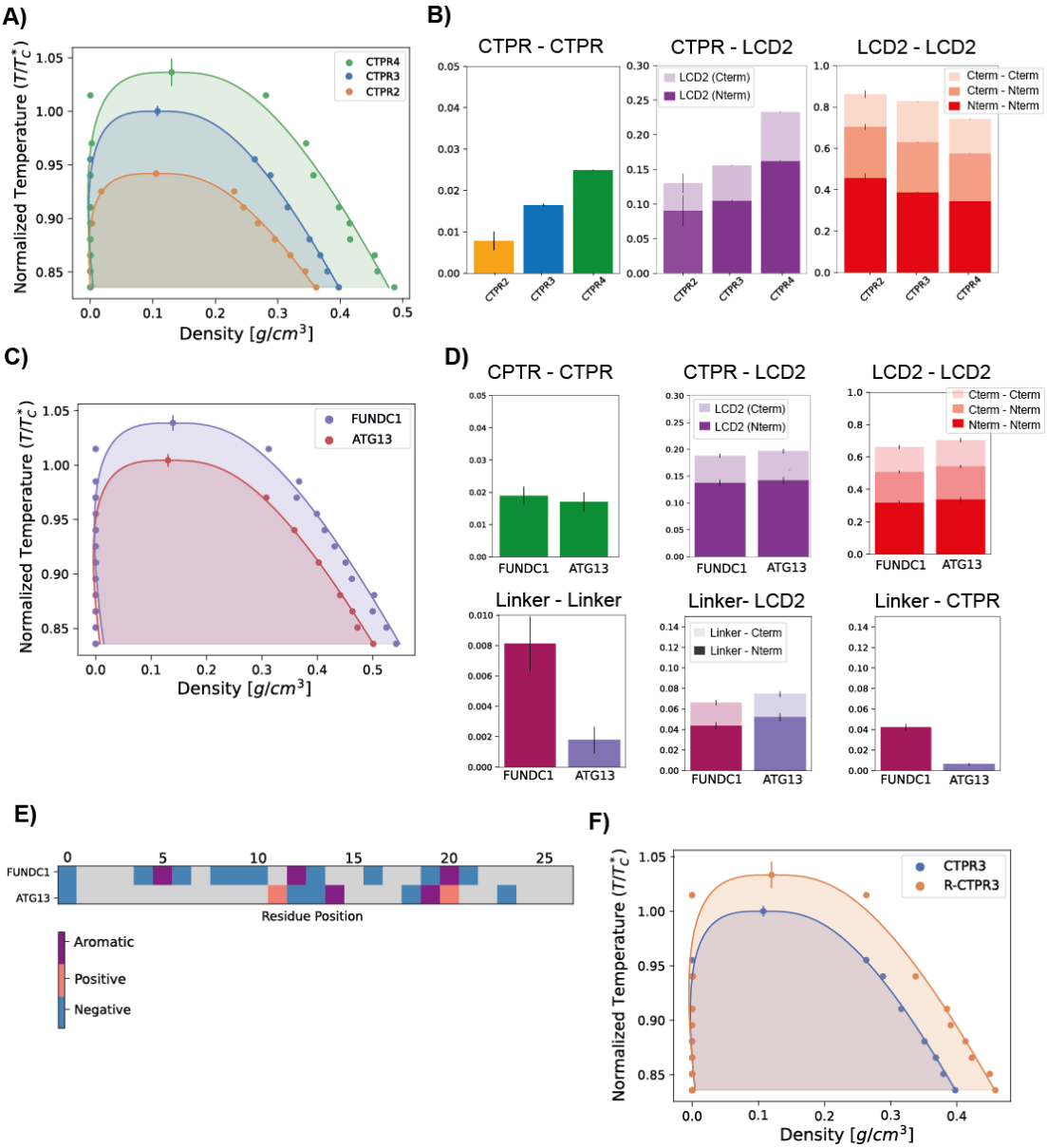

Figure 5: *In silico* analysis of phase-separation of LCD2-CTPR variants. A) Regulation of critical parameters based on the number of CTPR repeats. Binodal in the Temperature-Density phase space of LCD2-CTPR2, LCD2-CTPR3 and LCD2-CTPR4. Temperature in the y-axis is normalized by T*c, the critical temperature of LCD2-CTPR3, to aid in comparison. B) Contact analysis of different protein regions in the condensate: 'Repeat' stands for CTPR module, and the LCD2 has been divided into C-termini LCD2 and N-termini LCD2. The contact value represents the percentage of the intermolecular contacts that are of each correspondent type, as an average across simulation frames, and the error bars, the standard error of the average. C) Regulation of critical parameters according to linker identity. Binodal in the Temperature-Density phase space of LCD2-CTPR4-FUNDC1 and LCD2-CTPR4-ATG13. D) Contact analysis of the different protein regions in the condensate, including the linker region. E) Analysis of the amino acids in the linkers FUNDC1 and ATG13. F) Regulation of critical parameters by selective arginine mutations. Binodal in the Temperature-Density phase space of LCD2-CTPR3 and LCD2-R-CTPR3.

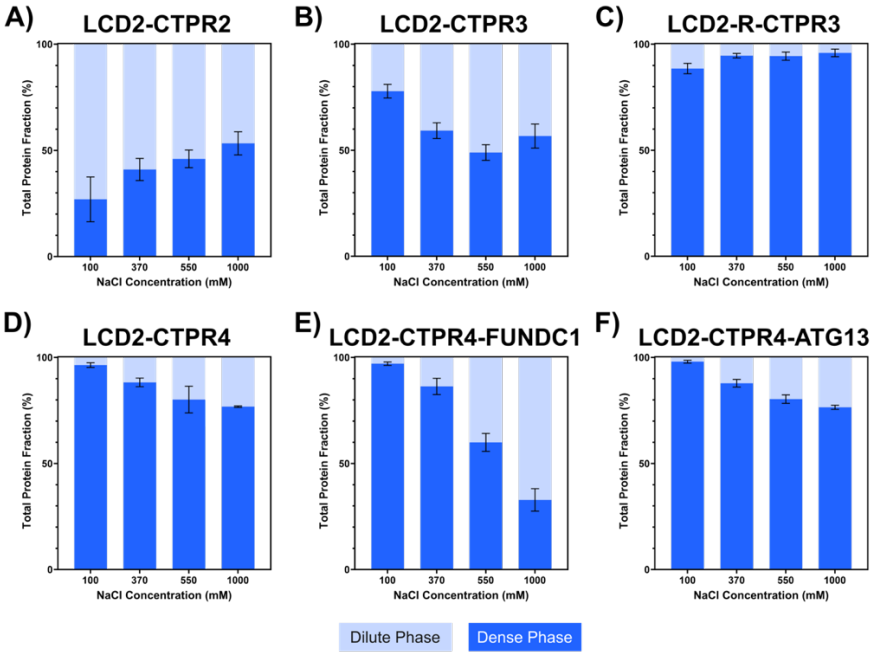

Figure 6: A-F) Sedimentation assay for LCD2-CTPR variants. Plots show the molar proportions of the dilute and dense phases in the total 2 μM protein solutions at different salt concentrations with 50 mM Tris-Cl, pH 7, 300 mM urea. Error is standard deviation (n = 3).

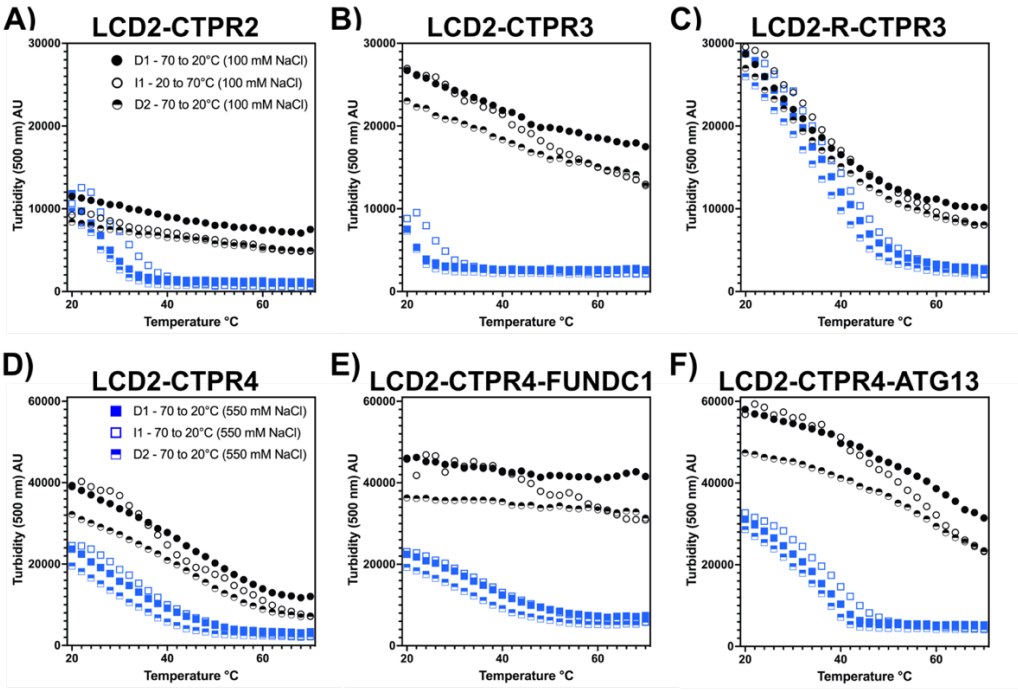

Figure 7: Turbidity assay for LCD2-CTPR variants. Condensates were induced using standard dilution methods and heated to 70°C. The turbidity (at 500 nm) was monitored incrementally from 70 → 20°C (D1: solid symbols), followed by an incremental increase from 20 → 70 °C (I2: open symbols) and finally another incremental decrease from 70 → 20°C (D2: half-filled symbols). The results for condensates induced in low salt (100 mM NaCl; black) and moderate salt (550 mM NaCl; blue) are shown. Samples contained 2 μM protein, 50 mM Tris-HCl pH 7, 300 mM urea and the indicated salt concentrations. Data is representative traces from n=3 experimental repeats.

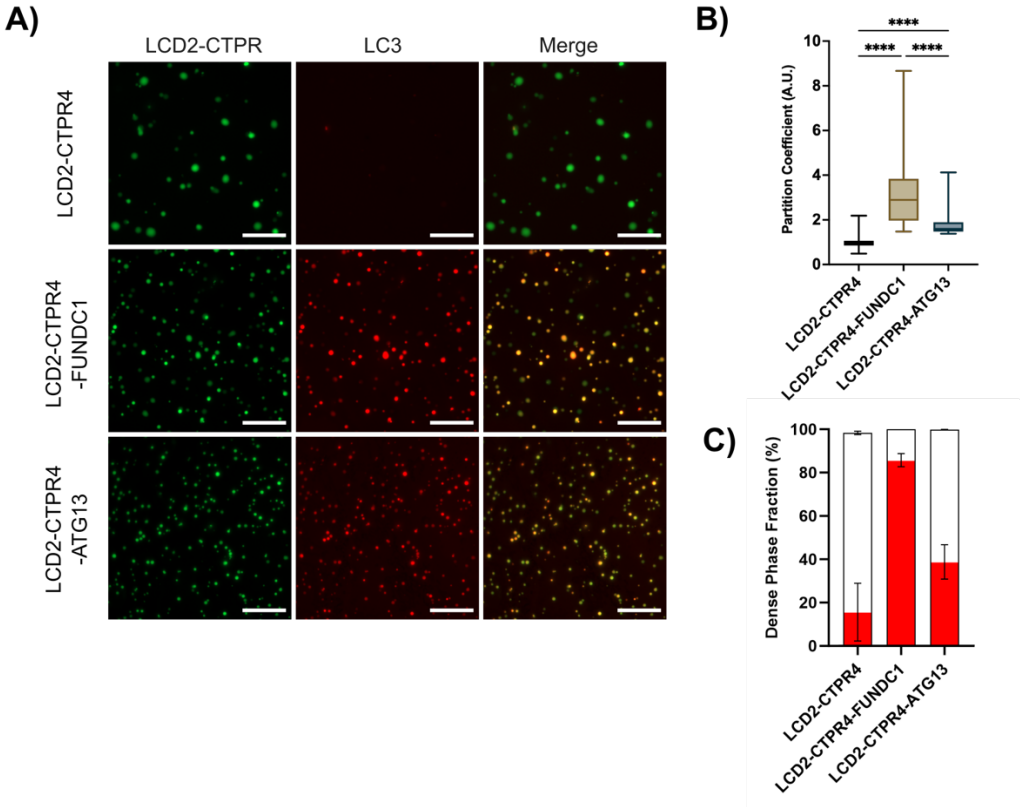

Figure 8: Recruitment of LC3 to LCD2-CTPR4-LIR condensates.

A) Fluorescence microscopy with FlAsH-labelled LCD2-CTPR4 variants (50 mM Tris, pH 7, 100 mM NaCl, 300 mM urea), showing that ROX-labelled LC3 can colocalise to those that contain an LIR. Scale bars are 20 µm. B) Partition coefficient data was measured with unlabelled LCD2-CTPR4 variant condensates (50 mM Tris-HCl, pH 7, 100 mM NaCl, 100 mM urea) co-localised with ROX-labelled LC3. Significance determined by 1-way analysis of variance (ANOVA); **** $P < 0.0001$, *** $P < 0.0002$, ** $P < 0.0021$, * $P < 0.0332$, ns < 0.1234. C) Sedimentation assay data showing extent of LC3 co-localisation to LCD2-CTPR4 condensates (50 mM Tris-HCl, pH 7, 100 mM NaCl, 100 mM urea). Empty bars are proportion of 2 µM LCD2-CTPR in the dense phase (close to 100% for all variants), filled red bars are proportion of 2 µM LC3 in the dense phase. Error is standard deviation (n = 3).

Table 1: C_{sat} values with standard deviation for the different LCD2-CTPR variants (50 mM Tris, pH 7, 300 mM urea, 1 M NaCl, room temperature)

Protein	C_{sat} (μM)
LCD2-CTPR2	0.95 ± 0.05
LCD2-CTPR3	1.01 ± 0.15
LCD2-R-CTPR3	0.27 ± 0.13
LCD2-CTPR4	0.36 ± 0.17
LCD2-CTPR4-FUNDC1	2.25 ± 0.13
LCD2-CTPR4-ATG13	0.51 ± 0.10