3 May 2026

Peptide heterogeneity enables and stabilises peptide/nucleic acid coacervates

Fatma Zohra Mihoubi[1,2], Karina Kinuyo Nakashima[1], Roman Staňo[3], Muhammad Ghufran Rafique[1], Fidan Rahmatova[1,2], Robin Kryštůfek[4], Kieran Russell[3], Václav Verner[4], Klara Hlouchova[5,4], Rosana Collepardo-Guevara[6,3], Claudia Bonfio[1]

1. Department of Biochemistry University of Cambridge
2. Institut de Science et d’Ingénierie Supramoléculaire, CNRS UMR 7006, University of Strasbourg
3. Yusuf Hamied Department of Chemistry University of Cambridge
4. Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences
5. Department of Cell Biology Charles University
6. Department of Genetics University of Cambridge

Abstract

Prebiotic environments were likely dominated by chemically heterogeneous mixtures of short peptides and oligonucleotides, yet the consequences of this diversity for the emergence of primitive compartments remain poorly understood. Here, we show that peptide length, sequence, and compositional complexity critically shape coacervation with short nucleic acids under prebiotically relevant conditions. Longer peptides form more stable coacervates, but mixed systems often outperform their pure components, revealing that heterogeneity can strengthen rather than destabilise phase separation. Sequence-specific effects further tune assembly: arginine-rich peptides outcompete lysine analogues, C-terminal glycine weakens stability, and even negatively charged residues can promote coacervation when multivalency and hydrogen bonding are preserved. Most strikingly, a chemically diverse tripeptide library readily coacervates with a range of DNA and RNA oligonucleotides, while retaining robust biomolecular partitioning and enhanced tolerance to Mg^{2+} stress. Together, these findings indicate that prebiotic chemical imperfection does not hinder protocell formation, but instead provides a powerful route to functional primitive coacervates.

Keywords

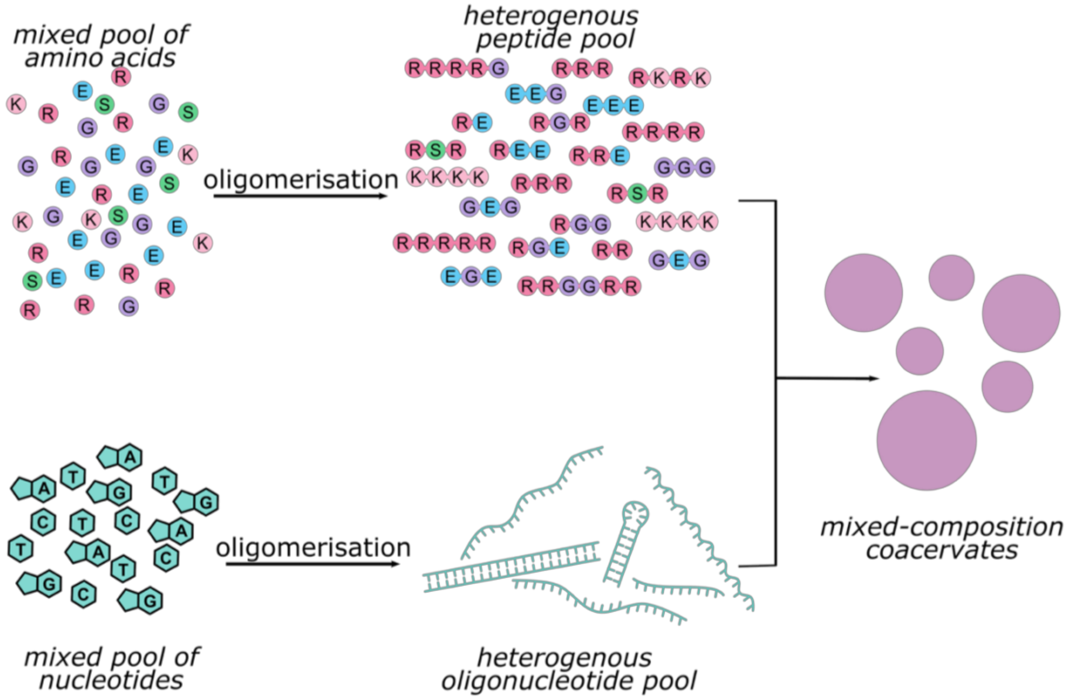
protocells, coacervates, peptides, origins of life, systems chemistry

Peptide heterogeneity enables and stabilises peptide/nucleic acid coacervates

Mihoubi F. Z.,[1,2,‡] Nakashima K. K.,[1,‡] Staňo, R.,[3] Rafique M. G.,[1] Rahmatova F.,[1,2] Kryštůfek R.,[4] Russell K.,[3] Verner V.,[4] Hlouchova K.,[4,5] Collepardo-Guevara R.,[3,6,*] Bonfio C.[1,*]
[1]Department of Biochemistry, University of Cambridge, CB2 1GA Cambridge, UK
[2]Institut de Science et d'Ingénierie Supramoléculaire, CNRS UMR 7006, University of Strasbourg, 67000, Strasbourg, France
[3]Yusuf Hamied Department of Chemistry, University of Cambridge, CB2 1EW Cambridge, UK
[4]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czech Republic
[5]Department of Cell Biology, Faculty of Science, Charles University, Prague, Czech Republic
[6]Department of Genetics, University of Cambridge, CB2 3EH, Cambridge, UK
[‡]These authors contributed equally
*Corresponding authors: cb2036@cam.ac.uk , rc597@cam.ac.uk

Abstract

Prebiotic environments were likely dominated by chemically heterogeneous mixtures of short peptides and oligonucleotides, yet the consequences of this diversity for the emergence of primitive compartments remain poorly understood. Here, we show that peptide length, sequence, and compositional complexity critically shape coacervation with short nucleic acids under prebiotically relevant conditions. Longer peptides form more stable coacervates, but mixed systems often outperform their pure components, revealing that heterogeneity can strengthen rather than destabilise phase separation. Sequence-specific effects further tune assembly: arginine-rich peptides outcompete lysine analogues, C-terminal glycine weakens stability, and even negatively charged residues can promote coacervation when multivalency and hydrogen bonding are preserved. Most strikingly, a chemically diverse tripeptide library readily coacervates with a range of DNA and RNA oligonucleotides, while retaining robust biomolecular partitioning and enhanced tolerance to Mg^{2+} stress. Together, these findings indicate that prebiotic chemical imperfection does not hinder protocell formation, but instead provides a powerful route to functional primitive coacervates.



Introduction

One of the enduring challenges in origins-of-life research is elucidating how dilute, heterogeneous mixtures of prebiotic molecules assembled into organised, functional systems capable of supporting life-like processes. Whether through extraterrestrial delivery or terrestrial non-enzymatic polymerisation of amino acids and nucleotides, life likely originated in chemically heterogeneous mixtures of short peptides and oligonucleotides of varying lengths and compositions.[1–3] These primitive polymers could, in turn, influence the organisation of their local environment by promoting compartmentalisation, thereby generating distinct microenvironments for further chemical evolution.[4] The emergence of the first living cells must therefore have relied on, and possibly arisen from, primitive compartments that tolerated, and potentially harnessed, chemical diversity to facilitate molecular organisation.[5]

Biomolecular phase separation is a spontaneous thermodynamic process in which a solution of biomolecules demixes into a dense phase, known as biomolecular condensate, in coexistence with a dilute phase. In mixtures of oppositely charged molecules, such as polycations and polyanions, phase separation can give rise to liquid-like droplets known as complex coacervates. Coacervates share key physicochemical features with intracellular biomolecular condensates, which efficiently sequester proteins and nucleic acids and can prevent their degradation whilst preserving function.[6,7] Beyond their relevance to extant biology, coacervates formed from simple peptides and oligonucleotides have been explored as plausible primitive compartments that could concentrate prebiotic solutes and support non-enzymatic reactions, including peptide-bond formation and RNA elongation.[8,9] More recently, coacervates comprising short, mixed-sequence oligonucleotides and model polyarginines were shown to sustain aptamer reconstitution and templated RNA chemistry, demonstrating that even chemically minimal nucleic acid systems can promote function.[10]

Compared to nucleic acids, peptides offer far greater combinatorial and chemical diversity, even at minimal polymer lengths. The assembly of a tripeptide from a restricted prebiotic amino acid alphabet can already offer hundreds of sequences spanning a wide range of charge, polarity, and hydrophobicity profiles.[11,12] This diversity gives rise to distinct patterns of interaction with oligonucleotides and provides a powerful means of tuning coacervate assembly and physicochemical properties.[13,14] Given this potential, the influence of peptide sequence and length on coacervate formation has mostly been investigated in the presence of long, often homopolymeric nucleic acids, and in biological contexts, with modern amino acid compositions.[13–17] Prebiotic peptide synthesis would have inevitably yielded peptides of varied lengths and sequences,[18] raising the question of how such molecular heterogeneity influences phase separation and the emergence of primitive coacervates.

To address this question, here we systematically investigate how peptide heterogeneity governs phase separation with short nucleic acids under prebiotically relevant conditions. We varied peptide length to identify the minimal polymer size capable of driving coacervate formation, and peptide sequence to evaluate how individual residues influence phase behaviour. We further compared individual peptides with peptide mixtures to assess whether compositional heterogeneity enhances or disrupts assembly. Our results reveal that coacervate stability depends strongly on peptide length and sequence, but that mixtures can exhibit comparable or even higher phase separation propensities than their individual components. Moreover, a tripeptide library mixing sequences with net charges from −3 to +3 at ratios to achieve overall charge-balance, readily forms coacervates with short, heterogeneous DNA and RNA oligonucleotides, illustrating the intrinsic potential of chemically diverse peptide/nucleic acid ensembles to self-organise in prebiotic environments.

Results and Discussion

Length

Proteins and peptides are typically soluble only within specific concentrations, pH, temperature, and ionic-strength windows.[19] Outside these ranges, they readily condense into amorphous precipitates, amyloid fibrils, or other nano- and microstructures. Above threshold saturation concentrations and under certain ranges of pH, temperature, and ionic conditions, positively charged proteins and short peptides can also assemble into liquid-like coacervates together with polyanions. This behaviour depends strongly on the presence of positively charged and aromatic residues that foster multivalent intermolecular interactions and enable the formation of condensate-spanning networks that stabilise the coacervate.[20] In particular, arginine (Arg, R) side chains can engage in electrostatic, hydrogen-bonding, cation–π and π–π interactions with nucleotides, providing stronger and more versatile interaction modes than other residues and making Arg-rich sequences key drivers of RNA/protein condensation.[21]

To assess the minimal molecular requirements for the emergence of primitive coacervates, we systematically mapped the conditions under which positively charged short peptides mixed with oligonucleotides undergo phase separation, in HEPES pH 7.5. Mixed-sequence single-stranded (ss) DNA oligomers served as model polyanions to avoid base-specific effects, while Arg peptides with unprotected *N*- and *C-termini* provided polycationic counterparts. Under these conditions, most peptide/nucleic acid mixtures formed coacervates (Fig. 1a, extended data in Supplementary Fig. S1). When both partners exceeded a certain length (R_5 or DNA_{30}), precipitation into solid aggregates occurred instead, whereas Arg monomers (R_1) and dimers (R_2) produced soluble mixtures irrespective of DNA length. Overall, peptide length exerted a stronger influence on the onset of phase separation than DNA length, highlighting the dominant role of Arg-mediated multivalent interactions, as the guanidinium group enables particularly strong cation–π and electrostatic interactions with nucleic acids. For instance, mixtures of R_3 and DNA_8 remained soluble at all concentrations tested, whereas coacervation occurred when four nucleobases were added (DNA_{12}); conversely, a one-residue increase in the peptide (R_4) was sufficient to induce phase separation with DNA_8.

To evaluate how peptide characteristics affect coacervate stability, we next quantified salt-dependent dissolution using turbidity measurements. The critical salt concentration (CSC), *i.e.*, the highest ionic strength tolerated before droplet dissolution, was calculated for each peptide–DNA combination of amino acid (aa) and nucleotide (nt) concentrations (Fig. 1b, extended data in Supplementary Table S2 and Fig. S2). In line with previous reports[10], we found that, for a fixed oligonucleotide length, longer Arg peptides produced higher CSC values (Fig. 1b). At matched total charge concentrations (20 mM), the longer oligomer (R_8) exhibited significantly higher CSCs than its shorter analogue (R_4), indicating that chain elongation enhances coacervate stability, improves charge efficiency, and reduces the peptide amount required to achieve comparable salt resistance.

To gain molecular insight into the mechanisms underlying coacervate stability, we complemented experiments with direct coexistence molecular simulations of the same coacervates using the Mpipi-Recharged model,[22,23] where each amino acid and nucleotide is parameterised to capture sequence-dependent phase behaviour, including electrostatic, cation–π and π–π interactions. Notably, the model reproduces the dominant role of arginine residues in driving phase separation.[24–26] Through simulations at increased salt concentrations, we were able to determine the critical salt concentrations of different systems (Fig. 1c, Supplementary Fig. S3).

This approach enables direct quantification of coacervate phase diagrams and intermolecular connectivity inside the coacervates, in particular, the valency of peptides and nucleic acids, defined as the number of distinct binding partners engaged by each molecule. Analysis of simulation trajectories reveals that increasing peptide length enhances the peptide valency, allowing longer peptides to simultaneously engage multiple nucleic acid chains and thereby promote the formation of a percolated intermolecular network (Fig. 1d). This increase in multivalent connectivity provides a mechanistic explanation for the higher stability of coacervates formed by longer Arg peptides observed experimentally and the clear length dependence on coacervate stability (Supplementary Figs. S4-5).

Based on the shape of the binodal in Fig. 1b (and Supplementary Fig. S5, obtained via molecular dynamics simulations), we estimate that the critical salt concentration (CSC*) is attained at ~20 mM [aa]: 5 mM [nt]. Beyond that point, it seems the coacervates are able to withstand charge excess while maintaining integrity (Supplementary Figs. S6-7). Within the framework of classical Flory-Huggins polymer phase-separation theory, CSC* can be described as a function of polymer length and a residual, non-electrostatic interaction parameter (χ_r; see Equation (1) in Methods). In this framework, the standard entropic contribution associated with polymer mixing, as well as mean-field interaction effects, are already explicitly accounted for. The parameter χ_r therefore captures additional non-electrostatic contributions to the free energy of phase separation, including short-range interactions (e.g., cation–π, hydrogen bonding, stacking) and residual entropic effects. In our case, because peptide and oligonucleotide lengths differ, we estimated χ_r for a given peptide by varying only oligonucleotide length. The CSC* for R_4/DNA_N coacervates is consistent with a $\chi_r \sim 0.3$ (Fig. 1e), while for R_8/DNA_N coacervates $\chi_r \sim 0.5$, indicating that while electrostatics dominate, non-electrostatic contributions become increasingly important as peptide length increases. This trend is consistent with the enhanced multivalency of longer peptides observed in the simulations, which enables the formation of a more densely connected liquid network.[27] Importantly, this effect does not arise solely from an increased number of binding sites in longer peptides, but also from their covalent linkage along the same chain, which promotes intermolecular bridging and enhances network connectivity, thereby increasing the thermodynamic stability of the coacervates.

Despite overall qualitative agreement between experiment and coarse-grained simulations, the computed CSC values were consistently lower, and a few systems, including R_4/DNA_8, did not phase separate in simulations, albeit forming coacervates experimentally (Fig. 1a). This discrepancy likely reflects the model being parameterised using fully protected amino acids, rather than the uncapped peptides used experimentally here. In short peptides, terminal groups represent a significant fraction of the total charge and can introduce additional electrostatic and dipolar interactions that stabilise the coacervate phase[28]. Consistent with this interpretation, experimentally protecting the peptide termini reduces the CSC by ~33% for $Ac\text{-}R_4\text{-}NH_2/DNA_8$ and ~16% for $Ac\text{-}R_4\text{-}NH_2/DNA_{20}$ (Supplementary Table S1).

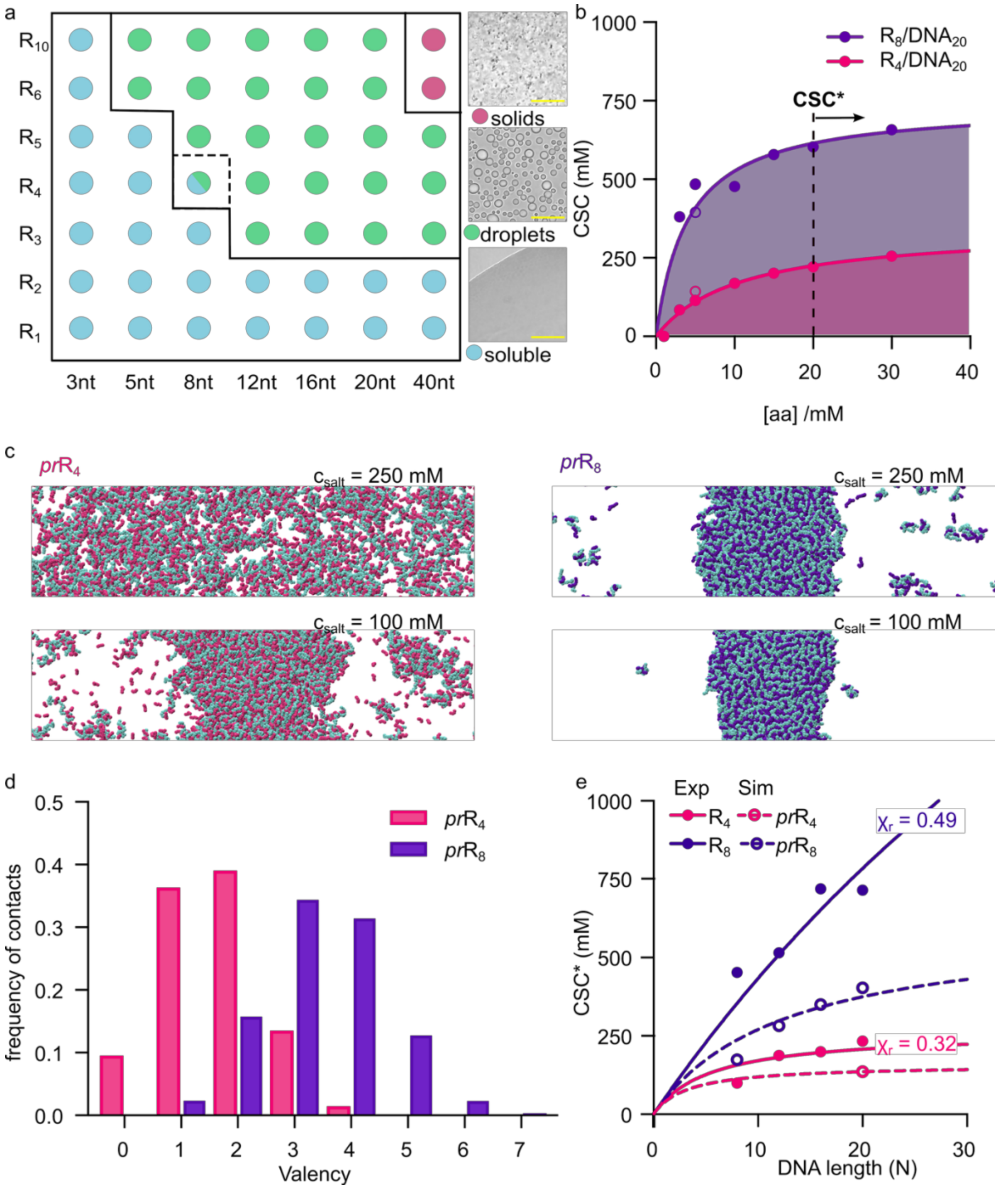


Figure 1. Effect of peptide length on phase separation with DNA and RNA. (a) compositional matrix indicating R_n peptide length, DNA_n length and the respective phase behaviour. The half blue/half green circle represents a mismatch between simulations and experiments, where simulations predict solubility and experiments lead to coacervates. Scale bar = 10 µm. (b) Critical salt concentration (CSC) of R_n/DNA_{20} mixtures at varying amino acid concentration ([aa]) and fixed [nt] = 5 mM; hollow circles represent CSC in silico. The dotted line represents the [aa]:[nt] ratio at which the CSC stabilises, referred to as CSC. The data were fitted to a one-site binding hyperbola. (c) Snapshots of coarse-grained simulations at different salt concentrations. Left:* pr*R_4/RNA_{20} mixture at 100 mM with poor phase separation (top) and 250 mM at a single-phase system (bottom). Right:* pr*R_8/RNA_{20} mixture at 100 mM (top) and 250 mM (bottom) showing robust phase separation in both cases. "pr" indicates peptide protection at the N and C termini, require dfor simulations. (d) Histogram of the frequency of contacts in silico for different valencies in the* pr*R_4 (magenta) and* pr*R_8 (purple) systems. The* pr*R_4 histogram peaks at n=2 while* pr*R_8 peaks at n=3, with a broader distribution. (e) CSC* of R4- and R8- based coacervates, varying DNA length measured experimentally (filled dots) and via simulations (hollow circles). The data were fitted according to Flory-Huggins to extract the non-electrostatic parameter χ_r, as detailed in Methods.*

Sequence

Charge density

In chemically heterogeneous prebiotic mixtures, the amino acid pool available for peptide formation was likely dominated by uncharged species, making highly charged sequences comparatively rare.[18] Under such conditions, the ability of short peptides to promote coacervate formation would depend not only on their net charge but also on how that charge is distributed along the chain and what other interactions the sequence can support.[29,30] We therefore hypothesised that (i) peptides containing fewer charged residues could still drive phase separation if additional non-ionic interactions, such as hydrophobic contacts, hydrogen bonding or π-interactions, stabilised the dense phase, and (ii) peptides of similar charge but different length, and hence charge spacing, might display distinct phase behaviour due to differences in the number and strength of intermolecular contacts.

Building on recent efforts to compare homo- and heteropeptides-based coacervates,[16,31] we examined how peptide sequence modulates coacervation by testing a library of peptides containing a fixed number of arginine residues and an increasing number of glycine spacers (Gly, G). All peptides carried a net positive charge (+4) but differed in length and therefore in charge density (σ).

The $R_2G_nR_2$ series (n = 0, 2, 8) consistently formed stable coacervates with DNA_7 but not DNA_6, regardless of peptide length, indicating that net charge rather than chain length governs the phase-separation threshold (Supplementary Fig. S8). For a more accurate estimation of the thresholds for coacervation, we used a linearisation of the plot of CSC* Vs 1/DNA-length (Fig. 1c) and obtained the minimal length as the x-intercept: R_6 (+6 charge, 6 residues) required only a tetramer (DNA_4) for coacervation, while R_4 (+4 charge, 4 residues) needed DNA_6 (Fig. 2a). Critically, $R_2G_2R_2$ (+4 charge, 6 residues), despite matching R_6 in length, also required DNA_6, confirming that net charge determines the minimal DNA length for phase separation more strongly than peptide length or charge density (Fig. 2a). The same approach can be applied to determine the minimal DNA length required for other peptides, using experimentally or computationally obtained CSCs (Supplementary Figures S3-4).

While net charge thus sets the phase-separation threshold, salt stability appears more sensitive to charge density. When $R_2G_2R_2$ (σ = 0.67) was mixed with DNA_{20}, we observed little to no change in the resulting coacervate salt stability compared to the R_4/DNA_{20} mixture, despite the two systems containing the same total concentration of Arg residues. At lower charge densities, *i.e.*, with longer Gly spacers, a slight decrease in salt stability was detected (Fig. 2b), in agreement with computational predictions. These results suggest that increased peptide length can compensate for reduced charge density down to a critical threshold, below which coacervate salt stability is expected to decrease more substantially.

Stability against NaCl is a well-established method for determining and comparing the relative stability of coacervates. However, the response to divalent salts does not necessarily follow the same trend. Mg^{2+} is a key player in several prebiotic and biological RNA processes, including folding, elongation and replication. In the presence of 20 mM $MgCl_2$, we observe R_4/DNA_{20} coacervate clustering under brightfield microscopy, which is significantly less present in low-density $R_2G_nR_2$ peptides (Supplementary Fig. S9). At 50 mM $MgCl_2$, R_4/DNA_{20} coacervates dissolve, whereas coacervates comprising $R_2G_2R_2$ or $R_2G_8R_2$ remain present (Supplementary Fig. S10). Taken together, these results support previous reports that lower charge-density coacervates could support Mg^{2+}-dependent catalysis inside the dense phase through better partitioning of Mg^{2+}.[16] They may also reduce the amount of free Mg^{2+} in the dilute phase, thereby limiting Mg^{2+}-induced clustering and preserving coacervate integrity under divalent cation stress.

Overall, these results suggest three prebiotic lessons. First, the net charge determines coacervation propensity over length and charge density. Second, at a given charge concentration, reduced charge density does not substantially compromise salt stability, indicating that heterogeneous peptides can still form robust coacervates if they retain enough positive charge. Third, more heterogeneous peptide compositions appear to confer improved resilience to Mg^{2+}, reducing clustering and helping maintain phase integrity under divalent cation stress. Together, these findings support the idea that chemically diverse peptide mixtures could still self-organise into stable coacervates in prebiotic environments.

Amino acid composition

To systematically probe how peptide sequence governs coacervate formation and function, we first examined patterning effects (with charge, length, and charge density held fixed), then explored the impact of amino acid identity across different chain lengths and charges, and finally assessed the functional consequences for biomolecular partitioning.

We compared the coacervation behaviour of arginine- (Arg, R) and lysine- (Lys, K) peptides in the presence of DNA_{20}. As previously reported for peptide-peptide mixtures[32], coacervates formed with K_4 were less stable than those with R_4, likely owing to the lower frequency of cation-π interactions between cationic Lys residues and DNA nucleobases.[14,33] These findings indicate that charge alone does not determine coacervation propensity; the molecular structure of the charged side chain also plays a critical role.[34]

We looked at how the introduction of a single Gly at different positions in a pentapeptide affects the relative stability of peptide/nucleic acid coacervates. Having a Gly at the *C*-terminus undermined coacervate stability more than at any other position, with R_4G_1-based coacervates having a 16% lower CSC than those formed with other Gly-containing peptides (Supplementary Fig. S11). We suggest that the presence of an Arg residue at the *C*-terminus neutralises the negative charge of the unprotected end, likely reducing repulsive interactions with the negatively charged DNA backbone. On the other hand, when Gly is introduced, such neutralisation does not occur, resulting in reduced overall DNA-peptide interactions, and hence coacervate stability.[35]

We next explored sequence effects in shorter tripeptides. To identify the influence of single amino acid modifications on peptide-driven coacervation, we tested RXR/DNA_{12} mixtures where X represents one of five amino acids: R, G, E, D, or S. Each peptide was evaluated for coacervation propensity by measuring the turbidity onset, *i.e.*, the amino acid concentration at which turbidity increases significantly due to droplet formation (Fig. 2c). As expected from our charge density studies, RGR showed the lowest propensity, while RRR formed coacervates most readily. Surprisingly, RER and RDR (each with a +1 net charge) phase-separated more readily than RGR and even RSR (+2 net charge). However, coacervation propensity did not correlate with stability, as RSR formed the most salt-resistant droplets, second only to RRR (Supplementary Table S1). These results demonstrate that coacervates with the lowest onset are not necessarily the most stable, and that increasing net charge does not always enhance phase separation. Instead, residue identity and sequence order play decisive roles in determining both the onset and stability of coacervation.

We then tested the limits of sequence variations with dimers. We previously reported that Arg dimers (R_2) would not form coacervates with DNA of any length up to 20 nucleotides (Fig. 1a). We thus sought to determine whether other dimers with different charge and composition could undergo phase separation with short oligonucleotides. Inspired by previous work showing that RER-polyphosphate coacervates are more stable than R_2- and R_3-polyphosphate coacervates,[36] we synthesised and tested peptide dimers: RG, GR, RE (Supplementary Figs. S12-15). We found that all dimers form coacervates when mixed with RNA_{12}. RG and GR, with a net charge of +1, required a concentration of 40 mM (amino acid), higher than that used in other experiments. In contrast, RE, with a net charge of 0, could form coacervates under the same conditions as Arg trimers and tetramers at 20 mM [aa] (Supplementary Table S3 and Supplementary Fig. S16). We reason that at short lengths, R_2 could efficiently coat a single DNA strand rather than generate and support an extended network of interactions with other DNA oligonucleotides that would ultimately enable phase separation of DNA. For longer peptides like R_3 and R_4, conformational effects can allow a single peptide to interact with multiple DNA strands via salt-bridge interactions, thereby favouring coacervation.[37] RE, although bearing a net charge of 0, can participate in strong peptide-peptide interactions, as compared to RR, with the glutamate residue available for H-bonding with other oligonucleotides and peptides.

To further explore the differences imparted to coacervates by these differently charged peptides, we evaluated the partitioning of various biomolecules. We systematically measured the partition coefficient (K_p) of coacervates made of RRR (+3 net charge), RSR (+2) and RER (+1) using four different probes, one positively charged probe (R_8-FITC), and three negatively-charged probes of different size and nature (fluorescein, DNA_8-FAM and FITC-E_8) (Fig. 2d, Supplementary Fig. S17). The polycationic probe (R_8) partitioned similarly across all coacervates, whereas anionic probes (fluorescein, DNA_8, and E_8) were progressively incorporated more efficiently as peptide charge decreased. A possible explanation is that reducing the net charge of the peptide minimises repulsive interactions with the scaffold, allowing better uptake of guest molecules.[38] A potential explanation may also be that reducing the charge density of the peptide reduces the interactions with the negative nucleic acid scaffold, thus allowing for more interactions with the probe.

Collectively, these findings reinforce that short, heteropolymeric peptides, plausibly present in prebiotic pools, are highly effective at driving coacervation. Beyond simple phase separation, these peptides form condensates that recruit nucleic acids more efficiently than homopolymers, while maintaining sufficient stability. Coupled with their previously reported enhanced Mg^{2+} tolerance and sequestration, such coacervates could both concentrate reaction components and buffer the dilute phase against disruptive divalent cations, creating favourable microenvironments for prebiotic RNA chemistry.

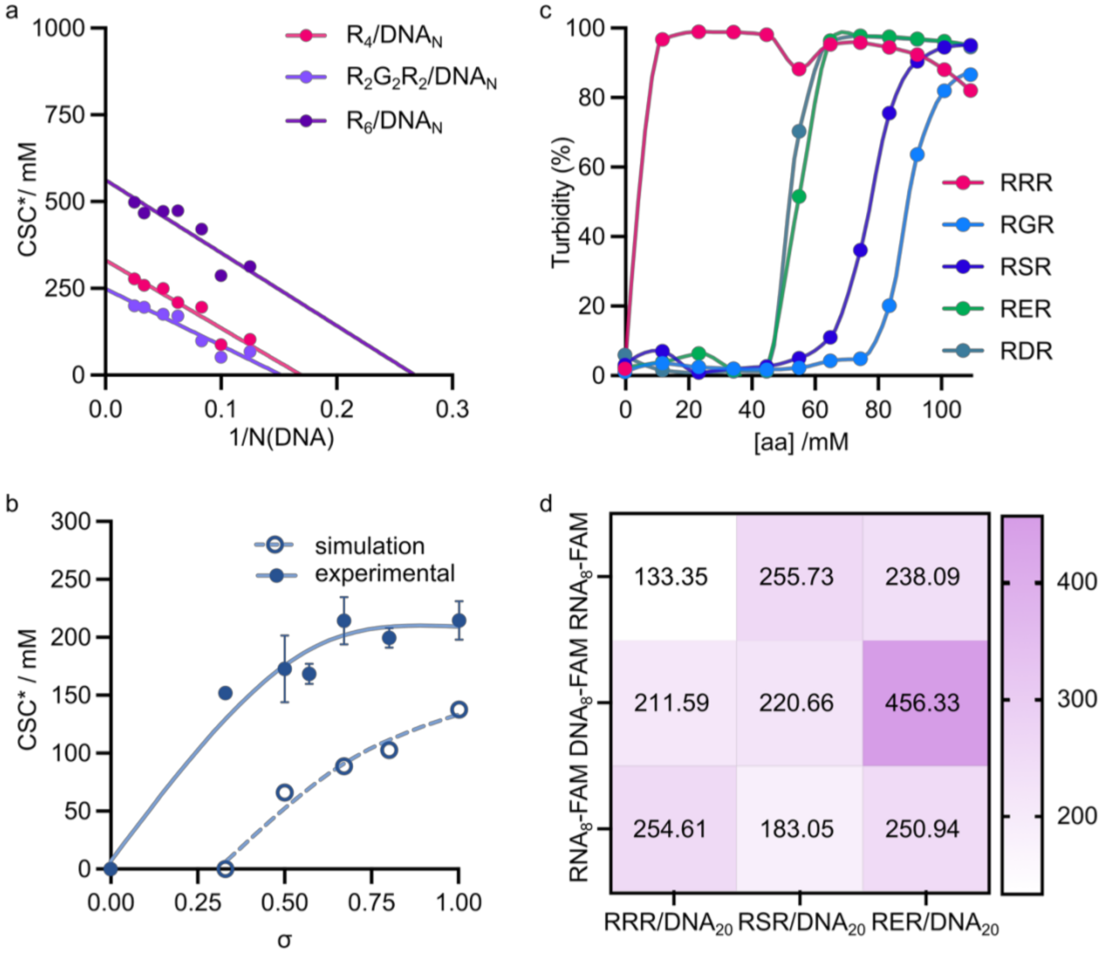


Figure 2. Effect of peptide sequence on formation and stability of coacervates with DNA. (a) Phase diagram of peptide/nucleic acid coacervates against charge density (b) Critical salt concentration (CSC) of peptide/nucleic acid mixtures as a function of peptide charge density. The curves represent a spline and are added as a guide to the eye. (b) Turbidity profile of DNA_{12} titrated with different RXR trimers ([nt] = 5 mM). (c) Partition coefficient (K_p) of R_8-FITC, E_8–FITC, DNA_8-FAM and RNA_8-FAM in RXR/DNA_{20} coacervates at [aa] = 100 mM and [nt] = 5 mM. R_8 = FITC-RRRRRRRR, E_8 = FITC-EEEEEEEE, DNA_8-FAM= FAM-ACTGACTG, RNA_8-FAM = FAM-ACUGACUG.

Mixtures

Prebiotic peptide synthesis likely produced a diversity of sequences, lengths, and charge patterns, dictated by the variable reactivity of each amino acid toward polymerisation.[3,39,40] Early Earth was thus characterised by messy chemical soups far richer than our model homo- or heteropeptides. Yet, if such heterogeneous peptide mixtures could still support coacervation with oligonucleotides, the emergence of primitive compartments, and thus primitive cells, would not have been a process requiring molecular specificity, but rather a statistical outcome of chemical complexity. We therefore began by testing the effect of simple additives (Arg monomer or dimer) on the stability of R_4/DNA_{20} coacervates. Next, we investigated whether defined mixtures of peptides representing realistic diversity (varying length, sequence and charge density) could undergo phase separation with DNA_{20}. Finally, in this section, we tested whether a complex tripeptide library, generated by reacting G, E, and R to form all 27 possible combinations, could likewise drive coacervation.

We began by examining the impact of simple amino acid monomers and dimers on the stability of pre-existing coacervates. In binary mixtures, replacing 25% of the amino acid content of R_4 with Arg monomer (R_1) (i.e. 15 mM [aa] from R_4 and 5 mM [aa] from R_1) had no significant effect on the CSC of R_4/DNA_{20} coacervates. However, at a 1:1 R_4:R_1 ratio of [monomer] (i.e. 10 mM [aa] from R_4 and 10 mM [aa] from R_1), the CSC decreased by 42% compared to R_4 alone (Supplementary Figs. S18-19). In other words, at high concentrations, Arg monomers behave like a salt, weakening interactions in R_4/DNA_{20} coacervates. A similar role was observed for glycine in nucleophosmin and rRNA condensates.[41] Similarly, turbidity onset measurements with DNA_{20} showed that titrating R_4 in the presence of R_1 raised the onset (lowering coacervation propensity) compared to R_4 alone, whereas adding R_2 lowered the onset. These results demonstrate that, while charged monomers may have destabilised peptide/nucleic acid coacervates, even their shortest polymerisation products, e.g., dimers, could rescue, if not enhance, the coacervation potential of peptide mixtures.

We next prepared three-component mixtures varying length (R_3, R_4, R_5), sequence (R_4, RKRK, K_4), and charge density (R_4, R_2GR_2, $R_2G_2R_2$) (Fig. 3a) at equimolar concentrations (6.67 mM amino acid each) and tested their ability to coacervate with DNA_{20} (5 mM nucleotide). Remarkably, each mixture exhibited a CSC higher than the average of its individual components at equivalent total amino acid concentration (Fig. 3b, Supplementary Fig. S20). Specifically, the length, sequence, and charge-density mixtures showed CSC values 23%, 53%, and 9% higher, respectively, than the average of the CSC values obtained for their individual constituents. These results demonstrate that coacervates not only readily formed with mixed-sequence, positively charged peptides and short oligonucleotides, but also benefited from the recruitment of shorter or less coacervation-promoting peptides. Importantly, peptide mixtures enhanced coacervate stability while requiring lower abundances of highly charged or longer peptides, which are least plausible prebiotically, as their formation requires higher polymerisation.

Simulations of these mixtures confirmed that the most coacervation-promoting components partition preferentially into the coacervate phase (Fig. 3c, Supplementary Figs. S21-26). Thus, the observed salt stability of mixture-based coacervates must account for the differential partitioning of all components. For example, R_5 ($K_p = 230$) partitions more strongly than R_3 ($K_p = 44.3$) in their respective, individual coacervates with DNA_{20}. Intriguingly, the K_p values of highly coacervation-promoting species like R_4 and R_5 were significantly reduced in mixtures: in R_3/R_4/R_5 coacervates with DNA_{20}, the K_p of R_5 was only one-fourth that in pure R_5 coacervates. Although peptides of higher multivalency still partition more than their shorter, less charged counterparts, this heterogeneity-induced crowding effect is most pronounced in the charge density mixture (C), where R_4 (normally the strongest partitioner) shows drastically reduced K_p compared to pure R_4/DNA_{20} coacervates. Similarly, we found that mixtures B and C were more resistant to Mg^{2+} than mixtures A and R_4/DNA_{20}, with coacervates still present at 50 mM $MgCl_2$ (Supplementary Fig. S27).

To explore an even more prebiotically plausible scenario, we expanded the screened sequence space to include peptides that were neutral or even negatively charged. This approach was supported by our earlier dipeptide screening, which showed that charge neutrality, or even negative charges on peptides, does not hinder coacervation but can, in fact, favour it. We thus generated a random tripeptide library by reacting Arg, Gly, and Glu, yielding up to 27 possible sequences with net charges ranging from -3 to +3. We then tested its ability to drive coacervation with various oligonucleotides: (ACTG) repeat sequences of ssDNA (DNA_n), its corresponding double-stranded DNA ($dsDNA_n$), and a random pool of oligos of unknown sequence ($\{A, C, T, G\}_n$) (Fig. 3d, Supplementary Figs. S28-30). Although this approach offers little control over exact sequence composition, it closely mimics prebiotic peptide diversity. With short strands (DNA_{12}), the peptide library remained soluble even at high concentrations. However, longer, heterogeneous DNA (DNA_{20}, DNA_{40}) and randomised 20-mer mixtures ($\{A,C,T,G\}_{20}$) induced robust droplet formation (Fig. 3e). Consistent with our previous report that RNA coacervates more readily than DNA,[10] RNA_{12} formed coacervates at only 20 mM total amino acid (Supplementary Fig. S31).

The strong interactions between peptides and DNA, combined with the rigidity of dsDNA compared to ssDNA, lead to precipitation at low salt concentration rather than coacervation.[42] In our system, we observed a similar behaviour with R_n/$dsDNA_{20}$ coacervates. However, when combining the RGE library with $dsDNA_{20}$, the mixture formed coacervates rather than solid aggregates, even though the system required twice as much peptide as with ssDNA. This finding suggests that the heterogeneity and lower charge density of the peptide library, compared to homogeneous, highly charged peptides, modulate interactions and allow the formation of fluid compartments rather than solid aggregates (Supplementary Figs. S32-33). To work with more prebiotically plausible mixtures, we decided to move from specific DNA sequences to random pools of fixed length, mirroring the reasoning behind the peptide library. A random pool of 20-mer DNA oligomers (N_{20}) was combined with the library. The mixtures formed coacervates at concentrations comparable to the DNA_{20} (control sequence, $(ACTG)_5$). Importantly, the CSC of N_{20} coacervates was about 4-fold higher than that of the DNA_{20} coacervates (CSC = 80.0 mM and 20.0 mM, respectively, Fig.3f). More than being evidently more prebiotically plausible, mixed oligo pools have been shown to reduce copying bias in non-enzymatic RNA replication fidelity.[43] In this context, the enhanced stability of coacervates from a mixed peptide library and a mixed oligo pool points towards a concordance of system complexity, coacervate formation and essential prebiotic processes.

Turbidity onset measurements in water revealed coacervation beginning at 31.9 mM amino acid, corresponding to ~10 mM arginine (one-third of the library's residues) (Supplementary Fig. S34). In HEPES buffer (25 mM, pH 7.5), the onset increased by 33% due to the increased ionic strength, corresponding to ~14 mM arginine, a value still lower than most individual Arg-based trimers tested earlier. Partitioning experiments with the library-generated coacervates showed that both positively and negatively charged probes partitioned to a comparable extent as in individual peptide-based coacervates (Supplementary Fig. S35). These results indicate that multi-component coacervates remain effective hubs for recruiting biomolecules from the dilute phase, irrespective of their compositional heterogeneity.

Taken together, these findings demonstrate the remarkable robustness of peptide/nucleic acid coacervation to compositional heterogeneity. Simple monomers destabilise pre-existing coacervates by acting as salts, but short dimers restore stability. Defined three-component mixtures outperform their individual components in salt stability while maintaining robust biomolecular partitioning, and the charge-balanced 27-tripeptide library coacervates efficiently with diverse oligonucleotides at low arginine-equivalent concentrations. Rather than requiring "molecular perfection", the emergence of primitive compartments appears statistically favoured by chemical diversity, rendering liquid-liquid phase separation a robust, accessible outcome of prebiotic chemistry.

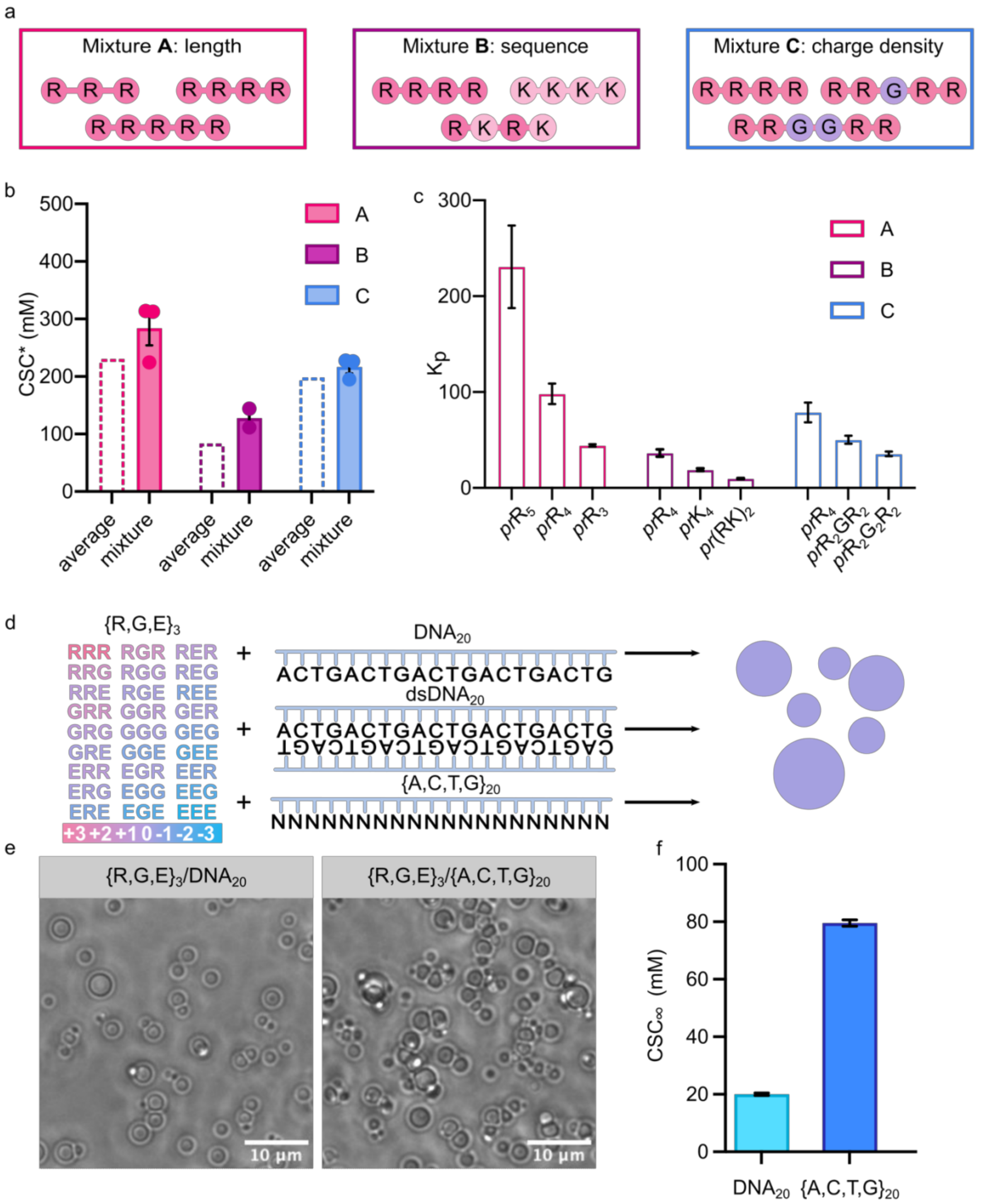


Figure 3. Effect of peptide diversity on coacervate stability. (a) Scheme of the three mixtures tested varying peptide: length (A), sequence (B) and charge density (C). (b) Critical salt concentration (CSC) of the individual components at the same charge ratio. The average CSC of the three peptides in the mixture is shown in dotted columns, next to the CSC of the mixture. (c) Partitioning of peptides in the dense phases of mixtures at 50 mM salt obtained via coarse-grained simulations. (d) Generation of a random library of trimers from a mixture of three amino acids. Multi-component coacervates form upon the addition of oligonucleotides. (b) Brightfield microscopy of the {R, G, E}$_3$ library mixtures and DNA$_{20}$ (left, [aa] = 80 mM) and with {A, C, T, G}$_{20}$ (bottom, [aa] = 80 mM); [nt] = 5 mM in 25 mM HEPES pH 7.5, scale bar: 10 µm. (f) CSC$_\infty$ of {R, G, E}$_3$ peptide library and DNA$_{20}$ or {A, C, T, G}$_{20}$ at 80:5 [aa]:[nt], in 25 mM HEPES pH 7.5. Error bars represent standard error, n = 3.

Conclusions

In this work, we aimed to understand how chemically heterogeneous prebiotic mixtures could have self-organised into functional compartments. While pure homopolymers reliably form coacervates, prebiotic reality demanded messy peptide ensembles of diverse lengths, sequences, and charges coacervating with heterogeneous oligonucleotides. This study systematically dissects how peptide physicochemical properties govern coacervation with short nucleic acids, and demonstrates that prebiotic "chemical imperfection" not only permits but systematically enhances protocell formation.

Peptide length emerges as the primary determinant of both coacervation onset and droplet stability: R_8 significantly outperforms R_4 at matched charge concentrations, while DNA length exerts a secondary, near-linear influence on the salt stability of coacervates. Sequence identity exerts fine control: arginine outperforms lysine despite equivalent charge due to more efficient cation-π interactions with nucleobases, while charge-neutralising *C*-terminal glycine dramatically reduces stability by eliminating carboxyl-end screening of DNA repulsion. Even more strikingly, negatively charged residues (Glu, Asp) paradoxically lower nucleation thresholds despite reducing net charge, revealing that multivalency and hydrogen-bonding capacity can outweigh electrostatics.

Compositional heterogeneity, inherent to prebiotic peptide pools, transforms rather than disrupts coacervation. Simple Arg monomers destabilise pre-existing R_4/DNA_{20} droplets by acting as salts, but dimers restore coacervate stability. Defined three-component mixtures systematically outperform their constituents, driven by selective partitioning of coacervation-promoting components despite heterogeneity-induced crowding effects. Most compellingly, an unbiased {G, E, R} tripeptide library spanning net charges from -3 to +3 efficiently enables coacervation with diverse oligonucleotides (DNA_{20}, $\{A, C, T, G\}_{20}$, RNA_{12}), matching or underscoring the performance of pure homopeptides. These "messy" systems retain robust biomolecular partitioning and, in line with previous studies,[44] have a higher Mg^{2+} tolerance, buffering the dilute phase against divalent cation stress. Short polymers rescue destabilising monomers, mixtures amplify coacervate stability beyond that of their pure components, and complex libraries self-organise despite their compositional heterogeneity. All in all, this work extends our prior demonstration that oligonucleotide heterogeneity enables the emergence of functional coacervates[10], revealing the vastly richer prebiotic peptide alphabet as an even more potent engine of emerging organisation.

Far from requiring molecular perfection, the formation of primitive coacervates appears statistically inevitable under prebiotically realistic conditions. Liquid-liquid phase separation emerges as a generic physicochemical response to polymer diversity, robust to the chemical messiness that defined early Earth. Future studies should test whether these heterogeneous coacervates catalyse peptide synthesis and peptide-driven processes or enable Mg^{2+}-dependent RNA chemistry more effectively than model systems, potentially revealing how chemical imperfection scaffolded the molecular precision of living cells.

References

(1) Kvenvolden, K.; Lawless, J.; Pering, K.; Peterson, E.; Flores, J.; Ponnamperuma, C.; Kaplan, I. R.; Moore, C. Evidence for Extraterrestrial Amino-Acids and Hydrocarbons in the Murchison Meteorite. *Nature* **1970**, *228* (5275), 923–926. https://doi.org/10.1038/228923a0.

(2) Verlander, M. S.; Orgel, L. E. Analysis of High Molecular Weight Material from the Polymerization of Adenosine Cyclic 2', 3'-Phosphate. *J Mol Evol* **1974**, *3* (2), 115–120. https://doi.org/10.1007/BF01796557.

(3) Fox, S. W.; Nakashima, T. The Assembly and Properties of Protobiological Structures: The Beginnings of Cellular Peptide Synthesis. *Biosystems* **1980**, *12* (3–4), 155–166. https://doi.org/10.1016/0303-2647(80)90013-1.

(4) Fox, S. W. *Self-Assembly of the Protocell from a Self-Ordered Polymer*; 1967.

(5) Szostak, J. W.; Bartel, D. P.; Luisi, P. L. Synthesizing Life. *Nature* **2001**, *409* (6818), 387–390.

(6) Aumiller, W. M. Jr.; Pir Cakmak, F.; Davis, B. W.; Keating, C. D. RNA-Based Coacervates as a Model for Membraneless Organelles: Formation, Properties, and Interfacial Liposome Assembly. *Langmuir* **2016**, *32* (39), 10042–10053. https://doi.org/10.1021/acs.langmuir.6b02499.

(7) Drobot, B.; Iglesias-Artola, J. M.; Le Vay, K.; Mayr, V.; Kar, M.; Kreysing, M.; Mutschler, H.; Tang, T.-Y. D. Compartmentalised RNA Catalysis in Membrane-Free Coacervate Protocells. *Nat Commun* **2018**, *9* (1), 3643. https://doi.org/10.1038/s41467-018-06072-w.

(8) Wang, J.; Abbas, M.; Wang, J.; Spruijt, E. Selective Amide Bond Formation in Redox-Active Coacervate Protocells. *Nat Commun* **2023**, *14* (1), 8492. https://doi.org/10.1038/s41467-023-44284-x.

(9) Poudyal, R. R.; Guth-Metzler, R. M.; Veenis, A. J.; Frankel, E. A.; Keating, C. D.; Bevilacqua, P. C. Template-Directed RNA Polymerization and Enhanced Ribozyme Catalysis inside Membraneless Compartments Formed by Coacervates. *Nat Commun* **2019**, *10* (1), 490. https://doi.org/10.1038/s41467-019-08353-4.

(10) Nakashima, K. K.; Mihoubi, F. Z.; Saraya, J. S.; Russell, K. O.; Rahmatova, F.; Robinson, J. D.; Maristany, M. J.; Huertas, J.; Rubio-Sánchez, R.; Collepardo-Guevara, R.; O'Flaherty, D. K.; Bonfio, C. Differential Stability and Dynamics of DNA-Based and RNA-Based Coacervates Affect Non-Enzymatic RNA Chemistry. *Nat Commun* **2025**, *16* (1), 9296. https://doi.org/10.1038/s41467-025-64335-9.

(11) Patel, B. H.; Percivalle, C.; Ritson, D. J.; Duffy, C. D.; Sutherland, J. D. Common Origins of RNA, Protein and Lipid Precursors in a Cyanosulfidic Protometabolism. *Nature Chem* **2015**, *7* (4), 301–307. https://doi.org/10.1038/nchem.2202.

(12) Makarov, M.; Sanchez Rocha, A. C.; Krystufek, R.; Cherepashuk, I.; Dzmitruk, V.; Charnavets, T.; Faustino, A. M.; Lebl, M.; Fujishima, K.; Fried, S. D.; Hlouchova, K. Early Selection of the Amino Acid Alphabet Was Adaptively Shaped by Biophysical Constraints of Foldability. *J. Am. Chem. Soc.* **2023**, *145* (9), 5320–5329. https://doi.org/10.1021/jacs.2c12987.

(13) Chang, L.-W.; Lytle, T. K.; Radhakrishna, M.; Madinya, J. J.; Vélez, J.; Sing, C. E.; Perry, S. L. Sequence and Entropy-Based Control of Complex Coacervates. *Nat Commun* **2017**, *8* (1), 1273. https://doi.org/10.1038/s41467-017-01249-1.

(14) Fisher, R. S.; Elbaum-Garfinkle, S. Tunable Multiphase Dynamics of Arginine and Lysine Liquid Condensates. *Nat Commun* **2020**, *11* (1), 4628. https://doi.org/10.1038/s41467-020-18224-y.

(15) Abbas, M.; Lipiński, W. P.; Nakashima, K. K.; Huck, W. T. S.; Spruijt, E. A Short Peptide Synthon for Liquid–Liquid Phase Separation. *Nat. Chem.* **2021**, *13* (11), 1046–1054. https://doi.org/10.1038/s41557-021-00788-x.

(16) Iglesias-Artola, J. M.; Drobot, B.; Kar, M.; Fritsch, A. W.; Mutschler, H.; Dora Tang, T.-Y.; Kreysing, M. Charge-Density Reduction Promotes Ribozyme Activity in RNA–Peptide Coacervates via RNA Fluidization and Magnesium Partitioning. *Nat. Chem.* **2022**, *14* (4), 407–416. https://doi.org/10.1038/s41557-022-00890-8.

(17) Baruch Leshem, A.; Sloan-Dennison, S.; Massarano, T.; Ben-David, S.; Graham, D.; Faulds, K.; Gottlieb, H. E.; Chill, J. H.; Lampel, A. Biomolecular Condensates Formed by Designer Minimalistic Peptides. *Nat Commun* **2023**, *14* (1), 421. https://doi.org/10.1038/s41467-023-36060-8.

(18) Forsythe, J. G.; Petrov, A. S.; Millar, W. C.; Yu, S.-S.; Krishnamurthy, R.; Grover, M. A.; Hud, N. V.; Fernández, F. M. Surveying the Sequence Diversity of Model Prebiotic Peptides by Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114* (37). https://doi.org/10.1073/pnas.1711631114.

(19) Creighton, T. E. *Proteins: Structures and Molecular Properties*, 2. ed., 8. print.; Freeman: New York, 2006.

(20) Saar, K. L.; Morgunov, A. S.; Qi, R.; Arter, W. E.; Krainer, G.; Lee, A. A.; Knowles, T. P. J. Learning the Molecular Grammar of Protein Condensates from Sequence Determinants and Embeddings. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118* (15), e2019053118. https://doi.org/10.1073/pnas.2019053118.

(21) Paloni, M.; Bussi, G.; Barducci, A. Arginine Multivalency Stabilizes Protein/RNA Condensates. *Protein Science* **2021**, *30* (7), 1418–1426. https://doi.org/10.1002/pro.4109.

(22) Tejedor, A. R.; Gonzalez, A. A.; Maristany, M. J.; Chew, P. Y.; Russell, K.; Ramirez, J.; Espinosa, J. R.; Collepardo-Guevara, R. Chemically-Informed Coarse-Graining of Electrostatic Forces in Charge-Rich Biomolecular Condensates. July 27, 2024. https://doi.org/10.1101/2024.07.26.605370.

(23) 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* (11), 732–743. https://doi.org/10.1038/s43588-021-00155-3.

(24) Wang, J.; Choi, J.-M.; Holehouse, A. S.; Lee, H. O.; Zhang, X.; Jahnel, M.; Maharana, S.; Lemaitre, R.; Pozniakovsky, A.; Drechsel, D.; Poser, I.; Pappu, R. V.; Alberti, S.; Hyman, A. A. A Molecular Grammar Governing

the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **2018**, *174* (3), 688-699.e16. https://doi.org/10.1016/j.cell.2018.06.006.

(25) Krainer, G.; Welsh, T. J.; Joseph, J. A.; St George-Hyslop, P.; Hyman, A. A.; Collepardo-Guevara, R.; Alberti, S.; Knowles, T. P. J. Reentrant Liquid Condensate Phase of Proteins Is Stabilized by Hydrophobic and Non-Ionic Interactions. *Biophysical Journal* **2021**, *120* (3), 28a. https://doi.org/10.1016/j.bpj.2020.11.426.

(26) 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 Behaviours of Disordered Prion-like Domains. *Nat. Chem.* **2022**, *14* (2), 196–207. https://doi.org/10.1038/s41557-021-00840-w.

(27) Espinosa, J. R.; Joseph, J. A.; Sanchez-Burgos, I.; Garaizar, A.; Frenkel, D.; Collepardo-Guevara, R. Liquid Network Connectivity Regulates the Stability and Composition of Biomolecular Condensates with Many Components. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117* (24), 13238–13247. https://doi.org/10.1073/pnas.1917569117.

(28) Kolchina, N.; Khavinson, V.; Linkova, N.; Yakimov, A.; Baitin, D.; Afanasyeva, A.; Petukhov, M. Systematic Search for Structural Motifs of Peptide Binding to Double-Stranded DNA. *Nucleic Acids Research* **2019**, *47* (20), 10553–10563. https://doi.org/10.1093/nar/gkz850.

(29) 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* (6478), 694–699. https://doi.org/10.1126/science.aaw8653.

(30) Sathyavageeswaran, A.; Pandey, P. K.; Holmlund, N.; Kaushik, P.; McIntosh, S.; Mokdad, R.; Perry, S. L. Role of Charge Patterning and Hydrophobicity in Peptide-Based Complex Coacervates. *Biomacromolecules* **2025**, *26* (10), 6574–6585. https://doi.org/10.1021/acs.biomac.5c00773.

(31) Valdes-Garcia, G.; Gamage, K.; Smith, C.; Martirosova, K.; Feig, M.; Lapidus, L. J. The Effect of Polymer Length in Liquid-Liquid Phase Separation. *Cell Reports Physical Science* **2023**, *4* (5), 101415. https://doi.org/10.1016/j.xcrp.2023.101415.

(32) Cakmak, F. P.; Choi, S.; Meyer, M. O.; Bevilacqua, P. C.; Keating, C. D. Prebiotically-Relevant Low Polyion Multivalency Can Improve Functionality of Membraneless Compartments. *Nat Commun* **2020**, *11* (1), 5949. https://doi.org/10.1038/s41467-020-19775-w.

(33) Kumar, K.; Woo, S. M.; Siu, T.; Cortopassi, W. A.; Duarte, F.; Paton, R. S. Cation–π Interactions in Protein–Ligand Binding: Theory and Data-Mining Reveal Different Roles for Lysine and Arginine. *Chem. Sci.* **2018**, *9* (10), 2655–2665. https://doi.org/10.1039/C7SC04905F.

(34) Abbas, M.; Lipiński, W. P.; Wang, J.; Spruijt, E. Peptide-Based Coacervates as Biomimetic Protocells. *Chem. Soc. Rev.* **2021**, *50* (6), 3690–3705. https://doi.org/10.1039/D0CS00307G.

(35) DeRouchey, J. E.; Rau, D. C. Role of Amino Acid Insertions on Intermolecular Forces between Arginine Peptide Condensed DNA Helices. *Journal of Biological Chemistry* **2011**, *286* (49), 41985–41992. https://doi.org/10.1074/jbc.M111.295808.

(36) Dai, S.; Xie, Z.; Wang, B.; Ye, R.; Ou, X.; Wang, C.; Yu, N.; Huang, C.; Zhao, J.; Cai, C.; Zhang, F.; Buratto, D.; Khan, T.; Qiao, Y.; Hua, Y.; Zhou, R.; Tian, B. An Inorganic Mineral-Based Protocell with Prebiotic Radiation Fitness. *Nat Commun* **2023**, *14* (1), 7699. https://doi.org/10.1038/s41467-023-43272-5.

(37) Yang, B.; Jin, S.; Park, Y.; Jung, Y. M.; Cha, H. J. Coacervation of Interfacial Adhesive Proteins for Initial Mussel Adhesion to a Wet Surface. *Small* **2018**, *14* (52), 1803377. https://doi.org/10.1002/smll.201803377.

(38) Sathyavageeswaran, A.; Izzo, M.; Sabadini, J. B.; Habeeba, U.; Belden, J.; Perry, S. L. Tuning Protein Uptake through Charge and Hydrophobicity in Peptide-Based Complex Coacervates. *Biomacromolecules* **2026**, acs.biomac.5c02639. https://doi.org/10.1021/acs.biomac.5c02639.

(39) Canavelli, P.; Islam, S.; Powner, M. W. Peptide Ligation by Chemoselective Aminonitrile Coupling in Water. *Nature* **2019**, *571* (7766), 546–549. https://doi.org/10.1038/s41586-019-1371-4.

(40) Verlander, M. S.; Lohrmann, R.; Orgel, L. E. Catalysts for the Self-Polymerization of Adenosine Cyclic 2',3'-Phosphate. *J Mol Evol* **1973**, *2* (4), 303–316. https://doi.org/10.1007/BF01654098.

(41) Xu, X.; Van Haren, M. H. I.; Smokers, I. B. A.; Visser, B. S.; White, P. B.; Jansen, R. S.; Spruijt, E. Amino Acids Bind to Phase-Separating Proteins and Modulate Biomolecular Condensate Stability and Dynamics. *Nat Commun* **2025**, *16* (1), 8723. https://doi.org/10.1038/s41467-025-63755-x.

(42) Fraccia, T. P.; Jia, T. Z. Liquid Crystal Coacervates Composed of Short Double-Stranded DNA and Cationic Peptides. *ACS Nano* **2020**, *14* (11), 15071–15082. https://doi.org/10.1021/acsnano.0c05083.

(43) Duzdevich, D.; Carr, C. E.; Colville, B. W. F.; Aitken, H. R. M.; Szostak, J. W. Overcoming Nucleotide Bias in the Nonenzymatic Copying of RNA Templates. *Nucleic Acids Research* **2024**, *52* (22), 13515–13529. https://doi.org/10.1093/nar/gkae982.

(44) Agrawal, A.; Radakovic, A.; Vonteddu, A.; Rizvi, S.; Huynh, V. N.; Douglas, J. F.; Tirrell, M. V.; Karim, A.; Szostak, J. W. Did the Exposure of Coacervate Droplets to Rain Make Them the First Stable Protocells? *Sci. Adv.* **2024**, *10* (34), eadn9657. https://doi.org/10.1126/sciadv.adn9657.

Acknowledgements

The authors acknowledge funding from the NWO (Dutch Research Council) via a Rubicon Fellowship (019.222EN.011 to K.K.N.), the Human Frontier Science Program Organization (HFSPO) *via* an Early Career Research Grant (RGY00062/2022, to C.B.), the ERC (Starting Grant) under the European Union's Horizon Europe research and innovation programme (GA 101162933 to C.B.), the UKRI Future Leaders Fellowship (UKRI2316/G130052 to C.B.), the Federation of European Biochemical Society *via* a FEBS Excellence Award (to C.B.), the Agence Nationale de la Recherche via an ANR AAPG JCJC 2022 (to C.B.), the CSC Graduate School funded by the Agence Nationale de la Recherche (CSC-IGS ANR-17-EURE-0016 for doctoral funding to F.R.), the University of Strasbourg Institute for Advanced Study (USIAS) via a USIAS Fellowship (to C.B.), the Foundation Jean-Marie Lehn, the Ernest Oppenheimer Fund *via* an Early Career Research Fellowship (to M.G.R.) and the UKRI EPSRC under the UK Government's guarantee scheme (EP/Z002028/1 to R.C.G.), following funding by the ERC (Consolidator Grant) under the European Union's Horizon Europe research and innovation programme. 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 (EP/R029407/1 to R.C.G.).

Contributions

F.Z.M., K.K.N. and C.B. designed and conceptualised the study. F.Z.M., K.K.N. and G.R. performed all experiments aimed at the systematic characterisation of primitive coacervates. F.R. performed peptide synthesis and characterised the peptide material. C.B. supervised peptide synthesis and coacervate experiments. R.S. and K.O.R. conceptualised and performed simulations. R.S. and K.O.R. analysed simulation data. R.C.G. supervised simulation work. R.K. and V.V. prepared the peptide library under the supervision of K.H.. K.K.N., R.C.G. and C.B. acquired funding. All authors discussed the results, contributed to the writing of the first draft of the manuscript, provided comments and approved the final version of the manuscript.

Peptide heterogeneity enables and stabilises peptide/nucleic acid coacervates - Supplementary information

Mihoubi F. Z.,[‡,1,2] Nakashima K. K.,[‡,1] Staňo R.,[3] Rafique M. G.,[1] Rahmatova F.,[1,2] Kryštůfek, R.,[4] Russell K.,[3] Verner, V.,[4] Hlouchova K.,[4,5] Collepardo-Guevara R.,*[,3,6] Bonfio C.*[,1]
[1]Department of Biochemistry, University of Cambridge, CB2 1GA Cambridge, UK
[2]Institut de Science et d'Ingénierie Supramoléculaire, CNRS UMR 7006, University of Strasbourg, 67000, Strasbourg, France
[3]Yusuf Hamied Department of Chemistry, University of Cambridge, CB2 1EW Cambridge, UK
[4]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czech Republic
[5]Department of Cell Biology, Faculty of Science, Charles University, Prague, Czech Republic
[6]Department of Genetics, University of Cambridge, CB2 3EH, Cambridge, UK
[‡]These authors contributed equally
*Corresponding authors: cb2036@cam.ac.uk , rc597@cam.ac.uk

Table of Contents

Methods

Materials

Reagents were purchased from Merck and Thermo Fisher and used without further purification unless otherwise stated. Oligonucleotides were purchased from Integrated DNA Technologies (IDT) or synthesised in-house when indicated. Peptides were purchased as TFA salts from GenScript or synthesised in-house when indicated. Sep-Pak C18 classic cartridges were purchased from Waters (Milford, MA). Samples containing DNA or RNA were prepared with Milli-Q water 18 MΩ.

Fmoc solid phase peptide synthesis (Fmoc-SPPS) was carried out on an induction heating-assisted PurePrep® Chorus synthesiser (Gyros Protein Technologies) pressurised with 4.5 N_2 and equipped with two independent reaction vessel slots with both induction heating and a UV-monitoring detector. Reverse-phase high-performance liquid chromatography (RP-HPLC) purifications on peptides were performed using an Agilent semi-preparative HPLC system equipped with a 1260 Infinity II binary pump, 1260 Infinity II variable wavelength detector with 3 mm preparative cell, and a 1290 Infinity II preparative open-bed sampler/collector with a 20 mL injection loop on a ReproSil Pur 120 C18-AQ 250 ×25 mm 5 µm particle size column from DrMaisch GmbH. pH was monitored with a Mettler Toledo FiveEasy pH meter, and adjustments were made with aqueous solutions of NaOH or HCl as appropriate. The turbidity of mixtures was determined using a BMG Labtech CLARIOstar plus. Concentrations were calculated using the Beer-Lambert equation (molar extinction coefficients were estimated using the OligoAnalyzer™ Tool (IDT)).

Microscopy images were processed using Fiji (http://rsb.info.nih.gov/ij/). Quantification of oligonucleotides using a ClarioStar Lvis microplate, reading absorption at 260 nm at least 3 times and correcting each value by a blank measurement.

Average and standard error values refer to $n \geq 3$ replicates. Statistical significance was determined using unpaired t-tests (ns $P > 0.05$; *$P \leq 0.1$; **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$).

Stock solutions

Peptide stocks were prepared in MilliQ water at a 100 mM monomer concentration, based on the molecular weight of the TFA salt. Solutions were stored at −20 °C and vortexed for 1 min before use. Single-strand DNA and RNA oligonucleotide stocks were prepared in DNase/RNase-free water at ~1 mM strand concentration. To facilitate the solubilisation of the oligonucleotides, solutions were heated to 50 °C for 5 min and cooled down to room temperature before the measurement. The concentration was determined from a diluted solution (250–500×) by measuring absorbance at 260 nm.

Peptide synthesis

Dipeptides were synthesised according to previously published SPPS procedures by using Fmoc-protected amino acids. Specifically, 0.50 mmol of commercially available, pre-loaded Wang resin was added to a plastic reactor equipped with a fritted plastic insert. The resin was allowed to swell in DMF for 30 min. For the deprotection step, 20% piperidine in DMF (5 mL/0.5 mmol) was added to the resin. The resin was left to react for 2 min before solvent removal. The treatment was repeated with 20% piperidine in DMF and left shaking for 15 min. The solvent was removed, and the resin was washed with DCM/DMF (5 × 5 mL). For the coupling step, Fmoc-protected amino acids (Fmoc-AA-OH) (3.0 equiv. relative to the resin loading) were dissolved in dry DMF. A solution of 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 3.8 equiv.) was added to the Fmoc-AA-OH solution, followed by *N,N*-Diisopropylethylamine (DIPEA, 6.0 equiv.), and added to the resin. The resulting mixture was agitated on a laboratory shaker for 45 min. DMF washes of the resin (5 ×5.0 mL) were performed before deprotection. Cycles of coupling and deprotection steps were performed to obtain the desired peptide sequence. After the final Fmoc removal, the resin was washed with DMF (3×), DCM (3×), and MeOH (3×), then left to dry under high vacuum overnight.

For cleavage, the resin was treated with the cleavage solution (trifluoroacetic acid (TFA):H2O:triisopropyl silane 95:2.5:2.5 volume ratio) for 2 h. TFA-peptide solutions were collected, and the resin was washed with TFA (2 × 3 mL). The collected fractions were concentrated under nitrogen flow and then added to cold diethyl ether, resulting in the precipitation of the peptide. The precipitate was centrifuged for 5 min at 5000 × g and washed with cold diethyl ether (10 mL). The resulting peptide was dissolved in acetonitrile:water (1:5, 10 mL) and lyophilised. Peptides were purified by RP-HPLC. Elution was performed at a flow rate of 10 mL/min using a linear gradient of acetonitrile and ultrapure water (both containing 0.1% LCMS grade formic acid). The gradient ranged from 20% to 80% acetonitrile over 1 h. UV absorption at 220 and 254 nm was used to monitor the collection of unprotected peptides. Fractions containing the target product were identified by mass spectrometry and lyophilised.

R_2: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 4.44-4.35 (m, 1H), 4.09 (t, J = 6.4 Hz, 1H), 3.25 (t, J = 6.8 Hz, 2H), 2.58–2.44 (m, 2H), 2.28–1.61 (m, 6H).
RG: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 4.14–3.98 (m, 3H), 3.25 (t, J = 6.8 Hz, 2H), 2.03–1.91 (m, 2H), 1.80–1.64 (m, 2H).
GR: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 4.47-4.40 (m, 1H), 3.90 (s, 2H), 3.23 (t, J = 6.8 Hz, 2H), 2.02–1.59 (m, 4H).
RE: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 4.48-4.32 (m, 1H), 4.12 (t, J = 5.9 Hz, 1H), 3.24 (q, J = 7.2 Hz, 3H), 2.19–1.43 (m, 8H).

Coacervate formation

Coacervates were formed at a 10-100 µL scale. Samples were prepared by adding, in the following order: MilliQ, HEPES buffer (25 mM, pH 7.5), peptides (from a 100 mM [peptide] stock solution), and DNA (from a 1-2 mM [DNA] stock solution). Fluorescent dyes (FITC-R_8, DNA_8FAM, RNA_8FAM, FITC-E_8, Fluorescein) were added to the coacervates as a final step, unless otherwise stated. Coacervate formation was confirmed by turbidity observed under the naked eye and by light microscopy.

Determination of the critical salt concentration (CSC)

The critical salt concentration was determined using turbidity measurements. Turbidity was indirectly measured on a plate reader, reading the absorbance at 600 nm and using the relation:

$$Turbidity = 100 - Transmittance\_\% = 100(1 - 10^{(Abs_{blank} - Abs)})$$

Samples of 100 µL were prepared in a COSTAR 96-well plate and titrated with concentrated NaCl stocks (1, 3, or 5 M). The concentration of the salt stock was chosen to minimise sample dilution during titrations (to a maximum of 20%) and maximise the number of points measured during the steep decay of absorbance. At the end of the titration, all mixtures reached the turbidity corresponding to the blank (100 µL of MilliQ). The CSC was calculated as follows:

- At each point, turbidity was plotted against the exact NaCl concentration;
- The curve was fitted every three points with the following equation: $Turbidity = a \times [NaCl] + b$;
- The linear fit with the steepest slope was used to identify the inflexion point i. The CSC was thus calculated as $CSC = -\frac{a_i}{b_i}$.

CSC vs [aa] data was fitted to a one-site binding hyperbola using the GraphPad curve fit function.

To determine the residual interaction parameter (Xr), the data were fitted using a least squares regression method to the following equation, derived from the classic mean-field Flory-Huggins model for polymer mixtures:[1]

$$CSC^* = \frac{A^2\alpha^4}{l^3\left(1 - 2\chi_r + \frac{2}{\sqrt{N}}\right)^2} \tag{1}$$

Where CSC^* represents the CSC of a given R_n/DNA_n system at 20:5 mM [aa]:[nt] A, the electrostatic interaction constant; α, the degree of ionisation; l, the characteristic length scale of a monomer; N, the DNA length; and χ_r the non-electrostatic interaction parameter. The following constants were fixed: $A = 1.0$; $\alpha = 0.95$; $l = 2\,nm$ and the data was fitted to extract χ_r.

Turbidity onset

We define the coacervation onset as the amino acid concentration at which each peptide forms coacervates in the presence of oligonucleotides, as assessed by turbidity measurements. Turbidity measurements were performed by monitoring absorbance at 600 nm in a plate reader during titration of the oligonucleotide solution with a concentrated peptide stock until the absorbance reached a maximum. As previously discussed, absorption was converted to turbidity, and the onset concentration corresponds to the amino acid concentration for turbidity >20%.

Preparation of observation chambers

Glass coverslips of 1.5 thickness were used to image coacervates under brightfield microscopy and confocal microscopy. Coverslips were passivated using polyvinyl alcohol (PVA, 13-23k, 87-89% hydrolysed). To do so, a 5% PVA solution was prepared by sonicating PVA in MilliQ water until fully dissolved. The solution was deposited on the coverslips for 1 hour, then rinsed and dried with compressed air. To make observation chambers, 4 mm-wide holes were punched in 3M double-sided tape (GPT-020F, 0.2 mm thickness) and sealed with a 10 mm-diameter circular coverslip.

Brightfield microscopy

Coacervates were imaged using a Nikon Eclipse TS2 inverted epifluorescence microscope equipped with a Moment A21K635003 camera (0.63× adaptor) and a Plan Fluor 40x air objective (Numerical Aperture = 0.75).

Confocal microscopy

Confocal microscope images were acquired on a Leica SP5 II, equipped with an HCX PL APO CS 63.0x1.40 oil objective lens and an Argon laser (488 nm, 9.52 mW).

Partition coefficients

Partitioning of fluorescent dyes in coacervates (Kp) was calculated using the following equation:

$$K_p = \frac{I_{\text{droplet}} - I_{\text{dark}}}{I_{\text{dilute phase}} - I_{\text{dark}}}$$

The fluorescence intensity within the droplet (Idroplet) was averaged across all droplets in the field of view (FOV) using a particle analysis plugin in ImageJ, with a low threshold to prevent underestimation. The intensity of the dilute phase was averaged for the entire FOV after droplets were removed. I_{dark} corresponds to the intensity measured in a sample lacking any fluorophore at the same laser power used for the respective sample.

Library design

Synthesis of pooled RGE trimer library

1 g of Rink Amide AM resin (0.64 mmol/g) was swelled in 20 ml of DMF for 30 min, and deprotected by first washing the resin with 10 ml of 20% (v/v) piperidine in DMF on a centrifugal filter, and then incubating with an additional 10 ml piperidine solution for 30 min. Resin was washed on a centrifugal filter unit (Biorad 7326204) 8 times with 10 ml of DMF, split into 10 equal portions as a DMF slurry in 15 ml conical tubes, and the residual DMF was drained with a pipette tip. Each portion was then coupled with 10 mL of a solution containing 300 mM of the AAs building block (Fmoc-L-Arg(Boc)2-OH, Fmoc-L-Glu(OtBu)-OH, or Fmoc-Gly-OH), 375 mM Oxyma Pure, and 375 mM *N,N'*-diisopropylcarbodiimide. Couplings were then placed on a carousel at RT and rotated at 20 RPM, protected from light. The following

day, all the split reactions were pooled, washed 4 times on a filter unit with 10 ml of DMF, and the next cycle commenced with deprotection with piperidine as described above.

After final Fmoc cleavage, the resin was washed 8 times with DCM and dried by spinning at 1000 × g for 10 min. Resin was then treated with 20 ml of 94.5:0.5:5 TFA:TIS:water mix for 2 h on a carousel at 20 RPM, protected from light. Resin was removed by filtration, and the collected filtrate was evaporated to dryness. Residue was then chilled to -80°C and briefly (~10 s) triturated with 3 × 10 ml of diethyl ether. The washed residue was allowed to dry in the open for 1 h, then resuspended and collected into a single tube in 20 ml of 1 mM HCl. Suspension was then lyophilised o/n, and the following day the material was fully dissolved in 20 ml of 1 mM HCl, and lyophilised again for a total of 4 times.

Commercially available HPLC-grade ACN, DMF, DIC and building blocks were used without further purification. Chemicals were obtained from Sigma-Aldrich, Lach-Ner, Novabiochem, and Iris Biotech. The molecular weight distribution of the combinatorial peptide library was confirmed by mass spectrometry using an UltrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) and an InfinityLab Single Quadrupole MS System (Agilent Technologies, Santa Clara, USA) according to standard procedures. Amino acid analysis was carried out using a Biochrom 30 (Harvard Bioscience, Holliston, USA) with ninhydrin derivatisation and hydrolysis in 6 N HCl at 110 °C in a vacuum-sealed ampoule for 20 hours.

Data analysis

Experimental and computational data were plotted using GraphPad Prism 10.6.

Computer Simulations

Microscopic Model

To represent the coacervate in silico, we use the coarse-grained model Mpipi-Recharged,[2] which is a particle-based bead-spring model derived for accurate simulations of the phase behaviour of peptides and unstructured single-stranded RNA. Each amino acid and each nucleotide is represented as a single spherical particle, while both peptide bonds and RNA bonds are emulated with harmonic potentials. The non-bonded interactions between the particles are governed by a combination of the Wang-Frenkel[3] potential and the screened electrostatic Yukawa potential.[4] The pair-wise force constants of both potentials were derived from the atomistic simulations as previously reported. [2,5] Both solvent and mineral ions are modelled only implicitly as a continuum, and their effect enters the energy balance only through the relative permittivity and screening length.

Simulation Method

To solve the model, we use the OpenMM 8.1.26 implementation of Molecular dynamics with the Langevin thermostat, with a coupling of 0.01/ps and a timestep of 10 fs. To simulate the phase behaviour of coacervates, we used the direct-coexistence method,[6] in which a slab of the coacervate representing the biomolecule-rich (liquid) phase is co-existing with the biomolecule-poor phase (vapour) phase within the same rectangular simulation box with typical dimensions 150 nm = Lx = 6Ly = 6Lz and with periodic boundary conditions. To build the system, we first initialise all the peptides and RNA molecules as rigid rods with no mutual overlaps. We then simulate the system for 10 ns, applying an external force on each molecule to pull them toward the centre of the simulation box, thereby creating a dense slab. We then remove the pulling force and equilibrate the system for 500 ns at T = 300K and the chosen salt concentration, followed by the production simulations spanning up to 1000 ns. All simulations are long enough to observe Rouse-like relaxation of chains by many radii of gyration, both in and out of the slab, and many events of vaporisation/condensation of individual peptides and RNA chains. A typical simulation contains between 25000 and 100000 particles, translating into thousands of individual chains, ensuring that the slab is wide enough in the y and z dimensions to prevent chain self-interaction and wide enough in the x dimension to minimise the interface-to-bulk ratio.

To quantify phase coexistence, we calculate particle density profiles along the x dimension and identify concentrations in the two spatial domains. To obtain the critical salt concentration, we plot the phase diagrams in the density-salt concentration plane and use the law of rectilinear diameters[7] with the same parameters as previously reported[2] to extrapolate the binodals towards the critical point, in whose vicinity our direct simulations would not be reliable.

In the simulations shown in the snapshot of Fig. S6, we used a cubic box, resulting in a coacervate phase in a droplet-like geometry. This choice was motivated by the need to match the experimental peptide concentrations, which would require a very long simulation box if rectangular. Diffusing the chains across the box and sampling the liquid-phase composition would be expensive under such conditions. It is known that, unlike in slab geometry, in droplet geometry the measured coexistence densities cannot be extrapolated to the macroscopic phase diagrams;[8] therefore, we do not attempt to estimate the CSC in such simulations. The purpose of these simulations is to show that the local densities of the nucleic acids and amino acids within the droplet are approximately the same.

Note on the discrepancy between the simulations and the experiments

While the computer simulations qualitatively reproduce the phase diagram in Fig. 1a, the computational model systematically underestimates the propensity for coacervation and predicts lower critical salt concentrations across all the studied systems. There are two main sources of these discrepancies. First, our model incorporates solvent and salt ions only implicitly via averaged effective interactions. Nevertheless, in highly charged systems like the one studied here, it is expected that electrostatic correlations, counterion condensation, and solvation would play a role, which is not fully captured by our model. Second, in contrast to the experiment, the computational model assumes protected peptide ends, thereby excluding the dipolar contribution of the peptide zwitterionic ends, which stabilises the coacervate phase. The comparison between the experiment and theory underlines the non-negligible contribution of these higher-order electrostatic effects. Another discrepancy arises when looking at how varying peptide length varies the slope of CSC *vs* 1/N. In experimental data, longer peptides yield steeper slopes, whereas in computational data, plots of different peptides show near-parallelism. This result could be due to the virtual protection of the *N-* and *C-termini* in the coarse-grained model.

Another potential source of the discrepancy is that in the simulations, we use [nt]:[aa] = 1, while the experiments use excess of the peptide. Figs. S6 and S7 where we vary [aa] while keeping [nt] constant shows that the density of the nucleotides inside of the coacervate droplet is approximately the same as the density of amino acids, showing that within our simulation model, the coacervation tends to the electroneutrality, and excess of peptide yields condensates of similar compositions across [aa] range. Nevertheless, it is possible that the current one-bead model under-represents the intrinsic valency of arginine, which could otherwise form simultaneous interactions with both the phosphate backbone and nucleobases of RNA, and in experiment, likely an excess of arginine could be partitioned into the coacervate and the electroneutrality would be restored by small ions which would be able to buffer the overall charge in the condensate, and which are represented only implicitly in our simulation.

References

(1) Lu, T.; Nakashima, K. K.; Spruijt, E. Temperature-Responsive Peptide–Nucleotide Coacervates. *J. Phys. Chem. B* **2021**, *125* (12), 3080–3091. https://doi.org/10.1021/acs.jpcb.0c10839.

(2) Tejedor, A. R.; Gonzalez, A. A.; Maristany, M. J.; Chew, P. Y.; Russell, K.; Ramirez, J.; Espinosa, J. R.; Collepardo-Guevara, R. Chemically-Informed Coarse-Graining of Electrostatic Forces in Charge-Rich Biomolecular Condensates. July 27, 2024. https://doi.org/10.1101/2024.07.26.605370.

(3) Wang, X.; Ramírez-Hinestrosa, S.; Dobnikar, J.; Frenkel, D. The Lennard-Jones Potential: When (Not) to Use It. *Phys. Chem. Chem. Phys.* **2020**, *22* (19), 10624–10633. https://doi.org/10.1039/C9CP05445F.

(4) YUKAWA, H. On the Interaction of Elementary Particles. I. THE PHYSICAL SOCIETY OF JAPAN , The Mathematical Society of Japan 1935. https://doi.org/10.11429/ppmsj1919.17.0_48.

(5) 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* (11), 732–743. https://doi.org/10.1038/s43588-021-00155-3.

(6) Ladd, A. J. C.; Woodcock, L. V. Triple-Point Coexistence Properties of the Lennard-Jones System. *Chemical Physics Letters* **1977**, *51* (1), 155–159. https://doi.org/10.1016/0009-2614(77)85375-X.

(7) Rowlinson, J. S. Physics of Liquids: Are Diameters Rectilinear? *Nature* **1986**, *319* (6052), 362–362. https://doi.org/10.1038/319362a0.

(8) Nilsson, D.; Irbäck, A. Finite-Size Shifts in Simulated Protein Droplet Phase Diagrams. *The Journal of Chemical Physics* **2021**, *154* (23), 235101. https://doi.org/10.1063/5.0052813.

Supplementary tables

Yield summary and MS:

library	tare (mg)	gross (mg)	yield (mg)	$M+H^+$ calcd.	MALDI + found	ESI + found
3R-3	11306,11	11384,03	77,92	189,09 – 486,33	n.d.	189,15 – 486,22

AAA:

AA	3R-3 n/n calcd.	3R-3 n/n found
Glu	33,33%	34,77%
Gly	33,33%	35,16%
Arg	33,33%	30,07%

RMSD **2,31%**

Possible 3R-3 masses (variations omitted due to mass overlap)

Combination	Ion formula	$M+H^+$	#
GGG	$C_6H_{13}N_4O_3^+$	189,0982	1
GGE	$C_9H_{17}N_4O_5^+$	261,1193	3
RGG	$C_{10}H_{22}N_7O_3^+$	288,1779	3
GEE	$C_{12}H_{21}N_4O_7^+$	333,1405	3
RGE	$C_{13}H_{26}N_7O_5^+$	360,199	6
RRG	$C_{14}H_{31}N_{10}O_3^+$	387,2575	3
EEE	$C_{15}H_{25}N_4O_9^+$	405,1616	1
REE	$C_{16}H_{30}N_7O_7^+$	432,2201	3
RRE	$C_{17}H_{35}N_{10}O_5^+$	459,2786	3
RRR	$C_{18}H_{40}N_{13}O_3^+$	486,3372	1

Note: Intensity does not meaningfully correlate with calculated abundance due to large differences in ionisation efficiency

Table S1. Critical salt concentrations (CSCs) were measured for peptide mixtures (and respective controls) and Arg-containing heteropeptides with oligonucleotides (5 mM nucleotide) at 25 mM HEPES, pH 7.4, and room temperature.

Composition	**[aa] (mM)**	**CSC (mM)**
R$_3$/DNA$_{20}$	6.67	19.8
R$_3$/DNA$_{20}$	20	85.0
R$_4$/DNA$_{20}$	6.67	174.8
R$_4$/DNA$_{20}$	20	207.4 ± 27.2
R$_5$/DNA$_{20}$	20	348.1
R$_3$/R$_4$/R$_5$/DNA$_{20}$	20	230.5
R$_2$GR$_2$/DNA$_{20}$	25	204.6
RG$_2$R/DNA$_{20}$	13.3	0
R$_4$/RG$_2$R/DNA$_{20}$	13.3	147.1
R$_2$G$_2$R$_2$/DNA$_{20}$	30	237.8
R$_2$G$_4$R$_2$/DNA$_{20}$	20	143.4
R$_2$G$_4$R$_2$/DNA$_{20}$	40	163.3
R$_2$G$_8$R$_2$/DNA$_{20}$	20	85.2
R$_2$G$_8$R$_2$/DNA$_{20}$	60	140.8
R$_2$GR$_2$/DNA$_{20}$	20	163.3
R$_2$G$_2$R$_2$/DNA$_{20}$	20	170.0
R$_4$/R$_2$GR$_2$/R$_2$G$_2$R$_2$/DNA$_{20}$	20	200.1
RKRK/DNA$_{20}$	6.67	50.3
RKRK/DNA$_{20}$	20	91.5
R$_4$/RKRK/DNA$_{20}$	13.3	182.0
K$_4$/DNA$_{20}$	6.67	0
K$_4$/RKRK/DNA$_{20}$	13.3	82.8
R$_4$/K$_4$/DNA$_{20}$	13.3	141.5
RGR/DNA$_{20}$	13.3	0
R$_3$/RER/DNA$_{20}$	13.3	32.1
R$_3$/RER/RGR/DNA$_{20}$	20	40.0
R$_1$/DNA$_8$	20	0
R$_4$/R$_1$(3:1)/DNA$_8$	20	84.3
R$_4$/R$_1$(1:1)/DNA$_8$	20	40.4
R$_1$/DNA$_{20}$	20	0
R$_4$/R$_1$(3:1)/DNA$_{20}$	20	201.9
R$_4$/R$_1$(1:1)/DNA$_{20}$	20	139.8
Ac-R$_4$-NH$_2$/DNA$_8$	20	66.1
Ac-R$_4$-NH$_2$/DNA$_{20}$	20	174.5
{R, G, E,}$_3$/DNA$_{20}$	20	20.0
{R, G, E,}$_3$/{A, C, T, G}$_{20}$	20	81.0
R$_4$/DNA$_{20}$	100	464.2
R$_8$/DNA$_{20}$	100	1160.5
R$_8$/DNA$_8$	100	606.8

Table S2. Critical salt concentrations (CSCs) were measured for mixtures of peptides and oligonucleotides (phase diagram studies) at 25 mM HEPES, pH 7.4, and room temperature. Concentrations refer to amino acid (aa) or nucleotide (nt) concentrations.

Mixture	[aa] (mM)	[nt] (mM)	CSC (mM)
R_4/DNA$_8$	2	5	0
	3	5	22.3
	5	5	49.4
	10	5	67.3
	20	5	99.3 ± 4.7
	40	5	110.0
	20	1	0
	20	2	51.8
	20	5	99.3 ± 4.7
	20	10	122.7
	20	15	118.6
	20	20	140.1
R_3/DNA$_{12}$	5	5	0
	10	5	0
	15	5	25.0
	20	5	38.8 ± 4.7
	30	5	36.7
	40	5	20.7
	60	5	47.0

Mixture	[aa] (mM)	[nt] (mM)	CSC (mM)
R_4/DNA$_{20}$	3	5	93.0
	5	5	123.5
	10	5	175.9
	15	5	213.8
	20	5	201.6 ± 19.3
	30	5	270.3
R_8/DNA$_{20}$	3	5	391.5
	5	5	503.8
	10	5	503.0
	15	5	609.0
	20	5	684.2 ± 55.4
	30	5	674.8

Table S3. Minimal amino acid concentration for the coacervation of dipeptides with oligonucleotides of different lengths. *N/A* stands for ‘non-applicable’.

Peptide dimer	**Oligonucleotide**	**Phase**	**Amino acid concentration required**
R_2	DNA_8	Soluble	*N/A*
	DNA_{12}	Soluble	*N/A*
	DNA_{20}	Soluble	*N/A*
	RNA_8	Soluble	*N/A*
	RNA_{12}	**Droplets**	60 mM
	RNA_{20}	**Droplets**	40 mM
RG, GR	DNA_8	Soluble	*N/A*
	DNA_{12}	Soluble	*N/A*
	DNA_{20}	Soluble	*N/A*
	RNA_8	**Droplets**	40 mM
	RNA_{12}	**Droplets**	40 mM
	RNA_{20}	**Droplets**	20 mM
RE	DNA_8	Soluble	*N/A*
	DNA_{12}	**Droplets**	40 mM
	DNA_{20}	**Droplets**	40 mM
	RNA_8	**Droplets**	20 mM
	RNA_{12}	**Droplets**	20 mM
	RNA_{20}	**Droplets**	20 mM

Supplementary Figures

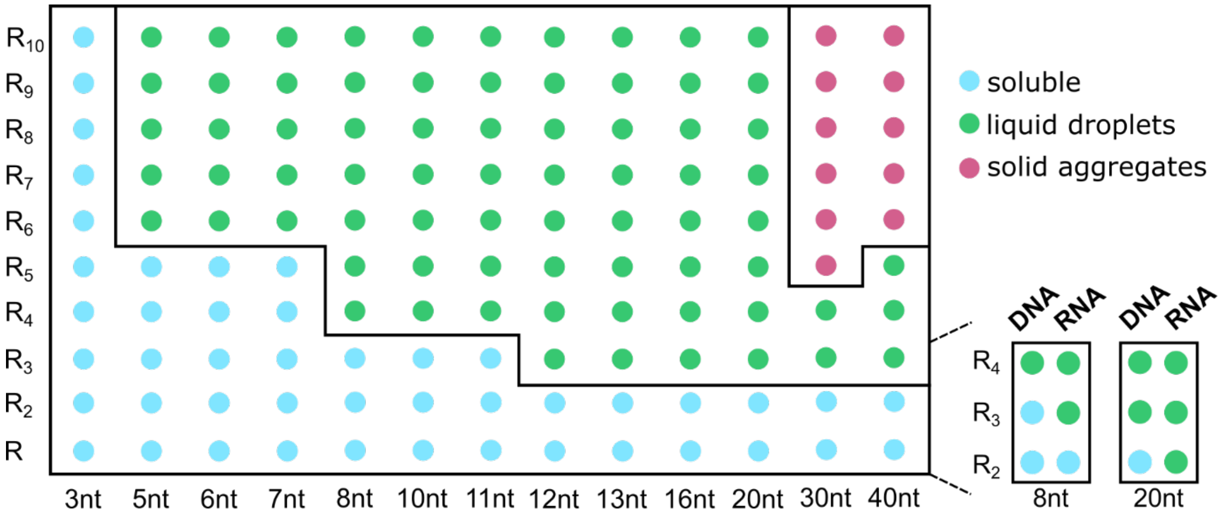


Fig. S1. Extended screening of peptide/DNA mixtures obtained by varying both peptide (vertical axis) and DNA lengths (horizontal axis). DNA sequences follow the motif ACTG, except for the 3nt-motif, which is dA_3. The inset to the right shows the screening results with RNA oligonucleotides alongside those obtained with DNA oligonucleotides of the same length (and motif). The “liquid droplets” region (green) expands as RNA increases. All mixtures were screened at 20 mM amino acid, 5 mM nucleotide, 25 mM HEPES (pH 7.4), and room temperature.

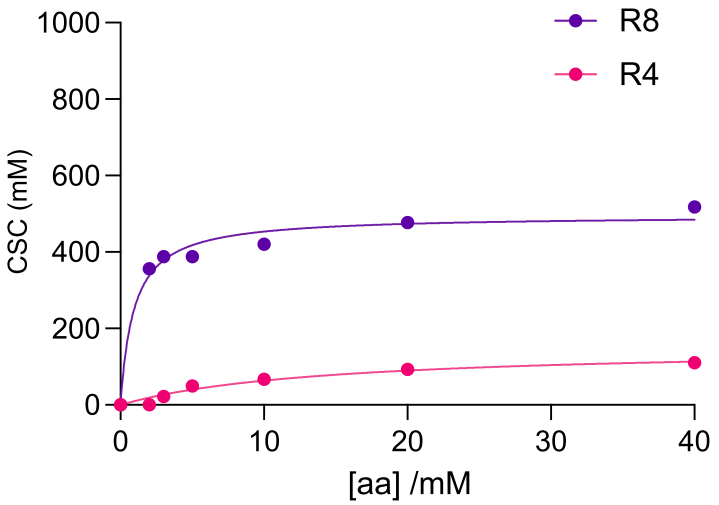


Fig. S2. Experimental CSC of R_4 and R_8/DNA_8 coacervates as a function of amino acid concentration ([aa]). The data were fitted to a one-site binding hyperbola using the GraphPad curve fitting tool.

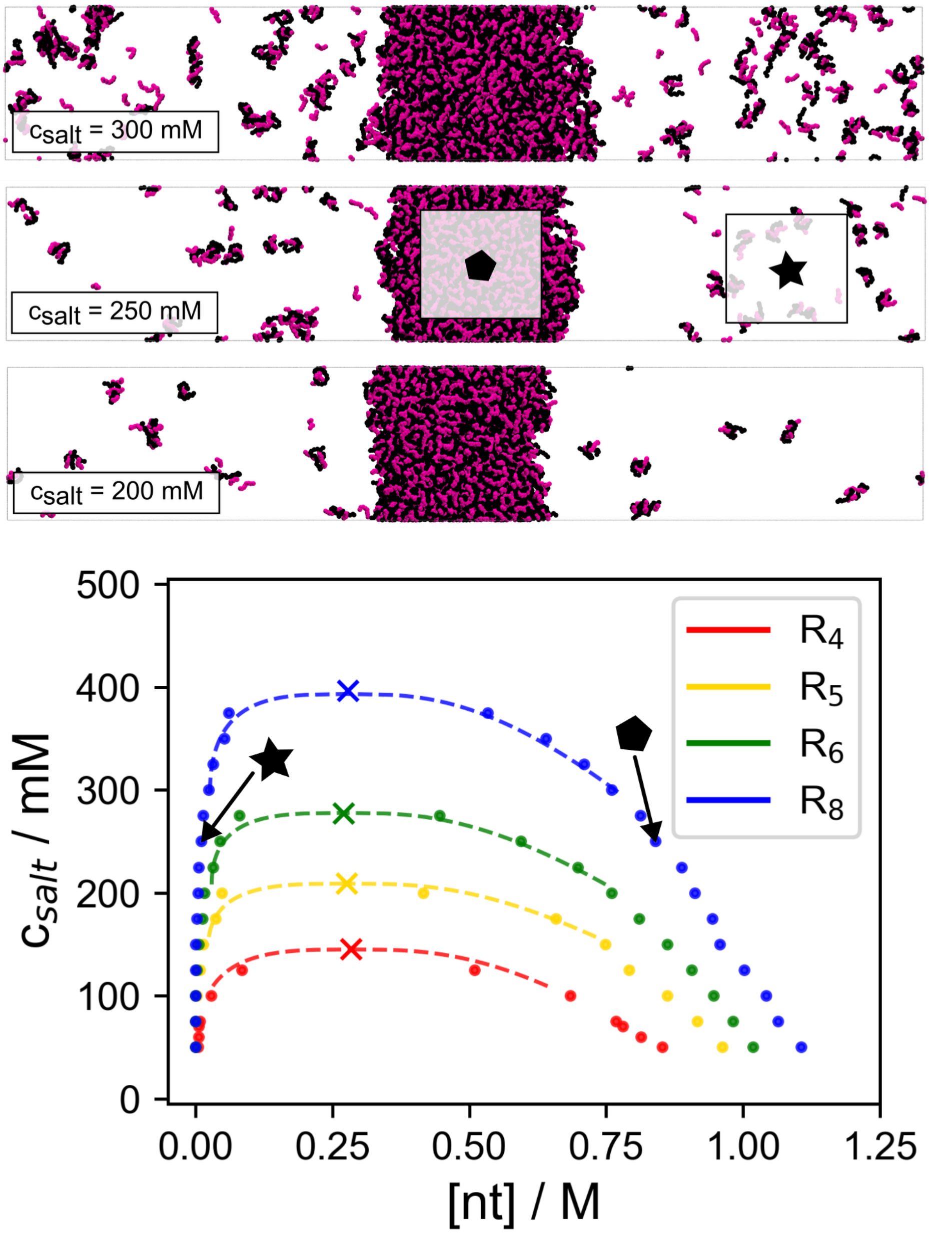


Fig. S3. Top: Snapshots of the simulations for assessing the critical salt concentrations, the two transparent boxes inside 250mM schematically depict the part of the space which corresponds to the coacervate and the supernatant respectively; bottom: phase diagram obtained from the simulation (points), fits (dashed lines) extrapolating the data towards the critical point (cross), the two points for 250mM correspond to the dense phase (pentagon) and dilute phase (star) in the snapshots above.

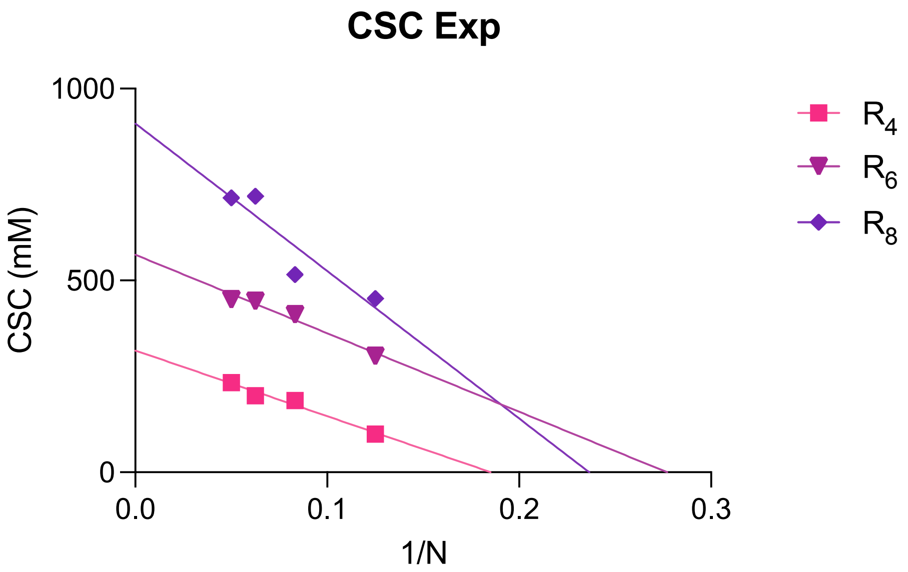


Fig. S4. Experimentally obtained critical salt concentration (CSC) for fixed peptide length and varying DNA length (linearised form, 1/N). The data was fitted to a linear equation.

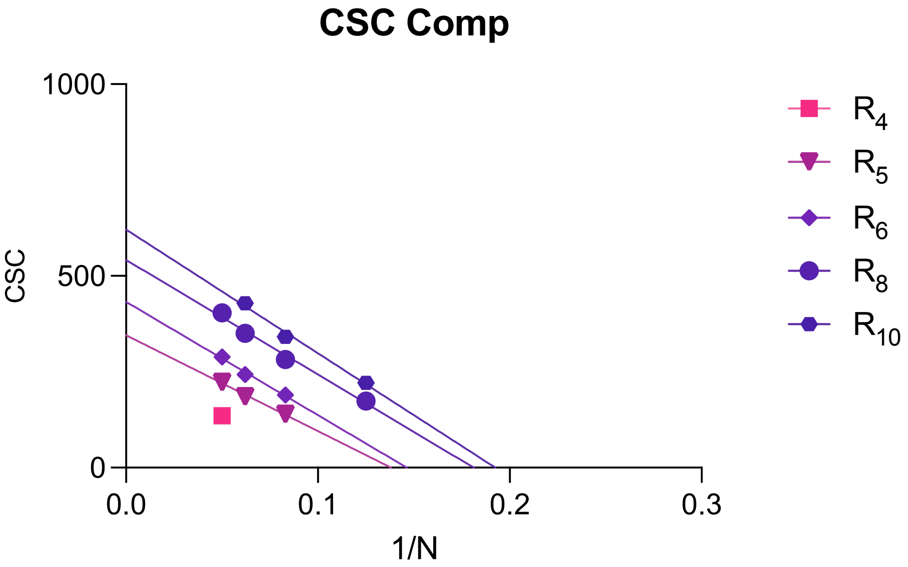


Fig. S5. Computationally obtained CSC for fixed peptide length and varying DNA length (1/N in the linearised plot).

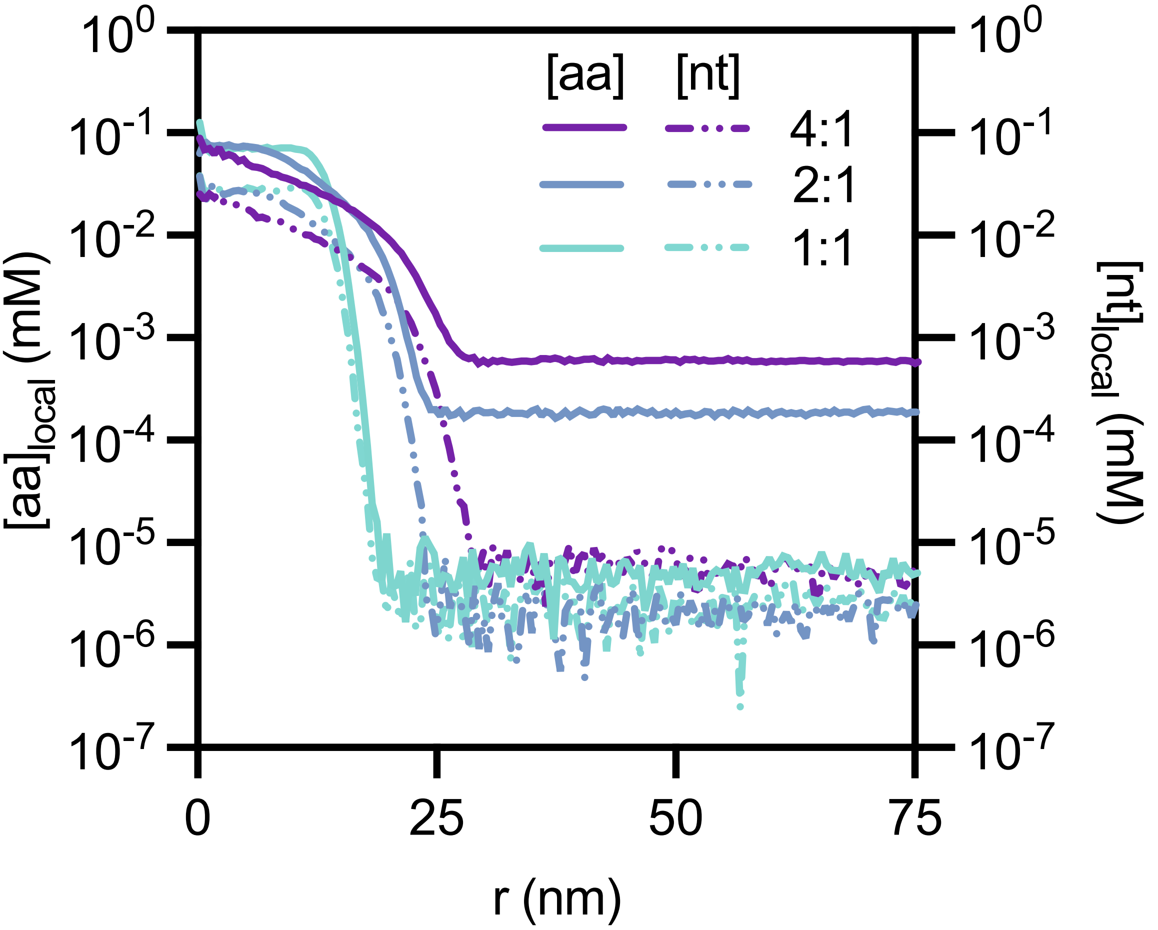


Fig. S6. Profiles of local concentration of nucleotides in the distance R from the centre of the spherical droplets in coacervates at different ratios (1:1, 2:1 and 4:1 [aa]:[nt]).

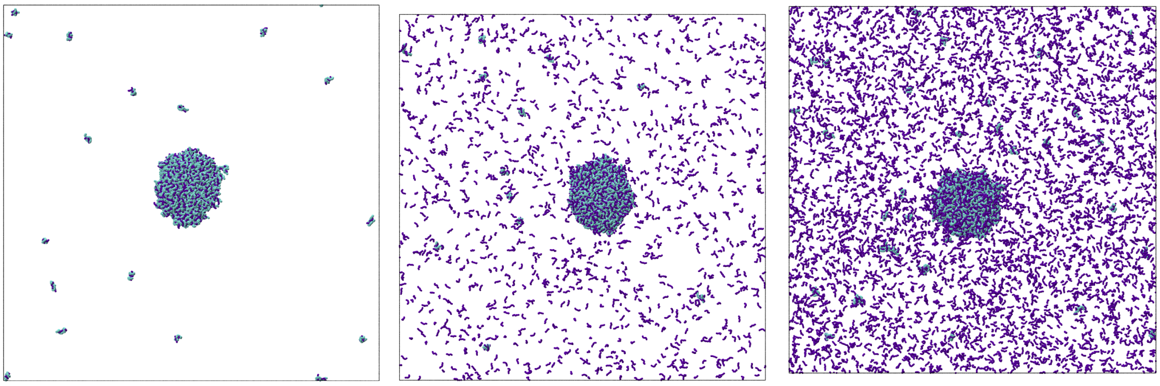


Fig. S7. Caption: Snapshots of coacervate droplets of RNA_20 (blue) and R_8 (violet) at c_salt = 50mM, [nt] = 5mM and [aa] = 5mM, 10mM and 20mM on the respective panels.

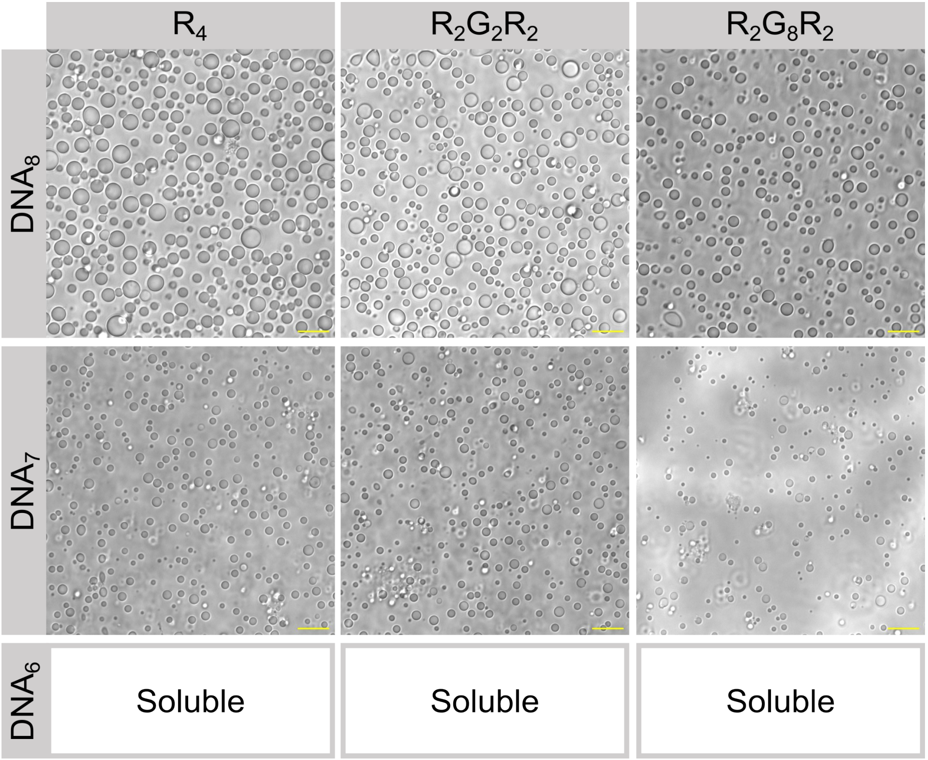


Fig. S8. Empirical screening of minimal DNA oligonucleotide length required to form coacervates with a series of $R_2G_NR_2$ peptides (N = 0, 2, 8). Scale bar = 10 μm, brightfield microscopy.

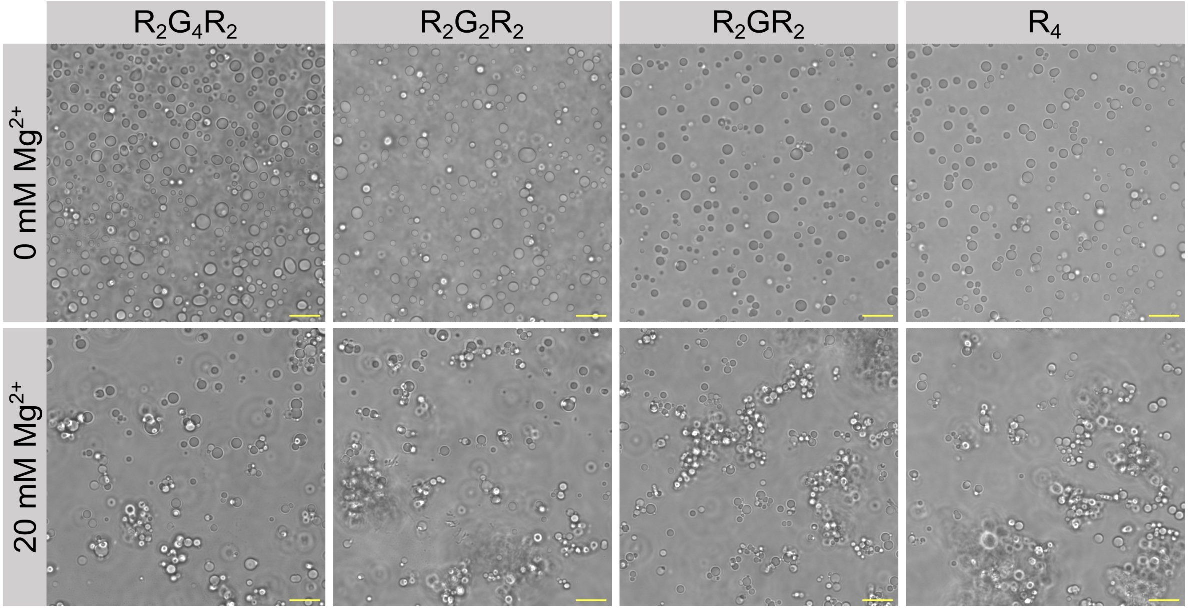


Fig. S9. $R_2G_NR_2/DNA_{20}$ coacervates at 0 and 20 mM Mg^{2+} concentrations. Scale bar = 10 μm, brightfield microscopy.

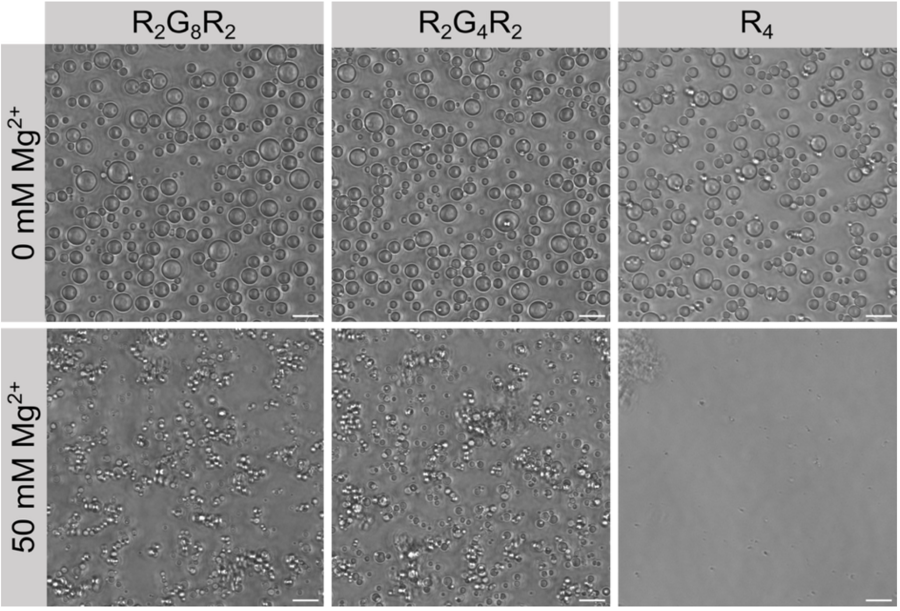


Fig. S10. $R_2G_NR_2$/DNA_{20} coacervates without and at 50 mM Mg^{2+} concentration. Scale bar = 10 µm, brightfield microscopy.

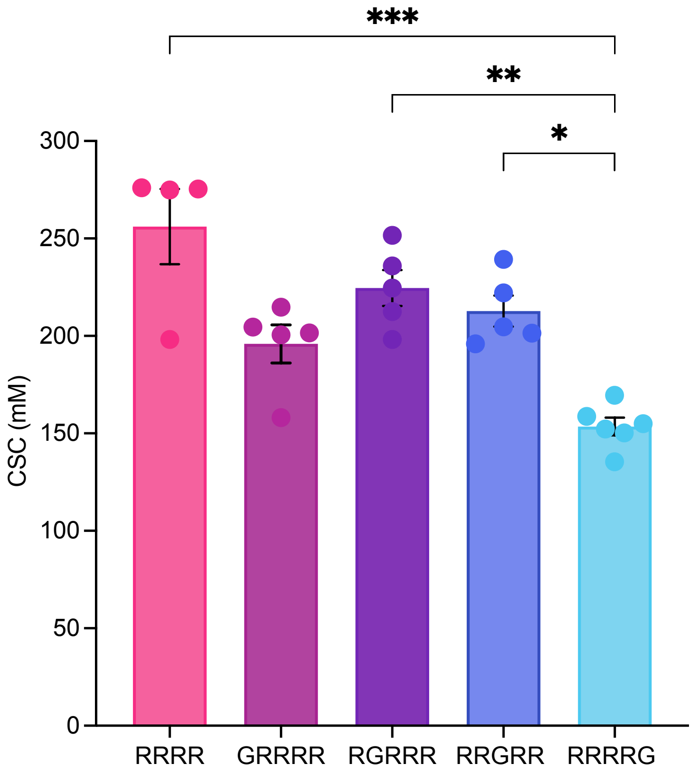


Fig. S11. CSC of R_4G_1/DNA_{20} coacervates varying the position of the G from N-terminus to C-terminus at various places. Only t-tests showing $p<0.05$ are shown. [aa] = 20 mM, [nt] = 5 mM in HEPES 25 mM pH 7.5

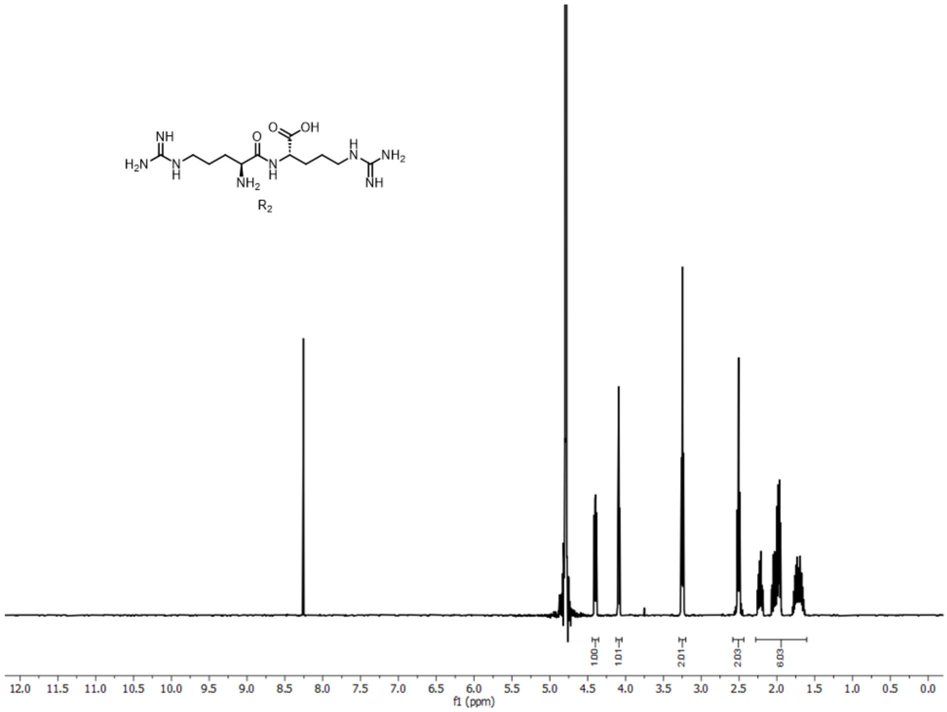


Fig. S12. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **R_2** synthesised in-house.

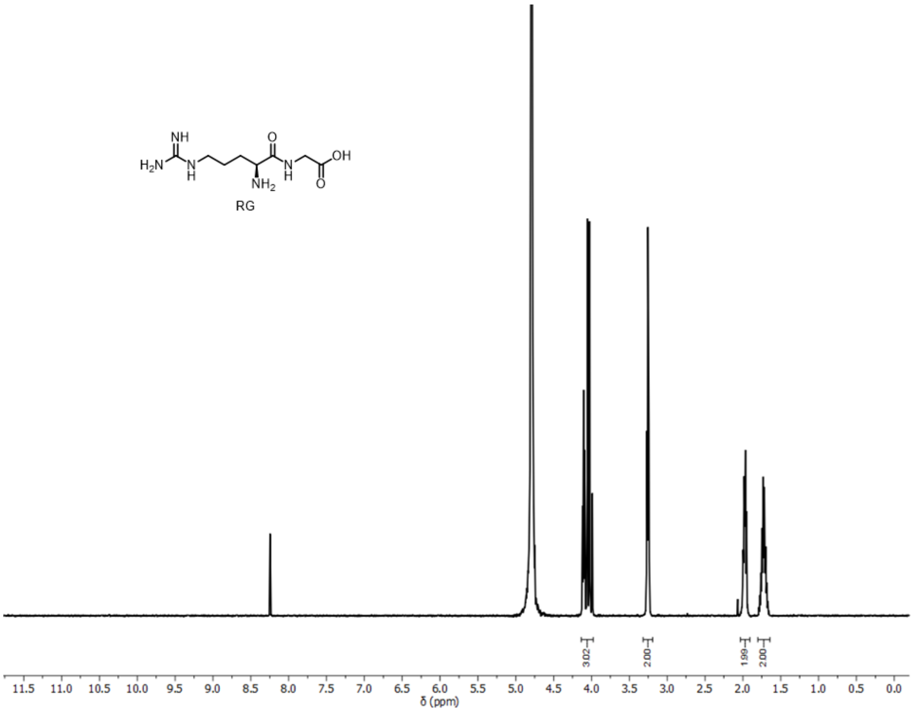


Fig. S13. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **RG** synthesised in-house.

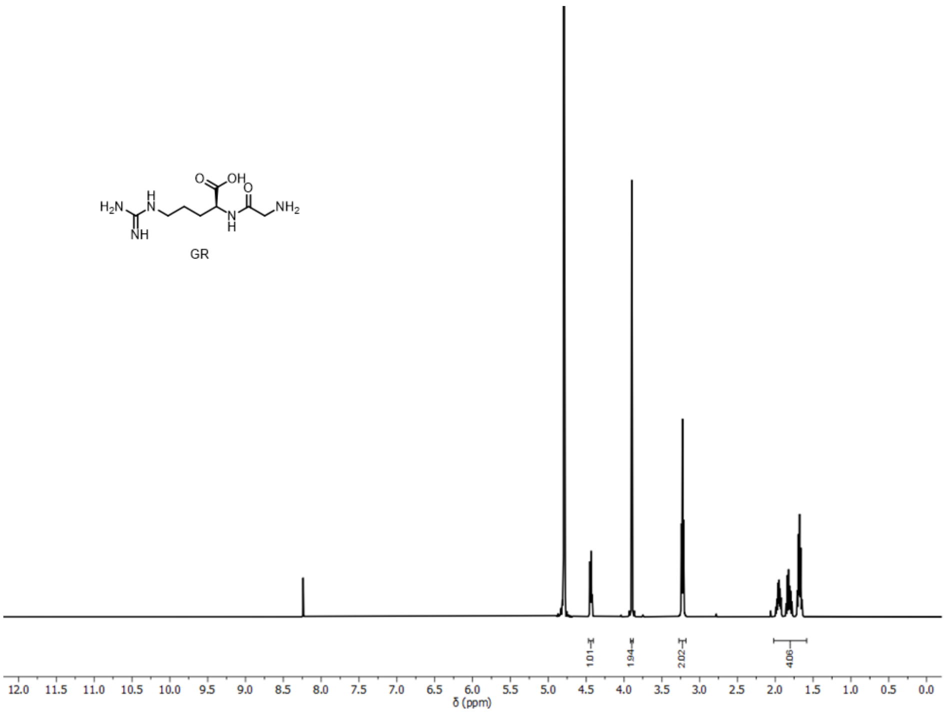


Fig. S14. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **GR** synthesised in-house.

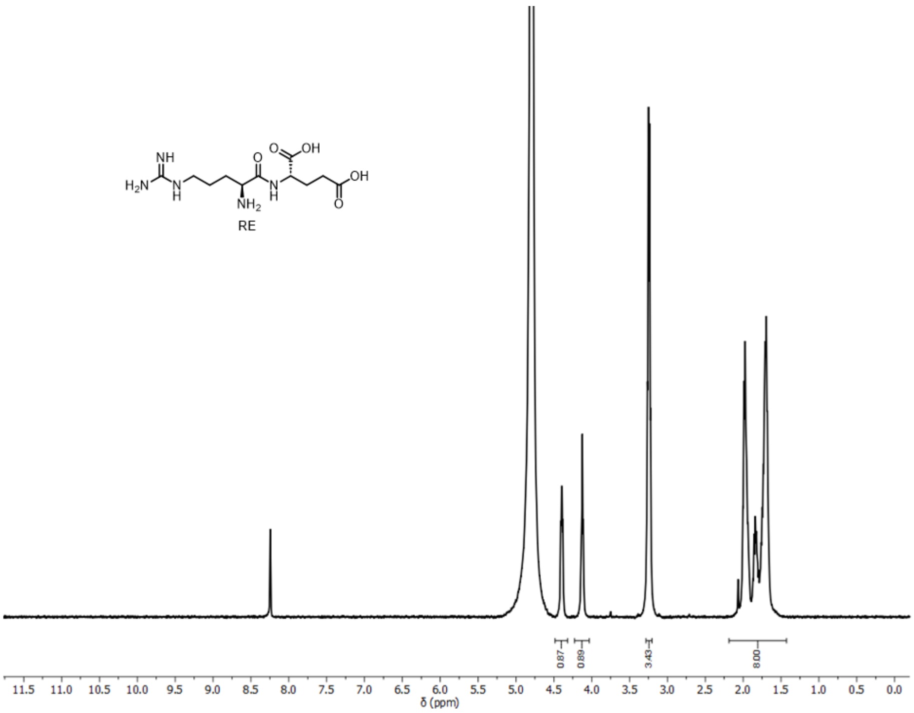


Fig. S15. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **RE** synthesised in-house.

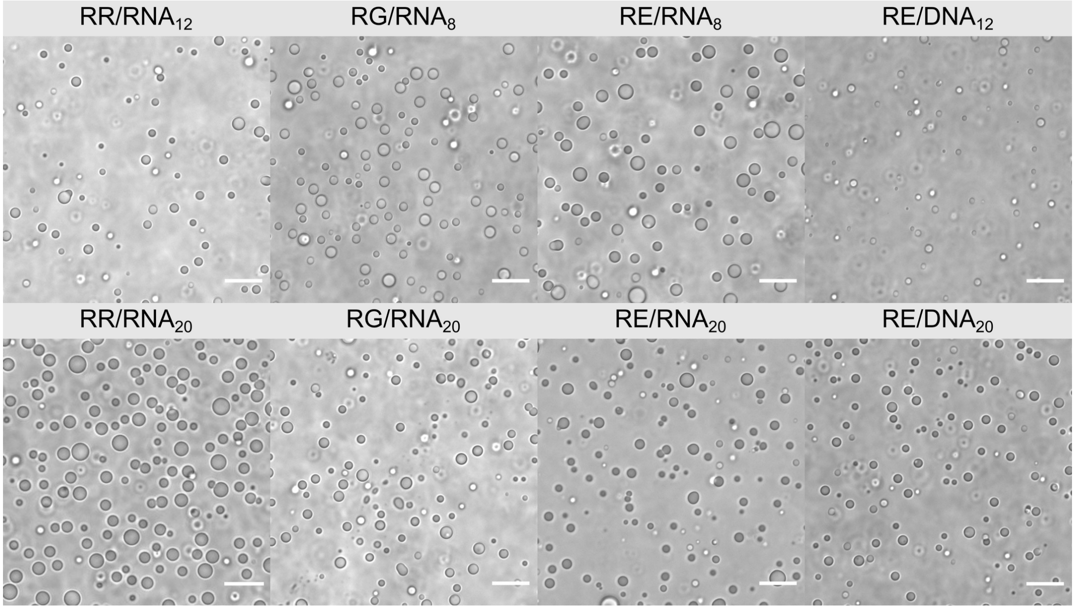


Fig. S16. Brightfield microscope images of coacervates composed of peptide dimers and oligonucleotides (**DNA** or **RNA**, 8-20 nt). Required concentrations of the components are listed in **Table S3**. Scale bar = 10 µm.

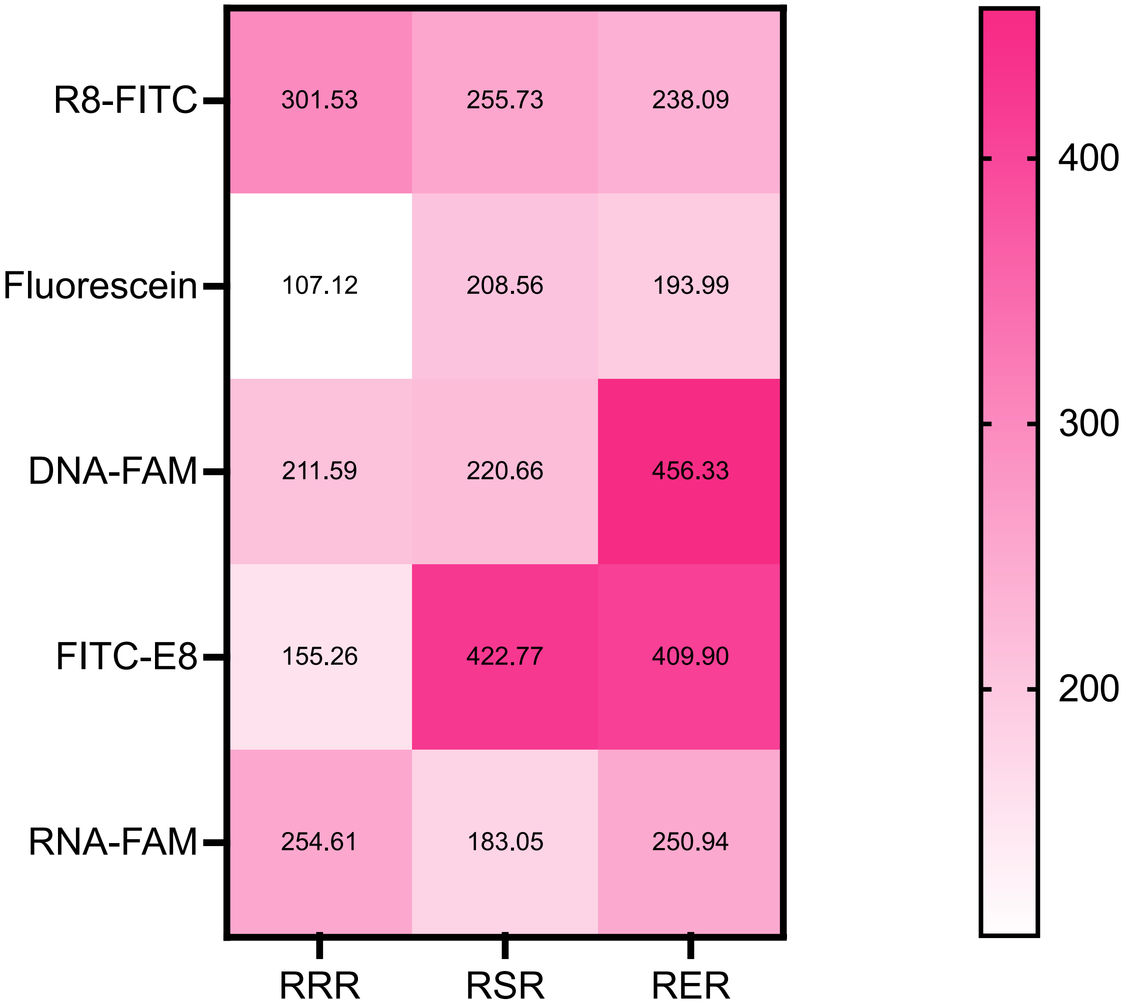


Fig. S17. Kp of various probes (R8-FITC, Fluorescein, DNA-FAM, E8-FITC, RNA-FAM) in RXR/DNA$_{20}$ coacervates. (X = R, S, E). E_8 = FITC-EEEEEEEE; R_8 = FITC-RRRRRRRR; $R_2G_4R_2$ = FITC-RRGGGGRR; RNA$_8$ = FAM-ACUGACUG; DNA$_8$ = FAM-ACTGACTG

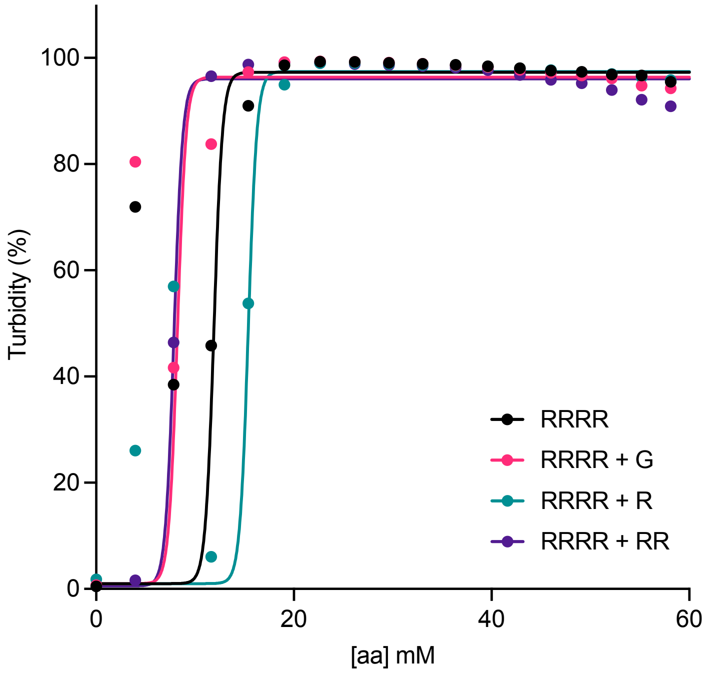


Fig. S18. Turbidity of onset measurements of R_4 in solution containing DNA_{20} and either Glycine, Arginine or di-arginine. The data were fitted to a simple sigmoidal using the GraphPad Prism curve-fitting tool.

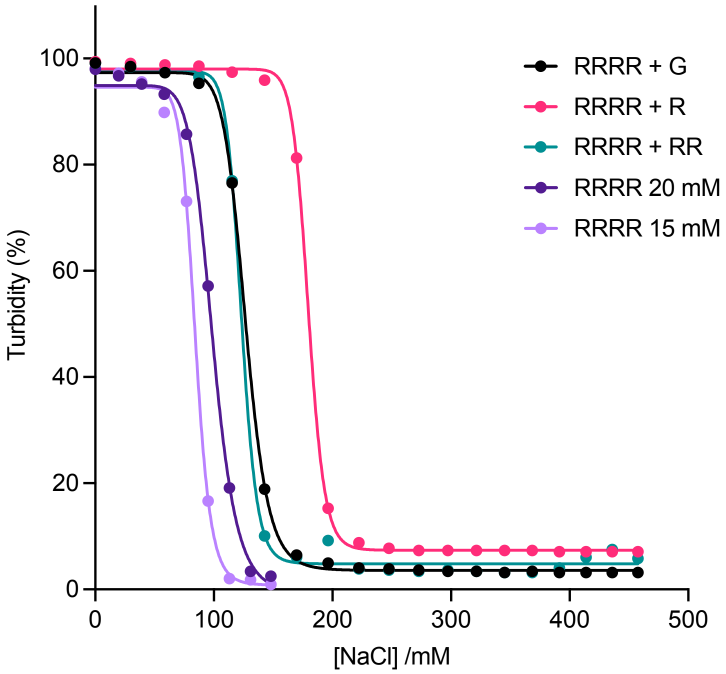


Fig. S19. Turbidity profile for CSC measurements of R4/DNA20 in the presence of either Glycine, Arginine or di-arginine. The data were fitted to a simple sigmoidal using the GraphPad Prism curve-fitting tool.

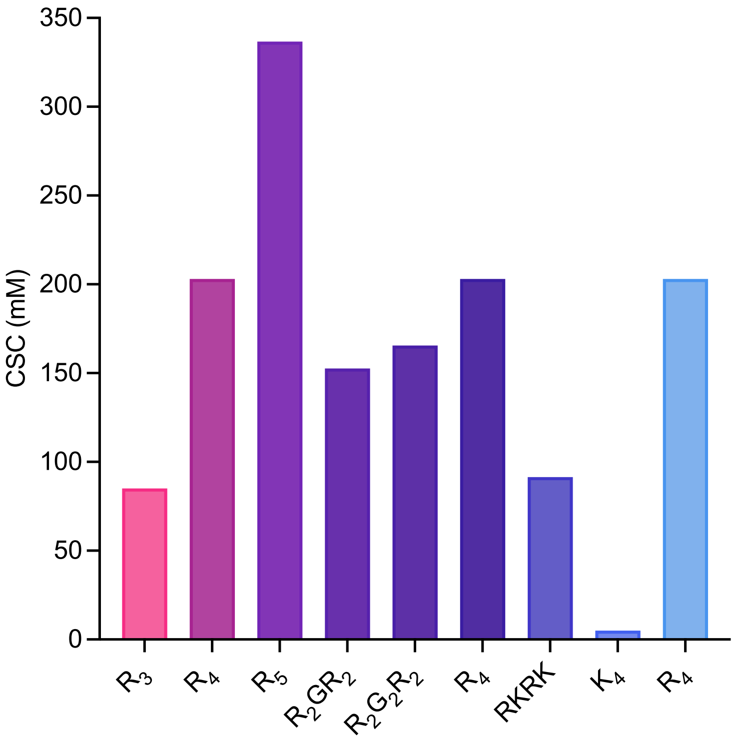


Fig. S20. CSC of peptide/DNA coacervates with the peptides corresponding to the composition of mixtures A, B and C. These values were used to calculate the numerical average of CSC in Fig. 3b. [aa] = 20 mM, [nt] = 5 mM.

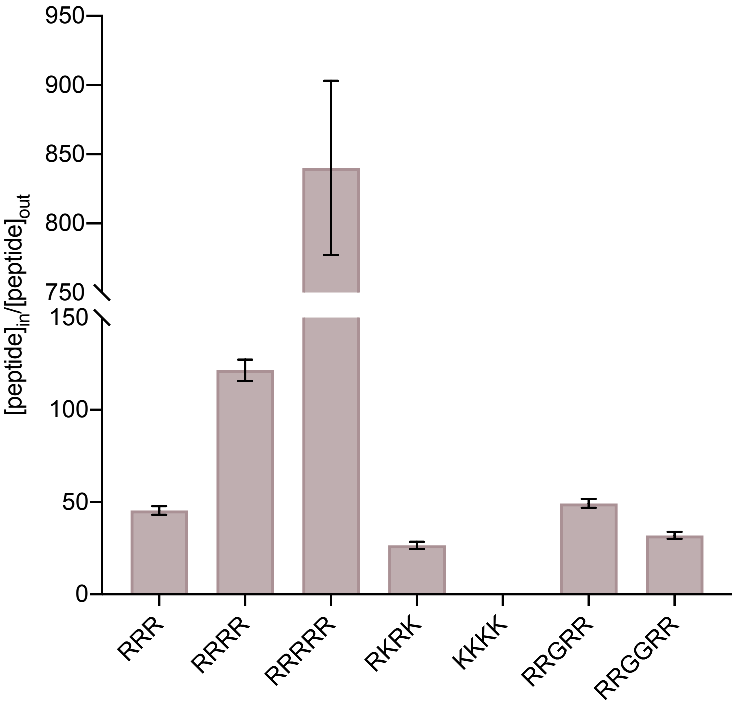


Fig. S21. Computationally obtained partition coefficients of peptides in coacervates composed of the respective peptide and DNA.

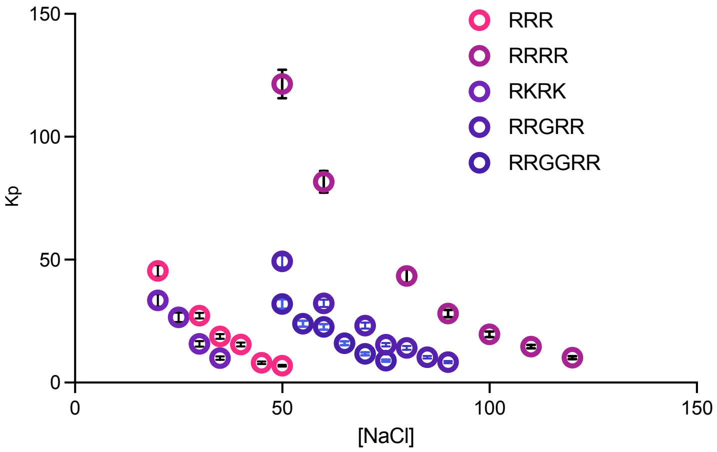


Fig. S22. Computationally obtained partition coefficient of peptides in single peptide/DNA systems as a function of salt concentration.

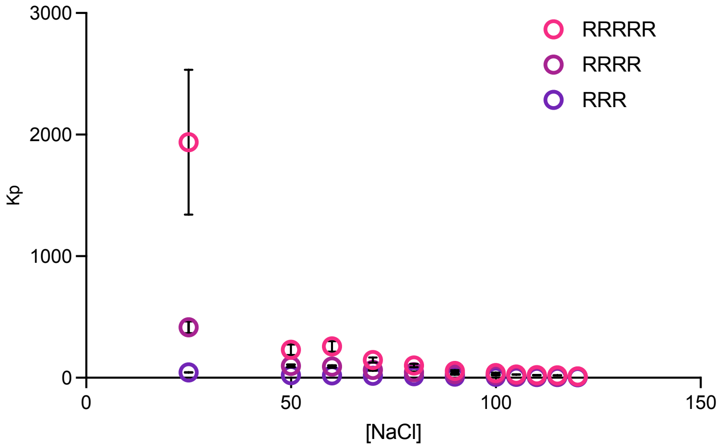


Fig. S23. Computed partitioning of peptides in mixture A.

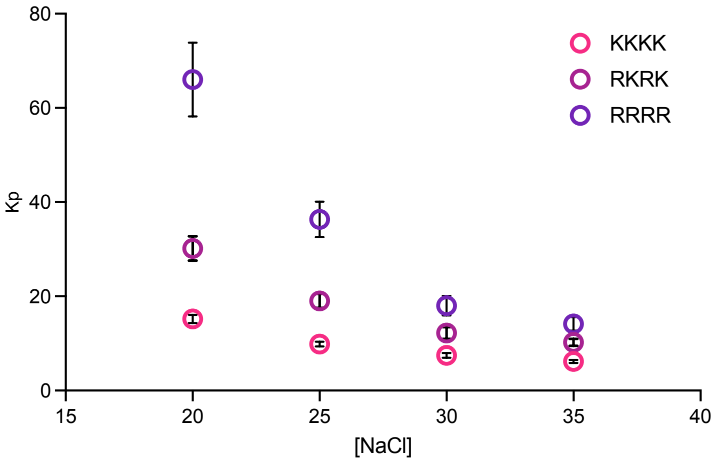


Fig. S24. Computed partitioning of peptides in mixture B.

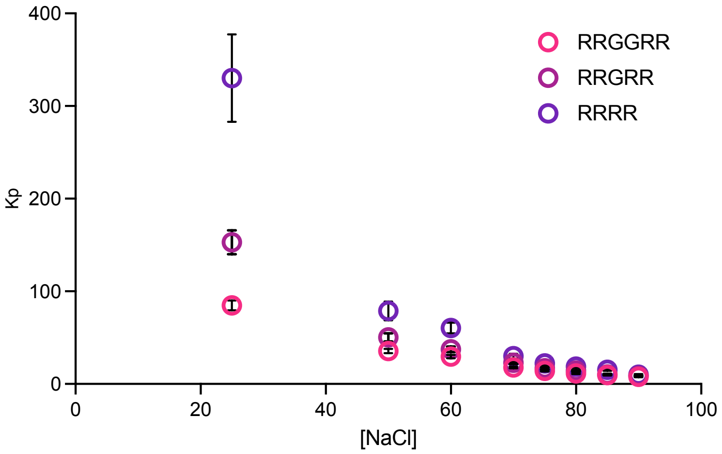


Fig. S25. Computed partitioning of peptides in mixture C.

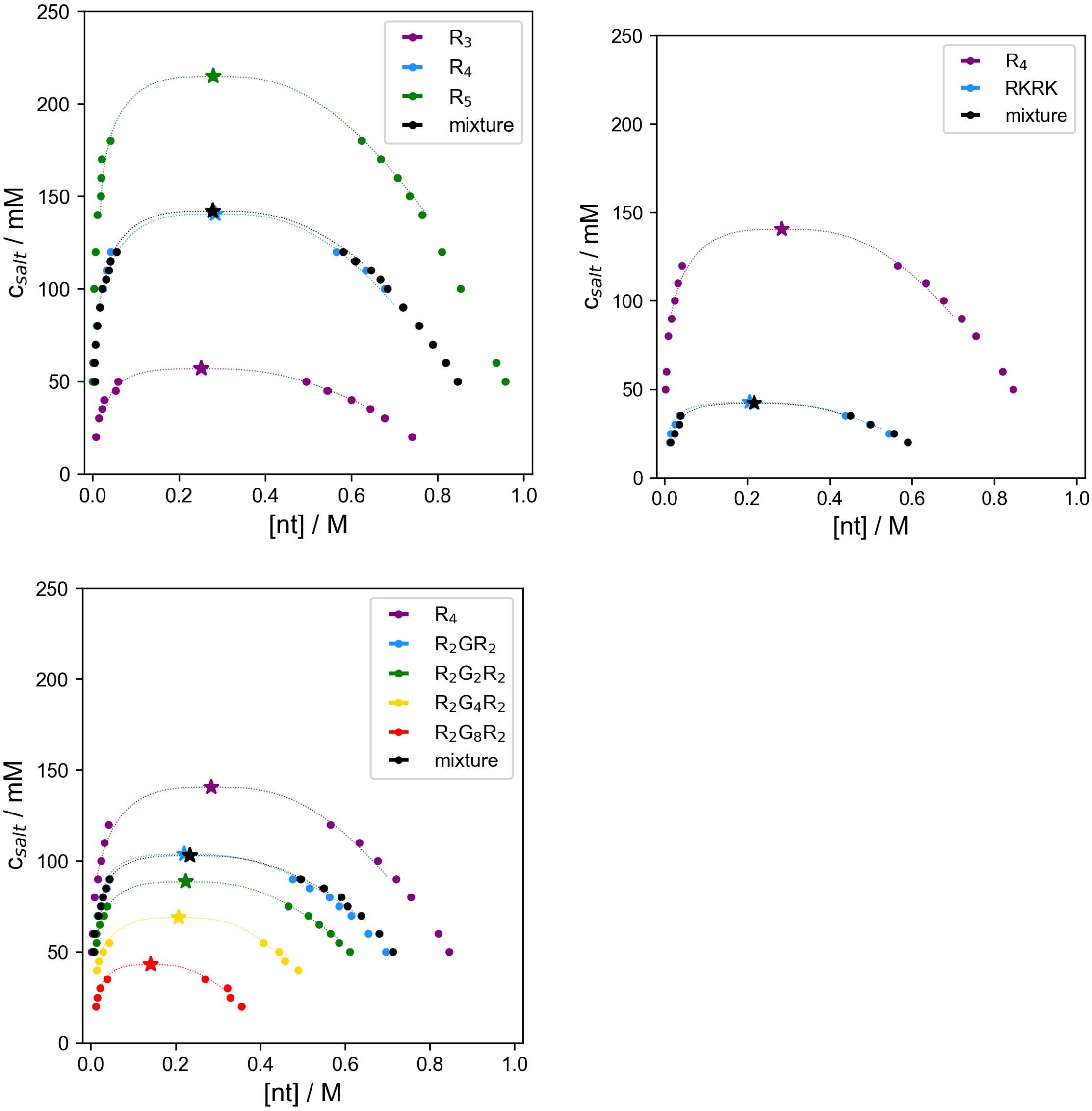


Fig. S26. Phase diagrams of RNA with peptide mixtures and individual constitutive peptides. Points are the co-existing points from the simulations, dotted lines are binodal envelopes as fitted using the law of rectilinear diameters, and the star denotes the critical point.

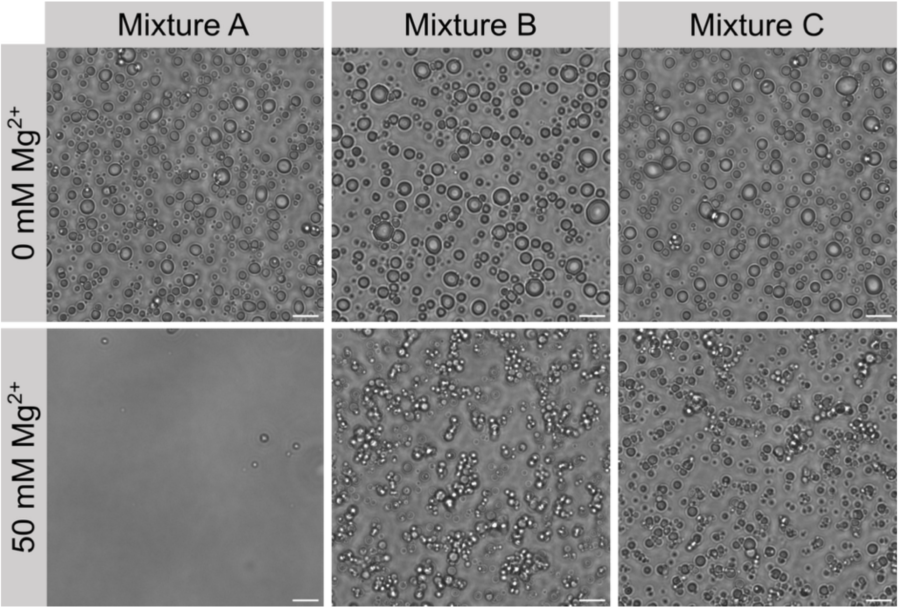


Fig. S27. Brightfield microscope images of mixtures A, B, and C in the presence of 0 and 50 mM $MgCl_2$. Scale bar =10 µm

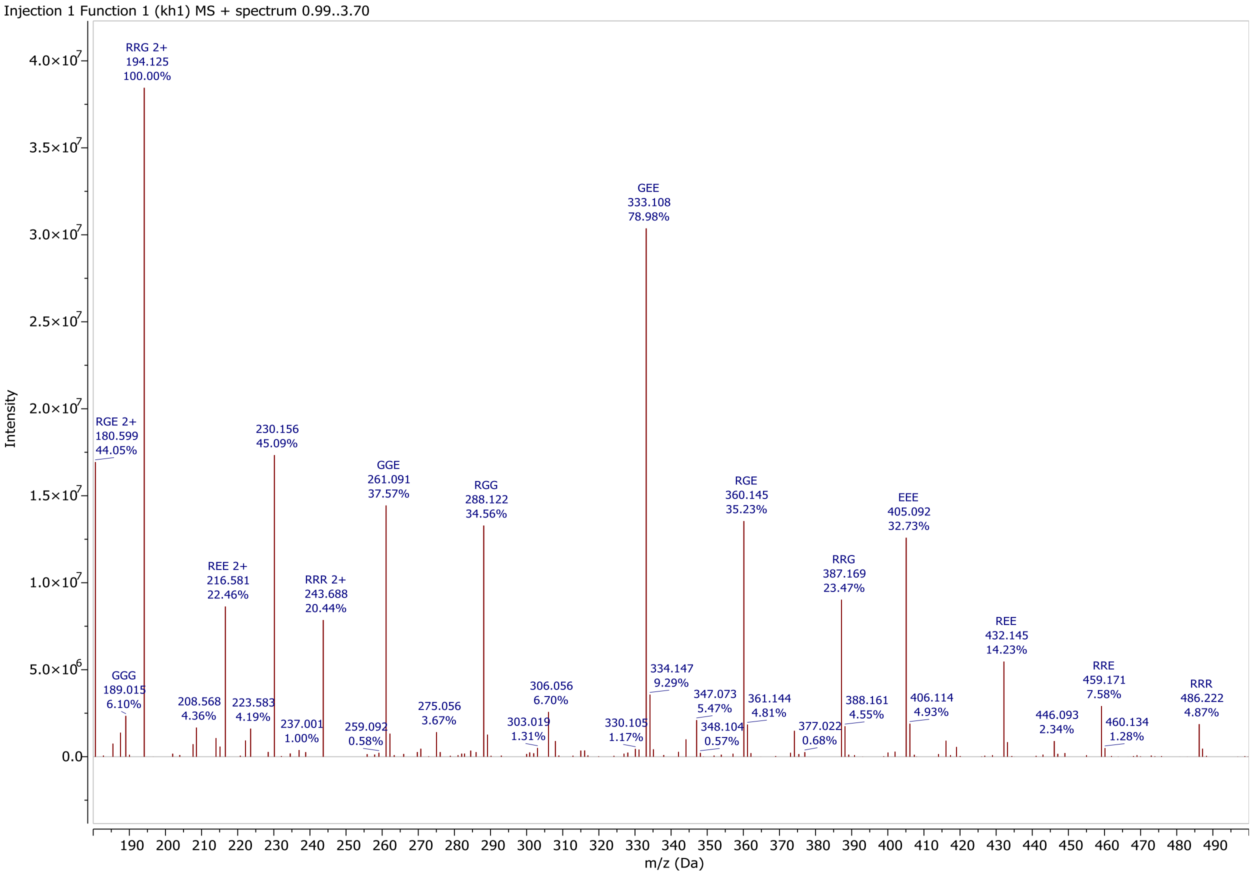


Fig. S28. Mass spectrum from the library in the 180-490 Da range, showing peaks corresponding to expected peptides in the library

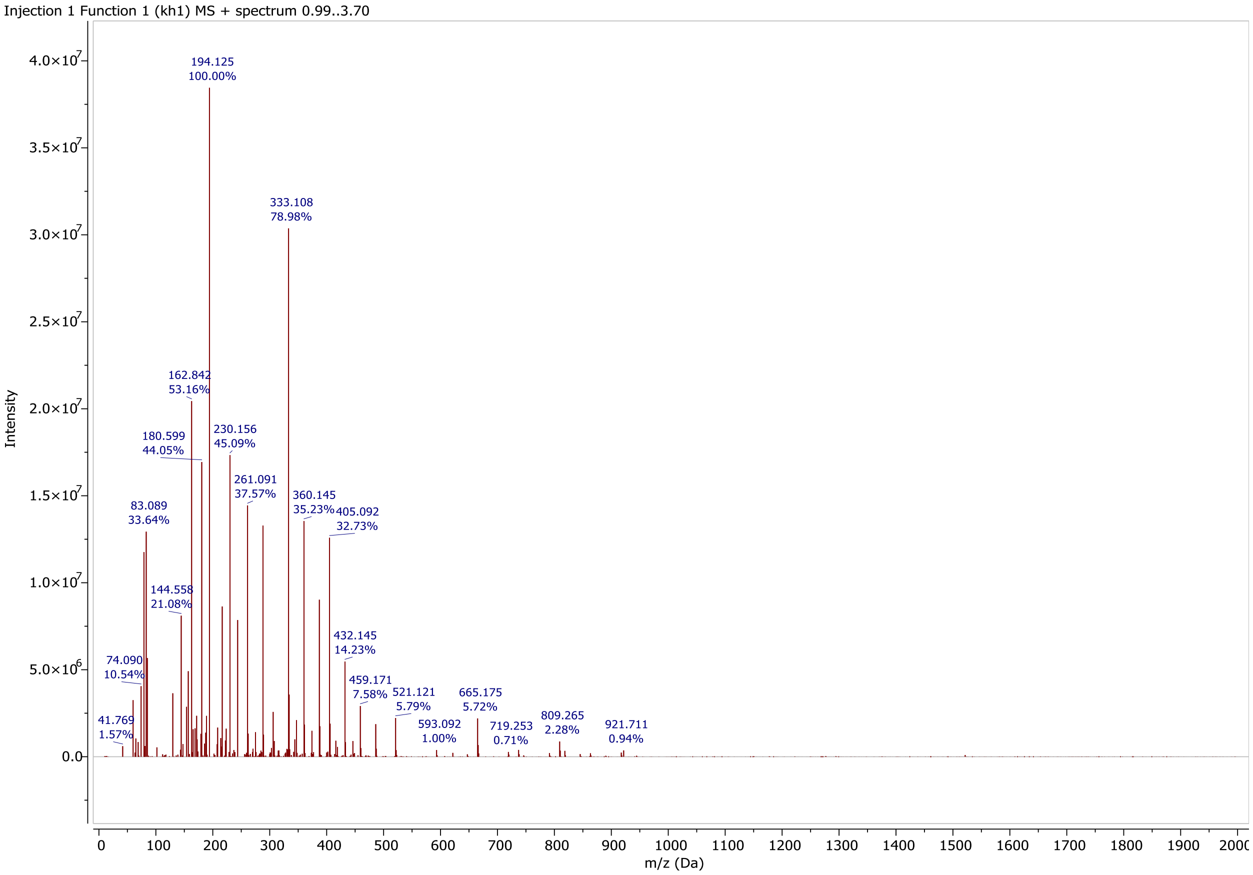


Fig. S29. Full-length mass spectrum of the peptide library.

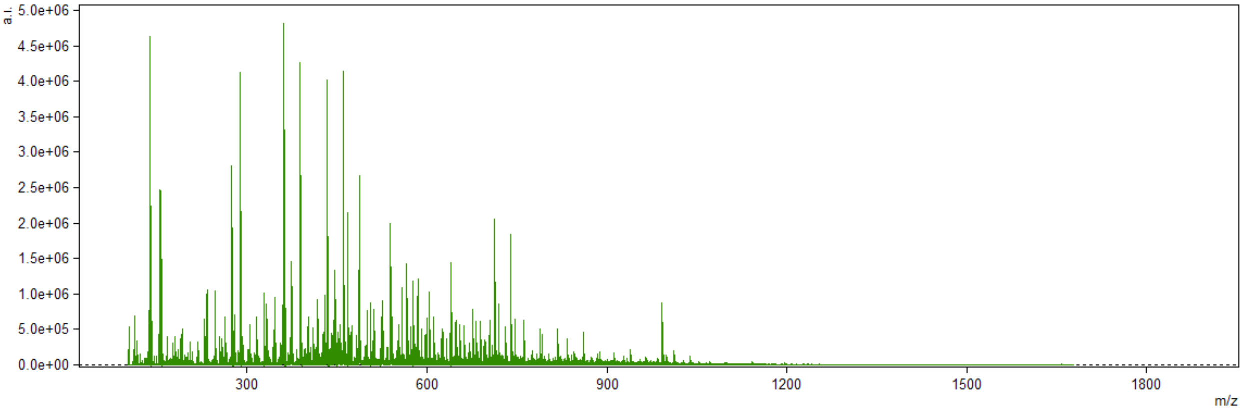


Fig. S30. GC analysis of the library

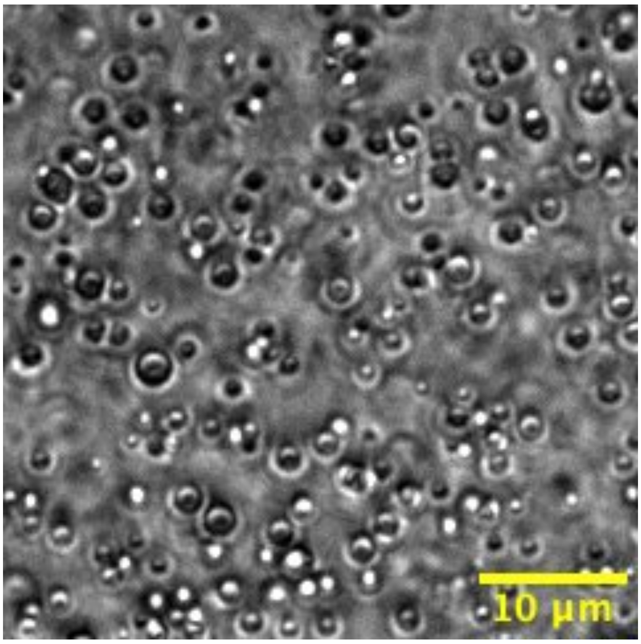


Fig. S31 – Brighfield microscope image of aggregates forming with R_3/RNA_{12} at 20:5 [aa]:[nt] in 25 mM HEPES pH 7.5. Scale bar =10 µm

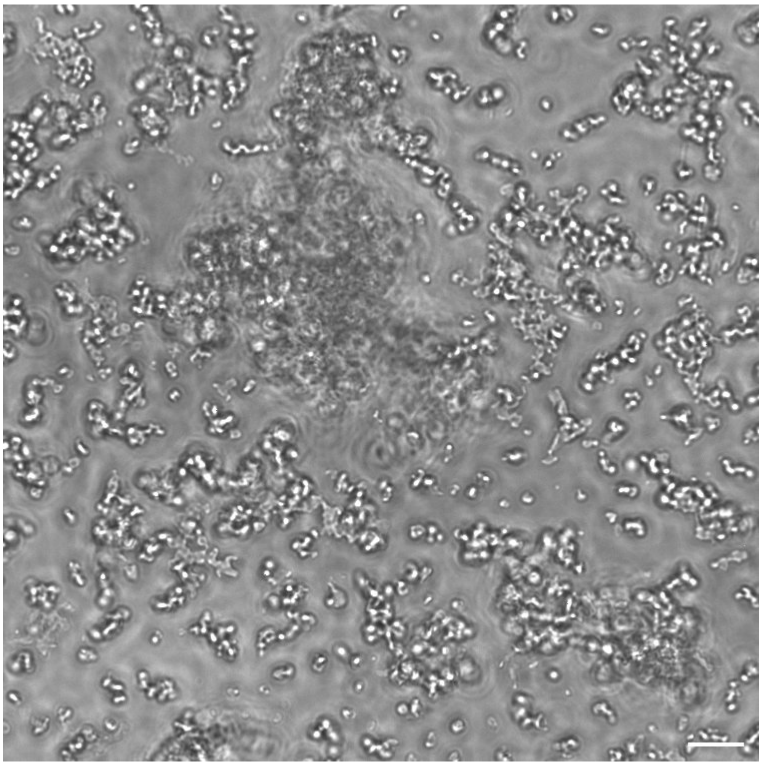


Fig. S32 – Brighfield microscope image of aggregates forming with R_3/$dsDNA_{20}$ at 20:5 [aa]:[nt] in 25 mM HEPES pH 7.5. Scale bar =10 µm

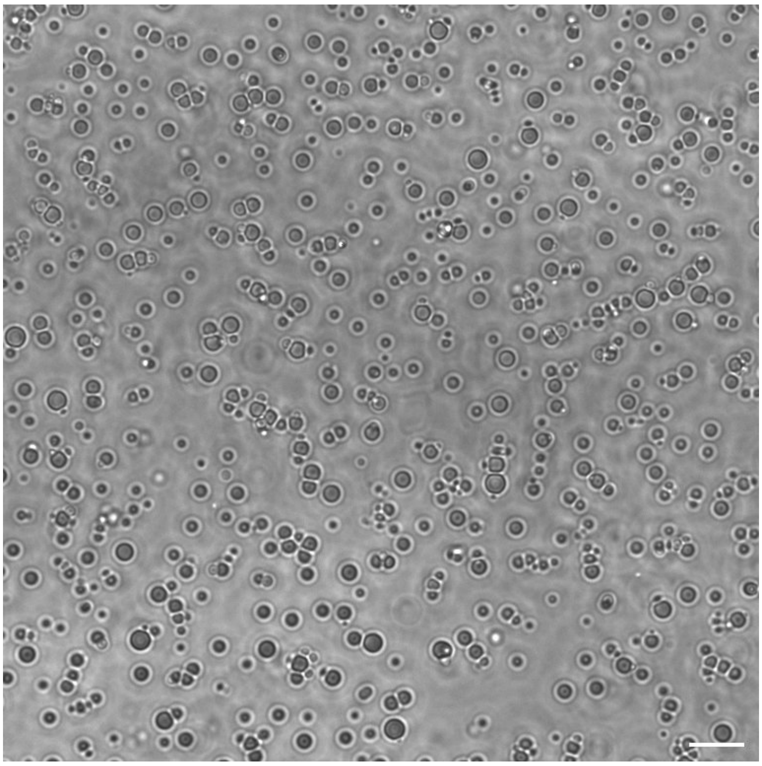


Fig. S33 – Brighfield microscope image of coacervates forming with {R, G, E}$_3$/dsDNA$_{20}$ at 160:5 [aa]:[nt] in 25 mM HEPES pH 7.5. Scale bar =10 µm

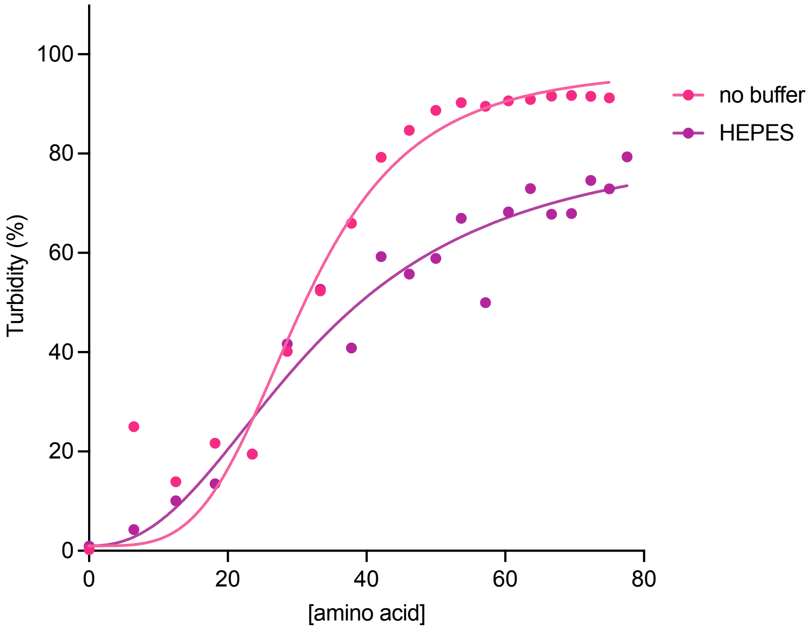


Fig. S34. Turbidity of onset measurements of {R, G, E} peptide library in solution containing DNA$_{20}$ at 5mM [nt] in the presence or absence of HEPES 25 mM, pH 7.5. The data were fitted to a simple sigmoidal using the GraphPad Prism curve-fitting tool.

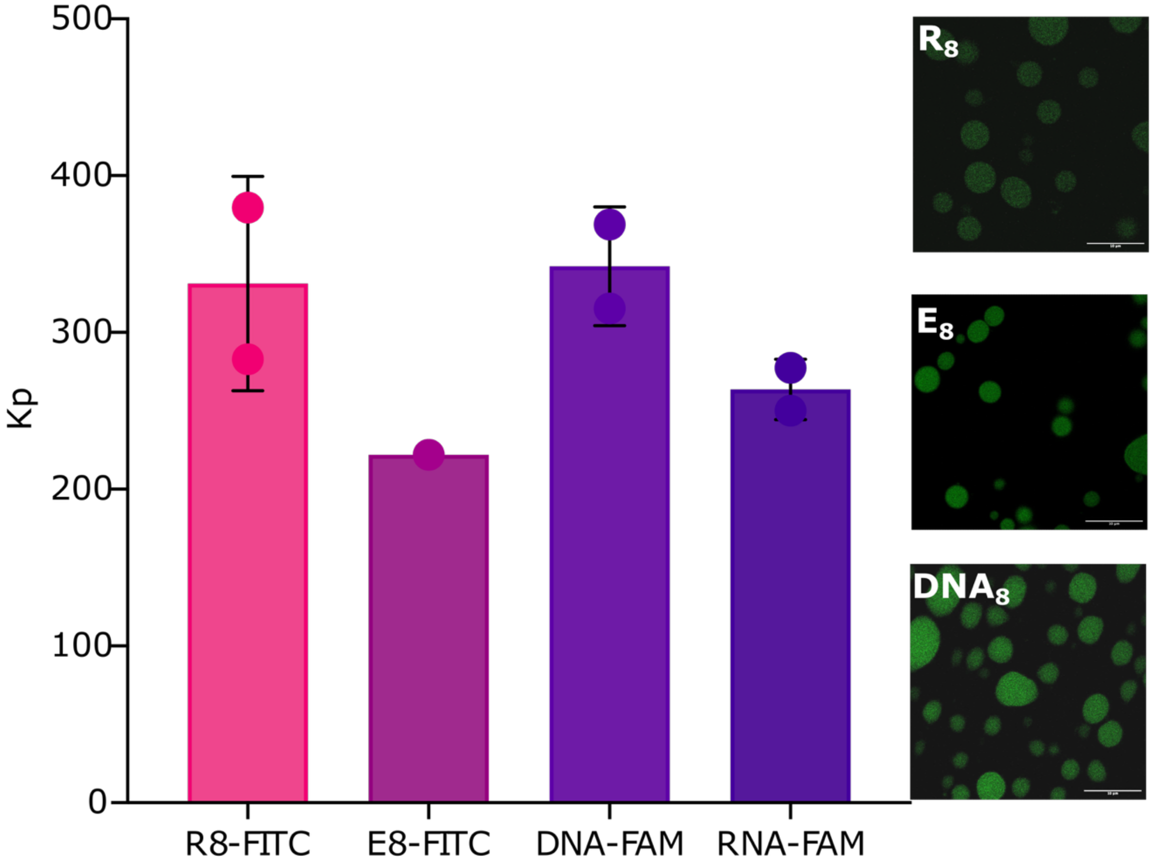


Fig. S35. Partitioning of different probes in {R, G, E}$_3$/DNA$_{20}$ coacervates. Left: Kp values for R$_8$-FITC, E$_8$-FITC, DNA$_8$-FAM and RNA$_8$-FAM in {R, G, E}$_3$/DNA$_{20}$ coacervates. Right: Confocal microscope images of R$_8$-FITC, E$_8$-FITC, DNA$_8$-FAM in {R, G, E}$_3$/DNA$_{20}$ coacervates showing higher fluorescence intensity y within the droplet. [aa] = 100 mM, [nt] = 5 mM, [probe] = 5 µM, HEPES 25 mM pH 7.5. Error bars represent SD for n = 2. Scale bar = 10 µm