Compositional diversity of minimal coacervates in a nucleic acid-peptide world

Karina K. Nakashima,[1,2,†] Fatma Zohra Mihoubi,[1,2,†] Kieran O. Russell,[3] Fidan Rahmatova,[1,2] Jagandeep S. Saraya,[4] James D. Robinson,[4] Maria Julia Maristany,[3,5] Jan Huertas,[3,6] Roger Rubio-Sánchez,[7] Derek K. O'Flaherty,[4] Rosana Collepardo-Guevara,[3,5,6*] and Claudia Bonfio[1,2,*]

[1]Institut de Science et d'Ingénierie Supramoléculaires, CNRS UMR 7006, University of Strasbourg, 67000 Strasbourg, France

[2]Department of Biochemistry, University of Cambridge, CB2 1GA Cambridge, UK

[3]Yusuf Hamied Department of Chemistry, University of Cambridge, CB2 1EW Cambridge, UK

[4]Department of Chemistry, University of Guelph, ON N1G 2W1, Guelph, Canada

[5]Cavendish Laboratory, Department of Physics, University of Cambridge, CB3 0HE, Cambridge, UK

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

[7]Department of Chemical Engineering and Biotechnology, University of Cambridge, CB3 0AS Cambridge, UK

[†]these authors contributed equally

[*]corresponding authors: rc597@cam.ac.uk, cb2036@cam.ac.uk

Abstract

The early co-evolution of RNA and peptides is at the core of the RNA-peptide world hypothesis. Recent studies suggest that nucleotides and amino acids could have formed and polymerised non-enzymatically under prebiotic conditions, generating short oligonucleotides and peptides capable of non-enzymatic RNA replication and peptide synthesis. However, whether the cooperation between nucleic acids and peptides stems from their co-localisation in primitive compartments is unclear. Here we demonstrate the early and likely inevitable emergence of primitive coacervates *via* liquid-liquid phase separation of prebiotic heterogeneous mixtures of short non-coded oligonucleotides and peptides. We show that peptide/nucleic acid coacervates are more prone to form than peptide/peptide coacervates, and that peptide/RNA coacervates are remarkably more stable than peptide/DNA coacervates. Atomistic simulations confirm that the more extended and less structured conformation of RNA over DNA enables more contact points with peptides. The more abundant interactions in peptide/RNA coacervates lead to enhanced salt and thermal stability, yet reduced fluidity compared to their DNA counterparts - which are in turn capable to fully preserve RNA secondary structure upon partitioning. Our findings suggest that peptide/oligonucleotide co-localisation *via* coacervation would have inevitably occurred at an early stage of a more holistic nucleic acid-peptide world scenario. Both RNA and DNA would have thus been required to ensure the emergence of coacervates with balanced stability and fluidity to host non-enzymatic RNA chemistry.

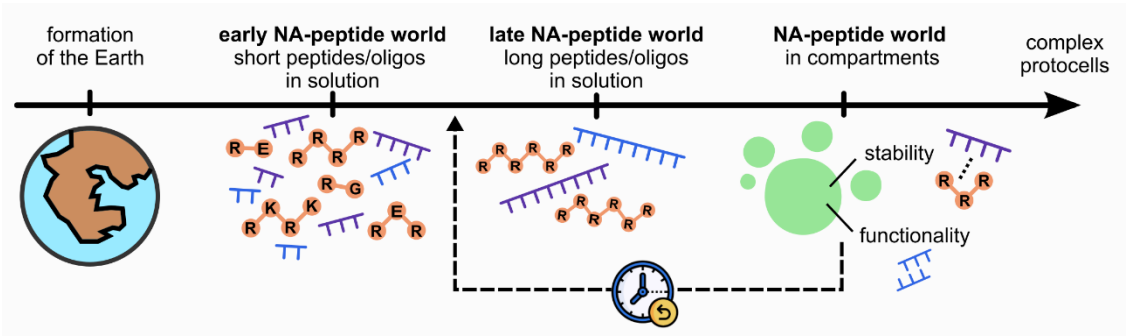
Introduction

The RNA-peptide world hypothesis postulates an early co-evolution of RNAs and peptides, from which RNA replication and peptide synthesis may have emerged.[1,2] It was recently shown that RNA nucleotides and amino acids form non-enzymatically, alongside DNA nucleotides, in prebiotic conditions[3,4] and polymerise into short RNA and DNA oligomers, and peptides.[5–7] While no defined prebiotic role was proposed for DNA so far, canonical and non-canonical RNAs were reported to template DNA polymerisation[8] and direct peptide synthesis,[1,2,9] and short peptides derived from the ribosomal core were

shown to enhance ribozyme activity.[10] Yet, how the primordial synergy between nucleic acids and peptides - ubiquitously preserved in the central dogma of modern biology - originated remains unknown. An intriguing hypothesis relies on the ability of the building blocks of life to cooperate upon co-localisation *via* compartmentalisation at an early stage of the evolutionary timeline.

Given their ability to spatially regulate cellular biochemistry,[11] biomolecular condensates, generated through the liquid-liquid phase separation of RNA and proteins, have been proposed as vestiges of primitive cells.[12] Employed as *in vitro* models of biomolecular condensates, complex coacervates comprising homopeptides and functional oligonucleotides, e.g., ribozymes, were shown to uptake dilute solutes, enable diffusion within and between the compartment and its environment, and host prebiotic reactions, e.g., ribozymatic activity.[12–18] Complex coacervation results from electrostatic interactions between oppositely-charged polymers, including positively-charged polyarginines or polylysines, and negatively-charged inorganic polyphosphates, polyglutamates, polyaspartates, sequence-specific DNA or RNA oligonucleotides.[12] However, in any prebiotic scenario, non-coded oligomerisation pathways would have likely afforded complex mixtures of peptides, DNA and RNA oligomers of limited length and high compositional heterogeneity.[5,19,20] Whether primitive coacervates would have spontaneously emerged from such a diverse assortment of prebiotic molecules, thus enabling their co-localisation early on in a prebiotic setting, or instead relied upon the synthesis of long, coded, functional polymers (*i.e.*, homopeptides and ribozymes) has been vastly overlooked.

To gain insight into the timeline for the emergence of coacervates on early Earth and their plausibility as primitive cells, here we delineate the minimal molecular requirements for prebiotic coacervation. Through experimental and computational studies, our findings indicate that coacervation would have inevitably occurred early on in the evolutionary timeline, likely simultaneously with the emergence of a holistic nucleic acid-peptide world, and modulated prebiotic RNA chemistry. Importantly, we find that RNA-based coacervates are more prone to form but less fluid, thus less suitable to host RNA activity, than DNA-based coacervates, opening up the possibility for different roles of DNA and RNA on early Earth. Our work allows for a reconsideration of the timeline and significance of primitive coacervates and supports the integration of compartmentalisation with replication and translation in a more holistic "compartmentalised nucleic acid-peptide world", marking a critical step on the trajectory towards life as we know it.



Scheme 1. A schematic representation of the evolutionary timeline for the nucleic acid-peptide (NA-peptide) world hypothesis. At an early stage, short peptides and oligonucleotides would have existed in solution; at a later stage, polymerisation pathways (e.g. ribozymatic RNA replication) would have led to polymers with the degree of complexity required for self-assembly. We question this assumption by suggesting that compartments, in particular coacervates, should be considered at an earlier stage of the evolutionary timeline of the NA-peptide world.

Results and discussion

Short peptides enable coacervation.

Arginine (Arg, R) residues are known to interact with negatively-charged monomers such as nucleotides *via* electrostatic and cation-π interactions.[21–23] Such contacts are among the most frequent in RNA/protein complexes.[24] Recently, Arg decamers were shown to undergo liquid-liquid phase separation (LLPS) in the presence of negatively charged molecules of low multivalency, e.g., nucleotide phosphates and decamers of glutamic acid (Glu, E) or aspartic acid (Asp, D).[16] Despite the suitability of long homopolymeric peptides to generate biomimetic coacervates, their abundance on early Earth would imply highly selective incorporation of certain amino acids during polymerisation or an environment enriched in a single amino acid. However, there is no indication for either scenario - libraries of early amino acids plausibly include arginine among ten canonical amino acids,[3] and non-coded prebiotic peptide syntheses are mostly non-selective.[5] We thus sought to improve the prebiotic plausibility of coacervates by assessing the role that peptide length, sequence and charge play to direct phase separation.

We systematically screened the phase space for mixtures of *coacervating* (*i.e.*, enabling coacervation)[25] peptides and oligonucleotides of varied length (1-10 Arg residues and 3-40 nucleotides). We employed mixed-sequence single-stranded (ss) DNA oligomers as model polyanions, to prevent any potential bias derived from a given nucleobase, and *N*- and *C-termini* unprotected peptides as polycationic counterparts (Supplementary Tables 1-2 and Supplementary Figures 1-9). Most peptide/nucleic acid combinations led to coacervation, as assessed by light microscopy (Figure 1a and Supplementary Figure 10). Precipitation of solid aggregates was detected when both types of polymers reached a certain length (8 Arg residues and 40 nucleotides). On the contrary, Arg (R_1) or Arg dimers (R_2) result in soluble mixtures regardless of the length of the DNA strand. Notably, we found that peptide length has greater influence than DNA length on the phase behaviour of peptide/nucleic acid mixtures. In particular, while a mixture of R_3 and DNA_8 is soluble, four extra nucleobases (DNA_{12}) enable coacervation with R_3; yet, phase separation with DNA_8 only requires increasing the peptide length by one unit (R_4).

To investigate how the molecular features of *coacervating* peptides influence the stability of coacervates derived thereof, we varied peptide length, sequence and charge in the presence of mixed-sequence DNA_{20} (5 mM nucleotide concentration). Turbidity measurements upon titration of NaCl allowed us to determine the critical salt concentration (CSC) of peptide/DNA mixtures (Figure 1b, Supplementary Table 3-4 and Supplementary Figure 11). Defined as the highest NaCl concentration tolerated before the complete dissolution of the droplets, CSC is conventionally taken as an indication of coacervate robustness.

Increasing the length of an Arg homopeptide (20 mM amino acid concentration) consistently led to higher CSC values of the corresponding coacervates, *i.e.*, from 92 mM for R_3 to 740 mM for R_8, indicating a correlation between peptide length and the strength of their interactions with DNA (Figure 1b). Partially or fully replacing Arg residues with lysines (Lys, K) resulted, respectively, in a 50% reduction in the CSC or dissolution of the coacervates, likely due to the lower frequency of cation-π interactions of nucleobases with Lys over Arg (Figure 1b).[26,27] Peptides composed of alternating D-Arg and L-Arg also led to coacervates with lower CSCs than those of their homochiral counterparts (Supplementary Table 3).

Inspired by a first attempt to compare homo- and heteropeptides in coacervates,[14] we next probed the influence of peptide sequence in governing coacervation by testing a library of peptides containing four

Arg residues and an increasing number of glycines (Gly, G). All peptides employed in this screening possess fixed, positive charge (+4), but varying charge density. When $R_2G_2R_2$ was mixed with DNA_{20} in place of R_4, we observed a 18% increase in salt stability of the resulting coacervates, likely due to the increased peptide length compensating for the 33.3% decrease in charge density of the peptide (Figure 1b and Supplementary Table 4). However, further reducing the charge density upon incorporating more Gly residues (e.g., $R_2G_4R_2$ or $R_2G_8R_2$) led to a decrease in salt stability. Our findings suggest that in a prebiotic setting where heteropeptides would have likely been more abundant than Arg homopeptides, the stability of primitive coacervates would have greatly depended on both sequence and length,[26] but less on the stereochemistry, of their peptide components.

To provide a more quantitative assessment of the effect of peptide charge on phase separation, we determined the CSC for coacervates made of a series of peptides (R_4, R_6 and $R_2G_2R_2$) with oligonucleotides of various lengths (DNA_N). In contrast with approximations predicted for long polymers,[28] we observed a linear relationship between the CSC and the inverse of polymer length (1/N) for primitive coacervates. We estimated the minimal oligonucleotide length required for coacervation with any given peptide to delineate the precise co-existence boundaries in the related phase space (Figure 1c and Supplementary Table 5). For example, we calculated that R_6 undergoes coacervation only if a DNA tetramer is present ($N \geq 3.7$). Conversely, a shorter or less charged peptide (e.g., R_4 or $R_2G_2R_2$) would require longer DNA oligomers ($N \geq 6$) to undergo phase separation, suggesting that peptides with different sequences and lengths, but the same charge (e.g., R_4 and $R_2G_2R_2$), have similar molecular requirements for coacervation. In support of this hypothesis, we assessed coacervation of $R_2G_nR_2$ peptides (n = 0, 2, 8). In all cases, stable coacervates were observed with DNA_7, but not with DNA_6 (Supplementary Figure 12), confirming that net charge has a more prominent role than length in modulating coacervation.[16]

In parallel, we studied the propensity of peptides with the same length, but different charge (e.g., R_3, RGR and RER), to phase separate in the presence of oligonucleotides (Figure 1d and Supplementary Figure 13). For these peptides, we measured the turbidity of peptide/DNA_{16} mixtures upon peptide titration (at fixed 10 mM nucleotide concentration). The minimal concentration of R_3 required to enable coacervation was found to be 4.8 mM; replacing an Arg residue with Gly or Glu induces a 9-fold and 3.5-fold concentration increase, respectively (Figure 1d). Elongating the peptide, while maintaining constant its charge, also increases the amount of peptide required for coacervation (Supplementary Figure 13). All in all, these findings indicate that substituting arginine with glycine is more detrimental to coacervation than the substitution with glutamic acid, in agreement with previous studies on inorganic polyphosphates.[29] Although replacing one arginine with glutamic acid lowers the peptide charge, it also potentially allows for more hydrogen bonding interactions with itself and DNA_{16}, which would explain the lower coacervation onset for RER than RGR trimers.

Without a genetic code to determine peptide sequences, a widely varied mixture of peptides would have formed on early Earth, subject to the reactivity of each amino acid towards polymerisation. We thus investigated whether compositionally diverse peptide pools would undergo coacervation in the presence of oligonucleotides (Figure 1e). We studied three types of mixtures in which peptides differ in their length (R_3, R_4, R_5), sequence (R_4, RKRK, K_4) or charge density (R_4, R_2GR_2, $R_2G_2R_2$). All mixtures were prepared at an equimolar concentration of their components (6.67 mM amino acid concentration) and added to a solution of DNA_{20} (5 mM nucleotide concentration). Remarkably, we found that all mixtures systematically had higher CSCs than the average value for their individual components, even when peptides uncapable of

coacervation on their own are present in the mixtures (e.g., K_4) (Supplementary Table 4). These results demonstrate that coacervates would have readily formed with a wide range of mixed-sequence peptides resulting from non-enzymatic polymerisation processes[5,19,30] and benefitted from recruiting shorter, *non-coacervating* peptides.

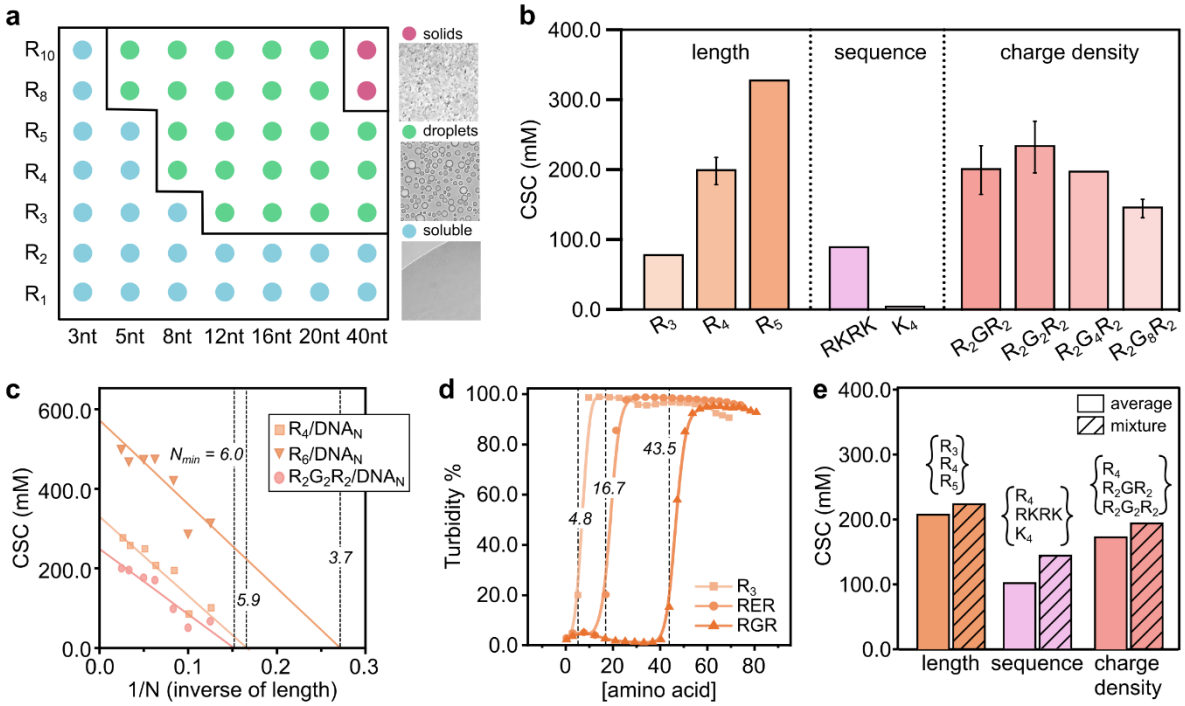


Figure 1. The compositional diversity of peptides influences the phase behaviour of primitive peptide/DNA coacervates. (a) Phase diagram of peptide/DNA coacervates, outlining the minimal peptide and DNA oligomer lengths required for coacervation. DNA sequences comprise the motif $(ACTG)_n$, except for 3nt, which is dA_3. (b) Critical salt concentrations (CSCs) of coacervates as a function of peptide length, sequence and charge density. Experiments were performed in the presence of DNA_{20} ([positively-charged amino acid] = 20 mM). (c) Estimated minimal length for DNA oligomers to undergo coacervation with peptides of the same charge (R_4 and $R_2G_2R_2$) and length (R_6 and $R_2G_2R_2$). N_{min} is obtained for CSC = 0 from the linear regression discussed in the Supplementary Information. (d) Titration curves of a DNA_{16} solution ([nt] = 10 mM) with R_3, RER or RGR peptides. The dotted lines indicate the onset of coacervation (*i.e.*, the concentration above which coacervation is observed). (e) Critical salt concentrations (CSCs) of equimolar mixtures of peptides compared to the average of the individual CSCs of the mixture components. The three mixtures tested in the presence of DNA_{20} ([nt] = 5 mM) are: R_3, R_4 and R_5 (length mixture); $R_2G_2R_2$, R_2GR_2 and R_4 (charge density mixture); and K_4, RKRK and R_4 (sequence mixture). Each peptide in the mixture was present in equimolar amino acid concentration ([amino acid] = 6.67 mM). Abbreviations: R = Arginine, G = Glycine, E = Glutamic acid, K = Lysine, nt = nucleotide.

RNA coacervates are highly stable.

Ribozymes or long nucleic acids (NAs) have been reported to undergo phase separation with positively-charged ions, molecules and peptides.[13–15,31–33] However, prebiotic polymerisation processes would have mainly produced short, non-functional oligonucleotides, for which the *coacervating* propensity has yet to be determined. In view of the prebiotic plausibility of both ribonucleotides and deoxyribonucleotides[34,35] and the lack of systematic comparative studies between different nucleic acids, we investigated the propensity of both single-stranded (ss) DNA and RNA oligomers to undergo coacervation. Although

evidence suggests that DNA would have been present in an RNA world,[36–39] its role remains unclear until the advent of a genetic takeover.

We first assessed the salt stability of coacervates made of Arg tetramers (R_4) with DNA_8 ($(ACTG)_2$) or RNA_8 ($(ACUG)_2$) and compared it with that of coacervates comprising Glu decamers (E_{10}) as polyanions (Supplementary Table 2). CSC values, measured at increasing concentrations of arginine (1-40 mM) and fixed concentrations of glutamic acid (10 mM) or nucleotide (5 mM) (Figure 2a, Supplementary Table 6 and Supplementary Figure 14) or, vice versa, by varying the anion concentration (1-20 mM) at a fixed concentration of arginine (20 mM) (Supplementary Table 6 and Supplementary Figure 15-16), were plotted to delineate the phase diagram of peptide/peptide and peptide/nucleic acid mixtures. In line with previous observations,[16] we observed that R_4/E_{10} does not form coacervates. When a longer positively charged peptide was used (R_{10}) with E_{10}, the maximum salt stability of the resulting coacervates was obtained when the two peptides were present in equimolar charge concentrations. A re-entrant tendency in the phase behaviour was observed in excess of R_{10}, suggesting that peptide/peptide mixtures can lead to coacervation only when the mismatch between the charge concentrations of the two peptides is minimal (Figure 2a). When E_{10} was instead replaced by DNA_8 and RNA_8, we observed coacervation in a broader range of conditions. Upon increasing the concentration of arginine, salt stability curves reached a plateau at 8:1 [Arg]:[nucleotide] ratio, exhibiting no re-entrant behaviour even at high polymer charge mismatch and generating wider phase co-existence regions than those of peptide/peptide coacervates (Figure 2a). The propensity of DNA oligonucleotides to undergo phase separation with peptides in mismatched charge concentrations suggests that the peptide/DNA coacervates were more likely to occur in a prebiotic setting than their peptide/peptide analogues. DNA-based coacervates were also found stable in more prebiotic buffers, including imidazole and phosphate, in a relatively wide pH range (6.2-7.8) and at high magnesium concentrations (up to 250 mM) (Supplementary Figures 17-19).

Surprisingly, we observed that the salt tolerance of R_4/RNA_8 is 2.3 times higher compared to the R_4/DNA_8 mixture, rising from 110.0 mM to 249.0 mM NaCl at 8:1 [Arg]:[nucleotide] ratio. A greater tendency of RNA oligomers to form coacervates over their DNA counterparts was also confirmed by measuring the minimal concentration of oligonucleotide and peptide required for coacervation, which is 2-fold lower for R_4/RNA_8 mixtures compared to R_4/DNA_8 mixtures (Figure 2b). Intrigued by the enhanced salt stability of peptide/RNA coacervates, we turned to hot-stage epifluorescence microscopy[40] to evaluate their temperature susceptibility compared to analogous peptide/DNA coacervates (Figure 2c and Supplementary Figure 20). Along the heating ramp, full dissolution of the R_4/DNA_8 coacervates was observed at ~45°C. Conversely, R_4/RNA_8 coacervates showcase greater thermal stability, dissolving only at ~60°C. In both cases, cooling led to coacervation, confirming the reversibility of the assembly process. While an additional hydroxyl group was shown to mildly increase the CSC of coacervates comprising small metabolites,[41] the unprecedented difference in the salt and thermal stability of DNA and RNA coacervates is suggestive of stronger interactions between RNA and peptides than between DNA and peptides.

The differences in thermal and salt stability observed for R_4/DNA_8 and R_4/RNA_8 coacervates suggest distinct minimal coacervation requirements with respect to the length of peptides when RNA oligomers are used instead of DNA analogues. We found that RNA_8, but not DNA_8, forms coacervates with Arg trimers (R_3) (CSC = 54.2 mM) (Supplementary Table 3). Similarly, droplets were observed when RNA_{20} was mixed with R_2 (Supplementary Table 7 and Supplementary Figure 21). We next computed the minimal peptide length required for coacervation for a series of oligonucleotides (DNA_8, RNA_8, DNA_{12} and RNA_{12}) (Figure 2d and

Supplementary Figure 22). In close agreement with the qualitative screening (Figure 1a), we confirmed that at least an Arg tetramer is required to form coacervates with DNA_8 ($N \geq 3.5$), whereas coacervation occurs with RNA_8 and the shorter R_3 ($N \geq 2.7$) (Supplementary Table 5). Because chimeric RNA-DNA oligonucleotides would have likely emerged from a prebiotic pool of ribonucleotides and deoxyribonucleotides,[34,35,42] we also tested an oligonucleotide comprising 50% RNA and 50% DNA nucleotides (HNA_8) and observed an N_{min} value similar to that obtained for RNA_8 ($N \geq 2.8$) (Figure 2d). An analogous trend was observed for longer oligonucleotides, with RNA_{12}, hybrid strands (HNA_{12}) and mixed DNA-RNA oligomers predicted to form coacervates with Arg trimers (Supplementary Figure 23). These results suggest that the effect of ribonucleotides or RNA oligomers in a heterogeneous mixture with Arg peptides would have likely overcome that of deoxyribonucleotides and DNA oligomers and led to the emergence of coacervates with minimal length requirements and salt stability similar to those of a pure peptide/RNA system.

Homopolymeric DNA and RNA sequences have been widely studied for their ability to form coacervate models.[28,43] Yet, as purines are only slightly more reactive than pyrimidines in template-free non-enzymatic RNA polymerisation,[44] heteropolymeric oligonucleotide sequences would have likely been more abundant than homopolymeric counterparts on early Earth. We thus investigated how oligonucleotide sequence and charge influences coacervation. When homopolymeric DNA decamers were employed, we observed solid-like aggregates in the presence of polycytosine and polyguanine; polyadenine and polythymine instead formed coacervates with substantially lower CSCs than those measured using heteropolymeric DNA sequences (Supplementary Figure 24). Similarly, the minimal oligonucleotide length required for coacervation with R_6 is almost 2-fold higher for polyadenine ($polyA_N$) compared to mixed-sequence oligonucleotides (N_{min} = 6.6 *vs* 3.7, respectively) (Supplementary Figure 25). These results indicate that short, mixed-sequence oligonucleotides exhibit a higher propensity towards coacervation than less prebiotic, homopolymeric strands. Conversely, we found that increasing the oligonucleotide charge by means of phosphate groups on the 5′ and 3′ ends causes the formation of clusters of coacervates and, in time, solid-like aggregates (Supplementary Figure 26), potentially due to the additional electrostatic interactions made available by the more exposed phosphate groups.

Following the observations that coacervation with low charge density heteropeptides is possible and that an R_2/RNA_{20} mixture forms coacervates, we investigated whether *N*- and *C-termini* unprotected peptide heterodimers (R_2, RG and RE) could undergo phase separation with short RNAs (Supplementary Table 7 and Supplementary Figure 21). As expected, none of the peptide dimers phase separated with DNA_8 and DNA_{12}. Surprisingly, we observed coacervation for RG and RE, but not R_2, when mixed with RNA_8. While RG (charge = +1) requires an amino acid concentration of 40 mM to form coacervates, RE (charge = 0) undergoes phase separation in the same conditions employed for R_3 and R_4 (20 mM amino acid concentration), likely due to the ability of RE to participate in a wider range of interactions, e.g., hydrogen bonding between glutamic acid and other peptides or oligonucleotides. Our results indicate that, even in a prebiotic scenario where short, heterogeneous peptides, RNA and DNA oligomers were present, phase separation would have inevitably occurred and possibly impacted the chemistry taking place at the dawn of a nucleic acid-peptide world.

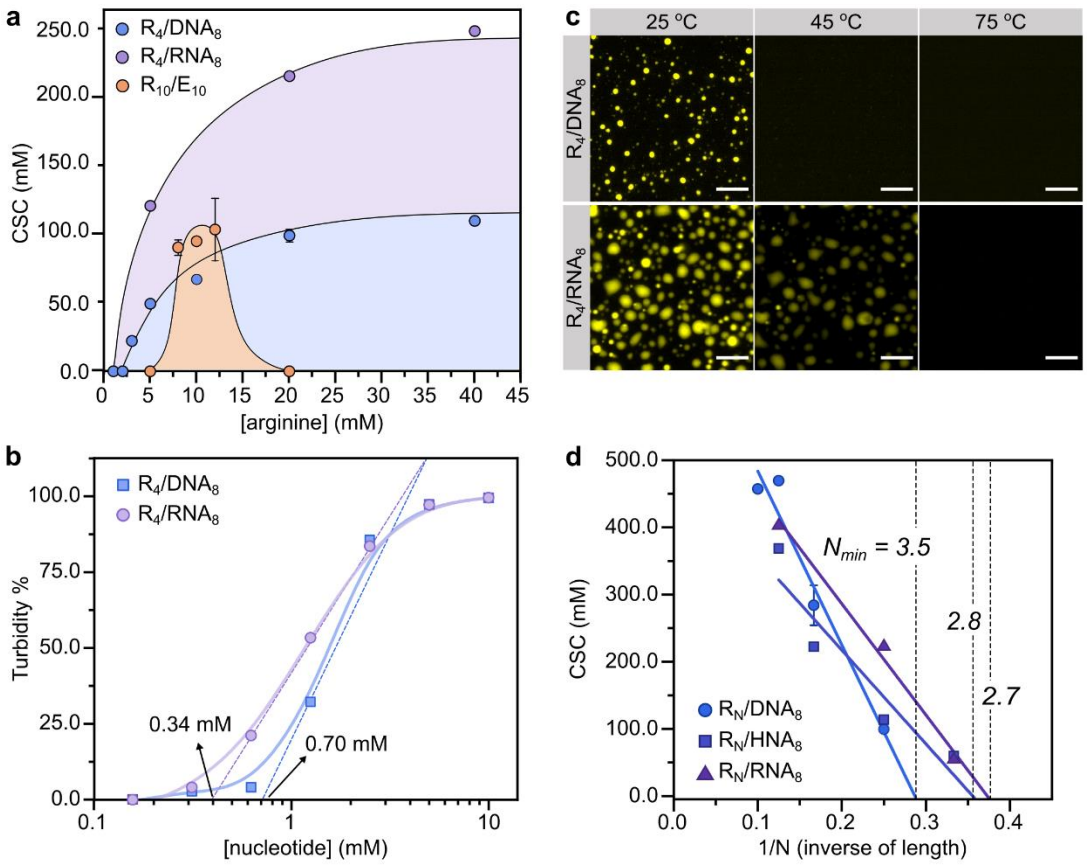


Figure 2. Peptide/RNA coacervates exhibit higher robustness than peptide/DNA coacervates. (a) Salt stability of peptide/peptide and oligonucleotide/peptide coacervates. Critical salt concentrations (CSCs) were measured through turbidity measurements of peptide/peptide and oligonucleotide/peptide solutions upon titration with NaCl in 25 mM HEPES, pH 7.5 and at room temperature. In all experiments, the anion concentration was kept constant ([nt] = 5 mM and [glutamic acid] = 10 mM). (b) Turbidity curves for R_4/DNA_8 and R_4/RNA_8 as a function of nucleotide concentration. The dotted lines are tangents to the inflection point, used to determine the minimal concentration required for coacervation (indicated in the graph). (c) Thermal stability of R_4 coacervates with DNA_8 and RNA_8. 1% Cy_3-$(TGAC)_2$ was used for visualization. Scale bars are 10 µm. (d) Estimation of the minimal peptide length (N_{min}) required for coacervation for a given nucleic acid composition. The employed oligonucleotides are DNA_8 = $(ACTG)_2$, HNA_8 = ArCrUGArCrUG and RNA_8 = $(ACUG)_2$. Abbreviations: R = Arginine, E = Glutamic acid, nt = nucleotide.

Peptides interact more with RNA than DNA.

To elucidate the distinct features of the interactions between peptides and DNA or RNA molecules at a molecular level, thus rationalising the different salt and thermal stabilities of the coacervates resulting thereof, we carried out all-atom simulations of mixtures representatives of our systems. Our models contain eight single-stranded (ss) DNA_8 or RNA_8 molecules with thirty-six Arg peptides (either R_3 or R_4) in explicit solvent and ions. While arginine is known to interact with RNA through multiple modes,[45] any distinction between RNA and DNA oligomers in undergoing coacervation has remained unexplored, due to the focus on probing the differences between double-stranded and single-stranded DNA (for their significance in genomic function),[46,47] and the assumption that RNA preceded DNA on early Earth.[48,49]

We performed simulations of four mixtures: R_3/DNA_8, R_4/DNA_8, R_3/RNA_8 and R_4/RNA_8. For each mixture, we analysed the trajectories to quantify the frequency of intermolecular contacts between Arg

peptides and oligonucleotides. Three main intermolecular interaction modes were identified: electrostatic (here defined as non-hydrogen bonding contacts between the positively-charged sidechain of Arg and the backbone phosphate group in both oligonucleotides), hydrogen bonding, and π-π/cation-π stacking interactions between the positively-charged Arg sidechain and the nucleobases in both oligonucleotides (Figure 3a). Across all interaction classes, RNA_8 consistently forms more contacts with Arg peptides than DNA_8, with differences being most pronounced in stacking interactions (96% increase in contact points for RNA_8 over DNA_8) and hydrogen bonding *via* the nucleobase (Figure 3b, Supplementary Table 8 and Supplementary Figures 27-28). Despite their relatively low frequency, the enhanced strength of cation-π stacking interactions between Arg residues and nucleobases demonstrated through quantum mechanical calculations of model systems, relative to electrostatic or hydrogen bonding in aqueous media,[50] suggests that even minor variations in their occurrence can have a significant energetic impact. As such, the higher number of stacking interactions present in the RNA_8 systems are expected to be a key contributor to the higher thermodynamic stability of RNA_8-based coacervates, as also shown by the studies on the thermal stability of coacervates (Figure 2c).

The higher frequency of intermolecular interactions observed for RNA_8 over DNA_8 can be attributed to conformational differences between the two nucleic acids, likely due to the additional hydroxyl moiety in ribonucleotides. Specifically, in comparison to DNA_8, RNA_8 adopts an unfolded structure when interacting with Arg peptides (Figure 3c), which enables the nucleobases in RNA_8 to engage more readily in intermolecular hydrogen bonding and stacking interactions with Arg residues. Expansion of polymers within the coacervate phase has been shown to increase the density of intermolecular connections, the enthalpic gain for coacervation and the thermodynamic stability of coacervates.[51] Our simulations reveal that the higher propensity of RNA_8 versus DNA_8 to expand upon interacting with Arg peptides is contributed by the significantly higher number of intermolecular stacking interactions it establishes. This observation aligns with previous structural analyses revealing that Arg exhibits stronger and more frequent π-π contacts with RNA nucleobases than DNA nucleobases.[52] Several other factors likely contribute to the higher propensity of RNA_8 versus DNA_8 to expand upon interaction with Arg peptides, including the weaker stacking interactions of uracil with other nucleobases compared to thymine[53] and the bias towards compact helical conformations observed in DNA force fields - initially developed to ensure adequate predictions of double-stranded DNA structures.[54]

The total valency per oligonucleotide, defined as the total number of intermolecular contacts that RNA_8 or DNA_8 form with Arg peptides (Figure 3d, Supplementary Table 9 and Supplementary Figure 29), correlates well with experimentally observed phase separation propensity (Figure 2). Indeed, we see that the R_3/DNA_8 system, the only simulated mixture for which we do not observe coacervation, exhibits the lowest number of inter-molecular contacts. Interestingly, elongating the peptide chain by one Arg residue (from R_3 to R_4) results in an 18% increase in hydrogen bonding and a 16% increase in electrostatic interactions (Supplementary Figure 27). This increase in interactions for the R_4/DNA_8 system aligns with our experimental finding that R_4 is the minimum peptide length required for coacervation with DNA_8 (Figure 2d). In contrast, R_3/RNA_8 makes a similar number of total intermolecular contacts to R_4/DNA_8 due to the increased hydrogen bonding and stacking interactions with the shorter peptide. This observation, combined with the fact that in all cases excess peptide remains in the simulation box at equilibrium (Figure 3e, Supplementary Table 10 and Supplementary Figure 30), suggests that the intrinsic physicochemical

differences between DNA and RNA can explain the different sensitivity to peptide length we observe in experiment.

Our simulations reveal striking differences in how Arg peptides interact with DNA and RNA, offering molecular insights that can rationalize the macroscopic differences in the phase separation behaviour observed experimentally. The RNA_8 systems exhibit a notably higher frequency of stacking and hydrogen bonding (Figure 3b), which likely underpins the enhanced thermodynamic stability of RNA_8-based coacervates. This stability manifests in their increased resilience to both salt concentration (Figure 2a) and temperature (Figure 2b), aligning our molecular-level findings with macroscopic observations.

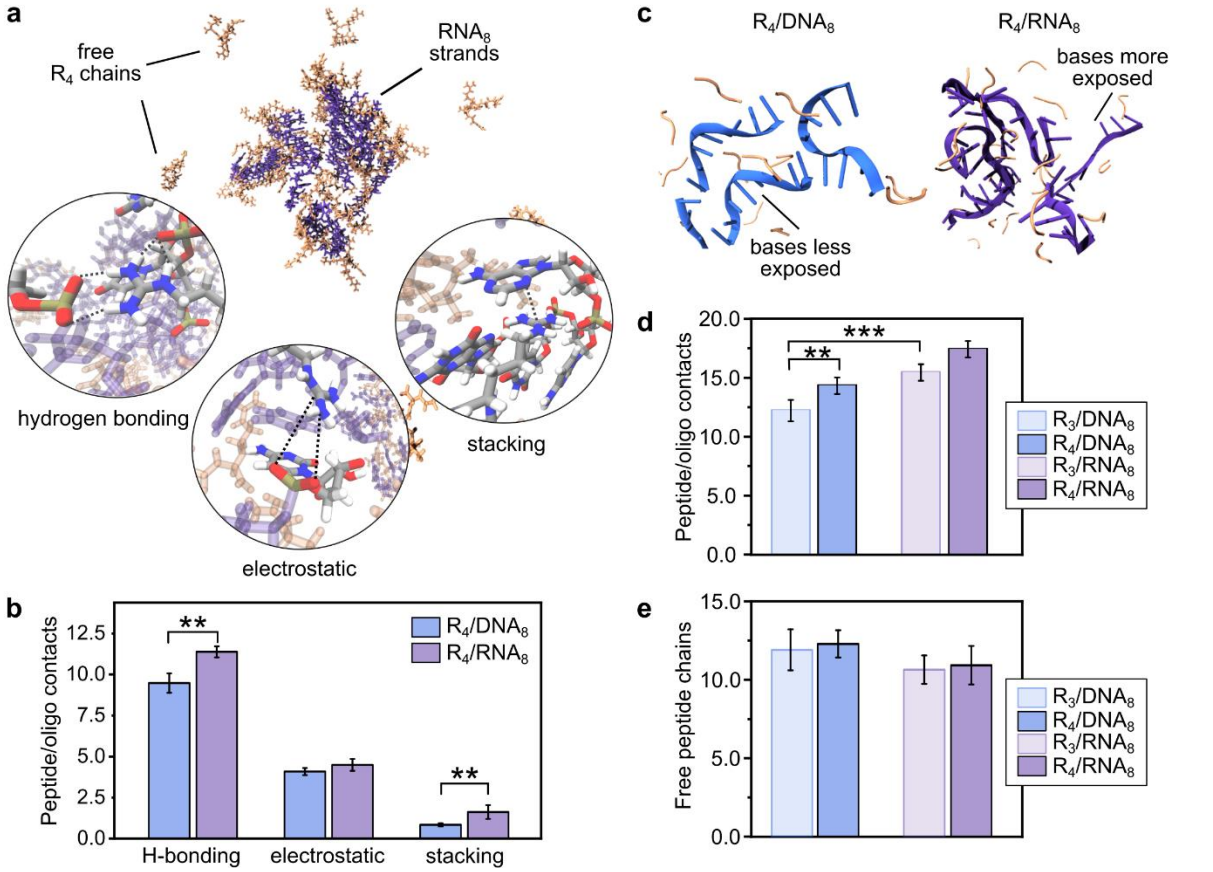


Figure 3. *In silico* investigations reveal contact modes and frequency of interactions in peptide/nucleic acid coacervates. (a) Representative simulation snapshot of the R_4/RNA_8 mixture, showing a cluster of RNA and peptide, and unbound peptide in excess. Inset magnifications highlight the investigated interaction modes, e.g., hydrogen bonding, electrostatic interactions, and π-π/cation-π stacking. (b) Comparison between the number of DNA and RNA interactions with arginine peptides (per frame, per nucleotide), separated into three categories: hydrogen bonding, electrostatic interactions and stacking. (c) Simplified rendering of R_4/DNA_8 (blue) and R_4/RNA_8 (purple) clusters, showing the helical, structured conformation acquired by DNA strands and the more disordered folding acquired by RNA strands, which leaves ribonucleotides exposed to interact with peptides. (d) Number of peptide/oligonucleotide contacts (all modes of interaction), which represents the total number of intermolecular contacts that one molecule of RNA_8 or DNA_8 forms with R_3 or R_4. (e) Excess peptide remaining in the simulation box at equilibrium for RNA_8 or DNA_8 coacervates with R_3 or R_4. Abbreviations: R = Arginine.

DNA coacervates preserve aptamer folding.

Model coacervates have been studied for their ability to increase the local concentration of dilute solutes, including oligonucleotides and potentially facilitate, among other processes, replication reactions.[15,55] However, whether the composition of the minimal coacervates described herein would influence their ability to partition solutes remains unexplored. In particular, we investigated whether the different propensity of DNA and RNA to undergo coacervation is reflected in a different tendency of the resulting coacervates to recruit peptides and oligonucleotides. The partition coefficients for a number of fluorescently-labelled probes (FITC-R_8, FAM-DNA_8 and FAM-RNA_8) were measured by confocal microscopy (**Error! Reference source not found.**Supplementary Table 11). In line with our atomistic results (Figure 3), FITC-R_8 exhibited a 1.5 times higher partition coefficient in R_8/RNA_8 than in R_8/DNA_8 coacervates (Figure 4a), due to the greater number of contacts occurring between Arg peptides and RNA. While FAM-DNA_8 and FAM-RNA_8 were found to partition similarly in R_4/DNA_8 coacervates, R_4/RNA_8 coacervates showcased a 1.3-fold enhanced tendency to recruit FAM-RNA_8 instead of FAM-DNA_8 (Figure 4b). The observed difference in partition coefficients for RNA-based coacervates likely results from the higher energetic cost of recruiting a conformationally-rigid and less interacting DNA probe into RNA coacervates.[56] Importantly, we observed that also oligonucleotides that are too short to undergo coacervation are efficiently recruited in primitive coacervates (Supplementary Figure 31).

Besides concentration, diffusion rate within coacervates has been shown to be a crucial aspect for nucleic acid reactivity, e.g., ribozymatic activity and non-enzymatic RNA polymerisation. Specifically, coacervates comprising long RNA strands (more than 50 nucleotides)[14] and Arg homopeptides[15] are known to have high viscosity and to inhibit ribozymatic activity. We thus characterised the fluidity of primitive coacervates focusing on the diffusion of recruited oligonucleotides by means of Fluorescence Recovery After Photobleaching (FRAP).

Firstly, we varied the length of the oligonucleotide probe partitioned into R_4/DNA_8 coacervates, employing three non-complementary Cy3-labelled probes made of 8, 16 and 32 nucleotides (Figure 4c). Empirical recovery times (τ) for all probes range between 5 and 95 seconds, indicating highly fluid primitive coacervates (Supplementary Table 12). In contrast, coacervates made of $(RG_2)_4$ and 76-nucleotide-long RNA recovers the fluorescence signal on a time scale of several minutes.[14] As expected, the empirical recovery time is proportional to the length of the coacervate components, which is correlated to the viscosity of the coacervate (Figure 4d and Supplementary Figure 32-33). Interestingly, peptide length has a stronger effect on coacervate fluidity than oligonucleotide length: the recovery time of the 8-nucleotide long probe in R_4/DNA_8 coacervates is nearly half of that in R_4/DNA_{16} coacervates (10 s and 18 s, respectively), whereas employing R_8 instead of R_4 for coacervation induces a 3-fold increase in the recovery time of the probe. These results align with those reported for coacervate stability and its stronger dependence on peptide rather than oligonucleotide length (Figure 1a and Supplementary Table 3).

Secondly, we investigated the potential correlation between coacervate stability and probe diffusion. The recovery times in R_4/DNA_8, R_4/DNA_8:RNA_8 (1:1 ratio) and R_4/RNA_8 coacervates confirm low viscosity for all systems. Yet, we found that the probe is more mobile in DNA-based than in RNA-based coacervates (empirical recovery times of 10 s and 62 s, respectively) (Figure 4e), suggesting that higher viscosity, and hence slower probe diffusion, correlates with the frequency of interactions between coacervate components determined in the all-atom simulations.

Thirdly, we performed FRAP using a probe (Cy3-dA_{11}) capable of base-pairing with the DNA molecule (dT_{10}) involved in coacervation (Supplementary Table 12 and Supplementary Figure 34-36). In comparison

with the non-complementary R_{10}/dA_{10} system (τ = 3 s), fluorescence recovery is slower in R_{10}/dT_{10} coacervates likely due to hybridisation between the probe and the *coacervating* oligonucleotide (10 s). Thus, although base-pairing occurs in primitive coacervates, it does not significantly hinder probe mobility.[57]

The RNA-peptide world posits an evolutionary period where primitive lifeforms relied heavily on the catalytic properties and information carrying capabilities of functionally-folded RNA oligonucleotides. We thus investigated whether primitive coacervates would have supported RNA folding. We assessed the impact that coacervates have on the secondary structure of RNA, by adding a shorter split version of the Broccoli aptamer[40] to a solution containing R_4/RNA_8, R_4/DNA_8 or R_4/DNA_{16} coacervates and measuring the fluorescence intensity of the bound fluorogenic probe, 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI) (Supplementary Figure 37-38). The fluorescence of the aptamer, and thereby its secondary structure, is fully preserved within R_4/DNA_8 coacervates, but only partially maintained in R_4/RNA_8 coacervates or in coacervates comprising sufficiently long DNAs, where interfacial partitioning is observed (Figure 4f-g). These findings indicate that coacervates made of short DNA oligonucleotides and Arg peptides - characterised by weaker and less abundant interactions than their RNA-based counterparts (Figure 3b) -, enable more efficient nucleic acid folding. Overall, while previous studies demonstrated that ribozyme activity and RNA primer extension are sensitive to polycation identity and heavily inhibited in polyarginine/RNA coacervates,[14,15] the full preservation of RNA secondary structure in Arg peptides/DNA coacervates (Fig. 4g) is unique, and highly promising for future studies on prebiotic RNA chemistry. What's more, the consistent differences in biophysical properties between DNA and RNA coacervates, inevitably generated from heterogeneous prebiotic mixtures on early Earth, point towards a more holistic nucleic-acid-peptide world scenario where all the precursors of the central dogma of biology would have played key, yet distinct, chemical roles to sustain the emergence of life.

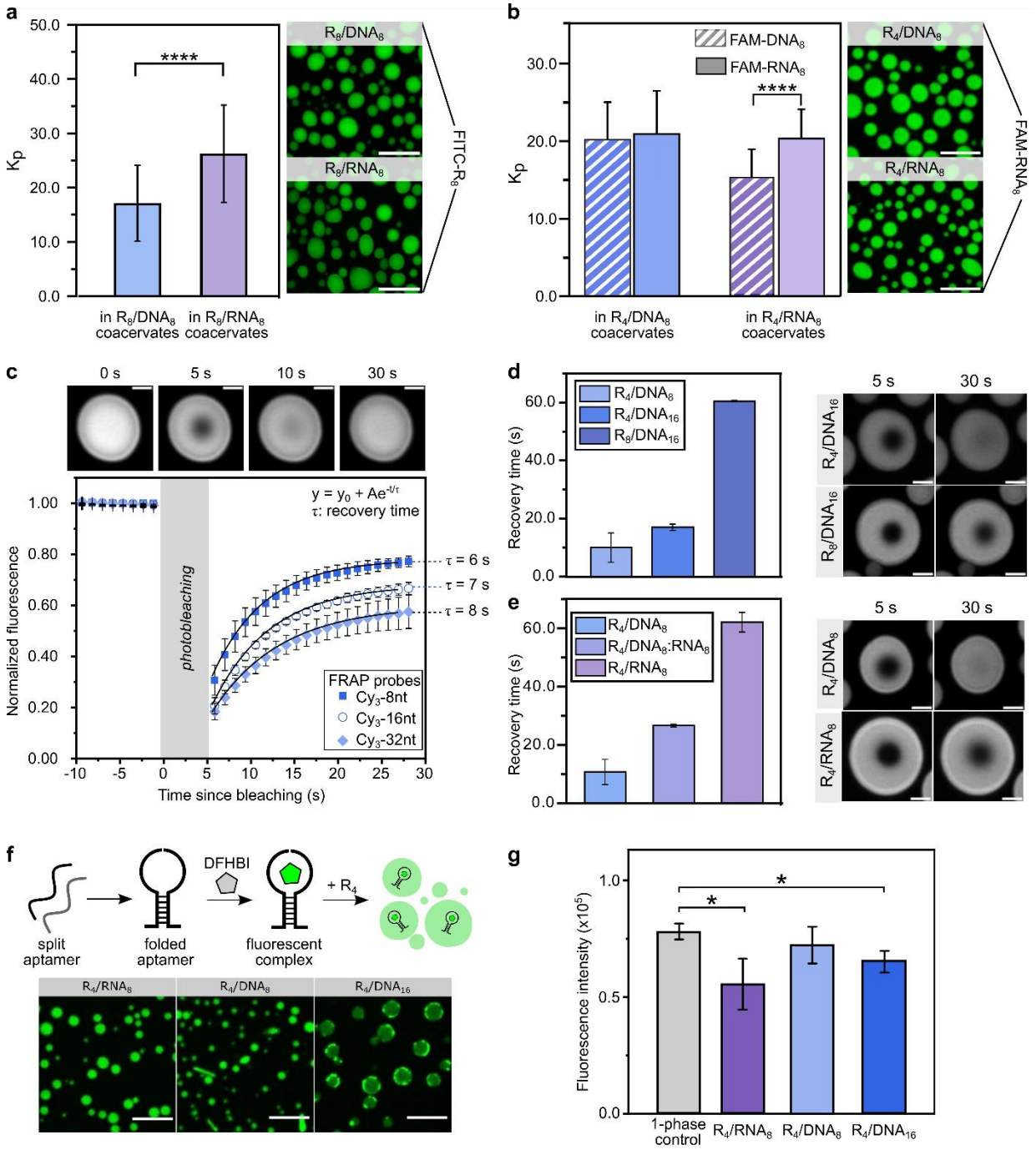


Figure 4. Biophysical properties of primitive peptide/NA coacervates are modulated by their composition. (a) Partitioning of a labelled peptide (1% FITC-R_8) in peptide/nucleic acid coacervates. Scale bar: 10 µm. (b) Partitioning of labelled oligonucleotides (1% FAM-DNA_8 or 1% FAM-RNA_8) in peptide/nucleic acid coacervates. Scale bar: 10 µm. (c) Example of FRAP profiles for the investigated peptide/nucleic acid coacervates. The fit of three probes in R_4/DNA_8 coacervates is included for clarity. Scale bar: 2 µm. (d) Recovery time for the probe Cy3-8nt in coacervates of varying peptide and DNA length: R_4/DNA_8, R_4/DNA_{16} and R_8/DNA_{16}. (e) Recovery time of the probe Cy3-8nt in coacervates comprising R_4 and DNA_8, RNA_8 or a DNA_8:RNA_8 (1:1 ratio) mixture. (f) Schematic representation and confocal micrographs of the split Broccoli aptamer reconstitution in primitive coacervates. (g) Total DFHBI emission in Broccoli aptamer samples after addition of R_4 (to trigger coacervation). In the 1-phase control, no oligonucleotide strand other

than the split aptamer was present, and no peptide was added. Abbreviations: R = Arginine, nt = nucleotide, τ = empirical recovery time.

Conclusions

Complex coacervates, formed upon liquid-liquid phase separation of oppositely-charged polymers, have long been suggested as models of primitive cells. Yet, the low prebiotic plausibility of the coacervating components, commonly designed to maximise coacervate stability or functionality, has led to the assumption that the emergence of coacervates succeeded the synthesis of long, sequence-specific, functional polymers (i.e., homopeptides and ribozymes). Our work challenges this assumption by demonstrating that short, mixed-sequence peptides and oligonucleotides inevitably undergo phase separation, generating coacervates that likely impacted prebiotically-relevant RNA chemistry.

The prebiotic plausibility of short peptides, RNA and DNA oligomers, and their intertwined role in the central dogma of biology, suggest their cooperation, likely due to co-localisation, early on in the evolutionary timeline. Our work shows that primitive coacervates can be generated by liquid-liquid phase separation of short, heterogeneous peptides and oligonucleotides (*i.e.*, peptide dimers and trimers, RNA and DNA octamers), likely indicating that compartmentalisation *via* coacervation would have occurred simultaneously to the early stages of non-coded amino acid and nucleotide polymerisation. In contrast with peptide-peptide coacervates, these minimal nucleic acid-peptide coacervates showcase enhanced stability to high concentration mismatch of their components and elevated salinity, thus loosening the chemical constraints on the compatible prebiotic environments that could have accommodated coacervation. The seemingly inevitable tendency of short heterogeneous peptides and oligonucleotides to undergo phase separation suggest that coacervates were unlikely selected as a fitness advantage at a late evolutionary stage, but rather were an inevitable consequence of prebiotic molecular composition in an early nucleic acid-peptide world.

Primitive coacervates can be effectively described by all-atom simulations of peptide/nucleic acid condensation. Mixtures comprising RNA oligonucleotides are characterised by a higher number of contacts between arginine residues and nucleotides compared to DNA-based counterparts, likely due to the more extended and less structured conformation acquired by RNA over DNA upon phase separation. While both nucleic acid/peptide coacervates exhibit enhanced stability and fluidity over previously reported models, the chemical diversity of RNA and DNA is mirrored in the diverse properties of the resulting coacervates. The observation that DNA-based coacervates more efficiently preserve RNA secondary structures suggests that DNA oligonucleotides might have played a key biochemical role in a nucleic acid-peptide world where both nucleic acids would have been present. All in all, the inevitability of coacervation invites to revisit prebiotic chemistry for its compatibility and efficacy in phase-separated environments; and the unique ability of DNA coacervates to efficiently preserve functional RNA folding offers a novel trajectory for the early evolution of primitive cells with sequence-dependent phenotypes.

Acknowledgements

We thank Jean-Daniel Fauny and Romain Vauchelles (IBMC, Unistra) for the access to the confocal microscopy facility, Valentin Bauer (ISIS, Unistra) for support with SPPS and peptide purification, Tiemei Lu for input on the initial experimental setup. Prof. Pietro Cicuta and Dr. Jurij Kotar for support in obtaining a

dataset on the temperature stability of minimal coacervates. We further thank Dr David A. Russell (University of Cambridge) for insightful comments on the manuscript and relevant literature provided.

The authors acknowledge funding from NWO (Dutch Research Council) *via* a Rubicon Fellowship (KKN, 019.222EN.011), Human Frontier Science Program Organization (HFSPO) *via* an Early Career Research Grant (C.B. and D.K.O. RGY00062/2022), Federation of European Biochemical Society *via* a FEBS Excellence Award (to C.B.), 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 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 Biotechnology and Biological Sciences Research Council *via* a BBSRC Discovery Fellowship (BB/X010228/1 to R.R.S.), the UKRI EPSRC under the UK Government's guarantee scheme (EP/Z002028/1), following funding by the ERC (Consolidator Grant) under the European Union's Horizon Europe research and innovation programme, the Winton Programme for Physics of Sustainability (for doctoral funding to M.J.M.), the NSERC *via* a NSERC Discovery Grant (RGPIN 2020-05043 to D.K.O.) and a NSERC Alliance Catalyst Grant (ALLRP 57555822 to D.K.O.). This project made use of time on HPC granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim (http://hecbiosim.ac.uk), supported by EPSRC (grant no. EP/R029407/1).

References

1. Müller, F. et al. A prebiotically plausible scenario of an RNA–peptide world. Nature 605, 279–284 (2022).
2. Su, M., Roberts, S. J. & Sutherland, J. D. RNA-directed peptide synthesis across a nicked loop. Nucleic Acids Research 11415–11422 (2024) doi:10.1093/nar/gkae702.
3. 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 7, 301–307 (2015).
4. Becker, S. et al. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. Science 366, 76 LP – 82 (2019).
5. Canavelli, P., Islam, S. & Powner, M. W. Peptide ligation by chemoselective aminonitrile coupling in water. Nature 571, 546–549 (2019).
6. Gibard, C., Bhowmik, S., Karki, M., Kim, E.-K. & Krishnamurthy, R. Phosphorylation, oligomerization and self-assembly in water under potential prebiotic conditions. Nature Chemistry 10, 212–217 (2018).
7. Forsythe, J. G. et al. Ester-Mediated Amide Bond Formation Driven by Wet-Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. Angewandte Chemie - International Edition 54, 9871–9875 (2015).
8. O'Flaherty, D. K., Zhou, L. & Szostak, J. W. Nonenzymatic Template-Directed Synthesis of Mixed-Sequence 3′-NP-DNA up to 25 Nucleotides Long Inside Model Protocells. J. Am. Chem. Soc. 141, 10481–10488 (2019).
9. Jash, B., Tremmel, P., Jovanovic, D. & Richert, C. Single nucleotide translation without ribosomes. Nat. Chem. 13, 751–757 (2021).
10. Tagami, S., Attwater, J. & Holliger, P. Simple peptides derived from the ribosomal core potentiate RNA polymerase ribozyme function. Nature Chem 9, 325–332 (2017).
11. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol 18, 285–298 (2017).
12. Abbas, M., Lipiński, W. P., Wang, J. & Spruijt, E. Peptide-based coacervates as biomimetic protocells. Chem. Soc. Rev. 50, 3690–3705 (2021).

13. Drobot, B. et al. Compartmentalised RNA catalysis in membrane-free coacervate protocells. Nat Commun 9, 3643 (2018).
14. Iglesias-Artola, J. M. et al. Charge-density reduction promotes ribozyme activity in RNA–peptide coacervates via RNA fluidization and magnesium partitioning. Nat. Chem. 14, 407–416 (2022).
15. Poudyal, R. R. et al. Template-directed RNA polymerization and enhanced ribozyme catalysis inside membraneless compartments formed by coacervates. Nat Commun 10, 490 (2019).
16. 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 11, 5949 (2020).
17. Nakashima, K. K., Van Haren, M. H. I., André, A. A. M., Robu, I. & Spruijt, E. Active coacervate droplets are protocells that grow and resist Ostwald ripening. Nat Commun 12, 3819 (2021).
18. Fraccia, T. P. & Martin, N. Non-enzymatic oligonucleotide ligation in coacervate protocells sustains compartment-content coupling. Nat Commun 14, 2606 (2023).
19. Fox, S. W. & Nakashima, T. The assembly and properties of protobiological structures: The beginnings of cellular peptide synthesis. Biosystems 12, 155–166 (1980).
20. Verlander, M. S. & Orgel, L. E. Analysis of high molecular weight material from the polymerization of adenosine cyclic 2′, 3′-phosphate. J Mol Evol 3, 115–120 (1974).
21. Krainer, G. et al. Reentrant Liquid Condensate Phase of Proteins is Stabilized by Hydrophobic and Non-Ionic interactions. Biophysical Journal 120, 28a (2021).
22. Bremer, A. et al. Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. Nat. Chem. 14, 196–207 (2022).
23. Joseph, J. A. et al. Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. Nat Comput Sci 1, 732–743 (2021).
24. Jones, S. Protein-RNA interactions: a structural analysis. Nucleic Acids Research 29, 943–954 (2001).
25. Blocher McTigue, W. C. & Perry, S. L. Protein Encapsulation Using Complex Coacervates: What Nature Has to Teach Us. Small 16, 1907671 (2020).
26. Fisher, R. S. & Elbaum-Garfinkle, S. Tunable multiphase dynamics of arginine and lysine liquid condensates. Nat Commun 11, 4628 (2020).
27. Kumar, K. et al. Cation–π interactions in protein–ligand binding: theory and data-mining reveal different roles for lysine and arginine. Chem. Sci. 9, 2655–2665 (2018).
28. Lu, T., Nakashima, K. K. & Spruijt, E. Temperature-Responsive Peptide–Nucleotide Coacervates. J. Phys. Chem. B 125, 3080–3091 (2021).
29. Dai, S. et al. An inorganic mineral-based protocell with prebiotic radiation fitness. Nat Commun 14, 7699 (2023).
30. Verlander, M. S., Lohrmann, R. & Orgel, L. E. Catalysts for the self-polymerization of adenosine cyclic 2′,3′-phosphate. J Mol Evol 2, 303–316 (1973).
31. Le Vay, K. K., Salibi, E., Ghosh, B., Tang, T. D. & Mutschler, H. Ribozyme activity modulates the physical properties of RNA–peptide coacervates. eLife 12, e83543 (2023).
32. Samanta, A., Sabatino, V., Ward, T. R. & Walther, A. Functional and morphological adaptation in DNA protocells via signal processing prompted by artificial metalloenzymes. Nat. Nanotechnol. 15, 914–921 (2020).
33. Codispoti, S. et al. The interplay between peptides and RNA is critical for protoribosome compartmentalization and stability. Nucleic Acids Research gkae823 (2024) doi:10.1093/nar/gkae823.
34. Xu, J. et al. Selective prebiotic formation of RNA pyrimidine and DNA purine nucleosides. Nature 582, 60–66 (2020).
35. Xu, J., Green, N. J., Russell, D. A., Liu, Z. & Sutherland, J. D. Prebiotic Photochemical Coproduction of Purine Ribo- and Deoxyribonucleosides. J. Am. Chem. Soc. 143, 14482–14486 (2021).

36. Kim, S. C., O'Flaherty, D. K., Giurgiu, C., Zhou, L. & Szostak, J. W. The Emergence of RNA from the Heterogeneous Products of Prebiotic Nucleotide Synthesis. J. Am. Chem. Soc. 143, 3267–3279 (2021).
37. Nakata, M. et al. End-to-End Stacking and Liquid Crystal Condensation of 6– to 20–Base Pair DNA Duplexes. Science 318, 1276–1279 (2007).
38. Dirscherl, C. F. et al. A heated rock crack captures and polymerizes primordial DNA and RNA. Phys. Chem. Chem. Phys. 25, 3375–3386 (2023).
39. Fraccia, T. P. et al. Abiotic ligation of DNA oligomers templated by their liquid crystal ordering. Nat Commun 6, 6424 (2015).
40. Rubio-Sánchez, R. et al. Thermally Driven Membrane Phase Transitions Enable Content Reshuffling in Primitive Cells. J. Am. Chem. Soc. 143, 16589–16598 (2021).
41. Smokers, I. B. A., Van Haren, M. H. I., Lu, T. & Spruijt, E. Complex Coacervation and Compartmentalized Conversion of Prebiotically Relevant Metabolites. ChemSystemsChem 4, e202200004 (2022).
42. Bhowmik, S. & Krishnamurthy, R. The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA. Nat. Chem. 11, 1009–1018 (2019).
43. Aumiller, W. M., 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 32, 10042–10053 (2016).
44. Costanzo, G. et al. Non-Enzymatic Oligomerization of 3', 5' Cyclic AMP. PLoS ONE 11, e0165723 (2016).
45. Paloni, M., Bussi, G. & Barducci, A. Arginine multivalency stabilizes protein/RNA condensates. Protein Science 30, 1418–1426 (2021).
46. Renger, R. et al. Co-condensation of proteins with single- and double-stranded DNA. Proceedings of the National Academy of Sciences 119, e2107871119 (2022).
47. Alberts, B. et al. Molecular Biology of the Cell. (Garland Science, 2002).
48. Joyce, G. F. & Szostak, J. W. Protocells and RNA Self-Replication. Cold Spring Harbor perspectives in biology 10, a034801–a034801 (2018).
49. Szostak, J. W. The eightfold path to non-enzymatic RNA replication. Journal of Systems Chemistry 3, 1–14 (2012).
50. Gallivan, J. P. & Dougherty, D. A. A Computational Study of Cation–π Interactions vs Salt Bridges in Aqueous Media: Implications for Protein Engineering. J. Am. Chem. Soc. 122, 870–874 (2000).
51. Garaizar, A., Sanchez-Burgos, I., Collepardo-Guevara, R. & Espinosa, J. R. Expansion of Intrinsically Disordered Proteins Increases the Range of Stability of Liquid–Liquid Phase Separation. Molecules 25, 4705 (2020).
52. Wilson, K. A., Kung, R. W., D'souza, S. & Wetmore, S. D. Anatomy of noncovalent interactions between the nucleobases or ribose and π-containing amino acids in RNA–protein complexes. Nucleic Acids Research 49, 2213–2225 (2021).
53. Brown, R. F., Andrews, C. T. & Elcock, A. H. Stacking Free Energies of All DNA and RNA Nucleoside Pairs and Dinucleoside-Monophosphates Computed Using Recently Revised AMBER Parameters and Compared with Experiment. J. Chem. Theory Comput. 11, 2315–2328 (2015).
54. Oweida, T. J., Kim, H. S., Donald, J. M., Singh, A. & Yingling, Y. G. Assessment of AMBER Force Fields for Simulations of ssDNA. J. Chem. Theory Comput. 17, 1208–1217 (2021).
55. Choi, S., Meyer, M. O., Bevilacqua, P. C. & Keating, C. D. Phase-specific RNA accumulation and duplex thermodynamics in multiphase coacervate models for membraneless organelles. Nat. Chem. 14, 1110–1117 (2022).
56. Nakashima, K. K., Vibhute, M. A. & Spruijt, E. Biomolecular Chemistry in Liquid Phase Separated Compartments. Front. Mol. Biosci. 6, (2019).

57. Vieregg, J. R. et al. Oligonucleotide–Peptide Complexes: Phase Control by Hybridization. J. Am. Chem. Soc. 140, 1632–1638 (2018).

Compositional diversity of minimal coacervates in a nucleic acid-peptide world

Karina K. Nakashima,[1,2,†] Fatma Zohra Mihoubi,[1,2,†] Kieran O. Russell,[3] Fidan Rahmatova,[1,2] Jagandeep S. Saraya,[4] James D. Robinson,[4] Maria Julia Maristany,[3,5] Jan Huertas,[3,6] Roger Rubio-Sánchez,[7] Derek K. O'Flaherty,[4] Rosana Collepardo-Guevara,[3,5,6*] and Claudia Bonfio[1,2,*]

[1]Institut de Science et d'Ingénierie Supramoléculaires, CNRS UMR 7006, University of Strasbourg, 67000 Strasbourg, France

[2]Department of Biochemistry, University of Cambridge, CB2 1GA Cambridge, UK

[3]Yusuf Hamied Department of Chemistry, University of Cambridge, CB2 1EW Cambridge, UK

[4]Department of Chemistry, University of Guelph, ON N1G 2W1, Guelph, Canada

[5]Cavendish Laboratory, Department of Physics, University of Cambridge, CB3 0HE, Cambridge, UK

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

[7]Department of Chemical Engineering and Biotechnology, University of Cambridge, CB3 0AS Cambridge, UK

[†]these authors contributed equally

[*]corresponding authors: rc597@cam.ac.uk, cb2036@cam.ac.uk

Supplementary Information

Contents

Materials and Methods

Reagents were purchased from Merck and Thermo Fisher and used without further purification unless otherwise stated. *N*-benzoyl-dA, *N*-isobutyryl-dG, *N*-acetyl-dC and dT phosphoramidites, and 2′-*O*-TBDMS protected, *N*-benzoyl-rA, *N*-isobutyryl-rG, *N*-acetyl-rC and rU phosphoramidites, and 6-FAM amidite (CLP-9777) were purchased from ChemGenes (Wilmington, MA). Oligonucleotides were purchased from Integrated DNA Technologies (IDT) and Eurofins or synthesised in-house when indicated. Peptides were purchased as TFA salts from GenScript or synthesised in-house when indicated. Sep-Pak C18 classic cartridge was purchased from Waters (Milford, MA). Water coming into contact with DNA/RNA oligomers was 18 MΩ grade.

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 x 25 mm 5 µm particle size column from DrMaisch GmbH. Purification of oligonucleotides was performed using a DNApac™ PA200 column with a Vanquish™ analytical purification high-performance liquid chromatography (HPLC) system. DNA and RNA melting temperatures and base pairing probabilities were assessed using NUPACK 4.0 (https://www.nupack.org/). pH monitoring was performed using 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)).

Coacervates were imaged using a Nikon Eclipse TS2 inverted epifluorescence microscope equipped with a Moment A21K635003 camera (0.63× adaptor) and a 60× oil immersion objective. Alternatively, coacervates were imaged using a Zeiss Axio Observer Z1 spinning disk confocal microscope equipped with a Yokogawa CSU confocal head and a 63× oil immersion objective. Images were processed using Fiji (http://rsb.info.nih.gov/ij/). Thermal studies were performed using a home-built Nikon Eclipse Ti-E inverted microscope equipped with a 20× objective lens (Nikon, Plan Fluor, N.A. 0.75) and a Grasshopper3 GS3-U3-23S6M camera (Point Gray Research). The illumination was provided by single-colour light-emitting diodes (LEDs) using a filter set for Texas Red. Temperature ramps were performed using a custom-built script, enabling precise manipulation of the instrument in terms of time, temperature, and illumination as required. FRAP experiments were performed using a Leica TCS SP5 laser scanning confocal microscope (Cavendish Laboratory, Cambridge) equipped with an HCX PL Apo 40× DRY (NA 0.85) objective lens and a HeNe laser (633 nm, 10 mW). DNA and RNA oligonucleotides (ONs) were synthesised using an ABI-394 DNA synthesiser. UV measurements on oligonucleotides were taken at 260 nm using an Agilent BioTek Epoch Microplate Spectrophotometer, reading each sample at least 3 times and correcting each value by a blank measurement. Polyacrylamide gels were imaged on an Amersham TYPHOON using the Cy2 laser at 25-50 µm pixel size.

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

Sample Preparation

Stock solutions. Peptide stocks were prepared in MilliQ water at a concentration of 100 mM based on the molecular weight of the TFA salt. The pH of E_{10} (glutamic acid decamer) was adjusted with ammonia for complete dissolution. Solutions were sonicated, stored at -20°C, and vortexed for 1 minute before use. Single-strand DNA and RNA oligonucleotide stocks were prepared in DNase/RNase-free water at a strand concentration of approximately 1 mM. To facilitate the solubilisation of the oligonucleotides, solutions were heated to 50°C for 5 minutes and cooled down to room temperature before the measurement. The concentration was checked on a diluted solution (250-500×), measuring the absorbance at 260 nm.

Coacervate preparation. Coacervates were prepared in a 10-100 µL scale by adding, respectively, MilliQ water, HEPES buffer (from a 500 mM stock buffer solution, pH 7.4), DNA or RNA (~1 mM oligonucleotide stock) and peptide (100 mM stock). Aptamer and fluorescent probes were added at last unless otherwise stated. Mixing was done by gently tapping the microtube to avoid reducing droplet size for imaging. Mixtures were assessed by light microscopy to confirm the presence of liquid droplets. Note: in FRAP and aptamer experiments, the peptide was added at last to enable the fast incorporation of fluorescent dyes.

Preparation of observation chambers. A passivated glass coverslip #1.5 was used as the observation surface in all experiments. Glass passivation was performed to prevent wetting. A 5 wt% solution of partially hydrolysed polyvinyl alcohol (PVA, 13-23k) was spread on top of clean coverslips and let adsorb for 1 hour inside a covered petri dish. The coverslips were rinsed thoroughly with distilled water and once with MilliQ before being dried with compressed air. For long imaging experiments (FRAP, K_p measurements, thermal ramps), 2-6 µL chambers were prepared using double-sided 3M tape (GPT-020F, 0.2 mm) and a hole-punch (2-4 mm Ø) and sealed using 10 mm Ø coverslips to prevent evaporation.

Experimental Methods

Fmoc solid-phase peptide synthesis (Fmoc-SPPS). R_4, R_2, RER, RG, GR, RE, $R_2E_2R_2$, $(D\text{-}R\text{-}L\text{-}R)_2$ and $(D\text{-}R\text{-}L\text{-}R)_4$ were synthesised according to previously published SPPS procedures by using Fmoc-protected amino acids.[1] 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 minutes. 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 minutes before the removal of the solvent. The treatment was repeated with 20% piperidine in DMF and left shaking for 15 minutes. The solvent was removed, and the resin was washed with DCM/DMF (5 x 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 minutes. DMF washes of the resin (5 x 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 (3x), DCM (3x), and MeOH (3x) and left to dry under a high vacuum overnight.

For cleavage, the resin was treated with the cleavage solution (trifluoroacetic acid (TFA):H_2O:triisopropyl silane 95:2.5:2.5 volume ratio) for 2 hours. TFA-peptide solutions were collected,

and the resin was washed with TFA (2x3 mL). The collected fractions were concentrated under nitrogen flow and added to cold diethyl ether, leading to the precipitation of the peptide. The precipitate was centrifuged for 5 minutes at 3000 rpm 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 hour. UV absorption at 220 nm and 254 nm was used to monitor the collection of unprotected peptides. Fractions containing the target product were identified by mass spectrometry and lyophilised.

$R_2E_2R_2$: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 4.29–4.15 (m, 4H), 3.97 (t, J = 6.4 Hz, 1H), 3.25–2.89 (m, 8H), 2.47–2.18 (m, 4H), 2.02–1.30 (m, 20H).
(D-R-L-R)$_2$: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 4.30-3.20 (m, 2H), 4.05-3.95 (m, 1H), 3.50–3.02 (m, 9H), 1.97–1.28 (m, 16H).
(D-R-L-R)$_4$: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 4.44–3.94 (m, 7H), 3.33–2.86 (m, 16H), 2.05–1.28 (m, 32H).
R_4: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 3.95-3.86 (m, 3H), 3.19-3.10 (m, 8H), 1.95–1.79 (m, 8H), 1.73–1.51 (m, 8H).
RER: ^{1}H-NMR (500 MHz, D_2O) δ (ppm) 4.34-4.28 (m, 1H), 4.28-4.22 (m, 1H), 3.97 (t, J = 6.4 Hz, 1H), 2.42 (t, J = 7.4 Hz, 2H), 2.08–1.64 (m, 6H), 1.63–1.44 (m, 4H).
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).

Oligonucleotide solid-phase synthesis. DNA and RNA oligonucleotides were assembled using standard reagents and standard manufacturer protocols on a 1 µmol scale. DMTr-removal reagent consisted of 3% trichloroacetic acid in dichloromethane, the activator consisted of 0.25 M 5-ethylthio tetrazole in acetonitrile, the oxidiser consisted of a 0.02 M solution of iodine in pyridine:water:tetrahydrofuran (8:16:76 volume ratio), and the capping reagents consisted of (Cap A) a solution of acetic anhydride:pyridine:tetrahydrofuran (10:10:80 volume ratio) and (Cap B) a 10% (v/v) solution of *N*-methylimidazole in tetrahydrofuran. All oligonucleotides were deprotected from the solid support using 25% ammonium hydroxide:ethanol 4:1 volume ratio (1mL total volume) for 17 h at 55°C and concentrated in a Savant SC 110A SpeedVac® Plus to a pellet. Oligonucleotides were then purified by ion exchange chromatography.

RNA oligomers were desilylated in DMSO:triethylamine trihydrofluoride 2:3 volume ratio (100 µL:150 µL) for 2 hours at 65°C and then precipitated in cooled 1-butanol for 1 hour. Upon centrifugation, the pellet was recovered, and the supernatant was discarded. The pellet was further washed with 200 µL of 1-butanol. Deprotected oligonucleotides were purified by Strong Anion-Exchange (SAX) HPLC with solvent A (50 mM Tris buffer pH 7.6, 10% v/v MeCN) and solvent B (50 mM Tris pH 7.6, 1 M NaCl, 10% v/v MeCN),

with a standard gradient of 0-75% over 15 minutes. Purified samples were desalted using Sep-Pak C18 Classic Cartridge (WaterTM). The Sep-Pak C18 cartridge was conditioned with 10 mL of MeCN, 10 mL of MeCN:water 1:1 volume ratio and 10 mL of 100 mM pH 7 NaOAc. The purified oligo was diluted to at least 2% v/v MeCN (1:4 dilution with water) and flowed through the cartridge at least twice for column loading. The bound oligonucleotide was washed with water (~25 mL), eluted from the column with 4 mL of MeCN:water 1:1 volume ratio, and concentrated into a pellet using a DNA concentrator.

Determination of the critical salt concentration (CSC). The robustness of complex coacervates is commonly assessed by their stability to salt, typically NaCl. The critical salt concentration corresponds to the highest NaCl concentration tolerated before the complete dissolution of coacervates. Turbidity was indirectly measured on a plate reader, reading the absorbance at 600 nm and using the relation:

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

Samples of 100 µL (or 20 µL in the case of peptide/RNA mixtures) were prepared in 96-well plates (or 384-well plates) and titrated with concentrated stocks of NaCl (1, 3 or 5 M). The concentration of the salt stock was chosen to minimise the dilution of the sample during titrations (20% maximum dilution) and maximise the number of points measured during the steep decay of absorbance. At the end of the titration, all mixtures reached the turbidity of the blank (100 µL of MilliQ). The titration curves have a sigmoidal shape, and the CSC was calculated as follows: (i) the exact concentration of NaCl was calculated at each point, taking into account the total volume in the well; (ii) the curve (turbidity *vs* NaCl concentration) was fitted every three points with a linear equation; (iii) the linear fit with the highest linear coefficient (absolute value) was used to identify the tangent at the inflection point ($y = ax + b$). The CSC was thus calculated as $CSC = -\frac{a}{b}$.

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

Minimal complex concentration for coacervation. We define the minimal complex concentration as the minimal concentration of peptide:oligonucleotide 4:1 concentration ratio required for coacervation. Coacervates were prepared as in previous experiments (20 mM amino acid concentration and 5 mM nucleotide concentration, 20 µL samples), then serially diluted in a 384-well plate. Absorbance at 600 nm was converted to turbidity, and the minimal complex concentration for peptide/oligonucleotide mixtures was determined as the intercept between the x-axis and the tangent to the inflection point of the sigmoidal curve.

Temperature stability with hot stage epifluorescence microscopy. Borosilicate glass capillaries (internal section of 2 x 0.2 mm) were passivated using the same protocol as the coverslips. One capillary end was sealed with optical glue and cured under UV light (λ = 365 nm) for 5 minutes. Peptide/oligonucleotide

mixtures containing 1% of Cy3-$(TGAC)_2$ were introduced in the capillary (approx. 30 µL), which was then completely sealed with a two-component epoxy resin and hardener glue. Glass capillaries were placed on a coverslip and subsequently on a copper plate connected to a Peltier element, enabling fine control over temperature.

Fluorescence Recovery After Photobleaching (FRAP). Cy3-labelled DNA oligonucleotides (Cy3-$(TGAC)_2$, Cy3-$(TGAC)_4$ and Cy3-$(TGAC)_8$, labelled on the 5′) were chosen as FRAP probes for peptide/oligonucleotide coacervates. Coacervates were prepared as described previously, with the peptide added last to the microtube. Imaging was done ca. 30 minutes after sample preparation and placement in the observation chamber.

For each measurement, a droplet was chosen in the centre of the field of view (512x512 px) and imaged for 10 frames (every 1.117 s). A circular region of interest (ROI), selected inside the droplet (smaller than the droplet) before the acquisition, was bleached using the 633 nm laser line at 100% intensity. Post-bleaching images were collected at the same framerate until ROI intensity reached a plateau, which for our samples varied between 30-250 s (all profiles available in SI). Pre- and post-bleaching imaging was performed using the 633 nm laser line at 4-6% intensity and pinhole size set to 1 AU. A standard photomultiplier tube was used as a detector (480-720 nm). Three droplets in different FOVs were bleached for each sample, and the recovery curves were averaged.

Partition coefficients. Partitioning of fluorescent client molecules was quantified using the equation[3]

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

The fluorescence intensity inside the droplet, $I_{droplet}$, was averaged among all droplets in the field of view (FOV) using a particle analysis plugin from ImageJ and 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. Client molecules used include: Cy3-$(TGAC)_2$, Cy3-$(ACTG)_2$, FITC-r$(ACUG)_2$, Cy3-dA_{11} and Magnesium Green.

Broccoli aptamer reconstitution. A minimal version of the Broccoli aptamer was split in strand A (5′-r(GCGGAGACGGUCGGGUCCAGAUA), 23nt) and strand B (5′-r(UAUCUGUCGAGUAGAGUGUGGGCUCCGC), 27nt) and its reconstitution was followed by DFHBI fluorescence in the presence of KCl. A 2000× DFHBI stock was prepared in DMSO and diluted 100× in 25 mM HEPES buffer before being added to the sample. The samples were prepared to ensure that coacervation takes place after aptamer reconstitution by mixing in the following order (unless otherwise stated): MilliQ, 25 mM HEPES buffer, 10 mM KCl), DNA_8/RNA_8/DNA_{16} (5 mM nt), strand A (10 µM), strand B (10 µM), 5 mM DFHBI and peptide (20 mM amino acid). Measurements were performed before and after adding the peptide; fluorescence was recorded every 15 minutes for 1 hour. For microscopy, coacervates containing the Broccoli aptamer and DFHBI were prepared and left to incubate for 30 minutes in sealed microscopy chambers.

Computational Methods

Atomistic Force-Field simulations. The simulations were performed using the Amber14SB force field for peptides,[4] OL3 parameters for RNA,[5] and bsc1 for DNA.[6] All systems were solvated using the TIP4P-FB water model,[7] and compatible ion parameters were applied for sodium (Na^+) and chloride (Cl^-) ions.

Temperature control was managed by a Langevin thermostat, set to 298 K with a friction coefficient of 1 ps^{-1}. For simulations conducted in the isothermal-isobaric (NPT) ensemble, the pressure was maintained at 1 atmosphere using a Monte Carlo barostat[8] with updates applied every 25 steps.

Using the LFMiddle discretisation scheme,[9] the Langevin integrator was employed to propagate the system dynamics. Hydrogen mass repartitioning was used, enabling a time step of 4 fs during the production simulations, which was further supported by constraining all bonds involving hydrogen atoms using the CCMA algorithm.[10] Non-bonded interactions were computed with a cutoff distance of 0.9 nm. Long-range electrostatics were handled using the Particle Mesh Ewald (PME) method.[11] All simulations were performed using OpenMM 8.1.2,[12] leveraging the CUDA platform in mixed precision mode. Energy minimisation was performed using OpenMM's built-in local energy minimiser, which utilises the L-BFGS optimisation algorithm[13] until it converged to a tolerance of 10 kJ mol^{-1} $nanometer^{-1}$.

Monomer Preparation. Initial monomer structures were built using PyMOL 2.5.7.[14] Single-stranded RNA and DNA 8-mers were constructed in an extended conformation approximating B-form dihedral angles. Peptides composed of polyarginine were prepared in an extended conformation. Peptides were modelled with protonated N-termini (NH_3^+) and deprotonated C-termini (COO^-) to reflect physiological conditions. Nucleotides were prepared without the 3′ phosphate group to match experimental conditions. Each monomer's initial configuration was solvated in a cubic box with a minimum of 5 Å between any solute atom and the box edge. The systems were neutralised, and ionic strength adjusted to 30 mM NaCl. Energy minimisation was performed, followed by a 100 ns NVT equilibration at 298 K. Evenly spaced configurations from the last 80 ns of the monomer simulations were extracted to build multi-chain systems.

Multi-Chain System Preparation. Multi-chain systems were constructed by placing monomers using Packmol 20.14.4,[15] enforcing a minimum distance of 10.0 Å between any two atoms to prevent overlaps. Each system consisted of 8 nucleotides (either RNA or DNA) and 36 polyarginine peptides, with one nucleotide and four peptides placed randomly within each octant of a cubic box with a side length of 140.0 Å. The assembled systems were solvated in a cubic box 145.0 Å per side, ensuring a minimum of 5 Å between any solute atom and the box edge under periodic boundary conditions, before being neutralised and brought to an ionic strength of 100 mM NaCl.

Equilibration Protocol. After minimisation, the systems were relaxed through the following steps:

- 250 ps of NPT simulation at 298 K and 1 atm with heavy atom positional restraints of 15 kcal mol^{-1} $Å^{-2}$ applied to all peptide and nucleotide heavy atoms, using a 2 fs time step.
- 250 ps of unrestrained NPT simulation at 298 K and 1 atm, with a 2 fs time step.
- 500 ps of unrestrained NVT simulation using a 2 fs time step.

Production Simulations. Production simulations were carried out in the NVT ensemble for 800 ns per replicate, with 5 independent replicates for each system, totalling 3.0 μs of simulation time per system.

Each replicate began from an independently prepared configuration and used different random number seeds to ensure statistical independence.

Trajectory frames were saved every 0.8 ns. The last 200 ns of each simulation (corresponding to 500 frames) were used for contact analysis. Contacts between molecules were defined based on a cutoff distance of 0.45 nm between heavy atoms and were analysed using a custom Python script utilising CuPy.

Interaction Analysis. Trajectories were analysed using the Python package MDTraj.[16] Hydrogen bonds were assessed using the Wernet-Nillson criteria.[17] Electrostatic interactions were defined as occurring when the CZ atom from an arginine sidechain approaches OP1 or OP2 atoms from a phosphate group closer than 0.6 nm without a hydrogen bond being established between the two residues. Following previous work,[18] Arg-nucleobase stacking interactions were defined as occurring when the CZ atom from an arginine sidechain approaches the centre-of-geometry of a nucleobase ring with the angle between the planes of the guanidium group and the nucleobase ring less than 30°.

Supplementary Tables

Supplementary Table 1. Peptide sequences used in the present work. All sequences are written *N*- to *C*-termini, are not protected (*i.e.*, H-peptide-OH) and used as TFA salts (SIH = Synthesised in-house).

Peptide sequence	Net charge	Charge density	Source
Dimers			
RG	1	0.5	SIH
GR	1	0.5	SIH
RE	0	0	SIH
R_2	2	1	SIH
Trimers			
R_3	3	1	GenScript
RGR	2	0.67	GenScript
RER	1	0.33	SIH
RYR	2	0.67	GenScript
Tetramers			
R_4	4	1	SIH, GenScript
K_4	4	1	GenScript
RKRK	4	1	GenScript
$(D\text{-}R\text{-}L\text{-}R)_2$	4	1	SIH
RG_2R	2	0.5	GenScript
R_4-HCl	4	1	GenScript
Pentamers			
R_5	5	1	GenScript
R_2GR_2	4	0.8	GenScript
Hexamers			
$R_2G_2R_2$	4	0.67	GenScript
R_6	6	1	GenScript
$R_2E_2R_2$	2	0.33	SIH
Other lengths			
$R_2G_4R_2$	4	0.5	GenScript
$R_2G_8R_2$	4	0.33	GenScript
R_7	7	1	GenScript
R_8	8	1	GenScript
$(D\text{-}R\text{-}L\text{-}R)_4$	8	1	SIH
R_9	9	1	GenScript
R_{10}	10	1	GenScript
E_{10}	-10	1	GenScript

Supplementary Table 2. Oligonucleotide acronyms and sequences used in the present work (SIH = synthesised in-house).

Name	Sequence	Source
dA_3	5′-AAA-3′	SIH
DNA_5	5′-ACTGA-3′	IDT
DNA_6	5′-ACTGAC-3′	IDT
DNA_7	5′-ACTGACT-3′	IDT
DNA_8	5′-ACTGACTG-3′	IDT
DNA_{10}	5′-ACTGACTGAC-3′	IDT
DNA_{12}	5′-ACTGACTGACTG -3′	IDT
DNA_{16}	5′-ACTGACTGACTGACTG-3′	IDT
DNA_{20}	5′-ACTGACTGACTGACTGACTG-3′	IDT
DNA_{40}	5′-ACTGACTGACTGACTGACTGACTGACTGACTGACTG-3′	IDT
dA_{10}	5′-AAAAAAAAAA-3′	IDT
dT_{10}	5′-TTTTTTTTTT-3′	IDT
dC_{10}	5′-CCCCCCCCCC-3′	IDT
dG_{10}	5′-GGGGAGGGGA-3′	IDT
Cy3-5nt	5′-TCAGT-Cy_3-3′	IDT
Cy3-8nt	5′-Cy3-TGACTGAC-3′	IDT
Cy3-16nt	5′-Cy3-TGACTGACTGACTGAC-3′	IDT
Cy3-32nt	5′-Cy3-TGACTGACTGACTGACTGAC-3′	IDT
Cy3-dA_{11}	5′-Cy3-AAAAAAAAAA-3′	IDT
Cy3-10nt	5′-/5Cy3/TGTGCCAGTA-3′	IDT
DNA_8* [a]	5′-CAGTCAGT-3′	IDT
HNA_8	5′-ArCrUGArCrUG-3′	IDT
RNA_8	5′-r(ACUGACUG)-3′	IDT, SIH
HNA_{12}	5′-ArCrUGArCrUGArCrUG-3′	IDT
RNA_{12}	5′-r(ACUGACUGACUG)-3′	IDT, SIH
RNA_{20}	5′- r(ACUGACUGACUGACUGACUG)-3′	IDT, SIH
dA_{11}	5′-AAAAAAAAAAA-3′	Eurofins
dA_{21}	5′-AAAAAAAAAAAAAAAAAAAAA-3′	Eurofins
dA_{31}	5′-AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA-3′	Eurofins
dA_{41}	5′-AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA-3′	Eurofins
dA_{51}	5′-AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA-3′	Eurofins
dA_{12}	5′-AAAAAAAAAAAA-3′	SIH
dA_{16}	5′-AAAAAAAAAAAAAAAA-3′	SIH
rA_{12}	5′-r(AAAAAAAAAAAA)-3′	SIH
rA_{16}	5′-r(AAAAAAAAAAAAAAAA)-3′	SIH
FAM-RNA_8	5′-FAM-r(ACUGACUG)-3′	SIH
FAM-DNA_8	5′-FAM-ACTGACTG-3′	SIH
DNA_{11}-Phos	5′-ACTGACTGACT-Phos-3′	IDT
Phos-DNA_{10}-Phos	5′-Phos-CTGACTGACT-Phos-3′	IDT
Broccoli aptamer A	5′-r(GCGGAGACGGUCGGGUCCAGAUA)-3′	Eurofins

Broccoli aptamer B	5′-r(UAUCUGUCGAGUAGAGUGUGGGCUCCGC)-3′	Eurofins

[a] DNA_8* is the complementary sequence to DNA_8

Supplementary Table 3. Critical salt concentrations (CSCs) measured for mixtures comprising 20 mM amino acid and 5 mM nucleotide (unless otherwise stated) in 25 mM HEPES pH 7.4 and room temperature.

Peptide	Oligonucleotide	CSC (mM, NaCl)	Peptide	Oligonucleotide	CSC (mM, NaCl)
R_4	DNA_8	99.3 ± 4.7	R_8	DNA_8	469.3
R_4	DNA_{12}	196.4	R_{10}	dA_{10}	258.2
R_4	DNA_{16}	212.6	R_{10}	dT_{10}	279.0
R_4	DNA_{20}	201.6 ± 19.3	R_{10}	DNA_{10}	583.8 ± 37.1
R_4	DNA_{40}	278.1	R_6	DNA_6	143.6
R_4	polyU	492.9	R_7	DNA_7	369.3
R_3	DNA_{12}	39.9	R_8	DNA_8	536.7
R_4	DNA_{12}	196.4	$(D\text{-}R\text{-}L\text{-}R)_2$	DNA_8	78.3
R_6	DNA_{12}	409.8	$(D\text{-}R\text{-}L\text{-}R)_2$	DNA_{12}	152.8
R_8	DNA_{12}	556.5	$(D\text{-}R\text{-}L\text{-}R)_4$	DNA_8	424.9
R_{10}	DNA_{12}	634.2	$(D\text{-}R\text{-}L\text{-}R)_4$	DNA_{12}	509.8
R_6	DNA_8	313.7	R_4	RNA_8	215.9
R_6	DNA_{12}	421.9	R_4	RNA_{12}	379.2
R_6	DNA_{16}	475.1	R_4	RNA_{20}	430.7
R_6	DNA_{20}	473.8	R_3	RNA_{12}	205.2
R_6	DNA_{40}	500.0	R_6	RNA_{12}	693.8
R_6	dA_{10}	99.9	R_3	RNA_8	54.2
R_6	dA_{11}	123.4	R_6	RNA_8	210.7
R_6	dA_{15}	153.5	R_8	RNA_8	402.6
R_6	dA_{31}	243.5	R_3	HNA_8	59.7
R_6	dA_{41}	256.6	R_4	HNA_8	113.7
R_6	dA_{51}	269.8	R_6	HNA_8	222.1
R_8	DNA_{12}	647.0	R_8	HNA_8	368.6
R_8	DNA_{16}	743.5	R_3	HNA_{12}	52.7
R_8	DNA_{20}	739.5	R_4	HNA_{12}	237.8
R_{10}	E_{10}	95.2 ± 0.4[a]	R_6	HNA_{12}	431.9

[a] Measured at 10 mM [Arg] and 10 mM [Glu]

Supplementary Table 4. Critical salt concentrations (CSCs) measured for peptide mixtures at 25 mM HEPES pH 7.4 and room temperature.

Composition	[aa] (mM)	[nt] (mM)	CSC (mM)
R_3/DNA_{20}	20	5	85.0
R_4/DNA_{20}	20	5	207.4 ± 27.2
R_5/DNA_2	20	5	348.1
$R_3/R_4/R_5/DNA_{20}$	20	5	230.5
R_2GR_2/DNA_{20}	25	5	204.6
$R_2G_2R_2/DNA_{20}$	30	5	237.8
$R_2G_4R_2/DNA_{20}$	40	5	163.3
$R_2G_8R_2/DNA_{20}$	60	5	140.8
$R_4/R_2GR_2/R_2G_2R_2/DNA_{20}$	20	5	200.1
R_2GR_2/DNA_{20}	20	5	163.3
$R_2G_2R_2/DNA_{20}$	20	5	170.0
$R_2G_4R_2/DNA_{20}$	20	5	143.4
$R_2G_8R_2/DNA_{20}$	20	5	85.2
$RKRK/DNA_{20}$	20	5	91.5
$R4/RKRK/DNA_{20}$	13.3	5	182.0
R_4/DNA_{20}	6.67	5	174.8
$RKRK/DNA_{20}$	6.67	5	50.3
K_4/DNA_{20}	6.67	5	0
$K_4/RKRK/DNA_{20}$	13.3	5	82.8
$R_4/K_4/DNA_{20}$	13.3	5	141.5
$R_4/RGGR/DNA_{20}$	13.3	5	147.1
R_3/DNA_{20}	6.67	5	19.8
$R_3/RER/DNA_{20}$	13.3	5	32.1
$R_3/RER/RGR/DNA_{20}$	20	5	40.0
$R_4/R_1(3:1)/DNA_8$	20	5	84.3
$R_4/R_1(1:1)/DNA_8$	20	5	40.4
$R_4/R_1(3:1)/DNA_{20}$	20	5	201.9
$R_4/R_1(1:1)/DNA_{20}$	20	5	139.8
$R_3/DNA_{12}:RNA_{12}$ (1:1)	20	5	44.7
$R_4/DNA_{12}:RNA_{12}$ (1:1)	20	5	258.5
$R_6/DNA_{12}:RNA_{12}$ (1:1)	20	5	422.4

Supplementary Table 5. Parameters calculated from the linear fit (CSC = a(1/N) + b) for the CSCs of peptides and oligonucleotides when one (or both) lengths are varied.

Peptide	Oligonucleotide	a	b	R^2	Calculated N_{min} (-a/b)	Empirical N_{min}
$R_2G_2R_2$	DNA_N	-1640.8	256.5	0.89	6.4	7
R_4	DNA_N	-1963.3	332.1	0.89	5.9	7
R_6	DNA_N	-2107.8	564.4	0.81	3.7	5
R_6	$polyA_N$	-2105.7	308.3	0.99	6.8	11
R_4	RNA_N	-2906.4	59.22	0.95	4.9	8[a]
R_N	DNA_{12}	-2561.6	866.9	0.99	2.9	3
R_N	DNA_{20}	-2951.5	1007.6	0.90	2.9	3
R_N	DNA_N	-7477.5	1420.2	0.94	5.3	6
R_N	DNA_8	-2588.7	742.7	0.96	3.5	4
R_N	HNA_8	-1610.3	564.4	0.89	3.0	3
R_N	RNA_8	-1653.8	616.9	0.99	2.7	3
R_N	RNA_{12}	-2835.8	1104.2	0.96	2.6	3
R_N	HNA_{12}	-2275.5	809.7	0.99	2.8	3
R_N	RNA_{12}:DNA_{12} 1:1	-2266.0	808.4	0.99	2.8	3

[a] Shortest RNA oligonucleotide tested

Supplementary Table 6. Critical salt concentrations (CSCs) measured for mixtures of peptides and oligonucleotides (phase diagram studies). Concentrations refer to amino acid (aa) or nucleotide (nt) concentrations.

Mixture	[aa] (mM)	[nt] (mM)	CSC (mM)	Mixture	[aa] (mM)	[nt] (mM)	CSC (mM)
R_4/DNA_8	2	5	0	R_4/DNA_{20}	3	5	93.0
	3	5	22.3		5	5	123.5
	5	5	49.4		10	5	175.9
	10	5	67.3		15	5	213.8
	20	5	99.3 ± 4.7		20	5	250.5
	40	5	110.0		30	5	270.3
	20	1	0	R_8/DNA_{20}	3	5	391.5
	20	2	51.8		5	5	503.8
	20	5	102.2		10	5	503.0
	20	10	122.7		15	5	609.0
	20	15	118.6		20	5	628.8
	20	20	140.1		30	5	674.8
R_3/DNA_{12}	5	5	0	R_{10}/E_{10}[a]	5	10	0
	10	5	0		8	10	90.6 ± 5.6
	15	5	25.0		10	10	95.2 ± 0.4
	20	5	32.6		12	10	103.8 ± 22.9
	30	5	36.7		20	10	0
	40	5	20.7		10	5	0
	60	5	47.0		10	8	81.9 ± 12.9
	20	2.5	30.7		10	12	107.5 ± 6.1
	20	5	44.0		10	15	93.7 ± 5.7
	20	10	46.4		10	20	84.8
	20	15	37.8		10	40	0
	20	20	23.2				

[a] Concentrations reported are [Arg] and [Glu]

Supplementary Table 7. Amino acid concentrations required 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 Table 8. Number of contacts per oligonucleotide strand with arginine residues, as computed with atomistic simulations. Results are categorised by mode of interaction for a given peptide/nucleic acid combination. Each value in a repeat is an average value for all oligonucleotide chains of that peptide/oligonucleotide mixture over time.

Number of contacts	Mixture			
	R_3/DNA_8	R_3/RNA_8	R_4/DNA_8	R_4/RNA_8
	H-bonding			
Repeat 1	8.32	10.30	9.35	11.06
Repeat 2	7.28	9.06	10.48	10.94
Repeat 3	8.82	11.17	8.71	11.81
Repeat 4	8.15	9.38	9.16	11.68
Repeat 5	7.56	10.58	9.69	11.42
Average	8.02	10.1	9.48	11.4
Standard deviation	0.55	0.8	0.59	0.3
	Electrostatic			
Repeat 1	3.19	3.51	4.28	3.95
Repeat 2	3.38	4.37	4.30	4.24
Repeat 3	4.04	4.18	3.70	4.98
Repeat 4	3.99	4.28	4.09	4.71
Repeat 5	3.05	3.52	4.06	4.57
Average	3.53	3.97	4.09	4.49
Standard deviation	0.41	0.38	0.22	0.36
	Stacking			
Repeat 1	0.68	1.59	0.71	1.70
Repeat 2	0.72	1.30	0.77	1.48
Repeat 3	0.95	1.40	0.92	0.88
Repeat 4	0.53	1.34	0.97	2.01
Repeat 5	0.74	1.59	0.83	2.03
Average	0.726	1.44	0.840	1.62
Standard deviation	0.135	0.13	0.097	0.42

Supplementary Table 9. Number of total contacts per oligonucleotide strand made with arginine residues, as computed with atomistic simulations. The number of contacts established with unique arginine residues or unique peptide chains is also specified. Each value in a repeat is an average value for all oligonucleotide chains of that peptide/oligonucleotide mixture over time.

Number of contacts	Mixture			
	R_3/DNA_8	R_3/RNA_8	R_4/DNA_8	R_4/RNA_8
	Total contacts			
Repeat 1	12.19	15.40	14.33	16.71
Repeat 2	11.37	14.72	15.55	16.66
Repeat 3	13.81	16.75	13.34	17.68
Repeat 4	12.66	15.00	14.23	18.40
Repeat 5	11.35	15.70	14.58	18.01
Average	12.3	15.5	14.4	17.5
Standard deviation	0.9	0.7	0.7	0.7
	With unique arginine residues			
Repeat 1	6.30	7.43	7.46	8.25
Repeat 2	6.09	6.93	8.35	7.51
Repeat 3	7.36	7.61	7.38	8.61
Repeat 4	6.90	7.03	7.46	8.55
Repeat 5	5.83	6.90	7.81	8.45
Average	6.50	7.18	7.69	8.28
Standard deviation	0.56	0.29	0.36	0.40
	With unique peptide chains			
Repeat 1	3.88	4.42	3.99	4.01
Repeat 2	3.97	4.14	3.96	3.90
Repeat 3	4.45	4.55	3.73	4.61
Repeat 4	4.35	4.28	3.54	4.40
Repeat 5	3.52	4.10	3.72	4.22
Average	4.03	4.30	3.79	4.23
Standard deviation	0.34	0.17	0.17	0.25

Supplementary Table 10. Number of free peptide chains for a given mixture, as computed with atomistic simulations. Each value in a repeat is an average value for all oligonucleotide chains of that peptide/oligonucleotide mixture over time.

Number of free peptides	Mixture			
	R_3/DNA_8	R_3/RNA_8	R_4/DNA_8	R_4/RNA_8
Repeat 1	12.5	11.4	13.7	10.6
Repeat 2	14.2	8.9	11.2	11.1
Repeat 3	11.1	11.2	12.8	12.3
Repeat 4	10.6	11.0	12.1	11.8
Repeat 5	11.2	10.7	11.7	8.8
Average	11.9	10.6	12.3	10.9
Standard deviation	1.3	0.9	0.9	1.2

Supplementary Table 11. Partition coefficients calculated from confocal fluorescence microscopy images. (* = 40mM [aa]). K_p values were calculated from fluorescence intensities measured using Fiji.

Peptide	Oligonucleotide	Probe	I_{in}	I_{out}	K_p	N
R_4	DNA_8	FAM-DNA_8	12672.9	898.9	20.3 ± 4.85	201
R_4	RNA_8	FAM-DNA_8	10117.3	1805.5	15.3 ± 3.6	308
R_4	DNA_8	FAM-RNA_8	15148.5	1767.7	21.0 ± 5.5	214
R_4	RNA_8	FAM-RNA_8	11540.0	1449.1	20.5 ± 3.6	363
R_8	DNA_8	FITC-R_8	40.4	2.06	19.9 ± 9.4	519
R_8	RNA_8	FITC-R_8	37.0	1.5	24.9 ± 10.4	365
R_4	DNA_8	Broccoli aptamer	1917.9	125.6	15.3 ± 8.3	59
R_4	RNA_8	Broccoli aptamer	2704.9	249.1	11.6 ± 3.0	366
R_4	DNA_{16}	Broccoli aptamer	2797.3	224.4	11.1 ±5.8	307
R_{10}	E_{10}	Broccoli aptamer	4825.2	175.6	30.8 ± 11.3	321
R_3	DNA_{12}	Cy_3-A_{11}	5062.7	262.1	19.1 ± 5.9	120
		Cy_3-A_{31}	12193.2	517.5	23.5 ± 7.8	180
		Cy_3-A_{51}	12837.7	642.3	20.1 ± 7.2	197
R_4	DNA_8	Cy_3-A_{11}	14126.5	547.5	25.8 ± 10.3	331
		Cy_3-A_{31}	18301.1	672.2	26.8 ± 9.0	330
		Cy_3-A_{51}	14905.3	480.5	32.5 ± 11.5	324
R_4	DNA_8	Mg Green, 0 mM Mg^{2+}	3848.0	6651.0	0.58	-
R_4	RNA_8	Mg Green, 0 mM Mg^{2+}	2756.4	4383.3	0.63	-
R_{10}	E_{10}	Mg Green, 0 mM Mg^{2+}	786.1	3409.5	0.23	-
R_4	DNA_8	Mg Green, 5 mM Mg^{2+}	4122.8	8821.8	0.47	-
R_4	RNA_8	Mg Green, 5 mM Mg^{2+}	2617.9	6242.7	0.42	-
R_{10}	E_{10}	Mg Green, 5 mM Mg^{2+}	443.3	3919.1	0.11	-
R_3	DNA_{12}	Cy3-8nt	29.6	0.1	133.3 ± 54.9	133
		Cy3-16nt	49.3	0.4	108.3 ± 67.2	171
		Cy3-32nt	47.6	1.1	52.9 ± 16.9	209
R_4	DNA_8	Cy3-8nt	72.3	1.8	38.9 ± 13.3	106
		Cy3-16nt	64.8	1.6	46.2 ± 17.8	131
		Cy3-32nt	38.8	1.7	33.9 ±15.1	175
R_4	DNA_{16}	Cy3-8nt	32.2	1.8	19.1 ± 4.6	124
		Cy3-16nt	26.9	0.7	45.1 ± 19.5	89
		Cy3-32nt	25.4	1.8	13.4 ± 4.6	124
R_8	DNA_{16}	Cy3-8nt	26.4	0.4	61.4 ± 21.3	122
		Cy3-16nt	23.2	2.2	11.9 ± 4.7	209
		Cy3-32nt	18.3	0.4	42.6 ± 27.1	135
R_4	RNA_8	Cy3-8nt	40.5	1.3	30.4 ± 10.5	98
R_4	dsDNA_8[a]	Cy3-8nt	33.4	0.7	48.5 ± 16.6	40
R_{10}	dA_{10}	Cy3-dA11	58.4	0.7	83.1 ± 31.1	115
R_{10}	dT_{10}	Cy3-dA11	53.2	2.1	25.4 ± 8.2	172
R_{10}	dC_{10}	Cy3-dA11	60.1	1.6	36.9 ± 12.5	181
R_{10}	DNA_{10}	Cy3-dA11	138.1	2.4	15.9 ± 4.5	139

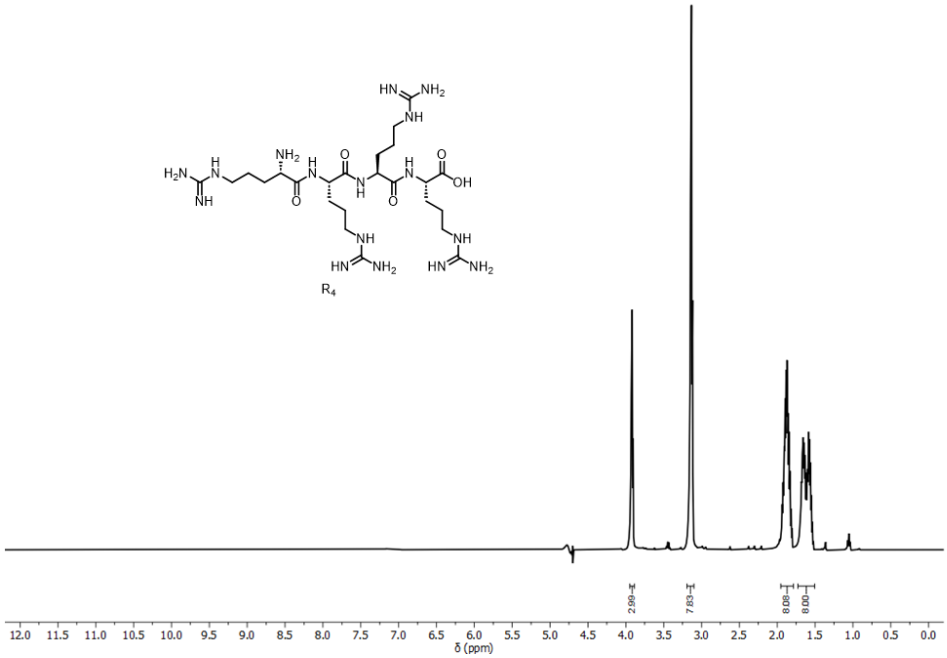
[a] ds denotes double-stranded DNA, prepared with DNA_8 and DNA_8* oligonucleotides

Supplementary Table 12. Parameters obtained from fitting FRAP profiles to the exponential decay: $y = y_0 + A_1e^{(-t/\tau)}$, where y is the normalised fluorescence and t is the time since photobleaching.

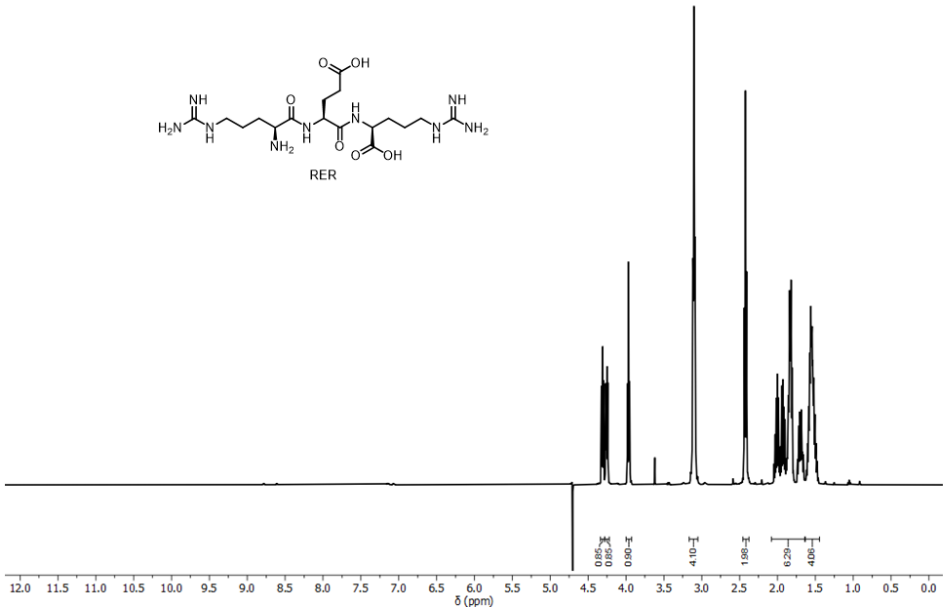
Peptide	Oligonucleotide	Probe	y_0	A_1	τ (s)	R^2
R_3	DNA_{12}	Cy3-8nt	0.96	-1.13	5.86	0.996
		Cy3-16nt	0.78	-0.79	6.73	0.994
		Cy3-32nt	0.55	-0.84	5.17	0.999
R_4	DNA_8	Cy3-8nt	0.78	-1.16	10.1 ± 5.0[a]	0.999
		Cy3-16nt	0.68	-1.08	6.64	1.00
		Cy3-32nt	0.60	-0.85	7.83	1.00
R_4	DNA_{16}	Cy3-8nt	0.79	-0.76	18.2	1.00
		Cy3-16nt	0.69	-0.60	26.1	1.00
		Cy3-32nt	0.66	-0.54	26.0	0.996
R_8	DNA_{16}	Cy3-8nt	0.60	-0.41	61.9	0.999
		Cy3-16nt	0.53	-0.37	80.6	0.999
		Cy3-32nt	0.47	-0.30	94.7	0.999
R_4	RNA_8	Cy3-8nt	0.56	-0.46	61.5 ± 4.7	0.998
R_4	DNA_8/RNA_8 (1:1)	Cy3-8nt	0.72	-0.66	26.24	0.995
R_4	$dsDNA_8$	Cy3-8nt	0.96	-0.92	8.94	0.999
		Cy3-16nt	0.55	-0.35	10.6	0.980
R_{10}	dA_{10}	Cy3-dA11	0.69	-1.18	3.36	0.994
R_{10}	dT_{10}	Cy3-dA11	0.64	-0.53	9.70	0.995
R_{10}	dC_{10}	Cy3-dA11	0.63	-0.63	6.15	0.999
R_{10}	DNA_{10}	Cy3-dA11	0.72	-0.60	28.5	0.998

[a] n = 6 replicates

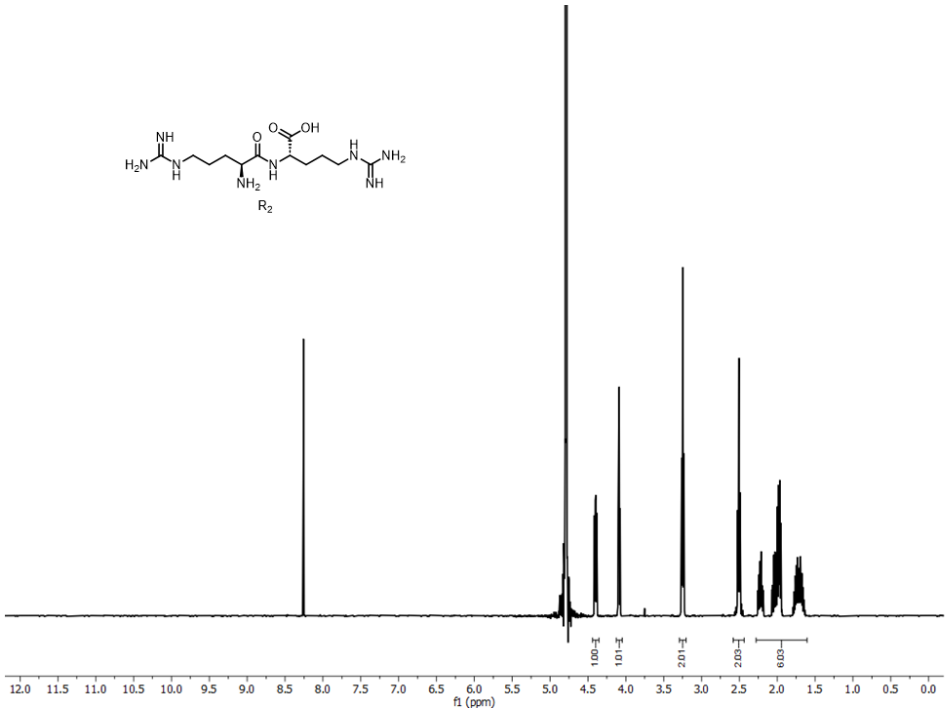
Supplementary Figures



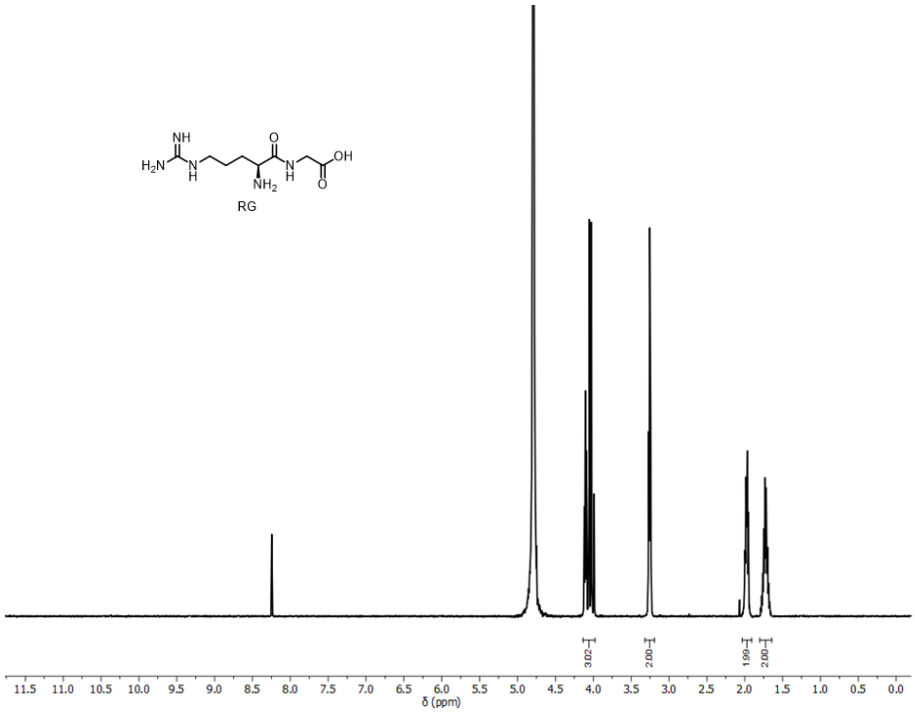
Supplementary Figure 1. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **R_4** synthesised in-house.



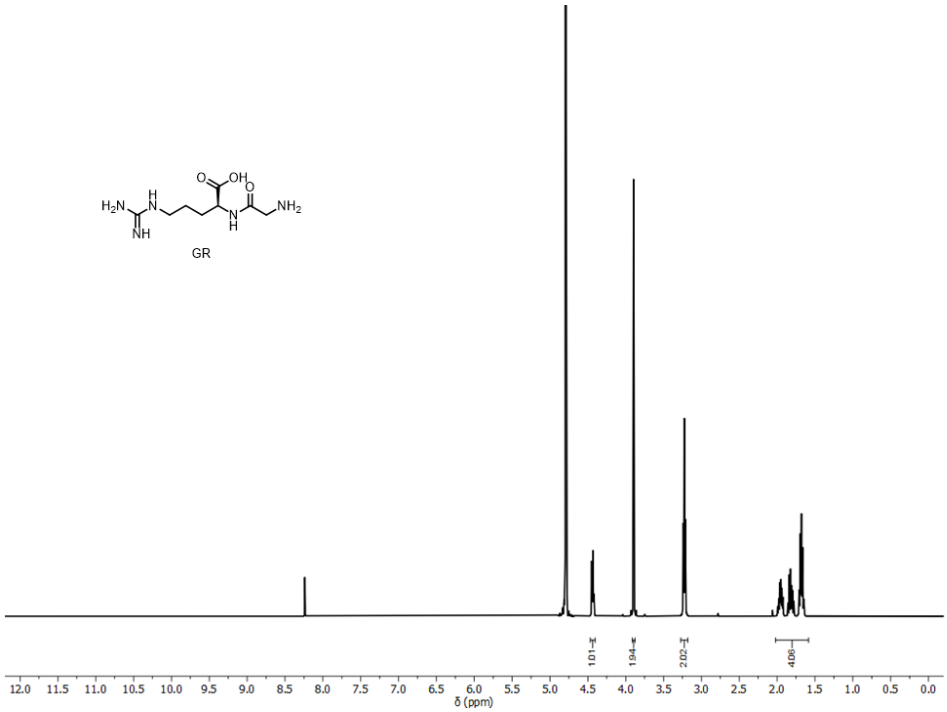
Supplementary Figure 2. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **RER** synthesised in-house.



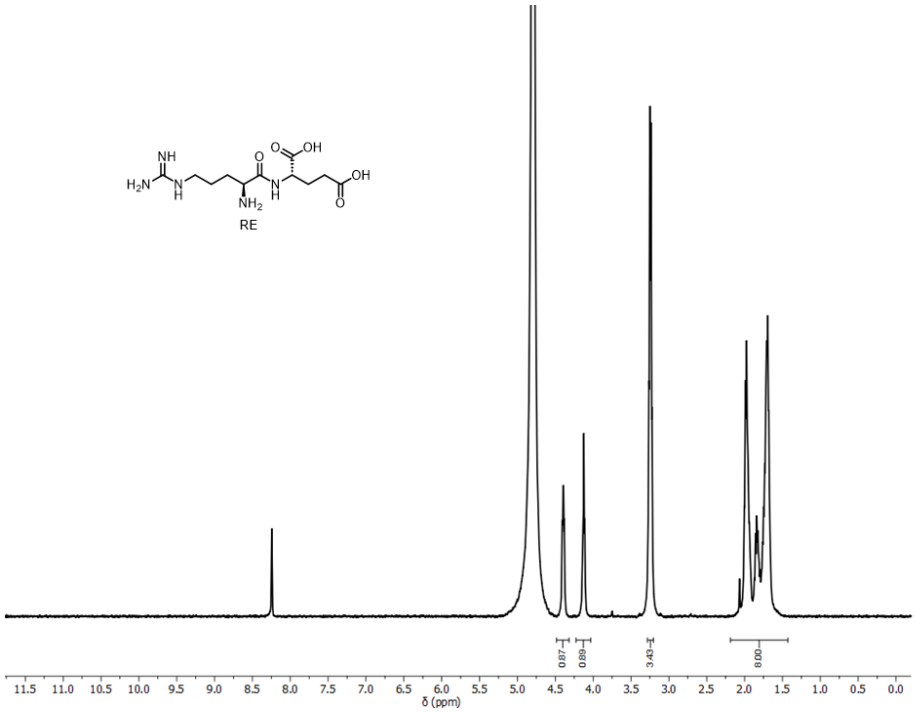
Supplementary Figure 3. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **R_2** synthesised in-house.



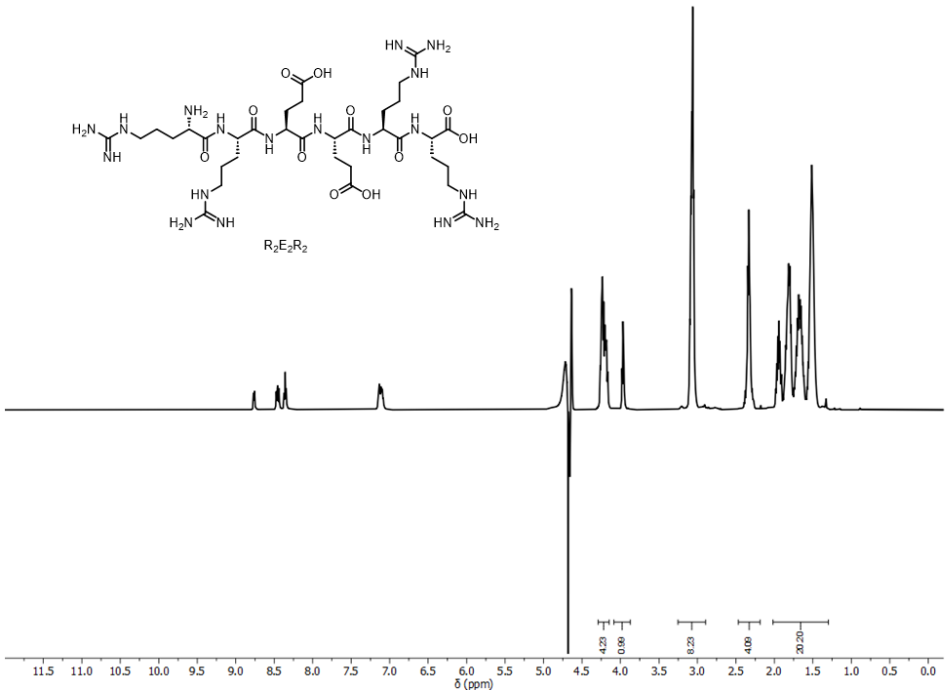
Supplementary Figure 4. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **RG** synthesised in-house.



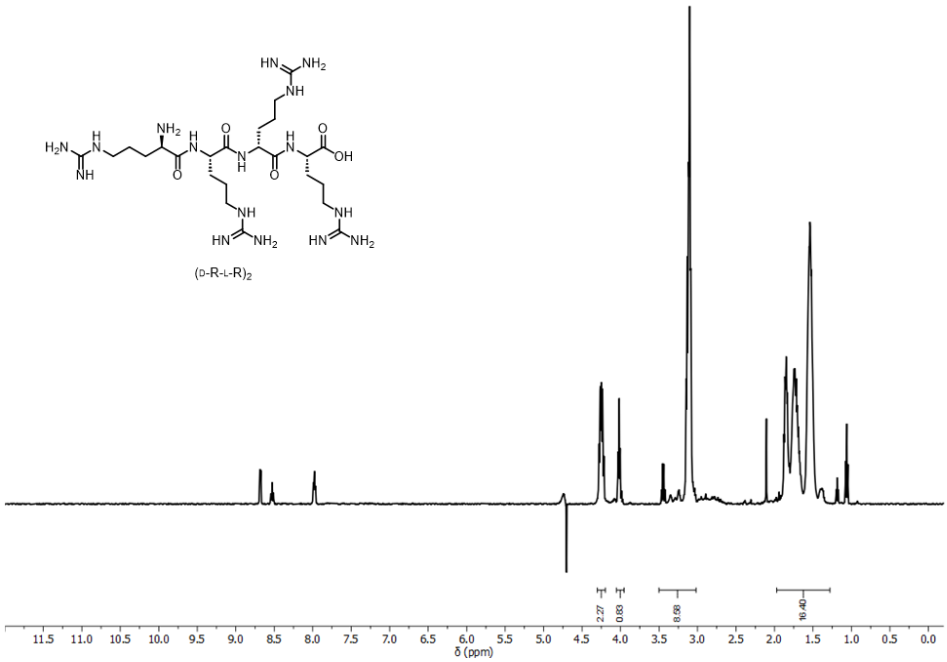
Supplementary Figure 5. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **GR** synthesized in-house.



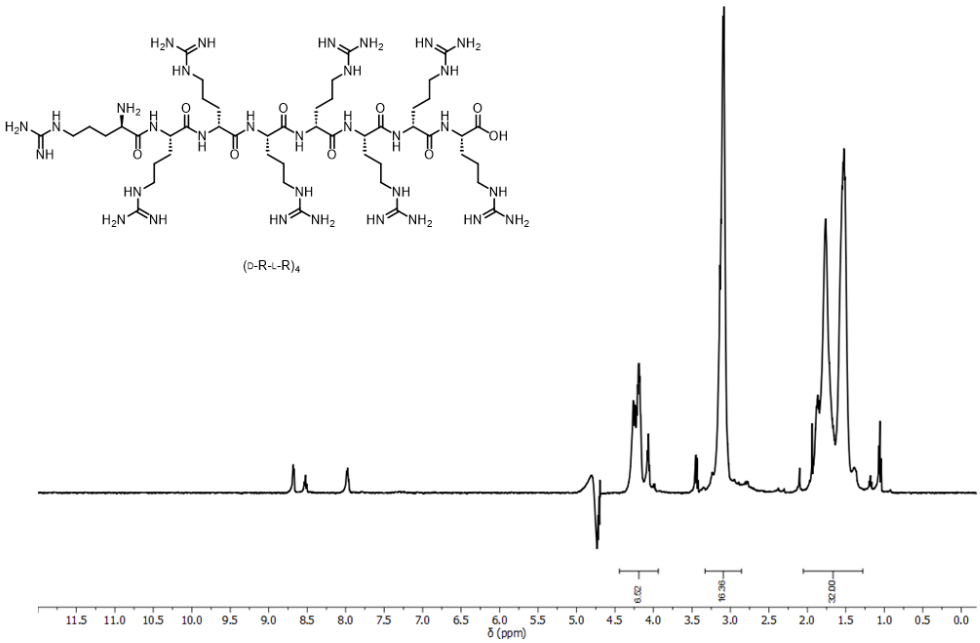
Supplementary Figure 6. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of **RE** synthesised in-house.



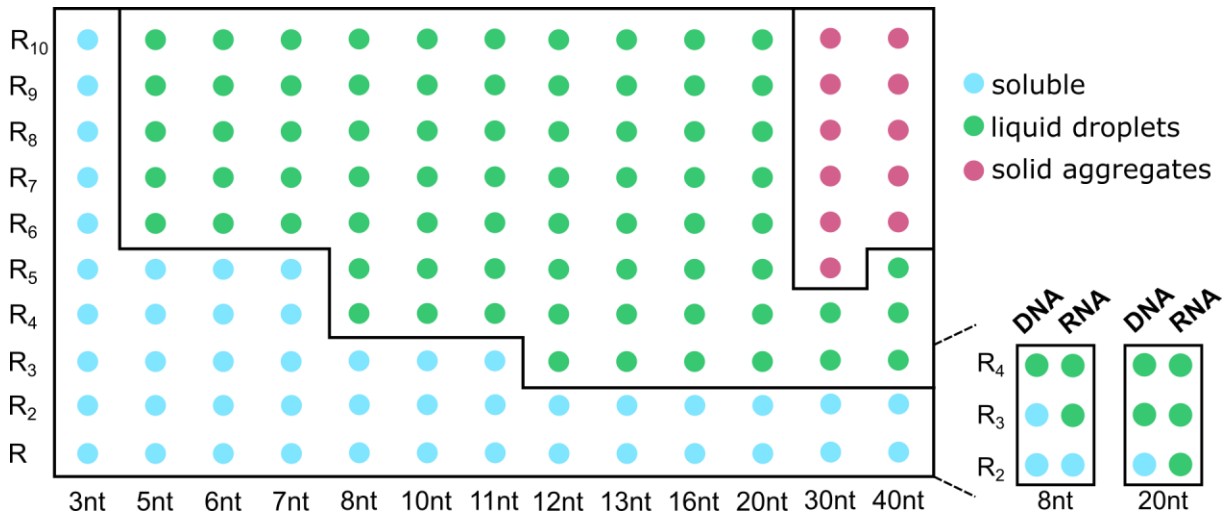
Supplementary Figure 7. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of $R_2E_2R_2$ synthesised in-house.



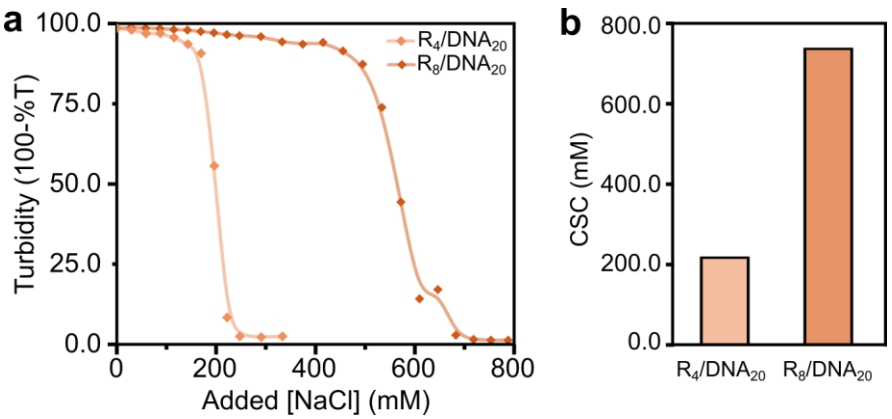
Supplementary Figure 8. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of (D-R-L-R)$_2$ synthesised in-house.



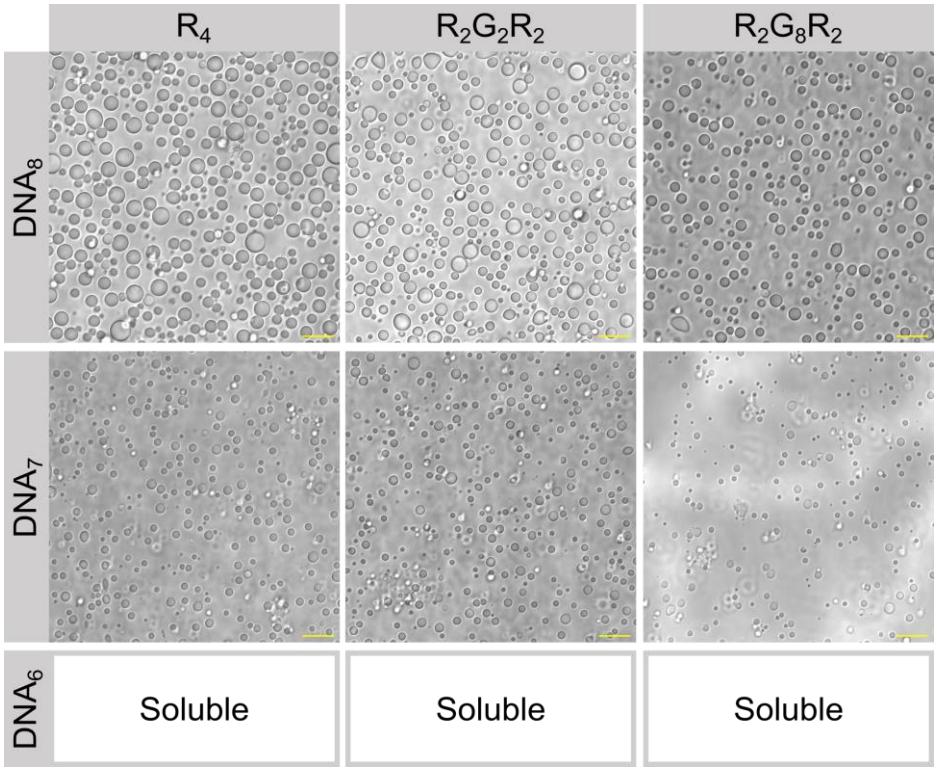
Supplementary Figure 9. ^{1}H-NMR (500 MHz, D_2O:H_2O 9:1) spectrum of (D-R-L-R)$_4$ synthesised in-house.



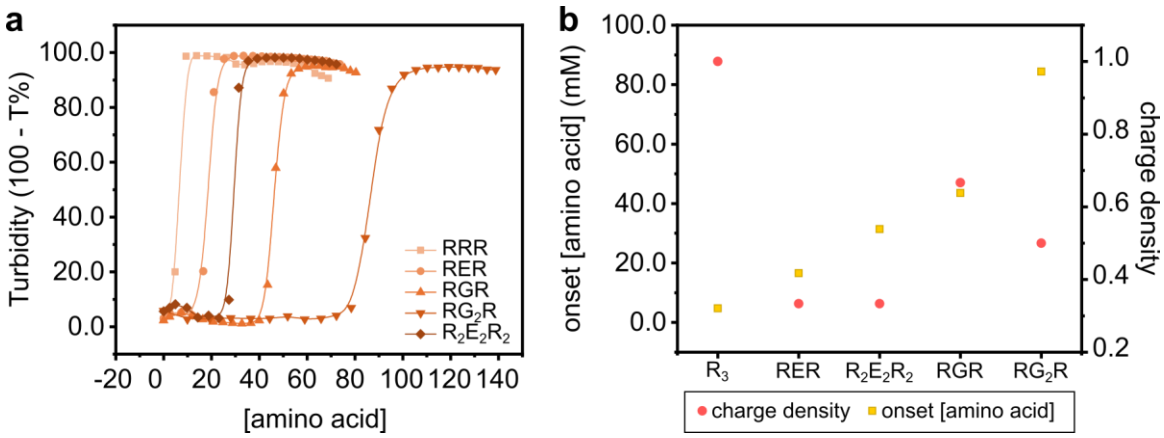
Supplementary Figure 10. Extended screening of peptide/DNA mixtures obtained by varying both peptide (vertical axis) and DNA lengths (horizontal axis). DNA sequences follow the motif ACTG, except for 3nt, which is dA_3. The inset to the right shows the screening results with RNA oligonucleotides alongside those obtained with DNA oligonucleotides of the same length (and motif). The "liquid droplets" region (green) expands with RNA. All mixtures were screened at 20 mM amino acid, 5 mM nucleotide, 25 mM HEPES pH 7.4 and room temperature.



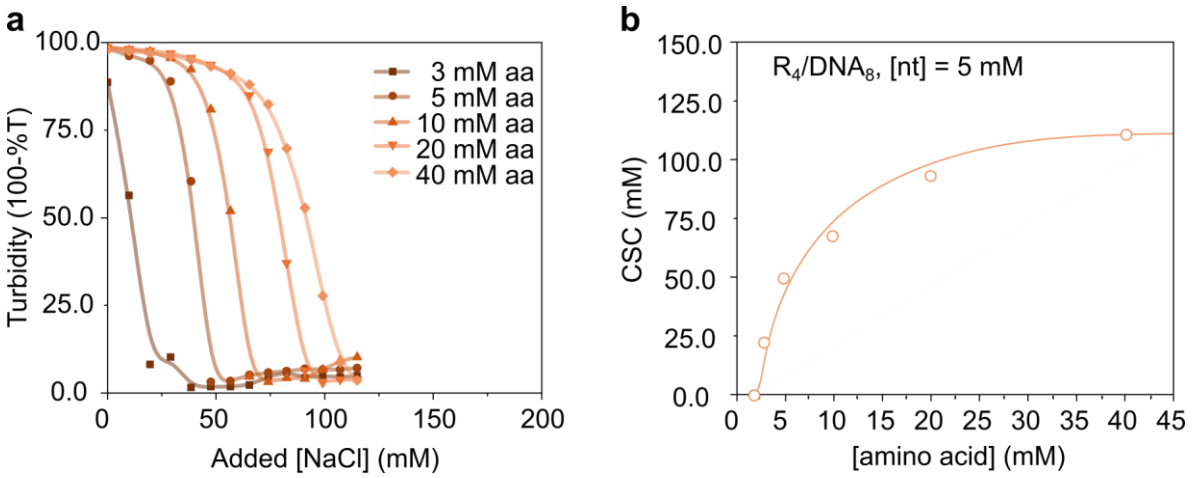
Supplementary Figure 11. (a) Salt titration curves used to determine the CSC values **(b)** for R_4/DNA_{20} and R_8/DNA_{20} at 20 mM arginine and 5 mM nucleotide concentrations.



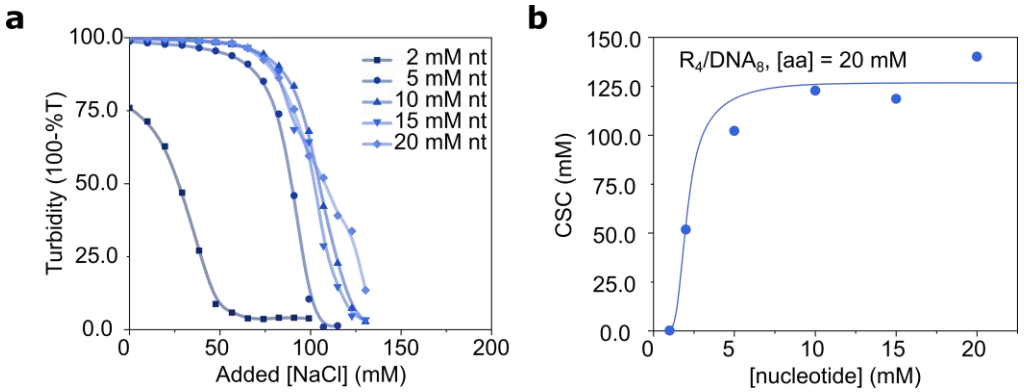
Supplementary Figure 12. Screening of minimal DNA oligonucleotide length required to form coacervates with a series of $R_2G_NR_2$ peptides (N = 0, 2, 8). Scale bar = 10 µm, bright-field microscopy.



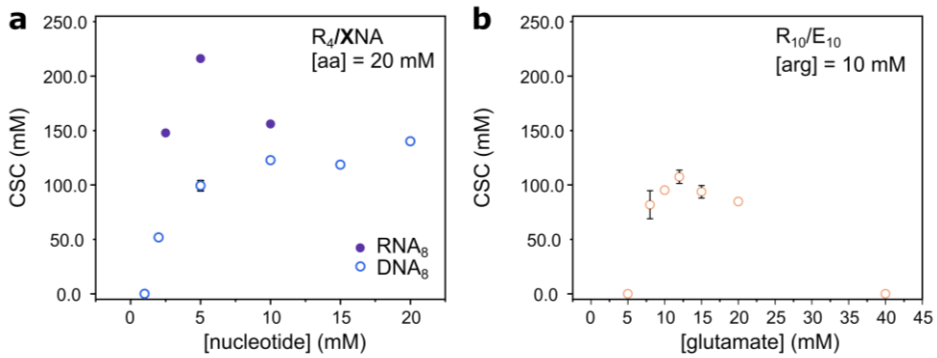
Supplementary Figure 13. (a) Extended data for the amino acid concentration leading to an onset of turbidity and, therefore, coacervation. All curves were recorded titrating a DNA_{16} solution ([nt] = 10 mM) with concentrated stocks of t peptides in the legend. **(b)** Relation between the amino acid concentration at which turbidity sharply increases (onset [amino acid]) and peptide charge density. Onset amino acid concentrations measured are 4.8 mM (R_3), 16.6 mM (RER), 31.4 mM ($R_2E_2R_2$), 43.5 mM (RGR) and 84.4 mM (RG_2R).



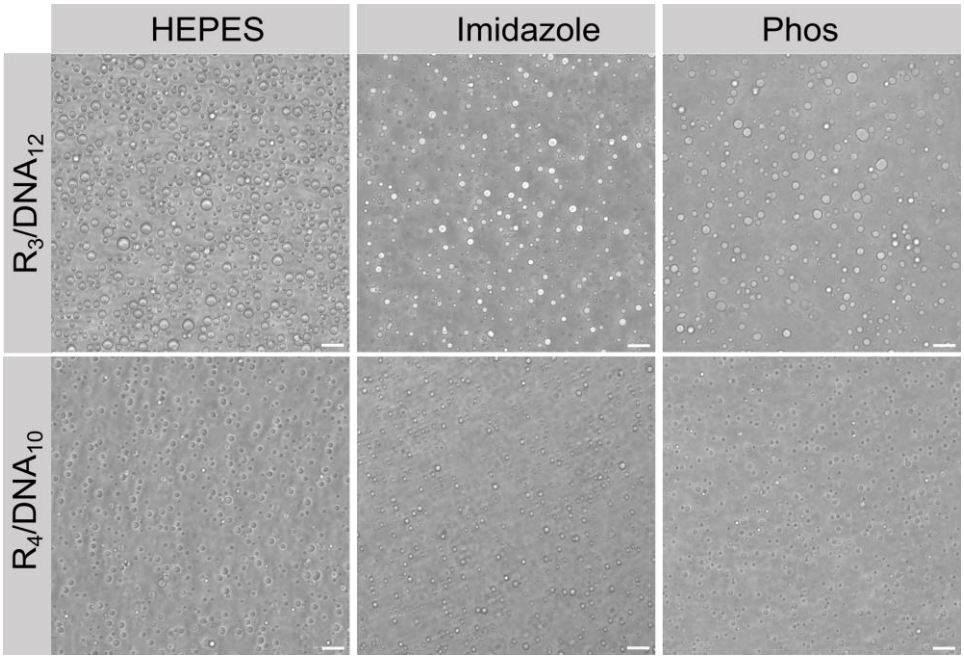
Supplementary Figure 14. (a) Salt titration curves of R_4/DNA_8 mixtures at 5 mM nucleotide. These curves are used to determine **(b)** the phase diagram of the mixture. CSC values were measured at different amino acid concentrations.



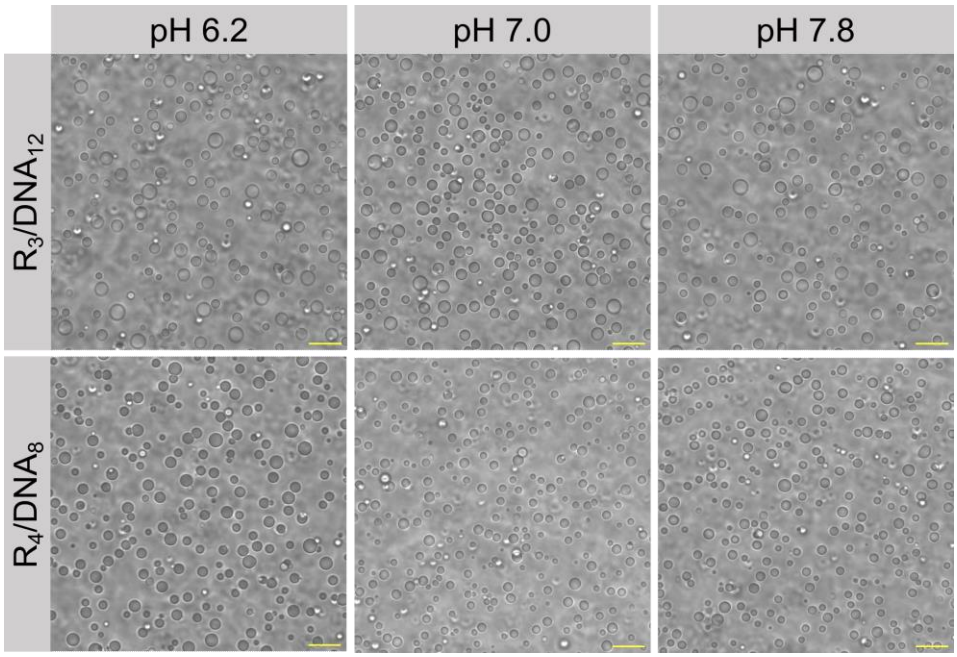
Supplementary Figure 15. (a) Salt titration curves of R_4/DNA_8 mixtures, all at 20 mM amino acid. These curves are used to determine **(b)** the phase diagram of the mixture. The CSC values were measured at different nucleotide concentrations.



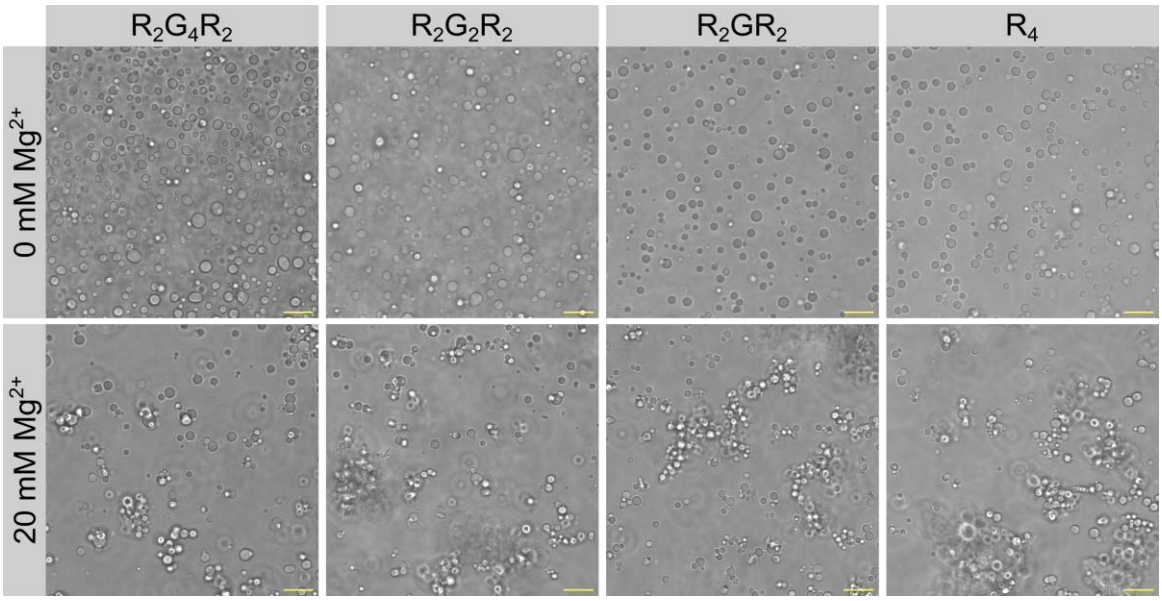
Supplementary Figure 16. Phase diagram of the mixtures in Figure 2 (main text), obtained by varying the concentration of the anionic monomer: nucleotide in the case of peptide/oligonucleotide mixtures **(a)**; glutamic acid in the case of the peptide/peptide coacervates **(b)**.



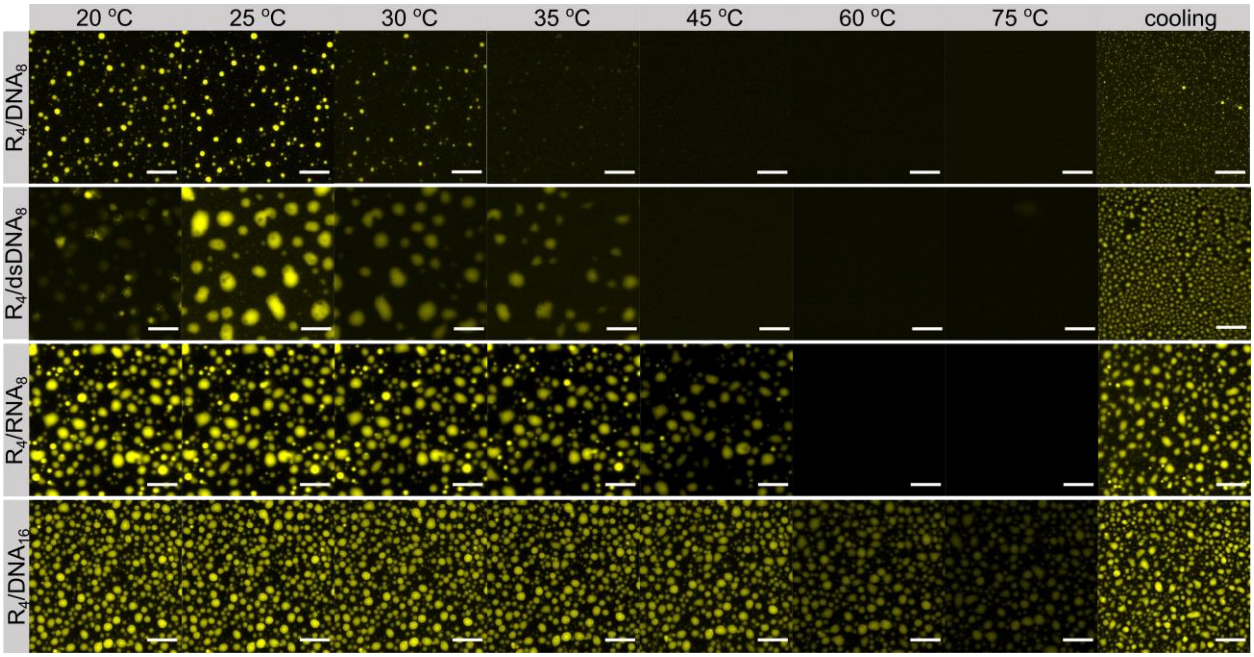
Supplementary Figure 17. Bright-field images of minimal coacervates in 25 mM HEPES pH 7.4, 25 mM imidazole pH 7.5 or 25 mM phosphate buffer pH 7.5. Scale bar = 10 µm.



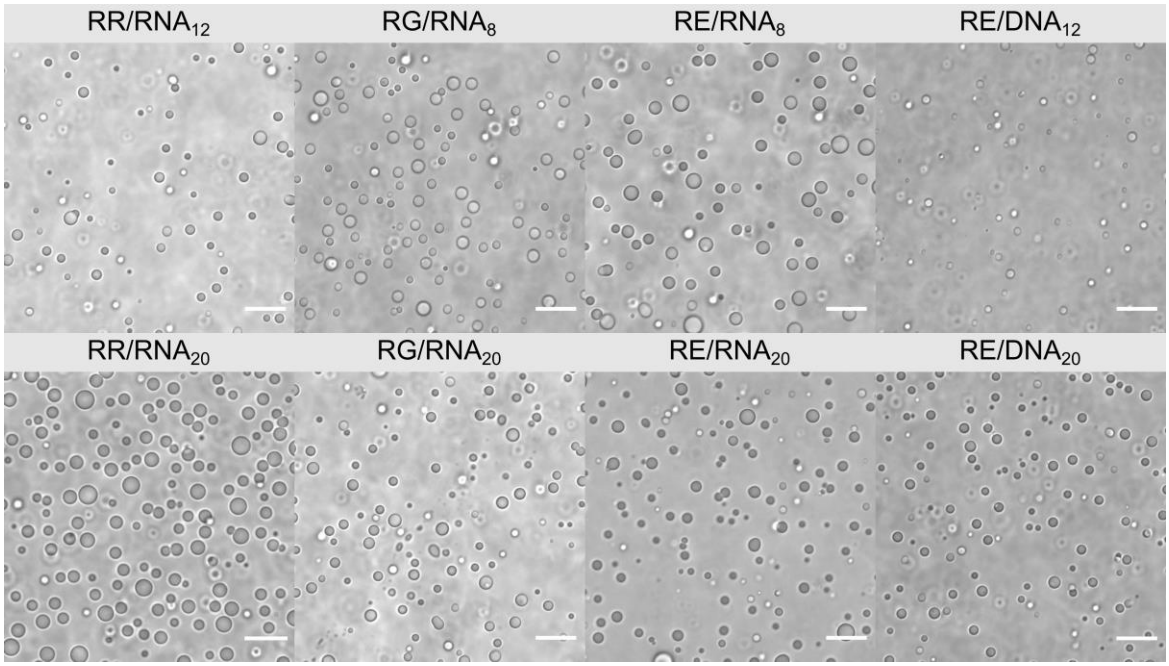
Supplementary Figure 18. Minimal coacervates in 25 mM imidazole buffer at different pH values, observed by bright-field microscopy. Scale bar = 10 µm.



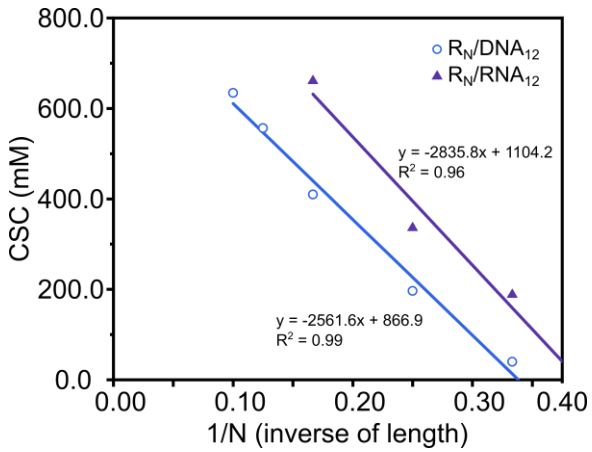
Supplementary Figure 19. $R_2G_NR_2/DNA_{20}$ coacervates at different Mg^{2+} concentrations. Scale bar = 10 μm, bright-field microscopy.



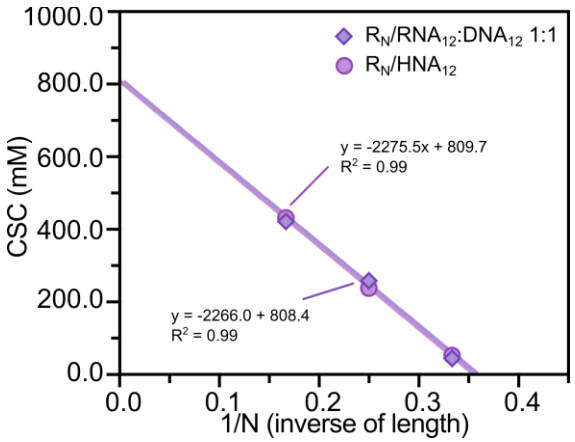
Supplementary Figure 20. Extended dataset for the thermal stability of peptide/oligonucleotide mixtures as in Figure 2 (main text). The R_4/$dsDNA_8$ mixture is added for comparison and shows an additional phase transition before room temperature. R_4/DNA_{16}, despite its similar CSC to R_4/RNA_8, did not fully dissolve in the heating ramp. All mixtures reassemble into droplets upon cooling, and coacervate fluorescence is recovered. Scale bar: 10 µm, fluorophore: Cy3-$(TGAC)_2$.



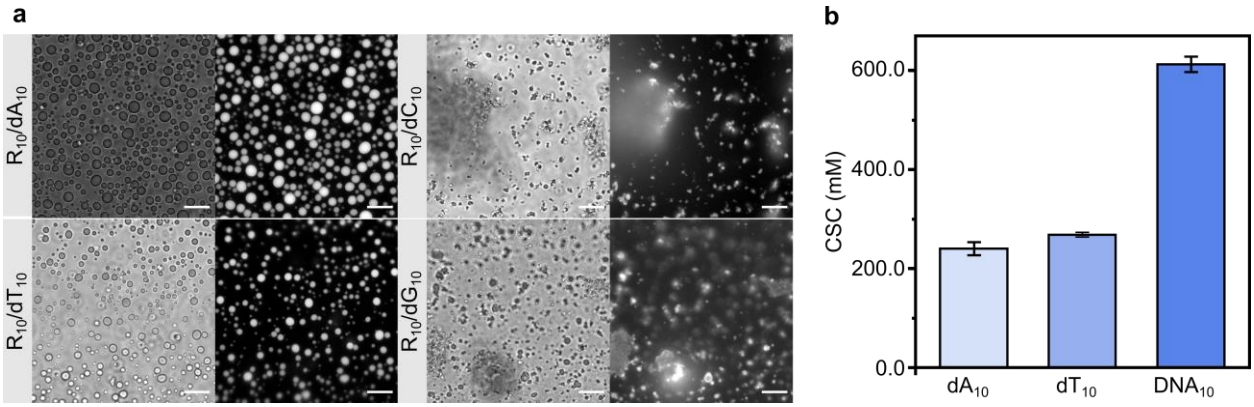
Supplementary Figure 21. Coacervates composed of peptide dimers and oligonucleotides (8-20 nt). Required concentrations of the components are listed in **Supplementary Table 7**. Scale bar = 10 µm, bright-field microscopy.



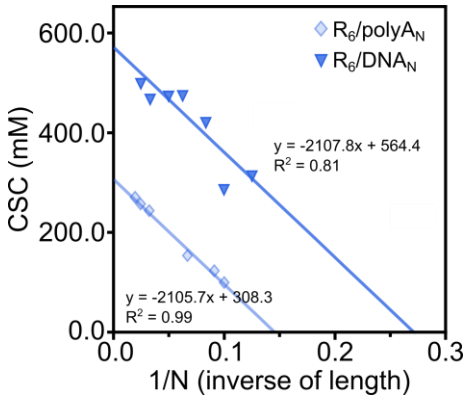
Supplementary Figure 22. CSC dependence on the inverse of the length of different peptides and oligonucleotides enabling the prediction of the shortest peptide to form coacervates with DNA_{12} or RNA_{12} (motif ACTG, open circles; or ACUG, purple triangles). The details of the linear regression are shown to calculate the length for CSC > 0.



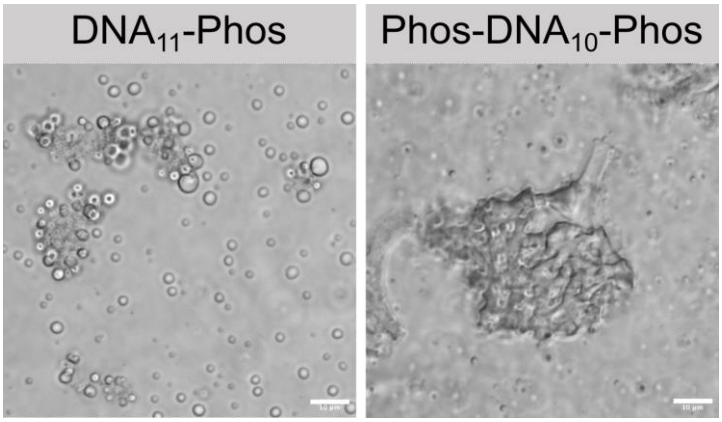
Supplementary Figure 23. CSC dependence on the inverse of the length of different peptides and oligonucleotides enabling the prediction of the shortest peptide to form coacervates with a mixture of DNA_{12} and RNA_{12} (purple diamonds) or with the hybrid strand HNA_{12} (magenta circles). The linear fits overlap, and details are shown in **Supplementary Table 5**. The details of the linear regression are shown to calculate the length for CSC > 0.



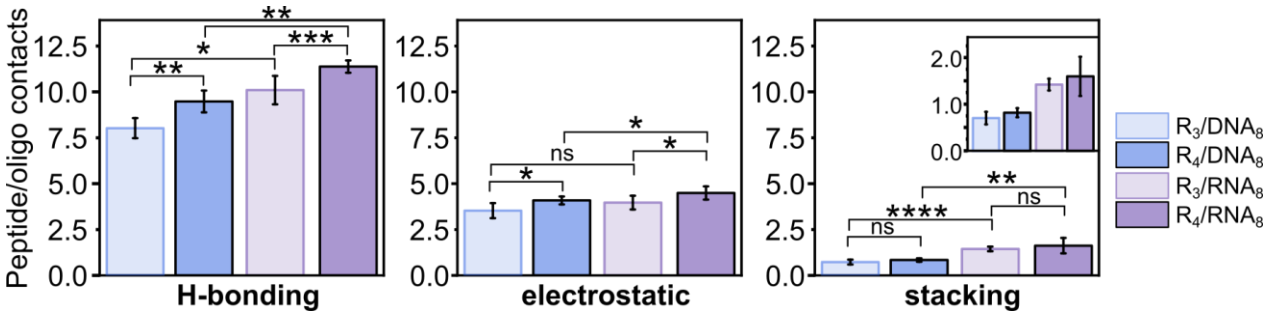
Supplementary Figure 24. (a) Mixtures of R_{10}/dX_{10} imaged by epifluorescence microscopy (probe: Cy3-10nt). Decamers of C and G, the bases capable of three hydrogen bonds, lead to solid aggregation instead of liquid droplets. **(b)** Critical salt concentrations (CSCs) of R_{10}/X_{10} mixtures that form coacervates, measured at 20 mM amino acid and 5 mM nucleotide, without any labelled oligo added. The CSC of solid aggregates is not a defined property.



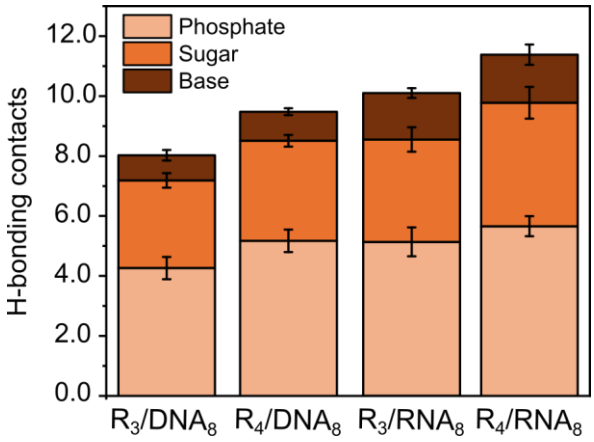
Supplementary Figure 25. CSC dependence on the inverse of the length of different peptides and oligonucleotides enabling the prediction of the shortest homopolymeric DNA (polyA$_N$, diamonds) to form coacervates with peptide R$_6$, in comparison to when a heteropolymeric DNA sequence is used (inverted triangles). The details of the linear regression are shown to calculate the length for CSC > 0.



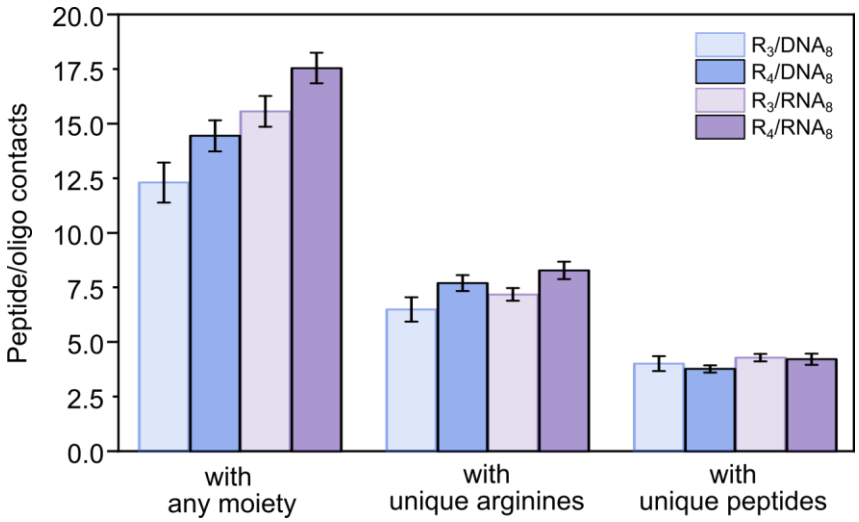
Supplementary Figure 26. R_4-based coacervates made with phosphate-modified oligos. Scale bar = 10 µm, bright-field microscopy.



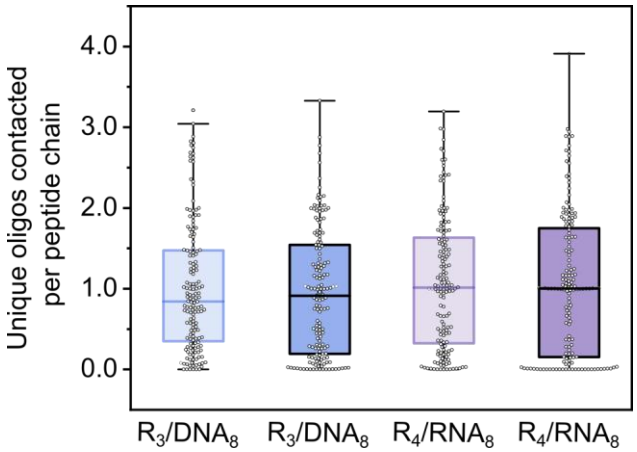
Supplementary Figure 27. Number of contacts per nucleotide through all three interaction modes for the four systems simulated atomistically. Error bars represent the standard deviation across 5 repeats.



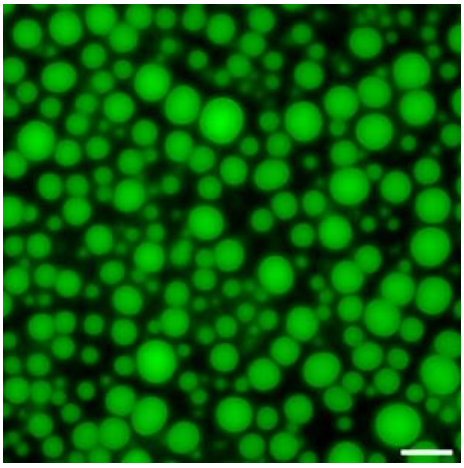
Supplementary Figure 28. Number of contacts *via* H-bonding, per nucleotide, established through the phosphate, sugar and base moieties. Error bars represent the standard deviation across 5 repeats.



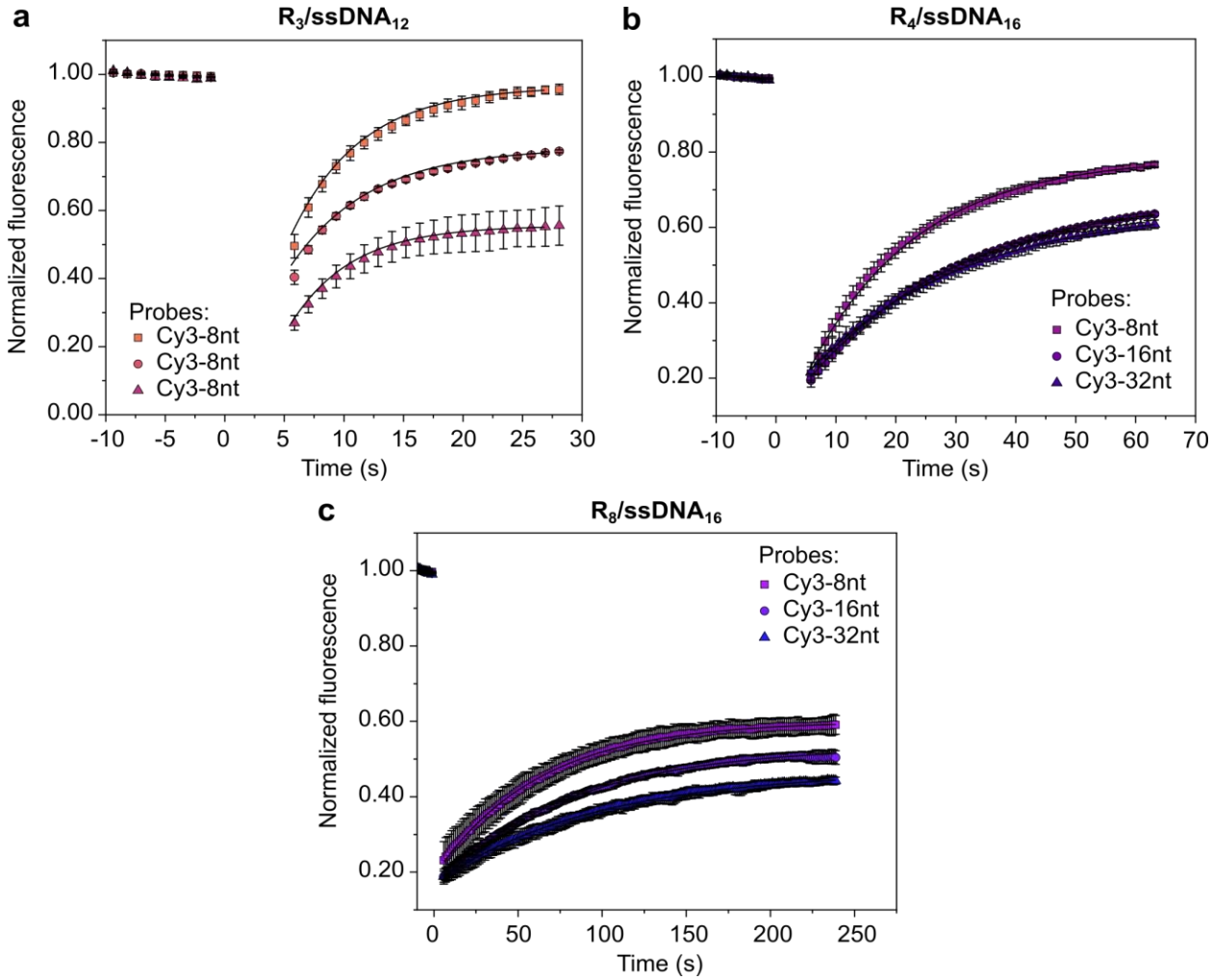
Supplementary Figure 29. Number of contacts established per oligonucleotide, grouped into three types of contact: with any peptide moiety, with unique arginine residues, or with unique peptide chains. Error bars represent the standard deviation across 5 repeats. The number of contacts with any peptide moiety is the "total valency" represented in Figure 3 (main text).



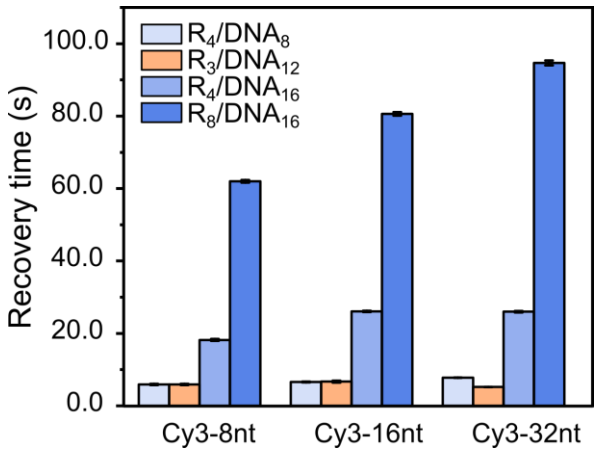
Supplementary Figure 30. Unique oligonucleotide chains contacted by a peptide chain for the different mixtures simulated atomistically. The mixture R_3/DNA_8 is the only one with a median <1, *i.e.*, on average, the peptide R_3 is free in the presence of DNA_8 chains. In this case, we represent each peptide chain sampled (36 chains, 5 repeats) instead of an average. Boxplots contain 50% of the data points measured for the 36 peptide chains in the simulation, averaged over time per simulation repeat (5x). The horizontal line represents the median.



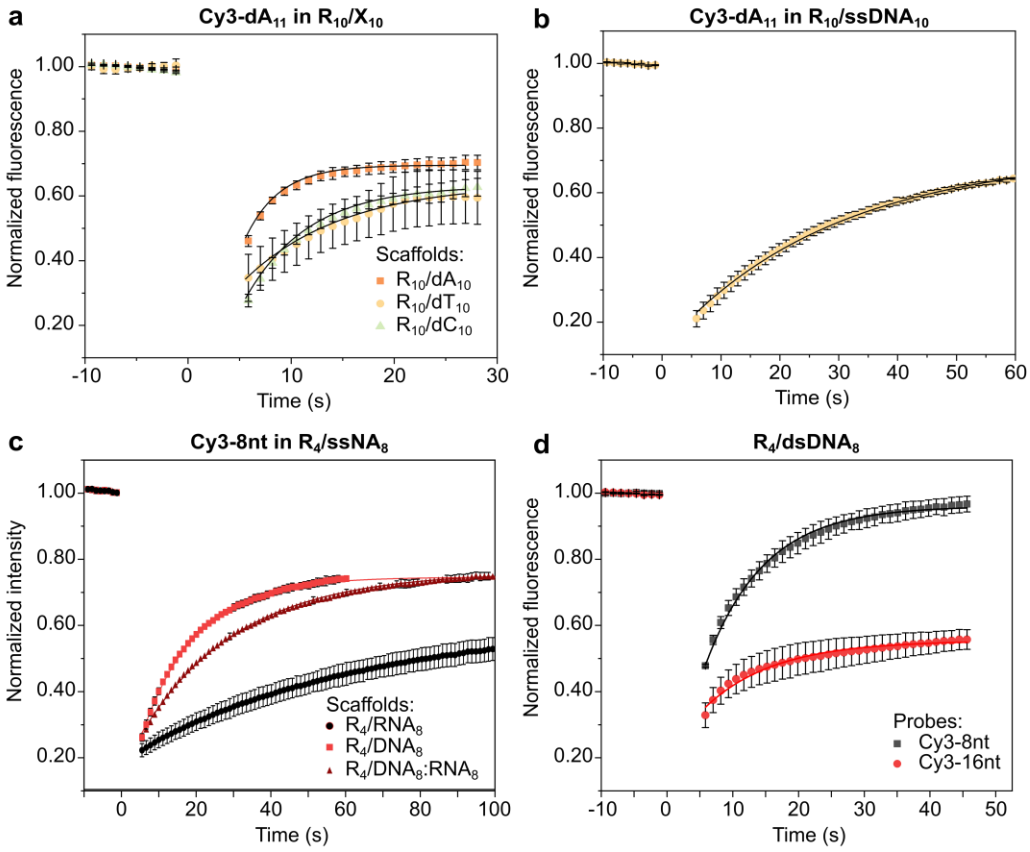
Supplementary Figure 31. Cy_3-DNA_5 in R_4/DNA_{20} coacervates, showing that a DNA pentamer is recruited in the droplets. Scale bar = 10 µm, epifluorescence microscopy.



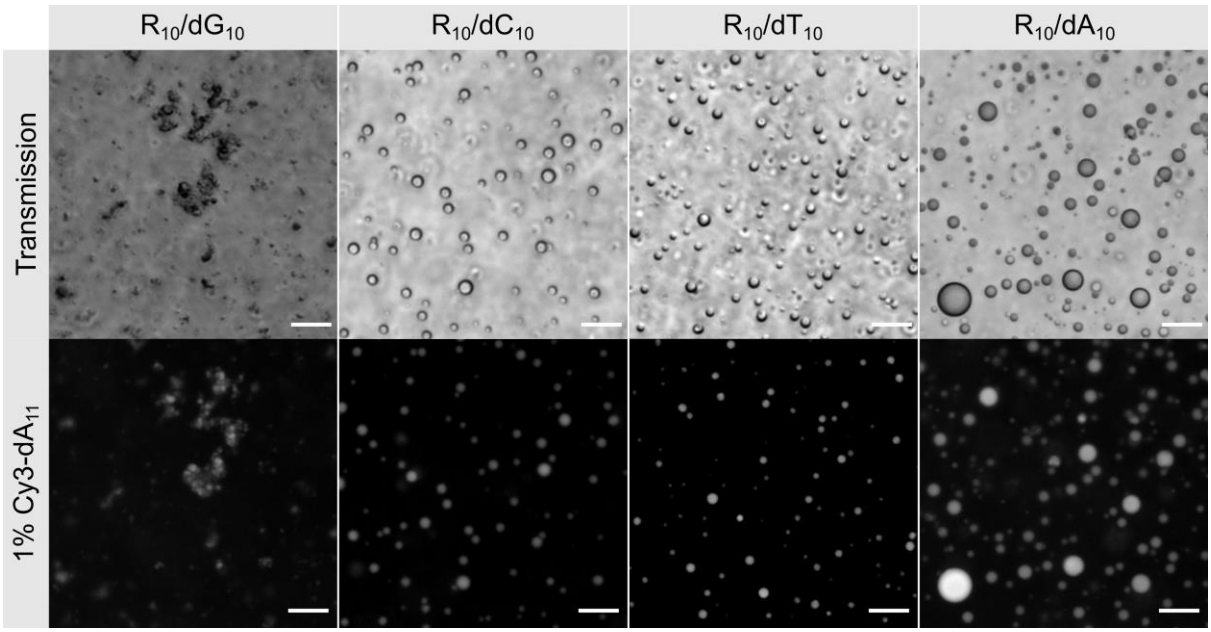
Supplementary Figure 32. FRAP profiles of a series of coacervates where peptide and DNA length were varied. Non-complementary probes of three lengths were tested for each coacervate system: 8, 16 and 32nt.



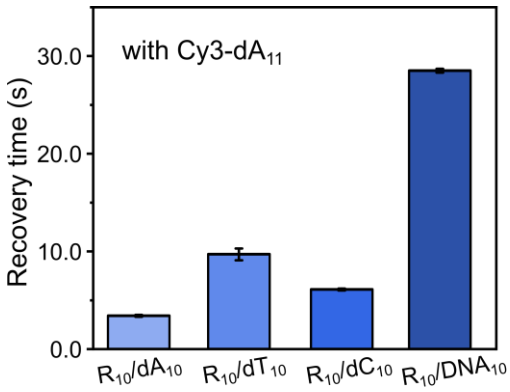
Supplementary Figure 33. Expanded dataset of FRAP recovery times for peptide-, DNA- and probe-length series as shown in Figure 4 (main text).



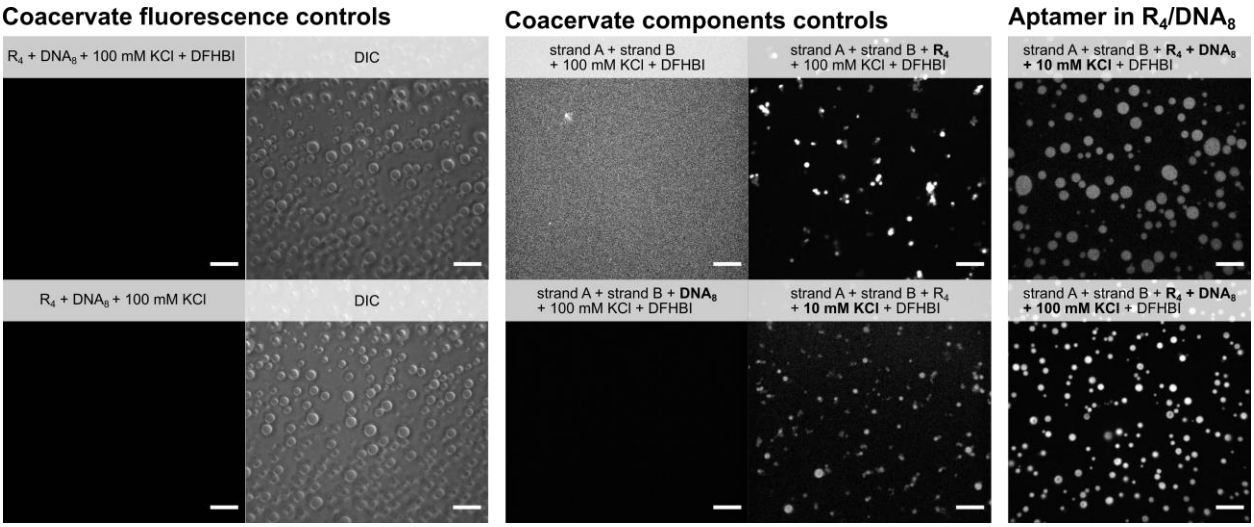
Supplementary Figure 34. FRAP profiles of coacervates showcasing the effect of interactions between probes and scaffold. (a) and (b) refer to the effect of homopolymers and heteropolymers of DNA. (c) and (d) probe the effect of DNA, RNA and dsDNA as scaffold strands.



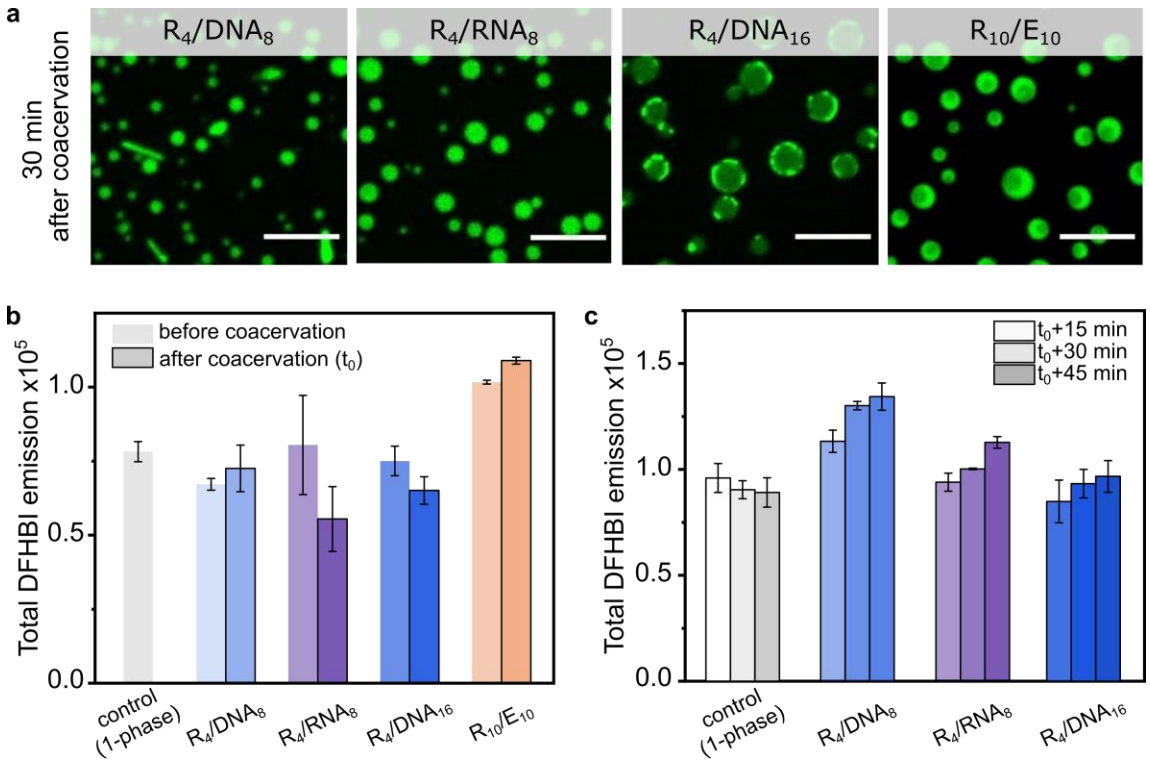
Supplementary Figure 35. Transmission and epifluorescence micrographs of R_{10}/X_{10} coacervates labelled with 1% Cy3-dA_{11}, as used in FRAP measurements. Scale bars are 10 μm.



Supplementary Figure 36. Expanded graph of FRAP recovery times for Cy3-dA_{11} in R_{10}/X_{10} coacervates. FRAP in R_{10}/dG_{10} was not measured due to aggregation under all conditions.



Supplementary Figure 37. Confocal micrographs of aptamer reconstitution in coacervates. (a) Confocal fluorescence micrographs of different controls and test samples under constant irradiation conditions in the 488 nm channel (DFHBI). The DIC channel is only shown for (1) and (2) to confirm the presence of coacervates. (b) Quantification of the total DFHBI emission in the field of view shown in (a). Error bars come from measurements in triplicate. (c) DFHBI emission in the presence of coacervates, now separating emission from all droplets in the FOV and the background (dilute phase). Samples were prepared adding DFHBI as the last component.



Supplementary Figure 38. DFHBI/Broccoli aptamer emission measured in a plate reader (bulk), in contrast with the data from **Figure S11**, measured with confocal microscopy detection. (a) Confocal micrographs of coacervates containing the DFHBI/Broccoli aptamer complex imaged ca. 30 min after preparation. Scale bars are 10 µm. (b) Expanded dataset shown in Figure 5 (main text), including the effect of R_4/DNA_{16} and R_{10}/E_{10} coacervates on emission, measured before and immediately after coacervation. Samples were prepared by adding the peptide as the last component, *i.e.*, coacervation occurs in the presence of the reconstituted aptamer. (c) Time course of DFHBI emission (15, 30 and 45 min after mixing). In this setup, the different partitioning of the aptamer among the coacervates tested is evident.

References

1. Jensen, K. J.; T. S., Pernille; Pedersen, Søren L. *Peptide Synthesis and Applications*. 252 (Humana Press, Totowa, NJ, 2013).
2. Robinson, J. D., Sammons, S. R. & O'Flaherty, D. K. Preparation of 2-Aminoimidazole-Activated Substrates for the Study of Nonenzymatic Genome Replication. *Current Protocols* **4**, e1119 (2024).
3. Nakashima, K. K., André, A. A. M. & Spruijt, E. Chapter Thirteen - Enzymatic control over coacervation. in *Methods in Enzymology* (ed. Keating, C. D.) vol. 646 353–389 (Academic Press, 2021).
4. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
5. Zgarbová, M. *et al.* Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **7**, 2886–2902 (2011).
6. Ivani, I. *et al.* Parmbsc1: a refined force field for DNA simulations. *Nat Methods* **13**, 55–58 (2016).
7. Wang, L.-P., Martinez, T. J. & Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **5**, 1885–1891 (2014).
8. Åqvist, J., Wennerström, P., Nervall, M., Bjelic, S. & Brandsdal, B. O. Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chemical Physics Letters* **384**, 288–294 (2004).
9. Zhang, Z., Liu, X., Yan, K., Tuckerman, M. E. & Liu, J. Unified Efficient Thermostat Scheme for the Canonical Ensemble with Holonomic or Isokinetic Constraints via Molecular Dynamics. *J. Phys. Chem. A* **123**, 6056–6079 (2019).
10. Eastman, P. & Pande, V. S. Constant Constraint Matrix Approximation: A Robust, Parallelizable Constraint Method for Molecular Simulations. *J. Chem. Theory Comput.* **6**, 434–437 (2010).
11. Essmann, U. *et al.* A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577–8593 (1995).
12. Eastman, P. *et al.* OpenMM 8: Molecular Dynamics Simulation with Machine Learning Potentials. *J. Phys. Chem. B* **128**, 109–116 (2024).
13. Liu, D. C. & Nocedal, J. On the limited memory BFGS method for large scale optimisation. *Mathematical Programming* **45**, 503–528 (1989).
14. PyMOL | pymol.org. https://pymol.org/.
15. Martínez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* **30**, 2157–2164 (2009).
16. McGibbon, R. T. *et al.* MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal* **109**, 1528–1532 (2015).
17. Wernet, Ph. *et al.* The Structure of the First Coordination Shell in Liquid Water. *Science* **304**, 995–999 (2004).
18. Paloni, M., Bussi, G. & Barducci, A. Arginine multivalency stabilises protein/RNA condensates. *Protein Science* **30**, 1418–1426 (2021).