Co-transcriptional production of programmable RNA condensates and synthetic organelles

Giacomo Fabrini[1,2,3], Sabrina Pia Nuccio[1], Jaimie M. Stewart[4], Shiyi Li[5], Anli Tang[6], Paul W. K. Rothemund[4], Elisa Franco[5,6], Marco Di Antonio[1], Lorenzo Di Michele[3,2,1*]

[1]Department of Chemistry, Imperial College London, 82 Wood Lane, London, W12 0BZ, UK.
[2]fabriCELL, Imperial College London, 82 Wood Lane, London, W12 0BZ, UK.
[3*]Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge, CB3 0AS, UK.
[4]Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA.
[5]Department of Bioengineering, University of California at Los Angeles, Los Angeles, CA 90024, USA.
[6]Department of Mechanical and Aerospace Engineering, University of California at Los Angeles, Los Angeles, CA 90024, USA.

*Corresponding author(s). E-mail(s): ld389@cam.ac.uk;

Abstract

Condensation of RNA and proteins is central to cellular functions, and the ability to program it would be valuable in synthetic biology and synthetic cell science. Here we introduce a modular platform for engineering synthetic RNA condensates from tailor-made, branched RNA nanostructures that fold and assemble co-transcriptionally. Up to three orthogonal condensates can form simultaneously and selectively accumulate guest molecules. The RNA condensates can be expressed within synthetic cells to produce membrane-less organelles with controlled number, size, morphology and compositions, and that display the ability to selectively capture proteins. The *in situ* expression of programmable RNA condensates could underpin spatial organisation of functionalities in both biological and synthetic cells.

Keywords: Condensates, RNA, Nanotechnology, LLPS, Organelles, Synthetic Cells

1 Introduction

Membrane-less compartmentalisation emerges from the condensation of proteins and RNA, and is recognised as a primary mechanism through which cells dynamically control biochemical processes [1–4]. By co-localising nucleic acids, enzymes and metabolites, membrane-less organelles (MLOs) such as nucleoli, Cajal bodies and stress granules are believed to regulate biogenesis, transcription, post-transcriptional modification and degradation, and generally improve cellular fitness [4–7]. The emergence of biomolecular condensates has also been linked to disease, particularly to neurodegeneration [8].

The ability to express "designer condensates" with prescribed properties would be a valuable tool to program cellular behaviour [9, 10] and engineer synthetic cells [11]. Remarkable examples based on engineered peptides [9, 10] or natural condensate-forming RNA sequences [12, 13], have highlighted the feasibility of this concept. The generality of these strategies, however, is hampered by challenges linked to protein engineering and the limited programmability of natural RNA repeats.

Leveraging the algorithmic toolkit of nucleic acid nanotechnology [14–16], here we introduce a systematic method for expressing designer biomolecular condensates from synthetic RNA nanostructures. Our elementary motifs consist of star-shaped junctions, or nanostars, which fold co-transcriptionally and assemble thanks to programmable base-pairing interactions, forming liquid or gel-like condensates. Thanks to the selectivity of base pairing, up to three co-existing but fully distinct condensate types can be produced. Expressing the condensates in model synthetic cells generates MLOs with controlled size, number, morphology and composition. Finally, including RNA aptamers enables selective capture of small molecules and proteins, imitating the ability of natural MLOs to recruit clients.

Because the RNA nanostars are transcribed *in situ* from DNA templates, rather than assembling from pre-synthesised strands, our platform could be directly applied to engineering living cells, besides its immediate use for creating functional MLOs in synthetic cells. In this context, exploring the vast design space of RNA nanostars will allow for fine tuning of condensate properties and self-assembly behaviour, as shown by Stewart *et al.* [17].

2 Results

The four-armed RNA nanostars, shown in Fig. 1a, consist of a single RNA strand that folds co-transcriptionally into the intended shape. The star-shape was inspired by well characterised DNA nanostars [18–20]. However the RNA motifs interact *via* self-complementary (palindromic) HIV-type Kissing-Loops (KLs) present at the end of each arm [21], rather than *via* the single-stranded (ss) overhangs or hydrophobic modifications adopted for DNA designs [18–20]. We tested three RNA nanostar designs, labelled A, B and C, featuring mutually orthogonal KL sequences (see

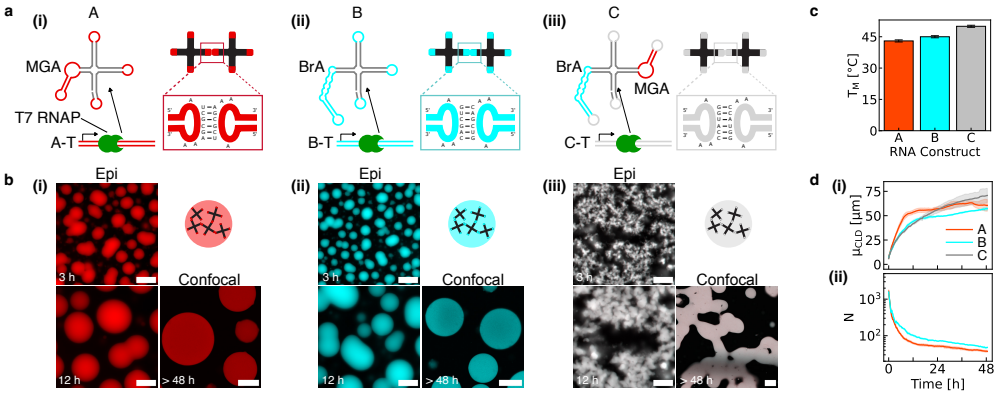

Fig. 1 Condensation of co-transcriptionally folding RNA nanostars. (a) Structure of the RNA motifs. (i) A-type RNA nanostars include a Malachite Green Aptamer (MGA), (ii) B-type include a Broccoli Aptamer (BrA), and (iii) C-type include both aptamers. Variants feature mutually-orthogonal, self-complementary (palindromic) Kissing Loops (KLs), whose sequences are shown in the insets. RNA nanostars are transcribed from linear dsDNA templates by T7 RNA polymerase (T7 RNAP). (b) Epifluorescence and confocal micrographs showing condensate formation and coarsening for all three designs in (a) at different timepoints of an *in vitro* transcription reaction. Epifluorescence micrographs have been linearly re-scaled to enhance contrast (see SI Methods, section 1.4). Pristine micrographs are shown in Fig. S3. Scale bars are 50 μm. (c) Melting temperatures of A-C condensates, determined as discussed in the SI Methods and Figs S5 and S6. Errorbars ($\pm$ 0.5 °C) stem from the discrete 1 °C step between consecutive measurements. (d) (i) Mean of the chord-length distribution, μ_{CLD}, and (ii) average number of condensates per microscopy field-of-view, N, as a function of time. Full chord-length distributions, as extracted from image segmentation, are shown in Fig. S16 (see SI Methods, section 1.4). N is not computed for system C, which does not form discrete aggregates. Data are shown as mean (solid lines) $\pm$ standard deviation (shaded regions) of three field-of-views within one sample.

zoomed-in boxes in Fig. 1a). In designs A and B, one of the dsRNA arms includes a Fluorescent Light up Aptamer (FLAP): Malachite Green Aptamer (MGA) for A [22, 23] and Broccoli Aptamer (BrA) for B [24] (Figs 1a (i) and (ii)). FLAPs yield a fluorescent signal upon binding their cognate fluorophores (malachite green for MGA and DFHBI for BrA), enabling characterisation *via* fluorescence microscopy and fluorimetry. In design C, both MGA and BrA were included in non-adjacent arms (Fig. 1a (iii)). The fluorescence characteristics of all RNA constructs are reported in Fig. S1. The arms that do not host FLAPs were designed to be 25 base-pairs long, and separated by an unbound uracil residue at the junction to improve flexibility [25]. The motifs were designed to be transcribed with T7 RNA Polymerase (T7 RNAP) from fully double-stranded DNA (dsDNA) templates, labelled as A-T, B-T and C-T for designs A, B and C, respectively (see Fig. S2). Details on sequence design and *in vitro* transcription protocols are reported in SI Methods, sections 1.1-1.3. All sequences are reported in Tables S1-S3.

As shown in Fig. 1b with representative epifluorescence and confocal snapshots, all three designs formed aggregates during *in vitro* transcription experiments. Additional examples are provided in Fig. S3, Fig. S4 (top) and video S1 (top). Variants A and

B formed spherical condensates that nucleated and grew, with frequent coalescence events, confirming that these designs yield liquid-like condensates. Conversely, design C formed a gel-like percolating structure which failed to produce discrete condensates but still grew over time. The higher apparent viscosity of design C when compared to A and B correlates with the thermal stability of the materials. Indeed, C exhibited the highest melting temperature, followed by B and A (Figs 1c, S5 and S6, and video S2). Regardless of their morphology, all aggregates displayed the intended fluorescence output, namely in the MGA channel for A (red), BrA channel for B (cyan) and both channels for C (white).

Degradation of the condensates resulting in bubble formation was observed over time, particularly for A-type aggregates (Fig. S3). This is ascribed to the action of environmental nucleases [26] or photodegradation.

The specificity of KL interactions was confirmed by non-sticky, control designs where KLs were replaced with scrambled sequences ($\bar{\text{A}}$, $\bar{\text{B}}$). These designs did not yield condensates and only produced diffused fluorescence, as shown in Figs S4 (bottom) and S7, and video S1 (bottom).

Besides microscopy visualisation, aptamer fluorescence was also used to monitor the rate of synthesis of RNA building blocks with fluorimetry, as reported in Fig. S8. All designs displayed a rapid initial increase, whose rate scaled (nearly) linearly with template concentration (Figs S9 and S10). After the initial phase of rapid RNA production, synthesis slowed down significantly or plateaued depending on the specific design, due to loss of polymerase activity and/or nucleotide depletion [27–29]. A peak was noted for condensate-forming motifs (A, B and C), which was not present in non-sticky designs ($\bar{\text{A}}$ and $\bar{\text{B}}$), ascribed to sedimentation of the fluorescent aggregates below the measurement plane. Differences in plateauing behaviours could be due to variations in the kinetics of aptamer folding and/or complexation with the respective fluorophores [30–32].

We quantified the aggregate growth dynamics for all designs by measuring the temporal evolution of the mean chord length μ_{CLD}, as extracted from segmentation of epifluorescence timelapses (see SI Methods, section 1.4). μ_{CLD} provides a measure of the typical lengthscale of the fluorescent features, and was chosen because it is applicable and meaningful regardless of aggregate shape [33–35]. As shown in Fig. 1d (i), aggregate size increased rapidly at early stages for all designs, before plateauing for A and B. Condensate size plateauing occurred roughly when fluorimetry traces also saturated (Fig. S8). The correspondence of these timescales hints at a scenario in which condensate growth slows down or stops when monomer transcription ceases, and coalescence of large condensates becomes rare, owing to the lack of substantial Brownian motion (video S1). For design C, the increase in typical aggregate lengthscale continued at later times, driven by coarsening of the percolating RNA network. Full chord-length distributions (CLDs) are shown in Fig. S11.

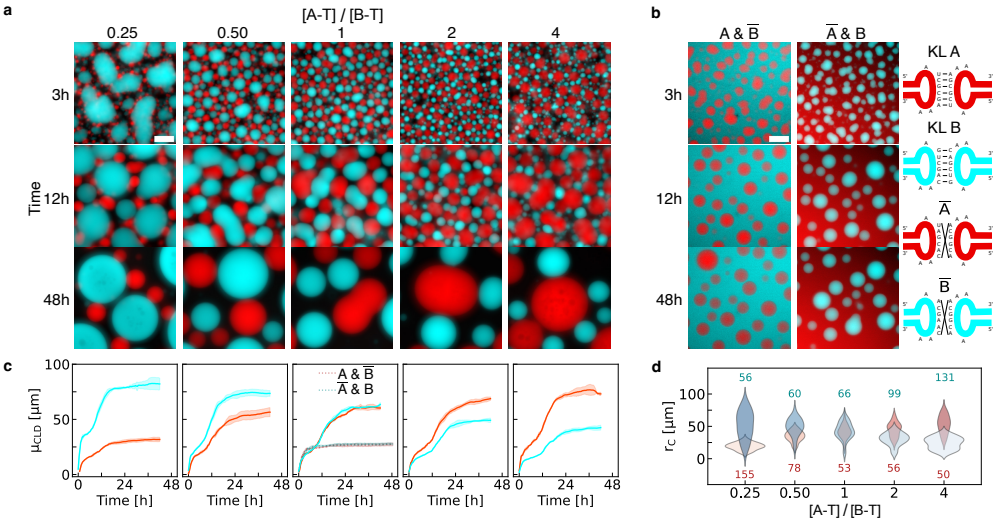

Fig. 2 Co-transcribed orthogonal RNA nanostars form immiscible condensates of controlled size. (a) Epifluorescence micrographs of binary systems of A and B RNA nanostars (see Fig. 1) at various timepoints during the transcription transient. Different ratios between the concentrations of the two templates, A-T and B-T, are tested, while keeping [A-T]+[B-T] constant. (b) Epifluorescence snapshots analogous to panel (a), but where either A or B is replaced by its non-sticky variant, namely $\bar{\mathrm{A}}$ or $\bar{\mathrm{B}}$. Sketches on the right hand side show examples of scrambled non-binding KL sequences. (c) Time evolution of the mean chord-length, μ_{CLD}, for samples in (a) and (b), the latter shown in the central panel as dashed lines. See Fig. S16 for full chord-length distributions. Red and cyan curves are relative to A and B condensates, respectively. Data are shown as mean (solid/dashed lines) $\pm$ standard deviation (shaded regions) of three field-of-views within one sample. (d) Distribution of the radii, r_{c}, of A (red) and B (cyan) condensates as a function of the template concentration ratio [A-T]/[B-T]. Epifluorescence micrographs in (a) and (b) have been linearly re-scaled to enhance contrast (see SI Methods, section 1.4). Pristine micrographs are shown in Fig. S12 and S14. All scale bars are 50 μm.

Trends consistent to those noted for μ_{CLD} were found for the average number, N, of A- or B-type condensates counted in each microscopy field-of-view. Data summarised in Fig. 1d (ii) show that N monotonically decreased following condensate coalescence. The number of aggregates was not studied for design C given its tendency to form a single percolating structure.

KL orthogonality enables the simultaneous transcription and assembly of designs A and B, that readily formed distinct, co-existing condensates, as shown in Fig. 2a (see also Figs S12, S13, and video S3). Consistently, if one of the RNA motifs was rendered non-sticky, condensates of one species co-existed with dispersed RNA nanostars of the other, as shown in Figs 2b, S14, S15, and video S4.

A simple strategy to control the relative size of A and B condensates is to tune the ratio between the concentrations of the corresponding DNA templates ([A-T] and [B-T], respectively), which determines the production rate of each nanostar, as demonstrated in Figs S9 and S10. Relative-size control is demonstrated visually in

Fig. 2a (see also Figs S12 and S13), through time-dependent μ_{CLD} analysis in Fig. 2c (see Fig. S16 for full CLDs), and through the distribution of the final radii r_c of the condensates, shown in Fig. 2d (see SI Methods, section 1.4). We observed condensate number to anti-correlate with their size. For instance, many small A-type condensates were formed when [B-T]>[A-T] (see Fig. S17).

As shown in Fig. 2c, condensates were found to grow in two stages in all binary systems of sticky nanostars: after an initial increase, a brief intermediate plateau was reached, followed by another growth phase before saturation. This behaviour was not observed in single component systems (Fig. 1d), nor in binary systems where one nanostar population was rendered non-sticky, for which the temporal evolution of μ_{CLD} is included in the central panel of Fig. 2c. Inspection of the relevant microscopy videos (videos S3 and S4) reveals that the intermediate size plateau was reached when condensates became temporarily hindered in their ability to make contact with same-type assemblies, due to being "caged" by nearest-neighbours condensates of the opposite type. Coalescence events that still managed to occur, however, reduced lateral crowding given that merged condensates occupy less space in the horizontal plane, triggering a positive-feedback process that rapidly unjammed the system and accelerated coalescence. In fact, Fig. 2c shows that condensates reached a larger overall size in systems where both A and B were sticky, compared to $\bar{\text{A}}$ & B and A & $\bar{\text{B}}$ mixtures (Figs 2b, S14 and S15), indicating that steric encumbrance from non-binding condensates of the opposite phase ultimately facilitates coalescence.

Two-step decreasing trends, consistent to those seen in μ_{CLD}, were observed in the number of condensates *per* field-of-view (Fig. S17).

Condensate co-assembly is also possible with all three RNA species A, B, and C, as shown in Figs S18, S19 and video S5. After allowing sufficient time for relaxation, all three species formed distinct spherical condensates, including C, which was unable to do so in single-component samples (Fig. 1b (iii)). The difference in morphology is likely due to A and B stars hindering the formation of a percolating C network in favour of smaller aggregates that relax more readily.

Addressable self-assembled RNA condensates could be extremely valuable to engineer compartmentalisation in synthetic or living cells, where they could operate as membrane-less organelles capable of recruiting target compounds and underpinning spatial separation of functionalities. To demonstrate this, we transcribed our condensates within basic synthetic cells constructed from water-in-oil emulsion droplets (Fig. 3a). All designs formed condensates that equilibrated into a single spherical organelle in each synthetic cell (Figs 3b and S20), including design C that generated only extended networks when produced in bulk. The different morphology is readily rationalised by noting that, when confined to the synthetic cell, C aggregates only need to relax over much smaller lengthscales. Yet, shape relaxation was slower for C compared to A and B, as confirmed by the time evolution of μ_{CLD} in Fig. 3c (see also micrographs in Fig. S21 (top) and full CLDs in Fig. S22) and video S6 (top).

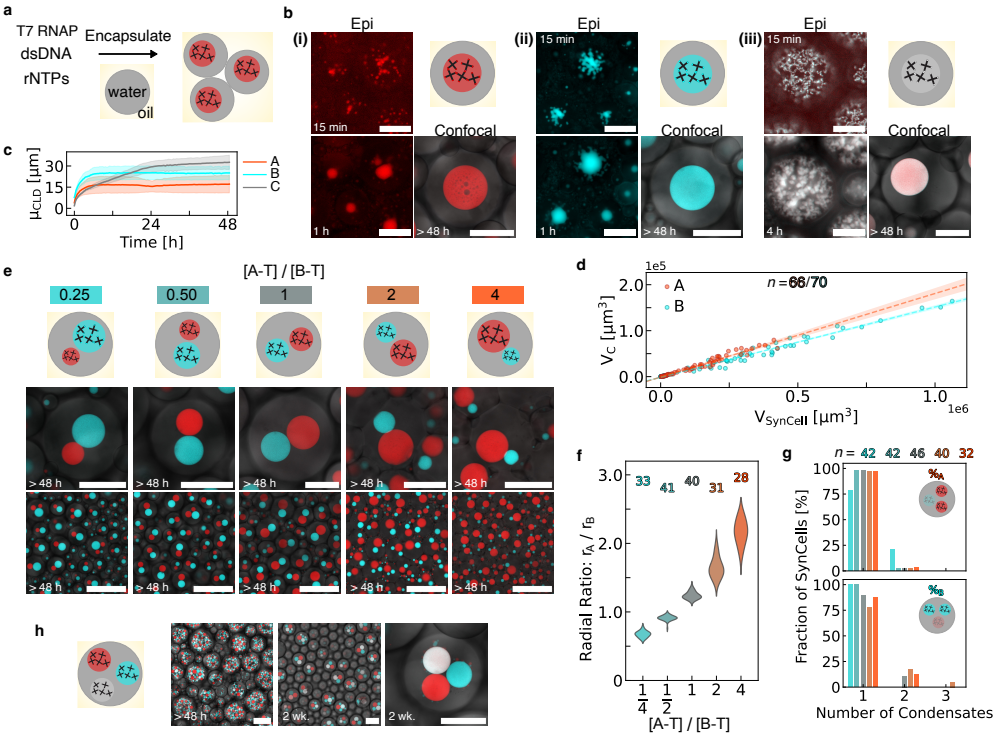

Fig. 3 Membrane-less RNA organelles expressed in synthetic cells. (a) Diagram showing membrane-less organelles (MLOs) formed in synthetic cells consisting of water-in-oil emulsion droplets encapsulating transcription machinery and DNA templates. (b) Epifluorescence and confocal micrographs showing MLO formation over time in synthetic cells expressing (i) A-type, (ii) B-type, and (iii) C-type RNA nanostars (see Fig. 1). Epifluorescence micrographs have been linearly re-scaled to enhance contrast (see SI Methods, section 1.4). Pristine images are shown in Fig. S20, alongside images relative to additional timepoints. (c) Time-dependent mean of the chord-length distribution, μ_{CLD}, computed as discussed in the SI Methods (section 1.4). Data are shown as mean (solid line) $\pm$ standard deviation (shaded region) from three fields-of-view within one sample. (d) Scatter plot of condensate volume (V_{C}) *vs* synthetic cell volume (V_{SynCell}) for samples in (b)-(i) and (b)-(ii). Dashed lines indicate best fits to linear regression models, with 95% confidence intervals shown as shaded regions (SI Methods, section 1.4). MLOs occupy $18.2 \pm 0.5\%$ and $15.3 \pm 0.3\%$ of the volume of the synthetic cells for A and B systems, respectively. (e) Zoomed-in (top) and larger field-of-view (bottom) confocal micrographs depicting synthetic cells co-expressing A and B-type condensates, with different template concentration ratios [A-T]/[B-T] (compare Fig 2a). (f) Distribution of the ratio between the radii of A and B MLOs ($r_{\mathrm{A}}/r_{\mathrm{B}}$) as a function of [A-T]/[B-T] for samples in (e) (see SI Methods). (g) Percentage of synthetic cells containing a given number of A-type (top) or B-type (bottom) MLOs. The percentages of synthetic cells containing exactly one A and one B MLOs are 78.57%, 97.62%, 86.96%, 77.50%, 87.50% for [A-T]/[B-T] = 0.25, 0.50, 1, 2, 4, respectively. Colour-codes in panels (f) and (g) match those in (e). Numbers in (f) and (g) indicate sampled synthetic cells. (h) Confocal micrographs showing synthetic cells expressing three orthogonal MLO-forming RNA nanostars (A, B and C in Fig. 1) at different timepoints. Scale bars in panel (e), bottom, and panel (h), left and centre, are 150 μm. All other scale bars are 50 μm.

Larger variation in condensate size reflects the variability in the size of droplet-based synthetic cells produced *via* emulsion methods. We indeed observed that the final

volume of the RNA condensates scales linearly with the volume of the surrounding emulsion droplet, as presented in Fig. 3d.

Control experiments with non-sticky designs revealed uniform fluorescence within the synthetic cells, confirming that the assembly specificity previously noted in bulk samples persists in confinement (Figs S21 (bottom), S23, and video S6 (bottom)). Like in bulk experiments, fluorimetry can be used to monitor RNA synthesis rates, as shown in Fig. S24, where the delayed growth in the MGA signal (A component) is due to initial accumulation of the malachite green fluorophore in the oil phase [36, 37], rather than to a slower growth of the condensates (compare with Fig. 3c). Also consistent with trends recorded in the bulk, initial transcription rates as determined *via* fluorimetry were found to scale (nearly) linearly with template concentration (Figs S25 and S26).

We obtained synthetic cells with two distinct synthetic organelles by encapsulating both A-T and B-T templates, as shown in Figs 3e, S27-S31, and videos S7-S8 with microscopy, and Fig. S32 with fluorimetry. Like in the bulk experiments, we could control the relative size of the organelles by changing the template ratio, as quantified in Fig. 3f for synthetic cells containing exactly one A and one B condensate. In the vast majority of cases, in fact, each synthetic cell contained one condensate of each type (Fig. 3g). When including component C, we obtained 3 distinct phases, as demonstrated by microscopy in Fig. 3h, (see also Fig. S33, video S9) and fluorimetry (Fig. S35), exemplifying the possibility for scaling up the number of addressable organelles. Consistent with bulk experiments (Figs S18 and S19), relaxation timescales were significantly slower for the ternary system compared to binary ones (Figs 3e and S33). It is also evident that the synthetic cells often failed to produce exactly three distinct MLOs, likely due to steric effects and the intrinsic slow relaxation of phase C. When replacing component C with $\tilde{\mathrm{C}}$, which features identical KLs but lacks any FLAPs, synthetic cells with exactly three MLOs were more common (Fig. S34), suggesting that the presence of the aptamers has an effect on the coarsening kinetics of the material.

We further expanded the possible organelle architectures in the synthetic cells by introducing linker RNA nanostars, dubbed L, that could control the mixing between the A and B designs upon their simultaneous transcription. Nanostar L is "chimeric", as it features two A-type [38] and two B-type KLs (Fig. 4a). As shown in Fig. 4b, low fractions of the linker template ([A-T]:[L-T]:[B-T]=10:1:10, or Linker Fraction = 1/21) produced grape-like clusters by inducing adhesion between condensates rich in A and B, while blocking their relaxation into two large condensates. Arrested coarsening is arguably due to inter-phase adhesion limiting the ability of the condensates to slide past each other. At higher Linker Fractions, the two opposite-type droplets coarsened into bigger domains, with Janus, capsule-like, and Russian-doll morphologies emerging at [A-T]:[L-T]:[B-T] = 5:1:5 (Linker Fraction = 1/11), 3:1:3 (Linker Fraction = 1/7) and 2:1:2 (Linker Fraction = 1/5), respectively (see also Figs S36-S40 and videos S10, S11).

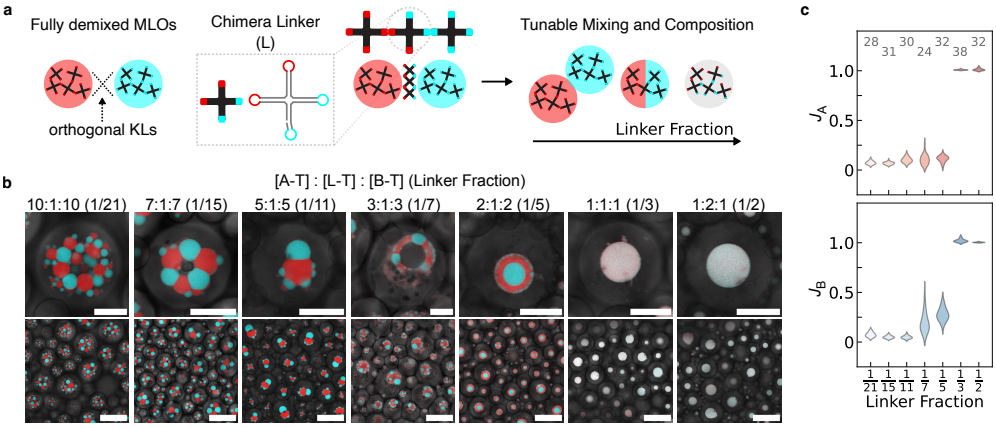

Fig. 4 Controlling morphology and composition of membrane-less organelles with linker RNA nanostars. (a) Chimeric RNA linker nanostars (L), with two A and two B KLs, enable control over mixing in systems of A and B nanostars by varying the relative concentration of the linker template L-T. (b) Zoomed-in (top) and larger field-of-view (bottom) confocal micrographs of synthetic cells producing A, B and L nanostars with different ratios between DNA templates ([A-T]:[L-T]:[B-T]). The Linker Fraction, shown in brackets, is computed as [L-T]/([A-T]+[L-T]+[B-T]). All scale bars are 50 μm. (c) Distributions of mixing indices J_A and J_B of the MLOs computed as discussed in the SI Methods (section 1.4) and shown as a function of the Linker Fraction for samples in (b). Low J_A and J_B are indicative of purer A-rich and B-rich phases, while $J_A, J_B \approx 1$ indicate complete mixing of the two RNA species. Numbers indicate examined synthetic cells.

For the latter, the fact that the A-phase formed the outer shell is consistent with the observation of a lower melting point compared to B, indicating weaker KL interactions and arguably lower interfacial tension with the bulk solution. For linker fractions of 1/3 and above, single phase spherical condensates were observed.

The abundance of linker nanostars also influences the degree of mixing of the two phases, assessed by measuring indices J_A and J_B, computed as the ratios between intensity of the fluorescent signal from the minority RNA component (*e.g.* signal from component A in B-rich condensates) over that recorded for the intensity of the majority phase (*e.g.* signal from A in A-rich condensates), as outlined in the SI Methods (section 1.3). The data, collated in Fig. 4i show a little mixing ($J_A, J_B \ll 1$) for low linker fraction, followed by moderate increase in mixing when more linker was produced, and by an abrupt jump to $J_A, J_B \approx 1$ upon reaching the threshold for complete mixing.

While the data in Figs 1-4 demonstrate that RNA condensates can selectively sequester small molecules, namely the fluorophores associated to MGA and BrA, imitating the functions of natural MLOs requires capturing larger and more functionally relevant macromolecules, particularly proteins. To this end, we modified designs A and B to include a 5′ overhang, to which a protein-binding RNA aptamer can connect *via* base pairing (Fig. 5a (i) and (ii)). In particular, the new nanostructures A_{YFP} and

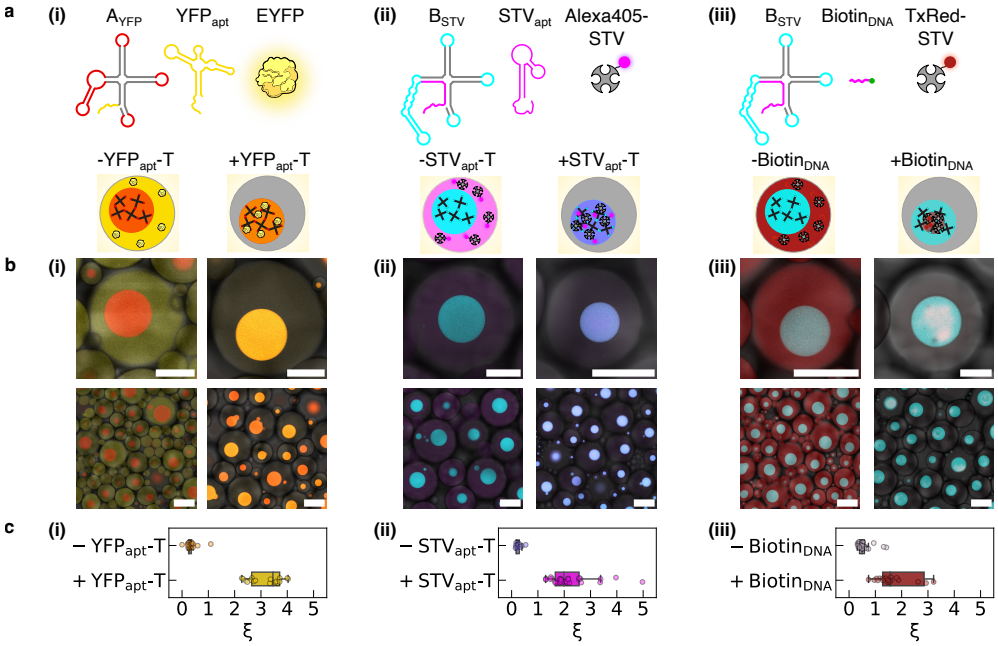

Fig. 5 Selective protein capture in designer RNA MLOs. (a) RNA nanostar designs are modified to include a single-stranded 5′ overhang that can connect to a protein-binding moiety. Design A_{YFP} is identical to nanostar A (Fig. 1) but can connect to YFP-binding aptamer (YFP_{apt}) (i). Design B_{STV} is identical to B but can connect to either a streptavidin (STV)-binding aptamer (STV_{apt}) (ii) or a biotinylated DNA oligonucleotide ($Biotin_{DNA}$) (iii). The templates of protein-binding aptamers (YFP_{apt}-T and STV_{apt}-T) are transcribed in synthetic cells alongside the associated RNA nanostar templates (A_{YFP}-T and B_{STV}-T, respectively), while pre-synthesised $Biotin_{DNA}$ is encapsulated alongside template B_{STV}-T. (b) Diagrams (top) and confocal micrographs (bottom) of synthetic cells expressing RNA MLOs in the absence (left) or in the presence (right) of protein-binding moieties for the systems presented in (a). When protein-binding aptamers are expressed or the biotinylated DNA oligonucleotide is included, target proteins are recruited in the MLOs. Scale bars are 50 μm (top) and 150 μm (bottom). (c) Protein partitioning parameter, ξ, computed from confocal micrographs in (b) as the ratio between the fluorescence signal of the target proteins recorded within and outside the MLOs for individual synthetic cells (see SI Methods, section 1.4). Data are shown for the three systems in (a) and (b), with and without protein-binding moieties. In the box plots, the central line marks the median, the box marks the Q1-Q3 interquartile range (IRQ) and the whiskers enclose datapoints within Q1$-1.5\times$IQR and Q1$+1.5\times$IQR. All datapoints are shown except for four outliers with $\xi > 5$ in (i) and (iii), omitted for ease of visualisation.

B_{STV} were designed to connect to a YFP-binding aptamer (YFP_{apt}) [39], and a streptavidin (STV)-binding aptamer (STV_{apt}) [40], respectively (see also Fig. S41). For both designs, when co-expressing the modified nanostars and their partner aptamers from distinct templates in synthetic cells (Fig. S42), the target proteins (EYFP and Alexa405-STV) readily partitioned within the formed MLOs (Figs 5b (i) and (ii), S43, S44a, and video S12). If the aptamer template was omitted, the target proteins remained in the lumen of the droplet-based synthetic cell. Protein partitioning was quantified through the parameter ξ, computed from image segmentation as the ratio between the fluorescence intensity of the protein recorded within or outside the

MLOs for individual synthetic cells (SI Methods section 1.3). When protein-binding aptamers were present, the median ξ was ~ 3.5 and ~ 2 for YFP and STV, respectively, while it fell below 0.5 when no aptamers were present (Fig. 5c (i) and (ii)).

As an alternative strategy for STV capture, we replaced $\mathrm{STV_{apt}}$ with a biotinylated DNA oligonucleotide complementary to the overhang in nanostar $\mathrm{B_{STV}}$ (Figs 5a (iii), S43, S44 and video S12). This approach still enables B-type MLOs to capture STV (Figs 5b (iii) and c (iii)), with the difference that the protein was distributed non-uniformly within the RNA condensates. Indeed, we observed irregularly-shaped clusters with solid-like appearance, reminiscent of multi-phase condensates found in cells [41–43]. The non-uniform protein distribution likely results from the finite amount of the biotinylated DNA anchor available, which is all sequestered at early transcription stages and thus accumulates at the centre of the condensates. The solid-like look of the protein-rich material may be a consequence of the tetravalent STV, which can irreversibly crosslink up to four RNA nanostars and thus make the material more viscous.

3 Discussion

Our platform enables the expression of synthetic condensates and membrane-less organelles with prescribed size, number, morphology, composition, and the ability to capture small guest molecules and proteins. The elementary building blocks, RNA nanostars, were designed from first principles utilising the rule-based approaches of nucleic acid nanotechnology, which provide extensive opportunities for further design variations aimed at programming arbitrary characteristics of the designer MLOs. Among the many design features that can be straightforwardly controlled are nanostar valency, flexibility and arm-length, all known to predictably influence self-assembly in analogous DNA systems [18, 20, 25, 30, 44]. Further RNA aptamers might be embedded to recruit molecular guests, including enzymes and metabolites, while ribozymes [45] could confer catalytic properties to the synthetic MLOs. Because the condensates can be expressed from DNA templates, their formation could be controlled through standard transcription regulation pathways in both synthetic cells and living cells. Owing to their open-ended programmability, we expect that the RNA-nanostar condensates will constitute a valuable new solution for the toolkit of synthetic biology.

Supplementary information. Supplementary information available: Methods, Figures, Tables and Videos.

Data availability. Raw data underpinning these results are available from the corresponding author during the review process. Upon possible publication the data will be shared on a freely accessible permanent repository, for which the DOI link will be provided.

Code availability. All data analysis code is available from the corresponding author.

Competing interests. The Regents of University of California has filed a patent application in the U.S. Patent and Trademark Office which includes disclosure of inventions described in this manuscript, Provisional Application Serial No. 63/588,142, filed on October 5, 2023, and entitled: SINGLE STRANDED RNA MOTIFS FOR IN VITRO COTRANSCRIPTIONAL PRODUCTION OF ORTHOGONAL PHASE SEPARATED CONDENSATES. Inventors: L. Di Michele, G. Fabrini, E. Franco. S. Li, A. Tang.

Acknowledgments. LDM acknowledges support from the European Research Council (ERC) under the Horizon 2020 Research and Innovation Programme (ERC-STG No 851667 – NANOCELL) and a Royal Society University Research Fellowship (UF160152, URF\R\221009). GF acknowledges funding from the Department of Chemistry at Imperial College London. MDA acknowledges support from a Biotechnology and Biological Sciences Research Council (BBSRC) David Phillips Fellowship (BB/R011605/1) and a Lister Institute Research Prize. SPN acknowledges support from the Engineering and Physical Sciences Research Council (EPSRC) (EP/S023518/1). JMS is a Merck Awardee of the Life Sciences Research Foundation. EF acknowledges support from the US NSF through CAREER award 1938194 and FMRG: Bio award 2134772, and from the Sloan Foundation through award G-2021-16831.

References

[1] Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).

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

[3] Lyon, A. S., Peeples, W. B. & Rosen, M. K. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **22**, 215–235 (2021).

[4] Feric, M. *et al.* Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686–1697 (2016).

[5] Hirose, T., Ninomiya, K., Nakagawa, S. & Yamazaki, T. A guide to membraneless organelles and their various roles in gene regulation. *Nat. Rev. Mol. Cell Biol.* **24**, 288–304 (2023).

[6] Gomes, E. & Shorter, J. The molecular language of membraneless organelles. *J. Biol. Chem.* **294**, 7115–7127 (2019).

[7] Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).

[8] Patel, A. *et al.* A Liquid-to-Solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066–1077 (2015).

[9] Hastings, R. L. & Boeynaems, S. Designer condensates: A toolkit for the biomolecular architect. *J. Mol. Biol.* **433**, 166837 (2021).

[10] Dai, Y., You, L. & Chilkoti, A. Engineering synthetic biomolecular condensates. *Nature Reviews Bioengineering* 1–15 (2023).

[11] Buddingh', B. C. & van Hest, J. C. M. Artificial cells: Synthetic compartments with life-like functionality and adaptivity. *Acc. Chem. Res.* **50**, 769–777 (2017).

[12] Guo, H. *et al.* Spatial engineering of e. coli with addressable phase-separated RNAs. *Cell* **185**, 3823–3837.e23 (2022).

[13] Xue, Z. *et al.* Targeted RNA condensation in living cells via genetically encodable triplet repeat tags. *Nucleic Acids Research* **51**, 8337–8347 (2023).

[14] Seeman, N. C. & Sleiman, H. F. DNA nanotechnology. *Nature Reviews Materials* **3**, 1–23 (2017).

[15] Geary, C., Rothemund, P. W. K. & Andersen, E. S. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science* **345**, 799–804 (2014).

[16] McRae, E. K. S. *et al.* Structure, folding and flexibility of co-transcriptional RNA origami. *Nat. Nanotechnol.* (2023).

[17] Stewart, J. M. *et al.* Modular rna motifs for orthogonal phase separated compartments. *bioRxiv* (2023). URL https://www.biorxiv.org/content/early/2023/10/06/2023.10.06.561123.

[18] Biffi, S. *et al.* Phase behavior and critical activated dynamics of limited-valence DNA nanostars. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15633–15637 (2013).

[19] Sato, Y., Sakamoto, T. & Takinoue, M. Sequence-based engineering of dynamic functions of micrometer-sized DNA droplets. *Sci Adv* **6**, eaba3471 (2020).

[20] Brady, R. A., Brooks, N. J., Cicuta, P. & Di Michele, L. Crystallization of amphiphilic DNA C-Stars. *Nano Lett.* **17**, 3276–3281 (2017).

[21] Liu, D. *et al.* Branched kissing loops for the construction of diverse RNA homooligomeric nanostructures. *Nat. Chem.* **12**, 249–259 (2020).

[22] Kolpashchikov, D. M. Binary malachite green aptamer for fluorescent detection of nucleic acids. *J. Am. Chem. Soc.* **127**, 12442–12443 (2005).

[23] Yerramilli, V. S. & Kim, K. H. Labeling RNAs in live cells using malachite green aptamer scaffolds as fluorescent probes. *ACS Synth. Biol.* **7**, 758–766 (2018).

[24] Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.* **136**, 16299–16308 (2014).

[25] Brady, R. A., Brooks, N. J., Foderà, V., Cicuta, P. & Di Michele, L. Amphiphilic-DNA platform for the design of crystalline frameworks with programmable structure and functionality. *J. Am. Chem. Soc.* **140**, 15384–15392 (2018).

[26] Saleh, O. A., Jeon, B.-J. & Liedl, T. Enzymatic degradation of liquid droplets of DNA is modulated near the phase boundary. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 16160–16166 (2020).

[27] Chamberlin, M. & Ring, J. Characterization of t7-specific ribonucleic acid polymerase: I. general properties of the enzymatic reaction and the template specificity of the enzyme. *J. Biol. Chem.* **248**, 2235–2244 (1973).

[28] Arnold, S. *et al.* Kinetic modeling and simulation of in vitro transcription by phage T7 RNA polymerase. *Biotechnol. Bioeng.* **72**, 548–561 (2001).

[29] Akama, S., Yamamura, M. & Kigawa, T. A multiphysics model of in vitro transcription coupling enzymatic reaction and precipitation formation. *Biophys. J.* **102**, 221–230 (2012).

[30] Agarwal, S., Osmanovic, D., Klocke, M. A. & Franco, E. The growth rate of DNA condensate droplets increases with the size of participating subunits. *ACS Nano* **16**, 11842–11851 (2022).

[31] Han, K. Y., Leslie, B. J., Fei, J., Zhang, J. & Ha, T. Understanding the photophysics of the spinach-DFHBI RNA aptamer-fluorogen complex to improve live-cell RNA imaging. *J. Am. Chem. Soc.* **135**, 19033–19038 (2013).

[32] Li, X., Kim, H., Litke, J. L., Wu, J. & Jaffrey, S. R. Fluorophore-Promoted RNA folding and photostability enables imaging of single Broccoli-Tagged mRNAs in live mammalian cells. *Angew. Chem. Int. Ed Engl.* **59**, 4511–4518 (2020).

[33] Levitz, P. Off-lattice reconstruction of porous media: critical evaluation, geometrical confinement and molecular transport. *Adv. Colloid Interface Sci.* **76-77**, 71–106 (1998).

[34] Testard, V., Berthier, L. & Kob, W. Influence of the glass transition on the liquid-gas spinodal decomposition. *Phys. Rev. Lett.* **106**, 125702 (2011).

[35] Di Michele, L. *et al.* Multistep kinetic self-assembly of DNA-coated colloids. *Nat. Commun.* **4**, 2007 (2013).

[36] Hansch, C., Leo, A. & Hoekman, D. *Exploring QSAR: hydrophobic, electronic, and steric constants* Vol. 2 (American Chemical Society Washington, DC, 1995).

[37] Franco, E. *et al.* Timing molecular motion and production with a synthetic transcriptional clock. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E784–93 (2011).

[38] Jeon, B.-J., Nguyen, D. T. & Saleh, O. A. Sequence-Controlled adhesion and microemulsification in a Two-Phase system of DNA liquid droplets. *J. Phys. Chem. B* **124**, 8888–8895 (2020).

[39] Shui, B. *et al.* RNA aptamers that functionally interact with green fluorescent protein and its derivatives. *Nucleic Acids Res.* **40**, e39 (2012).

[40] Leppek, K. & Stoecklin, G. An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. *Nucleic Acids Res.* **42**, e13 (2014).

[41] Yamazaki, T., Nakagawa, S. & Hirose, T. Architectural RNAs for membraneless nuclear body formation. *Cold Spring Harb. Symp. Quant. Biol.* **84**, 227–237 (2019).

[42] Schuster, B. S. *et al.* Biomolecular condensates: Sequence determinants of phase separation, microstructural organization, enzymatic activity, and material properties. *J. Phys. Chem. B* **125**, 3441–3451 (2021).

[43] Yewdall, N. A., André, A. A. M., Lu, T. & Spruijt, E. Coacervates as models of membraneless organelles. *Curr. Opin. Colloid Interface Sci.* **52**, 101416 (2021).

[44] Brady, R. A., Kaufhold, W. T., Brooks, N. J., Foderà, V. & Di Michele, L. Flexibility defines structure in crystals of amphiphilic DNA nanostars. *J. Phys. Condens. Matter* **31**, 074003 (2019).

[45] Tanner, N. K. Ribozymes: the characteristics and properties of catalytic RNAs. *FEMS Microbiol. Rev.* **23**, 257–275 (1999).