Smart nucleic acid condensates for bottom-up synthetic biology

Giacomo Fabrini

Department of Chemistry
Imperial College London

This dissertation is submitted for the degree of
Doctor of Philosophy

October 2023

Statement of Originality

I, *Giacomo Fabrini*, hereby declare that, except where the work of others is specifically referenced, this dissertation and its contents are my original work. Any experimental or intellectual contribution from collaborators to the work here presented has been duly recognised in the text, Preamble and Acknowledgments. The contents and results presented in this thesis have not been submitted, fully or partially, for any other degree or qualification in this, or any other university. This dissertation contains fewer than 100,000 words, bibliography and appendices excluded as per College guidance.

Giacomo Fabrini
October 2023

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).
Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.
When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.
Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

I would like to dedicate this thesis to my family, for their unwavering support, and to my partner, Sabrina, for her continuous encouragement and enthusiasm for science.

Acknowledgements

First of all, I would like to thank my supervisor, Dr. Lorenzo Di Michele, whose guidance and scientific enthusiasm were paramount to the development of the projects in this thesis. Thank you for always asking the questions I had not even imagined of asking - I will always take with me your playful yet rigorous scientific curiosity.
I would then like to thank my (now official) co-supervisor, Dr. Marco Di Antonio, for his collaboration and for all the chats we had during the first few months after lockdown when I was pretty much isolated from the rest of my group.
Thanks to all the people in the DiMichele Lab, the DiAntonio Group and the Membrane Biophysics Group. In particular, special thanks go to Will, who sparked my interest in machine learning, Roger, for always being encouraging and tremendously supporting in tough times, and Michal, for being the best and funniest Diamond trip mate I could have wished for. Last but not least, thanks to my mate Souro, for walking through this Ph.D. journey together, from one plate reader run to another, and for all the game nights with Anisha and Sabrina - here's to so many more in years to come!
Thanks to Dr. Aisling Minard and Dr. Ryan Brady for their collaboration and insights. Thanks also to Stephen Rothery from the Facility for Imaging by Light Microscopy for patiently helping me in my countless hours at the confocal.
I would like to thank the Chemistry Department at Imperial College London for funding my PhD and for having me as a Graduate Teaching Assistant for the past four years.
Thanks to Jacopo, for always being a piece of home and family in the UK.
The most special thanks go to Sabrina, for being my partner and the best collaborator I could have hoped for. You are my biggest inspiration, in life and in science, and the most precious gift of this Ph.D.
Finally, no acknowledgments would be complete without massively thanking my family. Thanks to my parents, Angela and Enrico, who have always supported me in all possible ways, even when they did not agree with the path I had chosen. Thanks for being there and for always being up to travel the world with me. Thanks also to Chiara and Valerio for putting up with me never being in touch and always being available for a board game night.

Abstract

Bottom-up synthetic biology aims at rationally engineering life-like behaviours, such as gene expression, energy production and communication, in synthetic entities assembled from bio-active components, termed artificial cells. The complex functions of biological cells often emerge from their heterogeneous sub-compartmentalised architecture, from the outer lipid membrane to inner organelles. Therefore, the capability to engineer programmable micro-compartments is key to developing biomimetic artificial cells.
While lipid, polymer and protein-based approaches have been proposed to replicate the semi-permeable cellular membrane, solutions relying on coacervates, hydrogels and biomolecular condensates are gaining traction as versatile and tunable mimics of cellular membrane-less organelles. Nucleic acid nanotechnology has emerged as a prime route to develop biomimetic artificial cellular platforms thanks to the control it enables over structure and dynamics at the nano- and microscale. In particular, branched nucleic acid nanostructures ("nanostars") have been extensively researched and exploited to build programmable nanomaterials, including liquid-like condensates, hydrogels and crystals, with potential applications in bottom-up synthetic biology.

In this thesis I investigate two approaches to engineer smart artificial cellular compartments based on nucleic acid condensates. The first strategy modularly combines the self-assembly robustness of amphiphilic DNA nanostars with the cation-responsiveness of non-canonical DNA G-quadruplexes. The resulting nanostructures can isothermally self-assemble into amphiphilic hydrogels upon addition of physiologically-relevant concentrations of potassium, and disassemble upon its chelation. The second approach exploits biological machinery to produce RNA nanostars from the transcription of DNA templates. The RNA nanostructures fold co-transcriptionally and self-assemble into condensates, both in bulk and in simple synthetic cells. The resulting RNA organelles can selectively recruit small molecules and proteins, and boast programmable number, morphology and mixing/demixing behaviours.
Overall, my findings contribute to demonstrating the largely untapped potential that nucleic acid-based synthetic condensates hold as a modular platform for artificial cell engineering, potentially paving the way to advanced biomimetic functionalities.

Preamble

Publications

Some of the ideas discussed in Chapter 1 regarding the potential of amphiphilic DNA nanotechnology for bottom-up synthetic biology and constructing biomimetic minimal model systems have been published in the following Feature Article:

R. Rubio-Sánchez[†], **G. Fabrini**[†], P. Cicuta, and L. Di Michele, Amphiphilic DNA nanostructures for bottom-up synthetic biology, *Chemical Communications*, 2021, 57, 12725.

The work on cation-responsive Quad-Stars presented in Chapter 3 has been published in:

G. Fabrini, A. Minard, R. A. Brady, M. Di Antonio and L. Di Michele, Cation-Responsive and Photocleavable Hydrogels from Noncanonical Amphiphilic DNA Nanostructures, *Nano Letters*, 2022, 22, 2, 602-611.

The results on co-transcriptional self-assembling RNA organelles discussed in Chapters 4-6 have been summarised in a manuscript, preprinted on bioRxiv and soon to be submitted for peer-reviewed publication. In particular, while all experiments and analysis shown here have been performed by me and (when mentioned) Sabrina Pia Nuccio, this manuscript has been co-authored with two other groups, one at UCLA (Franco Lab) and the other at Caltech (Rothemund Lab), who have conducted related research (not discussed in this thesis):

G. Fabrini, S.P. Nuccio, J.M. Stewart, S. Li, A. Tang, P.W.K.R. Rothemund, E.Franco, M. Di Antonio and L. Di Michele, Co-transcriptional production of programmable RNA condensates and synthetic organelles, preprint on bioRxiv and soon to be submitted for peer review.

[†]: These authors contributed equally to the work.

Collaborations and external contributions

During these four years, I have been lucky to receive help from and entertain insightful discussions with several collaborators, some of which have made direct contributions featured in this Ph.D. thesis and are thus listed below.
The results presented in Fig. 3.4 emerged from measurements performed jointly with Dr. Aisling Minard.
None of the results presented in Chapters 4-6 would have been possible without the expertise of Sabrina Pia Nuccio, who helped with optimising the PCR amplification protocols and patiently gel purified all the DNA templates used in downstream experiments. Moreover, Sabrina entirely carried out the preliminary work on synthetic RNA organelles expressed in *E. coli* featured in Chapter 7, for which I performed the imaging.

Abbreviations

A – Adenine
AFM – Atomic Force Microscopy
ATP – Adenosine triphosphate
bp – Base pair
BrA – Broccoli Aptamer
BSA – Bovine Serum Albumin
C – Cytosine
CD – Circular Dichroism
CFP – Cyan Fluorescent Protein
CHT – Circle Hough Transform
CLD – Chord Length Distribution
CTP – Cytidine triphosphate
DLS – Dynamic Light Scattering
DMEM – Dulbecco's Modified Eagle Medium
DNA – Deoxyribonucleic Acid
dsDNA/dsRNA – double-stranded DNA/RNA
(E)YFP – (Enhanced) Yellow Fluorescent Protein
FBS – Fetal Bovine Serum
FLAPs – Fluorescent Light-Up Aptamers
FOV(s) – Field(s)-of-view (microscopy)
FRET – Förster/Fluorescence resonance energy transfer
G – Guanine
G4 – G-quadruplex
GFP – Green Fluorescent Protein
GTP – Guanosine triphosphate
kb – kilobase
KL(s) – Kissing Loops(s)
MG – Malachite Green
MGA – Malachite Green Aptamer
NMM – N-methyl Mesoporphyrin IX
nt – Nucleotide
PBS – Phosphate-buffered Saline

px – Pixel
RNA – Ribonucleic Acid
rNTPs – Ribonucleotide triphosphates
ROI – Region of interest
ssDNA/ssRNA – single-stranded DNA/RNA
STV – Streptavidin
SynCell – Synthetic cell
T – Thymine
T7 RNAP – T7 RNA Polymerase
TBE – Tris/Borate/EDTA buffer
TE – Tris/EDTA buffer
TFO – Triplex Forming Oligonucleotide
TMR – Tetramethylrosamine
TMSD – Toehold-Mediated Strand Displacement
TTS – Triplex Target Site
U – Uracil
UTP – Uridine triphosphate
UV – Ultraviolet

Table of contents

List of figures **21**

List of tables **25**

1 Introduction **27**
1.1 Bottom-Up Synthetic Biology: the quest towards building Artificial Cells . 27
1.1.1 The importance of compartmentalisation 28
1.2 Nucleic Acids and Nucleic Acids Nanotechnology 29
1.2.1 Structure of nucleic acids 30
1.2.2 Thermodynamics of hybridisation 32
1.2.3 Biological role of DNA and RNA 36
1.3 DNA Nanotechnology 38
1.3.1 Structural DNA Nanotechnology 38
1.3.2 Dynamic DNA Nanotechnology 40
1.3.3 Non-canonical DNA motifs in DNA Nanotechnology 42
1.3.4 DNA nanotechnology for bottom-up synthetic biology 42
1.4 RNA Nanotechnology 45
1.4.1 Advantages and motivations to pursue RNA Nanotechnology 45
1.4.2 RNA Aptamers 48
1.4.3 Assembling RNA nanostructures from conserved motifs 52
1.4.4 Transcription and co-transcriptional folding 54
1.5 Objectives and outline of this thesis 59

2 Materials and Methods **63**
2.1 Design 63
2.2 Material handling 65
2.3 DNA nanostructure preparation 67
2.4 RNA nanostructure preparation 68
2.4.1 In Vitro Transcription: kit and protocol 69
2.4.2 Sample preparation 69
2.5 Agarose Gel Electrophoresis 70

2.6 Dynamic Light Scattering (DLS) . . . 71
2.7 Circular Dichroism (CD) . . . 73
2.8 Fluorimetry . . . 76
2.8.1 Excitation and emission scans . . . 76
2.8.2 Fluorescence Intensity Kinetics . . . 77
2.9 Bright-field and epifluorescence optical microscopy . . . 80
2.10 Laser Confocal Scanning microscopy . . . 84
2.11 Numerical methods . . . 85
2.11.1 Gel Image Analysis . . . 85
2.11.2 Fitting DLS Size Distributions . . . 85
2.11.3 Data analysis and fitting of Quad-Stars fluorescence kinetics . . . 86
2.11.4 Data analysis and fitting of RNA nanostar fluorescence kinetics . . . 86
2.11.5 General image processing for micrographs . . . 87
2.11.6 Segmentation of Quad-Star aggregates . . . 88
2.11.7 Isothermal assembly/disassembly area fraction profiles . . . 88
2.11.8 Rhodamine B cargo loading in Quad-Star aggregates . . . 88
2.11.9 Photo-induced disassembly . . . 89
2.11.10 Growth of RNA condensates in bulk/synthetic cells . . . 89
2.11.11 Calculation of Kissing Loop Melting Temperature . . . 91
2.11.12 CHT for radial distribution of orthogonal RNA condensates in bulk 91
2.11.13 Scaling of RNA organelle volume with synthetic cell volume . . . 92
2.11.14 Distribution of radial ratio of orthogonal RNA organelles in emulsion droplets . . . 92
2.11.15 Number of orthogonal RNA organelles within synthetic cells . . . 93
2.11.16 Evaluating linker-induced mixing . . . 94
2.11.17 Evaluation of protein partitioning with protein-binding constructs . 95

3 Cation-responsive amphiphilic non-canonical DNA nanostructures 97
3.1 Introduction . . . 97
3.2 Quad-Star Design . . . 98
3.3 Characterisation . . . 100
3.3.1 Characterisation of non-cholesterolised Quad-Stars . . . 101
3.3.2 Characterisation of cholesterolised Quad-Stars . . . 111
3.4 Reversible isothermal self-assembly and disassembly . . . 114
3.4.1 Aggregate formation kinetics . . . 116
3.4.2 Disassembly by potassium ion chelation through cryptand . . . 117
3.4.3 Assembly reversibility . . . 118
3.5 Cargo-loading capabilities and photo-induced irreversible disassembly . . . 118
3.5.1 Rhodamine B uptake and triggered release . . . 118
3.5.2 Biostability of Quad-Star aggregates . . . 120

3.5.3 NMM-enabled photodisassembly . . . 122
3.6 Chapter summary and outlook . . . 123

4 RNA condensates from co-transcriptionally folding RNA nanostars 127
4.1 Introduction . . . 127
4.2 RNA nanostar Design . . . 129
4.2.1 Kissing Loops . . . 131
4.2.2 DNA Template and PCR Primer Design . . . 134
4.3 Characterisation . . . 135
4.3.1 Characterisation of PCR amplified DNA templates . . . 135
4.3.2 Characterisation of RNA transcripts and aptamer folding . . . 135
4.4 Co-transcriptionally folding RNA nanostars form liquid condensates . . . 139
4.4.1 Growth kinetics . . . 140
4.4.2 Kissing Loop interactions are programmable . . . 142
4.4.3 Melting of bulk condensates sheds light on KL interaction strength . 142
4.5 Programmable KL interactions enable orthogonal condensate formation . . 143
4.5.1 Tuning the relative size of orthogonal condensates *via* DNA template ratio in binary systems . . . 143
4.5.2 Growth kinetics in binary systems . . . 145
4.5.3 Ternary systems . . . 145
4.6 Chapter summary and outlook . . . 147

5 RNA organelles 149
5.1 Introduction . . . 149
5.2 Synthetic RNA organelles in water-in-oil droplets . . . 150
5.2.1 Emulsion droplets . . . 151
5.2.2 RNA nanostars form synthetic organelles in emulsion droplets . . . 151
5.2.3 Comparison with fluorimetry measurements . . . 154
5.3 Multiple tunable RNA-based compartments in synthetic cell mimics . . . 156
5.3.1 Binary RNA nanostar systems . . . 156
5.3.2 Increasing the number of compartments: ternary systems . . . 158
5.4 Controlling mixing/demixing to unlock complex organelle morphologies . . 162
5.4.1 Tuning RNA organelle morphology *via* linker template fraction . . 162
5.4.2 Quantitatively evaluating mixing with varying Linker Fraction . . . 163
5.5 Chapter summary and outlook . . . 164

6 Towards functional RNA-based organelles in artificial cells 167
6.1 Introduction . . . 167
6.2 Design and characterisation of protein-capturing RNA constructs . . . 168
6.2.1 Nanostructure design: adding sticky ends . . . 168

6.2.2 Characterisation . . . 171
6.3 RNA organelles can capture target proteins *via* simple design modifications 174
6.3.1 Homogeneous protein enrichment *via* protein-binding RNA aptamers 174
6.3.2 Inhomogeneous STV capture *via* receptor-ligand interactions . . . 175
6.3.3 Quantification of protein uptake . . . 177
6.3.4 Simultaneous capture of distinct target proteins . . . 178
6.4 Chapter summary and outlook . . . 182

7 Conclusions & future research directions 185
7.1 Summary & discussion . . . 186
7.2 Outlook and directions for future research . . . 187
7.2.1 Encapsulation of RNA MLOs in more realistic artificial cells . . . 188
7.2.2 Photo-activated transcription in synthetic cells . . . 189
7.2.3 Multiple protein/enzyme colocalisation for enzymatic cascades . . . 189
7.2.4 Synthetic MLOs to influence the behaviour of living agents . . . 190

References 193

Appendix A Nanostructure and oligonucleotide sequences 229
A.1 Chapter 3 . . . 229
A.2 Chapters 4 and 5 . . . 231
A.3 Chapter 6 . . . 235

Appendix B Supporting Figures 237
B.1 Chapter 3 . . . 237
B.2 Chapter 4 . . . 241
B.3 Chapter 6 . . . 248

List of figures

1.1 Synthetic cells are rationally engineered to replicate life-like behaviours. . . 28
1.2 Chemical structure of DNA and RNA 31
1.3 Non-canonical nucleic acid motifs. 35
1.4 Structural DNA nanotechnology enables the construction of complex supramolecular assemblies with nanoscale resolution. 39
1.5 Toehold-mediated strand displacement and dynamic DNA nanotechnology. 41
1.6 Non-canonical DNA motifs for stimuli-responsiveness and gene expression regulation. 43
1.7 DNA-based cytoskeleta and condensates for artificial cells. 44
1.8 Amphiphilic DNA frameworks for bottom-up synthetic biology. 46
1.9 Structural variety of conserved RNA motifs. 47
1.10 Functional variety of RNA motifs. 48
1.11 Malachite Green aptamer. 50
1.12 Broccoli and Spinach aptamers. 51
1.13 RNA nanostructures from the modular composition of conserved motifs. . . 53
1.14 Mechanism and structure of T7 RNA Polymerase. 55
1.15 Co-transcriptional folding of complex RNA nanostructures. 58

2.1 Epifluorescence imaging setup. 83

3.1 Design of cation-responsive amphiphilic DNA Quad-Stars. 99
3.2 Self-assembly of non-cholesterolised Quad-Star variants *via* agarose gel electrophoresis. 101
3.3 Self-assembly of non-cholesterolised Quad-Star variants *via* DLS. 103
3.4 Circular Dichroism (CD) spectra confirm G4-mediated self-assembly. . . . 104
3.5 Kinetics of G4-dependent self-assembly of non-cholesterolised Quad-Stars *via* NMM fluorimetry. 105
3.6 Influence of monovalent cations on the self-assembly of non-cholesterolised Quad-Stars. 110
3.7 Influence of Mg^{2+} ions on the self-assembly of non-cholesterolised Quad-Stars. 111
3.8 Influence of annealing cationic conditions on cholesterol micelle formation. 112

3.9 Kinetics of G4-dependent self-assembly of cholesterolised Quad-Stars *via* NMM fluorimetry. 114
3.10 Isothermal assembly and hydrogel aggregate formation. 115
3.11 Quad-Stars and their hydrogel aggregates can reversibly assemble and disassemble upon addition or chelation of K^+ ions. 117
3.12 Amphiphilic Quad-Star hydrogel aggregates can uptake and release model hydrophobic small molecule cargo, rhodamine B. 119
3.13 Stability of Quad-Star hydrogel aggregates under model physiological conditions. 121
3.14 NMM-mediated, irreversible and highly localised photo-induced disassembly of Quad-Star hydrogel aggregates. 123

4.1 Topology and design of co-transcriptionally folding RNA nanostars. 130
4.2 Agarose gel electrophoresis confirms PCR amplified DNA templates display the expected sizes. 136
4.3 Agarose gel electrophoresis elucidates folding and size of RNA transcripts. 137
4.4 Fluorogenic aptamers enable monitoring of the co-transcriptional folding of RNA nanostars. 138
4.5 Programmable Kissing Loop interactions drive the self-assembly of RNA condensates. 141
4.6 Thermal ramps reveal different Kissing Loop interaction strength. 142
4.7 Simultaneous transcription and condensate co-assembly in bulk binary RNA nanostar systems 144
4.8 Simultaneous transcription and condensate co-assembly in bulk ternary systems of RNA nanostars. 146

5.1 Self-assembly of RNA condensates within water-in-oil droplets. 152
5.2 RNA-enriched cytoplasm from non sticky nanostars and fluorimetry characterisation of transcription in emulsion droplets. 154
5.3 Co-assembly of orthogonal binary RNA organelles with controlled size and number in droplet-based synthetic cells. 157
5.4 Co-assembly with non sticky RNA nanostar variants and fluorimetry characterisation of binary systems in emulsion droplets. 159
5.5 Co-assembly of three orthogonal RNA condensates in droplet-based synthetic cells. 160
5.6 Control over mixing/demixing behaviours *via* linker nanostars unlocks complex organelle morphologies. 161

6.1 Design of co-transcriptionally folding RNA nanostars for target protein capture. 169

6.2 Agarose gel electrophoresis confirms DNA templates encoding for protein-capturing constructs display the expected sizes. 172
6.3 Agarose gel electrophoresis elucidates folding and size of protein-capturing RNA transcripts. 173
6.4 RNA organelles can capture target proteins through protein-binding constructs. 176
6.5 Simultaneous capture of distinct proteins in addressable domains of multi-phase RNA organelles. 179
6.6 Protein-capturing RNA nanostars form multi-phase organelles despite KL orthogonality. 181

7.1 Forming RNA MLOs in more realistic and biomimetic artificial cells with temporal transcriptional control. 188
7.2 Synthetic RNA MLOs in bacteria. 191

B.1 Molecular structure and fluorescence emission spectra of NMM. 237
B.2 Self-assembly kinetics of non-cholesterolised Quad-Stars with varying $[Oligo]_0$ and [KCl]. 238
B.3 Raw fluorescence kinetics profiles for cholesterolised Quad-Stars control samples. 239
B.4 Kinetics of rhodamine B cargo release upon cryptand-induced disassembly of Quad-Star hydrogel aggregates. 239
B.5 UV-triggered and NMM-mediated photo-induced disassembly of Quad-Star aggregates. 240
B.6 Kinefold analysis of RNA nanostars A, B and C. 241
B.7 Chord Length Distribution of orthogonal organelles formed by binary RNA nanostar systems in bulk. 242
B.8 Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 0.25 in bulk. 243
B.9 Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 0.50 in bulk. 244
B.10 Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 1 in bulk. 245
B.11 Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 2 in bulk. 246
B.12 Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 4 in bulk. 247
B.13 Self-assembly of protein-capturing RNA organelles. 248
B.14 Simultaneous transcription and co-assembly of distinct protein-capturing RNA organelles. 249

List of tables

A.1 Sequences of Quad-Star strand components. 229
A.2 Sequences of RNA nanostars used in Chapters 4-5. 231
A.3 Sequences of DNA templates used in Chapters 4-5. 233
A.4 Sequences of PCR primers used in Chapters 4-5. 234
A.5 Sequences of RNA nanostars used in Chapter 6. 235
A.6 Sequences of DNA templates used in Chapter 6. 236
A.7 Sequences of single-stranded DNA oligonucleotides used in Chapter 6. . . . 236

Chapter 1

Introduction

1.1 Bottom-Up Synthetic Biology: the quest towards building Artificial Cells

The last two decades have seen the progressive emergence and consolidation of the field of synthetic biology [1]. On one hand, *top-down* synthetic biology [2, 3] aims at artificially enhancing biological organisms by granting them non-natural capabilities through genetic engineering [4]. While such route has produced organisms with intriguing behaviours [5, 6], it also proved challenging. For one, the starting organisms often experience excessive burden [7, 8] due to having to carry out the externally-introduced functions while still devoting resources and energy to staying alive. Moreover, the artificial behaviours can result in toxic by-products which the organism is not fit to handle [9], or require further modifications to achieve the desired goal, such as optimising conditions for recombinant protein folding [10]. Understandably, the resulting process is expensive, highly regulated and linked to a variety of biological risks, further slowing its progress and limiting its applicability. On the other hand, the more recent approach of bottom-up synthetic biology [11, 12] strives to engineer minimal model systems capable of replicating life-like behaviours by combining together bio-active molecular building blocks and/or biologically-evolved machinery. As such, these inanimate entities, termed artificial or synthetic cells and schematically represented in Fig. 1.1, are intrinsically decoupled from the biological complexity of the top-down approach, enabling us to gain a deeper understanding of biological phenomena by learning how to construct them from scratch [13]. Inspired by cellular behaviours, researchers have developed minimal synthetic platforms capable of replicating gene expression and protein synthesis [14], energy production [15, 16], proto-cellular communication pathways [17], compartmentalised and controlled reactions [18], vesicular transport [19] as well as triggered self-assembly of actin [20] and actin-mimicking cytoskeleta [21]. While most artificial cell platforms to date are only capable of simultaneously performing a few of these functions at most, they clearly hold considerable potential in a variety of technological areas and industries, ranging from

biosensing to biosynthesis, environmental bioremediation and, of course, biomedicine and tissue engineering [22–25].

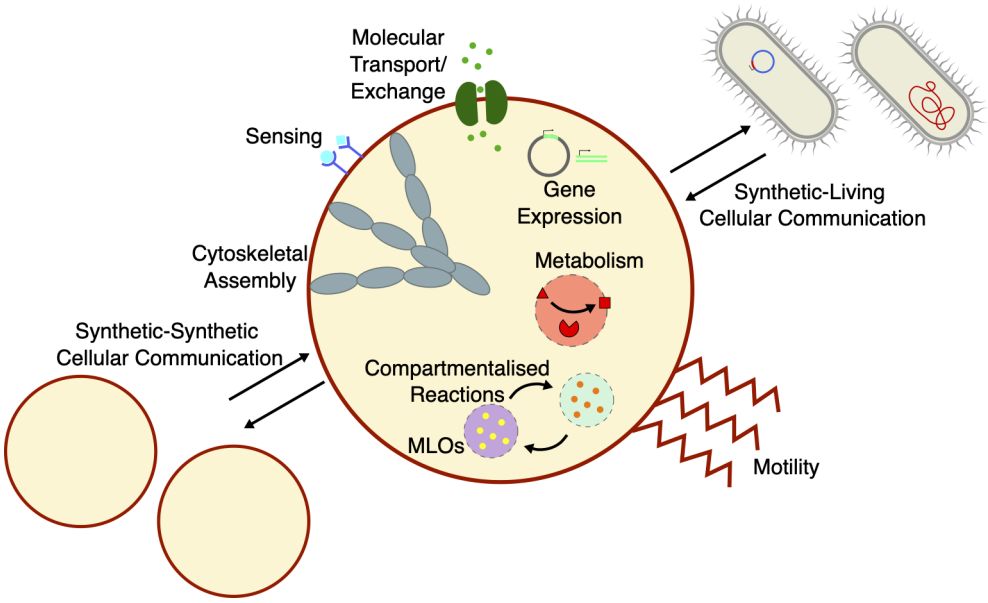


Fig. 1.1 **Synthetic cells are rationally engineered to replicate life-like behaviours.** Synthetic/artificial cells are minimal model systems built from the bottom-up from bio-active components and/or biological machinery, capable of replicating life-like behaviours and functionalities. The range of achieved biomimetic properties and mechanisms range from cytoskeletal self-assembly and inner sub-compartmentalisation to gene expression, biosensing, molecular transport, metabolism, motility and inter-cellular communication, either with other synthetic cells or living agents, such as bacteria.

1.1.1 The importance of compartmentalisation

The daunting and exquisitely regulated complexity of cellular phenomenology is made possible by the isolation of biochemical processes and reaction pathways in hierarchically compartmentalised environments [24, 26]. First, cells themselves are enclosed in a semi-permeable lipid membrane which enables them to maintain a considerably different inner biochemical milieu compared to the surrounding environment. Second, eukaryotic cells contain a myriad of internal compartments, both membrane-bound, such as the nucleus, the mitochondria and the endoplasmic reticulum [27, 28], and membrane-less, such as the nucleolus, Cajal bodies, processing (P) bodies, stress granules and nuclear speckles [29, 30, 28]. While the former are delimited by lipid membranes, the latter are assemblies of proteins and nucleic acids, particularly RNA, which mostly emerge *via* liquid-liquid phase separation (LLPS). Such LLPS is usually driven by scaffold components, such as intrinsically disordered proteins [29, 31, 32, 28] and architectural RNAs [33, 34], which can later recruit other client macromolecules which grant functionality to the assembly and can even produce subcompartmentalisation [34–36]. MLOs have been observed to play crucial roles ranging from responding to stress conditions, regulating RNA transport and protein sorting, as well as contributing to the biogenesis of biological machinery, including ribosomes

and the spliceosome [28]. Due to being membrane-less, they are highly dynamic and can quickly exchange material with the surrounding environment, as they are mostly liquid or hydrogel-like [29]. Interestingly, while prokaryotes lack membrane-bound organelles, they have recently been found to feature MLOs [37, 38], contributing to the growing interest [39] in these compartments, which have been termed the "organisers of cellular biochemistry" [30] due to their functional variety [28] and pervasiveness. This universality further proves that compartmentalisation is, in fact, key to ensuring cellular functionality, *i.e.* that biomolecular reactions happen at the right time, in the right place, with the right rate and with minimal cross-talk [29].
Considering that there is only so much that can be accomplished in a single biochemical environment, it is not surprising that hierarchical compartmentalisation is so well conserved in both eukaryotes and prokaryotes, enabling them to regulate molecular transport, maintain highly heterogeneous biomolecular distribution and simultaneously sustain otherwise incompatible reactions. It thus follows that, in order to replicate such behaviours and functionalities of living cells, artificial cells assembled from the bottom-up would benefit from a similar subcompartmentalised internal structure [24, 40]. Taking inspiration directly from nature, most synthetic cells are compartmentalised using a semi-permeable lipid [41, 42], protein [17, 43] or polymer-based membrane [42, 44]. While each approach as advantages and drawbacks, lipid-based Giant Unilamellar Vesicles (GUVs) boast the more veritable biomimicry [11] but suffer in terms of longevity. Encapsulation is also challenging, particularly for charged biomolecules, and often leads to poor yields [45, 46] without a microfluidic platform. Therefore, other compartmentalisation approaches, particularly water-in-oil emulsion droplets, are often used for rapid prototyping [20, 21]. Membrane-less compartments based on hydrogels [47–49] and coacervates [50, 51, 48] are also becoming increasingly popular, especially as platforms to replicate the behaviour of biological MLOs. In order to make such MLO mimics truly biomimetic, programmability and control over material properties, macromolecular partitioning and transport are key [13, 24, 52, 32]. It is for this reason that nucleic acids have been recognised as exceptonal candidates for the realisation of artificial cell platforms [53–55].

1.2 Nucleic Acids and Nucleic Acids Nanotechnology

DNA and RNA nanotechnology exploit the programmable and specific interactions found in DNA and RNA, and the versatility granted by the resulting structural, chemical and mechanical properties, which I will briefly discuss below.

1.2.1 Structure of nucleic acids

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are among the most well known and extensively studied biomolecules, making up two of the three components of the "central dogma" of molecular biology (see Section 1.2.3 below), and forming the basis for life as we know it. As they are both nucleic acids, DNA and RNA display several structural similarities, but at the same time possess distinctive features. As shown in Fig. 1.2A(i), both DNA and RNA are biopolymers composed of monomers termed *nucleotides*. Each nucleotide comprises a pentose sugar, either ribose for RNA or deoxyribose – that is lacking a hydroxyl group – for DNA, bound to a hydrophobic nitrogen base also called a *nucleobase* ("Base") on the 1' carbon, as well as a phosphate group (brown) on the 5' one. Nucleotides in nucleic acids are covalently linked *via* phosphodiester bonds between the hydroxyl group on the 3' carbon of the sugar in one monomer and the 5'-end phosphate group of the following one. This creates an alternating sugar-phosphate backbone, where consecutive monomers are held together by two (C-O) ester bonds through a phosphate group, hence the name "phosphodiester" linkage. This specificity of binding locations establishes the 5'-3' asymmetric directionality [56] of nucleic acids that is key to biological processes such as replication, transcription and translation [57].

Both DNA and RNA feature a four letter nucleobase (and nucleotide) vocabulary, three of which are shared, namely adenine (A), cytosine (C) and guanine (G), whereas thymine (T) in DNA is replaced by uracil (U) in RNA, which simply lacks the methyl group on the number 5 carbon. Nucleobases are divided in pyrimidines (C, T/U) and purines (A, G), featuring one and two aromatic rings, respectively. Due to their multiple hydrogen donors and acceptors, pyrimidines and purines can interact *via* hydrogen bonds in specific arrangements, known as Watson-Crick base pairing [62]. In particular, G pairs with C, forming three hydrogen bonds, and A pairs with T/U in DNA/RNA, forming two hydrogen bonds, as depicted in Fig. 1.2A(ii-iv), respectively. Differently from DNA, RNA also frequently features non-canonical G-U wobble-pairs, which display similar stability to A-U base pairs and are presented in Fig. 1.2A(v).

This base pairing mechanism serves as the foundation of nucleic acid hybridisation, which results in their double helical structure, first discovered for duplex DNA by Watson, Crick, Wilkins, Franklin and Gosling [62, 63]. Due to the directionality of nucleic acids, it follows that two strands can hybridise only if they are reverse complementary, *i.e.* if their nucleobases can hydrogen bond in anti-parallel orientation, one sequence read from 5' to 3' and the other from 3' to 5'. This produces a double-stranded structure, with the anionic phosphate groups on the outer surface, and the base paired nitrogen bases on the inside channel [63]. On top of specific base pairing interactions, duplex DNA and RNA are further stabilised by base stacking, that is $\pi - \pi$ stacking of the hydrophobic aromatic rings within consecutive nucleobases along the helical axis, *i.e.* orthogonally to the nucleobases themselves [64]. In order to maximise the $\pi - \pi$ overlap and the Van der Waals interactions among them,

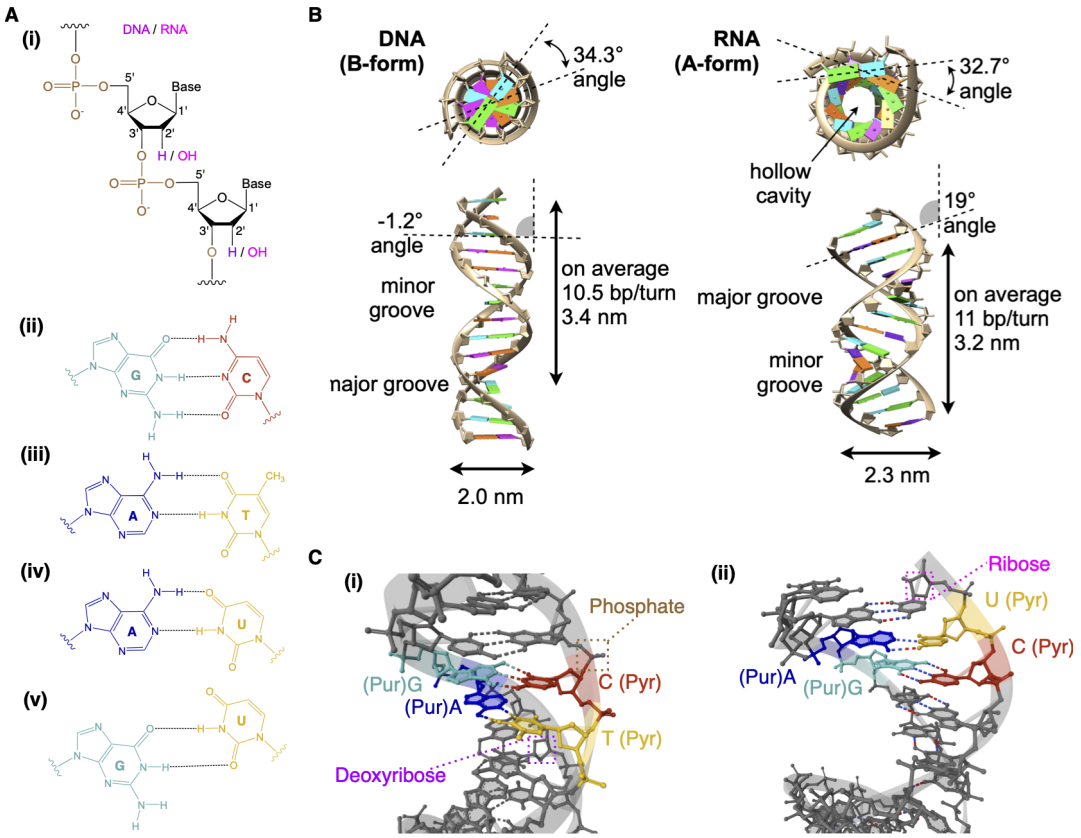


Fig. 1.2 **Chemical structure of DNA and RNA. A: (i)** Chemical structure of the alternating sugar-phosphate backbone in nucleic acids. Phosphate groups are coloured in brown, and the difference in hydroxyl groups between deoxyribose (DNA) and ribose (RNA) is highlighted in purple and magenta colours, respectively. Wavy bonds on top/bottom connect to other nucleotides in the nucleic acid chain. Watson-Crick base pairing interactions are depicted between **(ii)** G-C, **(iii)** A-T in DNA and **(iv)** A-U in RNA. **(v)** Non-canonical G-U wobble-pair in RNA. Nucleobases are colour-coded as follows: A in dark blue, C in red, G in cyan, T/U in yellow. Wavy bonds in **(ii-v)** indicate binding sites to sugar units. All structures were drawn in ChemDraw. **B:** Structural differences between *B*-form DNA on the left and *A*-form RNA on the right. Reproduced from ref. [58] in accordance to CC BY-NC-ND 3.0 license. **C:** 3D structure of **(i)** *B*-form DNA and **(ii)** *A*-form RNA, rendered *via* RCSB PDB at rscb.org [59], using models 1BNA and 413D from refs. [60, 61], respectively. Representative examples of nucleobases, phosphate groups and deoxyribose/ribose sugars are highlighted using the colour-coding from **A**.

nucleobases adopt rotated and tilted arrangements which confer both duplex DNA and RNA their helical structures.

Interestingly, despite the similarities between their building blocks, base pairing [65, 66] and base-stacking [67] interactions are (slightly) stronger in RNA compared to DNA. This is linked to marked structural differences between DNA, which adopts the well-known *B*-form conformation, and RNA, which prefers the ribbon-like *A*-form. As shown in Fig. 1.2B-C, both conformations are right-handed double-helices. However, there are two major geometrical discrepancies. First, while the nucleobases are mostly centred around the helical axis in the case of *B*-form DNA, they are considerably displaced towards the major groove of the helix in *A*-form RNA, leaving a more open cylindrical core [58]. Second, the additional hydroxyl group in the ribose sugar of RNA induces a change in its endoconformation, from

C2 in *B*-form DNA to C3 in *A*-form RNA. Combined, these dissimilarities result in *B*-form DNA being narrower and taller than *A*-form RNA, with a helix diameter of 2 nm compared to 2.3 nm, and an average distance between consecutive base pairs of 0.34 nm compared to 0.28-0.29 nm. Consistently, the helical pitches resulting from rotation and tilting to maximise base stacking are also different, equal to 10.5 and 11 base pairs (bp) for DNA and RNA [64, 68, 58], respectively.
The mechanical properties of the resulting duplexes are also different. The bending stiffness of polymers is often evaluated using the persistence length, l_{P}, defined as the decorrelation length of the direction of the polymer chain. While single-stranded DNA (ssDNA) and RNA (ssRNA) both display low values of persistence length around 1-2 nm [69–72], double-stranded DNA (dsDNA) and RNA (dsRNA) are considerably stronger, with persistence lengths of about 50 nm [69] and 62-64 nm [73], respectively.

1.2.2 Thermodynamics of hybridisation

Not only have the structure and chemistry of DNA and RNA been extensively characterised, but their physico-chemical properties have been proved to be robust and highly predictable. Binding affinities between oligonucleotides are generally expressed in terms of free energy of hybridisation. We shall consider the hybridisation of two short oligonucleotides, for which the probability of partially paired intermediates is low. Under this assumption, we can reasonably study their hybridisation using a simple two-state model, where single-stranded *A* and *B* can reversibly hybridise/dissociate into/from the duplex *AB* through second-order kinetics, as per the following reaction:

$$A + B \underset{k_{\mathrm{dis}}}{\overset{k_{\mathrm{hyb}}}{\rightleftharpoons}} AB \tag{1.1}$$

where $k_{\mathrm{hyb}}/k_{\mathrm{dis}}$ are the hybridisation/dissociation rate constants, respectively. The equilibrium constant of this reaction, K_{eq}, is simply given by:

$$K_{\mathrm{eq}} = \frac{k_{hyb}}{k_{dis}} = \frac{[AB]}{[A][B]} \tag{1.2}$$

where [AB], [A], [B] are the equilibrium concentrations of duplex *AB* and single-stranded *A* and *B* oligonucleotides, respectively. The standard hybridisation free energy, indicated as $\Delta G^{\circ}_{\mathrm{hyb}}$, comprises both enthalpic, $\Delta H^{\circ}_{\mathrm{hyb}}$, and entropic, $\Delta S^{\circ}_{\mathrm{hyb}}$, contributions as per the following equation:

$$\Delta G^{\circ}_{\mathrm{hyb}} = \Delta H^{\circ}_{\mathrm{hyb}} - T\Delta S^{\circ}_{\mathrm{hyb}} \tag{1.3}$$

where T is the temperature in Kelvin. The enthalpic term stems from favourable base pairing and stacking interactions, thus contributing with a term $\Delta H^{\circ}_{\mathrm{hyb}} < 0$. Conversely, the entropic term takes into account the reduction in degrees of freedom of the two individual

A and *B* oligonucleotides when constrained to being co-localised into the double-helical conformation of *AB* [74, 75]. As such, $\Delta S^\circ_{\text{hyb}} < 0$ and $-T\Delta S^\circ_{\text{hyb}} > 0$. Therefore, if we assume $\Delta H^\circ_{\text{hyb}}$ and $\Delta S^\circ_{\text{hyb}}$ to be temperature-invariant, $\Delta G^\circ_{\text{hyb}}$ is an increasing function of temperature: while duplex *AB* is more stable at lower temperatures, thermal energy and the related entropic term prevail over base pairing and stacking at higher temperatures, favouring denaturation, also termed melting, of *AB* into the single-stranded components *A* and *B*.
Such denaturation and the thermal stability of dsDNA and dsRNA are usually characterised *via* the so-called melting temperature, T_{M}, defined as the temperature value at which half of the oligonucleotides in solution are single-stranded and half are hybridised in duplexes. First of all, we can express the free energy of denaturation, $\Delta G^\circ_{\text{den}}$, in terms of the equilibrium constant defined above using Eq. 1.4 below:

$$\Delta G^\circ_{\text{den}} = -\Delta G^\circ_{\text{hyb}} = -RT\ln\left(\frac{1}{K_{\text{eq}}\text{P}^\circ}\right) = -RT\ln\left(\frac{[A][B]}{[AB]\text{P}^\circ}\right) \tag{1.4}$$

where R is the molar gas constant and $\text{P}^\circ = 1\,\text{M}$ is a reference concentration. When $T = T_{\text{M}}$, it follows from the definition that $[A] = [B] = [AB] = [AB]_{\text{in}}/2$, where $[AB]_{\text{in}}$ is the initial duplex concentration prior to denaturation. Substituting such values in Eq. 1.4, one obtains the following:

$$\Delta G^\circ_{\text{den}}(T_{\text{M}}) = -RT_{\text{M}}\ln\left(\frac{\frac{[AB]_{\text{in}}}{2}\frac{[AB]_{\text{in}}}{2}}{\frac{[AB]_{\text{in}}}{2}\text{P}^\circ}\right) = -RT_{\text{M}}\ln\left(\frac{[AB]_{\text{in}}}{2\text{P}^\circ}\right). \tag{1.5}$$

The melting temperature can thus be expressed in terms of the free energy as:

$$T_{\text{M}} = -\frac{\Delta G^\circ_{\text{den}}(T_{\text{M}})}{R\ln\left(\frac{[AB]_{\text{in}}}{2\text{P}^\circ}\right)} \tag{1.6}$$

or, using the definition in Eq. 1.3, in terms of enthalpic and entropic contributions as:

$$T_{\text{M}} = \frac{\Delta H^\circ_{\text{den}}}{\Delta S^\circ_{\text{den}} - R\ln\left(\frac{[AB]_{\text{in}}}{2\text{P}^\circ}\right)}. \tag{1.7}$$

$\Delta G^\circ_{\text{hyb/den}}$ and T_{M} can either be obtained experimentally or computationally predicted using the Nearest Neighbour (NN) model and empirical parameters such as those derived by SantaLucia [76, 74] for DNA, and Serra and Turner for RNA [77] through differential scanning calorimetry and melting curves. Such parameters were obtained by fitting hybridisation free energy measurements for a wide range of DNA and RNA motifs to a model taking into account base pairing, base stacking between adjacent base pairs – called nearest neighbours (NNs) – as well as correction terms or penalties. In particular, the model computes the free energy of hybridisation as:

$$\Delta G^\circ_{\text{hyb}} = \sum_i n_i \Delta G^\circ_{\text{hyb,i}} + \Delta G^\circ_{\text{init}} + \Delta G^\circ_{\text{sym}} \tag{1.8}$$

where n_i and $\Delta G^\circ_{\text{hyb,i}}$ are the number of occurrences and free energy contribution (base pairing and stacking) of the i^{th} among the 10 possible NNs, and $\Delta G^\circ_{\text{init}}$ and $\Delta G^\circ_{\text{sym}}$ are initiation and symmetry penalties considering terminal base pairs (A-T or G-C) and C2 symmetry of self-complementary oligonucleotides, respectively. Importantly, such value of $\Delta G^\circ_{\text{hyb}}$ refers to standard monovalent salt concentration, that is 1 M NaCl. Empirical salt correction factors [74, 78] have been introduced to compute $\Delta G^\circ_{\text{hyb}}$ in different ionic conditions. In the case of DNA [74], for instance, the salt correction term for monovalent Na^+ ions is proportional to $-N\ln([Na^+])$, where N is the number of base pairs and $[Na^+]$ the concentration of sodium ions. It follows that T_{M} depends on the oligonucleotide length and sequence, as well as on salt concentration (both monovalent and divalent), and other buffer parameters, such as pH.
Additional correction terms have been proposed to better model mismatches, bulges and loops, particularly in RNA [77], as well as DNA/RNA hybrids [79, 80] and more complex buffer conditions including cellular-like crowding [81]. Overall, the NN model enables the sequence-dependent prediction of hybridisation free energy and melting temperature by looking up the relevant parameters in precompiled tables.

Non-canonical nucleic acids motifs

While the *B* and *A*-form double-helices are the most common multi-stranded conformations of DNA and RNA, respectively, other secondary structures have been discovered. As these alternative secondary structures heavily rely on interactions other than Watson-Crick base pairing, called non-canonical Hoogsteen interactions or base pairs [82, 83], they are broadly called non-canonical nucleic acid motifs. The most studied non-canonical DNA motifs are presented in Fig. 1.3, but the RNA counterparts exist as well. I will first focus in more detail on the most well known motif, the G-quadruplex, as it will be central to part of this thesis, and then briefly discuss i-Motifs and triple helical structures (triplexes).
Fig. 1.3A presents DNA G-quadruplexes (G4s), non-canonical secondary structures found in guanine-rich sequences. Both DNA and RNA G4s are highly heterogeneous. Regarding their stoichiometry, they can be tetramolecular, bimolecular or unimolecular, as depicted in Fig. 1.3A(ii-vi), respectively. They can also differ in strand orientation, with strands being all parallel, all antiparallel or mixed parallel/antiparallel (hybrid) with respect to the adjacent ones, as presented in Fig. 1.3A(iv-vi) for unimolecular G4s, respectively. In particular, unimolecular G4s have been found to follow the generalised sequence $G_nX_mG_nX_oG_nX_pG_n$ where $n \geq 2$ is the number of consecutive guanines in each stretch, X can be any nucleotide within the loops, including guanines, and m, o, p are the loop lengths, usually in the range 1-7 nucleotides (nt), but they can be even longer [84].
Regardless of this diversity [85], the fundamental unit of all G4s is the G-quartet, also called G-tetrad, a semi-planar arrangement of four guanines in Hoogsteen hydrogen-bonding through interactions between the N7 of one guanine and the extracyclic amino group of the

adjacent one, as depicted in Fig. 1.3A(i). The $\pi - \pi$ stacking of several quartets yields a G4. As shown in Fig. 1.3A(i), G-quartets centrally coordinate a metal cation in the inner channel of the G4. To better visualise the geometry of G4s, Fig. 1.3B presents the 3D rendered structure of a tetramolecular parallel DNA G4, as this motif will be crucial to the work I will present in Chapter 3.

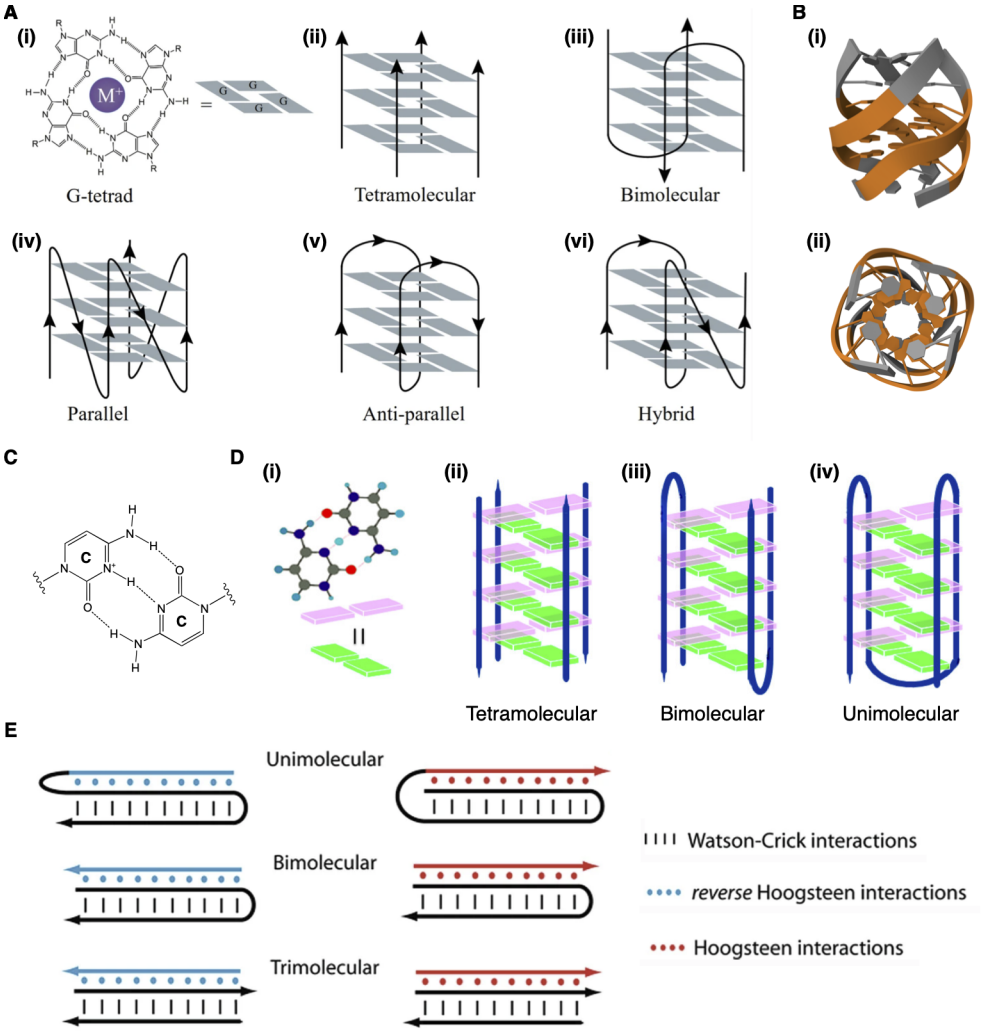


Fig. 1.3 **Non-canonical nucleic acid motifs. A:** Schematic diagrams of strand arrangements in guanine-rich DNA G-quadruplexes (G4s). **(i)** Planar assembly of Hoogsteen base paired guanines termed G-tetrad or G-quartet, coordinated with a central metal ion (M^+), accompanied with simplified sketch which is used in the remaining sub-panels. Simplified schematics of **(ii)** tetramolecular, **(iii)** bimolecular, and **(iv)** parallel, **(v)** anti- parallel and **(vi)** hybrid unimolecular DNA G4s, respectively. Adapted from ref. [86] with permission from Springer Nature Switzerland. **B:** 3D structure of tetramolecular parallel DNA G4 formed by $(T\,GGGG\,TT)_4$, in **(i)** frontal and **(ii)** top views, as rendered *via* RCSB PDB at rscb.org [59], using model 139D from ref. [87]. **C:** Chemical structure of hemi-protonated C-C^+ Hoogsteen base pair, generated using ChemDraw. **D:** Same as **A** for cytosine-rich DNA i-Motifs. **(i)** Sketch of hemi-protonated C-C^+ base pairs used to represent **(ii)** tetramolecular, **(iii)** bimolecular, and **(iv)** unimolecular i-Motifs, respectively. Adapted from ref. [88] with permission from the American Chemical Society. **E:** Same as **A** for DNA triplexes relying on either normal or reverse Hoogsteen interactions. Notation for different base pairing interactions is included on the right-hand side. Adapted from ref. [89] with permission from Springer Science Business Media. Sketches depict DNA motifs, but RNA counterparts exist.

Notably, cationic environment is paramount for the formation and the stability of G4s, with alkali metal ions, particularly K^+, and Na^+ to a lesser extent, coordinating guanines, either in-plane or sandwiched among pairs of G-quartets depending on the ionic radius, thus templating and stabilising the formation of G-quartets and their stacking into G4s [85]. It has, in fact, been computed with *ab initio* calculations that coordination with metal cations contributes to G4 stabilisation more than hydrogen bonding and stacking interactions [90]. Interestingly, different metal cations exhibit widely varying degrees of G4-templating and stabilisation. While literature reports are not in perfect agreement with one another, the general consensus is that K^+ ions are the most effective, followed by Ca^{2+}, Na^+ – particularly for certain G4 conformations – and Mg^{2+} – sometimes reported as destabilising and promoting canonical dsDNA instead. Conversely, Li^+ is neutral and does not participate in G4 stabilisation [91–93]. Other less physiologically relevant cations have been investigated as well, with some found to highly stabilise only specific G4 motifs, such as Pb^{2+} for the thrombin aptamer [94]. While G4s emerge in guanine-rich nucleic acids in specific cationic environments, i-Motifs are found in cytosine-rich sequence domains, usually at acidic pH [95–97]. This is due to the specific Hoogsteen interactions they rely on, namely hemi-protonated C-C^+ base pairing, as reported in Fig. 1.3C. Despite most i-Motifs being stable at pH < 6 (most in the range 4.2-5.2 [98]), some have been observed even at neutral pH [99]. Similarly to G4s, i-Motifs can comprise four, two or a single nucleic acid molecule, as schematically depicted in Fig. 1.3D(ii-iv), respectively.

Finally, triplexes are, as the name suggests, triple helical secondary structures in which a third oligonucleotide, called Triplex Forming Oligonucleotide (TFO), binds a Watson-Crick paired duplex, usually consisting of a homopurine (A, G) and a homopyridine (C, T/U) strands [100–102]. In particular, the TFO binds the homopurine strand at the major groove, at the so called triplex target site (TTS). Depending on its sequence, the TFO will bind either parallel (T/U, C), antiparallel (G, A) or in either way (G, T/U) to the homopurine strand, either through normal or reverse Hoogsteen interactions, as shown in Fig. 1.3E. Triplex formation can be favoured/disfavoured by tuning the number and type of triplex binding sites as well as buffer conditions, such as ionic environment and pH, as some Hoogsteen interactions require protonation of nucleobases similarly to what observed in i-Motifs.

1.2.3 Biological role of DNA and RNA

As nucleic acids are central to life, it is important to discuss their biological role, particularly as it pertains to the so called “central dogma” of molecular biology. The “central dogma” was formulated by Francis Crick during a lecture held in 1957 and published in 1958 [103], later denoted as a turning point for biology [104, 105]. Despite its name, the “central dogma” is rather an assumption, according to which the information flow in biological organisms proceeds from DNA, which acts as a genetic guidebook, to proteins, the functional macromolecules which carry out all the functions necessary for life. The intermediate step of

this information flow is RNA, which is synthesised from DNA through transcription, and later acts as template for the synthesis of proteins through translation. Other information flows are possible, namely DNA → DNA (replication), RNA → DNA (reverse transcription, viruses) and RNA → RNA (RNA-dependent RNA polymerase synthesis, viruses) [106, 107]. However, the point of the "central dogma" is that once the information encoded in DNA has reached the protein stage, it cannot back-propagate to modify the sequence of the starting genetic material.

If we compare the genetic information of DNA to a guidebook as above, then the complete guidebook is termed genome, and the specific sets of instructions, for instance those coding for specific proteins, are called genes. In particular, DNA molecules encode genetic information in the specific sequence of nucleobases attached to the directional sugar-phosphate backbone. In eukaryotic cells, DNA is found in the cell nucleus, separated from the cytoplasm by the nuclear envelope, while prokaryotes store their genetic material directly in the cytoplasm in the form of a single coiled circular DNA molecule. Interestingly, the above mentioned reverse complementarity of the two strands in duplex DNA provides an exquisitely simple mechanism for biological inheritance through DNA replication *via* templated polymerisation mediated by proteins called DNA polymerases. Through this process, cells pass the genetic information stored in their genome to the daughter cells in which they divide.

Conversely, the "central dogma" depicts RNA as a simple intermediary. Despite this apparent simplicity, though, it is now known that the vast majority of the eukaryotic transcriptome – the collection of RNA transcripts – consists of non-coding RNAs (ncRNAs) [108], that is RNA molecules that do not code for any functional proteins. Far from being unavailing, ncRNAs are involved in a plethora of biological functions, ranging from assisting protein synthesis (transfer RNA/tRNA [109, 110]), to ribosome biogenesis (ribosomal RNA/rRNA [111, 112]), gene regulation (microRNA/miRNA [113], small interfering RNA/siRNA [114], long non-coding RNAs/lncRNAs [115], riboswitches [116]) and catalysis (ribozymes) [117, 118]. Only pre-messengerRNA transcripts mature into messenger RNAs (mRNAs) [119] *via* capping of the 5' end, polyadenylation (*i.e.* addition of a poly-A tail) of the 3' end and splicing, that is the removal of non-protein coding regions called introns, and are then translated to proteins. Such functional diversity is also accompanied by a higher structural complexity compared to DNA: RNA displays in fact a tertiary structure, which involves folding of the secondary structure just like in proteins [120]. As a result, RNA combines both the programmability and simplicity of DNA and the extraordinary versatility of proteins. Due to the additional complexity, structural prediction is considerably more challenging for RNA, but recent computational approaches [121, 122], including those based on deep learning [123], are making substantial progress in both secondary and tertiary structures.

Finally, one might wonder why nucleic acids might adopt the non-canonical secondary structures presented in Section 1.2.2. Despite their extensive characterisation, the biological role of G4s has not been conclusively proved yet. Nonetheless, bioinformatic analysis

have highlighted their enrichment in gene-promoters and telomeric [124] regions – such as the highly studied $(TTAGGG)_n$ telomeric repeat in mammals [125]. This strongly hints at G4s exerting regulatory functions [126] and, as such, makes them promising therapeutic targets [127]. Similarly, albeit even less well understood, i-Motifs have been propose to to play a role in gene expression regulation [128] and in the formation of centromeres, chromosomal regions crucial to cell division [97]. The same is true for triplexes [97], which have been observed to inhibit particular DNA binding proteins [129], regulate gene expression through non-coding RNA-genomic DNA triplexes [130] and affect chromatin structure [131].

Beyond the biological role of nucleic acids: Nucleic Acid Nanotechnology

As mentioned in Sections 1.2.1 and 1.2.2 above, the specific base pairing rules governing nucleic acid hybridisation makes them incredibly programmable materials. Thanks to extensive characterisation studies, the behaviour of nucleic acids is well understood and can be rationally programmed through sequence engineering, down to changes of a single nucleotide, as well as trough external conditions such as temperature, salt concentration and pH. Moreover, nucleic acids are intrinsically biocompatible [132, 27] and versatile materials and, depending on the specific motif, display good thermal stability [133].
Nucleic acids nanotechnology thus exploits DNA and RNA as programmable building blocks with highly specific and predictable interactions, outside of their original biological context. In the following sections, I will first present the more mature field of DNA nanotechnology, and later proceed to RNA nanotechnology.

1.3 DNA Nanotechnology

Despite the similarities between DNA and RNA, several factors resulted in DNA nanotechnology advancing much faster than RNA nanotechnology. On the one hand, DNA is more resistant to degradation compared to RNA [134]. On the other hand, the decrease in DNA synthesis cost [135–137] and the development of bioconjugation chemistry techniques enabling chemically modified and functionalised oligonucleotides [138] highly contributed to the growth of the field.
While such distinction is often blurred rather than clear-cut, DNA nanotechnology is conventionally divided in two sub-fields, namely *structural* and *dynamic*, which I will present below in more detail.

1.3.1 Structural DNA Nanotechnology

In 1982, inspired by the work “Depth” by M.C. Escher, Nadrian Seeman envisioned that DNA junctions could be used to build supramolecular scaffolds to spatially arrange proteins and aid in their crystallisation [139]. In doing so, he laid the foundation of structural DNA

nanotechnology, which, in fact, aims at constructing 2D and 3D supramolecular assemblies with nanoscale precision.

While the four-way Holliday junction was already well known as a key intermediate in homologous recombination [140, 141], Seeman was the first to introduce its synthetic immobile – *i.e.* stable, not prone to branch migration (see Section 1.3.2) – counterparts [139]. Such DNA motifs became known as branched DNA or DNA nanostars, as they feature a central junction with protruding arm-like helices. In particular, Seeman designed these arms to terminate in unpaired, single-stranded overhangs, known as *sticky ends*, to enable programmable and specific interactions with other DNA nanostars. However, simple DNA nanostars were found to be too flexible to sustain the long range order required to assemble crystalline lattices [142]. Therefore, more rigid motifs were developed, such as the naturally-inspired double-crossover (DX) tiles [143] and the tensegrity triangle, depicted in Fig. 1.4A, which allowed the assembly of 2D and 3D periodic arrays and lattices [144, 145].

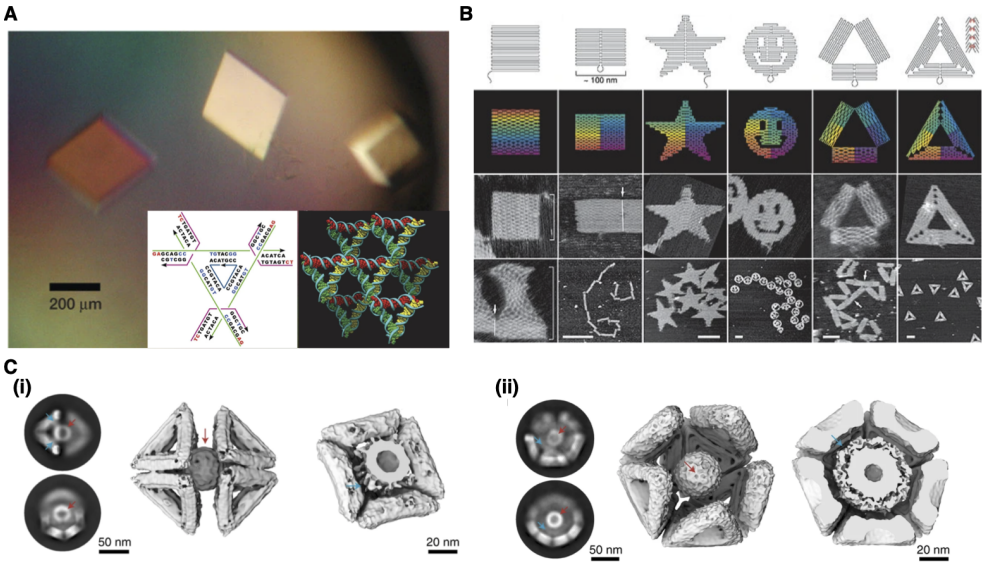


Fig. 1.4 **Structural DNA nanotechnology enables the construction of complex supramolecular assemblies with nanoscale resolution. A:** Optical micrograph depicting single DNA crystals self-assembled from tensegrity triangle subunits. Insets depict a single subunit (left) and the arragement of several interacting subunits (right). Scale bar is 200 μm. Adapted from ref. [145] with permission from Macmillan Publishers Limited. **B:** DNA origami enable the construction of near-arbitrarily shaped DNA nanostructures, as demonstrated by the programmed topologies on top and the AFM micrographs at the bottom. Scale bars are 1 μm (second column) or 100 nm (third column onwards). Adapted from ref. [146] with permission from Springer Nature Limited. **C:** Cryo-EM class averages (left, dark background) and reconstructions (right, light background) of **(i)** two half octahedral and **(ii)** one half icosahedral DNA origami shells capable of trapping hepatitis B virus (HBV) particles (spheres within the cages). Scale bars are 20 or 50 nm, as indicated. Adapted from ref. [147] with permission from Springer Nature Limited.

In 2006, Paul Rothemund introduced DNA origami [146], leading to a major breakthrough in the field of structural DNA nanotechnology. In DNA origami, a long phage-derived ssDNA, termed *scaffold*, is folded through hybridisation with hundreds of much shorter, *ad hoc* designed ssDNA strands, called *staples*, which effectively constrain the scaffold in the desired topology, as shown in Fig. 1.4B. DNA origami rapidly gained traction, and

researchers started designing increasingly complex 2D and 3D nanoarchitectures, achieving and surpassing the original goal of DNA-based lattices [148], such as large multi-scaffold designs [149] with sizes extending into the GigaDalton range [150] and cages that can trap viruses [147], as depicted in Fig. 1.4. DNA origami have also enabled researchers to develop a variety of functional nanomachines tackling nanomechanical actuation, from simple hinges and slider joints [151] to working ratchet motors [152].
A different way to unlock novel and more complex functionality in DNA nanostructures is exploiting chemical modifications enabled by bioconjugation chemistries [138]. Common functionalisations include fluorophores [55, 153, 154], proteins/enzymes [17, 155, 156] and hydrophobic moieties, ranging from lipids [157] to polymers [158, 159] and cholesterol molecules [160]. The latter are particularly interesting as they make the resulting DNA constructs amphiphilic.
On one hand, DNA amphiphiles are intriguing because they can be interfaced with synthetic and biological membranes. For instance, they have successfully been used to produce various biomimetic behaviours, from detecting lateral phase separation in membranes and mediating cargo transport [55], to replicating the behaviour of membrane-sculpting proteins through DNA origami line-actants [161]. On the other hand, amphiphilic DNA constructs combine the programmability, specificity and versatility of nucleic acids with the self-assembly robustness of natural amphiphiles, which relies on non-specific hydrophobic interactions, thus making it possible to produce a variety of interesting architectures [162]. These include discrete particles [157, 153], supramolecular assemblies [158, 159] and crystalline frameworks [160] (see Section 1.3.4 below), and exhibit a variety of functionalities, from controlled disassembly [153], to molecular partitioning [163] and cargo uptake and delivery [157].

1.3.2 Dynamic DNA Nanotechnology

Dynamic DNA nanotechnology, instead, exploits the reversibility of DNA hybridisation, as demonstrated by Eq. 1.1 in Section 1.2.2, to engineer out-of-equilibrium multi-state DNA circuits and nanomachines [134]. Most dynamic DNA nanotechnology relies on the process of *strand displacement*. As shown in Fig. 1.5A, one strand (domain a^*) in a duplex D could exchange binding partner with another complementary strand, either present unpaired in solution, as S' (domain a) in the diagram, or part of another duplex. This particular strand displacement proceeds uncatalysed, *i.e.* without any thermodynamic driver: the number of base pairs formed in both scenarios is identical. This is possible thanks to a) the just mentioned reversibility of base pairing and b) a phenomenon called *edge fraying* or *breathing* [164], whereby the terminal base pairs of a duplex are more prone to opening, due to the lack of stabilising base stacking on the other side, and forming another base pair with a different strand. If this opening and exchange repeats for all adjacent base pairs within the original duplex, a process called *branch migration*, then the strand in consideration will exchange binding partners from S to S'. However, branch migration is reversible as well, *i.e.*

it can proceed in both directions, which makes such mechanism not particularly effective to sustain directed cascade reactions through multiple states.

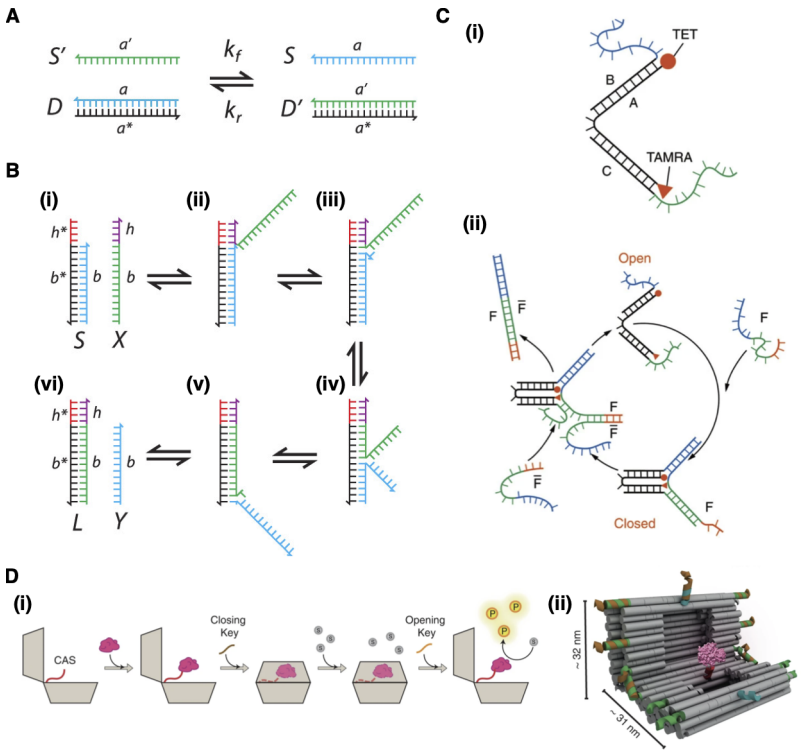


Fig. 1.5 **Toehold-mediated strand displacement and dynamic DNA nanotechnology. A:** Non-catalysed strand displacement *via edge fraying/breathing* of terminal base pairs. The number of base pairs formed is the same in both scenarios. **B:** Toehold-mediated (catalysed) strand displacement. Substrate strand is in black, with toehold in red (h^*b^*). Invader strand is in green, with toehold in purple (bh). Incumbent strand is in cyan (b). In **(i)** substrate is bound to incumbent strand. After toehold binding in **(ii)**, branch migration proceeds in **(iii-v)**. Finally, the incumbent strand is displaced in **(vi)**. **A, B** were reproduced from ref. [165] with permission from the American Chemical Society. **C:** Schematic drawing illustrating **(i)** DNA tweezers, the first dynamic DNA nanomachine relying on toehold-mediated strand displacement, and **(ii)** their working principle. Addition of fuel (F) closes the tweezers, while addition of antifuel ($\overline{\text{F}}$) re-opens them and releases the F$\overline{\text{F}}$ duplex as waste. Monitoring of actuation cycles is enabled by TET fluorescence, which is high when tweezers are open and decreases upon closing due to quenching from TAMRA. Adapted from ref. [166] with permission from Macmillan Magazines Ltd.. **D:** Control over enzymatic reactions using a reconfigurable DNA origami nanovault. **(i)** Schematic representation of the working principle: an enzyme (magenta) is tethered to a cargo-anchoring site (CAS, red). After the vault is closed using the closing key (brown), substrate molecules cannot interact with the enzyme. When the opening key (orange) re-opens the nanovault, substrate molecules interact with enzymes leading to product formation. **(ii)** 3D rendering of DNA nanovault (side locks - green, front lock - cyan). Adapted from ref. [156] in accordance to CC BY 4.0 license.

Such obstacle was overcome by introducing a thermodynamic driver, namely the formation of a larger number of base pairs, in a process called *toehold-mediated strand displacement* (TMSD), depicted in Fig. 1.5B. The substrate strand (h^*b^*) now features an unpaired single-stranded overhang (h^*), known as a *toehold*, which is absent in the current binding partner, termed the *incumbent* strand (b), but complementary to the toehold in the *invader* strand (bh). The process starts upon toehold binding (Fig. 1.5B(ii)), driven by the possibility of forming a larger number of base pairs, and proceeds through branch migration (Fig. 1.5B(iii-v)) until the incumbent strand is fully displaced (Fig. 1.5B(vi)). In particular, the toehold

overcomes the complete reversibility of branch migration, ensuring the invader strand stays in close proximity to the substrate for longer than in the absence of such toehold. The complementarity between substrate and invader does not need to be full, just more extended than the one between substrate and incumbent, thus allowing for multi-step cascade reactions through a series of discrete, progressively lower free-energy states [165]. Notably, the rate of toehold-mediated strand displacement can be tuned *via* the length of the toehold: the longer the toehold, at least until 6-7 nt [167], the faster the strand displacement reaction.
TMSD is a powerful tool, and it has enabled a wide range of complex logic circuits performing DNA-based computations [168–172], stimuli-responsive materials [173] and reconfigurable multi-state nano-devices. Some examples of the latter category are the well known DNA tweezers [166], presented in Fig. 1.5C, DNA walkers [174–176] and cargo-sorting robots [177], as well as a myriad of dynamic DNA origami [178, 179], such as the nanovault enabling control over enzymatic reactions [156] in Fig. 1.5D.

1.3.3 Non-canonical DNA motifs in DNA Nanotechnology

Another strategy to achieve dynamic reversibility in DNA nanostructures is exploiting the stimuli responsive non-canonical secondary structures introduced in Section 1.2.2 [180]. For instance, due to the crucial importance of cationic environment on G4 formation and stability, embedding DNA G4s in DNA nanostructures grants them cation-responsiveness [181–183], as exemplified by the G4-enabled actuation of a DNA origami arm depicted in Fig. 1.6A. Similarly, Fig. 1.6B proves that the inclusion of pH-sensitive i-Motifs as interaction domains between Y (three-armed) DNA nanostars [184] allows for the pH-dependent self-assembly/disassembly of DNA hydrogels. While triplexes relying on pH-dependent Hoogsteen interactions have been used to develop analogous pH-triggered nanoswitches [185] and hydrogels [186], they have also found application in synthetic gene circuits, consistently with their proposed biological role [130]. Fig. 1.6C provides an example of how triplexes can be exploited to achieve transcriptional control: the transcription of a DNA template coding for a TFO results in transcriptional suppression of another DNA template, coding for a Broccoli aptamer (see Section 1.4.2 below), which presents a TTS within the promoter region.

1.3.4 DNA nanotechnology for bottom-up synthetic biology

The described toolkit of DNA nanotechnology grants exquisite control over the designed systems and enables a wide range of biomimetic capabilities, thus making it a prime route for applications in bottom-up synthetic biology. Such applicability is further expanded by peculiar classes of nucleic acid motifs, such as *aptamers* [188, 189] (see Section 1.4.2) and catalytic DNAzymes [190], which enable a variety of biosensing and catalytic behaviours.
When it comes to structure-defining elements, considerable research effort has been devoted to the realisation of synthetic analogues of the actin cytoskeleton [191] and the cellular

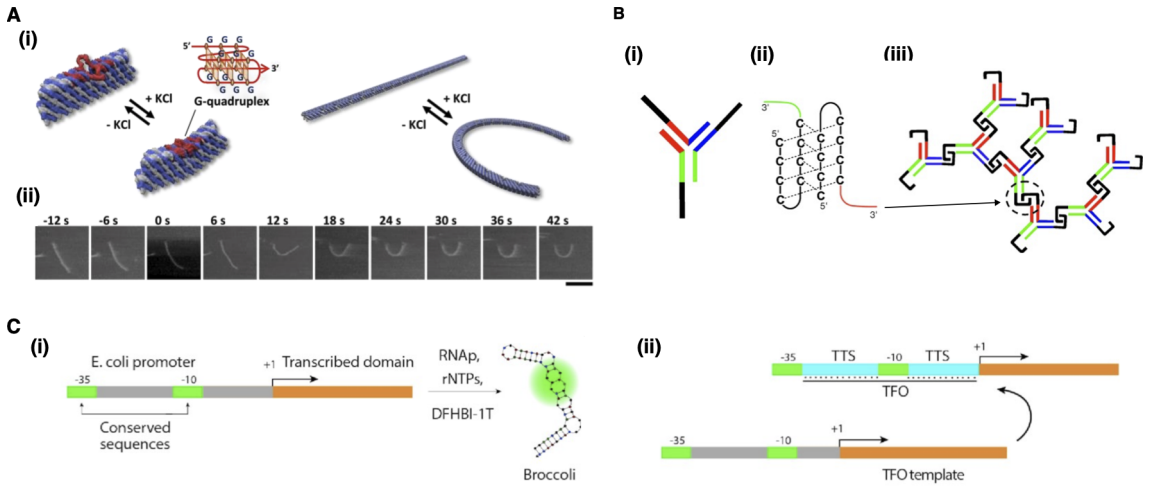


Fig. 1.6 **Non-canonical DNA motifs for stimuli-responsiveness and gene expression regulation. A: (i)** Diagram depicting the working principle for a G4-actuated DNA origami arm. Each module (left) comprises a DNA G4 motif, which enables its reversible contraction/relaxation in response to KCl addition/removal (middle). Serial concatenation of multiple modules (right) enables large deformation of the origami arm, as shown in micrographs in **(ii)**. Scale bars are 200 nm. Reproduced from ref. [187] with permission from Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. **B: (i)** Y DNA nanostars featuring **(ii)** i-Motif-forming domains at the end of each arm can reversibly assemble/disassemble **(iii)** into a DNA hydrogel. Reproduced from ref. [184] with permission from Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. **C: (i)** Design of a DNA template for transcription in *Escherichia coli* coding for a Broccoli aptamer, which becomes flurogenic upon binding DFHBI-1T (see Section 1.4.2). **(ii)** Mechanism behind triplex-enabled transcriptional suppression. The template in **(i)** is modified with two triplex target sites (TTS) in non-conserved domains of the promoter region. Transcription of a triplex forming oligonucleotide (TFO) leads to TFO binding to TTS, which competes with the polymerase for binding to the promoter, thereby suppressing transcription of the Broccoli aptamer. Reproduced from ref. [102] in accordance to CC BY 4.0 license.

inner compartments. A popular framework to mimic the former is based on fibre-like DNA nanotubes, as shown in Fig. 1.7A, which can exhibit reversible growth/disassembly within emulsion droplets [21], as well as bundling properties to enable active transport [192, 193]. Notably, one such implementation inactivated the tiles making up the DNA nanotubes and triggered their growth by integrating the system with *in vitro* transcription to produce the missing (RNA) sticky end interaction domains, which could later by degraded using RNAse H [21] granting the system high degrees of programmability and biomimicry. Other approaches rely instead on DNA-based or DNA-crosslinked hydrogels [194, 195, 49]. These can exhibit a mixture of behaviours, from cytoskeleton-like scaffolding to hosting distinct environments similarly to the cellular cytoplasm [49], as well as displaying features typical of inner cellular compartments, such as controlled pore size [196], molecular partitioning [197] and segregation of functionality in separate, addressable microenvironments [47]. Finally, some simpler cytoskeletal mimics were realised using branched DNA nanostars [198].
Interestingly, while unsuited to the formation of ordered lattices, DNA nanostars are still being extensively researched, particularly for the development of DNA-based condensates resembling synthetic MLOs [200]. Fig. 1.7B(i) presents an example of pure DNA-based condensates exhibiting programmed orthogonality thanks to orthogonal sticky end interaction

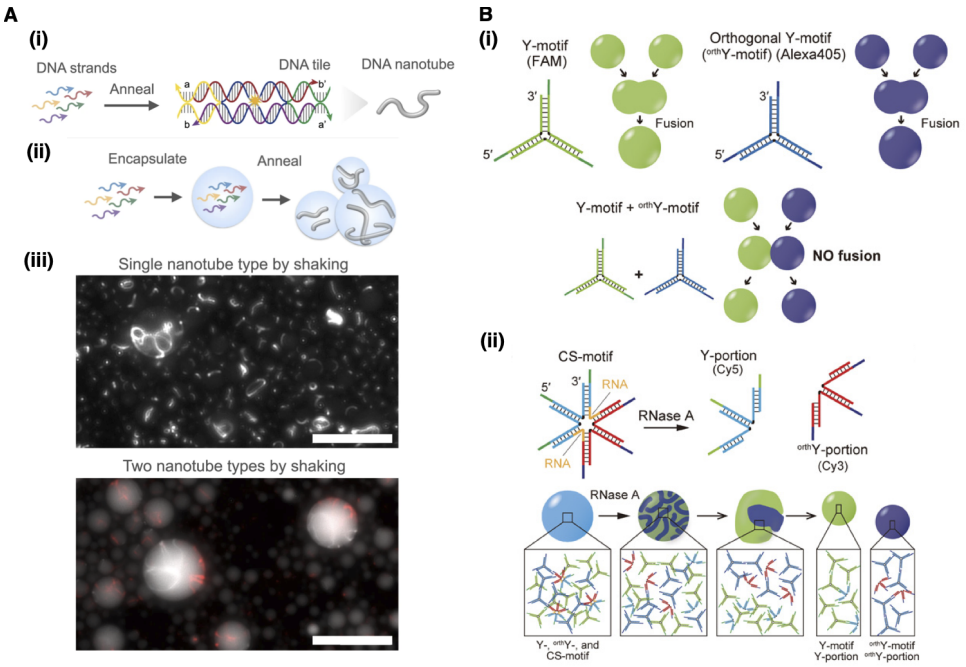


Fig. 1.7 **DNA-based cytoskeleta and condensates for artificial cells. A:** DNA nanotubes dynamically self-assemble into DNA-based cytoskeleta within emulsion droplet-based synthetic cells. **(i)** Annealing of DNA strands yields double-crossover (DX) tiles, which self-assemble into DNA nanotubes through sticky ends. **(ii)** Strands can be annealed even after encapsulation in emulsion droplets, producing a cytoskeleton comprising either **(iii)** a single type (top) or two orthogonal types (bottom) of DNA nanotubes. Scale bars are 30 μm. Reproduced from ref. [21] in accordance to CC BY 4.0 license. **B:** Orthogonal three-armed nanostars, Y and $^{\text{orth}}$Y, self-assemble into orthogonal DNA-based condensates through sticky ends. Only same-type condensates can fuse. **(ii)** Addition of a third, chimeric nanostar species, CS, leads to mixing of orthogonal condensates in **(i)**. However, CS includes RNA domains at the junction, which, when cleaved by RNAse A, separate the chimeric CS into a Y-portion and a $^{\text{orth}}$Y-portion, leading to demixing and division in orthogonal Y and $^{\text{orth}}$Y condensates as in **(i)**. Reproduced from ref. [199] in accordance to CC BY-NC 4.0 license.

domains [199]. Addition of a chimeric construct enables programmed mixing, as well as demixing upon its enzymatic degradation into two orthogonal hemi-nanostars, as depicted in Fig. 1.7(ii). Sequence engineering of the sticky end domains has also been proved to enable tuning of condensate fusion dynamics and their biophysical properties, such as surface tension and viscosity [201].

Artificial cell frameworks have also been developed using cholesterol-tagged amphiphilic DNA nanostars, termed C-Stars [160], as depicted in Fig. 1.8. Interestingly, the self-assembly of C-Stars is driven by the non-specific hydrophobic interactions among the cholesterol moieties, while the topology of the DNA nanostars simply dictates the spatial arrangement of the cholesterol clusters. Depending on the annealing rate (Fig. 1.8B), C-Stars self-assemble in either amorphous or crystalline frameworks, as depicted in Fig. 1.8C. The resulting assemblies are incredibly versatile and tolerant to design modifications, such as varying the number (3, 4 or 6 arms) or length of the nanostar arms. In particular, varying the nanostar arm length enables tuning of the lattice parameter and molecular partitioning behaviour. Including chemical modifications grants C-Star frameworks the capability to reversibly capture/release target molecules, and the inclusion of a bridge strand with a toehold enables disassembly

through TMSD [163].
C-Star condensates can also be patterned through reaction-diffusion mechanisms, as presented in Fig. 1.8D. Moreover, they can develop spatial segregation of functionalities, whereby a nucleus-like core hosts a DNA template with T7 promoter (see Section 1.4.4), which is transcribed by T7 RNA polymerase in fluorogenic Broccoli aptamers (see Section 1.4.2) which accumulate in the outer shell, as shown in Fig. 1.8E.
Beyond these mostly structural implementations, dynamic DNA circuits are increasingly being applied to synthetic biology [203] as well. In two notable examples, TMSD was exploited to implement enzyme-free proto-cellular communication strategies among proteinosomes, from detection to cascading, amplification and negative feedback [17] as well as controlled synthetic gene expression and CRISPR-Cas-based DNA processing [204]. The latter example, combined with the DNA/RNA nanotubes mentioned above [21] and the RNA-synthesising C-Stars condensates in Fig. 1.8E, hint at the considerable potential of integrating DNA nanotechnology with RNA nanotechnology to exploit the intimate links of RNA with biological machinery and naturally occuring MLOs.

1.4 RNA Nanotechnology

As previously discussed, RNA nanotechnology has historically trailed behind due to the much higher synthesis cost and sensitivity to nuclease degradation of RNA [205, 206]. The development of easy to use *in vitro* transcription systems enabling widespread enzymatic RNA synthesis from designer DNA templates and the growing interest in RNA therapeutics [207, 208] have recently led to a renewed popularity of the field. Below, I will discuss some of the advantages of RNA, as well as briefly mention both historical and state-of-the-art RNA nanotechnology approaches.

1.4.1 Advantages and motivations to pursue RNA Nanotechnology

Despite the two disadvantages mentinoed above, RNA has several tricks up its sleeve. On one hand, it benefits from higher stability at low pH, which should facilitate endosomal escape after receptor-mediated endocytosis thus making it a promising material for nucleic acids-based therapeutics [205]. Moreover, chemical modifications, such as replacement of the 2'-OH with 2'-Fluorine, can dramatically improve its thermal stability and resistance to nuclease degradation [205, 206] while preserving the *A*-form helical structure of dsRNA.
On the other hand, as mentioned in Section 1.2.1, RNA features a variety of non-canonical interactions, which lead to an impressively rich and diverse gallery of structural RNA motifs, some of which are presented in Fig. 1.9A. Frequent unpaired bulges and loops enable hairpins, 90°-kink (also called right-angle, RA) motifs and complex three-way loops. In particular, an interesting type of loop-loop interactions can form between two hairpins displaying Watson-

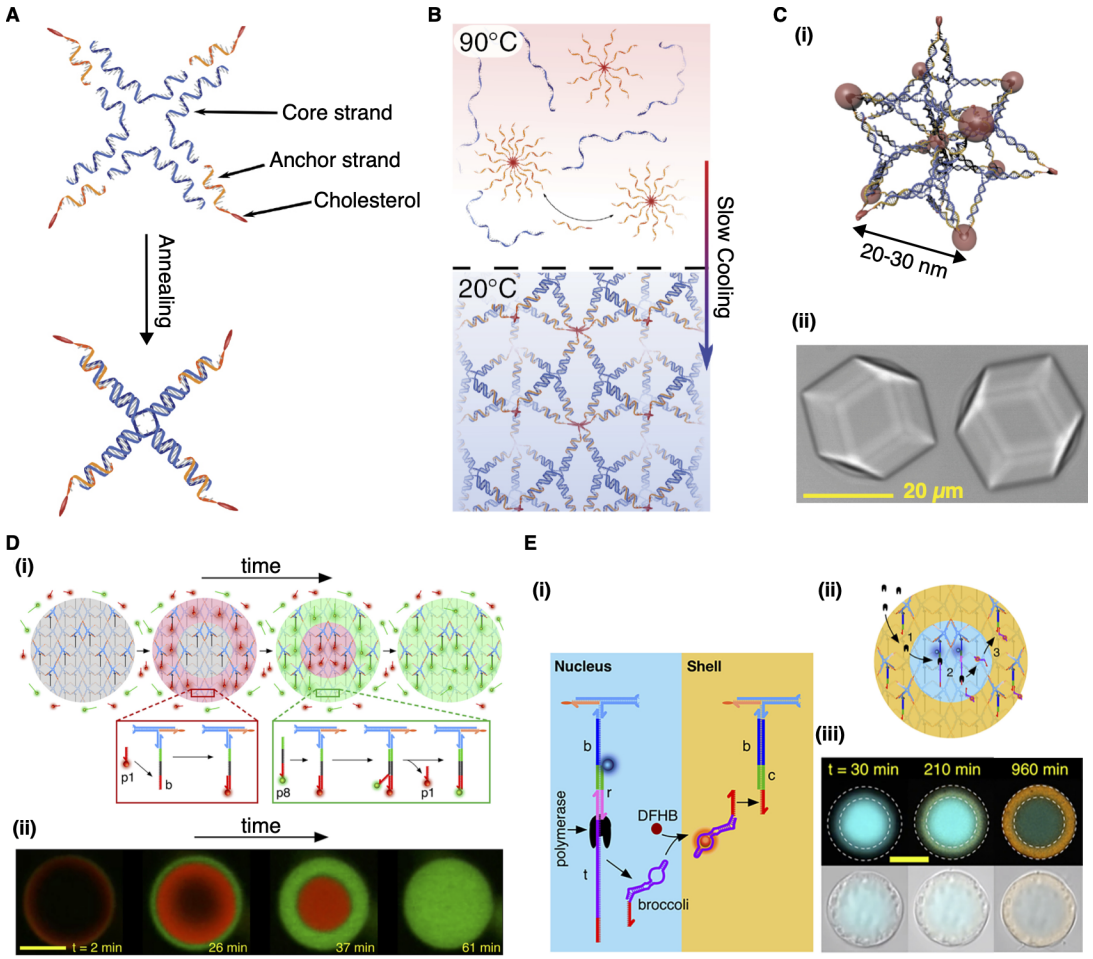


Fig. 1.8 **Amphiphilic DNA frameworks for bottom-up synthetic biology. A:** C-Stars comprise four "core" strands (blue) and four identical cholesterol-functionalised "anchor" strands (orange), which self-assemble into the programmed four-armed nanostar topology upon thermal annealing. **B:** Slow thermal annealing leads to the formation of C-Star crystalline assemblies. **C: (i)** Unit cell of C-Star crystalline framework, where cholesterol micelles occupy the vertices of a Body-Centred Cubic (BCC) lattice. **(ii)** Bright-field micrograph showing single rombic dodecahedral C-Star crystals. Scale bar is 20 μm. **A-C** were adapted from ref. [202] with permission from IOP Publishing, Ltd. **D:** Reaction-diffusion patterning of C-Star condensates. **(i)** Mechanism and **(ii)** representative confocal micrographs depicting the time-dependent patterning through fluorescently-tagged short (p1) and longer (p8) single-stranded oligonucleotides. Shorter oligonucleotides diffuse faster, but are later displaced by longer ones which can form more stable duplexes. **E:** RNA production in the nucleus of a DNA-based artificial cell. **(i)** T7 RNA polymerase (see Section 1.4.4) can transcribe the single-stranded template (double-stranded promoter) bound in the synthetic cell nucleus, thus producing fluorogenic Broccoli aptamers (see Section 1.4.2) which later accumulate in the shell/cytoplasm through hybridisation to complementary domains. This produces the spatial segregation of functionality depicted in **(ii)** and demonstrated by confocal micrographs in **(iii)**. Shell becomes fluorescent upon transcription and accumulation of Broccoli aptamers. Nucleus visualisation is enabled, instead, by fluorophore conjugation to the strand complementary to the DNA template. Scale bars in **D, E** are 15 μm. **D-E** were adapted from ref. [154] in accordance to CC BY 4.0 license.

Crick complementary loops, as depicted in Fig. 1.9B. Such hairpin dimer is known as a Kissing Loop (KL) complex, and it was originally discovered in the Dimerisation Initiaton Site (DIS) of the Human Immunodeficiency Virus-1 (HIV-1) genome [209–211], the sequence variants of which are presented in Fig. 1.9C. In the original KL complex, also known as 180° KL, the helices of the hairpin stems are coaxially stacked around the Watson-Crick

base paired loops [212, 213]. Conversely, a different, inverted loop-loop RNA motif, found to arise during the replication of the *ColE1* plasmid in *Escherichia coli (E. coli)*, leads to a 120°-bent KL complex. Finally, while immobile branched DNA junctions were purposefully designed, a similar RNA three-way junction (3WJ), rendered in Fig. 1.9D, was found to spontaneously form in the core of the packaging RNA (pRNA) which gears the bacteriophage phi29 packaging motor [214].

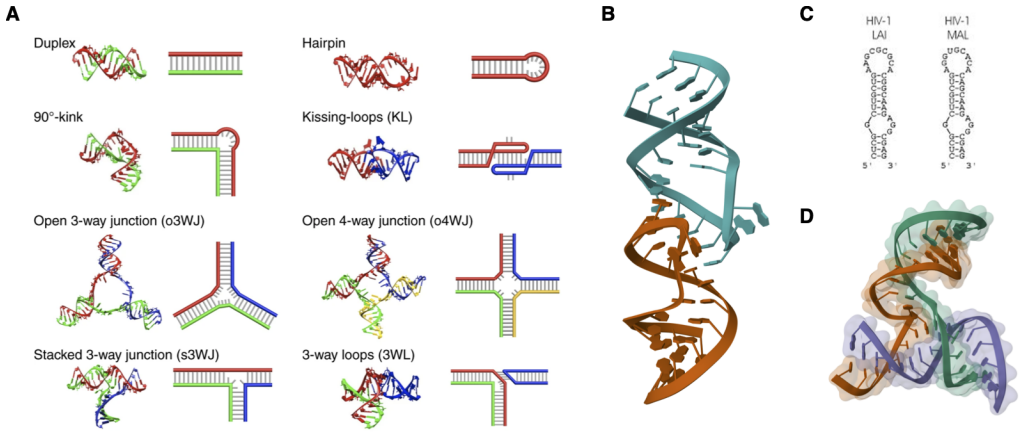


Fig. 1.9 **Structural variety of conserved RNA motifs. A:** Structural motifs found in biological RNAs and commonly adopted in RNA nanostructures, represented as 3D structures (left) and schematic diagrams (right). Solid thick coloured lines are sugar-phosphate backbones, while inner thinner grey ones are base pairs. Reproduced from ref. [215] in accordance to CC BY 4.0 license. **B:** 3D structure of the Kissing Loop complex in the Lai variant of the HIV-1 Dimerisation Initiation Site (DIS) rendered *via* RCSB PDB at rscb.org [59], using model 2B8R from ref. [210]. **C:** Sequences of KL-forming DIS variants in HIV-1 genome (Mal: African/Asian; Lai: European/American serotype). Adapted from ref. [216] with permission from Oxford University Press. **D:** 3D structure of the well known three-way junction (3WJ) formed by the core of the packaging RNA (pRNA) ring gearing the bacteriophage phi29 DNA packaging motor, rendered *via* RCSB PDB at rscb.org [59], using model 4KZ2 from ref. [214].

On top of such structural flexibility, RNA boasts an impressive functional versatility, almost comparable to that of proteins, from the high-binding affinity of aptamers (see Section 1.4.2 below) to the catalytic activity of ribozymes [117, 118, 217] and gene regulation capabilities of small interfering RNA (siRNA) [218] and riboswitches [116, 219, 220]. First, ribozymes are conserved or *de novo* designed RNA molecules displaying catalytic activity, such as the hammerhead ribozyme [221] depicted in Fig. 1.10A, which can cleave target RNA molecules which display complementarity to its arms. Second, siRNA are short, 20-25 bp-long dsRNA which emerge either from natural or synthetically programmed processing of longer dsRNA by Dicer [222]. Upon binding to proteins forming the RNA-induced silencing complex (RISC), the guide strand of the siRNA can bind to a complementary target, which is then cleaved and degraded, thus providing a convenient method for gene silencing [223], which is schematically presented in Fig. 1.10B. Finally, as illustrated in Fig. 1.10C, riboswitches [224] are gene regulatory constructs which hinder protein translation by sequestering the Ribosomal Binding Site (RBS). When specific conserved or designed RNA strands bind to the riboregulator, it undergoes a conformational switch which releases

the RBS and enables translation.

Given this incredible structural and functional versatility, it is not at all surprising that RNA plays such a central role in modern bio-engineering, biomedicine and synthetic biology [225, 226], from synthetic RNA circuits [227–230], RNA therapeutics (including the mRNA vaccines against COVID-19) [208] and biosensors based on RNA aptamers [226], which I will discuss in more detail below due to their important role in parts of this thesis.

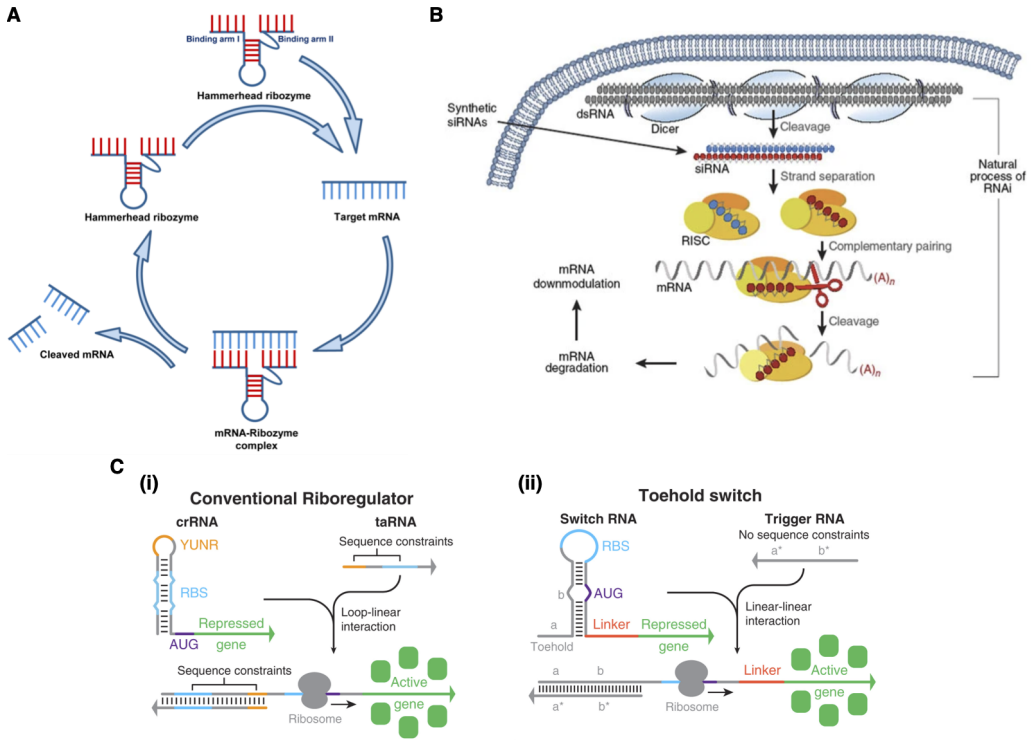


Fig. 1.10 **Functional variety of RNA motifs. A:** Schematic diagram depicting the catalytic activity of the hammerhead ribozyme. Reproduced from ref. [223] in accordance to CC BY license. **B:** Diagram illustrating RNA interference (RNAi) through small interfering RNA (siRNA) and RNA-induced silencing complex (RISC). Guide strand of siRNA binds RISC and cleaves the target mRNA. Reproduced from ref. [231] with permission from Springer Nature America. **C: (i)** Conventional riboregulators trap the Ribosomal Binding Site (RBS) and require binding to RNA with sequence constraints to open. Conversely, **(ii)** designer toehold riboswitches simply constrain within duplexes the regions before and after the start codon (AUG, in purple), and allow opening upon binding to RNA with no sequence constraints, just a complementary toehold. Reproduced from ref. [224] with permission from Elsevier, Inc.

1.4.2 RNA Aptamers

Aptamers are short nucleic acid motifs (both DNA and RNA) capable of binding particular targets, ranging from metal ions [232], to small molecules [233, 234], proteins [235, 236] and even whole cells [237]. Aptamers are usually discovered *via* a combinatorial chemistry technique known as SELEX. SELEX stands for Systematic Evolution of Ligands by EXponential enrichment [238], and selects high-affinity binders from a randomised pool of oligonucleotides through multiple stages consisting of target binding, elution and (in the case of RNA) reverse transcription. Candidate aptamers are then sequenced, and their binding

affinity and kinetics characterised. RNA aptamers have been successfully used in a variety of applications, from small molecule biosensing [239, 234, 240] *via* dual-aptamer FRET sensors [240] or split aptamer configurations [241, 242], to transcription monitoring [243] and even synthetic gene circuits inhibiting T7 RNA Polymerase (see Section 1.4.4) through the corresponding aptamer [244].
Fluorogenic RNA aptamers, also known as Fluorescence Light-Up Aptamers (FLAPs), are a particular class of RNA aptamers capable of binding dye molecules which, unbound, tend to be non fluorescent, but undergo considerable fluorescence emission enhancement upon binding to the cognate aptamers [245, 246]. They were originally developed as RNA mimics of fluorescent proteins to tag cellular RNAs [247]. Several FLAPs have been discovered, many of them binding to target ligands through non-canonical base triplets, quartets and G4s [248, 245]. Interestingly, even simple modifications of the target ligands result in large enough spectral emission shifts for multiplexed use [245].
In particular, two commonly used FLAPs, which will be used in Chapters 4-6 of this work, are the Malachite Green aptamer and the Broccoli aptamer, which I will introduce below.

Malachite Green aptamer

The Malachite Green aptamer (MGA) was first evolved *via* SELEX to bind the triphenylmethane dye Malachite Green (MG) and implement gene inactivation at the RNA level through RNA-CALI (Chromophore-Assisted Laser Inactivation) [249], as MG produces reactive radical species upon laser irradiation. Therefore, exposing to strong laser irradiation a mRNA molecule including a MG-binding RNA sequence in its 5' or 3' untraslated region (UTR) would lead to cleavage of the MGA-mRNA transcript, either leaving the mRNA without its 5' cap or its 3' poly-A tail, thus preventing translation and predisposing it to degradation pathways instead.
The minimal functional sequence of MGA is depicted in its secondary structure in Fig. 1.11A. The asymmetrical internal loop in the otherwise mostly dsRNA hairpin was identified as the key binding domain. A 2.8 Å-resolution crystal structure of MGA was obtained by using a Malachite Green analog, tetramethylrosamine (TMR), which lacks rotational mobility due to the oxygen bridging and locking in planar arrangement the two dimethlyamine-bearing phenyl rings, ultimately resulting in higher affinity and improved structure resolution. Chemical structures for both MG and TMR are presented in Fig. 1.11B. The crystal structure of MGA binding TMR (red) is reported in Fig. 1.11C, with colour-coded canonical and non-canonical base pairing and stacking interactions. Interestingly, there are no base-stacking interactions between the aptamer and the ligand, only electrostatic and steric complementarity. It is also intriguing how, despite not relying on a G4 as most fluorogenic RNA aptamers, MGA still displays a non-canonical, mixed base (CGGA) quadruple (blue) located on top of the bound ligand (red), confirming the almost ubiquitous presence of quadruplex structures in FLAPs. Despite reports of non-optimal affinity for MG ($K_d = 800$ nM), others have reported better

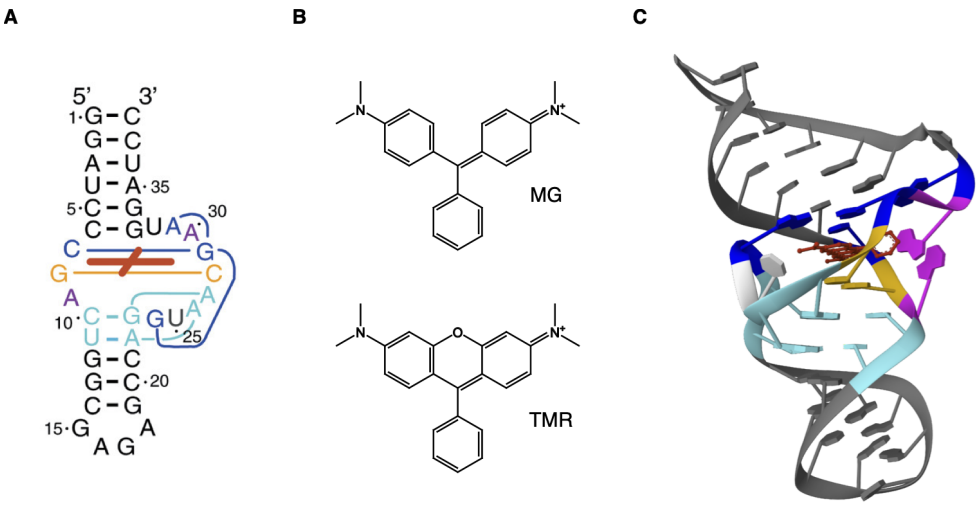


Fig. 1.11 **Malachite Green aptamer. A:** Secondary structure schematic of the minimal functional Malachite Green aptamer (MGA) evolved *via* SELEX. Annotated bases match the canonical and non-canonical base pairings coloured in **C**, while the red barred line is Malachite Green, the ligand. Adapted from ref. [250] with permission from Academic Press. **B:** Chemical structures of Malachite Green (MG) and tetramethylrosamine (TMR), produced using ChemDraw. MGA was evolved to bind MG, but its structure was resolved with TMR. **C:** 3D-rendered crystal structure of Malachite Green aptamer (*via* RCSB PDB at rscb.org [59], using model 1F1T from ref. [250]). TMR, the ligand, is coloured in red and is located in the centre of the binding pocket. The non-canonical quadruple defining the top of this binding domain is coloured in blue, while the two base-triples at the bottom are coloured in cyan. A central G-C canonical base pair is in yellow/orange, while stacked A bases are in magenta and the U base allowing for U-turn strand reversal is in light gray.

results in optimised buffer conditions ($K_d = 117$ nM) with up to $\approx$ 2000-fold fluorescence enhancement of MG bound to MGA compared to MG alone [251]. MGA and MG have been successfully used for biosensing in a dimeric split aptamer configuration [241], as well as for live cell and bacterial imaging in a multi-aptamer scaffold [252], since MG readily permeates through cell membranes. Excitation/emission maxima are around 610-630 nm/650-660 nm, respectively, making MGA suitable for multiplexing with other FLAPs in the green end of the spectrum.

Broccoli aptamer

The Broccoli aptamer (BrA) is the shortest (49 nt long) among the many RNA mimics of the Green Fluorescent Protein (GFP) that have been investigated over the years. Its secondary structure is schematically depicted in Fig. 1.12A. Similarly to the other GFP RNA mimics, BrA binds a synthetic mimic of the GFP chromophore called DFHBI [247], presented in Fig. 1.12B, and highly enhances its fluorescence emission. DFHBI and its derivatives are particularly promising due to their null to low cytotoxicity and their capability to permeate the cell membrane, making them suitable for live cell imaging [247, 253, 254]. DFHBI also displays slower photobleaching than both enhanced GFP (EGFP) and fluorescein, but it has been reported to quickly photoisomerise from *cis*- (binding ligand) to *trans*- (non binding), thus unbinding from the cognate aptamer [255, 256]. This issue can be overcome by working

in dye excess conditions, as photoconverted molecules can simply unbind, exchange, and convert back to *cis*- in the dark. Moreover, DFHBI derivatives with bulky substituents, such as the better performing DFHBI-1T [257], have been developed to considerably lessen photoconversion through steric hindrance, resulting in improved performance and higher affinity ($K_d \approx 300 - 360$ nM [253, 258], no K_d reported for DFHBI).

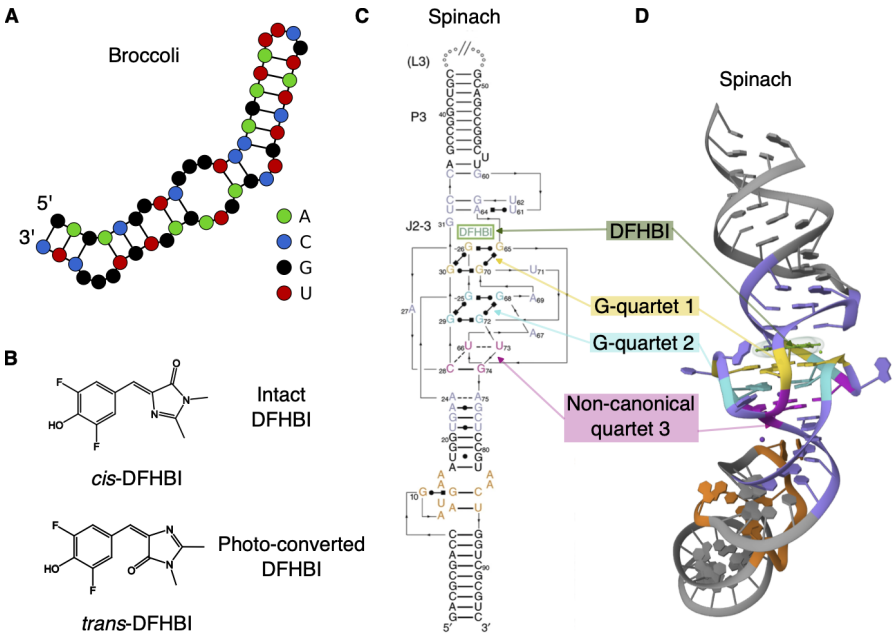


Fig. 1.12 **Broccoli and Spinach aptamers. A:** Secondary structure schematic of the minimal functional Broccoli aptamer discovered *via* SELEX-FACS evolution, as reported in ref. [253]. Colour-coding reflects base composition of the sequence. No crystal structure has been reported for Broccoli, but that of the parent aptamer, Spinach, which binds the same dye, has been characterised. **B:** Chemical structure of DFHBI (4-(3,5-difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one), the synthetic mimic of the Green Fluorescent Protein (GFP) chromophore (HBI, hydroxybenzylidene imidazolinone) bound by RNA aptamers such as Broccoli and Spinach [259]. DFHBI binds in Z (*cis*, intact) configuration, but it has been reported to photoisomerise into E (*trans*, photoconverted) [255, 256], quickly leading to unbinding from the RNA binding pocket. In dye-excess conditions, fluorescence is quickly recovered through dye exchange. Based on ref. [255], with chemical structures produced in ChemDraw. **C:** Secondary structure representation of the original DFHBI-binding RNA aptamer, Spinach, parent of Spinach2 and Broccoli [260]. Colour-coding is as follows: bases involved in G-quartets are in yellow (upper quartet) and cyan (lower quartet), those in a different non-canonical quartet in magenta, bases directly involved in interactions with DFHBI or in stabilising the DFHBI-binding pocket in light purple, other non-canonical motifs in orange. Adapted from ref. [260] with permission from Springer Nature America. **D:** 3D crystal structure of Spinach aptamer rendered *via* RCSB PDB at rscb.org [59], using model 4TS2 from ref. [260]. DFHBI is represented in light green (surrounded by a halo depicting the Gaussian volume) and annotated accordingly. Colour-coding matches the one in **C**. Small purple/green spheres are potassium/magnesium ions, respectively.

The first aptamer evolved to bind DFHBI was Spinach (96 nt long) [247], the structure of which is depicted in Fig. 1.12C-D due to the current lack of crystal structure reports for BrA. Despite its successful use for live cell imaging of intracellular RNAs, proteins and metabolites [261, 262], as well as in biosensors [263], Spinach suffers from poor folding in cellular environments, related to its low thermal stability and high dependence on Mg^{2+} ion availability. For this reason, several improved [264] and miniaturised [256] variants

have been proposed since. BrA was not explicitly derived from Spinach, rather it was evolved through a novel combination of early-stopped SELEX, fluorescence-activated cell sorting (FACS) and directed evolution in bacteria [253]. The result is a much shorter, high affinity binder capable of enhancing the fluorescence emission of DFHBI even in cellular environments thanks to improved folding, higher thermal stability and no dependence on Mg^{2+} ions.
The crystal structure for Spinach in Fig. 1.12D shows two A-form helices (light gray) stacked around a highly non-canonical binding domain. This binding pocket consists of an exquisitely intricate non-parallel three tetrad quadruplex, comprising two G-quartets (yellow, cyan), and one mixed-sequence tetrad (magenta) such as the one found in MGA, with two axially coordinated K^+ ions in between. The upper G-quartet (yellow) is perfectly planar, providing a flat hydrophobic surface for DFHBI (in light green) to $\pi - \pi$ stack on. A UAU base triple (light purple) interacts with DFHBI *via* $\pi - \pi$ stacking from the top, while an unpaired guanine residue co-planarly hydrogen bonds with the target ligand. This plethora of interactions, all found to be necessary to avoid deterioration or loss of fluorescence emission, effectively lock DFHBI in planar configuration and try to restrict its motion. Despite not being a direct Spinach derivative, the presence of shared conserved sequence stretches, common fluorescence excitation (447 nm peak)/emission (501 nm peak) signatures, and mutagenesis assays [265], strongly suggest that the binding mechanism of BrA to DFHBI might be similar. In that case, BrA might consist of an RNA G4 surrounded by flanking duplexes just like Spinach derivatives [265].

1.4.3 Assembling RNA nanostructures from modularly combined conserved motifs

Historically, RNA nanostructures have been designed according to one of three approaches [205]. The first is the direct analogue of DNA nanotechnology, and entails finding a minimal set of strands that can self-assemble into the designed architecture. Following such strategy, researchers were able to create RNA analogues of commonly used DNA nanostructures [58], such as RNA tiles assembling into RNA lattices [266] and nanotubes [267] capable of targeted delivery of siRNA [268]. The second approach is purely computational, using tools such as Kinefold [269], and involves scouring through structural databases [270–272] to optimise sequence requirements and resulting topologies [273, 271]. Finally, the third and most peculiar method relies on the modular composition through RNA-specific loop-loop interactions or similar of conserved, mostly biologically-derived domains, found in the same structural databases as the previous approach. Some nanostructures designed using this methodology are presented in Fig. 1.13.
The simplest example consists in the so called RNA LEGO, small nanostructures comprising two hairpins connected by a short single-stranded stretch which interact with identical or

different constructs through Kissing Loop interactions, as depicted in Fig. 1.13A. By varying the number of interacting constructs and their loop complementarities, it is possible to obtain linear and cyclic arrangements in the shape of simple polygons. Removal of KL-stabilising Mg^{2+} ions enables the programmed disruption of such assemblies.

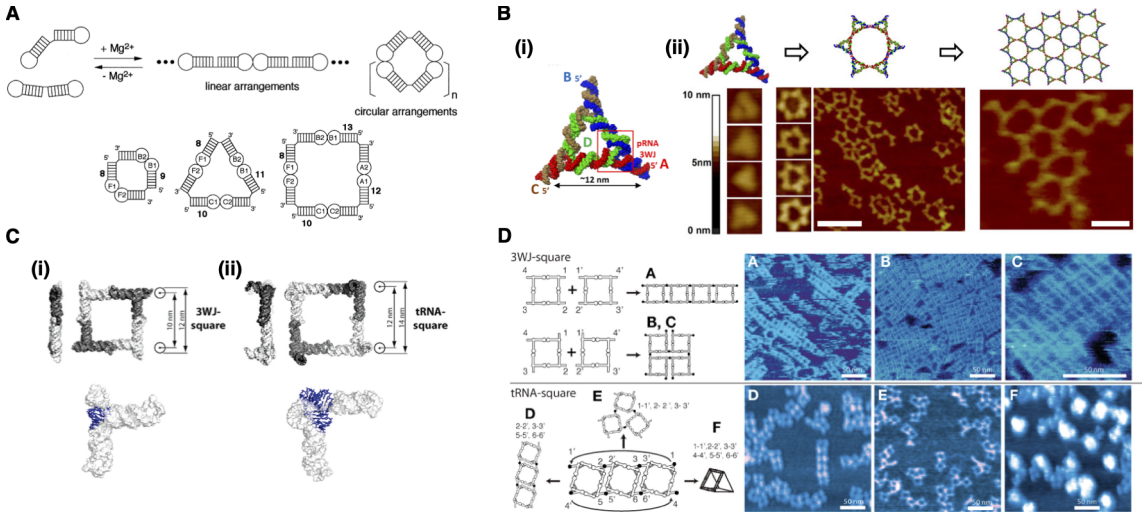


Fig. 1.13 **RNA nanostructures from the modular composition of conserved motifs. A:** RNA building blocks, termed RNA LEGO, self-assemble in linear or cyclic arrangements upon the addition of KL-stabilising Mg^{2+} ions. Different constructs and loops enable the formation of a range of simple polygonal motifs, such as those depicted at the bottom. Adapted from ref. [274] with permission from Cell Press. **B: (i)** RNA triangle nanostructures assembled from three phi29 pRNA 3WJ motifs can form **(ii)** discrete hexagons and extended 2D lattices, as demonstrated by AFM micrographs. Insets (left) are fields-of-view with 40 nm side. Scale bars are 100 nm (middle) and 35 nm (right). Adapted from ref. [275] (https://pubs.acs.org/doi/10.1021/nn5006254) with permission from the American Chemical Society (ACS), to which any further requests shall be directed. **C:** TectoRNAs: tectosquare structures derived by modularly composing **(i)** pRNA 3WJ and **(ii)** tRNA motifs (corners) through dsRNA (sides) and KL connectors, located in the middle of each side, accompanied by the 3D rendering of the corresponding conserved motif (bottom). **D:** Depending on the RNA subunit, tectosquares can assemble forming horizontally or vertically-aligned lattices as well as discrete multimers, as depicted in the AFM micrographs on the right-hand side. Scale bars are 50 nm. **C, D** were adapted from ref. [276] with permission from the American Chemical Society.

Fig. 1.13B shows a more complex implementation of the same underlying concept: the conserved domain is here the phi29 pRNA 3WJ motif, introduced in Fig. 1.9D, but the helices at its vertices have been extended to enable the formation of a triangular nanostructure out of three interacting 3WJ constructs. As demonstrated in Fig. 1.13B(ii), such triangles can further interact forming discrete hexagons or extended 2D lattices that can be made exceptionally thermally resistant through 2'-F chemical modification. This approach can be scaled to different geometries in both 2D (squares, pentagons, hexagons) [277] – which can be even interconverted [133] – and 3D (tetrahedra, prisms) [278, 279] to increase the multivalency of the RNA nanostructures, thus allowing for different functional RNA domains to be attached, such as FLAPs, siRNA or ribozymes [275, 279, 207].

Finally, Fig. 1.9C-D presents the simple, yet powerful approach of TectoRNAs (or RNA tectonics), which simply connects RNA non-canonical motifs, such as the phi29 pRNA 3WJ or the tRNA, with simple dsRNA struts connected through 180° KL domains, yielding

different kinds of nanostructures, such as the squares in Fig. 1.9C. Depending on the geometry of the specific interaction motif in their corners, the square subunits form horizontally or vertically-aligned fibre-like assemblies or discrete multimers with tilted interacting monomers, as shown in Fig. 1.9D. However, all the discussed RNA nanostructures still require thermal annealing just like most DNA nanostructures [280], thus failing to fully exploit the capabilities of RNA.

1.4.4 Transcription and co-transcriptional folding

In the quest to improve the biomimicry and utility of RNA nanostructures, researchers have thus turned to exploiting the related, naturally evolved biomolecular machinery, namely RNA polymerases, and processes, *i.e.* transcription and co-transcriptional folding of biological RNAs. As both topics are central to work presented in Chapters 4-6, I will discuss them below, starting from transcription using T7 RNA polymerase.

In vitro transcription and T7 RNA Polymerase

T7 RNA Polymerase (T7 RNAP) is one of the most extensively characterised DNA-dependent RNA polymerases, together with those from bacteriophages T3 [281] and SP6 [282]. First discovered in 1970 in *Escherichia coli* bacteria infected with the T7 bacteriophage [283], it consists of a single 99kDa protein subunit [284]. T7 RNAP exhibits remarkably high transcription efficiency, particularly in the optimal temperature range 37-42 °C, but also at slightly lower temperatures [285]. Mutants have also been engineered [286] to operate at higher temperatures (50-52 °C) [287].
Desirably, the transcription activity of T7 RNAP does not depend on other protein co-factors [288], and it is highly specific towards the corresponding double-stranded T7 promoter (consensus sequence of nontemplate strand: 5'-TAATACGACTCACTATA-3'), thus minimising undesired transcription products and cross-talk. Thanks to its simple structure, high activity and specificity, T7 RNAP is particularly amenable for minimal model systems, and, as such, it has been extensively characterised and has found applications including synthetic RNA circuits [289, 290] and metabolic pathway engineering [291].

Transcription Mechanism of T7 RNA Polymerase Differently from other viral polymerases such as SP6 [293], T7 RNAP can transcribe both single-stranded and double-stranded DNA templates [294–297], provided the T7 promoter region is double-stranded, as shown in Fig. 1.14A. The transcription process consists of three main stages: initiation, elongation and termination [298]. During initiation, T7 RNAP recognises the double-stranded T7 promoter region in the DNA template, binds to it, melts the DNA duplex at the start site, and begins synthesising from 5' to 3' the RNA strand complementary to the DNA template strand, traversing the latter in the 3' to 5' direction, as shown in Fig. 1.14B.

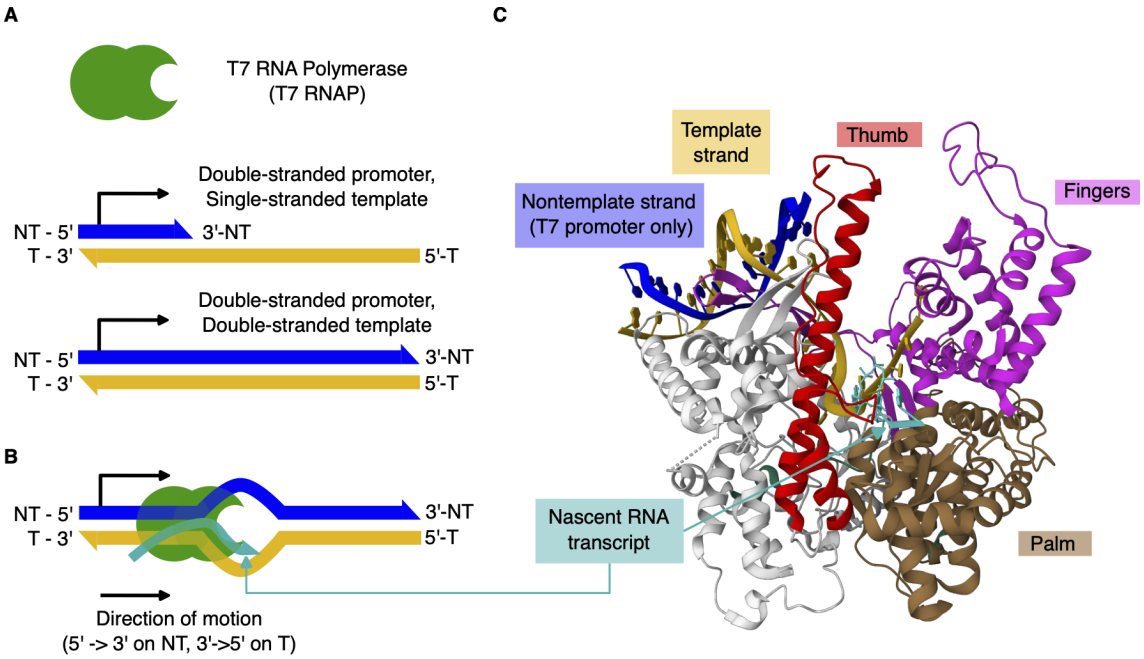


Fig. 1.14 **Mechanism and structure of T7 RNA Polymerase.** **A:** Simplified schematic of T7 RNA Polymerase (T7 RNAP) and the two possible DNA template architectures: partially and fully double-stranded, both with double-stranded T7 promoter region. Nontemplate/template DNA strands are labelled as "NT"/"T", respectively. **B:** Sketch of the transcription mechanism: after binding to the T7 promoter region, T7 RNAP melts the DNA template at the start site (here shown as a bubble in the double-stranded template), and synthesises the RNA transcript (cyan) from 5' to 3' while traversing the DNA template strand (yellow) in the opposite direction, *i.e.* from 3' to 5'. **C:** 3D structure of a transcribing T7 RNAP initiation complex rendered *via* RCSB PDB at rscb.org [59], using model 1QLN from ref. [292]. T7 RNAP is often simplistically described as a U-shaped hand grabbing the DNA template between its thumb (red) and remaining fingers (magenta), which, together with the underlying palm (brown) connecting them, define a space where the nascent RNA (cyan) grows during transcription.

In order to transition into elongation, T7 RNAP must undergo an initial synthesis phase called abortive cycling, where many short transcripts (2-10 nucleotides (nt) in length) are synthesised and released, while T7 RNAP unwinds and pulls on the downstream DNA template. The unwound template undergoes "DNA scrunching" [299, 300] within the T7 RNAP-DNA template complex. Scrunching stress accumulates progressively with the increase in length of the nascent RNA transcript, up to a length of around 14 nt. At that point, the combination of transcript length and accumulated scrunching stress highly facilitates the T7 RNAP in escaping the promoter domain [301]. Importantly, the sequence of the RNA transcript should start with a number of guanines (Gs) ranging from one to three for optimal transcription efficiency and fidelity [302–304]. Conversely, the presence of adenine (A) or uracil (U) in the first nucleotide (so called +1 position) usually results in either skipping of transcript nucleotides until a G is reached [303], or premature transcript release [305].

A different undesired phenomenon can occur during transcription initiation, called slippage or sliding [304], whereby, in the presence of at least three consecutive G-C base pairs among the nascent RNA transcript and the DNA template strand, the RNA strand slips backwards

relatively to the DNA template, triggering the addition of a further guanine residue [305]. Slippage can occur multiple times, resulting in a ladder of transcription products starting with varying lengths of 5'-poly-G stretches. However, it was found that stretches of two G-C base pairs (such as in transcripts starting with 5'-GGA) or less considerably reduce, if not prevent, slippage, limiting spurious product elongation.

Once it has escaped the T7 promoter region, the T7 RNAP-DNA template complex transitions to the elongation stage, which involves substantial conformational rearrangements of the T7 RNAP [306, 307]. T7 RNAP then proceeds to elongate the nascent RNA into the full-length product. The T7 RNAP-DNA template complex is extremely efficient (elongation rate ≈ 250 nt/s) and resilient, and it has been reported to transcribe past a variety of obstacles, including noncovalently bound proteins [288], nicks and gaps [308, 309], as well as branched junction motifs used in DNA nanostructures [310].

Finally, termination is reached when T7 RNAP reaches the end of the DNA template, or recognises a transcription terminator, a hairpin-like domain necessary for transcription to stop in circular plasmids. The RNA transcript can be released either due to co-transcriptional folding, thanks to local interactions usually resulting in the formation of hairpins, or upon a second passing of the T7 RNAP on the same DNA template, thus displacing the previous RNA transcript. Notably, despite correct synthesis and termination, transcription often results in aberrantly extended RNA transcripts. Under prolonged incubation conditions leading to high yield RNA synthesis, in fact, T7 RNAP has been shown to re-bind run-off RNA transcripts lacking a strong secondary structure at the 3' end [311], and elongate them according to self-templated additions. This usually results in spuriously longer transcripts, which exhibit a wide size distribution due to the variety of involved mechanisms, but mostly present the original 3' domains enclosed into hairpins resembling transcription terminators [284].

Finally, it is worth noting that, while T7 RNAP has been reported to transcribe correctly from both kinds of templates [294] as mentioned above, the nontemplate strand has been shown to play a role. This includes resuming elongation from a stalling state in the presence of template gaps [308], as well as ensuring efficient RNA synthesis and transcript release upon nontemplate and template strand re-annealing [312, 313]. This is particularly important for RNA transcripts featuring a complex secondary structure, such as the four-armed nanostars discussed in Chapters 4-6. The corresponding DNA template strands are the reverse complements of the DNA equivalent sequences of the RNA transcripts. As such, they will adopt a secondary structure similar to the one of the encoded RNA if left single-stranded (*e.g.* four-way branched topologies in the aforementioned scenario), even if some transcript base pairs are replaced with G-U wobble-pairs that do not correspond to valid DNA base pairs. This could result in slower or stalled transcription, as well as less efficient termination and transcript release as mentioned above. For these reasons, fully double-stranded templates will be used in Chapters 4-6.

Optimal transcription conditions Conditions for *in vitro* transcription can be optimised either by engineering the T7 RNAP itself as already mentioned, or by optimising the reaction buffer, mainly the concentration of ribonucleotides (rNTPs) and Mg^{2+} ions. While rNTPs are the building blocks for RNA synthesis, Mg^{2+} ions act as co-factors for the T7 RNAP [314–316] catalytic activity, and are needed to form the $MgNTP^{2-}$ substrate complexes for nucleotide incorporation by T7 RNAP.
In particular, it was reported [317] that optimal transcription efficiency requires high rNTP concentration (in the order of 12-40 mM), and subsaturating levels of Mg^{2+} ions compared to rNTPs. The latter condition avoids all rNTPs being available in the $MgNTP^{2-}$ substrate complex form, and thus prevents them from acting as competitive inhibitors with respect to the next rNTP to incorporate in the nascent transcript, which would ultimately stall T7 RNAP activity.
Even with these two key parameters optimised, though, transcription yield would be low without the addition of inorganic pyrophosphatase (iPPase, around 15 U/mL) to hydrolyse magnesium pyrophosphate and prevent its precipitation. Each transcription step, in fact, whereby a RNA strand of n nucleotides, $(RNA)_n$, is elongated into a $(n+1)$ nt-long construct, $(RNA)_{n+1}$, releases a pyrophosphate ion (PPi), *i.e.* $P_2O_7^{4-}$, as per the following Eq. 1.9 [314]:

$$(RNA)_n + MgNTP^{2-} \xrightarrow{\text{T7 RNAP, } Mg^{2+}} (RNA)_{n+1} + MgP_2O_7^{2-} + H^+. \tag{1.9}$$

After reaching supersaturation [316], PPi can precipitate as Mg_2PPi, *i.e.* $Mg_2P_2O_7$. As each PPi sequesters two Mg^{2+} ions, its precipitation considerably reduces the concentration of free Mg^{2+} ions in solution. This in turn leads to early stopping of transcription due to loss of activity of T7 RNAP in absence of Mg^{2+}, as well as lack of reaction substrates. Conversely, inorganic pyrophosphatase (iPPase) catalyses the following hydrolysis reaction:

$$MgP_2O_7^{2-} + H_2O \rightleftharpoons^{\text{iPPase}} 2HPO_4^{2-} + Mg^{2+} \tag{1.10}$$

thus preventing precipitation and avoiding sequestering of magnesium ions, allowing them to both maintain T7 RNAP catalytic activity, and replenish the pool of $MgNTP^{2-}$ substrates for RNA synthesis to proceed.
Common additions to *in vitro* transcription buffers further include: a) Dithiothreitol (DTT), which stabilises proteins with free sulfhydryl groups such as T7 RNAP post promoter binding; b) RNase inhibitors, which lessen RNA degradation; c) polycations such as spermine or spermidine, which facilitate the elongation phase of transcription [318, 319].

Co-transcriptional folding of RNA nanostructures

While transcription has been instrumental to RNA nanotechnology right from the start, as many RNA nanostructures were assembled from RNA strands obtained through *in vitro*

transcription, its role has become increasingly important in recent years. Once again, the inspiration came from biology. Biological RNAs, in fact, can adopt complex structures, and yet many reach such conformations through co-transcriptional folding [320], *i.e.* folding as transcription proceeds, thanks to local interactions, often resulting in hairpin formation, between already synthesised RNA domains. This insight sparked an interest in developing co-transcriptionally folding nanostructures: this in fact would have enabled their enzymatic synthesis from biologically-relevant dsDNA, rather than assembly from single-stranded components, and thus paved the way to synthesise the target nucleic acid nanostructures within living agents through molecular cloning.

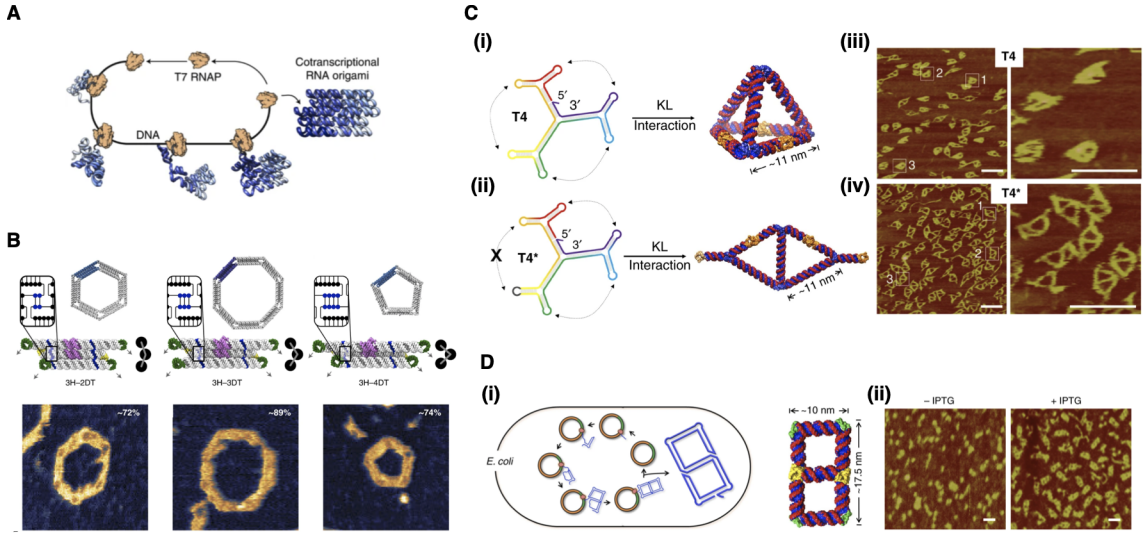


Fig. 1.15 **Co-transcriptional folding of complex RNA nanostructures. A:** Schematic representation illustrating the co-transcriptional folding of single-stranded RNA origami upon transcription of a DNA plasmid from a T7 RNA Polymerase [321]. **B:** Individual RNA origami components can be designed to interact through 120° Kissing Loop connectors [322], thus assembling into larger supramolecular architectures, such as (left) hexagonal, (middle) octagonal and (right) pentagonal rings. Curvature is tuned by changing the length of so-called dove-tail junctions (DT), nearby crossovers between three RNA helices [321, 323] Schematics (top) are accompanied by AFM micrographs (bottom) depicting 100 nm fields-of-view. **A, B** were adapted from ref. [323] with permission from Springer Nature Limited. **C:** Co-transcriptionally folding RNA nanostructures shaped as **(i)** 3D tetrahedra in the presence of all KL domains or **(ii)** flat double-triangular dumbbells with one KL domain scrambled, as demonstrated in AFM micrographs in **(iii-iv)**, respectively. Scale bars are 50 nm. **D: (i)** *In vivo* production of co-transcriptionally folding RNA double-squares through cloning of coding sequence in a plasmid followed by transformation in *E. coli*. **(ii)** AFM micrographs proving the lack/formation of the designed nanostructures in lysates from bacteria in the absence/presence of IPTG (isopropyl-β-D-1-thiogalactopyranoside) trigger, respectively. Scale bars are 20 nm. **C, D** were adapted from ref. [215] in accordance with CC BY 4.0 license.

While researchers did try to co-transcriptionally assemble multi-stranded architectures, this proved to be challenging [324] due to stoichiometry issues stemming from the different transcription rates of the DNA templates coding for the multiple strand components [295, 215, 325]. The introduction of co-transcriptionally folding single-stranded RNA origami [321, 323] signed the true paradigm shift in this research area. As shown in Fig. 1.15A, the individual origami subunits isothermally fold upon transcription, forming local interactions and exploiting inner 180° (pink) KLs to maintain structural integrity. Composing multiple

identical subunits, for instance through 120° (green) KL connectors, allows to construct even larger and more complex self-assembled architectures, both lattice-like and discrete, bypassing any stoichiometry issues, as demonstrated in Fig. 1.15B.
Such approach was soon extended to the design of simpler RNA nanostructures, such as squares, double-squares, pentagons and tetrahedra, using once again RNA-specific interaction domains such as KLs, as depicted in Fig. 1.15C. Most notably, cloning of the corresponding DNA templates in plasmids enabled expression and co-transcriptional folding of the very same synthetic RNA nanostructures in *E. coli*, as proved by micrographs in Fig. 1.15D. These recent advances pave the way for future breakthroughs in RNA therapeutics [207, 132], through *in vivo* production of the desired nucleic acids-based assemblies [326], and bottom-up synthetic biology [215], particularly regarding the expression of rationally designed and increasingly biomimetic artificial cell platforms within living agents.

1.5 Objectives and outline of this thesis

Advances in artificial cell engineering and bottom-up synthetic biology are bound to bring about breakthroughs in several technological areas, from biomedicine and tissue engineering [48] to biosynthesis and bioremediation [24, 327]. Moreover, in replicating cellular phenomenology from the bottom-up, such simplified minimal model systems will allow us to gain a much deeper understanding of biological phenomena in isolation from the surrounding biomolecular complexity, thereby leading to even further scientific and technological advancements.
Notably, since compartmentalisation is critical to maintaining distinct heterogeneous microenvironments exhibiting regulated transport and controlled macromolecular partitioning properties, engineering biomimetic compartments is crucial to advance the capabilities of artificial cells. As membrane-less approaches are rapidly gaining traction due to the renewed interest in biologically occurring membrane-less organelles (MLOs) and biomolecular condensates, nucleic acids have been recognised as prime candidates for developing biomimetic artificial cells thanks to their biocompatibility, programmability and ease of chemical modification [162]. Among the myriad of researched architectures, branched nucleic acid nanostars are simple, yet extremely versatile nanostructures which have been successfully used to build both pure [199] and amphiphilic DNA condensates [160, 163, 202] with tunable biophysical properties and macromolecular partitioning, making them promising building blocks to develop novel synthetic biology platforms.

In this Ph.D. thesis, aided by the toolkits of DNA and RNA nanotechnology, I focused on increasing the complexity and biomimicry of nucleic acids-based synthetic MLOs according to two strategies. On one end, I modularly included non-canonical DNA motifs in amphiphilic nanostars, enabling them to isothermally and reversibly self-assemble/disassemble

into hydrogel condensates in response to physiologically-relevant cations. On the other end, inspired from biological RNA-rich MLOs, I developed programmable RNA condensates with complex morphologies which self-assemble both in bulk and within synthetic cells upon transcription from designer DNA templates, *i.e.* in the presence of specific encoded information. Importantly, this thesis highlights the potential of modularity and integration of DNA and RNA nanotechnology with biologically evolved machinery to dramatically advance the biomimicry and versatility of next generation artificial cell platforms.

Chapter 2 summarises the experimental techniques and the data analysis pipelines I used and/or developed throughout this thesis.

In **Chapter 3** I investigate how the modular inclusion of non-canonical DNA G4s in amphiphilic DNA nanostars yields nanostructures responsive to physiologically-relevant cation identity and concentration. I show that these non-canonical nanostars isothermally and reversibly self-assemble/disassemble into physiologically-stable hydrogel aggregates upon addition/chelation of K^+ ions. These hydrogel particles can also uptake and release small hydrophobic molecules thanks to their intrinsic amphiphilicity, thus behaving as simplified synthetic counterparts of hydrogel-like biomolecular condensates.

Chapter 4 introduces a novel RNA-based platform. I prove how co-transcriptionally folding RNA nanostars can be transcribed from designer dsDNA templates and self-assemble through programmable Kissing Loop interactions into RNA condensates with controlled orthogonality, individual addressability and tunable morphology. This platform also benefits from straightforward fluorescence monitoring thanks to embedded fluorogenic aptamers, and direct integration with current DNA nanotechnology approaches.

In **Chapter 5**, I expand on this RNA framework by encapsulating it in emulsion droplets, producing RNA organelles in simple synthetic cells that scale in volume with the surrounding synthetic compartment similarly to cellular MLOs. I demonstrate that their relative size and number can be tuned, and that more complex morphologies, closely resembling the heterogeneity of biological MLOs, can be unlocked by controlled mixing/demixing through chimeric linker nanostars bridging otherwise orthogonal constructs.

Chapter 6 shows how simple design modifications confer RNA organelles the capability to capture target proteins. I prove that both homogeneous and non-homogeneous spatial protein partitioning profiles can be achieved. I also present an attempt at simultaneously capturing distinct proteins in orthogonal organelles, which, despite the need for future work, further substantiates the potential of this approach.

Finally, **Chapter 7** summarises the key findings of this thesis, highlighting the largely untapped potential of integrated DNA and RNA nanotechnology for bottom-up synthetic biology, and discusses future research avenues directly building on the work presented in this thesis.

Chapter 2

Materials and Methods

This Chapter summarises the materials, the experimental procedures and the data analysis methods used and/or developed for this work. First, I will discuss the design of DNA and RNA nanostructures, as well as material handling and sample preparation protocols. I will then present the characterisation techniques, followed by the data analysis methods.

2.1 Design

DNA Quad-Star design

Quad-Stars were designed using NUPACK [328] starting from C-Stars [163] with the same arm length (l = 28 base pairs (bp)). G4-forming domains were taken from literature reports [329]. The cholesterol-tagged strand was kept identical to the one used in C-Stars. All relevant oligonucleotide sequences are presented in the Appendix (Table A.1).

RNA nanostar design

RNA nanostars were designed starting from: a) RNA four-way junction and b) Kissing Loop sequences from literature [215]; c) Quad-Star arm sequences. Minimal functional sequences for RNA aptamers were obtained from literature as well: Malachite Green aptamer [330, 241], Broccoli aptamer [253], CFP/GFP/YFP (Cyan/Green/Yellow Fluorescent Protein) aptamer (sequence AP3-1) [235] and Streptavidin (STV) aptamer [236].
In particular, the minimum free energy configuration of all designs was evaluated using NUPACK (default "Serra and Turner, 1995" parameters and 1 M Na^+ salt conditions) [328]. Additionally, co-transcriptional folding of the designed nanostars was tested *via* stochastic folding simulations (Kinetic Monte Carlo [331]) performed using the Kinefold web server [269]. Stochastic RNA folding simulations rely on the assumption that single-strand diffusion and base-pair stacking/unstacking are much faster than helix folding/unfolding. As such, RNA folding is treated as a discrete series of partially-folded structures in quasi-

static equilibrium (with respect to the aforementioned faster phenomena), separated by rare (*i.e.* slow) transition events corresponding to helix formation/dissociation. Each of these transitions is associated with an Arrhenius-like rate depending on the free energy difference between the corresponding topologies (computed using the parameters by Serra and Turner mentioned above), according to which a certain event (helix addition/removal) is sampled. Since helix formation/dissociation is faster than the time it takes for the RNA polymerase to add a nucleotide to the growing RNA transcript [332], RNA can fold while transcription proceeds, *i.e.* co-transcriptionally. (The default value for T7 RNAP in Kinefold is one nucleotide added every 3 ms, corresponding to an elongation speed of ≈333 nt/s, which is higher than the consensus estimate reported in Section 1.4.4.) In co-transcriptional folding simulations, the RNA molecule of interest is synthesised (*i.e.* grows in length) over time, thus adopting different folding pathways compared to the renaturation fold of the entire molecule. While such simulations are not guaranteed to reach the minimum free energy structure due to their stochastic nature, they usually do so provided the RNA molecule is sufficiently short (<300 nt, true for all the designs discussed in Chapters 4-6) and the simulated molecular time is long enough, unless the folding trajectory gets stuck in a kinetically trapped state. Kinefold was thus used to elucidate whether the designed RNA nanostructures fold into the desired nanostar topologies or get trapped in misfolded secondary structures. This was evaluated by checking the lowest free-energy structures, the co-transcriptional folding movie and the "helix tracing graph", all provided as part of the simulation results. In particular, the "helix tracing graph" was used to check the stability of the helices within the nanostar arms (see Section 4.2). Simulations were run in batch mode (maximum computer simulation time = 1 day), selecting "Co-transcriptional Fold" and default settings (one nucleotide added every 3 ms by T7 RNAP, pseudoknots allowed, non crossing helix entanglements, automatically estimated molecular time). Due to their stochastic nature, at least three independent folding simulations (*i.e.* with different random seeds) were run *per* RNA design. Representative examples of the "helix tracing graph" obtained from co-transcriptional folding simulations of RNA nanostars A, B and C are provided in the Appendix (Fig. B.6).

Detailed explanations of the design procedure are provided in Section 4.2. RNA nanostar sequences are provided in the Appendix (Table A.2 for nanostars used in Chapters 4-5, and Table A.5 for protein-capturing nanostars used in Chapter 6).

DNA templates and primers for their PCR amplification were designed using NUPACK and the online NEB "T_m calculator", as discussed in Section 4.2.2. DNA template and oligonucleotide sequences, including primers, are provided in the Appendix (Tables A.3 and A.4 for constructs used in Chapters 4-5, and Tables A.6 and A.7 for protein-capturing constructs used in Chapter 6).

Generation of palindromic KL candidates

Once selected KL A from ref. [215], KL B and KL C were chosen among all possible palindromic loop candidates with the same length as KL A (6 nucleotides (nt)) using a Python3 script implementing the following steps. Given in input the palindromic loop length ($2N = 6$ nt) and the overall number of Gs and Cs in KL A ($2N_{GC} = 4$ bp), all possible N nt-long RNA half-loops were generated as permutations of the 4-letter nucleotide vocabulary (A, C, G, U) using itertools.product. Candidate half-loops were kept only if containing a total number of Gs and Cs amounting to $N_{GC} = 2$ nt. For each of the resulting half-loop sequences, the reverse complement was computed and appended to it, producing a $2N$ nt-long palindromic KL domain. The resulting candidates were further filtered by imposing a maximum overlap constraint (2-3 consecutive and/or identically spaced nucleotides) *via* simple string pattern matching, as discussed in Section 4.2.1.

2.2 Material handling

Saline buffers

Tris-EDTA 100× was purchased from Sigma-Aldrich and TE 1× buffer (10 mM Tris, 1 mM EDTA, pH 8.0) was prepared by simple dilution. KCl, LiCl, NaCl and $MgCl_2$ anhydrous were purchased from Sigma-Aldrich and dissolved in TE 1× buffer to produce saline buffers at the desired concentrations. Tris-Borate-EDTA (TBE) 1×, used as gel loading and running buffer, was prepared *via* dilution from TBE Buffer 10×, purchased from Thermo Scientific. All buffers were prepared fresh on a weekly basis to avoid degrading, and syringe-filtered prior to usage (Fisherbrand sterile PES syringe filters, 0.2 μm pore size, 33 mm diameter, Fisher Scientific).

Oligonucleotides

DNA oligonucleotides (also referred to as oligos/strands from here onwards) were purchased from Integrated DNA Technologies (IDT). Unless otherwise specified, all oligos were purified with standard desalting by the supplier. Functionalised strands (cholesterol-tagged "TerminalStrand_Chol", Alexa488-labelled "Arm_duplex_A_Alexa488" and biotin-conjugated "$\text{Biotin}_{\text{DNA}}$" used in Chapters 3 and 6, respectively) were instead purified *via* HPLC by the supplier. All double-stranded DNA templates were purchased as gBlocks, apart from "$\text{STV}_{\text{Apt}} - \text{T}$', which was purchased as a double-stranded, annealed and standard desalted Ultramer due to its sequence being shorter than the lower bound for gBlock synthesis.
All oligonucleotides were received lyophilised. All samples apart from PCR primers were resuspended in syringe-filtered (see Section 2.2) TE buffer 1×, usually at 100 μM concentration, following manufacturer instructions. PCR primers were resuspended at 100 μM in

nuclease-free water (UltraPure DNase/RNase-Free Distilled Water, Invitrogen) instead.
The concentration of reconstituted DNA oligos for Quad-Stars (Chapter 3), PCR primers for DNA template amplification and agarose gel purified DNA templates for RNA nanostars (Chapters 4-6) was determined by measuring A260, *i.e.* the absorbance at 260 nm, using a Thermo Scientific Nanodrop One Microvolume UV-Vis spectrophotometer. Three (for gel purified templates) or five (for all other samples) repeated measurements were performed and averaged. Unless otherwise specified, concentration in molar units was computed using the extinction coefficients provided as part of supplier specifications.
PCR primers (Chapters 4-6) were later diluted to 10 μM in nuclease-free water as per PCR kit instructions (see Section 2.4).

Proteins

Recombinant Enhanced Yellow Fluorescent Protein (EYFP) was purchased from RayBiotech (distributed by Cambridge Bioscience) at 10 μg scale, reconstituted at 0.67 mg/mL in nuclease-free water according to manufacturer's instructions and subsequently stored at -80 °C. Unfortunately, this resuspension protocol resulted in protein aggregation over time, leading to the emergence of bright fluorescent specks in all experiments apart from the first (see Section 6.3.4 and Fig. 6.5). TexasRed Streptavidin conjugate was purchased from Sigma Aldrich (CalBioChem), already resuspended at 1 mg/mL in 50 mM bicarbonate-borate buffer, 0.9% NaCl, 5 mg/ml BSA, pH 8.1, and stored at 4 °C according to specifications. Alexa Fluor 405 Streptavidin conjugate was purchased from Invitrogen at 1 mg scale, reconstituted at 1 mg/mL in Phosphate Buffered Saline pH 7.2 (Gibco) with addition of 5 mM sodium azide (using 0.1 M solution, Sigma Aldrich), as per supplier instructions, and stored at -20 °C. Work involving EYFP, TexasRed or Alexa Fluor 405 Streptavidin conjugates is presented in Chapter 6.

Dyes and other chemicals

N-methyl-mesoporphyrin IX (abbreviated as NMM), a (preferably parallel) G-quadruplex ligand used in Chapter 3 for both fluorimetry assays and irreversible photo-induced disassembly of Quad-Stars, was purchased from Cambridge Bioscience, Ltd. The as-received powder was dissolved in DMSO to prepare a stock solution at 10 mM concentration, which was stored at -25 °C. Solutions at lower concentrations were prepared *via* dilution in DMSO, and temporarily stored at 4 °C for less than one week. The chelating agent [2.2.2] cryptand (4,7,13,16,21,24-Hexaoxa-1,10-diazabicyclo[8.8.8] hexacosane), used in isothermal disassembly of Quad-Stars in Chapter 3, was purchased from Sigma-Aldrich (98 % purity). The as-received powder was dissolved in Milli-Q water, syringe-filtered (see Section 2.2) and stored at 4 °C. Rhodamine B, used as small hydrophobic cargo in Chapter 3, was purchased from Sigma Aldrich. A stock solution at 2.5 mM rhodamine B in TE 1× buffer supplemented

with 2 mM $MgCl_2$ was prepared and stored at 4 °C.
Malachite Green chloride (N,N,N',N'-Tetramethyl-4,4'-diaminotriphenylcarbenium chloride) and Malachite Green oxalate solution for microscopy (10 g/L in water), used as cognate dyes of the Malachite Green aptamer in Chapters 4-6 and 7, respectively, were purchased from Sigma-Aldrich. For the former, the as-received Malachite Green powder was dissolved in nuclease-free water (UltraPure DNase/RNase-Free Distilled Water, Invitrogen) to produce a 1 mM stock solution, which was stored at 4 °C.
DFHBI (3,5-difluoro-4-hydroxybenzylidene imidazolinone) and DFHBI-1T ((Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-1-(2,2,2-trifluoroethyl)-1H-imidazol-5(4H)-one), used as cognate dyes of the Broccoli aptamer in Chapters 4-6 and 7, respectively, were purchased from Sigma-Aldrich. In both cases, the as-received powder was dissolved in DMSO to produce a 10 mM stock solution, which was later entirely diluted to 1 mM in nuclease-free water and stored at -25 °C.
Pico-Surf fluorinated surfactant, 5% weight/weight in Novec™ 7500, and Novec™ 7500 fluorinated oil, used to produce water-in-oil emulsion droplet-based synthetic cells in Chapters 5-6, were purchased from Sphere Fluidics Ltd., and a 2% weight/weight Pico-Surf stock solution was made by simple dilution.
Materials for agarose gels are summarised in Section 2.5.

2.3 DNA nanostructure preparation

Both soluble, *i.e.* non-cholesterolised, and cholesterolised Quad-Star samples were prepared by stoichiometrically mixing the required oligonucleotides in Eppendorf tubes at final $[Oligo]_0 = 20\,\mu$M (with $[Oligo]_0$ being the starting oligonucleotide concentration), or equivalently [Quad-Star] = 5 μM. $[Oligo]_0$ was reduced to 5.6 μM for imaging purposes only.
When preparing cholesterolised Quad-Star samples, the cholesterol-bearing "TerminalStrand_Chol" strand was heated at 90 °C for around 3-5 min prior to mixing with other strands and buffer to break apart eventual aggregates. In order to make Quad-Stars soluble, the above mentioned "TerminalStrand_Chol" strand was replaced with "TerminalStrand_NoChol", which features identical nucleotide sequence but lacks the cholesterol moiety, thus preventing aggregation.
Resulting strand mixtures were annealed in a Bio-Rad C1000 Touch Thermal Cycler according to the following protocol: incubation at 95 °C for 3 min, slow cooling from 90 °C to 4 °C at $-0.25\,°\text{C}\,\text{min}^{-1}$, and a final incubation at 4 °C for 8-10 hours prior to usage. Annealed samples were stored at 4 °C and used within four days.

2.4 RNA nanostructure preparation

PCR Amplification of DNA templates

All gBlock DNA templates were amplified *via* Polymerase Chain Reaction (PCR), using the Q5 High-Fidelity DNA Polymerase from New England Biolabs. PCR mixtures were prepared on ice following manufacturer instructions, using 10 ng of DNA template per mixture, and annealed in a Bio-Rad C1000 Touch Thermal Cycler according to the following manufacturer protocol: pre-heating at 98 °C, initial denaturation at 98 °C for 30 s, 30 cycles amplification, final extension at 72 °C for 2 min, hold at 4 °C. Each cycle of amplification comprised the following steps: 98 °C for 10 s, 63-65 °C for 20 s (depending on primer melting temperature calculated using the online NEB "T_m calculator"), 72 °C for 7 s (stemming from recommended 20–30 s per kilobase (kb)). After annealing, PCR mixtures were stored at 4 °C and gel purified within one week.

Agarose gel purification of PCR products

See Section 2.5 for agarose gel electrophoresis settings. Gel bands corresponding to the expected construct size were cut using a scalpel under the UV illumination of the Syngene G:BOX Chemi XRQ gel documentation system using a UV-safe visor.
As multiple PCR mixtures were run for each DNA template to increase the final yield of purified DNA templates for downstream experiments, multiple bands were cut per sample, one per lane. The extracted gel bands were loaded in pairs in 2 mL Eppendorf tubes, treated with Monarch Gel Dissolving Buffer (New England Biolabs) at 4 μL per mg of gel band, and incubated at 50 °C until complete dissolution. The resulting mixtures were purified using the Monarch PCR & DNA Cleanup Kit (New England Biolabs), following manufacturer instructions for double-stranded DNA shorter than 2 kb. Finally, elution was performed with 12-14 μL per mixture, *i.e.* gel band pair.

Quantification of purified PCR amplified DNA templates

The concentration of purified DNA templates was determined as described in Section 2.2. Concentrations in ng/μL were converted to μM units using the Biomath Calculators by Promega, which implements the following calculation:

$$[\mathrm{DNA}]_{\mu\mathrm{M}} = \frac{10^6\,\mathrm{pg}}{1000\,\mathrm{ng}}[\mathrm{DNA}]_{\frac{\mathrm{ng}}{\mu\mathrm{L}}}\frac{1}{660\,\frac{\mathrm{pg}}{\mathrm{pmol}}\,\mathrm{N_{bp}}} \tag{2.1}$$

where N_{bp} is the number of base pairs in the double-stranded DNA template, and 660 $\frac{pg}{pmol}$ is the average molecular weight of a nucleotide pair. The result is in $\frac{pmol}{\mu L} = \frac{\mu mol}{L} = \mu M$.

2.4.1 In Vitro Transcription: kit and protocol

In vitro transcription was carried out using the CellScript T7-FlashScribe™ Transcription Kit, following manufacturer instructions and protocols. Each transcription mixture was supplemented with final [DNA template] = 40 nM, unless otherwise specified. In experiments involving binary and ternary systems, as well as linker-mixed systems, overall DNA template concentration was kept to 40 nM, with each template added according to the prescribed composition ratios.
After mixing the required components at room temperature, transcription mixtures for gel electrophoresis samples (see Section 2.5) were incubated in a Bio-Rad C1000 Touch Thermal Cycler at 30 °C (heated lid at 35 °C) for 16-18 h prior to treatment with DNase I (CellScript T7-FlashScribe™ Transcription Kit), as discussed in Section 2.5 below.
Similarly to most optimised *in vitro* transcription kits [317], the kit in use is likely to include inorganic pyrophosphatase (iPPase). Nonetheless, confocal micrographs in Fig. 4.7B, acquired after more than 48 h from the start of transcription, show some precipitate formation, particularly visible in bright-field, possibly hinting at magnesium pyrophosphate precipitation due to time-dependent loss of activity of the iPPase. This effect is much more pronounced in bulk assays compared to those in emulsion droplets.

2.4.2 Sample preparation

Bulk sample preparation

In vitro transcription mixtures were prepared as described in Section 2.4.1. Each bulk sample consisted of one 22 μL transcription mixture – 20 μL plus 1 μL 1 mM Malachite Green and 1 μL 1 mM DFHBI, yielding a final concentration of 45.45 μM for each dye.

Encapsulation in water-in-oil droplets

In vitro transcription mixtures were prepared as described in Section 2.4.1. For protein-capturing experiments presented in Chapter 6, transcription mixture volume was increased to 23 μL to accommodate Enhanced Yellow Fluorescent Protein (EYFP) and/or TexasRed or Alexa405-conjugated Streptavidin, at a final concentration of 1.25 μM each. For TexasRed-Streptavidin capture experiments *via* biotin-functionalised oligos, $\text{Biotin}_{\text{DNA}}$ and B_{OH2}-T were added to final concentrations of 10 μM and 30 nM, respectively. The latter concentration value was chosen to match the one used for other protein-capturing experiments involving simultaneously transcribed protein-capturing RNA aptamers, where [nanostar DNA template]/[aptamer DNA template] = 3 and [total DNA template] = 40 nM.
Encapsulation of *in vitro* transcription mixtures within water-in-oil emulsion droplets – also referred to as synthetic/artificial cells/cell mimics throughout Chapters 5-7 – was achieved *via* the "shaken" emulsion method, similarly to what described in ref. [21]. Briefly, 22-23 μL

of transcription mixture were added on top of 90 μL of 2% (weight/weight) Pico-Surf in Novec™ 7500 within an Eppendorf tube. The resulting mixtures were emulsified by vortexing at 2500 rpm (revolutions per minute) for 30 s and then left to equilibrate for 1-2 min before extracting the top layer containing the droplet-based synthetic cells.

2.5 Agarose Gel Electrophoresis

Agarose gels were prepared at either 1.5 % (for Quad-Stars) or 2.0 % (for DNA templates and RNA transcripts) weight/volume (w/v) agarose (for molecular biology, Sigma-Aldrich) in TBE 1× buffer, with the addition of 1× GelRed nucleic acid gel stain (3×, Biotium) apart from denaturing RNA gels. DNA reference ladders and loading dyes were purchased from Thermo Scientific (molecular biology-grade GeneRuler Ultra Low Range DNA Ladder, Trackit 6× Loading Buffers or TriTrack DNA Loading Dye 6x). RNA Gel Loading Dye (2×) and RiboRuler High Range RNA Ladder were purchased from Thermo Scientific.
Gel running settings are specified below. Imaging was performed using either a Syngene G:BOX Chemi XRQ gel documentation system or a Syngene Dyversity 4 imager equipped with a BFS-U3-63S4M-C USB 3.1 Blackfly® S, Monochrome Camera with a 12 mm UC Series fixed focal length lens (Edmund Optics) replacing the original camera and optics.

DNA nanostructures

Wells were loaded with ≈ 750-800 ng DNA samples using loading tips and dye. Gels were run at room temperature for 90 min at 70 V (7 V cm^{-1}).

PCR amplified DNA templates for gel purification

Wells were loaded with either one entire PCR reaction (50 μL plus TriTrack DNA Loading Dye 6x, for a total of 60 μL volume, requiring large gel combs) for amplified DNA templates or 1.5 μg DNA for other samples (Biotin$_{\mathrm{DNA}}$, STV$_{\mathrm{Apt}}$-T) and ULR reference ladder. Gels were run at room temperature for 90 min at 120 V (12 V cm^{-1}).

RNA transcripts

DNA templates were *in vitro* transcribed for 16-18 h at 30 °C. The resulting RNA transcripts were treated with 2 units (U) of DNase I (T7-FlashScribe™ Transcription Kit, CellScript, 1 U/μL) per 20 μL transcription mixture at 37 °C for 1 h in a Bio-Rad C1000 Touch Thermal Cycler.
For native gels, samples were briefly placed on ice, mixed with loading dye (TriTrack DNA Loading Dye 6x), and loaded into the wells (half a reaction, or 11 μL). As for DNA template gels, 1.5 μg DNA was loaded for the ULR reference ladder.

For denaturing gels, the formamide-containing RNA gel loading dye (2x) was added to the remaining half of the DNase I-treated transcription mixtures. The resulting samples were incubated at 70 °C for 30 min in a Bio-Rad C1000 Touch Thermal Cycler, then placed on ice and loaded into the wells. 10 μL of RiboRuler High Range RNA Ladder were loaded as reference ladder.
Both native and denaturing gels were run at room temperature for 75 min at 120 V (12 V cm^{-1}).

2.6 Dynamic Light Scattering (DLS)

Background

Dynamic Light Scattering (DLS) is a characterisation technique used to probe the size distribution of sub-micron colloidal particles in suspension or polymers in solution. A monochromatic light source, typically a laser, is passed through a polariser and shone on the sample, and the scattered light is collected after a second polariser onto a photomultiplier, producing a speckle pattern. Due to the Brownian motion of the particles in suspension, the speckle pattern changes with time as the relative distances among particles vary. It follows that DLS works best for dilute samples of small particles, where mostly single scattering events can be measured in the Rayleigh regime, *i.e.* uniform scattering in all directions from particles smaller than the wavelength of the light source in use, λ. These conditions will be thus taken as assumptions for the discussion below.
In particular, the time-dependent intensity trace of the speckle pattern, $I(t)$, is recorded, and its second-order autocorrelation function is computed, defined as:

$$g^2(\tau) = \frac{< I(t)\, I(t+\tau) >}{I(t)^2} \tag{2.2}$$

where τ is the time delay at which the autocorrelation is calculated, and $< ... >$ indicates the expected value. At short delays, particle motion is negligible and autocorrelation is high, whereas at increasing delays the autocorrelation decays, faster for smaller particles and slower for larger ones. Assuming spatially-coherent polarised chaotic light source (which is the case for lasers), the Siegert relation holds as follows:

$$g^2(\tau) = 1 + |g^1(\tau)|^2 \tag{2.3}$$

where $|g^1(\tau)|$ is the first-order, field-based autocorrelation function. Under the assumptions above and considering monodisperse samples, the first-order autocorrelation function can be well modelled as a single exponential decay, leading to:

$$g^2(\tau) = \alpha \left[1 + \beta e^{-2\Gamma\tau}\right] \tag{2.4}$$

where α takes into account any experimental deviation from the ideal baseline of 1, β is a scaling factor, and Γ is the decay rate. For samples of non-interacting particles (neither colliding nor electrostatically interacting) such as diluted samples of nucleic acids in saline buffers, the decay rate can be expressed in terms of the translational diffusion coefficient characterising the Brownian motion of the particles, D_t, as:

$$\Gamma = D_\mathrm{t}\, q = D_\mathrm{t} \frac{4\pi n_\mathrm{solv}}{\lambda} sin\left(\frac{\theta}{2}\right) \tag{2.5}$$

where q is the wave vector, n_solv is the solvent refractive index and θ is the scattering angle, *i.e.* the angle at which the detector is mounted. For weakly scattering samples, such as dilute samples of proteins or nucleic acids which are mostly transparent, side scattering – *i.e.* $\theta = 90^\circ$ – is usually adopted.

Fitting the autocorrelation function, it is possible to determine the translational diffusion coefficient. Assuming thermodynamic equilibrium, D_t is proportional to the ionic mobility, μ_ion, and in turn inversely proportional to the frictional coefficient in the surrounding viscous fluid, f, as in:

$$D_\mathrm{t} = \frac{\mu_\mathrm{ion} k_\mathrm{B} T}{q} = \frac{k_\mathrm{B} T}{f} \tag{2.6}$$

where k_B is the Boltzmann constant, T is temperature, and q is the ionic charge. Combining Stokes' law for a perfect sphere travelling through a fluid with viscosity η with Equation 2.6, the particle hydrodynamic/Stokes diameter, d_H, can be determined using the Stokes-Einstein equation:

$$d_\mathrm{H} = \frac{k_\mathrm{B} T}{3\pi\eta D_\mathrm{t}}. \tag{2.7}$$

Notably, the hydrodynamic diameter is the diameter of a hard sphere featuring the same D_t as the investigated particles (including solvent effects), and is thus a good approximation only for small optically and geometrically isotropic particles.

The model in Equation 2.4 can be extended to polydisperse samples by taking into account a sum of multiple exponential decays, Γ_i for the i^th population, as in:

$$g^2(\tau) = \alpha\left[1 + \beta e^{-2\Sigma_i \Gamma_i \tau}\right]. \tag{2.8}$$

In this case, numerical fitting is usually performed using either the cumulant fit [333, 334], an inverse Laplace transform-based method such as the CONTIN algorithm [335, 336] or a non-negative least squares algorithm with regularisation [337].

It is worth noting that the assumptions above should hold for dilute (μM range) and filtered (no large dust particles) samples of small DNA nanostructures (low refractive index contrast, size $<< \lambda$) thermally annealed (hence in thermodynamic equilibrium) in 100-300 mM monovalent salt, such as those investigated in Chapter 3.

Experimental

DLS measurements were carried out on a Malvern Zetasizer Ultra equipped with a 633 nm (red) light source, using the Malvern Low Volume Disposable Sizing Cell Kit (ZSU1002) and the corresponding 1 mm^2 glass capillaries (ZSU0003). Samples of both non-cholesterolised and cholesterolised Quad-Stars were syringe-filtered prior to annealing using 4 mm, 0.2 μm pore size, PES filters (Captiva Premium, Agilent Technologies). After annealing, each capillary was filled with about 4 μL of sample, and sealed at the bottom to prevent leakage using the clay sealing provided by the supplier.
The Low Volume Disposable Sizing Cell Kit only allows size measurements from side-scattering. Three to five measurements per sample were performed and averaged. Fitting of the autocorrelation function was performed automatically by the Malvern ZS Explorer software interface, with the solvent refractive index approximated with that of water, and the differential refractive index increment with that of dilute protein solutions. The same software interface was also employed to convert size distributions by intensity into size distribution by volume – under small particle assumption, *i.e.* Rayleigh scattering, intensity is proportional to d_H^6, while volume is proportional to d_H^3.

2.7 Circular Dichroism (CD)

Background

Circular Dichroism, often abbreviated as CD, is a spectroscopy technique which measures the difference in absorption of left-hand (LHC) and right-hand (RHC) circularly polarised light. A spectropolarimeter is used, which generates linearly polarised light and converts it into circularly polarised one using a photoelastic modulator, namely a piezoelastic element fused to a block of birefringent quartz. The spectropolarimeter shines LHC and RHC-polarised light onto the sample, measures the transmitted intensities, and computes the corresponding absorbances, $A_{L/R}$, as well as the differential absorbance, $\Delta A = A_L - A_R$. Such difference in absorbance stems from an asymmetry in molar extinction coefficients, namely $\varepsilon_L \neq \varepsilon_R$ for LHC and RHC-polarised light, which produces a non-zero molar circular dichroism (or differential molar absorptivity), defined as

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R = \frac{\Delta A}{C\,l} \tag{2.9}$$

where C is the concentration of the investigated substance in molar units, and l is the cell pathlength in cm.
It follows that CD is a characteristic feature of asymmetric, *i.e.* chiral, molecules. In particular, ultraviolet (UV) CD has historically been used to probe the secondary structure of macromolecules such as proteins and nucleic acids. In the latter case, the chirality stems

mostly from the relative orientation among phosphodiester-bonded nucleobases stacked in a helix, but also from asymmetric electronic transitions of individual nucleobases coupled to chiral ribose/deoxyribose sugars in RNA/DNA, respectively [338]. The resulting molar CD is positive for right-handed helices and negative for left-handed ones. Similarly, CD measurements can distinguish between α-helix-rich, β-sheet-rich or unfolded proteins.
On the molecular level, circular birefringence – or optical rotation –, whereby the velocity of light passing through the sample depends on its polarisation, stems from the combination of induced electrical and magnetic dipole moments in the measured macromolecules [339]. Such dipole moments are either in or out-of-phase depending on the light polarisation, and this ultimately results in the non-zero differential molar absorptivity mentioned above.
In the case of non-optically active materials, where $\varepsilon_{\mathrm{L}} = \varepsilon_{\mathrm{R}}$, equal intensities of incident LHC and RHC-polarised light result in equal transmitted intensities, which superimpose into linearly polarised light. In materials exhibiting circular birefringence, instead, equal intensities of incident LHC and RHC-polarised light are differentially absorbed and produce unequal transmitted intensities, which superimpose into elliptically polarised light. The ellipticity of polarisation, θ (in radians), is thus defined as:

$$tan\,(\theta) = \frac{E_{\mathrm{R}} - E_{\mathrm{L}}}{E_{\mathrm{R}} + E_{\mathrm{L}}} \tag{2.10}$$

where $E_{\mathrm{R/L}}$ is the amplitude of the electrical field vector in the RHC/LHC-polarised wave, respectively. The wavelength-dependent variation of this ellipticity of polarisation, historically reported in units of degrees (deg) or millidegrees (mdeg), defines the CD spectrum of a substance.
As the magnitude of CD is generally small, so is the resulting ellipticity, and it is possible to approximate $tan\,(\theta) \approx \theta$, thus leading to:

$$\theta \approx \frac{I_{\mathrm{R}}^{\frac{1}{2}} - I_{\mathrm{L}}^{\frac{1}{2}}}{I_{\mathrm{R}}^{\frac{1}{2}} + I_{\mathrm{L}}^{\frac{1}{2}}} \tag{2.11}$$

where $I_{\mathrm{R/L}}$ is the intensity of the RHC/LHC-polarised wave. By using Beer-Lambert law, namely $I = I_0\, e^{-A\,\ln(10)}$, where I and I_0 are the transmitted and incident light intensities and A is the absorbance, one obtains the following:

$$\theta \approx \frac{e^{-A_{\mathrm{R}}\frac{\ln(10)}{2}} - e^{-A_{\mathrm{L}}\frac{\ln(10)}{2}}}{e^{-A_{\mathrm{R}}\frac{\ln(10)}{2}} + e^{-A_{\mathrm{L}}\frac{\ln(10)}{2}}} = \frac{e^{\Delta A\frac{\ln(10)}{2}} - 1}{e^{\Delta A\frac{\ln(10)}{2}} + 1} \tag{2.12}$$

where ΔA is the experimentally measured differential absorbance defined above. Under the small CD assumptions above, $\Delta A << 1$, and first-order Taylor series expansion of Eq. 2.12 yields:

$$\theta \approx \Delta A \frac{\ln(10)}{4} \tag{2.13}$$

which can be readily converted in degrees as:

$$\theta(deg) \approx \Delta A \frac{\ln(10)}{4} \frac{180}{\pi}. \tag{2.14}$$

To remove the dependence from sample concentration and cell pathlength from Eq. 2.14, CD can be reported as molar ellipticity, defined as $[\theta] = \frac{100\theta}{Cl}$, where 100 compensates for l in cm. $[\theta]$ is usually expressed in units of deg·cm^2/dmol, and can be related to molar circular dichroism by taking into account Eq. 2.9 as per the Eq. 2.15 below:

$$[\theta] \approx 100 \frac{\theta}{Cl} = 100 \frac{\Delta A}{Cl} \frac{\ln(10)}{4} \frac{180}{\pi} \approx 3298.2 \Delta\varepsilon \tag{2.15}$$

which finally provides a clear link between the molecular-level descriptor, $\Delta\varepsilon$, the experimentally measured quantity, ΔA, and the ellipticities commonly used to report CD, θ and $[\theta]$.
As briefly mentioned above, the CD spectra of nucleic acids provide valuable information on their conformation. For instance, *B*-form DNA, which is a right-handed double-helix, displays a global minimum around 250 nm, and a broad positive peak at 260-280 nm. The more compact right-handed *A*-form, commonly adopted by RNA, exhibits a positive peak around 260-270 nm and a negative one around 210 nm. Conversely, the rarer left-handed *Z*-form double-helix has a global minimum at 200 nm, a maximum around 260 nm and a negative band around 290 nm [340]. Finally, non-canonical G-quadruplexes present varying CD spectra depending on their strand orientation, with parallel quadruplexes exhibiting a shift of the maximum/minimum towards 263 nm/240-245 nm, respectively, antiparallel ones presenting a negative band around 260-270 nm and a positive peak at 290-300 nm, and hybrid architectures a mix of the CD signatures above [341, 342].

Experimental

CD measurements were performed on a JASCO Corporation J-715 spectropolarimeter. For each Quad-Star sample ($[Oligo]_0$ = 20 μM or, equivalently, [Quad-Star] = 5 μM), 200 μL were loaded in a stoppered quartz cuvette featuring 1 mm pathlength.
Spectra were acquired with a 0.5 nm bandwidth and a 1 nm data pitch between 200 and 320 nm. For each sample, three to five spectra were accumulated and automatically averaged by the spectropolarimeter software. The spectropolarimeter in use directly reported CD spectra in the form of ellipticity of polarisation (mdeg). Data were then exported in TXT format and processed in Python3 as follows. Sample spectra were background corrected by subtracting the spectra of the corresponding saline buffers (TE 1× supplied with either 300 mM KCl or LiCl). Spectra were also cut at 220 nm due to saturation of the applied voltage leading to large introduced noise. Finally, spectra were filtered using a Savitzky-Golay filter (scipy.signal.savgol_filter) with window length equal to 17 datapoints and using a 3rd order polynomial prior to plotting.

2.8 Fluorimetry

Fluorescence excitation/emission scans and kinetics assays were performed using a BMG CLARIOstar Plus microplate reader equipped with dual injectors, using UV-Star 384 well-plates purchased from Greiner Bio-One. The Extended Dynamic Range mode of the microplate reader was enabled in all kinetics assays in order to avoid signal saturation.

2.8.1 Excitation and emission scans

NMM fluorescence emission scans

NMM fluorescence emission spectra were measured in 5G/6G samples (both with 9T spacer) after a 22 h self-assembly transient (see Section 2.8.2) in the presence or absence of 300 mM KCl injection. Concentration of NMM was adjusted to be stoichiometric with respect to expected G4 concentration, namely [NMM] = [G4] = [Quad-Star] = $[Oligo]_0/4 = 5\,\mu$M. Excitation was set to 388 ± 15 nm and fluorescence emission was recorded from 550 to 750 nm with 1 nm step. Bottom optics readout was used due to the optical adhesive required by long fluorimetry runs (see Section 2.8.2). The number of flashes per well was set to 20. Gain was automatically adjusted based on the brightest well. Data are reported after background-correction by subtracting the spectrum in plain TE 1× buffer (average of duplicates). Spectra in samples with/without added KCl are averages of triplicates/duplicates, respectively.

Excitation/Emission scans on RNA nanostars with fluorogenic aptamers

Fluorescence excitation and emission scans of co-transcriptionally folded RNA nanostars containing fluorescent light-up aptamers (FLAPs) were measured after 48 h of *in vitro* transcription, *i.e.* at the end of the kinetics assays described below in Section 2.8.2. The thermal stage was set to 30 °C. As described above, bottom optics readout was used due to the optical adhesive required by long fluorimetry runs (see Section 2.8.2). Similarly, the number of flashes per well was set to 20. Gain was automatically adjusted based on the brightest well in each channel. When scanning either excitation or emission, the resolution/step was set to 1 nm.

When collecting fluorescence excitation (Exc.) scans for DFHBI+BrA (Broccoli aptamer), fluorescence emission was recorded in the band 507 ± 8 nm while scanning excitation in the range 400-486 nm. For the Exc. scans of Malachite Green (MG) + MGA (Malachite Green aptamer), instead, emission was measured at 650 ± 8 nm while varying excitation in the range 520-630 nm. Exc. scans were collected for nanostars $\overline{\text{A}}$, A, $\overline{\text{B}}$ and B, as well as negative controls, namely dyes in nuclease-free water ("Water") or *in vitro* transcription mixtures ("IVTM").

Conversely, when acquiring fluorescence emission (Em.) scans for DFHBI+BrA, excitation was fixed to 447 ± 8 nm, *i.e.* the peak value recorded in the corresponding excitation scan,

and emission was recorded in the range 470-620 nm. For the Em. scans of MG+MGA, excitation was similarly set to the peak value of 607 ± 8 nm, and emission was measured in the range 630-720 nm. Em. scans were collected for nanostars $\overline{\text{A}}$, A, $\overline{\text{B}}$, B and C, as well as the negative controls mentioned above.

2.8.2 Fluorescence Intensity Kinetics

Self-assembly kinetics of Quad-Stars

For self-assembly and disassembly kinetics of Quad-Stars, NMM excitation was set to 388 ± 15 nm as for emission spectra. Since the two emission peaks of NMM at 608 nm and 670 nm were observed to provide comparable insights, NMM emission was only monitored in the range 670 ± 15 nm.
DNA samples were pre-annealed in TE 1× buffer supplied with either 300 mM LiCl (soluble Quad-Stars) or 2 mM $MgCl_2$ (cholesterolised Quad-Stars). Unless otherwise specified, all wells were filled with 32 μL mixtures, consisting of 30 μL of the described DNA samples and 2 μL NMM solution. NMM concentration was kept stoichiometric with respect to expected G4 concentrationn as mentioned above for emission spectra. Plain TE 1× buffer mixtures with the same [NMM] as samples were measured as blanks. Samples where KCl injection was replaced with plain TE 1× buffer, added prior to the start of run, were used as negative controls to ensure no fluorescence increase or any unexpected trends in the absence of the K^+ ion trigger.
Before starting any measurement run, the thermal stage of the microplate reader was set to 25 °C, the plate was inserted and allowed to equilibrate for 45 min. After thermalisation, fluorescence intensity was recorded for 15 min (soluble Quad-Stars) or 30 s (cholesterolised Quad-Stars) to obtain an initial baseline prior to injecting a 1.26 M KCl solution to induce Quad-Star assembly. Unless otherwise specified, final concentrations were the following: $[Oligo]_0$ = 20 μM, [KCl] = 300 mM. Such final [KCl] value required a 10 μL injection of the 1.26 M stock solution in use. Accordingly, final volumes were kept constant across the entire plate and equal to 42 μL.

Soluble Quad-Stars When investigating [KCl] dependency in soluble Quad-Stars, KCl concentration was varied by adjusting the injected volume. To ensure constant final volume in all wells, such injection reductions were replaced with an equivalent additional amount of plain TE 1× buffer, manually added prior to the start of the run.
Due to their slower assembly rates, kinetics for soluble Quad-Stars were measured in "plate kinetics" mode, where the entire plate is sequentially scanned at every timepoint. Shaking (600 rpm for 1-2 s) was enabled after every reading. Settling time post-shaking was set to 0.1 s. Runs were briefly paused after the KCl injection to seal the plate with a transparent adhesive film (MicroAmp optical adhesive films, Applied Biosystems, ThermoFisher Scientific), thus

reducing evaporation over the 22 h-long measurement runs. Due to this adhesive covering the top of the plate, measurements were collected using bottom optics readout. To better sample assembly kinetics, measurements were acquired as frequently as possible (30-40 s intervals) at first and more sparsely later, for a total of 1000 measurement cycles over the course of 22 h.
Data were acquired in triplicates for samples with varying [KCl], apart from 6G5T and 6G7T at 30 mM KCl, which were measured in duplicates only. Data for samples with varying $[Oligo]_0$ were collected in triplicates/quintuplicates for 5G/6G, respectively. Negative controls were measured in duplicates.

Cholesterolised Quad-Stars Owing to their much faster assembly, cholesterolised Quad-Stars (9T spacer only) were tested in "well kinetics" mode instead, where each well is injected and measured for the assigned number of cycles before proceeding to the next one. For cholesterolised Quad-Stars, $[Oligo]_0$ = 20 μM and [KCl] = 300 mM were used. Shaking was only performed immediately after the injection, using the same settings and settling time as above. Measurement interval was kept uniform and equal to 0.48 s for all 995 cycles. In order to prevent the cholesterol moieties in Quad-Stars from interacting with the hydrophobic plastic of the well-plates, wells were passivated with Bovine Serum Albumin (BSA, lyophylised powder, BioReagent, Sigma). Briefly, wells were filled with 70-100 μL of 1% w/v BSA solution in Milli-Q water. Coating was formed through overnight evaporation at 65 °C. Each well was thoroughly washed with Milli-Q water and nitrogen-dried prior to using the plate. Given the much shorter duration of such measurement runs, no optical adhesive was required. Therefore, given the bottom BSA coating, measurements were collected using top optics readout. Data for 5G/6G Quad-Stars were collected in four/ten repeats, respectively.
Experiments of subsequent assembly/disassembly transients were carried out on 6G Quad-Stars only. Experimental conditions were identical to folding experiments: "well kinetics" measurement mode, time step = 0.48 s, duration = 995 cycles. Stability of fluorescence signal prior to cryptand addition was ensured by delaying the injection of 2 M cryptand solution by $\approx$ 45 s from the start of the disassembly run. Injection volume was chosen to reach a final cryptand concentration equal to 320 mM, completely chelating the available 300 mM K^+ ions. Four repeats were performed.

Co-transcriptional folding and self-assembly kinetics of RNA nanostars

When monitoring co-transcriptional folding and self-assembly of RNA nanostars, both in bulk and within synthetic cells, excitation was set to 447 ± 5 nm/620 ± 5 nm , and emission was recorded at 501 ± 5 nm/655 ± 5 nm for DFHBI+BrA/MG+MGA channels, respectively. These values were chosen according to both literature values and peak excitation/emission obtained from the spectra discussed above. Dichroic was automatically set to 474 nm for DFHBI+BrA and 637.5 nm for MG+MGA. Well multichromatics were turned on, enabling

simultaneous monitoring of both DFHBI+BrA and MG+MGA channels in each well at every timepoint.

The thermal stage of the microplate reader was set to 30 °C at least 30 min prior to the start of each assay. The "plate kinetics" measurement mode was selected, as discussed above for soluble Quad-Stars. Since the long experimental time window (48 h) required use of the optical adhesive film (MicroAmp optical adhesive films, Applied Biosystems, ThermoFisher Scientific) to prevent evaporation, measurements were collected using bottom optics readout. Shaking was disabled, and the number of flashes per well per cycle was set to 20.

In order to ensure good sampling, in particular of the initial transcription burst and resulting assembly transient, measurement interval was adjusted to be denser at first (every 60 s for 500 cycles), and sparser later (every 180 s for 250 cycles, then every 400 s for the remaining 250 cycles). Measurement focal height was optimized for each assay (bulk/synthetic cells): sacrificial samples, within the same microplate and with the same sample volume, were prepared 48 h in advance, and an automatic focal height adjustment was performed to obtain the optimal focal plane.

In all assays, final sample volume was kept constant and equal to 22 μL across the entire microplate. Accordingly, negative controls, namely dyes in nuclease-free water ("Water") and dyes in *in vitro* transcription mixtures ("IVTM"), which lacked DNA template injection, received a nuclease-free water injection of identical volume prior to the start of the measurement run. Malachite Green and DFHBI dyes were added to a final concentration of 45.45 μM each to enable fluorescence read-out.

For kinetics assays in bulk, the measurement runs were paused after 11 cycles ($\approx$ 16 min) to inject the DNA templates, thereby triggering transcription. Injection was performed by manual pipetting due to the large number of distinct samples requiring a different DNA template. Care was taken to ensure proper mixing of the well contents upon injection, thus removing the need for shaking. As mentioned above, the microplate was covered with an adhesive film immediately after the injection to reduce evaporation. For kinetics assays in emulsion droplet-based synthetic cells, instead, due to the need to encapsulate the entire transcription mixture, including the DNA templates, prior to emulsification, the microplate was sealed before the start of the run, with no subsequent interruption. Unless otherwise specified, final [total DNA template] = 40 nM in both types of assays. For instance, when studying binary systems with stoichiometric composition in droplet-based synthetic cells, [A-T] = [B-T] = 20 nM. When investigating the influence of [DNA template] on initial transcription rates, [DNA template] was varied between 10, 20 and 40 nM by reducing the injected (bulk)/added (synthetic cells) DNA template volumes. In order to keep the final well volumes constant across the microplate as discussed above, such reductions were compensated by adding nuclease-free water instead.

2.9 Bright-field and epifluorescence optical microscopy

Timelapse imaging was performed *via* a Nikon Eclipse Ti2-E inverted microscope with Perfect Focus System (PFS), equipped with a Hamamatsu Orca-Flash 4.0 v3 camera, a Lumencor SPECTRA X LED engine, and the following objective lenses: CFI Plan Apochromat λ 10$\times$ (NA 0.45, WD 4000 μm), λ 20$\times$ (NA 0.75, WD 1000 μm), λ 40$\times$ (NA 0.95, WD 210 μm) and λ 60$\times$ Oil (N.A. 1.40, W.D. 130 μm) immersion lens, where NA is the numerical aperture and WD is the working distance.

DNA Quad-Stars

Isothermal self-assembly/disassembly of Quad-Stars

Cholesterolised Quad-Star samples were annealed overnight in TE 1$\times$ buffer supplemented with 2 mM $MgCl_2$, at $[Oligo]_0$ = 5.6 μM. $[Oligo]_0$ was reduced compared to fluorimetry assays to yield discrete aggregates and simplify downstream quantification *via* image segmentation. To enable fluorescence imaging, 25% of "Arm_duplex" strands were replaced with identical ones labelled with an Alexa488 fluorophore, namely "Arm_duplex_A_Alexa488", which was excited using a 510/25 nm (peak/bandwidth) SPECTRA X LED.
Unless otherwise specified, imaging was automated *via* the JOBS module in the Nikon NIS Elements software interface. PFS was enabled to ensure constant z-height during the timelapses. Samples were monitored for 2.5 h, with no delay between acquisitions. At each timepoint, both bright-field and epifluorescence z-stacks (42 μm height, 2 μm step) were captured. Three/four replicates were imaged for 5G/6G samples, with three non-overlapping fields-of-view (FOVs) monitored per sample well.
Unless otherwise specified, samples were imaged in FlexWell incubation chambers purchased from Grace Bio-Labs (6.5 mm $\times$ 6.5 mm $\times$ 3.2 mm, 8 wells). The incubation chambers were attached using the built-in adhesive strip on top of glass coverslips (24 mm $\times$ 60 mm, Menzel-Gläser). Initial sample volume was 90 μL. After 3 timepoints ($\approx$ 15-20 min), the run was briefly paused to trigger self-assembly by injecting 6.5 μL of TE 1$\times$ buffer supplemented with 4.5 M KCl and 2 mM $MgCl_2$. Chambers were sealed immediately after the injection by placing a second coverslip on top to limit sample evaporation.
Isothermal disassembly experiments were performed after the isothermal assembly described above. A 2 M solution of [2.2.2] cryptand in TE 1$\times$ buffer supplemented with 2 mM $MgCl_2$ was injected in each well, while monitoring several FOVs in either bright-field or Alexa488 channels for $\approx$1-2 min each at the highest possible framerate. In particular, the volume was divided in multiple injections, namely 2$\times$ 5 μL + 1$\times$ 10 μL + 1$\times$ 20 μL, to better monitor the disassembly.
For multiple cycles of reversible assembly-disassembly, the first round was performed as just described. Before triggering a new assembly transient with a 10/15 μL KCl injection

(second/third round), some of the supernatant (volume equal to the sum of previous KCl and cryptand injections, *i.e.* 46.5 μL/50 μL prior to second/third round) was carefully removed to prevent chambers from overflowing. The injected volume of 2 M cryptand solution was kept at 40 μL for each round. The multiple KCl/cryptand injections led to considerable sample dilution, *i.e.* ≈36.37%/38.78% for the second/third round with respect to the previous one, likely causing the decrease in aggregate formation both size- and number-wise. For the second/third cycle of reversible assembly/disassembly in 6G Quad-Stars, four/two repeats were imaged, respectively, with three non-overlapping FOVs being monitored in each.
Isothermal self-assembly of control samples, namely 2 mM $MgCl_2$ (negative control), 2 mM $MgCl_2$ + 300 mM LiCl, 300 mM LiCl + 300 mM LiCl, and 300 mM LiCl + 300 mM KCl, was performed with the same protocol described above. Injection of saline buffers – 2 M LiCl or 4.5 M KCl, both in TE 1× supplemented with 2 mM $MgCl_2$ – was administered after 2 cycles (≈ 10 min) from the start of the run. No buffer solution was injected in the 2 mM $MgCl_2$ negative control. Samples were imaged in incubation chambers obtained by carving circular holes (8 mm diameter) in a PDMS spacer and attached to a glass coverslip. Each sample was imaged in a single well (replicate), with three non-overlapping FOVs monitored. Only the 5G control in 300 mM LiCl + 300 mM KCl was monitored in two FOVs only. Each sample was imaged separately for 2.5 h every 5 min, using Nikon ND Acquisition.

Rhodamine B uptake and release in Quad-Star aggregates

Samples were prepared as described in Section 2.9. Before initiating assembly, though, rhodamine B was injected in the samples to a final concentration equal to ≈ 28 μM. Sample and KCl injection volumes, as well as incubation chambers were kept identical to those used in Section 2.9. Samples were imaged 3.5 h after the KCl injection with the 60× oil immersion lens. The fluorescence emission of rhodamine B was excited using a 550/15 nm SPECTRA X LED.
Rhodamine B release was triggered *via* a 3 μL injection of 2 M cryptand in TE 1× supplemented with 2 mM $MgCl_2$. The injection volume was reduced compared to Section 2.9 to slow down and better monitor the release process, imaging for 1-2 min at 2 fps. To the same aim, the cryptand injection was administered on the opposite side of the imaged FOV within each well.

UV-induced photo-disassembly of Quad-Star aggregates

Samples were prepared as discussed in Section 2.9 above. After 3.5 h from KCl injection, 6 μL of 105 μM NMM in DMSO were pipetted into the sample chambers, yielding [NMM]/[G4] ≈ 5. NMM fluorescence was excited *via* a 395/25 nm Spectra X LED. Time-lapses were acquired in NMM epifluorescence and/or bright-field channels at 2 fps, for a total of 2 min per FOV.

Biostability of Quad-Star aggregates

Samples were prepared as described in Section 2.9, albeit reducing the amount of Alexa488 fluorophore-labelled "Arm_duplex_A_Alexa488" to only 10 % of the overall amount of "Arm_duplex". As in Section 2.9, the fluorescence of Alexa488 was excited using s 510/25 nm SPECTRA X LED. Samples were imaged after 3-4 h from KCl injection, and were either left undisturbed in their assembly buffer (negative controls) or subjected to buffer exchange through supernatant removal and replacement. The latter was performed by carefully removing with a micropipette the top 75 μL out of the overall 96.5 μL and replacing with an equal amount of either Phosphate Buffered Saline (PBS) 1× (from PBS tablets, Sigma-Aldrich) or Dulbecco's Modified Eagle Medium (DMEM, high glucose, GlutaMAX, Gibco) 1× supplied with either 10 % or 30 % Fetal Bovine Serum (FBS, Gibco).
Samples were imaged every 20 min for 60 h, in both Alexa488 epifluorescence and bright-field (z-stacks), using an automated JOBS script in Nikon NIS Elements. Due to the long imaging duration, the PDMS incubation chambers were sealed with a transparent adhesive strip (FlexWell Seal Strips, Grace Bio-Labs) to reduce evaporation. Samples were imaged in triplicates, except for those in PBS 1×, which were imaged in duplicates only. Three FOVs were monitored for each sample well.

RNA nanostars

The following SPECTRA X LEDs were used to excite the corresponding fluorophores or fluorophore/aptamer complexes: 395/25 nm for Alexa Fluor 405 Streptavidin conjugate (Alexa405-STV), 470/24 nm for DFHBI+BrA, 550/15 nm for EYFP, 575/25 nm for TexasRed Streptavidin conjugate (TexasRed-STV), and 640/30 nm for MG+MGA.
Unless otherwise specified, [MG] = [DFHBI] = 45.45 μM in every sample, and [total DNA template] = 40 nM, as for fluorimetry experiments discussed in Section 2.8.2.
Samples were prepared as discussed in Sections 2.4.2 and 2.4.2, and imaged in hollow, rectangular, flat glass capillaries (0.20/0.40 mm × 4.00 mm × 50 mm, both purchased from VitroCom). Glass capillaries were chosen as they allow for a) substantially reduced evaporation during prolonged timelapse imaging (better sealing compared to incubation chambers), b) larger flat imaging area undisturbed by edge effects, and c) easier upside down imaging, necessary for use on a heating stage (see below). Capillaries were sealed and glued on top of glass coverslips (24 mm × 60 mm, Menzel Gläser) using a bi-component epoxy (Araldite Rapid). In the case of bulk samples, the sides of each capillary were capped with mineral oil to avoid the glue coming in contact with the water phase. This was not necessary for samples in emulsion droplets thanks to the surrounding oil phase. Glue was allowed to set for 30 min, during which samples were kept in a dark environment at room temperature. The glass coverslips were then taped to a peltier-controlled copper temperature stage (Temikra Ltd.). Temperature was controlled via custom XML scripts executed on a RaspberryPi 3

running RaspbianOS, with feedback enabled *via* a thermocouple attached with copper tape close to the samples. Due to the heatsink being positioned on top, the samples were imaged upside down, allowing for fluorescence imaging only. A schematic diagram and a photograph of the epifluorescence imaging setup are presented in Fig. 2.1.

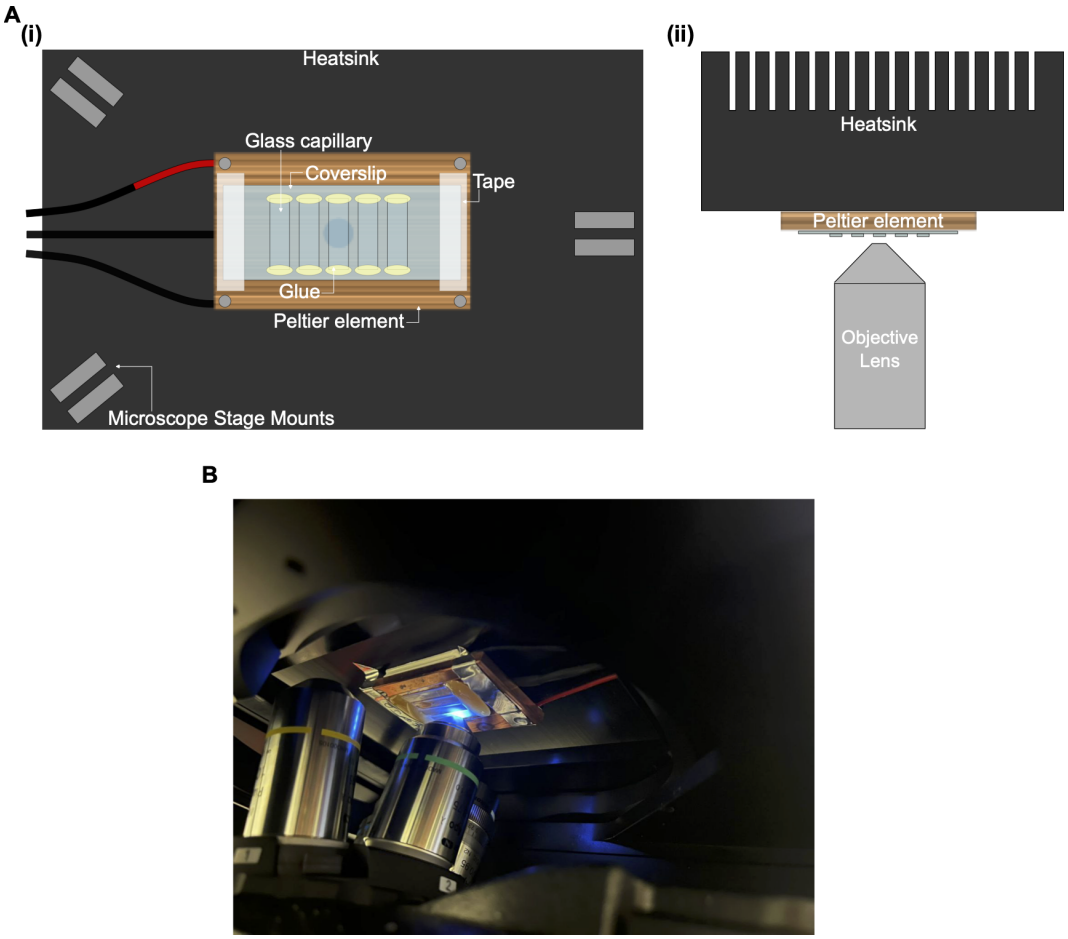


Fig. 2.1 **Epifluorescence imaging setup. A:** Schematic diagram of the epifluorescence imaging setup, **(i)** bottom and **(ii)** side views, respectively. **B:** Photograph of the epifluorescence imaging setup, depicting the objective lens, the glass capillaries glued to the glass coverslip (horizontally aligned, in this case) and taped (here with heat conductive tape) to the Peltier element of the heating stage. Photo credit: Diana Tanase (Di Michele group).

Imaging was automated *via* ND Acquisition in Nikon NIS Elements. PFS was enabled to ensure constant z-height during timelapse imaging. A single capillary was imaged per sample, with three non-overlapping FOVs monitored per capillary. In order to ensure proper focusing – given little to no signal at the start – z-stacks were captured at each timepoint, for a total height of 120-160 μm, ranging from -20/-60 to +100 μm with respect to the the PFS plane, with a 3.5-4 μm step depending on the imaging run.

Self-assembly of RNA condensates

Temperature was set to 30 °C. Samples were imaged first every 15 min for 10 h, then every 30 min for further 38 h. The only exceptions were binary and ternary nanostar systems in bulk, where a software issue led to early termination of the run at 42 h, and data at 48 h were collected manually.
As samples were sealed in capillaries, DNA templates had to be added prior to the start of imaging, rather than injected during the run. This led to imaging runs starting ≈1-1.5 h after adding the DNA templates to the corresponding transcription mixtures. Such delay is due to the time required for sample preparation, *i.e.* loading and sealing capillaries as well as waiting for the glue to set, and run setup, *i.e.* taping the capillaries, defining PFS offset and FOVs in each sample. Micrographs have been labelled according to the start of the imaging run, with $t = 0$ s being the start of the run rather than the start of transcription. When presenting data for samples imaged during different assays, care was taken to align the corresponding timepoints with respect to DNA template addition, *i.e.* start of transcription.

Melting of bulk RNA condensates

Samples were prepared as discussed in Section 2.9 above. Transcription was allowed to proceed for 48 h at 30 °C within a Bio-Rad C1000 Touch Thermal Cycler (heated lid at 35 °C). Samples were then kept in a dark environment at room temperature for less than 12 h. Upon imaging, temperature was initially set to 25 °C, and increased by 1 °C every 15 min up to 75 °C. Samples were imaged after 10 min of hold at each temperature.

2.10 Laser Confocal Scanning microscopy

Imaging was performed on a Leica Stellaris 8 (DMi8 CS Premium) inverted research microscope in the Facility for Imaging by Light Microscopy (FILM) at Imperial College London (Hammersmith). The microscope was equipped with a solid state 405 nm laser as well as a white light laser (WLL, 440-790 nm). The microscope was also equipped with four internal detectors (3 Power HyDS and 1 HyDX) plus a transmitted light detector (Trans PMT) for bright-field imaging. The following lenses were used: HC PL APO CS2 10× DRY (NA 0.40), HC PL APO CS2 20× DRY (NA 0.75) and HC PL APO CS2 63× OIL (NA 1.40, only used for RNA MLOs in *E. coli* in Chapter 7), where NA is the numerical aperture.
The following excitation/emission settings were used. For Alexa Fluor 405 Streptavidin conjugate (Alexa405-STV), the 405 nm solid state laser was used, and emission was recorded around 421 nm. For DFHBI(-1T)+BrA, the WLL excitation was tuned to 447 nm (472 nm) and emission was recorded around 501 nm (507 nm). For EYFP, the WLL excitation was tuned to 514 nm, with emission recorded around 527 nm. For TexasRed Streptavidin conjugate (TexasRed-STV), the WLL excitation was tuned to 595 nm, and emission was recorded

around 615 nm. For MG+MGA, the WLL excitation was tuned to 628 nm and emission was recorded around 650-660 nm. MG was not added to samples with TexasRed-STV to avoid spectral overlap. A line-sequential illumination mode was chosen, such as the following one: Trans PMT (bright-field), 405 nm (Alexa405-STV) and 628 nm (MG+MGA) in sequence 1, 447 nm (DFHBI+BrA) in sequence 2, 514 nm (EYFP) in sequence 3.
The pinhole was set to 1 Airy Unit (AU, computed automatically), and line-averaging was enabled and set to 2-3. Bi-directional scan mode along the x direction was enabled after carrying out phase calibration. Stage speed was set to 400 Hz for high-resolution single images (4096×4096 or $8192 \times 8192\,\mathrm{px}^2$ – px = pixel), and 1000 Hz for zoomed-in z-stacks acquired with the $20\times$ lens ($63\times$ lens only for MLOs in *E. coli* in Chapter 7). When capturing z-stacks, the zoom factor and the top/bottom planes were manually adjusted. Z-step was set to the automatically computed Nyquist-optimised value. Bit depth was set to the highest value, 16 bit. 2D orthogonal cross-sections (XY, XZ, YZ) were generated from z-stacks using the "Sections" plugin in the 3D module of the Leica Application Suite (LAS) X software interface.

2.11 Numerical methods

2.11.1 Gel Image Analysis

Lane profiling was performed *via* a semi-automated Python3 script requiring users to specify cropping – in the form of (x, y) coordinates of the upper and lower corners – and the number of lanes to profile. Lanes are identified by assuming constant lane-width within the same gel, and profiles are computed as the average intensity across a given lane-width along the (cropped) height of the gel. The script is publicly available on GitHub (https://github.com/GiacomoFabrini/Gel_Image_Analysis_1.0).
Unless otherwise specified in the figure captions, agarose gel images are reported with no contrast enhancement after tracing of the sample lanes found by lane profiling. Gels on DNA templates and RNA transcripts in Chapters 4-6 are reported after additional look-up table (greyscale) inversion for ease of visualisation.

2.11.2 Fitting DLS Size Distributions

Average hydrodynamic diameters, $< d_{\mathrm{H}} >$, were obtained by fit to either mono or bi-modal log-normal distributions, performed in Python3 using scipy.optimize.curve_fit with 10000 maximum function evaluations. Reported standard deviations on the hydrodynamic diameters were computed using the covariance matrix of the fit.

2.11.3 Data analysis and fitting of Quad-Stars fluorescence kinetics

Data preprocessing and fitting were performed in Python3. The mathematical model is discussed in detail in Section 3.3.1.
As for preprocessing, the recorded fluorescence intensity (F.I.) were background-corrected by subtracting blank profiles, namely those of TE 1× buffer supplied with the same [NMM] and receiving the same KCl injection. This is necessary as fluorescence emission of NMM is non-zero even in plain buffer. Time origin was also shifted so that $t_{\text{injection}} = 0\,\text{s}$.
Fitting was performed *via* batch minimisation of the Mean Squared Error (MSE) loss using the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm in the scipy.optimize.minimize implementation in Python3, with parameter bounds for the reaction order n (min = 2, max = 5).
Batch fitting was chosen to improve fitting results by taking into account data for all examined spacer lengths, as these have theoretically no impact on the reaction order n. In particular, for the kinetics of soluble Quad-Stars with varying [KCl] injections at fixed 20 μM $[Oligo]_0$, batches comprised all curves for designs with all spacer lengths sharing the same G-run length (5G separated by 6G) and the same [KCl] injection (*e.g.* 90 mM separated from 300 mM). This means that each batch was constrained to have the same fitted value of n regardless of spacer length, while all other parameters were allowed to vary. Conversely, for the kinetics of soluble Quad-Stars with different $[Oligo]_0$ at fixed 300 mM KCl, batches comprised all curves for designs with all spacer lengths sharing the same G-run length and the same $[Oligo]_0$. The resulting $t_{1/2}$ values were log-transformed ($\ln(t_{1/2})$) and linearly fitted with respect to $\ln([Oligo]_0)$, with datapoints weighted according to the associated standard deviations. The reaction order, n, was computed as $1 - m$, with m being the slope of such linear fit. Reported uncertainties were calculated from the covariance matrix of the fit.
Finally, for cholesterolised Quad-Stars, curves were fitted in batches simply corresponding to G-run length (5G/6G).

2.11.4 Data analysis and fitting of RNA nanostar fluorescence kinetics

Fluorescence kinetics were acquired in triplicates for all samples, including negative controls. Data preprocessing, analysis and visualisation were performed in Python3.
First of all, all profiles were min-max normalised, *i.e.* the raw minimum was subtracted and then they were divided by the range between the raw minimum and the raw maximum. For each sample, mean and standard deviation were computed from collected triplicates, both normalised and raw. In particular, unless otherwise specified, given a batch of samples to be visualised in the same panel or subplot (for instance B nanostars with varying [B-T]), normalised profiles (mean, standard deviation) were rescaled by multiplying them by the ratio between their raw average maximum and the highest one across the corresponding batch. This way, despite normalisation, relative intensities were preserved. The only exceptions

were binary nanostar systems within droplet-based synthetic cells: here the rescaling factor was computed between the raw average maximum of each sample and the one for the stoichiometric sticky nanostar system ([A-T] = [B-T]) to better showcase trends in fluorescence intensity resulting from varying [A-T]/[B-T] composition ratios.

Fitting initial transcription rate with varying DNA template concentration

To evaluate the influence of [DNA template] on the initial transcription rate, investigated through the proxy of initial increase in fluorescence intensity, 30 data points were extracted from the linear-like transient of the mean normalised profiles following DNA template addition. In the case of bulk *in vitro* transcription, such data points were taken from the time window immediately following manual injection. Conversely, for *in vitro* transcription within synthetic cell mimics, as DNA templates were added prior to emulsification and start of the measurement, those data points were extracted from the first time window exhibiting sufficiently linear-like behaviour, leading to possible artifacts. In both cases, resulting data points were linearly fitted with respect to the injected [DNA template] using scipy.optimize.curve_fit. Fitted slopes were plotted *vs* [DNA template] (markers), with errorbars displaying the standard deviations computed from the covariance matrix of the fit.

2.11.5 General image processing for micrographs

Raw epifluorescence timelapses from Nikon ND2 files were processed using Python3 scripts relying on the ND2Reader library to extract the in focus z-planes and save them in TIFF format for easier downstream usage. For experimental work on Quad-Stars described in Chapter 3, in focus z-planes were manually indicated, *i.e.* provided in input to the script. For work on RNA nanostars described in Chapters 4-6, instead, the optimal in focus planes were determined using the Normalised Variance criterion described in refs. [343, 344]. Briefly, the optimal in focus z-plane for each timepoint was identified by maximising the ratio between the image pixel intensity variance and its mean across the scanned z-height (using np.argmax). In the case of samples forming condensates in multiple fluorescence channels, the mean normalised variance across the corresponding channels was maximised instead. On some occurrences, normalised variance failed to detect the correct in focus planes at early timepoints in favour of very high z-levels, but later stabilised on the correct planes, close to the one locked by PFS. In that case, any plane exceeding 80% of the scanned height for the first few timepoints only was discarded and replaced with the first occurring lower, more stable z-height level. Finally, for samples failing to form any condensates, in focus planes were manually indicated by specifying the desired z-height range.
Epifluorescence micrographs in Chapter 3 are reported pristine/unprocessed, or after either background-subtraction *via* Rolling Ball (radius = 50-100 px, depending on aggregate size) or contrast-enhancement using stacked (whole timelapse) histogram stretching with 0.30%

of the pixels saturated, both carried out in FIJI (ImageJ) [345]. Image processing is specified in the corresponding figure captions.
Epifluorescence micrographs in Chapters 4-6 are reported after contrast-enhancement (0.25% pixels saturated) with normalisation enabled, *i.e.* each frame is processed independently, which allows to better resolve small fluorescent condensates at early timepoints where signal is low. Only channels exhibiting fluorescence signal are included in the composite micrographs in order to avoid excessive noise stemming from contrast-enhancement with normalisation.

2.11.6 Segmentation of Quad-Star aggregates

Timelapses in TIFF format were analysed using a FIJI macro implementing the following pipeline: background subtraction *via* Rolling Ball (radius = 50-100 px, depending on aggregate size) to remediate uneven illumination; denoising through subsequent median (radius = 5 px) and minimum (radius = 1 px) filters; segmentation by unsupervised thresholding (Li algorithm [346]) and conversion to binary mask; watershed to separate clusters of neighbouring aggregates; Particle Analysis with size ($> 100\,px^2$) and circularity ($> 60\,\%$) thresholds.

2.11.7 Isothermal assembly/disassembly area fraction profiles

The last step of the segmentation pipeline above produced a set of measurements including the area fraction (%) of microscopy FOV occupied by fluorescent Quad-Star aggregates, which was used as a proxy to investigate the kinetics of aggregate assembly and disassembly. Notably, all isothermal assembly timelapses were run with no acquisition delay, thereby leading to uneven time resolution (interval) when comparing samples imaged on different days, namely due to discrepancies in exposure time, z-stack height, and/or number of simultaneously imaged wells. Cubic spline interpolation (overall length equal to 2.5 h, number of equally spaced points set to 30) was used to overcome such time resolution unevenness. Profiles were aligned using the annotated injection timepoints, finally allowing for mean and standard deviation calculation. No interpolation was needed for reversible assembly/disassembly of 6G Quad-Stars nor for isothermal assembly of controls with different saline buffer injections.

2.11.8 Rhodamine B cargo loading in Quad-Star aggregates

In order to evaluate rhodamine B uptake in Quad-Star hydrogel aggregates, epifluorescence micrographs captured after 3.5 h from KCl-triggered assembly were processed using the segmentation pipeline in Section 2.11.6. Uptake was evaluated by computing the ratio between the average rhodamine B fluorescence intensity within Quad-Star aggregates and that in the

background through simple intersection with the computed binary masks (aggregates) and their reciprocal (background).
The same procedure was applied to epifluorescence timelapses of rhodamine B release upon cryptand injection. First of all, Quad-Star aggregates were segmented in the timepoints preceding the injection, and background was identified as the inverted selection of the obtained aggregate segmentation mask. Release profiles were then computed as the time-dependent evolution of the integrated rhodamine B fluorescence intensity within that background area (kept constant with progressing time). Area fraction profiles for the simultaneous disassembly of Quad-Star hydrogel aggregates were obtained as discussed in Section 2.11.7, and overlaid with release profiles with twin x axis (different y axis).
As the specific disassembly and cargo release timescales strongly depend on the proximity between the imaged FOV and the cryptand injection location, data were reported for single representative FOVs within independent 5G and 6G samples for both distant (main text) and close-by (Appendix) injections: as expected, closer proximity leads to faster and less temporally resolved profiles.

2.11.9 Photo-induced disassembly

Disassembly profiles were obtained using the segmentation pipeline in Section 2.11.6. The "IC" engraving in the 6G Quad-Star well was obtained by programming exposure to near UV LED (395/25 nm) at precise locations within the sample arranged along the desired engraving path. This was accomplished using the JOBS module in the Nikon NIS Elements software interface. After engraving, multiple wells were imaged using the CFI Planar APO λ 20$\times$ mentioned above, and stitched into the reported entire-well image using the automated "Large Image" stitching in NIS Elements.

2.11.10 Growth of RNA condensates in bulk/synthetic cells

Maximum Intensity Z-projection

Micrographs for growth kinetics analysis of RNA condensates in bulk and within synthetic cells were extracted similarly to what was reported in Section 2.11.5 for (semi-)automatic extraction of in focus planes from raw ND2 files using the Normalised Variance criterion, with two differences. First, the exported timelapses comprised maximum intensity z-projections as timepoints rather than single optimal focus planes. In particular, the projected z-volume consisted of 5 z-planes, spanning 2 planes on either side of the optimal focus plane, for a total projected volume of 7-8 μm around the mid plane. Second, the optimal in focus planes of each channel featuring condensate formation were determined independently, as the equatorial planes of differently sized condensates such as those in binary systems with extreme DNA template concentration ratios might differ considerably in z-height. For

condensates of RNA nanostar C, which appear fluorescent in both MG and DFHBI channels, this approach yields the same result as the fluorescent signals stem from the same objects occupying the same z-volumes.

Segmentation and Particle Analysis

Maximum intensity z-projected timelapses (see Section 2.11.10) investigating the self-assembly of RNA condensates were processed using a FIJI macro implementing the following pipeline: contrast-enhancement (0.25% pixels saturated, normalisation enabled – see Section 2.11.5); pixel (px) calibration using the px/μm scaling factor reported in the metadata of raw ND2 files; denoising *via* Gaussian Blur (sigma = 1); background-correction through two steps of Rolling Ball (sliding paraboloid enabled, smoothing disabled, radius = 1 px); automatic thresholding (see below); Particle Analysis with size thresholding ($> 10\,\text{px}^2$ in area). Segmentation masks and Particle Analysis results were saved in TIFF and CSV formats, respectively. In case of samples forming condensates in more than one fluorescence channel, each was analysed independently, resulting in one binary mask and Particle Analysis result file per channel.
As for thresholding, different unsupervised algorithms were used for assays in bulk – Li [346], which relies on minimum cross-entropy – and those within synthetic cells – Otsu [347], which minimises intra-class variance (*i.e.* maximises inter-class variance in the case of two classes only). This was done due to the tendency of Li's algorithm to incorrectly threshold signal from largely out-of-focus emulsion droplets, as well as over-segment neighbouring objects within them.
Particle Analysis results were processed in Python3 using pandas. The time-evolution of the number of A/B condensates in each FOV was obtained through a group-by statement counting the number of occurrences at each timepoint. Data are thus reported as mean number of objects per microscopy FOV, N, ± corresponding standard deviation from the FOV triplicates. Such analysis was carried out only for RNA nanostars A, B and their binary combinations, due to nanostar C forming mesh-like assemblies where Particle Analysis is not applicable.

Calculation of Chord Length Distribution (CLD)

Single channel segmentation masks obtained using the pipeline described in Section 2.11.10 were processed using a Python3 script relying on PoreSpy to compute the corresponding Chord Length Distributions (CLDs). In particular, the porespy.filters.apply_chords function, which relies on scipy and skimage, was used to return raw chords along both x and y directions, with spacing = 1 px, *i.e.* allowing for adjacent chords. Chords spanning objects touching the edges of the FOVs were kept setting the option trim_edges = False. Only channels showing the formation of condensates were analysed. For each sample, three independent FOVs within

the same capillary were imaged. For each channel of each FOV, chords along x and y were pooled together at each timepoint, converted from px to μm using the px/μm scaling factor reported in ND2 metadata, resulting in one time-dependent CLD per channel per FOV. From each of these replicates, the mean of the CLD – used as proxy for typical condensate size – was computed, resulting in a time-dependent sampling distribution of the mean comprising at each timepoint as many replicates as the number of FOVs, *i.e.* three. Data are reported as the time-dependent evolution of the mean of such mean replicates, μ_{CLD}, ± the standard error of the mean, estimated as the standard deviation of the sampling distribution given by the mean replicates. Full CLDs are displayed using ridge plots obtained by vertically stacking kernel density estimation (KDE) plots from representative timepoints, each produced with seaborn.kdeplot.

2.11.11 Calculation of Kissing Loop Melting Temperature

Timelapses were segmented according to the procedure in Section 2.11.10. Timepoints were converted to temperature steps by simply considering the starting condition (25 °C) and the temperature increment between consecutive frames (+1 °C). CLDs were extracted from the resulting binary masks as described in Section 2.11.10. As three non-overlapping FOVs per capillary and one capillary per sample were imaged, this resulted in three CLDs and corresponding mean replicates for each sample, from which μ_{CLD} and related standard error of the mean were computed. The melting temperature, T_{M}, of KL interactions was computed as the value at which μ_{CLD} dropped to 50% of its initial value at 25 °C.

2.11.12 Circle Hough Transform for radial distribution of orthogonal RNA condensates in bulk

Input data comprised maximum intensity z-projected epifluorescence micrographs, as discussed in Section 2.11.10, acquired with the CFI Plan Apochromat λ 20× objective lens after 48.5 h from the start of transcription. As mentioned above, three FOVs were imaged per sample. Micrographs were processed using a semi-automated Python3 script heavily relying on OpenCV (cv2) performing the following pipeline on one fluorescence channel at a time: conversion to 8 bit (skimage.util.img_as_ubyte); denoising by Median Blur (cv2.medianblur, radius = 3 px); greyscale to colour conversion (required for the following step); Circle Hough Transform (CHT, cv2.HoughCircles). In particular, the following parameters were used for CHT: image to accumulator resolution = 5, minDistance = 20 px, minRadius = 20 px, maxRadius = 500 px, method = HOUGH_GRADIENT_ALT. The script is semi-automated in that the two available threshold parameters (low and high passed to the Canny edge detector) were manually adjusted on a case by case basis. The script resulted in a list of circular objects found in the corresponding channel. Radii of the found circles were converted from px to μm using the px/μm scaling factor reported in the ND2 metadata.

Resulting circles were filtered to remove overlapping objects, leaving only the largest one. After filtering, the following numbers of condensates were kept, reported as A/B ([A-T]/[B-T] DNA template concentration ratio): 155/56 (0.25), 78/60 (0.50), 53/66 (1), 56/99 (2), 50/131 (4). Data were arranged in two pandas dataframe objects, one per channel, with X reporting the DNA template concentration ratio and Y the condensate radii. Data are reported in violin plots displayed using seaborn, with A/B condensates in red/cyan, respectively. Kernel width (along x) for violin plots was chosen to be proportional to the number of observations (scale = "count") and the bandwidth was set to 0.45.

2.11.13 Scaling of RNA organelle volume with synthetic cell volume

Input data comprised confocal micrographs acquired with the HC PL APO CS2 20× objective lens, depicting samples of pure A/B RNA organelles self-assembled within synthetic cells after > 48 h from the start of transcription. Emulsion droplets and the corresponding RNA organelles were manually annotated in FIJI using the circular selection tool, for a total of 66/70 datapoints for A/B organelles, respectively. The area of the annotated regions in μm^2 was computed in FIJI, saved in CSV format and processed in Python3 as described below. The radii of annotated emulsion droplets and RNA organelles were computed by assuming their shape to be perfectly spherical. This enabled the calculation of the corresponding volumes. Therefore, the scaling of RNA organelle volume with respect to emulsion droplet volume was simply evaluated by linear regression using scipy.stats.lingregress, which relies on ordinary least squares, and plotted using seaborn.regplot. The parameters of the optimal linear regression model, reported as mean ± standard error, as well as the R^2 and p-value with respect to the null hypothesis H0 of null slope are the following: A - slope = 0.182 ± 0.005, intercept: -974.321 ± 861.614, R^2 = 0.98, p-value = 2.20 10^{-45}; B - slope = 0.153 ± 0.003, intercept: -1462.931 ± 993.888, R^2 = 0.99, p-value = 2.72 10^{-59}.

2.11.14 Distribution of radial ratio of orthogonal RNA organelles in emulsion droplets

Input data were pixel to μm calibrated confocal micrographs comprising single equatorial planes, one FOV per sample, acquired with the HC PL APO CS2 20× objective lens after > 48 h from the start of transcription.

ROI selection

In order to compute the radial ratio for orthogonal organelles within the same emulsion droplet, an ROI (region of interest) selection step was added to the pipeline presented in Section 2.11.12. Briefly, the bright-field channel was contrast-enhanced as mentioned above (contrast-stretching with 0.20% pixels saturated, using skimage.exposure.rescale_intensity),

and used to produce a composite with MG and DFHBI channels (cv2.merge). Such composite was displayed to the user for ROI selection using cv2.selectROIs. Upon selection of a bounding box on the screen, the corresponding ROI was saved, cropped from both MG and DFHBI channels and passed to the rest of the pipeline as discussed below.

Circle Hough Transform (CHT)

ROI patches were processed using the pipeline in Section 2.11.12 with CHT parameters modified as follows: image to accumulator resolution = 1, minDistance = 20 px, minRadius = 10 px, maxRadius = 150 px, method = HOUGH_GRADIENT_ALT (kept unchanged). The two threshold parameters were kept constant and equal to: parameter1 = 150, parameter2 = 0.6.
As mentioned in Section 2.11.12, overlapping circles were removed and only the largest was kept. ROIs were only considered if an object, *i.e.* RNA organelle, was found in both MG and DFHBI channels. All detected objects were visually inspected to remove false positives, such as cases when an entire synthetic cell was detected rather than the enclosed RNA organelle. The resulting dataset was arranged in a pandas dataframe and saved in CSV format for later use.
Finally, only synthetic cells containing exactly one A and one B-type organelle were considered to compute the organelle radial ratio *vs* [A-T]/[B-T]. When multiple organelles of a given type are simultaneously present inside an emulsion droplet, in fact, they tend to be considerably smaller, thus inevitably skewing the resulting size ratio distribution. This resulted in the following numbers of synthetic cells being kept at each [A-T]/[B-T]: 33 (0.25), 41 (0.50), 40 (1), 31 (2), 28 (4). As reported in Section 2.11.12, data were arranged in a pandas dataframe, with X being the DNA template concentration ratio and Y the radial ratios of the corresponding organelles. Data were presented as violin plots in seaborn using similar kernel width parameters as above (scale = “count”, bw = 0.70).

2.11.15 Number of orthogonal RNA organelles within synthetic cells

The number of A/B organelles per synthetic cell with varying [A-T]/[B-T] was extracted from the dataset discussed in Section 2.11.14. In this case, the constraint of each synthetic cell containing exactly one A-type and one B-type organelle was lifted. This resulted in the following numbers of examined synthetic cells for each [A-T]/[B-T]: 42 (0.25), 42 (0.50), 46 (1), 40 (2), 32 (4). Data are reported as the fraction of synthetic cells presenting a certain number (1, 2 or 3) of A/B-type organelles, normalised with respect to the overall number of examined droplets for the same [A-T]/[B-T], and displayed in stacked (top: A, bottom: B) grouped bar charts.

2.11.16 Evaluating linker-induced mixing

Input data comprised confocal micrographs (single equatorial planes, pixel to μm calibrated), acquired with the HC PL APO CS2 10× objective lens after >48 h from the start of transcription. Each sample corresponded to a different Linker Fraction, defined as the DNA template concentration ratio [L-T]/([A-T]+[L-T]+[B-T]). Most samples were imaged in a single FOV, apart from Linker Fractions = 1/11 (or [A-T]:[L-T]:[B-T] = 5:1:5) and 1/5 (or [A-T]:[L-T]:[B-T] = 2:1:2), which were imaged in duplicate non-overlapping FOVs instead.

Segmentation of confocal micrographs

Both MG and DFHBI fluorescence channels of each micrograph were segmented using a FIJI macro implementing the following pipeline: denoising *via* Gaussian Blur (sigma = 3); background-subtraction *via* Rolling Ball (sliding paraboloid enabled, smoothing disabled, radius = 10 px); unsupervised thresholding *via* both Otsu and Li algorithms (separately, see below). This resulted in four segmentation masks per micrograph, two per channel, one processed *via* Otsu and the other using Li thresholding. All masks were saved in TIFF format.

ROI selection

Raw micrographs and masks from Section 2.11.16 above were passed to an ROI selection pipeline similar to the one described in Section 2.11.14. However, the images showed to the user for ROI selection were composites of MG and DFHBI fluorescent channels only (no bright-field). After ROI selection from the user in the form of a bounding box identifying an emulsion droplet, the corresponding ROI patches were extracted from the raw micrograph channels as well as the corresponding segmentation masks. All resulting patches were converted to 8 bit (skimage.util.img_as_ubyte) for the steps below.

Mixing evaluation

For each Linker Fraction value, the following quantities were computed: $I_{\mathrm{A}}^{\mathrm{A}}$, $I_{\mathrm{A}}^{\mathrm{B}}$, $I_{\mathrm{B}}^{\mathrm{A}}$, $I_{\mathrm{B}}^{\mathrm{B}}$, where $I_{\mathrm{X}}^{\mathrm{Y}}$ is the average intensity of fluorescent channel X within the binary mask obtained from channel Y. Such quantities were computed by averaging (np.mean) the result of the intersection of channel X and binary mask Y *via* bitwise AND (cv2.bitwise_and). The normalised mixed-term ratios, namely $J_{\mathrm{A}} = I_{\mathrm{A}}^{\mathrm{B}}/I_{\mathrm{A}}^{\mathrm{A}}$ and $J_{\mathrm{B}} = I_{\mathrm{B}}^{\mathrm{A}}/I_{\mathrm{B}}^{\mathrm{B}}$, termed mixing indices, were then used as proxies to evaluate the mixing between the otherwise orthogonal fluorescent channels and RNA nanostars bearing the corresponding aptamers.
In particular, when calculating the same-phase terms, *i.e.* $I_{\mathrm{A}}^{\mathrm{A}}$ and $I_{\mathrm{B}}^{\mathrm{B}}$, the bitwise AND was computed directly between micrograph patches and mask patches, as in:

$$I_{\mathrm{A/B}}^{A/B} = \mathrm{mean}(\mathrm{bitwiseAND}(\mathrm{patch}_{\mathrm{A/B}}, \mathrm{mask}_{\mathrm{A/B}})). \tag{2.16}$$

When computing the mixed-phase terms, *i.e.* $I_{\mathrm{A}}^{\mathrm{B}}$ and $I_{\mathrm{B}}^{\mathrm{A}}$, instead, the mask patches were adjusted depending on the Linker Fraction value. For partially mixed systems, namely those with Linker Fraction < 1/3, non-overlapping mask patches were used, *i.e.* the mask patch minus its intersection with the opposite-type mask patch, as in:

$$I_{\mathrm{A}}^{\mathrm{B}} = \mathrm{mean}(\mathrm{bitwiseAND}(\mathrm{patch_A}, \mathrm{mask_B} - \mathrm{mask_A} \bigcap \mathrm{mask_B})). \tag{2.17}$$

The opposite term was calculated analogously. This was done to prevent signal leakage from either phase. For fully-mixed systems, namely Linker Fraction > 1/3, instead, the union of opposite-type masks was used, as in:

$$I_{\mathrm{A}}^{\mathrm{B}} = \mathrm{mean}(\mathrm{bitwiseAND}(\mathrm{patch_A}, \mathrm{mask_A} \bigcup \mathrm{mask_B})) \tag{2.18}$$

and analogously for the opposite term. This was done to ensure consistent results regardless of discrepancies in segmentation masks stemming from differences in signal intensity across channels.

All ROIs and corresponding computations were visually inspected to discard any synthetic cell patches displaying either incorrect or extremely small segmentation binary masks in either fluorescent channel. To ensure independent datapoints, ROIs were also checked against overlap of the respective bounding boxes.

The mixing indices J_{A} and J_{B} were computed using binary masks obtained *via* both Li and Otsu thresholding by simply saving and reusing the ROIs. The resulting indices were then averaged to compensate the opposite tendencies of the two algorithms to under- (Otsu) and over-segment (Li). Data were arranged in a pandas dataframe, with Linker Fraction on one column and J_{A} or J_{B} on the other. Data were filtered as follows: NaN (not a number) values were dropped; datapoints either below $\mathrm{Q}_1 - 1.5$ IQR or above $\mathrm{Q}_3 + 1.5$ IQR were discarded as outliers, where $\mathrm{Q}_{1/3}$ is the first/third quartile and IQR = $\mathrm{Q}_3 - \mathrm{Q}_1$ is the interquartile range. This resulted in the following number of kept datapoints (synthetic cells) per Linker Fraction: 28 (1/21), 31 (1/15), 30 (1/11), 24 (1/7), 32 (1/5), 38 (1/3) and 32 (1/2). Data are presented in violin plots created using seaborn, with similar parameters for the kernel width as those in Section 2.11.14 (scale = "count", bw = 0.70).

2.11.17 Evaluation of protein partitioning with protein-binding constructs

Confocal micrographs acquired with the HC PL APO CS2 20× objective lens (single equatorial planes, pixel to μm calibrated) were exported in TIFF format. In focus emulsion droplets were manually annotated in FIJI using the circular selection tool. Binary masks were created from the annotated ROIs using the "Masks from ROIs" FIJI plugin [348] and saved in TIFF format. Micrographs and emulsion droplet masks were analysed through a Python3 script

performing the following pipeline: the channel corresponding to the nanostar-binding dye (MG/DFHBI) was denoised via Gaussian Blur (cv2.GaussianBlur, sigma = (3, 3)), intersected with the droplet annotation mask, and the RNA organelles were segmented using unsupervised Li algorithm (skimage.filters.threshold_li). This resulted in an inner mask comprising the RNA organelles (above threshold). The outer mask was obtained by inverting the inner mask (*i.e.* selecting below threshold), intersecting the result with the emulsion droplet annotation mask, and applying morphological opening (skimage.morphology.opening) to remove thin bright lines/halos around the previously segmented organelles.
The channel corresponding to the protein of interest (EYFP/Alexa405-STV/TexasRed-STV) was then intersected with, *i.e.* multiplied by, both inner and outer masks in the previous step, and the mean was computed in each case. This produced the average protein fluorescence intensity within the RNA organelles and outside them, *i.e.* in the droplet lumen. The ratios between inner and outer mean protein fluorescence intensity in the absence (-) and presence (+) of the protein-binding construct are presented in the form of three horizontal bar plots.

Chapter 3

Cation-responsive amphiphilic DNA nanostructures with non-canonical DNA G-Quadruplexes

3.1 Introduction

Biological cells display various degrees and hierarchies of compartmentalisation, through both membrane-bound [27] and membrane-less organelles [29, 349, 28]. These compartments enable cells to perform a variety of complex tasks, including, but not limited to, spatially organising macromolecules and metabolites [30, 350], handling processes and pathways [30, 351], as well as spatio-temporally controlling the rate of reactions by colocalising enzymes and substrates [30, 352, 351]. Differently from membrane-bound compartments, membrane-less organelles (MLOs) are very dynamic, both in their capability to quickly assemble and disassemble depending on external triggers (*e.g.* cellular stress for stress granules [350, 351]), and in the continuous exchange of their components with the surrounding environment [353, 354, 29]. Most MLOs, indeed, display a liquid-like behaviour, but some, both disease-related and physiological, are stiffer and hydrogel-like [355, 29, 356, 357, 349, 36]. It clearly follows that bottom-up assembled artificial cells require similar compartmentalisation capabilities [52]. As mentioned in Chapter 1, while membrane-bound compartments have been realised using lipid [41, 42], polymer [42, 44] or protein-based [17] membranes, MLOs often rely on coacervates [50, 51, 358, 359, 48] or hydrogels [359, 48, 199, 47, 49]. In this regard, DNA-based implementations have gained traction due to the superior biocompatibility, programmability and decreasing synthesis cost of DNA [136, 137]. Most DNA assemblies rely on canonical base pairing for structural programming, and on toehold-mediated strand displacement for actuation [166, 134, 360]. These materials can thus be easily made structurally responsive to the presence and concentration of nucleic acid molecules of specific sequences. However, for veritable biomimicry, it would be desirable for DNA-based synthetic

MLOs to reversibly assemble and disassemble in response to a broader range of biologically relevant, possibly physiological, stimuli. Such features would make the resulting DNA phases applicable beyond bottom-up synthetic biology, with avenues ranging from tissue engineering and cell scaffolding [361, 362] to small molecule biosensing [363, 364] and drug delivery and release [365, 366, 362].
In this Chapter, I combine the versatility, robustness and programmability of interactions of amphiphilic DNA nanostars [160, 163, 202, 154, 153, 367] with the cation-responsiveness of non-canonical DNA G-Quadruplex (G4) motifs to construct responsive hydrogel-like compartments capable of acting as synthetic organelles in nucleic acids-based artificial cells. I show that these hydrogel aggregates reversibly assemble and disassemble upon addition and chelation of G4-templating potassium (K^+) ions, and that their rate of formation can be tuned by varying the concentration of oligonucleotide components and/or available cations. Moreover, I demonstrate how the amphiphilicity of the building blocks allows for cargo uptake and triggered release of hydrophobic small molecules, producing a simplified mimic of dynamic target analyte enrichment upon a triggering stimulus often found in biological MLOs. Finally, I showcase how Quad-Stars hydrogels can undergo irreversible and highly localised photo-induced disassembly through the orthogonal mechanism of guanine photo-oxidation and cleavage upon near-UV irradiation in the presence of a G4-binding porphyrin photosensitiser [368–372].

The results and material presented in this chapter are based on the following Publication [373]:

G. Fabrini, A. Minard, R. A. Brady, M. Di Antonio and L. Di Michele, Cation-Responsive and Photocleavable Hydrogels from Noncanonical Amphiphilic DNA Nanostructures, *Nano Letters*, 22, 2, 602-611 (2022)

All the work presented in this Chapter was done by me, apart from measurements in Fig. 3.4 which were jointly performed with Dr. Aisling Minard (Di Antonio group). Moreover, DMEM + FBS buffers used in Fig. 3.13 were kindly prepared by Sabrina Pia Nuccio (Di Antonio group).

Sequences for DNA oligonucleotides used in this Chapter can be found in Table A.1 in the Appendix.

3.2 Quad-Star Design

The amphiphilic DNA nanostars, termed C-Stars, presented in Section 1.3.4 and developed by Brady *et al.* [160, 163, 202], are the basis for the cation-responsive DNA nanostructures presented in this Chapter. C-Stars, depicted in Fig. 3.1C, comprise four double-stranded DNA

(dsDNA) arms with tunable length around a central four-way junction. Each arm includes a strand with a covalently attached cholesterol moiety, enabling these nanostars to robustly self-assemble *via* non-specific hydrophobic interactions. Thanks to these interactions, C-Stars can form either crystalline frameworks or amorphous condensates depending on thermal annealing protocol, buffer conditions and design details [160, 163, 202].
Starting from the architecture of C-Stars, I developed modified amphiphilic DNA nanostructures comprising four, identical 28 base pairs (bp)-long dsDNA arms, each still terminating in a cholesterol moiety. I designed the arms to self-assemble into topologically analogous nanostars by means of a DNA G4, which replaces the original, canonical four-way junction. As such, I termed these nanostructures Quad-Stars.
As discussed in Section 1.2.2, DNA and RNA G-quadruplexes (G4s) can adopt a variety of topologies (parallel, anti-parallel, hybrid) and stoichiometries (unimolecular, bimolecular, tetramolecular). Despite this heterogeneity, they all consist of a varying number ($n \geq 2$) of stacked semi-planar tetrads of Hoogsteen hydrogen-bonded guanine residues, called G-quartets (or tetrads), arranged around an axially coordinated metal cation, as represented in Fig. 3.1A (top). Due to their tetravalent stoichiometry, Quad-Stars assemble *via* a tetramolecular parallel DNA G4, schematically depicted in Fig. 3.1A (bottom).

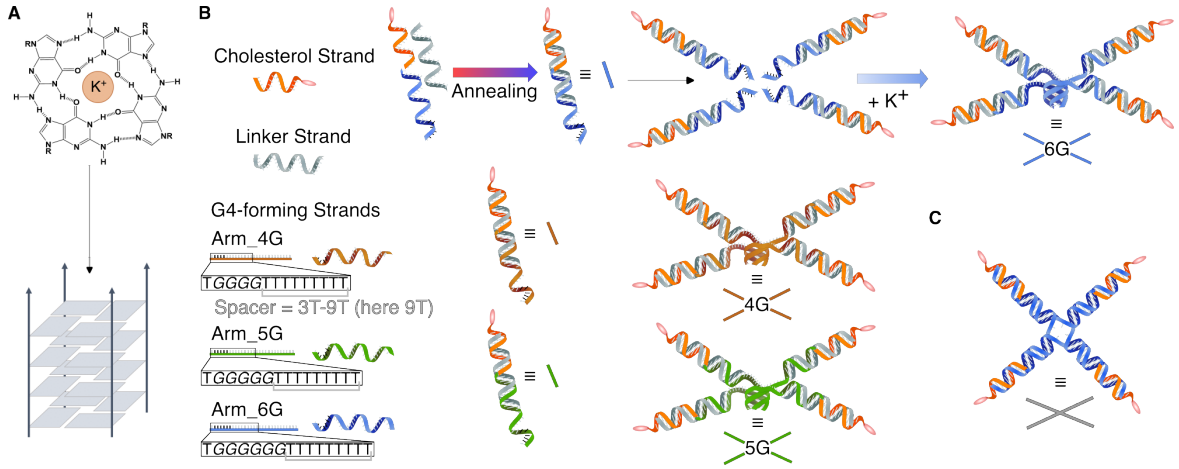


Fig. 3.1 **Design of cation-responsive amphiphilic DNA Quad-Stars. A:** Chemical structure of a G-quartet with a central coordinated K^+ ion (top) and simplified sketch of a parallel tetramolecular G-Quadruplex (G4) (bottom). **B:** Quad-Stars assemble from four identical 28 bp-long duplex arms, each comprising three strands: a cholesterol-bearing strand, a linker strand and a guanine-rich, G4-forming strand. During the annealing process, single-stranded DNA oligonucleotides hybridise into the double-stranded arms, but G4 formation and assembly of the designed Quad-Stars only occurs in K^+-rich environments. Three design variations are investigated, labelled as 4G (brown), 5G (green) and 6G (blue) displaying 4, 5 or 6 guanines in the G4-forming sequence, respectively. Moreover, several flexible oligo-thymine (oligo-T) spacers, separating the G4-forming region from the rest of the duplex arm, are tested, namely 3, 5, 7 and 9T, the latter (and longer) of which is schematically represented in this panel. Negative controls with 'no Gs', where guanines are replaced with thymines, are not represented. **C:** Schematic representation of C-Stars [160] with arm length $l = 28$ bp, identical to the one used in Quad-Stars. C-Stars and Quad-Stars share the same cholesterol-functionalised strands.

Fig. 3.1B includes a schematic representation of the DNA strands which make up Quad-Stars and the mechanism behind their self-assembly. Each of the four identical Quad-Star arms

self-assembles from three distinct DNA strands. The "Cholesterol Strand" bears a cholesterol moiety, attached *via* a triethylene glycol (TEG) spacer and is exactly identical to the one in C-Stars. The "Linker Strand" acts as a linker between the first and third strands. Finally, the "G4-forming Strand" terminates with a stretch of $n = 4$ (T GGGG T, "Arm_4G", brown), 5 (T GGGGG T, , "Arm_5G", green), or 6 (T GGGGGG T, , "Arm_6G", blue) consecutive guanines flanked by thymines, all sequences reported to self-assemble into tetramolecular parallel G4s under physiological conditions [329]. Since tetramolecular DNA G4s are bulkier and stiffer than canonical dsDNA, a flexible oligo-thymine (oligo-T) spacer was included between the duplex portion of the arm - formed by the cholesterol-bearing and the linker strands - and the G4-forming guanine stretch. In particular, oligo-T spacers with lengths of 3, 5, 7 and 9 nucleotides (nt) were tested. Accordingly, a generic construct with *n* consecutive guanines and a *m* nt-long oligo-T spacer will be referred to as *n*G*m*T, *e.g.* a construct with 6 guanines and a 9 thymine-long spacer is referred to as 6G9T.
As discussed in Section 1.2.2, metal cations are critical to the formation of G4s [91–93], with K^+ ions exerting the strongest stabilisation effect, followed by Na^+ and Mg^{2+}, while Li^+ ions do not contribute at all. Therefore, the sensitivity to cationic environment of DNA G4s should grant Quad-Stars a degree of cation-responsiveness. Similarly to C-Stars, Quad-Stars are expected to form nanoporous networks owing to the clustering of cholesterol moieties into micelle-like assemblies, coordinated and linked by the tetravalent DNA nanostar motifs. Unlike C-Stars, though, the replacement of the irreversible, canonical four-way junction with a tetramolecular DNA G4 should enable Quad-Stars and Quad-Star networks to respond to changes in ionic environment. In particular, Quad-Stars are designed to self-assemble at sufficiently high concentration of G4-promoting K^+ ions – as well as Na^+ and Mg^{2+} to a lower extent –, while failing to do so in the presence of other metal cations displaying lower to null affinity for the G4 microenvironment, typically exemplified by Li^+. As the folding and stability of G4s also depend on the number of stacked G-quartets – the more, the better [329] –, Quad-Star variants with longer guanine-stretches, *i.e.* 5G and 6G compared to 4G, are expected to self-assemble more promptly and with higher efficiency.

3.3 Characterisation

In this section, I will first go over the characterisation of non-cholesterolised nanostructures, where the cholesterol-tagged strands are replaced with non-functionalised ones of identical sequence but lacking the cholesterol moiety. This effectively prevents the nanostars from interacting with one another, thus making them soluble and enabling their characterisation with a variety of conventional molecular biology techniques. After that, I will move onto the characterisation of full cholesterolised Quad-Stars.

3.3.1 Characterisation of non-cholesterolised Quad-Stars

I first set out to probe the self-assembly of non-cholesterolised Quad-Stars into the desired branched architectures, *i.e.* four-way nanostars, at thermodynamic equilibrium. To this aim, samples of 4G, 5G and 6G designs with all possible oligo-T spacer lengths were slowly annealed in TE buffer containing 300 mM LiCl (non G4-promoting) or 300 mM KCl (G4-promoting) and characterised through a combination of agarose gel electrophoresis, Dynamic Light Scattering (DLS) and Circular Dichroism (CD). The results are presented in Fig. 3.2, 3.3 and 3.4.

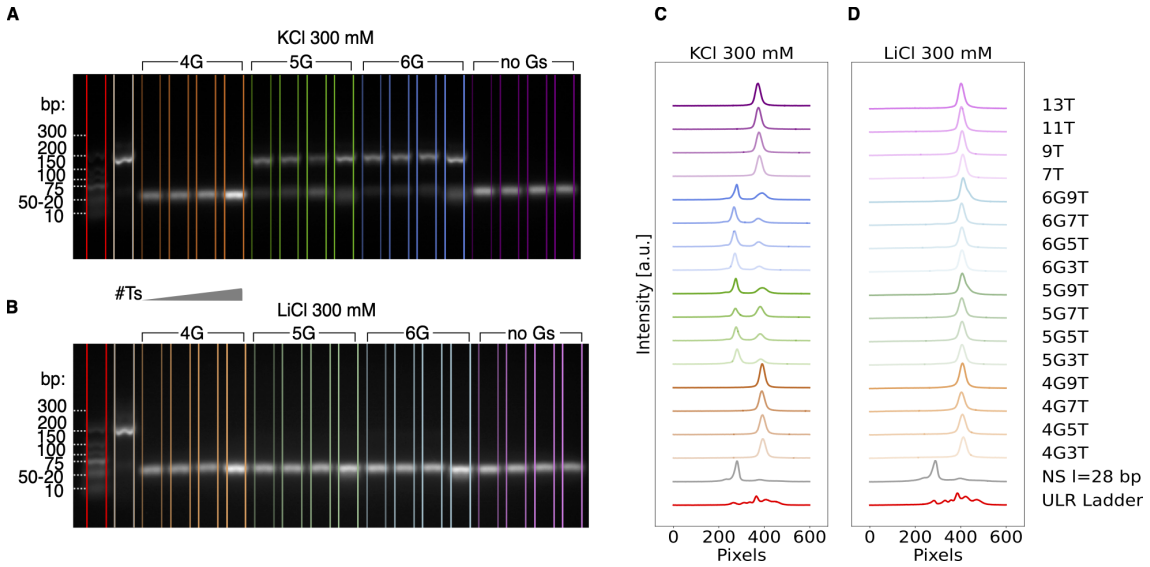


Fig. 3.2 **Self-assembly of non-cholesterolised Quad-Star variants *via* agarose gel electrophoresis. A, B:** Agarose gels of all investigated Quad-Star motifs annealed in either 300 mM KCl (**A**) or LiCl (**B**). Ultra-Low-Range (ULR) DNA reference ladder is used as electrophoresis control (red). Four-way nanostars (NS) with 28 bp arm length annealed in 300 mM NaCl are used as mobility reference for correctly self-assembled nanostars, *i.e.* positive control (light gray). Spacer lengths ("#Ts") increase from left (3T) to right (9T) within each sample group (4G, 5G, 6G and no Gs), as schematically depicted. **C, D:** Intensity profiles extracted from **A** and **B**, respectively. Pixels are numbered from top to bottom, along the height of the gel images. Visual inspection of the gels and the extracted lane intensity profiles reveals no nanostar motif self-assembles in 300 mM LiCl (**B**), while only 5G and 6G samples, at all oligo-T spacer lengths, correctly self-assemble in K^+-rich environments, indicated by slower migrating bands, completely absent in gel **B**, with mobility comparable to that of the NS control. Negative control samples with no Gs (7-13T) are incapable of forming G4s, and thus assembling into nanostars. Despite the potential for G4 formation, 4G samples fail to self-assemble at any spacer length. Samples for all illustrated measurements were prepared at 20 μM oligonucleotide concentration.

Influence of guanine stretch length

Agarose gels in Fig. 3.2A-B and related line profiles in Fig. 3.2C-D demonstrate that longer guanine stretches, capable of forming G4s with a higher number of stacked tetrads, grant Quad-Stars a higher self-assembly efficiency. In G4-promoting conditions, namely 300 mM KCl (Fig. 3.2A and C), both 5G (green lanes) and 6G (blue lanes) samples correctly self-assemble at all oligo-T spacer lengths, as evidenced by the occurrence of slower migrating

bands which are absent in Li^+-rich environments (Fig. 3.2B and D). Such slower bands, in fact, exhibit mobilities in excellent agreement with that of positive controls (between 150 and 200 bp markings of the reference ladder), namely regular four-way nanostars with the same arm length (l = 28 bp, light gray lane), thus indicating correct self-assembly into the designed topology. Conversely, all 4G samples (orange/brown lanes) fail to self-assemble in both cationic environments, rather they remain as disjoined duplex arms (between 50 and 75 bp markings of the reference ladder), as demonstrated by the corresponding bands, which display the same mobility as those of negative controls incapable of forming G4s (no Gs, purple lanes).

The volume-based size distributions obtained *via* DLS for all investigated motifs annealed in 300 mM KCl in Fig. 3.3A-B support the agarose gels in Fig. 3.2A, displaying a marked shift (from ≈4-5 nm to ≈6-8 nm) in mean hydrodynamic diameter ($< d_H >$) between correctly self-assembled 5G and 6G samples, and disjoined duplex arms in 4G constructs and no Gs negative controls (as in Fig. 3.2). Results in Fig. 3.3C further confirm the K^+-dependent self-assembly of 5G9T and 6G9T samples into the designed nanostars. When annealed in 300 mM KCl, in fact, they display $< d_H >$ values comparable to that of correctly assembled canonical four-way DNA nanostars with the same arm length (light gray), used as positive control similarly to Fig. 3.2. However, they fail to self-assemble in K^+-poor cationic conditions (300 mM LiCl), and rather exhibit $< d_H >$ values comparable to those of 4G designs and no Gs negative controls in Fig. 3.3A-B.

It should be noted that the G4-forming domains in 4G designs have previously been reported to form tetramolecular parallel G4s under physiological conditions [329], that is with lower K^+ ion availability compared to the conditions here investigated (300 mM KCl). However, such reports only examine the G4-forming domains themselves, which terminate on both sides in single flanking thymines, *i.e.* T GGGG T. In Quad-Stars, instead, each of those G4-forming domains is flanked on one side by a much bulkier 28 bp-long dsDNA arm, thus considerably increasing the steric and Coulomb repulsion experienced upon G4 formation. Such increase in repulsion is thus most likely enough to destabilise 4G designs, but not 5G or 6G ones, as G4 stability increases with higher numbers of stacked G-quartets.

Influence of oligo-T spacer length

Flexible oligo-T spacers with variable length were tested to see whether allowing for more flexibility between the stiffer tetramolecular G4 motif and the duplex arms would afford a higher self-assembly efficiency. Results in Fig. 3.2 and 3.3, combined, show that the influence of such flexible spacers is negligible, as confirmed by the summarised mean hydrodynamic diameters ($< d_H >$) for all sample variants *vs* the length of the oligo-T spacer in Fig. 3.3B.

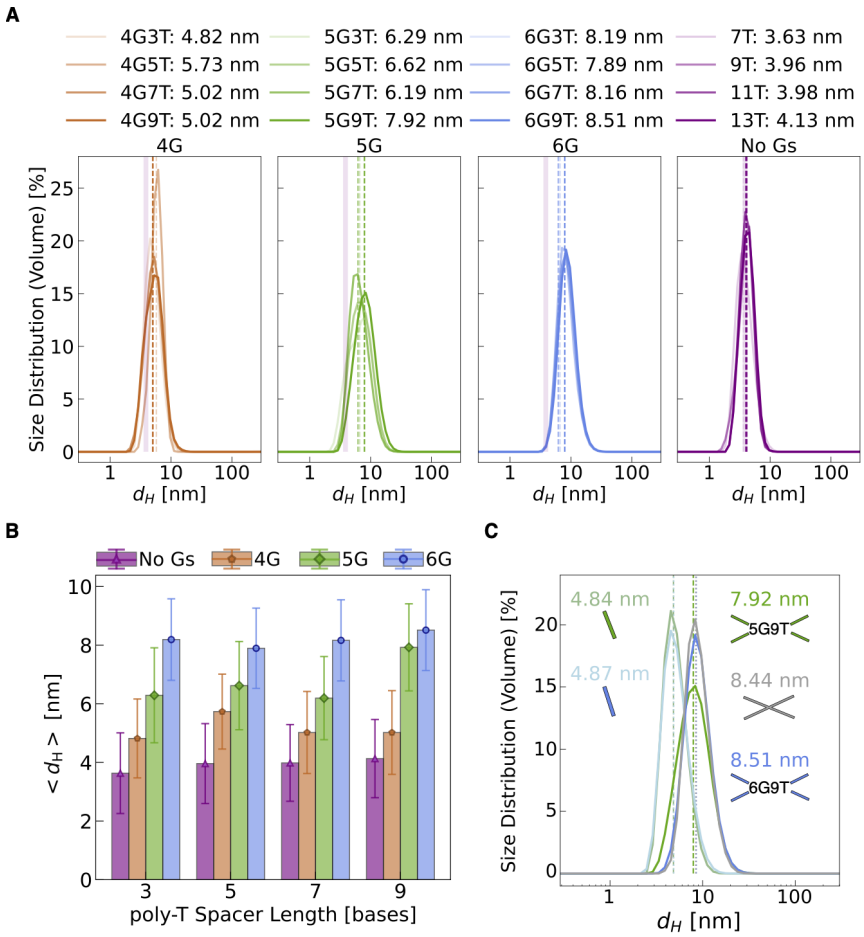


Fig. 3.3 **Self-assembly of non-cholesterolised Quad-Star variants *via* Dynamic Light Scattering (DLS). A:** Volume-based size distributions obtained *via* DLS for all Quad-Star variants annealed in 300 mM KCl support results from Fig. 3.2. Faint pink shaded regions span the range of the means of "No Gs" negative controls. **B:** Summary of $< d_H >$ computed from **A** for all variants annealed in 300 mM KCl against length of oligo-T spacers confirms a negligible influence of the spacer length on the self-assembly of Quad-Stars, as anticipated by Fig. 3.2. Data are presented as $< d_H >$ (bars and markers) $\pm$ standard deviation from covariance matrix of the fit (error bars). **C:** Volume-based size distributions obtained *via* DLS for 5G9T and 6G9T Quad-Star variants annealed in 300 mM LiCl or 300 mM KCl compared to correctly self-assembled four-way DNA nanostars with arm length l = 28 bp annealed in 300 mM NaCl (positive control). Combined, these data confirm that only 5G and 6G Quad-Stars (at all oligo-T spacer lengths) self-assemble in K^+-rich buffers, displaying average hydrodynamic diameters ($< d_H >$) close to that of canonical four-way nanostars ($< d_H >= 8.44$ nm, light gray in **C**). Even in K^+-rich buffers, in fact, samples with no Gs or 4G display consistently smaller $< d_H >$ values, close to those of 5G and 6G motifs in absence of K^+ ions (300 mM LiCl buffer, fainter profiles in **C**), as expected from disjoined dsDNA arms. Both in **A** and **C**, vertical dashed lines represent mean hydrodynamic diameters ($< d_H >$, also reported in the legends on top) obtained by fitting average profiles with lognormal distributions (fits not shown). Profiles in **A**, **C** (solid lines) are averages of at least three independent measurements. Samples for all illustrated measurements were prepared at 20 μM oligonucleotide concentration.

Circular Dichroism confirms G4-mediated self-assembly

Electrophoretic mobilities (Fig. 3.2) and average hydrodynamic diameters (Fig. 3.3) show a clear influence of the cationic environment on the self-assembly, or lack thereof, of Quad-Stars. Yet, they do not prove that self-assembled Quad-Stars comprise tetramolecular parallel G4 motifs. Circular Dichroism (CD) measurements in Fig. 3.4 unambiguously support the

self-assembly results in Fig. 3.2 and 3.3 and prove that the self-assembly of Quad-Stars is K^+-dependent due to being mediated by the formation of a parallel G4.

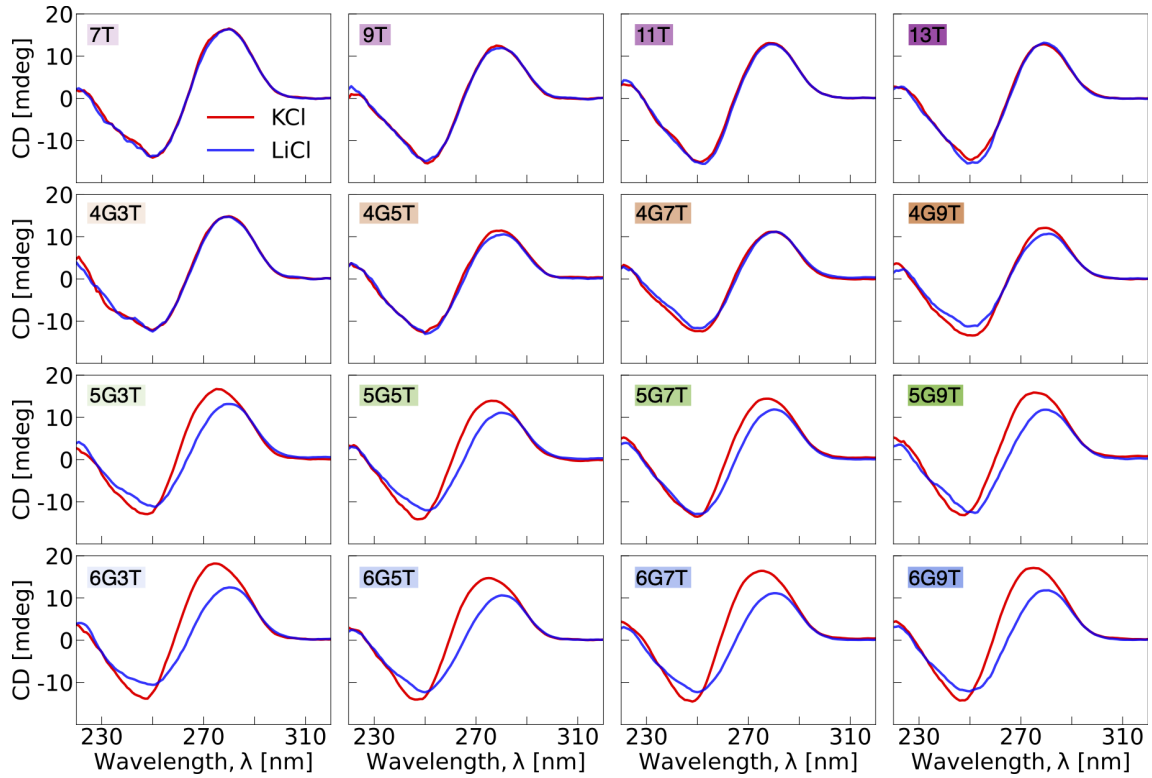


Fig. 3.4 **Circular Dichroism (CD) spectra confirm G4-mediated self-assembly.** Negative controls with guanines replaced by thymines (no Gs, first row) and 4G samples (second row) present the same signature in both 300 mM LiCl (blue) and KCl (red), with only a slight hint of shift in maximum and minimum appearing for 4G9T, thus confirming no G4 formation and no self-assembly as per results in Fig. 3.2 and 3.3. Only 5G (third row) and 6G (fourth row) samples display the clear, parallel G4 signature in KCl, with characteristic shift in maximum and minimum towards ≈ 263 and 240-245 nm, respectively, proving K^+-dependent assembly of Quad-Stars *via* tetramolecular parallel G4 formation, as per design. Quantity reported on the *y* axis and labelled "CD" is absolute (*i.e.* non molar-corrected) ellipticity of polarisation. All samples were prepared at 20 μM oligonucleotide concentration.

Negative controls (no Gs, first row) and 4G samples (second row) display the same CD signature in both 300 mM LiCl (blue) and KCl (red), with the exception of 4G9T, which starts displaying a slight shift in its maximum and minimum. All 5G (third row) and 6G (fourth row) designs, instead, regardless of oligo-T spacer length, display the clear, parallel G4 CD signature, with a characteristic shift in maxima and minima towards ≈ 263 and 240-245 nm, respectively, also accompanied by signal enhancement [374, 342, 341, 375]. Due to the conclusive evidence of no Gs and 4G samples failing to self-assemble upon annealing in K^+-rich buffers, only 5G and 6G motifs will be considered in the following sections.

G4 assembly kinetics *via* NMM fluorimetry

All measurements presented so far demonstrate that Quad-Stars self-assemble in a G4-mediated fashion when annealed in K^+-rich buffers, but do not provide any evidence regard-

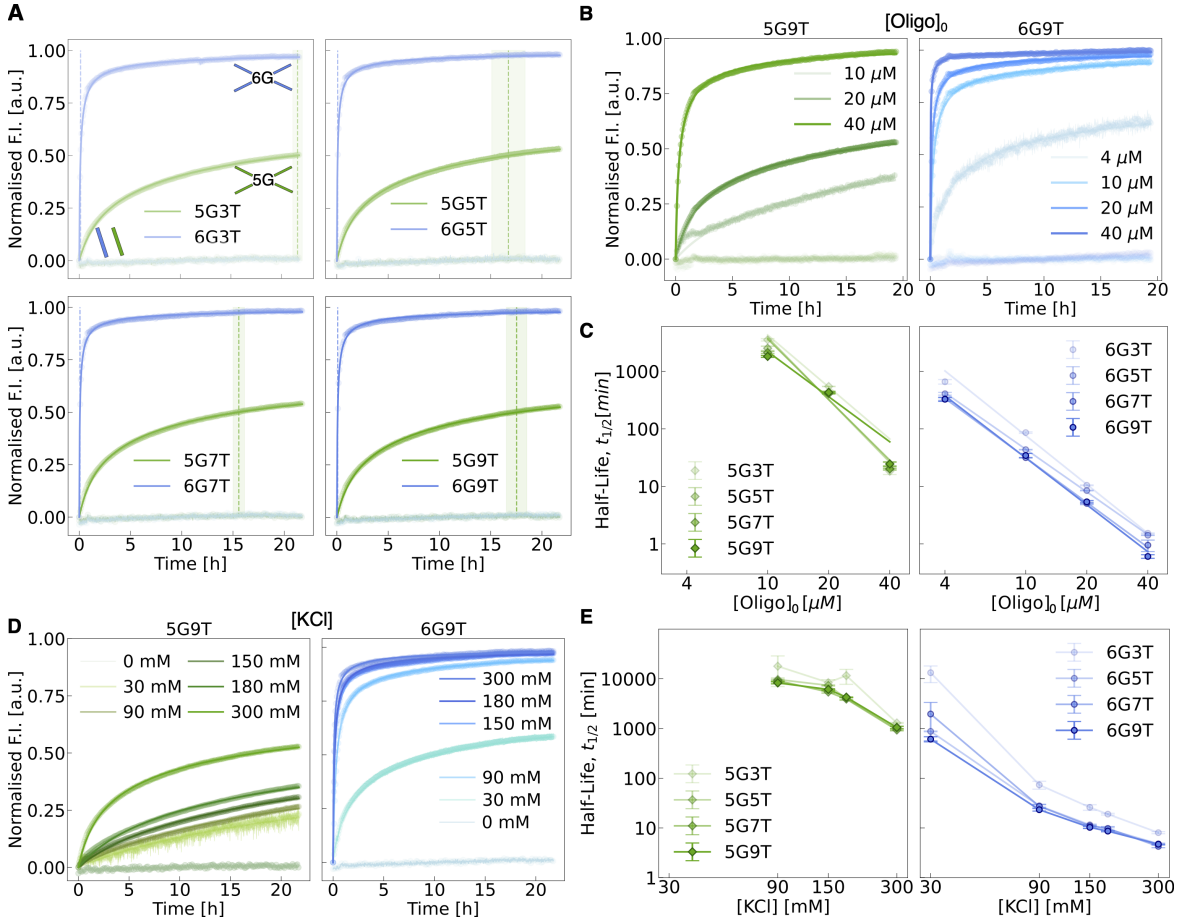


Fig. 3.5 **Kinetics of G4-dependent self-assembly of non-cholesterolised Quad-Stars *via* NMM fluorimetry. A:** Kinetic profiles for 5G and 6G samples, for all oligo-T spacer lengths, at target conditions, *i.e.* $[Oligo]_0$ = 20 μM, [KCl] = 300 mM, where $[Oligo]_0$ = [oligonucleotide] at $t = 0$ s. Disjoined duplex arms (start, annealed in Li^+-rich buffers) assemble into the sought nanostar topology (as schematised in the top right panel) upon KCl injection *via* tetramolecular G4 motifs, the formation of which enables NMM binding and fluorescence emission enhancement. **B:** Profiles for Quad-Stars with 9T spacers for all examined $[Oligo]_0$. **C:** Half-life of formation extracted from profiles in **B** (and analogous for shorter oligo-T spacer lengths, see Fig. B.2A) *vs* $[Oligo]_0$ at fixed [KCl] = 300 mM. Data have been linearly fitted (solid lines) in log-log scale, and the reaction order n extracted from the fitted slope, resulting in the following values: 4.005 $\pm$ 0.747 for 5G3T, 4.613 $\pm$ 0.667 for 5G5T, 4.501 $\pm$ 0.787 for 5G7T, 3.605 $\pm$ 0.740 for 5G9T, 3.823 $\pm$ 0.086 for 6G3T, 3.467 $\pm$ 0.039 for 6G5T, 3.590 $\pm$ 0.030 for 6G7T, 3.711 $\pm$ 0.127 for 6G9T. Reported uncertainties represent standard deviations from the covariance matrix of the fit. **D:** Profiles for Quad-Stars with 9T spacers for all examined [KCl]. **E:** Half-life of formation from profiles in **D** (and analogous for shorter oligo-T spacer lengths, see Fig. B.2B) *vs* [KCl] at fixed $[Oligo]_0$ = 20 μM. In **A**, **B** and **D**, curves with saturated colours (top) correspond to samples triggered with KCl injection, whereas fainter curves (bottom) are negative controls (no KCl). For samples, data are reported as mean (dots, one every 3 datapoints shown) $\pm$ standard deviation (shaded regions) and fits to theoretical tetramolecular kinetics (solid lines) [329]. For negative controls, only average, normalised datapoints are shown (see Section 2.8.2). Vertical lines are mean half-life values, $t_{1/2}$, with the surrounding shaded regions depicting the standard deviations. Data in **C** and **E** are reported as mean (markers) $\pm$ standard deviation (error bars). Data for negative controls, as well as 6G5T and 6G7T at 30 mM KCl have been averaged across independent duplicates. Remaining data have been averaged across at least three independent repeats (see Section 2.8.2).

ing their capability to do so isothermally in response to changes in the ionic environment, namely to the addition of G4-promoting K^+ ions. In order to assess this, non-cholesterolised Quad-Stars were first annealed in 300 mM LiCl, resulting in disjoined duplex arms, and then their assembly was triggered *via* rapid injection of KCl. The self-assembly transient was monitored by time-resolved fluorescence using N-methyl Mesoporphyrin IX (NMM), a porphyrin ligand whose quantum yield is dramatically enhanced upon preferential binding through π-π stacking to parallel G4s [376–378]. The molecular structure of NMM and its emission spectra are provided in the Appendix (Fig. B.1). It is worth noting that, since NMM is a G4 ligand, the resulting stacking interactions are expected to stabilise G4s within Quad-Stars. However, this would be true for any other G4-binding ligand enabling fluorescence monitoring of G4 formation. Moreover, this still enables relative comparisons of assembly kinetics, rather than absolute evaluations, as all samples are measured in the presence of NMM.

As presented in Fig. 3.5A, the resulting fluorescence intensity kinetic profiles monotonically increase after KCl injection, demonstrating NMM binding to formed G4s, thus reflecting the isothermal assembly of Quad-Stars in response to the desired change in cationic environment. Previous reports on the kinetics of tetramolecular G4 formation have proved that their assembly can be modelled as an irreversible reaction due to their considerable stability, *i.e.* long lifetimes [379, 329, 380], *post* formation, as in

$$4\,Oligo \xrightarrow{k_{on}} Q \tag{3.1}$$

with *Oligo* being the monomer oligonucleotide strand, Q the assembled tetramolecular G4, and k_{on} the folding/association rate constant. The concentration of monomer strand, $[Oligo]$, thus obeys the following ordinary differential equation:

$$\frac{d[Oligo]}{dt} = -k_{\mathrm{on}}[Oligo]^n \tag{3.2}$$

where t indicates time, and n is the reaction order. While Equation 3.1 would suggest an instantaneous tetramolecular reaction (n = 4), a four-body collision is highly unlikely. The aforementioned studies [379, 329, 380] have, indeed, reported n values ranging between 3 and 4 depending on the stability of (much shorter-lived) reaction intermediates with varying environmental conditions, particularly cation identity [379, 380]. For this reason, the reaction order is treated as a variable in Equation 3.2, rather than a known constant. The solution of Equation 3.2 provides a simple kinetic model to which the fluorescence intensity kinetic profiles in Fig. 3.5A can be fitted [329], namely

$$\alpha_{\mathrm{f}} = 1 - \alpha_{\mathrm{u}} = 1 - \frac{[Oligo]}{[Oligo]_0} = 1 - \{1 + [Oligo]_0^{(n-1)}\,(n-1)\,k_{\mathrm{on}}\,t\}^{\frac{1}{1-n}} \tag{3.3}$$

where α_{f} is the fraction of folded strands, *i.e.* participating in G4s, α_{u} is the fraction of unfolded (monomeric) strands, and $[Oligo]_0$ is the starting monomer concentration. Due

to fluorescence being proportional to the amount of formed G4s NMM can bind to, the recorded fluorescence intensity (F.I.) can be fitted, after the necessary pre-processing (see Section 2.11.3), as F.I.$= a \cdot \alpha_{\mathrm{f}}$, where a is a scaling parameter optimised through fitting that accounts for kinetic profiles potentially not reaching completion within the experimental time window. Precisely for this reason, assembly half-times, $t_{1/2}$, cannot be computed directly from the data, rather they must be extracted from the fitted k_{on} and n by computing the time needed to reach $\alpha_{\mathrm{f}} = \frac{1}{2}$, leading to

$$t_{1/2} = \frac{\left(\frac{1}{2}\right)^{1-n} - 1}{(n-1)\, k_{\mathrm{on}} \left[Oligo\right]_0^{(n-1)}}. \tag{3.4}$$

The half-life values obtained at the target conditions of $[Oligo]_0 = 20\,\mu$M and [KCl] = 300 mM clearly demonstrate that longer G-runs not only enhance G4 stability and promote higher folding efficiency (see Fig. 3.2), they also translate into faster assembly, as depicted in Fig. 3.5A. For instance, $t_{1/2} \sim 5$ min for 6G9T and $t_{1/2} \sim 15$ h for 5G9T, with similar discrepancies at all oligo-T spacer lengths. Such trend is in agreement with several literature reports [329, 380, 381], and it can be promptly rationalised in light of a nucleation-zipping assembly pathway [379]. The nucleation-zipping model [382] was originally proposed to explain experimental observations of negative apparent activation energies in the formation of N bp-long A_N-T_N dsDNA constructs, *i.e.* higher hybridisation rates at lower temperature. The model considers a rapid pre-equilibrium between the single-stranded oligonucleotides and an unstable dsDNA intermediate with $(N-1)$ bp, followed by the limiting step of nucleating the N^{th} base pair [383–385]. While later studies have highlighted that dsDNA hybridisation is a non-Arrhenius process and that the apparent activation energy is positive, albeit small, below the melting temperature [386, 387], variations of this model have been proposed to explain the apparent negative activation energy of association of tetramolecular quadruplexes. These either rely on a pre-equilibrium between monomers and dimers followed by a dimer/quadruplex limiting step [379, 388], or on a multitude of intermediates comprising dimers, trimers and tetramers with slipped strands [389], which ultimately serve to avoid the "all or nothing model" requiring the aforementioned unlikely four-body collision. Importantly, it follows from such models that longer stretches of consecutive guanines bestow increased stability upon transient dimers [388] and/or trimers [389] intermediates for G4 assembly, thus boosting the overall folding kinetics.

While n values are obtained as fitting parameters from Equation 3.3, these estimates are not guaranteed to hold throughout the whole range of starting monomer concentrations, $[Oligo]_0$, as it is instead expected of the reaction order. However, it follows from Equation 3.4 that assembly half-life values should follow a power-law behaviour such as

$$t_{1/2} = A \cdot \left[Oligo\right]_0^{1-n} \tag{3.5}$$

from which the reaction order, n, (as well as the constant A) can be more reliably estimated through fitting.

Figure 3.5B-C showcase the dependency of the kinetic profiles and the extracted $t_{1/2}$ values on $[Oligo]_0$. Expectedly, NMM fluorescence increases faster and plateaus earlier at higher $[Oligo]_0$ in both 5G9T and 6G9T samples, and $t_{1/2}$ values follow a decreasing power-law with increasing $[Oligo]_0$. Notably, fitting yields the following values for the reaction order: $n = 3.605 \pm 0.740$ for 5G9T and $n = 3.711 \pm 0.127$ for 6G9T, both closely matching the theoretical stoichiometry of 4. Similar trends can be observed at all oligo-T spacer lengths in Figure 3.5C. Notably, tetramolecular processes are rare in nucleic acids nanotechnology, where most reactions are bimolecular in nature. In particular, the higher reaction order bestows upon Quad-Star formation a much more dramatic dependence on the starting oligonucleotide concentration. In a bimolecular process, varying the starting monomer concentration by a factor of 4 ($10 \rightarrow 40\,\mu$M) would simply reduce the half-life of formation by the same factor ($t_{1/2} \propto [Oligo]_0^{-1}$), whereas in the case of Quad-Stars the half-life of formation drops by a factor of ≈ 64 ($t_{1/2} \propto [Oligo]_0^{-3}$).

G4 association rates have also been reported to highly depend on the identity and concentration of cations, increasing with increasing ion concentration with a rate that depends on the extent of stabilisation afforded by the specific ions, *e.g.* higher for K^+ than for Na^+ [379]). Therefore, I proceeded to investigate whether the self-assembly rate of Quad-Stars could be tuned by varying the concentration of injected K^+ ions. Figure 3.5D-E prove that, indeed, 5G and 6G Quad-Stars fold significantly faster in the presence of higher K^+ ion concentrations ([KCl]), with $t_{1/2}$ values dropping from ~ 10 h to ~ 5 min when $[K^+]$ is increased from 30 to 300 mM for 6G9T. A similarly strong dependence was found for 5G9T. Comparable trends are found to hold at all other oligo-T spacer lengths (see Fig. B.2 in the Appendix for corresponding profiles). It is worth mentioning that, due to the much slower formation rates, assembly kinetics of 5G designs at all oligo-T spacer lengths could not be characterised at the lowest oligonucleotide ($4\,\mu$M) and KCl (30 mM) concentrations used for 6G constructs, as it can be noted in Fig. 3.5E.

Overall, the combined behaviours displayed in Figure 3.5B-E demonstrate the capability of Quad-Stars to isothermally assemble in a cation- and oligonucleotide concentration-dependent fashion, meaning the assembly rates can be widely tuned by varying potassium and/or oligonucleotide concentrations.

Due to the negligible impact of the length of the flexible oligo-T spacer, as substantiated by results in Figs. 3.2, 3.3, 3.4 and 3.5, only 5G and 6G samples with the longest 9T spacer, *i.e.* 5G9T and 6G9T, will be examined from here onwards, and, for brevity, they shall simply be referred to as 5G and 6G.

Influence of monovalent cations: Li^+ *vs* Na^+ *vs* K^+

To assess the sensitivity and specificity of the self-assembly of Quad-Stars with respect to the cationic environment, two concentrations, namely 300 mM (as in previous measurements) and 100 mM, of monovalent cations Li^+, Na^+ and K^+ were tested, as showed in Fig. 3.6. The lower 100 mM concentration was chosen to more closely resemble physiologically-relevant ionic compositions, with intracellular K^+ and extracellular Na^+ ion concentrations both amounting to $\approx$ 100-150 mM. As agarose gel electrophoresis (Fig. 3.2), DLS (Fig. 3.3), CD (Fig. 3.4) and fluorimetry (Fig. 3.5) measurements have all demonstrated the lack of Quad-Star formation in both 5G and 6G non-cholesterolised samples annealed in 300 mM LiCl buffers, this condition will be used as negative control.
Agarose gel electrophoresis (Fig. 3.6A-B) and DLS measurements (Fig. 3.6C) reveal that samples annealed for 6 h (see Methods) in 300 mM NaCl exhibit a lower degree of self-assembly compared to KCl at the same concentration. Only 6G samples, in fact, exhibit the delayed, larger-sized band and hydrodynamic diameter shift corresponding to correctly assembled nanostars. When reducing the concentration to 100 mM, self-assembly is only observed for 6G motifs in 100 mM KCl. The apparent inconsistency with fluorimetry results for 5G samples in 100 mM KCl – with visible fluorescence signal in Fig. 3.5D, hinting at G4 formation and Quad-Star assembly – is likely due to the difference among annealing duration (6 h) and the time window monitored during fluorescence kinetics (tens of hours). It is also worth noting that the efficiency of Quad-Star formation, both 5G and 6G, in 300 mM KCl observed in Fig. 3.6A-B is considerably lower than the one displayed in Fig. 3.2 (as well as Fig. 3.7A-B below), *i.e.* the slower migrating band indicative of correctly assembled Quad-Stars is much fainter than the faster one corresponding to disjoined duplex arms. This might be due to issues with sample preparation, such as aggregation of G4-forming strands during storage prior to mixing and annealing.
These results point to the degree of quadruplex stabilisation afforded by Na^+ ions being intermediate between Li^+ and K^+, in agreement with literature [91], where Na^+ ions are often reported as being more effective in the stabilisation of anti-parallel or hybrid topologies [390] rather than the parallel one under investigation in this Chapter.

Influence of divalent cations: Li^+ *vs* Mg^{2+} *vs* K^+

Similar results were obtained when annealing non-cholesterolised Quad-Stars in 2 mM $MgCl_2$, tested due to the physiological relevance of Mg^{2+} ions, as well as their frequent use in DNA nanotechnology. Fig. 3.7 shows that 2 mM Mg^{2+} ions stabilise the G4 motifs in Quad-Stars to an extent comparable to 300 mM Na^+ and substantially lower than 300 mM K^+, with 5G designs failing to self-assemble and 6G ones displaying a fainter delayed band compared to 300 mM K^+ (easily observed in Fig. 3.7B). This points to a lower self-assembly efficiency in Mg^{2+}, *i.e.* lower fraction of correctly self-assembled Quad-Stars, compared to

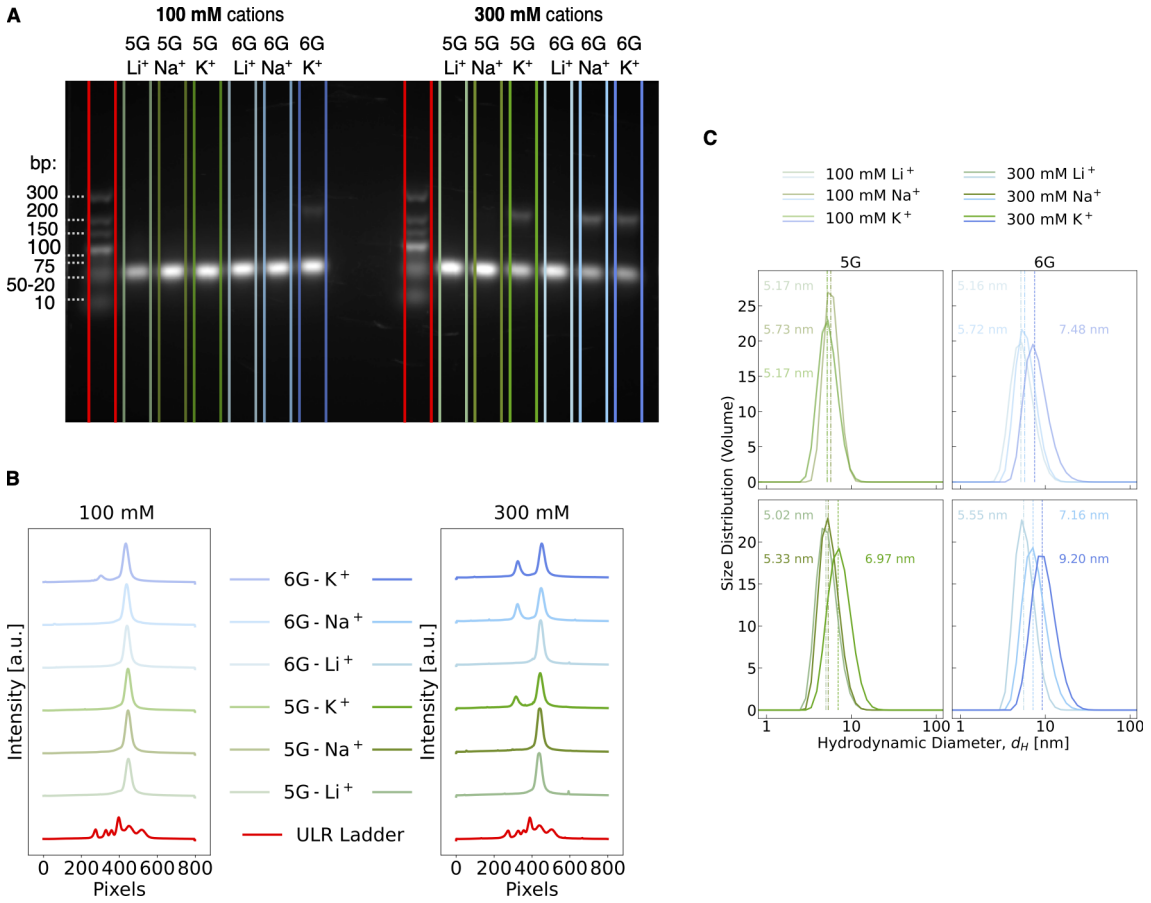


Fig. 3.6 **Influence of monovalent cations on the self-assembly of non-cholesterolised Quad-Stars. A:** Agarose gel comparing electrophoretic mobilities of 5G (green) and 6G (blue) Quad-Stars (both with 9T spacer) annealed in either 100 (right half) or 300 mM (left half) LiCl, NaCl or KCl, respectively. Ultra-Low-Range (ULR) DNA reference ladder is used as electrophoresis control (red). **B:** Intensity profiles extracted from agarose gel in **A**. Pixels are numbered from top to bottom, along the height of the gel images. Visual inspection of the gel shows that Na^+ ions support the self-assembly of Quad-Stars to a lesser extent than K^+ ones, with 6G Quad-Stars self-assembling only at the highest concentration (300 mM), and 5G constructs failing to do so at both examined concentrations. **C:** Volume-based size distributions obtained *via* DLS for samples in **A**. Average hydrodynamic diameters ($< d_H >$), obtained *via* fitting to lognormal distributions (fits not shown) are indicated by vertical lines. Reported $< d_H >$ values lie within one standard deviation from the ones presented in Fig. 3.3. Profiles (solid lines) are averages of at least four independent measurements. Mean hydrodynamic diameters and corresponding standard deviations (from covariance matrix of the fit) are hereby reported: 5G – 5.17 ± 1.30 nm in 100 mM LiCl, 5.73 ± 1.24 nm in 100 mM NaCl, 5.17 ± 1.30 nm in 100 mM KCl, 5.02 ± 1.31 nm in 300 mM LiCl, 5.33 ± 1.30 nm in 300 mM NaCl, 6.97 ± 1.37 nm in 300 mM KCl; 6G – 5.16 ± 1.35 nm in 100 mM LiCl, 5.72 ± 1.32 nm in 100 mM NaCl, 7.48 ± 1.36 nm in 100 mM KCl, 5.55 ± 1.30 nm in 300 mM LiCl, 7.16 ± 1.36 nm in 300 mM NaCl, 9.20 ± 1.38 nm in 300 mM KCl. All samples were prepared at 20 μM oligonucleotide concentration.

300 mM K^+, as confirmed by the bimodal size distribution obtained *via* DLS exhibited by 6G in 2 mM $MgCl_2$ depicted in Fig 3.7C. The downward shift in the larger sized peak can likely be ascribed to a higher degree of compactness, as often observed with DNA nanostructures in the presence of divalent Mg^{2+} ions. Analogously to Fig. 3.6 above, samples annealed in 300 mM LiCl are here used as negative controls.

Overall, the self-assembly, or lack thereof, of non-cholesterolised Quad-Stars in different

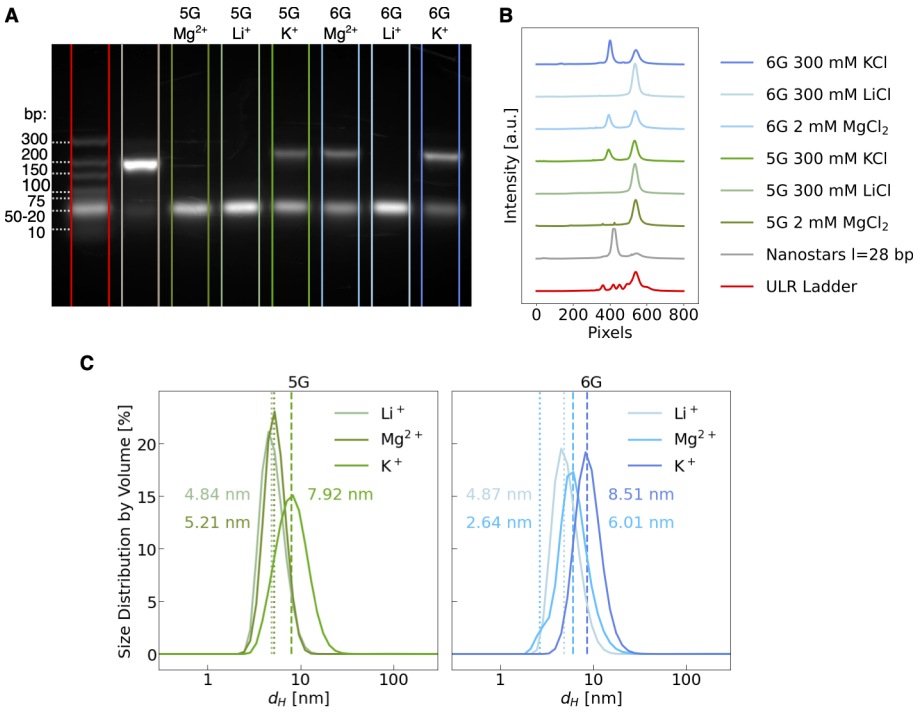


Fig. 3.7 **Influence of Mg^{2+} ions on the self-assembly of non-cholesterolised Quad-Stars. A:** Agarose gel comparing electrophoretic mobilities of reference four-arm nanostars with arm length l = 28 bp in 300 mM NaCl (light gray), 5G (green) and 6G (blue) Quad-Stars (both with 9T spacer) in 2 mM $MgCl_2$, 300 mM LiCl or 300 mM KCl, respectively. Ultra-Low-Range (ULR) DNA reference ladder is used as electrophoresis control (red). **B:** Intensity profiles extracted from agarose gel in **A**. Pixels are numbered from top to bottom, along the height of the gel images. Visual inspection of the gel demonstrates that 2 mM Mg^{2+} ions exhibit a degree of G4 stabilisation intermediate between Li^+ and K^+ ones, similarly to Na^+ ions, with 5G designs completely failing to self-assemble, and 6G designs showing a fainter slower migrating band, signalling a lower fraction of correctly assembled 6G Quad-Stars in $MgCl_2$ compared to KCl. **C:** Volume-based size distributions obtained *via* DLS for samples in **A**. Vertical lines show average hydrodynamic diameters ($< d_H >$), obtained *via* fitting to mono- or bi-modal lognormal distributions – fits not shown. As observed in **A**, 5G motifs fail to self-assemble in 2 mM Mg^{2+} ions, exhibiting $< d_H >$ similar to that in Li^+ buffers (disjoined duplex arms). Conversely, 6G in Mg^{2+} presents a bimodal distribution, with the larger-sized peak shifted towards the size of fully assembled Quad-Stars, albeit slightly smaller, possibly due to Mg^{2+}-induced compaction. While the smaller-sized peak for 6G is slightly shifted to lower size values compared to free dsDNA arms in Li^+ buffers, the two $< d_H >$ estimates lie within one standard deviation. Profiles (solid lines) are averages of at least five independent measurements. Mean hydrodynamic diameters and corresponding standard deviations (from covariance matrix of the fit) are hereby reported: 5G – 4.84 $\pm$ 1.33 nm in 300 mM LiCl, 5.31 $\pm$ 1.30 nm in 2 mM $MgCl_2$, 7.92 $\pm$ 1.49 nm in 300 mM KCl; 6G – 4.87 $\pm$ 1.36 nm in 300 mM LiCl, 2.64 $\pm$ 1.11 nm and 6.01 $\pm$ 1.40 nm in 2 mM $MgCl_2$, 8.51 $\pm$ 1.38 nm in 300 mM KCl. All samples were prepared at 20 μM oligonucleotide concentration.

cationic environments reflects literature trends, with K^+ ions being the most effective at promoting G4 formation and nanostar self-assembly, followed by Na^+ and Mg^{2+}, and finally by Li^+, which does not show any degree of G4 stabilisation.

3.3.2 Characterisation of cholesterolised Quad-Stars

After proving that individual non-cholesterolised Quad-Stars can isothermally self-assemble in response to G4-promoting K^+ ions, I proceeded to probe the self-assembly of choles-

terolised Quad-Stars into 3D percolating networks leading to macroscopic aggregate formation.

Influence of cationic environment on micelle formation

As observed for the cholesterol-tagged strands in C-Stars, Quad-Star duplex arms are expected to form cholesterol-induced micelles even when annealed in K^+-poor buffers, that is prior to G4-mediated Quad-Star assembly. Such micelle clusters would lead to strong colocalisation of the G4-forming domains/strands, ultimately influencing the formation rate of Quad-Stars and their assemblies. Therefore, I investigated the influence of the annealing cationic environment on the formation of such micelle clusters.

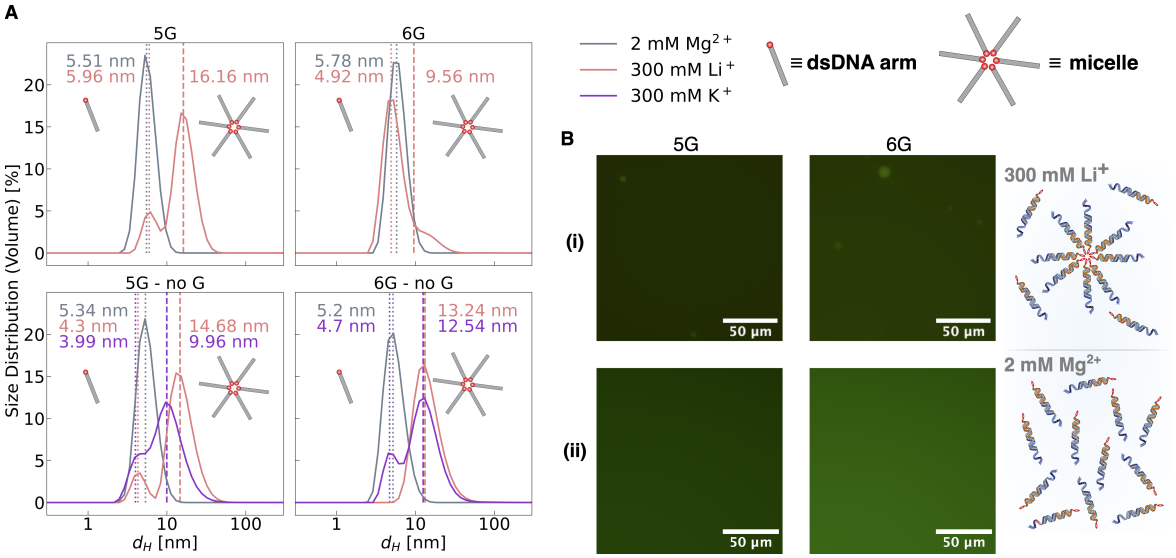


Fig. 3.8 **Influence of annealing cationic conditions on cholesterol micelle formation. A:** Volume-based size distributions obtained *via* DLS for cholesterolised 5G and 6G Quad-Stars (both with 9T spacer) annealed in either 2 mM $MgCl_2$ or 300 mM LiCl buffers, as well as negative controls (no Gs 14T and 15T, see Table A.1) annealed in either the same conditions or in 300 mM KCl buffer. Vertical lines are average hydrodynamic diameters ($< d_H >$), extracted *via* fit to mono- or bi-modal lognormal distributions (fits not shown). Quad-Star duplex arms fail to form micelle-like assemblies in lower ionic strength 2 mM $MgCl_2$ buffer, exhibiting $< d_H >$ values close to those of non-cholesterolised arms in Fig. 3.3A-B, much smaller than single assembled soluble Quad-Stars. Higher ionic strength buffers, such as 300 mM LiCl or KCl, allow for the formation of micelle-like clusters, as indicated by the larger-sized peaks at ≈ 10-15 nm, in equilibrium with some amount of free dsDNA arms (smaller-sized peaks closely matching the ones displayed by samples in 2 mM $MgCl_2$). Mean hydrodynamic diameters and standard deviations from covariance matrix of fit are hereby reported: 5G – 5.51 ± 1.29 nm in 2 mM $MgCl_2$, 5.96 ± 1.23 nm and 16.16 ± 1.36 nm in 300 mM LiCl; 6G – 5.78 ± 1.29 nm in 2 mM $MgCl_2$, 4.92 ± 1.29 nm and 9.56 ± 1.64 nm in 300 mM LiCl; 5G - no G – 5.34 ± 1.32 nm in 2 mM $MgCl_2$, 4.29 ± 1.20 nm and 14.68 ± 1.40 nm in 300 mM LiCl, 3.99 ± 1.24 nm and 9.96 ± 1.56 nm in 300 mM KCl; 6G - no G – 5.20 ± 1.34 nm in 2 mM $MgCl_2$, 13.24 ± 1.44 nm in 300 mM LiCl, 4.70 ± 1.21 nm and 12.54 ± 1.51 nm in 300 mM KCl. Profiles (solid lines) are averages of at least five independent measurements. **B:** Alexa488 epifluorescence micrographs depicting the formation, or lack thereof, of aggregates in 5G (left) and 6G (right) Quad-Star samples after annealing in **(i)** 300 mM LiCl or **(ii)** 2 mM $MgCl_2$ buffers, accompanied by sketches depicting formation and absence of micelles in such buffers, respectively, as evinced from DLS measurements in **A**. Micrographs are unprocessed and have been cropped for ease of visualisation. Scale bars = 50 µm. Samples for **A** and **B** were prepared at 20 µM and 5.6 µM oligonucleotide concentration, respectively.

DLS measurements in Fig. 3.8A confirm, indeed, the formation of micelle-like assemblies when cholesterolised Quad-Star components (5G - right, 6G - left) are annealed in 300 mM LiCl and 300 mM KCl – the latter condition only for negative controls with no Gs to avoid network formation, presented in the bottom row. However, no assemblies of similar size are observed in samples annealed in 2 mM $MgCl_2$, most likely due to the lower ionic strength not compensating the electrostatic repulsion among closely packed oligonucleotides.
As shown in the micrographs in Fig. 3.8(i), both 5G and 6G Quad-Stars annealed in Li^+-rich environments (300 mM LiCl) exhibit some small aggregates, made fluorescent thanks to an Alexa488-fluorophore conjugated to the linker strand, despite the lack of K^+ ions in solution. No such aggregates are observed, instead, in samples annealed in 2 mM $MgCl_2$, as demonstrated in micrographs depicted in Fig. 3.8(ii). Together with measurements in Figs. 3.3A and 3.7, these data hint at an aggregation mechanism where G4-forming oligonucleotides colocalise thanks to micelle formation, resulting in considerably enhanced local concentration of G4-forming strands. This, in turn, massively accelerates G4 formation kinetics due to the high dependency of tetramolecular G4 association rates on oligonucleotide concentration, as observed in Fig. 3.5B-C. While 300 mM LiCl is not enough to induce G4 formation in non-cholesterolised samples, it is capable of driving micelle formation. This leads to colocalisation of the component strands, which induces a certain degree of G4 formation and consequent low to moderate aggregation, despite the lack of G4-stabilising cations. Conversely, 2 mM $MgCl_2$ affords some degree of G4 stabilisation in non-cholesterolised designs, but fails to promote micelle formation, thus impeding strands from coming together into the designed nanostars and ultimately preventing aggregation. For this reason, samples annealed in 2 mM $MgCl_2$ were chosen as the starting condition for K^+ ion-triggered experiments, providing a blank slate to investigate the kinetics of G4 formation and Quad-Star aggregation.

NMM Fluorimetry

Similarly to non-cholesterolised samples, I annealed cholesterolised Quad-Stars in K^+-poor saline buffers, namely 2 mM $MgCl_2$, and monitored G4 formation and network assembly using time-resolved NMM fluorimetry upon injection of KCl ([KCl] = 300 mM). The resulting NMM fluorescence kinetics in Fig. 3.9A confirm that K^+ ion addition rapidly triggers G4 formation, with both 5G and 6G designs displaying substantially faster assembly compared to the non-cholesterolised counterparts in Fig. 3.5. Half-life values are dramatically reduced, as summarised in Fig. 3.9B, with 5G samples experiencing a 100-fold boost in assembly rate, *i.e.* decrease in half-life values. As expected, samples lacking KCl injection do not experience any fluorescence increase or aggregation, as depicted in Fig. 3.9C. These results are in agreement with the proposed assembly mechanism, whereby the colocalisation of G4-forming strands mediated by cholesterol micelles causes a dramatic boost in G4 formation rates. At first, in fact, the 2 mM $MgCl_2$ buffer used for annealing does not possess high enough ionic strength to promote micelle formation, as demonstrated by DLS measurements

in Fig. 3.8A. However, the addition of 300 mM KCl increases the ionic strength, thus inducing the formation of micelle-like assemblies, and also provides G4-promoting and stabilising cations, ultimately resulting in this considerable speed-up.

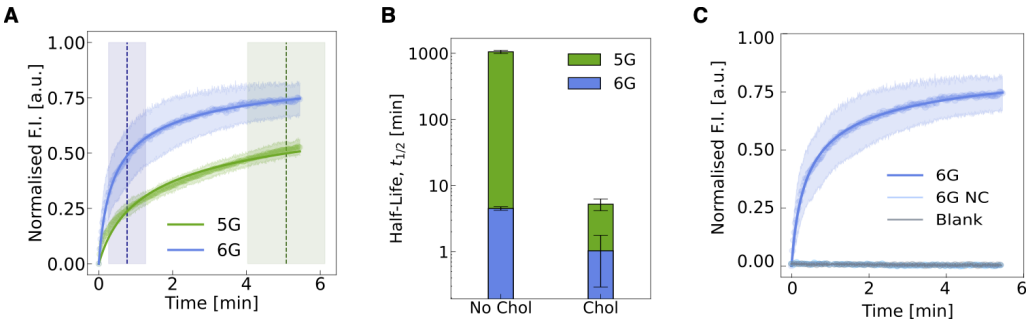


Fig. 3.9 **Kinetics of G4-dependent self-assembly of cholesterolised Quad-Stars *via* NMM fluorimetry.** **A:** Normalised fluorescence intensity kinetic profiles of G4-folding in cholesterolised 5G and 6G Quad-Stars (both with 9T spacer). Half-life values are depicted as mean (vertical dashed lines) ± standard deviation (shaded intervals). **B:** Half-life values, $t_{1/2}$, are dramatically reduced for cholesterolised Quad-Stars compared to soluble, non-cholesterolised ones (Fig. 3.5A.) Since tetramolecular G4 association rates highly depend on oligonucleotide concentration (Fig. 3.5B-C), such reduction in half-life values can most likely be ascribed to increased local concentration of G4-forming oligonucleotides caused by the formation of cholesterol-mediated micelles upon the ionic strength increase brought about by the KCl injection (see Fig. 3.8A). Data are reported as mean (filled bars) ± standard deviation (error bars). **C:** Comparison between normalised kinetic profiles for cholesterolised 6G Quad-Stars receiving KCl injection (bright curve, same as in **A**), negative controls (6G NC) lacking the KCl injection and blanks (plain TE buffer + NMM). Data for negative controls and blanks are averages of duplicates, normalised by mean subtraction and division (see Sections 2.8.2 and 2.11.3), and have been trimmed (time-wise) to match the time duration of the profile for 6G. Raw data for negative controls and blanks are provided in Fig. B.3. Data in **A** and **C** are presented as mean (datapoints, one marker every 3 points shown) ± standard deviation (shaded regions) from 4/10 (5G/6G, respectively) replicates. Solid lines in **A** are fits to theoretical tetramolecular kinetics as in Fig. 3.5 [329]. All samples were prepared at 20 μM oligonucleotide concentration.

3.4 Reversible isothermal self-assembly and disassembly

NMM fluorimetry assays in Fig. 3.9 monitor the isothermal folding of tetramolecular DNA G4 motifs within cholesterolised Quad-Stars, but they do not provide insights into their assembly into extended 3D networks and aggregate formation. In this section, I will demonstrate that it is possible to monitor such isothermal assembly and study its kinetics through epifluorescence microscopy timelapses. I will show that Quad-Stars form hydrogel aggregates, which can be disassembled upon chelation of G4-promoting K^+ ions through chelating agent [2.2.2]cryptand. Finally, I will prove that such hydrogel aggregates can be reversibly assembled and disassembled for up to three cycles *via* repeated K^+ ion addition and removal, *i.e.* cryptand addition. For these experiments, fluorescence imaging was enabled by the use of an Alexa488 fluorophore conjugated to the linker strand in the duplex Quad-Star arms, as previously seen in Fig. 3.8B. Such covalently-linked fluorophore was preferred to NMM to avoid NMM-mediated photo-induced oxidation and cleavage of G4-forming guanine residues upon strong illumination (see Section 3.5.3 below).

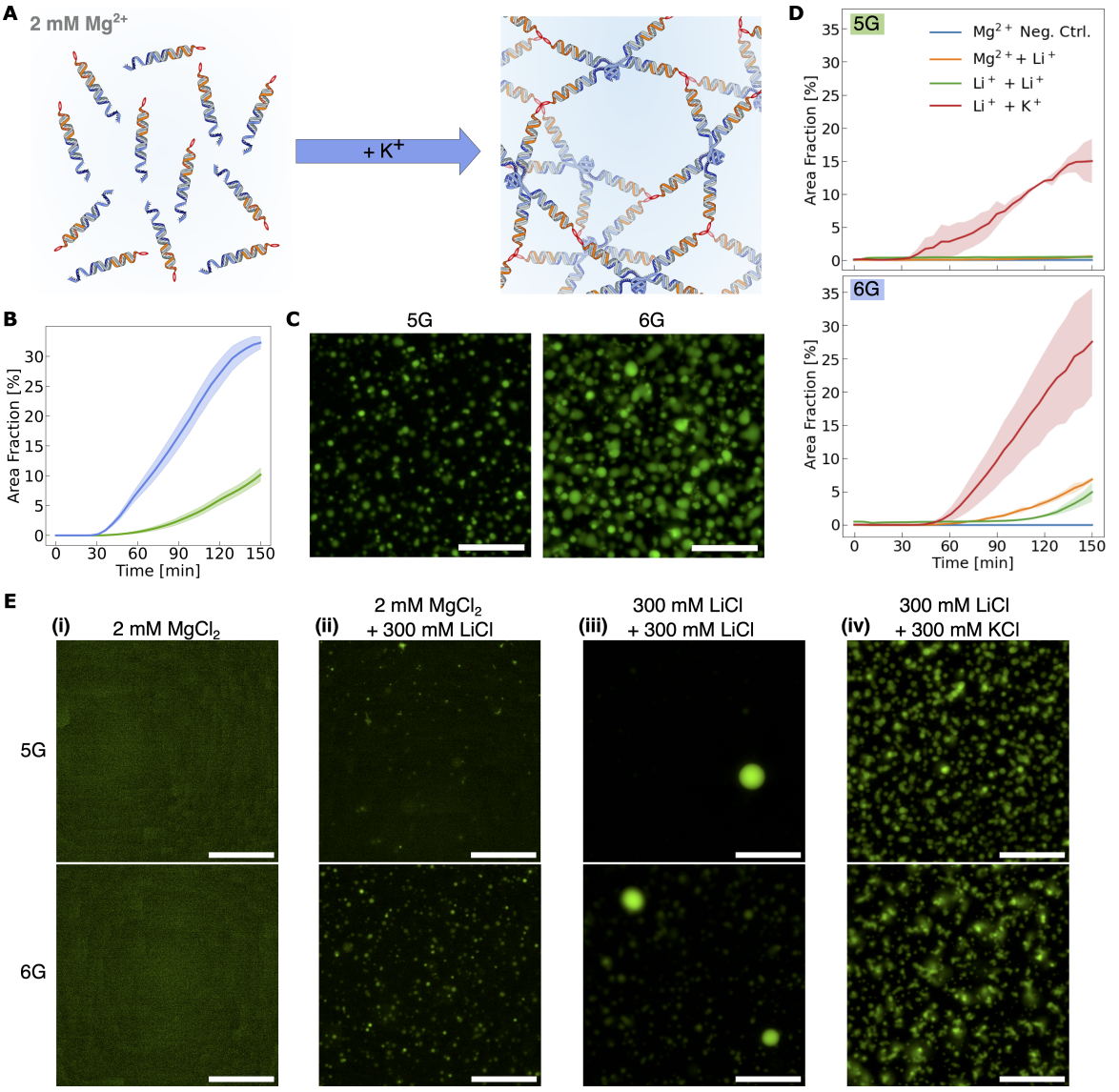


Fig. 3.10 **Isothermal assembly and hydrogel aggregate formation. A:** Schematic representation of K^+-induced Quad-Star assembly, from disjoined duplex arms, observed after annealing in 2 mM $MgCl_2$ (see Fig. 3.8), to a 3D percolating network leading to formation of hydrogel aggregates. **B:** Self-assembly kinetics of Quad-Star hydrogel aggregates (5G in green, 6G in blue, both with 9T spacer) represented by time-dependent area fraction profiles computed from epifluorescence microscopy timelapses (see Section 2.11.7). Data are presented as mean (solid lines) ± standard deviation (shaded regions) from 3/4 repeats for 5G/6G, respectively. **C:** Representative epifluorescence micrographs of 5G and 6G Quad-Star aggregates, formed after 2.5 h from 300 mM KCl injection. Samples were originally annealed in 2 mM $MgCl_2$-containing buffer. **D:** Area fraction profiles corresponding to control samples for 5G (top) and 6G (bottom) (both with 9T spacer) cholesterolised Quad-Stars: samples annealed in 2 mM $MgCl_2$ (blue); samples annealed in 2 mM $MgCl_2$ and supplemented with 300 mM LiCl (blue); samples annealed in 300 mM LiCl and supplemented with 300 mM LiCl (green); samples annealed in 300 mM LiCl and supplemented with 300 mM KCl (red). Data are presented as mean (solid lines) ± standard deviation (shaded regions) of three FOVs within single representative samples, apart from 5G - Li^+ + K^+ (red, left panel), reported instead as mean of two FOVs (solid line) ± interval among them (shaded region). **E:** Representative epifluorescence micrographs of 5G (top) and 6G (bottom) Quad-Star control samples from **D** after 2.5 h from the corresponding saline buffer injection (or lack thereof in **(i)**). Imaging is enabled by an Alexa488 fluorophore conjugated to the linker strand in the Quad-Star arms. Micrographs in **C**, **E** have been background-subtracted (see Section 2.11.5) and cropped for ease of visualisation. Scale bars are 50 μm. All samples were prepared at 5.6 μM oligonucleotide concentration.

3.4.1 Aggregate formation kinetics

The kinetics of Quad-Star assembly and aggregate formation, schematically depicted in Fig. 3.10A, can be investigated through time-dependent area fraction profiles, *i.e.* the fraction of a microscopy field of view (FOV) occupied by fluorescent Quad-Star aggregates *vs* time, as computed by image analysis (see Section 2.11.7). Such results are presented in Fig. 3.10B, which demonstrates, as expected from NMM fluorimetry measurements (Figs 3.5A and 3.9A), a much faster assembly rate for 6G designs compared to 5G ones. This is also confirmed by micrographs in Fig. 3.10C, which showcase round, fluorescent hydrogel-like aggregates after 2.5 h from 300 mM K^+ injection for both 5G and 6G constructs, with the latter exhibiting larger and more closely packed aggregates.

Influence of different annealing conditions and other cation injections

As it can be deduced from Fig. 3.8 and 3.9, KCl injection promotes Quad-Star aggregation by both increasing the ionic strength, thus facilitating micelle assembly and G4-forming strand colocalisation, and providing G4-promoting K^+ ions. To disentangle these two effects, 5G and 6G Quad-Stars were annealed in either 2 mM $MgCl_2$, as in Figs. 3.9 and 3.10A, or 300 mM LiCl, and then supplemented with either nothing, 300 mM LiCl or 300 mM KCl. The resulting time-dependent area fraction profiles are presented in Fig. 3.10D. Samples annealed in 2 mM $MgCl_2$ do not display any aggregation without further cation addition, as exemplified by the blue profile (Mg^{2+} Neg. Ctrl. = Negative Control) in Fig. 3.10D and accompanying micrographs after 2.5 h in Fig. 3.10E (i). Note that any patterns in the illumination of these micrographs are artifacts due to background subtraction in the absence of any fluorescent aggregates. When provided with a further 300 mM LiCl trigger injection (Fig. 3.10D, orange; Fig. 3.10E (ii)), some degree of aggregation is observed, particularly in 6G designs. However, this response is considerably less pronounced than the one depicted in Fig. 3.10B when samples are supplemented with the same concentration of G4-promoting K^+ ions, which still causes the same increase in ionic strength. This confirms the hypothesised interplay of ionic strength increase and degree of metal cation afforded G4-stabilisation in Quad-Star aggregate formation, and decouples the two effects: while 300 mM LiCl and KCl lead to the same increase in ionic strength, and equally facilitate micelle formation and boost in G4 assembly rates, G4-promoting K^+ ions result in much stronger and more effective aggregation.

Consistent results are obtained for samples annealed in 300 mM LiCl. While they present few small aggregates right after annealing as reported in Fig. 3.8B(i), they undergo further aggregate formation when supplemented with either 300 mM LiCl (Fig. 3.10D, green; Fig. 3.10E (iii)) or 300 mM KCl (Fig. 3.10D, red; Fig. 3.10E (iv)): the final ionic strength is the same, but the degree of aggregation is considerably larger in samples receiving a trigger injection of G4-promoting K^+ ions.

3.4.2 Disassembly by potassium ion chelation through cryptand

Potassium ions template the formation of G-quartets and stabilise the resulting G4 motifs by sitting in the central channel, sandwiched among pairs of adjacent G-quartets, thus axially coordinating with eight O6 atoms. Due to the crucial contribution of such coordination, potassium-stabilised G4s tend to structurally weaken [391] and rapidly unfold [392–395, 182, 187] when K^+ ions are removed. This can be achieved *via* chelation through [2.2.2]cryptand [393, 394, 182], a bicylic chelating agent with high affinity for alkali metal cations whose structure is reported in Fig. 3.11A. It should therefore be possibile to reverse the aggregation of Quad-Stars by chelating the G4-stabilising K^+ ions, thereby isothermally disassembling the hydrogel aggregates. This reversibility route was investigated through a combination of NMM fluorimetry and epifluorescence microscopy on 6G Quad-Stars,

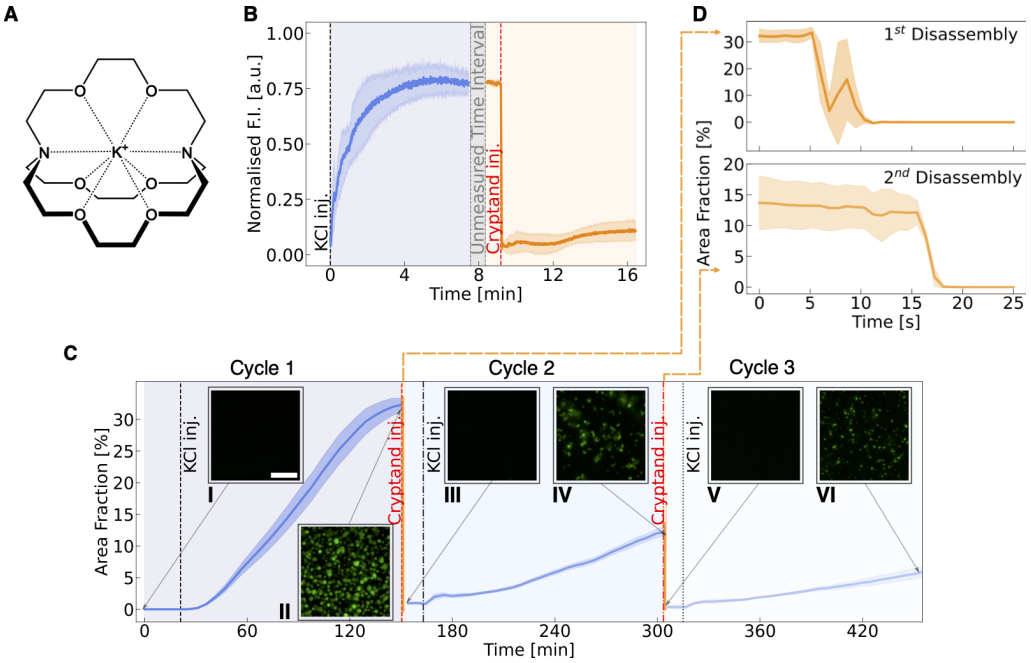


Fig. 3.11 **Quad-Stars and their hydrogel aggregates can reversibly assemble and disassemble upon addition or chelation of K^+ ions. A:** Chemical structure of [2.2.2]cryptand (or 4,7,13,16,21,24-Hexaoxa-1,10-diazabicyclo[8.8.8]hexacosane), represented in coordination with a K^+ ion. **B:** Reversibility of cholesterolised 6G (9T spacer) Quad-Star assembly assessed *via* NMM time-dependent fluorescence intensity kinetics. Addition of K^+ ions (black vertical line) triggers G4 folding, as presented in Fig. 3.9A, while their removal *via* addition of chelating agent [2.2.2]cryptand (red vertical line) induces fast disassembly. **C:** Reversibility of 6G (9T spacer) Quad-Stars hydrogel aggregate formation, monitored *via* time-dependent area fraction profiles as showed in Fig. 3.10B, is demonstrated by three consecutive assembly/disassembly rounds. Insets (I-VI) depict cropped, background-subtracted epifluorescence micrographs (see Section 2.11.5) at representative assembly/disassembly timepoints indicated by the corresponding arrows. All insets show the same FOV area, thus sharing scale bar, which is 50 μm. **D:** Zoomed-in time-dependent area fraction profiles from disassembly rounds in **C**. Time on the x-axis is zero-ed with respect to the corresponding cryptand injection ($t_{\text{injection}} = 0$ s). The reduction in starting area fraction with progressive assembly/disassembly round matches with the progressively slower assembly rates and lower end values of hydrogel aggregate area fraction displayed in **C**. Both are most likely caused by considerable sample dilution upon subsequent KCl/cryptand injections. Data in **B**, **C**, **D** are presented as mean (solid lines) $\pm$ standard deviation (shaded regions) from at least three repeats. Samples in **B** and **C**-**D** were prepared at 20 μM and 5.6 μM oligonucleotide concentration, respectively.

as presented in Fig. 3.11B-C. NMM fluorescence was monitored during the potassium-induced assembly transient, as presented in Fig. 3.9B, as well as immediately before and after [2.2.2]cryptand injection (see Section 2.9), resulting in almost instantaneous drop of fluorescence, as demonstrated by Fig. 3.11B, indicative of G4 and hydrogel aggregate disassembly.

3.4.3 Assembly reversibility

Such reversibility can be promptly extended to multiple subsequent rounds of assembly/disassembly by alternating KCl/cryptand addition in the same 6G Quad-Stars sample, respectively. Fig. 3.11C presents the resulting time-dependent area fraction profiles with Alexa488 epifluorescence micrographs from representative timepoints as insets. Visible hydrogel-like aggregates form upon each round of KCl injection, while they abruptly disassemble upon addition of K^+-chelating cryptand, as clearly exemplified by the zoomed-in profiles in Fig. 3.11D. Notably, the assembly rates and the end area fraction after each assembly transient both considerably decrease with increasing round number. This can be rationalised considering the high reaction order (tetramolecular) of Quad-Star formation and the progressive sample dilution upon multiple subsequent injections. Assuming perfect mixing at each KCl/cryptand injection, the 6G Quad-Star sample is diluted by ≈36.37%/38.78% upon the second/third assembly round with respect to the previous one, which, using the reaction order computed for non-cholesterolised 6G Quad-Stars (see Fig. 3.5), would slow down the assembly by ≈3.41/3.78-fold at each step. Conversely, the assembly kinetics observed through area fraction profiles point toward a slow down of ≈2.5-fold at each step, hinting at a dilution of ≈25-30% *per* assembly/disassembly round. This discrepancy is most likely related to the assumption of complete mixing in the calculation above, since cryptand was simply injected, rather than mixed by pipetting up and down, to better monitor the disassembly of the hydrogel aggregates at the bottom of the imaging chamber.

3.5 Cargo-loading capabilities and photo-induced irreversible disassembly

3.5.1 Rhodamine B uptake and triggered release

Quad-Star hydrogels can reversibly assemble and disassemble in response to changes in identity and concentration of physiologically-relevant cations, making them promising candidates for dynamic synthetic MLOs. However, membrane-less compartments in biological cells are also capable of capturing target analytes upon assembly in response to a specific stimulus, and later releasing them when such stimulus passes. Stress granules provide a prime example for this behaviour, in that, during celluar stress, they sequester mRNAs stalled in translation

initiation to prevent their aggregation, and then release them for translation to proceed when stress subsedes [351].

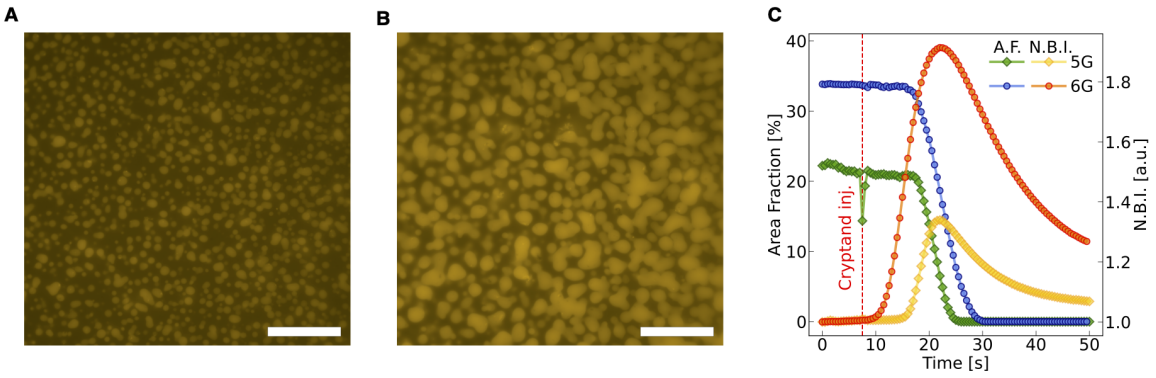


Fig. 3.12 **Amphiphilic Quad-Star hydrogel aggregates can uptake and release model hydrophobic small molecule cargo, rhodamine B. A, B:** Unprocessed epifluorescence micrographs of 5G (**A**) and 6G 5G (**B**) Quad-Star aggregates, formed after 3.5 h of K^+-triggered assembly in the presence of rhodamine B. In both cases, the hydrogel aggregates appear brighter than the background (27.2 $\pm$ 4.7 % brighter for 5G and 34.7 $\pm$ 3.4 % brighter for 6G), confirming the expected preferential segregation of rhodamine B within the cholesterol-rich hydrophobic pockets. Scale bars are 50 μm. **B:** Representative time-dependent profiles illustrating the disassembly of Quad-Star aggregates (A.F. = Area Fraction) upon cryptand injection (red vertical line), leading in turn to the release of the rhodamine B cargo (N.B.I. = Normalised Background Intensity). Profiles have been extracted from single FOVs of representative samples, where the distance between injection site and FOV allowed for sensible time resolution. All samples were prepared at 5.6 μM oligonucleotide concentration.

To mimic such behaviour in a simplified fashion, I exploited the amphiphilic nature of Quad-Stars, which, similarly to C-Stars [160, 163], allows them to uptake and become preferentially enriched in hydrophobic small molecules. This also grants Quad-Star hydrogels considerable potential as vectors for drug delivery and release, since many promising drug molecules are hydrophobic and challenging to deliver due to poor water solubility and bioavailability [396–399]. Thanks to the reversibility of Quad-Star aggregates, such cargo molecules can be released upon disassembly by chelation of G4-stabilising K^+ ions, or potentially through enzymes, such as biologically relevant helicases capable of selectively unfolding DNA G4s [400–403].

This cargo uptake and release mechanism is demonstrated in Fig. 3.12 using model hydrophobic small molecule rhodamine B, which can be easily visualised using fluorescence microscopy. Quad-Star arms were pre-annealed in 2 mM $MgCl_2$ and their isothermal assembly was triggered by KCl injection in the presence of rhodamine B. Micrographs in Fig. 3.12A-B demonstrate preferential segregation of hydrophobic rhodamine B within both 5G and 6G Quad-Star hydrogel aggregates, as proved by the higher inner fluorescence intensity compared to the background.

The rhodamine B cargo can be released into the surrounding solution by disassembling the host Quad-Star aggregates through K^+ ion chelation *via* [2.2.2]cryptand addition, as showed in Fig. 3.12C. The kinetics of such process are studied by simultaneously following Quad-

Star aggregate disassembly and cargo release through area fraction profiles and background fluorescence intensity, respectively. Upon cryptand injection, Quad-Star aggregates quickly disassemble as captured by the drop in the corresponding area fraction profiles, accompanied in turn by a sharp increase in background fluorescence, as rhodamine B is progressively released from the aggregates into the surrounding solution. Background fluorescence then slowly decreases due to diffusion of the excess dye out of the imaged FOV. Here, cryptand was injected far away from the examined FOV, allowing for sensible time-resolution and slower disassembly compared to what observed in Fig. 3.11 and analogous profiles in the Appendix (Fig. B.4).

3.5.2 Biostability of Quad-Star aggregates

To further substantiate the applicability of Quad-Stars as synthetic MLOs in artificial cells as well as drug delivery carriers, I tested their stability under model physiological conditions. In particular, Phosphate Buffered Saline (PBS) 1× was chosen to better mimic extracellular ionic environments [404], as it includes ≈140-150 mM Na^+ ions and ≈3-5 mM K^+ ions. While intracellular concentrations are approximately reversed, with ≈140 mM K^+ ions and ≈12 mM Na^+ ions, PBS 1× allows to investigate the capability of Quad-Star hydrogels to survive in the considerably less G4-stabilising extracellular ionic conditions. On the other hand, Dulbecco's Modified Eagle Medium (DMEM) supplied with either 10% or 30% Fetal Bovine Serum (FBS) was selected to reproduce conditions previously used in literature to assess the biostability of DNA nanostructures in tissue culture models [405–407]. DMEM is, in fact, a standard tissue culture medium containing similar cationic species and concentrations to PBS 1× (≈150 mM Na^+, 5 mM K^+) with additional divalent cations (≈0.8 mM Mg^{2+}, 1.8 mM Ca^{2+}), as well as 25 mM D-Glucose (high glucose DMEM) and a variety of vitamins and amino acids. Importantly, FBS introduces a plethora of additional macromolecules, from fatty acids and lipids, to proteins and enzymes, including proteases and nucleases (both DNAses and RNAses) capable of degrading nucleic acids, thus providing a challenging environment for the stability of Quad-Star hydrogel aggregates.

The isothermal assembly of 5G and 6G Quad-Star aggregates was triggered *via* K^+-ion injection as described above, and allowed to proceed for 3-4 h. Buffer exchange was then performed by supernatant removal and replacement with either of the buffers mentioned above. Samples were monitored *via* epifluorescence microscopy (through the Alexa488 fluorophore-conjugated strand used previously) for 60 h at room temperature, as summarised in Fig. 3.13. Quad-Star aggregates kept in their isothermal assembly, K^+-rich buffer continued growing and coarsening into a hydrogel layer nearly covering the entire bottom of the well (Fig. 3.13A and C for 5G and 6G, respectively). The clear reduction in fluorescence intensity can most likely be ascribed to photobleaching due to prolonged imaging. Conversely, buffer exchange with PBS 1× induced different responses in 5G and 6G samples. 5G hydrogel aggregates (Fig. 3.13B) partially disappeared, and the remaining ones underwent some degree of coars-

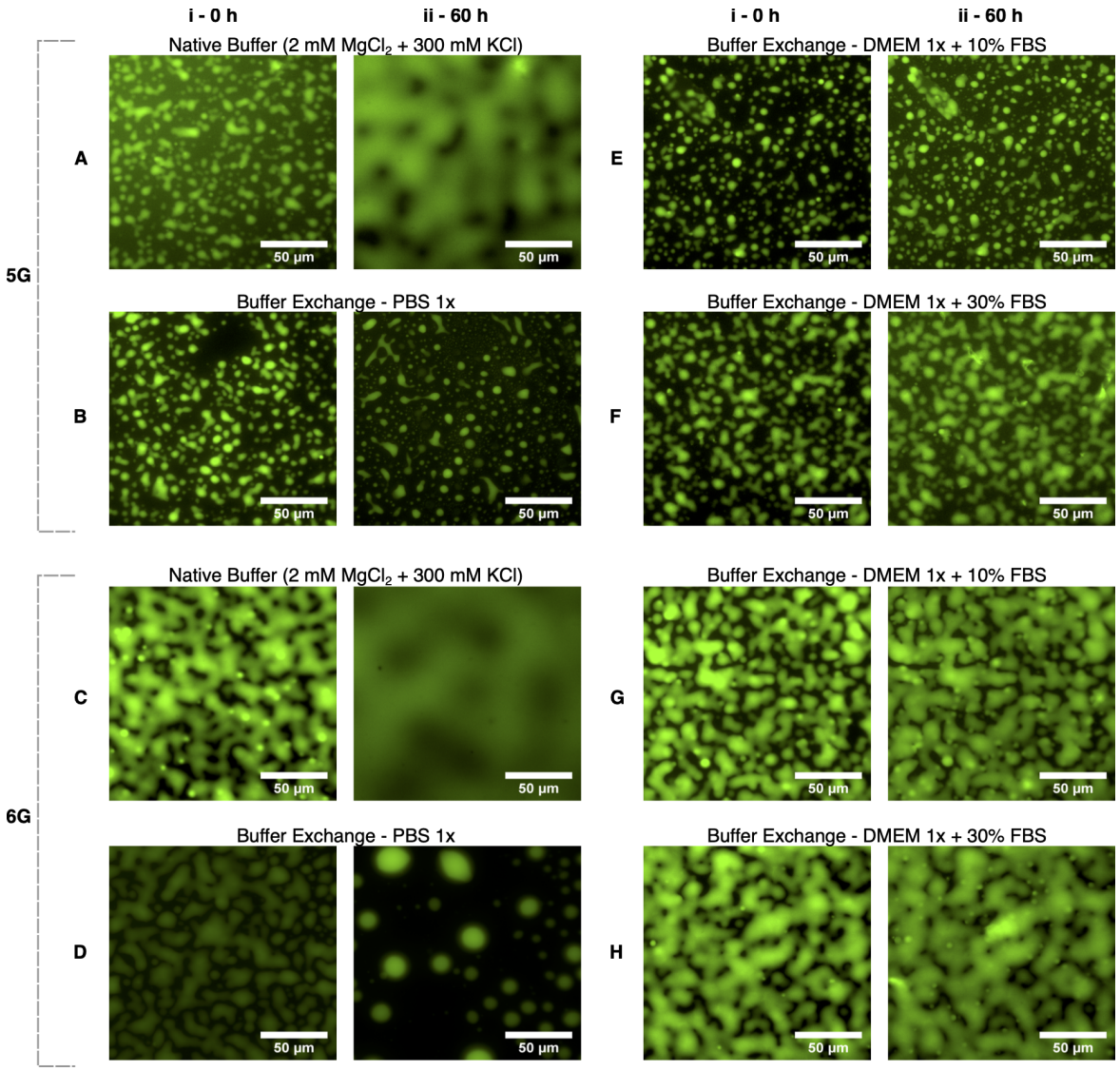


Fig. 3.13 **Stability of Quad-Star hydrogel aggregates under model physiological conditions. A, C:** Alexa-488 epifluorescence micrographs monitoring the stability of 5G (**A**) and 6G (**C**) Quad-Star hydrogel aggregates formed after 4 h of K^+-induced assembly, imaged at room temperature over the course of 60 h in their native assembly buffer (TE 1× supplied with 2 mM $MgCl_2$ and 300 mM KCl). **B, D:** Same as **A** for 5G (**B**) and 6G (**D**) Quad-Star hydrogel aggregates formed after 3 h of K^+-induced assembly, monitored at room temperature over the course of 60 h after buffer exchange with Phosphate Buffered Saline (PBS) 1× at time $t = 0$ h **(i)**. **E, G:** Same as **A** for 5G (**E**) and 6G (**G**) Quad-Star hydrogel aggregates formed after 3 h of K^+-induced assembly, monitored at room temperature over the course of 60 h after buffer exchange with Dulbecco's Modified Eagle Medium (DMEM) 1× supplied with 10% Fetal Bovine Serum (FBS) at time $t = 0$ h **(i)**. **F, H:** Same as **A** for 5G (**F**) and 6G (**H**) Quad-Star hydrogel aggregates formed after 4 h of K^+-induced assembly, monitored at room temperature over the course of 60 h after buffer exchange with Dulbecco's Modified Eagle Medium (DMEM) 1× supplied with 30% Fetal Bovine Serum (FBS) at time $t = 0$ h **(i)**. Buffer exchange was carried out *via* supernatant removal and replacement (see Section 2.9). Micrographs have been contrast-enhanced (see Section 2.11.5) and cropped for ease of visualisation. Scale bars are 50 μm. All samples were prepared at 5.6 μM oligonucleotide concentration.

ening. On the other hand, 6G Quad-Stars (Fig. 3.13D) did not experience visible growth or disassembly, and rater coarsened and coalesced – to a much fuller extent compared to 5G – into larger, discrete microgel particles. These behaviours both hint at enhanced fluidity

in buffers richer in Na^+ and poorer in K^+, with the reduction in G4-promoting K^+ ions impacting 5G samples more significantly most likely due to their reduced stability compared to 6G (see Figs. 3.5D and 3.6). Finally, complete retention of the initial morphology was observed in hydrogel particles transferred to DMEM + 10% (Fig. 3.13E and G for 5G and 6G, respectively) or 30% FBS (Fig. 3.13F and H for 5G and 6G, respectively). It is worth noting that the difference in area fraction occupied by hydrogel aggregates among Fig. 3.13A, C, F, H and B, D, E, G is due to the difference in assembly transient length prior to buffer exchange (4 *vs* 3 h, respectively).

Together, these result demonstrate the biostability of Quad-Star hydrogel aggregates, confirming their potential as reversible, responsive compartments in synthetic cells or drug delivery and release carriers.

3.5.3 NMM-enabled photodisassembly

So far, I have illustrated a single mechanism, namely K^+ ion chelation, to reversibly disassemble Quad-Star aggregates in a location-agnostic fashion, that is homogeneously in the entire sample. However, the non-canonical tetramolecular G4 that bridges Quad-Star arms also enables an orthogonal, irreversible and highly localised disassembly mechanism requiring near-UV exposure in the presence of the porphyrin photosensitiser NMM. Despite being typically used as a fluorescent probe for parallel G4s, as in Figs. 3.5 and 3.9, recent literature reports [368] have demonstrated that NMM can induce guanine photo-oxidation and cleavage upon strong longwave UV exposure capable of exciting its fluorescence emission (395-400 nm). To test whether such response could be elicited in Quad-Stars, I exposed 6G hydrogel aggregates to a near-UV 395 nm LED in the presence of NMM and monitored their integrity over time in both fluorescence (NMM emission, 670 nm) and bright-field, as depicted in Fig. 3.5.3A. After NMM injection (dashed vertical black line), Quad-Star aggregates quickly disassemble upon UV exposure, completely disappearing within 90 s from both fluorescence and bright-field channels of the reported composite micrographs. Notably, I found such disassembly to be irreversible, in agreement with literature reports of NMM-mediated, UV-induced selective photo-oxidation and cleavage of guanine residues at G4 sites [369–372, 368]. While the precise mechanism is still to be unveiled, it is clear that only G4s in proximity of UV-exposed NMM moieties will undergo damage and cleavage, making this orthogonal, photo-induced disassembly mechanism highly location-specific and thus allowing to trigger a spatially localised response. To demonstrate this, I engraved a pattern of interest ("IC" for Imperial College in Fig. 3.14B) in a large 6G Quad-Star sample by selectively disassembling aggregates within several FOVs along the corresponding path using the motorised stage of the microscope.

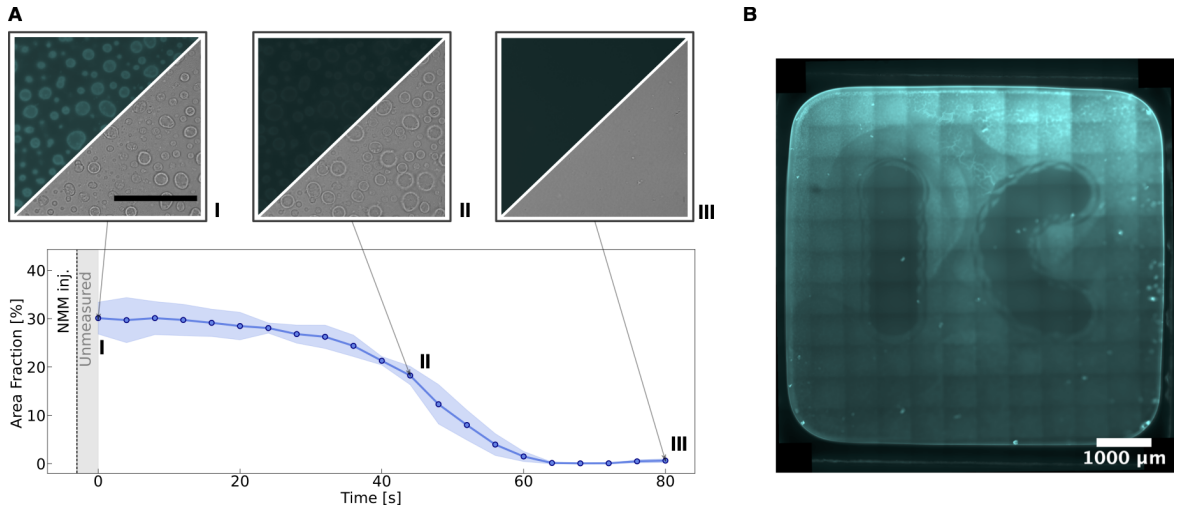


Fig. 3.14 **NMM-mediated, irreversible and highly localised photo-induced disassembly of Quad-Star hydrogel aggregates. A:** Near-UV (395 nm) exposure excites NMM fluorescence and leads to photo-induced guanine damage *via* oxidation and cleavage [368], ultimately causing irreversible and localised disassembly of Quad-Star aggregates. Depicted sample is 6G9T. Micrographs on top are composite images (unprocessed for epifluorescence, contrast-enhanced for ease of visualisation for bright-field) corresponding to timesteps (**I - III**) indicated by arrows on the Area Fraction profile at the bottom. Area Fraction data are reported as mean (circles) ± standard deviation (shaded interval) of three FOVs within the same sample. All insets depict the same FOV area, thus sharing the scale bar, which is 50 μm. Uncropped, unprocessed micrographs for timesteps **I - III**, as well as Alexa488 epifluorescence micrographs excluding NMM photobleaching as the cause of aggregates vanishing are provided in the Appendix (Fig. B.5). **B:** Photo-induced engraving in a 6G9T Quad-Star sample well, reading "IC" (abbreviation for Imperial College), proves highly localised nature of NMM-mediated photo-induced disassembly. The engraving was obtained *via* automatic exposure of several FOVs along the desired path with a 395 nm LED, thereby triggering location-selective photodisassembly. The visible underlying square-lattice pattern is an artifact resulting from slightly uneven illumination and automatic FOV stitching, and as such it does not represent a physical feature of the sample. Scale bar is 1000 μm. All samples were prepared at 5.6 μM oligonucleotide concentration.

3.6 Chapter summary and outlook

In this Chapter, I have introduced Quad-Stars, which combine the robustness and versatility of previously introduced amphiphilic DNA C-Stars [160] with the stimuli-responsiveness of non-canonical DNA G-quadruplexes. I have demonstrated that individual Quad-Stars self-assemble through tetramolecular parallel G4s in a K^+-dependent fashion thanks to the intrinsic cation-responsiveness of G4 motifs. I have also showed that Quad-Stars form discrete hydrogel aggregates, with sizes ranging from a few to a few tens of microns, through an amenable one-pot, enzyme-free and isothermal reaction.

The rates of individual motif self-assembly and overall hydrogel particle formation can be widely tuned *via* multiple parameters, namely K^+ ion and oligonucleotide concentrations, as well as the number of consecutive guanines in the G4-forming sequences, which dictate the maximum number of stacked G-quartets in the linking DNA G4 motifs.

Thanks to the cation-responsive DNA G4s, Quad-Star hydrogel aggregates can undergo controlled and reversible assembly and disassembly in response to addition and chelation *via* [2.2.2]cryptand of biologically relevant K^+ cations, here demonstrated for up to three

subsequent rounds. To further prove their potential for applications in synthetic biology and drug delivery and release, I demonstrated that Quad-Star hydrogel particles are stable under model physiological conditions. Moreover, they can readily uptake small hydrophobic molecules, here exemplified by rhodamine B, and later release them upon controlled disassembly through the chelation of G4-stabilising K^+ ions. An orthogonal stimulus for irreversible, highly localised near UV-induced disassembly mechanism in the presence of the G4 ligand and porphyrin photosensitiser NMM was demonstrated, further expanding the repertoire of control mechanisms for disassembly and cargo release.

Notably, the development of Quad-Stars was made possible by the tolerance to design changes exhibited by C-Stars. This characteristic made it possible to simply substitute a topological component – the four-way junction – with an analogous stimuli-responsive DNA motif – the tetramolecular parallel G4 – while retaining similar self-assembly behaviour. This modular design strategy could be further applied to expand the range of stimuli our nucleic acid constructs can respond to. To substantiate this claim, Walczak *et al.* recently made C-Stars responsive to another biologically-relevant stimulus, namely pH, by embedding cytosine-rich non-canonical i-Motifs [408].

While I have only investigated the bulk formation of Quad-Star hydrogel aggregates, it is evident that these could be either used as reversible scaffolds for hybrid synthetic cell-like agents [409, 410, 198, 52] or, even more straightforward, readily integrated in DNA-based artificial cells to yield cation-responsive compartments acting as synthetic MLOs. For instance, they could form a cation-responsive core in the multi-phasic C-Stars-based artificial cells recently demonstrated by Malouf *et al.* [411], where the difference in arm length among the C-Star constructs in use drives the formation of multiple compartments with varying responsiveness and functionality. The Quad-Star smart compartments could enable the creation of micro-environments capable of hosting active components whose action could in turn couple with the responsiveness of the host networks, implementing an intrinsically biomimetic synergy among structure and functionality [24, 412, 36].

To this aim, it should be possible to further expand the range of available stimuli for the assembly and disassembly of Quad-Star based MLOs. One promising route is resorting to naturally occurring enzymes, in particular biologically relevant helicases which are capable of selectively unfolding G4s [400–403]. These could be either added externally or produced *via* cell-free protein expression systems. Alternatively, one could leverage the wide array of chemical moieties that have recently been developed to tag, stabilise or de-stabilise G4s for applications in molecular biology and therapeutics [413].

Finally, while I focused on the characterisation of the thermodynamics and kinetics of assembly and disassembly of Quad-Stars, the nano- and micro-structural properties of Quad-Star aggregates are still to be elucidated. In particular, C-Stars, which constitute the founding architecture for Quad-Stars, can form, depending on their flexibility and size, discrete crystalline aggregates with Body Centred Cubic (BCC) symmetry and tunable lattice param-

eter [163, 202]. It remains to be discovered whether Quad-Stars are capable of doing the same, possibly affording novel crystal geometries due to the further geometrical and steric constraints dictated by the linking tetramolecular parallel DNA G4 motifs.

Chapter 4

Orthogonal RNA condensates from co-transcriptionally folding RNA nanostars

4.1 Introduction

Biological cells have evolved to operate far from equilibrium [414], while handling a variety of highly complex processes at constant temperature [415]. In our quest towards ever more advanced biomimicry, artificial cells would thus greatly benefit from the ability to perform functions and assemble components isothermally. Conversely, most synthetic DNA platforms, including the one discussed in Chapter 3, still rely on non biologically-relevant single-stranded DNA components, which further require some degree of thermal annealing to yield the desired architectures.

Beyond being the biological intermediate between DNA and proteins, RNA has showed significant potential for engineering biomimetic synthetic systems. On one hand, RNA displays similar programmability and versatility compared to DNA [206, 279, 54]. This should make it possible to systematically tune the behaviour and properties of the resulting architectures, as demonstrated by the wide variety of DNA nanotechnology-enabled approaches, from *de novo* designed DNA-based droplets [199, 201] to DNA-scaffolded protein lattices [416], as well as nanoparticle [417, 418], colloidal [419, 420] and liposome [421, 162] assemblies. On the other hand, RNA can be enzymatically synthesised from more biologically-relevant double-stranded DNA. Thanks to this naturally evolved mechanism, designer RNA constructs can be engineered to fold co-transcriptionally [320], *i.e.* fold while being transcribed from DNA templates. This removes stoichiometry-related issues stemming from differential transcription of different components, and thus considerably improves formation efficiency [295, 215, 325]. Co-transcriptional folding has successfully been applied to isothermally produce complex RNA origami [321, 323, 422] as well as smaller RNA nanostructures synthesised within

Escherichia coli (*E. coli*) [215].

The potential of RNA-based synthetic systems is clearly demonstrated by the role of RNA in naturally occurring biomolecular condensates [30, 423, 351] and membrane-less organelles (MLOs) [30, 28]. Despite the general view of protein-induced liquid-liquid phase-separation (LLPS) of MLOs [350, 424, 30, 38, 425–427, 28], it is becoming increasingly clearar that RNAs are not just passively recruited client molecules, rather they play an important role in MLO formation and maintenance. RNA-RNA and RNA-protein interactions have in fact been found to dictate the identity of protein condensates [428], prescribing the co-recruitment or partitioning of mRNAs encoding for functionally related or unrelated proteins in the same or demixed condensates, respectively. Moreover, a particular class of RNAs, called architectural RNAs (arcRNAs) [429, 33, 430–433], has been found to play a key structural role in MLOs: a prime example is the long non-coding RNA NEAT1_2 which templates the formation of paraspeckles, dictates their core-shell internal architecture, and precisely localises different proteins within such hierarchical assemblies [34].

To further substantiate the importance of RNA-RNA interactions in biomolecular condensate formation, biological RNAs containing tandem short nucleotide repeats have been found to undergo LLPS through pure RNA-RNA interactions, forming either liquid or gel-like assemblies both *in vitro* and *in vivo* [356]. Moreover, synthetic RNA motifs containing similar tandem repeats fused to a variety of protein-binding domains were engineered to produce spatially organised liquid-like synthetic MLOs within *E. coli* [39]. Despite such clear and promising results, control over LLPS *via* sequence engineering of the RNA building blocks is yet to be extensively researched.

In this Chapter, I introduce a novel, modular and programmable RNA nanostar-based platform capable of yielding RNA condensates from *de novo* designed DNA templates. I show that Kissing Loops, originally discovered in the Human Immunodeficiency Virus-1 (HIV-1) genome, can be used as interaction domains among the single-stranded RNA nanostructures, and engineered to produce multiple orthogonal nanostar motifs which can be simultaneously transcribed and co-assembled into distinct condensates with tunable size and growth rate. I also demonstrate that, despite the lack of precise design guidelines for Kissing Loop interactions, engineering of loop sequences and loop-loop interaction strength can profoundly influence the morphology of the RNA assemblies, producing both discrete liquid-like droplets and percolating network assemblies. The resulting RNA phases are enriched in target molecules *via* the modular embedding of fluorogenic aptamers within the RNA nanostar arms, which also enables fluorescence monitoring and imaging without the need for staining. Rather than only using input-limited biologically unavailable single-stranded DNA to produce nanostructures, this approach exploits naturally evolved transcriptional machinery, namely T7 RNA Polymerase, and biologically relevant double-stranded DNA to synthesise much larger amounts of RNA nanostructures compared to the DNA in input. Besides facilitating the design of sustained out-of-equilibrium behaviour in synthetic cells,

these results could unlock new strategies for introducing tailor-made, non-natural structures and organelles in living cells, modified to host the relevant template DNA within plasmids. Non-natural organelles could in turn be exploited for biochemical and metabolic engineering.

The results and material presented in this chapter are included in the following Publication [434]:

G. Fabrini, S.P. Nuccio, J.M. Stewart, S. Li, A. Tang, P.W.K.R. Rothemund, E.Franco, M. Di Antonio and L. Di Michele, Co-transcriptional production of programmable RNA condensates and synthetic organelles, preprint on bioRxiv and soon to be submitted for peer review.

All the work presented in this Chapter has been carried out by me, apart from DNA template purification steps, which were performed by Sabrina Pia Nuccio (Di Antonio group).

Sequences for oligonucleotides used in this Chapter can be found in Tables A.2 (RNA motifs), A.3 (DNA templates), A.4 (DNA PCR primers) in the Appendix.

4.2 RNA nanostar Design

Building on previous experience with designing four-armed DNA nanostars such as Quad-Stars, and considering the propensity of nanostars with such valency to easily undergo LLPS [435], the four-armed nanostar topology was adopted for co-transcriptionally folding RNA nanostructures as well. The designs are schematically presented in Fig. 4.1.
Due to the single-stranded nature of the co-transcriptionally folding RNA nanostars, a nick – that is, an interruption in the phosphate backbone – is present in each design. In order to improve the symmetry of the junction, the nick was placed along one of the arms, a feature that also results in strong secondary structure on the 3' end of the transcribed constructs, thus limiting aberrant transcript extension (see Section 1.4.4) and potentially facilitating correct transcription termination [311, 284].
All RNA nanostars were designed starting from the four-way junction found in the RNA tetra-square reported in ref. [215]. Arm length was chosen to be 25 base pairs (bp), similar to the one used in Quad-Stars (28 bp). Such arm length was preferred to shorter ones for the following reasons: a) longer arms and larger subunits lead to faster growth and larger final size of phase-separated condensates [436]; b) longer arms can directly accommodate fluorogenic aptamers for fluorescence monitoring and imaging, eliminating the need for staining; c) larger subunits imply higher porosity and potentially improved permeability and capability to host guest molecules.
Sequences for the nanostar arms were iteratively adjusted starting from those found in the

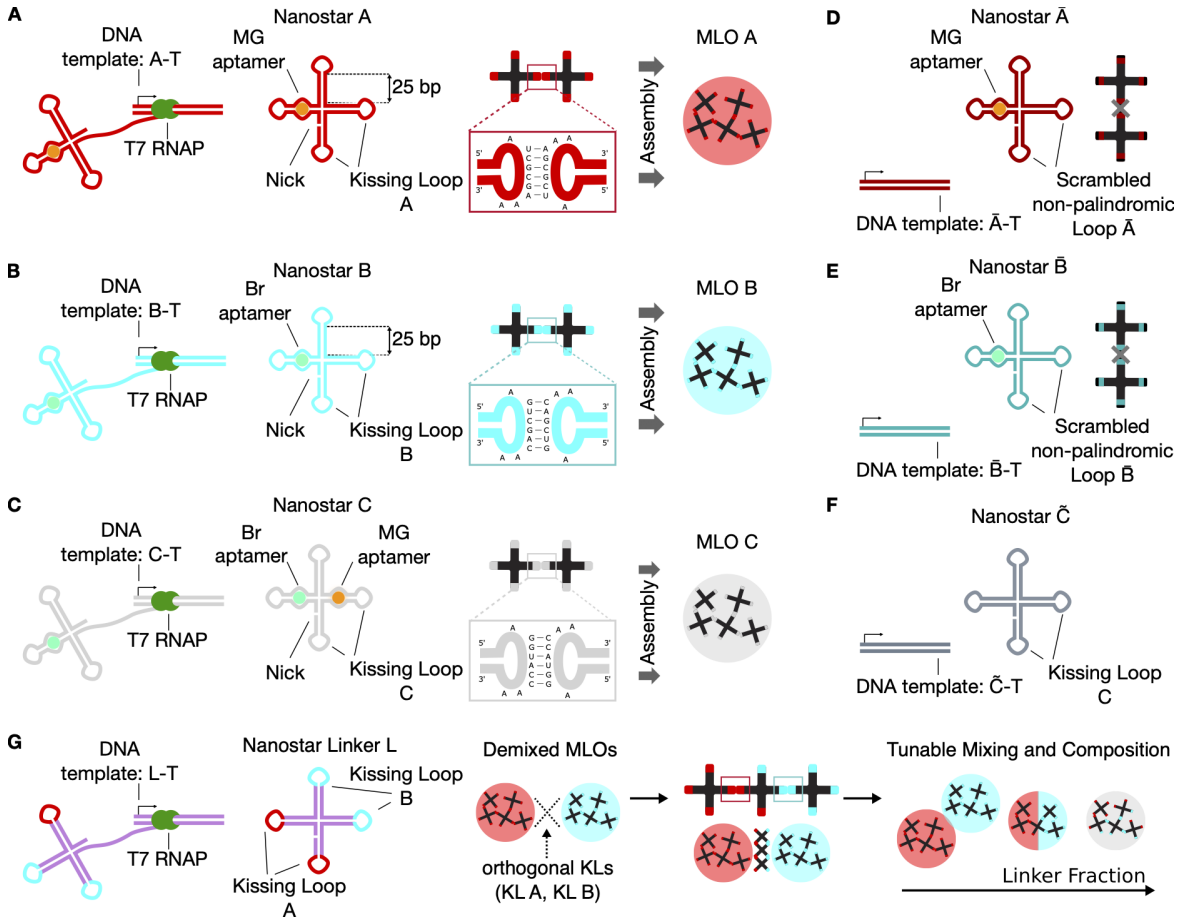


Fig. 4.1 **Topology and design of co-transcriptionally folding RNA nanostars.** Co-transcriptionally folding RNA nanostars are transcribed from fully double-stranded DNA templates indicated with the name of the nanostar followed by '-T'. **(A):** RNA nanostar A hosts a Malachite Green aptamer (MGA, see Section 1.4.2), binding Malachite Green, and can interact with other A nanostars *via* Kissing Loop A (KL A) domains to form condensate/membrane-less organelle (MLO) A. **(B):** RNA nanostar B hosts a Broccoli aptamer (BrA, see Section 1.4.2), binding DFHBI, and can interact with other B nanostars *via* KL B domains to form MLO B. **(C):** RNA nanostar C hosts both BrA and MGA on opposite arms, and can interact with other C nanostars *via* KL C domains to form MLO C. **(D, E):** RNA nanostars $\overline{\text{A}}$ and $\overline{\text{B}}$ are identical to nanostars A and B, respectively, but present scrambled, non-palindromic loop sequences, preventing condensate formation. Such lack of interactions is represented with a gray cross. **(F):** RNA nanostar $\tilde{\text{C}}$ is identical to nanostar C (including KL C loops), but lacks any fluorogenic aptamers. **(G):** RNA nanostar L lacks fluorogenic aptamers and presents pairs of adjacent KL A and KL B domains, enabling it to bridge the orthogonal A and B nanostar populations and the corresponding MLOs. Sequences for palindromic KLs are reported in zoomed insets. RNA aptamers are deliberately represented *via* simplified schematics, *i.e.* bubbles binding the corresponding ligands, due to their complex three-dimensional structure.

arms of Quad-Stars as well as sequences generated *via* the NUPACK Design module [328]. First of all, sequences were modified to follow the design rule outlined in ref. [215], which recommends replacing approximately 1 every 10 base pairs of double-stranded RNA (dsRNA) with a G-U wobble-pair to prevent too strong a secondary structure in the DNA template, which could hinder transcription. As a result, between 2 and 3 G-U wobble-pairs were inserted in each 25 bp-long dsRNA arm. Second, following the guidelines in Section 1.4.4, all transcripts were modified to start with either a GG or GC duplet, followed in most cases

by A or U residues.
The iterative adjustment proceeded until a) the minimum free energy structures predicted by NUPACK and Kinefold matched the designed topology, and b) the co-transcriptional folding trajectory in Kinefold showed stable formation of the helices within nanostar arms lacking aptamers, *i.e.* no major fluctuations indicative of helix dissociation, especially at long timescales (see Fig. B.6A). It is worth noting, in fact, that helices containing (or adjacent to) fluorogenic aptamer domains often display considerable fluctuations due to their partially single-stranded nature and complex structure, which is not well modelled by either NUPACK or Kinefold (see Fig. B.6B-C). On some occasions, the arm sequences were further refined to remove repeated nucleotide stretches that impeded their automated synthesis by the supplier (see Section 2.2).
Finally, while I did not perform experiments to prove co-transcriptional folding of the RNA nanostars, it is safe to assume that the designed nanostructures do, indeed, fold during transcription. As mentioned in Section 1.4.4, the elongation rate of T7 RNAP is ≈250 nt/s (although lower estimates have been proposed over the years), which amounts to one nucleotide being added to the nascent RNA transcript every 4 ms. This time scale is considerably larger than that of hairpin formation/folding (order of 10-100 μs) [384, 383, 437, 331, 269, 332, 438–440], and comparable with that of branch migration (order of ms) [167]. Taking into account that co-transcriptional folding of the designed nanostructures occurs through highly local interactions resulting in hairpin formation, it is reasonable to expect the nascent RNA transcripts to fold into the designed secondary structures while being synthesised.

4.2.1 Kissing Loops

In order for such single-stranded RNA nanostructures to interact with one another, it is not possible to resort to sticky ends as it is common in DNA nanotechnology, as there could only be two, at the 5' and 3' ends of each nanostructure, resulting in too low a valency for phase separation.
As mentioned in Section 1.4.1, Kissing Loops (KLs), originally found in the HIV-1 Dimerisation Initiation Site (DIS), can be used instead. KLs consist of two hairpins interacting *via* Watson-Crick complementary domains in their loops, resulting in an almost coaxial stacking of the two stem helices [212, 213] (in this work I have only investigated 180° KLs, not 120° KLs – see Section 1.4.1).
While KLs have been used in RNA nanostructures and found to be Mg^{2+}-dependent [274], precise design guidelines are yet to be formulated. Interaction strength is also not easily computed, due to the large stacking contribution from bases flanking the binding domain in the loops. However, it is clear that KLs can consist of both biologically derived and designer sequences, as long as Watson-Crick complementarity is maintained. It is also expected that a difference in KL interaction strength might result in different biophysical properties and

morphology of the resulting biomolecular condensates, similarly to what observed for DNA condensates held together by sticky ends [201].

Orthogonal palindromic Kissing Loops

Three orthogonal RNA nanostars were designed to investigate the programmability and specificity of KL interactions between RNA nanostructures, with the goal of nanostars interacting with constructs from the same population, while being orthogonal to different ones, thus resulting in multiple non-interacting assemblies. To guarantee interactions within the same population, KLs were made palindromic, taking inspiration from palindromic sticky-ends used in DNA nanotechnology [441, 442]. In this context, a nucleic acid sequence is defined palindromic if self-complementary, *i.e.* it reads the same as its own reverse complement. Three orthogonal palindromic KL domains were designed, namely KL A, KL B and KL C. Nanostar constructs will be referred to with the same name as the corresponding KL domain (*e.g.* nanostar A displays KL A) (Fig. 4.1A-C), unless otherwise specified.
KL A (5'-AUCGCGAAA-3') was adapted from the non-palindromic KL domain found in the bottom right arm of the T4 tetrahedron in ref. [215] by making it palindromic. The asymmetrical flanking bases (5'-A...AA-3') were kept due to their presence, albeit in reverse order, in the original KL found in the HIV-1 DIS (Lai variant [210, 211]), hinting at a possible role in supporting loop-loop interactions *via* base stacking.
KL B and KL C were designed to match the interaction strength of KL A while maximizing orthogonality towards it and among each other. In order to obtain domains with comparable interaction strength, only identical length (6 nt-long) palindromic sequences were considered, in the same 9 nt loop configuration with asymmetric flanking A bases, with the same base content as KL A (one A-U, two G-C base pairs). Candidate sequences were generated using a Python3 script (see Section 2.1) and accepted only if displaying an overlap with KL A of less than 2 consecutive and 2 identically spaced nucleotides. One of the resulting candidates (5'-AGGUACCAA-3') was adopted as KL C. The remaining candidates were discarded as they all followed the same pattern as KL C - two G/Cs at the start/end and one U-A pair in the middle - thus not guaranteeing enough orthogonality towards it. Therefore, the overlap constraints were relaxed to a maximum of 3 consecutive and 3 identically spaced nucleotides and candidate sequence generation was repeated. One of the resulting sequences (5'-AGUCGACAA-3') was selected as KL B due to its high similarity to the HIV-1 DIS KL (Mal variant [210], 5'-GUGCAC-3'), despite featuring inverted flanking bases (5'-A...AA-3' as opposed to 5'-AA...A-3' in HIV-1 DIS Lai variant [210, 211]).

Negative Controls with non-interacting loops

In order to further evaluate the programmability of KL interactions and their role in the formation of RNA-based condensates, negative control nanostar variants were designed to

be incapable of interacting, both within the same population and among orthogonal ones. In particular, the palindromic loops in nanostars A and B were scrambled and mutated to obtain, in each case, four non-palindromic and non-interacting loop domains. Arm and junction sequences were left untouched. The resulting non-interacting RNA nanostars will be indicated with an added overline, namely $\overline{\text{A}}$ (Fig. 4.1D) and $\overline{\text{B}}$ (Fig. 4.1E), respectively. Due to the non-interacting nature of such constructs, they will be referred to as non sticky, compared to designs with palindromic KLs (sticky).

Further RNA nanostar designs

A control modification of nanostar C, referred to as $\tilde{\text{C}}$, maintains the same KL C domain but lacks any fluorogenic aptamers (Fig. 4.1F), with the aim of probing whether results observed for C are to be ascribed to KL C interaction strength or to the higher complexity of the host nanostar, containing not one but two fluorogenic aptamers.
A linker nanostar construct, referred to as L (Fig. 4.1G), was designed to bridge otherwise orthogonal RNA nanostars A and B. This construct is a simple four-armed RNA nanostar, with no fluorogenic aptamers, comprising two adjacent KL A and two adjacent KL B domains in its loops. This chimeric structure will be used to control mixing between the two nanostar populations. Both $\tilde{\text{C}}$ and L will be used in the following Chapter.

Including minimal fluorogenic aptamer sequences in the nanostar arms

In the context of this Chapter and the following ones, Fluorescence Light-Up Aptamers (FLAPs), introduced in Section 1.4.2 play a crucial role, as they can be used to monitor the transcription of RNA constructs comprising them in their sequence [443, 243, 444]. In order to monitor multiple orthogonal RNA nanostar populations as discussed above, I chose the fluorescently orthogonal FLAPs Malachite Green aptamer (MGA) and Broccoli aptamer (BrA), introduced in Sections 1.4.2 and 1.4.2, respectively.
The minimal functional sequence for MGA was taken from refs. [330] and [241], which further report the final hairpin loop is inessential. This made it possible to easily embed MGA within the second transcribed arm of A nanostars, with the original loop replaced by KL A for nanostar-nanostar interactions, as depicted in Fig. 4.1A.
Similarly, the minimal sequence for BrA, taken from ref. [253], was included in the second transcribed arm of nanostar B, with the terminating loop replaced by KL B (Fig. 4.1B). Many RNA aptamers are known to suffer from misfolding [445, 264], which can impair their functionality. Fortunately, the folding of BrA has been reported to improve in branched scaffolds [446, 254], such as the phi29 pRNA 3WJ motif [446]. This is highly promising in view of embedding BrA within four-way nanostars, with the higher valency of such tetravalent branched junctions, compared to trivalent phi29 pRNA 3WJ motifs, allowing for more robust self-assembly through LLPS as mentioned above [435], as well as increased nanostructure

modularity, *e.g.* each nanostar could host multiple aptamers and other functional motifs such as siRNA. The formation of condensates by the nanostar scaffolds might also result in lower degradation from Dicer [222] or other cellular pathways, similarly to reports of lower degradation of RNA aptamers sequestered in RNA granules [254]. It is also worth noting that the crystal structure reported for the related Spinach aptamer in Fig. 1.12D reveals two helices in coaxial stacking around the non-canonical binding pocket, a geometrical feature which would have been very difficult to predict looking at its secondary structure. Therefore, embedding BrA in one of the RNA nanostar arms might very well lead to a coaxial arm structure, thus conserving symmetry around the junction despite the secondary structure schematic in Fig. 1.12A.
Finally, the same modified sequences were embedded in the second (BrA) and fourth (MGA) arms of nanostar C. As shown in Fig. 4.1C, the two aptamers were deliberately hosted in opposing arms to try and ensure the stability of the core nanostar motif upon co-transcriptional folding, given the intrinsically complex structure of the RNA aptamers. For the same reason, the included aptamer sequences were surrounded, where possible (*i.e.* not necessarily on both ends), by dsRNA flanking regions long enough to ensure the overall stability of the host arm helices and their quick folding upon transcription, as evaluated using Kinefold. Such dsRNA addition was limited in length with the aim of maintaining the highest possible symmetry across the four arms: the length of aptamer-hosting arms was adjusted to include 25 bp, *i.e.* not counting unpaired nucleotides, according to the minimum free energy secondary structure predicted by NUPACK.

4.2.2 DNA Template and PCR Primer Design

As mentioned in Section 1.4.4, single-stranded templates can hinder transcription and result in lower efficiency and yield, particularly in the case of transcripts with complex secondary structures such as the introduced RNA nanostars. Moreover, double-stranded DNA templates are more readily extendable to fragments cloned within a plasmid for bacterial transformation. Accordingly, I opted for a fully double-stranded, blunt-end DNA template architecture for Chapters 4-6 of this work.
For each of the designed RNA nanostructures, the nontemplate DNA strand was obtained by converting the RNA sequence to DNA (U $\rightarrow$ T) and adding a prefix comprising the consensus T7 promoter sequence (5'-TTCTAATACGACTCACTATA-3', with consensus T7 promoter highlighted in blue). This 'extended' promoter was taken from ref. [215]. The template DNA strand was obtained by simply computing the reverse complement of the nontemplate DNA strand. DNA templates will be referred to using the name of construct they code for followed by '-T'. For instance A-T is the DNA template coding for RNA nanostar A.
DNA primers for PCR amplification of the DNA templates were designed from the nontemplate and template strands, aiming for 40-60% GC content and primer length between 18 and 26 nt. DNA primer binding was verified using NUPACK, and the appropriate melting

temperature (T_M) for use with Q5 High-Fidelity DNA Polymerase was computed using the online NEB "T_m calculator", using the default 500 nM primer concentration and ensuring a difference in T_M for the two primers lower than the recommended limit (5 °C).

4.3 Characterisation

In this section, I will go over the characterisation of the presented RNA nanostars, first ensuring the correct PCR amplification of the corresponding DNA templates, and then verifying co-transcriptional folding and correct transcript size of the designed RNA constructs, all by means of agarose gel electrophoresis. The correct folding of the FLAPs within the RNA nanostars will also be assessed by fluorimetry. *In vitro* transcription in this Chapter and the following ones will be carried out at 30 °C.

4.3.1 Characterisation of PCR amplified DNA templates

DNA templates coding for RNA nanostars $\overline{\text{A}}$, A, $\overline{\text{B}}$, B, C, $\tilde{\text{C}}$ and L were PCR amplified and characterised *via* agarose gel electrophoresis to inspect their size *via* their electrophoretic mobility. As Fig. 4.2A shows, all amplification products display an electrophoretic mobility comparable to the one of double-stranded DNA fragments with the expected size (reported on top of each lane), as highlighted by the closest reference marking from the Ultra Low Range (ULR) DNA ladder in the left-most lane. This confirms the correct amplification of the starting DNA templates.
While all lanes display a single main band (or peak in the lane profiles in Fig. 4.2B), multiple spurious shorter bands can be spotted (most easily seen in the right-most lane for template L-T). Therefore, to avoid transcribing non-specific amplicons or truncated templates, the gel bands corresponding to the expected sizes were cut out, dissolved and eluted to obtain the purified DNA templates for *in vitro* transcription (see Section 2.4).

4.3.2 Characterisation of RNA transcripts and aptamer folding

Gel electrophoresis on RNA transcripts

PCR amplified and gel purified DNA templates were *in vitro* transcribed for 18 h at 30 °C using 40 nM of corresponding DNA template. The resulting transcripts were characterised in size *via* both native and denaturing agarose gel electrophoresis.
In denaturing gels, samples migrate according to their charge and primary structure: in the case of nucleic acids, denatured samples migrate according to their overall length in nucleotides (*i.e.* no base pairing). Denaturing gels can therefore confirm whether the obtained transcripts have the expected size or not. On the other hand, in native gels, samples migrate according to their charge and secondary structure: if a specific transcript folds into a nanostar

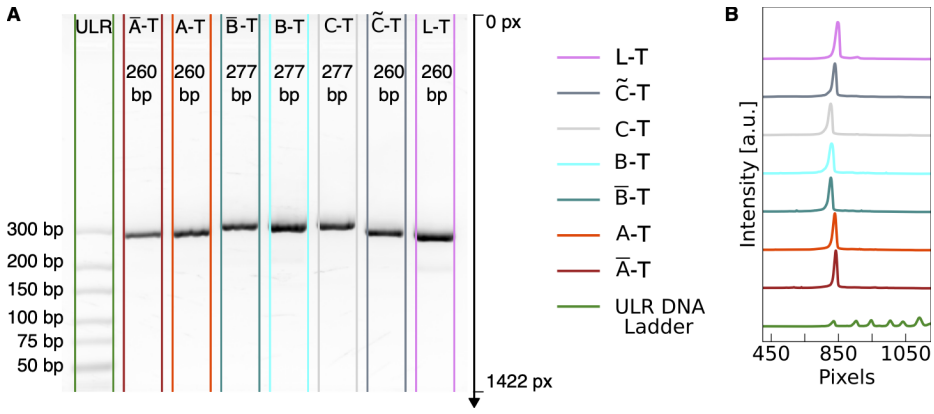


Fig. 4.2 **Agarose gel electrophoresis confirms PCR amplified DNA templates display the expected sizes.** **A:** Agarose gel representing the electrophoretic mobilities of PCR amplified DNA templates $\overline{\text{A}}$-T, A-T, $\overline{\text{B}}$-T, B-T, C-T, $\tilde{\text{C}}$-T and L-T. Ultra-Low-Range (ULR) DNA ladder is used as reference in the first lane, colour-coded in green. All templates follow the same colour-coding of the corresponding RNA constructs used in Fig. 4.1. Length in base pairs for ladder markings is reported to the side of the corresponding bands. Conversely, expected length in base pairs of amplified DNA templates is reported on top of the corresponding lane. **B:** Lane intensity profiles computed from gel in **A**. Visual inspection of the gel in **A** and the lane profiling in **B** confirms all PCR amplified templates display mobility comparable to the expected size.

topology, it will migrate as such, thus displaying an electrophoretic mobility different from single-stranded RNA with the same length.

Fig. 4.3A shows that RNA transcripts in native conditions display electrophoretic mobilities comparable to those of four-armed DNA nanostars and Quad-Stars (see Section 3.2), thus matching the size of the designed nanostar motifs, at least when examining the main band (apart from sample $\tilde{\text{C}}$ in the penultimate lane from the left, see below). Each lane, in fact, displays several shorter and sometimes longer bands and smears, but this is known to occur in gels run on crude RNA transcription mixtures, making the interpretation of both gels in Fig. 4.3A-B challenging and somewhat uncertain. It is worth noting that sample $\tilde{\text{C}}$, albeit displaying a faint band in alignment with the main one of all other transcripts, exhibits a different, faster moving main band, possibly hinting at a different folded secondary structure. Conversely, Fig. 4.3B investigates the electrophoretic mobility of the same samples in denaturing conditions, and confirms they all have consistent sizes with one another, as expected from the designed sizes: the maximum length deviation is 17 nt (transcripts $\overline{\text{B}}$, B and C) and is matched by slighlty slower and consistently positioned main bands in the corresponding lanes. As expected, all samples migrate much faster than the fastest observable ladder marking (1000 nt single-stranded RNA, due to the particular RNA reference ladder in use, see Section 2.5). Interestingly, sample $\tilde{\text{C}}$ migrates similarly to comparably sized transcripts ($\overline{\text{A}}$, A, L) under denaturing conditions, in deviation from the behaviour in the native gel in Fig. 4.3A(i). While further experiments are needed to elucidate this inconsistency, it might be explained by a correctly sized transcript, hence exhibiting a migration pattern consistent with other samples when denatured, featuring bistable secondary structures in

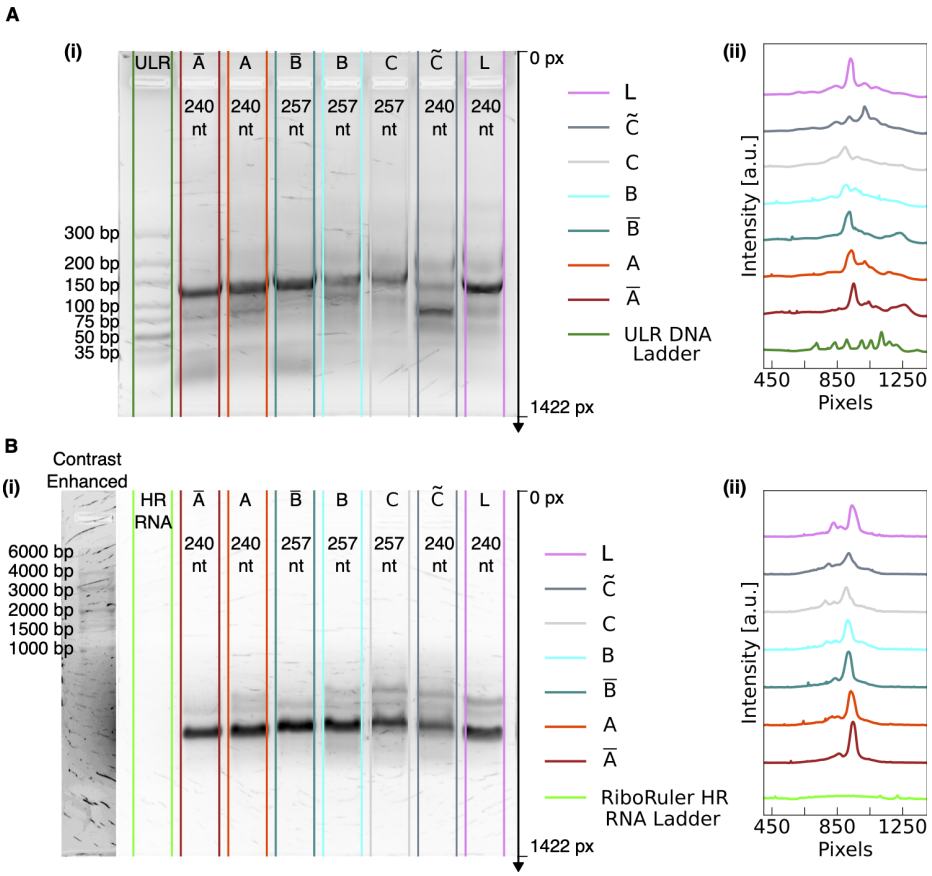


Fig. 4.3 **Agarose gel electrophoresis elucidates folding and size of RNA transcripts. A: (i)** Native agarose gel representing the electrophoretic mobilities of co-transcriptionally folded RNA nanostars $\overline{\text{A}}$, A, $\overline{\text{B}}$, B, C, $\tilde{\text{C}}$ and L. Ultra-Low-Range (ULR) DNA Ladder is used as reference in the first lane, colour-coded in green. **(ii)** Lane intensity profiles computed from gel in **A (i)**. Visual inspection of **(i)** and **(ii)** reveals a clear main band in (almost) all transcripts between 100 and 150 bp markings of ULR, comparable to the one observed in four-armed DNA nanostars and Quad-Stars in Fig. 3.2A. **B: (i)** Denaturing agarose gel representing the electrophoretic mobilities of the same RNA transcripts in **A**. RiboRuler High-Range (HR) RNA Ladder is used as reference in the first lane, colour-coded in lime green. Due to extremely low contrast, a contrast-enhanced reference lane is reported on the left-hand side. **(ii)** Lane intensity profiles computed from gel in **B (i)**. Visual inspection of **(i)** and **(ii)** confirms the presence of a main band at almost uniform length for all denatured transcripts, as expected from the designed motifs. Length in base pairs for ladder markings is reported to the side of the corresponding bands. Conversely, expected length in base pairs of RNA transcripts is reported on top of the corresponding lane. All samples apart from reference ladders follow the same colour-coding as in Fig. 4.1.

native conditions, namely the designed nanostar (expected slower, fainter band), and a more compact topology (faster, main band).

Fluorimetry confirms folding of fluorogenic aptamers in the arms of RNA nanostars

Fluorogenic RNA aptamers must properly fold to reach the conformation required for binding the target dyes and enhance their fluorescence emission [447]. In order to confirm proper folding and functionality of the aptamers embedded in the arms of the designed RNA nanostars, the DNA templates coding for nanostars A, B and C, as well as the controls with scrambled KL domains $\overline{\text{A}}$ and $\overline{\text{B}}$, were *in vitro* transcribed at 30 °C, and fluorescence

excitation and emission scans were performed after 48 h using a microplate reader.
Fig. 4.4A confirms the functionality of the aptamers within co-transcriptionally folding RNA nanostars, which all exhibit the excitation (dashed lines) and emission (solid lines) signatures expected from the corresponding aptamers. In particular, the top row shows the excitation and emission of nanostars A and $\overline{\text{A}}$, which host the MGA and clearly exhibit its characteristic excitation (peak at 616-630 nm) and emission (peak at 650-660 nm) spectra when bound to Malachite Green (MG) [241, 252], with no significant signal in the BrA+DFHBI band. The middle row presents the same information for nanostars B and $\overline{\text{B}}$, which contain the BrA and display its characteristic excitation (peak at 447 nm) and emission (peak at 500-510 nm) spectra when bound to DFHBI [253], with no significant signal in the MG+MGA band. Finally, the bottom row of Fig. 4.4A demonstrates that nanostar C presents, as expected, both the emission peaks of MGA bound to MG and BrA bound to DFHBI.

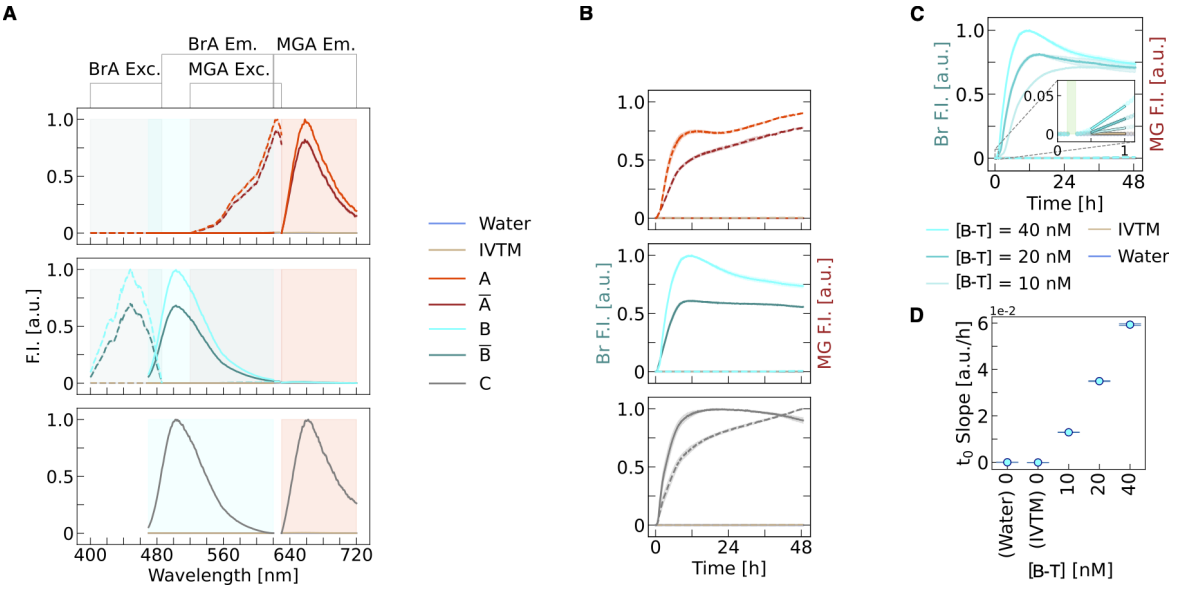


Fig. 4.4 **Fluorogenic aptamers enable monitoring of the co-transcriptional folding of RNA nanostars. A:** Fluorescence excitation (dashed lines) and emission (solid lines) scans for co-transcriptionally folded RNA nanostars in bulk: A and $\overline{\text{A}}$ (top row), B and $\overline{\text{B}}$ (middle row), C (bottom row, emission scans only). Negative controls, *i.e.* dyes in water (Water) and dyes in *In Vitro* Transcription Mixture (IVTM), are displayed in all rows. Excitation/emission scans in the DFHBI + Broccoli aptamer (BrA) band are outlined in cyan, while those in the Malachite Green (MG) + MG aptamer (MGA) band are outlined in red as per the labels on top. Data are shown as normalised mean (solid or dashed lines) ± standard deviation (shaded regions) from triplicates. **B:** Fluorescence intensity (F.I.) kinetics monitoring the co-transcriptional folding and self-assembly of samples in **A** (same row-wise arrangement), resulting from *in vitro* transcription carried out in bulk at 30 °C with [DNA template] = 40 nM. **C:** The initial transcription rate, examined through the proxy of initial fluorescence intensity increase rate, is proportional to the DNA template concentration (here [B-T] for RNA nanostar B). The characteristic peak-decrease feature indicative of condensate formation is more prominent at higher [B-T]. The zoomed inset shows the DNA template injection window (no injection in negative controls) shaded in light lime green, as well as the 30 timepoints used for linear fitting highlighted by surrounding boxes. In **B** and **C**, the left/right y-axis depicts fluorescence intensity in DFHBI + BrA (Br F.I.)/MG + MGA (MG F.I.) channels in solid/dashed lines, respectively. **D:** Initial fluorescence intensity increase rate computed *via* linear fitting of the 30 timepoints highlighted in **C** (inset). Data are presented as fitted slopes (circular markers) ± standard deviation from covariance matrix of the linear fit (error bars). Data in **B, C** are presented as normalised (see Section 2.11.4) mean (solid/dashed lines) ± standard deviation (shaded regions) from triplicates.

Fluorogenic aptamers enable fluorescence monitoring of co-transcriptional folding and self-assembly of RNA nanostars

The embedded FLAPs also enable fluorescence monitoring of the co-transcriptional folding and self-assembly of the RNA nanostars, as showed in Fig. 4.4B. Upon injection of the corresponding DNA templates, all samples, apart from negative controls (Water: dyes in water; IVTM: dyes in *In Vitro* Transcription Mixture), display an increase in fluorescence intensity due to aptamers being transcribed as part of the nanostars, folding, binding the cognate dyes and enhancing their fluorescence emission. After an initial phase of rapid RNA synthesis, transcription significantly slows down or plateaus depending on the specific RNA nanostar variant, due to loss of T7 RNAP activity and/or nucleotide depletion [448, 315, 316]. In particular, signal from the Broccoli aptamer reaches a plateau before 12 h, whereas the signal from the Malachite Green aptamer keeps increasing for the entire measurement timeframe. This difference in long-term behaviour can be observed in both single-aptamer (A, B) and double-aptamer (C) constructs, and as such it is likely related to the difference in binding affinity and/or photophysics of the two fluorogenic dyes-aptamer complexes [255, 252, 258], rather than to the host RNA nanostars.
Finally, samples with palindromic KL domains (A, B) display a characteristic peak followed by a decrease which is absent in corresponding samples with scrambled KL domains ($\overline{\text{A}}$, $\overline{\text{B}}$). This phenomenon is most likely due to fluorescent material being formed in the bulk (increase), forming discrete condensates and sinking to the bottom of the wells, below the plane of focus (decrease). This mechanism also explains why this phenomenon is not observed in construct C, which instead forms mesh-like assemblies extending in the three dimensional space of the microplate well (see Fig. 4.5A-B below).
The initial rate of transcription has been observed to be proportional to the concentration of DNA template. Figs. 4.4C, D and E confirm this observation holds for the RNA nanostar B system, with the initial fluorescence intensity increase rate (t_0 slope), which acts as a proxy for the initial transcription rate, increasing almost linearly with increasing DNA template concentration ([B-T] from 10 to 40 nM). It can also be observed in Fig. 4.4C that the characteristic peak-decrease feature observed in nanostar systems with palindromic KL domains becomes less and less evident with lower DNA template concentration, due to slower transcription and less sudden sinking of discrete condensates, hinting at the possibility of tuning the rate of condensate formation by varying the DNA template concentration in input.

4.4 Co-transcriptionally folding RNA nanostars form liquid condensates

The FLAPs hosted within the RNA nanostar arms also make it possible to image their self-assembly using fluorescence microscopy, thus observing whether the interactions among the

designed palindromic KL domains can indeed drive the formation of RNA-based condensates. Fig. 4.5A-B confirms that nanostars A, B and C all form RNA-based assemblies, which emerge from nucleation and growth processes before undergoing coalescence events. [449–451, 436, 452]. In particular, constructs A and B form discrete spherical liquid-like condensates, while nanostar C forms an extended percolating network. As expected from the aptamers in use within each nanostar, A condensates are enriched in Malachite Green (thus appearing red), B condensates in DFHBI (thus appearing cyan) and the C mesh-like assembly is enriched in both Malachite Green and DFHBI (thus appearing both red and cyan, which yields white in composite micrographs).
Notably, on long timescales (tens of hours), A condensates exhibit the formation of inner bubbles or vacuoles, *i.e.* dilute liquid phases which diffuse and undergo fusion with one another. The larger the A-type condensate, the larger and more numerous the vacuoles. Similar phenomena have been observed in both naturally occurring and synthetic biomolecular condensates, and they have been linked to compositional changes which happen much faster than the (comparatively slower) dynamics within the condensates. Such changes have been found to be driven by temperature [453], RNA synthesis [454, 455] or degradation processes [442, 456]. Given the constant experimental temperature and the limited temporal catalytic activity of the T7 RNAP, the latter option seems the most likely: on long timescales, possibly due to loss of activity of the RNase inhibitor, some RNA nanostars are degraded, and start accumulating within the condensates, inducing the formation of vacuoles due to the related osmotic pressure build-up [442, 456].
It is also worth noting that nanostars A, B and C share consistent portions of their sequences, apart from the aptamer and KL domains, and they all readily form condensates. In particular, C is identical to B apart from the fourth transcribed arm, which shares the same sequence as the second transcribed arm of nanostar A (MGA domain), while B (and C as just mentioned) shares the entire first arm and most of the third one (apart from two swapped bases) with nanostar A. This proves that the core RNA nanostar architecture is robust and highly tolerant to sequence modifications, making it modular and allowing for aptamer insertions and KL domain edits.

4.4.1 Growth kinetics

The growth kinetics of the RNA condensates can be investigated by image analysis and segmentation of the assembly timelapses, followed by Chord Length Distribution (CLD) calculation. Here, the mean of the CLD, μ_{CLD}, is used as a proxy for the typical object size [457–459, 420]. As showed in Fig. 4.5B(i), μ_{CLD} quickly increases at first for all constructs, then plateaus around 12 h for condensates A and B, while it keeps growing for the entire experimental timeframe (48.5 h) in the case of the percolating network formed by nanostar C. The correspondence of these plateau timescales with the ones observed *via* fluorimetry in Fig. 4.4B hints at a scenario where the condensate growth either slows down

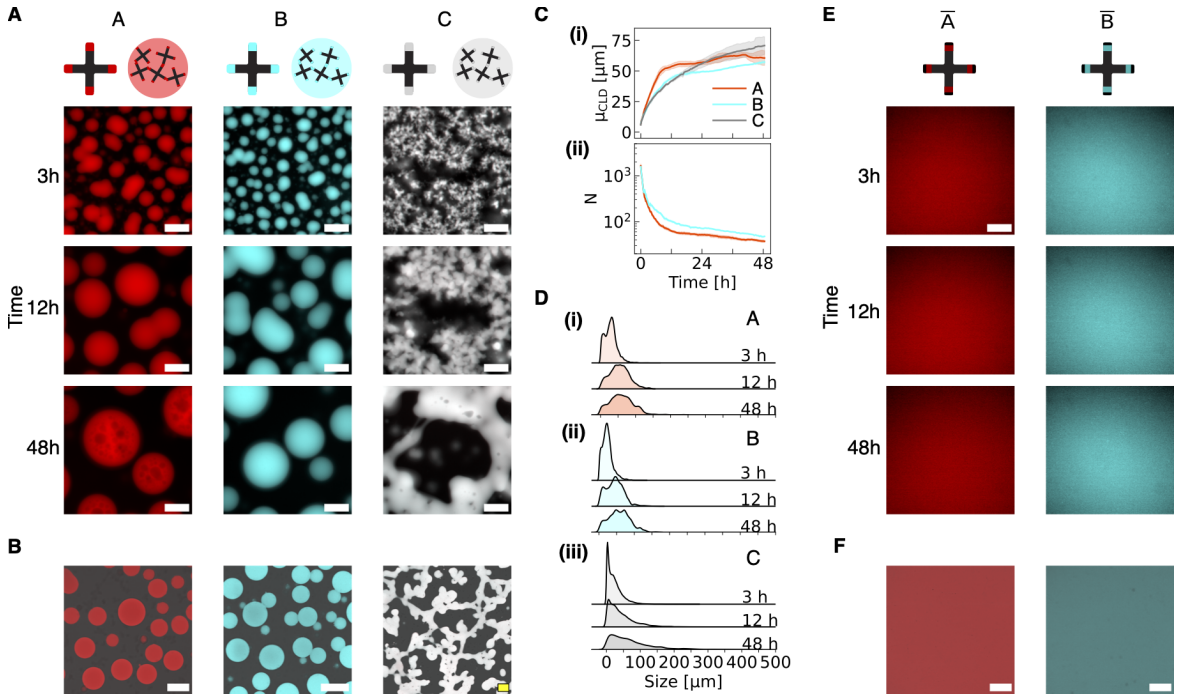


Fig. 4.5 **Programmable Kissing Loop interactions drive the self-assembly of RNA condensates. A:** Epifluorescence micrographs of representative timepoints (3, 12, 48 h, along vertical axis) depicting the bulk self-assembly at 30 °C of condensates from three nanostar variants A, B and C, respectively, accompanied by simplified schematics on top. **B:** Confocal micrographs depicting samples in **A** after more than 48 h from start of transcription. Scale bar for nanostar C assemblies is reported in yellow for ease of visualisation. **C: (i)** Evolution of condensate mean size *vs* time examined *via* the mean of the Chord Length Distribution (μ_{CLD}). μ_{CLD} increases with time, plateauing for species forming spherical condensates (A, B), but not for more rigid, mesh-like assemblies (C). **(ii)** Time dependence of the average number (N) of condensates (semi-log y scale) observed in a microscopy field-of-view (FOV), used to track growth kinetics. As condensates coalesce, their number progressively decreases. Due to nanostar C forming interconnected mesh-like assemblies spanning the whole sample, these data are only presented for A and B nanostars. Data are presented as mean (solid lines) ± standard deviation among replicates (shaded regions) from three non-overlapping FOVs within the same sample. **D:** Chord Length Distribution (CLD) ridge-plots at timepoints matching the ones in **A** for constructs A **(i)**, B **(ii)** and C **(iii)**, respectively. **E:** Epifluorescence micrographs of representative timepoints (3, 12, 48 h, along vertical axis) depicting the lack of condensate formation in the case of control nanostars $\overline{A}$ and $\overline{B}$, respectively, accompanied by simplified schematics on top. Lack of complementarity among loop domains leads to nanostars remaining disperse in solution and (almost) uniform fluorescence distribution. **F:** Confocal micrographs depicting samples in **E** after more than 48 h from start of transcription. Epifluorescence micrographs in **A, E** are contrast-enhanced (see Section 2.11.5). Scale bars are 50 μm. Confocal micrographs in **B, F** are pristine. Scale bars are 100 μm. All samples were obtained by *in vitro* transcription with [DNA template] = 40 nM.

or completely stops when monomer transcription ceases due to either loss of T7 RNAP activity or depletion of available rNTPs, and coalescence events among large droplets become increasingly rarer due to lack of substantial Brownian motion. Particle analysis and object counting for the discrete condensates formed by A and B reveals that, as expected, the increase in average size is accompanied by a corresponding sharp decrease in object number due to coalescence events, as depicted in Fig. 4.5.

Fig. 4.5C presents the entire CLDs for constructs A (i), B (ii) and C (iii) at the same representative timepoints depicted in Fig. 4.5A. Notably, the CLD for construct C is much

broader (with some chords exceeding 400 μm) and positively skewed compared to the ones for A and B due to the intrinsically asymmetrical and non convex shape of its percolating mesh-like assemblies, compared to the discrete spherical condensates observed in orthogonal A and B samples.

4.4.2 Kissing Loop interactions are programmable

While Fig. 4.5A-B show that sticky nanostars, *i.e.* with palindromic loops, form condensates, this alone does not prove that condensation occurs through the designed KL interactions. However, non sticky control nanostar variants $\overline{\text{A}}$ and $\overline{\text{B}}$, featuring scrambled non-interacting loops, fail to yield condensates, and rather remain dispersed in the bulk, as depicted in Fig. 4.5E-F. This unequivocally proves that RNA condensates formed by sticky nanostar designs A, B and C indeed rely on the designed loop-loop interactions, making KLs programmable binding domains in the presented RNA platform, just like regular sticky ends in multi-stranded designs.

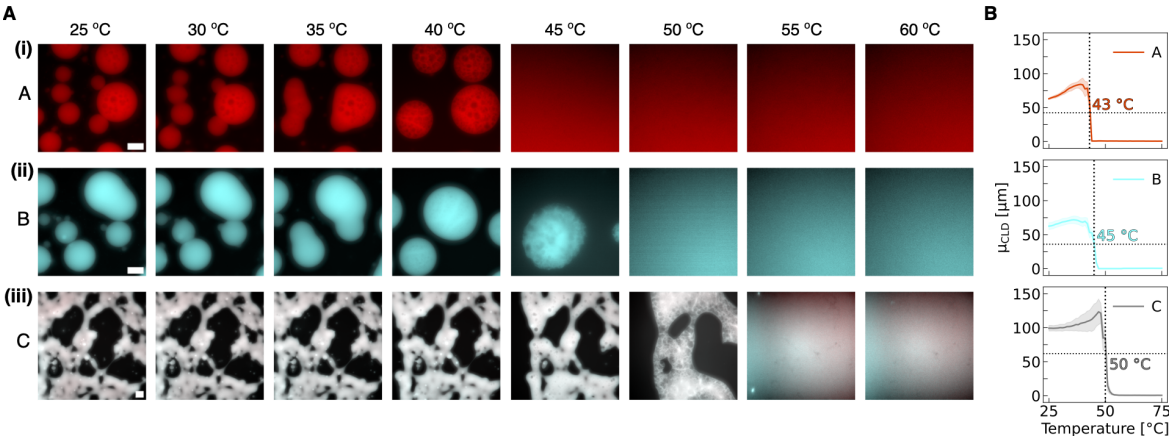


Fig. 4.6 **Thermal ramps reveal different Kissing Loop interaction strengths. A:** Contrast-enhanced epifluorescence micrographs of representative temperatures depicting melting of condensates from RNA nanostars A **(i)**, B **(ii)** and C **(iii)** in bulk. All micrographs in each series (i, ii, iii) show the same FOV area and, as such, share the same scale bar as the first micrograph in the series. Scale bars are 50 μm. **B:** Change in condensate size with increasing temperature is examined *via* μ_{CLD}. Vertical dotted lines indicate the temperature values (reported in text) at which condensates reach 50% of their original μ_{CLD} at 25 °C, indicated by the horizontal dotted line. Data are presented as mean (solid lines) ± standard deviation among mean replicates (shaded regions) from three non-overlapping fields-of-view within the same sample.

4.4.3 Melting of bulk condensates sheds light on KL interaction strength

The observed difference in morphology between the discrete liquid-like spherical condensates formed by nanostar variants A and B and the hydrogel-like percolating network assembly of nanostar C can be attributed to differences in the interaction strength across the orthogonal KL A, KL B and KL C domains. In order to test this, samples were allowed to self-assemble in bulk at 30 °C for 48 h, kept in a dark environment at room temperature for approximately

12 h, and finally subjected to thermal ramps from 25 °C to 75 °C at 1 °C increase every 15 min. Fig. 4.6A shows condensates first swelling, with bubbles forming on the inside, and then completely disassembling, revealing a clear trend in melting temperature of the Kissing Loop interactions, with KL A being the weakest, followed by KL B and KL C, respectively. Fig. 4.6B quantifies such melting temperature as the temperature at which the condensates reach half of their initial μ_{CLD} (*i.e.* at 25 °C). This observation matches with the different morphologies observed in Fig. 4.5A. Type A condensates are the most fluid, due to the transcription temperature ($T = 30°\mathrm{C}$) being close to the melting temperature of KL A interactions ($T_{\mathrm{M}} = 43°\mathrm{C}$), a factor which might contribute to the emergence of the bubbles/vacuoles discussed above. Type B condensates display slightly improved thermal stability of interactions ($T_{\mathrm{M}} = 45°\mathrm{C}$), still allowing them to relax into discrete liquid assemblies. Finally, KL C interactions are much stronger ($T_{\mathrm{M}} = 50°\mathrm{C}$), explaining the mesh-like morphology observed in Fig. 4.5A: assemblies start nucleating, and soon reach and bind to the surrounding neighbours prior to fully relaxing and coarsening, thus remaining stuck in a percolating network.

4.5 Programmable Kissing Loop interactions enable orthogonal condensate formation

The designed orthogonality among KL A and KL B domains enables the simultaneous transcription of RNA variants A and B, which readily form distinct, co-existing condensates, as presented in Fig. 4.7A-B. Consistently with previous results on nanostar control variants, micrographs in Fig. 4.7C-D show that if one of the simultaneously transcribed RNA motifs is made non sticky, condensates of the sticky species co-exist with dispersed RNA nanostars of the non sticky one (left: $\mathrm{A}\,\&\,\overline{\mathrm{B}}$, middle: $\overline{\mathrm{A}}\,\&\,\mathrm{B}$). If both variants are made non sticky, then no condensates are formed at all (right: $\overline{\mathrm{A}}\,\&\,\overline{\mathrm{B}}$).

4.5.1 Tuning the relative size of orthogonal condensates *via* DNA template ratio in binary systems

Notably, micrographs in Fig. 4.7A-B make it evident that varying the DNA template concentration ratio, [A-T]/[B-T], provides a simple strategy to tune the relative sizes of orthogonal condensates, as DNA template concentration influences transcription rate, as seen in Fig. 4.4C, and overall yield of the encoded RNA nanostars. Higher (relative) amounts of one DNA template, in fact, produce larger and faster coarsening condensates of the corresponding type, as depicted in Fig. 4.7E(i)-(ii), respectively. To further support this result, violin plots in Fig. 4.7F present the radial size distributions (after more than 48 h) of A (red) and B (blue) condensates with varying DNA template ratios, as determined by image segmentation and Circle Hough Transform (see Section 2.11.12). When the DNA template ratio [A-T]/[B-T]

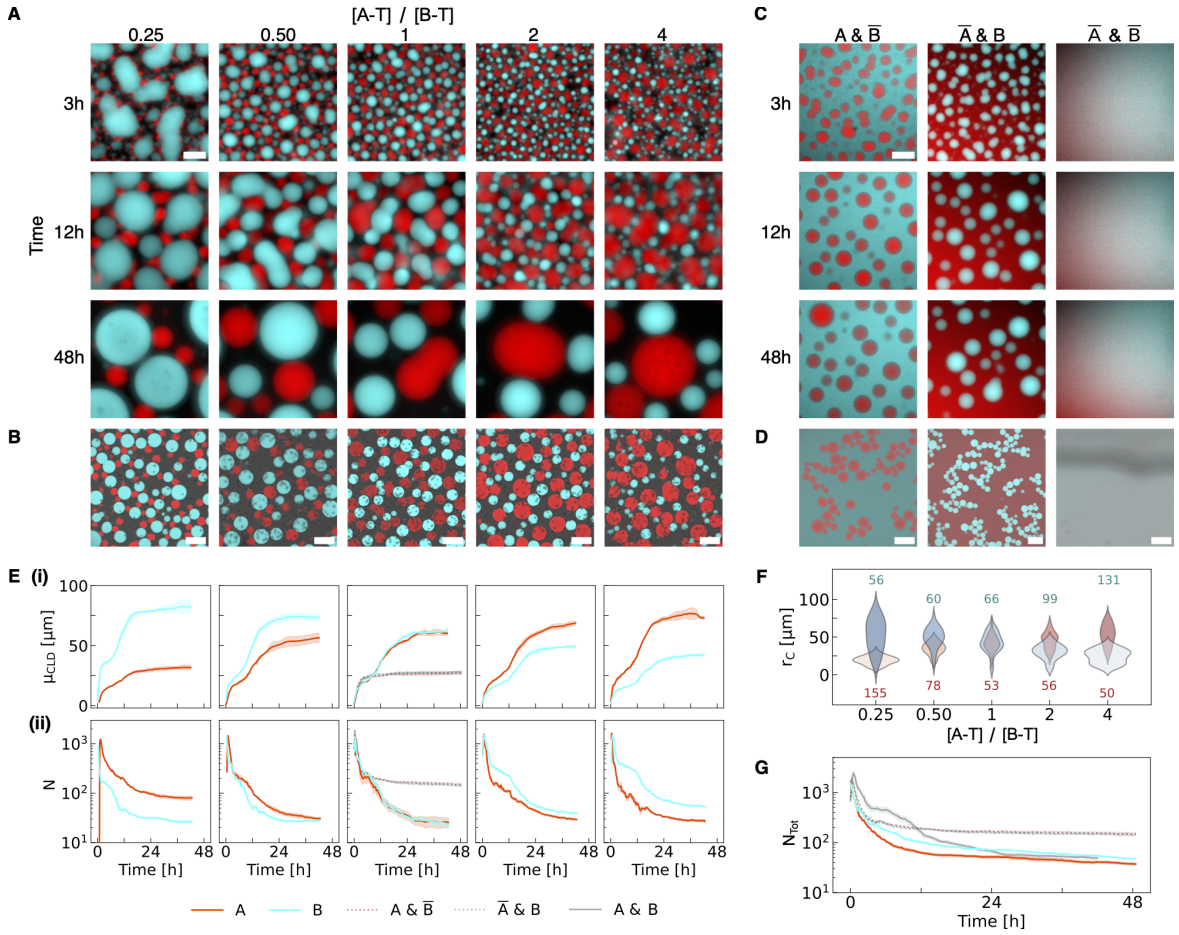


Fig. 4.7 **Simultaneous transcription and condensate co-assembly in bulk binary RNA nanostar systems.** **A:** Epifluorescence micrographs of representative timepoints (3, 12, 48 h, along vertical axis) depicting the bulk self-assembly of binary systems with different [A-T]/[B-T] ratios (horizontal axis). **B:** Confocal micrographs depicting samples in **A** after more than 48 h from start of transcription. Small spherical aggregates visible in bright-field are most likely magnesium pyrophosphate precipitates, as discussed in Section 1.4.4. **C:** Same as **A** for control stoichiometric mixtures of A-T and B-T where either or both code for non sticky nanostar variants (indicated with an overline: $\overline{\text{A}}$, $\overline{\text{B}}$). **D:** Confocal micrographs depicting samples in **C** after more than 48 h from start of transcription. **E:** Time-dependent evolution of **(i)** size and **(ii)** number (semi-log y scale) of A and B condensates in binary systems with varying [A-T]/[B-T] template ratio, as examined *via* μ_{CLD}. Stoichiometric binary mixtures where one nanostar species is made non sticky are included in the middle panel (dotted lines). Full CLDs at representative timepoints are provided in the Appendix (see Fig. B.7). **F:** Violin plots displaying the radial size distributions of condensates A (reds) and B (blues) *vs* [A-T]/[B-T] DNA template ratio after more than 48 h from start of transcription. Colour-coded text on top/bottom reports the number of examined condensates of each type. **G:** Comparison of the time dependence (in semi-log y scale) of the total number of condensates N_{Tot} (sum of A and B condensate numbers) in single and binary systems. Data for single species systems is the same as reported in Fig. 4.5C(ii). Data in **E, G** are presented as mean (solid/dotted lines) ± standard deviation among replicates (shaded regions) from three non-overlapping fields-of-view within the same sample. Epifluorescence micrographs in **A, C** are contrast-enhanced (see Section 2.11.5), and represent the same field-of-view area, thus sharing the scale bar with the corresponding top left micrograph. Scale bars are 50 μm. Confocal micrographs in **B, D** are pristine. Scale bars are 100 μm.

increases, the radial size distribution for B condensates shifts to progressively smaller values, while the one for A condensates shifts in an opposite, increasing direction, displaying a characteristic scissor-like trend.

4.5.2 Growth kinetics in binary systems

As for systems with a single RNA species, the growth kinetics of binary nanostar systems were investigated by considering the time evolution of μ_{CLD} and the average number of condensates per field-of-view, N, as depicted in Fig. 4.7E(i)-(ii), respectively. Fig. 4.7E(i) demonstrates that in binary systems of sticky nanostars, at all [A-T]/[B-T] DNA template ratios, condensates grow in two-stages: an initial sharp growth terminated by a short-lived, intermediate plateau, followed by a second, slower growth phase ultimately resulting in saturation. This additional growth phase is not observed in systems with a single RNA species (Fig. 4.5C(i)), nor in stoichiometric binary systems where one RNA species is non sticky (Fig. 4.7E(i), middle panel, dotted profiles).
Micrograph sequences from the timelapses of such binary sticky systems (see Figs. B.8, B.9, B.10, B.11, and B.12 in Appendix) reveal that the intermediate plateau is reached when condensates of each type become transiently hindered in their ability to reach and fuse with same-type neighbours due to caging among nearest-neighbour condensates of the opposite type. However, a few coalescence events still happen, thus markedly reducing lateral crowding, as the larger droplets resulting from coalescence occupy less space in the horizontal space provided the overall volume is conserved. This increase in motion freedom triggers a positive-feedback process that rapidly releases the system and drives further coalescence. When considering stoichiometric mixtures of sticky nanostars, indeed, both A and B condensates reach larger final sizes than the respective counterparts in systems where the other-type RNA species is non sticky, proving that such spatial crowdedness and the resulting steric hindrance from opposite-type non-binding condensates ultimately promotes a higher degree of coalescence. Consistently with these observations, Fig. 4.7E(ii) shows that two-stage growth in binary sticky mixtures is accompanied by a two-stage decreasing trend in the average number of condensates per field-of-view (N). It can also be observed by comparing Figs. 4.7E(i) and (ii) that condensates corresponding to the most abundant DNA template not only grow larger, as discussed in the previous section, but coarsen faster and to a fuller extent, reaching lower final numbers. Finally, Fig. 4.7G confirms the above-mentioned fuller degree of coalescence in binary sticky mixtures (solid grey line) compared to binary systems with one non sticky species (dotted lines), as evidenced by the lower overall condensate number, N_{Tot}, *i.e.* the sum of average A and B condensate numbers per microscopy field-of-view. Despite the intermediate hindrance to coalescence due to multiple orthogonal species, in fact, N_{Tot} for binary sticky mixtures still reaches values comparable to those of single sticky nanostar systems (solid red/cyan lines).

4.5.3 Ternary systems

Simultaneous transcription and condensate co-assembly can be extended to all three orthogonal RNA nanostar variants (A, B, C) as shown in Fig. 4.8A.

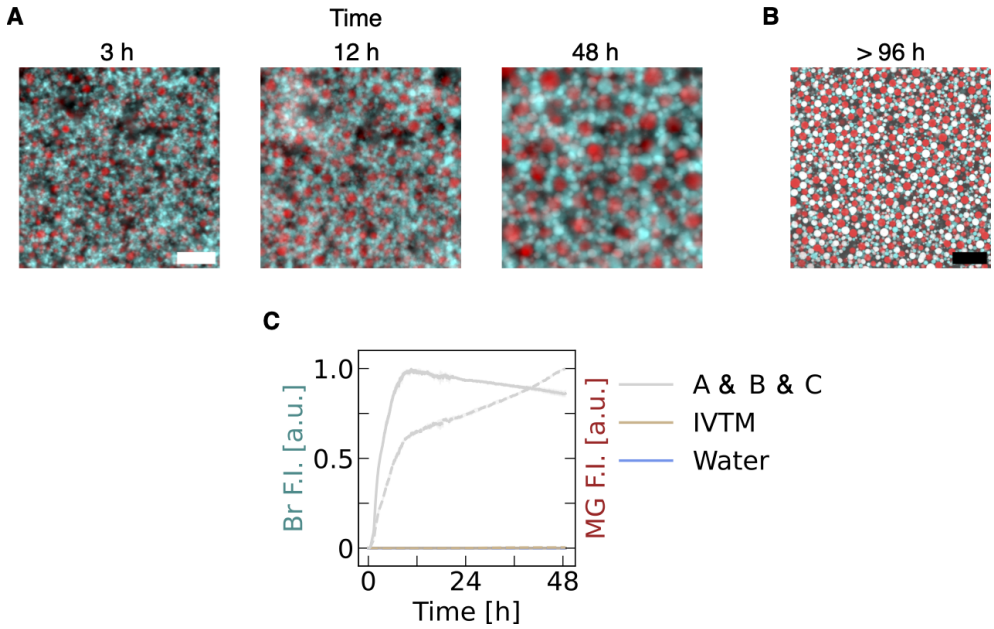


Fig. 4.8 **Simultaneous transcription and condensate co-assembly in bulk ternary systems of RNA nanostars. A:** Epifluorescence micrographs of representative timepoints (horizontal axis) depicting the bulk self-assembly of ternary stoichiometric mixtures of A-T, B-T and C-T (all coding for sticky nanostars). **B:** Confocal micrographs depicting samples in **A** after more than 96 h from start of transcription. **C:** Fluorescence intensity (F.I.) kinetics of co-transcriptional folding and self-assembly of ternary system in bulk at 30 °C with [A-T] = [B-T] = [C-T] ≈ 13.33 nM. The left/right y-axis depicts fluorescence intensity in DFHBI + BrA (Br F.I.)/MG + MGA (MG F.I.) channels in solid/dashed lines, respectively. Data in **C** are presented as normalised (see Section 2.11.4) mean (solid/dashed lines) ± standard deviation (shaded regions) from triplicates. Epifluorescence micrographs in **A** are contrast-enhanced (see Section 2.11.5), and represent the same field-of-view area, thus sharing the scale bar with the first micrograph. Scale bars are 50 μm. Confocal micrograph in **B** is pristine. Scale bar is 100 μm.

One notable difference is the much slower growth dynamics of condensates, with all species displaying very small spherical droplets at the 48 h mark. However, Fig. 4.8B confirms that, after allowing sufficient time for relaxation, *i.e.* approximately 96 h from the start of transcription, all three species form distinct spherical condensates, including nanostar C, which instead formed mesh-like assemblies in Fig. 4.5A-B. This difference in morphology is likely driven by A and B hindering the formation of the percolating C network in favor of smaller and more readily relaxing aggregates: when C condensates start nucleating, they cannot easily coalesce with same-type neighbours, and thus manage to relax to more regular and compact morphologies on the shorter length scales. As for single constructs, the self-assembly of stochiometric mixtures of A, B and C nanostars can be monitored *via* fluorimetry as well, as shown in Fig. 4.8C, which once again showcases the difference in long-term behaviour between the BrA+DFHBI and MGA+MG aptamer-dye complexes discussed in Fig 4.4B.

4.6 Chapter summary and outlook

In this chapter, I have presented a platform yielding programmable and orthogonal RNA condensates from co-transcriptionally folding RNA nanostar building blocks. These RNA nanostars fold upon transcription by T7 RNA Polymerase from designer double-stranded DNA templates using commercially available *in vitro* transcription kits. The single-stranded architecture dictated by co-transcriptional folding removes stoichiometry issues which often lower the assembly efficiency of multi-stranded nanostructures. Thanks to the variety of known RNA aptamers and the proved robustness of the RNA nanostars against design variations, fluorogenic aptamers can be easily embedded in a modular fashion within the arms of the nanostructures, making fluorescence monitoring and imaging straightforward, with no need for staining.

The introduced RNA nanostars form condensates thanks to Kissing Loop interactions among palindromic loop domains. By engineering the sequences of such KL domains, it is possible to obtain assemblies with varying relaxation timescales, morphologies and viscosities, from discrete liquid-like spherical condensates to more hydrogel-like percolating networks. It is also possible to programmatically prevent RNA nanostars from interacting altogether by scrambling the loop sequences and making them non-palindromic and non-interacting, thus leading to RNA nanostars remaining disperse in the bulk with no condensate formation.

Previous reports of RNA phase-separated assemblies focused on condensates either not entirely consisting of RNA (RNA and RNA-binding proteins [350, 428]) or formed by biologically occurring, tandem repeat-rich sequences [356, 460, 39] which leave no control on the condensate formation process apart from *a posteriori* disruption strategies. The introduced platform enables an unprecedented degree of control over the formation of RNA condensates, almost on par with the one the DNA nanotechnology community has developed for DNA-based condensates relying on nanostructures interacting through sticky ends. Similarly, KLs can be made orthogonal, yielding multiple RNA nanostar populations capable of simultaneous transcription and orthogonal condensate co-assembly. The relative size of co-assembled condensates can be tuned by simply varying the concentration ratio among the corresponding DNA templates. Combined with the proved influence of DNA template concentration on transcription rate, this result should also enable control over the condensate formation rate and the resulting size distribution in single nanostar systems by simply varying the corresponding DNA template concentration.

Most notably, these RNA nanostars are transcribed from designer DNA templates, thus allowing for direct integration with DNA-based synthetic cell frameworks and cytoskeleta, similarly to what recently showed in hybrid DNA-RNA approaches by Leathers *et al.* [154] and Malouf *et al.* [411]. This platform could also be used to produce synthetic MLOs in bacteria by integrating the DNA templates coding for the RNA building blocks in plasmids *via* molecular cloning, thus paving the way for hybrid synergistic approaches combining living and synthetic cell agents. In Chapter 5, I demonstrate how this platform can be encapsulated

in the simplest artificial cells, namely water-in-oil droplets obtained *via* emulsion methods, leading to the formation of RNA-based MLO mimics within primitive synthetic cells.

Chapter 5

RNA condensates as programmable and orthogonal organelles in droplet-based artificial cells

5.1 Introduction

Biologically occurring membraneless organelles (MLOs) are subcompartments within the cytoplasm or nucleus of cells. As they are not membrane-bound, they are incapable of separating cytotoxic reactions from the rest of the cell, carrying out homeostatic reactions or maintaining a pH gradient [30]. Nonetheless, they are involved in a multitude of regulatory processes [28], from facilitating enzymatic reactions to suppressing client molecule activity in the bulk cytoplasm, as well as promoting one biomolecular pathway over another, thus shifting metabolic flux. They have also been shown to suppress cellular noise [461, 39], for instance variations in transcriptional or translational levels, and increase overall cellular fitness [462], which explains why they are so widespread in eukaryotes.

Interestingly, such functional variety is directly matched by their structural complexity and hetereogeneity in a fascinating shape-function synergy: MLOs are rarely homogeneous, single-phase liquid droplets, rather they exhibit multiple liquid or hydrogel-like subcompartments arranged in a hierarchical architecture. This morphological complexity arises from differences in the material properties encoded in the sequence of their biopolymer subunits, hinting at a strategy for further tuning *via* sequence engineering.

The ubiquitous presence and functional variety of MLOs in eukaryotic cells clearly suggest that organelle-mimicking structures would greatly enhance the biomimicry of synthetic cell systems. However, a high degree of control and programmability is required for such organelle mimics to match the versatility, functionality and structural heterogeneity of their biological counterparts. Since fine-tuning the properties of naturally occurring MLOs *via* sequence engineering is challenging due to the sheer number of components involved [463],

de novo designed synthetic organelles relying on programmable nucleic acids are a highly promising approach, especially considering nucleic acids already play a role in biomolecular condensates.
The RNA nanostar platform introduced in Chapter 4 could, indeed, be ideal to build MLO-mimicking condensates in minimal artificial cell systems. However, in the previous Chapter, condensation was only achieved in bulk solution, differently from biologically occurring MLOs that emerge in confinement. Moreover, RNA condensates constructed in Chapter 4 are homogeneous and comprise a single phase, and thus fail to match the heterogeneity of biological MLOs.
In this Chapter, I expand on the RNA platform introduced in the previous Chapter to tackle these issues. On the one hand, I encapsulate the *in vitro* transcription mixtures producing the RNA nanostar building blocks within the simplest artificial cell mimics – water-in-oil droplets – resulting in the formation of synthetic RNA organelles. Simultaneous transcription of RNA nanostars with programmably orthogonal interaction domains leads to distinct, addressable and orthogonal organelles with tunable relative size and controllable number, paving the way for engineering even more veritably biomimetic compartments in artificial cell systems. On the other hand, I demonstrate how the simple introduction of a chimeric linker nanostar affords control over mixing/demixing behaviours of otherwise orthogonal condensates, and unlocks a whole spectrum of morphological architectures, capable of closely mimicking the complexity of biological MLOs.
The results and material presented in this chapter are included in the following Publication [434]:

G. Fabrini, S.P. Nuccio, J.M. Stewart, S. Li, A. Tang, P.W.K.R. Rothemund, E.Franco, M. Di Antonio and L. Di Michele, Co-transcriptional production of programmable RNA condensates and synthetic organelles, preprint on bioRxiv and soon to be submitted for peer review.

All the work presented in this Chapter has been carried out by me, apart from DNA template purification steps, which, as in Chapter 4, were performed by Sabrina Pia Nuccio (Di Antonio group).

Sequences for oligonucleotides used in this Chapter can be found in Tables A.2 (RNA motifs), A.3 (DNA templates), A.4 (DNA PCR primers) in the Appendix.

5.2 Synthetic RNA organelles in water-in-oil droplets

In this section, I will demonstrate that it is possible to yield single RNA condensates resembling primitive synthetic organelles within confined cell-like agents. In particular, this will

be achieved by encapsulating *in vitro* transcription reactions in the simplest artificial cell mimics, *i.e.* water-in-oil emulsion droplets. I will characterise the self-assembly and growth kinetics of such condensates by means of fluorescence microscopy and fluorimetry.

5.2.1 Emulsion droplets

Two of the most commonly used artificial cell mimics are water-in-oil emulsion droplets and Giant Unilamellar Vesicles (GUVs) [45, 11, 464–466, 40, 467, 468]. GUVs are clearly more "biomimetic" than emulsion droplets, thanks to their similarity to the cellular lipid membrane, but encapsulating biomolecules within them is considerably more challenging [469, 470]. Conversely, emulsion droplets provide a straightforward platform to encapsulate both charged and uncharged biomolecules in confined environments [21], provided biomolecule exchange among compartments is not required. Therefore, for the work described in this Chapter and the following one, I opted for water-in-oil emulsion droplets, stabilised by a fluorinated surfactant amidst a fluorinated oil phase, similarly to previous literature examples [21] (see Section 2.4.2). As exemplified in Fig. 5.1A, the *in vitro* transcription mixtures, comprising transcription buffer, T7 RNAP, ribonucleotides, and DNA template, are encapsulated by emulsification, resulting in transcription being carried out within confined water droplets, where RNA nanostars fold co-transcriptionally and self-assemble into organelle-mimicking condensates.

5.2.2 RNA nanostars form synthetic MLO-like condensates in emulsion droplets

Fig. 5.1B demonstrates, indeed, that the concept in Fig. 5.1A works: encapsulation of the *in vitro* transcription mixtures used in Chapter 4, supplied with a single DNA template coding for either RNA nanostar variant A **(i)**, B **(ii)** or C **(iii)**, yields single RNA condensates within the emulsion droplets.

Differently from bulk systems (Fig. 4.5), though, all RNA designs form spherical assemblies within the droplet-based synthetic cells, as depicted in confocal micrographs in Fig. 5.1C. This stems from the fact that the reduced size of emulsion droplets limits the length scales over which condensates have to relax. This effect is most noticeable for design C, which displays considerably different condensate morphologies, namely spherical-like in emulsion droplets *vs* percolating network-like in bulk. While its coarsening is still slower compared to A and B variants owing to the stronger KL C interactions (Fig. 4.6), the confined droplet environment dramatically reduces the length scale over which C phases have to relax (compared to entire bulk samples, *i.e.* capillary chambers). This in turn enables the emergence of the observed much more compact spherical-like condensates. These observations are supported by the time-evolution of μ_{CLD}, used again as proxy for condensate size, presented in Fig. 5.1D: μ_{CLD} first increases with time, and then plateaus for all RNA variants. It can be noticed that

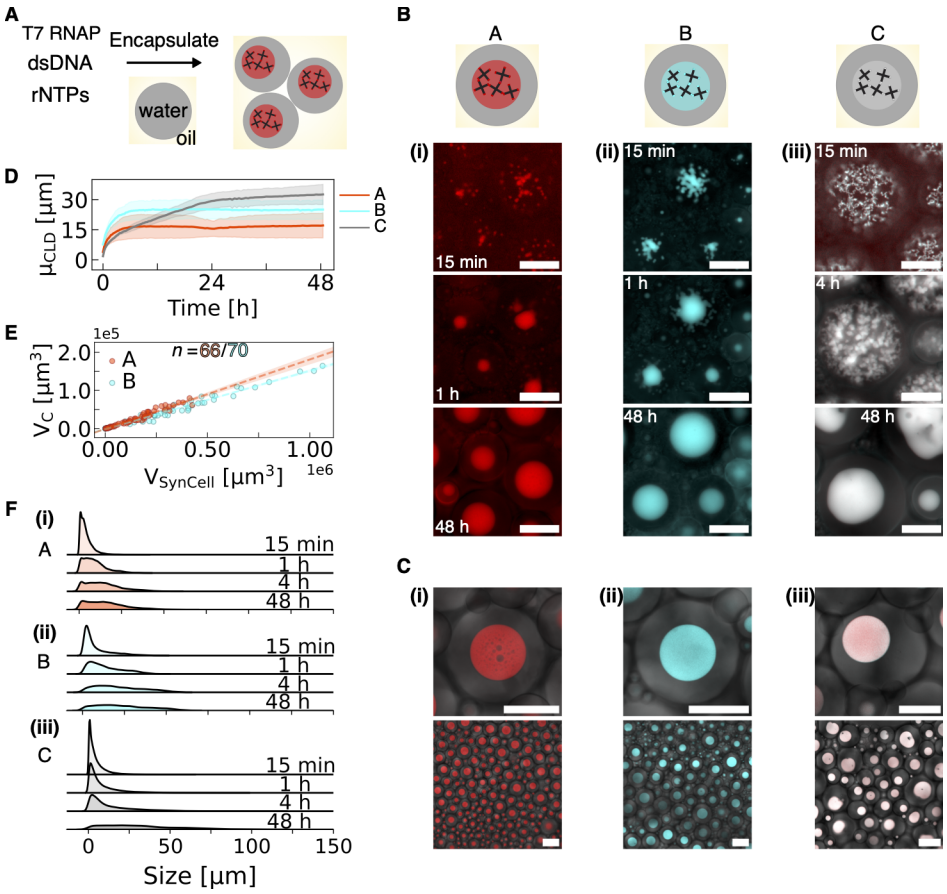


Fig. 5.1 **Self-assembly of RNA condensates within water-in-oil droplets. A**: Schematic of RNA condensate formation in water-in-oil droplets through encapsulation *via* emulsion method. **B**: Contrast-enhanced (see Section 2.11.5) epifluorescence micrographs of representative timepoints (15 min, 1 h or 4 h, 48 h, along the vertical axis) during the self-assembly of RNA condensates within water-in-oil droplets, from nanostars A **(i)**, B **(ii)** and C **(iii)**, respectively, accompanied by simplified schematics (top). **C:** (Top) zoomed-in and (bottom) larger field-of-view (FOV) confocal micrographs depicting samples in **B** after more than 48 h from start of transcription. **D:** Evolution of condensate mean size *vs* time examined *via* the mean of the chord length distribution, μ_{CLD}. Data are reported as mean (solid lines) ± standard deviation among mean replicates (shaded regions) from three non-overlapping FOVs within one sample. **E:** Scaling of condensate volume (V_{C}) with droplet-based synthetic cell volume (V_{SynCell}), as evaluated *via* linear regression (see Section 2.11.13). Number of examined condensate-droplet pairs are reported and colour-coded. Dashed lines are the lines of best fit, while shaded regions are 95% confidence intervals computed using 1000 bootstrap samples. **F:** Chord Length Distribution (CLD) ridge plots (see Section 2.11.10) at timepoints matching the ones in **A** for constructs A **(i)**, B **(ii)** and C **(iii)**, respectively. Confocal micrographs in **C** are pristine. Scale bars in **B** and **C**(top) are 50 μm. Scale bars in **C**(bottom) are 100 μm. All samples were obtained by *in vitro* transcription at 30 °C with [DNA template] = 40 nM.

size for A and B condensates plateaus much earlier than in bulk ($\approx$ 4 h *vs* 12 h), again due to the confined droplet environment, and earlier than C condensates. However, μ_{CLD} for C assemblies did not plateau at all in bulk, but does so in emulsion droplets for the discussed mechanism.

Fig. 5.1D also shows much larger variation in condensate size across all designs, as indicated by the larger standard deviation among μ_{CLD} replicates (shaded regions). This can be ascribed to the considerable variability in the size of droplets obtained *via* the emulsion

method, which in turn drives the scaling of condensate size with droplet size, as demonstrated for A and B condensates in Fig. 5.1E. Condensate volumes, in fact, scale linearly with droplet volumes, as indicated by the excellent agreement with the linear regression model (A: $R^2 = 0.98$, p-value = $2.2 \cdot 10^{-45}$; B: $R^2 = 0.99$, p-value = $2.7 \cdot 10^{-59}$; both p-values with respect to zero gradient null-hypothesis). In particular, A (red) and B (cyan) condensates reach a volume which is about 15-18% of the volume of the corresponding emulsion droplets (A: slope = 0.182 ± 0.005; B: slope = 0.153 ± 0.003; data are reported as mean ± standard error on the mean). We can try to rationalise such particular growth limit as follows. We can simplistically model each nanostar as a sphere with radius equal to the nanostar arm length $l_{\mathrm{ns}} = n_{\mathrm{bp}}\, d_{\mathrm{bp}}$, where n_{bp} is the number of base pairs in each arm, and $d_{\mathrm{bp}} = 0.28\,\mathrm{nm}$ [471] is the average distance among consecutive RNA base pairs. The number of base pairs in each arm, n_{bp}, was set to 28 bp to take into account the 25 bp-long double-stranded arm and half of the 6 bp-long coaxially stacked KLs upon interaction with neighbouring nanostars. In this simple model, each condensate of volume V_{C} will contain a number of nanostars, N_{ns}, equal to:

$$N_{\mathrm{ns}} = \frac{V_{\mathrm{C}}}{\frac{4}{3}\,\pi\, l_{\mathrm{ns}}^3}. \tag{5.1}$$

We can then evaluate the moles of ribonucleotide (rNTP) building blocks used up to build such condensate, $n_{\,\mathrm{rNTP,\,cond}}$, as:

$$n_{\,\mathrm{rNTP,\,cond}} = \frac{N_{\mathrm{ns}}\, N_{\mathrm{rNTP\,/\,ns}}}{N_{\mathrm{A}}} \tag{5.2}$$

where $N_{\mathrm{rNTP\,/\,ns}}$ is the length in nucleotides of the nanostar construct (A - 240 nt; B - 257 nt), and N_{A} is the Avogadro constant. This quantity can be compared with the initial rNTP reservoir in the corresponding droplet of volume V_{drop}, namely $n_{\,\mathrm{rNTP,\,drop}}$. Given the injected volume of each rNTP, $V_{\mathrm{rNTP}} = 1.8\,\mu\mathrm{L}$, its stock concentration, $C_{\mathrm{rNTP}} = 100\,\mathrm{mM}$, and the transcription reaction volume, $V_{\mathrm{reax}} = 22\,\mu\mathrm{L}$, this is simply equal to

$$n_{\,\mathrm{rNTP,\,drop}} = 4\frac{V_{\mathrm{rNTP}}\, C_{\mathrm{rNTP}}}{V_{\mathrm{reax}}} V_{\mathrm{drop}} \tag{5.3}$$

where the factor of 4 takes into account that ATP, CTP, GTP and UTP are all added stoichiometrically. When computing the ratio $n_{\,\mathrm{rNTP,\,cond}}/n_{\,\mathrm{rNTP,\,drop}}$ for all investigated A-type and B-type condensates in Fig. 5.1E, the results are as follow: A - 0.97 ± 0.20, B - 1.03 ± 0.21. While it is not possible to conclude that the organelle-mimicking RNA condensates stop growing because of rNTP depletion due to the intrinsic simplicity of the model, it is clear that both A-type and B-type condensates have incorporated a substantial fraction of the available rNTP building blocks.

The above-mentioned increase in condensate size variability can also be observed in the full CLDs presented in Fig. 5.1F, which confirm faster coarsening and overall broader distributions in all designs: looking at A and B, for instance, CLDs are steady from 4 h onwards and

much broader compared to the ones in Fig. 4.5D.
Consistently with bulk experiments in Fig. 4.5E, the programmability of Kissing Loop interactions holds in confined water-in-oil droplets as well, as demonstrated by Fig. 5.2A. Non sticky nanostar variants $\overline{A}$ and $\overline{B}$ fail to yield condensates, and rather remain dispersed within the emulsion droplets, creating a uniformly fluorescent RNA-rich lumen resembling a primitive cytoplasm.

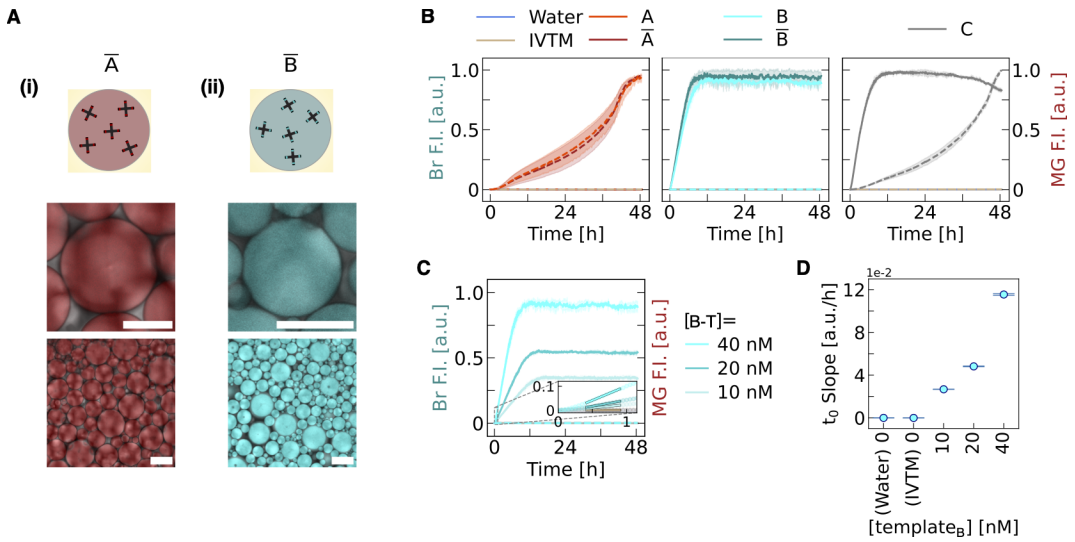


Fig. 5.2 **RNA-enriched cytoplasm from non sticky nanostars and fluorimetry characterisation of transcription in emulsion droplets. A:** Schematics (top) and confocal micrographs (middle: zoomed-in, bottom: larger FOV) depicting lack of condensate formation from non sticky design variants **(i)** $\overline{A}$ and **(ii)** $\overline{B}$, after more than 48 h from start of transcription. **B:** Fluorescence intensity (F.I.) kinetics monitoring the co-transcriptional folding and self-assembly of samples in **A** and Fig. 5.1B, namely A, $\overline{A}$ (left), B, $\overline{B}$ (middle), and C (right). Negative controls, *i.e.* dyes in water (Water) and dyes in *In Vitro* Transcription mixture (IVTM) with no added DNA template, are displayed in all panels. **C:** The initial transcription rate, examined through the proxy of initial F.I. increase rate, is proportional to the DNA template concentration (here [B-T] for RNA nanostar B). The zoomed inset depicts the 30 timepoints used for linear fitting highlighted by surrounding boxes. In **B** and **C**, the left/right y-axis depicts F.I. in DFHBI + BrA (Br F.I.)/MG + MGA (MG F.I.) channels in solid/dashed lines, respectively. Data in **B, C** are presented as normalised (see Section 2.11.4) mean (solid/dashed lines) ± standard deviation (shaded regions) from triplicates. **D:** Initial F.I. increase rate computed by linear fitting of the 30 timepoints boxed in **C** (inset). Data are presented as fitted slopes (circular markers) ± standard deviation from covariance matrix of the linear fit (error bars). Confocal micrographs in **A** are pristine. Scale bars in **A**(top) are 50 μm. Scale bars in **A**(bottom) are 100 μm. All samples were obtained by *in vitro* transcription at 30 °C with [DNA template] = 40 nM.

5.2.3 Comparison with fluorimetry measurements

Thanks to the use of fluorogenic aptamers, fluorimetry can once again be used to monitor the co-transcriptional folding of RNA nanostars, this time within droplet-based synthetic cell mimics, as showed in Fig. 5.2B. Similarly to bulk experiments, the fluorescence intensity of the corresponding FLAPs (MGA in A and $\overline{A}$, BrA in B and $\overline{B}$, both MGA and BrA in C) increases with time. Differently from bulk measurements, though, there is no evidence of the characteristic peak-decrease feature of sticky nanostar variants in bulk (see Fig. 4.4). This

agrees well with the previously introduced hypothesis of such feature reflecting fluorescent condensates forming and then sinking below the measurement plane, which indeed cannot happen in the case of condensates enclosed in confined water-in-oil droplets.
A second clear difference lies in the delayed double-stage growth exhibited by all MGA-containing nanostar designs (A, $\overline{\text{A}}$ and C, dashed profiles). This does not result from an actual delay in condensate growth, as Figs. 5.1C, D and F all prove growth kinetics are faster within emulsion droplets, and is rather due to initial preferential accumulation of MG in the oil outside of the droplets [472, 473]. The observed delay in fluorescence intensity increase in MG+MGA channel is thus due to the time needed for MG to diffuse from the oil phase into the water phase, where it can bind to the cognate aptamer contained in the transcribed RNA nanostars. This leads to a switch in the preferential enrichment of MG, from the naturally higher affinity oil phase to the otherwise disfavoured water phase, effectively demonstrating cargo uptake against phase affinity gradients thanks to the strong interactions among the dye and the cognate aptamer.
Similarly to bulk measurements, the initial rate of transcription, indirectly observed through the initial fluorescence intensity increase rate, strongly depends on DNA template concentration. This is demonstrated for RNA nanostar B in Fig. 5.2C-D: higher DNA template concentration promotes faster transcription, with initial transcription rates monotonically increasing with DNA template concentration (here [B-T]). It can be noted that the fitted initial transcription rates are approximately twice those measured in bulk (see Fig. 4.4C-D). However, as DNA templates are added prior to emulsification, *i.e.* not injected during the measurement run as in bulk, it is possible that the time window chosen for fitting is not equivalent to the one used for bulk measurements, despite showing a linear behaviour. Therefore, the observed increase in initial transcription rates is likely not indicative of a faster RNA synthesis.
Higher DNA template concentrations not only promote earlier plateauing, but also higher plateau levels. This is consistent with a scenario where RNA synthesis ceases when either the T7 RNAP loses catalytic activity or the rNTP building blocks are depleted. If rNTP depletion were to be the limiting factor, all curves would reach the same plateau, albeit at different times. As this is not the case, this points to a loss of T7 RNAP catalytic activity around 15 h and 17 h for 20 nM and 10 nM DNA template concentration, respectively. It is indeed plausible that lower DNA template concentration, resulting in lower RNA synthesis, grant a more prolonged enzymatic activity. At the highest examined DNA template concentration (40 nM), fluorescence intensity plateaus around 12 h. However, I previously noted that condensates seem to have incorporated approximately the same order of magnitude of rNTPs as the available reservoir. Therefore, it is possible that at higher DNA template concentration RNA synthesis approaches the limit of rNTP building block depletion, rather than being limited by loss of T7 RNAP activity. Due to not having measured even higher DNA template concentrations, though, this cannot be definitively proved and remains only an educated

guess.
These results, combined, hint at the possibility of controlling the overall RNA yield and the resulting condensate formation rate by simply tuning the corresponding DNA template concentration.

5.3 Multiple tunable RNA-based compartments in synthetic cell mimics

In this section, I will demonstrate the possibility of producing not just one, but multiple distinct self-assembled RNA condensates, interchangeably termed organelles from here onwards, within emulsion droplets. First, I will describe binary A/B systems with tunable relative size and robustly controlled organelle number in each droplet. I will then briefly discuss systems with three distinct types of orthogonal compartments produced by ternary mixtures of A, B and C nanostar motifs.

5.3.1 Binary RNA nanostar systems form two orthogonal phases (organelles / cytoplasm)

Binary sticky nanostar systems yield orthogonal RNA organelles with tunable size

Similarly to co-assembled orthogonal condensates in bulk (Fig. 4.7), simultaneous trancription of A-T and B-T DNA templates within emulsion droplets yields two orthogonal synthetic organelles thanks to the programmability of KL interactions, as demonstrated in Fig. 5.3A-B. As previously showed in bulk, the relative size of orthogonal A and B organelles within emulsion droplets can be tuned by simply varying the concentration ratio among the corresponding DNA templates, namely [A-T]/[B-T], as depicted in Fig. 5.3A-B and quantified in Fig. 5.3C. In particular, the confined environment of emulsion droplets allows to evaluate the ratio between the radii of A and B condensates (r_A/r_B) within the same droplet, which was previously not possible in bulk. When looking at droplet-based synthetic cells containing exactly one A and one B organelles, Fig. 5.3C shows that, as the [A-T]/[B-T] ratio increases, the r_A/r_B distribution shifts from low (large B organelles, small A ones) to progressively higher values (small B organelles, large A ones) following an almost linear trend. Varying the DNA template concentration ratio by a factor of 16 (from 0.25 to 4), the relative condensate size can be tuned by about 3-fold (27-fold volume-wise, looking at the mode of the size distributions).

Controlling the number of organelles in binary sticky nanostar systems

While most droplet-based synthetic cells contain exactly one organelle of each type in the long time limit, as showed in Fig. 5.3B(i), micrographs in Fig. 5.3B(ii) reveal that some

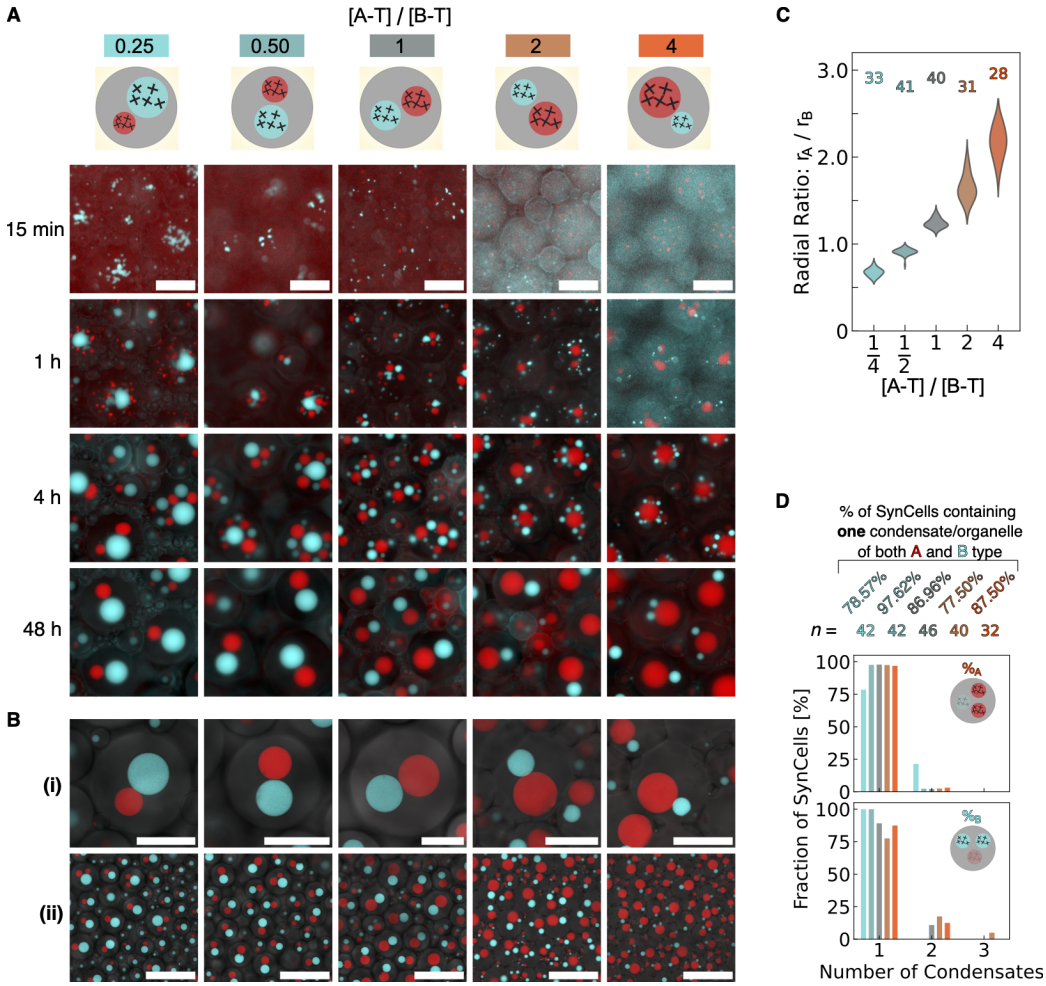


Fig. 5.3 **Co-assembly of orthogonal binary RNA organelles with controlled size and number in droplet-based synthetic cells. A:** Epifluorescence micrographs of representative timepoints (15 min, 1 h, 4 h and 48 h, along the vertical axis) during the self-assembly of binary (A, B) systems with varying [A-T]/[B-T] DNA template concentration ratios (from 0.25 to 4), as highlighted by schematic diagrams and colour-coding (top), which also applies to **C** and **D**. **B:** Zoomed-in **(i)** and larger FOV **(ii)** confocal micrographs depicting samples in **A** after more than 48 h. **C:** Distribution of A/B condensate size ratio (r_A/r_B) after 48 h with varying [A-T]/[B-T] for droplets containing exactly one A and one B condensates. **D:** Percentage of synthetic cells containing a certain number of A (top) and B (bottom) condensates after more than 48 h, regardless of the number of opposite-type condensates. Number of examined droplets per template ratio are reported on top in **C** and **D**. Further text on top of **D** reports the percentages of examined synthetic cells displaying exactly one organelle of each type (A/B) with varying [A-T]/[B-T] in the long time limit (>48 h). Micrographs in **A** are contrast-enhanced (see Section 2.11.5), and they all depict the same FOV area, thus sharing scale bars. Scale bars are 50 μm. Micrographs in **B** are pristine. Scale bars are 50 μm **(i)** and 150 μm **(ii)**. All samples were obtained by *in vitro* transcription at 30 °C with [total DNA template] = 40 nM.

emulsion droplets contain, indeed, different numbers of condensates, particularly in a 2:1 proportion or viceversa. Fig. 5.3D shows the percentage of emulsion droplets containing one, two or three condensate of each kind (A: top, B: bottom), respectively. Clearly, most droplets contain either a single A or a single B condensate, with higher number of either phases leading to lower fractions of observations. Merging the information about the number of A and B condensates, it is possible to determine the percentages of droplet-based synthetic cells comprising exactly one A and one B organelles after more than 48 h, namely 78.57%

([A-T]/[B-T] = 0.25), 97.62% ([A-T]/[B-T] = 0.50), 86.96% ([A-T]/[B-T] = 1), 77.50% ([A-T]/[B-T] = 2) and 87.50% ([A-T]/[B-T] = 4), respectively. This finding proves that we can not only tune the relative size of the orthogonal synthetic organelles within a single emulsion droplet, but also exert a high degree of control over their number, with more than 77% of droplet-based synthetic cells exhibiting exactly two RNA condensates (one of each kind) at all DNA template concentration ratios.

Binary systems with non sticky nanostars

Consistently with bulk observations, Fig. 5.4A-B (top: $A\&\overline{B}$, middle: $\overline{A}\&B$) prove that, if one of the DNA templates encodes for a non sticky nanostar variant, the droplet-based cell mimics will comprise a single RNA organelle, formed by the sticky nanostar species, surrounded by a dispersed phase of the non sticky one, similarly to a primitive nucleus in cytoplasm arrangement. If both templates produce non sticky nanostars upon transcription, then no synthetic organelles are formed altogether, as demonstrated by Fig. 5.4A-B (bottom: $\overline{A}\&\overline{B}$).

Fluorescence kinetics in binary systems

As with single components, fluorimetry measurements in Fig. 5.4C support what already observed in Figs. 5.3A-C and 5.4A-B. When normalising all profiles with respect to the stoichiometric system with both sticky nanostars ([A-T]/[B-T] = 1), it is evident that lower [A-T]/[B-T] ratios (left) lead to faster and stronger fluorescence intensity increase from DFHBI bound to BrA within B nanostars, whereas the same holds for MG bound to MGA within A nanostars at higher [A-T]/[B-T] ratios (right). Control stoichiometric mixtures where either or both nanostar species are made non sticky (middle panel) exhibit the same behaviour as mixtures of sticky nanostars with no clear trends in increase rate or endpoint intensity.

5.3.2 Increasing the number of compartments: ternary systems

Similarly to ternary systems in bulk, Fig. 5.5A shows that three orthogonal RNA nanostars can be simultaneously transcribed and co-assembled in spherical condensates of three distinct types within emulsion droplets. However, the stiffness of construct C (Fig. 5.5A (left)), combined with the increase in the number of non interacting RNA species, leads to much longer relaxation times, producing more numerous and smaller condensates within each droplet at the 48 h timepoint compared to Figs. 5.1B and 5.3A. Conversely, stoichiometric ternary mixtures where design C is replaced by $\tilde{C}$, identical to C but lacking fluorogenic aptamers, exhibit a higher degree of relaxation at the same 48 h timepoint, with fewer and larger A and B condensates, as $\tilde{C}$ (Fig. 5.5A (right)) cannot be detected in fluorescence channels. This observation suggests that the slower relaxation of construct C compared to

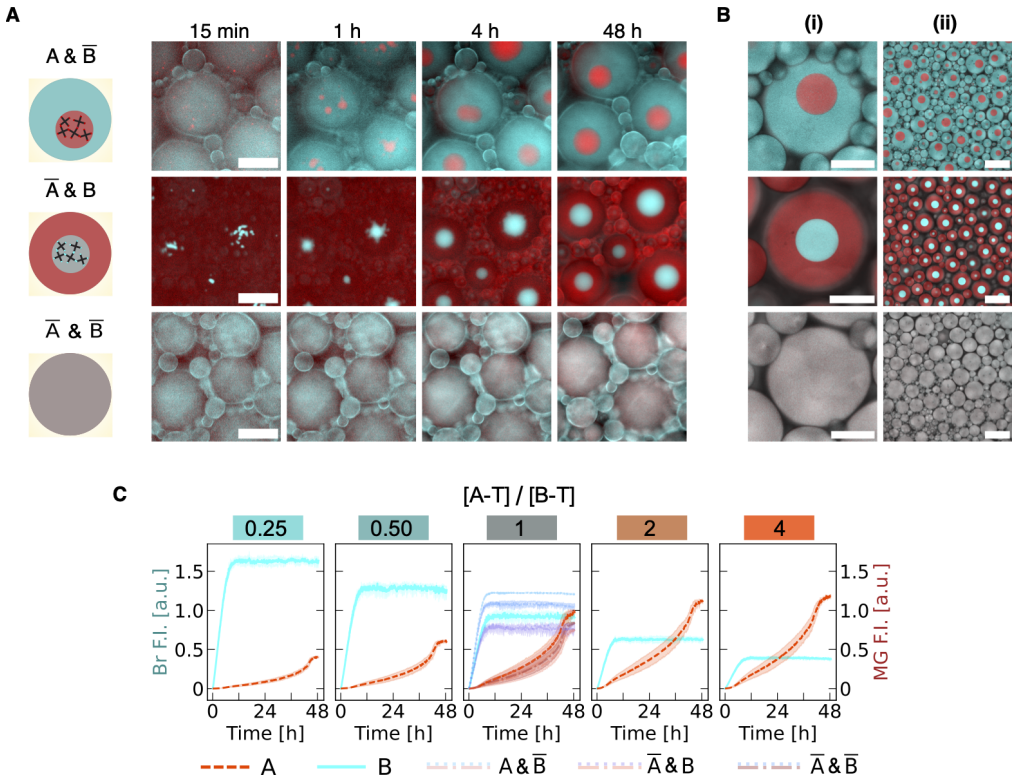


Fig. 5.4 **Co-assembly with non sticky RNA nanostar variants and fluorimetry characterisation of binary systems in emulsion droplets. A:** Epifluorescence micrographs of representative timepoints (15 min, 1 h, 4 h and 48 h, along the horizontal axis) during the self-assembly of stoichiometric binary systems where either or both A/B RNA designs are non sticky, as highlighted by the schematic diagrams on the left-hand side. **B:** Zoomed-in **(i)** and larger FOV **(ii)** confocal micrographs depicting samples in **A** after more than 48 h. **C:** Fluorescence intensity (F.I.) kinetics monitoring the co-transcriptional folding and self-assembly of binary systems in **A** and Fig. 5.3A. Left and right y-axes depict F.I. in DFHBI + BrA (Br F.I.) and MG + MGA (MG F.I.) channels in solid/dotted and dashed/dashed-dotted lines, respectively (the latter options referring to non sticky control mixtures, middle panel). Data in **G** are presented as mean (solid/dotted and dashed/dashed-dotted lines) ± standard deviation (shaded regions) from triplicates, normalised with respect to [A-T]/[B-T] = 1 curve (see Section 2.11.4). Micrographs in **A** are contrast-enhanced (see Section 2.11.5), and they all depict the same FOV area, thus sharing scale bars. Scale bars are 50 μm. Micrographs in **B** are pristine. Scale bars are 50 μm **(i)** and 150 μm **(ii)**. All samples were obtained by *in vitro* transcription at 30 °C with [total DNA template] = 40 nM.

A and B, already observed both in bulk and in emulsion droplets, is not only related to the stronger, more thermally stable KL C interactions (Fig. 4.6). Other contributing factors are likely the added co-transcriptional folding complexity brought about by hosting two RNA aptamers compared to one, as well as the resulting structure of the nanostar arms, which likely influences the viscosity of the resulting C-type assemblies. Fluorimetry measurements in Fig. 5.5B support the formation of nanostars containing MGA and/or BrA, similarly to what observed in bulk (Fig. 4.8C).

In both described ternary systems, Fig. 5.5A shows that 48 h is not enough to observe even a single occurrence of the ideal scenario presented in Fig. 5.5C, *i.e.* exactly three RNA organelles (one of each type) within an emulsion droplet. Nonetheless, Fig. 5.5D-E demonstrate that, given enough time (2 weeks), the orthogonal RNA organelles relax more

and more, with some emulsion droplets displaying, as desired, exactly three RNA-based compartments, one of each kind, as highlighted in Fig. 5.5D-E (iii). Despite the very low fraction of emulsion droplets matching the target scenario presented in Fig. 5.5C, all the examined droplets present three distinct types of individually addressable compartments: A enriched in MG - red; B in DFHBI - cyan; C in both dyes - gray / $\tilde{\text{C}}$ not enriched - non fluorescent. This demonstrates the promising capabilities of this platform to produce multiple addressable RNA-based synthetic organelles within artificial cells, each representing a distinct microenvironment with the potential of promoting different biomolecular reactions or pathways.

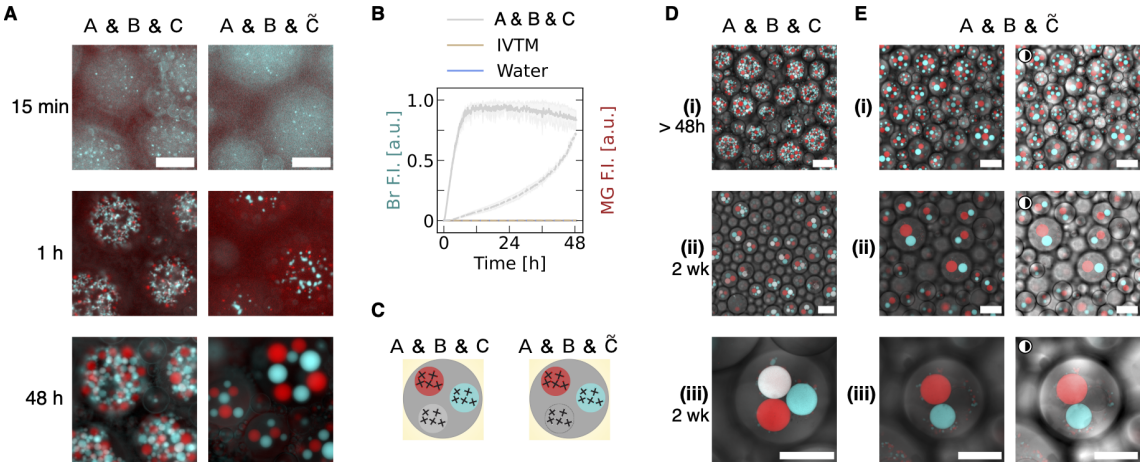


Fig. 5.5 **Co-assembly of three orthogonal RNA condensates in droplet-based synthetic cells. A:** Contrast-enhanced (see Section 2.11.5) epifluorescence micrographs of representative timepoints (15 min, 1 h, 48 h, along the vertical axis) depicting the simultaneous transcription and condensate co-assembly from ternary stoichiometric mixtures of A-T, B-T and C-T (left) or $\tilde{\text{C}}$-T (right). $\tilde{\text{C}}$-T codes for construct $\tilde{\text{C}}$, identical to C (including KL C domains) but lacking fluorogenic aptamers. Micrographs represent the same FOV area, thus sharing the scale bar with the first micrograph in each column. Scale bars are 50 μm. **B:** Fluorescence intensity (F.I.) kinetics monitoring co-transcriptional folding and self-assembly of ternary system with all fluorogenic constructs. Left/right y-axis depicts fluorescence intensity in DFHBI + BrA (Br F.I.)/MG + MGA (MG F.I.) channels in solid/dashed lines, respectively. Negative controls include dyes in water (Water) and dyes in *In Vitro* Transcription mixture (IVTM) with no added DNA template. Data are presented as normalised (see Section 2.11.4) mean (solid/dashed lines) ± standard deviation (shaded regions) from triplicates. **C:** Schematic representation of the two ternary systems, with $\tilde{\text{C}}$ (right) forming a non fluorescent phase indistinguishable from background (but visible in bright-field), in the ideal case of full relaxation of all three components into single organelles. **D:** Pristine confocal micrographs of ternary system with all fluorogenic constructs **(i)** after more than 48 h and after 2 weeks ("2 wk"), both **(ii)** large FOV and **(iii)** zoomed-in on an individual droplet. **E:** Same as **D** for mixture with non fluorogenic third component $\tilde{\text{C}}$, and additional contrast-enhanced (bright-field only, highlighted by semi-shaded circle) micrographs on the right-hand side for ease of visualisation of the third non fluorescent assembly. Scale bars in **D (i-ii), E (i-ii)** are 100 μm. Scale bars in **D (iii), E (iii)** are 50 μm. All samples were obtained by *in vitro* transcription at 30 °C, with [A-T] = [B-T] = [C-T or $\tilde{\text{C}}$-T] = 13.33 nM.

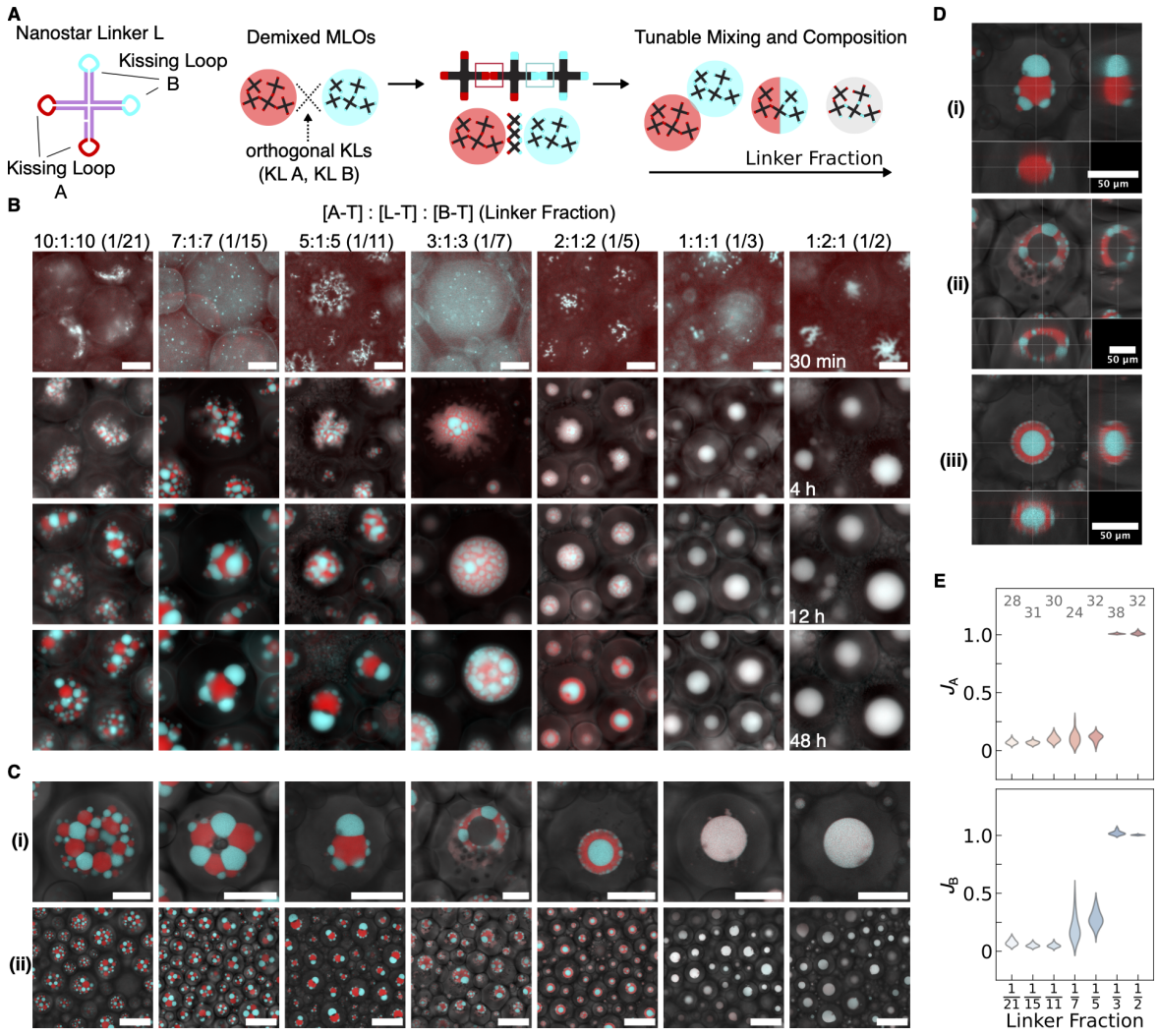


Fig. 5.6 **Control over mixing/demixing behaviours *via* linker nanostars unlocks complex organelle morphologies. A:** Diagram illustrating control over composition and mixing in A/B systems *via* chimeric RNA linker nanostars (L). Linker Fraction (L.F.) is defined as [L-T]/([A-T]+[L-T]+[B-T]). For all results in this Figure, [A-T] = [B-T]. **B:** Contrast-enhanced (see Section 2.11.5) epifluorescence micrographs of representative timepoints (30 min, 4 h, 12 h and 48 h, as reported in right-most column along the vertical axis) depicting the self-assembly within emulsion droplets of binary A/B systems with varying [A-T]:[L-T]:[B-T] DNA template proportions (horizontal axis, reported on top, with L.F. values in parentheses). All micrographs depict the same FOV area, thus sharing scale bars. **C: (i)** Zoomed-in and **(ii)** larger FOV confocal micrographs of samples in **B** after more than 48 h. **D:** XY (top left), XZ (top right), YZ (bottom left) 2D cross-sections from confocal z-stacks outlining the internal microstructure of synthetic organelles assembled within emulsion droplets with [A-T]:[L-T]:[B-T] = **(i)** 5:1:5 (L.F. = 1/11), **(ii)** 3:1:3 (L.F. = 1/7), and **(iii)** 2:1:2 (L.F. = 1/5), respectively. **E:** Distribution of mixing indices J_A and J_B *vs* Linker Fraction for samples in **B-C**. Numbers on top indicate examined synthetic cells per Linker Fraction. Data points exceeding the first and third quartiles by more than 1.5 times the interquartile range have been discarded as outliers (see Section 2.11.16). Scale bars in **B**, **C (i)**, **D** are 50 μm. Scale bars in **C (ii)** are 150 μm. All confocal micrographs in **C, D** are pristine. All samples were obtained by *in vitro* transcription at 30 °C, with [total DNA template] = 40 nM.

5.4 Controlling mixing/demixing to unlock complex organelle morphologies

As mentioned above, most naturally occurring biomolecular condensates and MLOs are not homogeneous, rather they display complex hierarchical morphologies, from the segregated core-shell architecture of paraspeckles and PML bodies, to the asymmetrical ball-socket arrangement of Cajal bodies and the multi-layered Russian doll [474] structure of nucleoli [28]. The platform presented so far only produces simple RNA condensates, each comprising a single, homogeneous phase. In this section, I expand the range of achievable architectures and morphologies, thus more closely mimicking the complexity of biological MLOs, by means of a linker (L) nanostar, briefly introduced in Section 4.2.1 and presented again in Fig. 5.6A for ease of access. This RNA nanostar exhibits a chimeric design, in that it features two KL A and two KL B domains, which enable it to act as a bridge or linker, hence the name, among *a priori* orthogonal A and B nanostars and their condensate assemblies.

5.4.1 Tuning RNA organelle morphology *via* linker template fraction

In view of the design of L, simultaneous transcription of the DNA templates coding for nanostars A, B and L should yield assemblies of A and B condensates with increasing degrees of adhesion and mixing, as the Linker Fraction (L.F.), defined as [L-T]/([A-T] + [B-T] + [L-T]), increases. Fig. 5.6B-C demonstrates that this is, indeed, the case. At low fractions of linker template ([A-T]:[L-T]:[B-T]=10:1:10, or equivalently L.F. = 1/21), A and B phases adhere to one another forming grape-like clusters. The inter-phase adhesion, in fact, likely limits sliding among the two phases, therefore preventing further coalescence and relaxation into two large opposite-type condensates adhering in the middle. Slightly higher linker fractions promote even more pronounced and extended inter-phase adhesion ([A-T]:[L-T]:[B-T]=7:1:7, L.F. = 1/15), ultimately leading to Janus-like – or Squirtle-like, for Pokémon fans – structures ([A-T]:[L-T]:[B-T]=5:1:5, L.F. = 1/11) with fuller relaxation and bigger orthogonal domains, as showed by the cross-sections in Fig. 5.6D(i).
Further increasing the linker template fraction leads to the emergence of multi-phase morphologies, such as hollow microcapsules ([A-T]:[L-T]:[B-T]=3:1:3, L.F. = 1/7) and core-shell/Russian-doll structures ([A-T]:[L-T]:[B-T]=2:1:2, L.F. = 1/5), as depicted in Fig. 5.6D(ii) and (iii), respectively. Notably, both conditions exhibit preferential segregation of the A-rich phase on the outside and the B-rich phase on the inside. This can be easily rationalised by taking into account that A nanostars exhibit a lower melting temperature of KL interactions compared to B nanostars, namely $T_{M,A} < T_{M,B}$. It follows that KL A interactions are weaker than KL B ones, *i.e.* A is more fluid. As the strength of interactions underlying LLPS influences the biophysical properties of the resulting phases [201], A will likely display a lower interfacial tension with respect to the bulk solution compared to B.

Such a scenario is in excellent agreement with previous studies on nucleoli, which have found that their hierarchical structure emerges due to the presence of two immiscible components, the more fluid of which encapsulates the stiffer second phase due to differences in surface tension [463]. The same has been found to hold in other biomolecular condensates [36] and multi-phase coacervate droplets [475].
For such morphologies to occur in our particular case, separate A and B condensates must exhibit a higher overall surface tension compared to B encapsulated in a larger A-rich phase surrounded by bulk solution, as described in Equation 5.4 below, derived from ref. [475]:

$$4\pi(R_{\mathrm{A}}^3+R_{\mathrm{B}}^3)^{\frac{2}{3}}\gamma_{\mathrm{A}}+4\pi R_{\mathrm{B}}^2\gamma_{\mathrm{AB}}<4\pi \mathrm{R}_{\mathrm{A}}^2\gamma_{\mathrm{A}}+4\pi \mathrm{R}_{\mathrm{B}}^2\gamma_{\mathrm{B}} \tag{5.4}$$

where R_{A}, R_{B}, γ_{A} and γ_{B} are the radii and surface tensions of A and B phases, and γ_{AB} is the A-B interfacial tension, respectively. Equation 5.4 produces the following condition:

$$\gamma_{\mathrm{AB}}<\gamma_{\mathrm{B}}-\alpha\gamma_{\mathrm{A}} \tag{5.5}$$

where $\alpha=(1+(\frac{R_{\mathrm{A}}}{R_{\mathrm{B}}})^3)^{\frac{2}{3}}-(\frac{R_{\mathrm{A}}}{R_{\mathrm{B}}})^2$ is a geometrical factor taking into account the increase in surface area of the A-rich phase upon encapsulating the B-rich one. Looking at the morphologies in Fig. 5.6C, it is clear that Equation 5.5 only holds when linker fractions are high enough to sufficiently lower the AB interfacial tension γ_{AB}, which would be very high in the absence of linker nanostars due to the orthogonality between KL A and KL B. Finally, raising the linker fraction to stoichiometric template concentration conditions ([A-T]:[L-T]:[B-T]=1:1:1, L.F. = 1/3) or even beyond ([A-T]:[L-T]:[B-T]=1:2:1, L.F. = 1/2) lowers the interfacial tension so much that it produces fully mixed single-phase spherical condensates.

5.4.2 Quantitatively evaluating mixing with varying Linker Fraction

The degree of mixing between A and B phases with varying Linker Fraction can be quantitatively evaluated by image analysis, as showed in Fig. 5.6E. The two depicted indices, J_{A} and J_{B}, one for each phase, measure the overlap between the fluorescent channels corresponding to phases A (MG+MGA) and B (DFHBI+BrA). In particular, they are computed as the ratios between the average intensity of the minority RNA component and the average intensity of the majority phase, exploiting binary segmentation masks (see Section 2.11.16). For instance, $J_{\mathrm{A}}=I_{\mathrm{A}}^{\mathrm{B}}/I_{\mathrm{A}}^{\mathrm{A}}$ where $I_{\mathrm{A}}^{\mathrm{B}}$ and $I_{\mathrm{A}}^{\mathrm{A}}$ are the mean intensity of A in B-rich phase and A in A-rich phase, respectively. As expected from micrographs in Fig. 5.6B-C, both mixing indices exhibit a sharp sigmoidal behaviour, with very low mixing (mostly adhesion) at low linker fractions, moderately increasing at progressively higher relative amounts of linker template, before abruptly jumping to full mixing at stoichiometric template conditions and beyond. It would be interesting to investigate whether such behaviour can be reproduced using classical mean-field models for mixing/demixing derived from polymer physics [476, 477], such as

the Flory-Huggins regular solution theory. For instance, one could model this system either as a quaternary mixture (A, B, L, solvent), or as a simpler ternary one (A, B, solvent) where the interaction parameter between the A and B phases, χ_{AB}, becomes more energetically favourable with increasing Linker Fraction.
Combined with the morphological observations above, these results effectively demonstrate the possibility of programming mixing/demixing behaviours and thus accessing a wide morphological spectrum in our model system. Not only does this dramatically increase the biomimicry of the produced condensate architectures, but it does so in a starkingly simple way, by introducing varying amounts of a chimeric linker capable of bridging otherwise orthogonal nanostars, taking inspiration from previous DNA nanotechnology approaches [199, 478]. Conversely, to achieve a similar result in biological MLOs, one would have to painstakingly vary the sequences of all involved biopolymer subunits one by one to try and alter the material properties they encode for. Once again, such simplicity and programmability further substantiate the potential of the discussed RNA platform for minimal synthetic cell systems.

5.5 Chapter summary and outlook

In this chapter, I have demonstrated how the RNA nanostar platform introduced in Chapter 4 can be encapsulated in water-in-oil emulsion droplets to yield addressable synthetic organelles capable of programmable mixing and demixing. Such an extension is incredibly straightforward and does not require any modification of the platform or protocols used in bulk, simply an additional encapsulation step *via* vortexing.
As observed in the previous chapter, nanostars can be designed to interact *via* palindromic KLs, thus forming organelle-like condensates through LLPS, or programmatically prevented from doing so by making their loops non-complementary, yielding a dispersed RNA-rich lumen mimicking a cytoplasm within the artificial cell mimics. Compared to bulk results, though, the confined environment of emulsion droplets accelerates coarsening by reducing relaxation length scales, thus leading to more compact and spherical-like assemblies across all tested nanostar motifs. Interestingly, the volumes of these synthetic organelles scale with those of the emulsion droplets containing them, similarly to how biomolecular condensates and MLOs scale in biological cells [354, 479].
Thanks to the programmability of KL interactions, multiple RNA nanostar populations can be simultaneously transcribed in each emulsion droplet, leading to the co-assembly of orthogonal, individually addressable organelles. The relative size of opposite-type organelles in binary systems can be easily tuned by varying the concentration ratio among the corresponding DNA templates. Despite such composition variations, this platform affords robust control over the number of synthetic organelles in each droplet-based synthetic cell. The number of distinct compartments can be further increased to three, despite the considerable loss in precise number control capabilities due to the much longer relaxation timescales.

Finally, the introduction of a chimeric linker nanostar capable of bridging orthogonal RNA motifs enables control over mixing/demixing behaviours of the resulting condensates by simply varying the fraction of the corresponding DNA template. In particular, varying the relative amounts of transcribed linkers tunes the interfacial tension between otherwise immiscible phases, and unlocks a wide range of morphologies, far more complex than simple homogeneous droplets, tremendously boosting the biomimicry of this approach.
In Chapter 6, I demonstrate how these synthetic RNA organelles can be made functional, enabling them to capture target proteins and thus act as crucibles/sponges, further bridging the gap towards the properties and functionality of MLOs in biological cells.

Chapter 6

Towards functional RNA-based organelles in artificial cells

6.1 Introduction

Biomolecular condensates and membrane-less organelles (MLOs) have been defined as the organisers of cellular biochemistry [30] due to their fundamental ability to concentrate a variety of macromolecules, including proteins and nucleic acids. MLOs can act as hubs, bringing together a variety of biomolecules to direct metabolic flux, or genomic loci and transcription factors to coordinately regulate the expression of several genes. Similarly, they can also function as crucibles, concentrating enzymes and substrates to boost the rate of biochemical reactions. Finally, they can behave as sponges, suppressing the activity of the sequestered biomolecules or slowing down related pathways in the surrounding solution [351, 28].
The synthetic cells we strive to develop would clearly benefit from a similar level of internal organisation of their biochemical milieu, which would make them more biomimetic and more functionally versatile. Specifically, it would be beneficial to drive accumulation of proteins, which often carry biological functionality.
In this Chapter, I demonstrate how simple design modifications of the RNA nanostars can grant the resulting RNA organelles explored in Chapter 5 the capability of capturing target proteins. In particular, I investigate two capture mechanisms, one based on simultaneously transcribed RNA nanostars and protein-binding RNA aptamers, and another relying on high affinity receptor-ligand-like interactions, where the protein (receptor) is bound *via* short ligand-conjugated oligonucleotides. Interestingly, it is possible to obtain both homogeneous and inhomogeneous protein enrichment profiles within the RNA organelles depending on the chosen capture mechanism. Both strategies rely on RNA nanostars and protein-binding constructs interacting through sticky ends, as often done in nucleic acid nanotechnology, further highlighting the ease of integration of this RNA platform with current methods.

Part of the results and material presented in this chapter are included in the following Publication [434]:

G. Fabrini, S.P. Nuccio, J.M. Stewart, S. Li, A. Tang, P.W.K.R. Rothemund, E.Franco, M. Di Antonio and L. Di Michele, Co-transcriptional production of programmable RNA condensates and synthetic organelles, preprint on bioRxiv and soon to be submitted for peer review.

All the work presented in this Chapter has been carried out by me, apart from DNA template purification steps, which, as in Chapters 4 and 5, were performed by Sabrina Pia Nuccio.

Sequences for oligonucleotides used in this Chapter can be found in Tables A.5 (RNA motifs), A.6 (DNA templates), A.7 (DNA PCR primers) in the Appendix.

6.2 Design and characterisation of protein-capturing RNA constructs

In this Section, I will first discuss the design of RNA nanostars and aptamers used to obtain protein-capturing RNA organelles. I will then proceed to characterise both the DNA templates encoding for such RNA motifs, and the RNA products obtained from their *in vitro* transcription.

6.2.1 Nanostructure design: adding sticky ends

In order to sequester target proteins within the RNA organelles discussed in Chapter 5, a protein-binding construct is needed. The simplest approach is to exploit protein-binding RNA aptamers. These could be either hosted within the (single) RNA strand forming the nanostar, as previously done for MGA and BrA, or transcribed separately and then connected to the nanostars through base-pairing. As for the first option, most protein-binding RNA aptamers are considerably larger – *i.e.* longer in sequence – and more complex than MGA and BrA, thus making it challenging to simply integrate them in nanostars without considerably altering their topology and size. Therefore, I opted to add sticky ends to both RNA nanostar designs A and B introduced in Chapter 4.
As mentioned in Section 1.4.4, T7 RNA polymerase tends to spuriously extend transcripts lacking a strong secondary structure on the 3' end. I thus positioned the single-stranded overhangs at the 5' end of the transcripts instead, right after the T7 promoter region of the corresponding DNA template. Each overhang comprises a 12 nucleotide (nt)-long binding domain as well as a three adenine-long spacer separating the sticky end from the start of the RNA nanostar sequence, granting flexibility. The resulting modified RNA nanostars

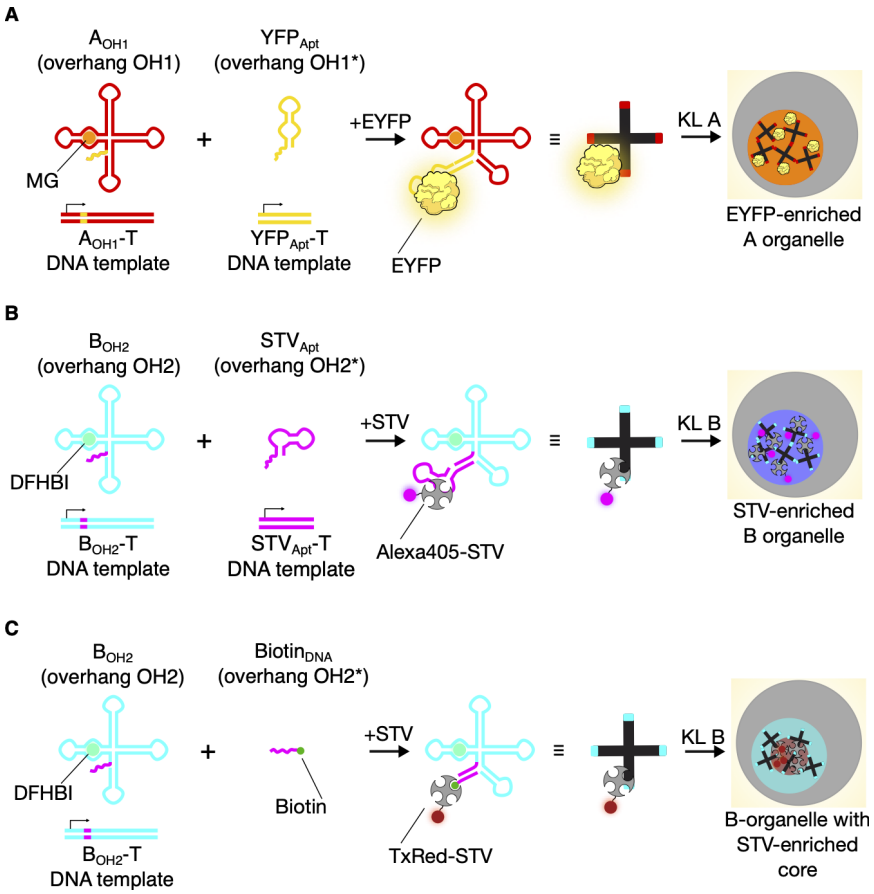


Fig. 6.1 **Design of co-transcriptionally folding RNA nanostars for target protein capture.** All RNA constructs are transcribed from fully double-stranded DNA templates, indicated with the construct name followed by '-T', following the notation introduced in Chapter 4. **A**: RNA nanostar A_{OH1} is identical to A with the addition of a 5' 15 nt-long sticky end (12 nt capture overhang OH1 and three adenine-long spacer). YFP_{Apt} is a CFP/GFP/YFP aptamer [235] extended at the 5' *via* a 16 nt-long sticky end comprising the complementary domain OH1*, preceded by an additional guanine for efficient transcription initiation, and followed by a three adenine-long spacer. Nanostar A_{OH1} is designed to bind YFP_{Apt} through the OH1-OH1* complementary overhang domains upon simultaneous co-transcriptional folding, thus enabling preferential enrichment of EYFP within the A_{OH1}-type RNA organelles. **B**: RNA nanostar B_{OH2} is identical to B with the addition of a 5' 15 nt-long sticky end (12 nt capture overhang OH2 and three adenine-long spacer). STV_{Apt} is a STV aptamer [236] extended at the 5' *via* a 15 nt-long sticky end comprising the complementary domain OH2* followed by a three uracil-long spacer. Nanostar B_{OH2} is designed to bind STV_{Apt} through the OH2-OH2* complementary overhang domains upon simultaneous co-transcriptional folding, thus enabling preferential enrichment of (Alexa405-tagged) STV within the B_{OH2}-type RNA organelles. **C**: STV can also be sequestered by exploiting high-affinity STV-biotin interactions through a biotin-conjugated single-stranded DNA oligonucleotide, $Biotin_{DNA}$, whose sequence is the DNA equivalent of the 15 nt overhang used in STV_{Apt} followed by a 3' biotin modification. Addition of $Biotin_{DNA}$ and (TexasRed-conjugated) STV moieties before co-transcriptional folding of B_{OH2} nanostars, while working at low protein concentration, is expected to produce B-type organelles with an enriched core/depleted shell architecture. Drawings in **A-C** are not to scale. EYFP graphical element in **A** was created with BioRender.com.

A_{OH1} (yellow sticky end, OH1 binding domain) and B_{OH2} (magenta sticky end, OH2 binding domain) are presented in Fig. 6.1A-B, respectively.

I chose to design two orthogonal binding domains (OH1, OH2) rather than one to grant A_{OH1} and B_{OH2} organelles the ability to sequester distinct target proteins, possibly at the same time within the same droplet-based synthetic cell. To simplify the detection of the protein

cargoes, both individually and simultaneously, I opted for Enhanced Yellow Fluorescent Protein (EYFP) and fluorophore-conjugated (either TexasRed or Alexa405) Streptavidin (STV) as targets. Not only are both proteins well-suited for imaging, but they have minimal fluorescence emission overlap, and cognate RNA aptamers have already been described in the literature (CFP/GFP/YFP aptamer [235], STV aptamer [236]). I thus modified the minimal aptamer sequences by prepending sticky ends complementary to those in A_{OH1} ($OH1^*$) and B_{OH2} ($OH2^*$), followed again by short 3 nt-long spacers, resulting in EYFP-binding YFP_{Apt} and STV-binding STV_{Apt} constructs, respectively.

In the proposed mechanism, simultaneous transcription of A_{OH1}-YFP_{Apt} and B_{OH2}-STV_{Apt} pairs in the presence of target proteins EYFP and STV results in both protein-aptamer and aptamer-nanostar binding, which yield EYFP-rich A_{OH1}-type and STV-rich B_{OH2}-type organelles, as schematically depicted in Fig. 6.1A-B, respectively.

Exploiting the sticky ends attached to B_{OH2} nanostars, I decided to investigate a second STV capture mechanism based on the high affinity interactions among biotin and streptavidin, as presented in Fig. 6.1C. In particular, I designed a short single-stranded DNA oligonucleotide featuring the same sequence as the overhang in STV_{Apt} and terminating in a 3' biotin-modification, termed $Biotin_{DNA}$. Notably, the same would hold for a biotin-conjugated RNA oligonucleotide of equivalent sequence – DNA was chosen simply due to its cost-effectiveness. In this second capture strategy, $Biotin_{DNA}$ is added to the transcription mixture prior to encapsulation in emulsion droplets, at the same time as STV. This effectively results in $Biotin_{DNA}$-STV complexes being formed prior to any significant synthesis of RNA nanostars. Due to the tetrameric nature of STV and the biotin excess conditions used in this work ($\approx$ 8-fold $Biotin_{DNA}$ excess), such complexes will likely be $(Biotin_{DNA})_4$-STV. Consequently, the first transcribed B_{OH2} nanostars will each bind $\frac{1}{4}$ of a STV moiety, *i.e.* each STV will link four RNA nanostars hybridised to the four biotinylated oligonucleotides it is complexed with. Such STV-mediated cross-linking will most likely lead to an increase in affinity between the nanostars, beyond what regularly afforded by KLs. Conversely, B_{OH2} nanostars transcribed at later times will lack cargo due to the low STV concentration in use, adopted to limit the steric hindrance stemming from the bulky protein cargo. Therefore, this alternative mechanism is expected to produce RNA organelles exhibiting inhomogeneous distribution of the target protein, with a STV-enriched core and a depleted shell, as schematically depicted in Fig. 6.1C.

To better distinguish the two approaches involving B_{OH2} nanostars in Fig. 6.1B-C, STV molecules with different conjugated fluorophores were used, namely Alexa405-tagged STV for the former (aptamer-based, Fig. 6.1B) and TexasRed-conjugated STV (TxRed-STV) for the latter (biotin-streptavidin interactions, Fig. 6.1C).

DNA templates and primers for PCR amplification of constructs A_{OH1}, B_{OH2} and YFP_{Apt} (A_{OH1}-T, B_{OH2}-T and YFP_{Apt}-T, respectively), were designed as discussed in Section 4.2.2. Due to the short sequence of STV_{Apt}, the corresponding template (STV_{Apt}-T) was pur-

chased as a different product, namely annealed duplex Ultramers, that did not require PCR amplification (see Section 2.2).

6.2.2 Characterisation

Characterisation of DNA templates

DNA templates A_{OH1}-T, B_{OH2}-T and YFP_{Apt}-T were PCR amplified, and their electrophoretic mobilities were characterised *via* agarose gel electrophoresis, together with those of the single-stranded biotin-conjugated oligonucleotide, $Biotin_{DNA}$ and the template for STV_{Apt} (STV_{Apt}-T). The gel in Fig. 6.2A (i-ii) shows that all DNA constructs display mobilities in excellent agreement with the expected sizes (reported on top of each lane) after 45 min of electrophoresis, as highlighted by the closest reference markings from the Ultra Low Range (ULR) DNA ladder in the left-most lane. The difference in band intensity between the first three and the last two lanes is simply due to the former samples having undergone PCR amplification, and thus presenting much higher concentrations than the latter ones.
In order to confirm the size of the templates for the two modified RNA nanostars and inspect the potential presence of spurious amplicons, electrophoresis was allowed to proceed for additional 45 min. The resulting gel in Fig. 6.2B (i-ii) shows that, as expected, the size of the modified nanostar templates is slightly below (A_{OH1}-T, 275 bp) or around (B_{OH2}-T, 292 bp) 300 bp, respectively. Notably, due to its high mobility, the short single-stranded $Biotin_{DNA}$ construct is not detectable in the gel in Fig. 6.2B (i) due to having overrun the available gel length.
Consistently with the procedure adopted in Chapters 4 and 5, all PCR amplified templates (A_{OH1}-T, B_{OH2}-T, YFP_{Apt}-T) were agarose gel purified, *i.e.* the main bands were cut, dissolved and eluted, as described in Section 2.4.

Gel electrophoresis on RNA transcripts

As for designs presented in Chapter 4, DNA templates for protein-capturing motifs were *in vitro* transcribed for 18 h at 30 °C ([DNA template] = 40 nM), and the resulting transcription products were characterised *via* native and denaturing agarose gel electrophoresis, as reported in Fig. 6.3A (i-ii) and B (i-ii), respectively.
Fig. 6.3A (i-ii) show that, in native conditions, both modified RNA nanostars A_{OH1} and B_{OH2} display several bands. These include the expected band corresponding to monomeric RNA nanostars located between 100 and 150 bp of the ULR reference ladder. This migration distance is consistent with the one noted in Chapter 4 (Fig. 4.3A (i)) for A and B nanostars, as single-stranded overhangs are not expected to considerably alter the mobility of the much larger nanostars. However, the monomeric band is not the dominant one this time. This likely reflects some misfolding events of transcribed nanostars induced by the addition of the 5' single-stranded overhangs: as these sticky ends are the first sequence stretch to be transcribed,

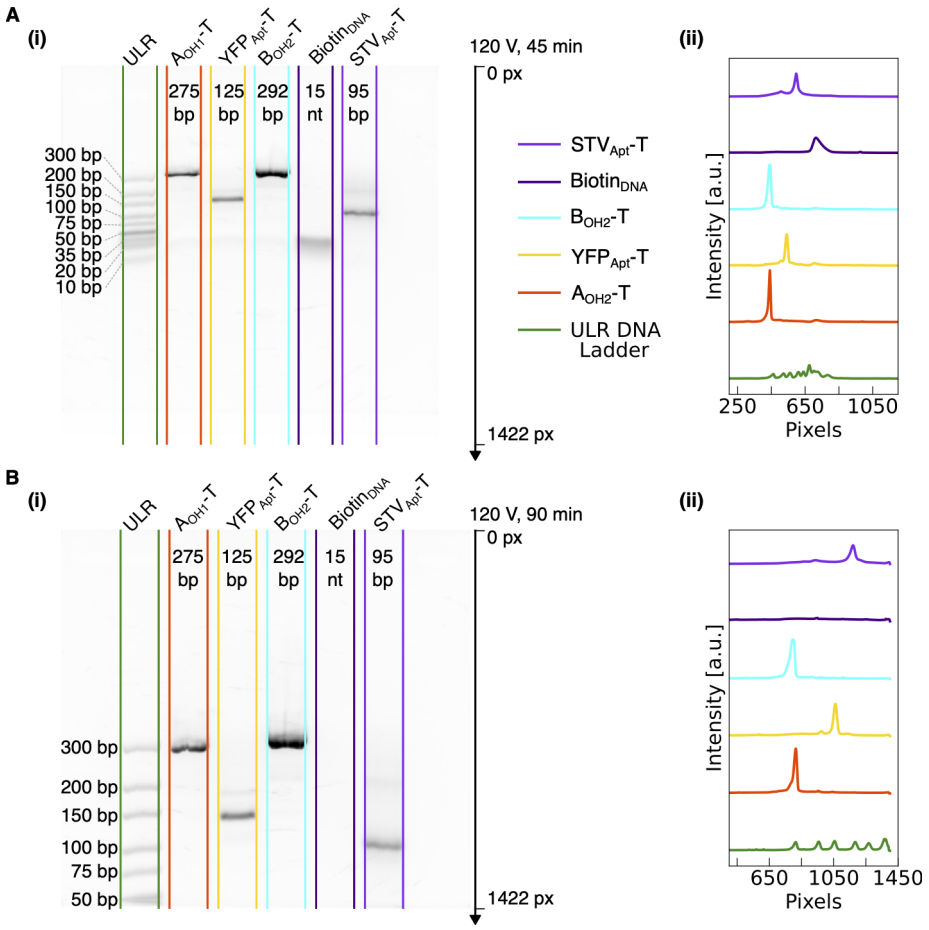


Fig. 6.2 **Agarose gel electrophoresis confirms DNA templates encoding for protein-capturing constructs display the expected sizes. A**: **(i)** Agarose gel representing the electrophoretic mobilities of PCR amplified DNA templates A_{OH1}-T, YFP_{Apt}-T, and B_{OH2}-T, as well as non amplified single-stranded oligonucleotide $Biotin_{DNA}$ and double-stranded template STV_{Apt}-T after a 45 min run at 120 V (see Section 2.5). Ultra-Low-Range (ULR) DNA ladder is used as reference in the first lane, coloured in green. **(ii)** Lane intensity profiles computed from gel in **A (i)**. Visual inspection of the gel in **A (i)** and the lane profiling in **A (ii)** confirms all DNA constructs display mobilities comparable to the expected sizes. **B**: **(i)** Same gel in **A (i)** after additional 45 min of electrophoresis. **(ii)** Lane intensity profiles computed from gel in **B (i)**. In both **A (i)** and **B (i)**, the length in base pairs for ladder markings is reported to the side of the corresponding bands, while the expected length of DNA templates and constructs is reported on top of the corresponding lane.

they have the potential to interact with downstream sequence domains, kinetically trapping them and ultimately preventing the designed topologies to be reached.

When examining the protein-binding RNA aptamers, YFP_{Apt} exhibits a single, faster migrating band, in accordance to its more compact geometry and overall shorter sequence. Conversely, STV_{Apt} displays two bands, with the expected one (slightly below YFP_{Apt}) not being, once again, the dominant band. This might be due to lower than expected DNA template purity, which was not PCR amplified nor gel purified due to the different production method (see Section 2.2). Nonetheless, as expected from the much shorter sequences and smaller construct sizes, both RNA aptamers migrate significantly faster than the corresponding RNA nanostars.

When comparing individual nanostars with those simultaneously transcribed with the ap-

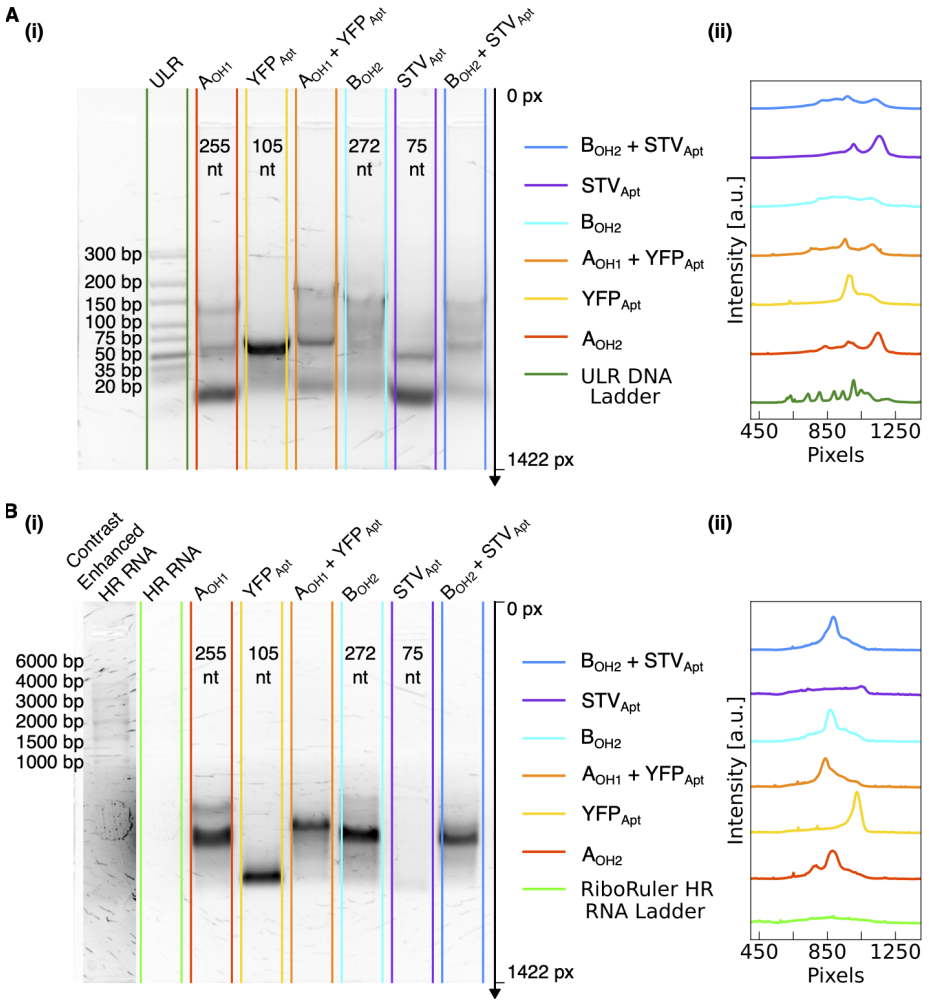


Fig. 6.3 **Agarose gel electrophoresis elucidates folding and size of protein-capturing RNA transcripts. A: (i)** Native agarose gel investigating the electrophoretic mobilities of co-transcriptionally folded RNA constructs A_{OH1}, YFP_{Apt}, B_{OH2}, STV_{Apt}, as well as the corresponding pairs ($A_{OH1} + YFP_{Apt}$, $B_{OH2} + STV_{Apt}$) upon simultaneous transcription in stoichiometric DNA template conditions. Ultra-Low-Range (ULR) DNA Ladder is used as reference in the first lane, coloured in green. **(ii)** Lane intensity profiles computed from gel in **A (i)**. **B: (i)** Denaturing agarose gel representing the electrophoretic mobilities of the same RNA transcripts in **A (i)**. RiboRuler High-Range (HR) single-stranded RNA Ladder is used as reference in the first lane, coloured in lime green. A contrast-enhanced reference lane is reported on the left-hand side for ease of visualisation given the extremely low contrast. **(ii)** Lane intensity profiles computed from gel in **B (i)**. In both **A (i)** and **B (i)**, the length in base pairs for ladder markings is reported to the side of the corresponding bands, while the expected length of RNA transcripts is reported on top of the corresponding lane. All samples apart from those simultaneously transcribed follow the colour-coding in Fig. 6.2.

tamer constructs featuring complementary sticky ends, we observe the emergence of bands corresponding to RNA aptamers alone, as well as, in the case of $A_{OH1} + YFP_{Apt}$, a slower migrating band, indicative of the formation of a larger complex. Such behaviour is not observed in $B_{OH2} + STV_{Apt}$, but this might be due to the resulting complex being close in mobility to the B_{OH2} nanostar alone, considering its large size and the relative compactness of STV_{Apt}.

Under denaturing conditions, instead, samples display fewer bands, as observed in Fig. 6.3B (i-ii). In particular, the protein-binding RNA aptamers display a single main band each. Nanos-

tars, on the other hand, still present either two bands (A_{OH1}) or a main band within a smear of products (B_{OH2}), likely due to the misfolding events hypothesised above. Finally, samples with simultaneously transcribed nanostars and corresponding protein-binding aptamers should exhibit lane intensity profiles roughly corresponding to the one of the nanostar and that of the aptamer superimposed, due to the sticky end interactions being denatured. While the bands of the RNA nanostars can be clearly identified in both such lanes, those of the protein-binding aptamers are less clearly visible. This is possibly due to a combination of lower aptamer transcription yield due to lower DNA template concentration compared to individually synthesised aptamers ([nanostar DNA template] = [aptamer DNA template] = 20 nM *vs* 40 nM), and different transcription efficiencies of the nanostar and aptamer constructs, which could result in more RNA nanostars being synthesised compared to protein-binding aptamers.
As expected, all samples migrate considerably faster than the smallest-sized ladder marking (1000 nt single-stranded RNA), which only provides a vastly overestimated upper bound for the nanostar size estimate. However, the main bands exhibited by STV_{Apt} (75 nt) and YFP_{Apt} (105 nt) can be used as supplementary reference markings, providing a useful lower bound for the size estimate of the RNA nanostars, both considerably slower than YFP_{Apt} in accordance with their larger sizes ($\approx$ by a factor of two).

6.3 RNA organelles can capture target proteins *via* simple design modifications

In this Section, I will demonstrate through confocal microscopy that the modified RNA nanostars described in Fig. 6.1 self-assemble into RNA organelles capable of sequestering EYFP and STV proteins as per the designed capture mechanisms. I will first discuss the systems involving protein-capturing RNA aptamers, and then the one relying on biotin-streptavidin interactions. The work presented in this Chapter directly builds on Chapter 5, in that transcription in carried out within emulsion droplet-based synthetic cells.

6.3.1 Homogeneous protein enrichment *via* protein-binding RNA aptamers

As outlined in Fig. 6.1A-B, nanostars A_{OH1} and B_{OH2} are designed to sequester EYFP and STV by hybridising to protein-binding constructs YFP_{Apt} and STV_{Apt}, respectively, *via* complementary stick ends. Since nanostars themselves lack any protein-binding domains, protein-capturing is only possible upon transcription of the aptamers. As such, omitting the corresponding DNA templates should impede preferential protein partitioning in the RNA organelles. Confocal micrographs in Fig. 6.4A and B prove, indeed, that the designed mechanism works for both A_{OH1}-YFP_{Apt} and B_{OH2}-STV_{Apt} systems, respectively.

Considering the A_{OH1}-YFP_{Apt} system first (Fig. 6.4A), RNA nanostars transcribed in the absence of the DNA template coding for YFP_{Apt} (-YFP_{Apt}-T) do form bright red organelles, as demonstrated for pristine A nanostars in Fig. 5.1B (i), but fail to sequester EYFP proteins, which rather remain dispersed in the droplet lumen, as observed in Fig. 6.1A (i-ii) (left). When simultaneously transcribed with YFP_{Apt} (+YFP_{Apt}-T), though, the assembled RNA organelles appear highly enriched in EYFP, as demonstrated by their orange (overlap of red and yellow)/yellow colour in the presence/absence (middle/right) of MG dye, respectively.
Similarly, Fig. 6.4B shows that RNA nanostars B_{OH2} synthesised in the absence of the DNA template coding for STV_{Apt} (-STV_{Apt}-T) form the expected cyan organelles, as observed for pristine B nanostars in Fig. 5.1B (ii). However, these RNA assemblies fail to preferentially partition the bulky Alexa405-STV proteins, the majority of which remains dispersed in the droplet lumen, as observed in confocal micrographs in Fig. 6.1B (i-ii) (left). Simultaneous transcription of nanostars and STV_{Apt} (+STV_{Apt}-T), instead, enables the self-assembled B_{OH2} organelles to uptake considerably more Alexa405-STV, as evidenced by their indigo colour (overlap of cyan and magenta, right).
Notably, these systems produce RNA organelles that appear homogeneously enriched in the target proteins. Such protein partitioning profiles are a direct consequence of i) the liquid-like behaviour of organelles and ii) the capture mechanism, which relies on simultaneous transcription of nanostars and protein-binding aptamers. At first, RNA nanostars and protein-binding aptamers are synthesised as time goes by – albeit the latter at a slower pace due to template concentration imbalance, namely $[A_{OH1}\text{-T}]/[YFP_{Apt}] = [B_{OH2}\text{-T}]/[STV_{Apt}] = 3$. As a result, target EYFP and Alexa405-STV proteins are progressively sequestered into the A_{OH1} and B_{OH2} RNA organelles, respectively. Due to the low concentration of protein moieties in use, at some point most (if not all) proteins will have been captured, but synthesis of RNA nanostars and protein-binding constructs will proceed as long as T7 RNAP is active and ribonucleotides are not depleted, leading to further RNA organelle growth. Thanks to the liquid-like behaviour of the organelles, though, the sequestered protein moieties can redistribute within them *via* diffusion of the nanostars. Alternatively, redistribution could also occur due to reversibility of protein-aptamer binding. Both these mechanisms ultimately lead to uniform enrichment profiles, as proved by the 2D cross-sections from confocal z-stacks in Fig. 6.5A-B (iii), respectively.

6.3.2 Inhomogeneous STV capture *via* receptor-ligand interactions

Biologically occurring MLOs often exhibit an asymmetrical distribution (sometimes RNA-induced [34]) of their protein components. Fig. 6.4C proves that it is possible to engineer capture mechanisms resulting in such inhomogeneous distribution of target proteins within the discussed synthetic RNA organelles, as demonstrated through the biotin-mediated capture of STV (here TexasRed-STV) inside B_{OH2} assemblies.
As per the design in Fig. 6.1C, in the absence of the biotinylated oligonucleotide (-$Biotin_{DNA}$),

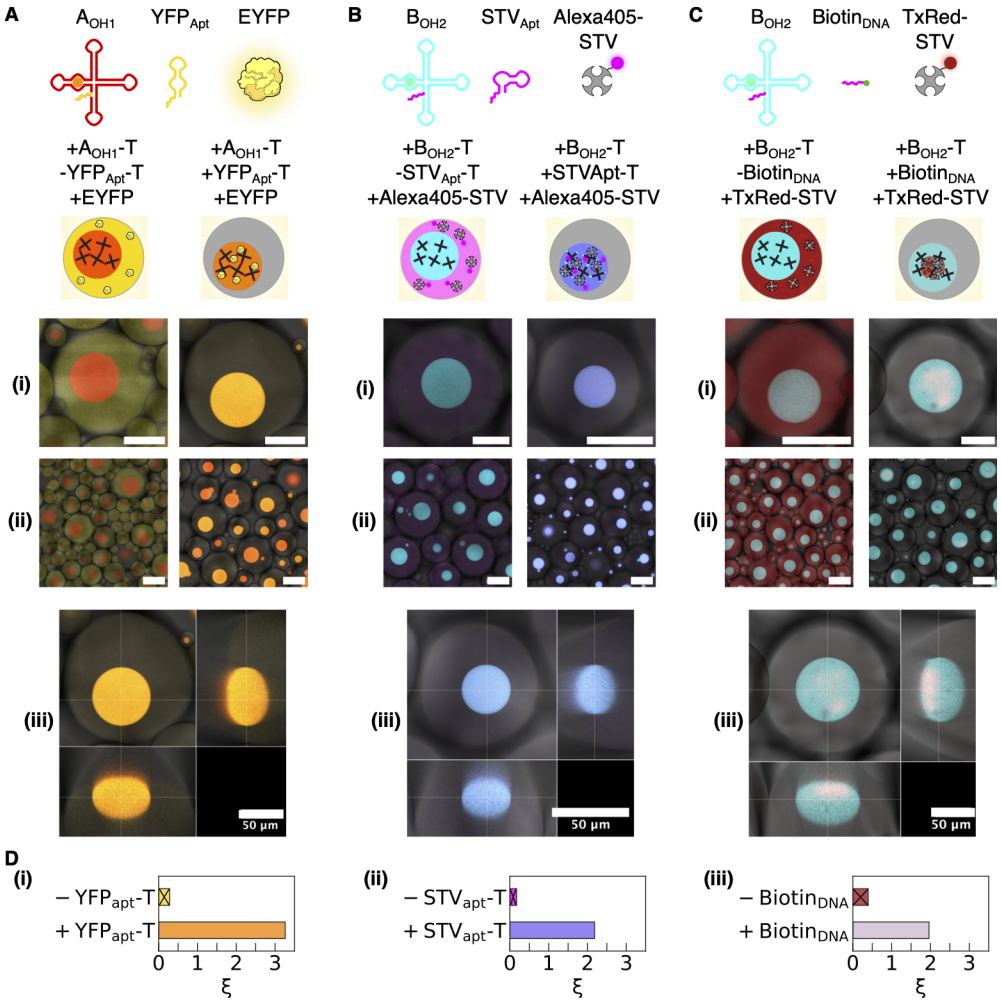


Fig. 6.4 **RNA organelles can capture target proteins through protein-binding constructs.** **A**: RNA nanostars A_{OH1} (red) fail to capture EYFP (yellow) in the absence of YFP_{Apt} (-YFP_{Apt}-T, left), while they succeed in doing so when simultaneously transcribed with it (+YFP_{Apt}-T, right), as demonstrated by **(i)** zoomed-in and **(ii)** large FOV confocal micrographs (left/middle and right, respectively). **(iii)** XY (top left), XZ (top right), YZ (bottom left) 2D cross-sections from confocal z-stacks outlining the internal microstructure and protein distribution within A_{OH1} organelles in the presence of YFP_{Apt} and MG (**(i-ii)**, middle). **B**: RNA nanostars B_{OH2} (cyan) fail to capture Alexa405-STV (magenta) in the absence of STV_{Apt}-T (-STV_{Apt}-T, left), while they succeed in doing so when simultaneously transcribed with it (+STV_{Apt}-T, right), as demonstrated by **(i)** zoomed-in and **(ii)** large FOV confocal micrographs (left/right, respectively). **(iii)** Same as **A (iii)** for B_{OH2} organelles in the presence of STV_{Apt} and DFHBI (**(i-ii)**, right). **C**: RNA nanostars B_{OH2} (cyan) fail to capture TexasRed-STV (TxRed-STV) in the absence of $Biotin_{DNA}$ (-$Biotin_{DNA}$, left), while they succeed in doing so in its presence (+$Biotin_{DNA}$, right), as demonstrated by **(i)** zoomed-in and **(ii)** large FOV confocal micrographs (left/right, respectively). Due to the limiting concentration of STV, the resulting organelles display a protein-enriched core/protein-depleted shell structure (**(i-ii)**, right). **(iii)** Same as **A (iii)** for B_{OH2} organelles in the presence of $Biotin_{DNA}$ and DFHBI (**(i-ii)**, right). **D**: Partitioning coefficient ξ, computed as the ratio between mean protein fluorescence intensity inside the RNA organelles and in the surrounding synthetic cell lumen (see Section 2.11.17), for the different investigated systems, as measured from the single confocal micrographs in **A-C (ii)**. For each system, the bar plot compares ξ in the absence (-) and presence (+) of the protein-binding construct, namely **(i)** YFP_{Apt} for **A**, **(ii)** STV_{Apt} for **B**, and **(iii)** $Biotin_{DNA}$ for **C**, respectively. Confocal micrographs in **A, C** are pristine, while those in **B** have been contrast-enhanced (bright-field only) to aid visualisation. Scale bars in **A-C (i)** and **(iii)** are 50 μm. Scale bars in **A-C (ii)** are 100 μm. Epifluorescence micrographs depicting the self-assembly of samples in **A-C** are provided in the Appendix (Fig. B.13). EYFP graphical element in **A** was created with BioRender.com.

the RNA nanostars fail to capture TexasRed-STV, resulting in bright cyan condensates surrounded by a uniformly red droplet lumen, as observed in confocal micrographs in Fig. 6.1C (i-ii) (left). Conversely, in the presence of $\text{Biotin}_{\text{DNA}}$ (+$\text{Biotin}_{\text{DNA}}$), the RNA nanostars bind the short biotin-functionalised oligonucleotides through complementary OH2 and OH2* sticky end domains, and $\text{Biotin}_{\text{DNA}}$ constructs capture TexasRed-STV *via* high-affinity biotin-streptavidin interactions. Interestingly, the resulting organelles exhibit a more complex architecture than the one observed in Fig. 6.4B, comprising two distinct microenvironments with different TexasRed-STV enrichment levels. In particular, confocal micrographs in Fig. 6.1C (i-ii) (right) and 2D cross-sections from confocal z-stacks in Fig. 6.1C (iii) reveal a STV-enriched core and a STV-depleted shell. Notably, the TexasRed-STV-rich core seems to be shifted to the lower portion of the B_{OH2} organelle, most likely due to density-induced sedimentation. Such inhomogeneity in the target protein distribution directly contrasts with the homogeneous partitioning observed in organelles relying on STV_{Apt} in Fig. 6.4B (iii), and is in excellent agreement with the hypothesised working principle of this alternative capture strategy, as presented in Fig. 6.1C. This inhomogenoeus protein partitioning profile stems, in fact, from i) the limiting concentration of target proteins, most of which get sequestered in the early stages of transcription as described above, and ii) the lack of structural rearrangement of the STV-rich core formed by the first RNA nanostars cross-linked by STV. The lack of geometrical symmetry of the STV-rich core observed in Fig. 6.1C (iii) hints, in fact, at this being a hydrogel-like kinetically arrested phase emerging from each STV moiety irreversibly coordinating four RNA nanostars. While further experiments, particularly Fluorescence Recovery After Photobleaching (FRAP), would be required to confirm this is the case, such a hydrogel-core/liquid-like-shell morphology is compelling, as multi-layered biological MLOs often feature a comparable architecture [34, 36].

6.3.3 Quantification of protein uptake

Segregation of the protein cargo within the designed RNA organelles can be evaluated by computing the partitioning coefficient, ξ. This is defined as the ratio between the mean protein fluorescence intensity inside the RNA organelles and the one in the surrounding synthetic cell lumen, *i.e.* within/outside the binary mask of the condensates as obtained through image segmentation (see Section 2.11.17). Fig. 6.4D compares such partitioning coefficients for all discussed systems in the presence or absence of the corresponding protein-binding constructs, namely YFP_{Apt} for EYFP in Fig. 6.4D (i), STV_{Apt} for Alexa405-STV in Fig. 6.4D (ii), and $\text{Biotin}_{\text{DNA}}$ for TexasRed-STV in Fig. 6.4D (iii).

As expected, in the absence of protein-binding constructs the RNA organelles are generally depleted in proteins compared to the surrounding solution, simply because of the large protein size, as demonstrated by the low corresponding ξ values for both EYFP and STV variants. In all systems, though, protein enrichment within RNA organelles is considerably enhanced

when the protein-binding constructs are present, as demonstrated by the dramatic increase in the partitioning coefficients: all values are around or above two, indicating that organelles are, on average, around twice as bright as the droplet lumen, as clearly visible in the micrograph in Fig. 6.4A-C (i-ii).
Notably, systems designed to capture STV, regardless of the approach, display lower partitioning coefficients, *i.e.* protein-capturing efficiencies, compared to the EYFP one. This can be easily rationalised considering the larger size of STV compared to EYFP (≈ 60 kDa *vs* ≈ 27 kDa), which clearly leads to lower incorporation of the bulky STV proteins within the RNA condensate porous microstructure due to the higher steric encumbrance. Finally, the slightly lower partitioning coefficient for biotin-mediated STV capture (Fig. 6.4D(iii)) compared to the STV_{Apt} mechanism (Fig. 6.4D(ii)) is likely linked to the the inhomogeneous protein distribution of the former. This inhomogeneity, in fact, skews the mean fluorescence signal inside the RNA organelles towards lower values, especially when performing the analysis on single confocal planes where various organelles are imaged at different focal heights. This can, in fact, fail to capture the hydrogel-like STV-rich cores which, as shown in Fig. 6.4C(iii), tend to sediment to the bottom of the RNA organelles.

6.3.4 Simultaneous capture of distinct target proteins *via* binary RNA nanostar systems

Thanks to the orthogonality among pristine A and B nanostars and the (intended) non complementarity of the 5' sticky ends, simultaneous transcription is expected to produce two orthogonal phases capable of capturing distinct target proteins. Therefore, I tested the possibility of selectively capturing both EYFP and STV in simultaneously transcribed and co-assembled A_{OH1} and B_{OH2} organelles.
Contrary to expectations, though, Fig. 6.5 reveals that all strategies yield emulsion droplets containing single multi-phase RNA organelles rather than two single-phase ones. Nonetheless, these multi-phase RNA assemblies all feature very distinct protein enrichment domains that still follow the design guidelines in Fig. 6.1. For example, Fig. 6.5A shows that simultaneous transcription of A_{OH1} and B_{OH2} nanostars, as well as YFP_{Apt} and STV_{Apt} constructs, produces organelles where A-rich/B-rich domains are preferentially enriched in EYFP/Alexa405-STV, respectively.
The resulting organelle morphologies closely resemble those of linker-mixed A/B systems with intermediate linker fractions (Fig. 5.6), thus clearly hinting at some degree of interaction among the participating constructs (nanostars, aptamers and proteins). Despite the expected construct orthogonality, in fact, control experiments presented in Fig. 6.6A-C reveal that some interactions are already observed in systems comprising nanostars only, which mostly display single large B_{OH2}-rich phases sandwiched between or enveloped within A_{OH1}-rich ones (Fig. 6.6C). As demonstrated by micrographs in Fig. 6.6D-F, though, systems presenting

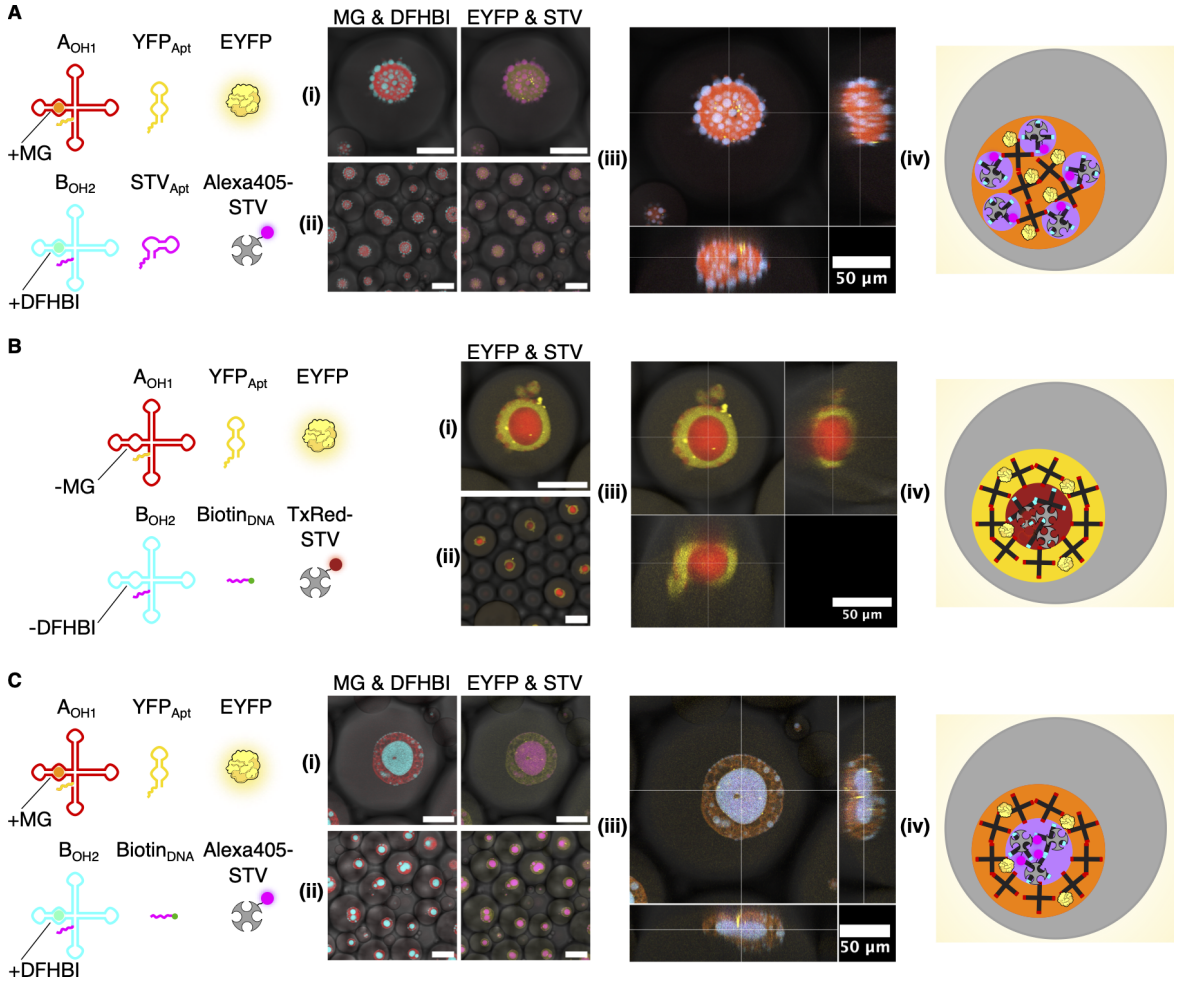


Fig. 6.5 **Simultaneous capture of distinct proteins in addressable domains of multi-phase RNA organelles.** **A**: Simultaneous transcription of nanostars A_{OH1} and B_{OH2} and aptamers YFP_{Apt} and STV_{Apt} yields multi-phase RNA organelles exhibiting distinct A_{OH1} (MG) and EYFP-rich and B_{OH2} (DFHBI) and Alexa405-STV-rich domains, as demonstrated by **(i)** zoomed-in and **(ii)** large FOV confocal micrographs. **(iii)** XY (top left), XZ (top right), YZ (bottom left) 2D cross-sections from confocal z-stacks (composite of all available channels) outlining the internal microstructure and protein distribution of samples in **(i-ii)**, accompanied by schematic representation in **(iv)**. **B**: Same as **A** for binary systems sequestering TexasRed-STV (TxRed-STV) through $Biotin_{DNA}$. MG and DFHBI dyes were here omitted. The different STV capture strategy leads to further coalescence of TexasRed-STV domains into a larger central core, as demonstrated by **(i)** zoomed-in and **(ii)** large FOV confocal micrographs. **(iii-iv)**. Same as **A (iii)** for samples in **B (i-ii)** and **B (iv)**. Samples in **A, B** were imaged after a week from start of transcription. **C**: Same as **B** but replacing TexasRed-STV with Alexa405-STV to confirm protein localisation in the distinct A_{OH1}/B_{OH2} domains. **(i)** Zoomed-in and **(ii)** large FOV confocal micrographs confirm co-localisation of MG (A_{OH1}) and EYFP, as well as DFHBI (B_{OH2}) and Alexa4-STV. **(iii-iv)** Same as **A (iii)** for samples in **C (i-ii)** and **C (iv)**. Samples in **C** were imaged after three weeks from start of transcription. Confocal micrographs in **A, B** are pristine, while those in **C** have been contrast-enhanced. Scale bars in **A-C (i)** and **(iii)** are 50 μm. Scale bars in **A-C (ii)** are 100 μm. Epifluorescence micrographs depicting the self-assembly of samples in **A-C** are provided in the Appendix (Fig. B.14). EYFP graphical element in **A-C** was created with BioRender.com.

both nanostars and protein-binding aptamers exhibit different morphologies, with more numerous and smaller B_{OH2}-rich domains dispersed into or adhering to a A_{OH1}-rich phase (Fig. 6.6F), identical to what observed for multi-phase organelles in Fig. 6.5A. Analysis of systems comprising both nanostars and/or both aptamers in NUPACK reveals, indeed, low

levels of undesired interactions leading to mixed A_{OH1}-B_{OH2} complexes (< 3 % of overall concentration), mediated in all cases by the single-stranded overhangs introduced in this Chapter. While NUPACK computations cannot realistically model complex RNA-RNA interactions, this spurious mixing is likely caused by either a) partial complementarity between the two 5' overhangs, or b) interactions between core nanostar sequence domains due to overhang-induced misfolding.

Mechanism a) is simpler, as it involves multiple direct, albeit weak, overhang (OH1) - overhang (OH2) interactions which lead to partial mixing of the two phases. Mechanism b), despite being more complex, is consistent with the aforementioned observation that individual RNA nanostars featuring such sticky ends (Fig. 6.3A (i)) exhibit lower folding efficiency compared to pristine ones (Fig. 4.3A (i)), and this can potentially be ascribed to misfolding. During RNA synthesis, the 5' overhangs are the first sequence stretches to be transcribed, and they lack any strong secondary structure, as they are single-stranded. As such, after being synthesised, they can bind to newly transcribed downstream sequence domains within the core of the same RNA nanostars, especially within the first nanostar arm, which remains single-stranded almost until the end of the synthesis of the entire nanostructure and could thus be prone to establishing spurious local interactions. These could lead to the corresponding nanostars being kinetically trapped in misfolded configurations, where some core sequence domains are exposed. Because A_{OH1} and B_{OH2} nanostars feature almost identical core sequences (apart from the fluorogenic aptamers, like A and B nanostars), these misfolded configurations may enable them to interact through their exposed cores, leading to unwanted binding.

Regardless of the specific mechanism, the analysis performed in NUPACK seems to point to A_{OH1}-B_{OH2} nanostar-nanostar interactions as the main contributor to this undesired mixing. However, Fig. 6.6D and Fig. 6.6F show that the presence of protein-binding aptamers does influence the final morphology, thus hinting at a role of aptamer-nanostar and/or aptamer-aptamer interactions, even though this could simply be due to increased steric and Coulomb repulsion among nanostars with attached aptamers.

Similarly to what observed in Fig. 6.5A, Fig. 6.5B demonstrates that capturing TexasRed-STV *via* biotin-conjugated oligonucleotides ($Biotin_{DNA}$) instead of STV_{Apt} still yields multi-phase organelles. However, the resulting RNA assemblies exhibit a different morphology, more closely resembling a core-shell or fried-egg-like architecture. This result is supported by Fig. 6.5C, where TexasRed-STV is replaced with Alexa405-STV in order to investigate colocalisation of proteins with nanostar phases through MG and DFHBI dyes, which is not easily achievable with TexasRed due to its considerable fluorescence overlap with MG+MGA. Such morphological difference can be explained by taking into account the one observed between B_{OH2} organelles capturing STV through STV_{Apt} and $Biotin_{DNA}$. In the former approach, B_{OH2} nanostars sequester STV progressively with time, as do A_{OH1} nanostars with EYFP. Domains start forming and, despite designed not to do so, interact/adhere to

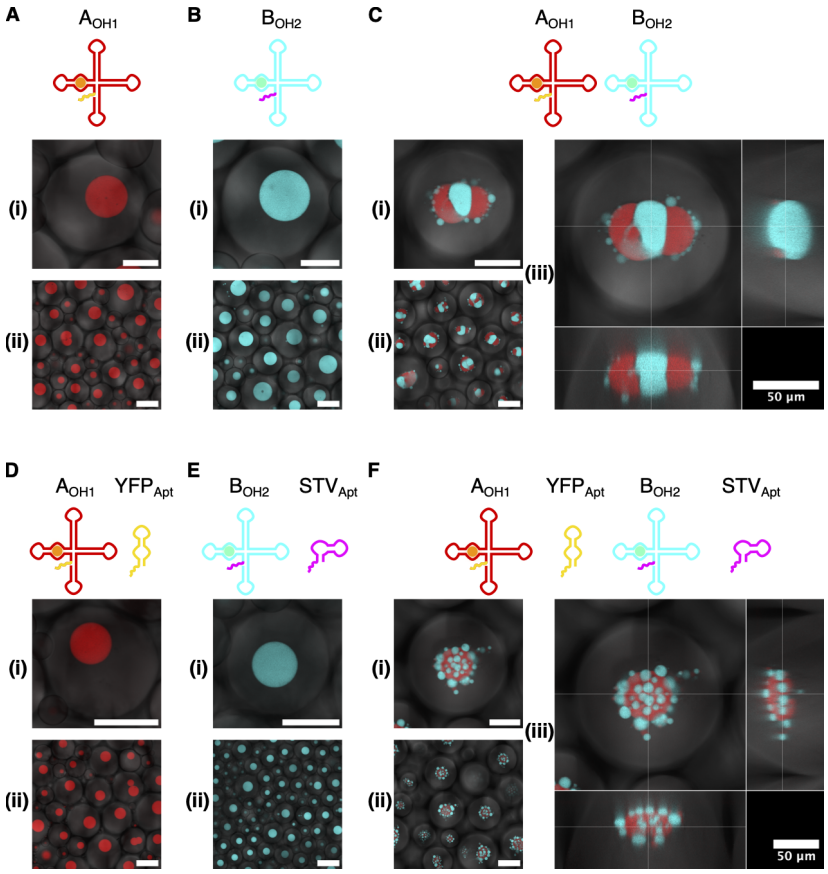


Fig. 6.6 **Protein-capturing RNA nanostars form multi-phase organelles despite KL orthogonality.** Uncropped, pristine confocal micrographs depicting RNA organelles formed by protein-capturing systems comprising one/two nanostar species (left and middle/right) in the absence/presence (top/bottom) of the corresponding protein-binding aptamers, respectively. **A**: A_{OH1} forms single MG-rich (red) organelles within emulsion droplets, as demonstrated by **(i)** zoomed-in and **(ii)** large field-of-view (FOV) confocal micrographs, and previously observed for unmodified A nanostars (Fig. 5.1B(i)). **B**: B_{OH2} forms single DFHBI-rich (cyan) organelles within emulsion droplets, as demonstrated by **(i)** zoomed-in and **(ii)** FOV confocal micrographs, and previously observed for unmodified B nanostars (Fig. 5.1B(ii)). **C**: Simultaneous transcription of A_{OH1} and B_{OH2} (with [A_{OH1}-T] = [B_{OH2}-T] = 20 nM) produces multi-phase RNA organelles, as demonstrated by **(i)** zoomed-in and **(ii)** FOV confocal micrographs. The internal microstructure of sample in **(i)** is illustrated in **(iii)** *via* XY (top left), XZ (top right), YZ (bottom left) 2D cross-sections from confocal z-stacks. **D**: Same as **A** with A_{OH1} co-transcribed with YFP_{Apt} (no EYFP protein added). **E**: Same as **B** with B_{OH2} co-transcribed with STV_{Apt} (no STV protein added). **F**: Same as **C** with co-transcribed YFP_{Apt} and STV_{Apt} (no proteins added). All samples were imaged after 48 h from the start of transcription. Scale bars in **A-F (i)** and **(iii)** are 50 μm. Scale bars in **A-F (ii)** are 100 μm. All constructs are simultaneously transcribed, with [nanostar DNA template]/[aptamer DNA template] = 3 for each pair in **D-F**, and [total DNA template] = 40 nM.

one another as discussed above. While A_{OH1} nanostars have weaker and more fluid KL A interactions and bind smaller EYFP proteins, B_{OH2} nanostars have stronger KL B interactions and sequester much bulkier STV proteins. Moreover, the accumulation of STV moieties induces a high steric hindrance to the coalescence of B_{OH2} domains, which remain scattered within a bulk A_{OH1} phase. Conversely, in the latter approach, most STV moieties are captured by the first transcribed B_{OH2} nanostars, when A_{OH1} still cannot fully bind all available EYFP due to the lack of sufficient YFP_{Apt}. As discussed above, the biotin excess leads to each STV moiety crosslinking four nanostars in a denser STV-rich hydrogel-like B_{OH2} organelle core

around which the more fluid A_{OH1} phase, hosting smaller EYFP proteins, can accumulate and assemble.
Comparing micrographs in Fig. 6.5B (samples imaged after one week) and Fig. 6.5C (samples imaged after three weeks) reveals the emergence of cavities/vacuoles in the A_{OH1} phase of the latter, as previously observed in A condensates in bulk (Fig. 4.5A) and emulsion droplets (Fig. 5.1C). This is likely due to either structural re-arrangement within the more fluid A phase, or degradation processes, both occurring on longer timescales. Fig. 6.5A-C also show bright yellow fluorescent specks, most likely due to EYFP aggregation. Despite following manufacturer instructions (see Section 2.2) to resuspend in plain nuclease-free water, this method clearly proved inadequate. Resuspending in Phosphate Buffered Saline (PBS) with 5 mM sodium azide as for fluorophore-conjugated STV will likely yield better results.
While the binary A_{OH1}/B_{OH2} system failed to form two large orthogonal organelles capturing distinct proteins in well separated phases, it still demonstrates the possibility of engineering distinct microscale domains selectively enriched in multiple proteins within a synthetic RNA MLO. Moreover, it is most likely still possible to achieve the original goal *via* one of the following routes. First, the single-stranded overhang sequences can be re-engineered to avoid a) overhang-overhang interactions between A/B nanostar populations, and b) complementarity to downstream core sequence domains within the same nanostar, particularly within the exposed portion of the first arm. Second, core sequences of either or both A and B nanostars could be modified to make A_{OH1} and B_{OH2} fully orthogonal upon co-transcriptional folding, for instance using the Design module in NUPACK with one of the two designs fixed. Third, the sticky ends could be re-designed to fold into hairpins hiding the capture domains. Such hairpins could end in a toehold, enabling overhangs displaying complementary capture domains to open in the presence of one another, *e.g.* hairpin with OH1 in A_{OH1} and hairpin with $OH1^*$ in YFP_{Apt}. This would dramatically reduce the risk of undesired interactions between the overhang domains and downstream nanostar sequence elements.

6.4 Chapter summary and outlook

In this Chapter, I have modified the RNA nanostar platform discussed in Chapters 4 and 5 to yield RNA organelles capable of selectively sequestering target proteins. This design modification is straightforward, and simply involves the addition of 5' sticky ends which enable hybridisation to complementary ones within protein-binding constructs. Moreover, this approach can be directly integrated with most state-of-the-art DNA and RNA nanotechnology approaches due to the ubiquitous use of sticky end interactions.
Importantly, I have demonstrated that it is possible to engineer both homogeneous and inhomogeneous protein distribution profiles within RNA organelles depending on the protein capture strategy. On the one hand, simultaneous co-transcriptional folding of RNA nanostars and protein-binding aptamers yields homogeneously protein-enriched synthetic

RNA organelles in droplet-based artificial cells. While this approach was only demonstrated for EYFP and STV, a wide range of RNA aptamers have already been engineered to bind biologically relevant proteins including, but not limited to, cytokines, proteases, kinases, cell-surface receptors and cell-adhesion molecules [480–482]. Novel RNA aptamers can further be discovered *via* SELEX, for instance directing the evolution towards constructs capable of binding the enzymes necessary to sustain a particular reaction cascade or pathway [483]. Moreover, while astoundingly simple, the ability to sequester biomolecules such as proteins in designer condensates is essential for their viability as a means to influence the behaviour of living agents such as bacteria in which such organelles could be expressed [484, 39, 31]. On the other hand, replacing a transcribed STV-binding aptamer with a short biotin-conjugated DNA construct added before the start of transcription produces RNA organelles with a protein-rich core/protein-depleted shell. Such inhomogeneous protein enrichment profile more closely resembles those observed in biologically occurring MLOs, where different protein components preferentially partition in distinct domains [34, 31, 32]. As layered architectures have been hypothesised to contribute to enzymatic reaction directionality [32], similar inhomogeneously enriched organelles could prove useful in metabolic engineering of the sought after artificial cell mimics.

However, the external addition of a biotin-conjugated DNA construct is less biologically relevant compared to the transcription of an encoded protein-binding RNA aptamer. This downside could be mitigated by replacing the DNA oligonucleotide with a biotin-tagged RNA transcript with identical sequence, which can be obtained through modified *in vitro* transcription [485]. In order to prescribe the spatial distribution of such protein-binding constructs, one could exploit a combination of switchable transcription systems [486] and reaction-diffusion patterning, as recently demonstrated for DNA-based condensates by Leathers *et al.* [154]. Such an approach could also be extended to protein-binding RNA aptamers, widening the range of compatible target proteins.

Finally, I have proved that binary RNA nanostar systems allow to simultaneously capture two target proteins. While the demonstrated system failed to yield two fully orthogonal phases enriched in distinct proteins, the obtained multi-phase RNA organelles exhibit distinct A_{OH1}/B_{OH2} phase domains programmably enriched in EYFP/STV, respectively. Further sequence engineering of the 5' overhangs and/or the core nanostar motifs directed to minimise undesired interactions upon co-transcriptional folding will likely yield completely orthogonal A_{OH1}-B_{OH2} organelles with markedly different protein partitioning. It is also easy to envision how this platform could be scaled up to capture multiple target biomolecules, for instance by simultaneous transcription of multiple protein-binding aptamers. This, in turn, would pave the way for more complex and programmable engineering of the inner workings of artificial cells, with some organelles sustaining enzymatic cascades through enzyme-substrate colocalisation, and others suppressing particular metabolic pathways by sequestering the

necessary biomolecules upon specified stimuli, bringing us closer to the ultimate goal of replicating life-like behaviours.

Chapter 7

Conclusions & future research directions

The quest to construct ever more biomimetic synthetic cell systems has attracted considerable interest in recent years, due to their potential in a range of scientific and technological areas. Artificial cell platforms have, in fact, been employed as minimal model systems to study cellular processes and mechanisms in a simplified setting, where they can be disentangled from the surrounding, often overwhelming, biological complexity [9]. They have also showed great promise for application in a wide variety of biomedical applications, from biosensing, theranostics and controlled drug delivery and release, to biosynthesis, bioremediation and tissue engineering [13]. Similarly to how biological cells comprise a range of membrane-bound and membrane-less organelles (MLOs), these artificial cell systems require varying degrees of compartmentalisation to organise their molecular components, catalyse biomolecular reactions and direct metabolic flux. While synthetic mimics of cellular membranes and membrane-bound organelles mostly rely on lipid [41, 42], polymer [42, 44] or protein [17] membranes, nucleic acids have been identified as a prime candidate to engineer synthetic MLOs from the bottom up. In particular, the programmability and versatility of nucleic acids are key to replicate the complex behaviour and internal architecture of biological MLOs.

In this thesis, I investigated two approaches to engineer membrane-less compartments in artificial cell systems in the form of synthetic DNA and RNA-based condensates designed from the bottom up. Thanks to the versatile toolkit of nucleic acids nanotechnology, I was able to design both hydrogel- and liquid-like nucleic acids condensates, both in bulk and within minimal model systems, capable of responding to stimuli such as changes in cationic environment or the presence of a specific promoter-bearing DNA template encoding for RNA building blocks. I also succeeded in engineering organelles exhibiting highly complex morphologies, and, despite the need for future work, capable of capturing target proteins according to distinct partitioning profiles, pushing the boundaries of nucleic acids-sustained biomimicry.

7.1 Summary & discussion

In **Chapter 3** I demonstrated that combining non-canonical DNA G4s and amphiphilic DNA nanostars produces cation-responsive nanostructures, termed Quad-Stars, which isothermally self-assemble into hydrogel aggregates/disassemble in response to the addition/removal of physiologically relevant K^+ ions, respectively. The hydrogel aggregates are stable under model physiological conditions, can uptake small hydrophobic molecules, and release them upon disassembly, thus simplistically mimicking the behaviour of hydrogel-like MLOs in biological cells.

In **Chapter 4** I have introduced a novel platform based on co-transcriptionally folding RNA nanostars capable of self-assembling into RNA condensates. Simultaneous transcription of RNA nanostars featuring non-interacting KLs yields orthogonal, individually addressable RNA condensates. Interactions are highly programmable, allowing to prescribe which designs form condensates and which fail to do so, and sequence engineering leads to different biophysical properties and morphologies of the RNA assemblies. Notably, all nanostars are transcribed from dsDNA templates, enabling direct integration with state-of-the-art DNA nanotechnology approaches.

In **Chapter 5** I proved that the RNA framework introduced in Chapter 4 can be encapsulated in simple emulsion droplet-based synthetic cells to produce RNA organelles resembling biologically occurring MLOs. These organelles scale in volume with the surrounding compartments and retain the programmability and addressability demonstrated in Chapter 4. Simultaneous transcription of non-interacting RNA nanostars produces orthogonal organelles – up to three distinct compartments per synthetic cell – with tunable relative size. Addition of a simple chimeric linker nanostar to otherwise orthogonal RNA nanostructures enables control over mixing/demixing behaviours and unlocks a whole new range of condensate morphologies, from Janus-like structures to hollow microcapsules and core-shell architectures.

Finally, in **Chapter 6** I modified the RNA organelles discussed in Chapter 5 to make them capable of capturing target proteins. I showed that it is possible to engineer the spatial protein enrichment profile within the RNA organelles through the choice of the specific capture mechanism, producing both homogeneous and inhomogeneous – enriched core-depleted shell – protein partitioning. While further work is required to ensure full orthogonality, I also proved that simultaneous transcription of binary RNA nanostar systems capable of sequestering two different proteins results in multi-phase organelles with distinct, programmably enriched microdomains.

Both platforms presented in this thesis provide strategies to design and produce nucleic acids-based condensates displaying biomimetic behaviours. Depending on the approach and

the tuning of material properties through sequence design and experimental conditions, it is possible to produce smart condensates exhibiting either hydrogel- or liquid-like behaviour. These platforms can respond to stimuli or information, be it physiologically relevant cation identity and concentration, or the presence of T7 promoter-bearing DNA templates encoding for RNA building blocks.

On one hand, Quad-Stars demonstrate that modularity can really be feature-enabling in DNA nanotechnology: simply replacing the canonical four-way Holliday junction with a DNA G4 grants Quad-Stars and the resulting hydrogel assemblies the ability to reversibly assemble/disassemble at constant temperature in response to K^+ ion addition/chelation. The modularity of Quad-Stars has already inspired a similar approach, recently presented in the literature, this time exploiting i-Motifs to achieve pH-responsiveness [408], and the same principles could be applied to the broader spectrum of DNA nanotechnology.

On the other hand, I believe that the novelty and biomimicry of the introduced RNA nanostar platform truly make it a game-changer in the bottom-up synthetic biology landscape. Compared to previous approaches, in fact, it works at constant temperature, removing any need for annealing, and it exploits biologically evolved machinery and mechanisms, namely the T7 RNAP enzyme and transcription, as well as co-transcriptional folding, as observed in biological RNAs. Most notably, this platform produces synthetic organelles within artificial cell mimics that are based on RNA, not a foreign material, rather a highly relevant one, already involved in the formation of the biological features, namely MLOs, that our synthetic constructs aim at mimicking. Finally, the RNA nanostars are transcribed from fully double-stranded DNA templates, enabling direct integration with other nucleic acids nanotechnology-based synthetic biology frameworks. For these reasons, I am confident this work will pave the way for much more complex, versatile and biomimetic minimal cell systems relying on profoundly interconnected synthetic designer constructs and biologically derived machinery and mechanisms.

7.2 Outlook and directions for future research

As mentioned throughout this thesis, the capability to rationally engineer synthetic MLO-mimicking condensates from first principles would be very valuable in the quest to produce ever more biomimetic artificial cells. The RNA nanostar-based platform discussed in Chapters 4-6 is particularly promising to these aims, as it yields *de novo* designed, addressable and orthogonal condensates with tunable size and controllable mixing/demixing. In this Section, I briefly highlight some directions for future work that directly build on Chapters 4-6, some supported by preliminary results and some still at the conceptual stage.

7.2.1 Encapsulation of RNA MLOs in more realistic artificial cells

First of all, it should be possible to synthesise and assemble the RNA organelles discussed in Chapter 5 in more biomimetic artificial cell mimics, particularly Giant Unilamellar Vesicles (GUVs). The most common approach to encapsulate charged biomolecules, such as those within *in vitro* transcription mixtures, within GUVs is phase transfer *via* the inverted emulsion method, schematically depicted in Fig. 7.1A. In this case, the aqueous phase comprises the transcription mixture with added T7 RNAP and DNA template, as well as a suitable amount of sucrose ensuring the density gradient necessary for sedimentation. As phase transfer is known to lead to poor yield [45] and suboptimal encapsulation efficiency [46], a microfluidic setup could be used instead, which would also improve the size control and monodispersity of the resulting artificial cell mimics [487].

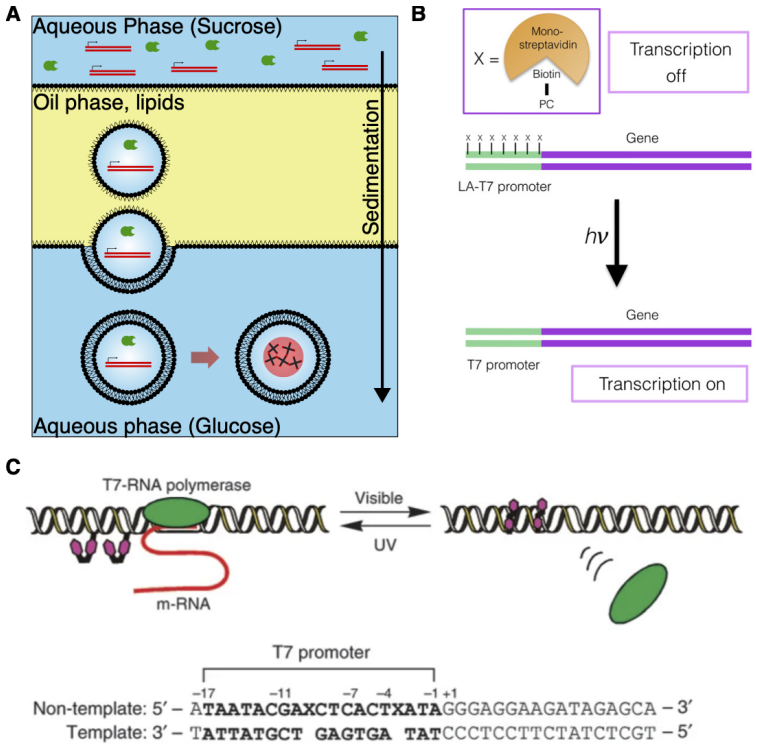


Fig. 7.1 **Forming RNA MLOs in more realistic and biomimetic artificial cells with temporal transcriptional control. A**: Schematic representation of RNA organelle formation in GUVs through inverted emulsion method. Sketch is largely based on the one in ref. [488]. **B**: Photo-switchable transcription system exploiting T7 promoters featuring modified thymine residues conjugated to photo-cleavable (PC) linkers terminating in biotin moieties. In the presence of streptavidin, the T7 promoter results inaccessible for the T7 RNAP. Upon UV exposure, the PC linkers are cleaved and transcription can proceed. Reproduced from ref. [486] in accordance with CC BY-NC 4.0 license. **C**: Photo-switchable transcription system relying on T7 promoters containing azobenzene moieties. Upon UV exposure, the azobenzene moieties undergo *trans*→*cis* photoisomerisation, turning the T7 promoter single-stranded and impeding transcription. Exposure to visible light induces the opposite isomerisation, and restores the double-stranded conformation of the T7 promoter, allowing transcription to start. Reproduced from ref. [489] with permission from Springer Nature Limited.

7.2.2 Photo-activated transcription in synthetic cells

Phase transfer in GUVs takes considerably more time compared to encapsulation in emulsion droplets. Adding DNA templates prior to phase transfer would therefore lead to a substantial delay in imaging and characterisation compared to the start of transcription, which, albeit slower, still takes place at room temperature. This issue directly links with the overarching goal of explicitly controlling, both spatially and temporally, the formation of such synthetic RNA organelles, which is not possible in the implementation discussed in Chapter 5. While it is possible to slow down the transcription and self-assembly rate of the RNA organelles by reducing the corresponding DNA template concentration, this cannot be precisely switched on or off.

However, it is relatively easy to envision at least one solution to both mentioned challenges, namely switchable transcription. For instance, one could resort to a light-activated strategy similar to the one reported in ref. [486] and sketched in Fig. 7.1B, where some thymine residues in the T7 promoter of the DNA templates are functionalised with biotin moieties through photo-cleavable linkers (PC in Fig. 7.1B). In the presence of streptavidin, the T7 promoter is thus blocked and transcription is dramatically hindered, if not completely repressed. This should allow for encapsulation in GUVs through phase transfer with no significant RNA synthesis. Upon UV exposure, though, such linkers are cleaved and the blockade is lifted, exposing the T7 promoter region and enabling transcription, thus leading to the formation of the desired RNA organelles with precise temporal control. This approach is particularly promising, as a slight variant relying on photocleavable 2-nitrobenzyl groups has recently been used to successfully implement light-controlled synthetic cell-living cell communication exploiting quorum-sensing in bacteria [490].

Similar end results could also be obtained through different strategies, for instance by including light-dependent azobenzene moieties within the T7 promoter region, thus allowing to switch the T7 promoter conformation between single- (*cis*, off) and double-stranded (*trans*, on), as demonstrated in ref. [489] and schematised in Fig. 7.1C, thus impeding or permitting RNA synthesis, respectively.

7.2.3 Multiple protein/enzyme colocalisation for enzymatic cascades

Building on the results from Chapter 6, one can easily envision extending the platform to capture multiple proteins, particularly enzymes participating in the same enzymatic cascade [52], within the same RNA organelle. This could be achieved, for example, through simultaneous transcription of multiple protein-binding RNA aptamers, either already reported in the literature or discovered *via* SELEX, sharing the same capture domain which enables hybridisation to the complementary one in the RNA nanostars forming the organelles. This should lead to the partitioning of multiple proteins/enzymes within the same RNA organelles, provided the molecular weights are comparable. On one hand, this would be

amenable because enzyme colocalisation has been showed to greatly enhance the rate and overall yield of cascade reactions [351]. On the other hand, the three-dimensional nature of the (programmably) liquid- or hydrogel-like matrix of the RNA organelles would also enable regulated transport of substrate and product species, improving the overall substrate to product conversion efficiency as compared to surface-immobilised enzymatic cascades [491]. This in turn would make the designed synthetic cells promising candidates for applications in biosynthesis and biocatalysis [492].
Moreover, programmed colocalisation of specific enzyme-substrate pairs within hierarchically structured compartments would be highly desirable for metabolic engineering of the synthetic cell mimics, for instance to produce energy (in the form of ATP) *via* anaerobic glycolysis [493], or to shift the metabolic flux from one pathway to another [30]. To these aims, one could also exploit reaction-diffusion patterning, similar to that recently demonstrated by Leathers *et al.* in DNA-based condensates [154], combined with switchable transcription systems, as discussed above, to prescribe the spatial arrangement of enzymes and substrates, for instance through RNA aptamers differing in size and transcribed at different times. This would produce synthetic compartments capable of mimicking the enzymatic reaction directionality of naturally occurring biomolecular condensates [32], such as that observed in the ensemble of reactions leading to ribosome biogenesis in nucleoli [494, 28].

7.2.4 Synthetic MLOs to influence the behaviour of living agents

Finally, while this thesis has focused on bottom-up approaches, the historical goal of synthetic biology – also called top-down synthetic biology [495] – has been to reprogram living systems to enable behaviours and functionalities beyond their natural potential. To these aims, one could envision integrating the double-stranded DNA template fragments encoding for RNA nanostars discussed in Chapters 4-6 within circular plasmids for expression in bacteria, enabling the formation of synthetic, bottom-up self-assembled organelles with programmable shape and functionality within living agents, including bacteria themselves. This is particularly interesting for at least three reasons: i) organisation *via* MLOs emerging through LLPS has recently been recognised to play a major role in bacteria [37, 496, 497]; ii) MLOs have been found to increase bacterial fitness [497, 498] and modulate cellular noise [461, 39] depending on the phase of the cell cycle; and iii) the idea of integrating bottom-up self-assembled synthetic constructs with living cells to influence the behaviour of the latter is becoming increasingly popular [48, 499, 359].
While some examples of synthetic organelles in bacteria have already been reported in the literature [39, 500], these rely on biologically derived tandem repeat RNAs, such as 47× rCAG, known to undergo LLPS. This in turn removes the possibility of programmably tuning the biophysical properties of the resulting assemblies. The discussed RNA nanostar platform, instead, could enable precise control not only over orthogonality and addressability, but also hydrogel- or liquid-like behaviour, mixing/demixing and morphology of the resulting

organelles. For this reason, I set out to try a proof of concept implementation of this idea. The work presented below has been carried out entirely by Sabrina Pia Nuccio (Di Antonio group), while I performed the imaging.

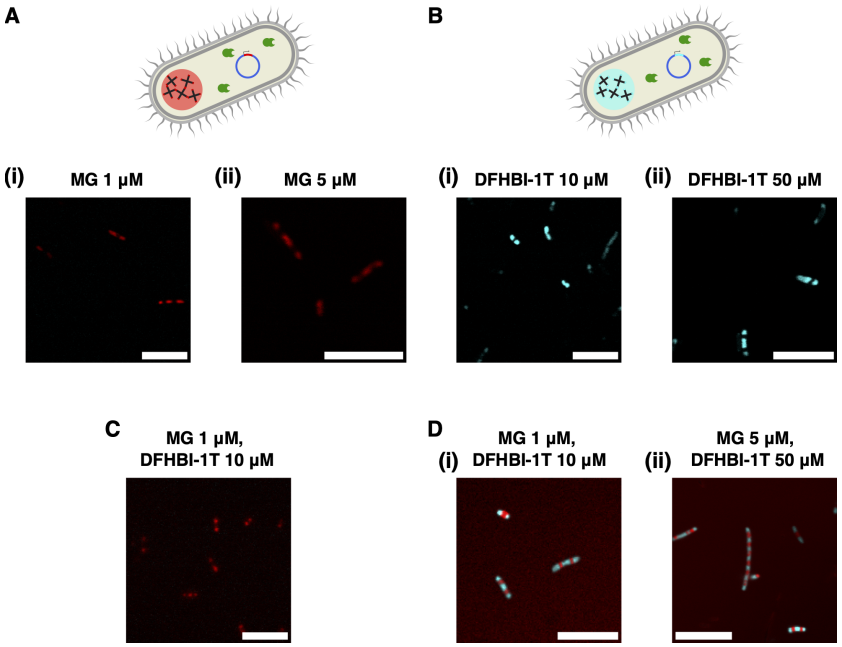


Fig. 7.2 **Synthetic RNA MLOs in bacteria. A**: Confocal micrographs representing BL21 (DE3) competent cells transformed with a plasmid including the DNA fragment coding for RNA nanostar A, indicated by the schematic above, after 2 h incubation with **(i)** 1 μM and **(ii)** 5 μM Malachite Green oxalate, respectively. **B**: Same as **A** for BL21 (DE3) competent cells transformed with a plasmid including the DNA fragment coding for RNA nanostar B, indicated by the schematic above, after 2 h incubation with **(i)** 10 μM and **(ii)** 50 μM DFHBI-1T, respectively. **C**: Same as **A** in the presence of both 1 μM Malachite Green oxalate and 10 μM DFHBI-1T. **D**: Same as **B** in the presence of **(i)** 1 μM Malachite Green oxalate and 10 μM DFHBI-1T, or **(ii)** 5 μM Malachite Green oxalate and 50 μM DFHBI-1T. Samples were imaged by sandwiching liquid droplets between two glass coverslips. Micrographs in **A-C** are pristine, while those in **D** have been contrast-enhanced for ease of visualisation. All scale bars are 10 μm.

The double-stranded template fragments A-T and B-T introduced in Chapter 4, coding for nanostars A and B, respectively, were modified by adding a transcription terminator, and cloned into an in-house plasmid already containing the T7 promoter. The two resulting plasmids (one for A-T, one for B-T) were transformed in *E. coli* (DH5α). Sanger sequencing of the plasmids obtained from bacterial lysis confirmed the correct integration of the desired fragments. Plasmids were then transformed *via* heat shock in BL21 (DE3) competent *E. coli*, which were then incubated overnight. The next day, one colony per construct was picked and further incubated overnight. After reaching OD_{600} = 0.6, bacteria were incubated with Malachite Green oxalate and DFHBI-1T for 2 h prior to imaging. Here, DFHBI-1T was chosen over regular DFHBI for its stronger fluorescence emission properties [257]. Some preliminary results are presented in Fig. 7.2. Bacteria transformed with plasmids coding for RNA nanostars A and B clearly exhibit bright fluorescent spots in Malachite Green and DFHBI channels, which resemble polar condensates, as showed in Fig. 7.2A-B, respectively. Incubation with both Malachite Green oxalate and DFHBI-1T of bacteria transformed with

plasmid coding for nanostar A still produces the expected behaviour in Fig. 7.2C. However, this is not the case for those transformed with plasmid coding for RNA nanostar B, which, as depicted in micrographs in Fig. 7.2D, clearly feature alternating red and cyan bright spots, despite the plasmid only coding for nanostar B (cyan) and not for nanostar A (red). While further experiments, including controls with non sticky nanostars, are needed to elucidate this phenomenon, this might be due to Malachite Green staining something we have not accounted for – it is commonly used, for instance, for bacterial endospore staining [501]. It is also worth noting that, in micrographs in Fig. 7.2D, bacteria seem unusually elongated, possibly hinting at slower or arrested division. This could in turn be related to the RNA transcripts themselves, due to bacterial burden induced by transcription of the related exogenous genes [502, 7], or to cytotoxic effects of the Malachite Green dye, a known antimicrobial [503] – despite working at concentrations below the cytotoxicity limit [252]. Finally, the presence of bright fluorescence spots within the bacteria is not informative of the nature of such assemblies nor does it prove the correct transcription and folding of the modified nanostars with attached transcription terminator domains. Nonetheless, the results in Fig. 7.2 serve as a stepping stone, clearly pointing at the feasibility and potential of this approach.

References

[1] Fankang Meng and Tom Ellis. The second decade of synthetic biology: 2010-2020. *Nat. Commun.*, 11(1):5174, October 2020.

[2] Steven A Benner and A Michael Sismour. Synthetic biology. *Nat. Rev. Genet.*, 6(7):533–543, July 2005.

[3] Jim Haseloff and Jim Ajioka. Synthetic biology: history, challenges and prospects. *J. R. Soc. Interface*, 6 Suppl 4(Suppl 4):S389–91, August 2009.

[4] Daniel G Gibson, John I Glass, Carole Lartigue, Vladimir N Noskov, Ray-Yuan Chuang, Mikkel A Algire, Gwynedd A Benders, Michael G Montague, Li Ma, Monzia M Moodie, Chuck Merryman, Sanjay Vashee, Radha Krishnakumar, Nacyra Assad-Garcia, Cynthia Andrews-Pfannkoch, Evgeniya A Denisova, Lei Young, Zhi-Qing Qi, Thomas H Segall-Shapiro, Christopher H Calvey, Prashanth P Parmar, Clyde A Hutchison, 3rd, Hamilton O Smith, and J Craig Venter. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329(5987):52–56, July 2010.

[5] Seth L Shipman, Jeff Nivala, Jeffrey D Macklis, and George M Church. CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria. *Nature*, 547(7663):345–349, July 2017.

[6] Cheng Kai Lim, Jing Wui Yeoh, Aurelius Andrew Kunartama, Wen Shan Yew, and Chueh Loo Poh. A biological camera that captures and stores images directly into DNA. *Nat. Commun.*, 14(1):3921, July 2023.

[7] Francesca Ceroni, Rhys Algar, Guy-Bart Stan, and Tom Ellis. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods*, 12(5):415–418, May 2015.

[8] Timothy Frei, Federica Cella, Fabiana Tedeschi, Joaquín Gutiérrez, Guy-Bart Stan, Mustafa Khammash, and Velia Siciliano. Characterization and mitigation of gene expression burden in mammalian cells. *Nat. Commun.*, 11(1):4641, September 2020.

[9] Yuval Elani. Construction of membrane-bound artificial cells using microfluidics: a new frontier in bottom-up synthetic biology. *Biochem. Soc. Trans.*, 44(3):723–730, June 2016.

[10] D L Hartley and J F Kane. Properties of inclusion bodies from recombinant escherichia coli. *Biochem. Soc. Trans.*, 16(2):101–102, April 1988.

[11] Petra Schwille and Stefan Diez. Synthetic biology of minimal systems. *Crit. Rev. Biochem. Mol. Biol.*, 44(4):223–242, 2009.

[12] Haiyang Jia and Petra Schwille. Bottom-up synthetic biology: reconstitution in space and time. *Curr. Opin. Biotechnol.*, 60:179–187, December 2019.

[13] Petra Schwille. Bottom-up synthetic biology: engineering in a tinkerer's world. *Science*, 333(6047):1252–1254, September 2011.

[14] Rita Sachse, Srujan K Dondapati, Susanne F Fenz, Thomas Schmidt, and Stefan Kubick. Membrane protein synthesis in cell-free systems: from bio-mimetic systems to bio-membranes. *FEBS Lett.*, 588(17):2774–2781, August 2014.

[15] Bruno Pitard, Peter Richard, Mireia Duñach, Guy Girault, and Jean-Louis Rigaud. ATP synthesis by the F0F1 ATP synthase from thermophilic bacillus PS3 reconstituted into liposomes with bacteriorhodopsin. 1. factors defining the optimal reconstitution of ATP synthases with bacteriorhodopsin. *Eur. J. Biochem.*, 235(3):769–778, February 1996.

[16] Keel Yong Lee, Sung-Jin Park, Keon Ah Lee, Se-Hwan Kim, Heeyeon Kim, Yasmine Meroz, L Mahadevan, Kwang-Hwan Jung, Tae Kyu Ahn, Kevin Kit Parker, and Kwanwoo Shin. Photosynthetic artificial organelles sustain and control ATP-dependent reactions in a protocellular system. *Nat. Biotechnol.*, 36(6):530–535, July 2018.

[17] Alex Joesaar, Shuo Yang, Bas Bögels, Ardjan van der Linden, Pascal Pieters, B V V S Pavan Kumar, Neil Dalchau, Andrew Phillips, Stephen Mann, and Tom F A de Greef. DNA-based communication in populations of synthetic protocells. *Nat. Nanotechnol.*, 14(4):369–378, April 2019.

[18] James W Hindley, Yuval Elani, Catriona M McGilvery, Simak Ali, Charlotte L Bevan, Robert V Law, and Oscar Ces. Light-triggered enzymatic reactions in nested vesicle reactors. *Nat. Commun.*, 9(1):1093, March 2018.

[19] Christine Schmitt, Anna H Lippert, Navid Bonakdar, Vahid Sandoghdar, and Lars M Voll. Compartmentalization and transport in synthetic vesicles. *Front Bioeng Biotechnol*, 4:19, February 2016.

[20] Kevin Jahnke, Marian Weiss, Cornelia Weber, Ilia Platzman, Kerstin Göpfrich, and Joachim P Spatz. Engineering Light-Responsive contractile actomyosin networks with DNA nanotechnology. *Adv Biosyst*, 4(9):e2000102, September 2020.

[21] Siddharth Agarwal, Melissa A Klocke, Passa E Pungchai, and Elisa Franco. Dynamic self-assembly of compartmentalized DNA nanotubes. *Nat. Commun.*, 12(1):3557, June 2021.

[22] J Craig Blain and Jack W Szostak. Progress toward synthetic cells. *Annu. Rev. Biochem.*, 83:615–640, March 2014.

[23] Can Xu, Shuo Hu, and Xiaoyuan Chen. Artificial cells: from basic science to applications. *Mater. Today*, 19(9):516–532, November 2016.

[24] Bastiaan C Buddingh' and Jan C M van Hest. Artificial cells: Synthetic compartments with life-like functionality and adaptivity. *Acc. Chem. Res.*, 50(4):769–777, April 2017.

[25] Wakana Sato, Tomasz Zajkowski, Felix Moser, and Katarzyna P Adamala. Synthetic cells in biomedical applications. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.*, 14(2):e1761, March 2022.

[26] Yuval Mulla, Anders Aufderhorst-Roberts, and Gijsje H Koenderink. Shaping up synthetic cells. *Phys. Biol.*, 15(4):041001, April 2018.

[27] Chuankai Zhou. The molecular and functional interaction between Membrane-Bound organelles and Membrane-Less condensates. *Front Cell Dev Biol*, 10:896305, April 2022.

[28] Tetsuro Hirose, Kensuke Ninomiya, Shinichi Nakagawa, and Tomohiro Yamazaki. A guide to membraneless organelles and their various roles in gene regulation. *Nat. Rev. Mol. Cell Biol.*, 24(4):288–304, April 2023.

[29] Anthony A Hyman, Christoph A Weber, and Frank Jülicher. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.*, 30:39–58, 2014.

[30] Salman F Banani, Hyun O Lee, Anthony A Hyman, and Michael K Rosen. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.*, 18(5):285–298, May 2017.

[31] Yifan Dai, Mina Farag, Dongheon Lee, Xiangze Zeng, Kyeri Kim, Hye-In Son, Xiao Guo, Jonathan Su, Nikhil Peterson, Javid Mohammed, Max Ney, Daniel Mark Shapiro, Rohit V Pappu, Ashutosh Chilkoti, and Lingchong You. Programmable synthetic biomolecular condensates for cellular control. *Nat. Chem. Biol.*, 19(4):518–528, April 2023.

[32] Yifan Dai, Lingchong You, and Ashutosh Chilkoti. Engineering synthetic biomolecular condensates. *Nature Reviews Bioengineering*, pages 1–15, April 2023.

[33] Takeshi Chujo, Tomohiro Yamazaki, and Tetsuro Hirose. Architectural RNAs (arcRNAs): A class of long noncoding RNAs that function as the scaffold of nuclear bodies. *Biochim. Biophys. Acta*, 1859(1):139–146, January 2016.

[34] Tomohiro Yamazaki, Shinichi Nakagawa, and Tetsuro Hirose. Architectural RNAs for membraneless nuclear body formation. *Cold Spring Harb. Symp. Quant. Biol.*, 84:227–237, 2019.

[35] Nicolette G Moreau, Nicolas Martin, Pierangelo Gobbo, T-Y Dora Tang, and Stephen Mann. Spontaneous membrane-less multi-compartmentalization via aqueous two-phase separation in complex coacervate micro-droplets. *Chem. Commun.*, 56(84):12717–12720, October 2020.

[36] N Amy Yewdall, Alain A M André, Tiemei Lu, and Evan Spruijt. Coacervates as models of membraneless organelles. *Curr. Opin. Colloid Interface Sci.*, 52:101416, April 2021.

[37] Anne-Marie Ladouceur, Baljyot Singh Parmar, Stefan Biedzinski, James Wall, S Graydon Tope, David Cohn, Albright Kim, Nicolas Soubry, Rodrigo Reyes-Lamothe, and Stephanie C Weber. Clusters of bacterial RNA polymerase are biomolecular condensates that assemble through liquid-liquid phase separation. *Proc. Natl. Acad. Sci. U. S. A.*, 117(31):18540–18549, August 2020.

[38] Vivian Yeong, Emily G Werth, Lewis M Brown, and Allie C Obermeyer. Formation of biomolecular condensates in bacteria by tuning protein electrostatics. *ACS Cent Sci*, 6(12):2301–2310, December 2020.

[39] Haotian Guo, Joseph C Ryan, Xiaohu Song, Adeline Mallet, Mengmeng Zhang, Victor Pabst, Antoine L Decrulle, Paulina Ejsmont, Edwin H Wintermute, and Ariel B Lindner. Spatial engineering of e. coli with addressable phase-separated RNAs. *Cell*, 185(20):3823–3837.e23, September 2022.

[40] Tatiana Trantidou, Mark Friddin, Yuval Elani, Nicholas J Brooks, Robert V Law, John M Seddon, and Oscar Ces. Engineering compartmentalized biomimetic micro- and nanocontainers. *ACS Nano*, 11(7):6549–6565, July 2017.

[41] Yuval Elani, Antony Gee, Robert V Law, and Oscar Ces. Engineering multi-compartment vesicle networks. *Chem. Sci.*, 4(8):3332–3338, 2013.

[42] Emeline Rideau, Rumiana Dimova, Petra Schwille, Frederik R Wurm, and Katharina Landfester. Liposomes and polymersomes: a comparative review towards cell mimicking. *Chem. Soc. Rev.*, 47(23):8572–8610, November 2018.

[43] Avik Samanta, Valerio Sabatino, Thomas R Ward, and Andreas Walther. Functional and morphological adaptation in DNA protocells via signal processing prompted by artificial metalloenzymes. *Nat. Nanotechnol.*, 15(11):914–921, November 2020.

[44] Caterina LoPresti, Hannah Lomas, Marzia Massignani, Thomas Smart, and Giuseppe Battaglia. Polymersomes: nature inspired nanometer sized compartments. *J. Mater. Chem.*, 19(22):3576–3590, 2009.

[45] Sophie Pautot, Barbara J Frisken, and D A Weitz. Production of unilamellar vesicles using an inverted emulsion. *Langmuir*, 19(7):2870–2879, April 2003.

[46] Pashiini Supramaniam, Zibo Wang, Stelios Chatzimichail, Christopher Parperis, Aditi Kumar, Vanessa Ho, Oscar Ces, and Ali Salehi-Reyhani. Measuring encapsulation efficiency in Cell-Mimicking giant unilamellar vesicles. *ACS Synth. Biol.*, 12(4):1227–1238, April 2023.

[47] Lukas Aufinger and Friedrich C Simmel. Artificial Gel-Based organelles for spatial organization of Cell-Free gene expression reactions. *Angew. Chem. Int. Ed Engl.*, 57(52):17245–17248, December 2018.

[48] Yuval Elani. Interfacing living and synthetic cells as an emerging frontier in synthetic biology. *Angew. Chem. Int. Ed Engl.*, 60(11):5602–5611, March 2021.

[49] Matthew E Allen, James W Hindley, Divesh K Baxani, Oscar Ces, and Yuval Elani. Hydrogels as functional components in artificial cell systems. *Nat Rev Chem*, 6(8):562–578, August 2022.

[50] Wiggert J Altenburg, N Amy Yewdall, Daan F M Vervoort, Marleen H M E van Stevendaal, Alexander F Mason, and Jan C M van Hest. Programmed spatial organization of biomacromolecules into discrete, coacervate-based protocells. *Nat. Commun.*, 11(1):6282, December 2020.

[51] Joseph R Simon, Nick J Carroll, Michael Rubinstein, Ashutosh Chilkoti, and Gabriel P López. Programming molecular self-assembly of intrinsically disordered proteins containing sequences of low complexity. *Nat. Chem.*, 9(6):509–515, June 2017.

[52] Renee L Hastings and Steven Boeynaems. Designer condensates: A toolkit for the biomolecular architect. *J. Mol. Biol.*, 433(12):166837, June 2021.

[53] Kerstin Göpfrich, Ilia Platzman, and Joachim P Spatz. Mastering complexity: Towards bottom-up construction of multifunctional eukaryotic synthetic cells. *Trends Biotechnol.*, 36(9):938–951, September 2018.

[54] Jongmin Kim and Elisa Franco. RNA nanotechnology in synthetic biology. *Curr. Opin. Biotechnol.*, 63:135–141, June 2020.

[55] Roger Rubio-Sánchez, Simone Eizagirre Barker, Michal Walczak, Pietro Cicuta, and Lorenzo Di Michele. A modular, dynamic, DNA-Based platform for regulating cargo distribution and transport between lipid domains. *Nano Lett.*, 21(7):2800–2808, April 2021.

[56] Claire Ashworth. DNA nanotechnology: Building big with DNA bricks. *Nature Reviews Materials*, 3(1):1–1, January 2018.

[57] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of the Cell, 6th edition*. Garland Science, Taylor and Francis Group, New York, 2015.

[58] Jaimie Marie Stewart and Elisa Franco. Self-assembly of large RNA structures: learning from DNA nanotechnology. *DNA and RNA Nanotechnology*, 2(1):23–35, February 2016.

[59] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, T N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic Acids Res.*, 28(1):235–242, January 2000.

[60] H R Drew, R M Wing, T Takano, C Broka, S Tanaka, K Itakura, and R E Dickerson. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U. S. A.*, 78(4):2179–2183, April 1981.

[61] Yoshiyuki Tanaka, Satoshi Fujii, Hidekazu Hiroaki, Takeshi Sakata, Toshiki Tanaka, Seiichi Uesugi, Ken-Ichi Tomita, and Yoshimasa Kyogoku. A'-form RNA double helix in the single crystal structure of r(UGAGCUUCGGCUC). *Nucleic Acids Res.*, 27(4):949–955, February 1999.

[62] James D Watson and Francis H Crick. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–738, April 1953.

[63] Rosalind E Franklin and Raymond G Gosling. Molecular configuration in sodium thymonucleate. *Nature*, 171(4356):740–741, April 1953.

[64] Christopher A Hunter and Jeremy K M Sanders. The nature of .pi.-.pi. interactions. *J. Am. Chem. Soc.*, 112(14):5525–5534, July 1990.

[65] Ioannis Vakonakis and Andy C LiWang. N1...N3 hydrogen bonds of A:U base pairs of RNA are stronger than those of A:T base pairs of DNA. *J. Am. Chem. Soc.*, 126(18):5688–5689, May 2004.

[66] Marcel Swart, Célia Fonseca Guerra, and F Matthias Bickelhaupt. Hydrogen bonds of RNA are stronger than those of DNA, but NMR monitors only presence of methyl substituent in uracil/thymine. *J. Am. Chem. Soc.*, 126(51):16718–16719, December 2004.

[67] Reid F Brown, Casey T Andrews, and Adrian H Elcock. 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(5):2315–2328, May 2015.

[68] Nadrian C Seeman. An overview of structural DNA nanotechnology. *Mol. Biotechnol.*, 37(3):246–257, November 2007.

[69] Gerald S Manning. The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force. *Biophys. J.*, 91(10):3607–3616, November 2006.

[70] Changbong Hyeon, Ruxandra I Dima, and D Thirumalai. Size, shape, and flexibility of RNA structures. *J. Chem. Phys.*, 125(19):194905, November 2006.

[71] Huimin Chen, Steve P Meisburger, Suzette A Pabit, Julie L Sutton, Watt W Webb, and Lois Pollack. Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc. Natl. Acad. Sci. U. S. A.*, 109(3):799–804, January 2012.

[72] Kotaro Hayashi, Hiroyuki Chaya, Shigeto Fukushima, Sumiyo Watanabe, Hiroyasu Takemoto, Kensuke Osada, Nobuhiro Nishiyama, Kanjiro Miyata, and Kazunori Kataoka. Influence of RNA strand rigidity on polyion complex formation with block catiomers. *Macromol. Rapid Commun.*, 37(6):486–493, March 2016.

[73] J A Abels, F Moreno-Herrero, T van der Heijden, C Dekker, and N H Dekker. Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys. J.*, 88(4):2737–2744, April 2005.

[74] John SantaLucia, Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.*, 95(4):1460–1465, February 1998.

[75] John SantaLucia, Jr and Donald Hicks. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33:415–440, 2004.

[76] John SantaLucia, Jr, Hatim T Allawi, and P Ananda Seneviratne. Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, 35(11):3555–3562, March 1996.

[77] Martin J Serra and Douglas H Turner. [11] predicting thermodynamic properties of RNA. In *Methods in Enzymology*, volume 259, pages 242–261. Academic Press, January 1995.

[78] Izabela Ferreira, Elizabeth A Jolley, Brent M Znosko, and Gerald Weber. Replacing salt correction factors with optimized RNA nearest-neighbour enthalpy and entropy parameters. *Chem. Phys.*, 521:69–76, May 2019.

[79] Vivianne Basílio Barbosa, Erik de Oliveira Martins, and Gerald Weber. Nearest-neighbour parameters optimized for melting temperature prediction of DNA/RNA hybrids at high and low salt concentrations. *Biophys. Chem.*, 251:106189, August 2019.

[80] Dipanwita Banerjee, Hisae Tateishi-Karimata, Tatsuya Ohyama, Saptarshi Ghosh, Tamaki Endoh, Shuntaro Takahashi, and Naoki Sugimoto. Improved nearest-neighbor parameters for the stability of RNA/DNA hybrids under a physiological condition. *Nucleic Acids Res.*, 48(21):12042–12054, December 2020.

[81] Saptarshi Ghosh, Shuntaro Takahashi, Dipanwita Banerjee, Tatsuya Ohyama, Tamaki Endoh, Hisae Tateishi-Karimata, and Naoki Sugimoto. Nearest-neighbor parameters for the prediction of RNA duplex stability in diverse in vitro and cellular-like crowding conditions. *Nucleic Acids Res.*, 51(9):4101–4111, May 2023.

[82] Karst Hoogsteen. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.*, 12(10):822–823, October 1959.

[83] Evgenia N Nikolova, Huiqing Zhou, Federico L Gottardo, Heidi S Alvey, Isaac J Kimsey, and Hashim M Al-Hashimi. A historical account of hoogsteen base-pairs in duplex DNA. *Biopolymers*, 99(12):955–968, December 2013.

[84] Pascale Hazel, Julian Huppert, Shankar Balasubramanian, and Stephen Neidle. Loop-length-dependent folding of G-quadruplexes. *J. Am. Chem. Soc.*, 126(50):16405–16415, December 2004.

[85] Sarah Burge, Gary N Parkinson, Pascale Hazel, Alan K Todd, and Stephen Neidle. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, 34(19):5402–5415, September 2006.

[86] Fang-Yuan Teng, Zong-Zhe Jiang, Man Guo, Xiao-Zhen Tan, Feng Chen, Xu-Guang Xi, and Yong Xu. G-quadruplex DNA: a novel target for drug design. *Cell. Mol. Life Sci.*, 78(19-20):6557–6583, October 2021.

[87] Yong Wang and Dinshaw J Patel. Solution structure of a parallel-stranded G-quadruplex DNA. *J. Mol. Biol.*, 234(4):1171–1183, December 1993.

[88] Yuanchen Dong, Zhongqiang Yang, and Dongsheng Liu. DNA nanotechnology based on i-motif structures. *Acc. Chem. Res.*, 47(6):1853–1860, June 2014.

[89] Andrea Idili and Francesco Ricci. Design and characterization of pH-Triggered DNA nanoswitches and nanodevices based on DNA triplex structures. In Giampaolo Zuccheri, editor, *DNA Nanotechnology: Methods and Protocols*, pages 79–100. Springer New York, New York, NY, 2018.

[90] Jiande Gu, Jerzy Leszczynski, and Manju Bansal. A new insight into the structure and stability of hoogsteen hydrogen-bonded g-tetrad: an ab initio SCF study. *Chem. Phys. Lett.*, 311(3):209–214, September 1999.

[91] Debmalya Bhattacharyya, Gayan Mirihana Arachchilage, and Soumitra Basu. Metal cations in G-Quadruplex folding and stability. *Front Chem*, 4:38, September 2016.

[92] Lokendra Poudel, Nicole F Steinmetz, Roger H French, V Adrian Parsegian, Rudolf Podgornik, and Wai-Yim Ching. Implication of the solvent effect, metal ions and topology in the electronic structure and hydrogen bonding of human telomeric G-quadruplex DNA. *Phys. Chem. Chem. Phys.*, 18(31):21573–21585, August 2016.

[93] Xin Tong, Lu Ga, Chaolu Eerdun, Ruiguo Zhao, and Jun Ai. Simple monovalent metal ion logical order to regulate the secondary conformation of G-Quadruplex. *ACS Omega*, 7(43):39224–39233, November 2022.

[94] Ivan Smirnov and Richard H Shafer. Lead is unusually effective in sequence-specific folding of DNA. *J. Mol. Biol.*, 296(1):1–5, February 2000.

[95] Kalle Gehring, Jean-Louis Leroy, and Maurice Guéron. A tetrameric DNA structure with protonated cytosine.cytosine base pairs. *Nature*, 363(6429):561–565, June 1993.

[96] K Snoussi, S Nonin-Lecomte, and J L Leroy. The RNA i-motif. *J. Mol. Biol.*, 309(1):139–153, May 2001.

[97] Hala Abou Assi, Miguel Garavís, Carlos González, and Masad J Damha. i-motif DNA: structural features and significance to cell biology. *Nucleic Acids Res.*, 46(16):8038–8056, September 2018.

[98] Jamie M Dettler, Robert Buscaglia, Jingjing Cui, Derek Cashman, Meredith Blynn, and Edwin A Lewis. Biophysical characterization of an ensemble of intramolecular i-motifs formed by the human c-MYC NHE III1 P1 promoter mutant sequence. *Biophys. J.*, 99(2):561–567, July 2010.

[99] Daekyu Sun and Laurence H Hurley. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-myc promoter: implications for drug targeting and control of gene expression. *J. Med. Chem.*, 52(9):2863–2874, May 2009.

[100] G Felsenfeld, David R Davies, and Alexander Rich. Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc.*, 79(8):2023–2024, April 1957.

[101] Yuwei Hu, Alessandro Cecconello, Andrea Idili, Francesco Ricci, and Itamar Willner. Triplex DNA nanostructures: From basic properties to applications. *Angew. Chem. Int. Ed Engl.*, 56(48):15210–15233, November 2017.

[102] Alessandro Cecconello, Massimiliano Magro, Fabio Vianello, and Friedrich C Simmel. Rational design of hybrid DNA-RNA triplex structures as modulators of transcriptional activity in vitro. *Nucleic Acids Res.*, 50(22):13172–13182, December 2022.

[103] Francis H Crick. On protein synthesis. *Symp. Soc. Exp. Biol.*, 12:138–163, 1958.

[104] Francis H Crick. Central dogma of molecular biology. *Nature*, 227(5258):561–563, August 1970.

[105] Matthew Cobb. 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol.*, 15(9):e2003243, September 2017.

[106] Susan Payne. Introduction to RNA viruses. *Viruses*, page 97, 2017.

[107] Sephra Rampersad and Paula Tennant. Chapter 3 - replication and expression strategies of viruses. In Paula Tennant, Gustavo Fermin, and Jerome E Foster, editors, *Viruses*, pages 55–82. Academic Press, January 2018.

[108] ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414):57–74, September 2012.

[109] Mahlon B Hoagland, Mary Louise Stephenson, Jesse F Scott, Liselotte I Hecht, and Paul C Zamecnik. A soluble ribonucleic acid intermediate in protein synthesis. *J. Biol. Chem.*, 231(1):241–257, March 1958.

[110] Patrick O'Donoghue, Jiqiang Ling, and Dieter Söll. Transfer RNA function and evolution. *RNA Biol.*, 15(4-5):423–426, 2018.

[111] Maciej Szymański, Mirosława Z Barciszewska, Volker A Erdmann, and Jan Barciszewski. 5 S rRNA: structure and interactions. *Biochem. J*, 371(Pt 3):641–651, May 2003.

[112] R Brimacombe and W Stiege. Structure and function of ribosomal RNA. *Biochem. J*, 229(1):1–17, July 1985.

[113] Jacek Krol, Inga Loedige, and Witold Filipowicz. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.*, 11(9):597–610, September 2010.

[114] Hassan Dana, Ghanbar Mahmoodi Chalbatani, Habibollah Mahmoodzadeh, Rezvan Karimloo, Omid Rezaiean, Amirreza Moradzadeh, Narges Mehmandoost, Fateme Moazzen, Ali Mazraeh, Vahid Marmari, Mohammad Ebrahimi, Mohammad Menati Rashno, Saeid Jan Abadi, and Elahe Gharagouzlo. Molecular mechanisms and biological functions of siRNA. *Int. J. Biomed. Sci.*, 13(2):48–57, June 2017.

[115] Luisa Statello, Chun-Jie Guo, Ling-Ling Chen, and Maite Huarte. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.*, 22(2):96–118, February 2021.

[116] Ronald R Breaker. Riboswitches and translation control. *Cold Spring Harb. Perspect. Biol.*, 10(11), November 2018.

[117] N Kyle Tanner. Ribozymes: the characteristics and properties of catalytic RNAs. *FEMS Microbiol. Rev.*, 23(3):257–275, June 1999.

[118] Nils G Walter and David R Engelke. Ribozymes: catalytic RNAs that cut things, make things, and do odd and useful jobs. *Biologist*, 49(5):199–203, October 2002.

[119] M Soller. Pre-messenger RNA processing and its regulation: a genomic perspective. *Cell. Mol. Life Sci.*, 63(7-8):796–819, April 2006.

[120] Samuel E Butcher and Anna Marie Pyle. The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc. Chem. Res.*, 44(12):1302–1311, December 2011.

[121] Haopeng Yu, Yiman Qi, and Yiliang Ding. Deep learning in RNA structure studies. *Front Mol Biosci*, 9:869601, May 2022.

[122] Jun Li and Shi-Jie Chen. RNAJP: enhanced RNA 3D structure predictions with non-canonical interactions and global topology sampling. *Nucleic Acids Res.*, 51(7):3341–3356, April 2023.

[123] Kengo Sato, Manato Akiyama, and Yasubumi Sakakibara. RNA secondary structure prediction using deep learning with thermodynamic integration. *Nat. Commun.*, 12(1):941, February 2021.

[124] Denise Liano, Ludovica Monti, Souroprobho Chowdhury, Federica Raguseo, and Marco Di Antonio. Long-range DNA interactions: inter-molecular G-quadruplexes and their potential biological relevance. *Chem. Commun.*, 58(92):12753–12762, November 2022.

[125] R V Reshetnikov, A M Kopylov, and A V Golovin. Classification of G-quadruplex DNA on the basis of the quadruplex twist angle and planarity of g-quartets. *Acta Naturae*, 2(4):72–81, October 2010.

[126] Jenna Robinson, Federica Raguseo, Sabrina Pia Nuccio, Denise Liano, and Marco Di Antonio. DNA G-quadruplex structures: more than simple roadblocks to transcription? *Nucleic Acids Res.*, 49(15):8419–8431, September 2021.

[127] Yuwei Chen and Danzhou Yang. Sequence, stability, and structure of G-quadruplexes and their interactions with drugs. *Curr. Protoc. Nucleic Acid Chem.*, Chapter 17:Unit17.5, September 2012.

[128] Tracy A Brooks, Samantha Kendrick, and Laurence Hurley. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.*, 277(17):3459–3469, September 2010.

[129] Louis J Maher, III, Barbara Wold, and Peter B Dervan. Inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation. *Science*, 245(4919):725–730, August 1989.

[130] Albino Bacolla, Guliang Wang, and Karen M Vasquez. New perspectives on DNA and RNA triplexes as effectors of biological activity. *PLoS Genet.*, 11(12):e1005696, December 2015.

[131] Fabian A Buske, John S Mattick, and Timothy L Bailey. Potential in vivo roles of nucleic acid triple-helices. *RNA Biol.*, 8(3):427–439, May 2011.

[132] Bárbara B Mendes, João Conniot, Aviram Avital, Dongbao Yao, Xingya Jiang, Xiang Zhou, Noga Sharf-Pauker, Yuling Xiao, Omer Adir, Haojun Liang, Jinjun Shi, Avi Schroeder, and João Conde. Nanodelivery of nucleic acids. *Nat Rev Methods Primers*, 2, April 2022.

[133] Emil F Khisamutdinov, Hui Li, Daniel L Jasinski, Jiao Chen, Jian Fu, and Peixuan Guo. Enhancing immunomodulation on innate immunity by shape transition among RNA triangle, square and pentagon nanovehicles. *Nucleic Acids Res.*, 42(15):9996–10004, September 2014.

[134] David Yu Zhang and Georg Seelig. Dynamic DNA nanotechnology using strand-displacement reactions. *Nat. Chem.*, 3(2):103–113, February 2011.

[135] Marvin H Caruthers. Chemical synthesis of DNA and DNA analogs. *Acc. Chem. Res.*, 24(9):278–284, September 1991.

[136] Alexandria N Marchi, Ishtiaq Saaem, Jingdong Tian, and Thomas H LaBean. One-pot assembly of a hetero-dimeric DNA origami from chip-derived staples and double-stranded scaffold. *ACS Nano*, 7(2):903–910, February 2013.

[137] Thorsten L Schmidt, Brian J Beliveau, Yavuz O Uca, Mark Theilmann, Felipe Da Cruz, Chao-Ting Wu, and William M Shih. Scalable amplification of strand subsets from chip-synthesized oligonucleotide libraries. *Nat. Commun.*, 6:8634, November 2015.

[138] Mikael Madsen and Kurt V Gothelf. Chemistries for DNA nanotechnology. *Chem. Rev.*, 119(10):6384–6458, May 2019.

[139] Nadrian C Seeman. Nucleic acid junctions and lattices. *J. Theor. Biol.*, 99(2):237–247, November 1982.

[140] Robin Holliday. A mechanism for gene conversion in fungi. *Genet. Res.*, 5(2):282–304, July 1964.

[141] Qinqin Song, Yuemiao Hu, Anqi Yin, Hongbo Wang, and Qikun Yin. DNA holliday junction: History, regulation and bioactivity. *Int. J. Mol. Sci.*, 23(17), August 2022.

[142] Lorenzo Rovigatti, Frank Smallenburg, Flavio Romano, and Francesco Sciortino. Gels of DNA nanostars never crystallize. *ACS Nano*, 8(4):3567–3574, April 2014.

[143] Fu Tsu Ju and Seeman Nadrian C. DNA double-crossover molecules. *Biochemistry*, 32(13):3211–3220, April 1993.

[144] Erik Winfree, Furong Liu, Lisa A Wenzler, and Nadrian C Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394(6693):539–544, August 1998.

[145] Jianping Zheng, Jens J Birktoft, Yi Chen, Tong Wang, Ruojie Sha, Pamela E Constantinou, Stephan L Ginell, Chengde Mao, and Nadrian C Seeman. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature*, 461(7260):74–77, September 2009.

[146] Paul W K Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, March 2006.

[147] Christian Sigl, Elena M Willner, Wouter Engelen, Jessica A Kretzmann, Ken Sachenbacher, Anna Liedl, Fenna Kolbe, Florian Wilsch, S Ali Aghvami, Ulrike Protzer, Michael F Hagan, Seth Fraden, and Hendrik Dietz. Programmable icosahedral shell system for virus trapping. *Nat. Mater.*, 20(9):1281–1289, September 2021.

[148] Ye Tian, Julien R Lhermitte, Lin Bai, Thi Vo, Huolin L Xin, Huilin Li, Ruipeng Li, Masafumi Fukuto, Kevin G Yager, Jason S Kahn, Yan Xiong, Brian Minevich, Sanat K Kumar, and Oleg Gang. Ordered three-dimensional nanomaterials using DNA-prescribed and valence-controlled material voxels. *Nat. Mater.*, 19(7):789–796, July 2020.

[149] Chao-Min Huang, Anjelica Kucinic, Joshua A Johnson, Hai-Jun Su, and Carlos E Castro. Integrated computer-aided engineering and design for DNA assemblies. *Nat. Mater.*, 20(9):1264–1271, September 2021.

[150] Klaus F Wagenbauer, Christian Sigl, and Hendrik Dietz. Gigadalton-scale shape-programmable DNA assemblies. *Nature*, 552(7683):78–83, December 2017.

[151] Carlos E Castro, Hai-Jun Su, Alexander E Marras, Lifeng Zhou, and Joshua Johnson. Mechanical design of DNA nanostructures. *Nanoscale*, 7(14):5913–5921, April 2015.

[152] Anna-Katharina Pumm, Wouter Engelen, Enzo Kopperger, Jonas Isensee, Matthias Vogt, Viktorija Kozina, Massimo Kube, Maximilian N Honemann, Eva Bertosin, Martin Langecker, Ramin Golestanian, Friedrich C Simmel, and Hendrik Dietz. A DNA origami rotary ratchet motor. *Nature*, 607(7919):492–498, July 2022.

[153] Michal Walczak, Ryan A Brady, Leonardo Mancini, Claudia Contini, Roger Rubio-Sánchez, William T Kaufhold, Pietro Cicuta, and Lorenzo Di Michele. Responsive core-shell DNA particles trigger lipid-membrane disruption and bacteria entrapment. *Nat. Commun.*, 12(1):4743, August 2021.

[154] Adrian Leathers, Michal Walczak, Ryan A Brady, Assala Al Samad, Jurij Kotar, Michael J Booth, Pietro Cicuta, and Lorenzo Di Michele. Reaction-diffusion patterning of DNA-based artificial cells. *J. Am. Chem. Soc.*, 144(38):17468–17476, September 2022.

[155] Avik Samanta, Maximilian Hörner, Wei Liu, Wilfried Weber, and Andreas Walther. Signal-processing and adaptive prototissue formation in metabolic DNA protocells. *Nat. Commun.*, 13(1):3968, July 2022.

[156] Guido Grossi, Mette Dalgaard Ebbesen Jepsen, Jørgen Kjems, and Ebbe Sloth Andersen. Control of enzyme reactions by a reconfigurable DNA nanovault. *Nat. Commun.*, 8(1):992, October 2017.

[157] Young Hoon Roh, Jong Bum Lee, Pichamon Kiatwuthinon, Mark R Hartman, Judy J Cha, Soong Ho Um, David A Muller, and Dan Luo. DNAsomes: Multifunctional DNA-based nanocarriers. *Small*, 7(1):74–78, January 2011.

[158] Christopher J Serpell, Thomas G W Edwardson, Pongphak Chidchob, Karina M M Carneiro, and Hanadi F Sleiman. Precision polymers and 3D DNA nanostructures: emergent assemblies from new parameter space. *J. Am. Chem. Soc.*, 136(44):15767–15774, November 2014.

[159] Pongphak Chidchob, Thomas G W Edwardson, Christopher J Serpell, and Hanadi F Sleiman. Synergy of two assembly languages in DNA nanostructures: Self-Assembly of Sequence-Defined polymers on DNA cages. *J. Am. Chem. Soc.*, 138(13):4416–4425, April 2016.

[160] Ryan A Brady, Nicholas J Brooks, Pietro Cicuta, and Lorenzo Di Michele. Crystallization of amphiphilic DNA C-Stars. *Nano Lett.*, 17(5):3276–3281, May 2017.

[161] Roger Rubio-Sánchez, Bortolo Matteo Mognetti, Pietro Cicuta, and Lorenzo Di Michele. DNA-Origami Line-Actants control domain organization and fission in synthetic membranes. *J. Am. Chem. Soc.*, 145(20):11265–11275, May 2023.

[162] Roger Rubio-Sánchez, Giacomo Fabrini, Pietro Cicuta, and Lorenzo Di Michele. Amphiphilic DNA nanostructures for bottom-up synthetic biology. *Chem. Commun.*, 57(95):12725–12740, November 2021.

[163] Ryan A Brady, Nicholas J Brooks, Vito Foderà, Pietro Cicuta, and Lorenzo Di Michele. Amphiphilic-DNA platform for the design of crystalline frameworks with programmable structure and functionality. *J. Am. Chem. Soc.*, 140(45):15384–15392, November 2018.

[164] Peter H von Hippel, Neil P Johnson, and Andrew H Marcus. Fifty years of DNA "breathing": Reflections on old and new approaches. *Biopolymers*, 99(12):923–954, December 2013.

[165] Friedrich C Simmel, Bernard Yurke, and Hari R Singh. Principles and applications of nucleic acid strand displacement reactions. *Chem. Rev.*, 119(10):6326–6369, May 2019.

[166] Bernard Yurke, Andrew J Turberfield, Allen P Mills, Jr, Friedrich C Simmel, and Jennifer L Neumann. A DNA-fuelled molecular machine made of DNA. *Nature*, 406(6796):605–608, August 2000.

[167] Niranjan Srinivas, Thomas E Ouldridge, Petr Sulc, Joseph M Schaeffer, Bernard Yurke, Ard A Louis, Jonathan P K Doye, and Erik Winfree. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.*, 41(22):10641–10658, December 2013.

[168] Georg Seelig, David Soloveichik, David Yu Zhang, and Erik Winfree. Enzyme-free nucleic acid logic circuits. *Science*, 314(5805):1585–1588, December 2006.

[169] Bingling Li, Andrew D Ellington, and Xi Chen. Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Res.*, 39(16):e110, September 2011.

[170] Lulu Qian, Erik Winfree, and Jehoshua Bruck. Neural network computation with DNA strand displacement cascades. *Nature*, 475(7356):368–372, July 2011.

[171] Kevin M Cherry and Lulu Qian. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature*, 559(7714):370–376, July 2018.

[172] Xiewei Xiong, Tong Zhu, Yun Zhu, Mengyao Cao, Jin Xiao, Li Li, Fei Wang, Chunhai Fan, and Hao Pei. Molecular convolutional neural networks with DNA regulatory circuits. *Nature Machine Intelligence*, 4(7):625–635, July 2022.

[173] Angelo Cangialosi, Changkyu Yoon, Jiayu Liu, Qi Huang, Jingkai Guo, Thao D Nguyen, David H Gracias, and Rebecca Schulman. DNA sequence-directed shape change of photopatterned hydrogels via high-degree swelling. *Science*, 357(6356):1126–1130, September 2017.

[174] William B Sherman and Nadrian C Seeman. A precisely controlled DNA biped walking device. *Nano Lett.*, 4(7):1203–1207, July 2004.

[175] Tosan Omabegho, Ruojie Sha, and Nadrian C Seeman. A bipedal DNA brownian motor with coordinated legs. *Science*, 324(5923):67–71, April 2009.

[176] Friedrich C Simmel. Processive motion of bipedal DNA walkers. *Chemphyschem*, 10(15):2593–2597, October 2009.

[177] Anupama J Thubagere, Wei Li, Robert F Johnson, Zibo Chen, Shayan Doroudi, Yae Lim Lee, Gregory Izatt, Sarah Wittman, Niranjan Srinivas, Damien Woods, Erik Winfree, and Lulu Qian. A cargo-sorting DNA robot. *Science*, 357(6356), September 2017.

[178] Hao Yan, Xiaoping Zhang, Zhiyong Shen, and Nadrian C Seeman. A robust DNA mechanical device controlled by hybridization topology. *Nature*, 415(6867):62–65, January 2002.

[179] Heini Ijäs, Sami Nummelin, Boxuan Shen, Mauri A Kostiainen, and Veikko Linko. Dynamic DNA origami devices: from Strand-Displacement reactions to External-Stimuli responsive systems. *Int. J. Mol. Sci.*, 19(7), July 2018.

[180] Shiliang He, Zhilei Ge, Xiaolei Zuo, Chunhai Fan, and Xiuhai Mao. Dynamic regulation of DNA nanostructures by noncanonical nucleic acids. *NPG Asia Materials*, 13(1):1–12, April 2021.

[181] Jason S Kahn, Alexander Trifonov, Alessandro Cecconello, Weiwei Guo, Chunhai Fan, and Itamar Willner. Integration of switchable DNA-Based hydrogels with surfaces by the hybridization chain reaction. *Nano Lett.*, 15(11):7773–7778, November 2015.

[182] Miu Shan Chan, Hoi Man Leung, Sze Wing Wong, Zihong Lin, Qi Gao, Tristan Juin Han Chang, King Wai Chiu Lai, and Pik Kwan Lo. Reversible reconfiguration of high-order DNA nanostructures by employing g-quartet toeholds as adhesive units. *Nanoscale*, 12(4):2464–2471, January 2020.

[183] Jiantong Dong, Michael P O'Hagan, and Itamar Willner. Switchable and dynamic G-quadruplexes and their applications. *Chem. Soc. Rev.*, 51(17):7631–7661, August 2022.

[184] Enjun Cheng, Yongzheng Xing, Ping Chen, Yang Yang, Yawei Sun, Dejian Zhou, Lijin Xu, Qinghua Fan, and Dongsheng Liu. A ph-triggered, fast-responding DNA hydrogel. *Angew. Chem. Int. Ed Engl.*, 48(41):7660–7663, 2009.

[185] Andrea Idili, Alexis Vallée-Bélisle, and Francesco Ricci. Programmable ph-triggered DNA nanoswitches. *J. Am. Chem. Soc.*, 136(16):5836–5839, April 2014.

[186] Jiangtao Ren, Yuwei Hu, Chun-Hua Lu, Weiwei Guo, Miguel Angel Aleman-Garcia, Francesco Ricci, and Itamar Willner. ph-responsive and switchable triplex-based DNA hydrogels. *Chem. Sci.*, 6(7):4190–4195, July 2015.

[187] Yuki Suzuki, Ibuki Kawamata, Kohei Mizuno, and Satoshi Murata. Large deformation of a DNA-Origami nanoarm induced by the cumulative actuation of Tension-Adjustable modules. *Angew. Chem. Int. Ed Engl.*, 59(15):6230–6234, April 2020.

[188] Andrew D Ellington and Jack W Szostak. In vitro selection of RNA molecules that bind specific ligands. *Nature*, 346(6287):818–822, August 1990.

[189] Luiz F M Passalacqua, Michael T Banco, Jared D Moon, Xing Li, Samie R Jaffrey, and Adrian R Ferré-D'Amaré. Intricate 3D architecture of a DNA mimic of GFP. *Nature*, 618(7967):1078–1084, June 2023.

[190] Ronald R Breaker and Gerald F Joyce. A DNA enzyme that cleaves RNA. *Chem. Biol.*, 1(4):223–229, December 1994.

[191] Andrew J Davidson and Will Wood. Unravelling the actin cytoskeleton: A new competitive edge? *Trends Cell Biol.*, 26(8):569–576, August 2016.

[192] Kevin Jahnke, Vanessa Huth, Ulrike Mersdorf, Na Liu, and Kerstin Göpfrich. Bottom-Up assembly of synthetic cells with a DNA cytoskeleton. *ACS Nano*, 16(5):7233–7241, May 2022.

[193] Pengfei Zhan, Kevin Jahnke, Na Liu, and Kerstin Göpfrich. Functional DNA-based cytoskeletons for synthetic cells. *Nat. Chem.*, 14(8):958–963, August 2022.

[194] Jason S Kahn, Yuwei Hu, and Itamar Willner. Stimuli-Responsive DNA-Based hydrogels: From basic principles to applications. *Acc. Chem. Res.*, 50(4):680–690, April 2017.

[195] Vinod Morya, Shanka Walia, Biman B Mandal, Chinmay Ghoroi, and Dhiraj Bhatia. Functional DNA based hydrogels: Development, properties and biological applications. *ACS Biomater Sci Eng*, 6(11):6021–6035, November 2020.

[196] Joshua Fern and Rebecca Schulman. Modular DNA strand-displacement controllers for directing material expansion. *Nat. Commun.*, 9(1):3766, September 2018.

[197] Julia Müller, Anna Christina Jäkel, Dominic Schwarz, Lukas Aufinger, and Friedrich C Simmel. Programming diffusion and localization of DNA signals in 3d-printed DNA-functionalized hydrogels. *Small*, 16(31):e2001815, August 2020.

[198] Chikako Kurokawa, Kei Fujiwara, Masamune Morita, Ibuki Kawamata, Yui Kawagishi, Atsushi Sakai, Yoshihiro Murayama, Shin-Ichiro M Nomura, Satoshi Murata, Masahiro Takinoue, and Miho Yanagisawa. DNA cytoskeleton for stabilizing artificial cells. *Proc. Natl. Acad. Sci. U. S. A.*, 114(28):7228–7233, July 2017.

[199] Yusuke Sato, Tetsuro Sakamoto, and Masahiro Takinoue. Sequence-based engineering of dynamic functions of micrometer-sized DNA droplets. *Sci Adv*, 6(23):eaba3471, June 2020.

[200] Sungho Do, Chanseok Lee, Taehyun Lee, Do-Nyun Kim, and Yongdae Shin. Engineering DNA-based synthetic condensates with programmable material properties, compositions, and functionalities. *Sci Adv*, 8(41):eabj1771, October 2022.

[201] Yusuke Sato and Masahiro Takinoue. Sequence-dependent fusion dynamics and physical properties of DNA droplets. *Nanoscale Adv*, 5(7):1919–1925, March 2023.

[202] Ryan A Brady, Will T Kaufhold, Nicholas J Brooks, Vito Foderà, and Lorenzo Di Michele. Flexibility defines structure in crystals of amphiphilic DNA nanostars. *J. Phys. Condens. Matter*, 31(7):074003, February 2019.

[203] Patrick J Grimes, Agostino Galanti, and Pierangelo Gobbo. Bioinspired networks of communicating synthetic protocells. *Front Mol Biosci*, 8:804717, December 2021.

[204] Shuo Yang, Alex Joesaar, Bas W A Bögels, Stephen Mann, and Tom F A de Greef. Protocellular CRISPR/Cas-Based diffusive communication using transcriptional RNA signaling. *Angew. Chem. Int. Ed Engl.*, 61(26):e202202436, June 2022.

[205] Peixuan Guo. RNA nanotechnology: engineering, assembly and applications in detection, gene delivery and therapy. *J. Nanosci. Nanotechnol.*, 5(12):1964–1982, December 2005.

[206] Peixuan Guo. The emerging field of RNA nanotechnology. *Nat. Nanotechnol.*, 5(12):833–842, December 2010.

[207] Peixuan Guo and Kirill A Afonin. *RNA Nanotechnology and Therapeutics*. CRC Press, July 2022.

[208] Eduarde Rohner, Ran Yang, Kylie S Foo, Alexander Goedel, and Kenneth R Chien. Unlocking the promise of mRNA therapeutics. *Nat. Biotechnol.*, 40(11):1586–1600, November 2022.

[209] E Ennifar, M Yusupov, P Walter, R Marquet, B Ehresmann, C Ehresmann, and P Dumas. The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure*, 7(11):1439–1449, November 1999.

[210] Eric Ennifar, Philippe Walter, Bernard Ehresmann, Chantal Ehresmann, and Philippe Dumas. Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. *Nat. Struct. Biol.*, 8(12):1064–1068, December 2001.

[211] Hansini Mundigala, Jonathan B Michaux, Andrew L Feig, Eric Ennifar, and David Rueda. HIV-1 DIS stem loop forms an obligatory bent kissing intermediate in the dimerization pathway. *Nucleic Acids Res.*, 42(11):7281–7289, June 2014.

[212] Song Cao and Shi-Jie Chen. Structure and stability of RNA/RNA kissing complex: with application to HIV dimerization initiation signal. *RNA*, 17(12):2130–2143, December 2011.

[213] Petr Šulc, Flavio Romano, Thomas E Ouldridge, Jonathan P K Doye, and Ard A Louis. A nucleotide-level coarse-grained model of RNA. *J. Chem. Phys.*, 140(23):235102, June 2014.

[214] Hui Zhang, James A Endrizzi, Yi Shu, Farzin Haque, Claude Sauter, Lyudmila S Shlyakhtenko, Yuri Lyubchenko, Peixuan Guo, and Young-In Chi. Crystal structure of 3WJ core revealing divalent ion-promoted thermostability and assembly of the phi29 hexameric motor pRNA. *RNA*, 19(9):1226–1237, September 2013.

[215] Mo Li, Mengxi Zheng, Siyu Wu, Cheng Tian, Di Liu, Yossi Weizmann, Wen Jiang, Guansong Wang, and Chengde Mao. In vivo production of RNA nanostructures via programmed folding of single-stranded RNAs. *Nat. Commun.*, 9(1):2196, June 2018.

[216] Nikolai Windbichler, Michael Werner, and Renée Schroeder. Kissing complex-mediated dimerisation of HIV-1 RNA: coupling extended duplex formation to ribozyme cleavage. *Nucleic Acids Res.*, 31(22):6419–6427, November 2003.

[217] Sabine Müller, Bettina Appel, Darko Balke, Robert Hieronymus, and Claudia Nübel. Thirty-five years of research into ribozymes and nucleic acid catalysis: where do we stand today? *F1000Res.*, 5, June 2016.

[218] Andrew Fire, Siqun Xu, Mary K Montgomery, Steven A Kostas, Samuel E Driver, and Craig C Mello. Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. *Nature*, 391(6669):806–811, February 1998.

[219] Tin Hoang Trung Chau, Dung Hoang Anh Mai, Diep Ngoc Pham, Hoa Thi Quynh Le, and Eun Yeol Lee. Developments of riboswitches and toehold switches for molecular Detection-Biosensing and molecular diagnostics. *Int. J. Mol. Sci.*, 21(9), April 2020.

[220] Wade W Grabow and Luc Jaeger. RNA self-assembly and RNA nanotechnology. *Acc. Chem. Res.*, 47(6):1871–1880, June 2014.

[221] Christian Hammann, Andrej Luptak, Jonathan Perreault, and Marcos de la Peña. The ubiquitous hammerhead ribozyme. *RNA*, 18(5):871–885, May 2012.

[222] Ian J Macrae, Kaihong Zhou, Fei Li, Adrian Repic, Angela N Brooks, W Zacheus Cande, Paul D Adams, and Jennifer A Doudna. Structural basis for double-stranded RNA processing by dicer. *Science*, 311(5758):195–198, January 2006.

[223] Kumari Asha, Prashant Kumar, Melvin Sanicas, Clement A Meseko, Madhu Khanna, and Binod Kumar. Advancements in nucleic acid based therapeutics against respiratory viral infections. *J. Clin. Med. Res.*, 8(1), December 2018.

[224] Alexander A Green, Pamela A Silver, James J Collins, and Peng Yin. Toehold switches: de-novo-designed regulators of gene expression. *Cell*, 159(4):925–939, November 2014.

[225] Farren J Isaacs, Daniel J Dwyer, and James J Collins. RNA synthetic biology. *Nat. Biotechnol.*, 24(5):545–554, May 2006.

[226] Peter B Dykstra, Matias Kaplan, and Christina D Smolke. Engineering synthetic RNA devices for cell control. *Nat. Rev. Genet.*, 23(4):215–228, April 2022.

[227] Eric A Davidson and Andrew D Ellington. Synthetic RNA circuits. *Nat. Chem. Biol.*, 3(1):23–28, January 2007.

[228] Satoshi Matsuura, Hiroki Ono, Shunsuke Kawasaki, Yi Kuang, Yoshihiko Fujita, and Hirohide Saito. Synthetic RNA-based logic computation in mammalian cells. *Nat. Commun.*, 9(1):4847, November 2018.

[229] Malvin L Pardi, Juanqi Wu, Shunsuke Kawasaki, and Hirohide Saito. Synthetic RNA-based post-transcriptional expression control methods and genetic circuits. *Adv. Drug Deliv. Rev.*, 184:114196, May 2022.

[230] Shigetoshi Kameda, Hirohisa Ohno, and Hirohide Saito. Synthetic circular RNA switches and circuits that control protein expression in mammalian cells. *Nucleic Acids Res.*, 51(4):e24, February 2023.

[231] David Bumcrot, Muthiah Manoharan, Victor Koteliansky, and Dinah W Y Sah. RNAi therapeutics: a potential new class of pharmaceutical drugs. *Nat. Chem. Biol.*, 2(12):711–719, December 2006.

[232] Jerzy Ciesiolka, Jessica Gorski, and Michael Yarus. Selection of an RNA domain that binds Zn^{2+}. *RNA*, 1(5):538–550, July 1995.

[233] Thorsten Dieckmann, Ellen Suzuki, Gerald K Nakamura, and Juli Feigon. Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA*, 2(7):628–640, July 1996.

[234] Annamaria Ruscito and Maria C DeRosa. Small-Molecule binding aptamers: Selection strategies, characterization, and applications. *Front Chem*, 4:14, May 2016.

[235] Bo Shui, Abdullah Ozer, Warren Zipfel, Nevedita Sahu, Avtar Singh, John T Lis, Hua Shi, and Michael I Kotlikoff. RNA aptamers that functionally interact with green fluorescent protein and its derivatives. *Nucleic Acids Res.*, 40(5):e39, March 2012.

[236] Kathrin Leppek and Georg Stoecklin. An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. *Nucleic Acids Res.*, 42(2):e13, January 2014.

[237] Xiaohong Fang and Weihong Tan. Aptamers generated from cell-SELEX for molecular medicine: a chemical biology approach. *Acc. Chem. Res.*, 43(1):48–57, January 2010.

[238] Craig Tuerk and Larry Gold. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, 249(4968):505–510, August 1990.

[239] R E Wang, Y Zhang, J Cai, W Cai, and T Gao. Aptamer-based fluorescent biosensors. *Curr. Med. Chem.*, 18(27):4175–4184, 2011.

[240] Mette D E Jepsen, Steffen M Sparvath, Thorbjørn B Nielsen, Ane H Langvad, Guido Grossi, Kurt V Gothelf, and Ebbe S Andersen. Development of a genetically encodable FRET system using fluorescent RNA aptamers. *Nat. Commun.*, 9(1):18, January 2018.

[241] Dmitry M Kolpashchikov. Binary malachite green aptamer for fluorescent detection of nucleic acids. *J. Am. Chem. Soc.*, 127(36):12442–12443, September 2005.

[242] Khalid K Alam, Kwaku D Tawiah, Matthew F Lichte, David Porciani, and Donald H Burke. A fluorescent split aptamer for visualizing RNA–RNA assembly in vivo. *ACS Synth. Biol.*, 6(9):1710–1721, September 2017.

[243] Katharina Höfer, Lisa V Langejürgen, and Andres Jäschke. Universal aptamer-based real-time monitoring of enzymatic RNA synthesis. *J. Am. Chem. Soc.*, 135(37):13692–13694, September 2013.

[244] Jongmin Kim, Juan F Quijano, Jeongwon Kim, Enoch Yeung, and Richard M Murray. Synthetic logic circuits using RNA aptamer against T7 RNA polymerase. September 2014.

[245] Saskia Neubacher and Sven Hennig. RNA structure and cellular applications of fluorescent light-up aptamers. *Angew. Chem. Int. Ed Engl.*, 58(5):1266–1279, January 2019.

[246] Lynda Truong and Adrian R Ferré-D'Amaré. From fluorescent proteins to fluorogenic RNAs: Tools for imaging cellular macromolecules. *Protein Sci.*, 28(8):1374–1386, August 2019.

[247] Jeremy S Paige, Karen Y Wu, and Samie R Jaffrey. RNA mimics of green fluorescent protein. *Science*, 333(6042):642–646, July 2011.

[248] Robert J Trachman, 3rd, Lynda Truong, and Adrian R Ferré-D'Amaré. Structural principles of fluorescent RNA aptamers. *Trends Pharmacol. Sci.*, 38(10):928–939, October 2017.

[249] Wade W Grabow, Paul Zakrevsky, Kirill A Afonin, Arkadiusz Chworos, Bruce A Shapiro, and Luc Jaeger. Self-assembling RNA nanorings based on RNAI/II inverse kissing complexes. *Nano Lett.*, 11(2):878–887, February 2011.

[250] Christopher Baugh, Dilârâ Grate, and Charles Wilson. 2.8 Å crystal structure of the malachite green aptamer. *J. Mol. Biol.*, 301(1):117–128, August 2000.

[251] Jeremy R Babendure, Stephen R Adams, and Roger Y Tsien. Aptamers switch on fluorescence of triphenylmethane dyes. *J. Am. Chem. Soc.*, 125(48):14716–14717, December 2003.

[252] V Siddartha Yerramilli and Kyung Hyuk Kim. Labeling RNAs in live cells using malachite green aptamer scaffolds as fluorescent probes. *ACS Synth. Biol.*, 7(3):758–766, March 2018.

[253] Grigory S Filonov, Jared D Moon, Nina Svensen, and Samie R Jaffrey. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.*, 136(46):16299–16308, November 2014.

[254] Grigory S Filonov, Christina W Kam, Wenjiao Song, and Samie R Jaffrey. In-gel imaging of RNA processing using broccoli reveals optimal aptamer expression strategies. *Chem. Biol.*, 22(5):649–660, May 2015.

[255] Kyu Young Han, Benjamin J Leslie, Jingyi Fei, Jichuan Zhang, and Taekjip Ha. Understanding the photophysics of the spinach-DFHBI RNA aptamer-fluorogen complex to improve live-cell RNA imaging. *J. Am. Chem. Soc.*, 135(50):19033–19038, December 2013.

[256] Nguyen Thuan Dao, Reinhard Haselsberger, Mai Thu Khuc, Anh Tuân Phan, Alexander A Voityuk, and Maria-Elisabeth Michel-Beyerle. Photophysics of DFHBI bound to RNA aptamer baby spinach. *Sci. Rep.*, 11(1):7356, April 2021.

[257] Wenjiao Song, Rita L Strack, Nina Svensen, and Samie R Jaffrey. Plug-and-play fluorophores extend the spectral properties of spinach. *J. Am. Chem. Soc.*, 136(4):1198–1201, January 2014.

[258] Xing Li, Hyaeyeong Kim, Jacob L Litke, Jiahui Wu, and Samie R Jaffrey. Fluorophore-Promoted RNA folding and photostability enables imaging of single Broccoli-Tagged mRNAs in live mammalian cells. *Angew. Chem. Int. Ed Engl.*, 59(11):4511–4518, March 2020.

[259] Hao Huang, Nikolai B Suslov, Nan-Sheng Li, Sandip A Shelke, Molly E Evans, Yelena Koldobskaya, Phoebe A Rice, and Joseph A Piccirilli. A G-quadruplex–containing RNA activates fluorescence in a GFP-like fluorophore. *Nat. Chem. Biol.*, 10(8):686–691, June 2014.

[260] Katherine Deigan Warner, Michael C Chen, Wenjiao Song, Rita L Strack, Andrea Thorn, Samie R Jaffrey, and Adrian R Ferré-D'Amaré. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat. Struct. Mol. Biol.*, 21(8):658–663, August 2014.

[261] Rita L Strack, Wenjiao Song, and Samie R Jaffrey. Using spinach-based sensors for fluorescence imaging of intracellular metabolites and proteins in living bacteria. *Nat. Protoc.*, 9(1):146–155, January 2014.

[262] Mingxu You, Jacob L Litke, and Samie R Jaffrey. Imaging metabolite dynamics in living cells using a spinach-based riboswitch. *Proc. Natl. Acad. Sci. U. S. A.*, 112(21):E2756–65, May 2015.

[263] Shuo-Fu Yuan and Hal S Alper. Improving spinach2-and broccoli-based biosensors for single and double analytes. *Biotechnology Notes*, 1:2–8, January 2020.

[264] Rita L Strack, Matthew D Disney, and Samie R Jaffrey. A superfolding spinach2 reveals the dynamic nature of trinucleotide repeat–containing RNA. *Nat. Methods*, 10(12):1219–1224, October 2013.

[265] Eman A Ageely, Zachary J Kartje, Kushal J Rohilla, Christopher L Barkau, and Keith T Gagnon. Quadruplex-Flanking stem structures modulate the stability and metal ion preferences of RNA mimics of GFP. *ACS Chem. Biol.*, 11(9):2398–2406, September 2016.

[266] Jaimie Marie Stewart, Hari K K Subramanian, and Elisa Franco. Self-assembly of multi-stranded RNA motifs into lattices and tubular structures. *Nucleic Acids Res.*, 45(9):5449–5457, May 2017.

[267] Jaimie Marie Stewart, Cody Geary, and Elisa Franco. Design and characterization of RNA nanotubes. *ACS Nano*, 13(5):5214–5221, May 2019.

[268] Jaimie Marie Stewart, Mathias Viard, Hari K K Subramanian, Brandon K Roark, Kirill A Afonin, and Elisa Franco. Programmable RNA microstructures for coordinated delivery of siRNAs. *Nanoscale*, 8(40):17542–17550, October 2016.

[269] A Xayaphoummine, T Bucher, and H Isambert. Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res.*, 33(Web Server issue):W605–10, July 2005.

[270] Eckart Bindewald, Robert Hayes, Yaroslava G Yingling, Wojciech Kasprzak, and Bruce A Shapiro. RNAJunction: a database of RNA junctions and kissing loops for three-dimensional structural analysis and nanodesign. *Nucleic Acids Res.*, 36(Database issue):D392–7, January 2008.

[271] Nicholas R Markham and Michael Zuker. UNAFold: software for nucleic acid folding and hybridization. *Methods Mol. Biol.*, 453:3–31, 2008.

[272] Bruce A Shapiro. Computational design strategies for RNA nanostructures. *J. Biomol. Struct. Dyn.*, 2009.

[273] Songchuan Guo, Nuska Tschammer, Sulma Mohammed, and Peixuan Guo. Specific delivery of therapeutic RNAs to cancer cells via the dimerization mechanism of phi29 motor pRNA. *Hum. Gene Ther.*, 16(9):1097–1109, September 2005.

[274] Satoru Horiya, Xianglan Li, Gota Kawai, Ryota Saito, Akira Katoh, Koh Kobayashi, and Kazuo Harada. RNA LEGO: magnesium-dependent formation of specific RNA assemblies through kissing interactions. *Chem. Biol.*, 10(7):645–654, July 2003.

[275] Emil F Khisamutdinov, Daniel L Jasinski, and Peixuan Guo. RNA as a boiling-resistant anionic polymer material to build robust structures with defined shape and stoichiometry. *ACS Nano*, 8(5):4771–4781, May 2014.

[276] Isil Severcan, Cody Geary, Erik Verzemnieks, Arkadiusz Chworos, and Luc Jaeger. Square-shaped RNA particles from different RNA folds. *Nano Lett.*, 9(3):1270–1277, March 2009.

[277] Daniel L Jasinski, Emil F Khisamutdinov, Yuri L Lyubchenko, and Peixuan Guo. Physicochemically tunable polyfunctionalized RNA square architecture with fluorogenic and ribozymatic properties. *ACS Nano*, 8(8):7620–7629, August 2014.

[278] Yi Shu, Farzin Haque, Dan Shu, Wei Li, Zhenqi Zhu, Malak Kotb, Yuri Lyubchenko, and Peixuan Guo. Fabrication of 14 different RNA nanoparticles for specific tumor targeting without accumulation in normal organs. *RNA*, 19(6):767–777, June 2013.

[279] Daniel Jasinski, Farzin Haque, Daniel W Binzel, and Peixuan Guo. Advancement of the emerging field of RNA nanotechnology. *ACS Nano*, 11(2):1142–1164, February 2017.

[280] Caroline Rossi-Gendron, Farah El Fakih, Laura Bourdon, Koyomi Nakazawa, Julie Finkel, Nicolas Triomphe, Léa Chocron, Masayuki Endo, Hiroshi Sugiyama, Gaëtan Bellot, Mathieu Morel, Sergii Rudiuk, and Damien Baigl. Isothermal self-assembly of multicomponent and evolutive DNA nanostructures. *Nat. Nanotechnol.*, July 2023.

[281] H K Majumder, U Maitra, and M Rosenberg. Termination of transcription by bacteriophage T3 RNA polymerase: homogeneous 3'-terminal oligonucleotide sequence of in vitro T3 RNA polymerase transcripts. *Proc. Natl. Acad. Sci. U. S. A.*, 76(10):5110–5113, October 1979.

[282] Eugene T Butler and Michael J Chamberlin. Bacteriophage SP6-specific RNA polymerase. i. isolation and characterization of the enzyme. *J. Biol. Chem.*, 257(10):5772–5778, May 1982.

[283] Michael Chamberlin, Janet McGrath, and Lucy Waskell. New RNA polymerase from escherichia coli infected with bacteriophage T7. *Nature*, 228(5268):227–231, October 1970.

[284] Yasaman Gholamalipour, Aruni Karunanayake Mudiyanselage, and Craig T Martin. 3' end additions by T7 RNA polymerase are RNA self-templated, distributive and diverse in character-RNA-Seq analyses. *Nucleic Acids Res.*, 46(18):9253–9263, October 2018.

[285] Bakhos A Tannous, Eleftheria Laios, and Theodore K Christopoulos. T7 RNA polymerase as a self-replicating label for antigen quantification. *Nucleic Acids Res.*, 30(24):e140, December 2002.

[286] Jean-Claude Boulain, Janie Dassa, Laurent Mesta, Alexandra Savatier, Narciso Costa, Bruno H Muller, Guillaume L'hostis, Enrico A Stura, Alain Troesch, and Frédéric Ducancel. Mutants with higher stability and specific activity from a single thermosensitive variant of T7 RNA polymerase. *Protein Eng. Des. Sel.*, 26(11):725–734, November 2013.

[287] Athanasios Dousis, Kanchana Ravichandran, Elissa M Hobert, Melissa J Moore, and Amy E Rabideau. An engineered T7 RNA polymerase that produces mRNA free of immunostimulatory byproducts. *Nat. Biotechnol.*, 41(4):560–568, April 2023.

[288] Srinivas S Sastry and Barbara M Ross. Nuclease activity of T7 RNA polymerase and the heterogeneity of transcription elongation complexes. *J. Biol. Chem.*, 272(13):8644–8652, March 1997.

[289] Manish Kushwaha and Howard M Salis. A portable expression resource for engineering cross-species genetic circuits and pathways. *Nat. Commun.*, 6:7832, July 2015.

[290] Jesus Fernandez-Rodriguez, Felix Moser, Miryoung Song, and Christopher A Voigt. Engineering RGB color vision into escherichia coli. *Nat. Chem. Biol.*, 13(7):706–708, July 2017.

[291] Karsten Temme, Rena Hill, Thomas H Segall-Shapiro, Felix Moser, and Christopher A Voigt. Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nucleic Acids Res.*, 40(17):8773–8781, September 2012.

[292] Graham M Cheetham and Thomas A Steitz. Structure of a transcribing T7 RNA polymerase initiation complex. *Science*, 286(5448):2305–2309, December 1999.

[293] Dragana Korencić, Dieter Söll, and Alexandre Ambrogelly. A one-step method for in vitro production of tRNA transcripts. *Nucleic Acids Res.*, 30(20):e105, October 2002.

[294] John F Milligan, Duncan R Groebe, Gary W Witherell, and Olke C Uhlenbeck. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, 15(21):8783–8798, November 1987.

[295] John F Milligan and Olke C Uhlenbeck. [5] synthesis of small RNAs using T7 RNA polymerase. In *Methods in Enzymology*, volume 180, pages 51–62. Academic Press, January 1989.

[296] Maribeth Maslak and Craig T Martin. Kinetic analysis of T7 RNA polymerase transcription initiation from promoters containing single-stranded regions. *Biochemistry*, 32(16):4281–4285, April 1993.

[297] Manli Jiang, Minqing Rong, Craig Martin, and William T McAllister. Interrupting the template strand of the T7 promoter facilitates translocation of the DNA during initiation, reducing transcript slippage and the release of abortive products. *J. Mol. Biol.*, 310(3):509–522, July 2001.

[298] Gary M Skinner, Christoph G Baumann, Diana M Quinn, Justin E Molloy, and James G Hoggett. Promoter binding, initiation, and elongation by bacteriophage T7 RNA polymerase. a single-molecule view of the transcription cycle. *J. Biol. Chem.*, 279(5):3239–3244, January 2004.

[299] Andrey Revyakin, Chenyu Liu, Richard H Ebright, and Terence R Strick. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*, 314(5802):1139–1143, November 2006.

[300] Achillefs N Kapanidis, Emmanuel Margeat, Sam On Ho, Ekaterine Kortkhonjia, Shimon Weiss, and Richard H Ebright. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*, 314(5802):1144–1147, November 2006.

[301] Luis G Brieba and Rui Sousa. T7 promoter release mediated by DNA scrunching. *EMBO J.*, 20(23):6826–6835, December 2001.

[302] Yiping Jia and Smita S Patel. Kinetic mechanism of GTP binding and RNA synthesis during transcription initiation by bacteriophage T7 RNA polymerase. *J. Biol. Chem.*, 272(48):30147–30153, November 1997.

[303] Iaroslav Kuzmine, Philip A Gottlieb, and Craig T Martin. Binding of the priming nucleotide in the initiation of transcription by T7 RNA polymerase*. *J. Biol. Chem.*, 278(5):2819–2823, January 2003.

[304] Thomas Conrad, Izabela Plumbom, Maria Alcobendas, Ramon Vidal, and Sascha Sauer. Maximizing transcription of nucleic acids with efficient T7 promoters. *Commun Biol*, 3(1):439, August 2020.

[305] Craig T Martin, Daniel K Muller, and Joseph E Coleman. Processivity in early stages of transcription by T7 RNA polymerase. *Biochemistry*, 27(11):3966–3974, May 1988.

[306] Kimberly J Durniak, Scott Bailey, and Thomas A Steitz. The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation. *Science*, 322(5901):553–557, October 2008.

[307] Hye Ran Koh, Rahul Roy, Maria Sorokina, Guo-Qing Tang, Divya Nandakumar, Smita S Patel, and Taekjip Ha. Correlating transcription initiation and conformational changes by a Single-Subunit RNA polymerase with near Base-Pair resolution. *Mol. Cell*, 70(4):695–706.e5, May 2018.

[308] Wei Zhou, Daniel Reines, and Paul W Doetsch. T7 RNA polymerase bypass of large gaps on the template strand reveals a critical role of the nontemplate strand in elongation. *Cell*, 82(4):577–585, August 1995.

[309] Minqing Rong, Russell K Durbin, and William T McAllister. Template strand switching by T7 RNA polymerase. *J. Biol. Chem.*, 273(17):10253–10260, April 1998.

[310] Samuel W Schaffter, Leopold N Green, Joanna Schneider, Hari K K Subramanian, Rebecca Schulman, and Elisa Franco. T7 RNA polymerase non-specifically transcribes and induces disassembly of DNA nanostructures. *Nucleic Acids Res.*, 46(10):5332–5343, June 2018.

[311] Francisco J Triana-Alonso, Marylena Dabrowski, Jörg Wadzack, and Knud H Nierhaus. Self-coded 3'-extension of run-off transcripts produces aberrant products during in vitro transcription with T7 RNA polymerase. *J. Biol. Chem.*, 270(11):6298–6307, March 1995.

[312] B A Donahue, S Yin, J S Taylor, D Reines, and P C Hanawalt. Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proc. Natl. Acad. Sci. U. S. A.*, 91(18):8502–8506, August 1994.

[313] Qing Guo and Rui Sousa. Multiple roles for the T7 promoter nontemplate strand during transcription initiation and polymerase release. *J. Biol. Chem.*, 280(5):3474–3482, February 2005.

[314] Jeffrey A Kern and Robert H Davis. Application of solution equilibrium analysis to in vitro RNA transcription. *Biotechnol. Prog.*, 13(6):747–756, November 1997.

[315] Sabine Arnold, Martin Siemann, Kai Scharnweber, Markus Werner, Sandra Baumann, and Matthias Reuss. Kinetic modeling and simulation of in vitro transcription by phage T7 RNA polymerase. *Biotechnol. Bioeng.*, 72(5):548–561, March 2001.

[316] Satoru Akama, Masayuki Yamamura, and Takanori Kigawa. A multiphysics model of in vitro transcription coupling enzymatic reaction and precipitation formation. *Biophys. J.*, 102(2):221–230, January 2012.

[317] Susan C Milburn, Marianna Goldrick, and Matthew Winkler. Compositions and methods for increasing the yields of in vitro RNA transcription and other polynucleotide synthetic reactions, October 1993.

[318] Anju Jain and Anil K Tyagi. Role of polyamines in the synthesis of RNA in mycobacteria. *Mol. Cell. Biochem.*, 78(1):3–8, November 1987.

[319] Magali Frugier, Catherine Florentz, Mir Wais Hosseini, Jean-Marie Lehn, and Richard Giegé. Synthetic polyamines stimulate in vitro transcription by T7 RNA polymerase. *Nucleic Acids Res.*, 22(14):2784–2790, July 1994.

[320] Daniel Lai, Jeff R Proctor, and Irmtraud M Meyer. On the importance of cotranscriptional RNA structure formation. *RNA*, 19(11):1461–1473, November 2013.

[321] Cody Geary, Paul W K Rothemund, and Ebbe S Andersen. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science*, 345(6198):799–804, August 2014.

[322] Anna J Lee and Donald M Crothers. The solution structure of an RNA loop-loop complex: the ColE1 inverted loop sequence. *Structure*, 6(8):993–1005, August 1998.

[323] Cody Geary, Guido Grossi, Ewan K S McRae, Paul W K Rothemund, and Ebbe S Andersen. RNA origami design tools enable cotranscriptional folding of kilobase-sized nanoscaffolds. *Nat. Chem.*, 13(6):549–558, June 2021.

[324] Maria L Kireeva, Kirill A Afonin, Bruce A Shapiro, and Mikhail Kashlev. Cotranscriptional production of chemically modified RNA nanoparticles. *Methods Mol. Biol.*, 1632:91–105, 2017.

[325] James C Taggart, Jean-Benoît Lalanne, and Gene-Wei Li. Quantitative control for stoichiometric protein synthesis. *Annu. Rev. Microbiol.*, 75:243–267, October 2021.

[326] Jong Bum Lee, Jinkee Hong, Daniel K Bonner, Zhiyong Poon, and Paula T Hammond. Self-assembled RNA interference microsponges for efficient siRNA delivery. *Nat. Mater.*, 11(4):316–322, February 2012.

[327] Pasquale Stano and Luisa Damiano. Synthetic cell research: Is technical progress leaving theoretical and epistemological investigations one step behind? *Front Robot AI*, 10:1143196, March 2023.

[328] Joseph N Zadeh, Conrad D Steenberg, Justin S Bois, Brian R Wolfe, Marshall B Pierce, Asif R Khan, Robert M Dirks, and Niles A Pierce. NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.*, 32(1):170–173, January 2011.

[329] Jean-Louis Mergny, Anne De Cian, Amar Ghelab, Barbara Saccà, and Laurent Lacroix. Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res.*, 33(1):81–94, January 2005.

[330] Dilara Grate and Charles Wilson. Laser-mediated, site-specific inactivation of RNA transcripts. *Proc. Natl. Acad. Sci. U. S. A.*, 96(11):6131–6136, May 1999.

[331] Hervé Isambert and Eric D Siggia. Modeling RNA folding paths with pseudoknots: Application to hepatitis delta virus ribozyme. *Proc. Natl. Acad. Sci. U. S. A.*, 97(12):6515–6520, June 2000.

[332] Hervé Isambert. The jerky and knotty dynamics of RNA. *Methods*, 49(2):189–196, October 2009.

[333] Dennis E Koppel. Analysis of macromolecular polydispersity in intensity correlation spectroscopy: the method of cumulants. *J. Chem. Phys.*, 1972.

[334] Barbara J Frisken. Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *Appl. Opt.*, 40(24):4087–4091, August 2001.

[335] Stephen W Provencher. *Makromol. Chem.*, 180(1):201–209, January 1979.

[336] Stephen W Provencher. CONTIN: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Comput. Phys. Commun.*, 27(3):229–242, September 1982.

[337] Zhenming Liu, Xiangke Zhang, Yajing Wang, Jin Shen, Xi Yuan, Tongtong Mu, Wei Liu, Changzhi Li, and Zongzheng Wang. Research on tikhonov regularization parameter selection in dynamic light scattering measurement of flowing particles. *J. Opt.*, 51(4):1038–1051, December 2022.

[338] Alison Rodger. Circular dichroism and chirality. In Gordon C K Roberts, editor, *Encyclopedia of Biophysics*, pages 311–313. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[339] Steven S Andrews and James Tretton. Physical principles of circular dichroism. *J. Chem. Educ.*, 97(12):4370–4376, December 2020.

[340] Tomoo Miyahara, Hiroshi Nakatsuji, and Hiroshi Sugiyama. Similarities and differences between RNA and DNA Double-Helical structures in circular dichroism spectroscopy: A SAC-CI study. *J. Phys. Chem. A*, 120(45):9008–9018, November 2016.

[341] Rafael Del Villar-Guerra, John O Trent, and Jonathan B Chaires. G-Quadruplex secondary structure obtained from circular dichroism spectroscopy. *Angew. Chem. Int. Ed Engl.*, 57(24):7171–7175, June 2018.

[342] Josué Carvalho, João A Queiroz, and Carla Cruz. Circular dichroism of G-Quadruplex: a laboratory experiment for the study of topology and ligand binding. *J. Chem. Educ.*, 94(10):1547–1551, October 2017.

[343] Frans C Groen, Ian T Young, and Guido Ligthart. A comparison of different focus functions for use in autofocus algorithms. *Cytometry*, 6(2):81–91, March 1985.

[344] O A Osibote, R Dendere, S Krishnan, and T S Douglas. Automated focusing in bright-field microscopy for tuberculosis detection. *J. Microsc.*, 240(2):155–163, November 2010.

[345] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, 9(7):676–682, June 2012.

[346] C H Li and P K S Tam. An iterative algorithm for minimum cross entropy thresholding. *Pattern Recognit. Lett.*, 19(8):771–776, June 1998.

[347] Nobuyuki Otsu. A threshold selection method from Gray-Level histograms. *IEEE Trans. Syst. Man Cybern.*, 9(1):62–66, January 1979.

[348] Laurent Thomas and Pierre Trehin. LauLauThom/MaskFromRois-Fiji: Masks from ROIs plugins for fiji - initial release, July 2021.

[349] Edward Gomes and James Shorter. The molecular language of membraneless organelles. *J. Biol. Chem.*, 294(18):7115–7127, May 2019.

[350] Amandine Molliex, Jamshid Temirov, Jihun Lee, Maura Coughlin, Anderson P Kanagaraj, Hong Joo Kim, Tanja Mittag, and J Paul Taylor. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*, 163(1):123–133, September 2015.

[351] Andrew S Lyon, William B Peeples, and Michael K Rosen. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.*, 22(3):215–235, March 2021.

[352] Benjamin S Schuster, Roshan Mammen Regy, Elliott M Dolan, Aishwarya Kanchi Ranganath, Nina Jovic, Sagar D Khare, Zheng Shi, and Jeetain Mittal. Biomolecular condensates: Sequence determinants of phase separation, microstructural organization, enzymatic activity, and material properties. *J. Phys. Chem. B*, 125(14):3441–3451, April 2021.

[353] Clifford P Brangwynne, Christian R Eckmann, David S Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A Hyman. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324(5935):1729–1732, June 2009.

[354] Clifford P Brangwynne. Phase transitions and size scaling of membrane-less organelles. *J. Cell Biol.*, 203(6):875–881, December 2013.

[355] Steffen Frey, Ralf P Richter, and Dirk Görlich. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science*, 314(5800):815–817, November 2006.

[356] Ankur Jain and Ronald D Vale. RNA phase transitions in repeat expansion disorders. *Nature*, 546(7657):243–247, June 2017.

[357] Stephanie C Weber. Sequence-encoded material properties dictate the structure and function of nuclear bodies. *Curr. Opin. Cell Biol.*, 46:62–71, June 2017.

[358] Yan Qiao, Mei Li, Richard Booth, and Stephen Mann. Predatory behaviour in synthetic protocell communities. *Nat. Chem.*, 9(2):110–119, February 2017.

[359] Camila Guindani, Lucas Caire da Silva, Shoupeng Cao, Tsvetomir Ivanov, and Katharina Landfester. Synthetic cells: From simple Bio-Inspired modules to sophisticated integrated systems. *Angew. Chem. Int. Ed Engl.*, 61(16):e202110855, April 2022.

[360] Marcello DeLuca, Ze Shi, Carlos E Castro, and Gaurav Arya. Dynamic DNA nanotechnology: toward functional nanoscale devices. *Nanoscale Horiz.*, 5(2):182–201, February 2020.

[361] Weiwei Guo, Xiu-Juan Qi, Ron Orbach, Chun-Hua Lu, Lina Freage, Iris Mironi-Harpaz, Dror Seliktar, Huang-Hao Yang, and Itamar Willner. Reversible ag(+)-crosslinked DNA hydrogels. *Chem. Commun.*, 50(31):4065–4068, April 2014.

[362] Jasmina Gačanin, Christopher V Synatschke, and Tanja Weil. Biomedical applications of DNA-based hydrogels. *Adv. Funct. Mater.*, 30(4):1906253, January 2020.

[363] Juewen Liu. Oligonucleotide -functionalized hydrogels as stimuli responsive materials and biosensors. *Soft Matter*, 7(15):6757–6767, 2011.

[364] Na Xu, Nannan Ma, Xiaotong Yang, Guixia Ling, Jia Yu, and Peng Zhang. Preparation of intelligent DNA hydrogel and its applications in biosensing. *Eur. Polym. J.*, 137:109951, August 2020.

[365] Yu Nishida, Shozo Ohtsuki, Yuki Araie, Yuka Umeki, Masayuki Endo, Tomoko Emura, Kumi Hidaka, Hiroshi Sugiyama, Yuki Takahashi, Yoshinobu Takakura, and Makiya Nishikawa. Self-assembling DNA hydrogel-based delivery of immunoinhibitory nucleic acids to immune cells. *Nanomedicine*, 12(1):123–130, January 2016.

[366] Nikhil Sood, Ankur Bhardwaj, Shuchi Mehta, and Abhinav Mehta. Stimuli-responsive hydrogels in drug delivery and tissue engineering. *Drug Deliv.*, 23(3):758–780, 2016.

[367] Michal Walczak, Ryan A Brady, Adrian Leathers, Jurij Kotar, and Lorenzo Di Michele. Influence of hydrophobic moieties on the crystallization of amphiphilic DNA nanostructures. *J. Chem. Phys.*, 158(8):084501, February 2023.

[368] Dominic McBrayer, Michelle Schoonover, Kimberly J Long, Ruby Escobedo, and Sean M Kerwin. N-Methylmesoporphyrin IX exhibits G-quadruplex-specific photocleavage activity. *Chembiochem*, 20(15):1924–1927, August 2019.

[369] Masayuki Morikawa, Katsuhito Kino, Takanori Oyoshi, Masayo Suzuki, Takanobu Kobayashi, and Hiroshi Miyazawa. Product analysis of photooxidation in isolated quadruplex DNA; 8-oxo-7,8-dihydroguanine and its oxidation product at 3'-g are formed instead of 2,5-diamino-4h-imidazol-4-one. *RSC Adv.*, 3(48):25694, 2013.

[370] Masayuki Morikawa, Katsuhito Kino, Takanori Oyoshi, Masayo Suzuki, Takanobu Kobayashi, and Hiroshi Miyazawa. Analysis of guanine oxidation products in double-stranded DNA and proposed guanine oxidation pathways in single-stranded, double-stranded or quadruplex DNA. *Biomolecules*, 4(1):140–159, February 2014.

[371] Keiko Kawauchi, Ryoto Urano, Natsuki Kinoshita, Shin Kuwamoto, Takeru Torii, Yoshiki Hashimoto, Shinya Taniguchi, Mitsuki Tsuruta, and Daisuke Miyoshi. Photosensitizers based on G-Quadruplex ligand for cancer photodynamic therapy. *Genes*, 11(11), November 2020.

[372] Jean Cadet and J Richard Wagner. DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Cold Spring Harb. Perspect. Biol.*, 5(2), February 2013.

[373] Giacomo Fabrini, Aisling Minard, Ryan A Brady, Marco Di Antonio, and Lorenzo Di Michele. Cation-responsive and photocleavable hydrogels from noncanonical amphiphilic DNA nanostructures. *Nano Lett.*, 22(2):602–611, January 2022.

[374] P Balagurumoorthy, S K Brahmachari, D Mohanty, M Bansal, and V Sasisekharan. Hairpin and parallel quartet structures for telomeric sequences. *Nucleic Acids Res.*, 20(15):4061–4067, August 1992.

[375] Bo Zhou, Yanyan Geng, Changdong Liu, Haitao Miao, Yaguang Ren, Naining Xu, Xiao Shi, Yingying You, Tunglun Lee, and Guang Zhu. Characterizations of distinct parallel and antiparallel G-quadruplexes formed by two-repeat ALS and FTD related GGGGCC sequence. *Sci. Rep.*, 8(1):2366, February 2018.

[376] Yingfu Li, Ronald Geyer, and Dipankar Sen. Recognition of anionic porphyrins by DNA aptamers. *Biochemistry*, 35(21):6911–6922, May 1996.

[377] Jinsong Ren and Jonathan B Chaires. Sequence and structural selectivity of nucleic acid binding ligands. *Biochemistry*, 38(49):16067–16075, December 1999.

[378] John M Nicoludis, Stephen T Miller, Philip D Jeffrey, Steven P Barrett, Paul R Rablen, Thomas J Lawton, and Liliya A Yatsunyk. Optimized end-stacking provides specificity of n-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA. *J. Am. Chem. Soc.*, 134(50):20446–20456, December 2012.

[379] Jacqueline R Wyatt, Peter W Davis, and Susan M Freier. Kinetics of g-quartet-mediated tetramer formation. *Biochemistry*, 35(24):8002–8008, June 1996.

[380] Caroline Bardin and Jean Louis Leroy. The formation pathway of tetramolecular G-quadruplexes. *Nucleic Acids Res.*, 36(2):477–488, February 2008.

[381] Jean-Louis Mergny and Dipankar Sen. DNA quadruple helices in nanotechnology. *Chem. Rev.*, 119(10):6290–6325, May 2019.

[382] Thomas E Ouldridge, Petr Sulc, Flavio Romano, Jonathan P K Doye, and Ard A Louis. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Res.*, 41(19):8886–8895, October 2013.

[383] Dietmar Pörschke. Cooperative nonenzymic base recognition II. thermodynamics of the helix-coil transition of oligoadenylic + oligouridylic acids. *Biopolymers*, 10(10):1989–2013, October 1971.

[384] Mary E Craig, Donald M Crothers, and Paul Doty. Relaxation kinetics of dimer formation by self complementary oligonucleotides. *J. Mol. Biol.*, 62(2):383–401, December 1971.

[385] Gnanapragasam Niranjani and Rajamanickam Murugan. Theory on the mechanism of DNA renaturation: Stochastic nucleation and zipping. *PLoS One*, 11(4):e0153172, April 2016.

[386] Chunlai Chen, Wenjuan Wang, Zhang Wang, Fang Wei, and Xin Sheng Zhao. Influence of secondary structure on kinetics and reaction mechanism of DNA hybridization. *Nucleic Acids Res.*, 35(9):2875–2884, April 2007.

[387] Gen He, Jie Li, Haina Ci, Chuanmin Qi, and Xuefeng Guo. Direct measurement of Single-Molecule DNA hybridization dynamics with Single-Base resolution. *Angew. Chem. Int. Ed Engl.*, 55(31):9036–9040, July 2016.

[388] Soma Mondal, Jyotsna Bhat, Jagannath Jana, Meghomukta Mukherjee, and Subhrangsu Chatterjee. Reverse Watson–Crick G–G base pair in G-quadruplex formation. *Mol. Biosyst.*, 12(1):18–22, 2016.

[389] Richard Stefl, Thomas E Cheatham, 3rd, Nad'a Spacková, Eva Fadrná, Imre Berger, Jaroslav Koca, and Jirí Sponer. Formation pathways of a guanine-quadruplex DNA revealed by molecular dynamics and thermodynamic analysis of the substates. *Biophys. J.*, 85(3):1787–1804, September 2003.

[390] Attila Ambrus, Ding Chen, Jixun Dai, Tiffanie Bialis, Roger A Jones, and Danzhou Yang. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.*, 34(9):2723–2735, May 2006.

[391] Ana Elisa Bergues-Pupo, J Ricardo Arias-Gonzalez, María Carmen Morón, Alessandro Fiasconaro, and Fernando Falo. Role of the central cations in the mechanical unfolding of DNA and RNA G-quadruplexes. *Nucleic Acids Res.*, 43(15):7638–7647, September 2015.

[392] Miguel Angel Aleman-Garcia, Ron Orbach, and Itamar Willner. Ion-responsive hemin-G-quadruplexes for switchable DNAzyme and enzyme functions. *Chemistry*, 20(19):5619–5624, May 2014.

[393] Ruth Aizen, Eyal Golub, Alexander Trifonov, Simcha Shimron, Angelica Niazov-Elkan, and Itamar Willner. G-quadruplex-stimulated optical and electrocatalytic DNA switches. *Small*, 11(30):3654–3658, August 2015.

[394] L Olejko, P J Cywiński, and I Bald. An ion-controlled four-color fluorescent telomeric switch on DNA origami structures. *Nanoscale*, 8(19):10339–10347, May 2016.

[395] Xia Liu, Junji Zhang, Michael Fadeev, Ziyuan Li, Verena Wulf, He Tian, and Itamar Willner. Chemical and photochemical DNA "gears" reversibly control stiffness, shape-memory, self-healing and controlled release properties of polyacrylamide hydrogels. *Chem. Sci.*, 10(4):1008–1016, 2019.

[396] I D Gulyakin, N A Oborotova, and V M Pechennikov. Solubilization of hydrophobic antitumor drugs (review). *Pharm. Chem. J.*, 48(3):209–213, June 2014.

[397] Swati Sareen, George Mathew, and Lincy Joseph. Improvement in solubility of poor water-soluble drugs by solid dispersion. *Int. J. Pharm. Investig.*, 2(1):12–17, January 2012.

[398] Ketan T Savjani, Anuradha K Gajjar, and Jignasa K Savjani. Drug solubility: importance and enhancement techniques. *ISRN Pharm.*, 2012:195727, July 2012.

[399] Sandeep Kalepu and Vijaykumar Nekkanti. Insoluble drug delivery strategies: review of recent advances and business prospects. *Acta Pharm Sin B*, 5(5):442–453, September 2015.

[400] Cyril M Sanders. Human pif1 helicase is a G-quadruplex DNA-binding protein with G-quadruplex DNA-unwinding activity. *Biochem. J*, 430(1):119–128, August 2010.

[401] Colin G Wu and Maria Spies. G-quadruplex recognition and remodeling by the FANCJ helicase. *Nucleic Acids Res.*, 44(18):8742–8753, October 2016.

[402] Markus Sauer and Katrin Paeschke. G-quadruplex unwinding helicases and their function in vivo. *Biochem. Soc. Trans.*, 45(5):1173–1182, October 2017.

[403] Alicia K Byrd, Matthew R Bell, and Kevin D Raney. Pif1 helicase unfolding of G-quadruplex DNA is highly dependent on sequence and reaction conditions. *J. Biol. Chem.*, 293(46):17792–17802, November 2018.

[404] Samuel O Thier. Potassium physiology. *Am. J. Med.*, 80(4A):3–7, April 1986.

[405] Arun Richard Chandrasekaran. Nuclease resistance of DNA nanostructures. *Nat Rev Chem*, 5(4):225–239, April 2021.

[406] Jaeseung Hahn, Shelley F J Wickham, William M Shih, and Steven D Perrault. Addressing the instability of DNA nanostructures in tissue culture. *ACS Nano*, 8(9):8765–8775, September 2014.

[407] Shih-Ting Wang, Melissa A Gray, Sunting Xuan, Yiyang Lin, James Byrnes, Andy I Nguyen, Nevena Todorova, Molly M Stevens, Carolyn R Bertozzi, Ronald N Zuckermann, and Oleg Gang. DNA origami protection and molecular interfacing through engineered sequence-defined peptoids. *Proc. Natl. Acad. Sci. U. S. A.*, 117(12):6339–6348, March 2020.

[408] Michal Walczak, Leonardo Mancini, Jiayi Xu, Federica Raguseo, Jurij Kotar, Pietro Cicuta, and Lorenzo Di Michele. A synthetic signalling network imitating the action of immune cells in response to bacterial metabolism. *Adv. Mater.*, page e2301562, May 2023.

[409] Steven D Perrault and William M Shih. Virus-inspired membrane encapsulation of DNA nanostructures to achieve in vivo stability. *ACS Nano*, 8(5):5132–5140, May 2014.

[410] Zhao Zhang, Yang Yang, Frederic Pincet, Marc C Llaguno, and Chenxiang Lin. Placing and shaping liposomes with reconfigurable DNA nanocages. *Nat. Chem.*, 9(7):653–659, June 2017.

[411] Layla Malouf, Diana A Tanase, Giacomo Fabrini, Miguel Paez-Perez, Adrian Leathers, Michael J Booth, and Lorenzo Di Michele. Sculpting DNA-based synthetic cells through phase separation and phase-targeted activity. March 2023.

[412] Sungwoo Jeong, Huong Thanh Nguyen, Chang Ho Kim, Mai Nguyet Ly, and Kwanwoo Shin. Toward artificial cells: Novel advances in energy conversion and cellular motility. *Adv. Funct. Mater.*, 30(11):1907182, March 2020.

[413] Federica Raguseo, Souroprobho Chowdhury, Aisling Minard, and Marco Di Antonio. Chemical-biology approaches to probe DNA and RNA G-quadruplex structures in the genome. *Chem. Commun.*, 56(9):1317–1324, January 2020.

[414] Satish Kumar Gupta and Ming Guo. Equilibrium and out-of-equilibrium mechanics of living mammalian cytoplasm. *J. Mech. Phys. Solids*, 107:284–293, October 2017.

[415] Benjamin D Knapp and Kerwyn Casey Huang. The effects of temperature on cellular physiology. *Annu. Rev. Biophys.*, 51:499–526, May 2022.

[416] Shih-Ting Wang, Brian Minevich, Jianfang Liu, Honghu Zhang, Dmytro Nykypanchuk, James Byrnes, Wu Liu, Lev Bershadsky, Qun Liu, Tong Wang, Gang Ren, and Oleg Gang. Designed and biologically active protein lattices. *Nat. Commun.*, 12(1):3702, June 2021.

[417] Robert Schreiber, Ngoc Luong, Zhiyuan Fan, Anton Kuzyk, Philipp C Nickels, Tao Zhang, David M Smith, Bernard Yurke, Wan Kuang, Alexander O Govorov, and Tim Liedl. Chiral plasmonic DNA nanostructures with switchable circular dichroism. *Nat. Commun.*, 4:2948, 2013.

[418] Chi Chen, Xin Luo, Alexander E K Kaplan, Moungi G Bawendi, Robert J Macfarlane, and Mark Bathe. Ultrafast dense DNA functionalization of quantum dots and rods for scalable 2D array fabrication with nanoscale precision. *Sci Adv*, 9(32):eadh8508, August 2023.

[419] Dmytro Nykypanchuk, Mathew M Maye, Daniel van der Lelie, and Oleg Gang. DNA-guided crystallization of colloidal nanoparticles. *Nature*, 451(7178):549–552, January 2008.

[420] Lorenzo Di Michele, Francesco Varrato, Jurij Kotar, Simon H Nathan, Giuseppe Foffi, and Erika Eiser. Multistep kinetic self-assembly of DNA-coated colloids. *Nat. Commun.*, 4:2007, 2013.

[421] Lucia Parolini, Jurij Kotar, Lorenzo Di Michele, and Bortolo M Mognetti. Controlling Self-Assembly kinetics of DNA-Functionalized liposomes using toehold exchange mechanism. *ACS Nano*, 10(2):2392–2398, February 2016.

[422] Ewan K S McRae, Helena Østergaard Rasmussen, Jianfang Liu, Andreas Bøggild, Michael T A Nguyen, Nestor Sampedro Vallina, Thomas Boesen, Jan Skov Pedersen, Gang Ren, Cody Geary, and Ebbe Sloth Andersen. Structure, folding and flexibility of co-transcriptional RNA origami. *Nat. Nanotechnol.*, February 2023.

[423] Simon Alberti, Amy Gladfelter, and Tanja Mittag. Considerations and challenges in studying Liquid-Liquid phase separation and biomolecular condensates. *Cell*, 176(3):419–434, January 2019.

[424] Avinash Patel, Hyun O Lee, Louise Jawerth, Shovamayee Maharana, Marcus Jahnel, Marco Y Hein, Stoyno Stoynov, Julia Mahamid, Shambaditya Saha, Titus M Franzmann, Andrej Pozniakovski, Ina Poser, Nicola Maghelli, Loic A Royer, Martin Weigert, Eugene W Myers, Stephan Grill, David Drechsel, Anthony A Hyman, and Simon Alberti. A Liquid-to-Solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell*, 162(5):1066–1077, August 2015.

[425] Wade Borcherds, Anne Bremer, Madeleine B Borgia, and Tanja Mittag. How do intrinsically disordered protein regions encode a driving force for liquid-liquid phase separation? *Curr. Opin. Struct. Biol.*, 67:41–50, April 2021.

[426] Masaru Yoshikawa, Tatsuyuki Yoshii, Masahiro Ikuta, and Shinya Tsukiji. Synthetic protein condensates that inducibly recruit and release protein activity in living cells. *J. Am. Chem. Soc.*, 143(17):6434–6446, May 2021.

[427] Mariane Gonçalves-Kulik, Pablo Mier, Kristina Kastano, Juan Cortés, Pau Bernadó, Friederike Schmid, and Miguel A Andrade-Navarro. Low complexity induces structure in protein regions predicted as intrinsically disordered. *Biomolecules*, 12(8), August 2022.

[428] Erin M Langdon, Yupeng Qiu, Amirhossein Ghanbari Niaki, Grace A McLaughlin, Chase A Weidmann, Therese M Gerbich, Jean A Smith, John M Crutchley, Christina M Termini, Kevin M Weeks, Sua Myong, and Amy S Gladfelter. mRNA structure determines specificity of a polyq-driven phase separation. *Science*, 360(6391):922–927, May 2018.

[429] John S Mattick. Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Rep.*, 2(11):986–991, November 2001.

[430] Takeshi Chujo and Tetsuro Hirose. Nuclear bodies built on architectural long non-coding RNAs: Unifying principles of their construction and function. *Mol. Cells*, 40(12):889–896, December 2017.

[431] Run-Wen Yao, Yang Wang, and Ling-Ling Chen. Cellular functions of long noncoding RNAs. *Nat. Cell Biol.*, 21(5):542–551, May 2019.

[432] Jitendra Thakur and Steven Henikoff. Architectural RNA in chromatin organization. *Biochem. Soc. Trans.*, 48(5):1967–1978, October 2020.

[433] Kensuke Ninomiya and Tetsuro Hirose. Short tandem Repeat-Enriched architectural RNAs in nuclear bodies: Functions and associated diseases. *Noncoding RNA*, 6(1), February 2020.

[434] Giacomo Fabrini, Sabrina Pia Nuccio, Jaimie Marie Stewart, Shiyi Lin, Anli Tang, Paul W K Rothemund, Elisa Franco, Marco Di Antonio, and Lorenzo Di Michele. Co-transcriptional production of programmable RNA condensates and synthetic organelles. *bioRxiv, doi: 10.1101/2023.10.06.561174*, October 2023.

[435] Silvia Biffi, Roberto Cerbino, Francesca Bomboi, Elvezia Maria Paraboschi, Rosanna Asselta, Francesco Sciortino, and Tommaso Bellini. Phase behavior and critical activated dynamics of limited-valence DNA nanostars. *Proc. Natl. Acad. Sci. U. S. A.*, 110(39):15633–15637, September 2013.

[436] Siddharth Agarwal, Dino Osmanovic, Melissa A Klocke, and Elisa Franco. The growth rate of DNA condensate droplets increases with the size of participating subunits. *ACS Nano*, 16(8):11842–11851, August 2022.

[437] Dietmar Pörschke. Thermodynamic and kinetic parameters of an oligonucleotide hairpin helix. *Biophys. Chem.*, 1(5):381–386, May 1974.

[438] Peinan Zhao, Wenbing Zhang, and Shi-Jie Chen. Cotranscriptional folding kinetics of ribonucleic acid secondary structures. *J. Chem. Phys.*, 135(24):245101, December 2011.

[439] Brittany Rauzan, Elizabeth McMichael, Rachel Cave, Lesley R Sevcik, Kara Ostrosky, Elisabeth Whitman, Rachel Stegemann, Audra L Sinclair, Martin J Serra, and Alice A Deckert. Kinetics and thermodynamics of DNA, RNA, and hybrid duplex formation. *Biochemistry*, 52(5):765–772, February 2013.

[440] Stefan Badelt, Ronny Lorenz, and Ivo L Hofacker. DrTransformer: heuristic cotranscriptional RNA folding using the nearest neighbor energy model. *Bioinformatics*, 39(1), January 2023.

[441] Nathaniel Conrad, Tynan Kennedy, Deborah K Fygenson, and Omar A Saleh. Increasing valence pushes DNA nanostar networks to the isostatic point. *Proc. Natl. Acad. Sci. U. S. A.*, 116(15):7238–7243, April 2019.

[442] Omar A Saleh, Byoung-Jin Jeon, and Tim Liedl. Enzymatic degradation of liquid droplets of DNA is modulated near the phase boundary. *Proc. Natl. Acad. Sci. U. S. A.*, 117(28):16160–16166, July 2020.

[443] Jeremy S Paige, Thinh Nguyen-Duc, Wenjiao Song, and Samie R Jaffrey. Fluorescence imaging of cellular metabolites with RNA. *Science*, 335(6073):1194, March 2012.

[444] Wenjiao Song, Rita L Strack, and Samie R Jaffrey. Imaging bacterial protein expression using genetically encoded RNA sensors. *Nat. Methods*, 10(9):873–875, September 2013.

[445] Robert E Martell, Joseph R Nevins, and Bruce A Sullenger. Optimizing aptamer activity for gene therapy applications using expression cassette SELEX. *Mol. Ther.*, 6(1):30–34, July 2002.

[446] Dan Shu, Emil F Khisamutdinov, Le Zhang, and Peixuan Guo. Programmable folding of fusion RNA in vivo and in vitro driven by pRNA 3WJ motif of phi29 DNA packaging motor. *Nucleic Acids Res.*, 42(2):e10, January 2014.

[447] Qian Hou and Samie R Jaffrey. Synthetic biology tools to promote the folding and function of RNA aptamers in mammalian cells. *RNA Biol.*, 20(1):198–206, January 2023.

[448] Michael Chamberlin and Janet Ring. 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(6):2235–2244, March 1973.

[449] I M Lifshitz and V V Slyozov. The kinetics of precipitation from supersaturated solid solutions. *J. Phys. Chem. Solids*, 19(1):35–50, April 1961.

[450] Daniel A Beysens. Kinetics and morphology of phase separation in fluids: The role of droplet coalescence. *Physica A: Statistical Mechanics and its Applications*, 239(1):329–339, May 1997.

[451] Celeste Sagui and Martin Grant. Theory of nucleation and growth during phase separation. *Phys. Rev. E*, 59(4):4175–4187, April 1999.

[452] Zsuzsa Sárkány, Fernando Rocha, Anna Bratek-Skicki, Peter Tompa, Sandra Macedo-Ribeiro, and Pedro M Martins. Quantification of surface tension effects and Nucleation-and-Growth rates during Self-Assembly of biological condensates. *Adv. Sci.*, page e2301501, June 2023.

[453] Nadia A Erkamp, Tomas Sneideris, Hannes Ausserwöger, Daoyuan Qian, Seema Qamar, Jonathon Nixon-Abell, Peter St George-Hyslop, Jeremy D Schmit, David A Weitz, and Tuomas P J Knowles. Spatially non-uniform condensates emerge from dynamically arrested phase separation. *Nat. Commun.*, 14(1):684, February 2023.

[454] Marina Feric, Azadeh Sarfallah, Furqan Dar, Dmitry Temiakov, Rohit V Pappu, and Tom Misteli. Mesoscale structure–function relationships in mitochondrial transcriptional condensates. *Proceedings of the National Academy of Sciences*, 119(41):e2207303119, 2022.

[455] Priya R Banerjee, Anthony N Milin, Mahdi Muhammad Moosa, Paulo L Onuchic, and Ashok A Deniz. Reentrant phase transition drives dynamic substructure formation in ribonucleoprotein droplets. *Angew. Chem. Int. Ed Engl.*, 56(38):11354–11359, September 2017.

[456] Omar A Saleh, Sam Wilken, Todd M Squires, and Tim Liedl. Vacuole dynamics and popping-based motility in liquid droplets of DNA. *Nat. Commun.*, 14(1):3574, June 2023.

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

[458] Pierre Levitz. Toolbox for 3D imaging and modeling of porous media: Relationship with transport properties. *Cem. Concr. Res.*, 37(3):351–359, March 2007.

[459] Vincent Testard, Ludovic Berthier, and Walter Kob. Influence of the glass transition on the liquid-gas spinodal decomposition. *Phys. Rev. Lett.*, 106(12):125702, March 2011.

[460] Jiahui Wu, Nina Svensen, Wenjiao Song, Hyaeyeong Kim, Sulei Zhang, Xing Li, and Samie R Jaffrey. Self-Assembly of intracellular multivalent RNA complexes using dimeric corn and beetroot aptamers. *J. Am. Chem. Soc.*, 144(12):5471–5477, March 2022.

[461] Adam Klosin, Florian Oltsch, Tyler Harmon, Alf Honigmann, Frank Jülicher, Anthony A Hyman, and Christoph Zechner. Phase separation provides a mechanism to reduce noise in cells. *Science*, 367(6476):464–468, January 2020.

[462] Vladimir N Uversky. Chapter 17 - MLOstasis: liquid–liquid phase separation and biomolecular condensates in cell competition, fitness, and aging. In Vladimir N Uversky, editor, *Droplets of Life*, pages 485–504. Academic Press, January 2023.

[463] Marina Feric, Nilesh Vaidya, Tyler S Harmon, Diana M Mitrea, Lian Zhu, Tiffany M Richardson, Richard W Kriwacki, Rohit V Pappu, and Clifford P Brangwynne. Coexisting liquid phases underlie nucleolar subcompartments. *Cell*, 165(7):1686–1697, June 2016.

[464] Abolfazl Akbarzadeh, Rogaie Rezaei-Sadabady, Soodabeh Davaran, Sang Woo Joo, Nosratollah Zarghami, Younes Hanifehpour, Mohammad Samiei, Mohammad Kouhi, and Kazem Nejati-Koshki. Liposome: classification, preparation, and applications. *Nanoscale Res. Lett.*, 8(1):102, February 2013.

[465] Nan-Nan Deng, Maaruthy Yelleswarapu, Lifei Zheng, and Wilhelm T S Huck. Microfluidic assembly of monodisperse vesosomes as artificial cell models. *J. Am. Chem. Soc.*, 139(2):587–590, January 2017.

[466] Nan-Nan Deng and Wilhelm T S Huck. Microfluidic formation of monodisperse coacervate organelles in liposomes. *Angew. Chem. Int. Ed Engl.*, 56(33):9736–9740, August 2017.

[467] Barbara Haller, Kerstin Göpfrich, Martin Schröter, Jan-Willi Janiesch, Ilia Platzman, and Joachim P Spatz. Charge-controlled microfluidic formation of lipid-based single- and multicompartment systems. *Lab Chip*, 18(17):2665–2674, August 2018.

[468] Valentin Romanov, John McCullough, Bruce K Gale, and Adam Frost. A tunable microfluidic device enables cargo encapsulation by cell- or Organelle-Sized lipid vesicles comprising asymmetric lipid bilayers. *Adv Biosyst*, 3(7), July 2019.

[469] Nelson F Morales-Penningston, Jing Wu, Elaine R Farkas, Shih Lin Goh, Tatyana M Konyakhina, Judy Y Zheng, Watt W Webb, and Gerald W Feigenson. GUV preparation and imaging: minimizing artifacts. *Biochim. Biophys. Acta*, 1798(7):1324–1332, July 2010.

[470] Peter Walde, Katia Cosentino, Helen Engel, and Pasquale Stano. Giant vesicles: preparations and applications. *Chembiochem*, 11(7):848–865, May 2010.

[471] Jan Lipfert, Gary M Skinner, Johannes M Keegstra, Toivo Hensgens, Tessa Jager, David Dulin, Mariana Köber, Zhongbo Yu, Serge P Donkers, Fang-Chieh Chou, Rhiju Das, and Nynke H Dekker. Double-stranded RNA under force and torque: similarities to and striking differences from double-stranded DNA. *Proc. Natl. Acad. Sci. U. S. A.*, 111(43):15408–15413, October 2014.

[472] Corwin Hansch, Albert Leo, and David Hoekman. *Exploring QSAR: hydrophobic, electronic, and steric constants*, volume 2. American Chemical Society Washington, DC, 1995.

[473] Elisa Franco, Eike Friedrichs, Jongmin Kim, Ralf Jungmann, Richard Murray, Erik Winfree, and Friedrich C Simmel. Timing molecular motion and production with a synthetic transcriptional clock. *Proc. Natl. Acad. Sci. U. S. A.*, 108(40):E784–93, October 2011.

[474] Sheng Mao, Derek Kuldinow, Mikko P Haataja, and Andrej Košmrlj. Phase behavior and morphology of multicomponent liquid mixtures. *Soft Matter*, 15(6):1297–1311, February 2019.

[475] Tiemei Lu and Evan Spruijt. Multiphase complex coacervate droplets. *J. Am. Chem. Soc.*, 142(6):2905–2914, February 2020.

[476] Joel Berry, Clifford P Brangwynne, and Mikko Haataja. Physical principles of intracellular organization via active and passive phase transitions. *Rep. Prog. Phys.*, 81(4):046601, April 2018.

[477] Lin-Ge Li and Zhonghuai Hou. Theoretical modelling of liquid-liquid phase separation: from particle-based to field-based simulation. *Biophys Rep*, 8(2):55–67, April 2022.

[478] Byoung-Jin Jeon, Dan T Nguyen, and Omar A Saleh. Sequence-Controlled adhesion and microemulsification in a Two-Phase system of DNA liquid droplets. *J. Phys. Chem. B*, 124(40):8888–8895, October 2020.

[479] Stephanie C Weber and Clifford P Brangwynne. Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Curr. Biol.*, 25(5):641–646, March 2015.

[480] Anthony D Keefe, Supriya Pai, and Andrew Ellington. Aptamers as therapeutics. *Nat. Rev. Drug Discov.*, 9(7):537–550, July 2010.

[481] Katherine Germer, Marissa Leonard, and Xiaoting Zhang. RNA aptamers and their therapeutic and diagnostic applications. *Int. J. Biochem. Mol. Biol.*, 4(1):27–40, March 2013.

[482] Nils Bjerregaard, Peter A Andreasen, and Daniel M Dupont. Expected and unexpected features of protein-binding RNA aptamers. *Wiley Interdiscip. Rev. RNA*, 7(6):744–757, November 2016.

[483] Aoi Mameuda, Masahiro Takinoue, and Koki Kamiya. Control of reversible formation and dispersion of the three enzyme networks integrating DNA computing. *Anal. Chem.*, 95(25):9548–9554, June 2023.

[484] Mikael V Garabedian, Wentao Wang, Jorge B Dabdoub, Michelle Tong, Reese M Caldwell, William Benman, Benjamin S Schuster, Alexander Deiters, and Matthew C Good. Designer membraneless organelles sequester native factors for control of cell behavior. *Nat. Chem. Biol.*, 17(9):998–1007, September 2021.

[485] Faqing Huang, Jun He, Yilin Zhang, and Yanlin Guo. Synthesis of biotin-AMP conjugate for 5' biotin labeling of RNA through one-step in vitro transcription. *Nat. Protoc.*, 3(12):1848–1861, 2008.

[486] Michael J Booth, Vanessa Restrepo Schild, Alexander D Graham, Sam N Olof, and Hagan Bayley. Light-activated communication in synthetic tissues. *Sci Adv*, 2(4):e1600056, April 2016.

[487] Naresh Yandrapalli, Julien Petit, Oliver Bäumchen, and Tom Robinson. Surfactant-free production of biomimetic giant unilamellar vesicles using PDMS-based microfluidics. *Commun Chem*, 4(1):100, June 2021.

[488] Gakushi Tsuji, Takeshi Sunami, and Norikazu Ichihashi. Production of giant unilamellar vesicles by the water-in-oil emulsion-transfer method without high internal concentrations of sugars. *J. Biosci. Bioeng.*, 126(4):540–545, October 2018.

[489] Hiroyuki Asanuma, Xingguo Liang, Hidenori Nishioka, Daijiro Matsunaga, Mingzhe Liu, and Makoto Komiyama. Synthesis of azobenzene-tethered DNA for reversible photo-regulation of DNA functions: hybridization and transcription. *Nat. Protoc.*, 2(1):203–212, 2007.

[490] Jefferson M Smith, Denis Hartmann, and Michael J Booth. Engineering cellular communication between light-activated synthetic cells and bacteria. *Nat. Chem. Biol.*, July 2023.

[491] Johanna Meyer, Lars-Erik Meyer, and Selin Kara. Enzyme immobilization in hydrogels: A perfect liaison for efficient and sustainable biocatalysis. *Eng. Life Sci.*, 22(3-4):165–177, March 2022.

[492] Christopher T Walsh and Bradley S Moore. Enzymatic cascade reactions in biosynthesis. *Angew. Chem. Int. Ed Engl.*, 58(21):6846–6879, May 2019.

[493] Gregory G Fuller and John K Kim. Compartmentalization and metabolic regulation of glycolysis. *J. Cell Sci.*, 134(20), October 2021.

[494] Denis L J Lafontaine, Joshua A Riback, Rümeyza Bascetin, and Clifford P Brangwynne. The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.*, 22(3):165–182, March 2021.

[495] M A J Roberts, R M Cranenburgh, M P Stevens, and P C F Oyston. Synthetic biology: biology by design. *Microbiology*, 159(Pt 7):1219–1220, July 2013.

[496] Christopher A Azaldegui, Anthony G Vecchiarelli, and Julie S Biteen. The emergence of phase separation as an organizing principle in bacteria. *Biophys. J.*, 120(7):1123–1138, April 2021.

[497] Xin Jin, Ji-Eun Lee, Charley Schaefer, Xinwei Luo, Adam J M Wollman, Alex L Payne-Dwyer, Tian Tian, Xiaowei Zhang, Xiao Chen, Yingxing Li, Tom C B McLeish, Mark C Leake, and Fan Bai. Membraneless organelles formed by liquid-liquid phase separation increase bacterial fitness. *Sci Adv*, 7(43):eabh2929, October 2021.

[498] Keren Lasker, Steven Boeynaems, Vinson Lam, Daniel Scholl, Emma Stainton, Adam Briner, Maarten Jacquemyn, Dirk Daelemans, Ashok Deniz, Elizabeth Villa, Alex S Holehouse, Aaron D Gitler, and Lucy Shapiro. The material properties of a bacterial-derived biomolecular condensate tune biological function in natural and synthetic systems. *Nat. Commun.*, 13(1):5643, September 2022.

[499] Vincent Mukwaya, Stephen Mann, and Hongjing Dou. Chemical communication at the synthetic cell/living cell interface. *Commun Chem*, 4(1):161, November 2021.

[500] Zhaolin Xue, Kewei Ren, Rigumula Wu, Zhining Sun, Ru Zheng, Qian Tian, Ahsan Ausaf Ali, Lan Mi, and Mingxu You. Targeted RNA condensation in living cells via genetically encodable triplet repeat tags. *Nucleic Acids Res.*, 51(16):8337–8347, September 2023.

[501] A Oktari, Y Supriatin, M Kamal, and H Syafrullah. The bacterial endospore stain on schaeffer fulton using variation of methylene blue solution. *J. Phys. Conf. Ser.*, 812(1):012066, February 2017.

[502] Michael Lynch and Georgi K Marinov. The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci. U. S. A.*, 112(51):15690–15695, December 2015.

[503] Simone Furgeri Godinho Vilela, Juliana Campos Junqueira, Junia Oliveira Barbosa, Marta Majewski, Egberto Munin, and Antonio Olavo Cardoso Jorge. Photodynamic inactivation of staphylococcus aureus and escherichia coli biofilms by malachite green and phenothiazine dyes: an in vitro study. *Arch. Oral Biol.*, 57(6):704–710, June 2012.

Appendix A

Nanostructure and oligonucleotide sequences

A.1 Chapter 3

Table A.1 All DNA strands used in Chapter 3 and their sequences according to 5'-3' convention. Abbreviations: 5Alex488N – 5' Alexa488 fluorophore modification; 3CholTEG – 3' CholesterolTEG modification, TEG – triethyleneglycol.

Strand	Sequence
Arm_duplex	CGACGCCGTGACGCGTTGGACGCACTCG
Arm_duplex_A_Alexa488	/5Alex488N/ACGACGCCGTGACGCGTTGGACGCACTCG
TerminalStrand_NoChol	GCGTCACGGCGTCGAA
TerminalStrand_Chol	GCGTCACGGCGTCGAA/3CholTEG/
Arm_NoGs_7T	TTTTTTTTTCGAGTGCGTCCAAC
Arm_NoGs_9T	TTTTTTTTTTTCGAGTGCGTCCAAC
Arm_NoGs_11T	TTTTTTTTTTTTTCGAGTGCGTCCAAC
Arm_NoGs_13T	TTTTTTTTTTTTTTTCGAGTGCGTCCAAC
Arm_NoGs_14T (5G-noG)	TTTTTTTTTTTTTTTTCGAGTGCGTCCAAC
Arm_NoGs_16T (6G-noG)	TTTTTTTTTTTTTTTTTCGAGTGCGTCCAAC

Continues on next page

Table A.1 – continued from previous page.

Strand	**Sequence**
Arm_4G3T	TGGGGTTTCGAGTGCGTCCAAC
Arm_4G5T	TGGGGTTTTTCGAGTGCGTCCAAC
Arm_4G7T	TGGGGTTTTTTTCGAGTGCGTCCAAC
Arm_4G9T	TGGGGTTTTTTTTTTCGAGTGCGTCCAAC
Arm_5G3T	TGGGGGTTTCGAGTGCGTCCAAC
Arm_5G5T	TGGGGGTTTTTCGAGTGCGTCCAAC
Arm_5G7T	TGGGGGTTTTTTTCGAGTGCGTCCAAC
Arm_5G9T	TGGGGGTTTTTTTTTTCGAGTGCGTCCAAC
Arm_6G3T	TGGGGGGTTTCGAGTGCGTCCAAC
Arm_6G5T	TGGGGGGTTTTTCGAGTGCGTCCAAC
Arm_6G7T	TGGGGGGTTTTTTTTCGAGTGCGTCCAAC
Arm_6G9T	TGGGGGGTTTTTTTTTTCGAGTGCGTCCAAC

A.2 Chapters 4 and 5

Table A.2 Sequences of all RNA nanostars used in Chapters 4 and 5, reported according to 5'-3' convention.

Strand	Sequence
$\overline{A}$	GCA CAG UGC UAU GAG UGU GCA CGG GAU CCC GAC UGG CGC CAA CGA ACA AGG UGC CAG GUA ACG AAU GGA UCC UGU GCU GCA CAU UAG AGU CGC UGU AUG ACC CAU AGC ACA AGG GUC GUA CAG CGG CUC UAG UGU GCU CGC GUG CCU CAG AGG ACC UGU CAC CAU GGC GAA AGG UGA UAG GUC CUU UGA GGU ACG CGU CAC UCG UAG CAU UGU GCC UGU CUC CAA CGG CAA AGG AGA UAG
A	GCA CAG UGC UAU GAG UGU GCA CGG GAU CCC GAC UGG CCG CAU CGC GAA AGU GGC CAG GUA ACG AAU GGA UCC UGU GCU GCA CAU UAG AGU CGC UGU AUG ACC CAU CGC GAA AGG GUC GUA CAG CGG CUC UAG UGU GCU CGC GUG CCU CAG AGG ACC UGU CAC CAU CGC GAA AGG UGA UAG GUC CUU UGA GGU ACG CGU CAC UCG UAG CAU UGU GCC UGU CUC CAU CGC GAA AGG AGA UAG
$\overline{B}$	GCA CAG UGC UAU GAG UGU CGC GAC GGA GAC GGU CGG GUC CAG AUA GGC CAU GGC GAA AGG UCU AUC UGU CGA GUA GAG UGU GGG CUC CGU CGC GUG CAC AUU AGA GUC GCU GUA UGA CCC AUA GCA CAA GGG UCG UAC AGC GGC UCU AGU GUG CUG CAC AGU GUC UGU GCG ACU GCA CCC AAC GAA CAA GGG UGU AGU CGC AUA GAC AUU GUG CUC ACU CGU AGC AUU GUG CCU GUC UCC AAC GGC AAA GGA GAU AG
B	GCA CAG UGC UAU GAG UGU CGC GAC GGA GAC GGU CGG GUC CAG AUA GGC CAG UCG ACA AGG UCU AUC UGU CGA GUA GAG UGU GGG CUC CGU CGC GUG CAC AUU AGA GUC GCU GUA UGC CAC AGU CGA CAA GUG GCG UAC AGC GGC UCU AGU GUG CUG CAC AGU GUC UGU GCG ACU GCA CCC AGU CGA CAA GGG UGU AGU CGC AUA GAC AUU GUG CUC ACU CGU AGC AUU GUG CCU GUC UCC AGU CGA CAA GGA GAU AG
C	GCA CAG UGC UAU GAG UGU CGC GAC GGA GAC GGU CGG GUC CAG AUA GGC CAG GUA CCA AGG UCU AUC UGU CGA GUA GAG UGU GGG CUC CGU CGC GUG CAC AUU AGA GUC GCU GUA UGC CAC AGG UAC CAA GUG GCG UAC AGC GGC UCU AGU GUG CUG CAC GGG AUC CCG ACU GGC CGC AGG UAC CAA GUG GCC AGG UAA CGA AUG GAU CCU GUG CUC ACU CGU AGC AUU GUG CCU GUC UCC AGG UAC CAA GGA GAU AG

Continues on next page

Table A.2 – continued from previous page.

Strand	Sequence
C̃	GCA CAG UGC UAU GAG UGU GCA CAG UGU CUG UGC GAC UGC ACG CAG GUA CCA AGC GUG UAG UCG CAU AGA CAU UGU GCU GCA CAU UAG AGU CGC UGU AUG ACC CAG GUA CCA AGG GUC GUA CAG CGG CUC UAG UGU GCU CGC GUG CCU CAG AGG ACC UGU CAG CAG GUA CCA AGC UGA UAG GUC CUU UGA GGU ACG CGU CAC UCG UAG CAU UGU GCC UGU CUC CAG GUA CCA AGG AGA UAG
L	GCA CAG UGC UAU GAG UGU GCA CAG UGU CUG UGC GAC UGC ACC CAU CGC GAA AGG GUG UAG UCG CAU AGA CAU UGU GCU GCA CAU UAG AGU CGC UGU AUG AGG GAU CGC GAA ACC CUC GUA CAG CGG CUC UAG UGU GCU CGC GUG CCU CAG AGG ACC UGU CAC CAG UCG ACA AGG UGA UAG GUC CUU UGA GGU ACG CGU CAC UCG UAG CAU UGU GCC UGU CUC CAG UCG ACA AGG AGA UAG

Table A.3 Sequences of all nontemplate strands of dsDNA templates coding for RNA nanostars used in Chapters 4 and 5, reported according to 5'-3' convention.

Strand	Sequence
$\overline{\text{A}}$-T	TTC TAA TAC GAC TCA CTA TAG CAC AGT GCT ATG AGT GTG CAC GGG ATC CCG ACT GGC GCC AAC GAA CAA GGT GCC AGG TAA CGA ATG GAT CCT GTG CTG CAC ATT AGA GTC GCT GTA TGA CCC ATA GCA CAA GGG TCG TAC AGC GGC TCT AGT GTG CTC GCG TGC CTC AGA GGA CCT GTC ACC ATG GCG AAA GGT GAT AGG TCC TTT GAG GTA CGC GTC ACT CGT AGC ATT GTG CCT GTC TCC AAC GGC AAA GGA GAT AG
A-T	TTC TAA TAC GAC TCA CTA TAG CAC AGT GCT ATG AGT GTG CAC GGG ATC CCG ACT GGC CGC ATC GCG AAA GTG GCC AGG TAA CGA ATG GAT CCT GTG CTG CAC ATT AGA GTC GCT GTA TGA CCC ATC GCG AAA GGG TCG TAC AGC GGC TCT AGT GTG CTC GCG TGC CTC AGA GGA CCT GTC ACC ATC GCG AAA GGT GAT AGG TCC TTT GAG GTA CGC GTC ACT CGT AGC ATT GTG CCT GTC TCC ATC GCG AAA GGA GAT AG
$\overline{\text{B}}$-T	TTC TAA TAC GAC TCA CTA TAG CAC AGT GCT ATG AGT GTC GCG ACG GAG ACG GTC GGG TCC AGA TAG GCC ATG GCG AAA GGT CTA TCT GTC GAG TAG AGT GTG GGC TCC GTC GCG TGC ACA TTA GAG TCG CTG TAT GAC CCA TAG CAC AAG GGT CGT ACA GCG GCT CTA GTG TGC TGC ACA GTG TCT GTG CGA CTG CAC CCA ACG AAC AAG GGT GTA GTC GCA TAG ACA TTG TGC TCA CTC GTA GCA TTG TGC CTG TCT CCA ACG GCA AAG GAG ATA G
B-T	TTC TAA TAC GAC TCA CTA TAG CAC AGT GCT ATG AGT GTC GCG ACG GAG ACG GTC GGG TCC AGA TAG GCC AGT CGA CAA GGT CTA TCT GTC GAG TAG AGT GTG GGC TCC GTC GCG TGC ACA TTA GAG TCG CTG TAT GCC ACA GTC GAC AAG TGG CGT ACA GCG GCT CTA GTG TGC TGC ACA GTG TCT GTG CGA CTG CAC CCA GTC GAC AAG GGT GTA GTC GCA TAG ACA TTG TGC TCA CTC GTA GCA TTG TGC CTG TCT CCA GTC GAC AAG GAG ATA G
C-T	TTC TAA TAC GAC TCA CTA TAG CAC AGT GCT ATG AGT GTC GCG ACG GAG ACG GTC GGG TCC AGA TAG GCC AGG TAC CAA GGT CTA TCT GTC GAG TAG AGT GTG GGC TCC GTC GCG TGC ACA TTA GAG TCG CTG TAT GCC ACA GGT ACC AAG TGG CGT ACA GCG GCT CTA GTG TGC TGC ACG GGA TCC CGA CTG GCC GCA GGT ACC AAG TGG CCA GGT AAC GAA TGG ATC CTG TGC TCA CTC GTA GCA TTG TGC CTG TCT CCA GGT ACC AAG GAG ATA G

Continues on next page

Table A.3 – continued from previous page.

Strand	Sequence
C̃-T	TTC TAA TAC GAC TCA CTA TAG CAC AGT GCT ATG AGT GTG CAC AGT GTC TGT GCG ACT GCA CGC AGG TAC CAA GCG TGT AGT CGC ATA GAC ATT GTG CTG CAC ATT AGA GTC GCT GTA TGA CCC AGG TAC CAA GGG TCG TAC AGC GGC TCT AGT GTG CTC GCG TGC CTC AGA GGA CCT GTC AGC AGG TAC CAA GCT GAT AGG TCC TTT GAG GTA CGC GTC ACT CGT AGC ATT GTG CCT GTC TCC AGG TAC CAA GGA GAT AG
L-T	TTC TAA TAC GAC TCA CTA TAG CAC AGT GCT ATG AGT GTG CAC AGT GTC TGT GCG ACT GCA CCC ATC GCG AAA GGG TGT AGT CGC ATA GAC ATT GTG CTG CAC ATT AGA GTC GCT GTA TGA GGG ATC GCG AAA CCC TCG TAC AGC GGC TCT AGT GTG CTC GCG TGC CTC AGA GGA CCT GTC ACC AGT CGA CAA GGT GAT AGG TCC TTT GAG GTA CGC GTC ACT CGT AGC ATT GTG CCT GTC TCC AGT CGA CAA GGA GAT AG

Table A.4 PCR primers used in Chapters 4 and 5, their sequences reported according to 5'-3' convention, as well as the DNA templates they amplify. The sequences of the corresponding DNA templates can be found in Table A.3.

Strand	Templates	Sequence
Primer f (forward)	$\overline{\text{A}}$-T, A-T, $\overline{\text{B}}$-T, B-T, C-T, C̃-T, L-T	TTCTAATACGACTCACTATAGCACAG
Primer rα (reverse)	$\overline{\text{A}}$-T, $\overline{\text{B}}$-T	CTATCTCCTTTGCCGTTGGA
Primer rβ (reverse)	A-T	CTATCTCCTTTCGCGATGGA
Primer rγ (reverse)	B-T, L-T	CTATCTCCTTGTCGACTGGAG
Primer rδ (reverse)	C-T, C̃-T	CTATCTCCTTGGTACCTGGAG

A.3 Chapter 6

Table A.5 Sequences of all RNA nanostars used in Chapter 6, reported according to 5'-3' convention.

Strand	**Sequence**
A_{OH1}	GGA UCC GGU UGC AAA GCA CAG UGC UAU GAG UGU GCA CGG GAU CCC GAC UGG CCG CAU CGC GAA AGU GGC CAG GUA ACG AAU GGA UCC UGU GCU GCA CAU UAG AGU CGC UGU AUG ACC CAU CGC GAA AGG GUC GUA CAG CGG CUC UAG UGU GCU CGC GUG CCU CAG AGG ACC UGU CAC CAU CGC GAA AGG UGA UAG GUC CUU UGA GGU ACG CGU CAC UCG UAG CAU UGU GCC UGU CUC CAU CGC GAA AGG AGA UAG
YFP_{Apt}	GGC AAC CGG AUC CAA AGC AGC UUC UGG ACU GCG AUG GGA GCA CGA AAC GUC GUG GCG CAA UUG GGU GGG GAA AGU CCU UAA AAG AGG GCC ACC ACA GAA GCU GCA
B_{OH2}	GCA CUA GAC CAC AAA GCA CAG UGC UAU GAG UGU CGC GAC GGA GAC GGU CGG GUC CAG AUA GGC CAG UCG ACA AGG UCU AUC UGU CGA GUA GAG UGU GGG CUC CGU CGC GUG CAC AUU AGA GUC GCU GUA UGC CAC AGU CGA CAA GUG GCG UAC AGC GGC UCU AGU GUG CUG CAC AGU GUC UGU GCG ACU GCA CCC AGU CGA CAA GGG UGU AGU CGC AUA GAC AUU GUG CUC ACU CGU AGC AUU GUG CCU GUC UCC AGU CGA CAA GGA GAU AG
STV_{Apt}	GUG GUC UAG UGC UUU AUG CGG CCG CCG ACC AGA AUC AUG CAA GUG CGU AAG AUA GUC GCG GGU CGG CGG CCG CAU

Table A.6 Sequences of all nontemplate strands of dsDNA templates coding for RNA nanostars used in Chapter 6, reported according to 5'-3' convention.

Strand	Sequence
A_{OH1}-T	TTC TAA TAC GAC TCA CTA TAG GAT CCG GTT GCA AAG CAC AGT GCT ATG AGT GTG CAC GGG ATC CCG ACT GGC CGC ATC GCG AAA GTG GCC AGG TAA CGA ATG GAT CCT GTG CTG CAC ATT AGA GTC GCT GTA TGA CCC ATC GCG AAA GGG TCG TAC AGC GGC TCT AGT GTG CTC GCG TGC CTC AGA GGA CCT GTC ACC ATC GCG AAA GGT GAT AGG TCC TTT GAG GTA CGC GTC ACT CGT AGC ATT GTG CCT GTC TCC ATC GCG AAA GGA GAT AG
YFP_{Apt}-T	TTC TAA TAC GAC TCA CTA TAG GCA ACC GGA TCC AAA GCA GCT TCT GGA CTG CGA TGG GAG CAC GAA ACG TCG TGG CGC AAT TGG GTG GGG AAA GTC CTT AAA AGA GGG CCA CCA CAG AAG CTG CA
B_{OH2}-T	TTC TAA TAC GAC TCA CTA TAG CAC TAG ACC ACA AAG CAC AGT GCT ATG AGT GTC GCG ACG GAG ACG GTC GGG TCC AGA TAG GCC AGT CGA CAA GGT CTA TCT GTC GAG TAG AGT GTG GGC TCC GTC GCG TGC ACA TTA GAG TCG CTG TAT GCC ACA GTC GAC AAG TGG CGT ACA GCG GCT CTA GTG TGC TGC ACA GTG TCT GTG CGA CTG CAC CCA GTC GAC AAG GGT GTA GTC GCA TAG ACA TTG TGC TCA CTC GTA GCA TTG TGC CTG TCT CCA GTC GAC AAG GAG ATA G
STV_{Apt}-T	TTC TAA TAC GAC TCA CTA TAG TGG TCT AGT GCT TTA TGC GGC CGC CGA CCA GAA TCA TGC AAG TGC GTA AGA TAG TCG CGG GTC GGC GGC CGC AT

Table A.7 PCR primers (and the DNA templates they amplify) and other DNA oligonucleotides used in Chapter 6, with sequences reported according to 5'-3' convention. Sequences of DNA templates corresponding to primers below can be found in Table A.6. Abbreviations: 3Bio – 3' biotin modification.

Strand	Templates	Sequence
Primer f2 (forward)	A_{OH1}-T	TTCTAATACGACTCACTATAGGATCCG
Primer r2 (reverse)	A_{OH1}-T	CTATCTCCTTTCGCGATGGA
Primer f3 (forward)	B_{OH2}-T	TTCTAATACGACTCACTATAGCACTAGA
Primer r3 (reverse)	B_{OH2}-T	CTATCTCCTTGTCGACTGGAGA
Primer f4 (forward)	YFP_{Apt}-T	TTCTAATACGACTCACTATAGGCA
Primer r4 (reverse)	YFP_{Apt}-T	TGCAGCTTCTGTGGTGG
$Biotin_{DNA}$	not applicable	GTGGTCTAGTGCTTT/3Bio/

Appendix B

Supporting Figures

B.1 Chapter 3

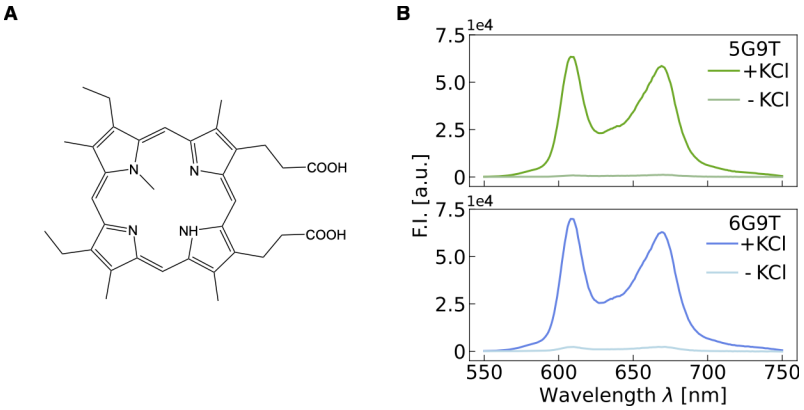


Fig. B.1 **Molecular structure and fluorescence emission spectra of NMM. A:** Molecular structure of N-methyl Mesoporphyrin IX (NMM), produced using ChemDraw. **B:** Fluorescence emission spectra of NMM bound to 5G9T (top) and 6G9T (bottom) after self-assembly transient in the presence (+KCl) or absence (-KCl) of 300 mM KCl injection (see Section 2.8.1). Samples which received a KCl injection display much stronger NMM fluorescence emission. NMM displays two emission peaks, one around 608 nm and the other around 670 nm. The latter was monitored for self-assembly kinetics assays. Spectra for samples which received KCl injection are average of triplicates, while those for samples lacking KCl injection and buffer blanks (used for background-subtraction) are average of duplicates.

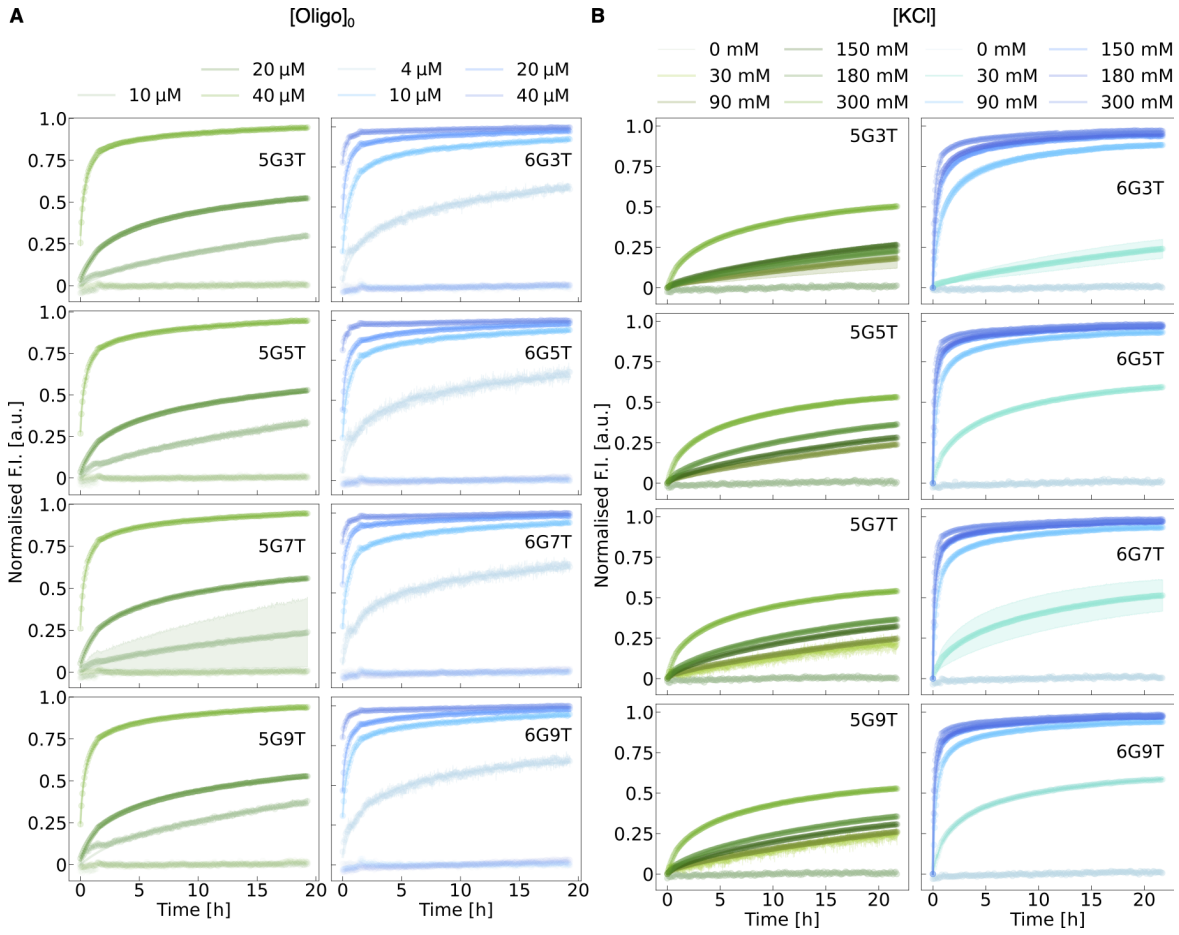


Fig. B.2 **Kinetics of G4-dependent self-assembly of non-cholesterolised Quad-Stars *via* NMM fluorimetry with varying oligonucleotide and KCl concentrations. A:** Kinetic profiles for all poly-T spacer lengths (increasing from top to bottom) and all examined $[Oligo]_0$ (starting oligonucleotide concentration), from which half-life values in Fig. 3.5C have been computed. **B:** Kinetic profiles for all poly-T spacer lengths (increasing from top to bottom) and all examined [KCl], from which half-life values in Fig. 3.5E have been computed. Curves with saturated colours (top) correspond to samples triggered with KCl injection, whereas fainter curves (bottom) are negative controls (no KCl). For samples, data are reported as mean (dots, one every 3 datapoints shown) $\pm$ standard deviation (shaded regions) and fits to theoretical tetramolecular kinetics (solid lines) [329]. For negative controls, only average, normalised datapoints are shown (see Sections 2.8.2 and 2.11.3). Data for negative controls, as well as 6G5T and 6G7T at 30 mM KCl have been averaged across independent duplicates. All other data have been averaged across at least three independent repeats (see Sections 2.8.2 and 2.11.3).

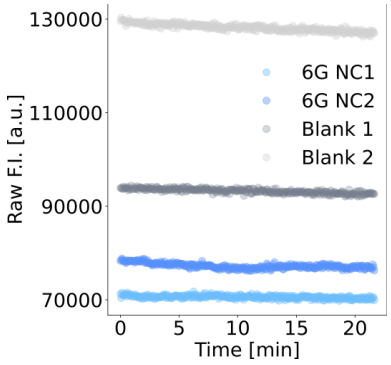


Fig. B.3 **Raw NMM fluorescence kinetics profiles for cholesterolised Quad-Stars negative controls and blanks.** Raw fluorescence intensity (F.I.) kinetics of 6G cholesterolised Quad-Stars negative controls ("NC"), *i.e.* samples annealed in 2 mM $MgCl_2$ and measured without any K^+ ion trigger for assembly, and blanks ("Blank"), *i.e.* plain TE buffer supplemented with NMM. As expected, samples without the cation trigger exhibit a flat-line behaviour with no fluorescence increase, just like blanks consisting of plain buffer and dye. These raw data directly supplement the same normalised profiles in Fig. 3.9C. For ease of visualisation, only one marker per three datapoints has been displayed as for all fluorimetry profiles.

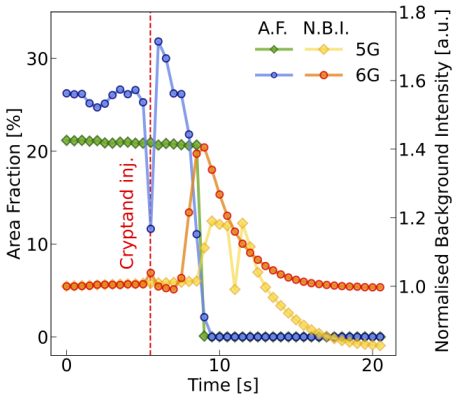


Fig. B.4 **Kinetics of rhodamine B cargo release upon cryptand-induced disassembly of Quad-Star hydrogel aggregates.** 5G and 6G Quad-Star aggregates disassemble upon cryptand injection (red dashed vertical line), as monitored *via* the corresponding area fraction profiles (A.F. = Area Fraction). This leads to the release of the rhodamine B cargo, as evidenced by the corresponding background intensity profiles (N.B.I. = Normalised Background Intensity). Curves for the same sample are displayed with the same markers (5G - circles, 6G - squares). Profiles are extracted from single FOVs of single samples. These profiles directly supplement Fig. 3.12C, proving that the specific timescales of disassembly and cargo release strongly depend on the proximity of the injection site to the imaged FOV. In the present case, the cryptand injection site was much closer to the imaged FOV than in Fig. 3.12C, thus inducing much faster (approximately 2-fold) and less temporally resolved disassembly and cargo release profiles.

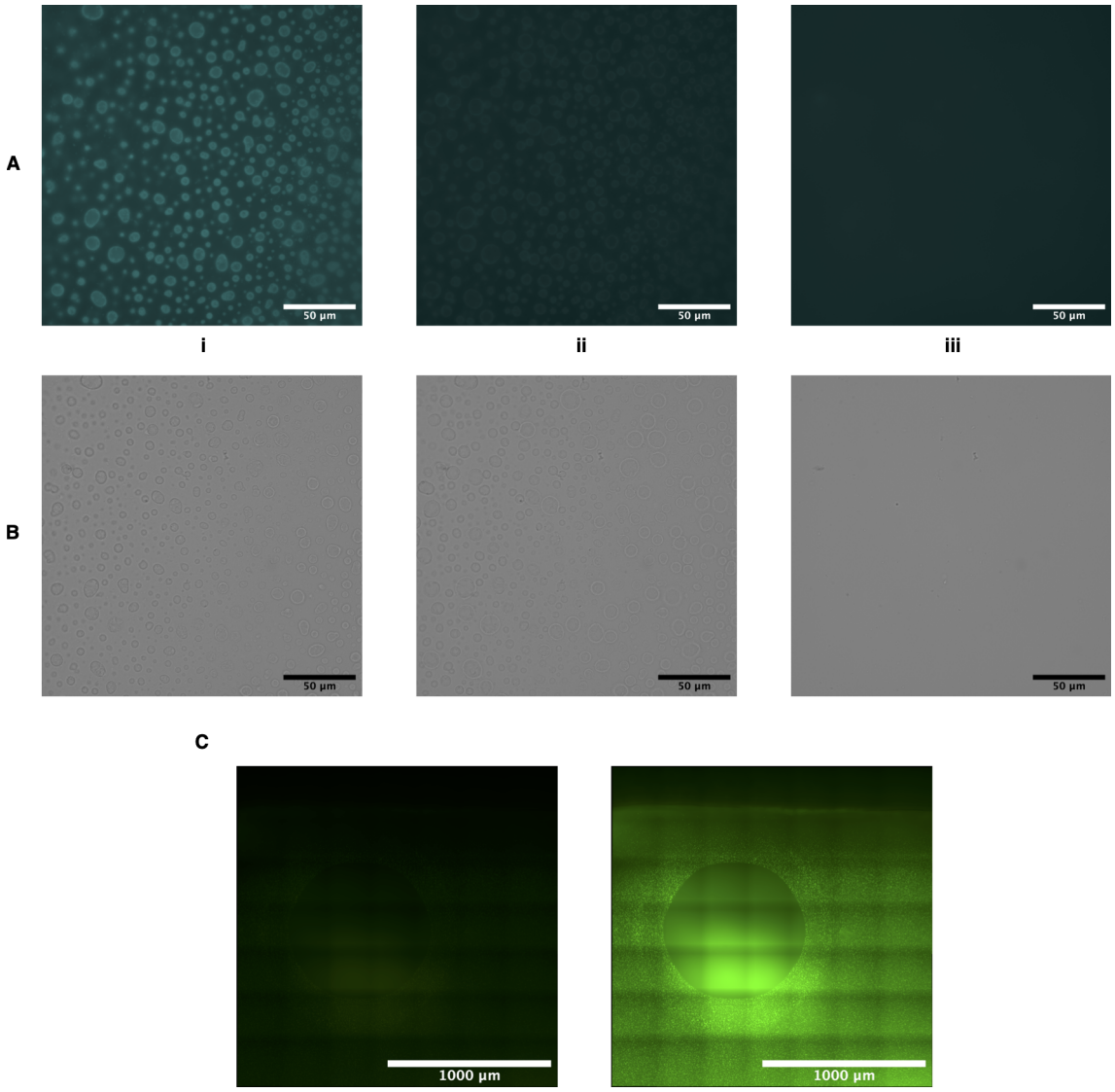


Fig. B.5 **UV-triggered and NMM-mediated photo-induced disassembly of Quad-Star aggregates. A, B:** Unprocessed and uncropped NMM epifluorescence and bright-field micrographs, respectively, corresponding to inset panels in Fig. 3.14A, depicting the irreversible near UV (395 nm)-induced disassembly of 6G Quad-Star aggregates in the presence of NMM. Micrographs in **A**, **B** are number-matched and follow the same uppercase roman numerals used in Fig. 3.14A. Uneven focus in **B** is most likely due to uneven positioning of the sample on the microscope stage. Scale bars are 50 μm. **C:** Unprocessed (left) and background-subtracted (right, *via* Rolling Ball algorithm in FIJI) micrograph depicting highly localised UV-induced disassembly of hydrogel aggregates in an Alexa488 fluorophore-labelled 6G Quad-Star sample in the presence of NMM. Only a circular portion was exposed to near UV, as it can be clearly observed from the circular hole carved by NMM-mediated photocleavage in the Alexa488 channel. This demonstrates that the disappearance of aggregates in the NMM channel of micrographs in **A** is due to actual disassembly and not to photobleaching of the porphyrin dye, as already supported by the same trend observed in bright-field (**B**). As for Fig. 3.14B, the underlying square-lattice pattern in **C** is an artefact stemming from the combination of a slightly uneven illumination and the automatic FOV stitching ("Large Image" in NIS Elements), and, as such, it does not represent a physical feature of the sample. Scale bars are 1000 μm.

B.2 Chapter 4

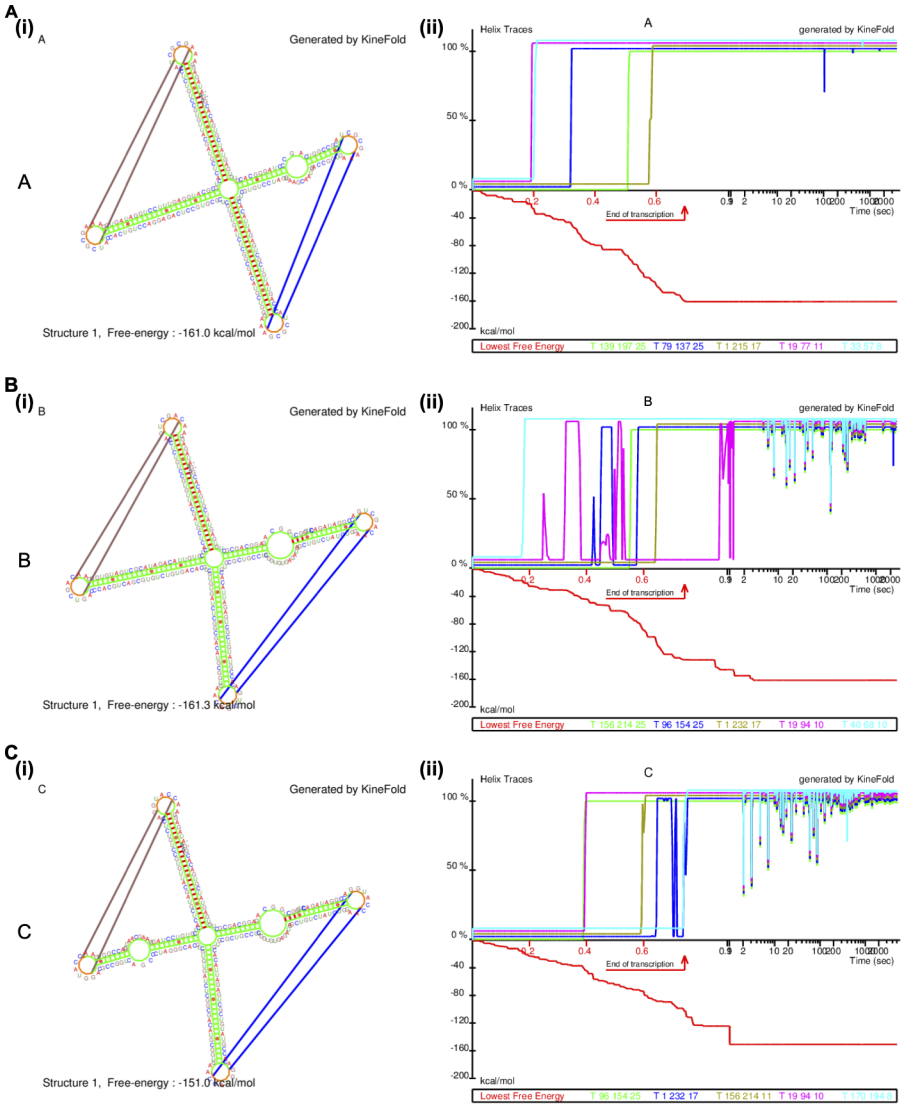


Fig. B.6 **Kinefold analysis of RNA nanostars A, B and C. A: (i)** Minimum free energy secondary structure of RNA nanostar A *post* co-transcriptional folding and **(ii)** corresponding "helix tracing graph", depicting the time-dependent evolution of the fraction of folded helices (0%-100%) for the five most stable helices in the nanostar (top) as well as the free energy over time (bottom, red). Time on the *x* axis is reported in linear/log scale during/after transcription, respectively. Colour-codes for the helices match the labels at the bottom, each reported as "T i j k", where i/j is the starting/final nucleotide, and k the number of base pairs in the helix. For instance, "T 79 137 25" is the fully double-stranded third arm (in order of transcription), comprising 25 base pairs between the 79^{th} and the 137^{th} transcribed nucleotides. **B:** Same as **A** for RNA nanostar B. The magenta profile (T 19 94 10) corresponds to the double-stranded portion of the second nanostar arm preceding the Broccoli aptamer, hence the major fluctuations. **C:** Same as **A** for RNA nanostar C. The blue profile (T 1 232 17) corresponds to the portion of the first nanostar arm close to the junction, *i.e.* the first transcribed stretch of the nanostructure. Its fluctuations prior to the end of transcription might be related to structural rearrangements upon formation of the remaining portion of the double-stranded arm. All panels were produced by representative Kinefold batch simulations (see Section 2.1).

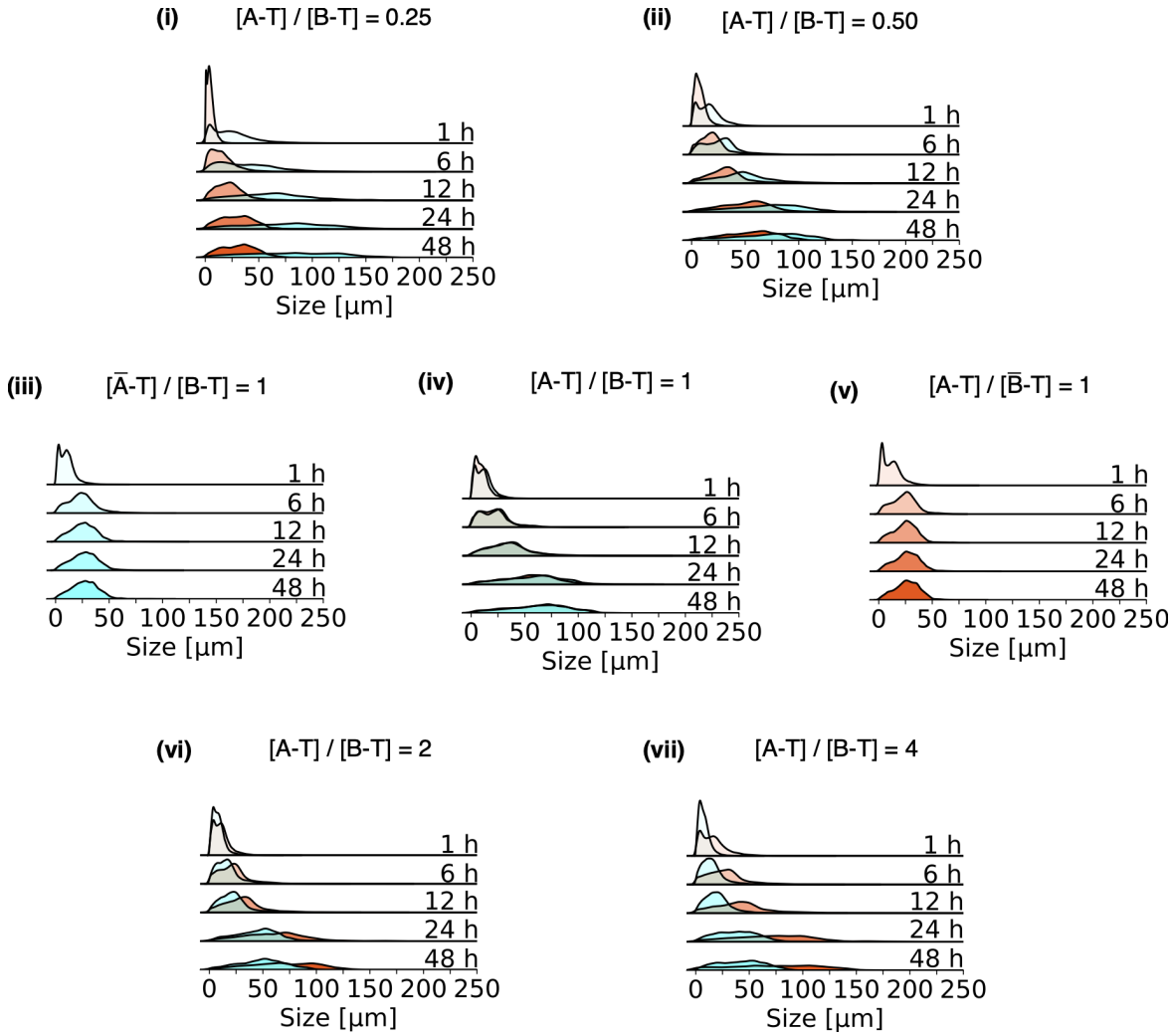


Fig. B.7 **Chord Length Distribution of orthogonal organelles formed by binary RNA nanostar systems in bulk.** Chord Length Distribution (CLD) ridge-plots at representative timepoints (1 h, 6 h, 12 h, 24 h and 48 h, vertical axis) for binary sticky nanostar systems with varying [A-T]/[B-T] DNA template concentration ratio, as well as stoichiometric binary systems with non sticky nanostars (middle row, left and right). The mean of these CLDs, μ_{CLD}, is presented in Fig. 4.7E to illustrate growth kinetics.

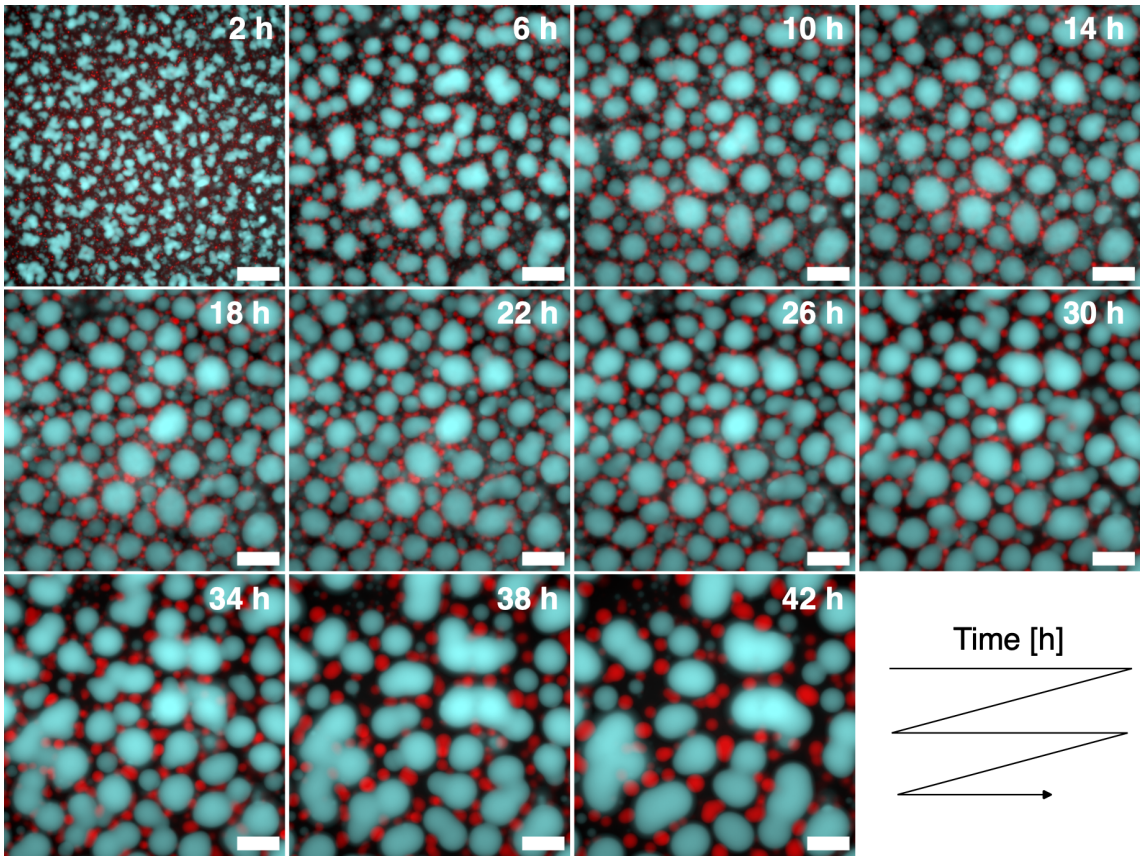


Fig. B.8 **Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 0.25 in bulk.** Uncropped, contrast-enhanced (see Section 2.11.5) epifluorescence micrographs depicting the bulk self-assembly of binary mixtures with [A-T]/[B-T] = 0.25 with time-step = 4 h. Time proceeds from left to right in each row, as schematically depicted in the bottom right corner. Scale bars are 100 μm.

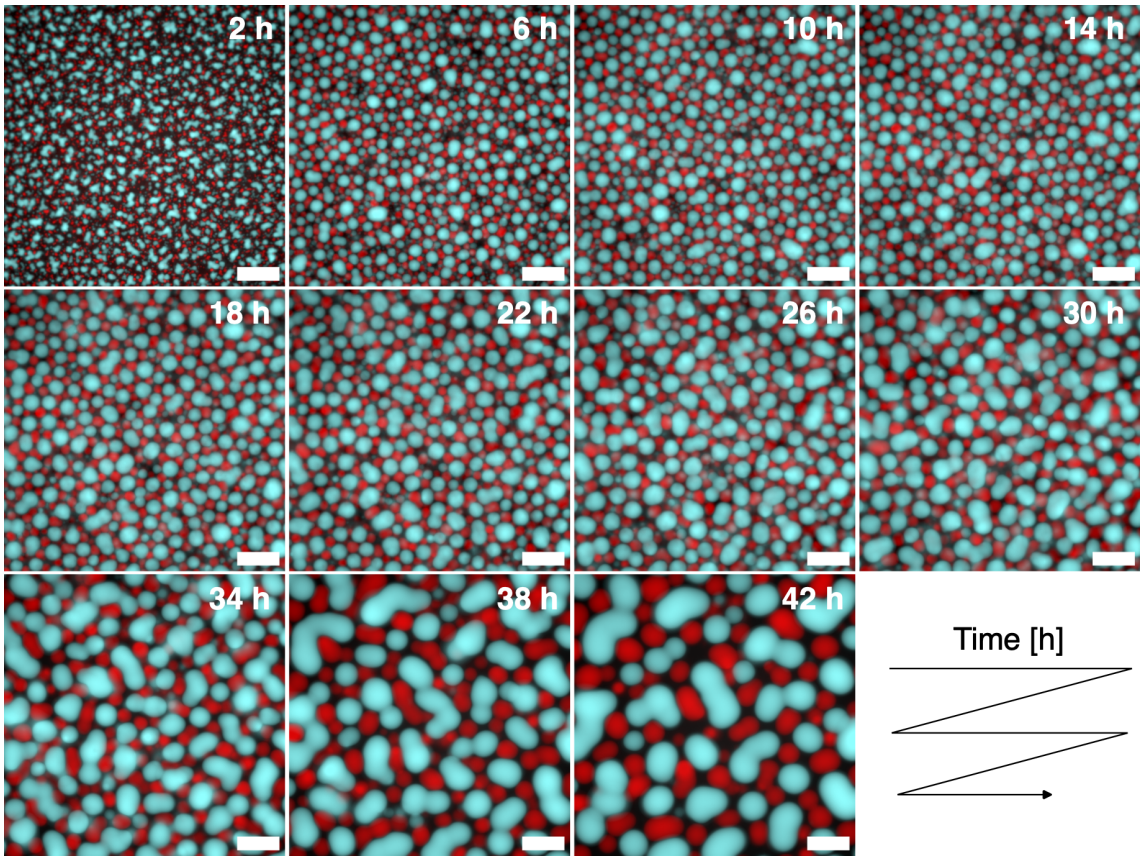


Fig. B.9 **Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 0.50 in bulk.** Uncropped, contrast-enhanced (see Section 2.11.5) epifluorescence micrographs depicting the bulk self-assembly of binary mixtures with [A-T]/[B-T] = 0.50 with time-step = 4 h. Time proceeds from left to right in each row, as schematically depicted in the bottom right corner. Scale bars are 100 μm.

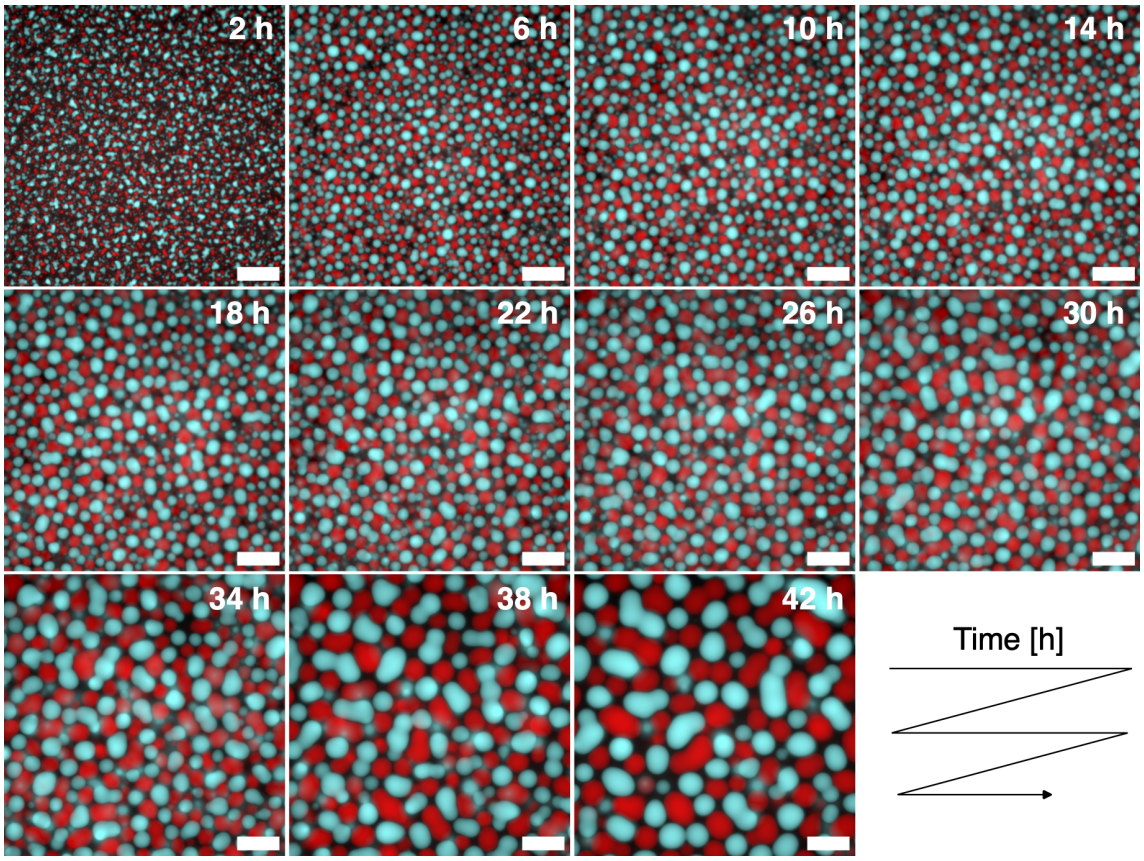


Fig. B.10 **Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 1 in bulk.** Uncropped, contrast-enhanced (see Section 2.11.5) epifluorescence micrographs depicting the bulk self-assembly of binary mixtures with [A-T]/[B-T] = 1 with time-step = 4 h. Time proceeds from left to right in each row, as schematically depicted in the bottom right corner. Scale bars are 100 μm.

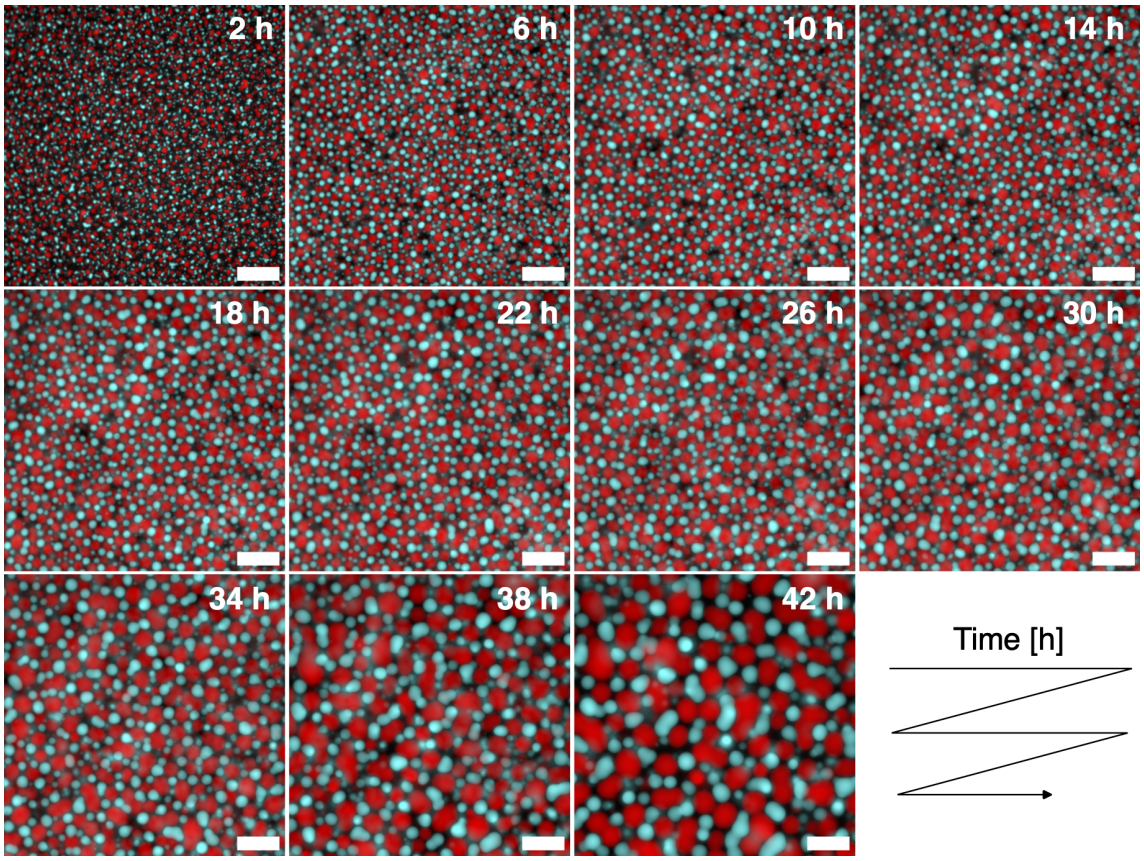


Fig. B.11 **Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 2 in bulk.** Uncropped, contrast-enhanced (see Section 2.11.5) epifluorescence micrographs depicting the bulk self-assembly of binary mixtures with [A-T]/[B-T] = 2 with time-step = 4 h. Time proceeds from left to right in each row, as schematically depicted in the bottom right corner. Scale bars are 100 μm.

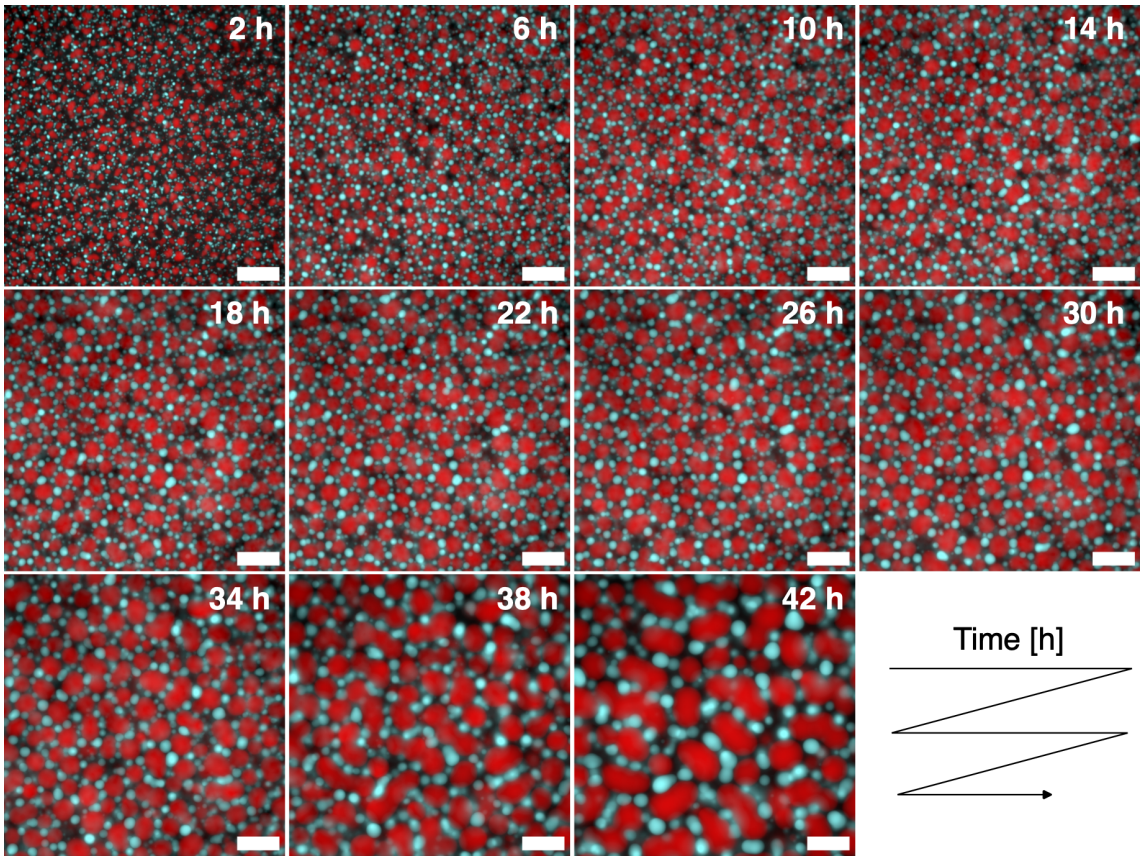


Fig. B.12 **Self-assembly of binary RNA nanostar systems with [A-T]/[B-T] = 4 in bulk.** Uncropped, contrast-enhanced (see Section 2.11.5) epifluorescence micrographs depicting the bulk self-assembly of binary mixtures with [A-T]/[B-T] = 4 with time-step = 4 h. Time proceeds from left to right in each row, as schematically depicted in the bottom right corner. Scale bars are 100 μm.

B.3 Chapter 6

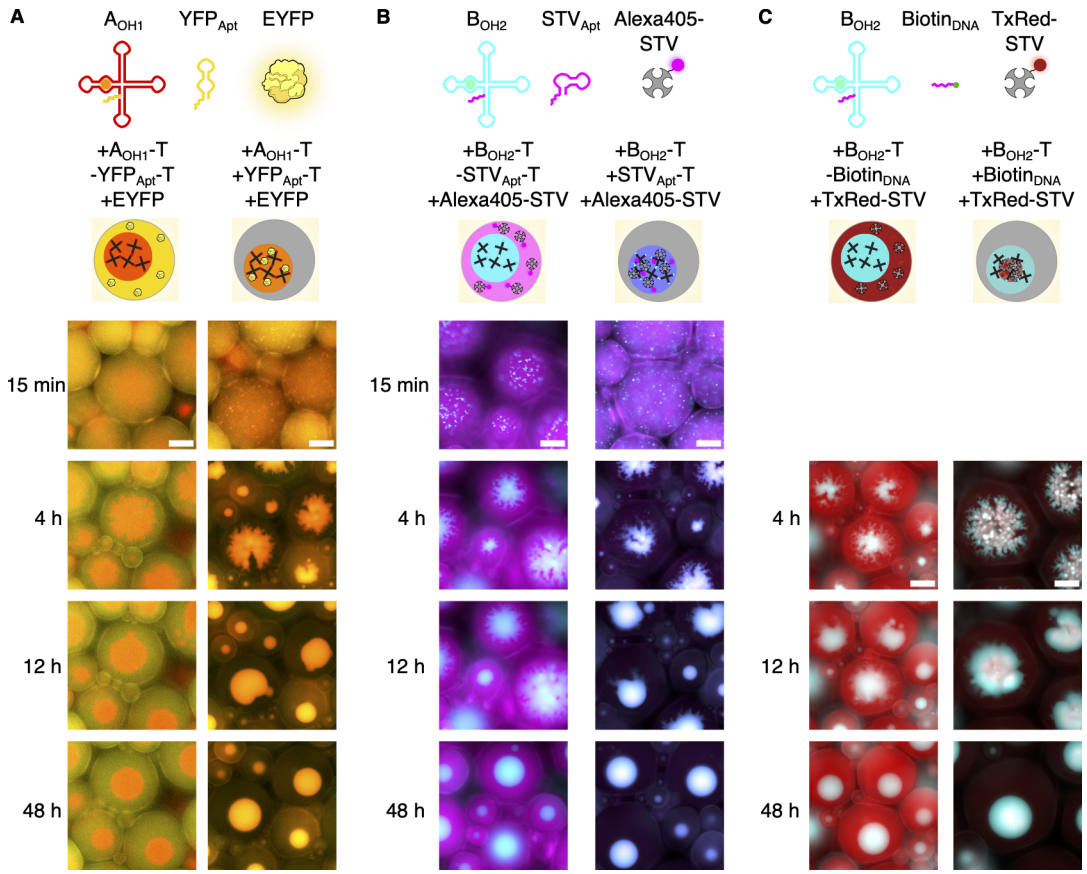


Fig. B.13 **Self-assembly of protein-capturing RNA organelles.** Epifluorescence micrographs depicting representative timepoints during the self-assembly of protein-capturing RNA organelles from single nanostar species within water-in-oil emulsion droplets. **A**: A_{OH1} (red) nanostars fail to capture EYFP (yellow) in the absence of YFP_{Apt} (-YFP_{Apt}-T, left), while succeed in doing so when simultaneously transcribed with it (+YFP_{Apt}-T, right). **B**: Same as **A** for B_{OH2} (cyan) nanostars capturing Alexa405-STV (magenta) *via* co-transcribed STV_{Apt} in the presence/absence of the corresponding DNA template (∓STV_{Apt}-T, left/right). **C**: Same as **B** for B_{OH2} (cyan) nanostars capturing TexasRed-STV (red) *via* short biotin-conjugated DNA oligonucleotide $Biotin_{DNA}$ ([$Biotin_{DNA}$] = 10 μM). In **C** the timepoint at 15 min is lacking due to software-related issues in the corresponding run resulting in usable data from 3.5 h onwards only. Micrographs have been cropped and contrast-enhanced for ease of visualisation (see Section 2.11.5), and all depict the same field-of-view area, thus sharing scale bars with the first micrograph in each series (column). Scale bars are 50 μm. In all systems, [nanostar DNA template]/[aptamer DNA template] = 3, [total DNA template] = 40 nM and [EYFP] = [STV] = 1.25 μM.

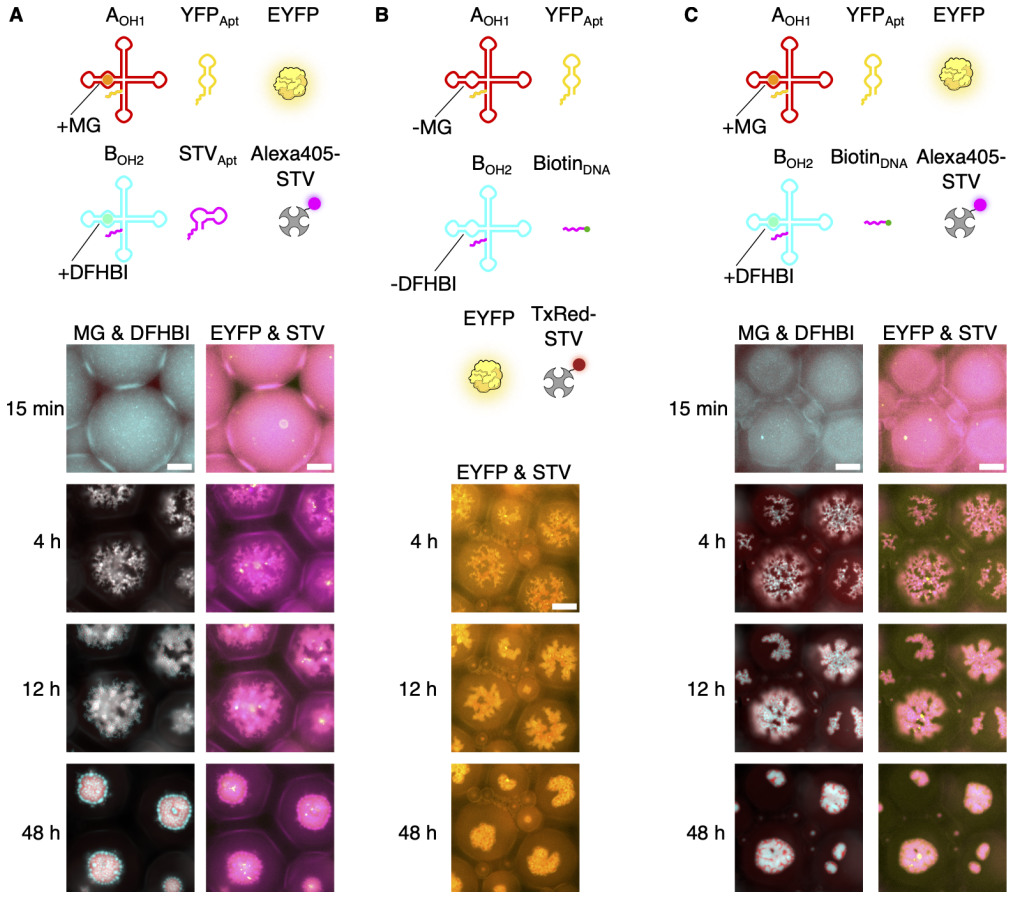


Fig. B.14 **Simultaneous transcription and co-assembly of distinct protein-capturing RNA organelles.** Epifluorescence micrographs depicting representative timepoints during the self-assembly of protein-capturing RNA organelles from simultaneously transcribed A_{OH1} and B_{OH2} nanostars within water-in-oil emulsion droplets. **A**: A_{OH1}/B_{OH2} (red/cyan) nanostars capture EYFP/Alexa405-STV (yellow/magenta) *via* YFP_{Apt}/STV_{Apt}. All constructs are simultaneously transcribed, with [nanostar DNA template]/[aptamer DNA template] = 3 for each pair. **B**: Same as **A** with B_{OH2} (cyan) nanostars capturing TexasRed-STV (red) *via* short biotin-conjugated DNA oligonucleotide $Biotin_{DNA}$ (added, not transcribed). Timepoint at 15 min is lacking due to software-related issues in the corresponding run resulting in usable data from 3.5 h onwards only. MG and DFHBI dyes were omitted from this run due to spectral overlap of MG+MGA with TexasRed. **C**: Same as **B** with TexasRed-STV replaced by Alexa405-STV to enable usage of nanostar-binding dyes to investigate colocalisation of target proteins and nanostar phases. In **B-C**, [$Biotin_{DNA}$] = 10 μM. Micrographs have been cropped and contrast-enhanced for ease of visualisation (see Section 2.11.5), and all depict the same field-of-view area, thus sharing scale bars with the first micrograph in each series (column). Scale bars are 50 μm. In all systems, [total DNA template] = 40 nM and [EYFP] = [STV] = 1.25 μM.