Reprogramming and manipulating DNA assembly using simple self-assembly principles

Nicole Avakyan

Department of Chemistry, McGill University

Montréal QC Canada

September 2018

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

To my parents, Inna and Levon,

Сидела-бы я здесь без вашей любви, терпения и ежедневной помощи?

... pour entrer dans le secret des choses, il faut d'abord se donner à elles.

- Simone de Beauvoir, Mémoires d'une jeune fille rangée (1958)

Abstract

Supramolecular chemistry is the "chemistry of molecular information". Information stored in the structural features of molecules is recognized and processed at the supramolecular level and drives the ever more complex organization of matter – from the formation of simple host-guest complexes, to the assembly of viral capsids and the emergence of life itself. The work presented in this thesis uses basic principles governing intermolecular interactions to manipulate the assembly of DNA, a molecule that exemplifies the concept of information storage in its structure. In Chapter 2, simple molecular recognition principles are applied to expand the DNA base-pairing alphabet. A small molecule, cyanuric acid, with three thymine-like faces, is determined to reprogram the assembly of poly(adenine) strands into micron-long fibers with a unique internal structure. Adenine residues and cyanuric acid molecules are shown to associate into hexameric rosettes that bring about the formation of poly(adenine) triplexes and subsequent polymerization into nanofibers. Fundamentally, this finding demonstrates that a small molecule can induce nucleic acid assembly, giving rise to a new type of structure from inexpensive, readily available materials. Further examination of the mechanisms involved in nanofiber formation in Chapter 3 uncovers that the supramolecular polymerization of poly(adenine) strands proceeds through a cooperative mechanism. A new methodology to analyse thermal hysteresis profiles provides a quantitative understanding of the kinetics and thermodynamics of the assembly process. In addition, key parameters influencing assembly are identified and can be tuned to modify nanofiber properties in potential nanomaterials applications. While hydrogen bonding and π-stacking interactions dominate the co-assembly of poly(adenine) and cyanuric acid described in Chapters 2 and 3, a different combination of intermolecular interactions is used to manipulate the long-range assembly of DNA nanostructures into ordered networks in Chapter 4. Blunt-ended DNA tiles, or tiles lacking complementary single-stranded portions at the duplex ends, with and without a cholesterol anchoring moiety, are interfaced with lipid bilayers, a soft support with composition-dependent properties. The balance of π-stacking interactions between DNA duplex ends, cholesterol-mediated nanostructure anchoring, electrostatic DNA binding to the bilayer surface, and the fluid nature of the lipid bilayer enable the formation of dynamic materials with long-range order and

tunable morphology on a cell membrane-like support. In turn, these lattices represent a tool for organizing materials such as nanoparticles or proteins at a biological interface, with potential applications in cellular recognition, development of model systems for study of membrane proteins or plasmonics. Altogether, this thesis demonstrates that information encoded in molecular building blocks and the interplay between intermolecular interactions can be used to manipulate the assembly and long-range organization of DNA to achieve new structures and higher complexity.

Résumé

La chimie supramoléculaire est aussi appelée la « chimie de l'information moléculaire ». L'information structurelle des molécules est reconnue et traitée au niveau supramoléculaire, entraînant ainsi la complexification de l'organisation de la matière – de la formation de simples composés d'inclusion à l'émergence de la vie elle-même. Le travail présenté dans cette thèse applique les principes fondamentaux régissant les interactions intermoléculaires pour manipuler l'assemblage de l'ADN, un archétype du stockage de l'information dans la structure moléculaire. Dans le Chapitre 2, des principes élémentaires de reconnaissance moléculaire sont appliqués pour élargir l'alphabet d'appariement de bases d'ADN. L'acide cyanurique, une petite molécule à trois faces de type thymine, est utilisé pour reprogrammer l'assemblage de brins de poly(adénine) en fibres d'une longueur de l'ordre du micromètre avec une structure interne unique. Nos résultats indiquent que l'association des unités d'adénine et d'acide cyanurique forme des rosettes hexamériques, entraînant ainsi la formation de triple-hélices de poly(adénine) et leur polymérisation subséquente en nanofibres. Fondamentalement, cette découverte démontre qu'une petite molécule peut induire l'assemblage de brins d'ADN, formant ainsi une nouvelle structure à partir de matériaux peu coûteux et abondants. Dans le Chapitre 3, des études visant les mécanismes impliqués dans la formation des nanofibres révèlent que la polymérisation supramoléculaire des brins de poly(adénine) avec l'acide cyanurique passe par un mécanisme coopératif. Une nouvelle méthodologie d'analyse de profils d'hystérèse thermique fournit une compréhension quantitative plus approfondie de la cinétique et de la thermodynamique du processus d'assemblage. Certains paramètres clés qui influencent l'assemblage sont également identifiés et peuvent être modifiés afin d'ajuster les propriétés des nanofibres et de développer des applications potentielles en tant que nanomatériaux. Alors que les liaisons hydrogène et les interactions entre cycles aromatiques dominent le co-assemblage de poly(adénine) et d'acide cyanurique décrit dans les Chapitres 2 et 3, une combinaison différente d'interactions intermoléculaires est utilisée pour manipuler l'assemblage de tuiles d'ADN en réseaux ordonnés à grande échelle dans le 4^{e} Chapitre. Ainsi, des tuiles d'ADN à extrémités franches, c'est-à-dire sans ajout de portions d'ADN complémentaire aux extrémités, sont modifiées avec une molécule de cholestérol servant d'ancre hydrophobe.

L’étude de l’interface entre ces tuiles, modifiées ou non, et des bicouches lipidiques, supports souples avec des propriétés dépendantes de la composition, est entreprise. L'équilibre des interactions entre cycles aromatiques joignant les extrémités des duplex, l'ancrage de la nanostructure médiée par le cholestérol et les liaisons électrostatiques de l'ADN à la surface de la bicouche permet de construire des matériaux dynamiques ordonnés à grande échelle, permettant aussi d’obtenir des morphologies ajustables. À leur tour, ces réseaux représentent un outil pour organiser des matériaux tels que des nanoparticules ou des protéines à une interface biologique, avec des applications potentielles dans la reconnaissance cellulaire, le développement de systèmes modèles pour l’étude de protéines membranaires ou les systèmes plasmoniques. En somme, cette thèse démontre que l’information structurelle codée dans des blocs de construction moléculaires et un juste équilibre d’interactions intermoléculaires peuvent être utilisés pour manipuler l'assemblage de l'ADN et son organisation à plus grande échelle.

Acknowledgements

I am grateful to the many people, in and out of the lab, who have helped me navigate the crests and troughs of academic life during "the PhD years".

I would like to thank my research supervisor Hanadi Sleiman, whose boundless enthusiasm and creativity never cease to amaze me. Hanadi's unique gift in communicating her love for research first brought me to her lab for a summer Honours project as an undergraduate student. Over the years, we have shared great discoveries and the frustrations that sometimes go into making them. I am most grateful for her support and patience during hard times, her insistence on staying positive, and her influence as a creative force who always prioritizes integrity.

Many thanks also to Professors Karine Auclair, Janine Mauzeroll and Nicolas Moitessier for insightful questions and valuable input during yearly review meetings.

I had the opportunity to collaborate with Andrea Greschner, Christopher Serpell, Violeta Toader, Anne Petitjean, Maciej Barlog and Hassan Bazzi on the biggest/longest project of my PhD. Thank you for sharing your unique talents and enthusiasm to solve this supramolecular mystery. Justin Conway was my partner in crimes committed against lipid bilayers and AFM imaging. The struggle was real man, but we made it through the sticky ends and unsaturated chains to the other side. Robert Harkness and Anthony Mittermaier introduced me to the world of log-log plots, curve fittings and mechanistic models. Thank you for having the patience to explain the same things to me over and over again. A special message of gratitude goes out across the Atlantic to Carlos Gonzalez, Irene Gomez-Pinto, Miguel Garavis, Israel Serrano and their colleagues at the Instituto de Química Física Rocasalano (CSIC) in Madrid for their generous welcome during my research stay in the Gonzalez lab. Because of your help, I can now count SPARKY among my friends.

Around the department and other facilities, research life was made easier with the particular expertise of the following people: Mohini Ramkaran, Beatrice Lego, Julia Del Re and Patricia Moraille, the AFM masters at McGill and UdeM; CSACS impresario Petr Fiurasek; oligo-LCMS wizard Alexander Wahba; and the incomparable Chantal Marotte. Thank you for dedicating your time and patience to help me through all sorts of science and administration conundrums.

I had the privilege to receive funding from the Chemistry Department, NSERC, FRQNT and CSACS over the years. Many thanks to these organizations and agencies for their support.

To my friends and colleagues of the Sleiman lab – you have been part of my life since the summer of 2010 – thank you for all the moments. Being a member of this group is not only about doing cutting-edge research in the company of good people (sometimes A LOT of people). The Sleiman lab is also Samosa Wednesday, Thirsty Thursday, SleimanSecretSinterSanta, music parties, weddings and childbirth – a place where life happens amidst fumehoods, benches and (sometimes broken) instruments. I've had the chance to work with you, share memories, share drinks, cry my eyes out, and laugh until my abs and lungs could no longer take it. I am grateful to have met people from all over that I still want to hang out with years from now.

Here are some of the people who have marked my stay in the Sleiman lab: Katie – the first lab member I met, my lab mom/my cat's mom's mom; Graham – kindly though smirkingly found intelligent answers to my dumbest questions; Andrea – all around fantastic human, Blender-master, MVPR (most valuable proof reader); Fiorita and Iltinho – the italo-brazilian party duo; Karina – aka sunshine personified; Mijalis and Anita – best most talented music party buddies, Dowine – a real Dutch bikes in any weather; Violeta – culture aficionado, thank you for enthusiastically attending my many choir concerts; Kimberlite – voice of reason, power-pointer extraordinaire; Amani – supplier of the best baklava; Johano – deep believer in keeping lunchtime conversation controversial and microagressive; Tom and JCON – religiously devoted to the brewing arts and to making sure that the party goes on late into the night; DB and HF – keep it suavemente bros; Alex "I got a sunburn" Prinzen; the inscrutable cool cat, Mikey the shark; vicomte Dodo de Rochambeau – can I visit your château sometime?; Chidchob the Dragon – sleepier than the name suggests, #1 nicest person in the lab; Tuanito – mischievous lab paparazzi; Xin – birthday twin, bench successor, I count on you to carry on my legacy of…. excellence? thanks; John – once upon a time AFM apprentice; Chris – once reluctant project taker-over; Casey – Calgarian baking/hockey star; Kamila – the Russian hacker; Oreo, Emparita and Kai – bossladies always ready to dance up a storm, sexual harassment perpetrators, oasis of femininity (lol) in the middle of the sausage party.

Last, but certainly not least – Katherine BEWJOLDE. I would have never made it through without you and am so grateful that you were by my side for 9 years (literally side by side, 8 hours a day, but who am I kidding, this is not a 9 to 5 job!). Thank you for being a friend even when I'm

annoying (Madame Katherine!). Thank you for your constant support, much mutual complaining, but most of all, the ridiculous fun times we shared over the years. I was devastated to miss your defense (AND YOUR WEDDING?! worst. friend. ever.), but still hope and pray that, one day in the future, you will forgive me, and we'll sit on your Magog porch sipping tea and eating strawberries in our grandma sweaters. Raf can come too, although he might prefer biking.

To the other Holmes Room crew members – Josh, Janet, Matt and David – the McGill journey would not have been the same without you. Chem semis, drunk lunches, study sessions that degenerated into $%*&@ knows what, yearly pool parties, the most ludicrous conversations – yes please can I have some more. We started in this together and although it is now much harder to gather everybody in the same place at the same time, it's a blast and a whirlwind of good memories whenever it happens. As expected, I'm the last person out of the exam room.

During the PhD years, choral singing became my escape activity of choice. I am grateful to have met such fantastic people in every choir – your friendship helped me get out of the science bubble and to restore my sanity by singing the loudest, highest notes (but you know, keeping it pretty). Special mention goes to Hanna and Elisa, the most harmonious metro/Otto Maass straircase duet/trio partners a girl could ask for. In the latter years, the St Lawrence Choir lady-crew made sure that every Wednesday night was a blast.

To Mengting Katherine Zhangmaster General DMD – thank you for the long-distance calls lasting 3 hours at a time. Great opportunities for both of us to vent frustrations and share our accomplishments across the kilometers separating us. And Amelia – we became friends before I spoke either French or English. Who knew that 18 years later we would both be about to become doctors? (the other kind of doctor). Thank you for the badminton, the baking and the moral support.

I found inspiration to pursue studies in Chemistry in Org. Chem. classes taught by Dr. Aniko Lysy at Marianopolis College. Dr. Lysy is the toughest, funniest and most passionate teacher I've ever had. I will never forget her kindness, her caring support and the phrase "My boy, if you protonate this nitrogen, I am going to jump into the wall!" she addressed to a fellow student in her indelible Hungarian accent.

My parents Inna and Levon have supported me in every-thing, every-day of my 29-year-long life so far. I could never express enough gratitude for your love and endless patience. Thank you for all of it, and for grossly exaggerating my qualities when you talk to people.

Table of contents

List of figures

List of tables and schemes

List of abbreviations

1D	One dimensional
2D	Two dimensional
3D	Three dimensional
A	Adenine
Å	Angstrom
Ac_2O	Acetic anhydride
AcMg	Acetate-magnesium
AFM	Atomic force microscopy
bp	Base pairs
C	Cytosine
CA	Cyanuric acid
$CaCl_2$	Calcium chloride
CD	Circular dichroism
CPG	Controlled pore glass
D	Diffusion coefficient
$d(A_n)$	Poly(deoxyadenosine)
DCA	Dicholoracetic acid
DNA	Deoxyribonucleic acid
DCM	Dichloromethane
DLS	Dynamic light scattering
DMF	Dimethylformamide
DMT	Dimethoxytrityl
DOPC	1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DOSY	Diffusion ordered spectroscopy
DPPC	1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine
ds	Double-stranded
DX	Double cross-over
EDTA	Ethylenediaminetetraacetic acid
E_T	FRET efficiency
ETT	5-(Ethylthio)-1H-tetrazole
FRET	Förster resonance energy transfer
FWHM	Full width at half maximum
G	Guanine
G4	Guanine quadruplex
H_2O	Water
HCl	Hydrochloric acid
HEG	Hexa(ethylene)glycol
HPLC	High performance (or pressure) liquid chromatography
I_2	Iodine
LCAA	Long chain alkylamine
LC-ESI-MS	Liquid chromatography-electrospray ionization mass spectrometry
L_β	Gel

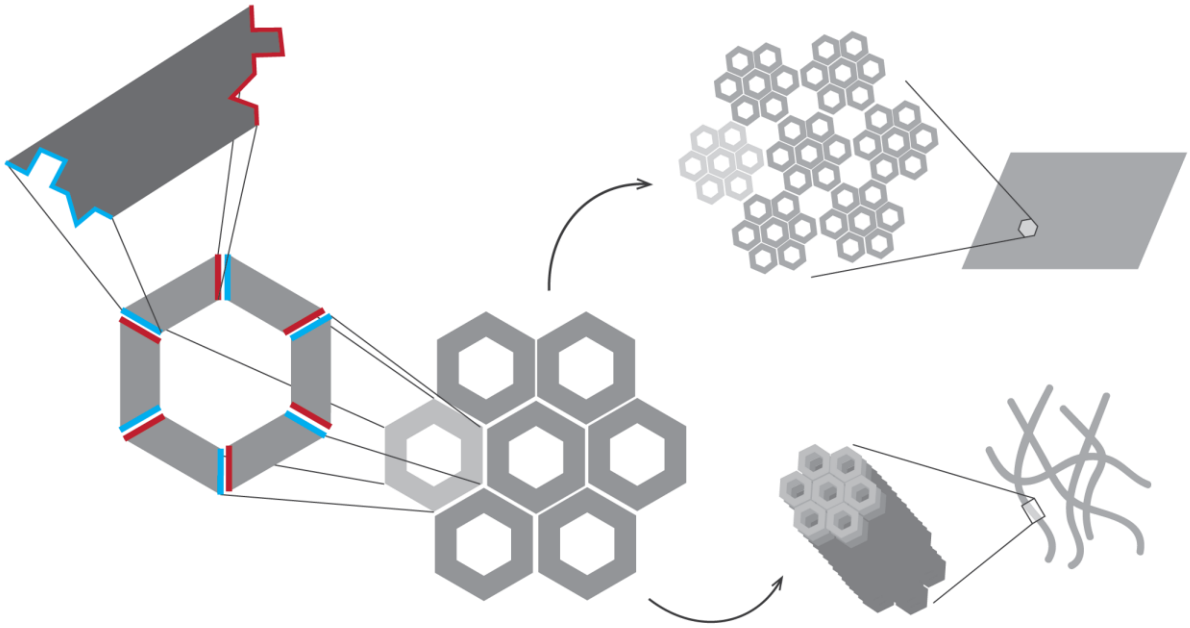
L_d	Liquid disordered
L_o	Liquid ordered
MeCN	Acetonitrile
MeIm	1-Methylimidazole
$Mg(OAc)_2 \cdot 4\ H_2O$	Magnesium acetate tetrahydrate
$MgCl_2 \cdot 6\ H_2O$	Magnesium chloride hexahydrate
$[Mg^{2+}]$	Magnesium ion concentration
µM	Micromolar
mM	Millimolar
min	Minute
M_N	Number average molecular weight
M_W	Weight average molecular weight
NaCl	Sodium chloride
$[Na^+]$	Sodium ion concentration
nm	Nanometer
nM	Nanomolar
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
PAGE	Polyacrylamide gel electrophoresis
$p(A_n)$	Polyadenine PNA
PDI	Polydispersity index
PNA	Peptide nucleic acid
ppm	Parts per million
$r(A_n)$	Poly(adenosine)
R_h	Hydrodynamic radius
RNA	Ribonucleic acid
rt	Room temperature
SLB	Supported lipid bilayer
STV	Streptavidin
TEM	Transmission electron microscopy
TEMED	Tetramethylethylenediamine
T	Thymine
TAMg	Tris-acetate-magnesium
TAP	2,4,6-Triaminopyrimidine
TBE	Tris-boric acid-EDTA
TEAA	Triethylammonium acetate
TEG	Triethylene glycol
TEMED	Tetramethylethylenediamine
THF	Tetrahydrofuran
Tris	Tris(hydroxymethyl)aminomethane
ss	Single-stranded
T_M	Melting temperature
TOCSY	Total correlation spectroscopy
UV-Vis	Ultraviolet-visible
VPO	Vapor pressure osmometry

Author contributions

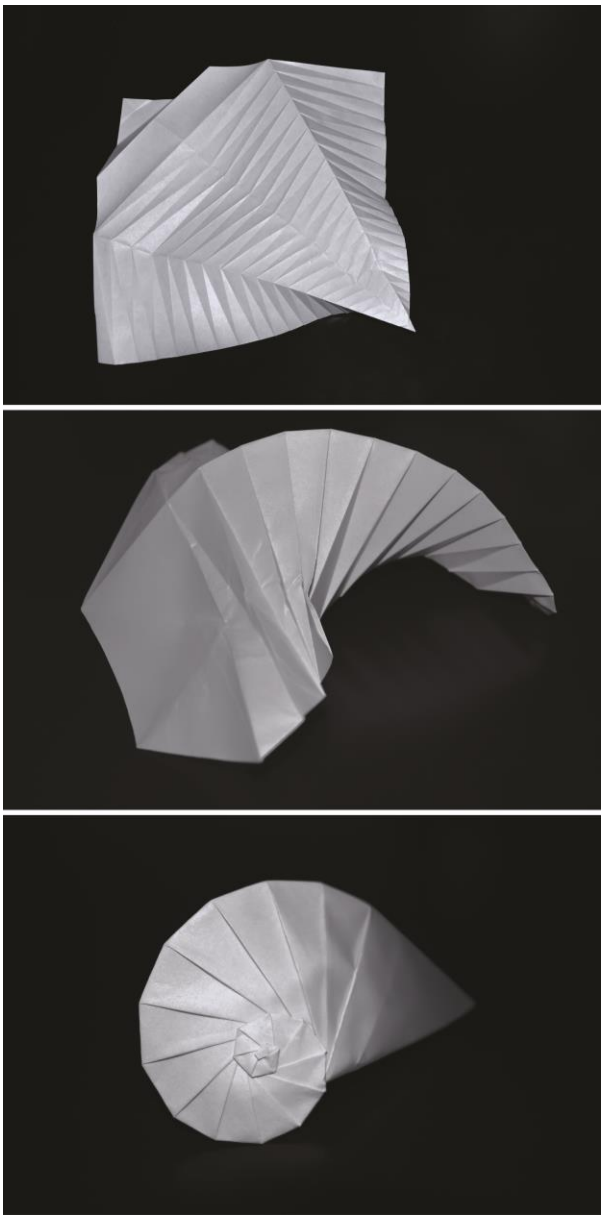
Prof. Hanadi F. Sleiman provided funding, research objectives, experimental design supervision and intellectual guidance for all the projects described in this thesis.

The specific contributions of co-authors are described at the start of each chapter.

1| Introduction



Portions of sections 1.4 and 1.5 have been adapted and expanded from "Long-range assembly of DNA into nanofibers and highly ordered networks" by KMM Carneiro, N Avakyan and HF Sleiman; Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol., ***5****, 266-285 (2013).*



Ab initio / Evolutio / Sicut est
From the beginning / Evolution / As it is

Model: Navel shell

Tomoko Fuse. SPIRAL Origami Art Design. Viereck Verlag (2012).

Folding, photography and image editing: Nicole Avakyan

1.1 Supramolecular chemistry and self-assembly: The chemistry of molecular information

Whether on the atomic or on the galactic scale, the emergence of complexity in the matter that surrounds us is driven by self-organization.[1] While at the level of star systems self-organization is operated by gravitational forces, in the realm of supramolecular chemistry, there is a collection of intermolecular interactions that determine how molecules acquire secondary structure, affect each other and assemble into larger complexes.

The father of supramolecular chemistry, Jean-Marie Lehn, has called this field 'the chemistry of molecular information'.[2] This information is stored in the structural features of every molecule. It is then read at the supramolecular level following principles of molecular recognition, taking into account electronic and steric effects, conformational flexibility, solvation (including the hydrophobic effect), etc.[2] The interplay of adhesive and repulsive forces[3] (such as hydrogen bonding, π-π stacking and van der Waals forces) defines interactions between molecules and drives the ever more intricate, hierarchical organization of matter. The spontaneous and reversible process by which molecules associate into larger and more complex structures according to the information stored in the molecules themselves is called self-assembly.[4] From the formation of a simple inclusion complex, to the folding of a protein serving a specific catalytic function, the assembly of a highly structurally complex biological entity such as a virus, and the emergence of life itself – the information stored in molecular structure drives ordering and complexity.

1.2 Self-assembly across length scales

The development of nanotechnology and nanomaterials, technologies and materials that operate and have specific size features on the nanometer scale, can be achieved by two main approaches – "top-down engineering" and "bottom-up synthesis".[5] On one hand, the top-down approach starts out with bulk materials and uses advanced techniques to engineer and manufacture miniaturized systems. On the other hand, the bottom-up method uses molecular synthesis and supramolecular interactions to produce functional nanoscale structures. Importantly, self-assembly can access the size range relevant to nanomaterials, from the molecular (measured in angstroms) to the top-down fabrication worlds (~ 100 nm). The added advantage of self-assembly is that

hierarchical organization can be integrated into the building blocks, providing access to ordered materials and systems at a larger scale, which is beautifully illustrated in the long-range organization and hierarchical complexity encountered in natural systems.

1.3 Natural and man-made long range self-assembly systems

The rational design of self-assembly systems is largely inspired by nature and assemblies found in biology. On the scale of isolated binding events, supramolecular interactions direct substrate binding to a receptor. At the level of small, discrete particles, viral genomes are sequestered for their protection inside protein shells that assemble according to exquisite symmetry through precise interactions between different subunits. Recently, artificial counterparts, protein cages composed of non-viral building blocks, have been rationally evolved[6] and were even able to encapsulate a viral genome.[7] In addition, countless drug delivery systems inspired by the cargo protection afforded by viral capsids have been developed in the last few decades to encapsulate small molecule drugs or nucleic acids for potential gene therapy applications.[8]

On a larger scale, supramolecular interactions drive the formation of complex materials with hierarchical organization that have a variety of fundamental functions. Biological membranes self-assembled from amphiphilic lipid molecules under the influence of the hydrophobic effect[9] have essential roles in creating compartmentalization in living organisms. They sequester different cellular components into separate compartments, thus giving rise to spatial organization and creating organelles with specialized functions. Furthermore, lipid membranes represent a fluid support to organize enzymes, receptors mediating communication with the outside environment, and tether scaffolding components. The scaffolding inside the cell, composed of dynamic yet robust supramolecular filaments, maintains the cell's shape, acts as a track for the transport of materials inside the cell, and can be reconfigured to direct cell motility.[10] There is a wide variety of supramolecular fibers in the extracellular environment as well, with structural and communication-related roles.[11,12]

It is no wonder that biological materials are a source of inspiration for long-range organization in technological applications. Compartmentalization afforded by supramolecular membranes and shells can be used to create a nanoreactor or to protect sensitive cargo during targeted drug delivery. Scaffolding materials can organize materials precisely in space, while

addressable surfaces have the added advantage of accelerating the kinetics of reactions by confining active components to 2D rather than 3D space. Altogether, the levels of controlled assembly achieved with hierarchical organization obtainable by self-assembly give rise to intricate structure that consequently enables greater functional complexity. The elegance and complexity of natural systems thus serve as a template for researchers to elaborate new functional materials.

1.3.1 Amphiphilic self-assembly and lipid bilayers

Amphiphilic molecules are among the most powerful building blocks for self-assembly. They are ubiquitous in biology (lipids), daily life (detergents) and advanced technological applications.[13] Amphiphiles contain both hydrophilic and hydrophobic portions which have opposing minimum energy requirements for solvation, defining their interactions with the surrounding medium. In water, amphiphilic self-assembly is driven by the hydrophobic effect, resulting in aggregation of hydrophobic portions away from the aqueous environment, and leaving hydrophilic groups exposed.[14]

Phospholipids are amphiphilic molecules which constitute the main components of biological membranes. Their structure is characterized by (i) a hydrophilic polar head group, which can be charged or zwitterionic, (ii) a linker, typically a small molecule such as glycerol, and (iii) hydrophobic tail groups, alkyl chains of variable length and degrees of unsaturation.[15] Glycero- and sphingophospholipids are the main types of phospholipids composing biological membranes (**Figure 1.1 a**).[16] Depending on the size and shape of their hydrophobic and hydrophilic portions (determined by the size and electrostatic properties of the head group, as well as the number, length and degree of unsaturation of the alkyl tails), phospholipids can self-assemble into various architectures such as micelles, extended bilayers or vesicles (**Figure 1.1 b**).

Within lipid bilayers, the structural features of phospholipids play additional important roles in long-range organization by influencing the bilayer phase behavior. Different phase behaviours (**Figure 1.1 c**) arise from lipid-lipid interactions that determine lipid packing and are highly dependent on the length and degree of unsaturation of the hydrophobic tails.[16] In typical biomembranes, liquid disordered (L_d), or fluid, phases tend be enriched in unsaturated glycerophospholipids, while solid-like gel phases are enriched in sphingomyelin, a saturated long-chain lipid. Gel phases exhibit much denser packing than do L_d phases which, in contrast, allow

for a high degree of translational mobility. Sterols, a non-polar lipid class, such as cholesterol (**Figure 1.1 a**), are also important biomembrane components. Cholesterol participates in the formation of liquid ordered (L_o), or 'raft'-like phases which display the high degree of ordering of a solid and high translational mobility of a fluid at the same time.[17] Artificial membranes can be generated from lipid mixtures where mixture composition and temperature allow fine control over phase behaviour.[18]

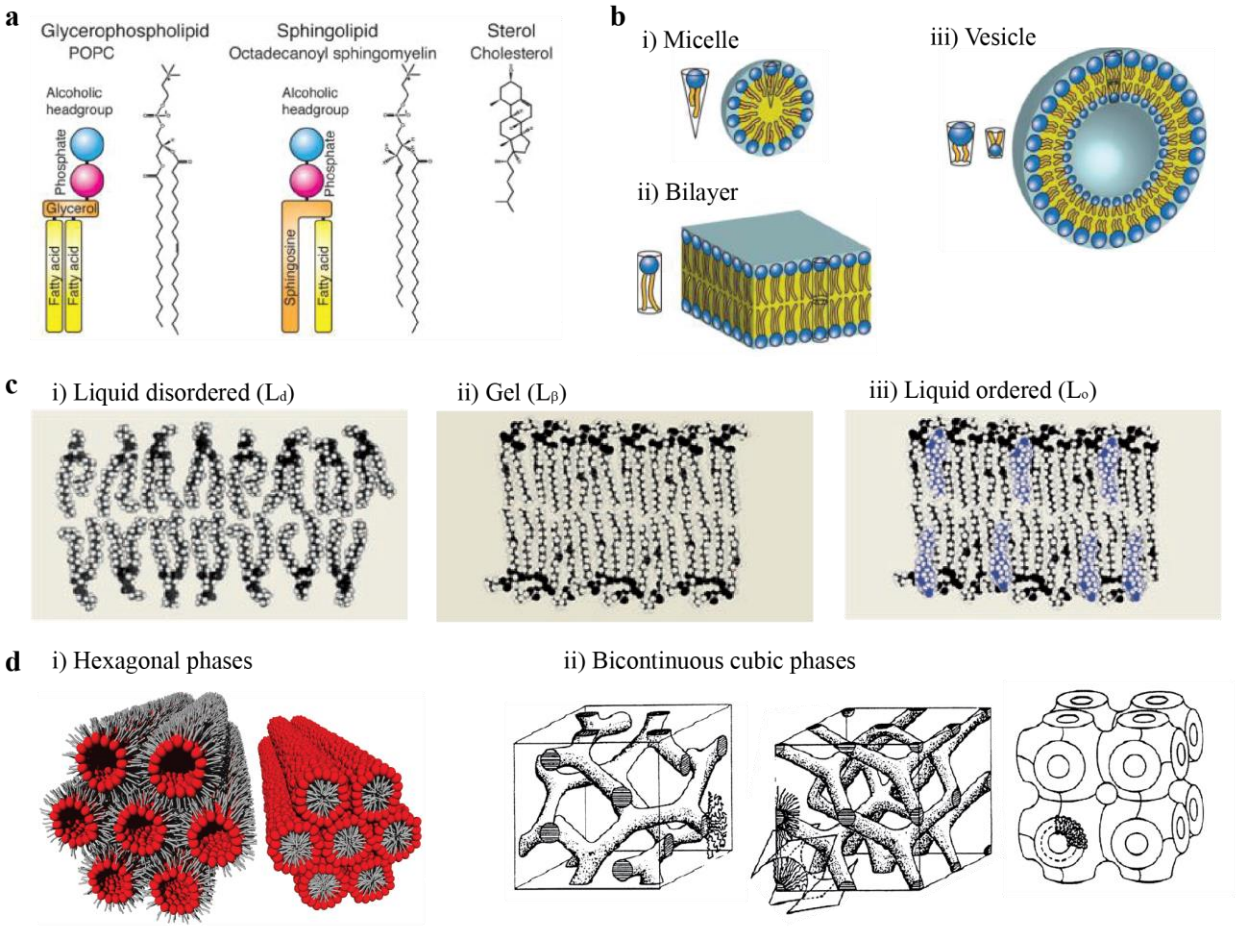


Figure 1.1| Biological lipids, their self-assembled structures and phase behaviours. a) Lipid types commonly found in biology: glycerophospholipid, sphingolipid and sterol. **b)** The shape of the lipid molecule highly influences the type of assembly it forms. Common lipid assemblies: **i)** Micelle, **ii)** Bilayer, and **iii)** Vesicle. **c)** Lipid bilayer phases: **i)** Liquid disordered (L_d), or fluid phase, **ii)** Gel phase (L_β), and **iii)** Liquid ordered (L_o), or 'raft' phase. **d)** Non-lamellar lipid structures: **i)** Hexagonal phases (inverted and micellar), and **ii)** Bicontinuous cubic phases. (Adapted with permission from: **(a)-(b)** Ref. 15, 2014 ACS; **(c)** Ref. 16, 2008 Springer Nature; **d)** Ref. 17, CC by 4.0)

Beyond bilayers or the lamellar phase (stacked bilayers separated by a thin water layer), biological lipids, surfactants, and their mixtures can organize into an array of non-lamellar long-range structures with 3D polymorphism (**Figure 1.1 d**).[17] Such structures have complex phase diagrams depending on molecular shape of the lipids, mixture composition, temperature and pressure. Hexagonal and cubic packing phases have lyotropic liquid crystal properties. Hexagonal phases are characterized by the formation of thin lipid cylinders which pack into a hexagonal lattice. On the other hand, cubic phases, such as a bicontinuous inverted phase, consist of continuous curved membranes folded into a cubic network and separating water compartments.[19] Both hexagonal and cubic non-lamellar phases are said to be ubiquitous in biological structures (e.g. mitochondria), and play a role in events such as membrane fusion.[17]

The properties of natural lipids and common surfactants have long served as an inspiration for the development of advanced materials. Beyond traditional amphiphiles composed of a small polar head group and alkyl chains, molecules incorporating π-stacking moieties,[20] metal binding complexes,[21] macrocyclic components,[22] peptides,[23] nucleic acids[24,25] or that are themselves products of self-assembly (supra-amphiphiles)[26] have generated much interest by expanding the repertoire of available amphiphilic building blocks for self-assembly. Furthermore, amphiphilic block-copolymers[27] are a class of materials with highly tunable self-assembly and stimuli-response properties. Overall, the still-growing class of amphiphilic building blocks and materials has broad application potential in, among others, drug delivery, catalysis, tissue engineering, photoelectric materials and stimuli-responsive surfaces.[19,28-31]

1.3.2 Supramolecular polymers and fibers

The term ‘supramolecular polymers’ first brings to mind their traditional counterparts, one-dimensional polymers built by covalently linking repeating units (monomers). Supramolecular polymers on the other hand arise following non-covalent assembly of the constituent monomers; the discussion here is limited to linear 1D assemblies. What supramolecular polymers lack in mechanical and chemical robustness commonly associated with conventional polymers, they make up in reversible and reconfigurable assembly, adaptive behaviour and stimuli-responsiveness – characteristics well-suited for the development of advanced materials with tunable properties, which can be recycled, and adapt to their environment.[32] Depending on monomer structure, the

collection of non-covalent interactions bringing the building blocks together, and the assembly conditions, supramolecular polymers can take on a variety of structures including ribbons, helical tapes and tubes.[33] The dynamic nature of polymer assembly also makes morphological transformations and hierarchical organization possible.[34]

Supramolecular polymers and fibers are ubiquitous in biological systems. Amyloid fibers,[35] cytoskeletal filaments such actin and microtubules,[10] extracellular matrix components such as collagen,[36] viral protein filaments[37] are all members of a vast supramolecular polymer family with important functions in biology.[38,39] For example, amyloids are a class of supramolecular polymers widely represented in nature. The formation of amyloid fibrils as a result of protein misfolding is most commonly associated with neurodegenerative diseases like Alzheimer's,[35] however amyloids also fulfil useful functions throughout biology, and represent a fundamental class of proteinaceous material.[40] Amyloid peptides, which can have a variety of amino acid sequences,[40] aggregate into fibrillar structures where they take on a cross-β conformation (**Figure 1.2**).[35] In this arrangement, the polypeptides are arranged perpendicular to the direction of fiber growth, and hydrogen bonding between adjacent chains results in β-sheet formation, driving longitudinal polymerization. Furthermore, fibrils exhibit higher order hierarchical organization through pairwise β-sheet packing into proto-filaments that then associate into polymorphic fibrils which vary in width and are typically several micrometers long.[41] Amyloid polymerization proceeds via a nucleation-growth cooperative mechanism, and is often irreversible owing to high kinetic barriers and/or large thermodynamic driving forces.[42]

In general, mechanisms of supramolecular polymerization are described by kinetic and thermodynamic aspects of the polymerization process and have a crucial influence on the structure of the resulting polymer.[32] Polymerization thermodynamics have been scrutinized at length, providing a solid theoretical foundation to describe the process.[43] The main mechanisms of polymerization are isodesmic and cooperative self-assembly. The isodesmic mechanism of supramolecular polymerization is similar to step-growth of covalently-linked polymers and implies that addition of each monomer to the growing chain is characterized by the same equilibrium constant K, such that the affinity of the polymer end is independent of chain length. Isodesmic polymerization typically gives rise to product populations with high polydispersity. On the other hand, the cooperative polymerization mechanism is characterized by at least two distinct stages. In the first step (nucleation), monomer association is described by linear growth with an

association constant K_n up to a total number of monomers s.[44] Beyond this point, various cooperative effects contribute to a change in the association constant to K_e (elongation), which is higher than K_n.[43,45] Cooperative polymerization is described as a nucleation-elongation mechanism due to the distinct stages of assembly.

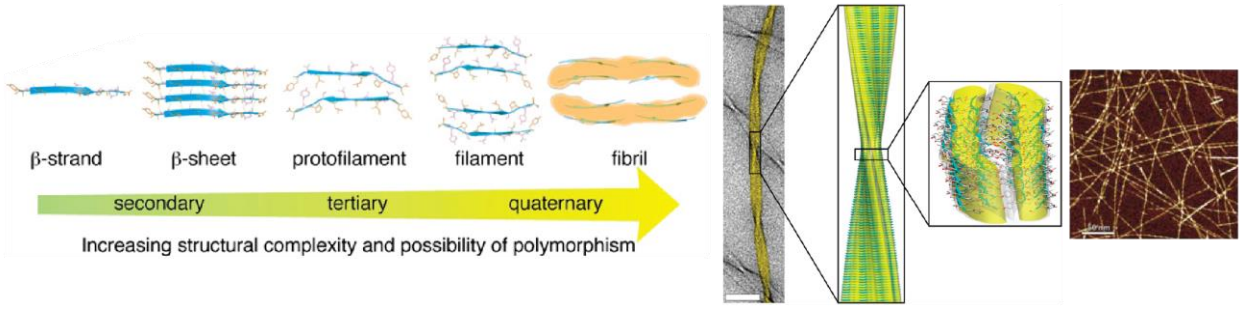


Figure 1.2| Hierarchical self-assembly of amyloid fibrils. The assembly of amyloid fibrils from short peptides displays hierarchical elements of architecture (secondary, tertiary and quaternary structure) mediated by non-covalent interactions. (Adapted with permission from Ref. 41, 2013 National Academy of Sciences.)

Hallmarks of the nucleation-elongation mechanism include (i) a time-dependent lag in polymer formation which (ii) can be eliminated with the addition of nuclei (seeding) at the onset of polymerization, and (iii) a critical monomer concentration or critical temperature associated with the onset of polymer growth.[43] Concentration- and temperature-dependent effects in general are very important in characterizing systems that undergo supramolecular polymerization.[43] For example, temperature plays different roles depending on whether the system undergoes assembly driven by entropy or by enthalpy. In entropy-driven polymerizations, there is a floor temperature below which no polymerization takes place. Entropy-driven polymerizations are often encountered in biological systems (e.g. amyloid fibril formation).[43] In contrast, polymerization processes driven by enthalpy are characterized by a ceiling temperature above which no polymerization is observed. Altogether, thermodynamic parameters of supramolecular polymerization have a profound impact on molecular structure, polymer robustness, chain length and length distribution in the product population, as well as responses to temperature or concentration changes.[32] Understanding this impact is essential in creating rationally designed supramolecular polymers with well-defined properties that are tunable to suit desirable functionalities.

Much of the rationally designed supramolecular polymerization processes described in the literature take place at equilibrium, under thermodynamic control. In many natural systems, on the other hand, advanced functionality is achieved by controlling assembly/disassembly kinetics and maintaining the process away from equilibrium, sometimes with continuous energy input, as in the case of remodeling of actin filaments.[46] The kinetics of supramolecular polymerization are generally less well described than the thermodynamics. However, there is increasing interest in examining and manipulating processes that take place under kinetic control, by harnessing kinetic barriers, taking advantage of seeding, hysteresis effects, pathway-dependent self-assembly outcomes or the coexistence of several metastable products.[32,46-54] In principle, systems assembled under kinetic control or far away from equilibrium with continuous energy supply have a vast potential for structural and functional diversity not accessible through more conventional self-assembly pathways.[54]

An great variety of synthetic supramolecular polymers has been developed over the last three decades. Although multiple non-covalent interactions tend to be involved in the polymerization process, it is convenient to classify the monomers that undergo supramolecular polymerization according to their primary assembly motifs. The Lehn group reported several landmark supramolecular polymerization studies of flexible and rigid building blocks based on molecular recognition mediated by hydrogen bonding between uridine and diacyl-amino-pyrimidine.[55,56] The ureido-pyrimidinone group capable of dimerizing via quadruple hydrogen bonding (**Figure 1.3 a**), developed by Meijer *et al.,* gave rise to random coil supramolecular polymers with mechanical properties similar to conventional polymers.[57] Inspired by DNA assembly, where hydrogen bonding associations are protected from competing interactions with water inside a stacked structure, building blocks containing the related ureido-triazine group as well as hydrophobic moieties that drive assembly due to solvophobic effects and solubilizing peripheral groups assembled in water and organic solvents into columnar aggregates with controllable helicity.[58,59] DNA structure (hydrogen-bonding recognition motifs and burial of hydrophobic surfaces within stacked assemblies) also served as a model for the development of helical rosette nanotubes, where the rosette refers to a six-membered hydrogen-bonded supermacrocycle (**Figure 1.3 b**).[60,61] The assembly of such nanotubes was shown to be entropically driven, a finding which be explained by invoking the hydrophobic effect, wherein burial of hydrophobic surfaces within the polymer releases ordered water molecules into bulk solvent for

an entropic gain.[62] The discotic benzenetricarboxamide (BTA) scaffold was used to investigate the effect of building block chirality on assembly cooperativity through the addition of chiral and achiral substituents (**Figure 1.3 c**),[63] and to examine supramolecular polymer dynamics using super-resolution microscopy.[49,64]

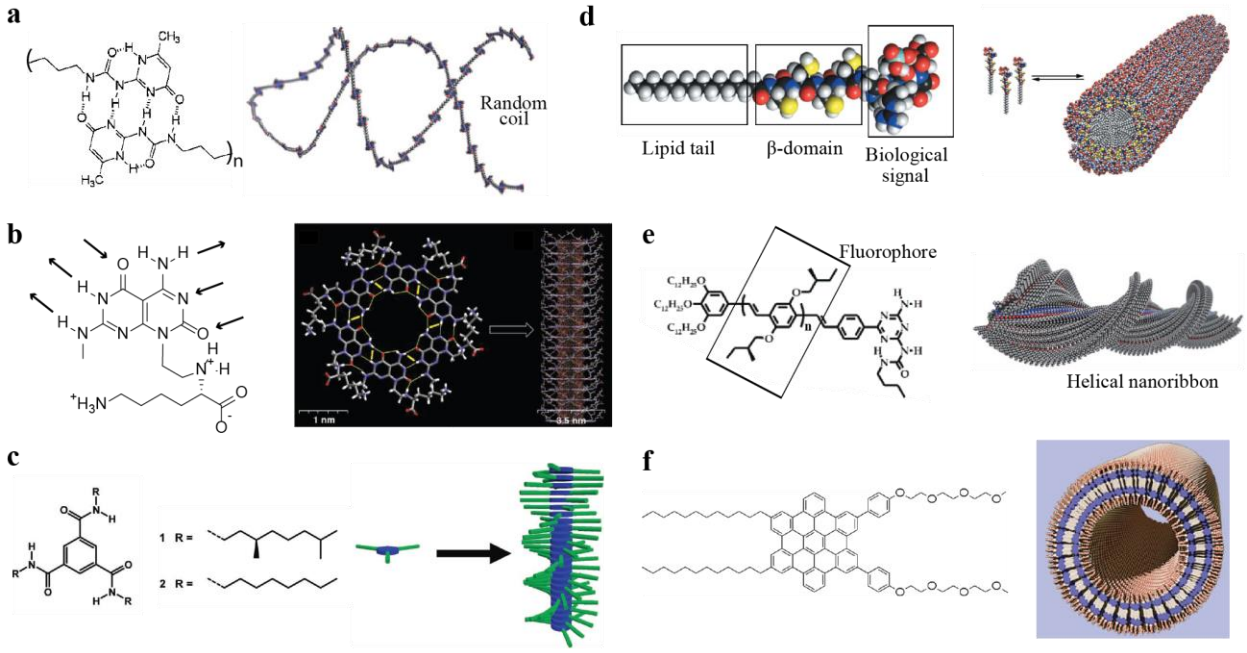


Figure 1.3| Rationally designed supramolecular polymers and fibers. a) Monomers with ureidopyrimidinone (UPy) termini assemble via quadruple hydrogen bonding into flexible random coil supramolecular polymers. **b)** Hydrogen-bonding monomers with recognition units resembling DNA bases form six-membered rosette arrays which stack into nanotubes with helicity directed by a chiral substituent. **c)** Monomers with a C_3-symmetric discotic benzenetricarboxamide (BTA) core assemble with a higher degree of cooperativity when modified with chiral side chains. **d)** Peptide amphiphiles containing three distinct structural regions (lipid tail, β-sheet forming peptide domain, and hydrophilic bioactive domain) form extended nanofibers with helical features directed by β-sheet conformation. **e)** Monomers containing oligo(*p*-phenylene vinylene) (OPV) fluorescent groups and aliphatic tails dimerize via hydrogen-bonding groups and then stack into helical ribbons. **f)** Hexabenzacoronene (HBC) based monomers functionalized with aliphatic chains on one side and hydrophilic oligoethylene glycol on the other form nanotubes with tunable diameter and wall thickness dependent on the size of the substituents. (Adapted with permission from: **(a)** Ref. 57, 1997 AAAS; **(b)** Ref. 60, 2001 ACS; **(c)** Ref. 63, 2008 ACS; **(d)** Ref. 65, 2001 AAAS; **(e)** Ref. 66, 2006 AAAS; **(f)** Ref. 67, 2004 AAAS; **(a)**, **(e)**, **(f)** Ref. 33, 2012 AAAS.)

As discussed above, amyloid fibril growth is highly dependent on β-sheet conformation of adjacent peptide chains through hydrogen bonding between their backbone amide groups.[41] Many

synthetic supramolecular polymerization systems based on peptide building blocks rely on β-sheet formation for longitudinal growth. Ghadiri *et al.* developed nanotubes based on the cooperative polymerization of cyclic polypeptides composed of alternating D- and L-amino acids.[68,69] With hydrophobic modification to their outside surface, they could be inserted into lipid bilayer membranes to serve as artificial ion channels.[70] Interestingly, nanotubes assembled from α- and 1,4-substituted ε-amino acids followed an anti-cooperative polymerization mechanism (where $K_n > K_e$, in contrast to cooperative polymerization).[71] Peptide amphiphiles, developed by the Stupp group, are a class of peptide-based building blocks composed of an aliphatic tail, a peptide β-sheet-forming domain and an adaptable hydrophilic domain which is used to incorporate bioactive cues in the assembly (**Figure 1.3 d**).[65,72] The self-assembly of these amphiphiles into long, highly organized fibers is driven by the microphase separation of the hydrophobic tails, and the formation of a network of hydrogen bonds between the β-sheet-forming portions of the peptide. By modifying the hydrophilic portion, the surface of the fibers can be functionalized with a high concentration of bioactive moieties or other functional groups making the nanofibers a highly tailorable material for tissue engineering and other advanced applications.[23,73,74]

The polymerization of building blocks with extended π-systems is another active area of research. By combining the ureido-triazine moiety mentioned above with oligo(*p*-phenylene vinylene) (OPV) groups bearing chiral side chains (**Figure 1.3 e**), Meijer *et al.* were able observe the cooperative assembly of helical tapes, this time incorporating long-range organization of photoactive groups that may be useful in the development of optoelectronic devices.[75,76] A detailed investigation of the nucleation-elongation mechanism of this system yielded important insights into the thermodynamics of its assembly, including the origin of the high degree of cooperativity which is related to a helical transition in the nucleating species, and the influence of solvent structure.[66] Furthermore, a time-resolved examination of the assembly process allowed quantification of kinetic parameters, and revealed the formation of metastable intermediates, illustrating pathway complexity in supramolecular polymerization.[77] The Aida group developed monomers with a π-extended hexabenzacoronene (HBC) core modified with aliphatic chains on one side, and hydrophilic oligoethylene glycol chains on the other (**Figure 1.3 f**). These monomers assembled into micrometer-long nanotubes with a consistent diameter and wall thickness dependent on the size of the peripheral chains.[67] The π-stacking of the HBC units within the nanotube wall endowed the structures with a "graphitic" character, exhibiting electrical

conductivity upon oxidation. Modification of the hydrophilic side of the monomer with an electron-accepting trinitrofluorenone (TNF) group, yielded nanotubes with electron-accepting TNF layers on their outside surfaces, and electron donating graphitic layers inside the walls, demonstrating spatial separation of charge carriers within a self-assembled nanostructure and producing a quick photoconductive response.[78] Furthermore, modification of the monomer with fullerene (C_{60}) gave rise to nanostructures displaying heterojunction behaviour (electron-transporting fullerene layers, and hole-transporting graphitic layers).[79]

The above discussion represents a small sampling of the still growing field of supramolecular polymerization. The advantages of these self-assembled materials include ease of preparation and processing, stimuli-responsive properties and reconfigurable, self-healing behaviour. Polymeric supramolecular materials of fibrous or rigid rod morphology can find a variety of applications, especially when they incorporate additional functionalities such as the display of bioactive cues, scaffolding of photoactive groups or organization of metallic or electronic components. Polymerized systems assembled under thermodynamic control can yield stable structures with well-defined properties, while recent interest in assembly under kinetic control or away from equilibrium opens the way for additional levels of structural complexity and adaptability to their environment in the materials which can be created using self-assembly.

1.4 DNA: Information-bearing molecule *par excellence*

Deoxyribonucleic acid (DNA) has been termed the molecule of life, as it carries genetic information in all living organisms from one generation to the next. DNA is an information-rich polymer, a molecule that exemplifies the principle of structure determining function. Under physiological conditions, DNA is most commonly found as a double helix composed of two complementary strands. The formation of this structure is driven by supramolecular interactions, such as hydrogen bonding between complementary base pairs – adenine (A) with thymine (T), and cytosine (C) with guanine (G), π-stacking between their surfaces, and electrostatic repulsion between the backbone phosphate units (**Figure 1.4 a**). The binding of complementary bases is perhaps the best-known example of molecular recognition. Structurally, DNA is most commonly found as a right-handed, antiparallel double-helix termed the "B-form", that has a 2 nm diameter and a 3.4 nm or 10.5 base pair pitch (height of a complete helix turn).[80] Other duplex types such

as A- and Z-form helices can be observed depending on the surrounding environment (**Figure 1.4 b**). Notably, double-stranded DNA can take on the right-handed A-form helix, a more compact structure with a wider diameter, in conditions of high salt and low humidity.[82] In the Z-form, DNA takes on a left-handed helix conformation that is most often encountered for sequences with alternating purine-pyrimidine bases such as $d(GC)_n$.[83]

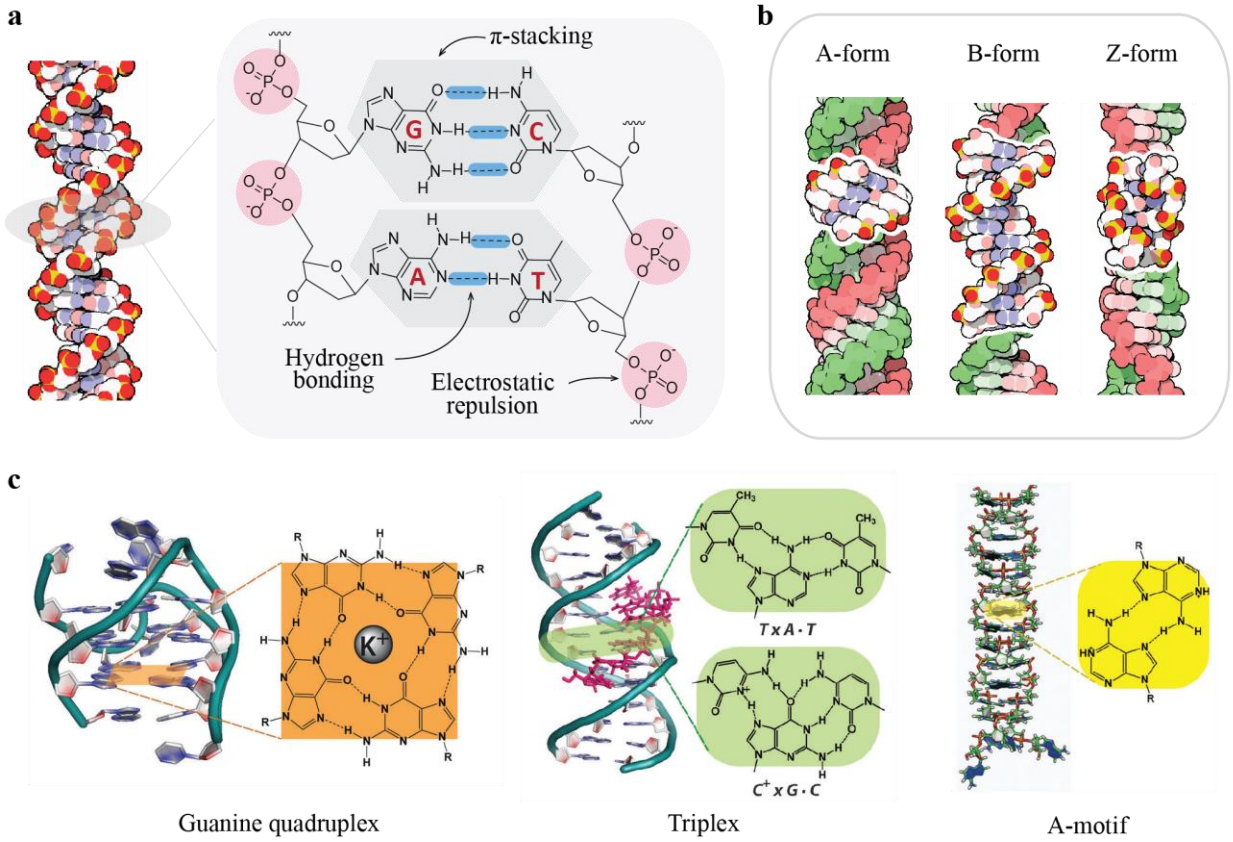


Figure 1.4| DNA, the molecule and its structures. a) The formation of a DNA double helix from complementary single strands occurs due to hydrogen bonding between complementary bases – adenine (A) binds to thymine (T), and guanine (G) to cytosine (C) – on their Watson-Crick faces, π-stacking between base-pair surfaces, and electrostatic repulsion between negatively charged phosphate groups in the backbone. **b)** A-, B- and Z-form DNA duplexes. **c)** Alternative naturally occurring DNA structures (from left to right): Guanine quadruplexes form from G-tetrad stacking following their assembly via Hoogsteen-face hydrogen bonding; DNA triplexes can form when purines A and G bind with their complement, T and C respectively, on both the Watson-Crick and Hoogsteen faces (C must be protonated); the A-motif is a parallel poly(adenine) duplex that forms in acidic conditions and involves Hoogsteen-face hydrogen bonding. (Adapted with permission from: **(a)-(b)** Ref. 81, 2001 RCSB PDB, **(c)** Ref. 83, 2010 RSC.)

Unmodified DNA strands are also known to assemble into a variety of alternative structures through non-standard base-pairing, in contrast to Watson-Crick base-pairing encountered in the DNA double helix.[83] Hoogsteen base-pairing is an alternative base-pairing mode that involves the N7 atom and C6 group of purines that adopt a *syn* rather than *anti* conformation (rotation around the glycosidic bond) and the Watson-Crick binding face of pyrimidines.[84] This diversity in structure often emerges in repetitive DNA sequences. Large portions of the genetic material that were once considered "junk DNA" contain various repetitive elements (terminal repeats, satellite DNA, transposable elements, etc.), many of which contain repetitive DNA sequences.[85] These areas of the genome are increasingly recognized for their structural and regulatory roles in genome function.[86] Formation of non-B-DNA structures from repetitive DNA sequences is thought to induce genetic instability, potentially contributing to disease states by promoting mutations or chromosome breakage.[87]

Examples of non-B-DNA structures include guanine quadruplexes, i-motifs, triple-helices, and the A-motif (**Figure 1.4 c**). Guanine-rich strands can fold into guanine quadruplexes which consist of π-stacked planar G-tetrads which are held together with a total of eight hydrogen bonds in the manner of Hoogsteen base-pairs.[88] The i-motif structure forms in C-rich regions in acidic conditions as two parallel-stranded C-C^+ hemi-protonated duplexes that intercalate with each other in an antiparallel orientation.[89] Both guanine quadruplexes and i-motif structures have been identified in human cells *in vivo*, supporting the notion that they and other non-B-DNA structures may have regulatory roles in the genome.[90,91] Among other alternative structures, prototypical DNA triplexes consist in a Watson-Crick base pair with an added Hoogsteen-paired base[92] while the A-motif is a parallel duplex formed from poly(A) DNA strands in acidic conditions.[93]

Ribonucleic acid (RNA) is the more structurally and functionally diverse counterpart of DNA.[94] It is important to note that RNA has an additional 2' hydroxyl group on the sugar moiety, and a G, C, A and U (uracil, a thymine analog lacking a methyl group at the C5 position of the pyrimidine ring) base-alphabet. As a duplex, RNA adopts the A-form. The importance of the roles of RNA in living organisms cannot be overstated, as it is involved in all steps of gene expression and regulation.[95]

Over the past three decades, investigations into DNA structure and self-assembly through sequence design and strand functionalization have expanded its role beyond that of a biological molecule.[96] DNA is an attractive building block for nanomaterials by virtue of its highly specific

molecular recognition properties and well-defined structure.[97] DNA's ability to base-pair with a complementary strand and its precise sequence programmability impart it with an unmatched level of molecular control. Due to its ubiquity among living organisms, DNA is a readily available material, and shorter nucleic acid strands (up to 200 bases) of any desired sequence can be obtained by automated synthesis at steadily decreasing costs. Furthermore, biological and chemical strategies can be used to functionalize and modify nucleic acids to imbue them with designer properties.

1.4.1 Chemical synthesis of DNA

Modern automated DNA synthesis relies on solid support chemistry, typically carried out on controlled-pore glass or, less frequently, polystyrene beads. The main advantages of solid support chemistry lie (i) in allowing reaction by-products to be washed away after each reaction step, contributing to increased yields and significantly simplifying purification, (ii) in driving reactions to completion by adding liquid-phase reactants in excess, and (iii) in the cyclic nature of the reaction sequence amenable to automation. Letsinger reported DNA synthesis on a solid support[98] around the same time as Merrifield demonstrated the success of the method for peptide synthesis.[99] The prevalent coupling approach in automated DNA synthesis uses phosphoramidite chemistry, a fast and high yielding method developed by Caruthers and Beaucage[100,101] following earlier work on more unstable nucleoside 3'-phosphodichloridite building blocks employed by Ogilvie in the first automated DNA synthesizer.[102]

The automated solid-support synthesis cycle of a DNA strand using phosphoramidite chemistry (**Figure 1.5**) consists of four steps, repeated for the addition of each base to yield a sequence controlled, monodisperse polymer product: i) Deprotection; ii) Coupling (with activation); iii) Capping; and iv) Oxidation. Synthesis is carried out from the 3' to the 5' end of a given strand. Deprotection removes the 5'-OH protecting group (usually a dimethoxytrityl group or DMT) under acidic conditions (dichloroacetic acid is used with DMT). Coupling of the next monomer, a 5'-DMT protected deoxynucleoside (2-cyanoethyl, N,N-diisopropyl) phosphoramidite, is carried out under mildly acidic conditions for activation purposes, typically with the addition of a tetrazole derivative. A capping step then takes unreacted strands out of the reaction cycle, ensuring sequence control and product monodispersity by passivating free 5'-OH

groups with an acetyl moiety. Oxidation with iodine converts phosphorus (III) to phosphorus (V), a more stable form, yielding a phosphate triester. The cycle can begin again for the next monomer addition, until the desired product is obtained, at which point the final DMT can be deblocked, followed by strand cleavage off the solid support and deprotection of the bases and backbone under basic conditions. The product can then be further purified to remove incomplete sequences. It is important to note that nucleic acids derived from biological sources or synthesized enzymatically are also widely used in nanotechnology.

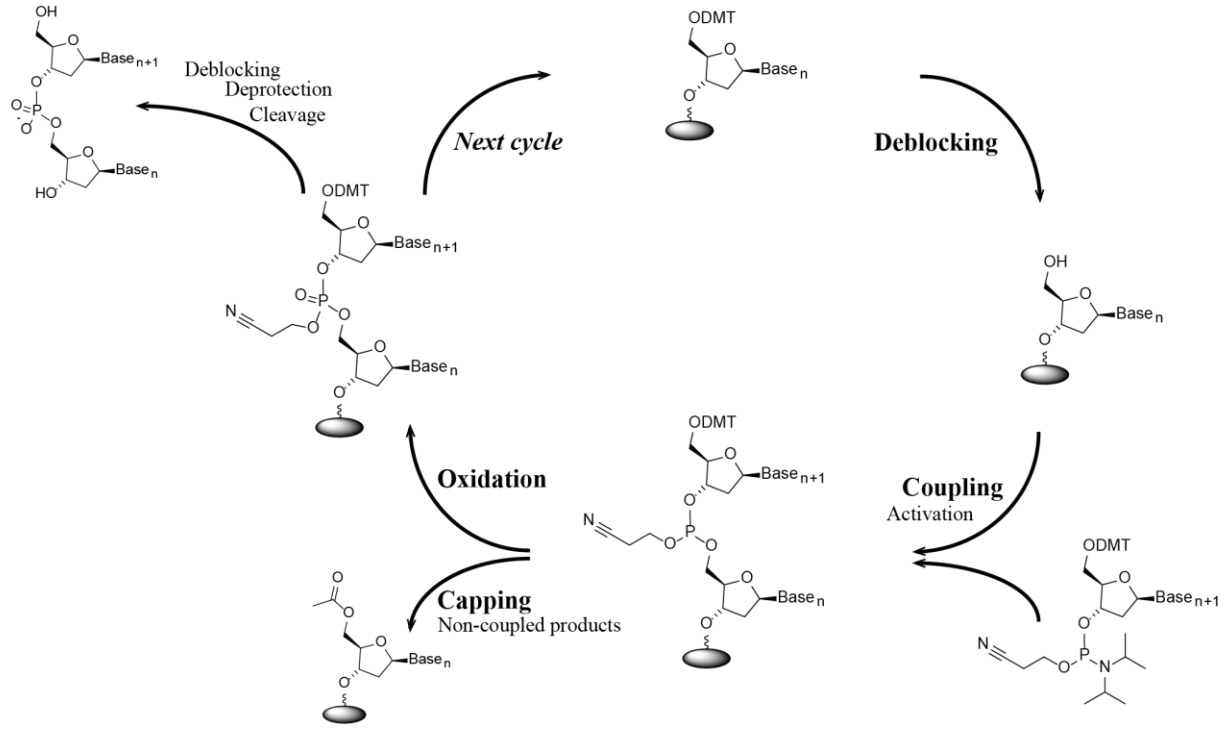


Figure 1.5| Automated phosphoramidite-based DNA synthesis on a solid support.

The development of automated DNA synthesis over the last six decades has had a transformational impact in areas such as molecular and synthetic biology, molecular diagnostics, sequencing technologies, DNA-based nanomaterials and computation, as well as oligonucleotide therapeutics.[103,104] Today, emerging technologies that scale up synthesis and facilitate recombination of gene fragments (such as the CRISPR-Cas9 system)[105] promise seemingly endless possibilities for nanotechnology and *de novo* genome engineering for 'designer' organisms among other innovative applications.[106]

1.4.2 Synthetically modified DNA

Beyond the canonical double helix, DNA is known to adopt a variety of alternative structures as outlined above. These structures highlight the versatility of the standard DNA base-pairing alphabet (composed of A, T, G and C). However, considering the wealth of intermolecular interactions available in supramolecular chemistry and the potential variety of structure and added function achievable with modified base pairing geometries and binding schemes, the natural base-pairing alphabet is limited. The field of synthetic DNA modification has expanded the genetic alphabet beyond hydrogen bonding interactions, and modifications to the DNA backbone have afforded new properties to oligonucleotides, including increased stability to degradation in physiological environments. Additional aims of the field include acquiring a better understanding of the principles behind nucleobase binding selectivity, DNA function and evolution.[107] The selective incorporation of new properties into DNA afforded by synthetic modification is highly relevant to synthetic biology, biotechnology and nanomedicine. Synthetic genetic polymers containing non-natural components such as alternative backbones, sugars or nucleobases have been termed "xeno-nucleic acids" (XNAs)[108] and enzymes facilitating their incorporation into sequences have been evolved.[109,110] Furthermore, from a nanomaterials standpoint, a larger DNA base-pairing alphabet opens the possibility of incorporating orthogonal interactions into DNA assembly, thus expanding the repertoire of achievable structures and increasing their complexity, while at the same time potentially reducing defects in assembly.

Alternative backbones to scaffold nucleobases have been developed since the late 1980s (**Figure 1.6 a**). The study of the etiology of nucleic acid structure undertaken by Eschenomoser, led to the development of a number of alternative sugar-based nucleic acids capable of hydrogen bonding.[111] Homo-DNA is a notable example, as it forms duplexes with higher stability than DNA while not being able to bind to DNA or RNA, thus representing the first orthogonal nucleic acid binding system.[112] On the other hand, locked nucleic acid (LNA), is a prime example of a nuclease-resistant nucleic-acid analog with high DNA binding affinity.[113] It was designed according to the principle of conformational restriction and has important therapeutic applications in gene silencing technologies.[114] Peptide nucleic acid (PNA) is another analog that was first developed for antisense applications.[115] It has a flexible achiral polyamide backbone and has been widely investigated for its nucleic acid recognition properties[116] in biotechnology, therapeutics and sensing applications.[117]

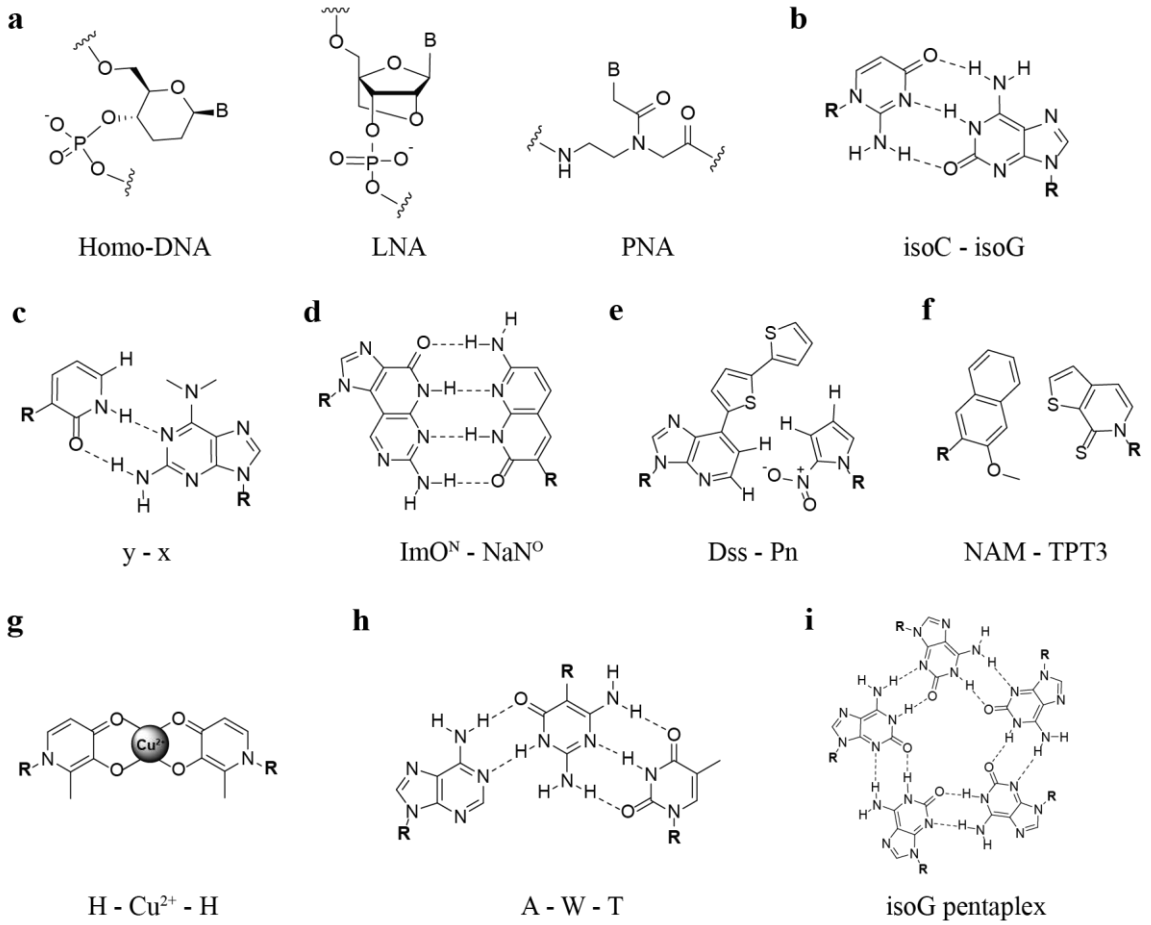


Figure 1.6| Alternative backbones and synthetic base pairs. a) Modified backbones: Homo-DNA, locked nucleic acid (LNA) and peptide nucleic acid (PNA). **b)** Base pair with alternative hydrogen-bonding pattern (isoC-isoG). **c)** Hydrogen-bonding base pair with shape complementarity (y-x). **d)** Size-expanded base pair with four hydrogen bonds (ImO^N-NaN^O). **e)** Shape complementary base pair without hydrogen bonding. The bases represent a fluorophore (Dss)-quencher (Pn) pair. **f)** Artificial third base pair incorporated in a semi-synthetic organism that is able to indefinitely store and use this new genetic information (NAM-TPT3). **g)** Cu^{2+} mediated artificial DNA base pair (H-Cu^{2+}-H). **h)** Synthetic Janus-Wedge type nucleobase binding to A on one side and T on the other (A-W-T). **i)** isoG pentaplex.

The modification of nucleobases has received significant attention in the last three decades. The Benner group pioneered base-pairs with alternative hydrogen bonding patterns (**Figure 1.6 b**) and worked toward ensuring their recognition by DNA polymerases, highlighting the utility of artificial base-pairs to expand the genetic code.[118,119] The incorporation of a stable and replicable additional base pair has been used in alternative codons to synthesize proteins with unnatural amino acids or designer functional groups.[120] In addition, base pairs that combine alternative

hydrogen bonding and shape complementarity considerations (**Figure 1.6 c**)[121] or size-expanded scaffolds (**Figure 1.6 d**)[122] have been developed. Bases with size-expanded scaffolds can have enhanced stability compared to natural DNA due to additional stacking interactions or hydrogen bonds. Aromatic size-expanded scaffolds can also have interesting fluorescence properties that are highly applicable in biophysical, biotechnological and nanomaterials contexts.[123,124]

The Kool group has introduced an entirely different class of base pairs by developing non-polar purine and pyrimidine isosteres that eliminate hydrogen bonding interactions. The first generation of non-polar bases were non-selective in binding to natural nucleobases, but could recognize their designed non-polar complements specifically.[125] Such structures represent excellent probes to examine the contributions of hydrogen bonding, polar and steric interactions to the properties of DNA and its recognition by polymerases.[126-128] Hirao and coworkers focused more specifically on shape complementarity in the absence of hydrogen bonding.[129,130] An interesting example of their work combines a fluorophore and a complementary quencher nucleobase analogs that can be used as beacons for real-time PCR among other potential applications (**Figure 1.6 e**).[131] Recently, the Romesberg group, after creating a number of hydrophobic, hydrogen-bonding-free base pairs over the last decade,[132-134] has achieved a milestone in expanding the functional base-pairing alphabet. The group developed a semi-synthetic organism (*e.coli* bacterium) with an expanded six-letter genetic alphabet (a third base pair),[135] further pushing the methodology so that the new base pair participates in additional codons that incorporate artificial amino acids into peptides produced by the organism, succeeding to indefinitely store and retrieve genetic information encoded in the additional base pair (**Figure 1.6 f**).[136,137]

Canonical Watson-Crick base pairs have also been replaced with metal coordinating complexes to impart new properties to DNA. Natural DNA mismatches can act as binding sites for metal cations, for example metal mediated T-Hg^{2+}-T[138] and C- Ag^{+}-C[139] base pairs have been reported and used for sensing and logic-gating.[140] Recently, an unmodified DNA duplex has been used to create a continuous hybrid silver nanowire through linear coordination of an Ag^{+} ion at every rung.[141] Shionoya and coworkers have developed a variety of metallo-base pairs adhering to important design principles such as ensuring size match to natural base pairs and maintaining planarity by using linear or square planar metal coordination geometry.[142] The hydroxypyridone-based Cu^{2+} binding H- Cu^{2+}-H base pair showcases these features (**Figure 1.6 g**).[143] Metal base-

pairing has also been incorporated into alternative DNA structures, for example three pyridine-bearing bases coordinated to Ag^+ can form a base triplet inside a triple helix.[144] In many cases, metallo-base pairs exhibit higher thermal stability than natural DNA, can introduce magnetic and conductive properties into DNA structures, be used in new nanoarchitectures, introduce dynamic stimuli-responsive function, and impart DNAzyme activity to structures.[142]

Designer nucleobases with multiple hydrogen-bonding binding faces have also been developed to access artificial alternative and higher order structures from DNA. Janus-Wedge nucleobases based on recognition motifs first described by Lehn[145] were organized on a PNA backbone to make a triple helix. The artificial nucleobases can wedge themselves into a natural base pair as they have two binding faces, recognizing A on one side and T on the other (**Figure 1.6 h**) (analogously for G and C).[146] Melamine has also been polymerized using a PNA backbone and bound symmetrically to poly(thymine) DNA strands in a 1:2 ratio forming a triplex.[147] Strands composed of isoG bases (**Figure 1.6 i**) and some of its analogs were shown to form quadruplexes and, interestingly, pentaplexes in the presence of certain cations.[148,149]

1.4.3 Self-assembly implications in the prebiotic origins of nucleic acids

There is some evidence supporting the hypothesis that life based on DNA genomes and protein enzymes was preceded by a simpler form based primarily on RNA – the so-called "RNA world hypothesis".[150,151] Within this framework, DNA may have evolved into its role as the information-storage polymer due to its superior chemical stability, while proteins may have acquired a multitude of catalytic and structural functions due to the variety of functional groups available in natural amino acids and the complex folding topology of polypeptides. Still today, RNA has a multitude of roles in cell function. For example, the peptidyl transferase center of the ribosome, the machinery of protein synthesis, is a ribozyme (an RNA-based enzyme), highlighting the central importance of RNA function in protein synthesis. It is viewed as a molecular fossil of a potential RNA world.[152]

In this context, the prebiotic emergence of RNA has been a fascinating problem for researchers. It is possible that RNA first emerged from the spontaneous non-enzymatic coupling of pre-existing molecular building blocks, but another potential scenario is that RNA evolved from pre-RNA polymers containing ancestral recognition units and alternative backbones.[153] Three

main models of RNA evolution have been advanced. The "classical" model proposes that nucleobases (or alternative recognition elements), sugars and phosphates successively coupled to each other into units that then polymerized. The "ribose-centric" model suggests that nucleobases formed on the ribose sugar from simpler components. Finally, the "polymer fusion" model, proposed by Hud, suggests that the first informational polymers resulted from a merger of supramolecular assembly of ancestral nucleobases with a pre-existing polymer that became a backbone.[154]

The last model was advanced following the discovery of several plausible early recognition elements, potential ancestors of our nucleobases. The first pair was 2,4,6-triaminopyrimidine (TAP) and cyanuric acid (CA), the latter is an accepted prebiotic molecule.[155] Derivatized for better solubility in water, they were shown to form hexameric rosettes that stacked into micron-long fibers with a 2 nm diameter through a highly cooperative and efficient mechanism in aqueous solution (**Figure 1.7 a**).[156] Later, condensation of TAP with ribose was demonstrated under prebiotic conditions (drying with mild heating in aqueous solution) and the resulting nucleoside (TARC) maintained its self-assembly properties with CA (**Figure 1.7 b**).[157] Furthermore, another complementary pair, melamine (M), also an accepted prebiotic molecule,[155] and barbituric acid (BA), were shown to spontaneously form glycosidic linkages with ribose and ribose-5-phosphate under prebiotic conditions, thus yielding proto-nucleosides and nucleotides. The co-assembly of the nucleotides also yielded long supramolecular fibers with an underlying rosette architecture (**Figure 1.7 c**).[158] The spatial ordering and increased local concentration resulting from the preorganization of plausible proto-nucleotide monomers and their precursors into supramolecular polymers could have facilitated their covalent linking into early information polymers (**Figure 1.7 d**). In addition, these new polymers may have acted as templates for hexad-mediated replication.[159] Although this is by no means the only potential model for the evolution of nucleic acids, it represents an interesting perspective illustrating self-assembly principles in the context of prebiotic chemistry.

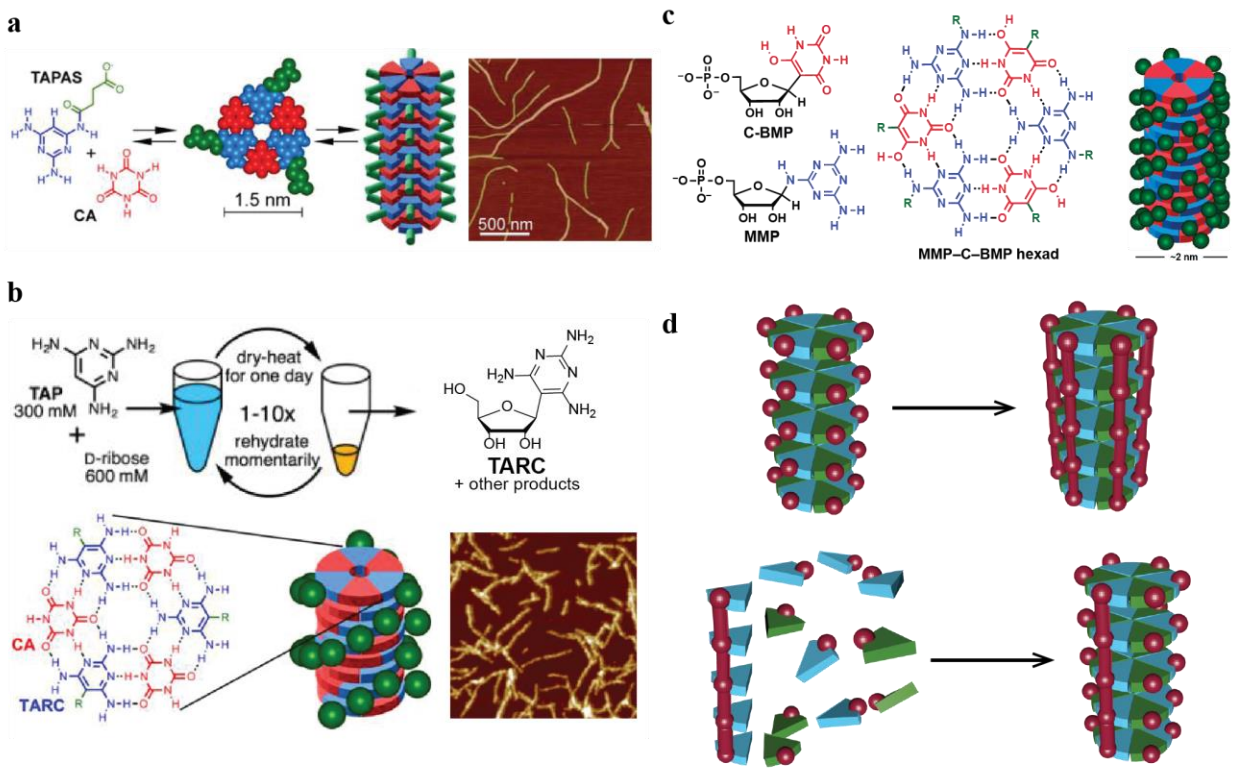


Figure 1.7| Self-assembly of plausible proto-nucleobases. a) TAP derivatized with succinic acid (TAPAS) assembles with CA into hydrogen-bonded hexads that stack into micron-long fibers with a 2 nm diameter. **b)** Spontaneous coupling of TAP to ribose (TARC) under prebiotic conditions. TARC forms supramolecular polymers with CA. **c)** Barbituric acid and melamine can form glycosidic bonds with ribose-5-phosphate under prebiotic conditions. The resulting complementary proto-nucleotides form hexads that stack into supramolecular polymers. **d)** Hypothetical conversion of the assemblies of monomeric proto-nucleosides into covalently linked polymers (top). Hypothetical hexad-mediated replication of early information polymers. (Adapted with permission from: **(a)** Ref. 156, 2013 ACS; **(b)** Ref. 157, 2014 ACS; **(c)** Ref. 158, 2016 Springer Nature.)

It should be noted that 1:1 co-assembly of M and CA is a very influential molecular recognition system in supramolecular chemistry, and forms the basis of many functional nanomaterials with long-range organization. Based on the complementary arrangement of their hydrogen bond donor and acceptor groups, M and CA form molecular aggregates of various geometries, such as hydrogen-bonded chains (or "tapes"), rings (also called "rosettes"), and infinite sheets (**Figure 1.8 a**) in solution and on surfaces.[160,161] The Whitesides group has demonstrated that the addition of different substituents on M and CA directs their assembly into specific architectures.[162,163] The aromatic planar ring structures of M and CA enable π-stacking

interactions between hydrogen-bonded aggregates to take assemblies into 3D (**Figure 1.8 b**). Although rosette formation is enthalpically favored in solvents such as chloroform due to the formation of 18 hydrogen bonds, the entropic change brought up by the association is highly unfavorable. Whitesides *et al.* have used preorganization (by tethering components to each other) and peripheral crowding (by adding bulky substituents), two key concepts in supramolecular chemistry, to remedy this effect, minimize the free energy of the assembly and promote rosette formation.[164,165] While in organic solvents M and CA are a canonical hydrogen-bonding pair, their assembly in aqueous solution is said to be largely driven by the enthalpically favorable surface area burial, a phenomenon also observed in nucleic acid recognition events.[166] Since the initial studies carried out in the early 1990s, the molecular recognition between M, CA and their derivatives has been used for the assembly of nanotubes, dynamic hydrogels, polymeric scaffolds, 2D surface assemblies and liquid crystalline materials for various applications such as small molecule encapsulation and sensing.[160]

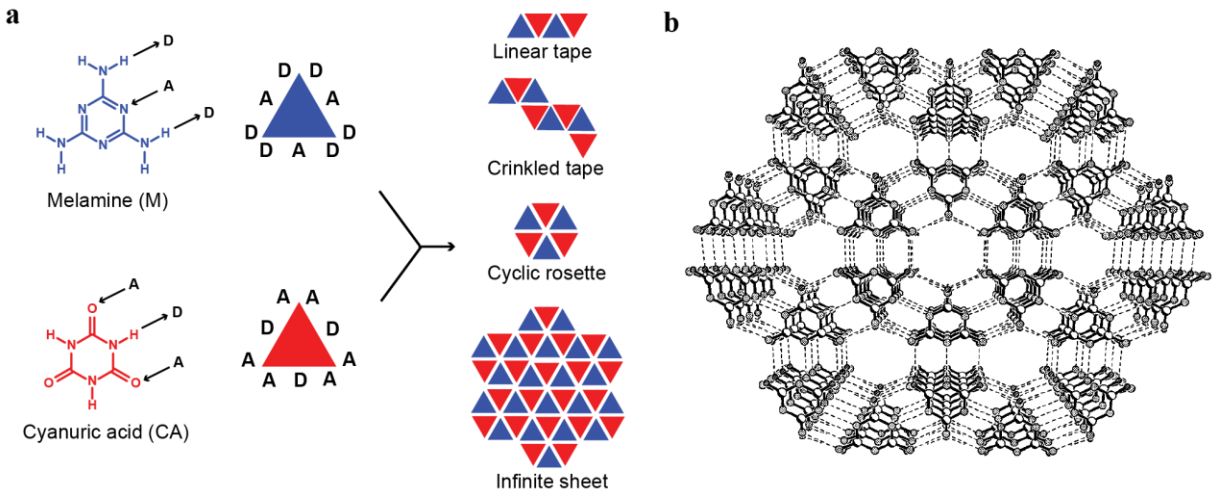


Figure 1.8| Self-assembly of melamine (M) and cyanuric acid (CA). a) Supramolecular architectures formed by M and CA based on the interactions between hydrogen bond donor (D) and acceptor (A) groups. Aggregates can be directed toward specific architectures with the addition of different substituents. **b)** 3D structure of a CA·M crystal obtained by hydrothermal synthesis. (Adapted with permission from: **(b)** Ref. 167, 1999 ACS.)

1.5 DNA nanotechnology: Using the molecule of life as a material

In the last three decades, the role of DNA in the consciousness of researchers spanning multiples fields, from crystallography to computer science, chemistry and photonics, has expanded beyond that of the carrier of genetic information to a molecular building block used to develop nanomaterials.[168] Since its inception, DNA nanotechnology, or self-assembly of highly ordered DNA materials based on sequence design and DNA functionalization, has yielded increasingly complex structures for a variety of potential applications.[168] Although DNA is inherently a one-dimensional (1D) object, its remarkable base-pairing fidelity, programmability, addressability and ease of functionalization can be harnessed to rationally design highly ordered nanomaterials in two- and three-dimensions (2D and 3D). Two main approaches, structural DNA nanotechnology and supramolecular DNA assembly, are used to create both discrete and extended structures. Structural DNA nanotechnology relies on both the precise DNA base pairing rules and the well-characterized structural features of DNA helices to build nanostructures.[97] On the other hand, by using DNA strands modified with organic or inorganic moieties, supramolecular DNA assembly harnesses not only base-pairing interactions, but also a wealth of supramolecular forces to drive the self-assembly of complex architectures and provide otherwise inaccessible geometry control and functionality within the designed structure.[169]

The discussion below mostly focuses on long-range assemblies, rather than smaller discrete structures from DNA. A number of recent publications review the assembly and applications of discrete DNA nanostructures.[168,170] Highly organized, long-range assemblies from DNA are of particular interest in the development of scaffolding materials and interfacing organized structures with different substrates. In addition, long-range DNA assemblies can help bridge the gap between top-down patterning and bottom up assembly methods in the quest toward materials with hierarchical structural properties and functionality built-in at each level of structure.

1.5.1 Tile-based assemblies

A large variety of highly ordered DNA networks, lattices and nanotubes has been developed using the principles of structural DNA nanotechnology. The Holliday junction, a four-arm intermediate in genetic recombination, is a prototype branched DNA construct found in

nature.[171] However, due to its instability and flexibility, this motif is not well suited for stable extended assembly. Seeman and coworkers first addressed this challenge by creating DNA branching units with immobile junction points.[172] Furthermore, the group pioneered the use of branched DNA constructs joined by cohesion via single-stranded DNA regions termed 'sticky ends' (**Figure 1.9 a**).[173,174] This method yielded a fully ligated DNA cube that sparked the development of the DNA nanotechnology field (**Figure 1.9 b**).[175]

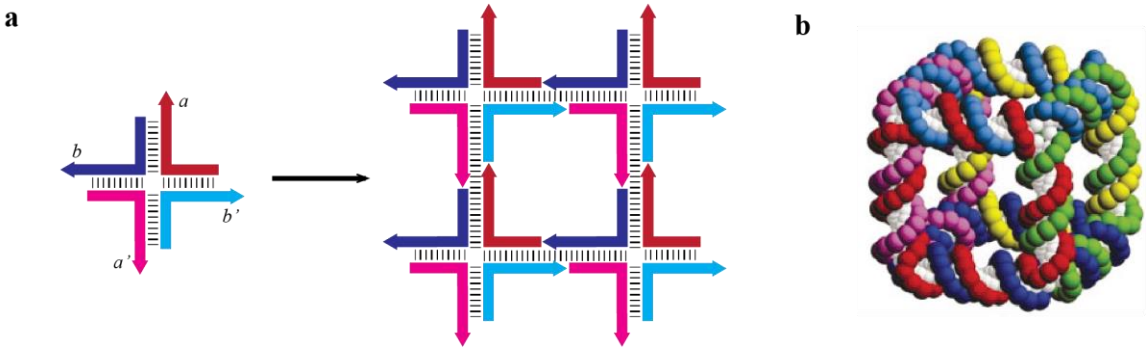


Figure 1.9| Assembly of branched DNA molecules. a) Assembly of branched DNA molecules through self-complementary single-stranded overhangs (sticky ends). Sticky end number and orientation determines assembly structure. In this case, four-way branched molecules with right angle orientation form a quadrilateral structure that is capable of further assembly through its unpaired sticky ends (sequence *a* complementary to *a'*, and *b* to *b*'). **b)** Ligated DNA cube. (Panel **(b)** adapted with permission from Ref. 174, 2003 Springer Nature.)

The next step toward stable long-range assembly was accomplished by creating crossover tiles, DNA constructs characterized by rigidity at the strand junction points. The double crossover (DX) molecule is composed of two double helical domains linked at two crossover points where a strand switches between adjacent helices, thus placing the two helices parallel to each other in a planar tile. DX tiles were used to make a 2D DNA crystal via sticky-end cohesion (**Figure 1.10 a**).[176] Following the initial DX tile design, DNA tiles grew in diversity. The main structural motifs can be categorized into planar tiles, branched tiles and helix bundles.[177] Planar tiles are based on the same principle as DX molecules, aligning a given number of DNA helices (3, 4, 6, 8, 12) side-by-side and linking them by strand crossover. Tiles obtained by this strategy formed 2D arrays with cavities of programmable size as determined by tile structure.[178-180] A development from these tiles, the DX triangle, was used to make arrays with pseudo-hexagonal geometry (**Figure 1.10 a**).[181]

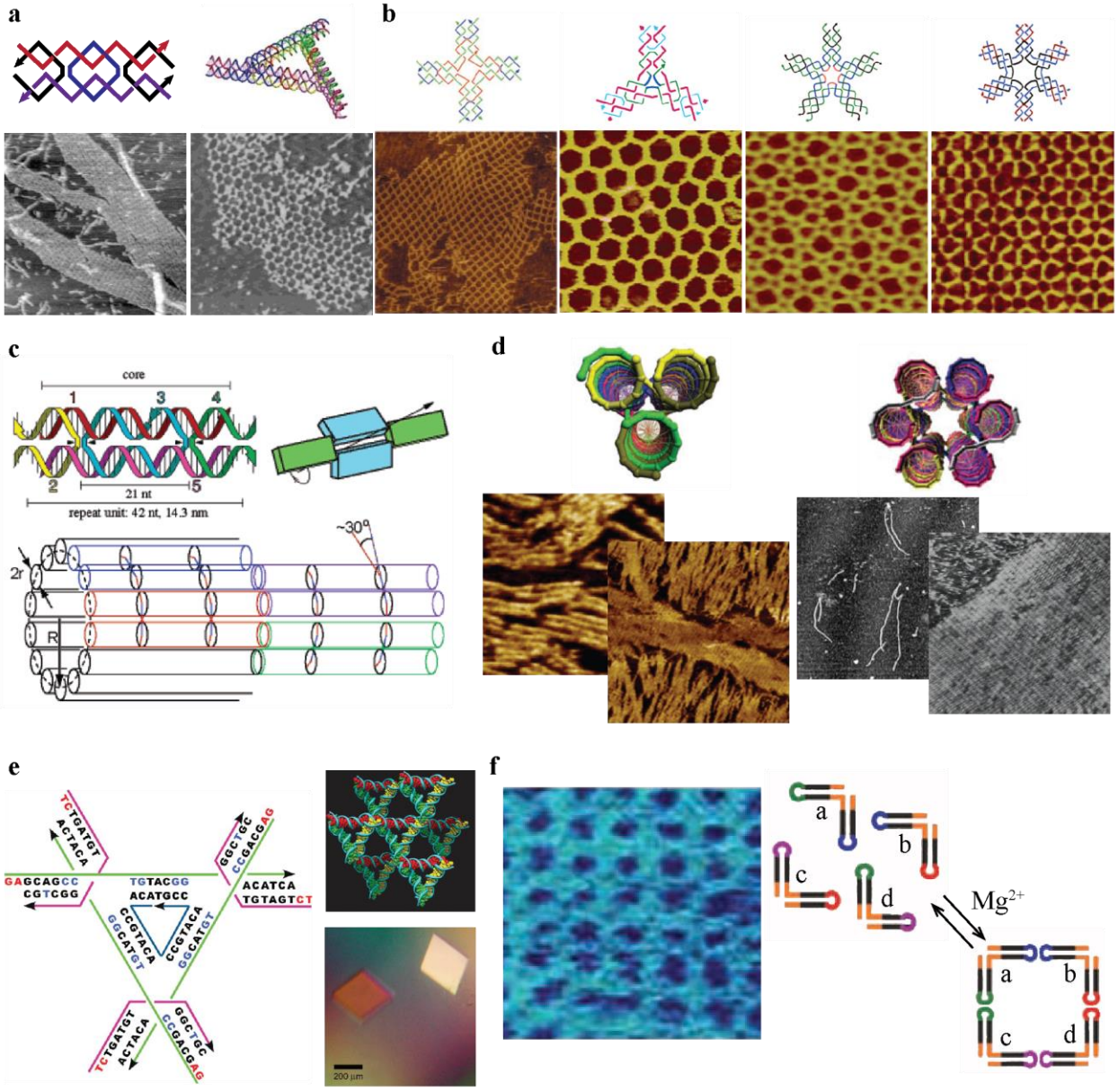


Figure 1.10| Tile-based assembly. a) Planar tiles: DX and DX triangle tiles. **b)** Branched tiles: Cross, 3-point star, 5-point star and 6-point star tiles. **c)** Nanotubes assembled from DX tiles by controlling crossover positioning and thus the dihedral angle between helices. **d)** Three-helix and six-helix bundles (3HB and 6HB). **e)** 3D macroscopic DNA crystals built from tensegrity triangle tiles. **f)** Assembly of lattices from RNA tectosquares. (Adapted with permission from: **(a** left**)** Ref. 176, 1998 Springer Nature; **(a** right**)** Ref. 181, 2004 ACS; **(b** left**)** Ref. 182, 2003 AAAS; **(b** second from left**)** Ref. 183, 2005 ACS; **(b** second from right**)** Ref. 184, 2008 National Academy of Sciences, U.S.A.; **(b** right**)** Ref. 185, 2006 ACS; **(c)** Ref. 186, 2004 AACS; helix bundle schemes in **(d)** Ref. 177, 2006 Wiley; AFM images in **(d** left**)** Ref. 187, 2005 ACS; AFM images in **(d** right**)** Ref. 188, 2005 ACS; **(e)** Ref. 189, 2009 Springer Nature; **(f)** Ref. 190, 2004 AAAS.)

Branched tiles have multiple rigid arms radiating from the center. Common morphologies include the cross motif[182] as well as 3-point,[183] 5-point[184] and 6-point[185] star motifs (**Figure 1.10 b**). Tile geometry once again determines cavity size and shape obtained upon lattice assembly. Yan and LaBean showed that branched cross-shaped tiles also formed nanotubes.[182] When connectivity between individual tiles oriented them in the same direction in the lattice plane, their intrinsic curvature accumulated resulting in nanotube rather than planar lattice formation. These tubes were used as scaffolds to produce uniform conductive silver nanowires up to 5 μm in length. T-junctions are a variation on the branched tile and can form lattices with high structural resolution due to their short interconnection length.[191] It is important to note that branched tiles can also be used to obtain discrete 3D assemblies in a modular fashion. The intrinsic flexibility of the central branching point in this type of tile can be controlled, by the same token modulating tile curvature. This control can be harnessed to obtain a series of discrete 3D polyhedra from 3-point star tiles[192] and precisely formed icosahedra from 5-point star tiles.[184] Furthermore, a greater structural variety has been achieved by combining tiles of different geometry and introducing directing agents, highlighting the versatility of the tile-based approach for building complex nanoscale objects.[193]

Although DNA nanotubes were observed along with 2D lattices in DX tile-based assembly, their formation was initially poorly understood. Rothemund *et al.* observed that with specific design of crossover points between DX tiles, tile arrays could form nanotubes up to 50 μm in length. The geometry of B-form DNA gives rise to non-planar tile juxtaposition depending on crossover point positioning, resulting in assembly curvature and tubular assembly (**Figure 1.10 c**).[186] Namely, separating crossover points by an integer number of helical half-turns promotes a planar arrangement of helices, whereas non-integer numbers break co-planarity between adjacent helices.

Furthermore, helix bundles, DNA tiles with built-in curvature, were developed by aligning a given number of DNA duplexes using multiple strand crossovers programmed to induce specific dihedral angles between adjacent helices. Three-[187] and six-helix bundles[188] (3HB and 6HB respectively) were the first examples of DNA tiles specifically designed to have non-planar geometry (**Figure 1.10 d**). For example, the 6HB design features a 7 nucleotide separation between strand crossover points producing 120° dihedral angles between helices, considering that a full helix turn is 10.5 bases long.[188] The helix bundle approach, as opposed to other tiling techniques that produce tubes of varied circumference, offers precise programmability of nanotube

circumference. Helix bundles were shown to form both nanotubes and 2D arrays using sticky end cohesion between building blocks.[187,188,194,195]

The Mao group introduced the concept of sequence symmetry to the field of structural DNA nanotechnology that previously relied on sequence uniqueness to promote the formation of a desired structure. Sequence symmetry minimizes the required sequence space and maximizes the number of times a given strand is used in the assembly, simplifying the sequence design process and reducing synthesis costs. An added advantage of this approach is the reduction of unpredictable interactions between strands and of distortions within the assemblies, promoting growth of 2D arrays to larger sizes. Systems designed with sequence symmetry were shown to grow networks up to 1 mm in size, orders of magnitude larger than their asymmetric equivalents.[196] Mao *et al.* also showed that a single strand can give rise to DNA nanotubes of uniform height and width. Two identical 52 nucleotide-long strands containing palindromic motifs associate to form DX-type tiles that assemble into planar arrays and further form tubular structures due to tile flexibility.[197]

Another strategy facilitating growth of extended networks is surface-mediated self-assembly. This method involves network formation directly on a mica surface where the constraint provided by surface interactions limits the flexibility of growing structures and prevents aggregate formation.[191,198] In a different sense, Yan and coworkers exploited the geometric symmetry of planar DNA tiles to design a cost-effective way to assemble finite-sized arrays. To achieve this, all tiles within the array were produced with the same core sequences and only sticky ends varied to precisely program connectivity between the tiles.[199]

A distinct type of triangle tile developed by Mao *et al.* on the principle of tensegrity, or the balance between flexible junctions and the rigid struts connecting them, is also known to form 2D arrays. The triangle is a stable structure composed of three DNA duplexes (rigid struts pushing outward) joined by four-arm junctions at the vertices (flexible junctions pushing inward) (**Figure 1.10 e**).[200] Remarkably, this construct was adapted by Seeman and Mao to build a DNA crystal of macromolecular dimensions (>250μm). More specifically, this achievement demonstrates that a DNA structure can be designed to self-assemble into a precise 3D crystalline lattice, one of the first aims for Seeman, the father of DNA nanotechnology.[189] More recently, a colour-changing crystal was developed, a first demonstration of a three-state 3D device that operates by isothermal strand displacement.[201]

RNA oligonucleotides can also be used as building blocks for the construction of nanostructures on their own, or in combination with DNA to make hybrid assemblies combining the programmability of DNA and the functional properties of RNA.[202] Jaeger and coworkers developed RNA structures termed "tectosquares" that self-assemble into squares through kissing loop regions and further form a lattice by sticky-end cohesion (**Figure 1.10 f**).[190] Aldaye, Silver *et al.* have shown *in vivo* assembly of RNA tile-based nanotubes and 2D scaffolds with well-defined protein docking sites, which associated with proteins in bacterial cells, serving as a spatial organization tool. This method demonstrated that rationally designed RNA structures can be built inside living cells and provide functional architectures therein.[203]

1.5.2 DNA origami and brick assembly

Scaffolded DNA origami, developed by Paul Rothemund, is an entirely different approach to structural DNA nanotechnology and has been transformative for the field since its introduction in 2006.[204] Using this method, virtually any arbitrary shape can be obtained in one step by folding a long genomic DNA strand with the help of short complementary strands (**Figure 1.11**). These short "staple" strands with computer-generated sequences hybridize to the scaffold introducing crossovers that effectively align and stitch together DNA duplexes to form the designed shape. In the initial report, a ~ 7000 nucleotide-long bacteriophage genome scaffold strand was used in combination with ~ 250 staple strands to produce 2D structures on the order of 100 nm in scale, with addressable spatial resolution of ~ 6 nm as determined by staple strand positioning.

The length of the scaffold strand is a limitation to the size of the structures produced by the DNA origami method. However, researchers have found ways to scale up assembly. Zhao *et al.* used planar eight-helix tiles equipped with single-stranded overhangs as "staple tiles" and quadrupled the surface area of the origami structures formed using the same scaffold strand as the traditional technique.[205] Furthermore, a strategy termed "superorigami" or "origami of origami" organizes origami tiles on a loose framework created from a single-stranded DNA scaffold held together by bridge strands.[206] On the other hand, a "jigsaw puzzle" strategy proposed by Sugiyama uses DNA origami tiles with sequence- and shape-complementarity to form larger structures.[207] Polymerization of rectangular tiles into nanoribbons with the addition single-stranded oligonucleotides to bridge an intermolecular scaffold seam has been described,[208] and Seeman *et*

al. have created crystalline 2D arrays of origami cross tiles using short complementary sticky ends at their interfaces.[209] Work from Rothemund later bypassed sequence-complementarity in the assembly of origami tiles into larger structures, demonstrating instead that π-stacking interactions between blunt-ended tile edges can bring them together.[210] Furthermore, the control of periodic long-range organization of origami structures on a mica surface has been achieved by mediating the surface diffusion of the structures with step-wise changes in counterion concentrations.[211] In addition, the generation and use of a ~ 51 000 nucleotide-long scaffold strand has been reported to yield origami structures seven times larger than the traditional method.[212]

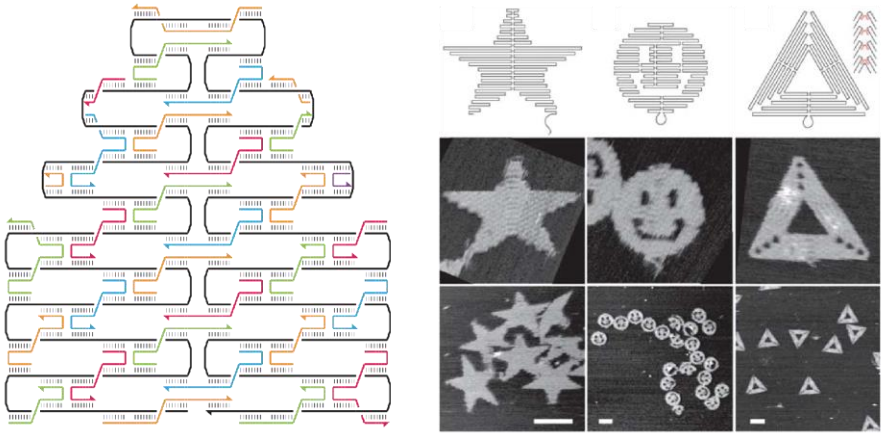


Figure 1.11| Scaffolded DNA origami in 2D. A long scaffold strand (left, black) is folded using short staple strands (left, coloured) into arbitrary 2D shapes (right). (Adapted with permission from Ref. 204, 2006 Springer Nature.)

The DNA origami approach has also been translated to the assembly of 3D objects. One approach to this problem was to stitch together flat origami sheets into hollow solids. Using this method, Gothelf and Kjems designed a lidded DNA box that can be unlocked upon addition of DNA 'key' sequences (**Figure 1.12 a**).[213] In another approach, the Shih group carefully adjusted the positions of crossover points between duplexes, demonstrating control of the curvature between helices, and thus taking origami structures into the third dimension by creating honeycomb helix lattices (**Figure 1.12 b**).[214] In a precursor to this method, six-helix bundle nanotubes were formed by scaffolded origami and used as a liquid crystal medium to align membrane proteins for NMR structure determination.[215] Furthermore, this approach led to the development of multilayered 3D DNA origami structures where helices adopt hexagonal, square or hybrid packing, an opportunity for more versatility in DNA assembly.[216,217]

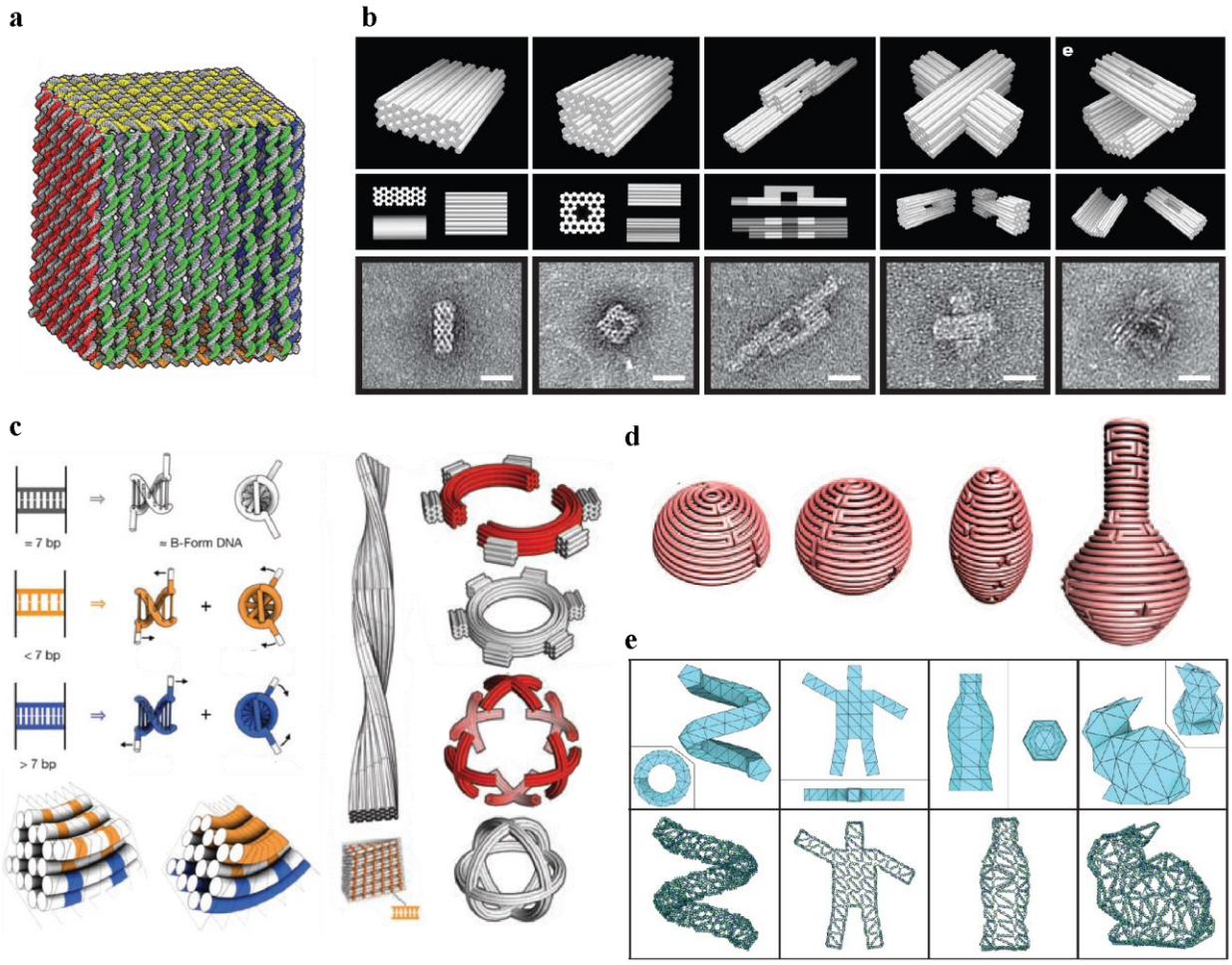


Figure 1.12| 3D DNA origami. a) Stitching flat origami sheets into a hollow box. **b)** 3D origami structures formed from honeycomb helix lattices and their transmission electron microscopy (TEM) characterization. **c)** Engineering origami structures with curvature and twist using base pair additions and deletions between crossover points in honeycomb helix lattices. **d)** Engineering curved origami structures by aligning helices in concentric rings in 3D. **e)** Using wireframe DNA origami meshes to make complex objects in 2D and 3D. (Adapted with permission from: **(a)** Ref. 213, 2009 Springer Nature; **(b)** Ref. 214, 2009 Springer Nature; **(c)** Ref. 218, 2009 AAAS; **(d)** Ref. 219, 2011 AAAS; **(e)** Ref. 220, 2015 Springer Nature.)

A further development in 3D DNA origami methodologies was the addition of curvature into the assembled structures. Targeted base pair deletions and additions between crossover points introduce strain, inducing prescribed curving and twisting of the structure (**Figure 1.12 c**).[218] A similar approach was recently used to form helical nanotubes with controlled diameter and chirality. [221] Furthermore, Yan and coworkers have assembled highly complex curved DNA origami objects such as a 'nanoflask' by orienting helices into concentric rings rather than packed

lattices (**Figure 1.12 d**).[219] Complex wireframe 3D structures have been built by folding the scaffold strand into bundles describing a polyhedral frame.[222] More recently, origami tripod structures were used to make the largest 3D polyhedra yet, yielding structures with 100 nm long edges and up to 60 megadaltons in molecular weight.[223] Additionally, using a wireframe rather than dense lattice building approach, a variety of complex 2D and 3D structures have been obtained by optimising the routing of the scaffold strand around the framework and introducing different junction types (**Figure 1.12 e**).[220,224]

DNA origami is undoubtedly a powerful method to build complex and varied structures on the nanoscale. However, it also has notable drawbacks.[225,226] Chief among them are complex structure and sequence design, the necessity to generate hundreds of unique strands, the large sequence pool this represents and the associated high start-up costs. Furthermore, error rates increase as strand uniqueness becomes more difficult to achieve, annealing times to prepare structures can be as long as a week and yields for complex structures are generally low. Efforts have been made to address these limitations. Sequence design automation can make origami more accessible,[227] folding pathways have been investigated to provide design tools that minimize misfolded products,[228] minimizing the sequence space by implementing reusable sequences is under investigation, and various adaptations of annealing and purification protocols[229] have been developed to simplify product preparation and isolation.[230,231]

In a unique tile-based approach that is similar to DNA origami in terms of DNA density as well as spatial resolution and addressability, Yin *et al.* proposed to use a single-stranded tile (SST), a DNA motif containing four modular domains with precise complementarity relationships. SSTs were initially used to produce nanotubes with monodisperse circumference of 4 to 20 DNA helices.[232] Aldaye and Silver adapted SSTs to make DNA ribbons and nanotubes and use them as the basis of artificial DNA/protein extracellular matrices with tunable properties to promote cellular adhesion, growth and modulate processes such as signal transduction.[233] More importantly, the SST approach was later proven to be a powerful method to assemble complex two-dimensional shapes from hundreds of distinct tiles selected from a master set that defines a molecular canvas (**Figure 1.13 a**).[234] With the tiles redubbed as "DNA bricks", the SST strategy was also adapted to build in 3D by introducing a 90° dihedral angle between tiles upon hybridization (**Figure 1.13 b**).[235] The 3D structures were shown to be reconfigurable with toehold-mediated strand displacement reactions (a single-stranded overhang, or toehold, appended to a

duplex can be used to displace a strand by pairing with an invading strand that is fully complementary to the strand that is removed).[236] Furthermore, brick assembly was used to produce microscale DNA crystals with prescribed depths and 3D nanoscale features such as cavities, a platform for various potential applications, such as well defined hosts to facilitate cryo-electron microscopy imaging of proteins.[237]

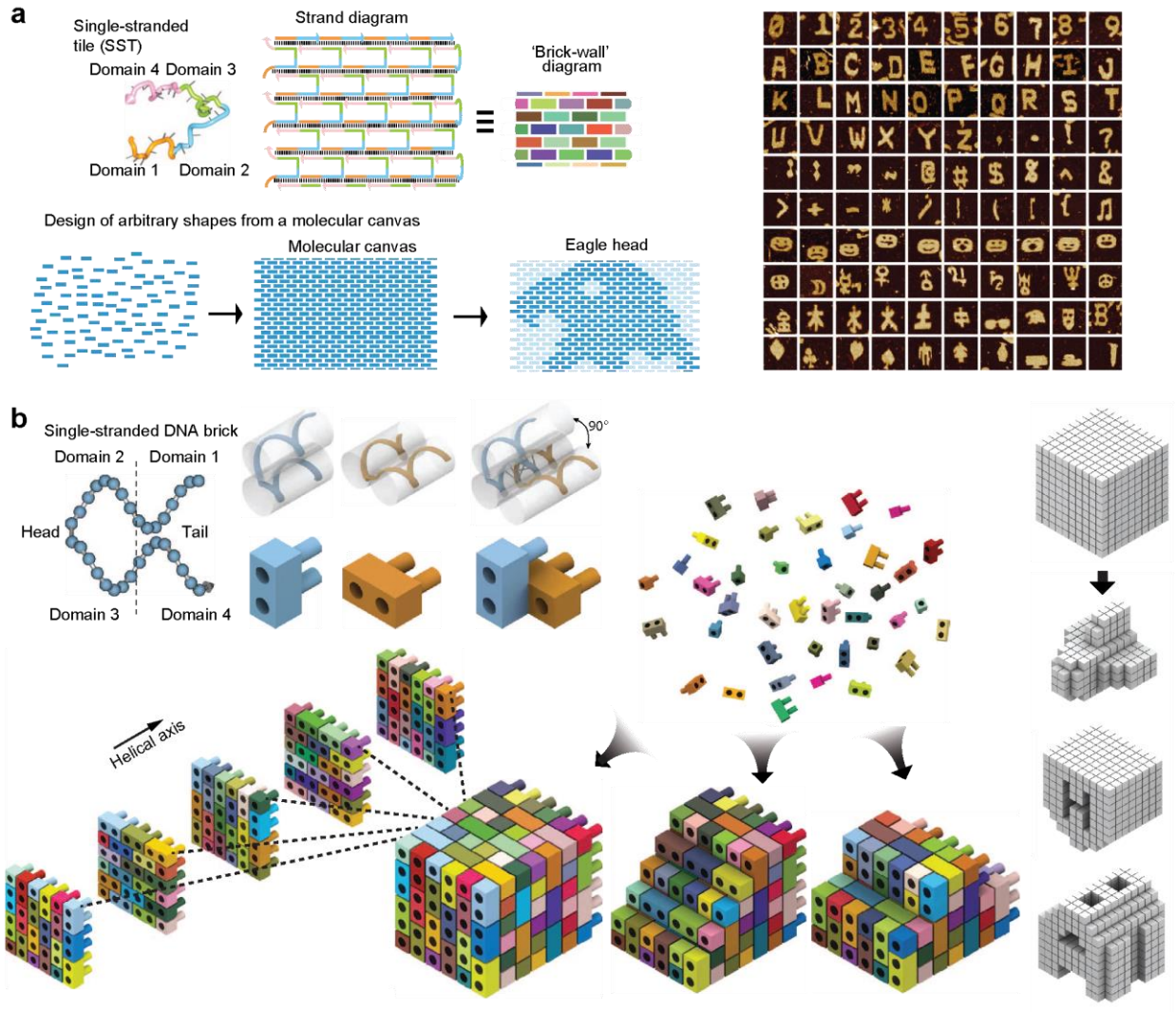


Figure 1.13| Single stranded tile (SST) and brick assembly. a) Design of arbitrary 2D shapes is carried out using a molecular canvas composed of SSTs. **b)** SST connectivity is altered to access complex structures in 3D via "brick assembly". (Adapted with permission from: **(a)** Ref. 234, 2012 Springer Nature; **(b)** Ref. 235, 2012 AAAS.)

A few important strategies that do not rely on Watson-Crick base pairing, long single stranded DNA scaffolds or short staple strands for programmable nanoscale assembly of oligonucleotides have been recently reported. The ssRNA origami method described by Rothemund *et al.* demonstrates co-transcriptional folding of a single RNA strand into prescribed nanostructures.[238] The nascent strand is designed to fold as it is synthesized using hairpins and RNA-based association motifs. Mao *et al.* have made an adaptation of the concept by expanding the types of naturally occurring RNA assembly motifs used within nanostructure design and avoiding unstable short seams.[239] Nanostructures produced following this methodology are well suited for *in vivo* production suggesting that large scales at a low cost are attainable. On the other hand, single stranded origami developed by Hao Yan and Peng Yin depends on standard base-pairing, but foregoes staple strands by rerouting single-stranded DNA and RNA sequences ranging from ~ 1000 to 10,000 nt in length into arbitrary shapes, forming the largest unimolecular synthetic nucleic acid structures reported to date.[240]

Recent developments in DNA origami and brick assembly have pushed the scale of the structures achievable by these methods to a remarkable extent. A redesign of the single-stranded DNA bricks, modifying domain length from 8 to 12 nucleotides enabled the formation of DNA objects up to 1 gigadalton in mmmolecular weight from up to 10,000 unique component strands.[241] Furthermore, an approach mimicking natural assembly principles of large biological structures, such as viral capsids, was used to build polyhedral gigadalton-scale structures on the order of 400 nm in diameter, by designing individual DNA origami modules and the precise interactions that bring about their hierarchical assembly.[242] The modules assemble via interactions between shape-programmable features on their surfaces that do not rely on traditional DNA base-pairing. Another advance in "scale-up" introduces a methodology for biotechnological mass production of DNA strands of virtually any length.[243] In this method, a single-stranded DNA precursor composed of the sequences of interest interspaced with self-excising DNAzyme cassettes is produced inside a bacteriophage, and cultured in bacteria for large scale production. This framework is shown to yield all necessary DNA strands for production of macroscopic amounts of DNA origami objects, with a projected cost three orders of magnitude lower than that of DNA strands produced using standard solid-phase synthesis. With the scale-up and mass production advances listed above, the field of DNA nanotechnology is poised for exponential growth and expanded applicability in various areas.

1.5.3 Supramolecular DNA assembly

As described above, DNA nanotechnology has exploited the highly programmable nature of Watson-Crick hybridization and well established structural parameters of the DNA duplex to build a panoply of structures of different morphologies and at a range of scales, from small cages to micro-scale 2D arrays and gigadalton polyhedra. However, there is a wider world of supramolecular interactions available in chemistry that is ripe for exploration to expand the structural diversity and functionality of DNA nanostructures. The field of supramolecular DNA assembly has focused on incorporating heterologous components and interactions orthogonal to Watson-Crick base-pairing into the DNA nanotechnology toolkit.[244] It has combined DNA hybridization with modifications such as insertion of synthetic organic moieties, inorganic components and polymeric additions to achieve new levels of structural and functional control over DNA assemblies, and to expand the possibilities of hierarchical organization.

Tile based assembly and DNA origami described previously produce DNA-dense structures which rely on strand crossovers and branching to take the linear DNA duplex into 2- and 3D. In contrast, introducing organic vertices into DNA strands via phosphoramidite chemistry[245,246] to modify their direction and preorganize the geometry of DNA building blocks enables the formation of discrete DNA-minimal wireframe polygonal and prismatic structures (**Figure 1.14 a**)[247-249] as well as extended nanotubes.[250] Such porous structures are well adapted for dynamic behaviour,[250] and can be used to host size-specific guests such as gold nanoparticles (AuNPs).[251,252] Furthermore, small organic insertions have been shown to stabilize DNA hybridization, increase the melting cooperativity of the structure, direct the distribution of assembly products and stabilize nanostructures from degradation in biological media.[170,253,254] Synthetic additions have also been used to induce curvature in higher order assembly, as demonstrated by Endo *et al*. who incorporated DNA-porphyrin connectors to drive formation of nanotubes from 2D tile arrays (**Figure 1.14 b**).[255] Synthetic modification of DNA tiles with a combination of thiol/amide, or alkyne/azide reactive groups is another strategy to induce curvature,[256] and can also create nanotubes stabilized by covalent crosslinking.[257,258] The incorporation of light-responsive azobenzene units within duplexes linking origami tiles into long-range assemblies was used as a photo-switch to modulate their rearrangement into prescribed 2D patterns.[259]

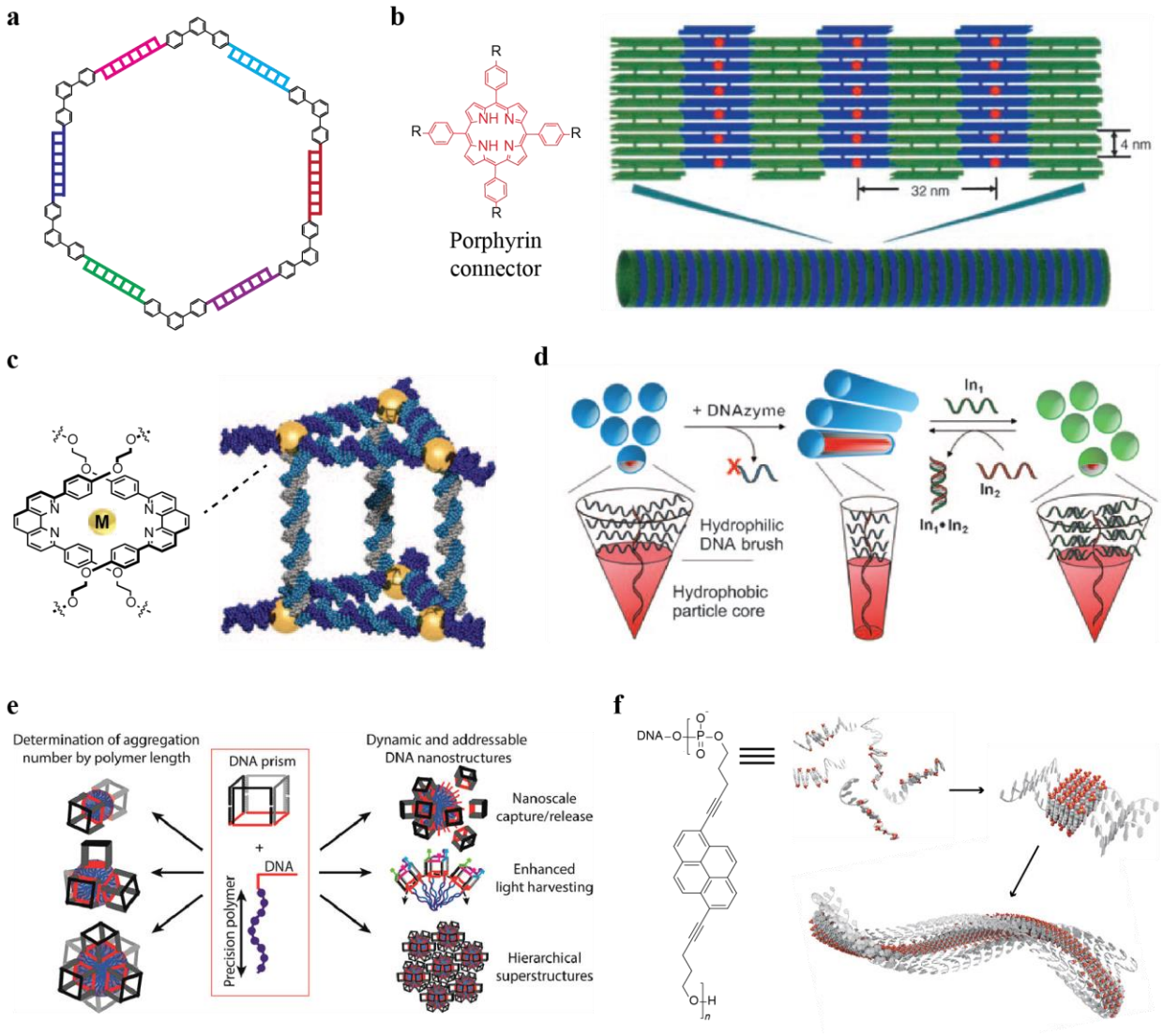


Figure 1.14| Supramolecular DNA assembly. a) Directed assembly of a DNA hexagon with organic corner units. **b)** Porphyrin connectors induce curvature driving nanotube formation from a 2D tile array. **c)** DNA cage with transition metal complexes at the vertices. **d)** Stimuli-responsive polymer-DNA assemblies. **e)** DNA cages decorated with sequence-defined DNA-hydrophobic polymer conjugates undergo assembly into quantized aggregates. **f)** Supramolecular polymerization of DNA-pyrene conjugates into helical ribbons. (Adapted with permission from: **(a)** Ref. 248, 2006 Wiley; **(b)** Ref. 255, 2005 Wiley; **(c)** Ref. 260, 2009 Springer Nature; **(d)** Ref. 261, 2010 Wiley; **(e)** Ref. 262, 2014 ACS; **(f)** Ref. 263, 2015 Wiley.)

As alluded to in section 1.4.2, the integration of transition metals into DNA endows the resulting molecules with new properties. An early example of transition metal binding complex incorporation into DNA strands used DNA hybridization to direct the formation of linear and

branched metal-organic oligomers to develop macromolecular structures that could form the basis of molecular circuits.[264] Concurrently, the field saw the development of a variety of branched DNA-metal ligand modules for hierarchical assembly, possessing geometric features prescribed by the position and valence of the built-in metal-binding vertices. The first cyclic metal-DNA nanostructures were formed based on ruthenium(II)-binding tris(bipyridine) components,[265] whereas triangular cyclic structures were assembled with bis(terpyridine)iron(II) complexes at the vertices.[266] The MacLaughlin group described branched junctions with four[267] and six DNA arms,[268] incorporating nickel(II) tetrahedral complexes and ruthenium(II) octahedral complexes respectively, that formed periodic metal-DNA assemblies. The Sleiman group reported a more dynamic multi-metallic triangular unit[269] which formed the basis for assembly of metal-DNA cages with precise 3D positioning of the metal centers at the vertices (**Figure 1.14 c**),[260] and later metal-DNA wireframe nanotubes.[270] The diphenyl-phenanthroline and terpyridine containing DNA-templated ligand environments reported by the Sleiman group were highly selective for the incorporation of specific transition metal ions.[271] Altogether, metal-DNA junctions represent orthogonal components capable of directing DNA assembly into new motifs, and can lend significant stabilizing effects to the resulting nanostructures.[142,260,271] DNA nanostructures containing metal binding complexes have potential applications in catalysis, sensing, electronic nanodevices and artificial photosynthesis.[272,273]

Synthetic insertion of hydrophobic components into DNA strands is another means of accessing orthogonal assembly modes in DNA nanomaterials. Molecular building blocks composed of DNA modified with hydrophobic small molecules, dendritic structures and polymers present amphiphilic characteristics akin to lipid molecules and amphiphilic block-copolymers. They are thus capable of assembly by microphase separation. This orthogonal assembly mode complements the DNA nanotechnology toolbox by introducing a powerful new strategy for long-range and hierarchical assembly. A discussion of DNA nanostructures modified with small lipid-like molecules and their interactions with lipid bilayers can be found in section 1.5.6. After initially investigating DNA structures modified with oligoethylene dendrons that undergo long-range assembly into fibers under certain solvent conditions,[274] the Sleiman group developed dendritic alkyl-DNA amphiphiles that formed micellar structures in aqueous conditions and showed unique assembly behaviour when positioned on a DNA-minimal cage scaffold.[275] Furthermore, DNA-polymer conjugates have given rise to DNA micelles with switchable character.[261,276] The

Gianneschi group has reported the stimuli-responsive morphological conversion of DNA-brush polymer spherical micelles into rod structures and back to micelles by modulating the size of the hydrophilic block of the DNA-polymer conjugate with DNAzyme cleavage (shrinks the hydrophilic block) and subsequent hybridization with complementary DNA strands (re-expansion of the hydrophilic block) (**Figure 1.14 d**).[261] The Sleiman group has also contributed to the field of DNA-polymer conjugates by developing sequence-defined polymers synthesized using standard phosphoramidite chemistry and automated solid-phase synthesis.[25] The precise control of monomer sequence and length of the polymeric block appended to DNA is instrumental in defining the self-assembly properties of the polymers. Sequence-controlled polymers synthesized according to this methodology were used to form micelles for drug delivery,[277] poly-fluorinated "DNA-Teflon" micelles,[278] nanoreactors,[279] and monodisperse DNA-imprinted polymer nanoparticles.[280] The decoration of DNA cages with sequence-controlled DNA-polymer conjugates produced quantized hierarchical assemblies with dynamic behaviour (**Figure 1.14 e**)[262] and highly tunable structural features dependent on the length and composition of the polymeric block.[281]

In the case of DNA building blocks modified with planar aromatic components, the contribution of π-stacking interactions between them is an additional driving force behind the formation of unique architectures with wide-ranging potential functionalities. For example, the stacking properties of chromophores have been used by Wang *et al.* to create hybrid junctions from foldable thermophilic DNA polymers.[282] The Häner group has synthesized DNA-oligo-pyrene chimera which underwent supramolecular polymerization into helical ribbons decorated with DNA strands as a result of stacking of the pyrene units (**Figure 1.14 f**).[263] In a demonstration of hierarchical organization, further hybridization of the DNA segments to complementary oligonucleotides triggered the assembly of individual ribbons into extensive networks mediated by the stacking interactions between the terminal bases of the DNA duplexes.[283] The same group also used branched DNA building blocks with anthracene modifications to create self-assembled networks that were converted to covalently linked 2D polymers following photochemical dimerization of the anthracene units.[284] DNA duplexes conjugated to phenanthrene oligomers formed self-assembled vesicles with light harvesting properties, highlighting a potential application of nanostructures with built-in chromophores.[285] Interestingly, the incorporation of a π-stacking cyanine dye in DNA-amphiphiles that normally form spherical micelles disrupted

micelle formation in favour of long fibers with stacked dyes at their core.[286] The variety of higher-order morphologies accessible by combining DNA with hydrophobic moieties, π-stacking and polymers represents a rich area of development for functional nanomaterials.

Alternate naturally-occurring DNA assembly motifs (guanine quadruplex, i-motif, triple helix, etc.) may seem like obvious candidates to expand the range of interactions used in DNA nanotechnology beyond standard base-paring, however they remain underrepresented in the field. Their unique geometries and properties can introduce otherwise inaccessible structural options and new functionalities to DNA architectures. Guanine-rich sequences have been shown to assemble into G-wires, or extended 1D stacks of G-quadruplexes.[287] Well-defined arrays composed of G-quadruplexes connected by duplex regions have combined Watson-Crick hybridization with non-B-DNA assembly and could be controlled by adjusting the concentration of monovalent cations in the surrounding environment.[288] Similarly to G-wires, long range assembly of cytosine-rich i-motif-forming sequences has been reported.[289] I-motif pillars decorated with DNA duplex branches can be used to organize materials such as gold nanoparticles along the pillar axis.[290] Both i-motif and triple helical elements have been incorporated in stimuli-responsive DNA-based hydrogels.[291,292] The pH dependent stability of DNA triplexes has been harnessed in the design of DX tile-based DNA nanotubes that disassemble in response to pH changes and can act as sensors.[293] Although there is growing interest in incorporating alternative DNA assembly motifs into DNA nanotechnology[294,295] there is still much work to be done, especially in the context of organized long-range assembly.

The addition of orthogonal interactions into the DNA assembly language through synthetic modification of oligonucleotides and alternate assembly modes has become an invaluable tool to increase the structural complexity and tunability available to DNA nanotechnology. It has enabled the development of structures with morphologies and functionalities that would be inaccessible with Watson-Crick base-pairing alone. The interplay of various supramolecular interactions that take over the assembly process once new components are introduced is complex and requires detailed study with each new system. However, the possibilities of incorporating designer properties into the resulting nanomaterials for advanced applications invite further investigation.

1.5.4 Long-range assembly via blunt-end stacking interactions

The hybridization of DNA strands is of central importance in life, biotechnology and nanotechnology. As described above, the programmable assembly of DNA nanostructures is achieved through the hybridization of carefully designed DNA sequences into prescribed structures, while long-range assembly of these building blocks into nanofibers and arrays is traditionally achieved by the intermediary of complementary sticky ends or staple strands.[176,209,237]

However, DNA has another valuable property that has so far been under-explored for hierarchical long-range assembly of DNA nanostructures. DNA duplexes with blunt ends (duplexes that do not have complementary single-stranded overhangs) are capable of spontaneous end-to-end aggregation via weak interactions that can lead to the formation of liquid crystal phases at high concentrations (**Figure 1.15 a**).[296,297] These adhesive forces are the result of stacking between the DNA base pairs located at the blunt duplex ends. A combination of π-stacking interactions, that depend on surface area of contact between bases and stacking geometry, electrostatic forces and hydrophobic effect make contributions to this weak interaction.[298]

Within the DNA duplex, it has been reported that base pair stacking, rather than hydrogen bond-mediated hybridization of complementary bases, makes the dominant contribution to double helix stability.[299] By studying long DNA duplexes with single nicks and gaps, Yakovchuk *et al.* have been able to determine the temperature and salt dependence of stacking free energy. In a wide range of temperature and salt conditions, they have found that, while A-T pairing via hydrogen-bonding almost always makes a destabilizing contribution, and G-C pairing contributes almost no stabilization, base stacking is always stabilizing. In fact, temperature and salt dependence of the stacking term fully determines the temperature and salt dependence of the duplex stability parameter.[299] Protozanova *et al.* have determined sequence dependent stacking free energy for different base pairs based on stacked-unstacked equilibria determined for nicked duplexes.[300] The stability of a DNA duplex is widely known to be sequence dependent and to scale with GC content.

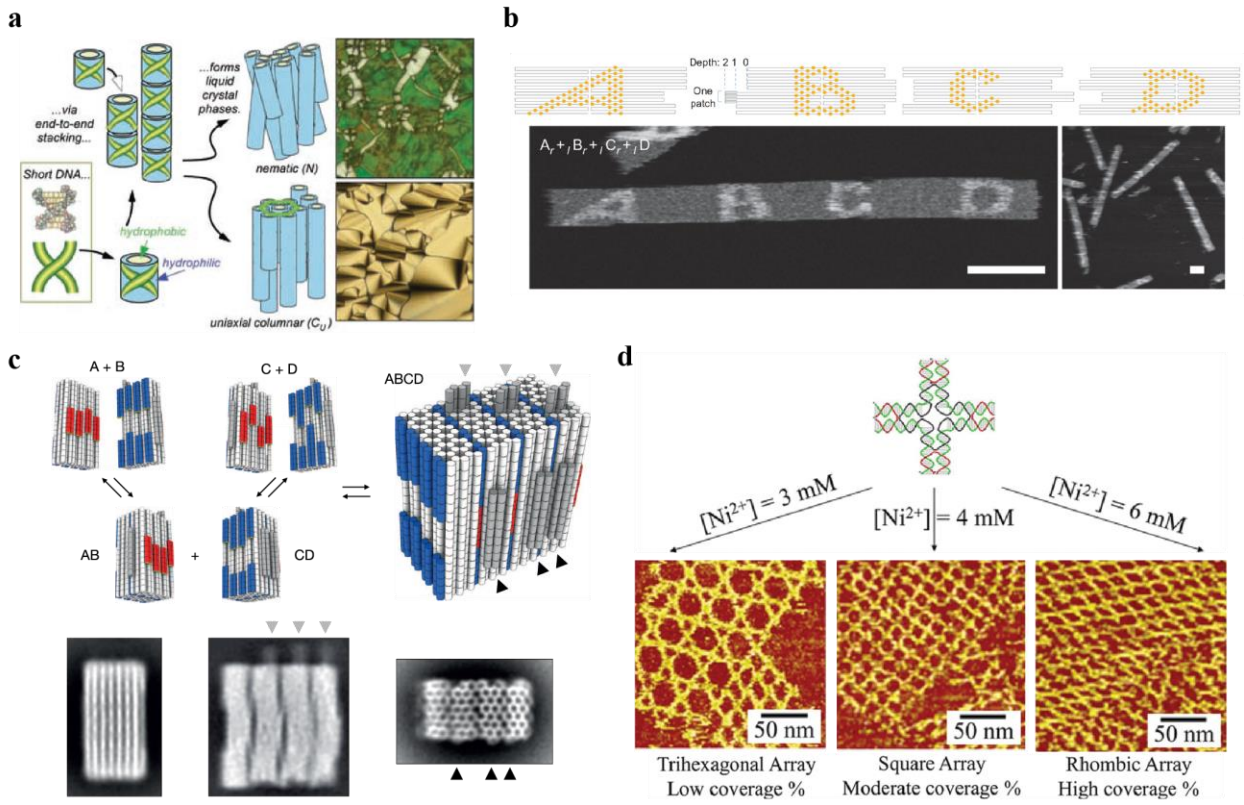


Figure 1.15| Long-range DNA assembly using blunt-end interactions. a) Aggregation of short DNA duplexes into nematic and columnar liquid crystal phases. **b)** DNA origami tiles with shape complementary and blunt-ended edges form a four-tile ribbon. **c)** 3D origami modules with blunt-ended shape-complementary interfaces assemble specifically into dimers that then combine to form a tetramer. Negative-stain TEM characterization is also shown. **d)** Cross-shaped tiles with blunt ends form lattices of varying morphology on mica depending on ambient cation concentrations. (Adapted with permission from: **(a)** Ref. 296, 2007 AAAS; **(b)** Ref. 210, 2011 Nature Springer; **(c)** Ref. 301, 2015 AAAS; **(d)** Ref. 302, 2017 Wiley.)

In long-range DNA assembly, the use of weakly adhesive DNA base-stacking interactions could promote dynamic rearrangement of building blocks toward regular patterning over larger surfaces, thus providing an error correction mechanism during the assembly process. In addition, this interaction can possibly give rise to morphologies differing from those obtained with traditional DNA base-pairing. Only a limited number of studies have taken advantage of this assembly mode so far. For example, one dimensional arrays of three-helix bundle DNA motifs have been produced from building blocks presenting blunt duplex ends at their edges.[303] The Sugiyama group combined the use of shape complementarity, blunt contacts and sticky-end cohesion to assemble origami 'jigsaw puzzle' pieces into programmable 1D ribbons[304] and 2D

arrays.[207] Soon afterward, Rothemund *et al.* eliminated the use of sticky ends, fully relying on shape complementarity and blunt end adhesion properties to build DNA origami arrays (**Figure 1.15 b**).[210] Furthermore, tuning of surface adhesion properties via cation addition gave rise to extensive periodic lattices on mica.[211] The Dietz group built complex, dynamic structures from 3D origami components by designing interfaces between components based on shape complementarity and blunt contacts as well (**Figure 1.15 c**).[301] More recently, the Mao group has combined blunt-end stacking and DNA-surface interactions mediated by cations to control long-range assembly of small branched and DX tiles (**Figure 1.15 d**), also finding interesting steric effects affecting assembly patterns with the addition of hairpin moieties at tile interfaces.[302]

1.5.5 Applications of extended DNA assemblies

Throughout its development, the DNA nanotechnology field has capitalized on the highly predictable nature of DNA base-pairing, complemented by a wealth of interactions employed in supramolecular DNA assembly, to enable construction of complex nanostructures and extended assemblies with nanoscale features. The ability to use DNA nanostructures as scaffolds to spatially organize components such as inorganic nanoparticles and proteins is one of the great strengths of DNA nanotechnology.[305,306] Both periodic and anisotropic positioning of heterogeneous components is possible given the addressability of DNA nanostructures, which is an important feature for increasing the complexity of hierarchical self-assembled materials. The vast applicability of such materials for fundamental studies of interactions between the scaffolded molecules, for the manipulation of biological or chemical cascades, the development of optical or electronic nanodevices, as well as for patterning on the nanoscale and lithographic applications justifies the high volume of research in this area.

Tile-based DNA assembly can give rise to periodically ordered arrays as discussed above, but more complex patterns are also attainable, as in the case of algorithmic assembly of Sierpinski triangles[307] using Wang tiles. Wang tiles have four different sides that can only assemble with their neighbours by juxtaposing “like” sides. In DNA Wang tiles, each side carries a different DNA sequence. Such tiles can form aperiodic lattices from a nucleating template, however off-template growth and tile misplacement represent common errors encountered in the growth process. Reducing the error rate should be possible with recently developed tools that improve robustness

in such systems.[308] DNA tiles can also potentially be used to generate self-replicating DNA materials.[309,310] Self-replication is an exciting avenue in materials fabrication as it could be used for autonomous amplification of components and their ordered assemblies, thus potentially enabling simplified methodologies for large scale production of functional materials. Furthermore, advanced algorithmic assembly has been achieved with DNA origami tiles. The Qian group has harnessed programmable disorder in random DNA origami tiling to make loop-, maze, and tree-patterned 2D tile arrays.[311] Building on the set of local assembly rules established in this study, the group went on to report hierarchical multistage assembly of origami tiles which have a fractal organization based on the interactions of different numbers and patterns of short sticky ends along different tile edges (**Figure 1.16 a**).[312] The resulting arrays sized up to 0.5 μm^2 are designed to render arbitrary patterns and images using almost 9000 pixels. An online software tool is available to convert user-defined patterns into DNA sequences, making the method more accessible and facilitating the assembly of complex 2D arrays which, in the future, could be functionalized to build circuits and devices.

Inorganic nanostructures such as metal or semiconductor nanoparticles and nanowires are an important class of materials for optical, electronic, catalytic and sensing applications.[313] The precise geometric arrangement of such components is an important requirement to harness their optical and electronic properties for functional circuits and plasmonics.[314,315] DNA nanotechnology is a powerful method to scaffold the assembly of inorganic nanostructures into complex architectures that take full advantage of this unique combination of bio- and inorganic materials.[168] Tile-based assembly can be used to pattern gold nanoparticles (AuNPs)[316,317] and quantum dots (QDs)[318] into extended, periodic 2D lattices with specific spacing between particles and the possiblity of incorporating particles of different sizes (**Figure 1.16 b**),[319] representing advances toward the development of nanoelectronic circuits. Positioning AuNPs into helical arrangements, sometimes with controllable chirality, has been facilitated by using extended nanotube and helix bundle scaffolds (**Figure 1.16 c)**,[320-322] as well as by stacking gold nanorods between planar DNA origami tiles bearing capture strands in specified orientations.[323] The precise organization afforded by DNA origami architecture has been used to build a nanoantenna, where the fluorescent response of a dye molecule positioned in a hotspot between two AuNPs was highly enhanced as a result of plasmonic coupling.[324] Furthermore, rigid DNA origami rods have served as scaffolds for hierarchical clusters composed of AuNPs, QDs and dye molecules.[325] Tunable

cluster stoichiometry and size, as well as their aggregation properties, allow detailed investigation of the collective optical properties emerging from the interaction of these components and invite the application of these nanostructures as energy funnels or porous scaffolds for catalysis. DNA base-pairing has mediated the assembly of 3D plasmonic crystals[326,327] and hollow DNA origami structures have served as moulds to cast inorganic particles of user-specified shape.[328] It should be noted that DNA nanostructures have found varied applications in the creation of other composite materials. DNA scaffolds have served as templates for polymerization[329,330] or for the creation of complex DNA-silica hybrid materials,[331] as supports to route individual polymer chains along pre-designed paths,[332] and for orienting carbon nanotubes into specified patterns.[333] The versatile collection of structures and material combinations displaying precise geometry and patterning of inorganic and polymer components described above, in addition to tools such as transfer of programmed patterns of DNA strands from a nanostructured template to nanoparticles,[334] promise much future advances in the development of nanoelectronic circuits, light harvesting platforms, sensors and other nanodevices.

In terms of the development of advanced biomaterials and biophysical tools, scaffolding protein assembly is another important direction for DNA nanotechnology. DNA-mediated nanostructures which precisely position proteins can become valuable tools to study protein-protein interactions, organize enzymatic cascades, and develop biomimetic materials.[168,335] One of the initial motivations behind the development of DNA nanotechnology was related to protein organization. DNA crystals, that were hypothetical at the time, were imagined to trap proteins inside a rigid framework and thus facilitate determination of their structure by X-ray crystallography.[173] Since then, the study of protein structure has indeed been enabled by a DNA nanotechnology tool: a liquid crystalline medium composed of DNA nanotubes was used to align membrane proteins for investigation using NMR spectroscopy.[215] Similarly to early scaffolding of metal nanoparticles, 2D arrays were some of the first ordered structures formed by interfacing proteins with DNA nanostructures using tile-based assembly,[182] even enabling imaging of a protein network by cryo-electron microscopy (**Figure 1.16 d**).[336] As site-specific protein-DNA conjugation remains challenging to this day,[335] protein immobilization on DNA lattices mostly relies on non-covalent interactions, such as biotin-streptavidin binding,[182] aptamer-directed capture,[337] protein recognition of Holliday junctions integrated into DNA tile design,[338] or adaptation of the His-tag method for protein purifcation.[336,339] DNA nanostructures have enabled

the construction of multi-enzyme complexes to study and tune the reactivity of enzymatic cascades.[340,341] Such systems facilitate inter-enzyme substrate diffusion and regulate interactions between reactive units.[342] An extended RNA scaffold directing the spatial organization of a hydrogen-producing enzymatic pathway has even been expressed *in vivo* (**Figure 1.16 e**) inside *e.coli* bacteria.[203] Furthermore, DNA nanostructures have been used to nucleate the assembly of amyloid fibrils which were later organized on a DNA origami support,[343] as well as to examine the opposing forces generated by motor proteins with different motion properties.[344] The Dietz group has described a novel approach for folding double stranded DNA using protein-DNA interactions. In this method, transcription activator-like (TAL) effector protein dimers bind to distant regions of a double-stranded DNA scaffold, thus bringing them in close proximity and stabilizing the folding of the DNA scaffold in predesigned configurations. The resulting hybrid DNA-protein nanostructures could represent a powerful new platform for protein organization.[345] Future development in DNA-mediated protein organization is expected to benefit from developments in DNA-protein conjugation chemistry.[335] Enabled by advances in synthetic biology, DNA or RNA-mediated protein scaffolding could be an important platform for biophysical studies and the development of biomimetic materials inside living organisms.[168]

Biophysical and biomedical applications of DNA nanotechnology are not limited to the organization of proteins. Super-resolution microscopy acquired a powerful new tool with the introduction of DNA points accumulation for imaging nanoscale topography (DNA-PAINT) technique, which relies on the programmability and specificity of DNA binding for transient labelling of molecules of interest.[346] In combination with DNA origami-based fluorescent barcodes,[347] DNA-PAINT can be used for multiplexed imaging of different targets.[348] Quantitative DNA-PAINT (qPAINT) is a more recent variation on the technique, which is capable of counting the number of imaged targets.[349] The development of discrete DNA nanostructures as drug delivery vehicles is a highly active area of research. Many considerations such as drug encapsulation, cellular uptake, specific targeting, stimuli-responsive behaviour, and stability in physiological media are under intense scrutiny by the community.[170,350] The use of DNA nanostructures in the development of materials for tissue engineering and regenerative medicine is another emerging area in DNA assembly. Supramolecular materials have long been of interest in this area of study,[351] and the power of DNA building blocks could lend new functions and levels of complexity to nanomaterials targeting this type of application. Tile-based DNA nanofibers with

structural tunability features have been used to mimic the extracellular matrix,[233] while DNA nanotubes functionalized with adhesion peptides have served as a bioactive substrate affecting stem cell differentiation.[352] In addition, interfacing DNA nanostructures with lipid bilayers is a research direction with potential applications in biophysical studies and biomaterial development. Advances in this area are discussed in section 1.5.6 below.

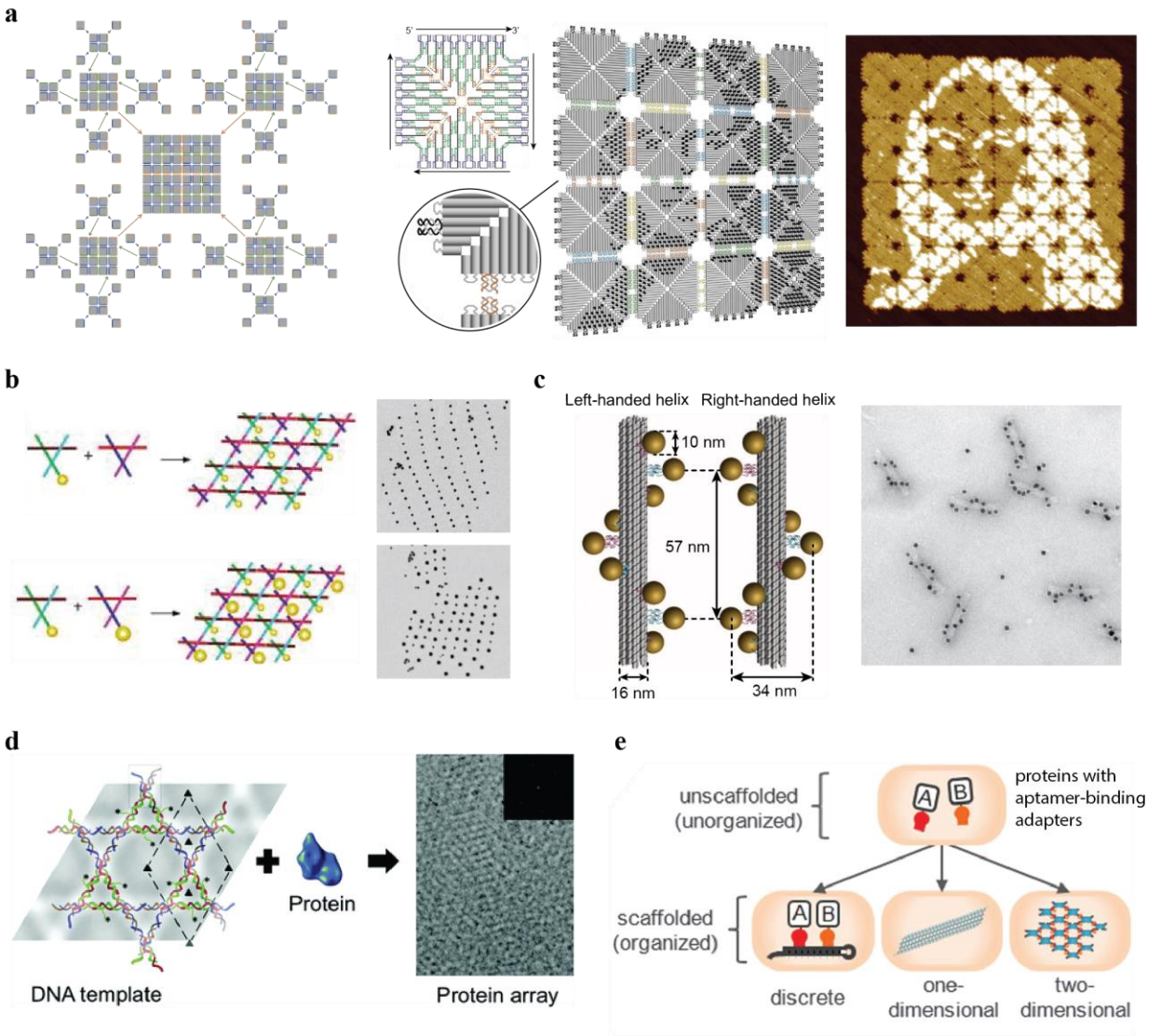


Figure 1.16| Applications of extended nucleic acid nanostructures. a) Fractal assembly of an DNA array bearing an arbitrary pattern from DNA origami tiles. **b)** 2D DNA tile lattice scaffolds periodic AuNP arrays. **c)** Chiral arrangement of AuNPs along a helix bundle. **d)** 2D protein array templated by DNA tile assembly. **e)** Polymerizable RNA tiles bearing protein-binding aptameric sequences which enable *in vivo* scaffolding of an enzymatic cascade. (Adapted with permission from: **(a)** Ref. 312, 2017 Springer Nature; **(b)** Ref. 319, 2006 ACS; **(c)** Ref. 321, Springer Nature; **(d)** Ref. 336, 2011 ACS; **(e)** Ref. 203, 2011 AAAS.)

The programmable nature of DNA self-assembly produces highly complex and addressable nanostructures built from the bottom up. With functionalization, a wide variety of applications becomes available. In addition, the dynamic nature of DNA assembly can be harnessed to create responsive structures that can act like nanorobots to undergo motion, transport cargo or perform algorithmic operations.[353-356] Furthermore, to achieve higher order organization of DNA nanostructures and additional components on surfaces, bottom-up assembly of DNA has been combined with and top-down fabrication through the use of lithographically patterned substrates.[357-359] Lithography techniques have the potential to extend the precise patterning of objects with nanoscale resolution to the micro- and millimeter length scales. Altogether, the ensemble of structures and techniques developed in DNA nanotechnology to build structures with extended organization show promise in several areas of application, such as production of nanocircuits, platforms for biophysical studies, biomimetic materials and plasmonic devices. Furthermore, with the scale-up and mass production advances outlined in section 1.5.2 above, high-scale production of DNA nanostructures becomes a realistic prospect in the near future, further expanding the applicability of DNA-based materials.

1.5.6 Interactions of DNA nanostructures with lipid membranes

As described above, biological lipid membranes compartmentalize contents, thus creating spatial organization and increasing the efficiency of reactions. Moreover, they serve as the interface for various functions such as recognition, transport and signalling, at the same time acting as a fluid framework for long-range organization, and as an anchoring support for membrane proteins. There is significant appeal in interfacing lipid bilayers with precisely designed structures built using DNA nanotechnology. By bringing these self-assembly worlds together, it is possible to probe biological phenomena with biomimetic DNA-based model systems, and to achieve higher-order, hierarchical organization of nanomaterials on a fluid, biologically relevant support.[15]

Lipoplexes, or mesoscale assemblies of lipids with DNA, are a highly documented class of complexes that have long held promise for gene therapy applications and transfection.[360,361] Cationic lipids are most common in the formation of lipoplexes due to the strong electrostatic binding of their headgroups to polyanionic nucleic acids, particularly at low salt concentrations which minimize electrostatic screening effects (**Figure 1.17 a**).[362] DNA also binds to zwitterionic

lipids, especially in the presence of divalent cations such as Mg^{2+} or Ca^{2+}, which are thought to bridge phosphate groups between neighbouring lipid molecules thus neutralizing their negative charges, and favoring electrostatic binding of the positive charges to the DNA backbone.[363,364] Interactions of unmodified DNA with lipid bilayers are highly dependent on bilayer composition and phase, with preferential binding reported for the gel and liquid ordered (L_o) phases, potentially due to a higher charge density which follows from closer lipid packing.[365,366] A more versatile approach to direct the binding of DNA molecules and nanostructures to lipid bilayers is by synthetically modifying DNA strands with hydrophobic moieties (**Figure 1.17 b**).[367] Modifications such as cholesterol,[368-370] tocopherol,[371,372] or porphyrins[373] are available as phosphoramidites allowing for simple preparation of conjugates. Hydrophobic polymers such as polypropeleneoxide (PPO) may also be converted to a phosphoramidite form for conjugation to DNA using standard solid phase synthesis methods.[374]

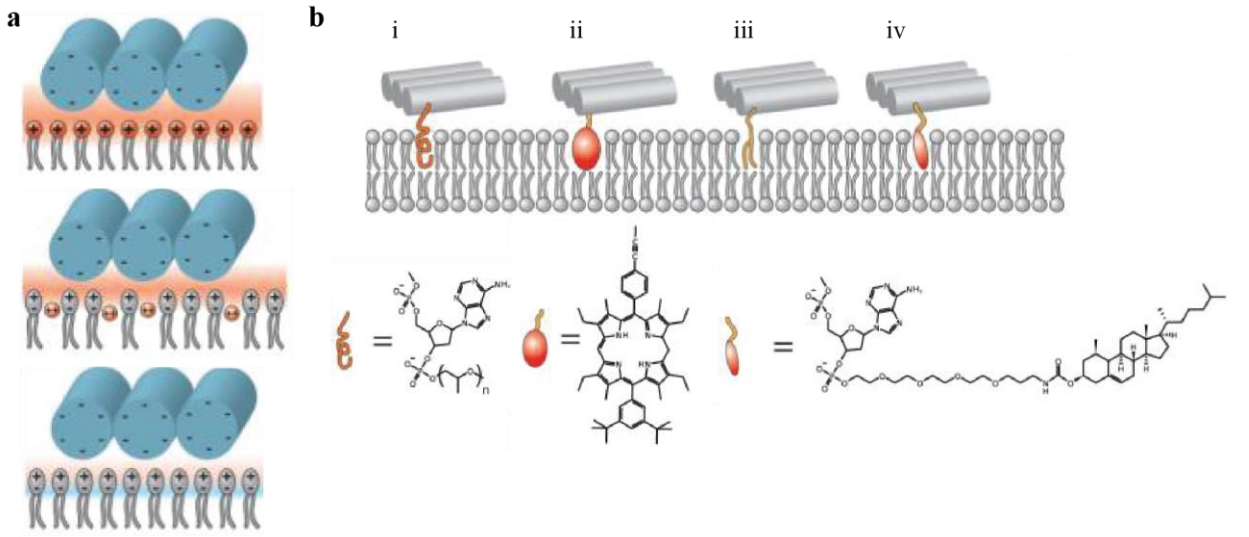


Figure 1.17| DNA-lipid interactions. a) Electrostatic binding of DNA to (from top to bottom) cationic lipids, zwitterionic lipids in the presence of divalent cations, and zwitterionic lipids in the absence of cations (weak binding). **b)** Amphiphilic DNA conjugates with **i)** a hydrophobic polymer, **ii)** a porphyrin, **iii)** alkyl chains and **iv)** cholesterol, where the hydrophobic moiety can embed into a lipid bilayer. (Adapted with permission from Ref. 15, 2014 ACS.)

Much work has looked into the binding properties of small DNA constructs such as single strands, duplexes and cages with lipidic modifications to vesicles and supported bilayers with focus on the orientation of constructs upon binding,[375,376] preferential localization to different lipid

phases,[368,371,372,377] diffusion on the bilayer upon binding,[373] and construct addressability through hybridization with additional DNA strands[369,378,379] or cleavage by enzymes.[372] Partitioning of constructs to different lipid phases can be achieved by modifying the anchoring hydrophobic moiety. For example, strands anchored with palmitoyl[372,377] and tocopherol[371,372,378] modifications have been shown to preferentially bind to liquid ordered (L_o) and liquid disordered (L_d) phases respectively. The partitioning of cholesterol modified constructs is affected by the presence of a spacer between the DNA portion and the cholesterol moiety, as well as by the number of cholesterol units, with localization either moslty on the L_d phase, mostly on the L_o phase, or spread over both depending on the anchor structure.[368,370,380]

One of the first examinations of a larger DNA nanostructure interacting with a lipid bilayer was carried out by the Schwille group, who used DNA origami nanorods modified with 8 cholesterol-triethylene glycol (TEG) anchors, all positioned on the same face of the structure.[381] This work examined the binding of the nanostructures to zwitterionic, cationic and anionic lipid vesicles, finding that the cholesterol anchors assured binding to all three, with strength of the interaction depending on cation concentration in the buffer. Furthermore, it reported switchable binding behaviour on phase-separated lipid membranes, with low Mg^{2+} conditions favoring binding to the L_d phase accompanied with crowding-dependent rotational and translational diffusion, and partitioning of the structures transitioning to the L_o phase in the presence of high Mg^{2+} concentrations. These findings were not reproducible with single stranded DNA-cholesterol conjugates, highlighting the specific properties of deposition of larger DNA nanostructures. The anchoring moiety is not the only determinant of deposition behaviour, as electrostatic interactions between the negatively charged DNA and lipid headgroups, which are highly dependent on headgroup charge, lipid packing density and counterion concentrations, take on an important role.

Following this work, the studies building on the concept of functional DNA nanostructures interacting with lipid bilayers multiplied. The Simmel group described a DNA origami “bilayer”. It formed when a rectangular tile modified with 35 cholesterol units folded upon itself to facilitate interactions between the hydrophobic moieties and protect them from the aqueous environment.[382] Interestingly, the DNA “bilayer” unfolded when exposed to a lipid bilayer and bound it, suggesting potential applications for this type of nanostructure in drug delivery applications. The Sugiyama group used a lipid bilayer support to study the reversible photoinduced assembly of hexagonal DNA origami tiles modified with cholesterol for bilayer anchoring, and azobenzene-bearing arms

for photodimerization (**Figure 1.18 a**).[383] The assembly/disassembly process was visualized by fast-scan AFM and demonstrated the potential for higher order assembly of DNA nanostructures on this fluid support. Cholesterol anchored DNA origami tiles modified with fluorophores were also used by the Yan group as membrane-exploring agents allowing for super-resolution fluorescence imaging of bilayer surface features.[384] Furthermore, tiles could be removed from the bilayer by displacing the cholesterol-modified anchoring strands, or assembled into linear oligomers with the addition of linking strands. This system represents a versatile tool in the development of membrane bound DNA nanodevices.

One of the early successes of the DNA nanotechnology field consisted in the creation of extensive, periodic arrays of DNA tiles. As discussed earlier, long-range organization of DNA origami tiles into lattices is also a focus of investigation. Combining this aim with the emerging understanding of the binding and dynamics of cholesterol-modified DNA nanostructures on lipid bilayers, the Liedl group reported lipid membrane-assisted growth of linear arrays and 2D lattices of DNA origami tiles (**Figure 1.18 b**).[385] The structures (multilayer rectangular DNA blocks and Y-shaped tiles) were modified with varying numbers of cholesterol units, deposited on DOPC (1,2-dioleoyl-*sn*-glycero-3-phosphocholine) fluid bilayers and assembled into networks via addition of complementary staple strands. This study examined the diffusion properties of the tiles, which were dependent on the size of the nanostructure as well as the number of anchoring units, finding that diffusion on the fluid bilayer support facilitated the formation of larger arrays than assembly in solution. The Sugiyama group applied an alternative strategy to achieve large organized arrays of DNA origami cross-shaped tiles (**Figure 1.18 c**).[386] They used unmodified DNA tiles with blunt-ended edges which bind electrostatically to zwitterionic DOPC in the presence of Mg^{2+} cations and form long range assemblies through an ensemble of weak blunt-end contacts. The free diffusion of tiles on the fluid support allows for dynamic processes such as array reorganization and defect filling observable using high speed AFM imaging. The same group further elaborated on this concept by describing hierarchical, lipid bilayer-assisted assembly of lattices in a two-step process.[387] The formation of an initial framework of cross-shaped tiles is followed by docking of square tiles in the framework cavities, which is mediated by electrostatic interactions facilitated by high Mg^{2+} concentrations, or can be directed with the addition of complementary strands at specific positions (**Figure 1.18 d**). In an additional study, the group examined the binding of unmodified cross-shaped tiles to phase separated bilayers and the effect of switching from Mg^{2+}

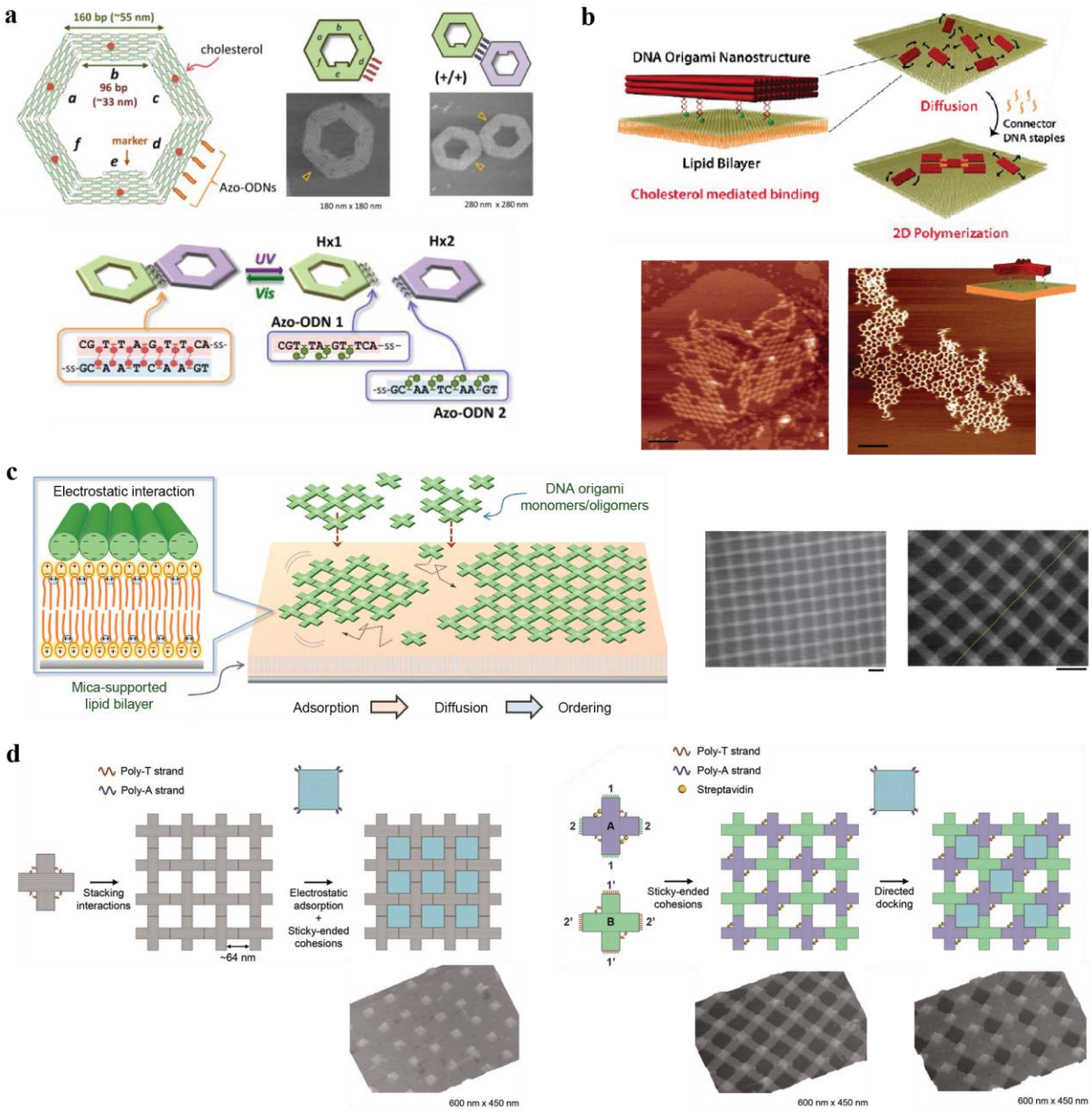


Figure 1.18| Long-range assembly of DNA nanostructures on lipid bilayers. a) Photoresponsive assembly/disassembly of hexagonal, cholesterol-modified DNA origami tiles on a lipid bilayer support. The linking duplexes are modified with azobenzene units which allow the duplex to form when in *trans* conformation and disrupt it in *cis* conformation. **b)** DNA origami lattices assembled from cholesterol-modified rectangular multilayer tiles or Y-shaped triskelion tiles with the addition connecting DNA staples. **c)** Membrane-assisted assembly of unmodified, blunt-ended DNA origami cross-tiles on a fluid bilayer support. **d)** Sequential self-assembly of tile arrays by directed docking. An initial DNA origami cross-tile lattice is formed as in **(c)** followed by the addition square tiles that are trapped in lattice cavities at high Mg^{2+} concentrations or via sticky-end cohesion. (Adapted with permission from: **(a)** Ref. 383, 2014 ACS; **(b)** Ref. 385, 2015 ACS; **(c)** Ref. 386, CC by 4.0; **(d)** Ref. 387, 2018 Wiley.)

to Na^+ containing buffer on their adsorption.[388] Dohno *et al.* have also examined the long-range organization of DX-tile constructs containing strands whose backbone was modified with aliphatic substituents capable of embedding in lipid membranes.[389] Altogether, this dynamic field of study shows promise for the development of mesoscale 2D stuctures with hierachical organization which can be used to scaffold functional moieties such as proteins or nanoparticles and respond to stimuli from the environment.

Another interesting area of development for DNA nanostructures at the interface with lipid membranes is in mimicking biological events and building model systems to study biological processes. For example, the hybridization between simple DNA-cholesterol constructs anchored to vesicles and supported lipid membranes has been used to mimic the action of SNARE proteins which facilitate vesicle fusion (**Figure 1.19 a**).[390,391] Such simplified artificial systems represent a tool to control and manipulate the composition of biological compartments for technological applications as well as to understand some of the mechanics of the complex fusion event in a simplified context. A more advanced platform composed of a lipid vesicle templated in a DNA origami ring and presenting SNARE protein fragments in a manner where both their number and orientation are controlled was used to study individual fusion events mediated by SNARE protein interactions.[392]

The development of DNA nanostructures as synthetic membrane channels is another growing area of biomimetic application. Examples of such nanostructures include a large multilayer DNA origami channel with a transmembrane stem and a barrel-shaped cap which anchors the structure to the bilayer through 26 cholesterol moieties (**Figure 1.19 b**),[393] as well as smaller helix bundle-type constructs anchored via porphyrin modifications,[394,395] or alkyl chains conjugated to the DNA backbone.[396] These artificial membrane channels have a range of properties, such as ion conductance similar to natural membrane channels,[393] voltage-gating,[395] ligand-controlled transport of cargo across the membrane,[397] the possibility to exert cytotoxic effects,[398] and enzymatic activity (mixing lipids between the inner and outer leaflets at the pore site).[399] These properties open the door for numerous applications, from sensing, to harnessing their antimicrobial and cytotoxic effects for therapeutics aims, and utilizing their ion conducting properties to drive molecular devices. Recently, DNA frameworks have also been utilized to organize protein fragments such as intrinsically disordered nucleoporins[400] and peptides derived from an oligosaccharide transporter[401] to form stable membrane-spanning pores. This new

scaffolding application of DNA nanostructures at the lipid interface represents a powerful tool for the investigation of membrane-bound biological structures and processes that are difficult to probe *in situ* in a simplified, controllable setting.

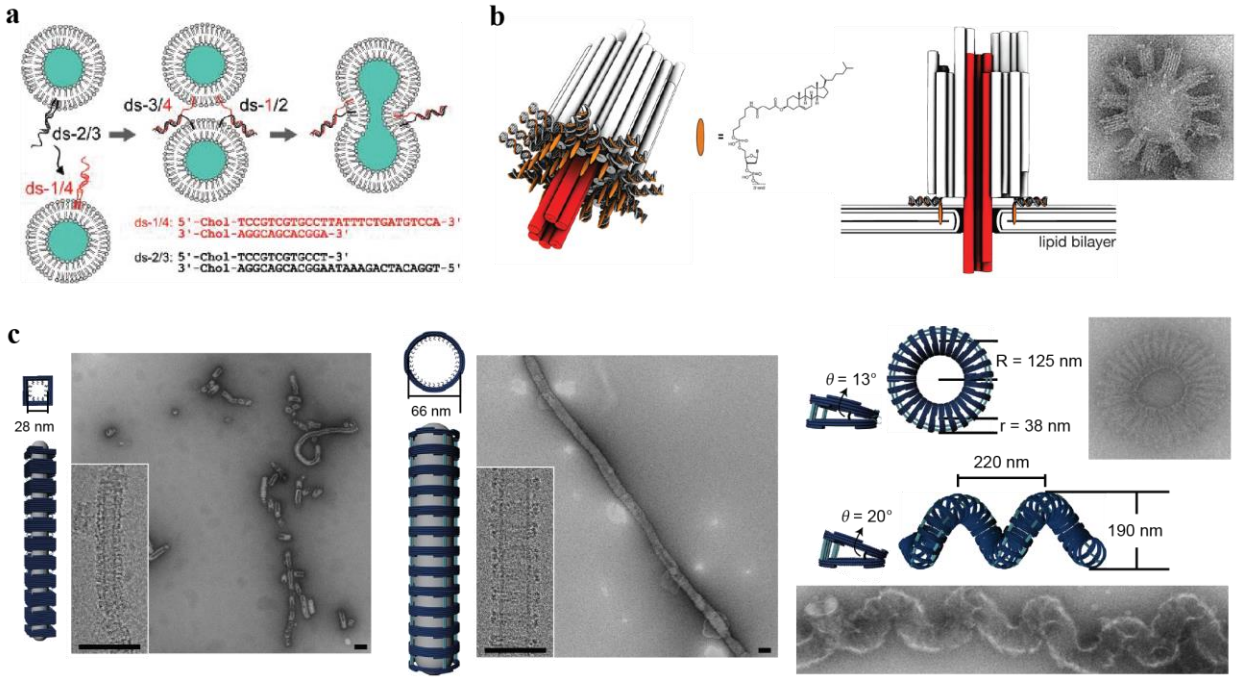


Figure 1.19| DNA nanostructures that mimic biological systems at lipid interfaces and shape lipid bilayers. a) DNA-induced lipid vesicle fusion. **b)** Artificial lipid membrane-spanning channel based on multilayer DNA origami. **c)** Designer liposomes of various shapes templated by reconfigurable DNA cages. (Adapted with permission from: **(a)** Ref. 390, 2007 ACS; **(b)** Ref. 393, 2012 AAAS; **(c)** Ref. 402, 2017 Springer Nature.)

DNA nanostructures can also be used as mimics of protein-based lipid bilayer-shaping cytoskeletons. The transformation of membrane morphology is one of the most fundamental cellular functions. Furthermore, the creation of vesicles of specific sizes and complex shapes found in cellular biology is an important goal in nanotechnology as it would provide better tools to study the interactions of different proteins with membranes of biologically relevant curvatures and compositions. The Schwille group has created DNA origami nanostructures imitating the action of coat-forming proteins which are able to sculpt and transform membranes. By designing either flat[403] or curved,[404] multiplayer DNA origami scaffolds capable of further self-assembly on vesicles, they were able to reproduce the effects of membrane-sculpting proteins of the BAR family. On the other hand, the Lin group, initially in collaboration with Rothman and Shih, has

used DNA origami rings modified with lipid handles to template the formation of lipid vesicles of specific size and variable composition.[405] By using a set of modular, reconfigurable DNA nanocages, the same group was able to form liposomes with designer geometries (linear tubes of various diameters, toroids, helices and tubular arrays) as well as remodel them following assembly by modifying the structure of the DNA scaffold (**Figure 1.19 c**).[402] The Lin group also designed spiral DNA scaffolds inspired by dynamin proteins and ESCRT machinery, two protein classes that polymerize into helical structures coating lipid tubules, to develop a strategy to create tubular liposomes from spherical vesicles.[406] This ensemble of studies demonstrates that rigid DNA nanostructures can template, bend and transform biological membranes in a controllable manner. These platforms have wide-ranging potential applications as tools for fundamental studies of protein-bilayer interactions, delivery vehicles and even as minimal biomimetic cellular machines, pushing the limits of nanotechnology into cellular biology.[407,408]

1.6 Context and scope of thesis research

Self-assembly is enabled by the information encoded in the structural features of molecules. It operates over a range of length scales and gives rise to organization and complexity in matter both in the natural world and in rationally designed systems and materials, many of which are inspired by structures found in biology. The DNA molecule exemplifies the concept of information storage both in its structure and in its biological function, which arises from the remarkable recognition properties of its base-pairing alphabet. The last few decades have seen an explosion in the fields exploring self-assembly based nanomaterials. Within this context, DNA nanotechnology has developed a rich ensemble of methodologies to build precise and highly addressable structures on the nanoscale, mostly by relying on standard Watson-Crick base-pairing, but also by expanding the assembly toolbox, introducing additional components and types of interactions into the nanostructures. Some of the current challenges in the development of DNA-based nanomaterials lie in further expanding the variety of orthogonal interactions that can be incorporated to achieve higher complexity and functionality. In addition, interfacing the nanostructures with different environments, be it the biological medium or substrates for organizing individual DNA components into larger assemblies, is crucial for many potential applications. The work presented in this thesis takes advantage of basic supramolecular chemistry

principles to incorporate alternative modes of DNA assembly into the nanotechnology toolbox and to achieve long range organization of DNA molecules through simple strategies.

In Chapter 2, basic molecular recognition principles are applied to expand the DNA base-pairing alphabet without making any synthetic modifications to the DNA strands themselves. A small molecule, cyanuric acid, which has three thymine-like faces, coaxes unmodified poly(adenine) strands into forming nanofibers. The cyanuric acid-mediated poly(adenine) assemblies are proposed to have a unique internal structure wherein adenine residues and cyanuric acid molecules associate into hexameric rosettes that bring about the formation of poly(adenine) triple helices which further polymerise into extended nanofibers. Fundamentally, this discovery demonstrates that a small molecule can induce nucleic acid assembly. This association gives rise to a non-canonical DNA structure which forms the basis of a new nanomaterial produced from inexpensive, readily available building blocks. In principle, small molecule-mediated DNA assembly is not unique to this nucleobase-small molecule combination, such that there is a vast potential for further expansion of the DNA alphabet through this methodology. Furthermore, the combination of this orthogonal assembly mode with traditional DNA nanotechnology based on standard Watson-Crick hybridization could be used to build DNA nanostructures with greater complexity.

Chapter 3 examines the mechanisms involved in cyanuric acid-mediated poly(adenine) nanofiber assembly as well as key parameters affecting the process. This investigation uncovers strand alignment within the nanofibers and determines that the supramolecular polymerization of poly(adenine) strands proceeds via a cooperative mechanism. Cyanuric acid concentration and poly(adenine) strand length are identified as the main parameters affecting the assembly process and nanofiber properties, most notably their thermal stability. Moreover, a novel methodology based on thermal hysteresis analysis is applied to acquire a quantitative representation of the kinetic and thermodynamic parameters underlying the cooperative supramolecular polymerization mechanism. Such insight into the mechanism and tunable assembly product behaviour in response to parameter variations is valuable for future development of applications of this unique DNA-based nanomaterial.

Chapter 4 departs from the context of basic DNA molecular recognition properties and supramolecular polymerization addressed in the previous two chapters by focusing on a different set of intermolecular interactions to generate long-range DNA assembly. Blunt-ended DNA tiles

(tiles lacking complementary single-stranded portions at the duplex ends) are modified with cholesterol, a hydrophobic anchoring moiety, and interfaced with lipid bilayers, a soft support with composition-dependent properties. Striking a balance between competing and complementary interactions (base-stacking between blunt DNA duplex ends, cholesterol anchor embedding into the hydrophobic portion of the bilayer, electrostatic DNA binding to the bilayer surface, and the fluid nature of the lipid bilayer) leads to the formation of dynamic lattices with long-range order and tunable morphology on a cell membrane-like support. In turn, these networks represent a tool for organizing materials such as nanoparticles or proteins at a biological interface, with potential applications in cellular recognition, development of model systems to study membrane proteins, plasmonics, or light harvesting.

The Appendix which closes the thesis takes another look at cyanuric acid-mediated poly(adenine) assembly. Nuclear magnetic resonance (NMR) spectroscopy studies on pre-organized poly(adenine) constructs are pursued to directly characterize the internal structure of cyanuric acid-mediated assembly to further validate the hexameric rosette architecture proposed in Chapter 2. Preliminary experiments are unsuccessful owing to the polymerization of constructs under investigation, however alternative strategies are suggested.

On the whole, this thesis presents how information encoded in simple molecular building blocks and the interplay between various intermolecular interactions can be directed to reprogram and manipulate DNA assembly and long-range organization. The resulting structures acquire unique features and higher complexity that suggest new possibilities for DNA-based nanomaterials.

1.7 References

1 Whitesides, G. M. & Grzybowski, B. Self-assembly at all scales. *Science* **295**, 2418-2421 (2002).

2 Lehn, J.-M. Toward complex matter: Supramolecular chemistry and self-organization. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4763-4768 (2002).

3 Steed, J. W. & Atwood, J. L. *Concepts* in *Supramolecular Chemistry*. 1-48 (John Wiley & Sons, Ltd, 2009).

4 Steed, J. W. & Atwood, J. L. *Self-assembly* in *Supramolecular Chemistry*. 10.1002/9780470740880.ch10 591-706 (John Wiley & Sons, Ltd, 2009).

5 Ozin, G. A., Arsenault, A. C. & Cademartiri, L. *Nanochemistry: A chemical approach to nanomaterials*. (Royal Society of Chemistry, 2009).

6 Hsia, Y., Bale, J. B., Gonen, S., Shi, D., Sheffler, W., Fong, K. K., Nattermann, U., Xu, C., Huang, P.-S., Ravichandran, R., Yi, S., Davis, T. N., Gonen, T., King, N. P. & Baker, D. Design of a hyperstable 60-subunit protein icosahedron. *Nature* **535**, 136-139 (2016).
7 Butterfield, G. L., Lajoie, M. J., Gustafson, H. H., Sellers, D. L., Nattermann, U., Ellis, D., Bale, J. B., Ke, S., Lenz, G. H., Yehdego, A., Ravichandran, R., Pun, S. H., King, N. P. & Baker, D. Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature* **552**, 415 (2017).
8 Farokhzad, O. C. & Langer, R. Impact of nanotechnology on drug delivery. *ACS Nano* **3**, 16-20 (2009).
9 Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature* **437**, 640 (2005).
10 Fletcher, D. A. & Mullins, R. D. Cell mechanics and the cytoskeleton. *Nature* **463**, 485-492 (2010).
11 Hynes, R. O. The extracellular matrix: not just pretty fibrils. *Science* **326**, 1216-1219 (2009).
12 Frantz, C., Stewart, K. M. & Weaver, V. M. The extracellular matrix at a glance. *Journal of cell science* **123**, 4195-4200 (2010).
13 Wang, C., Wang, Z. & Zhang, X. Amphiphilic building blocks for self-assembly: from amphiphiles to supra-amphiphiles. *Acc. Chem. Res.* **45**, 608-618 (2012).
14 Israelachvili, J. N., Mitchell, D. J. & Ninham, B. W. Theory of self-assembly of lipid bilayers and vesicles. *Biochim. Biophys. Acta, Biomembr.* **470**, 185-201 (1977).
15 Langecker, M., Arnaut, V., List, J. & Simmel, F. C. DNA nanostructures interacting with lipid bilayer membranes. *Acc. Chem. Res.* **47**, 1807-1815 (2014).
16 van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9**, 112-124 (2008).
17 Tresset, G. The multiple faces of self-assembled lipidic systems. *PMC Biophys.* **2** (2009).
18 Marsh, D. Cholesterol-induced fluid membrane domains: A compendium of lipid-raft ternary phase diagrams. *Biochim. Biophys. Acta, Biomembr.* **1788**, 2114-2123 (2009).
19 Fong, C., Le, T. & Drummond, C. J. Lyotropic liquid crystal engineering–ordered nanostructured small molecule amphiphile self-assembly materials by design. *Chem. Soc. Rev.* **41**, 1297-1322 (2012).
20 Kim, H.-J., Kim, T. & Lee, M. Responsive nanostructures from aqueous assembly of rigid-flexible block molecules. *Acc. Chem. Res.* **44**, 72-82 (2011).
21 Mayoral Munoz, M. J. & Fernandez, G. Metallosupramolecular amphiphilic π-systems. *Chem. Sci.* **3**, 1395-1398 (2012).
22 Jie, K., Zhou, Y., Yao, Y. & Huang, F. Macrocyclic amphiphiles. *Chem. Soc. Rev.* **44**, 3568-3587 (2015).
23 Cui, H., Webber, M. J. & Stupp, S. I. Self-assembly of peptide amphiphiles: From molecules to nanostructures to biomaterials. *Biopolymers* **94**, 1-18 (2010).
24 Kwak, M. & Herrmann, A. Nucleic acid amphiphiles: synthesis and self-assembled nanostructures. *Chem. Soc. Rev.* **40**, 5745-5755 (2011).
25 Edwardson, T. G. W., Carneiro, K. M. M., Serpell, C. J. & Sleiman, H. F. An efficient and modular route to sequence-defined polymers appended to DNA. *Angew. Chem., Int. Ed.* **53**, 4567-4571 (2014).
26 Kang, Y., Liu, K. & Zhang, X. Supra-amphiphiles: A new bridge between colloidal science and supramolecular chemistry. *Langmuir* **30**, 5989-6001 (2014).

27 Mai, Y. & Eisenberg, A. Self-assembly of block copolymers. *Chem. Soc. Rev.* **41**, 5969-5985 (2012).
28 Zayed, J. M., Nouvel, N., Rauwald, U. & Scherman, O. A. Chemical complexity - supramolecular self-assembly of synthetic and biological building blocks in water. *Chem. Soc. Rev.* **39**, 2806-2816 (2010).
29 Ramanathan, M., Shrestha, L. K., Mori, T., Ji, Q., Hill, J. P. & Ariga, K. Amphiphile nanoarchitectonics: from basic physical chemistry to advanced applications. *Phys. Chem. Chem. Phys.* **15**, 10580-10611 (2013).
30 Stuart, M. A. C., Huck, W. T. S., Genzer, J., Müller, M., Ober, C., Stamm, M., Sukhorukov, G. B., Szleifer, I., Tsukruk, V. V., Urban, M., Winnik, F., Zauscher, S., Luzinov, I. & Minko, S. Emerging applications of stimuli-responsive polymer materials. *Nat. Mater.* **9**, 101-113 (2010).
31 Soussan, E., Cassel, S., Blanzat, M. & Rico-Lattes, I. Drug delivery by soft matter: Matrix and vesicular carriers. *Angew. Chem., Int. Ed.* **48**, 274-288 (2009).
32 Krieg, E., Bastings, M. M., Besenius, P. & Rybtchinski, B. Supramolecular polymers in aqueous media. *Chem. Rev.* **116**, 2414-2477 (2016).
33 Aida, T., Meijer, E. W. & Stupp, S. I. Functional supramolecular polymers. *Science* **335**, 813-817 (2012).
34 Rest, C., Kandanelli, R. & Fernández, G. Strategies to create hierarchical self-assembled structures via cooperative non-covalent interactions. *Chem. Soc. Rev.* **44**, 2543-2572 (2015).
35 Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384-396 (2014).
36 Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annu. Rev. Biochem.* **78**, 929-958 (2009).
37 Durham, A. C. & Klug, A. Polymerization of tobacco mosaic virus protein and its control. *Nat. New Biol.* **229**, 42-46 (1971).
38 Fowler, D. M., Koulov, A. V., Balch, W. E. & Kelly, J. W. Functional amyloid – from bacteria to humans. *Trends Biochem. Sci.* **32**, 217-224 (2007).
39 Pollard, T. D. & Borisy, G. G. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **112**, 453-465 (2003).
40 Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884-890 (2003).
41 Fitzpatrick, A. W. P., Debelouchina, G. T., Bayro, M. J., Clare, D. K., Caporini, M. A., Bajaj, V. S., Jaroniec, C. P., Wang, L., Ladizhansky, V., Müller, S. A., MacPhee, C. E., Waudby, C. A., Mott, H. R., De Simone, A., Knowles, T. P. J., Saibil, H. R., Vendruscolo, M., Orlova, E. V., Griffin, R. G. & Dobson, C. M. Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5468-5473 (2013).
42 Wetzel, R. Kinetics and thermodynamics of amyloid fibril assembly. *Acc. Chem. Res.* **39**, 671-679 (2006).
43 De Greef, T. F., Smulders, M. M., Wolffs, M., Schenning, A. P., Sijbesma, R. P. & Meijer, E. W. Supramolecular polymerization. *Chem. Rev.* **109**, 5687-5754 (2009).
44 Goldstein, R. F. & Stryer, L. Cooperative polymerization reactions. Analytical approximations, numerical examples, and experimental strategy. *Biophys. J.* **50**, 583-599 (1986).
45 Zhao, D. & Moore, J. S. Nucleation-elongation: a mechanism for cooperative supramolecular polymerization. *Org. Biomol. Chem.* **1**, 3471-3491 (2003).

46 Sorrenti, A., Leira-Iglesias, J., Markvoort, A. J., de Greef, T. F. A. & Hermans, T. M. Non-equilibrium supramolecular polymerization. *Chem. Soc. Rev.* **46**, 5476-5490 (2017).
47 Wang, X., Guerin, G., Wang, H., Wang, Y., Manners, I. & Winnik, M. A. Cylindrical block copolymer micelles and co-micelles of controlled length and architecture. *Science* **317**, 644-647 (2007).
48 Gilroy, J. B., Gädt, T., Whittell, G. R., Chabanne, L., Mitchels, J. M., Richardson, R. M., Winnik, M. A. & Manners, I. Monodisperse cylindrical micelles by crystallization-driven living self-assembly. *Nat. Chem.* **2**, 566-570 (2010).
49 Albertazzi, L., van der Zwaag, D., Leenders, C. M. A., Fitzner, R., van der Hofstad, R. W. & Meijer, E. W. Probing exchange pathways in one-dimensional aggregates with super-resolution microscopy. *Science* **344**, 491-495 (2014).
50 Pal, A., Malakoutikhah, M., Leonetti, G., Tezcan, M., Colomb-Delsuc, M., Nguyen, V. D., van der Gucht, J. & Otto, S. Controlling the structure and length of self-synthesizing supramolecular polymers through nucleated growth and disassembly. *Angew. Chem., Int. Ed.* **54**, 7852-7856 (2015).
51 van der Zwaag, D., Pieters, P. A., Korevaar, P. A., Markvoort, A. J., Spiering, A. J. H., de Greef, T. F. A. & Meijer, E. W. Kinetic analysis as a tool to distinguish pathway complexity in molecular assembly: an unexpected outcome of structures in competition. *J. Am. Chem. Soc.* **137**, 12677-12688 (2015).
52 Ogi, S., Sugiyasu, K., Manna, S., Samitsu, S. & Takeuchi, M. Living supramolecular polymerization realized through a biomimetic approach. *Nat. Chem.* **6**, 188-195 (2014).
53 Kang, J., Miyajima, D., Mori, T., Inoue, Y., Itoh, Y. & Aida, T. A rational strategy for the realization of chain-growth supramolecular polymerization. *Science* **347**, 646-651 (2015).
54 Mattia, E. & Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **10**, 111-119 (2015).
55 Fouquey, C., Lehn, J.-M. & Levelut, A.-M. Molecular recognition directed self-assembly of supramolecular liquid crystalline polymers from complementary chiral components. *Adv. Mater.* **2**, 254-257 (1990).
56 Kotera, M., Lehn, J. M. & Vigneron, J. P. Self-assembled supramolecular rigid rods. *J. Chem. Soc., Chem. Commun.* DOI 10.1039/c39940000197, 197-199 (1994).
57 Sijbesma, R. P., Beijer, F. H., Brunsveld, L., Folmer, B. J. B., Hirschberg, J. H. K. K., Lange, R. F. M., Lowe, J. K. L. & Meijer, E. W. Reversible polymers formed from self-complementary monomers using quadruple hydrogen bonding. *Science* **278**, 1601-1604 (1997).
58 Hirschberg, J. H. K. K., Brunsveld, L., Ramzi, A., Vekemans, J. A. J. M., Sijbesma, R. P. & Meijer, E. W. Helical self-assembled polymers from cooperative stacking of hydrogen-bonded pairs. *Nature* **407**, 167-170 (2000).
59 Brunsveld, L., Vekemans, J. A. J. M., Hirschberg, J. H. K. K., Sijbesma, R. P. & Meijer, E. W. Hierarchical formation of helical supramolecular polymers via stacking of hydrogen-bonded pairs in water. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4977-4982 (2002).
60 Fenniri, H., Mathivanan, P., Vidale, K. L., Sherman, D. M., Hallenga, K., Wood, K. V. & Stowell, J. G. Helical rosette nanotubes: Design, self-assembly, and characterization. *J. Am. Chem. Soc.* **123**, 3854-3855 (2001).
61 Borzsonyi, G., Beingessner, R. L., Yamazaki, T., Cho, J.-Y., Myles, A. J., Malac, M., Egerton, R., Kawasaki, M., Ishizuka, K., Kovalenko, A. & Fenniri, H. Water-soluble J-

type rosette nanotubes with giant molar ellipticity. *J. Am. Chem. Soc.* **132**, 15136-15139 (2010).
62 Fenniri, H., Deng, B. L., Ribbe, A. E., Hallenga, K., Jacob, J. & Thiyagarajan, P. Entropically driven self-assembly of multichannel rosette nanotubes. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 6487 (2002).
63 Smulders, M. M., Schenning, A. P. & Meijer, E. W. Insight into the mechanisms of cooperative self-assembly: the "sergeants-and-soldiers" principle of chiral and achiral C3-symmetrical discotic triamides. *J. Am. Chem. Soc.* **130**, 606-611 (2008).
64 Baker, M. B., Albertazzi, L., Voets, I. K., Leenders, C. M. A., Palmans, A. R. A., Pavan, G. M. & Meijer, E. W. Consequences of chirality on the dynamics of a water-soluble supramolecular polymer. *Nat. Commun.* **6**, 6234 (2015).
65 Hartgerink, J. D., Beniash, E. & Stupp, S. I. Self-assembly and mineralization of peptide-amphiphile nanofibers. *Science* **294**, 1684-1688 (2001).
66 Jonkheijm, P., van der Schoot, P., Schenning, A. P. H. J. & Meijer, E. W. Probing the solvent-assisted nucleation pathway in chemical self-assembly. *Science* **313**, 80-83 (2006).
67 Hill, J. P., Jin, W., Kosaka, A., Fukushima, T., Ichihara, H., Shimomura, T., Ito, K., Hashizume, T., Ishii, N. & Aida, T. Self-assembled hexa-peri-hexabenzocoronene graphitic nanotube. *Science* **304**, 1481-1483 (2004).
68 Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E. & Khazanovich, N. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **366**, 324-327 (1993).
69 Hartgerink, J. D., Granja, J. R., Milligan, R. A. & Ghadiri, M. R. Self-assembling peptide nanotubes. *J. Am. Chem. Soc.* **118**, 43-50 (1996).
70 Ghadiri, M. R., Granja, J. R. & Buehler, L. K. Artificial transmembrane ion channels from self-assembling peptide nanotubes. *Nature* **369**, 301-304 (1994).
71 Horne, W. S., Stout, C. D. & Ghadiri, M. R. A heterocyclic peptide nanotube. *J. Am. Chem. Soc.* **125**, 9372-9376 (2003).
72 Hartgerink, J. D., Beniash, E. & Stupp, S. I. Peptide-amphiphile nanofibers: A versatile scaffold for the preparation of self-assembling materials. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5133-5138 (2002).
73 Silva, G. A., Czeisler, C., Niece, K. L., Beniash, E., Harrington, D. A., Kessler, J. A. & Stupp, S. I. Selective differentiation of neural progenitor cells by high-epitope density nanofibers. *Science* **303**, 1352-1355 (2004).
74 Webber, M. J., Tongers, J., Newcomb, C. J., Marquardt, K.-T., Bauersachs, J., Losordo, D. W. & Stupp, S. I. Supramolecular nanostructures that mimic VEGF as a strategy for ischemic tissue repair. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13438-13443 (2011).
75 Schenning, A. P. H. J., Jonkheijm, P., Peeters, E. & Meijer, E. W. Hierarchical order in supramolecular assemblies of hydrogen-bonded oligo(p-phenylene vinylene)s. *J. Am. Chem. Soc.* **123**, 409-416 (2001).
76 Jonkheijm, P., Hoeben, F. J. M., Kleppinger, R., van Herrikhuyzen, J., Schenning, A. P. H. J. & Meijer, E. W. Transfer of π-conjugated columnar stacks from solution to surfaces. *J. Am. Chem. Soc.* **125**, 15941-15949 (2003).
77 Korevaar, P. A., George, S. J., Markvoort, A. J., Smulders, M. M., Hilbers, P. A., Schenning, A. P., De Greef, T. F. & Meijer, E. W. Pathway complexity in supramolecular polymerization. *Nature* **481**, 492-496 (2012).

78 Yamamoto, Y., Fukushima, T., Suna, Y., Ishii, N., Saeki, A., Seki, S., Tagawa, S., Taniguchi, M., Kawai, T. & Aida, T. Photoconductive coaxial nanotubes of molecularly connected electron donor and acceptor layers. *Science* **314**, 1761-1764 (2006).
79 Yamamoto, Y., Zhang, G., Jin, W., Fukushima, T., Ishii, N., Saeki, A., Seki, S., Tagawa, S., Minari, T., Tsukagoshi, K. & Aida, T. Ambipolar-transporting coaxial nanotubes with a tailored molecular graphene–fullerene heterojunction. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21051-21056 (2009).
80 Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* **171**, 737-738 (1953).
81 Goodsell, D. S. RCSB Protein Data Bank. *Molecule of the month - DNA*, <10.2210/rcsb_pdb/mom_2001_11> (2001).
82 Leslie, A. G. W., Arnott, S., Chandrasekaran, R. & Ratliff, R. L. Polymorphism of DNA double helices. *J. Mol. Biol.* **143**, 49-72 (1980).
83 Choi, J. & Majima, T. Conformational changes of non-B DNA. *Chem. Soc. Rev.* **40**, 5893-5909 (2011).
84 Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **16**, 907-916 (1963).
85 International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860 (2001).
86 Shapiro, J. A. & von Sternberg, R. Why repetitive DNA is essential to genome function. *Biol. Rev.* **80**, 227-250 (2005).
87 Zhao, J., Bacolla, A., Wang, G. & Vasquez, K. M. Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* **67**, 43-62 (2010).
88 Davis, J. T. G-Quartets 40 years later: From 5′-GMP to molecular biology and supramolecular chemistry. *Angew. Chem., Int. Ed.* **43**, 668-698 (2004).
89 Gehring, K., Leroy, J. L. & Gueron, M. A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* **363**, 561-564 (1993).
90 Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182-186 (2013).
91 Zeraati, M., Langley, D. B., Schofield, P., Moye, A. L., Rouet, R., Hughes, W. E., Bryan, T. M., Dinger, M. E. & Christ, D. i-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* **10**, 631-637 (2018).
92 Vasquez, K. M. & Glazer, P. M. Triplex-forming oligonucleotides: principles and applications. *Q. Rev. Biophys.* **35**, 89-107 (2002).
93 Chakraborty, S., Sharma, S., Maiti, P. K. & Krishnan, Y. The poly dA helix: a new structural motif for high performance DNA-based molecular switches. *Nucleic Acids Res.* **37**, 2810-2817 (2009).
94 Mortimer, S. A., Kidwell, M. A. & Doudna, J. A. Insights into RNA structure and function from genome-wide studies. *Nat. Rev. Genet.* **15**, 469-479 (2014).
95 Sharp, P. A. The Centrality of RNA. *Cell* **136**, 577-580 (2009).
96 Aldaye, F. A., Palmer, A. L. & Sleiman, H. F. Assembling materials with DNA as the guide. *Science* **321**, 1795-1799 (2008).
97 Seeman, N. C. Nanomaterials based on DNA. *Annu. Rev. Biochem.* **79**, 65-87 (2010).
98 Letsinger, R. L. & Mahadevan, V. Oligonucleotide synthesis on a polymer support. *J. Am. Chem. Soc.* **87**, 3526-3527 (1965).

99 Merrifield, R. B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154 (1963).
100 Beaucage, S. L. & Caruthers, M. H. Deoxynucleoside phosphoramidites - A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **22**, 1859-1862 (1981).
101 McBride, L. J. & Caruthers, M. H. An investigation of several deoxynucleoside phosphoramidites useful for synthesizing deoxyoligonucleotides. *Tetrahedron Lett.* **24**, 245-248 (1983).
102 Alvarado-Urbina, G., Sathe, G., Liu, W., Gillen, M., Duck, P., Bender, R. & Ogilvie, K. Automated synthesis of gene fragments. *Science* **214**, 270-274 (1981).
103 Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **11**, 499 (2014).
104 Hughes, R. A. & Ellington, A. D. Synthetic DNA synthesis and assembly: Putting the synthetic in synthetic biology. *Cold Spring Harbor Perspect. Biol.* **9** (2017).
105 Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346** (2014).
106 Kohman, R. E., Kunjapur, A. M., Hysolli, E., Wang, Y. & Church, G. M. From designing the molecules of life to designing life: Future applications derived from advances in DNA technologies. *Angew. Chem., Int. Ed.* **57**, 4313-4328 (2018).
107 Wojciechowski, F. & Leumann, C. J. Alternative DNA base-pairs: from efforts to expand the genetic code to potential material applications. *Chem. Soc. Rev.* **40**, 5669-5679 (2011).
108 Pinheiro, V. B. & Holliger, P. The XNA world: progress towards replication and evolution of synthetic genetic polymers. *Curr. Opin. Chem. Biol.* **16**, 245-252 (2012).
109 Sefah, K., Yang, Z., Bradley, K. M., Hoshika, S., Jiménez, E., Zhang, L., Zhu, G., Shanker, S., Yu, F., Turek, D., Tan, W. & Benner, S. A. In vitro selection with artificial expanded genetic information systems. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1449-1454 (2014).
110 Taylor, A. I. & Holliger, P. Directed evolution of artificial enzymes (XNAzymes) from diverse repertoires of synthetic genetic polymers. *Nat. Protoc.* **10**, 1625-1642 (2015).
111 Eschenmoser, A. Chemical etiology of nucleic acid structure. *Science* **284**, 2118-2124 (1999).
112 Jürg, H., Hans-Jorg, R., Markus, B., Alfred, G., Ulf, D., Michael, G., Ravichandran, K., Bernhard, J., Christian, L. & Albert, E. Why pentose - and not hexose - nucleic acids? Part III. Oligo(2′,3′-dideoxy-β-D-glucopyranosyl)nucleotides. ('Homo-DNA'): base-pairing properties. *Helv. Chim. Acta* **76**, 259-352 (1993).
113 Koshkin, A. A., Singh, S. K., Nielsen, P., Rajwanshi, V. K., Kumar, R., Meldgaard, M., Olsen, C. E. & Wengel, J. LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* **54**, 3607-3630 (1998).
114 Veedu, R. N. & Wengel, J. Locked nucleic acids: Promising nucleic acid analogs for therapeutic applications. *Chem. Biodiversity* **7**, 536-542 (2010).
115 Nielsen, P., Egholm, M., Berg, R. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **254**, 1497-1500 (1991).
116 Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B. & Nielsen, P. E. PNA hybridizes to complementary

oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **365**, 566-568 (1993).
117 Nielsen, P. E. Peptide Nucleic Acids (PNA) in chemical biology and drug discovery. *Chem. Biodiversity* **7**, 786-804 (2010).
118 Piccirilli, J. A., Krauch, T., Moroney, S. E. & Benner, S. A. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **343**, 33-37 (1990).
119 Yang, Z., Hutter, D., Sheng, P., Sismour, A. M. & Benner, S. A. Artificially expanded genetic information system: a new base pair with an alternative hydrogen bonding pattern. *Nucleic Acids Res.* **34**, 6095-6101 (2006).
120 Bain, J. D., Switzer, C., Chamberlin, R. & Benner, S. A. Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code. *Nature* **356**, 537-539 (1992).
121 Hirao, I. Unnatural base pair systems for DNA/RNA-based biotechnology. *Curr. Opin. Chem. Biol.* **10**, 622-627 (2006).
122 Liu, H., Gao, J., Lynch, S. R., Saito, Y. D., Maynard, L. & Kool, E. T. A four-base paired genetic helix with expanded size. *Science* **302**, 868-871 (2003).
123 Krueger, A. T. & Kool, E. T. Fluorescence of size-expanded DNA bases: Reporting on DNA sequence and structure with an unnatural genetic set. *J. Am. Chem. Soc.* **130**, 3989-3999 (2008).
124 Teo, Y. N. & Kool, E. T. DNA-multichromophore systems. *Chem. Rev.* **112**, 4221-4245 (2012).
125 Schweitzer, B. A. & Kool, E. T. Aromatic nonpolar nucleosides as hydrophobic isosteres of pyrimidines and purine nucleosides. *J. Org. Chem.* **59**, 7238-7242 (1994).
126 Morales, J. C. & Kool, E. T. Minor groove interactions between polymerase and DNA: more essential to replication than Watson−Crick hydrogen bonds? *J. Am. Chem. Soc.* **121**, 2323-2324 (1999).
127 Kim, T. W., Delaney, J. C., Essigmann, J. M. & Kool, E. T. Probing the active site tightness of DNA polymerase in subangstrom increments. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15803-15808 (2005).
128 Matray, T. J. & Kool, E. T. Selective and stable DNA base pairing without hydrogen bonds. *J. Am. Chem. Soc.* **120**, 6191-6192 (1998).
129 Hirao, I., Kimoto, M., Mitsui, T., Fujiwara, T., Kawai, R., Sato, A., Harada, Y. & Yokoyama, S. An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods* **3**, 729 (2006).
130 Kimoto, M., Kawai, R., Mitsui, T., Yokoyama, S. & Hirao, I. An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules. *Nucleic Acids Res.* **37**, e14 (2009).
131 Kimoto, M., Mitsui, T., Yamashige, R., Sato, A., Yokoyama, S. & Hirao, I. A new unnatural base pair system between fluorophore and quencher base analogues for nucleic acid-based imaging technology. *J. Am. Chem. Soc.* **132**, 15418-15426 (2010).
132 Ogawa, A. K., Wu, Y. Q., McMinn, D. L., Liu, J. Q., Schultz, P. G. & Romesberg, F. E. Efforts toward the expansion of the genetic alphabet: Information storage and replication with unnatural hydrophobic base pairs. *J. Am. Chem. Soc.* **122**, 3274-3287 (2000).
133 Yu, C., Henry, A. A., Romesberg, F. E. & Schultz, P. G. Polymerase recognition of unnatural base pairs. *Angew. Chem., Int. Ed.* **41**, 3841-3844 (2002).

134 Matsuda, S., Leconte, A. M. & Romesberg, F. E. Minor groove hydrogen bonds and the replication of unnatural base pairs. *J. Am. Chem. Soc.* **129**, 5551-5557 (2007).
135 Malyshev, D. A., Dhami, K., Lavergne, T., Chen, T., Dai, N., Foster, J. M., Correa, I. R., Jr. & Romesberg, F. E. A semi-synthetic organism with an expanded genetic alphabet. *Nature* **509**, 385-388 (2014).
136 Zhang, Y., Lamb, B. M., Feldman, A. W., Zhou, A. X., Lavergne, T., Li, L. & Romesberg, F. E. A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 1317-1322 (2017).
137 Zhang, Y., Ptacin, J. L., Fischer, E. C., Aerni, H. R., Caffaro, C. E., San Jose, K., Feldman, A. W., Turner, C. R. & Romesberg, F. E. A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **551**, 644-647 (2017).
138 Miyake, Y., Togashi, H., Tashiro, M., Yamaguchi, H., Oda, S., Kudo, M., Tanaka, Y., Kondo, Y., Sawa, R., Fujimoto, T., Machinami, T. & Ono, A. Mercury(II)-mediated formation of thymine−Hg(II)−thymine base pairs in DNA duplexes. *J. Am. Chem. Soc.* **128**, 2172-2173 (2006).
139 Ono, A., Cao, S., Togashi, H., Tashiro, M., Fujimoto, T., Machinami, T., Oda, S., Miyake, Y., Okamoto, I. & Tanaka, Y. Specific interactions between silver(I) ions and cytosine-cytosine pairs in DNA duplexes. *Chem. Commun.*, 4825-4827 (2008).
140 Soo, P. K., Cheulhee, J. & Gyu, P. H. "Illusionary" polymerase activity triggered by metal ions: Use for molecular logic-gate operations. *Angew. Chem., Int. Ed.* **49**, 9757-9760 (2010).
141 Kondo, J., Tada, Y., Dairaku, T., Hattori, Y., Saneyoshi, H., Ono, A. & Tanaka, Y. A metallo-DNA nanowire with uninterrupted one-dimensional silver array. *Nat. Chem.* **9**, 956-960 (2017).
142 Takezawa, Y. & Shionoya, M. Metal-mediated DNA base pairing: alternatives to hydrogen-bonded Watson-Crick base pairs. *Acc. Chem. Res.* **45**, 2066-2076 (2012).
143 Tanaka, K., Tengeiji, A., Kato, T., Toyama, N. & Shionoya, M. A discrete self-assembled metal array in artificial DNA. *Science* **299**, 1212-1213 (2003).
144 Tanaka, K., Yamada, Y. & Shionoya, M. Formation of silver(I)-mediated DNA duplex and triplex through an alternative base pair of pyridine nucleobases. *J. Am. Chem. Soc.* **124**, 8802-8803 (2002).
145 Branda, N., Kurz, G. & Lehn, J.-M. Janus wedges: a new approach towards nucleobase-pair recognition. *Chem. Commun.*, 2443-2444 (1996).
146 Chen, H., Meena & McLaughlin, L. W. A Janus-wedge DNA triplex with A-W1-T and G-W2-C base triplets. *J. Am. Chem. Soc.* **130**, 13190-13191 (2008).
147 Zeng, Y., Pratumyot, Y., Piao, X. & Bong, D. Discrete assembly of synthetic peptide-DNA triplex structures from polyvalent melamine-thymine bifacial recognition. *J. Am. Chem. Soc.* **134**, 832-835 (2012).
148 Chaput, J. C. & Switzer, C. A DNA pentaplex incorporating nucleobase quintets. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10614-10619 (1999).
149 Jiang, D. & Seela, F. Oligonucleotide duplexes and multistrand assemblies with 8-aza-2'-deoxyisoguanosine: a fluorescent isoG(d) shape mimic expanding the genetic alphabet and forming ionophores. *J. Am. Chem. Soc.* **132**, 4016-4024 (2010).
150 Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418**, 214-221 (2002).
151 Robertson, M. P. & Scott, W. G. The structural basis of ribozyme-catalyzed RNA assembly. *Science* **315**, 1549-1553 (2007).

152 Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920-930 (2000).
153 Joyce, G. F., Schwartz, A. W., Miller, S. L. & Orgel, L. E. The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4398-4402 (1987).
154 Hud, Nicholas V., Cafferty, Brian J., Krishnamurthy, R. & Williams, Loren D. The origin of RNA and "My Grandfather's Axe". *Chemistry & Biology* **20**, 466-474 (2013).
155 Hayatsu, R., Studier, M. H., Moore, L. P. & Anders, E. Purines and triazines in the Murchison meteorite. *Geochimica et Cosmochimica Acta* **39**, 471-488 (1975).
156 Cafferty, B. J., Gallego, I., Chen, M. C., Farley, K. I., Eritja, R. & Hud, N. V. Efficient self-assembly in water of long noncovalent polymers by nucleobase analogues. *J. Am. Chem. Soc.* **135**, 2447-2450 (2013).
157 Chen, M. C., Cafferty, B. J., Mamajanov, I., Gallego, I., Khanam, J., Krishnamurthy, R. & Hud, N. V. Spontaneous prebiotic formation of a beta-ribofuranoside that self-assembles with a complementary heterocycle. *J. Am. Chem. Soc.* **136**, 5640-5646 (2014).
158 Cafferty, B. J., Fialho, D. M., Khanam, J., Krishnamurthy, R. & Hud, N. V. Spontaneous formation and base pairing of plausible prebiotic nucleotides in water. *Nat. Commun.* **7**, 11328 (2016).
159 Hud, N. Our odyssey to find a plausible prebiotic path to RNA: The first twenty years. *Synlett* **28**, 36-55 (2017).
160 Roy, B., Bairi, P. & Nandi, A. K. Supramolecular assembly of melamine and its derivatives: nanostructures to functional materials. *RSC. Adv.* **4**, 1708-1734 (2014).
161 Wang, Y., Wei, B. & Wang, Q. Crystal structure of melamine cyanuric acid complex (1:1) trihydrochloride, MCA·3HCl. *J. Crystallogr. Spectrosc. Res.* **20**, 79-84 (1990).
162 Zerkowski, J. A., Seto, C. T. & Whitesides, G. M. Solid-state structures of rosette and crinkled tape motifs derived from the cyanuric acid melamine lattice. *J. Am. Chem. Soc.* **114**, 5473-5475 (1992).
163 Zerkowski, J. A., MacDonald, J. C., Seto, C. T., Wierda, D. A. & Whitesides, G. M. Design of organic structures in the solid state: Molecular tapes based on the network of hydrogen bonds present in the cyanuric acid·melamine complex. *J. Am. Chem. Soc.* **116**, 2382-2391 (1994).
164 Seto, C. T. & Whitesides, G. M. Self-assembly based on the cyanuric acid-melamine lattice. *J. Am. Chem. Soc.* **112**, 6409-6411 (1990).
165 Whitesides, G. M., Simanek, E. E., Mathias, J. P., Seto, C. T., Chin, D., Mammen, M. & Gordon, D. M. Noncovalent synthesis: using physical-organic chemistry to make aggregates. *Acc. Chem. Res.* **28**, 37-44 (1995).
166 Ma, M. & Bong, D. Determinants of cyanuric acid and melamine assembly in water. *Langmuir* **27**, 8841-8853 (2011).
167 Ranganathan, A., Pedireddi, V. R. & Rao, C. N. R. Hydrothermal synthesis of organic channel structures: 1:1 hydrogen-bonded adducts of melamine with cyanuric and trithiocyanuric acids. *J. Am. Chem. Soc.* **121**, 1752-1753 (1999).
168 Seeman, N. C. & Sleiman, H. F. DNA nanotechnology. *Nat. Rev. Mater.* **3**, 17068 (2017).
169 McLaughlin, C. K., Hamblin, G. D. & Sleiman, H. F. Supramolecular DNA assembly. *Chem. Soc. Rev.* **40**, 5647-5656 (2011).
170 Bujold, K. E., Lacroix, A. & Sleiman, H. F. DNA nanostructures at the interface with biology. *Chem.* **4**, 495-521 (2018).

171 Meselson, M. S. & Radding, C. M. A general model for genetic recombination. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 358-361 (1975).
172 Seeman, N. C. & Kallenbach, N. R. Design of immobile nucleic acid junctions. *Biophys. J.* **44**, 201-209 (1983).
173 Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **99**, 237-247 (1982).
174 Seeman, N. C. DNA in a material world. *Nature* **421**, 427-431 (2003).
175 Chen, J. & Seeman, N. C. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **350**, 631-633 (1991).
176 Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394**, 539-544 (1998).
177 Lin, C., Liu, Y., Rinker, S. & Yan, H. DNA tile based self-assembly: building complex nanoarchitectures. *ChemPhysChem* **7**, 1641-1647 (2006).
178 LaBean, T. H., Yan, H., Kopatsch, J., Liu, F., Winfree, E., Reif, J. H. & Seeman, N. C. Construction, analysis, ligation, and self-assembly of DNA triple crossover complexes. *J. Am. Chem. Soc.* **122**, 1848-1860 (2000).
179 Ke, Y., Liu, Y., Zhang, J. & Yan, H. A study of DNA tube formation mechanisms using 4-, 8-, and 12-helix DNA nanostructures. *J. Am. Chem. Soc.* **128**, 4414-4421 (2006).
180 Reishus, D., Shaw, B., Brun, Y., Chelyapov, N. & Adleman, L. Self-assembly of DNA double-double crossover complexes into high-density, doubly connected, planar structures. *J. Am. Chem. Soc.* **127**, 17590-17591 (2005).
181 Ding, B., Sha, R. & Seeman, N. C. Pseudohexagonal 2D DNA crystals from double crossover cohesion. *J. Am. Chem. Soc.* **126**, 10230-10231 (2004).
182 Yan, H., Park, S. H., Finkelstein, G., Reif, J. H. & LaBean, T. H. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* **301**, 1882-1884 (2003).
183 He, Y., Chen, Y., Liu, H., Ribbe, A. E. & Mao, C. Self-assembly of hexagonal DNA two-dimensional (2D) arrays. *J. Am. Chem. Soc.* **127**, 12202-12203 (2005).
184 Zhang, C., Su, M., He, Y., Zhao, X., Fang, P., Ribbe, A. E., Jiang, W. & Mao, C. Conformational flexibility facilitates self-assembly of complex DNA nanostructures. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 10665-10669 (2008).
185 He, Y., Tian, Y., Ribbe, A. E. & Mao, C. Highly connected two-dimensional crystals of DNA six-point-stars. *J. Am. Chem. Soc.* **128**, 15978-15979 (2006).
186 Rothemund, P. W. K., Ekani-Nkodo, A., Papadakis, N., Kumar, A., Fygenson, D. K. & Winfree, E. Design and characterization of programmable DNA nanotubes. *J. Am. Chem. Soc.* **126**, 16344-16352 (2004).
187 Park, S. H., Barish, R., Li, H., Reif, J. H., Finkelstein, G., Yan, H. & LaBean, T. H. Three-helix bundle DNA tiles self-assemble into 2D lattice or 1D templates for silver nanowires. *Nano Lett.* **5**, 693-696 (2005).
188 Mathieu, F., Liao, S., Kopatsch, J., Wang, T., Mao, C. & Seeman, N. C. Six-helix bundles designed from DNA. *Nano Lett.* **5**, 661-665 (2005).
189 Zheng, J., Birktoft, J. J., Chen, Y., Wang, T., Sha, R., Constantinou, P. E., Ginell, S. L., Mao, C. & Seeman, N. C. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **461**, 74-77 (2009).
190 Chworos, A., Severcan, I., Koyfman, A. Y., Weinkam, P., Oroudjev, E., Hansma, H. G. & Jaeger, L. Building programmable jigsaw puzzles with RNA. *Science* **306**, 2068-2072 (2004).

191 Hamada, S. & Murata, S. Substrate-assisted assembly of interconnected single-duplex DNA nanostructures. *Angew. Chem., Int. Ed.* **48**, 6820-6823 (2009).
192 He, Y., Ye, T., Su, M., Zhang, C., Ribbe, A. E., Jiang, W. & Mao, C. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452**, 198-201 (2008).
193 Tian, C., Li, X., Liu, Z., Jiang, W., Wang, G. & Mao, C. Directed self-assembly of DNA tiles into complex nanocages. *Angew. Chem., Int. Ed.* **126**, 8179-8182 (2014).
194 Kuzuya, A., Wang, R., Sha, R. & Seeman, N. C. Six-helix and eight-helix DNA nanotubes assembled from half-tubes. *Nano Lett.* **7**, 1757-1763 (2007).
195 Wang, T., Schiffels, D., Martinez Cuesta, S., Kuchnir Fygenson, D. & Seeman, N. C. Design and characterization of 1D nanotubes and 2D periodic arrays self-assembled from DNA multi-helix bundles. *J. Am. Chem. Soc.* **134**, 1606-1616 (2012).
196 He, Y., Tian, Y., Chen, Y., Deng, Z., Ribbe, A. E. & Mao, C. Sequence symmetry as a tool for designing DNA nanostructures. *Angew. Chem., Int. Ed.* **44**, 6694-6696 (2005).
197 Liu, H., Chen, Y., He, Y., Ribbe, A. E. & Mao, C. Approaching the limit: can one DNA oligonucleotide assemble into large nanostructures? *Angew. Chem., Int. Ed.* **45**, 1942-1945 (2006).
198 Sun, X., Hyeon Ko, S., Zhang, C., Ribbe, A. E. & Mao, C. Surface-mediated DNA self-assembly. *J. Am. Chem. Soc.* **131**, 13248-13249 (2009).
199 Liu, Y., Ke, Y. & Yan, H. Self-assembly of symmetric finite-size DNA nanoarrays. *J. Am. Chem. Soc.* **127**, 17140-17141 (2005).
200 Liu, D., Wang, M., Deng, Z., Walulu, R. & Mao, C. Tensegrity: construction of rigid DNA triangles with flexible four-arm DNA junctions. *J. Am. Chem. Soc.* **126**, 2324-2325 (2004).
201 Hao, Y., Kristiansen, M., Sha, R., Birktoft, J. J., Hernandez, C., Mao, C. & Seeman, N. C. A device that operates within a self-assembled 3D DNA crystal. *Nat. Chem.* **9**, 824-827 (2017).
202 Ko, S. H., Su, M., Zhang, C., Ribbe, A. E., Jiang, W. & Mao, C. Synergistic self-assembly of RNA and DNA molecules. *Nat. Chem.* **2**, 1050-1055 (2010).
203 Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333**, 470-474 (2011).
204 Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302 (2006).
205 Zhao, Z., Yan, H. & Liu, Y. A route to scale up DNA origami using DNA tiles as folding staples. *Angew. Chem., Int. Ed.* **49**, 1414-1417 (2010).
206 Zhao, Z., Liu, Y. & Yan, H. Organizing DNA origami tiles into larger structures using preformed scaffold frames. *Nano Lett.* **11**, 2997-3002 (2011).
207 Rajendran, A., Endo, M., Katsuda, Y., Hidaka, K. & Sugiyama, H. Programmed two-dimensional self-assembly of multiple DNA origami jigsaw pieces. *ACS Nano* **5**, 665-671 (2011).
208 Jungmann, R., Scheible, M., Kuzyk, A., Pardatscher, G., Castro, C. E. & Simmel, F. C. DNA origami-based nanoribbons: assembly, length distribution, and twist. *Nanotechnology* **22**, 275301 (2011).
209 Liu, W., Zhong, H., Wang, R. & Seeman, N. C. Crystalline two-dimensional DNA origami arrays. *Angew. Chem., Int. Ed.* **50**, 264-267 (2011).
210 Woo, S. & Rothemund, P. W. K. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nat. Chem.* **3**, 620-627 (2011).

211 Woo, S. & Rothemund, P. W. Self-assembly of two-dimensional DNA origami lattices using cation-controlled surface diffusion. *Nat. Commun.* **5**, 4889 (2014).
212 Marchi, A. N., Saaem, I., Vogen, B. N., Brown, S. & LaBean, T. H. Toward larger DNA origami. *Nano Lett.* **14**, 5740-5747 (2014).
213 Andersen, E. S., Dong, M., Nielsen, M. M., Jahn, K., Subramani, R., Mamdouh, W., Golas, M. M., Sander, B., Stark, H., Oliveira, C. L. P., Pedersen, J. S., Birkedal, V., Besenbacher, F., Gothelf, K. V. & Kjems, J. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459**, 73-76 (2009).
214 Douglas, S. M., Dietz, H., Liedl, T., Hogberg, B., Graf, F. & Shih, W. M. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414-418 (2009).
215 Douglas, S. M., Chou, J. J. & Shih, W. M. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6644-6648 (2007).
216 Ke, Y., Douglas, S. M., Liu, M., Sharma, J., Cheng, A., Leung, A., Liu, Y., Shih, W. M. & Yan, H. Multilayer DNA origami packed on a square lattice. *J. Am. Chem. Soc.* **131**, 15903-15908 (2009).
217 Ke, Y., Voigt, N. V., Gothelf, K. V. & Shih, W. M. Multilayer DNA origami packed on hexagonal and hybrid lattices. *J. Am. Chem. Soc.* **134**, 1770-1774 (2012).
218 Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325**, 725-730 (2009).
219 Han, D., Pal, S., Nangreave, J., Deng, Z., Liu, Y. & Yan, H. DNA origami with complex curvatures in three-dimensional space. *Science* **332**, 342-346 (2011).
220 Benson, E., Mohammed, A., Gardell, J., Masich, S., Czeizler, E., Orponen, P. & Hogberg, B. DNA rendering of polyhedral meshes at the nanoscale. *Nature* **523**, 441-444 (2015).
221 Maier, A. M., Bae, W., Schiffels, D., Emmerig, J. F., Schiff, M. & Liedl, T. Self-assembled DNA tubes forming helices of controlled diameter and chirality. *ACS Nano* **11**, 1301-1306 (2017).
222 Zhang, F., Jiang, S., Wu, S., Li, Y., Mao, C., Liu, Y. & Yan, H. Complex wireframe DNA origami nanostructures with multi-arm junction vertices. *.Nat, Nanotechnol.* **10**, 779-784 (2015).
223 Iinuma, R., Ke, Y., Jungmann, R., Schlichthaerle, T., Woehrstein, J. B. & Yin, P. Polyhedra self-assembled from DNA tripods and characterized with 3D DNA-PAINT. *Science* **344**, 65-69 (2014).
224 Han, D., Pal, S., Yang, Y., Jiang, S., Nangreave, J., Liu, Y. & Yan, H. DNA gridiron nanostructures based on four-arm junctions. *Science* **339**, 1412-1415 (2013).
225 Pinheiro, A. V., Han, D., Shih, W. M. & Yan, H. Challenges and opportunities for structural DNA nanotechnology. *Nat. Nanotechnol.* **6**, 763-772 (2011).
226 Castro, C. E., Kilchherr, F., Kim, D.-N., Shiao, E. L., Wauer, T., Wortmann, P., Bathe, M. & Dietz, H. A primer to scaffolded DNA origami. *Nat. Methods* **8**, 221 (2011).
227 Veneziano, R., Ratanalert, S., Zhang, K., Zhang, F., Yan, H., Chiu, W. & Bathe, M. Designer nanoscale DNA assemblies programmed from the top down. *Science* **352**, 1534 (2016).
228 Dunn, K. E., Dannenberg, F., Ouldridge, T. E., Kwiatkowska, M., Turberfield, A. J. & Bath, J. Guiding the folding pathway of DNA origami. *Nature* **525**, 82 (2015).

229 Niekamp, S., Blumer, K., Nafisi, P. M., Tsui, K., Garbutt, J. & Douglas, S. M. Folding complex DNA nanostructures from limited sets of reusable sequences. *Nucleic Acids Res.* **44**, e102-e102 (2016).
230 Jungmann, R., Liedl, T., Sobey, T. L., Shih, W. & Simmel, F. C. Isothermal assembly of DNA origami structures using denaturing agents. *J. Am. Chem. Soc.* **130**, 10062-10063 (2008).
231 Bellot, G., McClintock, M. A., Lin, C. & Shih, W. M. Recovery of intact DNA nanostructures after agarose gel–based separation. *Nat. Methods* **8**, 192-194 (2011).
232 Yin, P., Hariadi, R. F., Sahu, S., Choi, H. M. T., Park, S. H., LaBean, T. H. & Reif, J. H. Programming DNA tube circumferences. *Science* **321**, 824-826 (2008).
233 Aldaye, F. A., Senapedis, W. T., Silver, P. A. & Way, J. C. A structurally tunable DNA-based extracellular matrix. *J. Am. Chem. Soc.* **132**, 14727-14729 (2010).
234 Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485**, 623-626 (2012).
235 Ke, Y., Ong, L. L., Shih, W. M. & Yin, P. Three-dimensional structures self-assembled from DNA bricks. *Science* **338**, 1177-1183 (2012).
236 Wei, B., Ong, L. L., Chen, J., Jaffe, A. S. & Yin, P. Complex reconfiguration of DNA nanostructures. *Angew. Chem.* **126**, 7605-7609 (2014).
237 Ke, Y., Ong, L. L., Sun, W., Song, J., Dong, M., Shih, W. M. & Yin, P. DNA brick crystals with prescribed depths. *Nat. Chem.* **6**, 994-1002 (2014).
238 Geary, C., Rothemund, P. W. K. & Andersen, E. S. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science* **345**, 799-804 (2014).
239 Li, M., Zheng, M., Wu, S., Tian, C., Liu, D., Weizmann, Y., Jiang, W., Wang, G. & Mao, C. In vivo production of RNA nanostructures via programmed folding of single-stranded RNAs. *Nat. Commun.* **9**, 2196 (2018).
240 Han, D., Qi, X., Myhrvold, C., Wang, B., Dai, M., Jiang, S., Bates, M., Liu, Y., An, B., Zhang, F., Yan, H. & Yin, P. Single-stranded DNA and RNA origami. *Science* **358** (2017).
241 Ong, L. L., Hanikel, N., Yaghi, O. K., Grun, C., Strauss, M. T., Bron, P., Lai-Kee-Him, J., Schueder, F., Wang, B., Wang, P., Kishi, J. Y., Myhrvold, C., Zhu, A., Jungmann, R., Bellot, G., Ke, Y. & Yin, P. Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components. *Nature* **552**, 72-77 (2017).
242 Wagenbauer, K. F., Sigl, C. & Dietz, H. Gigadalton-scale shape-programmable DNA assemblies. *Nature* **552**, 78-83 (2017).
243 Praetorius, F., Kick, B., Behler, K. L., Honemann, M. N., Weuster-Botz, D. & Dietz, H. Biotechnological mass production of DNA origami. *Nature* **552**, 84-87 (2017).
244 Chidchob, P. & Sleiman, H. F. *Supramolecular chemistry with DNA* in *Macrocyclic and supramolecular chemistry*. doi:10.1002/9781119053859.ch2 (ed R. M. Izatt) (2016).
245 Shi, J. F. & Bergstrom, D. E. Assembly of novel DNA cycles with rigid tetrahedral linkers. *Angew. Chem., Int. Edit. Engl.* **36**, 111-113 (1997).
246 Scheffler, M., Dorenbeck, A., Jordan, S., Wüstefeld, M. & von Kiedrowski, G. Self-assembly of trisoligonucleotidyls: the case for nano-acetylene and nano-cyclobutadiene. *Angew. Chem., Int. Ed.* **38**, 3311-3315 (1999).
247 Aldaye, F. A. & Sleiman, H. F. Modular access to structurally switchable 3D discrete DNA assemblies. *J. Am. Chem. Soc.* **129**, 13376-13377 (2007).
248 Aldaye, F. A. & Sleiman, H. F. Sequential self-assembly of a DNA hexagon as a template for the organization of gold nanoparticles. *Angew. Chem., Int. Ed.* **45**, 2204-2209 (2006).

249 Aldaye, F. A. & Sleiman, H. F. Dynamic DNA templates for discrete gold nanoparticle assemblies: control of geometry, modularity, write/erase and structural switching. *J. Am. Chem. Soc.* **129**, 4130-4131 (2007).
250 Aldaye, F. A., Lo, P. K., Karam, P., McLaughlin, C. K., Cosa, G. & Sleiman, H. F. Modular construction of DNA nanotubes of tunable geometry and single- or double-stranded character. *Nat. Nanotechnol.* **4**, 349-352 (2009).
251 Lo, P. K., Karam, P., Aldaye, F. A., McLaughlin, C. K., Hamblin, G. D., Cosa, G. & Sleiman, H. F. Loading and selective release of cargo in DNA nanotubes with longitudinal variation. *Nat. Chem.* **2**, 319-328 (2010).
252 Lo, P. K., Altvater, F. & Sleiman, H. F. Templated synthesis of DNA nanotubes with controlled, predetermined lengths. *J. Am. Chem. Soc.* **132**, 10212-10214 (2010).
253 Eryazici, I., Yildirim, I., Schatz, G. C. & Nguyen, S. T. Enhancing the melting properties of small molecule-DNA hybrids through designed hydrophobic interactions: an experimental-computational study. *J. Am. Chem. Soc.* **134**, 7450-7458 (2012).
254 Greschner, A. A., Toader, V. & Sleiman, H. F. The role of organic linkers in directing DNA self-assembly and significantly stabilizing DNA duplexes. *J. Am. Chem. Soc.* **134**, 14382-14389 (2012).
255 Endo, M., Seeman, N. C. & Majima, T. DNA tube structures controlled by a four-way-branched DNA connector. *Angew. Chem., Int. Ed.* **44**, 6074-6077 (2005).
256 Liu, D., Park, S. H., Reif, J. H. & LaBean, T. H. DNA nanotubes self-assembled from triple-crossover tiles as templates for conductive nanowires. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 717-722 (2004).
257 Wilner, O. I., Henning, A., Shlyahovsky, B. & Willner, I. Covalently linked DNA nanotubes. *Nano Lett.* **10**, 1458-1465 (2010).
258 Cassinelli, V., Oberleitner, B., Sobotta, J., Nickels, P., Grossi, G., Kempter, S., Frischmuth, T., Liedl, T. & Manetto, A. One-step formation of "chain-armor"-stabilized DNA nanostructures. *Angew. Chem., Int. Ed.* **54**, 7795-7798 (2015).
259 Yang, Y., Endo, M., Hidaka, K. & Sugiyama, H. Photo-controllable DNA origami nanostructures assembling into predesigned multiorientational patterns. *J. Am. Chem. Soc.* **134**, 20645-20653 (2012).
260 Yang, H., McLaughlin, C. K., Aldaye, F. A., Hamblin, G. D., Rys, A. Z., Rouiller, I. & Sleiman, H. F. Metal-nucleic acid cages. *Nat. Chem.* **1**, 390-396 (2009).
261 Chien, M.-P., Rush, A. M., Thompson, M. P. & Gianneschi, N. C. Programmable shape-shifting micelles. *Angew. Chem., Int. Ed.* **49**, 5076-5080 (2010).
262 Serpell, C. J., Edwardson, T. G. W., Chidchob, P., Carneiro, K. M. M. & Sleiman, H. F. Precision polymers and 3D DNA nanostructures: emergent assemblies from new parameter space. *J. Am. Chem. Soc.* **136**, 15767-15774 (2014).
263 Vyborna, Y., Vybornyi, M., Rudnev, A. V. & Häner, R. DNA-grafted supramolecular polymers: Helical ribbon structures formed by self-assembly of pyrene–DNA chimeric oligomers. *Angew. Chem., Int. Ed.* **54**, 7934-7938 (2015).
264 Gothelf, K. V., Thomsen, A., Nielsen, M., Cló, E. & Brown, R. S. Modular DNA-programmed assembly of linear and branched conjugated nanostructures. *J. Am. Chem. Soc.* **126**, 1044-1046 (2004).
265 Mitra, D., Di Cesare, N. & Sleiman, H. F. Self-assembly of cyclic metal–DNA nanostructures using ruthenium tris(bipyridine)-branched oligonucleotides. *Angew. Chem., Int. Ed.* **43**, 5804-5808 (2004).

266 Choi, J. S., Kang, C. W., Jung, K., Yang, J. W., Kim, Y.-G. & Han, H. Synthesis of DNA triangles with vertexes of bis(terpyridine)iron(II) complexes. *J. Am. Chem. Soc.* **126**, 8606-8607 (2004).
267 Stewart, K. M. & McLaughlin, L. W. Four-arm oligonucleotide Ni(II)−cyclam-centered complexes as precursors for the generation of supramolecular periodic assemblies. *J. Am. Chem. Soc.* **126**, 2050-2057 (2004).
268 Stewart, K. M., Rojo, J. & McLaughlin, L. W. Ru(II) tris(bipyridyl) complexes with six oligonucleotide arms as precursors for the generation of supramolecular assemblies. *Angew. Chem., Int. Ed.* **43**, 5808-5811 (2004).
269 Yang, H. & Sleiman, H. F. Templated synthesis of highly stable, electroactive, and dynamic metal-DNA branched junctions. *Angew. Chem., Int. Ed.* **47**, 2443-2446 (2008).
270 Yang, H., Altvater, F., de Bruijn, A. D., McLaughlin, C. K., Lo, P. K. & Sleiman, H. F. Chiral metal-DNA four-arm junctions and metalated nanotubular structures. *Angew. Chem., Int. Ed.* **50**, 4620-4623 (2011).
271 Yang, H., Rys, A. Z., McLaughlin, C. K. & Sleiman, H. F. Templated ligand environments for the selective incorporation of different metals into DNA. *Angew. Chem., Int. Ed.* **48**, 9919-9923 (2009).
272 Clever, G. H., Kaul, C. & Carell, T. DNA–metal base pairs. *Angew. Chem., Int. Ed.* **46**, 6226-6236 (2007).
273 Yang, H., Metera, K. L. & Sleiman, H. F. DNA modified with metal complexes: Applications in the construction of higher order metal–DNA nanostructures. *Coordination Chemistry Reviews* **254**, 2403-2415 (2010).
274 Carneiro, K. M. M., Aldaye, F. A. & Sleiman, H. F. Long-range assembly of DNA into nanofibers and highly ordered networks using a block copolymer approach. *J. Am. Chem. Soc.* **132**, 679-685 (2009).
275 Edwardson, T. G. W., Carneiro, K. M. M., McLaughlin, C. K., Serpell, C. J. & Sleiman, H. F. Site-specific positioning of dendritic alkyl chains on DNA cages enables their geometry-dependent self-assembly. *Nat. Chem.* **5**, 868-875 (2013).
276 Ding, K., Alemdaroglu, F. E., Borsch, M., Berger, R. & Herrmann, A. Engineering the structural properties of DNA block copolymer micelles by molecular recognition. *Angew. Chem., Int. Ed.* **46**, 1172-1175 (2007).
277 Bousmail, D., Amrein, L., Fakhoury, J. J., Fakih, H. H., Hsu, J. C. C., Panasci, L. & Sleiman, H. F. Precision spherical nucleic acids for delivery of anticancer drugs. *Chem. Sci.* **8**, 6218-6229 (2017).
278 de Rochambeau, D., Barlog, M., Edwardson, T., Fakhoury, J., Stein, R., Bazzi, H. & Sleiman, H. "DNA-Teflon" sequence-controlled polymers. *Polymer Chemistry* **7**, 4998-5003 (2016).
279 Trinh, T., Chidchob, P., Bazzi, H. & Sleiman, H. DNA micelles as nanoreactors: efficient DNA functionalization with hydrophobic organic molecules. *Chem. Commun.* **52**, 10914-10917 (2016).
280 Trinh, T., Liao, C., Toader, V., Barłóg, M., Bazzi, H. S., Li, J. & Sleiman, H. F. DNA-imprinted polymer nanoparticles with monodispersity and prescribed DNA-strand patterns. *Nat. Chem.* **10**, 184-192 (2017).
281 Chidchob, P., Edwardson, T. G. W., Serpell, C. J. & Sleiman, H. F. Synergy of two assembly languages in DNA nanostructures: self-assembly of sequence-defined polymers on DNA cages. *J. Am. Chem. Soc.* **138**, 4416-4425 (2016).

282 Wang, W., Wan, W., Zhou, H.-H., Niu, S. & Li, A. D. Q. Alternating DNA and π-conjugated sequences. Thermophilic foldable polymers. *J. Am. Chem. Soc.* **125**, 5248-5249 (2003).
283 Vyborna, Y., Vybornyi, M. & Häner, R. From ribbons to networks: Hierarchical organization of DNA-grafted supramolecular polymers. *J. Am. Chem. Soc.* **137**, 14051-14054 (2015).
284 Yu, H., Alexander, D. T. L., Aschauer, U. & Häner, R. Synthesis of responsive two-dimensional polymers via self-assembled DNA networks. *Angew. Chem., Int. Ed.* **56**, 5040-5044 (2017).
285 Bösch, C. D., Jevric, J., Bürki, N., Probst, M., Langenegger, S. M. & Häner, R. Supramolecular assembly of DNA-phenanthrene conjugates into vesicles with light-harvesting properties. *Bioconjugate Chem.* **29**, 1505-1509 (2018).
286 Bousmail, D., Chidchob, P. & Sleiman, H. F. Cyanine-mediated DNA nanofiber growth with controlled dimensionality. *J. Am. Chem. Soc.* **140**, 9518-9530 (2018).
287 Marsh, T. C., Vesenka, J. & Henderson, E. A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy. *Nucleic Acids Res.* **23**, 696-700 (1995).
288 Dutta, K., Fujimoto, T., Inoue, M., Miyoshi, D. & Sugimoto, N. Development of new functional nanostructures consisting of both DNA duplex and quadruplex. *Chem. Commun.* **46**, 7772-7774 (2010).
289 Ghodke, H. B., Krishnan, R., Vignesh, K., Kumar, G. V. P., Narayana, C. & Krishnan, Y. The i-tetraplex building block: rational design and controlled fabrication of robust 1D DNA scaffolds through non-Watson–Crick interactions. *Angew. Chem., Int. Ed.* **46**, 2646-2649 (2007).
290 Yang, Y., Zhou, C., Zhang, T., Cheng, E., Yang, Z. & Liu, D. DNA pillars constructed from an i-motif stem and duplex branches. *Small* **8**, 552-556 (2012).
291 Cheng, E., Xing, Y., Chen, P., Yang, Y., Sun, Y., Zhou, D., Xu, L., Fan, Q. & Liu, D. A pH-triggered, fast-responding DNA hydrogel. *Angew. Chem.* **121**, 7796-7799 (2009).
292 Ren, J., Hu, Y., Lu, C.-H., Guo, W., Aleman-Garcia, M. A., Ricci, F. & Willner, I. pH-responsive and switchable triplex-based DNA hydrogels. *Chem. Sci.* **6**, 4190-4195 (2015).
293 Green, L. N., Amodio, A., Subramanian, H. K. K., Ricci, F. & Franco, E. pH-Driven reversible self-assembly of micron-scale DNA scaffolds. *Nano Lett.* **17**, 7283-7288 (2017).
294 Chandrasekaran, A. R. & Rusling, D. A. Triplex-forming oligonucleotides: a third strand for DNA nanotechnology. *Nucleic Acids Res.* **46**, 1021-1037 (2018).
295 Dong, Y., Yang, Z. & Liu, D. DNA nanotechnology based on i-motif structures. *Acc. Chem. Res.* **47**, 1853-1860 (2014).
296 Nakata, M., Zanchetta, G., Chapman, B. D., Jones, C. D., Cross, J. O., Pindak, R., Bellini, T. & Clark, N. A. End-to-end stacking and liquid crystal condensation of 6- to 20-base pair DNA duplexes. *Science* **318**, 1276-1279 (2007).
297 Maffeo, C., Luan, B. & Aksimentiev, A. End-to-end attraction of duplex DNA. *Nucleic Acids Res.* **40**, 3812-3821 (2012).
298 Kool, E. T. Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 1-22 (2001).
299 Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **34**, 564-574 (2006).

300 Protozanova, E., Yakovchuk, P. & Frank-Kamenetskii, M. D. Stacked-unstacked equilibrium at the nick site of DNA. *J. Mol. Biol.* **342**, 775-785 (2004).
301 Gerling, T., Wagenbauer, K. F., Neuner, A. M. & Dietz, H. Dynamic DNA devices and assemblies formed by shape-complementary, non–base pairing 3D components. *Science* **347**, 1446-1452 (2015).
302 Liu, L., Li, Y., Wang, Y., Zheng, J. & Mao, C. Regulating DNA self-assembly by DNA-surface interactions. *ChemBioChem* **18**, 2404-2407 (2017).
303 Wang, R., Kuzuya, A., Liu, W. & Seeman, N. C. Blunt-ended DNA stacking interactions in a 3-helix motif. *Chem. Commun.* **46**, 4905-4907 (2010).
304 Endo, M., Sugita, T., Katsuda, Y., Hidaka, K. & Sugiyama, H. Programmed-assembly system using DNA jigsaw pieces. *Chem. Eur. J.* **16**, 5362-5368 (2010).
305 Wang, Z.-G. & Ding, B. Engineering DNA self-assemblies as templates for functional nanostructures. *Acc. Chem. Res.* **47**, 1654-1662 (2014).
306 Hong, F., Zhang, F., Liu, Y. & Yan, H. DNA origami: Scaffolds for creating higher order structures. *Chem. Rev.* **117**, 12584-12640 (2017).
307 Rothemund, P. W. K., Papadakis, N. & Winfree, E. Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biol.* **2**, e424 (2004).
308 Evans, C. G. & Winfree, E. Physical principles for DNA tile self-assembly. *Chem. Soc. Rev.* **46**, 3808-3829 (2017).
309 Wang, T., Sha, R., Dreyfus, R., Leunissen, M. E., Maass, C., Pine, D. J., Chaikin, P. M. & Seeman, N. C. Self-replication of information-bearing nanoscale patterns. *Nature* **478**, 225-228 (2011).
310 Kim, J., Lee, J., Hamada, S., Murata, S. & Ha Park, S. Self-replication of DNA rings. *Nat. Nanotechnol.* **10**, 528-533 (2015).
311 Tikhomirov, G., Petersen, P. & Qian, L. Programmable disorder in random DNA tilings. *Nat. Nanotechnol.* **12**, 251-259 (2017).
312 Tikhomirov, G., Petersen, P. & Qian, L. Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns. *Nature* **552**, 67-71 (2017).
313 Lohse, S. E. & Murphy, C. J. Applications of colloidal inorganic nanoparticles: From medicine to energy. *J. Am. Chem. Soc.* **134**, 15607-15620 (2012).
314 Klinkova, A., Choueiri, R. M. & Kumacheva, E. Self-assembled plasmonic nanostructures. *Chem. Soc. Rev.* **43**, 3976-3991 (2014).
315 Liu, N. & Liedl, T. DNA-assembled advanced plasmonic architectures. *Chem. Rev.* **118**, 3032-3053 (2018).
316 Le, J. D., Pinto, Y., Seeman, N. C., Musier-Forsyth, K., Taton, T. A. & Kiehl, R. A. DNA-templated self-assembly of metallic nanocomponent arrays on a surface. *Nano Lett.* **4**, 2343-2347 (2004).
317 Sharma, J., Chhabra, R., Liu, Y., Ke, Y. & Yan, H. DNA-templated self-assembly of two-dimensional and periodical gold nanoparticle arrays. *Angew. Chem., Int. Ed.* **45**, 730-735 (2006).
318 Sharma, J., Ke, Y., Lin, C., Chhabra, R., Wang, Q., Nangreave, J., Liu, Y. & Yan, H. DNA-tile-directed self-assembly of quantum dots into two-dimensional nanopatterns. *Angew. Chem., Int. Ed.* **47**, 5157-5159 (2008).
319 Zheng, J., Constantinou, P. E., Micheel, C., Alivisatos, A. P., Kiehl, R. A. & Seeman, N. C. Two-dimensional nanoparticle arrays show the organizational power of robust DNA motifs. *Nano Lett.* **6**, 1502-1504 (2006).

320 Sharma, J., Chhabra, R., Cheng, A., Brownell, J., Liu, Y. & Yan, H. Control of self-assembly of DNA tubules through integration of gold nanoparticles. *Science* **323**, 112-116 (2009).
321 Kuzyk, A., Schreiber, R., Fan, Z., Pardatscher, G., Roller, E.-M., Högele, A., Simmel, F. C., Govorov, A. O. & Liedl, T. DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response. *Nature* **483**, 311-314 (2012).
322 Shen, X., Song, C., Wang, J., Shi, D., Wang, Z., Liu, N. & Ding, B. Rolling up gold nanoparticle-dressed DNA origami into three-dimensional plasmonic chiral nanostructures. *J. Am. Chem. Soc.* **134**, 146-149 (2012).
323 Lan, X., Lu, X., Shen, C., Ke, Y., Ni, W. & Wang, Q. Au nanorod helical superstructures with designed chirality. *J. Am. Chem. Soc.* **137**, 457-462 (2015).
324 Acuna, G. P., Möller, F. M., Holzmeister, P., Beater, S., Lalkens, B. & Tinnefeld, P. Fluorescence enhancement at docking sites of DNA-directed self-assembled nanoantennas. *Science* **338**, 506-510 (2012).
325 Schreiber, R., Do, J., Roller, E.-M., Zhang, T., Schüller, V. J., Nickels, P. C., Feldmann, J. & Liedl, T. Hierarchical assembly of metal nanoparticles, quantum dots and organic dyes using DNA origami scaffolds. *Nat. Nanotechnol.* **9**, 74-78 (2013).
326 Nykypanchuk, D., Maye, M. M., van der Lelie, D. & Gang, O. DNA-guided crystallization of colloidal nanoparticles. *Nature* **451**, 549 (2008).
327 Park, S. Y., Lytton-Jean, A. K. R., Lee, B., Weigand, S., Schatz, G. C. & Mirkin, C. A. DNA-programmable nanoparticle crystallization. *Nature* **451**, 553–556 (2008).
328 Sun, W., Boulais, E., Hakobyan, Y., Wang, W. L., Guan, A., Bathe, M. & Yin, P. Casting inorganic structures with DNA molds. *Science* **346**, 1258361 (2014).
329 Tokura, Y., Harvey, S., Xu, X., Chen, C., Morsbach, S., Wunderlich, K., Fytas, G., Wu, Y., Ng, D. Y. W. & Weil, T. Polymer tube nanoreactors by DNA-origami templated synthesis. *Chem. Commun.* **54**, 2808-2811 (2018).
330 Tokura, Y., Harvey, S., Chen, C., Wu, Y., Ng, D. Y. W. & Weil, T. Fabrication of defined polydopamine nanostructures by DNA origami-templated polymerization. *Angew. Chem., Int. Ed.* **57**, 1587-1591 (2018).
331 Liu, X., Zhang, F., Jing, X., Pan, M., Liu, P., Li, W., Zhu, B., Li, J., Chen, H., Wang, L., Lin, J., Liu, Y., Zhao, D., Yan, H. & Fan, C. Complex silica composite nanomaterials templated with DNA origami. *Nature* **559**, 593-598 (2018).
332 Knudsen, J. B., Liu, L., Bank Kodal, A. L., Madsen, M., Li, Q., Song, J., Woehrstein, J. B., Wickham, S. F. J., Strauss, M. T., Schueder, F., Vinther, J., Krissanaprasit, A., Gudnason, D., Smith, A. A. A., Ogaki, R., Zelikin, A. N., Besenbacher, F., Birkedal, V., Yin, P., Shih, W. M., Jungmann, R., Dong, M. & Gothelf, K. V. Routing of individual polymers in designed patterns. *Nat. Nanotechnol.* **10**, 892 (2015).
333 Maune, H. T., Han, S.-p., Barish, R. D., Bockrath, M., Iii, W. A. G., Rothemund, P. W. K. & Winfree, E. Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates. *Nat. Nanotechnol.* **5**, 61-66 (2009).
334 Edwardson, T. G. W., Lau, K. L., Bousmail, D., Serpell, C. J. & Sleiman, H. F. Transfer of molecular recognition information from DNA nanostructures to gold nanoparticles. *Nat. Chem.* **8**, 162-170 (2016).
335 Yang, Y. R., Liu, Y. & Yan, H. DNA nanostructures as programmable biomolecular scaffolds. *Bioconjugate Chem.* **26**, 1381-1395 (2015).

336 Selmi, D. N., Adamson, R. J., Attrill, H., Goddard, A. D., Gilbert, R. J. C., Watts, A. & Turberfield, A. J. DNA-templated protein arrays for single-molecule imaging. *Nano Lett.* **11**, 657-660 (2011).
337 Chhabra, R., Sharma, J., Ke, Y., Liu, Y., Rinker, S., Lindsay, S. & Yan, H. Spatially addressable multiprotein nanoarrays templated by aptamer-tagged DNA nanoarchitectures. *J. Am. Chem. Soc.* **129**, 10304-10305 (2007).
338 Malo, J., Mitchell, J. C., Vénien-Bryan, C., Harris, J. R., Wille, H., Sherratt, D. J. & Turberfield, A. J. Engineering a 2D protein–DNA crystal. *Angew. Chem., Int. Ed.* **44**, 3057-3061 (2005).
339 Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. & Séraphin, B. A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030-1032 (1999).
340 Wilner, O. I., Weizmann, Y., Gill, R., Lioubashevski, O., Freeman, R. & Willner, I. Enzyme cascades activated on topologically programmed DNA scaffolds. *Nat. Nanotechnol.* **4**, 249-254 (2009).
341 Fu, J., Yang, Y. R., Johnson-Buck, A., Liu, M., Liu, Y., Walter, N. G., Woodbury, N. W. & Yan, H. Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nat. Nanotechnol.* **9**, 531-536 (2014).
342 Fu, J., Yang, Y. R., Dhakal, S., Zhao, Z., Liu, M., Zhang, T., Walter, N. G. & Yan, H. Assembly of multienzyme complexes on DNA nanostructures. *Nat. Protoc.* **11**, 2243-2273 (2016).
343 Udomprasert, A., Bongiovanni, M. N., Sha, R., Sherman, W. B., Wang, T., Arora, P. S., Canary, J. W., Gras, S. L. & Seeman, N. C. Amyloid fibrils nucleated and organized by DNA origami constructions. *Nat. Nanotechnol.* **9**, 537-541 (2014).
344 Derr, N. D., Goodman, B. S., Jungmann, R., Leschziner, A. E., Shih, W. M. & Reck-Peterson, S. L. Tug-of-war in motor protein ensembles revealed with a programmable DNA origami scaffold. *Science* **338**, 662-665 (2012).
345 Praetorius, F. & Dietz, H. Self-assembly of genetically encoded DNA-protein hybrid nanoscale shapes. *Science* **355**, eaam5488 (2017).
346 Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nat. Protoc.* **12**, 1198-1228 (2017).
347 Lin, C., Jungmann, R., Leifer, A. M., Li, C., Levner, D., Church, G. M., Shih, W. M. & Yin, P. Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA. *Nat. Chem.* **4**, 832 (2012).
348 Jungmann, R., Avendaño, M. S., Woehrstein, J. B., Dai, M., Shih, W. M. & Yin, P. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat. Methods* **11**, 313 (2014).
349 Jungmann, R., Avendaño, M. S., Dai, M., Woehrstein, J. B., Agasti, S. S., Feiger, Z., Rodal, A. & Yin, P. Quantitative super-resolution imaging with qPAINT using transient binding analysis. *Nat. Methods* **13**, 439-442 (2016).
350 Hu, Q., Li, H., Wang, L., Gu, H. & Fan, C. DNA nanotechnology-enabled drug delivery systems. *Chem. Rev.* 10.1021/acs.chemrev.7b00663 (2018).
351 Boekhoven, J. & Stupp, S. I. 25th anniversary article: supramolecular materials for regenerative medicine. *Adv. Mater.* **26**, 1642-1659 (2014).

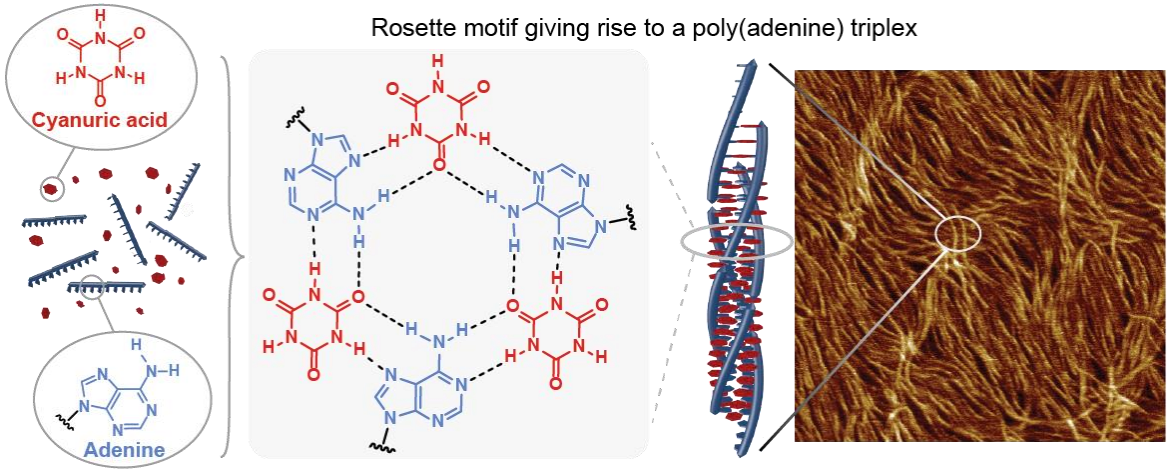
352 Stephanopoulos, N., Freeman, R., North, H. A., Sur, S., Jeong, S. J., Tantakitti, F., Kessler, J. A. & Stupp, S. I. Bioactive DNA-peptide nanotubes enhance the differentiation of neural stem cells into neurons. *Nano Lett.* **15**, 603-609 (2015).
353 Lund, K., Manzo, A. J., Dabby, N., Michelotti, N., Johnson-Buck, A., Nangreave, J., Taylor, S., Pei, R., Stojanovic, M. N., Walter, N. G., Winfree, E. & Yan, H. Molecular robots guided by prescriptive landscapes. *Nature* **465**, 206-210 (2010).
354 Gu, H., Chao, J., Xiao, S.-J. & Seeman, N. C. A proximity-based programmable DNA nanoscale assembly line. *Nature* **465**, 202-205 (2010).
355 Thubagere, A. J., Li, W., Johnson, R. F., Chen, Z., Doroudi, S., Lee, Y. L., Izatt, G., Wittman, S., Srinivas, N., Woods, D., Winfree, E. & Qian, L. A cargo-sorting DNA robot. *Science* **357**, eaan6558 (2017).
356 Kopperger, E., List, J., Madhira, S., Rothfischer, F., Lamb, D. C. & Simmel, F. C. A self-assembled nanoscale robotic arm controlled by electric fields. *Science* **359**, 296-301 (2018).
357 Lin, C., Ke, Y., Liu, Y., Mertig, M., Gu, J. & Yan, H. Functional DNA nanotube arrays: bottom-up meets top-down. *Angew. Chem., Int. Ed.* **119**, 6201-6204 (2007).
358 Kershner, R. J., Bozano, L. D., Micheel, C. M., Hung, A. M., Fornof, A. R., Cha, J. N., Rettner, C. T., Bersani, M., Frommer, J., Rothemund, P. W. K. & Wallraff, G. M. Placement and orientation of individual DNA shapes on lithographically patterned surfaces. *Nat. Nanotechnol.* **4**, 557-561 (2009).
359 Hung, A. M., Micheel, C. M., Bozano, L. D., Osterbur, L. W., Wallraff, G. M. & Cha, J. N. Large-area spatially ordered arrays of gold nanoparticles directed by lithographically confined DNA origami. *Nat. Nanotechnol.* **5**, 121-126 (2010).
360 Wasungu, L. & Hoekstra, D. Cationic lipids, lipoplexes and intracellular delivery of genes. *J. Controlled Release* **116**, 255-264 (2006).
361 Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. & Danielsen, M. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7413-7417 (1987).
362 Pozharski, E. & MacDonald, R. C. Thermodynamics of cationic lipid-DNA complex formation as studied by isothermal titration calorimetry. *Biophys. J.* **83**, 556-565 (2002).
363 Gromelski, S. & Brezesinski, G. DNA condensation and interaction with zwitterionic phospholipids mediated by divalent cations. *Langmuir* **22**, 6293-6301 (2006).
364 McManus, J. J., Rädler, J. O. & Dawson, K. A. Does calcium turn a zwitterionic lipid cationic? *J. Phys. Chem. B* **107**, 9869-9875 (2003).
365 Malghani, M. S. & Yang, J. Stable binding of DNA to zwitterionic lipid bilayers in aqueous solutions. *J. Phys. Chem. B* **102**, 8930-8933 (1998).
366 Kato, A., Tsuji, A., Yanagisawa, M., Saeki, D., Juni, K., Morimoto, Y. & Yoshikawa, K. Phase separation on a phospholipid membrane inducing a characteristic localization of DNA accompanied by its structural transition. *J. Phys. Chem. Lett.* **1**, 3391-3395 (2010).
367 Schade, M., Berti, D., Huster, D., Herrmann, A. & Arbuzova, A. Lipophilic nucleic acids — A flexible construction kit for organization and functionalization of surfaces. *Adv. Colloid Interface Sci.* **208**, 235-251 (2014).
368 Beales, P. A. & Vanderlick, T. K. Partitioning of membrane-anchored DNA between coexisting lipid phases. *J. Phys. Chem. B* **113**, 13678-13686 (2009).

369 Conway, J. W., Madwar, C., Edwardson, T. G., McLaughlin, C. K., Fahkoury, J., Lennox, R. B. & Sleiman, H. F. Dynamic behavior of DNA cages anchored on spherically supported lipid bilayers. *J. Am. Chem. Soc.* **136**, 12987-12997 (2014).
370 Banchelli, M., Betti, F., Berti, D., Caminati, G., Bombelli, F. B., Brown, T., Wilhelmsson, L. M., Nordén, B. & Baglioni, P. Phospholipid membranes decorated by cholesterol-based oligonucleotides as soft hybrid nanostructures. *J. Phys. Chem. B* **112**, 10942-10952 (2008).
371 Kurz, A., Bunge, A., Windeck, A.-K., Rost, M., Flasche, W., Arbuzova, A., Strohbach, D., Müller, S., Liebscher, J., Huster, D. & Herrmann, A. Lipid-anchored oligonucleotides for stable double-helix formation in distinct membrane domains. *Angew. Chem., Int. Ed.* **45**, 4440-4444 (2006).
372 Schade, M., Knoll, A., Vogel, A., Seitz, O., Liebscher, J., Huster, D., Herrmann, A. & Arbuzova, A. Remote control of lipophilic nucleic acids domain partitioning by DNA hybridization and enzymatic cleavage. *J. Am. Chem. Soc.* **134**, 20490-20497 (2012).
373 Börjesson, K., Lundberg, E. P., Woller, J. G., Nordén, B. & Albinsson, B. Soft-surface DNA nanotechnology: DNA constructs anchored and aligned to lipid membrane. *Angew. Chem., Int. Ed.* **50**, 8312-8315 (2011).
374 Rodríguez-Pulido, A., Kondrachuk, A. I., Prusty, D. K., Gao, J., Loi, M. A. & Herrmann, A. Light-triggered sequence-specific cargo release from DNA block copolymer lipid vesicles. *Angew. Chem., Int. Ed.* **52**, 1008-1012 (2013).
375 Lundberg, E. P., Feng, B., Saeid Mohammadi, A., Wilhelmsson, L. M. & Nordén, B. Controlling and monitoring orientation of DNA nanoconstructs on lipid surfaces. *Langmuir* **29**, 285-293 (2013).
376 Baldelli Bombelli, F., Betti, F., Gambinossi, F., Caminati, G., Brown, T., Baglioni, P. & Berti, D. Closed nanoconstructs assembled by step-by-step ss-DNA coupling assisted by phospholipid membranes. *Soft Matter* **5**, 1639-1645 (2009).
377 Loew, M., Springer, R., Scolari, S., Altenbrunn, F., Seitz, O., Liebscher, J., Huster, D., Herrmann, A. & Arbuzova, A. Lipid domain specific recruitment of lipophilic nucleic acids: A key for switchable functionalization of membranes. *J. Am. Chem. Soc.* **132**, 16066-16072 (2010).
378 Bunge, A., Kurz, A., Windeck, A.-K., Korte, T., Flasche, W., Liebscher, J., Herrmann, A. & Huster, D. Lipophilic oligonucleotides spontaneously insert into lipid membranes, bind complementary DNA strands, and sequester into lipid-disordered domains. *Langmuir* **23**, 4455-4464 (2007).
379 Peng, R., Wang, H., Lyu, Y., Xu, L., Liu, H., Kuai, H., Liu, Q. & Tan, W. Facile assembly/disassembly of DNA nanostructures anchored on cell-mimicking giant vesicles. *J. Am. Chem. Soc.* **139**, 12410-12413 (2017).
380 Bunge, A., Loew, M., Pescador, P., Arbuzova, A., Brodersen, N., Kang, J., Dähne, L., Liebscher, J., Herrmann, A., Stengel, G. & Huster, D. Lipid membranes carrying lipophilic cholesterol-based oligonucleotides - characterization and application on layer-by-layer coated particles. *J. Phys. Chem. B* **113**, 16425-16434 (2009).
381 Czogalla, A., Petrov, E. P., Kauert, D. J., Uzunova, V., Zhang, Y., Seidel, R. & Schwille, P. Switchable domain partitioning and diffusion of DNA origami rods on membranes. *Faraday Discuss.* **161**, 31-43 (2013).
382 List, J., Weber, M. & Simmel, F. C. Hydrophobic actuation of a DNA origami bilayer structure. *Angew. Chem., Int. Ed.* **53**, 4236-4239 (2014).

383 Suzuki, Y., Endo, M., Yang, Y. & Sugiyama, H. Dynamic assembly/disassembly processes of photoresponsive DNA origami nanostructures directly visualized on a lipid membrane surface. *J. Am. Chem. Soc.* **136**, 1714-1717 (2014).
384 Johnson-Buck, A., Jiang, S., Yan, H. & Walter, N. G. DNA–cholesterol barges as programmable membrane-exploring agents. *ACS Nano* **8**, 5641-5649 (2014).
385 Kocabey, S., Kempter, S., List, J., Xing, Y., Bae, W., Schiffels, D., Shih, W. M., Simmel, F. C. & Liedl, T. Membrane-assisted growth of DNA origami nanostructure arrays. *ACS Nano* **9**, 3530-3539 (2015).
386 Suzuki, Y., Endo, M. & Sugiyama, H. Lipid-bilayer-assisted two-dimensional self-assembly of DNA origami nanostructures. *Nat. Commun.* **6**, 8052 (2015).
387 Suzuki, Y., Sugiyama, H. & Endo, M. Complexing DNA origami frameworks through sequential self-assembly based on directed docking. *Angew. Chem., Int. Ed.* **57**, 7061-7065 (2018).
388 Sato, Y., Endo, M., Morita, M., Takinoue, M., Sugiyama, H., Murata, S., Nomura, S. M. & Suzuki, Y. Environment-dependent self-assembly of DNA origami lattices on phase-separated lipid membranes. *Adv. Mater. Interfaces* **5**, 1800437 (2018).
389 Dohno, C., Makishi, S., Nakatani, K. & Contera, S. Amphiphilic DNA tiles for controlled insertion and 2D assembly on fluid lipid membranes: the effect on mechanical properties. *Nanoscale* **9**, 3051-3058 (2017).
390 Stengel, G., Zahn, R. & Höök, F. DNA-induced programmable fusion of phospholipid vesicles. *J. Am. Chem. Soc.* **129**, 9584-9585 (2007).
391 van Lengerich, B., Rawle, Robert J., Bendix, Poul M. & Boxer, Steven G. Individual vesicle fusion events mediated by lipid-anchored DNA. *Biophys. J.* **105**, 409-419 (2013).
392 Xu, W., Nathwani, B., Lin, C., Wang, J., Karatekin, E., Pincet, F., Shih, W. & Rothman, J. E. A programmable DNA origami platform to organize SNAREs for membrane fusion. *J. Am. Chem. Soc.* **138**, 4439-4447 (2016).
393 Langecker, M., Arnaut, V., Martin, T. G., List, J., Renner, S., Mayer, M., Dietz, H. & Simmel, F. C. Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science* **338**, 932-936 (2012).
394 Burns, J. R., Göpfrich, K., Wood, J. W., Thacker, V. V., Stulz, E., Keyser, U. F. & Howorka, S. Lipid-bilayer-spanning DNA nanopores with a bifunctional porphyrin anchor. *Angew. Chem., Int. Ed.* **52**, 12069-12072 (2013).
395 Seifert, A., Göpfrich, K., Burns, J. R., Fertig, N., Keyser, U. F. & Howorka, S. Bilayer-spanning DNA nanopores with voltage-switching between open and closed state. *ACS Nano* **9**, 1117-1126 (2015).
396 Burns, J. R., Stulz, E. & Howorka, S. Self-assembled DNA nanopores that span lipid bilayers. *Nano Lett.* **13**, 2351-2356 (2013).
397 Burns, J. R., Seifert, A., Fertig, N. & Howorka, S. A biomimetic DNA-based channel for the ligand-controlled transport of charged molecular cargo across a biological membrane. *Nat. Nanotechnol.* **11**, 152-156 (2016).
398 Burns, J. R., Al-Juffali, N., Janes, S. M. & Howorka, S. Membrane-spanning DNA nanopores with cytotoxic effect. *Angew. Chem., Int. Ed.* **53**, 12466-12470 (2014).
399 Ohmann, A., Li, C.-Y., Maffeo, C., Al Nahas, K., Baumann, K. N., Göpfrich, K., Yoo, J., Keyser, U. F. & Aksimentiev, A. A synthetic enzyme built from DNA flips 107 lipids per second in biological membranes. *Nat. Commun.* **9**, 2426 (2018).

400 Fisher, P. D. E., Shen, Q., Akpinar, B., Davis, L. K., Chung, K. K. H., Baddeley, D., Šarić, A., Melia, T. J., Hoogenboom, B. W., Lin, C. & Lusk, C. P. A programmable DNA origami platform for organizing intrinsically disordered nucleoporins within nanopore confinement. *ACS Nano* **12**, 1508-1518 (2018).
401 Spruijt, E., Tusk, S. E. & Bayley, H. DNA scaffolds support stable and uniform peptide nanopores. *Nat. Nanotechnol.* **13**, 739-745 (2018).
402 Zhang, Z., Yang, Y., Pincet, F., C. Llaguno, M. & Lin, C. Placing and shaping liposomes with reconfigurable DNA nanocages. *Nat. Chem.* **9**, 653-659 (2017).
403 Czogalla, A., Kauert, D. J., Franquelim, H. G., Uzunova, V., Zhang, Y., Seidel, R. & Schwille, P. Amphipathic DNA origami nanoparticles to scaffold and deform lipid membrane vesicles. *Angew. Chem.* **127**, 6601-6605 (2015).
404 Franquelim, H. G., Khmelinskaia, A., Sobczak, J.-P., Dietz, H. & Schwille, P. Membrane sculpting by curved DNA origami scaffolds. *Nat. Commun.* **9**, 811 (2018).
405 Yang, Y., Wang, J., Shigematsu, H., Xu, W., Shih, W. M., Rothman, J. E. & Lin, C. Self-assembly of size-controlled liposomes on DNA nanotemplates. *Nat. Chem.* **8**, 476-483 (2016).
406 Grome, M. W., Zhang, Z., Pincet, F. & Lin, C. Vesicle tubulation with self-assembling DNA nanosprings. *Angew. Chem., Int. Ed.* **57**, 5330-5334 (2018).
407 Czogalla, A., Franquelim, Henri G. & Schwille, P. DNA nanostructures on membranes as tools for synthetic biology. *Biophys. J.* **110**, 1698-1707 (2016).
408 Pugh, G. C., Burns, J. R. & Howorka, S. Comparing proteins and nucleic acids for next-generation biomolecular engineering. *Nat. Rev. Chem.* **2**, 113-130 (2018).

2| Reprogramming the assembly of unmodified DNA with a small molecule



This chapter is mostly composed of work published as "Reprogramming the assembly of unmodified DNA with a small molecule" by N Avakyan, AA Greschner, F Aldaye, CJ Serpell, V Toader, A Petitjean & HF Sleiman; Nat. Chem. ***8****, 368-376 (2016). The short section on cyanuric acid analogs has not been previously published.*

Helical

Models adapted from: Origami DNA, Your Genome
(https://www.yourgenome.org/activities/origami-dna)

Folding, photography and image editing: Nicole Avakyan

Contributions

NA designed and performed all experiments unless listed below and analysed the results obtained. **Faisal Aldaye** and **Andrea A. Greschner** participated in the conception of the project, predating the work presented in this chapter. **AAG** also synthesized the 7-deazaadenosine-containing strand (section 2.4.18), the strands composing the intramolecular system (section 2.4.19), contributed to the design of several figures (by building figure elements using the program Blender) and provided insight over the course of the project. **Dr. Christopher J. Serpell** obtained the adenosine-cyanuric acid co-crystal, solved its X-ray diffraction pattern and contributed the figure showing rearrangement of the crystal tape structure (section 2.4.11 and Figure 2.8). **Dr. Violeta Toader** synthesized the TIPDS-dA and hex-CA derivatives (sections 2.4.12 and 2.4.13). **Prof. Anne Petitjean** performed the vapor pressure osmometry studies and analyzed the results obtained (section 2.4.14). **Dr. Maciej Barlog** synthesized TEG-CA and pr-CA (sections 2.4.22 and 2.4.23).

2.1 Introduction

The standard DNA base-pairing alphabet gives rise to both the familiar double helix and a variety of alternate assembly motifs such as the guanine quadruplex, i-motif and triple helix.[1,2] DNA's stability, remarkable recognition properties and diversity in tertiary structure are not only relevant to biological function. Three decades of DNA nanotechnology research have revealed its power as a building block in developing nanomaterials.[3] Considering the wealth of intermolecular interactions exploited in supramolecular chemistry, the DNA base-pairing code is limited to just two hydrogen bonding motifs that harness only four letters of a vast potential alphabet. Efforts to expand this genetic 'alphabet' have been driven by several goals. First, information about the principles behind the selectivity, function and evolution of DNA can be obtained by replacing its bases or backbone with artificial counterparts.[4-6] Second, selective incorporation of new properties relevant to materials science, synthetic biology, biotechnology and medicine can be achieved by using unnatural DNA analogs.[7-11] Finally, compared to the standard A-T/G-C duplex code, an expanded DNA alphabet grants access to increasingly complex DNA nanostructures, with potentially reduced defects.[12,13]

The creation of artificial nucleic acid motifs has relied on synthetic incorporation of nucleobases with altered hydrogen-bonding patterns,[14,15] hydrophobic interactions,[16,17] or metal coordination.[18-20] Synthetic incorporation of unnatural bases which present more than one hydrogen-bonding 'face' in order to access higher-order DNA structures has been of great interest.[21-25] 'Janus'[26] monomers (complementary to two nucleobases simultaneously) were attached to peptide nucleic acids (PNA) that associated into triplexes with the addition of two complementary DNA strands.[27] In another example, *iso*-guanine strands formed a DNA pentaplex in the presence of large alkali ions.[28,29] In the studies cited above, changing the DNA self-assembly code required synthetically laborious attachment of new monomers onto nucleoside sugar moieties, their incorporation into DNA strands and assembly of these new strands into altered structures. Intercalation of crescent-shaped alkaloids such as coralyne has been shown to induce the formation of an antiparallel duplex from poly(adenine) sequences,[30-34] and addition of π-extended molecules to guanine-rich strands can chaperone the formation of quadruplex structures.[35]

Adenine contains two hydrogen-bonding (Watson-Crick and Hoogsteen) faces, each with a two-point donor-acceptor motif. In the presence of poly(thymine), poly(adenine) can form double and triple helices (**Figure 2.1 a**). On its own, poly(adenine) assembles into a parallel homoduplex in acidic conditions,[36,37] whereas an antiparallel homoduplex can be formed at neutral pH with the addition of small molecule intercalators such as coralyne.[30,31] In principle, we can reprogram poly(adenine) interactions to induce assembly of a higher-order structure if we combine adenine with a small symmetric molecule such as cyanuric acid (CA), that contains three complementary thymine-like faces (**Figure 2.1 b**). CA and its analogs have been extensively used in supramolecular chemistry, most notably to form hydrogen-bonded tape and hexameric rosette architectures with melamine and its analogs.[38-40]

This chapter describes how the simple addition of CA can reprogram the assembly of unmodified poly(adenine) strands into nanofibers. Microscopy, spectroscopy and light scattering techniques are used to demonstrate that assemblies generated from poly(adenine) are indeed dependent on the presence of CA. Investigation into the structure of the assemblies supports the association of CA and adenine units into a proposed hydrogen-bonded hexameric rosette that gives rise to a DNA triplex. Additionally, the versatility of the assembly is tested by changing the poly(adenine) backbone to RNA and PNA, examining assembly with CA analogs and by

decorating nanofibers with streptavidin molecules. Fundamentally, this report illustrates that small hydrogen-bonding molecules can be used, in water, to induce the assembly of nucleic acids into new forms. This orthogonal assembly mode expands the 4-letter DNA alphabet and increases the range of possible architectures in DNA nanotechnology. In addition, the structures described here can be generated in large quantities from biocompatible components, allowing prospective applications in tissue engineering and drug delivery.

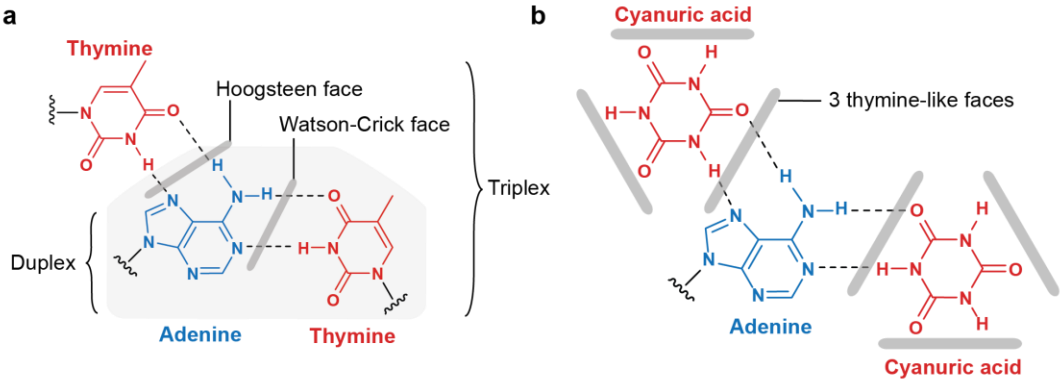


Figure 2.1| Adenine hydrogen bonding to thymine or cyanuric acid. a) The Watson-Crick and Hoogsteen binding faces of adenine can participate in the formation of duplex or triplex structures with thymine. **b)** Cyanuric acid, a small molecule with three thymine-like faces, could form a non-canonical DNA assembly motif by interacting with adenine.

2.2 Results and discussion

2.2.1 Evidence of CA-mediated assembly of d(A_{15})

CA-mediated assembly of poly(adenine) DNA (d(A_n)) was first examined by atomic force microscopy (AFM) in air. Deposition of a mixture of d(A_{15}), excess CA and sodium chloride (NaCl) in water led to the formation of fibers on the mica surface upon drying (**Figure 2.2 a**). Drying effects, such as variations in local concentrations of mixture components in different areas on the surface, affect the density of fibers observed and their appearance. However, in general, a network is formed from fibers of uniform 2.5 ± 0.1 nm height (**Figure 2.2 b**). In contrast to the assemblies obtained for the d(A_{15}) mixture with CA, no fiber formation was present in control experiments where d(A_{15}) or CA were deposited alone under matching pH and salt conditions (**Figure 2.2 c**).

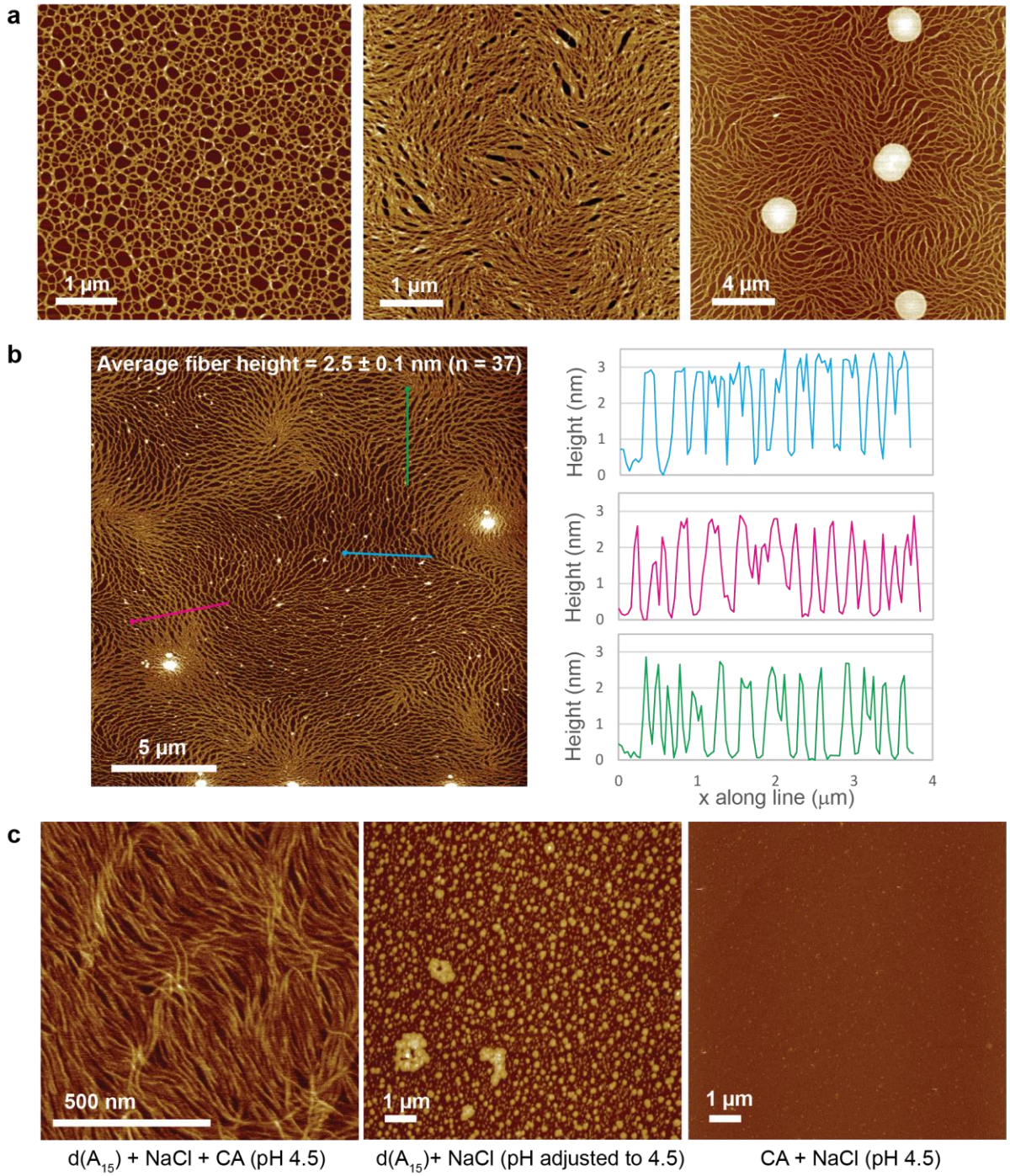


Figure 2.2| AFM topography images in air for CA-mediated assembly of d(A_{15}). a) Upon drying, a mixture of d(A_{15}), CA and NaCl yields extensive fiber networks. Drying effects influence the density and appearance of the features on the surface. **b)** Topography image and feature height traces used to calculate average fiber height. **c)** AFM height images in air for CA-mediated assembly of d(A_{15}) into nanofibers (left); neither d(A_{15}) (center) nor CA alone (right) show fiber growth under the same pH and salt conditions.

AFM imaging in fluid also revealed well-defined nanofibers sometimes reaching microns in length (height of 2.0 ± 0.1 nm), suggesting that CA-mediated assembly of d(A_{15}) is a solution-based phenomenon that does not depend on the infinitely high concentrations of mixture components that are achieved by drying it down on the mica surface (**Figure 2.3 a**). Furthermore, imaging in solution confirmed that no fiber assembly takes place in the absence of CA. Additional evidence for the role of CA in mediating fiber formation was acquired by dynamic light scattering (DLS). CA addition to d(A_{15}) brought on a marked increase in scatter intensity (**Figure 2.3 b**) representing the presence of larger particles in the sample, in keeping with the formation of long fibers observed by AFM. DLS measurements confirmed that aggregates freely diffusing in solution were non-spherical, as the mean hydrodynamic radius obtained varied with changing detection angle. Altogether, these findings indicate that using CA, a small molecule with binding features complementary to a natural DNA base, is a means to produce self-assembled nanofibers from simple, inexpensive components.

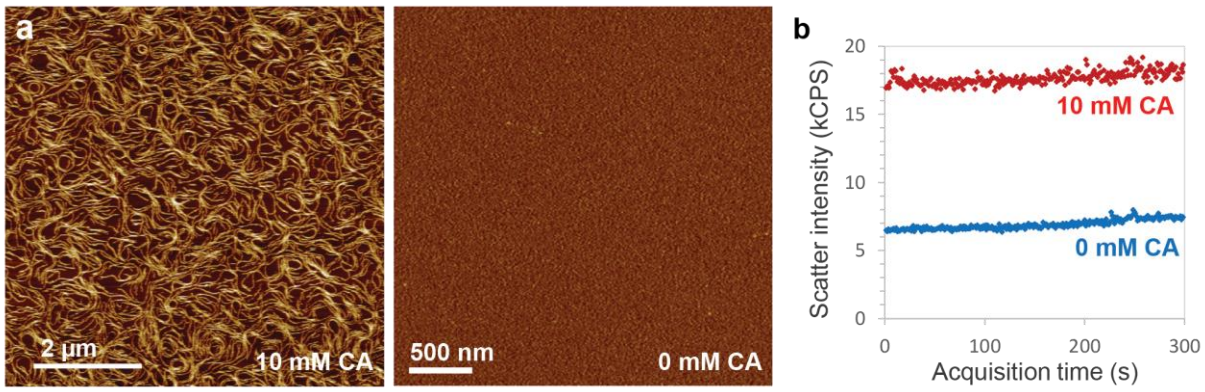


Figure 2.3| CA-mediated assembly of d(A_{15}) in solution. Extensive assembly in the presence of CA strongly suggests that both d(A_{15}) and CA are required for fiber formation. **a)** AFM imaging of 5 µM d(A_{15}) in fluid. Fiber formation in the presence of 10 mM CA (left); no fiber formation in the absence of CA (right). **b)** Scatter intensity signal from DLS for d(A_{15}) in the presence of 10 mM (red) or 0 mM (blue) CA.

The structural nature of these fibers was studied by circular dichroism (CD) and ultraviolet-visible (UV-Vis) absorbance spectroscopies. UV-Vis absorbance showed hypochromicity and a slight red shift of the DNA absorbance peak at 260 nm for d(A_{15}) in the presence of CA and NaCl (pH 4.5) (**Figure 2.4 a**). Furthermore, a rise in absorbance was detected in the 285 nm region, a feature associated with formation of J-aggregates as a result of ordered chromophore stacking.[41]

Under the same conditions, the CD spectrum for d(A_{15}) changed dramatically in the presence of CA, with the appearance of strong negative bands at 212 and 252 nm and of weak positive bands at 267 and 285 nm (**Figure 2.4 b**). The CD spectrum is distinct from that of a B-DNA duplex and bears similarities to the spectrum of a poly(A)-poly(T)-poly(T) triplex, with a slight red-shift.[42] Control CD and UV-Vis experiments showed that CA addition has no such effect on DNA strands of random sequence (**Figure 2.4 c-d**), an observation confirmed by AFM imaging in air (data not shown).

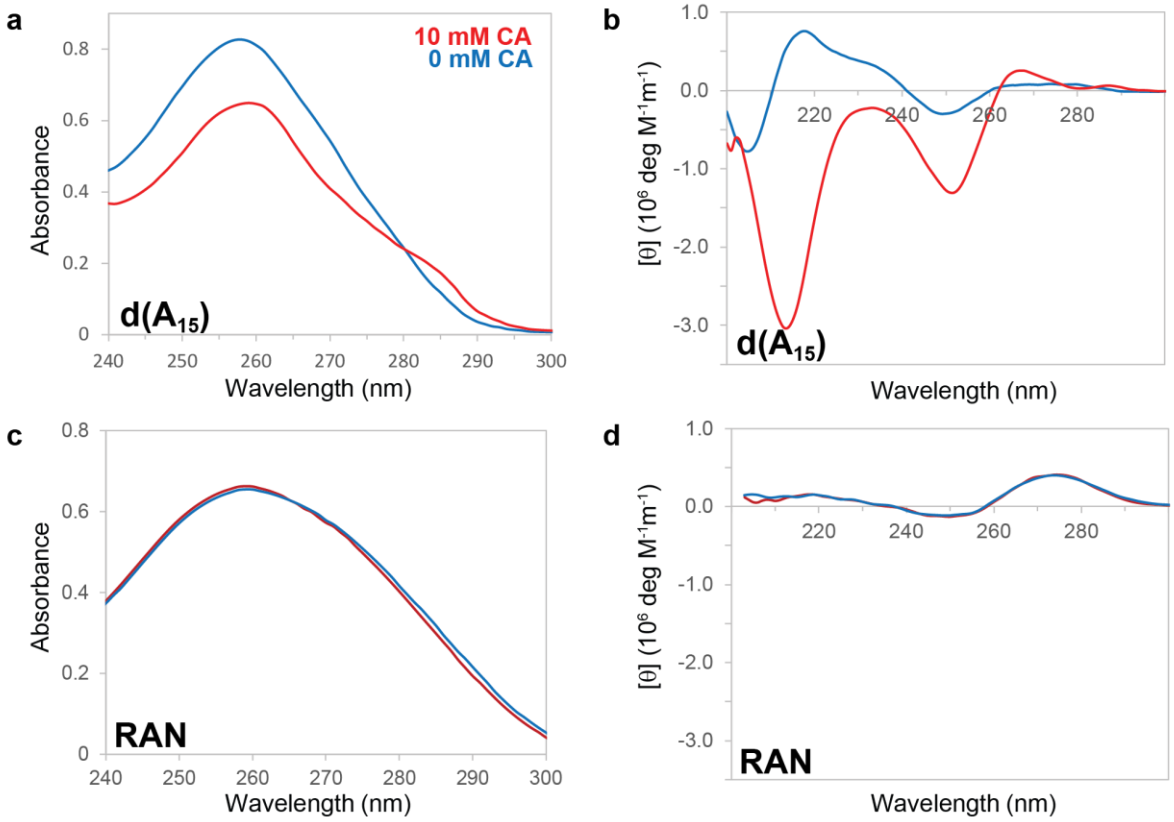


Figure 2.4| Spectra for d(A_{15}) and a random sequence control RAN in the presence of 10 mM (red) or 0 mM CA (blue). a) UV-Vis absorbance spectra for d(A_{15}), **b)** CD spectra for d(A_{15}), **c)** UV-Vis spectra for RAN and **d)** CD spectra for RAN.

Thermal denaturation of the CA-mediated assembly was monitored by CD, UV-Vis and DLS (**Figure 2.5 a-c**). The curves showed a well-defined, cooperative transition with a sigmoidal shape akin to the thermal denaturation of a DNA double helix. This observation is a sign of uniform assembly that does not go through stable intermediate stages during the melting process. Scatter

intensity for CA-mediated d(A_{15}) structures monitored by DLS similarly showed a sharp drop with increasing temperature, suggesting that large aggregates disassembled directly to their constituent components (**Figure 2.5 c**). This transition can also be qualitatively observed by AFM in solution while imaging at an ambient temperature of 29°C (**Figure 2.5 d**). With continuous imaging of a region of the sample at these conditions, fibers initially deposited on the surface disassembled over time due to instability at this temperature. Furthermore, a kinetic barrier to assembly was evident in the hysteresis detected between the melting and annealing curves obtained by CD (**Figure 2.5 e**), an observation often associated with cooperative self-assembly processes[43] that will be discussed in Chapter 3.

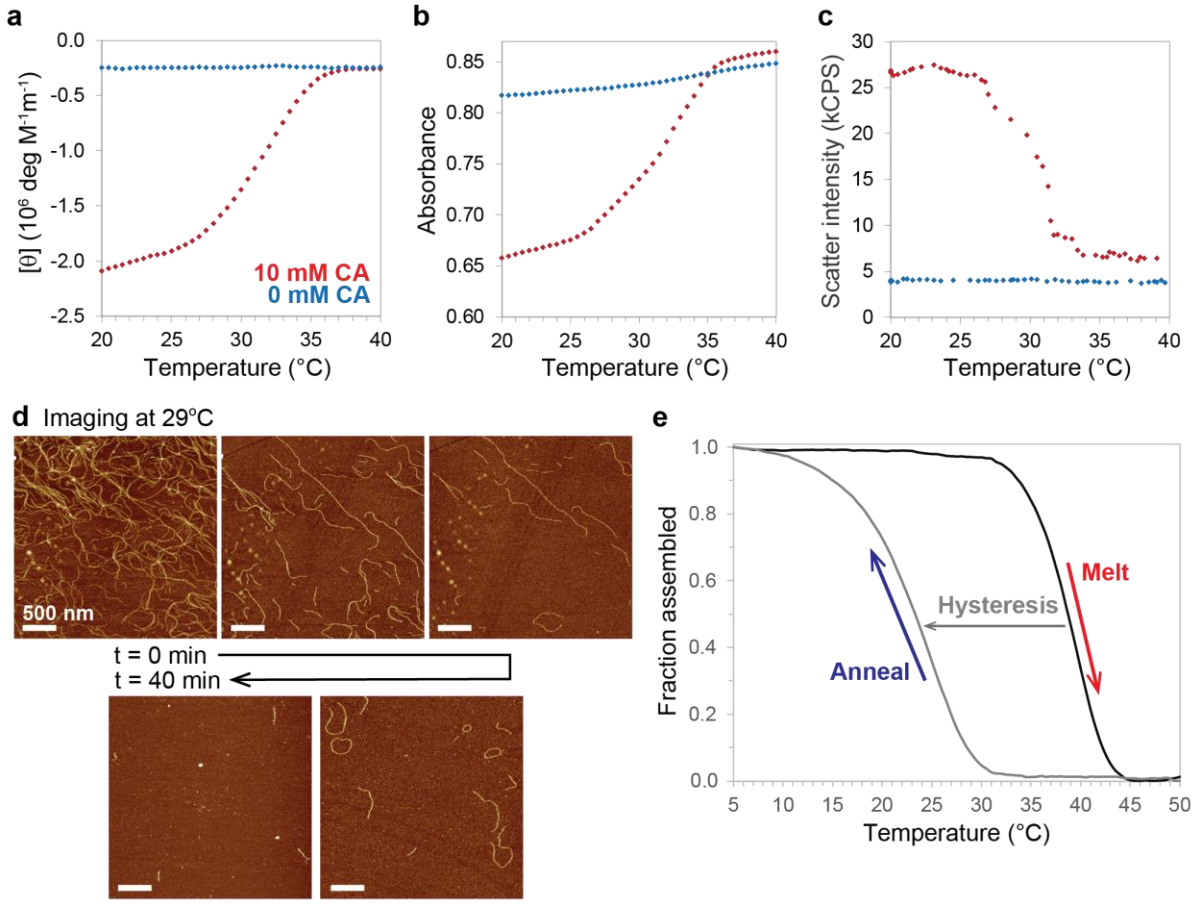


Figure 2.5| Thermal denaturation of d(A_{15}). Melting curves for d(A_{15}) in the presence of 10 mM (red) and 0 mM (blue) CA by **(a)** CD (monitored at 252 nm)**, (b)** UV-Vis absorbance (monitored at 260 nm) and **(c)** DLS. **d)** Qualitative representation of fiber melting followed by AFM in solution. Imaging at an ambient temperature of 29 °C, scale bar: 500 nm. **e)** Normalized melting and annealing curves for d(A_{15}) in the presence of 18 mM CA obtained by CD (monitored at 252 nm).

The same CD features characteristic of assembly were obtained in the presence of Mg^{2+} rather than Na^{+} and at lower ion concentrations. Just as in the case of standard DNA hybridization, where cations are required to counter the dense negative charge of the DNA phosphate backbone, counterions are essential for the formation of CA-mediated d(A_{15}) assemblies, with divalent cations having a more potent stabilizing effect than monovalent ones.[44-46] Compared to Na^{+}, much lower concentrations of Mg^{2+} ions were required to observe the CD spectrum characteristic of CA-mediated assembly of d(A_{15}) (**Figure 2.6 a-b**). No assembly was reported below 20 mM Na^{+} whereas with Mg^{2+} concentrations as low as 0.25 mM, assembly still occurred, as monitored by CD and UV-Vis absorbance. Above 20 mM magnesium, the onset of sample gelation was apparent with the naked eye and light scattering from the formation of large aggregates distorted the CD signal. The nature and concentration of counterions thus have an important effect on the extent of assembly. As expected, the T_M (melting temperature) values for CA-mediated assemblies of d(A_{15}) increase with increasing counterion concentration (**Figure 2.6 c**). For further experiments, counterion concentration was maintained at 7.6 mM Mg^{2+}.

Under highly acidic conditions (pH 3), poly(adenine) strands (both RNA and DNA) self-associate into parallel homoduplexes as a result of protonation of the adenine units.[36,37,47,48] Based on major differences in the CD spectra obtained for the poly(adenine) duplex formed in acidic conditions and assemblies obtained in the presence of CA, it was ascertained that CA-mediated assembly of d(A_{15}) produces a distinct species. At pH 3, the CD spectrum of d(A_{15}), both with CA added or in the absence of the small molecule, was dominated by the parallel poly(dA) duplex signal, whereas at pH 8 the spectrum indicated the presence of single-stranded d(A_{15}) (**Figure 2.7 a-b**). From pH 4 to 7, the spectra obtained were characteristic of CA-mediated assembly (**Figure 2.7 a**). In this range, both adenine bases (pK_{aH} ~3.6)[49] and CA (pK_a ~6.88)[49] are expected to be mostly uncharged, suggesting that uncharged components are necessary for the formation of the higher-order structures. Spectra of d(A_{15}) samples at pH 4 to pH 8 in the absence CA resembled the spectrum for d(A_{15}) in the presence of CA at pH 8 (**Figure 2.7 a**, blue trace), indicating no assembly.

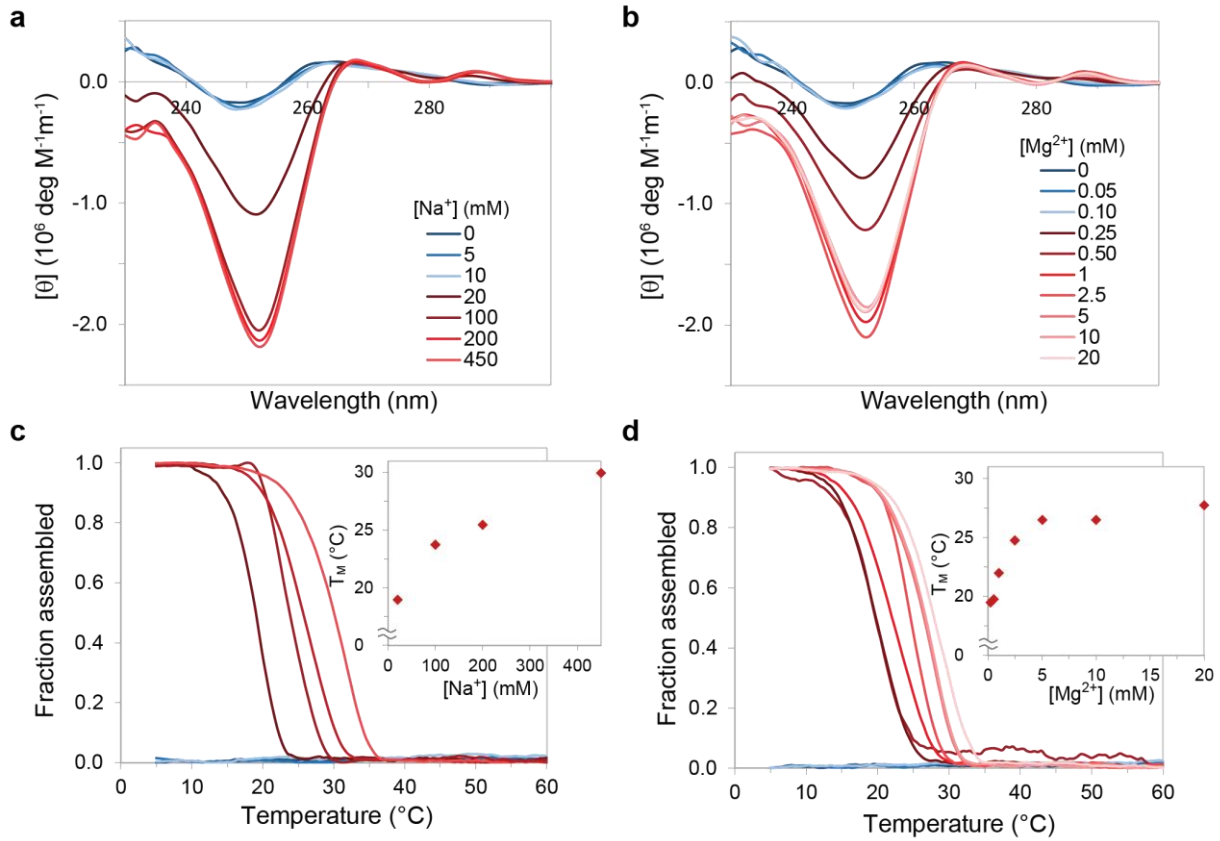


Figure 2.6| Effect of counterions on CA-mediated assembly of d(A_{15}) (10mM CA). a) CD spectra for d(A_{15}) in a range of Na^+ concentrations (0, 5, 10, 20, 100, 200 and 450 mM). **b)** CD spectra for d(A_{15}) in a range of Mg^{2+} concentrations (0, 0.05, 0.10, 0.25, 0.50, 1.00, 2.50, 5.00, 10.00 and 20.00 mM). **c-d)** Normalized thermal denaturation curves for d(A_{15}) in a range of Na^+ **(c)** and Mg^{2+} **(d)** concentrations (as listed in **(a)** and **(b)**). Insets show the trend of T_M dependence on counterion concentration. Traces in shades of red are attributed to samples showing CA-mediated assembly, whereas traces in shades of blue appear to show no assembly.

It is possible to draw parallels between the T_M dependence on CA concentration and subtle changes in stability dependent on sample pH (**Figure 2.7 c**). From pH 4 to 7, where a CD spectrum associated with CA-mediated assembly was observed, thermal denaturation studies following the CD signal yielded similar T_M values indicating comparable stability, such that change in pH did not alter the assembly process. The variation in T_M values obtained in this pH range seems to stem from the concentration of fully protonated CA available for assembly in the samples. For samples at pH 5 and 6, the protonated CA concentration varied from 9.8 to 9.3 mM (based on CA pK_a equal to 6.88 and equilibrium calculations from experimentally determined sample pH) and samples had equivalent T_M values, whereas for the pH 7 sample, the concentration dropped down to 6.3 mM

and was accompanied by a 3°C drop in T_M. The T_M value also dropped for the pH 4 sample, possibly due to a larger extent of adenine protonation disrupting CA-mediated assembly.

Overall, AFM imaging, DLS, CD and UV-Vis absorbance results reveal that d(A_n) assembly is mediated by CA in a unique fashion, giving rise to a structure with features distinct from the previously reported poly(A) duplexes that form in acidic conditions.

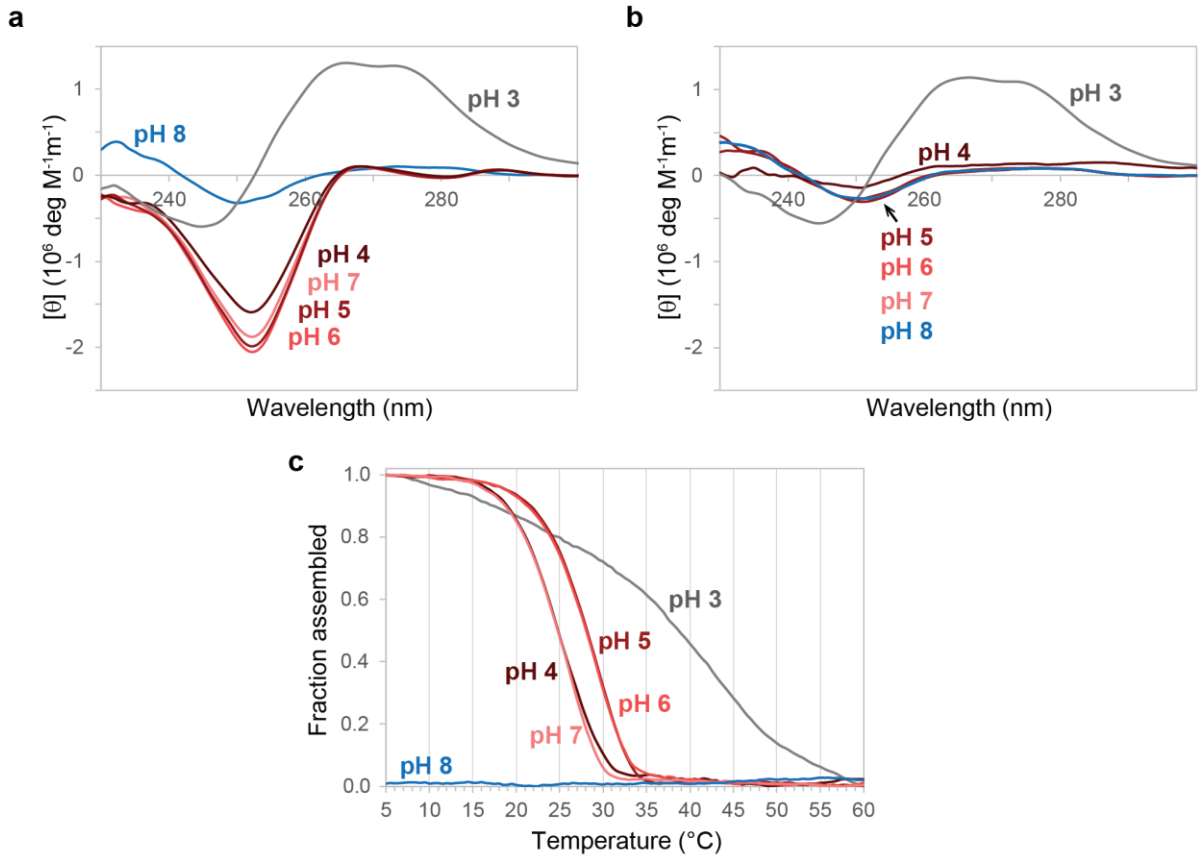


Figure 2.7| Effect of pH on CA-mediated assembly of d(A_{15}). CD spectra for d(A_{15}) in the presence of 10 mM **(a)** and 0 mM **(b)** CA in a pH range (pH 3, 4, 5, 6, 7 and 8). **c)** Thermal denaturation curves for d(A_{15}) in the presence of 10 mM CA in the same pH range, obtained by CD (monitored at 266 nm for pH 3 and at 252 nm for all other samples). pH is controlled by a Tris-acetate magnesium (TAMg) buffer; pH 3 (grey traces), pH 8 (blue traces), pH 4 – 7 (traces in shades of red).

2.2.2 Investigation of fiber structure

The ability of a small molecule to mediate hydrogen-bonded interactions in water may seem counterintuitive, but is not unprecedented.[50] Hud *et al.* demonstrated self-assembly of CA with hydrophilic derivatives of the hydrogen bonding molecule 2,4,6-triaminopyrimidine (TAP), a melamine analog, in water, including a monomeric TAP-ribose derivative.[51] These mixtures produced hexameric rosettes that stacked into long fibers with a similar diameter to the d(A_n):CA structure described here (as measured from AFM imaging and predictive calculations). Comparable to our studies, Hud *et al.* observed the absence of intermediates in the assembly process and ascribed this property to the large hydrophobic surface of the rosette (> 1 nm^2) that imposes a high energetic cost (+ 27 kcal/mol) for its exposure to the aqueous solvent in intermediate structures. Furthermore, in the d(A_n):CA system, the adenine units are tethered by a phosphodiester backbone. This results in significant preorganization in the d(A_n) single strands,[36] which is expected to reduce the entropic cost of assembly.

In contrast to melamine and TAP, adenine presents its hydrogen bond donor and acceptor groups on an asymmetric scaffold. As an initial step towards examining the geometry of interaction between adenine and CA, crystallization of an adenosine:CA complex was pursued. A single crystal of the 1:1 adenosine:CA complex was obtained from an equimolar mixture and the X-ray diffraction pattern was solved to reveal an alternating adenosine:CA tape structure (**Figure 2.8**). The experimentally determined structure, representative of the angles and distances between the interacting units, was used as a starting point for rudimentary geometrical modelling. By rotating adenosine locations by 120° around CA (i.e. moving them from one face to the next chemically identical face of CA), the nucleosides became perfectly aligned for formation of a hexameric rosette (**Figure 2.8 b**). This is consistent with the tendency of CA to form hexameric rosettes when assembled with melamine as well as TAP derivatives.[6,39,40,51] The hydrogen bonding patterns and geometries in this model are preserved on a local scale, with only the symmetry operations having been modified. No other rosette structure can be constructed from these data. In this model the sugar units are perpendicular to the plane of the rosette, providing the foundations for a phosphate backbone, thus not prohibiting rosettes from stacking upon each other, if they were connected to a phosphate backbone as in an oligonucleotide.

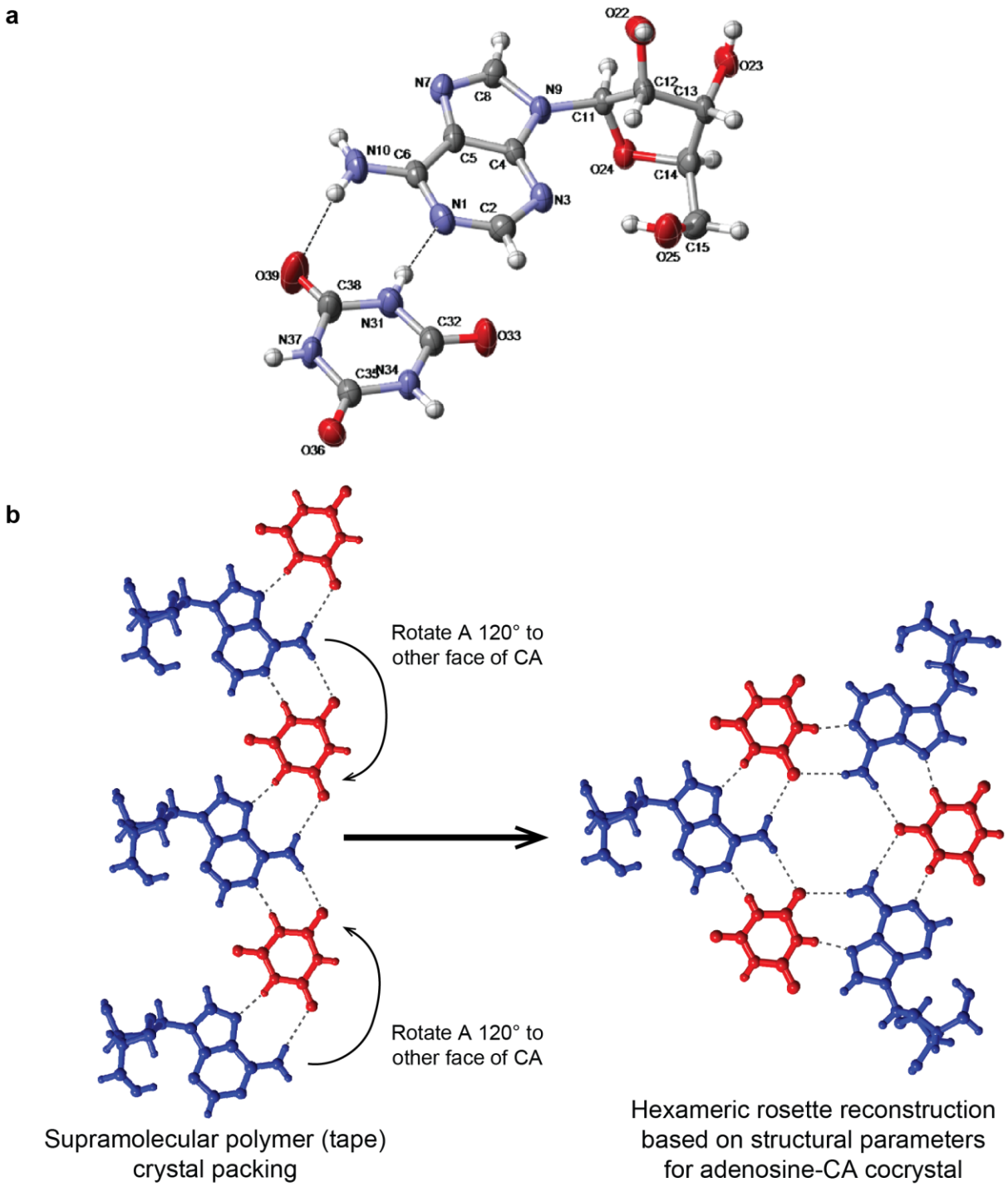


Figure 2.8| X-ray structure of adenosine-cyanuric acid co-crystal. a) Asymmetric unit of the 1:1 CA:A crystal structure. Thermal ellipsoids shown at 50% probability. **b)** The geometric and hydrogen bonding requirements of the tape crystal components are also consistent with a hexameric CA:A rosette.

To verify that monomeric adenine is capable of forming rosette structures with CA in solution, we prepared deoxy-tetraisopropyldisiloxane-adenosine (TIPDS-dA) and hexyl-CA (Hex-CA) derivatives and tested their assembly in toluene (**Figure 2.9 a**).[52] The parent molecules dA and CA were modified with hydrophobic substituents for improved solubility in organic solvents. The bulky TIPDS group was appended to dA to favour formation of a closed structure, rather than oligomeric ribbons, upon assembly with Hex-CA.

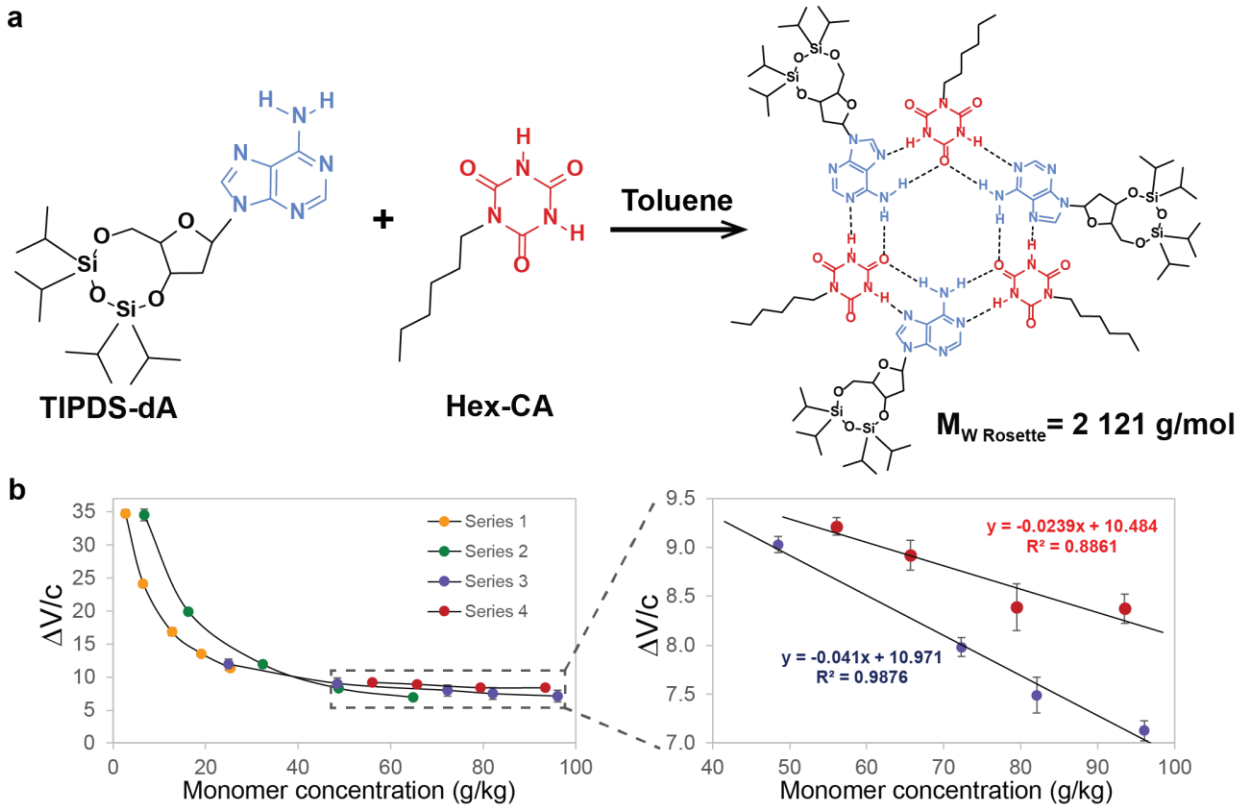


Figure 2.9| Assembly of TIPDS-dA and Hex-CA in toluene. a) Structure of TIPDS-dA and Hex-CA and illustration of hexameric rosette structure that could form in a 1:1 mixture of these molecules. The calculated M_w of the rosette would be 2121 g/mol. **b)** Measured ΔV/c vs. c curves for sample Series 1-4 (left). Plot on the right shows a magnification of the Series 3 and 4 high concentration range with corresponding linear regression fits and associated equations.

Vapor pressure osmometry (VPO) was used to study the products of interaction of a 1:1 TIPDS-dA:Hex-CA mixture in a range of concentrations. The ΔV/c *vs*. c (where ΔV is the VPO response and c is concentration in g/kg) curves for the four series of measurements show

concentration-dependent self-assembly, with strong curvature at lower concentrations, suggesting a self-associating system (**Figure 2.9 b**, left). At higher concentrations (50-100 g/kg, equivalent to 60-120 mM), the VPO response was fairly linear (**Figure 2.9 b**, right), which may suggest the formation of a well-defined species. However, at these concentrations, the VPO response is sensitive to drifting, such that the formation of oligomers of different sizes cannot be unambiguously ruled out. In addition, long equilibration times (up to 15 minutes) were required at higher concentrations, contributing to uncertainty on the absolute value of the signal. The molecular weight (M_W) ranges calculated from Series 3 and Series 4 values were of 1 950 - 2 050 g/mol and 2 040 - 2 100 g/mol respectively, based on the benzil and PS2500 standards. An average M_W of 2 050 ± 150 g/mol can be estimated (VPO commonly carries an intrinsic 10% error in M_W determination), which is well in line with the 2 121 g/mol M_W expected for a 1:1 TIPDS-dA:Hex-CA hexameric rosette.

Furthermore, the 1:1 TIPDS-dA:Hex-CA mixture was investigated by variable temperature ^{1}H NMR spectroscopy (**Figure 2.10**). The spectrum for the 1:1 mixture (100 mM concentration in monomers) at 25 °C showed broad signals for the adenosine protons and the N-CH_2 protons of Hex-CA (labelled α), with the NH protons of Hex-CA as the most downfield signals detected (**Figure 2.10 b**). With increasing temperature, signals became sharper and the NH protons of Hex-CA as well as NH_2 protons of TIPDS-dA moved upfield (NH: 13.60 to 11.82 ppm; NH_2: 7.44 to 6.82 ppm). In contrast, the differences in chemical shifts observed for TIPDS-dA alone in this temperature range were much smaller, except for the NH_2 signal (5.44 to 4.87 ppm) (**Figure 2.10 c**). These observations suggest hydrogen bonding between TIPDS-dA and Hex-CA in the 1:1 mixture is disrupted at high temperatures. Due to the broad peak shape in the spectra for the 1:1 mixture, it is not possible to unambiguously conclude whether a single closed structure or a mixture of oligomers are present in solution.

Diffusion ordered NMR spectroscopy (DOSY) was used to address this uncertainty. The DOSY spectrum of an analogous 1:1 TIPDS-dA:Hex-CA mixture showed the formation of a single slow-diffusing species with a diffusion coefficient $D_{Mix} = 2.74 \times 10^{-10}$ m^2/s, rather than a mixture of oligomers with different diffusion coefficients (**Figure 2.11 a**). The only other species identified in the sample is the solvent toluene. The diffusion coefficient for the TIPDS-dA monomer at the same concentration was $D_{TIPDS\text{-}dA} = 4.8 \times 10^{-10}$ m^2/s (**Figure 2.11 b**).

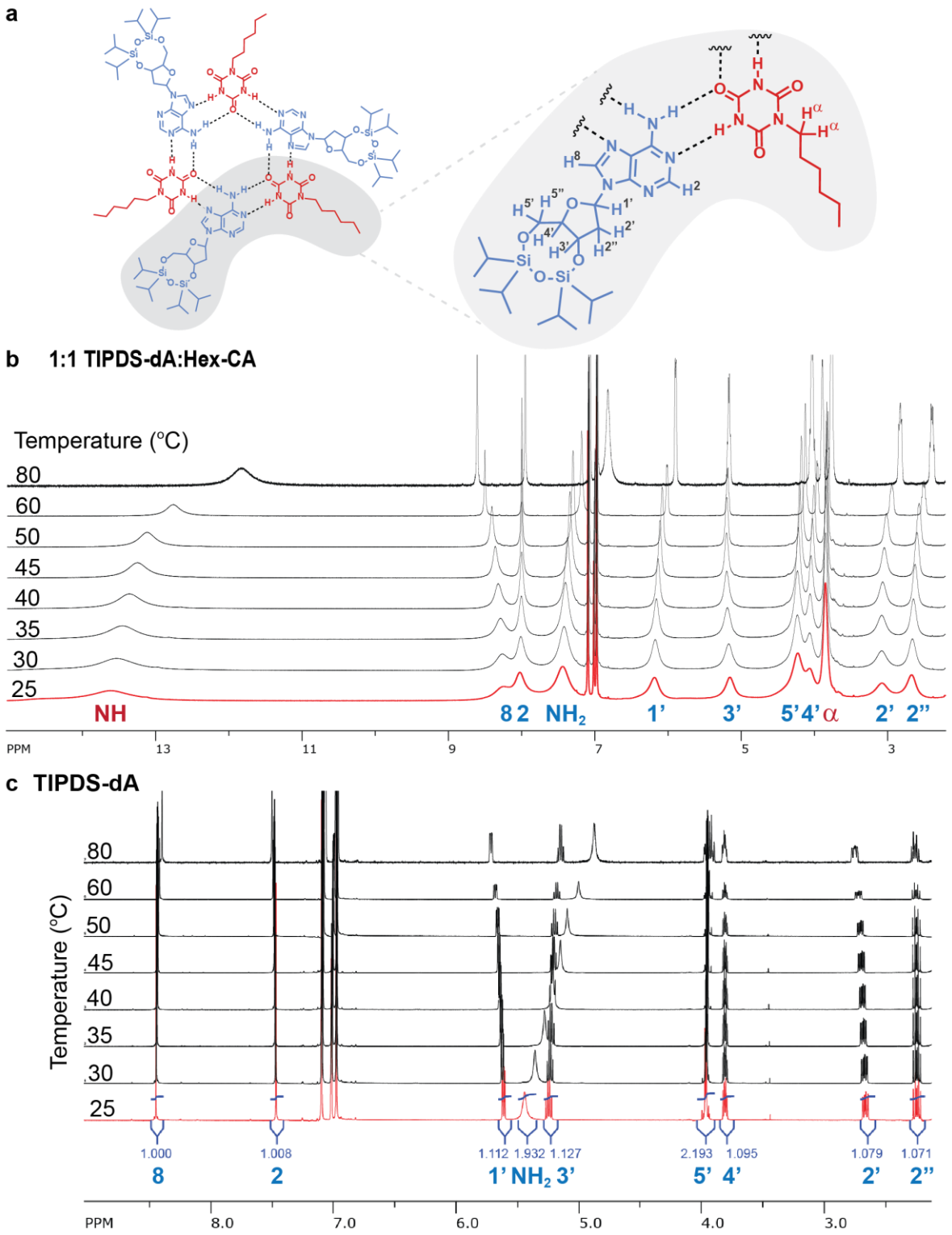


Figure 2.10| Variable temperature ^{1}H NMR in toluene-d$_8$ (temperature values: 25, 30, 35, 40, 45, 50, 60 and 80 °C). a) Schematic representation of 1:1 TIPDS-dA:Hex-CA hexameric rosette. Blue: TIPDS-dA; red: Hex-CA (left). Zoom in of an asymmetric TIPDS-dA:Hex-CA unit shows proton labelling legend (right). **b)** 1:1 TIPDS-dA:Hex-CA, **c)** TIPDS-dA alone.

Equation 2.1[53] was used to estimate the M_W of the rosette based on known quantities: $M_{W\ TIPDS\text{-}dA}$ = 493.75 g/mol, $D_{TIPDS\text{-}dA}$ and D_{Mix}. Based on this equation, an estimated molecular weight range for the species in the mixture is between 1 520 and 2 665 g/mol. Given these observations and the VPO results, reporting a M_W consistent with the rosette structure, it is possible to conclude that a hexameric hydrogen bonded rosette forms in a 1:1 TIPDS-dA:Hex-CA mixture.

$$\sqrt[3]{\frac{M_{W\ TIPDS-dA}}{M_{W\ Mix}}} \leq \frac{D_{Mix}}{D_{TIPDS-dA}} \leq \sqrt{\frac{M_{W\ TIPDS-dA}}{M_{W\ Mix}}} \qquad \text{(Equation 2.1)}$$

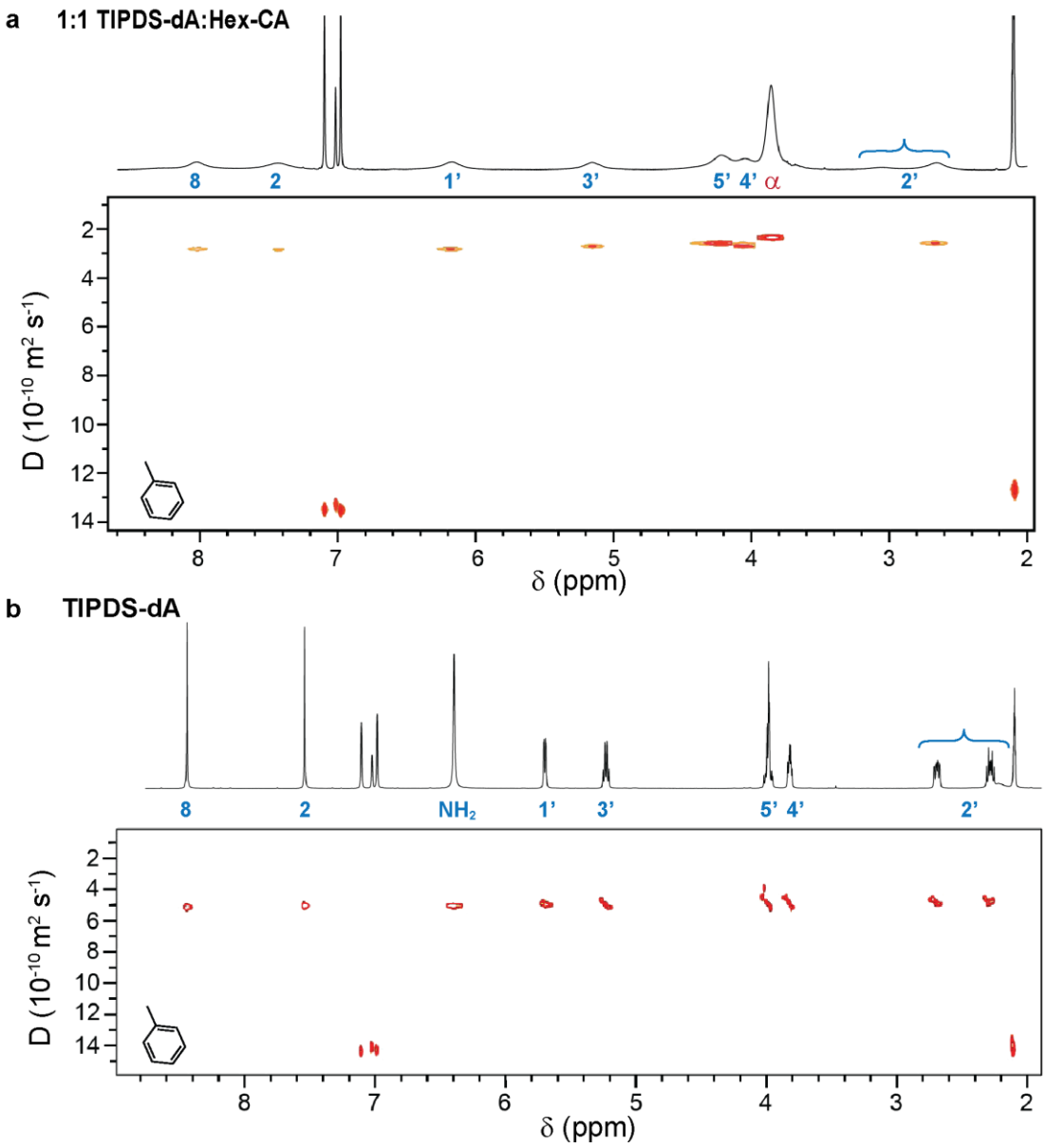


Figure 2.11| DOSY spectra in toluene-d8. a) 1:1 mixture of TIPDS-dA:Hex-CA, **b)** TIPDS-dA alone. Proton labelling as in the scheme illustrated in **Figure 2.10 a**.

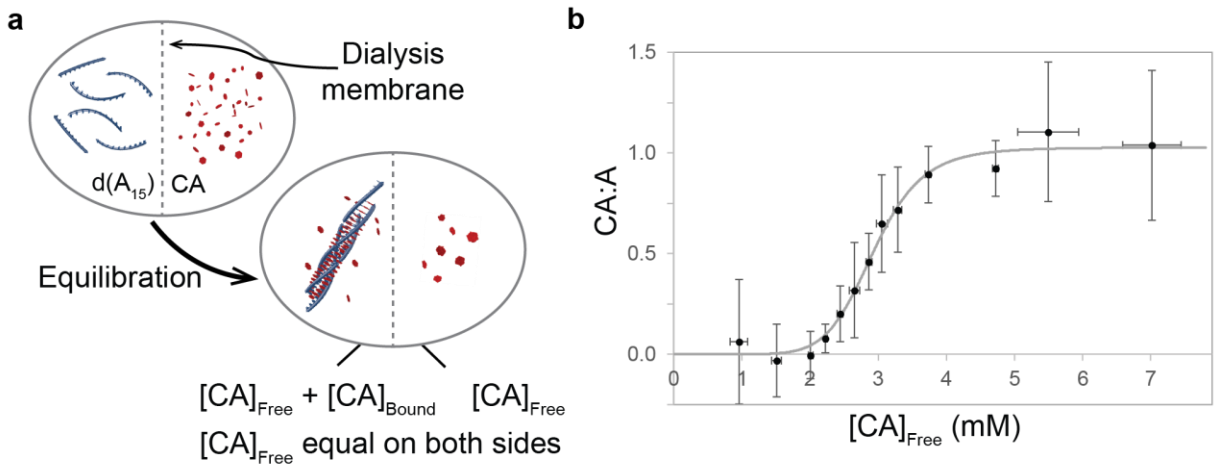


Figure 2.12| Determination of CA:A stoichiometry by equilibrium dialysis. a) The equilibrium dialysis unit has two chambers separated by a dialysis membrane (dotted line). Initially, the sample side contains free d(A_{15}) and the buffer side contains CA solution (in a range of concentrations). With time, CA equilibrates across the membrane and interacts with d(A_{15}) on the sample side. At equilibrium, $[CA]_{Free}$ is equal on both sides, while the sample side contains an additional $[CA]_{Bound}$. **b)** Binding curve displaying CA:A stoichiometry *vs.* $[CA]_{Free}$ within the CA mediated d(A_{15}) assembly (n = 12, 95 % CL).

Having demonstrated that modified adenosine and CA monomers are capable of forming hexameric rosettes in an organic medium, the native poly(A):CA assembly was re-examined in aqueous solution. In order to determine CA:A stoichiometry in the CA:d(A_n) structure, equilibrium dialysis experiments were carried out.[54,55] In this method, a macromolecule and small molecule ligand are separated by a dialysis membrane, permeable only to the small molecule in two chambers of equal volume. In this particular case, the sample chamber initially contained d(A_{15}), and the concentration of CA in the buffer chamber was varied. (**Figure 2.12 a**). With time, CA diffused across the membrane and assembled with d(A_{15}). Once assembly and equilibration were complete, the $[CA]_{Free}$ was the same on both sides of the membrane, with the sample chamber containing an additional amount of CA that participated in assembly with d(A_{15}) as $[CA]_{Bound}$. The contents of both chambers were analyzed by high performance liquid chromatography (HPLC) and the integrated area under the CA peak was used to calculate $[CA]_{Free}$ and $[CA]_{Bound}$ which were then used to determine CA:A stoichiometry and build a binding curve (**Figure 2.12 b**). The curve revealed CA:A saturation at 1, consistent with a 1:1 CA:A ratio in the structure. A complication of the equilibrium dialysis experiment is that CA binding to d(A_{15}) is coupled to a structural change

of the single-stranded DNA into a new assembly, as well as its elongation into fibers.[33] This experiment is different from binding studies of small molecules with double stranded DNA, in which the duplex structure is relatively maintained throughout the dialysis experiment.[56] The error bars associated with the data points (despite 12 repetitions of the experiment) are likely the result of these complications, and also arise from precipitation of growing assemblies and adhesion of mixture components to surfaces of the dialysis chamber.

Additionally, confirmation that both the Watson-Crick and Hoogsteen faces of adenine participate in CA-mediated assembly in solution, and thus, that a closed 1:1 CA:A structure can form, was pursued. An oligonucleotide, d(deaza-A_{12}), composed of adenine analog 7-deazaadenine bases, which lack the hydrogen bonding N^7 atom, thus eliminating the Hoogsteen binding face of adenine, was synthesized. The d(deaza-A_{12}) strand, incubated with 15 mM CA (TAMg pH 4.5 and 6), did not display the change in CD signal characteristic with CA-mediated assembly. At pH 4.5 (below the 5.3 pK_a of 7-deazaadenine) the spectrum remained unchanged, whereas at pH 6, it showed features reminiscent of Z-DNA (**Figure 2.13**).[42] In both pH conditions, the spectral features were distinct from those observed for CA-mediated assembly of d(A_n) suggesting that both hydrogen bonding faces of adenine are indeed required for CA-mediated assembly.

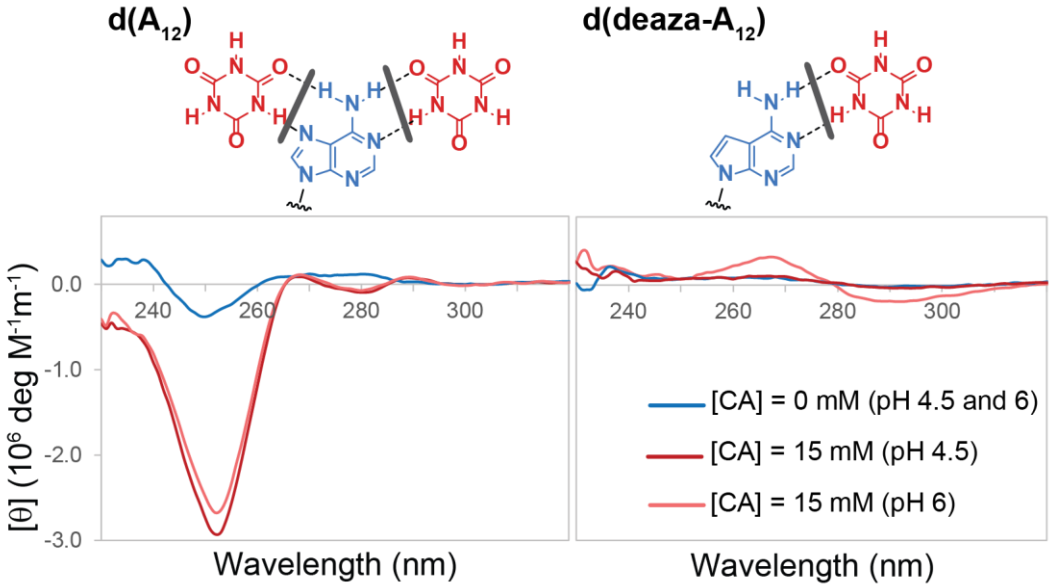


Figure 2.13| CD spectra for d(A_{12}) (left) *vs.* d(deaza-A_{12}) (right). Spectra obtained at pH 4.5 and pH 6, either with 0 mM CA or 15 mM CA.

Based on the VPO and DOSY studies with modified monomeric adenosine and CA units in toluene, the 1:1 CA:A ratio in the CA-mediated structure of d(A_{15}) in aqueous solution determined by equilibrium dialysis, the participation of both the Watson-Crick and Hoogsteen faces of adenine in CA-mediated assembly and AFM height measurements, a hexameric CA:A rosette which leads to the growth of d(A_n) triplex fibers is the proposed structure for CA-mediated d(A_n) assemblies (**Figure 2.14**).

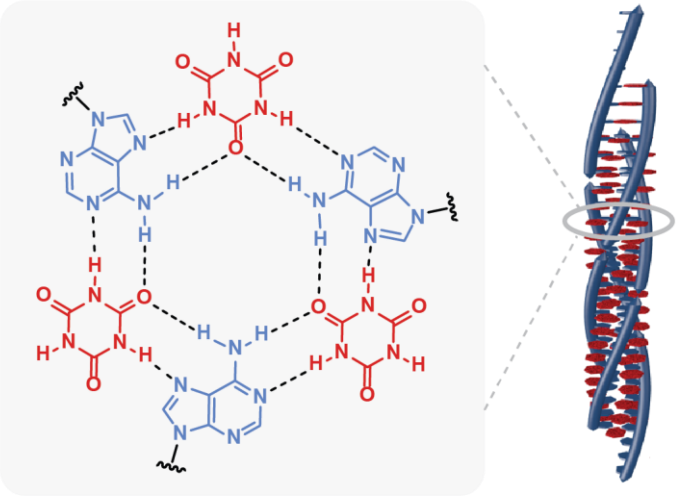
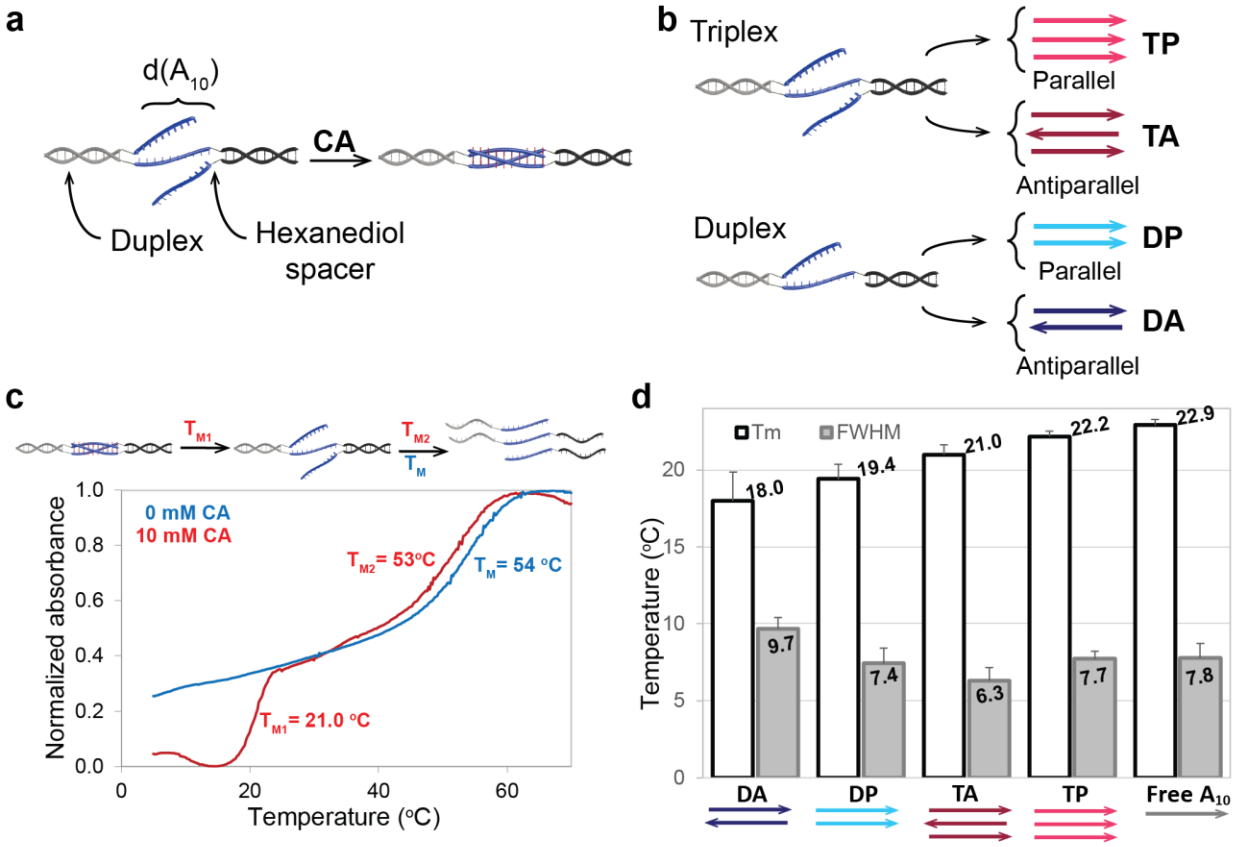


Figure 2.14| Proposed 1:1 CA:A hexameric rosette structure leading to d(A_n) triplex formation.

To further examine possible strand orientations and potential alternative molecularity for the assembly (a tetrameric rather than hexameric structure), a preorganized, intramolecular test system incapable of fiber elongation was designed. In this system, either two or three d(A_{10}) regions are spatially tethered into position using clipping duplexes. These clipping strands are linked to each of the poly(adenine) regions via a hexane-diol spacer to alleviate any potential structural strain that might occur during the self-assembly process (**Figure 2.15 a**). The flanking DNA duplexes template four distinct d(A_{10}) environments (**Figure 2.15 b**): (i) 3 parallel strands (TP); (ii) 3 antiparallel strands (TA); (iii) 2 parallel (DP) and (iv) 2 antiparallel strands (DA). d($CA_{10}C$) was used as a non-preorganized control strand, termed "Free d(A_{10})". DLS results showed that, in contrast to "Free A_{10}", TP did not undergo aggregation in the presence CA (**Table 2.2** in section 2.4.19). CD spectra for constructs in the presence of CA showed both features for duplex DNA and the d(A_n):CA distinct peaks.



e

T_M	DA	DP	TA	TP
DP	0.0434			
TA	0.0023	0.0020		
TP	0.0007	0.0003	0.0002	
Free A_{10}	0.0004	0.0001	0.0001	0.0031

f

FWHM	DA	DP	TA	TP
DP	0.0014			
TA	0.0002	0.0218		
TP	0.0006	**0.3249**	0.0034	
Free A_{10}	0.0023	**0.3317**	0.0074	**0.8197**

Figure 2.15| Preorganized system to study potential strand orientation and alternative molecularity for CA-mediated assembly. a) Clipping duplex regions are used to preorganize d(A_{10}) regions. Hexanediol spacers separate the duplex portions from the poly(A) stretches. **b)** Four different poly(A) microenvironments are created by the preorganized system: TP (3 parallel strands), TA (3 antiparallel strands), DP (2 parallel strands) and DA (2 antiparallel strands). **c)** Thermal denaturation followed by UV-Vis spectroscopy (TA construct shown) with 0 (blue) or 10 mM (red) CA. Denaturation of the preorganized d(A_{10}) regions interacting with CA (T_{M1}) is independent of the melting of duplex portions which occurs at a higher temperature (T_{M2} and T_M). **d)** Comparison of T_M (white) and FWHM (grey) values for the preorganized systems compared to "Free A_{10}". **e)** and **f)** Evaluation of statistical significance of differences between the thermal denaturation results for different preorganized constructs Tables of p-values comparing pairs of constructs; **(e)** T_M values and **(f)** FWHM values. Values in bold type ($p > 0.05$) are for pairs that are not statistically different.

Thermal denaturation of the constructs was monitored by CD and UV-Vis at 252 nm, a wavelength that tracks both the d(A_n):CA and duplex environments. Two distinct melting transitions were noted, the first for the d(A_n):CA structure, and the second for the DNA duplexes at the extremities of the structure (**Figure 2.15 c**, T_{M1} and T_{M2}). The separation between the T_M values enables the analysis of the designed poly(A) environments. While T_M is a reflection of the stability of the structure, full width at half maximum (FWHM), determined from the peak of the derivative of the melting curve, illustrates the cooperativity of the melting transition. Of the four environments, TP had the highest T_M that was also closest to the T_M of CA-mediated assembly of "Free A_{10}", whereas the DA construct showed the lowest T_M value (**Figure 2.15 d**). TA and DP had intermediate and decreasing T_M values ($T_{M\,TA} > T_{M\,DP}$). Although the T_M values for the various preorganized constructs are numerically close, statistical analysis shows that differences between them are indeed significant (**Figure 2.15 e**). Additionally, for FWHM values, those for the parallel constructs and "Free A_{10}" overlap, pointing to a similarity in cooperative behaviour between these systems (**Figure 2.15 f**). These observations agree with the proposed triplex-forming rosette structure, and suggest that invoking a structure of molecularity higher than a triplex to explain the thermal denaturation results for free d(A_{10}) is unwarranted.

2.2.3 Probing the versatility of CA-mediated assembly

We have uncovered the assembly of d(A_n) with CA and examined its underlying structure. We could then propose a hexameric CA:A hydrogen-bonded array as the base unit that brings about the formation of a poly(A) triplex. A look at this assembly motif using modified components was next sought in order to probe its versatility. Poly(A) strands with alternate backbones, small molecule CA-analogs and biotin-modified d(A_{15}) replaced standard d(A_n) and CA to examine the consequences of these substitutions on assembly and the potential applications of the nanomaterial produced.

Initially, nucleobase oligomers with alternative backbones were examined. The interaction of CA with A_{15} RNA (r(A_{15})) (TAMg pH 4.5) was studied by CD. In the presence of 15 mM CA, the CD spectrum for r(A_{15}) acquired the same features as those characteristic of d(A_{15}) assembly with CA (**Figure 2.16 a**) and had similar thermal denaturation properties. In the absence of CA at this pH, RNA has a stronger tendency than DNA to form a parallel homoduplex,[36,47] as evidenced

by a positive peak in the 250-300 nm region in the CD spectrum. Thermal denaturation of r(A_{15}) with CA was similar to that of d(A_{15}), but slightly broader and shifted to a lower T_M (**Figure 2.16 b**). AFM imaging of r(A_{15}) with CA in solution showed extensive formation of fibers (**Figure 2.16 c**), as in the case of d(A_{15}) under similar conditions. With its numerous biological functions and diverse self-assembly motifs, RNA is emerging as another exciting nucleic acid building block.[57-59] In addition, messenger RNA (mRNA) in eukaryotic cells has a 3'-poly(A) tail composed of ~200-300 adenosine residues.[60,61] This tail is essential for mRNA translation and plays an important role in its stability and maturation. As polyadenylation is an important step in the production of mature messenger RNA, CA-mediated poly(A) assembly can potentially find applications in probing or modulating protein expression.

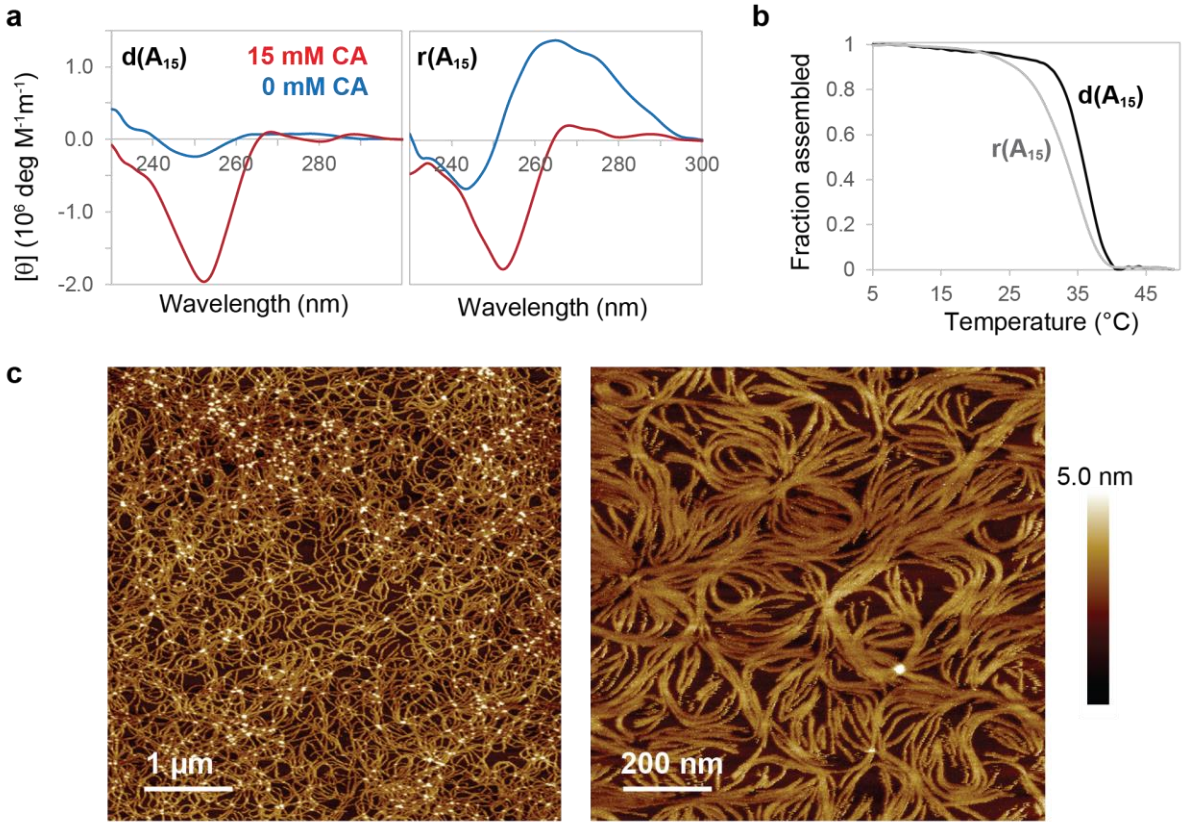


Figure 2.16| CA-mediated assembly r(A_{15}). a) CD spectra of d(A_{15}) (left) and r(A_{15}) (right) with 0 (blue) or 15 (red) mM CA in TAMg at pH 4.5. **b)** Thermal denaturation of d(A_{15}) (black) and r(A_{15}) (grey) with 15 mM CA obtained by CD (monitored at 252 nm). **c)** AFM imaging of r(A_{15}) with 15 mM CA in fluid.

Peptide nucleic acid (PNA), is an uncharged, achiral nucleic acid analog. The assembly of PNA A_7 ($p(A_7)$) with CA was monitored by UV-Vis spectroscopy, as PNA does not have a CD signature. Upon addition of CA to $p(A_7)$ (in water or TAMg pH 4.5), a hypochromic effect and slight red-shift were observed for the nucleobase signal at 260 nm (**Figure 2.17 a**). Short fibers, mostly around 100 nm in length and 2.0 ± 0.1 nm in height were observed by AFM (**Figure 2.17 c**). On the whole, the assembly of poly(A) DNA, RNA and PNA can be robustly achieved with the simple external addition of CA to the system, and produces long-range fibers of uniform, well-defined structure.

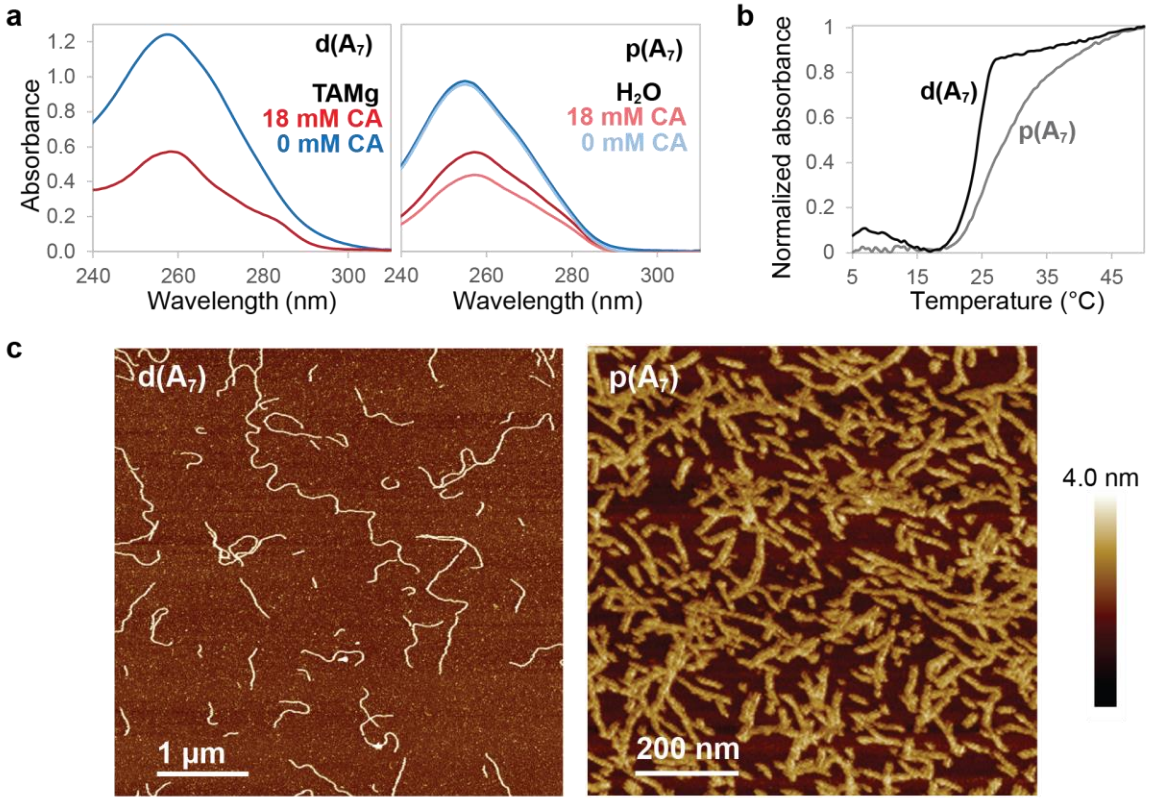


Figure 2.17| CA-mediated assembly p(A_7). a) UV-Vis absorbance spectra of d(A_7) (left) and p(A_7) (right) with 0 (blue) or 18 (red) mM CA. d(A_7) in TAMg pH 4.5; p(A_7) in TAMg at pH 4.5 and in H_2O (lighter shades of blue and red). **b)** Thermal denaturation of d(A_{15}) (black) and r(A_{15}) (grey) with 15 mM CA obtained by UV-Vis absorbance (monitored at 260 nm). **c)** AFM imaging of d(A_7) (left) and p(A_7) (right) with 18 mM CA in fluid.

Furthermore, d(A_{15}) assembly was screened with several CA analogs, to potentially expand the library of small molecules capable of mediating poly(A) assembly. The library included the

following molecules: triethylene glycol CA (TEG-CA), propyl-CA (pr-CA), uracil (U), thymine (T), uric acid (UA), xanthine (X), barbituric acid (BA), glycine anhydride (GA) and 5,6-dihydrouracil (DHU) (**Figure 2.18 a**). CA was initially of interest for assembly with poly(A) due to its three symmetric sides presenting hydrogen bond donor-acceptor motifs similar to T. TEG-CA and pr-CA were synthetically modified CA analogs, where one hydrogen-bond donating nitrogen is substituted with triethylene glycol or a propyl chain respectively to mask one binding face. U and T are pyrimidines and natural A-binders that have two binding faces. UA and X are purines with multiple binding faces. BA, GA and DHU are CA analogs with only two binding faces and, as they contain sp^3 carbons, do not present a flat π-surface. These molecules have a range of pK_a values and solubilities.

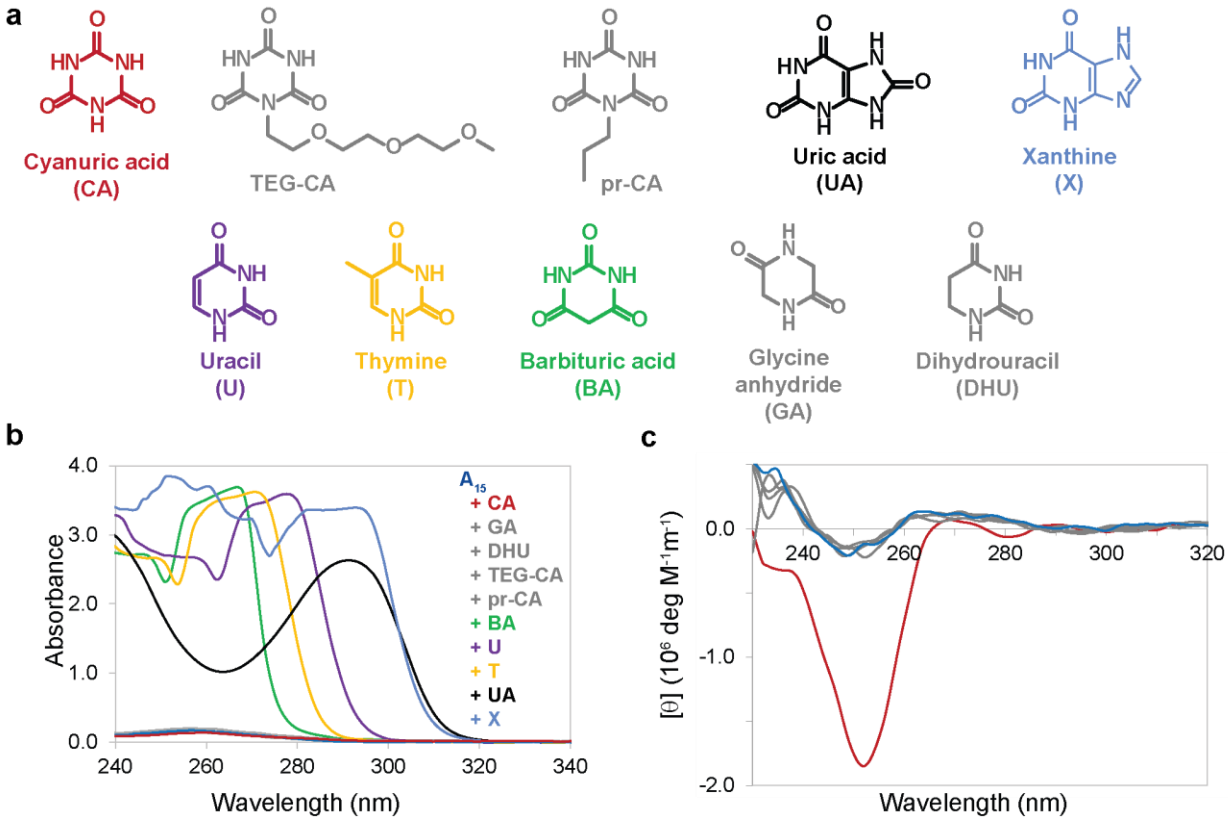


Figure 2.18| Screening for small molecule-mediated assembly of d(A_{15}). a) The library of small molecules used in the screening. **b)** UV-Vis absorbance spectra for d(A_{15}) alone (blue) and in the presence of a small molecule. The absorbance of BA (green), T (yellow), U (purple), UA (black) and X (light blue) is much higher than the absorbance of d(A_{15}), which prevents analysis of these molecules by CD. **c)** CD spectra for d(A_{15}) with CA, TEG-CA, pr-CA, GA and DHU (non-absorbing analogs).

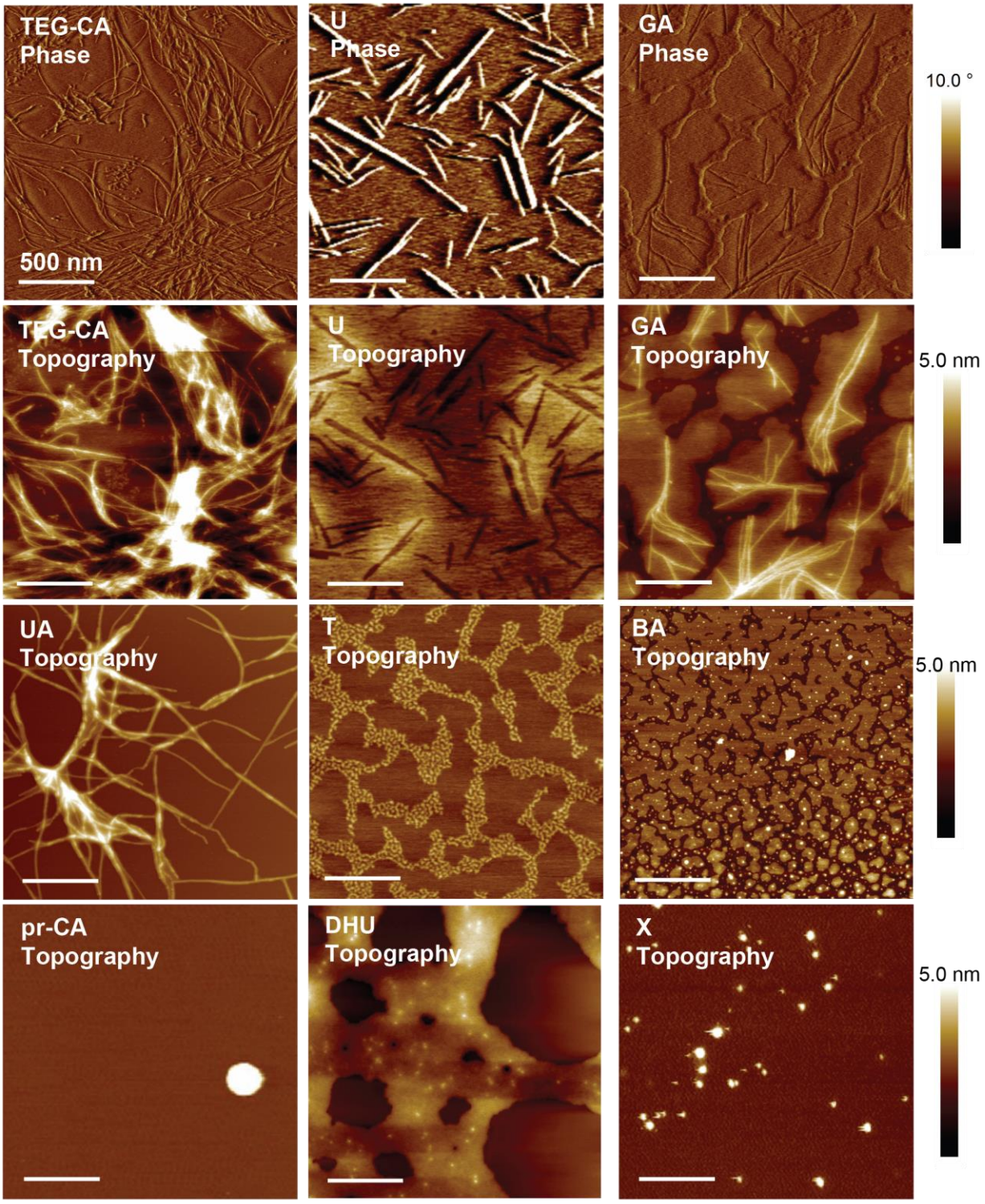


Figure 2.19| AFM images of d(A_{15}) in the presence of CA analogs in air. Top row: Phase images for d(A_{15}) with TEG-CA, U and GA. Second row: Topography images of the same sample regions depicted in the top row (TEG-CA, U and GA). Third and fourth rows: Topography images of d(A_{15}) with UA, T, BA, pr-CA, DHU and X. Scale bar: 500 nm.

An advantage of studying CA-mediated assembly of d(A_n) is that CA absorption of UV light does not overlap with the absorption profile of nucleic acids, which has a maximum around 260 nm. In contrast, CA analogs U, T, UA, X and BA are all highly absorbing in this region, which prevents the study of their assembly with d(A_{15}) by UV-Vis spectroscopy (**Figure 2.18 b**). As the CD signal also depends on absorption, these molecules cannot be studied by CD spectroscopy either, as their absorption (since they are present in large excess) oversaturates the spectrometer's photomultiplier detector, completely masking any features of the d(A_{15}) CD spectrum with noise. The UV-Vis and CD traces for d(A_{15}) with TEG-CA, pr-CA, GA and DHU (which are not absorbing in the same region as DNA), remain unchanged compared to d(A_{15}) alone, such that assembly cannot be detected by these methods.

Although investigation of CA-analog-mediated assembly of d(A_{15}) in solution was not fruitful, AFM studies in air did indicate that interactions with some of the analogs lead to fiber formation upon drying of the mixture on a mica surface. In the presence of TEG-CA, GA and UA, long fibers could be detected, whereas with U, rod structures were observed (**Figure 2.19**). Images for pr-CA, T, X, BA or DHU did not show organized assembly, although with T, small structures organized in patterns were observed. The influence of drying effects may have an impact on the features observed in all imaging in air. Oddly, interaction with TEG-CA or GA does not have an effect on the UV-Vis or CD spectra of d(A_{15}), although fibers are detected in the presence of these molecules by AFM in air. The increasing concentrations of materials achieved with sample drying on the surface and surface interactions may have a bearing on the assembly process. In summary, there is potential in studying CA-analog-mediated assembly of d(A_n) as some small molecules in the library mediate fiber formation once dry on a surface, however, studies in solution are complicated by the absorbance properties of these molecules.

To probe whether the CA-mediated d(A_{15}) fibers can act as templates for the organization of other materials, CA-mediated assembly of 5'-biotinylated d(A_{15}) (B-d(A_{15})) and unmodified d(A_{15}) mixtures (TAMg pH 6) at different molar ratios and surface loading were tested. The presence of a biotin modification did not alter fiber formation, as shown by AFM in solution. Addition of biotin-binding protein streptavidin (STV) in stoichiometric excess of 100% labelled B-d(A_{15}) fibers visibly altered the fibers, resulting in fiber bundling, as evidenced in the increased height of the topography features (**Figure 2.20 a**). Fiber bundling is most likely a consequence of

STV molecules, which have four biotin-binding sites, interacting with biotin moieties on adjacent fibers.

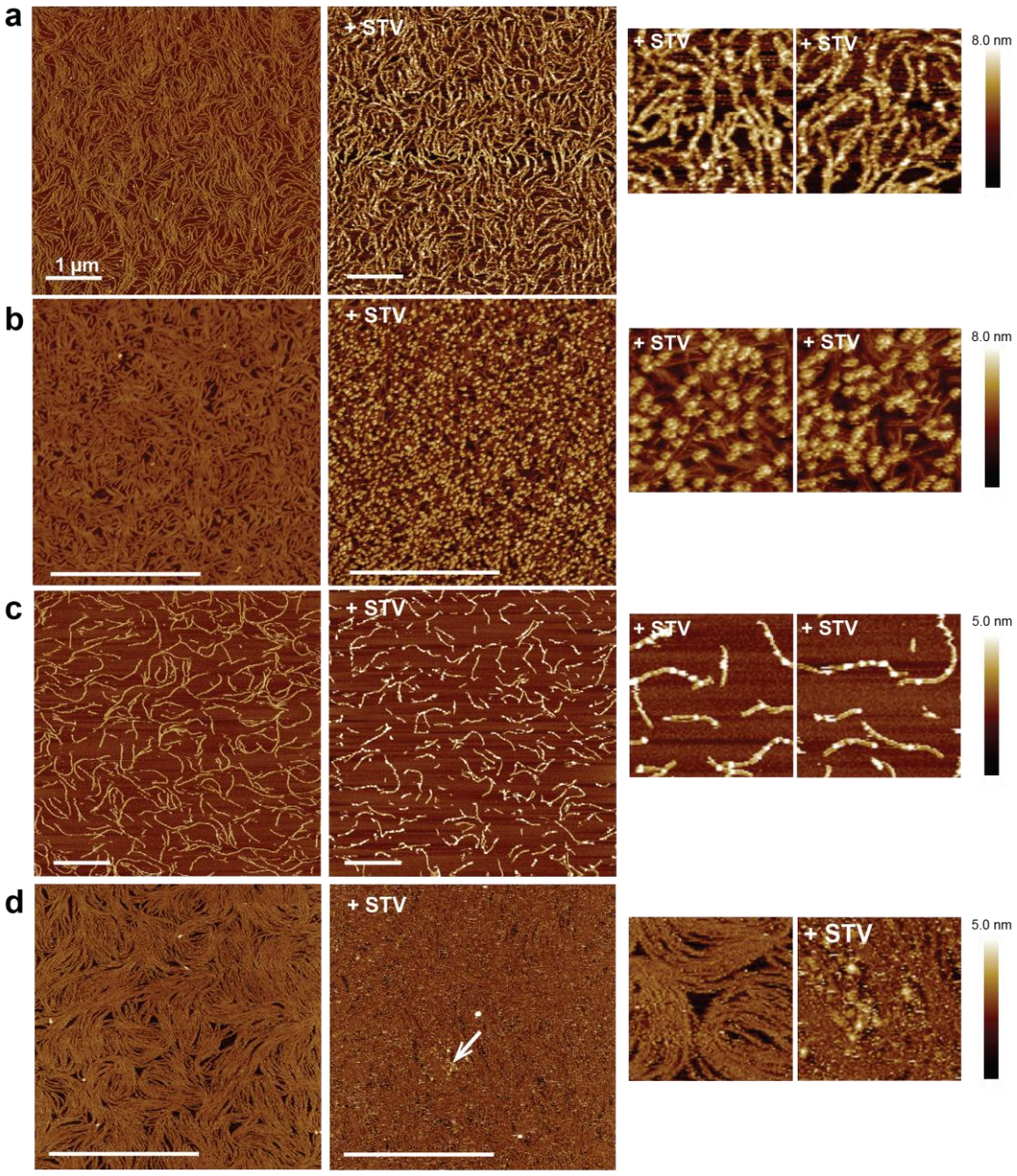


Figure 2.20| STV binding to CA-mediated fibers of B-d(A_{15}):d(A_{15}) mixtures studied by AFM in solution. (Left) before STV addition; (Center) after STV addition; (Right) zoomed-in features of interest. **a)** 7.5x sample dilution, 100% B-d(A_{15}) fibers, addition of 7:1 stoichiometric excess of STV: B-d(A_{15}). **b)** 7.5x sample dilution, 25% B-d(A_{15}) fibers, addition of 7:1 stoichiometric excess of STV: B-d(A_{15}). **c)** 50x sample dilution, 50% B-d(A_{15}) fibers, addition of 10:1 stoichiometric excess of STV: B-d(A_{15}). **d)** 7.5x sample dilution, 0% B-d(A_{15}) fibers, addition of 10:1 stoichiometric excess of STV:d(A_{15}). Arrow on middle image indicates area of bare mica where STV molecules have settled on the surface. Zoom in shows the same region before and after STV addition.

With lower ratios of B-d(A_{15}):d(A_{15}) (25% B-d(A_{15}), **Figure 2.20 b**) or with sample dilution (reduced loading) (50 times sample dilution, 50% B-d(A_{15}), **Figure 2.20 c**), individual STV molecules could be detected on the fibers. With STV addition to unlabelled CA-mediated assemblies of d(A_{15}) (25% B-d(A_{15}), **Figure 2.20 d**), no accumulation of protein was detected on the fibers. Instead the protein landed on bare mica areas between the fibers. With the addition of a large STV excess, imaging quality quickly degraded (likely due to the imaging probe's interactions with STV in solution), explaining increased noise in the image and loss in imaging resolution after STV addition. This experiment demonstrated that CA-mediated d(A_n) fibers can be decorated with STV, a step toward organization of biotinylated functional components on the nanofibers.

2.3 Conclusions

The work presented in this chapter tracks the discovery that cyanuric acid (CA), a small molecule with three thymine-like faces, can induce the assembly of unmodified poly(adenine) strands into a new non-canonical motif. This results in the growth of supramolecular fibers microns in length. The assembly is solution-based and the resulting fibers are consistent and abundant, as determined by AFM (in air and in solution) and DLS. CD spectroscopy reveals spectral features distinct from both B-DNA and previously reported poly(A) duplexes. Investigations into fiber structure support a proposed hexameric rosette, hydrogen-bonded, 1:1 CA:A array as the underlying assembly motif. AFM height measurements, assembly experiments with monomeric adenine, equilibrium dialysis, and elimination of a binding face of adenine by using an analog all support the formation of this motif, which brings about the formation triplex composed of d(A_n) strands. This phenomenon is observed for adenine oligomers with DNA, RNA, and PNA backbones and allows for functionalization of the oligonucleotide with another moiety.

This study has important consequences in a number of areas. The method described represents a facile way to create new nucleic acid-based nanomaterials. The CA-mediated nanofibers described here can be generated in large quantities. These structures are made from inexpensive and biocompatible units and are stable over a range of pH values. The CA-mediated poly(A) fibers may have interesting applications in tissue engineering and drug delivery. Our method introduces a new orthogonal interaction to the nucleic acid-based nanomaterial toolkit, expanding the 4-letter DNA alphabet in the process. Furthermore, guest-induced assembly of DNA

strands by a small molecule with multiple hydrogen bonding faces may be relevant in chemical biology. Long poly(A) strands are added to messenger RNA before its translation to a protein and poly(purine) stretches in general are biologically important.[62-64] Finally, the fact that a small molecule such as CA can organize poly(A) strands into long-range supramolecular polymers in aqueous media may have implications in prebiotic chemistry, specifically in answering the question of how individual nucleobases originally assembled onto a nucleic acid polymeric backbone.

2.4 Experimental

2.4.1 Materials

CA, tris(hydroxymethyl)aminomethane (Tris), magnesium chloride hexahydrate ($MgCl_2 \cdot 6 H_2O$), sodium chloride (NaCl), glacial acetic acid, urea, ethylenediaminetetraacetic acid (EDTA), formamide, ammonium hydroxide, Stains-All, trimethylamine, HPLC grade acetonitrile (MeCN) and toluene-d_8 were used as purchased from Sigma-Aldrich. Boric acid, hydrochloric acid (HCl) and GelRed were obtained from Fisher Scientific and used as supplied. Acrylamide/bis-acrylamide (40% 19:1) solution, ammonium persulfate and tetramethylethylenediamine (TEMED) were used as purchased from BioShop Canada Inc. Sephadex G-25 (super fine, DNA grade) was purchased from Glen Research. Streptavidin (STV) was purchased from Invitrogen.

Unless otherwise indicated, desalted DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT) and used without further purification. A_{15} RNA oligonucleotide was purchased from IDT with reverse phase HPLC purification. A_7 PNA was purchased from Panagene (Korea).

For in-house DNA synthesis, the reagents and materials listed below were used. Anhydrous acetonitrile, dA-CE, dC-CE, dmf-dG-CE, and dT-CE phosphoramidites, nucleoside derivatized and universal 1000 Å LCAA CPG pre-packed columns (loading densities of 25-40 μmol/g), activator (0.25 M ETT in acetonitrile), oxidizer (0.02 M I_2 in THF/pyridine/H_2O), Cap A mix (THF/lutidine/Ac_2O), Cap B mix (16% MeIm in THF), and deblock solution (3% DCA/DCM) were used as purchased from BioAutomation. 7-deaza-dA-CE phosphoramidite was purchased

from Glen Research. Reverse 3'-DMT deoxy adenosine (n-bz) 5'-CED and DMT-hexane-diol phoaphoramidites were purchased from ChemGenes.

Muscovite Ruby mica sheets (grade 2) were used as substrate for all AFM imaging studies. 1xTBE (Tris-boric acid-EDTA) buffer was composed of 45 mM Tris, 45 mM boric acid and 2 mM EDTA at pH 8.3. 1xTAMg buffer was composed of 40 mM Tris, 7.6 mM $MgCl_2 \cdot 6\ H_2O$ with pH adjusted to the desired value (3, 4, 4.5, 5, 6, 7, and 8) with glacial acetic acid. All buffers were prepared with Milli-Q water and samples were prepared with autoclaved Milli-Q water.

2.4.2 Instrumentation

UV-Vis DNA quantification measurements were performed on a BioTek Synergy HT microplate reader or a Nanodrop Lite spectrophotometer from Thermo Scientific. Circular dichroism (CD) and UV-Vis studies were performed using a 1 mm path length quartz cuvette on a Jasco-810 spectropolarimeter equipped with a xenon lamp, a Peltier temperature control unit and a water recirculator. A Mermade MM6 DNA synthesizer from Bioautomation was used for in-house DNA synthesis. Gel electrophoresis experiments were carried out on a 20 x 20 cm vertical acrylamide Hoefer 600 electrophoresis unit. Thermal annealing procedures were carried out using an Eppendorf Mastercycler Pro 96 well thermocycler. All pH measurements were performed on a Thermo Scientific Orion 3-Star Plus pH meter using a Mettler Toledo InLab Ultra-Micro combination pH electrode. Electrophoresis gels were scanned with either a CanoScan 4200F scanner by Canon or a Chemidoc MP Imaging System from BioRad. Equilibrium dialysis was performed using single use DispoEquilibrium Dialyzers (5000 Dalton molecular weight cutoff) from Harvard Apparatus. Reverse phase HPLC (RP-HPLC) analysis was conducted using a Hamilton PRP-1 column (150 mm x 4.1mm, 100 μm particle size) on an Agilent Infinity 1260 system.

2.4.3 General procedures

Unless otherwise indicated, the following general procedures were used. Samples were thermally annealed using a 50 → 5 protocol (50 °C for 10 minutes, followed by cooling to 5 °C at

a rate of 0.5 °C/min) and incubated overnight at 5 °C prior to any measurements. CD and UV-Vis spectra were obtained at 5 °C (230 nm to 350 nm range, 100 nm/min scan rate, 1 nm bandwidth, 2 accumulations). Ellipticity data (θ) obtained by CD were converted to molar ellipticity [θ], using Equation 2.2, where *c* is the molar concentration of DNA strands and *l* is the path length (in cm). In some cases, UV-Vis absorbance values were converted to molar absorptivity ε in the interest of comparing samples obtained at different DNA concentrations. Thermal denaturation of assemblies was recorded by following CD and UV-Vis signals at 252 nm over the 5 to 60 °C temperature range at 0.5 °C intervals, with a heating rate of 1 °C/min. The CD and UV-Vis melting data were converted to fraction of bases assembled (f) *vs.* temperature using Equation 2.3, where S is the signal at a given temperature, S_A is the signal for assembled d(A_n) species and S_F is the signal for free d(A_n) in solution. The melting temperature values (T_M) reported represent the temperature at which half-dissociation (fraction assembled f = 0.5) was recorded. First derivatives of the CD signal versus temperature curves were generated to verify the T_M. The values determined from both methods differed by less than 1.0 °C. FWHM was determined from the first derivative peak.

$$[\theta] = {}^{100\,\theta}/_{c\,l} \qquad \text{(Equation 2.2)}$$

$$f = {}^{S - S_F}/_{S_A - S_F} \qquad \text{(Equation 2.3)}$$

2.4.4 Atomic force microscopy (AFM) imaging in air

Samples and controls were prepared in water and had a total volume of 5 µL for deposition on mica. The sample composition was 0.1 nmol d(A_{15}), 10 mM CA and 5 mM NaCl (measured pH 4.5). The pH-adjusted control in the absence of CA was prepared by combining 0.1 nmol d(A_{15}) with 5 mM NaCl and adjusting the pH to 4.5 with HCl. The control in the absence of d(A_{15}) contained 10 mM CA and 5 mM NaCl (measured pH 4.5). Samples were incubated at ambient temperature for 30 minutes before deposition on freshly cleaved mica and left at 5 °C 4-6 hours. They were then placed in a desiccator and dried under vacuum overnight.

AFM imaging was carried out under ambient conditions in air using a Multimode 3 scanning probe microscope and Nanoscope IIIa controller (Digital Instruments, Santa Barbara, CA). Topography and phase contrast were simultaneously acquired in tapping mode with silicon probes (AC160TS from Olympus, nominal spring constant 42 N/m, resonant frequency of 300

kHz and tip radius < 10 nm). All images were captured at a 1 Hz scan rate and a resolution of 512 x 512 pixels.

2.4.5 AFM imaging in solution

Samples contained 10 μM d(A_{15}), 7.6 mM $MgCl_2$ and 10 mM CA. The samples were filtered using a 0.45 μm nylon syringe filter prior to thermal annealing and overnight incubation at 5 °C. For imaging, a 5 μL drop of the sample was deposited on a freshly cleaved mica surface and incubated for 2 minutes followed by addition of 50 μL of filtered 10 mM CA solution with 7.6 mM $MgCl_2$ to the closed MTFML (Bruker, Santa Barbara, CA) fluid cell and imaging.

Topography images were acquired by AFM carried out under ambient conditions using a MultiMode 8 microscope with a Nanoscope V controller (Bruker) in ScanAsyst mode. Silicon nitride levers with nominal spring constant of 0.7 N/m, resonant frequency of 150 kHz and tip radius < 20 nm were used (ScanAsyst Fluid). All images were captured at a 1.11 Hz scan rate and a resolution of 512 x 512 pixels. Average fiber height was determined to be 2.0 ± 0.1 nm.

2.4.6 Dynamic light scattering (DLS)

Samples contained 40 μM d(A_{15}), 7.6 mM $MgCl_2$ and 10 mM CA in water (pH 4.5). pH-controlled samples containing d(A_{15}) in the absence of CA (pH adjusted to 4.5 with addition of HCl) as well as controls in the absence of d(A_{15}) were prepared in 1 mL volumes. All samples were prepared in triplicate, filtered through 0.45 μm nylon syringe filters and incubated at 5 °C overnight.

DLS experiments were performed on a Brookhaven photon correlation spectrometer equipped with a BI9000 AT digital correlator. A Compass 315M-150 laser (Coherent Technologies, TX) was used at 532 nm as the incident light source. A refrigerated recirculator was used to control sample temperature. Autocorrelation functions were analyzed using the CONTIN algorithm. Borosilicate glass sample vials were purchased from Canadawide Scientific.

For constant temperature measurements, temperature was maintained at 20 °C and samples were allowed to equilibrate at this temperature for 15 minutes before any data were collected.

Scatter intensity data was acquired for 10 minutes with the detector set at 30, 50, 70, 90, 110 and 130° with respect to the incident light source. To obtain melting curves by DLS, scatter intensity was monitored while manually adjusting the temperature at a rate of 1 °C/min from 20 to 40 °C.

2.4.7 Circular dichroism (CD) and UV-Vis absorbance spectra

Samples containing 40 μM (A_{15}), 10 mM CA and 200 mM NaCl were prepared in water (pH 4.5). A pH controlled sample containing d(A_{15}) in the absence of CA (pH adjusted to 4.5 with addition of HCl) as well as a blank in the absence of d(A_{15}) were prepared. All samples were incubated overnight at 5 °C. CD and UV-Vis spectra were obtained (200 nm to 350 nm range, 100 nm/min scan rate, 1 nm bandwidth, 2 accumulations) at 20 °C. The samples were then thermally denatured at a rate of 1 °C/min following the CD signal at 252 nm and the UV-Vis signal at 260 nm, at 0.5 °C intervals over the 20 to 40 °C range.

For the melting/annealing curves illustrating hysteresis, a sample containing 40 μM d(A_{15}), 18 mM CA and 200 mM NaCl was prepared in water (pH 4.5) for an additional thermal denaturation study. The samples were incubated overnight at 5 °C and then thermally denatured at a rate of 0.5 °C/min following the CD signal at 252 nm, at 0.5 °C intervals over the 5 to 50 °C range. After a 5 minute equilibration at 50 °C, the sample was re-annealed at the same rate.

2.4.8 Random sequence control

Samples contained 40 μM of a random sequence DNA 20-mer control strand with no poly(adenine) stretches (**RAN**: 5' – CTTCAGTGGCATCAAGACGT – 3'), 10 mM CA and 200 mM NaCl, prepared in water (pH 4.5) along with a pH controlled samples containing RAN in the absence of CA (pH adjusted to 4.5 with addition of HCl), annealed and incubated overnight. CD and UV-Vis spectra were obtained (200 nm to 350 nm range, 100 nm/min scan rate, 1 nm bandwidth, 2 accumulations) at 5 °C.

2.4.9 Effect of counterions

Samples containing 40 μM d(A_{15}) and 10 mM CA were prepared in water with various concentrations of either NaCl or $MgCl_2$ added. For NaCl, the concentrations were: 5, 10, 20, 100, 200, 450 mM. For $MgCl_2$, the concentrations were: 0.05, 0.10, 0.25, 0.50, 1, 5, 10 and 20 mM. Samples were thermally annealed and incubated overnight. CD and UV-Vis spectra were obtained, followed by thermal denaturation.

2.4.10 Effect of pH on CA-mediated d(A_n) assembly

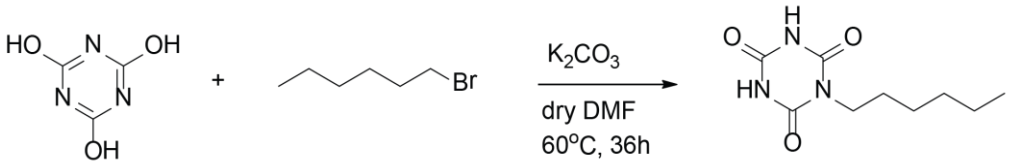
Samples containing 40 μM d(A_{15}) and 10 mM CA were prepared in 1xTAMg buffer of pH 3, 4, 5, 6, 7 and 8, annealed and incubated overnight. Controls with 0 mM CA were prepared in the same buffers. CD and UV-Vis spectra were obtained, followed by thermal denaturation. For the pH 3 samples, CD and UV-Vis signals were monitored at 266 nm instead of 252 nm (used for all other samples), due to a difference in CD signal maxima.

2.4.11 Adenosine-cyanuric acid co-crystallization

A single crystal of the 1:1 adenosine:CA complex was obtained by slow evaporation of an equimolar solution of the components in methanol. Single crystal X-ray diffraction data were collected using graphite monochromated Mo Kα radiation (λ = 0.71073 Å) on a Bruker APEX2 diffractometer at room temperature. Series of ω-scans were performed in such a way as to cover a sphere of data to a maximum resolution of 0.77 Å. Cell parameters and intensity data (including inter-frame scaling) were processed, and the structure solution was obtained using the APEX2 package. The structure was refined using full-matrix least-squares on F^2 within the CRYSTALS suite.[65] All non-hydrogen atoms were refined with anisotropic displacement parameters. The H atoms could be seen in the difference map, but those attached to carbon atoms were repositioned geometrically. Q-peaks for protic H atoms were confirmed by examining hydrogen bonding requirements. The H atoms were initially refined with soft restraints on the bond lengths and angles to regularise their geometry (C-H in the range 0.93-0.98 Å, N-H in the range 0.86-0.89 Å and isotropic displacement factors in the range 1.2-1.5 times U_{eq} of the parent atom), after which the positions were refined with riding constraints. After the construction of a stable, physically

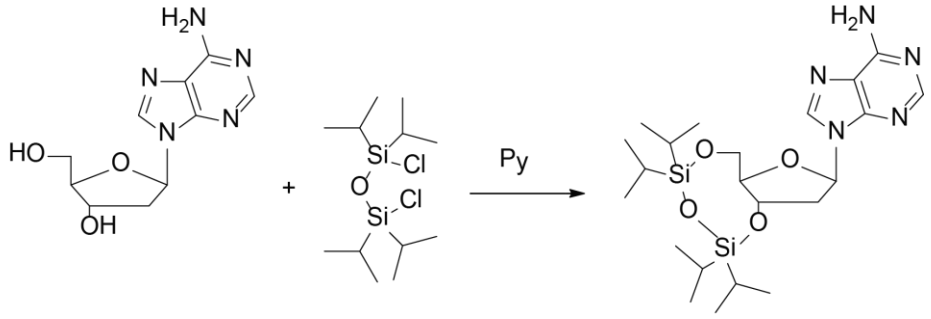
reasonable model, the weights were optimised,[66] leading to convergence of the refinement. IUCr CheckCIF/PLATON was used to validate the final structure.

2.4.12 Synthesis of 1-hexyl-1,3,5-triazinane-2,4,6-trione (hex-CA)



CA (2.14 g, 24.4 mmol, 3 equiv.) was suspended in 15 mL dry DMF under an inert atmosphere. 1-bromo hexane (1.34 g, 8.11 mmol, 1 equiv.) was added followed by K_2CO_3 (1.12g, 8.11 mmol, 1 equiv.). The reaction mixture was heated at 60ºC for 36 h. After the reaction was cooled at room temperature, ethyl ether was added and the solids were removed by filtration. The precipitate on the filter was washed several times with ethyl ether and the filtrate was concentrated under reduced pressure. The crude reaction mixture was purified by precipitation from hot hexanes. The cooling of the solution gave in time 400 mg of 1-hexyl-1,3,5-triazinane-2,4,6-trione (23 % yield). ^{1}H-NMR (400 MHz, DMSO-d_6) δ_H (ppm): 11.36 (s, 2H, C(O)N*H*C(O)), 3.6 (t, J=6.8 Hz, 2H, N-C*H*$_2$-CH$_2$~), 1.48 (t, J=6.4 Hz, 2H, N-CH$_2$-C*H*$_2$~), 1.27 (s (broad), 6H, ~(C*H*$_2$)$_3$CH$_3$), 0.85 (t, J= 6.4 Hz, (CH$_2$)$_3$*CH*$_3$). ^{13}C-NMR (75.4 MHz, DMSO-d_6) δ_C (ppm): 150.25, 149.05, 40.81, 31.33, 27.69, 26.222, 22.41,14.31.

2.4.13 Synthesis of 2'-deoxy-3',5'-O-(tetraisopropyldisiloxane-1,3-diyl)adenosine (TIPDS-dA)[67]



2'-Deoxyadenosine (Sigma Aldrich, 99+ purity) (0.2 g, 0.80 mmol, 1 equiv.) was co-evaporated twice with dry pyridine then dissolved in 4 mL dry pyridine under an inert atmosphere. To the resulting solution, 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (0.28 g, 0.88 mmol, 1.1 equiv.) was added dropwise and the reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted with dichloromethane and extracted with saturated aqueous sodium bicarbonate. The organic phase was separated, dried over magnesium sulfate and concentrated under reduced pressure. The crude product was purified by column chromatography (silica, dichloromethane/methanol 19:1 (v:v)). The fractions with $R_f = 0.22$ were separated to give 0.297 g of 2'-deoxy-3',5'-O-(tetraisopropyldisiloxane-1,3-diyl)adenosine as a white solid (75 % yield). ^{1}H-NMR (400 MHz, $CDCl_3$) δ_H (ppm): 8.32 (s, 1H, H-C_2), 8.06 (s, 1H, H-C_6), 6.30 (q, 1H, H-$C_{1'}$), 5.85 (s(broad), 2H, NH_2), 4.94 (q, 1H, H-$C_{4'}$), 4.05 (q, 2H, H-$C_{5'}$), 3.89 (m, 1H, H-$C_{3'}$), 2.69 (m, 2H, H-$C_{2'}$), 1.05 (m, 28H, ^{i}Pr) ^{13}C-NMR (75.4 MHz, DMSO-d_6) δ_C (ppm): 155.44, 152.91, 149.04, 138.91, 120.22, 85.16, 83.15, 69.80, 61.75, 40.02, 17.15, 12.97.

2.4.14 Vapor pressure osmometry (VPO) studies on TIPDS-dA:hex-CA assembly

VPO measurements were carried out on a Knauer K-7000 vapor pressure osmometer calibrated with benzil (M_w = 210.23 g/mol) and polystyrene PS2500 (M_w = 2,500 g/mol, M_w/M_n = 1.09) in HPLC grade toluene at 37 °C ± 2 °C. Two series of samples were initially prepared (Series 1 and Series 2) by adding solvent to the two mixed TIPDS-dA and hex-CA solids (1:1 molar ratio), gentle sonication and very short gentle heating. Two more sample series (Series 3 and Series 4)

were later prepared from samples recovered from the VPO chamber (dried under high vacuum and re-dissolved in fresh solvent). A concentration range from 2.7 to 96 g/kg was covered, corresponding to up to 117 mM in monomers. Variable equilibration times were used to obtain a stable response from the instrument, with the longest equilibration times (up to 15 minutes) required at the highest concentrations.

The M_W of species in the sample mixture was determined using Equation 2.4, where K_{VP} is a calibration factor determined from ΔV/c vs. c curves obtained for standards of known molecular weight ($K_{VP\ Benzil}$ = 21,400 and $K_{VP\ PS2500}$ = 22,000 under the given temperature, solvent and signal gain conditions) and b is the intercept for the linear extrapolation of the high-concentration ΔV/c vs. c sample response (**Figure 2.9 b**, right).

$$M_W = K_{VP}/b \qquad \text{(Equation 2.4)}$$

2.4.15 Temperature dependent NMR studies of TIPDS-dA:Hex-CA mixture

^{1}H-NMR spectra were acquired on a 500 MHz Varian VNMRS spectrometer equipped with a variable temperature unit with a dual broadband probe. A 100 mM 1:1 TIPDS-dA:Hex-CA mixture was prepared by dissolving pre-weighed solids in toluene-d_8 followed by sonication and gentle heating. Some solid remained undissolved which may have contributed to line broadening in NMR spectra. TIPDS-dA in 100 mM concentration in toluene-d_8 was used as a control. A Hex-CA control was not prepared as solubility of the compound in toluene-d_8 was not sufficient in the absence of TIPDS-dA. Spectra were obtained at 25, 30, 35, 40, 45, 50, 60 and 80 °C sequentially, with a 5 minute equilibration period at each temperature point. Subsequent cooling of the sample to 25 °C led to the recovery of the original spectra, demonstrating reversible structure formation.

2.4.16 DOSY NMR of TIPDS-dA:Hex-CA mixture

The 100 mM samples of the 1:1 TIPDS-dA:hex-CA mixture and TIPDS-dA described in section 2.4.15 were used to obtain DOSY spectra. A 500 MHz Varian VNMRS spectrometer with a dual broadband probe was used to acquire the spectra at 25°C. The spectra were acquired with DgcsteSL sequence using a diffusion delay of 50 ms, using 10 values for the diffusion gradient

strength from 1 000 to 30 000. A recycle delay of 2 ms was used for the accumulation of 64 scans for each value of the diffusion gradient strength. Phase and baseline corrections were applied before generating DOSY spectra.

2.4.17 Equilibrium dialysis

Equilibrium dialysis was used to establish the CA:A stoichiometry in the CA-mediated assembly of $d(A_n)$. Equal volumes (25 µL) of $d(A_{15})$ (40 µM) and CA (2, 3, 4, 4.5, 5, 5.5, 6, 6.5, 7, 8, 10, 12 and 15 mM) were allowed to equilibrate at 4 °C for 24 hours with gentle agitation, separated by a dialysis membrane (5000 Dalton molecular weight cut-off). Both solutions were then analyzed by RP-HPLC (solvents: 50 mM TEAA pH 8 and acetonitrile; gradient: 5:95 MeCN:TEAA for 5 minutes (min), 11 min gradient reaching 15:85 MeCN:TEAA, 2 min gradient reaching 95:5 MeCN:TEAA, 95:5, decrease to 5:95 over 2 min) to assess the equilibrium concentration of CA on either side of the membrane. The free concentration of CA, $[CA]_{free}$, and the concentration of CA bound, $[CA]_{Bound}$, were calculated based on the integrated area under the HPLC peak monitored by UV-Vis absorbance at 225 nm (retention time t = 1.52 min) (**Figure 2.21 a**) and a previously obtained calibration curve based on CA absorbance (**Figure 2.21 b**). The stoichiometry of CA bound per A, CA:A was determined from $[CA]_{Bound}$ and the initial $[d(A_{15})]$ at each CA concentration point. CA:A was plotted *vs.* $[CA]_{Free}$ to build a binding curve. The analysis was performed in triplicate in 4 separate experiments (n = 12) and all results were compiled. Error (CL 95%) is reported for both $[CA]_{Free}$ and CA:A.

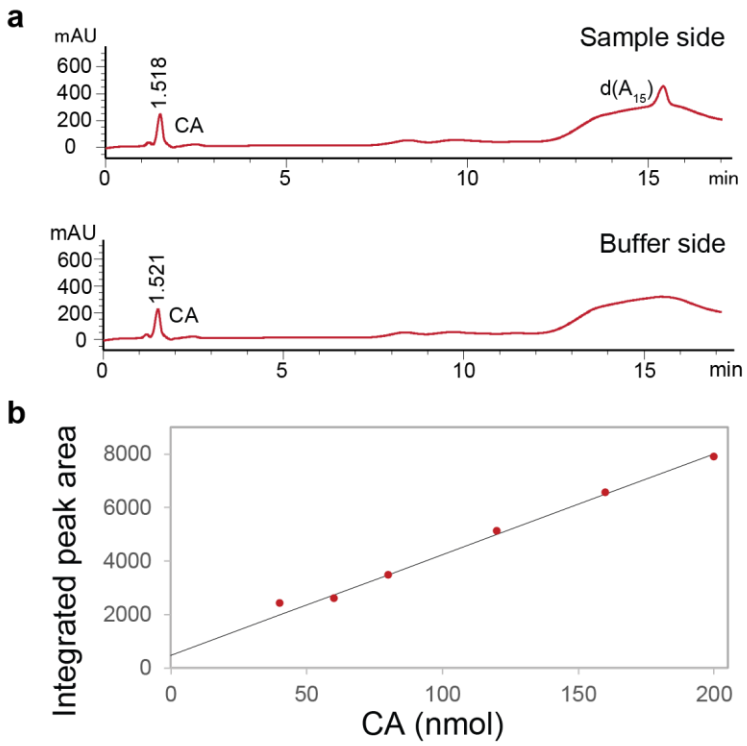


Figure 2.21| CA:A stoichiometry determination from HPLC traces. a) Sample HPLC traces monitoring UV-Vis absorbance at 225 nm used to determine $[CA]_{Free}$ and $[CA]_{Bound}$ in both chambers. **b)** Calibration curve used to calculate amount of CA from integrated peak area from UV-Vis absorbance HPLC traces.

2.4.18 7-Deazaadenine oligonucleotide assembly with CA

A DNA oligonucleotide composed of 12 7-deazaadenine bases (d(deaza-A_{12})) was synthesized in-house using 7-deaza-dA-CE phosphoramidite on a 1 µmol scale. Completed strands were deprotected in a 28 % ammonium hydroxide solution for 16 hours at 60 °C. d(A_{12}) was purchased from IDT. Both strands were purified by PAGE under denaturing conditions (20 % polyacrylamide with 8 M urea, using 1xTBE running buffer) at constant voltage (30 mA, at 250 V for the first 30 minutes, followed by 1 hour at 500 V). Following electrophoresis, the bands of purified oligonucleotides were excised and recovered from the gel (crushed, covered with water, vigorously shaken, flash frozen with liquid nitrogen and kept in a 60°C water bath for 16 hours). The solutions were then dried down to 1 mL and desalted using size exclusion chromatography (Sephadex G-25).

Samples containing 12.5 µM d(deaza-A_{12}) or d(A_{12}) in pH 4.5 and pH 6 1xTAMg in the presence of 15 mM CA as well as controls free of CA were incubated overnight at 4 °C. CD spectra were obtained at 5 °C.

2.4.19 Preorganized system with controlled strand orientation and molecularity

DNA synthesis was performed in house on a 1 µmol scale, starting from the required nucleotide-modified solid support. Coupling efficiency was monitored after removal of the DMT 5'-OH protecting groups. Strands Tri, AA, AB and B were synthesized using standard phosphoramidites in the 3'-5' direction. For strand PA and PB, the mixed sequence portion and DMT-hexane-diol phosphoramidite coupling were carried out in the 3'-5' direction, and synthesis of the d(A_{10}) portion in the 5'-3' direction was synthesized using reverse amidites. See **Table 2.1** for schematic strand representations and their full sequences. Completed sequences were cleaved from the solid support and deprotected as described in the previous section.

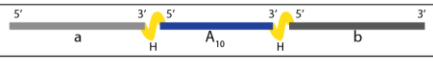
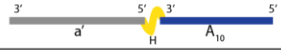
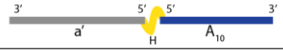
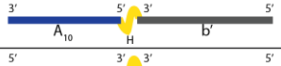
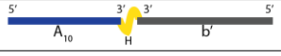
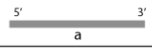
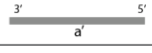
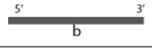
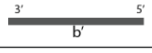
Name	Scheme	Sequences
Tri		5'-ATGACACTGACG-**H**-AAAAAAAAAA-**H**-TGAGGCGATGAT-3'
AA		3'-TACTGTGACTGC-**H**-AAAAAAAAAA-5'
PA		3'-TACTGTGACTGC-**H**-5'-5'-AAAAAAAAAA-3'
AB		5'-ATCATCGCC TCA-**H**-AAAAAAAAAA-3'
PB		5'-ATCATCGCC TCA-**H**-3'-3-AAAAAAAAAA-5'
a		5'-ATGACACTGACG-3'
a'		5'-CGTCAGTGTCAT-3'
b		5'-TGAGGCGATGAT-3'
b' = B		5'-ATCATCGCCTCA-3'

Table 2.1| Sequences and schematic representations of intramolecular system strands and control duplexes. Hexane-diol spacers (**H**) were used to give the system flexibility to ensure unencumbered CA-mediated assembly of d(A_{10}) segments.

Assembly of preorganized constructs was carried out in 1xTAMg pH 4.5 buffer. The constructs giving rise to different poly(A) environments (TP, TA, DP and DA) were prepared by combining equimolar amounts of component strands as shown in **Figure 2.22 a**. The double stranded portions of the construct were annealed first (95 °C for 5 minutes, slow cooling down to 4 °C over 4 hours). Quantitative construct formation was verified by 8% native PAGE, run at 5 °C with 1xTAMg pH 4.5 running buffer for 150 min at 250V (**Figure 2.22 b**). Upon addition of CA in 1xTAMg pH 4.5 up to a final concentration of 10 mM CA, the samples were subjected to another annealing protocol to allow the assembly of d(A_{10}) portions with CA to take place (30 °C for 10 minutes, followed by cooling to 5 °C at a rate of 0.5 °C/min). This annealing step was designed such that the maximum temperature was well below the denaturation temperature of the pre-annealed dsDNA portions of the construct. Total d(A_{10}) concentration in all samples was 10 μM, thus the final concentration of individual strands was 3.333 μM for triplex systems (TP and TA) and 5 μM for the duplex systems (DP and DA). A "Free A_{10}" control was prepared by annealing 10 μM d($CA_{10}C$) strand in 1xTAMg and 10 mM CA (30 °C for 10 minutes, followed by cooling to 5 °C at a rate of 0.5 °C/min). All samples were kept at 5 °C overnight before further studies were performed. Blanks containing only the random sequence duplex portions (aa' and bb') were prepared following the same procedures and subtracted from spectra obtained for full constructs.

To confirm that preorganized constructs do not show aggregation due to CA-mediated assembly of poly(A) segments, DLS experiments were carried out using a DynaPro™ Instrument from Wyatt Technology on TP and "Free A_{10}" with 0 or 10 mM CA. Samples were filtered through a 0.2 μm nylon filter before annealing (using the method described above). A cumulants fit model was used to calculate the particle size in the sample (assumes spherical objects).

CD and UV-Vis spectra were obtained for TP, TA, DP and DA and "Free A_{10}" to assess CA-mediated assembly based on the presence of the characteristic negative peak at 252 nm in the CD spectrum. The samples were then thermally denatured at a rate 0.5 °C/min, following the CD signal at 252 nm and the UV-Vis signal at 260 nm at 0.5 °C intervals from 5 to 60 °C. Error intervals for the triplicate measurements were calculated at 95% confidence. Statistical significance of the melting experiment results was assessed by examining p-values obtained from an unpaired t-test analysis of T_M and FWHM values for the different constructs (calculated using GraphPad QuickCalcs online software).

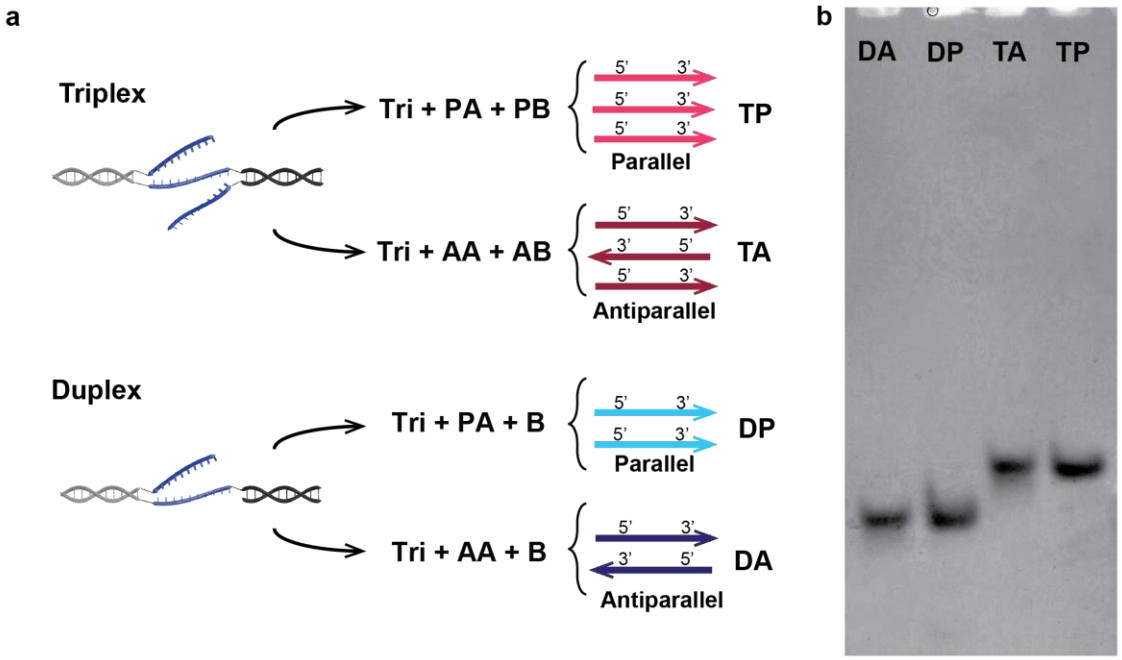


Figure 2.22| Assembly of preorganized constructs. a) Strand combinations giving rise to preorganized environments with different d(A_{10}) strand number and orientation. The preorganized system can be used to place A_{10} stretches in four different environments: triplex parallel (TP), triplex antiparallel (TA), duplex parallel (DP) and duplex antiparallel (DA). Refer to **Table 2.1** for the strand labels. **b**) 8% native PAGE of the preorganized constructs. Lane description from left to right: DA, DP, TA and TP. Loaded 0.01 OD_{260} total DNA per lane.

	[CA] (mM)	R_H (nm)
Free A_{10}	0	0.4 ± 0.2
	10	66.3 ± 9.6
TP	0	4.2 ± 0.3
	10	4.2 ± 0.5

Table 2.2| Summary of mean hydrodynamic radii (R_H) from DLS measurements. All samples were measured at least in triplicate (error CL 95%).

2.4.20 r(A_{15}) assembly with CA

Samples containing 40 µM r(A_{15}) or d(A_{15}) in 1xTAMg pH 4.5 in the presence of 0 and 15 mM CA were incubated overnight at 4 °C. CD and UV-Vis spectra were obtained at 5°C, followed by thermal denaturation.

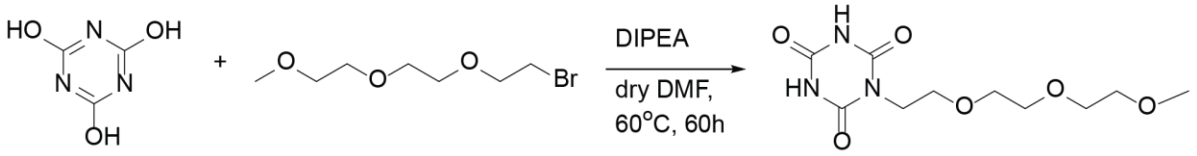
A 1 µL drop of 1:40 diluted sample (using 1xTAMg pH 4.5 with 15 mM CA as diluent) was deposited on a freshly cleaved mica surface treated with a $NiCl_2$ solution (20 µL of 20 mM $NiCl_2$ solution deposited for 1 minute and dried with a stream of nitrogen gas) and incubated for 4 minutes followed by addition of 80 µL of filtered 1xTAMg pH 4.5 with 15 mM CA to the fluid cell. Images were acquired as described in section 2.4.5. Average fiber height was measured to be 2.1 ± 0.1 nm.

2.4.21 p(A_7) assembly with CA

Samples containing 10 µM of d(A_7) or p(A_7) were incubated overnight at 4 °C in 1xTAMg pH 4.5 in the presence of 0 or 18 mM CA. p(A_7) was also prepared in water in the absence of counterions. UV-Vis spectra were obtained at 5°C, followed by thermal denaturation.

AFM imaging was carried out in solution. A 1 to 3 µL drop of sample was deposited on a freshly cleaved mica surface with $NiCl_2$ treatment (as described in section 2.4.20) and incubated for 2 minutes followed by addition of 75 µL of filtered 7.6 TAMg with 18 mM CA to the fluid cell. Images were acquired as described in section 2.4.5 with temperature maintained at 12 °C (using a mobile air conditioning unit with a large hose directed inside the noise isolation box housing the AFM instrument).

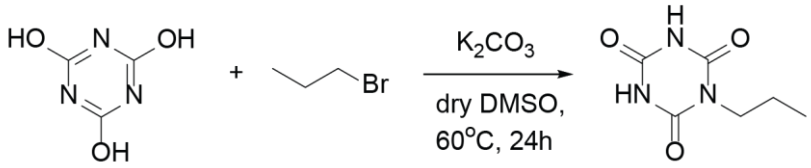
2.4.22 Synthesis of 1-(2-(2-(2-methoxyethoxy)ethoxy)ethyl)-1,3,5-triazinane-2,4,6-trione (TEG-CA)[68]



Cyanuric acid (2.0022 g, 15.51 mmol, 5.3 equiv.), 1-bromo-2-(2-(2-methoxy)ethoxy)ethane (0.6663 g, 2.92 mmol) were dissolved in 35 mL dry DMF. Diisopropyl ethyl amine (DIPEA) (0.4906g, 3.796 mmol, 1.3 equiv.) was added drop wise and the reaction mixture was heated at 60 °C for 60 h. After the reaction was cooled at room temperature, the solids

were removed by filtration and the filtrate was concentrated under reduced pressure. To the residue, dichloromethane was added and the slurry was subjected to a second filtration. The halogenated solvent was removed and the residue was taken in hot hexanes. The cooling of the solution gave in time 60 mg of 1-(2-(2-(2-methoxyethoxy)ethoxy)ethyl)-1,3,5-triazinane-2,4,6-trione (TEG-CA) (8% yield) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 3.38 (s, 3H, O*CH*$_3$), 3.56 (t, J=3.6 Hz, 2H), 3.68 – 3.63 (m, 6H), 3.75 (t, J=5.2 Hz, 2H), 4.06 (t, J=5.2 Hz, 2H), 9.57 (s, 2H, C(O)N*H*C(O)). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm): 40.61, 59.04, 67.50, 70.06, 70.39, 70.59, 72.00, 148.86, 149.72.

2.4.23 Synthesis of 1-propyl-1,3,5-triazinane-2,4,6-trione (pr-CA)[69]



Cyanuric acid (3.14 g, 24.4 mmol, 3 equiv.) and anhydrous K_2CO_3 (1.04 g, 8.13 mmol, 1 equiv.) were suspended in dry DMSO (15 mL), and heated to 60 °C. N-propyl bromide (1 g, 8.13 mmol, 1 equiv.) was added in one portion and sealed flask heated at 60 °C for 24 h. The reaction mixture was diluted with diethyl ether (50 mL) and washed with water (50 mL), then with brine (3 x 50 mL). Ether solution was dried with Na_2SO_4, and evaporated to approx. 10 mL. The product was precipitated with n-hexane as a white powder (279 mg, 20%).^{1}H NMR (400 MHz, DMSO-*d*$_6$) δ(ppm): 0.84 (t, *J* = 7.4 Hz, 3H), 1.46-1.58 (m, 2H), 3.56-3.61 (m, 2H), 11.39 (br. s, 2H). ^{13}C NMR (100 MHz, DMSO- *d*$_6$) δ (ppm): 11.4, 21.1, 42.4, 149.1, 150.3.

2.4.24 Screening for d(A_{15}) assembly with CA analogs

The analog library included: Cyanuric acid (CA, pK_a = 6.88), TEG-CA, pr-CA, thymine (T, pK_a = 9.94), uracil (U, pK_a = 9.5), uric acid (UA, pK_a = 5.4, 11.3), xanthine (X, pK_a = 7.4, 11.1), glycine anhydride (GA, pK_a = 11.73), 5,6-dihydrouracil (DHU, pK_a = 11.73) and barbituric acid (BA, pK_a = 4.01). UA and X stocks required addition of NaOH to promote dissolution.

Samples containing 10 μM A_{15} in pH 4.5 1xTAMg buffer (or unbuffered 7.6 mM $MgCl_2$) and 10 mM CA or analog were annealed (50 °C for 10 minutes, followed by cooling to 5 °C at a rate of 0.5 °C/min) and incubated overnight at 5 °C. CD and UV-Vis absorbance spectra were obtained at 5 °C for samples with d(A_{15}) and CA analog blanks.

For AFM imaging a 5 μL drop of sample prepared as described above was deposited on freshly cleaved mica. Samples were left at 5 °C for 4-6 hours and later dried under vacuum overnight. For samples containing CA, TEG-CA, pr-CA, U, T, GA, DHU and BA, AFM imaging was carried out under ambient conditions in air using tapping mode as described in section 2.4.4. Imaging of samples containing UA and X was performed on a MultiMode 8 microscope with a Nanoscope V controller (Bruker) in ScanAsyst mode. Silicon nitride levers with a nominal spring constant of 0.4 N/m, resonant frequency of 70 kHz and a 2 nm tip radius were used (ScanAsyst Air).

2.4.25 Streptavidin (STV) binding to biotinylated CA-mediated d(A_{15}) fibers

HPLC purified, 5'-biotinylated A_{15} (B-d(A_{15})) was purchased from IDT. d(A_{15}) was purchased from IDT and purified by PAGE. Samples containing either B-d(A_{15}), d(A_{15}) or a mixture of the two (B-d(A_{15}):d(A_{15}) ratios: 0:100, 25:75, 50:50, 75:25 and 100:0) were prepared with total DNA concentration of 5 μM and 18 mM CA in 1xTAMg pH 6, annealed and incubated overnight.

For AFM imaging, a 1 μL drop of sample (sample dilution was also used to vary fiber coverage of the surface) was deposited on a freshly cleaved mica surface with $NiCl_2$ treatment (described in section 2.4.20) and incubated for 2 minutes followed by addition of 80 μL of filtered 1xTAMg pH 6 with 18 mM CA to the closed liquid cell. Images were acquired as described in section 2.4.5, with temperature maintained at 12 °C. STV was added either in stoichiometric 1:1 amount with respect to biotin content in the sample or in large excess (>10:1), followed by immediate image acquisition.

2.5 References

1 Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738 (1953).

2 Choi, J. & Majima, T. Conformational changes of non-B DNA. *Chem. Soc. Rev.* **40**, 5893-5909 (2011).

3 Aldaye, F. A., Palmer, A. L. & Sleiman, H. F. Assembling materials with DNA as the guide. *Science* **321**, 1795-1799 (2008).

4 Egholm, M., Buchardt, O., Nielsen, P. E. & Berg, R. H. Peptide nucleic acids (PNA) - oligonucleotide analogs with an achiral peptide backbone. *J. Am. Chem. Soc.* **114**, 1895-1897 (1992).

5 Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B. & Nielsen, P. E. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **365**, 566-568 (1993).

6 Chen, M. C., Cafferty, B. J., Mamajanov, I., Gallego, I., Khanam, J., Krishnamurthy, R. & Hud, N. V. Spontaneous prebiotic formation of a beta-ribofuranoside that self-assembles with a complementary heterocycle. *J. Am. Chem. Soc.* **136**, 5640-5646 (2014).

7 Malyshev, D. A., Dhami, K., Lavergne, T., Chen, T., Dai, N., Foster, J. M., Correa, I. R., Jr. & Romesberg, F. E. A semi-synthetic organism with an expanded genetic alphabet. *Nature* **509**, 385-388 (2014).

8 Yang, Z., Chen, F., Chamberlin, S. G. & Benner, S. A. Expanded genetic alphabets in the polymerase chain reaction. *Angew. Chem., Int. Ed.* **49**, 177-180 (2010).

9 Teo, Y. N. & Kool, E. T. DNA-multichromophore systems. *Chem. Rev.* **112**, 4221-4245 (2012).

10 Winnacker, M. & Kool, E. T. Artificial genetic sets composed of size-expanded base pairs. *Angew. Chem., Int. Ed.* **52**, 12498-12508 (2013).

11 Sessler, J. L. & Jayawickramarajah, J. Functionalized base-pairs: versatile scaffolds for self-assembly. *Chem. Commun.* 10.1039/b418526a, 1939-1949 (2005).

12 Woo, S. & Rothemund, P. W. K. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nat. Chem.* **3**, 620-627 (2011).

13 Yang, H., McLaughlin, C. K., Aldaye, F. A., Hamblin, G. D., Rys, A. Z., Rouiller, I. & Sleiman, H. F. Metal-nucleic acid cages. *Nat. Chem.* **1**, 390-396 (2009).

14 Piccirilli, J. A., Krauch, T., Moroney, S. E. & Benner, S. A. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **343**, 33-37 (1990).

15 Kimoto, M., Kawai, R., Mitsui, T., Yokoyama, S. & Hirao, I. An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules. *Nucleic Acids Res.* **37**, e14 (2009).

16 Ogawa, A. K., Wu, Y. Q., McMinn, D. L., Liu, J. Q., Schultz, P. G. & Romesberg, F. E. Efforts toward the expansion of the genetic alphabet: Information storage and replication with unnatural hydrophobic base pairs. *J. Am. Chem. Soc.* **122**, 3274-3287 (2000).

17 Wu, Y. Q., Ogawa, A. K., Berger, M., McMinn, D. L., Schultz, P. G. & Romesberg, F. E. Efforts toward expansion of the genetic alphabet: Optimization of interbase hydrophobic interactions. *J. Am. Chem. Soc.* **122**, 7621-7632 (2000).

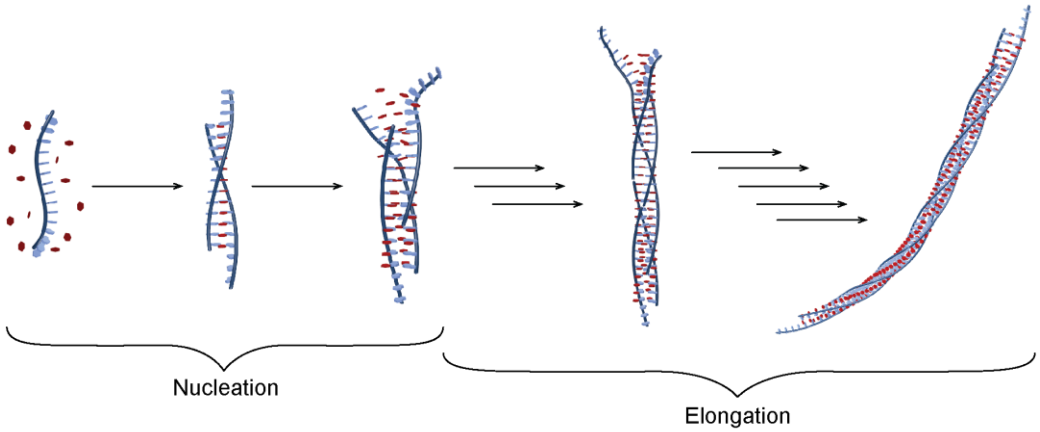
18 Meggers, E., Holland, P. L., Tolman, W. B., Romesberg, F. E. & Schultz, P. G. A novel copper-mediated DNA base pair. *J. Am. Chem. Soc.* **122**, 10714-10715 (2000).

19 Shionoya, M. & Tanaka, K. Synthetic incorporation of metal complexes into nucleic acids and peptides directed toward functionalized molecules. *Bull. Chem. Soc. Jpn.* **73**, 1945-1954 (2000).
20 Martin-Ortiz, M., Gomez-Gallego, M., Ramirez de Arellano, C. & Sierra, M. A. The selective synthesis of metallanucleosides and metallanucleotides: a new tool for the functionalization of nucleic acids. *Chem. Eur. J.* **18**, 12603-12608 (2012).
21 Chen, H., Meena & McLaughlin, L. W. A Janus-wedge DNA triplex with A-W1-T and G-W2-C base triplets. *J. Am. Chem. Soc.* **130**, 13190-13191 (2008).
22 Shin, D. & Tor, Y. Bifacial nucleoside as a surrogate for both T and A in duplex DNA. *J. Am. Chem. Soc.* **133**, 6926-6929 (2011).
23 Yang, H. Z., Pan, M. Y., Jiang, D. W. & He, Y. Synthesis of Janus type nucleoside analogues and their preliminary bioactivity. *Org. Biomol. Chem.* **9**, 1516-1522 (2011).
24 Pan, M.-Y., Hang, W., Zhao, X.-J., Zhao, H., Deng, P.-C., Xing, Z.-H., Qing, Y. & He, Y. Janus-type AT nucleosides: synthesis, solid and solution state structures. *Org. Biomol. Chem.* **9**, 5692-5702 (2011).
25 Largy, E., Liu, W., Hasan, A. & Perrin, D. M. Base-pairing behavior of a carbocyclic Janus-AT nucleoside analogue capable of recognizing A and T within a DNA duplex. *ChemBioChem* **14**, 2199-2208 (2013).
26 Marsh, A., Silvestri, M. & Lehn, J. M. Self-complementary hydrogen bonding heterocycles designed for the enforced self-assembly into supramolecular macrocycles. *Chem. Commun.* Doi 10.1039/Cc9960001527, 1527-1528 (1996).
27 Zeng, Y., Pratumyot, Y., Piao, X. & Bong, D. Discrete assembly of synthetic peptide-DNA triplex structures from polyvalent melamine-thymine bifacial recognition. *J. Am. Chem. Soc.* **134**, 832-835 (2012).
28 Chaput, J. C. & Switzer, C. A DNA pentaplex incorporating nucleobase quintets. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10614-10619 (1999).
29 Jiang, D. & Seela, F. Oligonucleotide duplexes and multistrand assemblies with 8-aza-2'-deoxyisoguanosine: a fluorescent isoG(d) shape mimic expanding the genetic alphabet and forming ionophores. *J. Am. Chem. Soc.* **132**, 4016-4024 (2010).
30 Polak, M. & Hud, N. V. Complete disproportionation of duplex poly(dT)·poly(dA) into triplex poly(dT)·poly(dA)·poly(dT) and poly(dA) by coralyne. *Nucleic Acids Res.* **30**, 983-992 (2002).
31 Persil, Ö., Santai, C. T., Jain, S. S. & Hud, N. V. Assembly of an antiparallel homo-adenine DNA duplex by small-molecule binding. *J. Am. Chem. Soc.* **126**, 8644-8645 (2004).
32 Jain, S. S., Polak, M. & Hud, N. V. Controlling nucleic acid secondary structure by intercalation: effects of DNA strand length on coralyne-driven duplex disproportionation. *Nucleic Acids Res.* **31**, 4608-4615 (2003).
33 Çetinkol, O. P. & Hud, N. V. Molecular recognition of poly(A) by small ligands: an alternative method of analysis reveals nanomolar, cooperative and shape-selective binding. *Nucleic Acids Res.* **37**, 611-621 (2009).
34 Park, K. S. & Park, H. G. A DNA-templated silver nanocluster probe for label-free, turn-on fluorescence-based screening of homo-adenine binding molecules. *Biosens. Bioelectron.* **64**, 618-624 (2015).
35 Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discovery* **10**, 261-275 (2011).

36 Rich, A., Davies, D. R., Crick, F. H. & Watson, J. D. The molecular structure of polyadenylic acid. *J. Mol. Biol.* **3**, 71-86 (1961).
37 Chakraborty, S., Sharma, S., Maiti, P. K. & Krishnan, Y. The poly dA helix: a new structural motif for high performance DNA-based molecular switches. *Nucleic Acids Res.* **37**, 2810-2817 (2009).
38 Zerkowski, J. A., Seto, C. T. & Whitesides, G. M. Solid-state structures of rosette and crinkled tape motifs derived from the cyanuric acid melamine lattice. *J. Am. Chem. Soc.* **114**, 5473-5475 (1992).
39 Rakotondradany, F., Palmer, A., Toader, V., Chen, B. Z., Whitehead, M. A. & Sleiman, H. F. Hydrogen-bond self-assembly of DNA-analogues into hexameric rosettes. *Chem. Commun.* Doi 10.1039/B511984g, 5441-5443 (2005).
40 Roy, B., Bairi, P. & Nandi, A. K. Supramolecular assembly of melamine and its derivatives: nanostructures to functional materials. *RSC. Adv.* **4**, 1708-1734 (2014).
41 Jelley, E. E. Spectral absorption and fluorescence of dyes in the molecular state. *Nature* **138**, 1009-1010 (1936).
42 Berova, N., Nakanishi, K. & Woody, R. *Circular dichroism: principles and applications*. (Wiley-VCH, New York, 2000).
43 De Greef, T. F., Smulders, M. M., Wolffs, M., Schenning, A. P., Sijbesma, R. P. & Meijer, E. W. Supramolecular polymerization. *Chem. Rev.* **109**, 5687-5754 (2009).
44 Tan, Z.-J. & Chen, S.-J. Nucleic acid helix stability: effects of salt concentration, cation valence and size, and chain length. *Biophys. J.* **90**, 1175-1190 (2006).
45 Sissoëff, I., Grisvard, J. & Guillé, E. Studies on metal ions-DNA interactions: Specific behaviour of reiterative DNA sequences. *Prog. Biophys. Mol. Biol.* **31**, 165-199 (1978).
46 Owczarzy, R., Moreira, B. G., You, Y., Behlke, M. A. & Walder, J. A. Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry* **47**, 5336-5353 (2008).
47 Safaee, N., Noronha, A. M., Rodionov, D., Kozlov, G., Wilds, C. J., Sheldrick, G. M. & Gehring, K. Structure of the parallel duplex of poly(A) RNA: evaluation of a 50 year-old prediction. *Angew. Chem., Int. Ed.* **52**, 10370-10373 (2013).
48 Krishnan, Y. & Simmel, F. C. Nucleic acid based molecular devices. *Angew. Chem., Int. Ed.* **50**, 3124-3156 (2011).
49 Lide, D. R. *CRC handbook of chemistry and physics*. (CRC press, 2004).
50 Ma, M. & Bong, D. Determinants of cyanuric acid and melamine assembly in water. *Langmuir* **27**, 8841-8853 (2011).
51 Cafferty, B. J., Gallego, I., Chen, M. C., Farley, K. I., Eritja, R. & Hud, N. V. Efficient self-assembly in water of long noncovalent polymers by nucleobase analogues. *J. Am. Chem. Soc.* **135**, 2447-2450 (2013).
52 Whitesides, G. M., Simanek, E. E., Mathias, J. P., Seto, C. T., Chin, D., Mammen, M. & Gordon, D. M. Noncovalent synthesis: using physical-organic chemistry to make aggregates. *Acc. Chem. Res.* **28**, 37-44 (1995).
53 Timmerman, P., Weidmann, J.-L., Jolliffe, K. A., Prins, L. J., Reinhoudt, D. N., Shinkai, S., Frish, L. & Cohen, Y. NMR diffusion spectroscopy for the characterization of multicomponent hydrogen-bonded assemblies in solution. *J. Chem. Soc., Perkin Trans. 2* 10.1039/B003968N, 2077-2089 (2000).
54 Davies, R. J. & Davidson, N. Base pairing equilibria between polynucleotides and complementary monomers. *Biopolymers* **10**, 1455-1479 (1971).

55 Davies, R. J. Complexes of poly(adenylic acid) with complementary monomers. *Eur. J. Biochem.* **61**, 225-236 (1976).
56 Pyle, A. M., Rehmann, J. P., Meshoyrer, R., Kumar, C. V., Turro, N. J. & Barton, J. K. Mixed-ligand complexes of ruthenium(II): factors governing binding to DNA. *J. Am. Chem. Soc.* **111**, 3051-3058 (1989).
57 Chworos, A., Severcan, I., Koyfman, A. Y., Weinkam, P., Oroudjev, E., Hansma, H. G. & Jaeger, L. Building programmable jigsaw puzzles with RNA. *Science* **306**, 2068-2072 (2004).
58 Hao, C., Li, X., Tian, C., Jiang, W., Wang, G. & Mao, C. Construction of RNA nanocages by re-engineering the packaging RNA of Phi29 bacteriophage. *Nat. Commun.* **5** (2014).
59 Geary, C., Rothemund, P. W. K. & Andersen, E. S. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science* **345**, 799-804 (2014).
60 Weill, L., Belloc, E., Bava, F.-A. & Mendez, R. Translational control by changes in poly(A) tail length: recycling mRNAs. *Nat. Struct. Mol. Biol.* **19**, 577-585 (2012).
61 Di Giammartino, D. C., Nishida, K. & Manley, J. L. Mechanisms and consequences of alternative polyadenylation. *Mol. Cell* **43**, 853-866 (2011).
62 Sarafianos, S. G., Das, K., Tantillo, C., Clark, A. D., Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H. & Arnold, E. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* **20**, 1449-1461 (2001).
63 Goñi, J. R., de la Cruz, X. & Orozco, M. Triplex-forming oligonucleotide target sequences in the human genome. *Nucleic Acids Res.* **32**, 354-360 (2004).
64 Huppert, J. L. & Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **35**, 406-413 (2007).
65 Betteridge, P. W., Carruthers, J. R., Cooper, R. I., Prout, K. & Watkin, D. J. CRYSTALS version 12: software for guided crystal structure analysis. *J. Appl. Crystallogr.* **36**, 1487 (2003).
66 Carruthers, J. R. & Watkin, D. J. A weighting scheme for least-squares structure refinement. *Acta Crystallogr. Sect. A* **35**, 698-699 (1979).
67 Sigmund, H. & Pfleiderer, W. Nucleotides. Part LXXI. *Helv. Chim. Acta* **86**, 2299-2334 (2003).
68 Lee, J., Jeong Jeong, E. & Kim, J. Selective and sensitive detection of melamine by intra/inter liposomal interaction of polydiacetylene liposomes. *Chem. Commun.* **47**, 358-360 (2011).
69 Berl, V., Krische, M. J., Huc, I., Lehn, J.-M. & Schmutz, M. Template-induced and molecular recognition directed hierarchical generation of supramolecular assemblies from molecular strands. *Chem. Eur. J.* **6**, 1938-1946 (2000).

3| Mechanisms and parameters underpinning cyanuric acid-mediated poly(adenine) assembly



This chapter is composed of work published as "Reprogramming the assembly of unmodified DNA with a small molecule" by N Avakyan, AA Greschner, F Aldaye, CJ Serpell, V Toader, A Petitjean & HF Sleiman; Nat. Chem. ***8****, 368-376 (2016) (sections 3.2.1-3.2.3), work towards "Mapping the energy landscapes of supramolecular assembly by thermal hysteresis" by RW Harkness V, N Avakyan, HF Sleiman and AK Mittermaier; Nat. Commun.* ***9****, 3152 (2018) (section 3.2.5) and additional unpublished results (in sections 3.2.3, 3.2.4 and 3.2.5).*

Piece by piece

Model inspired by modular 3D origami based on triangular units

Folding, photography and image editing: Nicole Avakyan

Contributions

NA designed experiments, obtained all the experimental data used in this chapter, and performed analysis and interpretation of results. **Prof. Anthony K. Mittermaier** initiated the project aiming to use thermal hysteresis analysis to study supramolecular assembly mechanisms, provided research goals, experimental supervision, guided methodology development, data analysis and interpretation of results (results section 3.2.5 and experimental sections 3.4.10 through 3.4.12). **Robert W. Harkness, V** developed the thermal hysteresis data analysis methodology, performed the calculations (experimental sections 3.4.10 through 3.4.12) and contributed to interpretation of results (section 3.2.5), while **NA** acquired all experimental data and contributed results interpretation.

3.1 Introduction

The discovery of CA-mediated assembly of unmodified d(A_n) strands into nanofibers was described in Chapter 2. The internal structure of the assembly is proposed to be composed of stacked hexameric rosette arrays, formed by hydrogen bonding association between CA and adenine residues in a 1:1 stoichiometry. The material generated through CA-mediated assembly of d(A_n) strands not only has a unique internal structure, its robust elongation into micron-long fibers encourages its investigation for application as a nanomaterial.

CA-mediated assembly can be described as a supramolecular polymerization, or the non-covalent assembly of molecular building blocks into an extended 1D structure. Supramolecular polymers are ubiquitous in nature and have important biological functions. Cytoskeletal filaments such as actin or microtubules control cell shape and mechanics;[1,2] fibers composing the extracellular matrix, including collagen, provide structural support to cells and tissues, as well as possess important adhesive and signalling properties;[3,4] amyloid fibrils are famously involved in the development of neurodegenerative diseases such as Alzheimer's, but can also be functional in some cases where their formation is tightly controlled;[5,6] and protein filaments encase viral genomes to protect them.[7] Inspired by natural systems and the desire to build materials with new advanced functions, over the last thirty years researchers have developed a large variety of rationally designed self-assembling building blocks capable of forming multiple types of

supramolecular polymers both in organic and aqueous environments.[8,9] By nature of their non-covalent formation, supramolecular polymers have many attractive properties such as adaptive behaviour, stimuli-responsiveness and hierarchical organization.[9-11] Some of the common extended 1D morphologies for supramolecular polymers include random coils,[12] and shape-persistent highly organized architectures such as cylindrical aggregates,[13] helical ribbons[14,15] and nanotubes.[16-18] The potential of using such assemblies as functional materials keeps expanding with the development of new structures and assembly methodologies. Promising areas of application for functional nanostructures based on supramolecular organization include bioactive fibrous scaffolds and hydrogels for tissue engineering and regenerative medicine,[19,20] as well as semiconducting and photoconductive nanowires for advanced electronic devices.[21,22]

Understanding the mechanisms of supramolecular polymerization is critical to the development of functional nanomaterials based on these systems. Properties such as polymer size distribution, robustness and stability to changing conditions in the medium of assembly are highly dependent on assembly mechanism, which itself is largely determined by monomer structure.[8,23,24] As changes in these properties affect the functionality of the material, and thus its applicability in a given setting, investigating the thermodynamic and kinetic aspects of assembly, as well as the influence of various parameters on them represents key insight in the development of functional materials. Isodesmic and cooperative growth are the two main mechanisms of supramolecular polymerization.[25] From the point of view of thermodynamics, the isodesmic mechanism is described by a single association constant K, such that each monomer addition to the growing chain occurs with equal affinity and there is poor control of molecular weight distribution. The cooperative mechanism, often referred to as nucleation-growth, is characterized by two assembly phases: a nucleation phase which proceeds with an association constant K_n until a certain pre-polymer length is reached, and an elongation phase, often triggered by structural changes in the nucleus, which proceeds with a more favorable association constant K_e.[8,26] In systems exhibiting cooperative polymerization, parameters such as monomer concentration or temperature have a dramatic influence on assembly outcome. In addition, cooperative systems often exhibit kinetic effects such as thermal hysteresis, a lag in equilibration of forward and reverse reactions resulting from unmatched reaction rates due to a kinetic barrier such as homogeneous nucleation.[8,27-30] In recent years, the kinetic aspects of supramolecular polymerization processes have come under

increased scrutiny, as there is untapped potential to achieve high levels of assembly outcome control by exercising kinetic control on the system.[31-35]

Chapter 3 explores several aspects of the mechanism of CA-mediated fiber growth along with the parameters that influence the assembly process. Strand alignment within the fiber was examined by using a simple experiment based on the fluorescence energy transfer between two chromophores. Furthermore, through a series of spectroscopic studies where the concentrations of both assembly components, d(A_n) strands and CA, were systematically varied to influence the extent of assembly, the formation of CA-mediated products was shown to proceed following a cooperative, or nucleation-elongation, supramolecular polymerization mechanism. CA concentration emerged as a crucial parameter determining assembly outcome, as a threshold concentration was required to trigger fiber formation. The significant influence of CA concentration was also observed in its impact on thermal stability of CA-mediated d(A_n) structures. A second important parameter, d(A_n) strand length, also influenced thermal stability and, interestingly, multiple melting transitions were observed with strands longer than d(A_{15}). The collection of observations regarding variations in CA concentration and d(A_n) strand length suggests that these parameters profoundly affect product distribution and thermal stability, and thus can be used to tune the properties of the CA-mediated nanofibers.

A new data analysis methodology developed by Harkness and Mittermaier and harnessing a common property among supramolecular systems, thermal hysteresis, provided a deeper, quantitative representation of the supramolecular polymerization process. By analysing thermal hysteresis profiles obtained at a series of different melting and annealing rates, maps of monomer (d(A_n) strand) consumption rates with respect to monomer concentration and temperature with a high level of detail could be built. These maps then revealed effective reaction orders over a range of temperatures in a model-free manner. Global fitting of this ensemble of data to a cooperative polymerization mechanism model was used to determine nucleation and elongation rate constants, activation energies, and to build quantitative energy diagrams to describe the assembly pathway. This treatment revealed insights about nucleus molecularity and shed new light on the impact of CA concentration and d(A_n) length both on kinetic and thermodynamic aspects of the assembly process.

3.2 Results and discussion

3.2.1 Assessment of strand alignment within nanofibers

The first aspect of the mechanism of fiber formation to consider is the organization of d(A_n) strands within the structure. There are two potential strand arrangements in the long-range organization of short oligonucleotides into fibers (**Figure 3.1**). The first possibility is the end-to-end π-stacking of finite structures (blunt-end stacking), driven by the tendency to bury the large hydrophobic surface of the CA:A rosette exposed at the blunt end of a fiber seed. The other possible mechanism presupposes a phase shift between the participating DNA strands during the growth process. The term 'phase shift' refers to staggered oligonucleotide assembly, producing overhangs that are further capable of recruiting strands and CA molecules and bringing them together longitudinally into fibers. This mechanism would be statistically favored, allowing for a variety of strand alignment conformations.

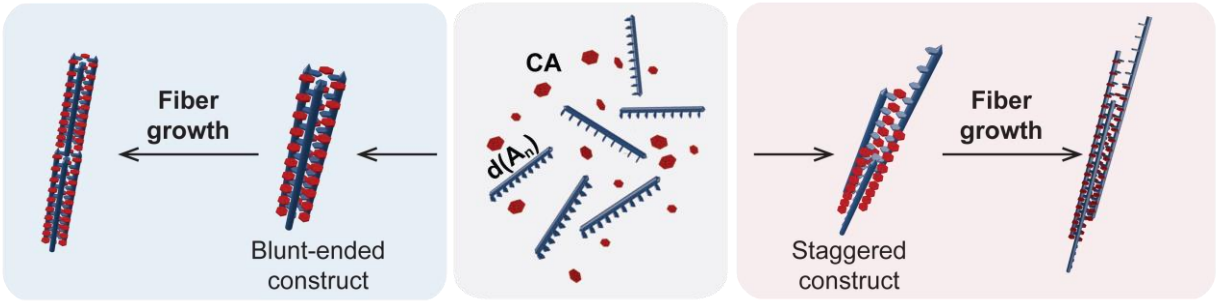


Figure 3.1| Potential modes of oligonucleotide organization into fibers. Fiber growth can occur through π-stacking of blunt-ended constructs (left) or through the formation of staggered constructs where fiber growth is further mediated by the available strand overhangs (right).

In order to examine the type of strand organization in CA-mediated fiber growth, an experiment based on fluorescence resonance energy transfer (FRET) was conceived. **d(A_{20})-2F**, a test d(A_n) strand labelled with a Cy3 dye at its 5'-end, and a Cy5 dye in the middle of the strand (after the 10th adenosine unit) (**Figure 3.2 a**), was designed. Depending on whether fiber growth proceeds via blunt-end stacking or staggered assembly, different FRET behaviors were predicted for the dye pair in **d(A_{20})-2F** upon CA-mediated assembly. Due to increased proximity between donor and acceptor molecules in the case of staggered assembly, FRET efficiency (E_T) was

expected to be higher than in the case of blunt-ended stacking, where identical dye molecules would be positioned next to each other and the Cy3 and Cy5 dyes would remain separated by 10 bases (**Figure 3.2 a**). To set benchmarks for the two possible cases, fluorescently labeled mixed sequence DNA duplexes, where the donor and acceptor molecules were preorganized in distinct blunt and staggered arrangements (**Figure 3.2 b**), were prepared as controls.

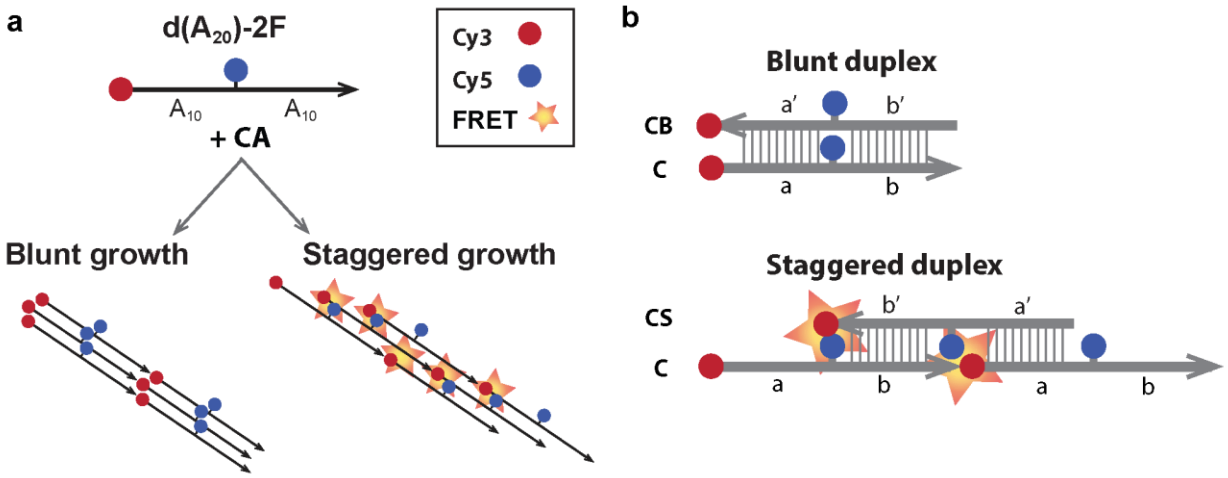


Figure 3.2| Test system and control duplexes for determination of oligonucleotide alignment in CA-mediated d(A$_n$) assemblies. a) Schematic representation of dually dye-labelled strand **d(A$_{20}$)-2F** and its potential FRET behavior upon CA-mediated assembly via blunt-ended stacking or staggered fiber growth. **b)** Schematic representations of fluorescently labeled strands composing the control duplexes preorganized to assemble into a blunt or a staggered construct.

The dye modified test strand **d(A$_{20}$)-2F**, as well as the strands composing the control duplexes (**C**, **CB** and **CS**) were characterized by denaturing polyacrylamide gel electrophoresis (PAGE) (**Figure 3.3 a**) and liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) (**Figure 3.3 b**). Furthermore, the assembly of the test duplexes was verified by native PAGE (**Figure 3.3 c**). The native gel confirmed the formation of a single discrete duplex for the blunt duplex control (Lane 1) and a range of products of different lengths for the staggered control duplex (Lane 2). Strand sequence design in the staggered duplex control allows polymerization of the construct by exposing complementary sticky ends, however the positioning of the dye molecules with respect to each other remains consistent regardless of the length of the assembly product.

b

Name	Expected Mass	Measured Mass
d(A_{20})-2F	7237.684	7237.469
C	7234.582	7234.375
CB	7069.597	7069.375
CS	7069.507	7069.281

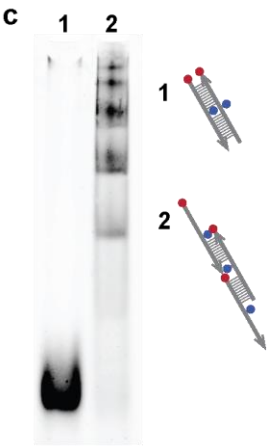


Figure 3.3| Characterization of fluorescently labelled strands and control duplexes. a) Denaturing PAGE analysis of dye-labelled strands. Lane description: (1) **d(A_{20})-2F**, (2) **C**, (3) **CB** and (4) **CS**. **b)** Table listing expected MW of dye-labelled strands and their MW measured by LC-ESI-MS analysis. **c)** Native PAGE analysis of control duplex constructs. Lane description: (1) Blunt duplex (**C** + **CB**) and (2) Staggered duplex (**C** + **CS**) as illustrated on the right.

Fluorescence emission spectra upon Cy3 excitation (λ_{Ex} = 548 nm) were obtained for **d(A_{20})-2F**, as well as for blunt and staggered duplexes with and without added CA (**Figure 3.4 a**). The fluorescence spectra for both blunt and staggered duplexes indicate that the presence of 10 mM CA does not alter the fluorescence response for these constructs. In the case of **d(A_{20})-2F**, the addition of 10 mM CA results in a drop in the fluorescence emission intensity across the whole spectrum, which could be attributed to fluorescence self-quenching with dyes coming in close proximity upon assembly. Similarly, self-quenching explains the magnitude of the fluorescence signal for the control duplexes which drops compared to the signal magnitude observed for single strands. The distinctive negative peak at 252 nm in the CD spectrum for **d(A_{20})-2F** with 10 mM CA confirms that the dye modified strand undergoes CA-mediated assembly (**Figure 3.4 b**, left). Furthermore, the spectra for the control duplexes are not affected by CA addition (**Figure 3.4 b**, center and right), as expected from previous studies with random sequence DNA strands. It is important to note that due to the lower sensitivity of the CD instrument, the CD spectra were obtained at 5 µM total strand concentration in contrast to the 0.5 µM concentration used in fluorescence studies, where the lower concentration was selected to avoid inner filter effects due to the attached dyes.

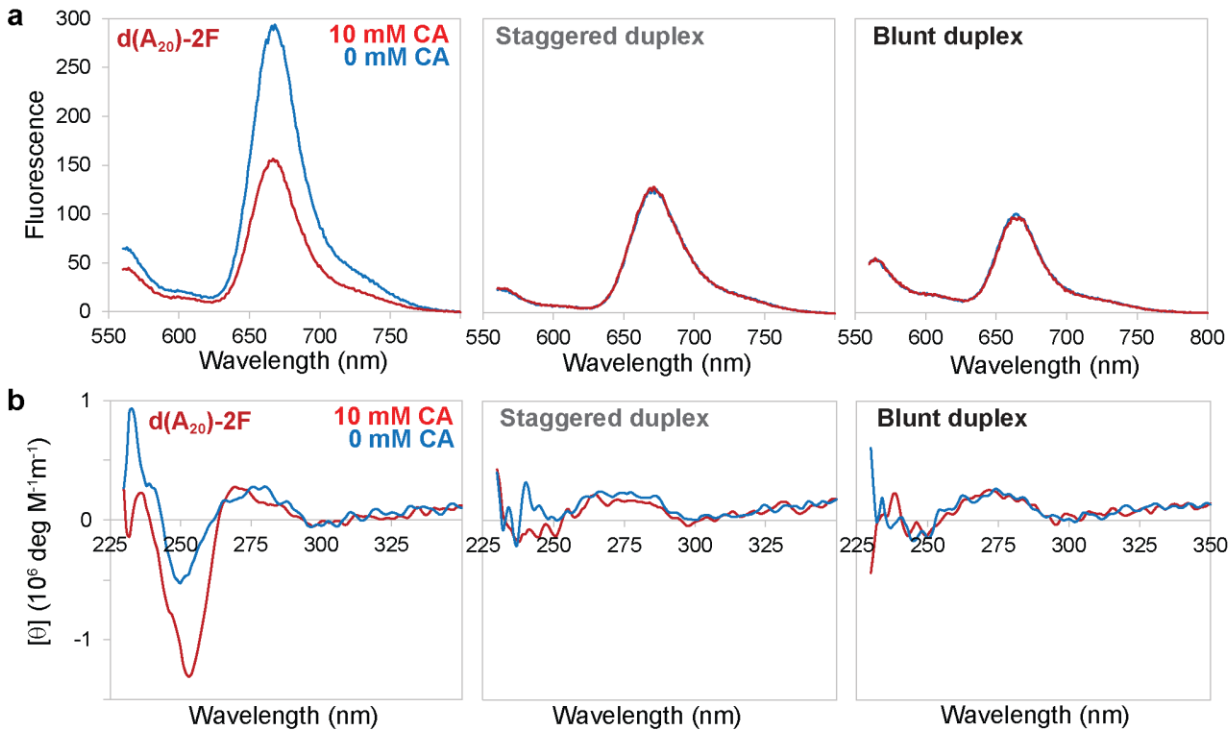


Figure 3.4| Fluorescence emission and CD spectra for test system and control duplexes. a) Fluorescence emission spectra acquired upon Cy3 excitation (λ_{Ex} = 548 nm) and **b)** CD spectra for **d(A$_{20}$)-2F** (left), staggered duplex (center) and blunt duplex (right); with 0 mM (blue) or 10 mM CA (red).

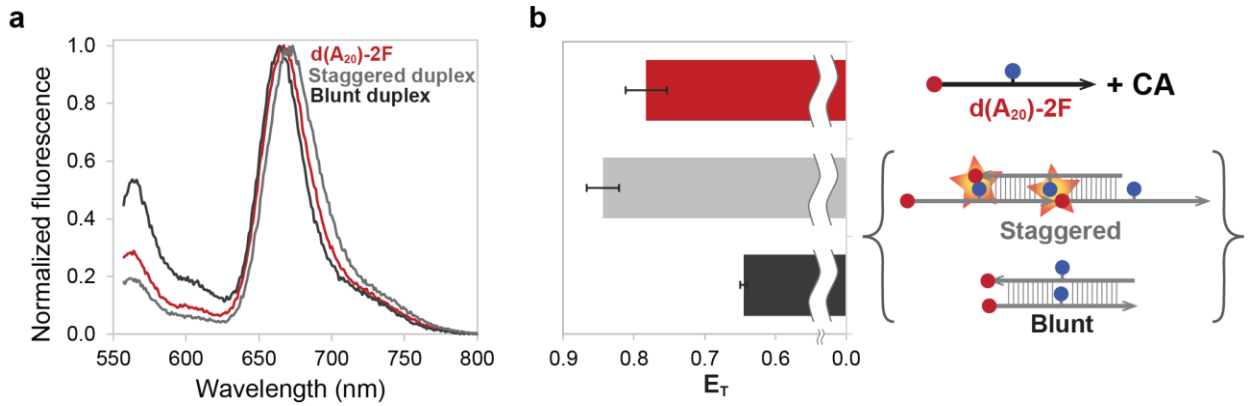


Figure 3.5| Determination of fiber strand alignment in CA-mediated assembly of d(A$_n$) by FRET. a) Normalized fluorescence spectra for **d(A$_{20}$)-2F** (red), staggered duplex (grey) and blunt duplex (black) with 10 mM CA. **b)** FRET efficiency comparison between the CA-mediated assembly of **d(A$_{20}$)-2F** and staggered or blunt duplexes (n = 3, 95 % CL).

The spectra for the different constructs, normalized to the emission intensity of Cy5, illustrate the differences in FRET behaviour (**Figure 3.5 a**). As predicted, the E_T calculated for the

blunt duplex control is significantly lower than for the staggered duplex control (**Figure 3.5 b**). The E_T for **d(A_{20})-2F** incubated with 10 mM CA (TAMg pH 6), is much higher than for the blunt control, although still significantly different from the E_T for the staggered construct. In comparison to the FRET behaviour of the control duplexes, CA-mediated assembly of **d(A_{20})-2F** behaviour suggests that CA-mediated fiber growth is more likely to occur through a staggered rather than blunt strand alignment within the fiber.

3.2.2 Supramolecular polymerization mechanism determination

Having examined the strand alignment within the fiber, it was next necessary to determine the process by which the CA-mediated supramolecular polymerization of d(A_n) occurs. Supramolecular polymerization may proceed via one of two main mechanisms: isodesmic (step-growth), or cooperative (nucleation-growth) polymerization (**Figure 3.6**).[8] The isodesmic mechanism describes step-growth polymerization, where fibers would grow progressively with increasing CA or d(A_n) concentration ([CA] and [d(A_n)]). In contrast, the cooperative supramolecular polymerization mechanism is characterized by non-linear growth, is often nucleated and would display an onset of fiber growth above a threshold or critical concentration of component materials ($[CA]_{cr}$ and $[d(A_n)]_{cr}$).

To investigate the mechanism of polymerization, the assembly of d(A_{15}) was monitored with varying concentrations of both d(A_{15}) and CA (pH 4.5). CD (**Figure 3.7 a**) and UV-Vis (**Figure 3.7 b**) spectra were obtained for d(A_{15}) at concentrations ranging from 5 to 75 μM, incubated with 0 to 15 mM CA. Molar ellipticity traces obtained from CD show that at low [CA], the spectrum remains unchanged relative to d(A_{15}) in the absence of CA. However, with increasing [CA], the intensity of the negative peak at 252 nm characteristic of CA-mediated assembly increases sharply. This effect is more dramatic at higher d(A_{15}) concentrations ([d(A_{15})] = 25-75 μM). The intensity of the peak grows with increasing [CA] and then stabilizes at the highest [CA] measured. As described in **Chapter 2**, the appearance of the strong negative peak at 252 nm is accompanied by spectral changes in the 260 to 300 nm range, where two weak positive bands emerge. Parallel [CA]-dependent observations can be made by following the intensity of the molar absorptivity traces obtained from UV-Vis absorbance at 260 nm. A hypochromic shift is detected with increasing [CA] after an initial lag, however the transition is not as distinct as the one seen in

the molar ellipticity spectra. Furthermore, a hyperchromic effect is seen at 285 nm, revealing an isosbestic point at 280 nm.

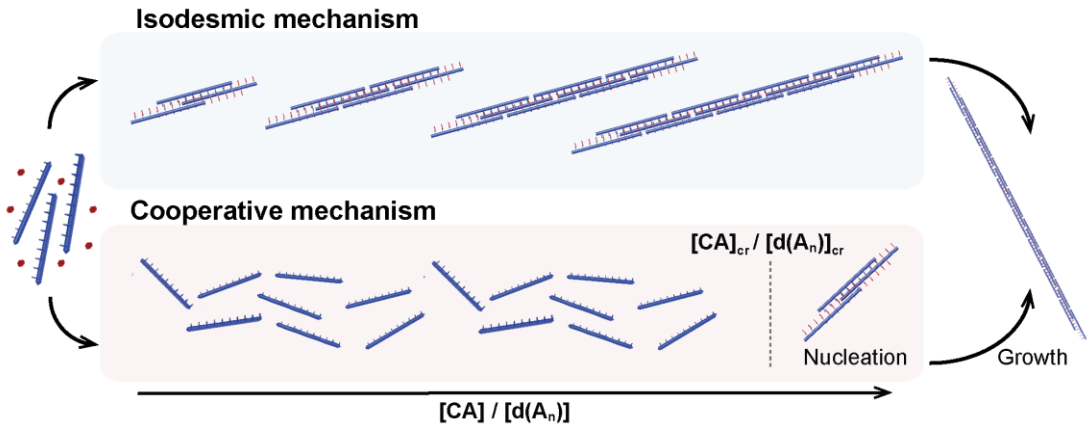


Figure 3.6| Representation of isodesmic *vs.* cooperative supramolecular polymerization mechanisms. The isodesmic polymerization mechanism (top) describes a process where fibers grow progressively with [CA] or [d(A_n)]. In contrast, the cooperative supramolecular polymerization mechanism (bottom) is characterized by non-linear growth with the onset of fiber growth occurring above a critical concentration, $[CA]_{cr}$ and $[d(A_n)]_{cr}$, following an initial nucleation step.

The magnitude of the distinctive CD peak at 252 nm or the absorbance at 260 nm can be converted to a fraction of assembled d(A_{15}) to examine the dependence of the extent of assembly on [CA] (**Figure 3.8 a**). The plots of fraction assembled *vs.* [CA] obtained from CD have a non-sigmoidal shape that clearly show that d(A_{15}) assembly is detected only above a certain threshold $[CA]_{cr}$. In the range of 25-75 μM d(A_{15}), assembly occurs above a $[CA]_{cr}$ of 3 mM, whereas at lower d(A_{15}) concentrations, the $[CA]_{cr}$ shifts to 4 mM and the transition becomes less sharp. This behavior suggests a cooperative polymerization mechanism (**Figure 3.8 b**), whereby elongation of fiber nuclei (likely small seed structures composed of staggered strands) to the observed higher-order assembly only occurs above $[CA]_{cr}$. Hud *et al.* demonstrated that dissociation of coralyne-induced poly(dA) duplexes into single-strands upon dilution occurs with a sharp transition as a function of concentration, indicating a similar cooperative assembly onset at a critical concentration of the small molecule.[36] The fraction assembled *vs.* [CA] plots obtained from UV-Vis absorbance (**Figure 3.8 c**) are more difficult to interpret, in part due to a lower signal to noise ratio associated with a smaller change in the signal characteristic of assembly, however the general shape of the curves remains non-sigmoidal.

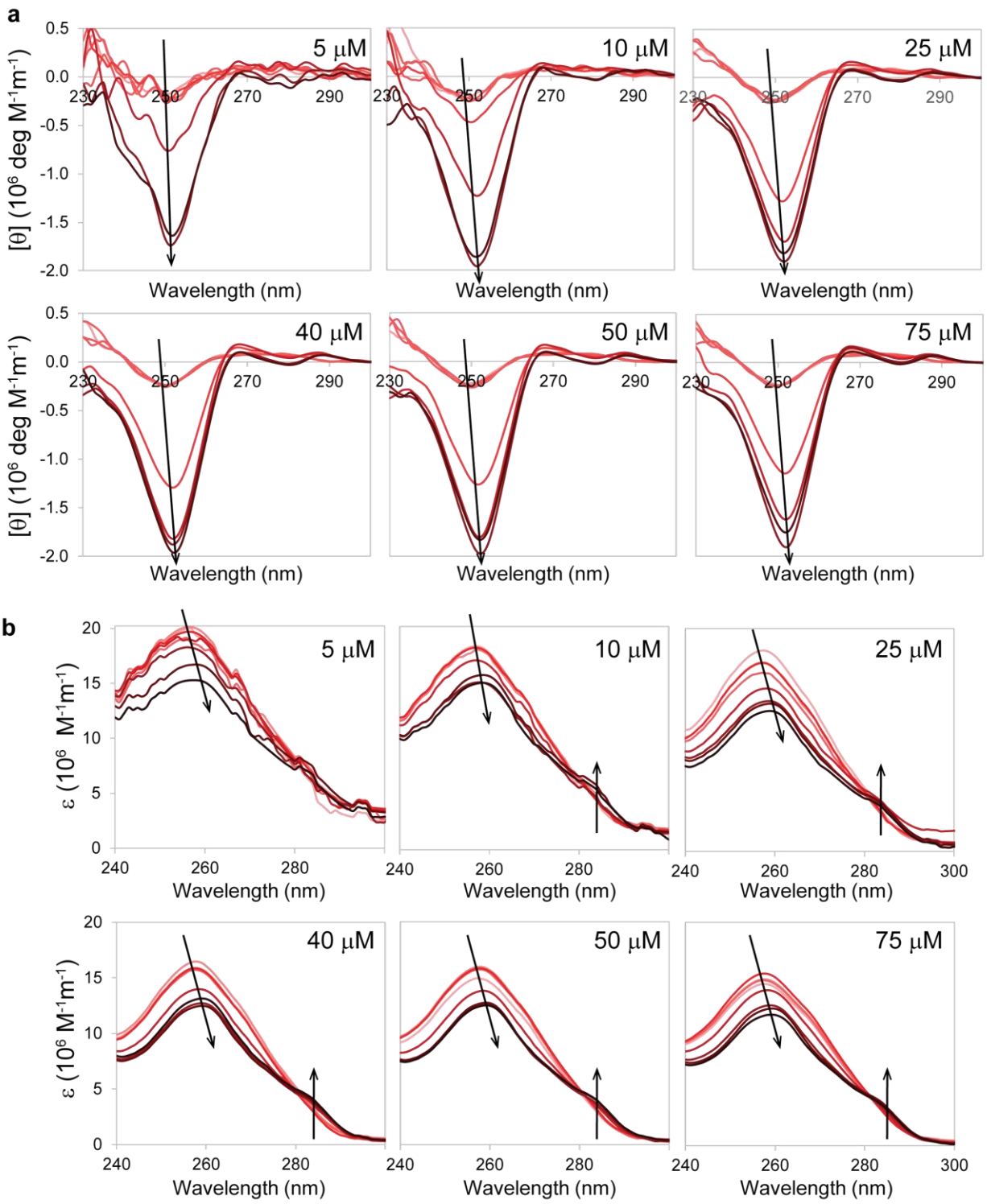


Figure 3.7| Combined effect of d(A_{15}) and CA concentrations on the extent of CA-mediated assembly. Spectra for d(A_{15}) (5, 10, 25, 40, 50 and 75 μM) mixtures with CA (0, 1, 2, 3, 4, 5, 10 and 15 mM) in TAMg pH 4.5. **a)** CD spectra (shown in normalized molar ellipticity units). **b)** UV-Vis absorbance spectra (shown in normalized molar absorptivity units). For **(a)** and **(b)** arrows indicate evolution of absorbance signal intensity with increasing [CA]. The colours of the curves change from light pink to deep burgundy with increasing [CA].

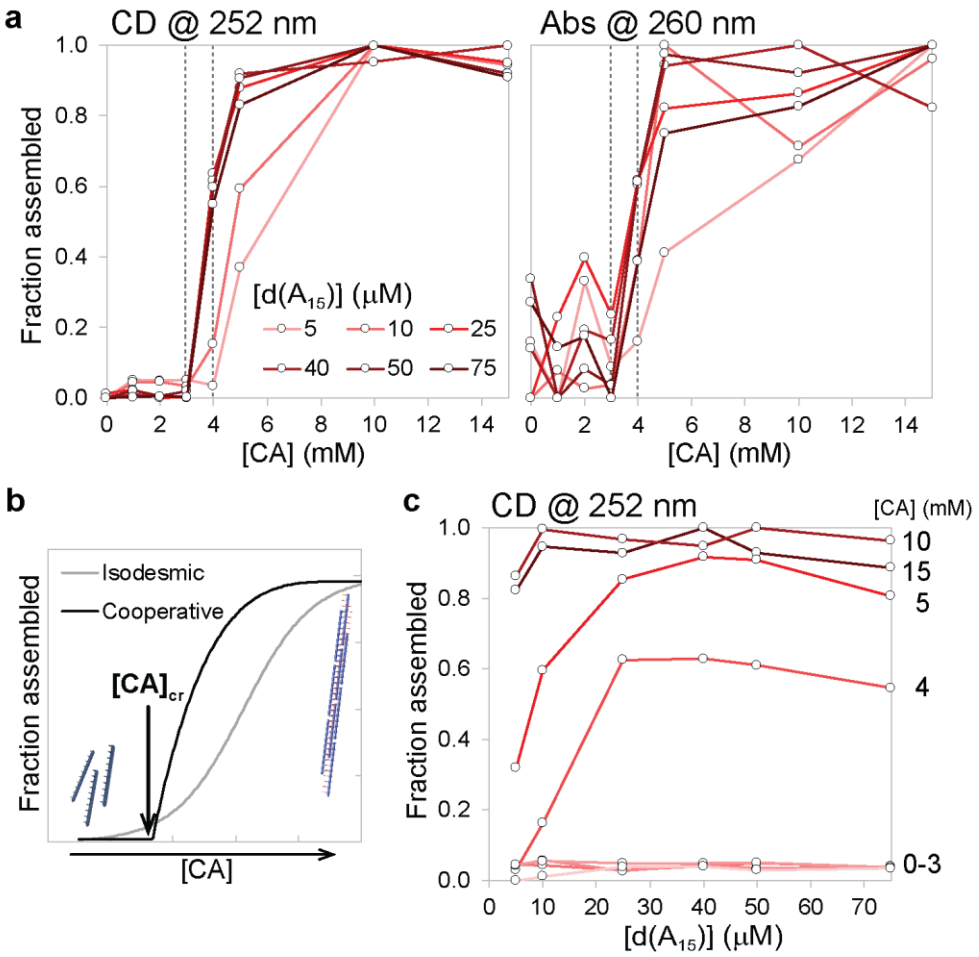


Figure 3.8| Determination of the mechanism of supramolecular polymerization of CA-mediated d(A_{15}) assembly. a) Fraction of d(A_{15}) assembled *vs.* [CA] curves at different d(A_{15}) concentrations (5, 10, 25, 40, 50 and 75 μM). Curves were obtained from CD spectra monitoring the peak at 252 nm (left) and from UV-Vis absorbance spectra at 260 nm (right). The colours of the curves change from light pink to deep burgundy with increasing [d(A_{15})]. **b)** Theoretical fraction of d(A_{15}) assembled vs. [CA] curves for isodesmic (grey) and cooperative (black) polymerization mechanisms. **c)** Fraction of d(A_{15}) assembled *vs.* [d(A_{15})] curves at different [CA] concentrations (0, 1, 2, 3, 4, 5, 10, 15 mM). Curves were obtained from CD spectra monitoring the peak at 252 nm. The colours of the curves change from light pink to deep burgundy with increasing [CA].

A $[CA]_{cr}$ of 3 mM is in excess of the stoichiometric 1:1 CA:A ratio featured in the proposed hexameric rosette structure, rather corresponding to a ratio of 2.67:1 CA:A (in the case of 75 μM d(A_{15})) or greater (for lower d(A_{15}) concentrations). It is possible that at very high d(A_{15}) concentrations (starting in the vicinity of 330 μM d(A_{15}), which is equivalent to 5 mM in adenine residues) assembly could be observed for 1:1 CA:A mixtures, as has been previously the case for

equimolar mixtures of derivatized TAP and CA,[37] however this prediction has not been experimentally verified due to instrumental limitations.

On the other hand, fraction of d(A_{15}) assembled plotted *vs.* [d(A_{15})] at different [CA] (**Figure 3.8 c**) is a less telling representation. The range of [d(A_{15})] that could be effectively studied by CD is insufficient to build a complete binding curve and so confirmation of the cooperative mechanism with respect to d(A_{15}) and determination of $[d(A_{15})]_{cr}$ could not be made. The plot does show that independent of [d(A_{15})], no assembly is detected for [CA] from 0 to 3 mM and that for [CA] of 10 and 15 mM, the fraction of assembled d(A_{15}) lies from 0.8 to 1. At 4 and 5 mM CA, the curves show in part the sharp transition from disassembled to assembled material, although a disassembled baseline cannot be captured due to insufficient data points in the low [d(A_{15})] range. Interestingly, the curves level off at intermediate values for fraction of d(A_{15}) assembled, indicating that d(A_{15}) fully participates in assembly only at higher [CA].

The [CA]-dependent switch from free d(A_{15}) to CA-mediated assembly observed by spectroscopic methods was mirrored by PAGE experiments (**Figure 3.9 a**). At low [CA] only single-stranded d(A_{15}) was observed, but with concentrations above 3 mM CA, the band shifted to lower mobility, indicating onset of assembly that can be associated with the formation of nuclei. There is a further slight decrease in the mobility of the band at 15 mM CA (rightmost lane). Although much larger structures are expected based on AFM and DLS results (see **Chapter 2**), extended structures are not observed by PAGE. It is possible that the short, staggered overhangs that allow for fiber elongation are disrupted as the assembly travels down the gel, leaving behind only the smallest assembly units.

The mobility of 15-mer duplex (**DS**) and triplex (**TS**) studied under the same PAGE conditions revealed that there is only a small difference in the gel mobility of these structures (**Figure 3.9 b**). Lane **DS+NB**, where **NB** is a non-binding 15-mer strand, shows the difference in mobility between sample 15-mer monomer and duplex structures. Considering these observations, the lower mobility band that appears in the PAGE experiment for d(A_{15}) in a range of [CA] (**Figure 3.9 a**), could represent either a duplex or triplex structure. A further mobility decrease (rightmost lane) could indicate the addition of another strand to this structure. Further experiments examining the mechanism of the assembly process, focusing on the stoichiometry of the nucleation and elongation stages would provide a better understanding of the structures involved. As described above, both PAGE and CD spectroscopy show that assembly only occurs above a [CA] threshold.

Moreover, results of the [CA]-dependent CD studies are consistent with a nucleation-elongation supramolecular polymerization mechanism.

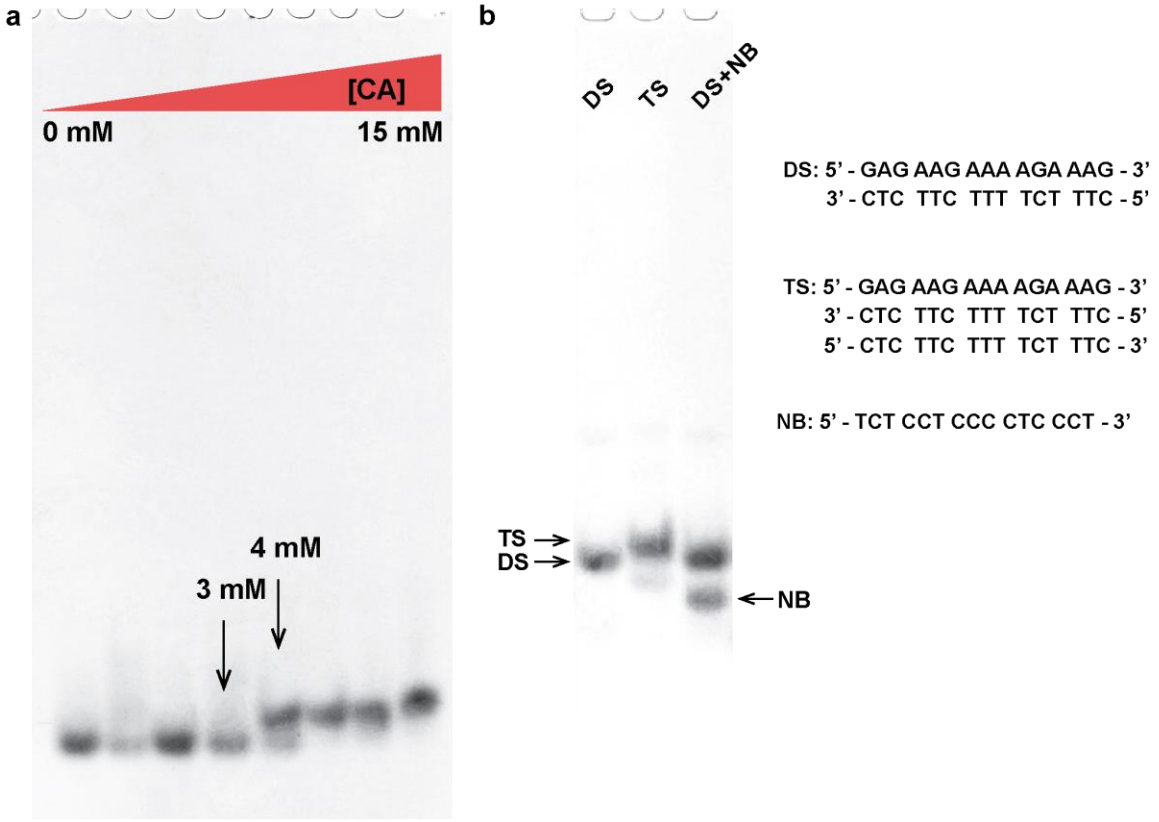


Figure 3.9| PAGE analysis of CA-mediated assemblies of d(A_{15}) in a range of CA concentrations and 15-mer triplex-forming controls. a) 50 μM d(A_{15}) in a range of CA concentrations (0, 1, 2, 3, 4, 5, 10 and 15 mM) on 8% native PAGE. **b)** Sequences and gel mobility of 15-mer triplex forming controls: DS (double stranded), TS (triple stranded), DS + NB (double stranded with non-binding third strand) on 8% native PAGE.

Thermal denaturation of samples used in determining the mechanism of supramolecular polymerization ([CA] ≥ 4 mM, where assembly was detected) indicated that the T_M of assemblies is influenced by both [CA] and [d(A_{15})], but to a different extent. The T_M values at a given [CA] were weakly affected by changes in [d(A_{15})], with variations of only 0.5 to 3.5 °C across the examined concentration range ([d(A_{15})] from 5 to 75 μM, or a 15-fold increase) (**Figure 3.10**). On the other hand, the effect of changing [CA] from 4 to 15 mM (only a 3.75-fold increase) produced a dramatic change in average T_M across the [d(A_{15})] range from 13.7 to 33.5 °C. This observation may be explained by the stoichiometric relationship of CA and d(A_{15}) strands in the system.

Although each rung of the proposed triplex structure (the hexameric rosette unit) requires a 1:1 CA:A ratio, the CA:d(A_{15}) strand ratio is 15:1 over the length of the fiber. Overall, the amount of CA molecules participating in fiber formation is 15 times larger than the amount of d(A_{15}) molecules, so increasing [CA] drives fiber formation by increasing the probability of filling the available binding sites along each strand in the emerging structure. Thus, [CA] appears to be the most important parameter affecting the extent and stability of assemblies.

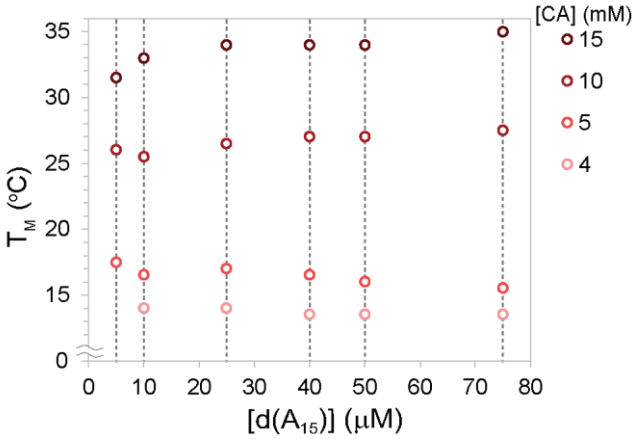
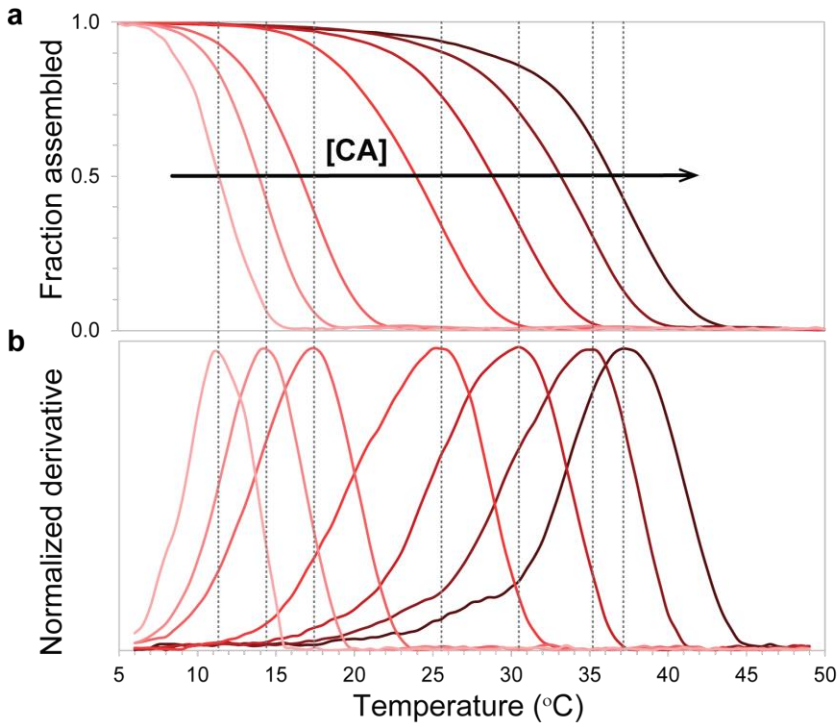


Figure 3.10| T_M values for CA-mediated assemblies of d(A_{15}) in a range of d(A_{15}) and CA concentrations. T_M values obtained from thermal denaturation following the CD signal at 252 nm. [d(A_{15})]: 5, 10, 25, 40, 50 and 75 μM; [CA]: 0, 1, 2, 3, 4, 5, 10, 15 mM. The colours of the points change from light pink to deep burgundy with increasing [CA].

3.2.3 Additional effects of [CA] and d(A_n) strand length on CA-mediated assembly

A closer examination of the effect of [CA] on the thermal denaturation of CA-mediated assemblies of d(A_{15}) was made by looking at the behavior of 40 μM d(A_{15}) using more [CA] data points. Above $[CA]_{cr}$, the T_M values determined from the melting curves, collected at 252 nm by CD spectroscopy, increased with increasing [CA] (**Figure 3.11 a** and **c**). The correlation between [CA] and T_M seemingly follows a logarithmic relationship (R^2 = 0.99). The stabilization effect occurring with increasing [CA] even at very high concentrations, suggests that, in addition to the impact of CA:d(A_{15}) ratio discussed above, the high CA content suppresses, or delays the melting process. Furthermore, CA molecules may be participating in additional stabilizing interactions

with the fibers, possibly through non-specific binding to their exterior or stacking within the central cavity of the triplex.



c

[CA] (mM)	3.6	4.8	6.0	9.0	12	15	18
T_M (°C)	11.3 ± 0.7	14.3 ± 0.7	17.3 ± 0.7	25.5 ± 1.2	30.5 ± 0.6	35.2 ± 1.4	37.2 ± 1.9
FWHM (°C)	4.9 ± 0.1	5.9 ± 0.1	7.2 ± 0.2	9.6 ± 0.6	9.6 ± 0.3	9.1 ± 0.5	8.4 ± 1.6

Figure 3.11| Thermal denaturation studies of d(A_{15}) structures assembled above $[CA]_{cr}$ monitored by CD at 252 nm. a) Melting curves for CA-mediated assemblies of d(A_{15}) in a range of CA concentrations above $[CA]_{cr}$ (3.6 mM to 18 mM), where assembly can be detected. **b)** Normalized derivatives for the melting curves shown in **(a)**. **c)** Summary of T_M and FWHM values from melting curves shown in **(a)** (n = 3, 95 % CL). The colours of the curves change from light pink to deep burgundy with increasing [CA].

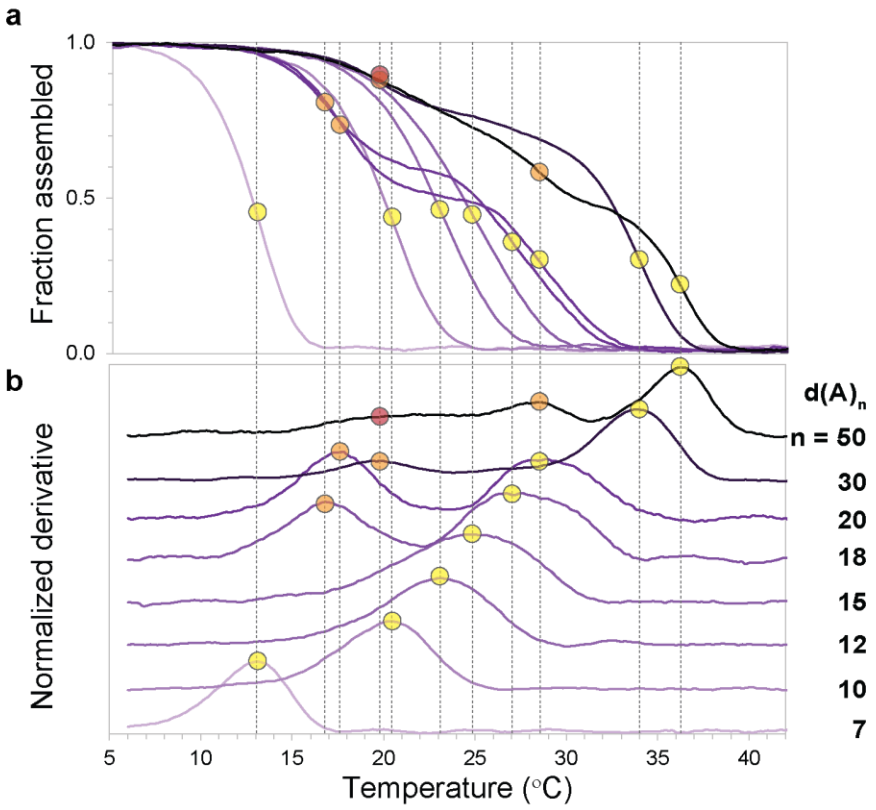
The changes in T_M values with increasing [CA] were accompanied with a broadening of the melting transition best observed in the normalized derivatives of the melting curves and full width at half maximum (FWHM) values extracted from them (**Figure 3.11 b** and **c**). FWHM is a marker of the cooperativity of the melting transition, as well as the activation energy required for melting to occur.[38] In the present case, where a mixture of supramolecular polymer products is

studied, it can also be interpreted as an indicator of polydispersity. The melting curve derivative broadens as [CA] increases from 3.6 to 9.0 mM. With further increases in [CA], the peak acquires a shoulder towards the lower temperature side, potentially indicating that a main population of fibers coexists with another smaller population with a lower T_M.

This phenomenon can be explained by considering the influence of [CA] on the nucleation-elongation mechanism. With [CA] as the main driving force behind fiber formation, the rate of nucleation at low [CA] is expected to be slow and accelerate with increasing [CA]. Although slow nucleation is generally associated with a broad product distribution,[8] and thus a wide melting transition, it is important to further consider the impact of [CA] on the stability of nuclei. On the way to nuclei formation, $d(A_n)$ strands may bind to each other in a range of registers, creating regions of overlap of variable length, and by consequence variable stability. Enthalpic stabilization of the nuclei is highly dependent both on hydrogen bonding between adenosine and CA units, and on the accumulation of π-stacking interactions between consecutive newly-formed adenine-CA rosettes. In this context, at low [CA], only nuclei with optimal strand overlap are stable enough to elongate rather than fall apart. This situation creates a population of fibers with a uniform internal structure that originates from a narrow selection of possible nuclei, which is reflected in a narrow melting transition. In contrast, the abundance of CA molecules at high [CA] conditions stabilize nuclei with a greater variety of strand alignments which then proceed to elongate into fibers with variable internal structure and thus melting behaviour. The consequence is a broader melting transition and the emergence of a shoulder in the melting derivative peak observed for [CA] ≥ 12 mM. In summary, [CA] affects both the thermal stability of $d(A_n)$ fibers and the distribution of fiber products.

The length of the $d(A_n)$ oligonucleotide also influences thermal denaturation properties of CA-mediated assembly. Melting curves were obtained for $d(A_n)$ strands with lengths ranging from n = 7 to n = 50 upon incubation with 10 mM CA (**Figure 3.12 a**), where the concentration of adenine residues was kept constant at 150 μM (such that strand concentration varied depending on strand length but the CA:A ratio remained constant). The derivatives of the melting curves (**Figure 3.12 b**) clearly show that for $d(A_n)$ strands where $n \leq 15$, the structures displayed monophasic melting transitions. However, with increasing strand length ($n \geq 18$), multistep transitions emerged. Across the range of strand lengths examined, the highest temperature melting transitions (yellow markers) increased progressively with $d(A_n)$ length, seemingly following a logarithmic

dependence (R^2 = 0.96). Importantly, assemblies formed with all d(A_n) strands with multiphasic melting transitions retain the same spectral CD features as the shorter, monophasic d(A_n) strands (**Figure 3.13**). This suggests that the internal structure within the assemblies remains unchanged, independent of strand length.



c

n	7	10	12	15	18	20	30	50
T_M (°C)	13.1 ± 1.3	20.5 ± 1.6	23.1 ± 1.7	24.9 ± 0.8	16.7 ± 1.0	17.6 ± 0.9	19.7 ± 1.9	19.9 ± 4.0
					27.1 ± 1.6	28.6 ± 1.3	34.1 ± 1.3	28.5 ± 0.6
								36.2 ± 1.0
FWHM (°C)	4.6 ± 0.2	5.6 ± 0.4	6.6 ± 1.2	8.8 ± 0.5	4.6 ± 0.8	4.8 ± 0.3	4.7 ± 1.3	4.1 ± 0.4
					7.6 ± 1.1	6.1 ± 0.9		

Figure 3.12| Thermal denaturation studies of d(A_n) assemblies monitored by CD at 252 nm. a) Melting curves for CA-mediated assemblies of d(A_n) (n = 7, 10, 12, 15, 18, 20, 30 and 50) with 10 mM CA. Yellow, orange and red circles denote different melting transitions. **b)** Normalized derivatives for the melting curves shown in **(a)**. **c)** Summary of T_M and FWHM values from melting curves shown in **(a)** (n = 3, 95 % CL). The colours of the curves change from light to deep purple with increasing d(A_n) length. Markers for melting transitions: yellow marks the highest temperature transition, orange – a lower one, and red an even lower one (only used for n = 50).

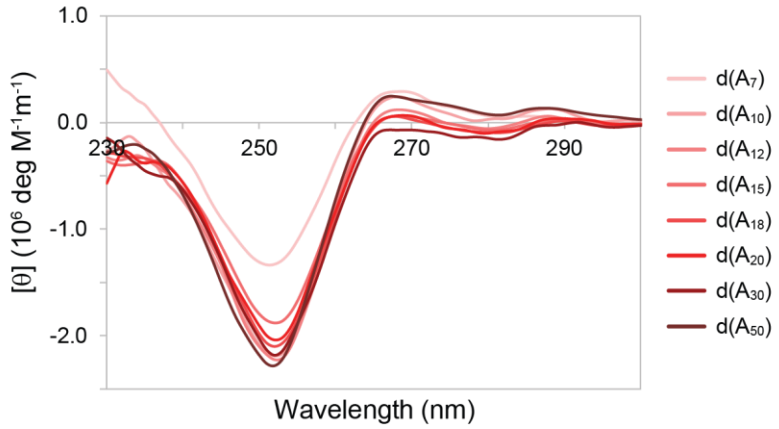


Figure 3.13| CD spectra (shown in normalized molar ellipticity units) for d(A_n) strands (n = 7, 10, 12, 15, 18, 20, 30 and 50) with 10 mM CA.

Several lines of reasoning may be invoked to justify the observed d(A_n) length-dependent trends in T_M and FWHM values. First, similarly to duplex DNA, longer regions of uninterrupted overlap between strands should provide greater enthalpic stabilization to the assembled structure because of π-stacking interactions between rosette surfaces within the fiber and hydrogen bonding, such that the highest T_M would be observed with the longest strands. Furthermore, in order to ensure polymerization with short strands, strands in the nuclei are likely shifted with respect to each other uniformly, creating a population of fibers that melt around the same temperature. The uniform binding register would ensure that the remaining overhangs are long enough to promote elongation. With longer strands, there are more possibilities of binding register during the nucleation stage. The resulting collection of nuclei contains structures with a range of different regions of overlap that are capable of elongation into a fiber population that now presents a range of melting behaviours. The trends observed in T_M and FWHM values (**Figure 3.12 c**), where an initial broadening with increasing strand length is followed by the splitting of the peak into multiple melting transitions, reflect the shift in the internal makeup of the fiber population. Similarly to the nucleus stabilization effect of increasing [CA], longer regions of overlap available with increasing d(A_n) strand length stabilize a larger variety of nuclei so they can elongate into fibers and increase the top melting temperature for a given strand length.

Interestingly, shorter fibers were observed by AFM with increasing d(A_n) length (**Figure 3.14**). This observation can once again be related to differential, strand length-dependent nucleus stabilization, with long strands initially forming a greater number of nuclei stable enough to

elongate. Proportionally, as the number of nuclei increases for long strands, fiber length decreases. Another possibility in line with melting transition observations is that longer strands do not participate in rosette units for their entire length, such that they remain partially free of interactions with CA and dangle on the outside of the fibers, potentially destabilizing them, increasing polydispersity, giving rise to lower temperature melting transitions and introducing shorter fiber populations. Furthermore, as longer strands are more flexible, they may be able to participate in intramolecular capping interactions that would also produce greater numbers of short fibers. However, this situation may be less likely due to the evidence for a preferred all parallel d(A_n) strand orientation within the CA-mediated structure (see **Chapter 2**).

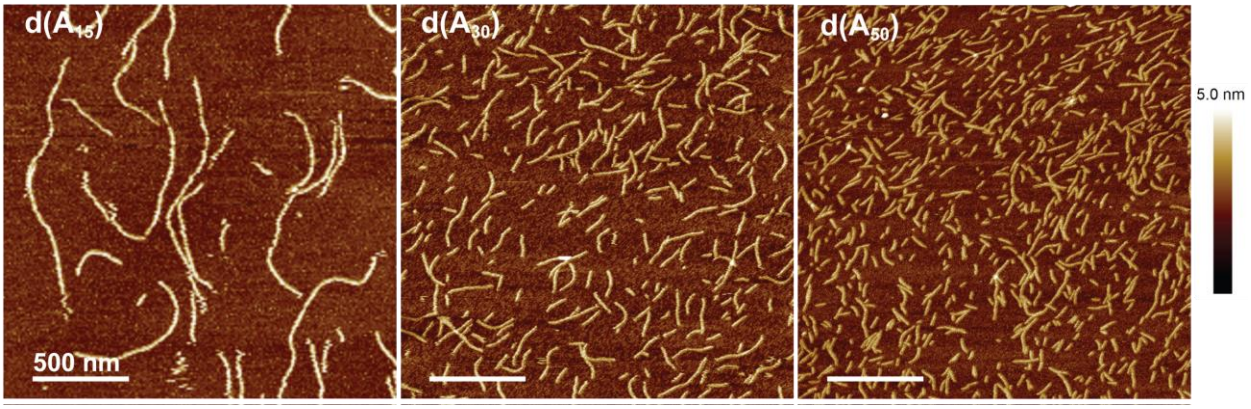


Figure 3.14| AFM images of CA-mediated assemblies prepared from d(A_n) strands of different lengths. Representative images obtained in solution for d(A_{15}) (left), d(A_{30}) (centre) and d(A_{50}) (right).

The combined effect of [CA] and d(A_n) length on $[CA]_{cr}$ and T_M trends was also examined. Fraction assembled *vs.* [CA] curves were obtained for d(A_n) strands with various lengths (n = 7 to 30) incubated in a range of [CA] concentrations (**Figure 3.15**). DNA concentration was normalized to a total [A] of 150 μM to maintain the CA:A ratios constant across strands of different lengths. Overall, there was a decrease in $[CA]_{cr}$ with increasing strand length. For short strands, d(A_7) and d(A_{10}), assembly was only observed above 7.5 and 5 mM CA respectively. The results for d(A_{20}) lie somewhat outside of this trend, as assembly was observed to a larger extent at 4 and 5 mM CA compared to the longer strand d(A_{30}). In this context again, the significant impact of [CA] on nucleus stability is evident in the increased $[CA]_{cr}$ values required for polymerization of short d(A_n) strands.

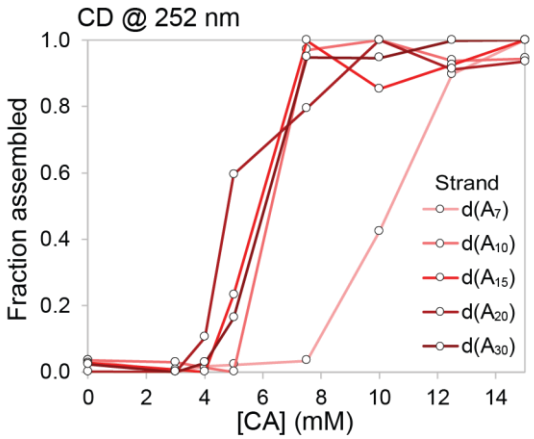


Figure 3.15| Fraction of d(A_n) assembled *vs.* [CA] curves for different d(A_n) strands (n = 7, 10, 15, 20 and 30). Curves obtained from CD spectra monitoring the peak at 252 nm. The colours of the curves change from light pink to deep burgundy with increasing d(A_n) length. Total [A] in each sample normalized to 150 µM (equivalent to 10 µM d(A_{15})). [CA] = 0, 3, 4, 5, 7.5, 10 and 15 mM in TAMg pH 4.5

The trends in thermal denaturation for strands of various lengths at different [CA] were replicated from the previous experiment, and new observations emerged (**Figure 3.16**). Multiple melting transitions were indeed recorded for d(A_n) strands with n ≥ 15, however not at all [CA]. For example, d(A_{20}) had a single melting transition at 5 mM CA, which significantly broadened and shifted to a higher temperature at 7.5 mM CA, only to split into 2 distinct transitions at even higher [CA]. At each [CA], the T_M for the top melting transition for each strand increased progressively with strand length. Interestingly, at [CA] from 10 to 15 mM, the main lower temperature melting transitions for d(A_{20}) and d(A_{30}) consistently nearly matched the T_M of d(A_{10}). This observation suggests that populations of fibers formed from d(A_{10}), d(A_{20}) and d(A_{30}) may have regions of overlap of similar length inside the structure. It is possible that clustering of T_M values thus emerges from thermal denaturation of regions of overlap of similar length within the fiber. Such regions may be triplex stretches within the fiber uninterrupted by nicks. We refer to these seemingly quantized overlap regions as "packets". The consistent match between T_M values observed for d(A_{10}) and the lower temperature transitions for d(A_{20}) and d(A_{30}) suggests that certain packet sizes may be preferred during the nucleation process for yet unknown structural reasons (here there seems to be a dependence on 10-mer repeats) and have a higher probability of elongation. Given the thermal denaturation results, the populations of fibers assembled from longer strands such as d(A_{20}) and d(A_{30}) seem to contain large packets that are strand length dependent,

giving rise to the top melting transition, and smaller packets, that give rise to lower temperature transitions and may be matching in length for the d(A_n) studied here. It should be noted that the T_M values of the same packets increased with increasing [CA]. In addition, the larger packet sizes seem to be favored with increasing [CA], as observed with strands that display multiphase-transitions (d(A_{20}) and d(A_{30})), where the top transition became more populated with increasing [CA]. Overall, the parallel variation of d(A_n) length and [CA] in thermal denaturation experiments has revealed additional clues on the distribution of fiber populations and their tunable characteristics.

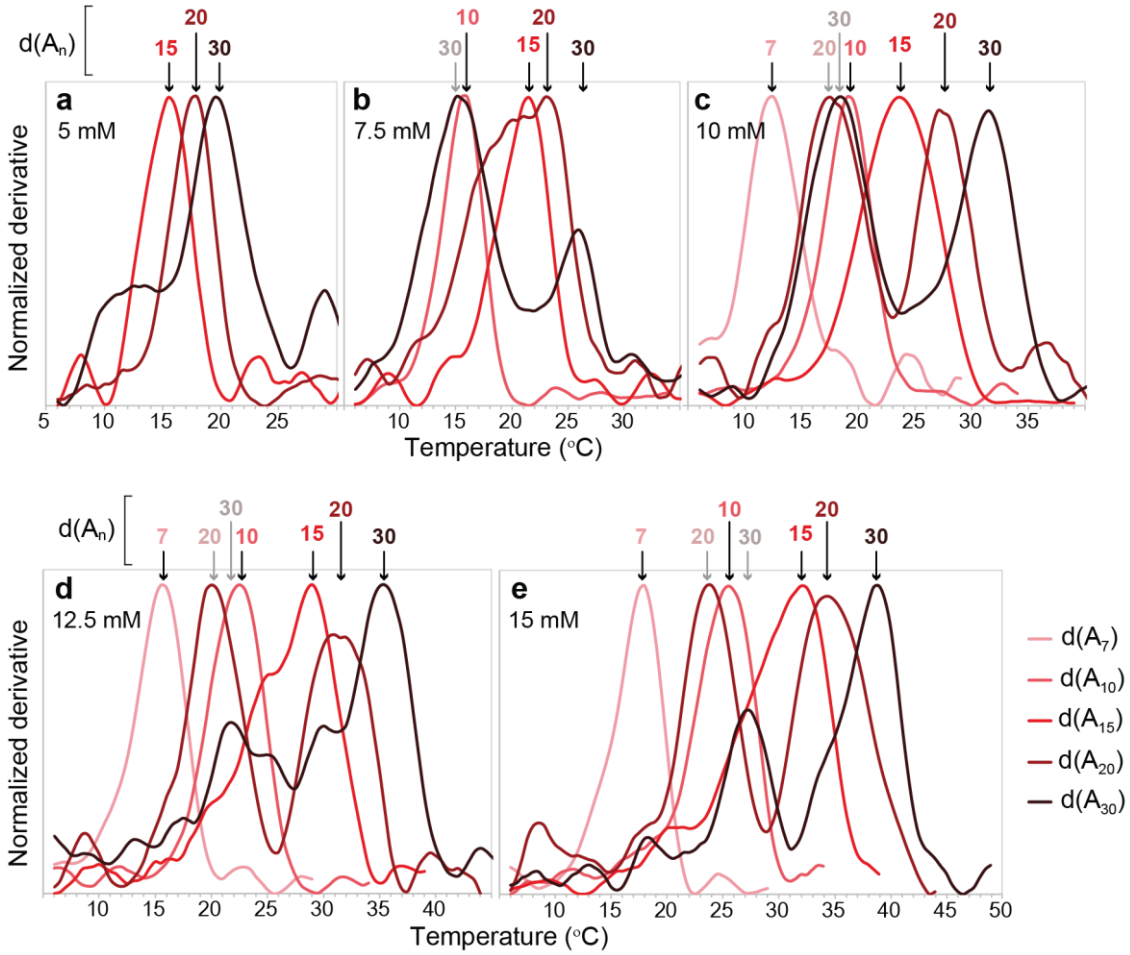


Figure 3.16| Thermal denaturation studies of d(A_n) using strands of different length (n = 7, 10, 15, 20 and 30) in range of [CA] ([CA] = 5, 7.5, 10, 12.5 and 15 mM). Normalized derivatives of thermal denaturation curves monitored by CD at 252 nm. Each panel shows a set of curves for assemblies formed from strands of different lengths for which melting transitions were observed. [CA]: **a)** 5, **b)** 7.5, **c)** 10, **d)** 12.5 and **e)** 15 mM. The colours of the curves change from light pink to deep burgundy with increasing d(A_n) length. Peaks are identified with corresponding strand length *n*. Total [A] in each sample was normalized to 150 μM (equivalent to 10 μM d(A_{15})).

The ensemble of studies examining the effect of [CA] concentration and d(A_n) length on CA-mediated assembly highlight how simple parameters can be used to tune the outcome and properties of the structures obtained. The thermal stability and length of nanofibers can be adjusted to modify the properties of potential functional nanomaterials based on this assembly principle.

3.2.4 Kinetics of CA-mediated d(A_n) assembly

The control of self-assembly kinetics is of fundamental importance to achieve advanced functionality in both natural systems and synthetic supramolecular materials.[21] Understanding of the mechanism, time-course of assembly, intermediate species and the influence of different parameters on the equilibrium populations of supramolecular polymers are essential in the development of such materials for specific applications. Following the qualitative assessment of the mechanism of supramolecular polymerization and identification of parameters influencing the outcome of CA-mediated d(A_n) assembly, we sought to get a deeper understanding of the polymerization process.

As a first step, isothermal kinetics measurements were performed for d(A_{15}) with 10 and 15 mM CA, and d(A_{10}) with 15 mM CA at different temperatures. These measurements monitored the accumulation of assembled material over time at a constant temperature following a jump from an elevated temperature where all d(A_n) was in a monomeric state. The curves acquired illustrate the effects of changing [CA] and strand length on the kinetics of assembly (**Figure 3.17**). Qualitatively, the curves show that polymerization accelerates with cooling, an indicator of enthalpy-driven assembly.[8] In addition, reducing [CA] (10 mM vs. 15 mM) and strand length (d(A_{10}) vs. d(A_{15})) considerably slows down assembly.

Unfortunately, fitting these curves reliably to obtain reaction orders and kinetic parameters proved impossible due to experimental flaws. The curves were obtained using a temperature jump method, where the sample was heated to denature the assembly before being placed at the target temperature in the CD instrument to record the kinetic run. With this method, the temperature of the sample ramped down to the target temperature over the first ~ 100 seconds such that this initial portion of the run could not be included in the fit. As a crucial portion of the reaction progress occurs in the first ~ 100 seconds, the data couldn't in fact be fitted due to the significant temperature changes occurring during this time.

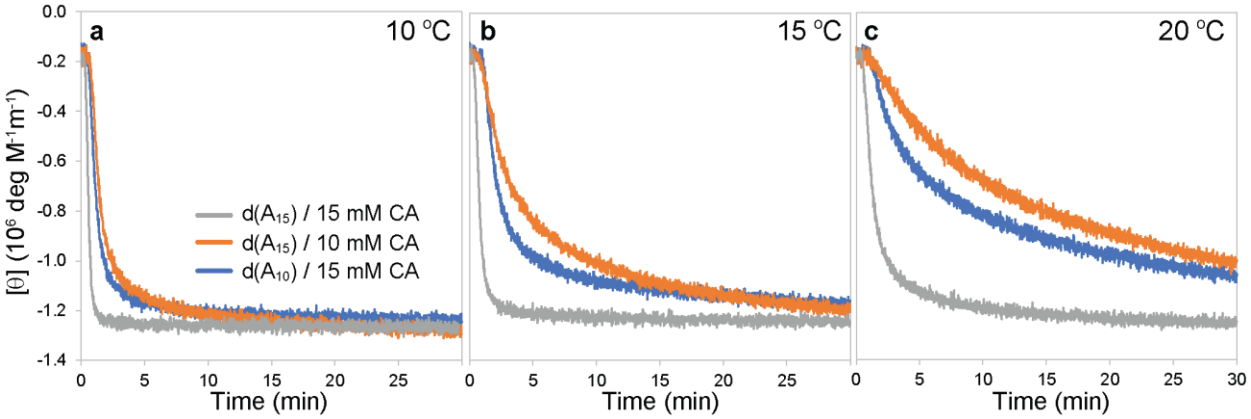


Figure 3.17| Isothermal kinetics for CA-mediated d(A_n) assembly obtained using a temperature jump method. Molar ellipticity acquired at 252 nm at **a)** 10, **b)** 15, and **c)** 20 °C. Time t = 0 is the moment the sample cell is placed at the target temperature following denaturation of the assembly at a raised temperature. Samples: d(A_{15}) with 15 mM CA (grey), d(A_{15}) with 10 mM CA (orange) and d(A_{10}) with 15 mM CA (blue).

A procedure mimicking a stopped-flow experiment was also tested. A small volume of concentrated, free d(A_n) was manually injected into a CA containing solution at the target concentration and temperature followed by immediate recording of the CD signal. However, this method did not yield reproducible curves and similarly lacked reliable signal in the early stages of assembly because of manual mixing and rapid response of the sample at the temperatures tested. Due to these limitations, isothermal kinetics experiments were not used to further examine reaction orders and determine kinetic parameters for this system.

3.2.5 Analysis of CA-mediated assembly of d(A_n) based on thermal hysteresis

An alternative approach to study the kinetics of supramolecular assembly was recently developed by Harkness and Mittermaier (RW Harkness V, N Avakyan, HF Sleiman and AK Mittermaier. *Mapping the energy landscapes of supramolecular assembly by thermal hysteresis*. Nat. Commun. **9**, 3152 (2018)). This methodology enables reaction rate mapping as a function of both monomer concentrations (*[M]*) and temperature based on spectroscopic thermal hysteresis (TH) analysis. TH is a kinetic phenomenon occurring when forward and reverse reactions fail to equilibrate.[8] In supramolecular assembly, it is observed as a lag between annealing and melting curves that are obtained at a given heating/cooling rate, implying that the processes are occurring

far from equilibrium and thus exhibit a kinetic delay. CA-mediated assembly of d(A_n) exhibits TH, as demonstrated in **Chapter 2** (**Figure 2.5 e**).

In the past, TH monitoring has been used to study the folding and unfolding rates of biomolecules in simple two-state unimolecular systems,[39-41] or by making *a priori* assumptions about the mechanism of assembly in a few multimeric systems.[28,42] However, Harkness and Mittermaier saw an opportunity to use spectroscopic TH data to elucidate complex supramolecular assembly mechanisms *de novo*. The TH experiments entail acquiring spectroscopic assembly and disassembly profiles of supramolecular systems with heating and cooling at different rates. The lag caused by TH varies with rate, such that faster heating/cooling displaces the melting/annealing curves further from equilibrium than do slower rates. The rates of assembly and disassembly in the system can then be calculated based on the lag from equilibrium. The TH-based method examines variations in reaction rates over a range of *[M]* and temperature to acquire detailed representations of assembly pathways. This is achieved by creating 3D maps of reaction rate as a function of both *[M]* and temperature from melting/annealing data acquired at a set of different scan rates. The surfaces are used to track apparent reaction order as a function of temperature in a model-free manner, providing a level of kinetic detail that cannot be achieved by traditional methods. Finally, based on the observed reaction order information, explicit mechanistic models are proposed and globally fit to TH datasets. The global fitting procedure permits extraction of the kinetic and thermodynamic parameters that govern the assembly process in question.

The model-free analysis is based on the different lag times associated with different heating/cooling rates in melting/annealing sets of TH curves, and the establishment of low/high temperature baselines where all the material is assumed to be assembled/disassembled respectively. Each trace is used to calculate the fraction of dissociated monomers (θ_U) as a function of temperature and heating/cooling rate so that the rate of monomer release or consumption ($d[M]/dt$) can be calculated from the slopes of the curves ($d\theta_U/dT$), the rate of temperature change (dT/dt) and the total monomer concentration (C_T) using Equation 3.1.[39]

$$\frac{d[M]}{dt} = C_T \frac{d\theta_U}{dT}\frac{dT}{dt} \qquad \text{(Equation 3.1)}$$

Given that multiple heating/cooling rates are used in acquiring the TH profiles, the data set samples a range of *[M]* and *d[M]/dt* values at any temperature within the transition region from

disassembled to assembled material (**Figure 3.18**). Equation 3.2 describes how the rate *d[M]/dt* varies in relation to *[M]*, C_T, the number of monomers in the assembly (*N*), effective reaction orders *n* and *m* (for assembly and disassembly respectively), and rate constants k_{on} and k_{off}.

$$\frac{d[M]}{dt} = -k_{on}[M]^n + k_{off}N\left(\frac{Ct-[M]}{N}\right)^m \qquad \text{(Equation 3.2)}$$

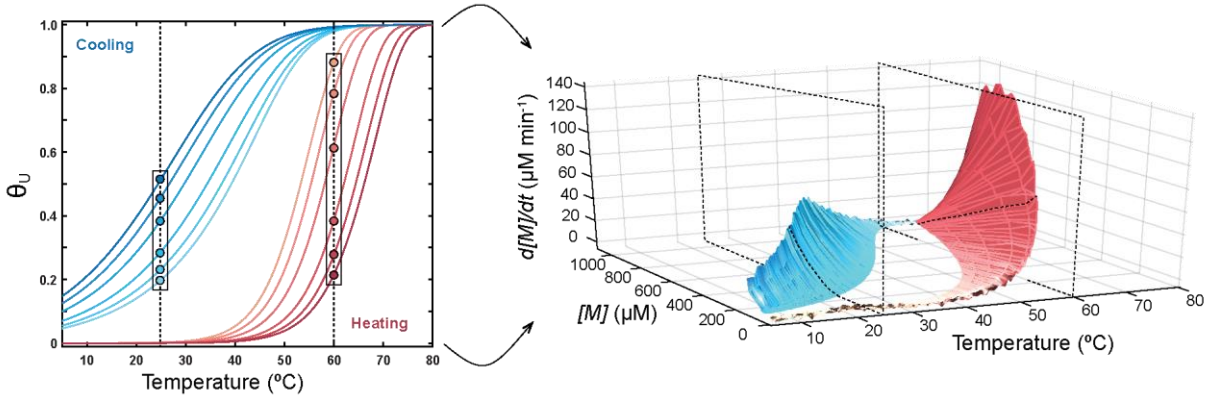


Figure 3.18| Model-free supramolecular assembly analysis from TH profiles. Sample TH dataset obtained using multiple annealing/melting rates (left) is used for generating a 3D supramolecular assembly map (right). Dashed lines/rectangles indicate temperature slices at which *[M]* and *d[M]/dt* are calculated using each curve (different annealing/melting rate). Light/dark blue and red indicate slow/fast scan rates for annealing and melting respectively.

In contrast to a system exhibiting little to no TH, where both terms in Equation 3.2 are of similar magnitude and thus values of *N*, *n* and *m* must be known to extract k_{on} and k_{off},[28,42] a system with large TH allows the equation to be simplified. When TH is large, the first term of Equation 3.2 dominates during annealing, while the second term dominates during melting, such that Equations 3.3 and 3.4 can be applied in the respective calculations. Given these approximations, it is possible to extract *n*, *m*, k_{on} and k_{off} from the linear plots afforded by taking the log of both sides of Equations 3.3 and 3.4. A plot of log(- *d[M]/dt*) *vs*. log(*[M]*) will have a slope of *n* and a y-intercept of log(k_{on}) for the annealing data, while in the case of melting, a plot of log(*d[M]/dt*) *vs*. log(C_T - *[M]*) will have a slope of *m* and y-intercept of (1 - *m*)log(*N*) + log(k_{off}).

Annealing: $$\frac{d[M]}{dt} \approx -k_{on}[M]^n \qquad \text{(Equation 3.3)}$$

Melting: $$\frac{d[M]}{dt} \approx k_{off}N\left(\frac{Ct-[M]}{N}\right)^m \qquad \text{(Equation 3.4)}$$

The values of n and m represent model-free estimates of apparent reaction orders at each temperature point over the transition range. The values of k_{on} and k_{off} are phenomenological constants describing the rate of the reaction, but are not meaningful in the supramolecular pathway analysis as they contain contributions from a number of processes. The approximations are considered valid if the ratio between the slopes of the melting and annealing curves at the slowest scan rate is roughly 3-fold in the middle of the annealing transition. The subsequent global fitting analysis requires a minimum ~ 1-2 °C of difference between the annealing/melting profiles in the transition region.

To apply the TH-based analysis method to CA-mediated d(A_n) assembly, TH profiles were acquired for three sets of conditions, examining the effect of variation in [CA] and d(A_n) strand length on assembly and disassembly kinetics: d(A_{15}) with 15 mM CA is compared to d(A_{15}) with 10 mM CA (reduced [CA]) and d(A_{10}) with 15 mM CA (reduced strand length). The plots of θ_U *vs.* temperature obtained from UV-Vis absorbance curves are shown in **Figure 3.19**. As the melting scans largely produced overlapping curves, indicating rapid dissociation at these temperatures, the disassembly of d(A_n) fibers could not be characterized further using the TH-based method.

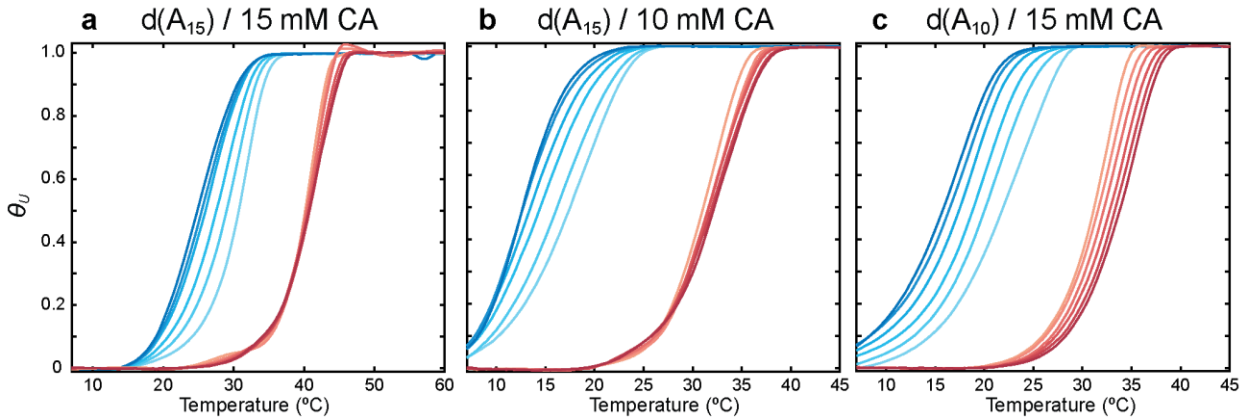


Figure 3.19| Fraction unfolded (θ_U) TH profiles as a function of temperature at different melting/annealing rates. a) d(A_{15}) with 15 mM CA, **b)** d(A_{15}) with 10 mM CA and **c)** d(A_{10}) with 15 mM CA. Data acquisition parameters are listed in **Table 3.3** in section 3.4.9.

On the other hand, the annealing scans displayed a marked cooling rate dependence and were subjected to in depth analysis (**Figure 3.20 a**). The rates of d(A_n) consumption $d[M]/dt$ were mapped as a function of both $[M]$ and temperature (**Figure 3.20 b**). The apparent reaction orders

(n) were calculated from the slopes of the linear plots obtained by taking log-log slices of the maps at sequential temperatures (at 0.5 °C intervals) across the transition region. The n values for all three d(A_n) systems under investigation remained close to 3 across the transition region (**Figure 3.20 c**). In polymerization reactions, because monomers are integrated into polymers by adding to an ensemble of growing fibers of different lengths, the apparent reaction order n does not reflect the molecularity of a single rate-limiting barrier.[43] Rather, n values can offer mechanistic insight as to the molecularity of the nucleus.[43,44]

The next step in the analysis consists of proposing an explicit mechanistic model corresponding to the calculated apparent reaction orders to quantify the supramolecular assembly pathway both in terms of kinetic and thermodynamic parameters. The temperature dependent annealing curves and reaction orders were fit using the Goldstein-Stryer cooperative supramolecular polymerization model,[43] which explicitly tracks populations of oligomers up to a certain number of monomer units (N = 100 in **Figure 3.20 a**), while populations of longer fibers are accounted for using the fibril pool approximation from Korevaar *et al*.[31,34] For monomers and oligomers smaller than the critical nucleus size s (critical number of d(A_n) strands), the association and dissociation constants were termed k_{n+} and k_{n-} respectively (nucleation phase), while oligomers containing more than s strands were described using rate constants k_{e+} and k_{e-} (elongation phase). For the critical nucleus itself, the association and dissociation constants were taken as k_{e+} and k_{n-} respectively. The forward rate constants k_{n+} and k_{e+} were held as equal, a simplification applied in other systems,[31,43,45] and assembly activation enthalpies were allowed to vary with temperature (change in heat capacity at constant pressure $\Delta C_p \neq 0$).[46]

The optimal critical nucleus size was determined by systematically varying s in the global fit. While excellent fits were obtained with s values from 2 to 4, s = 3 gave the best fit for d(A_{15}) with either 10 or 15 mM CA. It is important to point out that although the effective reaction order ($n \approx 3$) matches the optimal nucleus size for the d(A_{15}) systems, the relationship is generally less straightforward. Harkness and Mittermaier have shown that the effective order of a polymerization reaction can be interpreted quantitatively in terms of weighed averages of fluxes of individual assembly steps.[47] The effective orders extracted from TH analysis are rich in information and linked to critical nucleus size, however in supramolecular polymerization this relationship is complex and warrants further theoretical study. The experimental values of n between 2.5 and 3.5 observed for d(A_{10}) with 15 mM CA deviate from $s = 2$ which yields the best fit for this set

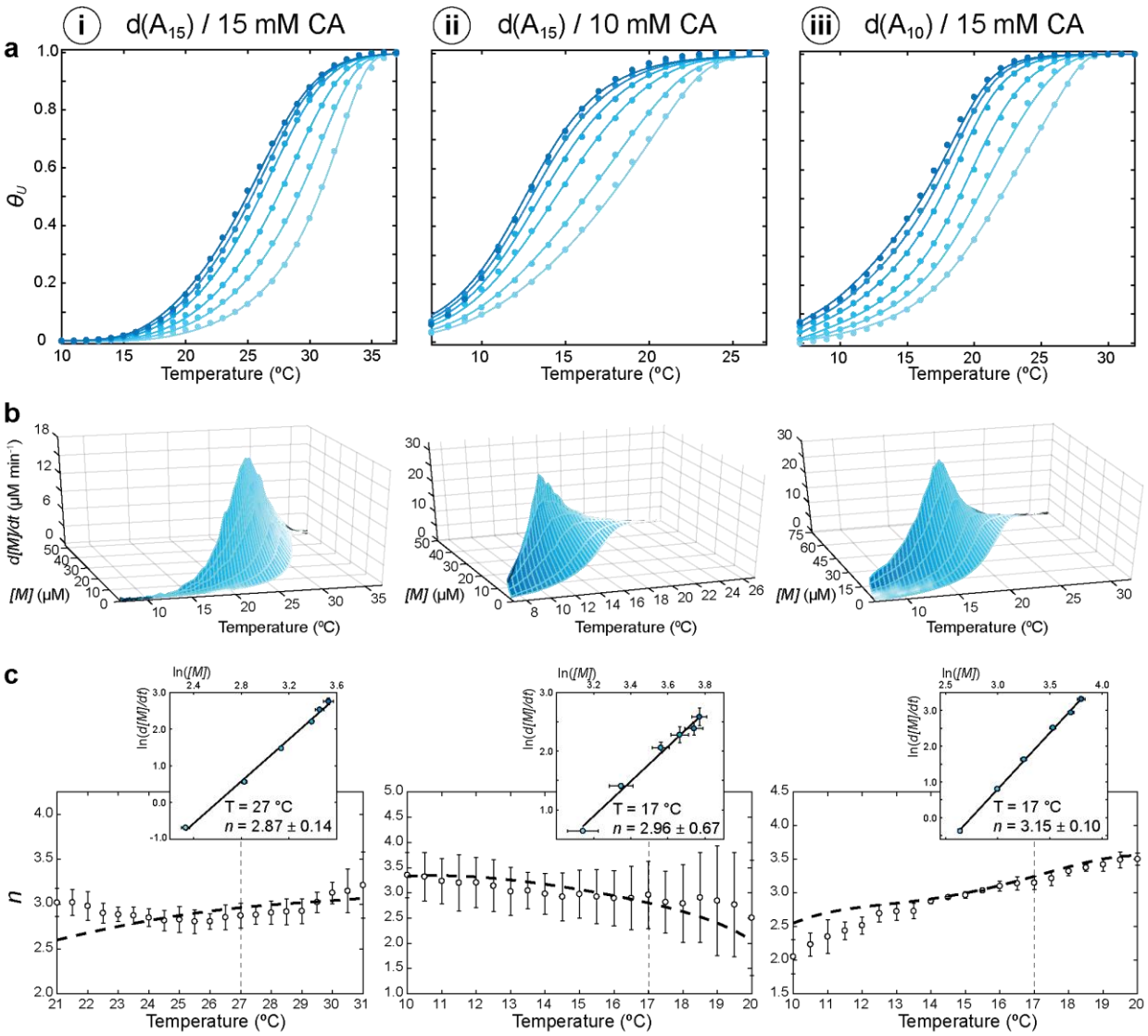


Figure 3.20| Model-free analysis of TH profiles for CA-mediated assembly of d(A_n). Columns correspond to samples: **i)** d(A_{15}) with 15 mM CA, **ii)** d(A_{15}) with 10 mM CA, and **iii)** d(A_{10}) with 15 mM CA. Rows correspond to: **a)** Fraction unfolded θ_U TH profiles as a function of temperature at different annealing rates. Circles show the experimental data and lines correspond to the globally fit Goldstein-Stryer model for cooperative supramolecular assembly assuming a nucleus size of **i)** 3, **ii)** 3 and **iii)** 2. Every 2nd experimental point is shown for clarity. Light to dark blue correspond to slow to fast annealing rates. **b)** 3D supramolecular assembly maps of rate *d[M]/dt vs. [M]* and temperature. **c)** Effective assembly orders *n* as functions of temperature through the annealing transition. Circles and dashed lines represent *n* obtained from model-free analysis of experimental and globally fitted data respectively. Insets illustrate model-free analysis of the maps in **(b)** at the given temperatures, where *n* is the slope of the linear ln(*d[M]/dt*) *vs.* ln(*[M]*) plot. Error bars correspond to the standard deviation of values obtained from analysis of triplicate TH experiments.

of parameters. This finding highlights the complexity of cooperative polymerization, where monomers are consumed at each step of the assembly process and elongation can proceed indefinitely, such that no single energy barrier is entirely rate determining. Nevertheless, the observed range of *n* values suggests that the most significant contributions to the flux occur in the early stages of the assembly process (the average flux gives an order ~ 2-3).

In cooperative supramolecular polymerization, the nucleus size can be defined as the number of monomers in the largest species which is more likely to dissociate than to grow.[45] This is well illustrated in quantitative free energy diagrams calculated at different temperatures to emphasize the differences in free energy of intermediates in the nucleation and elongation stages of assembly (**Figure 3.21**). For d(A_{15}) with 15 mM CA (**Figure 3.21 a**), the free energy diagram calculated at 20 °C shows that the trimeric nucleus (s = 3) is the least stable intermediate, while each subsequent monomer addition produces a lower energy oligomer. A nucleus composed of 3 strands is physically reasonable in the system given a proposed triple helical structure based on hexameric CA:A rosettes. It suggests that a minimum of 3 strands is required to complete the rosette arrangement and that binding of a fourth strand effectively stabilizes the nascent fiber in an arrangement that favors the addition of further strands. This finding also stands for d(A_{15}) with 10 mM CA (**Figure 3.21 b**), however it is best illustrated at a lower temperature (10 °C). For the system with reduced CA at 20 °C, the free energy differences between monomer additions and the heights of the energy barriers show that disassembly of the nucleus is much more likely to occur, as the heights of the energy barriers correspond directly to transition probabilities, and smaller barriers thus indicate that more molecules cross the barrier in a given direction per unit time. This seems to be the case mostly up to the nucleus, where disassembly barriers are small, implying a rapid equilibrium between monomer and nucleus.

In the case of d(A_{10}) with 15 mM CA (**Figure 3.21 c**), the best fit critical nucleus size switches to s = 2, which is also reflected in the free energy diagram, where the dimer is the highest energy intermediate. A possible explanation for this strand length-dependent change may have to do with a reduced number of possible binding registers for the shorter d(A_{10}) strand compared to d(A_{15}). With a nucleus composed of two strands, binding of a third d(A_{10}) strand has a greater probability of completing more hexameric rosettes to drive elongation, producing a reduction in free energy for the trimer intermediate and beyond. In contrast, statistically, two d(A_{15}) monomers can bind to each other in a greater number of slipped registers so that a third strand may not

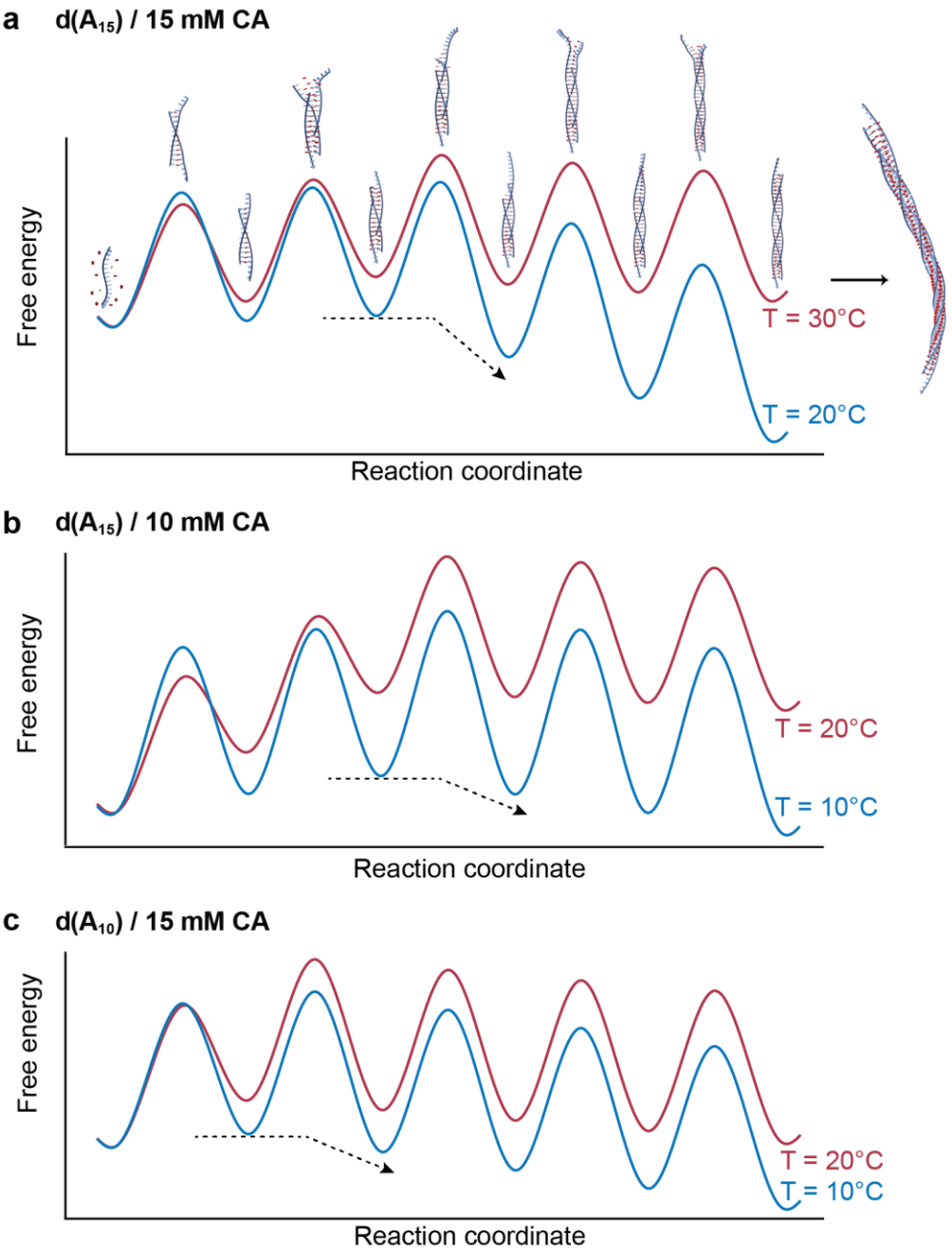


Figure 3.21| Quantitative free energy diagrams for CA-mediated cooperative of d(A_n) polymerization. a) d(A_{15}) with 15 mM CA at 20 at 30 °C; **b)** d(A_{15}) with 10 mM CA at 10 at 20 °C; **c)** d(A_{10}) with 15 mM CA at 10 at 20 °C. Model structures in **(a)** illustrate each intermediate stage in the polymerization both in ground and transition state. For **(a-c)**, lower temperature is in blue and higher temperature in red; the reaction coordinate is truncated at the hexamer, longer fibers form by stepwise monomer association; dashed arrows highlight the highest energy step for each system before elongation proceeds with further monomer addition.

complete enough hexamers to drive elongation, thus an additional strand becomes necessary to enter that phase of assembly. This finding suggests that there may be a correlation between d(A_n) length and nucleus size. Extrapolation to longer strands would indicate that the switch from nucleation to elongation may occur upon association of a larger number of monomers. Formation of larger nuclei would imply a longer delay for the onset of elongation, build-up of pre-elongation intermediates and populations of nuclei with strands bound in a variety of registers that then grow at different rates giving rise to fiber populations that display multiple melting transitions as observed with strands longer than d(A_{15}) in section 3.2.3.

Direct comparison of the free energy diagrams for the three different sets of d(A_n) assembly conditions seems to indicate that a reduction in [CA] has a more pronounced effect on the free energy changes along the pathway than does reduced strand length. At 20 °C, d(A_{15}) with 15 mM CA (**Figure 3.21 a**) steadily proceeds through elongation with a downhill trajectory in the free energy of intermediates. In contrast, the magnitude of the steps from one intermediate to the next at the same temperature is smaller for d(A_{10}) with 15 mM CA (**Figure 3.21 c**), while for d(A_{15}) with 10 mM CA (**Figure 3.21 b**), nucleation is much more unfavorable and the free energy steps in elongation do not follow a sharp downhill trajectory. Furthermore, at 10 °C, the elongation phase for the d(A_{10}) system is more favorable than it is for d(A_{15}) with 10 mM CA.

These differences can also be examined by comparing dissociation constants (K_D) for nucleation and elongation at different temperatures (**Table 3.1**). The K_D values were calculated by using rate constants and activation energy values obtained at reference temperatures in the global fit (listed in **Table 3.4** in section 3.4.12). The K_D values for nucleation show sharp temperature-dependent changes as expected from $K = e^{-\Delta G/RT}$. In addition, allowing for $\Delta C_p \neq 0$ means that the relationship between ln(K) *vs.* 1/T is non-linear, in contrast to a case where $\Delta C_p = 0$. (see **Table 3.4** for ΔC_p values obtained from global fitting). At 20 °C, nucleation K_Ds in the millimolar range for d(A_{15}) with 10 mM CA (reduced [CA]) and d(A_{10}) with 15 mM CA (reduced strand length) indicate that nuclei in these systems are highly unstable at this temperature, never more populated than the monomeric state. In comparison to d(A_{15}) with 15 mM CA, nucleation in the reduced [CA] system becomes favorable at a lower temperature than for a system where d(A_n) strand length is reduced. For elongation, the K_D values remain favorable across the 10 to 25 °C range for all three systems, meaning that fibers spontaneously elongate in these temperature and concentration

conditions. Elongation of d(A_{15}) with 15 mM CA is orders of magnitude more favorable in this temperature range than the systems with reduced [CA] and strand length.

	K_D (μM)							
	Nucleation				Elongation			
Temperature (°C)	**10**	**15**	**20**	**25**	**10**	**15**	**20**	**25**
50 μM d(A_{15}) / 15mM CA	0.1	1.0	10.0	101.5	3.9×10^{-8}	5.3×10^{-5}	1.4×10^{-2}	0.9
50 μM d(A_{15}) / 10mM CA *Reduced [CA]*	38.3	740.0	4.8×10^{3}	1.1×10^{4}	2.0	9.8	28.2	48.9
75 μM d(A_{10}) / 15mM CA *Reduced strand length*	17.6	62.4	1.7×10^{3}	3.1×10^{5}	1.9	8.1	25.1	55.7

Table 3.1| Calculated dissociation constants (K_D) for nucleation and elongation. K_D values (in μM units) calculated at 10, 15, 20 and 25 °C for 50 μM d(A_{15}) with 15 mM CA, 50 μM d(A_{15}) with 10 mM CA (system with reduced [CA]), and 75 μM d(A_{10}) with 15 mM CA (system with reduced d(A_n) strand length) based on reference values listed in **Table 3.4** in section 3.4.12. For nucleation, $K_D = k_{n-}/k_{n+}$, and for elongation, $K_D = k_{e-}/k_{e+}$,

TH-based analysis of the kinetics and thermodynamics of CA-mediated d(A_n) assembly brings to the forefront aspects of the process that could not be probed earlier. It has revealed valuable information on nucleation stoichiometry that appears to be d(A_n) strand length dependent. Furthermore, it highlights the impact of parameters such as [CA] and d(A_n) strand length on rates of assembly and stability of intermediate species, providing a quantitative and mechanistic framework for trends identified earlier using thermal denaturation experiments. Within the limited range of parameters examined here, changes in [CA] remain the most impactful parameter in the assembly process. It is important to point out that [CA] affects both the kinetics and thermodynamics of assembly. First, [CA] has an effect on the overall rate of monomer consumption ($k_{obs}[M] = k[M][CA]^x$, where k_{obs} is the observed rate constant and x is the number of CA molecules joining a monomer), such that increases in [CA] affect the rates of both nucleation and elongation as CA participates in both phases of assembly. In addition, changes in [CA] affect the magnitudes of the rate constants. For example, at 20 °C, the k_{n-} and k_{e-} for d(A_{15}) with 15 mM CA are four orders of magnitude smaller than for d(A_{15}) with 10 mM CA, while rate constants for

the forward reactions are more similar. Finally, the [CA] effect on rate constants impacts equilibrium constants making both nucleation and elongation more thermodynamically favorable. In combination with the increased nucleation and elongation rates, higher CA concentrations push the annealing and melting temperatures of assembly to higher values (increased thermal stability).

Overall, the insights obtained from TH-based analysis above provide a better understanding of the mechanism of CA-mediated d(A_n) assembly and quantify the process for three different sets of parameters highlighting once again the impact of [CA] and d(A_n) strand length. With further work, this new quantitative framework could be used to better control the outcome of assembly in terms of fiber population distributions and the properties of the supramolecular nanomaterial, such as nanofiber length and thermal stability behaviour.

3.3 Conclusions

This chapter has examined some aspects of the mechanism of CA-mediated d(A_n) assembly into nanofibers, and identified key parameters affecting the properties of the assembly products. We have established that d(A_n) strands likely take on a staggered arrangement within the nanofibers based on FRET experiments. Furthermore, spectroscopic studies using a range of d(A_{15}) and CA concentrations were used to show that the supramolecular polymerization of d(A_n) strands proceeds following a cooperative, nucleation-elongation mechanism. The most important parameter regulating the assembly outcome appears to be [CA], as it is a critical factor triggering the cooperative polymerization mechanism. Above $[CA]_{cr}$, increases in [CA] impact the thermal denaturation behaviour of d(A_{15}) assemblies, leading to dramatic changes in T_M for the structures. In contrast, [d(A_{15})] does not significantly affect melting behaviour. The length of d(A_n) strands is another crucial parameter. Assemblies formed from short strands (d(A_7) to d(A_{15})) melt in a single transition, while structures formed using longer strands (d(A_{18}) to d(A_{50})) display multi-phase melting, highlighting differences in product distributions and underlying effects potentially related to nucleation. In combination, variation of d(A_n) length and changes in [CA] have a profound impact on product distribution and thermal stability of the structures, enabling the production of nanofibers with tunable characteristics.

Examination of assembly kinetics using standard isothermal experiments was stalled by flaws in experimental design. However, a new approach developed by Harkness and Mittermaier

based on spectroscopic thermal hysteresis (TH) analysis revealed a quantitative representation of the assembly process, providing insight into both kinetic and thermodynamic facets of the mechanism. Based on TH profiles, maps of monomer consumption rates with respect to *[M]* and temperature were analyzed to reveal effective reaction orders over a range of temperatures. This ensemble of data was then globally fit to a cooperative polymerization mechanism model to calculate nucleation and elongation rate constants, activation energies, and use these values to build energy diagrams describing the assembly pathway. Importantly, this quantitative treatment has revealed new insights into the molecularity of the nucleus, which appears to be affected by strand length, switching from a trimeric to a dimeric nucleus when going from d(A_{15}) to d(A_{10}). Furthermore, the impact of [CA] and d(A_n) strand length on assembly was reassessed, highlighting again the dramatic effect of changes in [CA], this time in quantitative kinetic and thermodynamic terms.

Overall, these studies emphasize the tunability of CA-mediated assemblies and contribute mechanistic insights of interest to the field of supramolecular polymerization in general.

3.4 Experimental

3.4.1 Materials

Cyanuric acid (CA), tris(hydroxymethyl)aminomethane (Tris), magnesium chloride hexahydrate ($MgCl_2 \cdot 6\ H_2O$), sodium chloride (NaCl), glacial acetic acid, urea, ethylenediaminetetraacetic acid (EDTA), ammonium hydroxide, Stains-All, trimethylamine, HPLC grade acetonitrile (MeCN) and toluene-d_8 were used as purchased from Sigma-Aldrich. Boric acid and GelRed were obtained from Fisher Scientific and used as supplied. Acrylamide/bis-acrylamide (40% 19:1) solution, ammonium persulfate and tetramethylethylenediamine (TEMED) were used as purchased from BioShop Canada Inc. Sephadex G-25 (super fine, DNA grade) was purchased from Glen Research.

For in-house DNA synthesis, the reagents and materials listed below were used. Anhydrous acetonitrile, nucleoside derivatized and universal 1000 Å LCAA CPG pre-packed columns (loading densities of 25-40 μmol/g), activator (0.25 M ETT in acetonitrile, oxidizer (0.02 M I_2 in THF/pyridine/H_2O), Cap A mix (THF/lutidine/Ac_2O), Cap B mix (16% MeIm in THF), and

deblock solution (3% DCA/DCM) were all used as purchased from Glen Research. dA-CE, dC-CE, dmf-dG-CE, and dT-CE, cyanine 3 (Cy3) and cyanine 5 (Cy5) phophoramidites were purchased from Glen Research.

$d(A_n)$ (n = 7, 10, 12, 15, 18, 20, 30 and 50) oligonucleotides were purchased from Integrated DNA Technologies (IDT), purified by PAGE under denaturing conditions (15-24% polyacrylamide with 8 M urea, using 1xTBE running buffer) at constant voltage (30 mA, at 250 V for the first 30 minutes, followed by 30-60 minutes at 500 V) and desalted with Sephadex G-25. Triplex forming oligonucleotides were purchased from IDT desalted and used without further purification.

Muscovite Ruby mica sheets (grade 2) were used as substrate for all AFM imaging studies.

1xTBE buffer was composed of 45 mM Tris, 45 mM boric acid and 2 mM EDTA at pH 8.3. 1xTAMg buffer was composed of 40 mM Tris, 7.6 mM $MgCl_2 \cdot 6\ H_2O$ with pH adjusted to the desired value with glacial acetic acid. 1xAcMg buffer was composed of 40 mM acetic acid, 7.6 mM $MgCl_2 \cdot 6\ H_2O$, with pH adjusted 4.5. All buffers were prepared with Milli-Q water and samples were prepared with autoclaved Milli-Q water.

3.4.2 Instrumentation

A Mermade DNA synthesizer from Bioautomation was used for in-house DNA synthesis. Gel electrophoresis experiments were carried out on a 20 x 20 cm vertical acrylamide Hoefer 600 electrophoresis unit. Thermal annealing procedures were carried out using an Eppendorf Mastercycler Pro 96 well thermocycler. Electrophoresis gels were scanned with either a CanoScan 4200F scanner by Canon or a Chemidoc MP Imaging System from BioRad.

UV-Vis DNA quantification measurements were performed on a BioTek Synergy HT microplate reader or a Nanodrop Lite spectrophotometer from Thermo Scientific. Circular dichroism (CD) and UV-Vis studies were performed using a 1 mm path length quartz cuvette on a Jasco-810 spectropolarimeter equipped with a Peltier temperature control unit and a water recirculator. Temperature verification on the CD instrument was performed with a handheld digital thermometer (Oakton) equipped with a fine gage thermocouple (Omega).

3.4.3 General procedures

Unless otherwise indicated, the following general procedures were used. Samples were thermally annealed using a 50 → 5 protocol (50 °C for 10 minutes, followed by cooling to 5 °C at a rate of 0.5 °C/min) and incubated overnight at 5 °C prior to any measurements. CD and UV-Vis spectra were obtained at 5 °C (230 nm to 350 nm range, 100 nm/min scan rate, 1 nm bandwidth, 2 accumulations). Ellipticity data (θ) obtained by CD were converted to molar ellipticity [θ], using Equation 3.5, where *c* is the molar concentration of DNA strands and *l* is the path length (in cm). In some cases, UV-Vis absorbance values were converted to molar absorptivity ε in the interest of comparing samples obtained at different DNA concentrations. Thermal denaturation of assemblies was recorded by following CD and UV-Vis signals at 252 nm over the 5 to 60 °C temperature range at 0.5 °C intervals, with a heating rate of 1 °C/min. The CD and UV-Vis melting data were converted to fraction of bases assembled (f) *vs.* temperature using Equation 3.6, where S is the signal at a given temperature, S_A is the signal for assembled $d(A_n)$ species and S_F is the signal for free $d(A_n)$ in solution. The melting temperature values (T_M) reported represent the temperature at which half-dissociation (fraction assembled f = 0.5) was recorded. First derivatives of the CD signal versus temperature curves were generated to verify the T_M. The values determined from both methods differed by less than 1.0 °C. FWHM was determined from the first derivative peak. The T_M value from the derivative alone was used for the analysis of melting curves for the poly(adenine) strands of varying lengths.

$$[\theta] = {}^{100\,\theta}/_{c\,l} \qquad \text{(Equation 3.5)}$$

$$f = {}^{S - S_F}/_{S_A - S_F} \qquad \text{(Equation 3.6)}$$

3.4.4 Determination of fiber elongation mechanism by FRET

Synthesis of fluorescently labelled oligonucleotides was performed in house on a 1 μmol scale. Coupling efficiency was monitored after removal of the DMT 5'-OH protecting groups. Cy3 and Cy5 phosphoramidites were coupled manually off-column under inert atmosphere in a glove box. Completed sequences were deprotected in a 28% ammonium hydroxide solution for 16 hours at room temperature, purified by PAGE under denaturing conditions (20% polyacrylamide with 8

M urea, using 1xTBE running buffer) at constant voltage (30 mA, at 250 V for the first 30 minutes, followed by 1 hour at 500 V). Following electrophoresis, the bands of purified oligonucleotides were excised and recovered from the gel (crushed, covered with water, vigorously shaken, flash frozen with liquid nitrogen and kept at room temperature for 16 hours). The solutions were then dried down to 1 mL and desalted using size exclusion chromatography (Sephadex G-25). **Table 3.2** shows sequences for test strand **d(A_{20})-2F** and strands composing the blunt and staggered control duplexes.

All strands were characterized by denaturing PAGE. DNA was loaded at 0.003 OD_{260} per lane on 20% denaturing PAGE, run at room temperature with TBE running buffer for 30 min at 250 V followed by 45 min at 500 V and stained with GelRed. The oligonucleotides were analyzed by LC-ESI-MS using a Dionex Ultimate 3000 coupled to a Bruker Maxis Impact QTOF in negative ESI mode. Samples were run through an Acclaim RSLC 120 C18 column (2.2 µM 120A 2.1 x 50 mm) using a gradient of 98% mobile phase A (100 mM HFIP and 5 mM TEA in H2O) and 2 % mobile phase B (MeOH) to 40 % mobile phase A and 60% mobile phase B in 8 minutes. The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1.

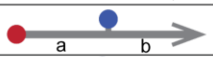
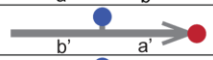
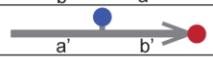
Name	Scheme	Sequences
d(A_{20})-2F		5'- **Cy3** - AAAAAAAAAA - **Cy5** - AAAAAAAAAA -3'
C		5'- **Cy3** - ACAGTGAAAG - **Cy5** - TCGGCAGAAT -3'
CB		5'- ATTCTGCCGA - **Cy5** - CTTTCACTGT - **Cy3** -3'
CS		5'- CTTTCACTGT - **Cy5** - ATTCTGCCGA - **Cy3** -3'

Table 3.2| Schematic representations and sequences of dye-labeled test strand d(A_{20})-2F and strands composing the control duplexes. Arrow direction is from 5' to 3'. Red circle represents a Cy3 modification; blue circle represents a Cy5 modification. a-a' and b-b' are complementary 10-mer sequences.

Samples for fluorescence studies were prepared with 0.5 µM total strand concentration to ensure that dye absorbance did not exceed 0.1 absorbance units. 1xTAMg pH 6 buffer was used for all fluorescence studies. Samples (test strand **d(A_{20})-2F**, blunt duplex composed of strands C+CB and staggered duplex with strands C+CS, all with 0 or 10 mM CA) were thermally annealed from 95 °C to 5 °C over 1 hour and incubated overnight at 5 °C prior to analysis. The blunt and staggered duplex constructs were examined by native PAGE. DNA was loaded 0.06 nmol per lane

on 8% native PAGE, run at 5°C with 1xTAMg pH 6 running buffer for 150 min at 250 V and stained with GelRed.

A Varian Cary Eclipse fluorescence spectrophotometer equipped with a temperature control unit was used for fluorescence measurements. Fluorescence spectra for FRET efficiency value determination were acquired at 10 °C, with both excitation and emission slit widths set at 5 nm to optimize signal intensity, using λ_{Ex} = 548 nm for Cy3 and collecting Cy3 and Cy5 emission from 555 nm to 800 nm with a scan rate of 600 nm/min. Samples were analyzed in triplicate.

FRET efficiency (E_T) was calculated using Equation 3.7, where F_A is fluorescence intensity for the acceptor (Cy5, λ_{Em} = 667 nm) and F_D is fluorescence intensity for the donor (Cy3, λ_{Em} = 564 nm) collected upon excitation of the donor dye. Whereas E_T for both the blunt and staggered duplexes remained unchanged with the addition of CA, the E_T for **d(A_{20})-2F** increased with the addition of CA from 0.71 ± 0.02 ([CA] = 0 mM) to 0.78 ± 0.03 ([CA] = 10 mM). The high E_T value for **d(A_{20})-2F** in the absence of CA can be attributed to the flexibility of the single strand.

$$E_T = \frac{F_A}{F_A + F_D} \qquad \text{(Equation 3.7)}$$

3.4.5 Combined effect of d(A_{15}) and CA concentration on assembly

For each of the following d(A_{15}) concentrations – 5, 10, 25, 40, 50 and 75 μM – a series of samples containing different CA concentrations (0, 1, 2, 3, 4, 5, 10, 15 mM) was prepared in 1xTAMg pH 4.5. Samples were annealed and incubated overnight. The CD and UV-Vis spectra were obtained followed by thermal denaturation.

For each series at a given d(A_{15}) concentration, molar ellipticity [θ] (from CD signal at 252 nm) and molar absorptivity ε (from UV-Vis absorbance at 260 nm) were used to calculate the fraction of DNA assembled (f) at each CA concentration point. The calculation was based on an assumption of full assembly (f = 1) at the minimal value in the series (reflecting intensity of characteristic CD peak at 252 or more pronounced hypochromicity associated with assembly from the absorbance peak) or no assembly (f = 0) at the maximal value.

3.4.6 Native PAGE analysis of CA-mediated assembly of d(A_{15})

Native PAGE experiments were performed on the 50 μM d(A_{15}) series in the CA concentration range described in section 3.4.5 above. A total of 0.2 nmol d(A_{15}) was loaded. Samples were diluted to a loading volume of 10 uL with 1xTAMg pH 4.5 containing the corresponding CA concentration. Samples were kept on ice prior to loading and running of the 8% native acrylamide gel at 4 °C (250 V for 150 minutes, stained with Stains-All). It is important to note that the gel was polymerized using TAMg pH 8 buffer and then pre-run with 1xTAMg pH 4.5 buffer containing 10 mM CA as a running buffer (250 V for 90 min), before samples were loaded. The same running buffer was used. This procedure was chosen to ensure no spontaneous dissociation of assemblies as a result of dilution in a low [CA] environment (preliminary experiments showed complete dissociation of assemblies when running buffer without CA was used). Triplex forming oligonucleotide (TFO) DNA controls showing differences in gel mobility for 15-mer duplex vs. triplex were also analyzed by PAGE. The TFO sequences (**Figure 3.9 b**) were adapted from Torigoe *et al*.[48] Samples were prepared in 6.67 μM total DNA concentration in 1xTAMg pH 4.5 and annealed (95 °C for 5 minutes, slow cooling down to 4 °C over 4 hours). A total of 0.24 nmol DNA was loaded per lane on an 8% native gel that was run at 250 V for 150 minutes with 1xTAMg pH 4.5 with 10 mM CA as running buffer and stained with Stains-All.

3.4.7 Effect of CA concentration on the T_M of CA-mediated assemblies of d(A_{15})

To further investigate the dependence of CA-mediated assembly of d(A_{15}) on CA concentration, another series of samples was prepared. Samples containing 40 μM d(A_{15}) in pH 4.5 1xTAMg buffer in a range of CA concentrations (0, 1.2, 2.4, 3.6, 4.8, 6, 9, 12, 15 and 18 mM) were prepared in triplicate, annealed and incubated overnight. CD and UV-Vis spectra were obtained followed by thermal denaturation.

3.4.8 Effect of strand length on CA-mediated assembly of d(A_n)

Samples containing d$(A)_n$ strands (n = 7, 10, 12, 15, 18, 20, 30 and 50) were prepared with 10 mM CA in pH 4.5 1xTAMg buffer, were annealed and incubated overnight. DNA concentration was normalized to 150 μM total d(A) bases (concentration in d(A_n) strands: 21.4, 15, 12.5, 10, 8.3,

7.5, 5 and 3 μM for n = 7, 10, 12, 15, 18, 20, 30 and 50 respectively). CD and UV-Vis spectra were obtained. The samples were then thermally denatured at a rate 0.5 °C/min, following the CD and UV-Vis signals at 252 nm at 0.2 °C intervals from 5 to 60 °C. For samples with a single melting transition (d(A_n) where n = 7, 10, 12 and 15), T_M was the temperature at which half-dissociation (fraction assembled f = 0.5) of the assembly occurred. First derivatives of the CD signal vs. temperature curve were generated to verify the T_M. The values determined from both methods differed by less than 1 °C. In the case of melting curves with multiple melting transitions ((d(A_n) were n = 18, 20, 30 and 50), only values from the peaks of the first derivative traces were used to determine T_M values.

For samples with various [CA] (0, 3, 4, 5, 7.5, 10, 12.5 and 15 mM CA) in combination with d$(A)_n$ strands (n = 7, 10, 15, 20 and 30), the same strand concentrations as above were used. CD spectra were obtained at 5 °C. Samples were thermally denatured at a rate 1 °C/min, following the CD signal at 252 nm at 0.5 °C intervals from 5 to 50 °C (or lower for short strands). Calculations of fraction assembled were carried out as described in section 3.4.5.

d(A_{15}), d(A_{30}) and d(A_{50}) samples were imaged by AFM in solution. A 5 μL drop of sampe diluted in 5x in 1xTAMg pH 4.5 with 10 mM CA was deposited on a freshly cleaved mica surface and incubated for 2 minutes followed by addition of 50 μL of filtered 1xTAMg pH 4.5 with 10 mM CA to the closed MTFML (Bruker) fluid cell. Topography images were acquired by AFM carried out under ambient conditions using a MultiMode 8 microscope with a Nanoscope V controller (Bruker) in ScanAsyst mode. Silicon nitride levers with nominal spring constant of 0.7 N/m, resonant frequency of 150 kHz and tip radius < 20 nm were used (ScanAsyst Fluid). All images were captured at a 1.11 Hz scan rate and a resolution of 512 x 512 pixels.

3.4.9 Acquisition of TH profiles and isothermal kinetics measurements

Samples contained 50 μM d(A15) and 10 or 15 mM CA, and 75 μM d(A10) and 15 mM CA in 1xAcMg pH 4.5. CD and UV-Vis absorbance were monitored at 252 nm for annealing and melting, **Table 3.3** summarizes the start and end temperature conditions as well as rates of temperature change used for each sample set. The melting/annealing rates were selected to ensure good separation between the curves. Samples were maintained at the top temperature for 5 minutes before annealing and at 2 °C for 5 minutes before melting. A layer of silicon oil was applied on

top of the sample solution to minimize evaporation. A stream of nitrogen gas was supplied to the sample chamber to prevent condensation on the cuvette with heating. Curves were obtained in triplicate. Sample TH profiles obtained by CD and UV-VIS are shown in **Figure 3.22**.

	Strand	[d(A_n)] (µM)	[CA] (mM)	Top and bottom temperatures (°C)	Scan rates (°C/min)	Isothermal kinetics temperatures (°C)
1	d(A_{15})	50	15	65 / 2	± 0.5, 1, 2, 3, 4, 5	10, 15, 20, 25
2	d(A_{15})	50	10	55 / 2	± 0.2, 0.5, 1, 2, 3, 4	5, 10, 15, 20
3	d(A_{10})	75	15	50 / 2	± 0.2, 0.5, 1, 2, 3, 4	10, 15, 17.5, 20

Table 3.3| Sample set description, melting/annealing scan parameters and temperatures used for isothermal kinetics measurements.

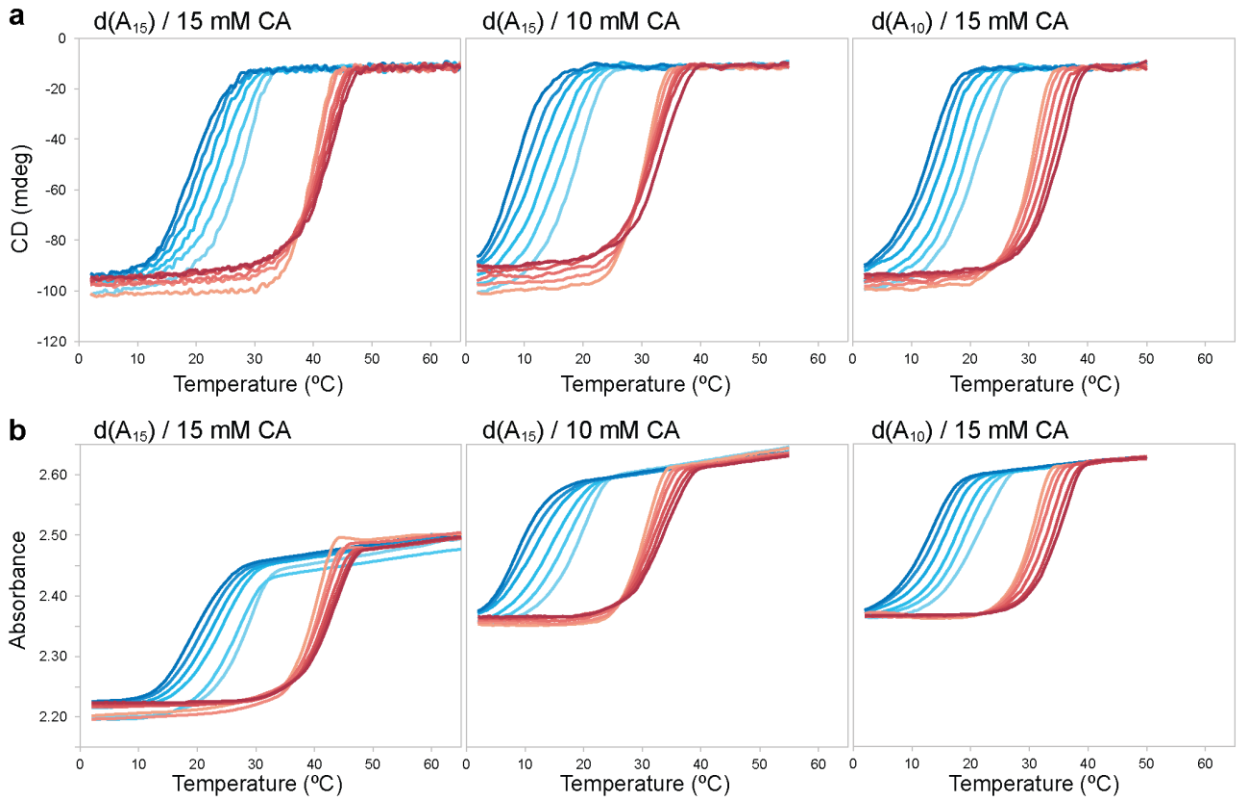


Figure 3.22| Sample TH profiles obtained by CD and UV-Vis absorbance at 252 nm. a) Profiles obtained by CD. **b)** Profiles obtained by UV-Vis absorbance. From left to right: d(A_{15}) with 15 mM CA; d(A_{15}) with 10 mM CA; d(A_{10}) with 15 mM CA. Parameters for data acquisition are listed in **Table 3.3**.

Isothermal assembly kinetics experiments were performed at the temperature settings listed in **Table 3.3** using a temperature jump method. The sample was heated at 65 °C for 5 minutes before being placed directly at the target temperature in the CD sample chamber. CD and UV-Vis absorbance signals at 252 nm were monitored at 1 s intervals for 30 minutes. Curves were obtained in triplicate.

3.4.10 Thermal correction of TH profiles

The temperature control unit of the CD instrument used to obtain thermal melting and annealing data adjusts and reports the temperature of the sample block housing the sample cuvette during the experiment. There is no internal check verifying that sample temperature matches the one recorded by the instrument. Discrepancies in temperature vary with experimental scan rate and scan direction (heating vs. cooling). To quantify the temperature discrepancy and correct the measured block temperatures, the temperature of a buffer solution was monitored with a digital thermocouple during heating and cooling scans between 5 and 65 °C (heating block temperature) at various scan rates (± 0.5, 1, 2, 3, 4 and 5 °C min^{-1}, where positive and negative values denote heating and cooling scans respectively).

Strongly linear correlations between solution and block temperatures (represented by Equation 3.8, R^2 = 0.99 to 1.00) were measured for all scan rates (**Figure 3.23**), where *m* is the slope of the temperature correlation and *b* is the intercept corresponding to the temperature offset. The values of *m* were found to be ~ 0.985 across all scan rates (**Figure 3.24 a**). The values for *b* were found to have a linear correlation with scan rate (**Figure 3.24 b**), allowing interpolation of the offsets at slow scan rates (such as ± 0.2 °C min^{-1}) using Equation 3.9.

$$T_{Solution} = m\,T_{Block} + b \qquad \text{(Equation 3.8)}$$

$$b = -0.5578\frac{dT}{dt} + 1.9421 \qquad \text{(Equation 3.9)}$$

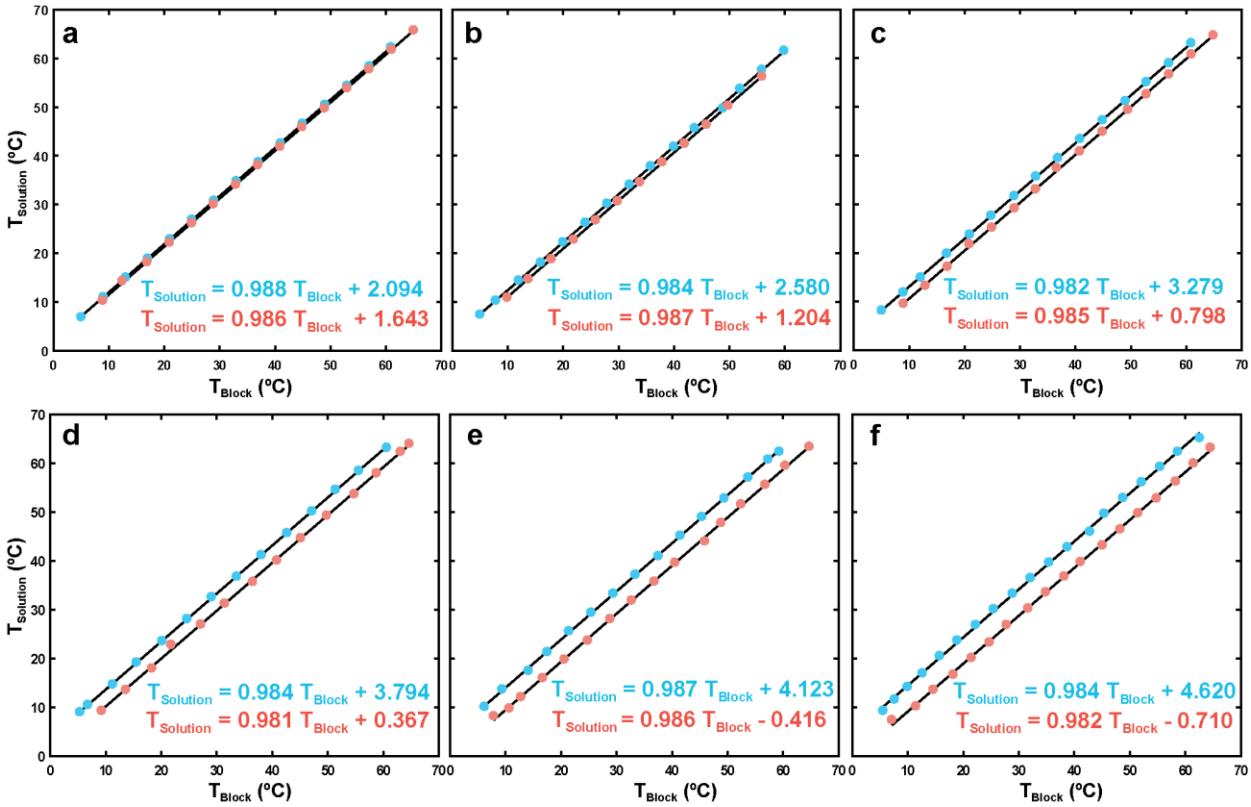


Figure 3.23| Correlations between block and solution temperatures at different scan rates. Scan rates: **a)** ± 0.5 °C min-1, **b)** ± 1 °C min-1, **c)** ± 2 °C min-1, **d)** ± 3 °C min-1, **e)** ± 4 °C min-1, **f)** ± 5 °C min-1. Data points for cooling (-) and heating (+) scans are shown as blue and red circles respectively. Linear fits to the correlations are shown as black lines, with the equations describing the fit displayed in the corresponding colour.

For all experimental thermal melting and annealing data, the block temperature was corrected by applying Equation 3.8, using m = 0.985 and b = temperature offset at a given scan rate. This correction shifted the TH profiles to different temperature domains depending on scan rate, such that a linear interpolation was necessary to place them in the same temperature sampling domain, with data sampling at the same interval as the uncorrected experimental data (0.5 °C). Panels **c** and **d** of **Figure 3.24** show an example of TH profiles after the UV-Vis absorbance data set has been temperature-corrected and interpolated (a replicate of d(A_{15}) with 15 mM CA data set is shown). The same treatment was applied to all data for CA-mediated assembly of d(A_n). The temperature-corrected and interpolated UV-Vis data were used for all following analyses.

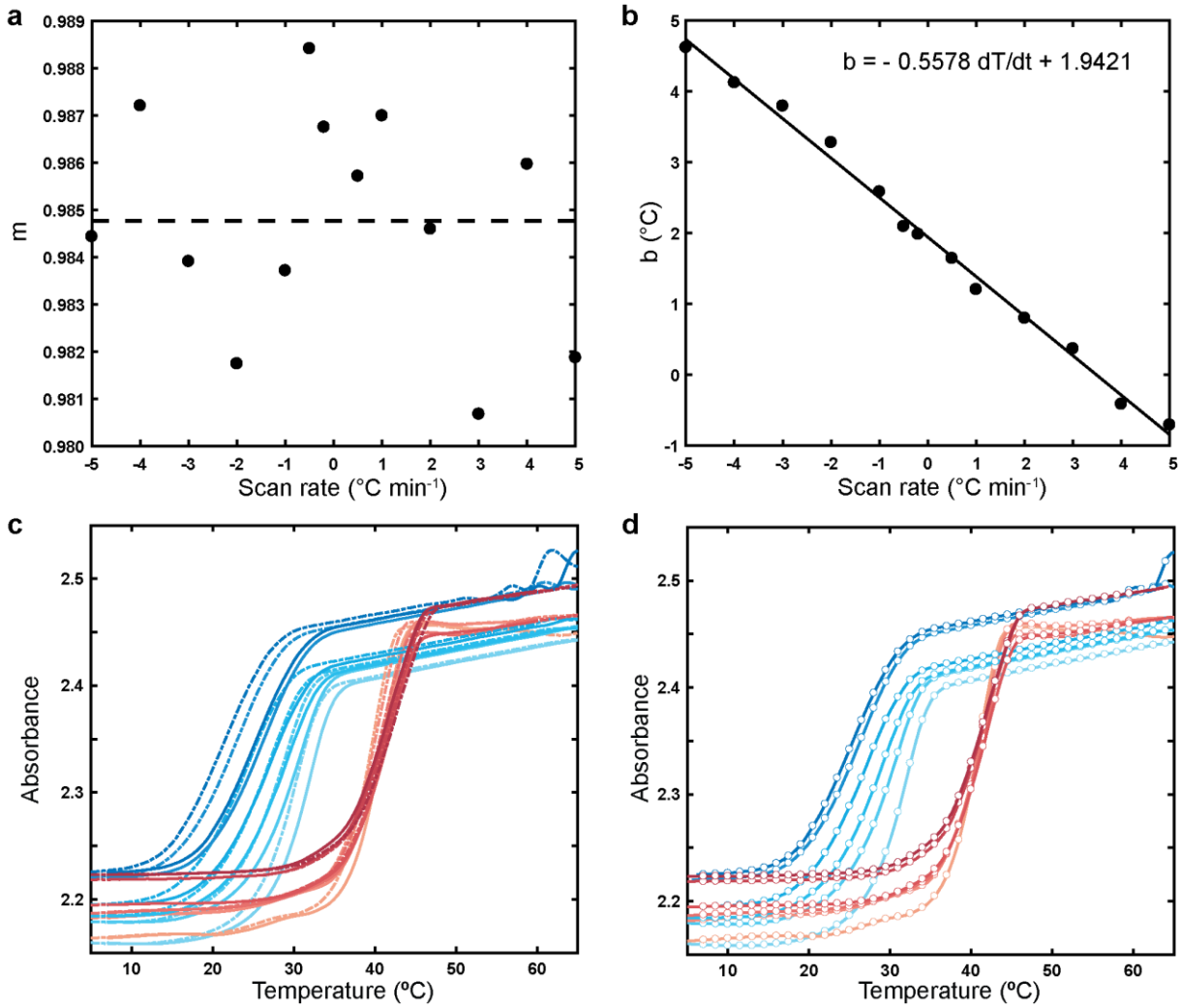


Figure 3.24| Temperature correction of TH profiles. a) Slopes from the correlations in **Figure 3.23** as a function of scan rate. The mean slope m = 0.9848 is indicated by the dashed black line. **b)** Solution temperature offset with respect to block temperature obtained from correlations in **Figure 3.23** as a function of scan rate (dT/dt). The linear fit is shown as a black line. **c)** Uncorrected (dashed lines) and temperature-corrected (blue and red lines for annealing and melting respectively) TH profiles for d(A_{15}) with 15 mM CA. **d)** Linearly interpolated TH profiles (blue and red circles for annealing and melting respectively) overlaid with temperature corrected profiles from panel **c**. Every third interpolated point is shown for clarity. The interpolation was performed to place the corrected data in the same temperature domain at a 0.5 °C sampling frequency.

3.4.11 Model-free analysis of TH profiles

The analysis was performed by **RWH**. The temperature-corrected and interpolated UV-Vis absorbance TH profiles obtained at the given annealing/melting rates (**Table 3.3**) were first fit

with linear baselines for assembled (A_F) and monomeric (A_U) signals according to Equations 3.10 and 3.11, where m_F/b_F and m_U/b_U are the slope/intercept for the folded and monomeric baselines respectively. Using the baselines, the TH profiles were converted to fraction unfolded θ_U using Equation 3.12, where *A(T)* is an absorbance data point on the melting/annealing curve. $\theta_U(T)$ takes on limiting values of 0 and 1 respectively at low and high temperatures, corresponding to fully assembled and fully disassembled states. At each temperature point, the total d(A_n) concentration in the sample C_T, is related to the concentrations of the monomeric *[M]* and assembled states *[F]* according to Equation 3.13, where *N* is the number of monomers within the assembly. θ_U was used to calculate *[M](T)* using Equation 3.14.

$$A_F(T) = m_F T + b_F \quad \text{(Equation 3.10)}$$

$$A_U(T) = m_U T + b_U \quad \text{(Equation 3.11)}$$

$$\theta_U(T) = \frac{A(T) - A_F(T)}{A_U(T) - A_F(T)} \quad \text{(Equation 3.12)}$$

$$C_T = [M](T) + N[F](T) \quad \text{(Equation 3.13)}$$

$$[M](T) = \theta_U(T) C_T \quad \text{(Equation 3.14)}$$

The slopes of monomer concentration with respect to temperature $\frac{d}{dT}[M](T)$ were calculated numerically using rolling window regression where the derivative of a 3rd order polynomial fit to the calculated *[M](T)* in a centered five point moving window (experimental increment of 0.5 °C between data points) is used to calculate local slope (*movingslope*[49] function in MATLAB). The rate of change of *[M](T)* with time $\frac{d}{dt}[M](T)$ was obtained from the slopes and heating/cooling rate according to Equation 3.15 (rate is positive for heating and negative for cooling). The heating/cooling rate dependent $\frac{d}{dt}[M](T)$ and *[M](T)* were used to build 3D surface maps of supramolecular assembly with temperature as the 3rd axis.

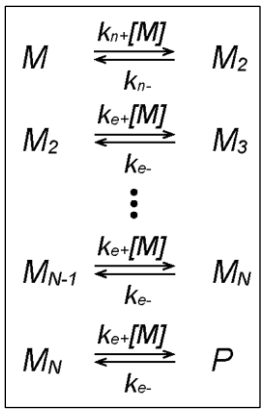
$$\frac{d}{dt}[M](T) = \frac{dT}{dt}\frac{d}{dT}[M](T) \quad \text{(Equation 3.15)}$$

To determine apparent reaction orders through the annealing/melting transitions, slices of the surfaces were taken at sequential temperature points and log-log analysis was applied using Equations 3.1-3.4. The intercepts correspond to effective rate constants k_{on} and k_{off}, however, they

contain contributions from a number of processes and are not meaningful for supramolecular pathway analysis. The method for extraction of effective reaction orders using model-free analysis outlined above requires: (i) adequate separation between curves obtained at different melting/annealing rates, i.e. melting/annealing profiles acquired at different melting/annealing rates must differ from each other substantially more than the scatter due to experimental noise; (ii) assembly and disassembly portions of TH data occurring independently of each other (a three-fold difference in calculated annealing/melting rates in the middle of the annealing transition at the slowest rate of temperature change is deemed sufficient); (iii) analysis is restricted to the ~10-90% θ_U region for the rate *vs*. concentration plots for better accuracy, away from the baseline regions.

3.4.12 Global fit of TH profiles

The global fits were carried out by **RWH**. The TH profiles were globally fit with the Goldstein-Stryer model for cooperative polymerization (Scheme **3.1**),[43] assuming reversible, cooperative stepwise association of monomers M to form nuclei M_s, which then elongate into fibers M_N. There are two distinct phases in the model: (i) Pre-nucleus equilibria governed by nucleation rate constants k_{n+} and k_{n-}; (ii) Post-nucleus equilibria governed by elongation rate constants k_{e+} and k_{e-}. To limit the number of equations that must be numerically integrated, only fibers up to size N are explicitly described. All larger structures are treated as a reversibly formed fibril pool P, an approximation from Korevaar *et al*. that was shown to improve numerical accuracy compared to truncation at a fiber length of N.[31,34] The rate equations presented in **Scheme 3.2** describe the model.



Scheme 3.1| The Goldstein-Stryer model for cooperative self assembly with reversible fibril pool approximation from Korevaar. The monomer M associates in a step-wise manner to form a nucleus of a defined size M_s which then elongates to give large assemblies. The two regimes are defined by nucleation and elongation rate constants. The case for a nucleus size s = 2 is shown (M_2). Post-nucleus oligomers elongate up to an explicitly described size of N beyond which they are treated as a fibril pool P according to the approximation by Korevaar et *al*.

Monomer:

$$\frac{d}{dt}[M] = -k_{n+}[M]\left(2[M] + \sum_{i=2}^{s-1}[M_i]\right) - k_{e+}[M]\left(\sum_{i=s}^{N}[M_i] + [P]\right) + k_{n-}\left(2[M_2] + \sum_{i=3}^{s}[M_i]\right) + k_{e-}\left(\sum_{i=s+1}^{N}[M_i] + [P]\right)$$

Pre-nucleus oligomers:

$$\frac{d}{dt}[M_i] = k_{n+}[M]([M_{i-1}] - [M_i]) + k_{n-}([M_{i+1}] - [M_i])$$

Nucleus:

$$\frac{d}{dt}[M_s] = k_{n+}[M][M_{s-1}] - k_{e+}[M][M_s] + k_{e-}[M_{s+1}] - k_{n-}[M_s]$$

Post-nucleus fibers:

$$\frac{d}{dt}[M_i] = k_{e+}[M]([M_{i-1}] - [M_i]) + k_{e-}([M_{i+1}] - [M_i])$$

Fiber length *N*:

$$\frac{d}{dt}[M_N] = k_{e+}[M]([M_{N-1}] - [M_N]) + k_{e-}((1-\alpha)[P] - [M_N])$$

Fibril number concentration:

$$\frac{d}{dt}[P] = k_{e+}[M][M_N] - k_{e-}(1-\alpha)[P]$$

Fibril mass concentration:

$$\frac{d}{dt}[Z] = k_{e+}[M]((N+1)[M_N] + [P]) - k_{e-}([P] + N(1-\alpha)[P])$$

Where α is given by:

$$\alpha = 1 - \left(\frac{[P]}{[Z] - N[P]}\right)$$

Scheme 3.2| Ensemble of rate equations describing the Goldstein-Stryer cooperative polymerization model, including the reversible fibril pool from Korevaar *et al*.

The global fits assumed that $k_{n+} = k_{e+}$,[31] and the temperature dependence of rate constants are given by Equation 3.16, where k_0 is the rate constant at a reference temperature T_{ref} and E_a is the activation energy. The fits also allowed for non-zero nucleic acid ΔC_ps[46,50] in the nucleation and elongation steps by including the $\Delta C_p^{\ddagger}$ parameter in calculating temperature-dependent activation energies according to Equation 3.17, where $E_{n+}{}^0$ is the activation energy at the reference temperature. Only the forward equation is shown for brevity. The differential rate equations from **Scheme 3.2** were numerically integrated as a function of temperature using inverse heating/cooling rate dt/dT and fractions of unfolded polymer θ_U and polymerized state θ_F according to Equations 3.18 and 3.19 to calculate absorbance profiles using Equation 3.20.

$$k(T) = k_0 e^{\frac{E_a}{R}\left(\frac{1}{T_{ref}} - \frac{1}{T}\right)} \quad \text{(Equation 3.16)}$$

$$E_{n+}(T) = E^0_{n+} + \Delta C_p^{\ddagger}(T - T_{ref}) \quad \text{(Equation 3.17)}$$

$$\theta_U = \frac{[M]}{C_T} \quad \text{(Equation 3.18)}$$

$$\theta_F = \frac{\sum_{i=2}^{N} i[M_i] + [Z]}{C_T} \quad \text{(Equation 3.19)}$$

$$A(T) = A_F(T)\theta_F(T) + A_U(T)\theta_U(T) \quad \text{(Equation 3.20)}$$

Global fits to TH profiles were carried out by varying kinetics parameters to minimize the residual sum of squares (*RSS*) between experimental and fitted absorbance data according to Equation 3.21, where $A_j^{exp}(T_k)$ and $A_j^{calc}(T_k, \xi)$ are the j^{th} experimental and fitted absorbance respectively, T_k is the k^{th} experimental temperature and ξ is the set of rate constants, activation energies and ΔC_ps describing the supramolecular polymerization.

$$RSS = \sum_{j=1}^{N} \sum_k \left(A_j^{exp}(T_k) - A_j^{calc}(T_k, \xi)\right)^2 \quad \text{(Equation 3.21)}$$

The nucleus size *s* was varied from 1 to 7 (1 represents an isodesmic polymerization $k_{e+} = k_{n+}$ and $k_{e-} = k_{n-}$) to optimize the fit and agreement with experimentally determined apparent reaction orders *n*. As mentioned in section 3.2.4, *s* = 3 gave the best fit for d(A_{15}) with 15 mM CA (**Figure 3.25**). An *s* = 3 nucleus size also gave the best fit for d(A_{15}) with 15 mM CA (*s* varied from 2 to 3) and *s* = 2 gave the best fits for d(A_{10}) with 15 mM CA (*s* varied from 2 to 4). Changing the fiber size value *N* (*N* = 50, 100, 200) was not shown to affect global fit results, highlighting the utility of the reversible fibril pool approach from Korevaar *et al*.[31,34] The global fit parameters at given reference temperatures are presented in **Table 3.4**.

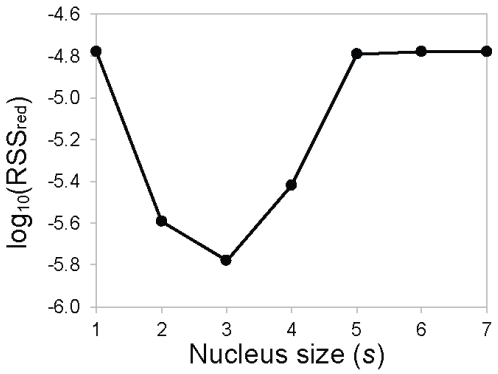


Figure 3.25| Global fit quality as a function of nucleus size *s* for d(A15) with 15 mM CA. The reduced RSS (RSS_{red}) was calculated as RSS_0/DF (DF = # points - # fit parameters).

Sample		d(A15) / 15mM CA	d(A15) / 10mM CA	d(A10) / 15mM CA
Nucleus size (*s*)		3	3	2
Reference temperature (°C)		25	15	18
Activation energies	$E_{n+} = E_{e+}$	14.5 ± 1.1	31.2 ± 1.9	28.5 ± 1.7
	E_{n-}	96.4 ± 2.0	110.6 ± 3.4	146.0 ± 4.2
	E_{e-}	133.0 ± 0.8	74.6 ± 2.6	65.2 ± 2.2
Rate constants	$k_{n+} = k_{e+}$	$(6.7 \pm 0.2) \times 10^4$	$(5.0 \pm 0.6) \times 10^5$	$(1.8 \pm 0.1) \times 10^4$
	k_{n-}	6.8 ± 0.3	$(3.7 \pm 0.7) \times 10^2$	6.4 ± 0.8
	k_{e-}	$(5.8 \pm 0.4) \times 10^{-2}$	4.9 ± 0.7	$(3 \pm 0.2) \times 10^{-1}$
Heat capacities	$\Delta C_p^{\ddagger}{}_{n+} = \Delta C_p^{\ddagger}{}_{e+}$	5.2 ± 0.3	1.0 ± 0.5	6.2 ± 0.3
	$\Delta C_p^{\ddagger}{}_{n-}$	5.5 ± 0.5	-2.3 ± 0.9	13.2 ± 0.6
	$\Delta C_p^{\ddagger}{}_{e-}$	0.5 ± 0.2	-0.6 ± 0.5	5.2 ± 0.3

Table 3.4| TH global fit parameters for CA-mediated d(An) assembly using the Goldstein-Stryer cooperative polymerization model at reference temperatures.[43] Activation energies (in kcal mol^{-1}) and rate constants (in M^{-1} min^{-1} and min^{-1} for forward and reverse steps respectively) are given at the listed reference temperatures. Activation heat capacities are given in kcal mol^{-1} K^{-1}.

Errors in the global fit parameters were calculated using a variance-covariance matrix[51] given by Equation 3.22, where RSS is the residual sum of squared differences between experimental and fitted data points, DF is the degrees of freedom in the fit (N data points minus Φ parameters of the global fit), $\widehat{W}$ is a diagonal matrix of fitting weights (all taken to be 1 in this case), $\hat{X}$ is a matrix of first derivatives of the differences between experimental and calculated data

points with respect to increments in each of the adjustable parameters. The diagonal elements in $\hat{V}$ are the variances of optimized parameters, while the off-diagonal elements are covariances between the errors of optimized parameters. Errors in fitted parameters were expressed as the square root of the variances.

$$\hat{V} = \frac{RSS}{DF}\left(\hat{X}\,\hat{W}\,\widehat{X^T}\right)^{-1} \qquad \text{(Equation 3.22)}$$

The activation free energies for the reaction coordinates were calculated with the Eyring equation. Sample equations for the conversion of monomer (1) to dimer (2) and back are given as Equation 3.23 and Equation 3.24 respectively, where k_B and h are the Boltzmann and Plank constants respectively, $[M]_{ave}$ is the average monomer concentration obtained from log-log plots, and the conversion factor of 60 accounts for the unit change from min^{-1} to s^{-1} for fitted rate constants. Analogous equations are used for subsequent steps in the polymerization, using k_{n+}/k_{n-} and k_{e+}/k_{e-} as applicable to each step along the reaction coordinate.

$$\Delta G^{\ddagger}_{1\rightarrow 2} = -RT\ ln\left(k_{n+}[M]_{ave} \Big/ 60\frac{k_B T}{h}\right) \qquad \text{(Equation 3.23)}$$

$$\Delta G^{\ddagger}_{2\rightarrow 1} = -RT\ ln\left(k_{n-} \Big/ 60\frac{k_B T}{h}\right) \qquad \text{(Equation 3.24)}$$

The monomer was assigned a free energy of 0 kcal mol^{-1} as the reference state and subsequent free energies were calculated according to Equation 3.25, where the sum runs to the final position along the trajectory (hexamer in this case). A cubic spline was fitted to the calculated free energy points to generate the smooth, quantitative reaction coordinates with distances between trough and peak corresponding to the activation energy for a given barrier.

$$G_j = G_1 + \sum_{i=2}^{11} \Delta G_i^{\ddagger} \qquad \text{(Equation 3.25)}$$

3.5 References

1 Fletcher, D. A. & Mullins, R. D. Cell mechanics and the cytoskeleton. *Nature* **463**, 485-492 (2010).

2 Pollard, T. D. & Borisy, G. G. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **112**, 453-465 (2003).

3 Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annu. Rev. Biochem.* **78**, 929-958 (2009).

4 Hynes, R. O. The extracellular matrix: not just pretty fibrils. *Science* **326**, 1216-1219 (2009).

5 Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384-396 (2014).

6 Fowler, D. M., Koulov, A. V., Balch, W. E. & Kelly, J. W. Functional amyloid – from bacteria to humans. *Trends Biochem. Sci.* **32**, 217-224 (2007).

7 Durham, A. C. & Klug, A. Polymerization of tobacco mosaic virus protein and its control. *Nat. New Biol.* **229**, 42-46 (1971).

8 De Greef, T. F., Smulders, M. M., Wolffs, M., Schenning, A. P., Sijbesma, R. P. & Meijer, E. W. Supramolecular polymerization. *Chem. Rev.* **109**, 5687-5754 (2009).

9 Krieg, E., Bastings, M. M., Besenius, P. & Rybtchinski, B. Supramolecular polymers in aqueous media. *Chem. Rev.* **116**, 2414-2477 (2016).

10 Yan, X., Wang, F., Zheng, B. & Huang, F. Stimuli-responsive supramolecular polymeric materials. *Chem. Soc. Rev.* **41**, 6042-6065 (2012).

11 Burnworth, M., Tang, L., Kumpfer, J. R., Duncan, A. J., Beyer, F. L., Fiore, G. L., Rowan, S. J. & Weder, C. Optically healable supramolecular polymers. *Nature* **472**, 334-337 (2011).

12 Sijbesma, R. P., Beijer, F. H., Brunsveld, L., Folmer, B. J. B., Hirschberg, J. H. K. K., Lange, R. F. M., Lowe, J. K. L. & Meijer, E. W. Reversible polymers formed from self-complementary monomers using quadruple hydrogen bonding. *Science* **278**, 1601-1604 (1997).

13 Hartgerink, J. D., Beniash, E. & Stupp, S. I. Peptide-amphiphile nanofibers: A versatile scaffold for the preparation of self-assembling materials. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5133-5138 (2002).

14 Schenning, A. P. H. J., Jonkheijm, P., Peeters, E. & Meijer, E. W. Hierarchical order in supramolecular assemblies of hydrogen-bonded oligo(p-phenylene vinylene)s. *J. Am. Chem. Soc.* **123**, 409-416 (2001).

15 Smulders, M. M., Schenning, A. P. & Meijer, E. W. Insight into the mechanisms of cooperative self-assembly: the "sergeants-and-soldiers" principle of chiral and achiral C3-symmetrical discotic triamides. *J. Am. Chem. Soc.* **130**, 606-611 (2008).

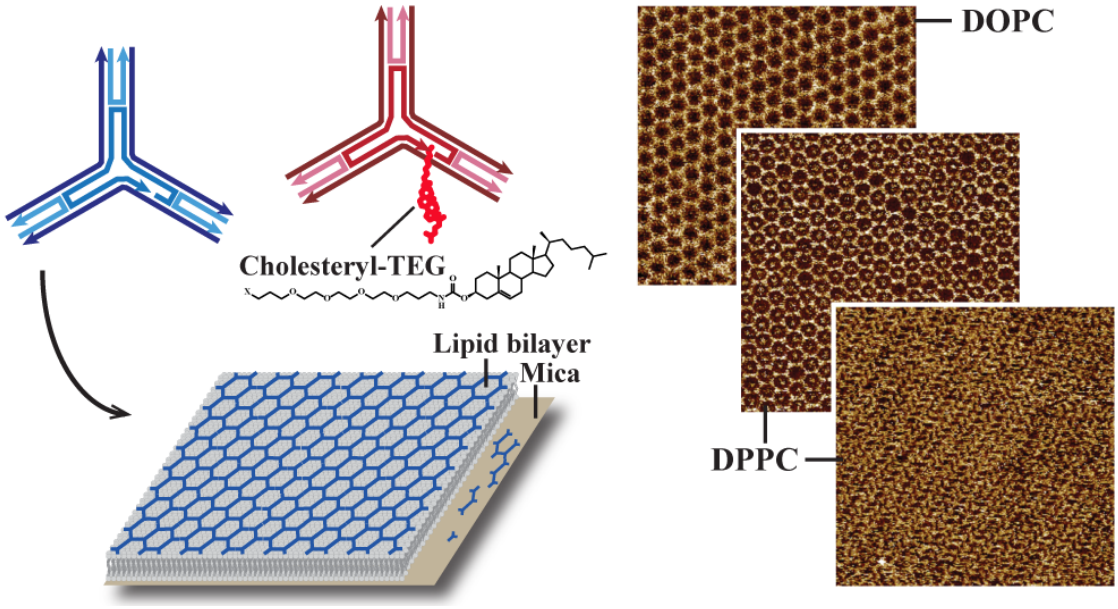
16 Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E. & Khazanovich, N. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **366**, 324-327 (1993).

17 Fenniri, H., Mathivanan, P., Vidale, K. L., Sherman, D. M., Hallenga, K., Wood, K. V. & Stowell, J. G. Helical rosette nanotubes: Design, self-assembly, and characterization. *J. Am. Chem. Soc.* **123**, 3854-3855 (2001).

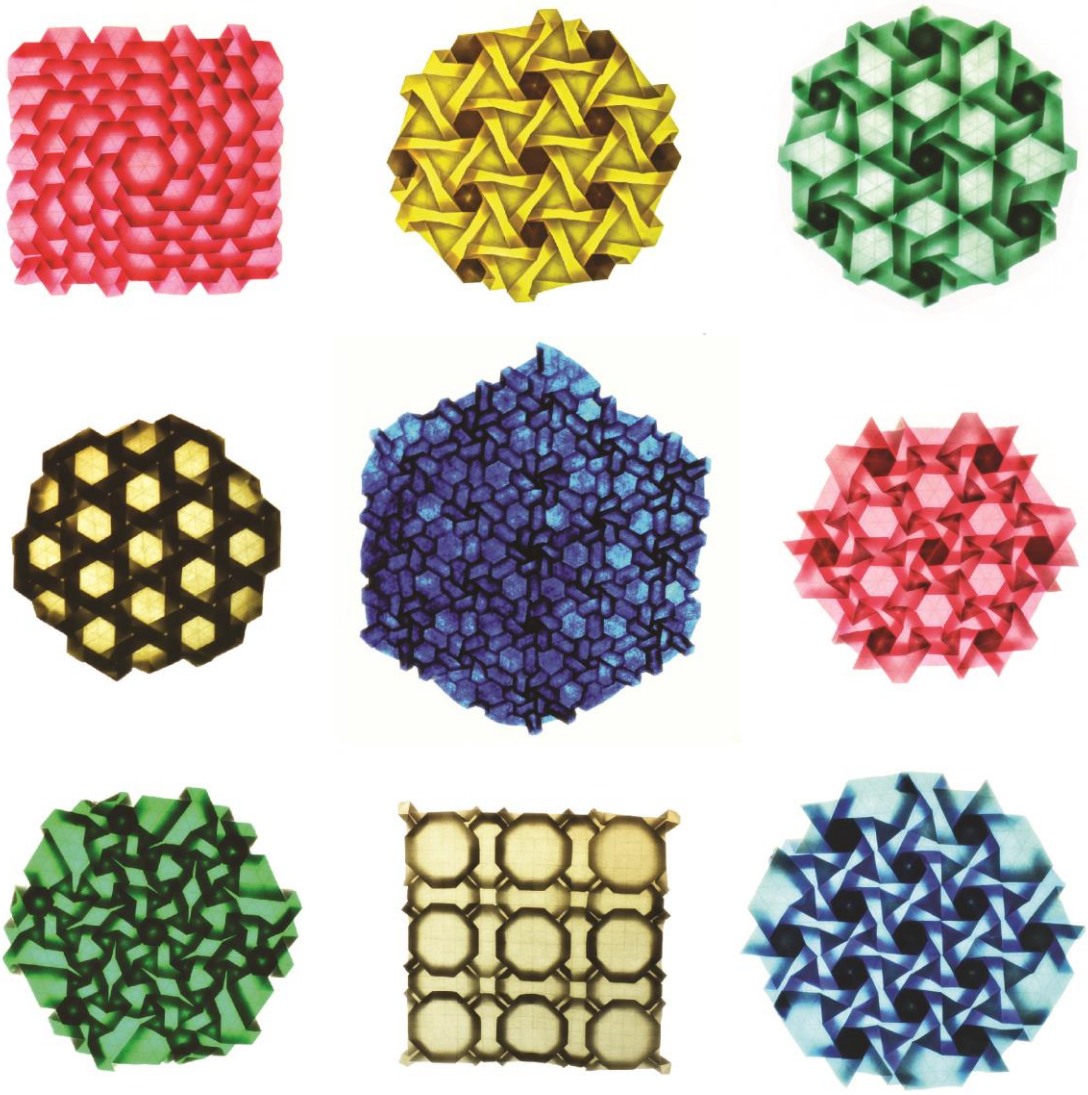
18 Hill, J. P., Jin, W., Kosaka, A., Fukushima, T., Ichihara, H., Shimomura, T., Ito, K., Hashizume, T., Ishii, N. & Aida, T. Self-assembled hexa-peri-hexabenzocoronene graphitic nanotube. *Science* **304**, 1481-1483 (2004).
19 Silva, G. A., Czeisler, C., Niece, K. L., Beniash, E., Harrington, D. A., Kessler, J. A. & Stupp, S. I. Selective differentiation of neural progenitor cells by high-epitope density nanofibers. *Science* **303**, 1352-1355 (2004).
20 Boekhoven, J. & Stupp, S. I. 25th anniversary article: supramolecular materials for regenerative medicine. *Adv. Mater.* **26**, 1642-1659 (2014).
21 Aida, T., Meijer, E. W. & Stupp, S. I. Functional supramolecular polymers. *Science* **335**, 813-817 (2012).
22 Whitesides, G. M. & Grzybowski, B. Self-assembly at all scales. *Science* **295**, 2418-2421 (2002).
23 Sangeetha, N. M. & Maitra, U. Supramolecular gels: Functions and uses. *Chem. Soc. Rev.* **34**, 821-836 (2005).
24 Serpe, M. J. & Craig, S. L. Physical organic chemistry of supramolecular polymers. *Langmuir* **23**, 1626-1634 (2007).
25 Smulders, M. M. J., Nieuwenhuizen, M. M. L., de Greef, T. F. A., van der Schoot, P., Schenning, A. P. H. J. & Meijer, E. W. How to distinguish isodesmic from cooperative supramolecular polymerisation. *Chem. Eur. J.* **16**, 362-367 (2010).
26 Zhao, D. & Moore, J. S. Nucleation-elongation: a mechanism for cooperative supramolecular polymerization. *Org. Biomol. Chem.* **1**, 3471-3491 (2003).
27 Van Gorp, J. J., Vekemans, J. A. & Meijer, E. W. C3-symmetrical supramolecular architectures: fibers and organic gels from discotic trisamides and trisureas. *J. Am. Chem. Soc.* **124**, 14759-14769 (2002).
28 Mizuno, K., Boudko, S. P., Engel, J. & Bachinger, H. P. Kinetic hysteresis in collagen folding. *Biophys. J.* **98**, 3004-3014 (2010).
29 Korevaar, P. A., Newcomb, C. J., Meijer, E. W. & Stupp, S. I. Pathway selection in peptide amphiphile assembly. *J. Am. Chem. Soc.* **136**, 8540-8543 (2014).
30 Schulman, R. & Winfree, E. Synthesis of crystals with a programmable kinetic barrier to nucleation. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15236-15241 (2007).
31 Korevaar, P. A., George, S. J., Markvoort, A. J., Smulders, M. M., Hilbers, P. A., Schenning, A. P., De Greef, T. F. & Meijer, E. W. Pathway complexity in supramolecular polymerization. *Nature* **481**, 492-496 (2012).
32 Ogi, S., Sugiyasu, K., Manna, S., Samitsu, S. & Takeuchi, M. Living supramolecular polymerization realized through a biomimetic approach. *Nat. Chem.* **6**, 188-195 (2014).
33 Kang, J., Miyajima, D., Mori, T., Inoue, Y., Itoh, Y. & Aida, T. A rational strategy for the realization of chain-growth supramolecular polymerization. *Science* **347**, 646-651 (2015).
34 van der Zwaag, D., Pieters, P. A., Korevaar, P. A., Markvoort, A. J., Spiering, A. J. H., de Greef, T. F. A. & Meijer, E. W. Kinetic analysis as a tool to distinguish pathway complexity in molecular assembly: an unexpected outcome of structures in competition. *J. Am. Chem. Soc.* **137**, 12677-12688 (2015).
35 Sorrenti, A., Leira-Iglesias, J., Markvoort, A. J., de Greef, T. F. A. & Hermans, T. M. Non-equilibrium supramolecular polymerization. *Chem. Soc. Rev.* **46**, 5476-5490 (2017).
36 Çetinkol, Ö. P. & Hud, N. V. Molecular recognition of poly(A) by small ligands: an alternative method of analysis reveals nanomolar, cooperative and shape-selective binding. *Nucleic Acids Res.* **37**, 611-621 (2009).

37 Cafferty, B. J., Gallego, I., Chen, M. C., Farley, K. I., Eritja, R. & Hud, N. V. Efficient self-assembly in water of long noncovalent polymers by nucleobase analogues. *J. Am. Chem. Soc.* **135**, 2447-2450 (2013).
38 Marky, L. A. & Breslauer, K. J. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **26**, 1601-1620 (1987).
39 Mergny, J. L. & Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **13**, 515-537 (2003).
40 Hatzakis, E., Okamoto, K. & Yang, D. Thermodynamic stability and folding kinetics of the major G-quadruplex and its loop isomers formed in the nuclease hypersensitive element in the human c-Myc promoter: effect of loops and flanking segments on the stability of parallel-stranded intramolecular G-quadruplexes. *Biochemistry* **49**, 9152-9160 (2010).
41 Mergny, J. L. & Lacroix, L. Kinetics and thermodynamics of i-DNA formation: phosphodiester versus modified oligodeoxynucleotides. *Nucleic Acids Res.* **26**, 4797-4803 (1998).
42 Prislan, I., Lah, J. & Vesnaver, G. Diverse polymorphism of G-quadruplexes as a kinetic phenomenon. *J. Am. Chem. Soc.* **130**, 14161-14169 (2008).
43 Goldstein, R. F. & Stryer, L. Cooperative polymerization reactions. Analytical approximations, numerical examples, and experimental strategy. *Biophys. J.* **50**, 583-599 (1986).
44 Powers, E. T. & Powers, D. L. The kinetics of nucleated polymerizations at high concentrations: amyloid fibril formation near and above the "supercritical concentration". *Biophys. J.* **91**, 122-132 (2006).
45 Frieden, C. & Goddette, D. W. Polymerization of actin and actin-like systems: evaluation of the time course of polymerization in relation to the mechanism. *Biochemistry* **22**, 5836-5843 (1983).
46 Tikhomirova, A., Taulier, N. & Chalikian, T. V. Energetics of nucleic acid stability: the effect of DeltaCP. *J. Am. Chem. Soc.* **126**, 16387-16394 (2004).
47 Harkness, R. W., Avakyan, N., Sleiman, H. F. & Mittermaier, A. K. Mapping the energy landscapes of supramolecular assembly by thermal hysteresis. *Nat. Commun.* **9**, 3152 (2018).
48 Torigoe, H., Nakagawa, O., Imanishi, T., Obika, S. & Sasaki, K. Chemical modification of triplex-forming oligonucleotide to promote pyrimidine motif triplex formation at physiological pH. *Biochimie* **94**, 1032-1040 (2012).
49 D'Errico, J. *Movingslope*, version 1.0 (MATLAB Central File Exchange, 2007).
50 Mikulecky, P. J. & Feig, A. L. Heat capacity changes associated with DNA duplex formation: salt- and sequence-dependent effects. *Biochemistry* **45**, 604-616 (2006).
51 Tellinghuisen, J. Statistical error propagation. *J. Phys. Chem. A* **105**, 3917-3921 (2001).

4| Long-range ordering of blunt-ended DNA tiles on supported lipid bilayers



This chapter is composed of work published as "Long-range ordering of blunt-ended DNA tiles on supported lipid bilayers" by N Avakyan, JW Conway & HF Sleiman, J. Am. Chem. Soc. 139, 12027-12034 (2017).



Crystalline

Models: *Spread hexagons, Pinwheel, Star twist*
Basket weave, Daisy chains, Roman church floor tiling
Double triangle sawtooth, Château-Chinon tiling, 3.6.3.6 tiling

Eric Gjerde. Origami tessellations: Awe-inspiring geometric designs. CRC Press (2008).

Folding, photography and image editing: Nicole Avakyan

Contributions

NA designed and performed all experiments described in this chapter and analysed the results obtained. **Justin W. Conway** initiated the project and performed preliminary experiments.

4.1 Introduction

DNA has emerged as a promising material to build extended networks that enable precise organization of nanomaterials[1-3] and proteins.[4-7] These DNA platforms are found predominantly as long-range 2D tile-based assemblies[8,9] and as networks of origami structures.[10,11] DNA tiles have been used to form 2D-crystalline networks of various geometries, taking advantage of the cohesion of short sticky ends to grow extended arrays on surfaces.[12] Various strategies have also been employed to enable long-range 2D assembly of DNA origami structures.[13] Whereas individual DNA origami nanostructures are built from a large number of unique strands, small tiles require only a few component strands[8,14] – an advantage of the tile-based approach that could simplify design and application, as well as improve scalability.

Traditionally, programmable and long-range assemblies from DNA have been achieved through DNA hybridization between building blocks, whether using short sticky ends or staple strands.[14-16] However, there has been increasing interest in using base stacking (π-stacking between nucleobases) at the blunt ends of DNA duplexes to build large structures. "Blunt-ended" structures do not have single-stranded complementary ends, thus they are not capable of assembling via base-pairing, but rather by π-stacking of their duplex ends. Base stacking, rather than hybridization through hydrogen bonding, is in fact thought to be the dominant stabilizing force in the formation of a DNA double helix from complementary strands.[17] The dynamic character of DNA base-stacking can ensure efficient error correction for long-range assembly and 2D-crystallization. In addition, this interaction can possibly give rise to morphologies differing from those obtained with traditional DNA base-pairing. One dimensional arrays of three-helix bundle DNA motifs have been produced from building blocks presenting blunt duplex ends at their edges.[18] Furthermore, the addition of shape complementarity as a design tool has extended the specificity of this strategy, was used to build programmable DNA origami arrays in 2D[19] and complex dynamic structures from 3D origami components.[20]

Of particular interest in DNA nanotechnology development is interfacing DNA self-assembly with lipid bilayer organization. Lipid membranes organize cell contents through compartmentalization, serve as the interface for various functions such as transport and signalling, while providing a fluid framework to anchor and organize membrane proteins. Moreover, lipid bilayers are highly ordered fluid materials that provide an ideal environment to direct DNA nanostructures to surfaces while maintaining their native structures.[21,22] To test the possibilities of such systems, a number of nanostructures, including DNA origami[23-28] discrete DNA scaffolds,[29-32] 3D-DNA cage structures[33,34] and tile networks[35] have been anchored to lipid bilayers with the addition of hydrophobic modifications such as porphyrins, lipid analogs, cholesterol or amphiphilic backbone conjugates.

Such structures represent potential tools to mimic or probe membrane-related cellular events,[36] but also to build long-range, dynamic platforms that organize functional molecules in 2D on a soft, fluid support. DNA binding to lipid membranes depends on lipid bilayer composition (length and unsaturation of hydrophobic tails, headgroup charge, presence of cholesterol, membrane proteins etc.), lipid bilayer phase behavior, and external factors such as counterion charge and concentration, that are relevant to the electrostatic binding of DNA to the bilayer surface.[22] Polymerization of cholesterol-anchored DNA origami tiles into organized networks, assisted by the addition of staple strands, has been reported.[27] In addition, the work by Suzuki *et al*. has notably demonstrated that unmodified, blunt-ended DNA origami tiles can bind electrostatically to a fluid lipid bilayer and form large lattices.[37] Furthermore, tile diffusion on this substrate promotes error correction during lattice formation and elimination of blunt-end contacts leads to tile rearrangement into a new packing mode.

This chapter reports a facile and scalable method to pattern DNA nanostructures on lipid bilayers. This method combines orthogonal interactions – base stacking between blunt-ended DNA duplexes in addition to cholesterol-mediated DNA anchoring and electrostatic binding to lipid bilayers – to achieve the long-range ordering of small DNA tiles on bilayers with minimal defects. We report on the formation of extensive, micrometer-sized 2D DNA tile arrays from blunt-ended three-point star (3PS)[8,38] tiles on phosphatidylcholine supported lipid bilayers (SLBs). The tiles were either unmodified (**U**) or conjugated to a cholesteryl-TEG (**CHOL**) anchor (**Figure 4.1 a**) and their deposition on SLBs was studied by atomic force microscopy (AFM). At room temperature, the DOPC (1,2-dioleoyl-sn-glycero-3-phosphocholine) and DPPC (1,2-dipalmitoyl-

sn-glycero-3-phosphocholine) lipid bilayers under investigation present fluid (liquid disordered) and gel-phase behavior respectively (**Figure 4.1 b**), leading to distinct interaction with **U** and **CHOL** tiles depending on bilayer composition. On fluid DOPC bilayers, the **CHOL** tiles form extensive hexagonal arrays, while **U** tiles do not bind. In contrast, on the gel-like DPPC bilayers, both the **U** and **CHOL** tiles assemble into long-range networks enabled by Mg^{2+}-mediated electrostatic binding. The divergent tile interactions with fluid and gel-phase bilayers are best observed for a 1:1 mixture of **CHOL** and **U** tiles in mixed bilayer DPPC:DOPC systems. Here, the mixture phase-separates into a DOPC phase with **CHOL** tiles only, and a DPPC phase with **CHOL** and **U** tiles.

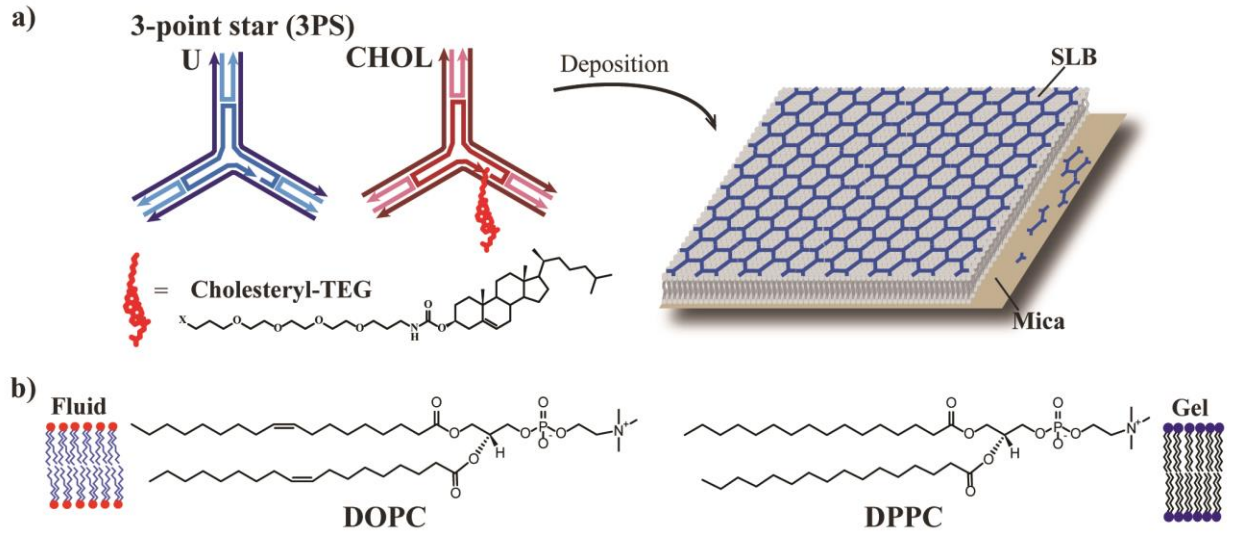


Figure 4.1| Three-point star (3PS) tile deposition on supported lipid bilayers (SLBs). a) Schematic representation of unmodified (**U**) and cholesteryl-TEG (**CHOL**) modified blunt 3PS tiles and their deposition on supported lipid bilayers. **b)** Structures of DOPC and DPPC, which adopt fluid and gel-phase behaviour respectively at ambient conditions.

Furthermore, on DPPC, new tile binding motifs emerge at higher DNA tile concentrations. Tile packing density and therefore array patterns are dependent on tile concentration, illustrating a delicate balance between base-stacking and bilayer-mediated tile-tile interactions. Balancing the dynamic interactions at play, tiles on the lipid bilayer surface lose blunt-end contacts and rearrange into progressively denser but still highly organized packings (*Honeycomb, Honeycomb + Filled Voids, Parallel, Ridges and Superpacked*), to maximize electrostatic contacts between the bilayer

surface and DNA. Disrupting the π-stacking interactions by introducing T-loops on the tile-arm ends prevents any tile organization on SLBs. Similar trends are observed with a tile of different geometry, the cross $(X)^4$ tile.

4.2 Results and discussion

4.2.1 3PS tile design and characterization

To study the deposition and higher order assembly of small DNA tiles on supported lipid bilayers, we have adapted the 3PS tile conceived by the Mao group.[8] The simple 3PS tile design follows the principle of sequence symmetry, thus minimizing the number of strands required for tile assembly and restricting the necessary sequence space, in contrast to larger and more complex DNA structures such as DNA origami.[39] The geometry of the tile, with 3 arms oriented symmetrically at 120° angles (**Figure 4.2 a**), was originally designed to produce hexagonal, or honeycomb, arrays through the interaction of short self-complementary sticky ends on each tile-arm.

To maximize the dynamic nature of the interactions between tiles on a fluid support such as a lipid bilayer, thus favoring tile rearrangement and error correction in the formation of large arrays, we have instead focused on blunt-ended tiles, which are expected to interact through π-stacking interactions between duplex ends.[19,37,40,41] Unmodified (**U**) and cholesteryl-TEG conjugated (**CHOL**) tiles were prepared (**Figure 4.1 a**). The cholesteryl modification was selected due to its well-documented interactions with lipid bilayers as an anchoring unit for DNA nanostructures.[21,22] The moiety was introduced at the 3'-end of the central tile strand because of considerations for spatial separation between the anchor and tile-arm ends, as well as ease of synthesis.

Strand purity was assessed by denaturing polyacrylamide gel electrophoresis (PAGE) analysis (**Figure 4.2 b**), and additionally by liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) for the modified strand S1 CHOL (**Figure 4.3**). Correct assembly of **U** and **CHOL** 3PS tiles was assessed by non-denaturing PAGE analysis (**Figure 4.4**). Blunt-ended tiles do not form aggregates detectable by PAGE, as no bands for structures larger than individual tiles were observed.

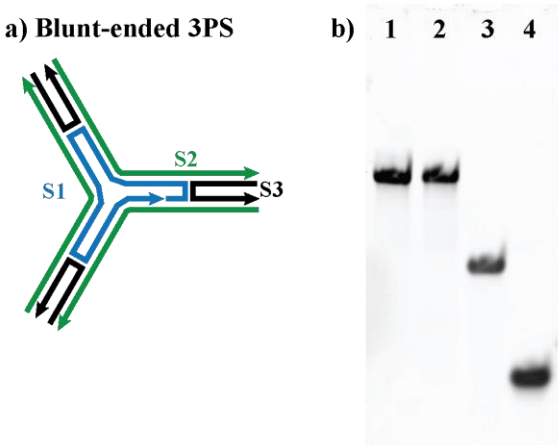


Figure 4.2| 3PS tile structure and strand characterization. a) Schematic representation of the **U** blunt-ended 3PS tile composed of strands S1, S2 and S3 in 1:3:3 stoichiometry. **CHOL** tiles are prepared with strand S1 CHOL. **b)** Denaturing PAGE (12%, 1xTBE, 30 min at 250 V followed by 45 min at 500 V) of 3PS tile component strands. Lane description: 1. S1, 2. S1 CHOL, 3. S2, and 4. S3.

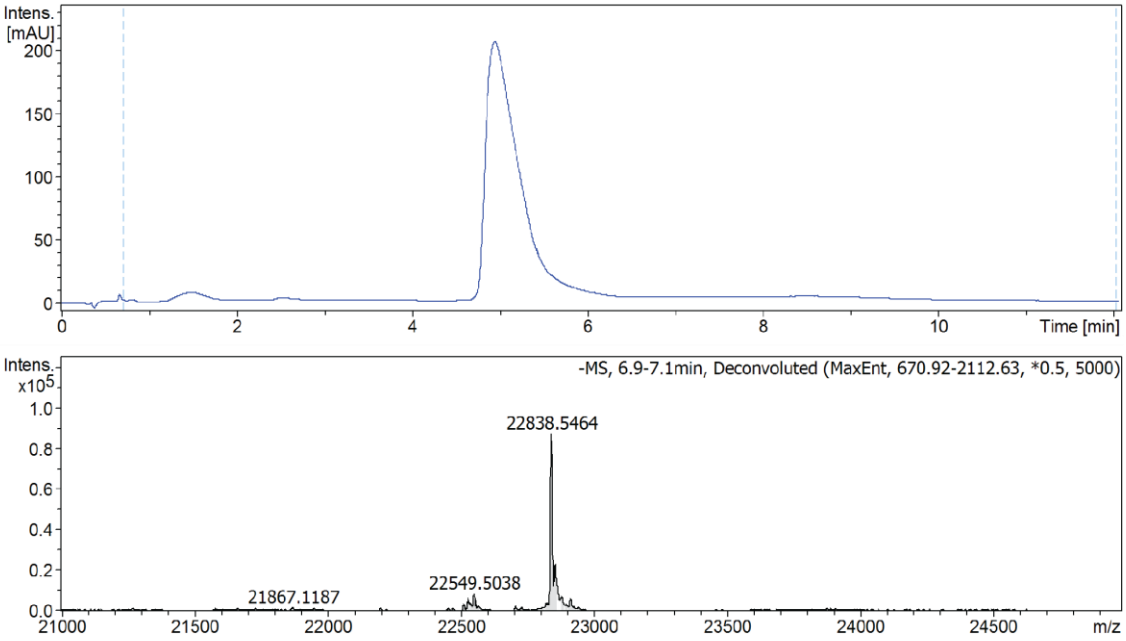


Figure 4.3| LC-ESI-MS characterization of cholesterol modified strand S1 CHOL. LC trace detected by absorbance at 260 nm (top) and mass spectrum (bottom) for strand S1 CHOL (calculated M_W = 22 839 g/mol).

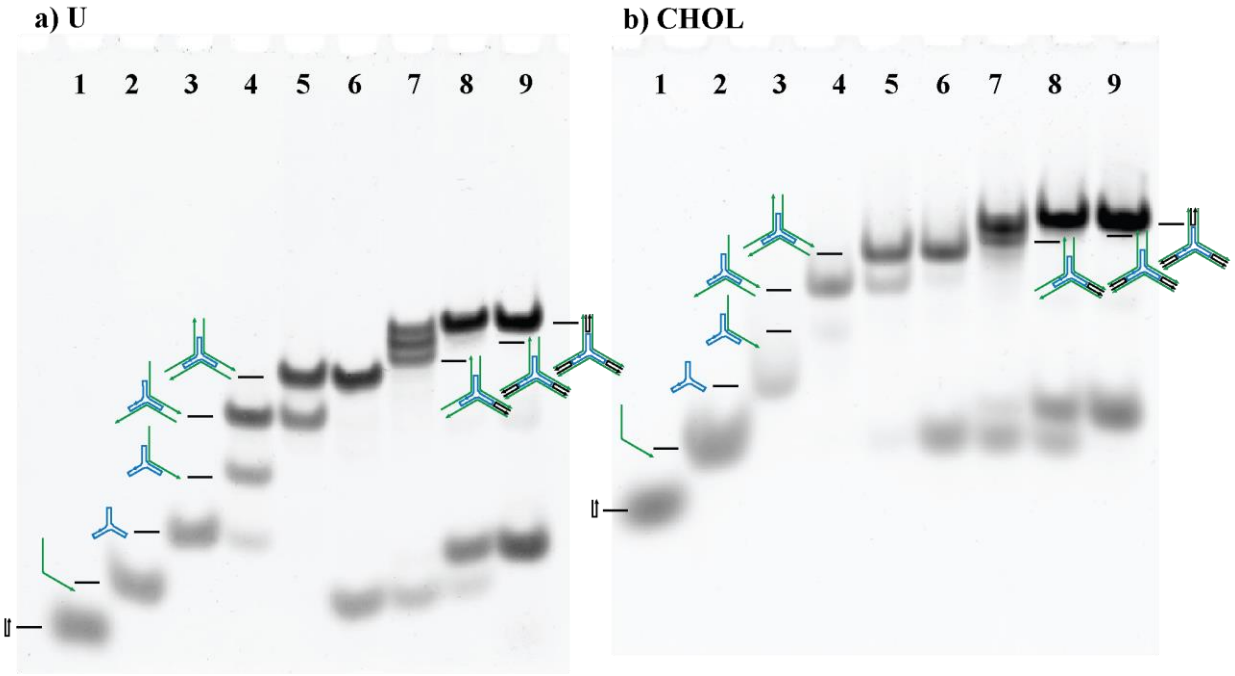


Figure 4.4| Assembly of blunt-ended 3PS tiles analyzed by native PAGE. a) U tile (composed of strands S1, S2 and S3). **b) CHOL** tile (composed of strands S1 CHOL, S2 and S3). Lane description for **(a)**: 1. S3, 2. S2, 3. S1, 4. S1+S2 (1:1), 5. S1+S2 (1:2), 6. S1+S2 (1:3), 7. S1+S2+S3 (1:3:1), 8. S1+S2+S3 (1:3:2), 9. S1+S2+S3 (1:3:3). For **(b)**, same lane description but using the S1 CHOL strand. Native PAGE 6%, TAMg running buffer, 90 min at 120 V and 5 °C.

4.2.2 3PS tile deposition on DOPC and DPPC

The deposition of **U** and **CHOL** 3PS tiles was studied on SLBs by AFM at ambient conditions in the presence of pH 8.0 Tris-acetate buffer containing 12.5 mM Mg^{2+} (TAMg), a standard buffer for the preparation of DNA nanostructures. The SLBs were produced from DOPC and DPPC, synthetic lipids with zwitterionic phosphatidylcholine headgroups (**Figure 4.1 b**). DOPC, the unsaturated variant, exists in a fluid (liquid disordered) phase at room temperature whereas DPPC, the saturated variant, is in a gel state. Stable anchoring of DNA nanostructures with cholesterol and other lipophilic modifications to DOPC bilayers has been previously reported[22] and unmodified DNA origami tiles are also known to bind to DOPC in the presence of Mg^{2+}.[37] The presence of Mg^{2+} also enables binding of DNA nanostructures to DPPC bilayers[35] and preferential binding to the liquid ordered domains in mixed bilayer systems (in bilayers composed

of phospholipids and cholesterol).[22] Such preferential partitioning of unmodified vs. cholesterol-anchored nanostructures has been reported for DNA origami rods.[42]

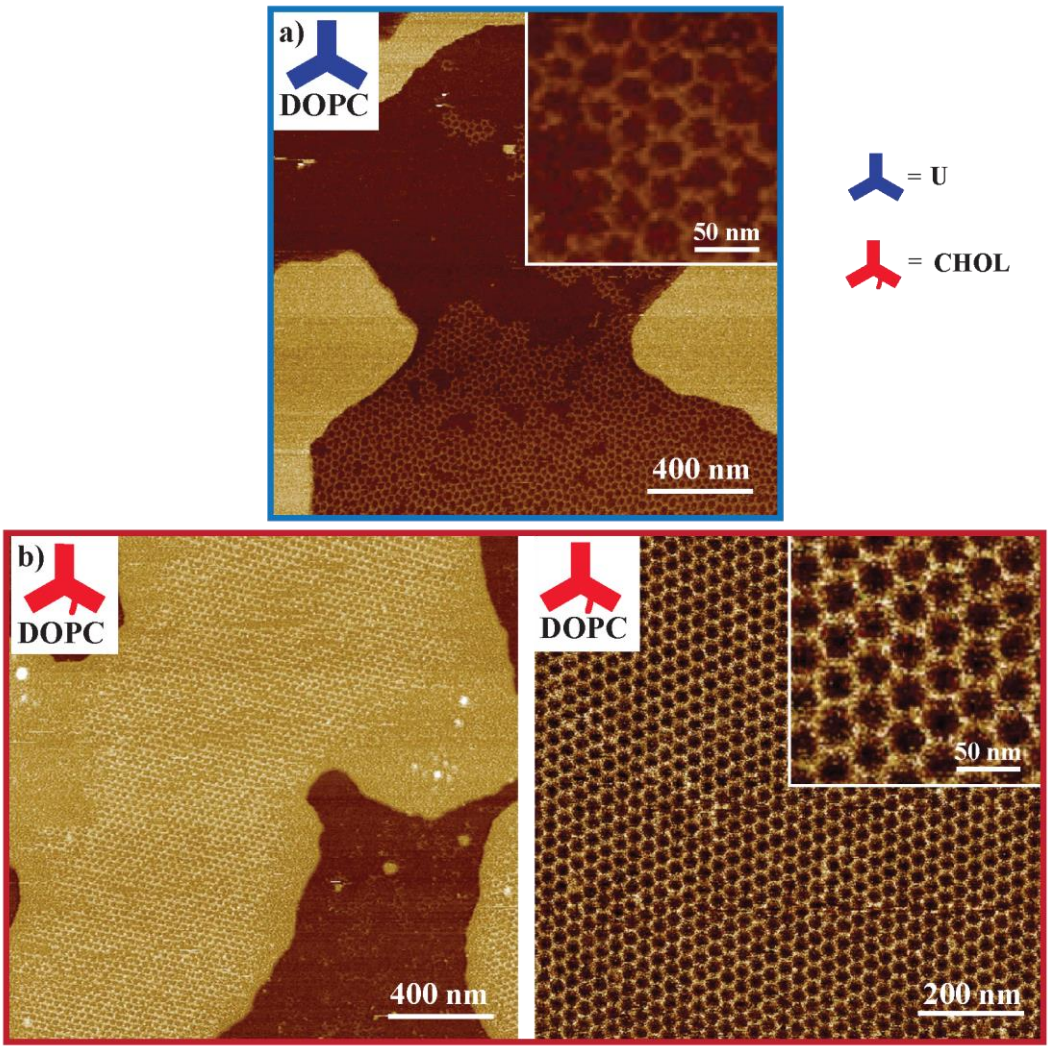


Figure 4.5| 3PS tiles on supported DOPC bilayers. AFM images of: **a) U** tiles; **b) CHOL** tiles, close-up of honeycomb tile array (right). V_{Dep} = 2 µL, 0.5 µM tile solution. Scale bars as indicated on the images.

When deposited on DOPC bilayers, **U** tiles were found to bind preferentially to areas of exposed mica rather than the lipid bilayer (**Figure 4.5 a**), forming small areas of honeycomb array with many defects. On the other hand, **CHOL** tiles were mostly found on the DOPC bilayer, forming much larger, micrometer-sized areas of highly ordered honeycomb array mediated by blunt-end stacking interactions between the tile arm ends (**Figure 4.5 b**). Previous work has shown that unmodified DNA nanostructures such as large origami tiles can bind to DOPC electrostatically

in the presence of Mg^{2+} cations.[37] In contrast, the smaller surface area and less DNA-dense structure of the 3PS tile might explain the absence of **U** tile binding to DOPC in the present case. The small tile size may further favor diffusion and reorientation of **CHOL** tiles preferentially landing on the fluid DOPC bilayer, leading to the formation of larger defect-free honeycomb arrays.

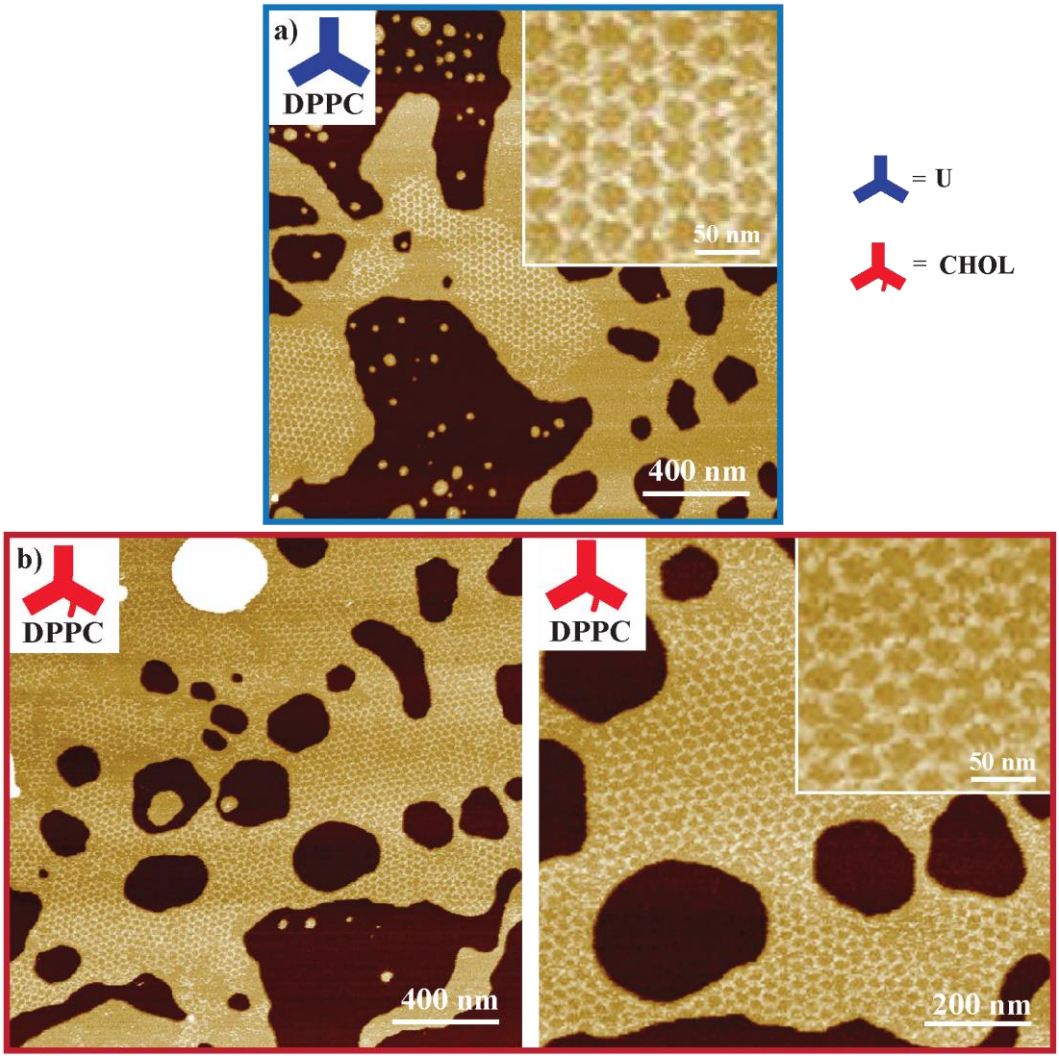


Figure 4.6| 3PS tiles on supported DPPC bilayers. AFM images of: **a) U** tiles, **b) CHOL** tiles, close-up of honeycomb tile array (right). V_{Dep} = 2 µL, 0.5 µM tile solution. Scale bars as indicated on the images.

In contrast to the divergent binding behaviour of **U** and **CHOL** tiles dependent on the anchoring cholesteryl moiety on DOPC, both **U** and **CHOL** tiles bound to the DPPC bilayer, with virtually no tiles detected on exposed mica areas (**Figure 4.6**). The electrostatic binding interaction

between zwitterionic headgroups of DPPC lipids and the negatively charged DNA backbone depends on the presence of divalent cations such as Mg^{2+} and Ca^{2+}.[43,44] The preferential binding of double-stranded DNA to liquid ordered (gel-phase) zwitterionic lipid bilayers over fluid bilayers has been reported and attributed to the closer packing of lipid molecules in the gel-phase and thus a higher charge density on the bilayer surface.[45] Both **U** and **CHOL** tiles bound to DPPC formed large honeycomb array domains. The surface diffusion of tiles bound to DPPC is expected to be lower than on DOPC due to the gel-phase behaviour of the DPPC bilayer.[42] However, it appears that the DPPC surface still allows for more diffusion than the mica surface in the same buffer conditions, as **U** and **CHOL** tiles deposited on bare mica form networks with more defects (Figure 4.7).

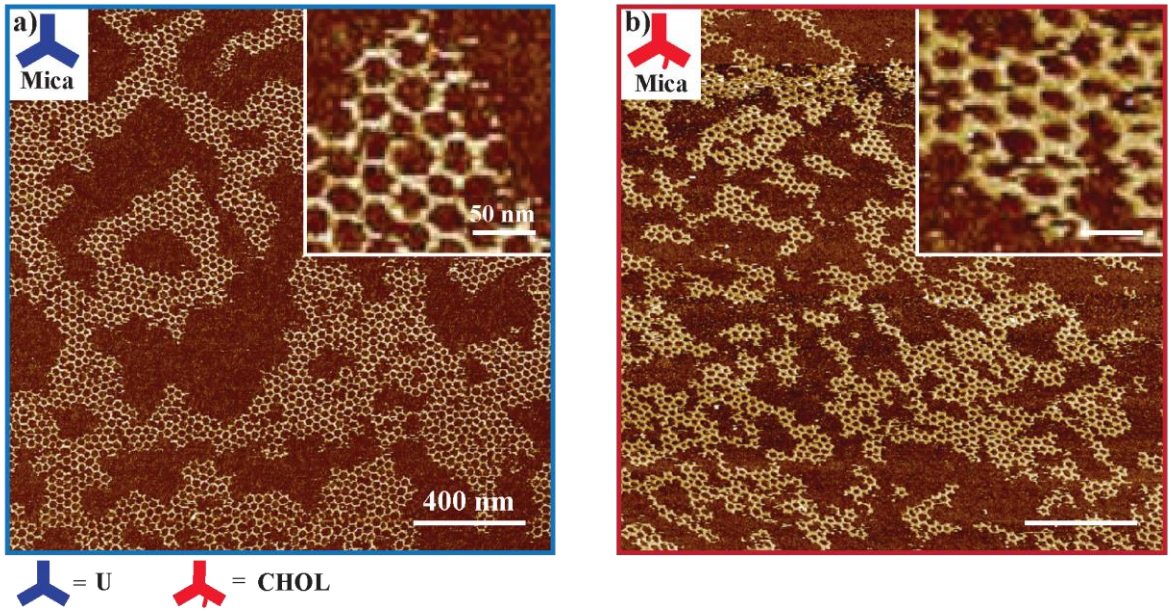


Figure 4.7| U and CHOL 3PS tiles deposited on mica. AFM images of: **a) U** tiles, **b) CHOL** tiles. V_{Dep} = 2 µL, scale bar = 400 nm, inset scale = 50 nm.

4.2.3 Volume of deposition (V_{Dep}) dependent 3PS tile packing on DPPC

3PS tiles do not only form hexagonal arrays on DPPC as higher-density ordered morphologies were also detected. This observation led us to further investigate how tile concentration, or more practically, the volume of tile deposition (V_{Dep}), influences tile packing on the bilayer surface. All AFM imaging was carried out using 0.5 µM tile concentration and V_{Dep}

describes the volume (in μL) of this stock solution deposited on a prepared supported lipid bilayer surface. By increasing V_{Dep} of 3PS tiles on DPPC, we could identify several distinct tile deposition patterns for both **U** and **CHOL** tiles (**Figure 4.8**).

Five different tile packing modes were detected (V_{Dep} from 1 μL to 10 μL), with tile density on the surface increasing progressively with increasing V_{Dep}. The packing stages (**Figure 4.9**) will be referred to as follows: *Honeycomb*, *Honeycomb + Filled Voids*, *Parallel*, *Ridges* and *Superpacking*. At low V_{Dep} (**U** tiles at 1 μL and **CHOL** tiles at 1, 2 and 4 μL V_{Dep}), only areas of honeycomb arrays were detected. In the next packing stage (**U** tiles at 3.5 μL V_{Dep}), the hexagonal voids in the *Honeycomb* network were occupied by individual tiles (*Honeycomb + Filled Voids*). In this arrangement, electrostatic contacts between the DNA tiles and bilayer surface are maximized without affecting the blunt-end stacking interactions that bring about the formation of long-range honeycomb assemblies.

However, with further increases to V_{Dep} (**U** and **CHOL** tiles at 5 μL and 10 μL V_{Dep}), the hexagonal array is disrupted in favour of denser packings where tiles form parallel lines, ridges and zig-zag arrangements (*Parallel*, *Ridges* and *Superpacking* respectively). These tile packing modes share some characteristics, namely that tiles form side by side parallel lines (**Figure 4.10**), possibly to take advantage of tile-shape complementarity and minimize void space in the resulting pattern. Although in the *Parallel* packing mode, it is not possible to detect the orientation of adjacent tiles, both the *Ridges* and *Superpacking* modes feature alternating orientations of adjacent tiles. *Ridges* refer to parallel lines formed of tile pairs of alternating orientation through blunt-end stacking interactions of one tile-arm. In the *Superpacking* mode, zig-zagging 1D tile assemblies use two of the 3PS tile-arms for blunt-end stacking interactions. These assemblies then stack next to each other, minimizing voids in the resulting DNA tile surface. The high-density packing arrangements (*Parallel*, *Ridges* and *Superpacking*) can coexist, however there is a trend toward the densest *Superpacking* state with increasing V_{Dep}. It is possible that *Parallel* patterns appear because of AFM imaging artefacts, such that *Parallel* and *Ridges* packing are not distinct stages; however, these morphologies were reproducibly observed. High density packing arrangements demonstrate that the driving force to maximize electrostatic interactions between DNA and DPPC are sufficiently strong to disrupt or prevent the formation of blunt-end interactions between tiles to maximize the surface area occupied by DNA and thus minimize voids.

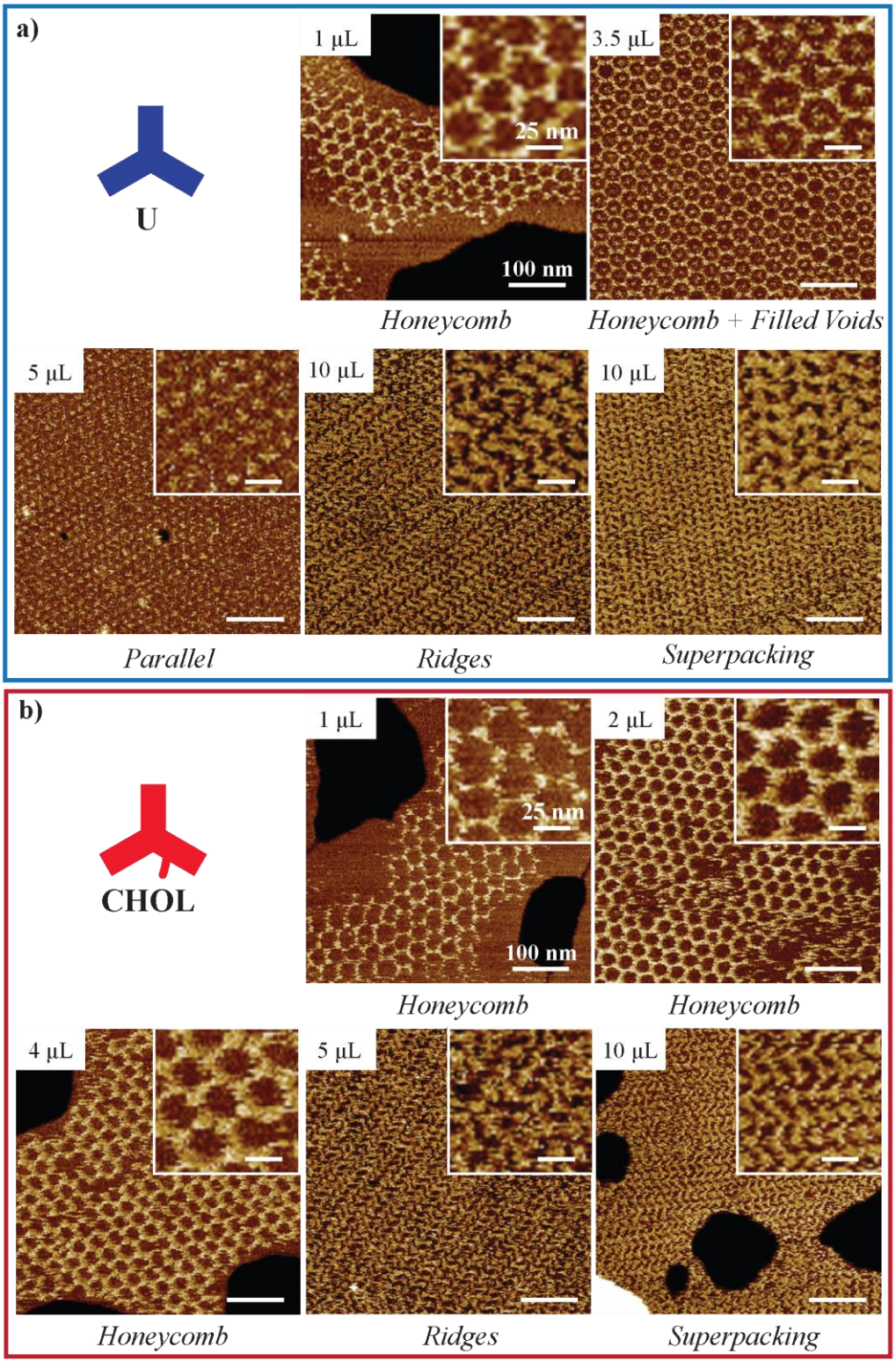


Figure 4.8| 3PS tiles on DPPC bilayers with increasing V_{Dep}. AFM images of: **a) U** tiles (V_{Dep}: 1, 3.5, 5 and 10 µL). With increasing V_{Dep}, the observed packing modes are *Honeycomb*, *Honeycomb + Filled voids*, *Parallel*, *Ridges* and *Superpacking*. **b) CHOL** tiles (V_{Dep}: 1, 2, 4, 5 and 10 µL). With increasing V_{Dep}, the observed packing modes are *Honeycomb*, *Honeycomb*, *Honeycomb*, *Ridges* and *Superpacking*. Scale bar = 100 nm, inset scale = 25 nm.

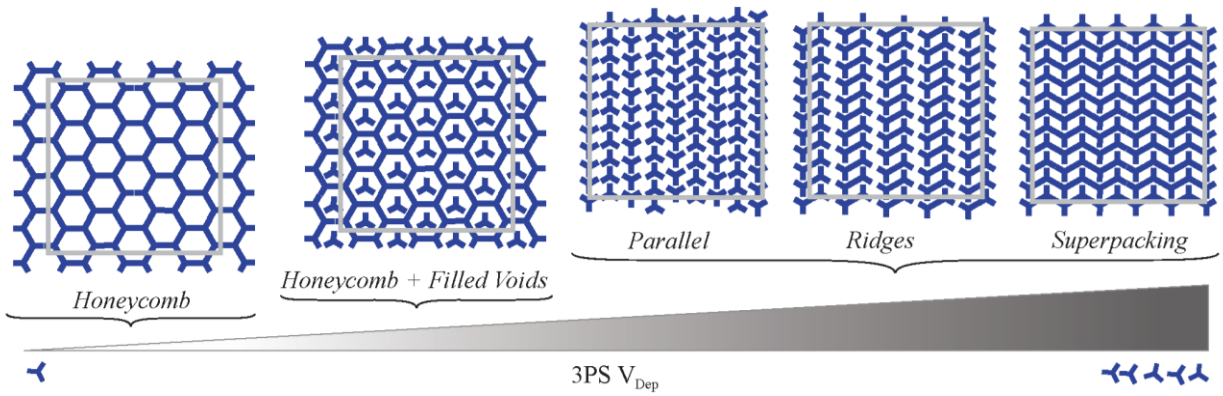


Figure 4.9| Schematic representation of the variation in 3PS tile deposition patterns observed with increasing V_{Dep}. The patterns reflect an increased density of 3PS tiles deposited per unit area and the following stages can be identified from low to high V_{Dep}: *Honeycomb*, *Honeycomb + Filled Voids*, *Parallel*, *Ridges* and *Superpacking*.

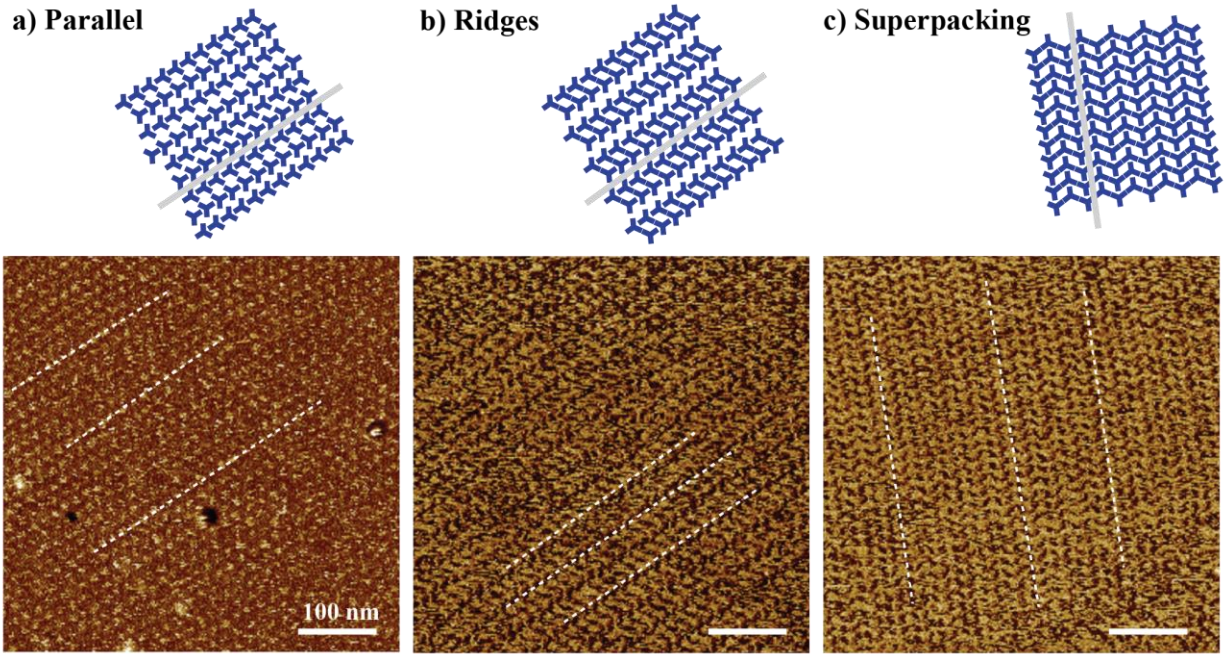


Figure 4.10| Illustration of high density 3PS tile packing arrangements on DPPC. Schematic representation (top) of the tile packing arrangement detected in the AFM images of **U** tiles on DPPC reproduced from **Figure 4.8 a** (bottom). Black and white lines highlight that tiles form parallel lines. **a)** *Parallel* (V_{Dep} = 5 µL), **b)** *Ridges* (V_{Dep} = 10 µL), **c)** *Superpacking* (V_{Dep} = 10 µL). Scale bar = 100 nm.

It is important to note that on DOPC, **CHOL** tiles exclusively form honeycomb arrays over the entire V_{Dep} range (**Figure 4.11**) and that the high density packing modes are not observed, although it is possible to note some tiles in the honeycomb voids on occasion. On DOPC, tile

binding is largely mediated by the cholesterol anchoring unit, and the overall binding affinity of **CHOL** tiles to DOPC is expected to be lower than to DPPC. This is due to weaker electrostatic binding to a fluid bilayer in contrast to a gel-phase bilayer as mentioned previously.[42] In addition, insertion of the rigid cholesteryl structure among unsaturated lipid tails interferes with their packing,[46] with associated entropic costs. These factors would not favor dense packing of cholesterol tiles on DOPC. Thus, once the honeycomb structure is formed from **CHOL** tiles on DOPC, any additional tiles land only transiently, and do not promote rearrangement of tile packing through the disruption of blunt-end stacking contacts. Additionally, **CHOL** tile binding to DOPC domains is suspected to depend on the correct tile orientation on the surface to allow for anchor embedding in the bilayer, before network formation can occur through tile diffusion and optimization of blunt-end contacts. In contrast, both **U** and **CHOL** tiles landing on DPPC domains bind electrostatically independent of anchor orientation, implying that high density binding on DPPC may be under kinetic control.

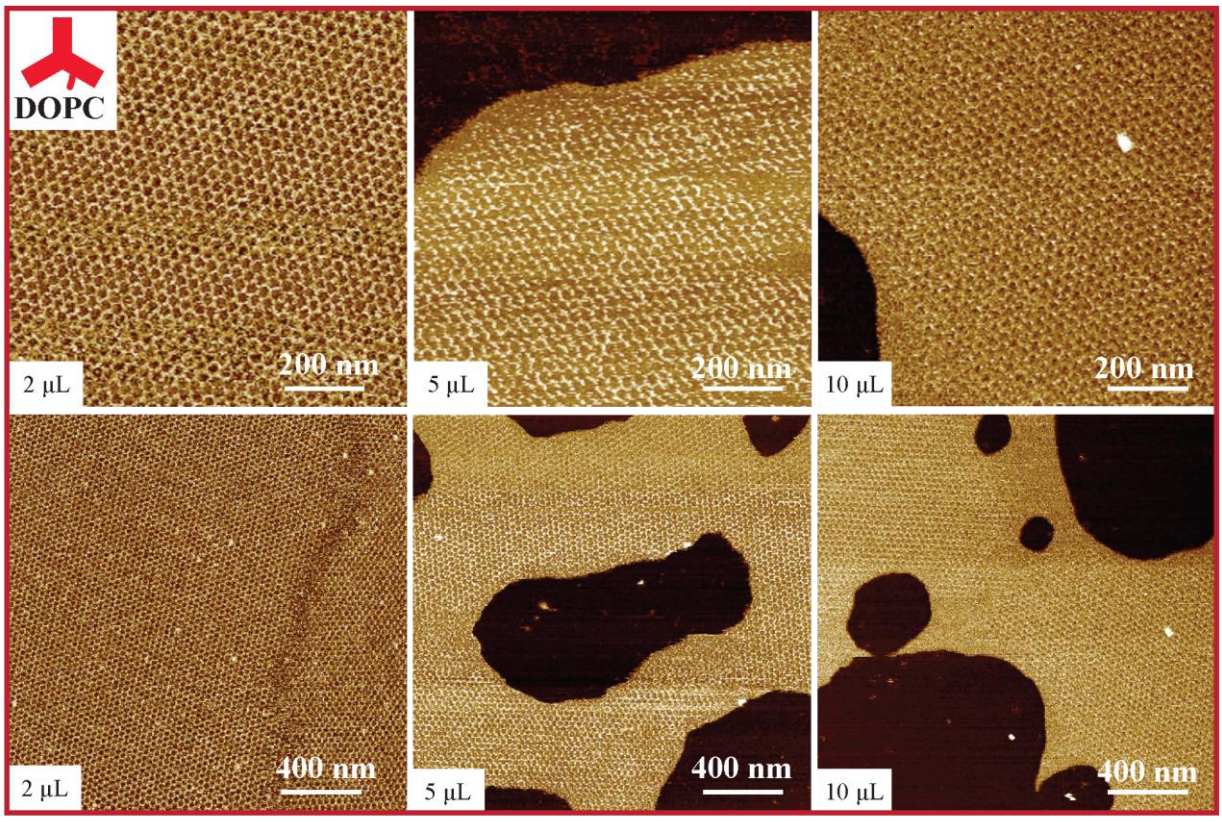


Figure 4.11| CHOL 3PS tiles on DOPC bilayers with increasing V_{Dep}. AFM images of **CHOL** tile deposition at V_{Dep} = 2, 5 and 10 µL (from left to right). V_{Dep} as indicated on the image. Stock tile concentration of 0.5 µM. Scale bars as indicated on the images.

4.2.4 Tile packing transitions on DPPC

The ability of 3PS tiles to rearrange on the DPPC surface once bound to the SLB has important implications for the formation of ordered arrays on this surface. With DOPC, the mobility of lipids composing the bilayer as well as of DNA structures bound to the surface is well documented.[27,37,42] In contrast, with gel-phase DPPC, the close packing of lipid molecules is associated with reduced mobility. In addition, we have observed a high density of DNA nanostructure binding to DPPC above, that is favored compared to binding to mica. It is unclear whether DNA tiles would tend to move on the DPPC surface, as rearrangements of DNA nanostructures on mica are favored in the presence of monovalent cations, rather than the Mg^{2+} used in this system.[47] Furthermore, the transition temperature from gel to fluid phase for DPPC occurs at approximately 42 °C (the transition can be broad[48]). The influence of temperature on tile diffusion and organization on the surface is another interesting aspect of the system.

To examine these aspects of tile organization, we first studied tile packing transitions on DPPC at room temperature with successive *in situ* additions of **U** tiles. An initial mixed packing stage combining *Honeycomb* and *Honeycomb + Filled Voids* domains (V_{Dep} = 2 µL) was observed to shift to a denser packing with further tile addition (additions of 1 µL at a time, reaching a total V_{Dep} of 5 µL), with blunt-end stacking interactions lost to allow a greater number of tiles to bind electrostatically to the bilayer (**Figure 4.12**). This finding suggests that tiles retain mobility on the gel-phase DPPC surface and are capable of rearranging to maximize electrostatic binding.

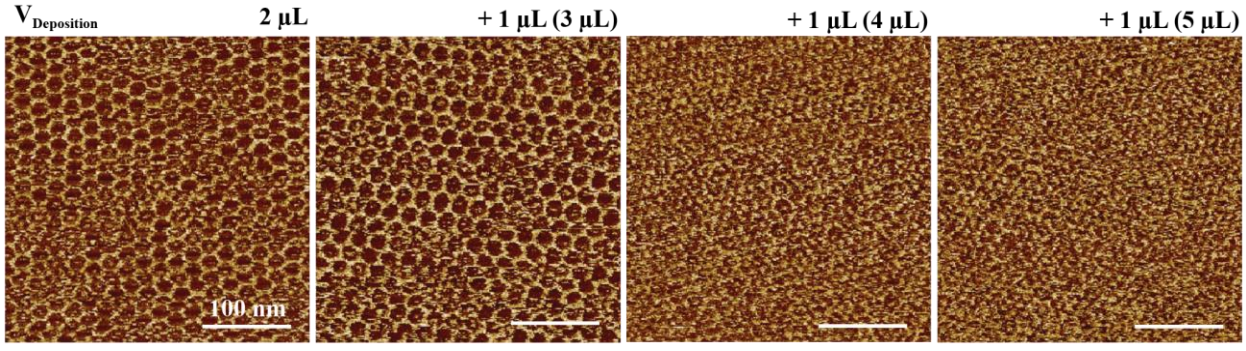


Figure 4.12| 3PS tile packing transitions with *in situ* tile addition on DPPC. 3PS **U** tile deposition on DPPC. V_{Dep} of 2 µL followed by sequential additions of 1 µL aliquots up to a total V_{Dep} of 5 µL. Tile stock concentration: 0.5 µM, scale bar = 100 nm.

The effect of temperature on 3PS tile binding and packing on DPPC was also investigated. **CHOL** tiles on DPPC were imaged with increasing temperature (from 25 to 50 °C) (**Figure 4.13**), examining the effect of lipid phase transition from gel to fluid state. Tiles that were initially organized into adjacent domains of mixed packing shifted to a more organized, denser packing state in the 35 – 45 °C temperature range, or around the transition temperature of DPPC (42 °C). It is possible that as lipid molecules in the bilayers become more mobile and the density of headgroups on the surface decreases, tiles bound to the surface are able to diffuse better to adopt a more regular, denser packing. With a further temperature increase to 50 °C, domains of organized tiles and individual tiles were no longer observed on the bilayer surface, either due to their high mobility on the fluid surface, dissociation from the surface or loss of hybridization within the tile structure (tile disassembly).

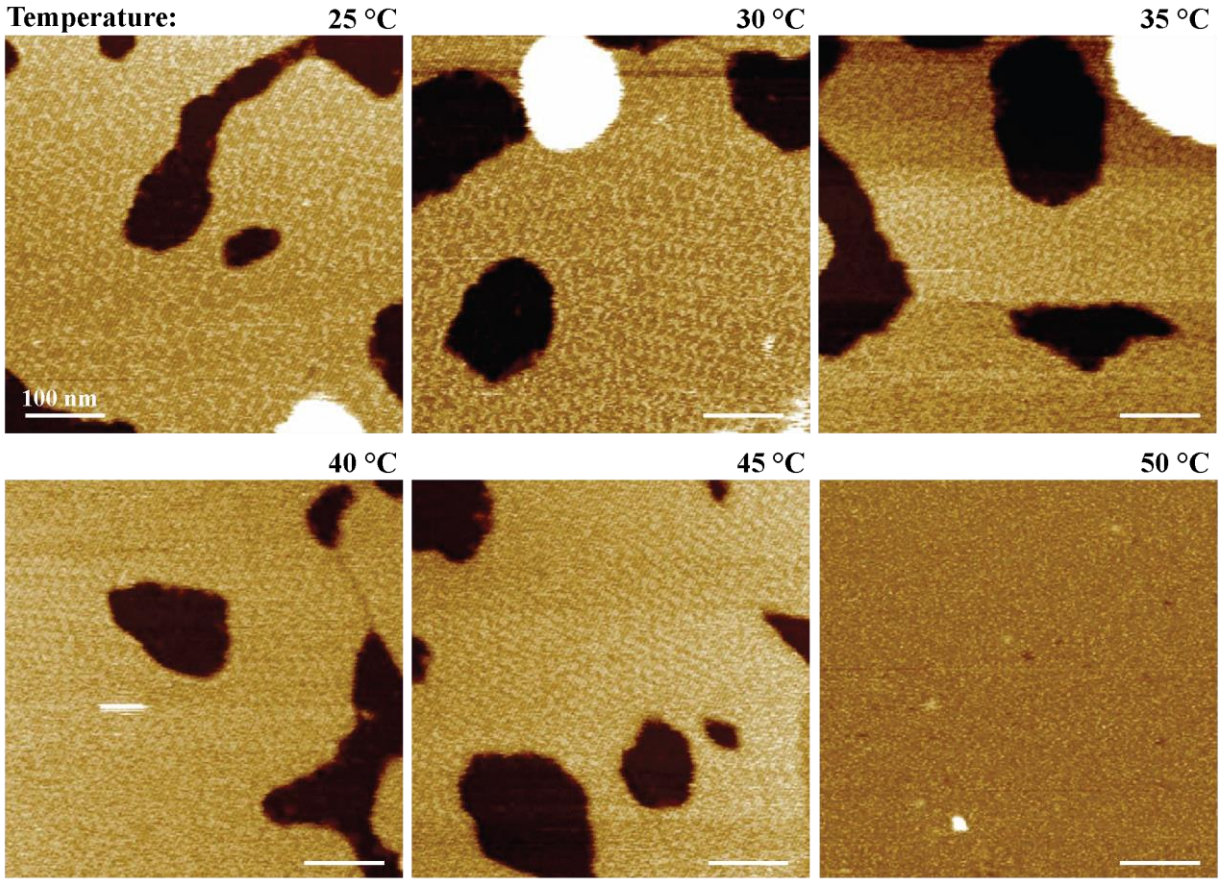


Figure 4.13| Temperature dependent imaging of 3PS tiles on DPPC. CHOL 3PS tiles (V_{Dep} = 10 µL) on DPPC imaged in a range of temperatures. Top row: 25, 30 and 35 °C sample temperatures. Bottom row: 40, 45 and 50 °C sample temperatures. Scale bar = 100 nm.

Furthermore, tile deposition on DPPC above the transition temperature was examined in comparison to deposition at room temperature. **U** tiles (V_{Dep} = 5 µL) were deposited on DPPC at 45 °C, followed by cooling to ambient conditions prior to imaging, effectively representing an annealing experiment. Qualitatively, more areas of dense, regular packing were observed in the annealed preparation (**Figure 4.14 a**), compared to tile deposition at ambient conditions (**Figure 4.14 b**), suggesting that annealing allows the system to maximize favorable electrostatic and blunt-end stacking interactions after initial binding to the fluid substrate.

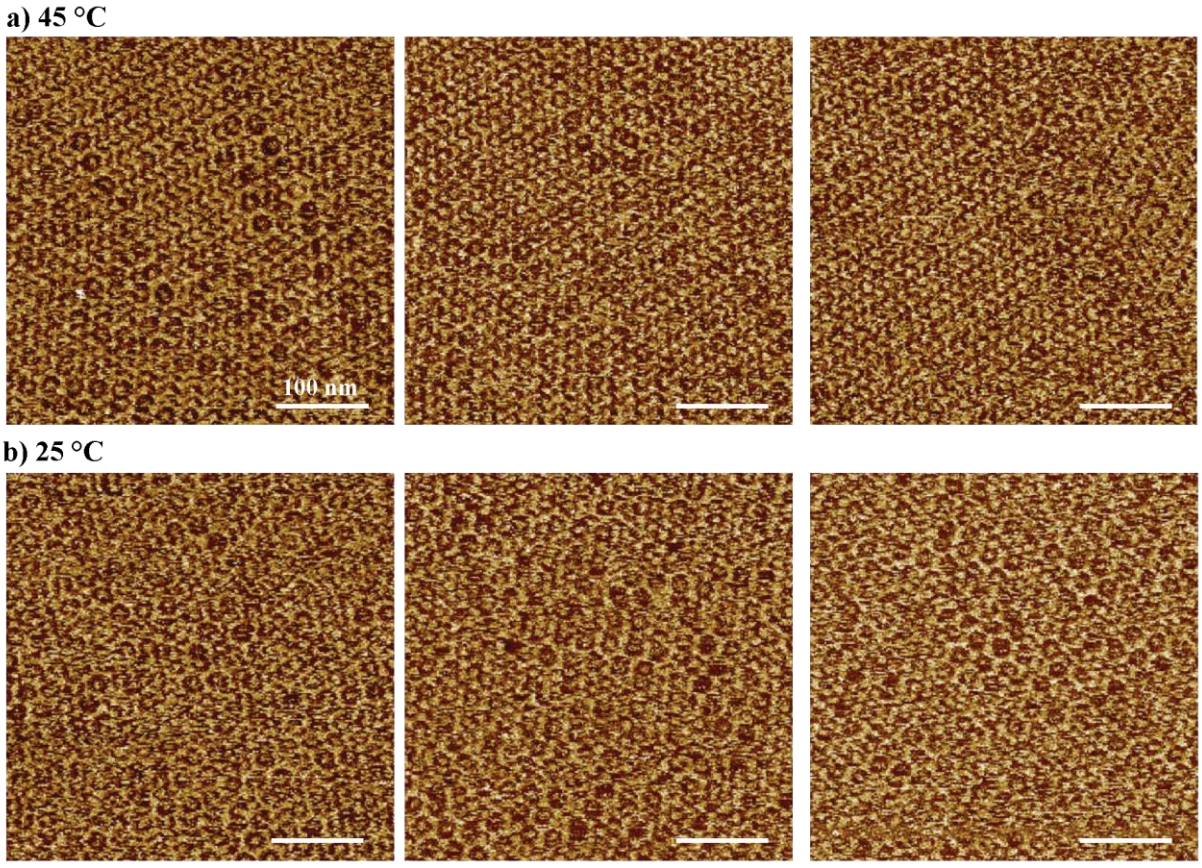


Figure 4.14| 3PS tile deposition above DPPC transition temperature. U 3PS tiles (V_{Dep} = 5 µL) deposited on DPPC at **a)** 45 °C and **b)** 25 °C. Imaging carried out at 25 °C. Scale bar = 100 nm.

4.2.5 Effect of blunt-end stacking disruption on long-range 3PS ordering

The long-range ordering of 3PS tiles observed above depends on either electrostatic or cholesteryl anchor-dependent interactions with the SLBs and blunt-end stacking between tile-arm

ends. The consequences of disrupting blunt-end stacking on tile ordering were probed by modifying tile-arm ends with non-interacting four-nucleotide T loops. Strand S3 was modified to incorporate four thymine (4T) residues at the end of each tile arm (S3Tloop strand) while maintaining full sequence complementarity with strands S1 and S2 otherwise (**Figure 4.15**). Assembly of 3PS **U** and **CHOL** tiles with T loop ends was verified by native PAGE (**Figure 4.16**).

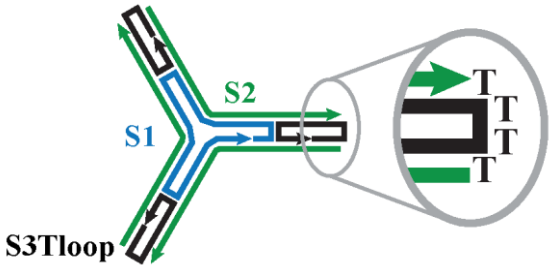


Figure 4.15| Schematic representation of the U 3PS tile with T loop ends. The tile is composed of strands S1, S2 and S3Tloop in a 1:3:3 stoichiometry.

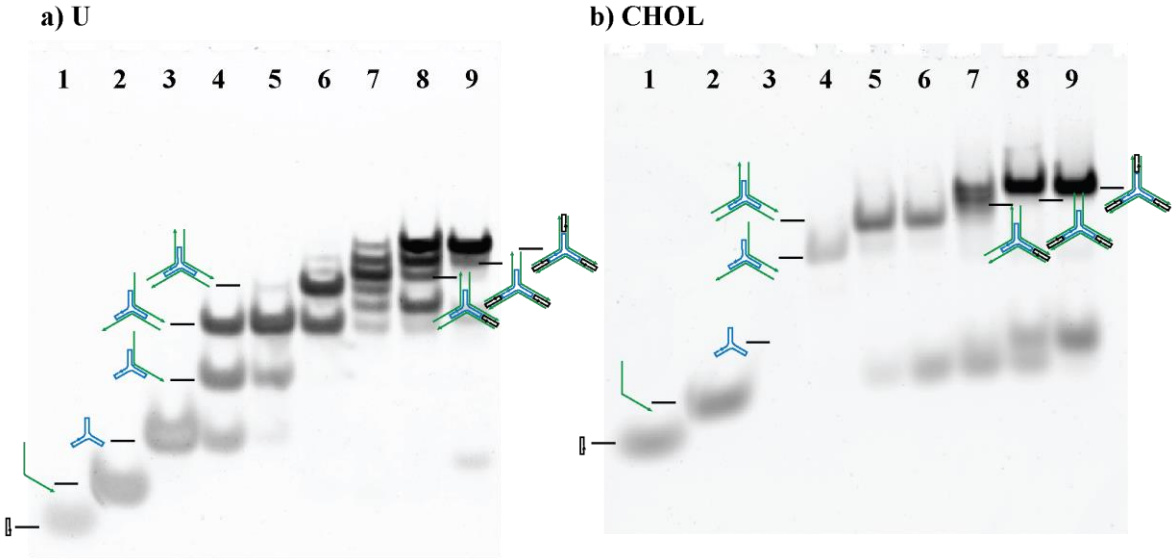


Figure 4.16| Assembly of 3PS tiles with T loop ends analyzed by native PAGE. a) U tile (composed of strands S1, S2 and S3Tloop). **b) CHOL** tile (composed of strands S1 CHOL, S2 and S3Tloop). Lane description for **(a)**: 1. S3Tloop, 2. S2, 3. S1, 4. S1+S2 (1:1), 5. S1+S2 (1:2), 6. S1+S2 (1:3), 7. S1+S2+S3Tloop (1:3:1), 8. S1+S2+S3Tloop (1:3:2), 9. S1+S2+S3Tloop (1:3:3). For **(b)**, same lane description but using the S1 CHOL strand. Native PAGE 6%, TAMg running buffer, 90 min at 120 V and 5 °C.

T loop **CHOL** tiles on DOPC did not show long range organization or even transient smaller assemblies (**Figure 4.17 a**). Although the bilayer surface did not appear smooth, suggesting tile binding to the surface, individual tiles with distinct 3PS shape could not be detected, possibly due to tile mobility on the fluid DOPC surface.

Both **U** and **CHOL** T loop tiles on DPPC (**Figure 4.17 b** and **c**) remained disorganized on the bilayer surface where, in contrast to DOPC, the 3PS shape of individual tiles could be clearly detected. Interestingly, even at high V_{Dep} (7.5 µL and 10 µL for **U** and **CHOL** tiles respectively), dense tile packing modes featuring tile organization into parallel lines and maximizing electrostatic tile binding to the DPPC surface were not detected. This observation suggests that the *Parallel* packing mode discussed above, where contacts between tile arms are not clearly visible, also relies on blunt-end stacking interactions. T loop 3PS tile deposition on DOPC and DPPC demonstrates that blunt-end stacking interactions are essential for long-range organization of these small tiles on the SLB surface.

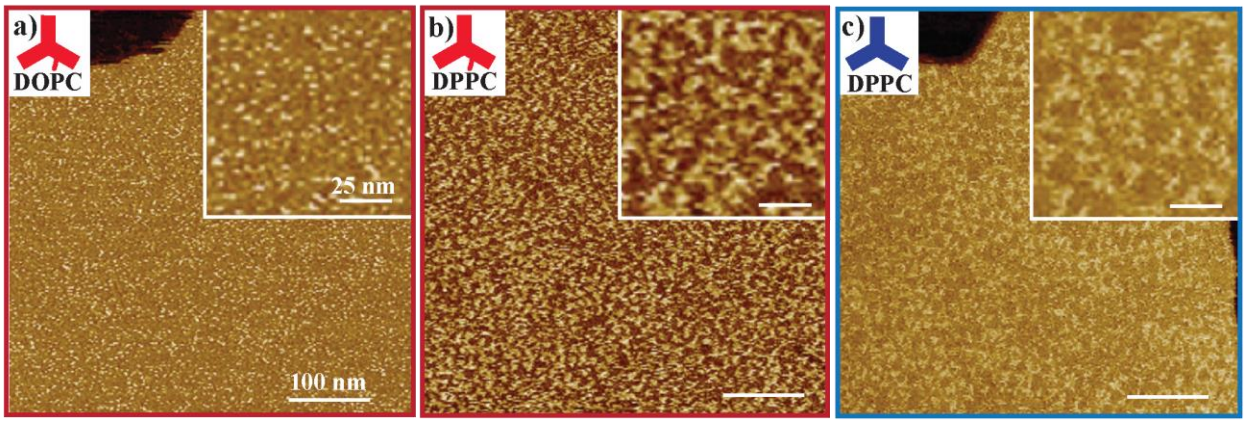


Figure 4.17| Consequences of disrupting blunt-end stacking interactions on long-range 3PS tile organization on SLBs. AFM images of 3PS tiles with T loop ends: **a) CHOL** tiles on DOPC (V_{Dep} = 2 µL). **b) CHOL** tiles on DPPC (V_{Dep} = 10 µL). **c) U** tiles on DPPC (V_{Dep} = 7.5 µL). Stock tile concentration of 0.5 µM, scale bar = 100 nm, inset scale bar = 25 nm.

4.2.6 3PS tile deposition on phase-separated mixed DPPC:DOPC bilayers

The observation of **U** and **CHOL** tile binding to phase-separated bilayers formed from a 70:30 DPPC:DOPC mixture provided further illustration of the dominant interactions driving tile organization on bilayer surfaces. DOPC and DPPC phase-separate due to structural differences in

their alkyl tails.[49-51] DPPC-rich and DOPC-rich domains within the phase separated bilayer can be easily identified by AFM imaging due to the consistent height difference of about 0.8 nm between the domains, where DPPC areas have a larger height than DOPC (**Figure 4.18**).

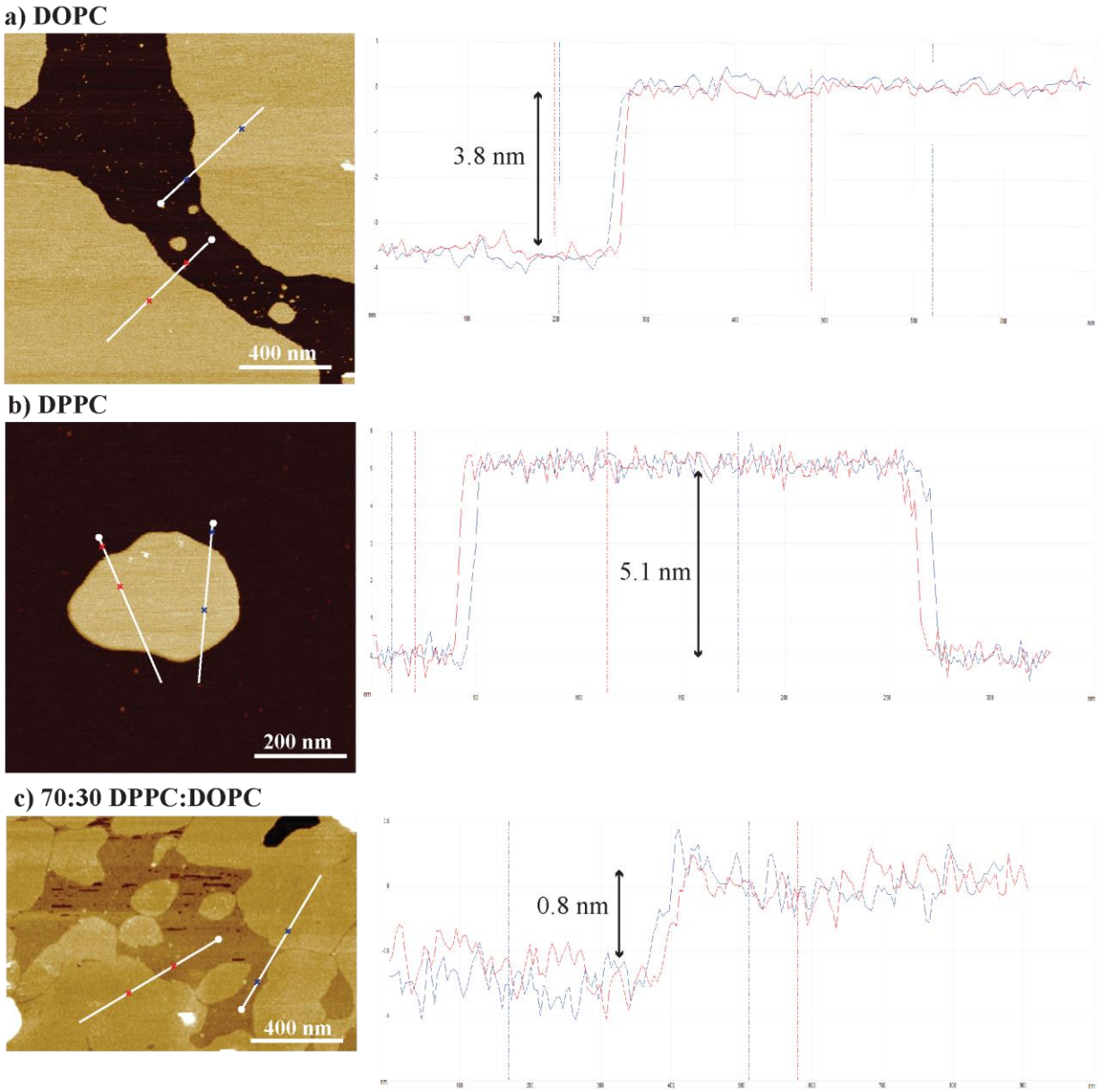


Figure 4.18| Lipid bilayer height analysis from AFM topography images. a) DOPC, **b)** DPPC and **c)** 70:30 DPPC:DOPC. For **(a)-(c)**, (left) AFM topography image and (right) topography trace and height measurements corresponding to the lines marked across the image. Scale bars as indicated on the images.

U tiles bound exclusively to DPPC domains (**Figure 4.19 a**), whereas **CHOL** tiles were found on both DOPC and DPPC-rich areas (**Figure 4.19 b**). **CHOL** tiles formed small areas of ordered honeycomb arrays on DOPC, while both **U** and **CHOL** tiles showed much denser packing on DPPC domains.

Furthermore, when a **1:1 U**:**CHOL** tile mixture was deposited on a mixed bilayer, large honeycomb arrays could be observed on DOPC. Considering the lack of binding of **U** tiles to DOPC, we assign these as **CHOL** tiles. In contrast, densely packed tiles were observed on DPPC domains (**Figure 4.20**), which is again consistent with more tile deposition on DPPC. Given the affinity of **U** and **CHOL** tiles to DPPC, these packed domains may be composed of both tile types. Thus, the mixture phase-separates into a DOPC-rich phase with **CHOL** tiles only, and a DPPC-rich phase, possibly with a mixture of **U** and **CHOL** tiles.

Collectively, the observations of **U**, **CHOL** or **1:1 U:CHOL** 3PS tile binding on mixed DPPC:DOPC bilayers suggest that the electrostatic interaction between DNA and the DPPC bilayer surface dominates in the deposition process, as the majority of both **U** and **CHOL** tiles are found to be densely packed on DPPC domains. Similarly, the preferential partitioning of cholesteryl-TEG and unmodified DNA origami rods to liquid ordered lipid domains in the presence of Mg^{2+} ions has been reported.[42] However, 3PS tile structures have a much smaller contact area with the bilayer, as previously reported origami rods measured approximately 420 nm in length in contrast to 3PS tiles where each arm is only 6.6 nm long.

Another remarkable observation regarding blunt-end interactions can be made from the seemingly continuous contacts between tiles located at the edges of lipid domains. For example, uninterrupted networks of **U** tiles are detected between adjacent DPPC domains on the mixed bilayer (white arrows, **Figure 4.19 a**). In the case of **CHOL** tiles, it appears that tiles located at the edge of a DPPC domain can make bridges with adjacent tiles at the edge of the surrounding DOPC (white arrows, **Figure 4.19 b**), thus appearing to "stitch them" together.

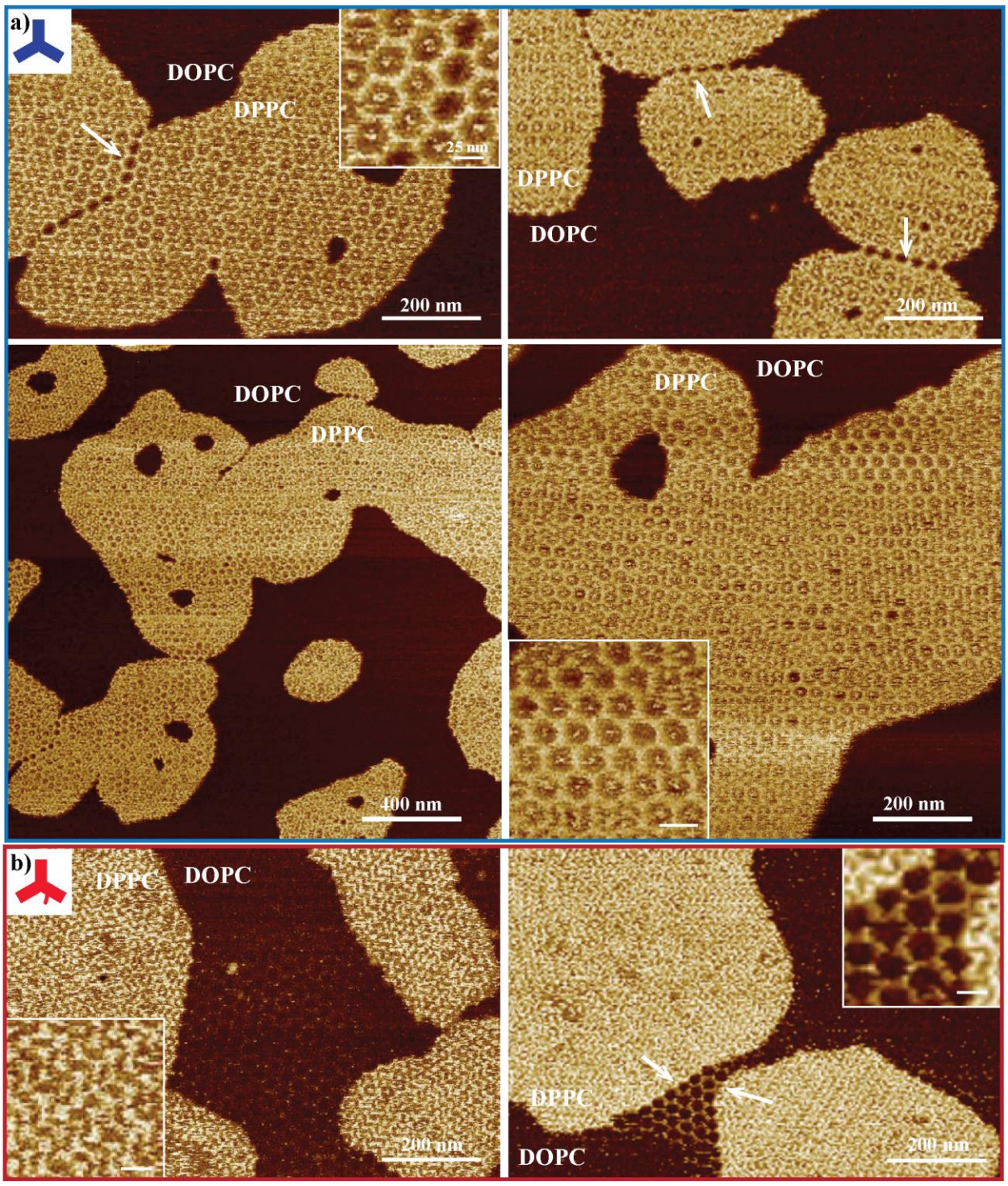


Figure 4.19| Deposition of 3PS tiles on 70:30 DPPC:DOPC mixed bilayers. a) U tiles, **b) CHOL** tiles. DOPC and DPPC regions as indicated on the images. White arrows show areas of blunt-end contacts between tiles across boundaries of adjacent lipid domains. Stock tile concentration of 0.5 μM, V_{Dep} = 2 μL, scale bars as indicated on the images (inset scale bar = 25 nm).

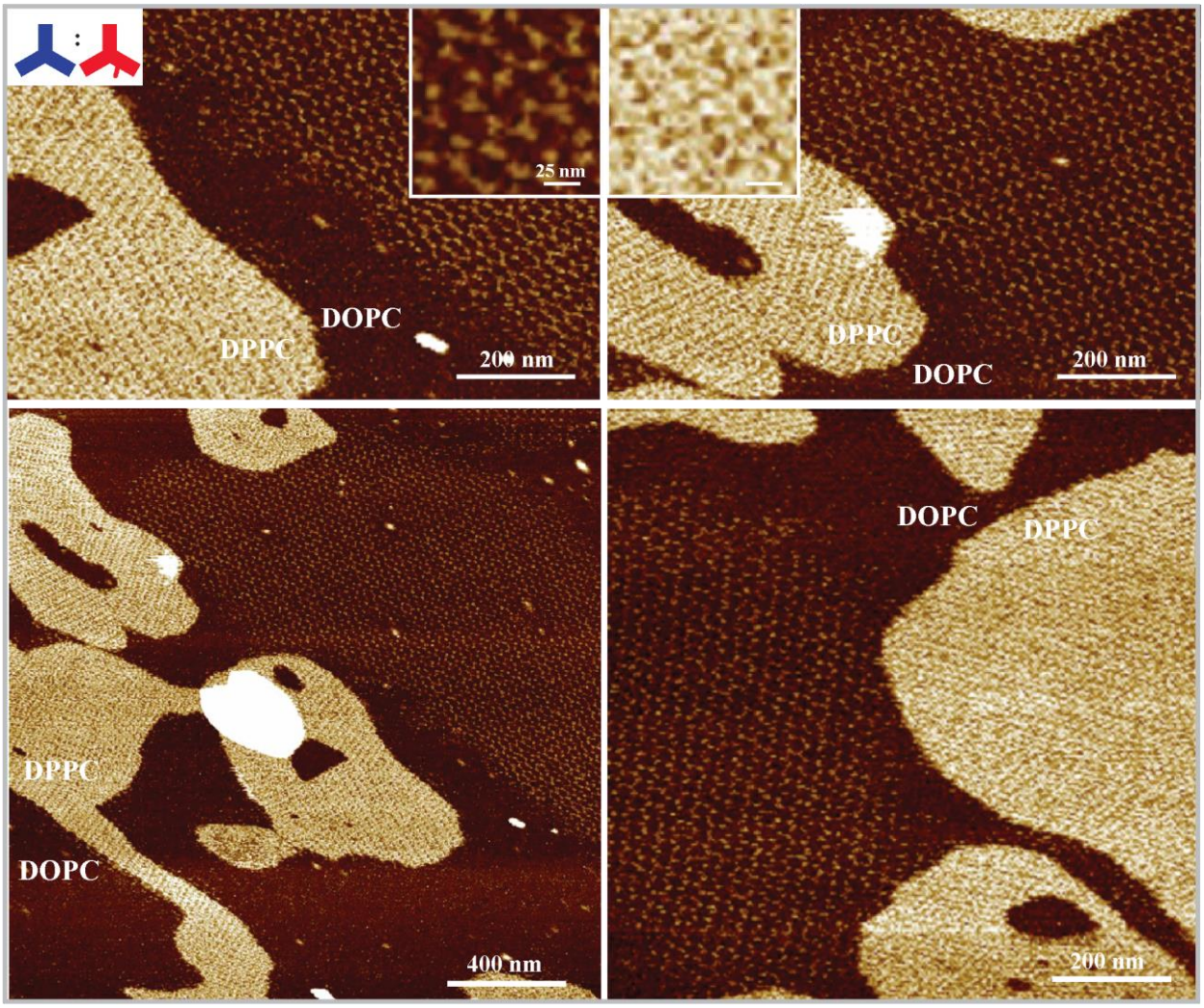


Figure 4.20| 1:1 U:CHOL 3PS tile mixture deposition on 70:30 DPPC:DOPC mixed bilayers. DOPC and DPPC regions as indicated on the images. Stock tile concentration of 0.5 μM, V_{Dep} = 2 μL, scale bars as indicated on the images (inset scale bar = 25 nm).

4.2.7 Cross (X) tile deposition on DOPC and DPPC

The deposition of a second type of DNA tile on SLBs was also investigated. The cross (X) tile, first reported by Yan *et al.*,[4] resembles the 3PS tile in crossover design principle, but instead has four arms oriented at 90° angles and uses unique sequences instead of sequence symmetry (**Figure 4.21 a**). X tiles were prepared in **U** and **CHOL** versions using the same type of modification as for 3PS tiles. Strand characterization and tile assembly are shown in **Figure 4.21 b, Figure 4.22** and **Figure 4.23**).

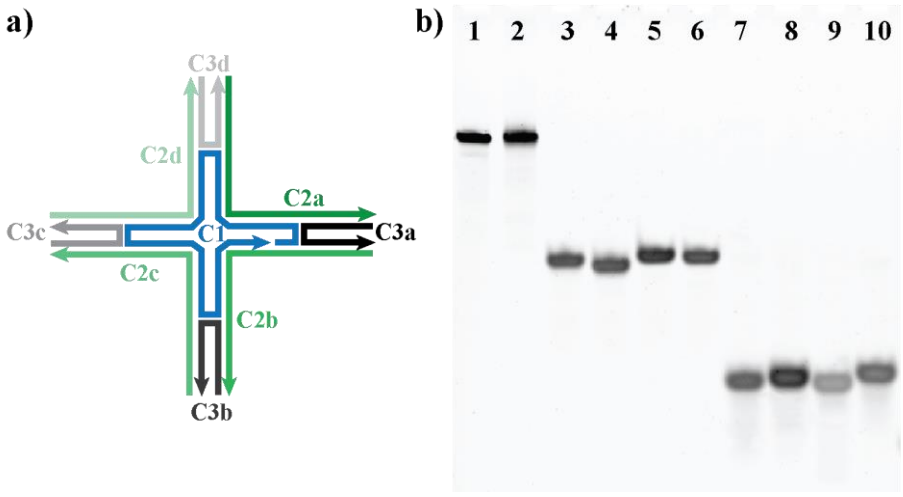


Figure 4.21| X tile structure and strand characterization. a) Schematic representation of the unmodified (**U**) blunt-ended X tile. CHOL tiles are prepared with strand C1 CHOL. **b)** Denaturing PAGE (12%, 1xTBE, 30 min at 250 V followed by 45 min at 500 V) of X tile component strands. Lane description: 1. C1, 2. C1 CHOL, 3. C2a, 4. C2b, 5. C2c, 6. C2d, 7. C3a, 8. C3b, 9. C3c and 10. C3d.

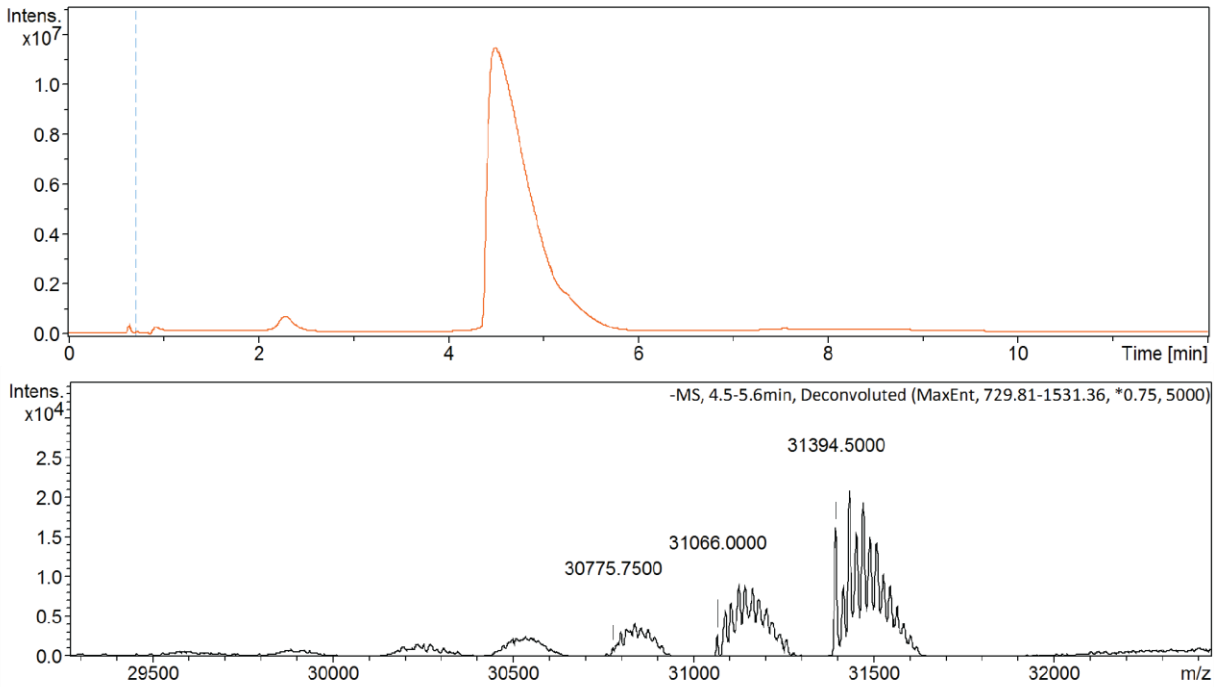


Figure 4.22| LC-ESI-MS characterization of cholesterol modified strand C1 CHOL. LC trace detected by absorbance at 260 nm (top) and mass spectrum (bottom) for strand C1 CHOL (calculated M_W = 31 395 g/mol).

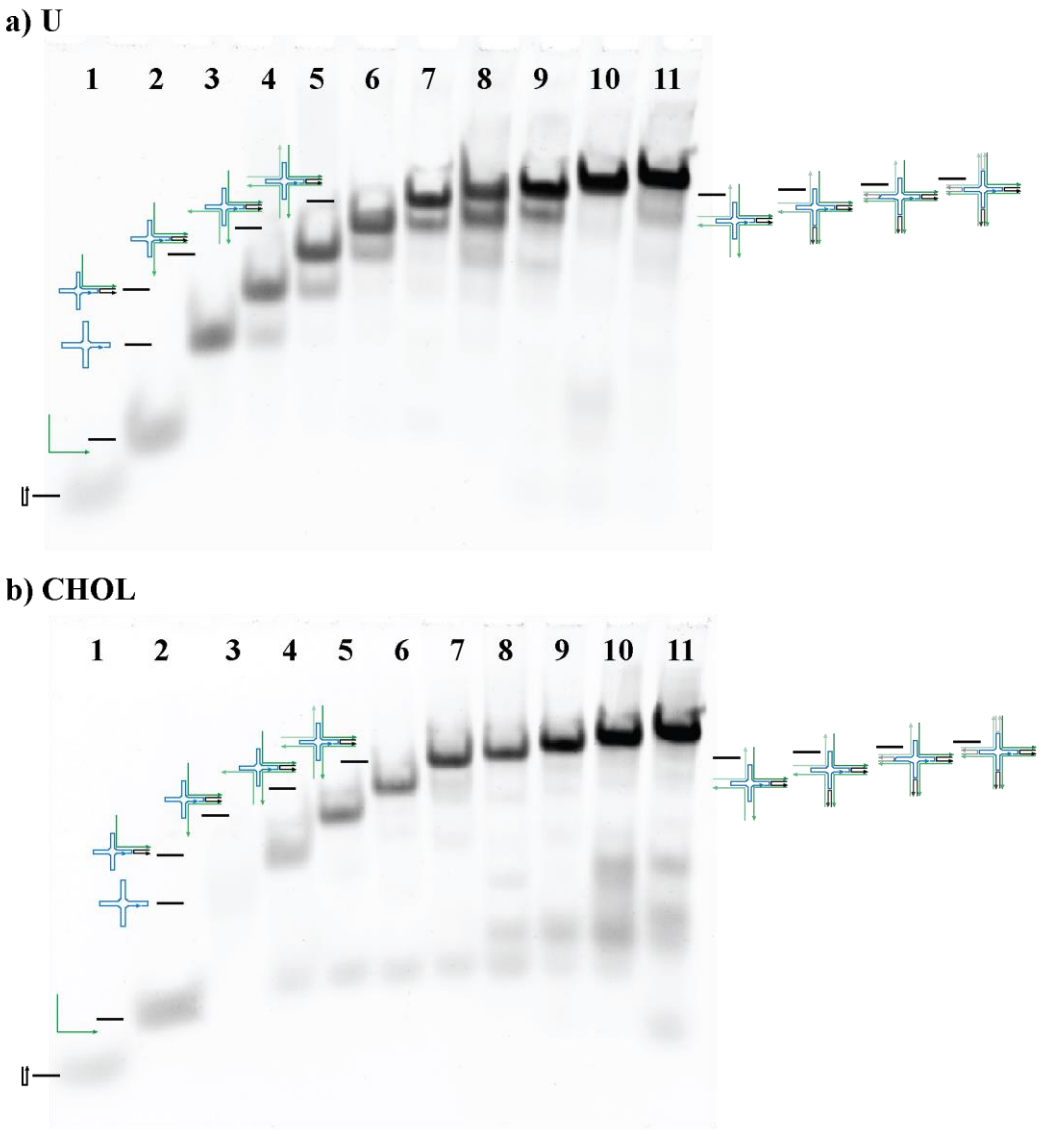


Figure 4.23| Assembly of X tiles analyzed by native PAGE. a) U X tile. **b) CHOL** X tile. Lane description for **(a)**: 1. C3a, 2. C2a, 3. C1, 4. C1+C2a, 5. C1+C2a+C2b, 6. C1+C2a+C2b+C2c, 7. C1+C2a+C2b+C2c+C2d, 8. C1+C2a+C2b+C2c+C2d+C3a, 9. C1+C2a+C2b+C2c+C2d+C3a+C3b, 10. C1+C2a+C2b+C2c+C2d+C3a+C3b+C3c, 11. C1+C2a+C2b+C2c+C2d+C3a+C3b+C3c+C3d. For **(b)**, same lane description but using C1 CHOL. Native PAGE 6%, TAMg running buffer, 90 min at 120 V and 5 °C.

Another distinctive feature of the X tile is the potential for greater flexibility of the tile arms due to the presence of 4 thymidine (4T) spacers in the central strand between the tile arms (**Figure 4.24 a**), in contrast to the 3T spacers in 3PS tiles used above. Variations on the 3PS tile

where longer T-spacers provide flexibility have been used to build closed 3D structures from 3PS tiles, however tiles with 3T spacers are too rigid to produce the required curvature.[38]

Increased X tile flexibility gives rise to variations in the angle between tile arms such that a range of lattice geometries are possible upon X tile assembly through blunt end interactions. The possible geometries include square (90° angles between tile arms), parallelogram (> 90°) and even hexagram (120°) lattices (**Figure 4.24 b**). The two extremes of this conformational range, square and hexagram lattices, were observed with X tile deposition on mica. Tiles assembled into small domains of both geometries (**Figure 4.24 c**).

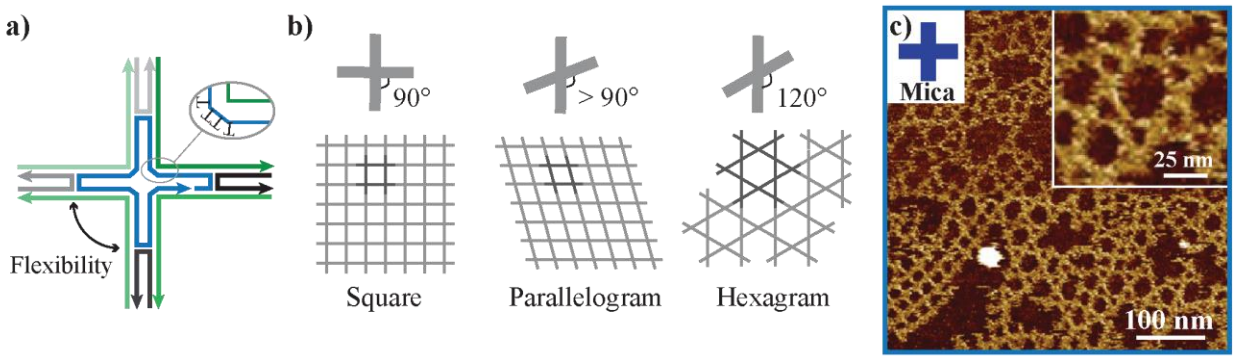


Figure 4.24| Cross (X) tile assemblies. a) Schematic representation of the X tile. The central strand contains 4T spacers between the arms giving the arms enough flexibility to adopt orientations with different angles. **b)** Depending on the angle between tile arms, the X tile can form lattices with different geometries. Square, parallelogram and hexagram lattices are illustrated (inter-arm angles of 90 °, > 90 ° or 120 ° respectively give rise to such packing). **c) U** X tile on mica. Domains of square and hexagram lattices are observed. Inset shows detail of hexagram repeat unit. Stock tile concentration of 0.5 µM, V_{Dep} = 2 µL, scale bar = 100 nm (inset scale bar = 25 nm).

On SLBs, the divergent binding characteristics observed with **U** and **CHOL** 3PS tiles were replicated with X tiles. **CHOL** X tiles bound to DOPC formed extensive parallelogram (distorted square) lattices (**Figure 4.25 a**). On the other hand, **U** tiles were found exclusively on mica (**Figure 4.25 b**), forming only small areas of organized material in contrast to the micrometer-scale organization of **CHOL** tiles on DOPC. The extreme angle distortion giving rise to hexagram lattices was not observed on SLBs.

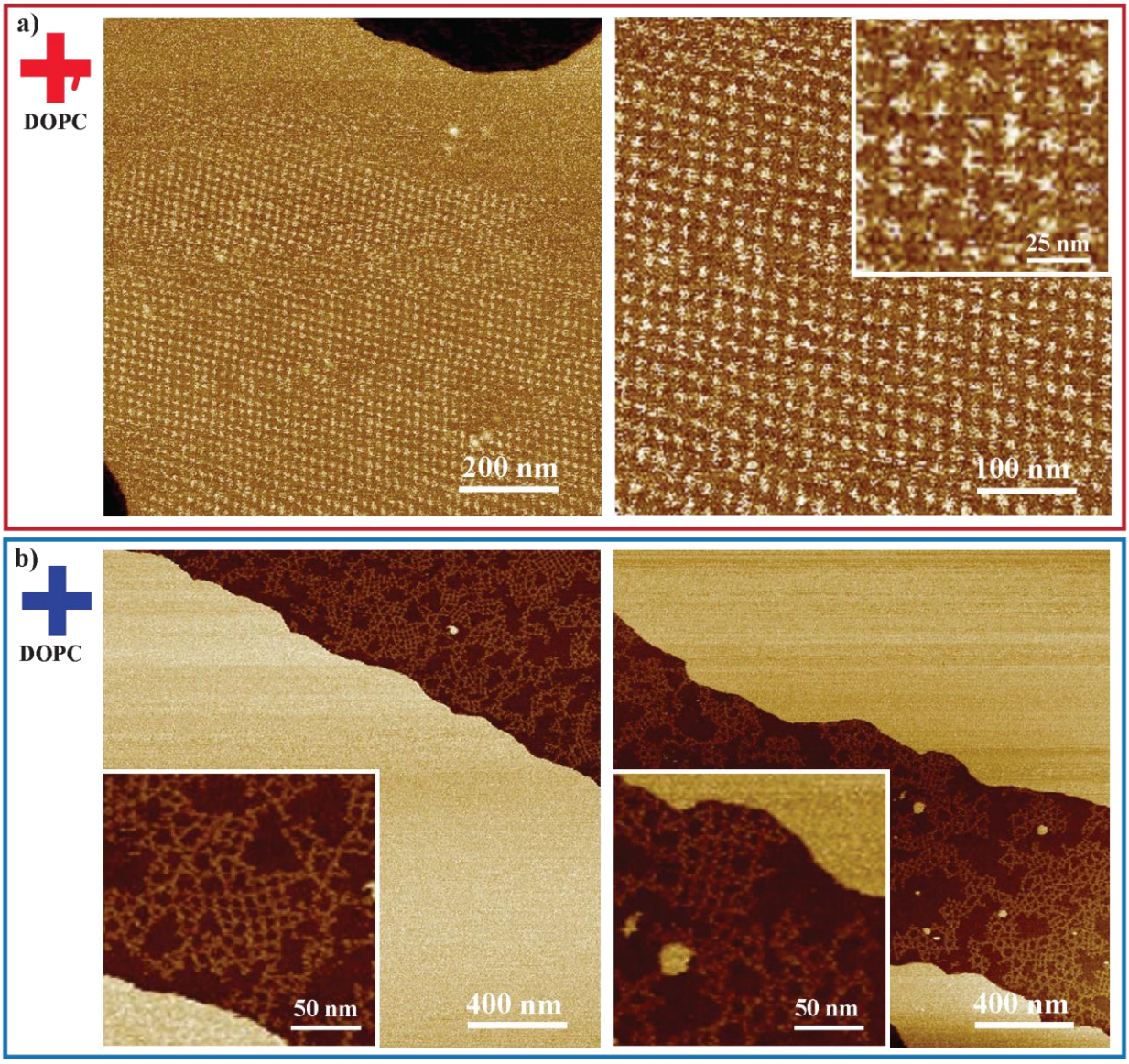


Figure 4.25| X tiles on DOPC. AFM images of: **a) CHOL** tiles and **b) U** tiles. V_{Dep} = 2 µL, scale bars as indicated on the images.

Furthermore, **CHOL** and **U** X tiles on DPPC showed V_{Dep} dependent tile packing. For **CHOL** tiles, a square lattice dominated at lower V_{Dep}, shifting to a dense packing where tile angles were highly distorted at high V_{Dep}, (**Figure 4.26 a**). **U** tiles showed angle distortion both at low and high V_{Dep}, with tile density on the surface increasing with V_{Dep} (**Figure 4.26 b**). A distinctive pattern composed of side-by-side parallel lines was observed for V_{Dep} = 10 µL, where the blunt-end connections between tile arms were difficult to resolve.

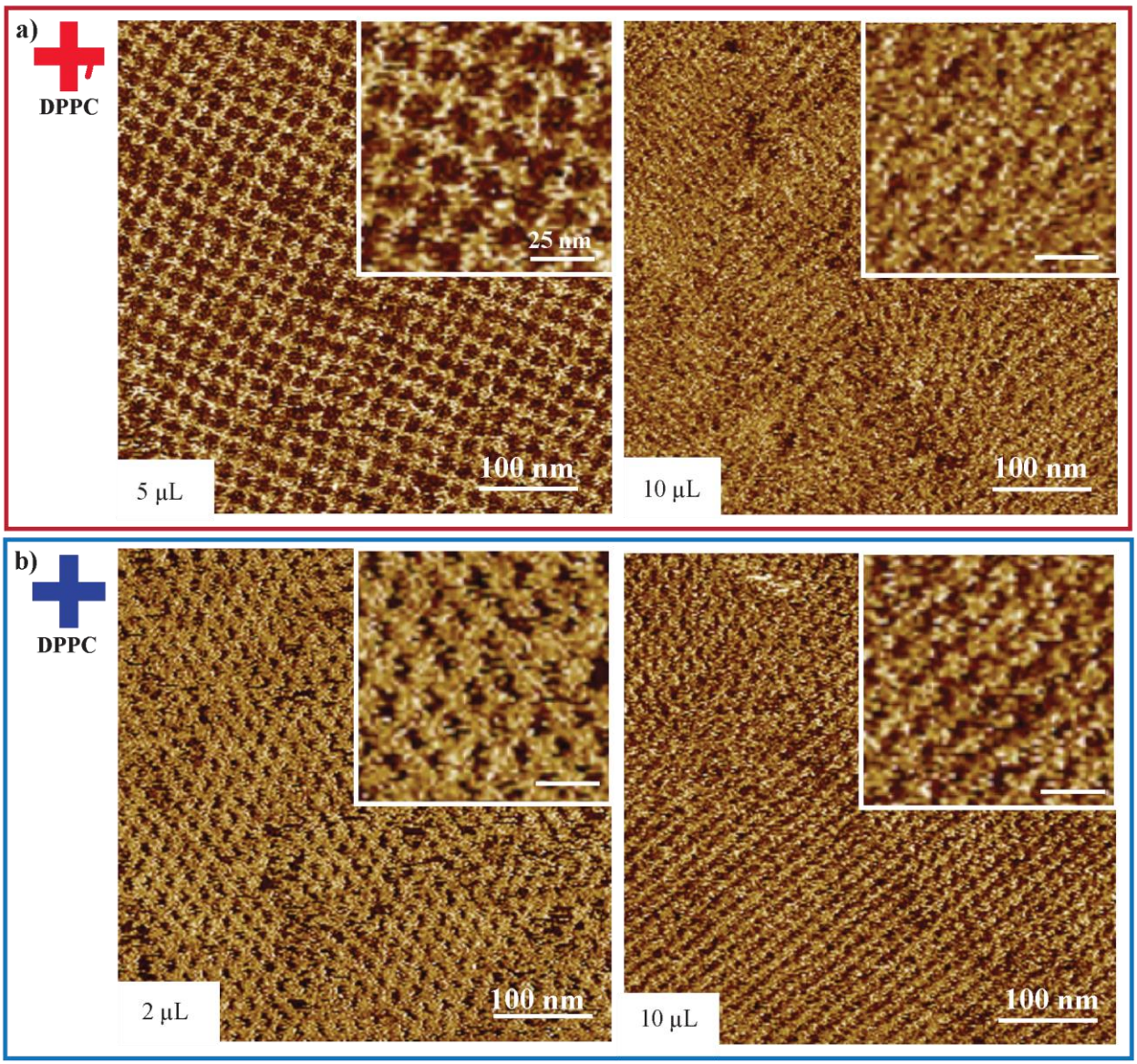


Figure 4.26| X tiles on DPPC. AFM images of: **a) CHOL** tiles (V_{Dep} = 5 µL – left, 10 µL – right) and **b) U** tiles (V_{Dep} = 2 µL – left, 10 µL – right). Scale bar = 100 nm (inset scale bar = 25 nm).

4.3 Conclusions

In this chapter, we have examined the long-range ordering of blunt-ended three-arm and four-arm crossover DNA tiles on fluid and gel-phase zwitterionic SLBs. A cholesterol anchor was required for lattice formation on the DOPC fluid bilayer, whereas both unmodified and cholesterol-anchored tiles showed extensive ordering on the DPPC gel-phase lipid. This divergent deposition behavior was attributed to the greater strength of electrostatic interactions between DNA and

zwitterionic lipid headgroups in the gel-phase bilayer, mediated by Mg^{2+} cations. This property also permitted selective patterning of different domains in a phase-separated lipid system. Furthermore, tile deposition on gel-phase SLBs was accompanied by a concentration-dependent reconfiguration of tile-packing morphology on the surface, a characteristic that could be employed for layer-by-layer deposition of tiles with various modifications to achieve a greater surface patterning complexity by simple means.

Of paramount importance to the variety of morphologies observed, as well as the long-range scale of organized assemblies, was the use of base stacking interactions between blunt tile-arm ends, rather than more traditional hybridization of sticky ends. The dynamic nature of the blunt-end stacking contacts allowed error correction to occur, giving rise to a defect-free tile network on the fluid bilayer. In addition, it facilitated the formation of dense tile packing modes where some blunt end interactions were disrupted in favor of maximizing electrostatic binding. The advantage of using lipid bilayers as a support for long-range organization of DNA is first in the rich variety of composition and phase properties that can be harnessed in combination with compatible anchoring moieties. Furthermore, the fluidity, composition and temperature dependent properties of lipid bilayers allow for surface diffusion of DNA building blocks, a property which promotes error correction and morphological transitions that may have desirable applications. Supported lipid bilayers represent a gateway for interfacing DNA nanostructures with a biologically relevant environment with many possible applications as well.

The principles on display, namely the balance of inter-tile base-stacking interactions with tile-bilayer interactions, although simple, are of fundamental interest and are broadly applicable. The resulting organized materials are composed of just a few inexpensive strands. They can find applications in a number of areas, such as patterned surfaces for selective cell recognition, tools for creating biological models, templates for organizing light-harvesting units, proteins and nanoparticles, as well as long-range ordered, fluid surfaces to interface DNA nanostructures with solid substrates to develop advanced materials with hierarchical organization.

4.4 Experimental

4.4.1 Materials

Tris(hydroxymethyl)aminomethane (Tris), magnesium acetate tetrahydrate ($Mg(OAc)_2 \cdot 4 H_2O$), sodium chloride (NaCl), glacial acetic acid, urea, ethylenediaminetetraacetic acid (EDTA), ammonium hydroxide, trimethylamine and HPLC grade acetonitrile (MeCN) were used as purchased from Sigma-Aldrich. Boric acid and GelRed were obtained from Fisher Scientific and used as supplied. Acrylamide/bis-acrylamide (40% 19:1) solution, ammonium persulfate and tetramethylethylenediamine (TEMED) were used as purchased from BioShop Canada Inc. Sephadex G-25 (super fine, DNA grade) was purchased from Glen Research. 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) and 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) were purchased from Avanti Polar Lipids as solutions in chloroform.

The reagents and materials listed below were used for DNA synthesis. Anhydrous MeCN, dA-CE, dC-CE, dmf-dG-CE and dT-CE phophoramidites, universal 1000 Å LCAA CPG pre-packed columns (loading densities of 25-40 μmol/g), activator (0.25 M ETT in acetonitrile, oxidizer (0.02 M I_2 in THF/pyridine/H_2O), Cap A mix (THF/lutidine/Ac_2O), Cap B mix (16% MeIm in THF), and deblock solution (3% DCA/DCM) were all used as purchased from BioAutomation. 3'-Cholesteryl-TEG CPG was purchased from Glen research.

Muscovite Ruby mica sheets (grade 2) were used as substrate for all AFM imaging studies. 1xTBE (Tris-boric acid-EDTA) buffer was composed of 45 mM Tris, 45 mM boric acid and 2 mM EDTA at pH 8.3. 1xTAMg buffer was composed of 40 mM Tris, 12.5 mM $Mg(OAc)_2$ with pH adjusted to 8.0 with glacial acetic acid. Tris/NaCl/$CaCl_2$ buffer contained 10 mM Tris, 100 mM NaCl and 3 mM $CaCl_2$ with pH adjusted to 7.4 with HCl. Tris/NaCl buffer contained 10 mM Tris and 100 mM NaCl (pH 7.4). All buffers were prepared with Milli-Q water and samples were prepared with autoclaved Milli-Q water.

4.4.2 Instrumentation

UV-Vis DNA quantification measurements were performed on a Nanodrop Lite spectrophotometer from Thermo Scientific. A Mermade MM6 DNA synthesizer from

BioAutomation was used for DNA synthesis. Gel electrophoresis experiments were carried out on a 20 x 20 cm vertical acrylamide Hoefer 600 and on 8.3 x 7.3 cm vertical acrylamide Mini-PROTEAN electrophoresis units (BioRad). Reverse phase HPLC (RP-HPLC) purification was conducted using a Hamilton PRP-1 column (150 mm x 4.1 mm, 100 μm particle size) on an Agilent Infinity 1260 system. Liquid chromatography electrospray ioization mass spectrometry (LC-ESI-MS) was performed on a Dionex Ultimate 3000 system coupled to a Bruker MaXis Impact QTOF. Thermal annealing procedures were carried out using an Eppendorf Mastercycler Pro 96 well thermocycler. Electrophoresis gels were imaged on a Chemidoc MP Imaging System from BioRad. A VCX 130 Vibra-Cell ultrasonic liquid processor equipped with a 3-mm stepped microtip (Sonics & Materials Inc.) was used for sonication of lipid suspensions.

AFM imaging was carried out on a MultiMode 8 microscope with a Nanoscope V controller (Bruker). An air-fluid heating element (Bio-Heater) connected to a thermal applications controller (Bruker) were used for temperature dependent AFM imaging. The temperature within the fluid cell was measured with a fine gage thermocouple (Omega) connected to a handheld digital thermometer (Oakton).

4.4.3 Oligonucleotide synthesis and characterization

Oligonucleotide synthesis was performed in house on a 1 μmol scale, using standard, automated, phosphoramidite solid-phase DNA synthesis. The cholesteryl-TEG modification in strands S1 CHOL and C1 CHOL was introduced at the 3' end of strands S1 and C1 by using a 3'-Cholesteryl-TEG modified CPG column. In all syntheses, coupling efficiency was monitored after removal of the DMT 5'-OH protecting groups. Sequences for strands composing three-point star (3PS) and cross (X) tiles are compiled in **Table 4.1** and **Table 4.2** respectively.

Completed sequences were deprotected in a 28% ammonium hydroxide solution for 16 hours in a 65 °C water bath, purified by PAGE under denaturing conditions (12, 15 or 20% polyacrylamide with 8 M urea, the percentage depending on strand length, using 1xTBE running buffer) at constant voltage (30 mA, at 250 V for the first 30 minutes, followed by 1 hour at 500 V). Following electrophoresis, the bands of purified oligonucleotides were excised and recovered from the gel (crushed, covered with water, vigorously shaken, flash frozen with liquid nitrogen and kept in a 65 °C water bath for 16 hours). The solutions were then dried down to 1 mL and

desalted using size exclusion chromatography (Sephadex G-25). Strand purity was assessed by denaturing PAGE analysis.

Strand	Sequence (5' - 3')
S1	AGGCACCATCGTAGG TTT CTTGCCAGGCACCATCGTAGG TTT CTTGCCAGGCACCATCGTAGG TTT CTTGCC
S1 CHOL	AGGCACCATCGTAGG TTT CTTGCCAGGCACCATCGTAGG TTT CTTGCCAGGCACCATCGTAGG TTT CTTGCC - **TEG-CHOL**
S2	ACTATGCAACC TGCCTGGCAAG CCTACGATGG ACACGGTAACG
S3	CGTTACCGTGT GGTTGCATAGT
S3Tloop	CATAGT TTTT CGTTACCGTGT GGTTG

Table 4.1| Sequences for strands composing 3PS tiles. The strand sequences were designed by the Mao group.[8] Non-nucleoside cholesterol (**TEG-CHOL**) modification was appended to strand S1 as indicated.

Strand	Sequence (5' - 3')
C1	AGGCACCATCGTAGG TTTT CGTTCCGATCACCAACGGAGT TTTT TCTGCCGTACACCAGTGAAGT TTTT CGATCCTAGCACCTCTGGAGT TTTT CTTGCC
C1 CHOL	AGGCACCATCGTAGG TTTT CGTTCCGATCACCAACGGAGT TTTT TCTGCCGTACACCAGTGAAGT TTTT CGATCCTAGCACCTCTGGAGT TTTT CTTGCC - **TEG-CHOL**
C2a	GAGCGCAACC TGCCTGGCAAG ACTCCAGAGG ACTACTCATCC
C2b	GACTGAGCCC TGCTAGGATCG ACTTCACTGG ACCGTTCTACC
C2c	GTTGGCTTCC TGTACGGCAGA ACTCCGTTGG ACGAACACTCC
C2d	GGATAGCGCC TGATCGGAACG CCTACGATGG ACACGCCGACC
C3a	GGTCGGCGTGT GGTTGCGCTC
C3b	GGATGAGTAGT GGGCTCAGTC
C3c	GGTAGAACGGT GGAAGCCAAC
C3d	GGAGTGTTCGT GGCGCTATCC

Table 4.2| Sequences for strands composing X tiles. The strand sequences were designed by the LaBean group.[4] Non-nucleoside cholesterol (**TEG-CHOL**) modification was appended to strand C1 as indicated.

The cholesterol modified strands S1 CHOL and C1 CHOL were further purified by RP-HPLC (50 mM triethylammonium acetate (TEAA, pH 8.0) and HPLC grade acetonitrile (MeCN) with elution gradient of 3-50% acetonitrile over 30 minutes at 60 °C), as the TEG-CHOL moiety can undergo cleavage such that the modification is lost. Following RP-HPLC purification, the relevant fractions were dried down and resuspended in water. The presence of the cholesterol modification was verified by LC-ESI-MS analysis in negative mode.

4.4.4 3PS and X tile assembly and characterization

Tile preparation for AFM studies was carried out in 1xTAMg buffer at a final tile concentration of 0.5 μM. For 3PS tiles, different variations of S1, S2 and S3 strands were combined at a molar ratio of 1:3:3. For X tiles, all component strands were combined in equimolar ratio. All tiles were subjected to an annealing protocol from 95ºC to 4ºC over 6 hours.

All tile constructs were characterized by native PAGE (6%, 1xTAMg buffer, 250 V for 90 minutes) to confirm correct assembly. Tile assembly was carried out at 0.1 μM for PAGE studies and 0.5 μM for AFM studies.

4.4.5 Preparation of supported lipid bilayers (SLBs)

DOPC, DPPC and 70:30 DPPC:DOPC SLBs were prepared by vesicle fusion, adapting the protocol from Mingeot-Leclercq *et al*.[52] Lipid stock in chloroform was placed in a dry glass vial, the solvent was removed by rotary evaporation and the dry lipid film was maintained under vacuum overnight. The lipid film was resuspended in Tris/NaCl/$CaCl_2$ buffer at a 1 mM total lipid concentration. In the case of DPPC, heating at 60 °C for 2 minutes was supplied to facilitate the hydration of the film. The suspension was vigorously vortexed for 5 minutes prior to sonication to obtain small unilamellar vesicles (SUVs). Sonication with a probe sonicator was carried out over ice to achieve solution clarity (130 W, 1 minute cycles with ON 20 s and OFF 5 s steps, repeated 5 to 10 times). SUV suspensions were filtered on 0.2 μm nylon filters to eliminate any metal particles from the probe sonicator.

SLBs were prepared by depositing 120 μL of SUV solution on freshly cleaved mica (15

mm diameter disc secured on a metal sample puck with epoxy adhesive). In the case of DOPC and 70:30 DPPC:DOPC mixed lipid systems, samples were incubated for 60 minutes at ambient temperature. For DPPC, samples were heated at 58 °C for 60 minutes with 120 µL additions of Tris/NaCl buffer made every 10 minutes to prevent surface dewetting. The sample was then cooled to room temperature over 10 minutes. For all SLBs, the subsequent preparation step consisted of rinsing with Tris/NaCl solution (120 µL, 2 rinses) followed by rinsing with 120 µL of 1xTAMg. Upon addition of 20 µL of 1xTAMg to the surface, the samples were ready for tile deposition.

4.4.6 Tile deposition on SLBs and AFM imaging procedure

A given volume of 0.5 µM tile solution (deposition volume V_{Dep} from 1 µL to 10 µL) was added to the surface and the sample was immediately placed on the AFM sample stage. A fluid cell was then positioned on the sample (MTFML fluid cell, Bruker) and 70 µL of 1xTAMg was injected into the cell.

The same sample deposition method was followed for tile imaging on bare mica.

Topography images were acquired at ambient conditions in ScanAsyst mode. Silicon nitride levers with a nominal spring constant of 0.1 N/m, resonant frequency of 38 kHz and a tip radius of 2 nm were used (MSNL-E, Bruker). The cantilevers were calibrated in solution using the thermal tune method. Images were captured at scan rates between 0.5 and 1.5 Hz at a resolution of 512 x 512 pixels. The z-limit was set at 1 µm and the peak force amplitude at 50 to 75 nm. The peak force setpoint was set automatically by the software and varied from 150 pN to 300 pN. Images were processed using Nanoscope Analysis 1.5 software (Bruker).

4.4.7 3PS tile packing transitions on DPPC with *in situ* tile additions

Transitions in 3PS tile packing on DPPC were studied by imaging the sample following *in situ* tile additions upon initial tile deposition on the surface. The DPPC SLB was prepared following the procedure described in section 4.4.5. An initial 2 µL aliquot of 0.5 µM **U** tile solution was added to the surface and the sample was imaged according to the procedure described in section 4.4.6. Sequential additions of 1 µL of 0.5 µM **U** tile solution in 10 µL of 1xTAMg buffer

to the sample cell were followed by a 5 minute equilibration period and imaging. The same region was imaged after each addition; however, some drifting of the sampled area did occur.

4.4.8 Temperature dependent studies of 3PS tile deposition on DPPC

The effect of temperature was first investigated by imaging **CHOL** 3PS tiles on DPPC in a range of temperatures. **CHOL** 3PS tiles (V_{Dep} = 10 μL) were deposited on DPPC at ambient conditions after standard bilayer preparation. The sample and fluid cell were positioned on the heating element on the AFM scanner. Imaging was carried out at 5 °C intervals from 25 °C to 50 °C with a 10 minute equilibration time at each temperature point. The temperature inside the cell was monitored with a thermocouple connected to a digital read thermometer and was consistent with the temperature controller setting throughout the experiment.

The effect of temperature on 3PS tile organization on DPPC was also examined by comparing tile deposition at ambient conditions to deposition above the DPPC transition temperature. The transition temperature for DPPC (from gel phase to liquid disordered, fluid state) is of 42 °C. **U** 3PS tiles were deposited on DPPC with the sample maintained either at 25 °C (ambient conditions) or at 45 °C (above the transition temperature).

After preparing the DPPC bilayers as described in section 4.4.5, the sample was heated to 45 °C for 10 minutes with 60 μL of 1xTAMg on the surface. For tile deposition at high temperature, 5 μL of **U** 3PS tile (0.5 μM) was added at this point, allowed to equilibrate for 2 minutes and cooled back to 25 °C over 10 minutes, representing an annealing process. For deposition at ambient conditions, the bilayer sample was cooled back from 45 °C over 10 minutes before tile addition and 2 minute equilibration period. Samples were imaged at ambient conditions as described in section 4.4.6.

4.4.9 Additional AFM images

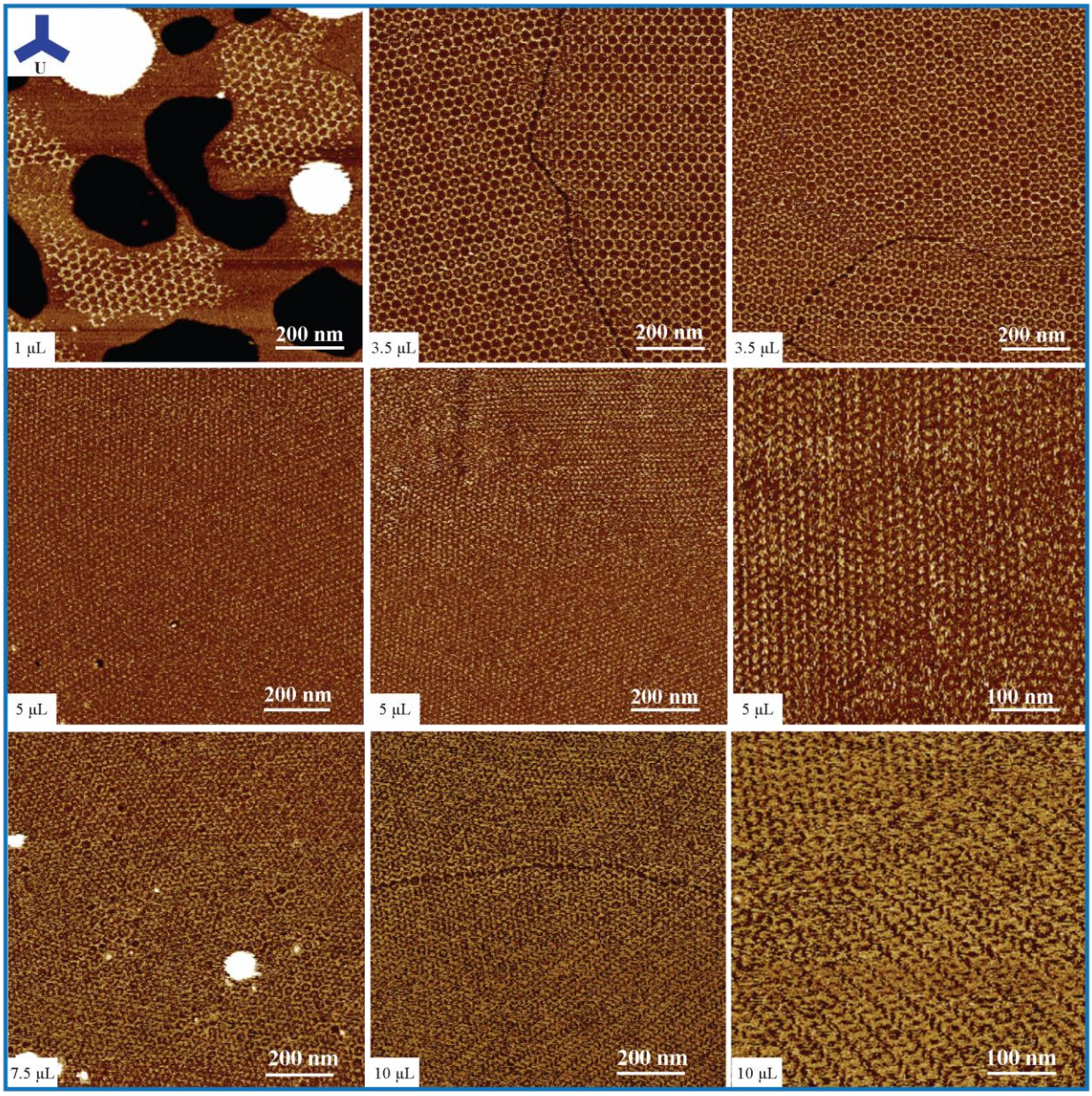


Figure 4.27| U 3PS tiles on DPPC with increasing V_{Dep}. V_{Dep} and scale bars as indicated on the images. Stock tile concentration of 0.5 µM.

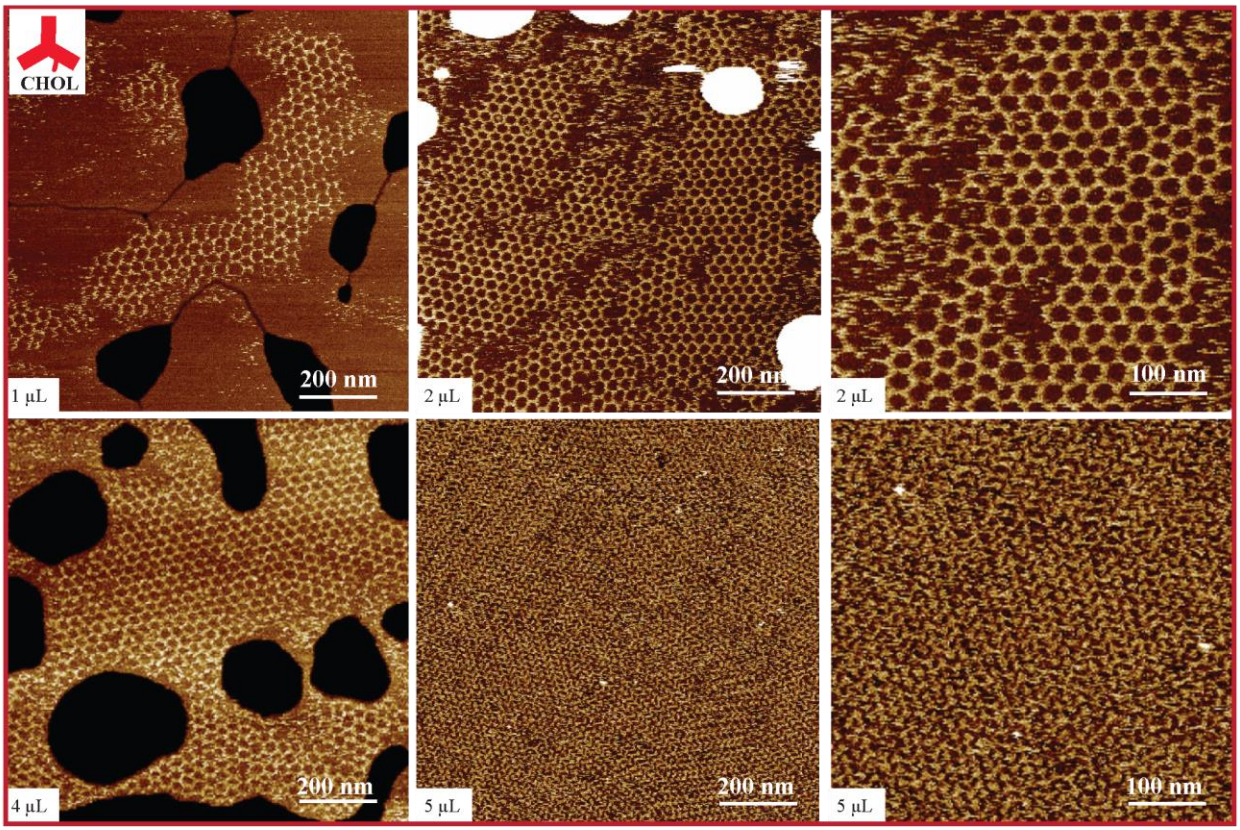


Figure 4.28| CHOL 3PS tiles on DPPC with increasing V_{Dep} V_{Dep} and scale bars as indicated on the images. Stock tile concentration of 0.5 µM.

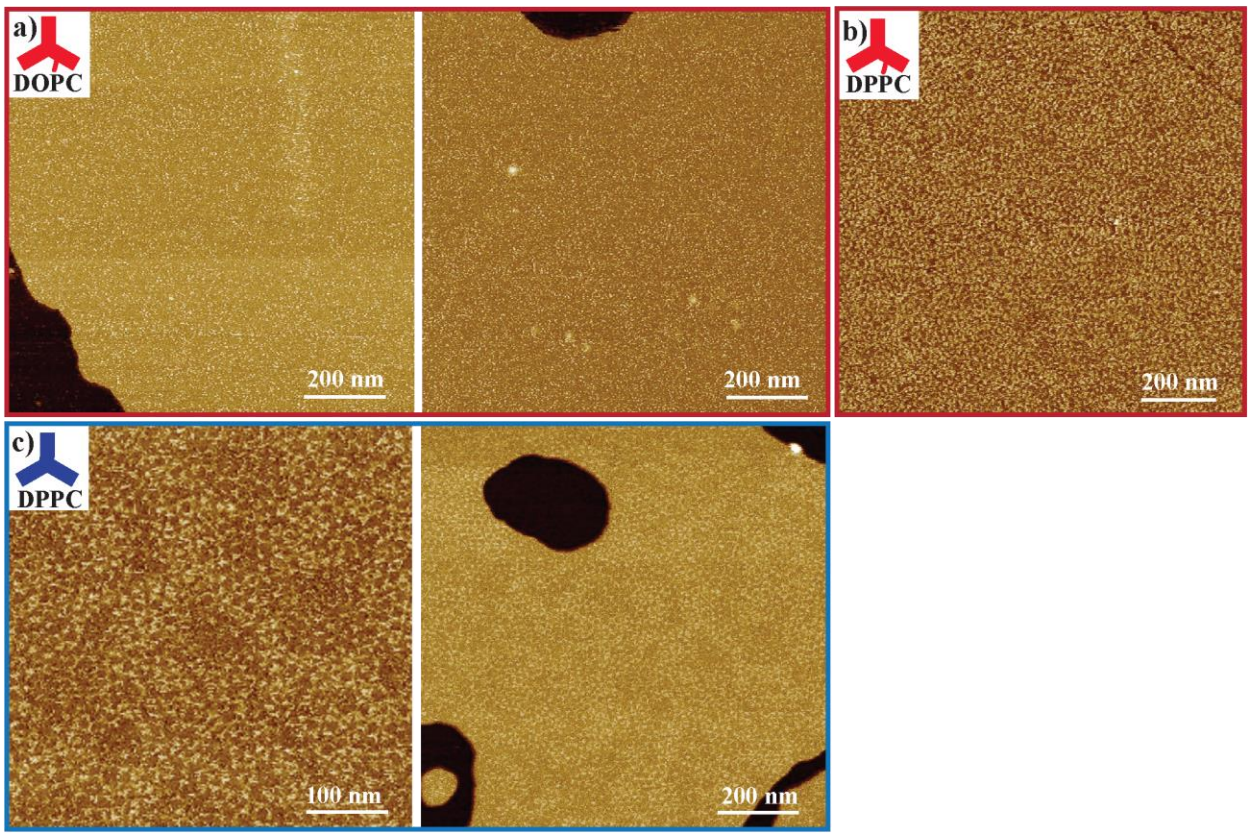


Figure 4.29| 3PS T loop tiles on SLBs. a) CHOL tiles on DOPC (V_{Dep} = 2 µL). **b) CHOL** tiles on DPPC (V_{Dep} = 10 µL). **c) U** tiles on DPPC (V_{Dep} = 7.5 µL). Stock tile concentration of 0.5 µM, scale bars as indicated on the images.

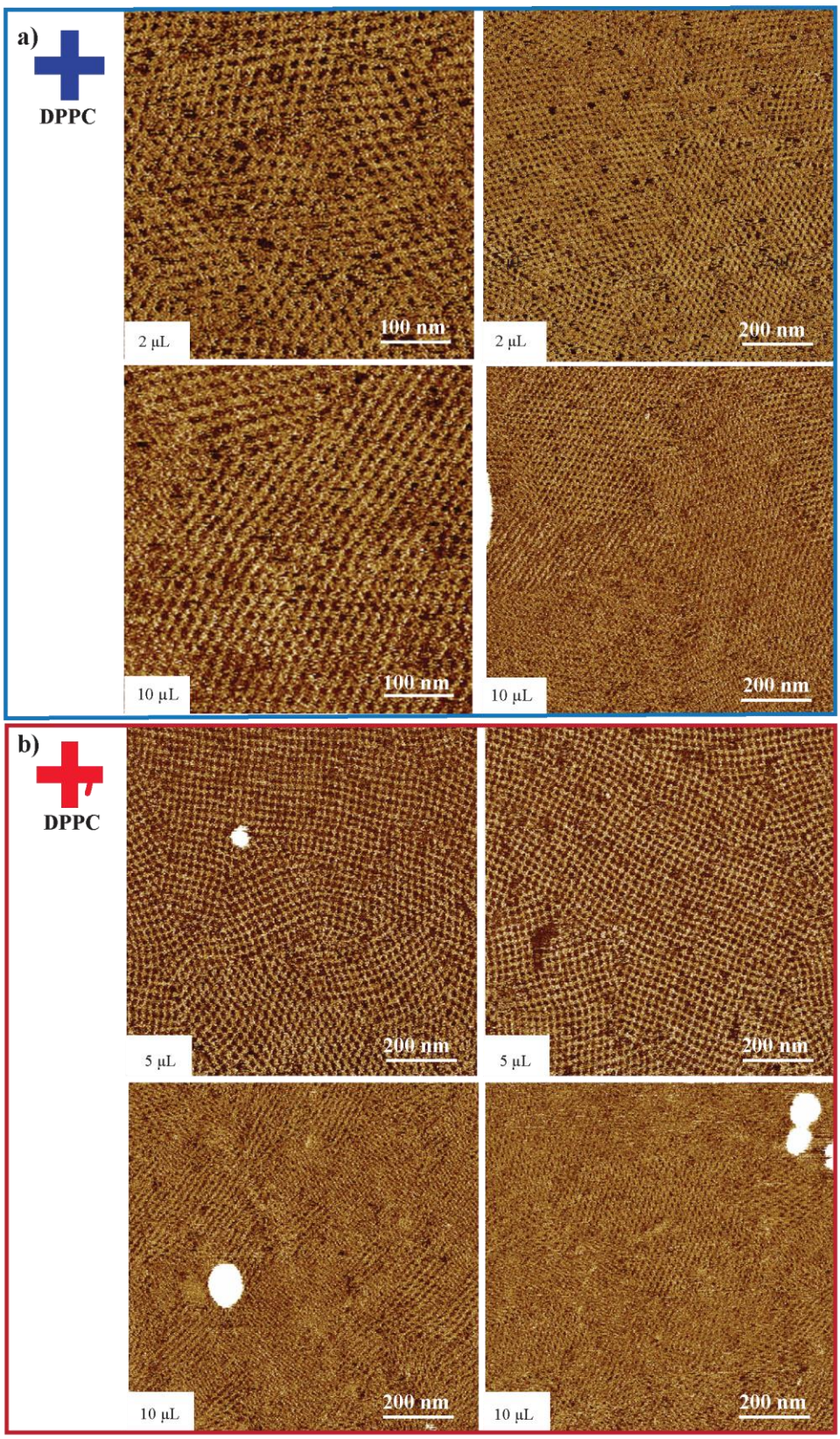


Figure 4.30| X tiles on DPPC. AFM images of **a) CHOL** tiles (V_{Dep} = 5 µL – top, 10 µL – bottom) and **b) U** tiles (V_{Dep} = 2 µL – top, 10 µL – bottom). Scale bars as indicated on the images.

4.5 References

1 Le, J. D., Pinto, Y., Seeman, N. C., Musier-Forsyth, K., Taton, T. A. & Kiehl, R. A. DNA-templated self-assembly of metallic nanocomponent arrays on a surface. *Nano Lett.* **4**, 2343-2347 (2004).

2 Zhang, J., Liu, Y., Ke, Y. & Yan, H. Periodic square-like gold nanoparticle arrays templated by self-assembled 2D DNA nanogrids on a surface. *Nano Lett.* **6**, 248-251 (2006).

3 Sharma, J., Ke, Y., Lin, C., Chhabra, R., Wang, Q., Nangreave, J., Liu, Y. & Yan, H. DNA-tile-directed self-assembly of quantum dots into two-dimensional nanopatterns. *Angew. Chem., Int. Ed.* **47**, 5157-5159 (2008).

4 Yan, H., Park, S. H., Finkelstein, G., Reif, J. H. & LaBean, T. H. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* **301**, 1882-1884 (2003).

5 He, Y., Tian, Y., Ribbe, A. E. & Mao, C. Antibody nanoarrays with a pitch of ~20 nanometers. *J. Am. Chem. Soc.* **128**, 12664-12665 (2006).

6 Chhabra, R., Sharma, J., Ke, Y., Liu, Y., Rinker, S., Lindsay, S. & Yan, H. Spatially addressable multiprotein nanoarrays templated by aptamer-tagged DNA nanoarchitectures. *J. Am. Chem. Soc.* **129**, 10304-10305 (2007).

7 Chandrasekaran, A. R. Programmable DNA scaffolds for spatially-ordered protein assembly. *Nanoscale* **8**, 4436-4446 (2016).

8 He, Y., Chen, Y., Liu, H., Ribbe, A. E. & Mao, C. Self-assembly of hexagonal DNA two-dimensional (2D) arrays. *J. Am. Chem. Soc.* **127**, 12202-12203 (2005).

9 Seeman, N. C. Nanomaterials based on DNA. *Annu. Rev. Biochem.* **79**, 65-87 (2010).

10 Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302 (2006).

11 Nangreave, J., Han, D., Liu, Y. & Yan, H. DNA origami: a history and current perspective. *Curr. Opin. Chem. Biol.* **14**, 608-615 (2010).

12 Carneiro, K. M. M., Avakyan, N. & Sleiman, H. F. Long-range assembly of DNA into nanofibers and highly ordered networks. *Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol.* **5**, 266-285 (2013).

13 Pfeifer, W. & Saccà, B. From nano to macro through hierarchical self-assembly: the DNA paradigm. *ChemBioChem* **17**, 1063-1080 (2016).

14 Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394**, 539-544 (1998).

15 Liu, W., Zhong, H., Wang, R. & Seeman, N. C. Crystalline two-dimensional DNA origami arrays. *Angew. Chem., Int. Ed.* **50**, 264-267 (2011).

16 Ke, Y., Ong, L. L., Sun, W., Song, J., Dong, M., Shih, W. M. & Yin, P. DNA brick crystals with prescribed depths. *Nat. Chem.* **6**, 994-1002 (2014).

17 Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **34**, 564-574 (2006).

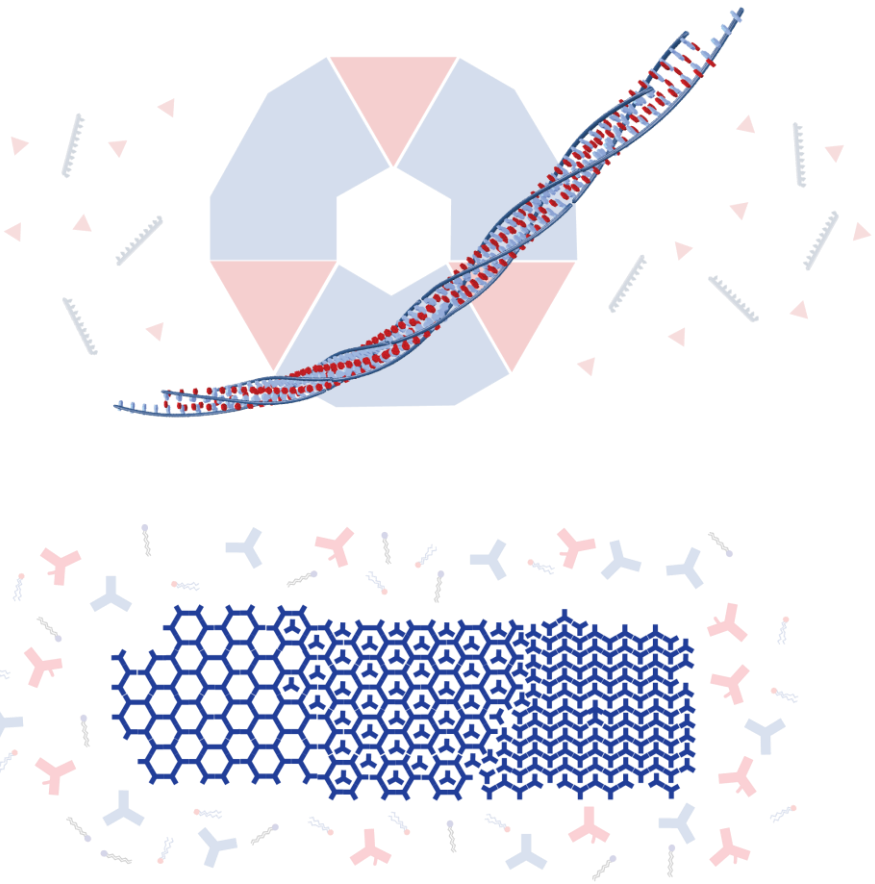
18 Wang, R., Kuzuya, A., Liu, W. & Seeman, N. C. Blunt-ended DNA stacking interactions in a 3-helix motif. *Chem. Commun.* **46**, 4905-4907 (2010).

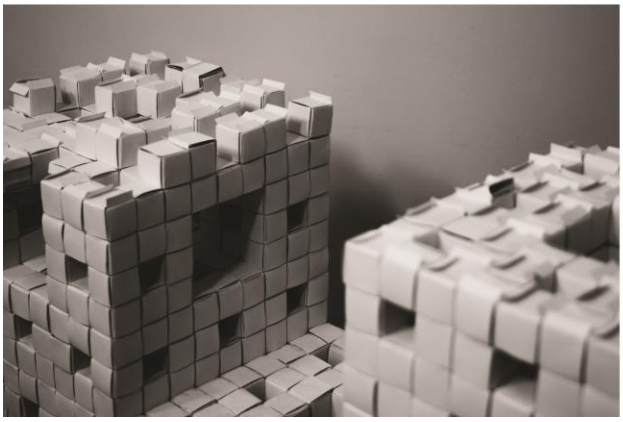
19 Woo, S. & Rothemund, P. W. K. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nat. Chem.* **3**, 620-627 (2011).

20 Gerling, T., Wagenbauer, K. F., Neuner, A. M. & Dietz, H. Dynamic DNA devices and assemblies formed by shape-complementary, non–base pairing 3D components. *Science* **347**, 1446-1452 (2015).
21 Schade, M., Berti, D., Huster, D., Herrmann, A. & Arbuzova, A. Lipophilic nucleic acids — A flexible construction kit for organization and functionalization of surfaces. *Adv. Colloid Interface Sci.* **208**, 235-251 (2014).
22 Langecker, M., Arnaut, V., List, J. & Simmel, F. C. DNA nanostructures interacting with lipid bilayer membranes. *Acc. Chem. Res.* **47**, 1807-1815 (2014).
23 Langecker, M., Arnaut, V., Martin, T. G., List, J., Renner, S., Mayer, M., Dietz, H. & Simmel, F. C. Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science* **338**, 932-936 (2012).
24 Burns, J. R., Göpfrich, K., Wood, J. W., Thacker, V. V., Stulz, E., Keyser, U. F. & Howorka, S. Lipid-bilayer-spanning DNA nanopores with a bifunctional porphyrin anchor. *Angew. Chem., Int. Ed.* **52**, 12069-12072 (2013).
25 Johnson-Buck, A., Jiang, S., Yan, H. & Walter, N. G. DNA–cholesterol barges as programmable membrane-exploring agents. *ACS Nano* **8**, 5641-5649 (2014).
26 Suzuki, Y., Endo, M., Yang, Y. & Sugiyama, H. Dynamic assembly/disassembly processes of photoresponsive DNA origami nanostructures directly visualized on a lipid membrane surface. *J. Am. Chem. Soc.* **136**, 1714-1717 (2014).
27 Kocabey, S., Kempter, S., List, J., Xing, Y., Bae, W., Schiffels, D., Shih, W. M., Simmel, F. C. & Liedl, T. Membrane-assisted growth of DNA origami nanostructure arrays. *ACS Nano* **9**, 3530-3539 (2015).
28 Yang, Y., Wang, J., Shigematsu, H., Xu, W., Shih, W. M., Rothman, J. E. & Lin, C. Self-assembly of size-controlled liposomes on DNA nanotemplates. *Nat. Chem.* **8**, 476-483 (2016).
29 Börjesson, K., Lundberg, E. P., Woller, J. G., Nordén, B. & Albinsson, B. Soft-surface DNA nanotechnology: DNA constructs anchored and aligned to lipid membrane. *Angew. Chem., Int. Ed.* **50**, 8312-8315 (2011).
30 Lundberg, E. P., Feng, B., Saeid Mohammadi, A., Wilhelmsson, L. M. & Nordén, B. Controlling and monitoring orientation of DNA nanoconstructs on lipid surfaces. *Langmuir* **29**, 285-293 (2013).
31 Gissot, A., Oumzil, K., Patwa, A. & Barthelemy, P. A hybrid lipid oligonucleotide: a versatile tool for supramolecular chemistry. *New J. Chem.* **38**, 5129-5134 (2014).
32 Dong, Y., Sun, Y., Wang, L., Wang, D., Zhou, T., Yang, Z., Chen, Z., Wang, Q., Fan, Q. & Liu, D. Frame-guided assembly of vesicles with programmed geometry and dimensions. *Angew. Chem., Int. Ed.* **53**, 2607-2610 (2014).
33 Conway, J. W., Madwar, C., Edwardson, T. G., McLaughlin, C. K., Fahkoury, J., Lennox, R. B. & Sleiman, H. F. Dynamic behavior of DNA cages anchored on spherically supported lipid bilayers. *J. Am. Chem. Soc.* **136**, 12987-12997 (2014).
34 Peng, R., Wang, H., Lyu, Y., Xu, L., Liu, H., Kuai, H., Liu, Q. & Tan, W. Facile assembly/disassembly of DNA nanostructures anchored on cell-mimicking giant vesicles. *J. Am. Chem. Soc.* 10.1021/jacs.7b07485 (2017).
35 Dohno, C., Makishi, S., Nakatani, K. & Contera, S. Amphiphilic DNA tiles for controlled insertion and 2D assembly on fluid lipid membranes: the effect on mechanical properties. *Nanoscale* **9**, 3051-3058 (2017).

36 Czogalla, A., Franquelim, Henri G. & Schwille, P. DNA nanostructures on membranes as tools for synthetic biology. *Biophys. J.* **110**, 1698-1707 (2016).
37 Suzuki, Y., Endo, M. & Sugiyama, H. Lipid-bilayer-assisted two-dimensional self-assembly of DNA origami nanostructures. *Nat. Commun.* **6**, 8052 (2015).
38 He, Y., Ye, T., Su, M., Zhang, C., Ribbe, A. E., Jiang, W. & Mao, C. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452**, 198-201 (2008).
39 He, Y., Tian, Y., Chen, Y., Deng, Z., Ribbe, A. E. & Mao, C. Sequence symmetry as a tool for designing DNA nanostructures. *Angew. Chem., Int. Ed.* **44**, 6694-6696 (2005).
40 Maffeo, C., Luan, B. & Aksimentiev, A. End-to-end attraction of duplex DNA. *Nucleic Acids Res.* **40**, 3812-3821 (2012).
41 Nakata, M., Zanchetta, G., Chapman, B. D., Jones, C. D., Cross, J. O., Pindak, R., Bellini, T. & Clark, N. A. End-to-end stacking and liquid crystal condensation of 6– to 20–base pair DNA duplexes. *Science* **318**, 1276-1279 (2007).
42 Czogalla, A., Petrov, E. P., Kauert, D. J., Uzunova, V., Zhang, Y., Seidel, R. & Schwille, P. Switchable domain partitioning and diffusion of DNA origami rods on membranes. *Faraday Discuss.* **161**, 31-43 (2013).
43 Ainalem, M.-L., Kristen, N., Edler, K. J., Höök, F., Sparr, E. & Nylander, T. DNA binding to zwitterionic model membranes. *Langmuir* **26**, 4965-4976 (2010).
44 Malghani, M. S. & Yang, J. Stable binding of DNA to zwitterionic lipid bilayers in aqueous solutions. *J. Phys. Chem. B* **102**, 8930-8933 (1998).
45 Kato, A., Tsuji, A., Yanagisawa, M., Saeki, D., Juni, K., Morimoto, Y. & Yoshikawa, K. Phase separation on a phospholipid membrane inducing a characteristic localization of DNA accompanied by its structural transition. *J. Phys. Chem. Lett.* **1**, 3391-3395 (2010).
46 Beales, P. A. & Vanderlick, T. K. Partitioning of membrane-anchored DNA between coexisting lipid phases. *J. Phys. Chem. B* **113**, 13678-13686 (2009).
47 Woo, S. & Rothemund, P. W. Self-assembly of two-dimensional DNA origami lattices using cation-controlled surface diffusion. *Nat. Commun.* **5**, 4889 (2014).
48 Leonenko, Z. V., Finot, E., Ma, H., Dahms, T. E. S. & Cramb, D. T. Investigation of temperature-induced phase transitions in DOPC and DPPC phospholipid bilayers using temperature-controlled scanning force microscopy. *Biophys. J.* **86**, 3783-3793 (2004).
49 Lawrence, J. C., Saslowsky, D. E., Michael Edwardson, J. & Henderson, R. M. Real-time analysis of the effects of cholesterol on lipid raft behavior using atomic force microscopy. *Biophys. J.* **84**, 1827-1832 (2003).
50 Connell, S. D. & Smith, D. A. The atomic force microscope as a tool for studying phase separation in lipid membranes (Review). *Mol. Membr. Biol.* **23**, 17-28 (2006).
51 Schade, M., Knoll, A., Vogel, A., Seitz, O., Liebscher, J., Huster, D., Herrmann, A. & Arbuzova, A. Remote control of lipophilic nucleic acids domain partitioning by DNA hybridization and enzymatic cleavage. *J. Am. Chem. Soc.* **134**, 20490-20497 (2012).
52 Mingeot-Leclercq, M.-P., Deleu, M., Brasseur, R. & Dufrene, Y. F. Atomic force microscopy of supported lipid bilayers. *Nat. Protocols* **3**, 1654-1659 (2008).

5| Conclusions





Order upon order

Model adapted from: Level 3 Menger sponge made from index cards

Mengermania by Nicholas Rougeux
(https://www.c82.net/mengermania/instructions.php)

Folding, photography and image editing: Nicole Avakyan

5.1 Conclusions and contributions to original knowledge

The work presented in this thesis has aimed to use simple principles of supramolecular chemistry and combinations of intermolecular interactions to expand the assembly modes of DNA and achieve long-range organization within DNA-based nanomaterials. The methods described here did not require intricate engineering or involve synthesis of complex building blocks, rather relying on simple components and interactions to promote higher-order assembly. Having access to a broader range of interactions within DNA nanostructures opens the door for greater control over long-range assembly through hierarchical organization. Intricate control of structure at the nanoscale is important in developing functional nanomaterials, thus the new, easily applicable methods to manipulate DNA assembly presented here may find broad application in the field of DNA-based nanomaterials.

Chapter 2 reported the discovery of a new mode of DNA assembly, wherein the addition of a small molecule with multiple hydrogen-bonding faces reprogrammed the assembly of unmodified DNA. Cyanuric acid (CA) was shown to coax poly(adenine) DNA ($d(A_n)$) into forming nanofibers with a unique internal structure. The association of adenine and CA units was proposed to form hydrogen-bonded hexameric rosettes, bringing together $d(A_n)$ strands into a triple helix, and further leading to polymerization of the components into nanofibers. A combination of indirect characterization methods to validate the proposed structure was described – a collection of evidence in support of the hexameric rosette structure. The versatility of CA-mediated assembly was probed by examining CA interaction with poly(adenine) RNA and PNA, both of which also produced nanofibers. Furthermore, the potential of further expanding the base-pairing alphabet via small-molecule mediated assembly was assessed by screening a library of CA analogs for assembly with $d(A_n)$. Finally, biotinylated CA-mediated $d(A_n)$ nanofibers were decorated with streptavidin to illustrate their potential for organizing functional moieties. Altogether, CA-mediated $d(A_n)$ assembly is a facile new method to create a nucleic acid-based nanomaterial which can be generated in large quantities from inexpensive, biocompatible building blocks. The nanofibers are stable over a range of pH values and counterion concentrations. Small molecule-mediated assembly of nucleic acids is, in principle, a broadly applicable method to introduce new letters into the DNA base-pairing alphabet that is not limited to the CA:$d(A_n)$ system.

Chapter 3 further examined CA-mediated d(A_n) assembly by focusing on aspects of the mechanism of nanofiber formation and key parameters affecting the process. Strands within the assembly were shown to take on a staggered arrangement and the supramolecular polymerization mechanism was consistent with a cooperative, or nucleation-elongation, process. A minimum threshold CA concentration ($[CA]_{cr}$) was required to initiate long-range assembly and CA concentration in general was identified as the most important parameter regulating both the assembly process and nanofiber properties such as thermal stability. In addition, d(A_n) strand length had interesting effects on the thermal denaturation profiles of CA-mediated assemblies, as multi-phase melting profiles were observed for strands longer than d(A_{15}). Both CA concentration and d(A_n) strand length impacted the kinetics of the polymerization process, as evidenced by the in-depth analysis of thermal hysteresis profiles using a novel method to study supramolecular assembly developed by Harkness and Mittermaier. This methodology used model-free analysis to map monomer consumption rates with respect to monomer concentration and temperature, enabling determination of effective reaction orders for assembly over a range of temperatures. This detailed information was then used to globally fit the data to a cooperative polymerization model, and to build quantitative energy diagrams describing the assembly process. Important insights as to the nucleation molecularity as well as a quantitative representation of the impact of CA concentration and d(A_n) length on kinetic and thermodynamic parameters were revealed by this analysis. Overall, the understanding of the mechanism and impact of tunable parameters on CA-mediated d(A_n) assembly described in Chapter 3 is valuable not only from a fundamental self-assembly point of view, but also for future development of applications of this unique DNA-based nanomaterial.

Chapter 4 studied the long-range ordering of small blunt-ended DNA crossover tiles on fluid and gel-phase zwitterionic supported lipid bilayers (SLBs). The addition of a cholesterol anchor was necessary for tile deposition and extended lattice formation on fluid DOPC bilayers, whereas both unmodified and cholesterol-anchored tiles could bind and form ordered assemblies on the DPPC gel-phase lipid. Increased density of Mg^{2+}-mediated electrostatic interactions between DNA and zwitterionic lipid headgroups in gel-phase bilayers, compared to fluid bilayers, explains the divergent deposition behavior. In accordance with these properties, selective patterning of different domains in a phase-separated lipid system was observed. Furthermore, tile packing on gel-phase SLBs depended on tile concentration and several packing modes of

increasing density, maximizing electrostatic contacts between lipid headgroups and DNA, were identified. Base stacking interactions between blunt tile-arm ends were of paramount importance to obtain the various morphologies and long-range tile organization. Their dynamic nature allowed error correction to occur, giving rise to defect-free tile networks on the fluid bilayer, and facilitated formation of dense packing modes on gel-phase bilayers. The balance of inter-tile base-stacking interactions and tile-bilayer interactions showcased once again how simple interactions can be harnessed to manipulate long-range DNA assembly of simple nanostructures composed of just a few inexpensive strands.

Overall, this thesis contributes to the broad field of self-assembly and, more specifically, to DNA nanotechnology by taking advantage of the information encoded in simple molecular building blocks to reprogram and manipulate DNA assembly and long-range organization. Just as in natural self-assembly systems, the interplay between simple intermolecular interactions was shown to drive emerging complexity in the nanostructure of the materials generated using the methods presented here. The resulting assemblies displayed unique features and complexity that suggest new avenues of development for DNA-based nanomaterials.

5.2 Future work and perspective

The report of CA-mediated d(A_n) assembly described in Chapters 2 and 3 has important consequences in several areas. There are a number of biologically relevant bifacial "Janus" molecules, such as 8-oxo-guanine or folic acid, among many others, which may also be able to form higher-order structures with nucleic acid strands. The supramolecular behavior of small hydrogen bonding molecules with nucleic acids has not been examined in detail to our knowledge, so there is much work to be done in this area. The small molecule-induced reprogramming of nucleic acid assembly illustrated with the CA-d(A_n) system may serve as a model for further expansion of the DNA alphabet by combining other compatible small molecules with natural DNA bases. It is worth noting that this plug-and-play methodology may give rise to alternative assembly molecularities and geometries. Their formation may also be favored in different pH ranges, depending on the properties of the small molecule. As the CA-mediated assembly motif was shown to coexist with standard Watson-Crick DNA base-pairing, it introduces an orthogonal interaction to the nucleic acid-based nanomaterials toolkit, expanding the four-letter DNA alphabet and the

repertoire of interactions available to increase the complexity of DNA nanostructures. This expansion fits into the field of supramolecular DNA assembly, which incorporates heterologous components into DNA nanostructures to introduce additional types of supramolecular interactions, diversify the language of DNA assembly, and achieve structures with new functions. The coexistence of Watson-Crick and small molecule-mediated assembly in DNA nanostructures is an element of hierarchical organization which may used to achieve long range assembly from smaller structures and incorporate stimuli-responsive character into the architectures. Furthermore, guest-induced assembly of DNA strands using a small molecule with multiple hydrogen bonding faces may be relevant in chemical biology. Long poly(adenine) extensions are added to messenger RNA before its translation to a protein and poly(purine) stretches in general are biologically important. Our study raises the possibility that CA (or derivatives) could be used to probe and potentially modulate protein translation. To move into this direction, steps must first be taken to stabilize assembly at physiological pH values and more elevated temperatures. Following that, it is worth exploring whether intracellular small-molecule mediated assembly is achievable and can be used to modulate protein expression. Finally, the fact that a small molecule such as CA can organize $d(A_n)$ strands into long-range supramolecular polymers in aqueous media may have implications in prebiotic chemistry, specifically in the context of how ancestral nucleobases became incorporated into pre-nucleic acid polymers. The evolution of natural nucleobases from plausible prebiotic recognition units such as melamine and cyanuric acid is an interesting question, as is the possibility of polymerization of building blocks following formation of supramolecular fibers from small molecules. In the context of CA-mediated $d(A_n)$ assembly, studying strategies to link individual CA molecules post-assembly with $d(A_n)$ strands into oligomers may be a worthwhile line of investigation.

Additionally, insight about the thermal denaturation properties of CA-mediated assembly products and the mechanisms behind their formation obtained in Chapter 3 are important first steps toward characterizing the fiber nanomaterial and understanding how to tune its properties in view of potential applications. The stimuli-responsive nature of the self-assembled nanofibers and their capacity to serve as a scaffold for functional moieties along the fiber axis invite the possibility of developing tunable, functional hydrogel materials which may find applications in tissue engineering and other areas. A systematic investigation of the influence of [CA], $d(A)_n$ strand length and environmental factors such as counterion concentration and temperature on fiber

populations and fiber length distribution would be valuable not only for a deeper fundamental understanding of the supramolecular polymerization process, but also for regulation of nanomaterial properties to a desired application. Given the nucleation-elongation nature of the fiber polymerization mechanism, it would be interesting to examine whether fiber length can be controlled by designing conditions to carry out "seeded" polymerization, a situation where supramolecular polymers acquire a "living" character. One of the main current challenges to application of the nanofibers as a biomaterial is their poor thermal stability. Strategies to improve their robustness include crosslinking or the development of CA-bearing oligomers, which may face a smaller entropic barrier to assembly and raise the melting temperature of the nanofibers. Another aspect of CA-mediated d(A_n) assembly that deserves further attention is the validation of the internal structure with a direct method such as X-ray crystallography or NMR spectroscopy. Preliminary studies using NMR spectroscopy are described in the Appendix.

It should be noted that application of the thermal hysteresis-based analysis of assembly mechanisms described in Chapter 3 is not limited to supramolecular polymerization. It is in fact a broadly applicable methodology to study complex self-assembly processes as it can track temperature-dependent reaction orders with a great level of detail from a dataset that is easy to obtain using commonly available instrumental equipment. The main requirements for the application of this analysis methodology is that the reaction under investigation is reversible and temperature-controlled, that the components remain soluble over the full temperature range studied, and that the extent of self-assembly correlates with a real-time response such as spectroscopic absorbance or ellipticity.

The dynamic long-range assembly of DNA tiles on supported lipid bilayers reported in Chapter 4 represents a simple way to organize DNA nanomaterials on the micron-scale at a biologically-relevant interface. The ability to create ordered networks on lipid bilayers can find applications in areas such as patterned surfaces for selective cell recognition or templates for organizing proteins and nanoparticles on a biocompatible surface. The simple principles described in this study can be extended to tiles with different anchoring moieties and bilayers of various compositions towards the development of hierarchical materials of higher pattern complexity. Different bilayer regions in mixed composition systems can be specifically targeted by network deposition, the periodic networks may serve as a framework for long-range organization of functional moieties such as proteins and nanoparticles tethered to individual tiles, and network

cavities may also serve as areas of deposition of additional components. Moreover, the concentration-dependent changes in network morphology may be harnessed for a layer-by-layer deposition strategy, a method to build complexity in nanomaterial patterning with successive additions of tiles bearing different functional components. Whereas the dynamic rearrangement properties of blunt tiles on a fluid support are of interest in building error-free, long-range, periodic networks, it is also worthwhile to consider strategies to freeze the assembled material in a given morphology to maintain its structure overtime. Designing methods of crosslinking the tiles in a desirable orientation is thus a potential development step for this platform. Tiles modified with metal-binding ligands or anthracene units capable of photo-crosslinking are possible candidates for this line of inquiry. Altogether, organized DNA nanomaterials on lipid bilayers may enable specific recognition between cells to drive the formation of designer cell-clusters, or act like a cytoskeleton mimic that is able to control the shape of a cell. Furthermore, networks with addressable long-range organization may serve as platforms to build circuits or light-harvesting farms, among other nanotechnological applications.

5.3 List of publications

1 KMM Carneiro,* N Avakyan* & HF Sleiman. Long-range assembly of DNA into nanofibers and highly ordered networks. *Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol.* **5**, 266-285 (2013).

* Equal contribution

2 N Avakyan, AA Greschner, F Aldaye, CJ Serpell, V Toader, A Petitjean & HF Sleiman. Reprogramming the assembly of unmodified DNA with a small molecule. *Nat. Chem.* **8**, 368-376 (2016).

3 N Avakyan, JW Conway & HF Sleiman. Long-range ordering of blunt-ended DNA tiles on supported lipid bilayers. *J. Am. Chem. Soc.* **139**, 12027-12034 (2017).

4 RW Harkness V, N Avakyan, HF Sleiman & AK Mittermaier. Mapping the energy landscapes of supramolecular assembly by thermal hysteresis. *Nat. Commun.* **9**, 3152 (2018).

Featured on the *Nature Communications* Editors' Highlights page.

A1| Investigation of CA-mediated d(A_n) assembly by NMR spectroscopy

Contributions

NA designed and synthesized the molecules, designed experimental conditions and analysed the results. **Donatien de Rochambeau** carried out an additional round of synthesis and purification of 3HA7H and 3HA10H strands to provide more material for NMR experiments. **Prof. Carlos González** and **Dr. Irene Gomez Pinto** performed the NMR experiments with assistance from **NA**, and guided interpretation of results.

A1.1 Introduction

The discovery of CA-mediated assembly of unmodified d(A_n) strands into supramolecular nanofibers was described in Chapter 2. The internal structure of the assembly is proposed to be composed of hexameric rosette arrays that form through hydrogen bonding associations between CA and adenine residues in a 1:1 stoichiometry. This proposed structure is supported by a number of experimental findings. Examination of the assembly of monomeric adenosine and CA analogs in organic solvent is consistent with the formation of rosette assemblies by VPO and DOSY NMR. In aqueous native environment, using equilibrium dialysis experiments with d(A_{15}) strands, we have established that CA-mediated assemblies have 1:1 CA:A stoichiometry. We have also shown that both the Watson-Crick and Hoogsteen faces of adenine residues are required for assembly to take place by investigating the association of CA with DNA strands composed of adenine analogs that lack a Hoogsteen binding face. In addition, the height of the nanofibers measured from AFM images is consistent with the estimated fiber diameter for d(A_n) triplexes with a hexameric rosette structure at their core.

Although these results paint a convincing picture supporting the proposed structure, they represent indirect evidence. A direct examination of the structure has proved elusive. A molecular

representation of the rosette structure would be achievable by X-ray analysis of a crystal structure, or, in solution, by investigating the CA-mediated association of d(A_n) strands using nuclear magnetic resonance (NMR) spectroscopy. NMR spectroscopy is a powerful tool to study nucleic acid structure in solution. Notably, NMR methods have been used to solve non-canonical DNA structures such as the i-motif[1] and the d(A_n) duplex,[2] to study small molecule binding to DNA[3] or the structures of DNA aptamers[4] among many revelatory applications.

Using a combination of 1D ^{1}H NMR, total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY), it is possible to assign the relevant proton signals for short strands. Labeling DNA with ^{13}C and ^{15}N nuclei provides additional tools for structural investigation. NOE signals are also used to calculate proton to proton distances (only short distances can be measured, in the range of 5-6 Å), essential parameters to determine intra- and inter-residue distances and local structure, which along with information on dihedral angles are later used as constraints for 3D structural simulations in solving a structure with atomic-scale resolution.[5]

In this appendix, a series of DNA strands composed of short d(A_n) stretches, capping moieties and branching elements providing preorganization were designed to investigate CA-mediated assembly using NMR spectroscopy. The systems were designed with the goal of ensuring that CA-mediated assembly takes place, but that polymerization is supressed, as it would create aggregation products undetectable by NMR. Although initial experiments suggested that branched molecules may prevent long range aggregation while still showing CA-mediated assembly features, temperature dependent 1D NMR spectra later showed that aggregation did take place, prohibiting investigation of the CA-mediated assembled species. Attempts to look at dynamic species exchange behaviour at intermediate temperatures in the middle of the melting transition did not reveal new information. A variety of temperature dependent 1D NMR studies illustrated further some known behaviours of CA-mediated d(A_n) assembly. These studies indicate that an alternative design of the preorganized constructs is required to prevent aggregation. Suggestions are given at the end of the Appendix.

A1.2 Results and discussion

A1.2.1 Preparation and characterization of linear and branched d(A_n) strands

Strand design choices were driven by requirements of NMR experiments. They can be summarized as follows: (i) minimal d(A_n) strand length and (ii) formation of discrete structures that avoids long-range assembly. A list of the strands that were synthesized is presented in **Table A1.1**, followed by a discussion of the choices made in strand design.

For these strands with repetitive d(A_n) sequences, minimal strand length is necessary to simplify the analysis of NMR spectra and avoid as much as possible overlap in signal for the different adenosine residues. On the other hand, a certain minimal number of residues is required for CA-mediated structure to assemble. Strands with 7 or 10 adenine residues were chosen based on previous experiments. d(A_n) strands with 10 residues were used in the design of the preorganized intramolecular system (Chapter 2), where they were shown to successfully assemble with CA, while DLS results indicated that long range structures were avoided. In the past, shorter d(A_7) strands were shown to successfully form CA-mediated structures at reasonable CA concentrations (10 mM), while shorter strands such as d(A_5) required nearly saturated CA solution (18 mM) for any assembly to be observed. In addition, structures formed from d(A_5) strands were highly thermally unstable and required long equilibration times for assembly, thus the shortest strand length chosen for the current design was d(A_7).

The second design requirement is the elimination of large supramolecular polymer formation, while maintaining the core CA-mediated structure intact over a short range. Essentially, we wish to arrest CA-mediated assembly at the nucleus stage, where we propose that a stack of CA:A rosettes mediates the formation of an all parallel d(A_n) triplex. This is a necessary requirement as preliminary NMR experiments with short d(A_n) strands had been unsuccessful because most of the DNA material participated in large structures that were invisible to NMR. Large structures tumble slowly in solution and the spin-spin relaxation time T_2^* decreases leading to substantial signal decay.[6] Our experience with the intramolecular preorganized system (Chapter 2) indicated that formation of large aggregates was avoided in that case. The preorganization of that system relied on the presence of 12-mer DNA duplexes flanked by flexible spacers. The presence of this additional DNA material would be detrimental to the design of an NMR-appropriate system, as it would mask the adenosine signals of interest. An alternative solution

proposed in this case is the use of branching phosphoramidites to grow d(A_n) stretches in a parallel orientation from the branching point. A symmetrical branching phosphoramidite can be used to preorganize 2 d(A_n) strands, while a trebler branching unit can preorganize 3 parallel d(A_n) strands. Although the use of the trebler produces a desirable strand arrangement given our current understanding of CA-mediated assembly, initial synthesis attempts showed that synthesis yields using this molecule were limited. For this reason, strands using the symmetrical branching units were synthesized as an alternative higher yielding design, as NMR experiments require a large amount of DNA material.

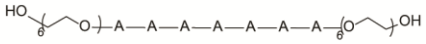
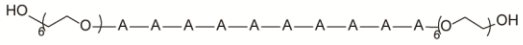
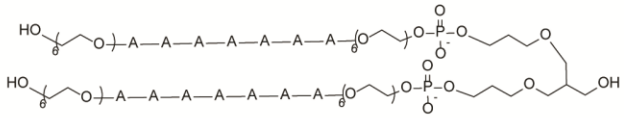
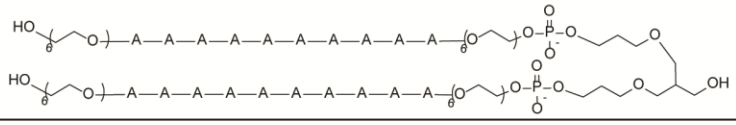
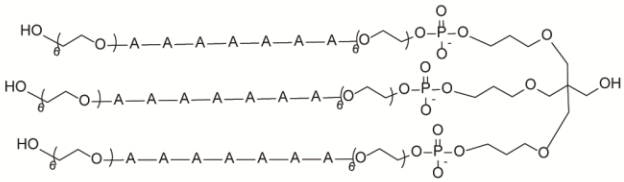
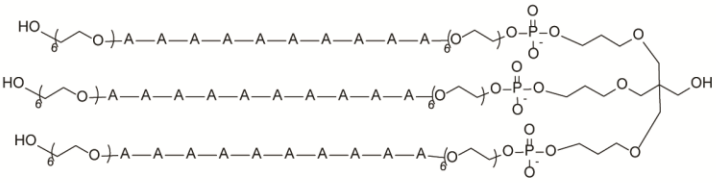
Strand	Structure (5' – 3')	Exact Mass	Molecular Weight
CA7C	CAAAAAAAC	2707.54	2708.87
HA7H		2817.69	2819.10
HA10H		3756.87	3758.73
2HA7H		5851.35	5854.22
2HA10H		7729.69	7733.48
3HA7H		8949.15	8953.57
3HA10H		11766.67	11772.46

Table A1.1| Linear and branched d(A_n) strands. The structure, exact mass and molecular weight (both in g/mol) are listed for each strand.

In addition to preorganization, capping moieties were included in the design to further discourage CA-mediated polymerization. Hexaethylene glycol (HEG) was chosen as the capping unit for most strands as its protons would produce ^{1}H shifts away from the signals of interest for adenosine. Capped versions of monomeric (unbranched) d(A_7) and d(A_{10}), HA7H and HA10H, were also synthesized. The strand CA7C, a d(A_7) strand without synthetic modifications simply capped with cytosine units was also investigated in preliminary experiments.

The strands listed in **Table A1.1** were synthesized according to standard automated solid-phase synthesis procedures and purified by using reverse-phase packed cartridges that rely on the presence of the final 5'-DMT protecting group. This choice in purification procedure was made because of poor yields from commonly used polyacrylamide gel electrophoresis (PAGE) purification. The yields of unbranched strands and strands synthesized using the symmetrical branching unit following cartridge purification were excellent. On the other hand, molecules containing the trebler branching unit were obtained only in much lower yields, partly because of incomplete coupling following trebler addition (strands with only one or two branches), but also due to less efficient cleavage of the final product from the solid support.

Strand characterization by denaturing PAGE was unsuccessful because of difficulties in staining short d(A_n) strands. The PAGE characterization aimed to verify whether incomplete products with missing branches formed as a result of failed coupling of one of the reactive groups at the branching stage of the synthesis were still present after purification using reverse-phase packed cartridges.

All modified strands were characterized by LC-ESI-MS. No side products were observed for strands HA7H and HA10H (**Figure A1.1** and **Figure A1.2**). Strands incorporating a symmetrical branching moiety did not elute as a single peak in liquid chromatography, indicating potential presence of multiple species, although the main peak in the mass spectrum corresponded to the correct structure. For strands containing the symmetrical branching unit (2HA7H and 2HA10H), small amounts of n-1 products as well as a product truncated at the branching point (only one d(A_n) branch present) were present (**Figure A1.3** and **Figure A1.4**). The intensity of the peak corresponding to the truncated single branch product was compared to the intensity of the full length product. For 2HA7H, the intensity of the peak for the truncated product was ~ 2.5% of the full product, and ~ 4% for 2HA10H.

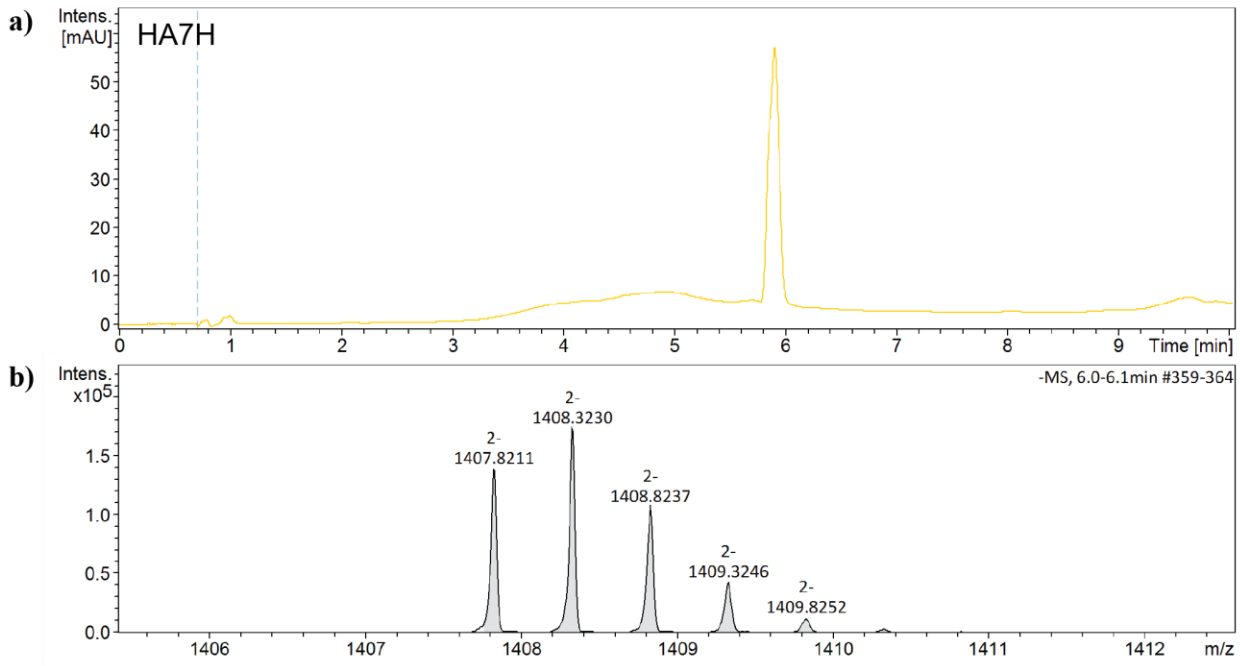


Figure A1.1| LC-ESI-MS analysis for HA7H. a) Chromatogram (UV-Vis absorbance at 260 nm). **b)** Mass spectrum (calculated exact mass = 2 817.69 g/mol, experimental mass from largest peak = 2 816.65 g/mol).

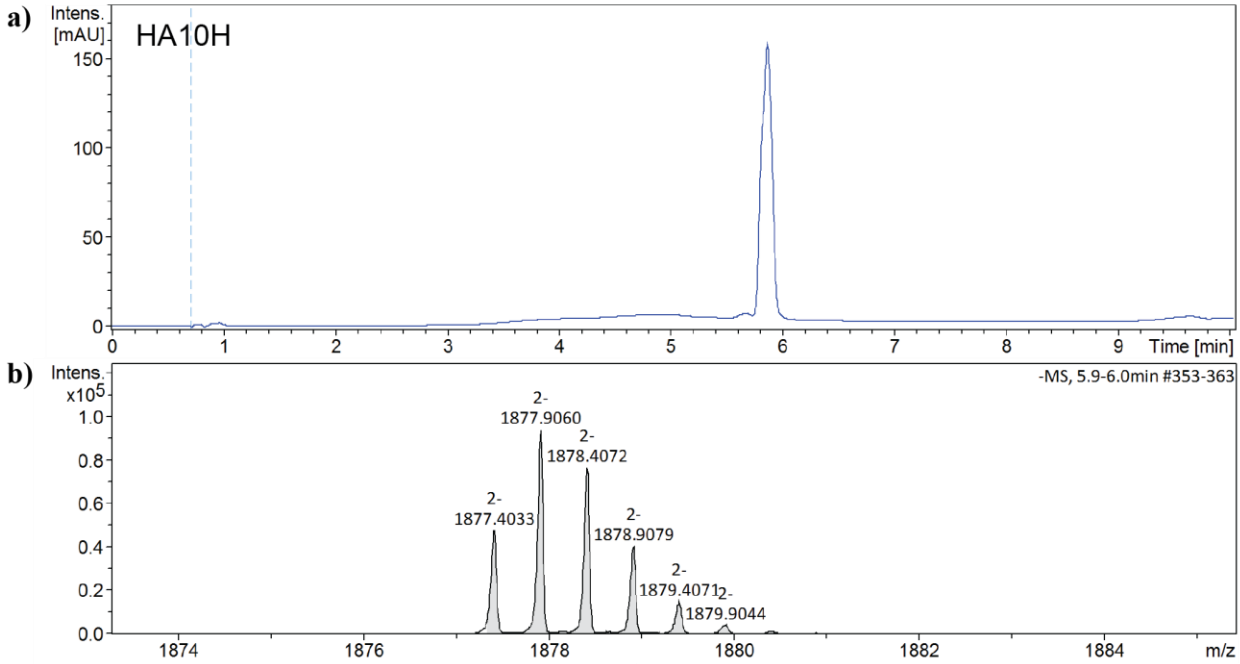


Figure A1.2| LC-ESI-MS analysis for HA10H. a) Chromatogram (UV-Vis absorbance at 260 nm). **b)** Mass spectrum (calculated exact mass = 3756.87 g/mol, experimental mass from largest peak = 3755.81 g/mol).

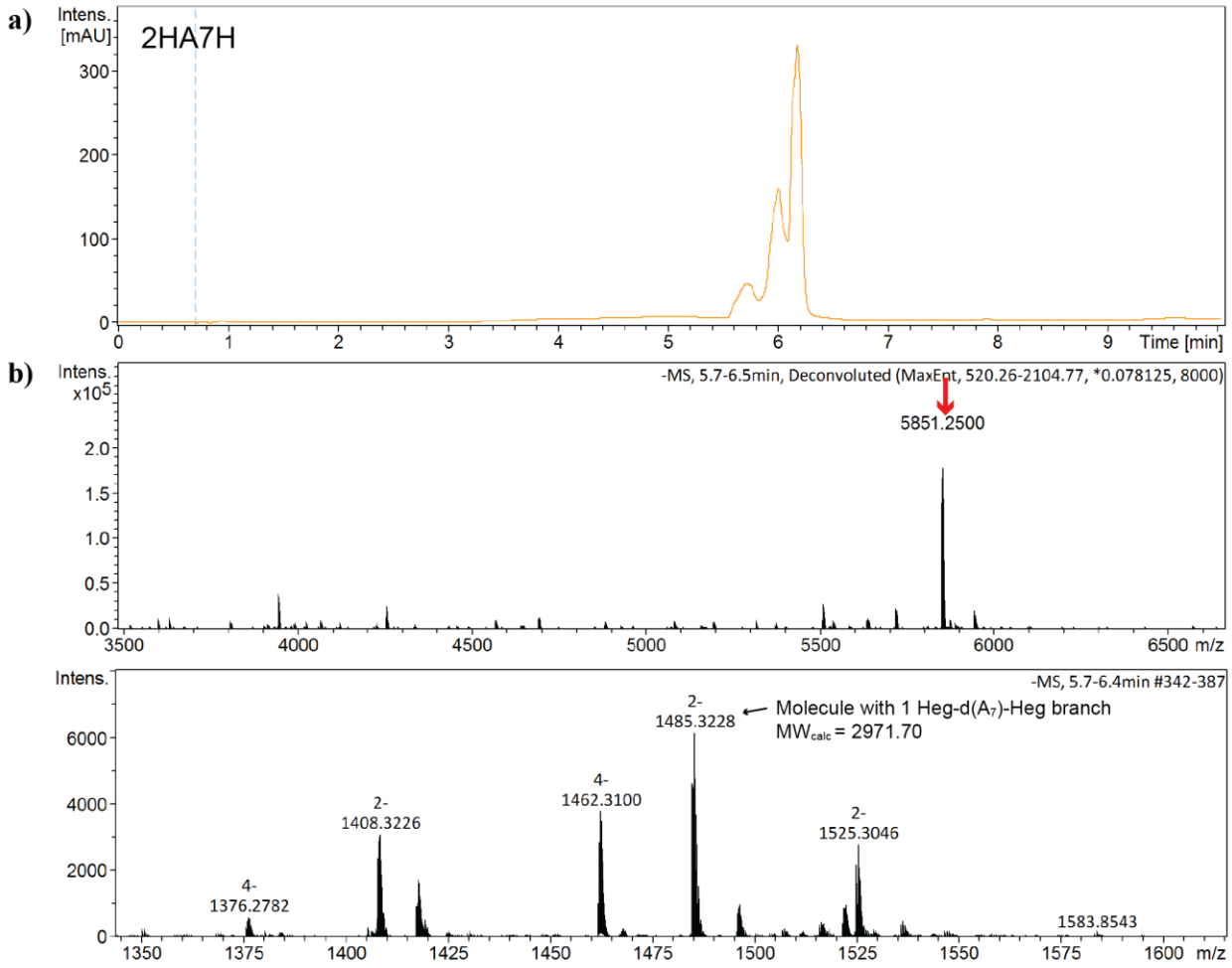


Figure A1.3| LC-ESI-MS analysis for 2HA7H. a) Chromatogram (UV-Vis absorbance at 260 nm). **b)** Deconvoluted mass spectrum (calculated exact mass = 5851.35 g/mol, experimental mass from main peak (red arrow) = 5851.25 g/mol). A lower mass range is also shown to show additional side products (bottom). Peak for species with 2971.79 g/mol mass corresponds to a molecule with only one complete d(A_7) branch grown from the symmetric branching unit. The intensity of this peak is about 2.5% of that of full product peak.

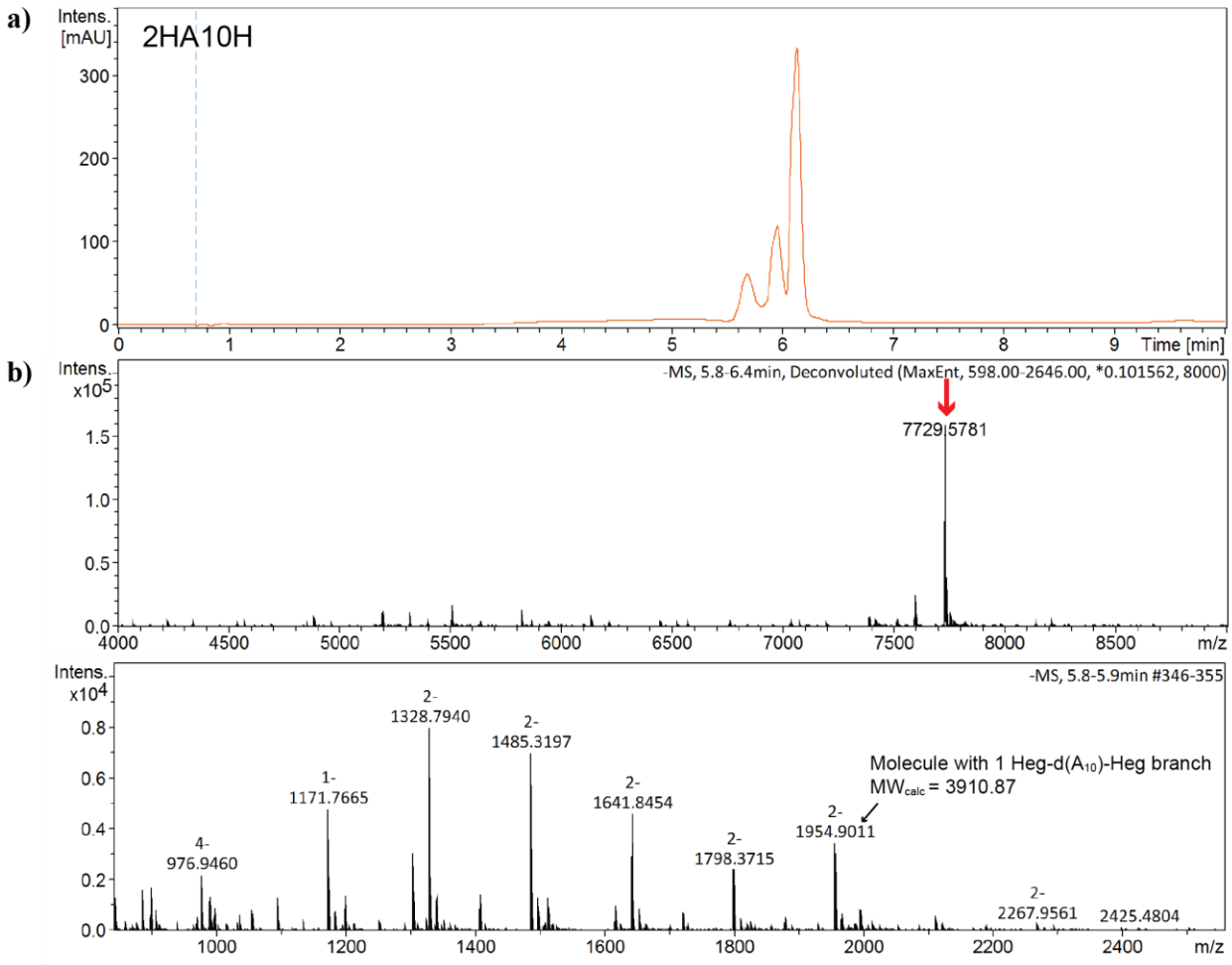


Figure A1.4| LC-ESI-MS analysis for 2HA10H. a) Chromatogram (UV-Vis absorbance at 260 nm). **b)** Deconvoluted mass spectrum (calculated exact mass = 7729.69 g/mol, experimental mass from largest peak (red arrow) = 7729.58 g/mol). A lower mass range is also shown to show additional side products (bottom). Peak for species with 3910.87 g/mol mass corresponds to a molecule with only one complete d(A_{10}) branch grown from the symmetric branching unit. The intensity of this peak is about 4% of that of the full product peak.

The distribution of products for strands 3HA7H and 3HA10H, containing the 3-way branching trebler moiety, indicates the presence of many products that did not reach full length in synthesis or were not fully deprotected (**Figure A1.5** and **Figure A1.6**). The main species other than the full-length product appearing in the mass spectra for 3HA7H and 3HA10H are the full-length products still conjugated to the solid support linker (+ 262 g/mol), that failed to cleave off the molecule during deprotection. These products were present in significant proportion (~ 60% and ~30% of main product based on peak height in mass spectra for 3HA7H and 3HA10H

respectively). Truncated products were also observed for 3HA7H (two full branches out of three, ~16% compared to height of main peak product). Although many other lower mass products were present both for 3HA7H and 3HA10H, truncated products with one out of three full branches were not identified for either construct. Furthermore, for 3HA10H, the two-branch truncated product was not detected either. Due to time constraints and limited amounts of material, the trebler-containing strands were not purified further.

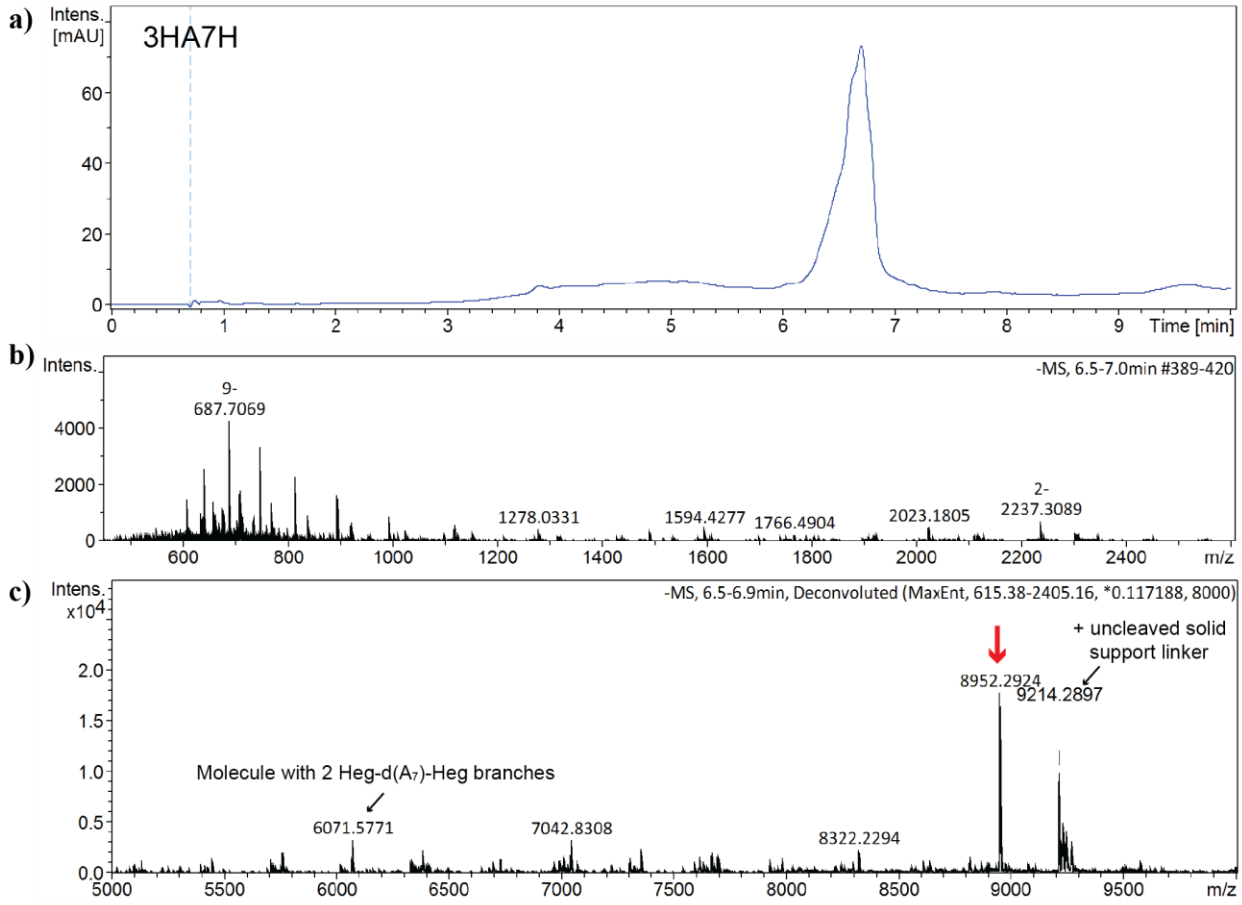


Figure A1.5| LC-ESI-MS analysis for 3HA7H. a) Chromatogram (UV-Vis absorbance at 260 nm). **b)** Non-deconvoluted mass spectrum. **c)** Deconvoluted mass spectrum (calculated molecular weight = 8953.57 g/mol, experimental mass from largest peak = 8952.29 g/mol). Peak for species with mass 9214.29 corresponds to the full product with an uncleaved solid support linker; it's intensity is about 60% of the main product peak. Peak for species with 6071.5771 g/mol mass corresponds to a molecule with only two complete d(A_7) branches grown from the trebler branching unit (and to peak with mass 687.71 g/mol and charge -9 in the spectrum in **(b)**); it's intensity is about 16% of the main product peak. Although many other side products are present, a species with only one complete d(A_7) branch (mass 3189.85 g/mol) could not be identified.

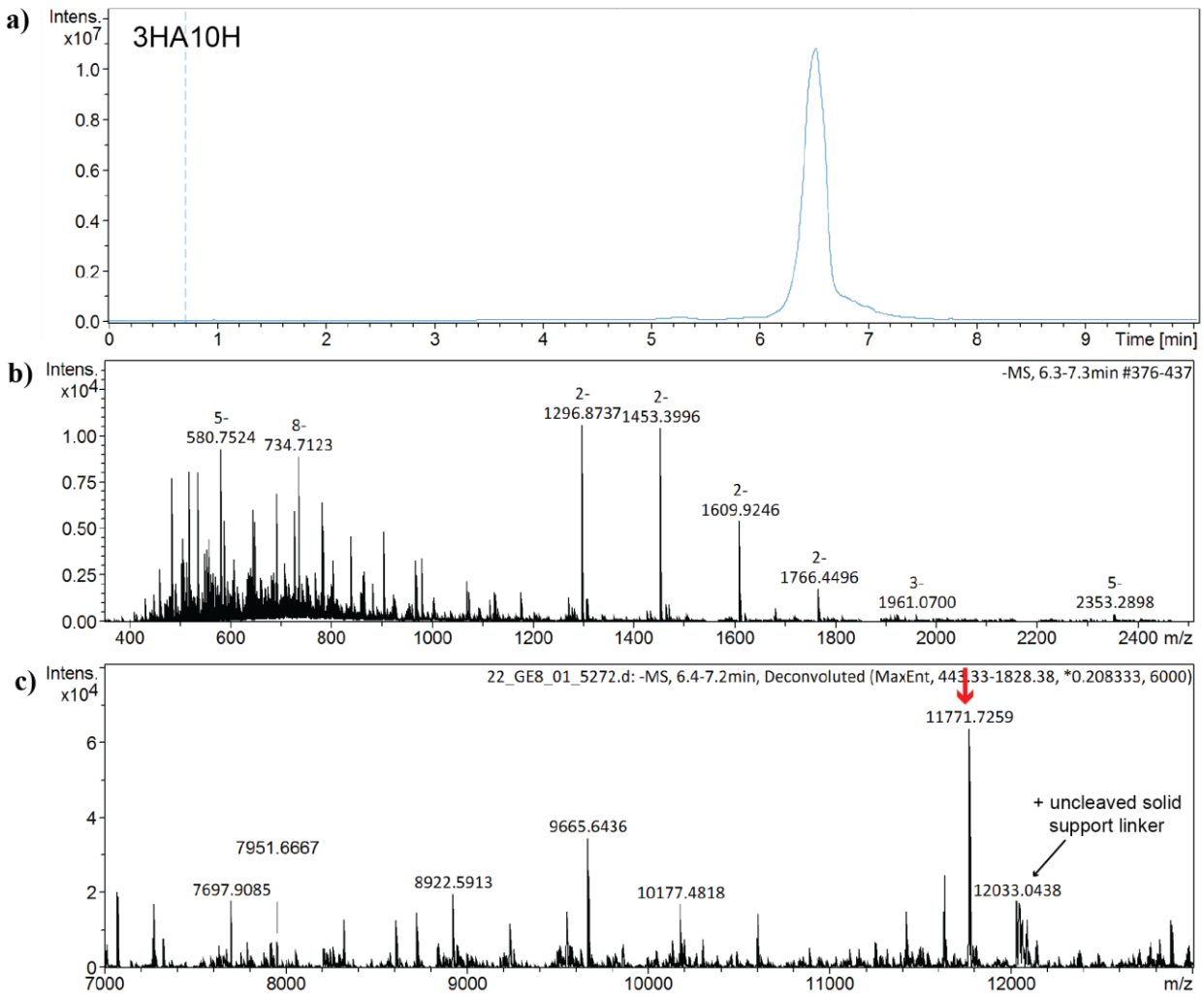
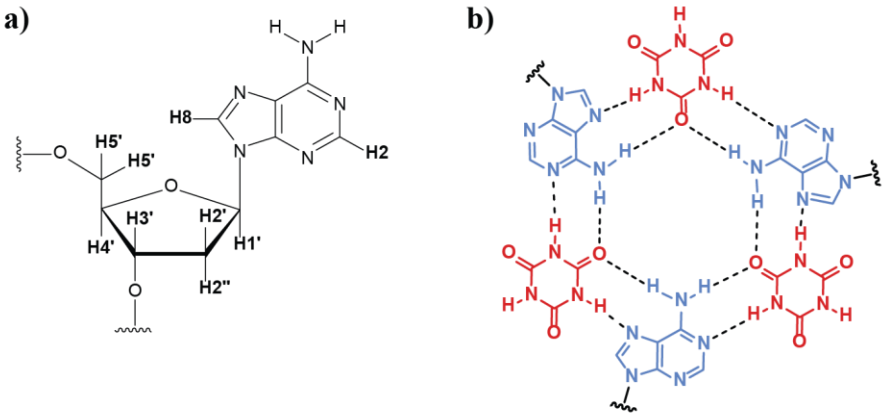


Figure A1.6| LC-ESI-MS analysis for 3HA10H. a) Chromatogram (UV-Vis absorbance at 260 nm). **b)** Non-deconvoluted mass spectrum. **c)** Mass spectrum (calculated molecular weight = 11772.46 g/mol, experimental mass from largest peak = 11771.75 g/mol). Peak for species with mass 12033.04 corresponds to the full product with an uncleaved solid support linker; it's intensity is about 30% of the main product peak. Although many other side products are present, species with only one or two complete d(A_{10}) branches (mass 4129.02 and 7947.85 g/mol respectively) could not be identified.

A1.2.2 NMR investigation of CA-mediated assembly of d(A_n) molecules

Both 1D and 2D experiments were performed to investigate d(A_n) structure upon assembly with CA. The labelling of deoxyadenosine protons is shown in **Scheme A1.1 a** for reference. According to the proposed hydrogen bonded 1:1 CA:A hexameric array structure (**Scheme A1.1**

b), CA is expected to interact with both hydrogen bonding faces of adenine and thus to influence the chemical shifts of adenine protons H2 and H8 in ^{1}H NMR.



Scheme A1.1| Adenosine and CA-mediated hexameric rosette structure. a) Labelling of deoxyadenosine protons as referred to in NMR spectra assignment. **b)** Proposed 1:1 CA:A hydrogen bonded hexameric array structure.

Preliminary DLS experiments comparing how CA7C, HA7H and 3HA7H respond to addition CA addition were performed, examining assembly of these strands with 10 mM CA *vs.* no added CA (**Table A1.2**). DLS showed that capping the d(A_7) strand with cytosine residues (CA7C) does not prevent long-range assembly from taking place. This is reflected in elevated light scattering count rate for CA7C with 10 mM CA, which then translates to a large apparent hydrodynamic diameter (D_H) compared to samples without CA addition. In contrast, there was no change in count rate for HA7H upon CA addition, and only a small increase for 3HA7H. Due to the low count rates, D_H calculations for these samples are prone to a large variability making the results unreliable (very large error from replicate measurements). However, the qualitative conclusion is that the extent of CA-mediated assembly is lower for HA7H and 3HA7H compared to CA7C.

The same samples examined by CD revealed new information (**Figure A1.7**). Whereas both CA7C and 3HA7H showed characteristic spectral changes upon CA addition indicating CA-mediated assembly (negative peak at 252 nm), the spectrum for HA7H remained unchanged. In this case, the hexaethyleneglycol (HEG) caps at either end of the strand seemed to prevent CA-mediated assembly.

	Count rate (kcps)		D_H (nm)	
	10 mM CA	**0 mM CA**	**10 mM CA**	**0 mM CA**
CA7C	44.5 ± 18.7	2.5 ± 0.5	303.0 ± 218.7	1.4 ± 0.5
HA7H	1.4 ± 0.2	1.3 ± 0.1	1.5 ± 4.5	0.3 ± 0.2
3HA7H	4.6 ± 0.2	2.6 ± 0.4	0.5 ± 1.1	0.6 ± 2.0

Table A1.2| Summary of count rate (kcps) and hydrodynamic diameter (D_H) measurements from DLS for CA7H, HA7H and 3HA7H. Averages obtained from measurements obtained at 5 °C (CL 95%, n = 3). Samples contained a total [A] of 100 µM.

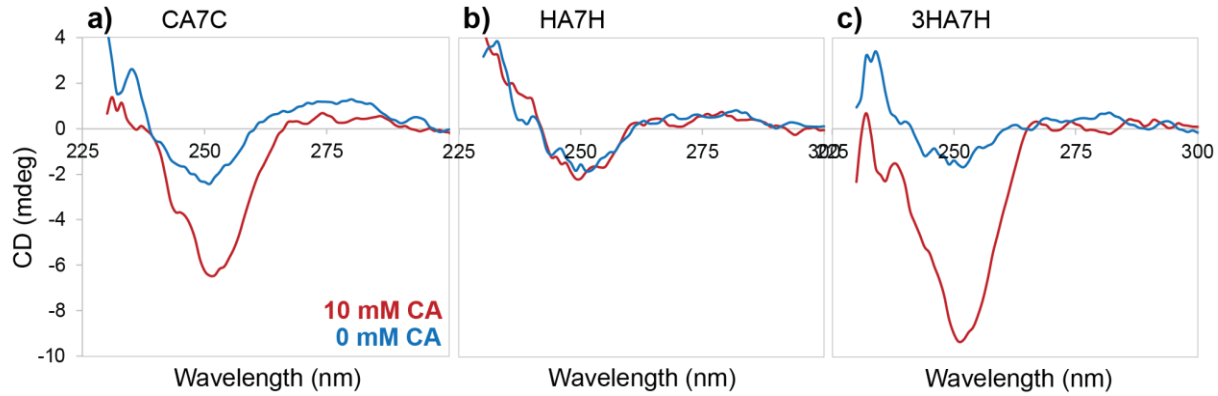


Figure A1.7| CD spectra for d(A_7) variants with 10 mM CA (red) *vs.* no added CA (blue). a) CA7C, **b)** HA7H, **c)** 3HA7H. Spectra collected at 5 °C (pH 4.5). Samples contained a total [A] of 100 µM.

Previous experiments with short d(A_n) strands indicated that they require higher minimal CA concentrations to ensure assembly. For example, no assembly was observed below 10 mM CA with d(A_7) strands. ^{1}H NMR spectra for HA7H were thus obtained with 15 mM CA, to potentially favor assembly with a higher CA concentration. However, ^{1}H NMR spectra for HA7H showed that the presence of 15 mM CA has no impact on the chemical shifts of adenine protons H2 and H8 (**Figure A1.8**), and the spectrum overlaps completely with the one obtained in the absence of CA. As slow kinetics of assembly for short strands have been observed in the past, incubation was extended in the hope of a different assembly outcome in the long term. Even after extensive incubation of HA7H with 15 mM CA (7 or 14 days at 4 °C), the spectrum remained unchanged, indicating that for the monomeric (unbranched) HA7H, the capping HEG groups at either end of the strand hinder CA-mediated assembly even at high [CA].

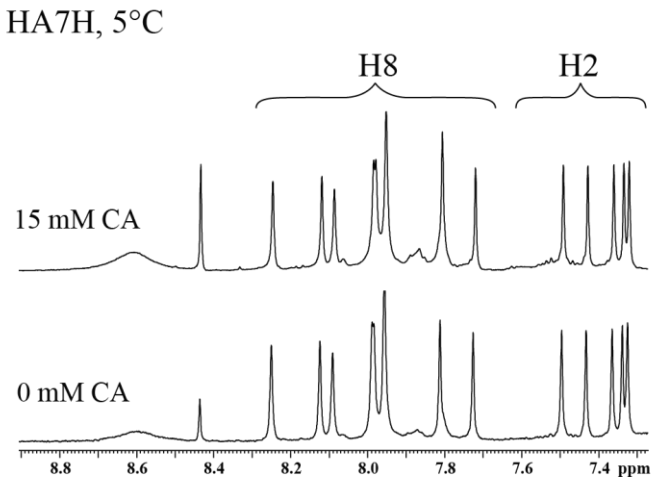


Figure A1.8| ^{1}H NMR spectra for HA7H with 15 mM CA (top) vs. 0 mM CA (bottom) at 5 °C (pH 4.5). Sample concentration 1 mM.

In contrast, for both branched strands with d(A_7) stretches, 2HA7H and 3HA7H, CD spectra acquired features characteristic of CA-mediated assembly (strong negative peaks at 212 and 252 nm) upon overnight incubation with 15 mM CA (**Figure A1.9 a**). The preorganization effect afforded by the branching units seemingly overcame the limitations that prevented assembly of monomeric HA7H to take place, and this despite also having capping HEG molecules as part of strand design for 2HA7H and 3HA7H. However, it is possible that the presence of truncated products (constructs with missing branches detected in LC-ESI-MS characterization) also contributed to the assemblies observed. In addition, the melting and annealing behaviour for these strands was different (**Figure A1.9 b**), with 2HA7H having a melting temperature (T_M) of 23 °C (4 °C higher than 3HA7H), with a narrower FWHM than 3HA7H, and an annealing temperature (T_A) of 10 °C (3 °C higher than 3HA7H), although with a wider FWHM in this case.

Unfortunately, the ^{1}H NMR spectra for 2HA7H with 15 mM CA indicated that the spectral signature observed by CD in the same conditions was in fact resulting from long-range CA-mediated assembly (**Figure A1.10**). At 5 °C, the intensity of NMR signals for adenosine protons dropped almost to zero, as had been observed in preliminary experiments with uncapped strands, pointing to the fact that most of the DNA material in the sample was now involved in large structures whose slow tumbling resulted in signal decay making them invisible to detection. Signal intensity was recovered by raising the temperature, causing assemblies to denature and release free

2HA7H strands back into solution. At 15 °C, the assemblies were partially denatured, and signal intensity continued to increase with a further temperature increase to 25 °C, a temperature where strands are mostly disassembled according to CD melting experiments (**Figure A1.9 b**). The changes in signal intensity were accompanied by small changes in the chemical shifts of the H8 and H2 adenosine protons.

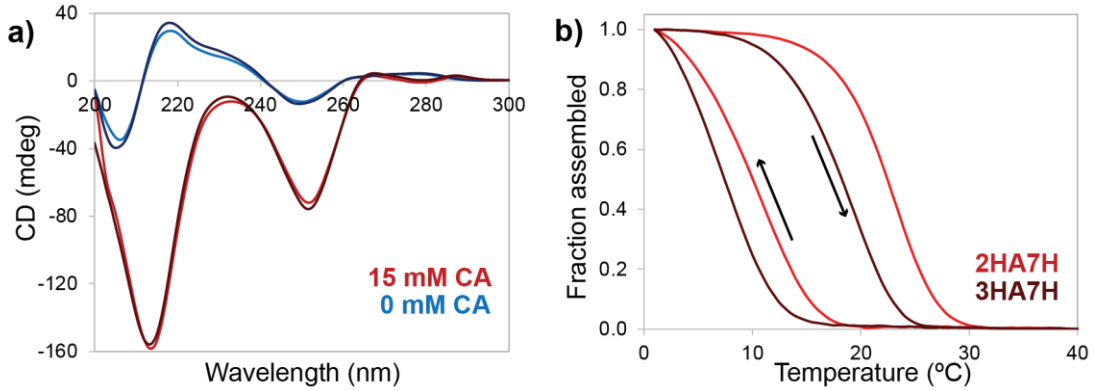


Figure A1.9| CA-mediated assembly of 2HA7H and 3HA7H studied by CD spectroscopy. a) CD spectra for 2HA7H and 3HA7H with 15 mM CA (red) or 0 mM CA (blue) at 5 °C. **b)** Thermal denaturation and reannealing curves for 2HA7H (bright red) and 3HA7H (burgundy) with 15 mM CA. Samples contained a [d(A_7)] of 100 μM.

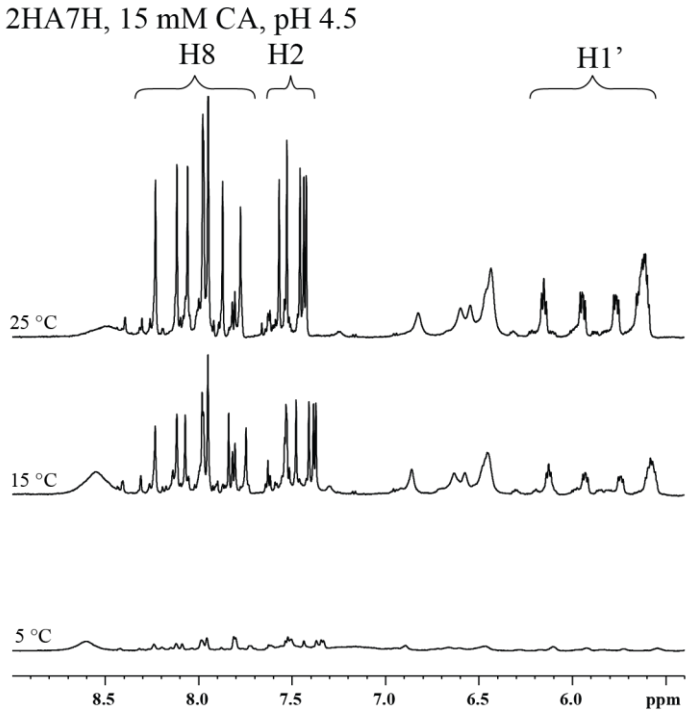


Figure A1.10| ^{1}H NMR spectra for the thermal denaturation of 2HA7H. 2HA7H with 15 mM CA (pH 4.5) at 5, 15 and 25 °C (bottom to top). Sample concertation 1 mM.

Proton assignment was carried out by using a combination of TOCSY and NOESY spectra acquired for HA7H and 2HA7H at pH 4.5 in the absence of CA at 5 °C. TOCSY is a 2D spectroscopic technique which displays cross peaks for protons correlated within a spin system. Signals in NOESY spectra on the other hand arise when protons are located in close proximity to each other. The relevant portions of the assigned NOESY spectra for HA7H and 2HA7H are presented in **Figure A1.11** and **Figure A1.12** respectively. Adenosine residue labeling in the assignment is from 1 to 7 in the 5' to 3' direction. Both intra-residue and sequential, inter-residue, assignments are shown. Sequential cross-peaks are reported here for correlations between H8 and the H1' or H2'/H2'' protons of the 5' neighbouring residue in the strand. The peak patterns for HA7H and 2HA7H are nearly identical, suggesting that d(A_7) strands preorganized via the symmetrical branching unit within HA7H do not take on a particular conformation that is different from monomeric HA7H. Within the grouping of H8 protons, the A7H8 proton is the most deshielded, followed by A1H8 and A2H8. A3H8 and A6H8 nearly overlap and are located more upfield, next to A4H8 and A5H8 which cannot be distinguished from each other due to overlap. The chemical shifts of the H1' protons follow a similar pattern related to residue position within the strand. It is possible that the protons closer to either end of the strand are more deshielded due to lower tendency for stacking with neighbouring residues, which is more likely to occur in the middle of the strand, for residues A4 and A5 for example, which in fact do have more shielded protons. The small changes in chemical shift seen for H8, H2 and H1' protons in **Figure A1.10** with increasing temperature are more pronounced for the upfield protons (in the middle of the strand), as the strand loses stacking between bases at elevated temperatures. In structure determination, the correct assignment of H2'/H2'' protons is important to determine sugar conformation.[7] Typically, when the base is in an *anti* configuration, the cross peak intensity between H1' and H2'' is stronger than for H1' and H2'. In contrast, cross peaks with H8 are stronger with H2' than H2''. This distinction was used to assign H2'/H2'' cross peaks with H1' and H8. Some weak cross peaks appear between H2 and H1' protons, or with exchangeable amino protons (NH_2), however they cannot be assigned with certainty. There is no indication for inter-strand contacts in the cross peaks observed in these spectra.

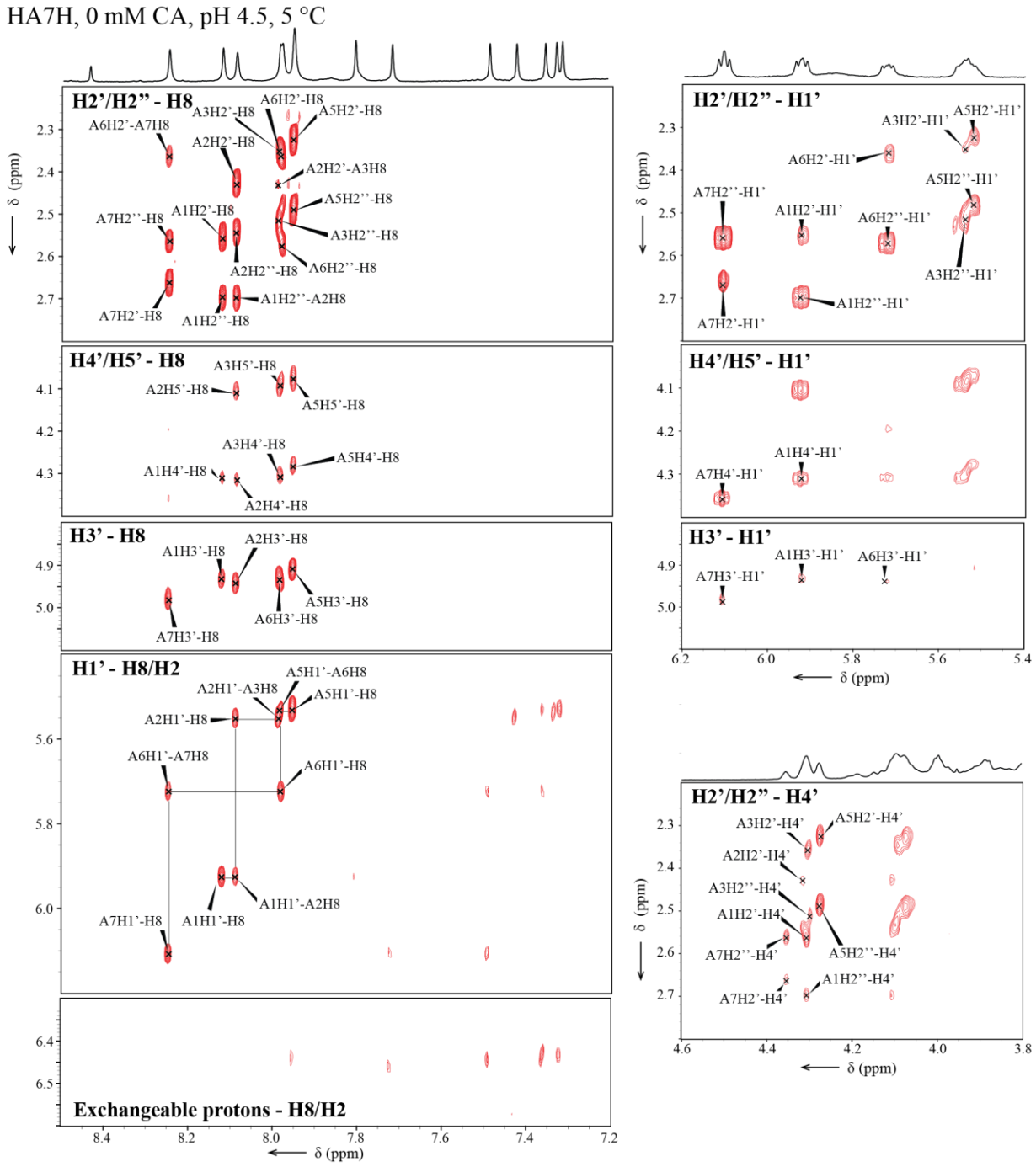


Figure A1.11| Assigned NOESY spectrum for HA7H at pH 4.5 and 5 °C (250 ms mixing time). Sample concentration 1 mM.

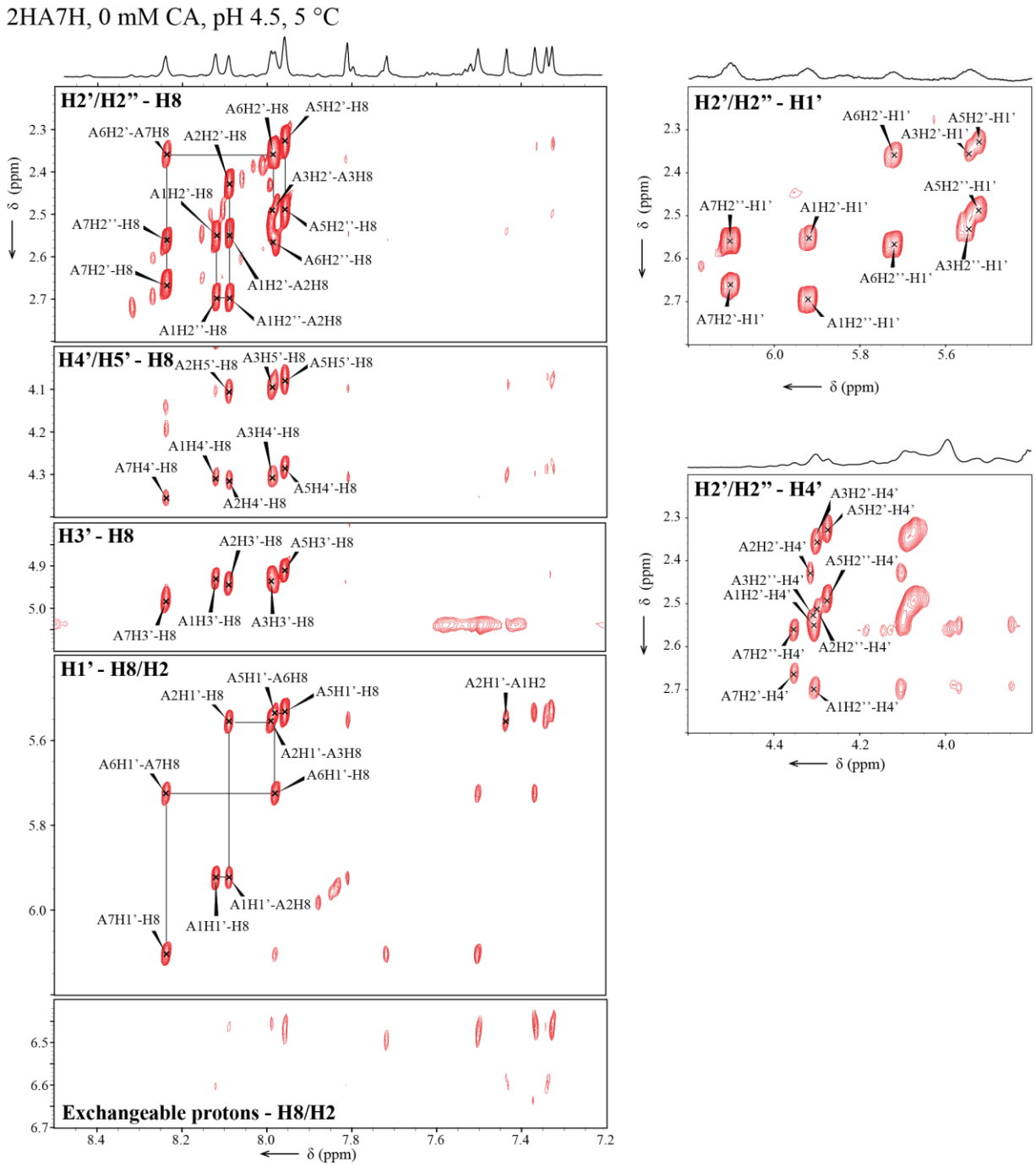


Figure A1.12| Assigned NOESY spectrum for 2HA7H at pH 4.5 and 5 °C (250 ms mixing time). Sample concentration 1 mM.

NOESY spectra were also obtained for 2HA7H with 15 mM CA at 5, 15 and 25 °C (**Figure A1.13**). The patterns of the chemical shifts remained largely the same for the cross peaks in the

H2'/H2''-H8 and H1'-H8 regions. The cross peaks in other regions were weak. As observed in 1D ^{1}H NMR spectra for 2HA7H (**Figure A1.10**), at 5 °C most of the signal decays due to the participation of strands in large assemblies. The 15 and 25 °C temperature conditions for additional NOESY experiments were selected to probe an intermediary temperature region during the transition from long-range CA-mediated assembly to monomeric material which occurs with melting. However, as mentioned above, the patterns of chemical shifts in these conditions remain largely unchanged. In addition, reduced mixing times were used for these experiments to potentially capture transient exchange of species from the assembled to monomeric state by looking for cross peaks close to the diagonal. However, this method was unsuccessful as relevant signals for exchangeable protons and shifts in H8 and H2 protons were not detected.

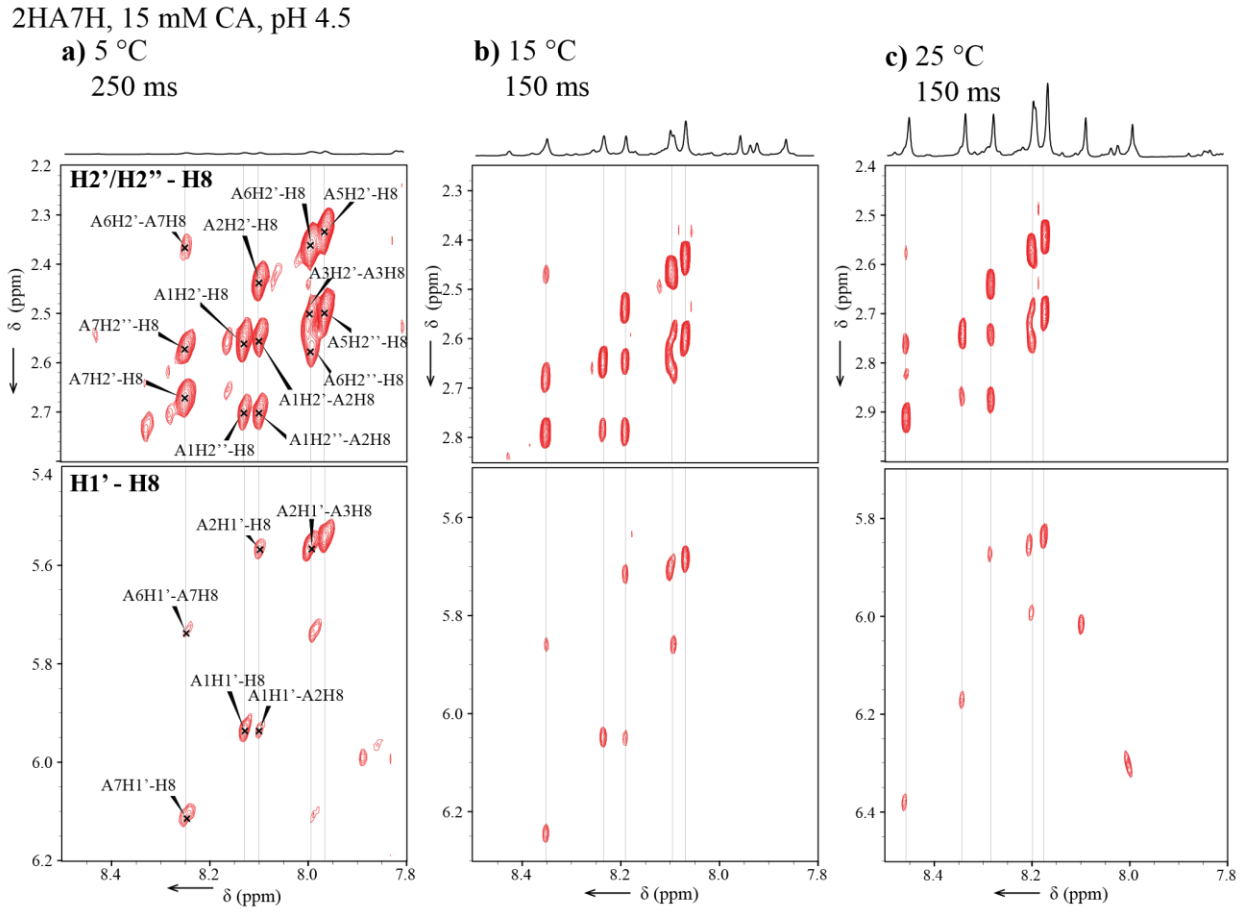


Figure A1.13| Temperature dependent NOESY spectra for 2HA7H with 15 mM CA at pH 4.5 and 5 °C. a) 5 °C (250 ms mixing time), b) 15 °C (150 ms mixing time) and c) 25 °C (150 ms mixing time). Sample concentration 1 mM. Note that chemical shifts do not match those observed in **Figure A1.10** because the NOESY spectra were not correctly referenced.

A1.2.3 Additional temperature dependent NMR experiments in the presence of CA

The experiments presented in the previous section revealed that, despite the design choices considering strand length, addition of capping groups and strand preorganization via branching molecules aiming to prevent aggregation following CA addition, the systems used did not meet an essential requirement for structure investigation by NMR by failing to prevent long-range assembly (considering that a small proportion of truncated products was present). The results in the following section were obtained concurrently, using branched strands that were not yet studied by NMR above (2HA10H, 3HA7H and 3HA10H). The 1D ^{1}H NMR spectra for trebler-containing strands 3HA7H and 3HA10H indicated the presence of side products to a larger extent than the monomeric HA7H and branched 2HA7H (although some small peaks associated with side products and impurities were also detected for the latter), most likely associated with the most abundant side product, a full-length product with an uncleaved solid support linker. The trebler-containing strands were not studied in detail with 2D experiments as they also showed aggregation (drop in signal in 1D NMR) with the addition of 10 mM CA. In addition, the NOESY spectrum for 2HA10H (data not shown) had a nearly identical pattern of cross peaks as 2HA7H for the most important regions (cross peaks between H1' and H8 as well as H2'/H2'' and H8), with more residues in the middle of the strand having overlapping signals, similarly to residues A4 and A5 in HA7H and 2HA7H.

As a result, these strands could only be used to make additional phenomenological observations related to temperature responsive assembly behaviour. The melting behaviour of 3HA7H and 3HA10H was compared using a series of ^{1}H NMR spectra at temperatures from 5 to 45 °C following incubation with 10 mM CA at 4 °C overnight (**Figure A1.14**). At 5 °C, signal intensity was low due to CA-mediated polymerization. As expected from previous studies, the aggregation of d(A_7) containing 3HA7H is less extensive than for d(A_{10}) containing 3HA10H, as strand length is known to influence extent of assembly at 10 mM CA. The known trends of strand length effect on melting behaviour are also on display as the transition from long range assembly to monomeric products detected by NMR (increase in signal intensity) occurs in different temperature ranges for 3HA7H and 3HA10H. For 3HA7H, the largest change in intensity is observed between 5 and 15 °C, whereas for 3HA10H, the transition continues until 25 °C, in line with previously observed trends, relating delay in thermal denaturation behaviour with increased strand length. The influence of side products present in trebler-containing samples is most apparent

in the H1' region where the peaks are much less well defined than for the monomeric or symmetrical branch-containing strands.

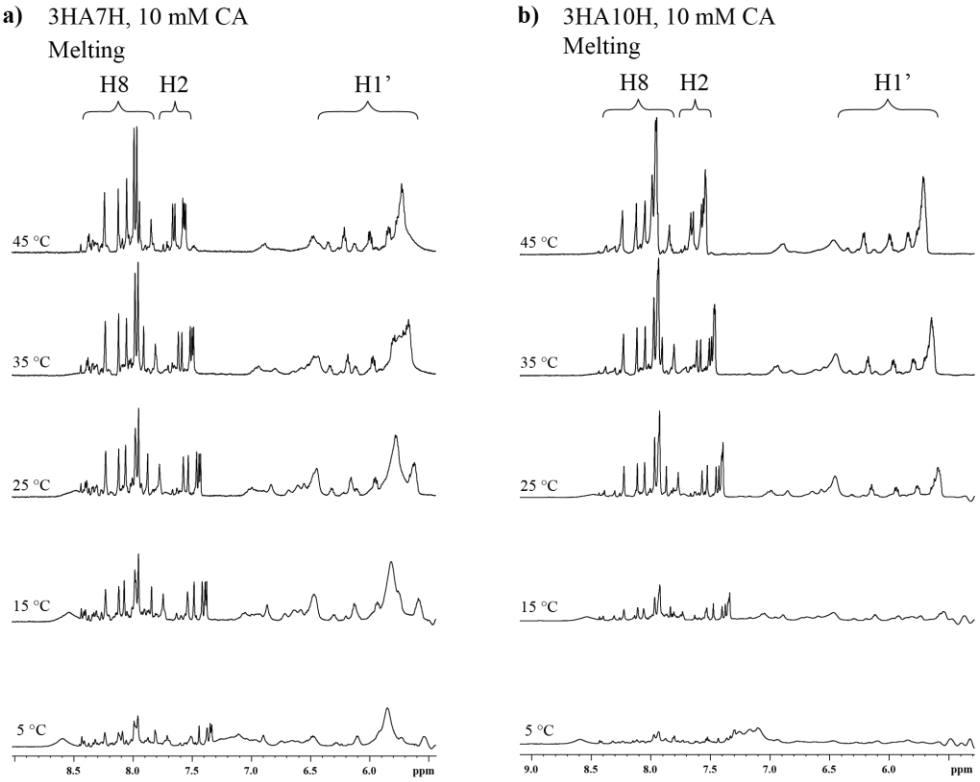


Figure A1.14| ^{1}H NMR spectra following the thermal denaturation of 3HA7H and 3HA10H with 10 mM CA. **a)** 3HA7H (pH 4.5) at 5 to 45 °C (bottom to top). **b)** 3HA10H (pH 4.5) at 5 to 45 °C (bottom to top). Sample concentration 0.5 mM.

Temperature dependent ^{1}H NMR spectra for 3HA10H with 10 mM CA also provide another illustration of the hysteresis observed between the annealing and melting behaviour of CA-mediated d(A_n) assembly (**Figure A1.15**). The difference in the processes is most apparent in spectra at 15 °C on annealing and on melting. There is a larger delay in recovering signal intensity (monomeric state) in melting, compared to the annealing process. The spectra in this case were acquired with long equilibration times at each temperature setting (30 minutes), such that the overall rate of temperature change was on average ~ 0.25 °C/min over the sampled temperature range. The hysteresis remains substantial even at this temperature change rate, mirroring earlier observations made with monomeric, uncapped d(A_{10}) strands in Chapter 3.

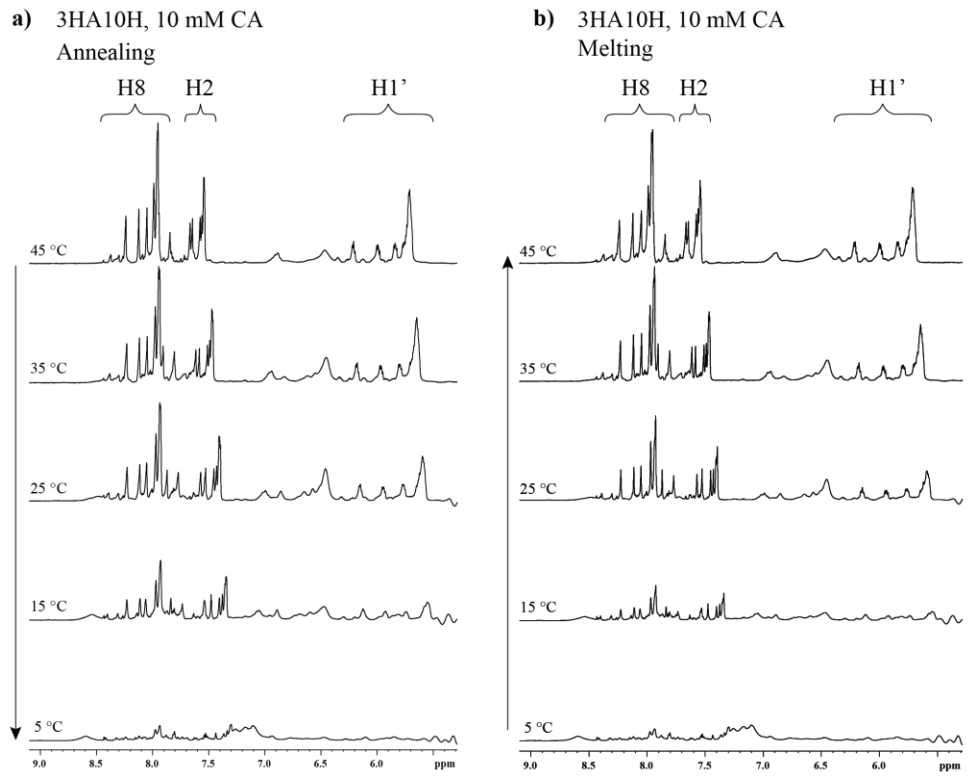


Figure A1.15| ^{1}H NMR spectra following annealing and melting with temperature of 3HA10H with 10 mM CA. **a)** Annealing (pH 4.5) at 5 to 45 °C (bottom to top). **b)** Melting (pH 4.5) at 5 to 45 °C (bottom to top). Sample concentration 0.5 mM. Arrows indicate direction of temperature change.

^{13}C labelled CA (CA-$^{13}C_3$) was also used to examine the assembly with 2HA10H using ^{13}C NMR. The ^{1}H spectra for 2HA10H with 15 mM CA-$^{13}C_3$ (**Figure A1.16 b**), once again show an increase in signal intensity with increasing temperature reflecting de-aggregation of assembled DNA material with melting. Similarly, the single peak appearing in the ^{13}C spectrum starts out at a low intensity at 5 °C, due to incorporation of much of the CA present in solution into higher-order assemblies with 2HA10H, and grows in intensity with increasing temperature, accompanied by a slight displacement to a higher chemical shift (**Figure A1.17 a**). The spectra for CA-$^{13}C_3$ alone show similar changes in chemical shift with temperature (**Figure A1.17 b**), suggesting that it is indeed the free CA remaining in solution which gives rise to the signal observed for the sample containing 2HA10H. Note that the splitting of the CA-$^{13}C_3$ peak seen in **Figure A1.17 b** likely results from the aging of the aqueous solution (CA can undergo keto-enol tautomerization in water, although the keto form dominates). In contrast, freshly dissolved CA-$^{13}C_3$ displays a single peak in the ^{13}C spectrum (**Figure A1.17 c**).

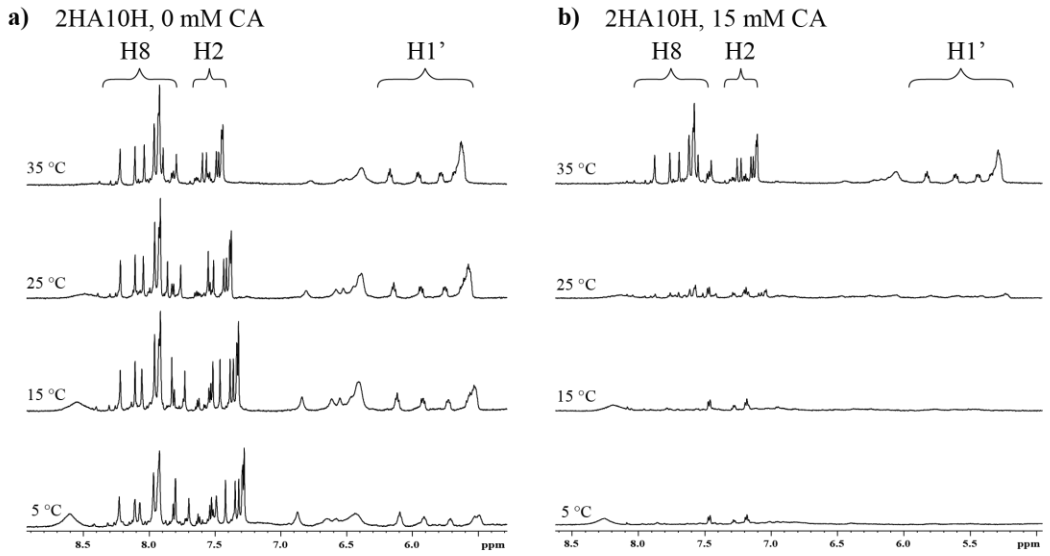


Figure A1.16| ^{1}H NMR spectra for the thermal denaturation of 2HA10H. a) 2HA10H without added CA (pH 4.5) at 5 to 35 °C (bottom to top). **b)** 2HA10H with 15 mM CA (pH 4.5) at 5 to 35 °C (bottom to top). Sample concentration 0.5 mM.

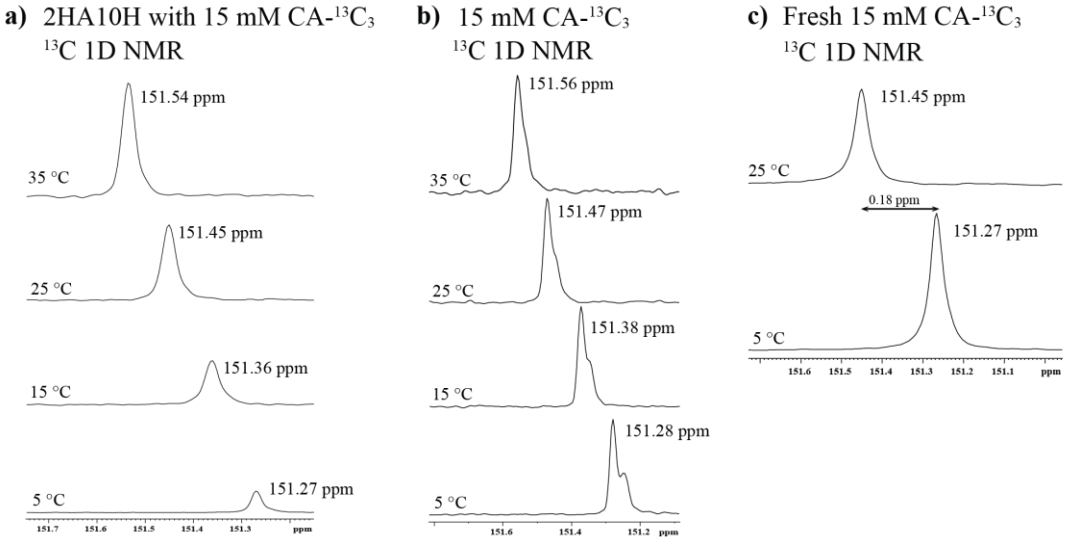


Figure A1.17 | ^{13}C NMR spectra for the thermal denaturation of 2HA10H with CA-$^{13}C_3$. a) 2HA10H with 15 mM CA-$^{13}C_3$ (pH 4.5) at 5 to 35 °C (bottom to top). **b)** 15 mM CA-$^{13}C_3$ (pH 4.5) at 5 to 35 °C (bottom to top). Sample concentration 0.5 mM. **c)** Freshly dissolved 15 mM CA-$^{13}C_3$ (pH 4.5) at 5 and 25 °C (bottom and top respectively).

A1.3 Conclusions

NMR spectroscopy is a powerful tool for solving the 3D structures of proteins and nucleic acids. However, this sensitive, technically complex method has multiple requirements in terms of

sample purity, high sample concentrations and size of the molecules under investigation, among others. In this appendix, we have aimed to harness the potential of NMR spectroscopy approaches to directly observe CA-mediated assembly of d(A_n), not merely to rely on indirect characterization methods to study this structure (Chapters 2 and 3). We report on a series of monomeric and branched d(A_n) molecules that were designed with the goal of ensuring that CA-mediated assembly takes place, but only on a short length scale, avoiding extensive polymerization that makes structures invisible to detection by NMR. The designs presented here did not achieve this goal, as CA addition led to aggregation reflected by a significant drop in d(A_n) signal for all systems showing characteristics of CA-mediated assembly. In the absence of CA, NOESY cross peaks for HA7H and 2HA7H strands could be assigned. Temperature-dependent experiments on branched strands illustrated melting and annealing behaviour trends observed previously with unmodified d(A_n) strands. In addition to ^{1}H NMR, ^{13}C NMR experiments were carried out using ^{13}C labelled CA and illustrated the incorporation of CA into the polymerized d(A_n) structures.

Future perspectives for explicit solving of the CA-mediated d(A_n) structure keep raising the same questions that were asked at the outset of this study, as the experiments reported here have not succeeded in addressing them. In order to use NMR for this purpose, strand designs should be rethought. Alternative capping groups and preorganization schemes, or, in an entirely different approach, the creation of a self-contained system that is capped by a branching unit at either end must be considered. Importantly, pure materials are required to ensure that observations are not related to the presence of truncated products and impurities. Alternatively, it is possible to pursue CA-mediated d(A_n) assembly crystallization, an unpredictable avenue that we have touched on, but not fully explored, that nonetheless could lead to a treasure trove of structural information.

A1.4 Experimental

A1.4.1 Materials

HPLC grade acetonitrile (MeCN), triethylamine (TEA), glacial acetic acid, trifluoroacetic acid (TFA), ammonium hydroxide (NH_4OH), sodium chloride (NaCl), magnesium chloride ($MgCl_2$), hydrochloric acid (HCl), sodium hydroxide (NaOH), sodium phosphate monobasic (NaH_2PO_4), sodium phosphate dibasic (Na_2HPO_4), cyanuric acid (CA) and cyanuric acid-$^{13}C_3$ (CA-$^{13}C_3$) were used as purchased from Sigma-Aldrich.

For in-house DNA synthesis, the reagents and materials listed below were used. Anhydrous MeCN, universal 1000 Å LCAA CPG pre-packed columns (loading densities of 25-40 μmol/g), activator (0.25 M ETT in MeCN), alternative activator (0.25M DCI in MeCN), oxidizer (0.02 M I_2 in THF/pyridine/H_2O), Cap A mix (THF/lutidine/Ac_2O), Cap B mix (16% MeIm in THF), and deblock solution (3% DCA/DCM) were all used as purchased from Glen Research. dA-CE and trebler phophoramidites were purchased from Glen Research. DMT-hexaethaloxy-glycol and symmetrical branching CED phosphoramidites were purchased from ChemGenes. Glen-Pak DNA purification columns were purchased from Glen Research. All phosphoramidites were dissolved in anhydrous MeCN at 0.1 M concentration for synthesis.

1xTBE (Tris-boric acid-EDTA) buffer was composed of 45 mM Tris, 45 mM boric acid and 2 mM EDTA at pH 8.3. 1xNaPiMg (sodium phosphate with added magnesium) buffer was composed 50 mM NaH_2PO_4 and 7.5 mM $MgCl_2$, pH was adjusted to 4.5 using HCl.

A1.4.2 Instrumentation

A Mermade MM6 DNA synthesizer from Bioautomation was used for in-house DNA synthesis. UV-Vis DNA quantification measurements were performed on a Nanodrop Lite spectrophotometer from Thermo Scientific. Liquid chromatography electrospray ioization mass spectrometry (LC-ESI-MS) was performed on a Dionex Ultimate 3000 system coupled to a Bruker MaXis Impact QTOF. Thermal annealing procedures were carried out using a heating black with adjustable temperature. DLS experiments were carried out on a Zetasizer Nano S90 from Malvern equipped with temperature control. CD experiments were performed on a Jasco J-810 spectropolarimeter equipped with a xenon lamp, a Pelletier temperature control unit and water circulation. NMR spectra were acquired on Bruker Avance spectrometers operating at 600 or 800 MHz.

A1.4.3 Synthesis of d(A_n) strands

Oligonucleotide synthesis was performed in-house on a 1 μmol scale, using standard automated solid-phase DNA synthesis on universal 1000 Å LCAA-CPG solid-supports. The coupling of the branching phosphoramidites was performed by combining equal volumes of

activator (0.25 M ETT in MeCN and 0.25M DCI in MeCN used for the symmetrical branch and the trebler coupling respectively) and phosphoramidite in anhydrous MeCN with an extended reaction time of 15 minutes under nitrogen atmosphere, followed by transfer of the column to the synthesizer for capping, oxidation and deblocking steps. In all syntheses, coupling efficiency was monitored after removal of the DMT 5'-OH protecting groups. The final DMT 5'-OH protecting group was not removed on the synthesizer to enable purification by oligonucleotide purification cartridges with reverse phase packing (Glen-Pak). Completed sequences were deprotected in a 28% NH_4OH solution for 16 hours in a 65 °C water bath, and purified by Glen-Pak cartridges following manufacturer-recommended procedure. Sequences for all modified strands are compiled in **Table A1.1**. Strand purity for modified strands was assessed by LC-ESI-MS analysis (**Figure A1.1** to **Figure A1.6**).

A1.4.4 Sample preparation and CD experiments

Samples for all studies were prepared by combining concentrated DNA strand stocks with CA (0, 10 or 15 mM) in 1xNaPiMg pH 4.5 buffer. Samples were heated to 50 °C for 5 minutes followed by cooling at room temperature for 30 minutes and incubation at 4 °C overnight.

For DLS studies, samples were filtered using 0.2 μm nylon syringe filters.

For NMR studies, samples contained 10% v/v D_2O and 1 μL of 5 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as a calibration standard.

CD spectra were obtained (200 or 230 nm to 350 nm range, 100 nm/min scan rate, 1 nm bandwidth, 2 accumulations) at 5 °C. The samples were then thermally denatured at a rate of 1 °C/min following the CD signal at 252 nm, at 0.5 °C intervals over the 1 to 40 °C range.

A1.4.5 NMR experiments

1D experiments were carried out at temperatures ranging from 5 to 45 °C using 5 or 10 °C intervals. Equilibration at each temperature setting was carried out for 5 minutes in most cases, except for the melting and annealing experiments with 3HA10H where a 30 minute equilibration time was used (slow annealing and melting conditions).

TOCSY spectra were recorded with standard MLEV17 spinlock sequence and with 80 ms mixing time. NOESY spectra in H_2O were acquired with mixing times ranging from 75 to 250 ms. All 2D experiments were performed in H_2O, where water suppression was achieved by including a WATERGATE[8] module in the pulse sequence prior to acquisition. The 2D experiments were carried out at temperatures ranging from 5 to 25 °C. All spectra were processed with Topspin software (Bruker), which was also used to generate 1D plots. The spectral analysis program Sparky[9] was used for assignment of the NOESY cross-peaks and 2D data plotting.

A1.5 References

1 Gehring, K., Leroy, J. L. & Gueron, M. A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* **363**, 561-564 (1993).

2 Chakraborty, S., Sharma, S., Maiti, P. K. & Krishnan, Y. The poly dA helix: a new structural motif for high performance DNA-based molecular switches. *Nucleic Acids Res.* **37**, 2810-2817 (2009).

3 Fedoroff, O. Y., Salazar, M., Han, H., Chemeris, V. V., Kerwin, S. M. & Hurley, L. H. NMR-based model of a telomerase-inhibiting compound bound to G-quadruplex DNA. *Biochemistry* **37**, 12367-12374 (1998).

4 Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A. & Feigon, J. Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3745-3749 (1993).

5 Wijmenga, S. S. & van Buuren, B. N. M. The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc.* **32**, 287-387 (1998).

6 Macomber, R. S. *A complete introduction to modern NMR spectroscopy*. (Wiley, 1998).

7 Wüthrich, K. *NMR of Proteins and Nucleic Acids*. (Wiley, 1986).

8 Piotto, M., Saudek, V. & Sklenář, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661-665 (1992).

9 Goddard, T. D. & Kneller, D. G. *SPARKY 3* (University of California, San Francisco).