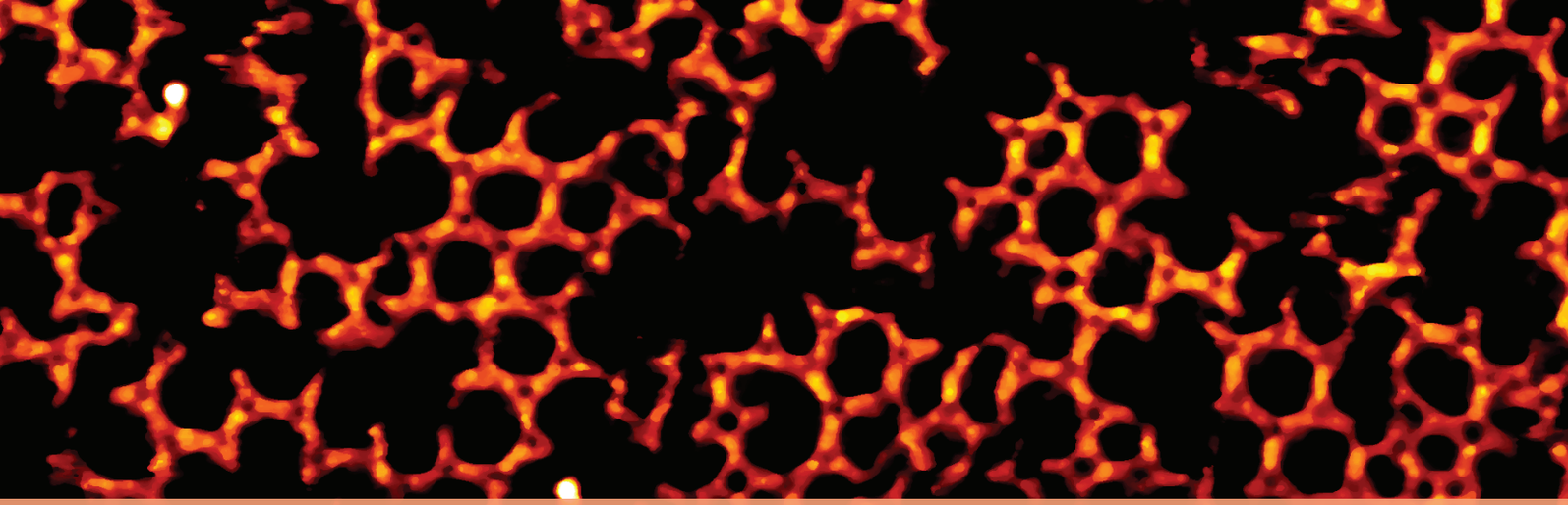
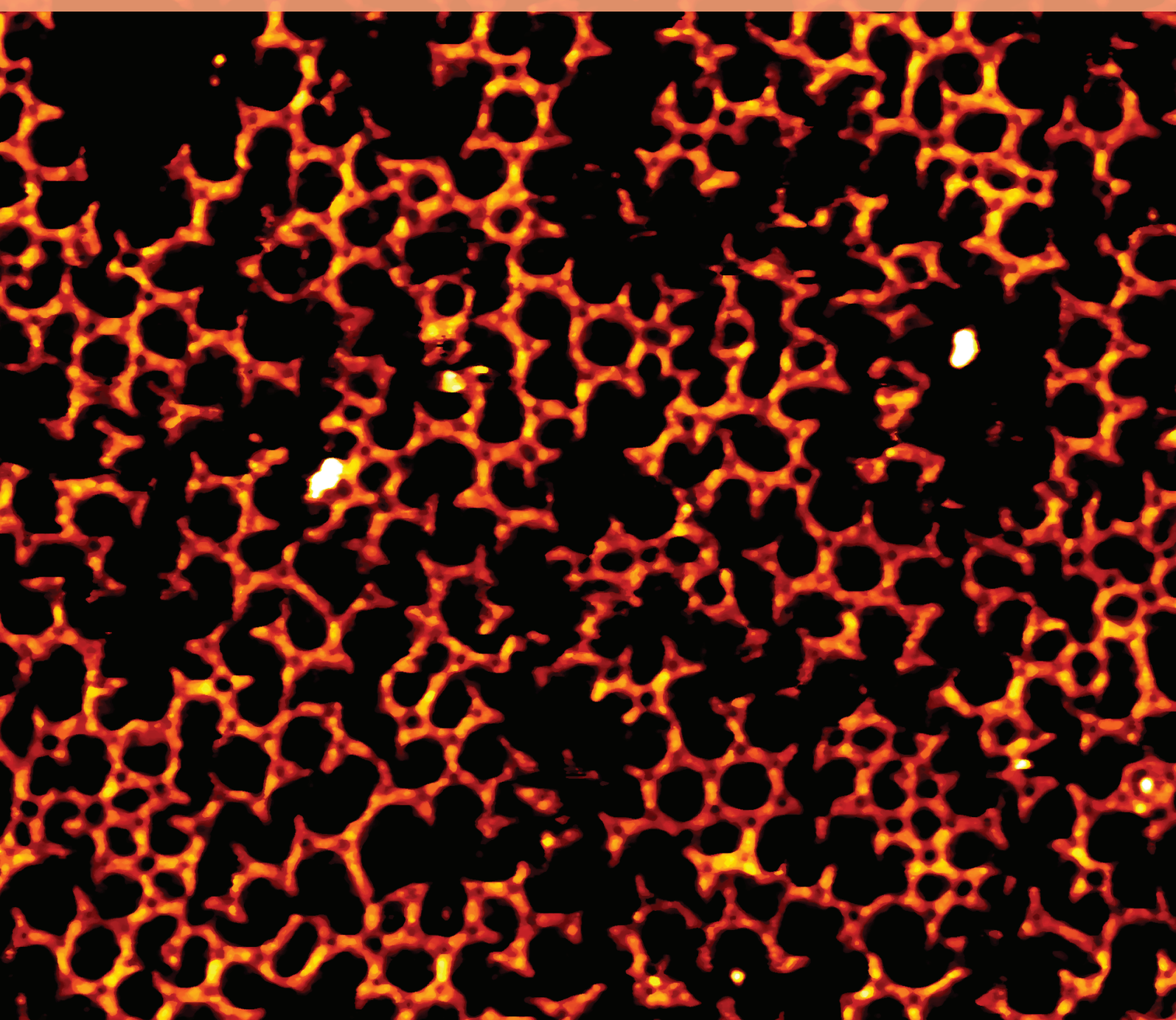


Nucleic acid nanostructures: From biomimetic assembly to *in vivo* applications

Aradhana Chopra



Technische Universität München
Fakultät für Physik
Physik synthetischer Biosysteme (E14)

Nucleic acid nanostructures: From biomimetic assembly to *in vivo* applications

Aradhana Chopra

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Martin Zacharias
Prüfer der Dissertation:

1. Prof. Dr. Friedrich C. Simmel
2. Prof. Dr. Tim Liedl

Die Dissertation wurde am 03.04.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 04.06.2018 angenommen.

Contents

Zusammenfassung **7**

Abstract **9**

List of Figures **11**

1 Introduction **15**

2 Fundamentals **19**

2.1 Nucleic acid structure 19

2.2 Properties of DNA and RNA 21

2.2.1 Polymeric properties 21

2.2.2 Chemical properties 22

2.2.3 Thermodynamic properties 24

2.3 Nucleic acid condensation 26

2.3.1 *In vivo* DNA condensation 26

2.3.2 *In vitro* DNA condensates 29

2.3.3 Physics of DNA condensation 31

2.4 Nucleic acid nanotechnology 33

2.4.1 DNA nanotechnology 34

2.4.2 RNA nanotechnology 38

3 DNA origami nanostructures folding using multivalent ions **45**

3.1 Introduction 45

3.2 Experimental details and methods 47

3.3 Results and discussion 49

3.3.1 Folding of DNA origami nanostructures using a trivalent condensing agent 49

3.3.2 Circular dichroism 51

3.3.3 Valence dependence and aggregation effects 51

3.4 Conclusion and outlook 55

4 Characterization and applications of Spd^{3+}-folded DNA origami structures **57**

4.1 Introduction 57

4.2 Experimental details and methods 58

4.3 Results and discussion 63

4.3.1 DNA nanostructures in electric fields 63

4.3.2 Modification and purification of Spd-nanostructures 66

4.3.3 Electrotransfection of DNA nanostructures in Jurkat cells . . 68

4.3.4 Cell viability assays 71

4.3.5 Stability of Spd-nanostructures in cell lysate 73

4.4 Conclusion and outlook 74

5 Towards 3D RNA nanostructures: design and synthesis **77**

5.1 Introduction 77

5.2 Experimental details and methods 79

5.3 Results and discussion 84

5.3.1 Design and folding of 3H-AE RNA tile 84

5.3.2 Modification of 3H-AE RNA tile with one aptamer 86

5.3.3 Modification of 3H-AE RNA tile with two aptamers 88

5.4 Conclusion and outlook 91

6 RNA scaffold production by rolling circle transcription (RCT) **93**

6.1 Introduction 93

6.2 Experimental details and methods 94

6.3 Results and discussion 96

6.3.1 RCT of PP7-MGA RNA aptamers 96

6.3.2 Confocal microscopy of transcription products 97

6.4 Conclusion and outlook 102

7 *In vivo* synthesis of catalytic DNA molecules **105**

7.1 Introduction 105

7.2 Experimental details and methods 107

7.3 Results and discussion 110

7.3.1 Design of MOSIC-DNAzyme and ss phagemid protocol . . . 110

7.3.2 Phagemid digestion 111

7.3.3 Catalytic activity of DNAzyme in DNA nanomotors 112

7.4 Conclusion and outlook 113

8 Conclusion and outlook **115**

References **117**

Publications **131**

Appendix **133**

Acknowledgement **161**

Zusammenfassung

Die natürliche Vielseitigkeit, vorhersehbare Basenpaarung und Biokompatibilität von Nukleinsäuren machen sie zu einem idealen Substrat für biotechnologische und nanotechnologische Anwendungen. Die Nucleinsäurenanotechnologie hat diese Merkmale zur Entwicklung von statischen sowie dynamischen Strukturen für Anwendungen in zahlreichen Forschungsbereichen genutzt.

In dieser Arbeit, wurden drei verschiedene Arten von Origami-Strukturen – ein flaches Blatt, eine hohle Röhre und ein kompakter Origami-Block in magnesiumfreiem Puffer, derniedrige (< 1 mM) Konzentrationen des natürlich vorkommenden Kondensationsmittels Spermidin (Spd) enthielt, gefalten. Vergleichbar mit natürlicher DNA Kondensation, muss auch hier die benötigte Menge an Spermidin proportional zur DNA Konzentration sein. Bei hohen Spd-Konezentrationen wurde Aggregation und das Ausfällen von Strukturen beobachtet. Im Gegensatz zu Origami-Strukturen, die in herkömmlichen Puffern gefaltet werden, sind Spd-Nanostrukturen beim Anlegen von hohen elektrischen Feldimpulsen, welche üblicherweise für Elektrotransfektionsexperimente verwendet werden, stabil. Es wurde weiter gezeigt, dass Säugetierzellen über Elektroporation mit Spd-stabilisierten Strukturen transfiziert werden können. Die Stabilität von Spd-Nanostrukturen wurde zudem in Zelllysat getestet.

Die Vielseitigkeit von RNA in ihrer Struktur und Modularität und ihre günstigen in vivo-Eigenschaften haben sie zu einem attraktiven Ersatz als Baumaterial auf der Nanoebene gemacht. In dieser Arbeit wurde ein antiparalleles drei Helix-RNA-Tile (3H-AE), basierend auf dem ssRNA-Origami-Ansatz, entworfen und konstruiert. Die Tiles interagieren mit den benachbarten Tiles über 120° Kissing Loops (KL) und formen dabei hexagonale Gitter. Die synthetische dreidimensionale (3D) RNA Nanostrukturen wurden von bestehenden Drei-Helix Tiles abgeleitet, wobei der interne Tetra Loop durch L-förmige RNA Motive, ersetzt wurde. Verschiedene Tiledesigns wurden kreiert. Ein oder zwei RNA Aptamere, die fluoreszierenden Moleküle binden, wurden senkrecht zur Tileebene positioniert. Im Falle von zwei Aptameren, wie hier bei einem Malachitgrün- (MGA) und Spinat- Aptamer, wurde sichergestellt, dass diese senkrecht zu Tileebene stehen und in gegensätzliche Richtung zeigen. Das Fluoreszenzsignal beider Aptamere erlaubte eine Echtzeitüberwachung der kotranskriptionalen Faltung der Tiles. Die

Aptamere an den entsprechenden Positionen können leicht durch andere Aptamere ersetzt werden. Darüber hinaus wurde die Rolling-Circle-Transkription (RCT) als alternative Strategie für die Synthese langer RNA-Wiederholungen untersucht und könnte möglicherweise als RNA-Scaffold ausgenutzt werden, um das Clustering von Enzymen zu untersuchen. Verglichen zur Erschaffung einzelner Strukturen wird das Co-clustering zur Anhäufung von Enzymen als bessere Strategie gesehen, um Reaktionsgeschwindigkeiten zu verbessern. wurde als günstigere Strategie angesehen. Als Grundsatzbeweis wurde ein RNA Scaffold, der aus einer Kombination von proteinbindenden Aptameren und Aptameren für fluoreszierende Moleküle besteht, konzipiert. Dieser wurde mittels RCT synthetisiert und seine Bindeeigenschaften wurden charakterisiert. Transkriptionale Aggregation und die Wirkung von Pyrophosphatase auf RNA Aggregate wurde ebenfalls diskutiert.

Ferner wurde die in vivo-Synthese einer in vitro selektierten RNA-spaltenden katalytischen DNA, des 10-23 DNAzym durchgeführt. Die katalytische Fähigkeit, welche in Zukunft als Komponente eines autonomen DNA-Motors nutzbar sein kann, wurden demonstriert.

Abstract

The natural versatility and predictable base-pairing of nucleic acids make them an ideal substrate for bioengineering and nanotechnological applications. Nucleic acid nanotechnology has exploited these features to develop static as well as dynamic structures for applications in numerous research areas.

In this work, three different types of origami structures - a flat sheet, a hollow tube, and a compact origami block were assembled in magnesium-free buffer solutions containing low (< 1 mM) concentrations of the naturally occurring condensing agent spermidine (Spd^{3+}). Similar to DNA condensation, the amount of spermidine required for origami folding was found to be proportional to the DNA concentration. At excessive amounts, aggregation and precipitation of structures was observed. In contrast to origami structures formed in conventional buffers, the resulting Spd-nanostructures were stable in the presence of high electric field pulses, that are commonly used for electrotransfection experiments. It was further shown that the spermidine stabilized structures could be delivered into mammalian cells via electroporation. The stability of Spd-nanostructures was also tested in cell lysate.

The adaptability of RNA in structure, modularity and conducive *in vivo* attributes have made it an attractive substitute as a building material. In this thesis, a novel three helix "antiparallel even" RNA tile (3H-AE) based on the ssRNA origami approach was designed and constructed. The tiles interacted with the neighbouring tiles via 120° kissing loops (KL) and assembled into hexagonal lattices. Synthetic three-dimensional (3D) RNA nanostructures were derived from the existing three helix tiles by the introduction of L-shaped RNA motifs with a 90° bend in place of the interior tetra loops. Different tile designs were generated with either one or two fluorescent RNA aptamers, malachite green aptamer (MGA) and spinach aptamer positioned perpendicular to the plane of the tile, extending in opposite directions. The fluorescence signal from the aptamers allowed real-time monitoring of cotranscriptional folding of the tile. The aptamer positions could be easily replaced by other aptamers of choice. In addition, rolling circle transcription (RCT) was explored as an alternate strategy for the synthesis of long repeats of RNA that could potentially be exploited as RNA scaffolds to study clustering of enzymes. Co-clustering of enzymes into agglomerates has been considered as a more favorable strategy to enhance the overall reaction flux rather than creating individual struc-

tures. As a proof of principle, an RNA scaffold encoding a combination of protein binding aptamer and fluorescent aptamer i.e. PP7 aptamer and MGA (PP7-MGA) respectively was synthesized by RCT and its binding to the target molecules was investigated. Transcriptional aggregation and the effect of pyrophosphatase on the RNA aggregates were also discussed.

Further, *in vivo* synthesis of an *in vitro* selected RNA cleaving catalytic DNA molecule, 10-23 DNAzyme was shown. Also, its catalytic abilities as a component of an autonomous DNA motor were demonstrated.

List of Figures

1 Components of nucleic acid 21

2 Denaturation and renaturation of DNA 24

3 DNA condensation in eukaryotes 27

4 Chemical structures and synthesis pathway of polyamines 28

5 Applications of polyamines in biology 28

6 Different morphologies observed during *in vitro* DNA condensation . 30

7 Foundations of nucleic acid nanotechnology 34

8 Representative examples of structural DNA nanotechnology 36

9 Examples of dynamic DNA nanotechnology 38

10 Naturally occurring RNA motifs. 39

11 Strategies for construction of RNA nanostructures. 41

12 Methods for production of RNA nanostructures. 42

13 Folding of DNA origami nanostructures in condensing agent spermidine (Spd^{3+}). 50

14 TEM micrographs of DNA origami nanostructures of different shapes and lattice types in Spd^{3+}. 50

15 Circular dichroism of Mg and Spd-nanostructures. 51

16 Relation between DNA origami and folding buffer concentration. . . 52

17 Folding reactions and corresponding TEM micrographs of 42HB at different concentrations of Spd^{3+}. 53

18 Applications of Spd-nanostructures. 56

19 Effect of different electric field strengths on the stability of 42HB folded in different buffers. 64

20 Stability of Mg/Spd nanostructures in electric fields favourable for electrotransfection in mammalian cells. 65

21 Purification of Atto647N labeled Spd-nanostructures (42HB variant) by size exclusion chromatography (SEC). 67

22 Schematic illustration of electrotransfer experiment. 68

23 Electrotransfection of Spd-nanostructures in Jurkat cells. 69

24 Electrotransfer of Mg-nanostructures in Jurkat cells. 70

25 Electroporation of with Atto647N labeled ssDNA strands in Jurkat cells. 70

26 CyQUANT cell viability assay. 72

27 PrestoBlue cell viability assay. 73

28 Stability of Spd-nanostructures (42HB) in Jurkat cell lysate. 74

29 Classical and modern view of RNA. 77

30 Schematic representation of the steps involved in design of ssRNA origami. 81

31 Schematic representation of steps involved in design and synthesis of ssRNA origami. 84

32 Folding of designed 3H-AE RNA tile. 85

33 AFM images of hexagonal lattice formed by 3H-AE RNA tiles. . . . 86

34 Modification of 3H-AE RNA tile with MGA. 87

35 Comparison of 3H-AE-MGapt V1 and V2. 88

36 Modification of 3H-AE-MG-apt V1 RNA tile with spinach aptamer. 89

37 Real-time fluorescence of DFHBI and MG during transcription of different RNA constructs. 90

38 Potential applications of synthetic RNA nanostructures. 92

39 RCT of PP7-MGA RNA aptamers. 97

40 Confocal micrographs showing co-localization of YPet and MG on different RNA scaffolds. 98

41 Transcriptional aggregates of PP7-MGA RCT RNA 99

42 Co-localization of transcription reactants and products in RNA aggregates. 100

43 Effect of PPiase on transcription using linear and circularized DNA templates (PP7-MGA). 101

44 RNA based clustering of multiple enzymes. 103

45 Design and synthesis of 10-23 DNAzyme using ss phagemid strategy. 111

46 Verification and digestion of ss phagemid containing 10-23 DNAzyme. 112

47 Catalytic reaction of IDT and MOSIC-motors. 113

48 Cadnano diagram of 42 helix bundle DNA nanostructures used in Spd^{3+}-folding and electrotransfection experiments. 133

49 Cadnano diagram of tube nanostructures used in Spd^{3+} mediated folding. 134

50 Cadnano diagram of rectangular DNA origami (tCRRO) used in Spd^{3+} mediated folding. 134

51 Screening of Spd^{3+} concentrations required for folding 50 nM DNA origami nanostructures (Tube and tCRRO). 135

52 Negative-stain TEM micrographs of 42HB folded in traditional buffer containing $MgCl_2$ and Spd^{3+}. 136

53 Negative-stain TEM micrographs of tube folded in traditional buffer containing $MgCl_2$ and Spd^{3+}. 137

54 Positive-stain TEM micrographs of tCRRO folded in traditional buffer containing $MgCl_2$ and Spd^{3+}. 138

55 Stability of Mg-nanostructures (42HB) subjected to a range of electric field strengths. 139

56 Stability of Spd-nanostructures (42HB) subjected to a range of electric field strengths. 140

57 Effect of different electric field strengths on the stability of Mg and Spd-nanostructures in high resistance sample buffers and smaller path length cuvettes (0.1 cm). 141

58 Fluorescent labeling of DNA nanostructures via hybridization. . . . 142

59 Different purification strategies used for Spd-nanostructures purification. . . . 143

60 Electrotransfection of Spd-nanostructures in Jurkat cells. . . . 144

61 Spd-nanostructures at different time points and z volumes in a single Jurkat cell. . . . 145

62 Lipofectamine-medated transfection of Spd-nanostructures in Jurkat cells. . . . 146

63 Verification of PCR products for different RNA tiles. . . . 147

64 Effect of EDTA on transcriptional aggregates of RCT products from PP7-MGA. . . . 147

65 Steps involved in synthesis of 10-23 DNAzyme by MOSIC method. . 148

66 Verification of strand E (or 10-23 DNAzyme) after electroelution. . . 148

67 42HB staple list. . . . 150

68 42HB staple list contd. . . . 151

69 Tube staple list. . . . 152

70 Tube staple list contd. . . . 153

71 tCRRO staple list. . . . 154

72 M13mp18 phage ssDNA scaffold variants. . . . 155

73 Thermal annealing ramp for folding. . . . 156

74 Primer sequences for amplifcation of DNA template. . . . 156

75 DNA gBlock sequences. . . . 157

76 RNA tile sequences. . . . 157

77 PCR reaction and conditions for DNA gBlocks. . . . 158

78 DNA sequences for RCT. . . . 158

79 DNA sequences for MOSIC-DNAzyme and DNA nanomotor. . . . 159

1 Introduction

The cell is the basic structural and functional unit of all known living organisms. One of the most important cellular components is the genetic material composed of nucleic acids – deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) that are responsible for transmission, expression and conservation of genetic information. It is this genetic material which allows the cell to replicate independently, making it the smallest unit of life. The discovery of the structure of DNA has opened a new era in biology, genetics, medicine, biotechnology, and biochemistry and provided further basis for studying and manipulating the biological processes [1, 2]. With novel biological functions and properties being discovered for both DNA and RNA, they are no longer perceived just as molecules for storage and transmission of genetic information. Rather, they have been applied in totally different contexts [3].

Nucleic acids offer the advantages of predictable base-pairing and information carrying features, which distinguishes them from other biomolecules and man-made polymers. This has led to the field of nucleic acid nanotechnology, which started out with using DNA as a primary building material, but over the years RNA has also joined the band wagon [4, 5]. Both these molecules have their own advantages and disadvantages. The greater stability of DNA over RNA makes it the material of choice in some situations. Whereas in others, the demand of specific features exhibited by RNA make it the desired material. They have been utilized both as building blocks to create nanoscale architectures and as versatile programmable scaffolds for the assembly of nanomaterials. Alongside the traditional WC-base pairing interactions, dynamic nucleic acid devices and machines that are responsive to different stimuli have also been exploited [6]. Also, several techniques such as SELEX (Systematic evolution of ligands by exponential enrichment) [7], which are capable of selecting nucleic acids from a plethora of molecules have laid the basis for applications of nucleic acids in various disciplines of science and technology.

The flexibility and predictable base-pairing of nucleic acids make them an ideal substrate for bioengineering and nanotechnological applications [6]. In this thesis, different aspects of nucleic acid nanostructures ranging from their biomimetic assembly and transfer to cells to design of novel synthetic nanostructures to their *in vivo* synthesis – have been studied. The potential applications of such structures in the field of nucleic acid nanotechnology and other related fields are also discussed.

Chapter 2 provides an overview of nucleic acids (structure and properties) and the field of nucleic acid nanotechnology. It highlights the major developments in the fields of DNA and RNA nanotechnology. Further, it illustrates the phenomenon of nucleic acid condensation with special emphasis on polyamine mediated condensation in an *in vitro* and *in vivo* context.

In chapters 3 and 4, concepts of DNA condensation (*in vitro* and *in vivo*) have been applied to DNA nanotechnology. Chapter 3 focusses on assembly and characterization of artificial DNA nanostructures, produced by the DNA origami technique, using the naturally occurring condensing agent– spermidine (Spd^{3+}) at its physiological concentration. This helps in circumventing the demands of non native buffer conditions for viable nanostructures. Further, the similarities between the folding process and DNA condensation have been addressed.

Chapter 4 provides a systematic characterization of the stability of DNA origami nanostructures folded in different buffers to a diverse set of electric field strengths. Further, it exploits the stability of the Spd-folded DNA nanostructures in the presence of an electric field for electrotransfection experiments in mammalian cells. The viability of cells in electric fields and uptake of DNA nanostructures are also discussed.

Chapter 5 and 6 concentrate on the design and synthesis of novel synthetic RNA based nanostructures and scaffolds. Chapter 5 aims at the design and synthesis of artificial three-dimensional (3D) RNA nanostructures assembled from a single RNA strand (ssRNA origami technique). It utilizes the exceptional folding capabilities and unique structural features of naturally existing RNA.

Chapter 6 exploits an alternate transcription technique – rolling circle transcription (RCT) for synthesis of long continuous multimeric RNA. Such RNA scaffolds could potentially be used for co-localization of multiple enzymes for effective substrate channeling and enzyme clustering effects. As a proof of concept, RNA encoding a protein binding domain (PP7 RNA aptamer) that binds to viral coat protein (PCP) part of the PCP-YPet fusion protein and a fluorescent aptamer (malachite green aptamer) have been synthesized by RCT. In addition, transcriptional aggregation and its causes and effects are discussed.

Chapter 7 describes *in vivo* production of an *in vitro* selected and synthesized catalytic DNA molecule called 10-23 DNAzyme. Further, its functioning as a DNA

nanomotor component and potential applications are discussed.

In chapter 8, an overall conclusion from the results obtained in different projects is provided. It includes the design, characterization and potential applications of nanostructures made from DNA and RNA. It not only provides insights to the questions asked here but also opens up new avenues for the applications of nucleic acid based nanostructures which could be further explored.

2 Fundamentals

This chapter provides an overview of nucleic acid structure and properties (sections 2.1 and 2.2). Further, it describes the process of *in vivo* nucleic acid condensation, *in vitro* studies carried out to understand such processes and their potential applications (section 2.3). In addition, it presents an overview of nucleic acid nanotechnology and its relevance (section 2.4).

2.1 Nucleic acid structure

Nucleic acids are biopolymers composed of repeating monomer units, called nucleotides. Each nucleotide is made of three components: a 5-carbon sugar, a phosphate group and a nitrogenous base (figure 1(A)). Nucleotides without the phosphate group is called a nucleoside. The nitrogenous bases are derivatives of purines and pyrimidines. Adenine (A), guanine (G) are the major purines; and cytosine (C), thymine (T) and uracil (U) are the major pyrimidine derivatives found in nucleic acids (figure 1(B)). DNA and RNA differ in the nature of pentose sugar - deoxyribose and ribose in DNA and RNA respectively; and presence of nitrogenous base U (in RNA) instead of T (in DNA). However, some methylated, hydroxymethylated and glucosylated forms of the major bases also exist in both DNA and RNA.

Nucleic acids can exist as double-stranded (ds) or single-stranded (ss) molecules. DNA is mostly present as a double-stranded molecule while RNA is single stranded. The nucleotides of both DNA and RNA are covalently linked by a phosphodiester linkage where $5'$ phosphate group of one nucleotide is connected to $3'$ hydroxyl group of the next nucleotide, thereby creating an alternating phosphate sugar backbone. The strands are antiparallel with their $5'$-$3'$ phosphodiester bonds running in opposite directions. They are not identical, instead are complementary to one another and interact via hydrogen bonding. The most common hydrogen-bonding patterns that predominate in dsDNA and RNA are those defined by Watson and Crick, in which A bonds specifically to T (or U) and G bonds to C (figure 1(C)) [1]. This specificity of hybridization has been the foundation of many molecular biology techniques including the polymerase chain reaction (PCR), DNA sequencing, southern blotting, and emerging applications such as microarray technology and in nucleic acid nanotechnology [2, 6]. The offset pairing of the two strands creates a major

groove and minor groove on the surface of the duplex. The presence of grooves make the nucleic acids accessible for interactions with proteins and other molecules.

Nucleic acids exist in different conformations depending on the properties of the helix and the conformation of sugar molecules present. Most commonly studied forms of DNA include A, B and Z forms with B form being the most stable and commonly occurring form of DNA in nature (table 1, figure 1(A)) and 1(D))). The deoxyribose sugar in DNA may alternate between 2′-endo and 3′-endo conformations, allowing DNA to switch between B and A form under appropriate conditions. Reversible transition of B to A form has been observed in prokaryotes in response to desiccation [8]. In addition, sequences forming Z DNA near promoter regions of actively transcribing genes in mammalian nuclei have been detected. It has been speculated that Z DNA is involved in gene regulation [9]. A form is the most preva-

Table 1: Comparison of A, B and Z forms of DNA [10].

	A form	B form	Z form
Helical sense	Right handed	Right handed	Left handed
Diameter	∼26Å	∼20Å	∼18Å
Base pair per helical turn	11	10.5	12
Helix rise per base pair	2.6Å	3.4Å	3.7Å
Base tilt normal to helix axis	20°	6°	7°
Sugar pucker conformation	C-3' endo	C-2' endo	C-2' endo and C-3' endo

lent form in RNA (figure 1(D)). Naturally occurring B form RNA has not yet been observed. It is speculated that the presence of 2′ OH locks the ribose into a 3′-endo chair conformation, eliminating the possibility of a stable B form. However, dinucleotides of RNA with B form conformation have been synthesized and crystallized [11]. RNA-DNA hybrids adopt an A form conformation. Other forms of nucleic acids also exist such as C-DNA, H-DNA (triple stranded DNAs via Hoogsteen base pairing) and G-tetraplexes.

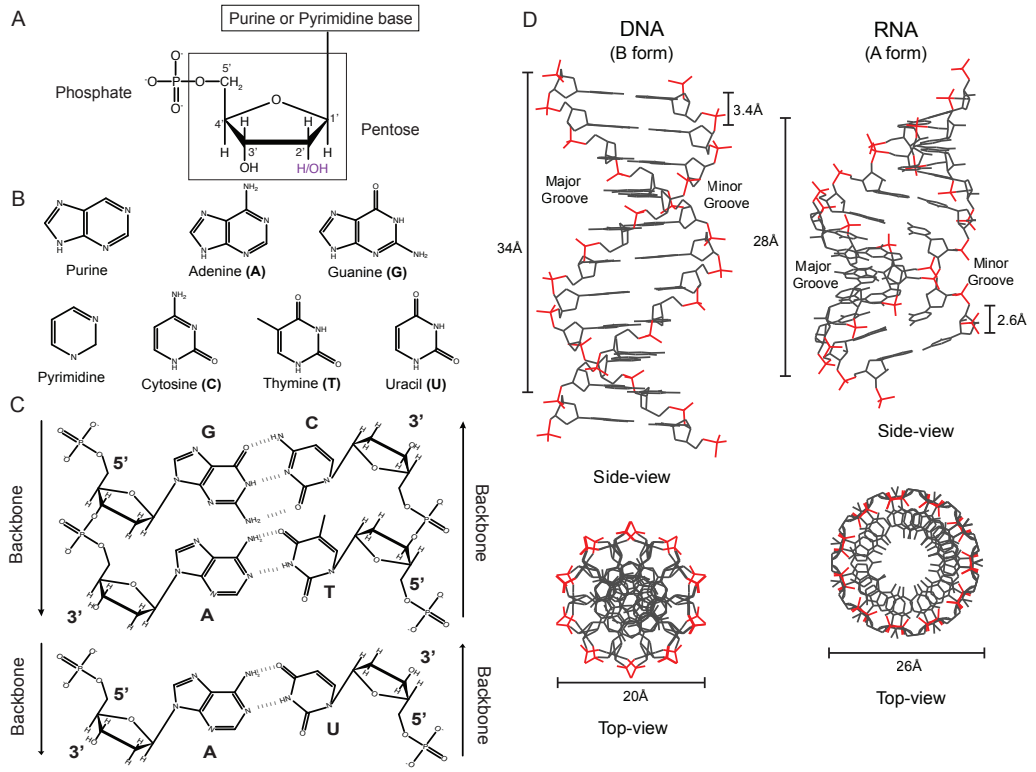


Figure 1: Components of DNA and RNA. (A) Basic structure of a nucleotide. The 2′ carbon of the pentose sugar can either have a H or an OH group (shown in purple), resulting in a deoxyribonucleotide or ribonucleotide respectively. (B) Nitrogenous bases and their origin. (C) Watson-crick (WC) base-pairing between adenine (A) and thymine (T) in case of DNA / uracil (U) in case of RNA stabilized by two hydrogen bonds and cytosine (C) and guanine (G) stabilized by three hydrogen bonds. (D) Side and top view of different forms of double helix (left: Right handed B form for DNA, Right: Right handed A form for RNA). Colour code for atoms: grey = C,H and N, red = O and P. The structure of B and A forms of double helix were created by make-na server, `http://structure.usc.edu/make-na/server.html` which uses nucleic acid builder (NAB) software [12]

2.2 Properties of DNA and RNA

2.2.1 Polymeric properties

Nucleic acids are amongst one of the stiffest known polymers. The measure of stiffness of a polymer is referred to as its persistence length, which characterizes the length over which the tangent vector to the polymer axis becomes uncorrelated. The persistence length of DNA and RNA duplexes are found to be ≈ 45-50 nm and 60-70 nm respectively [13] making RNA duplexes stiffer than DNA duplexes. The

persistence lengths of ssDNA and RNA are found to be in the range of 10-60 Å measured under different sets of conditions. Persistence length of DNA and RNA depends on a number of factors including ionic concentration, temperature etc [14].

2.2.2 Chemical properties

The overall structure and properties of nucleic acids are dependent on the properties of the component nucleotides. The alternating phosphate sugar backbone in nucleic acids is hydrophilic and faces the surrounding water. Nitrogenous bases are hydrophobic, aromatic molecules present towards the interior of the double helix and interact with the complementary base on the antiparallel strand via hydrogen bonding. On the same strand, the bases interact via hydrophobic stacking where they are posititioned with the planes of their aromatic rings almost parallel to each other and perpendicular to the helix. This also involves a combination of van der Waals and dipole-dipole interactions between the bases. Base stacking helps to minimize contact of the bases with water, and plays a significant role in stabilizing the structure of nucleic acids.

The negative charge of the phosphodiester bridge in the nucleic acid backbone repels nucleophilic species such as hydroxide ion (OH^-) making it much less susceptible to hydrolytic attack in comparison to other esters. This resistance is crucial for maintaining the integrity of information stored in nucleic acids. The presence of the $2'$ OH group in ribose confers special properties to RNA. On one hand, it locks the ribose sugar into a $3'$endo chair conformation, making it structurally favorable for the RNA double helix to adopt the A form; but on the other hand, it makes the RNA more suceptible to base-catalyzed hydrolysis and nuclease digestion in contrast to the $2'$ H group in DNA. Thus, the greater stability of DNA probably accounts for its evolution as the hereditary material in most cells and in many viruses.

Nucleic acids exhibit a characteristic absorption at wavelengths near 260 nm, because of their nitrogenous bases that absorb UV light. **Hypochromic effect** describes the decrease in the absorbance of dsDNA (or RNA) in comparison to its ss counterpart whereas **hyperchromic effect** describes the opposite effect, an increase in the absorbance of ssDNA (or RNA). Such effects are observed during the process of denaturation and renaturation.

Denaturation and renaturation. Even though the stability of ds-structure of nucleic acids is critically important, strand separation and annealing of complementary sequences are essential for a number of *in vivo* processes like DNA replication, DNA repair and transcription. *In vitro*, nucleic acids can undergo denaturation (separation) and renaturation (hybridization/annealing) cycles, in which the two strands of the double helix are separated and hybridized respectively under appropriate conditions (figure 2). The disruption of hydrogen bonding and base pairing interactions cause unwinding of the double helix to form single strands, completely or partially separated from each other. However, no covalent bonds are broken. The separation occurs over a temperature range and the midpoint is called the melting temperature (Tm). Tm is defined as the temperature at which half of the DNA strands are in the random coil or single-stranded state. It is influenced by the base composition, length of sequence, ionic concentration and presence of chemical reagents [15]. Nucleic acids with a higher GC content have a higher Tm than those with a higher AT content because of differential number of hydrogen bonds. Tm is an index of thermal stability, but the main factor responsible for the stability of a nucleic acid structure is base stacking.

Renaturation involves the formation of non-covalent, sequence specific interactions between two or more complementary strands of nucleic acids into a single most energetically preferred complex often referred to as a duplex. The process of denaturation and renaturation have been applied to a variety of biotechnological techiques such as PCR, hybridization probe based assays etc [2].

Duplexes of DNA, RNA and RNA-DNA hybrids (one RNA strand and one DNA strand) exhibit different thermal stabilities. RNA duplexes are more stable than DNA duplexes *in vitro*. At neutral pH, denaturation of a RNA duplex often requires temperatures that are 20K above those required for denaturation of a DNA duplex with a similar base composition [16]. The stability of an RNA-DNA hybrid generally lies in between the stabilities of RNA and DNA duplexes. This trend is primarily attributed to the A form versus B form conformational differences.

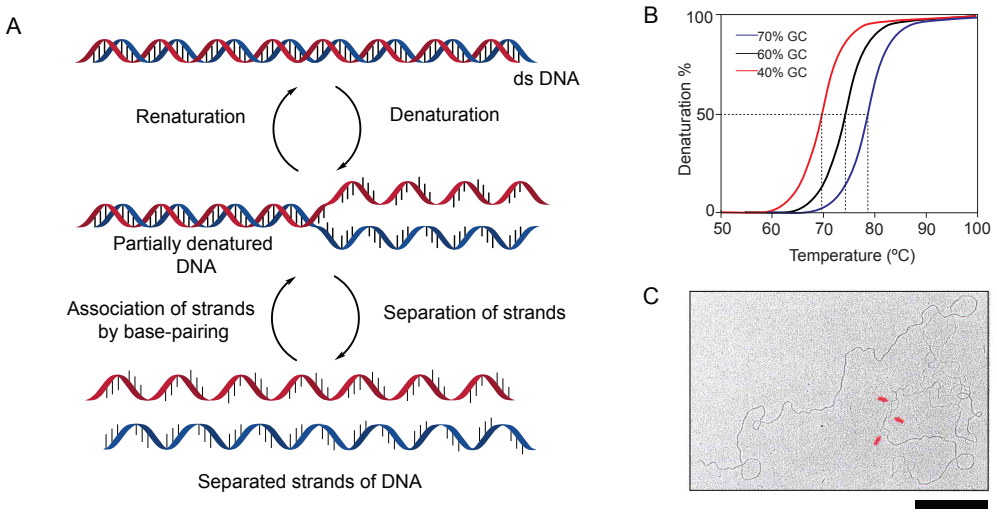


Figure 2: Denaturation and renaturation of DNA. (A) Denaturation and renaturation cycles under appropriate conditions. (B) Graph showing the relation of Tm with GC content in DNA (50 bp). Tm was calculated by Olioanalzyer tool from IDT [17]. (C) Transmission electron microscope (TEM) images of denaturing ds DNA [10].

2.2.3 Thermodynamic properties

Thermodynamic properties of nucleic acids play a very important role in understanding the biological function of nucleic acids such as DNA replication, mutation, repair and transcription [18]. They also contribute to the design of nucleic acid sequences for a variety of molecular biology techniques including PCR, cDNA expression profiling, genomic diagnostics [2, 19]. In addition, thermodynamic parameters are involved in secondary structure prediction and hybridization of nucleic acids. This is significant where the folding potential of nucleic acids is being exploited such as in *in vitro* SELEX, for generation of nucleic acid based catalysts and in areas of nucleic acid nanotechnology and computing [6, 7].

A number of factors including base pairing, base stacking interactions and entropic effects contribute to the total hybridization energy of a duplex. In general, the standard Gibbs free energy change (for one mole) of nucleic acid hybridization, ΔG° at a given temperature T can be calculated using the following equation:

$$\Delta G^\circ(T) = \Delta H^\circ - T\Delta S^\circ \tag{1}$$

where ΔH° is the standard enthalpy change, the thermodynamic measure of the total energy and ΔS° is the standard entropy change, a measure of the possible

states a system can assume. The ΔG° can be calculated for any temperature once ΔH° and ΔS° are known. The values of entropy, enthalpy and free energy based on the nearest neighbour (NN) model for a duplex formation have been generated. These parameters include influence of salts and correction factors for ends of a duplex strand, for the initiation of the hybridization and symmetry correction of a DNA sequence, which accounts for self dimerization [15, 20].

The total free energy can be calculated with [20]:

$$\Delta G^\circ_{total}(T) = \Delta G^\circ_{init}(T) + \Delta G^\circ_{sym}(T) + \sum \Delta G^\circ_{stack}(T) + \Delta G^\circ_{terminal} AT(T) \quad (2)$$

A database of nearest neighbour thermodynamic parameters for WC base pairs, internal mismatches, terminal mismatches, terminal dangling ends, coaxial stacking and loops including hairpins, bulges, internal loops, and multibranched loops has been accumulated for both DNA and RNA [20–22]. It has been incorporated into a number of softwares including Mfold and Vienna RNA package [23, 24].

When the thermodynamic parameters of a given nucleic acid sequence are known, the Tm can be predicted and vice-versa. The ΔH° and ΔS° can be calculated using the Van't Hoff's plot between ln(K) and 1/Tm. The slope and intercept of the graph give the ΔH° and ΔS° values respectively. K can also be substituted by the ratio of the concentration of products and reactants.

The thermodynamics of hybridization for a short nucleic acid sequence can be accurately described as a two-state process. In this approximation, the possibility of intermediate partial binding states in the formation of a double strand state from two single stranded oligonucleotides is neglected. As a result, there are two single strands, A and B form forming a duplex, AB.

$$AB \longleftrightarrow A + B \quad (3)$$

The equilbrium constant $K = \frac{[A][B]}{[AB]}$ According to Van't Hoff's equation, the relation between free energy and equilibrium constant is given by:

$$\Delta G^\circ = -RTlnK = -RTln\frac{[A][B]}{[AB]} \quad (4)$$

where R is the ideal gas law constant and T is the temperature. At T=Tm, $[A] = [B] = [AB] = \frac{[AB]_{initial}}{2}$ i.e. half the duplex has been dissociated.

$$Tm = -\frac{\Delta G^\circ}{Rln\frac{[AB]_{initial}}{2}} \quad (5)$$

ΔG° can be substituted by $\Delta H^\circ - T\Delta S^\circ$, then the above equation for Tm becomes:

$$Tm = \frac{\Delta H^\circ}{(\Delta S^\circ - \frac{[AB]_{initial}}{2})} \tag{6}$$

The terms ΔH° and ΔS° are given for association reaction, then the above equation becomes:

$$Tm = \frac{\Delta H^\circ}{\Delta S^\circ + Rln(\frac{[A]_{total}}{2} - \frac{[B]_{total}}{2})}, where [B]_{total} < [A]_{total}. \tag{7}$$

2.3 Nucleic acid condensation

2.3.1 *In vivo* DNA condensation

Nucleic acid condensation plays a crucial role in most biological systems, from viruses to higher more complicated eukaryotic organisms. It caters to the need of compaction of the genetic material within the small confines of the cell nucleus, sperm heads, the bacterial cytoplasm, or in viral capsids. In addition, it provides control and regulation over numerous biological processes making it essential for the sustainability and adaptability of the organisms [25]. Typical lengths of a single biological DNA molecule vary from a few hundred micrometers in T4 phage to metres in a human cell, or longer in other organisms [26]. Condensation of DNA *in vivo* is facilitated by different proteins such as histone proteins, protamines and polyamines. In eukaryotes, DNA is condensed mainly by histone proteins which form an octameric protein core. The DNA is wound around the octamer and is called the nucleosome core particle. A number of nucleosomes are connected via linker DNA and histone proteins (not observed in yeast) and it appears like "beads on a string" which further compacts into highly condensed chromosome structure with a series of intermediate packaging steps as shown in figure 3. In bacteria, DNA compaction and organization is found to occur at different length scales and is carried out by various proteins and other molecules such as integration host factor (IHF), sequence specific proteins SlmA, SeqA and MatP; and non coding RNA (ncRNA) etc [31]. Viral DNA is packaged by neutralization of negative charges with positively charged domains of capsid protein, ssDNA/RNA binding proteins, or with polyamines. Several studies have demonstrated that both herpes simplex viruses (HSV) and poxviruses package high concentrations of polyamines, sufficient to neutralize more than 40% of the negative charge on the DNA [32].

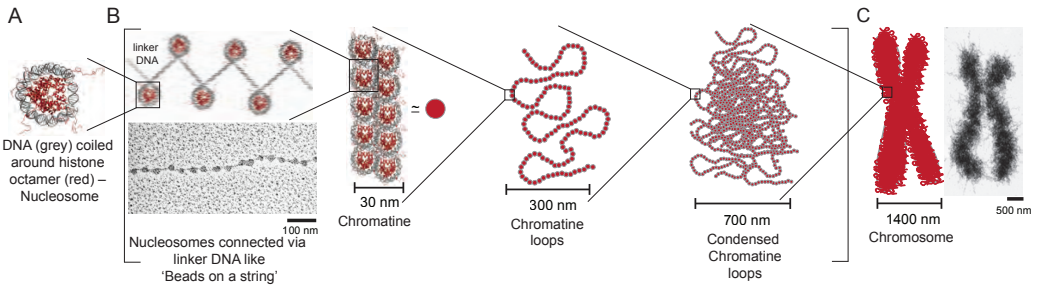


Figure 3: DNA condensation in eukaryotes. (A) Nucleosome core particle structure: dsDNA (grey) wound around octamer of histone proteins (red). (B) Intermediate structural packaging steps from "beads on a string" [27] to the condensed chromatin. (C) Highy condensed chromosome structure and corresponding TEM micrograph [27]. X-Ray structure of the nucleosome core particle [28], item 1KX5 from RCSB PDB (www.rcsb.org) [29] was created with UCSF Chimera [30]. (B) and (C) are adaptations of (A).

Alongside proteins, polyamines are known to play a significant role in condensation and packaging of the genetic material irrespective of cell type. Polyamines are multivalent molecules consisting of linear flexible carbon chains with amino groups that are fully protonated under physiological pH, and thus electrostatically interact with negatively charged biomolecules especially nucleic acids, proteins, and phospholipids. In eukaryotes, the most prevalent biogenic polyamines include putrescine, spermidine and spermine. Bacteria and viruses on the other hand have a more diverse repertoire of polyamines including spermidine, homospermidine, norspermidine, putrescine, cadaverine, 1,3-diaminopropane and agmatine etc [32]. Most of them are synthesized at intermediate steps via a single synthetic pathway as shown in figure 4 (B) or are modifications of the synthesized polyamines. The intracellular concentration of different polyamines has been found to vary with cell type and state. Most of the polyamines are present in a bound state and the concentration of free polyamines show fluctuations in response to intracellular signals [33]. For instance, typical concentration of spermidine range from 0.1-3 mM in mammalian cells, $\sim$ 7 mM in *Neurospora crassa* and 4-6 mM in bacteria [34, 35].

These compounds play an important role in the function of nucleic acids *in vivo* because of their nucleic acid binding and condensing abilities. Immunocytochemical studies have shown spermidine and spermine to be associated with mitotic chromosomes and involved in stabilization of the chromatin structure during cell cycle [32]. Besides DNA condensation, polyamines have been involved in diverse biological processes including replication, transcription, translation, post-translational

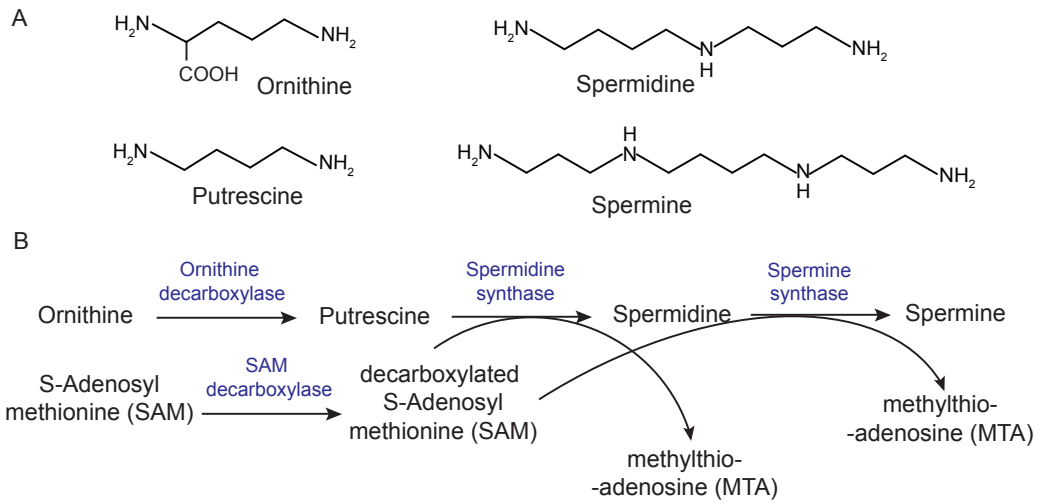


Figure 4: Chemical structures and synthesis pathway of polyamines. (A) Chemical structure of different polyamines. (B) Polyamine synthesis pathway.

modification, ion channel gating and membrane stability [32, 35] (figure 5). In an *in vitro* context, polyamines have also gathered considerable attention as condensing agents, potential radioprotective agents and as cancer chemopreventive agents [36].

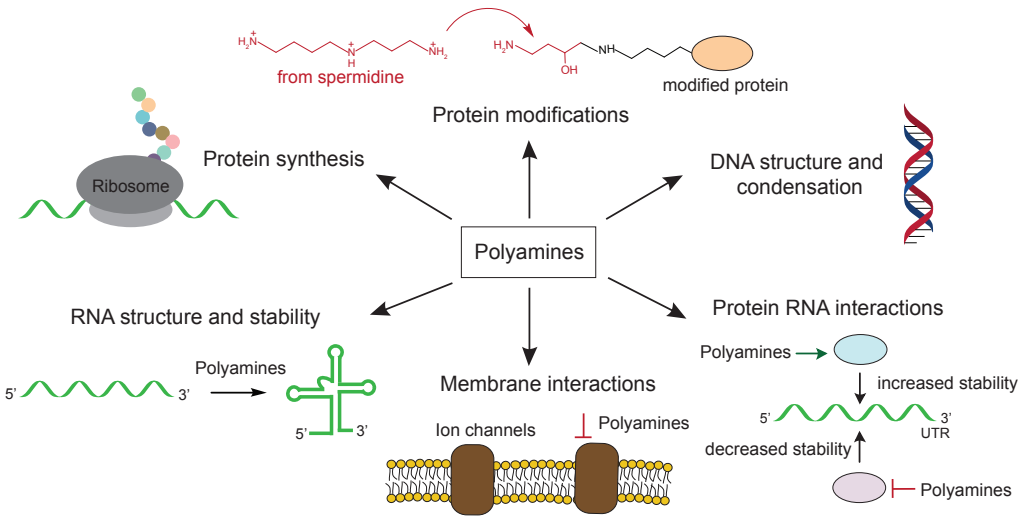


Figure 5: Applications of polyamines in biology.

2.3.2 *In vitro* DNA condensates

More detailed investigation of the biological effects of DNA condensation typically requires studies *in vivo*. However, model systems inducing DNA condensation under controlled laboratory conditions are of great interest for understanding the packaging of DNA and RNA in viruses, functional organization of DNA in prokaryotic nucleoids and eukaryotic nuclei. *In vitro* studies provide greater insights into the physical aspects and potential applications of DNA condensation.

DNA condensation can be induced *in vitro* either by applying an external osmotic force to bring the double helices together, or by inducing attractive interactions between the DNA segments. Different substances have been employed, which use either of the above mentioned approaches to induce the condensation of DNA. They include multivalent cations (e.g. polyamines, cationic polypeptides – polylysine, basic proteins – histones and inorganic metal cation complex – hexammincobalt $Co(NH_3)_6{}^{3+}$, alcohols (e.g. different percentages of ethanol, methanol and isopropanol alone or in mixture), neutral and anionic polymers (e.g. polyethylene glycol (PEG) at high concentrations and in presence of monovalent salt, anionic polymers – polyglutamate, polyaspartate and anionic peptides) and cationic liposomes [37]. Condensing agents generally function by decreasing the repulsions between DNA segments either by neutralizing the negative charge of the phosphate by cations, or by making DNA-solvent interactions less favorable, like adding ethanol, which is a poorer solvent than water for DNA, or by addition of another polymer which excludes volume such as PEG. Multivalent cations are the most thoroughly studied condensing agents. The condensation process can be reversed by addition of mono or divalent cations like sodium (Na^+) or magnesium (Mg^{2+}) to condensates produced by multivalent condensing agents. They compete with the multivalent ion for binding and also influence the effective concentration of multivalent ions required for condensation [38].

It has been found that the morphology of DNA condensates strongly depends on the DNA concentration along with the type and concentration of condensing agent used. Toroidal DNA particles were observed for the first time upon treatment of T7 phage DNA with spermidine [39] and $Co(NH_3)_6{}^{3+}$, and other multivalent amines thereafter (figure 6 [40]). Tightly packed toroidal DNA complexes have drawn maximum scientific interest because of their similarity in size and appearance to

DNA bundles produced by the lysis of T2 and T7 phages heads [41]. They have been investigated as a prototype to explain the high density packing of DNA in virus capsids and considered as a potential candidate for gene delivery applications [42, 43]. Other conformations such as rods and fibres, have been seen under different conditions of salt and alcohol ([38, 44]. Liquid crystalline arrays of DNA have been observed at very high concentrations. These liquid crystals have been shown to exist in different phases such as cholesteric and columnar hexagonal phase (figure 6 (H-I)). Naturally occurring DNA with periodic order has also been observed in a variety of viruses, prokaryotic and eukaryotic cells suggesting it to be a pertinent mode of organization *in vivo* [45] (figure 6).

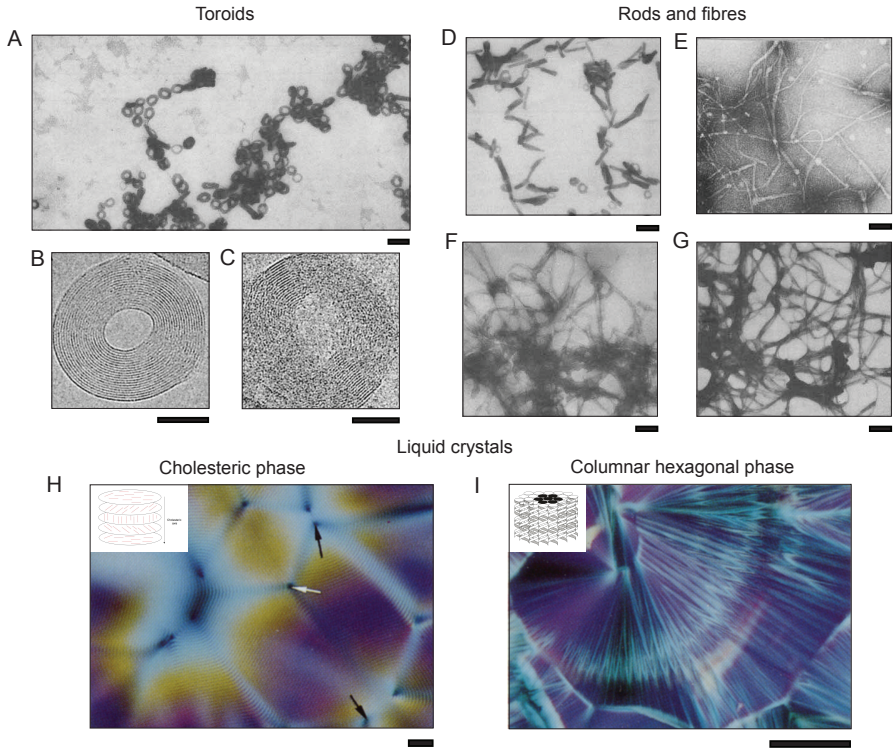


Figure 6: Different morphologies observed during *in vitro* DNA condensation. (A-C) DNA toroids condensed by $Co(NH_3)_6{}^{3+}$ and observed in transmission (A,[44]) and cryo-electron microscopy (B, C,[40]) respectively. (D-G) DNA rods and fibres, produced as a transition from toroids [44]. (H,I) DNA liquid crystals exhibiting cholesteric and columnar hexagonal phases respectively using polarizing microscopy [46]. Scale bar: 50 nm (B and C), 100 nm (A, D-G) and 10 µm (H and I).

Apart from contributing to diverse morphologies of DNA, condensing agents have also been found to impart special properties. For instance, polyamines have been

shown to enhance the stability of DNA duplexes with respect to heat denaturation and radiation [47, 48]. It has been proposed that both electrostatic interactions between polyamine amino-protonated groups and negatively charged backbone PO_2^- group, as well as direct interaction with DNA bases are responsible for such features. Further, polyamines also have the ability to induce DNA conformational transitions under certain conditions [49]. Different theories and models have been proposed for understanding the interactions between DNA and multivalent ions in DNA condensation, some of which are discussed in the next section.

2.3.3 Physics of DNA condensation

DNA condensation is defined as the collapse of long chains of DNA initially in a coil configuration into a compact globule state in presence of condensing agents. DNA is a highly charged polyelectrolyte with a linear charge density of -5.9 e/ nm. This gives rise to a variety of interesting biophysical effects such as counterion condensation [50, 51], electrostatic stiffening [52] and aggregation in the presence of condensing agents [37, 53], which are intimately linked to the biological functions of DNA. In aqueous solutions, monovalent and divalent cations are attracted by the negatively charged DNA polymer. These ions can either be associated with DNA or remain close to the surface and stay mobile. This results in formation of an ion cloud around DNA. Various polyelectrolyte theories have been given which describe the electrostatic behaviour of DNA-polycation systems. One such theory is the mean-field Poisson-Boltzmann (PB) theory. It is based on two assumptions – all charges are considered as point like and only interactions between charges considered is Coulombic interactions.

It allows the calculation of the electrostatic potential ψ in the neighbourhood of DNA, which describes the nature of interactions between DNA and other cations in solution. The electrostatic potential is given by the Poisson-Boltzmann equation–

$$\nabla^2\psi = \frac{-e}{\epsilon}\sum_{i=1}^{N} z_i n_0 exp(\frac{-ez_i\nabla}{kT}) \tag{8}$$

where ϵ is the dielectric constant of the medium, T is temperature, k is Boltzmann constant, n_0 is the average concentration of ions, and z_i is the valence.

According to this theory, the interaction between two DNA molecules is always

repulsive even in the presence of polycations. Hence, this theory cannot provide an explanation for the DNA condensation process.

For DNA condensation to occur, repulsive forces between highly charged DNA strands need to be overcome and attraction to take place. The physics behind DNA condensation can resolved by a correlated counterion fluctuation correction to the PB mean-field theory. The counterions not only screen coulombic repulsions between the DNA phosphates, they also produce attraction through correlated fluctuations of the ion atmosphere. A non-specific long-range attractive contribution to the force between two DNA molecules that varies inversely with the square of the inter-rod distance, was obtained by including correlations between the fluctuations of the two DNA [51]. Different forces are active during DNA condensation. They include bending, entropic forces, coulombic forces and hydration. Bending and entropic forces resist DNA condensation, as they involve bending of a stiff polymer like DNA into shapes; and reduction in number and compaction of molecules which are responsible for entropic effects. Coulombic and hydration forces are the only forces which favor condensation.

Because of the high charge of DNA, its condensation requires counterion compensation. According to Manning's counterion condensation theory, it is calculated that DNA condensation occurs when atleast 75% of the phosphate charge is neutralized. The percentage of charge neutralization depends on the valence of ion, thereby providing an explanation to condensation induced by ions of valence greater than or equal to three and not by monovalent ions. However, divalent cations can induce condensation, but only under certain conditions [50, 54]. Other mechanisms which explain the condensation phenomenon include, the direct ion-bridging of neighboring DNA chain segments by polycations based on short-range electrostatic interactions and the attraction force induced by long-range attractive hydration forces as a result of restructuring of water molecules between adjacent DNA molecules induced by counterions.

Recently, atomistic level molecular dynamic simulations have been applied to investigate the structure, inter-DNA forces and ion atmosphere in DNA condensates for a broad range of ionic conditions (Na^{+}, Mg^{2+} and spermine (Spm^{4+})). The results of the simulations were concordant with the experimental data and showed that DNA assembled in hexagonal lattices with spacing of 28-34 Å [55, 56]. The spacing was influenced by the valence of counterions. Further, it was shown that

the condensation process is driven by the electrostatics of polycations and there was no influence of the hydration forces. It was found that the cations carrying a higher charge localized near the surface of DNA. Non-contact bridging was suggested to be responsible for the attractive forces between DNA strands rather than direct-bridging by ions with valence ≥ 3 [57].

2.4 Nucleic acid nanotechnology

Nucleic acids have been best known as carriers of molecular information which is central to their role in biological systems. It is this information carrying feature that distinguishes both DNA and RNA from other biological and man-made polymers. Ever since the elucidation of the structure of DNA by Watson and Crick, it has largely been used as a standard in understanding the fundamental biological processes and exploiting that understanding to develop practical applications. This has resulted in a series of advancements in areas as diverse as basic biology, polymer physics, forensics and medicine.

In 1982, Nadrian Seeman, a chemist and crytallographer proposed the building of nucleic acid junctions and lattices by exploiting the unique molecular recognition properties and strict base pairing rules of DNA in a completely non-biological context [4]. It was inspired by the already existing nucleic acids junctions occurring in biology, that are recognized by WC base-pairing and sequence symmetry between DNA strands. Such junctions have been transiently observed and are important to various biological processes such as the replication junction formed during DNA replication, Holliday junctions in genetic recombination and cruciform structures in gene regulation [58, 59]. The sequence symmetry causes instability of the branch points in the junctions which impeded the physical characterization of junctions for a long time. Seeman suggested that maximization of WC base-pairing and lack of sequence symmetry of oligomeric nucleic acids will preferentially associate them to form migrationally immobile juctions, which could possibly result in formation of two (2D) or three dimensional (3D) lattices by intermolecular interactions (figure 7).

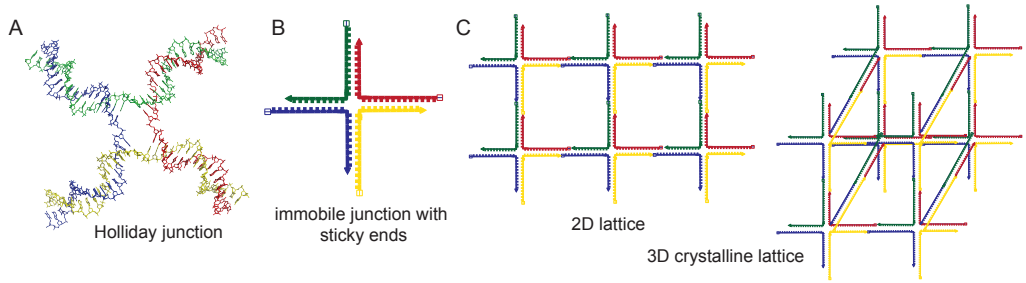


Figure 7: Foundations of nucleic acid nanotechnology. (A) X-Ray structure of a Holliday junction of bacteriophage P1, item 3CRX from RCSB PDB [29]. (B) Immobile DNA junction proposed by Seeman. (C) Self assembled 2D and 3D DNA lattices using immobile DNA junctions.

The initial goals of such lattices based on branched DNA motifs was to scaffold the assembly of proteins and other molecules in periodic arrays for crystallographic studies and memory-device construction [60, 61]. This has led to the emergence of the field of "nucleic acid nanotechnology" with nucleic acids serving as important building blocks for bottom-up nanotechnology. The central strength of nucleic acid nanotechnology is its ability to use the chemical information contained in programmed DNA and RNA sequences to organize either itself or other chemical species for specialized functions. Both DNA and RNA have been used as molecules for the self-assembly of molecular nanostructures and also as a material for building machine-like nanodevices. Alongside, numerous other milestones have been achieved and other new directions have been enunciated. Some of which are described in detail in the following sections.

2.4.1 DNA nanotechnology

The specificity and predictability of WC base-pairing have made DNA an ideal and multifaceted material for engineering and programmed self-assembly at the nanoscale. The availability of convenient and economical solid support synthesis for DNA, stability and ease of chemical modifications (such as biotin, lipids, fluorescent labels etc.) specific to function make it a molecule of choice. Moreover, the extensive amount of research done in the fields of chemistry, molecular biology and biotechnology have created a toolbox including a battery of enzymes, unnatural nucleotides etc. capable of manipulation of DNA [62, 63].

Structural DNA nanotechnology focuses on bottom-up assembly and characterization of DNA based nanostructures into a static, equilibrium end state. It employs synthetic immobile DNA junctions which have progressed into rigid DNA multiple cross-over molecular motifs more suitable for the assembly of periodic structures [64]. A detailed description of the different DNA nanostructures realised by synthetic branched junctions and multiple cross-over motifs made from DNA are provided in [6]. The shape and size of such structures has been precisely tweaked. However, a major limitation has been the requirement of perfect equimolar stoichiometry of the component oligonucleotides. As a consequence, structures with limited size, complexity and lower yield have been created. Several other multistranded DNA structures have been reported including DNA triplexes, G-quadruplexes and I-motif DNA structures. These structures provide new insights into helix stability and folding, loop conformations and cation interactions [65].

One of the most important developments in structural DNA nanotechnology since the introduction of the double-crossover (DX) motifs has been the use of a 'scaffold' DNA strand for the assembly of nanostructures. It has been previously demonstrated that a long ssDNA could be used to organize DX motifs into barcode-patterned lattices, and could assemble into a 3D wire-frame octahedron [66, 67]. However, the major breakthrough came with Paul Rothemund's 'DNA origami technique', where a long ss DNA scaffold (genomic DNA from M13mp18 phage, $\sim$ 7429 nucleotides long) could be folded into predefined shapes with the help of hundreds of short 'staple' strands, complementary to different regions of the scaffold [68]. The DNA origami technique has several important advantages in comparison to the traditional assembly strategies mentioned above. It avoids stoichiometric dependence and the folding proceeds without many errors. It results in fewer defects and higher yield of the final product. This enables the synthesis of DNA nanostructures with defined dimensions and more complex shapes, with full spatial addressability, facilitating the attachment of molecules at prescribed positions at nanometer-scale precision. Since then, the method has been generalized to construct a variety of DNA based architectures in 2D and 3D with customized curvatures, twists, bends etc [69–71]. The robustness of the technique has made it widely accessible.

More recently, techniques for modular scaffold-free fabrication [72, 73], 3D meshing, and wireframe-based approaches [74, 75], as well as shape-complementarity-based construction [76] of DNA objects have been introduced. Figure 8 shows the representative examples of different strategies used for the construction of artificial DNA nanostructures.

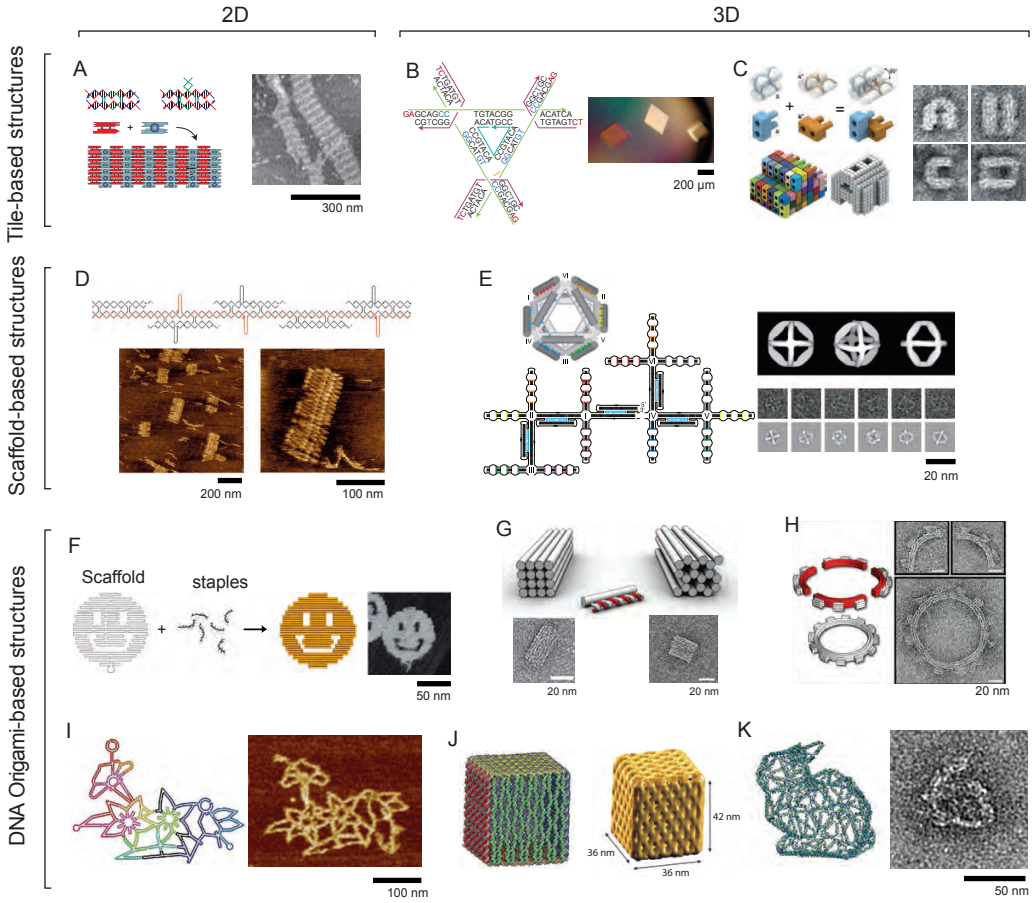


Figure 8: Representative examples of structural DNA nanotechnology. (A) 2D DNA crystalline arrays self-assembled from synthetic DX tiles, [77]. (B) Self-assembled 3D DNA crystal from tensegrity triangle motif, [78]. (C) 3D structures built with single-stranded tile DNA bricks, [72, 73]. (D) Self assembly of barcode lattice around scaffold DNA strand, [66]. (E) DNA octahedron made from ss DNA [67]. (F) 2D DNA origami,[68]. (G-H) 3D and curved DNA origami respectively, [71, 79]. (I) Wireframe DNA origami nanostructures, [74]. (J) Hollow DNA box with controllable lid [70]. (K) 3D DNA nanostructures with polyhedral meshes,[75].

In addition, powerful computational tools for designing and analyzing DNA nanostructures such as cadnano and CanDo have been developed, to design DNA nanoarchitectures for any application [80–82]. Alongside, effort have been put in the improvisation of the existing strategies of DNA synthesis and purification, and new methods have been developed [83–86]. Further, attempts have been made to synthesize DNA based nanostructures *in vivo* [87, 88].

In 1994, Adleman for the first time demonstrated the potential role of DNA in computing. He demonstrated the solution for a computational problem, Hamiltonian Path problem by using DNA and standard molecular biology techniques [89]. Similarly, algorithmic assembly proposed by Winfree added logical operations to the notion of nanoscale self-assembly [90]. Such diverse features and implementations of DNA have expanded the vision of DNA nanotechnology which primarily consisted of structural DNA nanotechnology (with DNA being used as a raw material for bottom-up assembly) to more dynamic non-equilibrium systems employing DNA, giving rise to the sub-fields of dynamic DNA nanotechnology and DNA computing. **Dynamic DNA nanotechnology** focuses on non-equilibrium systems, in which the DNA molecules undergo a series of conformational changes based on a chemical or physical stimulus. It exploits the base-pair interactions and DNA strand displacement reactions to control reaction kinetics in DNA based synthetic systems [91, 92]. As a consequence, a variety of dynamically reconfigurable devices, including synthetic molecular logic circuits, catalytic amplifiers, autonomous molecular motors, programmable DNA guides and reconfigurable DNA nanostructures have been constructed [49, 91], and complex computation using DNA have been realised [93]. Such structures are often called DNA "nanomachines" or DNA "nanodevices". These nanodevices have not only been developed *in vitro*, but also DNA nanomachines functioning as a pH sensor and molecular logic circuits inside the cell have been developed [94, 95]. Examples of dynamic DNA devices based on different principles are shown in figure 9.

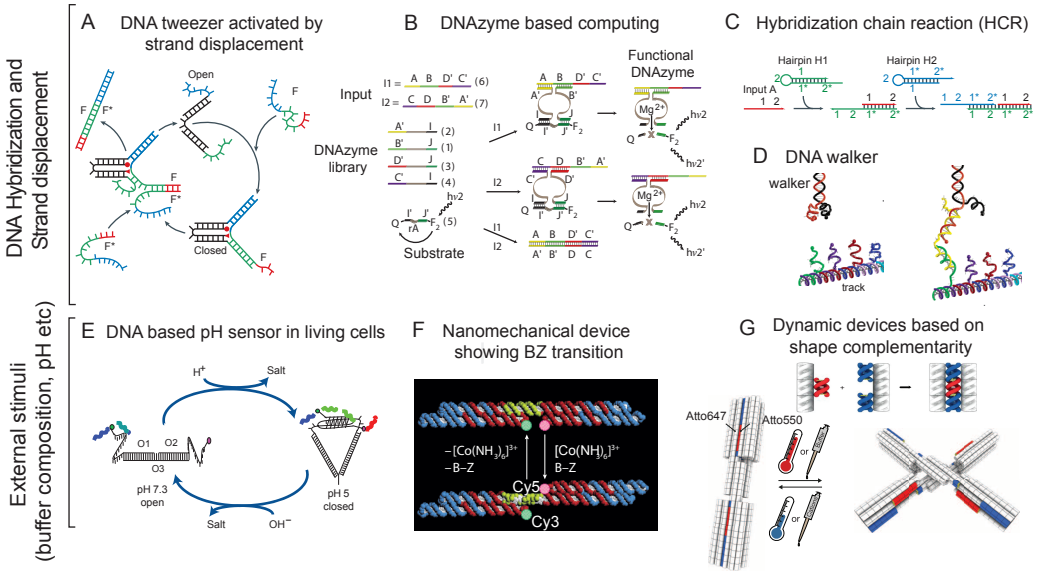


Figure 9: Examples of dynamic DNA nanotechnology. (A) DNA tweezers, [92]. (B) DNAzymes, [96]. (C) Hybridization chain reaction, [97]. (D) DNA walker, [98]. (E) DNA based pH sensor, [94], (F) Nanomechanical device from DNA, [49]. (G) Dynamic DNA devices [76].

DNA nanotechnology offers a plethora of design strategies and techniques that can be used independently or in combination with others, to construct increasingly complex systems for scientific and technological applications. This has led to a continuous outburst of potential applications of DNA nanotechnology in fields as diverse as molecular and cellular biophysics, diagnostics and therapeutics, drug delivery, biomimetic systems, electronics and plasmonics [99]. DNA nanotechnology has also stimulated basic sciences by raising many unasked fundamental questions that will lead to a greater understanding of the molecules at the center of life. These developments have provided ample support for the further advancements of the field of DNA nanotechnology.

2.4.2 RNA nanotechnology

While DNA served as the pioneer building material in the development of nucleic acid nanotechnology, RNA has come forth as an interesting alternative material with its own explicit advantages. The versatility of RNA in structure, thermostability, suitability for bottom-up self-assembly and useful *in vivo* attributes have made it an

attractive candidate. The structure and properties of RNA are discussed in sections 2.1 and 2.2.

Most naturally occurring RNAs are single-stranded and contain self-complementary sequences that facilitate its self folding process. In addition, rigid structural motifs consisting of canonical, noncanonical base-pairing and stacking of helices play a significant role in RNA folding. Analysis of natural RNA has shown a recurrence of similar and thermodynamically stable structural motifs defining localized regions of conserved sequences across different organisms, thereby hinting towards their preferential selection by nature. These motifs allow natural RNAs to fold into a variety of complex 3D structures exhibiting helices, loops, junctions, bulges, stems, hairpins, and pseudoknots [100, 101], for diverse biological functions including intermolecular recognition, catalytic, and regulatory functions [3]. In addition, the interactions between the negatively charged phosphate backbone with metal ions, ligands and proteins, also influence RNA folding, assembly and function. Divalent cations such as Mg^{2+} etc. have been found to be essential for folding and catalytic functions of RNA. Small molecule-binding refolds riboswitches, thereby allowing regulation of gene expression. This feature has been used to manipulate dynamic RNA molecules [102]. This remarkable folding capability and modularity of biological RNAs have provided insights and huge opportunities for self-assembly of novel RNA nanostructures. This has led to the emergence of the field of RNA nanotechnology. Some of the naturally occurring RNA motifs that have been used for construction of artificial RNA based nanostructures are presented in figure 10.

One of the initial works on artificial RNA based nanostructures was from Guo *et*

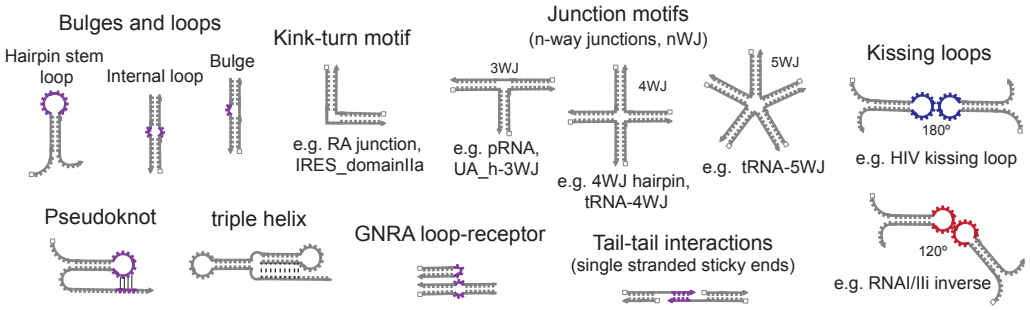


Figure 10: Naturally occurring RNA motifs.

al., where they reported the ability to engineer RNA (pRNA, packaging RNA from ϕ29 virus into precise constructs to build concise RNA architecture such as dimers,

trimers, and hexamers via bottom-up self-assembly, thus introducing the concept of RNA nanotechnology [103]. Since then, there has been a great deal of work done in the development of novel synthetic RNA nanostructures [5].

Seeking inspiration from the design principles found in nature, several assembly strategies have been developed for building a diverse set of RNA nanostructures. They include (1) **RNA architectonics**,(2) **RNA/DNA hybrid self-assembly** and (3) **single-strand RNA assembly**. **RNA architectonics** is based on the modular character of RNA, which can be disassembed and reassembled to create different RNA nanoarchitectures. Large libraries of thermodynamically stable modular RNAs, that include both structural and interacting motifs have been identified and characterized from natural RNA molecules. Further, they have been used as suitable "parts" for designing self-assembling RNA units (or tectoRNAs) [104]. The **RNA-DNA hybrid strategy** leverages the properties of both RNA and DNA. It has been used for producing large nucleic acid structures based on DNA origami technique, where a large RNA scaffold is folded with a number of complementary DNA/RNA staples [5]. **Single strand RNA assembly** relies on a number of RNA strands that are unstructured by themselves but when mixed together assemble into a structure. A technique known as **RNA origami** has been developed, in which a single RNA strand folds into predefined RNA tile that further assembles into hexagonal and rectilinear lattices [105]. It employs a variety of RNA tertiary motifs to mediate the intra and inter RNA tile interactions. It is described in greater detail in chapter 5. Schematic illustrations for each is shown in figure 11. New unimolecular DNA and RNA folding strategies have been developed based on minimizing the knotting complexity by employing parallel crosovers in design, that avoid kinetic trapping during folding. This has resulted in RNA nanostructures of a variety of shapes and up to 6000 nt in length [106].

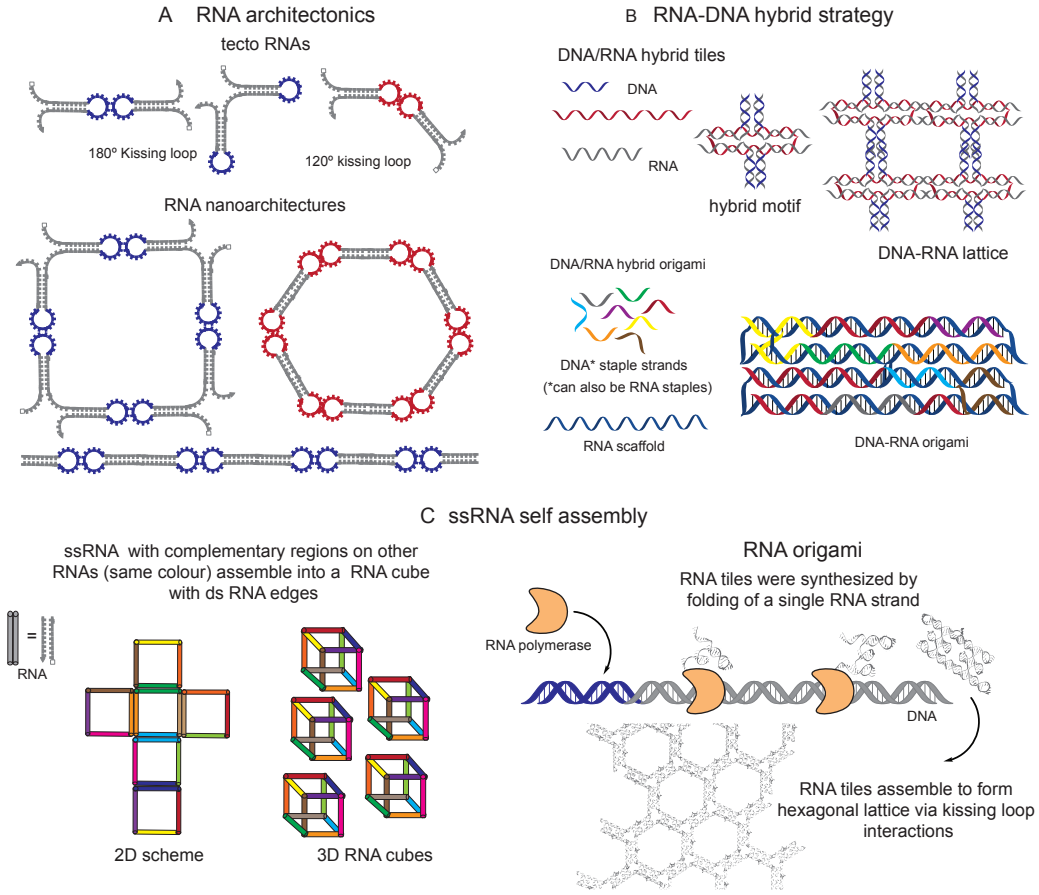


Figure 11: Strategies for construction of RNA nanostructures..(A) RNA architectonics. (B)DNA/RNA hybrid assembly. (C) ssRNA self-assembly.

Cotranscriptional self-assembly is an important complementary strategy in RNA nanotechnology and has been applied to the above mentioned formats of self-assembly. It couples RNA synthesis with RNA self-assembly and can be applied in both *in vitro* and *in vivo*. A unique example is rolling circle transcription (RCT), where long RNAs with periodic repeats are synthesized from a circular DNA template [107].

RNA nanostructures have been produced by transcription both *in vivo* and *in vitro*. RNA is genetically encoded and expressed in cells using their transcription machinery [108, 109]. Artificial RNAs that function as fluorescent tags have been produced inside cells [110, 111]. *In vitro*, purified RNA polymerases from *E.coli* or bacteriophages (T7, T3, SP6 etc.) have been used for the synthesis of RNA and RNA based

nanostructures (figure 12). The dependence on the cellular transcription system *in vivo* and *in vitro* enzyme catalyzed synthesis results in copious amounts of synthesized RNA. This is one of the major advantages of RNA based nanotechnology. There has been an increasing interest of RNA based nanotechnologies in various

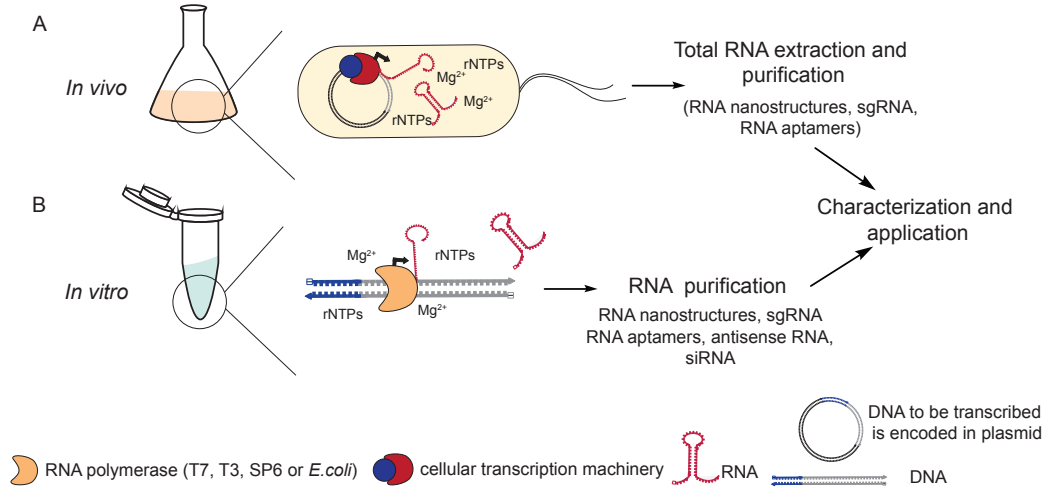


Figure 12: Methods for production of RNA nanostructures.(A) *In vivo.* (B) *In vitro.*

fields such as biochemistry, biophysics, chemistry, structural biology, nanomedicine, synthetic biology and biocomputation. Both structural and functional properties of RNA have been exploited. RNA-based therapeutics has gained considerable attention owing to the inherent abilities of RNA [112]. It includes siRNAs, miRNAs, antisense oligonucleotides (ASOs), RNA aptamers, ribozymes, riboswitches etc. Artificial RNA based nanostructures have found great potential especially in the field of RNA-based therapeutics. They function as nanoscaffolds for with controlled stoichiometric organization of mulitfunctional attributes including therapeutic RNA moieties, and other biomolecules such as enzymes for enhanced activities [109, 113, 114]. Other areas of potential applications of small catalytic and regulatory RNAs such as ribozymes, riboswitches include *in vitro* diagnostics, regulation of functional biological pathways and RNA computing [5].

A number of softwares and programs have been developed to understand RNA folding, which play a significant role in the design and construction of RNA nanostructures. Some of them include Mfold, Nupack, SimRNAweb, and Assemble [23, 81, 115, 116]. Also, databases such as RNAjunction and RNA bricks have been

developed that function as repositories of various RNA modules including n-way junctions, loops etc. [117, 118].

Alongside these developments, the field of RNA nanotechnology in general faces some challenges. These include the relative chemical instability of RNA in comparison to its DNA counterpart; kinetic and thermodynamic constraints associated with synthesis and folding of large RNA nanostructures, large scale production, purification and cost, and limited capabilities of endosomal escape. However, various strategies have been developed to overcome these challenges. For instance, chemical modifications of RNA have been shown to impart greater chemical and thermal stability to RNA; and endosomal escape thereby allowing its *in vivo* use. Specialized techniques for production and purification such as ultracentrifugation have been employed to reach high yields of functionalized RNA nanostructures [5]. The development of directed evolution techniques like SELEX have made it feasible to generate a collection of artificial RNAs to function as dynamic modules with novel binding and catalytic properties; and evolve more advanced structural motifs keeping a balance between the kinetic and thermodynamic aspects associated with RNA folding for construction of larger more complex RNA nano-assemblies.

3 DNA origami nanostructures folding using multivalent ions

3.1 Introduction

Recent advances in DNA nanotechnology have demonstrated that DNA in addition to its paramount biological role – can be successfully utilized as a programmable building material for the self-assembly of biomolecular nanostructures [119] and machine-like dynamical nanodevices [120]. DNA nanostructures have been used as platforms for the precise spatial organization of a variety of functional molecules including fluorophores [121], aptamers [122], antibodies [123], membrane receptors [124], lipids [125, 126] and inorganic nanomaterials [127] via different strategies, which has led to the emergence of a wide range of applications for DNA nanotechnology.

DNA origami has been one of the most robust and widely used methods for the design and assembly of different geometries of DNA nanostructures [68, 69]. It is based on the hybridization of a large number of short DNA "staple" strands - typically of $\approx$ 30-50 nucleotides (nt) in length (variable, and depends on the design) with a long ssDNA molecule termed the "scaffold", which is commonly derived from genomic DNA of M13mp18 phage and has a length between 7 and 8 kilobases (kb). This hybridization enables different parts of the scaffold to come together and fold into a structure with a predefined shape and size. The overall structure of the origami complexes can be freely chosen and is determined by the routing of the scaffold and thus the choice of the staple sequences. A completely folded DNA origami structure comprises of a large number of DNA double helices closely packed together and connected via multiple strand crossovers.

One of the potential disadvantages of DNA as a molecular building material is its polyanionic nature. dsDNA is one of the most highly charged polymers existing in nature with a linear charge density of $\approx$ 1e/0.17 nm (or 5.9e/nm). On the contrary, this feature of DNA is responsible for a number of fascinating biophysical effects, some of which are described in section 2.3. In context of structural DNA nanotechnology, the strong electrostatic repulsion between DNA duplexes has to be overcome to facilitate the assembly of compact DNA structures. In most folding protocols for DNA nanostructures, the divalent magnesium cation (Mg^{2+}) is thus

used at ≈ 10-30 mM concentrations to reduce the inter-helical repulsion [68–71] and package the DNA to fold into predefined structures of desired shapes.

Diverse applications require DNA nanostructures to adapt and be functional under conditions, distinct from their initial folding and characterization. For instance, high magnesium concentrations are detrimental for the function of enzymes, and may also lead to nanoparticle aggregation. In some cases, it is possible to replace the buffers of DNA origami structures from Mg^{2+}-containing into buffers containing high Na^+ or K^+ concentrations with reduced or zero magnesium content [126, 128]. Other strategies involve the covalent linkage of the DNA strands comprising the structures [129], or the utilization of wireframe assemblies with reduced inter-helical electrostatic interactions [74, 75]. Origami structures have also been folded in buffers containing high concentrations of monovalent cations (i.e., typically, $[Na^+] > 0.5$ M [130]), and even in the water-free eutectic solvent glycholine (a 4:1 mixture of glycerol and choline chloride) [131]. Additionally, DNA origami formation has also been controlled by dendrons (functionalized with tetravalent ions) that respond to different external stimuli such as pH, temperature, light, electric and magnetic fields [132]. Also, europium (Eu^{3+}) decorated DNA origami structures have been observed with preserved structural integrity, even in the absence of Mg^{2+} [133].

The high salt concentration requirements for the assembly of DNA nanostructures poses a particular challenge for potential applications of DNA nanostructures in cell biology and nanomedicine. Recently, the sensitivity of DNA origami nanostructures to *in vitro* cell culture media was studied [128]. It was found that the denaturation of the structures due to depletion of divalent cations was a major problem, and they also suffered from degradation by nucleases. These effects could be rectified, e.g., by inhibiting the nucleases and supplementing the media with magnesium. Various approaches to protect and stabilize DNA nanostructures under physiological conditions have been attempted, which involve the utilization of PEG [134] as a crowding agent, oligolysine-PEG complexes [135], decoration with viral coat proteins [136] or complexation with transfection agents such as lipofectamine, which also facilitate cellular uptake of the structures [136–139].

As a part of work done in this thesis, an alternative approach for folding of DNA origami structures under low ionic strength and conductivity conditions that employs a trivalent DNA condensing agent, spermidine (Spd^{3+}) was investigated. Spermidine is a biogenic polyamine found at concentrations ranging from 0.1-3 mM in

mammalian cells, with typical concentrations on the order of 2 mM, which may vary slightly depending on the cell state and type [34]. Spermidine along with other polyamines are known to play an important role in DNA condensation. A detailed description of the synthesis pathway and functions of spermidine and associated polyamines are described in section 2.3.

In this chapter, the "biocompatible" folding of three types of DNA origami structures with different shapes and helix packing densities at approximately physiological spermidine concentrations ($[Spd^{3+}] < 1$ mM) is demonstrated. Further, the effect of Spd^{3+} on the conformation of the structures, as verified by TEM and circular dichroism experiments is presented. In addition, the consequences of ratio of condensing agent to DNA and the concentration of condensing agent on folding of DNA nanostructures are shown.

3.2 Experimental details and methods

One pot synthesis of DNA origami nanostructures

Three different DNA origami nanostructures namely, the twist corrected regular Rothemund rectangle (tCRRO, a single-layered square lattice structure consisting of 24 parallel double helices with dimensions 90 nm × 65 nm × 2 nm (length × width × height) [68]), a DNA origami tube (a hollow honeycomb lattice structure with dimensions 83 nm × 23 nm × 29 nm (length × diameter 1 × diameter 2)) designed by J. List (Simmel Lab, TUM), and a 42 helix bundle (42HB, a compact honeycomb lattice structure with dimensions 60 nm × 26 nm × 16 nm [130]) were each assembled in one pot reactions using different variants of M13mp18 scaffold and their respective sets of DNA staple strands (used at 3× excess) in a variety of folding buffers – traditional buffer containing magnesium (1X TAE (pH 8), 10-20 mM $MgCl_2$, depending on the nanostructure) and solutions containing different concentrations of spermidine trihydrochloride ($C_7H_22N_3Cl_3$ or Spd.3HCl or Spd^{3+}) (Sigma). The structures folded in traditional buffer and $C_7H_22N_3Cl_3$ based buffers were referred to as Mg-nanostructures and Spd-nanostructures respectively. DNA staple strands were purchased from Eurofins MWG, M13 scaffold variants (p7560, p7249) were kindly provided by F. Praetorius (Dietz lab, TUM) [83]. The sample solutions were subjected to a thermal annealing ramp that cooled from 65°C to 20°C over a time course of several hours (different annealing times were used depending

on the target nanostructure. The details of the thermal annealing ramp is provided in the appendix (figure 73). Cadnano designs, nucleotide sequences of the different structures and scaffold sequences are provided in the appendix (figures 48-50, 67-71 and 72 respectively). All chemicals were purchased from Carl Roth, Germany unless otherwise specified.

Agarose gel electrophoresis (AGE)

Agarose gel electrophoresis was carried out to check for proper folding of structures. All the reaction products were analyzed in 2% agarose gels containing 1X TAE, 12.5 mM $MgCl_2$, at 100 V in a gel box immersed in an ice water bath for 2.5 hours. The gels were post-stained with SYBR gold (Invitrogen, USA) for 40 minutes and destained with buffer. The gels were scanned on a Typhoon FLV 9000 laser scanner and analyzed with the software package Fiji [140]. Samples were mixed with 6X purple loading dye (NEB) and loaded onto the gel. 2-Log DNA ladder (NEB) was also used as a marker.

Transmission electron microscopy (TEM)

For electron microscopy, purified DNA nanostructures were adsorbed on glow-discharged formvar-supported carbon-coated Cu400 TEM grids (Science Services, Munich, Germany) and stained using a 2% aqueous uranyl formate solution containing 25 mM sodium hydroxide. The grids were plasma cleaned with carbon side up before use. Plasma conditions: vacuum – 0.8 mbar, charge – 30 mA. The staining procedure involved three sequential steps – adsorption of sample (10-20 nM) (5 µl) on grid, for 30 seconds, removal of unadsorbed sample and washing with 5 µl stain (for negative staining) / water (for positive staining) ($\sim$ 1-2 seconds) and incubation with 20 µl stain for 40 seconds. The stain was removed with filter paper and the grids were dried for $\sim$ 30 minutes prior to imaging. Imaging was performed using a Philips CM100 electron microscope operated at 100 kV. Images were acquired at 28500X magnification using an AMT 4 Megapixel CCD camera. Class averaging of DNA nanostructures was performed by using the Xmipp protocol provided in the software package Scipion [141]. Negative staining was done for 42 helix bundle and tube structures; and positive staining was carried out for rectangles.

Circular dichroism (CD) measurement

Circular dichroism spectra of 42HB folded in traditional buffer containing magnesium and in Spd^{3+} were recorded from 200 nm to 400 nm with data being collected every 0.2 nm on a Jasco J-815 CD spectrometer.

3.3 Results and discussion

3.3.1 Folding of DNA origami nanostructures using a trivalent condensing agent

The trivalent condensing agent spermidine (Spd^{3+}) was used to facilitate the assembly of three DNA origami nanostructures with different shapes and helix packing densities: a twist-corrected regular Rothemund rectangle [68], a hollow DNA origami tube, and a 42 helix bundle (42HB) [130] in a one pot reaction.

Schematic representation of the folding process and the chemical structure of the condensing agent is given in figure 13 (A). The yield and success of the origami folding process was analysed by agarose gel electrophoresis (AGE). 42 helix bundles (50 nM) were assembled at varying Spd^{3+} concentrations ranging from 0-10 mM as shown in figure 13 (B) (subsection 3.2). As a positive control, nanostructures were also folded in conventional folding buffer with typical millimolar concentrations of $MgCl_2$ (C). For reactions with Spd^{3+}, concentrations up to 700 µM (well 9, figure 13 (B)), bands corresponding to unstructured or partially structured DNA complexes moving with the scaffold (S) were observed. Faster migrating species emerged at higher concentrations of condensing agent ($[Spd^{3+}] \approx$ 800–1000 µM), which corresponded to the correctly folded origami structures (red box in figure 13 (B)). Agarose gels corresponding to successful folding of DNA origami tubes and rectangles in Spd^{3+} buffers are shown in figure 51 (appendix).

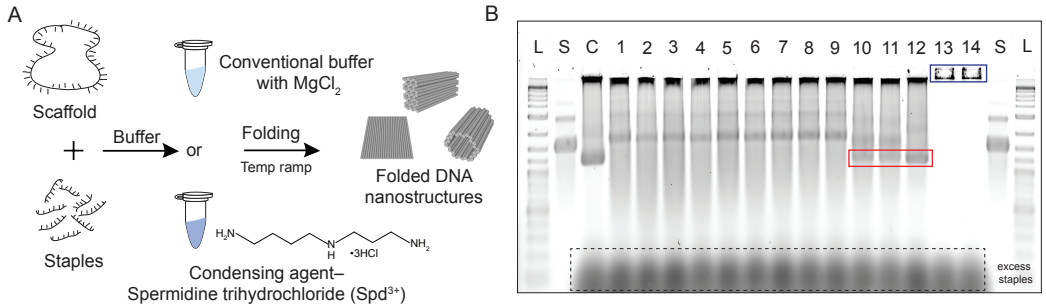


Figure 13: Folding of DNA origami nanostructures in condensing agent spermidine (Spd^{3+}). (A) Schematic illustration of the folding process using different buffers. (B) Agarose gel demonstrating the folding of 50 nM 42 helix bundle using different concentrations of Spd^{3+}. L: 2-Log ladder, S: scaffold, C: DNA nanostructures folded in $MgCl_2$ (control). Spd^{3+} concentrations in the lanes, Lane 1: 0 µM, 2: 50 µM, 3: 100 µM, 4: 200 µM, 5: 300 µM, 6: 400 µM, 7: 500 µM, 8: 600 µM, 9: 700 µM, 10: 800 µM, 11: 900 µM, 12: 1 mM, 13: 5 mM, 14: 10 mM.

Alongside gel electrophoresis, the correct assembly of the Spd-nanostructures (42HB, tube and rectangles) was verified by TEM. Different staining protocols were used for different structures depending on their shape, for their proper imaging (42HB and tube– negative staining; rectangles– positive staining) (figure 14 (A)-(C)) as described in the previous section.

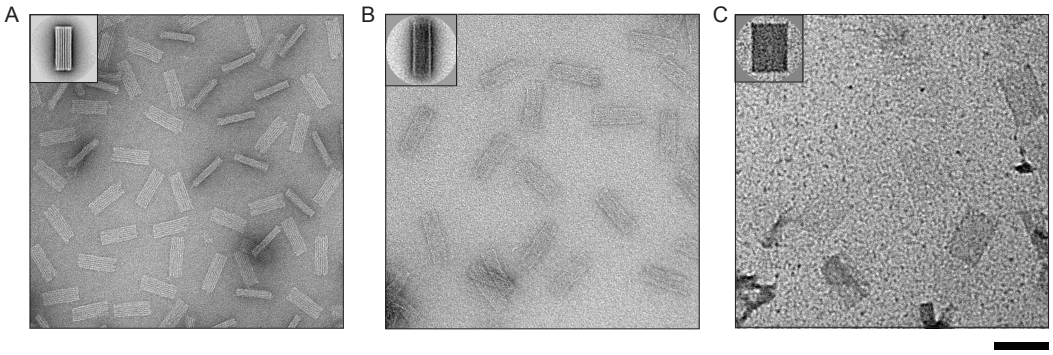


Figure 14: TEM micrographs of DNA origami nanostructures of different shapes and lattice types in Spd^{3+}). (A) 42 helix bundle. (B) Tube. (C) Rectangles. Inset: Class-averaged single particle image. Scale bar: 100 nm.

In addition, both Mg and Spd-nanostructure variants of the three DNA origami structures– 42HB, tube and rectangle were folded and analyzed by TEM as shown in figures 52-54 (appendix). The dimensions of each of the structures folded either in millimolar concentrations of $MgCl_2$ or micromolar concentrations of Spd^{3+} were

found to be the same and no other change was observed. These results confirm the proper assembly of DNA origami based nanostructures at low concentrations of trivalent condensing agents without any compromise in the size and shape of the structures. It has been demonstrated before that multilayered DNA origami nanostructures can be folded in buffers containing high concentrations of monovalent salt (i.e., Na^{+}), which are roughly 100 times higher than the magnesium concentrations contained in typical folding buffers [130]. The optimum Spd^{3+} concentrations found here are about ten times smaller than the usual magnesium concentrations, thereby suggesting the dependence of concentration of salt required for folding on the valence of ion.

3.3.2 Circular dichroism

Figure 15 shows the results obtained from circular dichroism (CD) measurements of 42HB folded in traditional Mg^{2+} and Spd^{3+} buffer. Both Mg and Spd-nanostructures resulted in the same CD signatures, suggesting that no major conformational distortion was induced by the presence of DNA condensing agent Spd^{3+}.

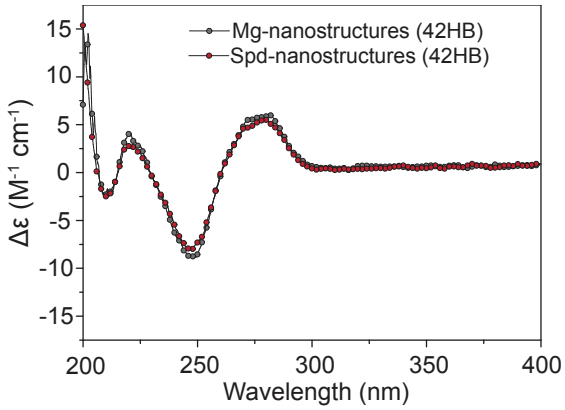


Figure 15: Circular dichroism of Mg and Spd-nanostructures.

3.3.3 Valence dependence and aggregation effects

It was observed that the ratio of DNA (scaffold and staples) and condensing agent had a strong influence on the folding of nanostructures. Figure 16 shows that the concentration of Spd^{3+} required for folding was approximately linearly proportional

to the concentration of the DNA origami structures to be folded. At a particular DNA concentration, folding occurred over a narrow range of Spd^{3+}. However, no such observation was made for Mg-nanostructures, and all of them assembled at the same magnesium concentration irrespective of the DNA concentration. This observation was consistently made for all the three DNA origami nanostructures– 42HB, tube and rectangles regardless of their shape and size. The data shown here is for 42 helix bundles. In most folding protocols, the concentration of the nanostructures folded is determined by the starting scaffold concentration. The 42HB were formed by $\approx$ 7500 bp using a three-fold excess of staple strands over scaffold during the assembly. For a scaffold concentration of 50 nM, this corresponds to a concentration of DNA phosphates of $\approx$ 1.5 mM in the folding solution. In this case, the optimum Spd^{3+} concentration for folding was found to be $\approx$ 800 µM (figure 13 (B), gel) (i.e., $[Spd^{3+}]/[\text{DNA-phosphate}] \approx 0.5$), slightly overcompensating the charge of the phosphates.

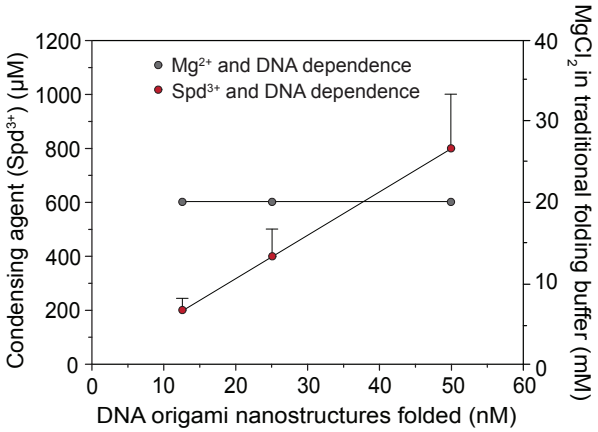


Figure 16: Relation between DNA origami and folding buffer concentration. The concentration of Spd^{3+} required for folding is roughly proportional to the concentration of DNA nanostructures to be folded (red circle symbols). Grey circle symbols show that folding of DNA origami nanostructures is independent of the concentration of DNA and folding occurs at the same concentration of $MgCl_2$ in traditional buffer (at $\approx$ 20 mM).

It was also observed that any deviation from the appropriate ratio of DNA and Spd^{3+} results in different behaviour. Figure 17 shows the various effects seen as a result of deviation from the optimum ratios. For a constant DNA concentration, at higher trivalent salt concentrations, a whitish precipitate was observed just after mixing even before the start of the folding process, suggesting aggregation and

potentially condensation of DNA. The presence of these aggregates was also visible in the agarose gels (blue box in figure 13 (B), wells (13-14) and TEM micrographs (figure 17 (C)). Further, no folding was observed at sub-optimal concentrations of Spd^{3+} and fixed DNA concentration (figure 17 (A)).

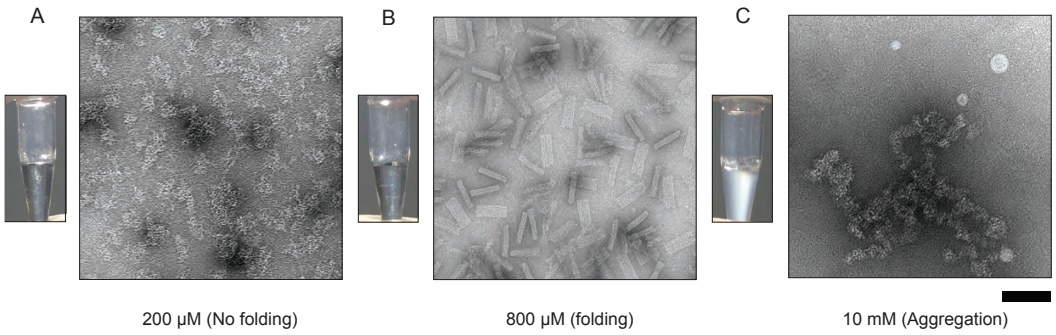


Figure 17: Folding reactions and corresponding TEM micrographs of 42HB at different concentrations of Spd^{3+}. Concentration of 42HB to be folded = 50 nM. Aggregation and precipitation of DNA scaffold and staples at higher concentrations of Spd^{3+} is observed. (A) 200 µM (No folding). (B) 800 µM (folding concentration for 50 nM 42HB). (C) 10 mM (aggregation of DNA even before thermal annealing cycle). Scale bar: 100 nm.

It is well known that multivalent cations (with charge $\geq +3e$) such as the polyamines spermidine and spermine interact differently with DNA than cations with smaller charge [37, 53]. Multivalent ions mediate like-charge attraction between DNA duplexes and can thus condense DNA into tightly packed structures whereas Na^+ and Mg^{2+} cannot induce DNA condensation in aqueous solution. Various models have been put forward to explain the peculiar interactions between polyamines and dsDNA molecules, which assume a correlated arrangement of the counterions along the charged DNA rods and also bridging between helices by the rod-like polyamines [142]. Experimental evidence from CD, IR (Infrared spectroscopy), and NMR (Nuclear magnetic resonance) studies [143, 144] as well as molecular dynamic (MD) simulations [145] suggest that Spd^{3+} preferentially interacts with the phosphates [143], but also with the bases in the major and minor groove [144].

Although, the experimental settings of both processes – DNA condensation and Spd^{3+}-origami folding are different, there exists some similarities between the two. In conventional DNA condensation experiments using polyamines, $\approx$ 150 bp long dsDNA molecules were mixed with the condensing agent with or without additional monovalent salt, and the concentration range in which precipitation of DNA oc-

curred, was determined [38]. The polyamine concentration inducing precipitation was found to increase smoothly with the "intermediate" DNA concentration (0.1-10 mM). At other DNA concentrations, a deviation from this linearity was observed. At low DNA concentrations, larger amounts of polyamines were required, when monovalent salt were added to the solution.

In case of Spd^{3+}-mediated DNA origami folding, $\approx$ 7.5 kb long ss scaffold strand and a large number of shorter staple strands at roughly millimolar (i.e., "intermediate" [38]) DNA monomer concentrations, are mixed in a Spd^{3+}-folding buffer (without additional salt) and subjected to a thermal annealing ramp. A trend similar to DNA condensation was also observed in case of Spd^{3+}-origami folding, where concentration of Spd^{3+} required for folding varied almost linearly with the concentration of DNA origami to be folded (figure 16). It is probable that Spd^{3+} initially supports annealing of the staples with the scaffold allowing them to be folded into structures with predefined shapes, and thereby becomes incorporated inside of the closely packed origami structures in a similar way as assumed for DNA condensates. Spd^{3+} will not only interact with the origami structures, but also with the excessive staple strands remaining free in solution. Thus, the ratio between Spd^{3+} and the phosphates in the origami structures cannot be determined precisely, but it is roughly on the same order as in DNA condensation. The most prominent feature of Spd^{3+}-origami folding is that the Spd^{3+} concentration required for folding exhibits the same linear trend with DNA concentration as in the condensation experiments. The aggregation and precipitation of DNA nanostructures observed at higher Spd^{3+} concentrations indicates that the optimum Spd^{3+} concentration for origami folding might be just below or right at the condensation transition.

Further, high resolution TEM and AFM studies have shown that the DNA helices in origami structures come close only at the crossover points, whereas they otherwise avoid each other due to electrostatic repulsions [68, 146, 147]. The estimated distance between neighboring DNA helices in origami structures is in the range of $\approx$ 2.5–3 nm, which is actually very similar to the spacing in DNA condensates, for which the DNA duplexes were found packed on a hexagonal lattice with a Bragg spacing of $\approx$ 3 nm [68, 146–149]. In addition, the overall size of the Spd-nanostructures is not changed with respect to the Mg-folded structures thus cannot distinguish between conventional origami folding or internal "condensation" (figures 52-54, appendix).

3.4 Conclusion and outlook

The results demonstrate an interesting approach towards the "biocompatible" folding of DNA origami nanostructures – 42HB, tube and rectangle with different helix packing densities, using a naturally occurring DNA condensing agent, spermidine. The folding takes place at approximately physiological concentration of spermidine ($[Spd^{3+}] < 1$ mM), which is at least ten-fold lower than typical Mg^{2+} concentrations required for folding in conventional buffers. The results suggest that an optimum concentration of Spd^{3+} required for folding is at about half-compensation of the phosphate charges in the origami structures. At much higher concentrations, aggregation and precipitation of the origami structures occurs. Such effects show the resemblance of Spd^{3+}-supported origami folding to *in vitro* DNA condensation.

Spd^{3+}-folding is an interesting alternative to high ionic strength folding buffers, especially in situations where the presence of large concentrations of mono- or divalent cations is detrimental. In addition, Spd-nanostructures can potentially be employed in mimicking and understanding the DNA condensation process *in vivo* (figure 18). Spd^{3+} along with other polyamines have also been used as a component of non-viral gene delivery systems – either as simple cationic oligomers or in conjugation with lipids. The compaction of DNA with polyamines has been found to enhance its cellular uptake and thus improve transfection efficiency [150]. Thus, the advantage of Spd-nanostructures to be able to fold at low concentrations of Spd^{3+}; and the inherent ability of Spd^{3+} to compact DNA and increase its cellular uptake can be used to exploit and develop alternative methods of cellular transfection of artificial DNA nanostructures for potential biological applications (figure 18), as described in greater detail in chapter 4.

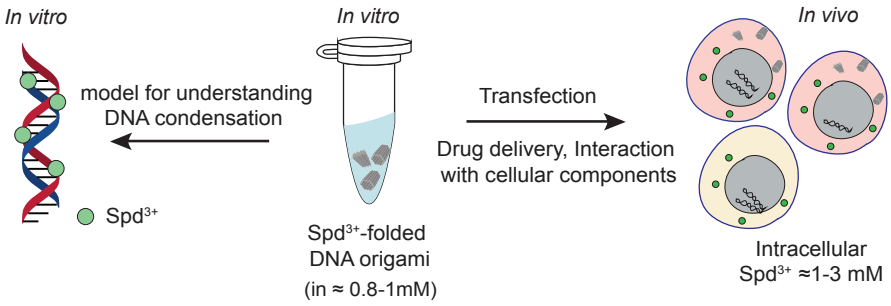


Figure 18: Applications of Spd-nanostructures.

4 Characterization and applications of Spd^{3+}-folded DNA origami structures

4.1 Introduction

The structural versatility and biocompatibility of DNA origami nanostructures have opened up ample opportunities for effective, tunable and targeted transport of molecules in cells for therapeutic and biosensing applications. In principle, a single DNA origami structure is capable of performing all tasks necessary for drug delivery or docking of molecules inside cells. The relative ease of DNA origami to adopt different configurations makes it easier to include and spatially organize different kinds of molecules which include cell-targeting ligands, drug molecules and functionalized sites. DNA nanostructures containing the intercalating anticancer drug doxorubicin and a logic-gated DNA origami "robot" that releases molecular payloads when triggered by cell surface proteins have been demonstrated [151, 152].

Numerous transfection experiments with DNA nanostructures have been performed via receptor mediated endocytosis, using nanostructures decorated with CpG sequences and viral proteins; and employing lipid based transfection agents [136–139]. However, efficient transfection is always challenging and there is a constant lookout for novel, more efficient and versatile approaches. In order to assess the fate of structures inside cells, stability assays which mimic the cellular environment *in vitro* have been performed and the endurance of structures was determined by techniques such as gel electrophoresis and microscopy [153].

In the context of this thesis, a well known technique– electroporation (or electrotransfection/electrotransfer) was explored for cellular transfection of DNA origami. It is defined as a process where electric fields are applied to cells to increase the permeability of the membrane by creation of transient pores that allow uptake of foreign material including chemicals, drugs or DNA. It offers the foreign material (mostly DNA) the advantage of escaping the endosome and going straight into the cytoplasm, as opposed to other methods which primarily work on the basis of receptor mediated endocytosis. This results in a more efficient uptake by cells [154]. One prerequisite for electroporation based experiments is the requirement of relatively low ionic strength and low conductivity buffers. In the previous chapter, we have shown the an alternative approach of assembling DNA origami nanostructures using

spermidine (Spd^{3+}), a trivalent DNA condensing agent. It was observed that Spd-nanostructures assembled at almost ten-fold lower ionic strengths in comparison to the DNA origami structure counterparts folded in conventional buffer containing Mg^{2+}, making Spd-nanostructures appropriate for electroporation experiments. In addition, Spd-nanostructures were assembled at physiological concentrations of Spd^{3+} found in mammalian cells as opposed to the high salt (Mg^{2+}) required in case of Mg-nanostructures, thereby suggesting Spd-nanostructures to be better suited for cellular transfection experiments in general.

In this chapter, the effect of different electric field strengths on the structural integrity of DNA origami structures (42HB) assembled in different buffers (Spd^{3+} and Mg^{2+}) is presented. Further, an alternative method of delivering DNA nanostructures to mammalian cells via electroporation is demonstrated. In the end, viability of cells to electroporation conditions and stability of DNA origami nanostructures in cell lysate are addressed.

4.2 Experimental details and methods

DNA nanostructures in electric fields

DNA nanostructures folded in Mg^{2+} containing buffers or in $C_7H_22N_3Cl_3$ (described in the previous chapter) and diluted to a concentration of 10 nM in the cell culture medium RPMI (Gibco, USA, conductivity $\approx$ 1.5 S/m) were subjected to different electric field pulses using a BTX electroporator ECM 399. The strength and duration of the pulses were varied by altering the length of the cuvettes, the voltage applied and the capacitance mode (C = 1050 µF or 36 µF) selected on the instrument. The electroporation instrument used was based on the capacitor discharge method [155]. In this method, a capacitor C (which can be selected to be 36 µF in high voltage (HV) mode and 1050 µF in low voltage (LV) mode) is first charged to a set voltage, and then discharged via a resistor R_i, which is connected in parallel to the sample cell (load resistance R_L). The voltage across the capacitor and thus the electric field in the sample cell therefore exponentially decays over time with a characteristic RC time τ, where the effective resistance R is given by $R = \frac{R_i R_L}{(R_i + R_L)}$.

Electroporation cuvettes with 0.1 cm electrode distance (electrode area 0.5 cm $\times$ 1.1 cm) were filled with a sample volume of 100 µl, while cuvettes with 0.4 cm

distance (electrode area 2 cm × 1.1 cm) were filled with 800 µl. Electric field pulse and electroporation experiments were performed within 5 minutes after dilution in order to avoid disintegration of the structures in the RPMI medium (with $[Mg^{2+}]$ ≈ 0.4 mM, RPMI is essentially Mg-free) [128].

Fluorescent labeling of DNA origami nanostructures

Both Mg- and Spd-nanostructures (42HB) were fluorescently labeled at n positions (n = 1, 7, 14, 21) via hybridization to Atto647N-modified strands. Assembled 42 helix bundle structures were incubated with 3X excess of modified strands. The sequences of the extension and dye labeled DNA strand are provided in the appendix. Excess DNA staples were removed by different purification strategies.

Purification of DNA origami nanostructures

Mg-nanostructures were purified by PEG purification using 15% PEG-8000 and 500 mM sodium chloride [85]. Spd-nanostructures were purified by size exclusion chromatography (SEC) using sephacryl resins (SR 300) (GE Healthcare Life Sciences) to remove excess DNA staples [86]. Sephacryl resins were equilibrated with MilliQ water and 50% v/v slurry was prepared. The resins were packed into Pierce centrifugal columns (0.8 ml, Invitrogen) by spinning the columns at 800 rcf for 3 minutes at 15°C. A volume of 100 µl unpurified Spd-nanostructures was subsequently passed sequentially through two spin columns packed with different resin volumes (400 and 260 µl respectively) and spun at 800 rcf for 3 minutes at 15°C. The maximum volume of sample to be purified was kept ∼ 25-30% of the first packed column. It resulted in a more efficient purification of the sample. The scale of purification was increased by using larger volume columns (5 ml), keeping all the other parameters constant. The purity of the product was verified by AGE.

Cell culture

Jurkat cells were grown as a suspension culture in tissue culture flasks (T-25, VWR) in growth medium (RPMI (Gibco, USA)) supplemented with 10% fetal bovine serum (FBS), 1% glutamax-I, and 1% penicillin and streptomycin. The cells were seeded

at a density of 3×10^5cells/ml (counted by hemocytometer using Trypan Blue (Thermofischer Scientific) cell viability assay [156]. The cells were grown at 37°C in a humidified atmosphere containing 5% CO2 until ∼ 70-80% confluency was attained. The cell culture was passaged every 2-3 days by replacing a portion of cells with fresh medium to lower the seeding density. This was carried out for 3-4 generations before the cells were used for transfection experiments.

Electrotransfer of DNA (ssDNA or Mg/Spd-nanostructures) in Jurkat cells

One day before electroporation, suspension culture of Jurkat cells was split in tissue culture flasks (T-25, VWR) in growth medium (RPMI (Gibco, USA)) supplemented with 10% FBS, 1% glutamax-I, and 1% penicillin and streptomycin. Cells were seeded in each flask at a density of 3×10^6 cells/ml. Cells were incubated overnight at 37°C in a humidified atmosphere containing 5% CO2. On the following day, cells were washed 2-3 times with PBS (Phosphate buffered saline) at 350 rcf for 5 minutes and resuspended in 650 µl RPMI (Gibco, USA) not supplemented with additional components and added to 0.4 cm electroporation cuvettes. A final concentration of 10 nM purified Atto647N labeled DNA nanostructures (Mg/Spd-nanostructures, n = 21) or Atto647N labeled single stranded DNA was added to cells. The cuvettes were kept on ice all the time. Electroporation was carried out at 200 V (low voltage mode, C = 1050 µF) using a BTX electroporator ECM 399. The electroporated cells were added to fully supplemented growth media as above and kept in an incubator at 37°C in a humidified atmosphere containing 5% CO2 for 3-4 hours to allow the cells to recover before imaging.

Lipofectamine-mediated transfection of Spd-nanostructures in Jurkat cells

One day before electroporation, suspension culture of Jurkat cells was split in tissue culture flasks (T-25, VWR) in growth medium (RPMI (Gibco, USA)) supplemented with 10% FBS, 1% glutamax-I, and 1% penicillin and streptomycin. Cells were seeded in each flask at a density of 3×10^6 cells/ml (counted by a hemocytometer). Cells were incubated overnight at 37°C in a humidified atmosphere containing 5% CO2. On the following day, cells were washed 2-3 times with PBS at 350 rcf for 5 minutes and resuspended in 650 µl Opti-MEM (serum free media, Gibco, USA)) not

supplemented with additional components. Lipofectamine-3000 (Invitrogen, USA) was diluted 50-fold in Opti-MEM (serum-free medium) (Gibco, USA) and mixed in 1:1 ratio with fluorescently labeled Spd-nanostructures that were also diluted in Opti-MEM. This lipofectamine-DNA mixture was left for 15 minutes at 37°C and then added to Jurkat cells already resuspended in Opti-MEM. The cells were kept in an incubator at 37°C in a humidified atmosphere containing 5% CO2 for 3-4 hours to allow the cells to recover before imaging.

Confocal fluorescence microscopy

Before imaging, the growth medium was removed and the cells were washed with PBS. The cells were treated with Benzonase nuclease (EMD Millipore) by incubating the cells for 30 minutes in a solution of PBS containing at least 0.5 U/ 1 µl of nuclease to dispense of DNA nanostructures outside of the cells prior to imaging. Cell organelles were stained in order to assess the intracellular localization of the DNA nanostructures introduced. LysoSensor Green (DND-189, Invitrogen, USA) and CellMask Orange Plasma membrane stain (Invitrogen, USA) were used to stain the lysosomes and cell membrane respectively according to the manufacturer's protocol. The lysosome stain was used at a final concentration of 100 nM in growth medium, RPMI. The cells were incubated in RPMI containing LysoSensor Green stain for 30 minutes. The cells were washed three times post staining with 1X PBS at 350 rcf for 5 minutes. Similarly, the cell membrane stain was used at a final concentration of 1X in RPMI. The staining was carried out for 5-10 minutes and the cells were washed three times in 1X PBS at 350 rcf for 5 minutes.

Post staining, Jurkat cells were supplied with fully supplemented growth medium and seeded in collagen coated IBIDI^{TM} chambers to facilitate their adherence for imaging. Live cells were imaged using a Leica SP II inverted fluorescence confocal microscope (Leica) in a sequential scan mode and analyzed using Fiji. Fluorescence filters used were Ex: 488 nm / Em: 510-540 nm for Lysosensor Green, Ex: 561 nm / Em: 565-600 nm for CellMask Orange Plasma membrane stain and Ex: 633 nm / Em: 660-700 nm Em for Atto647N labeled DNA nanostructures (Spd/Mg-nanostructures).

Cell viability assays

Two different assays, namely CyQuant Direct assay (Thermo Scientific) and PrestoBlue cell viability assay (Thermo Scientific) were used to evaluate the viability of Jurkat cells after being electrotransfected with DNA origami nanostructures. **CyQUANT Direct assay.** Jurkat cells (10^5 cells/ml) exposed to different conditions - no electroporation, electroporation without any DNA origami nanostructures, electroporation with 42HB, were seeded in triplicates (n = 3) by adding 100 µl of cell suspension into a 96-well plate for each time point (0, 24, 48, or 72 hours after seeding). At each time point, cell proliferation was measured by CyQUANT Direct assay following the manufacturer's protocol. Fluorescence was measured using a FLUOstar Omega plate reader at 530 nm.

PrestoBlue assay. Jurkat cells (10^6 cells/ml) subjected to different conditions as mentioned above were seeded in triplicates (n = 3) by adding 90 µl of cell suspension into a 96-well plate for each time point (0, 3, 6, or 24 hours after seeding). Cell viability was measured using PrestoBlue assay at each time point following the manufacturer's protocol. Fluorescence was measured on a using a FLUOstar Omega plate reader at 590 nm. The fluorescence observed at different time points for both the assays was normalized to their respective fluorescence readings at zero hour.

Stability of Spd-nanostructures in cell lysate

Cell lysis buffer (10X) was purchased from Cell Signaling Technonolgy (CST). Jurkat cell lysate (10^6 cells/ml) was prepared as per manufacturer's protocol. Cell lysates corresponding to different cell number was obtained by dilution of the lysate. The quality of cell lysate prepared was verified by standard BCA and Bradford assays. 20 nM 42HB Spd-nanostructures were incubated with lysate corresponding to different cell number for different durations (0 and 6 hours) at 37°C. The endurance of the Mg/Spd-nanostructures was then analysed by agarose gels and TEM.

4.3 Results and discussion

4.3.1 DNA nanostructures in electric fields

In order to exploit electrotransfer as an alternate strategy to introduce DNA origami into cells, the influence of a range of electric field strengths on the stability of DNA nanostructures was investigated. For electrotransfection experiments with mammalian cells, standard cell culture medium was used which has a relatively high conductivity of ≈ 1.5 S/m [157], but less than conductivity of standard folding buffers containing $MgCl_2$ and Spd^{3+} folding concentrations. Using the equation from above, this translates to resistance $R_L \approx 12\ \Omega$ and time constants of $\tau = 0.4$ milliseconds (ms) in HV mode and $\tau = 12$ milliseconds (ms) in LV mode. Similar conditions were employed for DNA origami structures. 42HB origami structures folded in $MgCl_2$ containing and Spd^{3+}-buffers were diluted in RPMI and subjected to high electrical field pulses generated by an electroporation apparatus based on the capacitor discharge method.

A comprehensive table showing the effect of an entire spectrum of electric field strengths on the stability of differently folded nanostructures is provided in figure 19. It shows the relation between different parameters which affect the electric field. The results were analysed by TEM and agarose gels. Spd-nanostructures sustained the exposure to different electric fields under all applied conditions, while Mg-nanostructures disintegrated in higher fields and for longer pulse durations as shown in figure 19. The range of stability of the structures was assessed using agarose gel electrophoresis, which clearly showed destruction of Mg-folded structures, while indicating that Spd-folded structures remained intact (figures 55-56, appendix).

According to figure 19 , Mg-folded 42HB structures were found to uphold the long ($\tau = 12$ ms) electric fields pulses only below $E \approx 125$ V/cm, and the short duration ($\tau = 0.4$ ms) pulses below $E \approx 4000$ V/cm. During these pulses, approximately the same amount of charge flows through the chamber, which is proportional to the product of the maximum field strength and the time constant, $E_{max} \cdot \tau$ (note that, roughly, $Q = \int I \cdot dt = \int j \cdot A \cdot dt = \int \kappa \cdot E \cdot A \cdot dt = \kappa \cdot E_{max} \cdot A \int \exp\left(\frac{-t}{\tau}\right) \cdot dt = \kappa \cdot A \cdot E_{max} \cdot \tau$, where j is the current density, κ is the conductivity and A is the electrode area).

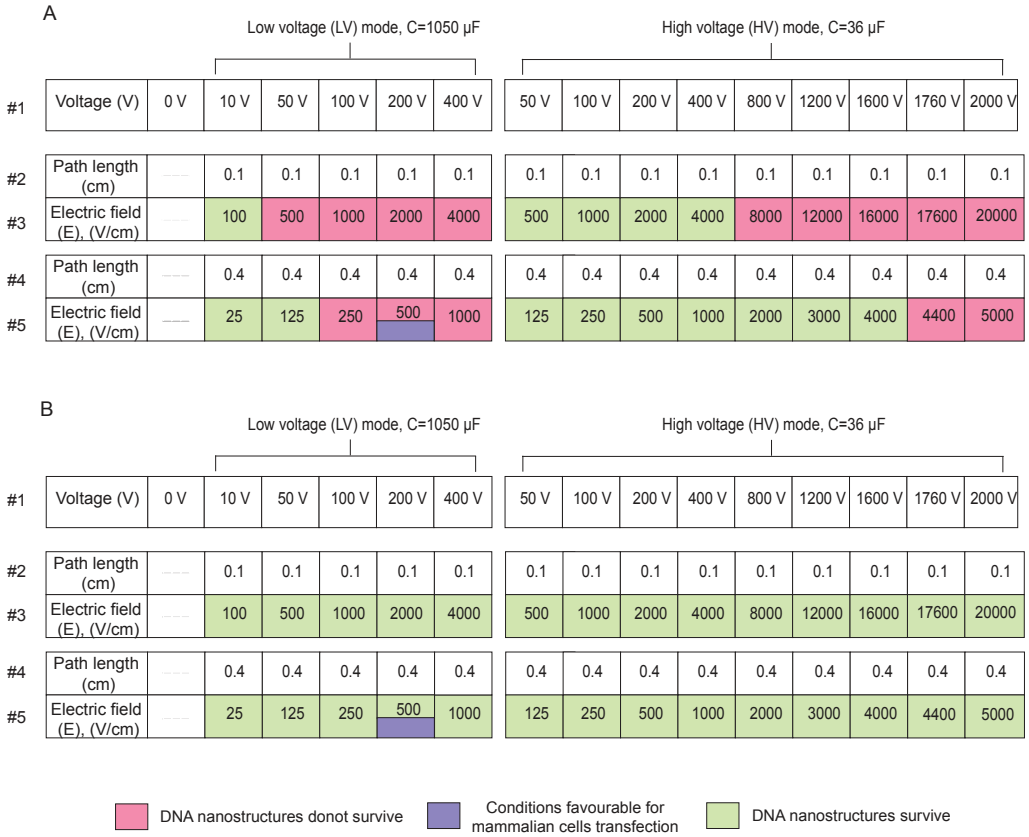


Figure 19: Effect of different electric field strengths on the stability of 42HB folded in different buffers. (A) Mg-nanostructures. (B) Spd-nanostructures. Mg/Spd-nanostructures diluted in RPMI and subjected to a range of electric field strengths both in LV and HV mode. The strength of the field is varied by using cuvettes of different path lengths, 0.1 and 0.4 cm. Voltages, path length and the corresponding electric field strengths analogous to the modes are specified in rows #1; #2, #4; #3, #5 respectively, in (A) and (B). The rows equivalent to electric field strengths (#3, #5) are color coded with respect to survival of nanostructures (pink/green) and conditions favorable for mammalian cell transfection (purple).

Mg-nanostructures failed to survive at conditions favorable for mammalian cell transfection. On the contrary, Spd-nanostructures survived all conditions. However, at electric field strengths of E $\approx$ 17600V/cm and higher their stability is compromised (figures 55-56, appendix). The electric field conditions favourable for electrotransfection of DNA into mammalian cells (e.g. Jurkat) i.e., using 0.4 cm electroporation cuvettes, relatively low voltages and longer pulse durations are shown in figure 20. As evidenced by TEM micrographs and corresponding agarose gels (figure 20 (B,F and D,H)) respectively, Mg-nanostructures indeed disintegrate

in the electric fields, while Spd-nanostructures seem to be unaffected. Bands corresponding to folded structures are observed for Spd-nanostructures (green box, 20 (H)) and not Mg-nanostructures (red box, 20 (D)) The tables in figure 20 (C and G) summarize the conditions used and the impact of the pulsed electric fields on the stability of Mg^{2+} and Spd^{3+}-folded 42HB.

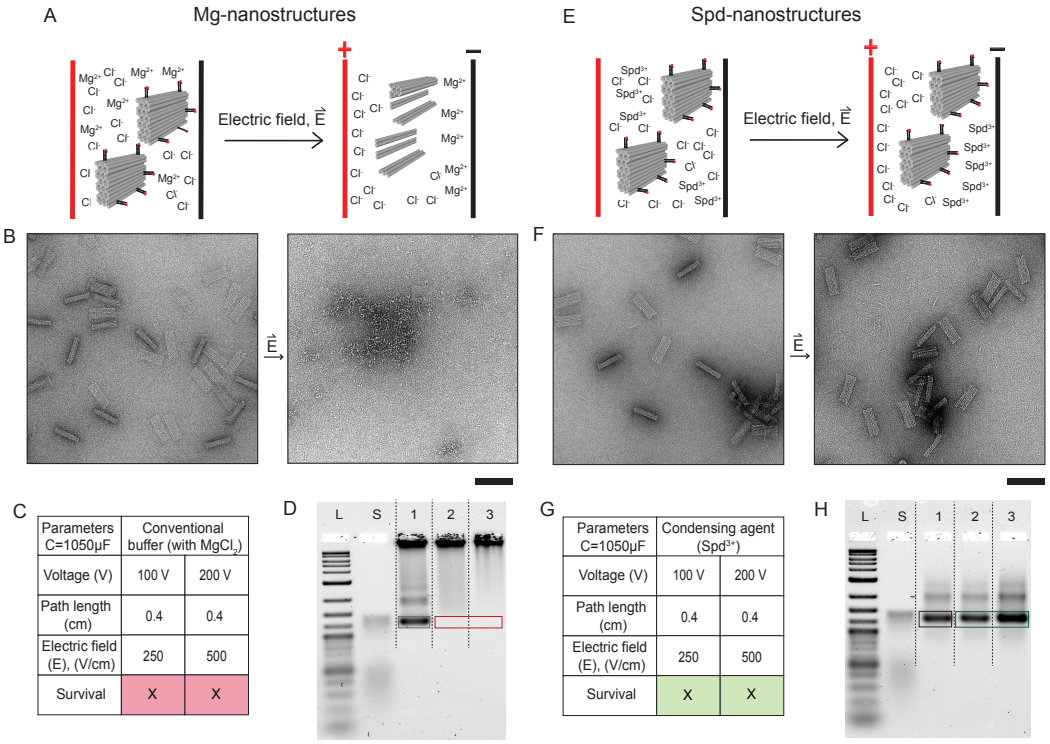


Figure 20: Stability of Mg/Spd nanostructures in electric fields favourable for electrotransfection in mammalian cells. (A, B) and (E, F). Schematic representation and illustrative TEM micrographs exhibiting the fate of DNA nanostructures (42HB) folded in $MgCl_2$ and Spd^{3+} respectively before and after application of electric field. (C, G). Tables with details of different electric fields applied, associated parameters and fate of the nanostructures. (D, H). Agarose gels showing the fate of Mg- and Spd-nanostructures (42HB variant) to different electric fields in comparison to controls. L: 2-Log ladder, S: scaffold, 1: control (no electric field), 2: E = 250 V/cm, 3: E = 500 V/cm. Scale bar: 100 nm.

In addition, electroporation parameters typically used for bacterial transformation were also investigated. Electroporation is known to have a higher efficiency of transformation in contrast to chemical methods in bacteria [158]. For bacteria, the conditions employed are slightly different in contrast to mammalian cells because of different cell wall composition (no cell wall in mammalian cells) and cell size.

Higher field strengths, using smaller 0.1 cm cuvettes, and shorter pulse lengths are applied. For the applications of strong electric field pulses, higher resistance samples are preferred to avoid large electrical currents or electrical breakthrough in the sample chamber, which could damage cells and instrumentation. So, Mg- or Spd-folded DNA origami structures (42HB) were diluted in a commonly used high resistance medium for bacterial electroporation (10% v/v glycerol in water) and subjected them to a range of electric field strengths, both in LV and in HV mode (figure 57, appendix). In the low conductivity case, $R \approx R_i$ and thus the time constants are increased to $\tau \approx 0.16$ seconds and $\tau \approx 5$ ms, respectively. As before, the Spd-folded structures sustained all pulse conditions, while Mg-nanostructures disintegrated at high field strengths and for longer pulse durations.

In the presence of an electric field, counterions are expected to flow around and also through the DNA origami structures. This effect has been previously studied by molecular dynamics simulations in the context of DNA-based membrane channels [159] and also for DNA origami plates attached to solid state nanopores [64] in a 50 mM $MgCl_2$/ 1 M KCl electrolyte solution. It is thus likely that for Mg-folded origami structures, the application of a strong electrical field pulse strips away the Mg^{2+} counterions that overcomes the negative charge thereby folding and stabilizing the structure. As a consequence, the structures disintegrate due to internal electrostatic repulsions, and the stripped counterions cannot be replaced by a sufficient amount of other divalent ions. By contrast, Spd^{3+} ions are more tightly attached to the DNA origami nanostructures and stabilize the structures even in the presence of a strong electric pulse. This is consistent with the view that the elongated polycation Spd^{3+} stabilizes the origami structures by making inter- and intrastrand polyelectrolyte bridges [53].

4.3.2 Modification and purification of Spd-nanostructures

Both Mg- and Spd-nanostructures (42HB) were fluorescently labeled with different numbers of fluorophores via hybridization to Atto647N-modified oligonucleotides to facilitate their detection and localization in eukaryotic cells (figure 58 (A)). Spd-nanostructures labeled with different number of fluorophores were analysed by agarose gel. Retarded migration of Spd^{3+}-origami was observed with the increase in number of modified positions on the structures (figure 58 (B)). 42HB Mg/Spd^{3+}-origami decorated with 21 fluorophores on each structure were further employed for

electrotransfection experiments.

It was found that the existing methods of DNA origami purification such as PEG precipitation [85], spin column filtration and gel extraction [86, 160] were not well suited for Spd-nanostructures. High sodium chloride concentrations used in PEG precipitation unfolded the Spd-nanostructures, potentially due to the competition between the monovalent and trivalent ions [38]. Spd-nanostructures interacted with and clogged spin column filter membranes, making them unsuitable for purification. Gel extraction resulted in contamination of the structures with ions contained in the gel buffer. A comparison of Spd-nanostructures purified by different protocols is given in figure 59 (appendix).

Size exclusion chromatography (SEC) was found to be most suitable for purification of Spd-nanostructures keeping them intact without changing the composition of the charge-compensating counterions. Schematic representation of the SEC process and corresponding agarose gel to check the quality and yield of purified labeled Spd-nanostructures is shown in figure 21.

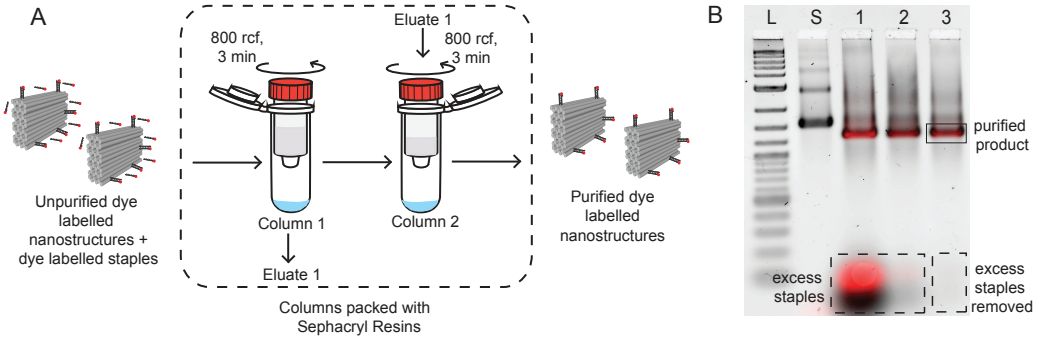


Figure 21: Purification of Atto647N labeled Spd-nanostructures (42HB variant) by size exclusion chromatography (SEC). (A) Schematic representation of SEC method using a series of two columns packed with different volumes of sephacryl resins. (B) AGE to check the quality and yield of purified products. Overlay of typhoon scanner image and SyBr gold stained image to identify the dye labeled DNA from unlabeled DNA). L-2log ladder, S-scaffold, 1: unpurified Spd-nanostructures, 2: Eluate 1, 3: Eluate 2. The recovery yield was found to be ≈ 85-90% (Eluate 2, band in well 3) (calculated from band intensities measured using Fiji [140].

4.3.3 Electrotransfection of DNA nanostructures in Jurkat cells

The stability of Spd-folded origami structures under electroporation conditions (E = 500V/cm, low voltage mode, C = 1050 µF) was further exploited for electrotransfection of fluorescently labeled 42 helix bundles (with Atto647N dye, n = 21) in Jurkat cells. Schematic representation of the electroporation process is shown in figure 22.

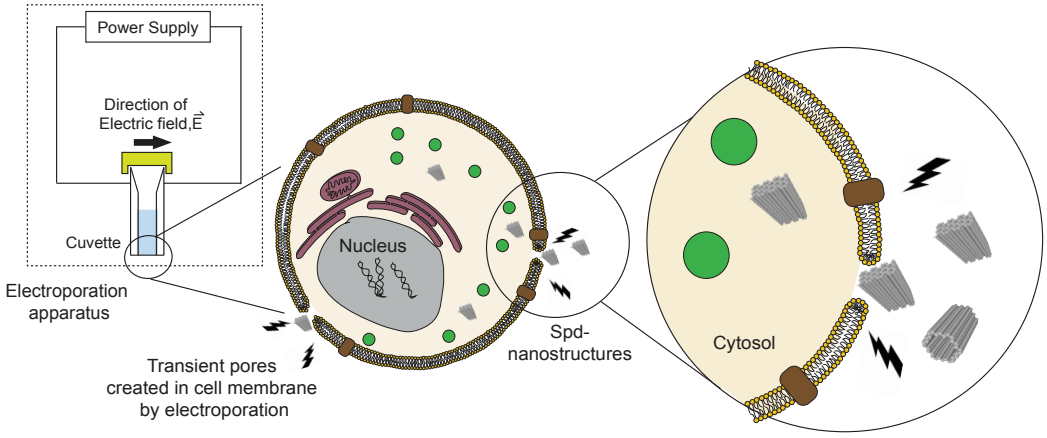


Figure 22: Schematic illustration of electrotransfer experiment.

Multiple co-localization of fluorophores favored better detection of signal by confocal microscopy. The confocal micrographs of live Jurkat cells electroporated with Atto647N-42HB at different z positions which represents the different volumetric sections of the cell are shown in figure 23 (B).

After electroporation, the nanostructures were found to be present in various subcellular volumes. The intracellular localization of the nanostructures was confirmed by cell organelle staining, in which lysosomes and cell membrane were stained in green and blue respectively. Blue blobs (shown by white arrows) correspond to tip the cell membrane which appears at different z positions for different cells due to variable cell size ($\sim$ 11-15 µm) of the cells and differential attachment of cells to the chamber. From the confocal images it was observed that Spd-nanostructures were found to be predominantly localized in the cytosol. Additional confocal microscopy images of live Jurkat cells electroporated with Atto647N-42HB at different z positions are provided in figure 60 (appendix). Also, internalization of Spd-nanostructures was observed at different time points and volumes in a single Jurkat cell (61, appendix).

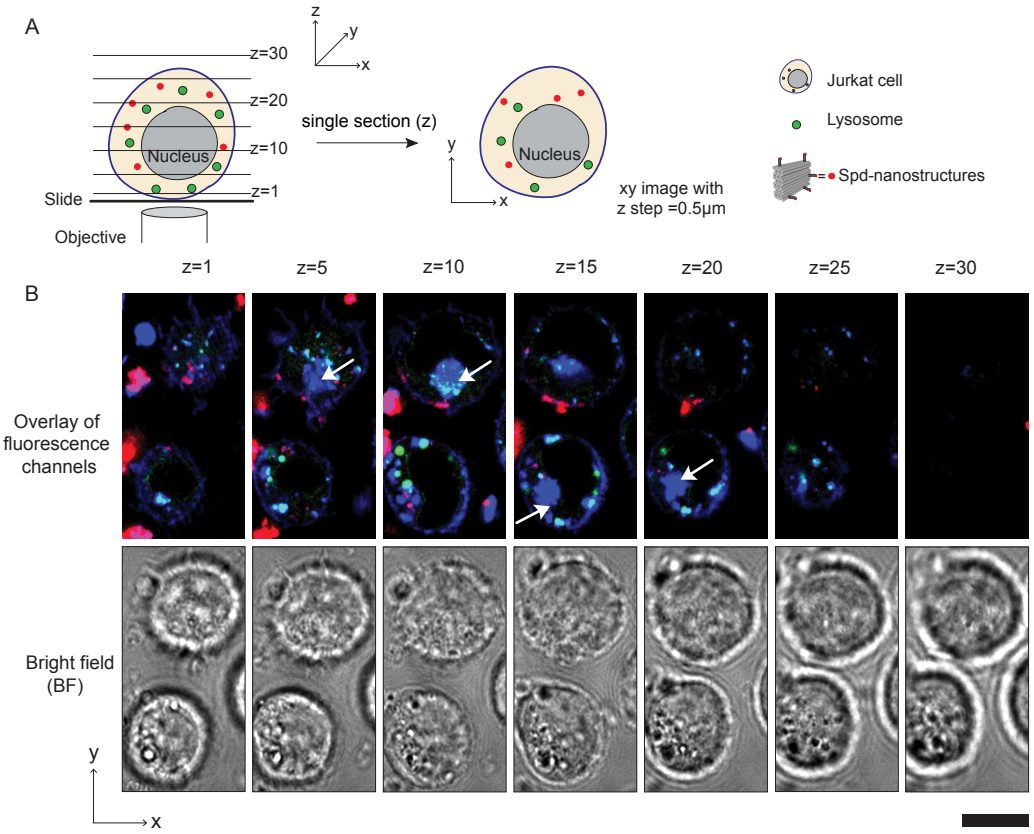


Figure 23: Electrotransfection of Spd-nanostructures in Jurkat cells. (A) Schematic representation illustrating the distribution of Spd-nanostructures at different subcellular z sections and typical images obtained by confocal microscopy. (B) Confocal micrographs of overlay of fluorescence channels (green, blue and red) and bright field (BF) were taken at different z sections (each z = 0.5 µm). The cell organelles lysosome (green) and cell membrane (blue) were stained to mark the boundaries of the cell and pinpoint location of the electroporated species (red, Spd-nanostructures). Blue blobs (shown by white arrows) correspond to tip the cell membrane which appears at different z positions for different cells due to variable cell size ($\sim$ 11-15 µm) of the cells and differential attachment of cells to the chamber. Scale bar: 10 µm.

By contrast, no incorporation of labeled Mg-origami structures (with Atto647N dye, n = 21) via electroporation was observed, suggesting that the nanostructures disintegrated when the electric field was applied (figure 24). In control experiments, when only dye labeled DNA strands were transfected using the aforementioned electrotransfection conditions, fluorescence of individual dispersed DNA strands could not be detected by confocal microscopy (figure 25).

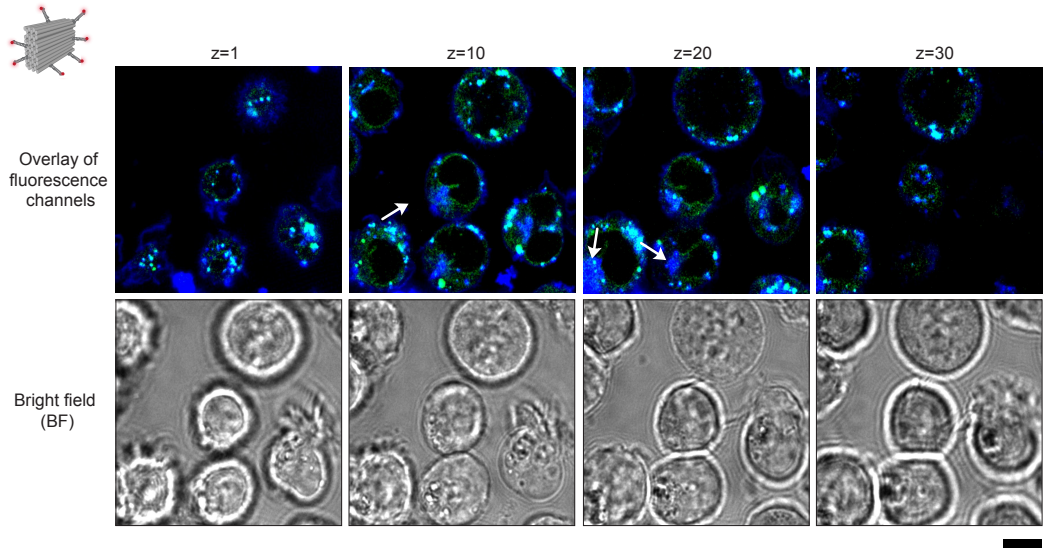


Figure 24: Electrotransfer of Mg-nanostructures in Jurkat cells. Confocal micrographs of overlay of fluorescence channels (green, blue and red) and bright field (BF) were taken at different z sections (each z = 0.5 µm). The cell organelles were stained as mentioned above to mark the boundaries of the cell and pinpoint location of the electroporated species (red, Mg-nanostructures). Scale bar: 10 µm.

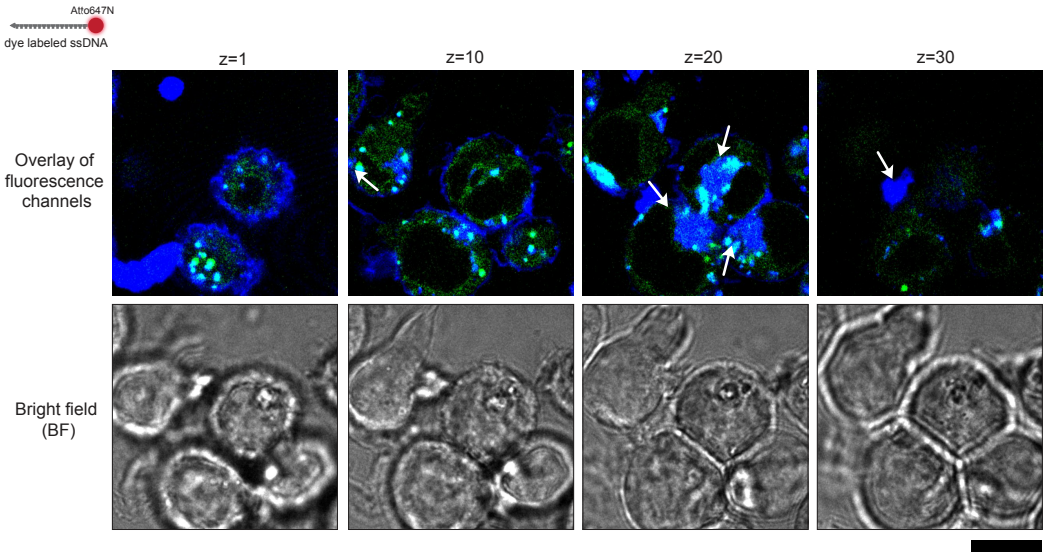


Figure 25: Electroporation of with Atto647N labeled ssDNA strands in Jurkat cells. Confocal micrographs of overlay of fluorescence channels (green, blue and red) and bright field (BF) were taken at different z positions. The cell organelles were stained as mentioned above to mark the boundaries of the cell and pinpoint location of the electroporated species (red, Atto647N labeled DNA strands). Scale bar: 10 µm.

In cases where cells did not survive due to irreversible rupture of the cell membrane, large amount of DNA nanostructures potentially binding to the ruptured membrane were observed.

Also, lipofectamine mediated transfection of Spd-nanostructures in Jurkat cells was carried out. Spd-nanostructures formed complexes with lipofectamine as seen in TEM images in figure 62 (A) (appendix) and were found to be present at different volumes of the cell as seen in figure 62 (B) (appendix). As mentioned in figure 23, the blue blobs (shown by white arrows in figures 24-25, 62) also correspond to tip the cell membrane which appears at different z positions for different cells due to variable cell size ($\sim$ 11-15 µm) of the cells and differential attachment of cells to the chamber.

4.3.4 Cell viability assays

Two methods were used to assay the viability of Jurkat cells exposed to the above mentioned conditions. Both the techniques determined the cell viability by orthogonal mechanisms. CyQuant direct assay was used to measure the viability of the cells over a period of 72 hours (or 3 days) and PrestoBlue was used for a time span of 24 hours.

CyQUANT Direct assay is based on a live cell-permeant DNA-binding dye (Ex = 480 nm, Em = 530 nm) in combination with an impermeable suppression reagent. The suppression agent blocks staining of dead cells and cells with compromised cell membranes, causing only staining of healthy cells with intact membranes. The increase in fluorescence of DNA-binding dye is proportional to the increasing cell number over time, thus confirming cell proliferation and growth. The schematic representation of the assay and effect of different experimental conditions on the fluorescence observed as a result of DNA binding over time is shown in figure 26. It was observed that the fluorescence signal increased with time when the cells were subjected to different experimental conditions (No electroporation and electroporation with and without origami). This suggested an increase in cell number confirming cell proliferation and growth. It validates that the cells survived blank electroporation as well as electroporation with origami structures. The fluorescence at different time points was normalized to the fluorescence values at zero hour.

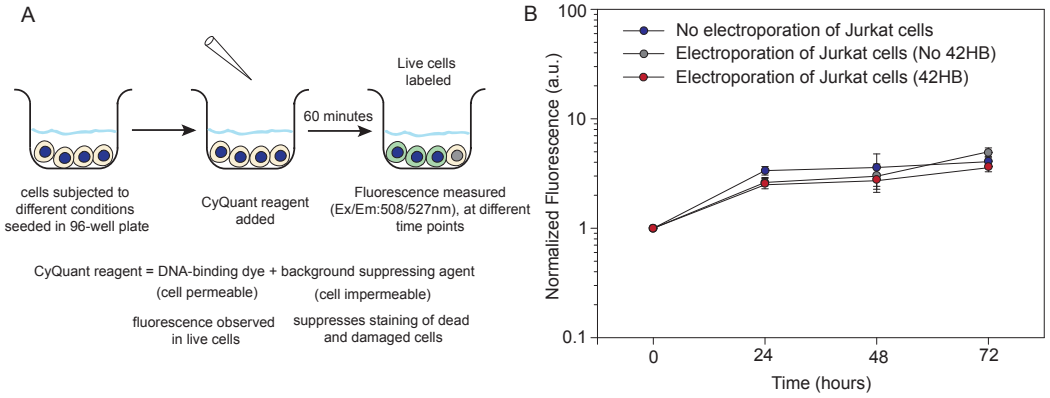


Figure 26: CyQUANT cell viability assay. (A) Schematic representation of the steps involved. (B) The viability of Jurkat cells, seeded at different densities for the assays – 10^4 cells/ml subjected to different conditions– no electroporation, electroporation without any DNA origami nanostructures and electroporation with 42HB was analyzed.

PrestoBlue assay uses a cell permeable non fluorescent, blue colored resazurin-based reagent (Ex = 569 nm / Em = 586 nm) which functions as a cell viability indicator. The PrestoBlue reagent gets modified to a red colored highly fluorescent compound, resorufin, by the reducing environment of the viable cells. The increase in fluorescence is proportional to the increasing cell number over time, thus corroborating cell viability. Figure 27 (A) shows a schematic representation of the chemical reaction on which the assay is based. Cell samples exposed to the same conditions as described in the previous assay were used. The fluorescence was measured over a period of 24 hours. It was observed that the fluorescence increased with time suggesting an increase in cell number, thereby confirming cell viability in the electrotransfected samples also (figure 27 (B)). The fluorescence at different time points was normalized to the fluorescence values at zero hour.

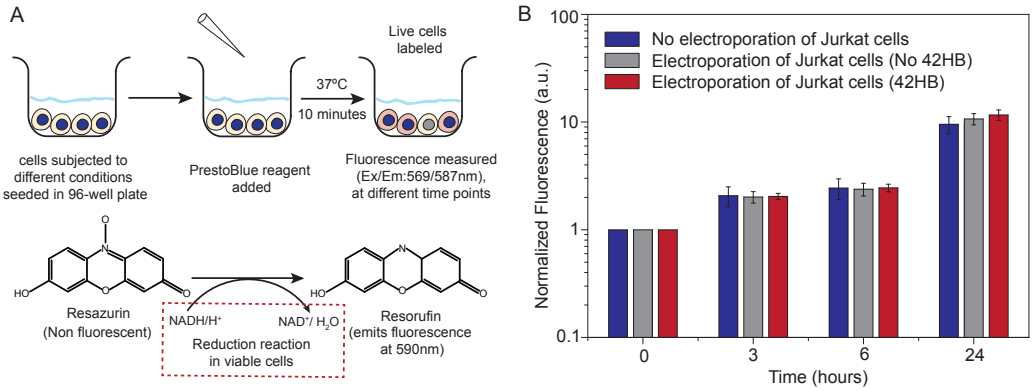


Figure 27: PrestoBlue cell viability assay. (A) Schematic representation of the steps involved. (B) The viability of Jurkat cells, seeded at different densities for the assays – 10^6 cells/ml subjected to different conditions– no electroporation, electroporation without any DNA origami nanostructures and electroporation with 42HB was analyzed.

Both the assays independently gave concordant results in terms of viability of cells. This confirmed that Jurkat cells survive electroporation pulses, electrotransfection of origami structures and remain viable.

4.3.5 Stability of Spd-nanostructures in cell lysate

The question of intracellular integrity of the Spd-nanostructures could not be addressed completely by confocal experiments. Therefore, stability of origami structures in cell lysate as previously described in [153] were performed. The lysate corresponding to cell densities 10^3 and 10^4 cells/ml were used to study the stability of Spd-nanostructures. The survival of DNA nanostructures was analysed by TEM and agarose gels. The nanostructures were in fact found to be stable in the cell lysate corresponding to different densities (cell number/ml) as seen in figure 28.

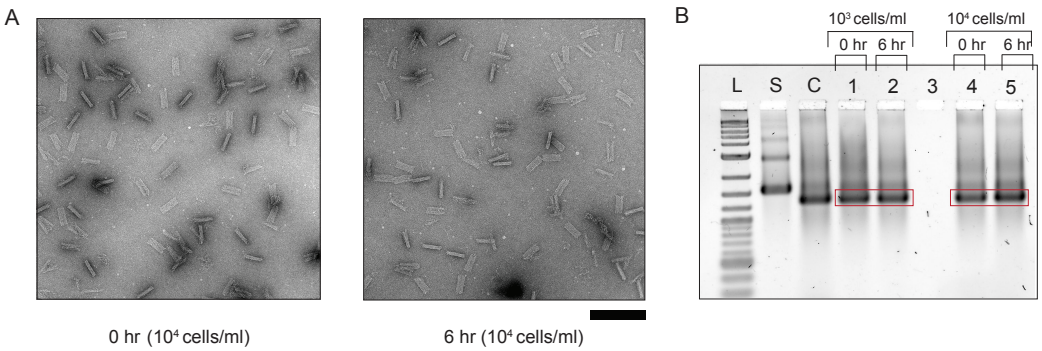


Figure 28: Stability of Spd-nanostructures (42HB) in Jurkat cell lysate. (A) TEM micrographs showing the stability of 42HB in Jurkat lysate at 2 different time points (0 hour and 6 hour). (B) Agarose gel showing the stability of nanostructures in cell lysates corresponding to different cell densities. L: 2log ladder C: Control Spd-nanostructures without lysate, 1: 0 hour incubation (10^3 cells/ml), 2: 6 hour incubation (10^3 cells/ml) 3: Jurkat lysate, 4: 6 hour incubation (10^4 cells/ml), 5: 0 hour incubation (10^4cells/ml). Scale bar: 200 nm.

4.4 Conclusion and outlook

An alternative strategy for transfection of DNA nanostructures using electric pulses is investigated. It exploits the feature of increased resistance of Spd-nanostructures to electric fields suitable for electroporation of DNA into mammalian cells. The results obtained convincingly show the robustness of Spd^{3+}-folded origami structures to different electric field pulses as they are typically applied in electrotransfection experiments. Spd-nanostructures exhibit an increased stability under all conditions investigated, whereas Mg-nanostructures disintegrate at all electric field strengths and pulse durations relevant for cellular delivery via electroporation. In addition, successful electrotransfection of Spd^{3+}-folded origami into Jurkat cells is demonstrated. Experiments confirming the stability of Spd-nanostructures in cell lysates further imply that these Spd-nanostructures should also be stable in intracellular environments; and could potentially be used for biological applications. Notably, the physiological concentrations of Spd^{3+} is on the same order as that used in origami folding experiments. Alongside, it is shown that the viability of cells is not compromised either by electric field pulses or by presence of Spd-nanostructures inside cells.

Spermidine appears to stabilize DNA origami structures in low conductivity buffers as well as in electrical fields, which could be of general interest for electric field

manipulation of DNA origami objects. The results demonstrate that with origami nanostructures folded in and stabilized by the polyamine spermidine, electroporation could be an interesting alternative approach towards intracellular delivery of such nanostructures, potentially providing access to other subcellular compartments than typically achieved with the other methods. Electrotransfection could be used for delivery of functionalized Spd-nanostructures (carrying drugs, other molecular cargo) which interact with intracellular components to perform predefined functions in cells. Further, Spd-nanostructures assemble at natural concentrations of Spd^{+3}, which could be potentially be exploited for creation of nucleic acid nanostructures *in vivo*.

5 Towards 3D RNA nanostructures: design and synthesis

5.1 Introduction

The classical view of biology envisions RNA primarily in three main roles, as a messenger (mRNA) that shuttles information between DNA and proteins, as an adaptor (tRNA) that translates the information stored in mRNA into functional proteins and as a structural component (rRNA) of the ribosome. Over the years, this conventional image of RNA has changed, as it has been found to be involved in many other essential functions in the cell. Non coding RNAs (ncRNAs) belonging to various classes including ribozymes, riboswitches, and RNA–protein complexes have been involved in regulatory functions at transcriptional and translational levels, genome-editing and chromatin remodelling etc [161] (Figure 29).

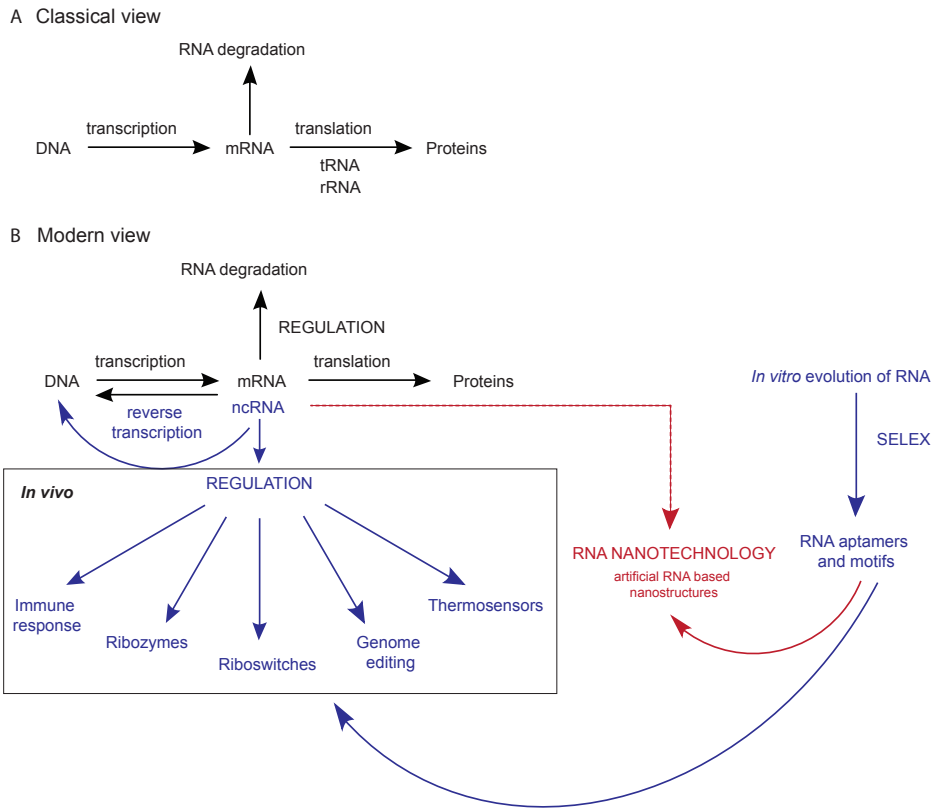


Figure 29: Classical and modern view of RNA.

These naturally occurring ncRNAs fold and assemble into complex 3D architectures via a plethora of secondary and tertiary interactions that endow RNA with the ability to perform a range of functions. Misfolding and mutations in RNA have been associated with many diseases, for e.g. Huntington's disease, myotonic dystrophy and cancer etc [162]. Therefore, an understanding of RNA structures and their dynamic regulation has become an integral aspect of interpreting RNA function.

This exceptional folding capability and modularity of biological RNAs have provided insights and created huge opportunities for construction and self-assembly of novel artificial RNA nanostructures. This has led to the emergence of the field of RNA nanotechnology, akin to the more established DNA nanotechnology. RNA offers several advantages over DNA: they can be transcribed in large amounts, they include a diverse variety of naturally occurring tertiary motifs and functional domains [100, 101], can also be genetically encoded and expressed in cells [108, 109]. Rationally designed RNA nanostructures and devices have great potential for applications in synthetic biology, metabolic engineering, and nanomedicine [5].

Numerous RNA nanostructures have been designed using different approaches as discussed in section 2.4.2. A technique known as ssRNA origami has been developed, that involves design of artificial RNA structures folding from a single strand [105]. It consists of precise organization of antiparallel RNA helices using a variety of RNA tertiary motifs to generate predefined ssRNA tiles and lattices and also employs a dove tail crossover pattern to extend the architecture to more than two helices making the method scalable. Different types of tiles have been created by varying the spacing between the two antiparallel crossovers – "antiparallel odd" (RNA-AO) or "antiparallel even" (RNA-AE) where odd and even refer to the number of helical half-turns between crossovers. RNA nanostructures up to six helices tall and 660 nt in length have been constructed using this approach. ssRNA origami tiles have been used in combination with fluorogenic RNA aptamers for the construction of nanoscale aptamer based FRET sensor [163]. Such sensors have been shown to function both *in vitro* and *in vivo*.

A considerable amount of effort has been put into RNA design and assembly strategies to increase the scale and variety of shapes constructed keeping a balance between the kinetic and thermodynamic influences of RNA folding and assembly. Novel "tecto RNA" units have been created comprising of two or more tertiary RNA motifs. More complex, responsive and distinctive shapes have been further

realised by combination of same or different newly derived "tecto RNA" units [164].

In this chapter, a novel three helix "antiparallel even" RNA tile (3H-AE) based on the ssRNA origami approach has been constructed that assembles into hexagonal lattices. Further, modifications in the three helix design have been introduced for extending the tiles into the third dimension. Two RNA motifs namely domain IIa and subdomain IIa of Interal ribosomal entry site (IRES) in Hepatitis C virus (HCV) have been used for almost perpendicular arrangement of malachite green aptamer (MGA) onto the RNA tile. Additionally, 3H-AE tiles with two aptamers (MGA and spinach) perpendicular and extending from either side of the tile have been designed and synthesized.

5.2 Experimental details and methods

Design of three-helix RNA tiles (3H-AE)

3H-AE and its subsequent modifications were designed by the ssRNA origami method that has been previously described in great detail in [105] and [165]. Briefly, using the 3D modeling programs Swiss-PdbViewer [166] and UCSF Chimera [30], three standard A-form RNA double helices were positioned over one another and rotated to create an optimal spacing for an anti-parallel even (AE) double crossover (DX) (or dovetail crossover). An internal 180° KL (HIV-1 DIS, PDB ID: 2B8R) was placed between the crossovers on helix number 1 (top). UUCG tetraloops (extracted from PDB ID: 1F7Y) were positioned at the ends and in between crossovers on helix number 2 (middle). 120° KL (RNA i/ii inverse loop, PDB ID: 2BJ2) were positioned at the ends of double helices number 1 and 3 for formation of hexagonal lattices. The 180° KL forms 6 base pairs between the two loops resulting in a coaxial stack and is in phase with the A-form helix of RNA whereas the 120° KL forms 7 base pairs between the loops resulting in a continuous, but curved coaxial stack. Modifications in 3H-AE design included additional RNA motifs. Subdomain IIa (PDB ID: 2PN4) and domain IIa (PDB ID: 1P5M) of IRES in HCV virus were used as connectors between the RNA tile and MGA (PDB ID: 1F1T) in modified versions 3H-AE-MGapt V1 and 3H-AE-MGapt V2 respectively. These domains consist of a 90° angle which allows for almost perpendicular arrangement of the added RNA to the tile. Additional RNA sequence encoding spinach aptamer (PDB ID: 4KZD)

was connected to 3H-AE-MGapt V1 by subdomain IIa to generate a new modified structure 3H-AE-MG-spinach-apt.

After the initial modeling, the 3D structures were ligated with a Perl script ("ligate.pl", which was available from `www.andersen-lab.dk`) and refined using a recursive geometric refinement function in Assemble [116]. 3D models in this work were rendered in UCSF Chimera. The designed structure was further traced using a Perl script ("trace.pl", which was also available from `www.andersen-lab.dk`) and an input was generated, that was used for the design of corresponding DNA sequences in Nupack [81]. The sequences of the 180° KL, 120°, tetraloops, subdomain IIa, domain IIa and MGA and spinach aptamer were chosen from the PDB files and added to the designs. Additionally, some base pairs in dovetail crossovers and adjacent base pairs were constrained to be strongly stacking G-C pairs in an attempt to specify which conformation of the junction would be preferred. Sequences were further constrained to contain at least one GU wobble pair per every eight continuous base pairs in order to avoid secondary structure in the RNA-encoding DNA template and simplification of its synthesis. Such positions were specified by using the letter 'K'. All the remaining positions were designed by Nupack. The $5'$ end of each sequence was constrained to begin with GGG, an optimal leader sequence for T7RNAP. Primers were generated specific to DNA sequence generated by Nupack and Tm was calculated using NEB Tm calculator (`www.tmcalculator.neb.com`). A detailed step by step protocol for the design of ssRNA origami structures is provided in figure 30.

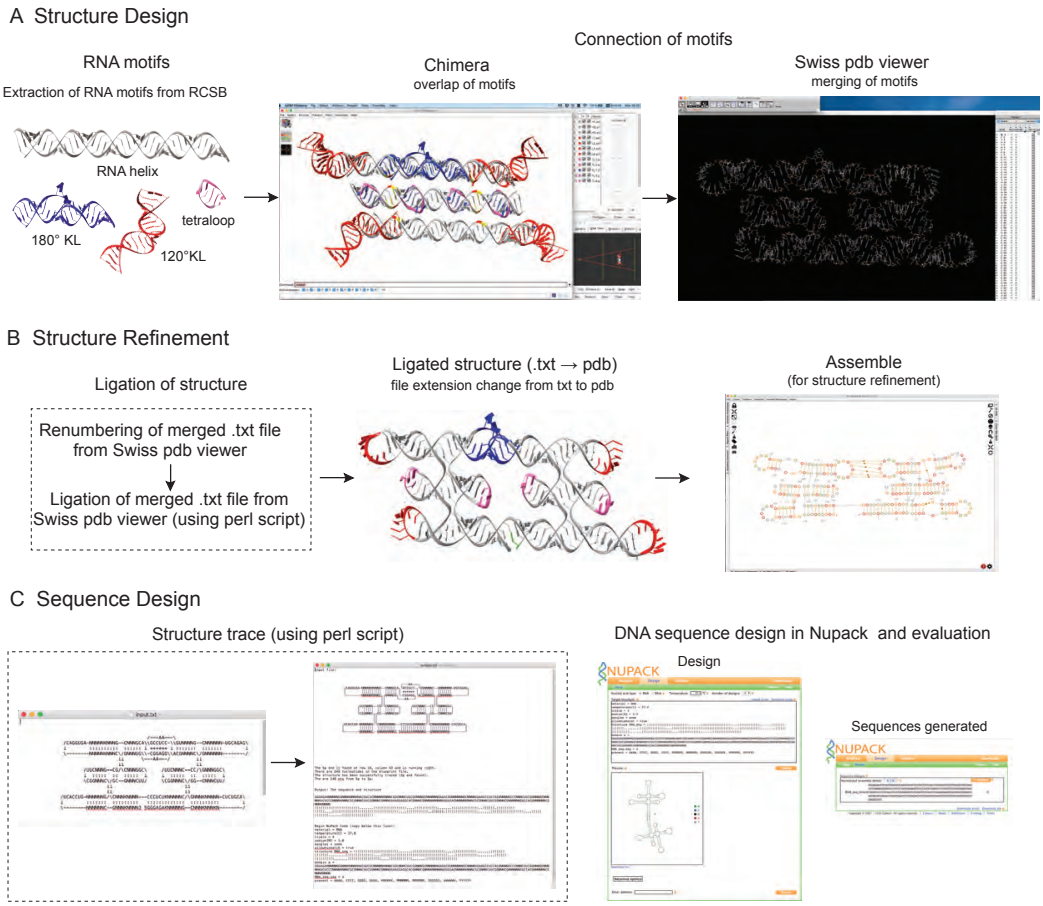


Figure 30: Schematic representation of the steps involved in design of ssRNA origami. (A) Structure design. (B) Structure refinement. (C) Sequence design.

Polymerase chain reaction (PCR)

DNA templates for all RNA designs were ordered as "custom gBlock" dsDNA oligos from Integrated DNA Technologies (IDT). Custom gBlocks were supplied as 250 ng of lyophilized DNA. DNA gBlocks were reconstituted in nuclease free water (nF water) to a concentration of 5 ng/1 µl and were amplified by Phusion PCR master mix (2X) from New England Biolab (NEB) over 25 cycles using primers specific to the 5′ and 3′ ends of the template in a Mastercycler (Eppendorf). Amplified DNAs were purified using a standard Monarch PCR Cleanup kit from NEB and suspended in nF water. PCR products were verified using 2% agarose gel containing SERVA DNA stain G (Serva). The gel was run in 1X TAE at room temperature

at 120 V for 45 minutes. Gel images were taken on a Typhoon FLV 9000 laser scanner and analyzed with the software package Fiji [140]. DNA gBlocks and primer sequences for all RNA designs along with the corresponding conditions used for PCR amplification are listed in the appendix (Figures 74-77).

In vitro transcription

RNA tiles for different experiments were transcribed *in vitro* using T7 RNA polymerase (T7RNAP) and purified. $\approx$ 100 ng/1 µl of template DNA was mixed in a transcription reaction containing 1X RNA polymerase buffer (RNAP buffer) and 16 mM rNTPs (4 mM each) from NEB. 15 mM $MgCl_2$ and RNase Inhibitor (NEB) ($\sim$ U/1 µl) were also added prior to addition of T7 RNAP provided by S. Sagredo (Simmel lab, TUM) ($\sim$ 200 nM). Reactions were carried out in 50 µl volumes at 37°C for 3-4 hours. For RNA designs with aptamer modifications, non-fluorescent chemical compounds Malachite green (MG) (10 µM) and DFHBI (3,5-difluoro-4-hydroxybenzylidene imidazolinone, 10 µM) respectively were also added that emitted fluorescence in presence of the respective MG and spinach aptamers. Additionally, 125 mM potassium chloride (KCl) was added during transcription of the RNA tile with spinach aptamer modification. Transcription of RNA tiles with modifications was monitored with time by measuring the fluorescence of MG and DFHBI using iQ5 real-time PCR detection system from Bio-rad using filters with excitation and emission wavelengths corresponding to the fluorphores (Ex = 475-495 nm, Em = 515-545 nm for DFHBI and Ex = 615-645 nm, Em = 669-699 nm for MG).

Purification

Prior to purification, the transcription reaction was stopped by addition of 0.08 U/1 µl DNaseI (NEB) and incubated for 30 minutes at 37°C. DNaseI stopped the reaction by degrading the template DNA. Purification was done by standard phenol-chloroform extraction using Phase Lock Gel (PLG) heavy columns from QuantaBio. RNAs were precipitated by 2-3 volumes 100% ethanol (EtOH) and 0.1 volume 5M sodium acetate (pH 5.5) for atleast one hour followed by subsequent washing in 70% EtOH, drying under vacuum and resuspension in 20 µl nF water. RNA concentration was estimated by denaturing polyacrylamide gel electrophoresis (denaturing PAGE).

Polyacrylamide gel electrophoresis (PAGE)

8% Acrylamide:bis (29:1) gel containing 8 M urea and 1X TBE buffer (denaturing PAGE) was used to analyse the purified RNA tiles and estimate their concentration. The gel was pre-run at 50°C for 30 minutes. The samples and RiboRuler Low Range Ladder (Thermofisher scientific) were prepared in 1:1 ratio with 2X RNA loading dye (Thermofisher scientific) and incubated at 80°C for 10 minutes prior to loading onto the gel. The gels were run at 80 V for 1 hour 45 minutes. After electrophoresis, the gels were stained with SYBR^{TM} Green II RNA Gel Stain (Thermofisher scientific) for 10 minutes and scanned on a Typhoon FLV 9000 laser scanner and analyzed with the software package Fiji [140].

Sample preparation and atomic force microscopy (AFM)

Freshly-cleaved mica affixed to a metal disk was placed inside an empty tip box containing a wet kimwipe and water (to maintain the humidity in the chamber). 60 µl drop of 1.25X AFM buffer (1X AFM buffer = Tris-Borate 1X (pH 8.1), 2 mM magnesium acetate ($Mg(OAc)_2$), 50 mM potassium chloride (KCl), 50 mM sodium chloride (NaCl)) was added to the mica and incubated at 37°C. Additional purification of RNA tiles was carried out to separate the desired product from abortive transcripts using Quantum Prep Freeze 'N Squeeze gel extraction spin columns from Bio-rad. Purified RNA tiles were diluted in pure water at a concentration calibrated to give coverage, but not large aggregates, on the mica surface ($\sim$ 75-100 nM for 3H tiles). Tiles were next subjected to a heat denaturation/renaturation protocol (95°C for 3 minutes and then snap-cooled to 4°C for 5 minutes). 15 µl of RNA sample was added to the mica on top of the pre-heated (37°C) buffer drop and incubated for 45 minutes to 1 hour to allow coverage of the surface and formation of hexagonal lattices. AFM images were collected in tapping mode under buffer using Asylum research's high resolution fast scanning AFM Cypher S. Olympus BL-AC40TS-C2 silicon nitride probes with a spring constant of $\sim$ 0.09 N/m and at a resonant frequency of $\sim$ 20-30 kHz were typically used for imaging, with a drive frequency of $\sim$ 5-8 kHz.

5.3 Results and discussion

5.3.1 Design and folding of 3H-AE RNA tile

A new antiparallel even RNA tile (RNA-AE) was designed and constructed. It consisted of three RNA helices placed over one another connected via dove tail crossovers and made into a cotranscriptionally folding continuous ssRNA using one 180° KLs, four 120° KLs and four tetraloops (figure 31 (A)). Further, modifications of 3H-AE have been created and discussed in sections 5.3.2 and 5.3.3. The schematic representation of the steps and softwares involved in design and transcription of 3H-AE (and its modifications) is provided in figure 31 (A and B). DNA gBlocks for 3H-AE RNA tile and its subsequent modifications were PCR amplified using Phusion High-Fidelity PCR Master mix (with HF or GC buffer). The purity and amplification was verified by 2% agarose gel (figure 63, appendix).

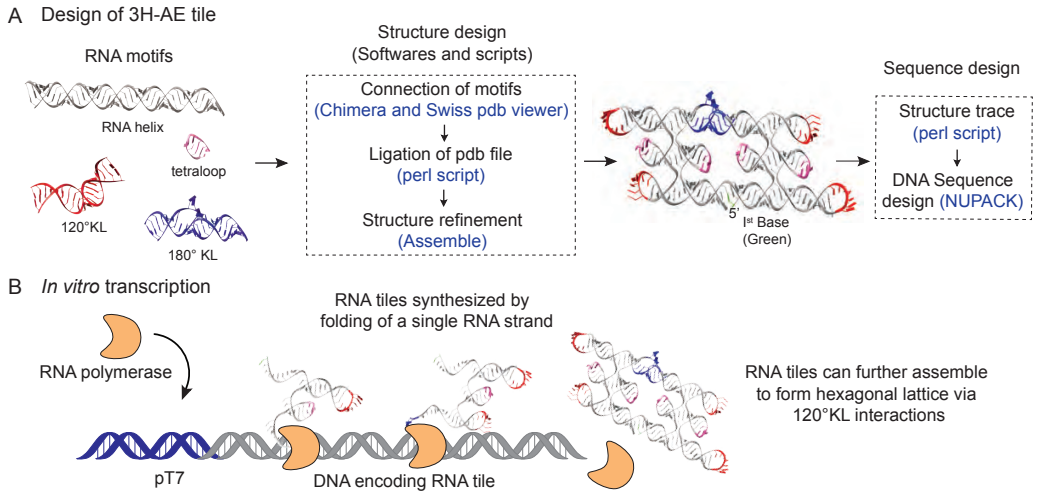


Figure 31: Schematic representation of steps involved in design and synthesis of ssRNA origami. (A) Design. (B) *In vitro* transcription of 3H-AE RNA tile.

Transcription and folding of RNA structures (3H-AE RNA tile, 248 nt, figure 32 (A)) was verified by denaturing PAGE and AFM respectively. The transcribed RNA was observed at the correct length i.e. 248 nt as shown in figure 32 (B). The precise folding of structure and subsequent interactions between 120° KLs resulting in formation of hexagonal lattices are shown in figures 32 (D) and 33 (B-E).

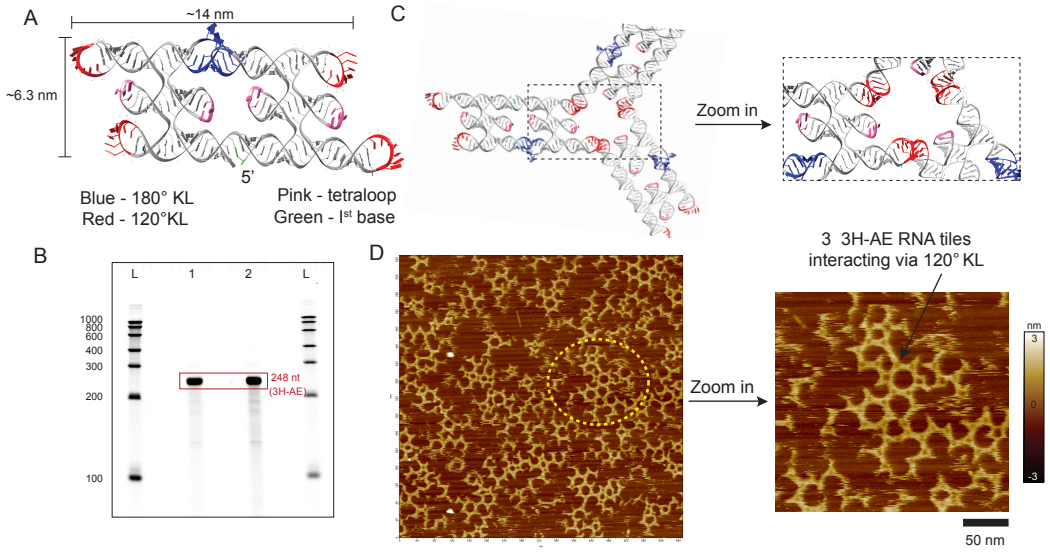


Figure 32: Folding of designed 3H-AE RNA tile. (A) Schematic representation of designed 3H-AE tile. (B) Denaturing PAGE gel showing the correct length of 3H-AE RNA tiles. L: LowRange Riboruler, 1-2: RNA transcribed from DNA template amplified from Phusion High-Fidelity Master mix with HF or GC buffer respectively. (C) Interaction of three tiles via 120° KL. (D) AFM images showing the correct assembly and interaction of tiles. Yellow dashed region further zoomed in to show the interaction of 120° KL.

120° KL interactions between individual tiles resulted in formation of hexagonal lattices ranging from few nm to hundreds of nm as observed by AFM imaging. The dimensions of the RNA tiles observed were also in accordance with the design (figure 33).

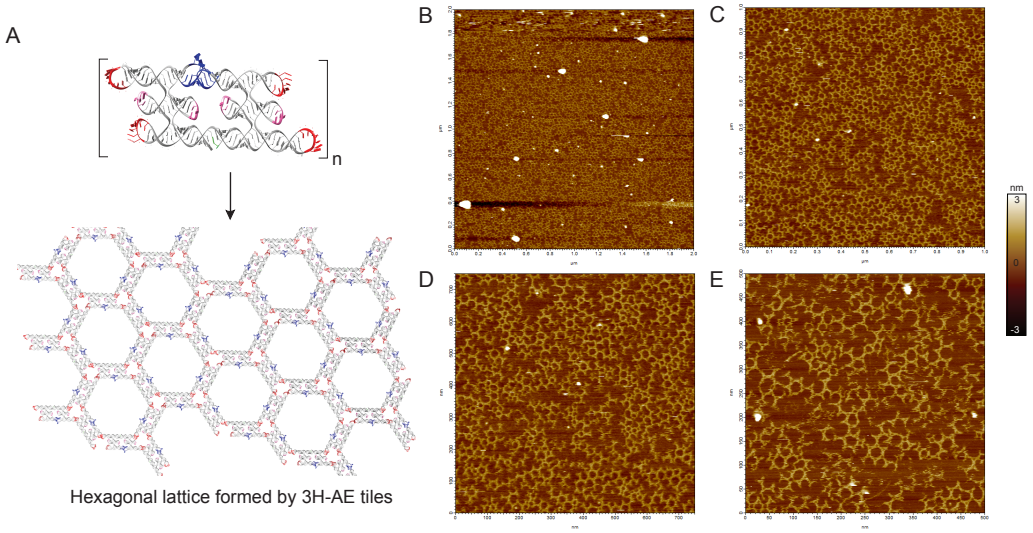


Figure 33: AFM images of hexagonal lattice formed by 3H-AE RNA tiles. (A) Schematic representation of hexagonal lattice formation. (B-E) AFM images showing hexagonal lattices, corresponding to different areas of imaging – 2 µm × 2 µm, 1 µm × 1 µm, 700 nm × 700 nm and 500 nm × 500 nm.

5.3.2 Modification of 3H-AE RNA tile with one aptamer

Modified versions of the 3H-AE RNA with one or more fluorescent RNA aptamers were also designed. The first version included the replacement of one of the two interior tetraloops by additional RNA motifs – subdomain IIa of IRES in HCV virus (PDB ID: 2PN4) and malachite green aptamer (MGA). The former was an L-shaped motif that provided a 90° turn and in turn facilitated attachment of an RNA sequence protruding perpendicular to the tile extending into 3D. The RNA sequence could encode a fluorescent aptamer, a protein binding aptamer or any other RNA aptamer.

In the first modification of the design, MGA was connected to domain IIa (3H-AE-MGapt V1) (figure 34 (A)) which resulted in an increase in the number of nucleotides in the tile from 248 to 297 nt, as verified by denaturing PAGE (figure 34 (B)). The presence of a fluorescent aptamer on RNA tile further enabled real-time monitoring of cotranscriptional folding of the tile (figure 34 (C)). The fluorescence was normalized to the maximum MG fluorescence intensity obtained. Additionally, AFM was also employed to check the formation of lattices as the 120° KLs remained intact despite the modifications in the tile (figure 34 (D)).

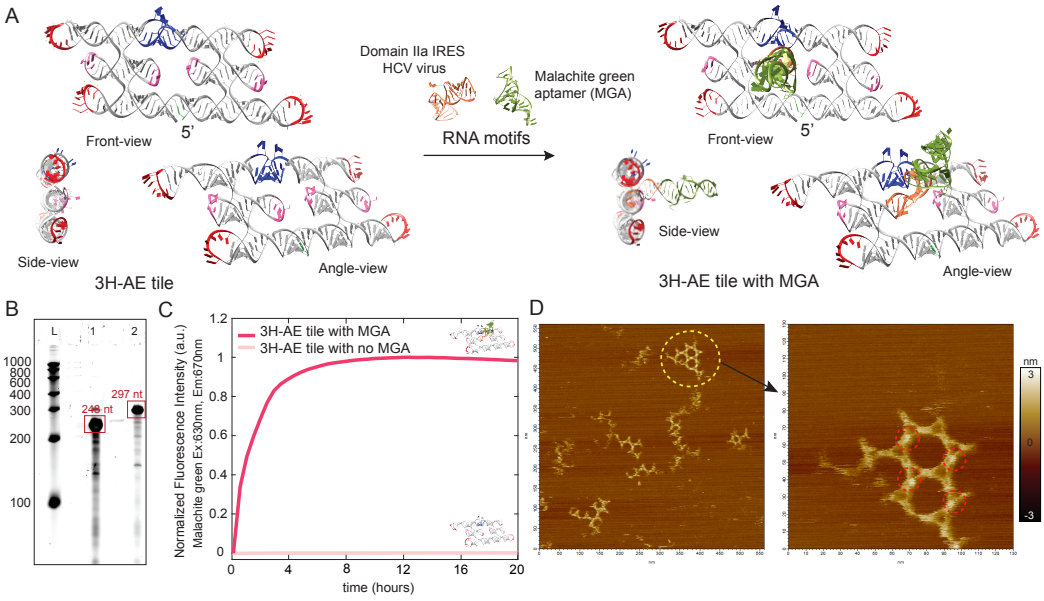


Figure 34: Modification of 3H-AE RNA tile with MGA. (A) Schematic representation of addition of MGA to 3H-AE RNA tile. (B) Denaturing PAGE gel showing the length of unmodified and modified RNA tiles. L: LowRange Riboruler, 1: 3H-AE RNA tile (248 nt), 2: 3H-AE-MGapt RNA tile (297 nt). (C) Real-time fluorescence. (D) AFM images showing hexagonal lattice formation. Yellow dashed encircled region is further zoomed in and the red dashed circles shown the position of MGA. Area of imaging - 550 nm × 550 nm, 130 nm × 130 nm respectively.

In order to exploit other motifs for positioning functional RNAs perpendicular to the tile arrangement, a variation in 3H-AE-MGapt V1 was made by replacing the connecting motif subdomain IIa (PDB ID: 2PN4) by domain IIa of IRES in HCV virus (PDB ID: 1P5M). Domain IIa includes a few additional unpaired bases in addition to the L-shaped subdomain IIa. As a result, the length of the tile was increased by a few bases from 297 nt to 303 nt (3H-AE-MGapt V2) (figure 35 (A) and (B)). There was almost no visible shift observed in migration of RNA. This could be attributed to the small nucleotide difference (6 nt) between 3H-AE-MGapt V1 and V2 that could not be resolved by the percentage of gel used. It was observed that the exchange of connection motifs did not enhance or compromise the correct folding and lattice forming abilities of the tiles as verified by increasing fluorescence of MGA over time and AFM (figure 35 (C) and (D) respectively). As mentioned above, the fluorescence was normalized with respect to maximum MG fluorescence. The only difference betwen V1 and V2 was a slight shift in the angle of MGA with the tile.

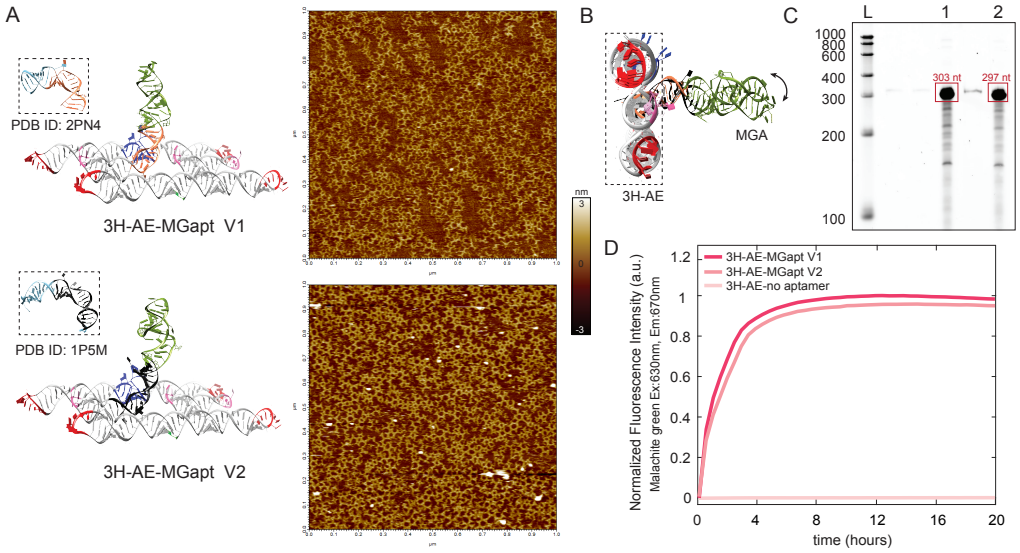


Figure 35: Comparison of 3H-AE-MGapt V1 and V2. (A) Schematic representation and corresponding AFM images of 3H-AE-MGapt V1 and V2. Area of imaging - 1 µm × 1 µm. MGA is connected to the tiles via subdomain IIa (PDB ID: 2PN4, coral) and domain IIa (PDB ID: 1P5M, black) of IRES in HCV virus. Inset shows the extracted motifs (coral (2PN4) and black (1P5M)) from the PDB file (sky blue). (B) Side-view of overlay of 3H-AE-MGapt V1 and V2. It shows a slight shift in the angle between MGA and RNA tile due to presence of different connection motifs (shown by arrow). (C) Denaturing gel showing the lengths of 3H-AE-MGapt V1 and V2. L: LowRange RiboRuler, 1: 3H-AE-MGapt V2 and 3: 3H-AE-MGapt V1. (D) Real-time fluorescence measurements demonstrating cotranscriptional folding of 3H-AE tiles with MGA.

5.3.3 Modification of 3H-AE RNA tile with two aptamers

Further, the complexity of the tiles was increased by addition of another aptamer on 3H-AE-MGapt V1 (297 nt) using subdomain IIa again. The second internal tetraloop was replaced by the spinach aptamer. It was connected via another subdomain IIa and extended perpendicular from the tile in the direction opposite to MGA (3H-AE-MG-spinach-apt) as seen in figure 36 (A). The length of the RNA tile increased to 388 nt which was checked by a denaturing PAGE gel (figure 36 (C)). Both MGA and spinach aptamers were added to the tile structure in such a way that each of them is present in one half of the tile, thereby acting as signals to ensure proper folding of 3H-AE-MG-spinach-apt. Hexagonal structures attempting to form lattices were observed in AFM as seen in figure 36 (D).

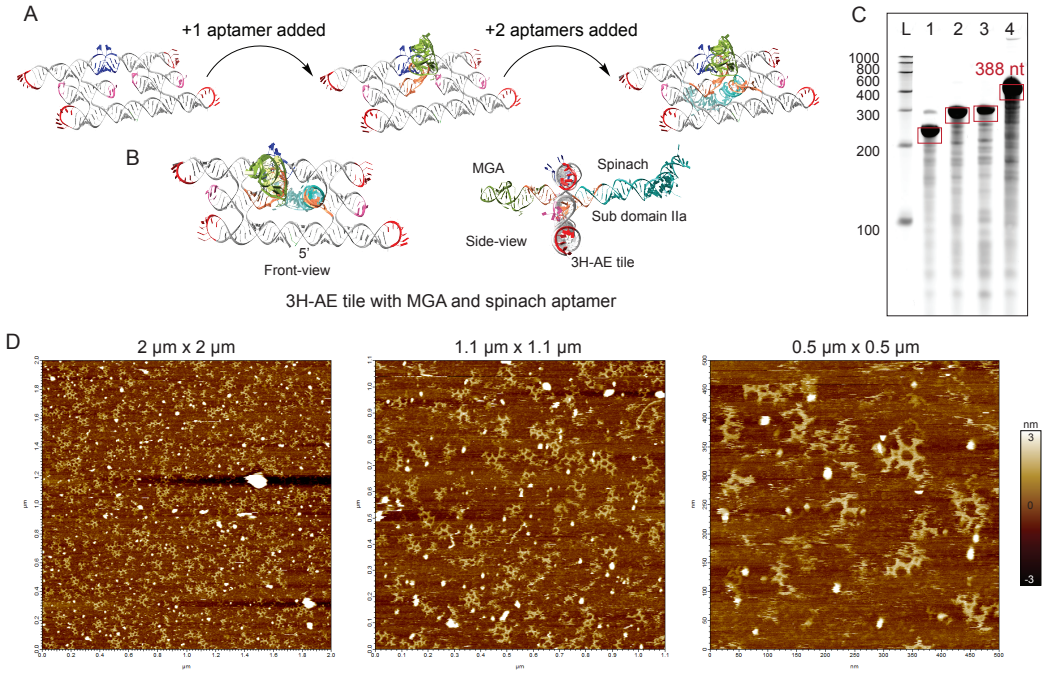


Figure 36: Modification of 3H-AE-MG-apt V1 RNA tile with spinach aptamer. (A) Schematic representation of addition of MGA and spinach aptamer to 3H-AE RNA tile. (B) Front and side-view of 3H-AE-MG-spinach-apt. (C) Denaturing PAGE gel demonstrating the length of unmodified and modified RNA tiles. L- LowRange Riboruler, 1: 3H-AE RNA tile (248 nt), 2: 3H-AE-MGapt V1 RNA tile (297 nt), 3: 3H-AE-MGapt V2 RNA tile (303 nt), 4: 3H-AE-MG-spinach-apt RNA tile (388 nt). (D) AFM images corresponding to different areas of imaging.

However, no well defined lattices were observed. This could be attributed to two factors – low concentration of RNA tiles making it difficult to form lattices. Secondly, it could be speculated that RNA tiles did not lie flat onto the mica surface due to the presence of MG and spinach aptamers coming out from either directions and not because of disruption of 120° KLs.

This was demonstrated by real-time monitoring of fluorescence of DFHBI and MG (figure 37). MG only, DFHBI only and both MG and DFHBI were added to all four versions of 3H-AE RNA tiles and the fluorescence was recorded in both fluorescence channels (Ex = 475-495 nm, Em = 515-545 nm for DFHBI and Ex = 615-645 nm, Em = 669-699 nm for MG). The fluorescence intensities were normalized with respect to maximum fluorescence obtained when DFHBI and MG were present separately regardless of fluorescence channel. In the presence of both DFHBI and MG, the fluorescence was normalized with respect to maximum fluorescence obtained in

each channel. MG fluorescence was observed in the DFHBI fluorescence channel for structures with MGA (i.e. 3H-AE-MGapt V1, V2 and 3H-AE-MG-spinach-apt) in presence of only MG or MG in combination with DFHBI. This could be due to the S2 fluorescence of MG at a lower wavelength of ≈ 470-530 nm which gets detected in the DFHBI fluorescence channel [167]. The combination of MGA and spinach aptamer were just used as detection signals to monitor structure formation. So, the fluorescence of MG in DFHBI channel was ignored. For specialised applications, the aptamers may be replaced by other RNA sequences.

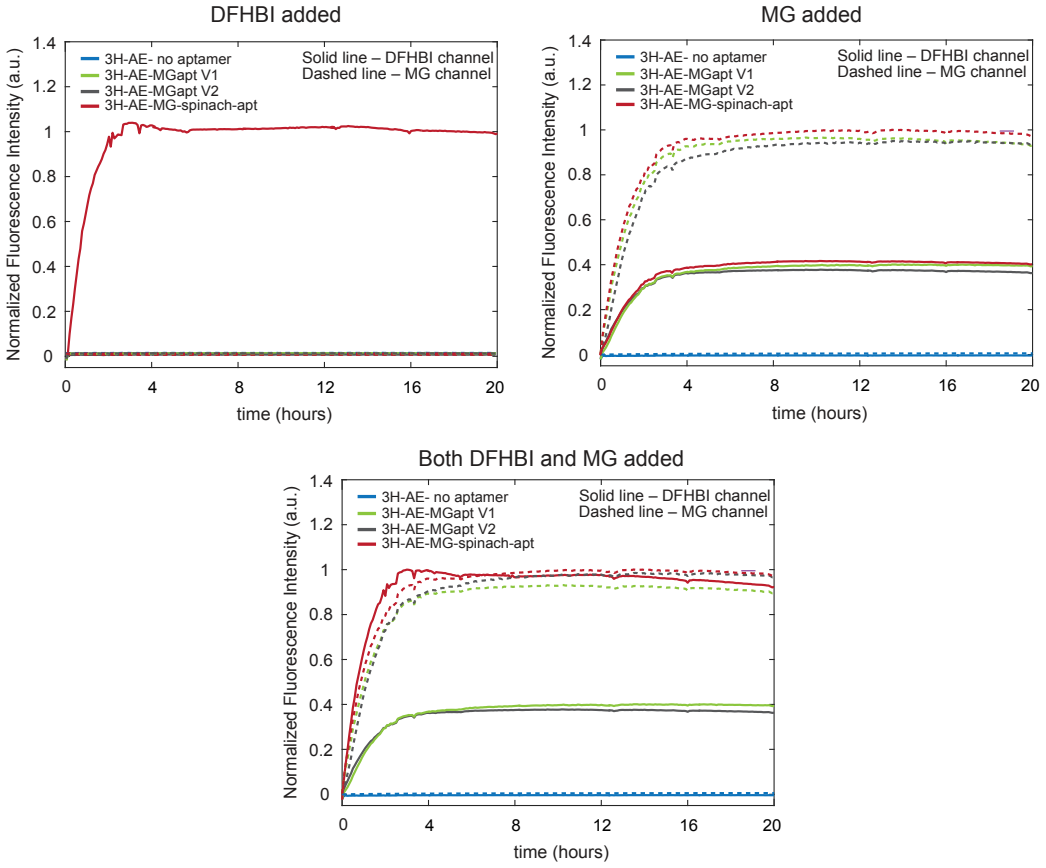


Figure 37: Real-time fluorescence of DFHBI and MG during transcription of different RNA constructs.

Other techniques such as TEM and cryo-EM may be used for better imaging of such 3D structures. 3H-AE RNA tile design and its modifications offers several advantages – domain IIa and subdomain IIa motifs is a relatively small in comparison to the bulkier motifs such as 3WJ of pRNA motif and allows addition of

RNAs perpendicular to the tile structure. Besides, it permits the addition of RNAs perpendicular to the tile and doesnot interfere with other features of the tile such as the ability of 120° KLs to form lattices. It offers greater insights into making more complex 3D RNA structures and lattices.

5.4 Conclusion and outlook

The results demonstrated here show the initial efforts towards design and synthesis of complex 3D RNA nanostructures. Three helix RNA tiles are folded from a single strand and consist of a variety of naturally occurring RNA motifs that assist in reaching the desired predefined shape. Of prime focus are the domain IIa and subdomain IIa present in IRES of HCV virus. Both these motifs contain a 90° bend which allows positioning of other RNAs such as fluorogenic RNA aptamers perpendicular to the tile that enables its extension in the third dimension. The structures 3H-AE-MGapt V1, V2 and 3H-AE-MG-spinach-apt are the first examples of 3D RNA structures folded from a single strand of RNA and capable of lattice formation.

As the lattice forming ability of RNA tiles is retained, the local concentration of the almost perpendicularly affixed RNA structures can be increased, that can potentially improve the overall effect. This addititve effect may be considered analogous to the property of avidity observed in some biomolecules such as antibodies etc. This property may be exploited for increased local interactions of RNA nanostructures for communication between compartments (figure 38). The structures can be further modifed in terms of increase in length of tiles by n number of half turns to enable extension of RNA from the same side or addition of other RNA molecules interacting with a variety of molecules and performing different functions such as lipid binding RNA aptamers, ribozymes etc respectively. The extension in the third dimension can also facilitate the addition of KL that can be used for construction of 3D arrangement of RNA hexagonal lattices over one another.

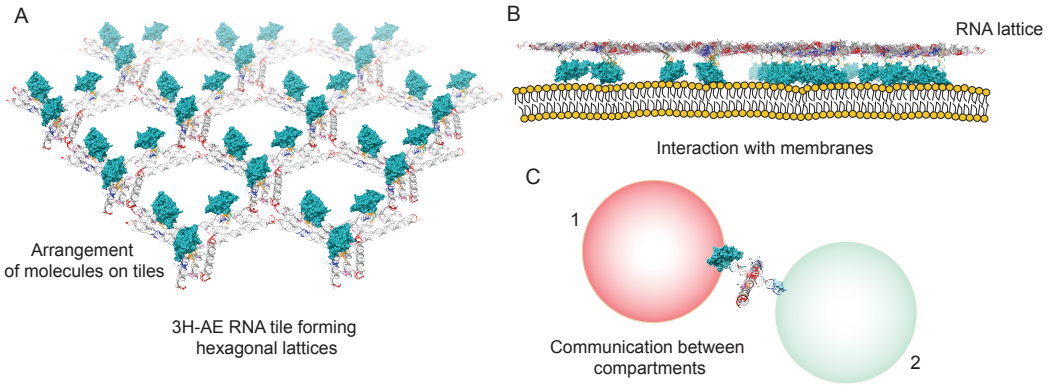


Figure 38: Potential applications of synthetic RNA nanostructures. (A) Arrangement of molecules at definite positions. (B) Interaction of RNA lattices with lipid membranes. (C) Communication between compartments.

6 RNA scaffold production by rolling circle transcription (RCT)

6.1 Introduction

Spatial organization of biochemical reactions is commonly observed in biological systems and it plays a key role in complex metabolic control. In nature, this organization is achieved by protein domain interactions, compartmentalization, organization of enzymes on different scaffolds (such as membranes in organelles and cells, proteins and nucleic acid scaffolds etc.), enzyme clustering (as in case of metabolons) or even separation of reactions in different cells [168]. This organization increases the local concentration of important molecules by bringing them in close proximity, which helps direct substrate flow between interacting enzymes, minimizing cross-talk and increasing the yields of metabolic reactions [169].

Compartmentalization and scaffolding of enzymes are effective strategies for the control of metabolic pathways in both natural and synthetic systems. The co-localization of enzymes in artificial enzyme aggregates or on synthetic scaffolds can lead to enzyme activities exceeding those of free enzyme systems [168, 170]. Novel scaffolds made of proteins, nucleic acids, and micro-compartments with increasingly complex architecture have recently been explored and applied [169]. Synthetic nucleic acid scaffolds serve as a versatile, rationally programmable alternative to protein-based scaffolding strategies for co-localization of multiple molecules. The structural properties and predictable base-pairing of nucleic acids allow a greater degree of control over the spatial organization and orientation of molecules. Both DNA and RNA based scaffolds have been used for different metabolic pathways with each having its own advantages and disadvantages [171].

RNA offers a variety of attractive features in comparison to DNA. The remarkable folding capability, modularity, regulatory and catalytic functions of cellular RNA have provided insights and inspiration for the synthesis of RNA based scaffolds and nanostructures [3, 100, 101]. Naturally occurring RNA motifs have been employed for design and synthesis of novel artificial RNA nanostructures (chapter 5). RNA with different enzymes scaffolded on them have been produced both *in vitro* and *in vivo* [108, 109]. It has been theoretically observed that diverse arrangements of organization are favourable for different enzymatic reactions. For fast enzymes, it

was found that a clustered arrangement was more favorable in contrast to a more distributed organization which was more advantageous in cases where the diffusion of intermediate products was faster than the reaction of the second enzyme [172]. Thus, a promising alternative to the construction of RNA scaffolds consisting only of single copies of the representatives of an enzyme cascade is the creation of multi-enzyme complexes on RNA that contain multiple copies of cooperating enzymes.

Here, we demonstrate the use of an alternate strategy known as rolling circle transcription (RCT) for the production of engineered RNA scaffolds for binding and spatially organizing multiple enzymes. RCT is primarily an *in vitro* technique for RNA replication and amplification. It is inspired by the naturally occurring rolling circle replication (RCR) phenomenon of nucleic acid replication around a circular template such as plasmids, genomes of bacteriophages, and circular RNA genome of viroids. RCT transcribes long repeats of RNA from circular DNA templates. It has been used for biosensing applications and to form siRNA concatamers used in RNA interference [107]. It offers the advantage of creating long repeats of RNA that can potentially be exploited to create RNA scaffolds to study clustering of enzymes. Along similar lines, rolling circle amplification (RCA) for isothermal amplification of DNA has been developed. RCA has been widely used for amplification of signals in biosensing devices and also as an alternate strategy for large-scale production of DNA [84, 173].

In this chapter, an RNA scaffold encoding a combination of protein binding aptamer and fluorescent aptamer i.e. PP7 aptamer (binding to PCP-YPet fusion protein (FP)) and MGA (PP7-MGA) respectively has been synthesized via RCT. The specific binding of the RNA scaffold (produced by RCT) decorated with aptamers to its target molecules is also investigated.

6.2 Experimental details and methods

Design and generation of circularized DNA

DNA template encoding for PP7-MGA aptamer combination was circularized by splint ligation. The sequence of interest (PP7-MGA, 96 nt) was flanked by a split promoter sequence (pT7 (22 nt), T7RNAP) on both sides. DNA was circularized by hybridization of a complementary pT7 sequence consisting of ssDNA encoding PP7-

MGA and ds pT7 regions. Schematic representation of the splint ligation strategy is given in figure 39 (A). A spacer sequence was added between the two aptamers to allow flexibility without interfering with folding of individual RNA. Additionally, a linear double-stranded counter part and a control sequences (circularized DNA) for each were also designed. The control sequence was identical in length and GC content to the circularized PP7-MGA DNA template. Both control and spacer sequences were designed by Nupack. The sequences for circularization (PP7-MGA and control) were ordered as 5′-Phosphorylated (5′-Phos) strands from IDT. All other sequences were ordered from Eurofins MWG. The DNA sequences are given in figure 78 (appendix). Additionally, Atto550 labeled DNA template corresponding to dBrocoli aptamer (dBroc) (110 nt) was generated by PCR using labeled primers and provided by L. Oesinghaus (Simmel lab, TUM).

Splint ligation: 1:1 ratio of pT7 (10 µM) (ssDNA) and 5′-Phos ssDNA (DNA Ultramer from IDT) (10 µM) strand encoding aptamers flanked by split promoter sequence of T7RNAP (pT7) on both sides was hybridized in T4 ligase buffer (NEB) using a thermal annealing ramp that cooled from 65°C to 20°C. Further, 1.5 µl T4 DNA ligase was added (400 U/ 1 µl) to the reaction (60 µl) and incubated for 2 hours at RT.

RCT and purification

The products of the ligation reaction were subsequently used for the transcription reaction. Different RNA were transcribed from their corresponding DNA templates for different experiments according to protocol mentioned in chapter 5. A limiting concentration of T7RNAP was used in comparison to the DNA template concentration (T7RNAP: 200nM, DNA: 250 nM). For RNA with aptamer modifications, non-fluorescent chemical compounds malachite green (MG) (10 µM) and DFHBI-1T ((Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-1-(2,2,2-trifluoroethyl)-1H-imidazol-5(4 H)-one, 10 µM) respectively were also added that emitted fluorescence in presence of the respective MG and dBrocoli aptamers. Additionally, 125 mM potassium chloride (KCl) was added during transcription of dBroc. PP7-MGA (linear and circular) fluorescence was monitored with time by measuring the fluorescence of MG as demonstrated in the previous chapter. All RNA transcribed were analysed by 8% denaturing PAGE gel. For experiments related to effect of pyrophosphate on transcription, 1 µl Inorganic Pyrophosphatase (PPiase) 0.1 U /1 µl (NEB)

was added to 30 µl transcription reactions. For transcription of dBroc, Atto647N labeled T7RNAP provided by S. Sagredo (Simmel lab, TUM) ($\sim$ 200 nM) was used.

Confocal microscopy of different RNA

Purified RNAs (5 times diluted) were next subjected to a heat denaturation/renaturation protocol (95°C for 3 minutes and then snap-cooled to 4°C for 5 minutes) and incubated with PCP-YPet 1 µl (16 µM) and 5 µl MG (60 µM) in a reaction volume of 30 µl for 30-45 minutes. 1 µl RNase Inhibitor (NEB) ($\sim$ 40 U/1 µl) was also added to prevent degradation. Post incubation, different RNA samples were seeded in ibiTreat IBIDI^{TM} chambers for imaging. RNA subjected to different conditions (EDTA, PCP-YPet and MG, labeled Atto550-DNA, Atto647N-T7RNAP and fluorescent product dBroc) were imaged using a Leica SP II inverted fluorescence confocal microscope (Leica) and analyzed using Fiji. Fluorescence filters used were Ex: 488 nm / Em:510-540 nm for dBrocoli and YPet fluorescence, Ex:561 nm / Em: 565-600 nm for Atto550 labeled DNA template and Ex: 633 nm / Em: 660-700 nm for Atto647N labeled T7RNAP and MG. PCP-YPet was provided by M. Schwarz-Schilling (Simmel lab, TUM).

6.3 Results and discussion

6.3.1 RCT of PP7-MGA RNA aptamers

Long continuous repeats of PP7-MGA RNA aptamer combination were produced by RCT. Figure 39 (B) shows a comparison of RNA produced from transcription of DNA templates of different morphologies – linear and circular. A single band corresponding to the correct length was observed for RNA synthesized from a linear template, whereas multiple bands of different lengths were obtained from transcription of circular template. This confirmed the cycling of T7RNAP around the circularized DNA resulting in long continuous repeats of the encoded aptamers. The length of different bands provided an approximate number of times T7RNAP cycled around the DNA and number of copies of aptamers present on the RNA. The intensity of the bands on the gel showed the relative concentration of different lengths of RNA transcribed. The longest band corresponded to $\approx$ 700nt which is equivalent to roughly 7 copies of RNA aptmer pair ($\approx$ 7 turns of T7RNAP around

the DNA, 1 turn produces ≈ 96nt). PP7-MGA was designed such that MGA was placed after the PP7 aptamer and fluorescence monitoring of MGA would ensure proper folding of both aptamers. The appropriate folding of the aptamers via RCT was further confirmed by real-time fluorescence measurements MG. A similar fluorescence signal was obtained for both linear as well as rolling circle transcription (figure 39 (C)). As a control, a sequence of same length and GC content was also circularized and transcribed. It was observed that the control sequence could be transcribed via RCT and did not show any fluorescence as expected.

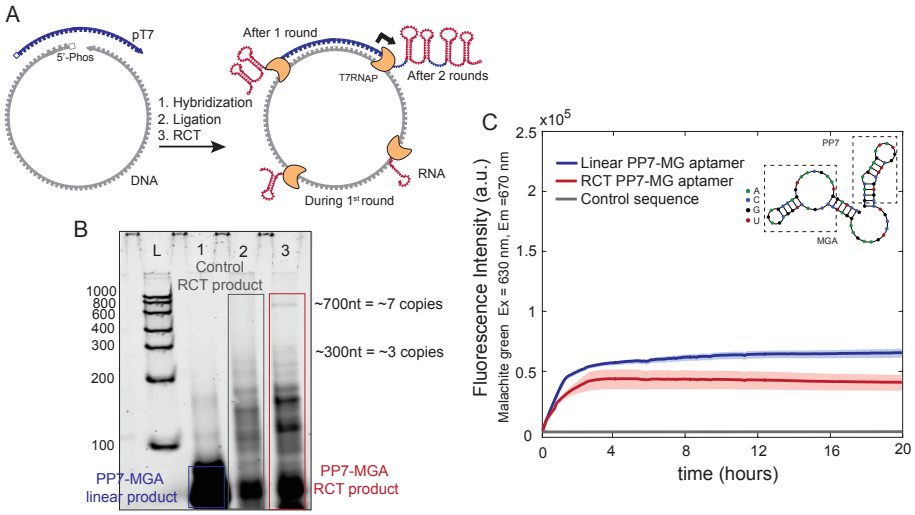


Figure 39: RCT of PP7-MGA RNA aptamers. (A) Schematic representation of RCT (dimensions of DNA template and T7RNAP: not to scale). (B) Denaturing gel showing purified RNA products from linear and circularized templates of PP7-MGA. L: LowRange RiboRuler, 1: linear product, 2: Control sequence RCT product, 3: PP7-MGA RCT product. (C) Real-time fluorescence measurements of MGA indicative of transcription of PP7-MGA using linear and circularized DNA (n = 3). Also, fluorescence from a circularized control sequence was measured. Inset: MFE structure of PP7-MGA from Nupack.

6.3.2 Confocal microscopy of transcription products

One of the main incentives of synthesizing RNA by RCT was to facilitate binding of multiple molecules in close proximity. This was verified by confocal microscopy of purified PP7-MGA RNA samples (linear, RCT and control) incubated with PCP-YPet FP and MG. Before incubation, RNA was also subjected to a temperature ramp to ensure proper folding. Visible co-localized areas of YPet and MG fluores-

cence was observed in PP7-MGA RCT product (figure 40, bottom row).

However, no such regions were obtained either in case of control RCT product (as expected, figure 40, middle row) or linear transcripts of PP7-MGA (figure 40, top row). This could be because of binding of one PCP-YPet and MG each in linear transcripts whereas in case of RCT multiple PCP-YPet and MG are present together as a chain on a single RNA.

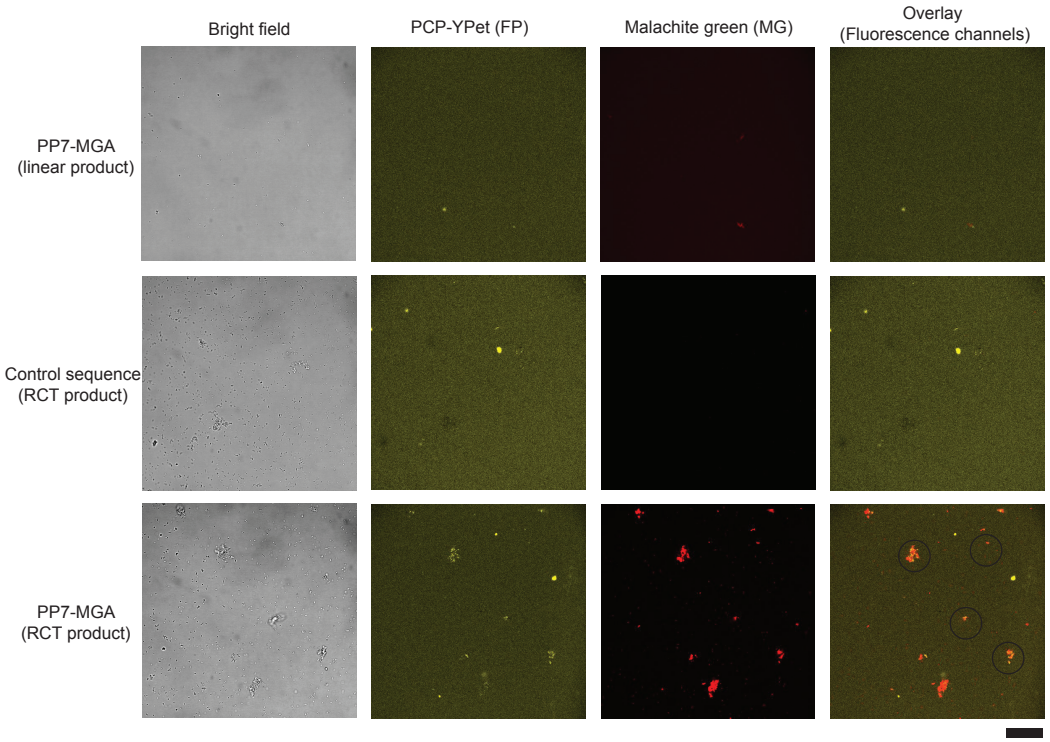


Figure 40: Confocal micrographs showing co-localization of YPet and MG on different RNA scaffolds. Encircled regions in PP7-MGA (RCT product) overlay image show co-localization. This is in constrast to linear product and control sequences. Scale bar: 40 µm.

The concentration of purified RCT RNA products could potentially be increased to observe greater co-localization. Also, methods such as gel extraction could be used to get rid of abortive transcripts that might interfere. Additionally, the concentration PCP-YPet could be kept in check in order to reduce background, as it will continue to fluoresce in presence or absence of RNA unlike MG.

The use of purified RNA samples is extremely important to rule out the possiblity of co-localization of due to transcriptional aggregation. Fluorescent aggregates of different sizes were observed when unpurified PP7-MGA RCT products incubated with MG were imaged (figure 41). Similar fluorescent aggregates were also observed when MG was added during the transcription reaction.

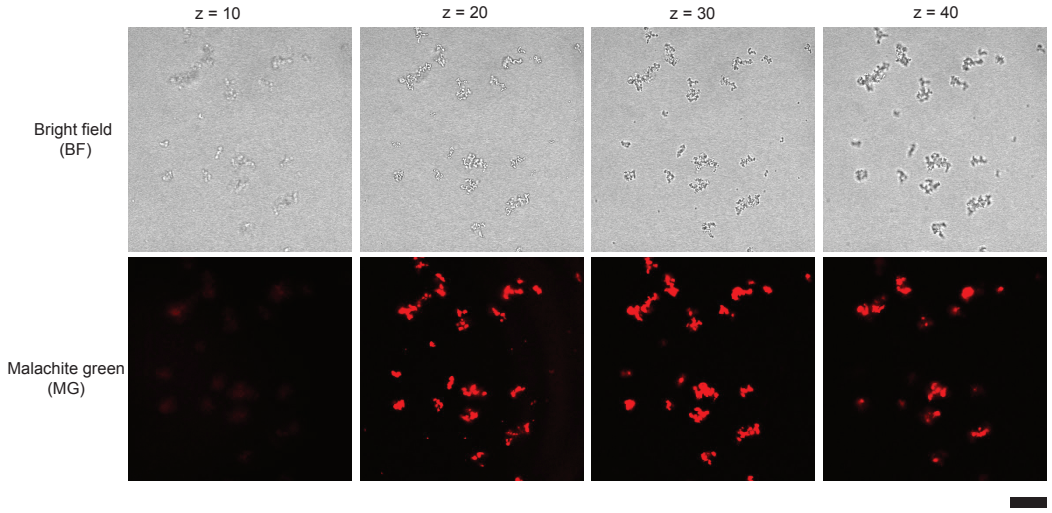


Figure 41: Transcriptional aggregates of PP7-MGA RCT RNA. Z stack of confocal images of aggregates ranging in size from few hundreds of nm to µm. Each z slice is 0.5 µm. Scale bar: 40 µm

The aggregates produced were complexes formed from Mg^{2+} and byproduct pyrophosphate (PPi) produced after the addition of an rNTP to a growing RNA strand during transcription [107]. This feature is universal to all transcription reactions irrespective of the nature of DNA template being employed (linear or circular). The aggregates not only include Mg^{2+}-PPi complexes and RNA product (dBroc aptamer) but also the DNA template and T7RNAP. This was verified using labeled linear DNA template (Atto550), T7RNAP (Atto647N) and fluorescent RNA during transcription (figure 42). This entrapment and decreased availability of free Mg^{2+} (speculated to affect the activity of T7RNAP [174]) could potentially affect the overall product synthesis.

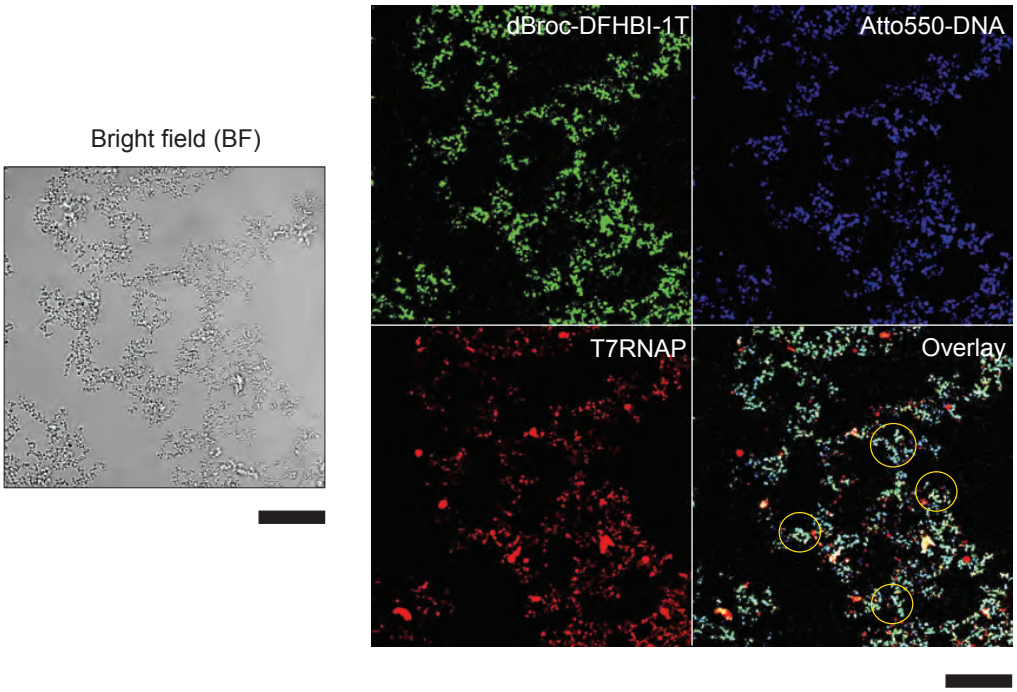


Figure 42: Co-localization of transcription reactants and products in RNA aggregates. BF and confocal images showing the entrapment of labeled reactants (Atto550-DNA (blue) and Atto647N-T7RNAP(red)) and fluorescent products (dBroc RNA aptamer (green)) of transcription in aggregates. Overlay of confocal micrographs shows co-localization (yellow circles). Scale bar: 100 µm

A typical transcription reaction is a polymerization type reaction where T7RNAP adds ribonucleotides (rNTPs) to a growing RNA chain and can be depicted by the following reaction:

$$(RNA)_n + rNTPs \xrightarrow[Mg^{2+}]{T7RNAP} (RNA)_{n+1} + PPi \tag{9}$$

where $(RNA)_{n+1}$ is the growing RNA chain, PPi is the pyrophosphate byproduct which complexes with Mg^{2+} to form insoluble Mg^{2+}-PPi complexes.

In order to increase the yield of RNA, PPiase was added to transcription reactions that degrades PPi and deters Mg^{2+}-PPi complex formation by breaking down PPi to 2Pi (orthophosphate) and releasing energy. The standard Gibb's free energy change, ΔG° for this reaction is -4.6kcal/mol [10].

$$PPi \xrightarrow{PPiase} 2Pi \tag{10}$$

This results in an overall shift of the transcription reaction in the foward direction favouring the formation of products $(RNA)_{n+1}$ and PPi. This also increases the overall Gibb's free energy change of the reaction ΔG. It can be calculated by the following equation:

$$\Delta G = \Delta G^{\circ} + RTlnQ \tag{11}$$

where Q is the reaction quotient, $Q = \frac{[RNA_{n+1}][PPi]}{[rNTPs][RNA_n]}$. A stark increase in fluorescence of MG (linear DNA template, PP7-MGA) was observed when PPiase was added during transcription. This was in contrast to the what was observed for circular DNA template as seen in figure 43.

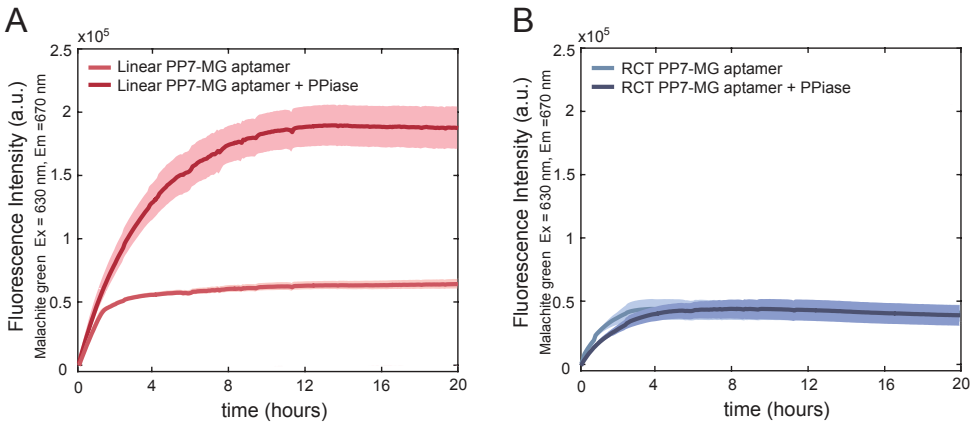


Figure 43: Effect of PPiase on transcription using linear and circularized DNA templates (PP7-MGA). Real-time fluorescence of MGA in the presence and absence of PPiase. (A) Linear template. (B) Circular template.

Since a limiting concentration of T7RNAP was used in the experiments with either morphologies of template DNA concentration (i.e [T7RNAP]<[DNA template]), it could be speculated that in case of the linear template, T7RNAP falls off from the template and hunts for the pT7 again. This "hunt" might get difficult with the increasing and growing Mg^{2+}-PPi complexes. This restricted movement might be overcome by the breaking of Mg^{2+}-PPi complexes by PPiase, thereby resulting in an increase in RNA production which is seen as increased fluorescence of MGA.

In case of circular DNA template, T7RNAP cycles around the same template so the accumulation of the complex does not influence it in the experimental settings tested. No significant change in the maximum fluorescence achieved was observed after PPiase addition.

Additionally, the effect of chelating agents like EDTA on the size and morphology of transcriptional aggregates produced from circularized PP7-MGA DNA were observed by DLS and confocal microscopy, respectively. The aggregates dissolved and their size reduced on addition of EDTA (figure 64, appendix).

6.4 Conclusion and outlook

Rolling circle transcription has been successfully used for generating long continuous repeats of PP7-MGA RNA and the assembly of YPet and MG was shown. Aggregates were observed in all transcription reactions irrespective of the nature of DNA template being employed.

From the results obtained, RCT could be forseen as a strategy for generation of RNA based multienzyme clusters. Recent studies suggest that co-clustering of enzymes into agglomerates is a more favorable strategy to enhance the overall reaction flux rather than creating individual structures. Clustering leads to an improved capture of reaction intermediates by increasing the probability that the product of one enzymatic reaction enters one of many neighboring active centers of the consecutive enzyme in the cascade [170, 172]. This could possibly increase the proximity and local concentration of the enzymes, thereby resulting in an improved transfer of reaction intermediates and enhanced the overall rate of reaction as seen in [170] (figure 44 (A)).

RNA structures of greater complexities such as RNA tiles with modifications as described in chapter 5 could be investigated via RCT. Such structures could be decorated with fusion proteins (peptide-enzyme) that interact via aptamer binding sites. Further improvements in design and experimental conditions (concentration of template and enzyme, duration) could be carried out to increase the yields. The transcriptional aggregates produced during RCT could be used as enzyme clusters, and their size could be tuned and optimized for different enzyme cascades. The clusters produced could also be encapsulated in larger microcompartments for optimal performance (figure 44 (B)).

In addition, the effect of PPiase on transcription of DNA templates of different morphologies was reported. The differential response of linear and circular DNA templates to PPiase during transcription is an interesting phenomenon and could be further explored.

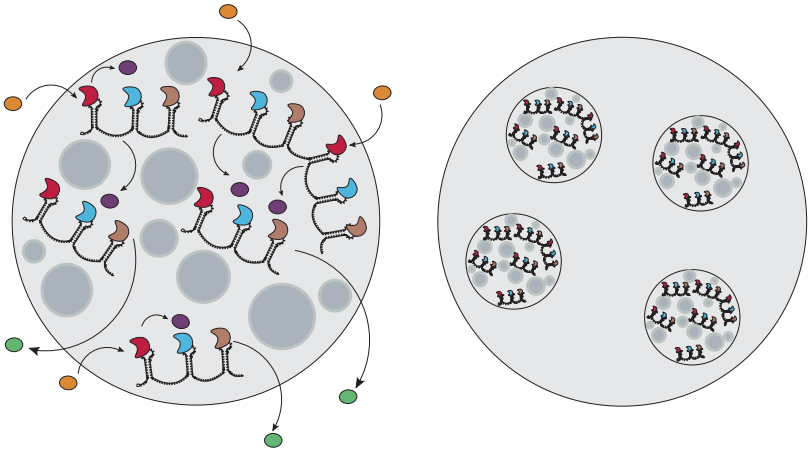


Figure 44: RNA based clustering of multiple enzymes. (A) RNA scaffold genereated by RCT used to organize multiple enzymes. Reactants and products capable of diffusion. Intermediates can have multiple encounters with subsequent enzymes of the cascade. (B) Encapsulation of multiple clusters.

7 *In vivo* synthesis of catalytic DNA molecules

7.1 Introduction

DNAzymes refer to a class of DNA molecules that are catalytically active and are synthesized via *in vitro* processes such as SELEX. They can be designed to sense a broad range of analytes and perform a series of functions including cleavage, ligation etc. under conditions desired for their intented applications. RNA-cleaving DNAzymes have attracted significant attention for both therapeutic and diagnostic applications due to their excellent programmability, stability, and activity. They have also been exploited in dynamic DNA nanotechnology and computing for construction of nucleic acid based circuits [95, 96]. In addition, DNAzymes have been used for synthesis of autonomous nanoscale motors that mimic biologically existing motors like myosin and kinesin [175]. So far, no natural DNAzymes have been discovered. Most DNA in cells are double-stranded while the catalytic core of DNAzymes is typically single-stranded. This is in contrast to the widespread occurrence of ribozymes (RNA exhibiting catalytic functions) and their roles in cellular systems [3].

10-23 DNAzyme is one of the most widely studied DNAzymes. It is an RNA cleaving DNAzyme and acts on phosphodiester bond between ribonucleoties showing the maximum efficiency for rA-rU. The 10-23 DNAzyme contains a 15-nucleotide catalytic core that is flanked by two substrate recognition domains. This DNAzyme cleaves complementary RNAs efficiently in a sequence specific manner between an unpaired purine and a paired pyrimidine. It has been widely used as an integral component of biosensors, microoarrays and hold great potential as therapeutics [176]. One example that employs the 10-23 DNAzyme has been in the construction of autonomous synthetic DNA motors which mimics the working of natural protein motors [175]. The DNA motor consisting of two strands (strand E/10-23 DNAzyme and strand F) with complementary regions has been designed in a way that its motion is dependent on the catalytic activity of the 10-23 DNAzyme. It continues to run as long as the fuel (RNA substrate) of the DNAzyme is available. The chemical energy released from the breaking of a covalent bond in the substrate, strand SH (by the DNAzyme) was converted into the mechanical motion of the motor.

The increasing demand of DNAzymes has also increased need for a larger scale of production and sequences with higher fidelity. This chapter focuses on large-scale biological production of 10-23 DNAzyme and its catalytic characterization.

Currently, chemically synthesized DNA is the main source of raw material for DNA nanotechnology. However, it is limited in scale and length of the sequences produced. Also, the errors in the DNA synthesis process, and the resulting faulty oligonucleotides, are believed to limit the performance of dynamic DNA devices by causing erroneous side reactions. Several methods have been developed for mass production of ssDNA of varied lengths that exploit either the bacterial replication machinery or enzymes such as phi29 polymerase with much lower error rates for synthesis ensuring the purity of the resulting DNA in comparison to oligonucleotides produced via solid-state synthesis [177]. Seelig et al. have shown that biologically generated DNA outperforms its synthetic PAGE purified counterparts in DNA based circuits [178]. A comparison of plasmid-derived and synthetic DNA gates has shown a higher turnover in chemical reactions for the biologically generated in contrast to the synthetic systems. It was speculated that higher purity of the plasmid-derived gates was most likely responsible for this behaviour.

A technique developed by Ducani et al. known as the 'monoclonal stoichiometric' (MOSIC) method produces DNA oligonucleotides with precisely controlled stoichiometric ratios using enzymes [84]. The DNA strands are produced either by *in vitro* rolling circle amplification (RCA) or by amplification of phagemid DNA in *E.coli*. DNA strands produced by this method have been used for different applications in DNA nanotechnology such as production of scaffold and staples of variable lengths for formation of DNA nanostructures and crystals. Recently, a new method has been developed that is an improvisation of the existing MOSIC method. ssDNA of arbitary length and sequences were produced by combining the self excising ability of DNAzyme I-R3 and ss phagemid DNA containing target strands produced by *E.coli* infected with helper bacteriophages [179]. This method further reduces the dependence on restriction enzymes (RE) for excision of target DNA strands. RE increase the cost of production and their functionality is temperature dependent.

In this chapter, *in vivo* synthesis of 10-23 DNAzyme using MOSIC method (or MOSIC-DNAzyme or strand E in motor) by amplification of phagemid DNA in *E.coli* is shown. It is used as a component for the construction of an autonomous DNA nanomotor capable of switching from one state to another in the presence of

a substrate and its enzyme activity is tested.

7.2 Experimental details and methods

Design of MOSIC-DNAzyme sequences

The sequences used in the construction of the DNA motor were taken from [175]. The pseudogene encoding the 10-23 DNAzyme sequence flanked with other sequences such as BseG1 or BtsC1 restriction enzyme sites was generated by MOSIC pseudogene generator (`http://www.hogberglab.net/mosic-generator/`) and ordered in pIDTBlue phagemid vector (containing two origin – pUC origin for replication as plasmid and f1 origin for replication as ssDNA). All strands and pIDTBlue phagemid were ordered from IDT. DNA sequences and the vector map are shown in figure 79 in the appendix.

ss phagemid protocol for production and harvest of DNAzyme

JM109 *E.coli* transformation and M13 helper phage infection. JM109 strain of *E.coli* was chemically transformed with pIDTBlue phagemid vector containing 10-23 DNAzyme sequence and plated on agar plates containing ampicillin. On following day, a single colony was picked and grown overnight in 5 ml luria broth (LB) medium containing 100 mg/ml ampicillin. Next, a 300 ml 2x YT medium (Sigma-Aldrich) was inoculated from the saturated overnight culture and 1.5 ml of $MgCl_2$ and 100 mg/ml ampicillin was added. The culture was grown at 37°C with shaking at 250 rpm. OD_{600} was monitored during cell growth. M13KO7 helper phage (NEB) with a multiplicity of infection of 20 (ratio of phage to cells) was added to the culture when the OD_{600} reached 0.45 corresponding to around 2.5×10^8 cells/ml. The volume of the helper phage added was dependent on the concentration of helper phage, multiplicity of infection and number of cells. After an additional shaking for 1-1.5 hours, kanamycin (50 mg/ml) was added to kill the bacteria. The culture was further shaken for about 3-4 hours.

Harvest of ssDNA. The cells were discarded by centrifuging twice at 5000 rcf for 20 minutes at 4°C. PEG precipitation (40 gm/litre PEG-8000 (Sigma-aldrich) and 30 gm/litre (Sigma-aldrich)) was carried out on ice for 1 hour with continuous stirring to precipitate phage particles from the supernatant. It was further

centrifuged at 15000 rcf for 30 minutes at 4°C. The supernatant was thrown, and pellet obtained was resuspended in 10 mM Tris pH 8.5. It was again centrifuged to get rid of any residual bacteria at 5000 rcf for 5 minutes at 4°C, and supernatant was transferred to fresh tubes. The phage proteins were removed by addition of 0.2 M NaOH and 1% SDS to the solution, mixed by gentle twirling and incubated at RT for 3 minutes. Further, 1.5 volumes of 3 M KOAc (pH 5.5, titrated with glacial acetic acid) was added, mixed by inversion and incubated in ice water bath for 10 minutes.

This mixture was centrifuged at 16500 rcf for 30 minutes at 4°C. The supernatant was decanted and 2 volumes of 100% ethanol was added and incubated in ice water bath for 1 hour. Another centrifugation step was performed at 16500 rcf for 30 minutes at 4°C. The obtained ssDNA (pellet), was washed with 70% ethanol and resuspended in 2 ml Tris (10 mM pH 8.5) and analyzed by agarose gel. An additional purification step similar to phenol chloroform extraction was performed using StrataClean Resin (Agilent) as per manufacturer's instructions. A representation of images that describe the sequence of steps involved in the producing ssDNA using phagemids is shown in figure 65 in appendix.

Agarose gel electrophoresis

ssDNA otained was verified using 1% agarose gel containing Ethidium Bromide (Thermofisher scientific) in 10 mM $MgCl_2$ and 0.5X TBE. The gel was run at 70 V for 1 hour 30 minutes. Gel images were taken on a Typhoon FLV 9000 laser scanner and analyzed with the software package Fiji [140]. 1 kb DNA Ladder (NEB) was used as a marker.

Restriction digestion

Strand E was extracted from the ssDNA phage using BseG1 (200 U/1 µl, Fermentas) in 1X Tango buffer by incubating at 55 °C for 1 hour followed by heat-inactivation at 80 °C for 20 minutes followed by analysis of digestion product by denaturing PAGE gel. BseG1 or its isoschizomer BtsC1 (NEB) can be used for digestion.

PAGE

Denaturing PAGE containing 12% Acrylamide:Bis (19:1), 20% formamide, 8 M urea and 1X TBE buffer was used to analyse the digested ssDNA and estimate their concentration. The gel was pre-run at 50°C for 30 minutes. The samples and Ultra Low Range DNA Ladder (Thermofisher scientific) were prepared in 1:1 ratio with 2X RNA loading dye (Thermofisher scientific) and incubated at 80°C for 10 minutes prior to loading onto the gel. The gels were run at 70 V for 1 hour 45 minutes. After electrophoresis, the gels were stained with SYBR^{TM} Gold Gel Stain (Thermofisher scientific) for 10 minutes and scanned on a Typhoon FLV 9000 laser scanner and analyzed with the software package Fiji [140]. Ultra Low Range DNA ladder (Thermofisher scientific) was used as a marker.

Native PAGE containing 20% Acrylamide:bis (19:1), 12.5 mM $MgCl_2$ and 1X TAE buffer was used to analyse the catalytic reaction of the 10-23 DNAzyme. The gels were run at 70 V for 1 hour 45 minutes. Ultra Low Range DNA ladder (Thermofisher scientific) was used as a marker.

Purification

BSeG1 digestion reaction was loaded and run on a 4% denaturing PAGE gel. The digested strand E was cut out from the gel by UV shadowing using a TLC plate. Next, it was eluted using gel dialyzing column (6-8 kDa, Millipore). The column was placed perpendicular to the direction of the field in an electrophoresis tank with 1X TBE buffer. A voltage of 140 V was applied for 1 hour such that the DNA comes out of gel into the solution and gets attached to the membrane. Further, the polarity was reversed and voltage was applied at 80V for 30 seconds that allowed the DNA to migrate in the buffer. The voltages and duration for electroelution is dependent on a number of factors including size of DNA, percentage of gel used etc. The column was kept in water overnight and concentrated using a concentrator (Eppendorf).

DNA nanomotor construction and catalytic activity

DNA nanomotor was constructed by combining complementary DNA strands – strand E (or 10-23 DNAzyme) and strand F (0.5 µM) in equimolar concentration in

1X TAE buffer containing 12.5 mM $MgCl_2$. For MOSIC-DNAzyme, digested but unpurified ss phagemid was mixed with strand F. The nanomotor was formed by slowly cooling the DNA using the following conditions: 95°C (3 minutes), 65 °C (30 minutes), 37°C (30 minutes) and 22°C (30 minutes). Both DNA motors along with substrate strand (strand S (control) or SH (DNA-RNA chimera, target) in different motor:substrate ratios (M:S/SH) i.e. 1:20, 1:60 were incubated in 1X TAE buffer containing 12.5 mM $MgCl_2$ at 22°C for 1 hour and analyzed by 20% native PAGE.

7.3 Results and discussion

7.3.1 Design of MOSIC-DNAzyme and ss phagemid protocol

The sequences used in the construction of the DNA motor were taken from [175] with minor modifications. All the sequences were synthesized commercially unless specified. The 10-23 DNAzyme sequence was designed by the MOSIC pseudogene generator (figure 45 (A)) and also synthesized by MOSIC method using ss phagemid protocol as described (figure 45 (C)). It allowed addition of BSeG1 or BtsC1 restriction enzyme sequences on either side of the sequence of interest i.e. 10-23 DNAzyme or strand E. In the phagemid protocol, pIDTBlue phagemid vector including the pseudogene was packaged as ssDNA in phage particles and extracted. This resulted in ds DNA hairpins produced by the BSeG1 restriction enzyme sequence flanking strand E. Schematic representation of the hairpin architecture is depicted in figure 45 (B). BseGI restriction site (outlined by vertical lines; locations of cuts are shown with black arrowheads) is flanked on one side by a GAA loop and a GCGC stem, and on the cutting side by two hybridized base pairs, where N is any nucleotide. These base pairs are formed by encoding two complementary bases to the last bases at the $3'$ end of 10-23 DNAzyme into the hairpin sequence. These two bases are encoded in the hairpin byproduct. Two additional TT's were added to the $5'$ and $3'$ end of the 10-23 DNAzyme sequence. The TT's at $3'$ end accounted for hybridization with N bases in the hairpin and at $5'$ end for proper folding and assembly of nanomotor with strand F (figure 46 (A)).

The nanomotor switches between open and closed states in the presence of a specific substrate. This specificity is attributed to the 10-23 DNAzyme component of the motor. Only one component of the nanomotor i.e. 10-23 DNAzyme was biologically

produced and the rest were commericallly synthesized. The substrate strands (S and SH) used in both motors were also commercially synthesized.

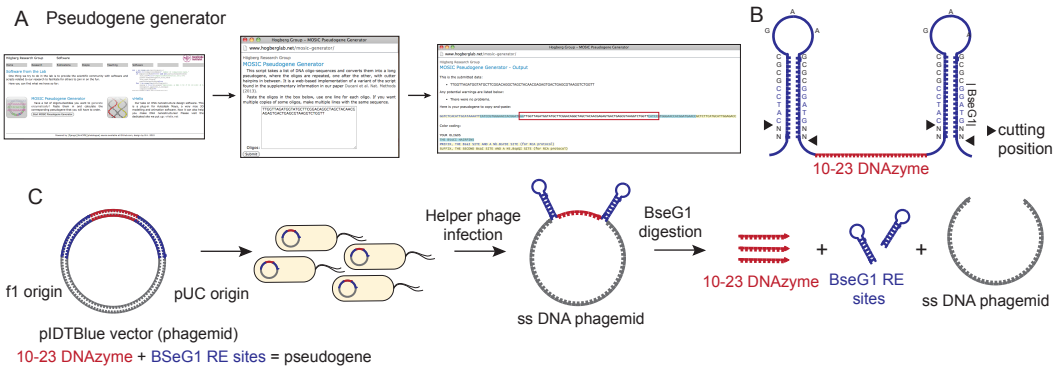


Figure 45: Design and synthesis of 10-23 DNAzyme using ss phagemid strategy. (A) Snapshots of pseudogene generator. (B) Schematic of BseG1 hairpin, (C) Outline of ss phagemid protocol.

7.3.2 Phagemid digestion

The intergrity and purity of ss phagemid harvested before and after application of StrataClean Resin was verified by agarose gel (figure 46 (B)). ss phagemid was further digested with BseG1 restriction enzyme to extract 10-23 DNA enzyme that was flanked by BseG1 recognition sequences. Figure 46 (C) shows digestion of ss phagemid using BseG1 at different concentrations. The digested phagemid ssDNA showed product bands of the same size as commercially procured strand E i.e., IDT-DNAzmye, thereby confirming the biological production of monodisperse strand E. Undigested products did not produce any bands of the the desired size.

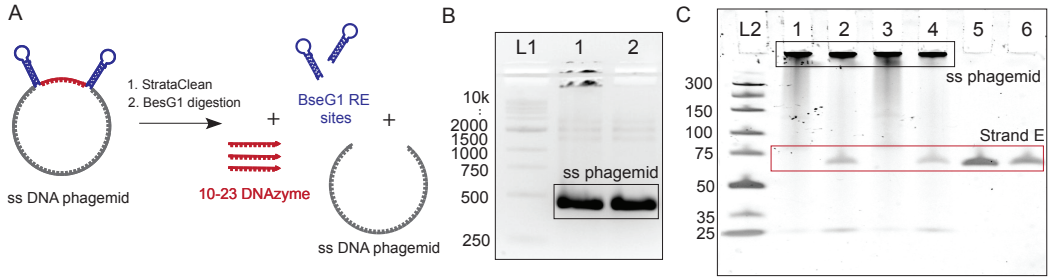


Figure 46: Verification and digestion of ss phagemid containing 10-23 DNAzyme. (A) Schematic illustration of ss phagemid digestion to generate 10-23 DNAzyme. (B) AGE, L1: 1 kb DNA Ladder, 1: ss phagemid with 10-23 DNAzyme without StrataClean, 2: ss phagemid with 10-23 DNAzyme with StrataClean. (C) Denaturing PAGE, L2: Ultra Low Range DNA Ladder, 1 and 2: ss phagemid without StrataClean undigested and digested respectively, 3 and 4: ss phagemid with StrataClean, undigested and digested respectively, 5 and 6: strand E from commercial vendor in the absence and presence of BseG1.

7.3.3 Catalytic activity of DNAzyme in DNA nanomotors

For inital experiments with MOSIC-motor, digested unpurified 10-23 DNAzyme was mixed with strand F to check the construction the of MOSIC-motor respectively. Bands corresponding to the correct length of motor were observed when compared to motor constructed with commercially synthesized DNAzyme (figure 47 (B), wells 10-13). This confirms the synthesis of 10-23 DNAzyme or strand E by MOSIC method using phagemid protocol and its ability to hybridize to the commercially synthesized strand F.

Further, the catalytic activity of the MOSIC-motor was verified by the incubation of substrate strands S (control) and SH (target) with the motor and the products were analysed by non denaturing PAGE gel as shown in the figure 47. As a control, the enzymatic activity of a motor constructed from commercially synthesized 10-23 DNAzyme and strand F was also performed (figure 47 (B), wells 5-8). The catalytic domain 10-23 DNAzyme was bound to DNA–RNA chimera substrate (strand SH) via WC base pairing, and cleaved only the phosphodiester bond between rA-rU ribonucleotides. This generated two short fragments ($\approx$ 7 nt) that were subsequently released. The process continues till all the substrate is cleaved. Since, unpurified DNAzyme was used remains from phagemid digestion including BseG1 hairpins and rest of the phagemid were also seen on the gel. The size of the hairpins was similar to substrate strands S and SH used. They were also present when only motor was

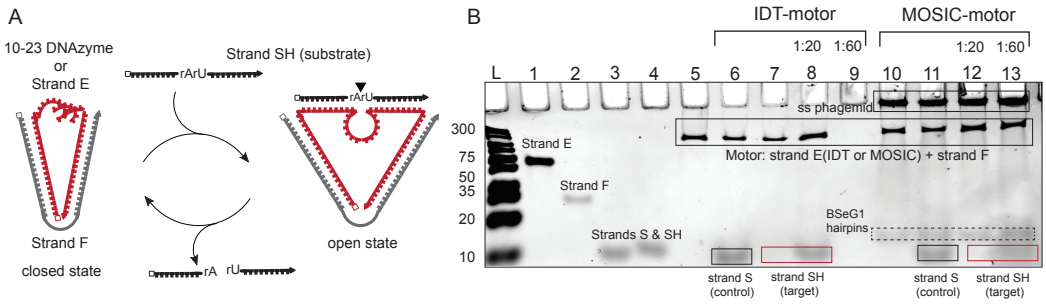


Figure 47: Catalytic reaction of IDT and MOSIC-motors with different substrate concentrations. (A) Schematic illustration of switching of the DNA nanomotor between closed and open states. (B) Nativ PAGE showing the operation of motors. L: Ultra Low Range DNA Ladder, 1-2: strand E and F motor components respectively, 3-4: Substrate strands S (control) and SH (target) respectively, 5: IDT-motor (strands E + F), 6: IDT-motor + strand S (1:60), 7: IDT-motor + strand SH (1:20), 8: IDT-motor + strand SH (1:60), 10: MOSIC-motor (strand E (produced by MOSIC) + F), 11: MOSIC-motor + strand S (1:60), 12: MOSIC-motor + strand SH (1:20), 13: MOSIC-motor + strand SH (1:60)

run (figure 47 (B), well 10). Both IDT and MOSIC motor cleaved substrate SH completely in 1 hour when M:SH ratio was 1:20 (no band observed in gel). At higher substrate concentrations, residual uncleaved substrate SH was observed for both motors. Different purification strategies for separation of digested strand E were exploited. Electroelution was carried out using a dialysing column (Millipore) for purification of strand E. The purity of 10-23 DNAzyme was verified by denaturing PAGE gel (figure 66, appendix).

7.4 Conclusion and outlook

ss phagemid strategy has been already applied for *in vivo* production of DNA strands for applications in DNA nanotechnology [84]. The results attained confirm the possibility of *in vivo* production of 10-23 DNAzyme which has been previously realised *in vitro* only. The length of 10-23 DNAzyme produced coincided with the sequence generated by solid-state synthesis. Also, the construction of a DNA motor with components of different origin (biological and commercially synthesized) shows the compatibility of interacting strands with each other. Both these observations were a clear evidence that the correct sequence was produced. However, more experiments with purified biologically produced strand E are required to validate its enzymatic potential and compare with its commercial counterpart. Chemical assays

such as methylation test for better characterization of biologically produced strands can be performed.

Biologically produced DNAzymes could provide as an alternative source of highly pure DNA and be used as components of DNA gates in DNA circuits. Also, they may be used as parts of biomimetic DNA based model systems of the biological protein-based counterparts to understand their functioning, and create more complicated DNA based devices. *In vivo* produced DNAzymes may also be biocompatible and easily taken up by cells and could be exploited for intracellular DNAzyme based biosensing applications.

8 Conclusion and outlook

In summary, different aspects of nucleic acids were explored in the context of nucleic acid based nanostructures. All projects discussed in this thesis seek inspiration from diverse features of nucleic acids. They vary from exploiting the Holliday junction as the basis for design of DNA origami nanostructures to assembling these structures in naturally occurring condensing agents, thereby mimicking *in vivo* DNA condensation. Naturally occurring ncRNA and RNA motifs were utilized for the design and construction of novel synthetic RNA nanostructures and scaffolds. Further, the cellular replication machinery was employed for the production of an *in vitro* designed and selected DNAzyme similar to DNA replication.

In the first part of the thesis, artificial DNA nanostructures were assembled using a naturally occurring trivalent condensing agent – spermidine. There was a striking similarity found between this folding of DNA structures and *in vitro* condensation of DNA. The folding was carried out at physiological conditions. The low salt requirement for folding and the ability to withstand electric fields of the Spd-nanostructures was further exploited for electrotransfection experiments using mammalian cells. Spd-folding of DNA nanostructures could be used as model systems for understanding DNA condensation *in vitro* and also provide insights into electrical manipulation of DNA nanostructures. Such Spd-folded nanostructures can overcome one of the major barriers of *in vivo* applications of DNA nanostructures.

In the second part, the structural and functional properties of biologically existing RNA were exploited. New synthetic 3D RNA nanoarchitectures were designed and folded using a single RNA strand. These structures consist of three helix tiles which were composed of a variety of naturally existing RNA motifs. The three helix tiles were further extended to the third dimension by addition of L-shaped 90° RNA motifs extracted from IRES of HCV virus. The tiles were designed in a way that L-shaped motifs pointed in opposite directions, thereby facilitating extension in different directions that were perpendicular to the tile. Different variations were constructed with either one or two extensions for positioning fluorescent RNA aptamers (MGA and spinach aptamer) perpendicular to the RNA tile. The fluorescence of the aptamer binding fluorophores also facilitated real-time progress of the folding reaction. The aptamer positions could be easily replaced with other ap-

tamers such as PP7 aptamer for PCP coat protein, BIV-TAR aptamer, ribozymes etc., customizing the RNA structures to perform specified functions. The structures created would provide a better understanding in designing more complex 3D RNA nanostructures with multiple functionalities.

The folding capability and predictable base-pairing of RNA have made it the popular material of choice for a synthetic scaffold for the control of metabolic pathways [108, 109]. Recently, it has been observed that co-clustering of enzymes into agglomerates is a more favorable strategy to enhance the overall reaction flux than creating individual structures [170, 172]. Rolling circle transcription could be employed as an alternative strategy for the synthesis of synthetic RNA scaffolds capable of organizing multiple copies of enzymes of metabolic pathways. As a proof of concept, long continuous repeats of PP7-MGA RNA were produced by RCT and the assembly of PCP-YPet fusion protein and MG fluorophore was shown.

In the last part, *in vivo* production of a single-stranded catalytic DNA molecule (DNAzyme) was carried out by single-stranded phagemid amplification. The DNAzyme produced was tested for its catalytic potential by using it as a component of a DNA nanomotor.

In conclusion, the results presented here provide a better understanding of artificial nucleic acid nanostructures, and open up unexplored domains for design and applications of novel nucleic acid based nanostructures. The modification of DNA nanostructures to sustain in physiological conditions is the key parameter for future applications. As demonstrated, the use of condensing agents for folding offers a viable solution. The need for mass production of nucleic acid based nanostructures points towards the direction of RNA based structures and *in vivo* synthesis of nucleic acid nanostructures.

References

[1] Watson, J. D., Crick, F. H., *et al.* Molecular structure of nucleic acids. *Nature* **171,** 737–738 (1953).

[2] Wetmur, J. G. DNA probes: applications of the principles of nucleic acid hybridization. *Critical reviews in biochemistry and molecular biology* **26,** 227–259 (1991).

[3] Serganov, A. & Patel, D. J. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nature Reviews Genetics* **8,** 776–790 (2007).

[4] Seeman, N. C. Nucleic acid junctions and lattices. *Journal of theoretical biology* **99,** 237–247 (1982).

[5] Jasinski, D., Haque, F., Binzel, D. W. & Guo, P. Advancement of the Emerging Field of RNA Nanotechnology. *ACS nano* **11,** 1142–1164 (2017).

[6] Seeman, N. C. *Structural DNA nanotechnology* (Cambridge University Press, 2016).

[7] Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *nature* **346,** 818–822 (1990).

[8] Whelan, D. R. *et al.* Detection of an en masse and reversible B-to A-DNA conformational transition in prokaryotes in response to desiccation. *Journal of the Royal Society Interface* **11,** 20140454 (2014).

[9] Oh, D.-B., Kim, Y.-G. & Rich, A. Z-DNA-binding proteins can act as potent effectors of gene expression in vivo. *Proceedings of the National Academy of Sciences* **99,** 16666–16671 (2002).

[10] Nelson, D. L., Lehninger, A. L. & Cox, M. M. *Lehninger principles of biochemistry* (Macmillan, 2008).

[11] Sedova, A. & Banavali, N. K. RNA approaches the B-form in stacked single strand dinucleotide contexts. *Biopolymers* **105,** 65–82 (2016).

[12] Macke, T. J. & Case, D. A. in (ACS Publications, 1998).

[13] Kebbekus, P., Draper, D. E. & Hagerman, P. Persistence length of RNA. *Biochemistry* **34,** 4354–4357 (1995).

[14] Tinland, B., Pluen, A., Sturm, J. & Weill, G. Persistence length of single-stranded DNA. *Macromolecules* **30,** 5763–5765 (1997).

[15] 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).

[16] Le Novere, N. MELTING, computing the melting temperature of nucleic acid duplex. *Bioinformatics* **17,** 1226–1227 (2001).

[17] Owczarzy, R. *et al.* IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Research* **36,** W163–W169 (2008).

[18] Chiu, W. L. A. K. *et al.* NTDB: thermodynamic database for nucleic acids, version 2.0. *Nucleic acids research* **31,** 483–485 (2003).

[19] Nazarenko, I., Pires, R., Lowe, B., Obaidy, M. & Rashtchian, A. Effect of primary and secondary structure of oligodeoxyribonucleotides on the fluorescent properties of conjugated dyes. *Nucleic acids research* **30,** 2089–2195 (2002).

[20] SantaLucia Jr, J. & Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **33,** 415–440 (2004).

[21] Freier, S. M. *et al.* Improved free-energy parameters for predictions of RNA duplex stability. *Proceedings of the National academy of Sciences* **83,** 9373–9377 (1986).

[22] Turner, D. H. & Mathews, D. H. NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic acids research* **38,** D280–D282 (2009).

[23] Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* **31,** 3406–3415 (2003).

[24] Hofacker, I. L. Vienna RNA secondary structure server. *Nucleic acids research* **31,** 3429–3431 (2003).

[25] DeRouchey, J. E. & Rau, D. C. Role of Amino Acid Insertions on Intermolecular Forces between Arginine Peptide Condensed DNA Helices implications for protamine-DNA packaging in sperm. *Journal of Biological Chemistry* **286,** 41985–41992 (2011).

[26] Hougaard, D. M. Polyamine cytochemistry: localization and possible functions of polyamines. *International review of cytology* **138,** 51–88 (1992).

[27] Ris, H. Preparation of chromatin and chromosomes for electron microscopy. *Methods in cell biology* **18,** 229–246 (1978).

[28] Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *Journal of molecular biology* **319,** 1097–1113 (2002).

[29] Berman, H. M. *et al.* in *International Tables for Crystallography Volume F: Crystallography of biological macromolecules* 675–684 (Springer, 2006).

[30] Pettersen, E. F. *et al.* UCSF Chimera: a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25,** 1605–1612 (2004).

[31] Qian, Z. *et al.* A new noncoding RNA arranges bacterial chromosome organization. *MBio* **6,** e00998–15 (2015).

[32] Mounce, B. C., Olsen, M. E., Vignuzzi, M. & Connor, J. H. Polyamines and Their Role in Virus Infection. *Microbiology and Molecular Biology Reviews* **81,** e00029–17 (2017).

[33] Veress, I., Haghighi, S., Pulkka, A. & Pajunen, A. Changes in gene expression in response to polyamine depletion indicates selective stabilization of mRNAs. *Biochemical Journal* **346,** 185–191 (2000).

[34] Davis, R. H., Morris, D. R. & Coffino, P. Sequestered end products and enzyme regulation: the case of ornithine decarboxylase. *Microbiological reviews* **56,** 280–290 (1992).

[35] Tabor, C. W. & Tabor, H. Polyamines. *Annual review of biochemistry* **53,** 749–790 (1984).

[36] Gerner, E. W. & Meyskens, F. L. Polyamines and cancer: old molecules, new understanding. *Nature Reviews Cancer* **4,** 781–792 (2004).

[37] Bloomfield, V. A. *et al.* DNA condensation by multivalent cations. *Biopolymers* **44,** 269–282 (1997).

[38] Raspaud, E., De La Cruz, M. O., Sikorav, J.-L. & Livolant, F. Precipitation of DNA by polyamines: a polyelectrolyte behavior. *Biophysical journal* **74,** 381–393 (1998).

[39] Wilson, R. W. & Bloomfield, V. A. Counterion-induced condensation of deoxyribonucleic acid. A light-scattering study. *Biochemistry* **18,** 2192–2196 (1979).

[40] Hud, N. V. & Downing, K. H. Cryoelectron microscopy of λ phage DNA condensates in vitreous ice: the fine structure of DNA toroids. *Proceedings of the National Academy of Sciences* **98,** 14925–14930 (2001).

[41] Moon, R. B. & Richards, J. H. Determination of intracellular pH by 31P magnetic resonance. *Journal of Biological Chemistry* **248,** 7276–7278 (1973).

[42] Baeza, I. *et al.* Possible prebiotic significance of polyamines in the condensation, protection, encapsulation, and biological properties of DNA. *Origins of Life and Evolution of Biospheres* **21,** 225–242 (1992).

[43] Lambert, O., Letellier, L., Gelbart, W. M. & Rigaud, J.-L. DNA delivery by phage as a strategy for encapsulating toroidal condensates of arbitrary size into liposomes. *Proceedings of the National Academy of Sciences* **97,** 7248–7253 (2000).

[44] Arscott, P. G., Ma, C., Wenner, J. R. & Bloomfield, V. A. DNA condensation by cobalt hexaammine (III) in alcohol–water mixtures: dielectric constant and other solvent effects. *Biopolymers* **36,** 345–364 (1995).

[45] Alberts, B. *et al.* Molecular biology of the cell. 3rd. *New York: Garland Pub* **43,** 67 (1994).

[46] Livolant, F. & Leforestier, A. Condensed phases of DNA: structures and phase transitions. *Progress in Polymer Science* **21,** 1115–1164 (1996).

[47] Tabor, H. The protective effect of spermine and other polyamines against heat denaturation of deoxyribonucleic acid. *Biochemistry* **1,** 496–501 (1962).

[48] Douki, T., Bretonniere, Y. & Cadet, J. Protection against radiation-induced degradation of DNA bases by polyamines. *Radiation research* **153,** 29–35 (2000).

[49] Mao, C., Sun, W., Shen, Z. & Seeman, N. C. A nanomechanical device based on the B–Z transition of DNA. *Nature* **397,** 144–146 (1999).

[50] Manning, G. S. Limiting laws and counterion condensation in polyelectrolyte solutions II. Self-diffusion of the small ions. *The Journal of Chemical Physics* **51,** 934–938 (1969).

[51] Oosawa, F. Interaction between parallel rodlike macroions. *Biopolymers* **6,** 1633–1647 (1968).

[52] Baumann, C. G., Smith, S. B., Bloomfield, V. A. & Bustamante, C. Ionic effects on the elasticity of single DNA molecules. *Proceedings of the National Academy of Sciences* **94,** 6185–6190 (1997).

[53] Teif, V. B. & Bohinc, K. Condensed DNA: condensing the concepts. *Progress in biophysics and molecular biology* **105,** 208–222 (2011).

[54] Fang, Y., Hoh, J. H. & Spisz, T. S. Ethanol-induced structural transitions of DNA on mica. *Nucleic acids research* **27,** 1943–1949 (1999).

[55] Todd, B. A., Parsegian, V. A., Shirahata, A., Thomas, T. & Rau, D. C. Attractive forces between cation condensed DNA double helices. *Biophysical journal* **94,** 4775–4782 (2008).

[56] Rau, D., Lee, B. & Parsegian, V. Measurement of the repulsive force between polyelectrolyte molecules in ionic solution: hydration forces between parallel DNA double helices. *Proceedings of the National Academy of Sciences* **81,** 2621–2625 (1984).

[57] Yoo, J. & Aksimentiev, A. The structure and intermolecular forces of DNA condensates. *Nucleic acids research* **44,** 2036–2046 (2016).

[58] Holliday, R. A mechanism for gene conversion in fungi. *Genetics Research* **5,** 282–304 (1964).

[59] Broker, T. R. & Lehman, I. Branched DNA molecules: intermediates in T4 recombination. *Journal of molecular biology* **60,** 131IN7137–136IN11149 (1971).

[60] Kallenbach, N. R., Ma, R.-I. & Seeman, N. C. An immobile nucleic acid junction constructed from oligonucleotides. *Nature* **305,** 829–831 (1983).

[61] Seeman, N. C. & Kallenbach, N. R. DNA branched junctions. *Annual review of biophysics and biomolecular structure* **23,** 53–86 (1994).

[62] Winz, M.-L., Linder, E. C., Andre, T., Becker, J. & Jäschke, A. Nucleotidyl transferase assisted DNA labeling with different click chemistries. *Nucleic acids research* **43,** e110–e110 (2015).

[63] Jahn, K. *et al.* Functional patterning of DNA origami by parallel enzymatic modification. *Bioconjugate chemistry* **22,** 819–823 (2011).

[64] Fu, T. J. & Seeman, N. C. DNA double-crossover molecules. *Biochemistry* **32,** 3211–3220 (1993).

[65] Gilbert, D. E. & Feigon, J. Multistranded DNA structures. *Current opinion in structural biology* **9,** 305–314 (1999).

[66] Yan, H., LaBean, T. H., Feng, L. & Reif, J. H. Directed nucleation assembly of DNA tile complexes for barcode-patterned lattices. *Proceedings of the National Academy of Sciences* **100,** 8103–8108 (2003).

[67] Shih, W. M., Quispe, J. D. & Joyce, G. F. A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature* **427,** 618–621 (2004).

[68] Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).

[69] Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459,** 414–418 (2009).

[70] Andersen, E. S. *et al.* Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459,** 73–76 (2009).

[71] Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325,** 725–730 (2009).

[72] Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485,** 623–626 (2012).

[73] Ke, Y., Ong, L. L., Shih, W. M. & Yin, P. Three-dimensional structures self-assembled from DNA bricks. *science* **338,** 1177–1183 (2012).

[74] Zhang, F. *et al.* Complex wireframe DNA origami nanostructures with multi-arm junction vertices. *Nature nanotechnology* **10,** 779–784 (2015).

[75] Benson, E. *et al.* DNA rendering of polyhedral meshes at the nanoscale. *Nature* **523,** 441–444 (2015).

[76] 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).

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

[78] Zheng, J. *et al.* From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **461,** 74–77 (2009).

[79] Han, D. *et al.* DNA origami with complex curvatures in three-dimensional space. *Science* **332,** 342–346 (2011).

[80] Douglas, S. M. *et al.* Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic acids research* **37,** 5001–5006 (2009).

[81] Zadeh, J. N. *et al.* NUPACK: analysis and design of nucleic acid systems. *Journal of computational chemistry* **32,** 170–173 (2011).

[82] Castro, C. E. *et al.* A primer to scaffolded DNA origami. *Nature methods* **8,** 221–229 (2011).

[83] Kick, B., Praetorius, F., Dietz, H. & Weuster-Botz, D. Efficient production of single-stranded phage DNA as scaffolds for DNA origami. *Nano letters* **15,** 4672–4676 (2015).

[84] Ducani, C., Kaul, C., Moche, M., Shih, W. M. & Högberg, B. Enzymatic production of'monoclonal stoichiometric'single-stranded DNA oligonucleotides. *Nature methods* **10,** 647–652 (2013).

[85] Stahl, E., Martin, T. G., Praetorius, F. & Dietz, H. Facile and scalable preparation of pure and dense DNA origami solutions. *Angewandte Chemie International Edition* **53,** 12735–12740 (2014).

[86] Shaw, A., Benson, E. & Högberg, B. Purification of functionalized DNA origami nanostructures. *ACS nano* **9,** 4968–4975 (2015).

[87] Elbaz, J., Yin, P. & Voigt, C. A. Genetic encoding of DNA nanostructures and their self-assembly in living bacteria. *Nature communications* **7** (2016).

[88] Lin, C. *et al.* In vivo cloning of artificial DNA nanostructures. *Proceedings of the National Academy of Sciences* **105,** 17626–17631 (2008).

[89] Adleman, L. M. Molecular computation of solutions to combinatorial problems. *Nature* **369,** 40 (1994).

[90] Winfree, E. *Algorithmic self-assembly of DNA* PhD thesis (California Institute of Technology, 1998).

[91] Zhang, D. Y. & Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature chemistry* **3,** 103–113 (2011).

[92] Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406,** 605–608 (2000).

[93] Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475,** 368–372 (2011).

[94] Modi, S. *et al.* A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nature Nanotechnology* **4,** 325–330 (2009).

[95] Groves, B. *et al.* Computing in mammalian cells with nucleic acid strand exchange. *Nature nanotechnology* **11,** nnano–2015 (2015).

[96] Elbaz, J. *et al.* DNA computing circuits using libraries of DNAzyme subunits. *Nature nanotechnology* **5,** 417–422 (2010).

[97] Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* **101,** 15275–15278 (2004).

[98] Shin, J.-S. & Pierce, N. A. A synthetic DNA walker for molecular transport. *Journal of the American Chemical Society* **126,** 10834–10835 (2004).

[99] Li, J., Green, A. A., Yan, H. & Fan, C. Engineering nucleic acid structures for programmable molecular circuitry and intracellular biocomputation. *Nature chemistry* **9,** 1056 (2017).

[100] Batey, R. T., Rambo, R. P., Doudna, J. A., *et al.* Tertiary motifs in RNA structure and folding. *Angewandte Chemie International Edition* **38,** 2326–2343 (1999).

[101] Leontis, N. B., Lescoute, A. & Westhof, E. The building blocks and motifs of RNA architecture. *Current opinion in structural biology* **16,** 279–287 (2006).

[102] Draper, D. E. RNA folding: thermodynamic and molecular descriptions of the roles of ions. *Biophysical journal* **95,** 5489–5495 (2008).

[103] Guo, P., Zhang, C., Chen, C., Garver, K. & Trottier, M. Inter-RNA interaction of phage φ29 pRNA to form a hexameric complex for viral DNA transportation. *Molecular cell* **2,** 149–155 (1998).

[104] Jaeger, L., Westhof, E. & Leontis, N. B. TectoRNA: modular assembly units for the construction of RNA nano-objects. *Nucleic acids research* **29,** 455–463 (2001).

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

[106] Han, D. *et al.* Single-stranded DNA and RNA origami. *Science* **358,** eaao2648 (2017).

[107] Lee, J. B., Hong, J., Bonner, D. K., Poon, Z. & Hammond, P. T. Self-assembled RNA interference microsponges for efficient siRNA delivery. *Nature materials* **11,** 316–322 (2012).

[108] 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).

[109] Sachdeva, G., Garg, A., Godding, D., Way, J. C. & Silver, P. A. In vivo co-localization of enzymes on RNA scaffolds increases metabolic production in a geometrically dependent manner. *Nucleic acids research* **42,** 9493–9503 (2014).

[110] Paige, J. S., Nguyen-Duc, T., Song, W. & Jaffrey, S. R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335,** 1194–1194 (2012).

[111] Song, W. *et al.* Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex. *Nature Chemical Biology* (2017).

[112] Khisamutdinov, E. F. *et al.* Enhancing immunomodulation on innate immunity by shape transition among RNA triangle, square and pentagon nanovehicles. *Nucleic acids research* **42,** 9996–10004 (2014).

[113] Afonin, K. A. *et al.* Multifunctional RNA nanoparticles. *Nano letters* **14,** 5662–5671 (2014).

[114] Zalatan, J. G. *et al.* Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160,** 339–350 (2015).

[115] Boniecki, M. J. *et al.* SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic acids research* **44,** e63–e63 (2016).

[116] Jossinet, F., Ludwig, T. E. & Westhof, E. Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. *Bioinformatics* **26,** 2057–2059 (2010).

[117] Bindewald, E., Hayes, R., Yingling, Y. G., Kasprzak, W. & Shapiro, B. A. RNAJunction: a database of RNA junctions and kissing loops for three-dimensional structural analysis and nanodesign. *Nucleic acids research* **36,** D392–D397 (2007).

[118] Chojnowski, G., Waleń, T. & Bujnicki, J. M. RNA Bricks: a database of RNA 3D motifs and their interactions. *Nucleic acids research* **42,** D123–D131 (2014).

[119] Seeman, N. C. Nanomaterials based on DNA. *Annual review of biochemistry* **79,** 65–87 (2010).

[120] Krishnan, Y. & Simmel, F. C. Nucleic acid based molecular devices. *Angewandte Chemie International Edition* **50,** 3124–3156 (2011).

[121] Acuna, G. *et al.* Fluorescence enhancement at docking sites of DNA-directed self-assembled nanoantennas. *Science* **338,** 506–510 (2012).

[122] Rinker, S., Ke, Y., Liu, Y., Chhabra, R. & Yan, H. Self-assembled DNA nanostructures for distance-dependent multivalent ligand–protein binding. *Nature nanotechnology* **3,** 418–422 (2008).

[123] Lin, C. *et al.* Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA. *Nature chemistry* **4,** 832–839 (2012).

[124] Shaw, A. *et al.* Spatial control of membrane receptor function using ligand nanocalipers. *Nature methods* **11,** 841–846 (2014).

[125] Yang, Y. *et al.* Self-assembly of size-controlled liposomes on DNA nanotemplates. *Nature chemistry* **8,** 476–483 (2016).

[126] Langecker, M. *et al.* Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science* **338,** 932–936 (2012).

[127] Schreiber, R. *et al.* Hierarchical assembly of metal nanoparticles, quantum dots and organic dyes using DNA origami scaffolds. *Nature nanotechnology* **9,** 74–78 (2014).

[128] Hahn, J., Wickham, S. F., Shih, W. M. & Perrault, S. D. Addressing the instability of DNA nanostructures in tissue culture. *ACS nano* **8,** 8765–8775 (2014).

[129] Cassinelli, V. *et al.* One-Step Formation of 'Chain-Armor'-Stabilized DNA Nanostructures. *Angewandte Chemie International Edition* **54,** 7795–7798 (2015).

[130] Martin, T. G. & Dietz, H. Magnesium-free self-assembly of multi-layer DNA objects. *Nature communications* **3,** 1103 (2012).

[131] Gállego, I., Grover, M. A. & Hud, N. V. Folding and Imaging of DNA Nanostructures in Anhydrous and Hydrated Deep-Eutectic Solvents. *Angewandte Chemie International Edition* **54,** 6765–6769 (2015).

[132] Eskelinen, A.-P., Rosilo, H., Kuzyk, A., Törmä, P. & Kostiainen, M. A. Controlling the formation of DNA origami structures with external signals. *Small* **8,** 2016–2020 (2012).

[133] Opherden, L., Oertel, J., Barkleit, A., Fahmy, K. & Keller, A. Paramagnetic decoration of DNA origami nanostructures by Eu^3+ coordination. *Langmuir* **30,** 8152–8159 (2014).

[134] Myhrvold, C., Dai, M., Silver, P. A. & Yin, P. Isothermal self-assembly of complex DNA structures under diverse and biocompatible conditions. *Nano letters* **13,** 4242–4248 (2013).

[135] Ponnuswamy, N. *et al.* Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nature communications* **8,** 15654 (2017).

[136] Mikkila, J. *et al.* Virus-encapsulated DNA origami nanostructures for cellular delivery. *Nano letters* **14,** 2196–2200 (2014).

[137] Schüller, V. J. *et al.* Cellular immunostimulation by CpG-sequence-coated DNA origami structures. *ACS nano* **5,** 9696–9702 (2011).

[138] Walsh, A. S., Yin, H., Erben, C. M., Wood, M. J. & Turberfield, A. J. DNA cage delivery to mammalian cells. *ACS nano* **5,** 5427–5432 (2011).

[139] Perrault, S. D. & Shih, W. M. Virus-inspired membrane encapsulation of DNA nanostructures to achieve in vivo stability. *ACS nano* **8,** 5132–5140 (2014).

[140] Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature methods* **9,** 676–682 (2012).

[141] De la Rosa-Trevín, J. *et al.* Scipion: a software framework toward integration, reproducibility and validation in 3D electron microscopy. *Journal of structural biology* **195,** 93–99 (2016).

[142] De La Cruz, M. O. *et al.* Precipitation of highly charged polyelectrolyte solutions in the presence of multivalent salts. *The Journal of chemical physics* **103,** 5781–5791 (1995).

[143] Deng, H., Bloomfield, V. A., Benevides, J. M. & Jr, G. J. T. Structural basis of polyamine–DNA recognition: spermidine and spermine interactions with genomic B-DNAs of different GC content probed by Raman spectroscopy. *Nucleic acids research* **28,** 3379–3385 (2000).

[144] Ouameur, A. A. & Tajmir-Riahi, H.-A. Structural analysis of DNA interactions with biogenic polyamines and cobalt (III) hexamine studied by Fourier transform infrared and capillary electrophoresis. *Journal of Biological Chemistry* **279,** 42041–42054 (2004).

[145] Van Dam, L., Korolev, N. & Nordenskiöld, L. Polyamine–nucleic acid interactions and the effects on structure in oriented DNA fibers. *Nucleic acids research* **30,** 419–428 (2002).

[146] Bai, X.-c., Martin, T. G., Scheres, S. H. & Dietz, H. Cryo-EM structure of a 3D DNA-origami object. *Proceedings of the National Academy of Sciences* **109,** 20012–20017 (2012).

[147] Kopperger, E., Pirzer, T. & Simmel, F. C. Diffusive transport of molecular cargo tethered to a DNA origami platform. *Nano letters* **15,** 2693–2699 (2015).

[148] Pelta, J., Durand, D., Doucet, J. & Livolant, F. DNA mesophases induced by spermidine: structural properties and biological implications. *Biophysical journal* **71,** 48–63 (1996).

[149] Raspaud, E., Durand, D. & Livolant, F. Interhelical spacing in liquid crystalline spermine and spermidine-DNA precipitates. *Biophysical journal* **88,** 392–403 (2005).

[150] Viola, J. *et al.* Fatty acid–spermine conjugates as DNA carriers for nonviral in vivo gene delivery. *Gene therapy* **16,** 1429–1440 (2009).

[151] Zhang, Q. *et al.* DNA origami as an in vivo drug delivery vehicle for cancer therapy. *Acs Nano* **8,** 6633–6643 (2014).

[152] Douglas, S. M., Bachelet, I. & Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335,** 831–834 (2012).

[153] Mei, Q. *et al.* Stability of DNA origami nanoarrays in cell lysate. *Nano letters* **11,** 1477–1482 (2011).

[154] Rosazza, C. *et al.* Intracellular tracking of single-plasmid DNA particles after delivery by electroporation. *Molecular therapy* **21,** 2217–2226 (2013).

[155] Neumann, E. & Rosenheck, K. Permeability changes induced by electric impulses in vesicular membranes. *The Journal of membrane biology* **10,** 279–290 (1972).

[156] Strober, W. Trypan blue exclusion test of cell viability. *Current protocols in immunology,* A3–B (2001).

[157] Mittal, N., Rosenthal, A. & Voldman, J. nDEP microwells for single-cell patterning in physiological media. *Lab on a Chip* **7,** 1146–1153 (2007).

[158] Dower, W. J., Miller, J. F. & Ragsdale, C. W. High efficiency transformation of E. coli by high voltage electroporation. *Nucleic acids research* **16,** 6127–6145 (1988).

[159] Yoo, J. & Aksimentiev, A. Molecular dynamics of membrane-spanning DNA channels: conductance mechanism, electro-osmotic transport, and mechanical gating. *The journal of physical chemistry letters* **6,** 4680–4687 (2015).

[160] Bellot, G., McClintock, M. A., Lin, C. & Shih, W. M. Recovery of intact DNA nanostructures after agarose gel–based separation. *Nature methods* **8,** 192 (2011).

[161] Mercer, T. R., Dinger, M. E. & Mattick, J. S. Long non-coding RNAs: insights into functions. *Nature Reviews Genetics* **10,** 155–159 (2009).

[162] Osborne, R. J. & Thornton, C. A. RNA-dominant diseases. *Human molecular genetics* **15,** R162–R169 (2006).

[163] Jepsen, M. D. *et al.* Development of a genetically encodable FRET system using fluorescent RNA aptamers. *Nature communications* **9,** 18 (2018).

[164] Geary, C., Chworos, A., Verzemnieks, E., Voss, N. R. & Jaeger, L. Composing RNA nanostructures from a syntax of RNA structural modules. *Nano letters* **17,** 7095–7101 (2017).

[165] Sparvath, S. L., Geary, C. W. & Andersen, E. S. Computer-Aided Design of RNA Origami Structures. *3D DNA Nanostructure: Methods and Protocols,* 51–80 (2017).

[166] Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *electrophoresis* **18,** 2714–2723 (1997).

[167] Laptenok, S. P., Addison, K., Heisler, I. A. & Meech, S. R. Excited state structural dynamics in higher lying electronic states: S2 state of malachite green. *Chemical Physics Letters* **607,** 43–46 (2014).

[168] Conrado, R. J., Varner, J. D. & DeLisa, M. P. Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Current opinion in biotechnology* **19,** 492–499 (2008).

[169] Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nature biotechnology* **27,** 753 (2009).

[170] Castellana, M. *et al.* Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nature biotechnology* **32,** 1011 (2014).

[171] Siu, K.-H. *et al.* Synthetic scaffolds for pathway enhancement. *Current opinion in biotechnology* **36,** 98–106 (2015).

[172] Buchner, A., Tostevin, F. & Gerland, U. Clustering and optimal arrangement of enzymes in reaction-diffusion systems. *Physical review letters* **110,** 208104 (2013).

[173] Mohsen, M. G. & Kool, E. T. The discovery of rolling circle amplification and rolling circle transcription. *Accounts of chemical research* **49,** 2540–2550 (2016).

[174] Arnold, S. *et al.* Kinetic modeling and simulation of in vitro transcription by phage T 7 RNA polymerase. *Biotechnology and bioengineering* **72,** 548–561 (2001).

[175] Chen, Y., Wang, M. & Mao, C. An autonomous DNA nanomotor powered by a DNA enzyme. *Angewandte Chemie International Edition* **43,** 3554–3557 (2004).

[176] Santoro, S. W. & Joyce, G. F. A general purpose RNA-cleaving DNA enzyme. *Proceedings of the National Academy of Sciences* **94,** 4262–4266 (1997).

[177] Kosuri, S. *et al.* Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nature biotechnology* **28,** 1295 (2010).

[178] Chen, Y.-J. *et al.* Programmable chemical controllers made from DNA. *Nature nanotechnology* **8,** 755 (2013).

[179] Praetorius, F. *et al.* Biotechnological mass production of DNA origami. *Nature* **552,** 84 (2017).

Publications

Electrotransfection of polyamine folded DNA origami structures. A. Chopra, S. Krishnan, and F.C. Simmel. *Nano letters, 16(10), 6683-6690*, 2016.

Appendix

DNA origami nanostructures folding using multivalent ions

Cadnano files for the DNA origami structures

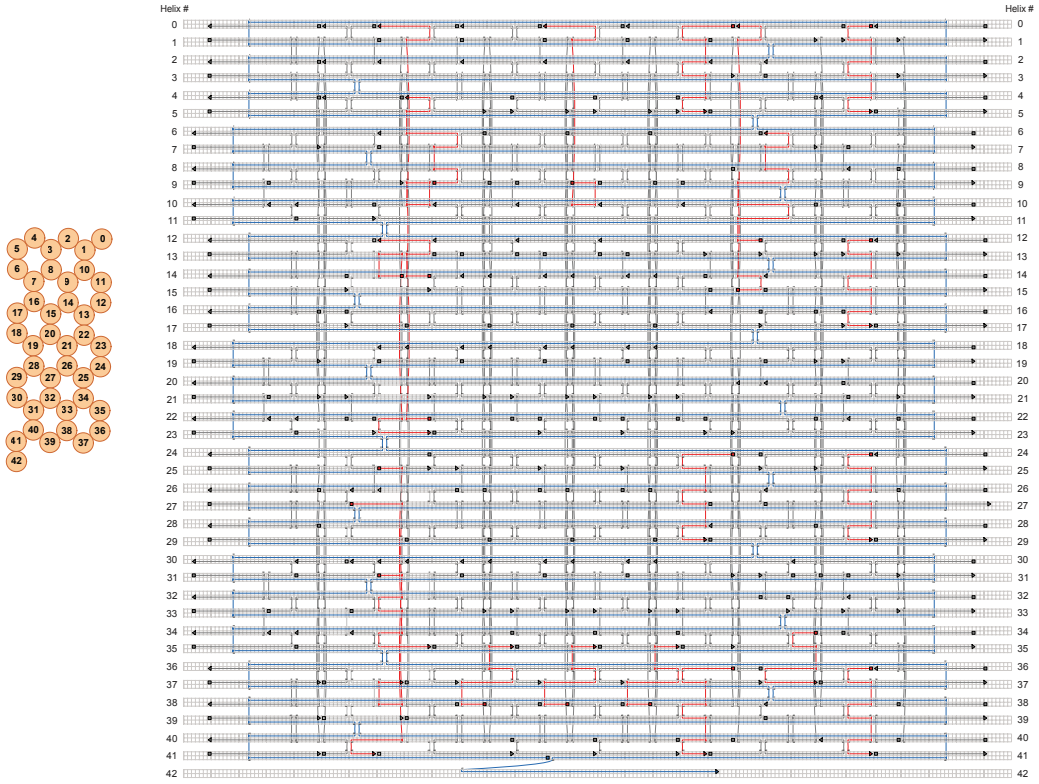


Figure 48: Cadnano diagram of 42 helix bundle DNA nanostructures used in folding and electrotransfection experiments. The scaffold strand is depicted in blue and staple strands are in black. The staple strands which are extended for hybridization with dye labeled complementary strands are shown in red.

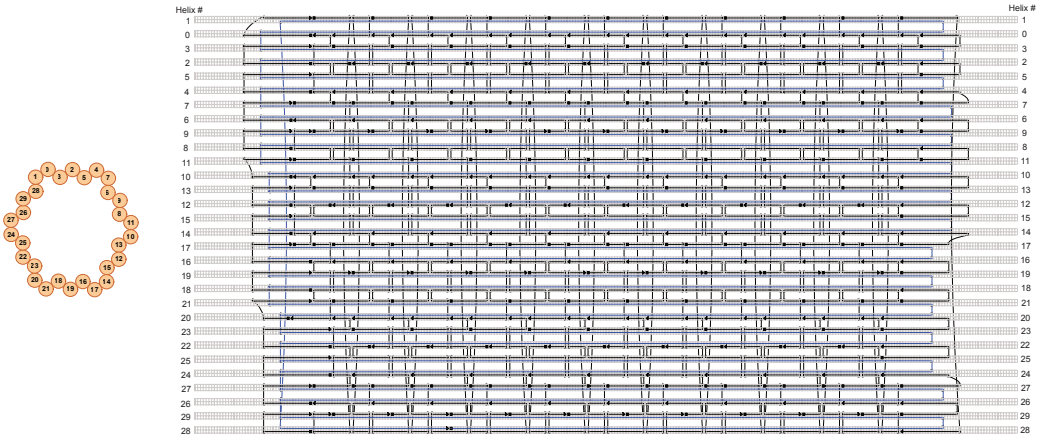


Figure 49: Cadnano diagram of tube nanostructures used in Spd^{3+} mediated folding.

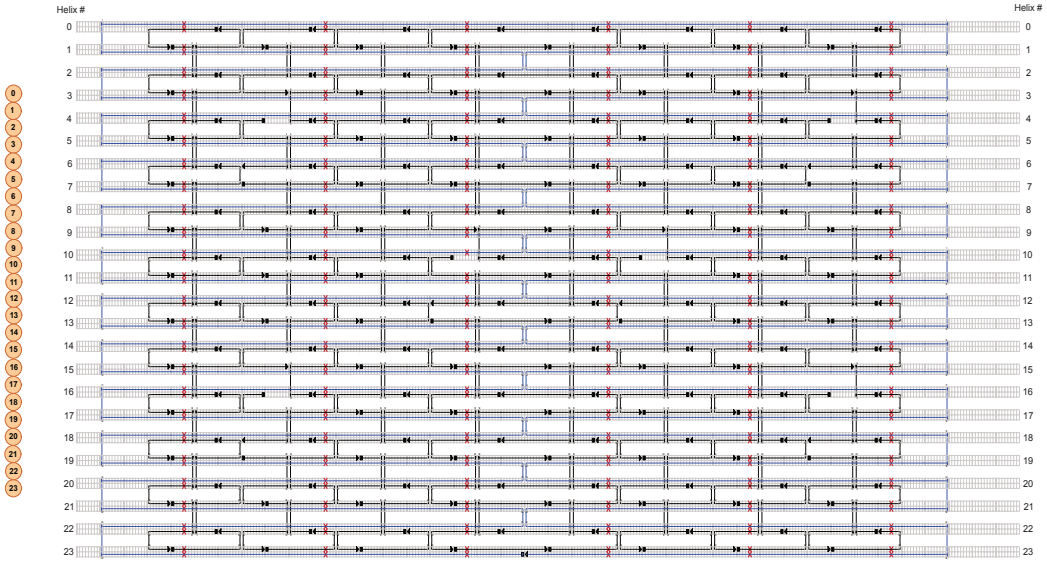


Figure 50: Cadnano diagram of rectangular DNA origami (tCRRO) used in Spd^{3+} mediated folding.

Screening of Spd^{3+} for folding of tube and rectangle DNA nanostructures

Agarose gels showing screening of Spd^{3+} concentration for folding of tube and rectangle (tCRRO) DNA origami structures. Image generated by sketchup and class averaged TEM image of each structure is shown.

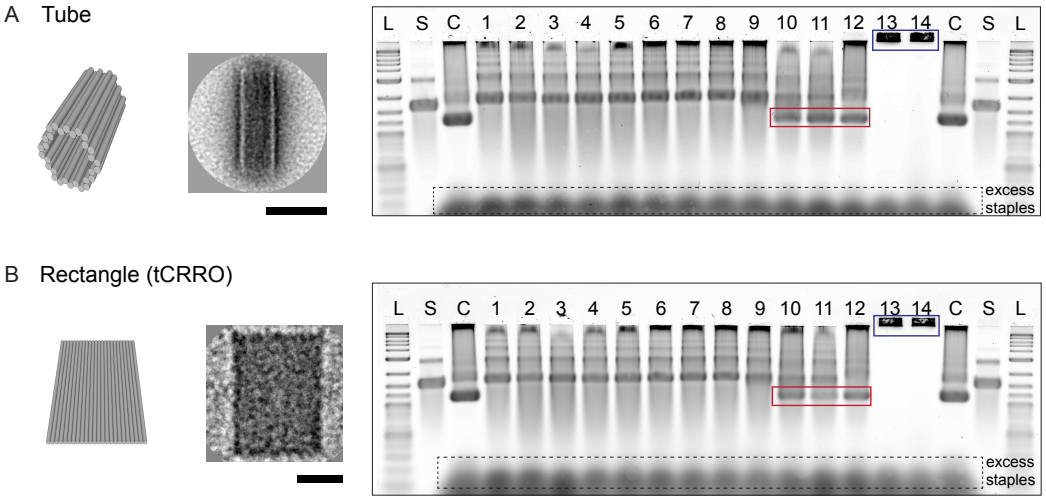


Figure 51: Screening of Spd^{3+} concentrations required for folding 50 nM DNA origami nanostructures (Tube and tCRRO) (A) Tubes, (B) tCRRO. L: 2-Log Ladder, C: DNA origami nanostructures folded in traditional buffer containing magnesium (control). Spd^{3+} concentrations in the numbered lanes are: 1: 0 µM, 2: 50 µM, 3: 100 µM, 4: 200 µM, 5: 300 µM, 6: 400 µM, 7: 500 µM, 8: 600 µM, 9: 700 µM, 10: 800 µM, 11: 900 µM, 12: 1 mM, 13: 5 mM, 14: 10 mM. Red boxes highlight bands corresponding to the correctly folded DNA origami nanostructures and blue boxes show aggregation. Scale bar: 50 nm.

TEM

Montages of TEM micrographs showing folding of DNA nanostructures in different buffers.

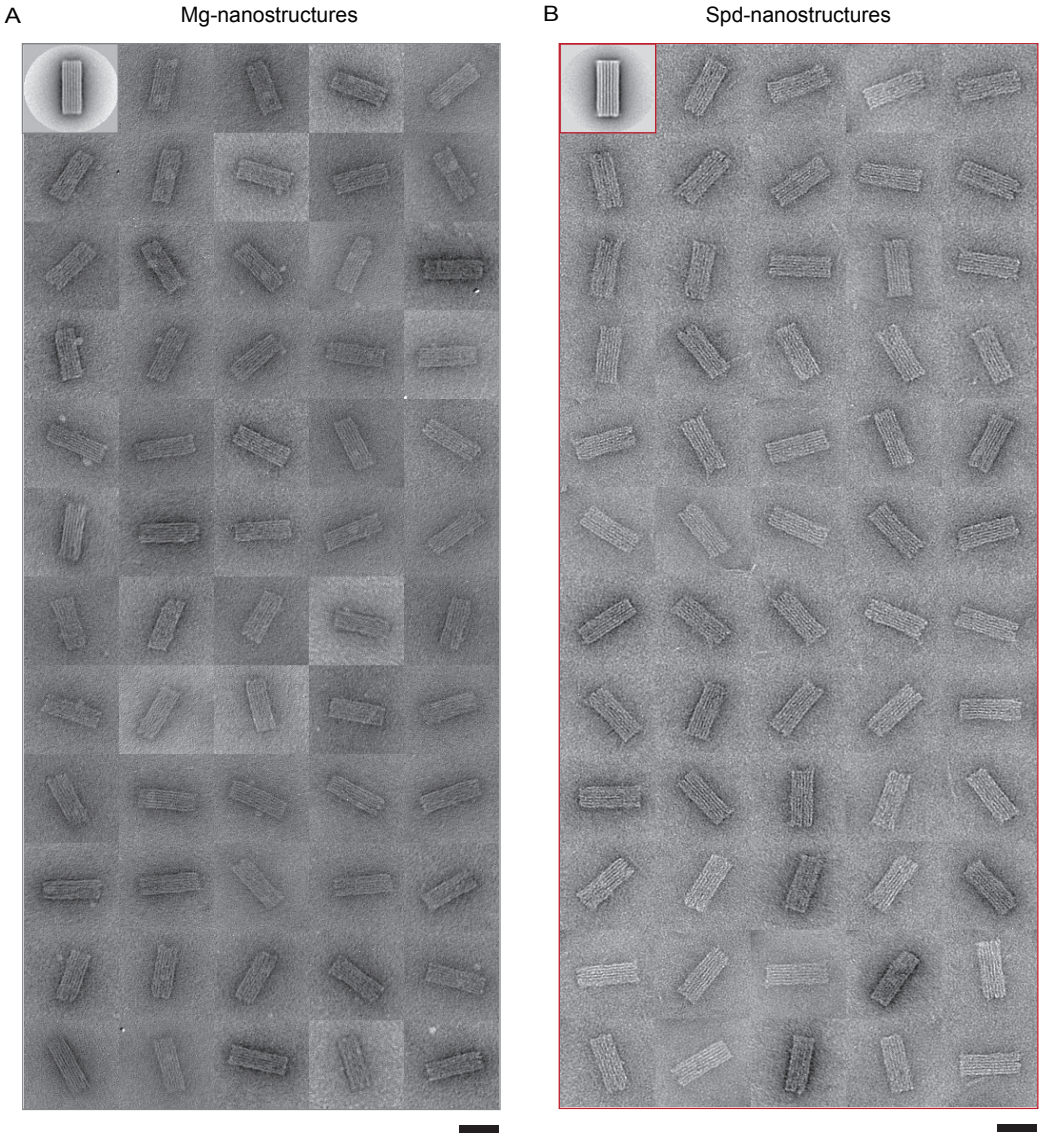


Figure 52: Negative-stain TEM micrographs of 42HB folded in traditional buffer containing $MgCl_2$ (A) and Spd^{3+} (B) respectively. Each micrograph is 100.5 nm x 100.5 nm. Top row, image #1 = class average micrograph. Scale bar: 50 nm. The dimensions of 42HB (length = 60 nm, width = 26 nm, height $\approx$ 16 nm) were found to be the same when folded in either magnesium containing buffer or Spd^{3+}.

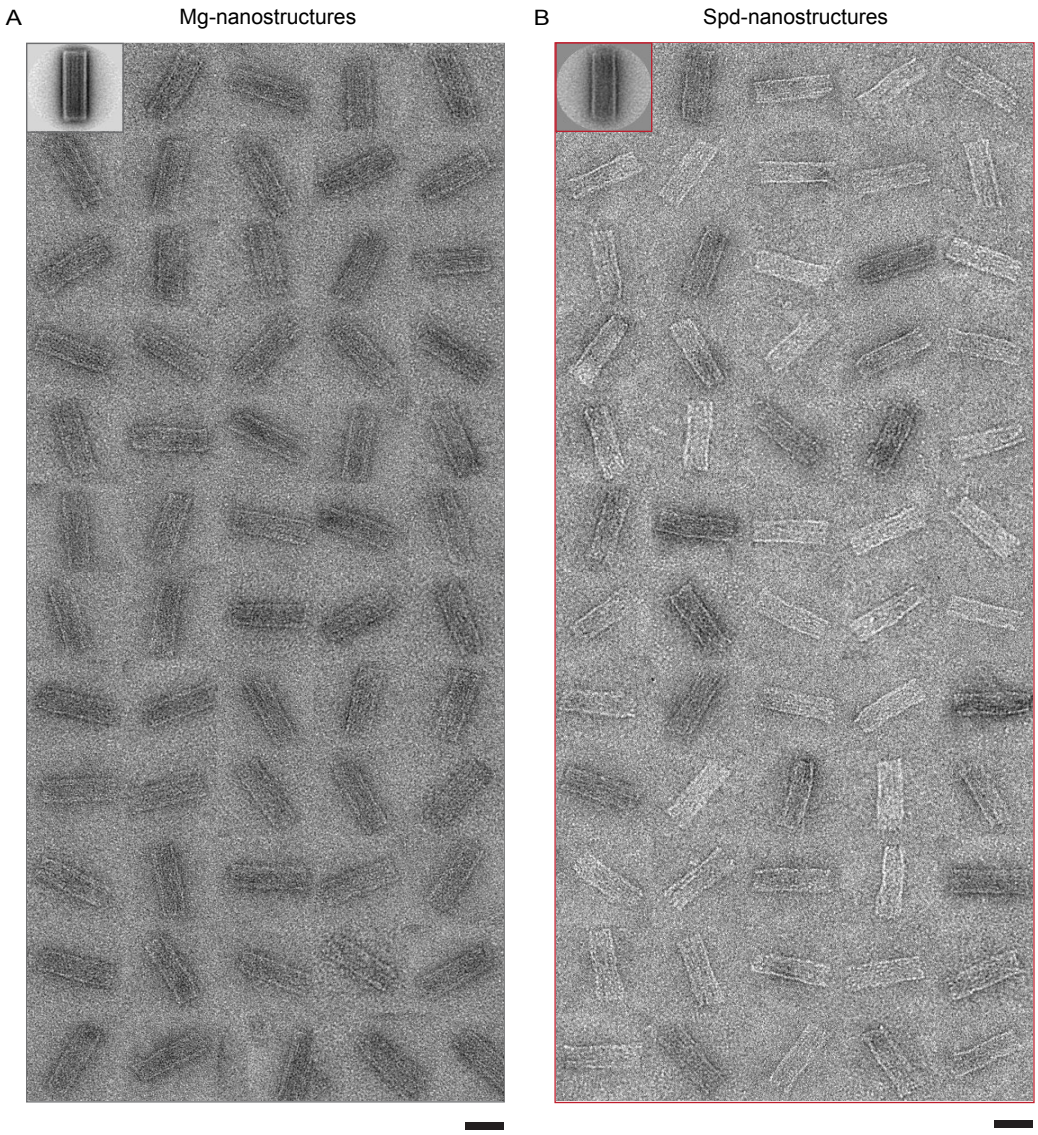


Figure 53: Negative-stain TEM micrographs of tube folded in traditional buffer containing $MgCl_2$ (A) and Spd^{3+} (B) respectively. Each micrograph is 106.6 nm x 106.6 nm. Top row, image #1 = class average micrograph. Scale bar: 50 nm. The dimensions of tube (length = 83 nm, diameter 1 $\approx$ 23 nm, diameter 2 $\approx$ 29 nm) were found to be the same when folded in either magnesium containing buffer or Spd^{3+}.

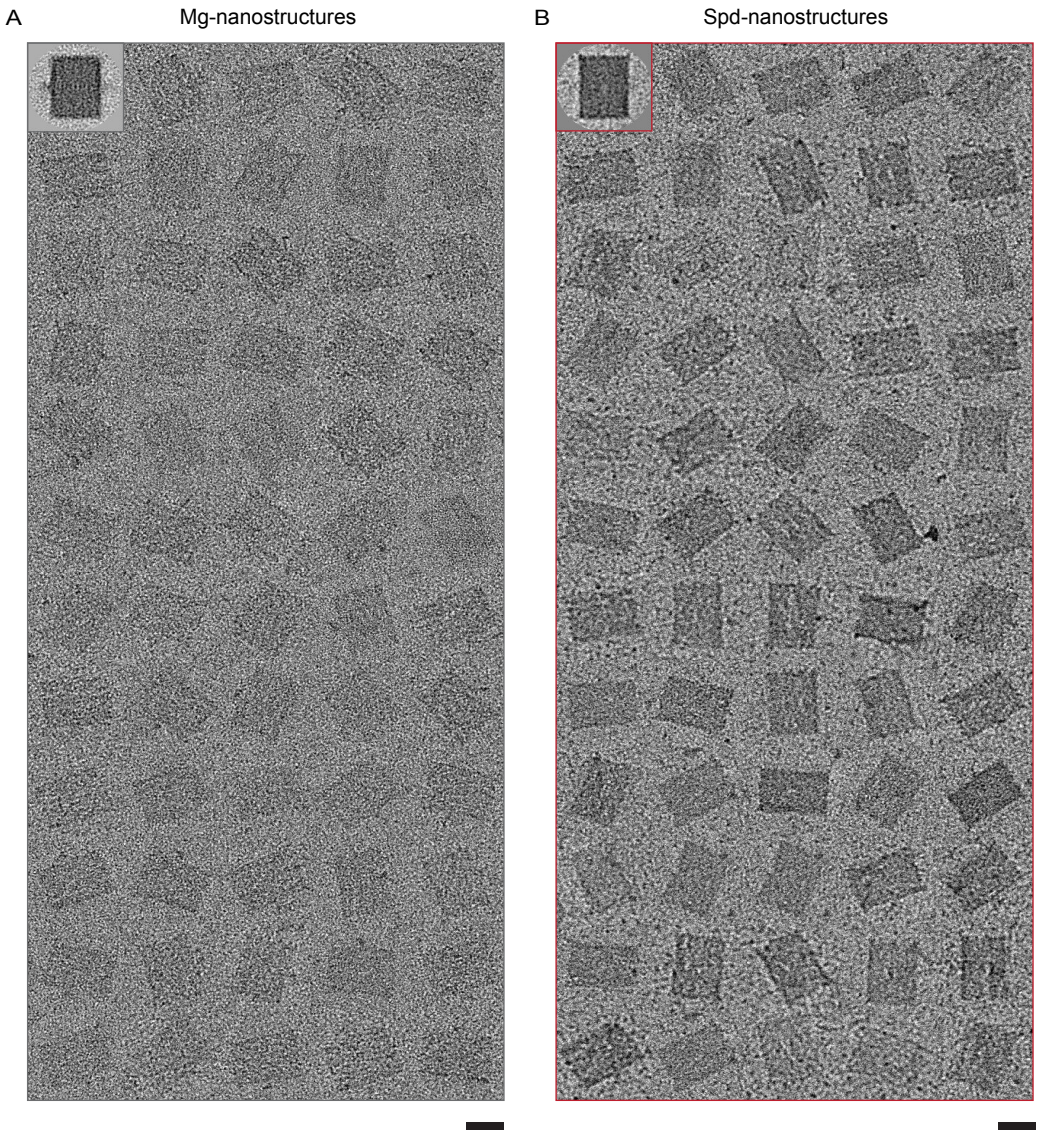


Figure 54: Positive-stain TEM micrographs of tCRRO folded in traditional buffer containing magnesium (A) and Spd^{3+} (B) respectively. Each micrograph: 125.7 nm x 125.7 nm. Top row, image #1 = class average micrograph. Scale bar: 50 nm. The dimensions of tube (length = 90 nm, width = 65 nm, height = 2 nm) were found to be the same when folded in either magnesium containing buffer or Spd^{3+}.

Characterization and applications of Spd^{3+}-folded DNA origami structures

Effect of electric fields on stability of DNA origami

Stability of Mg/Spd nanostructures (42HB) diluted in RPMI and subjected to a range of electric field strengths in both LV and HV mode, as verified by agarose gel electrophoresis.

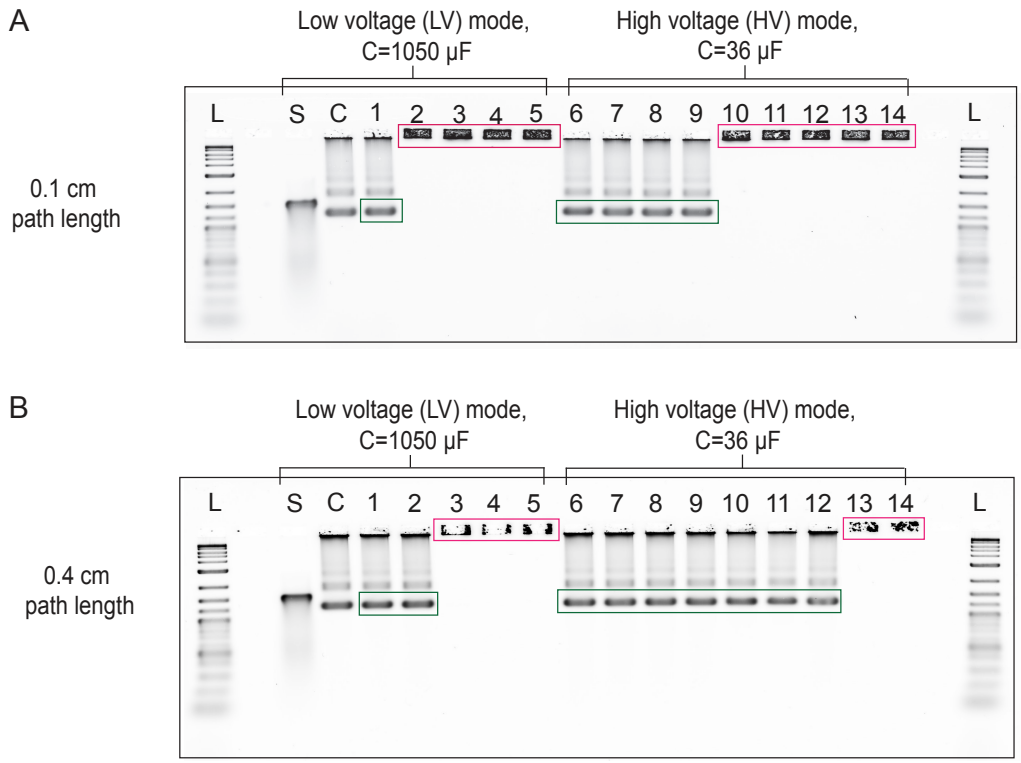


Figure 55: Stability of Mg-nanostructures (42HB) subjected to a range of electric field strengths. L: 2-Log Ladder, C: Control (No electric field). (A) 0.1 cm path length. Electric fields corresponding to the numbered lanes are: 1: 100 V/cm, 2: 500 V/cm, 3: 1000 V/cm, 4: 2000 V/cm, 5: 4000 V/cm in the low voltage mode (C = 1050 1050 µF) and 6: 500 V/cm, 7: 1000 V/cm, 8: 2000 V/cm, 9: 4000 V/cm, 10: 8000 V/cm, 11: 12000 V/cm, 12: 16000 V/cm, 13: 17600 V/cm, 14: 20000 V/cm in the high voltage mode (C = 36 µF). (B) 0.4 cm path length. Electric fields corresponding to the numbered lanes are: 1: 25 V/cm, 2: 125 V/cm, 3: 250 V/cm, 4: 500 V/cm, 5: 1000 V/cm in the low voltage mode (C = 1050 µF and 6: 125 V/cm, 7: 250 V/cm, 8: 500 V/cm, 9: 1000 V/cm, 10: 2000 V/cm, 11: 3000 V/cm, 12: 4000 V/cm, 13: 4400 V/cm, 14: 5000 V/cm in the high voltage mode (C = 36 µF). Green boxes: survived structures, red boxes: destroyed structures.

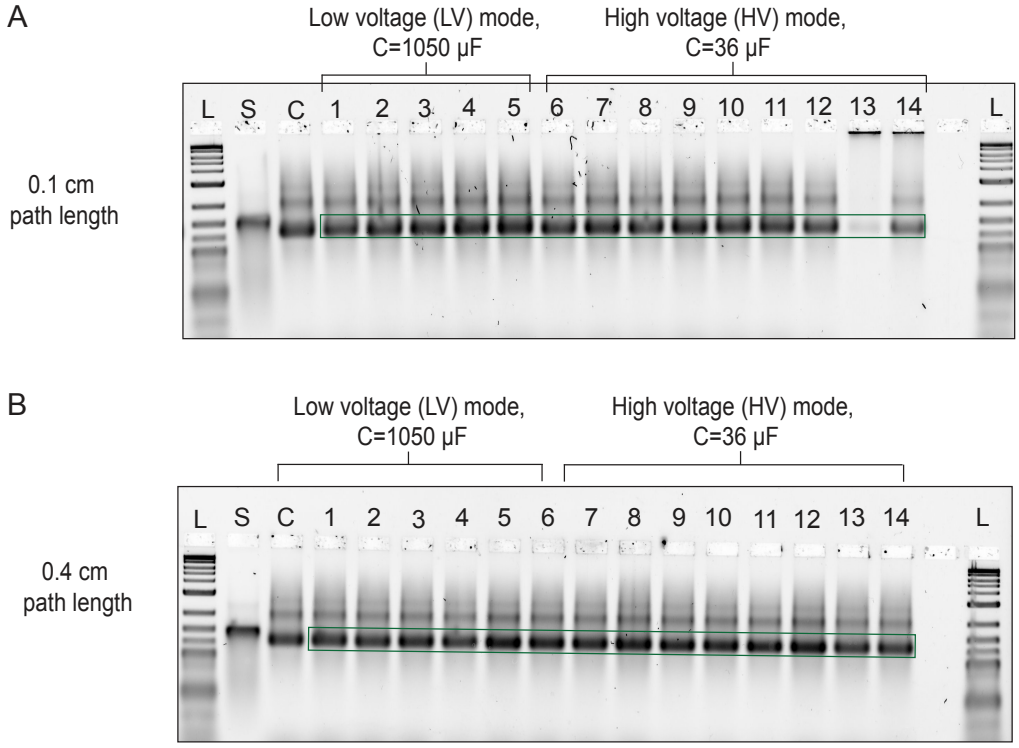


Figure 56: Stability of Spd-nanostructures (42HB) subjected to a range of electric field strengths. L: 2-Log Ladder, C: Control (No electric field) (A) 0.1 cm path length. Electric fields corresponding to the numbered lanes are: 1: 100 V/cm, 2: 500 V/cm, 3: 1000V /cm, 4: 2000 V/cm, 5: 4000 V/cm in the low voltage mode (C = 1050 µF) and 6: 500 V/cm, 7: 1000 V/cm, 8: 2000 V/cm, 9: 4000 V/cm, 10: 8000 V/cm, 11: 12000 V/cm, 12: 16000 V/cm, 13: 17600 V/cm, 14: 20000 V/cm in the high voltage mode (C = 36 µF). At electric field strengths of E ≈ 17600 V/cm and higher, bands appear in the pockets indicating that the stability of the structures is compromised in this case (wells 13, 14). (B) 0.4 cm path length. Electric fields corresponding to the numbered lanes are: 1: 25 V/cm, 2: 125 V/cm, 3: 250 V/cm, 4: 500 V/cm, 5: 1000 V/cm in the low voltage mode (C = 1050 µF) and 6: 125 V/cm, 7: 250 V/cm, 8: 500 V/cm, 9: 1000 V/cm, 10: 2000 V/cm, 11: 3000 V/cm, 12: 4000 V/cm, 13: 4400 V/cm, 14: 5000 V/cm in the high voltage mode (C = 36 µF). Green boxes: survived structures, red boxes: destroyed structures.

.

Effect of higher field strengths, using smaller 0.1 cm cuvettes, and shorter pulse lengths was investigated. The structures were diluted in high resistance medium i.e. 10% v/v glycerol in water). The survival of Mg/Spd- nanostructures is shown in figure 57.

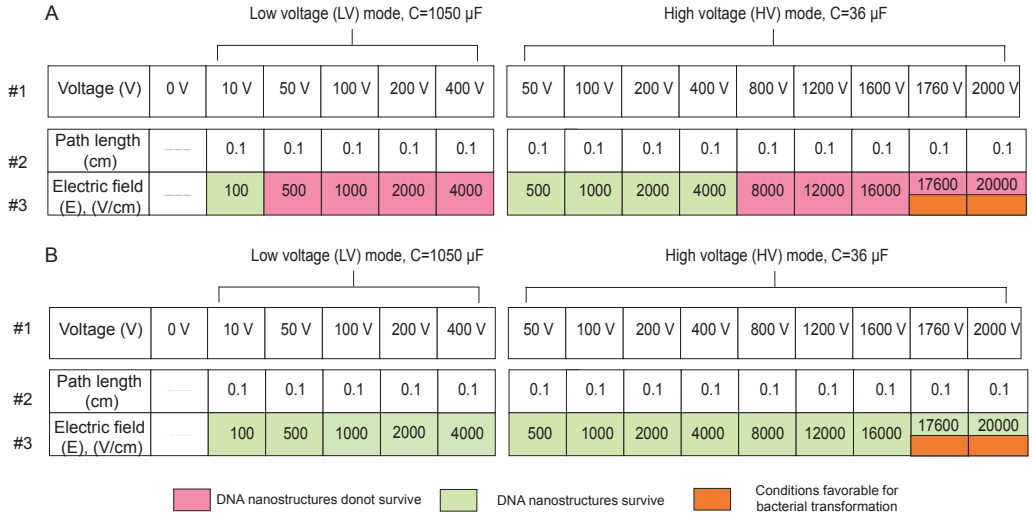


Figure 57: Effect of different electric field strengths on the stability of Mg/Spd-nanostructures in high resistance sample buffers (10% v/v glycerol in water) and using smaller path length cuvettes (0.1 cm). (A) Mg-nanostructures. (B) Spd-nanostructures. Voltages, path length and the corresponding electric field strengths analogous to the modes are specified in rows #1, #2 and #3 respectively. The rows corresponding to electric field strength (#3) are color coded with respect to survival of nanostructures (pink/green) and conditions favorable for bacterial transformation (orange).

.

Labeling and purification methods

Labeling was carried out by replacing the regular staples with the corresponding longer DNA staple strands, that were extended at the 3′ end with a 22 nt long sequence complementary to the 5′ Atto647N-labeled DNA strand (figure 58 (A)). The extension sequence is given in 8. Spd-nanostructures labeled with different number of fluorophores were analysed by agarose gel as seen in figure 58 (B).

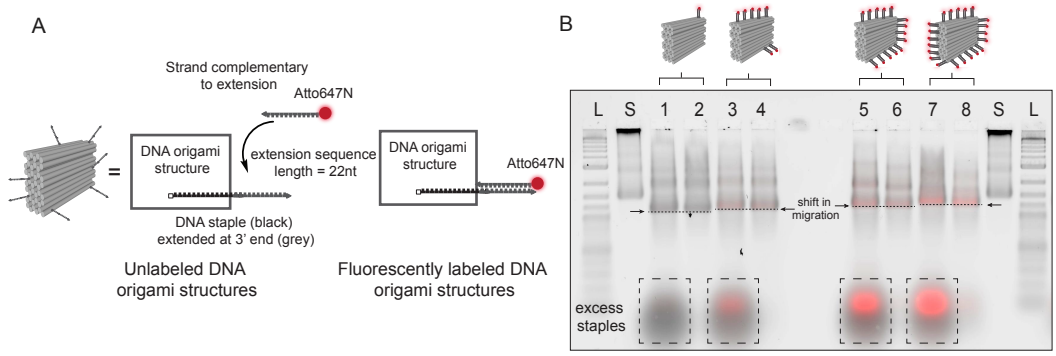


Figure 58: Fluorescent labeling of DNA nanostructures via hybridization. (A) Schematic representation. (B) Agarose gel showing 42HB Spd-nanostructures labeled with different number of fluorophores– unpurified and purified samples (by SEC). L: 2-Log Ladder, S: Scaffold, 1,2: unpurified and purified 42HB respectively with 1 fluorescent dye molecule, 3,4: unpurified and purified 42HB respectively with 7 fluorescent dye molecules, 5,6: unpurified and purified 42HB respectively with 14 Atto647N and 7,8: unpurified and purified 42HB respectively with 21 Atto647N dye molecules

.

Different purification protocols used for Spd-nanostructures. Figure 59 shows the different methods used with their protocol outline and representative agarose gel examples showing the degree of purification and shortcomings of each process.

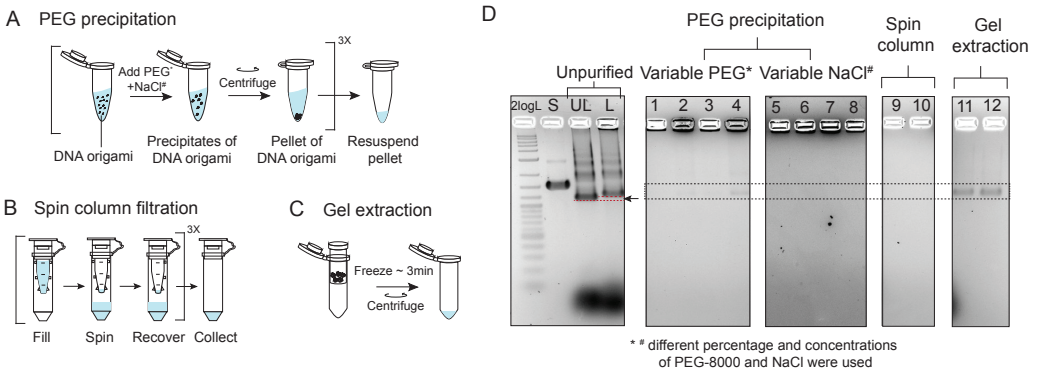


Figure 59: Different purification strategies used for Spd-nanostructures purification. (A-C) Schematic representation of different methods used for purification of Spd-nanostructures– PEG precipitation, spin column filtration and gel extraction. (D) Agarose gel showing the purity of Spd^{3+}-folded DNA origami purified by different methods. 2logL: 2-Log Ladder, S: scaffold, UL: unlabeled origami, L: labeled origami, 1-4: purification of labeled structures by PEG precipitation at constant NaCl concentration i.e. 500 mM [85] and variable PEG-8000 i.e. 1%, 4%, 10% and 15% respectively, 5-8: purification of labeled structures by PEG precipitation at constant PEG concentration i.e. 15% and variable NaCl i.e. 100 mM, 200 mM, 400 mM and 500 mM respectively, 9-10: purification using spin columns, no purified product was recovered, 11-12: gel extraction yielded low concentration of purified products

.

Electrotransfection of Spd-nanostructures in a single Jurkat cell

Electrotransfection of Spd-nanostructures in Jurkat cells. Internalized structures are encircled with white circles (figure 60). Figure 61 (B) shows internalized Spd-nanostructures in a single Jurkat cell at different time points and volumes. A single cell was monitored for 4 hours and confocal images (overlay of BF and red channel; Atto647N-labeled Spd-nanostructures) were taken every 0.5 hour for almost the entire volume of a cell ($\sim = 10\,\mu m$) with each step equal to $0.5\,\mu m$ (Total no: of steps = 20). It was observed that Spd-nanostructures were present at different z positions for different time points. This could either be that there are different structures present which get into focus at different time points as the adhesion of cells to the chamber might change (as seen in t = 0.5 and 2 at z = 15) or there is movement of the structures inside the cell. Dashed red boxes show the presence of origami at different z volumes and time points.

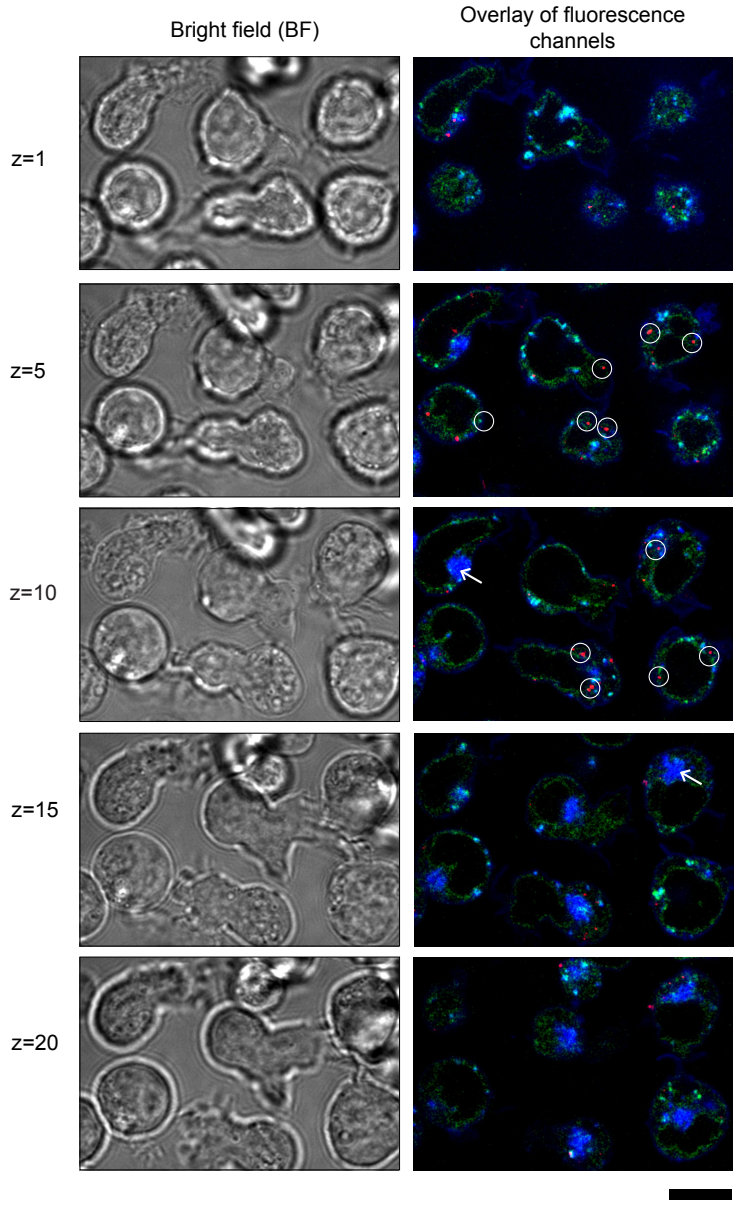


Figure 60: Electrotransfection of Spd-nanostructures in Jurkat cells. Confocal micrographs of overlay of fluorescence channels (green, blue and red) and bright field (BF) were taken at different z sections (each z = 0.5 µm). The cell organelles lysosome (green) and cell membrane (blue) were stained to mark the boundaries of the cell and pinpoint location of the electroporated species (red, Spd-nanostructures). Blue blobs (shown by white arrows) correspond to tip the cell membrane which appears at different z positions for different cells due to variable cell size ($\sim$ 11-15 µm) of the cells and differential attachment of cells to the chamber. Scale bar: 10 µm.

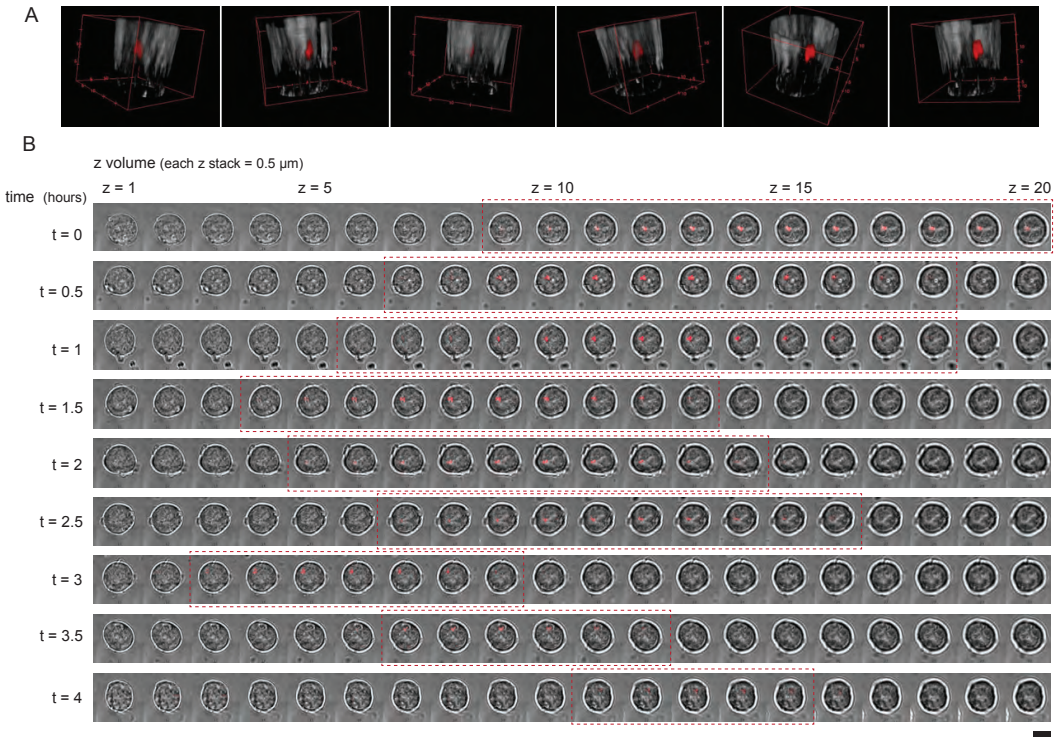


Figure 61: Spd-nanostructures at different time points and z volumes in a single Jurkat cell. (A) 3D reconstruction of a single cell containing Spd-nanostructures using Fiji [140] plugin 3D viewer. (B) Overlay of BF and red channel (corresponding to Atto647N-labeled Spd-nanostructures) confocal images at different time points and z steps. Scale bar: 5 µm.

Lipofectamine mediated transfection

Lipofectamine mediated transfection of Spd-nanostructures in Jurkat cells. Complex formation between Spd-nanostructures and lipofectamine were analysed by TEM as shown in figure 62 (A). Further, confocal micrographs of different z volumes of jurkat cells transfected with Spd^{3+}-origami structures with lipofectamine are provided (Figure 62 (B)). Internalization of Spd-nanostructures was observed (white circles).

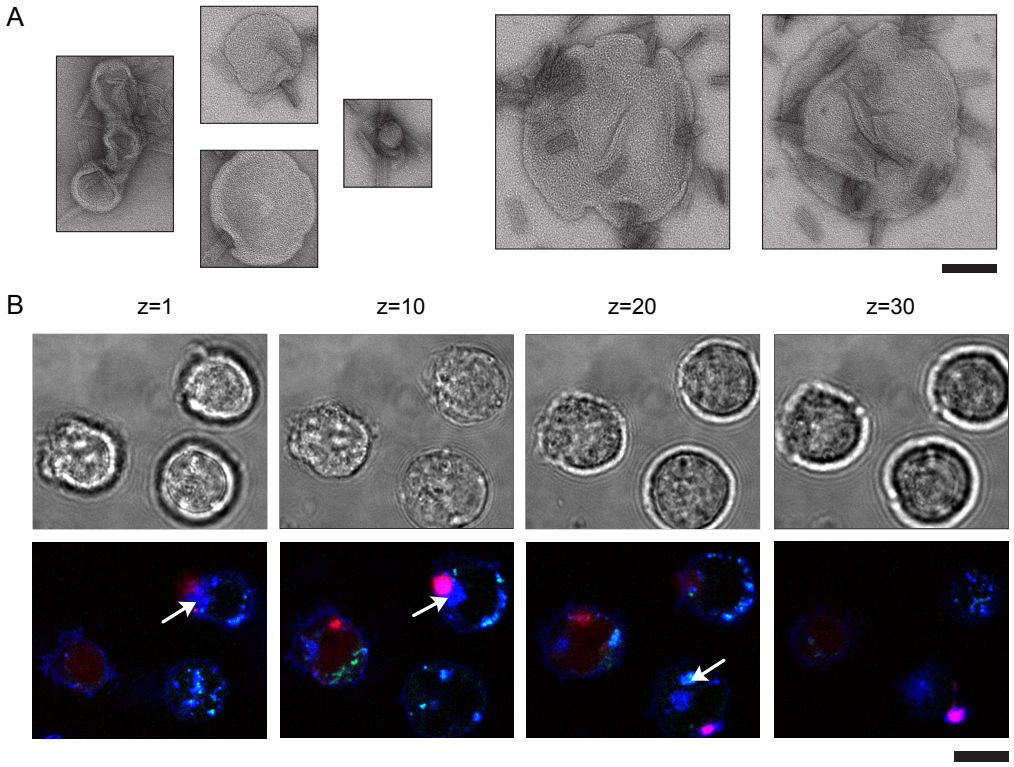


Figure 62: Lipofectamine-medated transfection of Spd-nanostructures in Jurkat cells. (A) TEM micrographs showing complexes formed between lipofectamine and Spd-nanostructures. Scale bar: 100 nm (B) Confocal micrographs of overlay of fluorescence channels (green, blue and red) and bright field (BF) were taken at different z sections (each z = 0.5 µm). The cell organelles lysosome (green) and cell membrane (blue) were stained to mark the boundaries of the cell and pinpoint location of the electroporated species (red, Spd-nanostructures). Blue blobs (shown by white arrows) correspond to tip the cell membrane which appears at different z positions for different cells due to variable cell size ($\sim$ 11-15 µm) of the cells and differential attachment of cells to the chamber. Scale bar: 10 µm.

Towards 3D RNA nanostructures: design and synthesis

PCR product after amplification using DNA gBlocks for RNA tiles were checked for the correctness of length and purity of the products onto an 2% agarose gel. Figure 63 shows PCR products corresponding to DNA templates for different modifications of RNA tile, amplified using HF or GC Phusion PCR master mix.

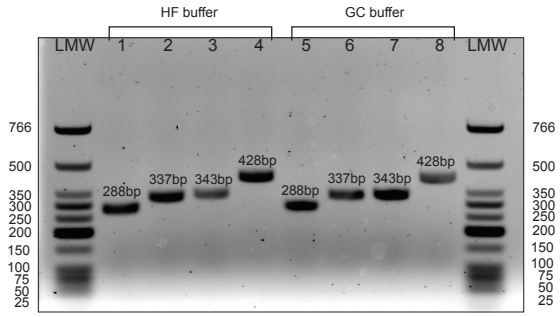


Figure 63: Verification of PCR products corresponding to different RNA tiles. LMW: Low Molecular Weight Ladder, 1-4: amplified DNA template (in HF Phusion PCR master mix) for 3H-AE, 3H-AE-MGapt V1, 3H-AE-MGapt V2 and 3H-AE-MG-spinach-apt RNA tiles respectively, 5-8: amplified DNA template (in GC Phusion PCR master mix) for 3H-AE, 3H-AE-MGapt V1, 3H-AE-MGapt V2 and 3H-AE-MG-spinach-apt RNA tiles respectively.

Rolling circle transcription (RCT)

Effect of EDTA on transcriptional aggregates of PP7-MGA RCT RNA.

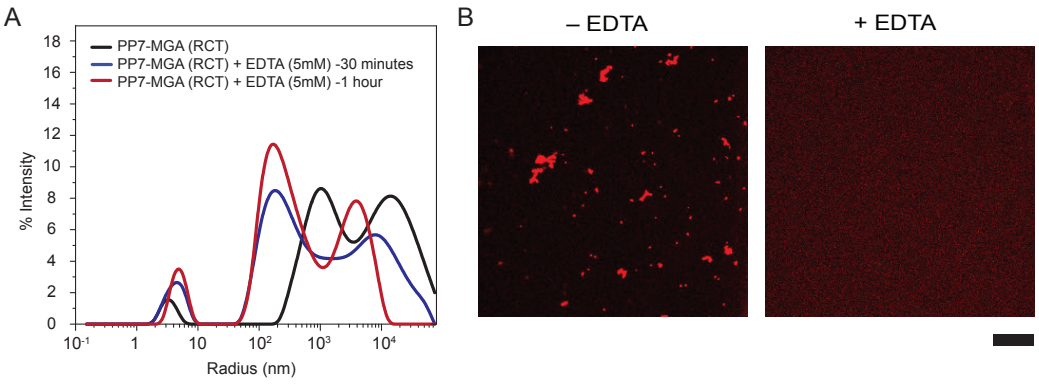


Figure 64: Effect of EDTA on transcriptional aggregates of RCT products from PP7-MGA. (A) DLS. (B) Confocal microscopy images. Scale bar: 40 µm.

In vivo synthesis of catalytic DNA molecules

Steps involved in synthesis and extraction of 10-23 DNAzyme by MOSIC method (figure 65) and denaturing gel verifying the electroeluted MOSIC-DNAzyme (figure 66).

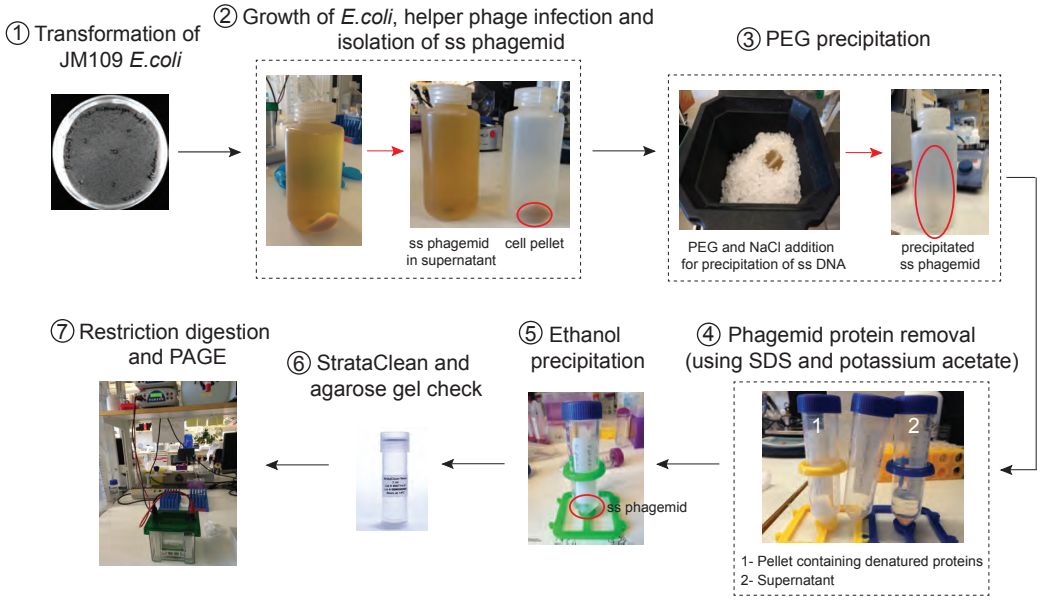


Figure 65: Steps involved in synthesis of 10-23 DNAzyme by MOSIC method.

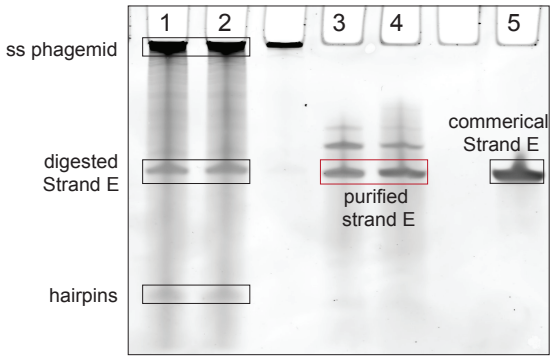


Figure 66: Verification of strand E (or 10-23 DNAzyme) after electroelution. 1-2: digested ss phagemid, 3-4: purified 10-23 DNAzyme by electroelution, 5: commercially procured strand E.

DNA and RNA sequences

This section includes the staple lists generated by cadnano for the DNA origami nanostructures used in chapter 3 (42HB, tube and tCRRO) and 4 (42HB only) respectively. In addition, it includes the DNA sequence which is labeled with Atto647N dye that hybridizes to complementary regions extended from some staples of 42HB (shown in red in the staple list in figure 67-71). Also, it includes the sequences of the two M13mp18 phage scaffolds used in chapter 3 and 4 (p7249 and p7560). Furthermore, all the primer sequences, DNA gBlocks, RNA tile sequences corresponding to different versions and RNA aptamer sequences used in chapter 5 are provided.

DNA sequences, staple lists and scaffold sequence corresponding to DNA origami structures

The staples which are extended at $3'$ end for hybridization with dye labeled complementary DNA strand are depicted in red and others in grey in the cadnano diagram and staple list. The extension sequence (Ext) is $5'$-TTATACATCTATTATACATCTA-$3'$ (22 nt) and the dye is labeled at $5'$ end of sequence $5'$-TAGATGTATAATAGATGTATAA-$3'$ (22 nt).

Start	End	Sequence (5'– 3')	Length	Color
0[111]	5[118]	CAGACGTAGGTGAACGGTCGCCCGTCGATAAACAGTACCAAGCGCGAAA	49	
0[167]	1[160]	CATTCCACAGACAGCCCTCATCAAGCAAGCCCAAT	35	
0[203]	0[175]	TTTTTTTTTTACCAGTACAAACTACAACG	29	
0[34]	0[6]	AGGAACAACTAAAGGAATTTTTTTTTTTT	29	
0[48]	9[55]	GAGTGAGAATCTCCAAGACAGTTAGACTCCTCAAGTCATAATGACGAGA	49	
0[69]	5[76]	CAACTTTAAAAGGAGATCGTCTTAGGATTGAAACAAAAGAATACACTAA	49	
0[90]	5[97]	TCTGTATCGGTTTAAGTTAAATCAGTACCTATTTCATCTTTGACCCCCA	49	
1[112]	12[105]	TTTCTTATTAATAAGACAAAACAGTATGTTAGCAA	35	
1[133]	13[132]	GATAGTTCGTTGGGCGATTGATGATTAAGACTCCTGCTTTAA	42	
1[49]	10[63]	AAAAAAAGGAATACTCCACGGAATAAGTTTATTTTTCAGTTG	42	
1[6]	2[6]	TTTTTTTTTTGCGAATAATAATTAGCAACGGCTATTTTTTTTTT	44	
1[70]	12[63]	GCCTTTAAGATTCAGTCACAAAACATATAAAAGAA	35	
1[91]	12[84]	TCAGCTTACAACATATATGGTAAATACATACATAA	35	
10[160]	13[160]	GAGCCAGATATTGAGAATACCCAAAAGAGACGAAGCCCTTTT	42	
10[181]	12[175]	ACTTGAGTAAAGGTCGAGGAA	21	
10[200]	1[203]	TTTTTTTTTCCGTCACCGCTGAGTTTCGTCTTTTTTTTT	41	
10[48]	2[35]	CACATTCTTGCCCTCATGACAAGAACCGAGACTTTCATCGGA	42	
11[2]	10[21]	TTTTTTTTTTATTACGAGGCATAGTACATAACG	33	
11[28]	12[6]	AGAGCAATAAAAACCAAAATAGCGAGTTTTTTTTT	36	
12[167]	8[168]	AATAACGCGGAAATTATTCATCCATTTGCATTACCCATCGAT	42	
12[203]	11[200]	TTTTTTTTTTCCAGAAGGAAACGAATTATCATTTTTTTTT	41	
12[48]	26[49]	TTACCAGGAGGGGGAAATATCGTCAAAAATGAAAACCTAATTCTGTAATG GATTGA	56	
13[105]	18[112]	TTCATTGCTATTATCGGGAGACCAGTTAAAGTACGGTGTCTGTAGTTTG	49	
13[112]	25[111]	AATCCCCAAATCACATCATAACGCAAGACAAAGAATGTTAAA	42	
13[6]	14[6]	TTTTTTTTTTAGGCTTTGCAAGCTTCAAAGCGATTTTTTTTTT	44	
13[63]	18[70]	TTAGACTAGAGGAACTTTACAATAAGAAGAATATAATGCTGTGTTTAGC	49	
13[70]	25[69]	GGATAGCTTAATGCACCGTGTAGTTAATTTCATCTATTGTAA	42	
13[84]	18[91]	ACTGCGGCGGATTGAAAACAGATATTATCATGTTTTAAATATTTCGCAA	49	
13[91]	25[90]	AATCGTCTCAACCGTAAATAAAAACTTTTTCAAATTTAAAAT	42	
14[132]	17[146]	AAAATCACTGAACAGTCTTTCCATTCCATATAACAGTTGATT	42	
14[203]	9[200]	TTTTTTTTTTCAATAATAAGAGAGGCCGGAATTTTTTTTT	41	
14[41]	0[35]	GCGTTTTATAAGGCAACTAATCGTTGAAAATAGAA	35	
15[133]	35[139]	AAGTCAGCAATAAATGAGTAAGTCTGGCAGGAACGCGAACTCGAACCAC	49	
15[147]	19[160]	TTGAGCGATCTTACTTAGCGAACCTCCCCTAAGAATTCATCG	42	
15[6]	16[6]	TTTTTTTTTTACCAGACCGGAACTTTAATTGCTCTTTTTTTTT	44	
16[146]	2[133]	CAACGCTCCACCAGGACAGAAAGGTGTAGACAACCAACCATCG	42	
16[160]	26[147]	ATCCTGACTAATATGTAATTCATGTAATATAGCTTAGATTAT	42	
16[167]	2[147]	CAATTTTCCGCCGCAGCAGCAAGGTTTACACCCTCATTTTCAGGGATAT	49	
16[203]	7[200]	TTTTTTTTTTATTAGTTGCTATGGTTGAGGCTTTTTTTTTT	41	
16[34]	21[34]	AGAGTACGCAAACTATTAAGCTGTACCA	28	
16[41]	4[35]	GGATTAGCAAGAGTAATCTAACTTTGAAAGAGGACAAGGCACTGAGGAA	49	
17[119]	2[112]	GAAGTTTCAGAGCCGAGCCACAGCGTCAGATATAAATAACCGATATATT	49	
17[175]	33[181]	TTTGAAGATAGAAGGCGCCCAAGATTTTGTAAATCGGTCAGTAATATCA	49	
17[56]	2[49]	AATTGCTACGATTTGAACCAGCATCTTTAGAAGGAACCCTCAGCAGCGA	49	
17[6]	19[34]	TTTTTTTTTTCTTTTGATAAGAGGTCATTCATCAATAATCATA	43	
17[77]	2[70]	AGCTCAATTATCCCCCGCCTCAGCCCCGTTTTGCGGCCGCTTTTGCGG	49	
17[98]	2[91]	GCAACTACAAAATACCCTCAGTTCATCGATAAGTGTGAGGCTTGCAGGG	49	
18[200]	17[203]	TTTTTTTTTTAAATCAGATCCTTAAATCAAGTTTTTTTTTT	41	
19[112]	33[118]	CCAAGTAATAACGGAATTACCCACTAACACGAATA	35	
19[147]	28[133]	TTTTATTCGCGAGGCGTTCCCAATTCTGAACAAGCCCTTTTA	42	
19[2]	20[2]	TTTTTTTTTTGCATTAACACTCAGAGCTTTTTTTTT	38	
19[49]	30[49]	AGCATGTATTTCAATTACCTGCGGCACCGCTTCTGTGCCAAG	42	
19[70]	33[76]	AATCGGCTCGCGCACATCAAGTTGAGGAGACAGTG	35	
19[91]	33[97]	ATTCCAATTTGAATTACATTTAGCCGTCGTAAGCA	35	
2[203]	1[181]	TTTTTTTTTTCTCAGAACCGCCACCCTCACGTAACA	36	
2[34]	10[28]	ACGAGGGTTTTTCAGCAGATA	21	
20[167]	1[167]	AGGTAAACAGAGAGATAGCAAGTAGCACGGAATTAAGGAACC	42	
20[200]	15[203]	TTTTTTTTTTAGAGAATATTTGAGTTAAGCCTTTTTTTTTT	41	
21[168]	17[167]	GAGGCATCGACAAACATTACCGCTTATCCGGTATTGACTTGC	42	
21[2]	22[2]	TTTTTTTTTTATAAAGCTAGCCTGAGAGTTTTTTTTTT	38	
21[21]	18[2]	AATCGGTAATAAAGTCCAATATCTACTAATAGTAGTATTTTTTTTTT	47	
22[104]	10[105]	ATGATATATAAATAACGTAGATTACCAGCGCCAAAAACGAAC	42	
22[111]	18[98]	CATCAATAACCCTCAACGCGCTATTAAAACCATTAGATACAT	42	
22[125]	10[126]	GACAGTCCTCAAATTATTACGGGGCGACATTCAACAAGAAAA	42	
22[132]	18[119]	GGCCGGACAATGCCCAACATGCATCGAGCGAACGAGTAGATT	42	
22[160]	25[160]	ACAACGCATACAAATTGGGTTATATAACAAAAGACGCTGAGA	42	
22[181]	24[175]	AATTGAGAAGCCAACTTTTTA	21	
22[200]	13[203]	TTTTTTTTTTGTAGGGCTTCGAACAAAGTTATTTTTTTTTT	41	
22[41]	11[48]	TACAAAGTTTGCCAACGACGACACTATCATAACCC	35	
22[48]	15[41]	AGAGATCTATGACCTAAGAATTAGCAAACCAACAG	35	
22[83]	10[84]	TGATAAAGTCCAATAGGTGGCTCAATAGAAAATTCTATTACA	42	
22[90]	18[77]	TTCTAGCCGCAAGGTAGATAACCTTATCATGGTCAATAACCT	42	
23[2]	22[21]	TTTTTTTTTTTCTGGAGCAAACAAGAGGTCATT	33	
23[28]	24[6]	GAATCGAATATGTACCCCGGTTGATATTTTTTTTTT	36	
23[63]	18[56]	AATACCGCGGAGAGCGGGAGAAAATAATAATCAATTATATTTTCATTTG	49	
24[146]	13[146]	TATATGTGTATCATATGCGTTCAGATTCAAAAGGGAGAAAAC	42	
24[203]	23[200]	TTTTTTTTTTCTGAGAGACTACCGCTCAACATTTTTTTTTT	41	
24[62]	26[42]	TCTGACCTAAATAACTAGCATATTGTATCCGTAAT	35	
25[6]	26[6]	TTTTTTTTTTATCAGAAAAGCCTGGTGTAGATGGTTTTTTTTT	44	
26[111]	30[112]	CAACATTTCATTTGATTCGCCTTGGGAAGGGCGATAAGGCGA	42	
26[118]	37[111]	CTTTCATTCAGCTCGTGTAATTCGCCATGTAATAAAAGGGACGAATC GG	49	
26[181]	25[203]	AGAATCCATTTATCAAAATCATAGGTTTTTTTTTT	36	
26[203]	21[200]	TTTTTTTTTTAATTTTCCCTTCCAGTAATATTTTTTTTT	41	
26[34]	34[28]	GTCACGTCCAAAAACGTAATC	21	
26[69]	30[70]	TGGGAACAAACAAAGAGGCGAAAACCAGGCAAAGCACGTTGT	42	
26[76]	37[69]	TTCTCCGACGTTAACCTCGATTGGCTATTGACCTGAAAGCGTCCGCT TT	49	
26[90]	30[91]	AACAACCAATTAATACCAAGTGCCATTCAGGCTGCGCCAGGG	42	
26[97]	37[90]	AGCGAGTTCGCATTTAAGCTACGAACTGCCAACAGAGATAGATGTCG TG	49	
27[133]	37[146]	CAGTACATATCTAACGGCTGATTTTCTTTTTGCGTATTGGGC	42	
27[147]	30[154]	ATATATGATACAGTGGTTAGAACCTACCACTTCTG	35	
27[6]	28[6]	TTTTTTTTTTGCGCATCGTAACGTATCGGCCTCATTTTTTTTTT	44	
28[160]	38[147]	GATGAATTGAGTGATTGAAAGGCATCACCTGGTAATATCCCA	42	
28[203]	19[200]	TTTTTTTTTTAAGAAATTGCGTATAGCAAGCTTTTTTTTTT	41	
28[34]	37[34]	GACGACACGTGCATGTCAACCTCTAATCATCCTGTGAGCTAA	42	
29[119]	14[119]	CGGTGCGAGAAACACCGCACTTTCAGCTGAACACCGGTCTTT	42	
29[133]	20[140]	TCGCTATTACGCAGAACAGTAAAGCCGTACGACGA	35	
29[175]	31[181]	TGCACGTAATCCTGTCAGATG	21	
29[56]	15[62]	GTGCCGGATTATTCAGAAACCATCCCATTAGCAGC	35	
29[6]	30[28]	TTTTTTTTTTGGAAGATCGCACTCCAGCCGCCAGGG	36	
29[77]	14[77]	GCCATTCTACAAAATGTCTTTGTCCTGATAACATACATCAAA	42	
29[98]	14[98]	GCAACTGTGATTGCGAACGGGCTGTTTACATTAGAAGTCAGA	42	
3[147]	7[160]	TCACCGTTAAGTTTAAGCCAGAATGGAAACAAATAACCAGAG	42	
3[6]	4[6]	TTTTTTTTTTCAGAGGCTTTGATTAAACGGGTAATTTTTTTTTT	44	
30[200]	29[203]	TTTTTTTTTTCATCAATATAAAACAGAAATATTTTTTTTTT	41	
30[41]	17[41]	AGGAGAAAGCTTTCAGGGGACCAGGCAAGGCAACGAGCTGAAAAG GTGGTTTGCGG	56	
31[112]	33[125]	TTATCATGCGCGTAGAGGCCACCGAGTAGTGAGACTAGGGGC	42	
31[154]	33[167]	TTATCATGCGTAACCGTTGTAGCAATACGGCCTTGCTTGCTG	42	
31[161]	13[167]	CATATTCTCAACAGATAACCTCATAGCGTTAGGCAATATTTATAAGAA A	49	
31[168]	19[167]	CTGATTAATTGTTTGGATTATATATCAAAACGTCATAGGAAT	42	
31[2]	32[2]	TTTTTTTTTTGGTTGTGAAATGTCCGTTTTTTTTTT	38	
31[28]	39[34]	CGCACGAGCCCGAGAAATCCC	21	
31[70]	33[83]	TCCTTTGTCCTCGTGAACGGTACGCCAGGAGAGAGCGGCCCT	42	
31[91]	33[104]	ATTTTAAACGAGCAAGAAGTGTTTTATTGCCCTTACTCGTC	42	
32[146]	20[147]	GGAAGGTTAAATCAAATTCGCTGTGTAGGTAAAACTGTCCAG	42	
32[153]	31[139]	GAATTGAAGCGGAAAATAATGGACAGCTGGCGAAAAAACCAC	42	
32[200]	27[204]	TTTTTTTTTTCAATATCTGTCGCTATTAATTTTTTTTTTT	42	
33[2]	34[2]	TTTTTTTTTTCCAAAATAATAGCTGTTTTTTTTTTTT	38	
33[21]	30[2]	CCCCGCTTTATGACATTCATGTGGATGTTCTTCTAAGTTTTTTTTTT	47	
33[42]	21[41]	GCTCGCCAATAGTGCTGAATTCTGCCAGGGGATAGAAAACAT	42	
34[104]	23[111]	CAGGGCTAAATTTTCGCGAGAGAATAAACACCGGA	35	
34[125]	23[132]	CTGGTTGATTTTTTATCCAATTTACTAGAAAAAGC	35	
34[167]	22[168]	GCAAATGAGAGTCAGCTTAGGTTCTTACCAGTATAAATCGCC	42	
34[200]	36[175]	TTTTTTTTTTCAGTGCCACGCTGAGATTAACACTTTTGAC	40	
34[83]	23[90]	CTTGTTATATTTTGATATTTTGATAAATAAGGCGT	35	
35[105]	30[98]	TAAAACAGAGTAAAGGTGGGCAACTAATAGTAACATTAAGTTGGGTA AC	49	
35[126]	30[119]	ACCGAACTGACCTCCTTGAATAATATCTAACAAAGGGGGGATGTGCT GC	49	
35[2]	34[21]	TTTTTTTTTTCCTGTGTGAAATTGTTATGGTCA	33	
35[28]	36[6]	ATCCGCTTAAAGTGTAAAGCCTGGGGTTTTTTTTTT	36	
35[63]	30[56]	TTTTGAAAAAGACGTGATACCTTTAGAATATTAAAAAAACGACGGCCA G	49	
35[84]	30[77]	TAATGCGCGTGGTGGCCATCTAATAGATGTTATTATTTTCCCAGTCAC G	49	
36[146]	25[146]	ATTGGCACAGCAGAAGATAAATAAAATCATTTCTCCCATCAA	42	
36[203]	35[200]	TTTTTTTTTTGAAATACCTACACGCCTGCAATTTTTTTTTT	41	

Figure 67: 42HB staple list.

37[161]	32[168]	CCGCCAGCGTCTGATGAGGCGGTCAGTAGCCAGCAAACCTCATGGCAAA	49	
37[35]	34[42]	CTCACATGGAAGCACACAATTCCACACAACCGAGC	35	

37[56]	39[55]	TCACTGCAAGAATACGTGGATACGAGCCTAATTGCCCCAGCATCAGGCCGATTAAA	56	
37[6]	38[6]	TTTTTTTTTTTGCCTAATGAGTTTGATGGTGGTTTTTTTTTTT	44	
38[111]	41[125]	AGCTGATAATCAGTCTATGGTCCCTAAAGGGAGCCCCCGATT	42	
38[181]	37[203]	GAACTCAAACAGGAAAAACGCTCATGTTTTTTTTT	36	
38[203]	33[200]	TTTTTTTTTTGCCTGAGTAGAAAACCCTCAATTTTTTTTT	41	
38[69]	41[83]	AAGCGGTTAGACAGTAGAATCCAACGTCAAAGGGCGAAAAAC	42	
38[90]	41[104]	TGGCCCTAATCCTGCGTATAACGTCTATCAACTAAATCGGAA	42	
39[147]	40[161]	AATTAACCACCACAGGCGAACGTGGCGAGAAAGGAACGCTGC	42	
39[35]	23[48]	TTATAAAGGCGAAATATTTACTCGAATTCAGGAAGGTCAATCTGAACGGTAATCGT	56	
39[56]	25[62]	GGGATTTCCACGCTGACTCTAGAGGATCTATTTAA	35	
39[6]	40[6]	TTTTTTTTTTCCGAAATCGGCAATAGGGTTGAGTTTTTTTTTT	44	
4[104]	14[91]	TTAATGCCCCCTGCCAGGCGGGCATTTTGTTTAATAGCAAAG	42	
4[125]	14[112]	TAACAGTGCCCGTAGAGGGTTGACTGTATTGTGAAACCCTGA	42	
4[160]	14[147]	CTGGTAAACTCAGGCCGTAATATCACCATAGCTATCTTACGA	42	
4[203]	3[181]	TTTTTTTTTTGTCATACATGGCTTTTGATCACCCTC	36	
4[34]	22[28]	GTTTCCAGGACTAAGATATTCTCAGTGAAATTCGAAAGAAGTGCTATCA	49	
4[55]	6[49]	ATTAAGAGGCTGCACAACCTAAAACGAATGACCAA	35	
4[83]	14[70]	GGAACCTATTATTCTAGCGGGTTATTAGTAGTAAAAAGATTA	42	
40[104]	21[104]	TGCTTTGAAGTTTGAGATTAGAACAATTAAATGTGTTTTTAG	42	
40[125]	21[125]	CTACAGGTTTGCGGTTAGGAGTTTTTTATAGCCAGTTAAATG	42	
40[139]	33[153]	CGCTTAACTGTCCATCACGCAGGGTGGTCGCATTTCACATAA	42	
40[181]	39[203]	CAAGTGTATTAGTAATAACATCACTTTTTTTTTTT	36	
40[203]	31[200]	TTTTTTTTTTCTAGGGCGCTGGATGGCAATTTTTTTTTTT	41	
40[83]	21[83]	CGTGCTTCCCGAACAATACATAAAACAACGTCGGAATTTCAA	42	
41[133]	31[146]	TGACGGGGAAAGCCCCCGCCGCAGAAGG	28	
41[175]	41[203]	AAAGCGAAAGGAGCGGGCGTTTTTTTTT	29	
41[35]	30[42]	AGAGTCCAAGAATACTTAAGTGTCCTACCTTTCTC	35	
41[49]	21[62]	AAGAACGTGGACTCAGAGCGGCAACTCGGTATTAGGATGATGAAACGGCACTTTTG	56	
41[6]	41[34]	TTTTTTTTTTGTTGTTCCAGTTTGGAACA	29	
5[133]	3[139]	GATTTGTATCATTATAACGGGGTCAGTGCCGGAAT	35	
5[175]	21[181]	ATTTACCGATTGGCGACAGGATTTGCACACAAGAAAAAGTACTTTCGAG	49	
5[6]	7[34]	TTTTTTTTTTAATACGTAATGCCACTACGAGATGAACCTTCAT	43	
6[118]	21[118]	CGACCTGACCCTCATAATTTGATTAACTAATGCAGATATATT	42	
6[146]	16[133]	ATCGCCTGATAAATAGAGCCGAACGAGC	28	
6[200]	5[203]	TTTTTTTTTTAGGTCAGACGTTCCAGTAAGCTTTTTTTTTT	41	
6[76]	21[76]	TCAATCAACCGGAAAATCCAAGAGAGAAACAAGAAAGCCTTT	42	
6[97]	21[97]	GCCGGAAGCCGCCAAACAGCCGGAAGCGTCAACAAATAAAAA	42	
7[168]	4[161]	CAGCATTCTTGATATTCACAAAGCGCAGGAGTGTA	35	
7[2]	8[2]	TTTTTTTTTTCGCATAGGCAAATCAACGTTTTTTTTTT	38	
7[49]	18[49]	ATCACCGTTTGTTTAACGTCAATGGCTTAGAGCTTGGGCGCG	42	
8[139]	23[139]	TCAAGTTTTTAAGATAAATCAACAGTTCTGAGAAACTGTTTA	42	
8[181]	13[181]	ATGAAACATTAGCACAAGAAAAGATAGC	28	
8[200]	3[203]	TTTTTTTTTTACGTCACCAAGAACCGCCACCTTTTTTTTT	41	
9[105]	5[104]	TTAATCAGCGCGTTAACCGCCCTCCATGTTACTTAGCGATTA	42	
9[119]	0[112]	TTACCTTATCTACGAACAGCTTTTTGTCGTCTTTC	35	
9[126]	5[125]	ATGCGATTGCCTTTCACCCTCTGTGTCGAAATCCGCAAAGTA	42	
9[2]	10[2]	TTTTTTTTTTTAACAAAGCCCAAAAGGATTTTTTTTTT	38	
9[21]	6[2]	TGCTCATATTACCCTGGCTGACGGTGTACAGACCAGGTTTTTTTTTT	47	
9[77]	0[70]	TTGGGCTGGTAGAAATTGTATGGGATTTTGCTAAA	35	
9[84]	5[83]	TGAGATGCGGTCATCCTCAGACGAGGCGCAGACGGAACACTC	42	
0[139]	5[132]	GTAACGATCTAAAGTGATACCCCCACGCGTATAGCCCTTGAGCAACGGA–Ext	49	
0[174]	5[174]	CCTGTAGCATGTACGAGCCACGTACCGCGATACAGTCTCTGA–Ext	42	
12[146]	0[140]	ACTGGCAGGGAGGGAAGGTAACATACCAGTCAGGAGCGCCGAAGTTAGC–Ext	49	
12[174]	17[174]	ACGCAATAGTAAGCCAATGAAATAACCCCCAGCTAGGGAGGT–Ext	42	
14[55]	23[62]	AGACTTCTAATAGTATTTTTGAATTAATGGTTTGA–Ext	35	
14[62]	0[49]	GCCCGAAAACACCAAGATTTAAGGCTCCCAACAGTTTCAGCG–Ext	42	
15[140]	6[147]	AGGGTAAATGACCAACTGGCTCATTAAACAGTAGCAACCACCAATCCTC–Ext	49	
22[62]	12[49]	GGTAGCTAAAATGTACGCAAAGACACGT–Ext	28	
24[139]	29[132]	AAATGCTGATGCAAAACCAATCTTCCTGATGGAAACATCGGGGGCCTCT–Ext	49	
24[174]	29[174]	ACCTCCGATAGTGATTGAAAATGCTTCTCAGGTTTAATTATT–Ext	42	
25[49]	35[62]	AAGCAAACCCGGGTACCACAGACAATAT–Ext	28	
27[42]	37[55]	TTTGAGCAAAAGAAACTTTACCTGGAGTGGTTTGCGTTGCGC–Ext	42	

31[49]	41[48]	AATTCGAGAGCTAAACAGGAAACTATTA–Ext	28	
34[160]	37[160]	AAAAATCACAGAGGAATGGATTATTTACGCAGAACAATATTA–Ext	42	
36[139]	41[132]	GATTCACCAGTCACGAGGCGGTTCACCAAAAGAGTTGCGCCGTAGAGCT–Ext	49	
36[174]	41[174]	GCTCAATCCATTGCAACTATCTTCTTTGAGCGGTCAGGGAAG–Ext	42	
38[118]	35[125]	GGGCAACCCAACGCGCGGGGAACGACCATAAAAAT–Ext	35	

38[76]	35[83]	TTGCAGCCCAGTCGGGAAACCACCCTTCTAGTCTT–Ext	35	
38[97]	35[104]	CACCGCCCCAGCTGCATTAATATTCTGGATAGCCC–Ext	35	
9[63]	4[56]	GAACGAGCGTTTGCAGCCACCTAAGGGAACCGAACAGAGGCATGAAAGT–Ext	49	
9[98]	0[91]	TTCAACTTAACGGAGCTTTCGTAGTAAATGAATTT–Ext	35	

Figure 68: 42HB staple list contd.

Start	End	Sequence (5'– 3')	Length	Color
15[252]	12[252]	AGGATTAGGATTAGCGGGGTTTTTTTTTTTTTAGCGCGTTTTCATCGGCA	50	
2[41]	3[41]	TAATACATTTGAGGATTTATTTTTTTTTTTTCATGGAAATACCTACATTT	50	
6[34]	7[34]	GTAATTTAGGCATTTTTTTTTTTTTATGGAAACAGTA	36	
0[41]	1[41]	TTGCAACAGGAAAAACGCTTTTTTTTTTTTTTAATCGGAACCCT	43	
19[252]	16[252]	ATGAACGGTGTATTTTTTTTTTTTCTTAAACAGCTT	36	
29[252]	26[252]	TGCAGCAAGCGGTCCATTTTTTTTTTTTATTACGCCAGCTGGC	42	
10[34]	11[34]	TATGGTTTATTTTTTTTTTTTAAAATAAACAGC	33	
24[48]	25[48]	CGAGTAACAACCCGTCGGATTTTTTTTTTTTAGAACCCTCATATATTTTA	50	
25[245]	22[245]	ACGGTAATCGTAAAACTAGTTTTTTTTTTTTAATAACCTGTTTAGCTATA	50	
20[33]	21[41]	CGGATTTTTTTTTTTTTAAGAAAAATCTACGTT	32	
3[259]	0[259]	CAGCAGAAGTTTTTTTTTTTTCTAAACAGG	30	
14[34]	15[34]	CCACAGACATTTTTTTTTTTTATAAATCCT	30	
26[41]	27[41]	CCGGAAGCATAATTTTTTTTTTTTTTCTCCGTGGGA	36	
16[41]	17[34]	AAATACGTAATGCCACTTTTTTTTTTTTGCCCTCATA	37	
23[245]	20[245]	ATACATTTCGCAAATGGTCTTTTTTTTTTTTTATGTTTAGACTGGATAGCG	50	
8[34]	9[34]	TTTGCCAGTTACTTTTTTTTTTTTGAGGCATTTTCG	36	
21[252]	18[252]	GGTAATAGTAAATTTTTTTTTTTTCAGACCAGGCGC	36	
13[252]	10[252]	TGCCTTTAGCGTCAGACTGTTTTTTTTTTTATCTTACCGAAGCCCTTTT	50	
7[259]	4[259]	AAGACAAAGAACTTTTTTTTTTTTTAAAACAGA	33	
1[252]	28[252]	AGAATCAGAGCGGGAGTTTTTTTTTTTCGCTGGTTTGCCCCAG	43	
4[41]	5[41]	ATTTGAATTACCTTTTTTATTTTTTTTTTTGAAGTATTAGACTTTACAA	50	
9[259]	6[259]	GTACCGCACTCATTTTTTTTTTTTGCGAGAAAACTT	36	
17[252]	14[252]	CGAGGTGAATTTTTTTTTTTTTTTTTTTGCTCAGTACCAGGCG	43	
28[41]	29[41]	CGTAAAGCACTATTTTTTTTTTTTTAGTGTAAAGCCT	36	
18[41]	19[41]	CAGTCAGGACGTTGGGTTTTTTTTTTTTTACGAAGGCACCAACC	44	
11[259]	8[259]	ATAGCAATAGCTTTTTTTTTTTTTTCGAGAACAAGC	36	
22[48]	23[48]	CGCAAGGATAAAAATTTTTTTTTTTTTTTTTTTAGCAAACTCCAACAGGTCA	50	
5[259]	2[259]	ATTTGCACGTTTTTTTTTTTTATAAAACAG	30	
27[252]	24[245]	GTGCGGGCCTCTTCGCTTTTTTTTTTTTCATGTCAATCATATGTACC	47	
12[34]	13[34]	CACAAACAATTTTTTTTTTTTTCCAGCGCCA	30	
17[63]	20[56]	TCTTTCCTCATGAGATACACTTGCGATTAACATTAATCGCGTTTTAATT	49	
7[49]	10[42]	TGTGAGTCATATTTGAATATAGAGCGTCCAATCCATCACAATCAATAGA	49	
7[245]	0[252]	GCAAATCCAATCGCAATAAAGCAAAATTAGGTGAGGAACCACAGGCCGA	49	
27[224]	0[217]	GCTGCGCGCGATTAGCCCTTCGGTGGTTCGAGCACACGGTACGCCAGAA	49	
27[98]	0[91]	GTAACCGCATGGTCTTCCAGTAAAACCGGGAAGGGGAGTAGAAGAACTC	49	
17[210]	20[203]	CCAAAAGTCGCCCAAATCATACAAGAACAAAATAGTCATTGAATCCCCC	49	
17[126]	20[119]	TCAGCGGCGAAAGAGAGATTTAACGAGTAGATACATATAGTCAGAAGCA	49	
17[42]	20[34]	TAACGATACGGGTATAAAACGATTATACAATAAAACAAAGCGAACCAGAC	50	
17[133]	10[140]	AGTGAGAATAGAAACGCCACCGGTCAGTGAACCGCGAATTAGACGCAGT	49	
17[238]	10[245]	TATCAGCTTGCTTTGATAAGTCAAGAGATTTTCGGTCAAGTTTAAGAAA	49	
17[105]	20[98]	GCTAAACGGGTAGCAGCGCGAGAGATGGAATACCATTGCATCAAAAAGA	49	
17[217]	10[224]	GAGCCTTTAATTGTTGATATATATTAAGTTAGCGTCCGTAATCGAACAA	49	
7[224]	0[231]	TAACTATATGTAAATTCAGGTAAGGGTTCACCGCCTCGCCATGACAGGA	49	
27[168]	20[175]	CTTTCCGGCACCGCAACGTTAATTTTTGTCCAATACTGGAAGGAAAACG	49	
17[154]	10[161]	CTAAAGGAATTGCGCAGAACCTGCCCGTCCCTCAGCCAGTAGGCATGAT	49	
7[77]	0[84]	TCGTCGCTATTAATTGATGAATAATTTAAATATCTCACCAGAAACTAT	49	
7[175]	10[168]	ATAGGTCGCGTTAATCCCATCTATTCTATCAGAGGACCCAAAAGAACTG	49	
27[147]	20[154]	GGAAGATCGCACTCTTCGCATTTAATGCGGCAAAGTATGCAATCAAAAA	49	
7[112]	10[105]	GAAAACACGTTATATCAGCTAATCAAGAAGAATAAACATACATAAAGGT	49	
27[42]	20[49]	ACAAACGGCGGATTATGTGAGAATGCAAATTTCAAGGATTAGCGAGCTT	49	
27[56]	0[49]	GACCGTACAATTCCGTGAGCTAGTTTTTATTTAGAGAACAATATTACCG	49	
17[70]	10[77]	AGACGTTAGTAAATCGTAACATCCAGTACCAGCATAAATATTAAAGACA	49	
27[105]	20[112]	TGCATCTGCCAGTTTAGGAACAAATCACAAAGCTACTTAATTAAGCGGA	49	
27[77]	0[70]	GTTGGTGTGTGAAAGCGTTGCCGTGAACGCCGGCGCGGCCTTGCTGGTA	49	
27[203]	0[196]	AAAGCGCAGGGTTTTGAGACGATCCCTTCGCGTACAAGTGTTTTTATAA	49	
17[196]	10[203]	CCAAAAAAAAGGCTTAGGTGTTATTATTTAATCAAATGAAACAACCGAG	49	
7[35]	0[42]	CATAAATCAATATACAATTTCACAATTCCAATAGATGACGCTCCAGCCA	49	
7[238]	10[231]	TGCTGATTAGTTAACAAGAACTTCATCGATAAGAGAGTAAGCAGATAGC	49	
7[70]	10[63]	CTGTAAATAGGGCTGGTAAAGCTGAATCTTTTGTTCCACGGAATAAGTT	49	
7[154]	10[147]	AATAGTGCGGAATCTCCTGAATTTAGCGAACTGAATAAGACTCCTTATT	49	
7[161]	0[168]	AATTTATCAAAATCGGATTCGTATCAGAGCTGAACGGCACAGAGAGTCT	49	
17[231]	20[224]	ATCGGTTCGCCGACACCAACTCTTCATCAAAAGAACTGCGGAATCGTCA	49	
17[91]	10[98]	CTGTATGGGATTTTGCCCAATCTTTTGAGAACCACATTAAAGGCAACA	49	
7[203]	0[210]	TCCGGCTTAGGTTGATATACAGATTATAGCTGAGACGAACTGTCCTGAG	49	
17[35]	10[35]	GTTAGCGTAGCATTCATTAAATGATATTAAGACAAAAATTCA	42	
7[98]	0[105]	CCCTTAGAATCCTTAATTACCCATTATCGGAATTGATAAAAGCTTGCCT	49	

	20[238]	AACTGTTGGGAAGGCCGGTTGTCGATGATTTTCATCCATTAGTCCAATA	49	
7[196]	10[189]	TTTAACCACCGACCTGTAGAAAGATATAATATCAGGAAACGCAATAATA	49	
27[182]	0[175]	TTCTGGTTTGTAAAGGTGGTTTAGCCCGGCTTAATGGCCACCGAGTAAA	49	
17[168]	20[161]	AATAATAGGGAGTTACCTGCTGCTCATTAACACTAAGAATGACCATAAA	49	
27[63]	20[70]	ATGGGATAGGTCACGCCAGCTGTAGGTAACTTTTGTGCTCCTTTCAAAT	49	
27[245]	0[238]	GCGATCGGAAAGGGAGAGAGTCAGGCGACCTCGTTTTAAAGGGATTTTA	49	
7[133]	10[126]	ATTAAGAAAGCCTGTGTTTATCGGGAGGAGCGCATATGTTAGCAAACGT	49	
17[189]	20[182]	AAAATCTATTCGGTCGGAACGCCAAATCTACCAGACTTTAAACAGTTCA	49	
27[119]	0[112]	TGAGGGGTACCGAGTGCCAGCGACTCCAAGCGGGCTAGTAATAACATCA	49	
27[210]	20[217]	CATTCGCCATTCAGCCCAAAAAGTCTGGGAAAAGGGAACGAGTAAATAT	49	
27[126]	20[133]	ACGACGACAGTATCATCAGCTTCAACCGAGCAATATGTAGCTTGACTAT	49	
17[112]	10[119]	AACTTTCAACAGTTCCTCATTACTGGTACCACCCTTCACCGAAGAAAAT	49	
7[91]	10[84]	TAATTTTTATAAAGCGACGACGCACCCAAAAATAGTATAAAAGAAACGC	49	
7[56]	0[63]	GAATAACCTTGCTTAACAAAAAATCCTTCAACTAAGGATTATATATCCA	49	
17[49]	10[56]	CTAAAGTTTTGTCGCAGTACACGCAGTCGGCAGGTACCGATTTATTTTG	49	
27[84]	20[91]	TAGATGGGCGCATCTTCCGGTTGAGAAAAAAACATATTTTTGTTAAGAG	49	
17[147]	20[140]	GGAACAACGGGATCTAAATTGTGCCCTGATTACGATCAGGTCTTTACCC	49	
27[161]	0[154]	CAGCCAGCAAGCTTGGTTTGCGTTGTTCCGTAACCGTCCATCACGCAAA	49	
7[182]	0[189]	TGAGAGACTACCTTTACATCGCAATATAAAAAATCTGGCTATTCAGTGA	49	
17[84]	20[77]	GAATTTTCTTTGAGTGACCCCTAATCATATTCATCGAAGCCCGAAAGAC	49	
27[140]	0[133]	GGCCTCACGACTCTGGCCAACAGTCCACAAGTGTATTGTAGCAATACTT	49	
7[119]	0[126]	TAGCGATAGCTTAGCAGAGGCAGAAACCTTGGCAAACAGAGACTTTGAT	49	
7[217]	10[210]	GGTTATACTAAATTGGCTGTCGCGCCCAAGAATTGAGTTACCAGAAGGA	49	
27[189]	20[196]	GCCGGAAACCAGGCTAAGCAAGCTATCACTAATAGACAGTTGTCAAATG	49	
7[140]	0[147]	CGCTGAGAAGAGTCATACCAAAATTATCCCTCAATCCTGAAATTAACCG	49	
17[175]	10[182]	ATTTTTTCACGTTGGAGGTTTCCCCCTGGAGCCACGCAAGGCACGGAAT	49	
29[49]	26[42]	CTAATGAACACAACATACGAG	21	
9[231]	6[224]	ATCATTCTTTCATCTTCTGAC	21	
13[105]	12[112]	GTGAATTATCACCGCAGAGCC	21	
3[238]	2[245]	TAAAAATACCGAACGCGGTCA	21	
23[161]	22[168]	CTAAAGTACGGTGTAATCATA	21	
19[161]	16[154]	ATCCGCGAAAGGCCGCTTTTG	21	
13[168]	12[175]	CACCATTACCATTACACCGGA	21	
29[70]	26[63]	ATTAATTTTGTTATCCGCTCA	21	
3[70]	2[77]	TTACATTGGCAGATTTTAGGA	21	
19[77]	16[70]	TCATCTTGACTAAAGACTTTT	21	
9[210]	6[203]	AATAATCTAATGGTTTGAAAT	21	
29[112]	26[105]	CCTGTCGCTCGAATTCGTAAT	21	
11[245]	9[251]	CAAGAAACAATGAAAAGCCGTTTTTATTGGGTATT	35	
2[223]	4[217]	GTGCCACCTTCTGAATAATGGTTAACGTCAGATGA	35	
1[210]	29[216]	TATGGTTGCTTTGACCGAAATCGGCAAAGGCAACA	35	
22[62]	24[56]	AGCCTTTTGCCTGAGTAATGTTTCATCAACATTAA	35	
19[182]	16[175]	ACTTAGCCGCTGAGGCTTGCA	21	
2[55]	4[49]	GAGCCGTGACAACTCGTATTATTAATTACATTTAA	35	
9[126]	6[119]	ACGCGCCTTTAGTATCATATG	21	
29[91]	26[84]	GCCCGCTATAGCTGTTTCCTG	21	
23[56]	22[63]	AGAGTACCTTTAATCGGGAGA	21	
2[76]	4[70]	GCACTAATGCCCGAACGTTATACAAACATCAAGAA	35	
13[63]	12[70]	GAGGGAGGGAAGGTTGACAGG	21	
13[84]	12[91]	GACGGAAATTATTCCACCAGA	21	
22[188]	24[182]	ATTAACAAGAGATCTACAAAGATATTTAAATTGTA	35	
22[125]	24[119]	AGAGCATCATCAATATGATATCATTTTTTAACCAA	35	
22[146]	24[140]	AAAATTATTCTAGCTGATAAATAAATTTTTGTTAA	35	
9[252]	6[245]	AAACCAATTTCAAATATATTT	21	
19[98]	16[91]	TATACCAAACGGCTACAGAGG	21	
3[133]	2[140]	TAGAACCCTTCTGACAATATC	21	
23[77]	22[84]	TTTGATAAGAGGTCTATGACC	21	
13[126]	12[133]	CTTGAGCCATTTGGCACCCTC	21	
2[139]	4[133]	TGGTCAGACCAGAAGGAGCGGGTTACAAAATCGCG	35	
2[160]	4[154]	ATCAAACATCATATTCCTGATCCTGATTGCTTTGA	35	
2[118]	4[112]	GTTGAAAATTTTGCGGAACAAGAATTATTCATTTC	35	
2[181]	4[175]	TCACCTTTGATGGCAATTCATGGAGAAACAATAAC	35	
29[133]	26[126]	ATGAATCAGAGGATCCCCGGG	21	
12[132]	14[126]	AGAGCCAATAAGTTTTAACGGCTCAGAGCCACCAC	35	
22[209]	24[203]	AATTCTAGGTCATTGCCTGAGACAGGAAGATTGTA	35	
1[231]	29[237]	GTATAACGTGCTTTAAATCCTGTTTGATACCGCCT	35	

Figure 69: Tube staple list.

19[245]	16[238]	AGGACAGGATACCGATAGTTG	21	
2[244]	4[238]	GTATTAAAGAACCTACCATATAAATTGCGTAGATT	35	
12[48]	14[42]	TTGGCCTGCCAGAATGGAAAGAACTACAACGCCTG	35	
22[167]	24[161]	CAGGCAACGGAGAGGGTAGCTATATTTTGTTAAAA	35	
13[189]	12[196]	CGGAAACGTCACCAAATCACC	21	
9[189]	6[182]	ACGAGCAGTGTGATAAATAAG	21	
23[182]	22[189]	TTTCATTCCATATATAGTAGC	21	
19[140]	16[133]	CGCCTGAGTCACCCTCAGCAG	21	
23[119]	22[126]	GCTGAATATAATGCAAGCCTC	21	
22[230]	24[224]	GCGAGCTAGCAAACAAGAGAAATAATCAGAAAAGC	35	
2[97]	4[91]	TTATCTAAAAAGTTTGAGTAATGAGCAAAAGAAGA	35	
19[119]	16[112]	TACAACGCAGCATCGGAACGA	21	
29[238]	26[231]	GGCCCTGGGATGTGCTGCAAG	21	
13[231]	12[238]	CAGTAGCGACAGAATCATAGC	21	
9[63]	6[56]	GACAAAATAATTGAGAATCGC	21	
12[195]	14[189]	GGAACCACCTATTTCGGAACCATCACCGTACTCAG	35	
12[90]	14[84]	GCCGCCGAGCGTCATACATGGAGGAACCCATGTAC	35	
11[35]	9[41]	CATATTATTTATCCTTTCCAGAGCCTAAAGCCAGT	35	
3[154]	2[161]	GCGTAAGAATACGTCTCAAAT	21	
19[224]	16[217]	CGAACTGAATGACAACAACCA	21	
22[104]	24[98]	TGTACCAGGCCGGAGACAGTCGCCATCAAAAATAA	35	
9[147]	6[140]	AGATAAGATAATTACTAGAAA	21	
12[216]	14[210]	CTTTTCACTGAAACATGAAAGAGTATAGCCCGGAA	35	
2[202]	4[196]	GCAAATGATCCTGATTGTTTGGTAACAGTACCTTT	35	
23[224]	22[231]	TAGATTTAGTTTGATTGGGGC	21	
19[56]	16[49]	CAAAAGAGAAGTTTCCATTAA	21	
3[49]	2[56]	CAATCGTCTGAAATTAGATTA	21	
29[175]	26[168]	GCGCCAGACGAGCCAGTGC	21	
13[147]	12[154]	AGCCAGCAAAATCAAGCCGCC	21	
29[217]	26[210]	GCTGATTAGTTGGGTAACGCC	21	
29[154]	26[147]	GAGAGGCGCATGCCTGCAGGT	21	
3[217]	2[224]	ATAGCCCTAAAACATGCAACA	21	
12[237]	14[231]	CCCCTTAAGGCTGAGACTCCTGCCGTCGAGAGGGT	35	
13[210]	12[217]	CATCGATAGCAGCATTGCCAT	21	
11[224]	9[230]	AGTTAAGCCCAATATAGGAATCATTACCTTTCCTT	35	
11[56]	9[62]	AATAAGAAACGATTTTACCAACGCTAACAAGTACC	35	
9[84]	6[77]	GTCCAGACCAACGCTCAACAG	21	
19[203]	16[196]	GACGGTCCGCATAACCGATAT	21	
12[153]	14[147]	ACCCTCAGCCTTGAGTAACAGGCCACCCTCAGAAC	35	
29[196]	26[189]	TCACCAGTCCCAGTCACGACG	21	
9[42]	6[35]	AATAAGAAACAACGCCAACAT	21	
1[42]	29[48]	AAAGGGAGCCCCCGTGGGGTCGAGGTGCGGGGTGC	35	
22[83]	24[77]	CTGTAATAAGATTCAAAAGGGCTGGCCTTCCTGTA	35	
23[140]	22[147]	CAACATGTTTTAAAAATTAGC	21	
3[112]	2[119]	GGACATTCTGGCCAATCAACA	21	
12[69]	14[63]	AGGTTGATCTGAATTTACCGTCTGAGTTTCGTCAC	35	
21[49]	19[55]	CGAACTAACGGAACTTAAGAACTGGCTCAAAGAGG	35	
12[111]	14[105]	GCCACCATGATACAGGAGTGTTTCAGGGATAGCAA	35	
23[98]	22[105]	CGGATGGCTTAGAGAATCGGT	21	
13[42]	12[49]	AAGGGCGACATTCACAGACGA	21	
9[105]	6[98]	AACATGTCAAATTCTTACCAG	21	
12[174]	14[168]	ACCGCCTATAAACAGTTAATGAGTACCGCCACCCT	35	
23[203]	22[210]	ATTCCCAATTCTGCTGGCATC	21	
3[175]	2[182]	ACAATATTTTTGAATAAAGCA	21	
9[168]	6[161]	AATAATAATAAGAATAAACAC	21	
3[196]	2[203]	TAGTCTTTAATGCGGCCAGCA	21	
21[91]	19[97]	AGTTGAGATTTAGGTTTAATTTCAACTTCAGCGAT	35	
3[91]	2[98]	TCACACGACCAGTAAGGAAGG	21	
21[196]	19[202]	CGACGATAAAAACCCGGATATTCATTACAGGCGCA	35	
21[112]	19[118]	CATTCAACTAATGCAGTAAATTGGGCTTAACAAAG	35	
11[182]	9[188]	GTAATTGAGCGCTAGAAGGCTTATCCGGCTAATTT	35	
1[126]	29[132]	GCTAGGGCGCTGGCTATTAAAGAACGTGTGCATTA	35	
21[175]	19[181]	TCATAACCCTCGTTAACGTAACAAAGCTCCATGTT	35	
21[238]	19[244]	GTTTTGCCAGAGGGATAGGCTGGCTGACTTGAAAG	35	
21[217]	19[223]	CGAGAGGCTTTTGCAAGAGTAATCTTGAAGGGAAC	35	
1[189]	29[195]	GCGCCGCTACAGGGATAAATCAAAAGAATTTCTTT	35	

11[140]	9[146]	TAGACGGGAGAATTAACCTCCCGACTTGCAACAAT	35	
1[63]	29[69]	GCTTGACGGGGAAACATCACCCAAATCAAACTCAC	35	
11[77]	9[83]	TAACGTCAAAAATGGCTACAATTTTATCTAATTCT	35	
1[147]	29[153]	GCGGTCACGCTGCGCAGTTTGGAACAAGGCGCGGG	35	
11[203]	9[209]	AGAGATAACCCACAATAGCAAGCAAATCACCAATC	35	
21[70]	19[76]	TTACAGGTAGAAAGTGTGAATTACCTTAAAAACAC	35	
21[133]	19[139]	TAACGCCAAAAGGAACGAGAAACACCAGGTATCAT	35	
11[98]	9[104]	CAGCCTTTACAGAGTTAGTTGCTATTTTAATAAAC	35	
1[84]	29[90]	AACGTGGCGAGAAATCTATCACCCACTAGCTCACT	35	
21[154]	19[160]	GGCATAGTAAGAGCCAGTGAATAAGGCTTGTCGAA	35	
1[105]	29[111]	AAGAAAGCGAAAGGACGTCAAAGGGCGACGGGAAA	35	
1[168]	29[174]	ACCACACCCGCCGCAGATAGGGTTGAGTGTATTGG	35	
11[161]	9[167]	CACCCTGAACAAAGAGAACGCGAGGCGTCAAGAAA	35	
11[119]	9[125]	CATAAAAACAGGGATTTTGAAGCCTTAAATGCAGA	35	

Figure 70: Tube staple list contd.

Start	End	Sequence 5'–3'	Length	Color
20[239]	18[240]	TGAGGAAGAACATTATCATTTTGCTATACTTC	32	
7[32]	9[31]	TTTCAACAACCCTCGTTTACCATTTGCAAA	30	
21[96]	23[95]	GAAACCTGGGCAACAGCTGATTGCGAACAAGA	32	
17[32]	19[31]	ATAATTCTGGTGCCGGAAACCAACTGTTGG	30	
18[207]	16[208]	CAAAATTACAATAACGGATTCGCCATAAATCA	32	
8[47]	6[48]	ACTATCATTTTAATCATTGTGAATATTCAGTG	32	
14[143]	13[159]	TTGCGGGAAAACATTATGACCCTGCGGAATCA	32	
13[224]	15[223]	CAACGCTATTTCATCTTCTGACATTTATCA	30	
11[64]	13[63]	TGATAAGAAGATACATTTCGCAAAAATCATAC	32	
16[143]	15[159]	CAATCATAACGGTAATCGTAAAACCCTTAGAA	32	
11[96]	13[95]	AGCTTAATTCTGCGAACGAGTAGAATTAAGCA	32	
16[175]	14[176]	AATCGTCGAACATAGCGATAGCTGATAAAT	30	
11[32]	13[31]	AGGATTATAGCTATATTTTCATCTACTAAT	30	
20[47]	18[48]	TTGTTATCGATCGGTGCGGGCCTCTTTCCGGC	32	
8[271]	6[272]	AACATAAACAATAATAACGGAATTATTACGC	30	
1[192]	3[191]	GGTCAGTGAGCGCAGTCTCTGAATATCACCGG	32	
21[256]	23[255]	AAATGAAACCAACAGAGATAGAACCGCTCAAT	32	
13[120]	15[127]	AATCGGTTGTACCAAGAAGCCTTTATTTCACAAGAGAA	38	
18[175]	16[176]	GAAATTGCACAGTAACAGTACCTTTCTGTA	30	
8[79]	6[80]	CAAAAGGAAAGAACTGGCTCATTCATTACC	30	
22[47]	20[48]	CCACGCTGTGAGTGAGCTAACTCAGTGTGAAA	32	
6[271]	4[272]	AGTATGTTTGGGAATTAGAGCCAAAGGCCG	30	
7[224]	9[223]	ATAGCCGTGAACACCCTGAACACAGTTACA	30	
10[207]	8[208]	CCGACTTGCCAGAGCCTAATTTGCAAGTCAGA	32	
21[32]	23[31]	TGCCTAAGTTTGCCCCAGCAGGGGCAAAAT	30	
8[143]	7[159]	TAGAAAGACTAACGGAACAACATTATAATAAG	32	
5[32]	7[31]	ATAAATTTTGCCCTGACGAGAATGGTTTAA	30	
9[64]	11[63]	TGTTTAGAATTAAGAGGAAGCCCGGCTCCTTT	32	
3[224]	5[223]	GAGCCGCCAGTAGCGACAGAATAATTATTC	30	
4[207]	2[208]	CCTTTAGCCCACCACCGGAACCGCTAAAGCCA	32	
2[143]	1[159]	ACGTTGAAGGAATTGCGAATAATATTGATGAT	32	
9[96]	11[95]	AATCGTCAAGTCAGAAGCAAAGCGTGGCTTAG	32	
23[32]	22[48]	CCCTTATAAATCAAAAGAATAGCCAAGCGGT	31	
6[111]	4[112]	TAATCTTGAAGGGAACCGAACTGAATACGTAA	32	
10[191]	12[184]	TTTGAAGCTAGAAACCAATCAATATGCAGAACGCGCCCTGT	40	
19[96]	21[95]	AAGGCGATCCCGGGTACCGAGCTCCCAGTCGG	32	
4[175]	2[176]	CGGCATTTCTTTTCATAATCAAATTACCGT	30	
2[111]	0[112]	AAGGAGCCAGCGGAGTGAGAATAGACCCTCAT	32	
6[143]	5[159]	GGCGCATAGACAGATGAACGGTGTCAGCGCCA	32	
16[47]	14[48]	AAAATTCGAAATTAATGCCGGAGAGGTAAAGA	32	
1[160]	2[144]	ACAGGAGTGCGTCATACATGGCTTATTTTTTC	32	
10[239]	8[240]	TCTAAGAACAGCCATATTATTTATAGACGGGA	32	
0[79]	1[63]	TGTACCGTAACACTGAGTTTCGTCAATGAATT	32	
19[160]	20[144]	ACTTTACACATTTGAGGATTTAGAACGGCCAG	32	
9[224]	11[223]	AAATAAACGCGAGGCGTTTTAGTATCATTC	30	
1[32]	3[31]	TGTCGTCAGTTGCGCCGACAATATTCGGTC	30	
8[111]	6[112]	ACCACATTTGGGAAGAAAAATCTAATCAAGAG	32	
16[239]	14[240]	ACCTTTTTAGGTCTGAAGAGACTACAAATATAT	32	
1[256]	3[263]	TATTATTCCAGGTCAGACGATTGGACCCTCAGAGCCACCA	40	
11[256]	13[255]	CATCGAGAAGTACCGACAAAAGGTTTGAGAAT	32	
15[224]	17[223]	AAATCATTAATGGAAACAGTACTGATTGCT	30	
12[207]	10[208]	TCAGCTAAATCGGCTGTCTTTCCTCGAACGTC	32	
12[239]	10[240]	TCTGTCCAGGGTATTAAACCAAGTTCCGGTAT	32	
16[63]	18[56]	ACGTTAATTAGCCAGCTTTCATCATCGCACTCCAGCCAGC	40	
22[111]	20[112]	GTGAGACGTCGTGCCAGCTGCATTTCGACTCT	32	
21[128]	23[127]	GGCCAACCAGGGTGGTTTTTCTCCAACGTC	30	
20[79]	18[80]	TAATCATGCTGGCGAAAGGGGGATCGGCCT	30	
1[64]	3[71]	TTCTGTATGAGGTGAATTTCTTAAGGCCGCTTTTGCGGGA	40	
23[192]	22[208]	ATTACCGCCAGCCATTGCAACAGGCACAGACA	32	
14[175]	12[176]	AAGGCGTTAGAAAAAGCCTGTTTTTATCAA	30	
0[207]	1[191]	GGAATAGGTGTATCACCGTACTCATTTAACGG	32	
16[271]	14[272]	AATTAATTTTAGGTTGGGTTATACGCAAGA	30	
12[271]	10[272]	GAATATAAACAAGCAAGCCGTTTAATAGCA	30	
10[127]	12[120]	TCAGGTCTGTAGCTCAACATGTTGTTTCATTCCATATA	38	
23[128]	23[159]	AAAGGGCGAAAAACCGTCTATCAACTCAAAC	31	
14[207]	12[208]	AATGGTTTTACAAATTCTTACCAGCAACATGT	32	
7[128]	9[135]	AAACGAATTCATCAGTTGAGATCCCCCTCAAATGCTT	37	
10[175]	8[176]	CAAGATTACTGAATCTTACCAACGAGATAA	30	
18[79]	16[80]	CAGGAAGAACATTAAATGTGAGCATAAGCA	30	
16[255]	18[248]	CAATTTCACGCAGAGGCGAATTATTCCTGATTGTTTGGAT	40	
5[192]	7[191]	GGAGGGAAGAATAAGTTTATTTTGTAGCTATC	32	
14[111]	12[112]	ATAAAAATTCAGAGCATAAAGCTAACAGTTGA	32	
3[96]	5[95]	CAGCATCGCGAAGGCACCAACCTAGCAGACGG	32	
7[56]	9[63]	TACCTTATGCGATTTTATTACGAGGCATAGTATAGTAAAA	40	
5[224]	7[223]	ATTAAAGCAACATATAAAAGAAGTAAGCAG	30	
18[47]	16[48]	ACCGCTTCGCGTCTGGCCTTCCTGATTTTGTT	32	
7[96]	9[95]	CAGGACGTCAACTAATGCAGATACTACTGCGG	32	
18[271]	16[272]	CAATATAATCATTTCAATTACCTAAAACAA	30	
15[96]	17[95]	CATTGCCTAAAACAGGAAGATTGTGAGTAACA	32	
12[47]	10[48]	AACCTGTTGAGAGTACCTTTAATTAAAGACTT	32	
20[175]	18[176]	AGATAATAAACAATTCGACAACTAAATAAA	30	
19[32]	21[31]	GAAGGGCCGCTCACAATTCCAGCCTGGGG	30	
8[175]	6[176]	CCCACAAGACAATGAAATAGCAATCACAAT	30	
10[271]	8[272]	AGCAAATCAACGATTTTTTGTTTAGAGAAT	30	
14[79]	12[80]	AATGCAATCAAAGAATTAGCAAATTTAGTT	30	
22[143]	21[159]	TTGGGCGCCGCGGGGGAGAGGCGGGCCATTAA	32	
13[160]	14[144]	TAATTACTAAATAAGAATAAACACTAATACTT	32	
23[64]	22[80]	GGGTTGAGTGTTGTTCCAGTTTGCCTTCAC	30	
4[255]	6[248]	AACCATCGACCGACTTGAGCCATTAGCAAACGTAGAAAAT	40	
19[248]	21[255]	GGAACAAAGAAACCACGGCAAATCAACAGTTGCCAGCAGC	40	
4[111]	2[112]	TGCCACTAGAACGAGGGTAGCAACGGCTCCAA	32	
23[224]	22[240]	TCATGGAAATACCTACATTTTGACCTTCTGA	31	
3[128]	5[127]	AGGCTTTCATTAAACGGGTAAACCAACTTT	30	

15[128]	17[127]	TCGATGATGTACCCCGGTTGATAAACGGCG	30	
6[175]	4[176]	CAATAGAAAGGGCGACATTCAACTTTTCAT	30	
2[239]	0[240]	TTCACAAAATGCCCCCTGCCTATTGGATAAGT	32	
14[271]	12[272]	CAAAGAACTTAACAACGCCAACAAATAAGA	30	
20[111]	18[112]	AGAGGATCTAAGTTGGGTAACGCCGCATCTGC	32	
20[207]	18[208]	CACTAACACCCGAACGTTATTAATTACCATAT	32	
7[248]	9[255]	GGAAACCGAGGAAACGAACAGGGAAGCGCATTCCCAATCC	40	
17[160]	18[144]	ATGAATATGTAGATTTTCAGGTTTACGTTGGT	32	
22[239]	20[240]	CCTGAAAGAGTGCCACGCTGAGAGAAAGGAAT	32	
8[207]	6[208]	GGGTAATTGCCCTTTTTAAGAAAAACGCAAAG	32	
23[256]	22[272]	CGTCTGAAATGGATTATTTACATAAAGGGA	30	
18[111]	16[112]	CAGTTTGAGATTCTCCGTGGGAACAATCAGAA	32	
14[239]	12[240]	TTTAGTTACAACAGTAGGGCTTAAAAAGTAAT	32	
10[47]	8[48]	CAAATATCTTGCCAGAGGGGGTAAAGAGCAAC	32	
2[175]	0[176]	TCCAGTAAGTACTGGTAATAAGTGGAGGTT	30	
6[207]	4[208]	ACACCACGGGTAAATATTGACGGACAAGTTTG	32	
23[96]	22[112]	GTCCACTATTAAAGAACGTGGACTTTTCACCA	32	
10[79]	8[80]	TCAAAAAGCTGGATAGCGTCCAAATAACGC	30	
2[47]	0[48]	ATACCGATTTTCCAGACGTTAGTAACCAGTAC	32	
0[239]	1[223]	GCCGTCGAGAGGGTTGATATAAGCCGTATAA	31	
6[239]	4[240]	AAAGGTGGGTGAATTATCACCGTCATAGCAGC	32	
16[79]	14[80]	AATATTTAGAGATCTACAAAGGCTATTTTA	30	
4[79]	2[80]	AGAGGCAATCGTCACCCTCAGCATATCAGC	30	
21[160]	22[144]	AAATACCGATAGCCCTAAAACATCTTTGCGTA	32	
12[79]	10[80]	TGACCATTGGTCATTTTTGCGGAGATTGCA	30	
22[207]	20[208]	ATATTTTTTGAGGCGGTCAGTATTTTTAGGAG	32	
17[128]	19[127]	GATTGACGCGCATCGTAACCGTAGGGTTTT	30	
23[160]	22[176]	TATCGGCCTTGCTGGTAATATCCTTAATGC	30	
1[128]	3[127]	AACTAAAAATCTCCAAAAAAAAGGCTACAG	30	
2[271]	0[272]	GGTTGAGGTGAAACATGAAAGTAAGGATTA	30	
0[47]	1[31]	AAACTACAACGCCTGTAGCATTCTAAAGTTT	31	
9[256]	11[255]	AAATAAGAAGATATAGAAGGCTTAACCGCACT	32	
4[47]	2[48]	TTGACCCCCTTGCAGGGAGTTAAAACAGCTTG	32	
17[192]	19[191]	GGGAGAAATTTGCACGTAAAACAGCGTATTAA	32	
14[47]	12[48]	TTCAAAGGCATTAACATCCAATATGGTCAAT	32	
1[224]	3[223]	ACAGTTACAAATAAATCCTCATCTCCCTCA	30	
5[96]	7[95]	TCAATCATACAAGAACCGGATATTATACCAGT	32	
22[79]	20[80]	CGCCTGGCCTCACTGCCCGCTTTGAATTCG	30	
21[192]	23[191]	AACAGAGGGAATGGCTATTAGTCTAGAACAAT	32	
20[271]	18[272]	GGTCAGTTCAGAAGGAGCGGAATAATTCAT	30	
18[143]	17[159]	GTAGATGGCGTAATGGGATAGGTCAACGTCAG	32	
0[143]	1[127]	GAACCGCCACCCTCAGAGCCACCAAAGGAAC	31	
13[32]	15[31]	AGTAGTAGGTGAGAAAGGCCGACCGTTCT	30	
21[64]	23[63]	GCGTTGCGCCTGAGAGAGTTGCAGCCGAGATA	32	
1[96]	3[95]	ACAGTTTCTTTAATTGTATCGGTTGCGAAAGA	32	
22[271]	20[272]	CATTCTGGAATCTAAAGCATCACAATATCT	30	
13[256]	15[263]	CGCCATATGCGAGAAAACTTTTTCCTTTTTAACCTCCGGC	40	
11[224]	13[223]	CAAGAACGACGACGACAATAAATATAAAGC	30	
16[111]	14[112]	AAGCCCCAGAGAGTCTGGAGCAAAACGCAAGG	32	
9[160]	10[144]	ATTTTATCGTTGCTATTTTGCACCCGAGAATG	32	
10[111]	8[112]	ACTATTATTAAATATTCATTGAATTTAGGAAT	32	
19[224]	21[223]	TTTGAGTGTTATCTAAAATATCAACACCGC	30	
2[207]	0[208]	GAATGGAACCTTGAGTAACAGTGCTATAGCCC	32	
12[143]	11[159]	GTCTGGAATTAAATATGCAACTAAAAATAATA	32	
21[224]	23[223]	CTGCAACCGTAAGAATACGTGGAAAAACGC	30	
13[64]	15[71]	AGGCAAGGGCCTGAGTAATGTGTAGGGTAGCTATTTTTGA	40	
22[175]	20[176]	GCGAACTGAACGAACCACCAGCACGTCAAT	30	
3[192]	5[191]	AACCAGAGGTCAGACTGTAGCGCGCGATTGAG	32	
4[239]	2[240]	ACCGTAATCACCCTCAGAACCGCCCCTTGATA	32	
5[128]	7[127]	GAAAGAGGGCTGGCTGACCTTCCGTTAATA	30	
19[128]	21[127]	CCCAGTCTTGCATGCCTGCAGGAATGAATC	30	
3[160]	4[144]	TTTGCCATTCGGTCATAGCCCCCTTTCATGAG	32	
8[239]	6[240]	GAATTAACAACAAAGTTACCAGAAACATACAT	32	
15[160]	16[144]	TCCTTGAACTATTAATTAATTTTCTAGCATGT	32	
12[175]	10[176]	CAATAGATTAATTTACGAGCATGCTTAAAT	30	
11[160]	12[144]	TCCCATCCAAGTCCTGAACAAGAAAGTACGGT	32	
0[271]	1[255]	GCGGGGTTTTGCTCAGTACCAGGCTCGGAACC	32	
7[192]	9[199]	TTACCGAAGAGCGCTAATATCAGAGCTAACGAGCGTCTTT	40	
12[111]	10[112]	TTCCCAATTGCTGAATATAATGCTTTACCCTG	32	
18[239]	16[240]	TGAATAATCCAAGTTACAAAATCGTTTGAATT	32	
10[143]	9[159]	ACCATAAATAAACAGTTCAGAAAACAGCTACA	32	
7[160]	8[144]	AGCAAGAAAATTGAGTTAAGCCCAATTACAGG	32	
19[56]	21[63]	TTCGCTATTACGCCAGGTCATAGCTGTTTCGTCATTAATT	40	
16[207]	14[208]	ATATATGTGAAGAGTCAATAGTGACTAAATTT	32	
15[192]	17[191]	GACGCTGAGAGTGAATAACCTTGCTTTACATC	32	
20[143]	19[159]	TGCCAAGCACGACGTTGTAAAACGAGTATTAG	32	
4[271]	2[272]	GAAACGTCCCCTCAGAGCCGCCAGACAGGA	30	
15[32]	17[31]	AGCTGATCATTAAATTTTTGTTCATCAAAA	30	
13[184]	15[191]	AGTATCATATGCGTTAGAAATACCGACCGTGTTAGATTAA	40	
17[96]	19[95]	ACCCGTCGGGGGACGACGACAGTATGTGCTGC	32	
6[79]	4[80]	CAAATCAATAGCCGGAACGAGGCAAACGAA	30	
2[79]	0[80]	TTGCTTTCGGGATTTTGCTAAACGAACCCA	30	
5[160]	6[144]	AAGACAAAAATTCATATGGTTTACACAGACCA	32	
17[224]	19[223]	TTGAATAGGAAGGGTTAGAACCTTTAAAAG	30	
6[47]	4[48]	AATAAGGCGTGTCGAAATCCGCGAACTCATCT	32	
0[175]	0[144]	TAGTACCGCCACCCTCAGAACCGCCACCCTCA	32	
9[32]	11[31]	AGAAGTTGCGTTTTAATTCGAGAACAGGTC	30	
0[111]	1[95]	TTTCAGGGATAGCAAGCCCAATAGAACTTTCA	32	
13[96]	15[95]	ATAAAGCCTTTTAGAACCCTCATATATCAGGT	32	
4[63]	6[56]	ACTAAAACCCTGCTCCATGTTACTCGTAACAAAGCTGCTC	40	
19[192]	21[191]	ATCCTTTGACTAATAGATTAGAGCGAAGATAA	32	
3[32]	5[31]	GCTGAGGCAGCGATTATACCAAATCGCCTG	30	
4[143]	3[159]	GAAGTTTCGAGGACTAAAGACTTTTATTAGCG	32	

Figure 71: tCRRO staple list.

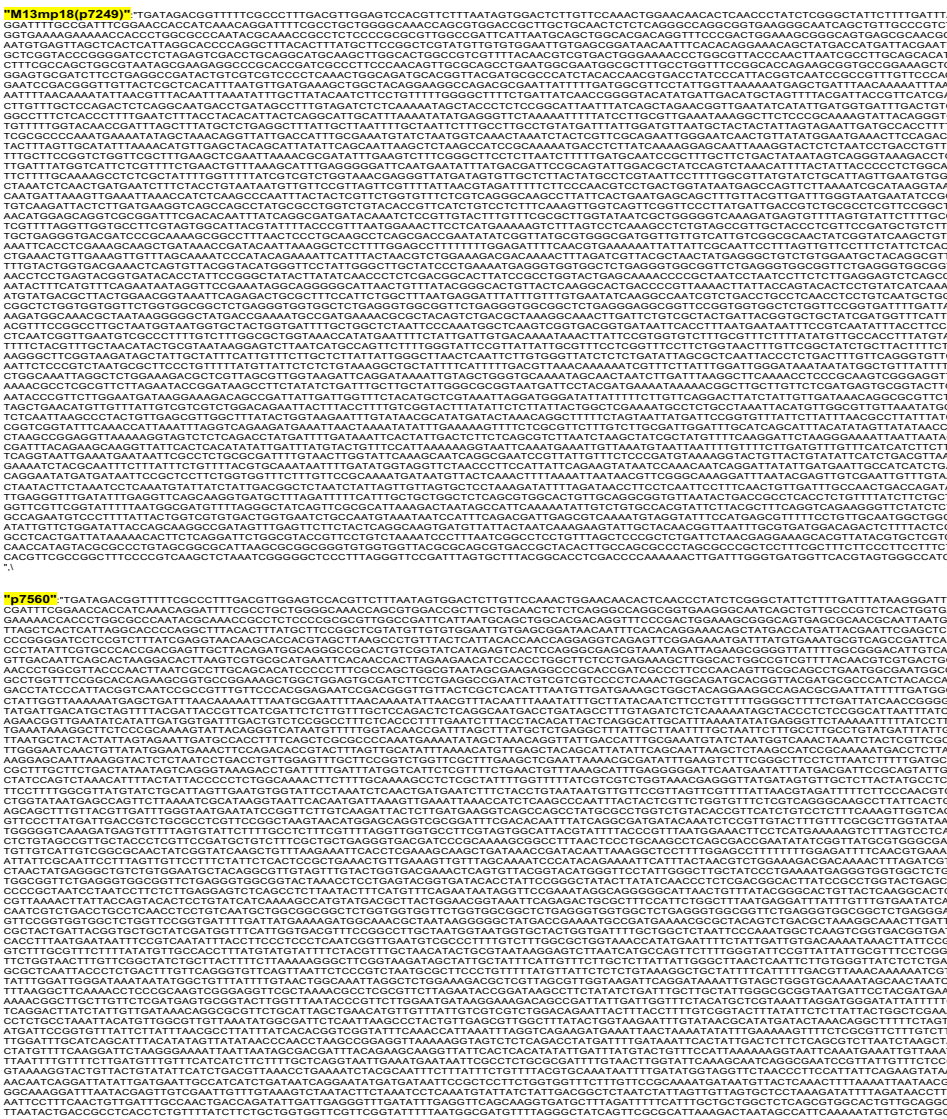


Figure 72: M13mp18 phage ssDNA scaffold variants. M13mp18 original sequence (p7249) used for folding tube and rectangles; and p7560 used for folding 42HB.

For 42HB

°C	Time/ °C
65	15 min
60-40	1K/ hour
20	Forever
Time	~21 hours

For tube and tCRRO

°C	Time/ °C
70	5 min
65-45	0.5K/min
20	Forever
Time	~1 hour

Figure 73: Thermal annealing ramp for folding.

DNA and RNA sequences corresponding to RNA nanostructures

Primer sequences used for all RNA tiles designed are given below.

	Length (nt)	Tm (°C)	GC (%)	Sequence (5'-3')
Forward Primer (FP)	22	72	59	AACTCTCGATCAGCGGACAGCG
Revere Primer (RP)	20	74	70	GCAGAGTGCGGACCGCAGTG

Figure 74: Primer sequences for amplifcation of DNA template.

DNA templates for PCR corresponding to 3H-AE RNA tiles and its modifications were ordered as gBlocks from IDT. DNA and RNA sequences are provided in figure 75 and 76 respectively. Red and green regions are analogous to the forward and reverse primers respectively. Blue regions code for the T7RNAP promoter binding region (pT7) and black regions encode for the RNA tile.

	Length (nt)	Sequence (5'-3')
3H-AE	288	AACTCTCGATCAGCGGACAGCGCTAATACGACTCACTATAGGGAGATACCTGGGAATGTGCTCACGTCTCGAGCATAT TCGGCCTGCTTCGGCAGGCGCTCGCGAGACGTGCGAGCGCAGGGTGAAGCCTCCACGCCCTGCCCTCACTTCGGTG AGCCAGGTGTCTCCCGCAGGGTGCGCGTGGCTTCGGCCACGACCTGAAGGAGGCACGGGTCGACGGTACCCAGTGG ACGGGTATCGTCGCGGAGCTTCGGCTCCGGACCGCGTCCACTGCGGTCCGCACTCTGC
3H-AE-MG-V1	337	AACTCTCGATCAGCGGACAGCGCTAATACGACTCACTATAGGGAGATACCTGGGAATGTGCTCACGTCTCGAGCATATTC GGCCTGCTTCGGCAGGCGCTCGCGAGACGTGCGAGCGCAGGGTGAAGCCTCCACGCCCTGCCCTCACTTCGGTGAGCC AGGTGTCTCCCGCAGGGTGCGCCGGCAGCCTCGGATCCCGACTGGCGAGAGCCAGGTAACGAATGGATCCGAGGAACT ACTGTCGGACCTGAAGGAGGCACGGGTCGACGGTACCCAGTGGACGGGTATCGTCGCGGAGCTTCGGCTCCGGACCGC GTCCACTGCGGTCCGCACTCTGC
3H-AE-MG-V2	343	AACTCTCGATCAGCGGACAGCGCTAATACGACTCACTATAGGGAGATACCTGGGAATGTGCTCACGTCTCGAGCATATTC GGCCTGCTTCGGCAGGCGCTCGCGAGACGTGCGAGCGCAGGGTGAAGCCTCCACGCCCTGCCCTCACTTCGGTGAGCC AGGTGTCTCCCGCAGGGTGCGCGCTGGCCCCTCCAGGATCCCGACTGGCGAGAGCCAGGTAACGAATGGATCCTGTGAG GAACTAGGCCGCGACCTGAAGGAGGCACGGGTCGACGGTACCCAGTGGACGGGTATCGTCGCGGAGCTTCGGCTCCGGA CCGCGTCCACTGCGGTCCGCACTCTGC
3H-AE-M(Spinach	428	AACTCTCGATCAGCGGACAGCGCTAATACGACTCACTATAGGGAGATACCTGGGAATGTGCTCACGTCTCGAGCATAT TCGGCCTGCTTCGGCAGGCGCTCGCGAGACGTGCGAGCGCAGGGTGAAGCCTCCACGCCCTGCCCGGCAGCCTCGA CGCGACCGAAATGGTGAAGGACGGGTCCAGTGCTTCGGCACTGTTGAGTAGAGTGTGAGCTCCGTAACTGGTCGCGT CGAGGAACTACTGTCGCCAGGTGTCTCCCGCAGGGTGCGCCGGCAGCCTCGGATCCCGACTGGCGAGAGCCAGGTAA CGAATGGATCCGAGGAACTACTGTCGGACCTGAAGGAGGCACGGGTCGACGGTACCCAGTGGACGGGTATCGTCGCGG AGCTTCGGCTCCGGACCGCGTCCACTGCGGTCCGCACTCTGC

Figure 75: DNA gBlock sequences.

	Length (nt)	Sequence (5'-3')
3H-AE	248	GGGAGAUACCUGGGAAUGUGCUCACGUCUCGAGCAUAUUCGGCCUGCUUCGGCAGGCGCUCGCGAGACGUG CGAGCGCAGGGUGAAGCCUCCACGCCCUGCCCUCACUUCGGUGAGCCAGGUGUCUCCCGCAGGGUGCGCGUGGC UUCGGCCACGACCUGAAGGAGGCACGGGUCGACGGUACCCAGUGGACGGGUAUCGUCGCGGAGCUUCGGCUCCG GACCGCGUCCACUGCGGUCCGCACUCUGC
3H-AE-MG-V1	297	GGGAGAUACCUGGGAAUGUGCUCACGUCUCGAGCAUAUUCGGCCUGCUUCGGCAGGCGCUCGCGAGACGUGCGA GCGCAGGGUGAAGCCUCCACGCCCUGCCCUCACUUCGGUGAGCCAGGUGUCUCCCGCAGGGUGCGCCGGCAGCC UCGGAUCCCGACUGGCGAGAGCCAGGUAACGAAUGGAUCCGAGGAACUACUGUCGGACCUGAAGGAGGCACGGGU CGACGGUACCCAGUGGACGGGUAUCGUCGCGGAGCUUCGGCUCCGGACCGCGUCCACUGCGGUCCGCACUCUGC
3H-AE-MG-V2	303	GGGAGAUACCUGGGAAUGUGCUCACGUCUCGAGCAUAUUCGGCCUGCUUCGGCAGGCGCUCGCGAGACGUGCGAGC GCAGGGUGAAGCCUCCACGCCCUGCCCUCACUUCGGUGAGCCAGGUGUCUCCCGCAGGGUGCGCGCUGGCCCCUCCAG GAUCCCGACUGGCGAGAGCCAGGUAACGAAUGGAUCCUGUGAGGAACUAGGCCGCGACCUGAAGGAGGCACGGGUCGAC GGUACCCAGUGGACGGGUAUCGUCGCGGAGCUUCGGCUCCGGACCGCGUCCACUGCGGUCCGCACUCUGC
3H-AE-M(Spinach	388	GGGAGAUACCUGGGAAUGUGCUCACGUCUCGAGCAUAUUCGGCCUGCUUCGGCAGGCGCUCGCGAGACG UGCGAGCGCAGGGUGAAGCCUCCACGCCCUGCCCGGCAGCCUCGACGCGACCGAAAUGGUGAAGGACGGGU CCAGUGCUUCGGCACUGUUGAGUAGAGUGUGAGCUCCGUAACUGGUCGCGUCGAGGAACUACUGUCGCCAGG UGUCUCCCGCAGGGUGCGCCGGCAGCCUCGGAUCCCGACUGGCGAGAGCCAGGUAACGAAUGGAUCCGAGGAA CUACUGUCGGACCUGAAGGAGGCACGGGUCGACGGUACCCAGUGGACGGGUAUCGUCGCGGAGCUUCGGCUCC GGACCGCGUCCACUGCGGUCCGCACUCUGC

Figure 76: RNA tile sequences.

PCR amplification conditions for DNA gBlocks

	Stock	final	Volume (µl)
FP	10 µM	500 nM	2.5
RP	10 µM	500 nM	2.5
gBlock	10 ng/µl	50 ng	5
Phusion master mix	2X	1X	25
nF water			15
Total			50

	Steps
1	98°C for 30sec
2	98°C for 30sec
3	72°C for 30sec
4	Repeat steps 2-3 for 24 cycles
5	72°C for 10min
6	4°C forever
Time	~ 1 hour

Figure 77: PCR reaction and conditions for DNA gBlocks.

DNA sequences for rolling circle amplification

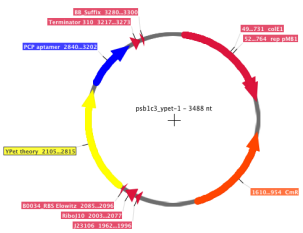
	Sequence (5'-3')
PP7-MGA (RCT)	TCGTATTAGAAGGATCCATTCGTTACCTGGCTCTCGCCAGTCGGGATCCTTCTTCAGGCAGGCACG AAGCCATATCTTCTGTGCCCCTATAGTGAG Cyan: Spacer, Red: MGA, Green: PP7 aptamer, Yellow: extra bases, Blue underlined: Split pT7.
PP7-MGA (Linear)	GGATCCATTCGTTACCTGGCTCTCGCCAGTCGGGATCCTTCTTCAGGCAGGCACGAAGCCATATCT TCTGTGCCCCTATAGTGAGTCGTATTAGAA
pT7 Splint	TTCTAATACGACTCACTATAGG
Control sequence	TCGTATTAGAATTGCTCGGCGTGCTTGCTCTCACGAACTTGACCTGGAGATCAAGGAGATGTTTCTT GTCGAACCGGACGGCGCCCCTATAGTGAG
dBroc	CTAATACGACTCACTATAGAGACGGTCGGGTCCATCTGAGACGGTCGGGTCCAGATATTCGTATCT GTCGAGTAGAGTGTGGGCTCAGATGTCGAGTAGAGTGTGGGCTC
Spacer sequences	TATAGTGAGTCGTATTAGAAAAAA
psb1c3 vector containing PCP-YPet	 Yellow- YPEt Fluorescent protein, Blue- PP7 coat protein (PCP)

Figure 78: DNA sequences for RCT.

DNA sequences for DNA nanomotor

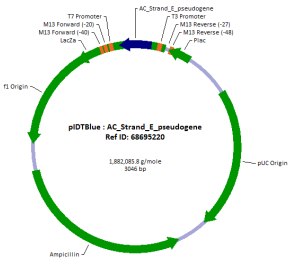
	Sequence (5'-3')
MOSIC-Strand E generated by pseudogene generator	GGTCTCACATTGCATAAAATTCATCCGTGGGAACCACGGATGAATTGGTTAGATGGTATGCTTCGGA CAGGCTAGCTACAACGAGAGTGACTGAGCGTAAGGTCTGGTTCATCCGTGGGAACCACGGATGAA CGGCTCTTCATGCATTGGAGACC Cyan and yellow highlighted regions denote restriction enzyme sequences and rest encode MOSIC-Strand E.
Strand E	TTGGTTAGATGGTATGCTTCGGACAGGCTAGCTACAACGAGAGTGACTGAGCGTAAGGTCTGGTT
Strand F unmodified	GCATACCATCTAACCAATAACCAGACCTTACGCTC
Strand F modified	FAM–GCATACCATCTAACCAATAACCAGACCTTACGCTC–TAMRA
Strand S	GTCACTCATGTCCGA
Strand SH	GTCATTCrArUGTCCGA
pIDTBlue vector containing Strand E	 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGA ATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGAT GGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGG GAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACC CGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGC TGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACT CACTATAGGGCGAATTGGGTACGATCGATGCGGCCTCGGTCTCCAATGCATGAAGAGCGGTTCATCCGTGGTTCCCACGGATGAACCAGACCTTACGCTCAGTCAC TCTCGTTGTAGCTAGCCTGTCCGAAGCATACCATCTAACCAATTCATCCGTGGTTCCCACGGATGAATTTTATGCAATGTGAGACCGCGCGTGATATGCAGCTCCAG CTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGC CGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGG CGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC CGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA TAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTAC GGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAA TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAAT AAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGT TACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCAC TGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCT GTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC Blue region codes for Strand E

Figure 79: DNA sequences for MOSIC-DNAzyme and DNA nanomotor.

Acknowledgement

I would like to express my sincere gratitude to my advisor Prof. Friedrich C. Simmel for giving me the opportunity to carry out my doctoral thesis in his lab. I would also like to thank him for his continuous support and motivation during my research and writing of this thesis.

I would also like to thank the European School of DNA Nanotechnology (EScoDNA) and SFB1032 for providing me with generous financial support to carry out my work. In addition, I would to thank Prof. Kurt Gothelf and Dr Lise Refstrup Linnebjerg Pedersen for coordinating and managing the EScoDNA program. Also, I would like to thank all the other PI's and student members of the program who made all the EScoDNA meetings and visits a memorable experience. A special mention for Erika Bischofs and Susanne Kinzel for their administrative assistance and Helene Budjarek for her technical support during my stay in the lab.

I would also thank Prof. Björn Högberg from Karolinska Institute and his group members for making me feel welcomed and comfortable during my visit to their lab for a secondment in 2015.

I would like to thank Prof. Ebbe Andersen from Aarhus University and his group members: Steffen Sparvath, Guido Grossi and Denis Selnihhin for their guidance and support when I was starting out with RNA work.

I thank my former as well as present labmates for their stimulating discussions and for all the fun we have had during my time in the lab. Also I thank my friends Swati, Daniela and Lukas Oesinghaus for their constructive discussions, proofreading my thesis and being a major source of support. I would like to thank my friends Garima, Prerna, Nishit, Aurore, Anna, Ali, Enzo, Günther, Matthaeus, Mario, Jonathan, Sandra and Vera for always being there for me.

I would also like to thank my friends Swati, Padma, Vishal, Sushe, Atul and Shreyas for their support and making Munich feel like home.

Last but not the least, I would like to thank my family: my parents, my grandmother (Nani), my aunt (Massi), my uncle and my brother (Arjun) for their constant support and encouragement throughout my PhD and while writing this thesis. A special thanks to my uncle for proofreading my thesis.

