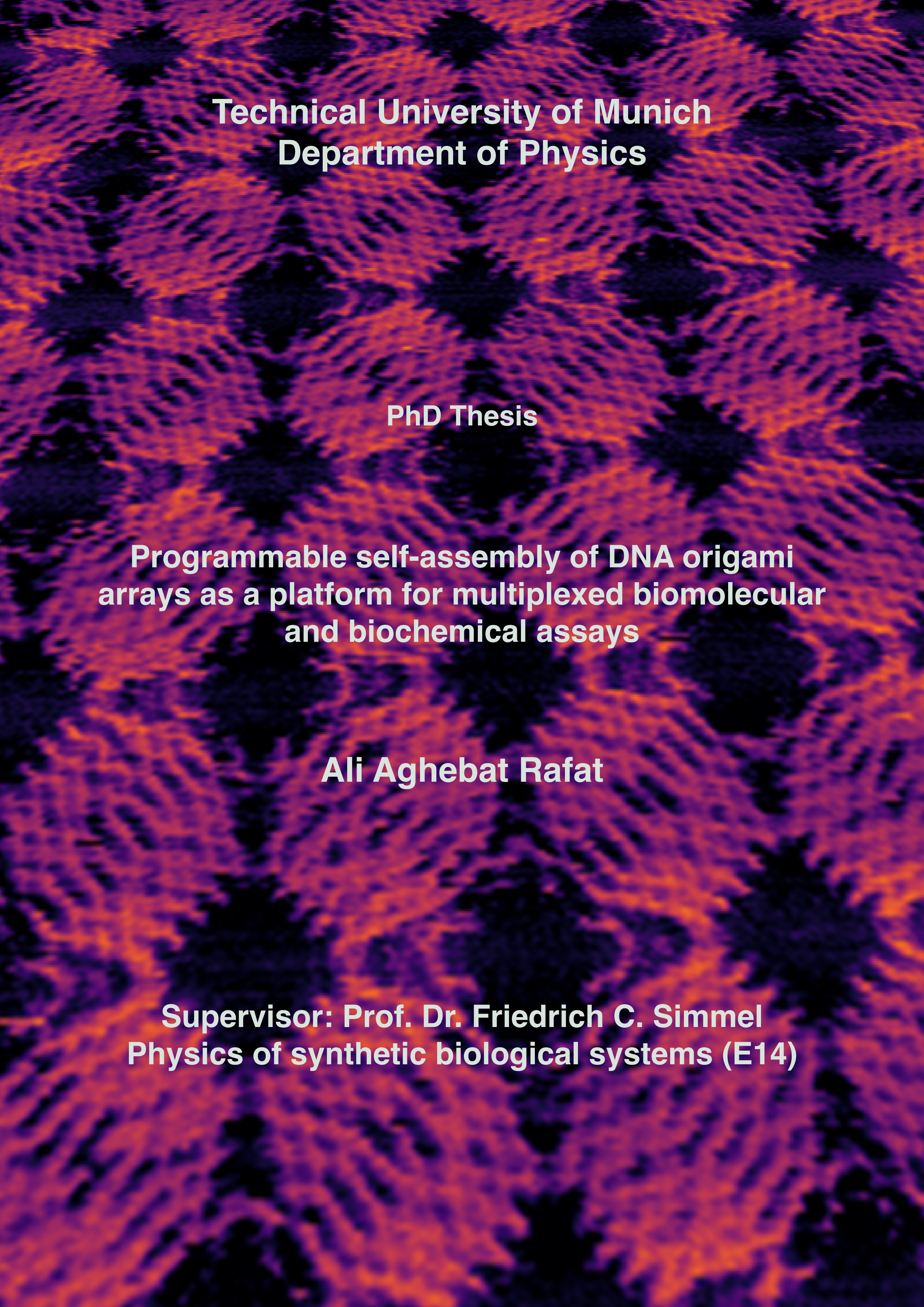


Technical University of Munich
Department of Physics

PhD Thesis

Programmable self-assembly of DNA origami arrays as a platform for multiplexed biomolecular and biochemical assays

Ali Aghebat Rafat

Supervisor: Prof. Dr. Friedrich C. Simmel
Physics of synthetic biological systems (E14)

Physics Department
Technical University of Munich

Physics of Synthetic Biological Systems E14

Programmable self-assembly of DNA origami arrays as a platform for multiplexed biomolecular and biochemical assays

Ali Aghebat Rafat

Vollständiger Abdruck der von der Fakultät für Physik
der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigten Dissertation.

Vorsitzende(r): Prof. Dr. Martin Zacharias

Prüfer der Dissertation:
1. Prof. Dr. Friedrich C. Simmel
2. Prof. Dr. Tim Liedl

Die Dissertation wurde am 07.09.2020
bei der Technischen Universität München eingereicht
und durch die Fakultät für Physik am 14.10.2020 angenommen.

Contents

List of Figures

Abstract

Bottom-up fabrication which employs the intrinsic properties of atoms and molecules to direct their self-organization is widely utilized to make nanostructures. The self-assembly of DNA using Watson-Crick base pairing has provided an attractive route toward this goal. A simple method, the so-called, "DNA origami technique" has facilitated self-assembly of long, single-stranded DNA molecules into relatively complex and arbitrary shapes. Sequence programmability of such molecules enables controlling their geometric and chemical properties with 5 nm precision. DNA nanostructures can be programmed to bind to a certain protein or a nanoparticle. Simultaneous binding of molecules by multiple binding partners is known to strongly reduce the apparent dissociation constant of the corresponding molecular complexes, and can be used to achieve strong, non-covalent molecular interactions. Based on these principles, efficient binding of proteins to DNA nanostructures has been achieved previously by placing several aptamers in close proximity to each other onto DNA scaffolds. Here, an approach is developed for exploring design parameters, such as the geometric arrangement or the mechanical properties of the binding sites, that use two-dimensional DNA origami-based nanocavities that bear aptamers with known mechanical properties at defined distances and orientations. The origami structures are labelled with barcodes, which enables large numbers of binding cavities to be investigated in parallel and under identical conditions, and facilitates a direct and reliable quantitative comparison of their binding yields. It is demonstrated that binding geometry and mechanical properties have a dramatic effect on origami-based multivalent binding sites, and that optimization of linker spacings and flexibilities can improve the effective binding strength of the sites substantially. Programmability of DNA origami structures is also used to design two rhombi that follow geometric properties of the Penrose tiles. Penrose-like domains are created by applying the so-called "Penrose connection rules" at the edges of these DNA origami structures.

Zusammenfassung

Die bottom-up Fertigung, bei der die intrinsischen Eigenschaften von Atomen und Molekülen genutzt werden, um ihre Selbstorganisation zu steuern, ist bei der Herstellung von Nanostrukturen weit verbreitet. Die Selbstorganisation von DNA mit Hilfe der Watson-Crick Basenpaarung hat einen attraktiven Weg zu diesem Ziel eröffnet. Eine einfache Methode, die so genannte "DNA Origami Technik", hat die Selbstorganisation von langen, einzelsträngigen DNA Molekülen zu relativ komplexen und willkürlichen Strukturen erleichtert. Die Sequenzprogrammierbarkeit solcher Moleküle ermöglicht die Kontrolle ihrer geometrischen und chemischen Eigenschaften mit einer Präzision von 5 nm. DNA Nanostrukturen können so programmiert werden, dass sie an ein bestimmtes Protein oder einen Nanopartikel binden. Es ist bekannt, dass die gleichzeitige Bindung von Molekülen durch mehrere Bindungspartner die Dissoziationskonstante der entsprechenden Molekülkomplexe stark herabsetzt, und kann zur Erzielung starker, nichtkovalenter molekularer Wechselwirkungen genutzt werden kann. Basierend auf diesen Prinzipien wurde eine effiziente Bindung von Proteinen an DNA Nanostrukturen bisher dadurch erreicht, dass mehrere Aptamere in unmittelbarer Nähe zueinander auf DNA Gerüsten platziert wurden. Hier wird ein Ansatz zur Erforschung von Designparametern wie der geometrischen Anordnung oder der mechanischen Eigenschaften der Bindungsstellen entwickelt, bei dem zweidimensionale DNA Origamibasierte Nanokavitäten verwendet werden, die Aptamere mit bekannten mechanischen Eigenschaften in definierten Abständen und Orientierungen tragen. Die Origami Strukturen sind mit Barcodes gekennzeichnet, wodurch eine große Anzahl von Bindungskavitäten parallel und unter identischen Bedingungen untersucht werden kann, und so ein direkter und zuverlässiger quantitativer Vergleich ihrer Bindungsausbeuten möglich ist. Es wird gezeigt, dass die Bindungsgeometrie und die mechanischen Eigenschaften einen großen Einfluss auf die multivalenten Bindungsstellen auf Origami-Basis haben und dass eine Optimierung der Linkerabstände und flexibilitäten die effektive Bindungsstärke der Stellen erheblich verbessern kann. Die Programmierbarkeit von DNA Origami Strukturen wird auch dazu verwendet, zwei Rhomben zu entwerfen, die den geometrischen Eigenschaften der Penrose-Kacheln folgen. Penrose-ähnliche Domänen werden durch Anwendung der sogenannten "Penrose-Verbindungsregeln" an den ändern dieser DNA Origami Strukturen erzeugt.

1 Introduction

DNA nanotechnology uses sequence-programmable selfassembly of DNA molecules to create nanoscale structures and devices [1]. In particular, the DNA origami technique [2, 3] has enabled the formation of almost arbitrarily shaped molecular assemblies, which can be decorated with functional molecules or nanoparticles with nanoscale precision. Origami structures are made of hundreds of short oligonucleotides or “staple strands” that are hybridised to a long single-stranded “scaffold” strand in a specific manner, which leads to the folding of the scaffold into the desired shape. Each of the staples can, in principle, be further chemically modified or extended, and can therefore be regarded as pixels (or voxels, in three dimensions) of the resulting structure with a size of $\sim$5 nm.

A wide variety of potential applications for origami structures have already been explored. The ability to place molecules or nanoparticles at precise distances and into specific geometries enables the biophysical characterization of biomolecular interactions under well-controlled conditions [4, 5, 6] or the creation of tailored nanophotonic and plasmonic systems composed of optically active components [7, 8]. In particular, origami structures that interact with proteins show great promise for applications in biosensing, theranostics, and biomedical nanorobotics [3, 9, 10].

Functionalization of origami structures has been achieved by covalent modification of staple strands with, for example, fluorophores, biotin [11], or DNA-binding proteins such as zinc finger proteins [12] or relaxases [13]. Alternatively, peptides, proteins, or nanoparticles of interest have been covalently attached to DNA linker strands [5, 14, 15, 16, 17], which were then hybridized to complementary staple strand extensions on the origami structures. An attractive approach for coupling DNA origami structures with functional protein components is based on the use of DNA-antibody conjugates [9, 18] or protein tags such as SNAP-tag or HaloTag [14, 19]. A different approach that does not require any covalent modification is based on DNA aptamers [20], which provides a “natural” link between DNA and protein nanotechnology.

However, typical dissociation constants for aptamer–protein interactions are on the order of $K_d = 1\text{–}100\ nM$, which is often not strong enough for specific applications. An obvious strategy to improve aptamer-mediated binding of proteins to origami or other DNA structures is to place several aptamers in close proximity to each other onto the scaffold, and thereby increase the “local concentration” of binding sites and use multivalency effects [21, 22]. For instance, in a systematic study of the influence of the distance between two thrombin-binding aptamers (that were placed on top of a flat origami sheet) on the binding of the protein [20], an estimated fifty-fold improvement in K_d was found compared to the K_d obtained with a single aptamer. The improvement in binding strength through multivalent binding to aptamers arranged on DNA origami scaffolds [20, 23, 24] or on tile-based DNA or RNA nanostructures [20, 25, 26] has been exploited in various contexts already, for instance in anti-coagulant applications [25, 27].

In previous work, the distance between two aptamers had been controlled primarily with relatively simple DNA scaffolds. By contrast, DNA nanostructures in principle permit the spatial modelling of more complex artificial binding sites, in which the geometry of the molecular interactions is controlled precisely. Moreover, the mechanical properties of DNA, in which single-stranded molecules are considerably more flexible than double-stranded ones, can be tuned to a certain degree. Apart from tuning the geometry of the binding site it is thus also possible to control its flexibility, which is expected to have a profound influence on the overall binding strength [6, 21, 22, 28].

In this work, to study these capabilities more systematically and reliably, origami structures with a central cavity are designed. Within the central cavity several aptamers can be placed at multiple locations in different geometric relationships, and with variable flexibilities. Using well-known aptamers for thrombin and streptavidin as examples, it is shown that these parameters affect the binding of proteins to the cavities dramatically, and that their optimisation results in unprecedented binding efficiencies.

In the first chapter of this work, some theoretical background knowledge on biophysics and biochemistry will be provided which enables better understanding of the thesis. Then the details on the experimental materials and methods, applied during the projects will be discussed. Eventually obtained results will be presented and discussed and followed with a conclusion.

2 Biophysical and biochemical background

In this chapter, some theoretical and biochemical knowledge for better understanding of this thesis is going to be provided. At first diffusion and Brownian motion will be explained. Following there will be a discussion on DNA, its charge and some of its secondary structures. Then a theoretical description for DNA-mica interaction will be presented. Afterwards, DNA origami method and the 2D assembly of the DNA origami structures will be demonstrated. In addition, tessellation and one specific quasicrystalline version of it, the so-called "Penrose tessellation" is going to be clarified. Eventually, there will be a brief discussion on avidity and multivalent binding.

2.1 Diffusion and Brownian motion

In 1828, a botanist called Robert Brown noticed that pollen grains suspended in water have a specific motion visible with his microscope. These pollen grains which are about 1 μm in diameter were large enough to be resolved by Brown's microscope. He naturally assumed that the motion was a sort of life process but he decided to check his assumption and figured out that the motion never stops even though the grains were kept for a long time in a sealed container. He also tried some totally lifeless particles with the similar size in water and still observed the motion at the same temperature. This random motion (random walk) is due to the collision of these particles with the fast-moving molecules in water. In 1905, Einstein presented his theoretical description for Brownian motion in one of his papers.
According to his theory, the mean-square displacement in a one-dimensional random walk increases linearly in time [29]:

$$<{X_N}^2> = 2Dt \tag{1}$$

Equation 1 is called the **one dimensional diffusion law** in which D is the diffusion constant, t is the total time and N is the number of steps in the random walk. The **diffusion constant** in such a random walk can be defined as following:

$$D = \frac{L^2}{2\triangle t} \tag{2}$$

where L is the step length and $\triangle t$ is the time between steps. The result for one-dimensional random walk can be extended to two or three dimensions:

$$<{R_N}^2> = 4Dt \tag{3}$$

will be the equation for diffusion in two dimensions and

$$<{R_N}^2> = 6Dt \tag{4}$$

the diffusion in three dimensions.

2.1.1 Friction's relation to diffusion, Einstein equation

The same collisions which cause the diffusion of particles in a fluid are also responsible for the friction (drag force). To study this friction, it will be assumed that there is a constant force f applied on a single particle that only moves in one dimension, in the time $\triangle t$ between each collision. Considering v_0 to be the particle velocity after a collision, the uniformly accelerated motion of the particle is:

$$\triangle x = v_0 \triangle t + \frac{f \triangle t^2}{2m} \tag{5}$$

Assuming that each collision will erase the memory of previous step, an average over v_0 will be equal to zero and the equation can be rewritten 5 as following:

$$<\triangle x> = \frac{f \triangle t^2}{2m} \tag{6}$$

and the net drift velocity will be:

$$v_{drift} = \frac{f}{\zeta} \tag{7}$$

where:

$$\zeta = \frac{2m}{\triangle t} \tag{8}$$

ζ is the viscous friction coefficient which is experimentally measurable like diffusion constant and is given by Stokes formula:

$$\zeta = 6\pi\eta a \tag{9}$$

where a is the radius of the particle and η is the viscosity of the fluid. Since both ζ and D are experimentally measurable, one can find $\triangle t$ (the duration of each random walk step) and L (the step length) using equations 8 and 2. Einstein noticed a third relation involving $\triangle t$ and L which comes from the ideal gas law. Considering the following:

$$\frac{L^2}{\triangle t^2} = {v_0}^2 \tag{10}$$

and the relation between the average value of v_0^2 in one dimension and K_BT from the ideal gas law is:

$$K_BT = m{<}{v_0}^2{>} \tag{11}$$

the famous **Einstein relation** is obtained which was the result of his work in 1905 [29]:

$$\zeta D = K_BT \tag{12}$$

where K_B is the Boltzmann constant and T is the temperature. According to this relation, the fluctuations in a particle position are linked to its frictional drag force. This quantitative connection is always equal to K_BT regardless of what sort of particle we study. For example a larger particle would face a higher drag force in fluid (ζ) but will diffuse less (D) in comparison to a smaller particle, therefore the outcome will always be the same. Combining the Einstein relation (equation 12) with the Stokes formula (equation 42) the so-called “**Stokes-Einstein relation**” can be written as:

$$D = \frac{RT}{6\pi\eta aN} \tag{13}$$

which relates the diffusion constant with the particle size. In the equation above R and N are the gas and Avogadro constant respectively. Employing this relation

Einstein was able to calculate the size of atoms and molecules plus the number of atoms in a mole of a gas using the measurable value of diffusion constant.

2.2 DNA

Deoxyribonucleic acid (DNA) is used as an information storage which encodes genetic information to produce proteins and RNA, employing its four bases, namely; Adenine (A), thymidine (T), Cytosine (C) and Guanine (G) (see figure 1A)
DNA is a polymer made of nucleotides as its monomers, which are sugar molecules connected to a phosphate, known as 2'-deoxyribose-5'-phosphate to which a base is attached (see figure 1B). In order to make the polymer, these nucleotides are connected by the 3'-hydroxyl of one 2'-deoxyribose via the phosphate (phosphodiester bonds) attached to the 5'-hydroxyl of another nucleotide 2'-deoxyribose [30] (see figure 1B). A single stranded DNA has an intrinsic orientation polarity due to the presence of phosphate and hydroxyl groups in its two ends. This orientation is usually shown by an arrow which starts from 5' hydroxyl group (connected to a phosphate group) and end to 3' hydroxyl group.(see figure 1B).

2.2.1 DNA charge

DNA is often considered to be a negatively charged polyelectrolyte because of deprotonation of the phosphate groups in DNA backbone under physiological conditions. An acidic molecule like DNA is neutral by itself, however, putting it in the water in presence of thermal motion ($K_B T$), causes some of the loosely attached atoms of DNA wander away and leave some electrons behind [31]. Therefore, the remaining molecule will have a net negative charge.
In fact without the thermal motion, DNA atoms would stay intact and the net DNA charge would be neutral. The positively charged atoms which at the first place were not covalently bound to DNA, will travel away and create a so-called counterion layer of a couple of nanometers around DNA. On one hand, the counterions want to travel away from DNA molecule to gain entropy, on the other hand, they would need lots of energy to completely pull away from home. In the end, the compromise between these to forces will specify the thickness of the counterion layer around DNA.

2.2.2 Single-stranded DNA

Single stranded DNA (ssDNA) is a flexible polymer which is typically described by the freely-jointed chain model. For the freely-jointed chain model, backbone bind angles are independent from each other and the polymer can form a random coil. In the ssDNA, nucleotides are freely connected and are able to have any orientations with respect to each other. Spacing between two nucleotides has a length of roughly 5-7 Å [32, 33].

The persistence length of ssDNA depends on the salt concentration and in high concentration of salt, can be approximated by the contour length of DNA [34] (table 1). Using the freely-jointed chain model, the mean square end to end distance of a ssDNA R with N nucleotides each having length l, can be calculated as following [35]:

$$<\overrightarrow{R}^2> = Nl^2 \tag{14}$$

and the radius of gyration for a ssDNA can be calculated with [35]:

$$R_G = \sqrt{\frac{N}{6}}l \tag{15}$$

In correlation with this model, pulling or pushing on a ssDNA would face a nonlinear entropic force and the ssDNA would behave like an entropic spring. There is an analytical approximation by Marko and Siggia for this entropic force which uses the worm-like chain model [36]:

$$F_{ssDNA}(x) = \frac{K_B T}{P}(\frac{x}{L} + \frac{1}{4(1-\frac{x}{L})^2} - \frac{1}{4}) \tag{16}$$

where K_B is the Boltzmann constant, T is the temperature, P is the persistence length and L=Nl is the contour length. This entropic force is an important parameter in DNA nanotechnology when ssDNA is used as a linker between biomolecules.

2.2.3 Double-Stranded DNA

Double-Stranded DNA (dsDNA) is made by hybridization of two ssDNA molecules which are complementary and anti-parallel. According to so-called Watson-Crick

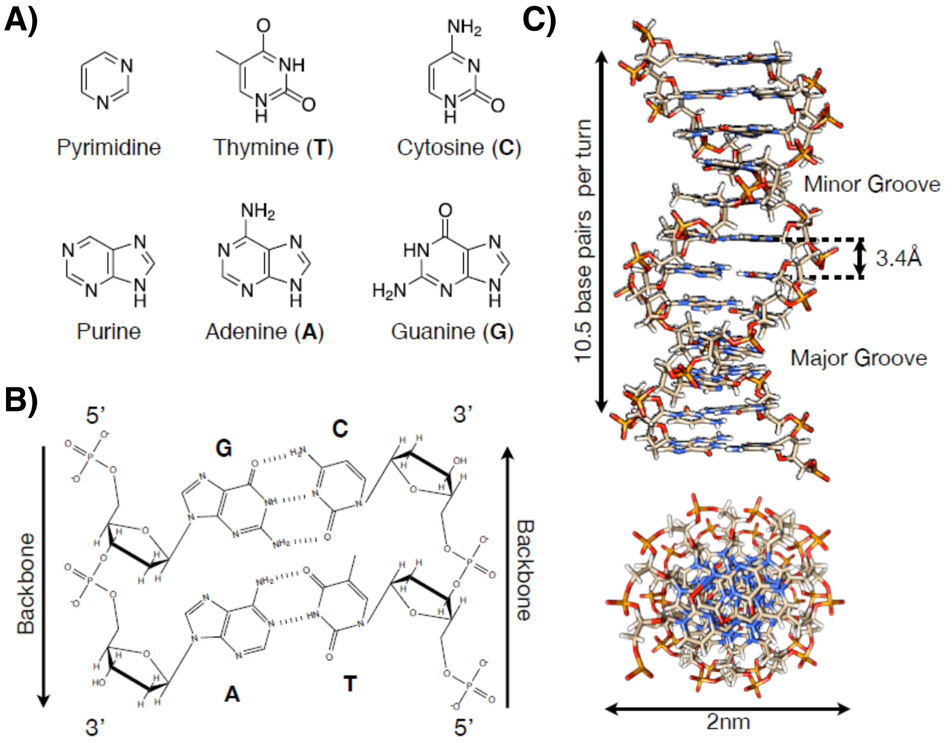


Figure 1: A) DNA bases, adenine (A), cytosine (C), guanine (G) and thymidine (T) and the purine and pyrimidine groups. B) Watson-Crick base-pairing of cytosine with guanine and thymidine with adenine on a dsDNA backbone. Phosphate connects the backbone in dsDNA. Polarity of DNA is from 5' (phosphate group) to 3' (hydroxyl group). C) Structure of B-form dsDNA with major and minor groves [37]. Number of dsDNA bases per turn are 10.5 bases and the diameter for a dsDNA equals to roughly 2 nm. Length of one nucleotide is approximately 3.4 Å. Atoms color codes: gray = C, white = H, red = O, orange = P and blue = N. Figure is taken from [38]

.

base paring, each thymidine (T) base can only connect to adenine (A) base and each cytosine (C) base can only connect to guanine (G) all by means of hydrogen bonds (see figure 1B). Combination of these hydrogen bonds plus stacking interaction between purine and pyrimidine (see figure 1A) groups in bases will form well known B-DNA (see figure 1C) which is a right-handed double helical molecule. The double helical structure of B-DNA was first determined by James D. Watson and Francis H. C. Crick in 1953 [39] with the help of an X-Ray diffraction pattern by Rosalind E. Franklin and Raymond G. Gosling [40]. In contrary to the general view that hydrogen bonds are mainly the responsible part to form and stabilizing double helical structure of DNA, it is stacking interaction [41] which makes a significant contribution to create the conformation of DNA. In fact, hydrogen bonds are aligning two strands of DNA next to each other to enable base stacking interaction.

Table 1: Comparison between ssDNA and dsDNA [38].

	ssDNA	dsDNA (B-form)
Spacing between two nucleotides	5-7 Å per nucleotide	3.4 Å per nucleotide
persistence length	0.5-5 nm [34]	40-50 nm [42]
diameter	1 nm	2 nm
shape	random coil	right-handed double helix, with major and minor groove
model	freely-jointed chain, worm-like chain	worm-like chain
Helical turn per nucleotide	nearly arbitrary	34.3°

Table 2: Polymer properties of different dsDNA forms [43]

	A-DNA	B-DNA	Z-DNA
Left or right handed	right handed	right handed	left handed
Base pairs per turn	11	10.5	12
Diameter	2.3	2	1.8
Rise per base (nm)	0.26	0.34	0.37

For standard B-form DNA, spacing between two nucleotides is considered to be about 3.4 Å, and the helical turn per nucleotide is approximately 34.3° [39, 40]. With respect to the angle for each nucleotide, number of bases for each turn in right-handed B-form DNA will be 10.5 base or 35.7 Å in length [44], The diameter of a DNA duplex in this conformation is roughly 2 nm (see figure 1C). Since there is a 120° angle between backbones complementary nucleotides, there are major and minor grooves in DNA structure [38] (see figure 1C).

There are other forms of dsDNA such as right-handed A-form DNA [45] or left-handed Z-form DNA [46]. Table 2 compares the properties of different dsDNA forms. In the projects which are going to be discussed in this thesis, only the B-form DNA was used.

Table 1 compares polymer properties of ssDNA and dsDNA. For dsDNA the spacing between two base pairs along the helix axis is approximately 3.4 Å and its salt dependent persistence length is around 40-50 nm. This means that dsDNA is a rigid molecule compared to ssDNA and a different model must be used to describe its properties. In fact, the rigidity of dsDNA is enabling the formation of almost arbitrary-shaped macromolecules using DNA self-assembly. Therefore, in this case, worm-like chain model [47] is used to calculate end-to-end distance and other features of dsDNA. The mean square end to end distance of dsDNA using the worm-like chain model is calculated as following [48, 42]:

$$<\overrightarrow{R}^2> = 2l_pL[1 - \frac{l_p}{L}(1 - \exp\frac{-L}{l_p})] \tag{17}$$

where l_p is the persistence length and L is the contour length. For a long dsDNA where $L \gg l_p$ the equation 17 can be written as:

$$<\overrightarrow{R}^2> = 2l_pL \tag{18}$$

2.2.4 Thermodynamics of Double-Stranded DNA

For double-stranded DNA, it is possible to calculate the hybridisation energies using the standard Gibbs free energy. Gibbs free energy is a thermodynamic potential which measures the maximum work obtainable from a thermodynamic system in constant pressure and temperature. This free energy is a combination of two parts, one is the number of possible states in the system and the other one is enthalpy of system or total energy of the system with SI units of J/mol.(see equation 19).

$$\triangle G = \triangle H - T\triangle S \tag{19}$$

The possibility to predict DNA double-helix structure formation has many biological and technological applications. For example, it enables us to predict the secondary structures of ssDNA which can be the cause for a disease [49] or carry a biological function [50, 51]. Moreover, it is important for some biotechnological applications like PCR, to design primer sequences with low secondary structures [52, 53].

There are four factors which dominate DNA hybridisation reactions, hydrogen bonds between the bases of complementary strands, stacking interactions between purine and pyrimidine groups of the bases, electrostatic repulsion of negatively-charged DNA backbones and the entropic effects. Thermodynamic properties of DNA strand hybridisations can be calculated employing nearest neighbour models [54, 55, 56]. The standard Gibbs free energy (Gibbs free energy for one mol) can be defined as a combination of base-pairing energy and base-stacking energy plus entropy of dsDNA for a constant temperature. According to a work by J. SantaLucia and D. Hicks [56], which has used so-called nearest-neighbor parameters for entropy, enthalpy and free energy, the lowest free energy is for Cytosine and Guanine base-pairing which has three hydrogen bonds compared to Adenine and thymidine base-pairing with two hydrogen bonds. This is the reason why usually the stability of DNA duplex is determined by number of GC pairs.

In general, it is not easy to separate the influence of base-pairing and base-stacking in DNA hybridization; however, there is a set of experimental data for base-stacking energies of different base pairs done by E. Protozanova et al. [41]. These data can be seen in the table 3.

Table 3: stacking free energies of different base pairs $\triangle G$ $[kcal/mol]$ [41]

	A	T	G	C
A	-1.11	-1.34	-1.06	-1.81
T	-0.19	-1.11	-0.55	-1.43
G	-1.43	-1.81	-1.44	-2.17
C	-0.55	-1.06	-0.91	-1.44

All these values are measured at the nick site of the DNA fragment. Figure 2A, demonstrates two state for DNA conformation. The state which is seen in the left is the stacked conformation of the DNA in the nick site. The loss of stacking between the base-pairs in the nick site induces a kink in DNA duplex and transition to the unstacked conformation (see figure 2A, right). Values in the table represent energy difference between these two states.

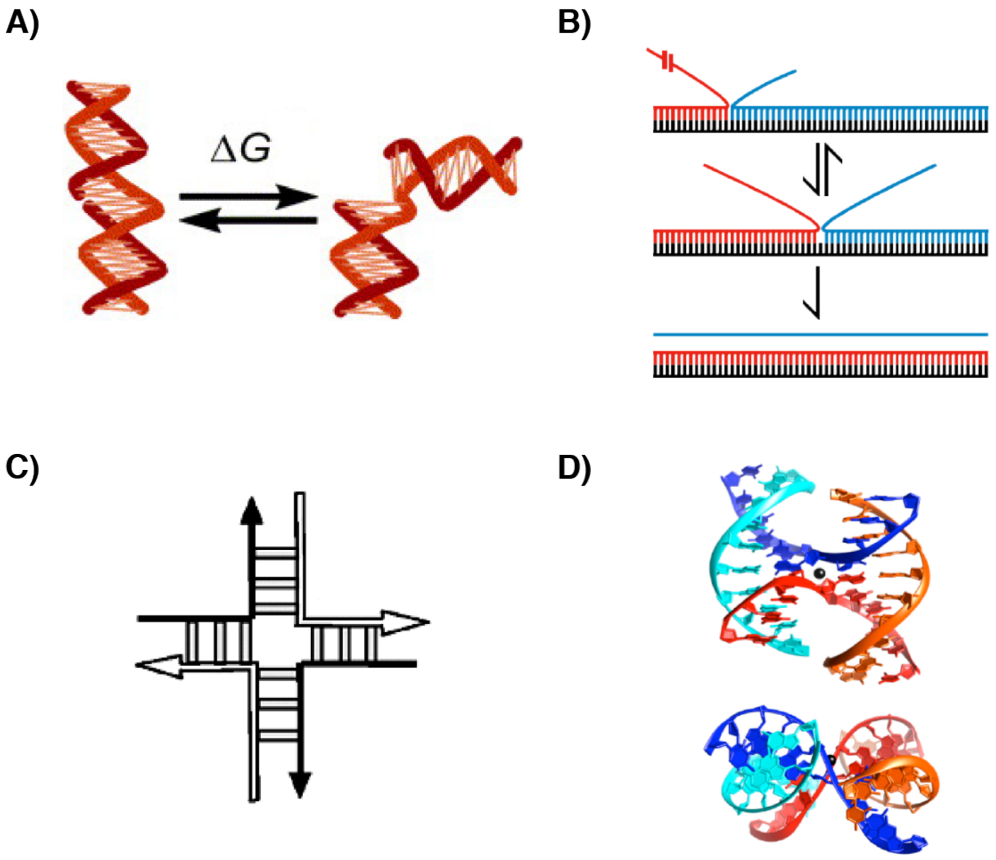


Figure 2: A) Schematic representation of the stacked to unstacked conformational transition of the DNA at the nick site which has the energy difference of $\triangle$G [41]. B) Strand displacement scheme. Red ssDNA is the complementary strand to the black one and is partially hybridized to it. Non-complementary part of black ssDNA to the blue ssDNA is the toehold. After hybridisation of the red strand to the toehold, the blue strand will be replaced by the red one following a random walk process. This reaction is reversible unless the red ssDNA has completely displaced the blue ssDNA [38]. C) Schematic design for parallel (or unstacked) conformation of Holliday junction [57]. D) Crystal structure of a DNA Holliday junction in its anti-parallel (or stacked) conformation in the presence of Mg^{2+} ions (black dot in the middle of the structure) [38].

2.2.5 Strand displacement reactions and Toehold

Branch migration or strand displacement happens when a ssDNA partially hybridises to its complement in the presence of another ssDNA (blue strand) which is connected to the other part of the complementary ssDNA (black strand) (see figure

2B). In figure 2B, the non-hybridised part of black ssDNA which establishes binding to the red ssDNA is so-called toehold [58, 59]. In a toehold strand displacement reaction, the replacing strand hybridises to the toehold region which results in its high local concentration. After the hybridisation is initiated, the replacing strand competes with the initial strand and removes it following a random walk. Since the replacing strand has a longer complementary region, it is thermodynamically more stable than the initial strand. Such reactions are frequently used to release and exchange DNA strands in DNA nanotechnology.
In DNA hybridisation reactions the on-rate of the reaction is mainly related to the diffusion or the strands concentration and the off-rate depends on the hybridisation energies. These on-rates and off-rates are very crucial for the speed of dynamic processes in DNA nanotechnology. For toehold-mediated strand displacement the toehold length has a key role in the reaction speed. In the work done by Zhang et al. [59] it is shown that up to a critical toehold length, 6-8 nucleotides, there is a strong dependency of the reaction speed on the toehold length; however, a toehold longer than that do not further increase the speed.

2.3 DNA secondary structures

Many different DNA conformations have been identified [60] which are often referred to as DNA secondary structures. DNA strands can have secondary structures by either hybridising to themselves to form simple structures or hybridising to other strands to create more complex structures. These so-called Non-B DNA structures are considered as genomic elements which have different roles in cell [61]. Among those are gene function and regulation [62], immune response [63], telomere maintenance [64], recombination [65], antigenic variation in human pathogens [66], and the generation of genomic diversity [67]. Here few examples of these secondary structures that are related to the projects in this thesis are discussed.

2.3.1 Holliday Junction

The holliday junction is a well known secondary structure of DNA in which four ssDNA are connected by Watson-Crick base-pairing (see figure 2C). The conformation of holliday junction is strongly salt dependent [57]. For example, by changing the concentration of Mg^{2+} ions in solution, it can change from parallel conformation to anti-parallel conformation (see figure 2D). In the assembly of DNA nanostructures crossovers between two helices are also referred as holliday junctions. This is why presence of magnesium ions in the folding of nanostructures is important.

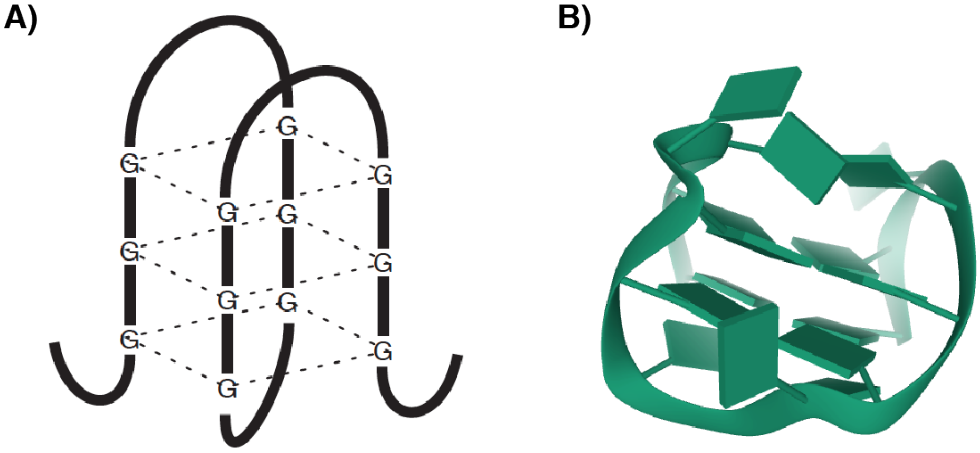


Figure 3: A) Schematic representation of the G-qadruplex [68]. B) NMR structure of thrombin binding DNA aptamer TBA taken from PDB `https://www.rcsb.org/3d-view/1RDE` [69].

2.3.2 G-quadruplex structure

Guanine-rich oligonucleotides typically show a strong tendency to form a supramolecular structure so-called G-quadruplex (see figure 3A). Because of its biological significance, the G-quadruplex structure is considered to be one of the most important secondary structures of DNA. For instance, it plays a crucial role in binding of DNA aptamers to their corresponding protein [70, 71].
There have been many ongoing structural studies on G-quadruplex motifs in the field of chemical biology of nucleic acids [72, 73, 74]. In the work by Arivazhagan Rajendran et al. is shown that presence of KCl is necessary for the formation and stability of the G-quadruplex structure [75].

2.3.3 DNA aptamers

Aptamers or sometimes referred to as chemical antibodies, are single stranded DNA or RNA molecules which are selected to specifically bind to a certain target (see figure 3B). Aptamers are selected through an iterative process so-called SELEX (selective evolution of ligand by exponential enrichment) [76, 77, 78]. The binding of aptamers to their target strongly depends on the folding of their 3D structure. Systematic studies show that the formation of the 3D structure is affected by the incubation conditions and buffer compositions [71]. Chemical antibodies or aptamers have a large number of potential applications in medicine and biotechnology. For example, with the current developments in cell-based and protein-based SELEX experiments, many aptamers have been identified which can target tumor cell-surface protein markers for both diagnostic and therapeutic applications [79].

2.4 DNA-Mica Interaction

Experiments by AFM show that DNA as a negatively charged polyelectrolyte (see section 2.2.1) can be strongly adsorbed on negatively charged mica in the presence of divalent cations. For two charged surfaces which have the same charge (e.g. DNA and mica), interaction between them is repulsive. The origin of this repulsive interaction is the so-called electrical double layer force [80]. This interaction comprises two parts, one part is related to the counterion cloud repulsion and the second part is coming from the thermal pressure. However, what is the origin of the attractive force between negatively charged DNA and mica?

This attractive originates in the correlations of counterions between DNA and mica. The important characteristics of this force are its short range and also its relevance to the competition between monovalent and divalent cations. another answer to the question could be a hydrophobic attractive force [81] between the mica surface and DNA polyelectrolyte. However, this force should not play a key role, because the attraction of DNA to mica is strongly dependent on the presence of divalent cations. In this section, first the repulsive interaction due to double layer electrical forces between DNA and mica will be discussed. As mentioned above, this force contains thermal pressure between DNA and mica as charged planes and also repulsive interaction between counterion clouds. In the second part, the origin of attraction between DNA and mica in the presence of divalent cations will be explained. This attraction comes from the correlations in counterion clouds for DNA and mica. Here, a simple model which considers DNA and mica as two lines of charges will be presented to explain this attractive force.

2.4.1 Double layer electrical forces between mica and DNA

Under certain conditions [82], DNA as well as mica can be modelled as a charged plane. In the presence of counterions, there will be a shell of counterions with the thickness of λ_z depending only on valency of counterions z and surface charge density σ.

$$\lambda_z = \frac{e}{4\pi\sigma l_b z} \tag{20}$$

In equation 20, e is the electron charge and l_b is the Bjerrum length:

$$l_b = \frac{e^2}{\varepsilon k_B T} \tag{21}$$

where ε is the dielectric constant, k_B the Boltzmann constant and T is the temperature.

In the presence of divalent cations, the pressure between mica and DNA surface with surface charge density of σ_a and σ_b respectively, can be approximated by using a Poisson-Boltzmann equation in one dimension:

$$\frac{d^2\phi(x)}{dx^2} + k^2 \exp(-\phi(x)) = \frac{l_b}{z} n(x) \tag{22}$$

In equation 22, k is a constant depending on boundary conditions, $\phi(x)$ is the normalized electrostatic potential $\phi(x){=}e\Psi(x)/k_B T$ and $n(x)$ is the external charge density. The boundary conditions for surface charge density σ_a at $x = 0$ and σ_b at $x = d$ are:

$$\left.\frac{d\phi(x)}{dx}\right|_{x=0} = -\frac{\sigma_a l_b}{ze} \ and \ \left.\frac{d\phi(x)}{dx}\right|_{x=d} = -\frac{\sigma_b l_b}{ze} \tag{23}$$

where d is the distance between mica and DNA. σ_a and σ_b are considered as the net charge, more specifically a sum over the adsorbed ion density and known surface charge density.

Then the pressure P(d) between two planes is given by equation 24 [83]:

$$P(d) = \frac{k_B T}{d l_b} \int_0^d dx((\frac{1}{2})(\frac{d\phi}{dx})^2 - \frac{d}{dx}(\frac{d\phi}{dx})) \tag{24}$$

The first term in the integral refers to thermal pressure and the second term is the electrostatic stress of the counterion clouds. Even for two planes which are

oppositely charged, thermal pressure can result in repulsion at a short range.
$P(d)$ is positive when the force is repulsive and is negative when it is attractive. After numerically solving this integral [82], it is seen that the ratio σ_a/σ_b is playing a key role on the resulting interaction. For the case in which DNA and mica are both negatively charged, this force is repulsive ($\sigma_a\sigma_b >> 0$) ;however, in case of pretreating the mica surface for example with $NiCl_2$, an inversion to surface charge density of mica can occur, which can cause attractive interaction in a long range (see figure 4A). The repulsive thermal force, which is always present for short range is due to entropy loss of counterion clouds and close to the surface is stronger than attraction between two planes.

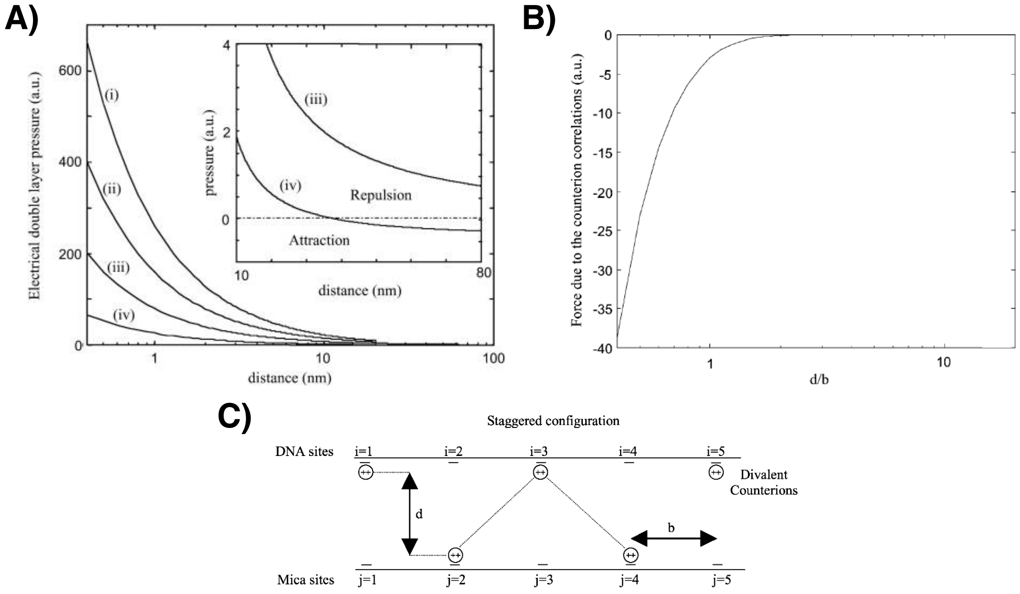


Figure 4: A) Electrical double-layer pressure between mica and DNA. (i), (ii), (iii) and (iv) are respectively for σ_a/σ_b =4, 2, 0.5 and -0.5. Repulsive interaction is weaker when mica is less charged and pretreated [82]. B) Attractive force between mica and DNA due to the correlation of a counterion cloud at $T = 0$. b is the distance between counterion sites and d is the distance between DNA and mica counterion layers [82]. C) Staggered configuration of counterions. i and j show the charge sites on DNA and mica, respectively [82].

2.4.2 Attraction between DNA and negatively charged mica

Like-charged surfaces such as mica and DNA in solution with counterions are able to attract each other through correlations in their shared counterion environment. To calculate this force, a simple model is used in which DNA and mica are considered as two lines of charges [84]. In addition, the only effect of counterion is a local renormalization of surface charge density and also counterions are considered as point-like charges [85]. With respect to these points, the Hamiltonian for unscreened electrostatic interaction between DNA and mica is [84]:

$$H = \frac{e^2}{2\varepsilon} \sum_i \sum_j \frac{(1 - z_i\phi_i)(1 - z_j\phi_j)}{\sqrt{x_{i,j}^2 + d^2}} \tag{25}$$

Where i is the label of mica sites and j for DNA sites. d is the distance between DNA and mica counterion layers, z_j is the valence of j^{th} ion and $\sqrt{x_{i,j}^2 + d^2}$ is the distance between j^{th} DNA site and i^{th} mica site. ϕ_i and ϕ_j are the occupation variables for the corresponding sites. For example, if j^{th} is empty $\phi_j = 0$ and if j^{th} is occupied, $\phi_j = 1$.

From the Hamiltonian, the force which is acting between DNA and mica can be calculated and is equal to:

$$F_c(d) = \frac{e^2 d}{\varepsilon} \sum_{i,j} \frac{(1 - z_i\phi_i)(1 - z_j\phi_j)}{(x_{i,j}^2 + d^2)^{3/2}} \tag{26}$$

After minimizing the free energy it turns out that a staggered configuration is favored for counterions on the surface [84] (see figure 4C). In this configuration, counterions have the furthest possible distance from each other. This seems reasonable due to the fact that repulsive interaction between charges will be minimized in this way. However, this is the case in which temperature $T = 0$. At non-zero temperatures and also high ionic strength, the staggered configuration and consequently the attraction between the two surfaces is disturbed [82].

At $T = 0$ (staggered configuration), considering the mean distance between charges in DNA and mica $b \cong 1\ nm$ [86] and assuming that only divalent cations participate in neutralization, $F_c(d)$ can be plotted as shown in figure 4B. It can be observed from the plot that the force is attractive and works for the short distances $d < b$. For a highly charged surface such as DNA, this interaction is expected to be short range, because b is very small.

2.4.3 Influence of competition between monovalent and divalent cations

In case of having monovalent cations, for example Na^+ in solution, there is a competition between monovalent and divalent cations to bridge between DNA and mica surface [82]. As can be seen from equation 25, for monovalent cations the terms $(1 - z_j\phi_j)$ and $(1 - z_i\phi_i)$ are always positive, which means they do not contribute to the attractive interaction. Therefore, the presence of monovalent cations in the buffer and their competition with divalent cations can weaken the adsorption of DNA molecules on the mica surface.

2.5 DNA Origami

Over the last three decades, the DNA molecule, which is typically employed to store genetic information in of living organisms, has found an ever-expanding contribution in creating materials on the nanoscale [87, 88, 89]. Particularly, the DNA origami technique, which was first proposed by Paul Rothemond in 2006, has enabled the formation of almost arbitrary 2D and 3D shapes (see figure 5 and 6) [2, 3]. This technique utilises a long single-stranded DNA (typically M13 phage genomic DNA), "the scaffold", [90] of around 7000 bases, which folds into a desired geometry using about 200 short ssDNA (staples).

Staples in DNA origami structures can be regarded as 5 nm pixels (or voxels in 3D). The possibility to extend or chemically modify these staples with a nanometer precision opens up many applications in different directions. Soon after the invention of DNA origami technique, a variety of DNA nanoconstructs with applications in nanomedicine [91, 92, 93, 94], biosensing [95, 96] and nanomaterials [97, 98, 7] were built.

2.5.1 2D assembly of DNA origami structures

DNA as a programmable material can be designed to self-assemble and create crystal-like patterns. These patterns are potential templates to form lattices of proteins [99] or nanoparticles, e.g. gold nanoparticles [100] which have possible applications in material science and biology. In addition, one can use these cryptal-like formations as stamps for molecular imprinting and transfer the patterns to other surfaces [101]. Moreover, it has been shown that one can create single-molecule nanoarrays of DNA origami structures using the lithography techniques [102, 103]. As a result, combining lithography as a top-down approach with the lattices of DNA origami structures as a bottom-up technique, could enable producing functionalized

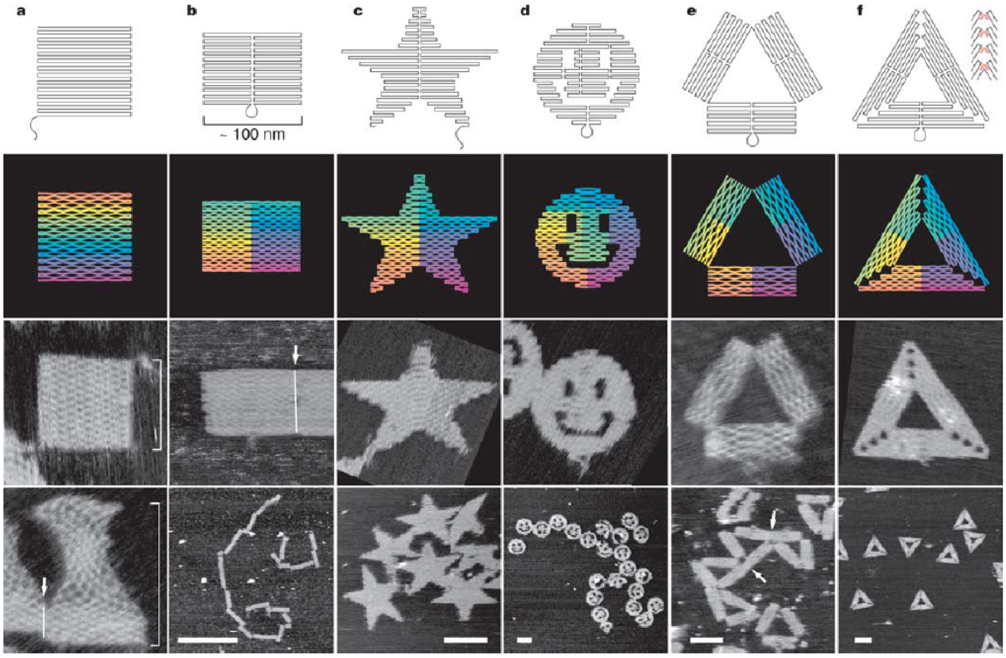


Figure 5: Schemes and AFM images of different 2D DNA origami structures. All images without scale bar are 165 nm by 165 nm and scale bars for lower images are b, 1 μm and c-f 100 nm. Copyright 2006, Springer Nature [2]. Figure have been taken with permission from the referenced papers.

DNA origami arrays on a micrometer with nm resolution.

There has been a variety of attempts to self-assemble DNA origami structures into lattices. These attempts are typically based on three different principles:

1) Base-pairing of unpaired nucleotides at the edge of origami structures

2) Base-stacking interactions at the edges and between blunt ends

3) Tuning the DNA-surface interaction to use the mobility of structures for formation of lattices

Different combinations of these parameters have resulted in micrometer-scale 2D lattices of DNA origami structures (see figure 7) [104, 105, 106, 99, 107].

2.5.2 Fractal assembly of DNA origami structures

Fractal assembly of DNA origami structures uses simple local assembly rules that are modified and applied recursively via a hierarchical multistage assembly process.

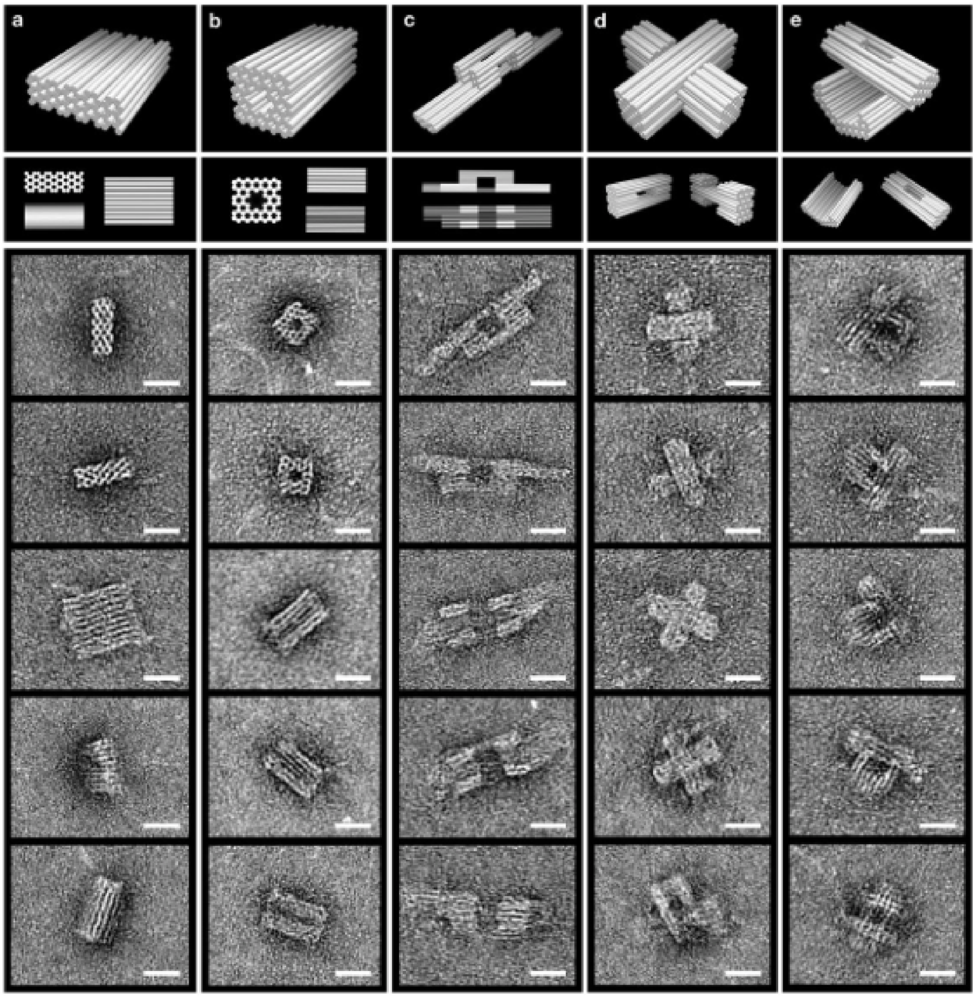


Figure 6: 3D models and TEM images of different 3D DNA origami structures, scale bars are 20 *nm*. Copyright 2009, Springer Nature [3]. Figure have been taken with permission from the referenced paper.

This technique creates DNA origami arrays with increasing size from single tiles (see figure 8 top) [108]. These single tiles can have their own unique modifications or features, since they are assembled separately. Employing this technique, DNA origami arrays with increasing numbers of uniquely addressable staples (or pixels) are formed that can be modified for different applications. Figure 8 shows Mona Lisa patterns, which have been created by extending several staples on the DNA origami arrays.

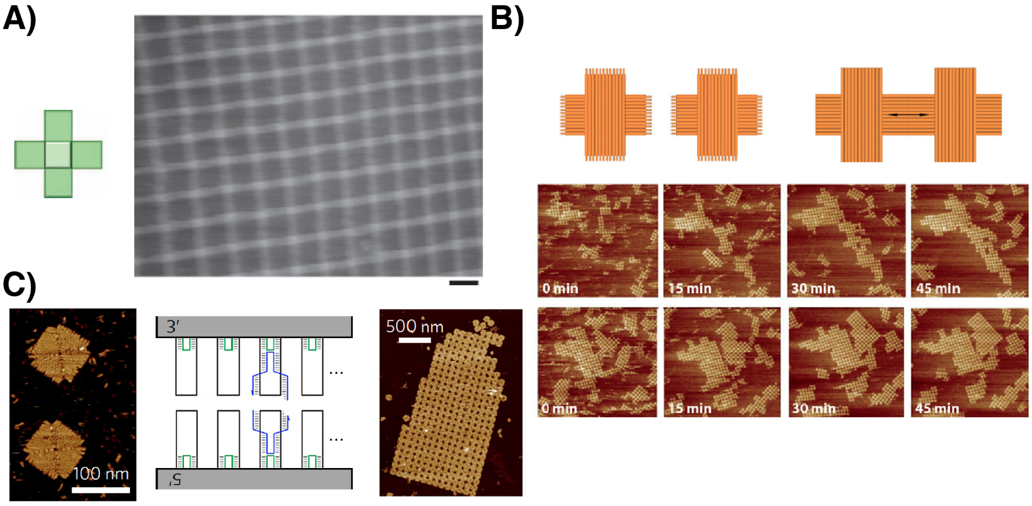


Figure 7: A) Lipid-bilayer-assisted self-assembly of cross-shaped DNA origami structures using base-stacking interaction. The scale bar is 100 nm [107]. B) 2D crystallisation of cross-shaped DNA origami tiles through time, by tuning DNA-mica interaction and base-stacking interaction between tiles. All AFM images are $3\mu m$ by $3\mu m$. Copyright 2014 WILEY-VCH Verlag GmbH and Co. KGaA, Weinheim [99]. C) 2D crystallisation of square DNA origami structure using combination of base-stacking and base-pairing as it is shown in the scheme. Copyright 2016, Springer Nature [106]. Figures have been taken by permission from referenced papers.

2.6 Tessellation

Tiling or tessellation is an interesting topic in mathematics, which deals with covering a whole surface or a plane, using one or more geometric shape, leaving no gaps or overlaps.

2.6.1 Periodic Tessellation

Periodic tessellation is a way of covering the whole surface by means of tiles with which at least there are two translational symmetry axes. It is possible to tile a surface periodically by only using triangles, tetragons or hexagons (figure 9A). However, using only Pentagons, it is not possible to cover a whole surface (figure 9B). There is another type of periodic tessellation in which uses more than one tile to cover the surface. This kind of tessellation is called Archimedean tiling [109]. For example, by using only octagon geometries, one would not be able to tile a sur-

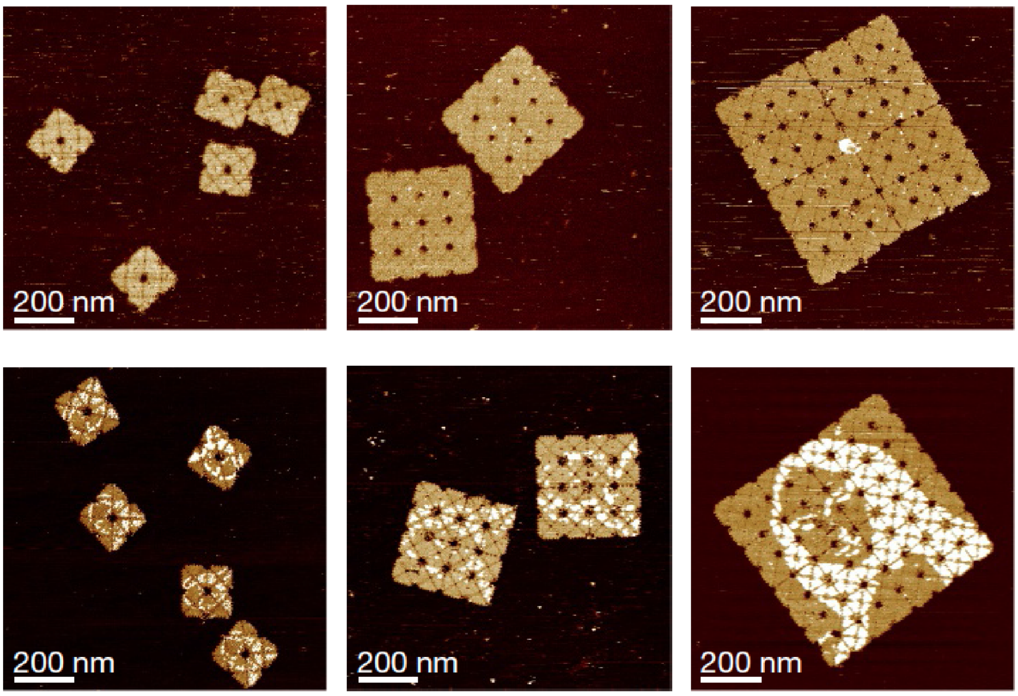


Figure 8: AFM images of plain arrays with and without a pattern. In these examples, the Mona Lisa pattern is used on the arrays. Figure is taken by permission from [108]. Copyright 2017, Springer Nature.

face completely. However, combining an octagon and a square enables a periodic tessellation (figure 9C).

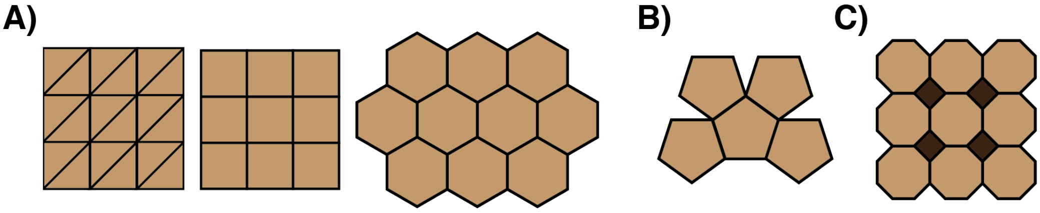


Figure 9: A) Schemes of triangles, squares and hexagons which periodically tile the plane. B) Pentagon scheme which shows that the tiling is impossible. C) An example of Archimedean tiling with octagons and squares.

2.6.2 Penrose Tessellation

As mentioned above, with Pentagons it is not possible to tessellate a surface. Further, it has been shown that it is impossible to tile a surface by means of only one shape with five-fold symmetry. In Penrose tessellation a surface is totally covered by use of an aperiodic set of prototiles which their combination also induce a five-fold symmetry in the surface pattern (see figure 10A). The name comes from the mathematician Roger Penrose who established this type of tessellation [110]. For Penrose tessellation, one possible combination of prototiles is two rhombi. These two rhombi have the same side length a, but different side angles. The rhombus with a sharp angle of 72° and an open angle of 108° is called wide rhombus, and another rhombus, which has a sharp angle of 36° and an open angle of 144°, is called slim rhombus (see figure 10B). The two rhombi must obey specific connection rules in order to produce a Penrose lattice. These connection rules are sketched in the figure 10B via the red and green marks on the sides. Two sides only connect when the colors of the marks and their distances from vertices match. Otherwise, the connection is not following the Penrose connection rules.

Φ the golden ratio is seen in the geometric construction of Penrose tiles. As demonstrated in equations 27 and 28, the ratio of sides and diagonals are equal to the golden ratio. b and c are the short and long diagonal lengths of the slim and wide rhombus, respectively (figure 10B).

$$\Phi = \frac{a}{b} = \frac{a}{2a\sin(18)} = \frac{1}{2}(1+\sqrt{5}) \tag{27}$$

For the slim rhombus and for the wide rhombus:

$$\Phi = \frac{c}{a} = \frac{2a\cos(36)}{a} = \frac{1}{2}(1+\sqrt{5}) \tag{28}$$

One more interesting characteristic of Penrose tiling is the convergence of the tile's ratio (ratio of the number of wide rhombi to the slim rhombi) to the golden ratio with ever larger patterns [111]:

$$\Phi = \lim_{N_{tiles}\to\infty} \frac{N_{wide}}{N_{slims}} = \frac{1}{2}(1+\sqrt{5}) \tag{29}$$

There are three main characteristics for Penrose tiling:

1) It is a non-periodic tiling with no translational symmetry.

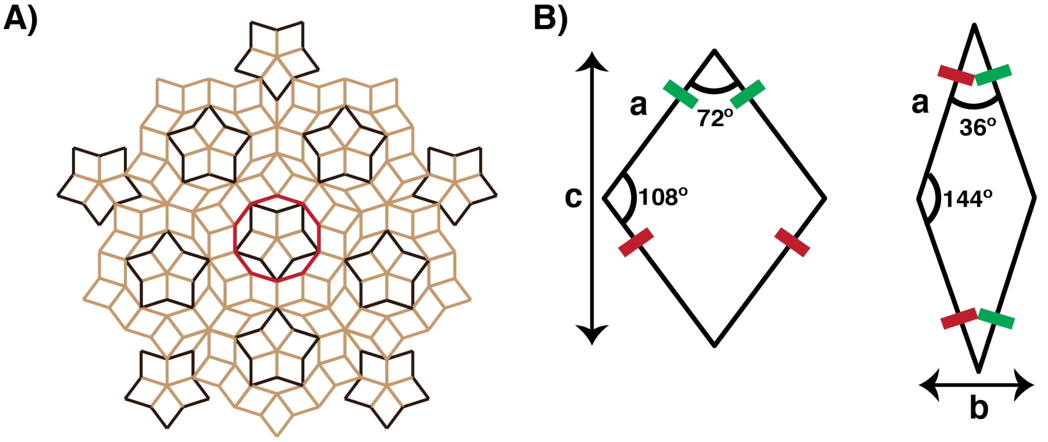


Figure 10: A) An example of Penrose tessellation. The five fold symmetry is clearly seen in the pattern and repeats itself at larger scales. Five-fold stars are highlighted in black to emphasise their repeating occurrence in the Penrose tiling. The red decagon in the middle represents the Penrose area which can be used for deflation-inflation of the tiles. B) Penrose tiles with the respective connection rules, represented by the red and green colored marks and their geometric properties [110, 112].

2) It is self-similar, i.e. the same pattern appears at larger and larger scales (see figure 10A).
3) It is a quasicrystal containing five-fold symmetry with a long range order [113]. Following Penrose matching rules, there are eight configurations which are called globally legal configurations [112]. These vertex stars are depicted in figure 11A. Each vertex is denoted by a set of integers (1,2,3,4) which specify the measures of the angles of the rhombi in the vertex, in multiples of $\theta = 36°$. Any tiling of the plane whose stars are limited to these configurations, can be continued using Penrose connection rules and form a Penrose tessellation [112]. These vertex stars can be seen in figure 10A. It is clear there that all the vertex stars are either of the eight configurations depicted in figure 11A. Figure 11B shows two example configurations, which are formed according to the Penrose matching rules but cannot be continued. Such configurations are considered as locally legal configurations, but if they occur in a tiling, that tiling is not a Penrose tiling.

Each Penrose tile T' (wide rhombus) and t' (slim rhombus) (figure 12) can be divided to smaller wide and slim rhombi T' and t' in which the ratio between the side length of the first generation to the next generation is the golden ratio Φ. In fact, by choosing a decagon in the Penrose tessellation (one example is the red

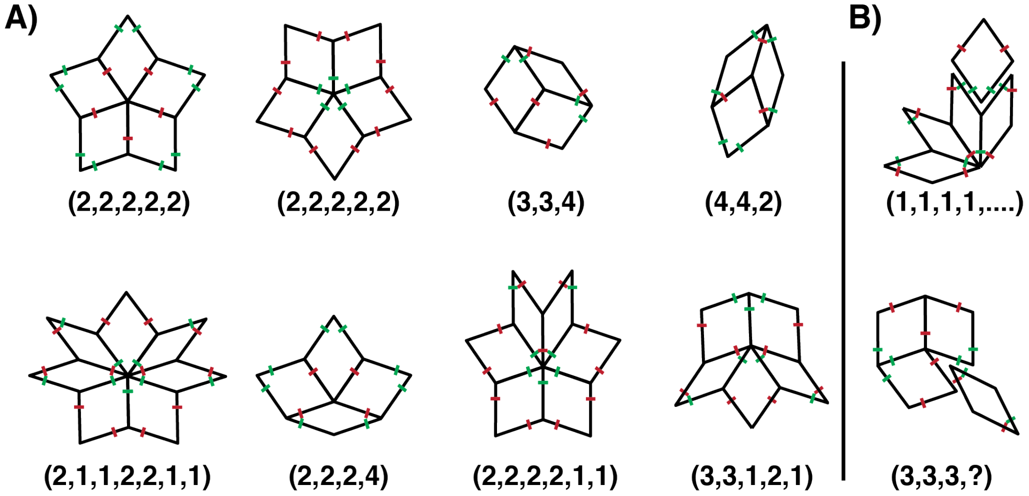


Figure 11: A) eight globally legal Penrose configurations which can be continued employing Penrose connection rules [112]. Integers are the measures of angles of the vertices, in the multiples of $\theta = 36°$. B) Two examples of locally legal Penrose configurations which cannot be continued using Penrose connection rules and are considered as dead ends [112].

decagon in the figure 10A) and doing infinite deflation-inflations to go to smaller generations of prototiles, one can prove that Penrose tiles cover the whole surface (see the reference video for visual illustration [114]). This tile division can be written by Fibonacci matrix as following [112]:

$$T' = 2T + t \tag{30}$$

$$t' = T + t \tag{31}$$

$$\begin{pmatrix} T' \\ t' \end{pmatrix} = M \begin{pmatrix} T \\ t \end{pmatrix} \tag{32}$$

$$M = P^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \tag{33}$$

where P is the fibonacci matrix:

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \tag{34}$$

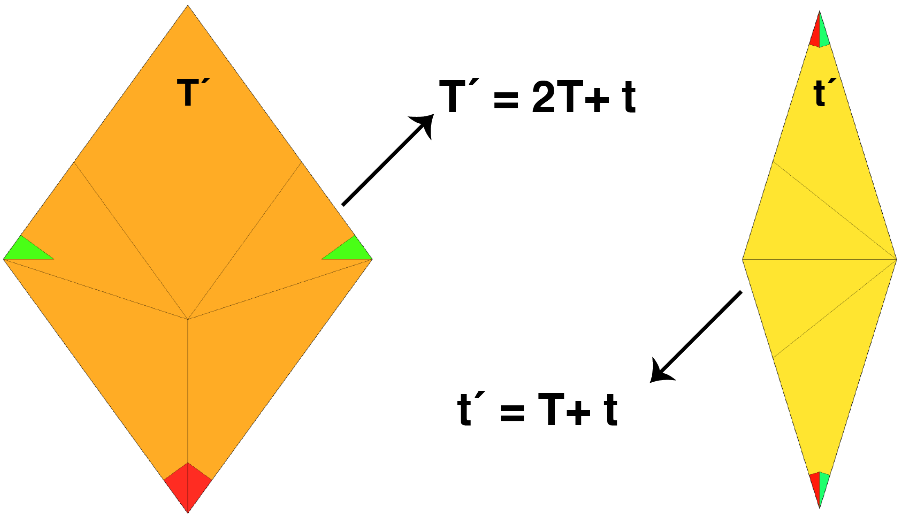


Figure 12: Schemes of Penrose tiles which show how tiles can be divided to smaller tiles following the Fibonacci matrix squared P^2 [112].

It is also possible to produce a periodic tiling by using different connection rules for Penrose rhombi. Another possibility is to produce non-quasicrastalline and non-periodic tilings by Penrose prototiles. This way is a variant tiling in which five-fold symmetry can exist. However, this five fold symmetry is not long range and consequently the tiling is not quasicrystalline.

2.7 Avidity or multivalent binding

Multivalent interactions, in which multiple ligands and multiple receptors bind to each other, are very common in nature. There are three general cases for receptor-ligand multivalent binding [115] :

1) monovalent ligand/bivalent receptor
2) Bivalent ligand/monovalent receptor
3) Bivalent ligand/bivalent receptor

Multivalent interactions have many applications in biology including immune responses such as viral and bacterial adhesion to tissue, antibody recognition, opsonisation of pathogens and immune system response to inflammation [21, 116, 117, 118].

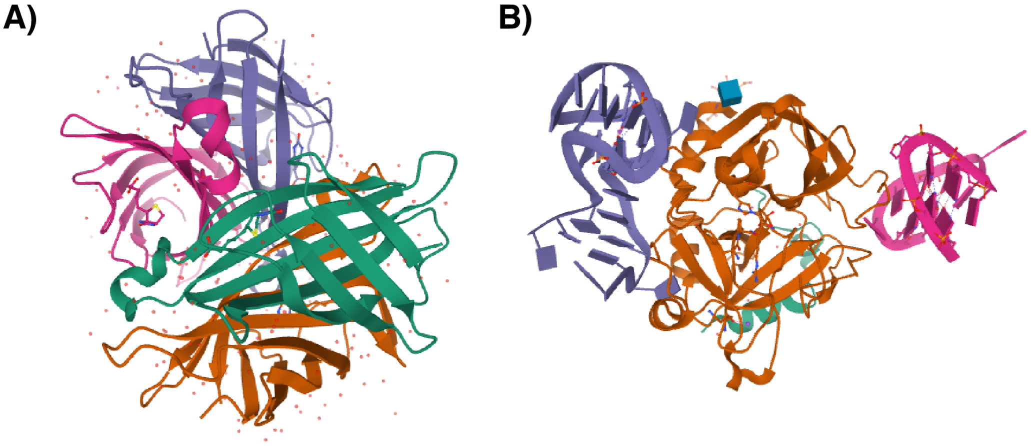


Figure 13: A) X-Ray diffraction structure of tetrameric streptavidin bound to biotin (taken from PDB `https://www.rcsb.org/3d-view/1NDJ/1` [119]). B) X-Ray diffraction structure of thrombin protein and its corresponding DNA aptamers. The DNA aptamers are coloured in purple and pink (taken from PDB `https://www.rcsb.org/3d-view/5EW2/1` [120]).

2.7.1 Streptavidin and thrombin as two examples for multivalent binding

Streptavidin is a 52.8 kDa protein which has four subunits, each of them with great affinity to biotin and can bind non-covalently to it (see figure 13A). Binding of streptavidin to biotin is one of the strongest non-covalent bindings in nature. The dissociation constant K_d for streptavidine bound to biotin is in fM range. In comparison, the dissociation constant for high affinity Antigen-antibody is on the order of 10^{-7} M.

Further, DNA aptamers have been identified, which can bind to streptavidin subunits. However, their K_d is typically in the range of 1-100 nM [121, 122]. Together, streptavidin and these aptamers can form a multivalent system. Although the binding affinity of streptavidin aptamers is low, a scaffold system which contains multiple aptamers could strongly bind to streptavidin [123].

Thrombin (36 kDa) is also a multivalent protein, with two identified DNA aptamers that can bind to specific sites on thrombin (see figure 13B) [124]. The dissociation constants for these two aptamers are also in the range of 1-100 nM, nevertheless, one can create high affinity systems by including both aptamers in one scaffold [20, 123].

3 Materials and Methods

In this chapter, general design principles of DNA origami structures will be discussed. Subsequently, information on Penrose tiles, the design features and the preparation and folding process of the structures will be provided. In the third section, the design characteristics of square-shaped DNA origami structure for protein binding assays, the 2D crystallisation of them and their 2 x 2 array formation will be explained. This section is followed by a discussion of the barcoding scheme and finally include the detailed protocols for sample preparation, purification, protein labelling and enrichment experiment. In the last two sections, two microscopy methods which have been used in this thesis, atomic force microscopy (AFM) and DNA PAINT will be presented.

3.1 DNA Origami design

The design of DNA origami structures can be divided into five steps [2]:
A) For the first step, it is necessary to make a geometric model for the desired shape (red line in figure 14A) and fill it with cylindrical segments which are double helices. Each cylindrical segment contains one helical turn with the length of 10.67 bases or 3.6 nm. Blue crosses in figure 14A are crossover positions that connect two double helices and keep them together. In these crossovers, one strand connects one double helix to the other one and continues there. In this thesis, the crossovers between two adjacent helices are occurring every 1.5 turns, consequently, the inter-helical distance between double helices is considered as 1 nm [2]. As a result, the height of the structure in y direction is a sum over helices diameter (2 nm) and the inter helical distances (1 nm).
B) After making the geometric model, the scaffold path for folding is specified (see figure 14B). In general, the scaffold path starts from one point, continues through the whole structure and ends at the same point (in this thesis circular phage genomes are used). To traverse between helices (cylinders), crossovers are made at the tangent point between the helices. In figure 14B, red crosses show scaffold crossovers. At these points, bases which connect two helices are facing in the same direction. The distance between two scaffold crossovers must be a multiple of 1.5 turns or 16 bases. As can be seen in the figure 14C, the result is a structure with a seam in the middle. This seam or line is where the scaffold does not cross.

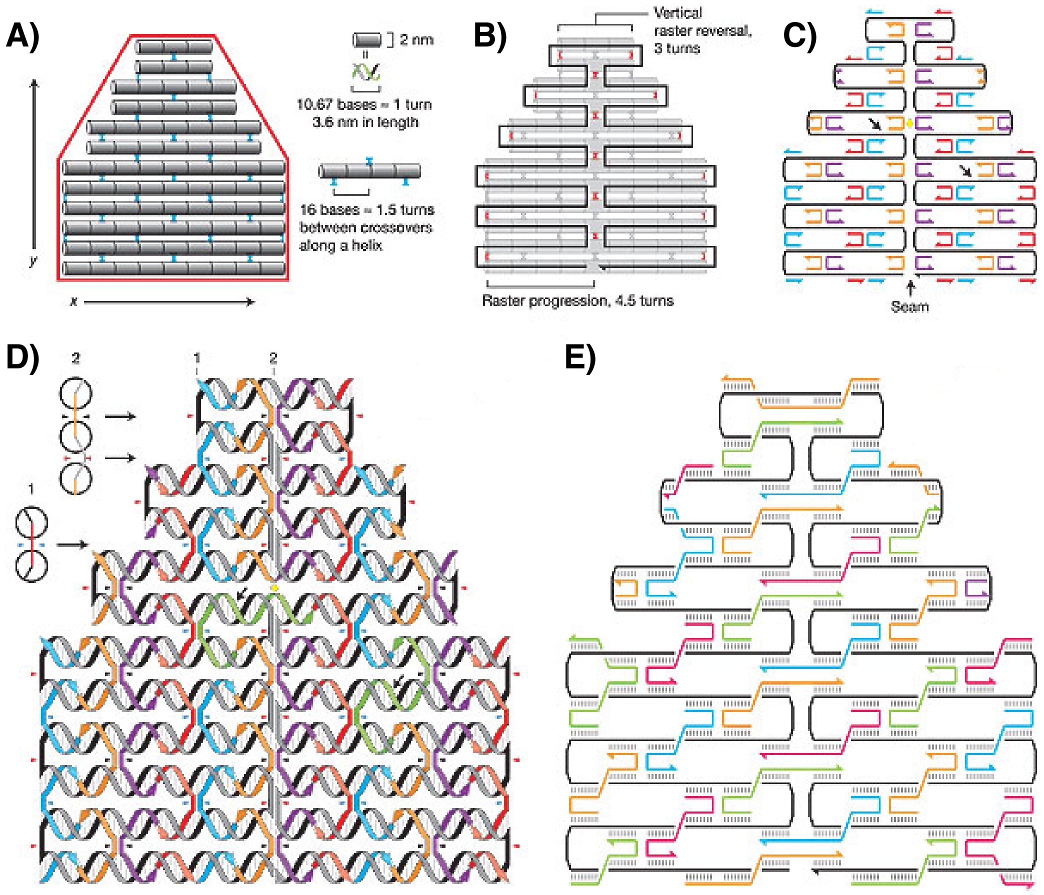


Figure 14: A) The red line is the geometric model of the desired shape filled with cylindrical segments each represents a DNA double helix with one complete turn. Blue crosses show crossover positions. B) The scaffold path with red crosses as scaffold crossovers. C) Preliminary staple design. Each staple is going only between two helices and is 16 bases. D) The same scheme in C only with helices drawn. Red triangles specify scaffold crossovers, black and blue ones specify staple crossovers with minor groove on top face and on bottom face. The top views on left show angles of staples in crossovers. E) Final design with extended staples which connect more than two helices and usually have 32 bases. Figure has been taken by permission from [2]. Copyright 2006, Springer Nature.

C) The third step is the staple design. To do so, the scaffold path can be drawn in the caDNAno software [125], as a base for staple design [126]. Staples are designed in a way that they do not leave scaffold bases unpaired which means they crossover between different helices and reverse their direction when they go from one helix to the other. In figure 14C, staples are connecting two helices and have a length of 16 bases.

D) Due to the asymmetric nature of DNA (major and minor grooves) and also the fact that a half-turn-helix contains a non-integer number of bases, a small strain will appear between bases at the crossovers. This strain can be prevented by precise positioning of crossovers. In figure 14D, red triangles show scaffold crossovers and black and blue triangles show crossovers with the minor groove on the top and bottom face. The left top views show the angle of staples at crossovers. Analysing these angles shows that sometimes by shifting and repositioning the crossover, strain in the structure can be avoided. For example in the top view left for number 2, close to red triangles there is a strain; however, close to the black triangles there is no strain.

E) Eventually, staples will be extended to have larger binding domains bases (not less than 8 bases). Moreover, they will connect more than two helices together. Most of these staples have 32 bases in length. In principle, staples which connect more than two helices are potential causes for kinetic traps. For example, a staple which connects three helices should not have a short middle segment with two long outer segments. This could cause that the middle segment never hybridises to the scaffold, which results in instability of the structure. Therefore, it would be better to use short-long-short segments in a staple design since the staples will not become too long and the possibility of having kinetic traps will be lower. In addition, it is better to avoid symmetric staple designs like long-long-long segments (in which one long segment alone has a high melting temperature) because staples are always in excess in solution and that can induce a forever-lasting competition between same staples and partial hybridisation of the staples to scaffold. After completion of the staple design, the list of staple sequences can be exported from caDNAno software and ordered for synthesis.

3.1.1 Three Dimensional DNA Origami Tiles

The same principles for designing 2D DNA origami tiles can be utilised to design 3D structures. The only difference here is scaffold routing. The caDNAno software provides two kinds of lattices which can be applied. One is the square lattice and the other is honeycomb lattice. These lattices specify the position of double helices with respect to their neighbors in 2D or 3D. In the square lattice each double helix in 3D has four nearest neighbors, however, in the honeycomb lattice each double helix has three nearest. In order to design 3D structures, scaffold routing must go through nearest neighbors of one double helix which are not in the same plane containing that double helix.

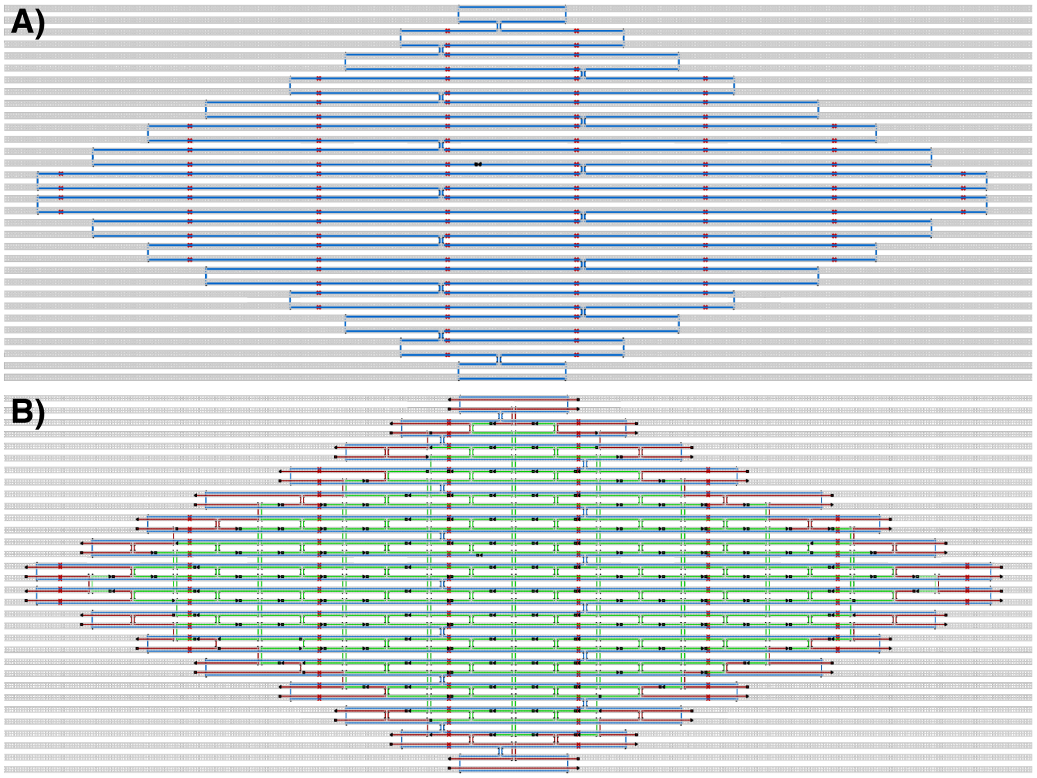


Figure 15: A) Scaffold routing in caDNAno for wide rhombus of Penrose in the square lattice. Red crosses are skipped positions for twist correction. B) Complete design of wide rhombus after staple breaking. Red staples are edge staples with 4 thymidine extensions to avoid base-stacking.

3.1.2 Twist correction

As it is discussed in section 2.2.3, for the B-form DNA a helical twist of 10.5 base pairs per turn is calculated. Designing in the square lattice using caDNAno software, the value which is assumed for the helical twist is 10.67 base pairs per turn which has been identified to cause a significant global twist in the DNA origami structures [127, 128]. To compensate this overall twist, the number of bases in a scaffold row is counted and divide by 10.67 to find the number of complete turns in that row in the caDNAno design. Then it is multiplied to 10.5 to find the number of bases for that row, assuming B-form DNA. Comparing this number to the value, which is obtained by caDNAno for the number of bases in that row assuming 10.67 base pairs per turn, the number of bases that needs to be deleted can be figured out [129]. These deletions will be distributed in the scaffold row and indicated as red crosses in the caDNAno design (see figure 15A and 15B). The deletion of bases is carried out in the same location for different scaffold rows which results in columns of red crosses in the design.

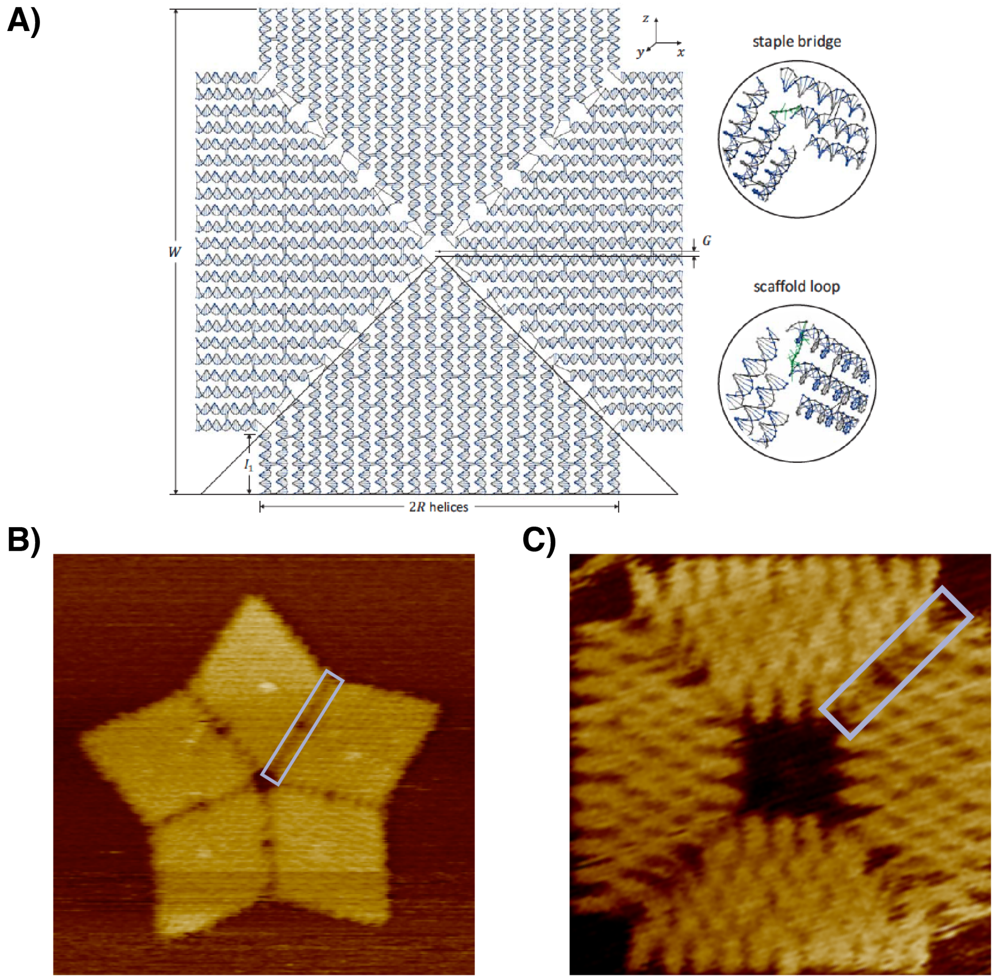


Figure 16: A) A three-dimensional model for calculating the lengths of single-stranded domains in staple bridges and scaffold loops. Figure has been taken by permission from [106]. Copyright 2016, Springer Nature. B) AFM image of a five-fold star. C) AFM image of a square-shaped origami structure. The grey rectangles in A and B identify connection regions between two double helices which should be bridged by single-stranded domains.

3.1.3 Calculation of single-stranded domains in the connection regions

Usually, in DNA origami designs, integer number of blocks (cylinders in figure 14A) are used for each helix row while it is possible to use arbitrary number of bases in a helix row. Figure 16A shows a three-dimensional model that is used in Grigory Tikhomirov et al. [106] to calculate the scaffold loops and staple bridges in their square DNA origami structure. In their design, they are not applying integer number of blocks, therefore, where the isosceles triangles connect the scaffold loops should contain single-stranded domains to fill gaps. The same is applied for the staples which bridge between isosceles triangles. This way the square structure will be composed closely together with minimised strain. In principle this method can be utilised for other structures or connecting regions that need single-stranded domains. Two examples of such regions which are related to this thesis are shown in figure 16B and 16C. The grey rectangles are representing the regions which need single-stranded domains.

In their model, a coordinate system is used in which the origin is the centre of the structure and the coordinate locations of where each base joins the backbone is calculated. To determine the single-stranded domains length, the Euclidean distances between all pairs of coordinates in staple bridges and scaffold loops are calculated and the length will be identified considering 0.4 *nm* for each nucleotide in single-stranded DNA (see Grigory Tikhomirov et al. [106] for a more detailed explanation). To do such calculations there are few assumptions which should be taken into account.

1) Length of each base pair in a double helix is 0.34 *nm*.

2) The dimeter for each double helix is 2 *nm*.

3) The inter helical distance which depends on the spacing between crossovers (1.5 or 2.5 turn). In case of 1.5 turns spacing between crossovers, the inter helical distance equals to 1 *nm*.

4) The value of the number of base pairs per 360° turn which depends on the twist correction. In their case the authors eliminated one base in each three columns (three 1.5 turns) which results in 10.44 base pairs per turn.

5) The assumption of 150° as the angle between the lines which connect the centre of helix to where staple and scaffold bases join the backbone [2].

Considering these values one can precisely calculate the domain lengths precisely and minimise the strain in the connecting regions in order to obtain a high yield of

DNA origami structures.

3.1.4 DNA origami structures designed for this thesis

Three structures have been designed for this thesis and all of them are designed in the caDNAno software using square lattice. AFM images of these structures are shown in figure 17. The two rhombi are designed during my master thesis [130]. In the experiments for this thesis, only the square-shaped DNA origami structure with the cavity in the centre is used.

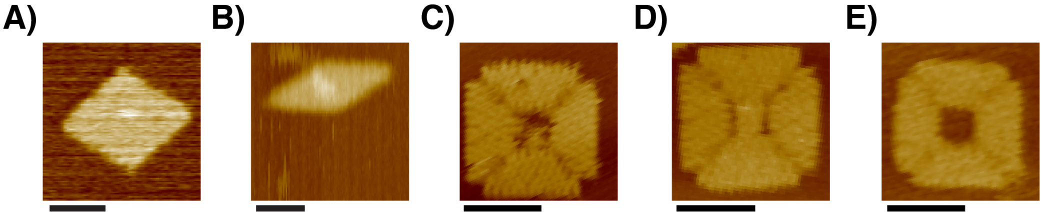


Figure 17: AFM images of A) Wide rhombus. The bright dot in the middle shows the scaffold loop. B) Slim rhombus. The bright dot shows the scaffold loop. These two structures are designed during my master thesis [130] C) Square-shaped structure with scaffold loop in the middle. The scaffold loop in the middle can be kept unfolded or folded to any desired sub-structure. D) Square-shaped structure with the scaffold loop folded to a sub-structure. E) Square-shaped structure with the scaffold loop left in the corner. In case of using cavity for binding assays this version is employed. All scale bars are 50 *nm*

3.2 Barcoded DNA origami assemblies for protein-binding assays

In this section, information regarding the square-shaped DNA origami structure is provided. This structures is used in the barcoded assemblies for protein-ligand binding assays. The project work is published in Ali Aghebat Rafat et al. [123].

The square-shaped DNA origami platform with an approximate four-fold rotational (C4) symmetry (figure 18A) is based on the design used previously for the assembly of extended DNA origami crystals [106]. In this design, DNA helices extend from the centre in two perpendicular directions (denoted by North-South (N-S) and West-East (W-E), appendix figures 59 and 60), which facilitates multimerisation between origami structures along the helix axis at all four edges by using a combination of sticky- and blunt-end interactions. The square-shaped DNA origami structure is designed with the possibility of either having a cavity or a substructure in the centre (see figure 17C, 17D and 17E). One possible application for a substructure in the centre would be forming of a platform for a dynamic arm [131]. Due to the four-fold symmetry of this structure one could potentially create 2D crystals of such arms in case it is needed.

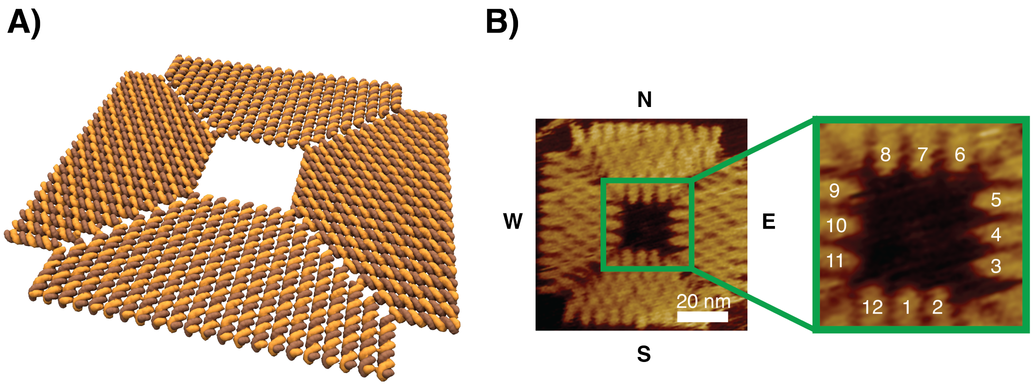


Figure 18: A) 3D Molecular model of the DNA nanostructure [123]. B) A detailed AFM image shows the origami structure and its cavity. For orientation, the outer edges of the structure are addressed with N, S, E, and W. The zoomed-in image of the symmetric cavity shows the 12 staples available for modification and their respective numbers. These staples are used to arrange the DNA aptamers in a controlled orientation and position [123].

The main goal to design the square-shaped DNA origami is formation of a cavity in its centre that can have multiple binding sites with different geometries and properties (see figure 18). Moving in this direction, this cavity is utilised to investigate binding properties of the thrombin and streptavidin to their known aptamers as multivalent proteins. Such a cavity with multiple binding sites would be a proper tool to study avidity of multivalent proteins to their ligands. The scaffold routing through the structure is done in a way that results in a symmetric square-shaped cavity in the centre. This cavity is lined with 12 uniquely addressable staples that can be modified arbitrarily and has a side length of approximately 24 nm (see figure 18B).

3.2.1 2D crystal formation

2D crystals of square-shaped origami structure are formed using two different methods (appendix figure 61). Both of them are in solution assembly methods which use combination of blunt-end interactions and base pairing (2-base truncation and

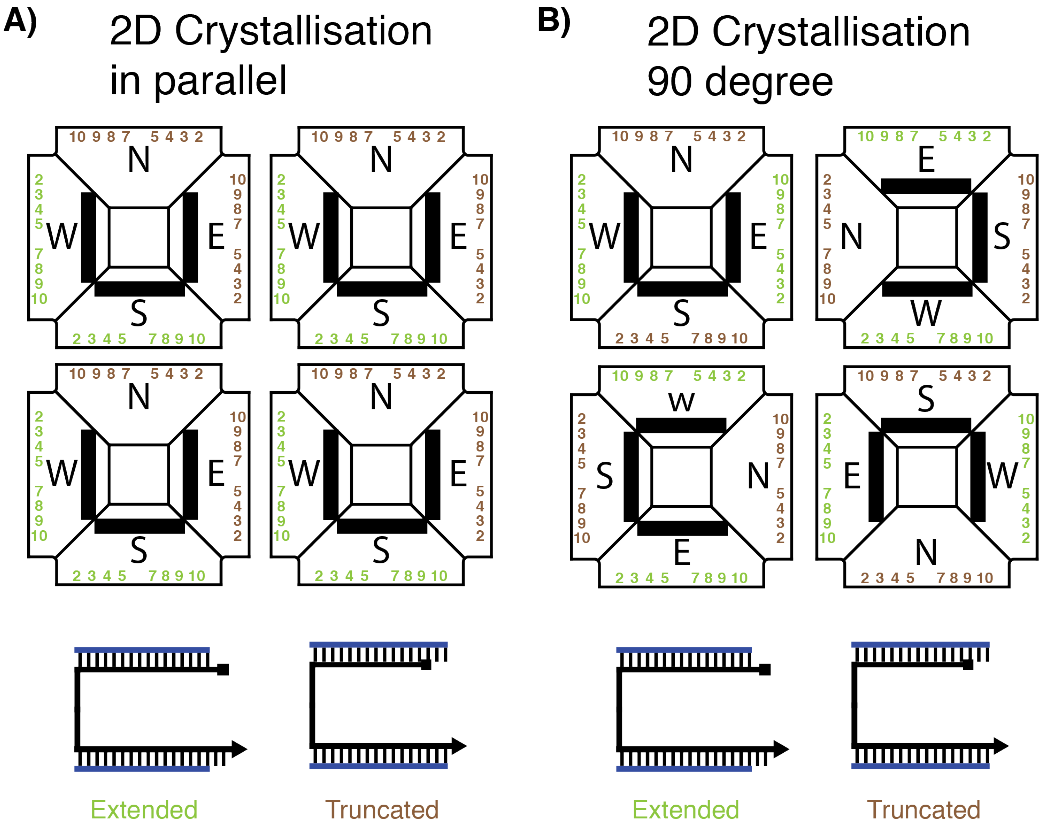


Figure 19: A) Scheme of the 2D crystallisation by using parallel sides (N-S, E-W). This type of connection resulted in elongated crystals, potentially due to the build-up of internal strain. Color codes and staple coordinates show the type of staple modification and the staples which are used in an edge respectively. B) Scheme of the 2D crystallisation with 90° rotation (N-E, W-S). Color codes and staple coordinates show the type of staple modification and the staples which are used in an edge respectively.

extension). Figure 19A and 19B represent the connection rules and type of staple modification which are applied to create these crystals. In case A, multimerisation is along the North-South and West-East axes which ended in formation of elongated crystals. In case B, North side of the structure is connected to its East side and the South side to the West side. This type of connection resulted in well-formed 2D crystals presumably due to the cancellation of the intrinsic twist in single structures.

3.2.2 2 x 2 array formation and barcodes

As it is shown in section 2.5.2, fractal assembly enables the creation of finite DNA origami arrays made of unique and independent structures that fold in separate pots [108]. Since the structures are folded separately, each of them can be equipped with different features that are uniquely addressable. In order to be able to distinguish between different single structures in AFM imaging, a barcoding system is established which uses dumbbell hairpin modified staples [2] to generate AFM height contrast between the structure and the barcode patterns. It is considered to create higher order fractal structures (4 x 4, 8 x 8 array) with less complexity in barcoding (where structures could be also identified via their position within the arrays), or to use only single structures or 2 x 2 arrays with a more complicated barcoding scheme. Due to the fact that the yield of fractally assembled structures drops dramatically for higher order assemblies, whereas here the yield of 2 x 2 arrays is reasonably good, it is decided to stick to 2 x 2 arrays and apply a barcoding scheme which not only enabled identifying each single structure within a 2 x 2 array, but also to distinguish between different 2 x 2 arrays. Therefore, the barcoding approach represents a balance between barcoding complexity and assembly yield. In figure 20A the connection rules for 2 x 2 array formation and the barcoding scheme for 2 x 2 array number 1 are shown.

Figure 20B shows three different types of staple modifications used at the edges of the tiles: 2-base truncation (brown), 2-base extension (green) and double hairpin passivation (black), which enable the structures to connect and create 2 x 2 arrays. For truncation and extension, 8 of 11 staples at the edge of a DNA origami structure are modified. Colors represent the type of modification for the corresponding staple coordinates. Within a 2 x 2 array, the four origami structures are numbered clockwise using Greek numbers from I to IV. The barcoding patterns are implemented in a way that resemble digital numbers. In the experiments here only a maximum of eight 2 x 2 arrays in a one-pot experiment is needed, but in principle, the barcoding scheme could be used to label several tens of distinct 2 x 2 arrays in one experiment. The barcodes are designed to follow a sufficiently asymmetric pattern in order to be able to distinguish arrays and single structures even when they landed on the mica substrate upside down or had defects. Considering these points, a zero-shape barcode is used on structure number I and no barcode at all on structure number IV

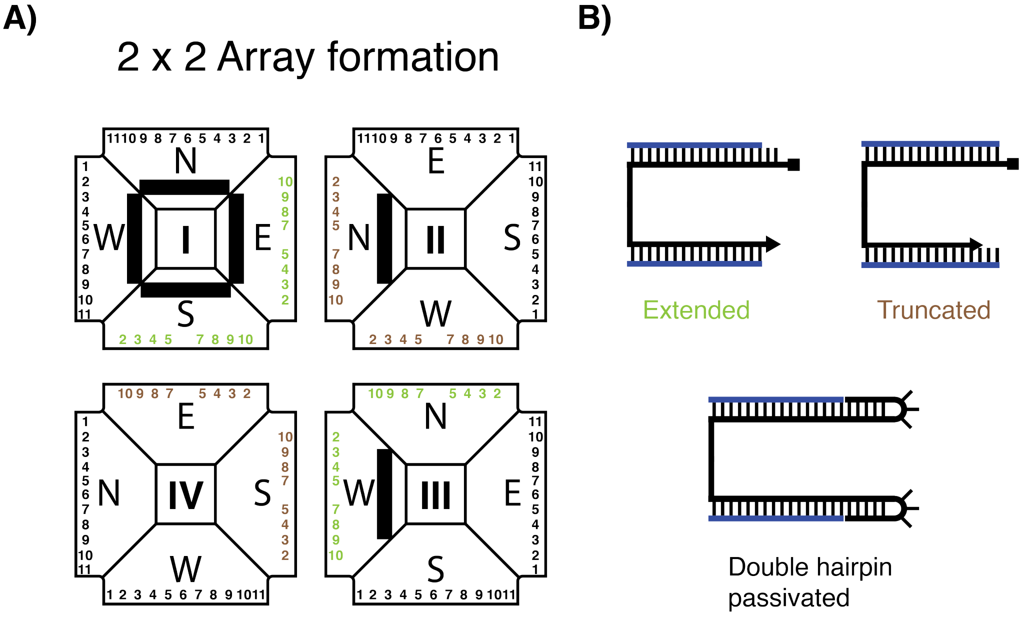


Figure 20: A) 2 x 2 array scheme which shows the connection rules between different sides. 1 to 11 are the staple coordinates which are at the edge of the DNA origami structure and their modification enable bounded and unbounded arrays. Color codes in addition to the staple coordinates show on each side of the square (North, South, East, West) which staples are used and what kind of modification they have. B) Different types of modification used for creation of 2 x 2 arrays (truncation, extension and double hairpin passivation).

which creates the necessary asymmetry for AFM imaging when using only one-digit numbers. Also when using two-digit numbers, enough asymmetry can be implemented to allow to distinguish different structures. Two-digit numbers with too high a degree of symmetry such as 11, 22, 33 ...99 or 89 and 98 and etc, cannot be used unless additional asymmetry is introduced elsewhere on the origami tiles. Here each trapezoid is used for one part of a digit to make sure that the patterns can be resolved even when the imaging quality and resolution is not very good. However, in principle one could also create a complete digit on each of the four trapezoids that would enable more barcoding possibilities, but identification by AFM imaging would then be more challenging. In the experiments here, the barcoding system is used to uniquely address each single structure within a 2 x 2 array and to specify which modifications are made in that structure (for example, the configuration of aptamers used or different flexibilities) when it is folded in a separate pot.

3.2.3 Passivation of blunt ends

In section 2.2.3, base-stacking interaction between DNA base pairs is discussed. Stacking interaction is a crucial attraction force which enables double helical conformation of dsDNA, which can also be used to connect DNA nanostructures. This short range attraction would strongly happen when two double helices at the edges of DNA origami structures are aligned in parallel to each other. For double helices that have an angle with respect to each other, in the presence of a hinge in the proximity, double helices can still bend and align parallel to each other and connect [129]. Otherwise nonparallel helices cannot interact. In cases that where is no intention to connect the sides of DNA origami tiles, the base-stacking interactions has to be passivated. This can be done by either leaving out the edge staples (consequently single-stranded scaffold will be left at the edge) or using extensions of thymidine bases (4T) (see figure 15B) [2].

For the square-shaped DNA origami structure, it is decided to use a different type of modification which is called double hairpin passivation [108] (figure 20B, 19A and 19B). Since this structure is used to study the binding of aptamers to their respective proteins and the aptamers have poly Adenine in their bases, unspecific interaction between aptamers and the edges are observed. Therefore, double hairpin passivated staples are implemented at the edges to prevent both base-stacking interaction and unspecific binding of aptamers there.

3.2.4 Preparation of scaffold and DNA aptamers

M13mp18 scaffold strand is provided in 100 nM aliquots in water by Florian Praetorius [90]. Staples are purchased from IDT in 1×TE buffer with a 100 μM concentration. DNA aptamers and modified staples in the cavity are ordered with HPLC or PAGE purification (see supplementary excel file of Ali Aghebat Rafat et al. [123]).

3.2.5 Sample preparation for 2 x 2 arrays

For 2 x 2 arrays, for each tile reaction mixtures of 50 μl total volume, containing 20 nM scaffold, 140 nM of unmodified staples (a mixture of all staples except barcodes or refills (unmodified staples for the barcode positions), edge staples and the staples in the cavity), 200 nM of either barcodes or their refills for each trapezoid (N, S, E, W), 200 nM to 1 μM of unmodified staples in the cavity, 350 nM of active edge staples for each tile, 100 nM to 180 nM of double hairpin passivation (DHP) edge staples for each side, 1 μM of each modified staple in the cavity which bind to DNA aptamers and 2 μM of each DNA aptamer in 1×TE, 12.5 mM $MgCl_2$ and 5 mM KCl (= 1×folding buffer = 1×FB) are annealed using Thermocycler.

2 x 2 arrays folded for streptavidin (STV) have 200 nM of modified staples in the cavity and 2 μM of the corresponding DNA aptamer. Since in the streptavidin case the same aptamer is used at four positions in the cavity we used a lower concentration for the modified staples, so that the aptamers are still present in excess per binding site. (Note that DNA aptamers should be always added in excess of the modified staples in the cavity in order to ensure complete occupation of all the aptamer binding positions in cavity). First, single DNA origami structures are folded using a temperature ramp which first held temperature at 70°C for 5 min and then decreased from 65°C to 45°C with a rate of 0.1°C per 12 seconds. After individual structures are folded, equal amounts of all the four structures comprising a 2 x 2 array are added to a new reaction tube (200 μl) and mixed properly using a vortex mixer. After mixing, the 200 μl volume is divided into two 100 μl fractions and put back into the Thermocycler. The temperature ramp used for annealing single structures into 2 x 2 arrays started at 55°C and decreased to 20°C with a rate of 0.1°C per 2 min.

3.2.6 Purification of 2 x 2 arrays

Folded 2 x 2 arrays are purified at 20°C using 100 kDa Amicon Ultra 0.5 ml filters (UFC5100BK). To wet the filters, first a one-step centrifugation is performed at 10000 rcf for 3 min using 400 μl of 1×TE, 12.5 mM $MgCl_2$ and 5 mM KCl (1×FB) buffer solution. Afterwards 30 μl of the 2 x 2 array sample plus 400 μl of 1×FB are added to the filter and centrifuged at 10000 rcf for 3 min. We repeated the same process several times by adding 30 μl aliquots of the 2 x 2 array sample to the same filter until all the sample is purified. Finally, the filter is inserted upside down into a 0.5 ml Amicon tube and centrifuged at 10000 rcf for 3 min to recover the purified arrays. After purification, the sample concentration is determined by measuring its absorbance at 260 nm. We also used a PEG purification protocol adjusted from Stahl et al. [132]. In this case, a 1:1 ratio of sample and precipitation buffer (1×TE, 12.5 mM $MgCl_2$, 5 mM KCl, 10% PEG 8000) are mixed and centrifuged at 20°C with 16000 rcf for 30 min. Afterwards the supernatant is removed and the pellet is resuspended in 1×FB.

3.2.7 Preparation of 2D crystals

For the formation of 2D crystals, reaction mixtures of 50 μl total volume, containing 40 nM scaffold, 200 nM of unmodified staples (a mixture of all staples except barcodes or refills, edge staples and the staples in the cavity), 200 nM of barcodes or refill stock for each trapezoid (N, S, E, W), 250 nM edge staples for the formation of 2D crystals (N-E and W-S), 200 nM of unmodified staples in the cavity, 1 μM of each modified staple in the cavity which bind to the DNA aptamers (in case of STV 200 nM of each modified staple in the cavity) and 2 μM of each DNA aptamer, all in 1×TE, 12.5 mM $MgCl_2$ and 5 mM KCl are annealed using Thermocycler. Folding of individual structures and 2D crystallisation is performed in one pot. The reaction mixture is first held at 70°C for 5 min, then the temperature is ramped from 65°C to 50°C with a rate of 0.1°C per 12 seconds, followed by another temperature ramp from 50°C to 20°C with 0.1°C per 2 min or 0.1°C per 6 min.

3.2.8 Protein binding to the DNA origami structures

Human α-Thrombin (Haematologic Technologies) is diluted in 1×FB to a concentration of 270 nM. Streptavidin from *Streptomycesavidinii* (Sigma-Aldrich) is diluted in 1X×FB to a concentration of 900 nM. For the incubation of 2 x 2 arrays with Streptavidin, 2 x 2 arrays at a concentration of 5.5 nM are mixed with 72 nM Streptavidin in 1× FB and incubated for 30 min to 1 hour at 37°C. For the incubation of α-Thrombin, 2 x 2 arrays at a concentration of 5 nM are mixed with 67 nM α-Thrombin in 1×FB and incubated for 30 min to 1 hour at 37°C. For the incubation of 2D crystals with Streptavidin, DNA origami crystals (25 nM) are mixed with Streptavidin at a concentration of 330 nM in 1×FB and incubated at 37°C for 30 min to one hour. For the incubation of 2D crystals with α-Thrombin, DNA origami crystals (13 nM) are mixed with α-Thrombin at a concentration of 180 nM in 1×FB and incubated at 37°C for 30 min to one hour. After incubation and before imaging, buffer solution is added to adjust the final concentration of DNA origami and proteins to 3 nM and 40 nM, respectively. For the 2D crystals, the final buffer solution has 120 mM to 150 mM NaCl in addition which facilitates the AFM imaging of crystals on mica substrate.

3.2.9 Enrichment of scaffolded aptamer configurations using streptavidin microbeads

Selection of origami-scaffolded aptamer configurations is performed using Dynabeads MyOne Streptavidin C1 (ThermoFisher Scientific) with a diameter of 1 μm. 100 μl of magnetic beads are washed three times with Washing Buffer (5 mM Tris-HCl, 0.5 mM EDTA, 1 M NaCl), and two times with 1×FB (1×TE, 12.5 mM MgCl2, 5 mM KCl) in low DNA binding 1.5 ml tubes (Eppendorf) using magnetic separation. Reactions are prepared by adding 5 nM of each origami configuration and 109 microparticles in a total volume of 100 μl of 1×FB. Samples are incubated at RT for 30 min, then placed in the magnetic separator for 1 min, after which the supernatant is carefully removed. Samples are washed 5 times in 100 μl 1×FB. We briefly incubated the samples with 100 μl of increasing concentrations of biotin (800 nM, 8 μM, and 80 μM) (Sigma-Aldrich) in 1×FB to elute the bound nanostructures from the streptavidin coated beads. Samples are further analysed under the AFM microscope.

3.3 Penrose tessellation

In section 2.6.2 Penrose tessellation is discussed. One way of creating a Penrose pattern (see figure 10A) is to use two rhombi with a certain geometry. Employing DNA origami technique these two rhombi are designed and folded (figure 17A and 17B and appendix figure 57). in the following section, information regarding design and folding of the single structures will be provided. Furthermore, the folding ramps for the creation of Penrose configurations and Penrose-like patches will be presented.

3.3.1 Design of Penrose tiles

According to section 2.6.2, for Penrose tiling, two rhombi with different geometries and same side lengths are required (see figure 10B). One rhombus has a sharp angle of 72° and an open angle of 108°. For the other rhombus, a sharp angle of 36° and an open angle of 144°. It is seen from the geometry of tiles that the surface area of slim rhombus and the wide rhombus is different. For the wide rhombus, the surface area is about twice as large. To implement this difference in the surface area, for the slim rhombus, scaffold routing is done in a way that the tile has two completely identical layers (see appendix figure 57A). By having two identical layers, first of all, the scaffold length which is needed for the slim rhombus will be twice as long which enables using of the same scaffold length for both structures. Secondly, the desired geometry of two tiles will be reached. For the wide rhombus, a scaffold path which results in a two dimensional DNA nanostructure is made. For both DNA origami tiles, scaffold length of 7249 nucleotides is used (see appendix figure 57B).

3.3.2 Passivation of blunt ends

For the Penrose tiles, staples which do not contribute to the multimerisation process are passivated with four thymidine extensions. In figure 21, red staples at the edges of structure are the passivated ones with four thymidine bases.

3.3.3 Five-fold star

For the design of the five-fold star which is an example of globally legal Penrose configurations, five wide rhombi must be connected via multimerisation staples according to the Penrose connection rules (see figure 10B). These staples start from one side of the wide rhombus and end at the other side of it (appendix figure 58). As shown in figure 16B, connecting the two sides of of wide rhombus comes along with a gap in the connection regions which should be filled using single-stranded domains. Such gaps will occur in all vertex stars of a Penrose tessellation. Two versions of the five-fold stars are designed, one during my master thesis (appendix figure 58A) and an updated version which implemented a more precise calculation for the single-stranded domain lengths. In addition, the structure design is modified since AFM analysis revealed that the wide rhombus needs to be resized to have its geometry closer to the Penrose tile (appendix figure 58B).

3.3.4 Sample preparation for single structures and five-fold stars

M13mp18 scaffold strand is provided in 100 nM aliquots in water by Florian Praetorius [90]. All single structures are folded in 1×TE (40 mM Tris, 20mM Acetic acid and 1mM EDTA) plus 12.5 mM $MgCl_2$. 50 nM scaffold is folded into DNA origami structures (which would result in 50 nM structures) using 3× excess of unmodified staples and 5 to 10× excess of multimerisation staples. To fold single structures the complete mix is heated to 70°C and hold for 5 min. Then temperature is decreased to 65°C and the sample is cooled down to 45°C at a rate of 0.1°C per 12 seconds. Eventually the temperature is quickly decreased and hold to 20°C.

For the formation of five-fold stars, folding of single structures and their assembly to five-fold star is done in one pot. The wide rhombi and the multimerisation staples are added together in one tube and the sample goes through the same temperature ramp used to fold single structures.

3.3.5 Towards Penrose tessellation

To create Penrose tessellation, Penrose connection rules are implemented in Penrose tiles using multimerisation staples with 8 base complementary regions (see figure 21). Since the rhombi sides in Penrose tiling are not aligned in parallel relative to each other, Base-stacking will not be the applicable interaction, moreover, shorter complementary regions such as 4 did not form any Penrose-like domain. Therefore, here eight bases of so-called “sticky ends” are used to connect the sides.

In section 2.6.2 locally legal configurations in the Penrose tessellation which cause dead ends (figure 11B) is discussed. To prevent these configurations or to decrease their probability, connection rules are implemented in a way that two sides of slim rhombus do not interact. In figure 21 top, the upper sides a and $a*$ connect to each other. However, the two lower sides, $b*$ and $c*$ stay inert. As a consequence, the locally legal configuration which is created by slim rhombi will be completely avoided (see figure 11B top). To decrease the possibility of the configuration in which three wide rhombi can connect according to figure 11B bottom, interaction between sides b and c in the wide rhombus (figure 21 bottom) is weakened using 4 bases of sticky ends. This interaction is represented by $b'$and $b'*$. For other connection rules 8 bases of sticky ends are used and implemented in the staples. Due to the orthogonality of Penrose connections, different types of extensions and truncations are used. In some cases truncated scaffold and extended staples interact and in the other cases truncated and extended staples hybridise to each other (dark blue ellipse in figure 21 bottom). In figure 21, letters and their colors specify the corresponding edges which connect according to the Penrose connection rules (schemes on the top

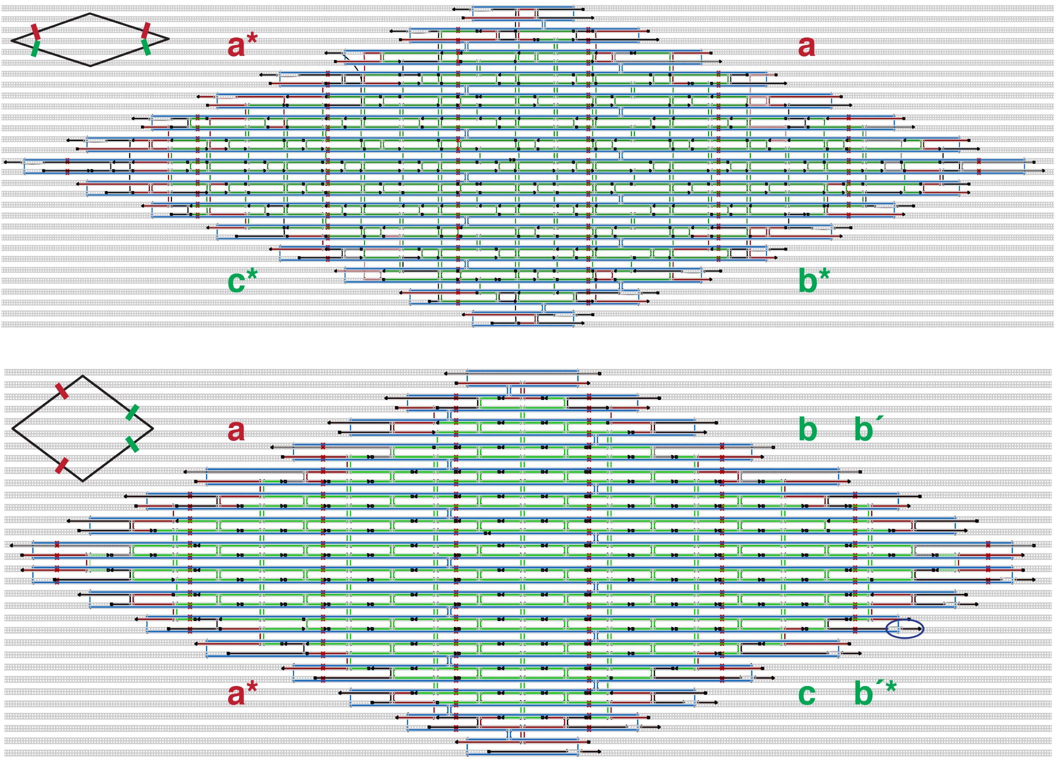


Figure 21: caDNAno designs and respective schemes of slim (top) and wide (bottom) rhombus. black staples represent the multimerisation staples. Letters and their colors specify the corresponding edges which connect according to the Penrose connection rules (schemes on the top left of designs). Star signs on letters stand for complementarity. Red staples represent non-connecting edge staples and the extensions showing 4 thymidine bases used to prevent base-stacking. The dark blue ellipse shows the type of connection in which the last 4 or 8 bases of the original staple have been left out and 8 new bases are added to implement the connection rules. $b'b'^*$ is the weak connection between two sides of wide rhombus in which first 4 bases of the 8 base-extensions in side b are complementary to the truncated part of scaffold. This way the locally legal configuration formed by wide rhombus can be prevented. Moreover, only two sides of slim rhombus can connect to prevent the locally legal configuration formed by slim rhombus.

left of designs). Star sign on letters note the complementarity of its respective side to a side with a similar letter but without star. Detailed caDNAno designs for the rhombi and their connections can be seen in figure 21.

3.3.6 Temperature ramps for Penrose-like domains

Forming Penrose-like domains is carried out in two steps. First single structures without multimerisation staples (sticky ends) are folded using the temperature ramp explained above. Subsequently, multimerisation staples are added following a ramp from 50°C to 5°C at a rate of 1°C per 1.5 hour. One can start from higher than 50°C. However, the melting point of single structures should not be reached. About 28 nM of wide rhombus and 16 nM of slim rhombus are mixed with multimerisation staples for the second step. All reactions are performed in 1×TE (40 mM Tris, 20mM Acetic acid and 1mM EDTA) plus 12.5 mM $MgCl_2$. For the multimerisation staples one can use 5 to 10 times excess of them compared to the respective DNA tile. To prepare the sample with NaCl for imaging, about 5 nM of the respective sample is prepared in a total volume of 50 μl which is in 1×TE (40 mM Tris, 20mM Acetic acid and 1mM EDTA) plus 12.5 mM $MgCl_2$ supplemented with 120 mM NaCl. The concentration of NaCl can be varied between 100 to 150 mM.

3.4 Atomic force microscopy

The microscopy method which is mainly used in this thesis to observe and analyse DNA origami structures is atomic force microscopy (AFM). AFM is a significantly helpful instrument to study biological samples and their evolution in the liquid, moreover, it is able to probe the surface dynamically and shows change in time [133]. In principle, AFM is able to probe a surface dynamically; this is done by a sharp tip (several nm in diameter) at the end of a cantilever. There are mainly two modes which are used in AFM, contact mode and AC mode (or intermittent contact mode). In contact mode, the tip is dragged on the surface and the deflection of the cantilever represents the surface-tip profile. In AC mode, tip and surface are not in permanent contact; instead, the cantilever is driven mechanically close to its resonance frequency and will oscillate close to the surface in order to be influenced by short range forces like van-der-Waals or dipole-dipole interactions and also material properties. Typically short range forces between uncharged atoms can be estimated using a Lennard-Jones potential V_{LJ} as function of the distance r between the atoms (equation 35). This potential consists of a negative, attractive term describing van-der-Waals interaction and a positive, repulsive term approximating the potential arising from the Pauli exclusion principle [134]. ϵ stands for the minimum potential at the distance r_m.

$$\frac{V}{\epsilon} = \left(\frac{r_m}{r}\right)^{12} - 2\left(\frac{r_m}{r}\right)^{6} \tag{35}$$

In practice, cantilever height in AFM is adjusted by using a piezoelectric actuator

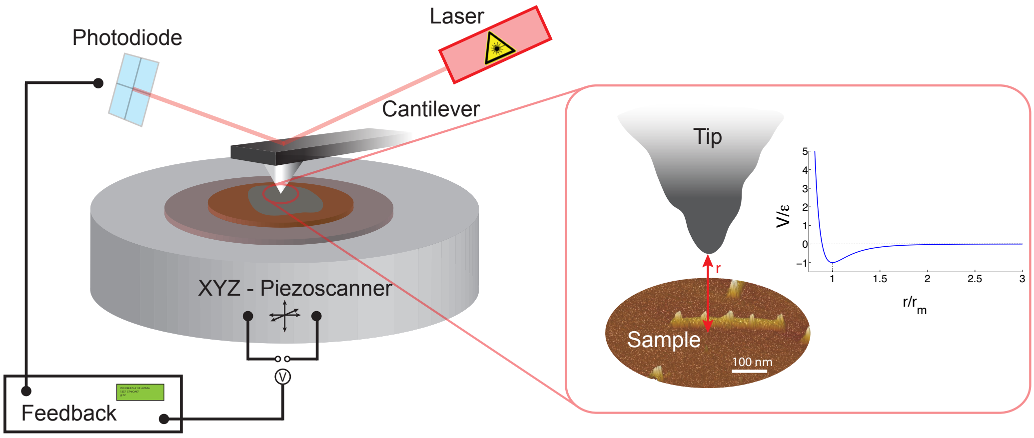


Figure 22: Working principle of an atomic force microscope: a tip is scanned over a sample surface. Thereby tip and cantilever are deflected by forces acting between atoms of the tip and the sample surface. These forces can be approximated by a Lennard-Jones potential (graph on the right). Figure is taken from [135].

which serves to set the oscillation amplitude for the cantilever.

In order to sense the deflection of the cantilever, a laser is reflected at the back of the cantilever to a photodiode consisting of four segments. Then deflection of the cantilever (angular displacement) will be collected by the photodiode and sent as an output signal to the detector (see figure 22).

3.4.1 Photothermal Excitation

For experiments in this work, AC mode in liquid has been used. Usually in AFM, cantilevers are mechanically driven at their resonance frequency by using a piezo. In liquid this will cause two problems, first of all, the mechanical excitation which is done by piezo will be coupled with the ambient environment and cause a drift and distortion in the resonance frequency of the cantilever. Secondly, there will be an instability in the response of the cantilever and also in the amplitude set point which means the amplitude set point has to be changed during imaging (see figure

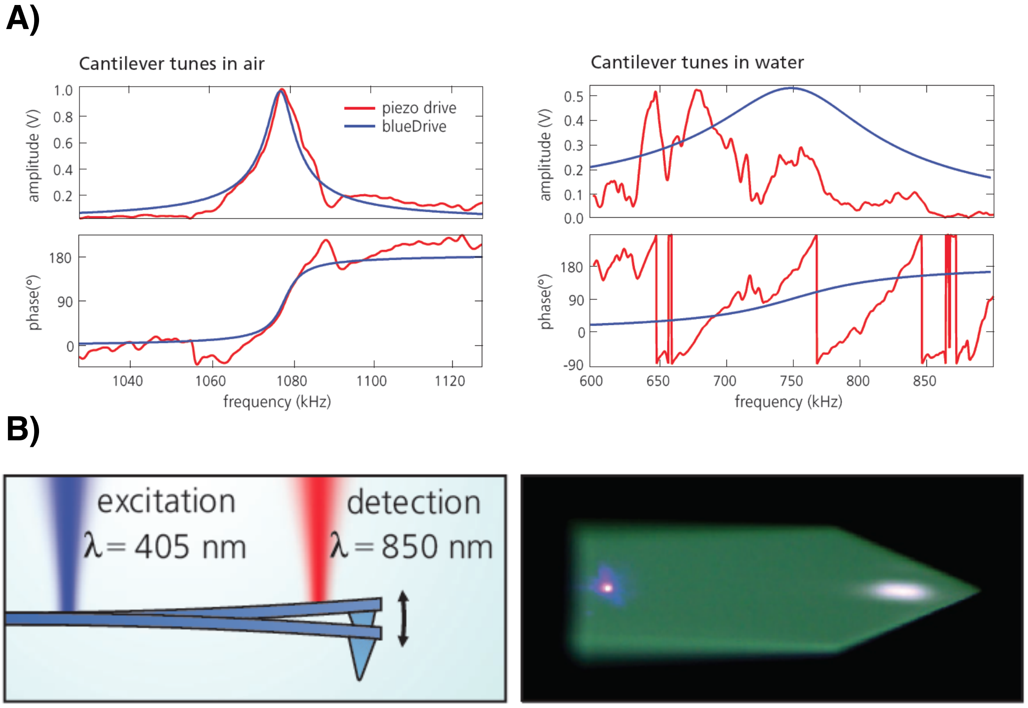


Figure 23: A) Cantilever response of an ArrowUHFAuD cantilever measured for amplitude and phase by using piezo drive and blueDrive in air and in water (red is the response for piezo and blue is for blueDrive). While in water the piezo drive response is far from ideal, the blueDrive response closely matches the thermal response. B) blueDrive concept, left sketch shows actual top-view of the optical image and the right image shows the laser focus positions. Images and plots are taken from Oxford instrument website: [136].

23A).

To overcome these issues, cantilever is thermally excited in its resonance frequency. For the thermal excitation, another laser is used in the set up which is focused near the base of the lever (see figure 23B). In fact, thermal energy of the laser causes mechanical stress at the cantilever. Thus by oscillation of laser intensity at the resonance frequency of the cantilever, it will also oscillate the cantilever at its resonance [137, 138]. Therefore, only cantilever will be excited and as a result all the problems related to the mechanical excitation in liquid which induces unwanted forces on the cantilever will disappear. Asylum Research Cypher AFM which is the atomic force microscope used in this work, has this feature that is called blue drive. BlueDrive is a blue laser which drives the cantilever and it can be moved independently from the primary red laser (see figure 23B). As it can be seen from the tune plots in figure 23A, tuning with the blueDrive is much cleaner than using piezo drive.

3.4.2 Imaging of DNA Nanostructures and Analysis

To image single structures, five-fold stars and the 2 x 2 arrays, 5 μl of the prepared sample (3 to 5 nM) after incubation is dropped onto a freshly cleaved mica substrate and 60 μl of 1×FB (respective 1×FB for Penrose and barcoded assembly) is added on top for AFM imaging in liquid. Images are collected immediately afterwards. In the case of origami crystals and Penrose-like domains, the whole 50 μl sample which is prepared after incubation and supplemented with NaCl (120 to 150 mM) is added onto the freshly cleaved mica and imaged immediately. To image Penrose-like domains without NaCl, 5 μl of 3 to 5 nM annealed samples are dropped onto mica and 60 μl of 1×FB is added on top. AFM imaging of the samples is performed using an Asylum Research AFM, Cypher ES (Oxford Instruments, Abingdon, UK). Two types of cantilevers are used for imaging, FASTSCAN-D from BRUKER and BL-AC40TS-C2 from OLYMPUS. In order to image the barcodes, the FASTSCAN-D cantilevers appeared to perform better due to their higher resonance frequency. Moreover, it seemed helpful to image with lower scan rates to be able to resolve barcodes nicely. Typically, the scan rate is set to 3 Hz.

The statistical analysis for the binding of DNA aptamers to thrombin and streptavidin, 200 tiles for each binding configuration are counted from different mica positions in an AFM session. To test our results we also tried different AFM sessions for same samples and the resulting errors are quite low. The error calculations will be discussed in detail in the result section.

3.5 DNA PAINT

PAINT (point accumulation for imaging in nanoscale topography) is first described by Sharonov et al. [139] which is a subdiffraction imaging based on targeting the surface of objects by using diffusive fluorescent probes in solution. In 2010 Jungmann et al. [140] used this concept by means of diffusing short fluorescently labeled DNA probes in solution (imager strands) that can transiently bind to docking strands protruding from DNA origami structures.

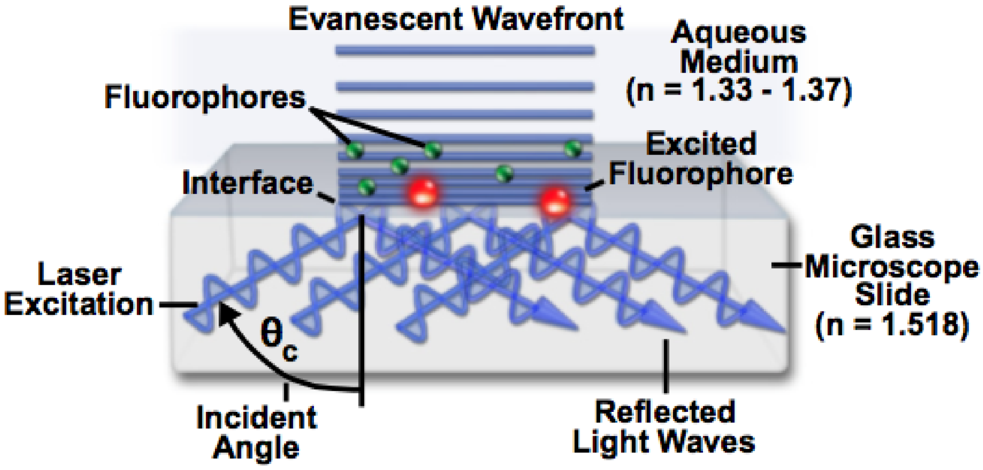


Figure 24: Detailed scheme for total internal reflection fluorescence microscopy (TIRF). The scheme is taken form Nikon website: [141]

To do DNA PAINT, total internal reflection fluorescence microscopy (TIRF) illumination is used that has the advantage of only activating dye-labeled imager strands in the TIRF volume. This reduces the background noise coming from imager strands in solution, in addition, depleted dyes from the TIRF volume can be always exchanged with fresh dyes and avoid bleaching. TIRF utilises properties of induced evanescent waves next to the interface between two media with different refractive indices [141] (see figure 24).

A collimated light beam propagating through a medium and reaching an interface, is either refracted as it enters the second medium or reflected at the interface depending on its incident angle and the refractive indices of two media. Total internal reflection would only happen when the beam propagates from a medium with higher refractive index to a lower one (see figure 24). According to the Snell's law the relationship between the angle of incident light and the refractive indices is as following:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{36}$$

in which n_1 and n_2 are the respective refractive indices for the two environments and θ_1 and θ_2 are the angle of incident beam and the refracted beam compared to the normal of the interface respectively. when θ_1 is greater than an angle which is so-called the critical angle θ_c, the incident light cannot pass through the interface and will be completely reflected.

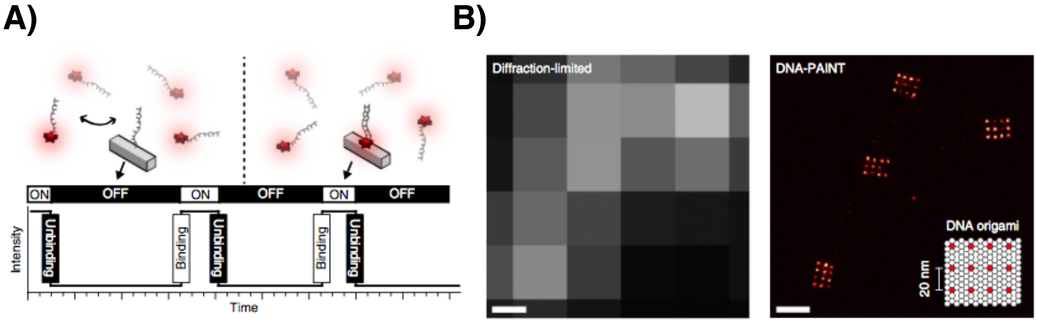


Figure 25: Figure is taken by permission from [142]. Copyright 2017, Springer Nature. A) DNA-PAINT concept. Transient binding of dye-labeled DNA strands (imagers) to their complementary target sequence (docking site) attached to a molecule of interest. The transient binding of imager strands is detected as blinking, illustrated by the intensity versus time trace. B) Diffraction-limited (left) and super-resolved DNA-PAINT images (right) of DNA origami nanostructures. Each structure consists of 12 docking strands that are arranged in a 20-nm grid (scheme in lower right corner).

$$n_1 \sin\theta_c = n_2 \tag{37}$$

Although the incident beam will not pass through but the reflected beam will create evanescent waves very close to the interface in the medium with lower refractive index. These electromagnetic waves can penetrate up to a few hundred nm and will decay exponentially in intensity with distance from the interface. Therefore, in a TIRF setup excitation of fluorescent dyes will be limited to a region which has a thickness of a few hundred nm.

As it is mentioned above, DNA PAINT monitors short dye-labeled DNA strands that transiently bind to docking strands protruding from DNA origami structures. Since the DNA origami structures are attached in the interface, labeled imager strands which bind to them will be excited by evanescent waves and create a blinking event. Figure 25A shows the intensity vs time trace for binding and unbinding of imager strands. Following individual hybridisation events on DNA origami structures in real time, one can calculate k_{on} association rate and k_{off} dissociation rate [140]:

$$k_{on}c = \frac{1}{\tau_d} \tag{38}$$

$$k_{off} = \frac{1}{\tau_b} \tag{39}$$

where τ_d and τ_b are dissociation time (off time) and binding time (on time) respectively and c is the imager strands concentration. k_{off} is related to the length of duplex formed by the imager and docking strands. Therefore we can adjust these parameters to pass the diffraction limit and get down to 5 nm resolution [143]. Figure 25B is an example image of 20 nm grids with DNA PAINT method [144].

Figure 26 shows the workflow in a DNA PAINT experiment [143]. In such an experiment DNA origami structures are attached to a glass slide via BSA-biotin, streptavidin, biotin bridge and illuminated using a total internal reflection setup (see figure 26A). To analyse the binding events and do single-molecule localisation, each binding event in a frame will be fitted to a 2D Gaussian point spread function (PSF) (figure 26B). The localisation precision is related to the number of photons as following [145]:

$$\triangle x \propto \frac{1}{\sqrt{N_{photons}}} \tag{40}$$

Therefore, getting maximum number of photons per localisation narrow the gaussian and increase the localisation precision. One could achieve this by using higher exposure times per frame, stronger laser power or dyes with higher quantum efficiency, however, it should be noted that having a very high exposure time in comparison to the blinking frequency of dyes would also result in having long off times in a frame which decreases the signal to noise ratio.

In the end, all these parameters need to be carefully tuned. Another important parameter that influences the resolution is the number of frames. Having higher number of frames will increase the number of localisations per binding event and improves the signal to noise ratio.

Figure 26C shows an example of double-blinking event or false localisation. False localisation occurs when two blinking events happen at the same time in a diffraction limited area.

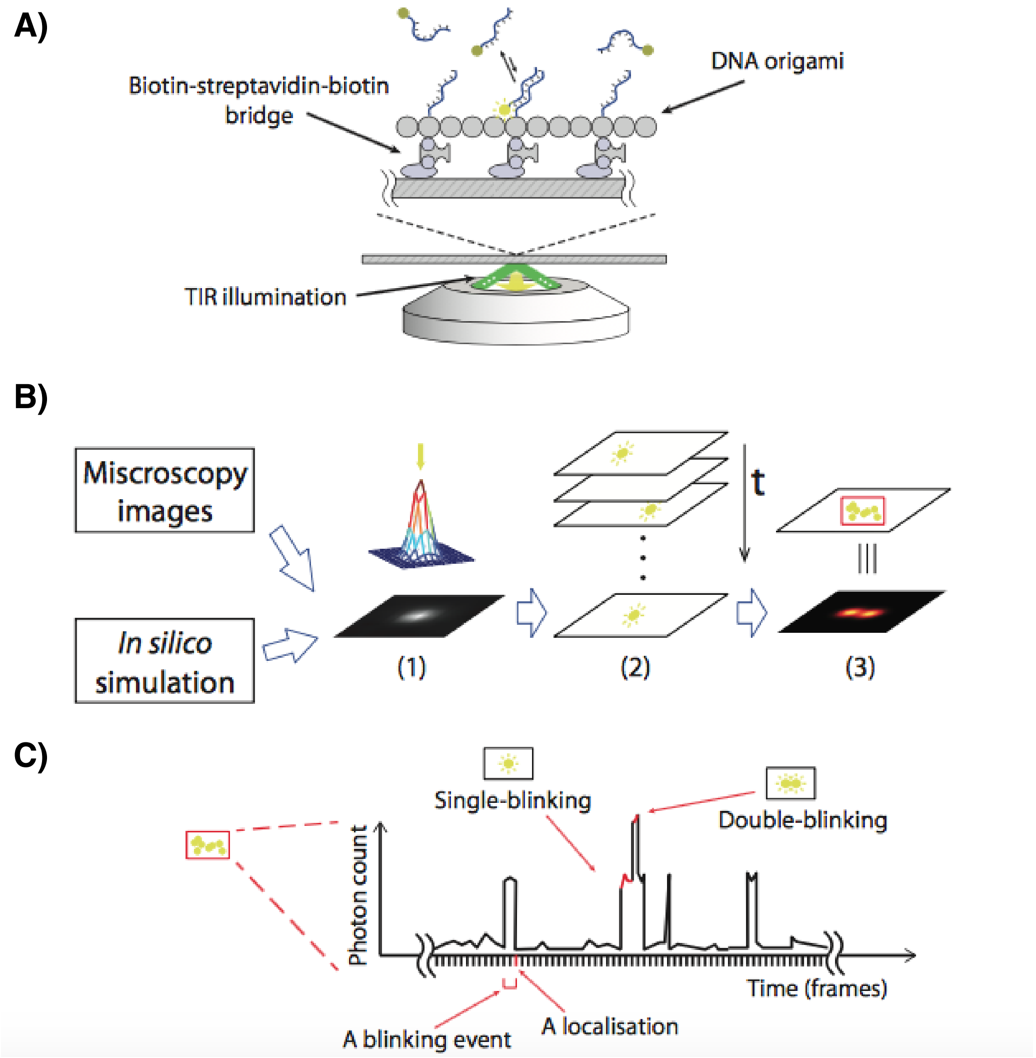


Figure 26: Figure is taken by permission from [143]. Copyright 2016, Springer Nature. A) DNA-PAINT experiment setup with DNA origami samples, illustrated in cross-section view. DNA origami nanostructures are attached to glass slide surface via BSA-biotin, streptavidin, biotin bridge, and illuminated under objective total internal reflection (TIR) setup. B) Schematic illustration of super-resolution image analysis workflow. Super-resolution movie is first analysed in three steps, (1) single-molecule localisation, where each localisation is fit to a 2D Gaussian point spread function (PSF), (2) super-resolution rendering, where, all super-localised centres from all imaging frames are combined and rendered as the final super-resolution image, (3) imaging analysis, histogram projection and kinetic analysis are performed on the super-resolution image. C) Illustration of a typical blinking trace within a certain area in the sample plane. Each blinking event produces a series of consecutive localisations. Double-blinkings from close-by targets give higher photon counts per localisation, and result in false localisations.

High concentration of imager strands which results in shorter τ_d could cause this and thats why it is very important to tune this parameter to improve the localisation precision.

3.5.1 Parameters and setup

For the experiments during this thesis, following parameters are applied:
-Number of frames: 10000 to 15000
-Exposure time for each frame: 400 to 500 ms
-Laser powers: 100 to 140 mW for red and 20 to 30 mW for green
Regarding the concentration of 2 x 2 rings, about 600 pM of them with roughly 300 pM of 20 nm grids as drift markers is used. In general for any sample combination (20 nm grids with 10 nm grids or rings) between 600 pM to 1 nM total concentration seemed to give good results. The imager strand concentration used is in most of the cases 1 nM. It is important to point out that for each setup and experiment, it is crucial to tune each of these parameters. Furthermore using low DNA binding Eppendorf tubes and pipette tips for better reproducibility is strongly recommended. For the image analysis Picasso software is used that has been published in Joerg Schnitzbauer et al. [142]. For the setup, an Olympus IX71 microscope with a 100× 1.45 NA objective and an Andor IXON 897 emCCD camera are used. A realtime autofocus system is also integrated in the setup which is necessary for long imaging durations and details about it can be found in Enzo Kopperger's PhD thesis [146].
Our DNA PAINT measurements are done using the following imager strand sequences:
P1: TAGATGTAT-cy3b
P2: ATGTAGATC-cy3b
P3: TAATGAAGA-Atto655
The usual sample preparation for DNA PAINT measurements is used here.

4 Results and Discussion

In recent years, DNA origami has been used as a bottom-up approach to build programmable functional materials with almost arbitrary geometries in 2D and 3D [2, 3]. In addition, the sequence programability of DNA origami structures enables creating extended crystalline and non-crystalline lattices. DNA origami structures can be regarded as scaffold objects that provide the possibility of chemical modification with 5 *nm* precision. The modification flexibility and precision of such structures, make them great candidates to design molecular binding assays for proteins and inorganic nanoparticles.

In this chapter, first the programmability of DNA origami structures is used to build DNA-based cavities with the potential to precisely control geometry and mechanical properties of multiple binding sites to design molecular binding assays. To this end, a square-shaped DNA origami structure with a cavity in its centre will be presented. Furthermore, the methods to assemble this structure into bounded and unbounded two-dimensional arrays will be discussed. Moreover, a barcoding scheme will be established with which molecular binding assay experiments are multiplexed. Eventually the assembled arrays of DNA-binding cavities along with the barcoding scheme are applied to study two protein model systems.

In the end of this chapter, an attempt towards creating the Penrose tessellation employing programmability of DNA origami structures is explained.

4.1 Barcoded DNA origami structures for multiplexed optimization and enrichment of DNA-based protein-binding cavities

Results in this section are mainly published in Ali Aghebat Rafat et al. [123] and the section is based on that.

DNA nanostructures in principle permit the spatial modelling of more complex artificial binding sites, in which the geometry of the molecular interactions is controlled precisely. Moreover, the mechanical properties of DNA, in which single-stranded molecules are considerably more flexible than double-stranded ones, can be tuned to a certain degree. Apart from tuning the geometry of the binding site it is thus also possible to control its flexibility, which is expected to have a profound influence on the overall binding strength [6, 21, 22, 28].

4.1.1 Design and assembly of square-shaped DNA origami structures

DNA origami structures with a central cavity are designed (figure 27), within which several aptamers can be placed at multiple locations, in different geometric relationships, and with variable flexibilities. The routing of the scaffold strand through the structure creates in its centre a symmetric square-shaped cavity, which has a side length of approximately 24 *nm*. The cavity is lined with 12 uniquely addressable staples that can be modified with arbitrary DNA aptamers. Using well-known aptamers for thrombin and streptavidin as examples, it is going to be shown that these parameters dramatically impact the binding of proteins to the cavities, and that their optimization results in unprecedented binding efficiencies.

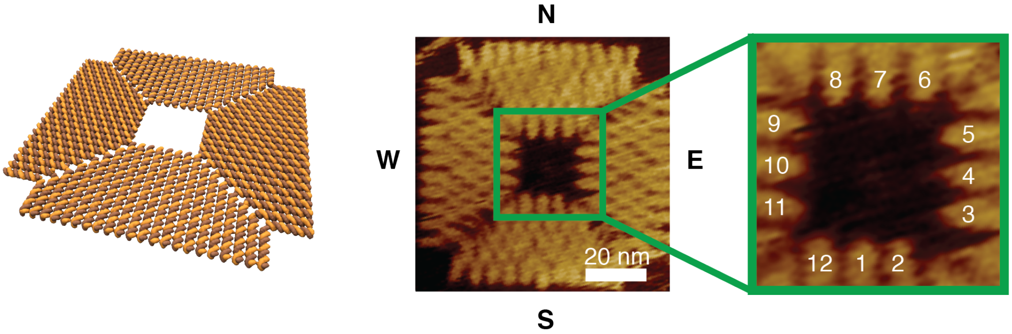


Figure 27: Molecular model of the DNA nanostructure and a detailed AFM image shows the origami structure and its cavity. For orientation the outer edges of the structure are addressed with North, South, East and West. The zoomed-in image of the symmetric cavity shows the 12 staples available for modification, and their respective numbers. These staples are used to arrange the DNA aptamers in a controlled orientation and position.

4.1.2 Assembly of square-shaped DNA origami structures into crystals

The square-shaped origami platform with an approximate four-fold rotational (C4) symmetry (figure 27), is based on a design previously used for the assembly of extended DNA origami lattices [106]. In this design, DNA helices extend from the centre in two perpendicular directions (further denoted by North–South (N–S) and West–East (W–E)), which facilitates multimerisation between origami structures along the helix axis via all four edges, using a combination of sticky- and blunt-end interactions (see section 3.2.1 and 3.2.2).

Two different methods are used here to create two-dimensional (2D) crystals of DNA origami structures. In one method, edges are multimerised along the North–South and East–West axes which resulted in only elongated, one-dimensional assemblies (figure 28A). By contrast, 2D crystal formation proceeded well when the N edge of one origami tile is connected to the E edge of the neighbouring tile and, correspondingly, the W edge to the S edge, which presumably results in a cancellation of the intrinsic twist of the monomers (28B). Later the well-formed 2D crystals are used to create crystal patterns of the thrombin and streptavidin.

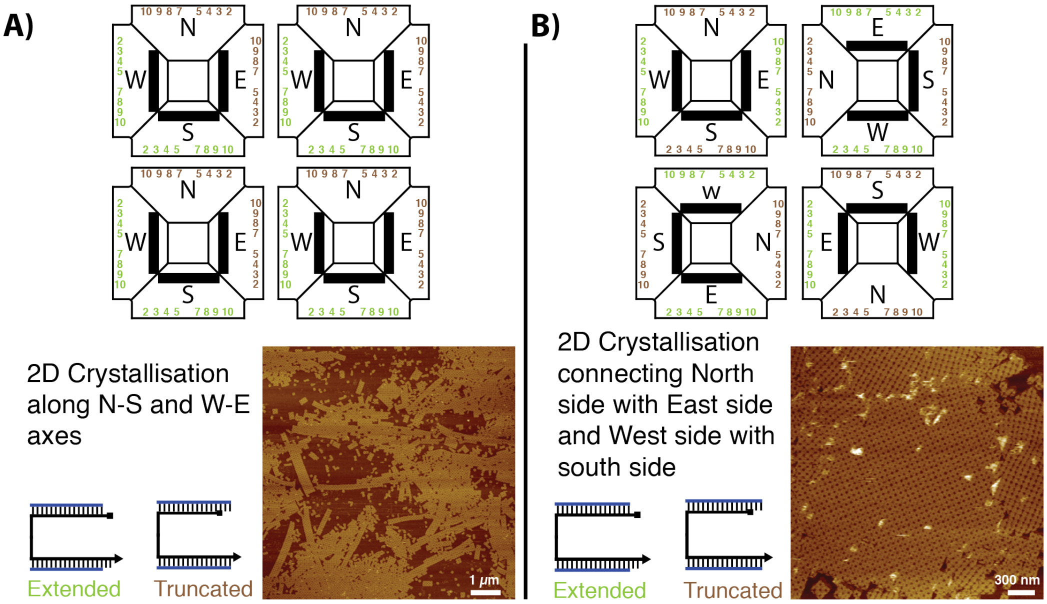


Figure 28: A) Scheme and AFM image of the 2D crystallization by using parallel sides (N–S, E–W). As can be seen in the AFM image, elongated crystals are formed using this type of connection rule, potentially due to the build-up of internal strain. Color codes and staple coordinates show the type of staple modification (truncation and extension) and the staples which are used in an edge respectively. B) Scheme and AFM image of the 2D crystallization with 90° rotation (N–E, W–S). Color codes and staple coordinates show the type of staple modification and the staples which are used in an edge respectively. Black lines on the trapezoids are signs to show how the structures are multimerised.

4.1.3 Assembly of square-shaped DNA origami structures in 2 x 2 arrays

In the binding assay experiments here, structures that comprised multiple cavities, in particular finite 2 x 2 arrays with four cavities are used. The fractal assembly strategy [108] (see section 2.5.2) is adopted to generate 2 x 2 arrays that can potentially contain different combinations of aptamers within their four cavities. The advantage of employing the fractal assembly technique is that each single tile is separately folded in a different tube. Therefore, single origami structures can have their own specific property without intervening each other. This enables direct comparison of the binding yield of different aptamer configurations under identical conditions in a single experiment. In figure 29A, Connection rules and type of staples that are applied to create 2 x 2 arrays are shown (detailed discussion in section 3.2.2). Example AFM images of the 2 x 2 arrays are shown in figure 29B. The yield of 2 x 2 arrays calculated from the AFM images is about 70%.

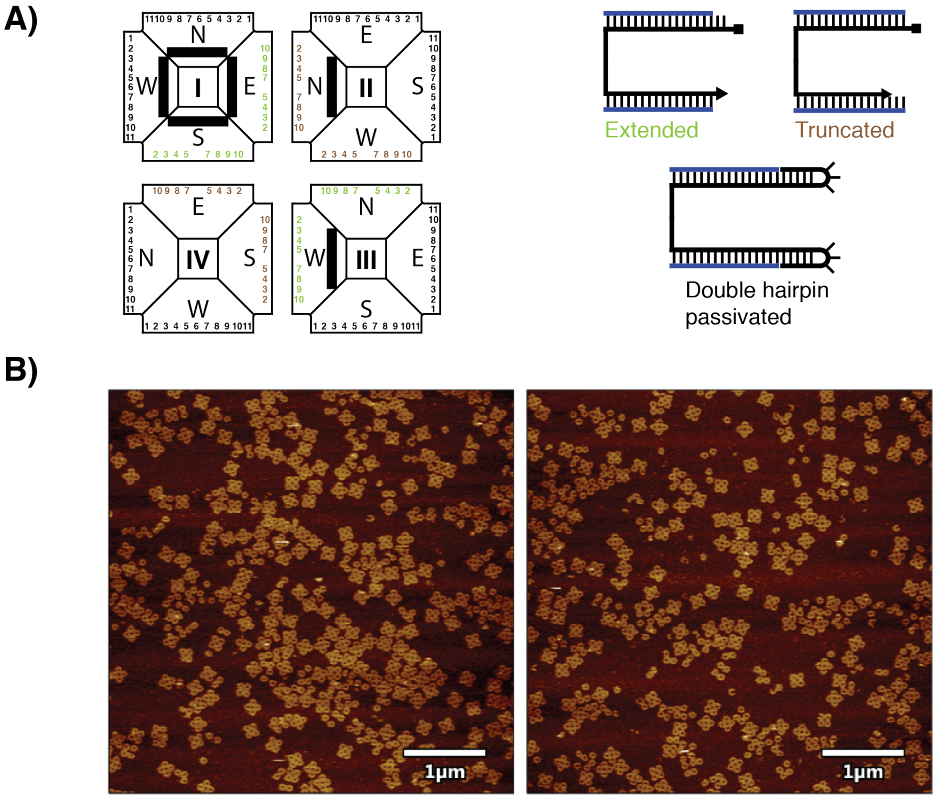


Figure 29: A) 2 x 2 array scheme which shows the connection rules between different sides. Color codes in addition to the staple coordinates show on each side of the square (North, South, East, West) which staples are used and what kind of modification (truncation, extension and double hairpin passivation) they have. Black lines on the trapezoids are representing a barcode example. B) AFM images representing 2 x 2 arrays. The yield of the formation of 2 x 2 arrays derived from these images is about 70%.

4.1.4 Barcodes and patterns on DNA origami structures

Staples in DNA origami structures can be considered as pixels with a size of 5 *nm*. It is possible to modify these staples to create certain patterns or barcodes with 5 *nm* resolution. Such patterns or barcodes can be used to either recognise DNA origami structures in an imaging technique or create specific artistic patterns on them [108]. In figure 30A, a U-shaped pattern is created on single DNA origami structures which enables visualisation of the connection between monomers in the crystal.

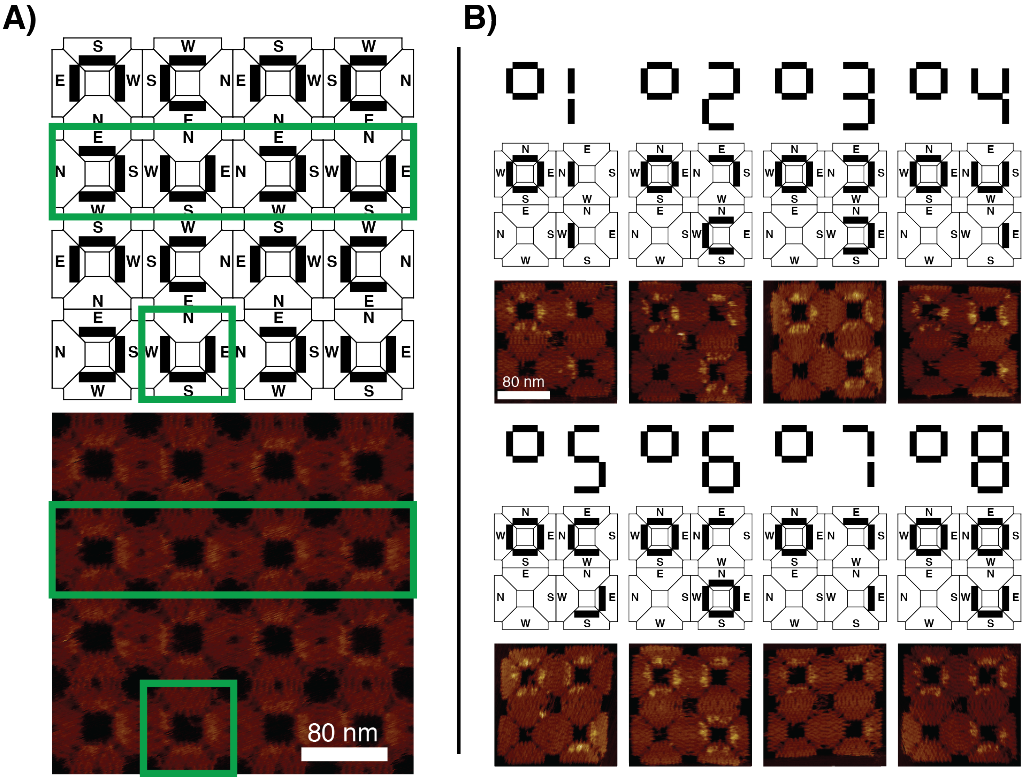


Figure 30: A) Scheme and AFM images of a 2D crystal generated from square-shaped DNA origami structures. U-shaped barcodes are used to visualise the connection between monomers in the crystal. The green boxes in the scheme and small AFM image highlight a single structure and a line of structures within the crystalline assembly B) Schemes and AFM images of the 2 x 2 arrays and their respective barcoding. The barcodes represent digital numbers that label the uniquely addressable arrays.

For the molecular binding assays, to increase the throughput, a barcoding scheme is introduced to distinguish between up to eight different 2 x 2 arrays and their single monomers in one AFM experiment (figure 30B). Ten or eleven staples on each of the four trapezoids that comprise each square are modified with dumbbell hairpins [2] to provide contrast during the AFM imaging. The barcode labels are designed in an asymmetric fashion, which enables differentiation between structures that are oriented face up or face down on the mica substrate during AFM imaging. (a detailed description of the barcoding scheme can be found in section 3.2.2). In principle more barcoding schemes are possible to be made but here only eight of them are required.

Combining the barcoding scheme together with the 2 x 2 arrays enables doing different experiments in parallel under identical conditions. Prior to the work here, for such binding experiments usually different measurements are done in different sessions and then their results are compared. The problem is that random errors in experiments can result in less reliable comparisons, moreover, enormous workload is needed to try all different parameters in different experimental sessions.

DNA PAINT to image patterns and color them As it is discussed in section 25, DNA PAINT technique employs fluorescently labeled DNA probes in solution to image below the diffraction limit [140]. This process happens by stochastic transient binding of DNA probes to the docking strands on the surface of DNA origami structures. It is also possible to use the DNA PAINT method to image patterns on DNA origami structures.

Figure 31A shows the AFM image of rectangular origami structures. As an example of imaging patterns on DNA nanostructures using DNA PAINT, 20 *nm* and 10 *nm* grids are created on the rectangles. To design these grids on the structures, staples which have 10 *nm* or 20 *nm* distance from each other are extended with the complementary sequences to the DNA probes. Figure 31B represents the schemes and super-resolution images of the grids on the rectangles. As it is clear from the DNA PAINT images not all the docking strands or extensions seem to be present. This issue is usually observed in the DNA PAINT measurements and there can be different reasons for it. One possible reason is the quality of both probes and the docking strands that are ordered. For example, ordering HPLC purified docking strands improved the number of localised points observed in the DNA PAINT image. This originates from sensitivity of DNA PAINT to the sequence of docking strands. If the sequence has a defect, the binding kinetics will be very different.

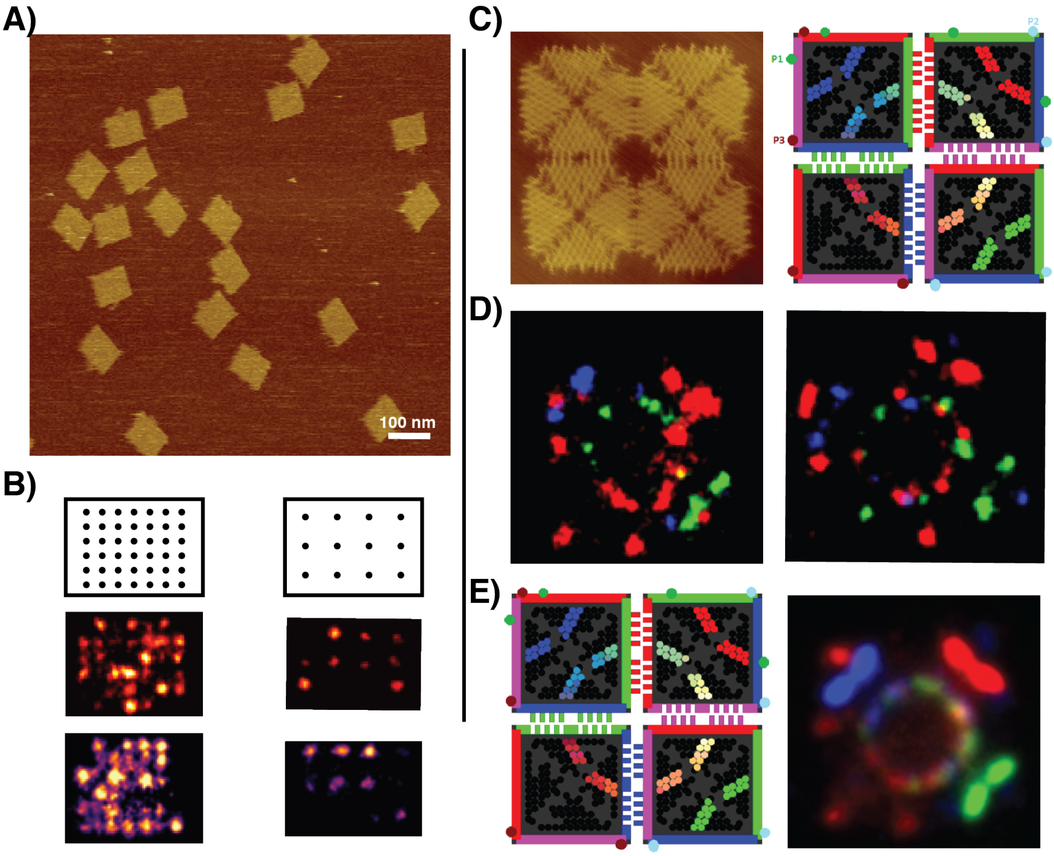


Figure 31: A) AFM image of rectangular DNA origami structures [2, 130]. B) Schemes and DNA PAINT super-resolution image of 10 and 20 *nm* grids created on rectangular origami structures. C) AFM image and scheme of a 2 x 2 array made in Qian lab containing to create a rainbow ring on the array. The coloured arcs and spots at the corners and edges are used for alignment. The scheme is made by Philip Petersen [147]. D) Example DNA PAINT super-resolution images of two arrays imaged with P1 (green), P2 (blue) and P3 extensions (red). For each ring, the three colors are aligned using the fiduciary markers using the Picasso software [142]. E) Scheme and aligned DNA PAINT images of the ring obtained from tens of 2 x 2 arrays imaged with different extensions using Picasso software.

Therefore, such docking strands do not have sufficient photon counts and will not be observed in the image. More detailed discussion on accessibility and incorporation of docking strands is presented in Maximilian T. Strauss et al. [144].

Coloring patterns with DNA PAINT It is shown in section 2.5.2 that using fractal assembly one can create different size of arrays with certain patterns on them. As a collaboration with Qian lab at Caltech, we decided to color these pat-

terns using DNA PAINT. The motivation behind these experiments is quantification of the dynamic behaviour of molecular robots at the single molecule level. Imaging with DNA PAINT in different channels (red, green and blue) could potentially enable analysis of different robots in parallel in their distinct location. To this end, 2 x 2 arrays with DNA PAINT docking strands on them are used (AFM image in figure 31C). As it is shown in the scheme in figure 31C, 68 staples which represent the rainbow ring in the scheme can have different P1, P2 or P3 extensions (see section 25 for details on extensions). Different probabilities for presence of each extension in a staple position is used to produce the specific color in that position.
Using only three types of receptors enable humans to identify different colors. In trichromatic RGB graphic systems, each pixel includes 8 bits for each color which results in 24-bit RGB pixels (or 256 divisions for each color) that can encode roughly 16 million different colors [148]. Each arbitrary color in such a system can be encoded as a combination of three primary colors (R,G,B). For example red:(255,0,0), green:(0,255,0) and yellow:(255,255,0). The amount of each primary color represents its intensity. When all three colors are at their max:(255,255,255), the resulting color will be white and if they are all zero the result is black. Equal proportions of less intensity is going to create different shades of grey.
In our system on the 2 x 2 arrays, we used a system with 6 divisions. each pixel in our case is one staple which can have P1 extension (green), P2 extension (blue), P3 extension (red) and no extension (P1,P2,P3,no extension). In our system red:(5,0,0,10), green:(0,5,0,10) and yellow:(5,5,0,5). Staples with no extension are needed to maintain the concentration ratios. Pipetting all these combinations and folding the 2 x 2s (the pipetting is done by Echo robot) result in stochastic rings that have different extensions in different positions. The sample is imaged for P1 and P2 in the green channel and for P3 in the red channel. Here 20 nm grids are used as drift markers during imaging.
Since the DNA PAINT image is going to have different rings in different channels that need to be aligned relative to each other, surface fiduciaries are designed on the corners and at the edges of 2 x 2s (scheme in figure 31C). Using the surface fiduciaries, first of all rings from one channel are alined with respect to each other and then the resulted aligned rings from the three channels are aligned to each other to create the final image. Figure 31D represents examples of 2 x 2 arrays or rings that are imaged in three channels. Tens of rings from each channel are taken and aligned using the picasso software [142] to produce the final rainbow ring in figure

31E. As it is clearly seen from the average aligned DNA PAINT image, the result fits quite good to the scheme. In principle, imaging more rings could potentially give a closer average DNA PAINT image to the primary model. One would need to explore the number of 2 x 2 arrays or rings that are needed to be imaged in order to obtain a sufficiently precise final DNA PAINT image.

4.1.5 Studying aptamer orientation in thrombin-binding cavities

Square-shaped DNA origami structures with a central cavity were previously discussed in this chapter. It is shown in figure 27 that the cavity includes 12 uniquely addressable staples that can be modified with arbitrary DNA aptamers and bind to a protein target. Using these DNA-binding cavities along with the fractal assembly method and barcoding, multiplexed molecular binding assays are designed.

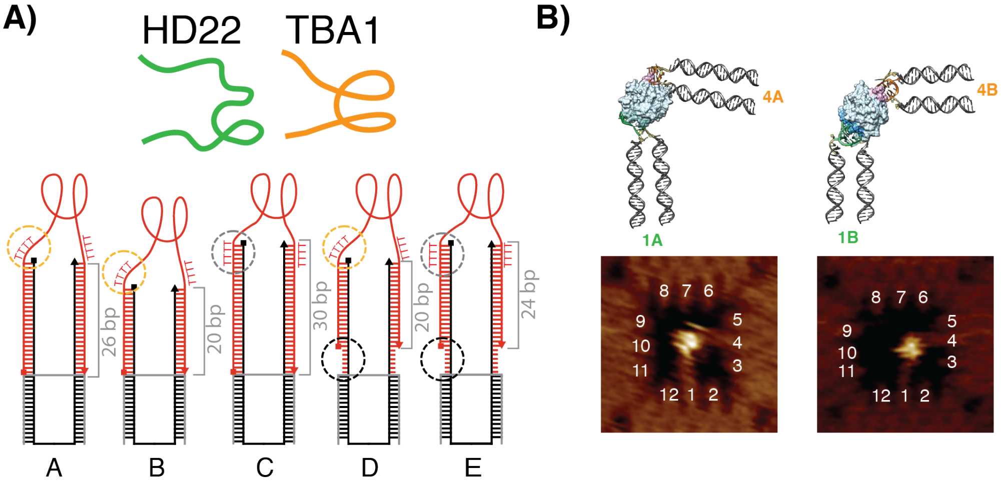


Figure 32: A) Configurations to assess the influence of spacer lengths and flexibilities with HD22 (shown in green) and TBA1 (in orange). The spacers are composed of a rigid double-stranded stem and optional flexible linkers to the origami and the aptamer structure (the staple and aptamer strands are in black and red, respectively). The length of the stem is varied from 20 bp to 30 bp. The connection between the stem and the aptamer is made through four thymidines either as a single-stranded “aptamer linker” (cases A, B, D; orange circle) or incorporated into the stem (cases C, E; grey circle). Cases D and E also have six unpaired nucleotides between the origami structure and the stem to introduce a flexible “stem linker” (black circle). B) Chimera models and AFM images of the cavity with HD22 and TBA1 designed according to cases A and B and placed in positions 1 and 4, respectively. Both AFM images and Chimera models show that aptamers in case A need to reconfigure themselves to enable binding to their respective sites on thrombin.

At first, the effect of geometry and flexibility of the binding site by using thrombin as a model target is investigated (figure 32A). Thrombin has two well-characterized aptamers named TBA1 [149] and HD22 [150] (figure 32A), which target the spatially separated thrombin exosites I and II, respectively (figure 32B). In one set of experiments, HD22 aptamer is placed on binding position 1 of the cavity and systematically changed the position of the TBA1 aptamer by attaching it to each of the other 11 available positions (figure 33B). This enables screening different orientations of the two aptamers with respect to each other in two dimensions.

To accommodate the target protein in the cavity, the aptamers are attached to the origami frame through DNA spacers, which consisted of a rigid, double-stranded stem and optional flexible linkers between origami and stem (the “stem linker”) and between stem and aptamer (the “aptamer linker”). In the same experiments, it is therefore possible to test different spacer lengths and flexibilities at the stem and in the vicinity of the aptamer loop to examine how these parameters influenced the binding of thrombin to the cavity (figure 33C, appendix figures 44 and 45).

For instance, in case A, aptamers are attached to the origami through a 26-bp-long, rigid stem, whereas the aptamers are connected to the stem through a flexible stretch of four (deoxy) thymidines (figure 32A). By contrast, case B had a stem of only 20 bp (shorter by approximately one half-turn of a double helix), which also results in a rotation of the aptamer connection with respect to the helix axis of the stem. A molecular model constructed with UCSF Chimera software suggests that when the TBA1 aptamer is placed at position 4, in case A the stem would have to bend outwards to facilitate binding of the aptamers to the thrombin exosites, whereas in case B the bending would be slightly inwards (figure 32B). Remarkably, the structures that are expected from these models are confirmed by the AFM images. To determine the best binding configuration for case A, the results for the 11 different configurations with HD22 fixed at position 1 (figure 33B) are compared. Statistical analysis of more than ≈200 tiles per configuration showes that an angle between the aptamers of approximately 90° in the cavity plane (positions 1 and 4) is the best, with a binding yield of 39%. Regardless of the design of the spacer (cases A–E listed in figure 32A), it is found that a 90° configuration (positions 1 and 4) always lead to a higher yield than the corresponding 180° configuration (positions 1 and 7) (figure 33C).

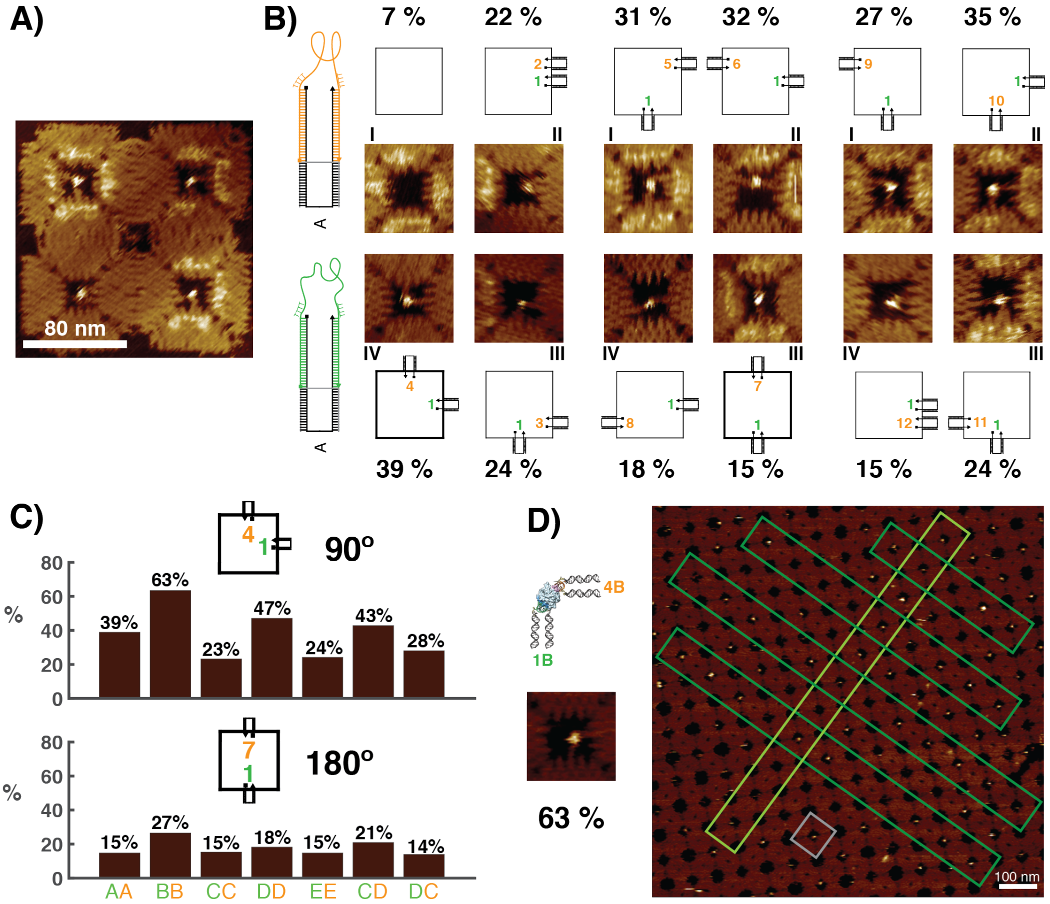


Figure 33: A) Example of an AFM image of a 2 x 2 array with thrombin bound in the cavities. B) Schemes and AFM images of the 11 different configurations of aptamers, arranged in groups of four. HD22 is always in position 1 and the TBA1 position in the cavity varies. All aptamers are attached as in case A. Statistical analysis shows that the 90° (1,4) configuration results in the best binding yield of these configurations. The binding assay for all the configurations is performed in a single pot and all are imaged simultaneously. C) Yield of thrombin binding to its aptamers in 90° (1,4) and 180° (1,7) configurations. The effect of aptamer flexibility at the stem and in the vicinity of the aptamer loop for these configurations are compared in the histogram. The 90° (1B,4B) configuration has the overall best yield. D) Formation of a 2D crystalline array of the DNA origami structures with the best configuration for binding to thrombin (a scheme and small AFM image of this configuration is shown on the left). The grey box highlights a single structure in the lattice and the dark green and light green boxes indicate linear arrangements within the crystal. All percentages are derived from single AFM experiments to enable direct comparison under identical conditions (no error bars are given because n = 1).

For the analysis and comparison of binding yields, single AFM samples are deliberately in focus, which are prepared from a single batch of origami cavities and proteins. To improve the statistics, the same sample is imaged at multiple locations. Identification of the barcoded structures then allows a direct comparison of different binding configurations and spacer designs under identical conditions. The reproducibility of the binding yields is also assessed that are obtained from different sample preparations and AFM imaging sessions. A variability on the order of only a few percent is found (absolute standard error of the mean typically $\leq 3\%$; see appendix figure 46).

4.1.6 Influence of aptamer linker length and flexibility on binding strength

Next, the effect of stem length and different flexibilities is studied by comparing various corresponding configurations. For example, the influence of flexibility at the aptamer linker may be assessed by comparing results for cases AA (both aptamers connected through A-type spacers) and CC (both aptamers connected through C-type spacers), or for cases DD and EE. For the 90° configuration, flexibility at the aptamer loop improves the binding yield by about 15–20% (absolute) in each case. To assess the influence of stem linker flexibility, the results for cases AA and DD, or CC and EE, can be compared. Cases CC and EE have a very similar binding yield, whereas case DD appears to perform better (leads to a higher yield) than AA by about 8% (figure 33C and appendix figure 44).

Collectively, these results suggest that stem linker flexibility is helpful only when there is also flexibility at the aptamer linker. The cases of mixed linker flexibilities for the aptamersHD22 and TBA1 are also investigated. The results suggest that the mode of attachment of HD22 does not change the binding yield substantially (variations of $\sim$5%). However, introducing flexibility into the TBA1 aptamer linker results in a 20% improvement in binding yield.

To understand these findings on a more detailed geometric level, the distance between the positions at which the TBA1 aptamer loop connects to the terminal bases of the stem duplexes is calculated (figure 34A and 34B). For a stem length of 20 bp (as in case B), this distance is around 2.7 nm, which is sufficiently close to the optimal distance for the TBA1 aptamer to form ($\sim$1.3 nm [124]) when given enough flexibility by the additional thymidine bases (each thymidine is around 0.4 nm).

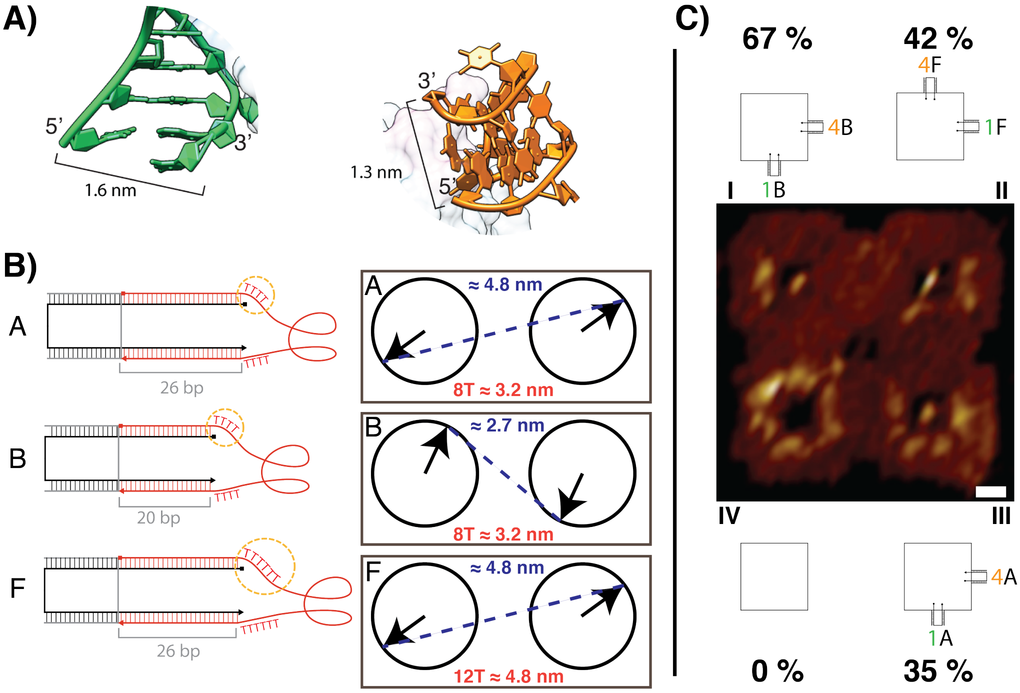


Figure 34: A) Distances between the 5' and 3' ends of the aptamers. 3D models rendered in Chimera using as model PDB IDs: 5EW1 [124] and 1HUT [151]. HD22 (green) and TBA1 (orange). B) The length of the spacer stem varied from 20 bp to 26 bp. The single-stranded linker between the stem and the aptamer consists of four (case A and B) or six (case F) thymidines. The boxes show the DNA orientation in the double helix and the distance between them. This distance is larger in the case of a 26 bp spacer. C) Percentages denote the binding yield in each configuration and the negative control for the nanostructure number IV. Schemes and AFM image of the three configurations of aptamers are shown. Structure number IV has no aptamer as a negative control. HD22 is always in position 1 and TBA1 in position 4.

By contrast, for cases with a 26 bp stem (cases AA and CC), this distance is around 4.8 nm, which cannot be compensated by four thymidines alone. As a consequence, mechanical stress is applied on the aptamer fold, which then results in a reduced binding yield. To test this hypothesis further, a structure with six thymidines for each stem in case A is created, which indeed improves the binding by about 7% (figure 34C). In addition, 2D crystals from square cavities equipped with the best-performing aptamer configuration is created (the 90° 1B,4B configuration), which enabled the arrangement of thrombin proteins into a crystalline pattern at a high yield (figure 33D).

Note that when using the 2 x 2 arrays, there is no attempt to determine apparent

Kd values for the various configurations. As in previous studies [20, 24], the aptamer binding yield observed in AFM on DNA origami structures is considerably lower than the yield that is expected based on the Kd measured on isolated aptamers by using other techniques. For instance, with a previously determined Kd $\approx 0.5\ nM$ for the HD22 aptamer, one would naively expect a binding yield of ∼100% under the conditions used here; however, a negligible binding in AFM is observed when a single HD22 aptamer in the origami cavity is used (figure 35).

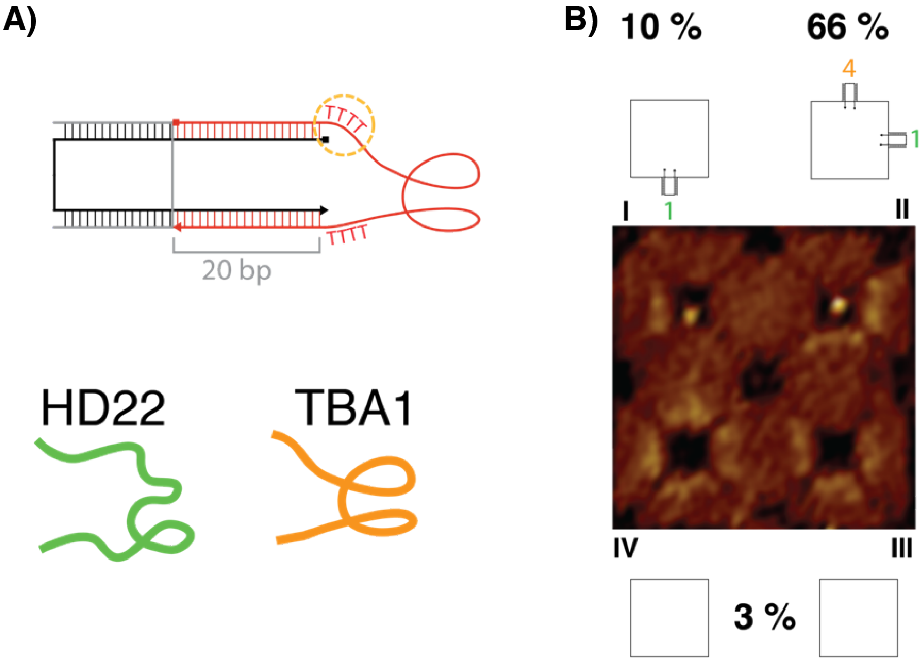


Figure 35: A) the length of the spacer is 20 bp. The single-stranded linker between the stem and the aptamer is composed of four thymidines. B) Percentages denote the binding yield in the two configurations and for the nanostructures number III and IV negative controls. Schemes and AFM image of the two configurations of aptamers are shown. HD22 is always in position 1 and the TBA1 position 4.

By contrast, the barcoded 2 x 2 arrays enable a direct comparison of different configurations under identical experimental conditions, which enables ranking the aptamer configurations reliably in terms of binding yields independent of their absolute values, and therefore to optimise the binding cavities. This also avoids problems that arise from uncertain effective protein concentrations and qualities, which are found to be a source of considerable variability in overall binding yield ($> 10\%$ for different protein batches).

In general, the findings here correspond well with the expectations for the role of rigidity and flexibility in molecular interactions. Rigid linkers are preferred when the distance between the binding partners is well matched and does not require extensive stretching or bending. However, if the geometry of the binding site does not match the optimal molecular distances for the binding interactions, flexible linkers may facilitate binding, whereas rigid linkers will not. A rigid binding site with ideal distances would perform even better, but a sub-optimal binding site can be improved by introducing flexibility.

4.1.7 Streptavidin-binding origami cavities with four identical aptamers

Later the molecular binding assay is applied to streptavidin as another model system, using the aptamer-functionalized square cavities. As with thrombin, first streptavidin and DNA-based cavities are incubated free in solution before depositing the nanostructures on mica. As streptavidin is a tetrameric protein, a total of four DNA aptamers are incorporated into the cavity, each of which is expected to interact with a single monomer. Several laboratories had previously selected aptamers for streptavidin, most of which shared a stem-dumbbell structure as a common secondary structural motif that is surmised to bind specifically to the same site on the protein. A minimal, 40-nucleotide-long, streptavidin-binding motif (termed streptavidin aptamer SAA1 here) is evolved by non-homologous random recombination and has a reported Kd of 62 ± 6.6 nM and 105 ± 21 nM with and without flanking primers, respectively [122].

The binding of streptavidin to SAA1 aptamers (without flanking primers) that are attached to the cavity through 26-bp-long (case A) and 20-bp-long (case B) stems and with 4-nucleotide-long aptamer linkers is investigated (figure 36A). As in the experiments with thrombin, 2 x 2 arrays are folded with their respective barcodes, each representing a different spatial configuration of the four aptamers with different spacers (an example is shown in figure 36B; appendix figures 47, 48 and 49 show images of all eight barcodes).

Comparison of the results for cases A and B shows that for identical spatial configuration, an attachment with a 20-bp-long stem always results in a higher binding yield (figure 36C). As with thrombin, the greater distance between the connection points between aptamer and 26 bp stems may result in mechanical stress that disfavours proper folding of the aptamer. When the short aptamer stem is used, the

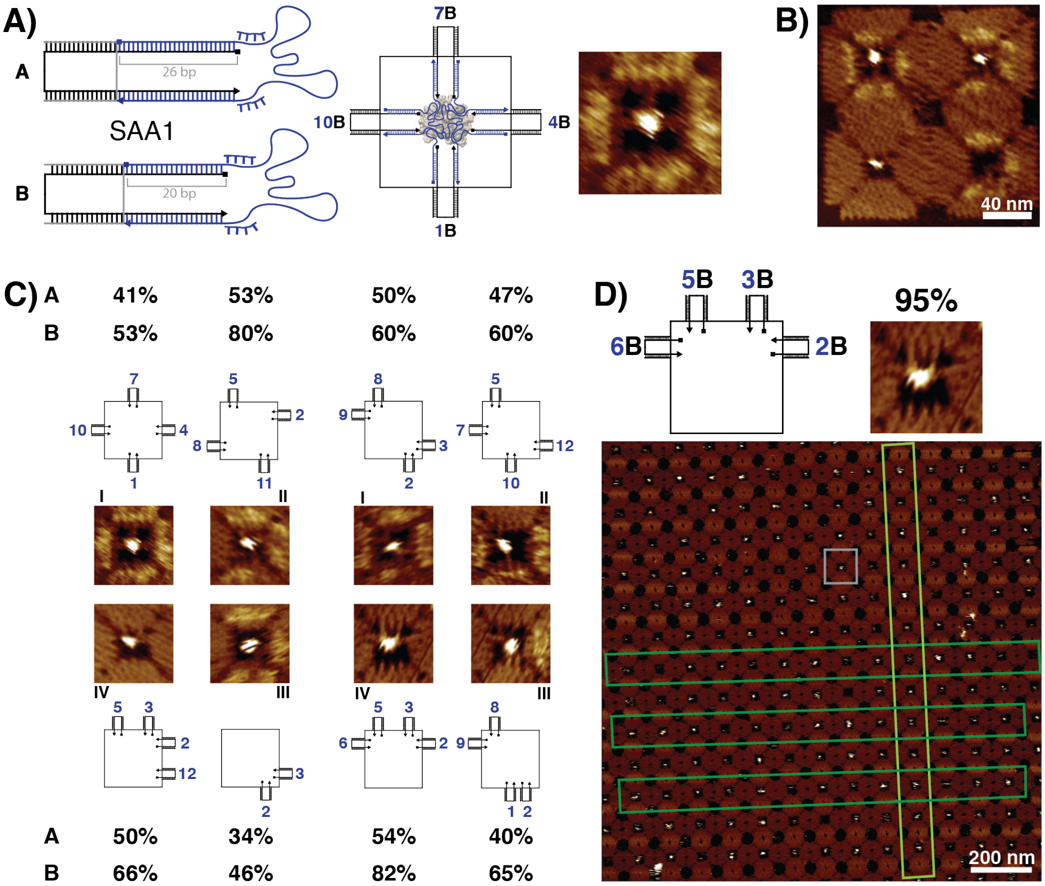


Figure 36: A) Scheme of the streptavidin aptamer SAA1 attached through spacers, each composed of a 26 bp (case A) or a 20 bp (case B) stem and additional four thymidines as aptamer linkers. Right, a 2D scheme and a zoomed-in AFM image of the (1,4,7,10) configuration for SAA1 in case B. B) Example of an AFM image of a 2 x 2 array (barcode 3) that includes different configurations of four SAA1 aptamers in four cavities. C) AFM images and schemes for eight configurations of four SAA1 aptamers binding streptavidin, displayed in two groups of four. The binding yields for both long and short stems for these different configurations are given. Typical errors of the percentages are $\leq$3%. D) Crystal formation of DNA origami structures with the best configuration (2B,3B,5B,6B) for binding to streptavidin. This configuration is shown in the scheme on top and in the small exemplar AFM image, which also appears in panel C. In the lattice shown in the large AFM image, the grey box highlights a single origami structure, and the dark green and light green boxes indicate linear arrangements of structures within the crystal. In the context of the crystal, we observe a binding yield of ~95%.

binding yield ranges from about 50% to up to 80%, which is comparable to results reported previously for streptavidin–biotin interactions on origami [11]. The binding yield for cavities that contained only two aptamers, which nevertheless resulted in a binding yield of 46% is also investigated.

4.1.8 Organizing streptavidin in 2D DNA origami crystals

In the next step, 2D DNA origami crystals by using cavities equipped with the best-performing SAA1 configuration are created. The crystals have an observed yield of ∼82%; this results in a DNA lattice-mediated arrangement of streptavidin into an almost crystalline pattern (figure 36D). Interestingly, at ∼95% the binding yield observed in the context of the origami lattice is higher than that for the isolated origami cavities. To verify the replicability of this result, 2D crystals from origami cavities with a different aptamer configuration are generated. In this case also, a higher binding yield for the crystals (again ∼95%) than with equipped cavities within 2 x 2 arrays is found (79%, appendix figure 50). This result is somewhat surprising because with an inter-origami distance of ∼50 *nm* one would not expect any cooperative action by neighbouring origami cavities within the lattice. Conceivably, the differences are caused by different binding or unbinding kinetics for isolated origami cavities and origami lattices, which would result in different apparent binding yields for the specific incubation and waiting times used in the experiments. Notably, such a difference is not observed for lattices of thrombin-binding cavities.

In addition, experiments that began with deposition of empty DNA origami lattices (equipped with the best-performing aptamer configuration) onto a mica substrate, followed by addition of proteins and incubation in situ are performed (figure 37). It is found that in this case, the binding yield is reduced slightly compared to DNA lattices incubated in solution, but is still substantially higher than for single, solution-incubated origami cavities.

To assess the replicability of the results obtained with streptavidin-binding cavities, binding yields obtained from independent AFM experiments are compared (figure 37) and also the variability within single AFM sessions are studied (appendix figure 51). A statistical Student's t-test showed that the differences measured for different configurations are indeed statistically significant, as standard errors of the binding yields are found to be ∼3% or lower. Although considerable variations among experiments with different protein batches is seen (comparing results in figure 36 and 37), the yields and trends are always consistent within the same batch.

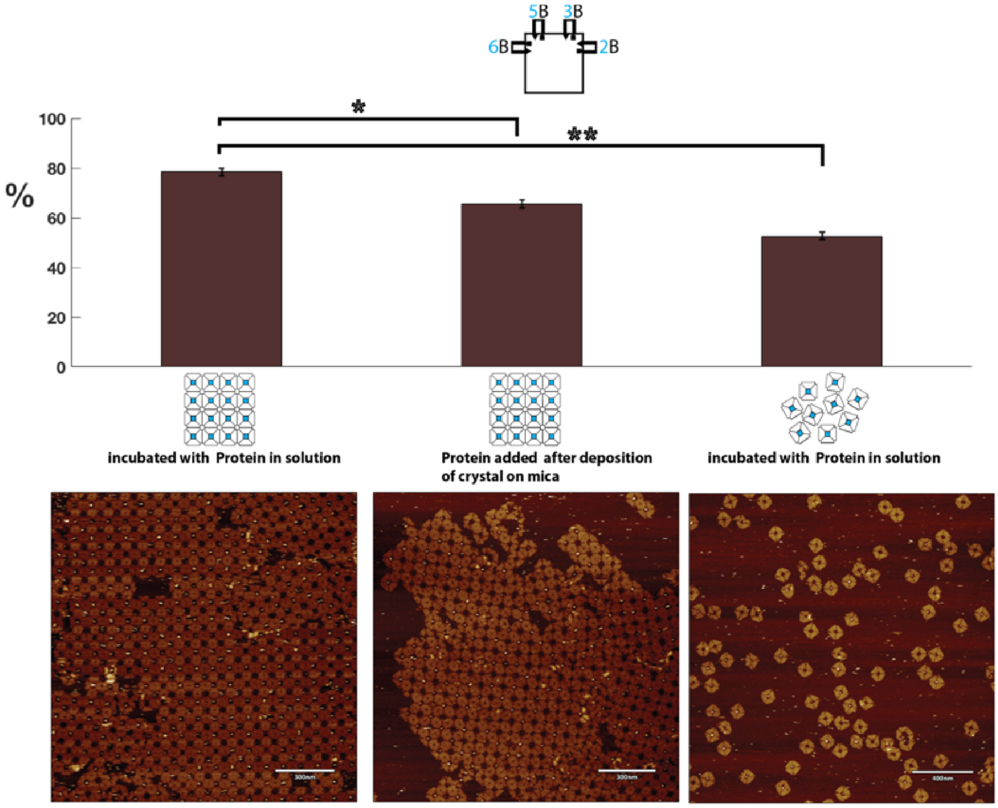


Figure 37: Comparison between the binding yield of streptavidin incubated either with 2D crystal/cavity structures in solution or on mica. The bar graph compares the binding yield for different incubation procedures for the SAA1 aptamer configuration (2, 3, 5, 6) shown on the top: The bar on the left corresponds to 2D crystals, which are incubated with streptavidin for 30 min at 37°C and then deposited on mica for imaging. The bar in the middle shows the binding yield of 2D crystals deposited on mica in the absence of protein, followed by addition of 5 μl of 40 nM STV solution in 1×FB and 60 μl of 1×folding buffer. The sample is then incubated at 37°C for 30 min and imaged afterwards. The right bar represents the binding yield of single structures incubated with streptavidin in solution for 30 min before imaging. The bar graphs state the mean ±SEM (n = 3). For the solution-incubated samples, three independent AFM sessions at different times are performed. For the crystal incubation with protein on the surface, three measurements are performed at three different positions on the mica surface in the same experiment. Analysis via a t-student test indicates statistically significant differences in the binding yields ($^{*}P < 0.05$, $^{**}P < 0.01$). Both 2D crystals give a higher binding yield than the isolated structures, while solution incubation appears to form better than surface incubation. Note that the experiment is performed with a different batch of streptavidin proteins, resulting in overall reduced binding yields. This gives an indication that the largest source of variability in the experiments probably is the protein quality and concentration. Importantly, comparison of barcoded structures under identical conditions is not affected by this variation and allows to make robust statements about binding yield differences.

4.1.9 Enrichment of optimised aptamer configurations from a small library of DNA structures

Ultimately, it will be interesting to explore whether in vitro selection and directed evolution experiments can be generalized to also address the geometric and mechanical properties of ligand-presenting scaffold structures. As a step towards this goal, cavities with good binding properties are "selected" from a pool that contains a mixture of different aptamer configurations.

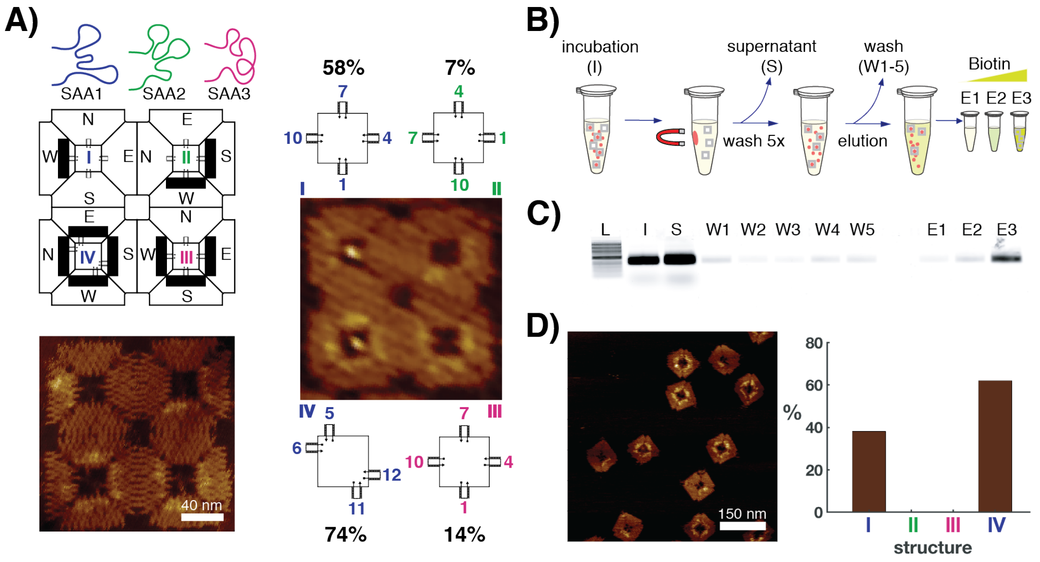


Figure 38: A) Schematic drawing of the three aptamers tested in the 2 x 2 array. Positions I and IV have the SAA1 aptamer located in configurations (1,4,7,10) and (5,6,11,12), respectively. Positions II and III have SAA2 and SAA3, respectively, in the symmetric configuration (1,4,7,10). AFM images of the array are with (right) and without (bottom left) streptavidin. The percentages refer to the binding assay done with different configurations (figure 36). B) Scheme of the selection process of the higher-affinity configuration with the use of streptavidin-coated magnetic particles. On the order of 109 microbeads are incubated with 20 *nM* nanostructures for 30 min at room temperature, and the supernatant is removed using a magnetic separator. Subsequently, beads are washed five times with evolution buffer (washes W1–W5) and eluted with increasing concentrations of biotin (E1–E3, with 0.8, 8, and 80 *μM* biotin, respectively). C) Agarose gel stained with SYBR Gold, showing each step of the selection protocol (L, 1 *kb* ladder supplied by New England Biolabs; I, initial sample; S, supernatant). D) AFM image of the biotin-eluted sample E3. The histogram shows the yield of the nanostructures purified from E3. Nanostructures in positions II and III of the 2 x 2 array are not selected by this method. Percentages are derived from single AFM experiments to enable direct comparison under identical conditions (no errors or error bars are given because n = 1).

First the minimal 40-nucleotide aptamer SAA1 is compared with two other 60-nucleotide-long aptamers SAA2 and SAA3 that had been selected previously by micromagnetic separation with reported dissociation constants of 25 nM and 50 nM, respectively [121] (figure 38A). As the binding position of these three aptamers on streptavidin is not known, for comparison simply a symmetric (1,4,7,10) cavity configuration is chosen for all of them. Remarkably, although SAA2 and SAA3 normally are better binders than SAA1, in the context of the aptamer configurations studied on 2 x 2 arrays, SAA1 clearly outperformed the two other aptamers.

Then an attempt is made to isolate high-affinity multivalent aptamer configurations by using a selection protocol that is based on streptavidin-coated magnetic beads. In the selection experiment, individual DNA cavities from the previous experiment are incubated with an excess of streptavidin-coated microparticles for 30 min, after which unbound nanostructures are removed by magnetic separation (figure 38B and 38C). After five washing steps, the origami cavities that still bind to the streptavidin beads are eluted by the addition of biotin, which is known to disturb aptamer binding through either direct competition for the binding site or induction of a conformational change of the protein [121].

Eluted DNA origami cavities are then imaged by AFM and identified by their barcode. Structures with barcodes that correspond to configuration numbers I and IV are observed exclusively (figure 38D), which is consistent with the results obtained from the quantitative analysis of binding to the 2 x 2 arrays with all four configurations (figure 38A).

In other words, the better binding cavities are enriched in this experiment by using a single round of affinity based origami selection. A similar selection strategy to enrich ligands that bind proteins only in a specific and controlled orientation is conceivable.

4.1.10 Discussion

It is demonstrated in this section how DNA-based scaffold structures can be used to systematically engineer multivalent binding cavities that present multiple aptamer ligands for a target protein. In contrast to previous work on DNA-based multivalent binders, not only the distance between the binding partners, but also their orientation and the flexibility of attachment is controlled. In the case of the two model aptamer systems used here (with aptamers for thrombin and streptavidin), the interplay of these parameters is shown to result in large variations in effective binding affinity. In the best configurations, a strong increase in binding yield is observed, which approached 95% in the case of streptavidin binding to extended lattices of binding cavities.

In the implementation of a multivalent binding scaffold, a flat, symmetric square-shaped DNA origami structure is used with a central square cavity in which the aptamers are presented. Together with a barcoding scheme for AFM characterization, this structure proves to be versatile and useful for the systematic investigation of geometric and mechanical parameters. In principle, other barcoding schemes with a larger address space could be implemented, and also different readout mechanisms; these features would be useful for screening larger numbers of parameters or even selecting multivalent binders de novo.

This approach does not need to be restricted to flat 2D cavities; it should also be applicable to three-dimensional (3D) origami binding pockets [5]. After an ideal combination of binders and binding parameters has been found, in principle the scaffold structure could be minimized to a much smaller structure, which preserves only the geometry of the binding pocket. This may prove useful in applications in which the relatively large size of standard origami structures is prohibitive. However, one should make sure that increasing the size of scaffold is not going to be at the cost of geometric rigidity of structures. An unstable geometry is not a reliable geometric scaffold for ligands to enable their binding to a specific target.

It is also shown that the best binding cavities can be selected from a mixture of cavities with different binding geometries of known aptamers. This presents the exciting possibility that also new binders could be selected that complement and support each other in the context of an origami cavity. Apart from providing the expected increase in binding affinity and selectivity, this would also enable the selection of binders that bind to protein targets at specific positions or in specific orientations. Proteins could be fixed in the origami frame with a known aptamer or an antibody at one position, and additional “helper aptamers” could be selected that bind to the protein at an alternative position. With this approach the specific targeting of active sites of enzymes or allosteric proteins should be possible, which would further permit modulation of their function.

4.2 Penrose tessellation

As it is discussed in section 2.6.2 and 3.3.1, one can create Penrose tessellation employing two rhombi with certain geometries. Previously during my master thesis [130], these two structures were designed and some attempts towards the Penrose tessellation were made. In this section, the current progress towards the Penrose tessellation during this thesis is going to be presented.

4.2.1 Penrose tiles

One way of creating Penrose tessellation is to use two rhombi with the same side length but different angles [110]. The rhombus with the sharp angle of 72° and the open angle 108° is called wide rhombus and the other rhombus with the sharp angle of 36° and the open one is 144° is called slim rhombus. Using DNA origami technique, two rhombi with approximate geometric properties of the Penrose tiles are designed. It should be pointed out that using DNA to create these geometries would limit the precision. Since DNA has specific polymer properties, one cannot create fully arbitrary shapes from it and the created structures are an approximation of their mathematical abstract. The AFM images of these two rhombi are shown in figure 39A and 39B.

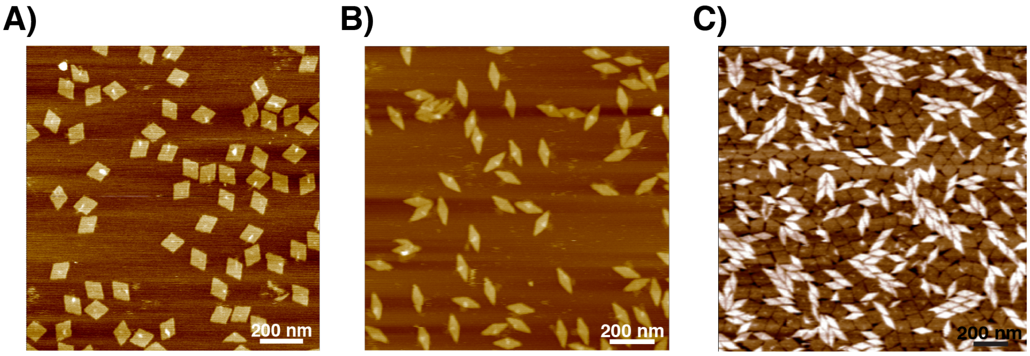


Figure 39: A) AFM image of the wide rhombus. B) AFM image of slim rhombus. C) Close packing of wide and slim rhombus on the mica surface using 80 mM NaCl. Since the slim rhombus consists of two layers and the wide rhombus one, slim rhombus is brighter in the AFM image. Bright dots on the structures are the scaffold loops.

The bright dots on structures are the unused scaffold part. In figure 39C, the two rhombi are closely packed on mica using NaCl. It is clear that the slim rhombus is brighter in the AFM image compared to the wide rhombus. As it is explained in section 3.3.1, the slim rhombus contains two identical layers which make the hight of this structure twice as high as the wide rhombus. As a result, the hight contrast will be different for the two structures and the slim rhombus looks brighter.

4.2.2 Connection rules with 4 base pair sticky ends

In section 2.6.2, it is shown that forming Penrose tessellation (figure 10A) requires special connection rules for connecting the two rhombi (see figure 10B). These connection rules can be applied using so-called “sticky ends”. Applying Penrose connection rules could potentially create locally legal configurations [112] that are shown in figure 40A. These configurations are considered as dead ends and cannot be followed to form the intended Penrose tessellation.

As a first try to implement connection rules, sticky ends with 4 base pairs are designed at the edges of the two rhombi. Figure 40B represents the AFM image of this experiment.

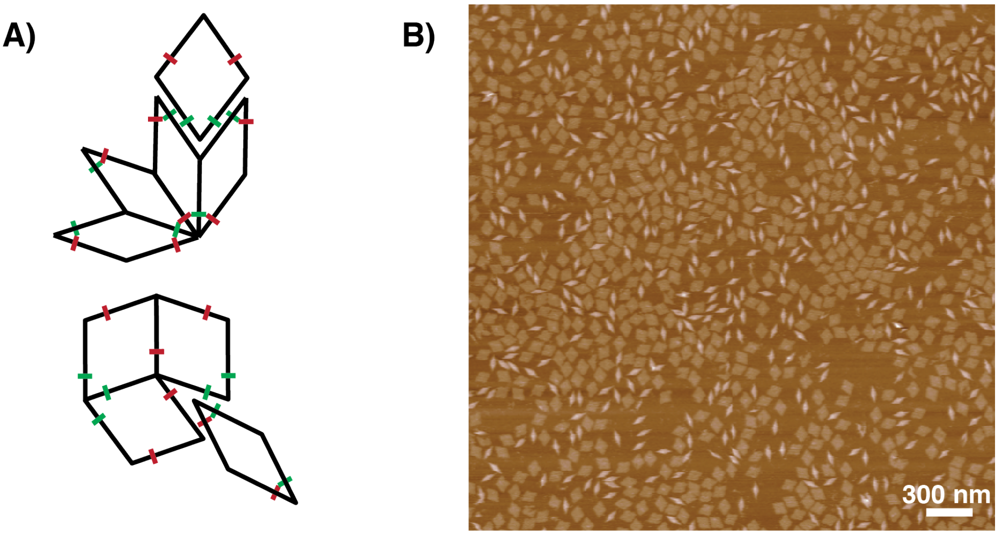


Figure 40: A) Schemes of locally legal configurations formed using Penrose connection rules. B) AFM image of Penrose tiles that have Penrose connection rules implemented in their sides using 4 base pair sticky ends. It is clear from the image that this type of connection is definitely not strong enough to multimerise DNA origami structures according to Penrose connection rules.

It is clear from the AFM image that sticky ends with 4 base pairs are not strong enough to multimerise the rhombi and form Penrose tessellation. Since for Penrose tessellation, the helices of rhombi do not align in parallel when they connect, stronger sticky ends must be used to fulfil the connections. The reason for applying weaker interactions at the first place is to avoid locally legal configurations showed in figure 40A. Using strong hybridisation interactions could simply result in formation of locally legal configurations and consequently kinetic traps that prevent Penrose tessellation.

4.2.3 Connection rules with 8 base pair sticky ends

To be able to both apply stronger interactions and avoid the locally legal configurations or decrease their probability, edge staples are modified as it is shown in figure 21. In this design, green connections between slim rhombi are eliminated to prevent the locally legal configuration formed by slim rhombi (figure 41A).

Looking at the example of Penrose tessellation and the vertex stars, lacking the green connection between slim rhombi do not prevent any of the eight vertex stars to form. Therefore, ignoring them completely should not harm the formation of vertex stars and the Penrose tessellation. In addition, the green connection between the two wide rhombi is weakened. 4 base pair sticky ends are used here to implement the green connection between the two wide rhombi ($b'$and $b'*$ in figure 41A). Considering the vertex stars for this case, one finds out that presence of green connections between wide rhombi to create Penrose tessellation is crucial. As a result, It is not possible to completely leave them out. All the other connections are applied with 8 base pair sticky ends and are shown by the schemes in figure 41A. The detailed discussion of theses connections can be found in section 3.3.5. In figure 41A, each side is represented by a letter and its complementary side has the same letter with a star on top (for instance side a connects to side $a*$). To form the Penrose-like domains, first single structures without the multimerisation staples are folded. Afterwards, the two different rhombi that are folded separately added together in a pot with multimerisation staples. Details on temperature ramps and foldings are explained in section 3.3.6. After the sample is annealed it is imaged with AFM.

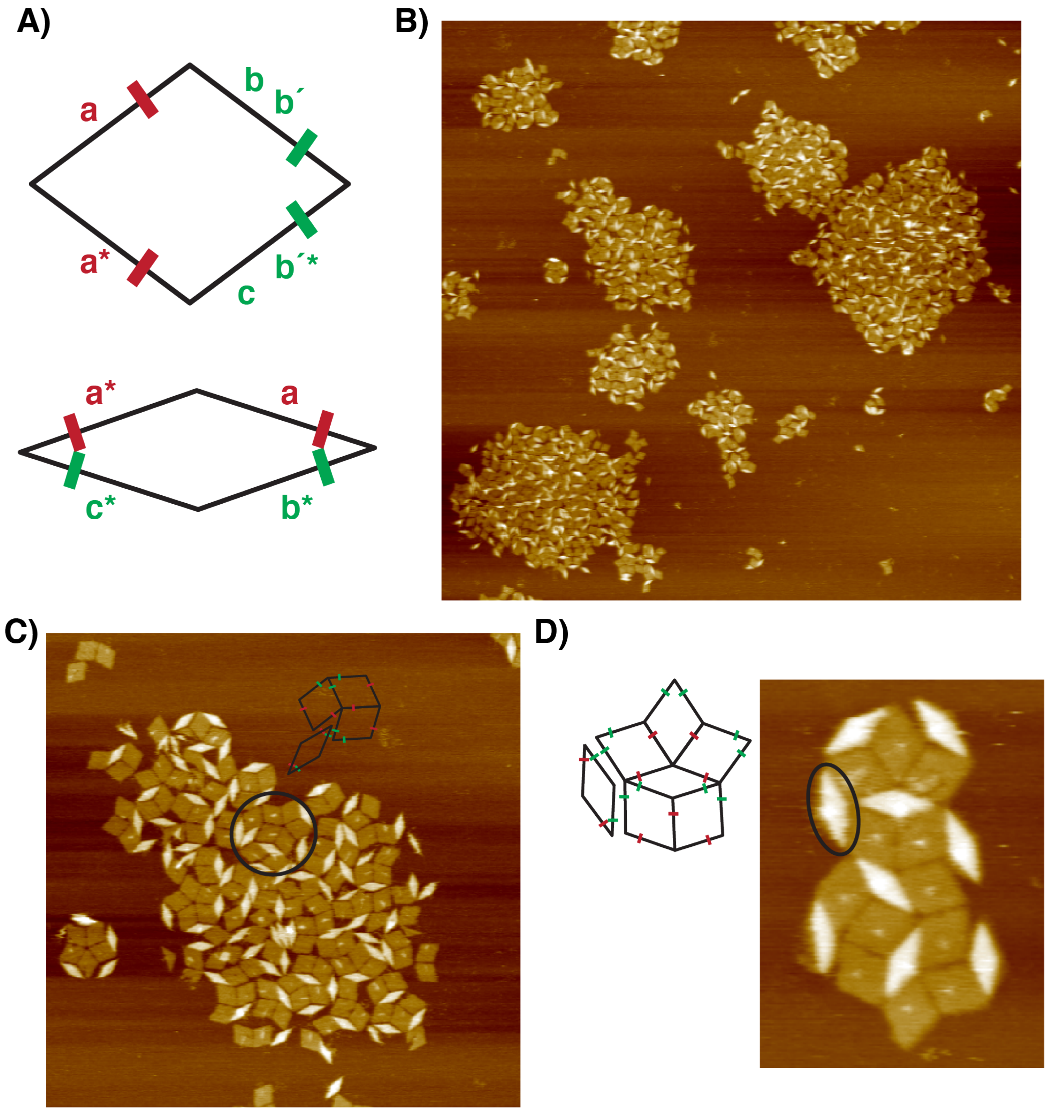


Figure 41: A) Schemes for connection rules. Detailed explanation in section 3.3.5. B) AFM image of the Penrose-like domains using 8 base pair sticky ends. The design for connection rules is according to figure 21. C) Zoomed in AFM image of one of the domains. The black circle shows the vertex stars that are connected via their wide rhombi and created the locally legal configuration made by wide rhombi. The scheme shows that the slim rhombus is only connected to one of the wide rhombi. D) A complete Penrose tessellation. Only the slim rhombus which is marked by the black ellipse is not following the connection rules. It can be seen from the scheme next to the AFM image that the slim rhombus is connected there with one of its sides. This means that the connections with 8 base pair sticky ends are very strong and can create partial mismatches in Penrose tessellation.

Figure 41B shows an AFM image of different Penrose-like domains on mica. Looking closer at these domains (figure 41C), it is clear that there are parts in which the two rhombi are connected exactly according to the Penrose connection rules. However, as it is highlighted by the black circle in figure 41C, locally legal configuration made by wide rhombi is also formed. Here, two vertex stars are connected via the green connection between the two wide rhombi and created the locally legal configuration. This also tells us that by only having vertex stars, one can create configurations that are not possible to be followed by the Penrose connection rules. In figure 41D, the AFM image of a Penrose domain is shown. The black ellipse highlights the slim rhombus which is connected there not following the connection rules. The reason for having such connections is due to the strong interactions used to connect the sides of rhombi. Although one side of the slim rhombus is complementary to the wide rhombus, the interaction is strong enough to keep the slim rhombus connected there. Leaving the marked slim rhombus out in figure 41D, all the vertex stars are members of the eight globally legal configurations showed in figure 11A. This is actually a good way to figure out if the tiling is a Penrose tiling.

4.2.4 Influence of NaCl on the Penrose lattice formation

In section 2.4, it is explained in detail that using NaCl in the AFM imagine buffer weaken the attraction of structures on mica. This results in the better mobility of structures on the mica substrate that enhances formation of arrays [99]. To this end, formed Penrose-Like domains are supplemented with NaCl and then imaged with AFM on mica. AFM images of this experiment can be seen in figure 42.

There are two motivations behind using NaCl in the imaging buffer. The first one is to help the growth of already formed Penrose seeds in solution, on the surface according to the Penrose connection rules. It is explained in section 2.1 that using **Stokes-Einstein relation**, the diffusion constant D of a particle can be related to its radius a as following:

$$D = \frac{RT}{6\pi\eta a N} \tag{41}$$

Due to the high viscous friction coefficient ζ which is written according to Stokes formula in equation 42:

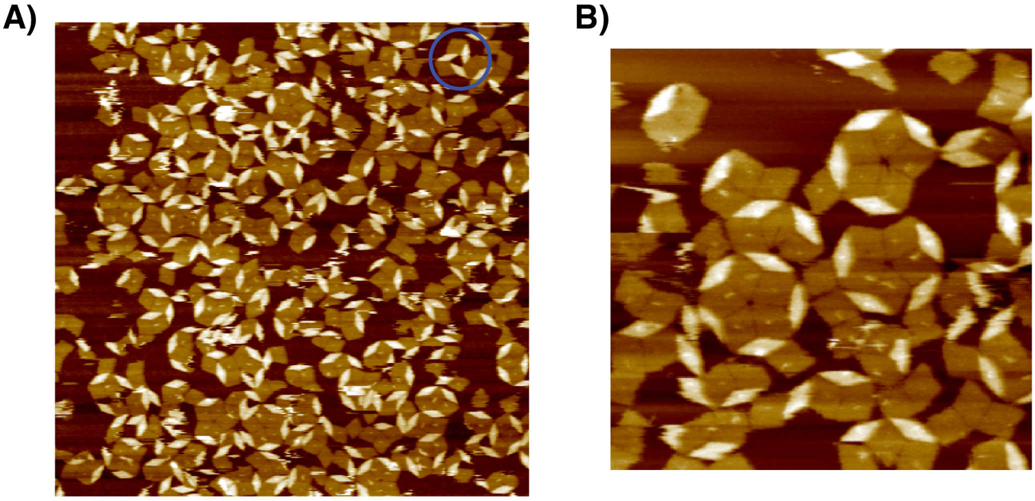


Figure 42: A) AFM image of Penrose tiles connecting via 8 base pair sticky ends supplemented with 120 mM NaCl. Blue circle shows an example of tiles connecting not via the Penrose connection rules that caused the empty gap in their centre. B) A zoomed in AFM image which represents Penrose patches including the members of globally legal vertex stars.

$$\zeta = 6\pi\eta a \tag{42}$$

larger Penrose seeds do not diffuse as fast as single tiles. Therefore, single Penrose tiles or smaller seeds can join the larger seeds on the surface and the Penrose seeds grow there. This way the mobility on the surface could be potentially used to enhance the formation of the larger Penrose lattices. The second motivation behind using the NaCl is to decrease unspecific interactions that are caused by strong adhesion of DNA origami structures on mica. As a result, interaction between structures will be mainly due to the sticky ends.

Looking at the AFM images in figure 42, it seems that the higher mobility on mica didn’t really improve the size of Penrose lattices though it helped to have less unspecific aggregation of smaller seeds. Comparing AFM images in figure 42 with the ones in figure 41 clarifies that using NaCl prevented large aggregations with empty regions on mica and smaller Penrose domains are distributed on the mica substrate. It is also clear that the formation of unwanted configurations due to the strong connections applied between the rhombi is not prevented (blue circle in figure 42A). Figure 42B represents AFM image of small Penrose lattices that contain members of globally legal vertex stars in Penrose tessellation.

4.2.5 Geometric modification of wide rhombus

To design Penrose tiles using DNA origami method, it is important to make a correct assumption regarding the inter helical distance between the double helices. Here for the first versions of the designs, the publication by Xiao-chen Bai et al. [152] is used. The DNA origami structure that they used there is a 3D structure designed in caDNAno using square lattice. The two rhombi for Penrose tessellation are designed in caDNAno software using square lattice too, however, the wide rhombus is a 2D structure that influences the assumption needed for the inter helical distance. Image analysis with AFM showed that indeed the assumption of 2.6 or 2.7 *nm* for each double helix diameter (0.6 or 0.7 *nm* for inter helical distance) in wide rhombus is not correct. Therefore, the old design of wide rhombus needed to be modified.

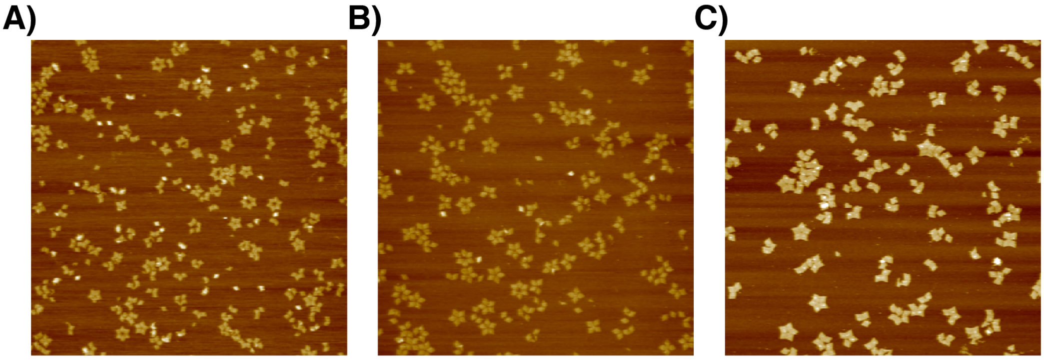


Figure 43: A) AFM image of five-fold star formed using the modified design of wide rhombus. Here, all the edge staples are used. B) AFM image of five-fold star formed using the modified design of wide rhombus. Here the two edge staples closer to the open angle of wide rhombus are left out to prevent the formation of six-fold stars. C) AFM image of five-fold stars using the old design of wide rhombus.

After analysis with AFM, it turned out that the better value to be used for the inter helical distance in the wide rhombus is 1 *nm*. This is the same value presented by Paul Rothemund in the first DNA origami paper [2]. Therefore, the old design of wide rhombus is modified to fulfil Penrose geometric properties assuming 1 *nm* as the inter helical distance. The modified design of the wide rhombus can be seen in the appendix figure 58.

Apart from the design modification, in the new version, a more precise calculation for the single-stranded domain lengths is done (see section 3.1.3 and figure 16). To test the influence of these modifications on the Penrose lattice formation, five-fold stars using different designs are created (figure 43). Figure 43A shows the AFM image of the new design in which all the staples at the edge of the structure are used to form five-fold stars. Here the calculated yield for five-fold stars after the analysis is about 15%. Closer look at the AFM image shows that six-fold stars are also formed. To decrease the probability of six-fold stars, two of the edge staples closer to the open angle of the wide rhombus are left out (figure 43B). The formation yield of five-fold stars for this case is about 21%. The same analysis is also done for the old design the calculated yield for the five-fold star is about 8% (figure 43C). Comparing the yield of five-fold stars with new modifications and the result for five-fold stars with the old design shows that the new design has indeed improved the formation of five-fold stars.

4.2.6 Discussion

Two rhombi with properties of the Penrose tiles are designed using DNA origami technique. To create Penrose tessellation, these tiles must follow the Penrose connection rules. Different types of sticky ends with 4 and 8 base pairs are used at the edges of structures to implement the connection rules. Only employing 4 base pair sticky ends results in no Penrose domains, however, applying 8 base pair sticky ends forces the structures to form Penrose-like domains. Formation of these domains happens at the cost of forming either a locally legal configuration (figure 41C) or partial mismatches (figure 41D).

Although, the design of connection rules with 8 base pair sticky ends here prevents the locally legal configuration made by slim rhombi, the locally legal configuration of wide rhombi is still formed. Considering the eight globally legal configurations that are the vertex stars in Penrose tessellation, eliminating the green connection between slim rhombi do not harm the formation of theses vertex stars, however, complete elimination of the green connection between wide rhombi prevents the formation of some of these stars. Therefore, the green connection between wide rhombi cannot be completely avoided and the locally legal configuration made by them is inevitable. One could think of creating all the eight globally legal configurations or Penrose vertex stars separately and then combining them to form Penrose lattices.

As it is shown in figure 41C, two vertex stars can also form a locally legal configuration. As a result, formation of these stars as building blocks dose not ensure creation of Penrose lattices.

Indeed there are two obstacles to overcome, one is complete prevention of locally legal configurations and the other is the partial mismatches occurring due to the strong interactions. To overcome them both one could think of better design of interaction energies in which kinetic traps would be less stable. The problem is that kinetic traps made by mismatches or locally legal configurations are not necessarily made by few connecting sides. They can involve many connecting sides with only one side of a rhombus not matching. Therefore, they are quite stable and can be regarded as favourable thermodynamic states in the assembly process. To avoid them it is needed to design extremely precise interaction energies which is going to be very challenging.

Using NaCl in the imaging buffer to make the structures more mobile on the surface did not help the formation of Penrose lattices. The size of Penrose lattices observed in this case are not larger than the case without NaCl. It only prevented the unspecific aggregation of structures due to their strong adhesion on mica.

Eventually it is very important to create DNA origami structures that have the closest geometry possible to the Penrose tiles. Moreover, the single-stranded domains used in the multimerisation staples are crucial to minimise the strain in the connections and improve the formation yield. It is shown here that the correction of these two parameters for the wide rhombus, improved the formation of five-fold stars by 13%.

Outlook

It is demonstrated here that using DNA origami technique, one can design certain geometries for different purposes. At first a square-shaped DNA origami structure with a central cavity was designed. This cavity that contains 12 uniquely addressable staples showed a great potential in molecular binding studies of multivalent targets. The scaffold routing of the square structure enables formation of large 2D lattices that is later used to arrange proteins. In addition to extended lattices, employing the fractal assembly method, finite size 2 x 2 arrays of these DNA-binding cavities with a high yield were produced. A novel barcoding scheme was established that together with the fractal assembly method, increased the throughput of the AFM meaturements and made it possible to multiplex experiments. Barcoded 2 x 2 arrays enable direct comparison of different binding configurations under identical conditions in the same experiment. As a consequence, reliable quantitative comparison of different cavities with different multiple binding sites is possible.

Then the DNA-binding cavities were used to study the binding of thrombin and streptavidin to their respective DNA aptamers. The staples in the cavity were modified with the DNA aptamers in variety of configurations to explore the influence of geometry and mechanical properties on the binding. Optimising these parameters in parallel resulted in unprecedented binding efficiencies (95% in case of streptavidin in the 2D lattices). Finally streptavidin-coated magnetic particles were used to select high-affinity multivalent aptamer configurations for streptavidin. The barcoding scheme facilitated tracing of high-affinity configurations during the selection experiment. This remarks the exciting possibility of identifying new aptamers and aptamer configurations for a certain position on the protein that has numerous diagnostic and therapeutic applications.

In the second part, sequence programmability of DNA origami structures was used to design two rhombi that almost satisfy geometric features of the Penrose tiles. Penrose connection rules were applied at the edges of these structures using sticky ends to form the Penrose tessellation. It turned out that applying the connection rules induces two types of kinetic traps that prevent formation of large Penrose lattices. The first one is due to the nature of connection rules in the Penrose tessellation and the second one due to the strength of interactions. Such kinetic traps seem to be very complicated to overcome. One possible solution would be precise design of interaction energies to make such kinetic traps less stable.

References

[1] Jones, M. R., Seeman, N. C. & Mirkin, C. A. Nanomaterials. programmable materials and the nature of the dna bond. *Science* (2015).

[2] Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* (2006).

[3] Douglas, S. M. *et al.* Self-assembly of dna into nanoscale three-dimensional shapes. *Nature* (2009).

[4] Funke, J. J. & Dietz, H. Placing molecules with bohr radius resolution using dna origami. *Nature Nanotechnology* (2015).

[5] Sprengel, A. *et al.* Tailored protein encapsulation into a dna host using geometrically organized supramolecular interactions. *Nat Commun* (2017).

[6] Shaw, A. *et al.* Binding to nanopatterned antigens is dominated by the spatial tolerance of antibodies. *Nat Nanotechnol* (2019).

[7] Kuzyk, A., Schreiber, R., Fan, Z., Pardatscher, G. & Roller…, E. DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response. *Nature* (2012).

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

[9] Shaw, A. *et al.* Spatial control of membrane receptor function using ligand nanocalipers. *Nat Methods* (2014).

[10] Li, S. *et al.* A dna nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo. *Nat Biotechnol* (2018).

[11] Voigt, N., Tørring, T., Rotaru, A. & Jacobsen…, M. Single-molecule chemical reactions on DNA origami. *Nature Nanotechnology* (2010).

[12] Nakata, E. *et al.* Zinc-finger proteins for site-specific protein positioning on dna-origami structures. *Angew Chem Int Ed Engl* (2012).

[13] Sagredo, S. *et al.* Orthogonal protein assembly on dna nanostructures using relaxases. *Angew Chem Int Ed Engl* (2016).

[14] Sacca, B. *et al.* Orthogonal protein decoration of dna origami. *Angew Chem Int Ed Engl* (2010).

[15] Grossi, G., Dalgaard Ebbesen Jepsen, M., Kjems, J. & Andersen, E. S. Control of enzyme reactions by a reconfigurable dna nanovault. *Nat Commun* (2017).

[16] Vogele, K., List, J., Simmel, F. C. & Pirzer, T. Enhanced efficiency of an enzyme cascade on dna-activated silica surfaces. *Langmuir* (2018).

[17] Linko, V., Eerikainen, M. & Kostiainen, M. A. A modular dna origami-based enzyme cascade nanoreactor. *Chem Commun (Camb)* (2015).

[18] Yamazaki, T., Heddle, J. G., Kuzuya, A. & Komiyama, M. Orthogonal enzyme arrays on a dna origami scaffold bearing size-tunable wells. *Nanoscale* (2014).

[19] Sacca, B. & Niemeyer, C. M. Functionalization of dna nanostructures with proteins. *Chem Soc Rev* (2011).

[20] Rinker, S., Ke, Y., Liu, Y., Chhabra, R. & Yan, H. Self-assembled dna nanostructures for distance-dependent multivalent ligand-protein binding. *Nat Nanotechnol* (2008).

[21] Mammen, M., Choi, S. & Whitesides, G. M. Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew Chem Int Ed Engl* (1998).

[22] Fasting, C. *et al.* Multivalency as a chemical organization and action principle. *Angew Chem Int Ed Engl* (2012).

[23] Tintore, M., Gallego, I., Manning, B., Eritja, R. & Fabrega, C. Dna origami as a dna repair nanosensor at the single-molecule level. *Angew Chem Int Ed Engl* (2013).

[24] Godonoga, M. *et al.* A dna aptamer recognising a malaria protein biomarker can function as part of a dna origami assembly. *Sci Rep* (2016).

[25] Rangnekar, A., Nash, J. A., Goodfred, B., Yingling, Y. G. & LaBean, T. H. Design of potent and controllable anticoagulants using dna aptamers and nanostructures. *Molecules* (2016).

[26] Rangnekar, A. *et al.* Increased anticoagulant activity of thrombin-binding dna aptamers by nanoscale organization on dna nanostructures. *Nanomedicine* (2012).

[27] Krissanaprasit, A. *et al.* Genetically encoded, functional single-strand rna origami: Anticoagulant. *Adv Mater* (2019).

[28] Liese, S. & Netz, R. R. Influence of length and flexibility of spacers on the binding affinity of divalent ligands. *Beilstein J Org Chem* (2015).

[29] Nelson, P. *Biological Physics: Energy, Information, Life, Student edition, Chapter 4* (2002).

[30] Watson, J. D. *et al.* *Molecular biology of the gene* (Pearson International Edition, 2008).

[31] Nelson, P. *Biological Physics: Energy, Information, Life, Student edition, Chapter 7* (2002).

[32] Murphy, M., Rasnik, I., Cheng, W., Lohman, T. & Ha, T. Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophysical Journal* (2004).

[33] Mills, J., Vacano, E. & Hagerman, P. Flexibility of single-stranded DNA: use of gapped duplex helices to determine the persistence lengths of poly (dT) and poly (dA). *Journal of Molecular Biology* (1999).

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

[35] Gedde, U. W. *Chain conformations in polymers. In Polymer Physics, Chapman and Hall* (1995).

[36] Marko, J. F. & Siggia, E. D. Stretching dna. *Macromolecules* (1995).

[37] Chandrasekaran, R. & Arnott, S. The structure of B-DNA in oriented fibers. *Journal of Biomolecular Structure and Dynamics* (1996).

[38] Martin, T. G. Functional synthetic dna nanostructures. *PhD Thesis* (2013).

[39] Watson, J. & Crick, F. Molecular structure of nucleic acids. *Nature* (1953).

[40] Franklin, R. & Gosling, R. Molecular configuration in sodium thymonucleate. *Nature* (1953).

[41] Protozanova, E., Yakovchuk, P. & D, F. M. Stacked-unstacked equilibrium at the nick site of DNA. *Journal of molecular biology* (2004).

[42] Wang, M., Yin, H., Landick, R., Gelles, J. & Block, S. Stretching DNA with optical tweezers. *Biophysical Journal* (1997).

[43] David L. Nelson, M. M. C. *Lehninger Principles of Biochemistry* (W.H. Freeman, 2008).

[44] Wang, J. Helical repeat of DNA in solution. *PNAS* (1979).

[45] Chandrasekaran, R., Wang, M. & He, R. A re-examination of the crystal structure of A-DNA using fiber diffraction data. *Journal of Biomolecular Structure and Dynamics* (1989).

[46] Wang, A., Quigley, G., Kolpak, F. & van der Marel, G. Left-handed double helical DNA: variations in the backbone conformation. *Science* (1981).

[47] Rubinstein, M. & Colby, R. *Polymer Physics.* (Oxford University Press, 2003).

[48] Baumann, C. G. *et al.* Stretching of single and collapsed dna and molecules. *Biophysical Journal Volume 78* (2000).

[49] Gacy, A. M., Goellner, G., Juranić, N., Macura, S. & McMurray, C. T. Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell* (1995).

[50] Caroline R. Astell, M. B. C. & Ward, D. C. Sequence analysis of the termini of virion and replicative forms of minute virus of mice dna suggests a modified rolling hairpin model for autonomous parvovirus dna replication. *Journal of virology* (1985).

[51] Ikoku, A. S. & Hearst, J. E. Identification of a structural hairpin in the filamentous chimeric phage m13gori1. *J Mol Biol* (1981).

[52] Mathews, D. H., Burkard, M. E., Freier, S. M., Wyatt, J. R. & Turner, D. H. Predicting oligonucleotide affinity to nucleic acid targets. *RNA* (1999).

[53] 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* (2002).

[54] SantaLucia, J., Allawi, H. T. & Seneviratne, P. A. Improved nearest-neighbor parameters for predicting dna duplex stability. *Biochemistry* (1996).

[55] SantaLucia, J. A unified view of polymer, dumbbell, and oligonucleotide dna nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences of the United States of America* (1998).

[56] SantaLucia, J. & Hicks, D. The thermodynamics of DNA structural motifs. *Annual Reviews* (2004).

[57] McKinney, S. A., Déclais, A.-C., Lilley, D. M. & Ha, T. Structural dynamics of individual holliday junctions. *Nature Structural Biology* (2002).

[58] Srinivas, N. *et al.* On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic acids research* (2013).

[59] Zhang, D. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *Journal of the American Chemical Society* (2009).

[60] Svozil, D., Kalina, J., Omelka, M. & Schneider, B. Dna conformations and their sequence preferences. *Nucleic Acids Res* (2008).

[61] J. Zhao, G. W., A. Bacolla & Vasquez, K. M. Non- b dna structure-induced genetic instability and evolution. *Cellular and Molecular Life Sciences* (2010).

[62] Belotserkovskii, B. P. *et al.* Mechanisms and implications of transcription blockage by guanine-rich dna sequences. *Proceedings of the National Academy of Sciences of the United States of America* (2010).

[63] Collier, D. A., Griffin, J. A. & Wells, R. D. Non-b right-handed dna conformations of homopurine.homopyrimidine sequences in the murine immunoglobulin c alpha switch region. *Journal of Biological Chemistry* (1988).

[64] Neidle, S. & Parkinson, G. N. The structure of telomeric dna. *Current Opinion in Structural Biology* (2003).

[65] Chin, J. Y., Schleifman, E. B. & Glazer, P. M. Repair and recombination induced by triple helix dna. *Frontiers in Bioscience* (2007).

[66] Meier, J. T., Simon, M. I. & Barbour, A. G. Antigenic variation is associated with dna rearrangements in a relapsing fever borrelia. *Cell* (1985).

[67] Smith, S. S. Evolutionary expansion of structurally complex dna sequences. *Cancer Genomics and Proteomics* (2010).

[68] List, J. J. *Reconfigurable DNA Nanostructures*. Ph.D. thesis, Technical University of Munich (2017).

[69] Mao, X., Marky, L. A. & Gmeiner, W. H. Nmr structure of the thrombin-binding dna aptamer stabilized by sr2+. *Journal of Biomolecular Structure and Dynamics* (2004).

[70] Lietard, J. *et al.* Mapping the affinity landscape of thrombin-binding aptamers on 2'f-ana/dna chimeric g-quadruplex microarrays. *Nucleic Acids Res* (2017).

[71] Baldrich, E., Restrepo, A. & O'Sullivan, C. K. Aptasensor development: Elucidation of critical parameters for optimal aptamer performance. *Anal. Chem.* (2004).

[72] Mashimo, T., Yagi, H., Sannohe, Y., Rajendran, A. & Sugiyama, H. Folding pathways of human telomeric type-1 and type-2 g-quadruplex structures. *Journal of the American Chemical Society* (2010).

[73] Koirala, D. *et al.* A single-molecule platform for investigation of interactions between g-quadruplexes and small-molecule ligands. *Nature chemistry* (2011).

[74] Phan, A. T., Modi, Y. S. & Patel, D. J. Propeller-type parallel-stranded g-quadruplexes in the human c-myc promoter. *Journal of the American Chemical Society* (2004).

[75] Rajendran, A. *et al.* Controlling the stoichiometry and strand polarity of a tetramolecular g-quadruplex structure by using a dna origami frame. *Nucleic Acids Res* (2013).

[76] Ellington, A. & Szostak, J. In vitro selection of rna molecules that bind specific ligands. *Nature* (1990).

[77] Robertson, D. & Joyce, G. Selection in vitro of an rna enzyme that specifically cleaves single-stranded dna. *Nature* (1990).

[78] Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: Rna ligands to bacteriophage t4 dna polymerase. *Science* (1990).

[79] Mercier, M.-C., Dontenwill, M. & Choulier, L. Selection of nucleic acid aptamers targeting tumor cell-surface protein biomarkers. *Cancers (Basel)* (2017).

[80] Israelachvili, J. N. Intermolecular and surface forces: revised third edition. *Academic press* (2011).

[81] Craig, V., Ninham, B. & Pashley, R. Study of the long-range hydrophobic attraction in concentrated salt solutions and its implications for electrostatic models. *Langmuir* (1998).

[82] Pastré, D., Piétrement, O., Fusil, S., Landousy, F. & Jeusset, J. Adsorption of dna to mica mediated by divalent counterions: a theoretical and experimental study. *Biophysical Journal* (2003).

[83] Lau, A. W.-C. *Fluctuation and Correlation Effects in Electrostatics of Highly-Charged Surfaces*. Ph.D. thesis, University of California Santa Barbara (2000).

[84] Arenzon, J., Stilck, J. & Levin, Y. Simple model for attraction between like-charged polyions. *The European Physical Journal B* (1999).

[85] Rouzina, I. & Bloomfield, V. Influence of ligand spatial organization on competitive electrostatic binding to DNA. *The Journal of Physical Chemistry* (1996).

[86] Pashley, R. Hydration forces between mica surfaces in electrolyte solutions. *Advances in Colloid and Interface Science* (1982).

[87] Seeman, N. Dna in a material world. *Nature* (2003).

[88] *DNA in Supramolecular Chemistry and Nanotechnology* (John Wiley and Sons, Inc, 2015).

[89] Stulz, E. Dna architectonics: towards the next generation of bio-inspired materials. *Chemistry a european journal* (2012).

[90] Kick, B., Praetorius, F., Dietz, H. & Weuster-Botz, D. Efficient production of single-stranded phage dna as scaffolds for dna origami. *Nano letters* (2015).

[91] Schüller, V., Heidegger, S. & Sandholzer..., N. Cellular immunostimulation by CpG-sequence-coated DNA origami structures. *ACS Nano* (2011).

[92] Walsh, A. S., Yin, H., Erben, C. M., Wood, M. J. & Turberfield, A. J. Dna cage delivery to mammalian cells. *ACS Nano* (2011).

[93] Li, J. *et al.* Self-assembled multivalent dna nanostructures for noninvasive intracellular delivery of immunostimulatory cpg oligonucleotides. *ACS nano* (2011).

[94] Douglas, S., Bachelet, I. & Church, G. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* (2012).

[95] Ke, Y., Lindsay, S., Chang, Y., Liu, Y. & Yan, H. Self-assembled water-soluble nucleic acid probe tiles for label-free RNA hybridization assays. *Science* (2008).

[96] Meyer, R. & Niemeyer, C. Orthogonal protein decoration of DNA nanostructures. *Small* (2011).

[97] Maune, H., Han, S., Barish, R. & Bockrath..., M. Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates. *Nature nanotechnology* (2009).

[98] Dutta, P., Varghese, R., Nangreave, J. & Lin..., S. DNA-directed artificial light-harvesting antenna. *Journal of the American Chemical Society* (2011).

[99] Aghebat Rafat, A., Pirzer, T., Scheible, M. B., Kostina, A. & Simmel, F. C. Surface-assisted large-scale ordering of dna origami tiles. *Angew Chem Int Ed Engl* (2014).

[100] Liu, W., Halverson, J., Tian, Y., Tkachenko, A. V. & Gang, O. Self-organized architectures from assorted dna-framed nanoparticles. *Nat Chem* (2016).

[101] Tian, C. *et al.* Dna nanostructures-mediated molecular imprinting lithography. *ACS Nano* (2017).

[102] Gopinath, A. & Rothemund, P. W. K. Optimized assembly and covalent coupling of single-molecule dna origami nanoarrays. *ACS Nano* (2014).

[103] Gopinath, A., Miyazono, E., Faraon, A. & Rothemund, P. W. K. Engineering and mapping nanocavity emission via precision placement of dna origami. *Nature* (2016).

[104] Liu, W., Zhong, H., Wang, R. & Seeman, N. C. Crystalline two-dimensional DNA-origami arrays. *Angewandte Chemie International Edition* (2011).

[105] 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* (2015).

[106] Tikhomirov, G., Petersen, P. & Qian, L. Programmable disorder in random dna tilings. *Nat Nanotechnol* (2017).

[107] Suzuki, Y., Endo, M. & Sugiyama, H. Lipid-bilayer-assisted two-dimensional self-assembly of dna origami nanostructures. *Nat Commun* (2015).

[108] Tikhomirov, G., Petersen, P. & Qian, L. Fractal assembly of micrometre-scale dna origami arrays with arbitrary patterns. *Nature* (2017).

[109] Conway, J., Burgiel, H. & Goodman-Strauss, C. The symmetries of things. *A K Peters, Ltd* (2008).

[110] *Algebraic theory of Penrose's non-periodic tilings of the plane.* (Department of Mathematics, Eindhoven University of Technology, 1981).

[111] Mackay, A. L. Crystallography and the penrose pattern. *Physica A: Statistical Mechanics and its Applications* (1982).

[112] Senechal, M. *Quasicrystals and Geometry* (Cambridge University Press, 1996).

[113] Kaliteevski, M. A. *et al.* Two-dimensional penrose-tiled photonic quasicrystals: from diffraction pattern to band structure. *Nanotechnology* (2000).

[114] URL `https://www.youtube.com/watch?v=Pyg0f27kKXw`.

[115] Jue, T. *Fundamental concepts in Biophysics* (Humana Press, a part of Springer Science+Business Media, LLC, 2009).

[116] Kiessling, L. L., Gestwicki, J. E. & Strong, L. E. Synthetic multivalent ligands as probes of signal transduction. *Angew Chem Int Ed Engl* (2006).

[117] Hunter, C. A. & Anderson, H. L. What is cooperativity? *Angew Chem Int Ed Engl* (2009).

[118] Krishnamurthy, V. M., Estrofi, L. A. & Whitesides, G. M. Multivalency in ligand design. *fragment-based Approaches in Drug Discovery, Edited by W. Jahnke and D. A. Erlanson, WILEY-VCH Verlag GmbH and Co. KGaA, Weinheim* (2006).

[119] Le Trong, I. *et al.* Streptavidin mutant y43f with biotin at 1.81a resolution. *Protein Data Bank* (2003).

[120] Pica, A., Russo Krauss, I., Parente, V. & Sica, F. Human thrombin sandwiched between two dna aptamers: Hd22 and hd1-deltat12. *Protein Data Bank* (2016).

[121] Oh, S. S. *et al.* Generation of highly specific aptamers via micromagnetic selection. *Anal Chem* (2009).

[122] Bittker, J. A., Le, B. V. & Liu, D. R. Nucleic acid evolution and minimization by nonhomologous random recombination. *Nat Biotechnol* (2002).

[123] Aghebat Rafat, A., Sagredo, S., Thalhamer, M. & Simmel, F. C. Barcoded dna origami structures for multiplexed optimization and enrichment of dna-based protein-binding cavities. *Nature chemistry* (2020).

[124] Pica, A. *et al.* Through-bond effects in the ternary complexes of thrombin sandwiched by two dna aptamers. *Nucleic Acids Res* (2017).

[125] URL `https://cadnano.org/`.

[126] Douglas, S. *et al.* Rapid prototyping of 3d dna-origami shapes with cadnano. *Nucleic Acids Research* (2009).

[127] Dietz, H., Douglas, S. M. & Shih, W. M. Folding dna into twisted and curved nanoscale shapes. *Science* (2009).

[128] Ke, Y. *et al.* Multilayer dna origami packed on a square lattice. *Jornal of the American Chemical Society* (2009).

[129] Woo, S. & Rothemund, P. W. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nature chemistry* (2011).

[130] Aghebat Rafat, A. *Surface-Assisted Large-Scale Ordering of DNA Origami Structures*. Master's thesis, Technical University of Munich (2014).

[131] Kopperger, E. *et al.* A self-assembled nanoscale robotic arm controlled by electric fields. *Science* (2018).

[132] Stahl, E., Martin, T. G., Praetorius, F. & Dietz, H. Facile and scalable preparation of pure and dense dna origami solutions. *Angew Chem Int Ed Engl* (2014).

[133] null & null. Atomic force microscope. *Physical Review Letters* **56** (1986).

[134] Jones, J. On the determination of molecular fields. i. from the variation of the viscosity of a gas with temperature. *Proceedings of the Royal Society of London* (1924).

[135] Pardatscher, G. *Single molecule studies based on lithographically arranged DNA origami structures*. Master's thesis, Technical University of Munich (2013).

[136] URL `https://afm.oxinst.com/assets/uploads/products/asylum/documents/blueDrive-Photothermal-Excitation-Data-Sheet.pdf`.

[137] Ratcliff, G., Erie, D. & Superfine, R. Photothermal modulation for oscillating mode atomic force microscopy in solution. *Applied Physics Letters* (1998).

[138] Labuda, A. *et al.* Comparison of photothermal and piezoacoustic excitation methods for frequency and phase modulation atomic force microscopy in liquid environments. *AIP Advances* (2011).

[139] Sharonov, A. & Hochstrasser, R. M. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *PNAS* (2006).

[140] Jungmann, R. *et al.* Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano letters* (2010).

[141] URL `https://www.microscopyu.com/techniques/fluorescence/total-internal-reflection-fluorescence-tirf-microscopy`.

[142] Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with dna-paint. *Nat Protoc* (2017).

[143] Dai, M., Jungmann, R. & Yin, P. Optical imaging of individual biomolecules in densely packed clusters. *Nat Nanotechnol* (2016).

[144] Strauss, M. T., Schueder, F., Haas, D., Nickels, P. C. & Jungmann, R. Quantifying absolute addressability in dna origami with molecular resolution. *Nat Commun* (2018).

[145] Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal* (2002).

[146] Kopperger, E. *DNA-based Nanoelectromechanical Systems, Technical University of Munich, 2018.* Ph.D. thesis, Technical University of Munich (2018).

[147] (2017). URL `http://qianlab.caltech.edu/FracTileCompiler/`.

[148] Murray, J. & Vanryper, W. *Encyclopedia of Graphics File Formats* (O'Reilly and Associates, Sebastopol, CA, 1994).

[149] Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H. & Toole, J. J. Selection of single-stranded-dna molecules that bind and inhibit human thrombin. *Nature* (1992).

[150] Tasset, D. M., Kubik, M. F. & Steiner, W. Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J Mol Biol* (1997).

[151] Padmanabhan, K., Padmanabhan, K. P., Ferrara, J. D., Sadler, J. E. & Tulinsky, A. The structure of alpha-thrombin inhibited by a 15-mer single-stranded dna aptamer. *The journal of biological chemistry* (1993).

[152] null, Scheres, S. & Dietz, H. Cryo-EM structure of a 3D DNA-origami object. *PNAS* (2012).

Publication Record

1) Ali Aghebat Rafat, Sandra Sagredo, Melissa Thalhammer, Friedrich C. Simmel. Barcoded DNA origami structures for multiplexed optimization and enrichment of DNA-based protein-binding cavities. Nature Chemistry (2020). `10.1038/s41557-020-0504-6`.

2) S. Sagredo, T. Pirzer, A. Aghebat Rafat, M. A. Götzfried, G. Moncalian, F. C. Simmel, F. de la cruz, Orthogonal protein assembly on DNA nanostructures Using relaxases, Angewandte Chemie int (2016). `10.1002/anie.201510313`

3) Aghebat Rafat, A., Pirzer, T., Scheible, M. B., Kostina, A. & Simmel, F. C. Surface-Assisted Large-Scale ordering of DNA origami tiles. Angewandte Chemie International Edition (2014). `http://dx.doi.org/10.1002/anie.201403965`.

Appendix

Figures

Binding yields for different configurations and different experiments

DNA origami designs

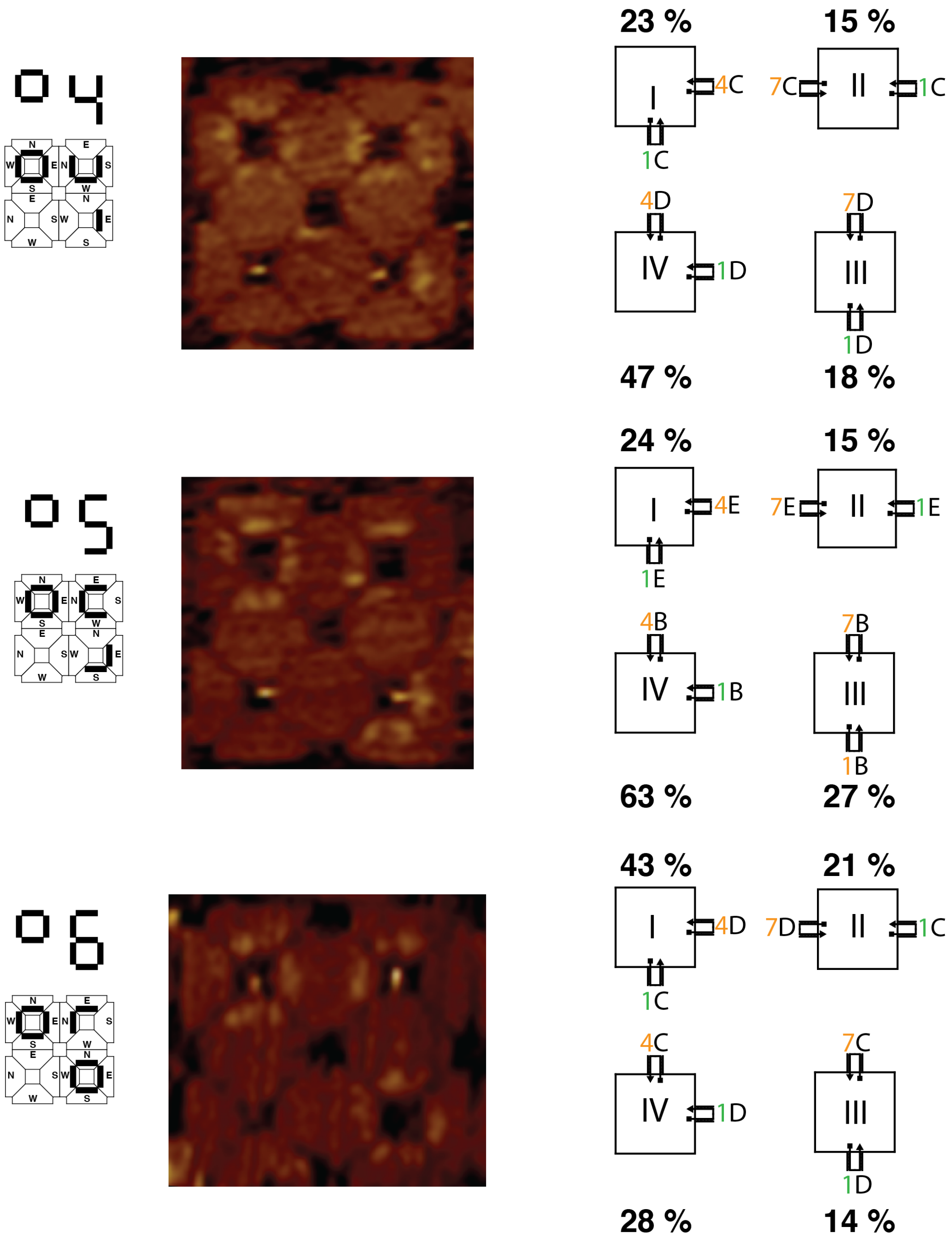


Figure 44: DNA arrays used to study thrombin binding while modifying spacer length and flexibility. Left, schematic illustration of the configurations used in 2 x 2 arrays barcoded with number 4, 5 and 6. Center, represents the AFM image of 2 x 2 arrays with thrombin bound in some of the cavities. Right, binding yield in % obtained from over ≈200 tiles for each configuration measured in a single AFM experiment (one sample of a set of barcoded cavities is imaged at multiple locations). Green denotes HD22, always in position 1, and orange denotes TBA1 aptamer in position 4 (90°) or 7 (180°). The letter indicates the spacer length and flexibility.

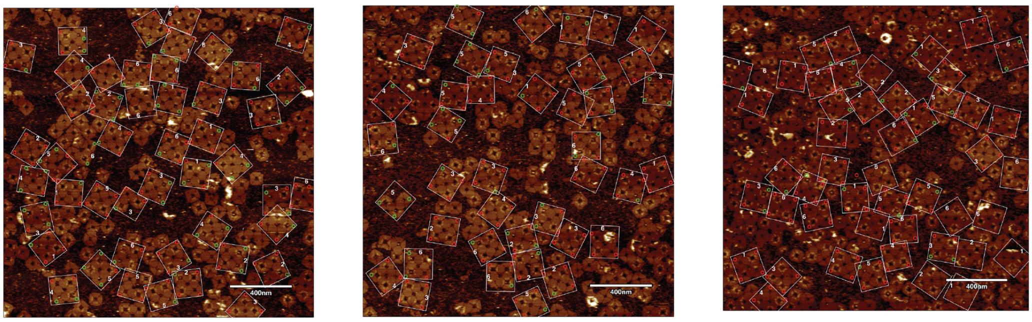


Figure 45: AFM quantification of thrombin binding to 2x2 arrays. AFM images of 2 x 2 arrays with barcodes from 1 to 6 incubated with alpha-thrombin. White squares indicate the 2 x 2 array, and the numbers indicate the barcode number. Green circles indicate origami with protein bound in the central cavity; red circles indicate empty nanostructures. Single origami structures or ambiguous cases are not considered in the statistics.

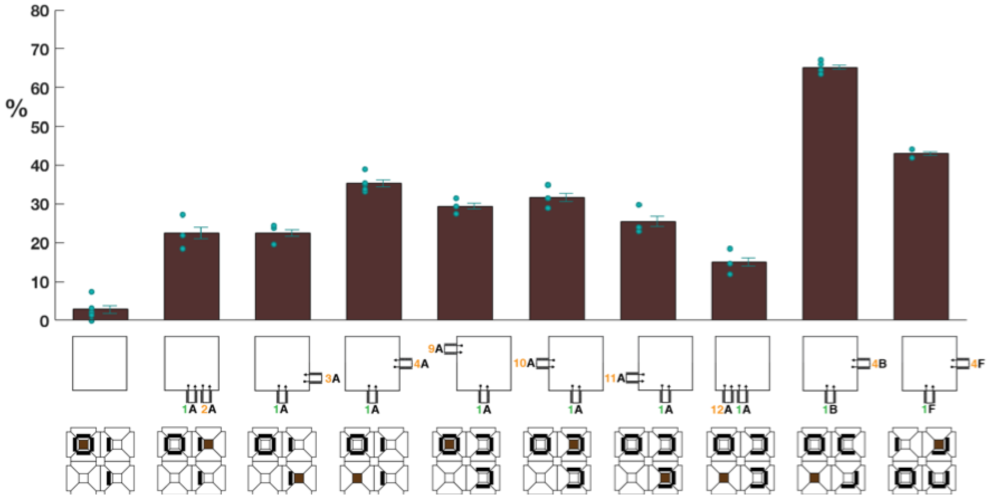


Figure 46: Comparison of binding yields for ten different thrombin aptamer configurations and spacer designs measured in independent AFM sessions (different sessions performed on different days). Bar graphs represent the mean ±SEM for each case obtained from n = 2-6 experiments, which are $\leq 3\%$. Green dots represent individual points (yield percentage obtained from independent AFM sessions). The schematic configuration of aptamers and the illustration of the position of the different configurations in 2 x 2 arrays (brown) plus their respective barcoded 2 x 2 are shown. SEM = standard error of the mean.

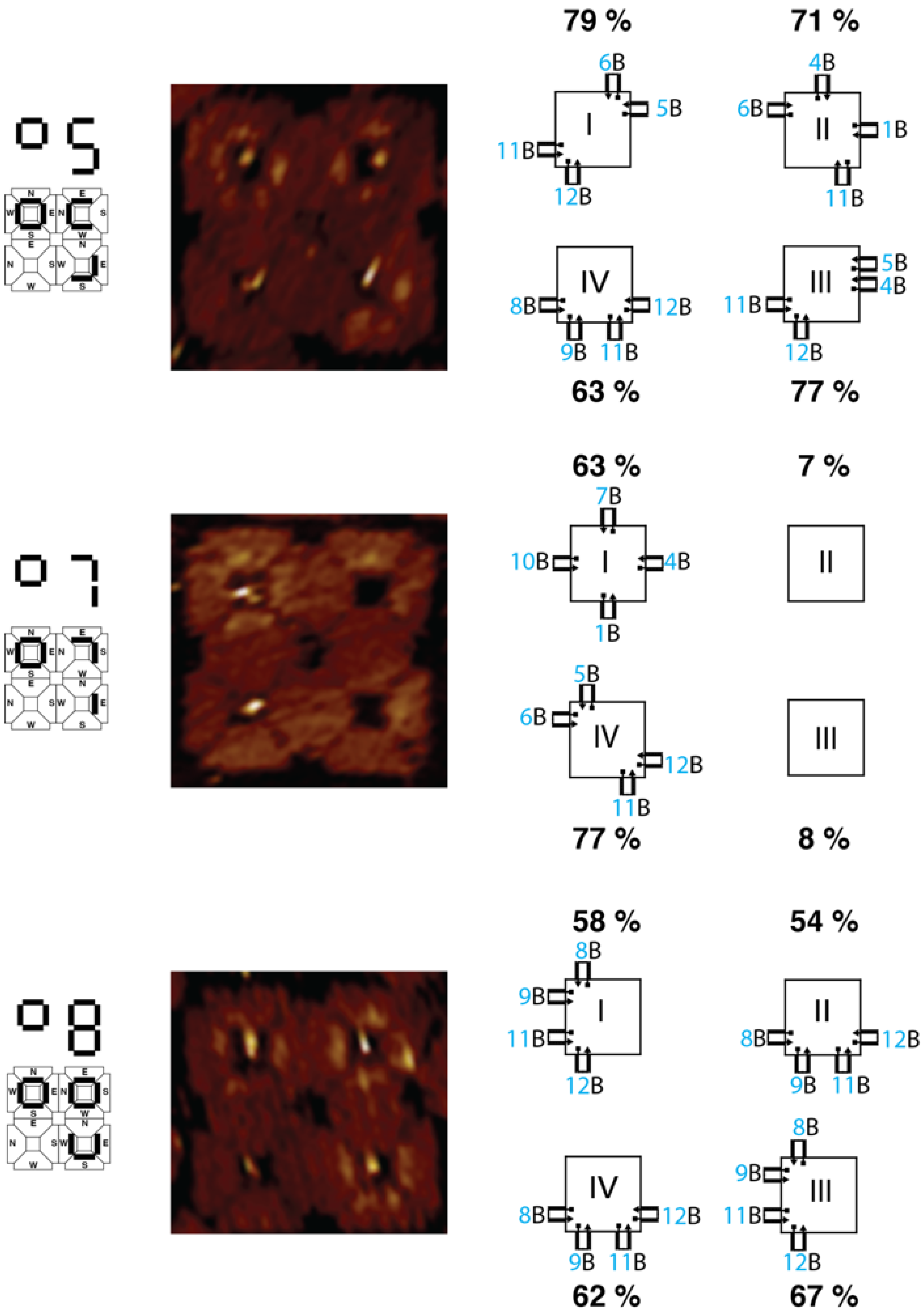


Figure 47: DNA arrays 5, 7 and 8, which are used to study streptavidin binding to SSA1 aptamers with different aptamer configurations. Left, schematic illustration of the configurations used in 2 x 2 arrays barcoded with number 5, 7 and 8. Center, representative AFM images of 2 x 2 arrays with streptavidin bound in some of the cavities. Right, Percentages denote the binding yield determined for each configuration. The configuration of the aptamers inside each cavity is represented as a number, that indicates the position of the aptamer, and a letter, that indicates the spacer length and flexibility (case B, 20 bp spacer and 4 thymidines as an aptamer linker). The number in blue indicates that all aptamers are the same SAA1 structure. Array 5 has the same configuration as 2 and 4 but with different staples.

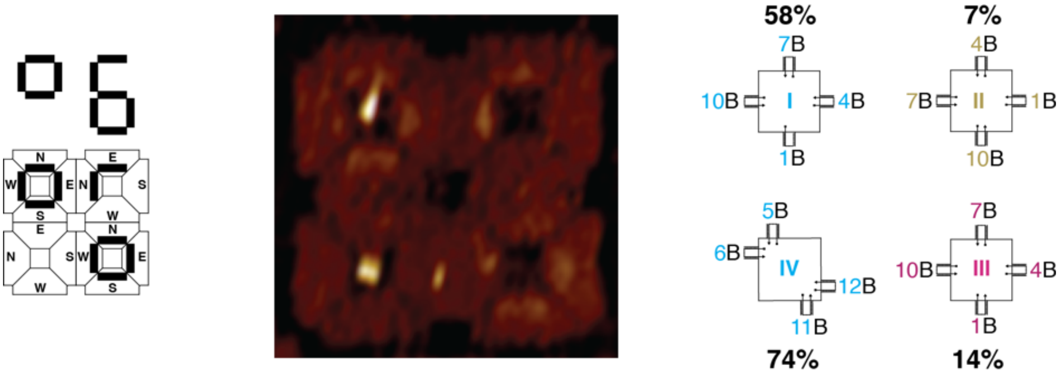


Figure 48: DNA array used to compare streptavidin binding to SAA1, SAA2 and SAA3 aptamers. Left, schematic illustration of the configurations used in 2 x 2s barcoded with number 6. Center, representative AFM image of a 2 x 2 array with streptavidin bound in two of the cavities. Right, Percentages denote the binding yield in each configuration. The configuration of aptamers inside each cavity is represented as a number, which indicates the position of the aptamer, and a letter, which indicates the spacer length and flexibility. Blue codes for SAA1 aptamer, khaki for SAA2 and pink for SAA3.

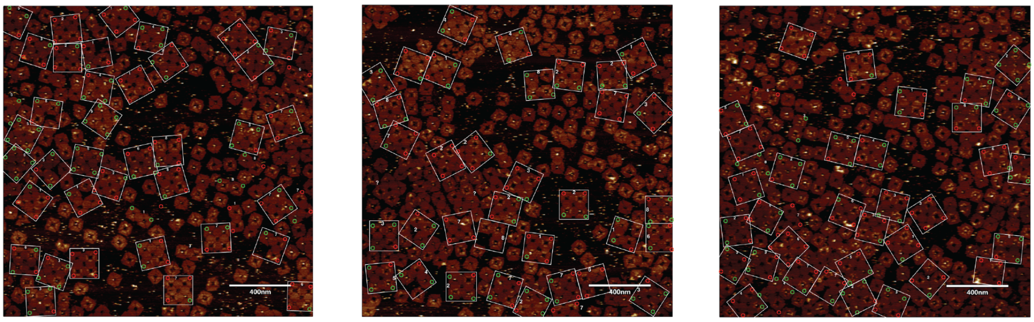


Figure 49: AFM quantification of streptavidin binding to 2x2 arrays with barcodes from 1 to 8. White squares indicate the 2x2 arrays, and the numbers indicate the barcode number. Green circles indicate origami structures with protein bound in the central cavity; red circles indicate empty nanostructures. Question marks indicate structures, which could not be classified. Single origami structures or ambiguous cases are not considered in the statistics.

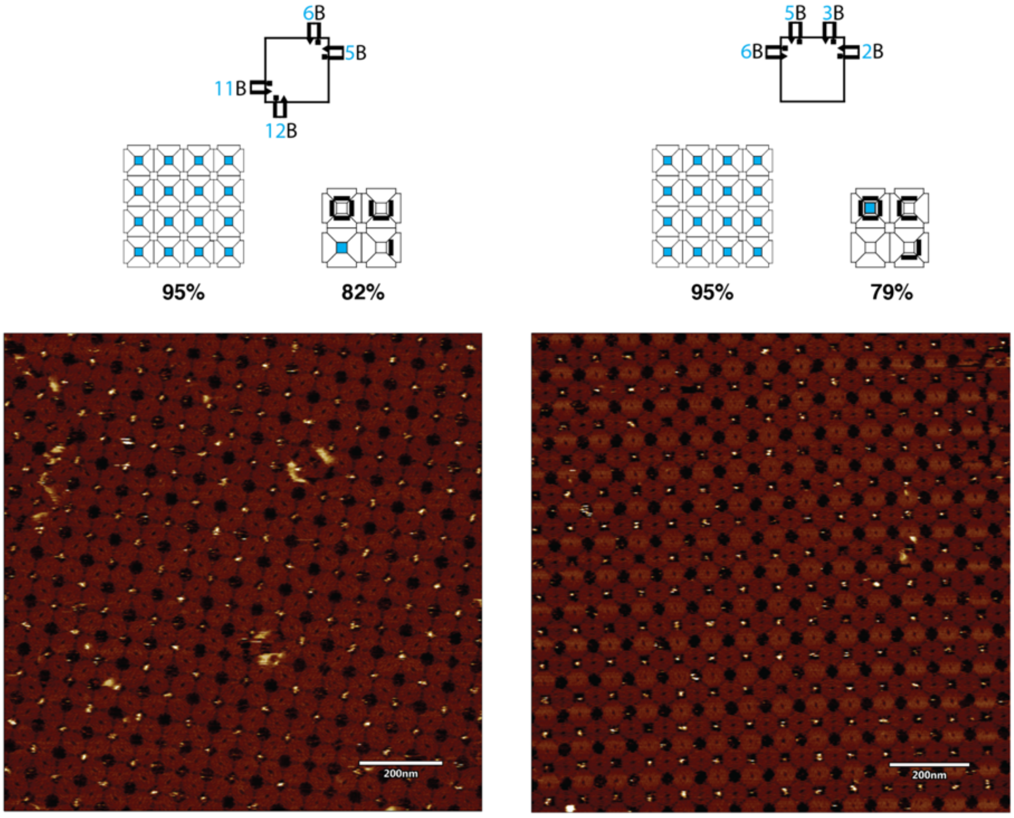
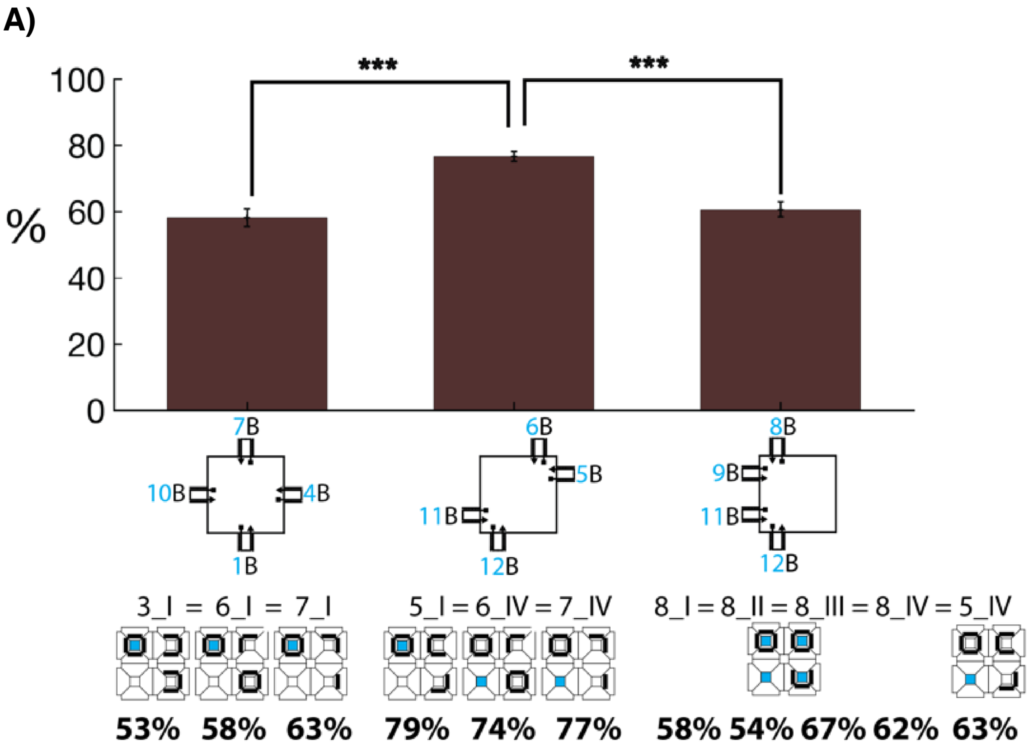
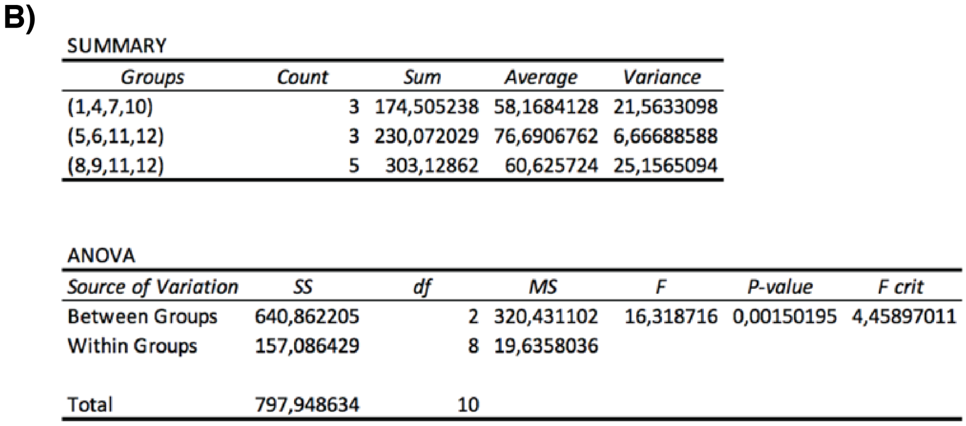


Figure 50: Comparison between the binding yields of streptavidin to cavities with two different aptamer configurations both in crystals and 2 x 2 arrays. Top: Schematic representation of the SAA1 aptamer in the cavity of a 2D crystal and 2 x 2 array (highlighted in blue) and the corresponding binding yields. Bottom: AFM images of 2D crystalline arrays formed from DNA origami structures with the two different aptamer configurations.





SUMMARY

Groups	Count	Sum	Average	Variance
(1,4,7,10)	3	174,505238	58,1684128	21,5633098
(5,6,11,12)	3	230,072029	76,6906762	6,66688588
(8,9,11,12)	5	303,12862	60,625724	25,1565094

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	640,862205	2	320,431102	16,318716	0,00150195	4,45897011
Within Groups	157,086429	8	19,6358036			
Total	797,948634	10				

Figure 51: Comparison between three streptavidin aptamer binding configurations to estimate binding yield uncertainty within the same experiment. Analysis of variance (ANOVA) and a t-student test reveal statistically significant differences in binding yield. A) Schematic representations of three SAA1 aptamer configurations with corresponding binding yields. As indicated, the same configuration is present in different cavities of barcoded 2x2 arrays. The percentages given below the structures show the binding yield of the cavities highlighted in blue. The bar graphs represent the mean ±SEM of the three configurations (n = 3, 3 and 5). P-values are calculated with student t-test (***$P < 0.001$). B) Summary of the data analysed using a one-way ANOVA test to assess statistical significance.

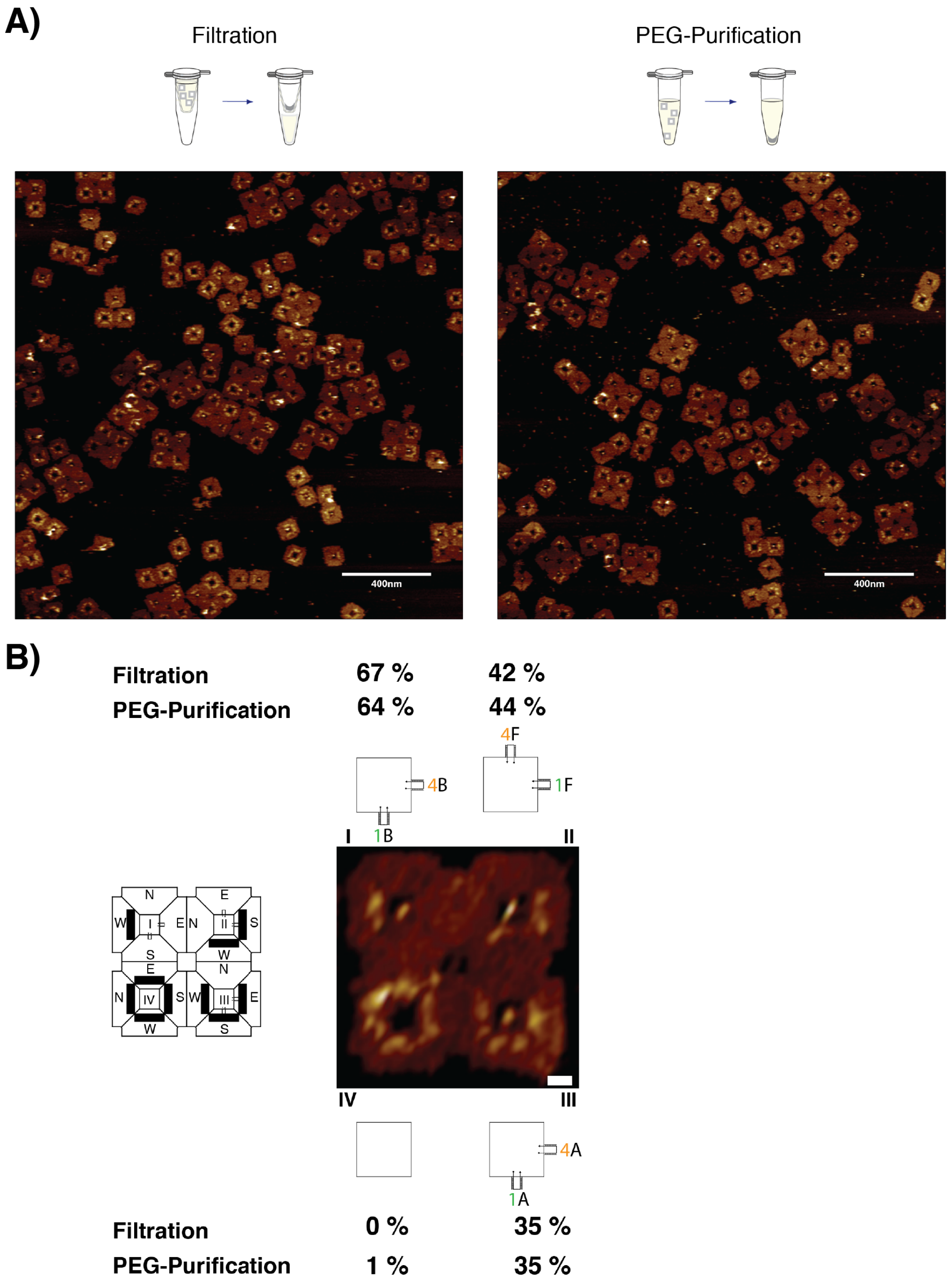
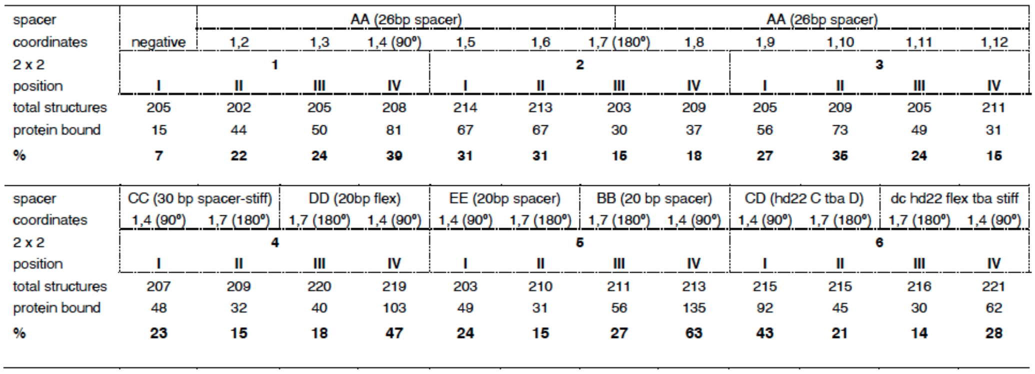


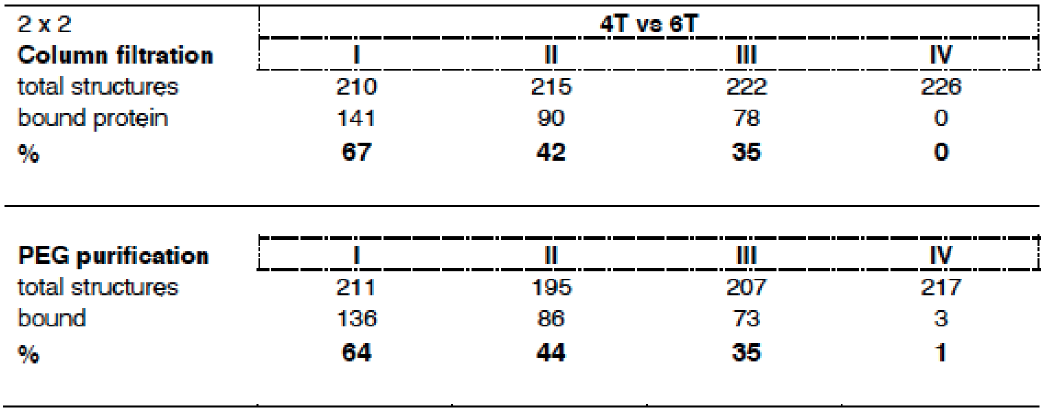
Figure 52: A) Binding yield comparison for nanostructures purified with either column-filtration (left) or PEG-purification (right). A) Scheme and AFM image of the 2 x 2 arrays incubated with protein for 30 min. In the case of PEG-purification, we observed slightly more unbound protein on the mica surface. B) Scheme and binding yield of the proteins in the four configurations.



spacer		AA (26bp spacer)					AA (26bp spacer)					
coordinates	negative	1,2	1,3	1,4 (90°)	1,5	1,6	1,7 (180°)	1,8	1,9	1,10	1,11	1,12
2 x 2	**1**				**2**				**3**			
position	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**
total structures	205	202	205	208	214	213	203	209	205	209	205	211
protein bound	15	44	50	81	67	67	30	37	56	73	49	31
%	**7**	**22**	**24**	**39**	**31**	**31**	**15**	**18**	**27**	**35**	**24**	**15**

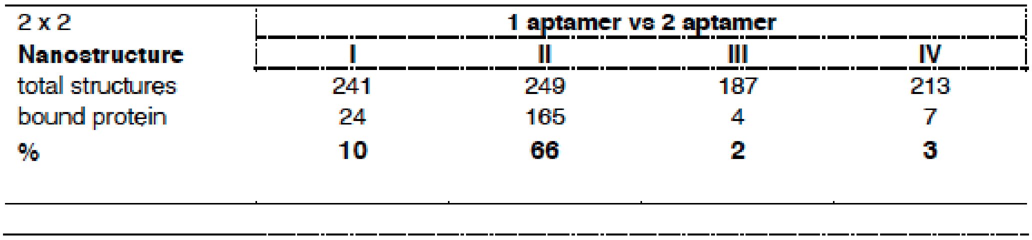
spacer	CC (30 bp spacer-stiff)		DD (20bp flex)		EE (20bp spacer)		BB (20 bp spacer)		CD (hd22 C tba D)		dc hd22 flex tba stiff	
coordinates	1,4 (90°)	1,7 (180°)	1,7 (180°)	1,4 (90°)	1,4 (90°)	1,7 (180°)	1,7 (180°)	1,4 (90°)	1,4 (90°)	1,7 (180°)	1,7 (180°)	1,4 (90°)
2 x 2	**4**				**5**				**6**			
position	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**
total structures	207	209	220	219	203	210	211	213	215	215	216	221
protein bound	48	32	40	103	49	31	56	135	92	45	30	62
%	**23**	**15**	**18**	**47**	**24**	**15**	**27**	**63**	**43**	**21**	**14**	**28**

Figure 53: Quantification of binding yields of α-Thrombin to 2 x 2 arrays 1-6. The length and type of spacer is indicated with characters A-E. HD22 is located in coordinate 1 and the position of the TBA1 aptamer changed in the 24 configurations.



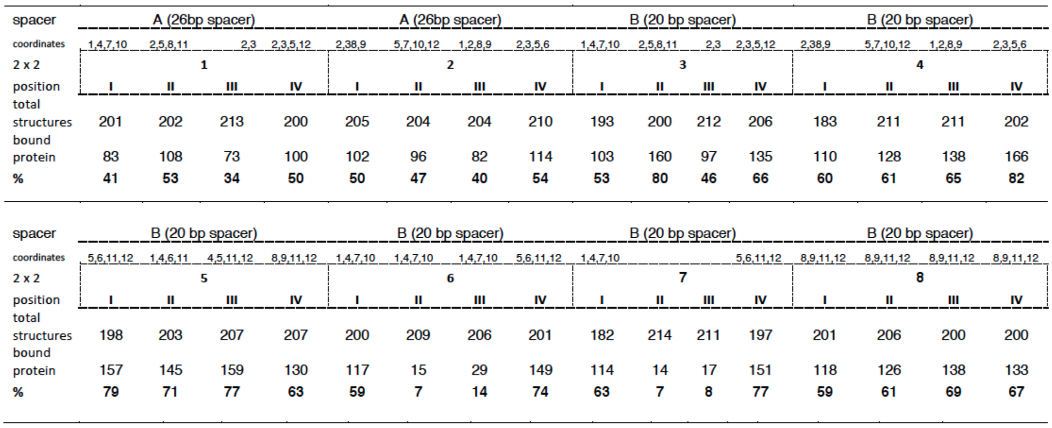
2 x 2	**4T vs 6T**			
Column filtration	**I**	**II**	**III**	**IV**
total structures	210	215	222	226
bound protein	141	90	78	0
%	**67**	**42**	**35**	**0**
PEG purification	**I**	**II**	**III**	**IV**
total structures	211	195	207	217
bound	136	86	73	3
%	**64**	**44**	**35**	**1**

Figure 54: Quantification of binding yields of α-Thrombin to 2 x 2s with 4T or 6T linker between the stem and the aptamer. The length of the spacer is denoted with the A, B and F cases. The configuration tested was 90°, in which HD22 is located in coordinate 1 and TBA1 aptamer in 4.



2 x 2	1 aptamer vs 2 aptamer			
Nanostructure	**I**	**II**	**III**	**IV**
total structures	241	249	187	213
bound protein	24	165	4	7
%	**10**	**66**	**2**	**3**

Figure 55: Quantification of binding yields of α-Thrombin to 2 x 2s with only HD22 or HD22 and TBA1 aptamer. The length of the spacer stem is 20 bp. The single-stranded linker between the stem and the aptamer contains four thymidines.



spacer	A (26bp spacer)				A (26bp spacer)				B (20 bp spacer)				B (20 bp spacer)			
coordinates	1,4,7,10	2,5,8,11	2,3	2,3,5,12	2,38,9	5,7,10,12	1,2,8,9	2,3,5,6	1,4,7,10	2,5,8,11	2,3	2,3,5,12	2,38,9	5,7,10,12	1,2,8,9	2,3,5,6
2 x 2	**1**				**2**				**3**				**4**			
position	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**
total structures	201	202	213	200	205	204	204	210	193	200	212	206	183	211	211	202
bound protein	83	108	73	100	102	96	82	114	103	160	97	135	110	128	138	166
%	**41**	**53**	**34**	**50**	**50**	**47**	**40**	**54**	**53**	**80**	**46**	**66**	**60**	**61**	**65**	**82**

spacer	B (20 bp spacer)				B (20 bp spacer)				B (20 bp spacer)				B (20 bp spacer)			
coordinates	5,6,11,12	1,4,6,11	4,5,11,12	8,9,11,12	1,4,7,10	1,4,7,10	1,4,7,10	5,6,11,12	1,4,7,10			5,6,11,12	8,9,11,12	8,9,11,12	8,9,11,12	8,9,11,12
2 x 2	**5**				**6**				**7**				**8**			
position	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**
total structures	198	203	207	207	200	209	206	201	182	214	211	197	201	206	200	200
bound protein	157	145	159	130	117	15	29	149	114	14	17	151	118	126	138	133
%	**79**	**71**	**77**	**63**	**59**	**7**	**14**	**74**	**63**	**7**	**8**	**77**	**59**	**61**	**69**	**67**

Figure 56: Quantification of binding yields of streptavidin to 2 x 2s 1-8 in their certain configurations. The length of the spacer stem is denoted with A (26 bp) or B (20 bp). The coordinates show the 4 positions where the SAA1-3 protrude.

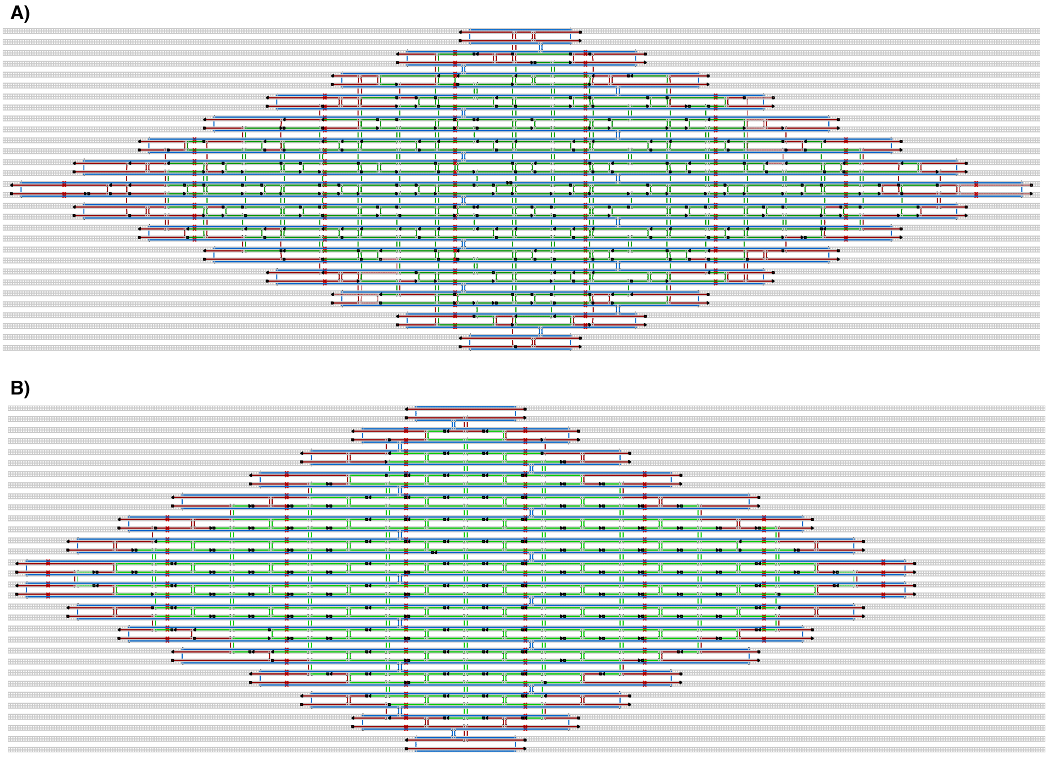


Figure 57: A) caDNAno design of slim rhombus. Red staples specify edge staples and the extensions showing 4 thymidine bases used to prevent base-stacking. This structure has two identical layers. In reality, every two scaffold rows which are next to each other are in a plane perpendicular to the design plane. B) caDNAno design of wide rhombus. Red staples represent edge staples and the extensions showing 4 thymidine bases used to prevent base-stacking.

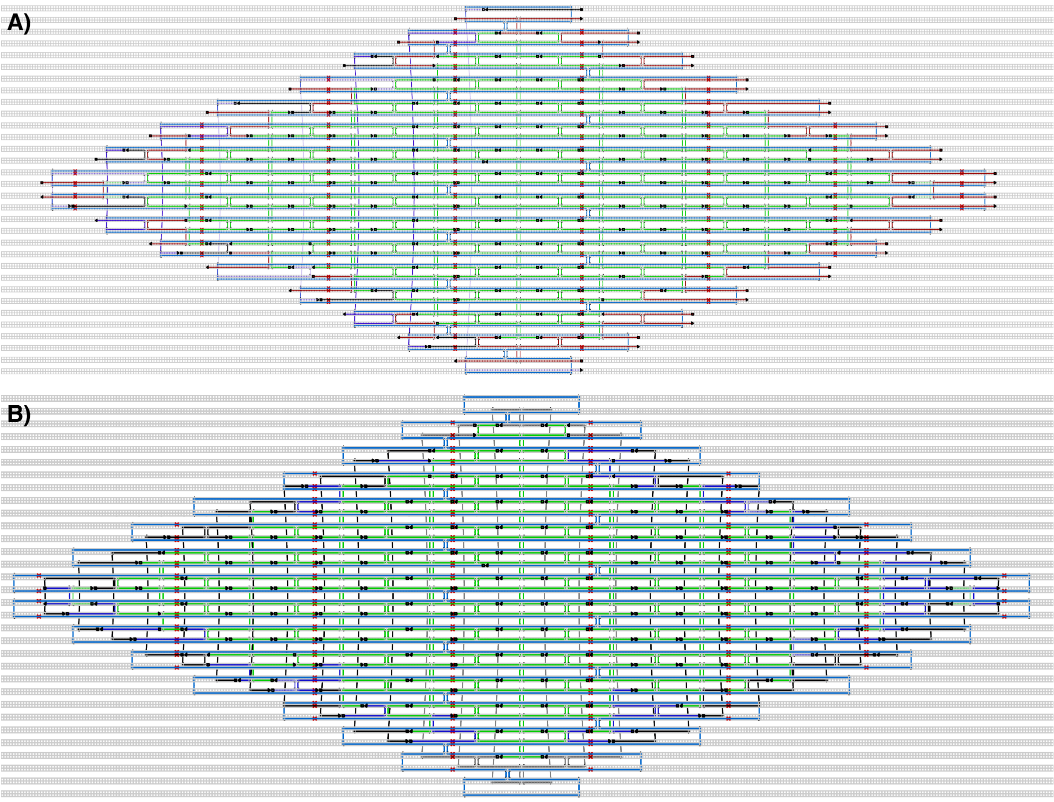


Figure 58: A) caDNAno design of star-shaped assembly of wide rhombus. Red staples specify edge staples and the extensions showing 4 thymidine bases used to prevent base-stacking. Purple staples showing polymerization staples for connecting wide rhombus origami tiles to form five-fold star. B) Modified caDNAno design of the wide rhombus. This design showed a better yield of five-fold starts. Black staples represent the connecting staples. Here the connections of both type of five-fold stars are made in the design.

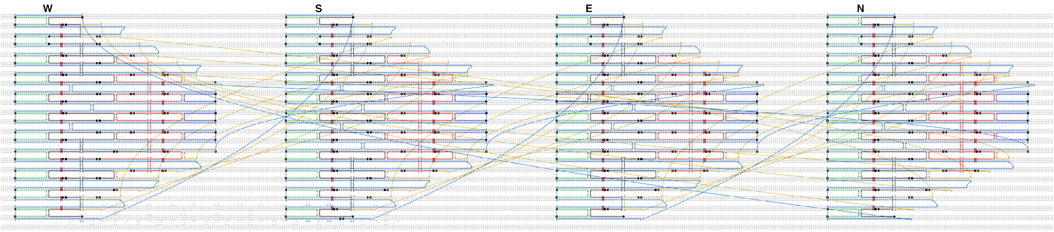


Figure 59: Cadnano design for the square-shaped nanostructure with the scaffold loop in the corner. This version would form a cavity in the centre. Staples used for barcoding are colored in red. From left to the right: trapezoids West, South, East and North. Blue staples are the cavity staples, yellow staples are the connecting staples between trapezoids.

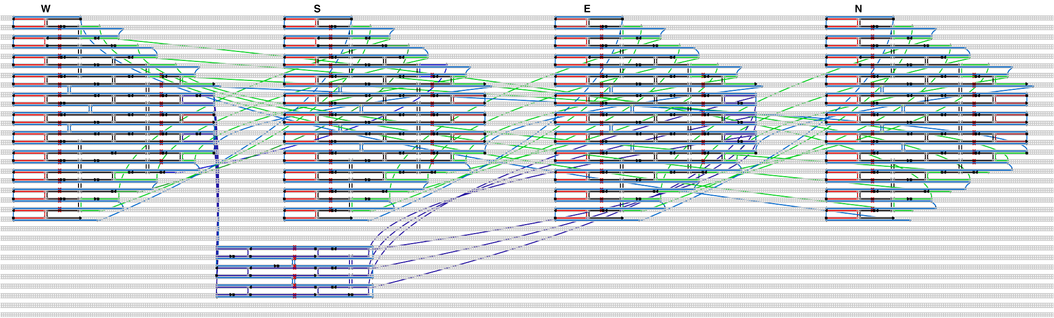


Figure 60: Cadnano design for the square-shaped nanostructure with the scaffold loop in the middle. This version can be used for applications in which having a substructure in the middle is desired. From left to the right: trapezoids West, South, East and North. Blue staples are the staples related to the substructure, green staples are the connecting staples between trapezoids.

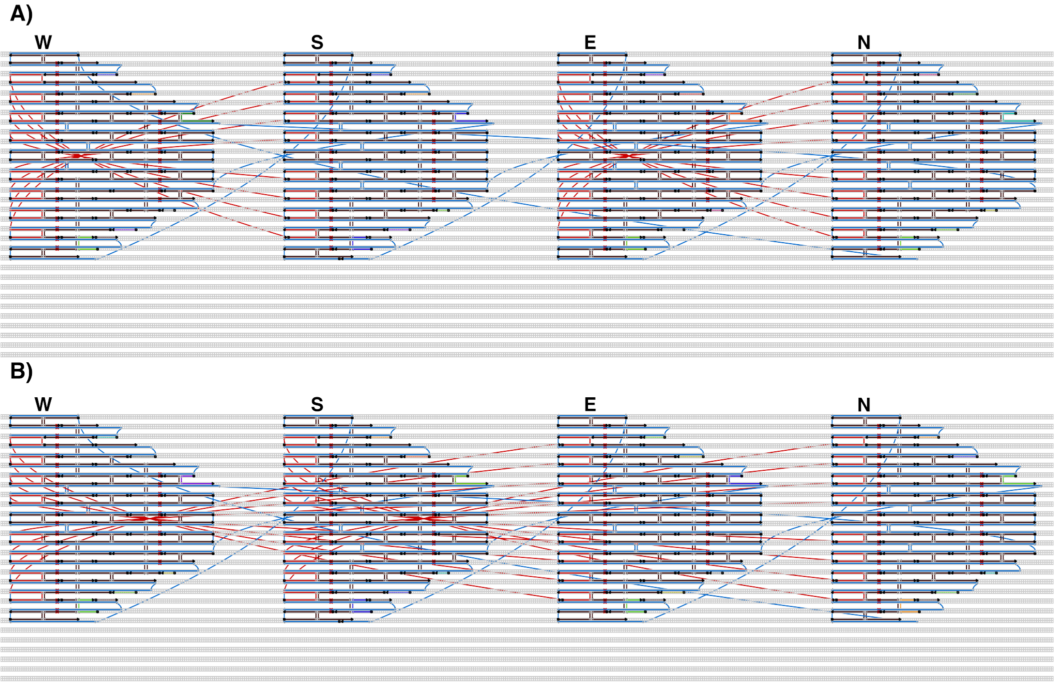


Figure 61: Cadnano design for the square-shaped nanostructure with the scaffold loop at the corner. A) Multimerisation by connecting West to South and North to East. Red staples show the connecting staples used for multimerisation. This case resulted in well-formed 2D crystals. B) multimerisation along the W-E and N-S axes. Red staples show the connecting staples used for multimerisation. In this case elongated crystals are formed. For better illustration of edge connections, connecting staples between trapezoids have been eliminated.

Acknowledgments

I would like to express my gratitude and very great appreciation to professor Friedrich C. Simmel for giving me the opportunity of working in his group for 8 years. These 8 years have been very important and helpful for thriving me in different life aspects.

I am very thankful of never-ending support by Helene, Susanne, Andrea and Tobias who keep the lab running properly and always provide helpful advices.

I would like to fully thank Sandra for the fruitful collaboration we had together that resulted in a successful paper.

Many thanks to my wonderful friends and colleagues, Aradhana, Aurore, Daniela, Swati, Enzo, Günther, Jonathan, Mario and Matthaus who not only provided me with great advices in different life matters but also a pleasant work environment.

I wish to acknowledge Qian lab, specifically professor Lulu Qian, Greg and Philip who I started working with since my stay abroad at Caltech and gained many valuable experiences.

I wanna give my special thank to Jungmann lab, in particular Max and professor Ralf Jungmann for their very helpful insights and collaboration during the DNA PAINT related projects.

My students, Melissa, Chris and Sebastian, it was always great to work with you on different topics and also a joy to overcome obstacles as a group together.

My admiration and thanks to staff members of chair E14 and E69 for all their helps and the nice atmosphere they created during the lab work.

Finally I am particularly grateful for the continuous assistance and support of my family specially my parents who, without them I would never reach the place I am today.

During my PhD work I was funded by IGSSE, ERC and NIM and would like to thank them for their support.