

Università degli Studi di Bologna

DOTTORATO IN SCIENZE CHIMICHE – XIX CICLO
SETTORE SCIENTIFICO-DISCIPLINARE: AREA 03 – CHIM/06
DIPARTIMENTO DI BIOCHIMICA “G.MORUZZI”

DESIGN, SYNTHESIS AND CHARACTERIZATION
OF DNA SUPRAMOLECULAR NANOSTRUCTURES

Il candidato

MARCO BRUCALE

Il relatore

PROF. MARGHERITA VENTURI

Il supervisore della ricerca

PROF. BRUNO SAMORÌ

Il coordinatore del corso di dottorato

PROF. VINCENZO BALZANI



When we have nothing to say, it is very hard to say nothing.

When we have nothing to do, it is very hard to do nothing.

[R. Fripp]

Preface

This thesis describes the results of three years of research at the Nanobioscience Lab of the Biochemistry Department “G. Moruzzi”, University of Bologna, under the supervision of Prof. Bruno Samorì and Dr. Giampaolo Zuccheri.

The research described here pertains to a field widely referred to as “structural DNA nanotechnology”. The basic idea of structural DNA nanotechnology is to use DNA molecules to fabricate nano-sized, geometrically well-defined constructs; the choice of using DNA as a nanoscale building block is due to its unique self-assembly capabilities. The spontaneous assembly of a group of DNA molecules in solution is strictly sequence-dependent, therefore it is possible to design sets of synthetic DNA strands, so that their sequences contain all the necessary information to build a specific nanoscopic object.

In this thesis, examples of novel DNA self-assembled systems of various nature are described. Most of the work presented herein has been published in widely accessible journals (see section 5.1). Each system was devised in an attempt to prototype a specific area of DNA self-assembly that we find promising for the future development of structural DNA nanotechnology.

Abstract

The geometrically deterministic self-assembly of nanoscale objects is a foremost goal of today's research. Biomolecules such as proteins or nucleic acids possess complex assembly capabilities surpassing by far in versatility and subtlety any artificial self-assembling system. It is possible to harness these intrinsic capabilities of biomolecules toward the construction of rationally designed nanostructures, using them as self-assembling building blocks. For its many unique characteristics, DNA is one of the most promising biological building blocks in this context. Since DNA helices are extensively characterized nano-objects enjoying a whole repertoire of interaction modes, DNA sequence design is also a form of "nano-architecture" that organizes matter according to a rational design in the nanoscale. In this thesis, I report the design, synthesis and characterization of various DNA self-assembling systems. Each system was designed and implemented in an attempt to answer a basic question about DNA self-assembly.

Although many reports exist on the results of DNA self-assembling systems, very few of them focus on the study of the dynamic processes leading to such results. Due to this, the mechanical knowledge of the involved processes is still fairly limited. To address an aspect of this issue, we designed and synthesized a series of DNA rhomboidal supramolecular tiles and studied their aggregation in various conditions. In this experimental effort, we show that a good degree of control can be achieved in the assembly of a 1D supramolecular polymer made of these rigid DNA tiles. Very rigid rod-like or circular structures can be obtained from the same tile thanks to the comprehension of the inner workings of the self-assembly of this system. The assembly of another similar system was also obtained by nucleation on a long ssDNA template instead than by tile aggregation, obtaining structures with extremely different characteristics.

To test the effect of the mechanical coupling of several flexible DNA nanostructures together, we experimentally compared the relative flexibility of an isolated four-way junction and that same junction inserted into a DNA parallelogram motif. Once the rigidity of the DNA rhombus motif was assessed, we designed and synthesized a rigid DNA parallelogram and employed it as a scaffold for the sequence-dependent positioning of proteins.

A dynamic DNA structure based on the formation and breakdown of an intramolecular cytosine–thymine (CT) motif DNA triple helix was designed and implemented. The switching can be performed repeatedly, quickly and independently of its local concentration without performance reduction over successive cycles. The controlled movement of the device was tested in solution and on the surface with both ensemble and single molecule experiments.

The design of a curved three-dimensional DNA origami prototype structure is also discussed.

Contents

CHAPTER I – INTRODUCTION	17
1 Towards Nanobiotechnology	19
1.1 What is Nanoscience? What is Nanotechnology?	19
1.1.1 Definition of Nanoscience and Nanotechnology	19
1.1.2 A (very) Brief History of Nanotechnology	20
1.1.3 Nanotechnology: Evolutionary and Revolutionary	24
1.2 How to do it?	25
1.2.1 Enabling Technology: Scanning Probe Microscopy	25
1.2.2 Bottom-Up and Top-Down Approaches	27
1.2.2.1 Top-Down	28
1.2.2.2 Bottom-Up	29
1.3 Nanobiotechnology	34
1.3.1 DNA-based fabrication in NanoBioTechnology	38
1.3.1.1 The Codes of DNA	38
1.3.1.2 Affinity vs. Specificity in DNA interactions	39
1.3.1.3 DNA in the Nanoscale: Shape, Dynamics, Flexibility	40
1.3.1.4 Sequence-dependent DNA-Surface Interactions	43
1.3.1.5 DNA for Bottom-Up Nanofabrication	45
1.3.1.6 DNA and Nanotubes	48
1.3.1.7 DNA and Metal Nanoparticles	51
1.3.1.8 Complex Self-assembled Architectures made of DNA	55

2	Structural DNA Nanotechnology	57
2.1	Static Architectures	58
2.1.1	Basic Element of Assembly I: DNA Junctions	59
2.1.2	Basic Element of Assembly II: “Sticky Ends”	61
2.1.3	Topological and Geometrical Dimensions of the Assemblies	64
2.1.3.1	Zero-Dimensional Topologies: Discrete Objects	66
2.1.3.2	One-Dimensional Topologies: Linear Arrays	68
2.1.3.3	Two-Dimensional Topologies: 2D Lattices	70
2.1.3.4	Three-Dimensional Topologies: 3D Crystals	76
2.1.4	Hierarchical Assembly	77
2.1.5	Sequence Design	80
2.1.6	Assembly Strategies for Aperiodic Structures	81
2.1.7	Enhancing the Structural Rigidity of the Nanostructures	86
2.1.8	Using Sequence Symmetry as a Tool	87
2.2	Dynamic Architectures	88
	<i>Chapter I Bibliography</i>	91

3	Novel DNA Self-assembled Nanostructures: Design, Synthesis, Characterization, and Preliminary Studies on Implementations	105
3.1	Structures based on the DNA rhombus motif	105
3.1.1	Design and synthesis of DNA parallelograms	106
3.1.2	Mono-dimensional arrays made of DNA parallelograms	112
3.1.2.1	Results	113
3.1.2.2	Discussion	121
3.1.2.3	Experimental	130
3.1.2.4	Conclusions	133
3.1.3	Assembly of the 1D array on a ssDNA template	134
3.1.3.1	The concept	134
3.1.3.2	Synthetic strategy	136
3.1.3.3	Production of the circular template for RCA	137
3.1.3.4	Rolling circle amplification of the template	139
3.1.3.5	Verifying the sequence of the RCA ssDNA product	140
3.1.3.6	Assembling the 1D parallelogram array on the template	142
3.1.3.7	Conclusions	143
3.1.4	Testing the mechanical rigidity of a DNA parallelogram	144
3.1.4.1	The concept	144
3.1.4.2	Assembling the structures	148
3.1.4.3	FRET experiments	149
3.1.4.4	Conclusions	150
3.1.5	A DNA parallelogram as a nanoscale scaffold for proteins	150
3.1.5.1	The concept	150
3.1.5.2	Design and synthesis of the DNA rhomboidal scaffold	151
3.1.5.3	Testing the incorporation of target sequences	154
3.1.5.4	Testing the incorporation of ssDNA/STV conjugates	157
3.1.5.5	Conclusion	158

3.2	A dynamic structure based on a DNA duplex-triplex transition	159
3.2.1	Rationale, design and synthesis	159
3.2.2	Experiments in solution: Static Characterization	162
3.2.3	Experiments in solution: Dynamic Characterization	164
3.2.4	Experiments on surfaces: Single Molecule Characterization	166
3.2.5	Experimental	168
3.2.6	Conclusion	170
3.3	A three-dimensional DNA Origami: preliminary studies	170
	<i>Chapter II Bibliography</i>	176

4	Conclusions	181
5	Appendixes	185
5.1	Publications produced during this PhD candidature	185
5.2	About me	188
5.3	Acknowledgements	189

I – INTRODUCTION

1 Towards Nanobiotechnology

1.1 What is Nanoscience? What is Nanotechnology?

1.1.1 Definition of Nanoscience and Nanotechnology

The etymology of the prefix “nano” stems from the Greek word “νάνος”, which means “dwarf”. Until quite recently, the main and perhaps only meaning of the prefix was “the billionth part of”, such as in “nano-second”. From the mid-70’s on however, the prefix started to acquire a different meaning, which is both more focused (it refers specifically to the length dimension) and looser (numerically). The most accurate depiction of the new meaning is in my opinion “pertaining to objects with smallest dimensions ranging from a few nanometers to less than one hundred nanometers”.^[1] It is a somewhat convoluted definition, but all of its parts are important to convey the correct concept. The part reading “pertaining to objects” means that “nano” does not apply to every *phenomenon* occurring in the scale of nanometers, but specifically to *objects*, and secondarily to those phenomena originated by nanoscale objects *only*. So for instance X-rays or UV radiation are not usually labeled as “nano”, but Surface Plasmon Resonance is. The “smallest dimension” part of the definition is important because it allows to encompass objects which have indefinitely large features in two spatial dimensions, but nanometers-sized features in the third. Finally, the numerical part of the definition “from *a few* nanometers to *less than one hundred* nanometers” intentionally limits “nano” to the range of sizes between the largest molecules and the smallest structures that can be fabricated by photolithography.

Although a fraction of the scientific community is still convinced of the contrary, nanoscience is not the billionth part of science, but the science of “nano” as per the second definition I gave in the paragraph above. Nanoscience can thus be viewed as an highly interdisciplinary area that can be tackled from the perspective of several classical scientific disciplines, including physics, chemistry, and biology.^[2] For a physicist, nanoscience could be for example the study of the behavior of electrons and photons in nanoscale structures. Nanostructures have a range of sizes in which quantum

phenomena are not “hidden behind” classical behavior, even at room temperature. Chemists routinely delved into nanoscience before it even existed: the chemistry and physics of polymers, micelles, colloids, phase-separated regions in block-copolymers is definitely nanoscience. Moving on to biology, nearly all of its most fascinating functional structures such as nucleic acids, viruses, proteins and sub-cellular organelles can be considered as nanostructures, and understanding their nanoscale mechanisms is a major goal of molecular biology.^[3]

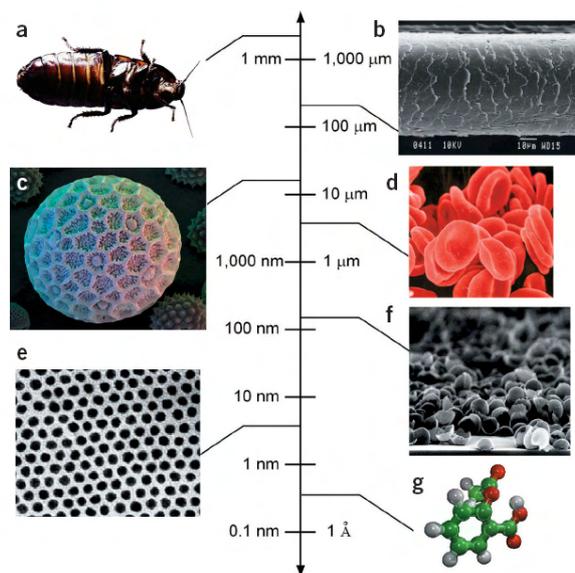


Figure 1 – Size comparison of progressively ‘small’ objects. (a) A cockroach (*Blattella Asahinai*) (b) A human hair (c) *Polygonum* pollen grain (d) Red blood cells (e) Cobalt nanocrystal superlattice (f) Aggregate of half-shells of Palladium (g) Acetylsalicylic Acid. [Image adapted from ref. 34]

Once nanoscience is unambiguously defined, an univocal definition of “nanotechnology” is straightforward: it can be defined as the application of nanoscientific knowledge for a practical purpose. However, during the last decade the focus of nanotechnological research has drifted considerably from the concepts that originally led to the emergence of the field.

1.1.2 A (very) Brief History of Nanotechnology

“What I want to talk about is the problem of manipulating and controlling things on a small scale.”

[Richard Phillips Feynman, Dec 29, 1959]

The first documented mention of some of the distinguishing concepts of nanotechnology predates the use of the term. On December 29, 1959, during an American Physical Society meeting at Caltech, physicist Richard P. Feynman gave a talk titled "There's Plenty of Room at the Bottom". In this talk, Feynman stated for the first time several ground-breaking concepts, most of which are still relevant today. I will briefly mention the main points in Feynman's talk. Firstly, he noted how the simple process of writing something on a surface can be in principle progressively miniaturized until the "letters" are merely a few atoms across:



Figure 2 – Richard Phillips Feynman [unknown photographer, ca. 1962]. Feynman is known for expanding the theory of quantum electrodynamics, the physics of the superfluidity of supercooled liquid helium, and particle theory. Feynman won the Nobel Prize in Physics in 1965, together with Schwinger and Tomonaga; he developed a way to understand the behavior of subatomic particles using pictorial tools (known as Feynman diagrams).

“And it turns out that all of the information that man has carefully accumulated in all the books in the world can be written in this form in a cube of material one two-hundredth of an inch wide--- which is the barest piece of dust that can be made out by the human eye. So there is plenty of room at the bottom! Don't tell me about microfilm!”

[Richard Phillips Feynman, Dec 29, 1959]

Feynman then suggested that an huge improvement in microscopy technology would be needed to embark in that sort of task. It is frankly hard not to be immediately reminded of the impetus nanotechnology received from the creation of new analytical tools such as the atomic force microscope and the scanning tunneling microscope.

“Unfortunately, the present microscope sees at a scale which is just a bit too crude. Make the microscope one hundred times more powerful, and many problems of biology would be made very much easier.”

[Richard Phillips Feynman, Dec 29, 1959]

Another key concept proposed by Feynman in this talk is that the very small level manufacturing is not limited to a static representation of data (as suggested by the writing analogy he used at the start of his talk). Rather, once in possess of the right tools, one could think of producing *active* entities, much like those found in biological systems, and build complex tools like computers in this tiny scale:

“Biology is not simply writing information; it is doing something about it. A biological system can be exceedingly small. Many of the cells are very tiny, but they are very active; they manufacture various substances; they walk around; they wiggle; and they do all kinds of marvelous things - all on a very small scale. Also, they store information. Consider the possibility that we too can make a thing very small which does what we want - that we can manufacture an object that manoeuvres at that level!”

[Richard Phillips Feynman, Dec 29, 1959]

Then, Feynman proposed that the process of miniaturized fabrication could be brought forward to its very extreme – the direct manipulation of atoms. He also described a process by which the ability to manipulate individual atoms and molecules might be developed, using one set of precise tools to build and operate another proportionally smaller set, so on down to the atomic scale.

“But I am not afraid to consider the final question as to whether, ultimately - in the great future - we can arrange the atoms the way we want; the very atoms, all the way down! What would happen if we could arrange the atoms one by one the way we want [...].”

[Richard Phillips Feynman, Dec 29, 1959]

Feynman also envisioned a new field of applied research that is very similar to something that material scientists are trying to do nowadays:

“What could we do with layered structures with just the right layers? What would the properties of materials be if we could really arrange the atoms the way we want them? [...] I can't see exactly what would happen, but I can hardly doubt that when we have some control of the arrangement of things on a small scale we will get an enormously greater range of possible properties that substances can have, and of different things that we can do.”

[Richard Phillips Feynman, Dec 29, 1959]

The first reported use of the word “nanotechnology” was in 1974, when Tokyo Science University Professor Norio Taniguchi in the paper “On the Basic Concept of ‘Nano-Technology’” defined it as “the processing of, separation, consolidation, and deformation of materials by one atom or one molecule”.[4]

In 1986, Dr. Kim Eric Drexler published a book entitled “Engines of Creation: the Coming Era of Nanotechnology”.[5] This book speculatively explored the concepts proposed by Feynman bringing them to a wide public attention. Drexler’s envisioned nanotechnology is often dubbed “molecular nanotechnology”, that is, the engineering of functional machines at the molecular scale designed and built atom-by-atom.

“Although inspired by biology (where nanomachines regularly build more nanomachines despite quantum uncertainty and thermal motion), Feynman’s vision of nanotechnology is fundamentally mechanical, not biological. Molecular manufacturing concepts follow this lead.”

[K. E. Drexler]

Drexler elaborated Feynman’s idea of using nanomachines to build complex products (including further nanomachines), by *positionally-controlled* mechanosynthesis guided by molecular machine systems (see Figure 3). Conventional chemistry employs stochastic processes driven toward some equilibrium to obtain stochastic results, and deterministic results are obtained only through complex enzyme-catalyzed reaction chains (such as in biological systems). In contrast, molecular nanotechnology would employ novel (and as yet unspecified) deterministic nanoscale processes to obtain deterministic results.

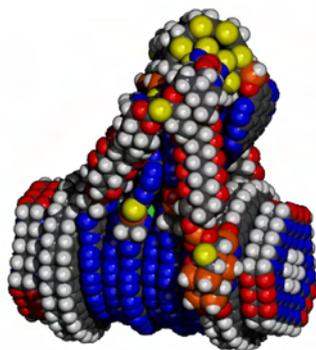


Figure 3 – Structure of an (hypothetical) nanoscale positional device with six degree of freedom proposed by Drexler and Merkle. A general-purpose molecular assembler arm should be able to move its “hand” by many atomic diameters, position it with fractional-atomic-diameter accuracy, and then execute finely-controlled motions to transfer one or a few atoms in a guided chemical reaction. [Copyright institute for molecular manufacturing www.imm.org]

Drexler's idea of molecular nanotechnology entails the placement of molecular moieties in specific positions to obtain desired chemical reactions. The most useful conceivable nanoscale machine, Drexler reflects, is a programmable "universal assembler" capable of building other structures. How to build this type of universal assembler is to date still unclear, and all Drexler's further elaboration ultimately takes for granted the feasibility of such a machine.^[6]

Several researchers have criticized the notion of universal assemblers since Drexler's proposal. A heated debate between Drexler and Nobel Prize winner Prof. Richard Erret Smalley was published in various journals, starting in 2001. Smalley was skeptical, in particular, about the very notion of mechanochemical synthesis:

"Much like you can't make a boy and a girl fall in love with each other simply by pushing them together, you cannot make precise chemistry occur as desired between two molecular objects with simple mechanical motion along a few degrees of freedom in the assembler-fixed frame of reference. Chemistry, like love, is more subtle than that. You need to guide the reactants down a particular reaction coordinate, and this coordinate treads through a many-dimensional hyperspace."

[R. E. Smalley, Chemical and Engineering News CENEAR 2003 (81) 37-42]

It is in my opinion very difficult to identify clear winner of the debate (as both contestants progressively tone down their respective claims). Nevertheless, Drexler's vision (although never scientifically refuted) seems to have fallen "out of fashion" during the last years. This could perhaps be because its tangible accomplishment is so beyond the current technological possibilities that it's even difficult to plan research focusing explicitly on Drexler's ideas.

During the last decade, the first viable nanotechnological products began to appear, and in some cases even to acquire a commercial status,^[7] thus taking *nanotechnology* out of the purely speculative.

1.1.3 Nanotechnology: Evolutionary and Revolutionary

As with every other technology, nanotechnology grows from applied science – in this case, nanoscience. As Professor George M. Whitesides notes in his paper "Nanoscience, Nanotechnology, and Chemistry"^[1] published in 2005, nanoscience has been with us for at least a decade, but technologies growing from it are still few, and the rate at which

they have appeared seems slower than that in areas such as, for example, biotechnology. Nevertheless, Whitesides notes, commercial nanotechnology exists and is in “the robust health of early childhood”.^[7] The type of currently marketable nanotechnologies however are still *evolutionary*, in the sense that they are based on products that already exist, and that have micrometer- and nanometer-scale features. Examples of evolutionary nanotechnologies are sub-100-nanometers electronics, several branches of materials science, and almost every aspect of applied chemistry.

The more interesting question is however when (or whether) there will be a *revolutionary* nanotechnology, based on fundamentally new nanoscience, and allowing us to build products that are as now unforeseeable.^[1] This type of nanotechnology will emerge perhaps from new nanostructured materials and their electronic properties, or innovative architectures of data storage systems, or artificial nanoscale devices mimicking actions performed by naturally occurring biological devices and machines.

1.2 How to do it?

1.2.1 Enabling Technology: Scanning Probe Microscopy

The wide diffusion of Scanning Probe Microscopy (SPM), opened in 1982 with the invention of Scanning Tunneling Microscope (STM),^[8] provided a great momentum to nanoscience research. Although other spectroscopic and microscopic techniques able to infer data about nanoscale structures were already available, the versatility and simplicity brought by the scanning probe approach aided its success. Moreover, since in SPM the probe is actually an object in physical proximity to the sample, SPM is particularly suited to be turned into a hybrid technique capable of both observing and modifying the nanometer-scale-sized sample. Using SPM to create nanoscale shapes and patterns is usually called Scanning Probe Lithography.

The basic concept of SPM is astonishingly simple. A physical probe of some sort is mechanically moved along a raster scan path over the sample, and the probe-sample interaction is recorded as a function of position. The type of probe-sample

interaction that is being monitored during the observation is usually called a “mode” of SPM.

SPM Mode	Signal	Probe	Resolution	Samples	Environment
STM	Tunneling current	Conducting wire	1 Å	Solid phase conductors	Vacuum, air
SFM (AFM)	Force	Flexible cantilever	1 Å	Surfaces	Vacuum, air
KPFM	Potential offset	Reference electrode	1 Å	Surfaces	Vacuum, air
SNOM	Transmission of photons	Wave guide	100 Å	Surfaces	Vacuum, air, liquid
SICM	Ion conductance	Micro-pipette	200 Å	Surfaces	Ionic solution
SECM	Faradic current	Micro-electrode	1000 Å	Surfaces	Ionic solution
SThM	Transmission of heat	Micro-thermocouple	1000 Å	Surfaces	Vacuum, air
SMFM	Magnetic force	Magnetized cantilever	100 Å	Magnetic Surfaces	Vacuum, air

Table 1 – List of most widespread SPM modes. Several other modes are fully established but are less widespread.

A respectable number of different SPM modes (see Table 1) are fully established today, and they cover a broad spectrum of information gathered, type of observable specimens and experimental conditions (including near-physiological conditions for most modes). Another feature that characterizes the various SPM techniques is the point-spread function limiting their resolution, that is, how they respond to a point object. In SPM, this corresponds to the volume of the imaginary shape that contains all the probe-sample interactions, and for some modes this amounts to just a few picometers. Combined together, these features allow to use SPM (in particular Atomic Force Microscopy)^[9] to analyze the nanoscale structure of biological samples in near-physiological conditions.^[10, 11]

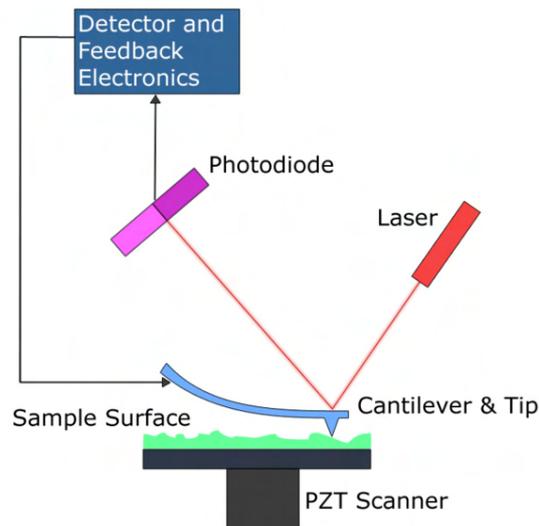


Figure 4 – Typical setup of a scanning probe microscope (in this case an AFM). The probe (or the sample under a stationary probe) is moved by a piezoelectric scanner, usually following a raster pattern. The sensor data forms an image of the probe-surface interaction. Feedback from the sensor is used to maintain the probe at a constant force or distance from the object surface. For atomic force microscopy the sensor is a position-sensitive photodetector that records the angle of reflection from a laser beam focused on the top of the cantilever.

Of course, the widespread ability to observe nanoscale structures and features on a surface rendered feasible many research projects that could be beyond reach prior to SPM. But for all its wonderful characteristics, SPM also reveals several crucial limitations when faced with the tasks required by nanotechnology. The single most obvious limit is that SPM is basically two-dimensional: it is great to infer information about surfaces or objects that can be deposited onto surfaces, but it cannot observe objects freely floating in a three-dimensional environment. Moreover, the acquisition rate is usually very slow compared to the kinetics of almost all interesting nanoscale processes. It is tempting to imagine what could happen if we had access to a fully three-dimensional SPM (or something similar), or if the acquisition time of a standard, 2D AFM was a few orders of magnitude faster.

1.2.2 Bottom-Up and Top-Down Approaches

The single most important question that needs an answer in nanotechnology is how to actually build nanoscale objects deterministically. All the possible strategies fall in one of two very broad classes. The two approaches are called “Top-Down” and “Bottom-

Up” after the type of action that is needed to go from the starting materials to the desired product.

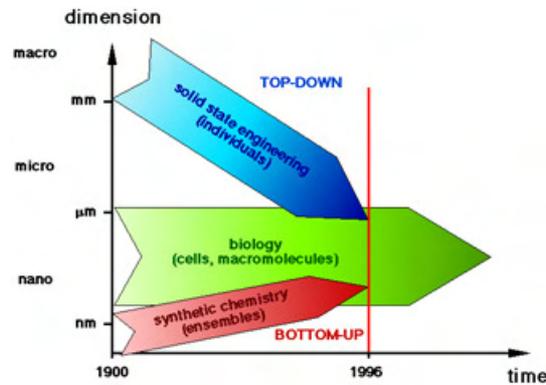


Figure 5 – Schematic depiction of the two established approaches toward nanoscale construction. [Image taken from the Swiss Federal Institute of Technology www.nanotechnology.ethz.ch]

1.2.2.1 Top-Down

The Top-Down approach represents the type of nanofabrication first imagined by Feynman and Drexler, and seeks to create nanoscale structures or devices by using larger, externally-controlled tools to direct their assembly. Examples of Top-Down nanofabrication include all types of lithography (photo-, electron beam, dip-pen, soft-, nanoimprint-) and etching techniques. In all these techniques, the generation of nanoscale features relies on the externally controlled relative movement of a tool and the substrate.

Many of these techniques are evolutions or refinements of pre-existing techniques that were traditionally used for microfabrication. The classic example of this is photolithography. In the late 1960s, state of the art photolithography allowed to trace features not smaller than 5 μm using visible light at 436 nm (Hg G-line). Continuous refinements improved the technology, and in 2006 IBM and JSR Micro Incorporated jointly demonstrated that High-Index Immersion Lithography using 193 nm light and special immersion fluids can be used to produce features of less than 30 nm.

Not all the definitions of the top-down approach found in literature are identical, and a quite common variation is to label as “top-down” a technique that specifically *removes* matter from a bulk material to fabricate nanoscale objects or features.

1.2.2.2 Bottom-Up

The Bottom-Up approach involves starting from constituent sub-components of the desired structure (e.g. a pile of bricks or molecules) and *let them assemble* together to get the final product (e.g. a house or a supra-molecular object). Of course, the idea is not to exert a direct control over the position of each component, but to exploit a specific physical property of the components in order to predict how their stochastic interactions will force them to assemble in a as deterministic as possible fashion. In other words, once the basic type of interaction between the sub-components are thoroughly understood, the strategy is to put them in a context where they can freely find each other and establish those interactions, without or with very limited external control. In this way, the result of the assembly is ultimately dependent on (and implicitly pre-determined by) exactly *how* sub-component can bind to each other. Since no direct action is taken to drive the components to the position assigned to them, this process is often called “self-assembly”. In the context of nanoscience and nanotechnology, the self-assembling components are obviously nano-sized objects, for examples molecules or colloidal particles. Bottom-up approaches based on self-assembly could in principle be able to produce the same structures accessible to top-down methods, but in a highly parallel (and supposedly much cheaper) fashion.

The key to successful nanoscale self-assembly is understanding which spontaneous process of aggregation between the “bricks” can be exploited at this scale.^[12-17] The obvious places to search to get the right inspiration are biological systems, where nature has harnessed chemical forces to create essentially all the structures needed by life. The types of interactions used by biological systems to construct nanostructures are chemical bonds of various nature. For this reason, chemistry could be considered the ultimate nanofabrication tool.

(i) Chemical Synthesis and Nanofabrication

Chemical synthesis is a sub-nanometer scale fabrication activity that can be performed on megaton scale, and it began to be systematically explored centuries ago. So one very legitimate question usually put forward by chemists is, “why look

elsewhere? isn't the organic synthesis we know already the best possible nanofabrication tool?" My personal opinion is that the answer could be something between yes and no. The problem with using traditional organic synthesis as a nanofabrication tool is that it's a too fine tool for today's needs. Nature knows how to synthesize a 50 kDa protein bond by bond, that then folds in a certain way to organize in space its reactive groups and perform further chemical synthesis. Unfortunately we're not that smart yet. It's very difficult to synthesize huge and complex molecules with hundreds of different functional groups, or to predict exactly how a very large molecule will fold or move around in solution. Today's nanoscientists usually want to build something that has a well-defined, persistent shape and that can be interfaced with micrometer-scale systems, at least to study it more easily.

Of course, that means structures usually a lot bigger than the biggest *individual* molecules afforded by rational synthesis, even using convergent strategies. Moreover, some unique functions such as light responsiveness, magnetic properties, catalytic activity and others usually appear only from groups of molecules, and not from individual, isolated ones. A solution to this is to use covalent polymers, but unfortunately the chemistry of polymers is almost completely statistical and does not (unless in the most peculiar circumstances) allow a direct and deterministic control of the structures it generates. So basically nanoscientists exploring the bottom-up nanofabrication approach need a specific type of chemistry that yields products reaching dimensions of tens of nanometers, permits self-assembly and allows for the precise positioning of each sub-component involved. The widespread response to this problem is supramolecular chemistry.^[12-17]

Supramolecular chemistry is an area of chemistry which focuses on forming reversible, non-covalent bonding interactions between molecules to afford multi-molecular products. While traditional organic synthesis focuses on the formation and breaking of covalent bonds, supramolecular chemistry utilizes far weaker and reversible non-covalent interactions, such as hydrogen bonding, metal coordination, hydrophobic effects, van der Waals forces, π - π interactions, and/or electrostatic effects to assemble molecules into multi-molecular complexes. Fundamental concepts such as host-guest chemistry, constitutional dynamic chemistry, self-assembly, and molecular recognition are easily demonstrated and explored by means of supramolecular chemistry.^[16-19]

A large part of the chemistry performed in biological systems is supramolecular chemistry. The quaternary structures of proteins, how nucleic acids can split and reassemble permitting access or protecting their code, or how a lipid bilayer forms, are all examples of supramolecular self-assembly occurring in cells. The ultimate goal of bottom-up nanotechnology is mimicking nature's ability of constructing a huge variety of structures, devices and machines of several disparate dimensions employing supramolecular self-assembly.

(ii) Self-Assembly in Action

Self-assembly is a universally pervasive phenomenon which generates structural organization on all scales from molecules to galaxies.^[20, 21] It is defined as a processes in which pre-existing parts or disordered components of a pre-existing system form structures of patterns. Self-assembly can be classified as either static or dynamic. Static self-assembly is when the ordered state is reached when the system is in equilibrium and does not dissipate energy. Dynamic self-assembly is when reaching the ordered state requires dissipation of energy. The types of self-assembly relevant to nanoscience are the molecular and supramolecular ones. The basic phenomenon behind molecular self-assembly in solution is that individual molecules will always seek the lowest energy level available to them, continuously reorienting their physical position in order to optimize the interactions with other molecules.

Successful self-assembly in a molecular context is dependant on four specific characteristics of the system: (i) the type of sub-components, (ii) the possible interactions between sub-components, (iii) the "adjustability" of the interactions once they are formed (this includes reversible interactions), (iv) the relative mobility of sub-components, in turn influenced by the environment in which the assembly takes place.

(i) The system must be formed by a group of molecules, or specific parts of a macromolecule, that have the potential to interact with one another in some determined way. Thus the formation of the interactions brings the system from a relatively more disordered state to a less disordered state – for example, from a solution of individual molecules to a linear polymer of the same molecules, or from a random-coiled linear polymer to a folded structure.

(ii) How sub-components interact to bind to each other is the most obvious parameter influencing the outcome of the assembly. Most natural self-assembling systems rely on weak interactions, where ‘weak’ in this context means comparable to thermal energies accessible to the system. Examples of interactions of this type are the van der Waals interactions, hydrogen bonding, π - π stacking, weak coordination bonds, hydrophobic interactions. The shape of the components and the relative position of their interacting parts is also crucial in determining the outcome of the assembly: a less than ideal compatibility of the components’ geometries can impair the proper formation of interactions.

(iii) The maximum amount of order is generated with molecular self-assembly if the system allows for a certain grade of error-correction. Practically and very generally, this means that the system must have the possibility of either adjusting the position of a sub-component once it is included in an aggregate, or continuously exchanging sub-components between bound and free state (this happens if aggregates are formed through reversible interactions). This implies that the energy of the interactions keeping the component in its assigned place in the assembly must be comparable to the energies of phenomena that tend to disrupt the assembly (see Figure 6a). For a collection of molecules in solution, these ‘destructive’ phenomena are generated mostly by thermal motion, and thus an optimal temperature exists for conducting the molecular self-assembly. This optimal temperature is the one at which the forces exerted on the aggregate (mostly by Brownian motion of the solvent) can barely disrupt the interactions that keep it together. If a collection of different interactions is used to build the aggregate, each with different energies, the best strategy would thus be to subject the system to a thermal scan from high temperature to low temperature (usually called an annealing). If the temperature is too low with respect to the interactions formed, components irreversibly bind every time they meet, there is no possibility of breaking the bonds and no increase of order is generated through self-assembly. An example of this is the formation of amorphous solids from rapidly cooled liquids, as opposed to the formation of crystals by the same liquids, but cooled slowly (see figure 6b-d).

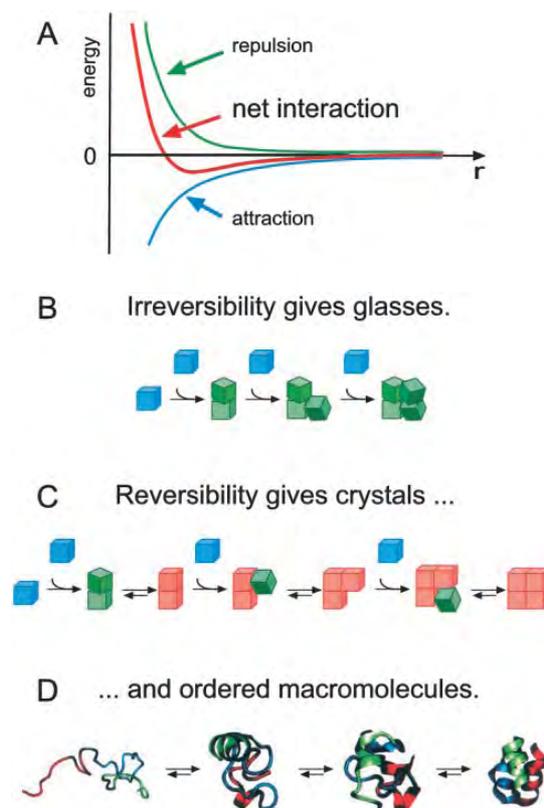


Figure 6 – (a) Aggregation occurs when there is a net attractive interaction from components, resulting from the balance of attraction and repulsion. (b,c) Effect of reversibility or irreversibility of the net interaction on the outcome of the assembly. Irreversibility leads to amorphous “glasses”, whereas reversibility leads to ordered self-assembled “crystals”. In the schemes, blue blocks are individual sub-components not included in the growing assembly. Pink blocks are components included in the assembly in the lowest-energy situation. Green blocks are components included in the assembly, but not in the lowest-energy form. (d) Many examples of biological self-assembling systems fall in the case described in panel ‘c’. Here, a linear polypeptide is reversibly folded into a complex asymmetric three-dimensional nanostructure (a protein). [Adapted from ref. 21]

(iv) For any assembly to occur, of the components must of course be able to find each other, i.e. be mobile. For molecules in solution, it is again the thermal motion that usually concurs mostly to provide the necessary mobility. Externally-induced convection or mechanical agitation can also influence the outcome of the assembly.

(iii) Complexity and Defects in Self-Assembly

The most simple conceivable molecular self-assembling system includes only one type of molecule, with only one mode of interaction with itself. In this case, the result of the self assembly depends on the ‘supramolecular valence’ of the molecule: for example if it can bind to two copies of itself, the self-assembly will generate a topologically linear, homogeneous construct; if it can bind to three the result will be a

topologically two-dimensional structure, and so on. The point is that even in this simple system, the shape of the assembly is implicitly stated in the mode of interaction available to the sub-component. If one wants to build more complex structures by means of molecular self-assembly, it is necessary to have more sub-components than just one, having more types of interaction available. Things start to get really interesting when the type of interactions between individual components are diverse enough to drive them to the formation a complex structure.

In this respect, one very important and still vastly unanswered question is, what is the inherent limit of complexity accessible with molecular self-assembly in solution. It seems reasonable to think that this limit will be determined by the interplay of error incorporation and error correction rates during self-assemblies. Nature has a number of strategies for limiting defects in its assemblies. One of these is the use of templates as guides toward the correct assembly in presence of many competing processes. The most spectacular example is probably that of chaperonin-assisted protein folding.^[22] In the last years, the group of Eric Winfree at Caltech started to investigate a number of different strategies for designing self-correcting, self-assembling systems, experimentally checking their theories on artificial DNA structures (see section 2.1.3.3).^[23-28]

1.3 Nanobiotechnology

“So biology and chemistry, not a mechanical engineering textbook, point in the direction we should look for answers [...] one should start with biology, which offers a cornucopia of designs and strategies that have been successful at the highest levels of sophistication. In tackling a difficult subject, it is sensible to start by studying at the feet of an accomplished master. Even if they are flagella, not feet.”

[G. M. Whitesides, “The Once and Future Nanomachine”, 16 Sept 2001 Scientific American]

Living systems are doubtlessly an example of very efficient bottom-up nanotechnology. Organisms routinely perform self-assembly tasks of awe-inspiring complexity, building devices capable of chemical synthesis, light responsiveness, motion on a disparate range of scales, fine sensing of analytes, self-repair, self-replication, memory and thought. When faced with the task of designing and implementing molecular self-assembling systems with some desired characteristics, it is usually a smart idea to check how

similar systems are assembled in living systems, and how they work; the very idea of a supramolecular combinatorial chemistry was probably inspired by reasoning about biological macromolecules.

One interestingly recurrent theme in self-assembled biological systems is the “smart” folding of a topologically linear polymer into a well-defined three-dimensional structure.^[29] This strategy is the biological response to the difficult three-dimensional pick-and-place fabrication methods employed by man-made macroscopic machinery, and hypothesized for Drexler’s universal assemblers. The typical example of this are proteins or nucleic acids, topologically linear polymers that can fold into geometrically defined three-dimensional structures. Typically, the biological function of proteins is critically dependent on their correct folding, and therefore the precision of the folding is crucial. The active sites of enzymes for example are capable of recognizing a specific host structure, perform chemo- regio- and stereo-selective synthesis, then release the modified host. For this to happen, specific functional groups must be arranged in three dimensions (see Figure 7), with specific orientations and positions, with sub-nanometer precision. The functional groups concurring to the functioning of the active site can be extremely far apart if the protein is linearized.

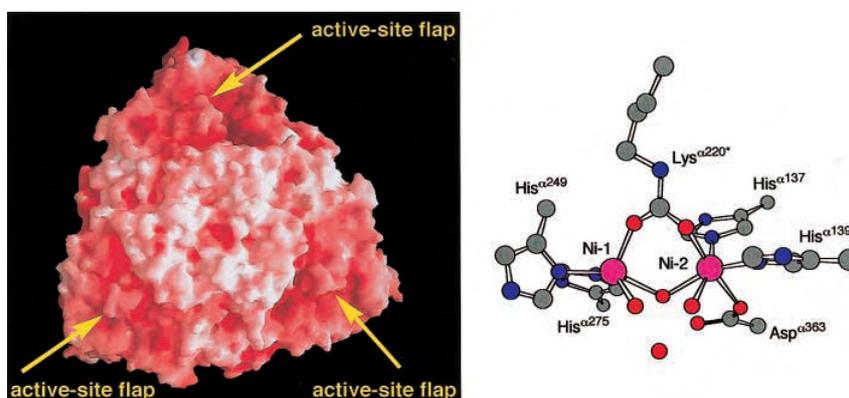


Figure 7 – (Left panel) Solid surface representation of the electrostatic potential of *B.pasteurii* urease. (Right panel) Coordination geometry of the active site nickel ions. Functional groups from several non-adjacent aminoacidic residues are driven by self-assembly to form this complex three-dimensional nanoscale structure capable of catalytic activity. The orientation and distance of each functional group is of paramount importance in allowing the protein to function appropriately. [Image taken from Musiani et al, J Biol Inorg Chem 2001 (6) 300-314]

But the question is, why is the strategy of folding 1D structures to obtain 3D structures so recurrent in biology? The reason lies in the fact that to be described properly, three-dimensional structures of this complexity need a lot of information. In

self-assembling systems, this information (the project of the assembly) must be intrinsically included in the system itself. The solution to this problem found in living systems is superlatively elegant and effective. Biomolecules involved in a self-assembly process are at the same time the physical repository of the project of the assembly, and the structural sub-units that are assembled according to this project. For this to actually work, nature employs only a small set of basic building blocks to build its most complex polymers- for example, peptides for proteins and nucleotides for nucleic acids. Since in a given environment (physiological conditions in a living organism) these building blocks have specific, determined modes of interaction, they can be used simultaneously as physical building blocks, and as *digital* bits of information about the structure of the assembly. Nature ultimately translates complex three-dimensional structures into numbers, just as Hilbert and Gödel translated complex logical statements into arithmetic. But a random collection of bytes does not contain any information (apart from their individual states) unless they have an *order* in which they must be read. And the most simple way to represent this order is to link them, one after another, in a one-dimensional matrix; hence, in terms of molecules, a linear polymer in which the beginning and end are different. In conclusion, linear polymers formed by a limited number of components in a specific order are common in nature because they are the most efficient way of storing *digitalized and compressed information*, including the shapes of three-dimensional structures and the materials needed to build them.

The bottom-up approach to nanofabrication has the ambition to learn how to use biological strategies for its ends. Unfortunately, we are currently unable to design from scratch complex information-containing polymers with the same efficiency displayed by nature, not to mention synthesizing them. So one slightly different approach is to use not only the basic strategy, but also the actual molecules found in biological systems, or specific portions of them. This approach is usually described as “nano-bio-technology”.^[30-34] The basic idea is to extrapolate biomolecules from the specific role they have in the cell, and use some of their unique characteristics towards a technological goal.

An especially elegant example of this approach was reported in 2002 by the group of Uri Sivan.^[35] Through the ingenious use of DNA and the RecA protein, they were able to produce self-assembling Field Effect Transistors operating at room

temperature. Multiple RecA proteins polymerize at a single stranded DNA (ssDNA) molecule. By homologous recombination, the resulting DNA-protein complex recognizes and binds to the complementary target sequence present in a DNA duplex. The DNA-protein complex was then treated with silver ions, which are reduced by aldehyde groups previously generated in the ds-DNA target. The resulting small silver grains are later used for the wet-chemical deposition of gold. This procedure leads to the formation of a conductive wire, with an insulating gap precisely at the position where RecA was bound. The information encoded in the DNA molecules thus replaces the masks used in conventional lithography, while the RecA protein serves as the resist. In this example, both the molecules and the specific functions they perform are the same found in living systems, but their capabilities are employed for a non-biological goal.

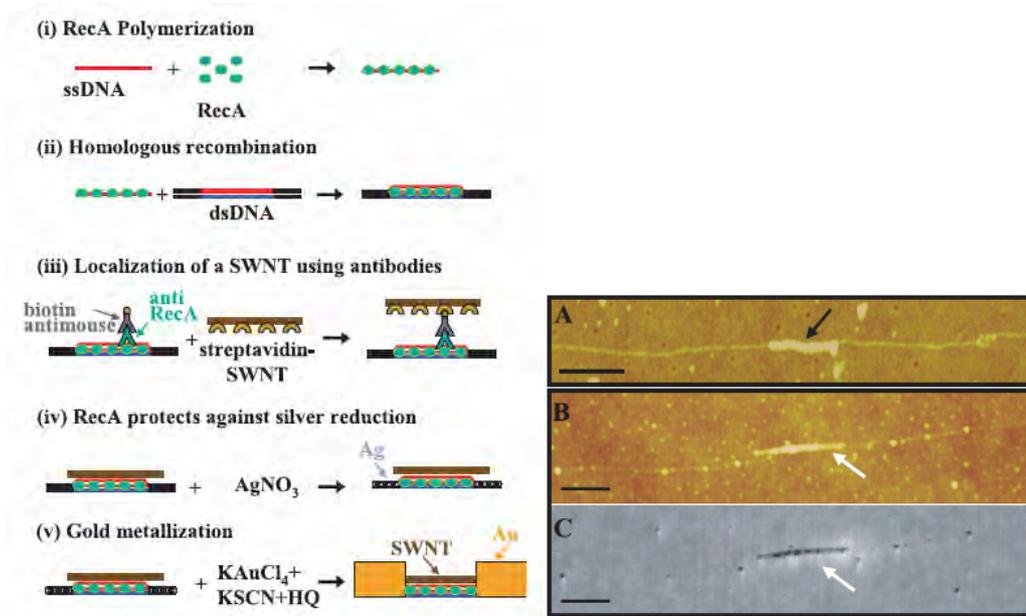


Figure 8 – (Left panel) Assembly of a DNA-templated FET and wires contacting it. Steps are as follows: (i) RecA monomers polymerize on a ssDNA molecule to form a nucleoprotein filament. (ii) Homologous recombination reaction leads to binding of the nucleoprotein filament at the desired address on an aldehydederivatized scaffold dsDNA molecule. (iii) The DNA-bound RecA is used to localize a streptavidin-functionalized SWNT, utilizing a primary antibody to RecA and a biotin-conjugated secondary antibody. (iv) Incubation in an AgNO_3 solution leads to the formation of silver clusters on the segments that are unprotected by RecA. (v) Electroless gold deposition, using the silver clusters as nucleation centers, results in the formation of two DNA-templated gold wires contacting the SWNT bound at the gap. (Right panel) Localization of a SWNT at a specific address on the scaffold dsDNA molecule using RecA. (A) An AFM image of a 500-base-long (~250 nm) RecA nucleoprotein filament (black arrow) localized at a homologous sequence on a λ -DNA scaffold molecule. Bar, 200 nm. (B) An AFM image of a streptavidin-coated SWNT (white arrow) bound to a 500-base-long nucleoprotein filament localized on a λ -DNA scaffold molecule. Bar, 300 nm. (C) A scanning conductance image of the same region as in (B). [Images and caption taken from reference 35].

1.3.1 DNA-based fabrication in NanoBioTechnology

Among all the biological molecules that can be employed for nanotechnology applications, DNA stands out for its several unique features, which makes it perhaps the most versatile nanoscale building block known to date. In the next few paragraphs I will review the ones most relevant to the context of nanobiotechnology.

1.3.1.1 The Codes of DNA

More than 50 years after the discovery of the structure of the DNA double-helix (see figure 9), our gratitude towards J. D. Watson and F. H. C. Crick is renewed daily, as the many wonders of this molecule never stop to suggest new pieces of research. On analyzing DNA fiber diffraction data, they described the canonical B structure, the most common in living organisms and in normal solution conditions. They had already envisaged that “the specific base pairing immediately suggests a possible copying mechanism for the genetic material” laying the foundations for the comprehension of the code underlying the functioning and heredity of all living organisms. DNA itself also controls the expression of codes written in its base sequence, for instance through the control of protein recognition mechanisms that are based on the modulation of its structure and dynamics along the chain. Ultrastructural characterization techniques methods led towards the discovery that the *codes* contained in the DNA base-sequence^[36] rule these interactive processes from the atomic scale of the single base-pair level to the nanometer and micrometer scale-lengths of its superstructures.

The term “code” was defined by Trifonov as “any pattern or bias in the sequence which corresponds to one or another specific biological (biomolecular) function or interaction.”^[37] The codes of DNA are generally chemical in nature, mostly structural: the stereochemistry of interaction between a pair of aromatic systems determines the base-pairing specificity. On a larger scale, the composition in space of many local chain deformations drives, for instance, the DNA wrapping around the histone proteins in nucleosomes in chromatin. The informational codes of DNA can provide also serve as a toolbox of assembly information that can be used to switch self-organization among different length and energy scales. The field of DNA nanotechnology has been so far

relying heavily on the Watson-Crick base-pairing code but there is room for much more, as other informational codes are also available to DNA.

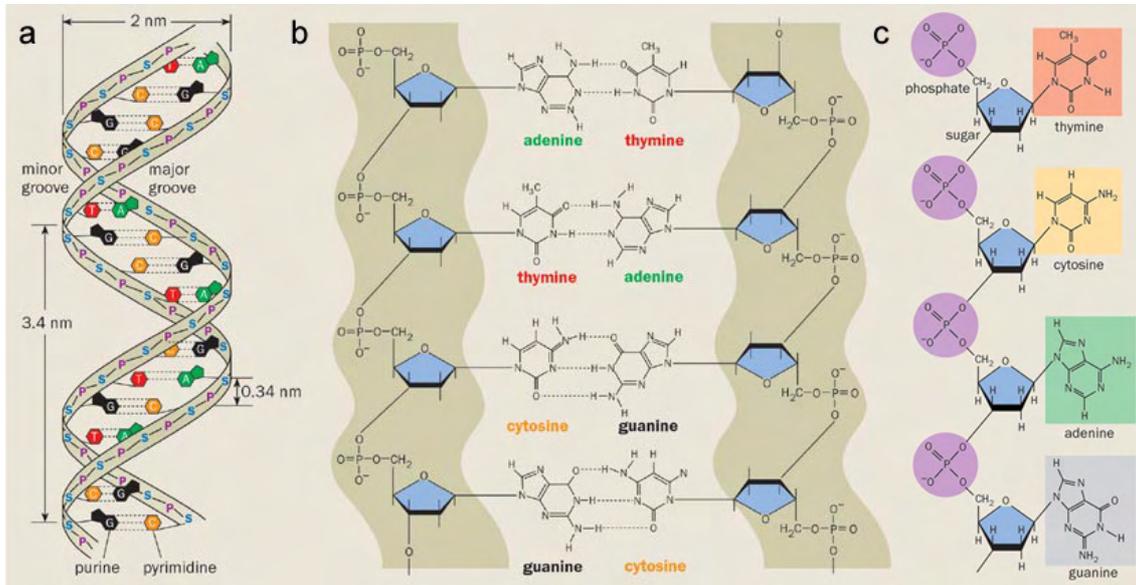


Figure 9 - (a) The DNA double helix has two polynucleotide chains that run in opposite directions and are twisted around each other. Hydrogen bonds hold complementary base pairs together - adenine (green) with thymine (red) and guanine (black) with cytosine (orange). The base pairs lie 0.34 nm apart axially, and there are approximately 10.5 of them within each full turn of the B-DNA double helix. (b) An untwisted view of the two chains. Each complementary base pair is held together by aligned hydrogen bonds and occupies almost exactly the same amount of space between the attachments to the ribose sugars, although the individual bases have different sizes and shapes. Due to this, only slight distortions are introduced into the helical structure. (c) Structure of a portion of a single strand of DNA. It is a polymeric chain, consisting of a sugar-phosphate backbone with bases attached to the sugar residues. In a single chain, the structure places no restrictions on the sequence in which the bases can occur.

1.3.1.2 Affinity vs. Specificity in DNA interactions

For all the applications of the DNA base-pairing, maximizing the affinity and the specificity of the Watson-Crick interaction code is particularly important. In those biomolecular interactions and recognitions that are based on shape complementarities, or steric fit between the two counterparts (enzyme-substrate, antigen-antibody, aptamer-small molecule complexes) both high specificity and high affinity can be achieved at the same time. A non-precise steric fit between two surfaces results in significant energetic penalties. The recognition mode and the association between two nucleic acid chains is based on a 1D nucleation-zipping mechanism instead.^[38] Steric fit and nucleation-zipping differ substantially. A strong zip with one irregular or missing link can still be fastened with high affinity leaving out the small mismatched part. The free energy loss

is in this case very small and an increase in affinity does not affect this loss because the mismatched complex will be stabilized comparably. This mechanism results in a gradual decrease of nucleic acid hybridization specificity with increasing the binding affinity. When nucleic-acid-based self-assembly is asked to generate a high degree of structural order, longer oligonucleotides are not always better, and that there might be alternative strategies for increasing the specificity and the affinity at the same time as brilliantly reviewed by Demidov and coworkers,^[38] such as using nucleic acid homologues or introducing additional energy penalties for a mismatched pairing. The first countermeasure usually taken when designing DNA molecules that need to pair efficiently is reducing the symmetrical and repetitive elements of the sequences. Computer software is available for this task, and it is commonly used for the design of DNA nanostructures (see section 2.1.5).

1.3.1.3 DNA in the Nanoscale: Shape, Dynamics, Flexibility

The base sequence of a DNA segment encodes for the average shape of DNA molecules and also for the dynamics of the chain. DNA is continuously morphing into shapes and structures that are slight modifications of the canonical B-form. Atomic Force Microscopy (AFM) micrographs can give hint of the apparently chaotic movements of single DNA molecules. A spread of DNA molecules on a surface (see figure 10) yields AFM images in which no two macromolecules have the same shape and conformation, in spite of having the identical chemical composition. Contrary to the first impression, the dynamics that leads to such a variety of shapes is not random.

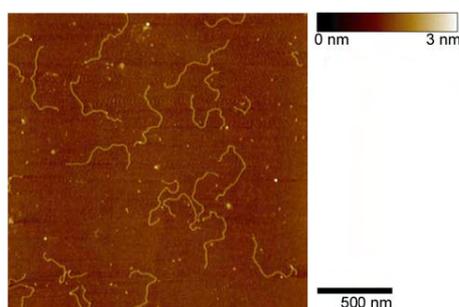


Figure 10 - AFM image of a specimen of double-stranded DNA molecules. All the molecules (besides some obvious fragments) have the same base sequence and length, but they display a different shape due to the intrinsic flexibility of the polymers.

A DNA molecule assumes a shape that is the result of the superimposition of the thermal fluctuations upon its intrinsic, lowest energy, structure, depending from its base sequence.^[39-41] The average structure of dsDNA is a result of the sequence since the chemical inhomogeneities imparted by the different base pairs along the chain give rise to modulations of the orientations of the average planes of the base pairs. These orientations are commonly expressed in terms of the base step orientational parameters: roll, tilt, twist (see figure 11). Considerable effort has gone into defining sets of these parameters corresponding to the lowest energy structures directed by the sequence. Donald Crothers has critically reviewed the recent achievements on the matter.^[42]

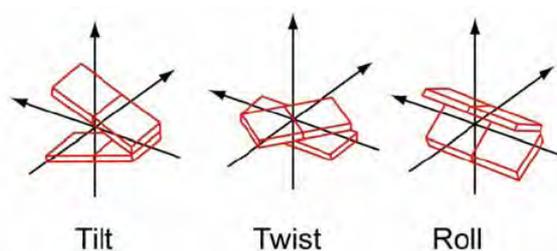


Figure 11 - The most important dinucleotide-step orientational angles: combinations of values of these along the helix will determine the average shape of a DNA molecule.

Variations in the roll or tilt angles give rise to bending of the double helical axis. These local bends might just lead to a local zigzag pattern of the chain axis, which would remain globally straight on a larger scale, unless the deformations are composed in phase with the repeat of the helical winding. In this latter case, they might give rise to extended curvatures that propagate from the Ångström to the nanometer scale.^[43] A notable natural example is the most highly curved DNA segment known in nature: the 211 bp segment from the kinetoplast DNA of the Trypanosomatidae Protozoan *Crithidia fasciculata*. Its sequence (see figure 12) is characterized by a periodical recurrence of tracts of 3 to 6 adenines; spaced by 10 or 11 bp, i.e. the average helical repeat of B-DNA. This distribution of the A-tracts, perfectly phased with the helical winding, makes this short DNA segment have its lowest conformational energy when wrapped in a circle. Electron microscopy evidence of such a large curvature was first presented by Jack Griffith.^[44]

Intrinsic curvatures have been monitored and studied by X-ray crystallography on very short double-stranded oligonucleotides,^[45] on longer DNA molecules by gel

retardation,^[46, 47] circularization kinetic,^[48-50] electron microscopy,^[51, 52] atomic force microscopy,^[53] and have been simulated by molecular dynamics.^[54] Often, these experiments were carried out with peculiar dsDNA constructs, i.e. on constructs with (i) anomalous flexibility sites, like single-stranded stretches,^[53] internal loops due to mismatches,^[47] a single nick,^[55] a double-stranded linker connecting triple-helix tracts,^[56] (ii) segments with very accurate phasing^[53] or unphasing,^[40, 49] of the adenine tracts with the helical periodicity. These experimental efforts usually derived conclusions on the local helical curvature from analyzing global parameters of the whole chain under investigation, like its persistence length, its end-to-end distance, or the cyclization J factor. The effect of defined sequence variants on the curvature was usually inferred from comparisons of the global parameters among sequences. A combinatorial approach has also been proposed.^[49]

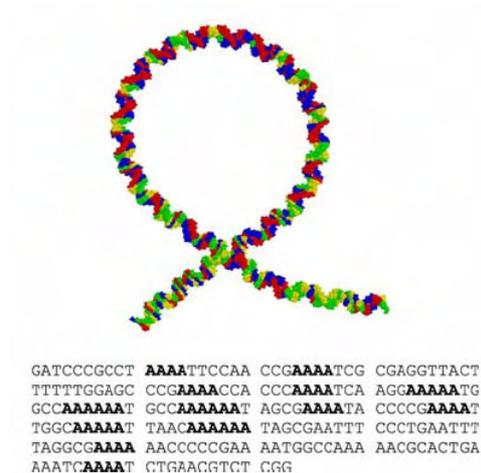


Figure 12 – Below, the base sequence of the curved section of the kinetoplast DNA of *Crithidia fasciculata*, characterized by the phased repetitions of A-tracts. Above, the depiction of a computer model of its lowest energy structure. It can be appreciated that the A-tracts (red portions) are segregated on the face of the molecule looking towards the observer.

The trajectory of the double-helical axis of individual dsDNA chains deposited on a substrate can be traced with the aid of the EM or the AFM. It is possible to set up methods to map the intrinsic curvature along the chain of a natural DNA of any sequence, just by gathering a collection of single-molecule data from high-resolution microscopy imaging. The curvature can be calculated from the local angular chain deflections along a large number of profiles, averaging all the values over the ensemble of profiles.^[57]

The intrinsic DNA curvatures can be theoretically predicted and experimentally evaluated, as described above. Nearest-neighbor methods for the computation of axial curvature of DNA are currently available and the different sets of parameters (results of different methods and optimizations) are in good agreement for their general results.^[42, 58]

Despite many efforts have been spent in this direction, the description of the origin of intrinsic curvatures at the atomic level remains somewhat disputed. Curvature is a long-range superstructural property that is more determined by the way the double helix composes and phases the local bends over different spatial scales, than by their individual values.^[36] A certain flexibility in the sequence is allowed without producing serious changes in the average shape of the molecule.

The base sequence determines not only the global and local average shape of a dsDNA molecule but also its response to the thermal fluctuations. In this way, the base sequence controls the formation of conformers. A conformation, even if poorly populated, can play an important role in a biological function: it can be recognized and selected to switch on processes that the most stable structures might not be able to activate. One of the experimental observables that gives insight on the accessible conformational space of a chain is its local axial flexibility, i.e. the tendency of the long axis of the double helix to deviate both locally and globally from a straight trajectory. While a significant agreement is found in the literature on the origin and determinants of DNA curvature, the issue of DNA flexibility is still under debate. Debate could be originating from the different experimental methods and so the different viewpoints. Evidence gathered at atomic resolution might give information on the atomic determinants of some of the possible chain motions, while evidence collected at a larger size-scale might give more information on the global behavior of a molecule, not being able to interpret fully the high-resolution determinants of the resultant flexibility.^[36]

1.3.1.4 Sequence-dependent DNA-Surface Interactions

Surface-DNA interactions can be sequence-dependent. Macromolecules can exert an exceptional degree of control over nucleation, phase stabilization, assembly, and pattern formation in inorganic structures.^[59-61] Peptides that can show selectivity for binding to

metal and metal oxide surfaces or that can recognize and control the growth of an inorganic semiconductor surface like that of GaAs have been selected.^[62, 63] Much interest is now attracted by the interfacing of inorganic and organic/biological materials and the full understanding of the rules for their interaction. At first sight, a straight DNA chain can rotate around its axis on the surface, so many azimuthal orientations are expected to be equally probable and the chemical interactions with the crystal surface is averaged to a cylindrical symmetry. On the basis of this consideration, it is not expected that DNA should exhibit any sequence preference in its binding to inorganic surfaces, nor that any azimuthal orientation should be preferred. On a more careful analysis, it can become evident that the surface density of charges on the outer surface of DNA can be modulated by the nucleobases hidden in the interior of the double-helix.^[64] On this ground, it is possible that, on interacting with a charged surface, there could be preferred azimuthal orientations of a straight DNA helix.

When a DNA helix is intrinsically curved, the curvature plane defines two faces for a DNA section. The chemical inhomogeneities that yield the curvature act also in such a way that the two resulting faces of DNA will expose sides of the molecule with a different chemical composition. If DNA curvature is induced by regular phasing of Adenine tracts (Atracts) with the helical periodicity, then the two defined faces will vary in the richness of adenines that they could expose to a surface adsorbed on that face. One face will turn out to be rich in adenines and the opposite one will turn out to be rich in thymines. By properly exploiting the internal symmetry of tailor-made DNA molecules, it was possible to show that the propensity in the adsorption of the different faces of highly curved DNA molecules can be significantly different, so that one face (the thymine rich, in our molecules) will adsorb up to 10 times more frequently than the other face on the surface of freshly-cleaved muscovite mica.^[65] This recognition effect seems to be directly dependent on the DNA curvature, itself related to the base sequence. It is reasonable to expect that specific surface-biomolecule recognition processes, such as the ones mentioned above, might play a role in the integration between self-assembled functional DNA nanostructures and microfabricated structures.

1.3.1.5 DNA for Bottom-Up Nanofabrication

In the specific context of nanoscale assembly, several unique features of DNA render it feasible to use it to build geometrically defined structures by self-assembly.

(i) DNA has easily predictable and programmable intra- and intermolecular interactions inherently encoded in its sequence. The most famous of those is the Watson-Crick base pairing leading to the formation of double helices, but other more unusual structures such as Hoogsteen triple helices,^[66] G-quadruplexes,^[67] I-motifs,^[68] curved tracts, and others,^[69] can be encoded in a DNA sequence (see Figure 13).

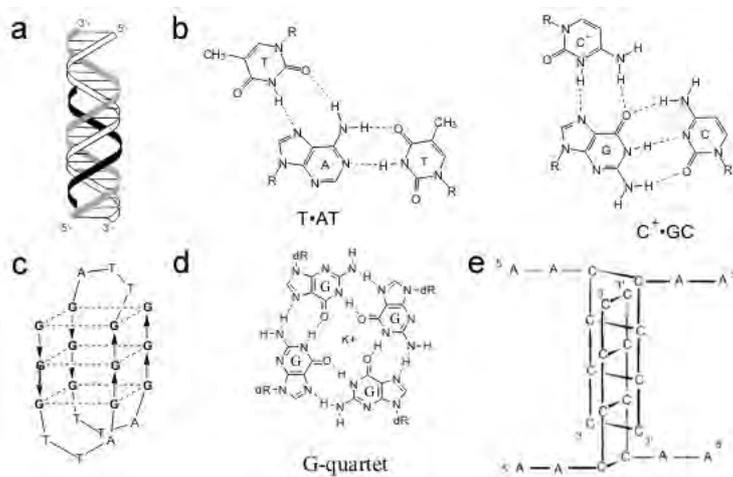


Figure 13 – (a) Schematic representation of a DNA Hoogsteen-type triple helix. The strand shown in black is usually called “triplex-forming strand” and binds to an homopurine strand in a standard duplex DNA via Hoogsteen bonds. Several other types of DNA triple helices exist. (b) Structure of the Hoogsteen bonds mentioned above. The triplet shown on the right needs the protonation of the cytosine in the triplex forming strand, and thus is only stable at acidic pH. (c) Simplified structure of a DNA quadruple helix formed by G-quartets. Other types of G-multiplets exist. (d) Structure of a potassium-ion stabilized G-quartet, leading to the formation of four-stranded structures like the one shown in panel c. (e) One example of another four-stranded structure. This one is called “i-motif” and is stable at acidic pH. The noncanonical base pairing involved is C⁺-C. Several other i-motif strand topologies are reported.

(ii) The hybridization energies of all these modes of interaction, and the geometrical shapes resulting from them are well known, and it is thus easy to predict the structure of a DNA strand in a given environment.

(iii) Just as RNA and proteins, DNA is a linear polymer with an extremely high informational content, but in contrast to what happens with other biomolecules, the informational richness of DNA is not responsible for a huge variety of structures. For example, with the exception of a few peculiar cases, most DNA double helices basically

have the same shape regardless of their sequences; therefore, it is a lot easier to predict the geometry of a DNA helix in a given context, than to predict the geometry of proteins and RNA. Thus, double stranded DNA in particular is easier to use as a building block.

(iv) The inter-molecular interaction energies between DNA molecules in solution are intrinsically encoded in their sequences; the sequences thus dictate which DNA strands will interact, and how. This means that DNA is capable of performing a strictly information-regulated self-assembling process that leads from a collection of DNA strands in solution to a supramolecular complex of known size and shape. This can be exploited to organize the intra- or inter-molecular self-assembly of an arbitrary number of natural or synthetic DNA molecules, and can thus be viewed as a very versatile form of programmable nanofabrication in which the program is defined by the set of the involved sequences.

(v) DNA has access to structures that are extremely varied in stiffness; the double helix for example has a persistence length of 50 nm^[70] (which is extremely rigid when observed at the scale of few nanometers, just as steel is rigid when observed at the scale of a few meters) and higher persistence lengths can be obtained with more unusual structures if needed. At the same time, DNA can include some very flexible parts such as single-stranded DNA. This means that DNA structures can easily balance stiffness and flexibility as needed by the project, and also that the flexible parts can tolerate small strains in the structure and allow imperfect designs to form.

(vi) Certain DNA sequences assume markedly different geometrical structures in response to specific external stimuli.^[71-73] These stimuli could be changes in the environment (for example its pH, ionic strength, temperature) or the insertion of molecules interacting with DNA (for example proteins, intercalators, other DNA molecules). This means that *ad-hoc* designed DNA structures are capable of dynamically assume well-defined structures and can thus work as nanoscale devices and motors.

(vii) DNA can be easily modified with extreme precision and versatility by synthetic chemistry or by taking advantage of the extensive toolbox provided by natural enzymes. A DNA molecule can be decorated with different species, such as metal nanoparticles, proteins, carbon nanotubes or organic dyes, and still retain its self-

assembly abilities,^[31, 74-82] thus providing a straightforward method for organizing an ample library of these nano-sized objects into well-defined structures.

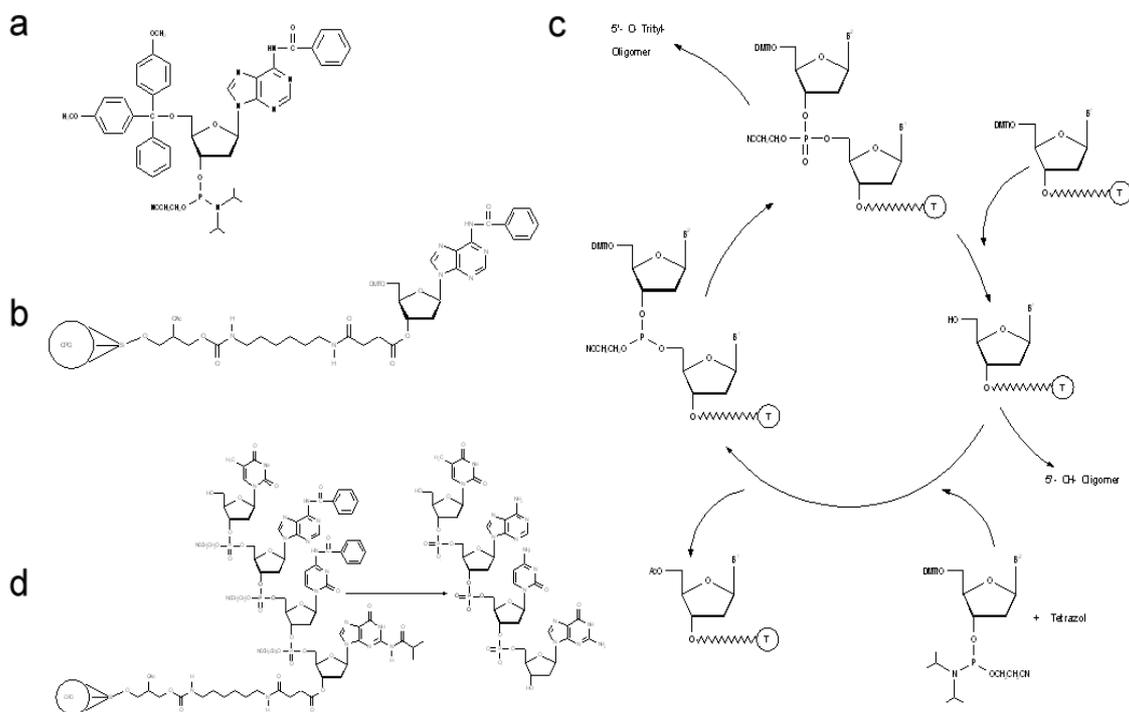


Figure 14 – (a) The most important raw materials for the solid-phase supported oligonucleotide synthesis are the nucleoside phosphoramidites. Here, a deoxy-adenine phosphoramidite is depicted. The reactive phosphoramidite group is located at the 3'-OH, the 5'-OH is blocked with the acid-labile dimethoxytrityl protection group (DMT). To prevent the exocyclic amino-functions of adenosine, guanosine and cytidine from undergoing unwanted side-reactions they have been blocked using acyl protecting groups. (b) To start the synthesis, a solid support (the most common are controlled pore glass and polystyrene) is linked to the first (3') nucleoside of the oligo to be synthesized via a base-labile linker. (c) The chemical synthesis of oligonucleotides is performed from 3' to 5'. During the key coupling step a 3'-phosphoramidite reacts with a free 5'-hydroxy group. Each cycle includes four main steps: (c1) Deblocking. The acid-labile 5'-dimethoxytrityl protecting group is cleaved from the base that is anchored to the solid phase support. A free reactive hydroxy function is obtained. Typical cleavage reagents include di- or tri-chloroacetic acid in dichloromethane. (c2) Coupling. The free 5'-OH group is now able to react with added phosphoramidite. As a result both nucleosides are linked by a phosphite bridge. The phosphoramidite has to be activated first using a weak acid. (c3) Capping. The coupling step has a yield of around 99% in an optimized system. Due to the large number of successive synthetic steps however, it is important to prevent remaining free OH-groups from reacting in the following steps, giving rise to non-specific sequences. Therefore in the capping step all free reactive groups are blocked using acetylation, effectively stopping any further reaction on them. (c4) Oxidation. The internucleotide phosphite group that has been created in the coupling step is oxidized to phosphate using iodine solution. These reactions are repeated until the desired oligonucleotide sequence has been produced. (d) After the required sequence is obtained, the following post-processing steps are required: (d1) Cleavage of the oligo from the solid support. By treating the support-bound oligonucleotide with concentrated ammonia solution, the ester bond between the 3'-OH of the oligo and the solid support is cleaved. (d2) Cleavage of the base protecting groups. To remove the protecting groups from the adenine, guanine and cytidine bases and to release the exocyclic amino function, the ammonia oligonucleotide solution is incubated at 50° C for a few hours.

(viii) Automated phosphoramidite chemistry^[83] offers a practical, economical and straightforward means for obtaining artificial DNA strands of almost any desired sequence in large quantities (see figure 14).

(ix) Compared to most other biomolecules, DNA is physically and chemically very stable, easier to handle and store. Structures built with it will thus be accordingly sturdier.

(x) There are well-established and straightforward methods for DNA purification and structural characterization.

All these features make DNA an exceptional building block that has been used to perform many types of nanoscale assemblies. DNA has been linked to proteins^[84-86] to regulate their activity or to confer to them self-assembly capabilities different from those they would naturally have.^[78] It has been used to direct the course of chemical reactions,^[87] to detect the hybridization of analytes of various nature in solution, to provide an internal standard in force measurements. It has been metallized to form conductive wires.^[88] Self-assembled monolayers functionalized with oligonucleotides have been formed on metals and silica substrates. The assembly of synthetic oligonucleotides was used to model calculations and logic gate operations (see section 2.1.3.3).^[89, 90] The conjugation of DNA with two nanoscale objects in particular was thoroughly investigated in the past decade: carbon nanotubes and metal nanoparticles.

1.3.1.6 DNA and Nanotubes

The remarkable characteristics of carbon nanotubes (CNTs) have elicited an escalating amount of research during the last years, gradually but firmly establishing CNTs in the scientific community opinion as a missing piece potentially capable of bringing novel and diverse technologies into existence, nanoelectronics ahead of the many competitors. Some applications are even ready for the market,^[91] as to keep research focused and lively.

Very recently, the two apparently disparate research fields of CNTs and nucleic acids have crossed paths as several research groups have tried to use DNA to functionalize carbon nanotubes, in an effort to drive their self-assembly into desired structures, or to bestow on them novel properties (such as an increased water solubility,

or the ability to act as a sensor for the presence of other DNA molecules, as needed for diagnostics).^[92-94]

CNTs can be chemically functionalized in many ways, and a number of these methods are viable for their functionalization with biomolecules such as DNA oligonucleotides.^[32, 77, 80, 94] Some of the methods lead to a chemical modification mainly located at the tips of the CNT, while others bring the functional group and the ensuing biomolecule to the sides. The preference in this localization is one of the hottest issues in CNT chemistry, and its precise control is a prerequisite for the production of univocally defined CNT-DNA networks (see figure 15).

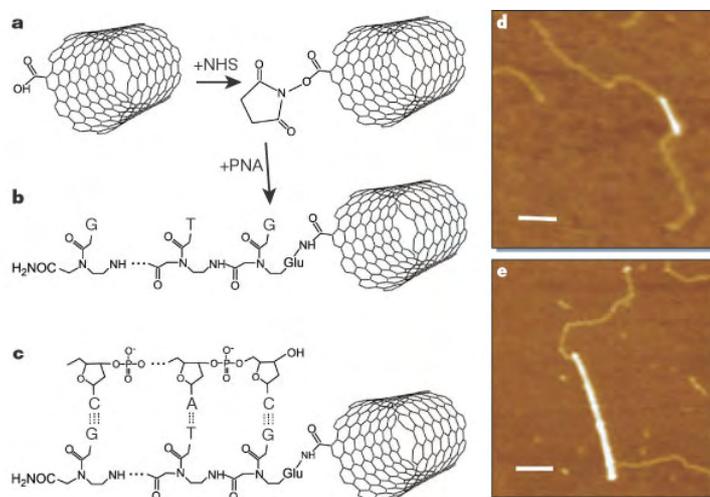


Figure 15 – (a,b) N-hydroxysuccinimide (NHS) esters formed on apically carboxylated, single-walled carbon nanotubes (SWNTs) are displaced by peptide nucleic acid (PNA), forming an amide linkage. (c) A DNA fragment with a single-stranded, sticky end hybridizes to the PNA-SWNT. (d,e) Tapping mode AFM images of DNA-PNA-SWNTs conjugates. SWNTs appear as bright lines; the paler strands represent bound DNA. Both scale bars are 100 nm. [Image taken from ref.77]

Taft and coworkers have recently showed their approach towards selective apical and side functionalizations of CNTs.^[95] They obtain apical functionalization by amide coupling of an amine-modified ss-DNA oligonucleotide with carboxylic groups created by nitric acid etching, and sidewall functionalization by the adsorption on the CNT of an hydrophobic pyrene moiety linked to a different ss-DNA oligonucleotide. While the chemistry they employ might not be original,^[96] they report that performing both apical and side functionalizations simultaneously enhances the selectivity of both, to the point of almost completely eliminating cross-reactivity. This selective functionalization of the tips and the sidewall of the CNT with two different DNA

oligonucleotides (oligos), and their subsequent hybridization with two complementary oligos bound to different sized Au nanoparticles, allowed the controlled placement of two types of nanoparticles on two zones of the same CNT. The oligonucleotides attached to the tips of the CNTs have been shown to be able to hybridize with their complementary oligos either via the attachment of oligo-functionalized gold nanoparticles, which are easy to detect,^[95] or via the attachment of the CNT onto oligo-functionalized electrodes.^[97, 98]

Using the same chemistry, Li and coworkers have recently shown that it is possible to use oligo-functionalized gold nanoparticles as cornerstones of a CNT assembly: several oligo-functionalized CNTs can bind to a central gold nanoparticle opportunely functionalized with thiol-modified oligos complementary to that on the CNTs. While the geometry and the stoichiometry of the assembly is not under control, the reported strategy is a simple way towards the directed assembly of multi-nanotube, multi-material constructions that could prove important in nanoelectronics. Li et al. also show how different types of nanotubes (single-wall and multi-wall CNTs) could be attached simultaneously to the same nanoparticle, although under statistical control only.

The chemistry of oligonucleotide grafting on gold nanoparticles allows however for even more complex structures: nanoparticles functionalized with multiple oligos can be obtained^[99] so that it could be possible to precisely direct the attachment of different CNTs functionalized with different oligos around a single nanoparticle. However, as mentioned above, the number and type of CNTs binding to a nanoparticle employing this strategy is only statistically dictated by nanoparticle size and surface density of functionalizations, and hence also the shape of the resulting structures is not truly under complete control. Furthermore, the presented strategy can lead to the construction of complex, multi-component nanostructures assembled from building blocks with recognized electronic properties, and could thus be regarded as a means to obtain nanoelectronic systems. However, in the resulting structures, the potentially conductive species (CNTs and Au nanoparticles) are de facto inescapably separated by the entire length of the rigid non-conductive DNA duplex formed by the hybridization of the oligos directing the self-assembly. This could perhaps render necessary another

synthetic step, namely an electroless metal deposition, if the end result should be used for electronic measurements.^[35, 100, 101]

Alternative approaches towards assembling and directing CNT with DNA has been presented by Keren and coworkers^[35] and Wooley and coworkers.^[102] In these authors' approaches, the sides of the nanotubes are functionalized or targeted by functional groups attached to the length of the DNA molecule. It is easy to align a double-stranded DNA molecule, if this is considerably long. It is so possible to localize and align CNT using DNA as a template. The localization can also be at a base-sequence specified place, as in Keren's method.^[35]

1.3.1.7 DNA and Metal Nanoparticles

The integration of metal or semiconductor nanoparticles with DNA can match the unique electronic, photonic, and catalytic properties of nanoparticles with the structural and recognition properties of DNA to create novel hybrid nanobiomaterials. Of particular interest for researchers are the use of biomolecule–nanoparticle assemblies for bioanalytical applications and for the fabrication of bioelectronic devices.

The groups of Chad Mirkin at Northwestern University and Paul Alivisatos at the University of California at Berkeley opened the way to the DNA-mediated nanoparticles assembly. Both groups work on the synthesis, characterization and assembly of DNA/AuNPs conjugates. AuNPs present fascinating aspects such as their size-related electronic, magnetic and optical properties and their possible applications to catalysis and biology.^[103]

The first example of assembly of gold nanoparticles in big aggregates with DNA was reported in 1996^[74] and an highly selective, colorimetric polynucleotide detection method based on optical properties of mercaptoalkyl oligonucleotide-modified gold nanoparticle probes was reported in 1997.^[104] Introduction of a single-stranded target oligonucleotide (30 bases) into a solution containing the appropriate probes resulted in the formation of a polymeric network of nanoparticles with a concomitant red-to-purple color change. A temperature increase over the duplex melting temperature causes the disassembly of the structure and the change of color from purple to red. The optical properties of macroscopic DNA-linked Au nanoparticle aggregates can be controlled

through choice of DNA linker length and the differences in the optical properties observed for the DNA-linked aggregates formed with the oligonucleotide linkers of different length are due not only to the interparticle distance but also to aggregate size.^[105] These evidences have important implications for the development of colorimetric detection methods based on gold nanoparticles.

Detection methods that rely on these materials show promise with respect to increased selectivity and sensitivity as compared with many conventional assays that rely on molecular probes.^[106, 107] In the case of target selectivity, the nanoparticles probes can be used to differentiate perfectly complementary targets from those with single-base mismatches, whereas the analogous assays based upon organic fluorophores do not offer such selectivity. The origin of this selectivity derives, in part, from the extraordinarily sharp melting profiles exhibited by duplex DNA structures formed between target strands of DNA and the nanoparticle probes.^[104] The melting properties of DNA linked nanoparticle aggregates are affected by a number of factors, including DNA surface density, nanoparticle size, interparticle distance, and salt concentration^[108] that should be combined in order to obtain the maximum level of selectivity.

Paul Alivisatos and coworkers have showed that nanocrystals modified with ssDNA can be arranged into homodimeric and homotrimeric assemblies^[109] and also in heterodimeric and heterotrimeric “nanocrystal molecules.”^[110] Phosphine stabilized gold nanoparticles can be isolated in an electrophoresis gel and they are stable for several cycles of separation and recovery. Thiol-modified single-stranded oligonucleotides can be incorporated into the protective phosphine shell and can react directly with the gold surface.

Single-stranded DNA/gold conjugates are assembled with different strategies. The first involved the use of two complementary ssDNA/conjugates to form double stranded nanocrystal structures. In the second approach, the DNA/nanoparticles conjugate are later assembled with template strands.^[110] The nanocrystals and the different hybrid conjugates can be characterized with TEM and gel electrophoresis.

A body of experimental investigation of the surface coverage of Au nanoparticles with DNA, the conformation of bound DNA^[111] and the role of nonspecific adsorption on the surface of gold nanoparticles is already available.

Christof Niemeyer and colleagues have reported the preparation and characterization of oligofunctional gold nanoparticle conjugates containing different DNA sequences: the bottom-up assembly of complex biomolecular functionalized nanoparticles is thus possible.^[99] A method based on strand displacement for the reversible sequence-specific switching of DNA/gold nanoparticles aggregation has also been reported by Niemeyer.^[112] This strategy takes advantage of linker oligonucleotide whose sequence is divided in three sections, two complementary to that of the oligonucleotides bound to 23 nm Au nanoparticles, and one which forms a dangling end in the nanoparticles aggregate. Linker introduction into a solution containing the appropriate conjugates results in the formation of a polymeric network of nanoparticles with the concomitant color change. The dangling end can later serve as a nucleation section to promote hybridization with a DNA sequence fully complementary to the linker: the complete pairing of this molecule induces aggregate disassembly.

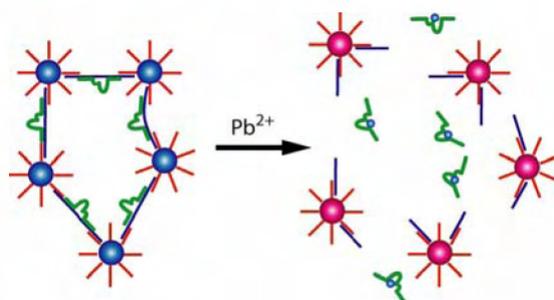


Figure 16 - Scheme of the DNAzyme-mediated disassembly of an assembly made of oligonucleotide-nanoparticle conjugates. The presence of Pb(II) in solution triggers the enzymatic cleavage of the RNA linkers that hold the nanoparticles connected. The solution color shifts from blue/purple to red upon disassembly of the nanoparticles. [Image taken from ref.113]

Starting from the early studies on DNA modified gold nanoparticles several applications have been reported on literature, in particular in the fields of analytical chemistry and biodiagnostic. Recently, numerous applications of modified gold nanoparticles in colorimetric sensors were published. Pb(II) detection can be realized with DNzyme driven disassembly of DNA/gold nanoparticles aggregates.^[113] A DNzyme (called also deoxyribozyme) is a catalytic DNA molecule of a particular sequence and 3-dimensional structure that can carry out specific chemical reactions, often with an efficiency comparable to that of protein enzymes. The DNA enzyme used by Liu and Lu is a RNA cleaving DNzyme activated by the presence of Pb(II) that was

obtained through an in vitro selection process.^[114] In the presence of Pb(II), this DNzyme can cleave a target molecule: if this cleaved molecule served as the linker between Au-oligonucleotides conjugates, then disintegration of the aggregate takes place and a change in color from blue to red can be observed (see figure 16).

Notable diagnostic applications are not lacking. Oligonucleotides have been successfully used as biochemical barcodes to measure the concentration of amyloid- α -derived diffusible ligands, a potential soluble pathogenic marker for Alzheimer's disease.^[115] The key to the bio-barcode assay is the homogeneous isolation of specific antigens by means of a sandwich process involving oligonucleotide-modified Au nanoparticles (NPs with biobarcode) and magnetic microparticles, both functionalized with specific antibodies to the antigen of interest. This system shows an extraordinary increase on sensitivity respect to conventional assay for early disease markers detection thanks to the very effective sequestration of antigen and in particular to the amplification process that occurs as a result of the large number of barcode DNA strands released for each antigen recognition and binding event. A similar bio-barcode assay has been previously employed towards the detection of the prostate-specific antigen reaching a sensitivity six orders of magnitude greater than the conventional ELISA assay for the same target.^[116]

The optical properties of colloidal gold have been recently employed towards the implementation of a "molecular ruler."^[117] When two nanoparticles are brought into proximity, the measured resonance in the wavelength of the plasmon resonance depends on the inter-particle separation. This effect has been applied to dynamics of DNA hybridization on the single-molecule level. Sonnichsen and coworkers have used plasmon coupling to monitor the directed assembly of functionalized 40 nm particle pairs through a 33-nucleotide ssDNA molecule. Liu and coworkers have designed a Au/DNA plasmonic molecular ruler that is able to measure length changes in the DNA molecules anchored to the nanoparticles by measuring shifts in the scattering spectrum of a single nanoparticle.^[118] Endonuclease-mediated shortening of DNA can be measured with base-pair resolution due to a reported wavelength change of 1.23 nm/bp.

1.3.1.8 Complex Self-assembled Architectures made of DNA

Finally, there is one branch of DNA-based nanobiotechnology that tries to expand the knowledge about the structures achievable with the self-assembly of DNA molecules. Since DNA helices are extensively characterized nano-objects with a whole repertoire of interaction modes, DNA sequence design is also a form of “nano-architecture” that organizes matter according to a rational design in the nanoscale. The goal of this line of research is to understand exactly how much is possible to obtain with DNA self-assembly, in terms of structural diversity, complexity, dimensionality and size. This exciting line of research is usually called “Structural DNA Nanotechnology”^[73, 119-125] and is the subject of the next chapter.

2 Structural DNA Nanotechnology

The birth of Structural DNA nanotechnology was in 1982, when professor Nadrian Seeman of New York University had the idea of utilizing DNA to build three-dimensional scaffolds with rationally designed dimensions and shapes.^[126] The intended use of these scaffolds was originally to enclose in their cavities large guest biomolecules, in order to understand whether this could facilitate crystallographic measurements on specimens that would hardly crystallize on their own.^[126] Seeman's first project of such a scaffold was a regular 3D array of DNA strands (see Figure 17) orthogonally arranged to form uniform cubic cavities.^[126] Although this specific design was never implemented experimentally, it was the precursor of a series of experiments that during the successive decade brought structural DNA nanotechnology to the attention of the worldwide scientific community.

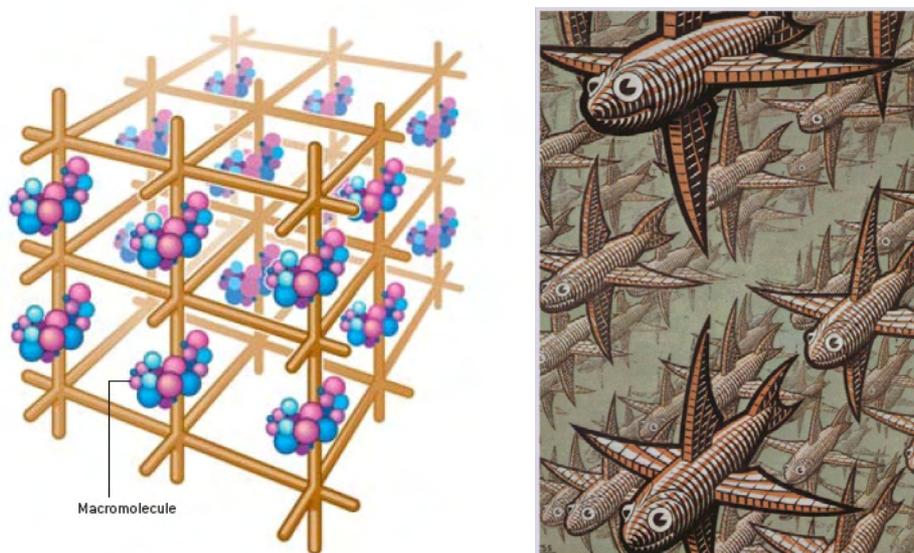


Figure 17 – (Left panel) Schematic representation of the three-dimensional DNA scaffold concept. Synthetic DNA oligonucleotides form six-arm branched junctions that self-assemble into a regular 3D lattice with a defined cavity size. The idea behind this design was to prove whether other biological macromolecules could be entrapped into these cavities and be included into an ordered crystal amenable to standard crystallographic measurements, even if they would not crystallize on their own. Unfortunately this was never even attempted because the formation of these 3D lattices proved extremely difficult. (Right panel) It was this wood engraving (“Depth” by M. C. Escher, 1955) that spurred N. C. Seeman imagination and inspired him to attempt the design and synthesis of the structures shown in the left panel, among many others.

Since its beginnings, the field of structural DNA nanotechnology witnessed an explosion of activity,^[119] and several academic groups around the world nowadays focus their research exclusively on it. In this chapter, I will try to outline the most important concepts and achievements of this fascinating discipline.

2.1 Static Architectures

The main goal of structural DNA nanotechnology is to provide an operatively simple way to control the positioning of matter in the nanoscale through the bottom-up self-assembly of DNA strands. The research on this field tries to find answers to several basic questions, such as:

- (i) What are the shapes obtainable with the self-assembly of DNA?
- (ii) What is the most complex ensemble of DNA molecules that can still self-assemble in solution?
- (iii) What is the largest dimension obtainable with self-assembled DNA structures?
- (iv) How can be DNA nanostructures interfaced with microscopic systems?
- (v) What practical strategies are the best to actually perform the self-assembly in solution?
- (vi) How can we minimize error incorporation in DNA self-assembling systems?
- (vii) How can we introduce the abilities of error-correction and self-healing to DNA self-assembling systems?
- (viii) How can we translate complex algorithms into the self-assembly of DNA molecules?
- (ix) What type of objects can we attach to DNA without losing neither the self-assembly capabilities of DNA nor the function of the attached objects?

Most of these questions have found an answer during the past two decades, but none of these answer has been proven to be the *best possible*. The precise potential of the structural DNA nanotechnology approach is still vastly unexplored.

2.1.1 Basic Element of Assembly I: DNA Junctions

The idea of realizing self-assembling DNA nanoarchitectures of arbitrary shape and complexity is without doubt captivating. However, there is one apparent conundrum right in its premises. How can one build complex structures using only a topologically linear polymer such as DNA?^[127]

“It is often useful to look at [DNA] from the contrary viewpoint expressed by Sherlock Holmes in Silver Blaze, when he remarked to the inspector on the ‘curious incident of the dog in the night-time.’ When the inspector replied, ‘The dog did nothing in the night-time,’ Holmes replied, ‘That was the curious incident.’ A similarly curious feature of DNA is its lack of branches: Insofar as we know, the helix axis of genomic DNA is topologically linear.”

[N. C. Seeman, Annu Rev. Biophys. Biomol. Struct. 1998 (27) 225-248]]

On the other hand, branched DNA molecules do appear as key intermediates in DNA metabolism, principally in the processes of replication, homologous recombination, and repair. An emblematic example is the Holliday junction.^[128-130]

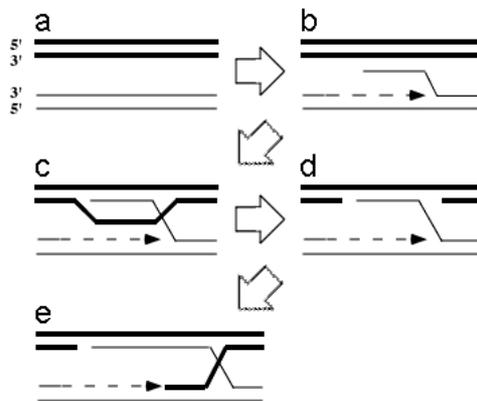


Figure 18 – The role of Holliday Junctions in homologous recombination. (a) Homologous DNA from the recipient (darker lines) and donor (lighter lines). Notice the relative orientation of the termini. (b) A single-strand nick in donor DNA is formed. DNA synthesis causes strand displacement. (c) The displaced ssDNA pairs with the recipient DNA forming a heteroduplex. The unpaired recipient DNA forms a “D-loop”. (d) Nucleases remove the unpaired D-loop DNA. (e) Strand-exchange produces a Holliday Junction (DNA crossover) which, due to the symmetry of the sequences involved, can branch-migrate to extend the length of the heteroduplex.

The Holliday junction (see figure 18) is a recombinational intermediate^[131] that contains four DNA strands arranged into four double-helical domains intersecting at a common branching point. The problem is that the intersection points of naturally

occurring Holliday junctions are flanked by homologous symmetric sequences. Because of this symmetry, the branching point can slide back and forth via branch migration isomerization.^[128]

This positional instability of the intersection rendered for a long time the physical characterization of Holliday junctions elusive. To overcome this problem, Seeman suggested that the branch point could be immobilized by eliminating the symmetry of the flanking sequences.^[129, 130] By using synthetic DNA oligonucleotides with sequences designed to have minimal symmetry, it was possible to obtain artificial, immobile four-way junctions (see figure 19). The advent of synthetic immobile DNA junctions has resulted in the accumulation of much physical data on branched DNA molecules.^[132]

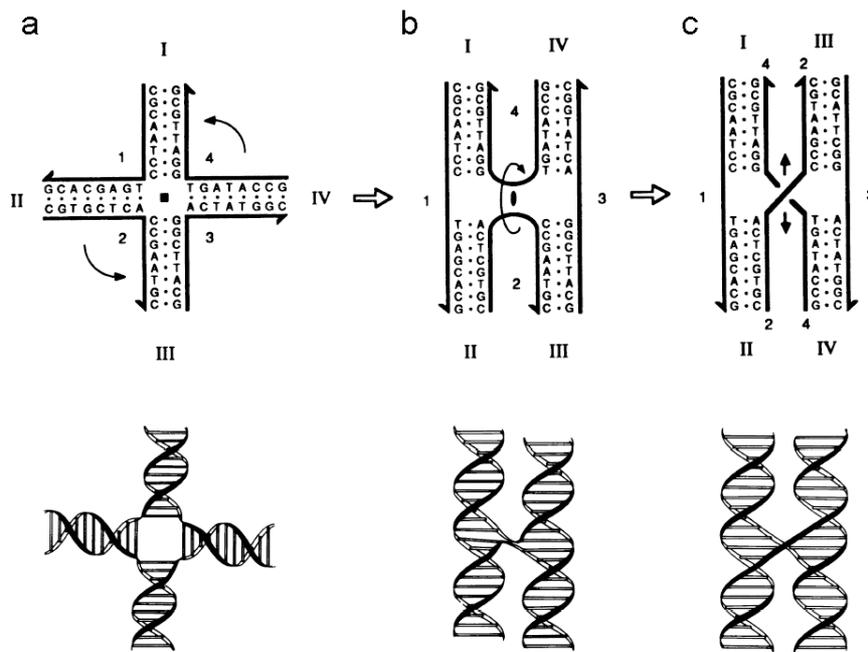


Figure 19 – Sequence and schematic representations of immobile junction or “J1 junction”. (Upper) Schemes showing different structural possibilities for the junction. (Lower) The same structures shown with a 10.5-fold helical representation of the DNA. (a) The structure is shown with a fourfold symmetric backbone in a crossroads structure. This junction is composed of four 16-mers. The strand numbering is indicated by arabic numerals, and the arms are numbered by roman numerals. Arrowheads are on the 3' ends of the strands. The naturally occurring Holliday junction has twofold sequence symmetry about the central branch point; this symmetry allows the structure to resolve to linear duplex DNA via the branch-point migration isomerization process (see figure 18). The J1 junction is stable in solution because this symmetry is eliminated. Backbone symmetry of the representation is indicated by conventional symbols. Curved arrows indicate transformation of the crossroads morphology into b and then c, both of which are twofold symmetric structures. [Image taken from M. E. Churchill et al, PNAS 1988 (85) 4653]

Seeman extended the structural characterization of artificial immobile DNA junctions to many variations of the Holliday Junction, including junctions with between three and six double-helical arms, and bulged versions of these.^[127] Conceptually, the possibility of introducing stable branching points at precise positions and orientations on a rigid linear object allows to build structures of virtually any shape; practically, the knowledge accumulated on the formation and structure of DNA multi-armed junctions allowed to implement this through the self-assembly of synthetic DNA oligonucleotides. The resulting structures are stick figures and networks, in which the edges consist of double-helical DNA and the vertices are the branching points of junctions.^[133, 134]

2.1.2 Basic Element of Assembly II: “Sticky Ends”

Practical considerations limit the potential of the above approach. First of all, since B-DNA is a helical structure, not all possible shapes are immediately available through the self-assembly of ssDNA strands- the assembly of some shapes can entail a topologically impossible winding of strands. Moreover, complex stick networks would need very long synthetic oligonucleotides, which are quite difficult to obtain in sufficient quantity and purity.

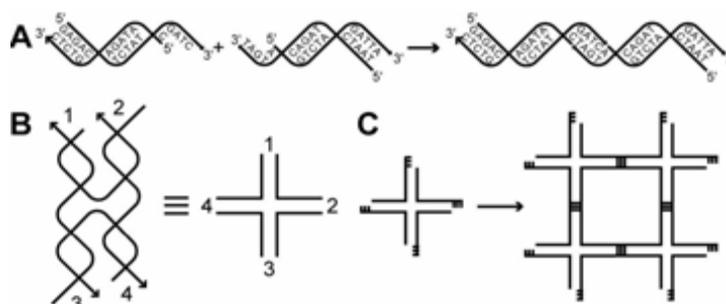


Figure 20 – (A) Cohesion of two double-helical fragments via “sticky-ends,” single-stranded overhangs protruding from a double helical segment. Adjoined sticky ends (with two nicks) can be further stabilized by chemical or enzymatic ligation to make the phosphodiester backbones continuous. (B) Schematic illustration of a stable Holliday Junction. Four oligonucleotides form a branched junction in which the sequences flanking the center are not symmetrical. The branching point cannot thus migrate as it does in the naturally occurring symmetric junctions. The angles and lengths in the scheme do not correspond to the actual structure. (C) Sticky-end cohesion of branched structures. Four-arm junctions bearing sticky-ends on the arms can form superstructures by binding to other similar junctions according to the sequences of the involved sticky ends. The shape depicted on the right, formed by just four junctions is often called ‘DNA parallelogram’ or ‘DNA rhombus motif.’ Further junctions can bind to the quadrilateral and eventually form a continuous lattice.

A way to overcome this is to first build “topologically allowed” sub-units of manageable size and complexity, and then assemble them in a deterministic way. The type of cohesion between sub-units, however, should have two fundamental requisites. It should have a geometrically well-defined structure as the rest of the architecture; and it should be able to form between specific parts of the sub-units through self-assembly.

Once again, DNA can cope with this requirements without any external help. The intermolecular organization between sub-units can be achieved with the sequence-dependant cohesion of single-stranded tracts protrusions (see figure 20). Such protrusions are usually called “cohesive ends” or, more affectionately, “sticky ends”.^[135] Sticky ends cohesion has all the requisites listed in the previous paragraph. Individual sticky ends can be positioned by design in specific places of the sub-units, and guide with complete specificity their intermolecular assembly. Most importantly, the correct pairing of complementary sticky ends results in the formation of a nicked B-DNA helix, without loss of structural continuity between the two linked parts. As reported in literature,^[135] the stacking forces between the nucleotides on either side of the nicks are sufficiently strong that there is no pronounced deformation of the helical structure, despite the interruption of the covalent backbone chain. Since the geometric properties of both the sub-units and the linking parts are known, the geometry of the complete architecture formed through sticky-end cohesion is known.

Sticky-end cohesion has one further feature that can be exploited advantageously. Since the interactions keeping together sticky ends are the same that keep together every other part of the construct, that is, Watson and Crick base pairing interactions, the relative stability of each portion of the structure, including sticky ends, are known. This allows to design structures and sticky ends so that during thermal annealing, some portions of the structure form before others in a known order.^[73] This allows to circumvent topologically-prohibited intermediates during structure formation, but also allows to plan a definite hierarchy of structures with different thermal stability ranges (see section 2.1.4).

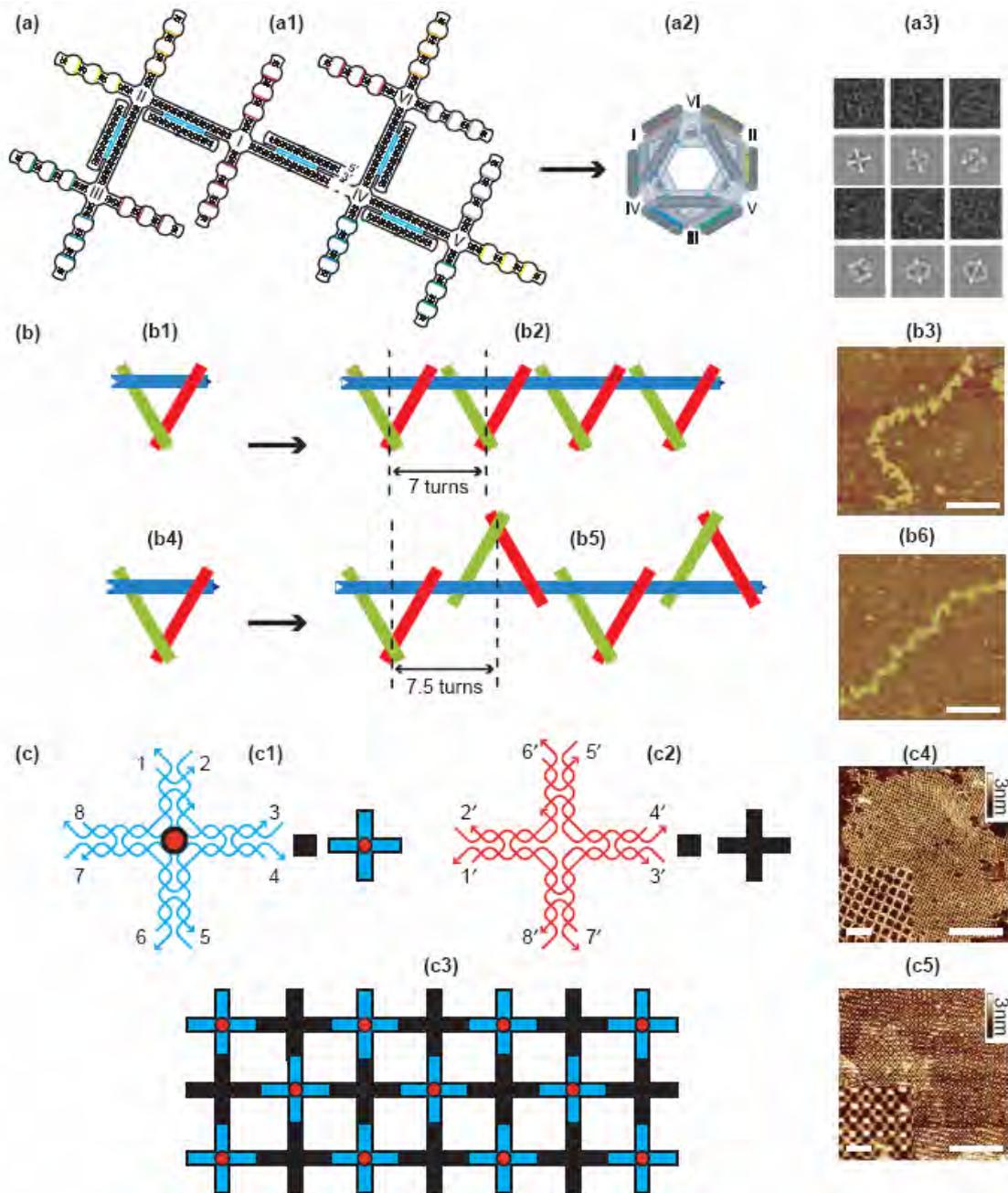


Figure 21 - Examples of DNA nanoarchitectures with different topological dimensionalities. (a) A zero-dimensional array (i.e. an individual DNA octahedral object). (a1) The ‘unfolded’ structure of the Octahedron. It consists of one long (1.7 kilobases) and five short strands. Five of the edges of the octahedron are double crossover (DX) motifs, formed by specific regions of the long strand and the five shorter-strands, here colored in blue. The remaining edges are subsequently formed by the paranemic cohesion of appropriately designed domains. The pairs of complementary domains are labeled with the same colors. Edges are joined by four-arm junctions (labeled with roman numbers). (a2) Representation of the structure resulting from the cohesion of all the domains. The vertices of the octahedron are the four-arm junctions and have the same numeration as in panel (a1). (a3) Cryo-electron microscopy imaging of the octahedron. The first- and the third-row panels are the raw images of individual particles, and the second and fourth are the result of projecting the experimental data onto a three-dimensional model of the folded octahedron, obtained by single-particle reconstruction techniques. (b) One-Dimensional DNA arrays made of triangular tiles. (b1) Schematic representation of a DNA triangle comprising three duplex domains (shown as rods of different colors). Vertices are four-way junctions, where the blue domain is equipped with a pair of complementary sticky ends on opposed termini. Each

domain is exactly seven-helical-turns long. (b2) One-dimensional array generated by the sticky-end cohesion of the triangular tile shown in panel (b1). Because the sticky ends are separated by an integer number of helical turns, each triangular tile is oriented in the same direction. (b3) An atomic force microscope (AFM) image of the array depicted in panel (b2). Size bar is 100 nm. (b4) Triangular tile similar to that represented in panel (b1) but with a 7.5-turn-long blue domain. (b5) One-dimensional array generated by the sticky-end cohesion of the triangular tile shown in panel (b4). Because the periodicity comprises an odd number of half-helical turns, tiles point alternately in opposite directions. (b6) An AFM image of the array depicted in panel (b5). Size bar is 100 nm. (c) Two-dimensional DNA array comprising two alternating tiles. (c1) Strand structure of one of the two tiles comprising the array. The tile is a DNA 'four-by-four' motif, consisting of four four-way junctions flanking a central cavity. Each external double-helical domain ends with a sticky-end, labeled with a different number. This tile incorporates a biotin group, represented here as a red dot, which does not impair the assembling ability of the tile. (c2) Strand structure of the second tile participating in the array, similar to that depicted in panel (c1) but without the biotin. Sticky ends labeled N0 are complementary to those labeled N in panel (c1). (c3) Schematic of the array resulting from the assembly of the tiles depicted in panels (c1) and (c2). The sticky ends in both tiles are cleverly arranged so that the array has a corrugation scheme that minimizes distortions out of the plane. (c4) An AFM image of the array schematized in panel (c3). Size bar is 300 nm in the main figure and 100 nm in the inset. (c5) An AFM image of the same array after functionalization with streptavidin (STV). STV molecules bind to the biotin present on one of the tiles and appear as white dots localized at alternated junctions, as expected from the model shown in panel (c3). Size bar is 300 nm in the main figure and 100 nm in the inset. [Image taken from Zuccheri G, Brucale M, Vinelli A, Samorì B "DNA-Based Artificial Nanostructures", in "Annual Review of Nanoresearch Vol.1" (Chapter 12) 2007, World Scientific Publishing]

2.1.3 Topological and Geometrical Dimensions of the Assemblies

DNA nanoarchitectures are characterized by their geometrical and topological properties. Although the geometrical dimensionality of an object is identified simply, the topological dimensionality of a structure is recognized less easily: in general terms, topology is the study of those properties of a structure that are preserved through its deformation. As a coarse but useful definition, the topological dimensionality of an object is the number of coordinates needed to univocally specify a point of the object. The characteristic way in which the points of an object are connected among themselves does not change with deformation and defines its topological attributes. As an example, a line is topologically one-dimensional, even if it is warped into a sinusoid or a helix and has a geometrical dimensionality higher than one. A flat surface, a curved surface or a tube are all two-dimensional objects topologically, regardless of their geometrical dimensionalities. These topological concepts are useful for describing and classifying DNA nanostructures without ambiguity.

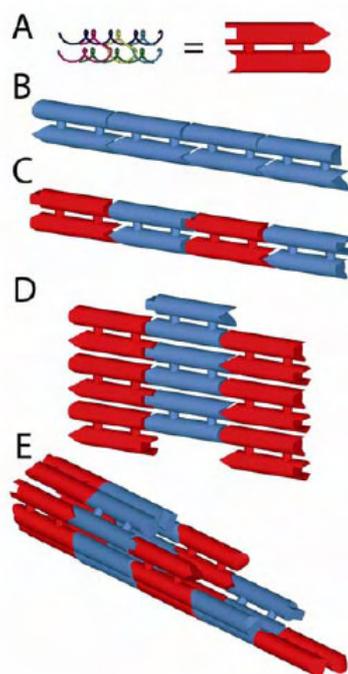


Figure 22 – Scheme of some of the possible spatial arrangements of a DNA tile. A DX tile (A) is schematized as two connected cylinders with two different types of joints at the ends. An example of (B) a 1D construct obtained with one type of tile and (C) a 1D construct obtained with two alternating tiles. An example of (D) a 2D array built of DX tiles. A DNA nanotube (E) can be built by connecting flat tiles at an angle. [Image taken from Zuccheri G, Brucale M, Vinelli A, Samori B “DNA-Based Artificial Nanostructures”, in “Annual Review of Nanoresearch Vol.1” (Chapter 12) 2007, World Scientific Publishing]

The most common approach is to base the description of structures on the way the branching points included in them are connected (see figure 22). This permits discussion of motifs and structures without the complication of considering their mechanical flexibility and, thus, the geometrical deformations to which they might be subjected. As an example, it is possible to discuss the design of DNA nanostructures having the connectivity of polyhedra,^[127] even if the exact geometry of such objects varies during time because of thermal conformational fluctuations.

An alternative approach is to consider every structure as an array of sub-components and take into account, exclusively, how the subunits (often referred to as tiles) constituting the array are connected among themselves regardless of their shape. For example, any array in which each tile is connected to exactly two other tiles is a topologically linear, one-dimensional array, even if the geometrical shape of the array has more than one dimension. An individual object not connected to any other is considered a zero-dimensional array.

One of the advantages of this alternative classification is that the geometrical shapes of the reported DNA nanoarchitectures are extremely varied, whereas there are only a limited number of different topological dimensionalities available to build DNA arrays (see figure 21). Moreover, the topological dimension of the array has an undeniably fundamental role in the design and implementation of DNA nanoarchitectures. This type of categorization is implied in numerous descriptions of DNA nanoarchitectures, even if it is seldom exploited as a means of classification.

2.1.3.1 Zero-Dimensional Topologies: Discrete Objects

The earliest achievements of structural DNA nanotechnology were the construction of discrete objects such as a DNA cube (see figure 23), a DNA truncated octahedron or DNA Borromean rings (three interlinked rings, whereby the linkage between any pair of rings disappears in the absence of the third).^[127, 136] These were not designed to interact with other molecules and form superstructures, thus they were topologically zero-dimensional, as described above.

The mainstay for this type of design was the DNA multiple-branched junction,^[134] which ensured complete sequence-dependent topological control but lacked the geometrical rigidity required to form objects with a reliably programmable shape. Recently however, Mao and co-workers proposed the possibility of implementing the concept of tensile integrity (tensegrity) in DNA nanostructures.^[137] Tensegrity is a property of objects, conferred by components that balance opposed tension and compression loads in a combination that yields a high mechanical resilience. By implementing this concept in the design of DNA nanoarchitectures a rigid zero-dimensional structure can be obtained, even from non-rigid components.

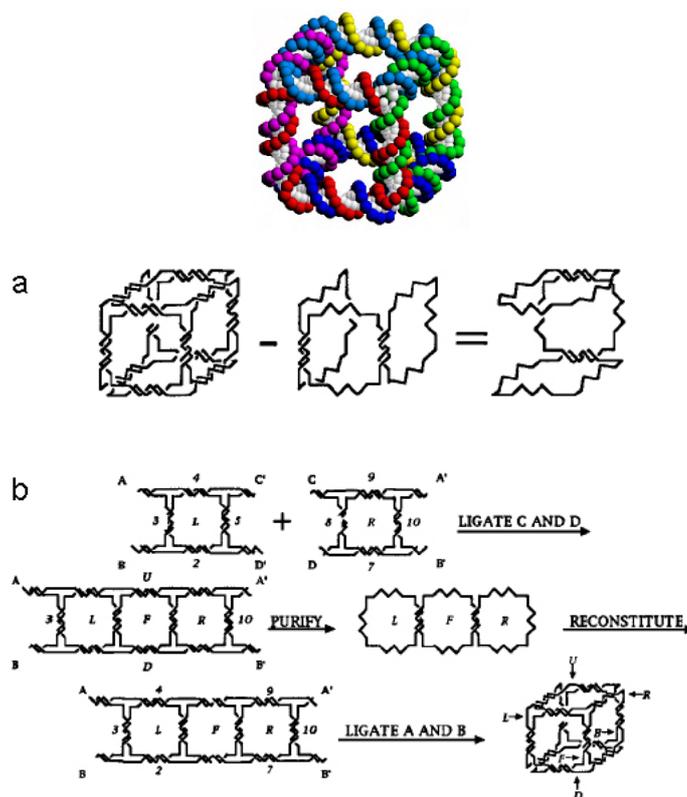


Figure 23 - (a) A DNA cube shown as the sum of two linear triple catenanes. Left, diagram representing the backbone of a DNA molecule whose helix axes have the connectivity of a cube. Middle, left-front-right linear triple catenane that would be removed if the cube were restricted on the left-front and right-front edges. Right, the product top-back-bottom linear triple catenane. (b) Synthesis of a DNA cube. Two ends of two quadrilaterals are ligated to form a belt-like molecule that was denatured and reconstituted to purify it from side-products. The belt-like molecule was then cyclized to form the cube-like molecule. [adapted from images in refs.121 and 127]

Other reported examples of discrete DNA nanostructures include a DNA tetrahedron (see figure 24),^[138, 139] and an octahedron (see figure 25) remarkably formed by a 1.7 kilobase single-stranded DNA that folds into the programmed shape with the aid of five shorter oligonucleotides.^[140]

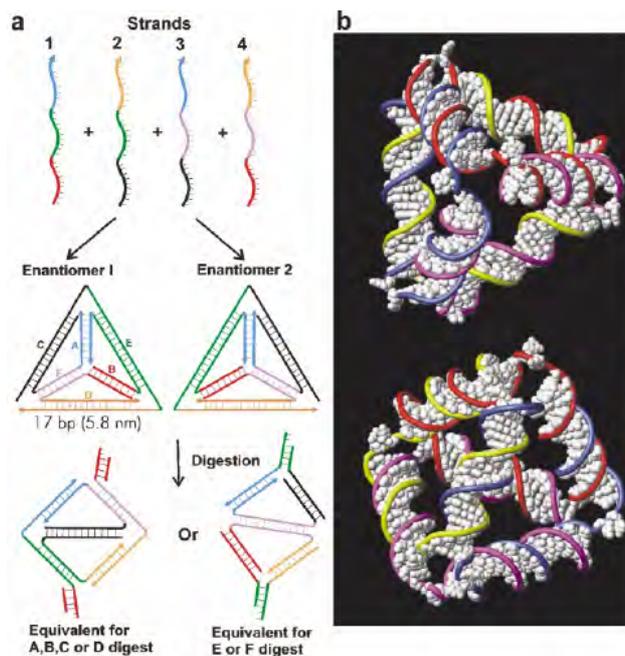


Figure 24 – (a) Synthetic scheme for a DNA tetrahedron, with Schlegel representations of the final product illustrating both possible enantiomers. Each edge is represented in a different colour. The products of edge digestions that cut the central (E, F) or end (A, B, C, D) subsequences are also illustrated. (b) Two different views of a space filling representation of a DNA tetrahedron with 17 bp edges and 2 bp hinges. The backbone of each strand is colored uniquely. [Image taken from ref.138]

2.1.3.2 One-Dimensional Topologies: Linear Arrays

To obtain a superstructure with a geometrically defined shape propagating in just one topological dimension, rigid constituent elements are required. The DNA rhombus^[141] and the DNA double crossover (DX)^[142, 143] were the first suitably rigid DNA motifs developed.

The simplest possible assembly of multiple sub-units is a one-dimensional periodic arrangement in which each unit binds to the successive by means of cohesive sticky-ends, leading to the formation of a long linear superstructure; this was achieved for the first time using the aforementioned structural motifs.^[121, 141, 144] Once this proof-of-concept was obtained, the research focus gradually shifted to the construction of arrays generated by DNA motifs capable of conferring specific features. Recently, a new type of DNA motif, the helical bundle, enabled the construction of one-dimensional, linearly arranged arrays reaching contour lengths of several micrometers.^[145]

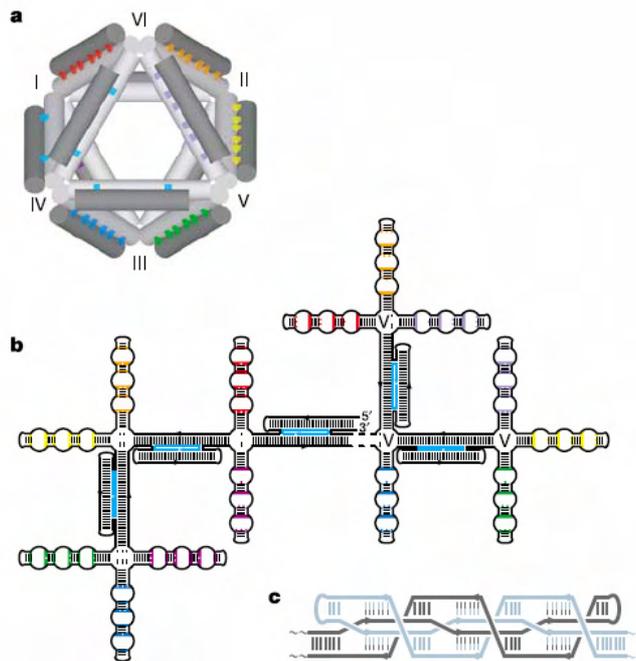


Figure 25 – Design of a DNA octahedron. (a) Three-dimensional structure involving twelve struts (octahedron edges) connected by six flexible joints (octahedron vertices). Five of the struts are DX motifs (cyan) and seven are PX motifs (rainbow colours). The joints are four-way junctions that connect the core-layer double helices of each strut. (b) Secondary structure of the branched-tree folding intermediate. The structure consists of a single heavy chain (black) and five unique light chains (cyan). Like colours indicate half-PX loops whose sequence-specific cross-association generates a strut that serves as an edge of the DNA octahedron. Coloured stripes coincide with strand crossover positions. Folding to the structure in the upper left is complete when all seven PX struts have formed. (c) Schematic of a PX strut. [Image taken from ref.140]

The simple option of having just one self-complementary constituent unit is not the only one suitable for building one-dimensional linear arrays. It is also possible to design a set of two or more sub-units with different complementarities, whereby the order of the successive components in the array is a consequence of the set of instructions represented by the complementary sub-units. This also permits us to go beyond the limits of a periodic succession of components and obtain a completely aperiodic, algorithmic self-assembly. The algorithmic self-assembly of DNA can be used not only for a molecular fabrication task but also as a physical model of computations.^[23, 24, 89, 146, 147] This approach led to the rapid convergence of DNA computing^[148] with structural DNA nanotechnology: both fields aim for full control over the assembly of individual units over disparate length scales in an attempt to create structures that can be programmed, bottom-up, from individual molecules.

2.1.3.3 Two-Dimensional Topologies: 2D Lattices

The structural units (e.g. branched junctions and sticky ends) and the self-assembly mechanisms described for one-dimensional structures are the same as those used to build two-dimensional arrays. A large variety of motifs were obtained from those structural units and used to design DNA tiles that are capable of self-assembling into two-dimensional periodic lattices (see figure 26), including the double- and triple-crossover (TX),^[28, 142, 149] the parallelogram or rhombus motif,^[141] the four-by-four structure,^[78] the three- and six-helix bundles^[145, 150, 151] and the DX triangle.^[152]

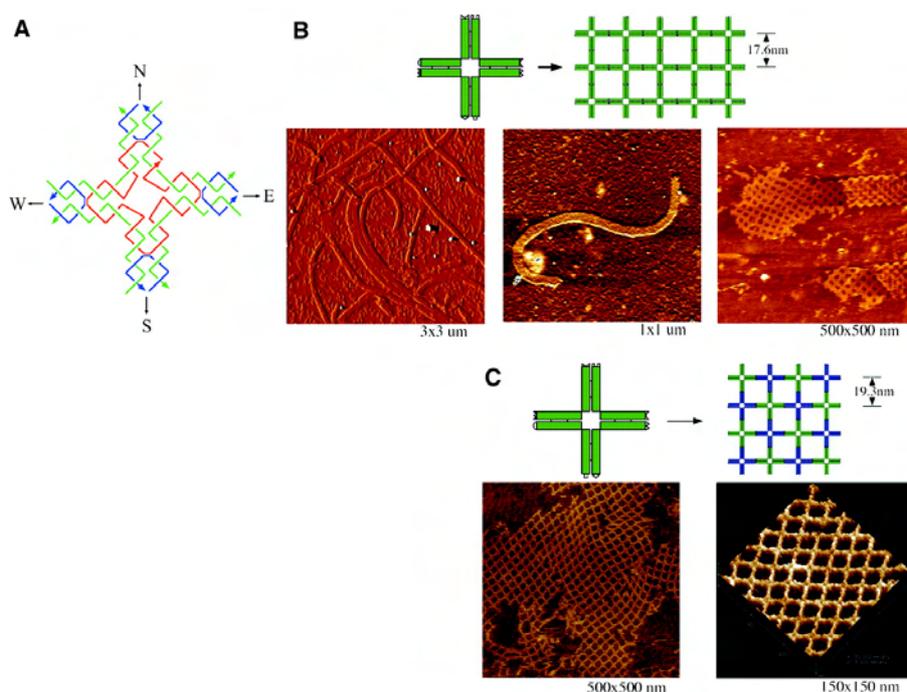


Figure 26 – An example of a periodic 2D DNA lattice. In this case, lattices are formed using a “4x4” DNA tile. (A) The 4x4 tile strand structure. (B) Self-assembly of nanoribbons. (Upper left) Double-helical domains are illustrated as rectangles, and paired rectangles represent four-arm junctions. Complementary sticky ends are shown as matching geometric shapes. (Upper right) Designed structure of self-assembled lattice. (Bottom) AFM images of the nanoribbons. The left panel shows an amplitude-mode image and the right two panels are AFM images in height mode. (C) Self-assembly of 2D nanogrids with corrugated design. (Upper left) The component tile is drawn similar to that in Fig. 1B; positions of sticky ends are changed. The tiles have two surfaces; one faces out of the plane and the other faces into the plane. Here the surface facing out of the plane is indicated in green; the other side (when visible) will be colored blue. (Upper right) Corrugated self-assembly. (Bottom) AFM images of the 2D lattices (nanogrids) formed from the corrugated design. The right panel is a surface plot of a magnified region from the left panel. [Image taken from ref.78]

Comparing the lists of motifs used to obtain one- and two-dimensional arrays, it is easy to notice that a few motifs are present in both, for example the DNA bundles and the DNA parallelogram. This is because the overall shape and properties of the assembled array is not dictated solely by the constituent motif but, more importantly, by the way each element binds to the neighboring ones. Thus, the shape of the array is ultimately dependent on the exact position, orientation and length of the cohesive sticky ends on the tiles.

Once the hurdle of designing the set of DNA tiles is overcome, two-dimensional lattices are usually better behaved than one-dimensional linear structures. One reason for this is that each tile in a regular two-dimensional lattice is linked to at least three other adjoining tiles instead of only two, as in linear arrays: this implies that the persistence of the whole structure does not depend, crucially, on the simultaneous stability of each tile–tile interaction. In linear arrays, the rupture of even one interaction between two adjacent tiles results in the breakdown of the whole structure into smaller parts. Conversely, in two-dimensional lattices, the rupture of one tile–tile bond is not sufficient to impair the overall structural integrity, and the removal of an internal tile is unlikely because it entails the simultaneous rupture of all the interactions with the neighboring tiles.

Two-dimensional lattices of DNA, in which the tiles include chemically modified oligonucleotides,^[31, 74, 76-79, 81, 82, 153] can be used to obtain programmable nanosized patterns of functional chemical moieties on surfaces (see figure 27). Through such chemical functionalizations these information-containing scaffolds can then direct the localization of further functional moieties on the nanoscale.^[31, 36, 74] For this reason, two-dimensional DNA arrays have several potential applications in molecular electronics, sensors and smart materials.

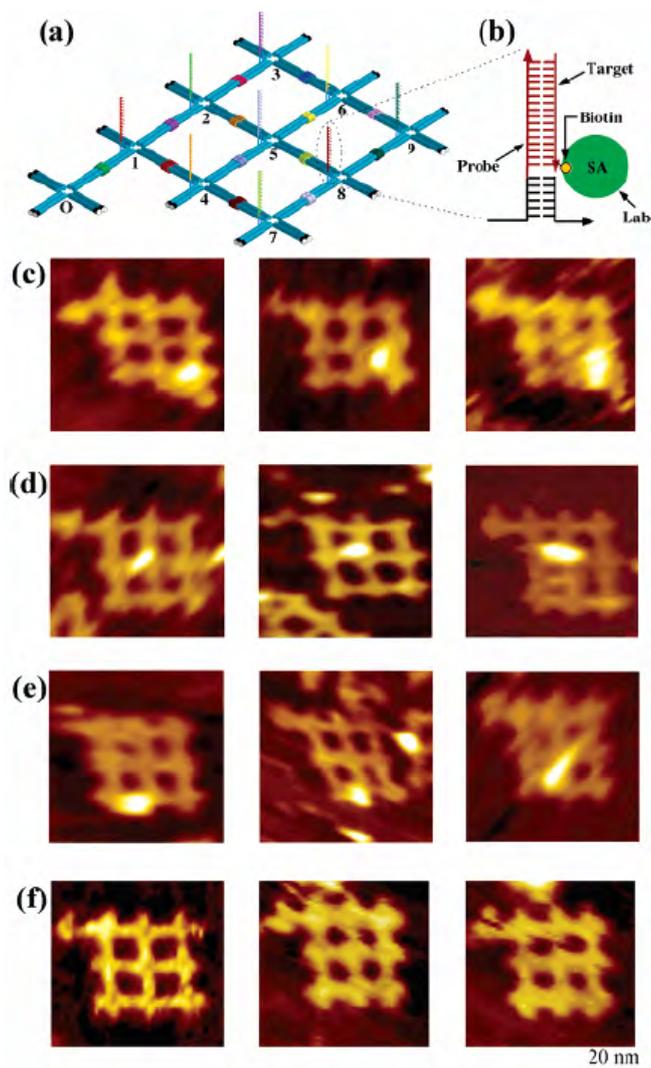


Figure 27 – An example of DNA 2D lattices (finite size structures in this case) bearing chemical moieties that are then used to deterministically immobilize single molecules on it. (a) An asymmetric array formed by ten 4x4 tiles. Its lack of symmetry elements allows to univocally identify each tile from an AFM image. Numbers 0-9 label each tile in the array. (b) Hybridization of the probe strand with the biotinylated target strand is labeled by streptavidin binding and detected by AFM as a bright spot at the probe position. (c) AFM images with expected signals for hybridization at tile 9. (d) AFM images with expected signals for hybridization at tile 5. (e) AFM images with expected signals for hybridization at tile 8. (f) The results of a control in which the arrays were exposed to biotinylated targets that were not complementary to any of the probes are shown as magnified images. [Image taken from K. Lund et al, J Am Chem Soc 2005 (127) 17606-7]

The increase of topological dimensionality from one- to two-dimensional arrays implies an increase in the complexity of the constituent network of tiles. This means that it is possible to code more complex algorithms into the self-assembly of two-dimensional lattices than into their one-dimensional counterparts. It has been theorized^[24] that a two-dimensional self-assembly of DNA tiles can physically model a

set of abstract “Wang tiles”.^[154] Two-dimensional Wang tiles (see figure 28) can be defined as equal sized squares with a color on each of its four sides. When laid down to tile a plane, each square must be placed so that adjoining sides have the same color; thus, each different tile set has its peculiar way of tiling the plane. All mathematical algorithms can be translated into a set of Wang tiles. A single Wang tile binding to neighboring tiles according to its color coding can be represented by a DNA tile spontaneously binding to other tiles as a result of the complementarity of its sticky ends. Accordingly, all algorithms can be encoded in a set of DNA Wang tiles. The result of the algorithm is the assembly of the tile set into a two-dimensional DNA ‘algorithmic crystal’.^[121, 155] However, any defect or mismatch occurring during the assembly impairs the output result; therefore, a considerable effort has been directed to design error-proof, self-assembling sets of tiles that minimize the occurrence of errors.^[146, 156]

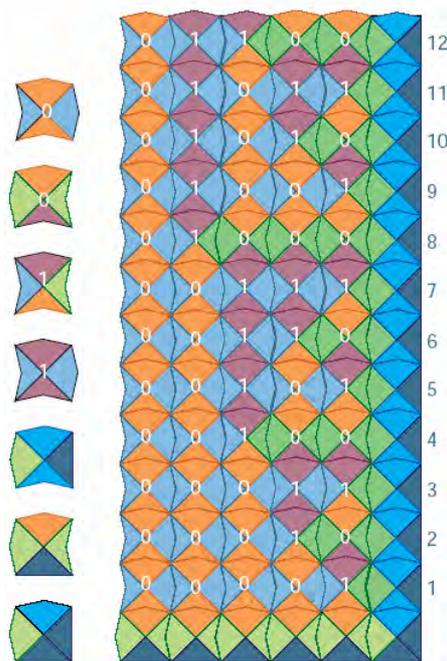


Figure 28 – An example of a set of “Wang tiles” (see below) is shown on the left. Each tile can only bind to another one if their adjoining edges share the same color. The result of the assembly of this set is shown on the right. It has been demonstrated that it is possible to perform Turing-universal computation with the algorithmic self-assembly of a set of two-dimensional Wang tiles. These tiles can be physically built with DNA, and the edge-joining properties in this case are the complementarities between sticky ends of DNA fragments. As an example, the right panel shows how it is possible to implement a simple algorithm with DNA tiles: the algorithm of counting. DNA tiles are represented by squares with coloured edges that are protruded or indented. Seven component tiles are shown on the left: three border tiles on the bottom and four tiles with the values 0 or 1. The array illustrates binary counting from 1 (bottom row) to 12 (top row). Assembly is assumed to proceed by forming the reverse L-shaped border first, followed by binding the tiles that fit into the sites containing two (but not one) edges. Thus, the border determines the 1 tile in its bend, then that 1 tile and the horizontal-border tile on its left determine the 0 tile that fits, while the 1 tile and the vertical-border tile above it determine the (different) 0 tile that fits. [Image taken from ref.121]

The degree of perfection of self-assembled DNA nanostructures depends on the emergence of defects or mismatches during the assembly, possibly impairing structure and function of the construct. DNA computing is particularly sensitive to any imperfection in the assembly and devising countermeasures to this inconvenience is an active field of research.^[146]

One approach for the minimization of growth errors is the optimization of the physical conditions under which the growth occurs. Mismatches occur when a DNA tile binds to the growing array in an energetically sub-optimal configuration. Under equilibrium regime, the energy landscape is thoroughly sampled over the course of time, and specific interactions are preferred on the basis of their thermodynamic stabilities: this process is controlled by association and dissociation rates. Arbitrarily low error-rates can be achieved by appropriate control of these contrasting rates, although this necessarily occurs at the cost of a significant slow down: decreasing error rates by a factor of 10 entails slowing down the self-assembly process by a factor of 100.^[26]

Erik Winfree and co-workers have proposed two main strategies to decrease the error rate in the formation of two-dimensional assemblies without slowing down the process. The first relies upon the concept of ‘proofreading tile sets’.^[26] These sets are designed so that if a mismatched tile is incorporated into the assembly there is no way to continue growth without making an additional error. The result is that, when a mismatch error occurs, the assembly process effectively stalls, giving time to the mismatched tile to detach from the assembly and be replaced by a correct tile.

The second strategy is aimed at the correction of spurious nucleation errors.^[157] This can be achieved if assembly originating from a seed tile (seeded growth) proceeds quickly, whereas those originating from a non-seed tile (unseeded growth) stops because their propagation is highly improbable. The main concept behind such a tile set is a predetermined sequence in which the tiles bind to the assembly during seeded growth. The tiles bind in this sequence because the correct growth continually provides the most favored binding site for the next tile, whereas out-of-sequence growth must continuously advance through energetically unfavorable structures (see figure 29 for the ‘zigzag’ implementation of this concept).

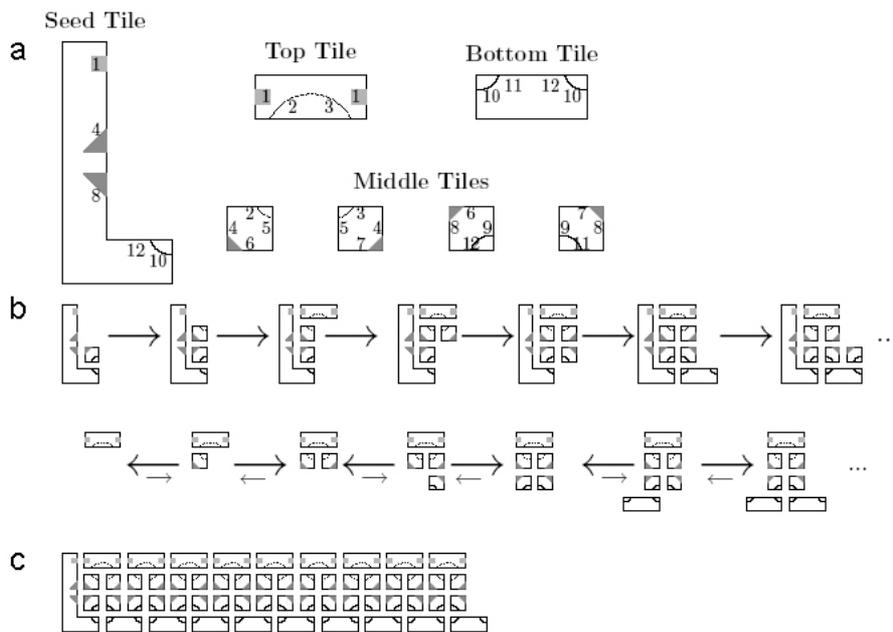


Figure 29 – (a) Schematic representation of a ‘zigzag’ self-assembling tile set. Each square, rectangle, and L shape represents a single tile. Excluding the seed tile, tiles are given unique bonds that determine where they fit in the assembly: each label has exactly one match on another tile. All correctly-matched bonds have the same strength. The geometric patterns shown on each tile identify them in subsequent figures. (b) Upper panel: seeded growth of the zigzag tile set. Lower panel: a possible series of steps by which the tiles could spuriously nucleate. The ‘correct’ assembly always provides the most energetically favorable structures. [Image taken from ref.157]

Although two-dimensional algorithmic self-assembly offers new capabilities for both computation and nano-construction compared with its one-dimensional counterpart,^[23, 156] the design is more challenging and was only successfully implemented, experimentally, recently.^[155]

If a periodic array is based on a tile that is not perfectly planar, or if the tiles are connected so that they are not coplanar, the deformations in the propagating structure will accumulate to an extent that might ultimately become a limiting factor in determining the extension of the lattice. Many published atomic-force microscopy images of two-dimensional lattices show ribbon-like structures rather than indefinitely wide periodic arrays. It has been proposed^[158] that this is because of the tendency of such lattices to form curved surfaces that eventually reach a tube-like shape, inhibiting further growth of the lattice. The ribbons seen in the images could be the result of discrete tubes forming in solution and then unrolling themselves during deposition on a flat surface, like that of mica. Recently, this feature has been rationalized and

intentionally included in the design of sets of tiles that self-assemble in DNA hollow nanotubes (see figure 30).^[145, 150, 151, 158-160]

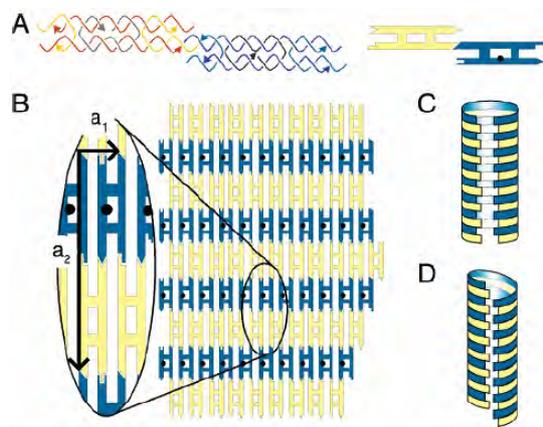


Figure 30 - Self-assembly of DNA tiles into sheets and tubes. (A) Structure of the double crossover tiles: arrowheads mark the 3' end of each oligonucleotide. The 6 nt single-stranded sticky ends on the first tile are complementary to those on the second tile; complementary shapes on the schematic representations of the tiles indicate complementary sticky ends. A 5' biotin label is present on the second tile, represented by a black dot. (B) The tiles tessellate to form extended two-dimensional arrays. The sheets of DNA tiles thus formed can fold forming either alternating rings (C) or nested helices (D) composed by the two tiles. [Image taken from ref.159]

2.1.3.4 Three-Dimensional Topologies: 3D Crystals

The original goal for building DNA nanoscale arrays came from Nadrian Seeman: his original idea was to use a regular DNA three-dimensional lattice as a guide to facilitate macromolecular crystallization.^[126] Unfortunately, the construction of such a lattice by means of the concepts developed for one- and two-dimensional structures proved to be an elusive task for decades.

Recently, however, Paukstelis et al reported the structure of a continuous DNA three-dimensional lattice held together by non-Watson–Crick interactions.^[161] The same design could be used to generate structures with varying lattice dimensions and, thus, in the future serve as the prototype of a DNA molecular sieve (see figure 31) or scaffold for the ordered inclusion of a variety of ‘guest’ molecules.

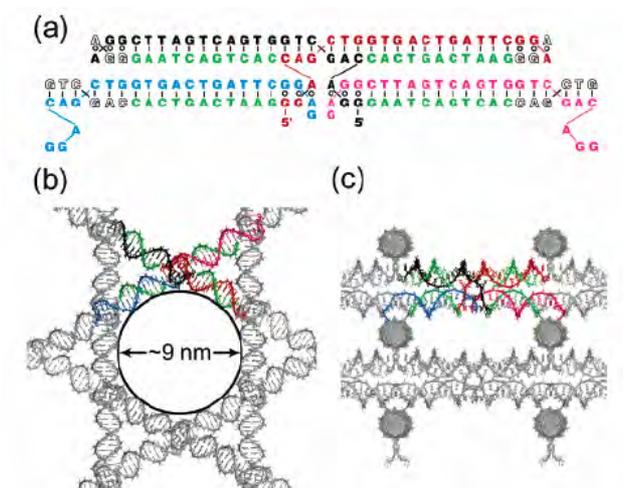


Figure 31 – The 3D DNA lattice discussed in reference 161. The secondary structure of the expanded DNA lattice is shown in (a). Four assembly strands are shown in different colors, and the spacer strands in green. The resulting model of the lattice structure as seen down the 6-fold axis is shown in (b) and orthogonal to the 6-fold in (c). Assembly and spacer strands are colored as in (a). [Image taken from ref.161]

Research is still going in the direction of trying to extend some of the concepts devised for 2D arrays in producing 3D arrays, and some exploratory results^[162] (and macroscopic crystals of DNA nanostructures) are coming out of the lab of Ned Seeman and his co-workers (personal communication).

2.1.4. Hierarchical Assembly

Nature can form large and complex functional aggregates from elementary building blocks that are often orders of magnitude smaller. However, the huge gap between the basic components and the complete assembly is seldom crossed in just one leap. Subunits combine into higher-order constructs that will, in turn, serve as the basic components of a next higher-order assembly until the final level of architecture and functionality is reached. In the cell, for example, nucleic acids are synthesized and later organized into progressively higher-order structures up to the chromatin and chromosome level. Such organization is made possible by several different hierarchical interactions among DNA and several classes of proteins.

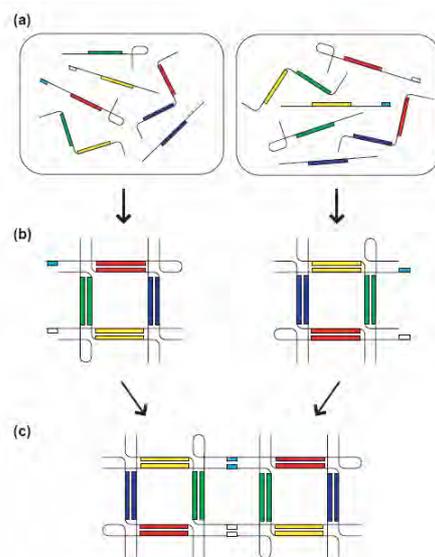


Figure 32 - A scheme for the possible stepwise assembly of DNA parallelograms arrays. Three hierarchical levels are depicted here: (a) single-stranded constituent oligonucleotides; (b) individual parallelograms; and (c) parallelogram-based superstructure. The colored stripes represent sequences, whereby stripes of matching color on different oligonucleotides are complementary sequences. Rounded rectangles represent physically separated vessels, in which the components are assembled during each step. Notice how the one-pot mixing of the oligonucleotides would not have resulted exclusively in the desired assembly. More generally, a system could be constituted by N different structural hierarchic levels (named $L_{(1)}, \dots, L_{(N)}$). The structures in a given level $L_{(n)}$ are completely stable (i.e. functionally acting as indivisible ‘monomers’) in a temperature range that is higher than that of the structures of level $L_{(n+1)}$ and not overlapping it: $L_{(n)}$ structures are effectively polymers, comprising several $L_{(n-1)}$ monomeric structures. The assembly of the final $L_{(N)}$ structure is obtained by the mixing of the $L_{(n)}$ monomers, individually synthesized in separate vessels, and the successive decrease in temperature to the $L_{(n+1)}$ stability region until the final hierarchic level is assembled. [Image taken from ref.73]

It is easy to recognize the levels of hierarchy (see figure 32) inherently present in the design of most of the DNA nano-assemblies described above. The lowest level is represented by the most basic components – the synthetic oligonucleotides fundamental to the formation of the desired structure. The next, and in some cases the last, hierarchical level is that of an individual supra-molecular object. These objects usually constitute oligonucleotides kept together by the pairing of relatively long stretches of bases (typically more than 10 nucleotides long). A further level of hierarchy, if present, is the combination of the objects into a larger superstructure. Typically, the constituent sub-structures are reciprocally bound by shorter cohesive sticky ends, in the range of 4–7 nucleotides-long,^[135] or other types of reversible interactions. These interactions do not disrupt the lower, pre-existing, levels; on the contrary, they need this structural integrity to be able to settle. Successive hierarchical levels are reached when discrete aggregates, formed at the previous level, merge into a new structure or when the

obtained structures are decorated with functional elements of a different chemical nature.

The preparation of even the most complex DNA structures can, in principle, be performed in a single step. This can be done by mixing all the constituent oligonucleotides at high temperature and then slowly cooling the mix in a near-equilibrium regime to maximize the number of interactions, thus converging through spontaneous self-assembly into the planned structure. However, it is probable that there might be an inherent complexity threshold in the successful self-assembly of a 'one-pot' mix of many oligonucleotides. The ruggedness of the energy landscape governing such assembly increases with the number of oligonucleotides involved.

Since the number of possible interactions increases exponentially with the number of oligonucleotides, so the time required for a blueprint-perfect assembly should increase accordingly. During this time, as the temperature decreases, the system has to sample all the possible interactions in search of the most stable ones at each temperature; soon, this becomes an impractically long time. Consequently, during the assembly of a large number of oligonucleotides, kinetic trapping events will occur that will inevitably drive the formation of a collection of unwanted structures if the assembly is conducted in a manageable time span.

To reach a level of complexity unattainable by simple one-pot methods, a stepwise procedure can be adopted. This strategy must be based on the design of structures with different, non-overlapping ranges of thermal stability corresponding to different hierarchic levels. Each substructure within the same level is assembled in separate vessels then brought into contact with the others and finally cooled to a temperature that stabilizes the higher level assembly.

During each step, only the sequences actually directing the assembly at that level (e.g. the appropriate sticky ends) are relevant to the process, whereas the sequences already included in the lower level (higher stability) structures are not disassembled. The result is that structures of the same hierarchic level can have almost identical sequences (with the exception of the sticky ends) because they are assembled separately and then never disassembled. A few examples of this multi-pot strategy have been already implemented.^[163, 164]

The practical limit of this approach has not been assessed but it could reasonably be reached when the level of structures gets big enough not to diffuse efficiently and find partner constructs belonging to the same hierarchy level, or when the number of levels is so high that it is almost impossible to obtain perfect segregation of thermal stability ranges between contiguous hierarchic levels.

2.1.5 Sequence Design

The key step in the construction of all these DNA architectures is the design of the sequences of the constituent oligonucleotides. The success or failure of the self-assembly is intrinsically determined by the interplay between the affinity and specificity of those sequences.^[38] Additional characteristics that could be required are the inclusion or exclusion of sub-sequences of biological or biochemical relevance (e.g. promoters, restriction sites and deoxyribozymes).^[165] The design problems of these sequences are not exclusive to structural DNA nanotechnology but also arise in other fields such as probe selection for DNA microarrays or primer design for PCR. However, the complexity of the self-assembly required to form even the most simple DNA architectures is usually much higher than that required by other such applications. The design of more than just a handful of sequences meeting the desired criteria makes the use of computer programs indispensable.^[166]

The pipeline for the design of a DNA nanostructure begins with the definition of the number, length and mutual connections of all the component oligonucleotides, keeping the intended topology of the assembly clear in mind. Then the base sequences are chosen, obeying a set of criteria most commonly based on the minimization of sequence symmetry^[126] and energy.^[129] Sometimes, however, sequence symmetry can be intentionally included and exploited advantageously.^[167] The ongoing development also tries to take into account the kinetic features of the energy landscape of the assembly in order to avoid trapping the assembling structures into unwanted stable by-products that would be alternative to the target structure.^[168]

It is easy to foresee the need, in the near future, for design tools that could also implement the concepts of hierarchical assembly by permitting the calculation of different stability regions for different hierarchy levels within one superstructure.

2.1.6 Assembly Strategies for Aperiodic Structures

The ability to build nanoscale scaffolds of arbitrary shapes with a command over their nanometer-resolution structure will probably be a key technological advance towards the construction of nanoscale circuitry, data storage units, sensor arrays or artificial molecular factories.^[119, 121, 169] Aperiodically patterned, self-assembled 2D or 3D DNA lattices are regarded as extremely promising candidates for playing the role of such nanoscale scaffolds. Three main strategies have been proposed and implemented in the past years for the creation of DNA aperiodic lattices: algorithmic self-assembly, stepwise assembly and directed nucleation assembly.^[170]

As discussed in section 2.1.3.3 for 2D-lattices, the algorithmic assembly of discrete DNA tiles is potentially a very simple and powerful approach for obtaining arbitrarily complex aperiodic structures.^[171] A set of DNA tiles with the ability of self-assemble through sticky-end mediated, information-driven cohesion (incorporating the algorithm of the assembly) can be designed to obtain potentially any shape.^[24, 28]

However, practical considerations hinder somewhat the seemingly unlimited potential of algorithmic assemblies of DNA tiles. Any complex shape would need a large number of individually synthesized DNA tiles.^[155, 164, 172] The result of their assembly would then be dictated not only by the sequences of their sticky ends, but also by the subtle interplay of assembly growth kinetics, mismatched pairings incidence, nucleation energies, concentration of the individual components and the temperature at which the assembly is performed.^[26, 146] Even prior to that, individual DNA tiles must be flawlessly formed by exact stoichiometric amounts of component strands to avoid competing coupling reactions of incomplete tiles. As mentioned above, due to these practical limitations successful experimental implementations of the algorithmic self-assembly of complex 2D structures was only recently reported.^[155, 163]

As also discussed in section 2.1.3.3, the complications arising from having multiple simultaneous processes competing during the self-assembly of the desired structure can be avoided by performing a stepwise, sequential assembly.^[172] This approach entails the formation of subsets of the structure which are then brought together in a stepwise fashion, thus making it possible to remove any excess of

unreacted species after each step. Another potential advantage of this approach in the context of building aperiodic DNA lattices is that the number of different tiles required for building a given structure is lower than that needed following the aforementioned algorithmic assembly strategy. This is because the final positions of the individual tiles in the assembly is not

dictated solely by their sticky-ends, but also by the order in which they are assembled. Identical tiles could thus be incorporated in the assembly at different positions. The drawback of this approach is that it needs extensive external input from the operator and its overall attractiveness is thus slightly lower than that of unmediated self-assembly.

The most successful strategy employed so far toward the assembly of complex aperiodic structures is that of the directed nucleation assembly. Proposed by LaBean, Winfree and Reif in 1999,^[173] this approach has been used for striking implementations in the following years.^[140, 170, 174, 175] The method entails the use of a long ssDNA template strand (see figure 33) encoding the pattern information of the complete structure. Several shorter strands are then designed to assemble at specified positions on the template strand, folding into the desired shape and completing the formation of the structure. Although this approach is apparently similar to the algorithmic assembly of individual tiles, its practical advantages are manifold.

Since the structure can only form around the template strand, its dimensions are defined with extreme precision and their persistence does not rely on the simultaneous stability of a series of tile-to-tile connective reversible interactions. Moreover, the result of the assembly is less prone to deviations from the desired shape caused by error incorporation. This is because most of the tile-to-tile connections present in an algorithmic array are substituted in a templated array by the irreversible, covalent connection present throughout the template strand backbone.

As proposed by Hao Yan et al in 2003,^[170] the directed nucleation approach is applicable to the formation of 2D planar aperiodic patterns. Recently, Paul Rothemund experimentally implemented this design^[174] and brought it a step further in its potential by demonstrating the formation of “DNA origami” shapes. In the DNA origami approach, a long ssDNA template is folded with the help of many short “staple” strands to produce an arbitrary planar shape. The shape is constituted by parallel double helices lying side-to-side in the same plane being connected by strand exchange crossovers

formed by the shorter strands. The long template strand traverses the shape from side to side in a raster-filling path, thus participating in each helix.

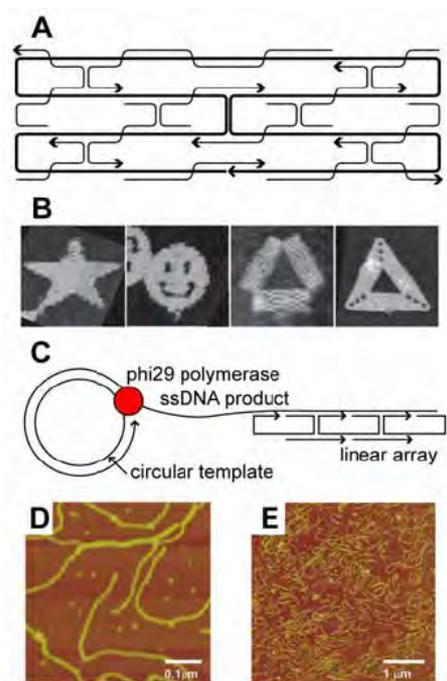


Figure 33 – Examples of the use of a long template in DNA nanoarchitectures. Schematic representation (A) of the design strategy of DNA origami. A long ssDNA template with a naturally occurring sequence is folded so that it fills completely the desired 2D figure (in this case simply a rectangle). Several short “staple” strands pin the long strand in place. Each staple strand binds to different domains of the template strands forming multiple crossovers. The spacing between successive crossovers is typically 1.5 full helical turns, so that they lie on the same plane forming a 180° angle. Examples (B) of shapes obtained by Paul Rothemund with the DNA origami design. All the AFM images are 165 nm x 165 nm. Schematic representation (C) of the strategy employed by Lubrich et al. for the production of long 1D arrays. Rolling Circle Amplification (RCA) of a circularized synthetic oligonucleotide yields a several kilobases long ssDNA template with a repetitive periodic sequence. Shorter synthetic oligonucleotides then assemble on the template yielding the desired periodic array. AFM images (D and E) of the linear arrays obtained with the strategy outlined above. [Image taken from Zuccheri G, Brucale M, Vinelli A, Samori B “DNA-Based Artificial Nanostructures”, in “Annual Review of Nanoresearch Vol.1” (Chapter 12) 2007, World Scientific Publishing]

Very interestingly, in Rothemund’s experiments the template strand is not an ad-hoc synthesized single-stranded polynucleotide, but the naturally occurring 7249-nt single-stranded genomic sequence of M13mp18 virus. This simple fact implies the shift of two of the main paradigms in the field of structural DNA nanotechnology: sequence symmetry minimization and exact stoichiometry determination. All previous DNA structural nanotechnology designs relied on the careful optimization of all the strand sequences involved in the assembly, typically (but not always, as exemplified by Mao and coworkers)^[167, 176, 177] investing particular care in the minimization of sequence

symmetry;^[178, 179] this was done in order to avoid errors in the pairing of the component strands. Furthermore, structures were always formed by as exact stoichiometric ratios of strands as possible, to maximize complete formation. In the origami approach, no sequence optimization is performed and exact stoichiometry determination is unnecessary.

Although long-range correlations are present in the sequences of naturally occurring DNA,^[180] at the scale involved in structural DNA nanotechnology a viral ssDNA strand can be regarded as having an essentially random sequence, and therefore more secondary structure and short-range sequence symmetry than an optimized synthetic strand. Nevertheless, perfectly assembled DNA origami do form with substantial yield, typically around 70 % in Rothemund's experiments. Since the short staple strands are not designed to bind to each other, their relative stoichiometry does not need to be precisely controlled. Most importantly, since complete structures can only form on template strands, it is possible to use a vast stoichiometric excess of staple strands (100-fold). This means that any undesired template-template intra- or inter-molecular interaction is very unlikely to remain stable in the assembly conditions. Likewise, imperfect staple-template interactions should be unlikely because the strand invasion equilibria at work during the assembly will favor the complete, correct pairings.

An additional advantage of the origami approach over the other strategies is that a less than ideal purity of the short staple strands does not impair the overall result, since these are used in large excess. This makes the time- and resources-consuming step of synthetic ODNs purification unnecessary. Taken together, these advantages make it possible to use the origami approach to build DNA aperiodic structures that are significantly larger and more complex than previously possible. Due to the low cost of unpurified short synthetic oligonucleotides, and the possibility to amplify the template strand with enzymatic methods, the origami approach also permits to conveniently increase the production scale of the constructs.

A particularly interesting technique demonstrated by Rothemund is using the staple strands to extrude short dsDNA hairpins out of the origami plane at specific points.^[174] This allows to emboss arbitrary shapes on the structure (see figure 34) with a theoretical resolution limit of 5.4 nm and 6 nm along the two orthogonal axes.

Substituting the dsDNA hairpins with sticky ends for directing the assembly of DNA-modified metal nanoparticles, nanotubes, fluorophores or proteins seems like a natural extension of this approach and will definitely represent a major step forward toward the bottom-up construction of functional nanoscale devices.



Figure 34 - Patterning DNA origami. (a) Model for a pattern representing DNA, rendered using hairpins on a rectangle. (b) AFM image of the actual DNA structure. One DNA turn in the depicted origami image (~ 100 nm) is 30 times bigger the size of an actual DNA turn (~ 3.6 nm) and the helix appears continuous when rectangles stack together appropriately. Letters are 30nm high, only 6 times larger than those written in 1990 positioning single atoms using STM; 50 billion copies rather than 1 were formed. (c,d) Model and AFM image, respectively, for a hexagonal pattern that highlights the nearly hexagonal pixel lattice used also in the previous figures. [image taken from ref.174]

Although blunt-end helix-stacking interactions are always a factor concurring to the behavior of DNA assemblies,^[164, 181] the large tile constructs obtained with the directed nucleation and origami approaches are especially prone to this type of edge-to-edge, blunt-end cohesion, due to the number of their constituent double-helices. Due to this, blunt-end cohesion can be exploited to design and obtain even larger assemblies (tens of megaDaltons for Rothemund's examples)^[174] or periodic lattices of shapes.

Shih and coworkers demonstrated how the long template strategy can be used to obtain a three-dimensional hollow octahedron.^[140] In their example, a 1669-nt ssDNA strand folds into a regular octahedron with an outer diameter of 22 nm, upon hybridization with five 40-nt helper strands. The long template strand was constructed by a polymerase chain reaction (PCR)-based assembly, using synthetic oligonucleotides.^[182] Their methodology is plausibly extensible to other shapes and is amenable to large-scale production since the bulk of the 3D structures is constituted by the template strand that can be conveniently amplified by polymerases once it is assembled from shorter synthetic oligonucleotides.

2.1.7 Enhancing the Structural Rigidity of the Nanostructures

Since the aim of structural DNA nanotechnology is to provide a facile bottom-up access to the accurate spatial arrangement of matter in the nanoscale, its success relies on the development of structurally well-defined DNA motifs.

Apparently, the most straightforward strategy to obtain the highest geometrical control over assemblies of DNA motifs is to remove any potential flexibility from the structures. Much effort has been invested in designing rigid motifs, including double crossovers,^[28] triple crossovers,^[149] “double-double” crossovers,^[183] 8- and 12- helix tiles,^[181] which have been polymerized to form ordered 2D arrays.

In the last years however, new motifs were reported^[78, 167, 184] which have the capability of forming large, ordered 2D arrays even if their structure should be relatively flexible. This prompted He and Mao to investigate the role of tile rigidity in the formation of large, well-ordered planar 2D arrays.^[185] The gist of their reasoning is that if the motif is too flexible, it will not preserve any direction of propagation in the plane and no large 2D arrays will form. On the other hand, if the motif is too rigid, any unpredicted distortion of the individual tile could induce a stress on the tile-to-tile connections that could also prevent the formation of large 2D arrays. So a subtle balance of flexibility and stress (and hence rigidity) in DNA nanostructures appears to be an important factor in determining the success of 2D planar assemblies, even if presently there is no design tool to predict how to attain the ideal balance. To prove this assertion, He and Mao synthesized a group of almost identical tiles, but having a range of different flexibilities. They found that only the middle motifs in the series could form 2D arrays, while neither the most flexible nor the most rigid did.^[185]

Following an orthogonal approach, Seeman and coworkers built a set of new motifs that can assemble into 2D arrays by means of adjoining pairs of sticky ends, and compared the results to these obtained with similar motifs with isolated sticky ends.^[162] Just as two joining sticky ends form a nicked double helix junction, double sticky end interactions result in the formation of a nicked DX structure, which appears to be sturdier and thus less susceptible to errors in twist between succeeding tiles. This means that lattices structures unattainable with single sticky-end cohesion are instead accessible with double cohesion tiles designs.

Taken together, these results seem to indicate that the best tiles for obtaining large, ordered DNA 2D lattices are those with rigid, strong cohesion points kept together by a moderately flexible core.

2.1.8 Using Sequence Symmetry as a Tool

Since its outset, structural DNA nanotechnology designs relied on the minimization of sequence symmetry in the component strands.^[179] The reasoning is that each section of the construct ought to have an unique sequence that should ideally pair only with its intended targets, hence promoting the formation of the desired structure. The more sequence symmetry is present in the involved strands, the more the chances of intra- or inter-molecular incorrect pairings during the self-assembly. Recently, Chengde Mao and coworkers started to investigate the possibility of using sequence symmetry as an asset rather than as a limitation.

The gist of the idea is to identify the symmetry elements of the desired structure backbone and use them to reduce the complexity of the sequences involved in the assembly. In Mao's first proof of concept, the structure under examination is a four-by-four DNA motif, formed by nine different strands, roughly having the shape of a four-arms cross with a fourfold rotational symmetry. Two strands are needed to form each arm, and one strand joins them together at the center. Since the nine strands have nine unique sequences, when the DNA bases are taken into account, the structure has in fact no symmetry elements. The same backbone shape, however, can be obtained with just three different sequences if the arms are designed to be identical (thus reusing the same two strands four times around the central strand).

It is important to note that sequence symmetry minimization is still employed in the design of each individual arm, but the arm is present in the structure more than once. Generalizing these concepts, the introduction of a degree of sequence symmetry in the design of DNA nanoarchitectures could bring on several advantages: (i) the sequence space needed for a given structure backbone is minimized, and thus sequence design is simplified; (ii) the number of different strands is reduced; and (iii) the unpredictable distortions present in most DNA tiles design self-cancel.^[167]

Liu et al tried to bring the consequences of the reasoning presented above to the limit. Employing sequence symmetry in the design, they were able to build large 2D crystalline arrays using only four different strands,^[177] and even micrometers-long DNA hollow nanotubes formed by just one short DNA oligonucleotide.^[176]

2.2 Dynamic Architectures

Another goal of structural DNA nanotechnology is to be able to dynamically alter the nanoscale structure of artificial self-assembled constructs. Objects that can modify their structure in response to a specific event or in accord with a predetermined program have the potential of functional utility; dynamic DNA nanostructures are thus promising nanoscale components for building functional nanodevices and machines. Proposed applications of DNA molecular devices include such materials science, nanoelectronics, biosensors, chemical synthesis^[87] and molecular therapy.

Many DNA self-assembled constructs capable of controlled motion have been reported in the last years.^[71, 72, 186] Several different principles are employed to obtain controlled movement in these constructs (see figure 35), including DNA conformational transitions,^[187-190] strand-displacement equilibria,^[191-198] and protein binding.^[199] Motion is either triggered by specific, externally provided stimuli, or is autonomous.^[195, 200, 201]

For example, under proper conditions, a GC-alternating sequence can undergo the B-Z transition that implies a reversal of handedness of the double-helix. Seeman and coworkers have employed this controlled and induced rotation to change the distance of separation of objects in space.^[187] When these objects were fluorescent dyes, the motion could be easily followed studying their photophysical properties. Many other examples of controllable dynamic objects made of DNA are nowadays found in the scientific literature.

One key feature of strand-displacement-based devices (prototyped by Yurke et al.)^[193] is that the sequence-dependent mechanisms for the actuation of the device allow simultaneous control of different devices in the same environment, or different parts of the same device. A number of ingenious variations of this idea have been reported, including DNA “walkers” that can move along a track step-by-step,^[196, 197] a 2D DNA

periodic lattice with tunable cavities,^[192] and a pair of DNA “molecular gears” that can revolve against each other.^[198]

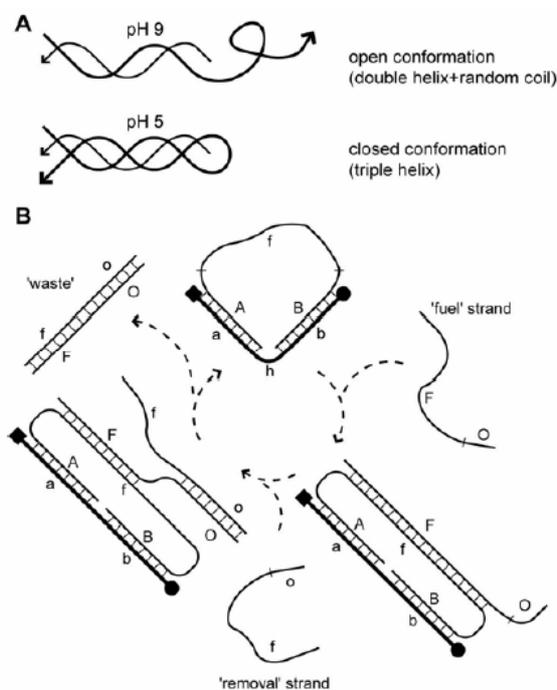


Figure 35 – Examples of strategies utilized for obtaining controlled motion of DNA nanostructures. A DNA conformational change (A) triggered by external input. One long and one short oligonucleotide combine to form a structure which comprises a double helical domain and a single-stranded overhang. At acidic pH, the overhang folds on the duplex by forming a cytosine motif triple helix. Cycling the pH between acidic and basic values causes repeated movement of the structure. Sequence-dependant addressing (B) of DNA “molecular tweezers.” Letters represent sequence tracts, the uppercase pairs with the lowercase. Addition of the “fuel” strand causes the tweezers to open. Its removal by a nucleation-zipping mechanism restores the system to its original open state. Each cycle produces a “waste” duplex product formed by the joining of one fuel and one removal strand. [Image taken from Zuccheri G, Brucalè M, Vinelli A, Samori B “DNA-Based Artificial Nanostructures”, in “Annual Review of Nanoresearch Vol.1” (Chapter 12) 2007, World Scientific Publishing]

In addition to “clocked” devices that respond directly to a change in their environment, autonomously running DNA nanomachines have been built. Yin et al. reported a walking device autonomously processing along a track by means of a sequence of enzymatic reactions.^[200] Mao and coworkers built a set of molecular tweezers containing an RNA-cleaving DNAzyme that can continuously cycle between its open and closed shape when the appropriate substrate is accessible. The cycling can be stopped by putting a “brake” on the device.^[201]

Another strategy to obtain free-running devices is to use “catalytic” strands that can unlock kinetically stable loops for the invasion of other strands.^[195, 202, 203] Two

further examples of autonomous devices was reported by Simmel and colleagues. In the first case, a pH-sensitive molecular construct is located in a reactor in which a non-equilibrium oscillatory chemical reaction is taking place.^[204, 205] The reaction produces pH variations that cause the device to cycle between its conformations. The second strategy entails the use of transcription of a designed sequence to control a set of molecular DNA tweezers.^[206-208]

Hybridization chain reaction^[209] can be exploited to obtain a ‘triggered’ self-assembly of static DNA nanostructures. Briefly, it is possible to store potential energy in locked conformations such as loops that are kinetically inaccessible at room temperature in the laboratory time scale, and then unlock them via a chain reaction of successive hybridizations initiated by a ‘catalyst strand.’ During the reaction, all the loops are opened one at a time and incorporated into a growing nanostructure.

Shu et al. recently reported an interesting use of DNA nanomotors on surfaces. They immobilized an ensemble of DNA motors on microfabricated silicon cantilevers and verified that the forces exerted by cycling the nanodevices can induce a surface stress capable of bending the cantilevers, demonstrating the translation of biochemical energy into micromechanical work.

In order to build complex functional nanodevices constituted by several different components working in concert, immobilizing all the different portions on some sort of common substrate or surface will probably be convenient. An immediate problem posed by this is that even if the functioning of a component is verified in solution, it might not work with the same efficiency, or even at all, when it is immobilized in a sterically hindered context such as on a surface. The functioning of DNA nanodevices immobilized on a nanoparticle surface was verified by Li and Tan.^[194]

Another issue is to understand the details of how the synthetic devices work at single molecule level. Most published data regarding DNA nanodevices motion refer to populations of devices, and almost nothing is known about the behavior of the individual in the population. Single-molecule studies of artificial DNA nanomotors are beginning to appear in the literature. Combining the two approaches, the single-molecule behavior of DNA nanodevices immobilized on a glass surface was recently verified^[210] by our group.

Chapter I Bibliography

- [1] G. M. Whitesides, *Vol. 1*, **2005**, pp. 172.
- [2] C. P. Poole, F. J. Owens, *Introduction to Nanotechnology*, John Wiley and Sons Inc, Hoboken, N.J., **2003**.
- [3] R. F. Service, *Science* **2002**, 298, 2322.
- [4] N. Taniguchi, in *Proc. Intl. Conf. Prod. Eng. Tokyo, Vol. Part II*, Japan Society of Precision Engineering, Tokyo, **1974**.
- [5] E. K. Drexler, *Engines of Creation: The Coming Era of Nanotechnology*, Bantam Doubleday Dell, **1986**.
- [6] E. K. Drexler, *Nanosystems: Molecular Machinery, Manufacturing, and Computation*, John Wiley & Sons, **1992**.
- [7] R. F. Service, *Science* **2004**, 304, 1732.
- [8] G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Applied Physics Letters* **1982**, 40, 178.
- [9] G. Binnig, C. F. Quate, C. Gerber, *Physical Review Letters* **1986**, 56, 930.
- [10] D. M. Czajkowsky, H. Iwamoto, Z. F. Shao, *Journal of Electron Microscopy* **2000**, 49, 395.
- [11] Z. F. Shao, J. Mou, D. M. Czajkowsky, J. Yang, J. Y. Yuan, *Advances in Physics* **1996**, 45, 1.
- [12] J. M. Lehn, *Macromolecular Symposia* **2001**, 174, 5.
- [13] J. M. Lehn, *Science* **2002**, 295, 2400.
- [14] J. M. Lehn, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, 4763.
- [15] J. M. Lehn, *Polymer International* **2002**, 51, 825.
- [16] J. M. Lehn, A. V. Eliseev, *Science* **2001**, 291, 2331.
- [17] J. M. Lehn, A. Singh, C. Fouquey, *Abstracts of Papers of the American Chemical Society* **1996**, 212, 205.
- [18] M. Kotera, J. M. Lehn, J. P. Vigneron, *Journal of the Chemical Society-Chemical Communications* **1994**, 197.
- [19] J. M. Lehn, *Abstracts of Papers of the American Chemical Society* **1990**, 199, 399.

- [20] C. N. R. Rao, *Current Science* **2001**, *81*, 1030.
- [21] G. M. Whitesides, M. Boncheva, *Vol. 99*, **2002**, pp. 4769.
- [22] D. Thirumalai, G. H. Lorimer, *Vol. 30*, **2001**, pp. 245.
- [23] E. Winfree, *Vol. 2006*.
- [24] E. Winfree, in *DNA Based Computers: DIMACS Workshop, Vol. 27* (Eds.: R. J. Lipton, E. B. Baum), American Mathematical Society, Providence, RI, **1996**, pp. 199.
- [25] E. Winfree, *Journal of Biomolecular Structure & Dynamics* **2000**, 263.
- [26] E. Winfree, R. Bekbolatov, *DNA Based Computers* **2003**, 9.
- [27] E. Winfree, R. Bekbolatov, *DNA Computing* **2004**, 2943, 126.
- [28] E. Winfree, F. Liu, L. A. Wenzler, N. C. Seeman, *Nature* **1998**, *394*, 539.
- [29] M. G. Rossmann, P. Argos, *Vol. 50*, **1981**, pp. 497.
- [30] C. M. Niemeyer, *Nature* **2004**, *430*, 20.
- [31] C. M. Niemeyer, *Current Opinion in Chemical Biology* **2000**, *4*, 609.
- [32] I. Willner, *Febs Journal* **2007**, *274*, 301.
- [33] L. Lederman, *Biotechniques* **2006**, *41*, 29.
- [34] G. M. Whitesides, *Nature Biotechnology* **2003**, *21*, 1161.
- [35] K. Keren, R. S. Berman, E. Buchstab, U. Sivan, E. Braun, *Science* **2003**, *302*, 1380.
- [36] B. Samori, G. Zuccheri, *Angewandte Chemie International Edition* **2004**, *44*, 1166.
- [37] E. N. Trifonov, *Bull. Math. Biol.* **1989**, *51*, 417.
- [38] V. V. Demidov, M. D. Frank-Kamenetskii, *Trends Biochem Sci* **2004**, *29*, 62.
- [39] J. A. Schellman, S. C. Harvey, *Biophysical Chemistry* **1995**, *55*, 95.
- [40] J. Bednar, P. Furrer, V. Katritch, A. Z. Stasiak, J. Dubochet, A. Stasiak, *Journal of Molecular Biology* **1995**, *254*, 579.
- [41] C. R. Calladine, H. R. Drew, *Journal of Molecular Biology* **1996**, *257*, 479.
- [42] D. M. Crothers, *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 15163.
- [43] M. A. El Hassan, C. R. Calladine, *Philosophical transactions-Royal Society of London. Physical Sciences and Engineering* **1997**, 355.

- [44] J. Griffith, M. Bleyman, C. A. Rauch, P. A. Kithchin, P. T. Englund, *Cell* **1986**, 46.
- [45] R. E. Dickerson, D. Goodsell, M. L. Kopka, *Journal of Molecular Biology* **1996**, 256, 108.
- [46] P. R. Hardwidge, R. B. Den, E. D. Ross, L. J. Maher, *Journal of Biomolecular Structure & Dynamics* **2000**, 18, 219.
- [47] J. D. Kahn, E. Yun, D. M. Crothers, *Nature* **1994**, 368, 163.
- [48] Y. Zhang, A. E. McEwen, D. M. Crothers, S. D. Levene, *Biophysical Journal* **2006**, 90, 1903.
- [49] Y. Zhang, D. M. Crothers, *Vol. 84*, **2003**, pp. 136.
- [50] D. Shore, J. Langowski, R. L. Baldwin, *Vol. 78*, **1981**, pp. 4833.
- [51] J. A. H. Cognet, C. Pakleza, D. Cherny, E. Delain, E. Le Cam, *Journal of Molecular Biology* **1999**, 285, 997.
- [52] G. Muzard, B. Theveny, B. Revet, *Embo Journal* **1990**, 9, 1289.
- [53] C. Rivetti, C. Walker, C. Bustamante, *Journal of Molecular Biology* **1998**, 280, 41.
- [54] W. K. Olson, V. B. Zhurkin, *Current Opinion in Structural Biology* **2000**, 10, 286.
- [55] E. Le Cam, F. Fack, J. Menissier-de Murcia, J. A. Cognet, A. Barbin, V. Sarantoglou, B. Revet, E. Delain, G. de Murcia, *J. Mol. Biol* **1994**, 235.
- [56] T. Akiyama, M. E. Hogan, *Biochemistry* **1997**, 36, 2307.
- [57] G. Zuccheri, A. Scipioni, V. Cavaliere, G. Gargiulo, P. De Santis, B. Samori, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, 98, 3074.
- [58] A. A. Gorin, V. B. Zhurkin, W. K. Olson, *Journal of Molecular Biology* **1995**, 247, 34.
- [59] N. C. Seeman, A. M. Belcher, *Proc Natl Acad Sci U S A* **2002**, 99 Suppl 2, 6451.
- [60] A. M. Belcher, X. H. Wu, R. J. Christensen, P. K. Hansma, G. D. Stucky, D. E. Morse, *Nature* **1996**, 381, 56.
- [61] G. Falini, S. Albeck, S. Weiner, L. Addadi, *Science* **1996**, 271, 67.
- [62] S. Brown, *Nature Biotechnology* **1997**, 15, 269.

- [63] S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara, A. M. Belcher, *Nature* **2000**, *405*, 665.
- [64] A. Scipioni, S. Pisano, A. Bergia, M. Savino, B. Samori, P. De Santis, *Chembiochem* **2006**, *7*, 1645.
- [65] B. Sampaolese, A. Bergia, A. Scipioni, G. Zuccheri, M. Savino, B. Samori, P. De Santis, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 13566.
- [66] G. Felsenfeld, D. R. Davies, A. Rich, *J. Am. Chem. Soc* **1957**, *79*, 2023.
- [67] J.-L. Mergny, C. Helene, *Nat Med* **1998**, *4*, 1366.
- [68] A. T. Phan, J.-L. Mergny, *Vol. 30*, **2002**, pp. 4618.
- [69] R. D. Wells, *Vol. 263*, **1988**, pp. 1095.
- [70] C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, *Science* **1994**, *265*, 1599.
- [71] C. M. Niemeyer, M. Adler, *Angew Chem Int Ed Engl* **2002**, *41*, 3779.
- [72] N. C. Seeman, *Trends Biochem Sci* **2005**, *30*, 119.
- [73] M. Brucale, G. Zuccheri, B. Samori, *Trends in Biotechnology* **2006**, *24*, 235.
- [74] C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff, *Nature* **1996**, *382*, 581.
- [75] C. M. Niemeyer, M. Adler, S. Gao, L. F. Chi, *Angewandte Chemie-International Edition* **2000**, *39*, 3055.
- [76] S. J. Park, A. A. Lazarides, C. A. Mirkin, R. L. Letsinger, *Angew Chem Int Ed Engl* **2001**, *40*, 2909.
- [77] K. A. Williams, P. T. M. Veenhuizen, B. G. d. l. Torre, R. Eritja, C. Dekker, *Nature* **2002**, *420*, 761.
- [78] H. Yan, S. H. Park, G. Finkelstein, J. H. Reif, T. H. LaBean, *Science* **2003**, *301*, 1882.
- [79] E. Katz, I. Willner, **2004**, *43*, 6042.
- [80] E. Katz, I. Willner, *Chemphyschem* **2004**, *5*, 1084.
- [81] R. Singh, D. Pantarotto, D. McCarthy, O. Chaloin, J. Hoebeke, C. D. Partidos, J. P. Briand, M. Prato, A. Bianco, K. Kostarelos, *J Am Chem Soc.* **2005**, *127*, 4388.
- [82] G. Zuccheri, M. Brucale, B. Samorì, *Small* **2005**, *1*, 590.
- [83] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, *22*, 1859.
- [84] L. Fruk, J. Muller, C. M. Niemeyer, *Chemistry* **2006**, *12*, 7448.

- [85] L. Fruk, C. M. Niemeyer, *Angew Chem Int Ed Engl* **2005**, *44*, 2603.
- [86] M. Lovrinovic, L. Fruk, H. Schroder, C. M. Niemeyer, *Chem Commun (Camb)* **2007**, 353.
- [87] Y. Chen, C. Mao, *J Am Chem Soc* **2004**, *126*, 13240.
- [88] E. Braun, Y. Eichen, U. Sivan, G. Ben-Yoseph, *Nature* **1998**, *391*, 775.
- [89] L. M. Adleman, *Science* **1994**, *266*, 1021.
- [90] L. M. Adleman, P. W. Rothmund, S. Roweis, E. Winfree, *J Comput Biol* **1999**, *6*, 53.
- [91] R. H. Baughman, A. A. Zakhidov, W. A. de Heer, *Science* **2002**, *297*, 787.
- [92] M. Zheng, A. Jagota, E. D. Semke, B. A. Diner, R. S. Mclean, S. R. Lustig, R. E. Richardson, N. G. Tassi, *Nature Materials* **2003**, *2*, 338.
- [93] M. Zheng, A. Jagota, M. S. Strano, A. P. Santos, P. Barone, S. G. Chou, B. A. Diner, M. S. Dresselhaus, R. S. McLean, G. B. Onoa, G. G. Samsonidze, E. D. Semke, M. Usrey, D. J. Walls, *Science* **2003**, *302*, 1545.
- [94] J. Wang, *Electroanalysis* **2005**, *17*, 7.
- [95] B. J. Taft, A. D. Lazareck, G. D. Withey, A. Yin, J. M. Xu, S. O. Kelley, *J Am Chem Soc* **2004**, *126*, 12750.
- [96] T. Lin, V. Bajpal, T. Ji, L. Dai, *Aust. J. Chem.* **2003**, *56*, 635.
- [97] M. Hazani, F. Hennrich, M. Kappes, R. Naaman, D. Peled, V. Sidorov, D. Shvarts, *Chemical Physics Letters* **2004**, *391*, 389.
- [98] M. Hazani, R. Naaman, F. Hennrich, M. M. Kappes, *Nano Letters* **2003**, *3*, 153.
- [99] C. M. Niemeyer, B. Ceyhan, M. Noyong, U. Simon, *Biochem Biophys Res Commun* **2003**, *311*, 995.
- [100] C. Dwyer, M. Guthold, M. Falvo, S. Washburn, R. Superfine, D. Erie, *Nanotechnology* **2002**, *13*, 601.
- [101] C. Dwyer, V. Johri, M. Cheung, J. Patwardhan, A. Lebeck, D. Sorin, *Nanotechnology* **2004**, *15*, 1240.
- [102] H. Xin, A. T. Woolley, *J Am Chem Soc* **2003**, *125*, 8710.
- [103] M. C. Daniel, D. Astruc, *Chem. Rev.* **2004**, *104*, 293.
- [104] R. Elghanian, J. J. Storhoff, R. C. Mucic, R. L. Letsinger, C. A. Mirkin, *Science* **1997**, *277*, 1078.

- [105] J. J. Storhoff, A. A. Lazarides, R. C. Mucic, C. A. Mirkin, R. L. Letsinger, G. C. Schatz, *Journal of the American Chemical Society* **2000**, *122*, 4640.
- [106] B. W. Kirk, M. Feinsod, R. Favis, R. M. Kliman, F. Barany, *Nucleic Acids Research* **2002**, *30*, 3295.
- [107] B. Schweitzer, S. Wiltshire, J. Lambert, S. O'Malley, K. Kukanskis, Z. R. Zhu, S. F. Kingsmore, P. M. Lizardi, D. C. Ward, *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 10113.
- [108] R. C. Jin, G. S. Wu, Z. Li, C. A. Mirkin, G. C. Schatz, *Journal of the American Chemical Society* **2003**, *125*, 1643.
- [109] A. P. Alivisatos, K. P. Johnsson, X. G. Peng, T. E. Wilson, C. J. Loweth, M. P. Bruchez, P. G. Schultz, *Nature* **1996**, *382*, 609.
- [110] C. J. Loweth, W. B. Caldwell, X. G. Peng, A. P. Alivisatos, P. G. Schultz, *Angewandte Chemie-International Edition* **1999**, *38*, 1808.
- [111] W. J. Parak, T. Pellegrino, C. M. Micheel, D. Gerion, S. C. Williams, A. P. Alivisatos, *Nano Letters* **2003**, *3*, 33.
- [112] P. Hazarika, B. Ceyhan, C. M. Niemeyer, *Angew Chem Int Ed Engl* **2004**, *43*, 6469.
- [113] J. W. Liu, Y. Lu, *Journal of the American Chemical Society* **2003**, *125*, 6642.
- [114] S. W. Santoro, G. F. Joyce, *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 4262.
- [115] D. G. Georganopoulou, L. Chang, J. M. Nam, C. S. Thaxton, E. J. Mufson, W. L. Klein, C. A. Mirkin, *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 2273.
- [116] J. M. Nam, C. S. Thaxton, C. A. Mirkin, *Science* **2003**, *301*, 1884.
- [117] C. Sonnichsen, B. M. Reinhard, J. Liphardt, A. P. Alivisatos, *Nature Biotechnology* **2005**, *23*, 741.
- [118] G. Liu, Y. Yin, S. Kunchakarra, B. Mukherjee, D. Gerion, S. D. Jett, D. G. Bear, J. W. Gray, A. P. Alivisatos, L. P. Lee, F. F. Chen, *Nature Nanotechnology* **2006**, *1*.
- [119] N. C. Seeman, *Methods Mol Biol* **2005**, *303*, 143.
- [120] N. C. Seeman, *Sci Am* **2004**, *290*, 64.
- [121] N. C. Seeman, *Nature* **2003**, *421*, 427.

- [122] K. V. Gothelf, T. H. LaBean, *Org Biomol Chem* **2005**, *3*, 4023.
- [123] H. Yan, *Science* **2004**, *306*, 2048.
- [124] H. Yan, *Science* **2004**, *306*, 2048.
- [125] K. Lund, B. Williams, Y. G. Ke, Y. Liu, H. Yan, *Current Nanoscience* **2006**, *2*, 113.
- [126] N. C. Seeman, *J Theor Biol* **1982**, *99*, 237.
- [127] N. C. Seeman, *Annu Rev Biophys Biomol Struct* **1998**, *27*, 225.
- [128] R. Holliday, *Genet. Res.* **1964**, *5*, 282.
- [129] N. R. Kallenbach, R. I. Ma, N. C. Seeman, *Nature* **1983**, *305*, 829.
- [130] N. R. Kallenbach, R. I. Ma, A. J. Wand, G. H. Veeneman, J. H. van Boom, N. C. Seeman, *J Biomol Struct Dyn* **1983**, *1*, 159.
- [131] N. D. F. Grindley, K. L. Whiteson, P. A. Rice, *Vol. 75*, **2006**, pp. 567.
- [132] T. J. Fu, Y. C. Tse-Dinh, N. C. Seeman, *J Mol Biol* **1994**, *236*, 91.
- [133] N. C. Seeman, *J Biomol Struct Dyn* **1988**, *5*, 997.
- [134] N. C. Seeman, *DNA Cell Biol* **1991**, *10*, 475.
- [135] H. Qiu, J. C. Dewan, N. C. Seeman, *J Mol Biol* **1997**, *267*, 881.
- [136] C. Mao, W. Sun, N. C. Seeman, *Nature* **1997**, *386*, 137.
- [137] D. Liu, M. Wang, Z. Deng, R. Walulu, C. Mao, *J Am Chem Soc* **2004**, *126*, 2324.
- [138] R. P. Goodman, R. M. Berry, A. J. Turberfield, *Chem Commun (Camb)* **2004**, 1372.
- [139] R. P. Goodman, I. A. Schaap, C. F. Tardin, C. M. Erben, R. M. Berry, C. F. Schmidt, A. J. Turberfield, *Science* **2005**, *310*, 1661.
- [140] W. M. Shih, J. D. Quispe, G. F. Joyce, *Nature* **2004**, *427*, 618.
- [141] C. Mao, W. Sun, N. C. Seeman, *J. Am. Chem. Soc* **1999**, *121*, 5437.
- [142] T. J. Fu, N. C. Seeman, *Biochemistry* **1993**, *32*, 3211.
- [143] X. Li, X. Yang, J. Qi, N. C. Seeman, *J. Am. Chem. Soc* **1996**, *118*, 6131.
- [144] N. C. Seeman, *Biochemistry* **2003**, *42*, 7259.
- [145] F. Mathieu, S. Liao, J. Kopatsch, T. Wang, C. Mao, N. C. Seeman, *Nano Lett* **2005**, *5*, 661.
- [146] S. Roweis, E. Winfree, *J Comput Biol* **1999**, *6*, 65.

- [147] S. Roweis, E. Winfree, R. Burgoyne, N. V. Chelyapov, M. F. Goodman, P. W. Rothmund, L. M. Adleman, *J Comput Biol* **1998**, *5*, 615.
- [148] C. Mao, T. H. LaBean, J. H. Reif, N. C. Seeman, *Nature* **2000**, *407*, 493.
- [149] T. H. LaBean, H. Yan, J. Kopatsch, F. Liu, E. Winfree, J. H. Reif, N. C. Seeman, *J Am Chem Soc* **2000**, *122*, 1848.
- [150] S. H. Park, R. Barish, H. Y. Li, J. H. Reif, G. Finkelstein, H. Yan, T. H. LaBean, *Nano Letters* **2005**, *5*, 693.
- [151] B. Wei, Y. L. Mi, *Biomacromolecules* **2005**, *6*, 2528.
- [152] B. Ding, R. Sha, N. C. Seeman, *J Am Chem Soc* **2004**, *126*, 10230.
- [153] S. H. Park, P. Yin, Y. Liu, J. H. Reif, T. H. LaBean, H. Yan, *Nano Letters* **2005**, *5*, 729.
- [154] H. Wang, *Bell System Technical Journal* **1961**, *40*, 1.
- [155] P. W. Rothmund, N. Papadakis, E. Winfree, *PLoS Biol* **2004**, *2*, e424.
- [156] E. Winfree, *Nanotechnology: Science and Computation (in press)*.
- [157] R. Schulman, E. Winfree, *DNA Computing* **2005**, *3384*, 319.
- [158] P. W. Rothmund, A. Ekani-Nkodo, N. Papadakis, A. Kumar, D. K. Fygenson, E. Winfree, *J Am Chem Soc* **2004**, *126*, 16344.
- [159] J. C. Mitchell, J. R. Harris, J. Malo, J. Bath, A. J. Turberfield, *J Am Chem Soc* **2004**, *126*, 16342.
- [160] A. Ekani-Nkodo, A. Kumar, D. K. Fygenson, *Phys Rev Lett* **2004**, *93*, 268301.
- [161] P. J. Paukstelis, J. Nowakowski, J. J. Birktoft, N. C. Seeman, *Chem Biol* **2004**, *11*, 1119.
- [162] P. E. Constantinou, T. Wang, J. Kopatsch, L. B. Israel, X. Zhang, B. Ding, W. B. Sherman, X. Wang, J. Zheng, R. Sha, N. C. Seeman, *Org Biomol Chem* **2006**, *4*, 3414.
- [163] A. Chworos, I. Severcan, A. Y. Koyfman, P. Weinkam, E. Oroudjev, H. G. Hansma, L. Jaeger, *Science* **2004**, *306*, 2068.
- [164] Y. Liu, Y. G. Ke, H. Yan, *Journal of the American Chemical Society* **2005**, *127*, 17140.
- [165] U. Feldkamp, H. Rauhe, W. Banzhaf, *Genetic Programming and Evolvable Machines* **2003**, *4*, 153.
- [166] A. Brenneman, A. Condon, *Theoretical Computer Science* **2002**, *287*, 39.

- [167] Y. He, Y. Tian, Y. Chen, Z. Deng, A. E. Ribbe, C. Mao, *Angew Chem Int Ed Engl* **2005**, *44*, 6694.
- [168] R. M. Dirks, M. Lin, E. Winfree, N. A. Pierce, *Nucleic Acids Res* **2004**, *32*, 1392.
- [169] N. C. Seeman, *Q Rev Biophys* **2005**, *38*, 363.
- [170] H. Yan, T. H. LaBean, L. Feng, J. H. Reif, *Proc Natl Acad Sci U S A* **2003**, *100*, 8103.
- [171] C. Lin, Y. Liu, S. Rinker, H. Yan, *Chemphyschem* **2006**, *7*, 1641.
- [172] S. H. Park, C. Pistol, S. J. Ahn, J. H. Reif, A. R. Lebeck, C. Dwyer, T. H. Labeau, *Angew Chem Int Ed Engl* **2006**, *45*, 6607.
- [173] T. H. LaBean, E. Winfree, J. H. Reif, in *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* (Eds.: E. Winfree, D. K. Gifford), Am. Math. Soc., Providence, RI, **1999**, pp. 123.
- [174] P. W. Rothmund, *Nature* **2006**, *440*, 297.
- [175] D. Lubrich, J. Bath, A. J. Turberfield, *Nanotechnology* **2005**, *16*, 1574.
- [176] H. Liu, Y. Chen, Y. He, A. E. Ribbe, C. Mao, *Angew Chem Int Ed Engl* **2006**, *45*, 1942.
- [177] H. Liu, Y. He, A. E. Ribbe, C. Mao, *Biomacromolecules* **2005**, *6*, 2943.
- [178] R. P. Goodman, *Biotechniques* **2005**, *38*, 548.
- [179] N. C. Seeman, *J Biomol Struct Dyn* **1990**, *8*, 573.
- [180] C. Vaillant, B. Audit, C. Thermes, A. Arneodo, *Eur Phys J E Soft Matter* **2006**, *19*, 263.
- [181] Y. G. Ke, Y. Liu, J. P. Zhang, H. Yan, *Journal of the American Chemical Society* **2006**, *128*, 4414.
- [182] W. P. C. Stemmer, A. Cramer, K. D. Ha, T. M. Brennan, H. L. Heyneker, *Gene* **1995**, *164*, 49.
- [183] D. Reishus, B. Shaw, Y. Brun, N. Chelyapov, L. Adleman, *Journal of the American Chemical Society* **2005**, *127*, 17590.
- [184] Y. He, Y. Chen, H. Liu, A. E. Ribbe, C. Mao, *J Am Chem Soc* **2005**, *127*, 12202.
- [185] Y. He, C. Mao, *Chem Commun (Camb)* **2006**, 968.
- [186] F. C. Simmel, W. U. Dittmer, *Small* **2005**, *1*, 284.
- [187] C. Mao, W. Sun, Z. Shen, N. C. Seeman, *Nature* **1999**, *397*, 144.

- [188] M. Brucale, G. Zuccheri, B. Samori, *Org Biomol Chem* **2005**, *3*, 575.
- [189] Y. Chen, S. H. Lee, C. Mao, *Angew Chem Int Ed Engl* **2004**, *43*, 5335.
- [190] X. Yang, A. V. Vologodskii, B. Liu, B. Kemper, N. C. Seeman, *Biopolymers* **1998**, *45*, 69.
- [191] H. Yan, X. Zhang, Z. Shen, N. C. Seeman, *Nature* **2002**, *415*, 62.
- [192] L. Feng, S. H. Park, J. H. Reif, H. Yan, *Angew Chem Int Ed Engl* **2003**, *42*, 4342.
- [193] B. Yurke, A. J. Turberfield, A. P. Mills, Jr., F. C. Simmel, J. L. Neumann, *Nature* **2000**, *406*, 605.
- [194] J. J. Li, W. Tan, *Nano Lett* **2002**, *2*, 315.
- [195] A. J. Turberfield, J. C. Mitchell, B. Yurke, A. P. Mills, Jr., M. I. Blakey, F. C. Simmel, *Phys Rev Lett* **2003**, *90*, 118102.
- [196] W. B. Sherman, N. C. Seeman, *Nano Lett* **2004**, *4*, 1203.
- [197] J. S. Shin, N. A. Pierce, *J Am Chem Soc* **2004**, *126*, 10834.
- [198] Y. Tian, C. Mao, *J Am Chem Soc* **2004**, *126*, 11410.
- [199] W. Shen, M. F. Bruist, S. D. Goodman, N. C. Seeman, *Angew Chem Int Ed Engl* **2004**, *43*, 4750.
- [200] P. Yin, H. Yan, X. G. Daniell, A. J. Turberfield, J. H. Reif, *Angew Chem Int Ed Engl* **2004**, *43*, 4906.
- [201] Y. Chen, C. Mao, *J Am Chem Soc* **2004**, *126*, 8626.
- [202] G. Seelig, B. Yurke, E. Winfree, *DNA Computing* **2005**, *3384*, 329.
- [203] G. Seelig, B. Yurke, E. Winfree, *Journal of the American Chemical Society* **2006**, *128*, 12211.
- [204] T. Liedl, M. Olapinski, F. C. Simmel, *Angewandte Chemie-International Edition* **2006**, *45*, 5007.
- [205] T. Liedl, F. C. Simmel, *Nano Letters* **2005**, *5*, 1894.
- [206] W. U. Dittmer, F. C. Simmel, *Nano Letters* **2004**, *4*, 689.
- [207] W. U. Dittmer, S. Kempter, J. O. Radler, F. C. Simmel, *Small* **2005**, *1*, 709.
- [208] W. U. Dittmer, A. Reuter, F. C. Simmel, *Angewandte Chemie-International Edition* **2004**, *43*, 3550.
- [209] R. M. Dirks, N. A. Pierce, *Proc Natl Acad Sci U S A* **2004**, *101*, 15275.

- [210] B. Kolaric, M. Sliwa, M. Brucale, R. A. L. Vallée, G. Zuccheri, B. Samori, J. Hofkens, F. C. De Schryver, *Photochemical & Photobiological Sciences* **2007**, DOI: 10.1039/b618689k.

II – RESULTS AND DISCUSSION

3 Novel DNA Self-assembled Nanostructures: Design, Synthesis, Characterization, and Preliminary Studies on Implementations

In this chapter, I will describe and discuss several new DNA nanostructures that I designed, synthesized and characterized during my PhD candidature. Section 3.1 includes the results pertaining to static structures based on the DNA parallelogram motif. In section 3.2 I report the results regarding a dynamic DNA structure based on a conformational transition from double to triple helix. Finally in section 3.3 I report the design and (yet to be completed) synthesis of a DNA helix bundle based on the “DNA origami” design (see section 2.1.6).

3.1 Structures based on the DNA rhombus motif

As discussed in section 2.1.1, a single immobile Holliday junction is quite flexible, and the efforts aimed at building geometrically and topologically defined structures using the Holliday junction alone were unsuccessful. However, a parallelogram made from four linked Holliday junctions^[1, 2] is quite robust due to the mechanical coupling of all the junctions (see figure 1 and 2). A mechanical force applied to the parallelogram must distort all the junctions simultaneously, making the whole construct approximately as stiff as the sum of its parts.

A DNA parallelogram (also called sometimes “rhombus motif”) consists of two parallel double-helical edges above its central plane and two other parallel dsDNA edges below the plane (see figure 2). The plane contains the crossover points of the four Holliday junctions. The structure has eight protruding arms that can end with sticky ends (see section 2.1.2). This makes it possible to build self-assembled arrays^[2-4] of DNA parallelograms (see figure 3).

The characteristics of the DNA parallelogram make it an especially interesting motif for building nanoscale objects. It is quite robust, it does not need a large number of different synthetic oligonucleotides to be formed, and most importantly its shape

allows the placement of sticky ends in nearly-orthogonal directions, allowing different lattice propagation directions when polymerized.

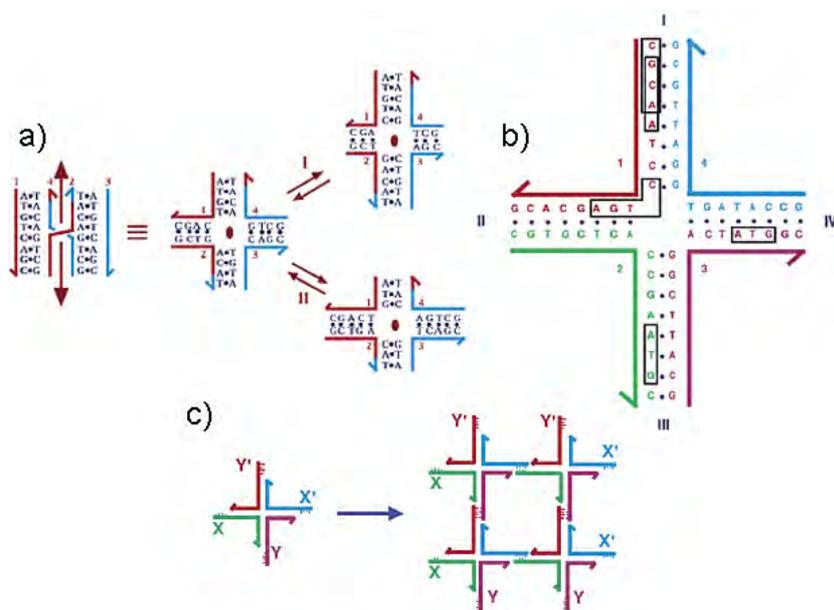


Figure 1 - (a) Branch migration. On the left, two homologous duplexes, red and blue, have undergone reciprocal exchange. There is a vertical dyad axis of sequence symmetry. In the center, the drawing has been rearranged so that the dyad is perpendicular to the plane of the page. Through the isomerization process of branch migration, the C's and G's flanking the junction exchange partners (process I), or the A's and T's exchange partners (process II). (b) An immobile branched junction. There is no dyad symmetry flanking the branch point; tetramers, such as the boxed sequences CGCA and GCAA are unique, and there is no TCAG to complement the CTGA flanking the corner. (c) Assembly of four immobile branched junctions to form a parallelogram. [Image taken from ref.2]

3.1.1 Design and synthesis of DNA parallelograms

The design of a DNA parallelogram starts with some basic decisions about the overall shape of the construct. With respect to size, an obvious constraint is that the double helixes separating the junctions must include an integer number of full 360° turns. If this condition is not met, each successive junction lies in a plane that is tilted with respect to the one containing the previous one, thus rendering the encounter of the first and fourth junction (and consequently the correct closure of the parallelogram) impossible. Since the average number of bases per full helical turn in B-DNA is around 10.5, it is usually safe to include a number of base pairs that is a multiple of 10.5.

So far, the parallelograms reported in literature have been designed to have multiples of two full helical turns (21 base pairs) between junctions. It is common to refer to DNA parallelograms indicating their size in terms of the number of full helical turns in their edges. Following this notation, the DNA parallelograms reported so far have internal sizes of 2×2 ^[5], 4×4 ^[11] (not to be confounded with the completely different “four-by-four” motif)^[6], 6×4 ^[7] and 6×6 ^[8]. The length of the edges of the parallelogram, influences the flexibility of the whole construct is, and thus its capability of self-assembling into a well-defined geometrical shape. Parallelograms with both pair of edges longer than 4 full helical turns seem to be rather untractable.^[8]

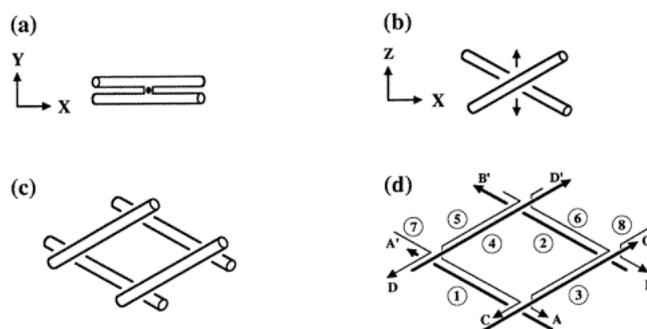


Figure 2 - Schematic representations of the main elements of a DNA parallelogram. (a) A view down the dyad axis of the Holliday junction. The dyad axis is indicated by the small circle. The upper helical domain is rotated 30 about the vertical so that its right end penetrates the page, and the lower helical domain is rotated 30 about the vertical so that its left end penetrates the page. The X and Y axes of a right-handed coordinate system are shown to help orient the reader. (b) A view with the dyad axis vertical. The molecule has been rotated 90 about the X axis, as indicated. The dyad axis is indicated by the double arrows. (c) The combination of four junctions into a rhombus-like motif. Four molecules, in the orientation of (b) are combined. There must be an integer number of full helical turns of DNA in each helix between crossover points; if this condition is not met the resulting structure cannot form correctly. (d) A possible strand structure for the rhombus-like motif. The strand structure of the molecule in (c) is shown. Strands that are geometrically continuous helices (numbered 1-4) are drawn with a dark line, and those involved in crossovers (numbered 5-8) are drawn with a thin line. Arrowheads indicate the 3' ends of the strands. Sticky ends are shown by the letters A, B, C, and D, and A', B', C', and D', respectively, represent their complements. [image taken from ref.1]

If the parallelogram is designed to self-assemble into larger superstructures, it must incorporate sticky ends; the obvious place to put these is on the remaining double helices protruding from the junctions toward the outside of the parallelogram. As many as eight sticky ends can be placed on these protruding arms to direct the self-assembly of a collection of parallelograms. However, it is important to note that the same constraint that apply for the edges (regarding the number of base pairs between junctions) must also be satisfied by the new parallelogram shapes that appear between

the junctions of two different parallelograms upon their adjoining. This means that the sum of complementary sticky ends must afford double helices having an integer number of full turns between the junctions that are joined. Also in this case, the “quantum” of full helical turns between sticky-end-joined junctions reported in literature is two.

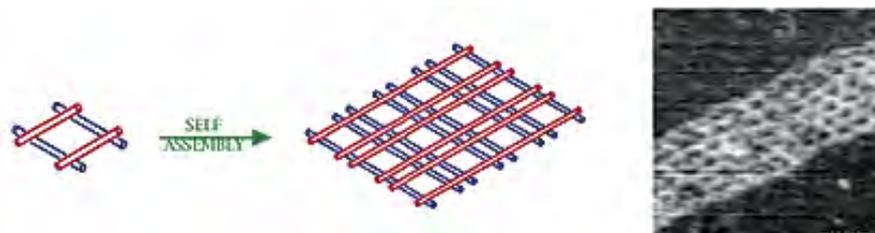


Figure 3 – An example of array build from DNA parallelograms. In this case, the parallelogram has sticky ends designed so that it can self-associate in two dimensions to form a 2D array. The size of the parallelogram cavities in this example are approximately 13 by 20 nanometers [image taken from ref.2]

For example, in the first parallelogram reported,^[1] six turns of DNA were used in each direction, with overhangs of one turn on each side, leading to four turns on each edge within the parallelogram, described by a notation of $(4+2) \times (4+2)$; thus, the periodic array contains a 4×4 turn parallelogram, a 4×2 turn parallelogram, a 2×4 turn parallelogram and a 2×2 turn parallelogram. The same paper reported a larger parallelogram, $(6+2) \times (4+2)$.^[1] Seeman and coworkers reported the partially successful formation of parallelograms of slightly larger dimensions.^[8] Arrays with sizes of $(4+4) \times (4+4)$, and $(6+2) \times (6+2)$ have been reported.^[8] Attempts to achieve larger arrays were unsuccessful when single helical domains are used for each direction.

Once the size of the parallelogram is decided, the length and position of each synthetic oligonucleotide in the scheme must be decided. Of course, shorter oligonucleotides are usually cheaper and can be purified more easily. However, since the size of the construct is already decided, using shorter oligonucleotides also means that a larger number of them is needed. This has two repercussions: it is more difficult to mix a large number of components in exact stoichiometric amount (and this is quite important for the correct formation of the parallelogram, see below), and the structure will have more nicks than it would have if it was composed by longer oligonucleotides. Therefore, DNA parallelograms are usually built using the longest synthetic oligonucleotides conveniently available that can be incorporated in the strand structure.

The actual number of oligonucleotides used to build a parallelogram varies depending on the strategy employed. The first strategy stems from the fact that a DNA parallelogram can be described as the juxtaposition of four J1 junctions, in which each arm terminates with a sticky end. To obtain the desired rhomboidal structure, the sticky-end-mediated self assembly of the junctions is followed by ligation (see figure 4). Since each J1 junction is formed by at least four synthetic oligonucleotides, this strategy requires the assembly of sixteen of them (four oligonucleotides per junction) prior to the enzymatic reaction.

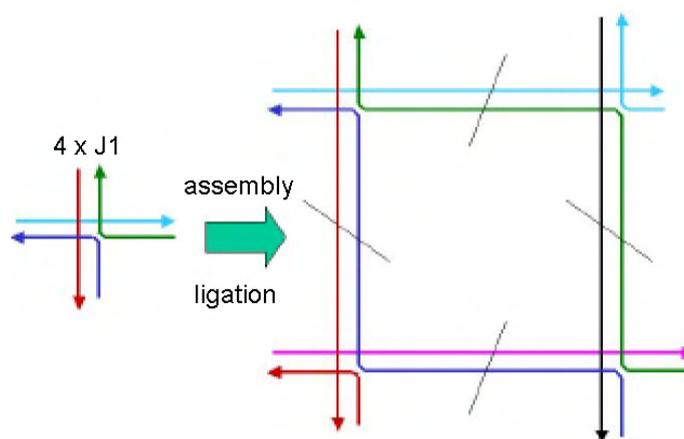


Figure 4 – Four J1 junctions are assembled via sticky ends, then ligated together to afford a DNA parallelogram. This strategy requires sixteen different oligonucleotides and an enzymatic reaction. The purification of a preformed DNA structure from the enzymatic reaction mixture is not usually trivial. Notice how, after the ligation step, only eight ssDNA strands are intertwined together to form the parallelogram [image courtesy of Dr. Andrea Giro, <http://www.puntogiro.com/>].

After the ligation, the parallelogram is composed by only eight continuous ssDNA strands (see figure 4). It is extremely more convenient to form directly the parallelogram from eight synthetic, longer oligonucleotides than to ligate sixteen shorter ones as detailed above.

It is possible to further reduce the number of different oligonucleotides needed to build the parallelogram if not all of the protruding arms are required to have sticky ends. If the two strands of a protruding arm must not form a sticky end, they can be “soldered” together to form an hairpin structure, thus effectively reducing the number of oligonucleotides by one for each sticky end eliminated (see figure 5).

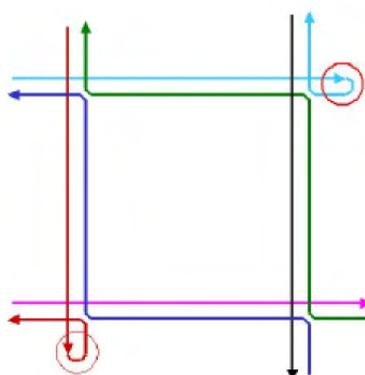


Figure 5 – Example of a parallelogram with a reduced number of constituent oligonucleotides (and consequently sticky ends). In this case, only six different oligonucleotides are needed to form the structure. Other combinations of “soldered” positions can be implemented, but no reported parallelogram is formed by less than six oligonucleotides. The reason for this is also that soldering some arms (for example the top left ones in the figure) would result in awkwardly long oligonucleotides [image courtesy of Dr. Andrea Giro].

The remaining step of the design is deciding the sequences of the oligonucleotides. The first feature to include in the sequences is the non-symmetry of the portions flanking each junction, to avoid branch migration. Usually, the first few bases around each junction are the same (see figure 6). The arrangement of bases depicted in the figure ensures that no branch migration is possible and leads to well-behaved, immobile nucleotide junctions.

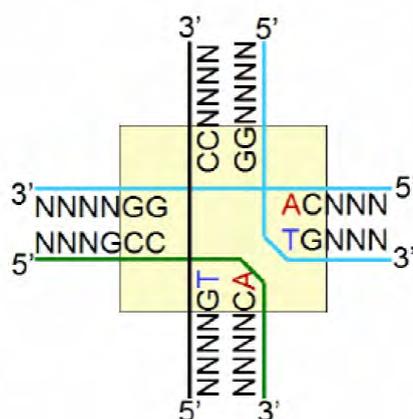


Figure 6 – The sequences used to stabilize each junction in the parallelogram motifs (usually called affectionately “the magic box”). The letter N indicates a generic nucleobase [image courtesy of Dr. Andrea Giro].

All the remaining nucleobases are selected so that no significant overlap of sequences exist between different oligonucleotides or tracts of the same oligonucleotide.

This step is necessarily performed with the aid of a computer; several good programs exist that can serve to the purpose. The first program developed to this end is SEQUIN,^[9] a sequence symmetry minimization algorithm optimized for the design of branched structures. While it was written almost twenty years ago and initially not intended for public use, SEQUIN is still perfectly functional and is used by many laboratories around the world. About half of the parallelograms reported in this thesis were designed using this software. A worthwhile competitor, especially with respect to user-friendliness, is the software NANEV.^[10] The other half of the parallelograms reported in this thesis were designed with NANEV. Other extremely well-designed software packages, useful for this step of design, are UNIQUIMER,^[11] and GIDEON.^[12]

Once the sequences have been defined, the oligonucleotides are synthesized via solid-phase supported phosphoramidite synthesis. The formation of the supramolecular DNA adduct is then very straightforward. All the constituent components are brought together in a buffer (frequently TAE) additioned with 12.5 mM of Mg²⁺. The high magnesium concentration is needed to stabilize the sterically overcrowded junctions.

Then, the temperature is raised to 90°C for a few minutes to ensure that no intra-strand secondary structure is still present when the assembly begins. The temperature is then lowered as slowly as conveniently possible (a cooling rate of 0.01°C/s is usually slow enough for the correct formation of most structures), until the thermal stability region of the complete structure is reached. This process is usually referred simply as “annealing” the structure.

It is very important that all the components are present at equimolar stoichiometry, since if one component is in excess with respect to the others, this will promote the formation of incomplete structures. To this end, titrations of complementary strands followed by gel electrophoresis are more useful than standard UV measurements.

An useful strategy to verify the correct formation of the parallelogram is to anneal together partial structures in which only a portion of the constituent oligonucleotides assemble together. An example of this strategy is reported in figure 7.

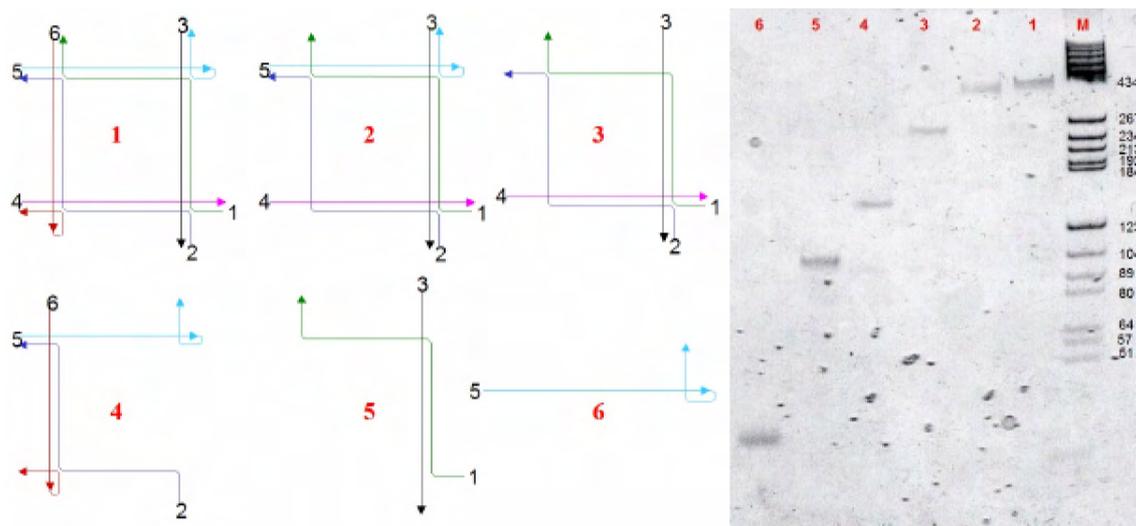


Figure 7 – Verifying the correct assembly of a DNA parallelogram composed by six different oligonucleotides. Oligo 1 and 2 are 104 bases long, oligo 3 and 4 are 68 bases long and oligo 5 and 6 are 78 bases long. The sub-structures shown on the left are annealed from 90°C to 4°C and then run on a 5% polyacrylamide gel at 7V/cm for 4h. In each lane, only a single band is visible, confirming that every oligo binds correctly to its intended partner and none of them is present in excess to the others. In lane one, the complete parallelogram runs as a single object [image courtesy of Dr. Andrea Giro].

3.1.2 Mono-dimensional arrays made of DNA parallelograms

As reported in Section 2, numerous different DNA structural motifs exist, and the literature reports many proofs that they can self-assemble with high efficiency. However, the mechanical knowledge of the involved processes is still fairly limited. To address an aspect of this issue, we studied the aggregation of a series of DNA rhomboidal supramolecular tiles in various conditions.

DNA nanoarchitectures can be designed to be monomeric or polymeric in nature. The former approach focuses on obtaining the desired shape directly in one hierarchical step from discrete oligonucleotides. Remarkable implementations of this approach appeared only recently in the literature.^[13, 14]

Most published DNA nanoarchitectures are however designed as polymers created by the assembly of monomeric repetitive units, often called “tiles,” capable of binding to each other in a programmed way towards higher and higher hierarchical levels of structural organization.^[15] The overall shape of the architecture is critically dependent on the topological connectivity of the tiles. This means that tiles that are structurally very similar can assemble into radically different architectures thanks to the choice among possible connectivities. For example a line

(1D connectivity) or a plane (2D connectivity) can be obtained with virtually the same tile.^[6, 16, 17]

Structures obtained by the 2D assembly of tiles (planes and tubes) comprise the majority of those reported in literature, while 1D assemblies are comparatively less abundant.^[1, 3, 16, 17]

At each level of hierarchy, all the aforementioned DNA architectures are bound by reversible interactions, thus they can also be viewed as supramolecular polymers.^[18] It has been shown that the polymerization of DNA-based monomers yields well-behaved, reversible polymers, the structure of which depends on a variety of factors including the strength and specificity of the association, the rigidity of the monomers, their concentration, and the environment in which the polymerization occurs.^[19, 20] This implies that DNA supramolecular polymers can be classified as ‘constitutionally dynamic materials,’^[18, 21] as their organization is never permanently defined but it can change in response to the environment.

In this section, I will present a family of DNA supramolecular architectures based on the polymerization of discrete DNA tiles having the shape of parallelograms and designed to have a one-dimensional inter-tile connectivity. One type of monomer can yield classes of supramolecular polymers of different shapes and size just by changing the polymerization conditions. Changing the base composition by just one base pair (among 142) in the monomer leads to dramatically different behavior. Moreover, this system adapts in response to external stimuli by interconverting between some of its possible states. In my opinion, this is an interesting model system for improving the knowledge on self-assembly processes.

3.1.2.1 Results

(i) Design of a structurally controlled supramolecular polymer based on nanometric DNA parallelogram tiles.

A mechanically rigid and structurally controlled supramolecular polymer is obtained thanks to the assembly of flat and rigid parallelogram-shaped tiles (see Figure 8) as first suggested by Seeman and coworkers.^[1] Each tile self-assembles

into a single possible structure from the designed interaction of six different oligonucleotides. The assembly occurs on cooling the solution from a state where no interaction is stable to one where individual tiles form (below the melting temperature, T_m , of the extended interactions amongst oligonucleotides). A slow cooling rate ensures that the assembly takes place at quasi-equilibrium conditions and only the most stable adduct forms.

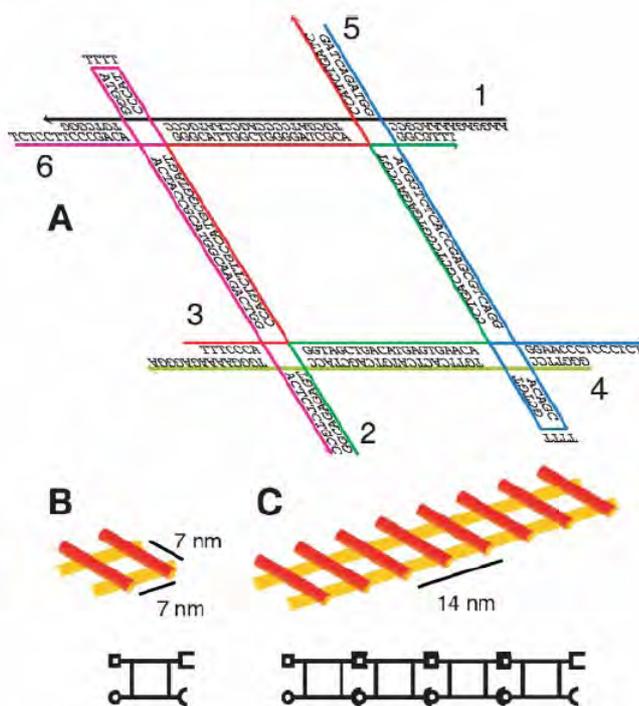


Figure 8 – (A) Oligonucleotide sequences and structure of a DNA parallelogram tile. The tile depicted here has two pairs of 7-nt-long sticky ends with different sequences and base composition and thus a different thermal stability. In this study, I also employed another parallelogram tile, identical to the previous one except for the sequence of one pair of sticky ends, which is in this case designed to have the same base composition of the other pair of sticky ends, but a different sequence. (B) 3D model of one parallelogram tile, where helices are represented as cylinders. Below, its schematic representation used in the next figures. (C) 3D model of a one-dimensional array of parallelograms linked to each other by means of both pairs of sticky ends (and schematic representation below).

As showed in Figure 8, each tile is made of 4 four-way junctions flanking two turns of double-stranded DNA (dsDNA). On two opposite sides, connections are predisposed through 7-nucleotide (nt) long sticky ends, protruding from 7-base pair (bp) long dsDNA segments. The T_m of these dsDNA segments is considerably lower than that characterizing the extended interactions that hold the tile together. The marked difference between these two regions of thermal stability makes the assembly hierarchical and enables the separation (in time or in space) of the two

assembly events. In our experiments, this feature allowed us to assemble the tiles in optimal conditions (with respect to concentration, temperature and cooling rate), and then study the polymerization in completely different conditions.

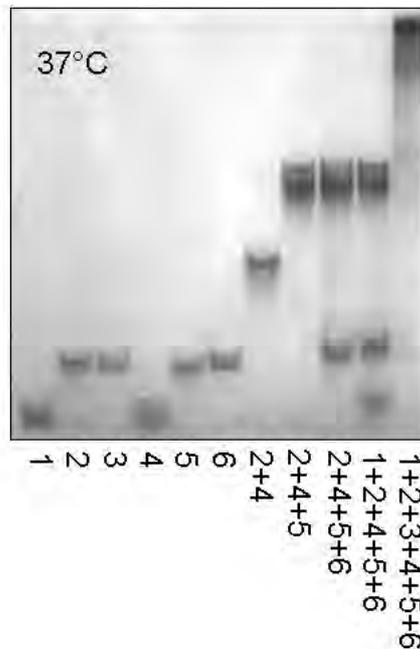
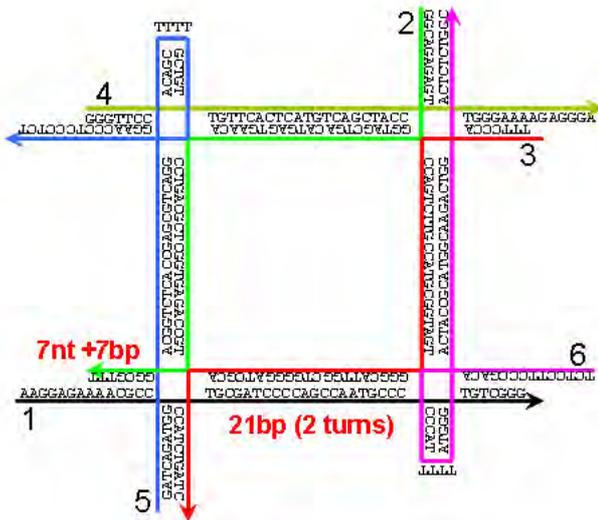


Figure 9 – Strand structure and sequences of the parallelogram discussed in this section, and electrophoretic run of constituent components at 37°C. Some sub-structures do not run as a single band at this temperature, but the complete parallelogram does (rightmost lane). This is because some incomplete structures are formed by portions linked only by a limited number of base pairs, but each oligonucleotide in the complete structure is bound to the others by more than forty base pairs.

As shown in Figure 8C, the assembly resulting from polymerization of one tile around its two couples of aligned sticky ends is a 1D rail-like structure that should be rigid due to the mechanical coupling among all the possible deformations. In-plane deformations seem to be particularly difficult. Compared to all previous implementations of this same DNA motif, we implemented a smaller tile, thus a smaller separation between the junctions (two turns compared to four) and longer, more stable sticky ends in order for the assembly to result more rigid and stably connected.

The resulting assembled 1D rail-like system has two rows of 4-way junctions separated by 7 nm along the longitudinal axis of the chain, and each couple of junctions is spaced 7 nm on the width of the chain. Each repeating unit thus occupies

14 nm of length along the chain (see figure 8). The joining of two tiles around only one of the two possible connections should guarantee that the minimum energy structure is still the same flat rail as for the doubly-connected system, with the significant difference that the structural flexibility would be so much higher that many different conformations be available to the incompletely connected polymers.

In the final, completely connected polymer, each tile is linked to another by up to 35 hydrogen bonds, plus the stacking energy of the newly-formed base pairs. The bonds are reversible, implying that the degree of polymerization can be changed by thermal treatment or modifications of the concentration of macromonomers.

(ii) Assembling parallelograms with two sticky ends having the same melting profile

Our first attempt towards the making of rigid and linear rail-like polyparallelogram structures was made by joining two consecutive parallelograms by means of two 7-bp long dsDNA segments with the same base sequence. The two sticky ends on both sides of the tiles have the same sequence but different orientation (one has a 5' overhang, the other a 3'-overhang).

We performed a thermal assembly of a 0.1 μM solution of all the constituent oligonucleotides from 90°C (where all interactions are unstable) to 15°C (where polymers should be stable) at a cooling rate of 0.01°C/s (the entire process takes slightly over 2 hours). The success and efficiency of the assembly was verified via electrophoretic analysis as detailed in section 3.1.1 (see figure 9).

The resulting DNA nanostructures were observed with the atomic force microscope (AFM) at room temperature after spreading at 20°C. The AFM analysis gives information on the structure and size of polymeric chains. As it can be seen in the AFM micrographs (a sample is shown in the upper row of panels in Figure 10), the assembly proceeds with high conversion in the conditions employed, as very few dot-like monomeric objects are observed (some dot-like single parallelogram objects have been characterized by us before).^[15]

The images show straight or slightly curved rod-like shapes, rods with short branchings and a very limited number of small ring-like cyclical objects (cyclical structures have an average contour length of 87 nm and represent as little as 2 % of

the imaged objects). The average dimension of the rod-like objects is 106 nm (66 % of the sample) thus, considering a repeating unit of 14 nm, 7.6 units. Branched or otherwise “imperfect” rail objects represent about 32 % of the population, while they probably represent the majority of the mass content of the sample, since they have a bigger size than rods and circular shapes.

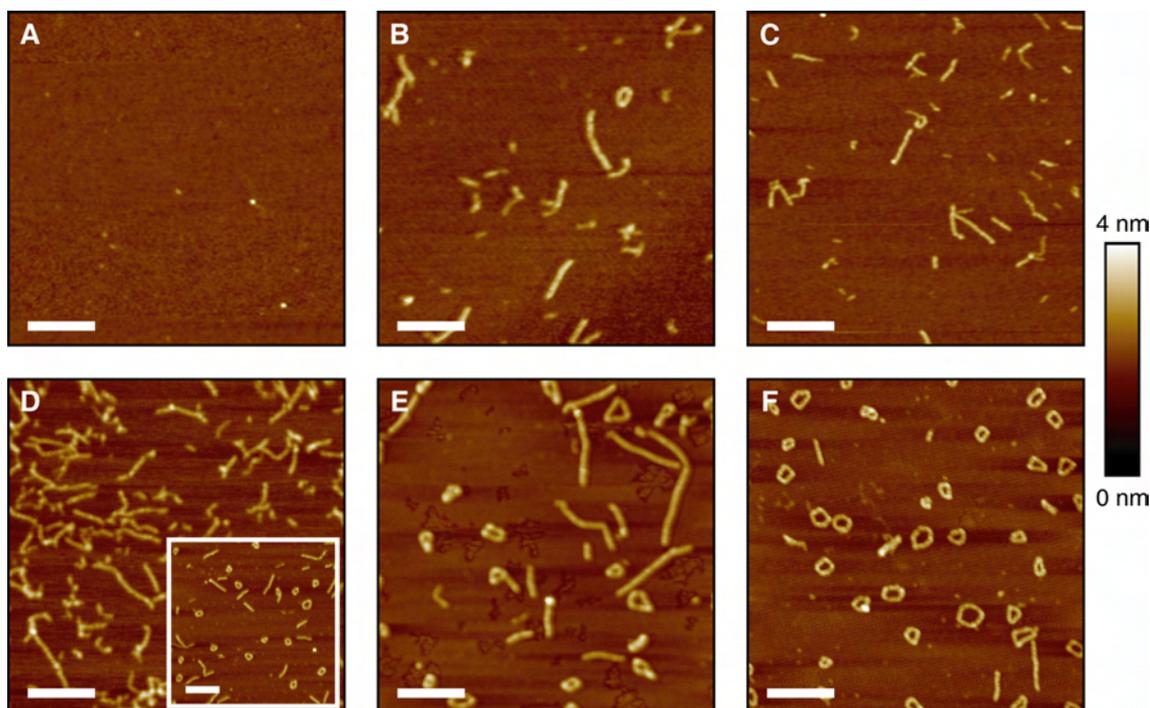


Figure 10 - AFM images representative of the result of the parallelogram polymerization in different conditions. (A-C) Annealing from 50 °C to 20 °C of the tile in which both pairs of sticky ends have the same thermal stability, at the same rate (0.01 °C/s) but different monomer concentrations: from left to right 1 nM, 10 nM, and 100 nM. (D-F) Annealing from 50 °C to 20 °C of the tile in which the two pairs of sticky ends have a different thermal stability at the same concentration (100 nM) but different cooling rates: from left to right 3.00 °C/s, 0.01 °C/s, approx. 0.0003 °C/s. The inset of panel D shows the result of a brief (5 min) heating of the same sample at 37 °C. All the scale bars of the images measure 200 nm.

All imaged objects show chains with a constant width compatible with the width of a parallelogram, demonstrating that, as designed, the connection among the tiles is one-dimensional.

(iii) Assembling parallelograms with isostable sticky ends at different concentration

The above-described assembly-by-polymerization was performed also at reduced concentrations. Taking advantage of the hierarchically separated assembly steps, the concentration of preformed but disjointed parallelogram tiles was adjusted

at 50°C and then the temperature was lowered to 20°C. Concentrations of 1, 10 and 100 nM were attempted. The results (reported in Figure 10A-C) demonstrate that at 1 nM concentration, tiles are so dilute that they do not assemble at all in the conditions of the experiments, while at 10 and 100 nM, the assembly produces virtually the same results (AFM imaging is performed immediately after dilution to the same concentration) with comparable resulting topology (branching) and chain size.

(iv) Slow Assembly of a DNA parallelogram with two different sticky ends

In order to improve and further characterize the assembly, we designed and implemented a polyparallelogram system based on a tile with two different pairs of sticky ends. The sticky ends have the same length while the base content is slightly different: one sticky end forms one G-C pair more than the other. This implies an approximately 2 °C melting temperature difference between them (see figure 8). The size and sequence of the rest of the parallelogram tiles is exactly the same as that of the previously reported experiments.

The obtained single parallelograms have been assembled with the same procedure as described above (100 nM oligonucleotide concentration, 0.01°C/s cooling rate from 90°C to 20°C). As evidenced by electrophoretic analysis, the assembly efficiency is comparable with the former case. The AFM study of the assembled products exhibits only two topologies of objects: rods and ring-like shapes of varying size (see Figure 9E). By careful examination of high-resolution micrographs, it appears that the ring-like cyclical objects are not smoothly circular, but instead have curved and straight sections of varying number and extension. For this reasons, we will from now refer to them as “taralli,” as they resemble the shape of these regional Italian bread-like snacks. By digitization of the AFM images, we could characterize the average size of the assemblies and the ratio between rods and taralli. Data have been obtained from 3441 objects, 31 % of which have a rod-like shape and a median length of 96 nm i.e. 6.9 repeating units, while the other 69 % are taralli, with a median length of 108 nm, i.e. 7.7 repeating units.

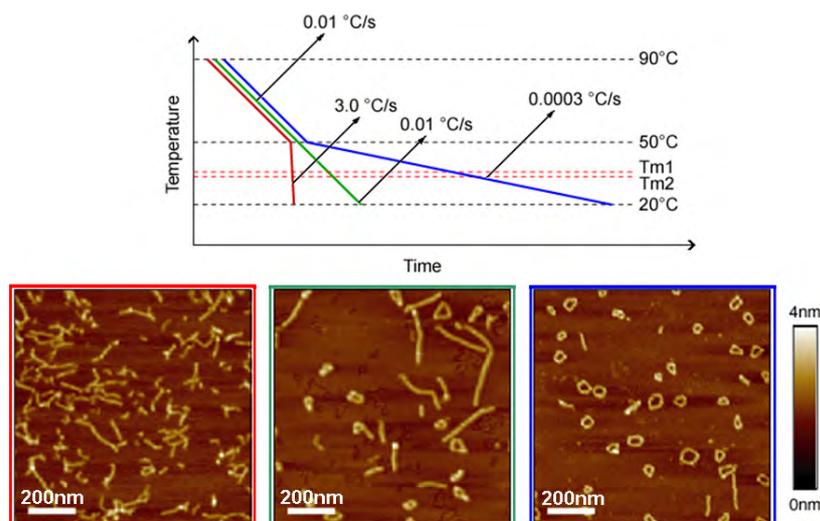


Figure 11 – Summary of the results of annealing the parallelogram with two different sticky ends at different cooling rates.

(v) Very slow assembly of a DNA parallelogram with two different sticky ends

The same experiment in the same conditions was repeated with an ultra-slow cooling rate. The high temperature mix was taken from 90°C to room temperature over a three-day time (0.0003 °C/s, approximately) by sealing a 2 l beaker of boiling water containing a floating tube with the mix in a thick styrofoam box and leaving it untouched at room temperature for 3 days.

The AFM imaging of this specimen revealed an overwhelming prevalence of taralli and very few rods (see Figure 12 for a sample AFM image). After digitization of the molecule profiles, only 70 rod-like object (11 % of the sample) out of 642 counted objects. Rod-like objects have a median length of 82 nm (i.e. 5.9 repeating units), while taralli shapes, 89 % of the sample, have a median length of 116 nm (i.e. 8.3 repeating units) thus slightly longer than with the faster cooling rate described above.

(vi) Very fast assembly of a DNA parallelogram with two different sticky ends

The constituted DNA parallelograms with different sticky ends were assembled at high rate through a cooling rate of 3°C/s. Curiously, a distribution of shapes and topologies very similar to the result of assembly of DNA parallelograms

with two sticky ends with the same stability and sequence is obtained. Branched and imperfect linear structures are obtained (Figure 10D). Very few cyclical taralli structures are obtained. An accurate quantitative description of this type of system is arduous: from the visual inspection of a few images, it appears that 47 % of the objects are rod-like, 3 % only are taralli, while 50 % of the objects are branched structures (comprising evidently imperfect rods).

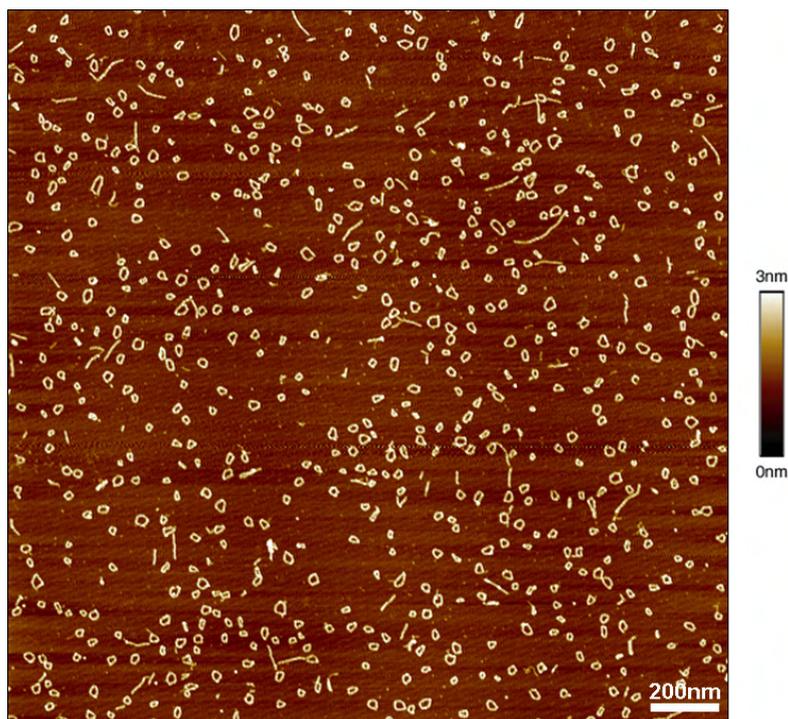


Figure 12 – AFM image of the parallelogram having two pairs of sticky ends with different thermal stabilities. Circularized polymers represent the 89% of the objects.

In a qualitative type of experiment, we treated the same specimen by keeping it at 37 °C for a few minutes. As evident from the inset of Figure 10D, the assemblies reorganize and the system then resembles the results of a slow assembly, with more perfectly assembled rods and a growing proportion of taralli, while branched structures tend to disappear. From the visual inspection of a few AFM images, 51 % of the images objects are rod-like and 37 % are taralli and only 12 % are left as branched shapes.

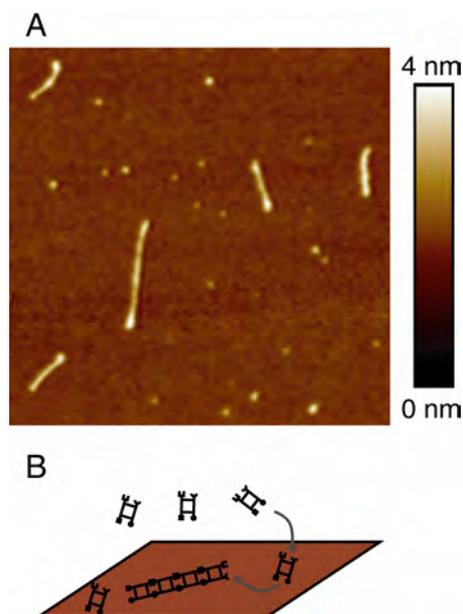


Figure 13 – (A) AFM image of the result of parallelogram polymerization on mica. A 1 nM solution of the tile with two different sticky ends in TAE/Mg²⁺ buffer was layered on a freshly cleaved mica surface at 20°C and incubated for 10 minutes. Interestingly, tile-to-tile assembly was never detected in solution at this concentration. As the deposition time increases, longer linear constructs are observed on the surface. Micrograph side is 1 μm. (B) Proposed scheme for the polymerization on the surface, showing monomers adsorbing on the mica surface and binding to each other by lateral diffusion.

(vii) Preliminary results on the assembly of parallelograms on a surface

A preliminary attempt at assembling the polymers directly on the surface of mica was performed by layering a solution of parallelograms at a concentration sub-critical for polymerization (1 nM, see above) at a temperature lower than the melting temperature of both sticky ends.

Over different intervals, the growth of polymers was quenched and characterized by AFM. The resulting structures are exclusively straight rods (no taralli or branched structures) of a size proportional, on average, to the allowed growth time. At any time, also individual monomeric objects are visible on the surface, together with longer polymeric objects (see Figure 13).

3.1.2.2 Discussion

The DNA rhombus or parallelogram motif has been described first by Seeman and coworkers and used for the creation of 1D and 2D arrays.^[1] It appears that this is not

one of the most efficient and treatable DNA structural motifs, and, especially, 1D arrays have seldom been obtained with high efficiency. In the last few years, this motif has been used less and less, in favour of several others that can yield the assembly of micrometer-long 1D and 2D objects.^[16, 22] Even with its inherent limitations, this tile, more than many others, allows a high degree of structural control, as with very little changes in the structure and base sequence it permits the assembly in 1D, 2D or branched topologies (and also in cyclical ones, as shown here). The concept of hierarchical assembly is also directly applicable.^[15]

From our reported data, it is evident that the assembly of DNA parallelogram macromonomers into supramolecular polymers is an efficient process, and while the size of the assembled objects does not compare to what recently achieved with helix bundles or various types of 2D systems, the conversion from single parallelograms to polymers seems complete under the investigated conditions. It is also apparent that the assembled structures are characterized by a high degree of structural rigidity, due to the designed mechanical coupling between all the possible deformations of the chain.

Even though the polymeric chain is highly nicked (as it is made of many separated oligonucleotides), it appears to have a persistence length in the order of at least several hundred nanometers, probably at least one order of magnitude higher than dsDNA. For even the longest imaged 1D polymers, it appears that the orientational correlation along the chains is not lost, and so the persistence length is higher than the chain contour length.

As data obtained in various conditions show, the self assembly mechanism is rather complex, possibly involving thermodynamic and kinetic effects, and it is possible to direct its results to the formation of linear or circular or branched structures only by the subtle change of such tested variables as the cooling rate, the oligonucleotide concentration, the insertion of point mutations in the sequence that forms the inter-tile connections.

(i) On the structural origin of rods and taralli

The design of the monomers implies that the maximization of the interactions among them affords a linear, rigid structure where each parallelogram is linked to

the successive by means of both of the sticky ends on one side. The occurrence of circular polymeric structures with straight sections (taralli) plausibly derives from the cyclization of a sufficiently long and flexible linear chain possibly formed by linking each parallelogram monomer only through one of the two sticky ends on each side (see Figure 14). This must happen before the second, unutilized couple of sticky ends binds, and so rigidifies the chain, impairing its circularization. Once a flexible chain of singly-connected tiles is circularized, the joining of a subset of all the free sticky ends is still possible, and it leads to straightened sections of rail-like structure within circular objects. The rigid rail-like domains are separated by more flexible and incompletely connected domains that are necessary for the chain to change direction in the circular shape.

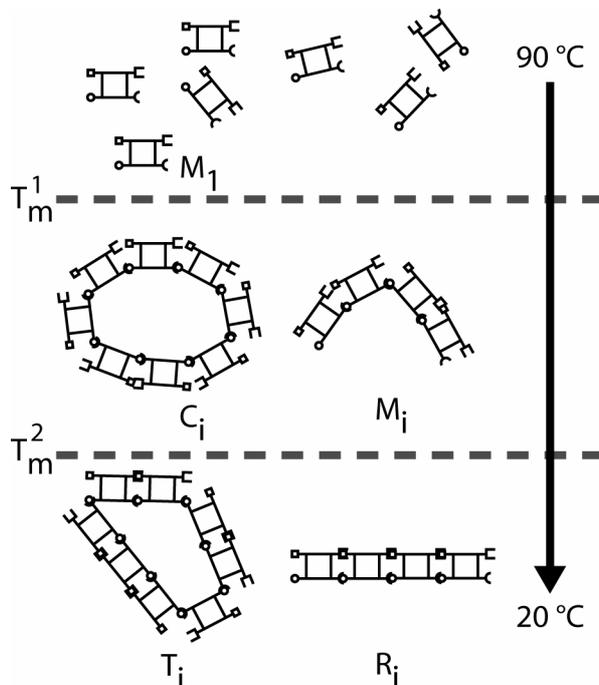


Figure 14 – Schematic representation of the polymerization of the parallelogram monomer with sticky ends having different melting temperatures, $T_m^1 > T_m^2$. During the annealing, the system spends a finite time during which only one of the connections can be stably joined. The only possible structures in this condition are flexible, singly-joined chains, which can remain linear (M_i) but will ultimately give rise to circular structures only (C_i). Upon approaching the second pair of sticky ends can join, transforming flexible linear chains M_i and flexible rings C_i into, respectively, rigid rods R_i and polygonal objects T_i . Since the amount of time elapsed at a temperature between T_m^1 and T_m^2 depends on the cooling rate, the occurrence of the different shapes in the final population of polymers observed at a temperature lower than T_m^2 is influenced strongly by the temperature program used to anneal the sample.

For the above to be a relevant phenomenon, there must be a finite time during the polymerization in which only one connection between each couple of tiles is set (and so the chain is flexible, see figure 14): this is certainly true when the two sticky ends on each side of a tile are different in sequence and so in thermal stability. In this case, the two melting temperatures of the two dsDNA sections formed are different: $T_m^1 > T_m^2$ and thus, during the cooling, a finite time is spent when one of the two sticky ends can be stably joined, while the other can not. For the same couple of sticky ends, the slower the cooling rate, the longer this interval and thus the more chains have a chance to cyclize. At $T_m^2 < T < T_m^1$ circularization is a favorable process for flexible chains, as it maximizes the number of base-pairs (see figure 15).

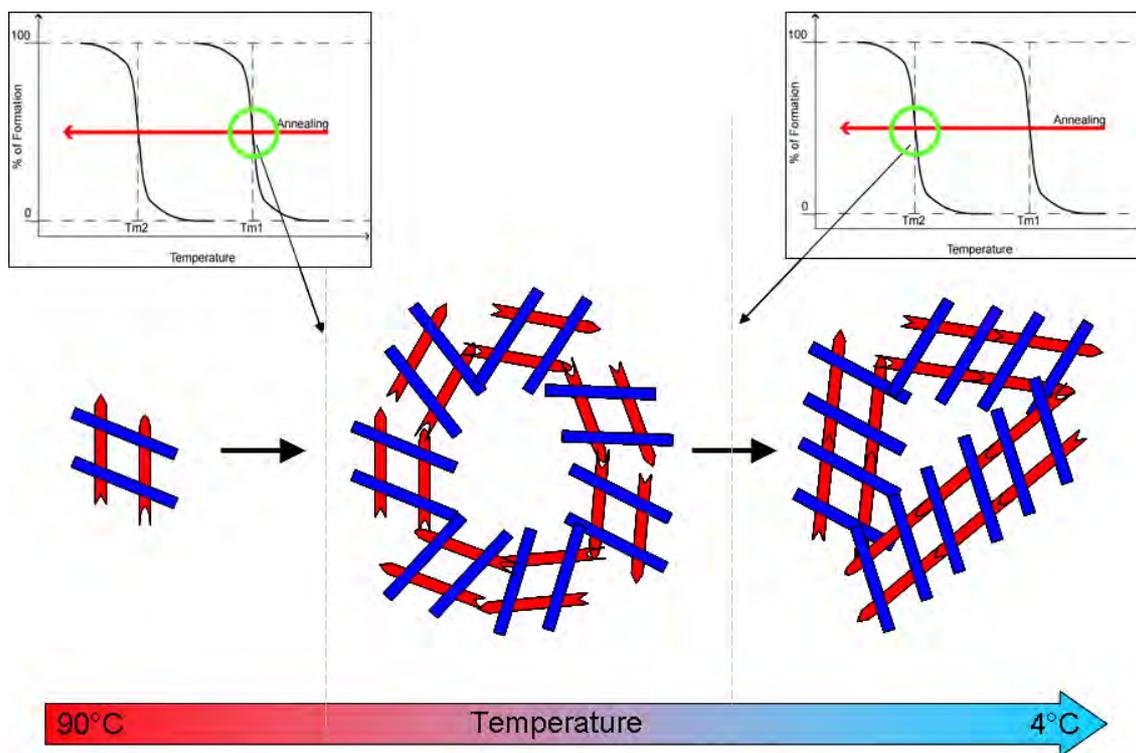


Figure 15 – The two pairs of sticky ends have different melting temperatures (the difference is about 2°C in the conditions used for the annealing). During the annealing, the two pairs of sticky ends are stabilized sequentially, so that when one interaction is stable the other is not. If the annealing is performed in quasi-equilibrium condition, all the growing shapes will have the time to maximize the number of interaction of the first type, and this means circularizing the flexible growing chains. When finally the second sticky end is stabilized, the circular shapes cannot be broken and the resulting objects are polygon where most of the second type, less stable sticky ends have joined. Increasing the annealing speed means giving the system less time to adapt to the intermediate region where only one sticky end can be formed and thus reduce the proportion of circular objects in the annealed sample.

When the temperature nears the lower melting temperature T_m^2 , the second type of connections becomes stable, leading to straight sections in the taralli and to rod-like linear chains, derived from the completion of the connections in the uncircularized fraction of the polymeric chains.

The qualitative results of a molecular-dynamics type of numerical simulation show the coexistence of linear chains and circular chains of different lengths when the simulation is run at a temperature where only one type of connection is stable (see figure 16A). When such molecular conformation is quenched to a temperature lower than both melting temperatures, then all chains tend to form extended portions of rail-like structures, leading to straight and rigid objects and to polygonal circular ones (see figure 16B), as observed in our experiments. The availability of a numerical simulation method can shed some light on the interplay of the different physical parameters in the quantitative determination of the polyparallelogram system (L. Rossi *et al.*, manuscript in preparation).

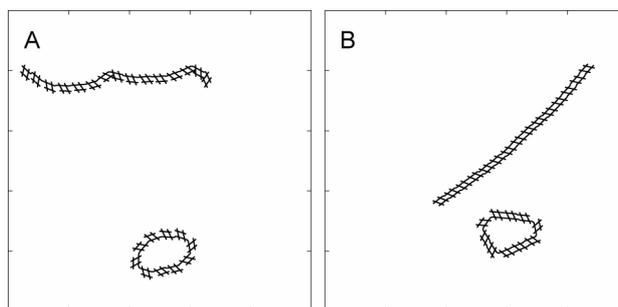


Figure 16 - Example snapshots of a molecular dynamics simulation of the parallelogram with sticky ends having different thermal stability. The side of the boxes is 350 nm. (A) Snapshot of a system composed by 25 parallelograms equilibrated at a temperature between T_m^1 and T_m^2 . (B) Snapshot of the same system after equilibration at a temperature lower than both T_m^1 and T_m^2 .

(ii) *The bifurcation between taralli and rods and factors influencing it*

As also derived on simpler polymeric systems,^[23] when the assembly is done through thermodynamic equilibrium states, only cyclical shapes should be obtained. Our experiments prove this assertion, as the slower the cooling rate, the higher the fraction of taralli with respect to the other accessible shapes (see Figure 10D-F).

A layout of the relevant concurrent reactions is proposed in figure 17. At a temperature higher than T_m^1 , only the single parallelogram tile (M_1) can be found in

solution, as no binding is thermodynamically accessible. At a temperature approaching T_m^1 , the chain growth $M_1 \rightarrow M_i \rightarrow M_n$ occurs together with the equilibrium cyclization $M_i \rightarrow C_i$, where C_i represents a flexible cyclical chain, the result of the cyclization of a linear flexible chain M_i made of i monomers. Successively, at a temperature around T_m^2 , each M_i can also rigidify to rod-like structures, R_i , while the C_i forms will turn into taralli shapes, described above (T_i). The thermodynamic constants of each equilibrium depend only of the temperature. Chain growth can only take place from linear chains (M_i or R_i).

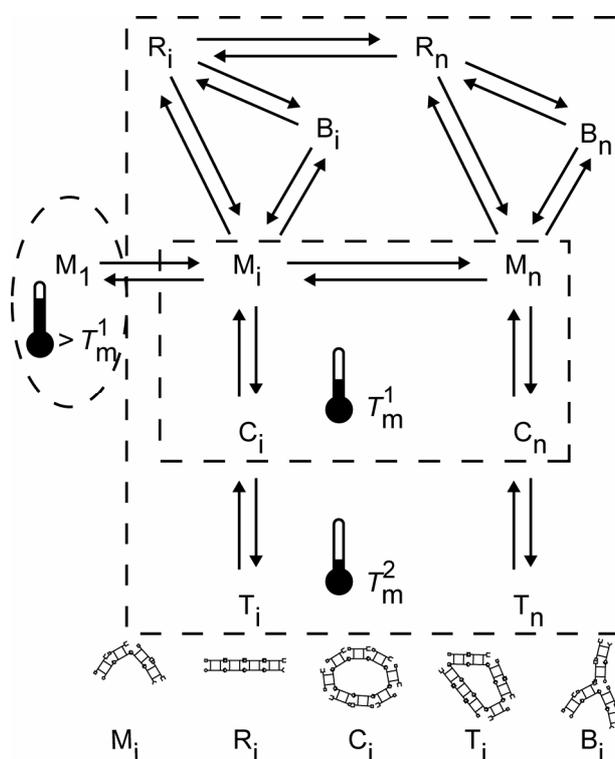


Figure 17 – Schematic outline of the polymerization reactions operating in solution. Dashed boxes include the sub-sets of reactions that are at the equilibrium at either T_m^1 or T_m^2 .

If the assembly is undertaken through quasi-equilibrium states, then chain growth will proceed to the highest molecular weights (M_n), and all linear chains will be converted to cyclical forms, as the temperature decreases, since the equilibria will be shifted in favor of the C_n shapes.

If, on the other hand, the assembly proceeds very fast, under kinetic control, irreversible ring closure can occur while chain size is still limited. As also the second type of connection becomes stable while the molecular weight of the chains

is still low, the $M_i \rightarrow R_i$ transformation prevents the formation of larger cyclical structures (C_n) and yields a larger fraction of rods, even longer ones.

Finally, if the assembly is done at intermediate cooling rates, the result depends on both kinetics and thermodynamics. In these conditions, a more varied distribution of types of polymers is found, and the size of rods and taralli is intermediate between the two extreme cases described above. Examples of these three cases are observed experimentally (compare Figures 10D-F and see the Results section for average sizes).

(iii) The emergence of branched structures

As of the reaction layout (figure 17), branched structures (B_i) can also arise. In order to obtain a branched structure in this system, it must occur that both connections are stable while in the presence of a high concentration of species that can bind to them. This can happen for parallelograms with two different sticky ends only if they are assembled under kinetic control. Under these conditions, mainly branched and straight linear forms will emerge, with a very small fraction of cyclical ones. This is experimentally witnessed in conditions of fast cooling (Figure 10D). Branched structures can, in principle, derive also from cyclical forms, even though such structures are hardly ever observed experimentally.

(iv) How the supramolecular system adapts to environmental conditions

All the described polymeric forms can be reset to the monomer state simply by heating the system to above both the T_m of the connections but below the temperature leading to disassembly of the parallelograms into oligonucleotides (in our case, 50°C was used). Moreover, as described in the Results section, the thermal equilibration of kinetically assembled poly parallelograms at a temperature intermediate between the two T_m , can lead to a distribution of forms similar to that obtained after an exclusively thermodynamically-controlled assembly (see figure 18). This system proves to be adaptive to the environment and its state can be interconverted repeatedly.

(v) Assembling polyparallelograms with sticky ends of the same thermal stability

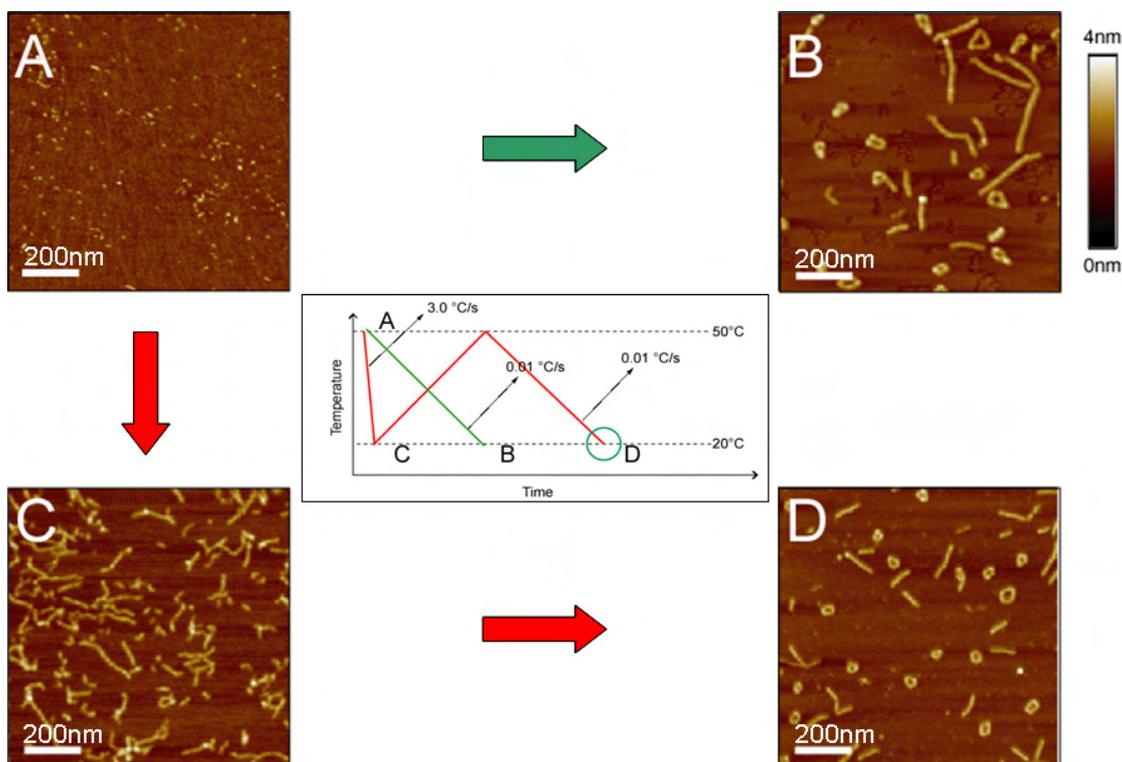


Figure 18 – Verifying the adaptivity of the system. The central panel shows two different traces (in red and green) of two different temperature programs to which a sample of the monomer was subjected. Different stations of the programs are evidenced with capital letters. The AFM images around the central panel are labeled with the same capital letters and are representative of samples taken in the stations identified by matching letters in the central panel. The arrows connecting AFM images have colors matching the traces in the central panel, showing which temperature program was used to obtain the subsequent samples. (Green Trace) Monomer at 50°C shown in AFM panel A. Cooling to polymer at 20°C with a negative slope of 0.01°C per second. Result of polymerization shown in AFM panel B. (Red Trace) Another aliquote of monomer at 50°C shown in AFM panel A. Annealing to polymer at 20°C with a cooling rate of 3.0°C per second. Result of polymerization shown in AFM panel C. Re-heating of the same aliquote to 50°C at random speed and then re-cooling to 20°C using the same slope that let from A to B in the green trace, that is, 0.01°C/s. Result of the new polymerization shown in D. Samples B and D were obtained with the same cooling rate, and statistics performed on the recorded images show that they have the same distribution of shapes.

All the above mentioned considerations withstanding, the entire reaction layout of figure 17 is accessible simultaneously for parallelograms with two sticky ends with the same melting temperature. Around T_m and below, the system will try to maximize all the possible interactions, leading to mainly branched and rigid linear structures. Only very small cyclical shapes are possible, as longer ones require a somewhat long singly connected chain to be present for a finite time. Due to this

behavior, in this system, there is a negligible difference among the assemblies obtained at slow or fast cooling rate. We confirmed this behavior on a qualitative basis (data not shown).

It is our understanding that in principle a reduction of the monomer concentration would allow the complete formation of the double connections among tiles and thus lead to a reduced branching. We thus assembled the parallelograms at a 10-fold reduced concentration, evidencing no significant improvement in the structural regularity. A further 10-fold reduction yielded no assembly in the time scale of an experiment. In our hands, this strategy towards a better structural control still lacks an experimental proof.

(vi) When assembly takes place in a system with reduced dimensionality

As the dimensionality of the environment of a system is reduced from 3 to 2 dimensions, the degrees of freedom of motion of its components are reduced, while their effective concentration is increased.^[24] If the encounter of two objects takes place through 2D diffusion from adsorbed states, numerous chain conformations are forbidden. As the parallelogram monomers and the perfectly assembled rail-like 1D chains are essentially flat objects, it is conceivable that they will be adsorbed in a condition propitious to their proper assembly.

As preliminary evidence, we layered a diluted solution of individual parallelograms on the surface of freshly cleaved mica kept at a temperature lower than the T_m of both sticky ends of the parallelograms. The concentration of monomers is such that no solution assembly is witnessed in the time scale of our longest experiments. The experimental evidence is that the growth of 1D rigid unbranched chains takes place with a good efficiency (see figure 13A), and chain length is proportional to the incubation time. No cyclical taralli are seen in this case, while monomers are still found on the surface at any stage of growth, as these continuously adsorb on the surface during chain growth (see figure 13B for a possible route to surface growth). Further experiments will be necessary to better characterize the growth on the surface.

(vii) How to drive the system towards a higher level of structural control

On the basis of the discussed experimental evidence and the hypothesized reaction system, we can propose guidelines towards a better control of the assembly of these supramolecular chains. A system where T_m^1 and T_m^2 define two well-separated stability domains appears far more controllable than one where only one T_m regulates the assembly of both connections. In this case, the relative abundance of rods and taralli can be determined by a fast or slow cooling rate.

This system also proves to be adaptive, as expected from supramolecular polymers:^[18] selective depolymerization of branched forms can be achieved, for instance by annealing at a temperature intermediate between T_m^1 and T_m^2 . By heating at around T_m^1 , the system is effectively reset and ready for a novel polymerization.

The control of the dimensionality of the assembly medium proves as an additional tool towards a finer level of structural control, as the growth of certain types of structures can be completely inhibited. This tool has never been used before in the context of DNA structural nanotechnology. Our polymer could provide a DNA implementation towards the type of constitutionally dynamic materials proposed by Jean-Marie Lehn.

3.1.2.3 Experimental

(i) Sequence design

All the sequences of the oligonucleotides used in this work were designed by applying the sequence symmetry minimization feature of the program NANEV 1.1^[10] run with a SSM search size of 4 and 1000 generations. The resulting sequences are shown in table 1.

Oligonucleotides were purchased from MWG (Ebersberg, Germany), HPLC purified and lyophilized by the supplier, and suspended in MilliQ H₂O (Millipore Simplicity) prior to use. The concentration of each strand was estimated by UV spectroscopy by measuring OD₂₆₀.

(ii) Nanostructure assembly

Parallelogram “Beta_1” - with two different sticky ends	
PAR_Beta1_1	<u>AAGGGA</u> AAACGCCTGCGATCCCCAGCCAATGCCCTGTCTGGG
PAR_Beta1_2	GGCAGAGAGTGGTAGCTGACATGAGTGAACACCTGACGCTCGGTGAGACCGTGGCGTTT
PAR_Beta1_3	TTTCCCACCACTCTTGCCATGCGGTAGTGGGCATTGGCTGGGGATCGCACCATCTGATC
PAR_Beta1_4	GGGTTCTGTTCACCTCATGTGCTGAGCTACCTGGGAAA <u>AGAGGGA</u>
PAR_Beta1_5	GATCAGATGGACGGTCTCACCGAGCGTCAGGACAGCTTTTGCTGTGGAACCC <u>TCCCTCT</u>
PAR_Beta1_6	<u>TCTCCTT</u> CCCGACACCCATTTTTATGGGACTACCGCATGGCAAGACTGGACTCTCTGCC
Parallelogram “Beta_1_Bis” - with sticky ends having the same base composition.	
PAR_Beta1_bis_1	<u>AGAGGGA</u> AAACGCCTGCGATCCCCAGCCAATGCCCTGTCTGGG
PAR_Beta1_bis_6	<u>TCCCTCT</u> CCCGACACCCATTTTTATGGGACTACCGCATGGCAAGACTGGACTCTCTGCC

Table 1 – Sequences used in the experiments described in this section. The parallelogram with two sets of sticky ends having the same melting temperature only differ in two of the six strands constituting it, the others are identical to the other parallelogram. Sequences are written from 5’ to 3’. Emphasized bases are those participating in sticky ends.

DNA parallelograms were assembled in solution by mixing stoichiometric quantities of each component strand in TAE/Mg²⁺ buffer (20 mM Tris, 2 mM EDTA, 12.5 mM MgCl₂, pH 8.4). The final concentration of each strand was 0.1 μM. The mixtures were then heated to 90 °C for 5 minutes and cooled to 20 °C following one of the protocols described below. The slowest, ‘quasi-equilibrium’ annealing was performed by placing the mixtures in a 2 l water bath at 90 °C in a styrofoam box and left to cool to room temperature (around 20 °C) over approximately 72 hours, resulting in a cooling rate of ≈ 0.0003 °C/s. Other annealings were performed in a PCR thermocycler (PCR Sprint, Thermo Electron Corp., Waltham, MA, U.S.A.) by cooling the mixtures at a rate of 0.01 °C/s from 90 °C to 50 °C to ensure correct parallelogram formation. Then the cooling to 20 °C

was completed at a rate of either 3.0 °C/s (fastest, ‘kinetic’ annealing) or 0.01 °C/s (intermediate case). When needed, the concentration of the monomer was changed only after performing the annealing from 90 °C to 50 °C at the usual strand concentration of 0.1 μM to ensure complete monomer formation.

Assemblies on the surface are performed after complete formation of the monomer at 0.1 μM in TAE/Mg²⁺. The monomer solution is then diluted to 1 nM in the same buffer and equilibrated at 20 °C for a few hours. Polymerization is performed by depositing 10 μl of this solution on a piece of freshly cleaved mica and leaving it at 20 °C for at least 10 minutes.

(iii) AFM imaging

Atomic Force Microscopy imaging was performed in tapping mode with PointProbe noncontact silicon probes (NanoSensors, Wetzlar-Blankenfeld, Germany) on a NanoScope IIIa SFM system equipped with a Multimode head and a type E piezoelectric scanner (Veeco, Santa Barbara, CA, U.S.A.). Raw SFM images have been processed only for background removal (flattening) by using the microscope manufacturer’s image-processing software. DNA molecule profiles have been measured from the SFM images with the software package ALEX.^[25] Annealing mixtures were deposited on freshly cleaved mica (Ruby Red Mica Sheets, Electron Microscopy Sciences, Fort Washington, USA) and left to adsorb for 2 min at room temperature (\approx 20 °C). The mica surface was then rinsed with \approx 500 μl of MilliQ H₂O (Millipore Simplicity) at the same temperature and dried with dry nitrogen.

(iv) Molecular Dynamics simulations of parallelogram polymerization

The basic elements used to perform the molecular dynamics simulations are mass points which interact by means of a shielded Coulomb potential and are confined within a box. Four of these “particles” are rigidly linked together to form the DNA parallelogram units. Additional particles represent the heat reservoir used

to set the temperature of the system. These are subjected to a type of dynamics described by the Langevin equation. Each of the four ends of a parallelogram unit can bind to the proper end of other units, with a probability that depends on the temperature. When two units bind together, restoring forces tend to optimize their conformation. It is possible to change the temperature during the simulation in order to investigate the emergence of different forms when the system is equilibrated above or below the melting temperature of one or both the sticky ends. Due to the simplifications introduced in the physical description of the system, it is possible to conveniently run simulations also on a personal computer.

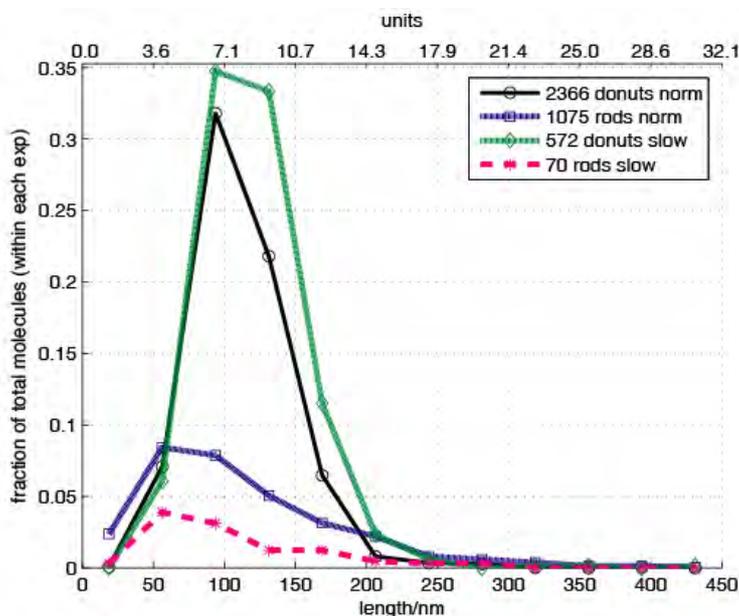


Figure 19 – Quantitative results of the polymerization at different speeds (see section 3.1.2.1)

3.1.2.4 Conclusions

In this experimental effort, we showed that a good degree of control can be achieved in the assembly of a 1D supramolecular polymer made of rigid DNA tiles. Very rigid rod-like structures or circular ones can be obtained from one tile only thanks to the comprehension of the inner workings of the self-assembly of this system.

As the tile design for the assembly of this 1D polymer enables further controlled branching or decoration with functional moieties, we believe that our results could prove useful towards the realization of functional nanoscale materials.

3.1.3 Assembly of the 1D array on a ssDNA template

3.1.3.1 The concept

As discussed in the previous section, the self-assembly of parallelogram tiles into an 1D continuous array can afford long nanoscale objects, entirely made of DNA, having a persistence length surpassing by far that of dsDNA. Such objects could be employed in strategies so far tested on linear dsDNA such as the metallization strategy reported by Keren et al detailed in section 1.3 and serve as scaffold for the positioning nanoscale components. In this context, the most important characteristic for DNA scaffolds is their rigidity, and parallelogram 1D rails do have this feature. However, other very useful features for such a scaffold is its length and stability. The longer the scaffold is, the easier is to locate it with SPM techniques and perhaps to interface it with microscopic objects such as microelectrodes and so on. Thermal stability is also very important to ensure that the scaffold can withstand all the processes of deposition, metallization, contacting without losing its structural integrity.

Following the reasoning above, the 1D linear railroads described in section 3.1.2 are not a particularly good choice, since although their rigidity when perfectly assembled is very high, their length is largely dependent on the experimental conditions. This is due to the fact that for the array to be continuously linear, without breaks or nicks, each and every pair of sticky ends must be perfectly joined throughout the whole structure. The longest 1D linear rail formed with this system ever observed was around 650 nm long, but the typical linear rod was around 90 nm long (see figure 19). This is unfortunately far from enough for using the DNA railroad as a scaffold for metallization.

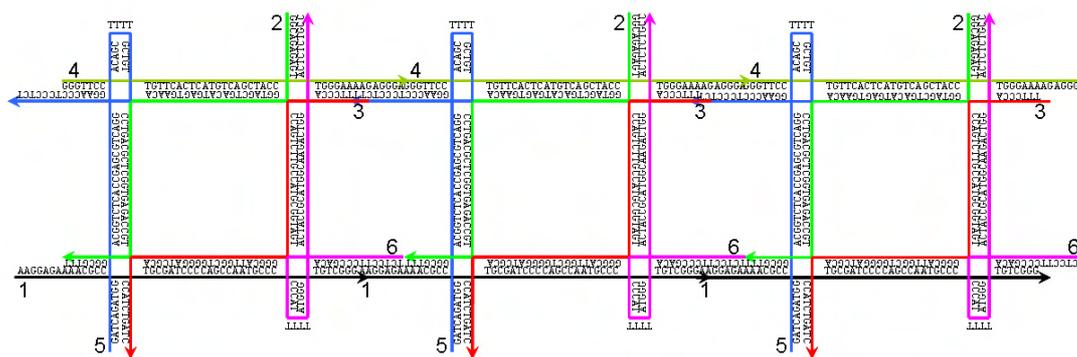


Figure 20 – Schematic representation of the 1D railroad described in section 3.1.2. When two tiles assemble, the 3' and 5' termini of oligo 4 are in contact with each other. This is also true for the termini of oligo 1. Both oligonucleotides could be substituted by a long ssDNA continuous template.

To solve this problem, we decided to follow the approach of templated polymerization. As shown in figure 20, two oligonucleotides in the parallelogram (oligo 1 and 4) do not directly participate in crossovers. They start and finish within a single linear DNA helix, while their respective complementary strands change at each junction position. They also participate in the two pairs of complementary sticky ends on two opposite sides of the parallelogram.

When the sticky ends are stabilized, parallelograms join to form the linear polymer, and two linear double helices are formed in the direction of polymerization propagation. These helices are highly nicked (at the sticky ends), but otherwise they are completely comparable to continuous dsDNA. One strand in these helices is constituted by multiple repetitions of oligo 4 or oligo 1, while the strand with opposite polarity is constituted by fragments of the other oligo. At the sticky ends, the 5' and 3' termini of oligo 4 and oligo 1 are in close proximity. In fact, their position is the same they would be in if they were part of a continuous long ssDNA strand; only the nick separates them.

If the 1D array of parallelograms was ligated at the sticky ends and then denatured, one would find among the products of ligation two extremely long ssDNA strands, with sequences formed by cyclic repetition of oligo 1 for one strand, and oligo 4 for the other.

Actually ligating together these parallelograms is not a practicable option. Each sticky end is flanked by two junctions separated by only 7 nm, and ligases would work at extremely low yields in these conditions. Moreover, as discussed previously, purifying assembled DNA structures from enzymatic reaction mixtures is not usually trivial. A much better option is to pre-synthesize a long, ssDNA strand with a repetitive

sequence of either oligo 1 or 4, and then assemble the linear 1D array on this template. Incorporating the continuous template into the construct would mean that the structural integrity of the linear 1D array would be guaranteed by the template itself, and not by the simultaneous stability of all the sticky ends in the construct. This would lead to the formation of longer, rigid and more stable 1D arrays.

3.1.3.2 Synthetic strategy

The ideal tool to produce long ssDNA strands with a repetitive sequence is the Φ 29 polymerase enzyme. Φ 29 is the replicative polymerase from the *Bacillus subtilis* phage Φ 29.^[26] This isothermal polymerase has the highest processivity and strand displacement activity among known DNA polymerases^[27] as well as an inherent 3'→5' proofreading exonuclease activity. It is not uncommon to obtain ssDNA with a length of several tens of kilobases from a typical Φ 29 reaction in optimal conditions. The amount of produced ssDNA is also extremely high; an amplification of approximately 10,000 times the mass of starting template is obtained in a few hours.

The high strand displacement capabilities of the Φ 29 polymerase means that if this enzyme binds to a circular template, it will be capable of displace the newly synthesized strand once it reaches the point at which the polymerization started. The enzyme will thus not stop the reaction but continue revolving around the circular template, generating a long ssDNA strand with a sequence constituted by the constant repetition of a sequence which is complementary to the template. This process is usually named Rolling Circle Amplification or RCA.

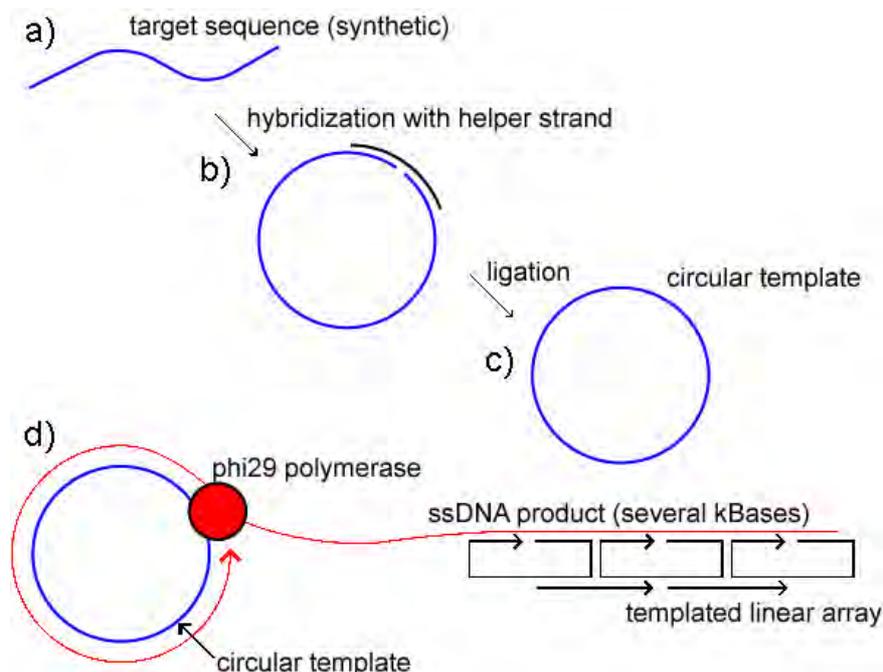


Figure 21 – Strategy for the production of a templated DNA one-dimensional array. (a) Synthesis of a 5'-phosphorylated oligonucleotide with a sequence complementary to the one that needs to be included in the array. In our case, the sequence is the complementary to oligo 4 in the parallelogram (oligo anti-4). (b) Oligo anti-4 is hybridized with a short helper strand, capable of keeping its two termini in contact. Only the construct where one anti-4 is circularized is shown in the figure, but the formation of constructs containing more than one oligo anti-4 and one helper strand is definitely possible and does not prejudice the success of the experiment. (c) Ligation with a DNA ligase yields circularized oligonucleotides, the shorter of which is just one circularized copy of oligo anti-4, but higher homologues containing multiple occurrences of oligo anti-4 can be generated. (d) Rolling Circle Amplification of the circularized templates. The product is a several kilobases long ssDNA strand constituted by the repetition of oligo 4 sequence. The DNA 1D array can then be assembled directly on the amplified strand.

An oligonucleotide containing several hundred copies of a designed base sequence can thus be produced by rolling circle replication of a circular synthetic oligonucleotide.^[28] An array assembled on this periodic template can extend up to the length of the template (see figure 21).

For testing this approach, we choose to synthesize the complementary oligonucleotide to oligo 4 in the parallelogram (oligo anti-4), circularize it and amplify the oligo 4 sequence by RCA.

3.1.3.3 Production of the circular template for RCA

Two oligonucleotides, with the sequences reported in table 2, were synthesized and HPLC purified. The first of them (oligo RCA_Anti-4) is the complementary to oligo 4 in the parallelogram (see figure 9). The second one (oligo RCA_Stitch) is a

short helper strand, which is complementary to the first eight and the last eight bases of oligo RCA_Anti-4. When hybridized together, oligo RCA_Stitch will bind to one tail and one head of oligo RCA_Anti-4, regardless whether or not they belong to the same oligo molecule or to two. In every case, a 5' and a 3' terminus of oligo RCA_Anti-4 will be kept together and ready to the successive ligation step.

RCA_Anti-4	TCCCTCTTTTCCCAGGTAGCTGACATGAGTGAACAGGAACCC (5' -phosphorylated)
RCA_Stitch	AAGAGGGAGGGTTCCT

Table 2 – Oligos used in the experiments described in this section. Sequences are written from 5' to 3'.

The two oligonucleotides were mixed in 1xTAE buffer at a final concentration of 10 μ M of each, and annealed to 90°C to 20°C in 3h to maximize adduct formation. Then 10x concentrated T4 Ligase reaction buffer (50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM Dithiothreitol, 25 μ g/ml BSA, pH 7.5 at 25°C) was added to a final concentration of 1x. Commercial T4 ligase was added (400 units for 40 picomoles of each oligo), then the mixture was incubated at 16°C overnight (approximately 15h). After the incubation, the sample was heat-inactivated at 65°C for 10 minutes. The result of the ligation was extracted twice with a (25:24:1) mixture of phenol : chloroform : isoamyl alcohol saturated with TE buffer at pH 7.8, then once with pure chloroform. The reunited aqueous phases were then purified on a Millipore microcon YM-3 centrifugal filter unit. An aliquote of the mixture was then resuspended in 1xTAE and analyzed by PAGE electrophoresis.

A representative gel is shown in figure 22 (refer to the figure caption for the experimental details). The first lane on the left is a pUC19 DNA/MspI (HpaII) marker. The second lane shows the result of the ligation. The stronger band, running as a linear dsDNA of approximately 50 bases, is the circularized RCA_Anti-4 monomer plus one hybridized RCA_Stitch. Heavier bands of decreasing intensity are the same type of

construct but with more than one copy of RCA_Anti-4. Oligo RCA_Anti-4 and RCA_Stitch were loaded in the rightmost lanes, in this order from left to right.

After desalting, it is not necessary to further purify the result prior to the RCA step. The oligo RCA_Stitch doubles its function as a primer for the Φ 29 enzyme in the successive reaction step. It is worth mentioning that the most abundant product of the ligation is the circularized monomer, which is only 42 nt long. It is possible that the short length of this template could hinder the Φ 29 amplification, to a limited extent. However, higher weight circularized objects are present in the template in significant amount and their length is more than adequate for optimal Φ 29 performance.

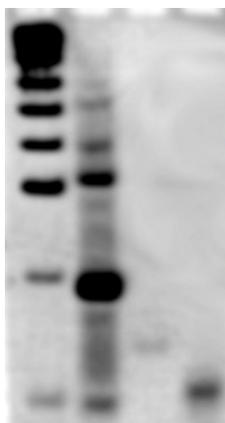


Figure 22 – PAGE electrophoresis of the result of the ligation. 10% PAGE (19:1 acrylamide : bisacrylamide) gel run in 1xTAE at 10V/cm for 2h in cold room at 4°C. The gel was stained with ethidium bromide for 10 minutes, then de-stained for 2 minutes in mQ water. The image was recorded with a digital camera, while illuminating the gel with UV light.

Please refer to the main text for a detailed description.

3.1.3.4 Rolling circle amplification of the template

The collection of circularized templates was resuspended in 1x Φ 29 reaction buffer (50 mM Tris-HCl, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM MgCl_2 , 4 mM Dithiothreitol, pH 7.5 at 25°C, supplemented with 200 $\mu\text{g}/\text{ml}$ Bovine Serum Albumin and 200 μM dNTPs). New England Biolabs Φ 29 polymerase (100 units) was added and the reaction mixture was incubated at 30°C overnight (approximately 15h). The enzyme was heat-inactivated with an incubation at 70°C for 10 minutes. The mixture was then purified in the same way described above at point (iii) for the purification after the ligation step.

The result of the Rolling Circle Amplification step were visualized with denaturing agarose gel electrophoresis (figure 23, the caption contains the experimental details). The rightmost lane of the gel contains a commercial (Fermentas) Lambda DNA/EcoRI+HindIII marker. The two central lanes contain two different aliquotes of the RCA polymerization experiment. The leftmost lane contains an aliquote identical to the previous ones but briefly (1 min) sonicated in water at room temperature prior to loading it on the gel.

As evidenced by the gel, the RCA polymerization produces a huge variety of ssDNA products (note the smear of the two central lanes, indicating the presence of products of highly varied molecular weight). Most of the products appear to be unable to enter this 1% agarose gel, as evidenced by the distinct bands at the wells. This means that the majority of the reaction products are extremely long ssDNA strands with lengths of several tens of kilobases. The brief sonication shatters the long ssDNA strands to shorter fragments, as evidenced by the fast moving band in the leftmost lane.

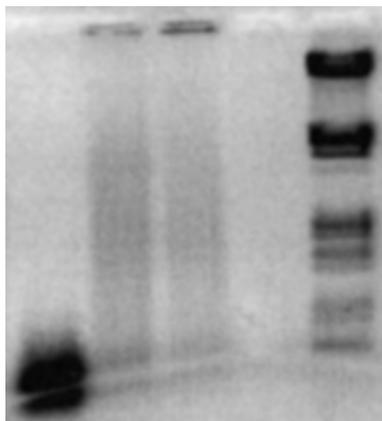


Figure 23 – Denaturing agarose gel electrophoretic run of the RCA reaction products. 1.2% agarose gel in water, soaked in 50 mM NaOH, 1 mM EDTA for 1h prior to the electrophoretic run. Samples were pre-incubated in the same buffer, with 10% Ficoll added. Bromocresol green was used as a tracking dye and a glass plate was kept on top of the gel while running to prevent its diffusion. Run performed at 30V for 6h, keeping the gel at room temperature. After the run the gel was equilibrated with 1 x TE buffer for 15 minutes, stained with ethidium bromide, illuminated with UV light then digitalized with a digital camera. Refer to the main text for the interpretation of the results.

3.1.3.5 Verifying the sequence of the RCA ssDNA product

To test whether the RCA product contained the required Par_Beta1_4 repeated sequence, an AFM imaging assay was performed. The RCA reaction mixture was

purified as detailed above, then diluted 40 times in 1x TAE/Mg²⁺ buffer. The diluted mixture was deposited on freshly cleaved mica at room temperature for 3 minutes, briefly rinsed with mQ water, dried under a nitrogen flow then visualized with tapping mode AFM in air.

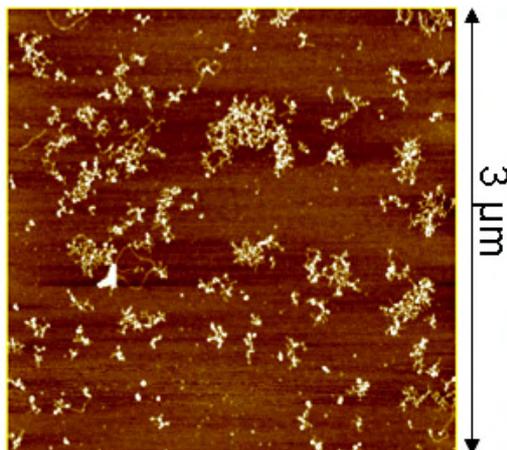


Figure 24 – AFM image of the RCA product. Image taken in tapping mode conducted in air, with a scan rate of 2 Hz. Refer to the main text for a discussion of the result.

Figure 24 shows a representative AFM image of the result. Several large agglomerates are visible, along with a few scattered smaller objects. The larger objects have the distinctive appearance of long ssDNA strands coiled on the surface. The size of these agglomerates seems to be highly varied. The long smears visible in the gel in figure 23 corroborates this observation.

Another aliquote of the purified RCA reaction mixture was added with an excess of oligo RCA_Anti-4 in 1xTAE buffer and annealed from 90°C to 20°C in one hour. The annealed mix was then diluted and AFM-imaged as detailed above. A representative AFM micrograph is shown in figure 25. The overall appearance of the sample is noticeably changed with respect to what it was prior to the addition of oligo RCA_Anti-4. Most objects in the images recorded have the characteristic appearance of dsDNA, even if some portions have still look like ssDNA. This observation confirms that the added oligo and the product of the RCA amplification have complementary sequences and form an highly nicked dsDNA strand upon hybridization. The length of the observed adducts is in some cases in the order of tens of microns.

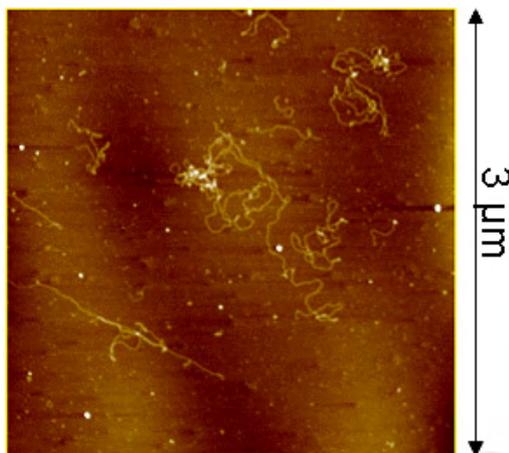


Figure 25 - AFM image of the RCA product after hybridization with oligo RCA_Anti-4. Image taken in tapping mode conducted in air, with a scan rate of 2 Hz. Please refer to the main text for a discussion of the result.

3.1.3.6 Assembling the 1D parallelogram array on the template

All the oligonucleotides composing parallelogram Beta1 (see section 3.1.2.3) with the exception of oligo PAR_Beta1_4 (the sequence of which is repeated multiple times in the ssDNA template produced by RCA) were mixed together in equimolar amounts in 1xTAE/Mg²⁺ buffer. An excess of this mixture was added to the ssDNA template solution and annealed from 90°C to 20°C at a cooling rate of 0.01°C/s. The annealed mixture was then diluted, deposited on mica and visualized by AFM imaging following the same procedure described above. A representative AFM image of the result is shown in figure 26.

Also in this case the overall appearance of the sample changed significantly from what was presented in figure 24 (RCA-produced ssDNA template alone). A lot of small fragments are visible scattered around which constitute the most abundant type of object in the sample. These are probably the excess of incomplete parallelogram that were annealed together with the long template and could not bind to it. These incomplete Beta1 parallelograms, although lacking oligo PAR_Beta1_4 and thus any structural rigidity, still have one functional sticky end (refer to figure 9). This means that they can coalesce, to a limited extent, into small amorphous aggregates, some of which are visible in figure 24. The second type of object present in the images are large linear or branched structures reaching lengths of several micrometers. All evidences point to the conclusion that these structures are indeed 1D arrays of Beta1-type

parallelograms formed directly on the ssDNA long template produced by RCA. It is worth mentioning that in this experiment no single oligonucleotide with a sequence complementary to the template was added to the annealing mixture, but the ssDNA compact aggregates visible in figure 24 have been nonetheless uncoiled and stretched to an appearance similar to that seen in figure 25. The only means available to make this happen in these conditions is the hybridization of all the constituent components of the parallelogram to the ssDNA template.

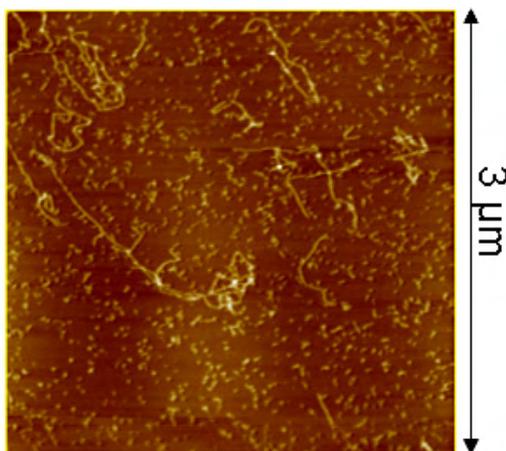


Figure 26 - AFM image of the RCA product after annealing with all the oligonucleotides composing the Beta1 parallelogram (see section 3.1.2.3) except for oligo PAR_Beta1_4, the sequence of which is included in the long ssDNA product template produced by RCA. Image taken in tapping mode conducted in air, with a scan rate of 2 Hz. Refer to the main text for a discussion of the result.

3.1.3.7 Conclusions

The experiments described in this section prove that the assembly of DNA rhombus motifs on a long linear ssDNA template constituting one of their edges is a viable synthetic route for building 1D arrays. The individual arrays produced in this way are longer, stiffer and more thermally stable than those described in section 3.1.2 due to the presence of the template linking all the tiles together.

Several steps of the protocol need to be improved, in particular the step of annealing the PAR_Beta1 oligonucleotides to the template. Having an excess of incomplete structures not bound to the template is highly undesirable in the context in which we planned to use these structures (scaffolding nanoelectronic components). A solution to this could be a solid-phase supported synthesis approach, in which a

biotinylated ssDNA template is bound to magnetic beads and hybridized with an excess of each PAR_Beta1 oligo in sequence. The oligonucleotides directly binding to the template would be added first, then those closing the rhombus on the opposite edge. Each successive hybridization step would be followed by removal of the excess of unhybridized oligonucleotides, affording a cleaner final product.

3.1.4 Testing the mechanical rigidity of a DNA parallelogram

3.1.4.1 The concept

As also discussed in section 3.1, an isolated Holliday Junction is a quite flexible object. The minor angle between its two helical domains oscillates around approximately 60° but this value cannot be relied upon for any geometrically-deterministic nanoscale construction since the junction itself does not have an high enough rigidity.

Even though DNA parallelograms like those described in the previous sections are simply the juxtaposition of four Holliday junctions, they do have an high enough rigidity to be used for geometrically-controlled assemblies. This extra rigidity is probably an effect of the mechanical coupling of all the junctions constituting the parallelogram. At the length scale of these parallelograms, DNA is a very rigid object (as also discussed in section 1.3.1.5, dsDNA has a persistence length of 50 nm, which is extremely rigid when observed at the scale of few nanometers, just as steel is rigid when observed at the scale of a few meters). Forces exerted on an individual junction in the parallelogram are dispersed through the rigid dsDNA struts connecting them to the remainder of the structure. This means that to distort the angle of a single junction in the parallelogram, one must impart enough force to distort all four of them at the same time.

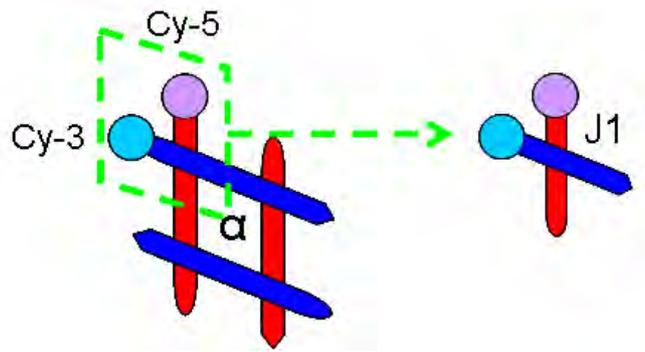


Figure 27 – Cartoon of the two structures used for the experiments reported in this section: the PAR_Alpha parallelogram and the J1_Alpha immobile Holliday junction.

To experimentally substantiate this hypothesis, we designed two different structures (shown in figure 27). The structure shown on the left is a complete DNA rhombus motif structure (called from now on “PAR_Alpha”) with edges of two full helical turns of dsDNA. The other structure (called from now on “J1_Alpha”) is one of the four Holliday junctions excerpted from the parallelogram, “cut” at exactly eight bases from the junction in each direction. The sequences of both structures are shown in table 3 and figure 28.

Since the sequences of the isolated junction are exactly the same present within the parallelogram around one of the junction (see figure 28), any difference in mechanical behavior between the junction and the parallelogram at that position must be ascribed solely to the mechanical coupling of the junction with the other ones in the parallelogram, or the absence thereof in the junction.

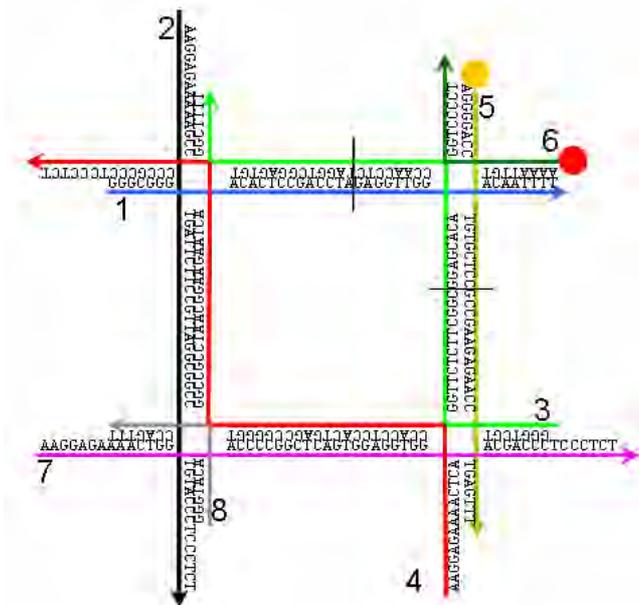


Figure 28 – Schematic representation of parallelogram “PAR_Alpha” and junction “J1_Alpha”, with sequences. Arrowheads represent 3’ termini. Red and yellow dots represent Cy5 and Cy3 dyes, respectively. The black lines show where the parallelogram is “cropped” to obtain the isolated junction.

Two oligonucleotides participating in the formation of the junction (both in the parallelogram and the isolated J1) were 5’-labeled with two organic dyes capable of performing Fluorescence Resonance Energy Transfer (FRET), an energy transfer mechanism between two fluorescent molecules. A fluorescent molecule (the donor) is excited by irradiation at a wavelength comprised in its excitation spectrum. By a long-range dipole-dipole coupling mechanism, this excited state is then non-radiatively transferred to a second molecule, (the acceptor). By doing this, the donor returns to the electronic ground state. In our case, Cy3 is the donor and Cy5 is the acceptor dye (see figure 29). The overall FRET efficiency (E) is determined by three parameters:

(A) The distance (r) between donor and acceptor. E decreases with an inverse sixth power law as r increases:

$$E = \frac{1}{1 + (r/R_0)^6}$$

where R_0 is a characteristic value of each (donor plus acceptor) pair called ‘Förster radius’ and describes the distance at which E is 0.5.

(B) The integral of spectral overlap between the donor emission spectrum and the acceptor absorption spectrum.

(C) The relative orientation between the donor emission dipole moment and the acceptor absorption dipole moment.

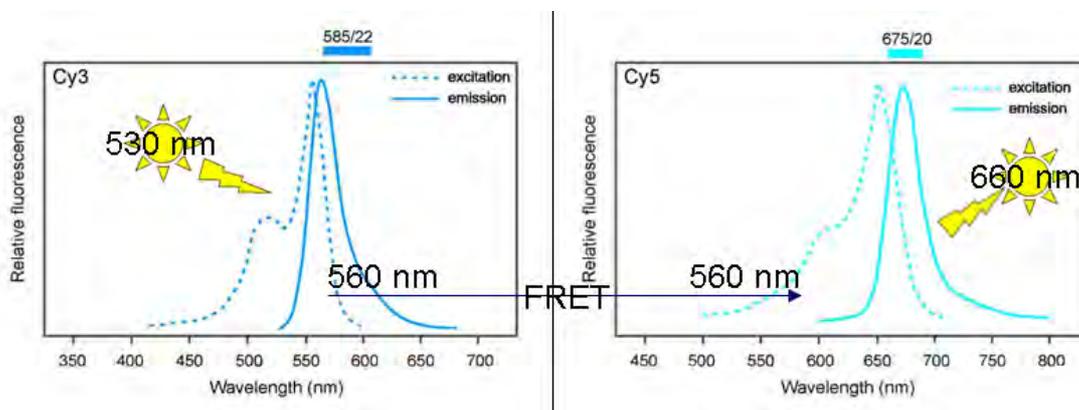


Figure 29 – Absorption and emission spectra of Cy3 and Cy5. The emission spectrum of Cy3 overlaps that of Cy5 and thus, if the two dyes are at an appropriate distance and their dipole moments are not blocked in an incompatible relative orientation, they can perform FRET (see above).

In our case, the two dyes meet all the required conditions: first of all they have a significant spectral overlap. Although they are not perfectly free to rotate in space, the relative orientations of their dipole moments varies constantly since they are tethered to the structure with quite flexible linkers (C_6 linkers). The structures to which they are tethered keep them at a distance of a few nanometers.

The key point is that the immediate surroundings of the dyes is identical in both the parallelogram and the isolated J1 junction. For example, these dyes are known to interact with DNA bases in not entirely predictable ways that could affect their performance. But in our structures, all the sequences that could reasonably affect the behavior of the FRET couple are identical. The linkers used to anchor the dyes to the DNA are also identical.

The average distance between the dyes should also be identical in the two structures. The only difference affecting the FRET couple tethered to the structures should be the different flexibility of the two junctions: the isolated one should be more flexible, the one inside the parallelogram should be more rigid. This means that although the average distance between dyes is the same in the two structures, this value could vary more around the average value in the junction, and comparatively less in the parallelogram. Since the FRET efficiency E varies with an inverse sixth power law on

the distance between dyes, the reasoning above would mean that E should be higher in the parallelogram structure than in the isolated junction in the same conditions.

Parallelogram “PAR_Alpha”	
PAR_Alpha_1	GGGCGGGACACTCCGACCTAG <u>GAGGTTGGACAATTT</u>
PAR_Alpha_2	AAAAAGCCTGATTCTTCCGTTAGGCGGCTGTACCCTCCCTCT
PAR_Alpha_3	GGGTCGTGGTTCTCTTCGGCG <u>GGAGCACACCAACCTC</u> TAGGTCGGAGTGTGGCTTTT
PAR_Alpha_4	AAGGAGAAAACCTACCACCTCCACTGAGCCGGGTGGCCGCTAACGGAAGAATCACCCG CCCTCCCTCT
PAR_Alpha_5	(5' Cy5) - <u>AGGGGACCTGTGCTCC</u> GCCGAAGAGAACCTGAGTTT
PAR/J1_Alpha_6	(5' Cy3) - <u>AAAATTGTGGTCCCT</u>
PAR_Alpha_7	AAGGAGAAAACCTGGACCCCGCTCAGTGGAGGTGGACGACCTCCCTCT
PAR_Alpha_8	GGGTACACCAGTTT
Holliday junction “J1_Alpha”	
J1_Alpha_1	<u>GAGGTTGGACAATTT</u>
J1_Alpha_3	<u>GGAGCACACCAACCTC</u>
J1_Alpha_5	(5' Cy5) - <u>AGGGGACCTGTGCTCC</u>
PAR/J1_Alpha_6	(5' Cy3) - <u>AAAATTGTGGTCCCT</u>

Table 3 - Sequences used in the experiments described in this section. Underlined bases are the ones flanking the junction and constitute the totality of the J1_Alpha structure. They are also entirely included in the parallelogram. The oligonucleotide PAR/J1_Alpha_6 is present in both the parallelogram and the isolated junction. Sequences are written from 5' to 3'.

3.1.4.2 Assembling the structures

All the sequences of the oligonucleotides used in this work were designed with the program SEQUIN run with a each size of 5. The resulting sequences are shown in table 3. Oligonucleotides were purchased from MWG (Ebersberg, Germany), HPLC purified and lyophilized by the supplier, and suspended in MilliQ H₂O (Millipore Simplicity) prior to use. The concentration of each strand was estimated by UV spectroscopy by measuring OD₂₆₀.

Both DNA structures were assembled in solution by mixing stoichiometric quantities of each component strand in TAE/Mg²⁺ buffer (20 mM Tris, 2 mM EDTA, 12.5 mM MgCl₂, pH 8.4). The final concentration of each strand was 0.1 μM. The

mixtures were then heated to 90 °C for 5 minutes and cooled to 10 °C in a PCR thermocycler (PCR Sprint, Thermo Electron Corp., Waltham, MA, U.S.A.) by cooling the mixtures at a rate of 0.01 °C/s. The correct formation of the structures was verified with native PAGE electrophoresis experiments (data not shown) in which the complete structures ran as single bands.

3.4.1.3 FRET experiments

The assembled structures solutions were put in a stirred thermostated cuvette, irradiated at 530 nm and their emission spectra were recorded at different temperatures (see figure 30) ranging from 20°C to 55°C.

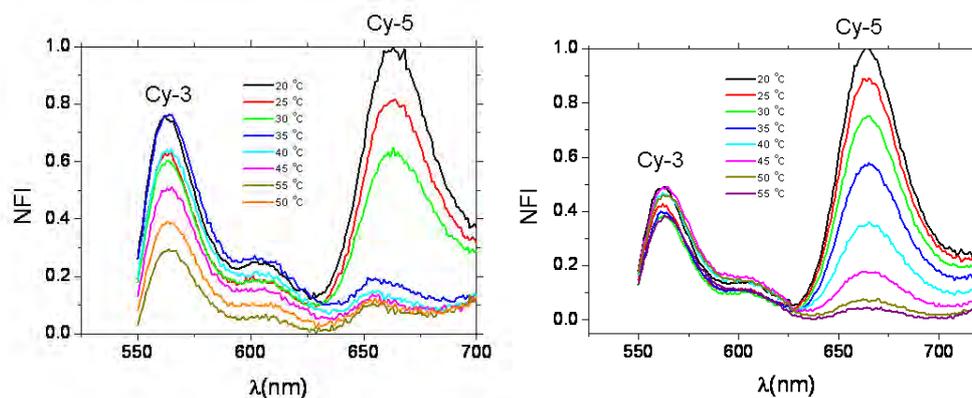


Figure 30 – (Left panel) emission spectra of the J1_Alpha structure in solution irradiated with 530 nm light. (Right panel) the same experiment repeated on the PAR_Alpha parallelogram. Please refer to the main text for discussion. [image courtesy of Dr. Branko Kolaric]

Since E is obviously not 100% in either case, the emission peak of the directly-irradiated dye (Cy3) is visible around 560 nm in both experiments. The Cy5 dye is excited by FRET and its emission is clearly visible around 660 nm. The recorded spectra were used to calculate the FRET efficiency E at each temperature for both the J1_Alpha and the PAR_Alpha. The results are shown in figure 31.

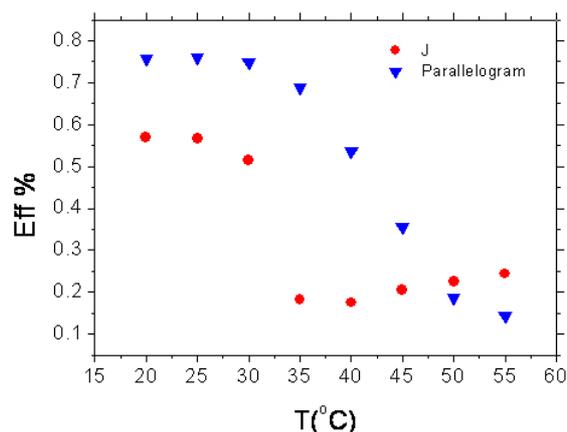


Figure 31 – Fret efficiencies as calculated from the series of spectra shown in figure 30. [image courtesy of Dr. Branko Kolaric]

3.4.1.4 Conclusions

As clearly evidenced by the data in figure 31, at 20°C the parallelogram induces a higher FRET efficiency than the J1 junction. Our conclusion is that the parallelogram is intrinsically more rigid than the isolated junction in these conditions. We increased the temperature in an attempt to induce more stochastic motion around the structures and thus try to stress their mechanical rigidity to an higher degree. However, the fluorescence yields of the dyes are severely limited by the increasing temperature, as visible in the spectra of figure 30. FRET drops abruptly in the J1 structure above 35°C, whereas it decreases more gradually in the parallelogram, ceasing only above 50°C. The overall conclusion from this set of data is that the parallelogram is more capable than the J1 junction to keep objects in place and function as a nanometer-scale scaffold for other objects.

3.1.5 A DNA parallelogram as a nanoscale scaffold for proteins

3.1.5.1 The concept

As also discussed in section 2, DNA nanoarchitectures can be used as a self-assembling scaffold for placing other objects deterministically in the nanoscale. The goal of the experiments presented in this section is to produce and characterize a rhomboidal,

mechanically rigid structure and use it for the precise positioning of proteins. As evidenced by the experiments presented in the previous section, the outward pointing arms of a small parallelogram are quite rigidly kept in a defined position; this would allow to keep proteins tethered to these arms at a known relative distance.

The idea behind such a structure is that knowing and modulating the relative distance between two or more proteins, known to interact in some way, would allow to test the effects of distance on this interaction. This would allow to study protein cross-talking systems and possibly enhance the amount of cross-talk by keeping the proteins at an optimal distance and not allowing to them to diffuse freely in solution. Verifying that this DNA-structure-induced spatial distribution of proteins can influence their activity would be the first step towards the design of synthetic “protein-based nanofactories”. Unfortunately we are still quite far from this result.

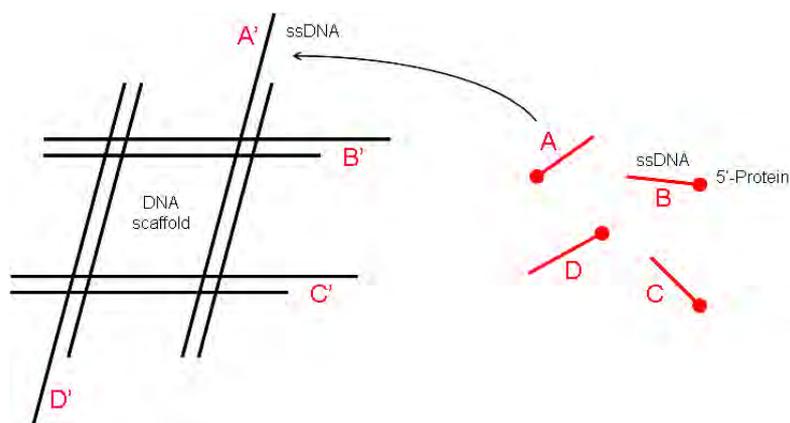


Figure 32 – Cartoon of the structure studied in this section. A DNA parallelogram has four sticky ends capable of recognizing synthetic ssDNA oligonucleotides functionalized at the 5' terminus with a protein.

3.1.5.2 Design and synthesis of the DNA rhomboidal scaffold

Figure 32 contains a schematic representation of the structure we designed to function as a nanoscale scaffold for the positioning of proteins. A parallelogram with identical dimensions to that studied in section 3.1.4 is equipped with four slightly longer outer arms, each one including a sticky end capable of hybridizing a ssDNA oligonucleotide functionalized at its 5' terminus with a protein. In this way, the proteins would be placed at exactly the desired position by the choice of their conjugated oligonucleotides, and then kept in place rigidly by the structure.

A very convenient way of producing ssDNA-protein conjugates is coupling 5'-biotinylated oligonucleotides with streptavidin-modified proteins.^[29] To test the feasibility of this approach, we decided to initially use biotinylated oligonucleotides conjugated with just a single streptavidin molecule.

One immediate problem posed by this approach is the following. The biotin-streptavidin bond is completely stable only below 65°C; subjecting it to higher temperatures would mean to lose some percentage of ssDNA-protein conjugation. DNA nanostructures however, are usually formed by annealing all the constituent components from high temperature (typically 90°C) to ensure that only the desired pairings form between all the constituent oligonucleotides. Thus the ‘annealing of all the components together’ approach and the ‘biotin-streptavidin conjugation’ approach seem to be mutually incompatible.

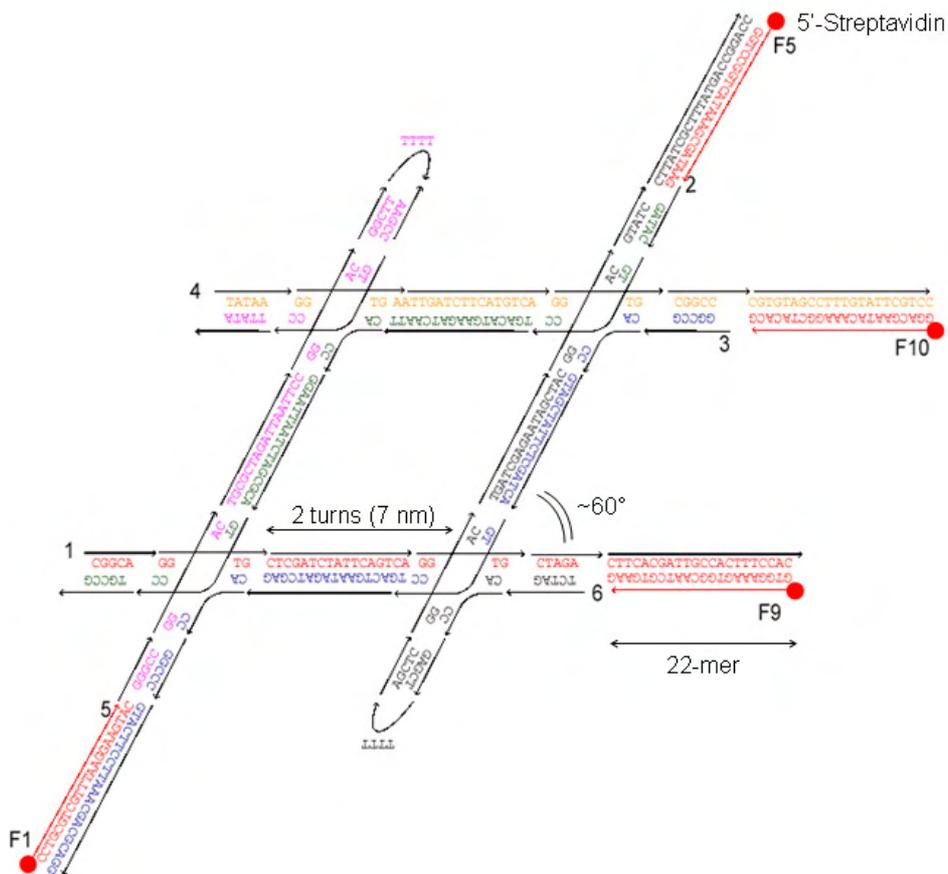


Figure 33 – Schematic representation of the structure and sequence used in the experiments reported in this section.

Quite recently, Niemeyer and coworkers reported^[30] a series of *in silico* designed short DNA sequences that are completely devoid of any noticeable secondary structure at 25°C. This means that they can specifically bind to a complementary strand even without annealing them together from high temperature, because no kinetic trapping in unwanted structures is likely to happen. These sequences are the ideal candidates to function as protein carriers in our design, since they could be added to the pre-formed DNA structure at room temperature, without subjecting them to high temperatures.

Parallelogram “PAR_PCN”	
PCN_1	CGGCAGGTGCTCGATCTATTTCAGTCAGGTGCTAGAC <u>CTTCACGATTGCCACTTCCAC</u>
PCN_2	GATACGTCCTGACATGAAGATCAATTCACCGGAATTAATCTAGCGCAGTCCTGCCG
PCN_3	GGCCGCACCGTAGCTATTCTCGATCAGTCCTGACTGAATAGATCGAGCACCGGCC <u>GTAC</u> <u>TTCCTTAAACGACGCAGG</u>
PCN_4	TATAAGGTGAATTGATCTTCATGTCAGGTGCGGCC <u>CGTGTAGCCTTTGTATTTCGTCC</u>
PCN_5	GGCCCGGACTGCGTAGATTAATTCGGACGGCTTTTTTAAGCCGTCCTTATA
PCN_6	TCTAGCACCGAGCTTTTTAGCTCGGACTGATCGAGAATAGCTACGGACGTATC <u>CTTATCG</u> <u>CTTTATGACCGGACC</u>
Carrier oligonucleotides PAR_PCN_F(n)s, <u>without</u> STV	
PCN_F1	CCTGCGTCGTTTAAGGAAGTAC
PCN_F5	GGTCCGGTCATAAAGCGATAAG
PCN_F9	GTGGAAAGTGGCAATCGTGAAG
PCN_F10	GGACGAATACAAAGGCTACACG
Carrier oligonucleotides PAR_PCN_F(n)*s, <u>with</u> 5'-STV	
PCN_F1*	5' - (STV) - (biotin) - CCTGCGTCGTTTAAGGAAGTAC
PCN_F5*	5' - (STV) - (biotin) - GGTCCGGTCATAAAGCGATAAG
PCN_F9*	5' - (STV) - (biotin) - GTGGAAAGTGGCAATCGTGAAG
PCN_F10*	5' - (STV) - (biotin) - GGACGAATACAAAGGCTACACG

Table 4 - Sequences used in the experiments described in this section. Sequences are written from 5' to 3'. Underlined sequences in the parallelogram are those complementary to the PCN_F(n).

Following this reasoning, four of the sequences reported by Feldkamp *et al*^[30] were chosen to be included in the structure the names of these sequences are PCN_F1,

PCN_F5, PCN_F9 and PCN_F10 (see table 4). Then, we designed a parallelogram structure in which four of the eight external arms are equipped with sticky ends with sequences complementary to those of each PNC_F(n) (see figure 33). The remainder of sequences of the structure were designed by applying the sequence symmetry minimization feature of the program NANEV 1.1^[10] run with a SSM search size of 4. Successive iterations were generated until the quality factor reached a value of 1756. All the resulting sequences are shown in table 4.

Oligonucleotides were purchased from MWG (Ebersberg, Germany), HPLC purified and lyophilized by the supplier, and suspended in MilliQ H₂O (Millipore Simplicity) prior to use. The concentration of each strand was estimated by UV spectroscopy by measuring OD₂₆₀ and fine-tuned performing complementarity-base titration as described in section 3.1.1.

The parallelogram was assembled in solution by mixing stoichiometric quantities of each component strand in TAE/Mg²⁺ buffer (20 mM Tris, 2 mM EDTA, 12.5 mM MgCl₂, pH 8.4). The final concentration of each strand was 0.1 μM. The structure was then annealed by placing the mixtures in a 2 l water bath at 90 °C in a styrofoam box and left to cool to room temperature (around 20 °C) over approximately 72 hours.

3.1.5.3 Testing the incorporation of target sequences

To test both the correct formation of the parallelogram and its ability to incorporate all the target sequences, the pre-formed structure was added with 1:1 stoichiometric amounts of all the PCN_F(n) sequences and incubated at 25°C for 30 minutes. The resulting structures were run on a 7,5% polyacrylamide (19:1 acrylamide:bisacrylamide) gel in TAE/Mg²⁺ at 5V/cm for 4h. The gel was then ethidium-bromide stained and imaged with a digital camera under UV illumination. The result of the electrophoretic run are shown in figure 34.

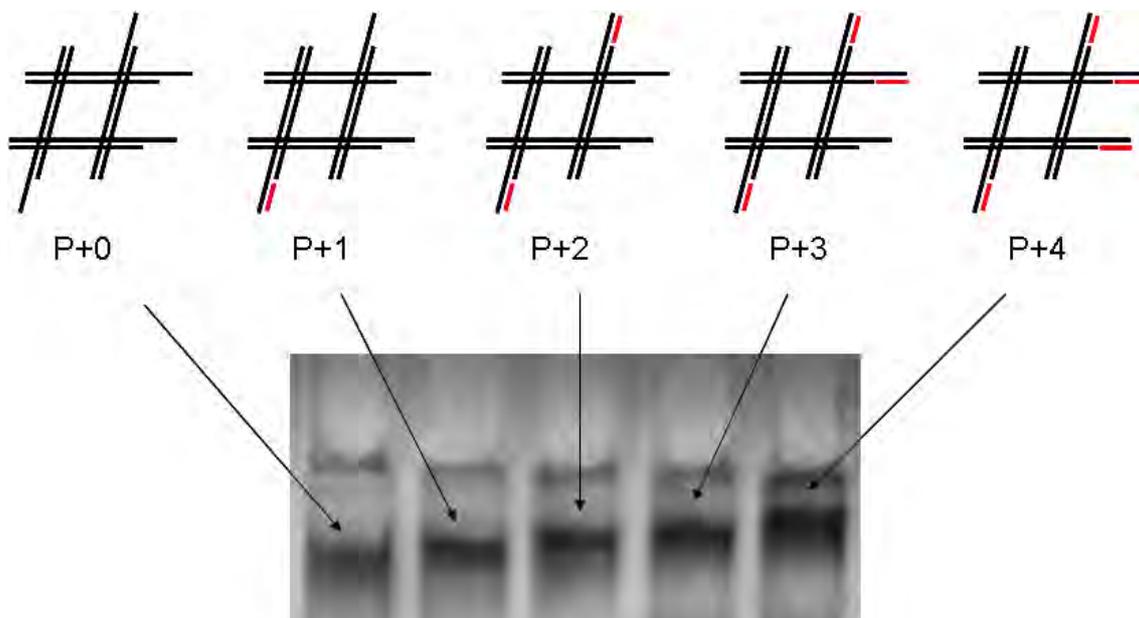


Figure 34 – Electrophoretic run of the PCN parallelogram added with zero, one, two, three or four of the target PCN_F(n) oligonucleotides. The parallelogram and the target oligonucleotides were mixed at room temperature in a 1:1 stoichiometric ratio, and incubated at 25°C for 30'. See the main text for details and discussion.

The first lane on the left contains the complete parallelogram, running as a single band (thus confirming its correct formation). The second lane contains the parallelogram plus PCN_F1. The corresponding single band is slightly retarded as an effect of increased molecular weight of the complex. In each successive lane, one PCN_F(n) is added, in this order: F5, F9, F10. Each complex runs as a single band with a slightly heavier mass than the preceding one, in accord with the good formation of all the complexes. The incorporation efficiency seems to be 100% for each PCN_F(n).

To test the specificity of target positioning within the template, the same experiment described above was performed, but annealing the parallelogram with a 50% stoichiometric defect of oligo PCN_4 (see figure 35). Since the annealing mixture only contains 50% of the needed amount of PCN_4, the first lane contains two bands: the upper one is the complete parallelogram incorporating all the available oligo PCN_4. The lower band is the incomplete parallelogram formed by the other five strands (without PCN_4). As stoichiometric amounts of PCN_F(n)s are added, the upper series of bands (the complete PCN) behaves exactly as in the previous experiment (compare with figure 34). The lower band also behaves similarly, showing mass increases for every successive addition save for the last one. This is due to the fact that the five-stranded structure lacks oligo PCN_4, the one that should bind to oligo PCN_F10.

When PCN_F10 is added to the structure lacking PCN_4, no new bond can form and the structure is not retarded any further.

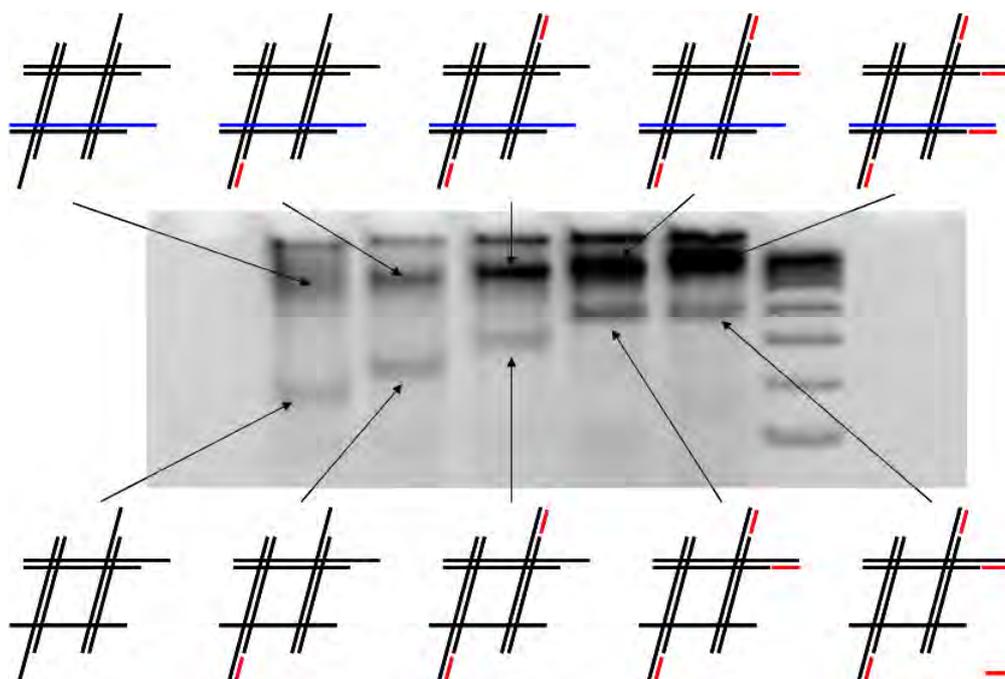


Figure 35 – An experiment identical to that described in figure 34, but with a 50% stoichiometric defect of oligo PCN_4. Oligo PCN_F10 cannot thus bind to the incomplete structure. The last lane on the right is a Lambda DNA/EcoRI+HindIII marker. See the main text for a more complete discussion.

To be sure that the incorporation efficiency of the target sequences PCN_F(n) was the highest possible, we repeated the experiment depicted in figure 34 but this time annealing all the sequences together from 90°C downwards, including the PCN_F(n) oligonucleotides. Then we run a native agarose gel comparing the result of complete annealing of the target oligonucleotides with the structures obtained by simply mixing the PCN_F(n)s and the complete pre-formed PCN at room temperature as described above. The results are shown in figure 36. Please refer to the figure caption for the experimental details. The two sets of bands are completely identical and no fast ‘leftover’ bands are visible, suggesting that the efficiency of incorporation is 100% in both cases and the ‘mixing at room temperature’ strategy is a viable route for preparing the structure.

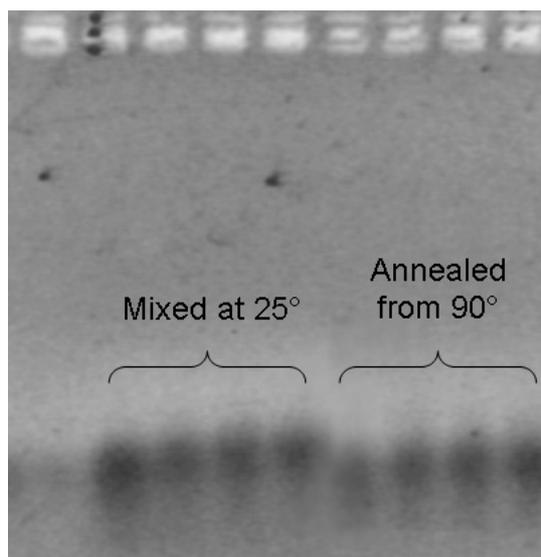


Figure 36 – Electrophoretic run of the PCN parallelogram either mixed at 25°C (left series) or annealed together (right series) with PCN_F(n)s. No difference is evident from the gel, suggesting that the incorporation is 100% efficient in both cases. 1% agarose gel in TAE/Mg²⁺ run at 5V/cm for 1h.

3.1.5.4 Testing the incorporation of ssDNA/STV conjugates

The incorporation of target oligonucleotides conjugated with biotin and bound to streptavidin (STV) was then attempted following the same protocol used for the gel shown in figure 34. The result of this attempt are shown in figure 37. Unfortunately the incorporation of functionalized oligonucleotides is far from being 100% efficient in the same conditions of stoichiometry used for the un-conjugated targets. A lot of the unreacted PCN is visible at each addition. Moreover, since the incorporation is not complete, multiple product bands are visible when more than one different PCN_F(n)* is added to the PCN.

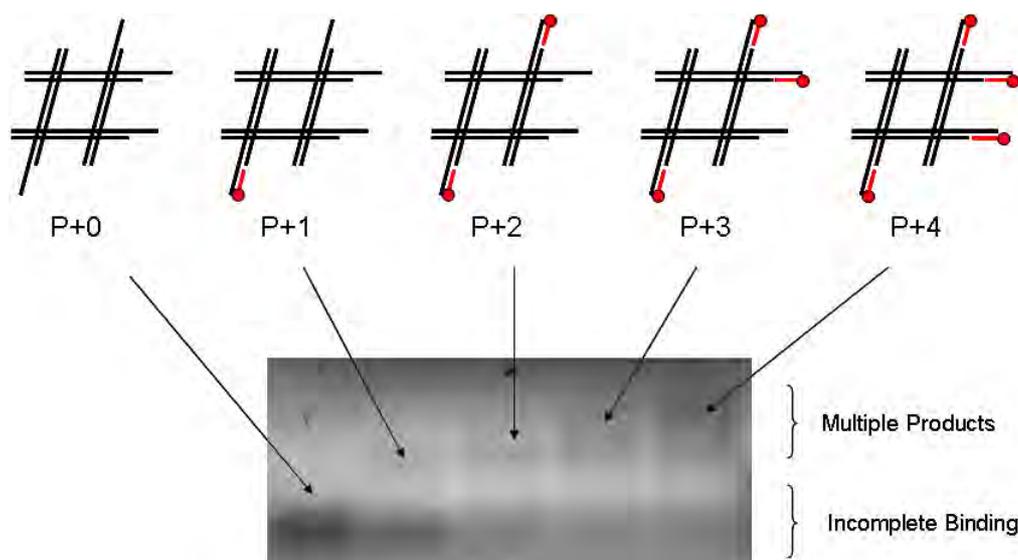


Figure 37 - An experiment identical to that described in figure 34, but using STV-DNA conjugates PCN_F(n)* instead of non-functionalized oligonucleotides.

In an attempt to push the hybridization efficiency, we tried to repeat the functionalization with higher stoichiometric PCN_F(n)* : PCN ratios. We tried ratios of 5:1 (data not shown) and 10:1 (see figure 38), noticing slight, but not decisive, increments in the hybridization efficiency. Increasing the ionic strength by adding 0.5 M NaCl was also attempted (data not shown), with no further results.

3.1.5.5 Conclusion

The experiments reported in this section can be viewed as a partial success. Parallelogram DNA tiles are capable of selectively bind at room temperature target sequences decorated with small organic moieties such as biotin, but if they are conjugated with bigger species such as proteins the efficiency of the incorporation is extremely reduced. It is difficult to understand whether this effect is just due to an increase of steric hindrance, but this seems unlikely considering the position in which the proteins should be tethered to our structure. To reinforce this point, the 100% efficient incorporation of these same sequences in more sterically hindered contexts was successfully implemented.^[30] Further experimenting with different proteins and DNA scaffolds is needed to answer this apparent inconsistency.

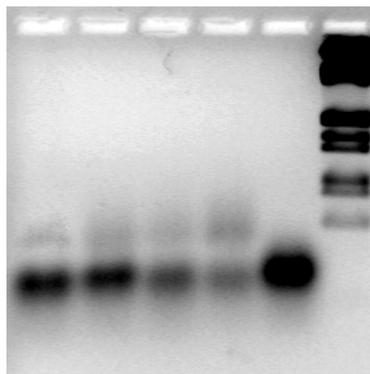


Figure 38 – Repetition of the experiment shown in figure 37, but with an higher stoichiometric ratio between target PCN_F(n)*s and PCN. The first lane on the left is PCN plus 10 equivalents of PCN_F1*. In successive lanes 10 equivalents of respectively PCN_F5*, F9*, and F10* were added to the previous sample. The fifth lane contains just the PCN parallelogram. The last lane is a Lambda DNA/EcoRI+HindIII marker.

3.2 A dynamic structure based on a DNA duplex-triplex transition

In this section, we report that the formation and breakdown of an intramolecular cytosine–thymine (CT) motif DNA triple helix can be performed repeatedly, quickly and independently of its local concentration without performance reduction over successive cycles; as a consequence, we propose that this set of characteristics makes the DNA duplex–triplex transition an ideal candidate to power simple nanometer-scale devices capable of maintaining effective performance regardless of their local concentration.

3.2.1 Rationale, design and synthesis

As also reviewed in section 2.2, a field of DNA nanotechnology research that has been thoroughly investigated recently is the design and construction of DNA nanomotors capable of moving nanoscale objects under some sort of external input. The most common strategies for obtaining controlled movement of the DNA nanostructures use competing hybridization equilibria which give rise to different topologies,^[31-36] or take advantage of the inherent capability of DNA to dynamically assume different conformations in response to changes in the environment. DNA conformational transitions utilized so far in pursuit of this goal include, among others, the B–Z^[37]

transition, the duplex–tetraplex^[38] transition, and the controlled DNA cruciform migration upon addition of an intercalator.^[39]

To be suitable for this kind of implementation, a DNA conformational transition has to satisfy the requirements of being reasonably fast and robust,^[33] that is, to respond to repeated exposure of an appropriate switching stimulus consistently, in the same way, without showing performance change or degradation. While most of the DNA conformational transitions employed in this field so far meet these requirements, several others remain uncharacterized with respect to their speed and robustness.

One such DNA conformation shift that has not been utilized so far* to generate controlled motion in the nanoscale is that between double helix and CT-motif triple helix. The CT-motif DNA triple helix is a well-documented structure^[41-44] formed by a target duplex and a homopyrimidinic ‘triplex forming oligonucleotide’ (TFO). The formation of this type of triplex is critically dependent on the protonation of the imino groups of the TFO cytosines and can therefore be driven dynamically by controlled pH changes.

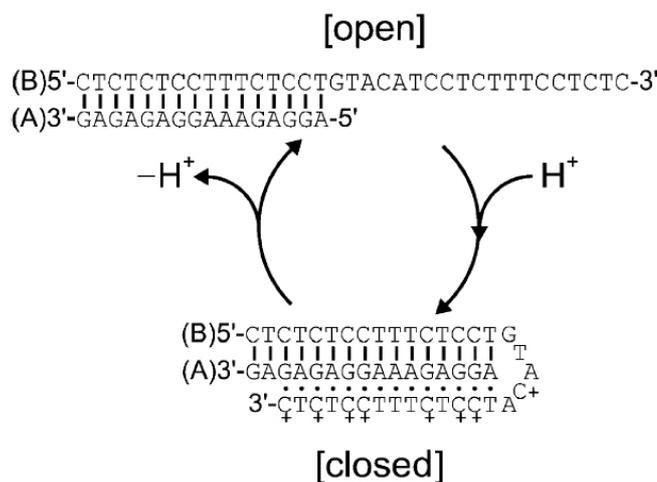


Figure 39 - Oligonucleotide sequences and conformations. See the main text for discussion.

* At the time in which the experiments described in this section were commenced (January 2004), the duplex-to-triplex transition was never used to power a nanoscale DNA actuator. During the publishing of our results, the group of Professor Chengde Mao presented similar experimental results obtained independently.^[40] Y. Chen, S. H. Lee, C. Mao, *Angew Chem Int Ed Engl* **2004**, *43*, 5335. To date, these two articles remain the only two published examples of DNA nanomachines powered by duplex-triplex transitions.

In order to study the dynamic behavior and the robustness of this transition we designed a DNA structure capable of alternately assuming two conformations by means of the pH-driven intramolecular formation and breakdown of a CT motif DNA triple helix, then performed static and dynamic characterization of the structure.

The structure under examination was generated by self-assembly from stoichiometric quantities of two synthetic DNA single-stranded oligonucleotides labeled A and B (see figure 39 and table 5). Oligo A is a homopurinic 16-mer. Oligo B is a 35-mer whose sequence is composed of three sections: a 16-nucleotide (nt) long homopyrimidinic section starting from the 5'-terminus, capable of forming a Watson and Crick duplex with oligo A; a central 'hinge' made of 5 random nucleotides; and a 14-nt long homopyrimidinic portion ending on the 3'-terminus the sequence of which was designed to act as an intramolecular CT-motif TFO targeting the constituted duplex. The two oligonucleotides form an adduct with a 16-bp duplex and a 19-nt long TFO single-stranded overhang.

DX-TX Dynamic Structure	
A	AGGAGAAAGGAGAGAG
B	CTCTCTCCTTTCTCCTGTACATCCTCTTTCCTCTC
B*	(5')-RhG-CTCTCTCCTTTCTCCTGTACATCCTCTTTCCTCTC-Dbc-(3')
C	CTCTCTCCTTTCTCCTGTACACATAGTACAGACAT

Table 5 - Sequences used in the experiments described in this section. Sequences are written from 5' to 3'. RhG = 5'-Rhodamine Green dye, Dbc = 3'-DabcyL.

As noted before, the conformation of this bipartite structure is critically dependent on the pH; in slightly acidic conditions (pH 5.0), the TFO section folds back on the duplex, binding parallel to its purine strand in the major groove, forming Hoogsteen-type triplets and bringing the two opposite termini of the adduct in close proximity. In moderately alkaline conditions (pH 9.0), the TFO section of oligo B cannot form a stable CT-motif triplex and assumes a random-coil conformation dangling from one end of the stable duplex (the open state). The two termini of oligo B in the adduct are thus kept further apart than in acidic conditions.

3.2.2 Experiments in solution: Static Characterization

The two states of the system were first characterized statically by circular dichroism (CD) spectroscopy, UV spectroscopy and electrophoretic mobility shift assay (EMSA).

CD spectra of the A + B adduct were recorded in both acidic (pH 5.0) and alkaline (pH 9.0) conditions (figure 40). The spectrum recorded at acidic pH shows a negative peak at ca. 215 nm that has been typically associated with the proposed triplex structure.^[45, 46]

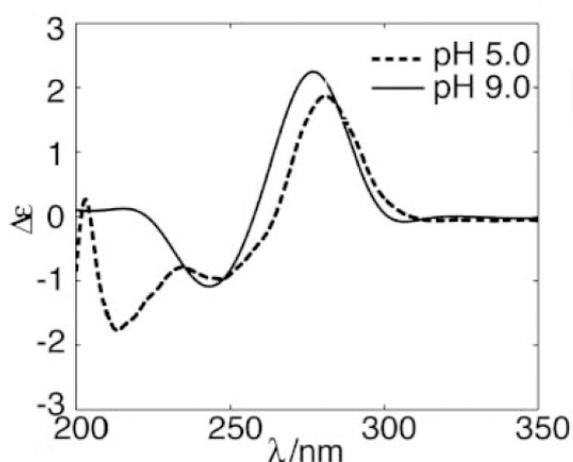


Figure 40 - CD spectra of the open and closed states. Open state: 10 μ M A + B in 20 μ M tris, 20 mM $MgCl_2$, 50 mM NaCl at pH 9.0. Closed state: 10 μ M A + B in 20 mM, 20 mM $MgCl_2$, 50 mM NaCl at pH 5.0.

The shift between open and closed conformations was also observed following ascending and descending pH titrations of the A + B adduct by UV spectroscopy (figure 41). A plot of the UV absorbance at 260 nm vs. the pH of the sample shows the stability zones of the two conformations, evidenced as zones of markedly different absorption separated by a sharp variation at pH 6.5. This can be explained owing to the different extent of the hypochromic effect associated with the two different structures.^[47, 48]

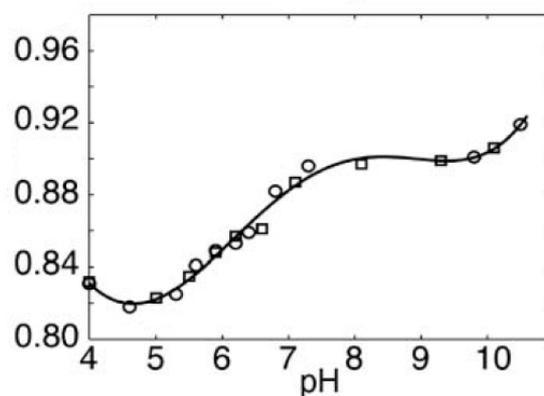


Figure 41 - Ascending (squares) and descending (circles) pH titration with 0.1 M HCl or NaOH of a 15 μ M A + B, 20 mM MgCl₂, 50 mM NaCl solution followed by UV absorption at 260 nm.

To further confirm the hypothesized structural change of the adduct, a modified version of oligo B was synthesized, with the TFO sequence portion replaced by a random sequence of the same length (oligo C). Stoichiometric adducts A + B and A + C were prepared, and then analyzed by non-denaturing polyacrylamide gel electrophoresis, both in basic (pH 9.0) and acidic (pH 5.0) buffer conditions along with their single-stranded constituents. At pH 9.0, the B and C oligonucleotides exhibit the same electrophoretic mobility. Similarly, the A + B and the A + C adducts also have the same mobility at this pH (see figure 42).



Figure 42 - Electrophoretic mobility of the B and C 35-mer single-stranded oligonucleotides and their adducts A + B and A + C at pH 5.0. Oligo B is slower than oligo C, however its adduct with oligo A is faster than the A + C adduct (see the main text for discussion). Gel run in a 20 mM acetate, 20 mM MgCl₂, 50 mM NaCl buffer at pH 5.0.

At pH 5.0, however, oligo C runs faster than oligo B, while the A + C adduct is slower than the A + B adduct. Oligo C includes 12 cytosines, whereas oligo B includes 16. Since these are protonated at acidic pH, the overall negative chargedensity of oligo C is expected to be higher than that of oligo B, explaining the observed mobility difference. The higher mobility of the A + B adduct with respect to the A + C can be rationalized by the increase in compactness upon triple helix formation that overrides the aforementioned charge density effect.

3.2.3 Experiments in solution: Dynamic Characterization

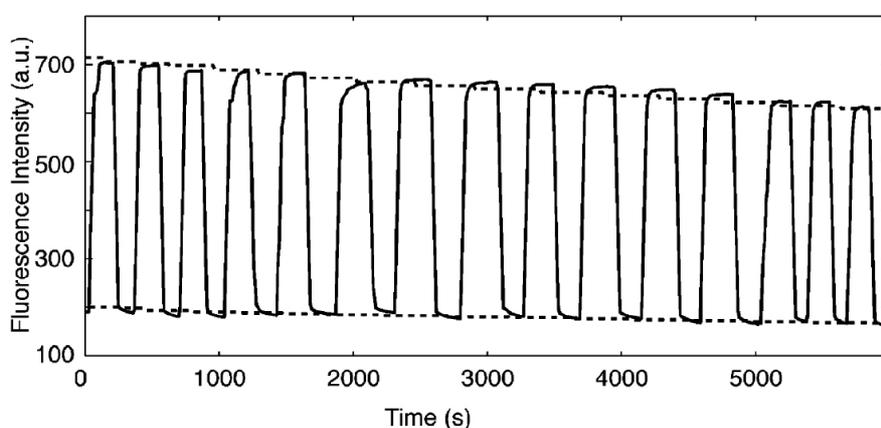


Figure 43 - Cycling of the A + B* nanomotor observed by fluorescence spectroscopy. Excitation was set at 500 nm and emission monitored at 530 nm. The initial A + B concentration was 0.33 μ M in a 20 mM MgCl₂, 50 mM NaCl solution. The pH was cycled between 5 and 9 by alternated addition of 0.1 M HCl or NaOH. Dashed lines show a plot of the expected peak intensity decrease resulting from dilution calculations. Switching between states is complete within a few seconds after the addition of acid or base.

The repeated switching back and forth between the open and closed states of the system was monitored in real time with fluorescence spectrophotometry. Following an established approach,^[31] a modified version of oligo B was synthesized, bearing a Rhodamine Green fluorophore at the 5' end and a Dabcyl quencher moiety at the 3' end (oligo B*). Rhodamine Green shows a strong fluorescence emission at ca. 530 nm when excited at 500 nm, and its fluorescence yield is pH-insensitive between pH 4.0 and 9.0. The fluorescence emission intensity of the A + B* system was measured while the pH of the sample was repeatedly cycled between 5.0 and 9.0 with controlled additions of concentrated HCl and NaOH. This

evidenced dramatic shifts in fluorescence synchronized with the induced pH changes (see figure 43).

The fluorescence emission changes are indicative of the system alternately assuming the open and closed state, because the fluorescent dye and the quencher on the two termini of the adduct are kept in close proximity in the closed state (thus allowing energy transfer from emitter to quencher)^[49] and further apart in the open state. The switching between the two states of the structure is completed within a few seconds from the acid or base addition and is therefore comparable with the switching speeds shown by published DNA transition-based molecular devices(see references of section 2.2) The fluorescence intensity consistently oscillates between two clearly defined states upon repeated exposure of the construct to the input stimuli, thus demonstrating the robustness of the observed transition.

The slight gradual reduction observed in the overall fluorescence intensity over the cycles is due to, and quantitatively correlates with, the dilution of the construct solution upon acid or base addition. This demonstrates that the molecular mechanism of this transition is not hampered (over the observed cycles) by the accumulating ‘waste product’ NaCl. The gradual increase of ionic strength due to the accumulating salt is not expected to significantly change the electrostatic potential near the surface of the DNA structure, and consequently its performance, at least until the salt reaches molar concentration.^[50]

The cycling of the examined structure does not entail any binding event involving two macromolecules, and should not therefore be influenced by the local concentration of the structure itself. The speed and the robustness of the cycling were proved to be concentration-independent by diluting the A + B adduct from micromolar (figure 43) to less than nanomolar (figure 44) and successfully repeating the cycling experiment, evidencing the same overall behavior.

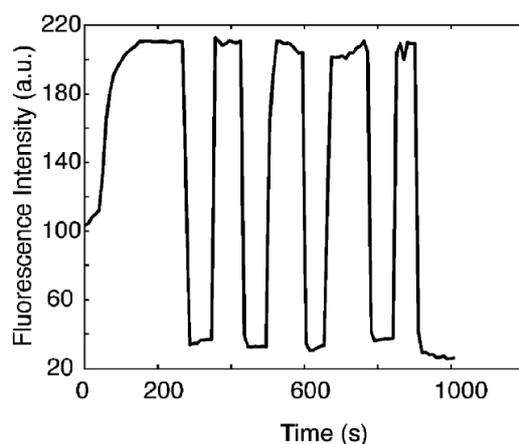


Figure 44 - Repetition of the experiment with an initial A + B concentration of 0.33 nM.

From a publicly available 3D structure of a CT-motif triplex,^[51] we could estimate the distance between the two termini of oligo B to be approximately 2 nm in the closed state. Employing a polymer statistics model developed for chains with sections of different flexibilities,^[25] we estimated that the root-mean-square distance between the two termini of oligo B is around 8 nm in the open state. This implies that the described repeated cycling of the pH generates a fully controllable relative motion of two termini that brings them to span on average a 6 nm long segment.

3.2.4 Experiments on surfaces: Single Molecule Characterization

To characterize the functioning of the device at the single molecule level, an extremely diluted solution of the device (picomolar range) was deposited on a glass surface. Figure 45 shows the fluorescence intensity time trace of one molecule at pH 9; the observed one-step photobleaching is a clear evidence that the fluorescence was collected from a single quantum system. The traces show irreversible photobleaching after approximately 15 seconds.

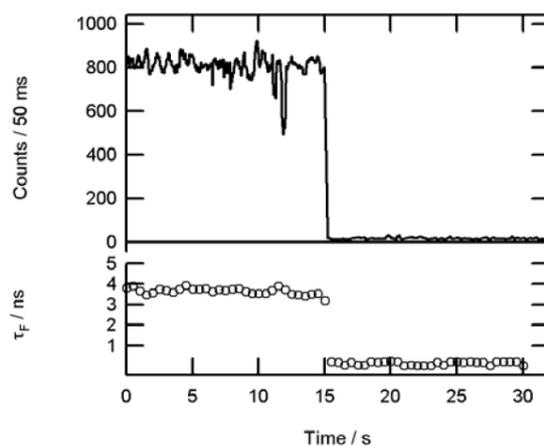


Figure 45 - Example of one fluorescence intensity trace of a single oligonucleotides construct at pH 9. The low panel show fluorescence lifetime (calculated each 0.5 s).^[52] [image taken from the same reference]

Figure 46 shows the fluorescence intensity curves and scanned images taken from single molecules at two pH values, showing a complete switching between the open and closed forms. At pH 5, the absence of any brighter spots proved that any single dye is completely quenched. The brighter spot is recovered at the same position once the pH value is increased again to 9. This demonstrates that the lack of intermolecular interactions between constructs lead to a complete quenching, thus the observation of a switching between the two states. This complete quenching was observed for more than 50 individual molecules.

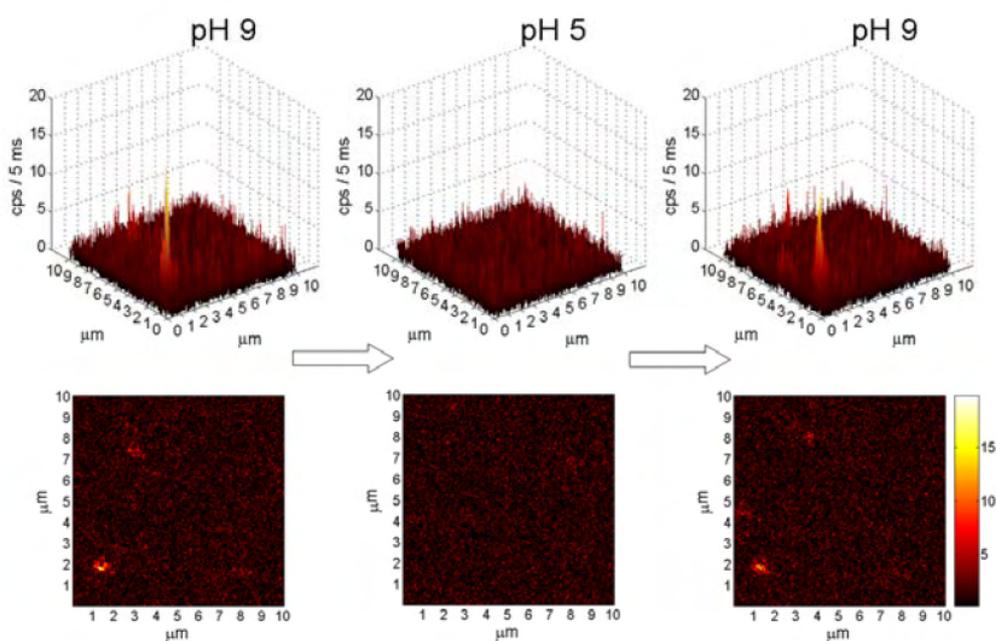


Figure 46 - Fluorescence emission of a single molecule oligonucleotides construct observed at two pH values, size of the images is 10 μm x 10 μm .^[52] [image taken from the same reference]

3.2.5 Experimental

Oligonucleotides were purchased from MWG (Ebersberg, Germany), HPLC purified and lyophilized by the supplier, and suspended in MilliQ H₂O (Millipore Simplicity) prior to use. Adducts in the desired conformations were assembled by mixing oligonucleotides in the appropriate buffer.

CD studies were conducted on a J-710 spectropolarimeter (Jasco, Easton, USA) in a 300 μl quartz cuvette with a 0.1 cm path length, containing a 10 μM solution of the nanomotor in either 20 mM acetate buffer at pH 5.0 or 20 mM Tris buffer at pH 9.0. Both samples included 50 mM NaCl and 20 mM MgCl₂.

UV spectra were recorded on a SmartSpec3000 UV-vis spectrophotometer (BioRad, Hercules, USA) in a 500 μl quartz cuvette with a 1 cm path length, loaded with a 15 μM A + B adduct, 20 mM MgCl₂, 50 mM NaCl solution.

FRET experiments were performed on a LSB-50 fluorescence spectrometer (Perkin Elmer, Boston, USA) in a 3 ml quartz cuvette equipped with a magnetic microstirrer and loaded with 2 ml of a 0.33 μM or 0.33 nM A + B* adduct, 50 mM NaCl and 20 mM MgCl₂ solution. The pH of the sample was cycled between 5 and 9

alternately adding 6 μl of 0.1 M HCl or 0.1 M NaOH aqueous solutions directly into the cuvette with a syringe micro-pump.

Polyacrylamide gel electrophoresis was performed in a BioRad Protean II Midi Gel Box. 1.5 mm thick native 15 % polyacrylamide gels were run in a 4°C cold room at 100 V for 24 hours in 20 mM acetate buffer at pH 5 or 20 mM Tris buffer at pH 9. Both buffers contained 20 mM MgCl_2 and 50 mM NaCl. Gels were then stained with stains-all dye, de-stained in distilled water to enhance contrast, and then photographed with a digital photcamera (Gel Doc 1000, BioRad, Hercules, USA).

Single-molecule fluorescence experiments were performed by Dr. Branko Kolaric and Dr. Michel Sliwa at Katholieke Universiteit Leuven, under the supervision of Prof. DeSchryver. Details of the experimental set-up and of the analysis procedure were published before by the group of Prof. Deschryver.^[53] In short, the a (488 nm, 8.18 MHz, 1.2 ps FWHM) light source was used. The excitation light, circularly polarized by use of a Berek compensator, was directed into the oil-immersion objective (Zeiss, 1.3 N.A., 100 \times) of an inverted microscope (Olympus IX70) by using a dichroic beam splitter. The excitation power was set to 1 μW at the entrance port of the microscope. The single molecule fluorescence was collected by the same objective, filtered and focused (via a 100 μm pinhole) onto an avalanche photodiode (SPCM-AQR-15, PerkinElmer) equipped with a time correlated single photon counting card (Becker & Hickl GmbH, SPC 630). This combination of the card used in FIFO mode and the APD, allows one to get the fluorescence signal and the decay profile of single dye molecules with an experimental instrument response in the order of 400 ps. A piezoelectric translation stage (Physik Instrument) is mounted on the microscope and allows one to get scanning images by recording the fluorescence intensity at each position. The maximum likelihood estimation (MLE) method was used to analyze the single molecule (SM) fluorescence decays.¹⁶ The oligonucleotides constructs were dissolved in tris buffer at a pM concentration and transferred on a cover slip (10 μL). The cover slip was then mounted at the bottom of a home made 1000 μL reaction chamber, which allows exposure of the oligonucleotides constructs to tris buffer solution while monitoring the fluorescence signal. The cycling between pH 5 and 9 was obtained by consecutive addition of 0.1 M NaOH and 0.1 M HCl respectively.

3.2.6 Conclusions

The collected evidence on the dynamic properties of the investigated transition permits us to add the duplex–triplex conformational transition to the repertoire of the possible DNA nanodevice mechanisms. A structure capable of performing the transition we discussed could easily be coupled with a variety of nano-objects to dynamically and precisely control their relative distance. Recent literature reports several examples of DNA nanomotors constituted by a DNA actuator capable of dynamic response to external stimuli, coupled to a rigid DNA structure^[54] (see also section 2.2).

Most DNA nanoscale devices published so far include in their cycles one or more events involving two macromolecules, usually two complementary oligonucleotides. This implies that their cycling is concentration dependent, and could suffer worsening if performed in the proximity of an obstacle capable of diminishing the access of macromolecules, such as a surface of some sort. Devices powered by the duplex–triplex transition we studied would have the advantage of not producing significant waste products and not being influenced by their concentration or the surroundings in which they would operate.

The transmission of the device opening and closing signals would be directly dependent on the mobility of very small and fast-diffusing species such as H^+ and OH^- . These characteristics would make the smooth functioning of the nanodevice possible in conditions that could hamper the performance of the DNA nanomotors reported so far, such as in nano-sized pores, or in systems where the mobility of the species involved in the cycling is critical.

3.3 A three-dimensional DNA Origami: preliminary studies

As summarized in section 2.1.6, DNA origami^[14] is an extremely promising design paradigm for building DNA nanoarchitectures in an efficient and straightforward method. An extremely advantageous feature of the approach is that it is highly generalizable, i.e. the same design approach can be applied to arbitrary shapes without changing the type of design.

For now, the DNA origami approach can be employed for the production of two-dimensional flat objects, with or without features. But there is in my opinion no apparent reason for not trying to extend the approach to other type of objects, for example three-dimensional objects or curved surfaces. At the time of writing however, no such structure was ever reported, even if the writer is reasonably sure that some will, before long.

To tackle this problem, we designed a small curved DNA origami object having the form of a six-helix bundle^[16] (see figure 47). The six-helix bundle motif was originally prototyped using short synthetic oligonucleotides^[16] but the topology of its constituent strands allows for a continuous, long ssDNA strand to participate in all its helices, going back and forth repeatedly through the structure. This strand could function as the ssDNA template on which DNA origami structures are assembled.

The template would never participate in the crossovers keeping the helices together. It would simply enter an helix, go through it doing an integer number of full helical turns and leave the helix at the opposite side, but in the same orientation with which it entered. Then the template strand would need to enter one adjacent helix in the bundle structure, reach the opposite side and so on. This would be repeated for each helix, until the structure was complete.

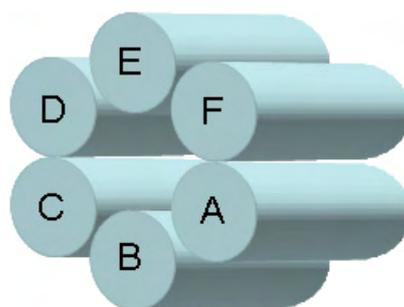


Figure 47 – Schematic representation of a DNA six-helix bundle. The lettering used in this figure to identify helices is also used in the following figures.

As the first step of design, we arbitrarily decided that the 5' terminus of the ssDNA template will be located in helix A, in the 'back' side of the structure (the side not visible in figure 47). Then it will reach the front side and enter helix B, reach the back of the structure through helix B, enter into helix C, and so on until all the helices are run through and the back side of helix F is reached.

The only problem posed by this design in addition to the ones discussed by Rothmund^[14, 55] for planar origami structures is the angle at which successive helices join. Supposing an one-nanometer separation between helices in the bundle, and identifying the points at which the template must enter and exit each helix, there are some large ‘jumps’ that the template strand must perform to reach the successive helix entry point (see figure 48). For example, the separation between the exit point of helix C and the entry point of helix D could be as much as 4.9 nm apart. The other ‘jumps’ are 2.6 nm at most.

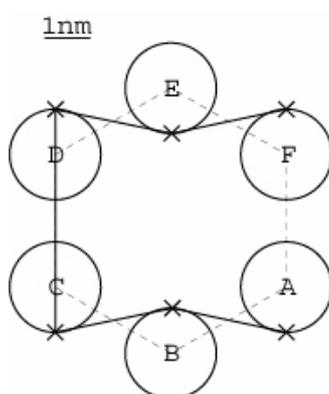


Figure 48 – Worst-possible scenario of helix separation in the bundle, drawn to scale. The Xs represent the points at which the ssDNA template enters and exits each helix. See the main text for discussion.

Taking these distances into account, we decided to include extra bases in the template at some of the positions between exit and entry points. For example, we decided to leave seven unpaired bases of ssDNA in the template after the last base participating in the formation of helix C and the first base constituting helix D. We included extra bases in the D-E and E-F jumps, and not in the A-B and B-C ones. We decided for this because, considering the helicity of the template strand in successive helices, some jumps could be favored by the partial unwinding of the template from the two helices. Some adjacent helices (A, B and C) are in a reciprocal polarity orientation such that unwinding bases from both creates an extra length of ssDNA between them. For the other helices (D, E and F) the opposite is true: they are in such a reciprocal orientation that no unwinding is possible at all, unless some extra length of ssDNA is introduced between exit and entry points. For this reason, we decided to put stretches of four unpaired bases at the D-E and E-F jump points.

For the considerations reported above, the needed template length is 519 bases. Although the DNA origami strategy should be largely independent of the actual sequences in the ssDNA template, we decided nevertheless to use, for our template, a shorter sub-set of the template employed in the only published example. We scanned the M13mp18 sequence employed by Rothemund to find a 519-bases tract with relatively weak intramolecular interactions, and designed the rest of the structure so that it would form on this tract. The result of the design is shown in figure 49.

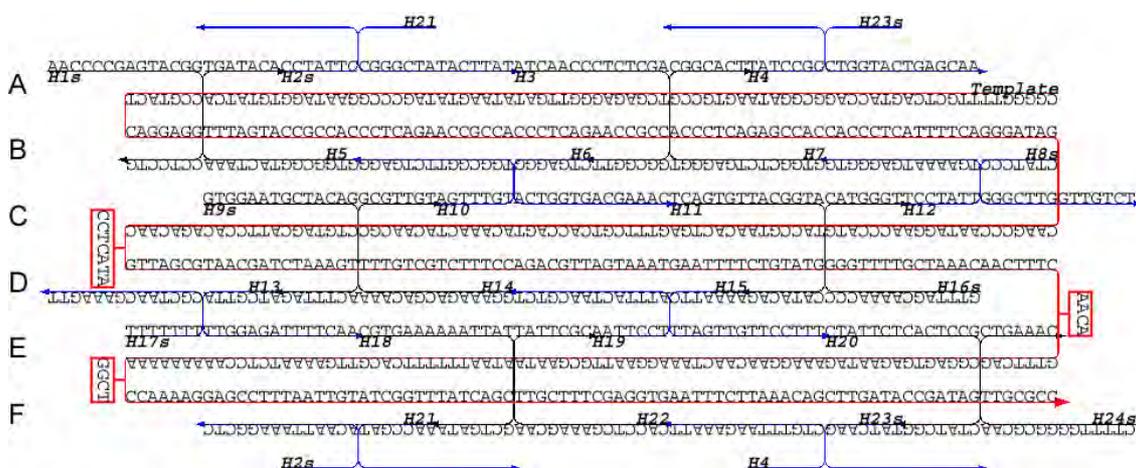


Figure 49 – Sequence and connectivity of the ssDNA template and the synthetic helper strands needed to form the structure. The structure showed here has sticky ends on both sides, to form 1D arrays. See the main text for discussion.

Twenty-four short ‘helper strands’ H(n) are needed to fold the template into the correct shape. The helper strands form the crossovers that keep the helixes together at the desired angles. Crossovers joining two adjacent helixes are contained in a same plane (because they are separated by an integer number of helical turns, four in this case). A third helix is joined to one of these with crossovers lying on another plane, forming an angle of 120° with the previous one. Each helical domain is exactly 84 bases (six full helical turns) long. All the sequences are shown in table 6.

We performed only very preliminary experiments on this design. We produced the single-stranded template by PCR amplification using one biotinylated primer (the one incorporated in the template sequence). Then we separated the biotinylated ssDNA template from its complementary strand by using magnetic beads. The design is now ready to be tested experimentally.

Origami Helix Bundle “6HB-O”					
Template		CGGGGTTTTGCTCAGTACCAGGCGGATAAAGTGCCGTGAGAGGGTTGATATAAGTATAGC CCGGAATAGGTGTATCACCGTACTCAGGAGGTTTAGTACCGCCACCCTCAGAACCGCCAC CCTCAGAACCGCCACCCTCAGAGCCACCACCCTCATTTTCAGGGATAGCAAGCCCAATAG GAACCCATGTACCGTAACTGAGTTTCGTACCCAGTACAACTACAACGCCTGTAGCAT TCCACAGACAACCCTCATAGTTAGCGTAAACGATCTAAAGTTTTGTCGTCTTCCAGACGT TAGTAAATGAATTTTCTGTATGGGGTTTTGCTAAACAACCTTCAACAGTTTCAGCGGAGT GAGAATAGAAAGGAACAATAAGGAATTGCGAATAATAATTTTTTTCAGGTTGAAAATCT CCAAAAAAAAAAGGCTCCAAAAGGAGCCTTTAATTGTATCGGTTTATCAGCTTGTCTTCGA GGTGAATTTCTTAAACAGCTTGATACCGATAGTTGCGCC			
Helper strands for a 6HB without sticky ends					
H1	AGTACGGCCTCCTG	H9	GTTGTCTGTGGAATGCTACAGA CTTTAG	H17	TTTTTTTCGCTAAC
H2	CCTATTCACAATTAAGGCTCC TTTTGG	H10	AGTTTGTGGCGGTTCTGAGG	H18	CGTGAAAAAATTATGCTGATA
H3	ATCAACCCTCTCGAGGCGGTT	H11	TCAGTGTACGGTACATACAG	H19	AATTCCTATTTACTAACGTCT
H4	TATCCGCCTGTTAAGAAATT	H12	TCCTATTTGAAAATGAGGGTG	H20	CTATTCTCACTCCGCTATCGG
H5	GTGGCGGTACTAAATGATACA	H13	ATCGTTATTGGAGATTTTCAA	H21	AACCGATCGGGCTATACTTAT
H6	CTGAGGGACTGGTGACGAAAC	H14	GGAAAGACGACAAAGCGTTGT	H22	CACCTCGAAAGCAATATTCGC
H7	GTGGCTCTGAGGGTCGGCACT	H15	AAAATTCTTAGTTGTTCCCTTT	H23	TATCAAGCTGGTACTGAGCAAA ACCCCG
H8	CTATCCCGGGCTTG	H16	GAAAGTTGTTTAGCAAAACCC CATGGGT	H24	GGCGCAACTGAAAC
Helper strands for a 6HB with sticky ends					
H1-S	AACCCCGAGTACGGCCTCCTG	H16-S	GTTTAGCAAAACCCCATGGGT		
H2-S	CCTATTCACAATTAAGGCTC	H17-S	TTTTTTTCGCTAACGAAAGTT		
H8-S	CTATCCCGGGCTTGGTTGTCT	H23-S	TATCAAGCTGGTACTGAGCAA		
H9-S	GTGGAATGCTACAGACTTTAG	H24-S	CTTTTGGGGCGCAACTGAAAC		

Table 6 – Sequences for a six-helix bundle built with the origami approach. See the main text for discussion.

In conclusion, during the design of this structure we did not find any structural reason that could impair the formation of origami helix bundles. If the experiments will confirm this conjecture, then this approach could be used to test the feasibility of several different tube-shaped structures. In particular, the non-strained DNA nanotubes with a definite number of helical domains described by Sherman and Seeman^[56] could be

experimentally implemented with the origami approach. The amount of synthetic oligonucleotides needed to form the larger non-strained tubes would make the building of such structures almost impossible without using the origami approach.

Chapter II Bibliography

- [1] C. Mao, W. Sun, N. C. Seeman, *J. Am. Chem. Soc* **1999**, *121*, 5437.
- [2] N. C. Seeman, *Biochemistry* **2003**, *42*, 7259.
- [3] R. Sha, F. Liu, D. P. Millar, N. C. Seeman, *Chem Biol* **2000**, *7*, 743.
- [4] R. Sha, F. Liu, N. C. Seeman, *Biochemistry* **2002**, *41*, 5950.
- [5] M. Brucale, G. Zuccheri, L. Rossi, A. Bazzani, G. Castellani, B. Samori, *Organic & Biomolecular Chemistry* **2006**, *4*, 3427.
- [6] H. Yan, S. H. Park, G. Finkelstein, J. H. Reif, T. H. LaBean, *Science* **2003**, *301*, 1882.
- [7] N. C. Seeman, *Trends Biotechnol* **1999**, *17*, 437.
- [8] P. E. Constantinou, T. Wang, J. Kopatsch, L. B. Israel, X. P. Zhang, B. Q. Ding, W. B. Sherman, X. Wang, J. P. Zheng, R. J. Sha, N. C. Seeman, *Organic & Biomolecular Chemistry* **2006**, *4*, 3414.
- [9] N. C. Seeman, *J Biomol Struct Dyn* **1990**, *8*, 573.
- [10] R. P. Goodman, *Biotechniques* **2005**, *38*, 548.
- [11] B. Wei, Z. Wang, Y. L. Mi, *Journal of Computational and Theoretical Nanoscience* **2006**, *3*.
- [12] J. J. Birac, W. B. Sherman, J. Kopatsch, P. E. Constantinou, N. C. Seeman, *J Mol Graph Model* **2006**.
- [13] W. M. Shih, J. D. Quispe, G. F. Joyce, *Nature* **2004**, *427*, 618.
- [14] P. W. Rothmund, *Nature* **2006**, *440*, 297.
- [15] M. Brucale, G. Zuccheri, B. Samori, *Trends in Biotechnology* **2006**, *24*, 235.
- [16] F. Mathieu, S. Liao, J. Kopatsch, T. Wang, C. Mao, N. C. Seeman, *Nano Lett* **2005**, *5*, 661.
- [17] S. H. Park, R. Barish, H. Y. Li, J. H. Reif, G. Finkelstein, H. Yan, T. H. LaBean, *Nano Letters* **2005**, *5*, 693.
- [18] J. M. Lehn, *Progress in Polymer Science* **2005**, *30*, 814.
- [19] E. A. Fogleman, V. R. Kempf, S. L. Craig, *Abstracts of Papers of the American Chemical Society* **2002**, *223*, D5.
- [20] E. A. Fogleman, W. C. Yount, J. Xu, S. L. Craig, *Angewandte Chemie-International Edition* **2002**, *41*, 4026.

- [21] J. M. Lehn, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 4763.
- [22] Y. He, Y. Tian, Y. Chen, Z. Deng, A. E. Ribbe, C. Mao, *Angew Chem Int Ed Engl* **2005**, *44*, 6694.
- [23] G. Ercolani, L. Mandolini, P. Mencarelli, S. Roelens, *Journal of the American Chemical Society* **1993**, *115*, 3901.
- [24] G. Adam, M. Delbrück, in *Structural Chemistry and Molecular Biology* (Eds.: A. Rich, N. Davidson), W. H. Freeman & Co., **1968**.
- [25] C. Rivetti, C. Walker, C. Bustamante, *Journal of Molecular Biology* **1998**, *280*, 41.
- [26] L. Blanco, M. Salas, *Proc Natl Acad Sci U S A* **1984**, *81*, 5325.
- [27] L. Blanco, A. Bernad, J. M. Lazaro, G. Martin, C. Garmendia, M. Salas, *J Biol Chem* **1989**, *264*, 8935.
- [28] D. Lubrich, J. Bath, A. J. Turberfield, *Nanotechnology* **2005**, *16*, 1574.
- [29] C. M. Niemeyer, *Trends Biotechnol* **2002**, *20*, 395.
- [30] U. Feldkamp, H. Schroeder, C. M. Niemeyer, *J Biomol Struct Dyn* **2006**, *23*, 657.
- [31] B. Yurke, A. J. Turberfield, A. P. Mills, Jr., F. C. Simmel, J. L. Neumann, *Nature* **2000**, *406*, 605.
- [32] F. C. Simmel, B. Yurke, *Physical Review E* **2001**, *6304*.
- [33] H. Yan, X. Zhang, Z. Shen, N. C. Seeman, *Nature* **2002**, *415*, 62.
- [34] K. Hamad-Schifferli, J. J. Schwartz, A. T. Santos, S. G. Zhang, J. M. Jacobson, *Nature* **2002**, *415*, 152.
- [35] L. Feng, S. H. Park, J. H. Reif, H. Yan, *Angew Chem Int Ed Engl* **2003**, *42*, 4342.
- [36] W. B. Sherman, N. C. Seeman, *Nano Lett* **2004**, *4*, 1203.
- [37] C. Mao, W. Sun, Z. Shen, N. C. Seeman, *Nature* **1999**, *397*, 144.
- [38] J. J. Li, W. Tan, *Nano Lett* **2002**, *2*, 315.
- [39] X. Yang, A. V. Vologodskii, B. Liu, B. Kemper, N. C. Seeman, *Biopolymers* **1998**, *45*, 69.
- [40] Y. Chen, S. H. Lee, C. Mao, *Angew Chem Int Ed Engl* **2004**, *43*, 5335.
- [41] G. Felsenfeld, D. R. Davies, A. Rich, *J. Am. Chem. Soc* **1957**, *79*, 2023.

- [42] J. S. Sun, T. Garestier, C. Helene, *Current Opinion in Structural Biology* **1996**, 6, 327.
- [43] J. L. Asensio, A. N. Lane, J. Dhesi, S. Bergqvist, T. Brown, *Journal of Molecular Biology* **1998**, 275, 811.
- [44] P. Alberti, P. B. Arimondo, J. L. Mergny, T. Garestier, C. Helene, J. S. Sun, *Nucleic Acids Research* **2002**, 30, 5407.
- [45] C. Gondeau, J. C. Maurizot, M. Durand, *Nucleic Acids Research* **1998**, 26, 4996.
- [46] L. E. Xodo, *Febs Letters* **1995**, 370, 153.
- [47] P. L. Husler, H. H. Klump, *Archives of Biochemistry and Biophysics* **1995**, 317, 46.
- [48] M. Mills, P. B. Arimondo, L. Lacroix, T. Garestier, C. Helene, H. Klump, J. L. Mergny, *Journal of Molecular Biology* **1999**, 291, 1035.
- [49] D. P. Millar, *Current Opinion in Structural Biology* **1996**, 6, 322.
- [50] J. L. Hecht, B. Honig, Y. K. Shin, W. L. Hubbell, *Journal of Physical Chemistry* **1995**, 99, 7782.
- [51] J. L. Asensio, T. Brown, A. N. Lane, *Nucleic Acids Research* **1998**, 26, 3677.
- [52] B. Kolaric, M. Sliwa, M. Brucale, R. A. L. Vallée, G. Zuccheri, B. Samori, J. Hofkens, F. C. De Schryver, *Photochemical & Photobiological Sciences* **2007**, DOI: 10.1039/b618689k.
- [53] T. Vosch, M. Cotlet, J. Hofkens, K. Van der Biest, M. Lor, K. Weston, P. Tinnefeld, M. Sauer, L. Latterini, K. Mullen, F. C. De Schryver, *Journal of Physical Chemistry A* **2003**, 107, 6920.
- [54] B. Ding, N. C. Seeman, *Science* **2006**, 314, 1583.
- [55] P. W. K. Rothmund, *Abstracts of Papers of the American Chemical Society* **2006**, 231.
- [56] W. B. Sherman, N. C. Seeman, *Biophys J* **2006**, 90, 4546.

III – CONCLUSION AND APPENDIXES

4 Conclusions and Perspective

In this thesis, I reported examples of novel DNA self-assembled systems of various nature. Most of the work presented herein has been published in widely accessible journals.

We showed how extremely different supramolecular constructs can be assembled with the same set of constituent sub-components by driving the self-assembly process toward only one of its possible outcomes. We tried to assemble a DNA 1D array directly on a long scaffold ssDNA template produced by Rolling Circle Amplification, and compared the resulting structures with those produced by tile polymerization. Scaffolded structures are markedly larger, but their self-assembly cannot be driven to different outcomes like the individual tiles system. We experimentally verified that the mechanical coupling of four flexible nanoscopic junctions yield an object that is more rigid than its individual parts, due to the distribution of stress to the entire structure. We designed and implemented a rigid nanoscopic DNA structure and immobilized proteins on it with sequence-dependent specificity. We designed and implemented a DNA dynamic structure, capable of moving its parts in response to a change in the pH of the environment. The functioning of this nanoscopic device was also tested at the single-molecule level by immobilizing it on a glass surface. I also discussed how the origami approach could be extended to produce curved, tube-like structures and designed a prototype DNA six-helix bundle using this approach.

All the systems described are attempts to study specific areas of structural DNA nanotechnology and DNA self-assembly that are still mostly unexplored. My personal opinion is that the most promising perspectives for this field lie in three main areas:

(i) Self-assembly processes

The first aspect to which I'd like to draw attention is the study of the dynamic processes that bring from the constituent components to the complete structure in a self-assembling system. Current structural DNA nanotechnology research focuses solely on the *results* of DNA self-assembly (many of which are amazing). DNA computation, on the other hand, focuses on the modeling of self-assembling behaviors, putting less emphasis on what exactly is being assembled. I think that there is an interesting region to study between these two approaches, that we attempted to prototype with the rhomboidal tile system reported in section 3.1.2. The possibility to dynamically alter the outcome of self-assembly systems can only stem from the detailed knowledge of the involved processes.

I think that the example offered by DNA origami is particularly exciting. The construction of sturdy, arbitrarily shaped objects with the DNA origami technique is hugely successful when compared to other reported techniques such as algorithmic aperiodic assemblies. This is *not* due to any peculiarity in the structure of DNA origami, but rather due to the peculiar nature of the process leading from all the constituent strands to those amazing shapes and patterns. If you consider carefully the process of how the helper strands bind to the ssDNA template (a summary of this can be found in section 2.1.6), the striking conclusion is that *nothing can go wrong*. All the sub-processes driving the main self-assembly process away from its intended outcome are so improbable as to be irrelevant. Now imagine what could be obtainable if one such system was designed so that it had several possible outcomes for assembly, and we had the knowledge of how to deterministically drive the process towards one of these. I think that in order to be able to design such a system, the scientific community will need a better understanding of assembly processes and, consequently, finer design tools than today's.

For example, tools that allowed us to reliably design structures with zones of different thermal stability, curvature or flexibility could lead to assembly processes in which different parts of a structure are stabilized in a given sequence. One could then design DNA 3D 'boxes' with 'lids' that could be opened or closed at will, and maybe capture guest molecules inside, or perform other actions.

(ii) The search for function

In his PhD thesis, Paul Rothmund observes that the field of DNA computing might appear to non-adepts like "an answer in search of a question". I think that same could be said about structural DNA nanotechnology. One direction of structural DNA nanotech research that I find will dramatically rise in importance in the next few years is trying to confer to DNA nanostructures specific functional capabilities emerging from the geometry of the structure itself.

The necessary tools all appear to be there – we know how to put proteins and other moieties in specific positions on DNA structures, and we know how to assemble even extremely complex DNA structures. The missing link is probably an extensive research on exactly which existing or new functional systems could benefit from a geometrically determined placement with nanoscopic precision. The simple systems reported in section 3.1.4 and 3.1.5 are attempts to move in this direction. We focused on creating a structurally simple system that could allow to test the effect of relative distance, and possibly orientation, of two or more proteins.

(iii) Interaction with living systems

There is another extremely exciting aspect of DNA nanotechnology that has been so far just barely touched by few research groups: devices and structures mainly composed by biomolecules have the intrinsic ability to interact directly with biological systems. Possible scenarios include the design of DNA computing systems that output specific biomolecules in response to recognized biomolecules. Or structures that release objects contained by them or attached to them in presence of certain species. Or artificial nanopores and channels, and ‘smart stoppers’ for natural pores and channels.

I guess that the dream of most researchers interested in structural DNA nanotechnology is to combine all the concepts outlined above in one coherent line of research. This research would be aimed at (i) accurately drive self-assembling processes to desired outcome structures, (ii) design such structures as to be functional nano-sized objects capable of performing a variety of tasks, and (iii) design structures capable of reading information from their environment and producing deterministic responses. All this is probably quite far from being accomplished. However, it also looks a lot like the most complex game nature eventually choose to play on this planet: life. I’d say it is a motivating precedent.

5 Appendixes

5.1 Publications produced during this PhD candidature

(i) Journals

Brucale M, Zuccheri G, Samori B, "The dynamic properties of an intramolecular transition from DNA duplex to cytosine-thymine motif triplex" *Org. Biomol. Chem.* 2005 3(4)575-577

Zuccheri G, Brucale M, Samori B, "The Tube or the Helix? This is the Question: Towards the Fully Controlled DNA-Directed Assembly of Carbon Nanotubes" *Small* 2005 1(6)590-592

Brucale M, Zuccheri G, Samori B, "Mastering the Complexity of DNA nanostructures" *Trends Biotechnol.* 2006 24(5)235-43

Brucale M, Zuccheri G, Rossi L, Bazzani A, Castellani G, Samori B "Characterization and modulation of the hierarchical self-assembly of nanostructured DNA tiles into supramolecular polymers" *Org. Biomol. Chem.* 2006 4(18)3427-34

Kolaric B, Sliwa M, Brucale M, Vallée RAL, Zuccheri G, Samori B, Hofkens J, De Schryver FC, "Single molecule fluorescence spectroscopy of pH sensitive oligonucleotides switches" *Photochemical & Photobiological Sciences* 2007, [in press, DOI: 10.1039/b618689k]

(ii) Book Chapters

Zuccheri G, Brucale M, Vinelli A, Samori B, "DNA-Based Artificial Nanostructures"
In: *Annual Review of Nanoresearch Vol.1 (Chapter 12)* 2007, World Scientific Publishing

(iii) Oral Presentations

Brucale M, Zuccheri G, Samori B, "A Surface-Bound, pH-controlled DNA Molecular Motor: Computational and Experimental Characterizations"
MMD Meeting Genova Jun 2005 proceedings p.45

Brucale M, Zuccheri G, Samorì B.

“Progress on the implementation of hierarchical assembly in DNA supramolecular structures”

SONS Meeting Nijmegen Apr 2005

Brucale M, Zuccheri G, Samorì B, “Linear and Circular Self-Assembled Supramolecular Polymers Made of DNA Parallelograms”

3rd National Conference on Nanoscience and Nanotechnology Trieste 22-24 May 2006

Brucale M, Zuccheri G, Samorì B.

“Progress on the single-molecule studies of artificial DNA nanostructures”

SONS Meeting Strasbourg Feb 2006

Brucale M, Zuccheri G, Samorì B.

“Progress on the design and construction of DNA artificial scaffolds for nanoelectronics”

NUCAN Meeting Potsdam Feb 2006

Brucale M, Zuccheri G, Samorì B, “From Biology to Nanotechnology: Building Nanostructures via the Hierarchical Self-Assembly of DNA Oligonucleotides”

Congresso della Società Chimica Italiana Firenze September 2006 proceedings p.428

Brucale M, Zuccheri G, Samorì B, “The first steps toward protein-based nanofactories: sequence-dependant immobilization of proteins on an artificial DNA scaffold”

SONS Meeting Cambridge Oct 2006

(iv) Other Communications at Congresses

Brucale M, Zuccheri G, Samorì B, “A DNA Nanomotor Based on a pH-sensitive conformational transition” INFMeeting Genova 2004

Brucale M, “A pH-Controlled DNA Nanomotor Based on a Triple Helix”

2nd Central European Conference “Chemistry towards Biology” 26 – 29 September 2004

Brucale M, Zuccheri G, Samorì B, “A DNA Molecular motor based on a pH-dependent conformational transition” DNANANOWIRES Modena 7 – 8 October 2005 [proceedings p. 11]

Zuccheri G, Brucale M, Samorì B, “Nanoscale Molecular assembly by design: building flexible or rigid, static or dynamic nanostructures thanks to the controlled self-assembly of DNA molecules”
DNANANOWIRES Modena 7 – 8 October 2005 [proceedings p.60]

Zuccheri G, Brucale M, Samorì B, "Nanoscale Molecular assembly by design: building flexible or rigid, static or dynamic nanostructures thanks to the controlled self-assembly of DNA molecules"
MMD Meeting Genova Jun 2005 proceedings p.51

Brucale M, Zuccheri G, Samorì B, “Self-Assembled Supramolecular Polymers Made of DNA Parallelograms” DNA-based Nanoscale Integration Jena May 18 – 20 2006

Brucale M, Zuccheri G, Samorì B, “A Self-Assembled Supramolecular Polymer Made of DNA Parallelograms” SONS Conference 2006, Pisa 29 June – 1 July 2006

Zuccheri G, Brucale M, Samorì B, “Building flexible or rigid, static or dynamic nanostructures thanks to the controlled self-assembly of oligonucleotides: one way towards nanoscale molecular assembly by design”
V Meeting “Materiali Molecolari avanzati per Fotonica ed Elettronica”, Arbatax 18-21 June 2006

Zuccheri G, Brucale M, Bergia A, Samorì B, “The interaction of DNA with surfaces and such other trifles” Marie Curie RTN Chextan Spring School (combined with the Netherlands Supramolecular Community Meeting): “Chiral expression and transfer at the Nanoscale: From Molecules to Machines”, Nijmegen 5-7 April 2006

Zuccheri G, Brucale M, Samorì B, “Supramolecular Polymeric Chains and Dynamic Objects of Nanometric Size Obtained by the Self-Assembly of DNA Oligonucleotides”
Congresso della Società Chimica Italiana Firenze September 2006 proceedings p.343

Zuccheri G, Brucale M, Samorì B, “DNA nanostructure adsorption and growth on inorganic surfaces”
American Vacuum Society 53rd International Symposium 12-17 November 2006

Zuccheri G, Brucale M, Sliwa M, Kolaric B, Vinelli A, Hofkens J, DeSchryver F, Samorì B, “The Programmed Self-assembly of DNA as a Tool for Controlling the Structure and Function of Matter on the Nanoscale” SONS Conference 2007, Strasbourg June 2007 (abstract accepted)

5.2 About me

(i) Contact Information

<i>Name</i>	Marco Brucale
<i>Birth place and date</i>	Bologna, 11/06/1976
<i>Email(s)</i>	marco.brucale@unibo.it mbrucale@isof.cnr.it marcobrucale@gmail.com

(ii) Education and Experience

2007	Post-doc work in Bruno Samorì's lab. <i>(I'm currently working on the loose-ends of the experiments presented in this thesis, like the 6HB-bundle and the parallelogram decorated with proteins)</i>
2004-2006	PhD in Prof. Samorì's lab. <i>(If you're reading this I suspect you know about that)</i>
2003-2004	Contract with “Biofin S.r.l” (MN, Italy) in Dr. Massimo Capobianco's lab at CNR. <i>(I hope I do not infringe the non-disclosure agreement by saying that I worked on the synthesis of daunomicyn-derivatized triplex-forming oligonucleotides)</i>
2003	Habilitation exam for the title of Chemist <i>(For the Italian law you're not a full-fledged chemist if you don't pass this exam)</i>
2002	Degree in Industrial Chemistry (University of Bologna) <i>(I worked as an undergraduate student in Prof. Alfredo Ricci's lab on the synthesis of chiral ferrocene-based ligands for homogeneous catalysis. My supervisor was Prof. Bianca Bonini)</i>
1994	Linguistic License (Liceo Copernico) <i>(Apart from English, I speak German semi-decently. I also understand most French and Spanish, but I can't really sustain a conversation. My Italian is pretty good too)</i>

(iii) Research Skills

I know all the basic techniques used in an organic synthesis lab, such as various preparative chromatography methods, IR, UV, CD, HPLC, GC, chiral LC, HPLC-MS, GC-MS, and so on. I have an user-level proficiency of recording and interpreting NMR spectra (H, C, and P). During my work at CNR I also learned the basics of operating an automated DNA synthesizer. During my PhD I acquired an extensive experience in AFM imaging, electrophoretic methods, and using software tools for nucleic acid sequence design. I also have a more limited experience in metal nanoparticle preparation and functionalization.

5.3 Acknowledgements (in strictly chronological order)

There's a huge lot of people that helped me during this work. They gave me access to their time, patience, wisdom, knowledge, and love. I know this is a completely inadequate way to repay them for what they gave me, but I'd like to list them all here nonetheless, and also give a big metaphorical hug to them. OK, here comes the big list:

My parents, Luigi and Anna Pia. That 'completely inadequate way to repay' bit above applies to them most. Thank you for all, I love you.

My brother Aldo, my uncle Cesare and Alfonso: you cannot choose relatives as you choose friends, but I'm so fortunate to have relatives that I would also have chosen as friends.

All the rest of the family, too numerous to be listed here. Love to you all.

Claudia, the single most important person in my life. She has this strange power to make me a better person. Everything would seem empty and dull without her.

Carla Boga: the first person who helped me after I got my degree, and purely out of her good heart.

Massimo Capobianco, my supervisor at CNR, who managed to be at the same time a boss and a friend. He's an incredibly good-hearted guy, and I wish him all the best. A big thank also to the rest of the lab, Umberto, Federico, Eleonora, Patrizia.

Bruno Samori and Giampaolo Zuccheri: for some reason, they made this crazy choice of believing in me right from the start. I'll never repay them for all the wealth of opportunities and experience they gave me. Thanks *a lot*, and I mean it.

Andrea Giro: he was the first to look after me when I arrived at the lab. You rule, Tvrno. I wish you all the best. May your dream come true.

All the people I've met at the lab in these years. They all contributed to make Bruno's lab a very interesting place to be. We all learned a lot from each other, working, discussing, laughing,

sometimes fighting, sometimes insulting each other. All in all, an improbable but extremely fine collection of half-crazy, very gifted people. In loose order they come: Sora Verzia, the beacon that saves you from the impending doom of the watery grave. Maurànte, mad as a hatter and perhaps slightly more, the biggest comical genius ever lived, and an intimidatingly good scientist. Dr. Sandal, *nanodontus extraordinaire*: the two of us have a lot of things in common, most of them pathological. No person can be considered wholly sane if she can follow a five-minutes (or hours) conversation between Dr. Sandal and me. Dr. Biggs, fellow armchair-sportsman, politician and extremely good chap. Giovanni: crafter of fine art, fluent speaker of one of the most beautiful dialects of planet Earth, thank you for the good time we had together. Claudio and “Gent.mo Ing. Guiducci”, the infiltrates from the Engineering dept. They’re both very competent and very funny, a remarkable feat for anybody, but almost impossible for an engineer... ha ha, wait, I’m just joking. Kind of. Oh, and the customary reprimand goes to Carlotta, for the huge +TAS. No wait, I’m joking again... kind of. Flaviana: she was with us for too little to grow any distinctive nickname... and if you believe this all the bridges are for sale, in fact, are on sale, to you. Nadia, the first undergraduate student I helped (or hindered horribly, can’t remember which): I’m quite fond of her. Dr. Valle, gentleman exquisite, violinist, fine wines connoisseur, slow-food affiliate: a very pleasant person and an outstanding scientist. Every place he’s in becomes a better place to be. Dr. Plasmon: a terrific guy, I wish I had one tenth of his energy, intelligence, knowledge, and faith. It’s been an honor for me to cross the philosophical swords with you. One thing we have in common is that we both are rabid worshippers of Johann Sebastian: thank you, Mr. Bach. You rule. OK, let’s move on: many thanks to Bo Chen, a nice, shy guy I wish I could know better before he had to go back to China. Alessandra, a tremendously nice gal, speaker of the alien idiom commonly called ‘Sardinian’, and the only female rugby player that does not scare me to death. Many thanks to Annina, another very nice lass I’m quite fond of. I wish you all the best. Many thanks to Muso, a rare and puzzling example of how one can be a good friend, a nice colleague, and an ogre at the same time. Do not hassle with the powerful demon Müüs. Thanks to Vijayender, a very nice chap, the last person who joined the lab: I wish you all the best. A special mention is deserved by Mr. Bartolucci. Does he know it? We’ll never know.

Many thanks to all the people who collaborated with us in the European consortiums: I cannot list them all, but let me send a special thank to the people at Leuven, Nijmegen, Dortmund, Paris, and Madrid. It was a great opportunity for me to know all these incredibly good scientists. I learned a lot from them, both as scientists and as ‘normal’ people.

A smorgasbord of thanks to all the pals at the NellArmadio Studios of Bologna: Simone, Alessandro, Federico, Veronica. A helluva bunch of extremely talented, mad people with a vision.

A *very blurred* vision, but whatever. You're great people and great friends. Thanks also to *the Akt* and the noble *Fauno Scorpionis*.

And finally, thanks to the “Venturi Caffè” coffee machine company. During my PhD candidature I pressed the ‘espresso’ button on your machine about 1300 times (rough estimate obtained counting the plastic coffee spoons collected in 365 days. Data not shown, and for your own sake. If anyone is interested, the collection of plastic spoons is still in the lab, and puzzles everyone immensely).

But most of my gratitude goes to thee, o unlikely reader of this thesis. If you arrived this far, I admire your great resolve. Unless, of course, you just jauntily flipped through the whole thing, to arrive unscathed to this paragraph. In that case, PEPEP.

And now, if you excuse me, I have that button to press, for the 1301st time.

Marco



UNIVERSITÀ DEGLI STUDI DI BOLOGNA
DIPARTIMENTO DI CHIMICA
"GIACOMO CIAMICIAN"
VIA SELMI, 2 - 40126 BOLOGNA (ITALY)

VINCENZO BALZANI
Professor of Chemistry

Tel +39 051 2099560 Fax +39 051 2099456
E-mail: vincenzo.balzani@unibo.it
<http://www.ciam.unibo.it/photochem.html>



DOTTORATO IN SCIENZE CHIMICHE – XIX Ciclo

Estratto del verbale della riunione del Collegio dei Docenti del Dottorato in Scienze Chimiche – XIX Ciclo dell'Università degli Studi di Bologna, tenuta il giorno 1 Dicembre 2006

Dott. Marco Brucale

Relatore prof. Bruno Samori

PROGETTAZIONE, SINTESI E CARATTERIZZAZIONE DI COSTRUTTI SUPRAMOLECOLARI COSTITUITI DA DESOSSILIGONUCLEOTIDI SINTETICI

Durante il periodo di dottorato, interamente svolto presso il laboratorio del Prof. Samori, Marco Brucale si è occupato della progettazione, sintesi e caratterizzazione di costrutti supramolecolari costituiti da desossiligonucleotidi sintetici. Il filone in cui si inserisce questa ricerca è quella della "DNA structural nanotechnology", così detta poiché mira a ottenere, tramite l'autoassemblaggio di frammenti di DNA, strutture la cui morfologia sia determinata con una risoluzione di pochi nanometri. Tali strutture possono essere considerate 'impalcature' per l'assemblaggio di altre specie (come proteine, fluorofori, nanotubi, nanoparticelle metalliche ed altro)] ed il loro scopo a lungo termine è quello di ottenere oggetti tecnologici funzionali come nanocircuiti elettronici, unità di immagazzinamento di dati o persino "nano-fabbriche" costituite da più componenti integrati.

Il primo costrutto progettato da Marco è stato un motore molecolare interamente costituito da DNA, che può assumere due conformazioni differenti tramite un riconoscimento di sequenza intramolecolare di tipo Hoogsteen. Questo tipo di meccanismo non era in precedenza stato sfruttato nell'ambito dei motori molecolari a DNA. In collaborazione con l'università di Leuven, il motore molecolare è stato poi ancorato ad una superficie di vetro ed il suo funzionamento è stato verificato a livello di singola molecola immobilizzata sulla superficie. In collaborazione con il Max Planck Institut di Mainz, lo stesso costrutto è stato fissato a cantilevers microfabbricati di nitruro di silicio e sono in corso esperimenti per misurare il lavoro meccanico esercitato sul cantilever dal movimento sincronizzato della popolazione di costrutti.

Un secondo tipo di costrutti studiati da Marco sono stati i parallelogrammi o romboidi di DNA. Il motivo strutturale romboidale era già noto in precedenza; quelli realizzati presso il nostro laboratorio sono i più piccoli di questo tipo fino ad ora riportati in letteratura. È possibile combinare 'tessere' monomeriche romboidali per formare dei polimeri supramolecolari dalla topologia mono- o bi-dimensionale. Una delle tessere romboidali da noi progettate è l'unica fino a ora pubblicata che può polimerizzare in maniera differente a seconda del tipo di trattamento termico subito; variando la velocità della scansione termica che guida la loro polimerizzazione si può fare entrare in gioco un controllo cinetico che può portare dallo stesso monomero a costrutti supramolecolari diversi. Un'altra tessera romboidale è stata usata come impalcatura per il posizionamento sequenza-dipendente di due, tre o quattro proteine sullo stesso costrutto e sono in corso esperimenti, in collaborazione con l'università di Dortmund, per caratterizzare le strutture risultanti. Il dott. Brucale si è anche occupato di misurare, tramite esperimenti FRET effettuati su parallelogrammi funzionalizzati con un'opportuna coppia di fluorofori, il grado di rigidità strutturale del motivo romboidale rispetto ai sotto-motivi costituenti, le giunzioni di Holliday immobilizzate dette anche giunzioni J1.

Negli ultimi mesi del suo dottorato Marco ha iniziato la progettazione e la realizzazione di strutture di tipo "DNA-Origami", cioè forme arbitrarie realizzate ripiegando una sequenza naturale di DNA tramite l'aiuto di numerosi oligonucleotidi sintetici. Le forme progettate, ed in via di realizzazione, sono (i) un nanotubo di DNA costituito da sei eliche parallele che potrebbe costituire il primo esempio di DNA-Origami a geometria tridimensionale e (ii) una struttura rettilinea ad alta rigidità basata sul motivo romboidale.

Vista l'intensa e produttiva attività del Candidato, il Collegio dei Docenti ritiene che il Dr. Marco Brucale sia altamente meritevole di conseguire il Dottorato di Ricerca in Scienze Chimiche.

Bologna, li 23 Febbraio 2007

Il Coordinatore

Prof. Vincenzo Balzani

A handwritten signature in black ink, appearing to read 'V. Balzani', with a long horizontal stroke extending to the right.