School of Computing

Synthetic DNA and RNA origami scaffolds, towards *in vitro* and *in vivo* applications

Alessandro Ceccarelli

Submitted for the degree of Doctor of Philosophy in the School of Computing, Newcastle University

February 2020

Declaration of Authorship

I hereby certify that this thesis has been composed by me and is based on my own work, unless stated otherwise. No other person's work has been used without due acknowledgement in this thesis. All references and verbatim extracts have been quoted, and all sources of information, including graphs and data sets, have been specifically acknowledged.

Signature: *Alessandro Ceccarelli*

Date: 23/02/2020

"An expert is a man who has made all the mistakes which can be made, in a narrow field."

Niels Bohr

Abstract

Nucleic acid self-assembly is a branch of nanotechnology that consists in the use of DNA and RNA, not only as genetic material, but also as structural medium. In the last 30 years this technique developed from the simplest structures into more refined constructions that could require thousands of interacting sequences.

DNA origami were presented by Rothemund as the evolution of DNA assemblies. Origami are nanostructures formed by a long single strand, the scaffold, held together by many shorter oligonucleotides, the staples. Later research showed how to functionalize the origami staples increasing the applicability.

Bacteriophage genomes were used as convenient source of single stranded DNA scaffold. Today the potential biotechnological applications require to be bio-orthogonal and easily transferable from academic research to industrial ones. To solve these problems we can rely on the powerful tools of synthetic biology.

I focused on the synthesis, analysis and application of synthetic scaffolds based on the combinatorial De Bruijn (DB) sequence. DB scaffolds are designed to be bio-orthogonal, uniquely addressable and thermodynamically optimized.

I demonstrated their folding efficiency in DNA origami and RNA/DNA hybrid origami.

DB scaffolds were analyzed *in vivo* in *E. coli* to asses their bio-orthogonality.

Finally I designed a research tool to improve our knowledge on origami folding dynamics. The system includes functionalized origami encapsulated in giant unilamellar vesicles. The system allows the analysis of the origami using high throughput technologies, and can simulate a simple cell-like environment facilitating the intermediate steps towards the *in vivo* origami folding.

With my interdisciplinary approach I contributed to the advancement of nucleic acid origami technology, taking advantage of tools and techniques from synthetic biology, origin of life, bioinformatics along with more classical scientific disciplines.

Acknowledgements

This Ph.D. has been a very interesting journey, one of those experiences made of new places, new situations and, above all, new people.

I would like to thank my supervisor, Natalio Krasnogor, for giving me the opportunity to make this journey and achieve this result, for his humanity and in particular, for showing me that science it's meant to overcome the limits, not setting them up. Because sometimes the craziest ideas are just early smart ideas.

I would like to acknowledge Dr. David Fulton, as my second supervisor during my first clumsy attempt at being a chemist.

I would like to thank all my colleagues at ICOS and CBCB, especially Jurek for designing the sequences and the origami but also Nunzia, Emanuela, Ben and Harold, for sharing the burden of science with me, for helping me every day with their knowledge, time and patience (and also for the BBQ parties). A special mention goes to Prof. Zenkin and Prof. Murray, for their insight on bacterial transcription.

A special thank goes to Prof. Martin Hanczyc for teaching me how to make vesicles, for hosting me at Univeristy of Trento during my second clumsy attempt at being a chemist and for his unlimited positive energy.

I would like to thank Prof. Stimming, Jingying and Baohua from the School of Chemistry for their guidance and support at the AFM.

I acknowledge the Faculty of Science, Agriculture and Engineering for funding my Ph.D. through the Doctoral Training Award program.

I would like to thank my family for supporting me even from one thousand miles away.

And finally I want to thank Alessia, the constant of my journey, for being always at my side and navigate through the good and the rough paths, for believing in me even when I didn't and for showing me that together we can always make it.

Index

List of Figures

List of Tables

Part I

Chapter 1

Introduction

In 1959 Richard Feynman envisioned the idea of nanotechnology in his lecture "There's plenty of room at the bottom". In his futuristic talk Feynman described the possibility to arrange atoms in a controlled manner, to build medical nanodevices ("swallowing the doctor") and miniaturized factories. Even if 20 years in advance, he managed to foresee the technological revolution that happened in the 1980s when finally scientists were able to control the matter down to the single atom, allowing us to write a book smaller than the head of a pin. In 1986 K. Eric Drexler added the concept of molecular nanotechnology to the equation, surely inspired by Feynman's talk and by the growing knowledge of biological systems. One of its most discussed aspects is the idea of a replicating nanorobot, or self-replicating nanomachine. Different interpretations of this idea can be found in disciplines once considered very distant: from the cellular self-replicating machinery in biology, to the Von Neumann universal constructor in computing science.

In my dissertation I will try to bring these concepts together using the tools of biotechnology, synthetic biology and nanotechnology with the support of computing science provided by our interdisciplinary team.

To do so I spent my Ph.D. working on one of the most exciting topics in the recent years: nucleic acid *origami*, a technology that uses nucleic acids as construction material to produce a wide array of designs in the nanometric scale. This technique, developed in 2006 by Paul Rothemund, permits the building of complex nanomachines using DNA and RNA in a bottom-up manner. It greatly improved the previous attempt to control DNA assembly and immediately gained the scientific spotlight on a Nature cover. Nucleic acid nanotechnology does not see the genetic material just as data storage system, it also considers

the chemical and physical characteristics of these versatile molecules: the high specificity of the Watson and Crick base pairing, the bio-compatibility, the controllable secondary structures and of course the growing opportunities in synthesis and modifications. The costs for the synthesis are following Moore's law and the costs for the sequencing are dropping even faster, making nucleic acids a promising material for the next breakthrough. Being a biological component, nucleic acids can also rely on the natural replication system inside the cell, a replication system to which we gained access and control through well established biotechnological tools. Compared to chemical synthesis, the cell replication machinery is even more efficient, granting a potential output far greater than our actual necessities.

The tools I just described are extremely powerful, nonetheless a key factor is still lacking to complete the picture, that is the implementation of the various components on a common platform. The long-term vision behind my dissertation is to use the bacterial cell as the platform, exploiting its internal systems as a nanodevice factory, integrating in its very own genome the tools to control it, to achieve a *biological universal constructor*. This idea reverses the common approach to the Von Neumann universal constructor, i.e. adding a replication mechanism to a cellular automata, and instead the goal becomes implementing a control system into an already replicating organism and transform it in a programmable universal constructor.

One of the greatest challenges is given by the extreme complexity of the cellular environment. Even though our knowledge on -omics is evolving at a fast pace, it is still very difficult to predict all the connection between cellular pathways. The solution designed to tackle this problem is a step-by-step approach, optimising the components *in vitro*, testing them into a synthetic system and finally integrate everything *in vivo*. This procedure will grant an easier way to identify problems and bottlenecks, it will also allow to use the matured knowledges in a modular fashion, as components for other applications.

1.1 Problem statement

Nucleic acid origami great potential is starting to be translated in industrial and medical applications. Nonetheless there are still many issues that could hinder its biotechnological value, amongst them the most relevant I can identify are: the widespread use of natural

genomes as scaffolds, the fact that molecular dynamics ruling its folding are still partially unclear and the lack of an *in vivo* origami synthesis route.

The first issue, the use of natural genetic material, derives from the difficulty of generating a long single strand DNA (ssDNA) molecule. M13 phage genomic DNA was the first source of scaffold used for DNA origami [1]. It is a native ssDNA molecule and this is its greatest advantage, it is also commercially available and it is very cheap to produce, the phage infects *E. coli* injecting it with its genome, and replicates it and its structural proteins inside the bacteria. All these characteristics are very good from a research-oriented point of view and makes it the most widespread scaffold. At the same time it carries great disadvantages when it comes to applicability. The main issue is given by the presence of many repetitions, these repeated sequences can be long up to 42 nucleotides and they cause misfolding of the nanostructures by presenting more than one binding site for a single staple [2]. When this happens it lowers the yield of the origami and sometimes it could be very difficult or impossible to separate the correctly formed origami from the misfolded ones. Natural sequences can also form secondary structures with very high thermodynamical stability that interfere with the staples binding. So far the best way to perform an origami folding using a natural scaffold requires a first step of denaturation at very high temperatures and then a slow gradual cooling down to avoid the formation of kinetic traps. This protocol is obviously not compatible with many living organisms, especially those of industrial interest. Finally, natural scaffolds are encoding active genetic material: genes and regulatory sequences. An M13 scaffold or other active genetic material can be a cause of safety concern if we want to design *in vivo* medical applications, and it is an even bigger problem if we want to insert the origami in living bacteria. Mutagenesis of these sequences can be only a partial solution because some of the genes, especially those related to the phagic replication, cannot be changed without compromising the scaffold synthesis. Other solutions considered the use of plasmid sequences as scaffold but in this case the production is time consuming, less efficient and still carries over some biologically active sequences like origin of replication and antibiotic resistances [3, 4]. An ideal solution would be to have an optimized fully synthetic sequence that could be tailored each time to the desired application. Sequences that would not interact with cellular components in unspecific ways. designed to perfectly fit the desired structure.

The second major issue is easily identifiable as an opportunity: a deeper knowledge of the thermodynamical processes involved in the folding is not only of great scientific interest but could also dramatically improve the transfer of the whole technology *in vivo*. Our group is already active in the optimization of synthetic scaffold sequences and the development of a computational origami folding model. To further improve this effort it is necessary to develop a biological system capable of generating collectible data, a system that could analyze the folding as it happens and detect the oddities and the unknown factors. An ideal tool would allow to characterize in which order and at what temperatures the staples are folding the scaffold.
The last critical step towards the *in vivo* folding is the understanding of the relationship between the origami components and the hosting cell. Optimal *in vitro* conditions are extremely different from those found in any cytoplasm, prokaryotic or eukaryotic, we cannot expect that the known protocols will work without major adjustments. The passage through these systems could take advantage from a synthetic biology tool that allows a progressive transposition granting the ability to work on the single issues without the "background noise" generated in a complete living cell.

1.2 Aims and objectives

The main goal of my dissertation is the advancement of nucleic acid origami technique towards the *in vivo* production and application, specifically into *E. coli* bacterial cell. To reach the final goal I proceeded in a series of secondary subtasks that would be necessary to, in the long run, achieve the "holy grail" of a universal constructor.

1. Application of a De Bruijn designed scaffold to fold a nucleic acid origami: demonstration of the feasibility and applicability of a fully synthetic scaffold sequence. This task addresses the issue of using a synthetic bio-orthogonal scaffold for advanced applications.

2. Production of an hybrid RNA scaffold / DNA staples origami: the RNA scaffold transcription is a step forward towards the *in vivo* synthesis. This task addresses both the issues of using a synthetic bio-orthogonal scaffold and the production of $in-vivo$ origami components.

3. Testing of the De Bruijn scaffold behaviour *in vivo*: the scaffold is tested inside an *E. coli* cell to prove the bio-compatibility addressing the necessity for bio-orthogonality.

4. Design and testing of a De Bruijn origami reporter: the origami carries a fluorescent molecular system that reports the folding in a real-time fashion. This experiment addresses the needs for a better knowledge of the thermodynamic laws ruling the origami folding.

5. Folding of the De Bruijn origami reporter inside a phospholipidic vesicle: the cell-like system allows the study of the molecular events mimicking the *in vivo* conditions but free of its perturbations. This complex task, in its variants, covers multiple issues: it will allow to improve the thermodynamics knowledge, it will help to confirm the bio-orthogonality of the synthetic sequences and it will accelerate the *in vivo* origami research.

1.3 Structure of the dissertation

My thesis is composed of 3 main parts.
The first part includes: the analysis of the previous research on nucleic acid nanotechnology and giant unilamellar vesicles, the literature review and a chapter on the experimental and computational techniques I mastered during my Ph.D..
The second part is composed of four chapters and is focused on the experimental results: one chapter about the DNA origami folded using a De Bruijn sequence generated scaffold, one about the RNA/DNA hybrid origami using a thermodynamically optimised De Bruijn scaffold, the third chapter contains the analysis of the previously used De Bruijn scaffolds, cloned and transformed in a set of *E. coli* strains and the fourth chapter describes the origami analysis tool based on the use of artificial liposomes.
The final part is the conclusion of my thesis, where I discuss the results and the future opportunities of research provided by my work.

1.4 Main contributions

This research is a the true connector between the fields of bioinformatics and biology. My work complements the PhD dissertation of Dr. Kozyra [5] which dealt primarily with the

computational aspects of the origami folding, e.g. optimization via genetic algorithms. My work is mainly concerned with the experimental challenges arising from Aims and objectives 1.2. The application of the theory confirmed furthermore the utility of non biological sequences, showing excellent folding ratios and improved thermodynamic profiles. These findings can easily be exploited both in research and industrial applications. My work advanced then in the *in vivo* field. Cloning the synthetic scaffold in a bacterial vector and expressing it inside *E. coli* is an important step towards the *in vivo* folding. Different scaffolds gave different results showing the presence of factors still unknown and worth of future research.

The vesicle-origami experiment generates a tool with a high potential research output. This system can improve the understanding of the thermodynamic processes involved in the origami folding giving a major contribute to synthetic biology, bioinformatics and physics fields. This is the fruit of a highly interdisciplinary work set-up in collaboration with Prof. Hanczyc at the University of Trento; connecting origami studies with the vesicle protocol pulls together the fields of synthetic and artificial biology, chemistry and origin of life.

1.5 Published work

Journal articles

1. Designing Uniquely Addressable Bio-orthogonal Synthetic Scaffolds for DNA and RNA Origami Jerzy Kozyra, Alessandro Ceccarelli, Emanuela Torelli, Annunziata Lopiccolo, Jing-Ying Gu, Harold Fellermann, Ulrich Stimming, and Natalio Krasnogor ACS Synthetic Biology 2017 6 (7), 1140-1149 DOI: 10.1021/acssynbio.6b00271

Others

1. Jerzy Kozyra, Chien-Yi Chang, Alessandro Ceccarelli, Harold Fellermann, and Natalio Krasnogor. "Programming synthetic scaffolds for DNA origami", Extended abstract printed for ECAL satellite workshop: Toward Programmable Biology, pp. 12-13, York, UK, 2015

2. Jerzy Kozyra, Alessandro Ceccarelli, Emanuela Torelli, Annunziata Lopiccolo, Harold Fellermann, Ben Shirt-Ediss and Natalio Krasnogor. "Toward *in vivo* nanostruc-

tures: synthetic origami scaffold based on De Bruijn sequence", Extended abstract printed for: Bio Origami - designed bionanostructures, from nucleic acids to proteins and beyond conference, p. 36, Ljubljana, Slovenia, 2017

3. Alessandro Ceccarelli, Jerzy Kozyra, Emanuela Torelli, Annunziata Lopiccolo, Harold Fellermann, Ben Shirt-Ediss and Natalio Krasnogor. "Toward *in vivo* nanostructures: hybrid origami folded on De Bruijn sequence scaffolds", Extended abstract printed for: Bio Origami - designed bionanostructures, from nucleic acids to proteins and beyond conference, p. 37, Ljubljana, Slovenia, 2017

Chapter 2

Background and related work

2.1 Introduction

In this chapter I will introduce the topics discussed in my dissertation and analyse the relevant literature summarising the state of the art.

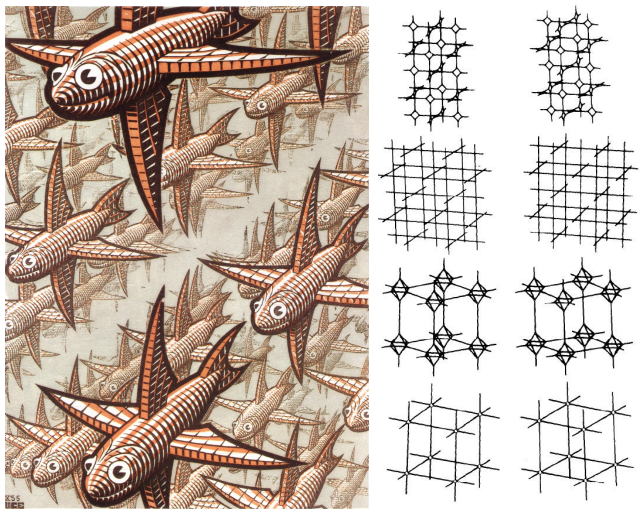
2.2 Literature survey

2.2.1 Nucleic acid assemblies

Nucleic acid engineering is a promising field, due to the intrinsic characteristics of the molecules themselves: biocompatibility, ease of replication and ease of further functionalising. In addition the generic advantages of the bottom-up approach [6], has made it one of the key topics of recent molecular research. In particular, nucleic acid assembly is one of the most interesting and fast developing technologies [7, 8]. Nucleic acid origami are the main focus of my thesis. The technique, first published in 2006 [1], has deep roots in the previous research on nucleic acid assemblies and nanostructures.

Going back to 1982, Nadrian Seeman published the first paper on nucleic acid junctions and the related lattices, inspired by the DNA replicational junction and the recombination process in the form of the Holliday junction [9]. In this paper Seeman shows how it is possible to control the shape of engineered nucleic acids molecules just exploiting the Watson and Crick pairing. The DNA sequence is not encoding "only" the genetic data, but it also encodes the shape that the molecule will assume. Inspired by Escher's painting "Depth", Seeman proposed also a theory of three-dimensional lattices, built using nucleic

acid junctions connected by sticky-ends oligos, it was the beginning of the DNA based nanoconstructs.



(a) (b)

Figure 2.1: 2.1a - M.C. Escher's painting "Depth". 2.1b - Seeman's DNA junctions lattices. From reference [9]

The technique was soon refined by designing immobile junctions to stabilize the whole structures [10]. The result was achieved by eliminating all the sequence symmetries at the centre of the junction, so the strands cannot migrate on the adjacent bases. It was in 1993 that Seeman described the DNA double crossover molecules [11]. A double crossover is a DNA structure where two Holliday junctions are connected by two helical arms [Fig. 2.2], thus, creating a much stabler DNA structure with a well defined structure and not prone to any branch migration effect. This paper marked the second milestone in the evolution of the technique and provided a push to the whole research field.

In 1996 the double crossover was applied as a nanoconstruction component [12]. In a few years a plethora of new designs was published and the techniques were refined [13]. The rigidity granted by the double crossover molecules allows the formation of stable DNA tiles, structures used as bricks in the building of biological self-assemblies [14]. At first the tiles permitted the growth of geometric and periodic lattices but in a few years, with new algorithm-driven designs, this technology allowed also the generation of aperiodic and well defined structures [15, 16, 17, 18, 19, 20, 21, 22].

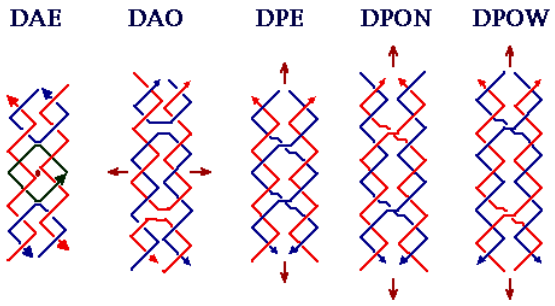


Figure 2.2: Graphic representation of different DNA double crossovers. From reference [11]

2.2.2 DNA origami

A major breakthrough in programmable nucleic acids engineering was made by Rothemund in 2006 with the first DNA origami. This assembly technique involves the use of a long single strand DNA molecule, the scaffold, which is folded in the desired shape by the annealing of multiple smaller single strands DNA oligonucleotides, the staples, thus recalling the Japanese art of folding paper [1] [Fig. 2.3]. Origami folding has many advantages over the previous assemblies, the most peculiar involving the reaction dynamics: the origami works with an excess of staples over the scaffold and does not suffer from any imbalance between the oligonucleotides, while the older DNA assembly protocols require a very strict stechiometric ratio between all the components. Moreover the folding is obtained following a simple one-pot reaction protocol: all the components are mixed together, heated up to a denaturing point and then slowly cooled to room temperature. With this technique it is easier to build nanostructures of arbitrary shape and pattern, and the fine placement of each staple grants a nanometric resolution, with the possibility of "decorating" each one of the staples.

Rothemund used M13mp18 genome, a phage, as scaffold for the origami. Being already single stranded, a viral genome was a convenient scaffold for the origami folding in terms of preparation and costs leaving to the synthesis only the shorter staples.

The design steps were computer-aided since the beginning and only the scaffold shape was designed "by hand", first creating a block diagram with each block representing one helix turn and then deciding the raster fill arrangement. All the staples related designs were completed by CAD.

Such a powerful tool gained a central spot in the nanotechnology scene and soon many

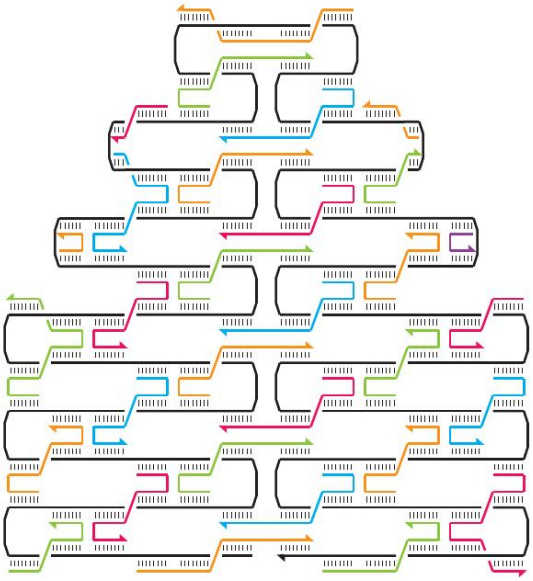


Figure 2.3: A graphic representation of an origami nanostructure. From reference [1]

groups started researching new upgrades and applications for this method.

Another groundbreaking evolution was published in 2009 when Douglas, in William Shih's group, folded the first 3D origami [23]. The 3D origami was folded by arranging the scaffold in a honeycomb pattern, formed by the antiparallel helices sequentially angled by 120°[Fig. 2.4]. To achieve this result, the same lab developed a new CAD software (caDNAno) that allows an easier design process for much more complex DNA structures [24]. The software became one of the most used in the publications that followed, for its intuitive and easy-to-learn user interface, and was soon integrated by a second software, CanDo [25], able to predict the shape and stability of an origami design.

It must be noted that other groups folded DNA origami 3D structures in the same year [26, 27] as the goal was considered of utmost importance for the evolution of the nanostructure applications. Soon many new designs were published, defining new 3D structures like nanotubules [28]

Shih's team soon managed to generate a twisted 3D origami structure acting on the number of base pairs between the crossing-overs [29]. DNA B helix has a 10.5 bp/turn conformation, it means that, in a honeycomb lattice arrangement, two anti-parallel strands can form a crossing-over every 7 nt (or every 120°). Reducing or augmenting this number forces the structures in a twisted conformation and could form angles up to 180°. Benson

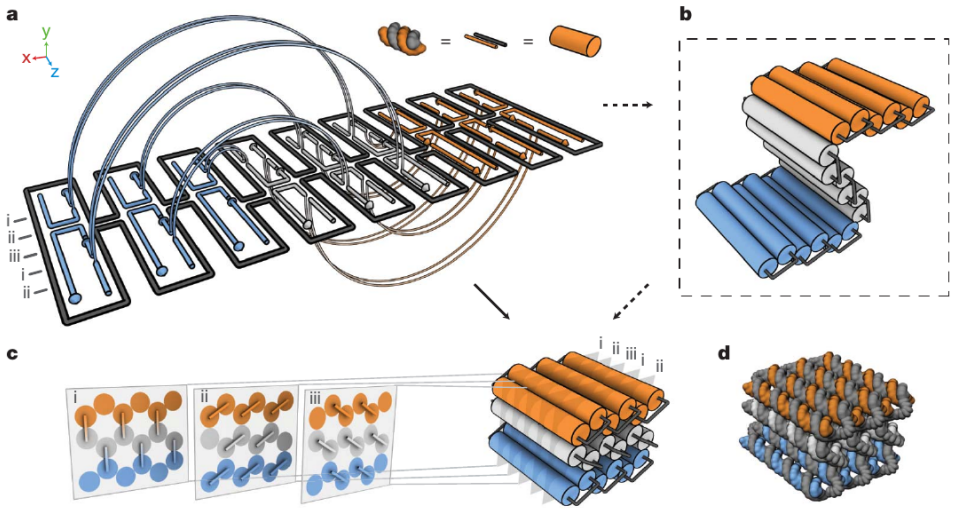


Figure 2.4: (a) 3D origami folding schematic. (b) and (c) Cylinder model of folding structure. (d) Atomistic model of 3D structure. From reference [23]

developed in 2015 an innovative 3D origami design [30]. Instead of using the classic honeycomb raster design, that generates bricks filling the volume, they used a triangle-based mesh that forms the surface of the origami leaving an empty volume on the inside. The algorithm is designed to fold the scaffold in a linear way, avoiding the use of staples connecting two distant points of the scaffold, and it can be considered as a viable option for an *in vivo* folding, allowing the folding of an RNA scaffold as soon as it is transcribed from the polymerase.

Over the subsequent years, many groups worked to improve origami folding protocols under every aspect, aiming to find a better scalability that could boost the interest in the applicability.

The original origami folding protocol can be considered quite extreme, especially if we take in consideration the high denaturing temperatures necessary at the beginning of the procedure, the particular ionic concentrations and the long times required by the temperature ramps, that can be longer than a week for 3D origami. To overcome these limitations many groups designed new protocols capable of solving one or more problems. The addition of denaturing agents, as formamide and urea, was one of the first attempts to achieve an isothermal folding [31]. On the other hand the purification steps proposed involved toxic reagents and dialysis, requiring longer times and more difficult procedures. Sobczak et al. managed to set up an isothermal protocol analysing the folding/unfolding ratio through fluorescent dyes [32]. With this new protocol a 3D origami can be folded

in two hours rather then two days, as previously required. This protocol achieved also a higher yield of origami reducing the misfolding and the degradation resulting by long incubation at high temperatures. In this research the initial denaturing step is still used and not investigated. A further evolution of this protocol can be found in the work of Kopielski et al., in their protocol the lowest starting temperature is found by a trial and error procedure, realising that there is no need for denaturation in a 2D origami folding using Sobczak protocol [33]. Furthermore, using the additive betaine, they managed to reduce the folding temperature to 37°. Chopra investigated the addition of different molecules other then magnesium or sodium salts. They folded different origami using spermidine, a molecule know for DNA aggregation inside cells and viruses, spermidine made the origami less salt dependent and resistant to electric fields as those used in electrotransfection [34]. Even if many different protocols are newly designed, the mechanics behind the origami folding are not yet explained. An experiment based on a scaffold with multiple correct foldings demonstrated though that the folding process follows a pathway [35]: a shape is preferred over the others depending on the interaction of the staples in a long range and locally.

The purification, ignored in the original publication by Rothemund [1], was initially performed through agarose gel electrophoresis extraction [36]. Later, new purification methods were developed, aimed at improving the quality of the final product: PEG precipitation [37, 38], ultracentrifugation [39], size exclusion by gel filtration [40], spin-column filtration [41], magnetic bead capture and FPLC [42]. Each of the methods presents pros and cons, and their suitability varies with the necessities.

DNA scaffold evolution

One of the most relevant aspects of the origami applicability passes through the scaffold design and production. When Rothemund invented the technique, M13mp18 genome was undoubtedly one of the cheapest and most convenient sources for a long single stranded DNA molecule [1]. A limited choice availability, to a single scaffold, limits the range of the features that one can design, and the use of genomic material also poses some limitation on the possible customizations of the sequence. Nonetheless phage DNA production scaling-up has been addressed, as the request of single strand DNA increased with the origami popularity [43].

Rolling circle amplification (RCA) is a replication system borrowed from phages. It has been used to scale up the amount of origami components without recurring to solid-phase synthesis. The first protocol has been designed just for staples synthesis. It can be performed both *in vitro* and *in vivo* just changing few steps [44]. Later a new protocol was0 published, capable of synthesizing both scaffold and staples at milligram scale, using a phagemid (a vector derived from phages) and a helper plasmid. The components are encoded in the phagemid and spaced out by self-cleaving DNAzymes cassettes. When the RCA synthesize the single strand DNA, the self-cleaving cassettes are activated by zinc salts and cut themselves out of the sequence leaving the clean staples and scaffold in solution. The whole system is cloned in *E. coli* and can be scaled up with the bacterial culture [45].

The use of an hybrid phage DNA, fusing λ phage with M13, brought the size of the scaffold up to 56 kilobases, though this scaffold is more prone to degradation than the much shorter M13 [46]. The limitations in the origami size were also tackled by the development of an advanced design algorithm capable of folding a fractal assembly of different tiles from the same scaffold [47]. Though it works only on 2D origami, to a maximum of 8 by 8 tiles assembly, it can allow to fold a structure up to 0.5 square micrometres with arbitrary patterns encoded on their surface.

Strategies that avoid the use of phage DNA have been also developed. The use of PCR products or enzymatic digestion of plasmids were amongst the first attempts. PCR allows the amplification a sequence of double stranded DNA from almost any template sequence up to 35 kilobases [48]. It requires a second step to generate a single strand scaffold, and different groups found different solutions to this problem, such as labeling one strand with biotin and removing it after chemical denaturation [3], or selectively digesting the complementary strand with a nuclease [4]. The digestion of a modified plasmid that encodes a nicking enzyme site [49] generates a single stranded circular scaffold. Using high copy numbers plasmids gives access to a scalable method to produce higher levels of starting material than PCR. The use of double stranded DNA as scaffold has been studied too. It has been used to fold two different origami, one for each strand [50], and as a single origami structure with the two scaffolds converging without regenerating the original double strand [51]. This strategy could drastically increase the availability of origami sources but it did not prove to be efficient enough for larger or more complex designs with a high

percentage of the scaffolds re-annealing back to the double strand conformation.
Many strategies were adopted to overcome the limitations imposed by the different scaffolds.
Supramolecular structures are one of the solutions to increase the size of the objects built using origami. For example a tripod, built on a slightly modified M13mp18 scaffold, was used as a brick to build regular polyhedra [52]. Interestingly, in the same article, a new method is described for the detection of large origami using fluorescent ultra-resolution microscopy.
Other tiles-like systems have been published for both 2D and 3D origami [53, 54]. Another solution involves the stacking of different origami exploiting non-base pairing at particular buffer conditions [55][Fig. 2.5]. With this method it is also possible to design a dynamic structure controlling its shape by varying cationic forces and temperature. The evolution of this technique, improving the rigidity of the DNA building blocks, allows the control the formation of structures up to 1.2 gigadaltons [56].

A different approach was taken using a DNA origami as frame to assemble a DNA tile system. With this method, the origami creates just a perimeter while the area is filled by the tiles. This hybrid allows to increase the area of the assembly compared to an origami with the same scaffold, at the same time the origami provide for the finite and controlled shape of the assembly [57].

2.2.3 TectoRNA and RNA origami

Around 20 years after Seeman's first publications on DNA assemblies the research on nanoconstructions opened also to RNA. This branch, called tectoRNA, exploits the many advantages associated with the RNA molecule. The most relevant difference with the majority of DNAs is the natural single stranded conformation, like the bacteriophage DNA, RNA is ready to participate in an assembly. RNA can take advantage of the greater interactivity typical of its secondary and tertiary structures and additional features can be added to the assemblies, like hairpins, junctions and kissing-loop interactions, while adapting the building strategies previously developed for DNA [58, 59, 60]. The advantages of RNA are not limited to its structural versatility, when it comes to applicability RNA shows a set of very attractive features as: an easy synthesis from a DNA template with a promoted transcription, reducing costs and maximizing yields; a larger nucleobases

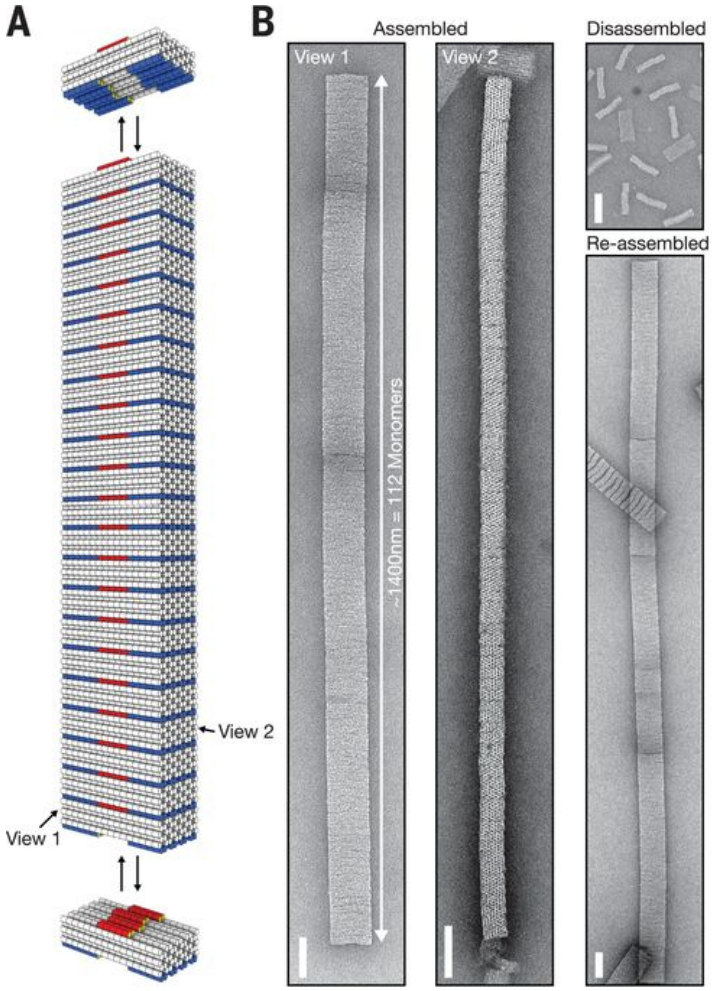


Figure 2.5: (a) 3D origami stack schematics exploiting non-base pairing. (b) TEM imaging of the stack. This technique takes advantage of the base stacking to connect multiple origami in a 3D superstructure. From reference [55]

pool that allows greater pairing flexibility; better thermodynamic and pH stability that increase the suitability for practical applications both *in vitro* and *in vivo*; and the opportunity to add biologically active sequences as ribozymes and sRNA [61, 62][Fig. 2.6].

Following the example of the DNA nanotechnology, new RNA assemblies were designed. Initially simpler 2D structures were inspired or copied by natural RNA motifs, like T and L junctions from transfer RNAs and ribosomial RNAs [63, 64]; soon evolving in more complex assemblies with 3D structures [65] and functionalized strands [66].
To overcome the limitations of the tile structures, the technology of origami was developed also for RNA. At first different groups arrived to the same result, an RNA scaffold folded using DNA staples. Endo used a section of Enhanced-GFP coding region as template for the scaffold [67], while Wang used a fragment of pUC19 plasmid [68]. Both groups

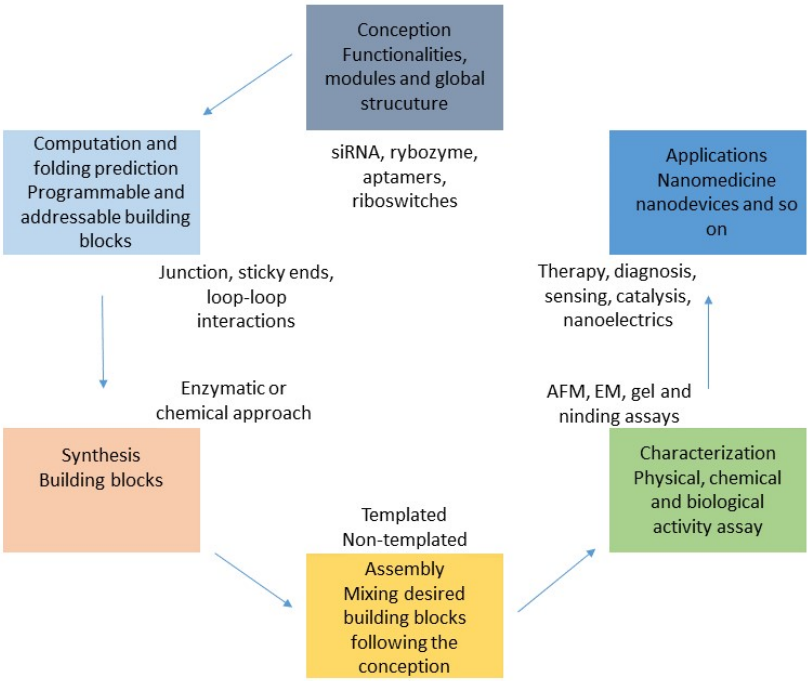


Figure 2.6: Design steps for RNA nanotechnological devices.

obtained the folding of the origami adapting the structure of the helix from the B conformation of DNA-DNA interaction (10.67 bp per turn) to the A helix of the RNA-DNA hybrid (11 bp per turn).

Endo prepared the first RNA origami [69], this time both the scaffold and the staples were transcribed by T7 polymerase, the short DNA templates were modified with biotin to remove them after the staple synthesis using streptavidin coated magnetic beads. The optimal folding conditions and origami size matched those of RNA-DNA hybrid origami.

2.2.4 Origami applications

Many of the applications stem from the origami technology: the more theoretical proof-of-concept studies, aimed to showcase the capability of the new techniques, like the DNA motor origami track [40] or the single nucleotide polymorphism visual detector [70], and the advanced biophysical and biomedical applications. These technologies take advantage of the extreme variability permitted by the functionalization of the staples in a single origami [71]. Various versions origami carriers for anti-cancer drugs are now in the spotlight [41, 72]. In these systems staples are modified to target a specific cell type, to control

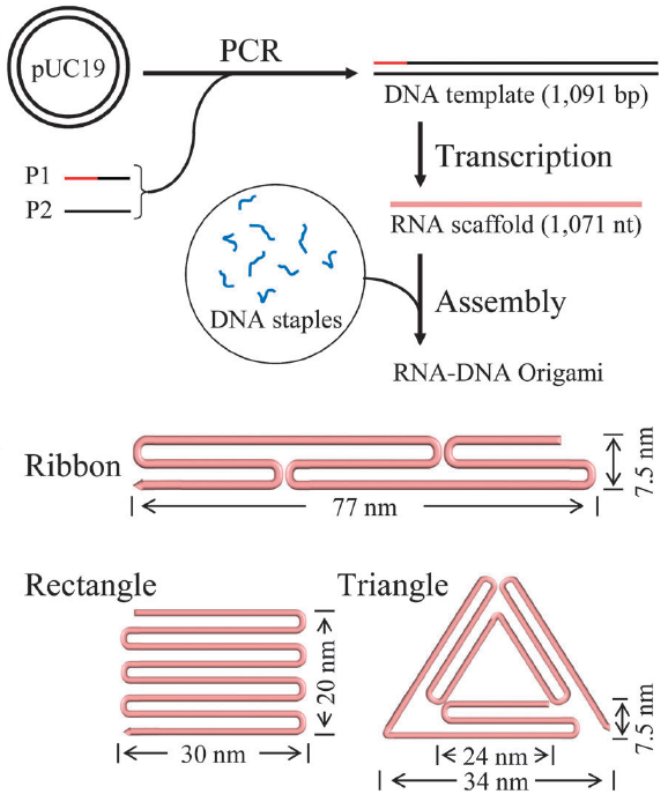


Figure 2.7: Scheme of the origami folding using an RNA scaffold and DNA staples. The scaffold was encoded in a pUC19 plasmid and amplified with a PCR, adding a T7 promoter through the primer. The PCR product was then transcribed to generate the RNA scaffold. DNA staples oligos were added to fold the final origami. From reference [68]

the opening and closing of the origami boxes, to bind and include the therapeutic drug inside the origami. These particular nanorobots take advantage of the biocompatibility of DNA to mask the drug presence to the body. The drug can not interact with the system until it's released, granting a double effect: the drug is targeted more finely to the right tissues, reducing undesired collateral effects, and at the same time cancer cells can not activate the drug-resistance systems until it's too late to overcome the therapeutic activity. Another origami interactive container has been designed by Takenaka [73]. In this case the box is capable of capturing gold nanoparticles from the buffer solution, and the origami lock is photoresponsive, it opens upon UV irradiation and closes with visible light. Other DNA containers have been designed to achieve different functions as compartmentalization or controlling enzyme reactions [26, 74, 75]. DNA origami has been designed also to interact with bacteria, in this case the structure is a DNA pyramid decorated with glutathione-protected gold nanocluster (fluorescent reporter) and intercalated

with Actinomycin D antibiotic. This delivery system increases the uptake of antibiotic in *E.coli* and *S.aureus* by 50% [76]. Torelli designed a nanorobot capable of exposing a payload, a DNAzyme, upon the hybridization to a single strand activating sequence [77].

Applications for origami can be found not only in the biological field but also in biophysics where the technique has been exploited to build a device that generates plasmonic resonance of gold nanoparticles arranged with nanometric precision [78]. Another study showed how DNA origami nanotubes could help the stabilization of membrane proteins during NMR [37]. Given their nanometric precision, origami have been used as standards and calibration tools in super-resolution microscopy and AFM [79].

Origami have been used also as platforms in nanotechnology and synthetic biology. He built a self replicating system based on DNA rafts [80], the rafts can dimerize starting from a single seed dimer, when cross-linked using UV sensitive molecules as linkers, the newly formed dimers act as seeds and the growth become exponential. Taking advantage of pH sensitive secondary structures, the system can be calibrated to prefer the dimerization of specific tiles in a mixed environment.

One of the main questions related to the possible *in vivo* applications is the behaviour of the origami inside a living cell. Mei et al. partially answered this question in 2011 showing how a DNA origami resists in a cell-free extract from a mammalian cell line, up to 12 hours [81]. The experiment showed also that the origami can retain a specific interaction activity with proteins. Different origami designs and cell lines were tested, the results support the applicability as a detector for proteomics and intracellular environments.

2.2.5 *In vivo* assemblies and origami

The idea of exploiting the origami *in vivo* propelled the research toward the nanostructure synthesis directly inside the cell, the first *in vivo* application of a nucleic acid assembly is achieved by Delebeque in 2011. This is an important breakthrough because the oligonucleotides are encoded in the cell through a plasmid and the nanostructures fold inside the cell. The assemblies are functionalized with protein binding aptamers that bind target oligopeptides fused to the enzymes of interest. The organizing activity is shown inside an *E.coli* living cell by coupling the two halves of a split GFP back to its functional structure, and augmenting an enzymatic pathway output spatially organizing two of its components on the same scaffold [82].

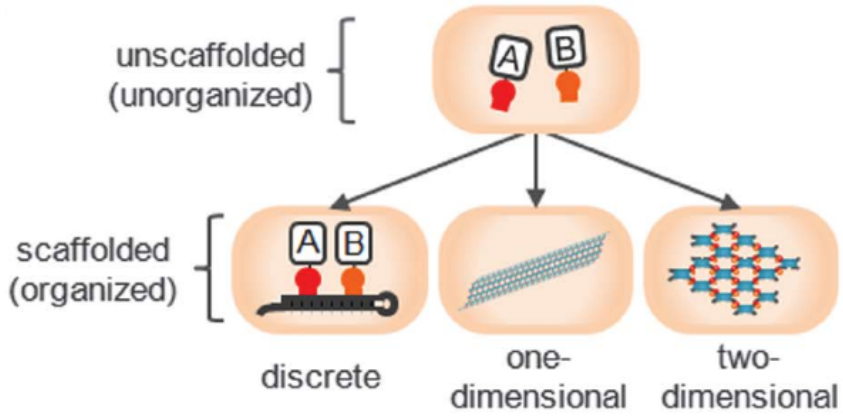


Figure 2.8: Graphic representation of the *in vivo* RNA scaffold. A and B are the two components organized on the assembly. The assembly can further generate one-dimensional (nanotubes) or two-dimensional (mesh) structures. From reference [82]

Based on this work, Sachdeva improved the system bringing the number of organized enzymes up to four, using larger RNA tiles and optimizing both the aptamers and ligands structures, obtaining an increase of 80% in a pentadecane production pathway in *E.coli* [83], both of these works use a tiles assembly system, bringing over its limitations compared to origami, most importantly: the necessity of a strict stechiometry amongst the different components and the generation of non-discrete structures.

The nature of RNA brought Geary and Andersen to speculate on the possibility of generating a tile-like structure just using a single self-interacting RNA scaffold. They designed different configurations of it [84] and tested them *in vitro* [85]. This assembly does not use any staple, instead it uses a single molecule of RNA that folds like a protein, as soon as a domain is transcribed it folds on itself forming a secondary structure, in this way the RNA molecule is more resistant to degradation. Once the secondary structures are all formed, they interact to generate the tertiary structure that can be 2D or 3D. Not requiring an initial denaturation, it is possible to perform a cotranscriptional folding in a single step. Based on the same concept, in 2018 another single molecule RNA assembly was designed, this time *in vivo*. The molecule is expressed through a plasmid in *E.coli*, then extracted by phenol or sonication and the structure is analysed by atomic force microscopy and cryo-EM [86]. Unfortunately no reporting system was used to demonstrate the folding inside the cell. These single molecule RNA structures are very different from

origami, their convenient folding procedure comes only after a molecular design that is specific for every single sequence and there are no softwares comparable to CADnano to simplify this procedure. Unlike the origami it is not possible to use the same scaffold to produce different shapes and every time it requires at least a rearrangement of the domains [87]. To make the design process easier, the critical junctions and structural corners are often copied from biological sequences found in various organisms, including viruses and bacteriophages [59]. On the contrary the origami shape can be changed with a new set of staples with minimal effort, or it is possible to change its function substituting only the functionalized staples and keeping the shape unaltered. The origami high number of components, if from one side makes the folding more difficult, on the other allows for a larger capacity of carrying active components, Elbaz designed a system to synthesize single stranded DNA filaments in *E. coli*, up to 205 nucleotides long, using HIV and murine leukemia reverse transcriptase. An assembly crossover made by four strands was folded both *in vitro* or directly *in vivo* [88].

2.2.6 Giant unilamellar vesicles

The final part of my Ph.D. involved the use of a technology used in a different research field than the previously described: the giant unilamellar vesicles (GUVs), a particular kind of artificial liposomes. The literature on vesicles and micelles is extremely vast and I will summarize only the aspects more relevant to my work.

This technology is of great interest in the study of bacterial division [89] and membrane proteins [90], the synthesis of cytomimetic systems reproducing cellular conditions [91, 92], the studies on the origin of life [93, 94] and the research in unconventional computing [95]. The first protocol for artificial lipid vesicles was published in 1969 [96], this method produces multilamellar vesicles, that consist in multiple concentric lipidic bilayers. In the following years new synthesis processes were developed, capable of producing unilamellar vesicles, single lipidic bilayers containing a controlled aqueous solution. Amongst them is the inverted emulsion protocol published in 2003 by Pautot [97]. The procedure consists in the passage of an inverted emulsion (aqueous droplets surrounded by a monolayer of lipids) through a lipid/aqueous interface, because of the amphiphilic behaviour of the lipids the second layer is captured around the droplet and the liposome is formed [Fig. 2.9]. A more detailed description of the protocol is found in section 3.8.

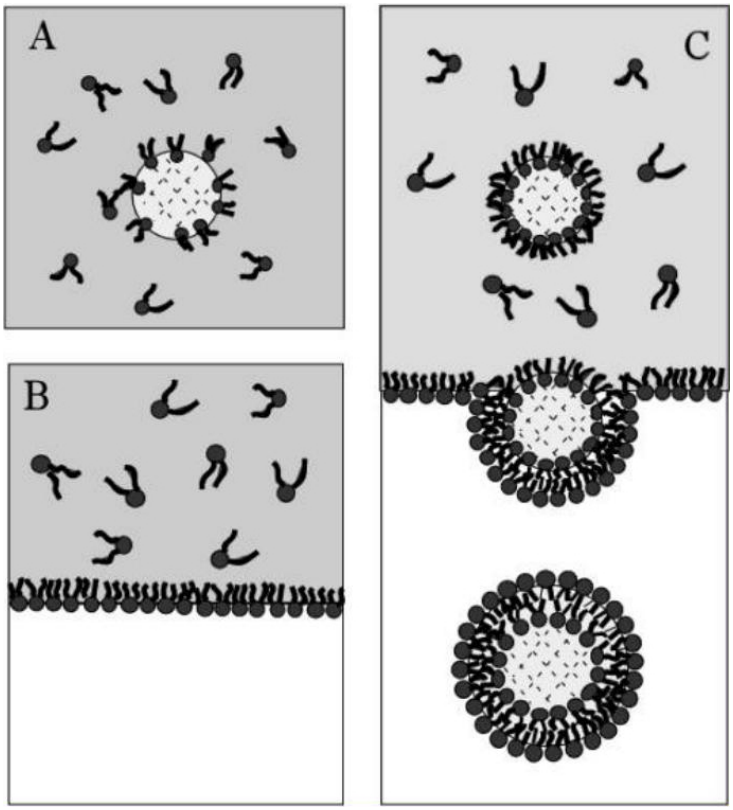


Figure 2.9: Giant Unilamellar Vesicles synthesis using the inverted emulsion protocol. A - The inverted emulsion is formed (droplet) preparing the first lipid layer. B - Preparation of the aqueous-lipidic interface with the second lipid layer. C - The droplet passes through the interface from the lipid to the aqueous phase, forming the lipid bilayer. From reference [97]

The main advantage of this technique is the simplicity of the procedure that does not require any special instrument. On the other hand all the steps require extra care and must be executed in a very precise manner or the vesicles will not form, also, small residues of the oil solution could contaminate the final product [89].

The product of the inverted emulsion protocol is a polydispersed mixture of GUVs with dimensions that vary from 10 μm to 100 μm. The size of GUVs allows the use of optical methods for their analysis.

Because of their composition and shape, GUVs has been used as compartments to perform *in vitro* experiments simulating a cell environment without the background activity of a living cell, Nishimura performed a GFP transcription inside a GUV containing a solution of mRNA and cell-free extract [91].

Artificial vesicles are also used for the study of the membrane fusion processes, many fusion protocols are described in literature, and I will only focus on those I consider relevant

for my research.

One of the first experiments involved the use of calcium ions to facilitate the vesicles fusion, with the simple incubation of the vesicles in a solution containing $Ca(Cl)_2$ [98]. Other metal ions were studied and characterized, in particular Mg^{2+} [99], that is of particular interest for its use in origami folding. A similar method was developed using a polyethylene glycol (PEG) solution [100]. In all these cases the vesicles seem to fuse for the dehydrating and effect of the fusing molecules and, in the case of metal ions, for the neutralizing effect on the repulsion of the headgroups charges, enhancing the contact between the membranes. The fusion between oppositely charged vesicles has also been observed and characterized through microscopy [101] and flow citometry [102]. These vesicles are synthesized using a mix of neutral and charged lipids, and the different compositions naturally fuse if close enough.

2.3 Technical background

2.3.1 Gel electrophoresis

The first analysis performed on the origami is usually gel electrophoresis. This technique allows the separation of nucleic acid molecules by charge and size: the sample is loaded in a gel that acts as a molecular sieve and the applied electric field moves the nucleic acids, negatively charged, towards the positive electrode. Smaller molecules will pass through the gel net more easily and will form bands at the bottom of the gel. It can be performed in agarose gels, for larger molecular weights, or polyacrylamide for smaller ones or to achieve higher resolutions. Once the sample has run long enough to be separated in clear bands, it can be visualized using fluorescent dyes that intercalates into the helix like ethidium bromide or Sybr® Gold.

The sample can also be recovered from the agarose gel using commercial kits, the band of interest is cut from the whole gel and the nucleic acid is retrieved chemically or mechanically.

2.3.2 Atomic force microscopy (AFM)

Atomic force microscopy is a scanning probe microscopy (SPM) technology that allows to scan samples at sub-nanometric resolution. The main components are: a reflecting piezoelectric cantilever which has a sharp tip at one extremity (the probe), a laser pointed at the reflective surface of the cantilever into a photodiode detector and a scanner that moves the sample. The instrument moves the surface under the tip in a raster pattern, when the tip engage the sample the cantilever is deflected by the interaction with the substrate, the piezoelectric current and the laser movements are collected by the detector and translated into a topographic image [Fig. 2.10].

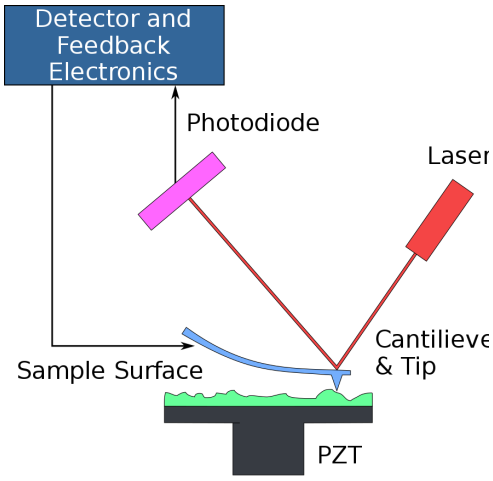


Figure 2.10: Diagram of AFM working principles. The sharp tip on the edge of the cantilever reacts with the sample surface, its movements are detected by piezoeletric components and by laser reflection, all the data are are electronically elaborated to reconstruct the sample topografy.

The nucleic acid sample is immobilized on a mica surface, a mineral formation composed of perfectly flat layers. Mica is negatively charged, to interact with nucleic acids it is necessary to add a positive divalent ion that acts as a bridge [Fig. 2.11] [103, 104], the most used are nickel or magnesium divalent ions.

AFM generates two-dimensional images from the surface so is recommended for 2D origami.

2.3.3 Transmission Electronic microscopy (TEM)

Transmission electron microscopy is another technique capable of visualizing a sample at nanometric scale. In this case the machine transmits a beam of electrons into a detector

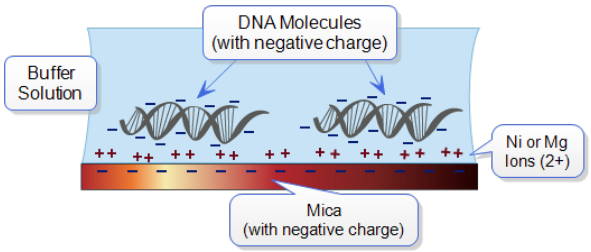


Figure 2.11: Interaction of DNA on mica in presence of positive divalent ions. From reference [104]

after it passed through the sample. The origami sample is suspended on a metal grid and pretreated with uranyl formate, a very electrondense salt that intercalates nucleic acids blocking the electron beam, the sample looks like a shadow over the background (positive staining technique).

With TEM it is possible to vary the depth of focus of the electron beam granting a better insight on 3D origami structures [Fig. 2.12].

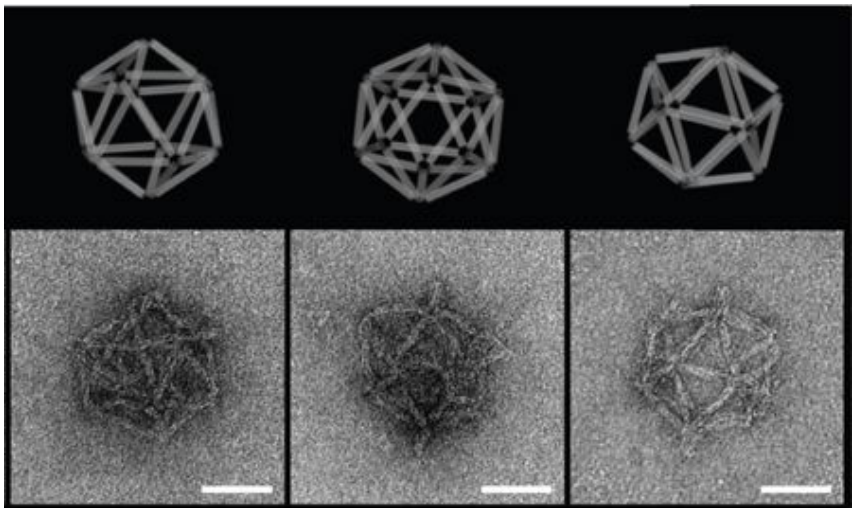


Figure 2.12: Example of TEM imaging on 3D origami. From reference [23]

2.3.4 Flow cytometry

Flow cytometry is the technology used to count and measure the fluorescence of *E. coli* cells during the reporter gene experiments. A flow cytometer can analyse individual small particles, like cells, in a high-throughput fashion. The main components of this machine are: the flow cell where the sheath fluid carries the bacteria and aligns them in a single

line; a laser system that excites the fluorophores and allows the measurement of light scattering; the detectors and photodiodes to collect the signal.

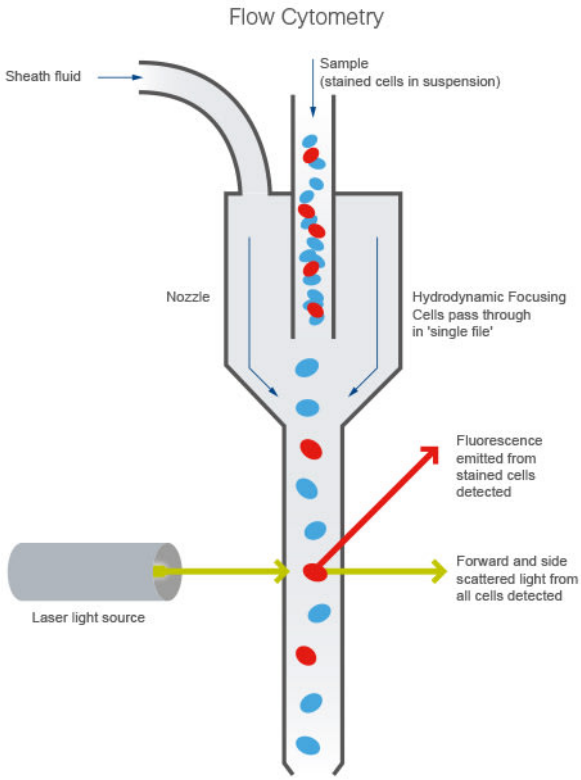


Figure 2.13: Representation of flow cytometry principles.

2.3.5 Microplate reader

The analysis of the functionalized origami fluorescence was performed using a microplate reader. This instrument, similarly to a flow cytometer, measures the intensity of the fluorescence exciting the fluorophore through a laser light and collecting the signal with a detector. In this case the sample is held in a microplate well and the operating volumes can be miniaturized down to 20 μl.

2.3.6 pUC19 vector

pUC19 is a commercial cloning vector often used as DNA scaffold for origami. It is easy to replicate and cheap to buy. The commercial version from New England Biolabs (NEB) was used during this research. pUC19 is designed to have a multicloning site containing different restriction enzyme sequences, a pMB1 origin of replication, a lacZ α gene and

an ampicillin cassette [Fig. 2.14]. It has a GC content of 52%.

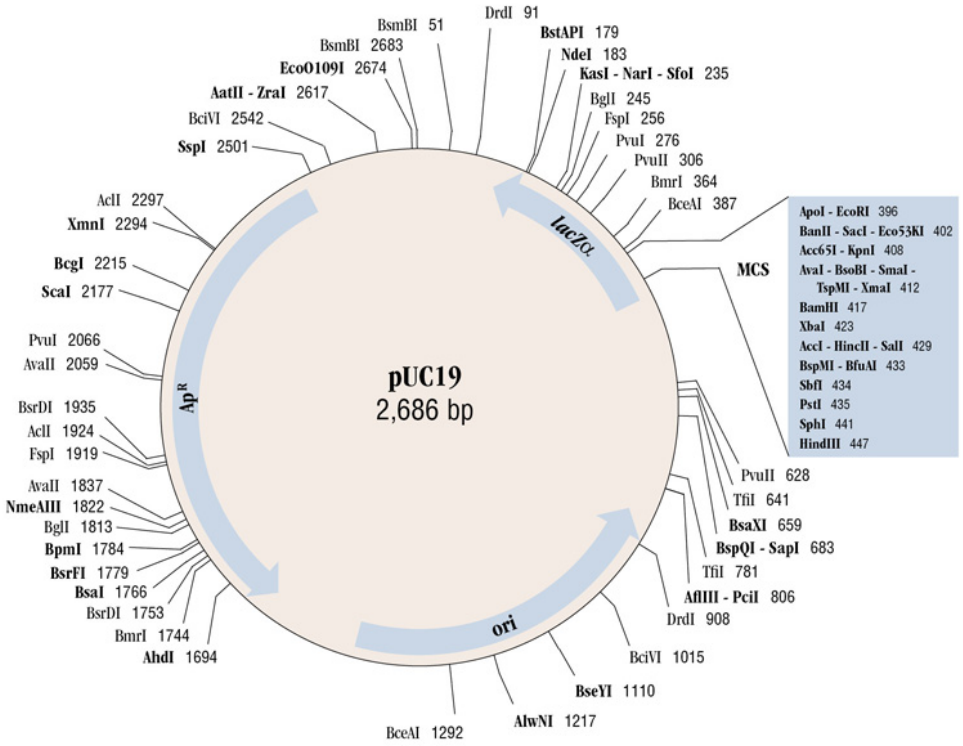


Figure 2.14: Map of pUC19 vector from NEB **??**

2.3.7 Commercial nucleic acid synthesis

Synthetic sequences were purchased from IDT and Eurogentec as follow: Standard PCR primers were purchased from Eurogentec. DB 2.4 biotinylated primers were ordered from IDT. DB Tri was synthesized by Life Technologies while DB 981 was ordered from IDT and both sequences were cloned in proprietary plasmids.

2.3.8 Nucleic acid concentration measurement

The concentration of nucleic acid was measured using a NanoDrop ND1000 spectrophotometer.

2.3.9 Nucleic acid storing

Nucleic acid products, purchased or synthesized in the lab, were divided in aliquots, storing part of them as stocks in the -80°C freezer for longer conservation, and the remaining

parts at -20°C for rapid use.

Chapter 3

Experimental and Computational methods

Dealing with an interdisciplinary project requires to master a set of techniques originating from very different research fields In this chapter I will describe the most important in order to understand the experiments.

3.1 Nucleic acid origami

The origami folding requires a two-steps approach: first a computer aided design and then the actual folding using wet-lab techniques. The next paragraphs will examine each single step in detail.

3.2 Origami computational design and De Bruijn sequence

The design step of an origami is a complex matter, especially for larger origami. Furthermore any thermodynamical optimization increases exponentially the workload required to reach a final design. During the years many research groups developed computer-aided-designing (CAD) tools aimed to make this step less requiring in terms of knowledge and more user-friendly. Two software packages were used for the basic design of the origami in this research: CADnano [23], and CanDo [25]. The first software allows the design of 2D and 3D DNA structures, while the second simulate the final structure showing its

stability.

The origami sequences optimization and the results are described in detail in Dr. Kozyra Ph.D. thesis [5].

The study focused on the properties of mathematical sequences, in particular the so called De Bruijn sequences, replacing the natural scaffolds so to improve the controllability of the folding process. By definition, De Bruijn (DB) sequences are "non-biological" and "highly addressable", hence programmable. "Non-biological" simply means that it has no comparable known sequences in nature, while "highly addressable" is a term borrowed from computer science and it means that for each staple sequence domain there is only one scaffold complementary target. A DB DNA or RNA sequence is composed of a four nucleotide alphabet, in predefined sub-sequences of k nucleotides length, that appear just once in the whole main sequence (no sequences long as k or longer are repeated), giving it a unique addressability while working with k long sub-units (or longer). The sequence can be described as $B(4, k)$ [Fig. 3.1]. By this rule, and using just the four natural nucleotides, a De Bruijn sequence can be long as 4^k (4096 nt).

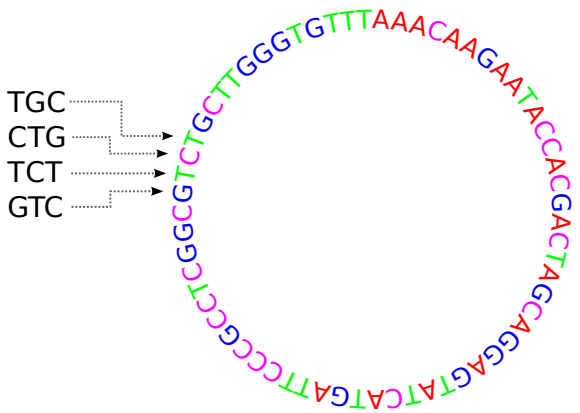


Figure 3.1: De Bruijn DNA sequence based on a 3 nucleotides subunit set, each triplet is included only once in the whole sequence. From reference [2]

The combinatorial system needed to design a long DB sequence requires a computing aid that was developed by Dr. Kozyra and used together with the standard softwares described above.

The DB scaffolds were furthermore improved and made bio-orthogonal by removing all the parts matching *E.coli* K12 sequences recorded in the PRODORIC database [105], and removing the most common restriction enzymes selected from the New England Biolabs

scaffold		NEB	
(length in knt)	PRODORIC	common	all
DBS (1)	0	0	27
DBS (2.4)	0	0	63
pUC19 (2.6)	28	9	66
M13mp18 (7.2)	42	9	89
λ-phage (48.5)	69	12	145

Table 3.1: The table shows that the De Bruijn sequences (DBS) have no common sequences with *E.coli* K12 and the most common NEB restriction enzymes, while the standard scaffold sequences match them extensively. From reference [2]

online database [Tab. 3.1]. The sequences were also checked against BLAST, showing no relevant alignments with any other recorded sequence.

The sequences also meet a series of predefined characteristics as thermodynamic stability, avoidance of secondary structures and transcription efficiency. The sequences were selected after a structural analysis performed with ViennaRNA package [106] and then refined removing the unwanted characteristics.
The design process involved a continuous dialogue between biological and computational groups, feedback from the laboratory results provided new inputs for the *in silico* analysis in a classical design-validation cycle. The final result consists in two DB sequences, one DNA scaffold of 2.4 kb and one RNA scaffold of 1Kb.

3.3 Molecular computational design

3.3.1 Plasmid design

All the plasmids were designed using the Benchling on-line platform (https://benchling.com) especially the Molecular Biology Design and Analysis Suite. This on-line lab note allows the design of biological parts, the annotation and further elaboration with convenient tools for PCR, digestions and sequence alignments.

3.3.2 Aptamer design

The analysis and design of the aptamers was performed on NUPACK (http://www.nupack.org), an online software suite used to predict the secondary structure of nucleic acids at the equilibrium. It can be used for both DNA and RNA and it allows to combine multiple strands with different sequences and concentrations.

3.4 Molecular laboratory methods

Here I will describe the technologies I used to build and analyse the genetic constructs used to study the scaffolds *in vivo* and inside the vesicles.

3.4.1 Plasmid transformations and extractions

The first necessary step, to produce a scaffold, is the transformation of the vectors in the bacteria. Each plasmid was transformed in *E. coli* 5-α competent cells (NEB), a standard cloning strain, the transformations were performed following the heat-shock protocol suggested by the NEB kit manual. The transformed cells were streaked on agar plates and incubated at 37°C overnight using the proper antibiotic for the selection: kanamycin 50 μl/ml for DB transformants and ampicillin 100 μl/ml for pUC19 transformants. The next day a colony was inoculated in 7 ml of LB broth and antibiotic and grew overnight. Two stocks for each strain were prepared mixing 750 μl of culture with 750 μl of glycerol 50% and stored them at -80°C. The remaining culture was used to extract the plasmids using a Qiagen plasmid mini kit or midi kit.

3.4.2 NEBuilder HiFi DNA Assembly cloning Kit

This commercial kit allows to clone the DB sequence in a plasmid to test the *in vivo* behaviour, it includes a 5' exonuclease, a DNA polymerase and a DNA ligase, all working in a one-pot reaction. All the experiments were performed following the manufacturer recommendations.

The vector and the DNA fragments to be cloned must have overlapping ends of about 20 nt, the exonuclease digests the 5' ends so the complementary sequences could anneal in the right order. The polymerase extends the 3' ends to rebuild the double helix and

fFinally the DNA ligase seals the construct, that is ready to be transformed in a bacteria.

3.4.3 Colony PCR

Colony PCR protocol: the tip of a loop was dipped in the target colony and dissolved it in 15 μl of water, the sample was heated to 95°C for 5 minutes in a thermal block and then centrifuged for 10 minutes at 13000 RCF. 2 μl of the supernatant were used as as template.

Colony PCR was always performed using a Phusion PCR kit (NEB) following the provided instructions.

3.4.4 *E. coli* T7 Express growth and induction protocol

2.4 kb insert *E. coli* T7 Express strains were streaked on agar plates with Kan and incubated overnight at 37°C in the shaker (shaking shelves in 37°C room) .
A colony from each plate was inoculated in 1,5 ml of LB + Kan (50 μg/μl) and grew overnight in the shaker.
In the morning each culture was refreshed in six tubes with 2 ml of LB + Kan (50 μg/μl), three of them were induced with IPTG 0,5 mM and three left uninduced. All incubated at 37°C in the shaker.

981 nt insert The strains were streaked in agar plate with Kan and incubated overnight at 37°C. A single colony from each of them was inoculated in 1 ml LB + Kan (50 μg/μl) and grew overnight at 37 °C in the shaker.
Each culture was refreshed in six tubes with 2 ml LB + Kan (50 μg/μl), three samples uninduced and three samples induced with IPTG 50 mM.

3.4.5 Flow cytometry

The flow cytometry experiments were performed with a BD FACSCanto II, on modified bacterial cultures to analyse the fluorescence generated by sfGFP expression systems. The strains were streaked in agar plate with Kan and incubated overnight at 37°C. A single colony from each of the strains was inoculated in 1 ml LB + Kan (50 μg/μl) and grew overnight at 37 °C in the shaker.

Each culture was refreshed in six tubes with 2 ml LB + Kan (50 μg/μl), three samples uninduced and three samples induced with IPTG 50 mM.
The fluorescence was analysed with a FACSCanto II cytometer after 4 and 6 hours of growth at 37°C in the shaker. For each sample, a 100 μl aliquote was taken and diluted in 900 μl of PBS just before the analysis. The analysis was controlled manually and each sample reached at least 35000 events.

3.5 Origami laboratory methods

3.5.1 pUC19 scaffold preparation

The pUC19 scaffold was generated through an enzymatic process. First, the plasmid was nicked using the Nt.BspQI enzyme (NEB) that cleaves only one strand of the DNA double stranded sequence. It recognizes only one sequence on pUC19 (position 690) and the cut unravels the plasmid removing the supercoiled band from the agarose electrophoresis gel [Fig. 4.2].

Once the plasmid was nicked it was digested with a mix of two nucleases: a T7 endonuclease (NEB) that can remove nucleotides from the nicked site, starting the reaction, and a λ exonuclease (NEB), very efficient at removing nucleotides in a 5'-3' direction, but that cannot start from gaps or nicks. The nuclease reactions were performed following the protocols provided by the producer and needed 20 hours to complete the digestion of the anti-scaffold [Fig. 4.3].

3.5.2 DNA Origami folding protocol

Initially, the original origami folding protocol was favoured over the previous assembly technologies for its simplicity. In detail, it was a one pot reaction, containing a long single strand DNA scaffold and complementary staples in excess to 100 fold, isomolarity of the oligonucleotides is not required. Ionic charges were added in the form of magnesium and sodium salts, at variable concentrations depending on the design, to stabilize the negative charges of the DNA backbone, their concentration was usually increased when folding 3D origami to overcome the tight packing of the helices. The solution was then heated up to 95°C to denaturate all the unwanted secondary structures. Finally a thermal ramp was

applied to reduce the temperature at 20°C in 2 hours [1].

pUC19 origami folding The origami folding was performed following the isothermal protocol developed by Sobczak et al. [32].

A one-pot reaction is prepared mixing: TE buffer,Mg(OAC)2 [8 mM], NaCl [5mM], origami staple mix [200nM]and DB 2.4 ssDNA scaffold [6 nM].

And an isothermal annealing is performed as follows:

a 95°C 30 seconds denaturation, followed by a 51°C 60 seconds isothermal annealing and a final 4°C hold

3.5.3 RNA/DNA hybrid origami folding protocol

DB Triangle origami folding The folding solution for the triangle origami is a classic one-pot reaction. The base folding mix remained the same for all the experiments and it is composed of TAE buffer 1X, Mg(OAC)2 [12.5 mM], Staple mix [100 nM] and DB Tri RNA scaffold [10 nM].

R1 origami folding The R1 origami folding investigation was performed preparing a mix of TAE buffer 1X, $Mg(OAC)_2$ [12.5 mM], R1 staple mix [100 nM] and DB Triangle short RNA scaffold [10 nM].

The origami was folded following an isothermal profile covering from 20°C to 69°C for 15 minutes, and holding the product at 4°C at the end of the folding.

The folding products were loaded and run into 1.5% w/v agarose gel electrophoresis in TAE buffer [Fig. 5.7].

DB square origami folding The R1 origami folding mix was composed of TAE buffer 1X, $Mg(OAC)_2$ [12.5 mM], R1 staple mix [100 nM] and DB 981 RNA scaffold [10 nM].

The R1 origami 50X-staples folding mix was in practice the same reaction with a higher staple concentration: TAE buffer 1X, $Mg(OAC)_2$ [12.5 mM], R1 staple mix [500 nM] and DB 981 RNA scaffold [10 nM].

3.5.4 Origami purification with Amicon Utra spin filters

500 μl of TAE buffer 1X were added to the column and centrifuged 5 minutes at 14000 rcf. The eluate was discarded. Folding buffer was added to the origami solution to 500 μl, loaded in the column and centrifuged for 1 minute at 14000 rcf discarding the eluate. The sample was washed three times ading 500 μl of folding buffer to the column and repeating the centrifugation step. The sample was retrieved reversing the column in a new tube and centrifuging it for 5 minutes at 1000 rcf.

3.5.5 Origami imaging

Transmission electron microscopy For TEM different protocols were tested changing reagent concentrations and incubations times. The most efficient experiment was the one showed in [Fig. 5.3,3.2].
To prepare DB triangle for TEM, the copper grid was incubated with 5 μl of $Mg(OAc)_2$ 0.5 M for 1 minute, then it was dried with filter paper from the side, 5 μl of sample were added and incubated for 4 minutes, the grid was dried again with filter paper, a 20μl drop of uranyl formate 2% was put on parafilm and the grid was positioned on it, face towards the liquid, for 40 seconds. The grid was dried one last time with filter paper and let it dry in air for 5 minutes before loading it in the microscope probe.

Atomic force microscopy AFM protocol for DB triangle origami is a modified version of the protocol used with DB 2.4 square. I found that nickel chloride forms a better ionic bridge between the mica and the RNA origami than magnesium acetate or magnesium chloride. It also works at much lower concentrations (10 mM against 125 mM of Mg^{++}).

AFM protocol for DB triangle: the mica was cleaved for every new sample using adhesive tape, 5 μl of the origami sample were added on the mica with 10 μl of $NiCl_2$ 10mM in TE buffer. $NiCl_2$ 10mM in TE buffer was used as scanning buffer for the fluid cell.

3.5.6 Origami AFM imaging

DNA origami imaging [Fig. 4.6] was obtained with the following protocol:

a new mica sheet was cleaned using adhesive tape, 5 μl of $Mg(OAc)_2$ 125 mM was

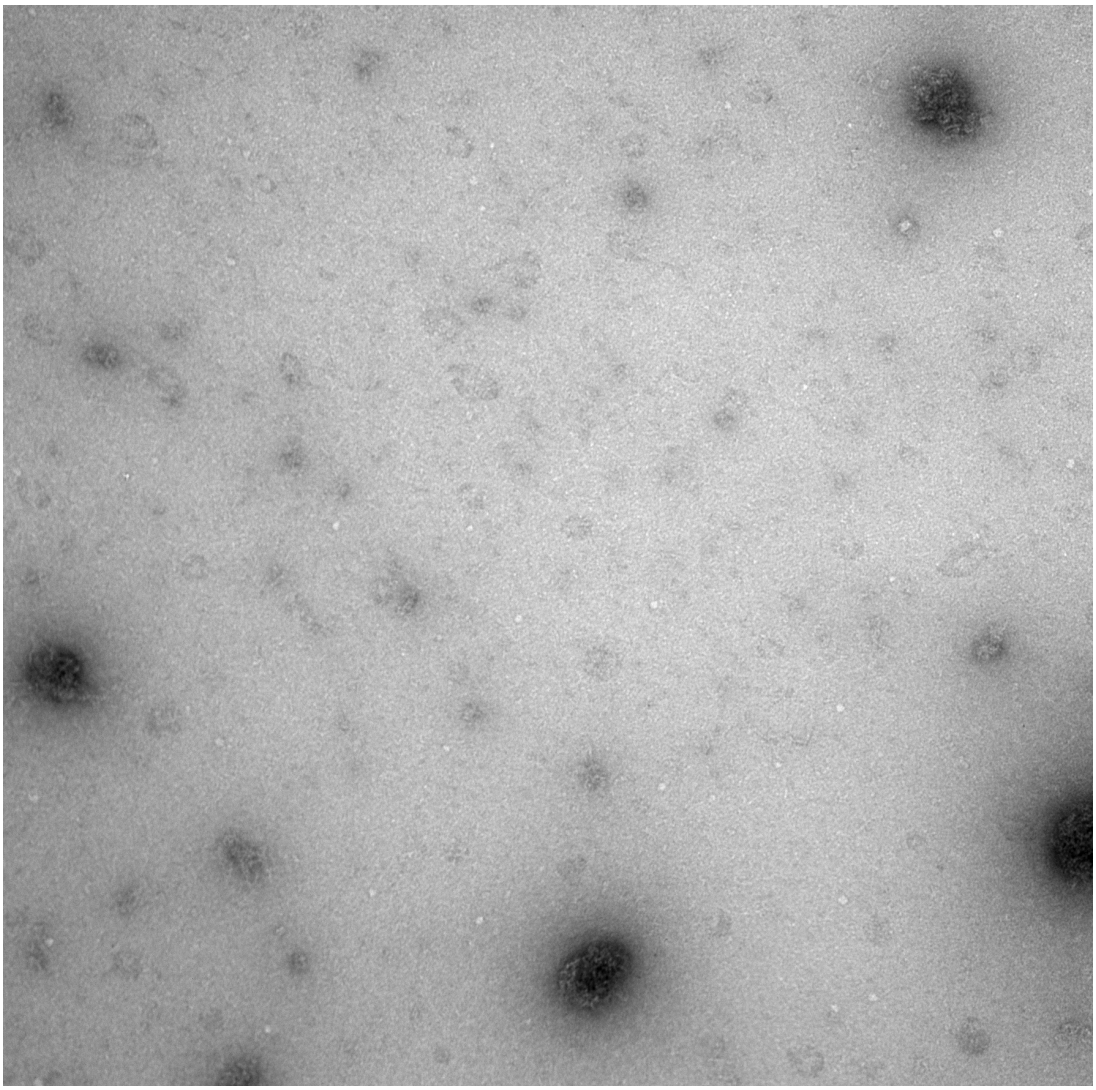


Figure 3.2: Transmission electron microscopy of the DB triangle origami. Full image of [Fig. 5.3].

added to the surface with 5 μl of DNA origami sample. The imaging was performed in magnesium buffer: TAE 1X, $Mg(OAc)_2$ 12.5 mM

DB R1 origami imaging For the imaging of DB R1 origami the mica was cleaved before every sample, 20 μl of $NiCl_2$ 10 mM in H_2O were added and incubated 3 minutes. $NiCl_2$ was removed using filter paper from the side of the mica to avoid contaminations, then 20 μl of the origami sample was applied to the surface. $Mg(OAc)_2$ 12,5 mM in TAE buffer was used as scanning buffer for the fluid cell.

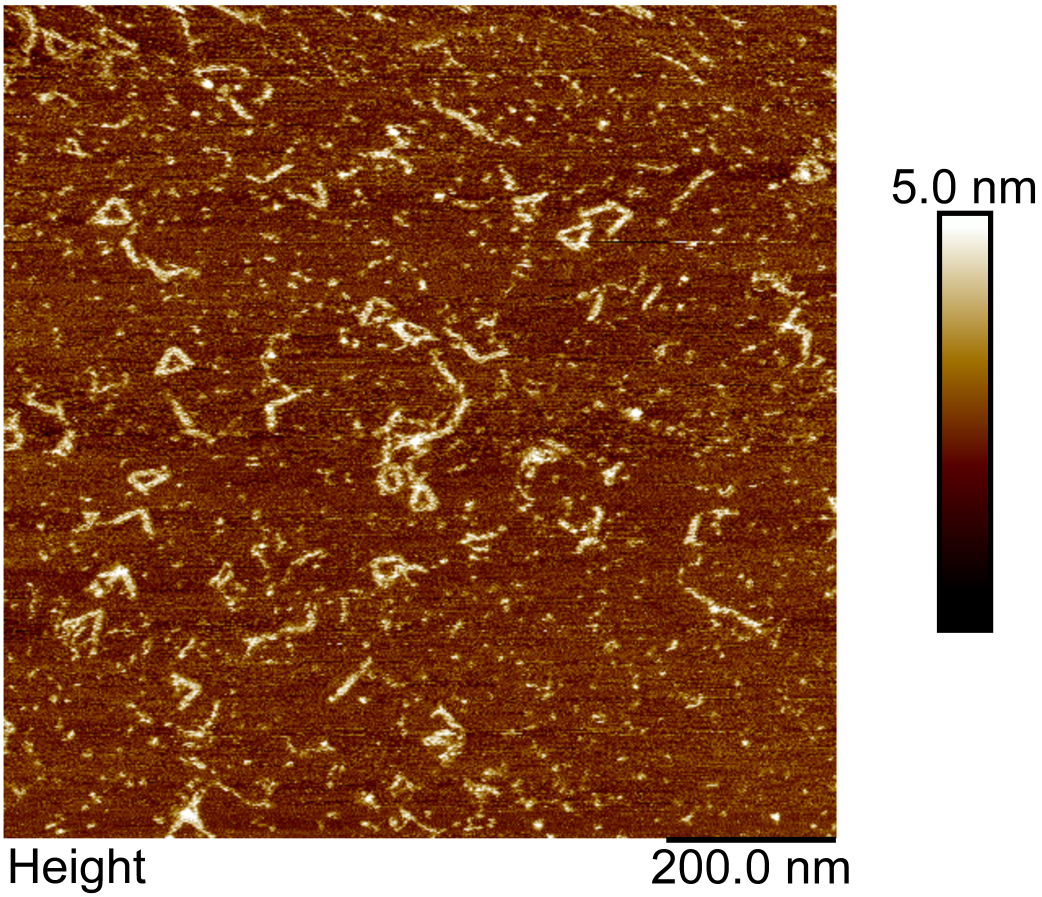


Figure 3.3: AFM image of the DB triangle origami. Full picture of [Fig. 5.4a].

3.6 *In vivo* DB scaffold study methods

3.6.1 PBAD33 reporter construct

pBAD33 was digested using HindIII-HF and KpnI-HF (NEB) restriction enzymes. An agarose gel electrophoresis purification was performed on the digested vector, cutting the gel band of interest and eluting the plasmid with a Monarch gel extraction kit (NEB).

The NEBuilder HiFi reaction mix was composed by HiFi buffer 1x, pBAD33-HindIII-KpnI [0,1 pmol], DB 2.4 HiFi [0,2 pmol] and RBS-$sfGFP$ HiFi [0,2 pmol]. The mix was incubated at 50°C for 30 minutes.
The new vector obtained, pBAD33-DB2.4-sfGFP [Fig. 3.4], was transformed in a *E. coli*-5α competent strain (NEB) following the heat-shock transformation protocol included in the kit: 2 μl of NEBuilder reaction were added to 50 μl of *E. coli*-5α competent cells and mixed gently. The solution was put in ice for 30 minutes and the heat-shock was performed at 42°C for 30 seconds. The cells were put in ice for 2 minutes then 950 μl of

SOC Outgrowth Medium (NEB) was added and the tube was incubated at 37°C for 60 minutes shaking at 180 rpm. After the incubation 100 μl of the cells were spread onto an agarose plate (Cam 25μg/μl) and incubated at 37°C overnight.

Figure 3.4: Map of pBAD33-DB2.4-sfGFP.

The plasmid was finally extracted with a midi-prep kit (Qiagen) and sent for sequencing (MRC PPU DNA Sequencing and Services) using overlapping primers to increase the precision.

pBAD33-DB2.4-sfGFP was also transformed in *E. coli* BL21 (DE3) and *E. coli* MG1655 strains, to check the absence of arabinose-induced fluorescence at the microscope (Nikon Eclipse Ti) (data not shown).

3.6.2 pET-28a reporter construct

pET-28a was digested with XbaI and HindIII restriction enzymes (NEB), an agarose gel electrophoresis was performed to check the digestion [Fig. 3.5], and to purify the products from the gel using a Qiaquick gel extraction kit (Qiagen).

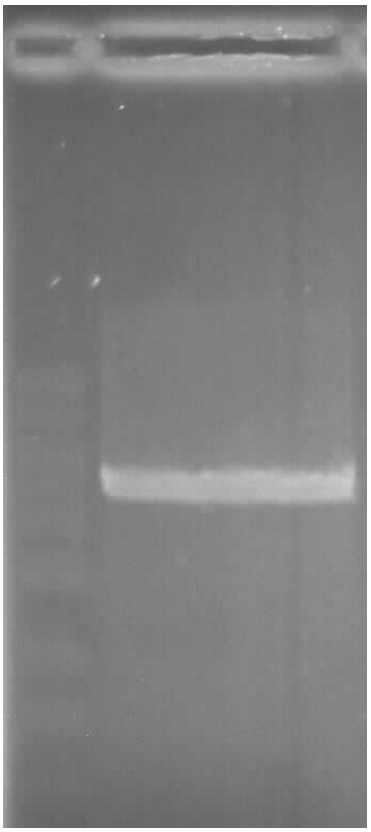


Figure 3.5: Agarose gel electrophoresis of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 - pET-28a digestion with XbaI and HindIII, as backbone for Nebuilder HiFi reaction. Low intensity fluorescence is due to a shorter EtBr staining step, to reduce the sample exposure.

The digested vector was used with the Nebuilder HiFi kit in a one-pot reaction including: HiFi buffer 1x, pET-28a-XbaI-HindIII [0,1 pmol], DB 2.4 HiFi [0,2 pmol] and RBS-$sfGFP$ HiFi [0,2 pmol].

The reaction was incubated at 50°C for 60 minutes.

The final product was transformed into *E. coli* T7 Express competent cells (NEB), a modified strain of *E. coli* BL21 expressing T7 polymerase.

3.7 Aptamers laboratory methods

3.7.1 Polyacrylamide gels

The aptamer folding experiments were analysed through polyacrylamide gel electrophoresis. All the gels used were polyacrylamide 10% in TBE from Invitrogen. The gels were run at 200 V for 45 minutes. The staining of the gels varied depending on the aptamer

analysed.

Broccoli aptamer The staining of the gel was composed of two separate steps, the first for DFHBI-1T and the second for nucleic acids.

After the electrophoresis the gel was washed in RNase free H_2O for 5 minutes, stained for 30 minutes in DFHBI-1T staining solution and visualized using a Typhoon laser scanner at 488 nm emission and 526 nm excitation.

The DFHBI-1T staining solution is composed of Hepes [40 mM], KCl [100 mM], $MgCl_2$ [1 mM] and DFHBI-1T [10 μM].

The gels were washed three times in RNase free H_2O for 5 minutes to remove the staining and then incubated for 10 minutes in SYBR Gold nucleic acid gel stain (Thermo Fisher) before visualizing it again with a Typhoon using the same wavelengths.

Malachite green aptamer The staining protocol is similar to the Broccoli experiments and followed this procedure: a first wash in RNase free H_2O for 5 minutes, stained for 30 minutes in MG staining solution and visualized in a Typhoon laser scanner using the 630 nm laser for the emission and 652 nm filter for the excitation. The Malachite green staining solution was a mix of Hepes [40 mM], KCl [100 mM], $MgCl_2$ [1 mM], Malachite green [10 μM]

The gels were washed three times in RNase free H_2O for 5 minutes to remove the staining and then incubated for 10 minutes in SYBR Gold nucleic acid gel staining (Thermo Fisher) before scanning them again with a Typhoon, using SYBR Gold wavelengths (488 nm/526 nm).

3.7.2 Agarose gels

The aptamers-functionalized origami were analysed by agarose gel (1,5% W/V in 1X TBE) electrophoresis and stained in a buffer composed of TAE 1X, $MgCl_2$ [12,5 mM] and Malachite green [5 μM] for 15 minutes.

3.8 Giant unilamellar vesicles methods

The protocol I used to generate the vesicles is a variant of the inverted micelle synthesis [97] described in 2.2.6. It can be performed using a microplate, for smaller reaction

volumes, or Eppendorf tubes for larger ones.

3.8.1 Labware silanization

The first step required is the silanization of the glassware and plasticware to avoid their interaction with the lipids. Tubes, microplates and microscope slides are treated with Repel-silane ES (GE Healthcare) as follow:

all the containers were filled with Repel-silane ES (or submerged in the case of microscope slides) and incubated for 10 minutes, Repel-silane was removed and washed first with ethanol and then water, repeating this passage three times, finally all the tools ere dried under compressed air.

Silanization of microplates: Microplates wells were filled Repel-silane ES and incubated for 10 minutes. Repel-silanewas removed and the wells were washed with water three times. 100 μl of DNA from salmon sperm was added per well and incubated 10 minutes, then washed with water three times and dried under compressed air.

3.8.2 Lipid mix preparation

The lipids used in the vesicle preparation were bought dried or in chloroform solution (Avanti polar lipids). The final product was a 200 μM lipids solution in mineral oil, formed by one or more lipid species.
Phospholipid solution protocol: the phospholipid powder was dissolved in chloroform to 5 or 10 mg/ml depending on the experiment. The lipids solutions were added accordingly to the final volume and ratio, in a sealable vial. The chloroform was evaporated using compressed air. Mineral oil was added to obtain a concentration of 200 μM. The vial was sealed tightly and heated at 75°C for 10 minutes in an oven or in a thermostatic bath, it was then vortexed for 1 minute. The heating and vortexing steps were repeated three times. The solution was then sonicated at 55°C for at least 30 minutes and incubated overnight at room temperature in a dark place to be used within three days.

It was noticed that sterile conditions improves the quality of the phospholipidic solution. It is also reported that air humidity contents over 50% could degrade the mineral oil quality disrupting this application [107].

3.8.3 Hosting and droplet solutions

The vesicle formation protocol requires the use of two aqueous solutions with different densities, the hosting solution (HS), lighter, and the droplet solution (DS), heavier. At the end of the vesicle preparation the HS is the extra-vesicular medium, while the DS is the internal content of the vesicles. To avoid the bursting of the vesicles due to the osmotic pressure, these two solutions must be equisomolar, within a tolerance range of 20 mOsm. Osmolarity was measured using a freezing point depression osmometer (Lser Type 15 Auto). Glucose (molecular weight 180.156 g/mol) and sucrose (molecular weight 342.2965 g/mol) were added to the hosting and droplet solutions respectively to a final osmolarity of 1000 mOsm, generating a large difference in density while producing a minimal osmotic pressure.
The solutions can have different compositions, the DS could contain all the components necessary to perform an experiment inside the vesicles, such as nucleic acids, fluorophores and buffers.

3.8.4 Vesicles preparation

The preparation of the vesicles is composed of three steps [Fig. 3.6]. While the passages are the same, the volumes changes if performed in microplates (M) or Eppendorf tubes (E).

Preparation of the interface The host solution (HS) was vortexed and added to the bottom of the vessel with a pipette (M 100 μl, E 500 μl) and spinned down in a centrifuge. The phospholipid solution (PS) was votexed and added on top of the HS, pouring it slowly from the side without touching the HS with the pipette (M 50 μl, E 200 μl). The mix was incubated for at least 30 minutes until the two solutions formed a flat aqueous/lipid interface.

Preparation of the inverted emulsion 6 μl of the droplet solution (DS) were added to an Eppendorf tube, the cap was locked securely shear forceswere applied passing the tube on the wells of a tube rack, repeating the movement 10 times or until the emulsion was formed, finally the emulsion was vortexed for few seconds.

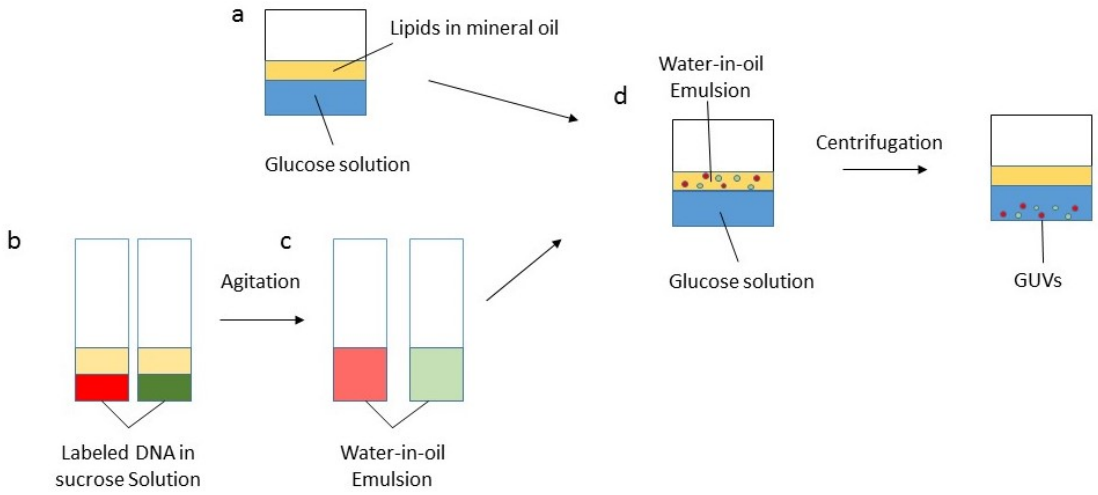


Figure 3.6: GUVs synthesis. A, B - Preparation of the interface. C - Inverted emulsion. D - Preparation of the second interface. E - Formation of the vesicles.

Vesicles assembly The inverted emulsion was placed on top of the aqueous/lipid interface pouring it slowly from the side of the vessel (M 100 μl, E 300 μl), the mix was incubated for 5 minutes or until the inverted emulsion dropped completely onto the interface. The samples were centrifugated 5 minutes at 500 RCF to generate a precipitate at the bottom of the vessel, if the precipitate was not visible the centrifugation was repeated at 1500 RCF to improve the yield. The mineral oil on top was removed with a vacuum aspirator and 150 μl of HS were added.The centrifugation and the aspiration steps were repeated to be sure that all the oil was removed. The vesicles were finally resuspended in the host solution by gentle pipetting.

POPC preparation protocol The droplet solution (DS) was prepared with sucrose [990 mM] and NaI [10 mM], the hosting solution (HS) with glucose [495 mM], sucrose [495 mM] and NaI [10 mM]. POPC was prepared dissolving the powder in chloroform to a final concentration of 5 mg/ml and evaporating the chloroform leaving the tube open in the fumehood. POPC was then dissolved in mineral oil to a final concentration 200 M. 100 μl of HS were added in a microplate well.Slowly 50 μl of POPC solution was added on top of the HS to form a layer, waiting for 30 minutes until the PS formed a flat interface with HS. In a new Eppendorf the droplet emulsion was prepared adding 20 μl of DS in 1000

μl of POPC and passing the tube on a tube rack surface 10 times to mechanically agitate the content. 100 μl of emulsion were added on top of the HS-PS interface incubating the sample 5 minutes to let the emulsion form a second flat interface.The vesicle reaction mix was centrifuged at 1500 RCF for 3 minutes at 4°C. The vesicles formation was checked on an inverted microscope just on visible light to confirm the formation of the vesicles. The top layer of the PS solution was removed using a vacuum pump. 150 μl of HS were slowly added on top of the vesicles and centrifuged again to wash the sample removing again the top layer of the solution.

A drop of 10 μl of vesicles solution was removed from the microplate well and put in an observation chamber, in some of the samples a droplet of Nile Red was added on top of the sample as phospholipidc bilayer staining dye. The observation chamber is formed by a silanized microscope slide and a siliconized glass circle coverslide (Hampton research) spaced by 1 mm handmade tape support and glued with transparent polish nail. The sealed chamber prevents the evaporation of the vesicles droplet, allowing to reverse the sample on the inverted microscope.

The vesicles were visualized using a Nikon Eclipse Ti inverted microscope. Brightfield, GFP and Nile Red fluorescence.

Part II

Chapter 4

DNA origami

In this chapter and in the next the results regarding the synthesis, folding and analysis of the synthetic nucleic acid origami will be discussed. A fully synthetic, non-biological derived nucleic acid sequence was folded into a complex nanostructure. The use of molecules fully based on artificial designs has its advantages: allowing the selection of the desired characteristics it is possible to overcome the natural limitations; at the same time the research moves in "uncharted territory" where no reference points are given because no natural homologous exists.

4.1 DeBruijn sequence design

The De Bruijn sequence used for this experiment was designed by Dr. Kozyra with the single purpose of folding a De Bruijn origami. The design of this sequence is based on the necessity to compare a DB scaffold with a classic pUC19 scaffold 2686 nucleotides (nt) long. In order to have a sequence this long, it was decided to have a minimal non-repeated sequence of 6 nucleotides. Applying this number to the De Bruijn rule 4^k it allows to have a sequence long up to 4^6, or 4096 nt. As explained in chapter 3.2 the results were filtered through different databases to remove the unwanted biological sequences like promoters, terminators and the most common restriction enzymes. The theoretical bio-orthogonality was then confirmed by an alignment analysis on BLAST. Once a set of possible sequences was computed, a further optimization was performed using the ViennaRNA package. This software allows the calculation of the minimum free energy (MFE) of DNA sequences and the definition of the stability of secondary structures. A lower value for MFE means that

the secondary structures are less stable, thus the annealing of the staples should happen more easily. For the final sequence, 2484 nt long and with a GC content of 50%, a MFE of -376.4 kcal/mol was measured, lower than -414.6 kcal/mol of pUC19 [2] [Tab. 4.1]. The sequence was synthesized and cloned in a commercial plasmid by Life Technologies that carries a kanamycin resistance cassette as described in 3.4.2.

TTTCTATGTC TGAGCCTGAA AATGCGATAT ATCGTCAGTC TCTGCGGCTG CCGAAACGCG	60
CTCAACCTCT ACCGTGGAAC CACGGGAAAG GCAACCGAAC CCTTTAAGGC TAATCGCGAG	120
CCGGGTCCCT AAGACAGCGG GATTACCCGG GCCGCGTCAC GCAGTCCTGT CTACTAAGCC	180
TACAGTGTAA AGAGAGCCAA GAGGTCTCGT GTCATGGTCG CACGCCTGGT TGAGTCAGGC	240
TTAGACTCTT GCATCCCCAG CAATAAGTAC ATTGACGTGC CGTTCACGTA CGTTTCCTGG	300
ACGCATGTGT GCGTAAGGTC ATAGAAGCCG ATCTCACCAA GCGCTTACAG AAGAGCTGGC	360
GACACGGATG GCGGTATACC GATACCCCCA TATAAAGTTC GTATAAGGGC AGGAGTTACC	420
TCGCGGTTCG GTGGCCGACG GCTCACTGGA TGTATAGTCC CACTTCCTCA GATGCACATC	480
CTCGAAGACT TCTGTTCGCA TTTTAGAAAA CTAACAGCTC TCCAGCCGCC CAAGTTAAAA	540
CGACCCTGTT TGGTCAATGA AGGTGGGAGT GCTTGCCGCA GGTAGCGAGG TACACTTACG	600
CCGGACCAAA TCTTTGGCCC GTGTATGGAT CATCCATAGC GCGAAGTGAC ACACTGCCCC	660
ACCTCATCTG ACTACGGTAA GTGCGGATTC GGCATGGGGA ACAAAGCTCA TTGGATAGCT	720
GAATAGCCAT ACTGAGGATA AACACTAGGA ATCGGGGGAT ATCCGTGAAG TTGACCATTA	780
CGGGCGCTAC CATGACCGAG GGATGACGAG ATTTAGGCAC GTTGTCCTAC TTAACCCCTT	840
GCGGTCGGAC TTTCGCGTGC TCTAATGACT CGATTTGGGA TCGTGGCGTT GGTGTAGAGC	900
GTATTGGCAC TGTTGCAATG TGAAATCGAA CATGGAGACG TTAGATGAGT GTGATCCACG	960
TGAGCTTTGC AGACAAAACA ATGGTGATAC TTCGTTGCTC AGGTGAGGCA TAAGATGGTA	1020
CTTGCTTATC GCAGCCTTAA AGCAGGGTCA GAGTCGGCTT CAGACCGGAA AAATTCAAAA	1080
GCGACTGTCG GTTATTCGCT CGCAATTATC TCGCTTTCAC CTGTACCCAA CAACGTATCT	1140
TCCCCGATTC ACTTTAGCCG TGCGACGCTT GTCGATAACG CTATCCTGCA CTTGAGAAAT	1200
TAAACCAGCG AATCTATACT ACTCGTAGCA GATTGCTGCG TTCGATCCCG GTGACCTAAC	1260
GGAGCTACAT CTAAGGAAGC GTCCTTTTGG ACTGACGGAA TTAGCTATGA CAATAGTAAC	1320
CGGCTATTAC ACGATAGTGG TTAAGAGTGA ACACGCGACC GCGCCCGAGT GGAGTACCAG	1380
GCGCGGCGAT TAAGTCTATT TATGGTTTCG ACTATGCTCG GCCCTTAGGA CTAGCATCTC	1440
TCTTATTTTG CTAAATACAA GGGAGATCAG TGAGTTGCCT CTTCATAAAT CACGAAGGGG	1500
CATTGCCCGG CACACAGCAT TAGGTCCAGG ACGACAAGAA TCAGAATTGC GTCTAAAAGT	1560
AAGCACGGCG GGTGTCGCTA ACCTGACATC GTTTTCTGCT GAGTAGAATA CTCAGTATAT	1620
ACATAATGGC AAATGAGCAT ATGGGAAGGA TGCGGGGTAG TCACTAAACT TCACACCTAC	1680
GCAAAGATCG ACATGTTCAG TTATGCGTGT GTCCCGTCGC GCGCATCGAG TTTGCCAGGG	1740
AATAATCTGT CAGCGTTTGT GTACGCGTTA ACTATAGGTT CAATTTCCGT CTGTAAGAAA	1800
CAGATAAGCG GTGCAAGACC TGGCTTGGCT ACGAGTAATC ATGAAAGTCG TAATGTCAAA	1860
TAGAGTTCCT AGGGACTCAT GCCTAGCCTC CCTGCGAGAC TAATACGATT GTGACGCGGG	1920
CTCGTCGGGT TAGCGGCCAA CTTGGAAGTA GTTGTGGCAT CAGGGCCACA AATTGAGCGA	1980
TCGGTAGGAG CAAGGAGAAC TTTGTCTCAG CTAAGTTTCA GGATTTTCCC TTCCGAGAGA	2040
CACCCTCGGT CACCGACTTA TACGCTGTCC GGTTTGAATG TACTCTGAAC GTCTCCTTCG	2100
CCAAAATCCG AAGCAAAAAC CGCAAGTGTC TTCGGATACA CATACGTGTT TCTTGTTTTG	2160
TTACTATTCT CAAAGTGGCT GACCCACACG TCATCGGCGT CGTGCATTCC AAGGTTACGA	2220
ACTAGAACAG TCGCCTATGG CTCTGGAATG CAACAGGAAA CTCACGGTTG GGGCGGCAGA	2280
GGCCCATGTC CAAAGGGTGA ATTTTTAATC CCTCACATTC TTCTTTCTCT AGGTAATAGG	2340
CTGGGTCGAA AAGGTATGCA GTAGGTGTGG ATTGGTTCTG GCAGTTTTAT AGACATTTGC	2400
GAACGCCCCC TGGGCTGTGA GACCCGCGAT GGGCAATCGT ACCTATAACA AGCCAGAAAG	2460
AAGGCGGACA TAGTTAGGGC GAAA	2484

Table 4.1: Sequence of the De Bruijn 2.4 Kb scaffold

4.2 Scaffold preparation, origami folding and comparison

4.2.1 Origami design

To compare the two sequences a similar square origami was designed by Dr. Kozyra using the CADnano software. The origami was folded by 70 staples that formed a squared shape with a side of 50 nm [Fig. 4.1]. Depending on the ionic concentration the dimensions can vary slightly, a high amount of positive charges relaxes the origami mesh, making it larger. To adapt the longer pUC19 sequence to the same design, the 200 nucleotides in excess were not folded and formed a dangling loop at the corner of the square.

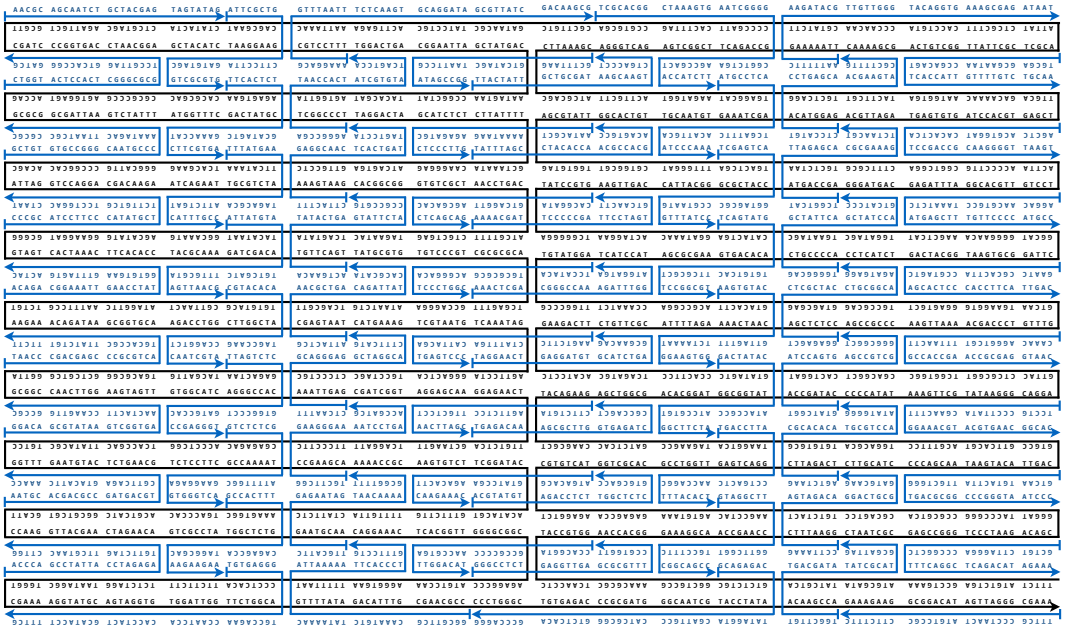


Figure 4.1: Design and sequence of DB 2.4 square origami.

4.2.2 Scaffold preparation

Single strand DNA production A DNA origami scaffold is a single strand DNA (ssDNA) molecule. The sequences to compare were encoded in a double stranded plasmid. Different strategies to remove the complementary strand were compared before characterizing the most efficient protocol for each sequence. Being two different products, a full plasmid in the case of pUC19 and an internal sequence for DB 2.4, two different techniques

were applied.

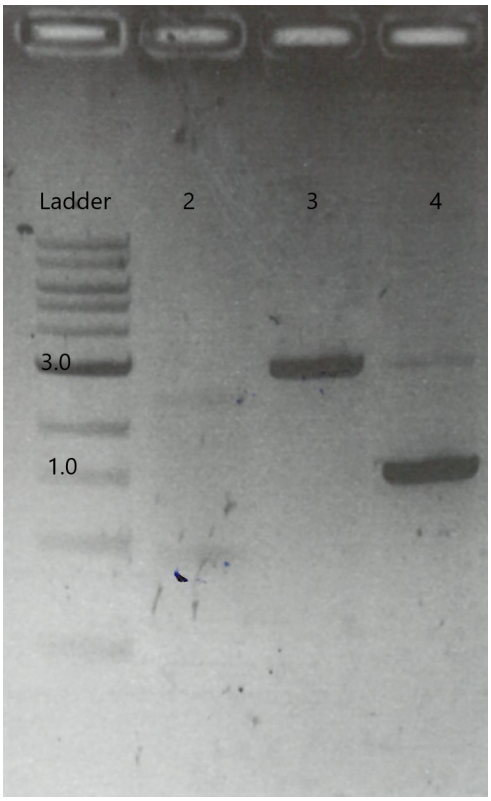


Figure 4.2: Agarose gel electrophoresis of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 - DB 2.4 digestion with λ exonuclease; lane 3 - pUC19 after nicking with Nt.BspQI; lane 4 - pUC19 undigested. There are two clear bands on DB 2.4, the upper one is the undigested dsDNA at 2400 nt, the lower one is the ssDNA scaffold around 900nt. The nicking efficiency of Nt.BspQI enzyme is clearly showed in lane 3 where no supercoiled pUC19 is detectable.

The preparation of DB 2.4 scaffold could not follow the standard pUC19 protocols described in 3.5.1, and different strategies were tested to identify an efficient method. In all the cases the first step required is a PCR (NEB Phusion) to amplify the DB sequence and have it as a linear fragment. The first attempt followed Zhang et al. paper [4] and involved the use of a phosphorylated reverse primer (IDT) at the 5' end of the anti-scaffold sequence. After the purification (Qiagen PCR purification kit), the PCR product is digested with λ exonuclease. This exonuclease can digest only the phosphorylated strand leaving the scaffold strand intact. Unfortunately, the reaction could not proceed to completion for reasons unknown, even when the incubation time was increased over the recommended time, probably because not all the primers were properly phosphorylated [Fig. 4.2 and 4.3].

The second procedure I tested starts also from a PCR method published by Pound et

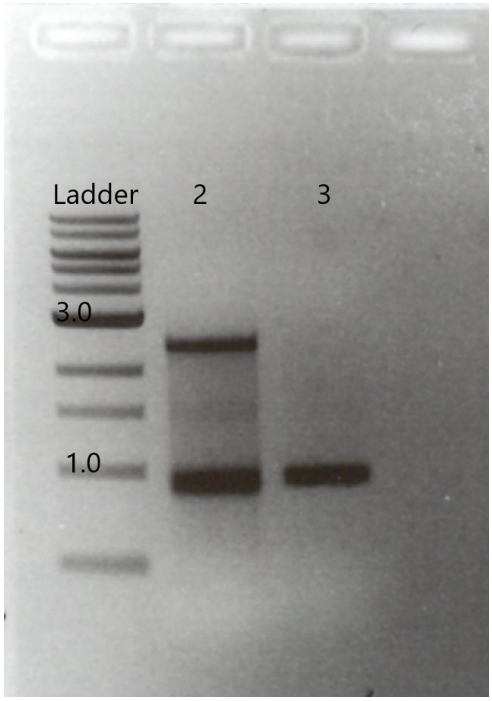


Figure 4.3: Agarose gel electrophoresis of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 - DB 2.4 digestion with λ exonuclease (higher starting amount, 4 hours long digestion); lane 3 - pUC19 after digestion from T7 endonuclease and λ exonuclease. It is clear, as in the previous gel [Fig. 4.2] that DB 2.4 is not completely digested. On the other lane pUC19 is detectable as a single band at the 1 kb mark.

al. [3]. In this case the reverse primer is functionalized with a biotin (IDT functionalized oligonucleotides). The biotinylated PCR product is then incubated with streptavidin-coated magnetic beads. Once the DNA is attached to the beads a solution of NaOH is added to denature the DNA. While the anti-scaffold remains attached to the beads, the scaffold floats freely in the solution. The beads can be collected at the bottom of the tube using a magnet and the scaffold solution is pipetted out to neutralize the pH in Na(OAc). The first attempt was performed using NEB hydrophilic streptavidin magnetic beads and a standard biotinylated primer. The yield of ssDNA was almost null. To improve the efficiency I used a primer with a TEG-biotin modification where the biotin is connected to the DNA 5' end through a longer triethylene glycol linker that grants more space for the biotin-streptavidin interaction [Fig. 4.4].

The performance of the system slightly increased, but not enough to produce an usable amount of scaffold. A breakthrough was finally achieved using a different kit of magnetic beads, the Dynabeads kilobaseBINDER kit (ThermoFisher scientific), coupled with TEG-biotinylated primers.

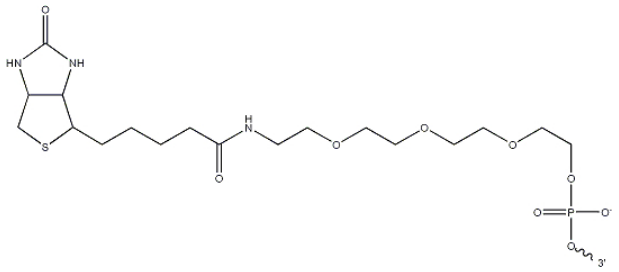


Figure 4.4: Structure of the biotin-TEG primer modification.

The DNA solution obtained was then purified by agarose gel electrophoresis to separate the scaffold from the dsDNA residues and the band was extracted with a Freeze'n'squeeze gel extraction kit (Bio-Rad) [Fig. 4.5]. This tailored protocol allowed the production of a sufficient amount of scaffold necessary to perform the origami folding.

Step 1: NEB Phusion PCR kit
Template DB 2.4 plasmid (Life Technologies)
Forward primer 5'-TTTCTATGTCTGAGCCTG-3'
Reverse primer Biotin-TEG/5'-TTTCGCCCTAACTATGTC-3'
Step 2: Dybeads kilbaseBINDER kit
starting from 5µg of PCR product use 100µg of magnetic beads (twice as recommended)
add TEG-biotinylated DB sequence to the beads and incubate 3 hours on a roller
denaturate the DNA adding 150µl of NaOH 0,2M
place the tube on a magnetic rack to separate the magnetic beads
collect the surnatant containing the ssDNA DB scaffold strand
neutralize the pH of the scaffold with 100µl of Na(OAc) 3M
Step 3: Agarose gel electrophoresis (1% agarose in TAE buffer)
extract the band using Freeze'n'squeeze gel extraction kit

Table 4.2: Workflow for De Bruijn 2.4 Kb scaffold synthesis. This protocol allows the production of a single DNA strand with an arbitrary sequence that can be used to fold a DNA origami.

4.2.3 Origami folding

The folding of DB 2.4 followed the same protocol as the control origami pUC19 2.4 described in 3.5.2.

To remove the staple excess the sample was brought to 500 μl with TE buffer 1X and filtered through an Amicon Ultra centrifugal filter unit with a cutoff of 100 KDa at 14,000 RCF for 1 minute, repeating the washing step three times.
An aliquot was run on agarose gel electrophoresis (2% agar in TAE, ice bath cooled) for an early confirmation of the origami formation [Fig. 4.5].

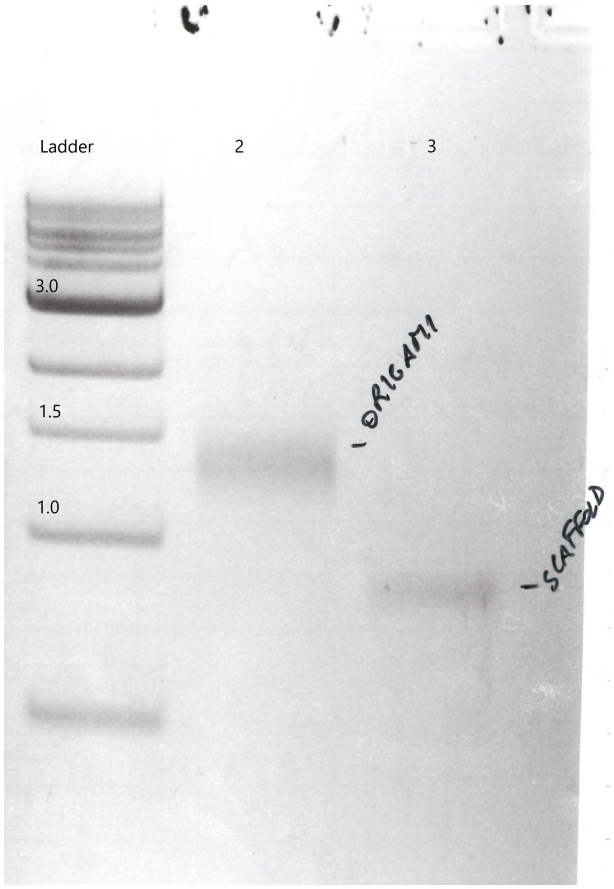
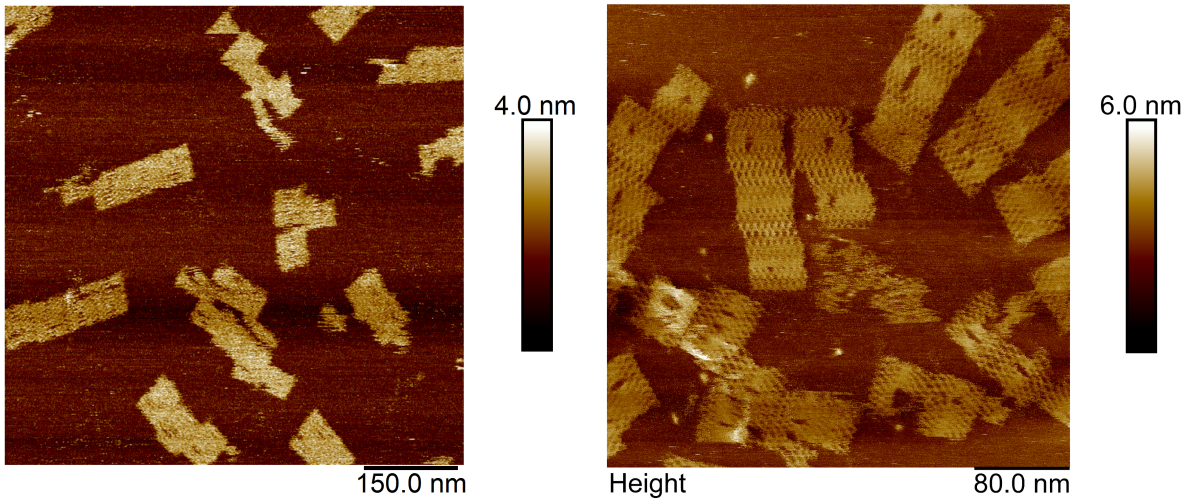


Figure 4.5: Agarose gel electrophoresis of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 - DB 2.4 DNA origami folded; lane 3 - DB 2.4 ssDNA scaffold.

4.2.4 AFM imaging

A definitive confirmation of the origami folding was obtained with the help of atomic force microscopy (AFM). I used a MultiMode 8 (Bruker) microscope in ScanAsyst mode (automatic settings enabled) in liquid.

As described in chapter 3.5.6, the sample was immobilized on a mica sheet using a solution of positive ions.



(a) (b)

Figure 4.6: AFM of DB 2.4 and pUC19 DNA origami in detail, it is possible to discern the formation of different origami structures. 4.6a - DB 2.4 DNA origami 4.6b - pUC19 DNA origami.

4.3 Summary

The results obtained from the comparison of the two origami shows clearly that the folding efficiency of the DB 2.4 scaffold is comparable with the commonly used pUC19, the origami are visible under the AFM microscope and the number of structures is comparable. It is important to remark again about the synthetic nature of the sequence, the total absence of encoded genes makes this technology a possible candidate for a future *in vivo* synthesis and, in theory, for medical applications. Albeit the excellent folding results and the theoretical bio-orthogonality given by the synthetic scaffold, this technology showed a rather important flaw, the workflow to obtain the ssDNA is rather time-consuming and expensive and it does not grant a large amount of final product per single reaction. A possible solution could be the synthesis of ssDNA through rolling circle amplification [45].

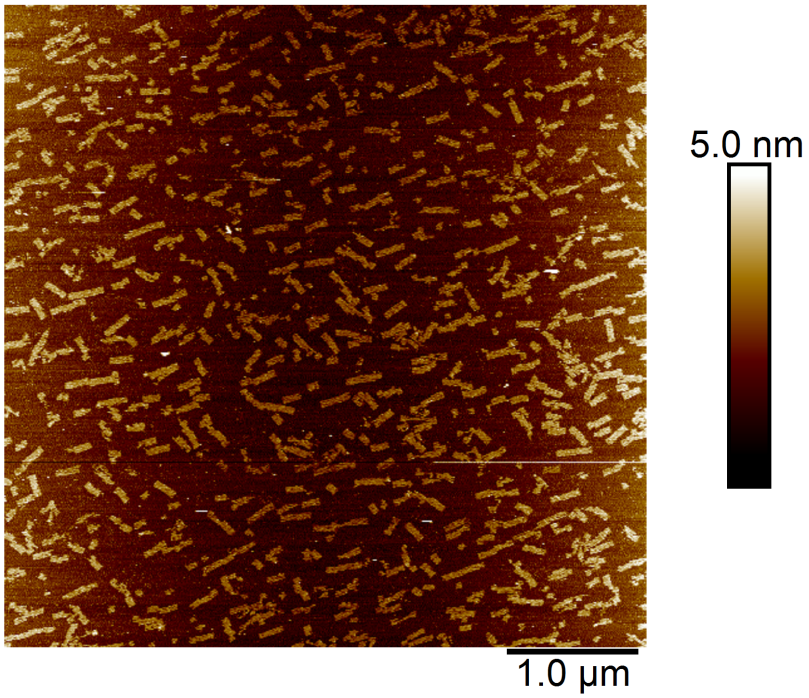


Figure 4.7: DB 2.4 DNA origami AFM panoramic scan (5 μm) of the final product. The high number of well formed structures demonstrates the efficient yield of the tailored protocols.

Chapter 5

RNA/DNA hybrid origami

In this chapter I will discuss the experiments involving the use of an RNA DB scaffold folded using DNA staples.
The use of RNA could be a viable option to prepare an origami *in vivo*. An RNA scaffold can be easily synthesized through the most common biotechnological tools both inside and outside the cell. RNA is a native single strand molecule, thus it is folding-ready. Moreover, the efficiency of the *in vitro* synthesis is much higher than the ssDNA preparation showed in chapter 4, making it a very good candidate for technological applications. The negative side of using RNA is the susceptibility to RNases, the molecule is prone to degradation as soon as the environment is not RNase free requiring extra care when working outside of a PCR cabinet.

5.1 DeBruijn RNA sequence design

The De Bruijn sequence designed for the RNA/DNA hybrid origami is 1026 nt long. A shorter scaffold was synthesized to make the folding less expensive, but long enough to fold an origami easily recognizable under the microscope. The sequence is optimized to the same degree of DB 2.4: the minimal non-repeated subsequence is 6 nt long and the most common biological sequences are filtered out to be bio-orthogonal [2].
The first version of the sequence was inserted in a larger construct designed for the *in vivo* transcription. It presents a T7 promoter followed by a LacO sequence at its 5' to have a strict control over the transcription. At the 3' there is a T7 terminator. The first origami designed uses only 981 nt of the whole sequence, leaving two shorts DB linkers

at the edges. Because of the origami triangular shape, the sequence was named DB Tri [Tab. 5.1].

TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCGGCGCACGGTTCTGTGATCGTGGCGGTCCAGCTAGC
AGGTTTGCGGCTCAGAAGAGCTGTTGTGTTTGTTTTCGACTACCAGAACGGAGTCTCTAGCGTGAGATAAGTAAGATTAG
GCTCGGAGAGTGTGAGGCTTCGTAATCGTACCACACACCAGGCGTAACCGCACTTAGACGCACAGGGTACAAGTGATAGG
TAAAGTTACGGCAGGACGCCCAAAAGTCTGGAGCACAAACGGGGCCCCGCTAGGGAAAACGCCGGGGTAACTATTGTTAT
AATTCAAGAATTAGAACTAAAAGGTAGTAGCACCACTCGGTGGGTTAAACTAGCTAAAGACACCGCTCCAACAGCCGAAA
GTGTACGCTGAATCACAGTCAAATTATACGGTGTTCGAGATCGCGAGTTTTGTGGGATTTGCACTCCAGATACCGATTCG
GTAGCTTTATCGTTCACTGTGTCACGCGCAGCGCCACCAAAGCTGAGACGTTCTCGAAATTCTAATTTCTACGATTAAGT
CCAAACAGAAAGCAATCTATTACACTGGAAGTCAGTAAAACAAAGGGATACAGATCCCGTGACGGCTAGTGCTGTGGTGT
CCGAAGTTGACTGTCAGAGAACAATCGCACCGGACAGTTCGTTGAGTTCCAGTTGCAATTGCGATAGTAATAATAGATAG
AGGCCGTGGAACCCCGTACTTCAGCGAGAAGTGGTCTTGGACTTGTACTGGGGCGAGCGGTGCGGGAACTCGTGTTGCCC
GCAAGCACTGCAACACAGCGGAAGGATAGCAACGATCACTCTTGCTTTGTCGGACTCAGTCTAGGAGCCGCCGAGCCAGT
CCCGCGCGTTCCCACGTTTCCGTAAACGTCCGCTTGGCCCGTCCACTGATATAGTTGGATCGGGAGAAATCGAAGCTCAC
GAACAGGAACGTAAGGCTGCTTGTTCTTTCACGGATCTCGGGCAGAATCTCAAACTCAATTACTCGATTTAGGTCGTCGC
AGTACAGCTTCCACGGGCTTGAAATAGCTCCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGTT
AATTAA

Legend:

T7 promoter
LacO
DB linker 1
DB981
DB linker 2
T7 terminator
PacI restriction site

Table 5.1: Sequence of the DNA template for DB Tri scaffold with legend.

Later, the template was modified to remove the unused parts, especially LacO and T7 terminator, to improve the efficiency for the *in vitro* transcription, obtaining a new sequence renamed DB 981 [Tab. 5.2].

TAATACGACTCACTATAGCGGTCCAGCTAGCAGGTTTGCGGCTCAGAAGAGCTGTTGTGTTTGTTTTCGACTACCAGAAC
GGAGTCTCTAGCGTGAGATAAGTAAGATTAGGCTCGGAGAGTGTGAGGCTTCGTAATCGTACCACACACCAGGCGTAACC
GCACTTAGACGCACAGGGTACAAGTGATAGGTAAAGTTACGGCAGGACGCCCAAAAGTCTGGAGCACAAACGGGGCCCCG
CTAGGGAAAACGCCGGGGTAACTATTGTTATAATTCAAGAATTAGAACTAAAAGGTAGTAGCACCACTCGGTGGGTTAAA
CTAGCTAAAGACACCGCTCCAACAGCCGAAAGTGTACGCTGAATCACAGTCAAATTATACGGTGTTCGAGATCGCGAGTT
TTGTGGGATTTGCACTCCAGATACCGATTCGGTAGCTTTATCGTTCACTGTGTCACGCGCAGCGCCACCAAAGCTGAGAC
GTTCTCGAAATTCTAATTTCTACGATTAAGTCCAAACAGAAAGCAATCTATTACACTGGAAGTCAGTAAAACAAAGGGAT
ACAGATCCCGTGACGGCTAGTGCTGTGGTGTCCGAAGTTGACTGTCAGAGAACAATCGCACCGGACAGTTCGTTGAGTTC
CAGTTGCAATTGCGATAGTAATAATAGATAGAGGCCGTGGAACCCCGTACTTCAGCGAGAAGTGGTCTTGGACTTGTACT
GGGGCGAGCGGTGCGGGAACTCGTGTTGCCCGCAAGCACTGCAACACAGCGGAAGGATAGCAACGATCACTCTTGCTTTG
TCGGACTCAGTCTAGGAGCCGCCGAGCCAGTCCCGCGCGTTCCCACGTTTCCGTAAACGTCCGCTTGGCCCGTCCACTGA
TATAGTTGGATCGGGAGAAATCGAAGCTCACGAACAGGAACGTAAGGCTGCTTGTTCTTTCACGGATCTCGGGCAGAATC
TCAAACTCAATTACTCGATTTAGGTCGTCGCAGTACAG

Legend:

T7 promoter
DB981

Table 5.2: Sequence of the DNA template for DB 981 scaffold with legend.

5.2 Origami triangle folding and analysis

Many combinations of origami design and laboratory techniques have been tested to obtain the best results. I will separate this section in paragraphs, each of them will deal with a different evolution of the scaffold and the origami.

5.2.1 DB triangle origami design

The first folding design is a hollow triangular shape formed by a sequence of three rectangles forming the sides, each 30 nm long. The first and the last rectangles are connected by two staples to close the origami in the right position [Fig. 5.1]. This origami is folded by 21 DNA staples and their sequences are optimized to reduce to a minimum the formation of secondary structures like hairpins, further facilitating the folding [2].

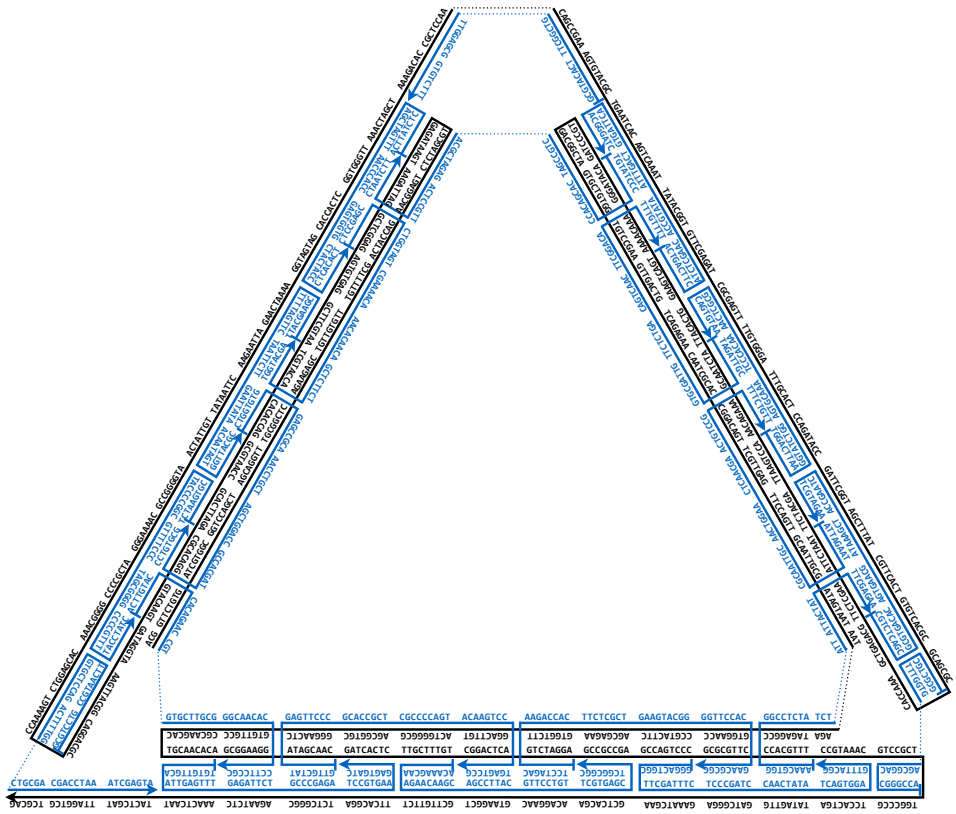


Figure 5.1: Design and sequence of the DB triangle origami. On the bottom left corner it is possible to locate the two staples connecting the open vertex.

5.2.2 DB Tri scaffold synthesis

The synthesis of the RNA scaffold started with the PCR amplification of its DNA template (Phusion High-fidelity DNA polymerase NEB). The PCR product was purified (Qiagen PCR purification kit) and used as the template for the RNA transcription (Ampliscribe T7 Flash transcription kit).

The RNA product was purified again: during the triangle origami experiments I used a phenol-chloroform-isoamyl alcohol mixture (Sigma-Aldrich) and ethanol precipitation [Fig. 5.2], with the newer origami (square and rectangle) I adopted the faster Monarch PCR and DNA cleanup kit (NEB), after assessing its good performances with RNA molecules.

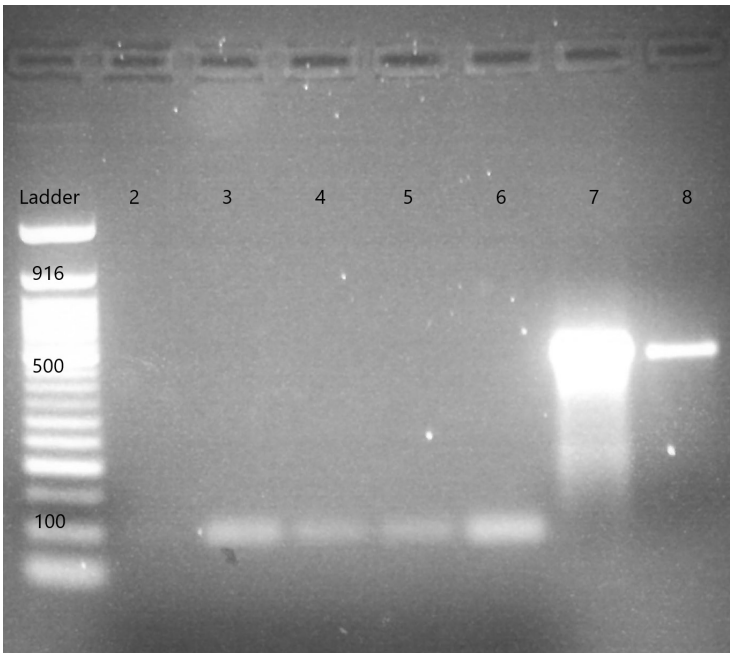


Figure 5.2: Agarose gel electrophoresis of: lane 1 - 50 bp DNA ladder (NEB); lane 7 - DB Tri RNA scaffold purified with phenol-chloroform-isoamyl alchool; lane 8 - control sequence.

5.2.3 DB Triangle origami folding

The investigation that led to the final folding protocol for the DB triangle has been extremely long and involved dozens of different attempts. It can be divided into three main parts:

1. Protocols based on the scaffold optimization theory: relying on the sequence opti-

mization I tried to fold the origami with a short thermal ramp from 37°C to 30°C.

2. Protocols for co-transcriptional folding: in the attempt to simulate an *in vivo* folding I prepared a one-pot reaction containing both T7 transcription and folding step components.

3. Protocols based on the classic origami folding: this protocol includes a slow descending thermal ramp from higher temperatures.

Protocols based on the scaffold optimization theory This protocol is the first folding attempt using *in vivo* compatible conditions. The origami is heated up to 37°C without the first step of denaturation, then the temperature is slowly reduced to 30°C in the attempt to stabilize the structure. The lack of experience in origami folding and microscopy techniques hindered the capacity to identify any definitive results. The only image of TEM showing the triangle origami could not be replicated but, with the experience matured in these years, I can say that the origami formed showing the same design limits of the subsequent attempts [Fig. 5.3]. Probably the difficulties encountered at the microscope stemmed from the lack of a purification step after the folding, with the free staples interfering with the sample deposition on the TEM grid.

Thermal profile: the sample was initially incubated at 37°C for 5 minutes, a thermal ramp of -0.01 °C/s was applied until the temperature of 30°C was reached and held for 5 minutes, the final hold was at 4°C.

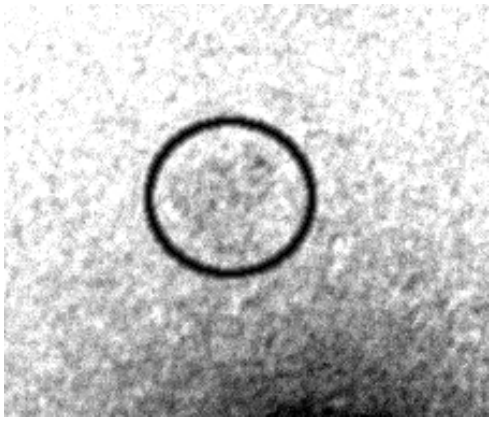


Figure 5.3: Transmission electron microscopy of the DB triangle origami. In this enhanced detail it is possible to see the nanostructure inside the black circle.

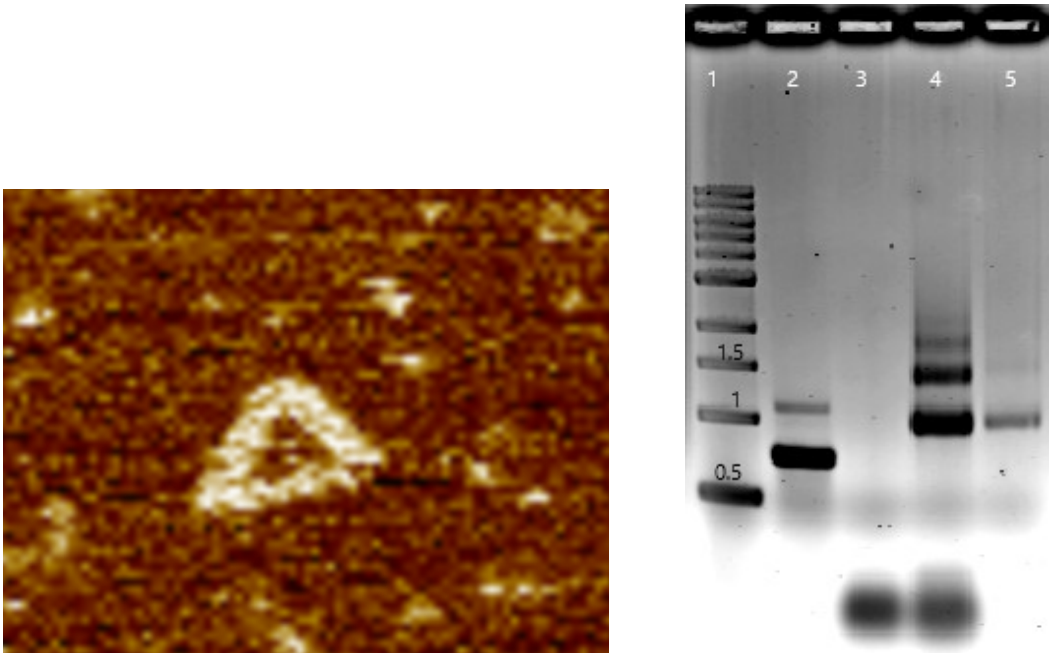
Protocols for co-transcriptional folding In parallel to the previous protocol I tried to prepare a more complicated process that unfortunately did not work. In this case, I prepared a one-pot reaction including both the T7 transcription components and the staple mix. The addition of magnesium, for the nucleic acid folding, probably compromised the activity of the T7 transcription. Even if this procedure did not work as expected, it could be interesting to investigate its applicability for an optimized origami designed for lower folding temperatures and ions concentration.

The transcription and folding mix was prepared as follows: T7 flash buffer 1X, nucleotides [8 mM], DTT [10 mM], Riboguard RNase inhibitor 50 U, T7 polymerase (kit), $Mg(OAc)_2$ [12.5 mM], Staple mix 10 nM and DB Tri DNA template 250 ng. A thermal annealing was applied starting at 42°C for 30 minutes, appliying athermal ramp of -0.01 °C/s to 30°C and holding the sample at 4°C.

Protocols based on the classic origami folding The last and most successful protocol can be described as a classic origami folding without denaturation step, in the attempt to reduce the thermal ramp time, and staying in a temperature range compatible with thermophile bacteria. The starting temperature is set to 65°C that is more than enough to remove any secondary structure or staple-staple interaction. The temperature is gradually lowered by 0.01 °C/s to allow the staples to anneal to the right scaffold domain, until it reaches 25 °C. After the thermal folding, the origami is purified using a filter column (Amicon centrifugal filter 100 KDa). The purified product is finally visualized through microscopy [Fig. 5.4a and 3.3].

5.2.4 DB triangle origami results

The results obtained with the origami triangle showed a problem in the folding process. I suspected that the incorrect folding of the two staples at the triangle vertex could cause a high number of misfolded origami. This idea originates from the fact that it is possible to identify a large amount of folded sides (rectangles) that are not closing into a triangle shape. It is also possible to see linear structures longer than a single open triangle, probably caused by the connection of two origami through the vertex staples [Fig. 3.3]. Another fundamental aspect of this origami, and all the RNA scaffolded structures, is



(a) (b)

Figure 5.4: 5.4a - AFM image of the DB triangle origami. The quality of the image is much clearer then TEM. 5.4b - An example of origami purification using Amicon Ultra centrifugal filter: lane 3 - staples; lane 4 - R1 origami 10 μl; lane 5 - purified R1 origami 1 μl. The staples have been removed

the degradation of the sample when it is not handled in an RNase free environment. Unfortunately, all the experiments were performed in facilities that are not designated as RNase free, reducing the available time for the imaging before the samples degraded on the microscope support.

5.3 DB square and DB rectangle origami folding and analysis

Even if the results with the triangle origami showed that it is possible to fold a De Bruijn RNA/DNA hybrid origami, we wanted to improve the technique in order to have a more reliable platform for the future work. To address the misfolding problem Dr. Kozyra designed a new origami based on the same scaffold.

The idea was to synthesize an origami that could be seen not only at the AFM but also inside a cell or a vesicle. To do so the origami needs to carry a modification that can be easily detected like a fluorescent molecule added to the staples.

Two more design were tested in parallel, using the same DB Tri scaffold as in the triangle origami. The first design is a square origami presenting linker staples on the perimeter to allow the formation of a super-structure. The idea was to create an origami assembly that would have been easier to see at the microscope. Also, a non-linked version of this origami was designed to use as a control if the super-structure would not form.
The other new design is a rectangle that does not form links, the rectangle shape has been chose because it makes it easier to identify the different sides of the structure.
In all cases, the staple sets were optimized to have a low minimum free energy to reduce the formation of secondary structures as in the previous RNA origami.

5.3.1 DB square origami design

This is a very simple design, the scaffold is folded in a raster configuration and the resulting square has a side of approximatively 30 nm.
In the assembly version there are 26 staples, of which six are linkers that are meant to bind two domains in two different origami. When the origami are in position to form the superstructure, the domains are adjacent and the linker staple can bind both of them. The linked version also has longer intra-origami staples (not involved in the origami link) that are binding three different domains forming two crossing-over.
In the non-linked version, the linker staples are separated in the two domains. Also the three-domains binding staples are separated in two shorter ones to facilitate the folding, bringing the staple number up to 31.
To perform the experiments two of the central staples were modified to carry a fluorophore at their 5', one with Alexa Fluor 488 (green) and the other with Alexa Fluor 594 (red). All the staples were synthesized by Eurogentec.

5.3.2 DB rectangle 1 origami design

This origami is a standard rectangle (R1) with the scaffold arranged in a raster pattern [Fig. 5.14]. Like the non-linked square, this design does not have any crucial staple that could disrupt the whole structure if missing. The rectangle measures 35 nm by 25 nm and is folded by 31 staples.

5.3.3 R1 origami and DB square folding

Denaturation issue In the attempt to improve the origami synthesis workflow I applied the isothermal folding protocol described by Sobczak et al. [32]. This procedure assures an effective folding of a 2D origami in less than 5 minutes when incubating it at its exact folding temperature, after an initial denaturation. My idea was to reduce to a minimum the time spent from the scaffold synthesis to the AFM imaging, so to increase the quality of the results reducing the chances of RNase degradation.
To find the proper temperature for these origami I set up multiple folding reactions in parallel following a thermal gradient. This experiment showed a peculiar scaffold behaviour that went unnoticed during the triangle folding: when the origami mix undergoes a denaturing step of 75°C or more, the resulting product will not move from the agarose gel wells during the electrophoresis. This effect is probably caused by the formation of an unpredicted agglomerate [Fig. 5.5].

Not being able to identify the problem at first, I tested all sort of protocol variations:

- all the isothermal folding temperatures from 20°C to 72°C
- classic thermal ramp protocol (a first denaturation and a slow temperature reduction over a long time)
- scaffold modification removing the T7 terminator (renamed DB Tri short) suspecting it could be part of some strong scaffold-scaffold interaction
- increasing the holding temperature after the folding up to room temperature (RT) to disrupt any weak interaction

As a control, an unfolded origami mix left at room temperature was loaded on an agarose gel electrophoresis [Fig. 5.6]. A band was visible at the 1 Kb mark, higher than the scaffold band alone, indicating that some folding process was acting but without forming the agglomerate that all the processed samples showed.

Origami folding Based on the effect shown by the control sample I repeated the process for R1 origami removing the denaturation step to detect the best folding conditions. Temperatures between 20°C and 69°C were tested in intervals of about 1.3°C.

A smaller experiment was also performed with the linked square origami testing the new protocol at 20°C, 39°C and 54°C following the same protocol [Fig. 5.8].

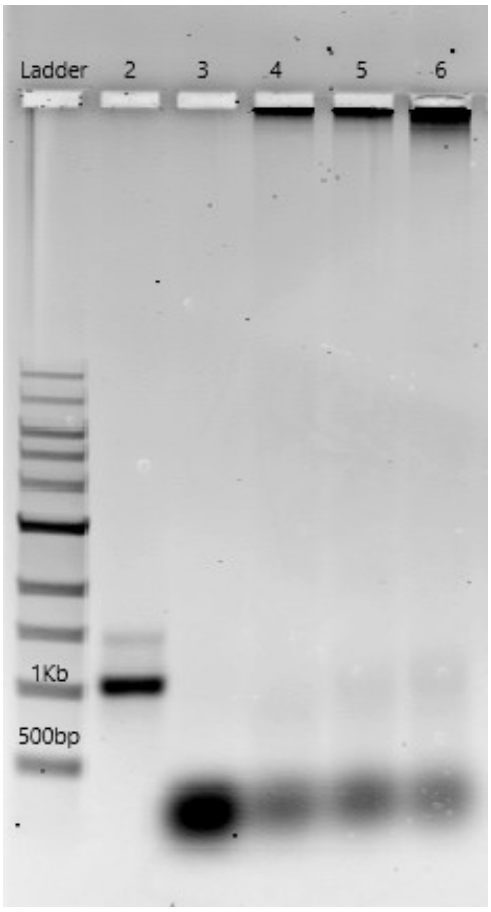


Figure 5.5: Agarose gel electrophoresis of: lane 1 - 1Kb DNA ladder (NEB); lane 2 - DB Tri RNA scaffold; lane 3 - R1 staple mix; lane 4 to 6 - R1 origami folding at 20°C - 37°C - 54°C post denaturation. It is possible to detect a strong signal on the lower edge of the origami wells.

DB981 R1 folding Before solving the denaturation problem the shortened version of the scaffold was ordered, named DB 981, suspecting the issue could be related to one of the regulating components. As previously described, this sequence lacks the LacO, T7 terminator and DB linkers surrounding the 981 nt used in the triangle folding. As one can notice the R1 origami is designed to cover a sequence including both the linkers (1026nt), this means that the first and last staples in R1 only bind one domain each, while the other acts as overhang.

The scaffold preparation proved to be more efficient, especially when amplifying the DNA template. Moreover, this is a true DB sequence without biological components [Fig. 5.9a]. The folding of the R1 origami was always performed with the isothermal short protocol without denaturation. Three temperatures were tested: 37°C for the *in vivo* possibility, 50°C and 55°C to check the temperatures around the theoretical optimal area [Fig. 5.9a,

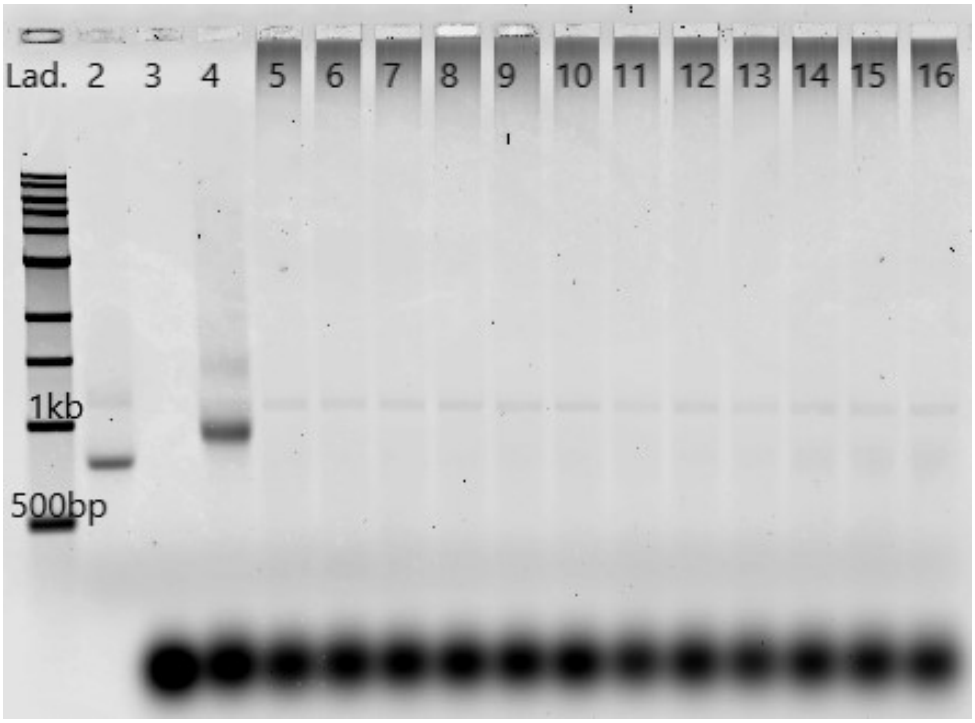


Figure 5.6: Agarose gel electrophoresis of: lane 1 - 1Kb DNA ladder (NEB); lane 2 - DB Tri RNA scaffold; lane 3 - R1 staple mix; lane 4 - R1 origami mix at RT; lane 5 to 16 - R1 origami folding from 22°C (5) to 38°C (16) post denaturation. The lane 4 band is higher than the scaffold in lane 2 suggesting an origami folding process.

5.9b].

After the folding, the samples were purified with Amicon Ultra centrifugal filters 100 KDa and visualized under the microscope.
A modified version of R1 folding was tested to investigate the effect of a 50:1 ratio of staples and scaffold to see if it could improve the yield.

I compared the isothermal folding at 55°C with the classic denaturation/thermal ramp protocol [Fig. 5.10].

The isothermal origami was also purified using Amicon Ultra centrifugal filters 100 KDa before imaging it at the AFM.

5.3.4 DB R1 origami imaging

The origami visualization at the AFM was performed improving the protocol used for the triangle origami. In this new version, the nickel chloride drop is added first to form the ionic bridges. After a brief incubation, it is removed with filter paper to reduce the

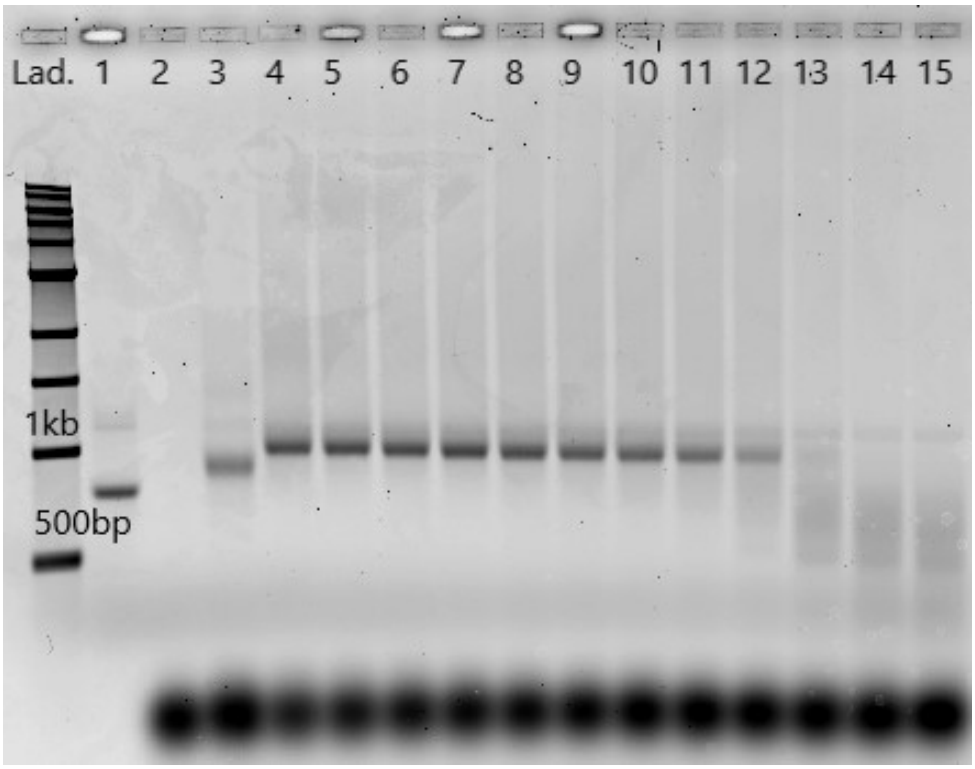


Figure 5.7: Agarose gel electrophoresis of: lane 1 - 1Kb DNA ladder (NEB); lane 2 - DB Tri RNA scaffold; lane 3 - R1 staple mix; lane 4 - R1 origami mix at RT; lane 5 to 16 - R1 origami folding from 53°C (5) to 69°C (16) post denaturation. In this agarose gel it is possible to see that the bands of the origami folded from 53°C (lane 5) to 64°C (lane 12) are clean and much higher than the control incubated at RT (lane 4). It is also possible to see how the origami starts to melt at temparatures higher than 65°C (lane 13).

background noise that is probably caused by the formation of nickel salt crystals. Also, nickel chloride is dissolved in water instead of TAE buffer to avoid chelation through EDTA. Finally, the AFM in scanning buffer used is the origami folding buffer instead of nickel chloride buffer.

It was clear from the AFM analysis that the only working protocol was the one with the folding set at 55°C. Nor 37°C or 50°C foldings formed any origami [Fig. 5.11, 5.12].

Folding the origami using the high staple ratio drastically increased the origami yield, but at the same time increased the background signal probably due to staple residues left over from the purification [Fig. 5.13].

5.3.5 DB R1 origami results

When using an isothermal folding with an incubation up to 15 minutes, R1 RNA/DNA hybrid origami correctly folds at 55°C, at the lower tested temperatures the folding is

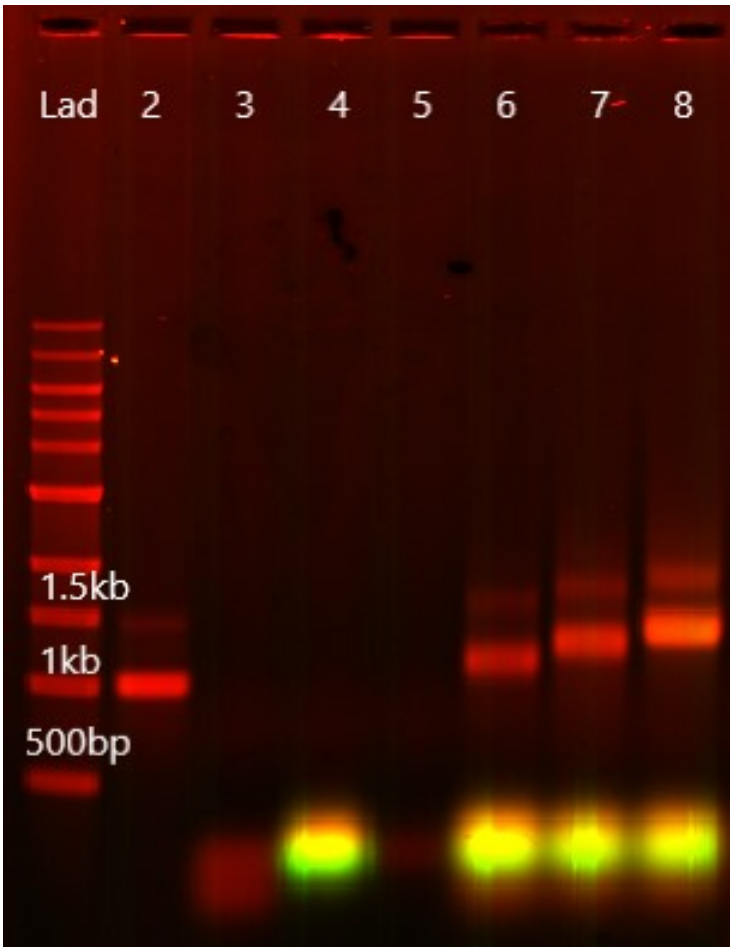


Figure 5.8: Agarose gel electrophoresis of: lane 1 - 1Kb DNA ladder (NEB); lane 2 - DB Tri RNA scaffold; lane 3 - square basic staple mix; lane 4 - square fluorescent staple mix; lane 5 - square linker staple mix; lane 6 - 20°C folding; lane 7 - 39°C folding; lane 8 - 54°C folding. As with R1 the higher band is located around 55°C.

partial or absent.

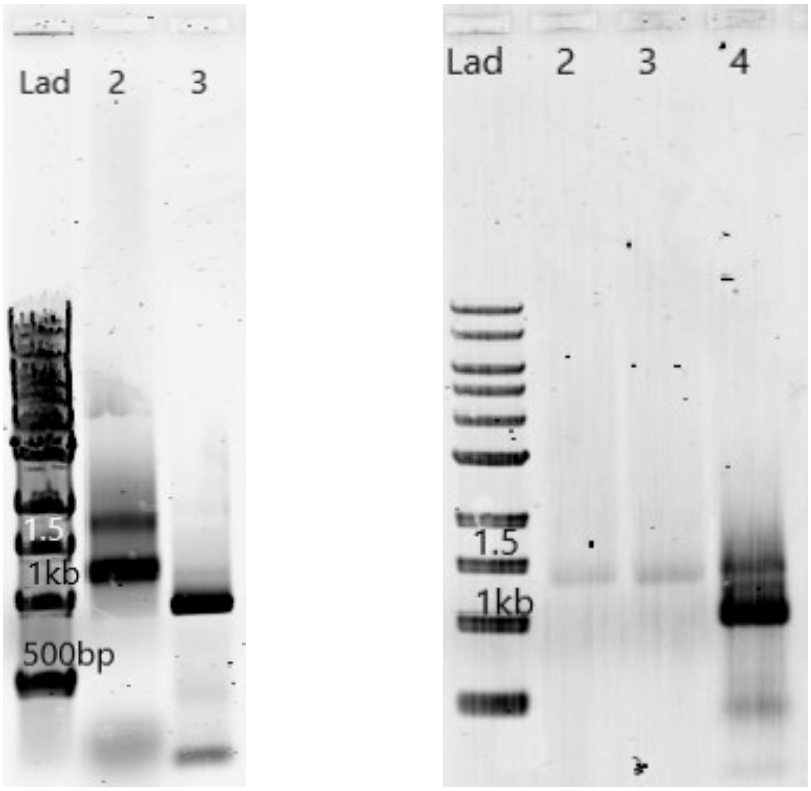
A higher staples concentration helps the folding increasing the yield of well-formed origami.

A peculiar behavior of all the RNA DB origami mix is detected after the denaturation step: an unknown assembly is formed and its physical properties do not allow it to run through an agarose gel.

5.4 Summary

These experiments demonstrate that it is possible to use a synthetic De Bruijn RNA scaffold to properly fold a hybrid RNA/DNA origami.

The results from this set of experiments confirmed Kopielski's paper [33] performed on



(a) (b)

Figure 5.9: 5.9a - Agarose gel electrophoresis of: lane 1 - 1Kb DNA ladder (NEB); lane 2 - DB 981 R1 origami 37°C folding; lane 3 - DB981 DNA template. 5.9b - Agarose gel electrophoresis of: lane 1 - 1Kb DNA ladder (NEB); lane 2 - DB 981 R1 origami 50°C folding, purified; lane3 - DB 981 R1 origami 55°C folding, purified; lane 4 - DB981 RNA scaffold.

phagic DNA origami. The optimizations obtained by the full sequence design proved efficient as expected, allowing the origami isothermal folding and producing a good yield of nanostructures. DB R1 origami also showed a predisposition for the folding at lower temperatures, even if it this feature was not designed on purpose. These findings are very promising in perspective of the *in vivo* folding, an aimed refinement of the scaffold sequence could probably allow a complete synthesis at life-compatible temperatures for the most part of the microorganisms. In the long term vision, a synthetic scaffold has a greater potential granted by the optimization capabilities without sacrificing any molecular characteristic.

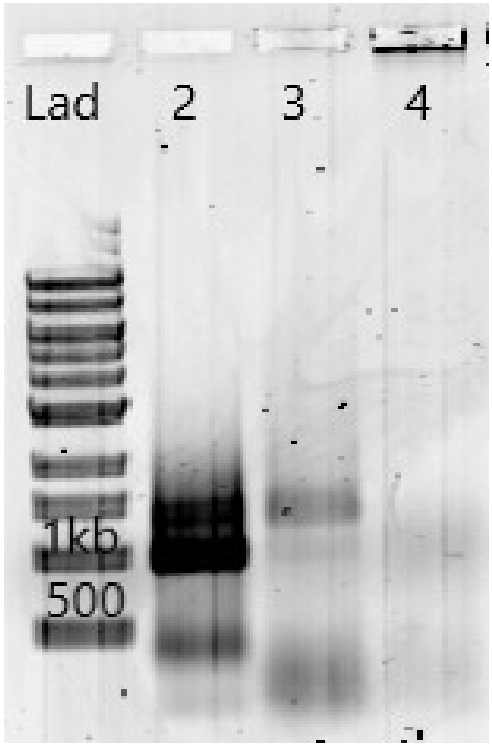
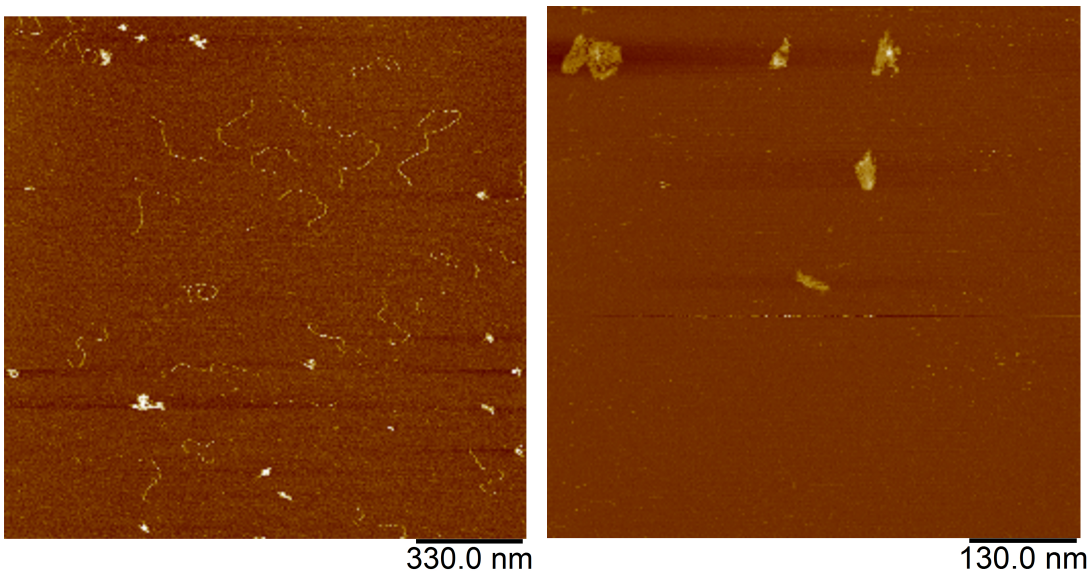


Figure 5.10: Agarose gel electrophoresis of: lane 1 - 1Kb DNA ladder (NEB); lane 2 - DB 981 RNA scaffold; lane 3 - R1 origami 50X staples isothermal 55°C folding; lane 4 - R1 origami 50X staples classic denaturation/thermal ramp folding. While the isothermal folding shows an origami band (lane 3), the classic folding forms the agglomerate in the gel well (lane 4).



(a) (b)

Figure 5.11: 5.11a - R1 origami 37°C folding shows only traces of unfolded scaffold; 5.11b - R1 origami 50°C folding shows a series of misfolded or partially folded origami.

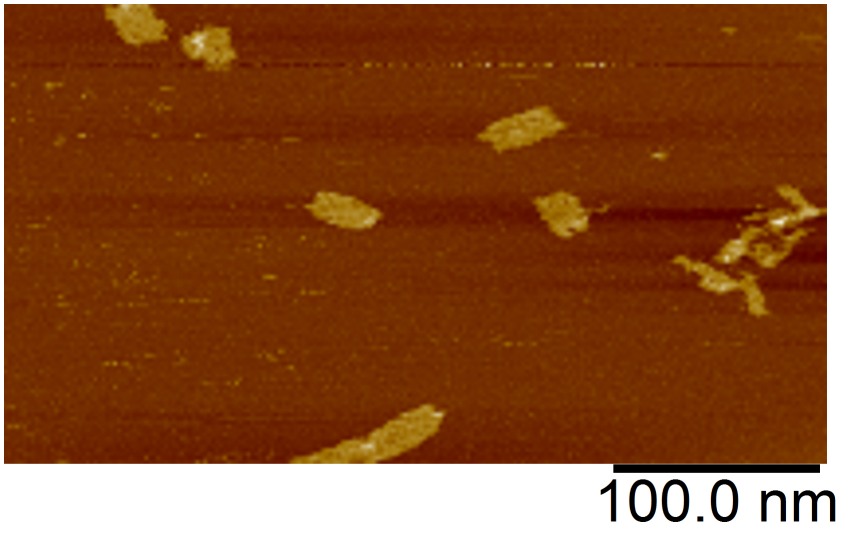
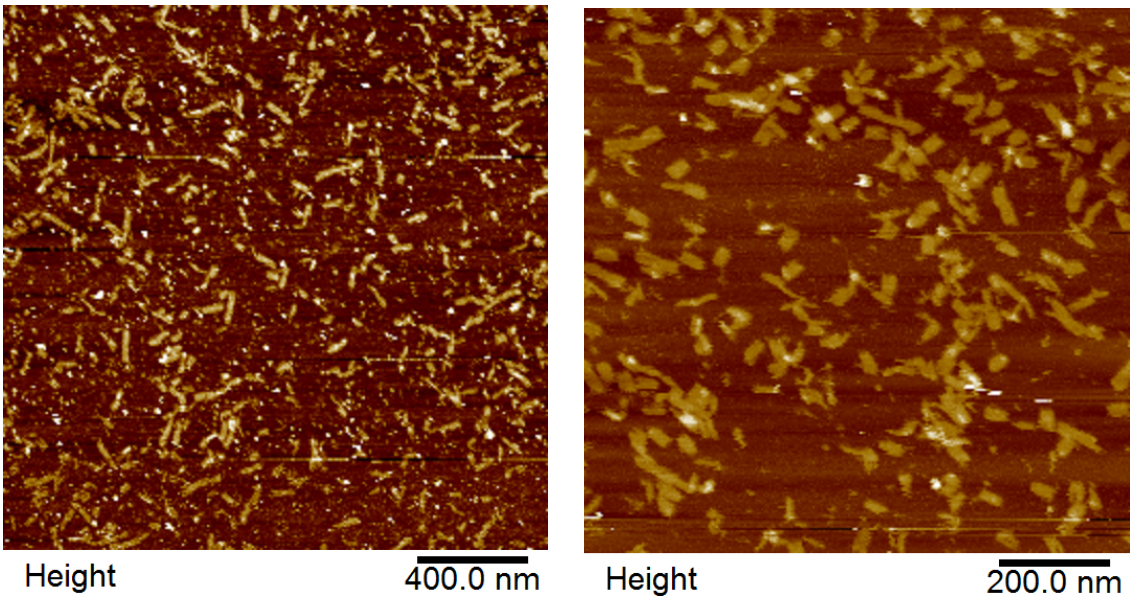


Figure 5.12: R1 origami 55°C folding shows a good number of well formed rectangular origami, on the right is visible an origami probably broken by the AFM probe contact.



(a) (b)

Figure 5.13: R1 origami 55°C folding using a 50:1 = staple:scaffold ratio. The folding efficiency is very high, a large amount of origami is visible; 5.13a - a panoramic scan with higher definition. 5.13b - a zoomed-in image with lower definition. Increasing the zoom means that the probe spends more time over a single object, the ScanAsyst system uses a lower strength on the Z-axis to reduce the chances of interaction between the probe and the substrate, reducing at the same time the definition of the image. This effect is amplified by the presence of unpaired staples.

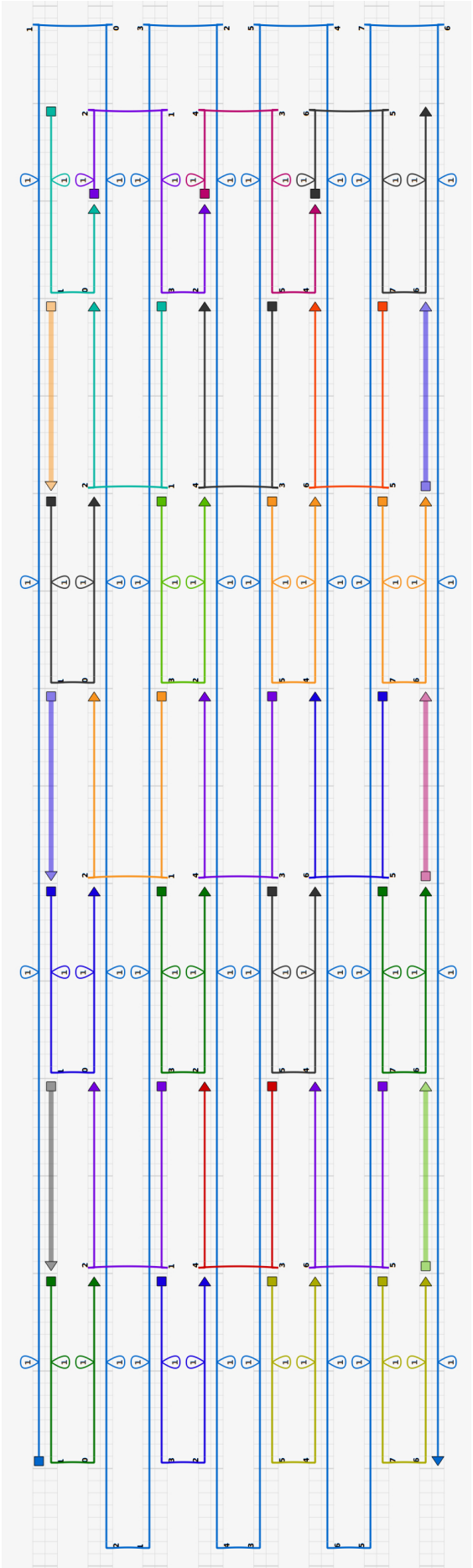


Figure 5.14: CADnano design of the R1 rectangular origami.

Chapter 6

In vivo DB scaffold study

The design of a DB synthetic scaffold relies on the key concepts of bio-compatibility and bio-orthogonality. In this chapter, I describe the experiments performed *in vivo* to characterize the impact that these sequences have on the cell transcription and translation systems. I designed a series of genetic constructs that could measure the impact of DB sequences (2.4 and 981) when inserted between the promoter and the ribosome binding site (RBS) of the superfolder Green Fluorescent Protein (sfGFP) gene. I compared the results obtained with these sequences with controls replacing each DB sequence with a section of *E. coli* polymerase gene *dnaE* of the same length. *dnaE* was selected for its dimension, allowing the selection and amplification of an internal sequence that does not contain known promoters or terminators, it was also chosen for being an *E. coli* gene, a common bacteria easily found in many microbiology stocks.

sfGFP is an engineered fluorescent protein encoded by a 714 nt long gene. It is a more robust version of the natural GFP and it conserves the same excitation (488 nm) and emission (530 nm) wavelengths [108].

The protocol for creating these constructs is generated using the NEBuilder Assembly tool, the online software that designs the proper primers for the NEBuilder HiFi DNA kit reaction (http://nebuilder.neb.com/). The kit allows the building of a genetic construct in a one-pot reaction containing all the linearized sequences that will compose the final plasmid.

6.1 pBAD33 based construct

The first attempt to create an expression system to analyse the effect of a DB sequence involved the use of pBAD33 plasmid as a backbone for the construct.
pBAD33 is a common cloning expression vector for *E. coli*, it contains a chloramphenicol (Cam) resistance cassette and the arabinose PBAD promoter [109].

6.1.1 De Bruijn 2.4 reporter gene construct

In order to clone DB 2.4 and RBS-$sfGFP$ into pBAD33 vector using NEBuilder HiFi DNA Assembly master mix (NEB) I followed the suggested kit protocol:
DB 2.4 and RBS-$sfGFP$ were amplified with a PCR reaction (Q5 PCR kit NEB)

DB 2.4 primers for HiFi reaction in pBAD33:

- DB 2.4 fwd: AGCGAATTCGAGCTCGGTACCTTTCTATGTCTGAGCCTG
- DB 2.4 rev: CTCTTTAATTTTTCGCCCTAACTATGTC

RBS-$sfGFP$ primers for HiFi reaction in pBAD33:

- (RBS)-$sfGFP$ fwd: TAGGGCGAAAAATT(AAAGAGGAGAAA)CCTAG
- $sfGFP$ rev: CTCATCCGCCAAAACAGCCAAGCTTTATTATTTGTAGAGCT-CATCC

The PCR products were cleaned using Monarch DNA purification kit (NEB) before the final assembly.

6.1.2 pBAD33-DB2.4-sfGFP expression

To test *E. coli*-5α with pBAD33-DB2.4-sfGFP six cultures were prepared, each in 2 ml LB + Cam 25μg/μl. Three were induced with arabinose 0.02% at the inoculation and three were left uninduced as control. After 5 hours The samples were analysed in the flow cytometer (Partec CyFlow space) to check the sfGFP fluorescence. The cytometer was controlled manually and all the samples reached at least 50000 events. The results showed that the induction had almost no effect [Fig. 6.1], meaning that the DB 2.4 completely disrupts the activity of PBAD promoter.

Suspecting that the non-coding DB sequence could have been the cause of a short processivity of the RNA polymerase, decoupling itself from the template before reaching the gene, the construct was cloned on a different plasmid.

6.2 pET-28a based constructs

pET-28a is an expression vector that is designed to work in a T7 RNA polymerase expression system. The relevant components of this plasmid are the T7lac promoter inducible with IPTG and the kanamycin (Kan) resistance cassette.

DB, *dnaE* and RBS-*sfGFP* sequences were cloned using the NEBuilder HiFi kit. As for the previous experiment the primers were designed using the online NEBuilder assembly tool.

6.2.1 De Bruijn 2.4 reporter gene construct

DB 2.4 and RBS-*sfGFP* inserts were amplified with Q5 PCR (NEB):

DB 2.4 primers for HiFi reaction in pET-28a:

- DB 2.4 fwd: GAGCGGATAACAATTCCCCTCTAGATTTCTATGTCTGAGCCTG
- DB 2.4 rev: CTCTTTAATTTTTCGCCCTAACTATGTC

RBS-*sfGFP* primers for HiFi reaction in pET-28a:

- (RBS)-*sfGFP* fwd: TAGGGCGAAAAATT(AAAGAGGAGAAA)CCTAG
- *sfGFP* rev: GGTGCTCGAGTGCGGCCGCAAGCTTTTTATTATTTGTAGAGCTCATCC

The components where assembled using the NEBuilder HiFi kit into pET-28a and transformed in *E. coli* T7 Express competent cells.
14 colonies of *E. coli* T7 Express were selected and inoculated in LB + Kan (50 μg/μl) to check for the presence of the right construct. The plasmids from each culture were extracted using a Miniprep plasmid extraction kit (Qiagen) and digested with XbaI and HindIII to cut out the insert and check if the dimensions were compatible with the desired

sequences [Fig. 6.2].

Plasmids from colony 14, the only plasmid of the right size, all the other plasmids presented a smaller size with an insert compatible with RBS-$sfGFP$ sequence only.
Plasmid 13 and 14 were sent for sequencing (MRC PPU DNA Sequencing and Services) confirming the agarose gel electrophoresis findings.
Plasmid 14 became pET-28a-DB 2.4-$sfGFP$ [Fig. 6.3].
Plasmid 13 was kept as positive control and renamed pET-28a-$sfGFP$ [Fig. 6.4].
The cultures containing the plasmids where finally conserved in stocks at -80°C.

6.2.2 *dnaE* 2.4 reporter gene construct

As a control sequence to compare with DB 2.4 I choose a fragment of *E. coli* MG1655 *dnaE* DNA polymerase III subunit alpha gene. This gene was chosen for its ready availability and its dimensions: being 3483 nt long it allowed to choose an internal sequence large as DB 2.4 excluding all the known bioactive sequences as promoters or terminators. The gene was conveniently amplified from a chromosomal DNA stock of our lab using a Q5 PCR kit.

dnaE primers for chromosome PCR (Eurogentec):

- *dnaE* 2.4 forward: CCGCACCGTTGGTAAAAAAG
- *dnaE* 2.4 reverse: CTTCGATGATGGCCTCAATC

The NEBuilder HiFi reaction was designed to add the *dnaE* product on the pET-28a-sfGFP control plasmid [Fig. 6.5]. The vector was digested with XbaI (NEB).
Primers for *dnaE* 2.4 NEBuilder reaction (Eurogentec):

- *dnaE* 2.4 HiFi forward: GAGCGGATAACAATTCCCCTCTAGACCGCACCGTTG-GTAAAAAAG
- *dnaE* 2.4 HiFi reverse: TCAGGCTCAGACATAGAAATCTAGACTTCGATGATG-GCCTCAATC

All the components were purified using Monarch DNA purification kit before setting up the NEBuilder HiFi reaction mix.

- HiFi buffer 1x
- pET-28a-XbaI-(RBS)sfGFP 0,1 pmol
- *dnaE*2.4 HiFi 0,2 pmol

The mix was incubated at 50°C for 60 minutes.

The plasmid obtained was transformed in *E. coli* T7 Express, and checked with a colony PCR [Fig. 6.6].

The two colonies presenting the proper insert were sent for sequencing for the final confirmation. As for the other samples the colonies were stocked at -80°C.

6.2.3 De Bruijn 2.4 and *dnaE* 2.4 *in vivo* expression

The comparison of the different constructs was performed in parallel and involved:

- *E. coli* T7 Express pET-28a-DB 2.4-sfGFP
- *E. coli* T7 Express pET-28a-*dnaE* 2.4-sfGFP
- *E. coli* T7 Express pET-28a-sfGFP

The cultures were analysed with flow cytometry after 4 hours and 7 hours from the induction, using the cytometer (Partec CyFlow space) and analysing 50000 events for each culture [Fig. 6.7].

The stand-alone sfGFP has the best expression overall. Unexpectedly, it is clear from the analysis that the vector carrying the DB 2.4 sequence is the one with the worst efficiency. *dnaE* 2.4 is placed halfway, clearly reducing sfGFP expression, although not to the same extent as the DB sequence.

6.2.4 De Bruijn 981 and *dnaE* 981 reporter gene constructs

To further investigate the bio-orthogonality characteristics I also compared DB 981 with a 981 nt fragment of *dnaE* [Fig. 6.8 and 6.9].

This time the protocol used for the synthesis of the two constructs was the same. pET-28a-$sfGFP$ was digested with XbaI (NEB) and purified with Monarch DNA purifi-

cation kit.
The two sequences for the cloning were amplified using a Q5 PCR kit (NEB).

Primers for DB 981 HiFi (Eurogentec):

- DB 981 HiFi forward: GAGCGGATAACAATTCCCCTCTAGAGCGGTCCAGCTAGCAGGT
- DB 981 HiFi reverse: TCAGGCTCAGACATAGAAATCTAGACTGTACTGCGACGACCTAAATCG

Primers for *dnaE* 981 HiFi (Eurogentec):

- *dnaE* 981 HiFi forward: GAGCGGATAACAATTCCCCTCTAGAGATTTCACCAACCTTTGTGGTCTG
- *dnaE* 981 HiFi reverse: TCAGGCTCAGACATAGAAATCTAGACAGTGAACCCGCACCGGA

Both the inserts were purified with Monarch DNA kit.
The NEBuilder HiFi kit was used as suggested by its protocol:

- HiFi buffer 1x
- pET-28a-XbaI-(RBS)sfGFP 0,1 pmol
- 981 nt insert HiFi 0,2 pmol

Both incubated at 50°C for 60 minutes. The new vectors were transformed in *E. coli* T7 express competent strains (NEB).

As for the other vectors, the correct synthesis was confirmed by colony PCR [Fig. 6.10] and sequencing of the chosen colony (MRC PPU DNA Sequencing and Services) after Miniprep extraction (Qiagen).
The strains containing the sequenced vectors were stocked at -80°C.

Primers for pET-28a colony PCR:

- pET28 forward: CTATAGGGGAATTGTGAGCGG
- pET28 reverse: CCCTCAAGACCCGTTTAGAGG

6.2.5 De Bruijn 981 and *dnaE* 981 *in vivo* expression

The same experimental procedures performed for DB 2.4 sequences were used to compare the strains containing the 981 nt samples.

- *E. coli* T7 Express pET-28a-DB 981-sfGFP
- *E. coli* T7 Express pET-28a-*dnaE* 981-sfGFP
- *E. coli* T7 Express pET-28a-sfGFP

The fluorescence was analysed with a FACSCanto II cytometer after 4 and 6 hours of growth at 37°C in the shaker. The analysis was controlled manually and each sample reached at least 35000 events.

This analysis shows that sfGFP alone is always the most efficient construct, but this time the DB 981 has a slightly better protein expression than the *dnaE* 981 control sequence [Fig. 6.11].

6.2.6 Summary

The *in vivo* experiments described in this chapter tried to assess the bio-orthogonality of the De Bruijn sequences I used as origami scaffolds. These experiments analysed the impact of the DB sequences in the transcription efficiency of a gene cloned at their 3'. Obviously, the conclusions that can be inferred by the data collected are limited to the transcription-translation process that happens on the vector transformed.

The first important conclusion is the necessity for a T7 RNA polymerase mediated transcription to produce a non-coding synthetic sequence. The use of the native transcription system was not functional, probably because the RNA polymerase loses its processivity when the target sequence does not interact with the translation machinery: this could

happen to save resources, because the bacteria does not transcribe a sequence that will not be translated. A T7 polymerase system could facilitate a tight control over the production, controlling the polymerase expression through an induction system, but on the other hand, it requires the use of dedicated strains as *E. coli* T7 Express. It also demonstrates that the non-coding DB sequence does not show any transcription activating feature.
As expected both the vectors with an additional cloned sequence have a lower efficiency compared to $sfGFP$ alone. The transcriptional burden is more than doubled with the 981 nt long sequences and almost four-fold higher with the 2.4 knt ones.
The data shows also how the two DB sequences behave differently compared to the controls. The larger DB 2.4 shows a higher negative impact on the final expression of sfGFP, while the protein synthesis when the shorter DB 981 is present, is much closer to the control strain. The causes for this behavior are not clear, as the database analysis does not find any match with other known sequences, they can only be hypothesized. It could depend on the activity of secondary structures due to the peculiar thermodynamic design, or on the presence of an unknown regulatory factor encoded in the sequence itself.

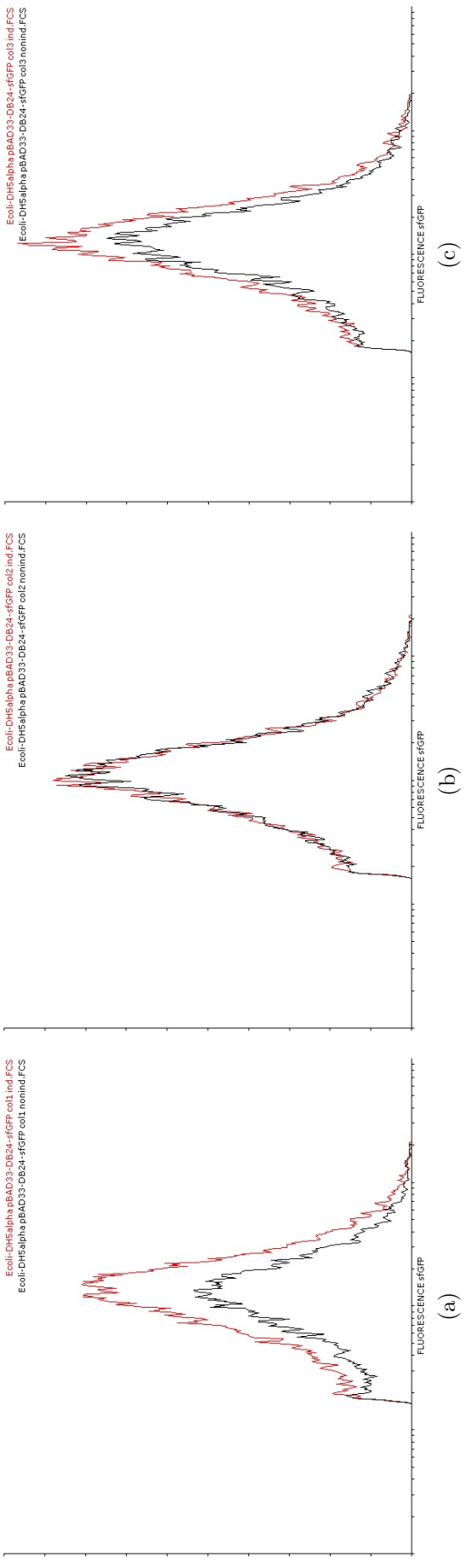
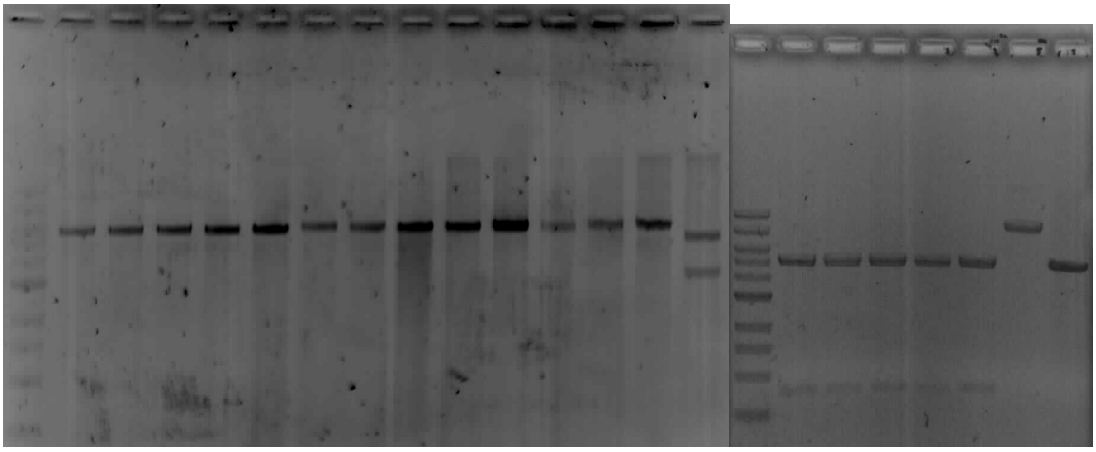


Figure 6.1: pBAD33-DB2.4-sfGFP flow cytometry results, three replicates: on the X-axis is the fluorescence intensity, on the Y-axis the cell count. The fluorescence intensity peaks are overlapping showing no differences between the induced and uninduced samples.



(a) (b)

Figure 6.2: 6.2a - Agarose gel electrophoresis of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 to 14 - plasmid extracted from colonies 1 to 13 digested with HindIII; lane 15 - plasmid extracted from colony 14 digested with XbaI and HindIII . 6.2b - Agarose gel electrophoresis of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 to 6 - plasmid extracted from colonies 9 to 13 digested with XbaI and HindIII; lane 7 - plasmid extracted from colony 14 linearized with HindIII; lane 8 - pET-28a double digestion with XbaI and HindIII as control.

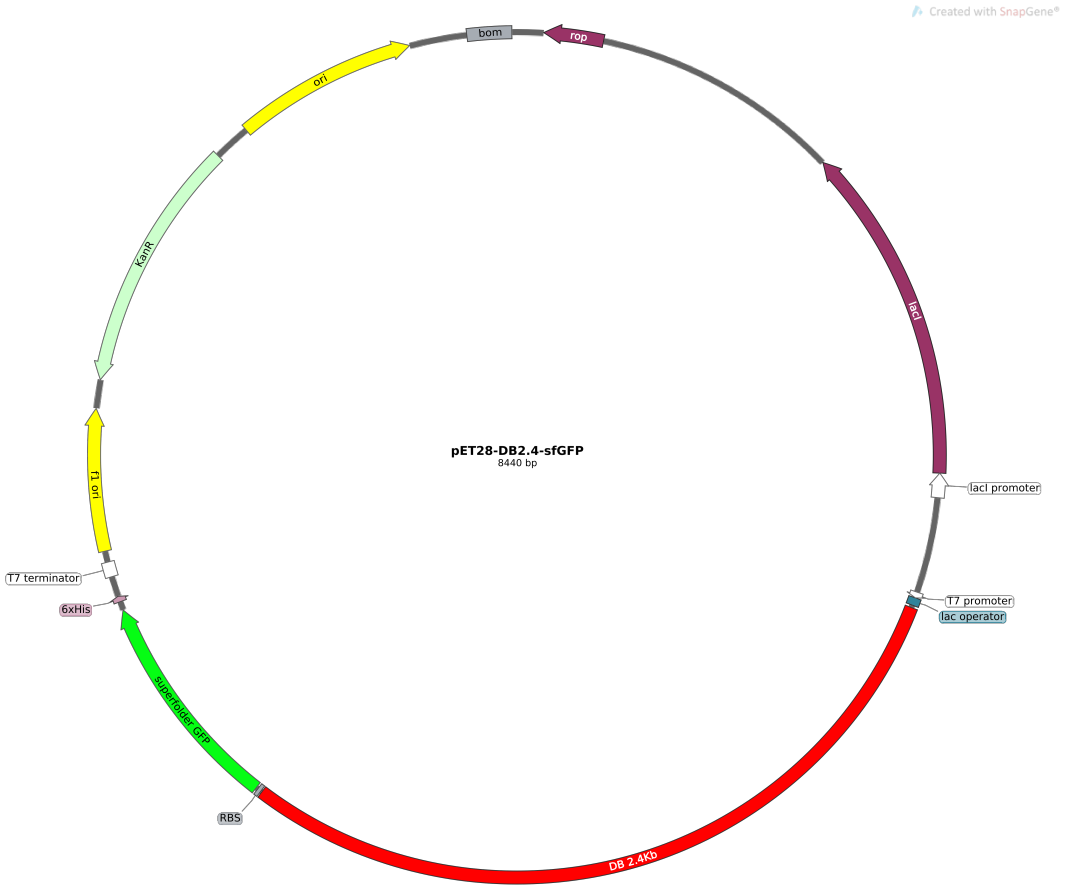


Figure 6.3: pET-28a-DB 2.4-sfGFP map.

Figure 6.4: pET-28a-sfGFP map.

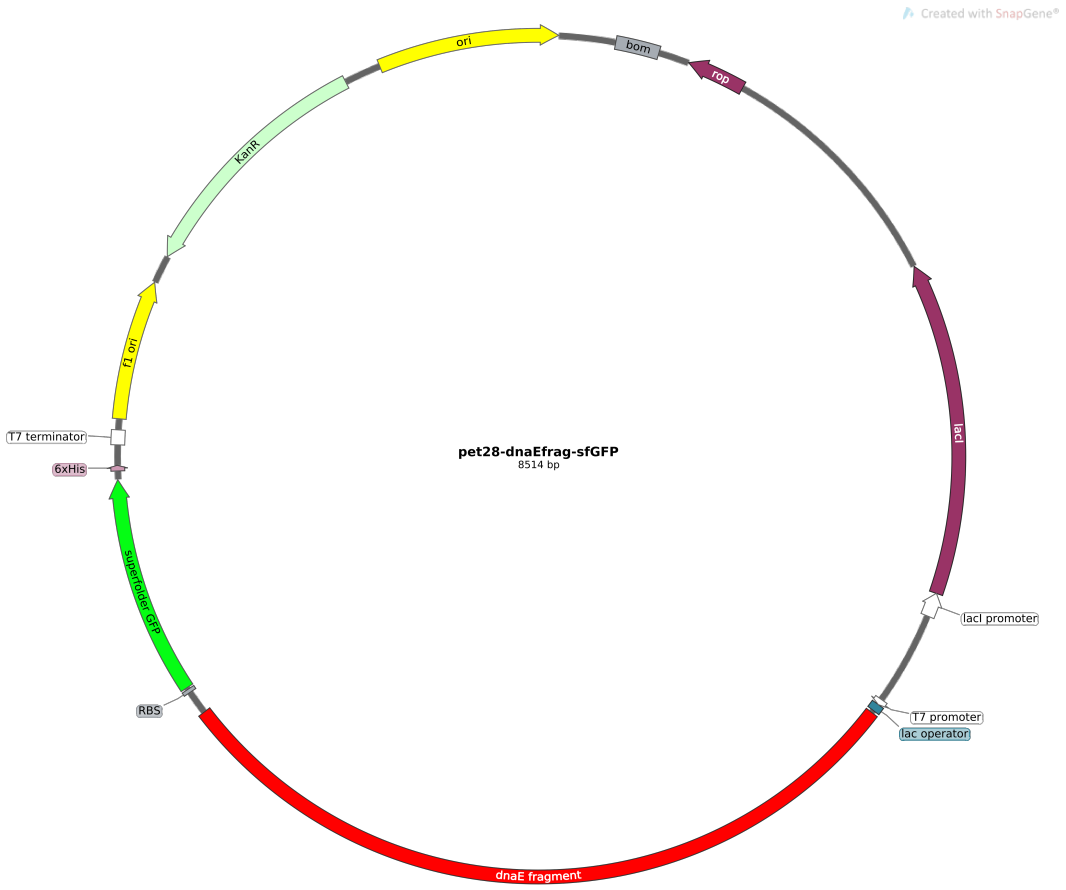


Figure 6.5: pET-28a-*dnaE* 2.4-sfGFP map.

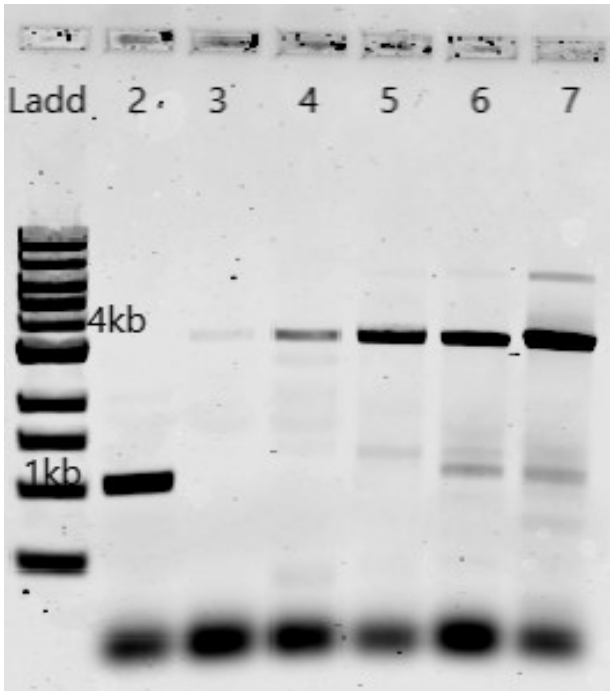


Figure 6.6: Agarose gel electrophoresis of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 - *dnaE* 2.4 control; lane 3 to 8 - *dnaE* 2.4 colony PCR from pET-28a-*dnaE* 2.4-sfGFP transformants. It is possible to see a band at control level only in lanes 6 and 7.

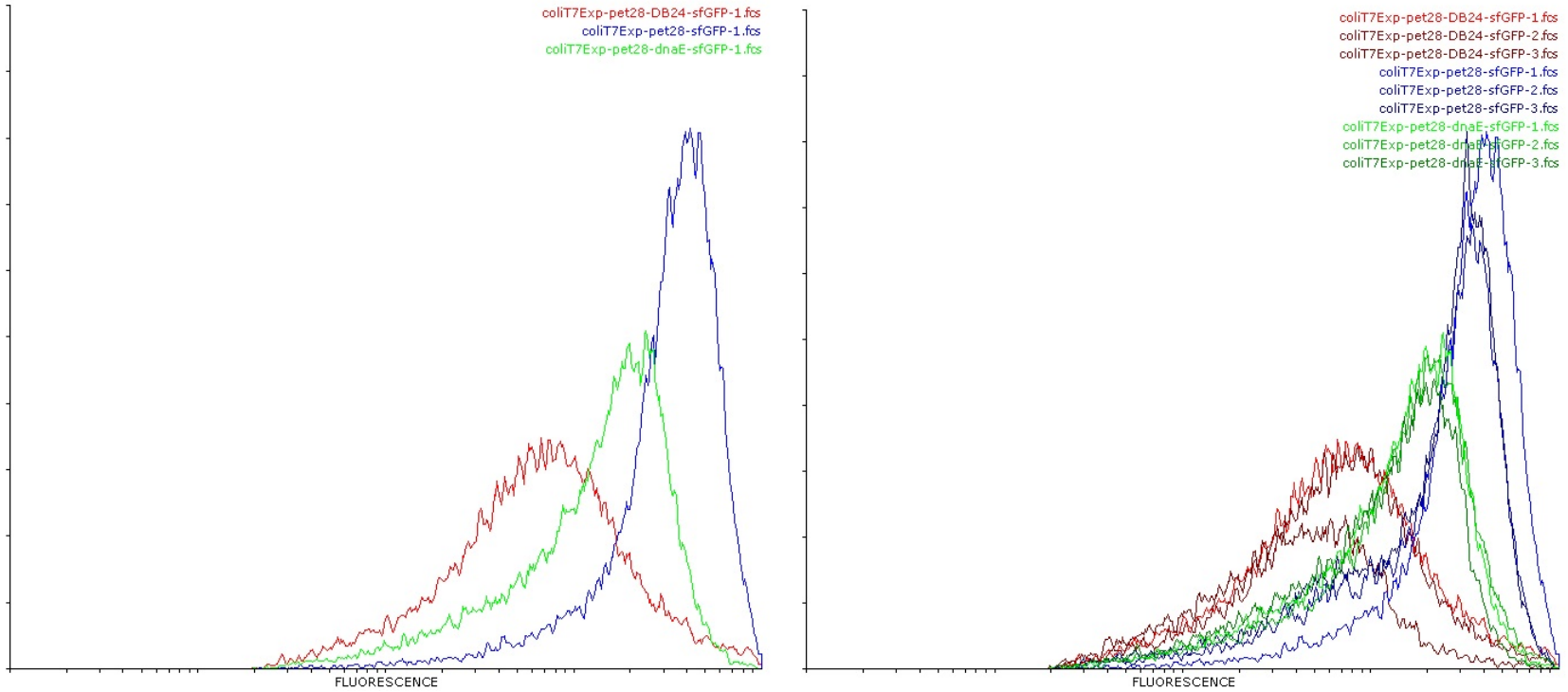


Figure 6.7: 6.7a - pET-DB 2.4-sfGFP (red), pET-*dnaE* 2.4-sfGFP (green), pET-sfGFP (blue) flow cytometry results, 7 hours after IPTG induction. On the X-axis is the fluorescence intensity, on the Y-axis the cell count. 6.7b - Three replicates of the experiment in one overlay graph.

Figure 6.8: pET-28a-DB981-sfGFP Map.

Figure 6.9: pET-28a-*dnaE* 981-sfGFP Map.

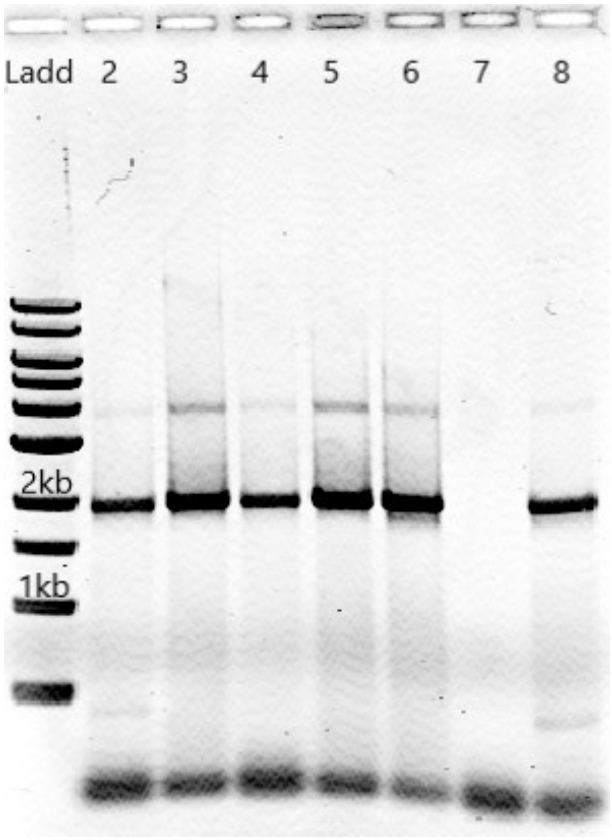
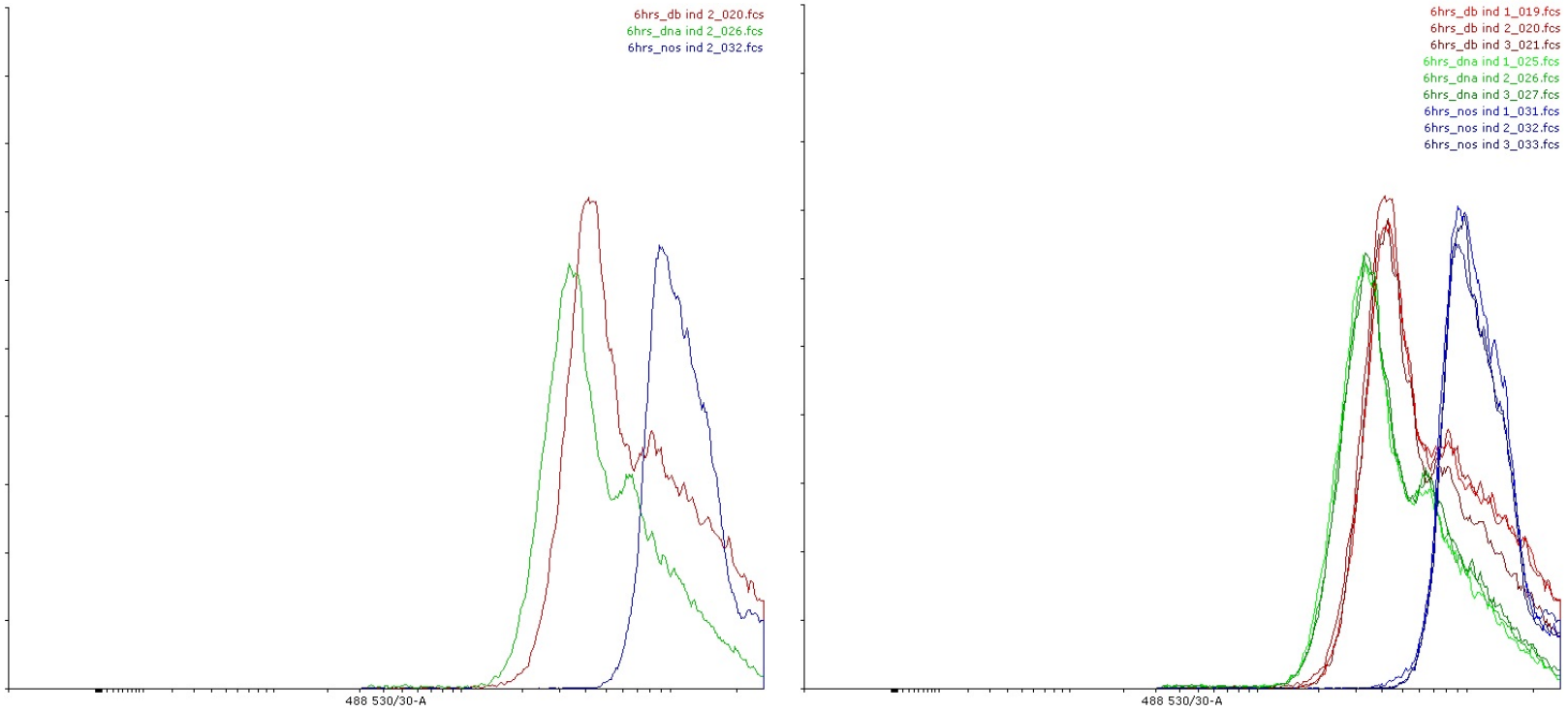


Figure 6.10: Agarose gel electrophoresis of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 to 5 - pET-28a-DB 981-sfGFP PCR ; lane 6 to 8 - pET-28a-*dnaE* 981-sfGFP PCR colony PCR. PCR in lane 7 is faulty for the lack of template. All the colonies tested showed a correct PCR product.



(a) (b)

Figure 6.11: 6.11a - pET-DB 981-sfGFP (red), pET-*dnaE* 981-sfGFP (green), pET-sfGFP (blue) flow cytometry results, 6 hours after IPTG induction. On the X-axis is the fluorescence intensity, on the Y-axis the cell count. 6.11b - Three replicates of induced cultures in one overlay graph.

Chapter 7

Origami folding thermodynamics analysis tool

This chapter is focused on the synthetic biology prototype toolset designed to study origami folding. These devices demonstrate nucleic acid origami and giant unilamellar vesicles technologies working in synergy. As vesicles can be used as protocells, they can provide part of the *in vivo* conditions avoiding the drawbacks of uncontrolled factors. Vesicles content can also be characterized through fluorescence analysis using high-throughput instruments. These capabilities can allow the detection and quantification of the folding of a functionalized origami. The origami staples are modified so they generate a signal upon the correct annealing on the scaffold. Different staples solutions were analysed to achieve a good folding detection, investigating different fluorescent aptamers and fluorophores. The origami folding reporter tool is not strictly bound to the vesicles system and can also be used in the classic *in vitro* experiments. The aim of these experiments was the creation of two tools that allow understanding better the origami folding process. In their simplest version, the origami components are inserted inside the vesicles allowing to follow the folding process using a fluorescence effect.

7.1 DB origami fluorescent reporters

The analysis of the origami folding requires a system that allows discriminating between a folded origami and an unfolded one.

To solve this problem I investigated the feasibility of different fluorescence complexes im-

plemented on the RNA/DNA hybrid R1 origami described in chapter 5.3.2. The concept behind this experiment is the use of the staples as folding reporters, when the marked staples are correctly annealed on the origami scaffold, they will interact emitting a signal. I tested three different systems:

- Broccoli aptamer [110]
- Malachite Green aptamer [66]
- Cy-3/Cy-5 FRET effect

7.1.1 Broccoli aptamer

The broccoli aptamer is an RNA sequence, derived from the Spinach aptamer, that forms an hairpin-like secondary structure. This structure is capable of binding a DFHBI-1T fluorescent molecule, considerably increasing its emission. The broccoli aptamer has been tested *in vitro* and *in vivo*. One of the most interesting applications is the use of a split aptamer, where the two halves of the sequence are extensions of two RNA staple oligonucleotides [66].

Broccoli sequence: GUAUGUGGGAGACGGUCGGGUCCAGAUAUAUCUGUCGAGUA-GAGUGUGGGCUCCCACAUAC
Broccoli split 1: GUAUGUGGGAGACGGUCGGGUCCAGAUA
Broccoli split 2: UAUCUGUCGAGUAGAGUGUGGGCUCCCACAUAC

When the two oligos are annealed to a common substrate, the two halves of the aptamer are adjacent and will reconstitute the aptamer activity. The experiment design started with the selection of three couples of staples to use as connectors for the aptamers halves. The first half extended from the 3' of one staple and the second from the 5' of the staple next to it. The couples selected were (aptamer sequence in brackets):

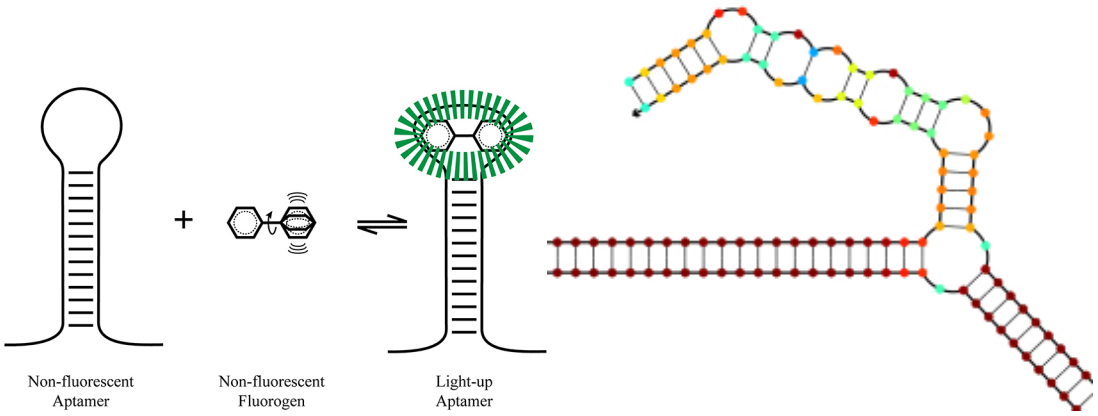


Figure 7.1: 7.1a - Broccoli aptamer can bind the molecule DHFBI-1 and stabilize it in its fluorescent state. 7.1b Broccoli structure at 25°C.

Staple 17 and 30

staple 17: GUUUAACCCACCGAGUCCUAUCACUUGUACCC (GUAUGUGGGAGACG-GUCGGGUCCAGAUA)
staple 30: (UAUCUGUCGAGUAGAGUGUGGGCUCCCACAUAC) AACACAACAGCU-CUUC

Staple 14 and 27

staple 14: ACGAAGCCUGACUGUGAUUCAGCGUAAAAACUCG (GUAUGUGGGA-GACGGUCGGGUCCAGAUA)
staple 27: (UAUCUGUCGAGUAGAGUGUGGGCUCCCACAUAC) CGAUCUCGAGAACUGUC-CGGUGCGAUAUUACUAU

Staple 4 and 6

staple 4: GAGUGAUCGUUGCUAUCGAGCUAUUUCAAGCCCG (GUAUGUGGGA-GACGGUCGGGUCCAGAUA)
staple 6: (UAUCUGUCGAGUAGAGUGUGGGCUCCCACAUAC) UGGAAGCUGUACUGCG

To space apart the signals I chose one staple couple from the 5' of the portion of the scaffold (17-30), one from the 3'(4-6) and one from the middle portion (14-27)[Fig.7.2].

For this experiment, only the aptamer staples were synthesized as RNA, all the other

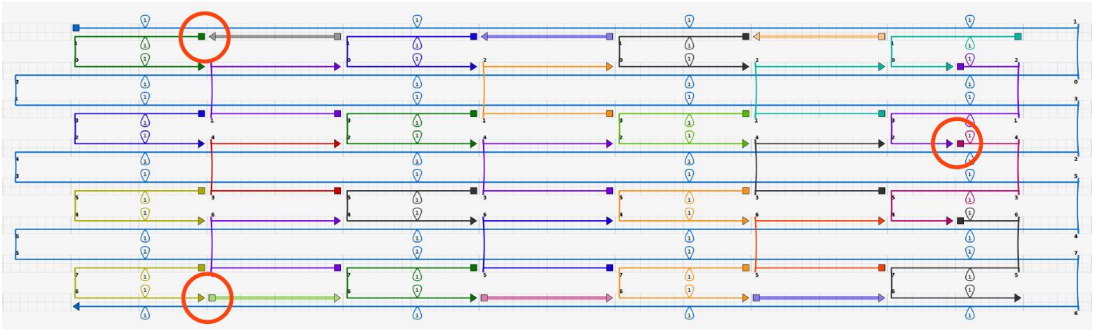


Figure 7.2: The split Broccoli aptamers location on R1 origami is marked with red circles. The Broccoli aptamer is split in two halves, each half is added as a staple extension on two consecutive staples protruding from the origami surface one next to the other. This configuration allows the regeneration of the functional aptamer structures when the origami is properly folded. Aptamers: "17-30" top-left circle; "4-6" bottom-left circle; "14-27" centre-right circle.

origami staples were the same DNA oligos from the original design.

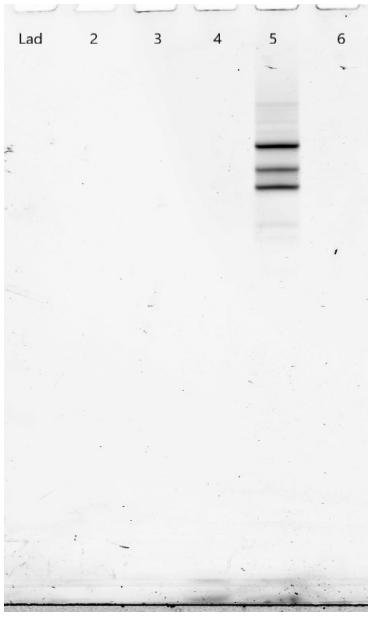
First, the modified staples were tested in a simpler assembly, using a target sequence complementary to the two staples domains to check the correct aptamer reconstitution. In these experiments, I mixed the three strands in a one-pot reaction following an origami isothermal folding protocol.

Broccoli aptamer regeneration protocol: the reaction is composed of: modified staple 1 [500 nM], modified staple 2 [500 nM], target sequence [500 nM], Hepes [10 mM], KCl [100 mM] and $MgCl_2$ [1mM].

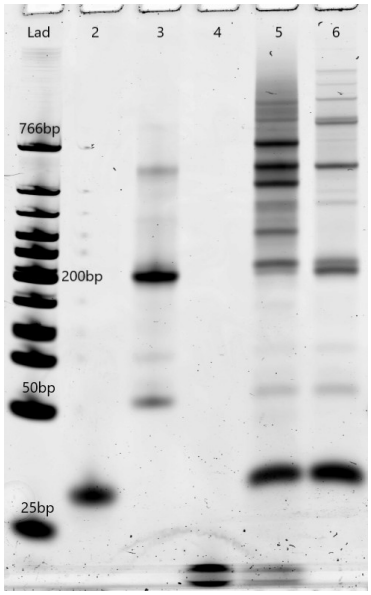
The thermal folding is a single step at 55°C for 15 minutes and a hold step at 4°C

The samples were loaded on a polyacrylamide gel.

The results show how all the couples regenerate the aptamer when the target sequence is added, but also that couples 14-27 and 4-6 regenerate it without the connector sequence in the control experiment [Fig. 7.3 and 7.4].

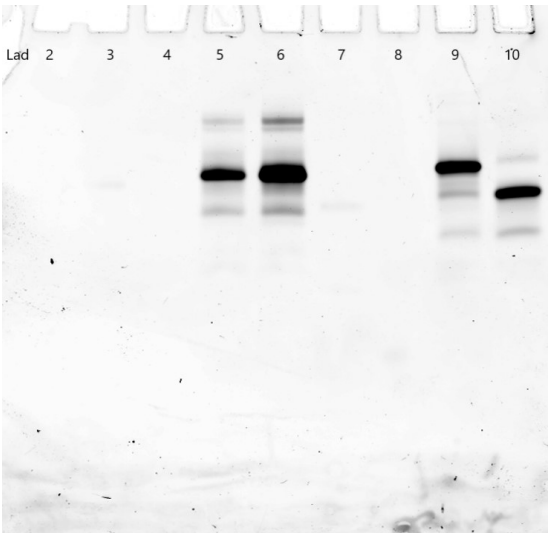


(a)

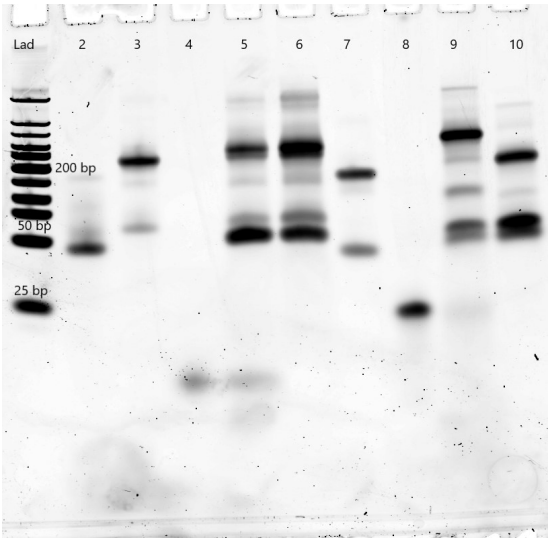


(b)

Figure 7.3: Polyacrylamide gel electrophoresis of: lane 1 - LMW DNA ladder (NEB); lane 2 - staple 17; lane 3 - staple 30; lane 4 - 17-30 target; lane 5 - staples 17-30 and target; lane 6 - staples 17-30 no target. 7.3a Gel after DFHBI staining. 7.3b Same gel after SYBR gold staining.



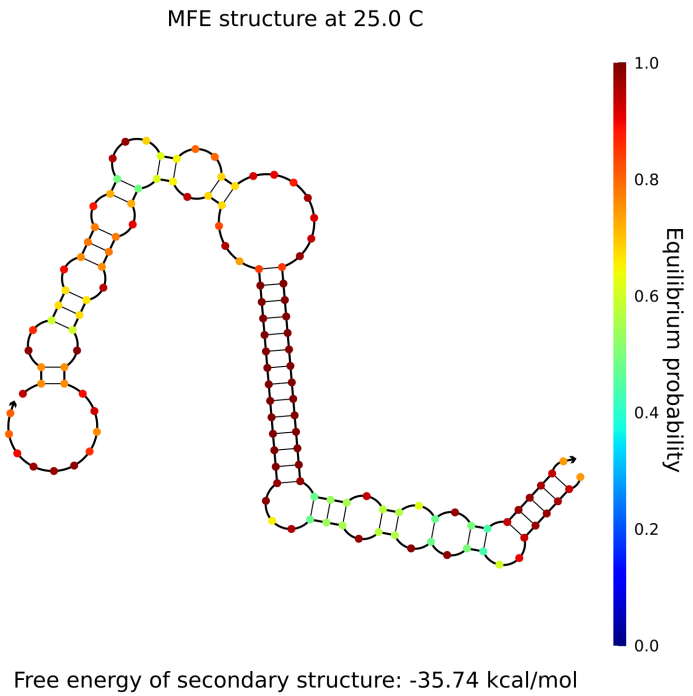
(a)



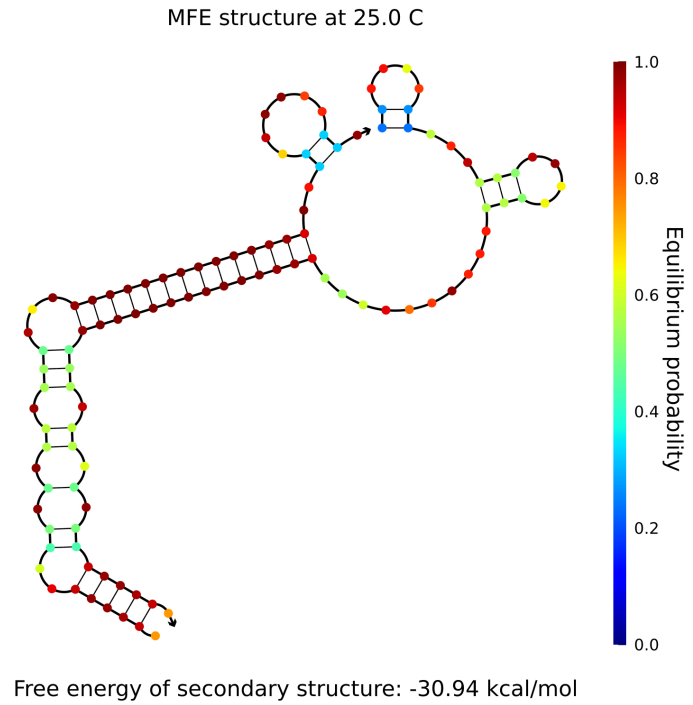
(b)

Figure 7.4: Polyacrylamide gel electrophoresis of: lane 1 - LMW DNA ladder (NEB); lane 2 - staple 14; lane 3 - staple 27; lane 4 - 14-27 target; lane 5 - staples 14-27 and target; lane 6 - staples 14-27 no target; lane 7 - staple 4; lane 8 - staple 6; lane 9 - 4-6 target; lane 10 - staples 4-6 and target; lane 11 - staples 4-6 no target . 7.4a Gel after DFHBI staining. 7.4b Same gel after SYBR gold staining.

To understand why the aptamers form also without target sequences, I analysed the staple combinations using the NUPACK software (http://www.nupack.org/). The analysis confirms that the aptamer structure is the most stable conformation that the oligos can assume when reacting in two or more copies [Fig. 7.5]. The analysis was performed considering only a maximum of four interacting oligos to reduce the computational time.



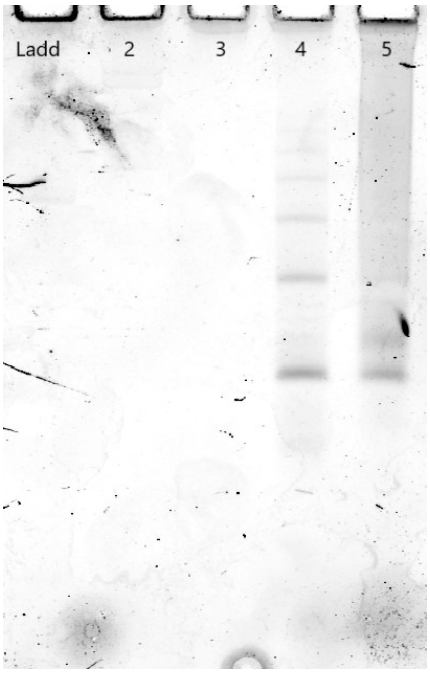
(a)



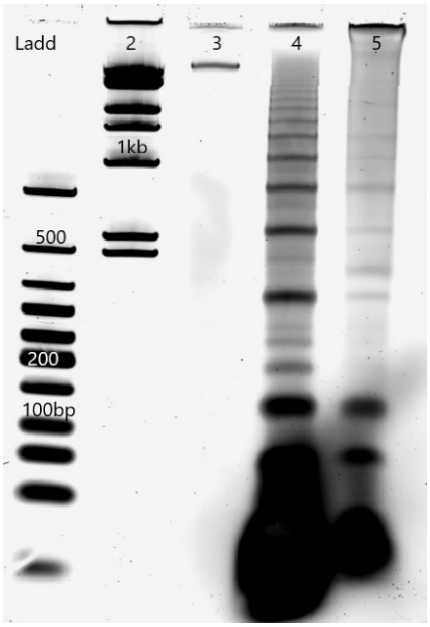
(b)

Figure 7.5: 7.5a - Staples 14-27 interaction. 7.5b - Staples 4-6 interaction. NUPACK analysis of two couples of oligonucleotides composing the Broccoli aptamer used in the experiments. In both cases the thermodynamic properties of the sequences used regenerate the aptamer even without the help of a target strand (or the origami scaffold). The aptamer structure is recognizable on the right on fig. 7.5a and on the left on fig. 7.5b

As the couple 17-30 showed an expected behavior, I performed an origami folding test, substituting the original 17-30 staples with the version carrying the aptamer parts. The electrophoresis result shows again the formation of multiple active aptamers at different band levels [Fig. 7.6]. Probably the presence of the other 29 staples generates a favorable condition for the aptamer formation.



(a)



(b)

Figure 7.6: Polyacrylamide gel electrophoresis of: lane 1 - LMW DNA ladder (NEB); lane 2 - 1 Kb DNA ladder (NEB); lane 3 - DB 1026 RNA scaffold; lane 4 - staple mix; lane 5 - DB 1026 R1 origami with Broccoli aptamer. 7.6a Gel after DFHBI staining. 7.6b Same gel after SYBR gold staining. A multiple aptamer formation is visible in both staple (lane 4) and origami (lane 5) lanes.

For the purpose of this experiment the thermodynamical optimization of the origami staples plays unfavorably. The staples, as a matter of fact, are designed to avoid the

formation of secondary structures and they find stability in the hairpin stems generated by the aptamer nucleotides. To overcome this problem I designed a new set of split broccoli aptamer staples, with a shorter hairpin stem with a lower stability.

7.1.2 Short broccoli aptamer

Using the NUPACK software I analysed a new design for the split Broccoli aptamer staples that has a shorter hairpin stem, removing 8 nt (Bros - Broccoli short), reducing the chances of the aptamer regeneration in the absence of the scaffold sequence. Testing the new aptamer on NUPACK for all the origami staple couples, only the couples 17-30 Bros and 24-20 Bros showed a behaviour compatible with the experiment. It is important to notice that the staples 24 and 20 are not in line but parallel, their 3' and 5' respectively are lying on a crossover spot, and so they are oriented in the same direction, increasing the chances of having the aptamer split sequences on the same face of the origami [Fig. 7.7].

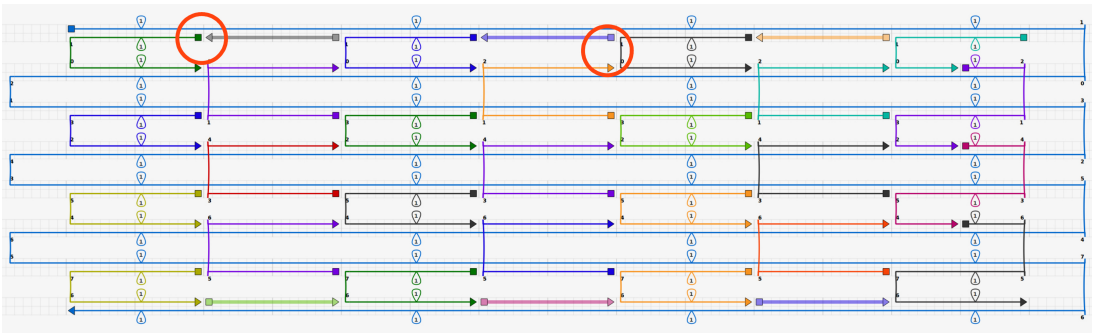
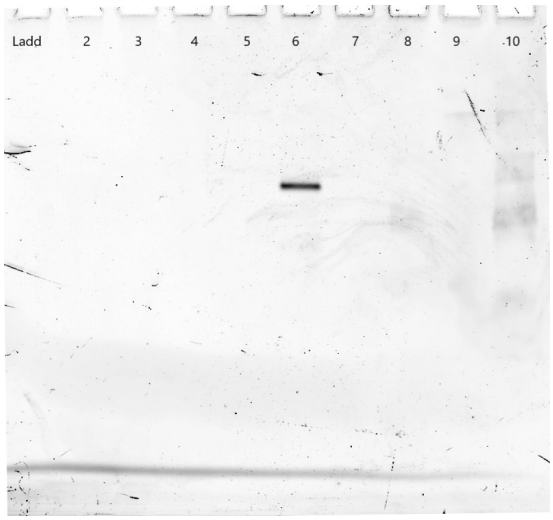
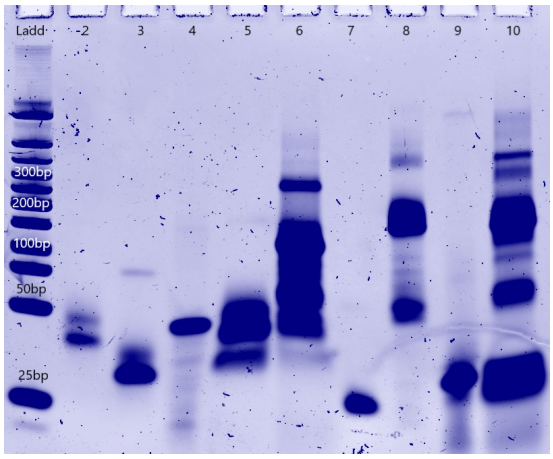


Figure 7.7: The shorter Broccoli aptamers location on R1 origami marked with red circles. This is another split Broccoli aptamer design, where each half of the aptamer is added as a staple extension on two consecutive staples protruding from the origami surface one next to the other. On the right cirlce is the 24-20 couple, the two halves are located on two different scaffold strands, at a cross-over point, coming out on the same side of the origami.

The new staples were tested following the same folding protocol as in the first design. Unfortunately, the couple 24-20 Bros showed no aptamer regeneration while the couple 17-30 Bros showed fluorescence as in the longer design [Fig. 7.8a].



(a)



(b)

Figure 7.8: Polyacrylamide gel electrophoresis of: lane 1 - LMW DNA ladder (NEB); lane 2 - staple 17 Bros; lane 3 - staple 30 Bros; lane 4 - target sequence for staples 17-30; lane 5 - 17-30 Bros; lane 6 - 17-30 Bros and target sequence; lane 7 - staple 24 Bros; lane 8 - staple 20 Bros; lane 9 - target sequence for staples 24-20; lane 10 - 24-20 Bros and target sequence. 7.8a Gel after DFHBI staining. 7.8b Same gel after SYBR gold staining. Only the staples 17-30 Bros show DFHBI-1T fluorescence (lane6).

A pilot experiment was then performed using a plate reader to identify any difference between a folded origami and an unfolded one. Theoretically, there should be a high difference in fluorescence between the samples. Unfortunately, the thermodynamics optimization of the origami design, combined with the high laboratory temperatures and the lack of cooling system in the plate reader, compromised the experiment: the unfolded control origami began the folding process before the analysis was completion, showing no difference in the fluorescence output against the folded samples (data not shown).

7.1.3 Malachite green aptamer

The second design for the folding reporter uses the malachite green (MG) split aptamer and was tested in parallel to the short Broccoli aptamer as an alternative option.
Malachite green sequence: UUUCCCGACUGGCCAGGUAACGAAUGGAUU
Malachite green split 1: UUUCCCGACUGG
Malachite green split 2: CCAGGUAACGAAUGGAUU

The NUPACK analysis was also performed on all the potential modified couples of the R1 origami to avoid the formation of unwanted secondary structures. MG aptamer is much smaller and with a lower stability still only the 4-6 couple simulation worked properly and was chosen in addition to 17-30 and 24-20 [Fig. 7.9].

The staple sequences are (aptamer sequence in brackets):

Staple 17 and 30

staple 17 MG: GUUUAACCCACCGAGUCCUAUCACUUGUACCC (UUUCCCGACUGG)
staple 30 MG: (CCAGGUAACGAAUGGAUU) AACACAACAGCUCUUC

Staple 14 and 27

staple 24 MG: AGCUGGACCGCCACGA (UUUCCCGACUGG)
staple 20 MG: (CCAGGUAACGAAUGGAUU) UCACAGAACCGUGCGCCGCGGGGC-CCCGUUUGUG

Staple 4 and 6

staple 4 MG: GAGUGAUCGUUGCUAUCGAGCUAUUUCAAGCCCG (UUUCCCGACUGG)
staple 6 MG: (CCAGGUAACGAAUGGAUU) UGGAAGCUGUACUGCG

The folding experiments using the target sequence were performed in the same aptamer buffer used with Broccoli.

The Malachite green aptamer regeneration protocol was composed of: modified staple 1 [500 nM], modified staple 2 500 [nM], target sequence [500 nM], Hepes [10 mM], KCl [100 mM] and $MgCl_2$ [1mM]. The oligos were annealed following a thermal folding at 55°C for 15 minutes, and holding the sample at 4°C

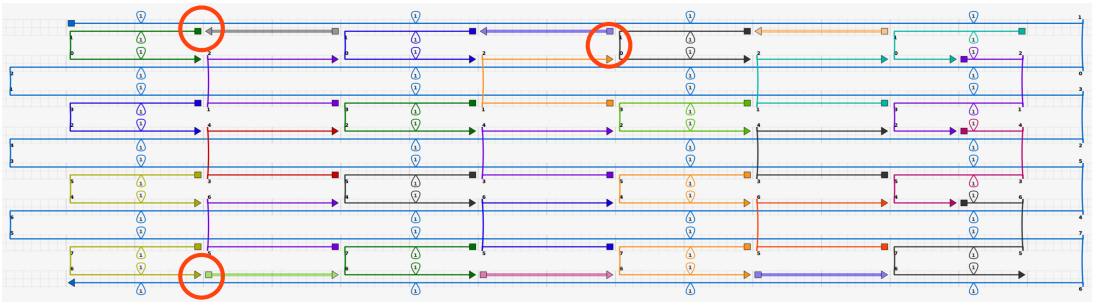


Figure 7.9: MG staples location on R1 origami marked with red circles. The split Malachite green aptamer halves are added as extensions on two adjacent staples allowing the regeneration of the full aptamer when the origami is correctly folded. 17-30 top-left circle; 4-6 bottom-left circle; 24-20 top-centre circle.

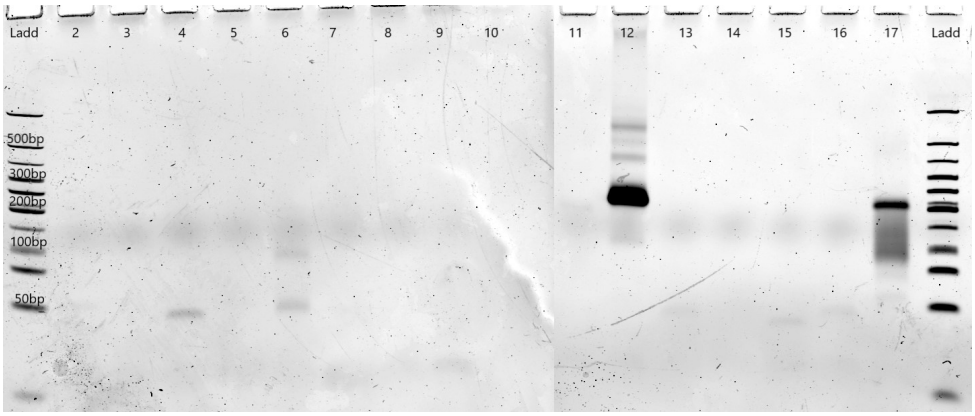
The samples were analysed in a 10% polyacrylamide gel in TBE.
The results show that the only staples not able to regenerate a functional aptamer are 17-30 MG, while 24-20 MG and 4-6 MG generate a strong fluorescence [Fig. 7.10].

All the MG staple couples were tested also in the origami folding, alone and in combination. The origami were folded using the isothermal protocol at 55°C for 15 minutes as for the other R1 origami, Using a staple mix concentration 10-fold higher than the scaffold. The samples were analysed through AGE and the MG gel staining was performed in folding buffer.

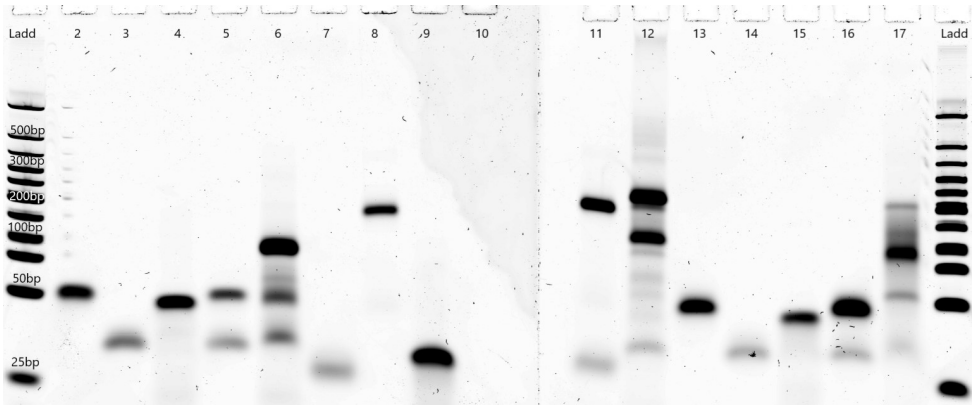
The results show that the strongest aptamer on the origami is the couple 4-6 MG, but at the same time it can reconstitute a functional aptamer just interacting with some other staple. Surprisingly the couple 17-30 works properly but shows a lower fluorescence. Finally, the couple 24-20 shows the weakest fluorescence of all [Fig. 7.11].
Another plate reader pilot experiment was performed to check the difference between a folded origami reporter and its unfolded control, the results showed again how the DB sequence optimization facilitates the origami folding at room temperature activating the aptamer signal [Fig. 7.12].

7.1.4 FRET

The last option developed for the construction of an origami-folding reporter take advantage of FRET effect. Two staples are functionalized with Cy-3 and Cy-5 modifications, when the staples are annealed to the right scaffold domain, the fluoroscopes are close to each other and the FRET interaction happens. When Cy-3 is excited with a 532 nm laser



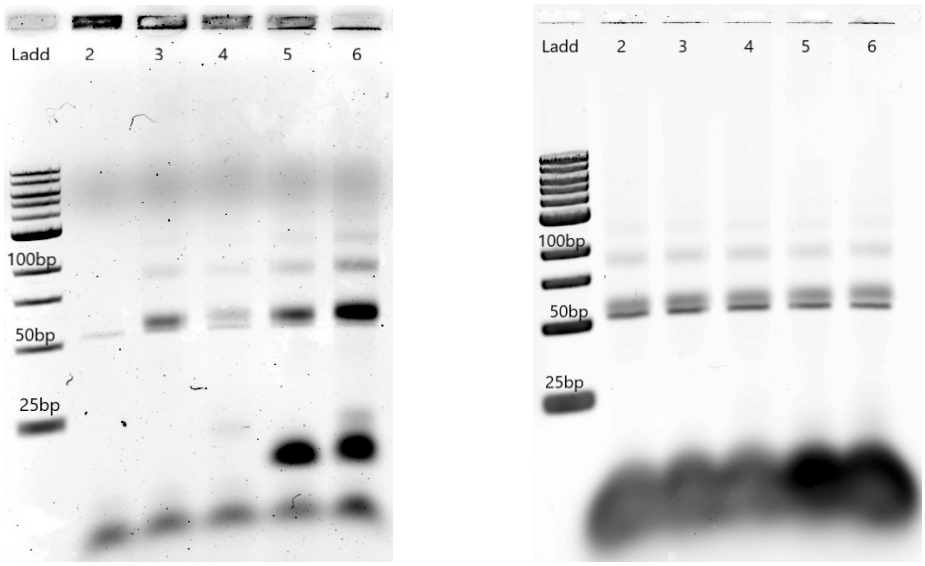
(a)



(b)

Figure 7.10: Polyacrylamide gel electrophoresis of: lane 1 - LMW DNA ladder (NEB); lane 2 - staple 17 MG; lane 3 - staple 30 MG; lane 4 - target sequence for staples 17-30; lane 5 - 17-30 MG; lane 6 - 17-30 MG and target sequence; lane 7 - staple 24 MG; lane 8 - staple 20 MG; lane 9 - target sequence for staples 24-20; lane 10 - empty; lane 11 - 24-20 MG; lane 12 - 24-20 MG and target sequence; lane 13 - staple 4 MG; lane 14 - staple 6 MG; lane 15 - target sequence for staples 4-6; lane 16 - 4-6 MG; lane 17 - 4-6 MG and target sequence; lane 18 - LMW DNA ladder (NEB). 7.10a Gel after MG staining. 7.10b Same gel after SYBR gold staining. Staples 24-20 MG (lane 12) and 4-6 MG (lane 17) can regenerate the aptamer, while staples 17-30 MG (lane 6) are not working.

it emits a fluorescence at 568 nm, this emission is sufficient to excite Cy-5 that will emit in turn at 666 nm. When the staples are not in the proper position the effect should be much lower or absent, exciting Cy-3 will not activate Cy-5. I based this experiment on the paper published by Niederholtmeyer et al. [111], where the two FRET fluorophores are positioned at 5 nt distance from each other. Considering that the functionalization can be added at a 5' or 3' end of the oligo, or internally on a thymine nucleotide, in the R1 origami only two staple couples are compliant with these parameters [Fig. 7.13].



(a) (b)

Figure 7.11: Polyacrylamide gel electrophoresis of: lane 1 - LMW DNA ladder (NEB); lane 2 - DB R1 origami; lane 3 - DB R1 origami 17-30 MG; lane 4 - DB R1 origami 24-20 MG; lane 5 - DB R1 origami 4-6 MG; lane 6 - DB R1 origami all MG staples . 7.11a Gel after MG staining. 7.11b Same gel after SYBR gold staining. Mg fluorescence is weak with 17-30 MG staples (lane3), almost absent with 24-20 MG (lane 4) and very strong with 4-6 MG (lane 5). Unfortunately, 4-6 MG generates a strong fluorescence also in the staple mix, indicating a non-specific aptamer regeneration.

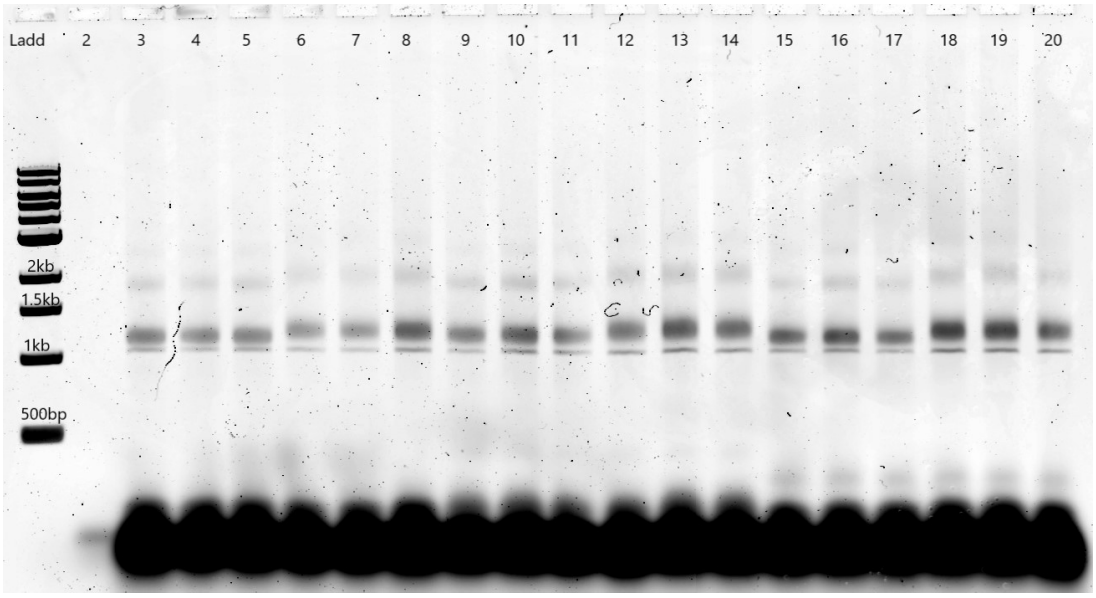


Figure 7.12: Mix of DB 1026 R1 origami after plate reader analysis. All the control samples started the folding process due to the high temperatures forming bands close to the origami sample.

Staple 13 and 3 FRET

staple 13 Cy-3: ACGCTAGAGAC(T-Cy3)CCGT

staple 3 Cy-5: Cy5-TCTGGTAGTCGAAAACATGTGCGTCTAAGTGCGG

Staple 26 and 25 FRET

staple 26 Cy-3: GTTCTAATTCT(T-Cy3)GAATCTCCAGACTTTTGGGC

staple 25 Cy-5: CY5-TATAACAATAGTTACCCAGAAATTAGAATTTCGA

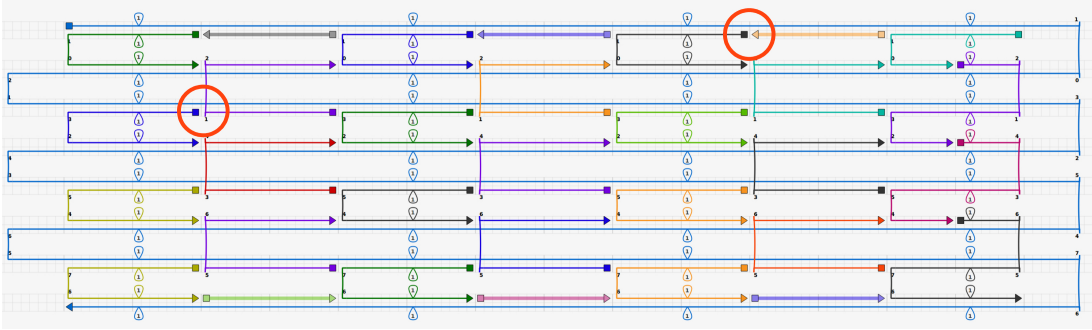
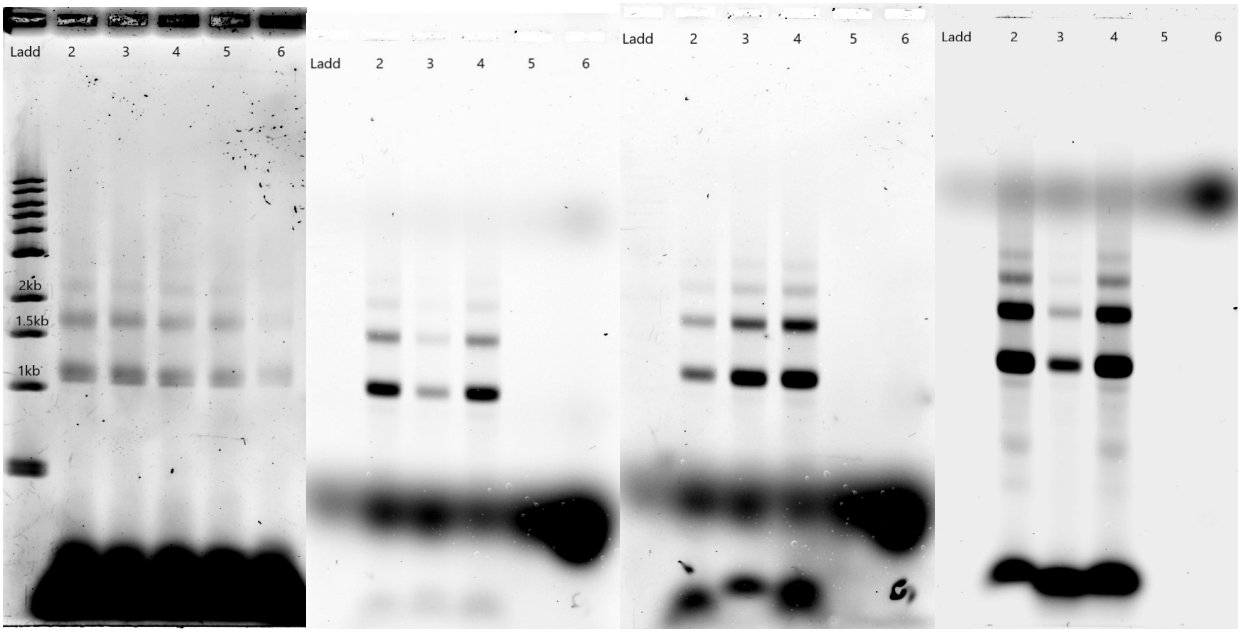


Figure 7.13: FRET staples location on R1 origami marked with red circles. The marked staples are modified to carry Cy3 and Cy5 fluorophores, when the origami is correctly folded they will interact with a FRET effect. 13-3 top-right circle; 26-25 centre-left circle.

These staples were tested directly in the origami folding. As the first experiments showed a quite strong background due to the ten-fold excess of staples over the scaffold (data not showed), it was decided to use an isomolar concentration of fluorescent staples and scaffold.

Th DB R1 FRET origami folding was performed in a one-pot reaction containing: TAE 1X, $MgCl_2$ [12,5 mM], R1 origami staples mix (without fluorescent staples) [200 nM], DB 981 scaffold [20 nM], 13-3 staples [20 nM] and 26-25 staples [20 nM]. The isothermal folding was performed at 55°C for 15 minutes and the sample was held at 4°C.

The folded origami were analysed through AGE 1% W/V in TBE. The results show a clear FRET effect from staples 13-3 and a weak one from staples 26-25 when exciting the sample for Cy3 and detecting Cy5 fluorescence, meaning that the staples are in FRET range when the origami is correctly folded. Staple 25 Cy-5 does not seem to anneal the scaffold with a very high efficiency, reducing the origami fluorescence and increasing the background signal. The origami carrying all the fluorescent staples has a signal output compatible with the sum of the two other samples 7.14.



(a) (b) (c) (d)

Figure 7.14: AGE of: lane 1 - 1 Kb DNA ladder (NEB); lane 2 - R1 with 13-3 FRET; lane 3 - R1 with 26-25 FRET; lane 4 - R1 with 13-3-26-25 FRET; lane 5 - R1 origami; lane 6 - RT origami mix. 7.14b - SYBR Gold staining. 7.14b - Cy-3/Cy-5 FRET. 7.14c Cy-3 filter. 7.14d Cy-5 filter.

To further confirm these findings I performed a plate reader analysis comparing the different staple combinations, unfortunately, also in this case, the origami optimization allowed the folding of the control samples. Nonetheless this time a sensible difference in fluorescence was measured [Tab. 7.1].

	13-3 FRET	26-25 FRET	All FRET	Ctrl
Fold	45797 (dev 3961)	35200 (dev 3344)	81415 (dev 3774)	4636
Unfold	28647 (dev 5129)	24804 (dev 7650)	42733 (dev 10969)	

Table 7.1: Plate reader experiment on R1 origami using FRET staples. In brackets the standard deviation. Ctrl refers to R1 origami folded without fluorescent staples.

7.2 Vesicles synthesis

7.2.1 Preliminary experiments

Artificial giant unilamellar vesicles (GUV) are phospholipidic liposomes that simulate a double layer cell membrane. I decided to use these vesicles as a carrier for the origami folding experiment. They allow replicating some of the cellular conditions and are used in origin of life studies. The protocol I use to synthesize them is derived from the inverted emulsion technique developed by Pautot et al. in 2003 [97]. The protocol as been slightly modified and adapted by Prof. Hanczyc from the University of Trento.
Using this protocol I managed to synthesize GUVs loaded with GFP for their visualization [Fig. 7.15]. I also generated vesicles loaded with a folded origami solution.

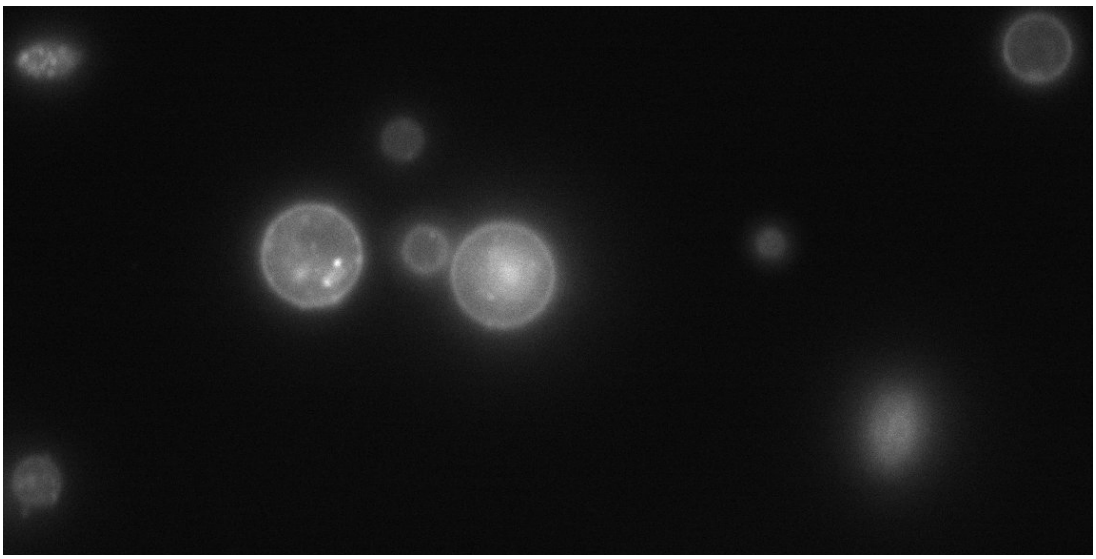


Figure 7.15: Giant unilamellar vesicles preliminary test, stained with nile red and visualized using a Nikon Eclipse Ti inverted microscope.

7.2.2 Giant unilamellar vesicles synthesis

In my experiments I synthesized GUVs using three different kinds of phospholipids: 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC, Avanti polar lipids) alone or in combination with the negatively charged 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (PEG2000 PE, Avanti polar lipids) or the positively charged 1,2-dioleoyl-3-trimethylammonium-propane (chloride salt) (DOTAP, Avanti polar lipids).

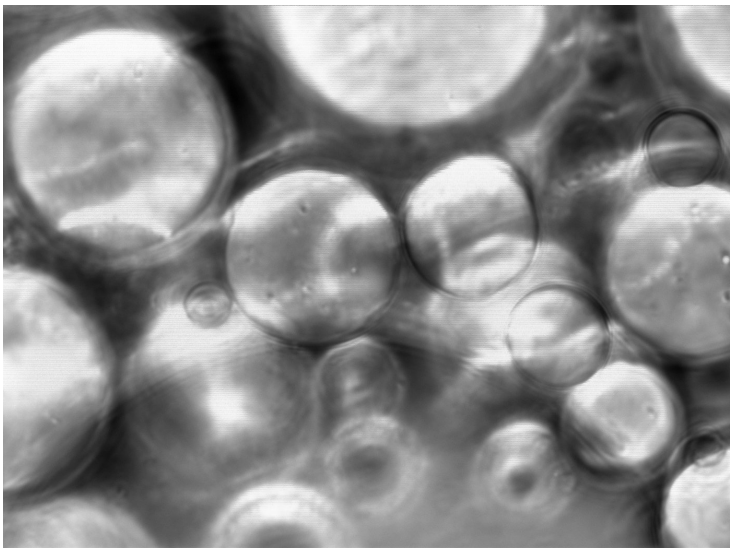
The first vesicles synthesis experiment, to familiarize with the technique, followed the general protocol described in section 3.8. The vesicles were produced with a solution of pure POPC. The formation was confirmed by microscopy.

The droplet solution (DS) was prepared with sucrose [990 mM] and NaI [10 mM], the hosting solution (HS) with glucose [495 mM], sucrose [495 mM] and NaI [10 mM]. POPC was prepared dissolving the powder in chloroform to a final concentration of 5 mg/ml and evaporating the chloroform leaving the tube open in the fumehood. POPC was then dissolved in mineral oil to a final concentration 200 M. 100 μl of HS were added in a microplate well.Slowly 50 μl of POPC solution was added on top of the HS to form a layer, waiting for 30 minutes until the PS formed a flat interface with HS. In a new Eppendorf the droplet emulsion was prepared adding 20 μl of DS in 1000 μl of POPC and passing the tube on a tube rack surface 10 times to mechanically agitate the content. 100 μl of emulsion were added on top of the HS-PS interface incubating the sample 5 minutes to let the emulsion form a second flat interface.The vesicle reaction mix was centrifuged at 1500 RCF for 3 minutes at 4°C. The vesicles formation was checked on an inverted microscope just on visible light to confirm the formation of the vesicles. The top layer of the PS solution was removed using a vacuum pump. 150 μl of HS were slowly added on top of the vesicles and centrifuged again to wash the sample removing again the top layer of the solution.

A drop of 10 μl of vesicles solution was removed from the microplate well and put in an observation chamber, in some of the samples a droplet of Nile Red was added on top of the sample as phospholipidc bilayer staining dye. The observation chamber is formed by a silanized microscope slide and a siliconized glass circle coverslide (Hampton research) spaced by 1 mm handmade tape support and glued with transparent polish nail. The sealed chamber prevents the evaporation of the vesicles droplet, allowing to reverse the sample on the inverted microscope.
The vesicles were visualized using a Nikon Eclipse Ti inverted microscope. Brightfield, GFP and Nile Red fluorescence confirmed the formation of the giant unilamellar vesicles [Fig. 7.15, 7.16].

Another experiment was performed to test the vesicles synthesis in Eppendorf tubes and with differently charged phospholipid mixes:

- POPC



(a)

(b)

Figure 7.16: Inverted microscope imaging of GUVs. 7.16a - Brightfield imaging of a high concentration spot to notice the high yield of GUVs synthesis. 7.16b - GFP imaging of GUVs of different sizes.

- POPC + PEG2000 PE (25:75 and 50:50 ratios)
- POPC + DOTAP (25:75, 50:50, 60:40, 40:60 and 75:25 ratios)

While the standard POPC behaved as expected, the most interesting results were obtained using the negatively charged POPC-PEG2000 PE mixes, capable of forming more resistant vesicles and higher yields, a side effect is that these vesicles fissate in smaller liposomes over time and cannot fuse with positively charged vesicles as the PEG2000 chains prevent the interactions. The POPC-PEG2000 vesicles were also generated using

an R1 origami solution as DS.
The vesicles containing DOTAP did not form when loaded with DNA, maybe because of the interaction between the positively charged lipids and the negatively charged nucleic acid was not favorable to the formation of lipoplexes [112].

7.3 Summary

7.3.1 Origami-folding reporter

The preliminary experiments on the origami-folding analysis tool highlighted a series of critical issues that can be used to better understand the nature of nucleic acid origami and how to apply this powerful technology.
The results described in section 7.1.4 show that R1 origami starts to fold at very low temperatures. Even if it is a non-complete process, it is sufficient to partially anneal the staples on the scaffold, activating the FRET effect between the fluorophores. This is clearly due to the scaffold and staples optimization of the RNA/DNA hybrid designs: the single components ca not form strong secondary structures providing any thermal barrier to the folding process.
The very same effect is also the cause of the split Broccoli aptamer regeneration described in section 7.1.1: the presence of the aptamer hairpin stem facilitates the formation of staple polymers, even in the absence of the scaffold. It is the opposite result of the experiment described by Afonin et al. [66], where the assembly folding regenerates the aptamer. Almost paradoxically, the over-optimization of a parameter (the folding) becomes sub-optimal in the optic of a new goal.
On the other end the Malachite green split aptamer is not very stable in the origami buffer and its fluorescence is extremely weak, for this reason it was discarded from the options. Moreover, the staples carrying the aptamer splits are RNA oligonucleotides more prone to degradation and more expensive to synthesize.
The project based on the FRET effect is the more promising for future applications. It works with DNA staples and it does not add any extra nucleic acid sequence. The limit I applied to the design, 5 nt distance between the fluorophores, can be loosened increasing the possibilities for multiple signals on a single origami. Based on the experimental results I also decided to discard the 26-25 FRET couple, keeping the 13-3 FRET staples

as reporter for the fluorescence experiments.
A nucleic acid origami is a highly technological machine, it requires a fine tuning calibrated on the task to perform, but at the same time its design can be optimized just by changing the set of staples, keeping the scaffold, the shape and any other characteristic unchanged.

7.3.2 Giant unilamellar vesicles

The experiments performed on the GUV synthesis, even if at a primordial state, do not show any concerning aspect. The vesicles formation allows the insertion of the origami, and few more experiments could confirm the feasibility of the folding analysis. More advanced experiments could involve the study of the origami folding inside a proto-cell at different conditions, such as folding buffers, cell-free extracts and raw bacterial extracts.

Part III

Chapter 8

Conclusions and discussion

The objective of my research project was to characterise the application of synthetic nucleic acid sequences as origami scaffolds. The hypothesis underlying this idea is that the *in vivo* synthesis of nucleic acid nanorobots will dramatically increase their biotechnological value, transforming these niche-technologies into powerful reprogrammable (and possibly reprogramming) biological machines. To better explain this statement it is necessary to define some of the features characterizing nucleic acid origami.

The definition of nanorobots, as nanometric automata performing a work, has been widely demonstrated by many applications. Origami can be used for their structural properties, like boxes, pores, struts or gears; they can be functionalized and used as nanoarrays, cheaper than classic microarrays; origami can also be programmed to respond to a stimulus, taking advantage of strand displacement techniques or the origami folding itself.

Origami are reprogrammable because a single scaffold can be used to fold different structures using different sets of staples, these features makes the scaffold reusable, it also opens up to possible applications where the origami structure changes depending on the environment.

Finally, the most biotechnological aspect of all, the origami can be "reprogramming". The activity of a nucleic acid origami inside a cell could change the behavior of its host, simpler assemblies have been introduced *in vivo* with good results, an origami can only expand these capabilities providing a control system expressed and folded by the cell itself.

The scaffold sequences here designed are based on combinatorial rules described by a Dutch mathematician, Nicolaas Govert de Bruijn. A nucleic acid De Bruijn (DB) se-

quence is formed by nucleotide subsequences of predetermined length that occur only once even for different reading frames, this means that every domain composing a DB origami scaffold can be unique by design.
Being arbitrarily designed, a synthetic sequence of this type can present many advantages over a natural one, in our case we wanted to obtain three main features:

- unique addressability
- bio-orthogonality
- origami folding optimization

The unique addressability is the direct effect of the DB algorithm, setting the subsequences length at 6 nt, every staple domain (8 or 16 nt) has at least two mismatching nucleotides if binding the wrong staple: in the case of competitive annealing, during the strand displacement phase, the correct staple will always have at least two stable nucleotides.
The bio-orthogonality, optimized in *E. coli*, was obtained cutting away all the biological sequences recorded in the most common genetic databases, avoiding the presence of positive and negative transcription factors and recurrent restriction enzymes.
The origami folding optimization was performed analysing the minimum free energy of the sequences obtained and choosing the one with the lower result. The scaffold obtained presents weaker secondary structures that could hinder the correct staple annealing even at lower temperatures. The unique addressability of the scaffold also contributes to the folding optimization, always favoring the annealing of the correct staples to the proper scaffold domain.
Two DB sequences were used as DNA and RNA origami scaffolds and some aspects of their bio-orthogonality were characterized inside *E. coli*.
Finally, I designed an application consisting of a functionalized origami designed to work inside phospholipidic giant unilamellar vesicles.

8.1 De Bruijn scaffold validation

8.1.1 DB 2.4 DNA origami

The DNA sequence designed using the De Bruijn algorithm folded the origami with a good yield, comparable, if not superior, to pUC19 control. Nonetheless, this scaffold presented a major challenge during the DNA single strand synthesis steps.

Two different protocols were tested to obtain an amount of scaffold sufficient to fold an origami batch.

The first protocol involving the digestion of the phosphorylated complementary strand proved completely inefficient. Even if a small part of the starting product was transformed in ssDNA, it was not sufficient to retrieve it after the necessary purification steps. Nor increasing the enzyme concentration or incubation time made a difference. The causes of this failure are not clear, but it is probably due to an imperfect phosphorylation of the amplification primers that caused only a partial digestion of the complementary strand, resulting in a very low yield of DNA scaffold.

The second protocol was instead based on the biotinylation of the complementary strand and its capture through streptavidin-coated magnetic beads. This protocol granted the synthesis of the necessary amount of ssDNA scaffold but still required micrograms of starting dsDNA. In terms of efficiency, although sufficient for the required experiments, this protocol is clearly not compatible with a future scale-up. It must be noted that, in the last years, new ssDNA synthesis protocols have been created, in particular the use of rolling circle amplification that allowed the *in vitro* production of grams of origami scaffold.

The origami folding procedure performed well, granting a good product yield, especially considering that it was the first time a DB origami was folded using an isothermal protocol. This experiment was also important as a testbed for AFM microscopy techniques, being DNA more stable than RNA, it allowed performing more extensive tests for longer times without degrading.

8.1.2 DB RNA/DNA hybrid origami

The study of RNA scaffold/DNA staples hybrid origami explored further after the first folding, generating a greater feedback in terms of DB sequences insight and origami fold-

ing. The issues tackled during this process involved the design, folding and imaging of the origami.

DB triangle The research on DB triangle hybrid origami mainly focused on the optimization of basic steps such as synthesis, folding and imaging processes.
The T7 RNA polymerase transcription solved the problem presented by the ssDNA scaffold, a single T7 reaction produces an amount of RNA scaffold in the order of micrograms and the required purification can be reduced to a single spin-column step, this method is clearly compatible with the hypothetical scaling-up required for industrial applicability.
On the other hand, working with RNA requires extra care to avoid its degradation. Ribonucleases are ubiquitous and the slightest contamination could degrade the sample, especially when working at folding temperatures (i.e. 40-60°C). Performing the experiments under a PCR cabinet solved the problem during the synthesis phase, but the lack of an RNase-free microscopy facility translated in an increased number of experiments to set-up the visualization protocols and reduced the final quality of the results. The only way to mitigate this problem consisted of the accurate cleaning of the instruments when possible. These precautions allowed the recording of AFM and TEM images. The microscopy analysis showed a better performance of AFM over TEM, electing the former as the main visualization technique.
Different folding protocols were tested for this origami: an attempt of cotranscriptional folding resulted in the absence of scaffold, probably related to the excess of ions in solution that blocked the T7 RNA polymerase, this path was not investigated further to focus on more rapid solutions; the classic folding consisting in a descending thermal ramp, derived from the original origami paper, performed sufficiently well allowing the origami formation. The DB triangle origami predates all the other origami experiments therefore no isothermal folding was tested on this design.
The final results on the DB triangle origami highlighted a flaw in the design that causes a misfolding of the vertex between the 5' and 3' scaffold ends: with the two locking staples not being stable enough, probably for the large distance between the domains, the only functional solution consisted in a different origami design that would not rely on an unpredictable locking system. A more classic rectangular design was produced to confirm the synthetic RNA scaffold functionality.

DB square and rectangle origami The square and rectangle origami designs establish a simpler raster-like scaffold disposition, in this design the staples domains are always close enough to avoid the problem presented by the DB triangle design.
The DB origami rectangle (R1) immediately showed an unexpected behavior when performing the classic folding protocol, forming a large misfolded agglomerate during the denaturation step. This issue brought to a deep investigation of the folding process that resulted in an improved isothermal protocol, adopted for all the later designs, that does not require a denaturation step. The RNA scaffold was also optimized removing the *lac* operon at the 5' and the T7 terminator at the 3', these sequences were initially included prospecting the *in vivo* application, providing only a further degree of complexity to the *in vitro* stage. The new sequence (DB 981) became the standard RNA scaffold for the hybrid origami. These solutions granted an excellent folding rate demonstrated by the AFM imaging. The tweaking of scaffold and staples ratio, brought up from 1:10 to 1:50, further improved the results at the expenses of resource economy.
The DB square was designed to test an origami super-structure, relying on the use of linker staples binding the origami in a chessboard assembly. While the gel results showed an encouraging pattern, the AFM imaging could not visualize a proper formation, it showed instead the deposition of large scale crystals of unknown origin. Due to the good results obtained with the R1 origami, it was chosen to branch out in that direction, postponing the super-structure enquiry until the folding mechanics will be properly investigated.

8.1.3 De Bruijn scaffold validation conclusions

The experiments on the origami folding clearly demonstrate that a DB sequence can be used as a scaffold. Moreover, it's clear that the thermodynamical optimization and unique addressability, exclusive features of an arbitrary synthetic sequence, can be a great advantage for working at *in vivo* compatible, lower temperatures: the origami folded using the synthetic scaffolds were capable of an isothermal folding without the necessity of an initial denaturation step as the biological sequences used before. With a proper sequence design it will be possible to perform all the folding steps at temperatures compatible with cellular cultures.

8.2 *In vivo* DB scaffold study

The last feature that needed confirmation was the bio-orthogonality, theoretically obtained by removing all the known biological sequences from the DB scaffolds.
Investigating the absence of general biological activities and interactions it is clearly not a self-contained procedure. It is possible, though, to set-up an experiment that covers some of the most important aspects for some of the species of interest. In this case the cloning of the DB sequences, as 5' untranslated regions between a promoter and a sfGFP reporter gene, showed that bio-orthogonality is a complex matter that cannot be fully predicted by the actual knowledge. The experiments were designed and performed on *E. coli*

8.2.1 pBAD33 construct

The first version of this experiment involved the use of a pBAD33 expression plasmid and different *E. coli* strains (*E. coli* 5α, MG1655 and BL21 (DE3)). Relying on *E. coli* native transcription machinery, the pBAD33-DB 2.4-sfGFP construct and its control (a random *dnaE* fragment of the same length) clearly showed no expression at all when analysed through flow cytometry. Probably, the extension of the non-coding sequence after the promoter reduced the RNA polymerase processivity, suppressing the sfGFP expression, at the same time it was noted that the DB 2.4 sequence had no transcription activating effects.

8.2.2 pET-28a constructs

A functional transcription-translation system was obtained by cloning the DB-sfGFP constructs in a pET-28a plasmid. This plasmid relies on a T7 RNA polymerase for its gene transcription and the DB sequences already contained the T7 promoter sequences by design for the *in vitro* synthesis. For these experiments, an *E. coli* T7 express strain was used.
Using pET-28a made immediately clear that the sfGFP expression was restored in both DB sequence and the control sequence constructs. Unexpectedly the first results showed also that DB 2.4 has a negative effect on the associated sfGFP expression compared to *dnaE* and sfGFP alone. Identifying the cause of the fluorescence reduction is not a trivial matter, the lack of similar recorded sequences makes any bioinformatic analysis

useless, and the absence of strong secondary structures, avoided by design, suggests that it does not depend on RNA polymerase dissociation by structural impairment. This analysis leaves a question mark on the true bio-orthogonality of DB 2.4 sequence and only analysing the single fragments composing it could shed a light on this unpredicted effect.

The same analysis performed on DB 981 produced a completely different result, in this case the DB sequence performed better than the corresponding *dnaE* fragment and slightly worst than sfGFP alone, pointing towards the initial hypothesis of bio-compatibility and bio-orthogonality.

8.2.3 *In vivo* DB scaffold study conclusions

The confirmation of bio-orthogonality provided a mixed set of results. Further investigations are necessary to understand the behavior expressed by DB 2.4, while DB 981 performances do not arise reasons for concern.

Though, the definitive proof of bio-orthogonality can only be provided through the extensive use of DB sequences in *in vivo* applications, over a larger number of organisms. At the moment only this kind of proof-of-concept experiments are viable, with their clear analytical limitations.

8.3 Origami folding thermodynamics analysis tool

Designing an origami application, that could help to better understand the thermodynamics of the origami folding and at the same time could pave the way towards its *in vivo* synthesis, seems a natural step forward, covering all the topics presented in my thesis. Combining a functionalized origami, that generates a fluorescent signal upon its folding, with the giant unilamellar vesicles synthesis, it is possible to create a cell-like system that allows the tracking of the origami folding using advanced instruments such as flow cytometers. At the same time, they can be used to simulate a simplified cellular environment in preparation for the final *in vivo* synthesis.

8.3.1 DB origami reporter

Three alternative solutions were tested for the generation of the DB origami folding reporter prototype. All of them are based on the DB R1 origami design, being the most efficient in terms of synthesis and folding.
The first two designs involve the use of staples functionalized with fluorescent split RNA aptamers that regenerate themselves upon the correct origami folding. The last one takes advantage of the FRET effect between Cy3 and Cy5 modified staples.

Broccoli aptamer The origami reporter using the Broccoli aptamer as fluorescence generator performed very well in regeneration tests showing a powerful signal. The DFHBI-1T molecule interacts only with the regenerated split aptamer and can permeate cellular membranes for *in vivo* experiments, furthermore Broccoli aptamer is compatible with the origami folding buffer. The molecule is potentially perfect for this role but it resulted incompatible with the R1 origami design, and more specifically, with the DB scaffolds thermodynamical optimizations: the aptamer structure includes a hairpin stem with a minimal free energy higher than any of the staple secondary structures, when in the folding mix, the annealing of the two aptamer halves happens spontaneously generating the signal without the actual origami folding.
A software analysis of the free energy of all the possible staple-aptamer combinations individuated two couples compatible with a shorter hairpin stem, 17-30 and 24-20. The shorter hairpin demonstrated to be sufficient to avoid the spontaneous regeneration and the two aptamers were tested on the origami folding, the folding results showed that only 17-30 split regenerates on the origami. The last couple left was finally tested in a pilot experiment: the fluorescence signals of the folded reporter origami and its unfolded version were measured using a plate reader, the unexpected results showed no difference between the samples. The reason is to be found again in the optimization of the DB sequence, it promotes a partial folding of the origami also at sub-optimal temperatures allowing the annealing of the aptamer-modified 17-30 staples and the consequent fluorescence.

Malachite green aptamer The second reporter relies on another split aptamer regeneration, the malachite green. The free energy simulation was performed in advance identifying three possible staple couples compatible with the split aptamer modification:

again 17-30 and 24-20, as for Broccoli, plus 4-6. Having a much more unstable sequence this aptamer did not regenerate spontaneously but the regeneration tests performed in agarose gel showed a good fluorescence/background ratio for 24-20 and 4-6 couples and no signal for 17-30. When tested in the origami using the folding buffer, with a lower ion concentration, all three couples worked again but the signal to background ratio became much worst and almost undetectable. This critical flaw obviously limits the applicability of the system, precluding the objective for which it is designed. Nonetheless, another plate reader experiment was performed, and again the unfolded control origami showed a similar signal to the folded one. To confirm the finding all the plate reader samples were loaded in an agarose gel electrophoresis, showing indeed the formation of a partially folded origami at room temperature

FRET The last DB origami folding reporter uses a completely different signal system: two couples of staples, 13-3 and 26-25, are functionalized with fluorescent molecules, each couple has one Cy-3 and one Cy-5 modification. When the origami folds the fluorescent molecules lie 5 nt away from each other and the Cy-3 emission activates Cy-5 fluorescence through the FRET effect. Because this system does not require any extra oligonucleotide modification the risk of a false positive is reduced and the buffer composition does not affect the efficiency of the reporters. Due to the intrinsic staple design, the two couples could only be tested in the full origami folding and not as single couples. The first attempt showed that a ratio of 10 fluorescent staples per 1 scaffold generates a disrupting background signal. Using a 1:1 ratio improved the signal to noise ratio by a greater extent but also showed a structural instability of the staple 25 Cy-5 producing a partial false negative result requiring the removal of the 26-25 couple. Also in this case, the final working design contemplated a single efficient signaling couple.

The plate reader experiment was performed also for this final reporter. Depending on the DB scaffold design, this system suffered again from the partial folding of the control origami. Still, the folded origami FRET signal remained higher than the partially folded control, probably because these fluorescent staples do not anneal efficiently at room temperature.

8.3.2 Giant unilamellar vesicles

The other component of this experiment is the giant unilamellar vesicle, an artificial liposome that can be loaded with a variety of different cargo and behaves as a simple proto-cell. I only had the chance to perform a few preliminary tests in order to set-up the first synthesis protocols.

In the first experiment, a neutrally charged POPC vesicle loaded with a GFP solution was generated and visualized under a fluorescence microscope. As expected the GUV showed an excellent synthesis yield, a good fluorescence signal and a prolonged stability.

In the second attempt, a mix of POPC and the negatively charged PEG2000 PE phospholipid was used to generate a set of vesicles loaded with a solution of folded DB R1 non-fluorescent origami. The PEG2000 PE addition further increased the vesicles efficiency, unfortunately, the unavailability of a fluorescent origami signal prevented the origami identification; it must be noted, though, that the synthesis protocol starts with the formation of the cargo solution droplets, making unavoidable the origami presence inside the final product.

The third and last version used a mix of POPC and DOTAP, a positively charged lipid. This time, while the empty vesicles formed, the origami loaded ones could not be synthesized. It is possible that the interaction between the positive lipid and the negative nucleic acids unbalanced the conditions for the vesicles formation.

8.3.3 Origami folding thermodynamics analysis tool conclusions

Being in the embryonic stage, few practical conclusions can be drawn from this set of experiments. It is clear that the DB R1 origami reporter is not the optimal structure to measure the folding process, the extreme folding optimization does not provide a controlled process, making the measurements very difficult, if not inaccurate. This demonstrates that the origami designs must be tailored for a specific goal. Under this point of view a DB synthetic sequence has a great advantage: the algorithm generates a large number of different sequences with different characteristics and the user can choose the one that suits its needs. On the contrary, a biological sequence is rather static and requires a great amount of modifications to radically change its behavior.

Both the aptamers and the FRET system showed a huge potential, coupled to a dedicated origami design they can easily generate more reliable reporters.

The giant unilamellar vesicles are a promising vessel for the origami analysis, being widely used in the last 40 years they are a flexible but reliable tool that works well with nucleic acids. The formation of vesicles loaded with origami is just the first step, but a relevant one, that supports the worth of this application.

8.4 Future work

Many are the research paths that could stem from my experiments, some of the results provided the inspiration for new potential applications and some of the questions still deserve a deeper investigation.

ssDNA synthesis The ssDNA DB scaffold synthesis could be tested on a rolling circle amplification *in vivo*, protocols for the generation of scaffolds and staples were already published but none of them has the capacity to fold at life-compatible temperatures and requiring an extraction and purification to be used *in vitro*. On the other hand, a DB scaffold could be designed for this purpose attempting the first *in vivo* folding. A DNA scaffold also has the advantage of being more stable and less susceptible to nucleases, while RNA will be probably limited to mutated bacterial strains with a reduced RNase activity.

Bio-orthogonality The DB scaffold still requires a better characterization of its bio-orthogonality. DB 2.4 DNA scaffold can be fragmented and analysed to detect which subsequence hinder the transcription of its construct. Also, more DB sequences can be tested to see if any pattern arises and which is the statistical incidence of interfering products. Obviously, the bio-orthogonality of a sequence is not only connected to the transcription machinery and, as said before, it will be necessary a wider application of *in vivo* DB sequences to have a complete picture.

DB origami folding reporter The folding reporter issues highlighted the necessity for a more focused development of applicable origami. A more reliable origami that folds at a predetermined temperature, not only solves this specific research problem, but can be also considered a proper molecular switch. Combining a temperature responsive origami with the split aptamer technology will make possible the creation of an artificial controllable

enzyme based on nucleic acids.

Origami folding analysis tool The experiments on the folding analysis tool are the most natural continuation of this research, as previously outlined there are two main possible paths involving GUVs and origami.
The first is the creation of a system that allows understanding how the origami folding happens. Different fluorescent reporters on the same origami could describe the temporal sequence of the staples annealing during the folding, the results can be used to improve or confirm the computational models. In this case, the vesicles could accomplish two purposes, they can allow the use of high throughput machines to analyse the origami, but they can also control the origami folding reaction: different GUVs carrying the separate origami components can be fused together in a controlled manner starting the folding only when required.
The second path considers the GUVs as proto-cells and uses them to gradually get closer to the *in vivo* origami synthesis. Instead of trying the folding directly inside a living cell, with many uncontrolled variables, smaller steps are performed inside the liposomes. Using buffers, cell-free extracts and raw extracts it could be possible to consider the potential variables individually or at least in smaller groups. Nucleic acid origami are still a young technology and their true potential is not yet expressed, the next big leap will coincide with their synthesis *in vivo* and the repercussions will involve all the biology fields, from the industrial biotechnology to the molecular medicine.

Bibliography

[1] Rothemund Paul W. K. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, mar 2006. 10.1038/nature04586.

[2] Jerzy Kozyra, Alessandro Ceccarelli, Emanuela Torelli, Annunziata Lopiccolo, Jing-Ying Gu, Harold Fellermann, Ulrich Stimming, and Natalio Krasnogor. Designing uniquely addressable bio-orthogonal synthetic scaffolds for DNA and RNA origami. *ACS Synthetic Biology*, 0(0):null, 0. PMID: 28414914.

[3] Elisabeth Pound, Jeffrey R Ashton, Héctor A Becerril, and Adam T Woolley. Polymerase chain reaction based scaffold preparation for the production of thin, branched DNA origami nanostructures of arbitrary sizes. *Nano Lett.*, 9(12):4302–5, December 2009.

[4] Honglu Zhang, Jie Chao, Dun Pan, Huajie Liu, Qing Huang, and Chunhai Fan. Folding super-sized DNA origami with scaffold strands from long-range pcr. *Chem. Commun.*, 48:6405–6407, 2012.

[5] Kozyra Jerzy Wieslaw. Computation and programmability at the nano-bio interface. *Newcastle University, Ph.D. Thesis*, 2017.

[6] GM Whitesides, JP Mathias, and CT Seto. Molecular self-assembly and nanochemistry: a chemical strategy for the synthesis of nanostructures. *Science*, 254(5036):1312–1319, 1991.

[7] Seeman Nadrian C. and Sleiman Hanadi F. DNA nanotechnology. *Nature Reviews Materials*, 3:17068, nov 2017.

[8] Li Jiang, Green Alexander A., Yan Hao, and Fan Chunhai. Engineering nucleic acid structures for programmable molecular circuitry and intracellular biocomputation. *Nature Chemistry*, 9:1056, sep 2017.

[9] Nadrian C. Seeman. Nucleic acid junctions and lattices. *Journal of Theoretical Biology*, 99(2):237–247, 1982.

[10] Kallenbach Neville R., Ma Rong-Ine, and Seeman Nadrian C. An immobile nucleic acid junction constructed from oligonucleotides. *Nature*, 305:829, oct 1983.

[11] Tsu Ju Fu and Nadrian C. Seeman. DNA double-crossover molecules. *Biochemistry*, 32(13):3211–3220, 1993. PMID: 8461289.

[12] Xiaojun Li, Xiaoping Yang, Jing Qi, and Nadrian C. Seeman. Antiparallel DNA double crossover molecules as components for nanoconstruction. *Journal of the American Chemical Society*, 118(26):6131–6140, 1996.

[13] Nadrian C. Seeman. DNA NANOTECHNOLOGY: Novel DNA Constructions. *Annual Review of Biophysics and Biomolecular Structure*, 27(1):225–248, 1998. PMID: 9646868.

[14] Winfree Erik, Liu Furong, Wenzler Lisa A., and Seeman Nadrian C. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394(6693):539–544, aug 1998. 10.1038/28998.

[15] Hao Yan, Thomas H. LaBean, Liping Feng, and John H. Reif. Directed nucleation assembly of DNA tile complexes for barcode-patterned lattices. *Proceedings of the National Academy of Sciences*, 100(14):8103–8108, 2003.

[16] Shih William M., Quispe Joel D., and Joyce Gerald F. A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature*, 427:618, feb 2004.

[17] Sung Ha Park, Constantin Pistol, Sang Jung Ahn, John H. Reif, Alvin R. Lebeck, Chris Dwyer, and Thomas H. LaBean. Finite-Size, Fully Addressable DNA Tile Lattices Formed by Hierarchical Assembly Procedures. *Angewandte Chemie International Edition*, 45(5):735–739, 2006.

[18] Wei Bryan, Dai Mingjie, and Yin Peng. Complex shapes self-assembled from single-stranded DNA tiles. *Nature*, 485:623, may 2012.

[19] Hamblin Graham D., Rahbani Janane F., and Sleiman Hanadi F. Sequential growth of long DNA strands with user-defined patterns for nanostructures and scaffolds. *Nature Communications*, 6:7065, may 2015.

[20] Kai Lin Lau and Hanadi F. Sleiman. Minimalist approach to complexity: Templating the assembly of DNA tile structures with sequentially grown input strands. *ACS Nano*, 10(7):6542–6551, 2016. PMID: 27303951.

[21] Ong Luvena L., Hanikel Nikita, Yaghi Omar K., Grun Casey, Strauss Maximilian T., Bron Patrick, Lai-Kee-Him Josephine, Schueder Florian, Wang Bei, Wang Pengfei, Kishi Jocelyn Y., Myhrvold Cameron, Zhu Allen, Jungmann Ralf, Bellot Gaetan, Ke Yonggang, and Yin Peng. Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components. *Nature*, 552:72, dec 2017.

[22] Wei Bryan, Ong Luvena L., Chen Jeffrey, Jaffe Alexander S., and Yin Peng. Complex reconfiguration of DNA nanostructures. *Angewandte Chemie*, 126(29):7605–7609, 2014.

[23] Douglas Shawn M., Dietz Hendrik, Liedl Tim, Hogberg Bjorn, Graf Franziska, and Shih William M. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature*, 459(7245):414–418, may 2009. 10.1038/nature08016.

[24] Shawn M. Douglas, Adam H. Marblestone, Surat Teerapittayanon, Alejandro Vazquez, George M. Church, and William M. Shih. Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Research*, 37(15):5001–5006, 2009.

[25] Castro Carlos Ernesto, Kilchherr Fabian, Kim Do-Nyun, Shiao Enrique Lin, Wauer Tobias, Wortmann Philipp, Bathe Mark, and Dietz Hendrik. A primer to scaffolded DNA origami. *Nat Meth*, 8(3):221–229, mar 2011. 10.1038/nmeth.1570.

[26] Ke Yonggang, Sharma Jaswinder, Liu Minghui, Jahn Kasper, Liu Yan, and Yan Hao. Scaffolded DNA Origami of a DNA Tetrahedron Molecular Container. *Nano Letters*, 9(6):24452447, 2009. doi: 10.1021/nl901165f.

[27] Endo Masayuki, Hidaka Kumi, Kato Takayuki, Namba Keiichi, and Sugiyama Hiroshi. DNA Prism Structures Constructed by Folding of Multiple Rectangular Arms. *Journal of the American Chemical Society*, 131(43):1557015571, 2009. doi: 10.1021/ja904252e.

[28] Masayuki Endo, Seigi Yamamoto, Tomoko Emura, Kumi Hidaka, Nobuhiro Morone, John E. Heuser, and Hiroshi Sugiyama. Helical DNA origami tubular structures with various sizes and arrangements. *Angewandte Chemie International Edition*, 53(29):7484–7490, 2014.

[29] Hendrik Dietz, Shawn M. Douglas, and William M. Shih. Folding DNA into twisted and curved nanoscale shapes. *Science*, 325(5941):725–730, 2009.

[30] Benson Erik, Mohammed Abdulmelik, Gardell Johan, Masich Sergej, Czeizler Eugen, Orponen Pekka, and Hgberg Bjrn. DNA rendering of polyhedral meshes at the nanoscale. *Nature*, 523:441, jul 2015.

[31] Jungmann Ralf, Liedl Tim, Sobey Thomas L., Shih William, and Simmel Friedrich C. Isothermal Assembly of DNA Origami Structures Using Denaturing Agents. *Journal of the American Chemical Society*, 130(31):1006210063, 2008. doi: 10.1021/ja8030196.

[32] Jean-Philippe J. Sobczak, Thomas G. Martin, Thomas Gerling, and Hendrik Dietz. Rapid Folding of DNA into Nanoscale Shapes at Constant Temperature. *Science*, 338(6113):1458–1461, 2012.

[33] Andreas Kopielski, Anne Schneider, Andrea Cski, and Wolfgang Fritzsche. Isothermal DNA origami folding: avoiding denaturing conditions for one-pot, hybrid-component annealing. *Nanoscale*, 7:2102–2106, 2015.

[34] Chopra Aradhana, Krishnan Swati, and Simmel Friedrich C. Electrotransfection of Polyamine Folded DNA Origami Structures. *Nano Letters*, 16(10):66836690, 2016. doi: 10.1021/acs.nanolett.6b03586.

[35] Dunn Katherine E., Dannenberg Frits, Ouldridge Thomas E., Kwiatkowska Marta, Turberfield Andrew J., and Bath Jonathan. Guiding the folding pathway of DNA origami. *Nature*, 525:82, aug 2015.

[36] Bellot Gatan, McClintock Mark A, Lin Chenxiang, and Shih William M. Recovery of intact DNA nanostructures following agarose-gelbased separation. *Nature methods*, 8(3):10.1038/nmeth0311–192, mar 2011.

[37] Douglas Shawn M, Chou James J, and Shih William M. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proceedings of the National Academy of Sciences of the United States of America*, 104(16):66446648, dec 2006.

[38] Stahl Evi, Martin Thomas G, Praetorius Florian, and Dietz Hendrik. Facile and Scalable Preparation of Pure and Dense DNA Origami Solutions. *Angewandte Chemie (International Ed. in English)*, 53(47):1273512740, aug 2014.

[39] Lin Chenxiang, Perrault Steven D, Kwak Minseok, Graf Franziska, and Shih William M. Purification of DNA-origami nanostructures by rate-zonal centrifugation. *Nucleic Acids Research*, 41(2):e40e40, oct 2012.

[40] Wickham Shelley F. J., Endo Masayuki, Katsuda Yousuke, Hidaka Kumi, Bath Jonathan, Sugiyama Hiroshi, and Turberfield Andrew J. Direct observation of stepwise movement of a synthetic molecular transporter. *Nature Nanotechnology*, 6:166, feb 2011.

[41] Shawn M Douglas, Ido Bachelet, and George M Church. A logic-gated nanorobot for targeted transport of molecular payloads. *Science*, 335(6070):831–4, February 2012.

[42] Alan Shaw, Erik Benson, and Bjrn Hgberg. Purification of functionalized DNA origami nanostructures. *ACS Nano*, 9(5):4968–4975, 2015. PMID: 25965916.

[43] Benjamin Kick, Florian Praetorius, Hendrik Dietz, and Dirk Weuster-Botz. Efficient Production of Single-Stranded Phage DNA as Scaffolds for DNA Origami. *Nano Lett.*, 15(7):4672–6, July 2015.

[44] Cosimo Ducani, Corinna Kaul, Martin Moche, William M Shih, and Björn Högberg. Enzymatic production of 'monoclonal stoichiometric' single-stranded DNA oligonucleotides. *Nat. Methods*, 10(7):647–52, July 2013.

[45] Praetorius Florian, Kick Benjamin, Behler Karl L., Honemann Maximilian N., Weuster-Botz Dirk, and Dietz Hendrik. Biotechnological mass production of DNA origami. *Nature*, 552:84, dec 2017.

[46] Marchi Alexandria N., Saaem Ishtiaq, Vogen Briana N., Brown Stanley, and LaBean Thomas H. Toward Larger DNA Origami. *Nano Letters*, 14(10):57405747, 2014. doi: 10.1021/nl502626s.

[47] Tikhomirov Grigory, Petersen Philip, and Qian Lulu. Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns. *Nature*, 552:67, dec 2017.

[48] W M Barnes. Pcr amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. *Proceedings of the National Academy of Sciences*, 91(6):2216–2220, 1994.

[49] Michael Erkelenz, Dennis M. Bauer, Rebecca Meyer, Christos Gatsogiannis, Stefan Raunser, Barbara Sacc, and Christof M. Niemeyer. A facile method for preparation of tailored scaffolds for DNAorigami. *Small*, 10(1):73–77, 2014.

[50] Hgberg Bjrn, Liedl Tim, and Shih William M. Folding DNA Origami from a Double-Stranded Source of Scaffold. *Journal of the American Chemical Society*, 131(26):91549155, 2009. doi: 10.1021/ja902569x.

[51] Yang Yang, Han Dongran, Nangreave Jeanette, Liu Yan, and Yan Hao. DNA Origami with Double-Stranded DNA As a Unified Scaffold. *ACS Nano*, 6(9):82098215, 2012. doi: 10.1021/nn302896c.

[52] Ryosuke Iinuma, Yonggang Ke, Ralf Jungmann, Thomas Schlichthaerle, Johannes B. Woehrstein, and Peng Yin. Polyhedra self-assembled from DNA tripods and characterized with 3d DNA-paint. *Science*, 344(6179):65–69, 2014.

[53] Liu Wenyan, Zhong Hong, Wang Risheng, and Seeman Nadrian C. Crystalline twodimensional DNAorigami arrays. *Angewandte Chemie International Edition*, 50(1):264–267, 2011.

[54] Risheng Wang, Kent Gorday, Colin Nuckolls, and Shalom J. Wind. Control of DNA origami inter-tile connection with vertical linkers. *Chem. Commun.*, 52:1610–1613, 2016.

[55] Thomas Gerling, Klaus F. Wagenbauer, Andrea M. Neuner, and Hendrik Dietz. Dynamic DNA devices and assemblies formed by shape-complementary, non–base pairing 3d components. *Science*, 347(6229):1446–1452, 2015.

[56] Wagenbauer Klaus F., Sigl Christian, and Dietz Hendrik. Gigadalton-scale shape-programmable DNA assemblies. *Nature*, 552:78, dec 2017.

[57] Li Wei, Yang Yang, Jiang Shuoxing, Yan Hao, and Liu Yan. Controlled Nucleation and Growth of DNA Tile Arrays within Prescribed DNA Origami Frames and Their Dynamics. *Journal of the American Chemical Society*, 136(10):37243727, 2013.

[58] Luc Jaeger and Neocles B. Leontis. Tecto-RNA: One-Dimensional Self-Assembly through Tertiary Interactions. *Angewandte Chemie International Edition*, 39(14):2521–2524, 2000.

[59] Eckart Bindewald, Robert Hayes, Yaroslava G. Yingling, Wojciech Kasprzak, and Bruce A. Shapiro. RNAJunction: a database of RNA junctions and kissing loops for three-dimensional structural analysis and nanodesign. *Nucleic Acids Research*, 36(Database-Issue):392–397, 2008.

[60] Kirill A. Afonin, Dennis J. Cieply, and Neocles B. Leontis. Specific RNA Self-Assembly with Minimal Paranemic Motifs. *Journal of the American Chemical Society*, 130(1):93–102, 2008. PMID: 18072767.

[61] Guo Peixuan. The emerging field of RNA nanotechnology. *Nat Nano*, 5(12):833–842, dec 2010. 10.1038/nnano.2010.231.

[62] Myhrvold Cameron and Silver Pamela A. Using synthetic RNAs as scaffolds and regulators. *Nature Structural &Amp; Molecular Biology*, 22:8, jan 2015.

[63] Arkadiusz Chworos, Isil Severcan, Alexey Y. Koyfman, Patrick Weinkam, Emin Oroudjev, Helen G. Hansma, and Luc Jaeger. Building programmable jigsaw puzzles with RNA. *Science*, 306(5704):2068–2072, 2004.

[64] Severcan Isil, Geary Cody, Verzemnieks Erik, Chworos Arkadiusz, and Jaeger Luc. Square-Shaped RNA Particles from Different RNA Folds. *Nano Letters*, 9(3):12701277, 2009. doi: 10.1021/nl900261h.

[65] Severcan Isil, Geary Cody, Chworos Arkadiusz, Voss Neil, Jacovetty Erica, and Jaeger Luc. A polyhedron made of tRNAs. *Nature chemistry*, 2(9):772779, jul 2010.

[66] Afonin Kirill A., Bindewald Eckart, Yaghoubian Alan J., Voss Neil, Jacovetty Erica, Shapiro Bruce A., and Jaeger Luc. In vitro assembly of cubic RNA-based scaffolds designed in silico. *Nature Nanotechnology*, 5:676, aug 2010.

[67] Masayuki Endo, Seigi Yamamoto, Koichi Tatsumi, Tomoko Emura, Kumi Hidaka, and Hiroshi Sugiyama. RNA-templated DNA origami structures. *Chem. Commun. (Camb.)*, 49(28):2879–81, April 2013.

[68] Pengfei Wang, Seung Hyeon Ko, Cheng Tian, Chenhui Hao, and Chengde Mao. RNA-DNA hybrid origami: folding of a long RNA single strand into complex nanostructures using short DNA helper strands. *Chem. Commun.*, 49:5462–5464, 2013.

[69] Endo Masayuki, Takeuchi Yosuke, Emura Tomoko, Hidaka Kumi, and Sugiyama Hiroshi. Preparation of chemically modified RNA origami nanostructures. *Chemistry A European Journal*, 20(47):15330–15333, 2014.

[70] Subramanian Hari K. K., Chakraborty Banani, Sha Ruojie, and Seeman Nadrian C. The Label-Free Unambiguous Detection and Symbolic Display of Single Nucleotide Polymorphisms on DNA Origami. *Nano Letters*, 11(2):910913, 2011. doi: 10.1021/nl104555t.

[71] Feiran Li, Haorong Chen, Jing Pan, Tae-Gon Cha, Igor L. Medintz, and Jong Hyun Choi. A DNAzyme-mediated logic gate for programming molecular capture and release on DNA origami. *Chem. Commun.*, 52:8369–8372, 2016.

[72] Li Suping, Jiang Qiao, Liu Shaoli, Zhang Yinlong, Tian Yanhua, Song Chen, Wang Jing, Zou Yiguo, Anderson Gregory J, Han Jing-Yan, Chang Yung, Liu Yan, Zhang Chen, Chen Liang, Zhou Guangbiao, Nie Guangjun, Yan Hao, Ding Baoquan, and Zhao Yuliang. A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo. *Nature Biotechnology*, 36:258, feb 2018.

[73] Tomohiro Takenaka, Masayuki Endo, Yuki Suzuki, Yangyang Yang, Tomoko Emura, Kumi Hidaka, Takayuki Kato, Tomoko Miyata, Keiichi Namba, and Hiroshi Sugiyama. Photoresponsive DNA nanocapsule having an open/close system for capture and release of nanomaterials. *Chemistry*, 20(46):14951–4, November 2014.

[74] Andersen Ebbe S., Dong Mingdong, Nielsen Morten M., Jahn Kasper, Subramani Ramesh, Mamdouh Wael, Golas Monika M., Sander Bjoern, Stark Holger, Oliveira Cristiano L. P., Pedersen Jan Skov, Birkedal Victoria, Besenbacher Flemming, Gothelf Kurt V., and Kjems Jrgen. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature*, 459:73, may 2009.

[75] Grossi Guido, Dalgaard Ebbesen Jepsen Mette, Kjems Jrgen, and Andersen Ebbe Sloth. Control of enzyme reactions by a reconfigurable DNA nanovault. *Nature Communications*, 8(1):992, 2017.

[76] Setyawati Magdiel I., Kutty Rajaletchumy Veloo, Tay Chor Yong, Yuan Xun, Xie Jianping, and Leong David T. Novel Theranostic DNA Nanoscaffolds for the Simultaneous Detection and Killing of Escherichia coli and Staphylococcus aureus. *ACS Applied Materials & Interfaces*, 6(24):2182221831, 2014. doi: 10.1021/am502591c.

[77] Emanuela Torelli, Monica Marini, Sabrina Palmano, Luca Piantanida, Cesare Polano, Alice Scarpellini, Marco Lazzarino, and Giuseppe Firrao. A DNA origami nanorobot controlled by nucleic acid hybridization. *Small*, 10(14):2918–2926, 2014.

[78] Robert Schreiber, Ngoc Luong, Zhiyuan Fan, Anton Kuzyk, Philipp C Nickels, Tao Zhang, David M Smith, Bernard Yurke, Wan Kuang, Alexander O Govorov, and Tim Liedl. Chiral plasmonic DNA nanostructures with switchable circular dichroism. *Nat Commun*, 4:2948, 2013.

[79] Schmied Jrgen J, Raab Mario, Forthmann Carsten, Pibiri Enrico, Wnsch Bettina, Dammeyer Thorben, and Tinnefeld Philip. DNA origamibased standards for quantitative fluorescence microscopy. *Nature Protocols*, 9:1367, may 2014.

[80] He Xiaojin, Sha Ruojie, Zhuo Rebecca, Mi Yongli, Chaikin Paul M., and Seeman Nadrian C. Exponential growth and selection in self-replicating materials from DNA origami rafts. *Nature Materials*, 16:993, sep 2017.

[81] Qian Mei, Xixi Wei, Fengyu Su, Yan Liu, Cody Youngbull, Roger Johnson, Stuart Lindsay, Hao Yan, and Deirdre Meldrum. Stability of DNA origami nanoarrays in cell lysate. *Nano Lett.*, 11(4):1477–82, April 2011.

[82] Camille J. Delebecque, Ariel B. Lindner, Pamela A. Silver, and Faisal A. Aldaye. Organization of Intracellular Reactions with Rationally Designed RNA Assemblies. *Science*, 333(6041):470–474, 2011.

[83] Sachdeva Gairik, Garg Abhishek, Godding David, Way Jeffrey C, and Silver Pamela A. In vivo co-localization of enzymes on RNA scaffolds increases metabolic production in a geometrically dependent manner. *Nucleic Acids Research*, 42(14):94939503, may 2014.

[84] Cody W. Geary and Ebbe Sloth Andersen. Design principles for single-stranded RNA origami structures. In Satoshi Murata and Satoshi Kobayashi, editors, *DNA Computing and Molecular Programming*, pages 1–19, Cham, 2014. Springer International Publishing.

[85] Cody Geary, Paul W. K. Rothemund, and Ebbe S. Andersen. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science*, 345(6198):799–804, 2014.

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

[87] Geary Cody, Chworos Arkadiusz, Verzemnieks Erik, Voss Neil R., and Jaeger Luc. Composing RNA Nanostructures from a Syntax of RNA Structural Modules. *Nano Letters*, 17(11):70957101, 2017. doi: 10.1021/acs.nanolett.7b03842.

[88] Elbaz Johann, Yin Peng, and Voigt Christopher A. Genetic encoding of DNA nanostructures and their self-assembly in living bacteria. *Nature Communications*, 7:11179, apr 2016.

[89] Mercedes Jimnez, AriaDNA Martos, Elisa J. Cabr, Ana Raso, and Germn Rivas. Giant vesicles: a powerful tool to reconstruct bacterial division assemblies in cell-like compartments. *Environmental Microbiology*, 15(12):3158–3168, 2013.

[90] Ida Louise Jørgensen, Gerdi Christine Kemmer, and Thomas Günther Pomorski. Membrane protein reconstitution into giant unilamellar vesicles: a review on current techniques. *European Biophysics Journal*, 46(2):103–119, Mar 2017.

[91] Nishimura Koji, Matsuura Tomoaki, Nishimura Kazuya, Sunami Takeshi, Suzuki Hiroaki, and Yomo Tetsuya. Cell-Free Protein Synthesis inside Giant Unilamellar Vesicles Analyzed by Flow Cytometry. *Langmuir*, 28(22):84268432, 2012. doi: 10.1021/la3001703.

[92] Hadorn Maik, Boenzli Eva, Srensen Kristian T., De Lucrezia Davide, Hanczyc Martin M., and Yomo Tetsuya. Defined DNA-Mediated Assemblies of Gene-Expressing Giant Unilamellar Vesicles. *Langmuir*, 29(49):1530915319, 2013. doi: 10.1021/la402621r.

[93] Pier Luigi Luisi, Peter Walde, and Thomas Oberholzer. Lipid vesicles as possible intermediates in the origin of life. *Current Opinion in Colloid and Interface Science*, 4(1):33 – 39, 1999.

[94] Kurihara Kensuke, Tamura Mieko, Shohda Koh-ichiroh, Toyota Taro, Suzuki Kentaro, and Sugawara Tadashi. Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated DNA. *Nature Chemistry*, 3:775, sep 2011.

[95] James Smaldon, Francisco J. Romero-Campero, Francisco Fernández Trillo, Marian Gheorghe, Cameron Alexander, and Natalio Krasnogor. A computational study of liposome logic: towards cellular computing from the bottom up. *Systems and Synthetic Biology*, 4(3):157–179, Sep 2010.

[96] John P. Reeves and Robert M. Dowben. Formation and properties of thin-walled phospholipid vesicles. *Journal of Cellular Physiology*, 73(1):49–60, 1969.

[97] Pautot Sophie, Frisken Barbara J., and Weitz D. A. Production of Unilamellar Vesicles Using an Inverted Emulsion. *Langmuir*, 19(7):28702879, 2003. doi: 10.1021/la026100v.

[98] T D Ingolia and D E Koshland. The role of calcium in fusion of artificial vesicles. *Journal of Biological Chemistry*, 253(11):3821–3829, 1978.

[99] Holopainen J M, Lehtonen J Y, and Kinnunen P K. Evidence for the extended phospholipid conformation in membrane fusion and hemifusion. *Biophysical Journal*, 76(4):21112120, apr 1999.

[100] Parente Roberta A. and Lentz Barry R. Rate and extent of polyethylene glycol-induced large vesicle fusion monitored by bilayer and internal contents mixing. *Biochemistry*, 25(21):66786688, 1986. doi: 10.1021/bi00369a053.

[101] D.P. Pantazatos and R.C. MacDonald. Directly observed membrane fusion between oppositely charged phospholipid bilayers. *The Journal of Membrane Biology*, 170(1):27–38, Jul 1999.

[102] Sunami Takeshi, Caschera Filippo, Morita Yuuki, Toyota Taro, Nishimura Kazuya, Matsuura Tomoaki, Suzuki Hiroaki, Hanczyc Martin M., and Yomo Tetsuya. Detection of Association and Fusion of Giant Vesicles Using a Fluorescence-Activated Cell Sorter. *Langmuir*, 26(19):1509815103, 2010. doi: 10.1021/la102689v.

[103] Pastr David, Pitrement Olivier, Fusil Stphane, Landousy Fabrice, Jeusset Josette, David Marie-Odile, Hamon Loc, Le Cam Eric, and Zozime Alain. Adsorption of DNA to Mica Mediated by Divalent Counterions: A Theoretical and Experimental Study. *Biophysical Journal*, 85(4):25072518, may 2003.

[104] Bruker Corporation. DNA on Mica. A Model Procedure for Nucleic Acid Binding. URL: goo.gl/nt7sZS, accessed: 20-07-2018.

[105] Mnch Richard, Hiller Karsten, Barg Heiko, Heldt Dana, Linz Simone, Wingender Edgar, and Jahn Dieter. PRODORIC: prokaryotic database of gene regulation. *Nucleic Acids Research*, 31(1):266269, sep 2002.

[106] Ronny Lorenz, Stephan H. Bernhart, Christian Höner zu Siederdissen, Hakim Tafer, Christoph Flamm, Peter F. Stadler, and Ivo L. Hofacker. ViennaRNA package 2.0. *Algorithms for Molecular Biology*, 6(1):26, Nov 2011.

[107] Thomas Litschel, Kristina A Ganzinger, Torgeir Movinkel, Michael Heymann, Tom Robinson, Hannes Mutschler, and Petra Schwille. Freeze-thaw cycles induce content exchange between cell-sized lipid vesicles. *New Journal of Physics*, 20(5):055008, 2018.

[108] Pdelacq Jean-Denis, Cabantous Stphanie, Tran Timothy, Terwilliger Thomas C, and Waldo Geoffrey S. Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology*, 24:79, dec 2005.

[109] Guzman L M, Belin D, Carson M J, and Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, 177(14):41214130, jul 1995.

[110] Khalid K. Alam, Kwaku D. Tawiah, Matthew F. Lichte, David Porciani, and Donald H. Burke. A fluorescent split aptamer for visualizing RNA-RNA assembly in vivo. *ACS synthetic biology*, 6(9):1710–1721, Sep 2017. 28548488[pmid].

[111] Henrike Niederholtmeyer, Ling Xu, and Sebastian J. Maerkl. Real-time mRNA measurement during an in vitro transcription and translation reaction using binary probes. *ACS Synthetic Biology*, 2(8):411–417, 2013. PMID: 23654250.

[112] Khatri Nirav, Baradia Dipesh, Vhora Imran, Rathi Mohan, and Misra Ambikanandan. Development and Characterization of siRNA Lipoplexes: Effect of Different Lipids, In Vitro Evaluation in Cancerous Cell Lines and In Vivo Toxicity Study. *AAPS PharmSciTech*, 15(6):16301643, aug 2014.