DNA nanopore as a signal transducer for the detection of short oligonucleotides

Luyan Yang

université de BORDEAUX

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

SPÉCIALITÉ : PHYSICO-CHIMIE DE LA MATIÈRE CONDENSÉE

Par **Luyan YANG**

UN NANOPORE D'ADN POUR LA DÉTECTION D'OLIGONUCLÉOTIDES COURTS

Sous la direction de : Juan ELEZGARAY

Soutenue le 25 Octobre 2023

Membres du jury :

M. Juan ELEZGARAY	Directeur de Recherche, CNRS	Directeur de thèse
M. Gaetan BELLOT	Chargé de recherche, INSERM	Rapporteur
M. Mathias WINTERHALTER	Professeur, Jacobs University Bremen	Rapporteur
Mme. Cécile ZAKRI	Professeur, Université de Bordeaux	Présidente
M. Alexis PREVOST	Directeur de Recherche, Sorbonne Université	Examinateur

Un nanopore d’ADN pour la détection d’oligonucléotides courts

Résumé

Les nanopores d'ADN incorporés dans des bicouches lipidiques artificielles peuvent être utilisés dans la biodétection et le séquençage. Dans cette thèse nous avons exploré la possibilité pour ces structures de détecter des séquences d’oligonucléotides courts, tels que des microARN. Les fluctuations de la structure membranaire causées par l'insertion et la désorption des nanopores sont utilisées pour effectuer des mesures précises de la conductance de la bicouche formée à l’intersection entre deux gouttes liquides (DIB) et pour mesurer la concentration d'oligonucléotides hybridés avec les nanopores d'ADN.

Le chapitre 1 fournit une introduction générale des bicouches lipidiques artificielles ainsi que trois types de nanopores. Un aperçu des méthodes de détection de microARN est fourni, ainsi que des informations détaillées sur les méthodes de détection des nanopores d'ADN.

Le chapitre 2 présente les méthodes expérimentales utilisées dans notre thèse.

Le chapitre 3 présente des expériences réalisées sur un nanopore formé par six hélices de 42 nucléotides de long et illustre la configuration expérimentale de l'approche DIB en manipulant mécaniquement la bicouche lipidique. Les résultats de simulations de dynamique moléculaire (oxDNA), des expériences de microscopie à force atomique (AFM), des expériences d'électrophorèse sur gel et des expériences de fluorescence sont rapportés. Et une discussion sur les mesures de conductance est donnée.

Le chapitre 4 décrit les perturbations dans la conductance d’une bicouche induites par un origami d'ADN (T1) combinant un nanopore à six hélices avec un origami d'ADN 2D rectangulaire planaire. Nous considérons également une autre structure optimisée origami d'ADN (T2) obtenue en générant un nanopore relié directement à une plate-forme 2D. Les résultats de caractérisation électriques (mesure de la conductance) ainsi que l'interaction avec une bicouche lipidique supportée à l’aide d’une microbalance de quartz avec mesure de la dissipation (QCM-D), sont également présentés.

Le chapitre 5 présente des conclusions et quelques perspectives.

Mots-clés : Nanopore d'ADN, Conductance, Détection de microARN, Bicouches d'interface gouttelettes

DNA nanopore as a signal transducer for the detection of short oligonucleotides

Abstract

DNA nanopores incorporated into artificial lipid bilayers are currently used in biosensing and sequencing. In this thesis, we investigate their ability to detect short oligonucleotides, such as DNA analogues of microRNA, fluctuations of lipid membrane structure caused by the insertion and desorption of nanopores bound to the target oligonucleotides are used to make precise measurements of the conductance of the bilayer formed at the intersection of two liquid droplets (DIB) and eventually monitor the concentration of target oligonucleotides.

Chapter 1 provides a general introduction of artificial lipid bilayers as well as three types of nanopores. An overview of microRNA detection methods is provided, along with extensive information on DNA nanopore detection approaches.

Chapter 2 introduces the experimental methods of this thesis.

Chapter 3 introduces laboratory work on the conformation of a 'cork' nanopore formed by the assembly of 42-nucleotides-long six-helix bundles and illustrates the experimental setup of the DIB approach (formation and characterization of DIBs). Results from coarse-grained oxDNA simulations, atomic force microscopy experiments (AFM), gel electrophoresis experiments, and fluorescent experiments to characterize this cork nanopore are reported. Also, a discussion about the conductance measurements in current recording experiments is given.

Chapter 4 describes a DNA origami (T1) structure which combines a six-helix-bundle nanopore with a planar rectangular 2D-DNA origami. In addition, current recordings in the absence and presence of targets are described. We also report a further optimized structure of DNA origami (T2) obtained by generating a nanopore and a 2D platform with a single scaffold. Results from experiments to characterize the structure and to investigate the bilayer conductance in current recording measurements, as well as the interaction with supported lipid bilayers in quartz crystal microbalance with dissipation monitoring (QCM-D) experiments, are presented.

Chapter 5 presents some conclusions and perspectives.

Keywords : DNA nanopore, Conductance, microRNA detection, Droplet interface bilayers

Unité de recherche

Centre de Recherche Paul Pascal (CRPP), CNRS, UMR 5031

115 Avenue du Dr. Albert Schweitzer, 33600 Pessac – France.

Acknowledgements

This thesis work was carried out at Chimie et Biologie des Membranes et Nanoobjets (CBMN, UMR 5248) and Centre de Recherche Paul Pascal (CRPP, UMR 5031), Université de Bordeaux, between October 2019 and October 2023, funded by Chinese scholarship Council (CSC), for which I am really grateful.

I would firstly like to express my deepest gratitude for my thesis supervisor, Juan Elezgaray for giving me a chance to join his group and start my study in France. He guides my research work by lots of insightful discussion and constantly suggestions. With his countless guidance in both academic and life, I am able to finish my phD study and this thesis.

My thanks also go to Carlos Drummond for in-depth comments and invaluable discussions in characterisation of my DNA nanostructures when using atomic force microscopy, Quartz crystal microbalance with dissipation monitoring and my topic, to Gille for giving me his kindful guide in lots of AFM experiments, and to Isabelle Broutin for her time and supportive work on Transmission electron microscopy. I am very much indebted to all the members in CBMN and CRPP, previous and present, I will always remember the nice time we spent together in our lab and university events: Rongyu, Andrea, Ilyses, Flavia, Luna, Lucas, Yarong, Yujie, Dandan, Jiangzhen, Longfei, Shaolin, Yaohui and Zhangjian, etc. Their encouragements and support keep me going further during my study and abroad life. Because all of them, I spent a happy life in Bordeaux.

Not quite last, but furthest from least, no words can express my appreciation to my grandparents, Wuxue and Shumin, my parents, Chongjun and Xia, my brother and his wife, Haosheng and Huihui, they support me in everything and stand in my side. And thank you to all my families for their constant love, support, and encouragement, I’m lucky to have all of them in my life, – throughout the sometimes rather unpredictable path that I have followed.

List of Abbreviations

2D	two-dimensional
3D	three-dimensional
4HB	4-helix bundle
6HB	6-helix bundle
52HB	52-helix bundle
AeL	aerolysin
AFM	Atomic force microscopy
α-HL	α-hemolysin
APS	ammonium persulfate
bis-acrylamide	N,N'-methylenebisacrylamide
bp	base pair
cDNA	complementary DNA
CHA	catalytic hairpin assembly
$CHCl_3$	chloroform
Chol	cholesterol
CNS	central nervous system
CRISPR	clustered regularly interspaced short palindromic repeats
DHB	Droplet on hydrated support bilayer
DIB	Droplets Interface Bilayer
DOPC	dioleoylphosphatidylcholine
DOPS	dioleoylphospatidylserine
DphPC	1,2-Diphytanoyl-sn-glycero-3-PC, $C_{48}H_{96}NO_8P$
EDTA	ethylenediaminetetraacetic acid
FAM	carboxyfluorescein
FITC	fluorescein isothiocyanate
FRAP	fluorescence recovery after photobleaching
FRET	fluorescence resonance energy transfer
GO	graphene oxide
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
LOD	limit of detection
miRNAs	microRNAs
ML	Monolayer
Mw	Molecular weight
NAF	nipple aspirate fluid

NPs	nanoparticles
nS	nanosiemens
nt	nucleotide
NUPACK	nucleic acid package
OPOE	n-Octylpolyoxyethylene, poly(ethylene glycol) octyl ether, $(CH_3(CH_2)_6CH_2(OCH_2CH_2)_nOH$,
PAGE	Polyacrylamide gel electrophoresis
PC	phosphocholine
PCR	polymerase chain reaction
PMMA	poly(methyl methacrylate)
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-PC, $C_{42}H_{82}NO_8P$
PPIs	protein–protein interactions
QCM	quartz crystal microbalance
QCM-D	quartz crystal microbalance with dissipation
QNM	quantitative nanomechanical mapping
rpm	revolutions per minute
RT	reverse transcriptase
RT-qPCR	real-time quantification of polymerase chain reaction
SCR	single channel recording
SiO_2	silicon dioxide
SLB	supported lipid bilayer
SMF	single molecule fluorescence
SVL	supported vesicular layer
TEM	transmission electron microscopy
TEMED	N,N,N’,N’-tetramethylethylendiamine
TRIS	tris(hydroxymethyl)aminomethane
UMIs	unique molecular identifiers
UV	ultraviolet

Content of Figures

Content of Tables

Table of Contents

1 Generalities and Context

1.1 Introduction of lipid bilayer membrane

1.1.1 Background of lipid bilayer membrane (BLM)

The intricate cellular membranes possess a remarkable degree of conformational flexibility, lateral tension, and electrostatic properties, which collectively present complex environments that can facilitate various biological and biochemical reactions for a multitude of functions.[1] Lipids, which are small molecules composed of hydrophobic or amphiphilic characteristics containing both non-polar and polar regions as in **Figure 1-1a**, play a vital role in this process. The glycerophospholipids, phosphatidylcholine (PC) and phosphatidylethanolamine (PE), are the major lipid components of eukaryotic and prokaryotic membranes. The cells leverage the lipid bilayers by controlling and tuning bilayer thickness, curvature, or physical constrains, thereby providing avenues for cellular function, particularly for membrane protein function (**Figure 1-1b**).[2,3]

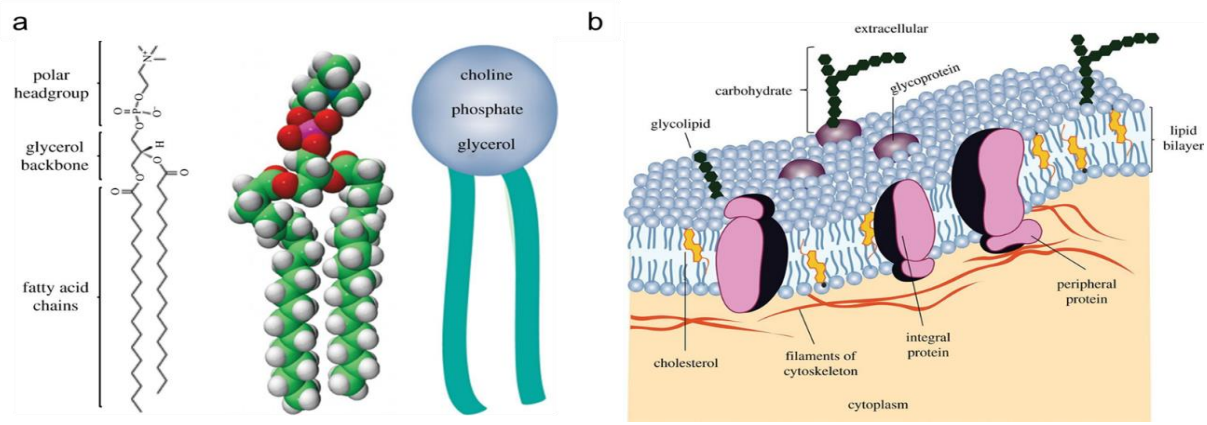


Figure 1-1. The schematic arrangement of phospholipid model and cell membrane.
(a) The presented figure illustrates the chemical, three-dimensional, and schematic structure of a phospholipid model composed of two fatty acid chains, a diglyceride backbone, and a phosphate headgroup (choline). It is noteworthy that while the phosphate heads exhibit hydrophilic properties, the fatty acid tails are hydrophobic. The arrangement of these polymers results in the formation of a lipid bilayer, which is a fundamental component of cell membranes, specifically the plasma membrane. (b) The cell membrane is a complex and dynamic system, characterized by the presence of diverse proteins that are either embedded within the membrane or transiently associated with the lipid bilayer. The images are reproduced from ref. [2].

To gain a comprehensive understanding of the cell membrane environment, researchers often turn to artificial bilayer systems. These systems, such as Droplet on hydrated support bilayer (DHB) and Droplets Interface Bilayer (DIB), as shown in **Figure 1-2**, allow for the manipulation of lipid membranes through techniques such as single channel recording (SCR) and single molecule fluorescence (SMF).[4-6] By utilizing these methods, researchers can study the behaviour of individual molecules within the membrane, as well as observe the interaction between the membrane and various substances. The use of artificial bilayer systems can provide valuable insights into the structure and function of the cell membrane, which is essential for a wide range of biological processes, and has greatly contributed to our understanding of the cell membrane environment.

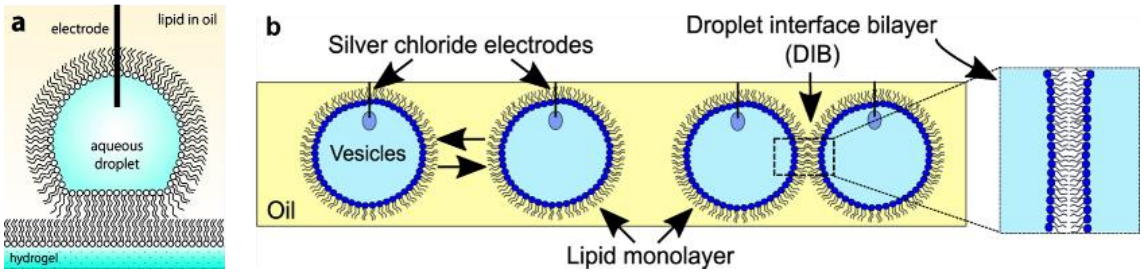


Figure 1-2. Schematic illustration of a droplet on a hydrated support bilayer (DHB) and two droplets forming a droplet interface bilayer (DIB).

(a) Schematic illustration of a droplet situated on a hydrated support bilayer (DHB). Lipid monolayers form spontaneously on aqueous surfaces when immersed in a lipid solution in hydrophobic oil. When the monolayer on the hydrogel and the monolayer of one aqueous droplet are brought into contact, they zip together to form a lipid bilayer. The image is reproduced from ref. [4]. (b) Two water droplets comprised of lipid vesicles are placed on the agar-coated tips of silver/silver chloride electrodes and placed into a well of oil. Following a brief incubation period, highly packed lipid monolayers assemble at the water–oil interface, culminating in the formation of a lipid bilayer when the droplets are manipulated into contact (as highlighted in the magnification). The image is reproduced from ref. [6].

Artificial black lipid membranes (BLMs) have been the focus of extensive research for over six decades, since Mueller et al. introduced the pioneering work in 1962. In their seminal study, lipid films were fabricated from hydrophobic pores coated with lipid-containing solutions.[7] After the formation of planar lipid bilayer is reported by Montal and Mueller using the variants of the Langmuir-Blodgett technique, where two air/water monolayers are raised past an aperture.[8] Subsequent studies by Wiese and Münstermann in 1998 utilized the ‘fold’ method to decrease the use of organic solvents, predominantly decane, in the membrane preparation. Lipids were spread twice with an organic solvent on top of the aqueous buffer.[9] These types of films, referred to as black lipid membranes, can be visualized as grey-black spot reflections under the microscope as they thin out.[10] The development of the droplet

interface bilayer (DIB) method in 2008 provided a superior platform for the formation of stable BLMs in an oil bath, thereby reducing the non-volatile oil present in the BLM.[5] Additionally, solvent-free lipid membranes have also been introduced in the literature. For instance, Teng et al. utilized semiconductor technology to create silicon (Si) chips with nano-tapered apertures, resulting in stable and robust oil-free BLMs.[11] These above black lipid membranes and solvent less lipid membranes belong to the category of free-standing BLMs. Except that, polymerized supported lipid membranes and supported lipid membranes on (i) metal [12] (ii) silicon (iii) fiberglass are classified to the supported lipid membranes into the third category.[13] Moreover, researchers have also developed BLMs doped with nanomaterials, such as fullerene-doped BLMs, based on hybrid membranes composed of lipids and fullerene derivatives that respond reversibly to light illumination.[14]

1.1.2 Lipid Bilayer applications

Lipid membranes are formed through the spontaneous, non-covalent assembly of lipids into bimolecular leaflets, which act as a permeability barrier surrounding aqueous volumes. These lipid bilayers possess the ability to store energy, and ion channels can facilitate the transport of ions by inputting voltage.[15] Despite the multifunctional advances in black lipid membrane (BLM) platforms, methods for measuring artificial BLM systems have remained largely unchanged over the past few decades. In the original BLM system, both sides of the BLM are filled with a buffer solution, and two electrodes immersed in the buffer solution apply a transmembrane voltage to the BLM to measure the transmembrane current (**Figure 1-3a**). Inspired by the evolution from two-terminal diodes to three-terminal transistors in semiconductor technology, Teng et al. improved the control of the BLM system by applying a voltage parallel to the BLM (lateral voltage) in addition to the traditional transmembrane voltage (**Figure 1-3b**).[11] They utilized lateral voltage to investigate the effects of fullerene-derivative-doped or ion-channel-incorporated lipid bilayer membranes and discovered the enhancement in transmembrane current measurements.

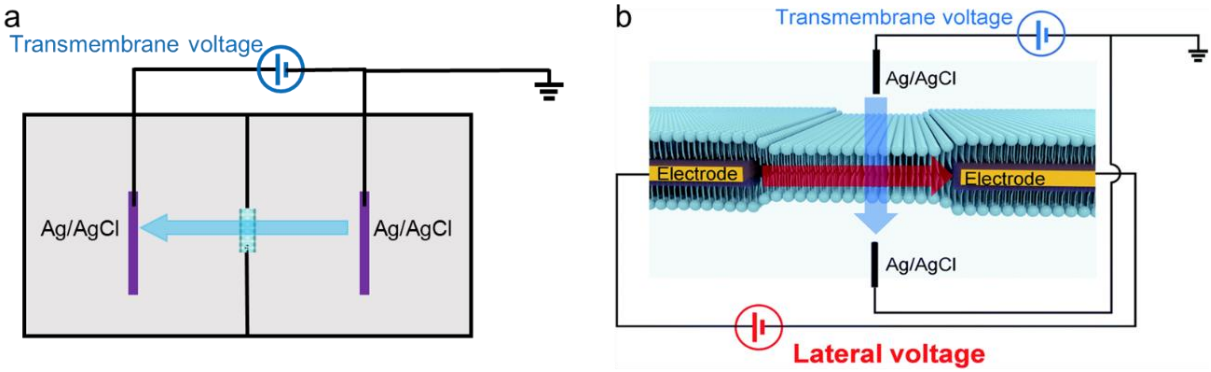


Figure 1-3. Scheme of BLM system applying traditional transmembrane voltage and a voltage parallel. *(a) Schematic representation of an original BLM system. (b) Schematic representation of a BLM system with two distinct input signals, namely lateral and transmembrane voltages. The image is reproduced from ref. [11].*

1.2 Fundamental concepts of nanopores

Nanopores, pores of nanometer dimensions, are ubiquitous in biology as pathways through which molecules traverse. Biological cells possess a variety of nanopores that control the transport of ions and molecules within and outside the cell and between subcellular structures. These nanopores could control the transport of messenger RNA from the nucleus into the nuclear membrane pore of the cytoplasm; the secretion of proteins across the pores of the cell membrane of the organelle and viruses also delivers the genome into the cells through the pores into cells. While many of these nanopore transmissions are actively regulated, some transmissions through holes are passive.[16,17] A notable milestone in nanopore detection was achieved by Deamer et al. in 1996,[18] who proposed the use of a solute for DNA sequencing. Recent advancements in nanotechnology have enabled the passage of individual polymer molecules through nanopores of only a few nanometers in size. The current research focuses on the phenomenon of nanopore transport, aiming to elucidate the physical processes behind it and use this phenomenon for multifunctional exploration.[19,20] The cell membranes and channel proteins found in eukaryotic cells are considered biological nanopores, and are believed to be the earliest form of selective biomolecule translocation devices. With the advancement of structural nanotechnology, there has been a growing interest in the creation of biomimetic nanotechnology in vitro. Consequently, numerous artificial synthetic nanopores have been developed, categorized into three main types: biological nanopores consisting of peptides, proteins, and DNA embedded in phospholipid membranes; solid nanopores synthesized from solid substances; and nanopores that are hybridized with both biological and solid nanopores.[21,22] And DNA nanopores are highlighted to be introduced in chapter 1.2.4.

1.2.1 Biological nanopores

Many biological processes depend on the presence of a single molecule passing through a nanopore within a membrane. Previous studies focused on naturally-occurring nanopores. Notably, biological nanopores, which are comprised of peptides [23] and proteins with narrow channels (generally 1–8 nm in diameter), such as Phi29, α-hemolysin (α-HL), MspA, etc. and proteins possessing narrow channels (typically ranging from 1 to 8 nanometers in diameter), such as Phi29, α-hemolysin (α-HL), and MspA, (**Figure 1-4**) are included in this category.[23]

Protein nanopores have the ability to generate highly sensitive and precise electrical signals when utilized in the analysis of small molecules, amino acid sequencing, and protein-protein interactions. For instance, Wu and Bayley [24] conducted an experiment using a mutant α-HL (**Figure 1-5a**) to detect mustards via thiol groups that were modified within the nanopore's lumen, resulting in a detection time of only 10 minutes at a concentration of 50μM. Similarly, Lu and his colleagues [25] employed a 5′-benzaldehyde modified poly(dA)$_4$ probe in conjunction with a K238Q mutant aerolysin (AeL) nanopore (with a diameter of approximately 1nm, **Figure 1-5b**) to discriminate cysteine and homocysteine in a mixture at the single-molecule level. The translocation behaviour of two amino acids in the nanopore produced duration distribution of signals, measuring 0.53±0.01 ms and 446.7±1.0 ms, respectively. Avinash and Liviu [26] have reported an engineered protein nanopore designed with a protein receptor (Inactivated RNase barnase, Bn) to detect another protein ligand (89-residue barstar, Bs), which is an inhibitor of Bn RNase (**Figure 1-5c**). When a protein ligand is present and binding to the protein receptor, transient protein–protein interactions (PPIs) can be monitored in complex biological fluids.

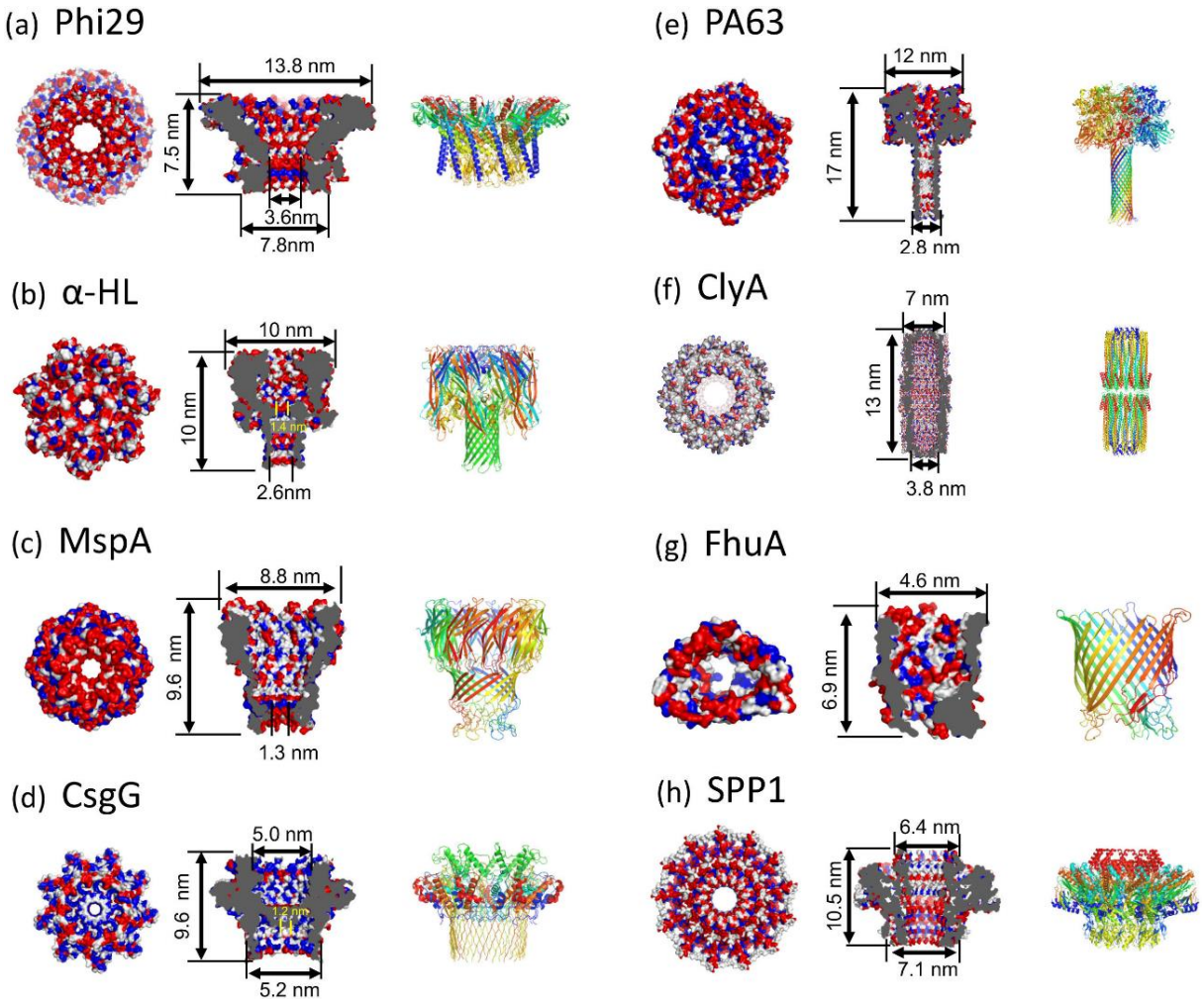


Figure 1-4. Structure of eight biological nanopores.
Bottom, cross-section and side views of: (a) phi29 connector (PDB: 1H5W), (b) α-hemolysin (PDB: 3ANZ), (c) MspA (PDB: 1UUN), (d) CsgG (PDB: 4UV3), (e) PA63 (PDB: 1V36), (f) ClyA (PDB: 2WCD), (g) FhuA (PDB: 1BY5), and (h) SPP1 connector (PDB: 2JES). The image is reproduced from ref. [23].

Researching the biological nanopores is helpful for understanding the fundamental mechanism of some biological processes and advancing our knowledge of biology. And also, they have potential applications in DNA sequencing, drug delivery and some other developing technologies. Moreover, some additional issues should be considered when using biological nanopores,[28] such as, the difficulty of modifying protein nanopores, the increased sensitivity changes of protein conformation, and activity in external environmental conditions, e.g., electrolyte pH or temperature.

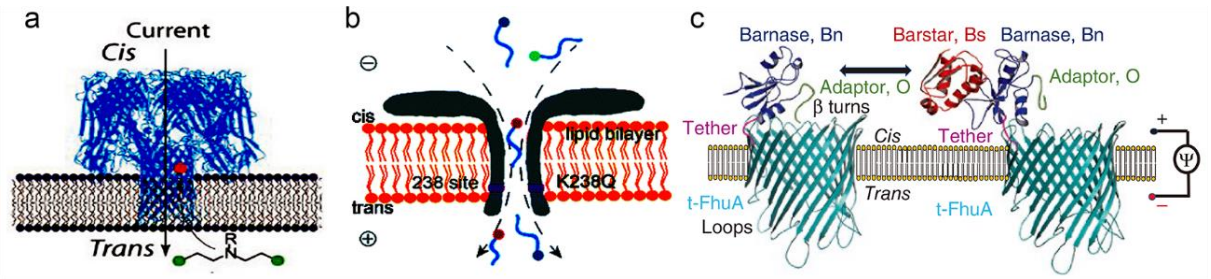


Figure 1-5. Scheme of three different biological nanopores.
(a) Scheme of the detection of mustard gas analogues through covalent reaction with engineered α-HL pores, where thiol groups are represented as red spots. The image is reproduced from ref. [24]. (b) A schematic of the K238Q AeL nanopore is depicted, which analyses the complex of cysteine and homocysteine binding with a probe. The image is reproduced from ref. [25]. (c) A schema of a protein pore-based nanostructure is presented for real-time sampling of transient protein–protein interactions. The image is reproduced from ref. [26].

1.2.2 Solid-state nanopores

As a result of the diverse range of solid-state nanopore materials available, including inorganic,[29-33] polymer,[34,35] one- and two-dimensional materials,[36] as shown in **Figure 1-6**, as well as their compatibility with semiconductor manufacturing techniques, solid-state nanopores have emerged as a viable alternative to biological nanopores. These various types of synthetic nanopores offer an exciting platform for numerous nanopore applications, such as DNA sequencing, protein analysis, and drug delivery etc.[37,38] Solid-state nanopores are fabricated using synthetic thin membranes and feature a small hole or channel with dimensions on the nanometer scale, typically ranging between 1–100 nm in dimeter and generated using drilling technologies. Solid-state nanopores can be damaged or degraded over times, which are more stable and durable compared to biological nanopore, with the added benefit of allowing for custom-made materials to suit specific applications.[39] Nevertheless, there remain challenges in improving the potential of functionalization in solid-state nanopores. These challenges include changing pore size, adjusting the hydrophilicity and hydrophobicity, modifying surface charge density, affecting ion transport, reducing ion current noise, improving molecular biology signal-to-noise ratio sensing,[40] and enhancing cost-effective and efficient batch fabrication, selectivity, and characterization methods.[39,41,42]

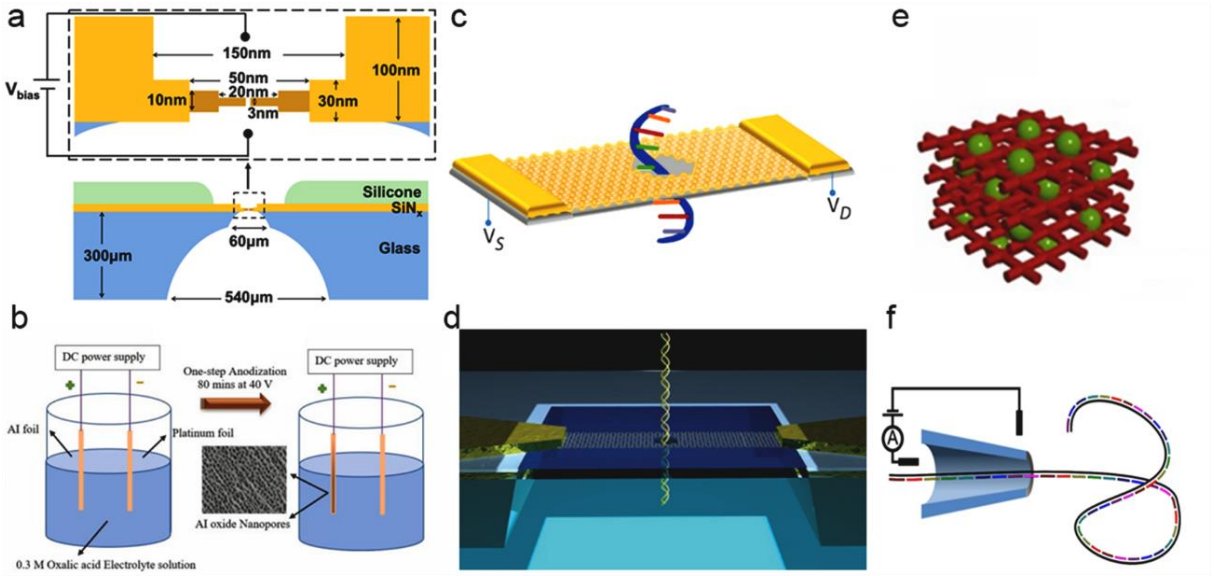


Figure 1-6. Scheme of six distinct categories of solid-state nanopores.

The solid-state nanopore are constructed by different materials. (a) Silicon nitride, the image is reproduced from ref. [29]. (b) aluminium oxide, the image is reproduced from ref. [30]. (c) gold, the image is reproduced from ref. [31]. (d) graphene or graphene oxide-based, which exhibit a range of exceptional mechanical and electrical properties, the image is reproduced from ref. [34]. (e) organic frameworks, the image is reproduced from ref. [35]. and (f) nano capillaries, the image is reproduced from ref. [43]. These materials can be employed for various applications in the field of nanotechnology.

1.2.3 Hybrid Nanopores integration with biological nanopores and synthetic membranes

Biological nanopores and synthetic membrane hybrid nanopores exhibit exceptional characteristics [21] that can alleviate the challenges of high environmental requirements for phospholipid membranes and integrate the benefits of both biological nanopores and solid-state nanopores. The pioneering work of Hall et al. [44] demonstrated the creation of a hybrid pore through the directed insertion of α-HL nanopores into SiN_x nanopores with diameters ranging from 2.4 to 3.6 nm (**Figure 1-7a**). More recently, Mojtabavi and his team [45] devised a novel hybrid nanopore by covalently immobilizing a bacteriophage portal protein within a solid-state nanopore, resulting in lipid-free and high-voltage (200–600 mV) biomolecular sensing, as shown in **Figure 1-7b**.

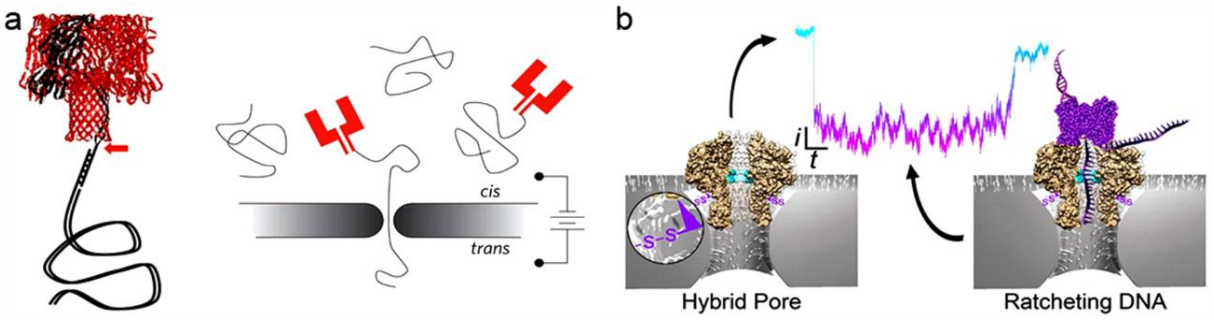


Figure 1-7. Scheme of two examples of hybrid nanopores.
(a) A 3kbp dsDNA to an α-HL nanopore via a 12-nucleotide oligomer, which is then electrophoretically translocated through a narrow SiNx nanopore. The image is reproduced from ref. [44]. (b) The immobilization of a G20c portal protein (an internal channel containing a 1.8nm constriction) within a chemically modified solid-state nanopore (7nm), which allows for motor-protein-mediated DNA ratcheting transport through the hybrid nanopore. The motor protein can slow down the DNA transport speed by five orders of magnitude compared to electrophoretically driven DNA transport. The image is reproduced from ref. [45].

The stability of synthetic membranes far surpasses that of phospholipid membranes, making their replacement a viable option for maintaining stability while retaining the high sequencing accuracy of proteins. Consequently, the essential properties required for the precise structure and engineering of biological protein pores can be combined with the potency needed for the fabrication of integrated systems. Biological nanopores possess a constant geometry that ensures sensitive and repeatable sensing, however, their structure may be impaired when inserted into a solid membrane.[46] This hybrid approach, which has yet to be extensively explored, involves the incorporation of a biological structure into a well-designed nanopore through the suppor.[47,48]

1.2.4 DNA nanopores

In recent years, there has been significant development in the area of DNA-based nanostructures that emulate naturally occurring membrane proteins.[49] The creation of biomimetic DNA nanotechnology through in vitro methods has been a captivating exercise for scientists, providing a rigorous platform for DNA nanopore design and synthesis. These DNA nanopores have been designed for various applications, including transmembrane ion/molecular channels, intracellular transport, molecular selective gating, label-free detection/sensing of biomolecules, DNA sequencing, and intercellular communications.[50] **Figure 1-8** illustrates several distinct shaped and functioning DNA nanopores.

Currently, DNA nanopore architectures have predominantly been dominated by helix bundle (HB) designs, such as 4HB, 6HB, 8HB, and 54HB.[51] For instance, our team designed a modified two 6-helix bundle (6HB) structure, known as the cork nanopore, described in chapter 3. In the same way, Gopfrich and his colleagues [52] developed a 4HB nanostructure that is similar in size to biological ion channels, with an exterior diameter of 5nm, an interior diameter of 0.8nm, and a molecular weight of 45kDa (**Figure 1-8a**). The structure is self-assembling in one minute due to its basic design and exhibits ion conduction through the lipid bilayer with gating and voltage switching characteristics, demonstrating its scalability for biological applications. Howorka group [53-56] has also developed several modified 6HB structures (**Figure 1-8b**) as a membrane-spinning gate to regulate transmembrane flux by modelling electric field-driven passage in synthetic nanopore systems. Similarly, Langecker et al. [57] used 26 cholesterol to develop self-assembled 2nm-wide DNA nanopores as nanoscale transmembrane channels in lipid bilayers (**Figure 1-8c**). The scaffold DNA origami comprises a stem that penetrates and crosses a lipid membrane and a barrel-shaped cap that attaches to the membrane. In their research, single-channel electrophysiological investigations revealed the conductance on the scale of 1 nanosiemens (nS), and single-molecule translocation experiments proved that synthetic channels may be utilized to distinguish individual DNA molecules. Thomsen and his colleagues [58] also describe a synthetic 9nm-wide DNA nanopore (**Figure 1-8d**) controlled by programmable lipid flaps and equipped with a size-selective gating system for macromolecular translocation.

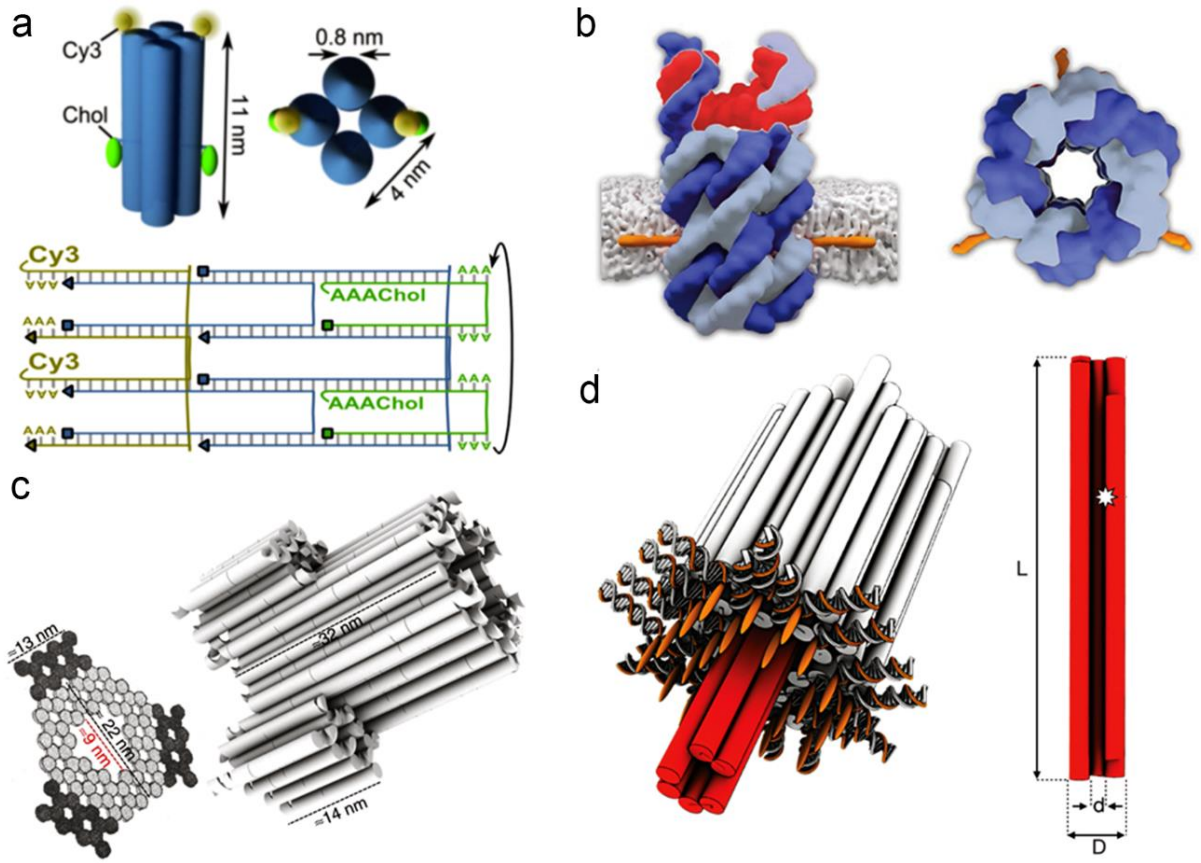


Figure 1-8. Here are four schematic examples of existing DNA nanopores.
(a) Schematic side view (left) and top view (right) of a 4HB structure composed of four interconnected duplexes represented as cylinders with two cholesterol-anchors in green and two Cy3 tags in yellow. Pathways of the eight strands (below) form the four duplexes and show positions of the Cy3 tags and cholesterol modifications. Squares represent the 5′ ends, triangles represent the 3′ ends, base pairs are indicated as vertical bars. The image is reproduced from ref. [52]. (b) Structural model of a 6HB composed of six DNA strands (alternately in dark blue and pale blue) with cholesterol-based membrane anchors (orange) on its outside. The image is reproduced from ref. [54]. (c) Structural model of a pseudosymmetric nanopore is based on a hexagonal origami lattice and has a 9-nm inner pore diameter, a 22-nm outer diameter, and a length of 32 nm. The image is reproduced from ref. [58]. (d) Structural model of a 54HB produced by 54 double-helical DNA domains packed on a honeycomb lattice, with cholesterol-modified oligonucleotides (orange ellipsoids) hybridizing to single-stranded DNA adaptor strands. Specifically, transmembrane stem is visible in red. The length L of transmembrane channel is 47nm, the tube diameter D is 6nm, and the inner diameter d is 2nm. The image is reproduced from ref. [57].

DNA nanopores have the potential to provide truly parallel, high-throughput analysis of DNA and proteins by creating nanopores in synthetic membranes. In particular, biomimetic DNA nanotubes have considerable potential in bioimaging, biosensing, and therapy due to their excellent biocompatibility and addressability. However, real-world biological applications remain challenging. The technology needs be enhanced in a more robust, reliable, and controllable way. Innovative simulations, for example, are frequently based on well-structured nanotubes, the folding process of which is still little understood, and research to

improve the efficiency, purity, and scalability of DNA nanopore manufacture is critical for their application. Furthermore, the processes by which DNA nanopores precisely regulate the release of their cargo, the ion leakage caused by nanochannel oscillations in membranes in DNA sequencing, and possibly the interaction of drug delivery vehicles with subcellular organelles are unknown, posing challenges for high-end applications. Despite these challenges, amounts of progress has been made in DNA nanotechnology and we expect biomimetic DNA nanopores to have a particular influence on fundamental biological research and real-world healthcare.

1.2.5 Analysis of membrane pores

Consider the previously mentioned nanopore classes, each with its own specific strengths and material conditions Therefore, when designing new pores for a specific application, it is essential to have a good understanding of the appropriate materials that impact the structural accuracy, size, and chemical variety of the matching nanopores (**Table 1-1**). Using various construction materials can lead to greater technological and scientific research and application. For instance, the use of graphene oxide as a construction material can enhance the sensing capabilities of the nanopores; the use of gold as a material can improve the stability and durability of the nanopores. Similarly, the use of DNA as a material can provide an easy way to define and modify nanopores. In conclusion, the design approach for constructing new nanopores is highly dependent on the knowledge of appropriate materials. Using a variety of materials can lead to significant advancements in scientific and technological research, which ultimately benefits society.

Table 1-1. Four classes of pores featuring the strengths and challenges of the building materials.

pores	strengths	challenges
peptide	+ Many non-standard amino acids + Non-protein structural folds + Design from scratch + Fast insertion kinetics	– Simple architectures – Lumen <1.5 nm
protein	+ Defined and stable scaffold + Simple engineering via amino acid changes + Targeted addition of synthetic components	– Re or de novo design difficult – Few defined pores wider >5 nm
DNA	+ Simple de novo design + Dedicated design software + Structures >20 nm accessible	– Limited chemical repertoire – Structural fluctuations and electronic leakiness of walls – Slow insertion kinetics
Synthetic organic	+ Widest chemical repertoire + Flexible design + Compact sizes <5 nm	– No unifying architectural principles – Challenging structural analysis – Inherent gating/closing of channel

Table adapted from ref. [59].

1.3 DNA nanostructure technologies

Over billions of years, DNA has been the primary means by which life on Earth encodes almost all forms and functions in the biochemical patterning of the four nucleotide bases.[60] In recent decades, the interdisciplinary integration of genomic chassis,[61] gene sequencing,[62] stereoscopic construction, clustered regularly interspaced short palindromic repeats (CRISPR),[63] and DNA origami technologies [64] has led to innovative design wisdom in DNA technology. Scientists have successfully recombined the genomes of simple life forms, designed genome chassis for artificial evolution platforms, and used gene editing technology CRISPR for genome correction; and precisely self-assembled of DNA nanomaterials like DNA origami.[65] The integration of construction principles and user-centered design techniques can facilitate the development of personalized bioinformatics.

DNA, being a versatile biomolecule,[49] is extremely important in life, leading to research focused on its existence circumstances, stability, detection techniques, and mutual conversion of diverse DNA configurations. Watson-Crick base pairing proposed more than thirty years ago [66] has been used to rationally design nanometre-scale structures from nucleic acids. DNA nanotechnology has allowed for increasingly intricate structures, particularly since the invention of the origami technique. DNA has been found to be a flexible building block for programmable DNA nano-object manufacturing.[49] A multiple-kilobase scaffold strand [67] is folded into an expected helice assays by interactions with hundreds of oligonucleotide staple

strands,[68] and their morphologies include single-stranded tiles, multi-crossover tiles, DNA origami, multi-rungs, planar, twisted and curved nanoscale,[69,70] 2D [71] and 3D [62,72] nanostructures. The bending and twisting of DNA origami structures can be tuned by the patterns of base pair insertions and deletions and chemical adducts (chemical adduct, e.g. ethidium bromide).[73] However, while general approaches for creating DNA origami polygonal meshes and design software are available, there are still constraints arising from DNA geometry and sense-antisense pairing, necessitating some manual adjustment during the design process. In essence, the nanopore experiment is a single-molecule detection method with ultra-sensitivity and high resolution, which can distinguish the differential difference between single nucleotide bases A, T, G and C, making it a promising tool for DNA research. Compared with the previous nanopore applications mainly based on gene sequencing, recent advances in nanopore technology have expanded to analyse a diverse range of molecular attributes, including the physical and chemical properties of nucleic acids, proteins, nanoparticles and ions.

1.3.1 Molecular dynamics simulation for DNA nanostructures

Molecular dynamics simulation has become an essential tool in the design and analysis of DNA nanostructures. These structures of DNA origami are often complex, and their behaviour can be difficult to anticipate without extensive practice. However, researchers can now use computer-aided, scriptable tools, such as scadnano by Doty et al. [74] or caDNAno by Douglas et al. [69] with greater precision and accuracy to design DNA nanostructures. One of the most significant advancements in DNA nanostructure design computation is the use of molecular dynamics simulations. These simulations enable researchers to obtain faster and less expensive feedback and insight on the behaviour of DNA nanostructure. Among them, coarse-grained molecular dynamics simulation tools, such as oxDNA,[75] CanDo,[76] and NAMD,[77] have substantially simplified the process of constructing and testing DNA nanostructures. With these tools, DNA nanostructures can be simulated and predicted based on their designing model and molecular size. This makes it easier for researchers to explore the behaviour of these structures in different environments, such as diffusion model or hybridization. Additionally, researchers can study the shape and stiffness properties of DNA nanostructures, which can have a significant impact on their function and stability. Overall, the use of molecular dynamics simulations has revolutionized the field of DNA nanostructure

design and analysis. These simulations provide researchers with a powerful tool to explore the behaviour of these structures in a variety of settings, enabling them to design more complex and sophisticated DNA nanostructures with greater accuracy and efficiency.

1.3.2 How to insert DNA nanopore into lipid bilayers

When DNA nanopores are insert into lipid bilayers, the initial step involves increasing the hydrophobic characteristics of the DNA structure. To achieve this, researchers have explored a medium of hydrophobic groups that encircle the surface of the nanostructure and are capable of binding to lipid membrane in terms of size and hydrophobicity, such as cholesterol, tocopherol,[78] or biotin [79] modifications. In this thesis, we employed the addition of cholesterol anchors to DNA nanostructures. These modifications of hydrophobic groups were modified at the end of single-stranded sequences extensions that are complementary to prolonged handles of the nanostructures (**Figure 1-9)**. It is important to distribute these positions throughout the structures to avoid hydrophobic interaction between adjacent anchors. By distributing of adaptors on the DNA nanostructures, the resulting structures can improve the incorporation to the membrane and decrease the tendency of aggregation. Furthermore, the introduction of cholesterol and saturated phospholipids can reduce membrane permeability and potentially strengthen liposome membranes.

In classic nanopore experiments,[5] two reservoirs are connected by a nanometer-sized pore, and an electric field is applied to nanopore to drive charged molecules through nanopore via electrophoretic force. Biomolecules of different sizes and charges can be distinguished according to parameters of translocation such as dwell time and current drop amplitude. The typical unitary conductance values of individual nanopores should correspond to theoretical values calculated using the following equation [55]:

$$\boldsymbol{G = \kappa \frac{\pi d^2}{4L + \pi d}}$$

(equation 1)

where κ is the electrical conductivity (equal to 10.86 S · m^{-1} for 1M KCl at 25°C), d is the predicted diameter of the pore, and L the length of the pore. However, this simple calculation should be used only as a rough guide for conductance predictions, because it assumes a constant mobility of electrolyte ions in negatively charged nanopores.

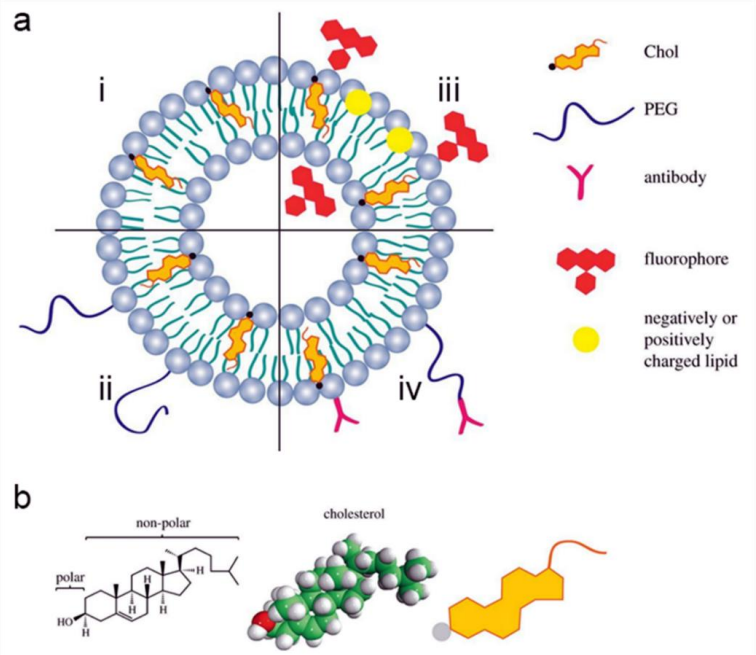


Figure 1-9. The scheme of (a) hydrophobic groups and (b) the intricate structure of cholesterol.
(a) We present a comprehensive analysis of the cross section of a liposome, which is characterized by spherical lipid bilayers. Specifically, we describe several types of liposomes: (i) conventional liposome; (ii) sterically stabilized liposomes (PEGylated surface provides a steric barrier that prevents the adsorption of proteins onto the liposome surface); (iii) ligand-targeted liposome; (iv) fluorescent liposome and charged liposomes. (b) Additionally, we provide a detailed examination of the chemical, three-dimensional and schematic structure of cholesterol (chol), which plays a crucial role in the formation and stability of lipid membranes. The image is reproduced from ref. [2].

1.4 MicroRNA detection

MicroRNAs (miRNAs) are a class of short single-stranded, noncoding RNAs in body fluids that are found in eukaryotic cells, with a length range of 19–22 nucleotides.[80] These miRNAs play a crucial role in regulating gene expression by binding to messenger RNA (mRNA) in the cell cytoplasm. The miRNAs are generated from hairpin-like precursors that can form base-pairing interactions with their target mRNAs within the RNA-induced silencing complex (RISC).[81] Instead of being translated quickly into a protein, this process results in the degradation or inhibition of mRNA translation, so if the level of a particular microRNA is underexpressed/overexpressed, the protein it normally regulates may be overexpressed/ underexpressed. Any underexpression or overexpression of specific miRNAs can cause corresponding changes in the protein expression levels of the target genes they regulate. In cancer cells, mutations can occur in the miRNA genes leading to abnormal expression levels of miRNAs, which in turn can cause aberrant gene expression and cancer

development and progression. Because of the complex regulatory functions of miRNA itself, it is closely related to some diseases. Researchers have addressed that plasma and salivary microRNAs can be considered reliable sources of biomarkers for cancer diagnosis and prognosis, eg. miR-15,[82] miR-21,[83] miR-34,[84,85] miR-155,[86-89] etc. Therefore, miRNA detection is of great significance in both disease diagnosis and miRNA function research. They are important biomarkers for non-invasive diagnosis, assessment, prediction and treatment of cancer.

However, due to the inherent low expression level, high sequence similarity, and short length of miRNA, there are some pre-analysis steps, including isolation and purification, must be taken before miRNA quantification. Traditional miRNA detection methods have limitations and cannot meet the current needs, and extensive efforts have been made to develop novel, high-precision, low-cost, high-practicability, checkpoint-multiplexed miRNAs. In vitro miRNA detection method. Such diagnostic methods are valuable for non-invasive diagnosis, assessment, and prediction of cancer.

1.4.1 MicroRNA detection methods in vitro

There are many methods for the detection of microRNA (miRNA), including real-time quantification of polymerase chain reaction (RT-qPCR), high-throughput microarray methods, RNA sequencing, DNA origami-based strategies, etc. Here, these four miRNA detection methods are introduced and discussed (**Table 1-2**).

In general, quantification of RNA molecules by RT-qPCR [90,91] involves two steps: the first step requires the use of a RNA-dependent DNA polymerase, also known as a reverse transcriptase, to copy RNA into complementary DNA (cDNA), the cDNA of the RNA target is synthesized using reverse transcriptase (**Figure 1-10a**). The second step then switches to the use of thermostable Taq polymerase and a pair of primers to amplify the cDNA as in a standard PCR test. During successful polymerisation, the probe is displaced and hydrolysed, separating fluorophore and quencher and releasing fluorescence. Amplification is monitored in real time by fluorescence, using dyes (such as SYBR safe) or specific fluorescent probes labelled with orthogonal dyes, which can be deduced the amount of PCR product generated.

Table 1-2. Comparison of four microRNA detection methods.

methods	Comments	Steps and time required	Total time	Sample required
RT-qPCR	·complicated ·long ·relatively accurate ·cost low ·high purity	1. Serum preparation, including blood clotting and centrifuging. (40 min) 2. RNA extraction. (1.5-3 hours) 3. Reverse transcription. (2-3 hours) 4. Real-time quantitative PCR. (1.5-3 hours)	6-10 hours	≥200μL
microarray	·complicated ·long ·cost high ·high purity	1. Serum preparation. (40-70 min) 2. RNA extraction. (1-2 hours) 3. miRNA target preparation and labelling. (2-3 hours) 4. miRNA array hybridisation, signal detection and array scanning. (24 hours)	28-32 hours	>1mL
RNA-seq	·complicated ·long ·cost low ·high purity	1. RNA extraction. (2-3 hours) 2. Ligation of 3' and 5' adapters. (18 hours) 3. Reverse transcription. (2 hours) 4. PCR amplification and purification. (2-3 hours) 5. Quality control and size selection. (3 hours) 6. Sequencing and data processing. (≥2 days)	>3 days	>1mL
Nanopore or DNA origami based method	·simple ·short ·cost low ·relatively low purity	1. Serum preparation, including blood clotting and centrifuging. (40 min) 2. Incubation nanopores with serum. (30-35 min) 3. Nanopore electrical assay. (20 min)	<1.5 hours	<1μL

Table adapted from refs. [92,93].

Microassays immobilizing by Catalytic hairpin assembly (CHA) components on the solid-phase chip surface allow enzyme-free isothermal amplification methods for simultaneous detection of various miRNAs in a single test. The target miRNA can trigger conformational transformations of hairpin-structured DNA probes on the chip and lead to the specific signal amplification, e.g., fluorescent signal (**Figure 1-10b**). Michael et al. [94] have reported an adapter ligation-based Small-seq method for microRNA sequencing. Custom-designed unique molecular identifiers (UMIs) were designed into adaptor ligation, capable of determining the number of small RNA molecules per cell. After using RT-qPCR amplification, Northern blot, gel extraction purification, and UMI enumeration, the DNA library pools can be analysed for miRNA sequencing.[95]

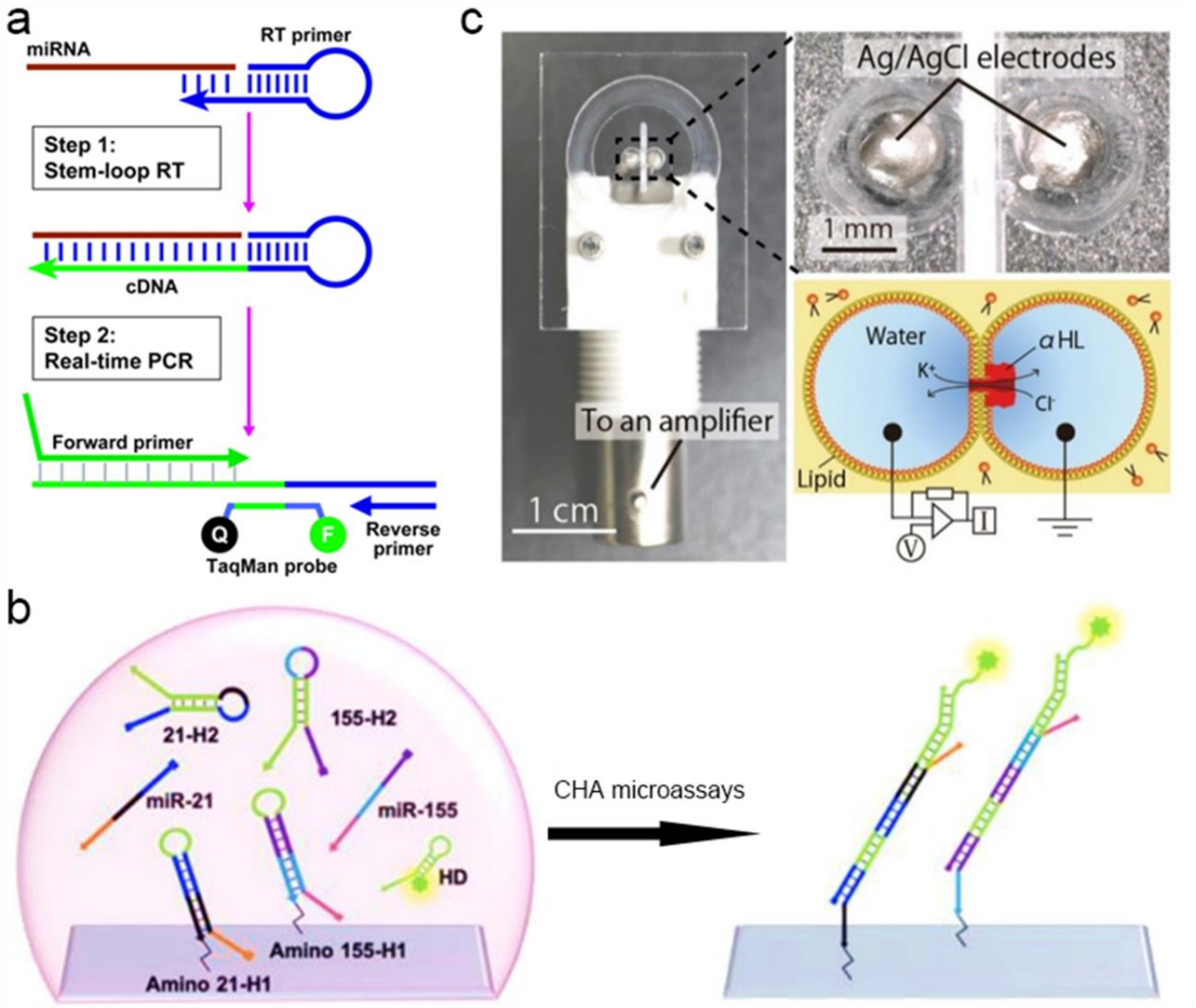


Figure 1-10. Scheme of several microRNA detection methods in vitro.

(a) Schematic description of TaqMan miRNA test. TaqMan-based real-time quantification of miRNAs includes two steps, stem–loop RT and real-time PCR. Stem–loop RT primers bind to at the 3′ portion of miRNA molecules and are reverse transcribed with reverse transcriptase. The resulting RT product is then quantified using conventional TaqMan PCR, which includes miRNA-specific forward and reverse primers, as well as a dye-labelled TaqMan probe. The purpose of tailed forward primer is to increase its melting temperature (Tm) depending on the sequence composition of miRNA molecules. The image is reproduced from ref. [90]. (b) Scheme of the microarray-based catalysed hairpin assembly (mi-CHA) biosensing for both miR-21 and miR-155. Mi-CHA, which is a solid support that contains many spots or probes that are designed to hybridize with specific miRNAs of interest. The CHA reaction release signal molecules and then can be detected by fluorescence. The image is reproduced from ref. [96]. (c) A photograph of the device for measuring channel currents. Lipid bilayers are prepared by the droplet contact method, and αHL is reconstituted in the lipid bilayer. The image is reproduced from ref. [97].

Electrical detection (**Figure 1-10c**) is based on the measurement of an electric current or potential difference [28,92,97-99], for instance, the change in system resistance when miRNA target hybridizes to a specialized probe molecule. Nanopores whose electric resistance is modulated by RNA channels can be used in such setups. The analysis method is based on the electrical signal amplification rather than the signal amplification of PCR. This method allows the uptake of single molecules in solution, such as miRNA, into DNA nanopores or the conformation change of DNA nanopore to study probes' structure, composition, and dynamic properties. It usually provides very low limit of detection (LOD) at low cost, and the highest resolution measurements may be performed at the single-molecule level, especially when determining the concentration of molecules in solution. Molecule counting through or binding with DNA nanopores can be utilized when the number of molecules is statistically discernible and the signals are above the noise limit. This capability focuses on the electrical detection and analysis of the current signal of a single molecule (microRNA) passing through a DNA nanopore.

1.4.2 What's the function of this 22nt microRNA, miR-21?

The utilization of microRNAs (miRNAs) has enabled the comparative evaluation of miRNA expression profiles in tumours and cell lines associated with cancer, with those of normal cells/tissues. Certain miRNAs, known as 'oncomiRs' [100], display varying levels of expression in cancer and possess the ability to influence cellular transformation, carcinogenesis and metastasis, functioning either as oncogenes or tumour suppressors. The dysregulated expression of miR-21 (**Figure 1-11**), one of the earliest miRNAs to be discovered in the human genome, has prominently emerged on its involvement in central nervous system (CNS) disorders,[101] chronic lymphocytic leukemia,[102] prostate cancer,[103] pancreatic cancer,[104,105] as well as lung cancer,[106] resulting in significant disability and mortality rates worldwide. Strikingly, microRNA-21 (miR-21) is a distinctive miRNA that has been found to be overexpressed in the vast majority of cancer types that have been analysed to date.[83,105]

Figure 1-11. The sequence of miR-21.
(a) Location of MIRN21 gene which codes for a 72-nt-long precursor miR-21 (pre-miR-21) on chromosome 17q23.1. (b) pre-miR-21 sequence and stem–loop structure folded using Mfold, mature miR-21 is shown in bold. (c) Sequences of mature miR-21, DNA analogue of miR-21, cDNA of miR-21. The image is reproduced from ref. [83].

1.4.3 How to use nanopore to detect oligonucleotides?

The pioneering work of single ion channels was conducted through electrophysiological experiments decades ago. These experiments enabled the measurement of ion flux through a single nanopore in the cell membrane.[107] By monitoring the current and force when molecules pass through nanopores, various phenomena involving DNA, RNA, and proteins can be investigated. To detect target molecules in low concentration or even impure clinical samples, we aim to use the nanopore method assisted by DNA strand replacement. Through a complex DNA assembly circuit system, low target signals can be amplified to produce significant current signals. For instance, Zhang et al. [108] have reported nanopore-based, label-free and amplification-free detection platform for detecting nipple aspirate fluid (NAF) sample for breast cancer diagnosis. When the probe binding to the biomarker in NAF, the conformational change of the engineered nanopore will lead to current blockage (**Figure 1-12**). And these peak signatures will enable us to simply and sensitively detect basal levels of biomarker from patients at single-molecule level.

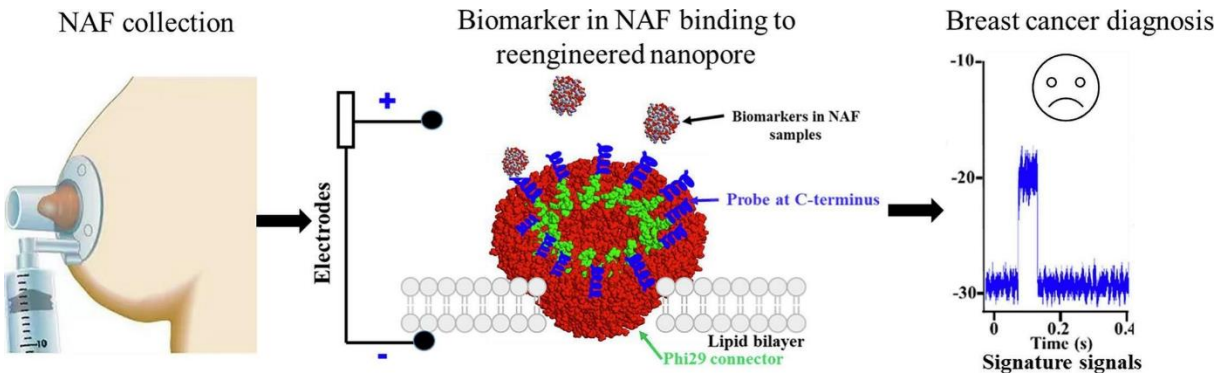


Figure 1-12. Schematic diagram of real-time detection breast cancer biomarkers from clinic nipple aspirate fluid (NAF) samples using the nanopore of Phi29 DNA-packaging motor.
The image is reproduced from ref. [108].

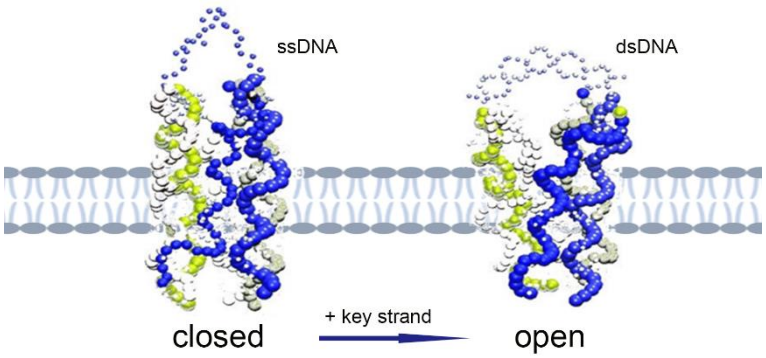


Figure 1-13. Coarse-grained model of the transition between closed and open conformations in absence and presence of 30-nt oligo targets.
On top of the nanopore, a single strand linking two helices across the lumen of the pore can bind to the key strand and change to a quite stiff double strand. The image is reproduced from ref. [109].

Previous work has reported that, except for protein channels, the 30-base nucleotide of DNA sequence [109] can be detected by DNA nanopore (**Figure 1-13**). Upon binding to its complementary sequence, the nanopore conformation changes and translates to a current increase. Here, the width of nanopore is theoretically increasing from 6nm (the width of the DNA nanopore) to ~ 10nm. However, for shorter length, such as 22-base nucleotides, nearly 7 nanometers, the conformation change from 6nm to 7nm can result in high testing errors in unreliable dimensions, while the current signal will not change much enough. To address this issue, the design strategy of nanopore structure needs to be considered. In the applying voltage state, the nanopore produces a linear current-voltage curve whose slope provides the ionic resistance (R_i) of nanopore. The R_i is related to following equation:

$$R_i = \frac{4\rho_i L_n}{\pi d_n}$$

(equation 2)

where ρ_i is specific ionic resistance of the electrolyte, L_n is the length of nanopore and d_n is the diameter of nanopore. By changing the length of nanopore through the combination of ssDNA and DNA nanopore, the current signals flowing through the DNA nanopore can be altered. This is because conductance is proportional to width and inversely proportional to length of nanopore. To decrease the length of nanotube and increase the signal, we cut four helix of DNA nanotube and kept one helix as hinge and another as stem loop, making 6HB two halves. The miRNA target can be complementary to the stem loop located between the two halves. With the help of stem loop, we design a cork-shaped nanopore containing open-closed states with 3–4 cholesterol anchors to extend the current signals and improve the sensitive and accuracy. This DNA nanopore sensing system is applied in the detection of short strands and introduced in chapter 3. This method can help the detection of miRNA in serum samples and potentially offers a simple and non-invasive liquid biopsy for diagnosing cancer.

To optimize the insertion yield of DNA nanostructures into lipid bilayers, is necessary to utilize hydrophobic groups as lipid anchors. However, the cork nanopore design has a limited number of cholesterol binding sites, only have 3–4 cholesterol anchors. In order to overcome this limitation, we were inspired by Krishnan's work and developed a T pore design that allows for a large number of hydrophobic functionalization.[110] Krishnan have designed a DNA T pore membrane channel using a double-layered top plate, a ~ 4nm lumen and a 27nm long stem attached to the centre of the plate (**Figure 1-14a**). The plate of this T pore can be functionalized with up to 57 hydrophobic modifications and thus allowed for stronger anchoring of the pore to the membrane. Similarly, one of the T-shape nanopore we described in chapter 4 is using cork-like 6HB as stem and a 2D DNA platform to facilitate modification chances and interaction area. So, an engineered 60nm × 90nm 2D DNA origami, with a ~6nm × 6nm square hole in the center of DNA tile, combines a cork-shaped nanopore by four DNA overhangs is designed like T pore, namely corkT1. The 2D DNA origami was designed by Hao Yan's group [111] as shown in **Figure 1-14b**, that so-called T1. This T1 structure arrangement and the details will be described in chapter 4.

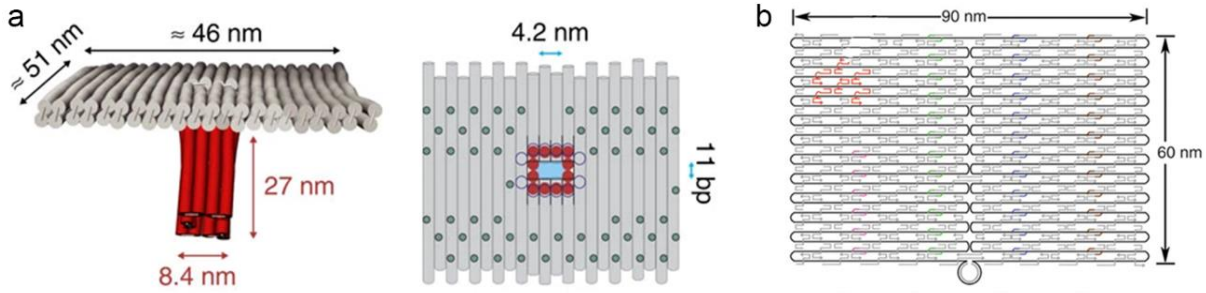


Figure 1-14. Design of the DNA T pore membrane channel.
(a) Left: side view of T pore, which is composed of a double-layered top plate and a 27nm long stem attached to the centre of the plate. The stem is formed by 12-helice bundle arranged in a square with 3.7×4.2 nm lumen. Right: top view of T pore. The plate provides a large area for membrane interactions. The green circles represent the distribution of the 57 possible positions for the introduction of membrane-interacting modifications (tocopherol or biotin). The image is reproduced from ref. [110]. (b) Schematic layout of the two-dimensional (2D) DNA tiles, which have dimensions of 90×60 nm. This simple, rectangular-shaped DNA origami tile was utilized, in which a circular single-stranded M13 viral DNA (black lines), composed of 7249 bases, is folded and stapled, with the help of >200 short synthetic DNA strands, to form the desired 2D tile. Helper strands (staples) are shown in grey or coloured lines. The image is reproduced from ref. [111].

Further experimental approaches (AFM and electrical methods) revealed the need for a higher yield of well-formed origami to characterize the T1 structure. To improve the efficiency of nanopore insertion and the stability of the DNA nanopore, we optimized the T pore design by using only one scaffold, a circular single-stranded M13 viral DNA, to form the stem and the plate. This newly designed structure, that so-called T2, is detailed in chapter 4. By increasing the number of cholesterols in the platform and enhancing the interconnections between the duplexes within the nanostructure, the efficiency of nanopore insertion can be improved and the stability of the DNA nanopore can be increased.

1.5 Conclusion and motivations

The incessant need for DNA nanostructure designs that facilitate a more profound comprehension of DNA nanopore structures, their compatibility in Droplet Interface Bilayers (DIBs), and their potential to enhance electrical recording testing in high-demand scenarios is a pressing concern. The development of nanopores is a prerequisite for indispensable tools in this regard. In the subsequent chapters 3 and 4, this these delves into the exploration of five models of DNA nanopore architectures and their characterization in lipid bilayers, followed by their potential for miRNA biosensing. All of these endeavors are geared towards enriching our current knowledge of DNA nanopores while simultaneously addressing the concerns regarding their sensor potential.

2 General Information on Characterization and Manipulation Methods

2.1 Introduction

The structure and function of pores are highly dependent on the properties of the lipid bilayers. Understanding the interaction between the two is critical in developing effective methods to modulate pore behaviour. Various experimental methods have been developed to assess this interaction, including droplet interface bilayers (DIBs), atomic force microscopy (AFM), and quartz crystal microbalance with dissipation monitoring (QCM-D). In this chapter, we will explore these experimental techniques and their protocols in detail.

Two computational tools, scaDNAno and oxDNA can be used to design, construct and simulate DNA structures for visualization, editing and analysis of simulation of DNAs before synthesis and practice. Additionally, Gel electrophoresis and Dynamic Light Scattering (DLS) experiments are used to characterize the size of output nanostructures.

Atomic force microscopy (AFM) is a high-resolution imaging technique for studying the structure and function of DNA nanostructures and lipid bilayers at the nanoscale level. AFM can be used to image the size and shapes of origami, and the surface of the bilayer at high resolution, allowing researchers to study their interaction with the surrounding membrane. The protocol for performing AFM involves the use of a cantilever with a sharp tip, which is used to scan the surface of the bilayer. By scanning a sharp tip over the surface of the membrane, researchers can create a detailed topographical map of the membrane and its pores. AFM can also be used to measure the mechanical properties of the membrane, from whose elasticity and stiffness, we can analyse AFM images in phase mode.

Quartz crystal microbalance with dissipation monitoring (QCM-D) is a highly sensitive technique for studying the interaction between lipid bilayers and pores. QCM-D involves the use of a quartz crystal that is coated with a thin layer of lipid membrane. As the membrane interacts with the pores, changes in the mass and viscoelastic properties of the membrane can be detected by monitoring the resonance frequency and dissipation of the quartz crystal. DIBs are a relatively new method for studying the properties of lipid bilayers. DIBs are formed by

the fusion of two water droplets, each containing a lipid monolayer, at an oil-water interface. The resulting bilayer is highly stable and can be used to study the properties of the lipid membrane, including its interaction with nanopores. The protocol for creating DIBs involves the use of fresh lipids in oil solutions and the careful manipulation of the droplets to ensure the formation of a stable bilayer. By measuring the electrical conductance of the bilayer, we can determine the properties of the pores within the bilayer. This method is particularly useful for studying ion channels.

In conclusion, the interaction between lipid bilayers and nanopores is a complex process that can be studied using a variety of experimental techniques. Droplet interface bilayers, fluorescent experiments, atomic force microscopy, and quartz crystal microbalance with dissipation monitoring are all powerful tools for studying this interaction. Each technique has its own unique strengths and weaknesses, and we can combine all methods for comprehensive analysis to obtain more objective analysis results. By utilizing these experimental methods and their respective protocols, researchers can gain a better understanding of the properties of lipid bilayers and the behaviour of pores within them.

Detection of short length of oligoes at low concentrations is a complex task. The experiments detailed in this chapter show that careful control and the use of multiple approaches are required to achieve the ultimate detection goal. The methods such as gel electrophoresis, DLS, AFM, DIBs and QCM-D, are mainly used for the design and characterization of nanopores, the adsorption of nanostructures on supported lipid bilayers, and the formation of DIBs in simple buffer systems. We will use them specifically in the next chapters.

2.2 Characterisation Methods

2.2.1 Scadnano software

Synthetic biology has grown tremendously with advances in DNA synthesis, sequencing, and editing technologies. In this field, designing and constructing complex DNA structures with high accuracy and efficiency is a necessary prerequisite for the synthesis of DNA structures. To address this challenge, a number of computational tools for designing and simulating DNA structures have been developed. We introduce scadnano (short for

"scriptable cadnano"), a computational tool similar to the most widely used cadnano software for designing synthetic DNA structures.[74] Compared to cadnano, scadnano offers several advantages over cadnano. Scadnano runs entirely in the browser, requiring no software to install. It can be manually edited, or it can also take advantage of Python scripting library to help automate tedious tasks. Its file format is human-readable, which can also help when debugging scripts or interfacing with other software.

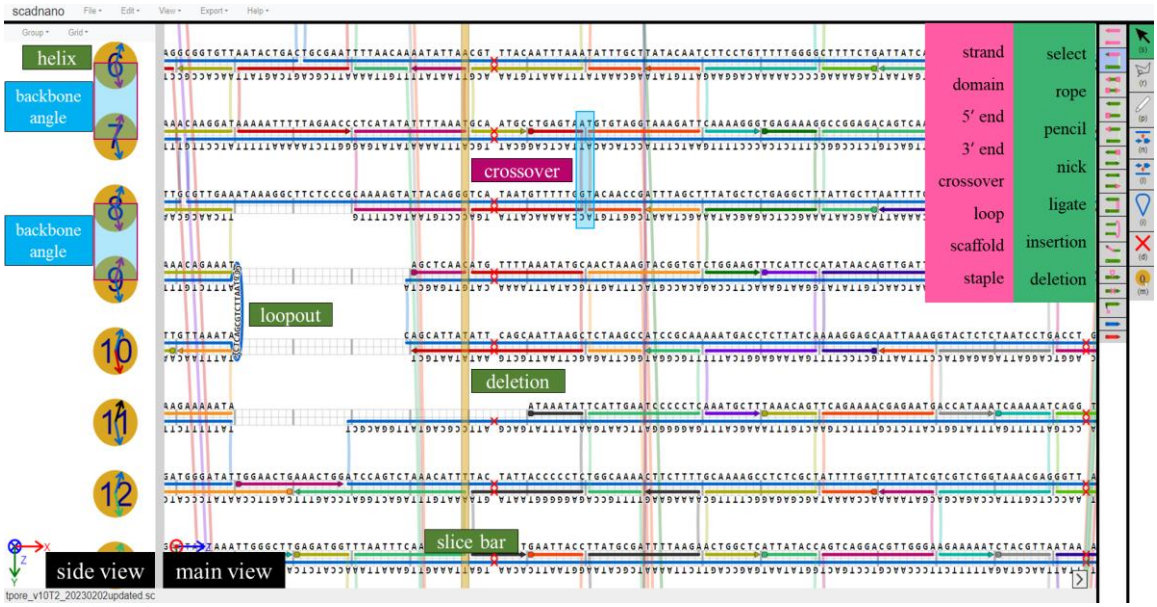


Figure 2-1. Screenshot of scadnano, showing most of the features.
The side of view shows the grid type, helix number and backbone angle. The main view shows the detailed sequences of DNA design. We can use the tool bar to edit DNAs.

Using scadnano, we can easily design 3D DNA nanostructures by defining grid type (square, honeycomb, hex, none lattice), DNA helices, domains, loops, crossovers and modifications (**Figure 2-1**). Most importantly, it does provide a primitive way to visualize the DNA backbone angles to help pick where to place crossovers. As blue highlighted in **Figure 2-1**, we use slice bar (yellow ruler) to pick one base position, and the side view displays the backbone angles at this base position of all helices to aid to choose the perfectly alignment of backbone angles. With its browser-based interface, visualization of the DNA backbone angles, accessibility of Python scripting library, and human-readable file format, scadnano offers researchers a powerful and collaborative tool for designing and simulating complex DNA structures.

2.2.2 oxDNA simulations

Before ordering DNA strands and to confirm the scadnano design, the 3D structure of new DNA nanopore designs can be checked by a simulation code called oxDNA.[75] This code is based on a coarse-grained force field, in which atoms are gathered into 'beads' with effective interactions that were parameterized to mimic as much as possible the thermodynamic properties of DNA and RNA. In the context of this work, the main limitation of this model is the absence of model interactions with lipid molecules. In its present state, oxDNA only provides with effective interactions between nucleotides, water and ions are only included in an implicit way. Despite these limitations, oxDNA simulations are useful in providing 3D views of the otherwise abstract representations of scadnano. These two tools help to ensure that the design will work as intended, and that there are less structural issues that could cause problems down the line.

A typical oxDNA simulation [112] includes three steps: (i) Minimization of the structure provided by scadnano. This step eliminates some initial steric clashes. Python scripts were used to generate initial configurations of the loop segments (set of unpaired nucleotides). (ii) Relaxation of the system. This step is often necessary before running a dynamics simulation in order to fix distances between nucleotides that are too far from equilibrium values. (iii) Simulation of dynamics. We used molecular dynamics to generate trajectories, instead of Monte Carlo methods [113]. The thermal fluctuation analysis of the structure is a movie to show the energetic implications of the design, which can be used to optimize its performance. Simulation time is on the order of 10^{-6}s.

2.2.3 Gel electrophoresis

Gel electrophoresis is a method for separation and analysis of various biomacromolecules (DNA, RNA, proteins, etc.) and their fragments, based on their size and charge, owing to low sample consumption, low cost, fast analysis, high throughput, and integration capability.[114] DNA gel electrophoresis has been the most important experimental tool to separate DNA fragments. Electrophoresis filled with free liquid solution can be suit to separate, analyse, identify, and purify DNA fragments by their charges, sizes, labeling, and shapes through agarose or polyacrylamide gels. On GE platform, small molecules (e.g. 6HB structures) move through the network freely in gel with a weak resistance, while

macromolecules (e.g. 2D DNA rectangular patterns) are retarded and migrate at slower rates so that we need to control the low temperature for the stability of macromolecules.[115]

Among them, agarose is a linear polymer composed of alternating residues of D-galactose and L-galactose.[116] Chains of agarose form helical fibres that aggregate into supercoiled structures with a radius of 20–30 nm. Gelation of agarose results in a three-dimensional (3D) mesh of channels which can be used for the separation of DNA fragments ranging from 20 base pairs to millions of bases. Besides agarose gel, native polyacrylamide gels can separate small DNA fragments (5–1000 base pairs) effectively with the advantages of the uniform pore size provided by the polyacrylamide gel.[117] Pore size is controlled by modulating the concentrations of acrylamide and bis-acrylamide (N,N'-methylenebisacrylamide) used in creating a gel. And the polymerization is initiated by the addition of ammonium persulfate (APS) along with N,N,N',N'-tetramethylethylendiamine (TEMED).

Generally, a final concentration of ~ 10nM, achieved by adding 20μL of folded constructs, is recommended for constructs that lack cholesterol modifications. The amount of scaffold and nanostructure loaded on a gel lane is critical. If the quantity is not enough, bands on the gel won't be visible. If there is too much number of oligoes, structures may collect in wells, smear during electrophoresis or get tangled creating false bands on the gel. Gels and buffer should contain appropriate magnesium solution.

This technique is simple, rapid to perform, and capable of resolving DNA fragments. The location of DNA bands within the gel can be determined directly by staining with low concentrations of fluorescent intercalating dyes, such as ethidium bromide or SYBR safe; bands containing as little as 20pg of double-stranded DNA can then be detected by direct examination of the gel in ultraviolet (UV) light. If necessary, these bands of DNA can be recovered from the gel. Overall, gel electrophoresis is an essential technique for analysing scaffold control lanes and nanostructures.

2.2.4 Dynamical Light Scattering experiments (DLS)

Dynamic light scattering (DLS) is an optical measurement technique that is utilized for the characterization of dispersed systems by evaluating high-frequency fluctuating scattered light which reflects the dynamics of microstructural processes. The most commonly used

application of DLS is for Brownian motion particle size analysis, in which it quantifies individual particles in liquids (**Figure 2-2**). DLS is capable of providing z-average hydrodynamic diameter, polydispersity index, and intensity weighted size distribution through cumulant analysis and distribution analysis. This analysis is reliant on the numerical conversion of spectral measurement signal values of representative samples into size distributions, and its main outcome is the size of individually moving entities, i.e., individual particles and particle aggregates or agglomerates, independent of the size of the constituent particles in the aggregate. Due to the fast analysis or low cost per measurement, DLS has become the most popular measurement technique for particle size analysis in the submicron range (1nm – 1μm).

The DLS measurements were conducted using a Malvern Zetasizer Instrument and its software, which allowed for simple and rapid analysis. Measurements were performed in the standard mode of acquisition using Eppendorf UVettes at 25°C. One of the limitations of the method is the fact that the liquid volume should be larger than 200μL. For nanoparticles dispersed in solvents, the DLS results are used to evaluate particle size and determine the presence or absence of agglomerates and aggregates.

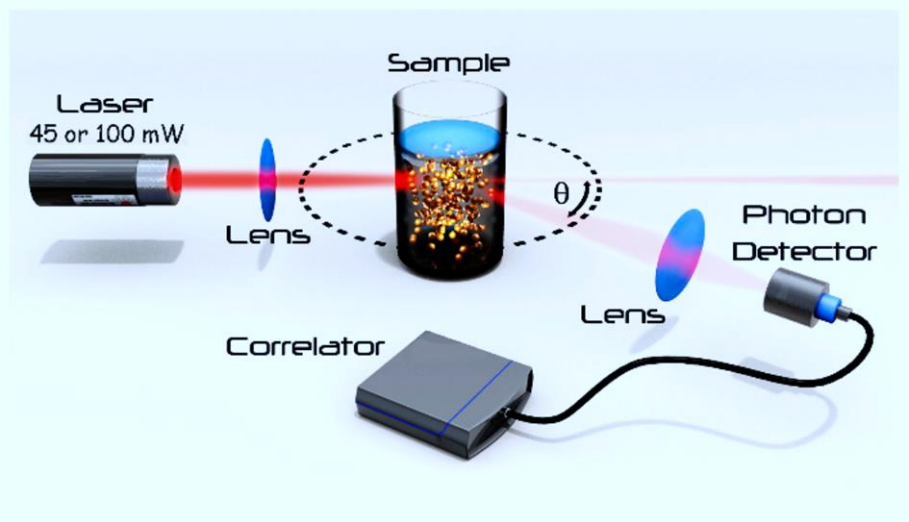


Figure 2-2. Depiction of the dynamic light scattering (DLS) apparatus with laser.
The sample is only schematically depicted to visualize the measurement principle. The image is reproduced from ref. [118]. Light scattered by a fine particle illuminated with a laser is measured with high time resolution under a defined angle θ; the fluctuation of the scattering signal reflects the dynamics of microscopic processes such as the particles' Brownian motion.

We have used DLS to determine the existence of aggregation when adding cholesterol moieties to DNA nanostructures. In general, DLS yields the hydrodynamic radius which is

usually different than the geometrical radius of gyration. Also, DLS measures the diffusion of moving objects which not only include (in our case) oligonucleotides but also the surrounding hydration layer, counterions and possibly surfactants. Our use of DLS has only been intended to discriminate between aggregated and dispersed phases, depending on the way cholesterol moieties are incorporated into DNA nanopores.

2.2.5 Atomic force microscopy experiments (AFM)

Atomic force microscopy (AFM) is a powerful technique which provides with surface characterisation of nanometer-to-micrometer-scale structures with nanometer resolution.[119] Therefore, AFM imaging is frequently used in a wide range of applications from surface physics to material characterization. It also provides unprecedented insights into the structure of biological macromolecules. Through the interpretation and modelling of high-resolution AFM images, we can infer the size and structure of the imaged structure.[120]

There are three basic imaging modes of AFM, which are contact mode, non-contact mode, and tapping mode. Contact AFM is a repulsive mode in which the probe tip makes "actual contact" with the sample softness and it is not suitable for studying biological macromolecules that may move or even deform. In non-contact mode, the detector detects long-range forces such as van der Waals forces and electrostatic forces that do not destroy the imaging sample, but it’s not suitable for imaging in liquids. Tapping mode in liquids has very little damage to the sample, and it’s suitable for soft, brittle and adhesive samples. The microcantilever is forced to vibrate near its resonant frequency, and the oscillating tip gently taps the surface, making intermittent contact with the sample. When the tip is not in contact with the sample, the microcantilever oscillates freely at maximum amplitude. When the two are in contact, the vibration amplitude of the cantilever decreases due to the space obstacle, at this time the feedback system controls the constant amplitude of the microcantilever. The tip thus moves up and down following the surface undulations to obtain topographical information.

In AFM, the sample is first deposited on top of an atomically flat surface, such as freshly cleaved mica. Then at a set of predefined locations, the probe approaches the sample up to a distance where some condition is satisfied. This condition depends on the type of AFM imaging (see below). By using a probe with a tip sharper than the object being observed, the

measured height profile will be related to the real shape of the sample (**Figure 2-3**). When imaging soft samples, such as DNA deposited on mica, the shape of the tip is prone to change due to adsorption of the DNA sample when scanning the sample, which will make the measurement reconstruction deteriorated or ineffective. As a result, we will observe that the clearer AFM images often appear when the AFM tip is a new one.

AFM experiments were performed on mica surfaces that bind nucleic acids with bridging magnesium ions. 30μL of samples was prepared at 0.5–1 μM of each DNA nanopore in 100μL 1×B buffer containing 1M KCl (pH 7.4), with 20mM $Mg(OAc)_2$ or TM buffer, pH 7.4. The analytes were incubated on freshly cleaved mica plates for 3−5 min. Imaging was performed in a liquid medium in order to approach as much as possible origami conformations in solution. We used both tapping mode and Quantitative mechanism in fluid (QNM) or tapping in fluid mode with a Bruker Dimension ICON AFM utilizing SNL-10 AFM probes. If necessary, the AFM images were processed by flattening to remove the background slope and adjusted for contrast and brightness. AFM experiments were performed to confirm the formation of DNA nanostructures.

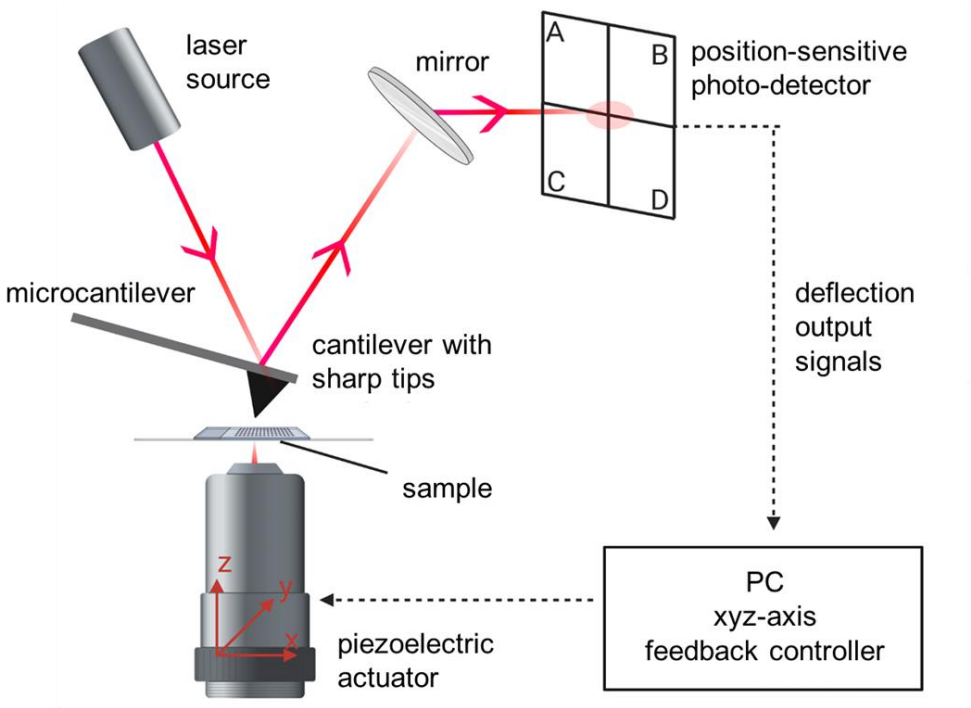


Figure 2-3. The scheme of how AFM works is depicted. The AFM imaging is based on a cantilever/tip assembly that interacts with the sample; this assembly is also commonly referred to as the probe.
The AFM probe interacts with the substrate through a raster scanning motion. The up/down and side to side motion of the AFM tip as it scans along the surface is monitored through a laser beam reflected off the cantilever. This reflected laser beam is tracked by a position-sensitive photo-detector (PSPD) that picks up the vertical and lateral motion of the probe.

Tapping mode yields height, amplitude and phase data of the sample surface, while Peak force mode (QNM) simultaneously generates height, adhesion, and modulus data.

Tapping mode in liquid: Tapping mode refers to a collection of AFM modes in which the cantilever oscillates at a frequency close to some resonance. Imaging can be obtained either by controlling the phase or the amplitude of the oscillation. In this thesis, we focused on monitoring amplitude. AFM images are obtained using an AFM imaging system (Bruker Inc.) with a SNL-10 silicon-tip on nitride cantilever (Vecco Inc.). Preliminary scanning is performed using ScanAsyst in air. After the AFM is setup, 5mL of 1×TM is added to the liquid cell and scanned under a tapping mode in liquid onto a freshly cleaved mica surface. An additional 500μL of 500pM DNA origami sample is injected to the fluid cell. The tapping amplitude set point was set to 9.4nm, the drive amplitude was 4500mV, and scanning rate was 0.8Hz. Good AFM images depends heavily on the quality of the AFM tip.

Quantitative mechanics standard in fluid mode (QNM) (also called Peak Force mode): In this mode, the height of the cantilever is monitored in such a way that its deflection (or equivalently, the force between the probe and the sample) is a given quantity, usually in the range 200–600 pN (**Figure 2-4**). AFM images were obtained on an AFM (Bruker Inc.) using a silicon nitride cantilever (SNL-10, Olympus). Samples are DNA nanostructure solutions prepared by the described method and need to be diluted with 1×TM buffer at room temperature. We then attempted to visualize nanostructure complexes by atomic force microscopy.

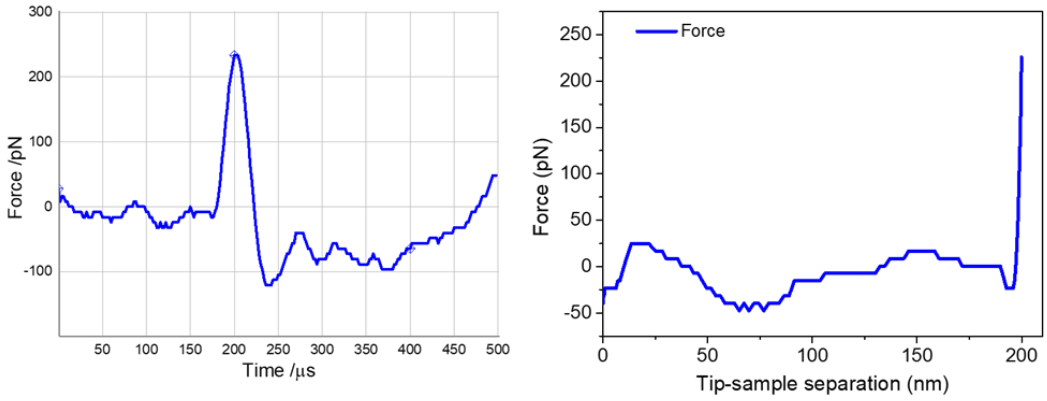


Figure 2-4. Peak force window and force monitor window in an AFM experiment.
(Left) The force is set at about 200–600 pN, depending on the sensitivity condition of SNL-10 cantilever. (Right) The image shown here is a force example of 200nm scan.

2.3 Manipulation Methods

2.3.1 Quartz Crystal Microbalance with Dissipation monitoring (QCM-D)

A quartz crystal microbalance (QCM) sensor is made of a thin disc of quartz crystal sandwiched between two electrodes based on inverse piezoelectric effect discovered by the Curies in the late 19th century.[121] As the core of QCM-D, quartz is a piezoelectric material, that is, a material in which mechanical deformation causes electrical charges in the material to move, and vice versa, where an applied electric field causes mechanical deformation. In fact, if you apply an AC voltage, the disk will oscillate back and forth in sync with the applied voltage.

QCM-D uses a so-called ring-down method, where the driving voltage is intermittently switched off and the decay in time of the oscillation is monitored. From the decay curve, the resonance frequency f and the energy dissipation D are extracted (**Figure 2-5**). The ubiquitous application of quartz crystals in oscillator circuits is based on their exceptional stability and very low energy dissipation. Depositing a film on the surface of a QCM crystal effectively increases the crystal thickness and therefore the wavelength of the standing wave. When the lipids formed a homogeneous rigid films and contacted with DNA nanostructures, it may either induce a small dissipation shift ($\Delta D \approx 0$) or a large one ($\Delta D > 0$).[121]

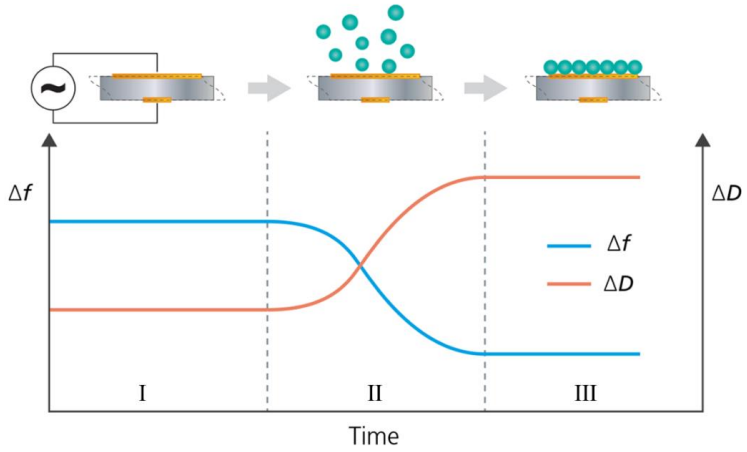


Figure 2-5. QCM-D provides information of frequency (f) and energy dissipation (D).
*The information from a single harmonic QCM-D can reveal whether the film is rigid or not, and if the Sauerbrey equation (**equation 3**) can be used for the quantification of mass. In the adsorption scenario and plot: (I) bare surface and stable baselines of Δf and ΔD; (II) molecules bind to the surface, and as a result, the frequency decrease and the dissipation increased, indicating mass uptake and increasing energy loss; (III) the surface uptake has been completed and the frequency and dissipation responses have stabilized. The image is adapted from https://www.biolinscientific.com.*

For a small dissipation shift ($\Delta D \approx 0$), the film can be approximated as rigid, and the Sauerbrey equation can be used to measure the area mass density of the film by the recording of oscillating frequency of a quartz crystal:

$$\boldsymbol{m_f = -C_{QCM}\frac{\Delta f_n}{n}}$$

(equation 3)

where m_f is the area mass density of the absorbed film, n is overtone number, Δf_n is resonance frequency changes observed at the overtone number n. Here, 1Hz deviation in the resonance frequency for a quartz oscillator with a 5MHz fundamental frequency corresponds to 17.7 ng cm^{-2} Hz^{-1} of C_{QCM} constant.

If $\Delta D > 0$, the film is sufficiently soft and sufficiently thick for the QCM to become sensitive to the mechanical or, more precisely, viscoelastic properties of the film. In this case, we treat the QCM-D response by the viscoelastic model based on the analysis of shear wave propagation in viscoelastic media:

$$\boldsymbol{\Delta f_n \approx -\frac{n}{C} m_f \left(1 - n\omega_F \rho_l \eta_l \frac{G_f'}{\rho_f\left(G_f'^2 + G_f''^2\right)}\right)}$$

$$\boldsymbol{\Delta D_n \approx \frac{2n}{C} m_f n \frac{\omega_F}{f_F} \rho_l \eta_l \frac{G_f'}{\rho_f\left(G_f'^2 + G_f''^2\right)}}$$

(equation 4)

Where m_f is the areal mass density of the film, f_F is the fundamental frequency, 5Hz here, $\omega_F = 2\pi f_F$ is the angular fundamental resonance frequency, η_l and ρ_l are the liquid's viscosity and density, ρ_f is the density of the film, the storage modulus G_f' describes material elasticity and the loss modulus G_f'' describes viscous energy dissipation in the material subjected to deformation, and C is the mass sensitivity constant depending on the material properties of the quartz crystal. Here, subscripts l and f refer to the liquid and the film, respectively. The **equation 4** is use to predict Δf and ΔD, yielding a well-defined equivalent thickness and z-averaged viscoelastic properties. Dedicated software for numerical fitting of QCM data in terms of film properties is available.

QCM-D can monitor the deposition of supported lipid bilayer to investigate binding events between molecules and lipid bilayer through frequency and dissipation signals from a piezoelectric quartz crystal oscillator sensor. The QCM-D crystal sensors were pre-coated with a 50nm thick Au film or an additional 30nm thick SiO_2 film on top of the gold, deposited by e-beam evaporation. Before use to decompose and remove organic contaminants on the

sensor, it was rinsed by pouring pure water and drying by nitrogen gas flow and then followed by UV/ozone treatment for 20 min. The QCM-D experiments were performed by injecting a solution of monodisperse lipid vesicles (DphPC: POPC = 7:3) in 1×B buffer, a solution of 1×BO buffer and solutions of different concentrations of nanopore samples. The measurements of the changes in frequency, Δf, and the changes in damping, ΔD (dissipation) were recorded as a function of time. The absorbed mass can be obtained by Sauerbrey equation (**equation 3**) and viscoelastic model described above. In this thesis, we have used QCM-D to have an independent method to test the interaction between nanopores and lipid bilayers (the scheme of lipid bilayer is shown in **Figure 2-6**. The idea is to prepare lipid vesicles that, upon interaction with the sensor surface will form a supported lipid bilayer (SLB). Once this is formed, nanopores are added and allowed to deposit.

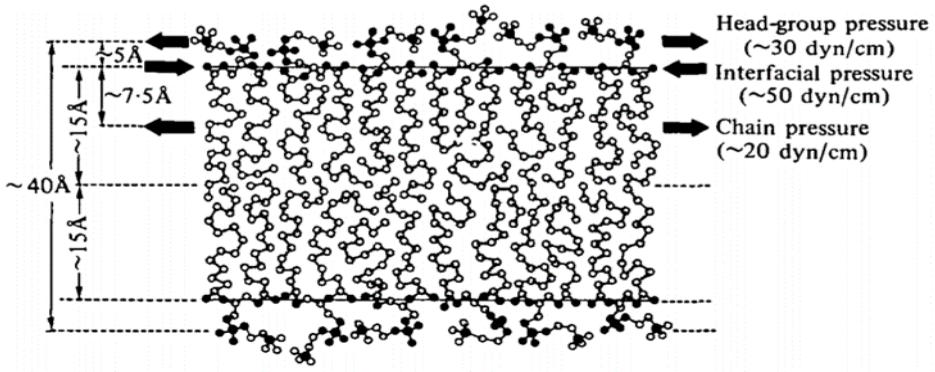


Figure 2-6. Scheme of lipid membrane.
The thickness of this lipid bilayer is ~ 4nm. And the image is reproduced from ref. [122].

2.3.2 Droplet Interface Bilayers (DIBs)

A crucial step in all the experiments of this thesis is the formation of a stable lipid bilayer. Several methods exist to perform this task, however most of them lack reproducibility and stability. The method called Droplet Interface Bilayer (DIB) is based on the simple fact that when bringing together two lipid monolayers, they spontaneously form a bilayer. In DIB, two water droplets are immersed in an oil bath containing lipids (mostly DphPC and POPC). When separated, each droplet is surrounded by a lipid monolayer. The position of each droplet is monitored by that of an electrode inserted into the droplet. For this, coating the tip of the electrode with an agarose gel is rather convenient. The bilayer formation takes place when the two droplets are brought together. There is no need of mechanical pressure for this formation to take place. The high stability of DIB bilayers comes from the fact that the lipid

reservoir is virtually infinite. This means that any fluctuation (pore formation) in the bilayer can automatically be 'repaired' by available lipids. This is to be compared to other methods, such as the Muntan-Muller method,[8] where the lipid reservoir is finite.

Under an applied electric field, as each molecule occupies the pore, a characteristic blockade of ionic current is produced. Information about length, composition, structure, and dynamic motion of the molecule can be deduced from modulations of the current blockade. Successful DIB formation requires a combination of parameters that I describe now.

Electrodes: these are silver filaments. In order to perform electric recordings, the composition of the tip of the electrode should be Ag/AgCl. This is simply obtained by soaking into sodium hypochlorite solution for 60 min. The tip then acquires a dark colour. We have tried 0.25mm and 0.125mm diameter silver wires (**Figure 2-7**), the thicker ones giving lower noise but being bulkier compared to water droplets. Silver electrodes are welded to copper wires, themselves being connected to ground and amplifier, respectively.

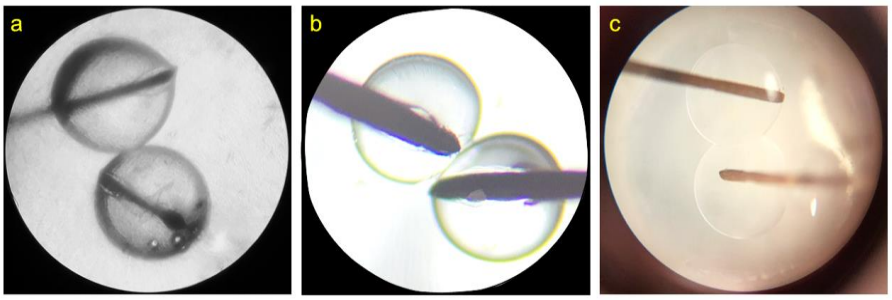


Figure 2-7. The microscope view of two droplets in DIB experiments.
Silver filaments of different thicknesses: (a) 0.125mm and (b, c) 0.25mm; and droplets of different sizes: (b) 0.3μL and (c) 3μL in DIBs experiments.

The Patchmaster and Intan Clamp software are versatile tools that can be as a driving program for amplifiers and data acquisition in electrophysiological experiments (**Figure 2-8**). They allow to monitor the currently active trace data displayed versus the time and measure amplitudes and durations.

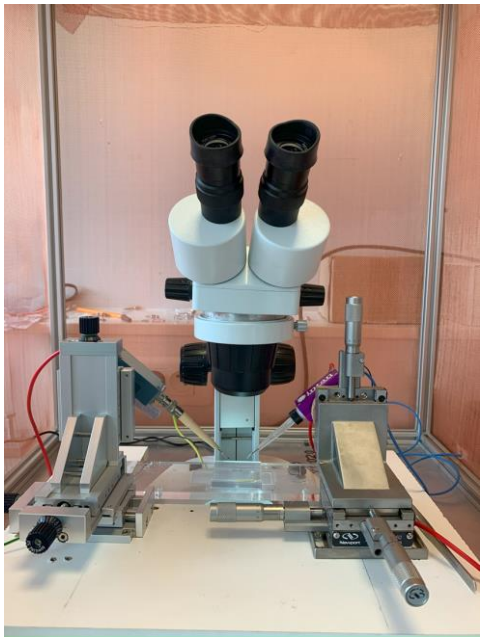


Figure 2-8. Schematic of patch clamp setup including Faraday cage, electrodes, amplifier, microscope, xyz positioners.

During a current recording experiment, a recording electrode with an electrolyte solution connected to an amplifier is brought into contact with the membrane of an isolated cell. Another electrode is placed in oil bath surrounding the cell as a reference ground electrode. An electrical circuit can be formed between the recording and reference electrode with the cell of interest in between.

The overall current recording procedure has four main steps: (i) start the configuration of amplifier, electrodes and software, and reset all the parameters. (ii) Next, the pulse test (10mV, 50Hz) is always applied to the pipette whenever we activate the amplifier. The purpose of pulse test is to provide the holding potential and applied to the pipette. When the pulse test runs, the current trace simultaneously displays, and then we can adjust the V_0 value to make the current baseline near to zero. The resistance of the substrate and the resistance and capacitance of the membrane are calculated from the current responses and displayed in the $R_{substrate}$, $R_{membrane}$ and $C_{membrane}$ fields. (iii) Set up a voltage stimulation pattern (negative potential, 10s and positive potential, 30s. and the voltage can be set to ± 30mV or ± 50mV in my thesis) and create a new file for data acquisition, note to write the sample name, oil, electrodes, sample concentration, voltage, date and time. Connect two electrodes putting in 300nL of two droplets in the 60μL of oil bath to the amplifier. To start data acquisition directly, the current responses are sampled and displayed. (iv) Export the data and parameters, and analyse the results by the python codes in Appendix A. code. The analysis allows us to display and calculate data that are based on the acquired Traces, thus giving us a fast overview over the results.

3 Characterization and Manipulation of cork nanopore

3.1 Introduction

Several studies suggest strong correlation between different types of cancer and the relative concentration of short circulating RNA sequences (miRNA). Because of short length and low concentration, miRNA detection is not easy. Standard methods such as RT-PCR require both the standard PCR amplification step and a preliminary additional step of reverse transcription.

In this chapter, we investigate the use of DNA nanopores as a tool to detect short oligonucleotide sequences at the single molecule level. These nanostructures show two different conformations depending on the presence of DNA analogues of miRNA sequences. By monitoring current across a lipid bilayer, we show that this change of conformation translates to different levels of conductance.

A ‘cork’ like DNA nanopore is synthesized: Upon cholesterol modification, this structure can interact with lipid bilayers, punching transient holes. The insertion of a single nanopore translates into a sharp conductance increase, which can be measured by electric recordings. Upon interaction with short DNA sequences, the nanopore changes conformation (‘closed’ to ‘open’ state), which increases the conductance. This change is significant enough to be detected.

3.2 Paper <Detection of Short DNA Sequences with DNA Nanopores>

3.2.1 Introduction

Micro-RNAs (miRNA) were first discovered in the nematode *C.elegans*,[123] and subsequently found in practically all eukaryotes. Despite their small size (length of the sequence between 19 and 24 nucleotides), these single strand, non-coding RNAs play an important role in the regulation of the genetic expression, through their capacity to hybridize with the 3′UTR of specific target mRNA (messenger RNA). Therefore, the specific function of each miRNA is strongly sequence-dependent, and the search of the associated targets is a

non-trivial problem. As shown in a recent review,[124] miRNAs are known to be associated with the normal development and function of the organism, but are also involved in diseases. Several lines of evidence highlight the putative implication of miRNA in the physiopathology of numerous cancers. This includes the miRNA dysregulations observed in tumors (over- or under-expression), that led to the concept of "oncomiR",[125] and to the use of miRNA for the classification of tumors origins.[126]

The attractiveness of extracellular miRNA as cancer biomarkers relies on their stability and their dysregulation in the diseased cells. However, because of their short sequence and low concentration, miRNA detection is intrinsically difficult. A large number of measurement methods have been explored, including PCR based methods, PCR free methods,[127] next-generation sequencing-based methods. Other less standard methods can be found in a recent review.[98] This work is based on a completely different approach, which in other contexts has been called 'stochastic detection'.[128] In this approach, two liquid media are separated by a membrane (lipid bilayer) in which channels such as transmembrane proteins can insert. Stochastic detection is based on the correlation between the presence of some analyte (here, miRNA) and the current across a single channel. A good example of stochastic detection is the method developed [129] for DNA sequencing: a single stranded DNA going through a single nanopore channel modulates the current through it which can be correlated with the DNA sequence.

In the past few years, DNA-based nanostructures [49,68,130] have been developed that mimic naturally occurring membrane proteins.[56,57,131] These nanostructures can also interact with lipid membranes, forming DNA nanopores. As compared to protein channels, they can be easily modified in terms of geometry or functionalization and are therefore good alternatives for stochastic detection. In ref. [109] we considered a DNA nanopore and showed that a 30nt long input sequence induced a measurable conductance change. Unfortunately, this strategy could not be applied to 22nt long sequences. In this paper, we modify this design to detect miR-21, a 22-nucleotide long miRNA involved in cancer (thyroid, breast and colorectal).[132] To facilitate sample preparation and handling, we will use the DNA analogue of miR-21, with sequence: TAGCTTATCAGACTGATGTTGA.

3.2.2 Methods and Results

Our goal is to obtain a nanotube composed of two barrels linked by a hinge and a locking mechanism (cf. **Figure 3-1**). The design of each barrel is inspired from that in ref. [133]. To design the locking mechanism, we take inspiration from ref. [134]. In this work, the goal was to trigger the opening and closing of a 3D origami box. The authors used two locks, each including a stem-loop structure with an 8nt loop. Upon addition of the opening key, an oligonucleotide complementary to a subset of the lock, the lock opened. We will use a similar mechanism here, as illustrated in **Figure 3-1**.

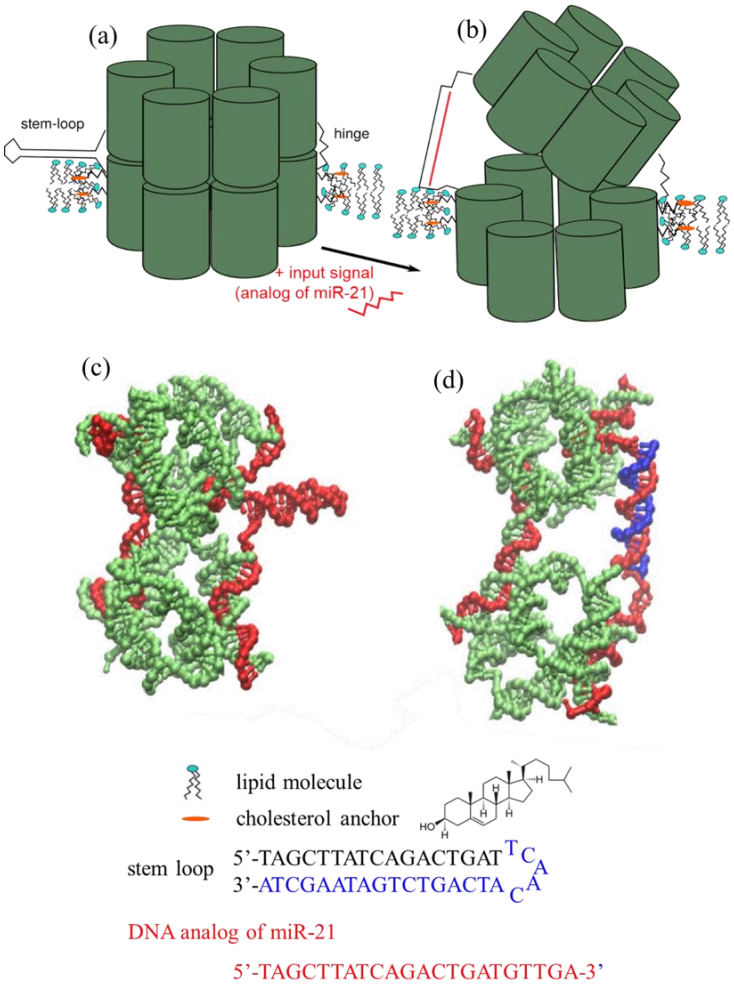


Figure 3-1.* (a) and (b): schematic representation of the opening mechanism.** ***Each cylinder represents a double helix. DNA nanopore is inserted into a lipid bilayer thanks to cholesteryl modifications (orange ellipses).
(a) Closed state: the stem loop imposes a short distance between two of the helices. (b) Open state: upon addition of miR-21, the stem loop unfolds giving rise to a mixed single and double stranded linker which pushes the two halves apart. (c) and (d) are oxDNA simulations of closed and open conformations, respectively. Strands that form the hinge or the stem-loop locking mechanism are in red. The input signal is in blue. Sequences of stem-loop and DNA analogue of miR-21 are also reported.

The nanopore is composed of two barrels. Each barrel is formed of six double helices arranged around a 2nm lumen with hexagonal cross-section. The staple design contains Holliday junction crossovers to increase the stability of the ensemble.[135] The stem of the locking mechanism is a 17nt double helix, the loop is 5nt long. Detailed description of sequences and binding topology (**Figure 3-6**) are given in the SI. Experimental protocols are given below. In the closed state, the stem-loop effectively imposes a short distance (∼2 nm) between two of the helices. The transition to the open state requires an input signal which can bind to the exposed nucleotides of the loop and open the stem through a strand displacement process. The stem loop becomes a long single stranded loop with a 22nt long double stranded section. This acts as an entropic spring, effectively increasing the distance between the two previously close helices. The length of the loop was optimized by computing (using Nupack[136]) the percentage of open structures. This percentage is maximum for loop length between 5 and 10 nucleotides. Note that a 22nt loop complementary to miR-21 yields a very unstable stem-loop structure. Here, we used the shortest of these loops (5nt), which ensures a high stability to the closed form, at the expense of sensitivity. Note also that miR-21 has some secondary structure: Nupack takes this fact into account in the computation of equilibrium distributions. Only pseudo-knot structures are excluded. However, these are very unlikely for 22nt strands.

In **Figure 3-1** are shown the results of two oxDNA [137] simulations, where the nanopore was simulated in closed and open states. Nanopore was assumed to be in solution, the cholesteryl modifications or any interaction with lipids were not taken into account. As shown in **Figure 3-1**, thermal fluctuations push the nanopore configuration far from the ideal hexagonal arrangement. Still, these simulations give support to the idea that the opening of the stem loop can significantly perturb the geometry of a nanopore even for short input signals such as miR-21. Also, the simulations suggest that the stem loop structure when attached to the nanopore is reasonably stable against thermal fluctuations. AFM images (**Figure 3-7**) confirm that the size of nanopores is close to the theoretical 14nm × 8nm size. The same picture shows that some misformed structures cannot be excluded. **Figure 3-8** shows that nanopores at higher concentrations (1μM) have some tendency to stack, forming filaments including more than one nanopore. Note that stacking could be enhanced by the interaction with the surface. PAGE electrophoresis experiments (**Figure 3-9**) further confirm the existence of a well-defined population migrating as a stable band.

To ascertain the opening mechanism and provide a proof of the detectability of the associated shape modification, we first used a fluorophore–quencher couple to monitor the FRET efficiency in the presence of the signal oligonucleotide. Two of the nanopore strands were modified with respectively Cy5 and BHQ2. In the absence of input signal, the nanopore should be in the closed state, in which BHQ2 quenches the fluorescence emission of Cy5. Initially, the three recordings in **Figure 3-2** show the same relaxation trend towards a closed state indicating that the system is sensitive to factors such as temperature gradients. Note that absorption by the walls could also contribute to this decrease. Adding miR-21 increased drastically the fluorescence recorded at 660 nm. The same illustration shows that the addition of two random sequences, 32nt long, had no effect on the nanopore opening. We conclude from this first experiment that, on the average, the nanopore behaves as expected when interacting with miR-21 strands.

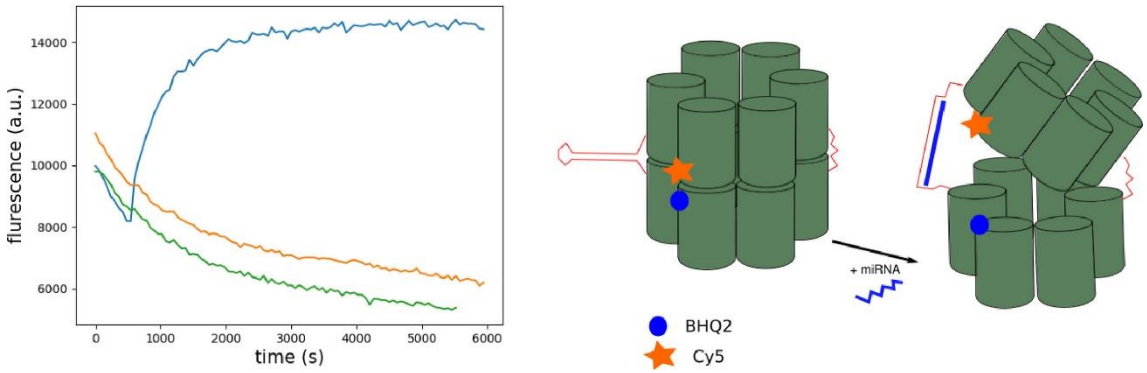


Figure 3-2. Left: Fluorescence (660 nm) recorded as a function of time for three different samples, all of them containing 100 nM solution of nanopore, modified with a Cy5–BHQ2 couple.
Before t= 600s, each sample only contained the nanopore. At t= 600s, an input signal was added. Blue: addition of miR-21. Orange and green: addition of two random sequences, 32nt long. Right: schematic representation of the quenched to fluorescent state transition.

DNA nanopore structures can interact with lipid bilayers when modified with hydrophobic moieties. In the pioneering work of Simmel and coll.,[57] the authors showed by TEM imaging how a large origami structure, modified with cholesteryl (cholesterol is attached to the deoxyribose via a six carbon spacer), was able to insert into a lipid vesicle. Subsequently, several teams also showed how similar structures could interact with locally planar bilayers by recording the current across bilayers. As compared to experiments with vesicles, the planar configuration offers the possibility to detect the insertion of single nanopores with electric recordings. For the planar configuration two main options can be distinguished. The formation of a black lipid layer has been used in refs. [53,133]. In this

configuration, two compartments are separated by a hydrophobic wall with a tiny hole. Painting lipids around the hole, then hydrating the system leads to the formation of a lipid bilayer. Alternatively, the so-called droplet interface bilayer (DIB) [138] configuration (**Figure 3-3**) deals with two aqueous droplets immersed in an oil bath containing lipids in solution. When the two droplets are not in contact, a lipid monolayer forms around each droplet. The position of each droplet can be monitored through electrodes connected to micromanipulators. If the two droplets are brought into contact, a lipid bilayer quickly forms at the intersection. As previously noticed,[139] the stability of this interface is remarkable, although it strongly depends on lipid and oil composition. We used previously a 'patch-clamp' approach [109] to obtain a bilayer, patching small pieces of giant unilamellar vesicles with a micropipette. As compared to this latter method, the use of DIBs is a much more robust approach with the disadvantage that not all lipid compositions can be explored.

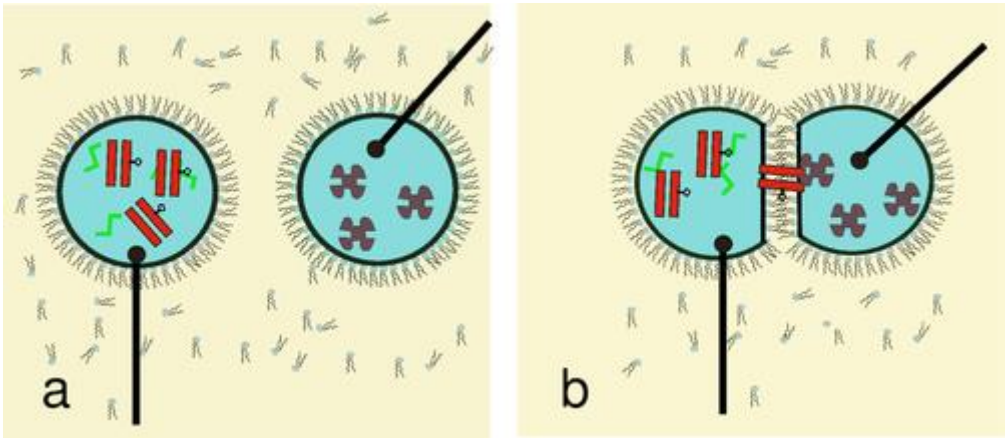


Figure 3-3. Illustration of the droplet interface bilayer (DIB) method.
Two aqueous droplets are immersed in an oil bath (yellow) containing lipids in solution. Droplet position is monitored through electrodes (black segments). (a) Before contact, a monolayer form at the surface of each droplet. (b) After contact a bilayer forms. Nanopores are contained in only one of these droplets (red rectangles). The other droplet contains streptavidin (brown).

We considered two approaches to enhance nanopore insertion into bilayer. Four strands were elongated with a common sequence, 15nt long, to which a complementary oligonucleotide modified with cholesteryl could bind. Experimentally, the insertion frequency of these structures into DIBs was very low. The second, more successful method, enhanced the insertion by adding two biotin modifications in one side (cf. **Figure 3-6**) of the nanopore in addition to the four cholesteryl modifications. In this second strategy, DIB system was asymmetric. One of the droplets, connected to ground, contained the nanopore. The other

droplet, connected to the probe (working) electrode, contained a solution of streptavidin. The biotin-streptavidin interaction is a classical biological tool to bind two partners. Our intuition was that in the event of a nanopore insertion, biotins would bind to streptavidin, thus maintaining nanopores in close proximity of the bilayer. We also hypothesized that the eventual transport of streptavidin across the DIB bilayer would be negligible.

A typical experiment started by the insertion of agarose coated electrodes on each of the droplets. Then, the droplets were brought into contact by monitoring electrode position. A lipid bilayer has a well-defined capacitive response to a short (10ms) 10 mV pulse. Its formation could be easily monitored, usually it took less than one minute after droplet contact. After stabilization of the bilayer's resistance, we imposed potential cycles (50s long) of alternatively negative (−30mV, 9s duration) and positive (30mV, 39s duration) potentials. During the positive potential phase, nanopores were expected to be driven towards the bilayer. Correspondingly, a negative potential would tend to remove them. **Figure 3-4** illustrates two typical situations we encountered. Current time recordings showed step-like profiles with an essentially stable baseline. The time between bilayer stabilization and appearance of current jumps was usually short (shorter than two or three potential cycles, i. e., two or three minutes). Most of the experiments actually displayed jumps from the first potential cycle. Each recording was stopped after 50 potential cycles (2500s). A few longer experiments were also carried out, showing that membrane stability and jump patterns could last longer. We observed several characteristic time intervals between jumps, with no evident link with experimental conditions. Fast transitions (jump frequency around 100 Hz) were much more frequent than the slower transitions (jump frequency around 10 Hz) displayed in panels **Figure 3-4a** and **Figure 3-4b**. The average success ratio, defined as the number of current recordings where jumps could be observed divided by the total number of current recordings, was rather low (less than 10 %, with a total number of 18 successful recordings). Unsuccessful recordings gave usually a flat signal (no jumps), or the interface was unstable leading to data difficult to interpret (with a vast majority of flat signals). In the absence of nanopores, no jumps were observed at all.

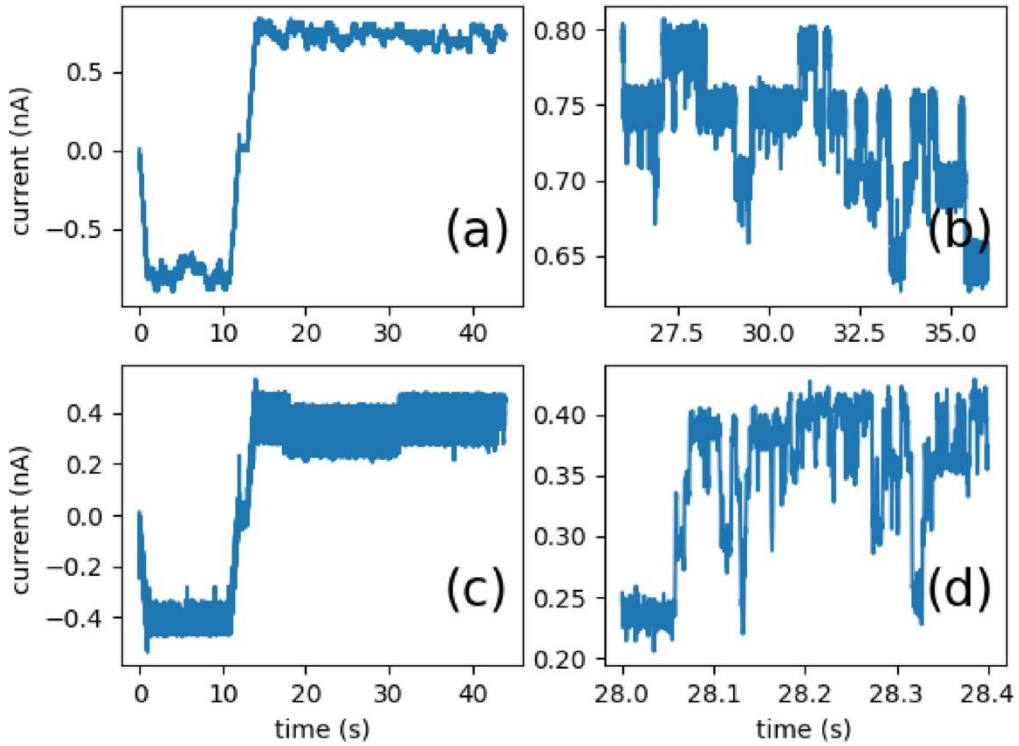


Figure 3-4**. **Time recording for a typical DIB experiment where voltage is kept constant at −30 mV (time interval [1, 10] s) and 30 mV (time interval [11, 50] s).
(a)(b) Slow dynamics regime (c)(d) Fast dynamics regime. (b) and (d) are zoomed images of (a) and (c), respectively.

To count the number and measure conductance jumps, we used a Hidden Markov Model (HMM).[140] Given a time recording, HMM approximates it by a sequence of N_s states, the values of which are optimized to minimize the difference between the sequence and the given time series. The number N_s is a free parameter of the model. HMM is well suited to approximate time series with step-like patterns, as those shown in **Figure 3-4**. An example of the HMM approximation is given in **Figure 3-10**. To avoid slow trends, conductance time series were divided into shorter intervals (0.1 s for fast dynamics recordings, 1 s for slow dynamic regime) on which HMM was applied.

A possible interpretation of the current recordings is as follows. Each time a nanopore inserts into the bilayer, its conductance increases by a fixed amount which depends mainly on the geometry of the nanopore. A simple estimate of the nanopore's conductance in its closed state using a geometrical model which ignores possible interactions between cations and nanopore's interior yields 1.3nS. In the absence of miR-21, we found conductance distribution centered around this value with a secondary peak close to 1.6nS, as shown in the histogram of **Figure 3-5**. Previous reports of similar structures [54] also yield values close to 1.6nS. As

proven by numerical simulations,[141] transport across DNA nanopores is not only through its lumen, cations can flow along the nanopore's outer surface or through the gaps in the DNA structure.[141] This would explain the fact that experimental values can be larger than theoretical ones. When the nanopore was incubated with the input signal, the stem-loop changed conformation as explained above and demonstrated by the coarse-grained simulations. Experimentally, this translated to the appearance of a second maximum in the distribution of conductances. Its value (2.6±0.2 nS) is less than twice the value of the closed state. This should be expected, as the open stem-loop pushes apart the two halves and at the same time hinders the entrance of the nanopore. Additional data (**Figure 3-11**) show that the existence of the high-conductance peak in the presence of miR-21 is a robust observation. They also show the existence of high conductance states with values larger than 2.6nS.

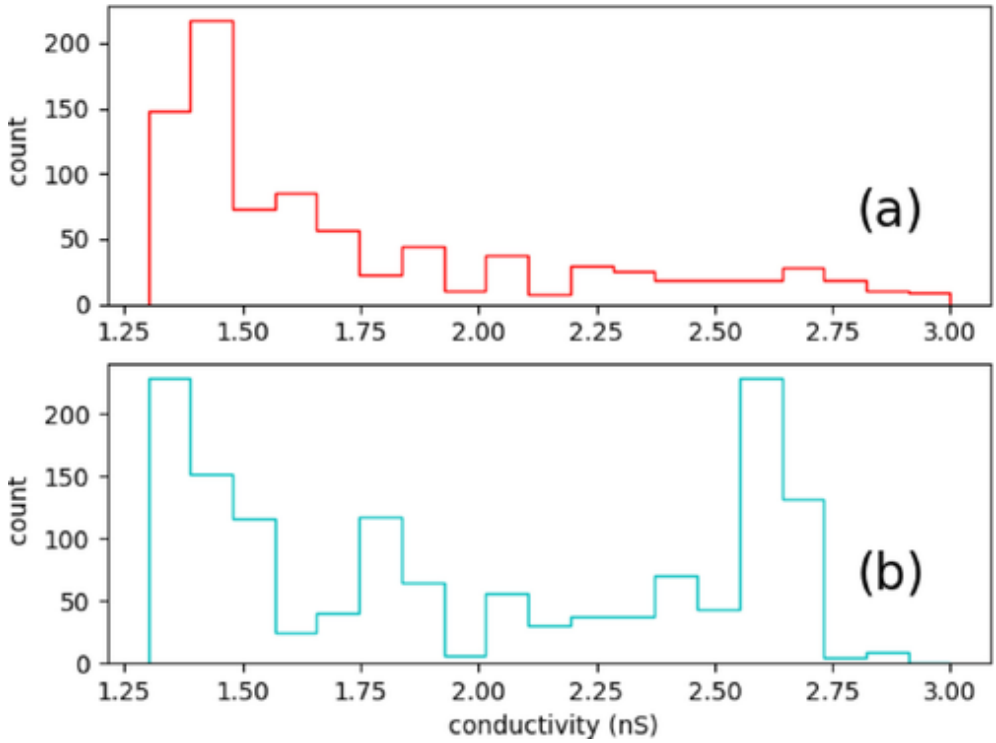


Figure 3-5. Conductance histograms obtained from two different current recordings, respectively. (a)[miR-21] = 0nM (N=915), (b) [miR-21] = 50nM. (N=1400).

From the present experiments, it is difficult to elucidate further the insertion mechanism of nanopores. A possible interpretation of the existence of transient states as those shown in **Figure 3-4** could be as follows: nanopores lay on one side of the bilayer inserting roughly half of the cholesterol modified strands. This metastable state has an energetic penalty due to the exposure of cholesterol to water. An alternative metastable state corresponds to a completely inserted nanopore where all the cholesterol moieties are in contact with the

bilayer's interior and, at the same time, the hydrophilic outer surface of the nanopore is also in contact with it unless a toroidal rearrangement of lipid heads (as sketched in **Figure 3-1**) takes place. Interaction with streptavidin probably lowers the energetic barrier between these two metastable states, which would explain the fast dynamics observed in many recordings. Slow insertion rates could then correspond to insertion in the absence of streptavidin. Note that an alternative interpretation would hold (streptavidin bound nanopores would correspond to the slow dynamic regime) if nanopores remained inserted inside bilayers, which seems unlikely.

3.2.3 Conclusions

In conclusion, sensing of short oligonucleotide sequences is potentially an important step in the early detection of diseases such as cancer. Developing portable, direct methods to perform such detection could considerably generalize the use of miRNA biomarkers. Compared to other single molecule detection procedures, nanopore based detection can benefit from miniaturization techniques used in semiconductor technology, which should provide eventually a compact, easy to use apparatus. In this report, we characterized a DNA nanopore structure able to change conformation upon binding with a DNA analogue of the miR-21 miRNA. The conformational change was characterized by fluorescence and electric recordings. In doing so, we have shown that detection of single miRNAs is a doable task when using the DIB configuration to generate stable and reproducible bilayers. The analysis of miRNA extracted from biological samples will require a preliminary purification step. This task is a common requirement to any method of miRNA quantification.[142] The major difficulty which remains to be solved is the low rate of insertion into bilayers. The possibility to detect low concentrations of miRNA depends on the feasibility of long electric recordings: the lower the miRNA concentration, the lower the number of possible open events. Reliable miRNA concentration measurements will therefore not only require parallel measurements but also a reasonable success rate in the detection of nanopores. This seems to be a major hurdle in the design of DNA based nanopores. A possibility explored by other groups [110] is to increase the number of hydrophobic moieties attached to the nanopore. This is only possible by embedding the nanopore structure into a larger platform.

3.2.4 Experimental Section

Fabrication of DNA Nanopores

DNA nanopores were fabricated in a one-pot reaction by stepwise cooling an equimolar mixture of staples (1 μM) in folding buffer (Tris-acetate-EDTA buffer, 20 mM $Mg(O\ Ac)_2$) from 90 to 20 °C in 3 h. Staple strands were designed using the scaDNAno [74] software. Before running DIB experiments, DNA nanopores were further diluted in a 1 M KCl, 10 mM $Mg(O\ Ac)_2$, 10 mM Hepes buffer containing 0.25 % OPOE (Poly(ethylene glycol) octyl ether) (Sigma). Cholesteryl functionalized DNA pores were produced by incubating the fully folded pores with cholesteryl modified strands (Eurogentec) for 45 min with 5 times excess. Before incubation, cholesteryl-modified oligonucleotides were heated to 60 °C for 45 min to avoid aggregation. Streptavidin (Sigma-Aldrich) was used without any further purification.

Lipid Preparation

POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) and DPhPC (diphytanoyl phosphatidylcholine) were purchased from Avanti Polar Lipids, hexadecane and silicon oil AR20 from Sigma. Lipids were stocked in chloroform with 10 mg/ml concentration. Before dissolution into a 7 : 3 hexadecane : silicon oil mixture, chloroform was evaporated in a vacuum dissicator for at least 1 h.[139] Dissolution of lipids into oil could require a mechanical stirring.

Electric Recording

Two 200nL droplets were deposited in a 60 μL well machined in poly(methyl methacrylate)(PMMA). The tip of two silver electrodes 100 μm in diameter were chlorinated overnight, then coated with agarose (2 % w/v) in 1×B buffer at 90°C. The agarose coating facilitated the insertion of electrodes inside the droplets. Electrodes were actuated through micromanipulators and connected to an electronic current amplifier (HEKA and Intan). Data were acquired at 5 kHz.

3.2.5 Supporting information

oxDNA simulations

Initial conditions for the open and close structures were obtained from the scadnano file, using the cadnano interface.py utility from oxDNA. These initial conditions were energy minimized, to avoid steric clashes and correct bond lengths. Once minimized, the structure was modified to become close to the expected stable configuration. For the closed form, this required an

additional distance restraining potential between nucleotides included in the stem-loop section. For the open form, the restraining potential was applied between nucleotides forming a double helix with the miR-21 strand. Once close to the target configurations (open and closed nanopore), the restraining potential was removed. The structures shown in the paper (**Figure 3-1**) correspond to the final configurations of trajectories 107 steps long.

Strand sequences

Table 3-1. Sequences in this paper given from 5' to 3'.

1	GCGGGGAGCGTATTAGAGTTG
2	TGTTCCAAATAGCCAAGCGGT-biotin
3	AGTGAGATGTCGTGACGTGGA
4	ATCGGCATTAAAGACCAGCTGCATTAATTTTTTCTCCTTCAC
5	CAACAGCATCCTGTTTCCGAA
6	TCCACTAAAATCCCCCCAGCAGGCGAAATGATTGCTTTCACC
7	TCCACGTTCTTTAATAGTGGA CTCTTGTTCCAAACTGGAACA
8a	GGCTATTCTTTTGATTTATAA
8b	GGGATTTTGCCGATTTCGGAA
9a	ACAGGATTTTCGCCTGCTGGG
9b	GCAAACCAGCGTGGACCGCTT-biotin
10	CAACTCTCTCAGGGCCAGGCGTAGCTTATCAGACTGATTCAACATCAGTCTGATAAGCTAGTGAAGGGCAATCAGCTGTTG
11a	TCTCACTGGTGAAAAGAAAAA
11b	CCACCCTGGCGCCCAATACGC-biotin
12a	TCCCCGCGCGTTGGCCGATTC
12b	ATTAATGCAGCTGGCACGACA
13i	TGTAATTGGTGATTTTTTTGTAATTGGTGATTTTTTTCCACGCTCCCTGAGGGGCGCC
13f	AGGGTGGGAATCGGACAAGAGTTTTTAGTTGGAGTTGATATTTTTTAGTTGGAGTTGATA
14	CGCCTGGGGTTTGCTTATAAATCAAAAGGTTTGGACCAACGC
magchol	cholesteryl-AATCACCAATTACA
tagchol	cholesteryl-TATCAACTCCAACT
miR-21	TAGCTTATCAGACTGATGTTGA
Random1	CAAGCCCAATAGGAACCCATGTACAAACAGTT
Random2	AATGCCCCGTAACAGTGCCCGTATCTCCCTCA

*Text in red corresponds to the stem-loop section. Text in blue corresponds to the cholesteryl binding sections. Random sequences correspond to the sequences used to check the specificity of opening mechanism in the fluorescence experiments (**Figure 3-2**).*

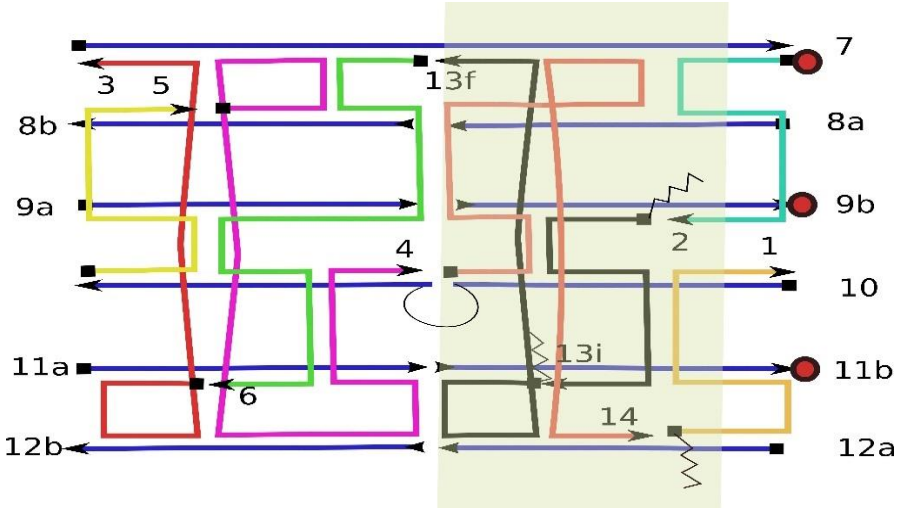


Figure 3-6: Schematic representation of the spatial arrangement of the 19 staples that form the nanopore.

Biotins are represented as red circles. Zigzag lines correspond to sequences where cholesterol may attach. The yellow rectangle corresponds to the zone where nanopore should interface with the lipid bilayer. Different colors were used to strengthen the difference between strands that, for each helix, play the role of scaffold (in blue) and strands that ensure the cohesion between different helices.

AFM imaging

AFM images (tapping mode for **Figure 3-7**, peak force mode for **Figure 3-8**) were made on nanopore samples in 20mM Mg^{2+}, 1×TAE buffer, deposited on freshly cleaved mica**. Figure 3-7** was made with 200nM nanopore sample, showing a heterogeneous population of well separated structures. **Figure 3-8** was made with a 1000nM sample.

In the case of **Figure 3-7**, to detect structures that could correspond to DNA nanopores, the following image analysis was performed:

- Isolate connected clusters of points with heights between 0.5nm and 3nm.
- For each cluster, compute the geometrical center, the principal and the secondary axis (weighted with the local height).
- For each cluster, determine the height profile along each principal axis. A height threshold h_t is required to define the border of each cluster. We considered the value h_t =0.1 (h_{max}−h_{min}). The length (respectively width) of the cluster is defined as the width of the height profile along the principal axis at h = h_t (respectively secondary axis). This provides with an estimate of the area of the cluster. A selection according to the cluster area allows to discard both small structures and potential nanopore aggregates.

Note that modelling AFM images of DNA nanopores as rectangles is an oversimplification of the real nanopore structure, as the stem-loop is not included.

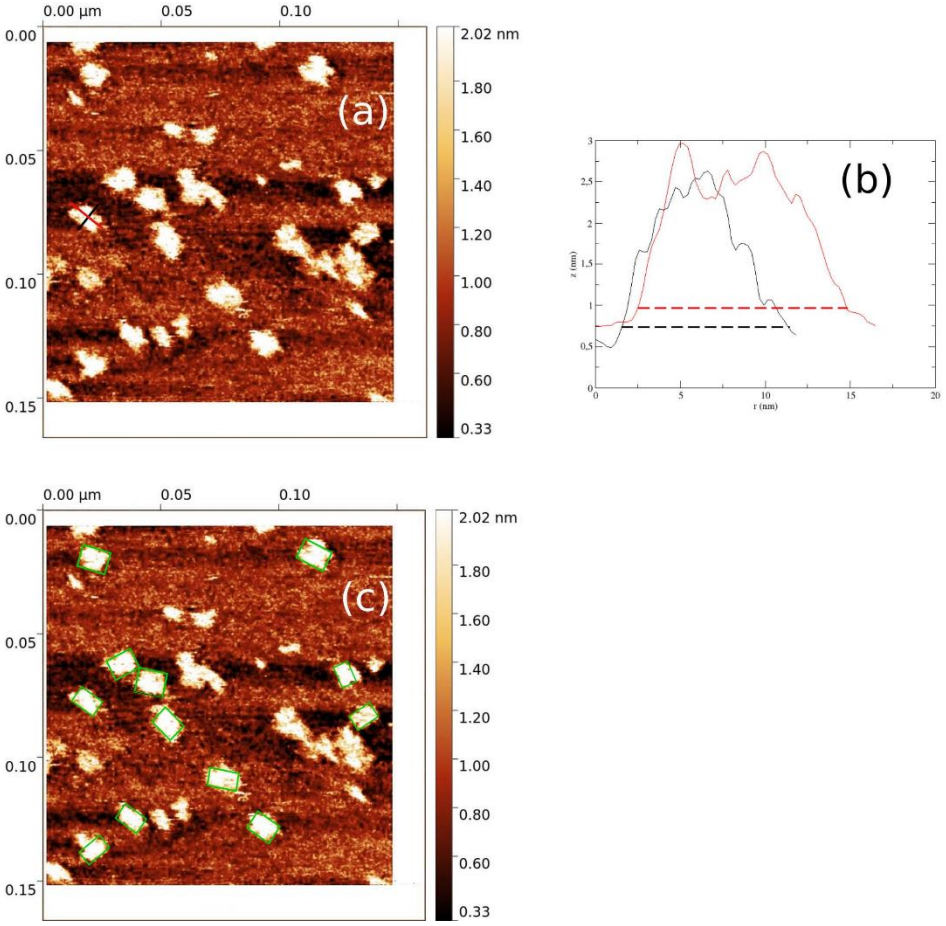


Figure 3-7. AFM images of nanopores.

(a) AFM image (tapping mode) of a 200nM nanopore sample in 20mM Mg^{2+}, 1×TAE buffer, deposited on freshly cleaved mica. This image shows a heterogeneous population of structures. (b) This illustration shows an example of height profiles measured along the red and black axes shown on the left part of (a). The center of this structure has a characteristic height of 2nm consistent with the fact that this structure is made of double stranded helices. Dashed lines correspond to the h_t threshold height (defined in text). (c) For each cluster of points, we performed the area estimation as explained in the main text. In this figure, only the clusters with area between 100nm^2 and 140nm^2 have been marked with a green rectangle. The average width and length of this set of rectangles are, respectively, 9.1 ± 1.2 nm and 13.7 ± 1.3 nm, quite compatible with the 8nm × 14nm theoretical prediction.

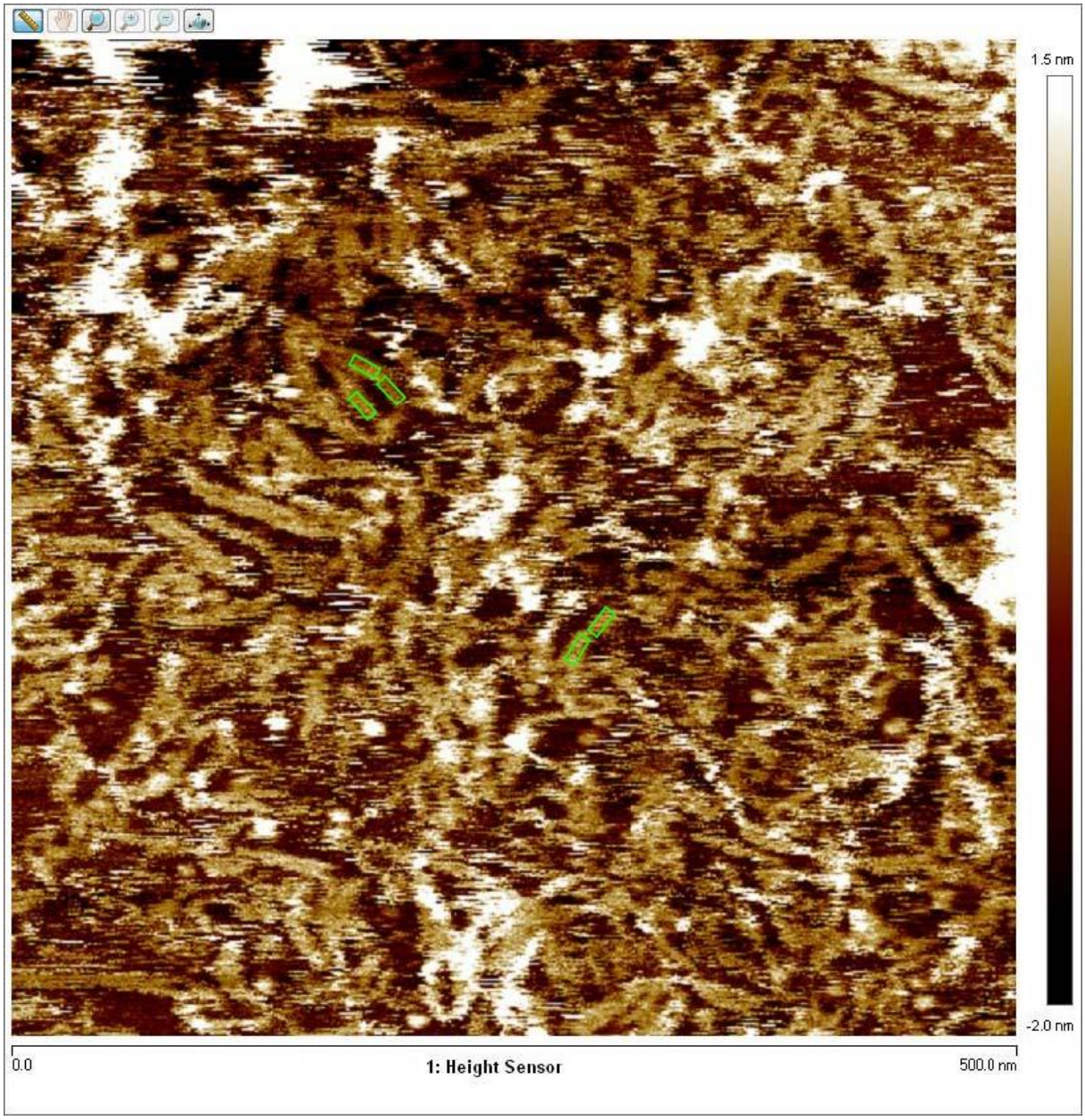


***Figure 3-8**. AFM image (peak force mode) of a 1μM nanopore sample in 20mM Mg^{2+}, 1×TAE buffer, deposited on freshly cleaved mica.*

Increasing the nanopore concentration enhances the stacking between nanopores. As a guide to the eye, several green rectangles with the expected nanopore size have been superimposed to the AFM image.

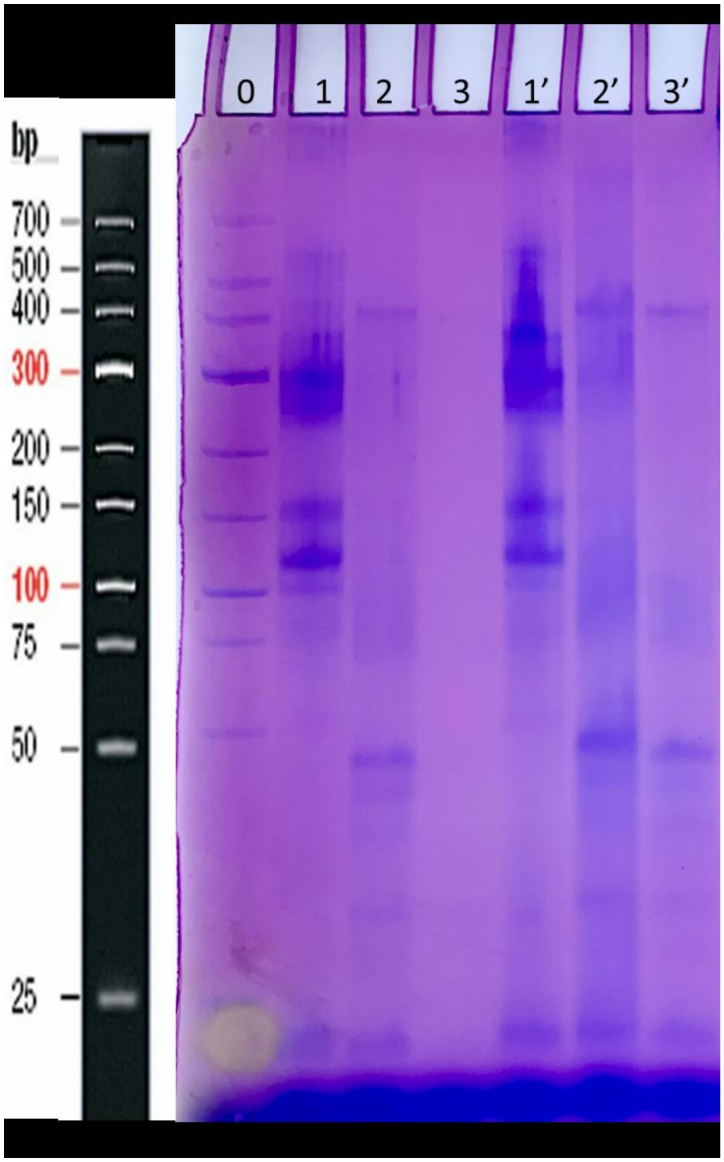


Figure 3-9. Gel electrophoresis run in a 12% PAGE gel.

Premigration step: 25V during 30s. Migration at 100V during 16h. Lane 0 is a DNA ladder. Lanes 1', 2' and 3' are repeats of lanes 1, 2 and 3 respectively. Lane 1: 2µM nanopore with linked staples (staple 4 linked to staple 5, staple 3 to staple 6, staple 14 with staple 1, staple 2 with staple 13), TAE buffer. Lane 2: 2µM nanopore with 12 staples, TAE buffer. Lane 3: same as lane 2 in TBE buffer. Nanopores are located around the 400bp position in lane 2, 2' and 3'. Nanopores with linked staples (lanes 1 and 1') seem to migrate slightly faster.

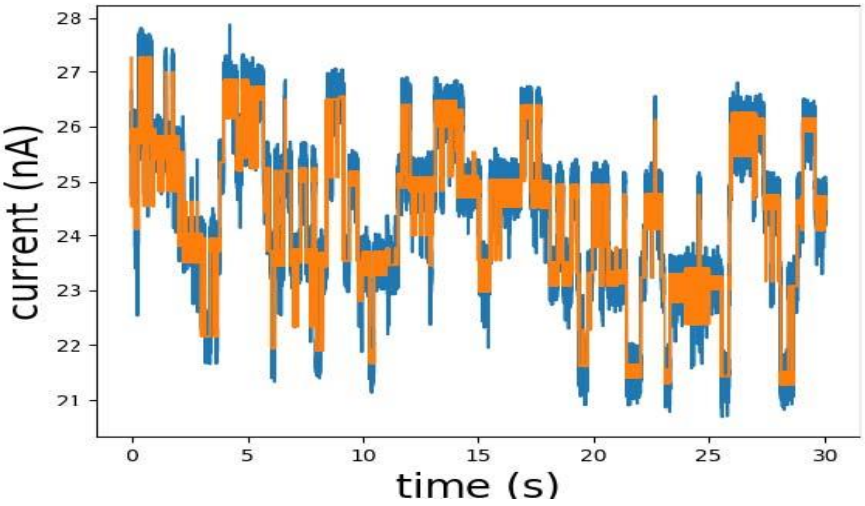


Figure 3-10. *An example of the HMM fitting result. In blue, current recording. In orange, HMM fit of the data.*

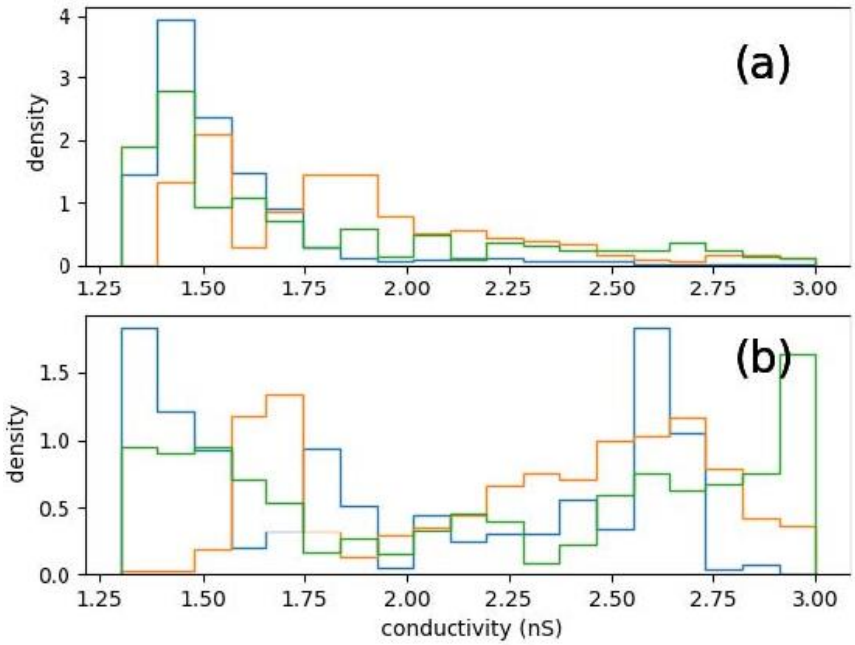


Figure 3-11. *(a) Conductance histograms obtained from three different current recordings [miR-21] = 0nM (b) Same as (a) with [miR-21] = 50nM. To take into account the differences in recording lengths, the probability density has been plotted: the integral of the density is normalized to one.*

3.3 Discussion

DNA nanopores have attracted much attention due to their ease of design and potential applications. Many books and review papers have given good introductions.[143-150] This chapter introduced laboratory work on the conformation of cork nanopore (42-nucleotides-long six-helix bundle) and illustrated the 22-nt oligoes detection through experimental DIB approach. Results from coarse-grained oxDNA simulations, atomic force microscopy experiments (AFM), gel electrophoresis experiments, DIB and fluorescent experiments of this cork nanopore characterization were reported. Here I discuss structural optimization studies aimed at aiding nanopore assembly and insertion capabilities. I introduce successively three versions of this cork like nanopore, named cork_short (the one presented so far), cork_long and cork_SA. They will be introduced subsequently, as well as their ability to interact with DIB bilayer.

3.3.1 Analysis of cork_long

Based on the obtained current recording results, we thought it would be helpful to improve the yield of nanopore structures during the annealing step. Cork_short design, as shown in **Figure 3-12**, has been previously presented, consists of 10 short scaffolds (in blue) and 9 short staples (in other different colours) to form the modified six helix bundle. We then focus on using less scaffolds to design the nanopore. The idea is the same as in the formation of large DNA origamis: the presence of a long scaffold guides the assembly of short staples, improving the overall yield. Therefore, we designed the structure of a new structure, cork_long (**Figure 3-13**) with a number of scaffolds reduced from 10 scaffolds (cork_short) to 3 scaffolds. Then we use the same 3-hour annealing protocol to prepare cork_long type of nanopore and study the oligo detection capabilities using by DIB measurements.

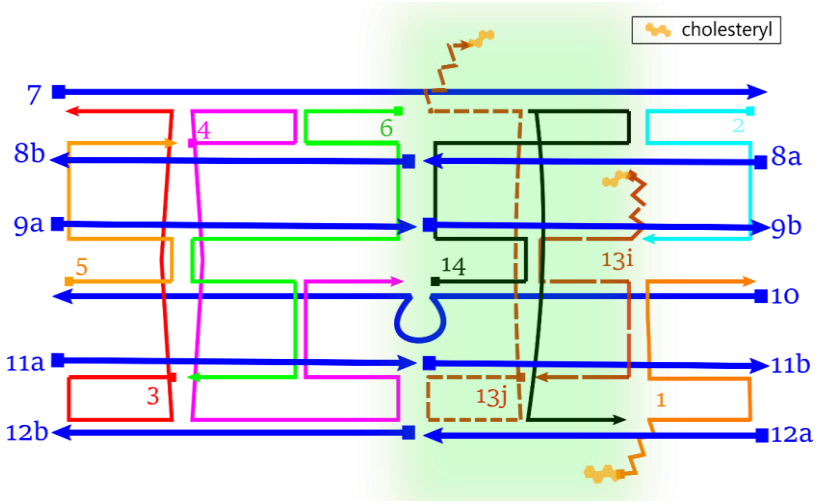


Figure 3-12. *Scheme representation of cork_short structure. This cork_short has ten separate scaffolds: 7, 8a, 8b, 9a, 9b, 10-stem loop, 11a, 11b, 12a, 12b. Nine staples are included: 1-chol, 2, 3, 4, 5, 6, 13i-chol, 13j-chol, 14. The zigzag lines are the complementary strands that can link cholesteryl to aid insertion into the lipid bilayer. The green shadow corresponds to the zone where nanopore should interface with the lipid bilayer.*

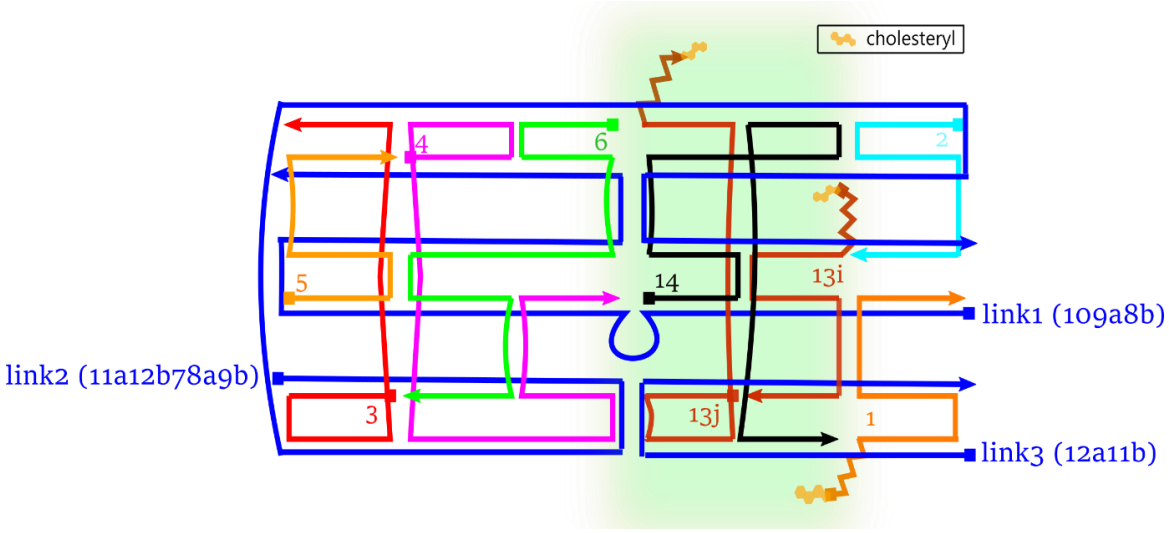


Figure 3-13. *Scheme representation of cork_long structure. This cork_long has three long scaffolds: link1, link2 and link3. In particular, scaffold link1 is obtained by the link of scaffold 10 with stem loop to capture input, scaffold 9a and scaffold 8b; scaffold link2 is obtained by the link of scaffold 11a, 12b, 7, 8a and 9b; scaffold link3 is obtained by the link of scaffold 12 and 11b. Then nine staples are included: 1-chol, 2, 3, 4, 5, 6, 13i-chol, 13j-chol, 14.*

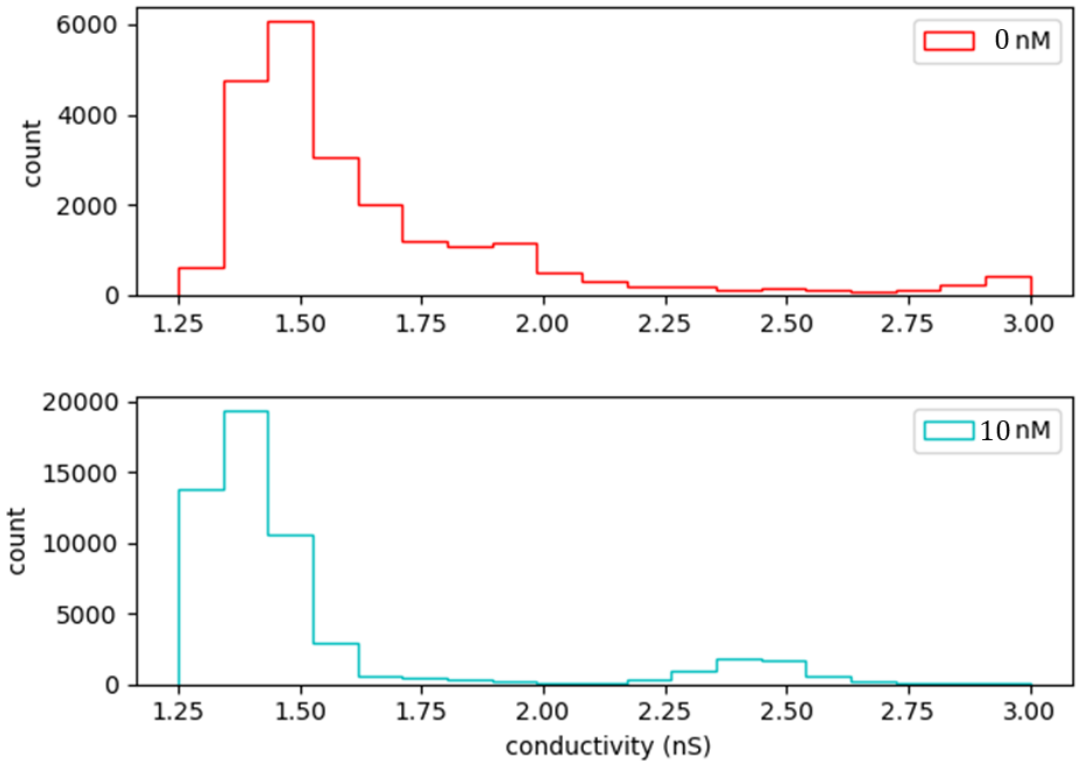


Figure 3-14. **Conductance histograms of cork_long obtained from two different current recordings, respectively.**
The conductance of them in absence of input signal is 1.47±0.01 nS ([input] = 0nM, N=22453) and in presence of input signal is 2.41±0.01 nS ([input] = 10nM, N=54366). [cork_long] = 20nM.

The conductance analysis of the cork_long nanopore reveals similar results to the cork_short nanopore as shown in **Figure 3-14**. The conductance in absence of input signal of cork_short and cork_long ranges from 1.3 to 1.4 nS, while in presence of input signal of cork_short and cork_long, it increases to 2.3–2.4 nS. This indicates that the reduction in scaffolds does not significantly affect the conductance properties of the nanopore.

To further characterize the cork_long structure, we employ tapping mode in liquids in Atomic Force Microscopy (AFM). This technique minimizes sample damage and allows us to observe the nanopore structures. AFM images of the cork_long nanopore were made with 200nM nanopore sample, showing a heterogeneous population of well separated structures. From **Figure 3-15**, two examples of cork_long nanopore (green block in (a) and zoomed image of (d)) exist the size of 8.9 × 14.9 nm, and 8.9 × 14.2 nm, respectively, which are congruent to the average size of a cork_short nanopore (9.1± 1.2 nm wide and 13.7 ± 1.3 nm long reported above). The phase panel (c) and (e) provide information about the material properties, showing different materials depending on the softness or hardness being probed by

the cantilever. The detected phase changes correspond to the height changes in (b) and (d), respectively.

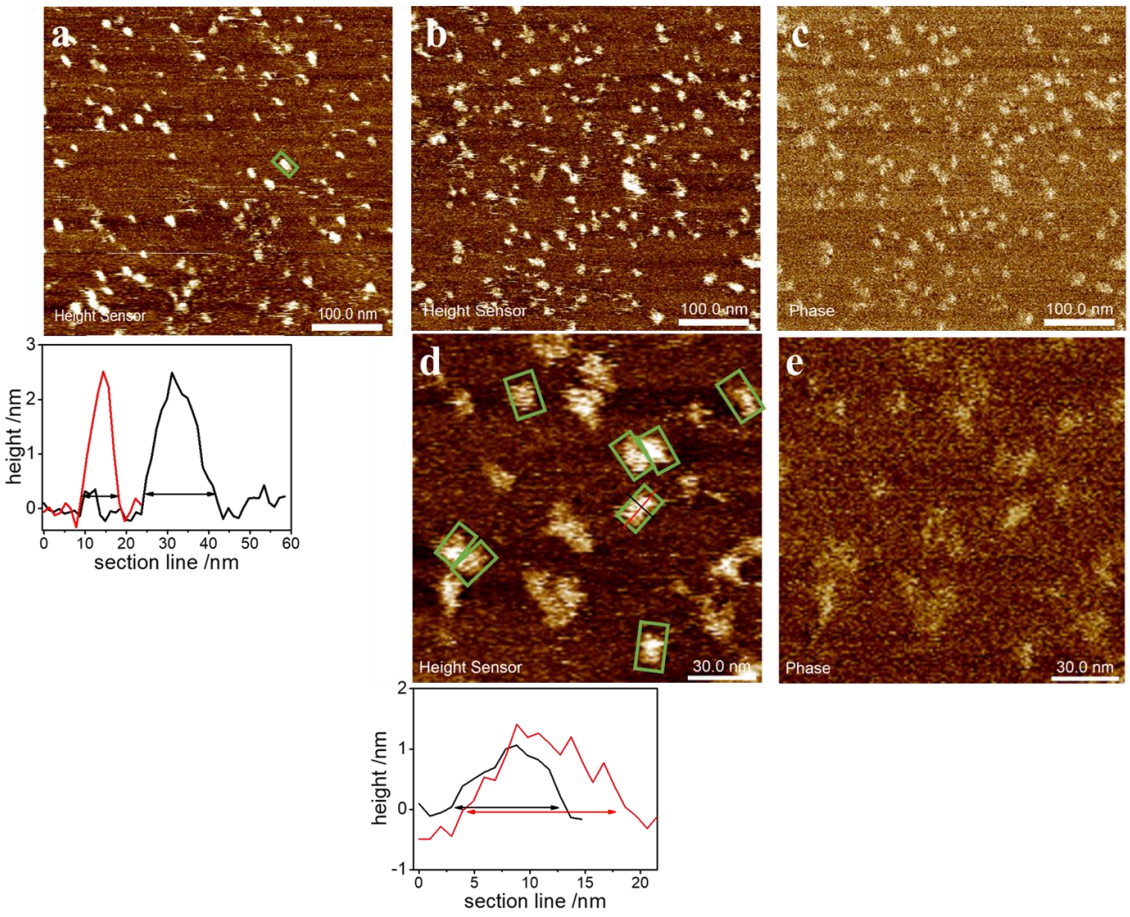


Figure 3-15. Atomic force micrographs of 200nM cork_long constructs. All images are obtained by tapping mode in liquid.

(a) This 500×500 nm micrograph has shown that most of clusters are separated to monomers in 1×TM buffer, pH 8, deposited on freshly cleaved mica. The green block is an example of cork_long structure, and the height profile at the bottom shows that it is 8.9nm wide and 14.9nm long. (d) The green block is another example of cork_long structure in zoomed graph, and the height profile at the bottom shows that it is 8.9nm wide and 14.2nm long. (c) and (e) are the phase panels of (b) and (d), separately.

3.2.2 Analysis of SA-biotin strategy

To further optimize the interaction of the lipid patches with DNA origami structures, a different design was used in this study. Inspired by Krishnan's work,[110] we investigated an alternative strategy for inserting and anchoring DNA nanopores into lipid membranes, which utilizes biotinylated DNA nanopores and lipid molecules.[151] To this end, we incorporated biotinylated anchor strands in the cork nanopore design as shown in **Figure 3-16** and introduced biotinylated POPC lipid molecules, which interacted via the binding of biotin and

streptavidin protein. It allowed specific membrane interactions of the nanopore using a protein-ligand interaction strategy as illustrated as chapter 3.2. Biotin-binding nanopore was intended as a control because it requires streptavidin (SA) as a bridge for coupling and this should only be presented on the biotinylated lipid patches.

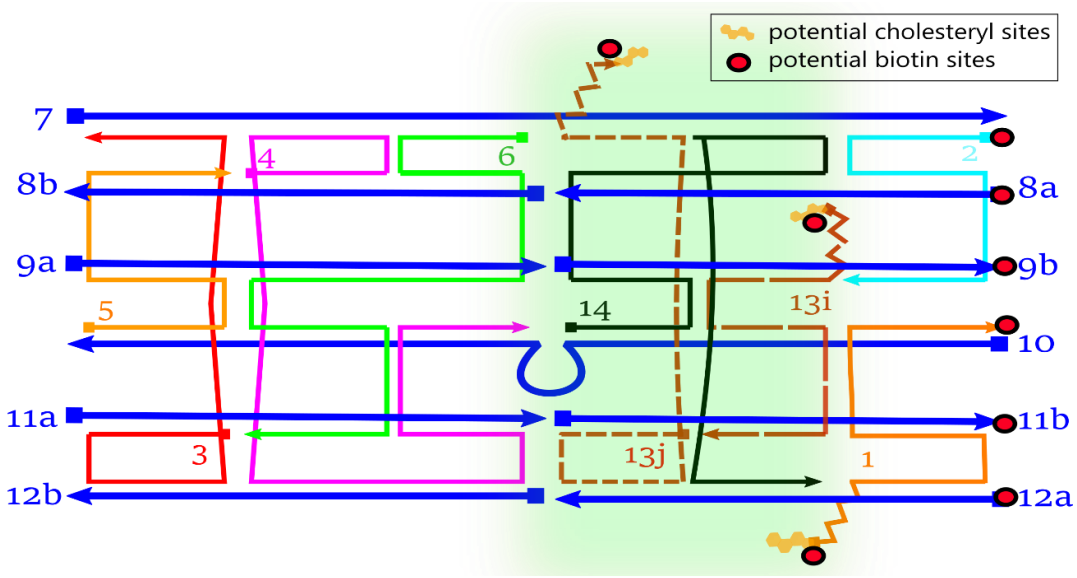


Figure 3-16. *Scheme of cork_SA structure. This cork_SA model has nine potential biotin modification sites at the bottom of the structure, including staple 1, 2, 13i, 13j, scaffold 8a, 9b, 11b, 12a. And the scaffolds and the staples are same like cork nanopore as described above.*

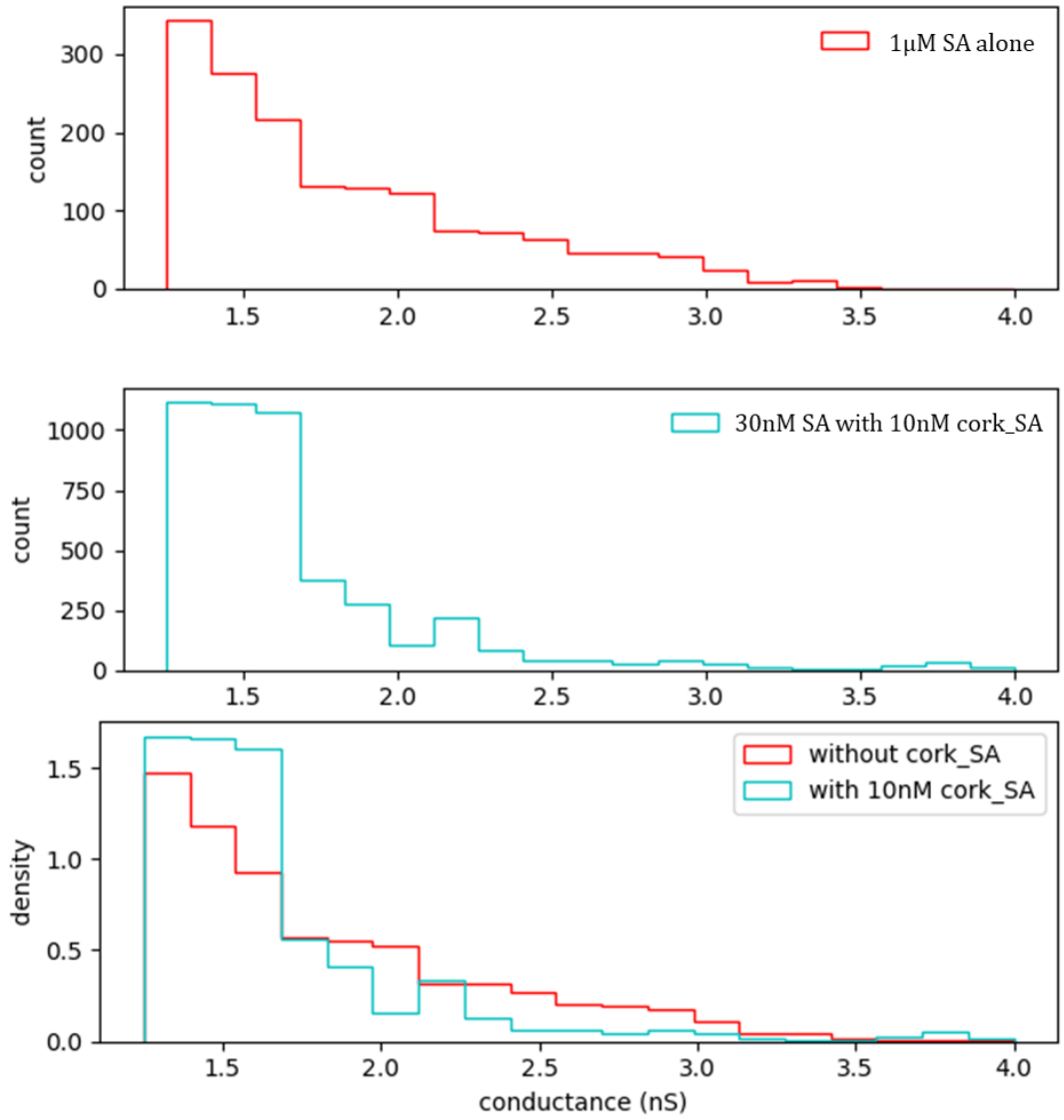


Figure 3-17. The histograms of conductance of SA (1μM) and cork_SA (10nM cork_SA with 30nM SA). *(a) and (b) The conductance of 1μM SA protein only is 1.31±0.01 nS (middle, N=1614), while the conductance of cork_SA and SA are 1.39±0.01 nS (bottom, N=4630), and on the histogram distribution of 1.25–1.7 nS has increased. (c) To take into account the differences in recording lengths, the probability density has been plotted: the integral of the density is normalized to one. Conductance histogram of SA protein with and without cork_SA nanopore.*

In this SA-biotin strategy, we perform experiments to detect input oligoes using the SA with concentrations ranging from 0.1nM, 1nM, 10nM, 100nM, 1μM, 3μM of SA in 1×B buffer. The cork nanopore concentration is set up from 0.001nM, 0.01nM, 0.05nM, 0.1nM, 0.5nM, 1nM, 5nM, 10nM, 20nM in 1×BO (0.25% OPOE) buffer. The results indicate that at SA concentrations lower than 100 nM, it becomes difficult to detect current jumps. The success ratio of current recording of cork_SA compared the cork_short we reported above is 6%, lower than 10%. However, when the SA concentration is increased to 100 nM and 1 μM,

current jumps are consistently observed, even in the absence of the cork_SA nanopore. Here, **Figure 3-17** showed the conductance changes can be detected only with 1μM SA protein. Compared with the histogram of the cork_SA nanopore together with SA protein, we found that the conductance is in large ranges of 1.3–3.5 nS. This suggests that, higher concentration of SA, protein alone, can insert into the biotin-modified lipid bilayer and then cause the current jumps (probably making protein nanopores). Although we used the interaction of streptavidin and biotin to design the binding of the nanopore structure to the phospholipid bilayer, it was found that a slightly higher concentration of streptavidin protein can obtain a stable and indiscriminate current signal in the current experiment. Both the addition of SA and nanopores will cause the current change signals, we can conclude that this strategy is not suitable to enhance nanopore insertion for detecting signals, leading us to abandon this approach.

In conclusion, our structural optimization studies involving the cork_long nanopore demonstrate that reducing the number of scaffolds does not significantly impact the conductance properties of the nanopore. AFM characterization confirms the presence of well-separated structures, supporting the successful formation of the optimized nanopore. Additionally, the SA-biotin strategy for nanopore insertion and anchoring proves ineffective in enhancing nanopore insertion for signal detection. These findings contribute to the understanding of DNA nanopore design and optimization, paving the way for further advancements in this field.

4 Characterization and Manipulation of nanopores inserted into an origami platform

4.1 Introduction

Over the past three decades, a diverse range of molecular devices [152] has emerged, offering promising applications in various fields such as medicine [153] and molecular robotics [154]. Among these devices, DNA (or RNA) structures have gained significant attention because of their predictable self-assembly behaviour (as compared to protein structures, for instance) and ease of synthesis. Abiotic DNA-programmed self-assembly, pioneered by Seeman,[155] has been a major focus of research, with numerous reviews highlighting different aspects of this field.[156-163]

DNA nanotechnology has demonstrated the potential to design molecular devices including motors [164] and logic circuits [165] by leveraging the self-assembly properties of nucleic acids. One of the interesting properties of such devices is the possibility to embed amplification circuits that create an interface between events at the molecular level and a macroscopic apparatus that provides with a measurable signal. In other words, the exquisite recognition or assembly properties of nucleic acids can be translated to a macroscopic signal. However, many if not all of them require a thresholding mechanism to avoid unwanted triggering of the amplification.

In the previous chapter, we have shown how a single recognition event between two complementary oligonucleotides can be translated to an electric signal. This method is fundamentally close to the so-called stochastic sensing [128] which in turn uses a widely spread recognition principle in biology. Many sensory systems rely on channels and pores that change conformation upon interaction with small molecules. In stochastic sensing, engineered nanopores are able to interact with specific molecules, which upon binding modify their electric conductance. These nanopores act as transducers, translating the molecular recognition to a physically measurable signal. Nanopore sequencing technology [128] is a good example of the far reaching applications of stochastic sensing. Here, we use the same

principle but, instead of protein nanopores, we develop DNA made nanopores that change conformation upon interaction with short oligonucleotide sequences.

A single stranded oligonucleotide is a floppy object (persistence length ~ 2nm), whereas a double stranded oligonucleotide is much stiffer (persistence length ~ 50nm). The conformational changes in DNA-made devices often rely on the mechanical properties of single-stranded and double-stranded oligonucleotides. Previous published works [109,166], also chapter 3 in this thesis, have explored two DNA nanopore designs that exhibit changes in conductance upon binding with short oligonucleotide sequences (called input in the sequel). This is not a purely academic problem. The detection of short RNA sequences called microRNAs plays an important role in the development of recent biomarkers. Existing methods such as standard reverse transcription-polymerase chain reaction (RT-PCR) display significant biases for short sequences. Even though alternative solutions to standard RT-PCR methods exist, there is still room for improvement in this field, particularly in the limit of low concentrations which characterize these oligonucleotides in body fluids.

Previous studies have utilized engineered nanopores which, upon binding, increased inner diameter [109] or reduced the effective length [166]. In both cases, this change of conductance could be detected using current recordings. Also, in both cases, the process is probably similar: DNA nanopores can insert into a lipid bilayer, making a hole. In a current recording, the insertion translates into a jump in conductance. As shown in **Figure 4-1**, in the absence of the input strand (brown line), the nanopore is in a ‘closed’ state with a corresponding conductance level (black curve). Upon binding with the input strand, the nanopore shifts to an ‘open’ state, leading to a different conductance value (red curve). Consequently, the histogram of conductance jumps exhibits two distinct peaks, allowing the inference of input concentration based on the relative heights of the peaks. As schematized in **Figure 4-1b**, depending on the nanopore design, there can be an overlap between the histograms of open and closed nanopores. This was actually the case in refs [109] and [166].

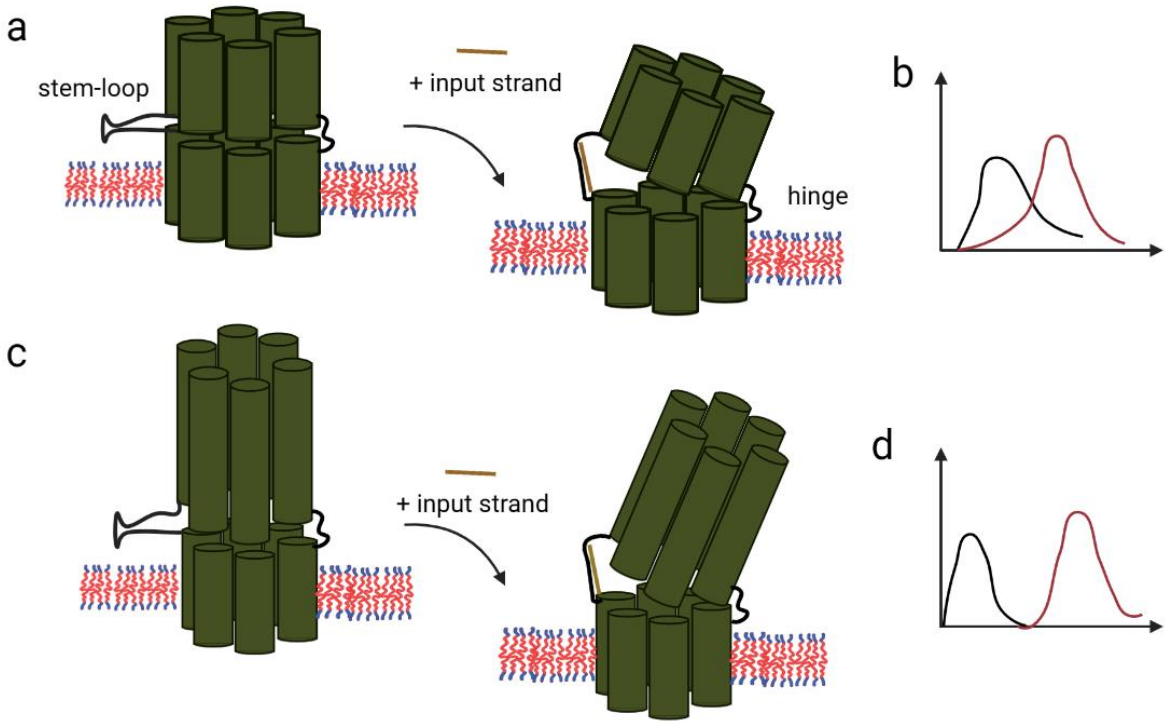


Figure 4-1. Principle of oligonucleotide detection.
(a) and (c) nanopore opening occurs when a short oligonucleotide, input strand, hybridizes to the loop of the stem-loop, the nanopore of (c) is asymmetric and has a longer length in the upper halve. (b) and (d) are histograms of conductance shifting from the black into red curve, representing the cases in (a) and (b), respectively.

Here, we want to push further this idea to develop new nanopores that unambiguously and efficiently respond to single, short oligonucleotides. For this, we will explore ways to improve three main deficiencies of our previous work. First, the insertion capabilities of DNA nanopores based on six helix bundle designs are poor. Second, nanopores in the 'closed' state have a conductance which is too close to that of 'open' nanopores, which implies that some of the insertion events cannot unambiguously be related to a 'closed' or an 'open' state. Finally, previous versions of nanopores had a strong tendency to leave the bilayer, which was reflected in the existence of positive and negative jumps. If nanopores can leave the bilayer, there is not a clear correspondence between the number of jumps (whether positive or negative) and the number of open nanopores. In the following, we explain how to fix these issues.

DNA based nanopores are hydrophilic structures. The insertion of any of these structures into lipid bilayers requires some hydrophobic moiety to be incorporated into the nanopore. The number of possible hydrophobic modifications and the method of attachment depend on the nanopore size. For instance, four or six helix bundles are quite small structures which can

bear a small number of hydrophobic modifications. On the other hand, nanopores embedded into large DNA platforms may display significantly larger number of hydrophobic modifications (up to 24–26) which in principle should facilitate the interaction with membranes.

DNA origami, pioneered by Dr. Paul Rothemund,[130] involves folding DNA into desired shapes using a single-stranded scaffold and short oligonucleotide staple strands. Here, we will focus on a planar rectangular 2D-DNA origami structure to hold hydrophobic moieties, specifically cholesteryl, in proximity to the nanopore. The arrangement and number of cholesterols will be optimized to enhance nanopore insertion efficiency. To address the overlap between open and closed states, we introduce longer and asymmetric nanopore designs, making it easier to monitor the two-fold difference in conductance (**Figure 4-1 c** and **d**).

Here, we study a DNA nanostructure consisting of a rectangular platform to which a perpendicular cylinder is attached. The cylinder is formed by six parallel double helices and is similar to that published elsewhere [166]. The rectangular platform is formed by 22 parallel double helices, arranged side by side. A long, circular, single stranded DNA (M13mp18) forms the common strand (scaffold) that runs along all the helices. For each double helix, the companion strand (usually called staple) is computed in such a way that several staple strands (belonging to neighbour helices) can be linked without distorting double helices. This method, due to P. Rothemund [130], is known as ‘DNA origami’.

The rectangular platform has a central empty area which is intended to be crossed by a cylinder, the active part of our device. This cylinder is made of six parallel double helices which may have in common the same scaffold as the platform. The device obtained by the assembly of platform and cylinder is intended to have three properties: (P1) be able to attach to a lipid bilayer (P2) be able to punch the lipid bilayer, creating a hole on it and (P3) be able to regulate the conductance across the cylinder. Our hypothesis is that the successful combination of these three functions should lead to a functional device able to regulate the conductance of lipid bilayers in a controlled way. Here, we rely on previously published work, as developed in chapter 3, to cope with P3. In short, four of the six double helices are split into two. A fifth helix is kinked and acts as a link between the two halves of the cylinder. Finally, a stem-loop structure acts as an effective length regulator. In its closed state, the

stem-loop acts as a lock, keeping the two halves at a close distance. In the presence of an input signal, a specific single stranded oligonucleotide able to hybridize to the loop of the stem-loop, the stem-loop changes conformation to a double helix linked to a single stranded section. The mechanical tension exerted by the 'open' conformation is higher than that of the 'closed' conformation, leading to the 'opening' of the cylinder.

In the following, two DNA origami platforms, T1 and T2, are introduced and discussed, differing in their incorporation of the nanopore and scaffold. T1 involves separate scaffolds for the nanopore and platform, while T2 utilizes a single scaffold traversing both the platform and the nanopore. Besides linking the nanopore to the platform, we also need to dictate the polarity of the assembly. To enforce the polarity, we will include additional staples linking the middle part of the cylinder to different positions on the platform. Overall, over 200 staples with average length 32-nt are necessary to build the device.

In summary, chapter 4 focuses on developing DNA-based nanopores embedded in DNA origami structures for improved sensing and detection of single short oligonucleotides. By optimizing nanopore insertion capabilities, conductance differences, and stability within the lipid bilayer, we aim to enhance the accuracy and efficiency of detection in low-concentration scenarios. These advancements have the potential to contribute to the development of biomarkers and improve molecular sensing technologies.

4.2 Methods and Results

4.2.1 Designing nanopores attached to a rectangular 2D DNA origami

In this section, we explore two possible strategies (T1 and T2) for vertically inserting a six-helix bundle (6HB) structure into a DNA rectangular tile, bearing a hole in the middle. The T1 strategy involves vertically inserting a six-helix bundle (6HB) structure into a DNA rectangular tile with a central hole, with separate designs for platform and nanopore. In the T2 strategy, the nanopore and rectangular platform share the same scaffold. We discuss the design methods for both the nanopore and platform structures and topologies.

4.2.1.1 Design of the 6HB nanopore structure

The 6HB structure is assembled by multi-crossover DNA staples. Following Seeman's work, [167] we arrange the helices in a honeycomb lattice, where individual crossovers are separated by seven nucleotide pairs. This arrangement allows B-DNA for an approximately 120° dihedral angle, corresponding to 2/3 and 4/3 turns for 7 and 14 nucleotide pairs, respectively. In principle, six helices, each rotated 120° out of the plane of the last pair, ought to produce a hexagonal bundle containing a hole down the middle. From **Figure 4-2**, it is clear that in its ideal form, the motif contains a central hole that is approximately the diameter of a DNA double helix, ~2 nm. We considered four different designs of the 6HB part. In our 6HB designing, we use scaDNAno to design 'cork', 'longcork' and 'cookie' nanopore for T1, and 'v7' nanopore for T2, as described in the following **Figure 4-3**. Thereinto, the structures of cookieT1 and v7T2 were designed with single-strand M13mp18 serving as a single scaffold to form the 'cookie' or 'v7' nanopores and corresponding T platform.

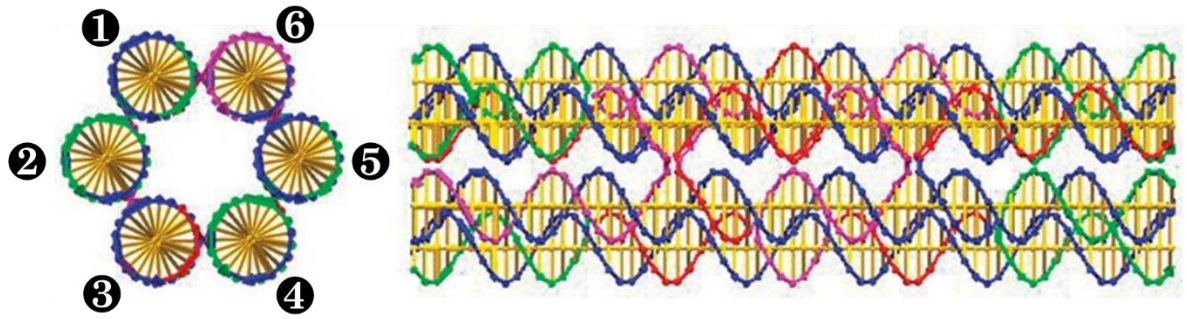


Figure 4-2. Schematic drawings of the six-helix bundle motif.
We use the staples of the motif with 7 nucleotide pairs between crossovers. (Left) A cross-sectional geometrical view of a six-helix bundle in which the helices are indicated by Roman numerals. (Right) A geometrical side view of a six-helix bundle in which the crossovers are separated by seven nucleotide pairs. The image is reproduced from ref. [167].

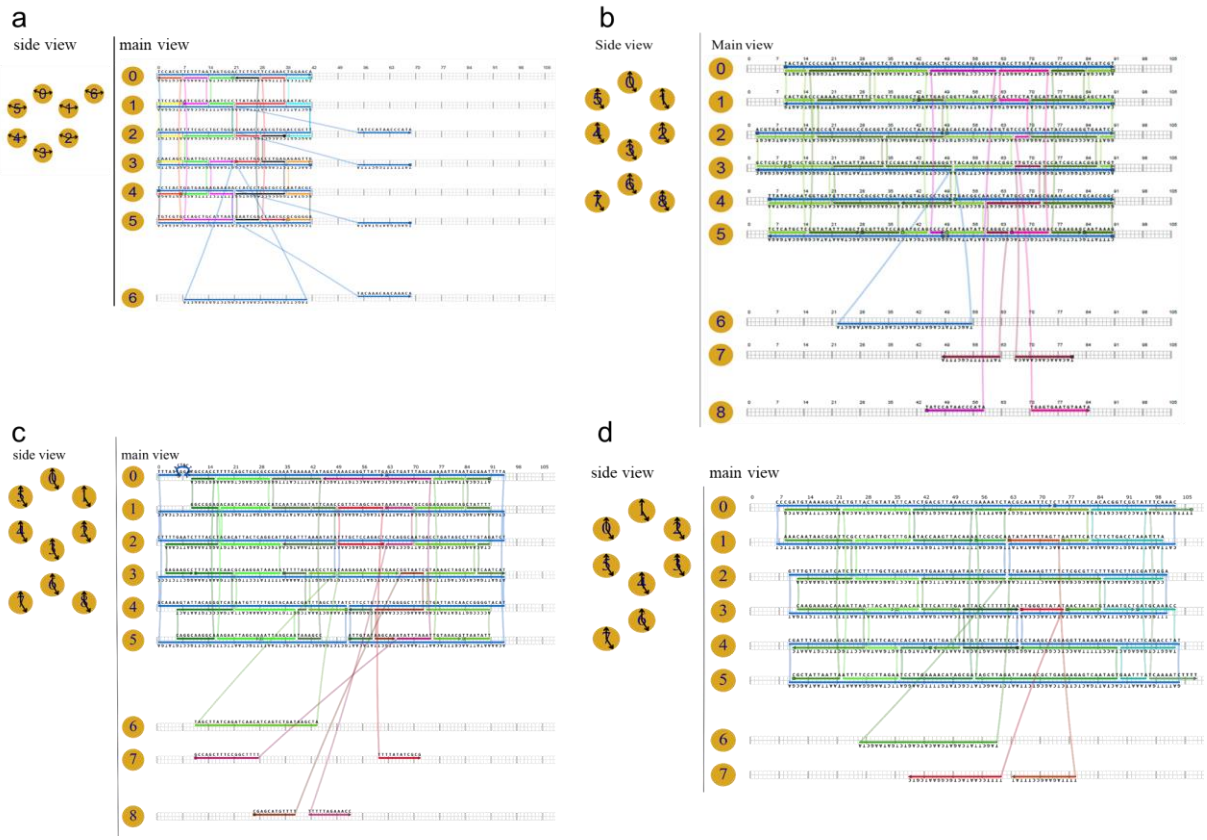


Figure 4-3. The side view and main view of 'cork' nanopore (a), 'longcork' nanopore (b) and 'cookie' nanopore (c) for T1, and 'v7' nanopore (d) for T2 in scaDNAno window.

There are six helices (helix 0-5) designed in honeycomb grids. The DNA sequences are shown in the main view. (a) Blue strands are scaffolds for 'cork' nanopore, eight other color strands are staples for cork nanopore. And three scaffold strands with four extensions are supplying four linkers between 'cork' nanopore and T1. The detailed sequences are shown in ***Table S3****. (b) Six blue strands are scaffolds for 'longcork' nanopore, seven green (light and dark) strands and three red strands are staples for 'longcork' nanopore and three red strands are supplying four linkers between 'longcork' nanopore and T1. The detailed sequences are shown in* ***Table S5****. (c) A single long blue strand is the scaffold (600-nt) from T1 for 'cookie' nanopore, eight green (light and dark) strands and four red strands are staples for 'cookie' nanopore, and four red strands are supplying four linkers between 'cookie' nanopore and T1. The detailed sequences are shown in* ***Table S9****. (d) A single long blue strand is scaffold (570-nt) from T2 for 'v7' nanopore, eight green (light and dark) strands and two red strands are staples for 'v7' nanopore, and two red strands are supplying four linkers between 'v7' nanopore and T2. The detailed sequences are shown in* ***Table S13****. The helical offset is set from 0 base to 105 bases. In honeycomb arrangement, every 7 bases are a block, as indicated by the major tick offset at zero helix.*

4.2.1.2 Interaction between an embedded DNA nanopore and a lipid bilayer

Figure 4-4 shows the expected size of each design. The difference between ‘cork’ and ‘longcork’ is the length of the half that is not directly bound to the rectangular platform and which acts as a ‘lid’. This part is longer for the ‘longcork’ design, which should translate into a lower conductance in the closed state, thereby separating the two conductance peaks as sketched in **Figure 4-1**. In the following, we indeed show that this is the case. Consequently, for the T2 design, we only considered an analogue of the ‘longcork’ design, named ‘v7’, and no analogue of the shortest ‘cork’ version. Finally, we also designed an analogue of the ‘v7’ structure using the T1 design and sharing the same scaffold with the nanopore (‘cookie’ nanopore). Unfortunately, the number of experiments done with this design was not large enough to reach any conclusion.

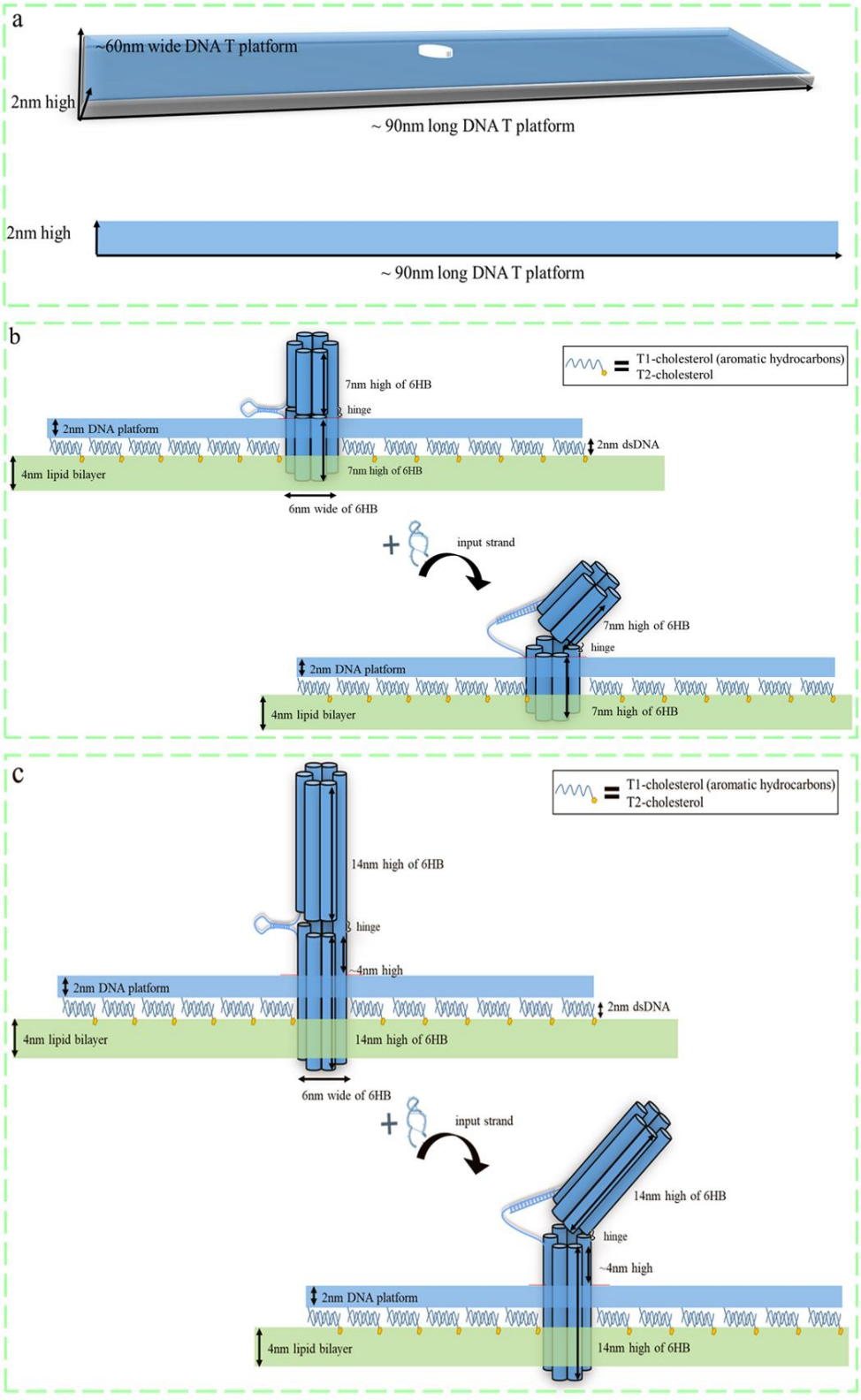


Figure 4-4. Scheme of T1 and T2 pore insertion into lipid bilayer by cholesterol anchors.
(a) The 3D diagram (up) and cross section (down) of DNA T platform with a hole in the middle. (b) The diagram of same short length of upper and bottom halves of 6HB, corkT1. (c) The diagram of longer length of two halves of 6HB, such as longcorkT1, cookieT1 and v7T2. This design increases the length of 6HB. In this way, the stem loop is ~ 4nm far away the lipid bilayer and the depth of insertion is increasing from 1nm to 4nm.

4.2.1.3 Design of the rectangular planar structure

For the design of the rectangular DNA tile, we also employ multi-crossover DNA staples inspired by Rothemund's work.[130] The helices are arranged in a square alignment to form the DNA tile, with crossovers separated by multiples of eight nucleotide pairs. This corresponds to approximately 3/4 and 3/2 turns for 8 and 16 nucleotide pairs, respectively. The folding path of the scaffold is chosen so that it passes through the whole area of the shape, running back and forth. The folding of the scaffold is ensured with the aid of many staple strands, usually binding to three adjacent helices either in an S-shaped or Z-shaped geometry. DNA origami motifs with straight edges sometimes stick together at the edges since the DNA base-pairs exposed at the edge are highly hydrophobic and tend to stack to each other. In order to prevent such aggregation, a single-stranded portion (typically T4) is often introduced to the staple strands located at the edges. Thus, the location of staple crossovers avoids strain on the helices, and the addition of single-stranded portions (typically TTTT) prevents aggregation at the edges of the DNA tile. Additionally, we can select possible cholesteryl (cholesterol modified single strands) binding staples, to ensure that the 3' ends are homogeneously oriented to the same side of the DNA rectangle.

We considered two different strategies (called T1 and T2) to link the nanopore to the platform. The DNA tile design is such that ~200 staples help to form a rectangular pattern as a nano-board with a central hole. The designs denoted as corkT1 and longcorkT1 share the same T1, as depicted in **Figure 4-5** and **Figure 4-6**, respectively. It is noteworthy that the T1 design for corkT1 and longcorkT1 differ from their connection sequences of 6HB and T1 solely in the middle of T1. We also considered two more combinations between 6HB and tile configurations sharing a single scaffold, resulting in the formation of cookieT1 and v7T2, as illustrated in **Figure 4-7** and **Figure 4-8**, respectively.

The length for a single turn (360°) of a B-DNA helix appears to be close to 10.5 base pairs or 3.4nm. Because of the existence of major and minor grooves, DNA backbones are not symmetrically spaced around the helix. However, in the origami strategy such as implemented in scadnano, to design the two-dimensional DNA nanotile, the 10.5 period is approximated either by 10 or 11 bases. For instance, the length between crossovers spaced by 1.5 turns is typically 16nt. In both T1 and T2, staples are typically 32nt long, with three concatenated parts —8nt + 16nt + 8nt— spanning three neighbouring helices. The central 16-nt stretch

binds to one helix, and each set of 8nt at the ends binds to the adjacent helices. Therefore, the application of this rule generates strained structures. This strain was only 'corrected' in the T2 platform, where deletions (denoted by red crosses in the scadnano files) were introduced. However, this correction is only partial, as can be observed, for instance, in the oxDNA simulations presented below, where the DNA tiles are twisted, and non-planar. The scaDNAno figures give illustration of the relative positions of the different elements that compose either T1 or T2 based nanopores.

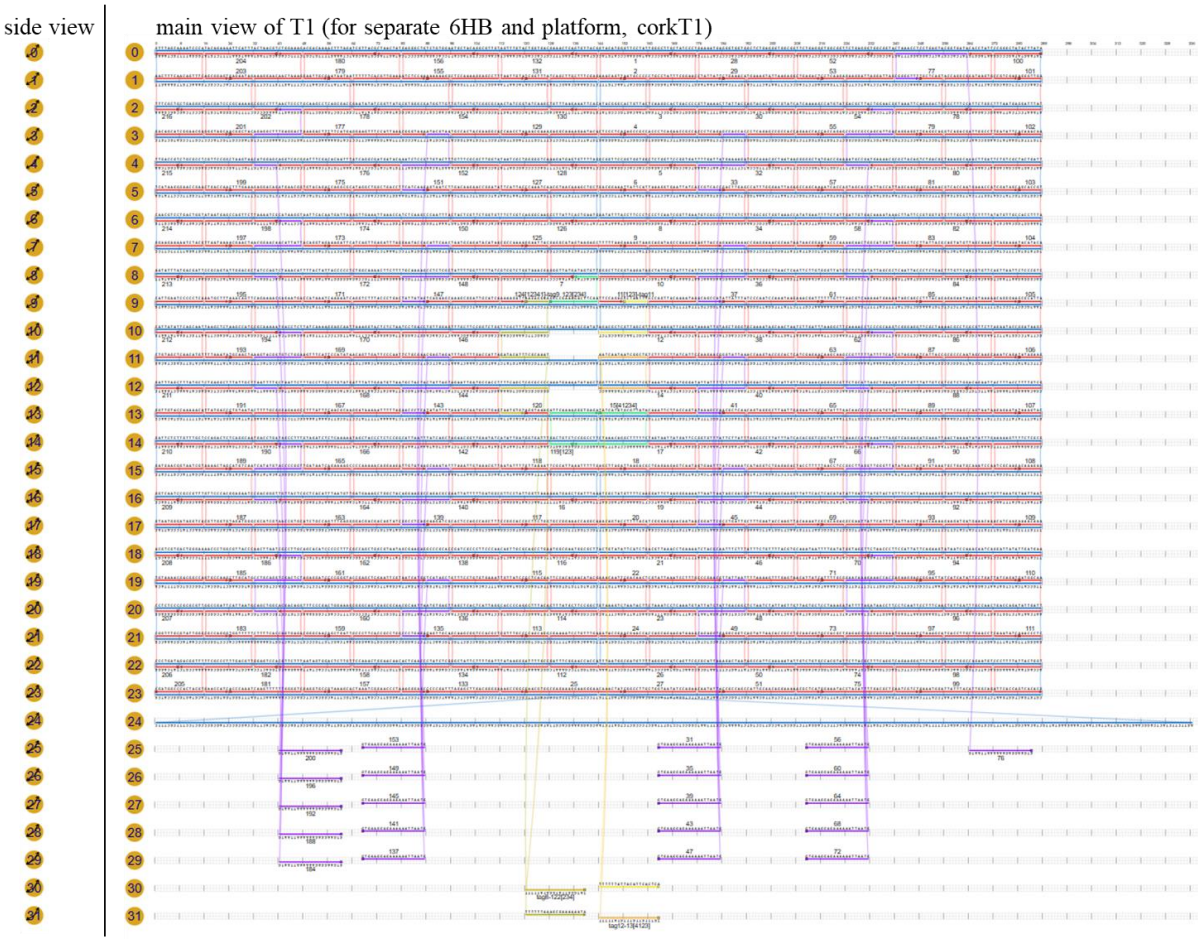


Figure 4-5. Scaffold / staple layout of the T1 object (for the assembly of corkT1).
*Generated with scaDNAno online. Scaffold-label is colored blue; possible cholesteryl-binding staple are colored purple. The four chains marked in yellow and three modified staples marked in green, where the four linkers can connect the four linkers of the 'cork' (**Figure 4-3a**) nanopores to construct corkT1. Other staples are colored red. The detailed sequences are shown in **Table S1** and **Table S2**.*

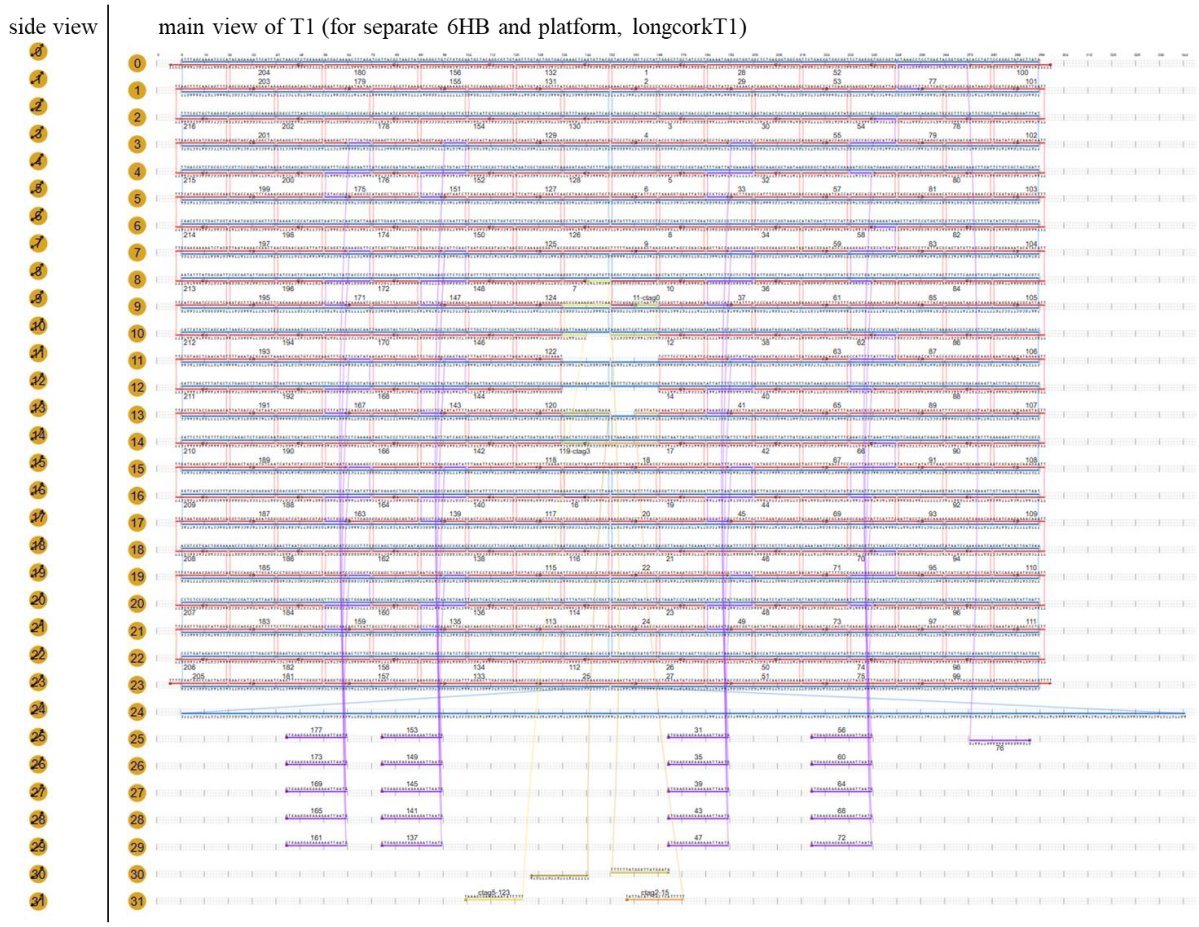


Figure 4-6. Scaffold / staple layout of the T1 object (for the assembly of longcorkT1).
*Generated with scaDNAno online. Scaffold-label is colored blue; possible cholesteryl-binding staple are colored purple. The four chains marked in yellow have four linkers that can connect the four linkers of the 'long cork' (**Figure 4-3b**) nanopores to construct longcorkT1. Other staples are colored red. The detailed sequences are shown in **Table S1** and **Table S2**.*

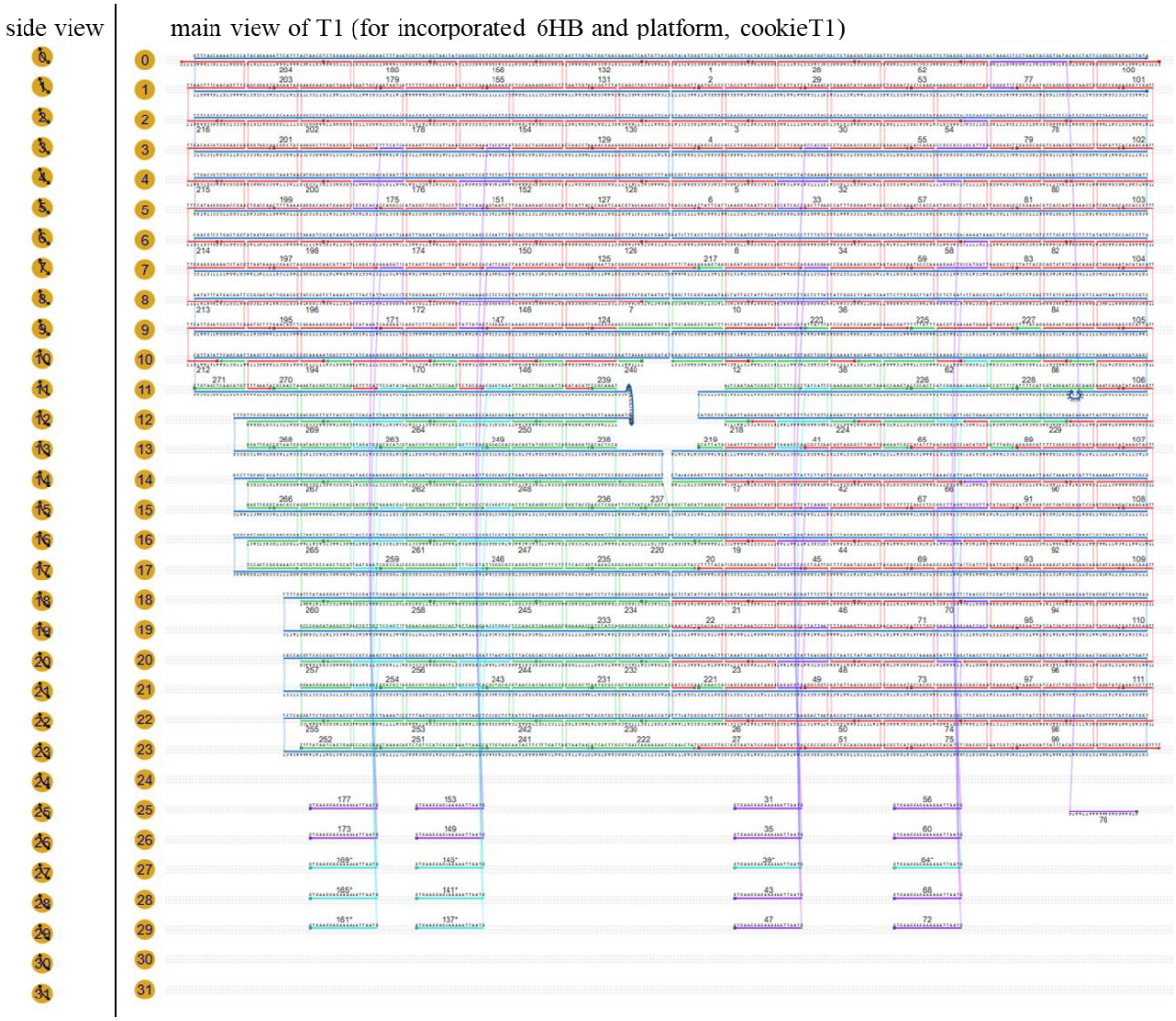


Figure 4-7. Scaffold / staple layout of the T1 object (for the assembly of cookieT1).
Generated with scaDNAno online. Scaffold-label is colored blue; possible cholesteryl-binding staple are colored purple and light blue; other staples are colored red and green. Light blue and green marked staples are modified staples compared to T1 in ***Figure 4-5****. The 600-nt loop-out in the middle of helix 11 and helix 12 is used as a scaffold to construct the 'cookie' nanopore as shown in* ***Figure 4-3c****, thus forming cookieT1. The detailed sequences are shown in* ***Table S7*** *and* ***Table S8****.*

Figure 4-8. Scaffold / staple layout of the T2 object (for the assembly of v7T2).
Generated with scaDNAno online. Scaffold-label is colored blue; fifty of possible cholesteryl-binding staples are colored purple; other 150 staples are colored orange. Nine green marked staples are the staples near the central pore of T2. The 570-nt loop-out in the middle of helix 9 and helix 10 is used as a scaffold to construct the 'v7' nanopore as shown in ***Figure 4-3d****, thus forming v7T2. The detailed sequences are shown in* ***Table S11*** *and* ***Table S12****.*

4.2.1.4 Linking the nanopore to the rectangular platform

Two strategies (T1 and T2) are considered to link the nanopore to the platform. In corkT1 and longcorkT1, four regions of the scaffold, located around the central hole, were left unbound to act as attaching points for the nanopore. Correspondingly, four staples of the 6HB design were elongated to hybridize to these sections. The scaffold of 'cookie' and 'v7' nanopore was designed to share the same scaffold as the tile. The staples of the 'cookie' and 'v7' nanopore also include four and two linkers (respectively) to enforce, as much as possible, that the 6HB is orthogonal to the tile.

A useful test of the correctness of the design is provided by molecular dynamics simulations. For this purpose, we use oxDNA software,[112] a collection of programs based on a coarse-grained description of DNA. In this approach, interactions with water and counter-ions are only included in an implicit way. Each nucleotide is described by a two-site bead, inter-bead interactions have been parameterized to approach as much as possible the mechanical properties of DNA. For instance, **Figure 4-9** shows two conformations of the v7T2 device, one close to the theoretical configuration where all double helices are parallel and stay in a plane, another one showing deformations induced by thermal fluctuations. Despite these fluctuations, the cylinder stays globally perpendicular to the platform. These simulations do not take into account any interaction with lipid bilayer. The latter being rather fluid, it could be expected that it adapts to the local curvature of DNA platform.

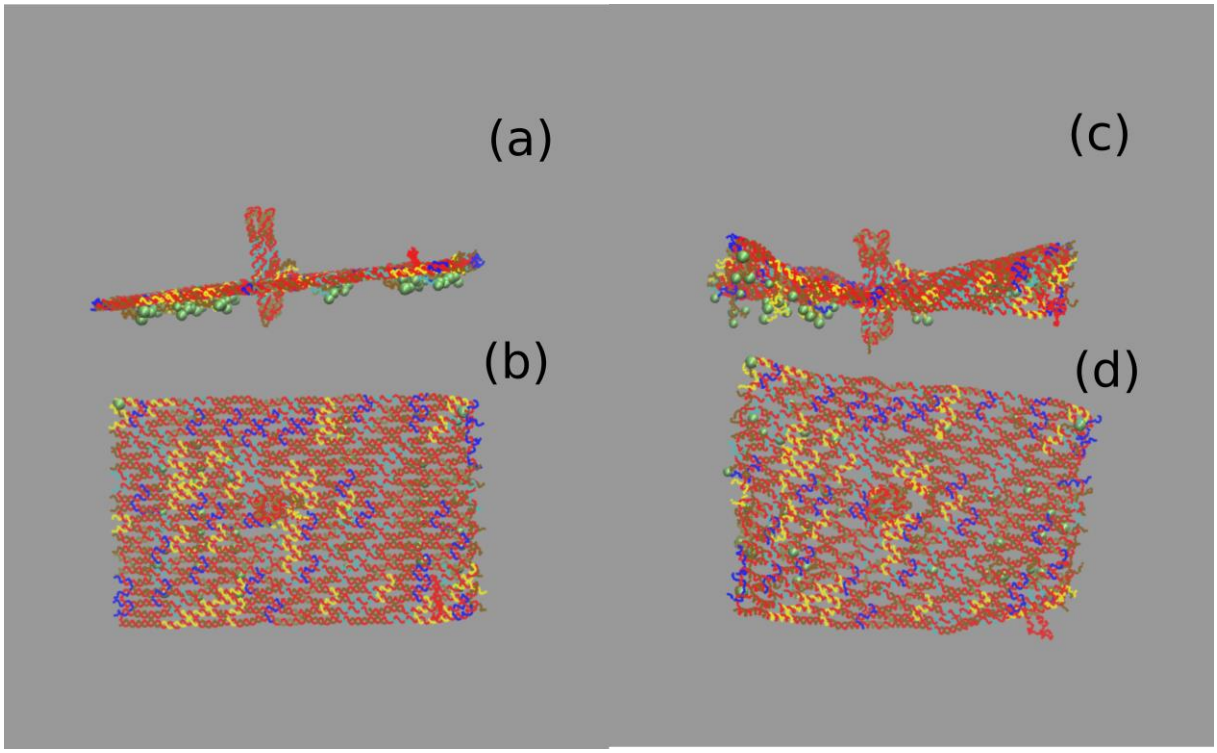


Figure 4-9. *(a) Lateral and (b) upper view of a device conformation obtained using oxDNA simulations. Each strand has a different color. Scaffold is in red. Lime spheres correspond to the positions where cholesteryl moieties are linked to the platform. Approximate lateral dimensions are 100nm × 60nm. (c) and (d) give the same views of a latter state of the simulation.*

In summary, our design approach involves the integration of a 6HB structure with a rectangular DNA tile, employing multi-crossover DNA staples. We have explored two strategies, T1 and T2, for attaching the nanopore to the platform, and performed molecular dynamics simulations to validate the design.

4.2.2 Characterization of DNA origami's structure

The correct formation of these four models of DNA nanostructures was confirmed using gel electrophoresis and AFM (peak force mode in liquid). For both T1 and T2, the expected AFM structure is a rectangle with a protrusion in the center.

4.2.2.1 Gel electrophoresis of T1 based structures

Figure 4-10 illustrates gel electrophoresis experiments showing the formation of the T1 platform in three different conditions. The assembled DNA origami structures folded efficiently in three different buffer conditions, as detected by 0.8% agarose gel. We observed

that the DNA origami structures of these T1 platforms all clearly appeared in the close distance to the pore (**Figure 4-10**). Their molecular weight is ~ 8kb in size, but because of their complex three-dimensional structure, compared with the simple double-stranded DNA of the ruler, the moving speed will be significantly slower.

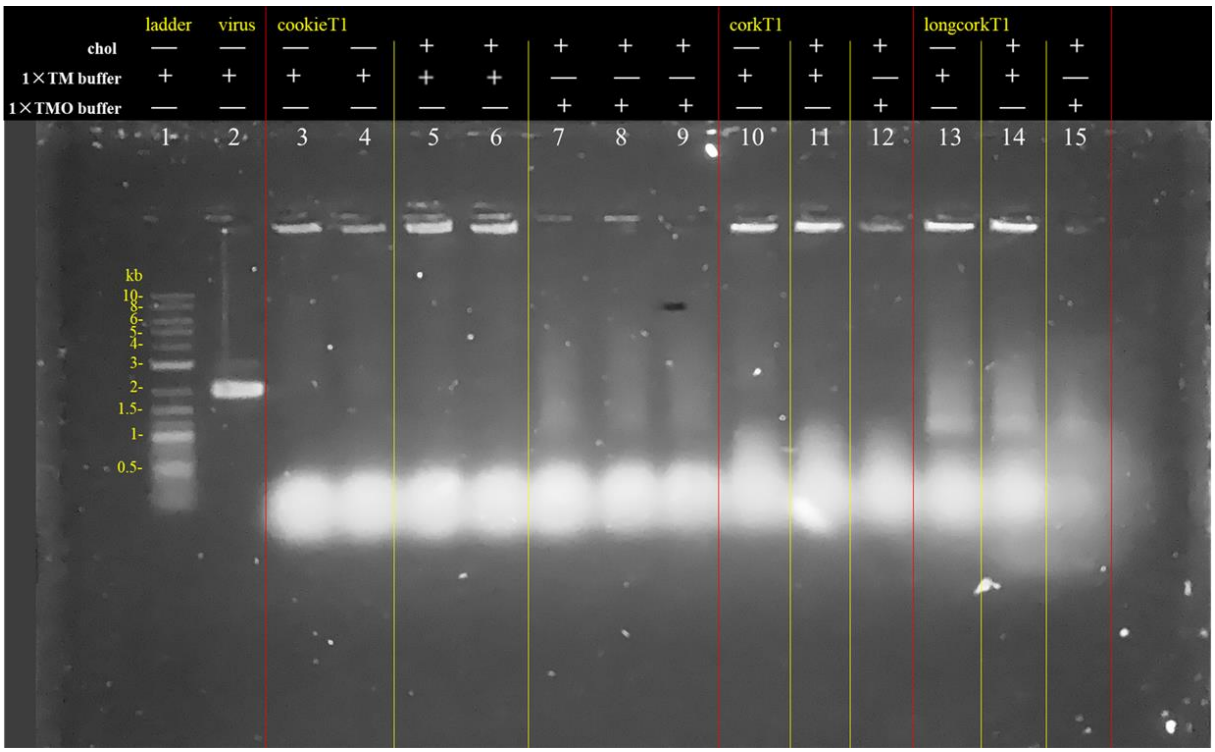


Figure 4-10. Agarose gel electrophoresis image. From left to right: 1kb DNA marker, M13mp18 virus, seven lanes of the mixture of cookieT1, three lanes of corkT1 and three lanes of longcorkT1.
Lane 1: 1kb DNA ladder; lane 2: scaffold virus only; lane 3–9: cookieT1 in different buffer conditions; lane 10–12: corkT1 in different buffer conditions; lane 13–15: longcorkT1 in different buffer conditions. Three different conditions are introduced: TM buffer: origami annealing in TAE/20mM $MgCl_2$ buffer; TM+chol: origami with cholesteryl annealing in TAE/20mM $MgCl_2$ buffer; TMO+chol: origami with cholesteryl annealing in TAE/20mM $MgCl_2$/0.1%OPOE buffer.

Using corkT1 as an example (lane 10–12), it is evident that the migration distance of the corkT1 sample, with added cholesterol in TM buffer (lane 11), is approximately equal to that of the sample without cholesterol (lane 10). To mitigate cholesterol aggregation, we subsequently introduced a surfactant (0.1% octylpolyoxyethylene, OPOE) to the folding buffer. However, it was observed that the corkT1 sample concentration decreased from ~ 20nM to approximately 6nM in TMO buffer (lane 12) compared to the samples annealed in TM buffer (lane 11), while a 'only positive' pattern of jumps was observed with the addition of OPOE in the current recording experiments. A noticeable difference can be observed in the brightness of the bands between lane 11 and lane 12. Agarose gel electrophoresis results exhibit similar bands for the nanopores in the absence and presence of cholesterol during the

annealing step. Additionally, the presence of the surfactant made it challenging to obtain higher concentration of the DNA structure.

4.2.2.2 AFM characterization of corkT1

Figure 4-11 shows large amounts of well-defined rectangles as well as some white irregular-like aggregates for the corkT1 structure in TM buffer (pH 7.4) with the 1:5 ratio of cork and T1 mixture in four distinct AFM images. The height and length profiles indicate that the rectangles have approximately 2nm height and 60×100 nm lateral dimensions, corresponding to the expected dimensions of single nanostructures as reported in Ke's work.[111] However, the number of rectangles with a white region in the center (correct assembly of T1 and cork 6HB) is quite low. Zoomed-in AFM images of some correctly folded single corkT1 nanostructures are displayed in **Figure 4-12 a** and **b**, where the enlarged 2D and 3D AFM image confirm the height of DNA tile is 2nm. The expected height of bright dots corresponding to the cork-like nanopore is ~ 14nm. However, the AFM images of these dots are only ~ 2nm higher than the assembled DNA nanostructures formed by scaffold and staples in the rectangle parts. Due to the sharpness of the nanopore (expected width ~ 6nm), it is possible that the AFM tip deforms the nanopore, which results in a less pronounced bright spot with increased width (~ 10–12 nm). Unexpectedly, the number of well-formed assemblies in TEM images for cork-T1 is much higher. Due to the small number of TEM experiments that we were able to perform, further conclusions would be speculative.

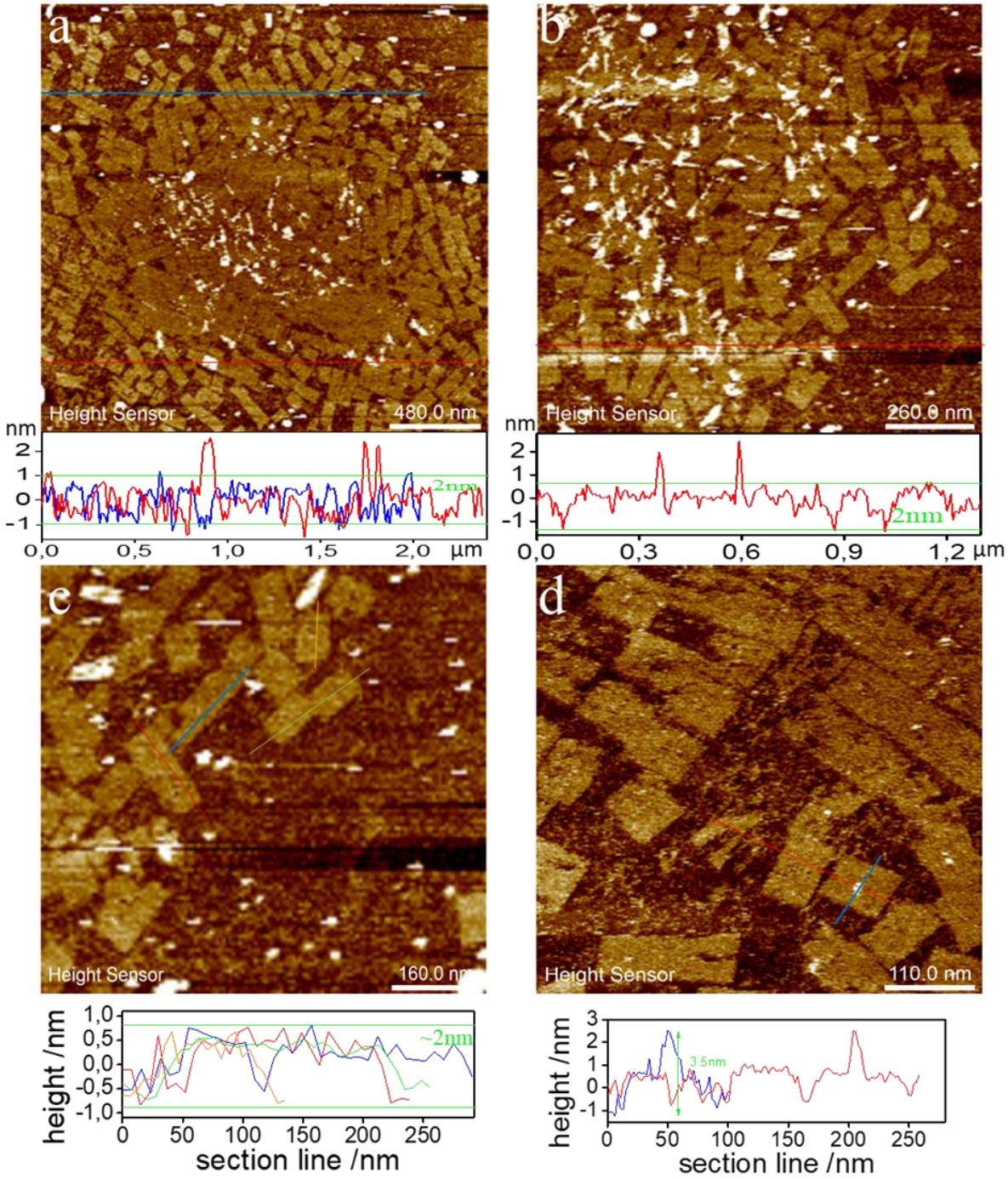


Figure 4-11. The AFM images of corkT1 assembled from the rectangular origami tiles with 6HB nanopore in the center. The addition of 6HB nanopore (cork) appear as a bright spot in the middle of the origami tile.

Based on the four distinct magnified AFM images, the thickness of DNA rectangular tiles was determined to be approximately 2nm, while the height of the nanopore located at the center was measured to be 4.3nm. Some white irregular-like aggregates were the overlapping of DNA origamis or the nanopore structures and their multimers. (a) AFM images are 2.4µm × 2.4µm. (b) AFM images are 2µm × 2µm. (c) AFM images are 820nm × 820nm. (d) AFM images are 550nm × 550nm.

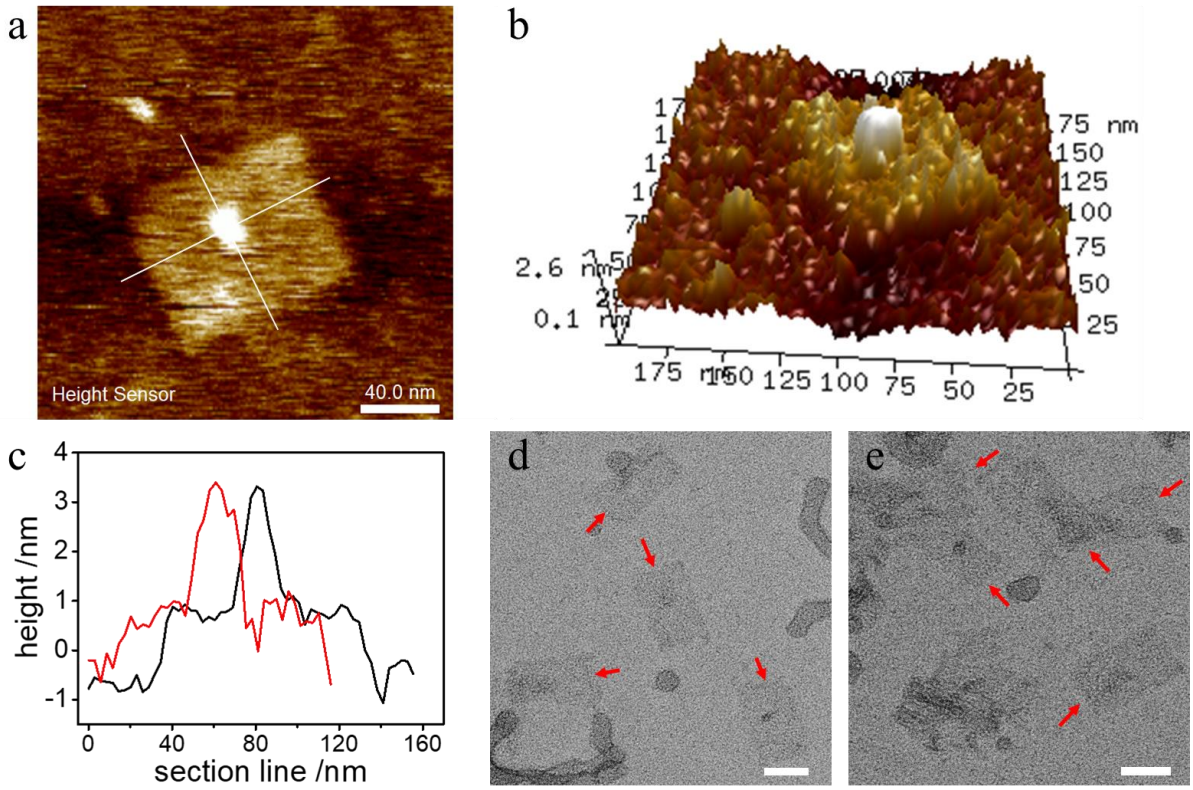


Figure 4-12. Zoom-in AFM images by peak force QNM mode in fluid and TEM images of 500pM corkT1 origami without the addition of T1-cholesteryl.

(a) 2D AFM micrograph (195nm × 195nm) of a single corkT1 structure. The addition of 6HB nanopore (cork) is shown as bright spot on the middle of origami tile. (b) The 3D AFM image for this corkT1. (c) The height profile of the white lines shows the width and length of the corkT1 are 72.9nm and 97.1nm, respectively. The height of the DNA platform is 2nm and the height of the nanopore in the center is 4.3nm. (d) and (e) are TEM images of corkT1, in which the dark spot corresponds to the nanopore. The TEM shows the width and length of the corkT1 are 60nm and 90nm, respectively. The scale bar of TEM images is 50nm.

4.2.2.3 AFM images of longcorkT1

Figure 4-13 shows AFM images obtained for the longcorkT1 structure. The height profile shows the height of DNA objects is 2nm. Through measuring the size of DNA objects in green blocks, the average width and length of rectangle are 77.0±2.7 nm and 96.5±2.9 nm, separately. But we can clearly see from the figure that there are many quarters and halves in the T1 structure. And no obvious insertion of the middle 6HB was observed.

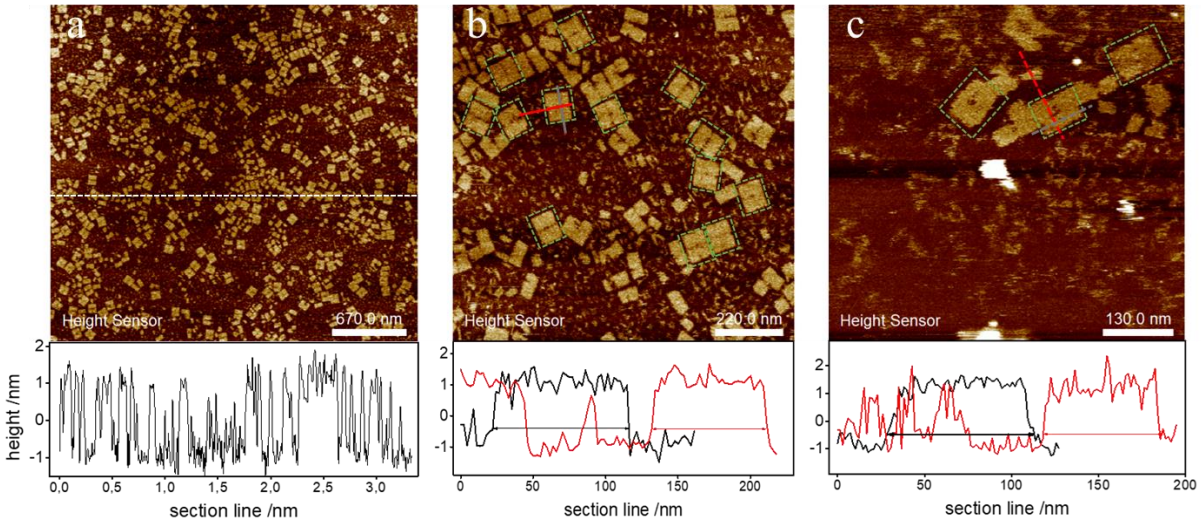


Figure 4-13. The AFM images of longcorkT1 assembled from the rectangular origami tiles with 6HB nanopore in the center. The addition of 6HB nanopore (longcork) should appear as a bright spot in the middle of origami tile.

(a) The height profile shows the height of DNA objects is 2nm. The AFM image is 3.3μm × 3.3μm. (b) The DNA objects in green blocks are rectangle shape and square shape. The average width and length of rectangle are 77.0±2.7 nm and 96.5±2.9 nm, separately. The average width and length of square are 81.6±0.5 nm and 83.0±1.7 nm, separately. The height profile is measured from the red and grey line. The size of object is 77nm wide and 95nm long. The AFM image is 1.1μm × 1.1μm. (c) The height profile of this zoomed micrograph is measured from the red and grey line. The size of object is 67.1nm wide and 91.0nm long. The AFM image is 665nm × 665nm.

4.2.2.4 AFM images of cookieT1

The AFM images in **Figure 4-14** clearly indicate the well-formation of cookieT1 nanopore. Comparing corkT1 and longcorkT1, almost each monomer of cookieT1 has a bright dot in the center. The height profile shows the height of DNA objects is 2nm. Through measuring the size of DNA objects (monomers), the average width and length of rectangle are 68.4±4.7 nm and 91.5±9.8 nm, separately. And we can clearly see from the figure that there are many bright dots in the cookieT1 structure. And the average height of bright dots representing 6HB in the middle is 5.8nm. The well-formed structures are observed clearly.

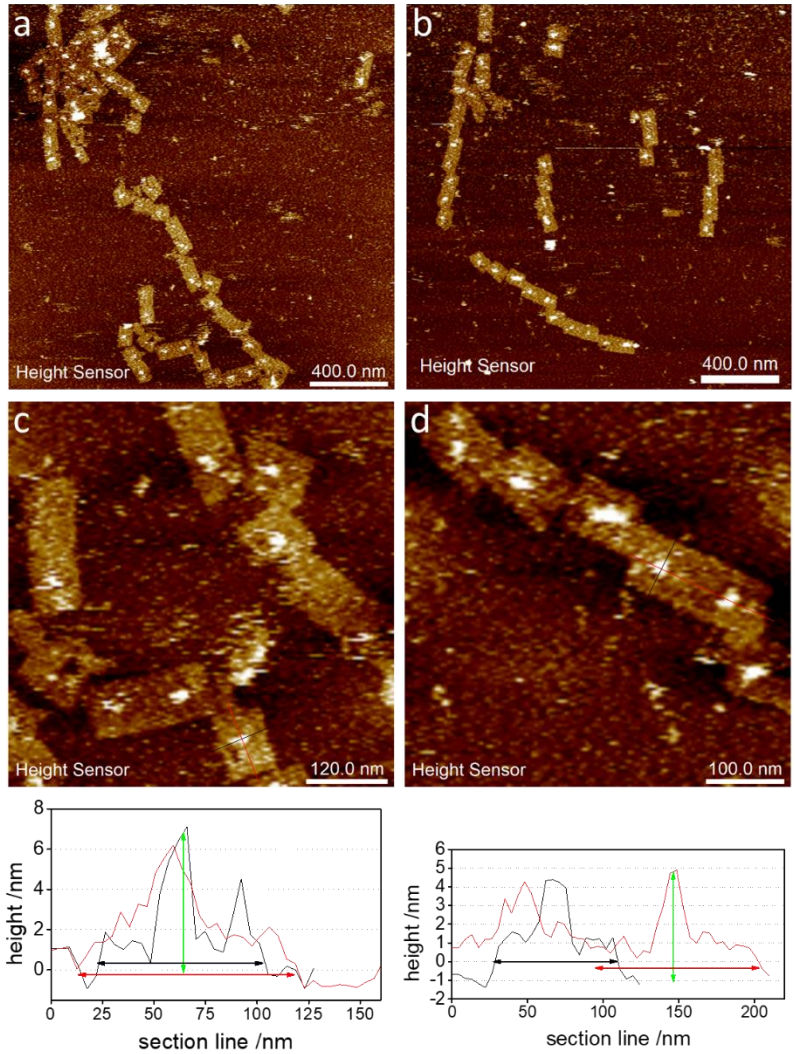


Figure 4-14. AFM images of cookieT1 assembled from the rectangular origami tiles with 6HB nanopore in the center. The formation of 6HB nanopore (cookie) should appear as a bright spot on the middle of origami tile.

(a) AFM image is 2μm × 2μm. (b) AFM image is 2μm × 2μm. (c) The height profile of this zoomed micrograph is measured from the red and black line. The size of object is 72.1nm wide and 100.1nm long. The height of DNA object and the bright dot in the middle are 2nm and 5.85nm, separately. The zoom-in AFM image is 570nm × 570nm. (d) The height profile of this zoomed micrograph is measured from the red and black line. The size of object is 68.8nm wide and 92,9nm long. The height of DNA object and the bright dot in the middle are 2nm and 5.87nm, separately. The zoom-in AFM image is 500nm × 500nm.

4.2.2.5 AFM images of v7T2

Figure 4-15 and **Figure 4-16** show that v7T2 structures are mostly well-formed. The height of DNA tiles is about 2nm same as T1. The size of T2 is 64.6×98.5 nm, which is slightly larger than T1 (60×90 nm), because of a modification of helix arrangement in T2. Comparing **Figure 4-15c** and **Figure 4-16c**, the height of bright dots is, respectively, 3.75nm

and 5.07nm. Although both images have been performed under similar conditions (buffer, cantilever type, imaging mode), the measured heights of the nanopore are significantly different. This is probably an indication of the influence of the state of the imaging tip when evaluating very irregular shapes, such as the protruding nanopore from the flat platform.

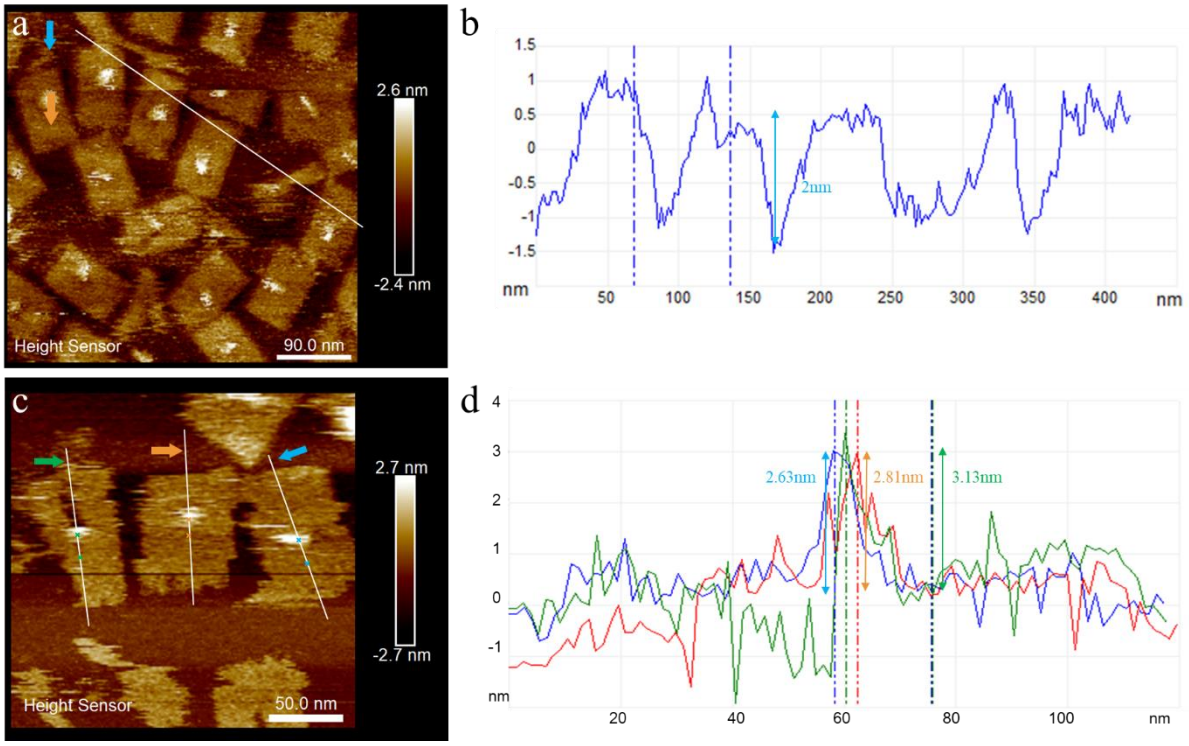


Figure 4-15. The AFM images of v7T2 assembled from the rectangular origami tiles with 6HB nanopore in the center, without the addition of T2-chol strands. The formation of 6HB nanopore (v7) should appear as a bright spot on the middle of origami tile.

(a) The height profile shows the height of DNA tile is ~2nm. AFM image is 430 × 430 nm. (c) The average height of 6HB nanopore (bright dot in the middle) is 3.75nm, however, the DNA tiles are damaged here. AFM image is 233 × 233 nm.

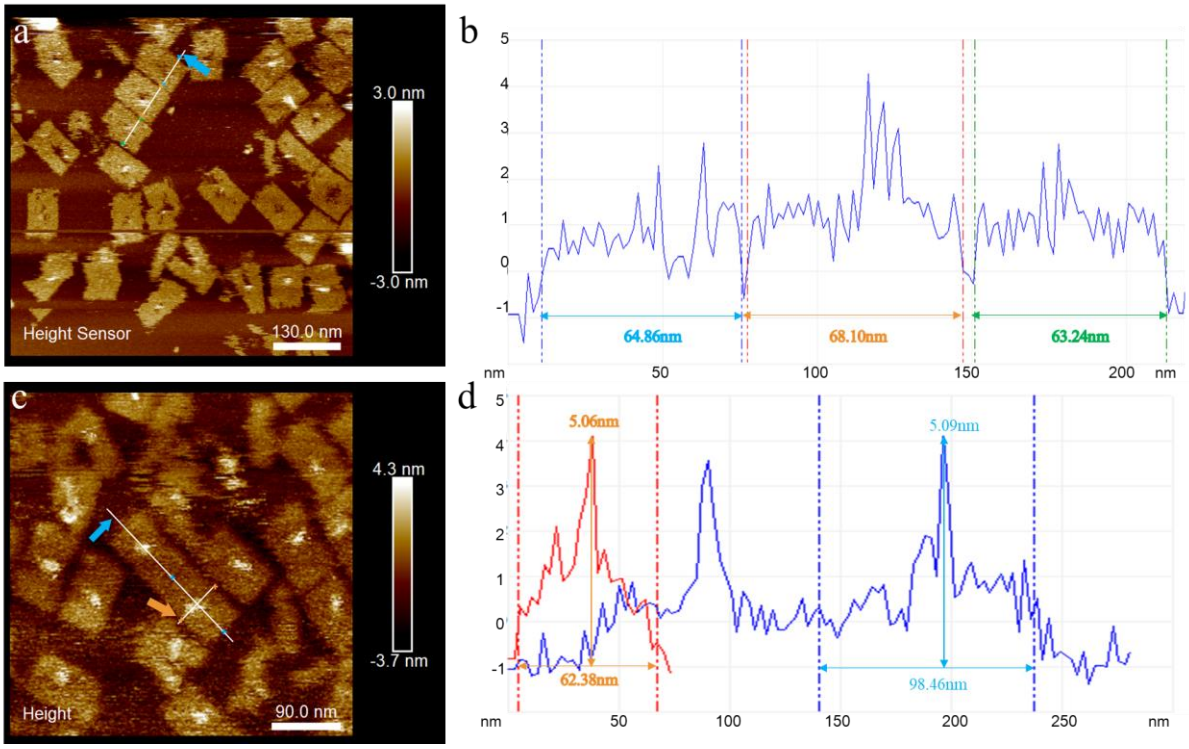


Figure 4-16. The AFM images of v7T2 assembled from the rectangular origami tiles with 6HB nanopore in the center, without the addition of T2-chol strands. The formation of 6HB nanopore (v7) should appear as a bright spot in the middle of origami tile.

(a) The sections of this image are to measure the width of DNA tiles. The average width of v7T2 is 64.6nm. AFM images are 451 × 451 nm. (b) Section two example of v7T2 in the middle of the view. The average leghth of v7T2 is 98.5nm. AFM images are 664 × 664 nm.

All these AFM images show DNA origami without the addition of cholesteryl binding. From section analysis of AFM images, the average width of three v7T2 structures is 65.4 ± 4.08 nm, and a typical size of v7T2 as shown in **Figure 4-16c** is 62.38nm wide and 98.46nm long with a 5.09nm high nanopore in the center. From a wider image view, the rectangle shapes are almost monomers and separate thoroughly. In contrast, from a zoomed view, because of the possible reason of the scanning damage by AFM probes, we choose three origami and found their section analysis have almost coincident characterization.

The overall conclusion from the AFM images performed on the different four structure types is that sharing the scaffold between the nanopore and the origami platform is the best way to obtain a well-formed nanopore embedded in a platform. Although this could seem to be obvious, our experience with current recordings is somewhat opposite to this observation, as explained later in this chapter.

4.2.2.6 Results of QCM-D experiments

We have previously explored the correct formation of the nanopore-platform complex using gel electrophoresis, AFM and TEM. However, all these techniques considered the nanopore-platform in isolation from any lipid bilayer. As explained in the introduction, our goal is to measure the interaction between nanopore-platforms and lipid bilayers. In this section, in order to get an independent estimate of the number of DNA devices that attach to a lipid bilayer, we present Quartz Crystal Microbalance with Dissipation monitoring (QCM-D) experiments. A QCM sensor is made of a thin disc of quartz crystal sandwiched between two electrodes. By measuring the frequency shift and dissipation of a crystal excited at its resonance frequency, it is possible to deduce the mass and viscosity of deposited samples.

To quantify the number of nanopores that can bind to a lipid bilayer, we need first to form a lipid bilayer on top of the quartz crystal. We have prepared vesicles from the same lipid mixture (DphPC: POPC = 7:3) as the one that will be used in current recordings. The QCM-D crystal sensors were precoated with a 50mm thick Au film or an additional 30nm thick SiO film on top of the gold, deposited by e-beam evaporation. Before use, to decompose and remove organic contaminants on the sensor, it was rinsed by pouring pure water and drying by nitrogen gas flow, followed by UV /ozone treatment for 20 min.

Preliminary QCM-D experiments were performed to determine optimal conditions to obtain supported lipid bilayers. The baseline measurement of QCM-D is firstly stabilized after flowing 1×B buffer for 20 min. The addition of DphPC: POPC (7:3, molar ratio) vesicles (**Figure 4-17**) showed a clear shift in frequency, indicating that these vesicles absorbed onto the substrate and formed a layer of unruptured vesicles. After the formation of a layer of unruptured vesicles, 1×B buffer wash with different concentrations of OPOE were performed. We found 0.1% –0.2% can promote the transition from vesicles to supported lipid bilayers (SLB).

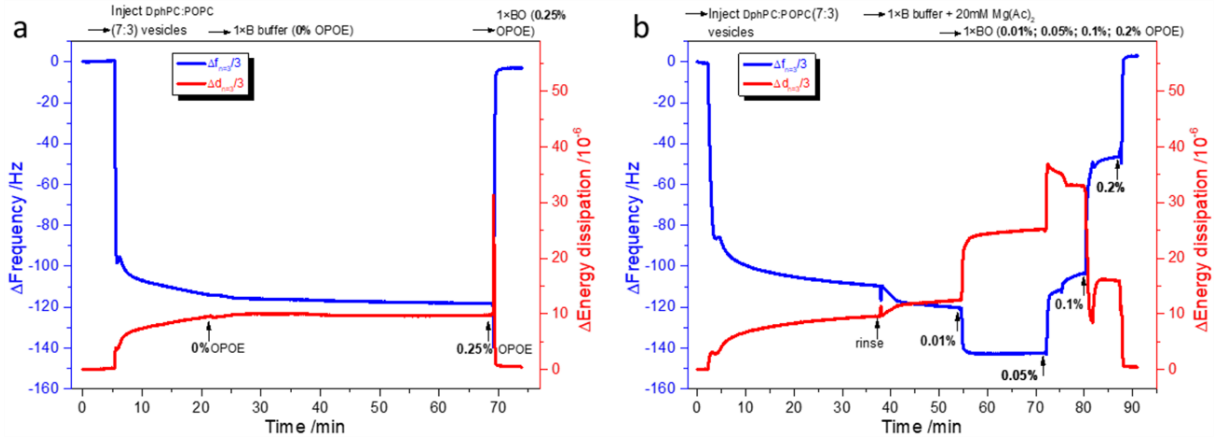


Figure 4-17. Frequency shift (blue) and dissipation shift (red) recorded on top of a SiO2-coated QCM-D sensor upon addition of lipid vesicles.

The baseline measurement is firstly stabilized after flowing 1×B buffer for 20 min. DphPC:POPC (7:3, molar ratio) vesicles (1mg/mL in 1×B buffer) were injected for 2 min at a speed of 0.31 mL/min and these vesicles were absorbed onto the substrate and formed a layer of unruptured vesicles. After the formation of a layer of unruptured vesicles, 1×B buffer wash with different concentrations of OPOE were performed. (a) 0% OPOE and 0.25% OPOE in 1×B were added, and the final frequency and dissipation values were -118Hz and 9.7×10-6, -3Hz and 0.45×10-6, respectively. Rinsing with 1×B (0% OPOE) only, the frequency and dissipation signals remain constant for a long time. Thereafter, rinsing with 1×BO (0.25% OPOE), the frequency and dissipation signals and the mass diversion of vesicles on the QCM-D sensor are significantly reduced. (b) Restart the measurement and continue to flow different buffer into the chamber, the frequency and dissipation stabilized at approximately -109Hz and 9.4×10-6 (lipid vesicles), -120Hz and 12.3×10-6 (1×B with 20mM MgAc2), -142Hz and 25×10-6 (1×BO (0.01% OPOE)), -103Hz and 33×10-6 (1×BO (0.05% OPOE)), -46Hz and 16×10-6 (1×BO (0.1% OPOE)), 2.8Hz and 0.3×10-6 (1×BO (0.2% OPOE)), respectively.

Besides, different sizes of vesicles were tested as shown in Figure 4-18**a** (57nm) and Figure 4-18**b** (142nm), respectively formed by sonication and extrusion. **Figure 4-18** shows the size distribution from DLS experiments that confirm that the input SUVs have a well-defined size. Upon 57nm and 142nm of DphPC/POPC (7:3) vesicles absorbed, the frequency decreases and equilibrates at a level of -44Hz and -114Hz, respectively. The dissipation increases monotonically and equilibrates at an elevated level of 4.9×10^{-6} and 16.2×10^{-6}. Subsequently, the addition of 0.1% OPOE in 1×B buffer triggers the SUV disruption and the formation of solid supported bilayers (SLB). The two plateau observed before and after rinsing with a solution containing 0.1% OPOE surfactant correspond respectively to vesicles and supported lipid bilayer. Note that the value Δf = - 20.2 Hz and Δf = -22 Hz of final frequency shift is mostly independent of the initial size of vesicles. To determine the bilayer thickness, the Sauerbrey equation (**equation 3**) can be used, yielding a value ~ 3.7nm (assuming lipid density is that of water). These data confirm that lipid vesicles burst upon addition of a minute concentration of surfactant, forming a lipid bilayer on top of the quartz

crystal. Therefore, we set up the parameters of 142nm DphPC: POPC (7:3, molar ratio) vesicles and 0.1% OPOE in 1×B buffer in further experiments.

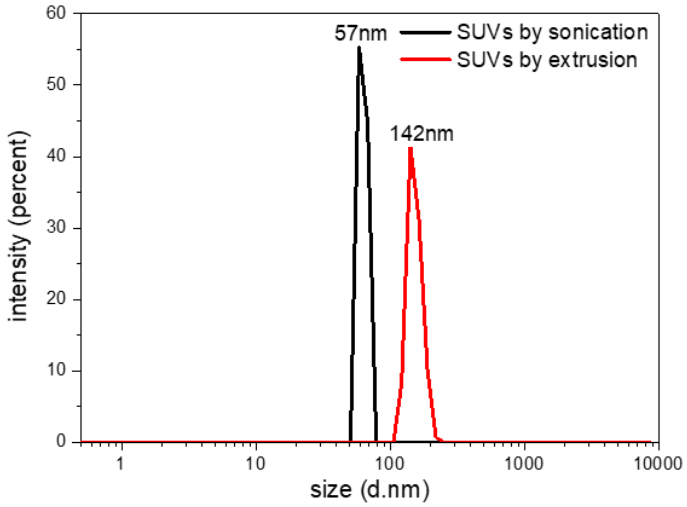


Figure 4-18. *DLS measurements for the SUVs obtained with two different methods: sonication and extrusion.*

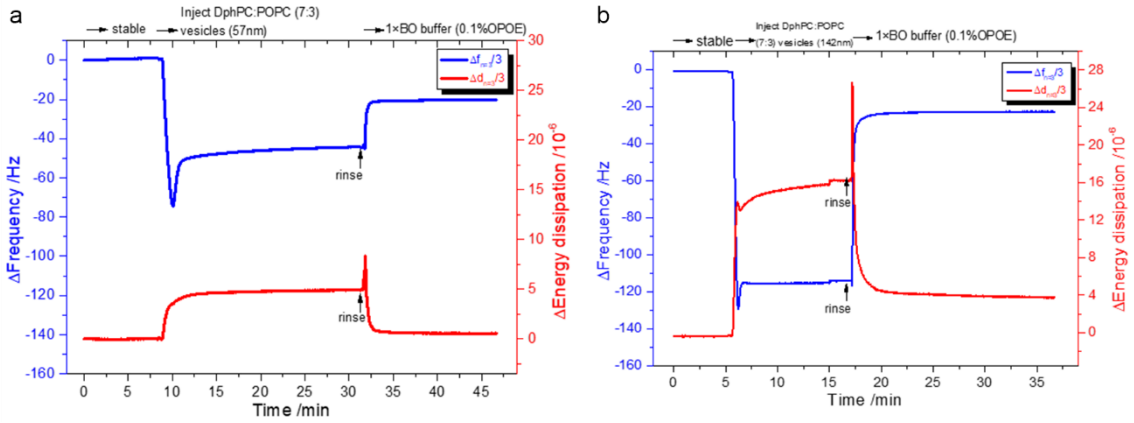


Figure 4-19. QCM-D responses for the deposition of different lipid vesicle sizes. Frequency shift (blue) and dissipation shift (red) recorded on top of a SiO2-coated QCM-D sensor.

Upon ~57nm (a) and ~142nm (b) of DphPC/POPC (7:3) vesicles absorbed, the frequency decreases and equilibrates at a level of -44Hz (a) and -114Hz (b). The dissipation increases monotonically and equilibrates at an elevated level of 4.9×10-6 (a) and 16.2×10-6 (b). These Δf and ΔD values are representative of the formation of a layer of supported vesicular layer (SVL), as reported previously by Nam-Joon group [168] for DOPC/DOPS (1:1) vesicles. In this case, vesicles only absorbed on the substrate but not rupture to form a supported lipid bilayer (SLB). Therefore, OPOE is used as a vesicle-destabilizing agent. In presence of 0.1% OPOE, 1×BO (0.1% OPOE) is injected to the measurement chamber, which is characterized by increases of frequency at -20.2Hz (a) and -22.7Hz (b) and decreases of dissipation at 0.57×10-6 (a) and 3.8×10-6 (b).

In **Figure 4-19** are represented the evolution of frequency shift and dissipation for three DNA samples that flow after the formation of SLBs, respectively 1nM random DNA with no cholesterol, 0.5nM and 10nM v7T2 DNA nanopore bound to cholesteryl. While the first two signals are difficult to interpret, the 10nM sample shows a clear trend in both frequency and dissipation corresponding to an absorption process. The slope of signal curve increases with the increase of nanopore concentration ranging from 0 to 10 nM, which suggests that the speed of binding events is proportional to the target concentration. While the concentration of nanopores is increasing, the speed of frequency shift in the first 1 min is increasing 9.1 times as illustrated in **Figure 4-20**.

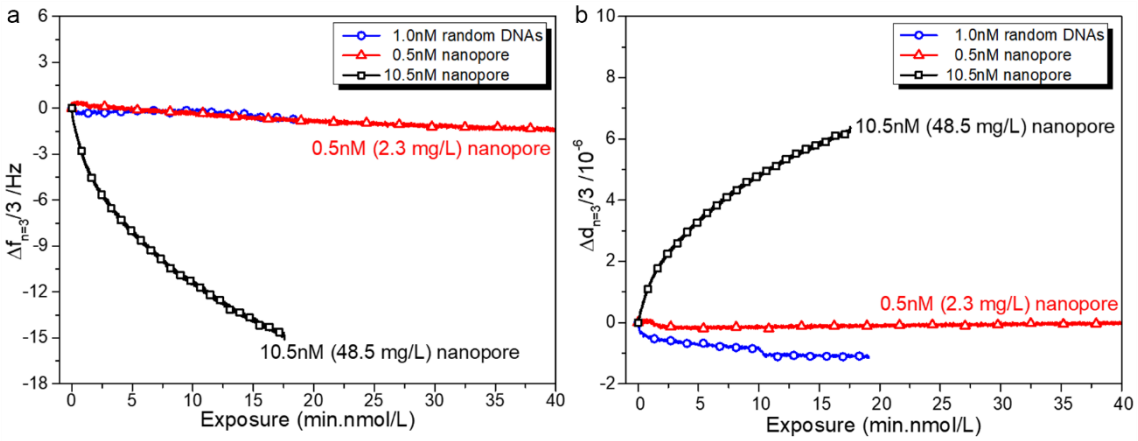


Figure 4-20. Frequency shift (a) and dissipation shift (b) were recorded on top of a SiO2-coated QCM-D sensor versus exposure of v7T2 origami at concentrations of 2.3 and 48.5 mg/L.
Upon absorption of 1.0nM random DNA staples (blue), 0.5nM v7T2 nanopore (red), and 10.5nM v7T2 nanopore (black) in the SLB (DphPC/POPC (7:3)), the frequency of the 10.5nM nanopore decreased by more than -15Hz (a), and the dissipation of the 10.5nM nanopore increased by more than 6.4×10^{-6} (b). The concentrations of the 0.5nM and 10.5nM v7T2 nanopore were 2.3 and 48.5 mg/L, respectively.

Figure 4-21 shows the evolution of frequency shift and dissipation of different overtone orders (n= 3, 5, 7, 9, 11, 13). By assuming that the the liquid density and viscosity are 1.5g.cm^{-3} and 0.7 mPas, a simultaneous fit of both data in (**Figure 4-21**) with viscoelastic model (**equation 4**) assuming shear modulus (G′, G″) of (41kPa, 132 kPa), yields a thickness estimate of 4.9nm for the absorbed film. The thickness difference with respect to a lipid only bilayer is therefore 1.2nm. This is compatible with a DNA nanopore device covering roughly half of the surface of the lipid bilayer, indicating that under these conditions, the binding of platform attached DNA nanopores is quite efficient. The QCM-D data of the addition of 10nM DNA are given in **Table 4-1**.

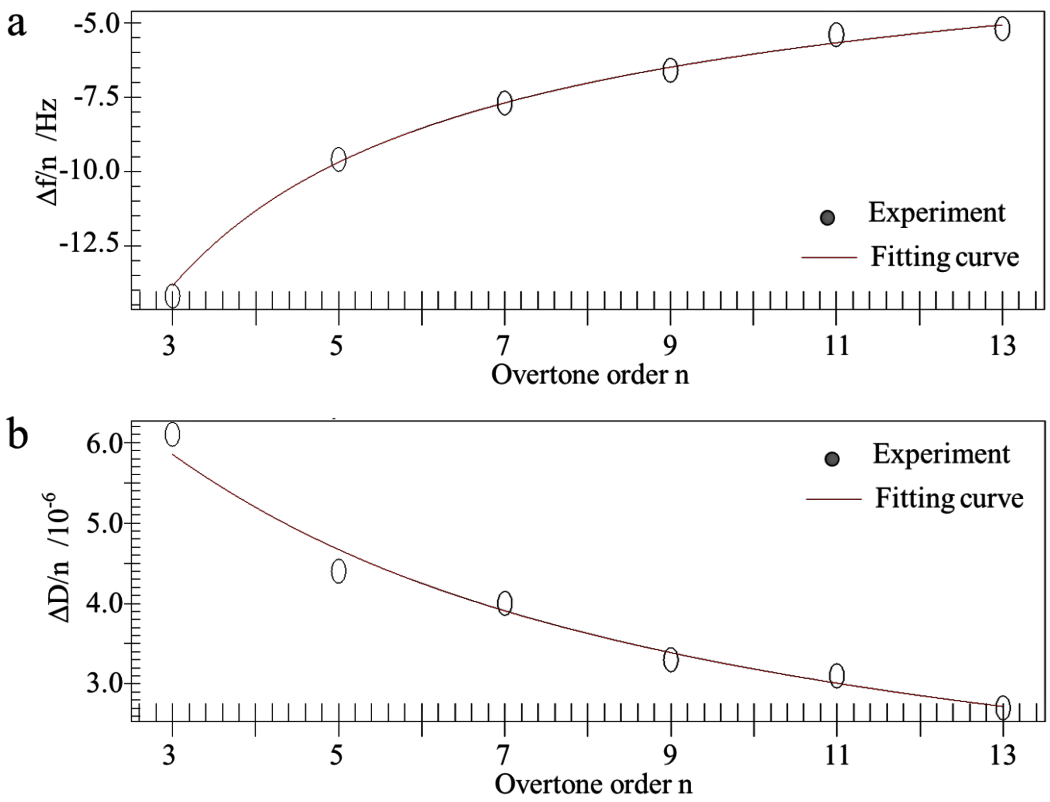


Figure 4-21. *Fitting of frequency and dissipation evolution of data in* ***Figure 4-20*** *and* ***Table 4-1*** *(DNA device concentration 10.5nM) using viscoelastic model, assuming a liquid density and viscosity of 1.5g.cm^{-3} and 0.7 mPas, shear modulus (G') and shear modulus imaginary (G'') of 41kPa and 132 kPa, and 4.9nm thickness of DNA origami in the absorbed film.*

In order to image the interaction of DNA origami structures with SLB membranes, we kept the chip after QCM-D experiments in buffer solution and used AFM (Tapping mode in fluid). **Figure 4-22** shows the phase signal of DNA aggregations for the v7T2 nanostructures in 1×BO (0.1% OPOE, pH 7.4) buffer. In the phase mode, stiff and soft materials show different phases in the AFM images (which is made at constant amplitude). The softer materials on the top of the QCM-D chip reveal the existence of rectangle like shapes with approximate width ~ 60nm and length ~100 nm, corresponding to DNA origami monomers. Note that the shapes observed in these images look blurred as compared to the ones obtained with cholesterol-less DNA nanopores. Therefore, besides the fact that the imaging here is performed on structures that are on top of a fluid substrate (the SLB), it is reasonable to think that part of them is also covered with cholesterol and/or OPOE surfactant.

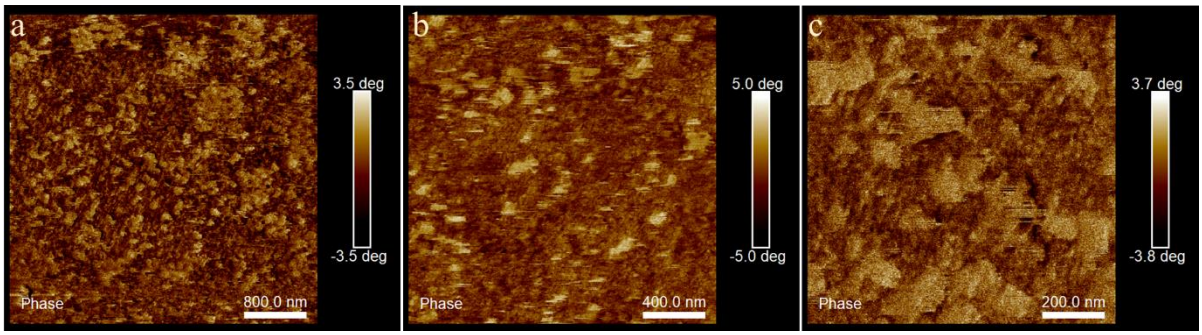


Figure 4-22. *After QCM-D experiments, the AFM imaging of Q-sensor chip with different resolutions. The AFM imaging of Supported lipid bilayer on the surface of quartz crystal.*

We also use AFM to study SLB formation on the mica and after add the DNA nanopores. **Figure 4-23** shows three examples of SLB on mica, it is clear that the thickness of 4nm SLB is constructed. But the DNA devices is preferred to attach to the mica surface than the SLB. It shows that the affinity of DNA devices and mica is higher than it with lipid bilayers. Maybe because the mica has positive charges but the DphPC and POPC lipids we choose to form lipid bilayer are neutral.

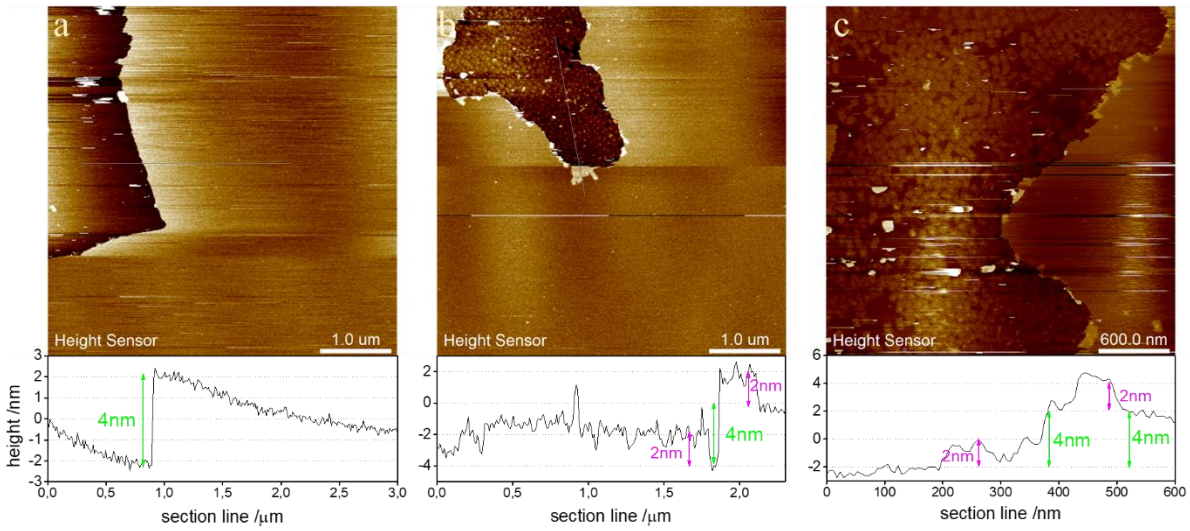


Figure 4-23. The AFM imaging of Supported lipid bilayer on mica.
(a) The height profile shows the height of lipid bilayer on mica is ~ 4nm. 1mg/mL DphPC/POPC (7:3) lipid vesicles are added on the mica. AFM image is 5 × 5 μm. (b) The height profile shows the height of v7T2 DNA tiles are ~2 nm and lipid bilayer on mica is ~ 4nm. After the bilayer formed in mica, 1nM of v7T2 nanopores are added and incubated with lipid bilayers. AFM image is 5 × 5 μm. (c) The zoom-in images of (b). AFM image is 3 × 3 μm.

Table 4-1. The resonance frequency f_n and energy dissipation D_n of recording data in Figure 4-21 (SLB with 10nM DNA device) in different overtone orders (n = 3, 5, 7, 9, 11 and 13).

n	f_0	f_n	Δf_n	D_0	D_n	ΔD_n
1	21	53.6	-32.6	-3.9	6	9.9
3	21.1	35.3	-14.2	0.1	6.2	6.1
5	20.8	30.4	-9.6	0.5	4.9	4.4
7	20.8	28.5	-7.7	0.2	4.2	4
9	20.6	27.2	-6.6	0.2	3.5	3.3
11	20.5	25.9	-5.4	0.4	3.5	3.1
13	20.6	25.8	-5.2	0.2	2.9	2.7

The overall conclusion from this QCM-D experiments is that DNA nanopore devices, coated with cholesterol, bind quite efficiently to supported lipid bilayers. Note that the flux conditions imposed in QCM-D experiments are not the same as the ones observed in DIB experiments. In the former, unbound structures are swept whereas in DIB experiments, structures diffuse freely, until an electric field is imposed. Despite this, we can conclude that the binding process should not be a limiting one in current recording experiments.

4.2.3 Controlled perforation of lipid bilayer

In this section, we discuss the results obtained in DIB experiments with the previously described DNA nanopore devices. In the following experiments, the basic configuration is always the same: after filtering, nanopores are kept at 4℃ until the experiment is done. Before the experiment, the nanopore is incubated with a known concentration of input strand. Two droplets are brought together, the left droplet only contains buffer (1M KCl), the right droplet contains the nanopore incubated with input strand, in the same buffer. A detailed description of the experimental conditions is given in section 4.3. It is also good to have in mind the predictions given by the simplest model for the resistance of a nanopore. This model[169] is based on the assumption that ions traverse a homogeneous cylinder with resistivity ρ, length L and radius a:

$$\mathbf{R}_{pore} = \frac{\boldsymbol{\rho}(\boldsymbol{L} + \frac{\boldsymbol{\pi a}}{\mathbf{2}})}{(\boldsymbol{\pi a^2})}$$

(*equation 5*)

In this expression, the L-dependent term corresponds to the resistance of the nanopore itself, and the correction πa/2 is the resistance due to the difficulty of accessing the nanopore

entrance (access resistance). In this approximation, the conductance (in nS) of a nanopore of 1nm inner radius and L nm length in a 1M KCl (conductivity = 10.9 S/m) salt solution, is given by the expression:

$$\boldsymbol{\sigma = \frac{10.9 \times \pi}{L + 1.5}}$$

(*equation 6*)

4.2.3.1 Analysis for corkT1

In this structure, the two halves on both sides of the stem-loop mechanism are quite symmetric (approximate length 7nm). Also, because the nanopore is only bound to the origami platform by Watson-Crick pairings, the proportion of well-formed structures is quite low (as demonstrated by AFM imaging). Despite this, this is the structure that we first studied and the one that gave, paradoxically, the highest number of conductance jumps. **Figure 4-24** shows the influence of adding a surfactant (OPOE) in the folding buffer. When the annealing is done in the absence of cholesterol strands, no surfactant is needed in the folding buffer. Cholesterol is simply added by incubating the filtered origamis with the cholesterol bound strands, previously heated, in the presence of OPOE. When the annealing is done in the presence of cholesterol strands, we noticed that the insertion rate (number of experiments where any jump is recorded versus total number of recordings) increased to ~30% provided OPOE was added in the folding buffer. Because the length of the binding section to cholesterol strands is 21nt, it is expected that the recognition and hybridization of cholesterol strands to the partially formed origami can take at rather high temperatures (~ 50°C). **Figure 4-24** shows that this hybridization is less efficient when the origami is already formed, even if cholesterol strands are dispersed (both by heating and surfactant).

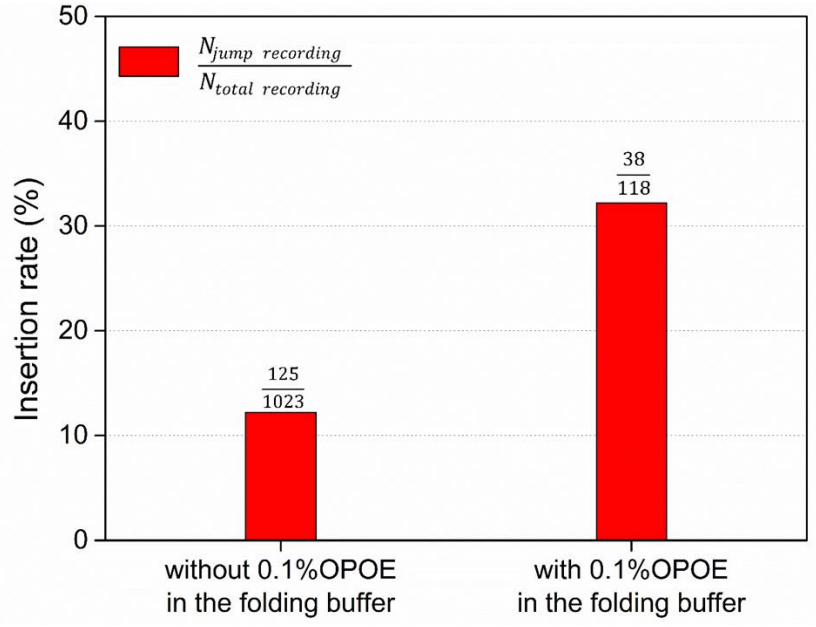


Figure 4-24. *A graph depicting the dependence of the insertion rate is presented. In the absence of 0.1% OPOE in the folding buffer, the insertion rate was 12.2% (N_{jump}/N_{total} = 125/1023). However, with the addition of 0.1% OPOE to the folding buffer, the insertion rate increased significantly to 32.2% (N_{jump}/N_{total} = 38/118).*

Figure 4-25 shows that not only the insertion ratio is influenced by the moment at which surfactant is introduced in the system. As previously stated, each time a nanopore inserts into the membrane, a jump in conductance is expected. **Figure 4-25 a** and **b** show both positive and negative jumps, whereas **Figure 4-25 c** and **d** show only positive jumps. Two possible mechanisms could explain negative jumps: a desorption of the nanopore-platform from the membrane, or a tilting of the nanopore, the platform remaining bound to the membrane. In terms of energy involved in both processes, the second seems a more plausible one. A third, intermediate scenario could be that of a nanopore-platform poorly bound to cholesterol strands, for which the desorption mechanism would not be very costly in terms of energy. In any case, the three possibilities seem to be ruled out by the use of surfactant in the folding buffer, because under these conditions, mostly positive jumps could be observed.

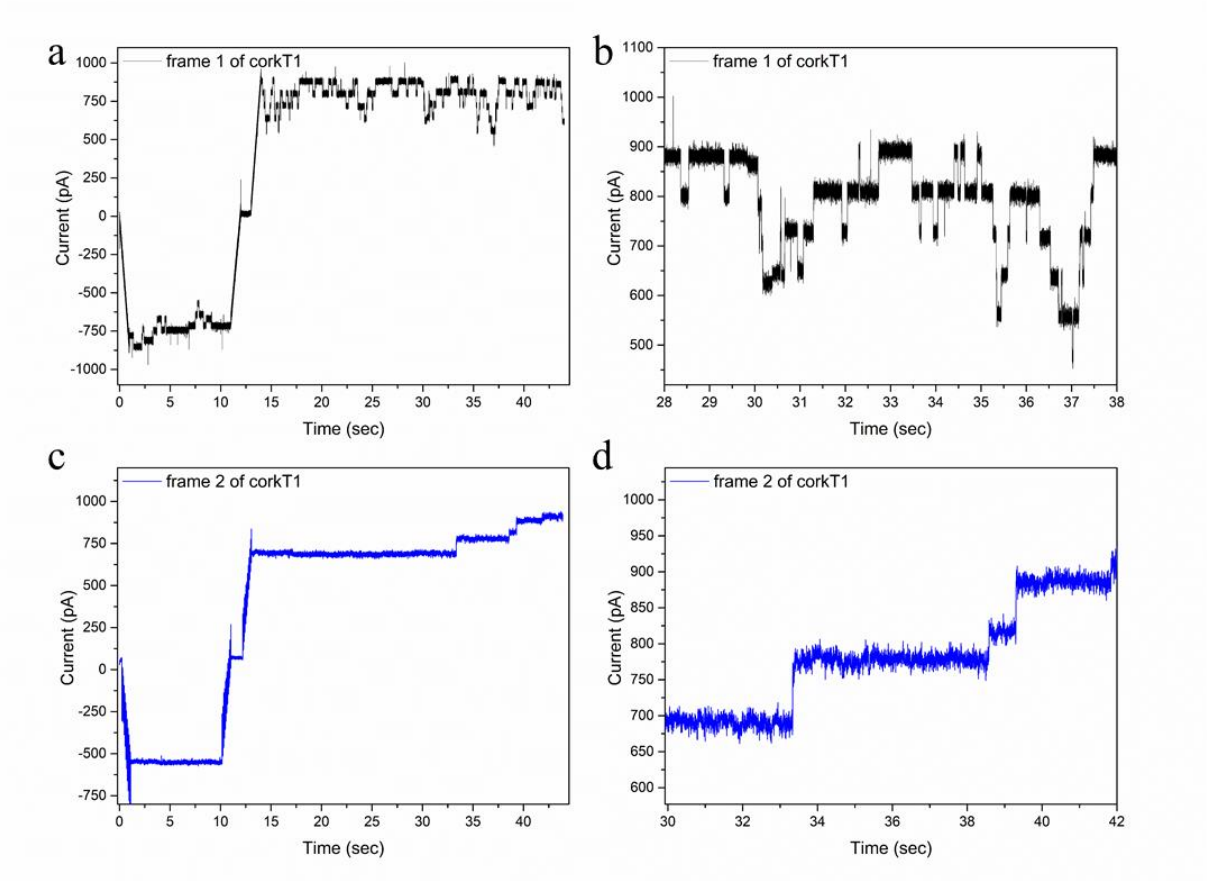


Figure 4-25. Current recording in a fixed potential experiment of the corkT1 structure.

(a) and (b) Surfactant was not included in the folding buffer. (c) and (d) Surfactant was included in the folding buffer. In these recordings, a negative potential (-50mV) is applied for the first ~12 s, a transition between -50mV and +50mV is then performed, finally +50mV are applied to the end of the recoding (~50s). The same cycle is applied repeteadly. (a) [corkT1] = 1nM; [input strand] = 0.1nM. (b) [corkT1] = 2nM; [input strand] = 1nM.

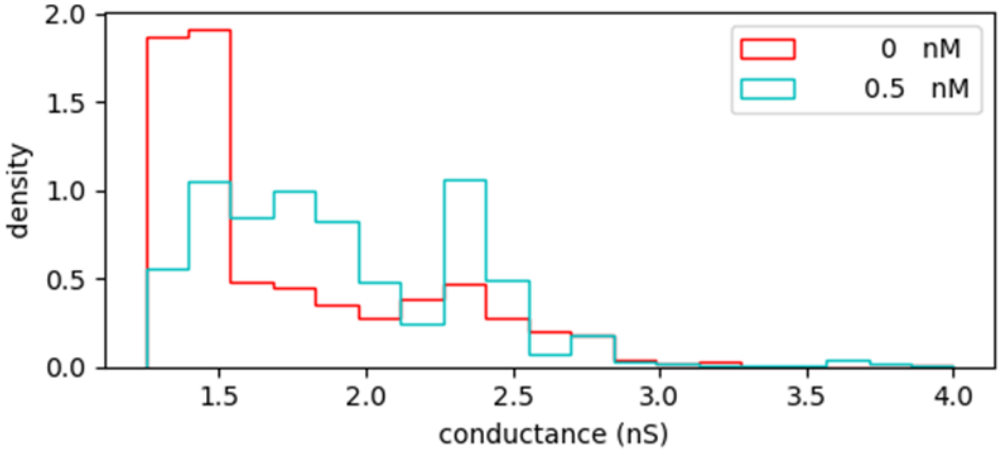


Figure 4-26. *Density histogram of the conductance jumps observed in several independent recordings of the corkT1 structure, in the absence (red line) or the presence (blue line) of input signal sequence. [corkT1] = 1nM.*

Figure 4-26 shows the conductance histogram obtained in the absence (or not) of the input signal, the input sequence. The shape of this histogram is very similar to the one we had in chapter 3, with the nanopore not embedded in a platform, that is, the overlap between open and closed states is very high. Note that the two main peaks observed here do not correspond well to the values given by equation (**equation 6**), respectively 2.2nS (L = 14nm) and 4nS (L = 7nm) for the closed and open states. This simple model clearly overestimates the ratio between open and closed conductances, as would be expected from the rough approximation done in this model, where the influence of the second half (the one not inserted into the platform) is completely overlooked.

4.2.3.2 Analysis for v7T2

The DNA device called corkT1 and the folding protocol where surfactant is included in the annealing buffer provide with two interesting properties: high(er) rate of insertion and predominance of positive jumps (as compared to mixtures of positive and negative jumps). However, the overlap between histograms when the input signal is present or not is still important. Here, we want to avoid this by increasing the length of the nanopore in such a way that, when closed, the conductance is lower. The scadnano diagrams (**Figure 4-3d**) give approximate length 30nm (respectively 12nm) for the nanopore in the ‘closed’ state (respectively ‘open’ state). The corresponding theoretical conductances, according to (**equation 6**) are 1.08nS (L = 30nm, 88nt long) and 2.53nS (L = 12nm, 35nt long). Note that the length of the upper half is the one that has been mostly elongated. Furthermore, taking

into account the difference in the folding success ratio between the T1 and T2 strategies, as shown by AFM images (**Figure 4-11** and **Figure 4-16**), here we preferentially explore the behaviour of the v7T2 structure. **Figure 4-27** illustrates again the different behaviour when surfactant is, or is not, included in the folding buffer.

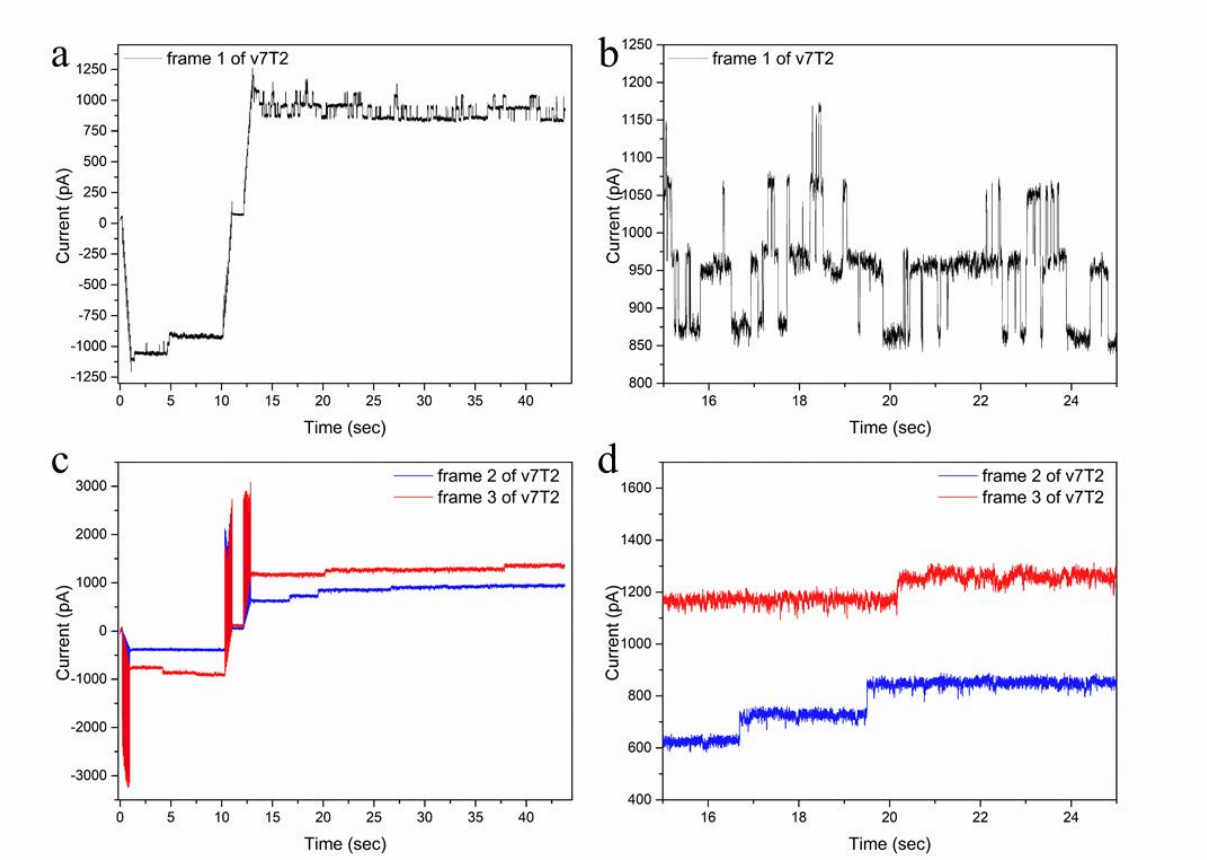


Figure 4-27. Current recording in a fixed potential experiment of the v7T2 structure.
(a) and (b) Surfactant was not included in the folding buffer. (c) and (d) Surfactant was included in the folding buffer. In these recordings, a negative potential (-50mV) is applied for the first ~12 s, a transition between -50mV and +50mV is then performed, finally +50mV are applied to the end of the recoding (~50s). The same cycle is applied repeteadly. (a) [corkT1] = 0.1nM; [input strand] = 0.5nM. (b) [corkT1] = 2nM; [input strand] = 1nM.

The first result is the difference in the number of observed jumps per frame (one frame corresponds to ~50s experiment), as shown in **Figure 4-28**: this rate is almost two orders of magnitude lower in the absence of input signal sequence. The second result is the absence of small jumps when the input is absent. Indeed, the size of these small jumps is comparable to that of recording noise. Therefore, although not perfect, we conclude that the v7T2 strategy is quite satisfactory, at least regarding the separation of conductance peaks.

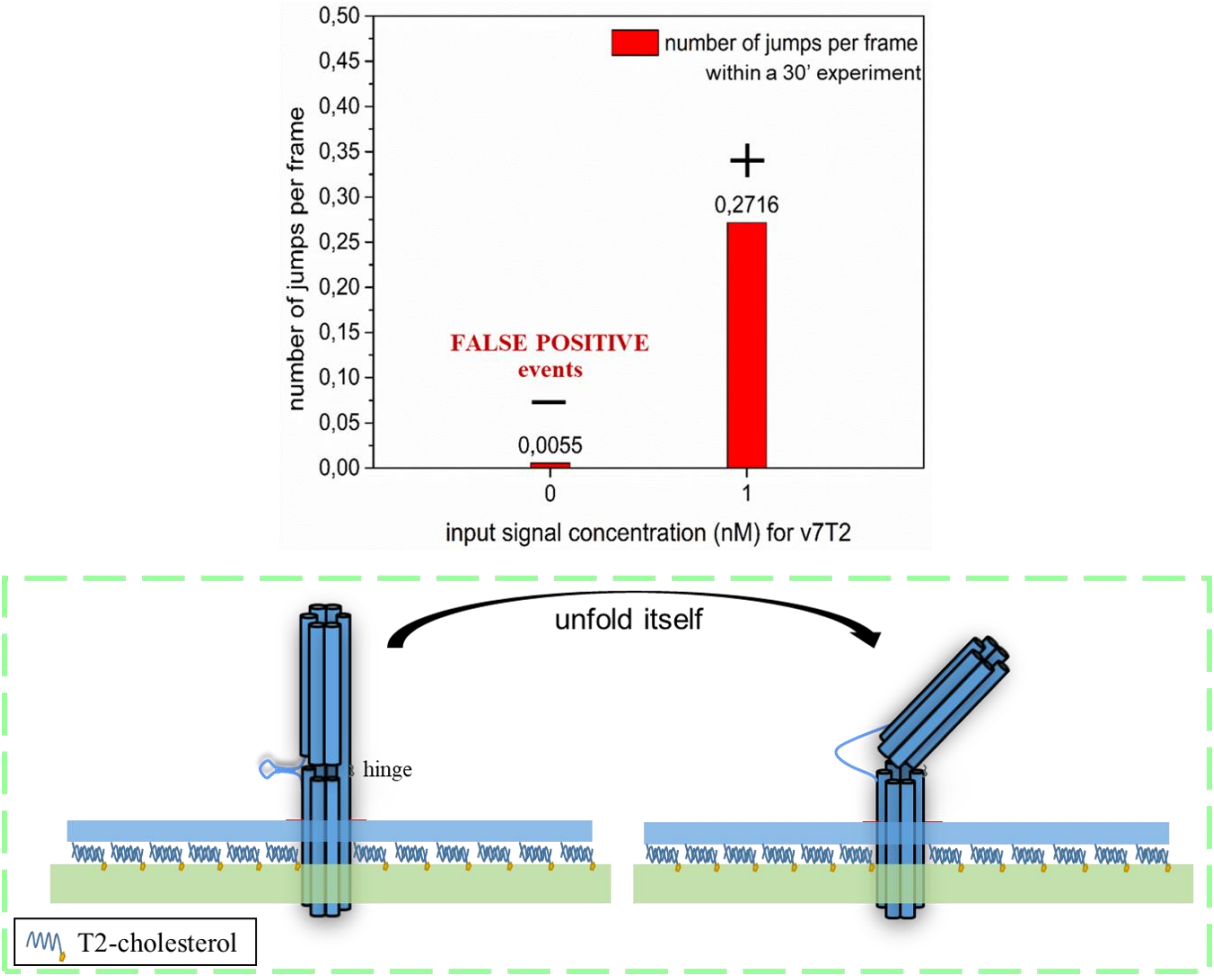


Figure 4-28. *(Up) Number of jumps per frame (~ 50s) for v7T2, [nanopore] = 1nM, left: [mi22] = 0nM; right: [mi22] = 1nM. (Bottom) The schematic diagram of FALSE POSITIVE events is the T2 nanopore at the closed-state can unfold itself.*

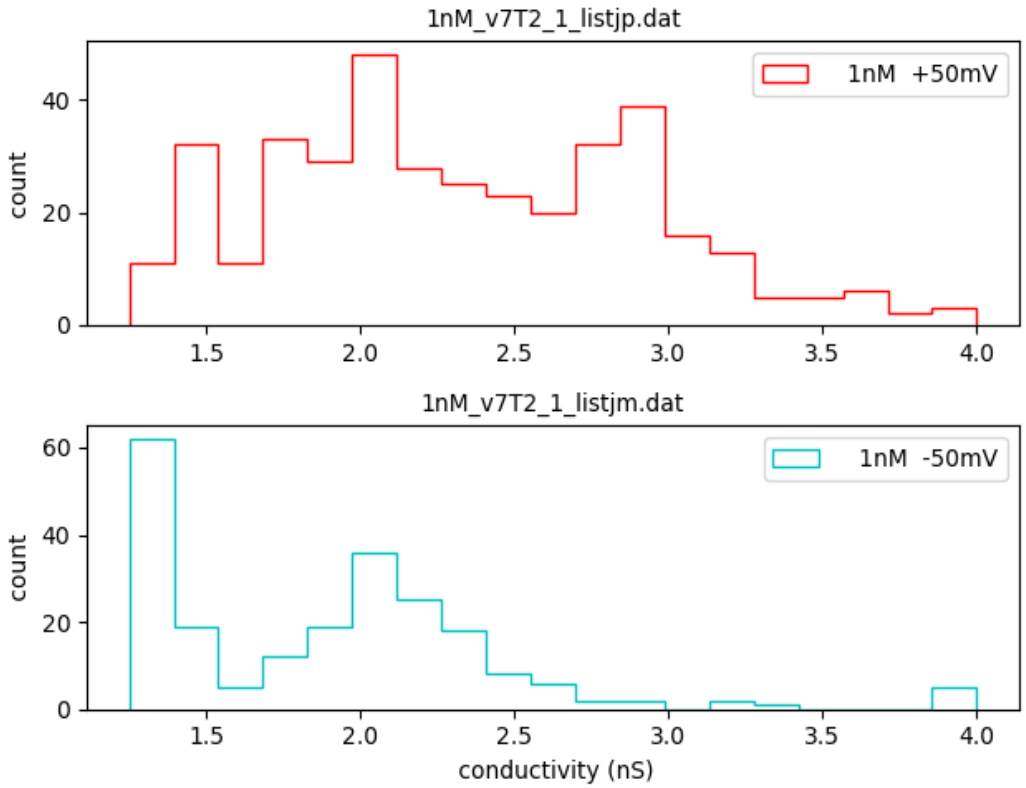


Figure 4-29. *The conductance histogram of v7T2 to detect different concentrations of input signal. (a) [input] = 0nM. (b) [input] = 0.5pM. (c) [input] = 0.5nM. (d) [input] = 5nM. The v7T2 concentration is 0.1nM.*

Let us now examine the distribution of values of the observed conductance jumps. These are shown in **Figure 4-29**, separating the values obtained for positive (+50mV) and negative (-50mV) holding potentials. For positive potentials, two well defined peaks are observed, respectively 2nS and 2.9 nS. The model (**equation 6**) gives a value 2.5nS in between the two observed peaks. Because the simple model can only be expected to give an upper value of the conductance, only the peak at 2nS, observed both for positive and negative holding potentials can be attributed unambiguously to the model structure. It can also be noted that the only two jumps observed in the absence of input signal were close to 2nS. At the moment, we do not have a clear explanation for the appearance of the 3nS peak.

4.3 Preparation protocols

4.3.1 Preparation of buffer

These chemical regents were supplied from Sigma-Aldrich Inc., unless noted otherwise. 0.22μm Millipore membrane filter units (Merck Millipore Ltd.).

Buffer preparation

1. 40mL, 500mM $MgCl_2$: measure 1.904g $MgCl_2$ in 40mL sterile milli-Q water and store at 4ºC.

2. 40mL, 2×B buffer: measure 5.964g KCl, 504.32mg Tris-HCl powder and add 3.2mL of 500mM $MgCl_2$ and 33mL sterile milli-Q water, use ~ 160μL of 1M NaOH to adjust pH to 7.4, adjust the volume up to 40mL. Then filter by 0.22μm Millipore membrane and store at 4ºC. The final concentration is 2M KCl, 40mM $MgCl_2$, 80mM Tris-HCl, pH 7.4.

3. 40mL, 2×TM buffer: add 8mL 10×concentrate TAE, 3.2mL of 500mM $MgCl_2$ and 26mL sterile milli-Q water, use ~ 480μL 1M acetic acid to adjust pH to 7.4, adjust the volume up to 40mL. Then filter by 0.22μm Millipore membrane and store at 4ºC. The final concentration is 2×TAE, 40mM $MgCl_2$, pH 7.4.

4. 10mL 1×B buffer: mix 5mL 2×B buffer and 5mL sterile milli-Q water and store at 4ºC. The final concentration is 1M KCl, 20mM $MgCl_2$, 40mM Tris-HCl, pH 7.4.

5. 10mL 1×BO buffer (0.1%): mix 5mL 2×buffer, 4.99mL sterile milli-Q water and 10μL OPOE (poly(ethylene glycol) octyl ether), and store at 4ºC. The final concentration is 1M KCl, 20mM $MgCl_2$, 40mM Tris-HCl, 0.1% OPOE (v/v), pH 7.4.

6. 10mL 1×BO buffer (0.2%): mix 5mL 2×buffer, 4.98mL sterile milli-Q water and 20μL OPOE, and store at 4ºC. The final concentration is 1M KCl, 20mM $MgCl_2$, 40mM Tris-HCl, 0.2% OPOE (v/v), pH 7.4.

7. 10mL 1×BO buffer (0.25%): mix 5mL 2×buffer, 4.98mL sterile milli-Q water and 20μL OPOE, and store at 4ºC. The final concentration is 1M KCl, 20mM $MgCl_2$, 40mM Tris-HCl, 0.25% OPOE (v/v), pH 7.4.

8. 10mL 2×BO buffer (0.5%): mix 10mL 2×buffer and 50μL OPOE, and store at 4ºC. The final concentration is 2M KCl, 40mM $MgCl_2$, 80mM Tris-HCl, 0.5% OPOE (v/v), pH 7.4.

9. 10mL 1×TM buffer: mix 5mL 2×TM buffer and 5mL sterile milli-Q water and store at 4ºC. The final concentration is 1×TAE, 20mM $MgCl_2$, pH 7.4.

10. 10mL 1×TMO buffer (0.1%): mix 5mL 2×TM buffer, 4.99mL sterile milli-Q water and 10μL OPOE, and store at 4ºC. The final concentration is 1×TAE, 20mM $MgCl_2$, 0.1% OPOE (v/v), pH 7.4.

4.3.2 Preparation of DNA solutions

Materials: For DNA nanobarrel pores, all the oligonucleotides were ordered on a 100-μmol scale in single separated tubes from Eurogentec Inc. For DNA origami structures, we ordered M13mp18 strand (7249 base long) supplied at 100nM from Tilibit Nanosystems Inc. and New England Biolabs Inc., and other oligonucleotides supplied of cartridge and PAGE grade (>80 bases) at 100μM from Eurogentec Inc. The vast majority of M13mp18 molecules are circular, with the remaining molecules being in the linear form which is determined by agarose gel electrophoresis. 100kDa MWCO Amicon centrifugal filter units supplied from Merck Millipore Ltd. 1.7mL safe-seal micro centrifuge tubes with low binding polymer technology supplied from Sorenson Bioscience Inc.

Table 4-2. The selection of ultrafiltration molecular weight cut-off (MWCO) for nucleic acid application. The source is from www.pall.com.

MWCO	Membrane nominal pore size	Base Pairs (double strand)	Nucleotides (single strand)
1kDa	–	5 – 16 bp	9 – 32 nt
3kDa	–	16 – 50 bp	32 – 95 nt
10kDa	–	50 – 145 bp	92 – 285 nt
30kDa	–	145 – 285 bp	285 – 950 nt
100kDa	10nm	475 – 1,450 bp	950 – 2,900 nt
300kDa	35nm	> 1,450 bp	> 2,900 nt

Solutions: All stock solutions of oligonucleotides were diluted in sterile milli-Q water to 100 μM. All virus should be stored at -20°C when not in use. Be very careful when aliquoting scaffold into new tubes to avoid contamination. Always label aliquoted tubes with original tube Lot # and opening date. Freeze–thaw cycles can damage the scaffold creating linear instead of circular scaffold which will result in lower complete and correct structures. Try to minimize freeze-thaw cycles as much as possible. The working buffer should be stored at 4°C. Always use aseptic techniques when performing experiments. This will minimize the chance of contamination. Always mix scaffold prior to use. Ensure purity, stability, correct sequence identity, and usability.

4.3.3 Fabrication of DNA nanochannel

Annealing

DNA nanobarrel folding: A folding reaction is heating and cooling equimolar mixtures of all 19 strands in 1×TM buffer for 3.2 hours. The annealing protocol used two different temperature ramps, from 90°C to 60°C in 1 h and from 60°C to 20°C in 2h, lid temperature is 60°C.

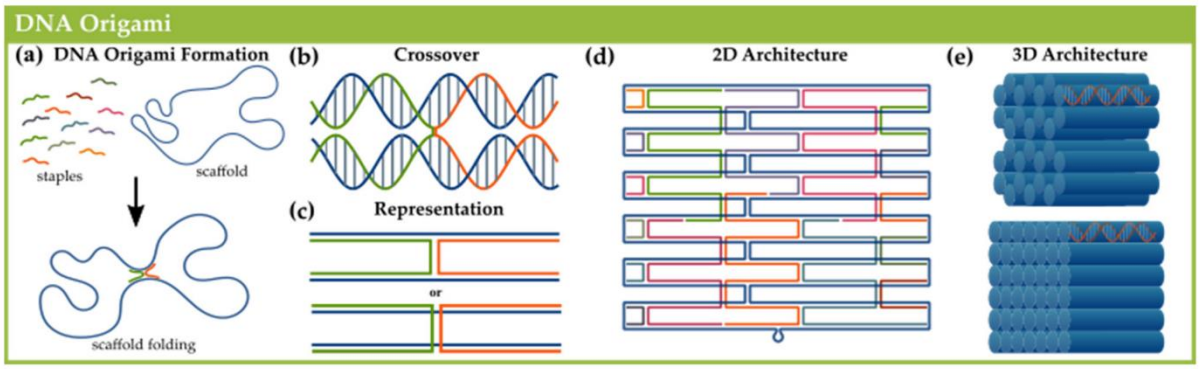


Figure 4-30. *DNA origami folding method: (a) DNA origami nanostructures are assembled from the directed folding of the scaffold strand (blue) by staple strands (multicolored); (b) Individual helical domains are connected by interhelix crossovers; (c) These interhelix crossovers are represented as straight vertical segments; (d) scaDNAno editing is done using a simple 2D visualization of a DNA origami nanostructure that folds into a 3D model; (e) These 3D DNA origami models represent the DNA double helix as a solid cylinder and generally follow a honeycomb (top) or square (bottom) lattice architecture (adapted from ref. [170]).*

DNA origami folding (**Figure 4-30**): A folding reaction is a mixture containing DNA scaffolds, DNA staples, cholesteryl strands, and 1×TMO buffer mastermix to stabilize pH and DNA stability with additional salt ions (typically magnesium). M13mp18 strand is used for assembling the scaffold of rectangular tile and six-helix bundle nanopore. DNA staples are used for assembling rectangular tile (mixture of ~187 or ~150 strands) and six-helix bundle nanopore (mixture of ~25 or ~10 strands), and for anchoring cholesteryl with ~20 extended nucleotides (mixture of ~21 or ~50 strands). Two tubes of DNA origamis are folded by a 20-hour annealing protocol. In the folding reactions, the DNA scaffold (M13mp18) was diluted to 5nM and the staple strands were diluted to 25nM each, corresponding to a 5× excess of each staple to the scaffold. The cholesteryl strands were diluted to 300nM, corresponding to a 6× excess among their binding sides of DNA staples, which can be added before or after the annealing protocol. The mixed sample was put on a thermal ramp starting with a heat denaturation at 80ºC for 20 min followed by cooling from 79ºC to 70ºC at a rate of 10 min/ºC,

69ºC to 41ºC at a rate of 25 min/ºC, 40–29 °C at a rate of 15 min/°C for over 16h, and slowly cooling from 35ºC to 22ºC for over 3h, and then hold at 4ºC, lid temperature is 60°C.

There are the components that can fold different types of DNA nanobarrel and DNA origami structures and the detailed information of their strands is shown in Appendix B.

(i) cork structure: The annealing sample include mixture of 9 scaffolds and 10 staples. 2μL of each strand and 162μL 1×TM buffer were added to a 300μL tube. Total volume is 200μL and final concentration is 1μM. Cholesteryl strands were incubated with nanobarrel before using.

(ii) corkT1 structure: The annealing sample include mixture of 187 staples for T1 rectangular tile, mixture of 21 possible cholesteryl binding sites, 42-nt cork-shape structure (10 scaffolds, 9 staples, 4 linkers to rectangle), and M13mp18 strand. 5.7μL 175nM M13mp18 strand, 18.7μL 534nM mixture of 187 staples, 2.5μL 4μM mixture of cork-shape nanopore, 2.1μL 4.76μM mixture of 21 possible cholesteryl binding sites and 6μL 10μM cholesteryl strands were added into 200μL 1× TMO buffer.

(iii) longcorkT1 structure: The annealing sample include mixture of 187 staples for T1 rectangular tile, mixture of 21 possible cholesteryl binding sites, 84-nt 'longcork' structure (9 scaffolds, 9 staples, 4 linkers to rectangle), and M13mp18 strand. 5.7μL 175nM M13mp18 strand, 18.7μL 534nM mixture of 187 staples, 2.5μL 4μM mixture of 'longcork', 2.1μL 4.76μM mixture of 21 possible cholesteryl binding sites and 6μL 10μM cholesteryl strands were added into 200μL 1× TMO buffer.

(iv) cookieT1 structure: The annealing sample include mixture of 182 staples for T1 rectangular tile, mixture of 21 possible cholesteryl binding sites, 80-nt 'cookie' structure (11 staples, 2 linkers to rectangle), and M13mp18 strand. 5.7μL 175nM M13mp18 strand, 18.2μL 549nM mixture of 182 staples, 2.1μL 4.76μM mixture of 'cookie', 2.1μL 4.76μM mixture of 21 possible cholesteryl binding sites and 6μL 10μM cholesteryl strands were added into 200μL 1× TMO buffer.

(v) v7T2 structure: The annealing sample include mixture of 150 staples for T2 rectangular tile, mixture of 48 possible cholesteryl binding sites, 80-nt 'v7' structure (10 staples and 3 linkers to rectangle), and M13mp18 strand. 5.7μL 175nM M13mp18 strand, 15μL 666nM mixture of 150 staples, 2.1μL 4.76μM mixture of 'v7', 5μL 2μM mixture of 50 possible cholesteryl binding sites and 6μL 10μM cholesteryl strands were added into 200μL 1× TMO buffer.

Purifying

In order to get rid of excess staple strands, the origami products were purified by ultrafiltration or gel extraction method.

Ultrafiltration purification (**Figure 4-31**): After self-assembly is complete, to get rid of excess staple strands, folding samples were filter-purified three times under 400μL 1×TM buffer by using 100kDa MWCO Amicon centrifugal filter units. The speed rate of spin is 13,000 rpm for three times for elution and then the speed rate of inversely spin is 10,000 rpm for 5 min for collection. Nanodrop was used to measure the concentration of purified solutions and its final concentration is at a range of 7–30 nM. This purified solution is stored at 4ºC for two weeks and then diluted to appropriate concentration ready for use.

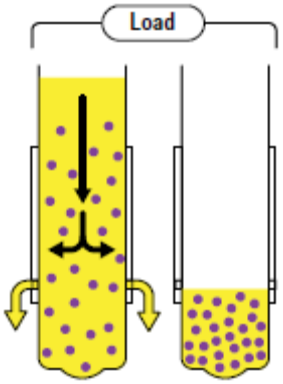


Figure 4-31. *Targets (purple dots) is dissolved in Buffer (yellow) during the concentration step. The image is adapted from the user guide of Amicon purification system.*

Gel extraction purification: after running 0.8% gel electrophoresis, we cut out of the target gel band from the agarose gel under UV light and make it small pieces and then placed in a tube. Then agarose cut-out entails simply squeezing at -20ºC for 10 min and then recovered by extraction with a porous filter membrane by centrifugation at 13,000× g for 5 min.[171]

4.3.4 Gel electrophoresis

Materials: Agarose powder ($C_{24}H_{38}O_{19}$), running buffer of 0.5× TAE and 12.5mM $MgCl_2$, 6× agarose loading dye, solution of DNA structures, ethidium bromide (2 droplets/50mL), stains-all dye (Sigma) or 10000× SYBR safe dye (20μL/200mL), 100bp and 1kb DNA ruler (BioRads) and an electrophoresis transformer used for agarose gel electrophoresis. Acrylamide solutions (for resolving & stacking gels), ammonium persulfate,

N,N,N',N'-tetramethylethylendiamine (TEMED), distilled water used for polyacrylamide gel electrophoresis.

Agarose gel electrophoresis: 150mL 0.8% agarose gel was prepared using running buffer (1×TAE buffer (40mM Tris-AcOH, 2mM Na_2EDTA, pH 8.5)). The SYBR safe, ethidium bromide (EtBr) or stain all dye was used as an oligonucleotide dye. The gel was run at 70V for 120 min in running buffer with loading of 10μL of the sample into each lane at room temperature. Afterward, it was photographed using the gel image analysis system.

Native polyacrylamide gel electrophoresis (Native PAGE): 40mL 8% gel was prepared by 18.5mL distilled water, 10.4mL 1×TAE buffer, 10.7mL 30% acrylamide mix, 0.4mL 10% ammonium persulfate and 0.024mL TEMED. The gel was run at 80 V for 180 min in running buffer with loading of 20μL of the sample into each lane at 4ºC. Afterward, the gel was treated by strains-all dye for 40 min and photographed using the gel image analysis system.
Note: For the scaffold control lane, we use 5μL of 100nM stock added to the control lane (0.5pmol). In general, 20μL, ~10nM (final concentration) of folded constructs yields reliable results for constructs without cholesterol modifications. The magnesium concentration in both the gels and buffer is also critical. During the first three trials, we observed blank bands in nanopores containing 20 mM $MgCl_2$ due to the high magnesium concentration in the gel. Following that, we adjusted the concentration to 12.5 mM for both the gel and buffer, resulting in clear bands.

4.3.5 AFM method

The sample was absorbed and spread on a mica plate, and then carefully used for the experiment. Scanning was performed in the same buffer solution using a QNM tapping fluid mode. Typical images obtained with sample of origami. Rectangular structures can display obviously. Nanopore structures were seen exclusively in the centre of the rectangle. From the AFM picture we can roughly evaluate that the nano-objects are about 90–110 nm long, 45–65 nm wide, meaning that the scaffold with mixtures of staples are folded well.

4.3.6 Transmission Electron Microscopy (TEM)

TEM is an excellent tool for characterizing nanoparticles with its resolution reaches about 0.07nm depending on thickness of the prepared sample and accelerating voltage for the

electron beam [172]. Dissolve 0.1g of uranyl acetate in 5mL of sterile milli-Q water and filter by 0.2μm Millipore membrane filter units to prepare 2% uranyl acetate solution. This solution is long lasting but the longer time for storage, the less efficient for treatment. Prepare glow-discharged, carbon-film-coated copper grids. Nanostructures were negatively stained. 5μL of samples suspension was dropped onto glow-discharged carbon-coated copper 300 mesh grids for 2 min. After stained with 2μL 2% uranyl acetate (w/v) solution for 15 seconds, we wait the grids dry at environmental conditions in dark. Images were acquired at 200kV using a transmission electron microscope (Tecnai and F20 FEG, FEI, ThermoFisher Scientific) and an Eagle 4k_4k and a Gatan USC1000 2k_2k camera (ThermoFisher Scientific). At least 10 locations on the TEM grid were examined. Image J (free software) was used for image analysis.

4.3.7 QCM-D method

Material: 1,2-Diphytanoyl-sn-glycero-3-PC ($C_{48}H_{96}NO_8P$, DphPC, MW 846.30 g/mol) and 1-palmitoyl-2-oleoyl-sn-glycero-3-PC ($C_{42}H_{82}NO_8P$, POPC, MW 760.10 g/mol) were purchased from Avanti Polar Lipids, USA. Poly(ethylene glycol) octyl ether (OPOE, n-Octylpolyoxyethylene, Octyl-POE, $CH_3(CH_2)_6CH_2(OCH_2CH_2)_nOH$) was purchased from Sigma Aldrich Inc. QCM-D sensor, QCM-D sensor crystals (5MHz) reactively sputter-coated with 50nm silicon oxide, plates of 11×11 mm^2 of silicon wafer were used in QCM-D experiments.

Vesicle preparation by sonication: 350μL 10mg/mL DphPC and 150μL 10mg/mL POPC mixed (DphPC: POPC = 7:3) and dried under a stream of nitrogen followed by drying in a vacuum desiccator for 2 hours. The lipid film was suspended in 1×B buffer to 1mg/mL and vortexed it for reconstitution. After milli-Q water pre-sonicated for 5 min three times, small unilamellar vesicles (SUVs) were obtained from 1mL lipids solution by sonication with a tip sonicator operated at 30% duty cycle for 30 min with refrigeration. After centrifuged at 13,000× g for 10 min to remove titanium particles, SUV suspensions were stored at 4°C before use.

Vesicle preparation by extrusion: 350μL 10mg/mL DphPC and 150μL 10mg/mL POPC mixed (the molar ratio of DphPC: POPC is 7:3) and dried under a stream of nitrogen followed by drying in a vacuum desiccator for 2 hours. The lipid film was suspended in 1×B buffer to 2mg/mL and vortexed it for reconstitution. Lipid mixtures were homogenized by five cycles

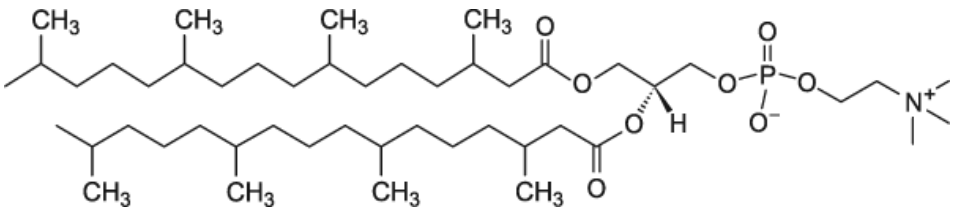
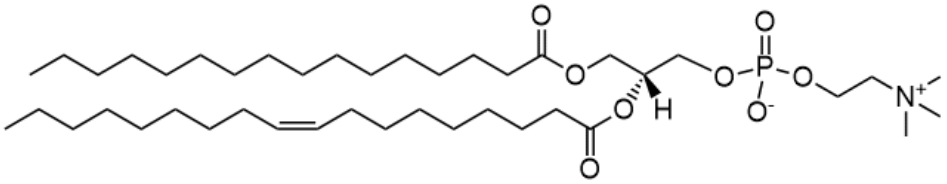
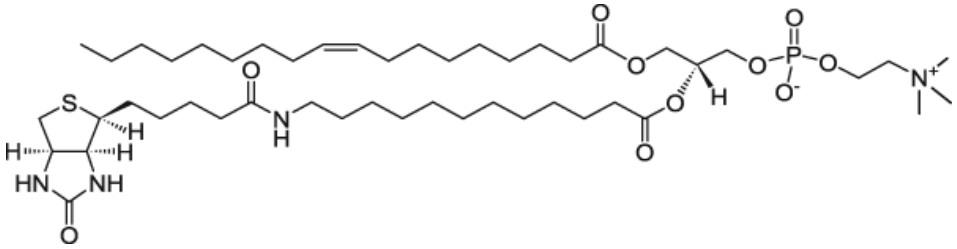
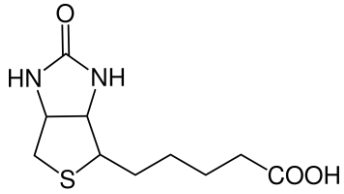
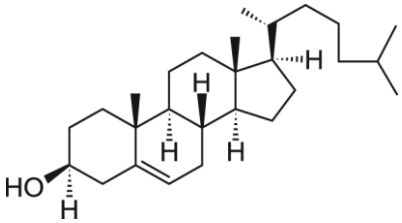
of freeze-thawing in nitrogen liquid and subsequent vortexing. After 11 times extrusion from 100nm extrusion pore membrane, small unilamellar vesicle (SUV) suspensions were stored at 4°C before use.

4.3.8 Droplet interface bilayers (DIBs)

Oil and lipids: the combination of hexadecane ($C_{16}H_{34}$, MW 226.44g/mol, melting point of 18ºC, density of 0.773g/mL) and 1,2-Diphytanoyl-sn-glycero-3-PC ($C_{48}H_{96}NO_8P$, DphPC, MW 846.30g/mol) is a hallmark of DIB experiments. Despite this being the published composition of many DIB experiments, in our hands this combination did not work. The addition of silicon oil AR 20 (polyphenyl-methylsiloxane, $[-Si(CH_3)(C_6H_5)O-]_n$, density of 1.102g/mL) turned out to be the solution to obtain stable bilayers. The experiments presented in chapter 3 (six helix bundle nanopores) showed that the addition of 1-palmitoyl-2-oleoyl-sn-glycero-3-PC ($C_{42}H_{82}NO_8P$, POPC, MW 760.10g/mol) significantly improved the insertion rate of nanopores. Additionally, the test of SA-biotin experiments presented in chapter 3 have mentioned 1-oleoyl-2-[12-biotinyl(aminododecanoyl)]-sn-glycero-3-phosphocholine ($C_{48}H_{89}N_4O_{10}PS$, 18:1-12:0 Biotin PC, MW 945.28g/mol). All lipids were purchased from Avanti Polar Lipids, USA.

All along this thesis, unless otherwise stated, I used the combination (75% hexadecane, 25% silicon oil) together with (70% DphPC and 30% POPC). This is probably not the optimal combination of oil/lipids. I noticed that increasing POPC yield too unstable bilayers, and decreasing POPC yield too stable bilayers. Optimizing this proportion is beyond the scope of this work, where we focused rather on the optimization of nanopore structure. It should also be mentioned that hexadecane becomes quite viscous below 18°C. In our setup, there was no tight temperature control.

Table 4-3. The structures of lipids, biotin and cholesterol.

Name	Structure
4ME 16:0 PC	
16:0-18:1(n9) PC	
18:1-12:0 biotin PC	
biotin	
cholesterol	

Gel coating: we have found that coating electrodes with a 2.4% agarose gel greatly improves the contact between electrodes and droplets. Otherwise, the monolayer's mechanical resistance is too high and electrodes cannot penetrate inside droplets. Gel coating is visible in **Figure 2-7**.

Detail of preparation steps

2.4% agarose solution (w/v): add 24mg agarose powder in 1mL 1×B buffer, need to dissolve at 90ºC and incubate at 65ºC for use.

10mg/mL DphPC: typically, measure 10mg DphPC powder and disperse in 1mL chloroform, and stored in a glass tube at -20ºC.

10mg/mL POPC: tyically, measure 10mg POPC powder and disperse in 1mL chloroform, and stored in a glass tube at -20ºC.

10mg/mL POPC-biotin: typically, measure 10mg POPC-biotin powder and disperse in 1mL chloroform, and stored in a glass tube at -20ºC.

500μL 7:3 oil: mix 350μL 10mg/mL DphPC and 150μL 10mg/mL POPC in a flask or in a glass tube. Rotate the flask in a fuming hood under N_2 flow with a low speed using a N_2 gas cylinder until a thin phospholipid film appears. Then, place the open flask into a desiccator that is connected to a vacuum pump. Aspirate the inside of the desiccator for at least 1 hour to remove the chloroform thoroughly using the vacuum pump. Later, add 375μL hexadecane and 125μL silicon oil into the dried flask and stored at room temperature. The mass ratio of DphPC and POPC is 7:3. And the glass tubes can be reused, but need to use milli-Q water and ethanol to clean for three times and dry in an electric constant temperature drying oven at 90ºC.

100μM input is diluted to 1μM in sterile milliQ-water, and then diluted to 25nM, 10nM, 1nM, 100pM, 25pM and 5pM in 1×TM buffer when used.

The tested samples are prepared before 10 min as shown in **Table 4-4** and **Table 4-5**.

Table 4-4. Preparation of testing samples in current recording measurements, in the condition of adding cholesteryl after the annealing protocol. Total volume is 50μL.

N0.5nM m0nM C20nM	48μL 0.52nM T1 or T2 nanopore, 0μL input, 1μL 1μM rigid-distal-T1-cholesteryl, 1μL sterile milliQ-water, in 1×B.
N0.5nM m0.0001nM C20nM	48μL 0.52nM nanopore, 1μL 5pM input, 1μL 1μM rigid-distal-T1-cholesteryl, in 1×B.
N0.5nM m0.0005nM C20nM	48μL 0.52nM nanopore, 1μL 25pM input, 1μL 1μM rigid-distal-T1-cholesteryl, in 1×B.
N0.5nM m0.001nM C20nM	48μL 0.52nM nanopore, 1μL 50pM input, 1μL 1μM rigid-distal-T1-cholesteryl, in 1×B.
N0.5nM m0.005nM C20nM	48μL 0.52nM nanopore, 1μL 250pM input, 1μL 1μM rigid-distal-T1-cholesteryl, in 1×B.
N0.5nM m0.01nM C20nM	48μL 0.52nM nanopore, 1μL 500pM input, 1μL 1μM rigid-distal-T1-cholesteryl, in 1×B.
N0.5nM m0.05nM C20nM	48μL 0.52nM nanopore, 1μL 2.5nM input, 1μL 1μM rigid-distal-T1-cholesteryl, in 1×B.
N0.5nM m0.1nM C20nM	48μL 0.52nM nanopore, 1μL 5nM input, 1μL 1μM rigid-distal-T1-cholesteryl, in 1×B.
N0.5nM m0.5nM C20nM	48μL 0.52nM nanopore, 1μL 25nM input, 1μL 1μM rigid-distal-T1-cholesteryl, in 1×B.

Table 4-5. Preparation of testing samples in current recording measurements, in the condition of adding cholesteryl in the annealing protocol. Total volume is 20μL.

N2nM m0nM	19μL 2.1nM T1 or T2 nanopore, 0μL input, 1μL sterile milliQ-water, in 1×B.
N2nM m0.001nM	19μL 2.1nM T1 or T2 nanopore, 1μL 20pM input, in 1×B.
N2nM m0.005nM	19μL 2.1nM T1 or T2 nanopore, 1μL 100pM input, in 1×B.
N2nM m0.01nM	19μL 2.1nM T1 or T2 nanopore, 1μL 200pM input, in 1×B.
N2nM m0.05nM	19μL 2.1nM T1 or T2 nanopore, 1μL 1nM input, in 1×B.
N2nM m0.1nM	19μL 2.1nM T1 or T2 nanopore, 1μL 2nM input, in 1×B.
N2nM m0.5nM	19μL 2.1nM T1 or T2 nanopore, 1μL 10nM input, in 1×B.
N2nM m1nM	19μL 2.1nM T1 or T2 nanopore, 1μL 20nM input, in 1×B.
N2nM m2nM	19μL 2.1nM T1 or T2 nanopore, 1μL 40nM input, in 1×B.
N2nM m5nM	19μL 2.1nM T1 or T2 nanopore, 1μL 100nM input, in 1×B.

Current recording experiments were performed with HEKA patch-clamp or Intan patch-clamp. Put the electrodes onto the micromanipulator holder, connect one of electrode to the head stage of a patch-clamp amplifier and the other one to the electrical ground. Set the electrodes to appropriate position above the stage of a microscope by manipulating the micromanipulator, allowing the tips of both Ag/AgCl electrode to soak into the electrolyte solution.

Place 60μL 7:3 oil used to fill the shallow oil PMMA well, and then place two droplets of 0.3μL electrolyte solution (buffer or sample) waiting to create lipid bilayer. View the formation of lipid bilayer by focusing the microscope at the bilayer area of two droplets. After that, measure the bilayer membrane electrical capacitance by applying a ramp potential (10mV/10ms followed by –10mV/10ms). Collect and analyse current recordings under ±30 mV or ±50 mV, and calculate the conductance by using python3 (Appendix A Code).

4.4 Conclusion

In this chapter, we have studied the consequences of inserting a DNA nanopore into a rectangular platform. Depending on the way the insertion was done (strategies T1 or T2, corresponding respectively to not sharing or sharing the scaffold between the platform and the nanopore), the results were very different in terms of success ratio of formation of sought structures. AFM imaging showed that the yield of T1 strategy was quite low, on the opposite the T2 strategy was quite successful from this point of view. Next, we considered the characteristics of the jumps in conductance, and particularly the possibility to avoid the existence of jumps which could not be attributed to the existence of an input signal. From this point of view, the elongation of the nanopore combined to the addition of surfactant in the

folding buffer proved to be a quite successful strategy. The addition of surfactant modified the characteristics of the insertion, yielding a characteristic 'only positive' pattern of jumps. This is particularly important from the point of view of quantifying the number of input signals associated with the conductance jumps. For instance, the difference in insertion frequency between the two protocols is striking when comparing jump patterns as shown in **Figure 4-25** or **Figure 4-27**. It can be concluded that many of the positive or negative jumps in **Figure 4-27a** are actually due to the same nanopore, a fact that hinders the application of these structures as a way to quantify concentrations of input signals.

Another conclusion can be drawn from the combined analysis of QCM-D, current recording and AFM experiments. QCM-D shows that the insertion capabilities of the combination nanopore + origami platform is quite efficient. The AFM experiments show however a clear difference between T1 and T2 strategies, in terms of formation of correctly formed structures. This means that the number of platforms that bind to the bilayer are similar (and high) for T1 and T2, however only the later is correctly formed. However, this is at odds with the observation that the frequency of conductance jumps is quite similar between T1 and T2 structures. A tentative conclusion would be that, in both cases, the 'punching' capabilities of the nanopore + platform are quite low, and not limited by the number of bound platforms.

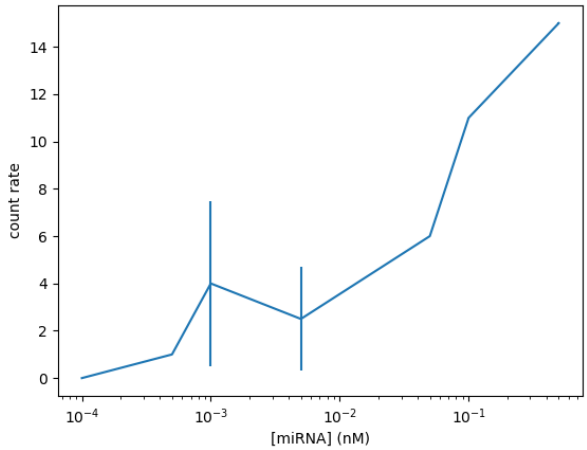


Figure 4-32. *Average count rate for the longcorkT1 structures as a function of the concentration of the DNA analogue of mi22 miRNA. [longcorkT1 nanopore] = 1nM.*

Finally, we would like to give (cf. **Figure 4-32**) the result of a large number of current recording experiments made with the T1 version of the longcork structure. These experiments were done before discovering the effect of adding surfactant in the annealing buffer. Therefore, the number of jumps was at the same time low (insertion rate low) and only partly related to concentration of input signal.

5 General Conclusion

This thesis delves into the field of DNA nanopores and their potential applications in detecting short oligonucleotide sequences, focusing on DNA analogue of miR-21. The initial two chapters serve as an introduction to nanopores, covering their fundamental properties and the detection methods employed in this study.

Chapter 3 introduces the 6-helix bundle DNA nanopore, a simple and lightweight cork-shaped nanopore that can be inserted into lipid bilayers by binding with 3–4 cholesterol molecules. The nanopore's morphology is characterized through AFM experiments, revealing dimensions of approximately 9.1 ± 1.2 nm width and 13.7 ± 1.3 nm length, which closely align with the theoretical size of 8nm × 14nm. Fluorescence experiments demonstrate that the nanopore's morphology undergoes changes based on the presence or absence of an input strand. And electrical recording experiments using the Droplet Interface Bilayer (DIB) method confirm the feasibility of detecting individual 22nt oligoes, resulting in an increase in conductivity from an initial 1.3 nS to 2.3–2.6 nS upon the addition of the input signal. Conductivity histograms in chapter 3 clearly depict this increase, highlighting the potential of nanopore-based detection for sensitive and specific miRNA analysis. Additionally, we explore improvements in the cork_short structure by using longer and fewer scaffolds in the cork_long structure, and replacing cholesterol groups with streptavidin-biotin binding in the cork_SA structure. Among these modifications, the cork_long nanopores exhibit a higher success rate in electrical signal experiments compared to cork_short.

However, a major ongoing challenge is the low insertion rate of the cork-shaped nanopore into the bilayer, which remains below 10% and directly affects the detection of low oligo concentrations.[166] Accurate concentration measurements of oligonucleotide chains depend not only on parallel measurements but also on reasonable success rates in nanopore detection. To address this challenge, ‘DNA origami’ technology [58] is considered, aiming to incorporate the nanopore into larger DNA platforms. This approach can increase the hydrophobic portion attached to the nanopore, enhancing the potential for efficient insertion. Additionally, increasing the length of the nanopore is targeted to improve its penetration through the lipid bilayer.

Chapter 4 investigates the insertion of DNA nanopores into rectangular platforms using two strategies. Firstly, the corkT1 and longcorkT1 in the T1 platform are inspired by Ke's report, connecting the cork-shaped nanopore with a previously reported DNA tile [68] through four double-stranded DNA linkers (separate strategy). Additionally, the cookieT1 in the T1 platform and the v7T2 in the T2 platform are designed using a single DNA strand (M13mp18) as a scaffold to construct the nanopore-platform structure (single-scaffold strategy). Significant differences in forming the desired structures are observed between the two strategies through AFM imaging, with the separate strategy showing relatively low yield of correctly formed structures compared to the single-scaffold strategy. This emphasizes the importance of optimizing nanopore design integrity. After, the QCM-D experiments demonstrate the efficient binding of T-shaped DNA nanopores to the lipid bilayer, as evidenced by a 1.2 nm increase in the binding membrane thickness upon adding 10 nM of the nanopore, corresponding to the nanopore covering half of the lipid bilayer surface.

Furthermore, electrical signal experiments reveal the appearance of high-conductance signals, and an unobserved "all-positive" jumping pattern. It is found that extending the length of the nanopore and adding surfactants in the folding buffer effectively achieve this unique "all-positive" jumping pattern, facilitating in the accurate quantification of input signals. By focusing on the characteristics of the conductivity jumps in the T platforms (such as, **Figure 4-25** and **Figure 4-27**), it is possible to distinguish jumps related to input signals from those unrelated to specific signals. Interestingly, the frequency of conductivity jumps between the two strategies shows some similarity (0.2 jumps per frame, **Figure 4-28**), indicating that the limited punching frequency is influenced not only by the number of binding platforms. However, under the same experimental conditions, we found that, without any input signal, most of the time there are no jumps at all (0.005 jumps per frame, **Figure 4-28**), while their size of jumps is identical to that in the presence of input signal. It shows that the appearance of FALSE POSITIVE events. So far, we do not know how to avoid this small effect.

Furthermore, we discuss the results of extensive current recording experiments using the longcorkT1 version of the nanopore structure were conducted before the discovery of the role of surfactants in the annealing buffer. Due to the low insertion rate, the number of jumps is also limited, and their relationship with the concentration of input strands is only partially observed (**Figure 4-32**).

Therefore, to compare the insertion events by QCM-D and DIB, we found that the number of detected insertion events (8 positive jumps per experiments from 43 DIB experiments) is very small compared to the number of available nanopores (millions). To try to understand this, we hypothesize that the T2 structure is not rigid enough to ensure an efficient insertion. In many cases, the interplay between the lipid bilayer and the nanopore leads to unwanted configurations, as shown in **Figure 5-1**.

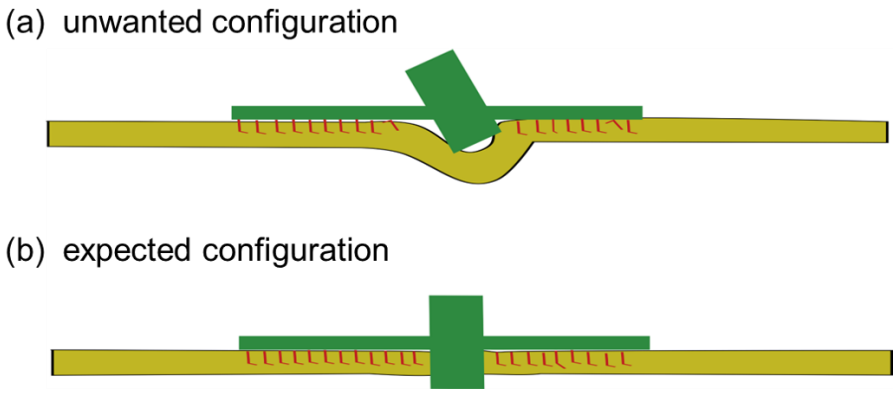
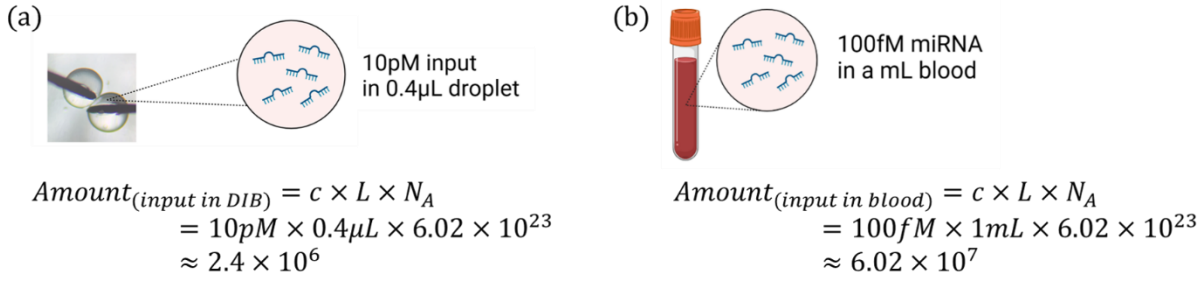


Figure 5-1. Hypothesis on possible interactions between nanopores and bilayers. *(a) one possible interaction: the T2 nanopore+platform structure is floppy and cannot penetrate the lipid bilayer. (b) Expected interaction: the T2 nanopore+platform structure is rigid enough to ensure efficient insertion.*

Still, the best results we have obtained so far is a ratio of ~8 insertions, recorded for 30' recording at a concentration of 0.1nM. Assuming a linear relationship between the concentration and the number of insertions, this leads to an estimate of 10pM concentration as the limit of detection for 30' long experiments. We can see from the calculation results (**Figure 5-2**), that the amount of input (miRNA) needed for this experiment is less than the amount that contained in a mL of 100fM concentration. They have shown a similar order of magnitude and the purified miRNAs from a mL blood can be utilized for 20–30 times testing, which is a really positive conclusion of my work.



$$Amount_{(input\ in\ DIB)} = c \times L \times N_A = 10pM \times 0.4\mu L \times 6.02 \times 10^{23} \approx 2.4 \times 10^{6}$$

$$Amount_{(input\ in\ blood)} = c \times L \times N_A = 100fM \times 1mL \times 6.02 \times 10^{23} \approx 6.02 \times 10^{7}$$

Figure 5-2. *The amount of input (miRNA) needed for DIB experiments (a) is less than the amount that contained in a mL of 100fM concentration in blood (b).*

These findings contribute to our understanding of DNA nanopores and their potential for detecting short oligonucleotide sequences. Despite challenges such as low insertion rates and limited punching capacity, further exploration and optimization of nanopore design hold promise for enhancing efficiency and reliability in various biosensing applications. The insights gained from this study pave the way for future research and development of advanced nanopore-based technologies for sensitive and specific miRNA analysis. With further improvements in DNA nanopore design, this method can be used not only for the detection of individual DNA molecules but also for the single-molecule detection of miRNA.

As mentioned earlier, microRNAs are a group of short (19-22 nt), single-stranded non-coding RNA molecules that play a role as post-transcriptional gene regulatory factors in various physiological processes. Abnormal expression levels of microRNAs have been closely associated with various diseases, including cancer, where specific types of microRNAs can act as oncogenes, tumor suppressors, or metastatic regulators. MicroRNA miR-21, a 22-nucleotide RNA, is one of the earliest discovered mammalian microRNAs and serves as a cancer biomarker in breast, prostate, liver, and colon cancers.

The detection of miRNAs holds great promise for early disease detection, including cancer, and the development of portable direct detection methods can greatly enhance the utilization of miRNA biomarkers in clinical settings. Nanopores have emerged as a remarkable tool in the fields of multi-molecular biophysics and biotechnology, offering significant advantages in nucleic acid detection and analysis of nanoparticle molecular conformations. Nanopore-based detection, utilizing semiconductor technology's miniaturization, provides prospects for creating compact and user-friendly miRNA analysis devices.

In conclusion, this thesis enhances our understanding of DNA nanopores and their potential for detecting short oligonucleotide sequences, particularly miRNA. Despite challenges such as low insertion rates and limited punching frequency, further exploration and optimization of nanopore design hold the promise for enhancing efficiency and reliability in various biosensing applications. The insights gained from this study pave the way for future research and development of advanced nanopore-based technologies for sensitive and specific miRNA analysis, ultimately advancing early disease detection and personalized medicine.

Bibliography

[1] Israelachvili JN, Marcelja S, Horn RG (**1980**) Physical Principles of Membrane Organization. Q. Rev. Biophys. 13 (2):121-200. doi:10.1017/s0033583500001645

[2] Monteiro N, Martins A, Reis RL, Neves NM (**2014**) Liposomes in Tissue Engineering and Regenerative Medicine. J. R. Soc. Interface 11 (101):20140459. doi:10.1098/rsif.2014.0459

[3] Sachs JN, Engelman DM (**2006**) Introduction to the Membrane Protein Reviews: The Interplay of Structure, Dynamics, and Environment in Membrane Protein Function. Annu. Rev. Biochem. 75:707-712. doi:10.1146/annurev.biochem.75.110105.142336

[4] Heron AJ, Thompson JR, Mason AE, Wallace MI (**2007**) Direct Detection of Membrane Channels from Gels Using Water-in-Oil Droplet Bilayers. J. Am. Chem. Soc. 129 (51):16042-16047. doi:10.1021/ja075715h

[5] Bayley H, Cronin B, Heron A, Holden MA, Hwang WL, Syeda R, Thompson J, Wallace M (**2008**) Droplet Interface Bilayers. Mol. Biosyst. 4 (12):1191-1208. doi:10.1039/b808893d

[6] Haylock S, Friddin MS, Hindley JW, Rodriguez E, Charalambous K, Booth PJ, Barter LMC, Ces O (**2020**) Membrane Protein Mediated Bilayer Communication in Networks of Droplet Interface Bilayers. Commun. Chem. 3 (1):77. doi:10.1038/s42004-020-0322-1

[7] Mueller P, Rudin DO, Tien HT, Wescott WC (**1962**) Reconstitution of Cell Membrane Structure in Vitro and Its Transformation into an Excitable System. Nature 194:979-980. doi:10.1038/194979a0

[8] Montal M, Mueller P (**1972**) Formation of Bimolecular Membranes from Lipid Monolayers and a Study of Their Electrical Properties. Proceedings of the National Academy of Sciences 69 (12):3561-3566. doi:10.1073/pnas.69.12.3561

[9] Wiese A, Munstermann M, Gutsmann T, Lindner B, Kawahara K, Zahringer U, Seydel U (**1998**) Molecular Mechanisms of Polymyxin B-Membrane Interactions: Direct Correlation between Surface Charge Density and Self-Promoted Transport. J. Membr. Biol. 162 (2):127-138. doi:10.1007/s002329900350

[10] Winterhalter M (**2000**) Black Lipid Membranes. Curr. Opin. Colloid Interface Sci. 5 (3-4):250-255. doi:10.1016/s1359-0294(00)00063-7

[11] Ma T, Sato M, Komiya M, Kanomata K, Watanabe T, Feng X, Miyata R, Tadaki D, Hirose F, Tozawa Y, Hirano-Iwata A (**2022**) Lateral Voltage as a New Input for Artificial Lipid Bilayer Systems. Faraday Discuss. 233 (0):244-256. doi:10.1039/d1fd00045d

[12] Nikoleli GP, Ibupoto Z, Nikolelis D, Likodimos V, Psaroudakis N, Tzamtzis N, Willander M, Hianik T (**2013**) Potentiometric Cholesterol Biosensing Application of Graphene Electrode with Stabilized Polymeric Lipid Membrane. Cent. Eur. J. Chem. 11 (9):1554-1561. doi:10.2478/s11532-013-0285-5

[13] Nikoleli GP, Nikolelis DP, Siontorou CG, Nikolelis MT, Karapetis S (**2018**) The Application of Lipid Membranes in Biosensing. Membranes 8 (4):108. doi:10.3390/membranes8040108

[14] Ikeda A, Sue T, Akiyama M, Fujioka K, Shigematsu T, Doi Y, Kikuchi J, Konishi T, Nakajima R (**2008**) Preparation of Highly Photosensitizing Liposomes with Fullerene-Doped Lipid Bilayer Using Dispersion-Controllable Molecular Exchange Reactions. Org. Lett. 10 (18):4077-4080. doi:10.1021/ol8015918

[15] Ehrenstein G, Lecar H (**1977**) Electrically Gated Ionic Channels in Lipid Bilayers. Q. Rev. Biophys. 10 (1):1-34. doi:10.1017/s0033583500000123

[16] Zhang L, Zhang Z, Negahban M, Jérusalem A (**2019**) Molecular Dynamics Simulation of Cell Membrane

Pore Sealing. Extreme Mech. Lett. 27:83-93. doi:10.1016/j.eml.2019.01.008

[17] Zhou Y, Kumon RE, Cui J, Deng CX (**2009**) The Size of Sonoporation Pores on the Cell Membrane. Ultrasound Med. Biol. 35 (10):1756-1760. doi:10.1016/j.ultrasmedbio.2009.05.012

[18] Kasianowicz JJ, Brandin E, Branton D, Deamer DW (**1996**) Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. Proc. Natl. Acad. Sci. U. S. A. 93 (24):13770-13773. doi:10.1073/pnas.93.24.13770

[19] Breckenridge LJ, Almers W (**1987**) Currents through the Fusion Pore That Forms During Exocytosis of a Secretory Vesicle. Nature 328 (6133):814-817. doi:10.1038/328814a0

[20] Man VH, Truong PM, Li MS, Wang J, Van-Oanh NT, Derreumaux P, Nguyen PH (**2019**) Molecular Mechanism of the Cell Membrane Pore Formation Induced by Bubble Stable Cavitation. J. Phys. Chem. B 123 (1):71-78. doi:10.1021/acs.jpcb.8b09391

[21] Nehra A, Ahlawat S, Singh KP (**2019**) A Biosensing Expedition of Nanopore: A Review. Sensors Actuators B: Chem. 284:595-622. doi:10.1016/j.snb.2018.12.143

[22] Ding T, Yang J, Pan V, Zhao N, Lu Z, Ke Y, Zhang C (**2020**) DNA Nanotechnology Assisted Nanopore-Based Analysis. Nucleic Acids Res. 48 (6):2791-2806. doi:10.1093/nar/gkaa095

[23] Melo MN, Ferre R, Castanho MA (**2009**) Antimicrobial Peptides: Linking Partition, Activity and High Membrane-Bound Concentrations. Nat. Rev. Microbiol. 7 (3):245-250. doi:10.1038/nrmicro2095

[24] Wu HC, Bayley H (**2008**) Single-Molecule Detection of Nitrogen Mustards by Covalent Reaction within a Protein Nanopore. J. Am. Chem. Soc. 130 (21):6813-6819. doi:10.1021/ja8004607

[25] Lu Y, Wu XY, Ying YL, Long YT (**2019**) Simultaneous Single-Molecule Discrimination of Cysteine and Homocysteine with a Protein Nanopore. Chem. Commun. 55 (63):9311-9314. doi:10.1039/c9cc04077c

[26] Thakur AK, Movileanu L (**2018**) Real-Time Measurement of Protein-Protein Interactions at Single-Molecule Resolution Using a Biological Nanopore. Nat. Biotechnol. 37 (1):96-101. doi:10.1038/nbt.4316

[27] Wang S, Zhao Z, Haque F, Guo P (**2018**) Engineering of Protein Nanopores for Sequencing, Chemical or Protein Sensing and Disease Diagnosis. Curr. Opin. Biotechnol. 51:80-89. doi:10.1016/j.copbio.2017.11.006

[28] Negahdary M, Angnes L (**2022**) Application of Electrochemical Biosensors for the Detection of Micrornas (Mirnas) Related to Cancer. Coord. Chem. Rev. 464:214565. doi:10.1016/j.ccr.2022.214565

[29] Chien CC, Shekar S, Niedzwiecki DJ, Shepard KL, Drndic M (**2019**) Single-Stranded DNA Translocation Recordings through Solid-State Nanopores on Glass Chips at 10 Mhz Measurement Bandwidth. ACS Nano 13 (9):10545-10554. doi:10.1021/acsnano.9b04626

[30] Kaviti AK, Akkala SR (**2023**) Influence of Anodization Time on Al2o3 Nanoporous Morphology and Optical Properties Using Energy Band Gap at Room Temperature. Results Eng. 17:100816. doi:10.1016/j.rineng.2022.100816

[31] Jena MK, Roy D, Pathak B (**2022**) Machine Learning Aided Interpretable Approach for Single Nucleotide-Based DNA Sequencing Using a Model Nanopore. J. Phys. Chem. Lett. 13 (50):11818-11830. doi:10.1021/acs.jpclett.2c02824

[32] Beamish E, Tabard-Cossa V, Godin M (**2019**) Programmable DNA Nanoswitch Sensing with Solid-State Nanopores. ACS Sens. 4 (9):2458-2464. doi:10.1021/acssensors.9b01053

[33] Fologea D, Gershow M, Ledden B, McNabb DS, Golovchenko JA, Li J (**2005**) Detecting Single Stranded DNA with a Solid State Nanopore. Nano Lett. 5 (10):1905-1909. doi:10.1021/nl051199m

[34] Traversi F, Raillon C, Benameur SM, Liu K, Khlybov S, Tosun M, Krasnozhon D, Kis A, Radenovic A (**2013**) Detecting the Translocation of DNA through a Nanopore Using Graphene Nanoribbons. Nat. Nanotechnol. 8 (12):939-945. doi:10.1038/nnano.2013.240

[35] Tian S, Bai Y, Li S, Chen Z, Zhang L, Li H, Zhou P, He Y (**2023**) Simple Preparation of Uio-66-Nh2-Modified Microsphere Layer/Nanofibrous Membrane by Coaxial Spinning for Purification of Complex Wastewater. J. Membr. Sci. 669:121291. doi:10.1016/j.memsci.2022.121291

[36] Ahmed T, Haraldsen JT, Zhu JX, Balatsky AV (**2014**) Next-Generation Epigenetic Detection Technique: Identifying Methylated Cytosine Using Graphene Nanopore. J. Phys. Chem. Lett. 5 (15):2601-2607. doi:10.1021/jz501085e

[37] Darvish A, Lee JS, Peng B, Saharia J, VenkatKalyana Sundaram R, Goyal G, Bandara N, Ahn CW, Kim J, Dutta P, Chaiken I, Kim MJ (**2019**) Mechanical Characterization of Hiv-1 with a Solid-State Nanopore Sensor. Electrophoresis 40 (5):776-783. doi:10.1002/elps.201800311

[38] Dekker C (**2007**) Solid-State Nanopores. Nat. Nanotechnol. 2 (4):209-215. doi:10.1038/nnano.2007.27

[39] Lin K, Chen C, Wang C, Lian P, Wang Y, Xue S, Sha J, Chen Y (**2022**) Fabrication of Solid-State Nanopores. Nanotechnology 33 (27). doi:10.1088/1361-6528/ac622b

[40] Albrecht T (**2019**) Single-Molecule Analysis with Solid-State Nanopores. Annu. Rev. Anal. Chem. (Palo Alto Calif.) 12 (1):371-387. doi:10.1146/annurev-anchem-061417-125903

[41] Acharya S, Jiang A, Kuo C, Nazarian R, Li K, Ma A, Siegal B, Toh C, Schmidt JJ (**2020**) Improved Measurement of Proteins Using a Solid-State Nanopore Coupled with a Hydrogel. ACS Sens. 5 (2):370-376. doi:10.1021/acssensors.9b01928

[42] Healy K, Schiedt B, Morrison AP (**2007**) Solid-State Nanopore Technologies for Nanopore-Based DNA Analysis. Nanomedicine 2 (6):875-897. doi:10.2217/17435889.2.6.875

[43] Bell NA, Keyser UF (**2015**) Specific Protein Detection Using Designed DNA Carriers and Nanopores. J. Am. Chem. Soc. 137 (5):2035-2041. doi:10.1021/ja512521w

[44] Hall AR, Scott A, Rotem D, Mehta KK, Bayley H, Dekker C (**2010**) Hybrid Pore Formation by Directed Insertion of Alpha-Haemolysin into Solid-State Nanopores. Nat. Nanotechnol. 5 (12):874-877. doi:10.1038/nnano.2010.237

[45] Mojtabavi M, Greive SJ, Antson AA, Wanunu M (**2022**) High-Voltage Biomolecular Sensing Using a Bacteriophage Portal Protein Covalently Immobilized within a Solid-State Nanopore. J. Am. Chem. Soc. 144 (49):22540-22548. doi:10.1021/jacs.2c08514

[46] Giovannini G, Ardini M, Maccaferri N, Zambrana-Puyalto X, Panella G, Angelucci F, Ippoliti R, Garoli D, De Angelis F (**2019**) Bio-Assisted Tailored Synthesis of Plasmonic Silver Nanorings and Site-Selective Deposition on Graphene Arrays. Adv. Opt. Mater. 8 (4):1901583. doi:10.1002/adom.201901583

[47] Sen P, Hoi H, Gupta M (**2021**) Low Noise Hybrid Nanopore with Engineered Ompg and Bilayer Mos(2). ACS Appl. Bio Mater. 4 (7):5416-5424. doi:10.1021/acsabm.1c00095

[48] Cressiot B, Greive SJ, Mojtabavi M, Antson AA, Wanunu M (**2018**) Thermostable Virus Portal Proteins as Reprogrammable Adapters for Solid-State Nanopore Sensors. Nat. Commun. 9 (1):4652. doi:10.1038/s41467-018-07116-x

[49] Seeman NC (**2010**) Nanomaterials Based on DNA. Annu. Rev. Biochem. 79:65-87. doi:10.1146/annurev-biochem-060308-102244

[50] Shen B, Piskunen P, Nummelin S, Liu Q, Kostiainen MA, Linko V (**2020**) Advanced DNA Nanopore Technologies. ACS Appl. Bio Mater. 3 (9):5606-5619. doi:10.1021/acsabm.0c00879

[51] Liu X, Zhao Y, Liu P, Wang L, Lin J, Fan C (**2019**) Biomimetic DNA Nanotubes: Nanoscale Channel Design and Applications. Angew. Chem., Int. Ed. Engl. 58 (27):8996-9011. doi:10.1002/anie.201807779

[52] Gopfrich K, Zettl T, Meijering AE, Hernandez-Ainsa S, Kocabey S, Liedl T, Keyser UF (**2015**) DNA-Tile Structures Induce Ionic Currents through Lipid Membranes. Nano Lett. 15 (5):3134-3138. doi:10.1021/acs.nanolett.5b00189

[53] Seifert A, Gopfrich K, Burns JR, Fertig N, Keyser UF, Howorka S (**2015**) Bilayer-Spanning DNA Nanopores with Voltage-Switching between Open and Closed State. ACS Nano 9 (2):1117-1126. doi:10.1021/nn5039433

[54] Burns JR, Seifert A, Fertig N, Howorka S (**2016**) A Biomimetic DNA-Based Channel for the Ligand-Controlled Transport of Charged Molecular Cargo across a Biological Membrane. Nat. Nanotechnol. 11 (2):152-156. doi:10.1038/nnano.2015.279

[55] Lanphere C, Offenbartl-Stiegert D, Dorey A, Pugh G, Georgiou E, Xing Y, Burns JR, Howorka S (**2021**) Design, Assembly, and Characterization of Membrane-Spanning DNA Nanopores. Nat. Protoc. 16 (1):86-130. doi:10.1038/s41596-020-0331-7

[56] Burns JR, Gopfrich K, Wood JW, Thacker VV, Stulz E, Keyser UF, Howorka S (**2013**) Lipid-Bilayer-Spanning DNA Nanopores with a Bifunctional Porphyrin Anchor. Angew. Chem., Int. Ed. Engl. 52 (46):12069-12072. doi:10.1002/anie.201305765

[57] Langecker M, Arnaut V, Martin TG, List J, Renner S, Mayer M, Dietz H, Simmel FC (**2012**) Synthetic Lipid Membrane Channels Formed by Designed DNA Nanostructures. Science 338 (6109):932-936. doi:10.1126/science.1225624

[58] Thomsen RP, Malle MG, Okholm AH, Krishnan S, Bohr SS, Sorensen RS, Ries O, Vogel S, Simmel FC, Hatzakis NS, Kjems J (**2019**) A Large Size-Selective DNA Nanopore with Sensing Applications. Nat. Commun. 10 (1):5655. doi:10.1038/s41467-019-13284-1

[59] Howorka S (**2017**) Building Membrane Nanopores. Nat. Nanotechnol. 12 (7):619-630. doi:10.1038/nnano.2017.99

[60] Fu D, Reif J (**2021**) 3d DNA Nanostructures: The Nanoscale Architect. Appl. Sci. 11 (6):2624. doi:10.3390/app11062624

[61] Broto A, Gaspari E, Miravet-Verde S, Dos Santos V, Isalan M (**2022**) A Genetic Toolkit and Gene Switches to Limit Mycoplasma Growth for Biosafety Applications. Nat. Commun. 13 (1):1910. doi:10.1038/s41467-022-29574-0

[62] Whelan FJ, Surette MG (**2017**) A Comprehensive Evaluation of the Sl1p Pipeline for 16s Rrna Gene Sequencing Analysis. Microbiome 5 (1):100. doi:10.1186/s40168-017-0314-2

[63] Wang B, Wang M, Zhang W, Xiao T, Chen CH, Wu A, Wu F, Traugh N, Wang X, Li Z, Mei S, Cui Y, Shi S, Lipp JJ, Hinterndorfer M, Zuber J, Brown M, Li W, Liu XS (**2019**) Integrative Analysis of Pooled Crispr Genetic Screens Using Mageckflute. Nat. Protoc. 14 (3):756-780. doi:10.1038/s41596-018-0113-7

[64] Strano MS (**2012**) Chemistry. Functional DNA Origami Devices. Science 338 (6109):890-891. doi:10.1126/science.1231024

[65] Madsen M, Gothelf KV (**2019**) Chemistries for DNA Nanotechnology. Chem. Rev. 119 (10):6384-6458.

doi:10.1021/acs.chemrev.8b00570

[66] Seeman NC (**1982**) Nucleic Acid Junctions and Lattices. J. Theor. Biol. 99 (2):237-247. doi:10.1016/0022-5193(82)90002-9

[67] Messing J (**1983**) New M13 Vectors for Cloning. Methods Enzymol. 101:20-78. doi:10.1016/0076-6879(83)01005-8

[68] Douglas SM, Dietz H, Liedl T, Hogberg B, Graf F, Shih WM (**2009**) Self-Assembly of DNA into Nanoscale Three-Dimensional Shapes. Nature 459 (7245):414-418. doi:10.1038/nature08016

[69] Douglas SM, Marblestone AH, Teerapittayanon S, Vazquez A, Church GM, Shih WM (**2009**) Rapid Prototyping of 3d DNA-Origami Shapes with Cadnano. Nucleic Acids Res. 37 (15):5001-5006. doi:10.1093/nar/gkp436

[70] Dietz H, Douglas SM, Shih WM (**2009**) Folding DNA into Twisted and Curved Nanoscale Shapes. Science 325 (5941):725-730. doi:10.1126/science.1174251

[71] Jun H, Zhang F, Shepherd T, Ratanalert S, Qi X, Yan H, Bathe M (**2019**) Autonomously Designed Free-Form 2d DNA Origami. Sci. Adv. 5 (1):eaav0655. doi:10.1126/sciadv.aav0655

[72] Ong LL, Hanikel N, Yaghi OK, Grun C, Strauss MT, Bron P, Lai-Kee-Him J, Schueder F, Wang B, Wang P, Kishi JY, Myhrvold C, Zhu A, Jungmann R, Bellot G, Ke Y, Yin P (**2017**) Programmable Self-Assembly of Three-Dimensional Nanostructures from 10,000 Unique Components. Nature 552 (7683):72-77. doi:10.1038/nature24648

[73] Xie C, Hu Y, Chen Z, Chen K, Pan L (**2022**) Tuning Curved DNA Origami Structures through Mechanical Design and Chemical Adducts. Nanotechnology 33 (40):405603. doi:10.1088/1361-6528/ac7d62

[74] Doty D, Lee BL, Stérin T (**2020**) Scadnano: A Browser-Based, Easily Scriptable Tool for Designing DNA Nanostructures. DNA 2020: Proceedings of the 26th International Conference on DNA Computing and Molecular Programming. doi:10.48550/arXiv.2005.11841

[75] Sulc P, Romano F, Ouldridge TE, Rovigatti L, Doye JP, Louis AA (**2012**) Sequence-Dependent Thermodynamics of a Coarse-Grained DNA Model. J. Chem. Phys. 137 (13):135101. doi:10.1063/1.4754132

[76] Kim DN, Kilchherr F, Dietz H, Bathe M (**2012**) Quantitative Prediction of 3d Solution Shape and Flexibility of Nucleic Acid Nanostructures. Nucleic Acids Res. 40 (7):2862-2868. doi:10.1093/nar/gkr1173

[77] Hudiyanti D, Radifar M, Raharjo TJ, Narsito N, Noegrohati S (**2014**) A Coarse-Grained Molecular Dynamics Simulation Using Namd Package to Reveal Aggregation Profile of Phospholipids Self-Assembly in Water. J. Chem. 2014:1-6. doi:10.1155/2014/273084

[78] Ohmann A, Gopfrich K, Joshi H, Thompson RF, Sobota D, Ranson NA, Aksimentiev A, Keyser UF (**2019**) Controlling Aggregation of Cholesterol-Modified DNA Nanostructures. Nucleic Acids Res. 47 (21):11441-11451. doi:10.1093/nar/gkz914

[79] Grubmuller H, Heymann B, Tavan P (**1996**) Ligand Binding: Molecular Mechanics Calculation of the Streptavidin-Biotin Rupture Force. Science 271 (5251):997-999. doi:10.1126/science.271.5251.997

[80] Moreno-Moya JM, Vilella F, Simon C (**2014**) Microrna: Key Gene Expression Regulators. Fertil. Steril. 101 (6):1516-1523. doi:10.1016/j.fertnstert.2013.10.042

[81] Pillai RS, Bhattacharyya SN, Filipowicz W (**2007**) Repression of Protein Synthesis by Mirnas: How Many Mechanisms? Trends Cell Biol. 17 (3):118-126. doi:10.1016/j.tcb.2006.12.007

[82] Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S,

Dono M, Rassenti L, Alder H, Volinia S, Liu CG, Kipps TJ, Negrini M, Croce CM (**2005**) Mir-15 and Mir-16 Induce Apoptosis by Targeting Bcl2. Proc. Natl. Acad. Sci. U. S. A. 102 (39):13944-13949. doi:10.1073/pnas.0506654102

[83] Selcuklu SD, Donoghue MT, Spillane C (**2009**) Mir-21 as a Key Regulator of Oncogenic Processes. Biochem. Soc. Trans. 37 (4):918-925. doi:10.1042/BST0370918

[84] Slabakova E, Culig Z, Remsik J, Soucek K (**2017**) Alternative Mechanisms of Mir-34a Regulation in Cancer. Cell Death Dis. 8 (10):e3100. doi:10.1038/cddis.2017.495

[85] Choi YJ, Lin CP, Ho JJ, He X, Okada N, Bu P, Zhong Y, Kim SY, Bennett MJ, Chen C, Ozturk A, Hicks GG, Hannon GJ, He L (**2011**) Mir-34 Mirnas Provide a Barrier for Somatic Cell Reprogramming. Nat. Cell Biol. 13 (11):1353-1360. doi:10.1038/ncb2366

[86] Kluiver J, Poppema S, de Jong D, Blokzijl T, Harms G, Jacobs S, Kroesen BJ, van den Berg A (**2005**) Bic and Mir-155 Are Highly Expressed in Hodgkin, Primary Mediastinal and Diffuse Large B Cell Lymphomas. J. Pathol. 207 (2):243-249. doi:10.1002/path.1825

[87] Kluiver J, Haralambieva E, de Jong D, Blokzijl T, Jacobs S, Kroesen BJ, Poppema S, van den Berg A (**2006**) Lack of Bic and Microrna Mir-155 Expression in Primary Cases of Burkitt Lymphoma. Genes Chromosom. Cancer 45 (2):147-153. doi:10.1002/gcc.20273

[88] Xiang X, Zhuang X, Ju S, Zhang S, Jiang H, Mu J, Zhang L, Miller D, Grizzle W, Zhang HG (**2011**) Mir-155 Promotes Macroscopic Tumor Formation yet Inhibits Tumor Dissemination from Mammary Fat Pads to the Lung by Preventing Emt. Oncogene 30 (31):3440-3453. doi:10.1038/onc.2011.54

[89] Yu F, Jia X, Du F, Wang J, Wang Y, Ai W, Fan D (**2013**) Mir-155-Deficient Bone Marrow Promotes Tumor Metastasis. Mol. Cancer Res. 11 (8):923-936. doi:10.1158/1541-7786.MCR-12-0686

[90] Zhu J, Feng X, Lou J, Li W, Li S, Zhu H, Yang L, Zhang A, He L, Li C (**2013**) Accurate Quantification of Microrna Via Single Strand Displacement Reaction on DNA Origami Motif. PLoS One 8 (8):e69856. doi:10.1371/journal.pone.0069856

[91] Cai S, Pataillot-Meakin T, Shibakawa A, Ren R, Bevan CL, Ladame S, Ivanov AP, Edel JB (**2021**) Single-Molecule Amplification-Free Multiplexed Detection of Circulating Microrna Cancer Biomarkers from Serum. Nat. Commun. 12 (1):3515. doi:10.1038/s41467-021-23497-y

[92] Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (**2005**) Real-Time Quantification of Micrornas by Stem-Loop Rt-Pcr. Nucleic Acids Res. 33 (20):e179. doi:10.1093/nar/gni178

[93] Bustin SA, Nolan T (**2020**) Rt-Qpcr Testing of Sars-Cov-2: A Primer. Int. J. Mol. Sci. 21 (8):3004. doi:10.3390/ijms21083004

[94] Hagemann-Jensen M, Abdullayev I, Sandberg R, Faridani OR (**2018**) Small-Seq for Single-Cell Small-Rna Sequencing. Nat. Protoc. 13 (10):2407-2424. doi:10.1038/s41596-018-0049-y

[95] Hagemann-Jensen M, Ziegenhain C, Chen P, Ramskold D, Hendriks GJ, Larsson AJM, Faridani OR, Sandberg R (**2020**) Single-Cell Rna Counting at Allele and Isoform Resolution Using Smart-Seq3. Nat. Biotechnol. 38 (6):708-714. doi:10.1038/s41587-020-0497-0

[96] Jin F, Xu D (**2021**) A Fluorescent Microarray Platform Based on Catalytic Hairpin Assembly for Micrornas Detection. Anal. Chim. Acta 1173:338666. doi:10.1016/j.aca.2021.338666

[97] Hiratani M, Ohara M, Kawano R (**2017**) Amplification and Quantification of an Antisense

Oligonucleotide from Target Microrna Using Programmable DNA and a Biological Nanopore. Anal. Chem. 89 (4):2312-2317. doi:10.1021/acs.analchem.6b03830

[98] Jet T, Gines G, Rondelez Y, Taly V (**2021**) Advances in Multiplexed Techniques for the Detection and Quantification of Micrornas. Chem. Soc. Rev. 50 (6):4141-4161. doi:10.1039/d0cs00609b

[99] Zhu CS, Zhu L, Tan DA, Qiu XY, Liu CY, Xie SS, Zhu LY (**2019**) Avenues toward Microrna Detection in Vitro: A Review of Technical Advances and Challenges. Comput. Struct. Biotechnol. J. 17 (904-916):904-916. doi:10.1016/j.csbj.2019.06.018

[100] Medina PP, Nolde M, Slack FJ (**2010**) Oncomir Addiction in an in Vivo Model of Microrna-21-Induced Pre-B-Cell Lymphoma. Nature 467 (7311):86-90. doi:10.1038/nature09284

[101] Bai X, Bian Z (**2022**) Microrna-21 Is a Versatile Regulator and Potential Treatment Target in Central Nervous System Disorders. Front. Mol. Neurosci. 15:842288. doi:10.3389/fnmol.2022.842288

[102] Rossi S, Shimizu M, Barbarotto E, Nicoloso MS, Dimitri F, Sampath D, Fabbri M, Lerner S, Barron LL, Rassenti LZ, Jiang L, Xiao L, Hu J, Secchiero P, Zauli G, Volinia S, Negrini M, Wierda W, Kipps TJ, Plunkett W, Coombes KR, Abruzzo LV, Keating MJ, Calin GA (**2010**) Microrna Fingerprinting of Cll Patients with Chromosome 17p Deletion Identify a Mir-21 Score That Stratifies Early Survival. Blood 116 (6):945-952. doi:10.1182/blood-2010-01-263889

[103] Stafford MYC, Willoughby CE, Walsh CP, McKenna DJ (**2022**) Prognostic Value of Mir-21 for Prostate Cancer: A Systematic Review and Meta-Analysis. Biosci. Rep. 42 (1):BSR20211972. doi:10.1042/BSR20211972

[104] Rezazadeh-Gavgani E, Oladghaffari M, Bahramian S, Majidazar R, Dolati S (**2023**) Microrna-21: A Critical Underestimated Molecule in Diabetic Retinopathy. iScience 859:147212. doi:10.1016/j.gene.2023.147212

[105] Dillhoff M, Liu J, Frankel W, Croce C, Bloomston M (**2008**) Microrna-21 Is Overexpressed in Pancreatic Cancer and a Potential Predictor of Survival. J. Gastrointest. Surg. 12 (12):2171-2176. doi:10.1007/s11605-008-0584-x

[106] Khodadoust A, Nasirizadeh N, Seyfati SM, Taheri RA, Ghanei M, Bagheri H (**2023**) High-Performance Strategy for the Construction of Electrochemical Biosensor for Simultaneous Detection of Mirna-141 and Mirna-21 as Lung Cancer Biomarkers. Talanta 252:123863. doi:10.1016/j.talanta.2022.123863

[107] Aksimentiev A (**2010**) Deciphering Ionic Current Signatures of DNA Transport through a Nanopore. Nanoscale 2 (4):468-483. doi:10.1039/b9nr00275h

[108] Zhang L, Burns N, Ji Z, Sun S, Deutscher SL, Carson WE, 3rd, Guo P (**2023**) Nipple Fluid for Breast Cancer Diagnosis Using the Nanopore of Phi29 DNA-Packaging Motor. Nanomedicine 48:102642. doi:10.1016/j.nano.2022.102642

[109] Mendoza O, Calmet P, Alves I, Lecomte S, Raoux M, Cullin C, Elezgaray J (**2017**) A Tensegrity Driven DNA Nanopore. Nanoscale 9 (27):9762-9769. doi:10.1039/c7nr01901g

[110] Krishnan S, Ziegler D, Arnaut V, Martin TG, Kapsner K, Henneberg K, Bausch AR, Dietz H, Simmel FC (**2016**) Molecular Transport through Large-Diameter DNA Nanopores. Nat. Commun. 7 (1):12787. doi:10.1038/ncomms12787

[111] Ke Y, Lindsay S, Chang Y, Liu Y, Yan H (**2008**) Self-Assembled Water-Soluble Nucleic Acid Probe Tiles for Label-Free Rna Hybridization Assays. Science 319 (5860):180-183. doi:10.1126/science.1150082

[112] Sengar A, Ouldridge TE, Henrich O, Rovigatti L, Sulc P (**2021**) A Primer on the Oxdna Model of DNA: When to Use It, How to Simulate It and How to Interpret the Results. Front. Mol. Biosci. 8:693710. doi:10.3389/fmolb.2021.693710

[113] Kroese DP, Brereton T, Taimre T, Botev ZI (**2014**) Why the Monte Carlo Method Is So Important Today. Wiley Interdiscip. Rev. Comput. Stat. 6 (6):386-392. doi:10.1002/wics.1314

[114] Ou X, Chen P, Huang X, Li S, Liu BF (**2020**) Microfluidic Chip Electrophoresis for Biochemical Analysis. J. Sep. Sci. 43 (1):258-270. doi:10.1002/jssc.201900758

[115] Duncombe TA, Herr AE (**2012**) Use of Polyacrylamide Gel Moving Boundary Electrophoresis to Enable Low-Power Protein Analysis in a Compact Microdevice. Anal. Chem. 84 (20):8740-8747. doi:10.1021/ac301875e

[116] Aaij C, Borst P (**1972**) The Gel Electrophoresis of DNA. Biochim. Biophys. Acta 269 (2):192-200. doi:10.1016/0005-2787(72)90426-1

[117] Forwood JK, Jans DA (**2006**) Quantitative Analysis of DNA-Protein Interactions Using Double-Labeled Native Gel Electrophoresis and Fluorescence-Based Imaging. Electrophoresis 27 (16):3166-3170. doi:10.1002/elps.200500872

[118] Aleandri S, Vaccaro A, Armenta R, Volker AC, Kuentz M (**2018**) Dynamic Light Scattering of Biopharmaceutics-Can Analytical Performance Be Enhanced by Laser Power? Pharmaceutics 10 (3):94. doi:10.3390/pharmaceutics10030094

[119] Changenet-Barret P, Emanuele E, Gustavsson T, Improta R, Kotlyar AB, Markovitsi D, Vayá I, Zakrzewska K, Zikich D (**2010**) Optical Properties of Guanine Nanowires: Experimental and Theoretical Study. J. Phys. Chem. C 114 (34):14339-14346. doi:10.1021/jp102106d

[120] Kolbeck PJ, Dass M, Martynenko IV, van Dijk-Moes RJA, Brouwer KJH, van Blaaderen A, Vanderlinden W, Liedl T, Lipfert J (**2023**) DNA Origami Fiducial for Accurate 3d Atomic Force Microscopy Imaging. Nano Lett. 23 (4):1236-1243. doi:10.1021/acs.nanolett.2c04299

[121] Reviakine I, Johannsmann D, Richter RP (**2011**) Hearing What You Cannot See and Visualizing What You Hear: Interpreting Quartz Crystal Microbalance Data from Solvated Interfaces. Anal. Chem. 83 (23):8838-8848. doi:10.1021/ac201778h

[122] Huang CY (**1982**) Determination of Binding Stoichiometry by the Continuous Variation Method: The Job Plot. Methods Enzymol. 87:509-525. doi:10.1016/S0076-6879(82)87029-8

[123] Wightman B, Ha I, Ruvkun G (**1993**) Posttranscriptional Regulation of the Heterochronic Gene Lin-14 by Lin-4 Mediates Temporal Pattern Formation in C. Elegans. Cell 75 (5):855-862. doi:10.1016/0092-8674(93)90530-4

[124] Hao J, Duan F-F, Wang Y (**2017**) Micrornas and Rna Binding Protein Regulators of Micrornas in the Control of Pluripotency and Reprogramming. Curr. Opin. Genet. Dev. 46:95-103. doi:10.1016/j.gde.2017.07.001

[125] Svoronos AA, Engelman DM, Slack FJ (**2016**) Oncomir or Tumor Suppressor? The Duplicity of Micrornas in Cancer. Cancer Res. 76 (13):3666-3670. doi:10.1158/0008-5472.CAN-16-0359

[126] Di Leva G, Croce CM (**2013**) Mirna Profiling of Cancer. Curr. Opin. Genet. Dev. 23 (1):3-11. doi:10.1016/j.gde.2013.01.004

[127] Neely LA, Patel S, Garver J, Gallo M, Hackett M, McLaughlin S, Nadel M, Harris J, Gullans S, Rooke J (**2006**) A Single-Molecule Method for the Quantitation of Microrna Gene Expression. Nat. Methods 3 (1):41-46.

doi:10.1038/nmeth825

[128] Bayley H, Cremer PS (**2001**) Stochastic Sensors Inspired by Biology. Nature 413 (6852):226-230. doi:10.1038/35093038

[129] Astier Y, Braha O, Bayley H (**2006**) Toward Single Molecule DNA Sequencing: Direct Identification of Ribonucleoside and Deoxyribonucleoside 5'-Monophosphates by Using an Engineered Protein Nanopore Equipped with a Molecular Adapter. J. Am. Chem. Soc. 128 (5):1705-1710. doi:10.1021/ja057123+

[130] Rothemund PW (**2006**) Folding DNA to Create Nanoscale Shapes and Patterns. Nature 440 (7082):297-302. doi:10.1038/nature04586

[131] Bell NAW, Engst CR, Ablay M, Divitini G, Ducati C, Liedl T, Keyser UF (**2012**) DNA Origami Nanopores. Nano Lett. 12 (1):512-517. doi:10.1021/nl204098n

[132] Peng Q, Zhang X, Min M, Zou L, Shen P, Zhu Y (**2017**) The Clinical Role of Microrna-21 as a Promising Biomarker in the Diagnosis and Prognosis of Colorectal Cancer: A Systematic Review and Meta-Analysis. Oncotarget 8 (27):44893-44909. doi:10.18632/oncotarget.16488

[133] Burns JR, Stulz E, Howorka S (**2013**) Self-Assembled DNA Nanopores That Span Lipid Bilayers. Nano Lett. 13 (6):2351-2356. doi:10.1021/nl304147f

[134] Zadegan RM, Jepsen MD, Thomsen KE, Okholm AH, Schaffert DH, Andersen ES, Birkedal V, Kjems J (**2012**) Construction of a 4 Zeptoliters Switchable 3d DNA Box Origami. ACS Nano 6 (11):10050-10053. doi:10.1021/nn303767b

[135] Arbona J-M, Aimé J-P, Elezgaray J (**2013**) Cooperativity in the Annealing of DNA Origamis. J. Chem. Phys. 138 (1):015105. doi:10.1063/1.4773405

[136] Dirks RM, Lin M, Winfree E, Pierce NA (**2004**) Paradigms for Computational Nucleic Acid Design. Nucleic Acids Res. 32 (4):1392-1403. doi:10.1093/nar/gkh291

[137] Ouldridge TE, Louis AA, Doye JPK (**2011**) Structural, Mechanical, and Thermodynamic Properties of a Coarse-Grained DNA Model. J. Chem. Phys. 134 (8):085101. doi:10.1063/1.3552946

[138] Hwang WL, Chen M, Cronin B, Holden MA, Bayley H (**2008**) Asymmetric Droplet Interface Bilayers. J. Am. Chem. Soc. 130 (18):5878-5879. doi:10.1021/ja802089s

[139] Taylor GJ, Venkatesan GA, Collier CP, Sarles SA (**2015**) Direct in Situ Measurement of Specific Capacitance, Monolayer Tension, and Bilayer Tension in a Droplet Interface Bilayer. Soft Matter 11 (38):7592-7605. doi:10.1039/C5SM01005E

[140] Rabiner LR (**1989**) A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proc. IEEE 77 (2):257–286. doi:10.1109/5.18626

[141] Yoo J, Aksimentiev A (**2015**) Molecular Dynamics of Membrane-Spanning DNA Channels: Conductance Mechanism, Electro-Osmotic Transport, and Mechanical Gating. J. Phys. Chem. Lett. 6 (23):4680-4687. doi:10.1021/acs.jpclett.5b01964

[142] Wright K, de Silva K, Purdie AC, Plain KM (**2020**) Comparison of Methods for Mirna Isolation and Quantification from Ovine Plasma. Sci. Rep. 10 (1):825. doi:10.1038/s41598-020-57659-7

[143] Baaken G, Sondermann M, Schlemmer C, Rühe J, Behrends JC (**2008**) Planar Microelectrode-Cavity Array for High-Resolution and Parallel Electrical Recording of Membrane Ionic Currents. Lab Chip 8:938-944. doi:10.1039/B800431E

[144] Burns JR, Howorka S (**2018**) Defined Bilayer Interactions of DNA Nanopores Revealed with a Nuclease-

Based Nanoprobe Strategy. ACS Nano 12:3263-3271. doi:10.1021/ACSNANO.7B07835

[145] Cai S, Sze JYY, Ivanov AP, Edel JB (**2019**) Small Molecule Electro-Optical Binding Assay Using Nanopores. Nat. Commun. 10:1797. doi:10.1038/S41467-019-09476-4

[146] Feng L, Li J, Sun J, Wang L, Fan C, Shen J (**2021**) Recent Advances of DNA Nanostructure-Based Cell Membrane Engineering. Adv. Healthc. Mater. 10 (6):2001718. doi:10.1002/adhm.202001718

[147] Göpfrich K, Ohmann A, Keyser UF (**2021**) Design and Assembly of Membrane-Spanning Dna Nanopores. In: Fahie MA (ed) Nanopore Technology (Methods in Molecular Biology), vol 2186. Humana Press Inc., Springer Nature, 33-48. doi:10.1007/978-1-0716-0806-7_4

[148] Ivanov A (**2011**) DNA Tunneling Detector Embedded in a Nanopore. Nano Lett. 11:279-285. doi:10.1021/nl103873a

[149] Jones SF, Joshi H, Terry SJ, Burns JR, Aksimentiev A, Eggert US, Howorka S (**2021**) Hydrophobic Interactions between DNA Duplexes and Synthetic and Biological Membranes. J. Am. Chem. Soc. 143:8305-8313. doi:10.1021/JACS.0C13235

[150] Keyser UF (**2016**) Enhancing Nanopore Sensing with DNA Nanotechnology. Nat. Nanotechnol. 11:106-108. doi:10.1038/NNANO.2016.2

[151] Hirtz M, Brglez J, Fuchs H, Niemeyer CM (**2015**) Selective Binding of DNA Origami on Biomimetic Lipid Patches. Small 11 (43):5752-5758. doi:10.1002/smll.201501333

[152] Li B, Ellington AD, Chen X (**2011**) Rational, Modular Adaptation of Enzyme-Free DNA Circuits to Multiple Detection Methods. Nucleic Acids Res. 39 (16):e110. doi:10.1093/nar/gkr504

[153] Dirks RM, Pierce NA (**2004**) Triggered Amplification by Hybridization Chain Reaction. Proc. Natl. Acad. Sci. U. S. A. 101 (43):15275-15278. doi:10.1073/pnas.0407024101

[154] Zhang DY, Turberfield AJ, Yurke B, Winfree E (**2007**) Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA. Science 318 (5853):1121-1125. doi:10.1126/science.1148532

[155] Seeman NC (**1998**) Nucleic Acid Nanostructures and Topology. Angew. Chem., Int. Ed. Engl. 37 (23):3220-3238. doi:10.1002/(SICI)1521-3773(19981217)37:23<3220::AID-ANIE3220>3.0.CO;2-C

[156] Wei W, Chen X, Wang X (**2022**) Nanopore Sensing Technique for Studying the Hofmeister Effect. Small 18 (23):e2200921. doi:10.1002/smll.202200921

[157] MacKenzie M, Argyropoulos C (**2023**) An Introduction to Nanopore Sequencing: Past, Present, and Future Considerations. Micromachines 14 (2):459. doi:10.3390/mi14020459

[158] Ueda H, Dasgupta B, Yu BY (**2023**) Rna Modification Detection Using Nanopore Direct Rna Sequencing and Nanodoc2. In: Arakawa K (ed) Nanopore Sequencing (Methods in Molecular Biology), vol 2632. Spring Nature, Humana, New York, 299-319. doi:10.1007/978-1-0716-2996-3_21

[159] Wen C, Bertosin E, Shi X, Dekker C, Schmid S (**2023**) Orientation-Locked DNA Origami for Stable Trapping of Small Proteins in the Nanopore Electro-Osmotic Trap. Nano Lett. 23 (3):788-794. doi:10.1021/acs.nanolett.2c03569

[160] Wick RR, Judd LM, Holt KE (**2023**) Assembling the Perfect Bacterial Genome Using Oxford Nanopore and Illumina Sequencing. PLoS Comput. Biol. 19 (3):e1010905. doi:10.1371/journal.pcbi.1010905

[161] Xia P, Rahman Laskar MA, Wang C (**2023**) Wafer-Scale Fabrication of Uniform, Micrometer-Sized, Triangular Membranes on Sapphire for High-Speed Protein Sensing in a Nanopore. ACS Appl. Mater. Interfaces 15 (2):2656-2664. doi:10.1021/acsami.2c18983

[162] Yu L, Kang X, Li F, Mehrafrooz B, Makhamreh A, Fallahi A, Foster JC, Aksimentiev A, Chen M, Wanunu M (**2023**) Unidirectional Single-File Transport of Full-Length Proteins through a Nanopore. Nat. Biotechnol. doi:10.1038/s41587-022-01598-3

[163] Zhou S, Ye J, Zhao X, Zhou Z, Dong Y, Shi Q, Liu N, Wu F (**2023**) A DNA-Schiff Base Functional Nanopore Sensing Platform for the Highly Sensitive Detection of Al(3+) and Zn(2+) Ions. Dalton Trans. 52 (6):1524-1532. doi:10.1039/d2dt03786f

[164] Yin P, Choi HM, Calvert CR, Pierce NA (**2008**) Programming Biomolecular Self-Assembly Pathways. Nature 451 (7176):318-322. doi:10.1038/nature06451

[165] Qian L, Winfree E (**2011**) Scaling up Digital Circuit Computation with DNA Strand Displacement Cascades. Science 332 (6034):1196-1201. doi:10.1126/science.1200520

[166] Yang L, Cullin C, Elezgaray J (**2022**) Detection of Short DNA Sequences with DNA Nanopores. Chemphyschem 23 (7):e202200021. doi:10.1002/cphc.202200021

[167] Mathieu F, Liao S, Kopatsch J, Wang T, Mao C, Seeman NC (**2005**) Six-Helix Bundles Designed from DNA. Nano Lett. 5 (4):661-665. doi:10.1021/nl050084f

[168] Cho N-J, Frank CW, Kasemo B, Höök F (**2010**) Quartz Crystal Microbalance with Dissipation Monitoring of Supported Lipid Bilayers on Various Substrates. Nat. Protoc. 5 (6):1096-1106. doi:10.1038/nprot.2010.65

[169] Hille B (**2001**) Ion Channels of Excitable Membrane. 3rd Edition, Sinauer Associates Inc., Sunderland. doi:10.1016/0306-3623(85)90172-7

[170] Babatunde B, Arias DS, Cagan J, Taylor RE (**2021**) Generating DNA Origami Nanostructures through Shape Annealing. Appl. Sci. 11 (7):2950. doi:10.3390/app11072950

[171] Wagenbauer KF, Engelhardt FAS, Stahl E, Hechtl VK, Stommer P, Seebacher F, Meregalli L, Ketterer P, Gerling T, Dietz H (**2017**) How We Make DNA Origami. Chembiochem 18 (19):1873-1885. doi:10.1002/cbic.201700377

[172] Kim HA, Seo JK, Kim T, Lee BT (**2014**) Nanometrology and Its Perspectives in Environmental Research. Environ. Health Toxicol. 29:e2014016. doi:10.5620/eht.e2014016

Appendix

Appendix A Code

A.1 npl.py

A procedure contains a function that can plot the capacitative current at a series of timepoints, data from HEKA Inc.

A.2 hmmheka.py

A procedure contains functions that convert and plot the capacitative current to capacitance, calculate specific capacitance at applied voltage offset, and count the number of capacitance jumps, data from HEKA Inc.

A.3 read.py

A procedure contains a function that convert the raw data (.clp) from Intan Clamp Inc to python-readable data (.dat).

A.4 unsersample.py

A procedure contains functions that can plot the capacitative current at a series of timepoints, data from Intan Clamp Inc.

A.5 hmm.py

A procedure contains functions that convert and plot the capacitative current to capacitance, calculate specific capacitance at applied voltage offset, and count the number of capacitance jumps, data from Intan Clamp Inc.

A.6 rj.py

A procedure contains a function that plots histogram of capacitance.

A.1 npl.py

```
# npl.py
import os, sys
#import pyqtgraph as pg
import numpy as np
import heka_reader
import matplotlib.pyplot as plt

file_with_data=sys.argv[1]


if(len(sys.argv) < 2):
   print " need file name "
   sys.exit(-1)

def open_heka_file(file_name):

    # Load a .dat file
    bundle = heka_reader.Bundle(file_name)

    colec = bundle.pul
    #print colec[0]
    nchildren = len(colec[0])
    print "nchildren = ",nchildren
    for i in range(nchildren):
      datos = bundle.data[0,i,0,0]
      print i,len(datos)
      plt.plot(datos)
      plt.show()

open_heka_file(file_with_data)
```

A.2 hmmheka.py

```
# hmmheka.py
#!/usr/bin/env python
# -*- coding: utf-8 -*-
from matplotlib.mlab import griddata
from matplotlib.pyplot import figure, show,colorbar
import matplotlib.pyplot as plt
import numpy as np
import sys
import string
from scipy.signal import bessel,filtfilt
import heka_reader
import os
import time

VOLT = 0.030
NAVER = 1
CONDUC_MAX = 40.
CNANO = 20.e-09

if(len(sys.argv) < 2):
   print " need file name "
   sys.exit(-1)

print "assume voltage == ",VOLT

"""try:
   DETAIL=int(raw_input('detailed (1) or not (0):'))
except ValueError:
   print "please, type 0 or 1"  """

def plot_sweep(c):
   x = np.arange(len(c))
   y = np.asarray(c)
   fig = plt.figure()
   #plt.ion()
   plt.plot((x),(y))
   plt.draw()
   plt.pause(1.)
   plt.close(fig)
```

```
40.
41.  def plot_2sweep(c,d):
42.     x = np.arange(len(c))
43.     y = np.asarray(c)
44.     #cmean = np.mean(y)
45.     fig = plt.figure()
46.     #plt.ion()
47.     plt.plot((x),(y))
48.     x = np.arange(len(d))
49.     y = np.asarray(d)
50.     #plt.plot((x),(y),'.')
51.     plt.plot((x),(y))
52.     plt.show()
53.     #plt.draw()
54.     #plt.pause(2.)
55.     #plt.close(fig)
56.
57.  def histo_diff(x):
58.    jump = []
59.    rest = []
60.    for i in range(len(x)-1):
61.      jump.append(abs(x[i+1]-x[i]))
62.    return rest,jump
63.
64.  def lin_fit(x,y):
65.     A = np.vstack([x, np.ones(len(x))]).T
66.     m, c = np.linalg.lstsq(A,y)[0]
67.     return m,c
68.
69.  def find_jump_hmm_rawjump(c):
70.     ff = open("hmmtmp.dat","w")
71.     for i in range(len(c)):
72.       ff.write("%g \n"%c[i] )
73.     ff.close()
74.     command = 'sage ../test_hmm.sage '+"%d "%int(sys.argv[1]) + "%d "%int(sys.argv[2])
75.     os.system(command)
76.     rest = []
77.     ff = open("hmmsol.dat","r")
78.     for line in ff.readlines():
79.       rest.append(float(line.split()[2] ) )
80.     ff.close()
81.     jump = []
82.     for i in range(len(rest)-1):
83.       if(rest[i+1] != rest[i]):
84.         dd = abs(rest[i+1]-rest[i])
85.         jump.append(dd)
86.         """if(dd > 1.8):
87.           print " jump > 1.8 ", i"""
88.
89.     return rest,jump
90.
91.  def find_jump_hmm(c):
92.     ff = open("hmmtmp.dat","w")
93.     for i in range(len(c)):
94.       ff.write("%g \n"%c[i] )
95.     ff.close()
96.     command = 'sage ../test_hmm.sage '+"%d "%int(sys.argv[1]) + "%d "%int(sys.argv[2])
97.     os.system(command)
98.     rest = []
99.     ff = open("hmmsol.dat","r")
100.    for line in ff.readlines():
101.      rest.append(float(line.split()[2] ) )
102.    ff.close()
103.    # first pass: look for ALL jumps
104.    jump = []
105.    tjump = []
106.    for i in range(len(rest)-1):
107.      if(rest[i+1] != rest[i]):
108.        dd = (rest[i+1]-rest[i])
109.        jump.append(dd)
110.        tjump.append(i)
111.        """if(dd > 1.8):
112.          print " jump > 1.8 ", i"""
113.    #second pass: compact jumps of same sign, separated by less than 2 (in time units)
```

```
114.
115.     ljump = len(jump)
116.     ij = -1
117.     while (ij < ljump-2):
118.       ij = ij+1
119.       if(jump[ij]*jump[ij+1] > 0.e0 and tjump[ij+1] - tjump[ij] < 3):
120.         jump[ij] =jump[ij] + jump[ij+1]
121.         jump[ij+1] = 0.
122.
123.     hmmj = []
124.     ff = open("hmmj.dat",'r')
125.     for line in ff.readlines():
126.       ll = line.split()
127.       hmmj.append((float(ll[0]),float(ll[1]) ))
128.
129.
130.     return rest,jump,hmmj
131.
132. def clean_ramp_moins_plus(c,DETAIL=0):
133.     print(" enter clean ramp ")
134.     ramp_m = []
135.     centro = 0.
136.     for i in range(64000,66000):
137.        centro = centro + c[i]
138.     centro = centro/2000.
139.
140.     conductance = 0.
141.     for i in range(5000,55000):
142.        if(abs(c[i]) > CNANO):
143.           jumps_p=[]
144.           jumps_m=[]
145.           rampa_ave_m=np.zeros(10)
146.           rampa_ave_p=np.zeros(10)
147.           return rampa_ave_m,rampa_ave_p, jumps_m,jumps_p
148.        #conductance = max(0.e0,-(c[i]-centro)*1.e09/VOLT)
149.        conductance = -c[i]*1.e09/VOLT
150.        ramp_m.append(conductance)  # this assumes that the conductivity is nS, and that V = -
     VOLT.
151.
152.     rest_m, jumps_m, hmmj_m= find_jump_hmm(ramp_m)
153.
154.
155.     plot_exception = 0
156.     for i in range(len(jumps_m)):
157.        if(jumps_m[i] > 1.6):
158.          plot_exception = 1
159.
160.     #DETAIL = 1
161.     if(DETAIL ): #and len(jumps_m) > 0 ): #or plot_exception > 0):
162.        plot_2sweep(ramp_m,rest_m)
163.        #plot_wt(rampa_ave_m)
164.
165.     ramp_p = []
166.     #for i in range(24000,44000):
167.     for i in range(70000,220000):
168.        if(abs(c[i]) > CNANO):
169.           jumps_p=[]
170.           jumps_m=[]
171.           rampa_ave_m=np.zeros(10)
172.           rampa_ave_p=np.zeros(10)
173.           return rampa_ave_m,rampa_ave_p, jumps_m,jumps_p
174.        #conductance = max(0.e0,(c[i]-centro)*1.e09/VOLT)
175.        conductance = c[i]*1.e09/VOLT
176.        ramp_p.append(conductance)  # this assumes that the conductivity is nS, and that V = 1
     00mV.
177.
178.     rest_p, jumps_p, hmmj_p = find_jump_hmm(ramp_p)
179.
180.
181.     plot_exception = 0
182.     for i in range(len(jumps_p)):
183.        if(jumps_p[i] > 1.6):
184.          plot_exception = 1
185.
186.     if(DETAIL ): #and len(jumps_p) > 0): # or plot_exception > 0):
187.        plot_2sweep(ramp_p,rest_p)
188.        #plot_wt(rampa_ave_p)
189.     return ramp_m,ramp_p, jumps_m,jumps_p, hmmj_m,hmmj_p
```

```
190.
191.
192.
193.  def aver(x,n):
194.     # take time series x, then average over n points, return sub-sampled time series
195.     ave = []
196.     nave = len(x)/n
197.     for i in range(0,nave):
198.        xx = 0.
199.        for j in range(n):
200.           xx = xx+ x[i*n+j]
201.        ave.append(xx/float(n))
202.     return ave
203.
204.
205.  g = open("histo_ramp.dat","w")
206.
207.
208.  def scan_file_onlyplot(filetoscan):
209.     f = open(filetoscan,"r")
210.     print("scanning file ",filetoscan)
211.     ns = -1
212.     c = []
213.
214.     bundle = heka_reader.Bundle(filetoscan)
215.
216.     colec = bundle.pul
217.     nchildren = len(colec[0])
218.     take = np.zeros(nchildren)
219.     ns = -1
220.     mode = 1
221.     while(ns < nchildren-1 and mode != 2 ):
222.        ns=ns+1
223.        c = bundle.data[0,ns,0,0]
224.        plot_sweep(c)
225.        try:
226.            mode=int(raw_input('take (1) not (0) stop (2):'))
227.            take[ns]=mode%2
228.        except ValueError:
229.            print("Not a number")
230.            sys.exit(-2)
231.     return take
232.
233.  def scan_file(filetoscan,take,iprint=0):
234.     f = open(filetoscan,"r")
235.     print("scanning file ",filetoscan)
236.     ns = -1
237.     c = []
238.     curr = []
239.     jumps_m = []
240.     jumps_p = []
241.     hmm_m = []
242.     hmm_p = []
243.     aver = []
244.     bundle = heka_reader.Bundle(filetoscan)
245.
246.     colec = bundle.pul
247.     nchildren = len(colec[0])
248.     ns = -1
249.     mode=1
250.     while(ns < nchildren-1 and mode !=2):
251.        ns=ns+1
252.        c = bundle.data[0,ns,0,0]
253.        if(take[ns] == 1):
254.                    clean_m, clean_p,jumps_loc_m,jumps_loc_p, hloc_m,hloc_p=clean_ramp_moins_p
    lus(c,DETAIL=iprint)
255.                    curr = curr + clean_m
256.                    curr = curr + clean_p
257.                    jumps_m = jumps_m + jumps_loc_m
258.                    jumps_p = jumps_p + jumps_loc_p
259.                    aver.append(np.average(np.asarray(clean_m)))
260.                    aver.append(np.average(np.asarray(clean_p)))
261.                    hmm_m = hmm_m + hloc_m
262.                    hmm_p = hmm_p + hloc_p
263.     f.close()
264.     return curr,jumps_m,jumps_p,aver,hmm_m,hmm_p
```

```
265.
266. if(len(sys.argv) < 4):
267.     print(" **************    OUUUUUPS ! **************")
268.     print("Usage is hmm n_states L_chunk data_file_1 data_file_2 ...")
269.     print(" Try again ")
270.     sys.exit(-2)
271.
272. nscan = 0
273. fig = plt.figure()
274.
275. ctot = []
276. jumps_mtot = []
277. jumps_ptot = []
278. cond_aver = []
279. take_or_not = []
280. hmm_mtot = []
281. hmm_ptot = []
282.
283. print("sys argv ",sys.argv)
284. # first, scan the frames to see which ones to compute
285. for i in range(3,len(sys.argv)):
286.    file_with_data=sys.argv[i]
287.    take_loc = scan_file_onlyplot(file_with_data)
288.    take_or_not.append(take_loc)
289.
290.
291. idetail = int(raw_input("now enter HMM evaluation. Print intermediate (1) or not (0)"))
292.
293. for i in range(3,len(sys.argv)):
294.    file_with_data=sys.argv[i]
295.    curr,jumps_m,jumps_p,curr_aver,hmm_m,hmm_p = scan_file(file_with_data,take_or_not[i-
3],iprint=idetail)
296.    ctot.append(curr)
297.    cond_aver=(curr_aver)
298.    jumps_mtot = jumps_mtot + jumps_m
299.    jumps_ptot = jumps_ptot + jumps_p
300.    hmm_mtot = hmm_mtot + hmm_m
301.    hmm_ptot = hmm_ptot + hmm_p
302.
303. print( " total of ",len(jumps_mtot),len(jumps_ptot)," jumps (m and p ) detected ")
304.
305.
306.
307. named = string.replace(file_with_data,".dat","_listjm.dat")
308. f = open(named,"w")
309. for i in range(len(jumps_mtot)):
310.    f.write("%14.5g %14.5g  \n"%(float(i),jumps_mtot[i]))
311. f.close()
312. named = string.replace(file_with_data,".dat","_listjp.dat")
313. f = open(named,"w")
314. for i in range(len(jumps_ptot)):
315.    f.write("%14.5g %14.5g  \n"%(float(i),jumps_ptot[i]))
316. f.close()
317.
318. named = string.replace(file_with_data,".dat","_hmm_m.dat")
319. f = open(named,"w")
320. for i in range(len(hmm_mtot)):
321.    f.write("%14.5g %14.5g  \n"%(hmm_mtot[i][0],hmm_mtot[i][1]))
322. f.close()
323. named = string.replace(file_with_data,".dat","_hmm_p.dat")
324. f = open(named,"w")
325. for i in range(len(hmm_ptot)):
326.    f.write("%14.5g %14.5g  \n"%(hmm_ptot[i][0],hmm_ptot[i][1]))
327. f.close()
328.
329. #print " mirna = ",mirna
330.
```

A.3 read.py

```
# read.py
import sys
import array
import numpy as np
import matplotlib.pyplot as plt
from scipy.signal import bessel,filtfilt,butter
import pickle

conversion_table = {'0': 0, '1': 1, '2': 2, '3': 3, '4': 4, '5': 5, '6': 6, '7': 7, '8': 8,
'9': 9, 'A': 10 , 'B': 11, 'C': 12, 'D': 13, 'E': 14, 'F': 15}

def hextodec(hexadecimal):
    decimal = 0

    #computing max power value
    power = len(hexadecimal) -1

    for digit in hexadecimal:
            decimal += conversion_table[digit]*16**power
            power -= 1
    return decimal
def length_wave(ff):
  length = 0
  for i in range(len(ff)):
    ll = ff[i]["End"] - ff[i]["Start"]
    length = length + ll+1
  return length

def bessel_filter(x,t,n):
   # apply bessel filter with frequency omega0/n where omega0 is the sampling frequency
   wn = 1./float(n)   # this is a fraction of the Nyquist frequency
   b, a = bessel(4, wn, 'low')
   #b, a = butter(8,0.0125,btype = 'low',analog = False)
   try:
     output_signal = filtfilt(b, a, x[:,0])
   except:
     output_signal = []

   return output_signal

def read_Intan_CLP_file(filename):

    # d = read_Intan_CLP_file(filename) reads the data from a CLAMP
    # file from the given filename (including path).
    #
    # d = read_Intan_CLP_file prompts for a filename.
    #
    # The returned 'd' is a data structure.
    #
    # Version 1.04, 23 October 2018
    #
    # Reads Intan Technologies CLAMP data file generated by controller
    # GUI.  Data are parsed and placed into a structure.
    #
    # Example:
    # >> d = read_Intan_CLP_file;
    # >> plot(d.Data.Time, d.Data.Measured)
    # >> d = read_Intan_CLP_file('myexperiment_AUX_160916_142731.clp');
    # >> plot(d.Data.Time, d.Data.ADC(2,:))


    fid = open(filename, 'rb')
    [header, ] = read_header(fid)
    #print("header = ",header)

    pos = fid.tell()
    if (pos != header["NumBytes"]):
        print('Header NumBytes doesn''t match number of bytes')

    fid.seek(header["NumBytes"])

    # this is the starting position of the data
    start_pos = fid.tell()
    fid.seek(0,2)   # go to end of file
    file_size = fid.tell()
    len_measure = file_size-
start_pos   # this is the total number of bytes with recording data. The actual length = len
_measure/16
```

```
76.
77.     if (header["Datatype"] == 0):
78.         samplrate = header["Settings"]["SamplingRate"]
79.         waveform = header["Settings"]["Waveform"]["Segments"]
80.         #print("wave = ",waveform)
81.         nonew = length_wave(waveform)    # number of data / waveform
82.         #Data = read_data(header, fid,samplrate);
83.     else:
84.         print("not implemented yet")
85.         #d = read_aux_data(d, fid, d.Header.NumADCs);
86.     """time = Data["Time"]
87.     measure = Data["Measured"]
88.     print("lenght of recording",len(time),len(measure))
89.     print("sampling rate",samplrate)"""
90.     #print("waveform",header["Settings"]["Waveform"])
91.     #for i in range(100):
92.     #  print(time[i],measure[i])
93.
94.     NAVER = 80
95.     nrecord = int(len_measure/16/nonew)+1
96.     print(" number of recordings ", nrecord," each frame is ",nonew)
97.     print(" will perform average on ",NAVER," frames ")
98.     i0 = 0
99.     i1 = min(nrecord,nonew)
100.
101.    file_short = filename.replace(".clp",".dat")
102.    data_short = []
103.    i = 0
104.    while(i < nrecord):
105.      len2read = min(nonew*16,len_measure - 16*nonew*i)
106.      try:
107.       data = read_data_block(fid,start_pos+16*i*nonew,len2read,samplrate)
108.       time = data["Time"]
109.       measure = data["Measured"]
110.       data_filtered = bessel_filter(measure,time,NAVER)
111.       print(" displaying frame number ",i+1)
112.       #plt.plot(time[i0:i1],measure[i0:i1])
113.       #plt.plot(time[::NAVER],data_filtered[::NAVER])
114.       #plt.show()
115.       data_short.append(data_filtered[::NAVER])
116.       del data
117.       del measure
118.      except:
119.       pass
120.      i = i+1
121.
122.    """display(sprintf('File contains %0.3f seconds of data sampled at %0.2f kS/s.\n', d.Dat
      a.Time(end), d.Header.Settings.SamplingRate / 1000));
123.
124.    if (datatype == 0)
125.        if (d.Header.Settings.IsVoltageClamp && d.Header.Settings.vClampX2mode)
126.            display(sprintf('This data was taken using 2x Voltage Clamp Mode.  Doubling clam
      p data to account for this.\n'));
127.            d.Data.Clamp = d.Data.Clamp * 2;
128.        end
129.    end"""
130.
131.    # Close data file.
132.    fid.close()
133.
134.    # dump short version of data
135.
136.    file_pick = open(file_short,"wb")
137.    pickle.dump(data_short,file_pick,protocol=pickle.HIGHEST_PROTOCOL)
138.    file_pick.close()
139. def read_header(fid):
140.
141.    magic = np.fromfile(fid,np.uint32,count=1)
142.
143.    themagic = hextodec('F3B1A481')
144.
145.
146.    if( magic[0] !=  themagic):
147.        print('Unrecognized file type.')
148.        sys.exit(-1)
```

```
149.
150.     header = {}
151.     # Read version number.
152.     Version = {}
153.     Version ["Major"] = np.fromfile(fid, np.int16, count=1)[0]
154.     Version [ "Minor" ] = np.fromfile(fid, np.int16, count=1)[0]
155.     header["Version"] = Version
156.
157.
158.     datatype = np.fromfile(fid, np.int16, count=1)[0]
159.     header["Datatype"] = datatype
160.     if (datatype == 0):
161.         print('Reading Intan Technologies CLAMP Data File')
162.
163.         NumBytes = np.fromfile(fid, np.uint16, count=1)[0]
164.         Date = {}
165.         Date["Year"] = np.fromfile(fid, np.int16, count=1)[0]
166.         Date["Month"] = np.fromfile(fid, np.int16, count=1)[0]
167.         Date["Day"] = np.fromfile(fid, np.int16, count=1)[0]
168.         Date["Hour"] = np.fromfile(fid, np.int16, count=1)[0]
169.         Date["Minute"] = np.fromfile(fid, np.int16, count=1)[0]
170.         Date["Second"] = np.fromfile(fid, np.int16, count=1)[0]
171.         header["Date"] = Date
172.         header["Version"] = Version
173.         header["NumBytes"] = NumBytes
174.         header["Chips"] = read_header_chips(fid)
175.         header["Settings"] = read_header_settings(fid)
176.     elif (datatype == 1):
177.         print('Reading Intan Technologies Aux CLAMP Data File')
178.
179.         NumADCs = np.fromfile(fid, np.uint16, count=1)[0]
180.         NumBytes = np.fromfile(fid, np.uint16, count=1)[0]
181.         Date = {}
182.         Date["Year"] = np.fromfile(fid, np.int16, count=1)[0]
183.         Date["Month"] = np.fromfile(fid, np.int16, count=1)[0]
184.
185.         Date["Day"] = np.fromfile(fid, np.int16, count=1)[0]
186.         Date["Hour"] = np.fromfile(fid, np.int16, count=1)[0]
187.         Date["Minute"] = np.fromfile(fid, np.int16, count=1)[0]
188.         Date["Second"] = np.fromfile(fid, np.int16, count=1)[0]
189.         header["Date"] = Date
190.         header["Version"] = Version
191.         header["NumBytes"] = NumBytes
192.         header["NumADCs"] = NumADCs
193.         header["SamplingRate"] =   np.fromfile(fid, np.float32, count=1)[0]
194.     else:
195.         error('Unrecognized data type.')
196.
197.     return [header ]
198.
199.
200. def read_header_settings(fid):
201.     settings = {}
202.     settings["EnableCapacitiveCompensation"] = np.fromfile(fid, np.uint8, count=1)[0]
203.     settings["CapCompensationMagnitude"] = np.fromfile(fid, np.float32, count=1)[0]
204.
205.     filtro = np.fromfile(fid, np.float32, count=1)[0]
206.     if (filtro > 0):
207.         settings["FilterCutoff"] = filtro
208.
209.     settings["PipetteOffset"] = np.fromfile(fid, np.float32, count=1)[0]
210.     settings["SamplingRate"] = np.fromfile(fid, np.float32, count=1)[0]
211.     settings["CellParameters_Rs"] = np.fromfile(fid, np.float32, count=1)[0]
212.     settings["CellParameters_Rm"] = np.fromfile(fid, np.float32, count=1)[0]
213.     settings["CellParameters_Cm"] = np.fromfile(fid, np.float32, count=1)[0]
214.     settings["IsVoltageClamp"] = np.fromfile(fid, np.uint8, count=1)[0]
215.     settings["vClampX2mode"] = np.fromfile(fid, np.uint8, count=1)[0]
216.     if (settings["IsVoltageClamp"] == 1):
217.         settings["VoltageClamp"] = read_header_voltage_clamp_settings(fid)
218.     else:
219.         settings["CurrentClamp"] = read_header_current_clamp_settings(fid)
220.
221.     settings["Waveform"] = read_waveform(fid)
222.     return settings
```

```
223.
224. def read_waveform(fid):
225.
226.     waveform = {}
227.     waveform["Interval"] = np.fromfile(fid, np.float32, count=1)[0]
228.     numSegments = np.fromfile(fid, np.uint16, count=1)[0]
229.     Segments = []
230.     for i in range(numSegments):
231.         Segments.append( read_segment(fid))
232.     waveform["Segments"] = Segments
233.
234.     return waveform
235.
236. def read_segment(fid):
237.
238.     segment = {}
239.     segment["WaveformNumber"] = np.fromfile(fid, np.uint8, count=1)[0]
240.     segment["TOffset"] =  np.fromfile(fid, np.uint32, count=1)[0]+ 1
241.     segment["Start"] = np.fromfile(fid, np.uint32, count=1)[0]+ 1
242.     segment["End"] = np.fromfile(fid, np.uint32, count=1)[0]+ 1
243.     segment["AppliedValue"] = np.fromfile(fid, np.float32, count=1)[0]
244.
245.     return segment
246.
247. def read_header_voltage_clamp_settings(fid):
248.     settings = {}
249.     settings["HoldingVoltage"] = np.fromfile(fid, np.float32, count=1)[0]
250.     settings["NominalResistance"] = np.fromfile(fid, np.float32, count=1)[0]
251.     settings["Resistance"] = np.fromfile(fid, np.float32, count=1)[0]
252.     settings["DesiredBandwidth"] = np.fromfile(fid, np.float32, count=1)[0]
253.     settings["ActualBandwidth"] = np.fromfile(fid, np.float32, count=1)[0]
254.     return settings
255.
256. def read_header_current_clamp_settings(fid):
257.     settings = {}
258.     settings["HoldingCurrent"] = np.fromfile(fid, np.float32, count=1)[0]
259.     settings["CurrentStepSize"] = np.fromfile(fid, np.float32, count=1)[0]
260.     return settings
261.
262. def  read_header_chips(fid):
263.     numChips =    np.fromfile(fid, np.uint16, count=1)[0]
264.     numChannels = np.fromfile(fid, np.uint16, count=1)[0]
265.     chips = []
266.     for i in range(numChips):
267.         chips.append( read_one_header_chip(fid, numChannels) )
268.
269.     return chips
270.
271. def read_one_header_chip(fid, numChannels):
272.     Channels = []
273.     for i in range(numChannels):
274.         Channels.append(  read_one_header_channel(fid) )
275.
276.     chip = {}
277.     chip["Channels"]=Channels
278.     chip["ChipRegisters"] = np.fromfile(fid, np.uint16, count=4)
279.     return chip
280.
281.
282. def read_one_header_channel(fid):
283.     channel = {}
284.     channel["Registers"] = np.fromfile(fid, np.uint16, count=14)
285.     channel["DifferenceAmpResidual"] = np.fromfile(fid, np.int32, count=1)[0]
286.     channel["VoltageAmpResidual"] = np.fromfile(fid, np.int32, count=1)[0]
287.     bestCalibration = np.fromfile(fid, np.uint8, count=16)
288.     """bestCalibration = reshape(bestCalibration, 2, 4, 2);
289.     channel.CoarseCalibration = reshape(bestCalibration(1,:,:), 4, 2);
290.     channel.FineCalibration = reshape(bestCalibration(2,:,:), 4, 2);"""
291.     channel["FeedbackResistors"] = np.fromfile(fid, np.float32, count=5)
292.     channel["DesiredBandwidth"] = np.fromfile(fid, np.float32, count=1)[0]
293.     return channel
294.
```

```
295.  def read_data(data, fid,SamplingRate):
296.      start_pos = fid.tell()
297.      fid.seek(0,2)   # go to enf of file
298.      file_size = fid.tell()
299.      fid.seek(start_pos) # go to beginning of file
300.      Data = read_one_batch(fid, file_size - start_pos, SamplingRate)
301.      return Data
302.
303.  def read_data_block(fid,start_pos,len2read,SamplingRate):
304.      # same as read_data, but goes directly to the frame number nblock
305.
306.      size_of_one = 4 + 4 + 4 + 4  # Timestep + Applied (Software) + Clamp Value + Measured
307.      length = int(len2read/size_of_one)
308.      fid.seek(start_pos) # go to beginning of file + shift of previous frames read
309.      data = {}
310.      byte_array = np.fromfile(fid, dtype=np.uint8, count=len2read)
311.      data_matrix = np.reshape(byte_array,(size_of_one,length),order='F')
312.      par = np.reshape(data_matrix[0:4,:],(1,4 * length),order='F')
313.
314.
315.      dpar = par.view( np.uint32)
316.      ldpar = len(dpar[0])
317.      data["Time"] = np.zeros(ldpar)
318.      for i in range(ldpar):
319.        data["Time"][i] = float(dpar[0][i])/SamplingRate
320.
321.      #print(" par time, shape = ",data["Time"].shape , "type = ",type(data["Time"]))
322.      #print(data["Time"])
323.      par = np.reshape(data_matrix[4:8,:],( 1, 4 * length),order='F')
324.      data["Clamp"] = (par.view(np.float32)).T
325.      par = np.reshape(data_matrix[8:12,:],( 1, 4 * length),order='F')
326.      data["TotalClamp"] = (par.view(np.float32)).T
327.      par = np.reshape(data_matrix[12:16,:],( 1, 4 * length),order='F')
328.      data["Measured"] = (par.view(np.float32)).T
329.      #print(" measured ",data["Measured"])
330.
331.      return data
332.
333.  def read_one_batch(fid, size, sampling_rate):
334.
335.      size_of_one = 4 + 4 + 4 + 4  # Timestep + Applied (Software) + Clamp Value + Measured
336.      length = size / size_of_one
337.
338.      data = {}
339.      #data["Measured"] = np.zeros(length)
340.
341.      # Read in the whole thing at once
342.      #byte_array = uint8(fread(fid, length * size_of_one, 'uint8'));
343.      #print(" reading ",length, " data ")
344.      byte_array = np.fromfile(fid, dtype=np.uint8, count=length*size_of_one)
345.      #print("byte_array = ",byte_array[0:30], type(byte_array))
346.
347.      #byte_array = byte_array_float.astype(np.uint8)
348.
349.
350.      #data_matrix = byte_array.reshape(size_of_one, length)
351.      data_matrix = np.reshape(byte_array,(size_of_one,length),order='F')
352.
353.
354.      # Now extract the pieces we need
355.      par = np.reshape(data_matrix[0:4,:],(1,4 * length),order='F')
356.
357.
358.      dpar = par.view( np.uint32)
359.      ldpar = len(dpar[0])
360.      data["Time"] = np.zeros(ldpar)
361.      for i in range(ldpar):
362.        data["Time"][i] = float(dpar[0][i])/sampling_rate
363.
364.      #print(" par time, shape = ",data["Time"].shape , "type = ",type(data["Time"]))
365.      #print(data["Time"])
366.      par = np.reshape(data_matrix[4:8,:],( 1, 4 * length),order='F')
367.      data["Clamp"] = (par.view(np.float32)).T
368.      par = np.reshape(data_matrix[8:12,:],( 1, 4 * length),order='F')
369.      data["TotalClamp"] = (par.view(np.float32)).T
370.      par = np.reshape(data_matrix[12:16,:],( 1, 4 * length),order='F')
371.      data["Measured"] = (par.view(np.float32)).T
372.      #print(" measured ",data["Measured"])
373.
374.      return data
```

```
"""
function data = read_aux_data(data, fid, numADCs)
    start_pos = ftell(fid);
    fseek(fid, 0, 'eof');
    file_size = ftell(fid);
    fseek(fid, start_pos, 'bof');

    data.Data = read_one_batch_aux(fid, file_size - start_pos, data.Header.Settings.SamplingRate, numADCs);
end

function data = read_one_batch_aux(fid, size, sampling_rate, numADCs)
    size_of_one = 4 + 2 + 2 + 2 * numADCs;  % Timestamps + DigIn + DigOut + ADCs
    length = size / size_of_one;

    % Read in the whole thing at once
    byte_array = uint8(fread(fid, length * size_of_one, 'uint8'));
    data_matrix = reshape(byte_array, size_of_one, length);

    % Now extract the pieces we need
    data.Time = double(typecast(reshape(data_matrix(1:4,:), 1, 4 * length), 'uint32')) / sampling_rate;
    data.DigitalIn = typecast(reshape(data_matrix(5:6,:), 1, 2 * length), 'uint16');
    data.DigitalOut = typecast(reshape(data_matrix(7:8,:), 1, 2 * length), 'uint16');
    for i=1:numADCs
        data.ADC(i,:) = 0.0003125 * (double(typecast(reshape(data_matrix((2*i+7):(2*i+8),:), 1, 2 * length), 'uint16')) - 2^15);
    end
end"""

#read_Intan_CLP_file("today_AUX_210209_165136.clp")
if __name__ == "__main__" :
  if(len(sys.argv) < 2):
   print( " need file name ")
   sys.exit(-1)
  file_with_data=sys.argv[1]
  read_Intan_CLP_file(file_with_data)
```

A.4 unsersample.py

```
# unsersample.py
import sys
import array
import numpy as np
import matplotlib.pyplot as plt
from scipy.signal import bessel,filtfilt,butter
import pickle

conversion_table = {'0': 0, '1': 1, '2': 2, '3': 3, '4': 4, '5': 5, '6': 6, '7': 7, '8': 8, '9': 9, 'A': 10 , 'B': 11, 'C': 12, 'D': 13, 'E': 14, 'F': 15}


def read_Intan_CLP_file(filename):

    # d = read_Intan_CLP_file(filename) reads the data from a CLAMP
    # file from the given filename (including path).
    #
    # d = read_Intan_CLP_file prompts for a filename.
    #
    # The returned 'd' is a data structure.
    #
    # Version 1.04, 23 October 2018
    #
    # Reads Intan Technologies CLAMP data file generated by controller
    # GUI.  Data are parsed and placed into a structure.
    #
    # Example:
    # >> d = read_Intan_CLP_file;
    # >> plot(d.Data.Time, d.Data.Measured)
    # >> d = read_Intan_CLP_file('myexperiment_AUX_160916_142731.clp');
    # >> plot(d.Data.Time, d.Data.ADC(2,:))

```

```
32.
33.     rr = open(filename,"rb")
34.     data = pickle.load(rr)
35.     nrecord = len(data)
36.
37.     for i in range(nrecord):
38.       plt.plot(data[i])
39.       plt.show()
40.
41.     # Close data file.
42.     rr.close()
43.
44.
45. def read_header(fid):
46.
47.     magic = np.fromfile(fid,np.uint32,count=1)
48.
49.     themagic = hextodec('F3B1A481')
50.
51.
52.     if( magic[0] !=  themagic):
53.         print('Unrecognized file type.')
54.         sys.exit(-1)
55.
56.     header = {}
57.     # Read version number.
58.     Version = {}
59.     Version ["Major"] = np.fromfile(fid, np.int16, count=1)[0]
60.     Version [ "Minor" ] = np.fromfile(fid, np.int16, count=1)[0]
61.     header["Version"] = Version
62.
63.
64.     datatype = np.fromfile(fid, np.int16, count=1)[0]
65.     header["Datatype"] = datatype
66.     if (datatype == 0):
67.         print('Reading Intan Technologies CLAMP Data File')
68.
69.         NumBytes = np.fromfile(fid, np.uint16, count=1)[0]
70.         Date = {}
71.         Date["Year"] = np.fromfile(fid, np.int16, count=1)[0]
72.         Date["Month"] = np.fromfile(fid, np.int16, count=1)[0]
73.         Date["Day"] = np.fromfile(fid, np.int16, count=1)[0]
74.         Date["Hour"] = np.fromfile(fid, np.int16, count=1)[0]
75.         Date["Minute"] = np.fromfile(fid, np.int16, count=1)[0]
76.         Date["Second"] = np.fromfile(fid, np.int16, count=1)[0]
77.         header["Date"] = Date
78.         header["Version"] = Version
79.         header["NumBytes"] = NumBytes
80.         header["Chips"] = read_header_chips(fid)
81.         header["Settings"] = read_header_settings(fid)
82.     elif (datatype == 1):
83.         print('Reading Intan Technologies Aux CLAMP Data File')
84.
85.         NumADCs = np.fromfile(fid, np.uint16, count=1)[0]
86.         NumBytes = np.fromfile(fid, np.uint16, count=1)[0]
87.         Date = {}
88.         Date["Year"] = np.fromfile(fid, np.int16, count=1)[0]
89.         Date["Month"] = np.fromfile(fid, np.int16, count=1)[0]
90.
91.         Date["Day"] = np.fromfile(fid, np.int16, count=1)[0]
92.         Date["Hour"] = np.fromfile(fid, np.int16, count=1)[0]
93.         Date["Minute"] = np.fromfile(fid, np.int16, count=1)[0]
94.         Date["Second"] = np.fromfile(fid, np.int16, count=1)[0]
95.         header["Date"] = Date
96.         header["Version"] = Version
97.         header["NumBytes"] = NumBytes
98.         header["NumADCs"] = NumADCs
99.         header["SamplingRate"] =   np.fromfile(fid, np.float32, count=1)[0]
100.    else:
101.        error('Unrecognized data type.')
102.
103.    return [header ]
104.
105.
106. def read_header_settings(fid):
107.    settings = {}
108.    settings["EnableCapacitiveCompensation"] = np.fromfile(fid, np.uint8, count=1)[0]
109.    settings["CapCompensationMagnitude"] = np.fromfile(fid, np.float32, count=1)[0]
110.
111.    filtro = np.fromfile(fid, np.float32, count=1)[0]
112.    if (filtro > 0):
113.        settings["FilterCutoff"] = filtro
```

```
114.
115.      settings["PipetteOffset"] = np.fromfile(fid, np.float32, count=1)[0]
116.      settings["SamplingRate"] = np.fromfile(fid, np.float32, count=1)[0]
117.      settings["CellParameters_Rs"] = np.fromfile(fid, np.float32, count=1)[0]
118.      settings["CellParameters_Rm"] = np.fromfile(fid, np.float32, count=1)[0]
119.      settings["CellParameters_Cm"] = np.fromfile(fid, np.float32, count=1)[0]
120.      settings["IsVoltageClamp"] = np.fromfile(fid, np.uint8, count=1)[0]
121.      settings["vClampX2mode"] = np.fromfile(fid, np.uint8, count=1)[0]
122.      if (settings["IsVoltageClamp"] == 1):
123.          settings["VoltageClamp"] = read_header_voltage_clamp_settings(fid)
124.      else:
125.          settings["CurrentClamp"] = read_header_current_clamp_settings(fid)
126.
127.      settings["Waveform"] = read_waveform(fid)
128.      return settings
129.
130.  def read_waveform(fid):
131.
132.      waveform = {}
133.      waveform["Interval"] = np.fromfile(fid, np.float32, count=1)[0]
134.      numSegments = np.fromfile(fid, np.uint16, count=1)[0]
135.      Segments = []
136.      for i in range(numSegments):
137.          Segments.append( read_segment(fid))
138.      waveform["Segments"] = Segments
139.
140.      return waveform
141.
142.  def read_segment(fid):
143.
144.      segment = {}
145.      segment["WaveformNumber"] = np.fromfile(fid, np.uint8, count=1)[0]
146.      segment["TOffset"] =  np.fromfile(fid, np.uint32, count=1)[0]+ 1
147.      segment["Start"] = np.fromfile(fid, np.uint32, count=1)[0]+ 1
148.      segment["End"] = np.fromfile(fid, np.uint32, count=1)[0]+ 1
149.      segment["AppliedValue"] = np.fromfile(fid, np.float32, count=1)[0]
150.
151.      return segment
152.
153.  def read_header_voltage_clamp_settings(fid):
154.      settings = {}
155.      settings["HoldingVoltage"] = np.fromfile(fid, np.float32, count=1)[0]
156.      settings["NominalResistance"] = np.fromfile(fid, np.float32, count=1)[0]
157.      settings["Resistance"] = np.fromfile(fid, np.float32, count=1)[0]
158.      settings["DesiredBandwidth"] = np.fromfile(fid, np.float32, count=1)[0]
159.      settings["ActualBandwidth"] = np.fromfile(fid, np.float32, count=1)[0]
160.      return settings
161.
162.  def read_header_current_clamp_settings(fid):
163.      settings = {}
164.      settings["HoldingCurrent"] = np.fromfile(fid, np.float32, count=1)[0]
165.      settings["CurrentStepSize"] = np.fromfile(fid, np.float32, count=1)[0]
166.      return settings
167.
168.  def  read_header_chips(fid):
169.      numChips =    np.fromfile(fid, np.uint16, count=1)[0]
170.      numChannels = np.fromfile(fid, np.uint16, count=1)[0]
171.      chips = []
172.      for i in range(numChips):
173.          chips.append( read_one_header_chip(fid, numChannels) )
174.
175.      return chips
176.
177.  def read_one_header_chip(fid, numChannels):
178.      Channels = []
179.      for i in range(numChannels):
180.          Channels.append(  read_one_header_channel(fid) )
181.
182.      chip = {}
183.      chip["Channels"]=Channels
184.      chip["ChipRegisters"] = np.fromfile(fid, np.uint16, count=4)
185.      return chip
186.
```

```
187.
188.  def read_one_header_channel(fid):
189.      channel = {}
190.      channel["Registers"] = np.fromfile(fid, np.uint16, count=14)
191.      channel["DifferenceAmpResidual"] = np.fromfile(fid, np.int32, count=1)[0]
192.      channel["VoltageAmpResidual"] = np.fromfile(fid, np.int32, count=1)[0]
193.      bestCalibration = np.fromfile(fid, np.uint8, count=16)
194.      """bestCalibration = reshape(bestCalibration, 2, 4, 2);
195.      channel.CoarseCalibration = reshape(bestCalibration(1,:,:), 4, 2);
196.      channel.FineCalibration = reshape(bestCalibration(2,:,:), 4, 2);"""
197.      channel["FeedbackResistors"] = np.fromfile(fid, np.float32, count=5)
198.      channel["DesiredBandwidth"] = np.fromfile(fid, np.float32, count=1)[0]
199.      return channel
200.
201.  def read_data(data, fid,SamplingRate):
202.      start_pos = fid.tell()
203.      fid.seek(0,2)   # go to enf of file
204.      file_size = fid.tell()
205.      fid.seek(start_pos) # go to beginning of file
206.      Data = read_one_batch(fid, file_size - start_pos, SamplingRate)
207.      return Data
208.
209.  def read_data_block(fid,start_pos,len2read,SamplingRate):
210.      # same as read_data, but goes directly to the frame number nblock
211.
212.      size_of_one = 4 + 4 + 4 + 4  # Timestep + Applied (Software) + Clamp Value + Measured
213.      length = int(len2read/size_of_one)
214.      fid.seek(start_pos) # go to beginning of file + shift of previous frames read
215.      data = {}
216.      byte_array = np.fromfile(fid, dtype=np.uint8, count=len2read)
217.      data_matrix = np.reshape(byte_array,(size_of_one,length),order='F')
218.      par = np.reshape(data_matrix[0:4,:],(1,4 * length),order='F')
219.
220.
221.      dpar = par.view( np.uint32)
222.      ldpar = len(dpar[0])
223.      data["Time"] = np.zeros(ldpar)
224.      for i in range(ldpar):
225.        data["Time"][i] = float(dpar[0][i])/SamplingRate
226.
227.      #print(" par time, shape = ",data["Time"].shape , "type = ",type(data["Time"]))
228.      #print(data["Time"])
229.      par = np.reshape(data_matrix[4:8,:],( 1, 4 * length),order='F')
230.      data["Clamp"] = (par.view(np.float32)).T
231.      par = np.reshape(data_matrix[8:12,:],( 1, 4 * length),order='F')
232.      data["TotalClamp"] = (par.view(np.float32)).T
233.      par = np.reshape(data_matrix[12:16,:],( 1, 4 * length),order='F')
234.      data["Measured"] = (par.view(np.float32)).T
235.      #print(" measured ",data["Measured"])
236.
237.      return data
238.
239.  def read_one_batch(fid, size, sampling_rate):
240.
241.      size_of_one = 4 + 4 + 4 + 4  # Timestep + Applied (Software) + Clamp Value + Measured
242.      length = size / size_of_one
243.
244.      data = {}
245.      #data["Measured"] = np.zeros(length)
246.
247.      # Read in the whole thing at once
248.      #byte_array = uint8(fread(fid, length * size_of_one, 'uint8'));
249.      #print(" reading ",length, " data ")
250.      byte_array = np.fromfile(fid, dtype=np.uint8, count=length*size_of_one)
251.      #print("byte_array = ",byte_array[0:30], type(byte_array))
252.
253.      #byte_array = byte_array_float.astype(np.uint8)
254.
255.
256.      #data_matrix = byte_array.reshape(size_of_one, length)
257.      data_matrix = np.reshape(byte_array,(size_of_one,length),order='F')
258.
259.
260.      # Now extract the pieces we need
261.      par = np.reshape(data_matrix[0:4,:],(1,4 * length),order='F')
262.
263.
```

```
    dpar = par.view( np.uint32)
    ldpar = len(dpar[0])
    data["Time"] = np.zeros(ldpar)
    for i in range(ldpar):
      data["Time"][i] = float(dpar[0][i])/sampling_rate

    #print(" par time, shape = ",data["Time"].shape , "type = ",type(data["Time"]))
    #print(data["Time"])
    par = np.reshape(data_matrix[4:8,:],( 1, 4 * length),order='F')
    data["Clamp"] = (par.view(np.float32)).T
    par = np.reshape(data_matrix[8:12,:],( 1, 4 * length),order='F')
    data["TotalClamp"] = (par.view(np.float32)).T
    par = np.reshape(data_matrix[12:16,:],( 1, 4 * length),order='F')
    data["Measured"] = (par.view(np.float32)).T
    #print(" measured ",data["Measured"])

    return data
"""
function data = read_aux_data(data, fid, numADCs)
    start_pos = ftell(fid);
    fseek(fid, 0, 'eof');
    file_size = ftell(fid);
    fseek(fid, start_pos, 'bof');

    data.Data = read_one_batch_aux(fid, file_size - start_pos, data.Header.Settings.Sampling
Rate, numADCs);
end

function data = read_one_batch_aux(fid, size, sampling_rate, numADCs)
    size_of_one = 4 + 2 + 2 + 2 * numADCs;  % Timestamps + DigIn + DigOut + ADCs
    length = size / size_of_one;

    % Read in the whole thing at once
    byte_array = uint8(fread(fid, length * size_of_one, 'uint8'));
    data_matrix = reshape(byte_array, size_of_one, length);

    % Now extract the pieces we need
    data.Time = double(typecast(reshape(data_matrix(1:4,:), 1, 4 * length), 'uint32')) / sam
pling_rate;
    data.DigitalIn = typecast(reshape(data_matrix(5:6,:), 1, 2 * length), 'uint16');
    data.DigitalOut = typecast(reshape(data_matrix(7:8,:), 1, 2 * length), 'uint16');
    for i=1:numADCs
        data.ADC(i,:) = 0.0003125 * (double(typecast(reshape(data_matrix((2*i+7):
(2*i+8),:), 1, 2 * length), 'uint16')) - 2^15);
    end
end"""

#read_Intan_CLP_file("today_AUX_210209_165136.clp")
if __name__ == "__main__" :
  if(len(sys.argv) < 2):
   print( " need file name ")
   sys.exit(-1)
  file_with_data=sys.argv[1]
  read_Intan_CLP_file(file_with_data)
```

A.5 hmm.py

```
# hmm.py
#!/usr/bin/env python
# -*- coding: utf-8 -*-
from matplotlib.pyplot import figure, show,colorbar
import matplotlib.pyplot as plt
#import matplotlib.use as usem
import numpy as np
import sys
import string
from scipy.signal import bessel,filtfilt
#import heka_reader
import os
import time
import pickle
#usem('TkAgg')
```

```
16.
17.  VOLT = 0.030
18.  NAVER = 1
19.  CONDUC_MAX = 40.
20.  CNANO = 20.e-09
21.
22.  N_sampling = 80
23.  # values when N_sampling = 20
24.  N_cen_0 = 30500
25.  N_cen_1 = 31500
26.  N_neg_0 = 3500
27.  N_neg_1 = 26500
28.  N_pos_0 = 35500
29.  N_pos_1 = 108500
30.
31.  if(N_sampling == 80):
32.    # values when N_sampling = 20
33.    N_cen_0 = 7000
34.    N_cen_1 = 7100
35.    N_neg_0 = 750
36.    N_neg_1 = 6750
37.    N_pos_0 = 9000
38.    N_pos_1 = 29000
39.
40.
41.  if(len(sys.argv) < 2):
42.     print( " need file name ")
43.     sys.exit(-1)
44.
45.  print( "assume voltage == ",VOLT)
46.
47.  """try:
48.     DETAIL=int(raw_input('detailed (1) or not (0):'))
49.  except ValueError:
50.     print "please, type 0 or 1"  """
51.
52.  def plot_sweep(c):
53.     x = np.arange(len(c))
54.     y = np.asarray(c)
55.     fig = plt.figure()
56.     #plt.ion()
57.     plt.plot((x),(y))
58.     plt.draw()
59.     plt.pause(1.)
60.     plt.close(fig)
61.
62.  def plot_2sweep(c,d):
63.     x = np.arange(len(c))
64.     y = np.asarray(c)
65.     #cmean = np.mean(y)
66.     fig = plt.figure()
67.     #plt.ion()
68.     plt.plot((x),(y))
69.     x = np.arange(len(d))
70.     y = np.asarray(d)
71.     #plt.plot((x),(y),'.')
72.     plt.plot((x),(y))
73.     plt.show()
74.     #plt.draw()
75.     #plt.pause(2.)
76.     #plt.close(fig)
77.
78.  def histo_diff(x):
79.    jump = []
80.    rest = []
81.    for i in range(len(x)-1):
82.      jump.append(abs(x[i+1]-x[i]))
83.    return rest,jump
84.
85.  def lin_fit(x,y):
86.     A = np.vstack([x, np.ones(len(x))]).T
87.     m, c = np.linalg.lstsq(A,y)[0]
88.     return m,c
```

```
89.
90.  def find_jump_hmm_rawjump(c):
91.     ff = open("hmmtmp.dat","w")
92.     for i in range(len(c)):
93.       ff.write("%g \n"%c[i] )
94.     ff.close()
95.     command = 'sage ../test_hmm.sage '+"%d "%int(sys.argv[1]) + "%d "%int(sys.argv[2])
96.     os.system(command)
97.     rest = []
98.     ff = open("hmmsol.dat","r")
99.     for line in ff.readlines():
100.      rest.append(float(line.split()[2] ) )
101.    ff.close()
102.    jump = []
103.    for i in range(len(rest)-1):
104.      if(rest[i+1] != rest[i]):
105.        dd = abs(rest[i+1]-rest[i])
106.        jump.append(dd)
107.        """if(dd > 1.8):
108.          print " jump > 1.8 ", i"""
109.
110.    return rest, jump
111.
112. def find_jump_hmm(c):
113.    ff = open("hmmtmp.dat","w")
114.    for i in range(len(c)):
115.      ff.write("%g \n"%c[i] )
116.    ff.close()
117.    command = 'sage ../test_hmm.sage '+"%d "%int(sys.argv[1]) + "%d "%int(sys.argv[2])
118.    os.system(command)
119.    rest = []
120.    ff = open("hmmsol.dat","r")
121.    for line in ff.readlines():
122.      rest.append(float(line.split()[2] ) )
123.    ff.close()
124.    # first pass: look for ALL jumps
125.    jump = []
126.    tjump = []
127.    for i in range(len(rest)-1):
128.      if(rest[i+1] != rest[i]):
129.        dd = (rest[i+1]-rest[i])
130.        jump.append(dd)
131.        tjump.append(i)
132.        """if(dd > 1.8):
133.          print " jump > 1.8 ", i"""
134.    #second pass: compact jumps of same sign, separated by less than 2 (in time units)
135.
136.    ljump = len(jump)
137.    ij = -1
138.    while (ij < ljump-2):
139.      ij = ij+1
140.      if(jump[ij]*jump[ij+1] > 0.e0 and tjump[ij+1] - tjump[ij] < 3):
141.        jump[ij] =jump[ij] + jump[ij+1]
142.        jump[ij+1] = 0.
143.
144.    hmmj = []
145.    ff = open("hmmj.dat",'r')
146.    for line in ff.readlines():
147.      ll = line.split()
148.      hmmj.append((float(ll[0]),float(ll[1]) ))
149.
150.    return rest,jump,hmmj
151.
152. def clean_ramp_moins_plus(c,DETAIL=0):
153.    print( " enter clean ramp ")
154.    ramp_m = []
155.    centro = 0.
156.
157.    for i in range(N_cen_0,N_cen_1):
158.       centro = centro + c[i]
159.    centro = centro/float(N_cen_1-N_cen_0)
160.
161.    conductance = 0.
162.    for i in range(N_neg_0,N_neg_1):
163.       if(abs(c[i]) > CNANO):
164.          jumps_p=[]
165.          jumps_m=[]
166.          rampa_ave_m=np.zeros(10)
167.          rampa_ave_p=np.zeros(10)
168.          return rampa_ave_m,rampa_ave_p, jumps_m,jumps_p
169.       #conductance = max(0.e0,-(c[i]-centro)*1.e09/VOLT)
170.       conductance = -c[i]*1.e09/VOLT
171.       ramp_m.append(conductance)  # this assumes that the conductivity is nS, and that V = -
VOLT.
```

```
172.
173.     rest_m, jumps_m, hmmj_m= find_jump_hmm(ramp_m)
174.
175.     plot_exception = 0
176.     for i in range(len(jumps_m)):
177.        if(jumps_m[i] > 1.6):
178.          plot_exception = 1
179.
180.     #DETAIL = 1
181.     if(DETAIL ): #and len(jumps_m) > 0 ): #or plot_exception > 0):
182.        plot_2sweep(ramp_m,rest_m)
183.        #plot_wt(rampa_ave_m)
184.
185.     ramp_p = []
186.     #for i in range(24000,44000):
187.     for i in range(N_pos_0,N_pos_1):
188.        if(abs(c[i]) > CNANO):
189.           jumps_p=[]
190.           jumps_m=[]
191.           rampa_ave_m=np.zeros(10)
192.           rampa_ave_p=np.zeros(10)
193.           return rampa_ave_m,rampa_ave_p, jumps_m,jumps_p
194.        #conductance = max(0.e0,(c[i]-centro)*1.e09/VOLT)
195.        conductance = c[i]*1.e09/VOLT
196.        ramp_p.append(conductance)  # this assumes that the conductivity is nS, and that V = 1
    00mV.
197.
198.     rest_p, jumps_p, hmmj_p = find_jump_hmm(ramp_p)
199.
200.
201.     plot_exception = 0
202.     for i in range(len(jumps_p)):
203.        if(jumps_p[i] > 1.6):
204.          plot_exception = 1
205.
206.     if(DETAIL ): #and len(jumps_p) > 0): # or plot_exception > 0):
207.        plot_2sweep(ramp_p,rest_p)
208.        #plot_wt(rampa_ave_p)
209.     return ramp_m,ramp_p, jumps_m,jumps_p, hmmj_m,hmmj_p
210.
211.
212.  def aver(x,n):
213.     # take time series x, then average over n points, return sub-sampled time series
214.     ave = []
215.     nave = len(x)/n
216.     for i in range(0,nave):
217.        xx = 0.
218.        for j in range(n):
219.           xx = xx+ x[i*n+j]
220.        ave.append(xx/float(n))
221.     return ave
222.
223.  g = open("histo_ramp.dat","w")
224.
225.  def scan_file_onlyplot(filetoscan):
226.     f = open(filetoscan,"rb")
227.     print( "scanning file ",filetoscan)
228.     ns = -1
229.     c = []
230.
231.     data = pickle.load(f)
232.     f.close()
233.     nchildren = len(data)
234.
235.     take = np.zeros(nchildren)
236.     ns = -1
237.     mode = 1
238.     while(ns < nchildren-1 and mode != 2 ):
239.        ns=ns+1
240.        c = data[ns]
241.        plot_sweep(c)
242.        try:
243.            mode=int(input('take (1) not (0) stop (2):'))
244.            take[ns]=mode%2
245.        except ValueError:
246.            print( "Not a number")
247.            sys.exit(-2)
248.     return take
```

```
249.
250. def scan_file(filetoscan,take,iprint=0):
251.    f = open(filetoscan,"rb")
252.    print( "scanning file ",filetoscan)
253.    ns = -1
254.    c = []
255.    curr = []
256.    jumps_m = []
257.    jumps_p = []
258.    hmm_m = []
259.    hmm_p = []
260.    aver = []
261.    data = pickle.load(f)
262.    f.close()
263.    nchildren = len(data)
264.    print(" will consider",nchildren,"frames")
265.    ns = -1
266.    mode=1
267.    while(ns < nchildren-1 and mode !=2):
268.       ns=ns+1
269.       c = data[ns]
270.       if(take[ns] == 1):
271.                   clean_m, clean_p,jumps_loc_m,jumps_loc_p, hloc_m,hloc_p=clean_ramp_moins_p
    lus(c,DETAIL=iprint)
272.                   curr = curr + clean_m
273.                   curr = curr + clean_p
274.                   jumps_m = jumps_m + jumps_loc_m
275.                   jumps_p = jumps_p + jumps_loc_p
276.                   aver.append(np.average(np.asarray(clean_m)))
277.                   aver.append(np.average(np.asarray(clean_p)))
278.                   hmm_m = hmm_m + hloc_m
279.                   hmm_p = hmm_p + hloc_p
280.    return curr,jumps_m,jumps_p,aver,hmm_m,hmm_p
281.
282. if(len(sys.argv) < 4):
283.     print(" **************    OUUUUUPS ! **************")
284.     print("Usage is hmm n_states L_chunk data_file_1 data_file_2 ...")
285.     print(" Try again ")
286.     sys.exit(-2)
287.
288. nscan = 0
289. fig = plt.figure()
290.
291. ctot = []
292. jumps_mtot = []
293. jumps_ptot = []
294. cond_aver = []
295. take_or_not = []
296. hmm_mtot = []
297. hmm_ptot = []
298.
299. print("sys argv ",sys.argv)
300. # first, scan the frames to see which ones to compute
301. for i in range(3,len(sys.argv)):
302.    file_with_data=sys.argv[i]
303.    take_loc = scan_file_onlyplot(file_with_data)
304.    take_or_not.append(take_loc)
305.
306.
307. idetail = int(input("now enter HMM evaluation. Print intermediate (1) or not (0)"))
308.
309. for i in range(3,len(sys.argv)):
310.    file_with_data=sys.argv[i]
311.    curr,jumps_m,jumps_p,curr_aver,hmm_m,hmm_p = scan_file(file_with_data,take_or_not[i-
    3],iprint=idetail)
312.    ctot.append(curr)
313.    cond_aver=(curr_aver)
314.    jumps_mtot = jumps_mtot + jumps_m
315.    jumps_ptot = jumps_ptot + jumps_p
316.    hmm_mtot = hmm_mtot + hmm_m
317.    hmm_ptot = hmm_ptot + hmm_p
318.
319. print( " total of ",len(jumps_mtot),len(jumps_ptot)," jumps (m and p ) detected ")
320.
321.
```

```
named = file_with_data.replace(".dat","_listjm.dat")
f = open(named,"w")
for i in range(len(jumps_mtot)):
   f.write("%14.5g %14.5g  \n"%(float(i),jumps_mtot[i]))
f.close()
named = file_with_data.replace(".dat","_listjp.dat")
f = open(named,"w")
for i in range(len(jumps_ptot)):
   f.write("%14.5g %14.5g  \n"%(float(i),jumps_ptot[i]))
f.close()

named = file_with_data.replace(".dat","_hmm_m.dat")
f = open(named,"w")
for i in range(len(hmm_mtot)):
   f.write("%14.5g %14.5g  \n"%(hmm_mtot[i][0],hmm_mtot[i][1]))
f.close()
named = file_with_data.replace(".dat","_hmm_p.dat")
f = open(named,"w")
for i in range(len(hmm_ptot)):
   f.write("%14.5g %14.5g  \n"%(hmm_ptot[i][0],hmm_ptot[i][1]))
f.close()

#print " mirna = ",mirna
```

A.6 rj.py

```
#!/usr/bin/env python
# -*- coding: utf-8 -*-
from matplotlib.mlab import griddata
from matplotlib.pyplot import figure, show,colorbar
import matplotlib.pyplot as plt
import numpy as np
import sys
import csv

def migrep(ff,cara):
  f = open(ff,"r")
  n = 0
  for l in f.readlines():
    if(cara in l):
      n = n+1
  f.close()
  return n

file_with_data=sys.argv[1]
gfile = sys.argv[2]
bincut = int(sys.argv[3])

if(len(sys.argv) < 4):
   print " need two file name and a bincut (< 20) "
   sys.exit(-1)

DeltaT = 0.001 # time in msec

f = open(file_with_data,"r")
g = open(gfile,"r")
"""fml = file_with_data.replace('_listj.dat','.dat')
gml = gfile.replace('_listj.dat','.dat')
nf = migrep(fml,'Series')
ng = migrep(gml,'Series')
print "numero de Series ",nf,ng"""

il = -1
c = []
for row in f.readlines():
   rows = row.split()
   c.append(float(rows[1].strip(',')))
nc = len(c)
d = []
for row in g.readlines():
   rows = row.split()
   d.append(float(rows[1].strip(',')))
nd = len(d)

```

```
cmin = 1.
cmax = 4.
histo, xhisto = np.histogram(np.asarray(c),bins=20,range=(cmin,cmax))
histod, xhistod = np.histogram(np.asarray(d),bins=20,range=(cmin,cmax))
sh = 0.
for i in range(len(histo)):
  sh = sh + histo[i]
shd = 0.
for i in range(len(histod)):
  shd = shd + histod[i]

p = 0.
for i in range(bincut,len(histo)):
   p = p+histo[i]
p = p/sh
pd = 0.
for i in range(bincut,len(histod)):
   pd = pd+histod[i]
pd = pd/shd
print " pis ",p,pd
print "histo",histo
print "histod",histod

plt.plot(xhisto[0:len(histo)-1],histo[0:len(histo)-1]/sh)
plt.plot(xhistod[0:len(histod)-1],histod[0:len(histod)-1]/shd)
plt.show()
f.close()
g.close()
f = open("rj.dat","w")
for i in range(len(xhisto)-1):
   f.write("%14.5g %14.5g %14.5g \n"%(xhisto[i],histo[i]/sh,histod[i]/shd))
```

Appendix B Sequences

B.1 Sequences of M13mp18

M13mp18 [length=7249] [version=09-MAY-2008] [topology=circular] Cloning vector M13mp18, complete sequence. M13 is a filamentous E. coli bacteriophage specific for male (F factor-containing) cells, and M13mp18 is a M13 lac phage vector. The single-stranded viral DNA is a circular molecule isolated from M13mp18. The vast majority of the molecules are circular, with the remaining molecules being in the linear form, as determined by agarose gel electrophoresis. The presence of linear DNA is inherent to the purification process. M13mp18 phage is propagated in E. coli ER2738. The phage is purified by polyethylene glycol precipitation. The single-stranded DNA is then extracted with phenol. The 40 µL of 250 µg/ml ssDNA is in pH 8, 10 mM Tris-HCl, 1 mM EDTA stored at -20 ºC.

Note:

600 bases of T1 were used to design nanopore marks as green marks (cookieT1).

570 bases of T2 were used to design nanopore, which marks as blue marks (v7T2).

TAACGAGGAAAGCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAGCGC
GGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCT
TCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGAT
TTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGA
TAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAAC
ACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGAACCACCATCAAACAGGA
TTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAAT
CAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCG
CGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAA
TTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAA
TTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTACCCGGGG
ATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAA
CCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCC
GCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGA
AGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAG
ATGCACGGTTACGATGCGCCCATCTACACCAACGTGACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCAC
GGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGA
ATTATTTTTGATGGCGTTCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAATGCGAATTTTAACAAA
ATATTAACGTTTACAATTTAAATATTTGCTTATACAATCTTCCTGTTTTTGGGGCTTTTCTGATTATCAACCGGG
GTACATATGATTGACATGCTAGTTTTACGATTACCGTTCATCGATTCTCTTGTTTGCTCCAGACTCTCAGGCAAT
GACCTGATAGCCTTTGTAGATCTCTCAAAAATAGCTACCCTCTCCGGCATTAATTTATCAGCTAGAACGGTTGA
ATATCATATTGATGGTGATTTGACTGTCTCCGGCCTTTCTCACCCTTTTGAATCTTTACCTACACATTACTCAGG
CATTGCATTTAAAATATATGAGGGTTCTAAAAATTTTTATCCTTGCGTTGAAATAAAGGCTTCTCCCGCAAAAG
TATTACAGGGTCATAATGTTTTTGGTACAACCGATTTAGCTTTATGCTCTGAGGCTTTATTGCTTAATTTTGCTA

ATTCTTTGCCTTGCCTGTATGATTTATTGGATGTTAATGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAG
CTCGCGCCCCAAATGAAAATATAGCTAAACAGGTTATTGACCATTTGCGAAATGTATCTAATGGTCAAACTAA
ATCTACTCGTTCGCAGAATTGGGAATCAACTGTTATATGGAATGAAACTTCCAGACACCGTACTTTAGTTGCAT
ATTTAAAACATGTTGAGCTACAGCATTATATTCAGCAATTAAGCTCTAAGCCATCCGCAAAAATGACCTCTTAT
CAAAAGGAGCAATTAAAGGTACTCTCTAATCCTGACCTGTTGGAGTTTGCTTCCGGTCTGGTTCGCTTTGAAGC
TCGAATTAAAACGCGATATTTGAAGTCTTTCGGGCTTCCTCTTAATCTTTTTGATGCAATCCGCTTTGCTTCTGA
CTATAATAGTCAGGGTAAAGACCTGATTTTTGATTTATGGTCATTCTCGTTTTCTGAACTGTTTAAAGCATTTGA
GGGGGATTCAATGAATATTTATGACGATTCCGCAGTATTGGACGCTATCCAGTCTAAACATTTTACTATTACCC
CCTCTGGCAAAACTTCTTTTGCAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTCTGGTAAACGAGGGTTATG
ATAGTGTTGCTCTTACTATGCCTCGTAATTCCTTTTGGCGTTATGTATCTGCATTAGTTGAATGTGGTATTCCTA
AATCTCAACTGATGAATCTTTCTACCTGTAATAATGTTGTTCCGTTAGTTCGTTTTATTAACGTAGATTTTTCTTC
CCAACGTCCTGACTGGTATAATGAGCCAGTTCTTAAAATCGCATAAGGTAATTCACAATGATTAAAGTTGAAAT
TAAACCATCTCAAGCCCAATTTACTACTCGTTCTGGTGTTTCTCGTCAGGGCAAGCCTTATTCACTGAATGAGC
AGCTTTGTTACGTTGATTTGGGTAATGAATATCCGGTTCTTGTCAAGATTACTCTTGATGAAGGTCAGCCAGCC
TATGCGCCTGGTCTGTACACCGTTCATCTGTCCTCTTTCAAAGTTGGTCAGTTCGGTTCCCTTATGATTGACCGT
CTGCGCCTCGTTCCGGCTAAGTAACATGGAGCAGGTCGCGGATTTCGACACAATTTATCAGGCGATGATACAA
ATCTCCGTTGTACTTTGTTTCGCGCTTGGTATAATCGCTGGGGGTCAAAGATGAGTGTTTTAGTGTATTCTTTTG
CCTCTTTCGTTTTAGGTTGGTGCCTTCGTAGTGGCATTACGTATTTTACCCGTTTAATGGAAACTTCCTCATGAA
AAAGTCTTTAGTCCTCAAAGCCTCTGTAGCCGTTGCTACCCTCGTTCCGATGCTGTCTTTCGCTGCTGAGGGTGA
CGATCCCGCAAAAGCGGCCTTTAACTCCCTGCAAGCCTCAGCGACCGAATATATCGGTTATGCGTGGGCGATG
GTTGTTGTCATTGTCGGCGCAACTATCGGTATCAAGCTGTTTAAGAAATTCACCTCGAAAGCAAGCTGATAAAC
CGATACAATTAAAGGCTCCTTTTGGAGCCTTTTTTTTGGAGATTTTCAACGTGAAAAAATTATTATTCGCAATTC
CTTTAGTTGTTCCTTTCTATTCTCACTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAAATCCCATACAGAAAATT
CATTTACTAACGTCTGGAAAGACGACAAAACTTTAGATCGTTACGCTAACTATGAGGGCTGTCTGTGGAATGCT
ACAGGCGTTGTAGTTTGTACTGGTGACGAAACTCAGTGTTACGGTACATGGGTTCCTATTGGGCTTGCTATCCC
TGAAAATGAGGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTACTAAACCT
CCTGAGTACGGTGATACACCTATTCCGGGCTATACTTATATCAACCCTCTCGACGGCACTTATCCGCCTGGTAC
TGAGCAAAACCCCGCTAATCCTAATCCTTCTCTTGAGGAGTCTCAGCCTCTTAATACTTTCATGTTTCAGAATAA
TAGGTTCCGAAATAGGCAGGGGGCATTAACTGTTTATACGGGCACTGTTACTCAAGGCACTGACCCCGTTAAA
ACTTATTACCAGTACACTCCTGTATCATCAAAAGCCATGTATGACGCTTACTGGAACGGTAAATTCAGAGACTG
CGCTTTCCATTCTGGCTTTAATGAGGATTTATTTGTTTGTGAATATCAAGGCCAATCGTCTGACCTGCCTCAACC
TCCTGTCAATGCTGGCGGCGGCTCTGGTGGTGGTTCTGGTGGCGGCTCTGAGGGTGGTGGCTCTGAGGGTGGCG
GTTCTGAGGGTGGCGGCTCTGAGGGAGGCGGTTCCGGTGGTGGCTCTGGTTCCGGTGATTTTGATTATGAAAAG
ATGGCAAACGCTAATAAGGGGGCTATGACCGAAAATGCCGATGAAAACGCGCTACAGTCTGACGCTAAAGGC
AAACTTGATTCTGTCGCTACTGATTACGGTGCTGCTATCGATGGTTTCATTGGTGACGTTTCCGGCCTTGCTAAT
GGTAATGGTGCTACTGGTGATTTTGCTGGCTCTAATTCCCAAATGGCTCAAGTCGGTGACGGTGATAATTCACC
TTTAATGAATAATTTCCGTCAATATTTACCTTCCCTCCCTCAATCGGTTGAATGTCGCCCTTTTGTCTTTGGCGCT
GGTAAACCATATGAATTTTCTATTGATTGTGACAAAATAAACTTATTCCGTGGTGTCTTTGCGTTTCTTTTATAT
GTTGCCACCTTTATGTATGTATTTTCTACGTTTGCTAACATACTGCGTAATAAGGAGTCTTAATCATGCCAGTTC
TTTTGGGTATTCCGTTATTATTGCGTTTCCTCGGTTTCCTTCTGGTAACTTTGTTCGGCTATCTGCTTACTTTTCTT
AAAAAGGGCTTCGGTAAGATAGCTATTGCTATTTCATTGTTTCTTGCTCTTATTATTGGGCTTAACTCAATTCTT
GTGGGTTATCTCTCTGATATTAGCGCTCAATTACCCTCTGACTTTGTTCAGGGTGTTCAGTTAATTCTCCCGTCT
AATGCGCTTCCCTGTTTTTATGTTATTCTCTCTGTAAAGGCTGCTATTTTCATTTTTGACGTTAAACAAAAAATC
GTTTCTTATTTGGATTGGGATAAATAATATGGCTGTTTATTTTGTAACTGGCAAATTAGGCTCTGGAAAGACGC
TCGTTAGCGTTGGTAAGATTCAGGATAAAATTGTAGCTGGGTGCAAAATAGCAACTAATCTTGATTTAAGGCTT

CAAAACCTCCCGCAAGTCGGGAGGTTCGCTAAAACGCCTCGCGTTCTTAGAATACCGGATAAGCCTTCTATATC
TGATTTGCTTGCTATTGGGCGCGGTAATGATTCCTACGATGAAAATAAAAACGGCTTGCTTGTTCTCGATGAGT
GCGGTACTTGGTTTAATACCCGTTCTTGGAATGATAAGGAAAGACAGCCGATTATTGATTGGTTTCTACATGCT
CGTAAATTAGGATGGGATATTATTTTTCTTGTTCAGGACTTATCTATTGTTGATAAACAGGCGCGTTCTGCATTA
GCTGAACATGTTGTTTATTGTCGTCGTCTGGACAGAATTACTTTACCTTTTGTCGGTACTTTATATTCTCTTATTA
CTGGCTCGAAAATGCCTCTGCCTAAATTACATGTTGGCGTTGTTAAATATGGCGATTCTCAATTAAGCCCTACT
GTTGAGCGTTGGCTTTATACTGGTAAGAATTTGTATAACGCATATGATACTAAACAGGCTTTTTCTAGTAATTA
TGATTCCGGTGTTTATTCTTATTTAACGCCTTATTTATCACACGGTCGGTATTTCAAACCATTAAATTTAGGTCA
GAAGATGAAATTAACTAAAATATATTTGAAAAAGTTTTCTCGCGTTCTTTGTCTTGCGATTGGATTTGCATCAG
CATTTACATATAGTTATATAACCCAACCTAAGCCGGAGGTTAAAAAGGTAGTCTCTCAGACCTATGATTTTGAT
AAATTCACTATTGACTCTTCTCAGCGTCTTAATCTAAGCTATCGCTATGTTTTCAAGGATTCTAAGGGAAAATTA
ATTAATAGCGACGATTTACAGAAGCAAGGTTATTCACTCACATATATTGATTTATGTACTGTTTCCATTAAAAA
AGGTAATTCAAATGAAATTGTTAAATGTAATTAATTTTGTTTTCTTGATGTTTGTTTCATCATCTTCTTTTGCTCA
GGTAATTGAAATGAATAATTCGCCTCTGCGCGATTTTGTAACTTGGTATTCAAAGCAATCAGGCGAATCCGTTA
TTGTTTCTCCCGATGTAAAAGGTACTGTTACTGTATATTCATCTGACGTTAAACCTGAAAATCTACGCAATTTCT
TTATTTCTGTTTTACGTGCAAATAATTTTGATATGGTAGGTTCTAACCCTTCCATTATTCAGAAGTATAATCCAA
ACAATCAGGATTATATTGATGAATTGCCATCATCTGATAATCAGGAATATGATGATAATTCCGCTCCTTCTGGT
GGTTTCTTTGTTCCGCAAAATGATAATGTTACTCAAACTTTTAAAATTAATAACGTTCGGGCAAAGGATTTAAT
ACGAGTTGTCGAATTGTTTGTAAAGTCTAATACTTCTAAATCCTCAAATGTATTATCTATTGACGGCTCTAATCT
ATTAGTTGTTAGTGCTCCTAAAGATATTTTAGATAACCTTCCTCAATTCCTTTCAACTGTTGATTTGCCAACTGA
CCAGATATTGATTGAGGGTTTGATATTTGAGGTTCAGCAAGGTGATGCTTTAGATTTTTCATTTGCTGCTGGCTC
TCAGCGTGGCACTGTTGCAGGCGGTGTTAATACTGACCGCCTCACCTCTGTTTTATCTTCTGCTGGTGGTTCGTT
CGGTATTTTTAATGGCGATGTTTTAGGGCTATCAGTTCGCGCATTAAAGACTAATAGCCATTCAAAAATATTGT
CTGTGCCACGTATTCTTACGCTTTCAGGTCAGAAGGGTTCTATCTCTGTTGGCCAGAATGTCCCTTTTATTACTG
GTCGTGTGACTGGTGAATCTGCCAATGTAAATAATCCATTTCAGACGATTGAGCGTCAAAATGTAGGTATTTCC
ATGAGCGTTTTTCCTGTTGCAATGGCTGGCGGTAATATTGTTCTGGATATTACCAGCAAGGCCGATAGTTTGAG
TTCTTCTACTCAGGCAAGTGATGTTATTACTAATCAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGAC
AGACTCTTTTACTCGGTGGCCTCACTGATTATAAAAACACTTCTCAGGATTCTGGCGTACCGTTCCTGTCTAAA
ATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATTC

B.2 Previously reported DNA rectangular shaped DNA origami

Schematic diagram and detailed sequences of previously reported DNA origami are presented here. We modified this structure to T1 structure and studied it in the chapter 4.2 of this thesis.

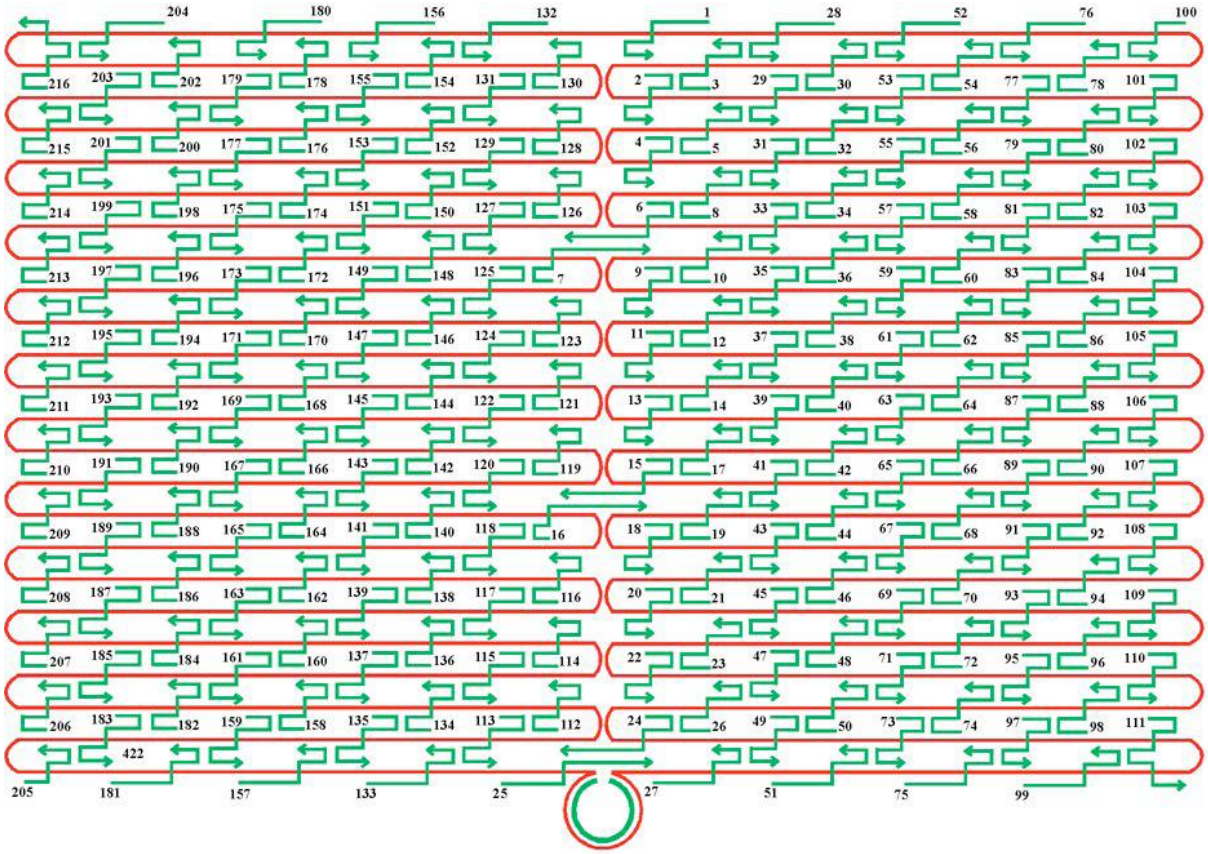


Figure S1. *The schematic diagram of the rectangular shaped DNA origami reported by Hao Yan's team.*[110]

Table S1. *This is the detailed original sequences of 216 staples for the above DNA tile from Hao Yan's work.*[110]

Name	Sequence
staple1	CAAGCCCAATAGGAAC CCATGTACAAACAGTT
staple2	AATGCCCCGTAACAGT GCCCGTATCTCCCTCA
staple3	TGCCTTGACTGCCTAT TTCGGAACAGGGATAG
staple4	GAGCCGCCCCACCACC GGAACCGCGACGGAAA
staple5	AACCAGAGACCCTCAG AACCGCCAGGGGTCAG
staple6	TTATTCATAGGGAAGG TAAATATT CATTCAGT
staple7	CATAACCCGAGGCATAGTAAGAGCTTTTTAAG
staple8	ATTGAGGGTAAAGGTG AATTATCAATCACCGG
staple9	AAAAGTAATATCTTACCGAAGCCCTTCCAGAG
staple10	GCAATAGCGCAGATAG CCGAACAATTCAACCG
staple11	CCTAATTTACGCTAAC GAGCGTCTAATCAATA
staple12	TCTTACCAGCCAGTTACAAAATAAATGAAATA
staple13	ATCGGCTGCGAGCATG TAGAAACCTATCATAT
staple14	CTAATTTATCTTTCCTTATCATTCATCCTGAA
staple15	GCGTTATAGAAAAAGC CTGTTTAG AAGGCCGG
staple16	GCTCATTTTCGCATTAAATTTTTGAGCTTAGA
staple17	AATTACTACAAATTCTTACCAGTAATCCCATC
staple18	TTAAGACGTTGAAAACATAGCGATAACAGTAC
staple19	TAGAATCCCTGAGAAG AGTCAATAGGAATCAT
staple20	CTTTTACACAGATGAA TATACAGTAAACAATT
staple21	TTTAACGTTCGGGAGA AACAATAATTTTCCCT
staple22	CGACAACTAAGTATTA GACTTTACAATACCGA
staple23	GGATTTAGCGTATTAA ATCCTTTGTTTTCAGG
staple24	ACGAACCAAAACATCG CCATTAAA TGGTGGTT
staple25	GAACGTGGCGAGAAAG GAAGGGAA CAAACTAT
staple26	TAGCCCTACCAGCAGA AGATAAAAACATTTGA
staple27	CGGCCTTGCTGGTAAT ATCCAGAACGAACTGA
staple28	CTCAGAGCCACCACCC TCATTTTCCTATTATT
staple29	CTGAAACAGGTAATAA GTTTTAACCCCTCAGA
staple30	AGTGTACTTGAAAGTA TTAAGAGGCCGCCACC
staple31	GCCACCACTCTTTTCATAATCAAACCGTCACC
staple32	GTTTGCCACCTCAGAG CCGCCACCGATACAGG
staple33	GACTTGAGAGACAAAA GGGCGACAAGTTACCA
staple34	AGCGCCAACCATTTGG GAATTAGATTATTAGC
staple35	GAAGGAAAATAAGAGCAAGAAACAACAGCCAT
staple36	GCCCAATACCGAGGAA ACGCAATAGGTTTACC
staple37	ATTATTTAACCCAGCT ACAATTTTCAAGAACG
staple38	TATTTTGCTCCCAATC CAAATAAGTGAGTTAA
staple39	GGTATTAAGAACAAGAAAAATAATTAAAGCCA
staple40	TAAGTCCTACCAAGTA CCGCACTCTTAGTTGC
staple41	ACGCTCAAAATAAGAA TAAACACCGTGAATTT
staple42	AGGCGTTACAGTAGGG CTTAATTGACAATAGA
staple43	ATCAAAATCGTCGCTATTAATTAACGGATTCG
staple44	CTGTAAATCATAGGTC TGAGAGACGATAAATA
staple45	CCTGATTGAAAGAAAT TGCGTAGACCCGAACG
staple46	ACAGAAATCTTTGAAT ACCAAGTTCCTTGCTT
staple47	TTATTAATGCCGTCAATAGATAATCAGAGGTG
staple48	AGATTAGATTTAAAAG TTTGAGTACACGTAAA
staple49	AGGCGGTCATTAGTCT TTAATGCGCAATATTA
staple50	GAATGGCTAGTATTAA CACCGCCTCAACTAAT
staple51	CCGCCAGCCATTGCAA CAGGAAAAATATTTTT
staple52	CCCTCAGAACCGCCAC CCTCAGAACTGAGACT
staple53	CCTCAAGAATACATGG CTTTTGATAGAACCAC
staple54	TAAGCGTCGAAGGATT AGGATTAGTACCGCCA
staple55	CACCAGAGTTCGGTCA TAGCCCCCGCCAGCAA
staple56	TCGGCATTCCGCCGCCAGCATTGACGTTCCAG
staple57	AATCACCAAATAGAAA ATTCATATATAACGGA
staple58	TCACAATCGTAGCACC ATTACCATCGTTTTCA
staple59	ATACCCAAGATAACCC ACAAGAATAAACGATT
staple60	ATCAGAGAAAGAACTGGCATGATTTTATTTTG

staple61	TTTTGTTTAAGCCTTA AATCAAGAATCGAGAA
staple62	AGGTTTTGAACGTCAA AAATGAAAGCGCTAAT
staple63	CAAGCAAGACGCGCCT GTTTATCAAGAATCGC
staple64	AATGCAGACCGTTTTTATTTTCATCTTGCGGG
staple65	CATATTTAGAAATACC GACCGTGTTACCTTTT
staple66	AATGGTTTACAACGCC AACATGTAGTTCAGCT
staple67	TAACCTCCATATGTGA GTGAATAAACAAAATC
staple68	AAATCAATGGCTTAGGTTGGGTTACTAAATTT
staple69	GCGCAGAGATATCAAA ATTATTTGACATTATC
staple70	AACCTACCGCGAATTA TTCATTTCCAGTACAT
staple71	ATTTTGCGTCTTTAGG AGCACTAAGCAACAGT
staple72	CTAAAATAGAACAAAGAAACCACCAGGGTTAG
staple73	GCCACGCTATACGTGG CACAGACAACGCTCAT
staple74	GCGTAAGAGAGAGCCA GCAGCAAAAAGGTTAT
staple75	GGAAATACCTACATTT TGACGCTCACCTGAAA
staple76	TATCACCGTACTCAGGAGGTTTAGCGGGGTTT
staple77	TGCTCAGTCAGTCTCT GAATTTACCAGGAGGT
staple78	GGAAAGCGACCAGGCG GATAAGTGAATAGGTG
staple79	TGAGGCAGGCGTCAGA CTGTAGCGTAGCAAGG
staple80	TGCCTTTAGTCAGACG ATTGGCCTGCCAGAAT
staple81	CCGGAAACACACCACG GAATAAGTAAGACTCC
staple82	ACGCAAAGGTCACCAA TGAAACCAATCAAGTT
staple83	TTATTACGGTCAGAGG GTAATTGAATAGCAGC
staple84	TGAACAAACAGTATGT TAGCAAACTAAAAGAA
staple85	CTTTACAGTTAGCGAA CCTCCCGACGTAGGAA
staple86	GAGGCGTTAGAGAATA ACATAAAAGAACACCC
staple87	TCATTACCCGACAATA AACAACATATTTAGGC
staple88	CCAGACGAGCGCCCAA TAGCAAGCAAGAACGC
staple89	AGAGGCATAATTTCAT CTTCTGACTATAACTA
staple90	TTTTAGTTTTTCGAGC CAGTAATAAATTCTGT
staple91	TATGTAAACCTTTTTT AATGGAAAAATTACCT
staple92	TTGAATTATGCTGATG CAAATCCACAAATATA
staple93	GAGCAAAAACTTCTGA ATAATGGAAGAAGGAG
staple94	TGGATTATGAAGATGA TGAAACAAAATTTCAT
staple95	CGGAATTATTGAAAGG AATTGAGGTGAAAAAT
staple96	ATCAACAGTCATCATA TTCCTGATTGATTGTT
staple97	CTAAAGCAAGATAGAA CCCTTCTGAATCGTCT
staple98	GCCAACAGTCACCTTG CTGAACCTGTTGGCAA
staple99	GAAATGGATTATTTAC ATTGGCAGACATTCTG
staple100	TTTT TATAAGTA TAGCCCGGCCGTCGAG
staple101	AGGGTTGA TTTT ATAAATCC TCATTAAATGATATTC
staple102	ACAAACAA TTTT AATCAGTA GCGACAGATCGATAGC
staple103	AGCACCGT TTTT TAAAGGTG GCAACATAGTAGAAAA
staple104	TACATACA TTTT GACGGGAG AATTAACTACAGGGAA
staple105	GCGCATTA TTTT GCTTATCC GGTATTCTAAATCAGA
staple106	TATAGAAG TTTT CGACAAAA GGTAAAGTAGAGAATA
staple107	TAAAGTAC TTTT CGCGAGAA AACTTTTTATCGCAAG
staple108	ACAAAGAA TTTT ATTAATTA CATTTAACACATCAAG
staple109	AAAACAAA TTTT TTCATCAA TATAATCCTATCAGAT
staple110	GATGGCAA TTTT AATCAATA TCTGGTCACAAATATC
staple111	AAACCCTC TTTT ACCAGTAA TAAAAGGGATTCACCA GTCACACG TTTT
staple112	CCGAAATCCGAAAATC CTGTTTGAAGCCGGAA
staple113	CCAGCAGGGGCAAAAT CCCTTATAAAGCCGGC
staple114	GCATAAAGTTCCACAC AACATACGAAGCGCCA
staple115	GCTCACAATGTAAAGC CTGGGGTGGGTTTGCC
staple116	TTCGCCATTGCCGGAA ACCAGGCATTAAATCA
staple117	GCTTCTGGTCAGGCTG CGCAACTGTGTTATCC
staple118	GTTAAAATTTTAACCA ATAGGAACCCGGCACC
staple119	AGACAGTCATTCAAAA GGGTGAGAAGCTATAT
staple120	AGGTAAAGAAATCACCATCAATATAATATTTT
staple121	TTTCATTTGGTCAATA ACCTGTTTATATCGCG
staple122	TCGCAAATGGGGCGCGAGCTGAAATAATGTGT
staple123	TTTTAATTGCCCGAAA GACTTCAAAACACTAT
staple124	AAGAGGAACGAGCTTCAAAGCGAAGATACATT

staple125	GGAATTACTCGTTTAC CAGACGACAAAAGATT
staple126	GAATAAGGACGTAACA AAGCTGCTCTAAAACA
staple127	CCAAATCACTTGCCCT GACGAGAACGCCAAAA
staple128	CTCATCTTGAGGCAAA AGAATACAGTGAATTT
staple129	AAACGAAATGACCCCC AGCGATTATTCATTAC
staple130	CTTAAACATCAGCTTG CTTTCGAGCGTAACAC
staple131	TCGGTTTAGCTTGATA CCGATAGTCCAACCTA
staple132	TGAGTTTCGTCACCAG TACAAACTTAATTGTA
staple133	CCCCGATTTAGAGCTT GACGGGGAAATCAAAA
staple134	GAATAGCCGCAAGCGG TCCACGCTCCTAATGA
staple135	GAGTTGCACGAGATAG GGTTGAGTAAGGGAGC
staple136	GTGAGCTAGTTTCCTG TGTGAAATTTGGGAAG
staple137	TCATAGCTACTCACATTAATTGCGCCCTGAGA
staple138	GGCGATCGCACTCCAG CCAGCTTTGCCATCAA
staple139	GAAGATCGGTGCGGGC CTCTTCGCAATCATGG
staple140	AAATAATTTTAAATTG TAAACGTTGATATTCA
staple141	GCAAATATCGCGTCTGGCCTTCCTGGCCTCAG
staple142	ACCGTTCTAAATGCAA TGCCTGAGAGGTGGCA
staple143	TATATTTTAGCTGATA AATTAATGTTGTATAA
staple144	TCAATTCTTTTAGTTT GACCATTACCAGACCG
staple145	CGAGTAGAACTAATAGTAGTAGCAAACCCTCA
staple146	GAAGCAAAAAAGCGGA TTGCATCAGATAAAAA
staple147	TCAGAAGCCTCCAACA GGTCAGGATCTGCGAA
staple148	CCAAAATATAATGCAG ATACATAAACACCAGA
staple149	CATTCAACGCGAGAGGCTTTTGCATATTATAG
staple150	ACGAGTAGTGACAAGA ACCGGATATACCAAGC
staple151	AGTAATCTTAAATTGG GCTTGAGAGAATACCA
staple152	GCGAAACATGCCACTA CGAAGGCATGCGCCGA
staple153	ATACGTAAAAGTACAACGGAGATTTCATCAAG
staple154	CAATGACACTCCAAAA GGAGCCTTACAACGCC
staple155	AAAAAAGGACAACCAT CGCCCACGCGGGTAAA
staple156	TGTAGCATTCCACAGA CAGCCCTCATCTCCAA
staple157	GTAAAGCACTAAATCG GAACCCTAGTTGTTCC
staple158	AGTTTGGAGCCCTTCA CCGCCTGGTTGCGCTC
staple159	AGCTGATTACAAGAGT CCACTATTGAGGTGCC
staple160	ACTGCCCGCCGAGCTC GAATTCGTTATTACGC
staple161	CCCGGGTACTTTCCAGTCGGGAAACGGGCAAC
staple162	CAGCTGGCGGACGACG ACAGTATCGTAGCCAG
staple163	GTTTGAGGGAAAGGGG GATGTGCTAGAGGATC
staple164	CTTTCATCCCCAAAAA CAGGAAGACCGGAGAG
staple165	AGAAAAGCAACATTAAATGTGAGCATCTGCCA
staple166	GGTAGCTAGGATAAAA ATTTTTAGTTAACATC
staple167	CAACGCAATTTTTGAG AGATCTACTGATAATC
staple168	CAATAAATACAGTTGA TTCCCAATTTAGAGAG
staple169	TCCATATACATACAGGCAAGGCAACTTTATTT
staple170	TACCTTTAAGGTCTTT ACCCTGACAAAGAAGT
staple171	CAAAAATCATTGCTCC TTTTGATAAGTTTCAT
staple172	TTTGCCAGATCAGTTG AGATTTAGTGGTTTAA
staple173	AAAGATTCAGGGGGTAATAGTAAACCATAAAT
staple174	TTTCAACTATAGGCTG GCTGACCTTGTATCAT
staple175	CCAGGCGCTTAATCAT TGTGAATTACAGGTAG
staple176	CGCCTGATGGAAGTTT CCATTAAACATAACCG
staple177	TTTCATGAAAATTGTGTCGAAATCTGTACAGA
staple178	ATATATTCTTTTTTCA CGTTGAAAATAGTTAG
staple179	AATAATAAGGTCGCTG AGGCTTGCAAAGACTT
staple180	CGTAACGATCTAAAGT TTTGTCGTGAATTGCG
staple181	ACCCAAATCAAGTTTT TTGGGGTCAAAGAACG
staple182	TGGACTCCCTTTTCAC CAGTGAGACCTGTCGT
staple183	TGGTTTTTAACGTCAA AGGGCGAAGAACCATC
staple184	GCCAGCTGCCTGCAGG TCGACTCTGCAAGGCG
staple185	CTTGCATGCATTAATG AATCGGCCCGCCAGGG
staple186	ATTAAGTTCGCATCGT AACCGTGCGAGTAACA
staple187	TAGATGGGGGGTAACG CCAGGGTTGTGCCAAG
staple188	ACCCGTCGTCATATGT ACCCCGGTAAAGGCTA

staple189	CATGTCAAGATTCTCC GTGGGAACCGTTGGTG
staple190	TCAGGTCACTTTTGCG GGAGAAGCAGAATTAG
staple191	CTGTAATATTGCCTGA GAGTCTGGAAAACTAG
staple192	CAAAATTAAAGTACGG TGTCTGGAAGAGGTCA
staple193	TGCAACTAAGCAATAA AGCCTCAGTTATGACC
staple194	TTTTTGCGCAGAAAAC GAGAATGAATGTTTAG
staple195	AAACAGTTGATGGCTT AGAGCTTATTTAAATA
staple196	ACTGGATAACGGAACA ACATTATTACCTTATG
staple197	ACGAACTAGCGTCCAA TACTGCGGAATGCTTT
staple198	CGATTTTAGAGGACAG ATGAACGGCGCGACCT
staple199	CTTTGAAAAGAACTGG CTCATTATTTAATAAA
staple200	GCTCCATGAGAGGCTT TGAGGACTAGGGAGTT
staple201	ACGGCTACTTACTTAG CCGGAACGCTGACCAA
staple202	AAAGGCCGAAAGGAAC AACTAAAGCTTTCCAG
staple203	GAGAATAGCTTTTGCG GGATCGTCGGGTAGCA
staple204	ACGTTAGTAAATGAAT TTTCTGTAAGCGGAGT
staple205	TTTT CGATGGCC CACTACGTAAACCGTC
staple206	TATCAGGG TTTT CGGTTTGC GTATTGGGAACGCGCG
staple207	GGGAGAGG TTTT TGTAAAAC GACGGCCATTCCCAGT
staple208	CACGACGT TTTT GTAATGGG ATAGGTCAAAACGGCG
staple209	GATTGACC TTTT GATGAACG GTAATCGTAGCAAACA
staple210	AGAGAATC TTTT GGTTGTAC CAAAAACAAGCATAAA
staple211	GCTAAATC TTTT CTGTAGCT CAACATGTATTGCTGA
staple212	ATATAATG TTTT CATTGAAT CCCCCTCAAATCGTCA
staple213	TAAATATT TTTT GGAAGAAA AATCTACGACCAGTCA
staple214	GGACGTTG TTTT TCATAAGG GAACCGAAAGGCGCAG
staple215	ACGGTCAA TTTT GACAGCAT CGGAACGAACCCTCAG
staple216	CAGCGAAAA TTTT ACTTTCA ACAGTTTCTGGGATTT TGCTAAAC TTTT

B.3 Detailed sequences for corkT1

Here are the sequences for ‘cork’ and T1 design in corkT1 origami. The modified strands will be illustrated in the following tables.

Table S2. *In this sheet, modified staple sequences with DNA extensions for cholesteryl attachment to the separated T1 tile are given.*

blank staple + extension	Sequences of modified staple (5’-3’)
staple31_choltag:	GTGAAGGAGAAAAAATTAATGGCCACCACTCTTTTCATAATCAAACCGTCACC
staple35_choltag	GTGAAGGAGAAAAAATTAATGGAAGGAAAATAAGAGCAAGAAACAACAGCCAT
staple39_choltag	GTGAAGGAGAAAAAATTAATGGGTATTAAGAACAAGAAAAATAATTAAAGCCA
staple43_choltag	GTGAAGGAGAAAAAATTAATGATCAAAATCGTCGCTATTAATTAACGGATTCG
staple47_choltag	GTGAAGGAGAAAAAATTAATGTTATTAATGCCGTCAATAGATAATCAGAGGTG
staple56_choltag	GTGAAGGAGAAAAAATTAATGTCGGCATTCCGCCGCCAGCATTGACGTTCCAG
staple60_choltag	GTGAAGGAGAAAAAATTAATGATCAGAGAAAGAACTGGCATGATTTTATTTTG
staple64_choltag	GTGAAGGAGAAAAAATTAATGAATGCAGACCGTTTTTATTTTCATCTTGCGGG
staple68_choltag	GTGAAGGAGAAAAAATTAATGAAATCAATGGCTTAGGTTGGGTTACTAAATTT
staple72_choltag	GTGAAGGAGAAAAAATTAATGCTAAAATAGAACAAAGAAACCACCAGGGTTAG
staple76_choltag	GTGAAGGAGAAAAAATTAATGTATCACCGTACTCAGGAGGTTTAGCGGGGTTT
staple137_choltag	GTGAAGGAGAAAAAATTAATGTCATAGCTACTCACATTAATTGCGCCCTGAGA
staple141_choltag	GTGAAGGAGAAAAAATTAATGGCAAATATCGCGTCTGGCCTTCCTGGCCTCAG
staple145_choltag	GTGAAGGAGAAAAAATTAATGCGAGTAGAACTAATAGTAGTAGCAAACCCTCA
staple149_choltag	GTGAAGGAGAAAAAATTAATGCATTCAACGCGAGAGGCTTTTGCATATTATAG
staple153_choltag	GTGAAGGAGAAAAAATTAATGATACGTAAAAGTACAACGGAGATTTCATCAAG
staple161_choltag	GTGAAGGAGAAAAAATTAATGCCCGGGTACTTTCCAGTCGGGAAACGGGCAAC
staple165_choltag	GTGAAGGAGAAAAAATTAATGAGAAAAGCAACATTAAATGTGAGCATCTGCCA
staple169_choltag	GTGAAGGAGAAAAAATTAATGTCCATATACATACAGGCAAGGCAACTTTATTT
staple173_choltag	GTGAAGGAGAAAAAATTAATGAAAGATTCAGGGGGTAATAGTAAACCATAAAT
staple177_choltag	GTGAAGGAGAAAAAATTAATGTTTCATGAAAATTGTGTCGAAATCTGTACAGA
pore_4f_cholesteryl also called T1-chol	CATTAATTTTTTCTCCTTCAC_cholesteryl
input strand	TAGCTTATCAGACTGATGTTGA

Note: we also ordered staple184, 188, 192, 196, and 200 with extensions, but the direction of these five staples is facing the other side of rectangle comparing to other modified staples.

Table S3. *The detailed 25 sequences and 8 near-by staples for corkT1.*

Name	Sequences (5'-3')
pore1-3T	GCGGGGAGCGTATTAGAGTTGTTT
pore2-3T	TTTTGTTCCAAATAGCCAAGCGGT
pore3	AGTGAGATGTCGTGACGTGGATTT
pore4	ATCGGCATTAAAGACCAGCTGCATTAATTTTTTCTCCTTCAC
pore5	TTTCAACAGCATCCTGTTTCCGAA
pore6	TCCACTAAAATCCCCCCAGCAGGCGAAATGATTGCTTTCACC
pore 7	TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACA
pore 8b	GGGATTTTGCCGATTTCGGAA
pore 9a	ACAGGATTTTCGCCTGCTGGG
pore 10_loop10	CAACTCTCTCAGGGCCAGGCGTAGCTTATCAGATCAACATCAGTCTGATAAG CTAGTGAAGGGCAATCAGCTGTTG
pore 11a	TCTCACTGGTGAAAAGAAAAA
pore 12b	ATTAATGCAGCTGGCACGACA
pore13	CCACGCTCCCTGAGGGGCGCCAGGGTGGGAATCGGACAAGAG
pore14	CGCCTGGGGTTTGCTTATAAATCAAAAGGTTTGGACCAACGC
pore8a-rtag	GGCTATTCTTTTGATTTATAATATCCATAACCCATA
rtag-pore9b	TATTTTTTCGCTTTAGCAAACCAGCGTGGACCGCTT
rtag-pore11b	TGAGTGAATGTAATACCACCCTGGCGCCCAATACGC
pore12a-rtag	TCCCCGCGCGTTGGCCGATTCTACAAACAACAAACA
staple11[1,2,3]*T5 (tag11)	CCTAATTTACGCTAACGAGCGTCTTTTTTTATTACATTCACTCA
(tag12)*T5 staple11[4]+13[1,2,3]	TGTTTGTTGTTTGTATTTTTAATCAATAATCGGCTGCGAGCATGTAGAAACC
stable13[4]+15_tpore	TATCATATGCGTTATAGAAAAAGCCTGTTTAGAAGGCCGG
staple119[1,2,3]_tpore	AGACAGTCATTCAAAAGGGTGAGA
(tag 8)*T5 stable122[2,3,4]	TATGGGTTATGGATATTTTTGGGGCGCGAGCTGAAATAATGTGT
staple123[2,3,4]_tpore	GCCCGAAAGACTTCAAAACACTAT
stable124+122[1]*T5_tag(9)	AAGAGGAACGAGCTTCAAAGCGAAGATACATTTCGCAAATTTTTTTTAAAGC GAAAAAATA
8 staples around nanopore, delete staple13 and staple121	
staple7	CATAACCCGAGGCATAGTAAGAGCTTTTTAAG
staple9	AAAAGTAATATCTTACCGAAGCCCTTCCAGAG
staple12	TCTTACCAGCCAGTTACAAAATAAATGAAATA
staple14	CTAATTTATCTTTCCTTATCATTCATCCTGAA
staple16	GCTCATTTTCGCATTAAATTTTTGAGCTTAGA
staple17	AATTACTACAAATTCTTACCAGTAATCCCATC
staple18	TTAAGACGTTGAAAACATAGCGATAACAGTAC
staple120	AGGTAAAGAAATCACCATCAATATAATATTTT

Table S4. *The annealing component for corkT1 structure.*

Component	Amount
M13mp18 virus	10μL 100nM
Mixture of 179 staples S1–S15 (no cholesterol binding staples)	9μL 558nM
arrangement of 21 cholesterol binding sites	1μL 4.7μM
25 cork pore strands and around 8 staples for corkT1	1.7μL 3μM
MilliQ water	78.3μL
2×TMO buffer [40mM Tris-acetate,1mM EDTA,20mM MgAc, pH 7.4]	100μL
Total Volume	200μL

Note:

The list of S1–S15 for corkT1 (179 strands (100 μM) for DNA rectangle were mixed with equimolar, termed as MR_558nM)	original sequence name from **Table S1**.
S1	1–6,8,10,19–22
S2	23–30,32–34,36
S3	37–38,40–42,44–46,48–51
S4	52–55,57–59,61–63,65–66
S5	67,69–71,73–75,77–81
S6	82–93
S7	94–105
S8	106–117
S9	118,125–135
S10	136,138–140,142–144,146–148,150–151
S11	152,154–160,162–164,166
S12	167–168,170–172,174–176,178–181
S13	182–193
S14	194–205
S15	206–216
The list of arrangement of 21 cholesteryl binding sites for corkT1	177 153 31 56 76 173 149 35 60 169 145 39 64 165 141 43 68 161 137 47 72
The list of 33 nanopore sequences for corkT1	‘cork’ scaffolds 7,8a-rtag,8b,9a,9b-rtag,10-loop10,11a,11b-rtag,12a-rtag,12b ‘cork’ staples 1,2,3,4,5,6,13,14
	linkers 11(1,2,3)-ctag11, ctag12-13(1,2,3), 13(4)15(1,2,3,4), 119(1,2,3), ctag8-122(2,3,4), 123(2,3,4), 124(1,2,3,4)-ctag9
	around 8 staples: staple7,9,12,14,16,17,18,120

B.4 Detailed sequences for longcorkT1

Here are the sequences for 'longcork' and T1 design in longcorkT1 origami. The staples that help form the isolated T1 platform are the same as corkT1, and will not be repeated here. And other modified strands will be illustrated in the following tables.

Table S5. *The detailed sequences for longcorkT1 structure.*

Name	Sequences (5'-3')
SCAF0[9]0[90]	TACTATCCCCGAATTTCATGAGTCTCTGTTATGAGCCACTCCTCCAGGGGGTTGACC TTGTAACGCGTCACCGTATCGTCGT
SCAF1[90]1[9]	CATAGCTGCCTAACTAATGCATAGAAGTGGAATGTTTAACCGCTCAATCAGCCCCA AGCAGAAAACAGGTTTTGGGTCAGTG
SCAF2[2]2[48]	TTTT ACGTGACTGTGGTATGCCGTAGGGCCCGCGGTCGTATCCTAATCTAG
SCAF2[49]2[90]	ACACGGCGATAATGTGTCCGTCCTAATACCCAGGGTGAATCC
SCAF3[90]3[2]	ACAACCCGTGGCGATGGACGGACAAGCCGTACACTTTGTATAGCTTATCAGATCAA CATCAGTCTGATAAGCTAACCCCTTCATAGTCGGACAGTTTAATGATTTCGGGCAGC GACAGCGAGC
SCAF4[5]4[50]	TTTTT TTATACCAATGGTGATGTTCTTCCGCGCTCGATACGTAGCCCTGGT
SCAF4[51]4[90]	TGACGGCAACGCTATGCCGTAGCGAAACACCTGCACCGGC
SCAF5[48]5[5]	GGGGCTGCATCTGGACAACGCAGCTAAATATACGGAGCATGAGA
SCAF5[90]5[49]	TTTT CTTTATTGCTCTTCGGTCTCGCCTACGGCCTGAATATTATGG
ST3[10]3[9]	TCGCTGCCCGAATCACCATTGGTATAATCTCATGCTCATAGTACACTGACCATACCA CAGTCACGTGCTCGCTG
ST5[28]5[27]	GCGTTGTCCAATCGAGCGCGGAAGAACAATCATTAAACTGACCGCGGGCCCTACGG CCAAAACCTGTTTTCCTCATGAAATTCGGGGCGTATATTTAGCT
ST1[42]4[38]	GATTGAGCTAGATTAGGATACGTCCGACTATGAAGGGGTACCAGGGCTACGT
ST5[38]1[41]	GATGCAGGCTCATAACAGAGATGCTTGGGGCT
ST5[83]5[82]	CAATAAAGGCCGGTGCAGGTGTTTCGTCCATCGCCACGGGTCACCCTGGGTATTAG GTGCATTAGTTAGGCTACGGTGACCCGAAGAG
ST0[90]3[90]	ACGACGAAGCTATGGGATTTGT TTTT
ST5[49]1[61]	CCATAATATTTGCCGTCATACAAAGTGTACGGCACACATTATCGCCGTGTCGGTTAA ACATTC
ST7[76]5[48]	TATCCATAACCCATAAACCCCCTGGAGGAGTGCCCC
ST5[65]8[55]	GTAGGCGAGAGCGTTACAAGGTCCACTTCTAACGGTGAGTGAATGTAATA
ST8[26]7[91]	TACAAACAACAAACATTGTCCGCTACGGCATAGCGTCAGGCCTATTTTTTCGCTTTA
staple11[1,2,3]-ctag0	CCTAATTTACGCTAACGAGCGTCTTTTTTTATGGGTTATGGATA
ctag5-staple123[1,2,3,4]	TAAAGCGAAAAAATATTTTTTTTTAATTGCCCGAAAGACTTCAAAACACTAT
ctag2-staple15[1,2,3,4]	TATTACATTCACTCATTTTTGCGTTATAGAAAAAGCCTGTTTAGAAGGCCGG
staple119[1,2,3]-ctag3	AGACAGTCATTCAAAAGGGTGAGATTTTTTGTTTGTTGTTTGTA
Ten staples around nanopore, delete staple13 and staple121	
staple7	CATAACCCGAGGCATA GTAAGAGC TTTTTAAG
staple9	AAAAGTAATATCTTAC CGAAGCCCTTCCAGAG
staple12	TCTTACCAGCCAGTTA CAAAATAAATGAAATA
staple14	CTAATTTATCTTTCCT TATCATTCATCCTGAA
staple16	GCTCATTTTCGCATTA AATTTTTG AGCTTAGA
staple17	AATTACTACAAATTCT TACCAGTAATCCCATC
staple18	TTAAGACGTTGAAAAC ATAGCGATAACAGTAC
staple120	AGGTAAAGAAATCACC ATCAATATAATATTTT
staple122	TCGCAAATGGGGCGCGAGCTGAAATAATGTGT
staple124	AAGAGGAACGAGCTTCAAAGCGAAGATACATT

Table S6. *The annealing component for longcorkT1 structure.*

Component	Amount
M13mp18 virus	10μL 100nM
Mixture of 179 staples S1-S15 (no cholesterol binding staples)	9μL 558nM
arrangement of 21 cholesterol binding sites	1μL 4.7μM
23 longcork pore and around 10 staples for longcorkT1	1.7μL 3μM
MilliQ water	78.3μL
2×TMO buffer [40mM Tris-acetate,1mM EDTA,20mM MgAc, pH 7.4]	100μL
total Volume	200μL

Note:

The list of S1–S15 for longcorkT1	original sequence name from **Table S1**.
S1	1–6,8,10,19–22
S2	23–30,32–34,36
S3	37–38,40–42,44–46,48–51
S4	52–55,57–59,61–63,65–66
S5	67,69–71,73–75,77–81
S6	82–93
S7	94–105
S8	106–117
S9	118,125–135
S10	136,138–140,142–144,146–148,150–151
S11	152,154–160,162–164,166
S12	167–168,170–172,174–176,178–181
S13	182–193
S14	194–205
S15	206–216
The list of arrangement of 21 cholesteryl binding sites for longcorkT1	177 153 31 56 76 173 149 35 60 169 145 39 64 165 141 43 68 161 137 47 72
The list of 33 nanopore sequences for longcorkT1	‘longcork’ scaffold 0,1,2a,2b,3,4a,4b,5a,5b ‘longcork’ staples t10,t16,t12,t11,t17,t15,t19,t18,t13,t14
	linkers 11-ctag0,ctag2-15,119-ctag3,ctag5-123
	around 10 staples staple7,9,12,14,16,17,18,120,122,124, delete 13,121

B.5 Detailed sequences for cookieT1

Here are the sequences for ‘cookie’ and T1 design in cookieT1 origami. The staples that help form T1 platform have some modifications, as shown in **Figure S2**. The blue labelled staples are colored in blue illustrated in the following tables.

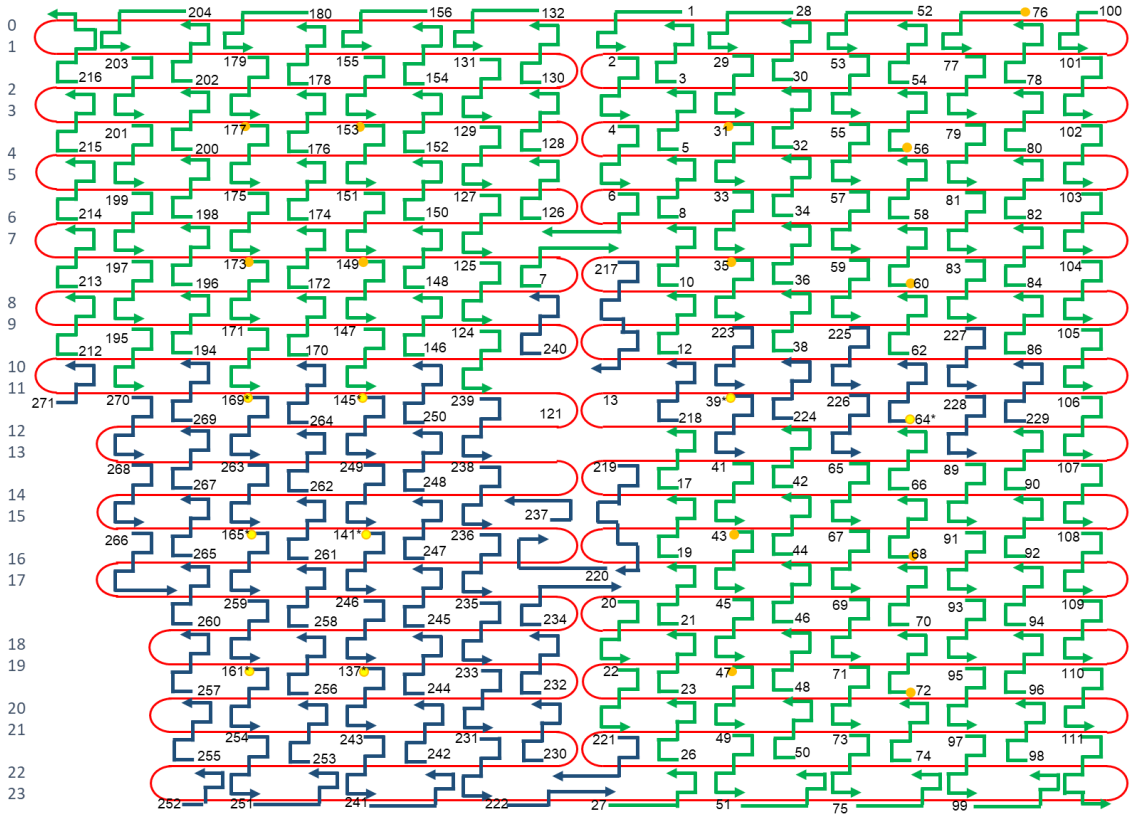


Figure S2. *Schematic of the rectangular DNA origami of cookieT1 structure.*
This is a representation of a simple Origami structure. In this figure, the continuous red strand represents the circular M13 viral genome. The previously designed auxiliary strands are all shown in green. The modified probe sequence is shown in dark blue. The arrow points to the end of the sequence. 3' end. Yellow circles indicate potential sites for binding cholesterol.

Table S7. *The detailed sequences of 127 staples without cholesteryl binding sites for cookieT1 platform.*

Mixture of every 12 tubes	Name	Sequence
E1	staple1	CAAGCCCAATAGGAACCCATGTACAAACAGTT
	staple2	AATGCCCCGTAACAGTGCCCGTATCTCCCTCA
	staple3	TGCCTTGACTGCCTATTTCGGAACAGGGATAG
	staple4	GAGCCGCCCCACCACCGGAACCGCGACGGAAA
	staple5	AACCAGAGACCCTCAGAACCGCCAGGGGTCAG
	staple6	TTATTCATAGGGAAGGTAAATATTCATTCAGT
	staple7	CATAACCCGAGGCATAGTAAGAGCTTTTTAAG
	staple8	ATTGAGGGTAAAGGTGAATTATCAATCACCGG
	staple217	AAAAGTAATATCTTACCGAAGCCCTTCCAGAGCCTAATTTACGCTAACGAGCGTCT
	staple10	GCAATAGCGCAGATAGCCGAACAATTCAACCG
	staple12	TCTTACCAGCCAGTTACAAAATAAATGAAATA
	staple218	CTAATTTAAATCAATAATCGGCTGATCCTGAA
E2	staple219	GCGTTATAGAAAAAGCCTGTTTAGAGCTTAGATTAAGACGTTGAAAAC
	staple220	ATAGCGATTAGCTGTTTCCTGTGTTTCGTAAT
	staple17	AATTACTACAAATTCTTACCAGTAATCCCATC
	staple19	TAGAATCCCTGAGAAGAGTCAATAGGAATCAT
	staple20	CTTTTACACAGATGAATATACAGTAAACAATT
	staple21	TTTAACGTTCGGGAGAAACAATAATTTTCCCT
	staple22	CGACAACTAAGTATTAGACTTTACAATACCGA
	staple23	GGATTTAGCGTATTAAATCCTTTGTTTTCAGG
	staple221	ACGAACCAAAACATCGCCATTAAATATGGTTG
	staple222	ATCACTTGCCTGAGTAGAAGAACTCAAACTAT
	staple26	TAGCCCTACCAGCAGAAGATAAAAACATTTGA
	staple27	CGGCCTTGCTGGTAATATCCAGAACGAACTGA
E3	staple28	CTCAGAGCCACCACCCTCATTTTCCTATTATT
	staple29	CTGAAACAGGTAATAAGTTTTAACCCCTCAGA
	staple30	AGTGTACTTGAAAGTATTAAGAGGCCGCCACC
	staple32	GTTTGCCACCTCAGAGCCGCCACCGATACAGG
	staple33	GACTTGAGAGACAAAAGGGCGACAAGTTACCA
	staple34	AGCGCCAACCATTTGGGAATTAGATTATTAGC
	staple36	GCCCAATACCGAGGAAACGCAATAGGTTTACC
	staple223	ATTATTTAACCCAGCTACAATTTTTCTTTCCT
	staple38	TATTTTGCTCCCAATCCAAATAAGTGAGTTAA
	staple224	TAAGTCCTCAAGAACGGGTATTAATTAGTTGC
	staple41	ACGCTCAAAATAAGAATAAACACCGTGAATTT
	staple42	AGGCGTTACAGTAGGGCTTAATTGACAATAGA
E4	staple44	CTGTAAATCATAGGTCTGAGAGACGATAAATA
	staple45	CCTGATTGAAAGAAATTGCGTAGACCCGAACG
	staple46	ACAGAAATCTTTGAATACCAAGTTCCTTGCTT
	staple48	AGATTAGATTTAAAAGTTTGAGTACACGTAAA
	staple49	AGGCGGTCATTAGTCTTTAATGCGCAATATTA
	staple50	GAATGGCTAGTATTAACACCGCCTCAACTAAT
	staple51	CCGCCAGCCATTGCAACAGGAAAAATATTTTT
	staple52	CCCTCAGAACCGCCACCCTCAGAACTGAGACT
	staple53	CCTCAAGAATACATGGCTTTTGATAGAACCAC
	staple54	TAAGCGTCGAAGGATTAGGATTAGTACCGCCA
	staple55	CACCAGAGTTCGGTCATAGCCCCCGCCAGCAA
	staple57	AATCACCAAATAGAAAATTCATATATAACGGA
E5	staple58	TCACAATCGTAGCACCATTACCATCGTTTTCA
	staple59	ATACCCAAGATAACCCACAAGAATAAACGATT
	staple225	TTTTGTTTAAGCCTTAAATCAAGAACCAAGTA
	staple62	AGGTTTTGAACGTCAAAAATGAAAGCGCTAAT
	staple226	CCGCACTCACGCGCCTGTTTATCAAGAATCGC
	staple65	CATATTTAGAAATACCGACCGTGTTACCTTTT
	staple66	AATGGTTTACAACGCCAACATGTAGTTCAGCT
	staple67	TAACCTCCATATGTGAGTGAATAAACAAAATC
	staple69	GCGCAGAGATATCAAAATTATTTGACATTATC
	staple70	AACCTACCGCGAATTATTCATTTCCAGTACAT
	staple71	ATTTTGCGTCTTTAGGAGCACTAAGCAACAGT

	staple73	GCCACGCTATACGTGGCACAGACAACGCTCAT
E6	staple74	GCGTAAGAGAGAGCCAGCAGCAAAAAGGTTAT
	staple75	GGAAATACCTACATTTTGACGCTCACCTGAAA
	staple77	TGCTCAGTCAGTCTCTGAATTTACCAGGAGGT
	staple78	GGAAAGCGACCAGGCGGATAAGTGAATAGGTG
	staple79	TGAGGCAGGCGTCAGACTGTAGCGTAGCAAGG
	staple80	TGCCTTTAGTCAGACGATTGGCCTGCCAGAAT
	staple81	CCGGAAACACACCACGGAATAAGTAAGACTCC
	staple82	ACGCAAAGGTCACCAATGAAACCAATCAAGTT
	staple83	TTATTACGGTCAGAGGGTAATTGAATAGCAGC
	staple84	TGAACAAACAGTATGTTAGCAAACTAAAAGAA
	staple227	CTTTACAGTTAGCGAACCTCCCGACCGTTTTT
	staple86	GAGGCGTTAGAGAATAACATAAAAGAACACCC
E7	staple228	ATTTTCATCGACAATAAACAACATATTTAGGC
	staple229	CCAGACGACGTAGGAATCGCAAGCAAGAACGC
	staple89	AGAGGCATAATTTCATCTTCTGACTATAACTA
	staple90	TTTTAGTTTTTCGAGCCAGTAATAAATTCTGT
	staple91	TATGTAAACCTTTTTTAATGGAAAAATTACCT
	staple92	TTGAATTATGCTGATGCAAATCCACAAATATA
	staple93	GAGCAAAAACTTCTGAATAATGGAAGAAGGAG
	staple94	TGGATTATGAAGATGATGAAACAAAATTTCAT
	staple95	CGGAATTATTGAAAGGAATTGAGGTGAAAAAT
	staple96	ATCAACAGTCATCATATTCCTGATTGATTGTT
	staple97	CTAAAGCAAGATAGAACCCTTCTGAATCGTCT
	staple98	GCCAACAGTCACCTTGCTGAACCTGTTGGCAA
E8	staple99	GAAATGGATTATTTACATTGGCAGACATTCTG
	staple100	TATAAGTATAGCCCGGCCGTCGAG
	staple101	AGGGTTGAATAAATCCTCATTAAATGATATTC
	staple102	ACAAACAAAATCAGTAGCGACAGATCGATAGC
	staple103	AGCACCGTTAAAGGTGGCAACATAGTAGAAAA
	staple104	TACATACAGACGGGAGAATTAACTACAGGGAA
	staple105	GCGCATTAGCTTATCCGGTATTCTAAATCAGA
	staple106	TATAGAAGCGACAAAAGGTAAAGTAGAGAATA
	staple107	TAAAGTACCGCGAGAAAACTTTTTATCGCAAG
	staple108	ACAAAGAAATTAATTACATTTAACACATCAAG
	staple109	AAAACAAATTCATCAATATAATCCTATCAGAT
	staple110	GATGGCAAAATCAATATCTGGTCACAAATATC
E9	staple111	AAACCCTCACCAGTAATAAAAGGGATTCACCAGTCACACG
	staple230	CTTTGACGGCTACAGGGCGCGTACCGTGAACC
	staple231	AATGCGCCAGCACGTATAACGTGCGTAATAAC
	staple232	ATCACCCAGGGCGATGGCCCACTACTTCACCG
	staple233	GTCTATCAAATCAAGTTTTTTGGGCCGCGCTT
	staple234	CCTGGCCCGCAACAGCTGATTGCCAACAGTAC
	staple235	TGAGACGGTGAGAGAGTTGCAGCAGAAAAACC
	staple236	AGCTCGAAGAAATTGTTATCCGCTTTCACCAG
	staple237	CATGGTCAACCGCTTCTGGTGC
	staple238	GCACTCCACGGAAACCAGGCAAAGGGGTACCG
	staple239	TCGCAAATTTTTAACCAATAGGAAGGAAGATC
	staple124	AAGAGGAACGAGCTTCAAAGCGAAGATACATT
E10	staple240	TTTTAATTGCCCGAAAGACTTCAAAACACTAT
	staple125	GGAATTACTCGTTTACCAGACGACAAAAGATT
	staple126	GAATAAGGACGTAACAAAGCTGCTCTAAAACA
	staple127	CCAAATCACTTGCCCTGACGAGAACGCCAAAA
	staple128	CTCATCTTGAGGCAAAAGAATACAGTGAATTT
	staple129	AAACGAAATGACCCCCAGCGATTATTCATTAC
	staple130	CTTAAACATCAGCTTGCTTTCGAGCGTAACAC
	staple131	TCGGTTTAGCTTGATACCGATAGTCCAACCTA
	staple132	TGAGTTTCGTCACCAGTACAAACTTAATTGTA
	staple241	GTTGTAGCAATACTTCTTTGATTATTTCCTCG
	staple242	TTAGAATCGTAACCACCACACCCGGTCGAGGT
	staple243	CGCTGCGCAGAGCGGGAGCTAAACAATTAACC
E11	staple244	GCCGTAAATCCAACGTCAAAGGGCAGCGGTCC
	staple245	ACGCTGGTCAGGGTGGTTTTTCTTCACAATTC
	staple246	TTGGGCGCTTGCCCCAGCAGGCGAATTAAAGA

	staple247	CACACAACACTCTAGAGGATCCCCCGCCATTC
	staple248	GCCATTCAGACAGTATCGGCCTCACGCCATCA
	staple249	GGGACGACGGCTGCGCAACTGTTGGCATGCCT
	staple250	AAAATAATTTTAGTTTGACCATTACCAGACCG
	staple146	GAAGCAAAAAAGCGGATTGCATCAGATAAAAA
	staple147	TCAGAAGCCTCCAACAGGTCAGGATCTGCGAA
	staple148	CCAAAATATAATGCAGATACATAAACACCAGA
	staple150	ACGAGTAGTGACAAGAACCGGATATACCAAGC
	staple151	AGTAATCTTAAATTGGGCTTGAGAGAATACCA
E12	staple152	GCGAAACATGCCACTACGAAGGCATGCGCCGA
	staple154	CAATGACACTCCAAAAGGAGCCTTACAACGCC
	staple155	AAAAAAGGACAACCATCGCCCACGCGGGTAAA
	staple156	TGTAGCATTCCACAGACAGCCCTCATCTCCAA
	staple251	GTAAAAGAGTCTGTCCATCACGCAAGGAGGCC
	staple252	TTTTATAATCAGTGAGCGGTACGC
	staple253	GATTAAAGGGGCGCTGGCAAGTGTCTAAAGGG
	staple254	GGGCGCTAGGATTTTAGACAGGAAGCCACCGA
	staple255	CAGAATCCGGAAGGGAAGAAAGCGGGAAAGCC
	staple256	AGCCCCCGGGAACAAGAGTCCACTAAATCCTG
	staple257	GGCGAACGGCCCGAGATAGGGTTGAAAATCCC
	staple258	TTTGATGGGCGCGGGGAGAGGCGGTAAAGTGT
E13	staple259	CGGCCAACTGGTTCCGAAATCGGCAGTGTTGT
	staple260	TTATAAATGTCGTGCCAGCTGCATAGCTAACT
	staple261	AAAGCCTGGGCCAGTGCCAAGCTTGGAAGGGC
	staple262	GATCGGTGTAACCGTGCATCTGCCTGTAGCCA
	staple263	GCGCATCGCGGGCCTCTTCGCTATGACGTTGT
	staple264	GCTTTCATACAGTTGATTCCCAATTTAGAGAG
	staple170	TACCTTTAAGGTCTTTACCCTGACAAAGAAGT
	staple171	CAAAAATCATTGCTCCTTTTGATAAGTTTCAT
	staple172	TTTGCCAGATCAGTTGAGATTTAGTGGTTTAA
	staple174	TTTCAACTATAGGCTGGCTGACCTTGTATCAT
	staple175	CCAGGCGCTTAATCATTGTGAATTACAGGTAG
	staple176	CGCCTGATGGAAGTTTCCATTAAACATAACCG
E14	staple178	ATATATTCTTTTTTCACGTTGAAAATAGTTAG
	staple179	AATAATAAGGTCGCTGAGGCTTGCAAAGACTT
	staple180	CGTAACGATCTAAAGTTTTGTCGTGAATTGCG
	staple265	CACATTAAGGGTTTTCCCAGTCACTACGCCAG
	staple266	TAACGCCATTGCGTTGCGCTCACTTCCAGTCGGGAAACCT
	staple267	CTGGCGAAGATAGGTCACGTTGGTCGAGTAAC
	staple268	CGTAATGGAGGGGGATGTGCTGCAAAGTTGGG
	staple269	AACCCGTCAAGTACGGTGTCTGGAAGAGGTCA
	staple270	TGCAACTAGGATTCTCCGTGGGAAGGATTGAC
	staple194	TTTTTGCGCAGAAAACGAGAATGAATGTTTAG
	staple195	AAACAGTTGATGGCTTAGAGCTTATTTAAATA
	staple196	ACTGGATAACGGAACAACATTATTACCTTATG
E15	staple197	ACGAACTAGCGTCCAATACTGCGGAATGCTTT
	staple198	CGATTTTAGAGGACAGATGAACGGCGCGACCT
	staple199	CTTTGAAAAGAACTGGCTCATTATTTAATAAA
	staple200	GCTCCATGAGAGGCTTTGAGGACTAGGGAGTT
	staple201	ACGGCTACTTACTTAGCCGGAACGCTGACCAA
	staple202	AAAGGCCGAAAGGAACAACTAAAGCTTTCCAG
	staple203	GAGAATAGCTTTTGCGGGATCGTCGGGTAGCA
	staple204	ACGTTAGTAAATGAATTTTCTGTAAGCGGAGT
	staple271	CTGTAGCTCAACATGTATTGCTGA
	staple212	ATATAATGCATTGAATCCCCCTCAAATCGTCA
	staple213	TAAATATTGGAAGAAAAATCTACGACCAGTCA
	staple214	GGACGTTGTCATAAGGGAACCGAAAGGCGCAG
	staple215	ACGGTCAAGACAGCATCGGAACGAACCCTCAG
	staple216	CAGCGAAAAACTTTCAACAGTTTCTGGGATTTTGCTAAAC

Table S8. *In this sheet, the modified staple sequences with DNA extensions (red) used to attach cholesterol to cookieT1 tiles are given. Chains marked in blue with an asterisk * were modified, other chains remained serial.*

blank staple + extension	Sequences of modified staple (5'-3')
staple31	GTGAAGGAGAAAAAATTAATGGCCACCACTCTTTTCATAATCAAACCGTCACC
staple35	GTGAAGGAGAAAAAATTAATGGAAGGAAAATAAGAGCAAGAAACAACAGCCAT
staple39*	GTGAAGGAGAAAAAATTAATGTATCATTCGAACAAGAAAAATAATTAAAGCCA
staple43	GTGAAGGAGAAAAAATTAATGATCAAAATCGTCGCTATTAATTAACGGATTCG
staple47	GTGAAGGAGAAAAAATTAATGTTATTAATGCCGTCAATAGATAATCAGAGGTG
staple56	GTGAAGGAGAAAAAATTAATGTCGGCATTCCGCCGCCAGCATTGACGTTCCAG
staple60	GTGAAGGAGAAAAAATTAATGATCAGAGAAAGAACTGGCATGATTTTATTTTG
staple64*	GTGAAGGAGAAAAAATTAATGAATGCAGAATCGAGAACAAGCAAGCTTGCGGG
staple68	GTGAAGGAGAAAAAATTAATGAAATCAATGGCTTAGGTTGGGTTACTAAATTT
staple72	GTGAAGGAGAAAAAATTAATGCTAAAATAGAACAAAGAAACCACCAGGGTTAG
staple76	GTGAAGGAGAAAAAATTAATGTATCACCGTACTCAGGAGGTTTAGCGGGGTTT
staple137*	GTGAAGGAGAAAAAATTAATGACGTGGACGCACTAAATCGGAACCAGCGGTCA
staple141*	GTGAAGGAGAAAAAATTAATGGCAGGTCGATACGAGCCGGAAGCATTTGCGTA
staple145*	GTGAAGGAGAAAAAATTAATGCGAGTAGATCGCGTCTGGCCTTCCAGTTTGAG
staple149	GTGAAGGAGAAAAAATTAATGCATTCAACGCGAGAGGCTTTTGCATATTATAG
staple153	GTGAAGGAGAAAAAATTAATGATACGTAAAAGTACAACGGAGATTTCATCAAG
staple161*	GTGAAGGAGAAAAAATTAATGTCCAGTTTATTTAGAGCTTGACGGAAAGGAGC
staple165*	GTGAAGGAGAAAAAATTAATGAAAACGACGGGTGCCTAATGAGTGTAATGAAT
staple169*	GTGAAGGAGAAAAAATTAATGTCCATATACAACATTAAATGTGAGGTAGATGG
staple173	GTGAAGGAGAAAAAATTAATGAAAGATTCAGGGGGTAATAGTAAACCATAAAT
staple177	GTGAAGGAGAAAAAATTAATGTTTCATGAAAATTGTGTCGAAATCTGTACAGA
pore_4f_cholesteryl also called T1-chol	CATTAATTTTTTCTCCTTCAC_cholesteryl
input strand	TAGCTTATCAGACTGATGTTGA

Table S9. *The detailed sequences of eleven staples to help form 'cookie' nanopore in cookieT1 structure.*

Name	Sequences (5'-3')
ST3[76]4[44]	AACTATATGAGAACGCGAGAAAACTTTTTCAATAGCTTATCAGATCAACATCAGTCTGATAA GCTAGAGGCGAATTATTCATTTCAATTTCATTTGAAAATCAA
ST3[65]7[38]	TTGGGTTATATTTTTCCAATACTGCGGAATCGTC
ST3[93]3[92]	GATGCAAACCTAGGTCTGAGAGACGAATTTAATACCGACCGTGTGTCTTCTGACCTAAATTT ACCAATCGCAAGACAATAAATGCT
ST1[75]0[62]	AATTTCAATAAATAAGAGAAATTGCGT
ST1[62]7[63]	ATATATTTTAGTTTTTTAGAAGCCTTTAT
ST1[54]4[65]	ATCGCGCAAGATTTTCTAGCTTAGATTAAGACGCTGAGAAGAGTCAATAGTTACCTTTTTAA CCTCCGGCTTAGG
ST4[64]3[64]	TGGAAACAGTACATATTACCTTTTTTAA
ST4[43]1[53]	TATATGTGTCCTTGAAAACATAGCGAAGGTTTAACGTCAGAATACCAAGTTACAAA
ST3[35]3[34]	TAACAATTACCTGAGCAAAAGACCTGATTGCTTTGATGAATATACAGTAACAGATTTTCCCT TAGAAAGTGAATATTACATT
ST5[10]4[10]	CGCTATTAATTATACCTTTTACATCGAAACAATAACGGATTCGAGATGATGAAACAACAAG AAAACAAAATTAAACCTTGCTTCTGTAAATC
ST0[107]5[107]	TTTTGTTTGAATCAAAATCTTTT

Table S10. *The annealing component for cookieT1 structure.*

Component	Amount
M13mp18 virus	10μL 100nM
Mixture of 182 staples E1-E15 (no cholesterol binding staples)	9μL 549nM
arrangement of 21 cholesterol binding sites	1μL 4.7μM
11 'cookie' pore for cookieT1 (dilute three times)	1.7μL 3μM
Milli-Q water	78.3μL
2×TMO buffer [40mM Tris-acetate, 1mM EDTA, 20mM MgAc, pH 7.4]	100μL
total Volume	200μL

Note:

The list of E1–E15 for cookieT1	original sequence name from **Table S7**
E1	1–6,8,217,10,12,218
E2	219,220,17,19–23,221,222,26,27
E3	28–30,32–34,36,223,38,224,41,42
E4	44–46,48–55,57
E5	58,59,225,62,226,65–67,69–71,73
E6	74,75,77–84,227,86
E7	228,229,89–98
E8	99–110
E9	111,230–239,124
E10	240,125–132,241–243
E11	244–250,146–148,150,151
E12	152,154–156,251–258
E13	259–264,170–172,174–176
E14	178–180,265–270,194–196
E15	197–204,271,212–216
The list of arrangement of 21 cholesteryl binding sites for cookieT1	177 153 31 56 76 173 149 35 60 169* 145* 39* 64* 165* 141* 43 68 161* 137* 47 72
The list of 25 'cookie' nanopore for cookieT1	'cookie' scaffold is from the virus 'cookie' staples 11 staples in **Table S9** around 7 staples 217,218,219,237,238,239,240

B.6 Detailed sequences for v7T2

Here are the sequences for 'v7' and T2 design in v7T2 origami. The modified strands will be illustrated in the following tables.

Table S11. Original sequences for T2 design (Figure 4-8).

The sequences from No.1 to No.151 are the staples without extensions for T2 assembly. Every 12 sequences of No.1 – No.150 staples will be prepared to a mixture, named T1, T2 … T12. And the following staples from No.152 to No.201 are the staples replaced with DNA extensions for cholesteryl strand binding for T2 assembly. The extensions are shown in bold and red in the table and are the complementary of T2-chol strands. Combining with the similar replaced sequences without extensions (No.202 to No.251), we can adjust the number and arrangement of cholesteryl binding sites.

Mixture of every 12 tubes	No.	Name	Sequences (5'-3')
T1	1	T2_h[0](0)h[2](16)	TGGCGAGAAAGGAAGGGAATTTTTTTTTTGGGCGCTAGGGTCACGCAAATTTTTTT
	2	T2_h[0](32)h[1](40)	ATTTAGAGCTTGACGGGGAAAGCCGGTCACGC
	3	T2_h[0](64)h[1](72)	CACTAAATCGGAACCCTAAAGGGCGCTTAAT
	4	T2_h[0](96)h[1](104)	ATCAAGTTTTTTGGGGTCGAGGTGTGGTTGCT
	5	T2_h[0](128)h[1](136)	GGCGATGGCCCACTACGTGAACCATCCTCGTT
	6	T2_h[0](160)h[1](168)	CAACGTCAAAGGGCGAAAAACCGGGCCAAC
	7	T2_h[0](192)h[1](200)	AACAAGAGTCCACTATTAAAGAACTTGGGCGC
	8	T2_h[0](224)h[1](232)	CGAGATAGGGTTGAGTGTTGTTCTGAGACGG
	9	T2_h[0](256)h[3](264)	CCCTTATAAATCAAAAGCCTGGCCCTGAGAGATGTTATCCGCTCACAAGCCAAGCT
	10	T2_h[0](288)h[1](296)	AAAATCCTGTTTGATGGTGGTTCACGCTGG
	11	T2_h[1](40)h[3](40)	TGCGCGTACAGTGAGGCCACCGAGAGTAATAA
	12	T2_h[1](72)h[3](72)	GCGCCGCTTACGCCAGAATCCTGATCAAACTA
T2	13	T2_h[1](104)h[3](104)	TTGACGAGGGCCGATTAAAGGGATAACAATAT
	14	T2_h[1](136)h[3](136)	ACGTGCCAAACCTGTGAATCAGAGAAACGCTC
	15	T2_h[1](168)h[3](168)	GCGCGGGGGTTGCGCTCACTGCCACGCCAG
	16	T2_h[1](200)h[3](200)	CAGGGTGGGCCTAATGAGTGAGCTAGGCGATT
	17	T2_h[1](232)h[3](232)	GCAACAGCGAGCCGGAAGCATAAACCAGTCAC
	18	T2_h[1](296)h[3](296)	TTTGCCCCAATCATGGTCATAGCTAGGATCCC
	19	T2_h[2](16)h[0](32)	TCTGTCCACGCTGGCAAGTGTAGCGGCGAACG
	20	T2_h[2](80)h[0](96)	GGAACGGACAGGGCGCGTACTACCGTAAAG
	21	T2_h[2](112)h[0](128)	AAACAGGACACGTATAACGTGCTTTCACCCAA
	22	T2_h[2](144)h[0](160)	AGTCGGGAGCTGCATTAATGAATCTCTATCAG
	23	T2_h[2](208)h[0](224)	CCTGGGGTTTTTTCTTTTCACCAGCAGTTTGG
	24	T2_h[2](304)h[1](319)	TTTTTTTTAATTCGTAGCAGGCTTTTTTTT
T3	25	T2_h[3](0)h[4](16)	TTTTTAACCGTTGTAGGACCTGAAAGCTTTTT
	26	T2_h[3](40)h[5](40)	CATCACTTGACATTCTGGCCAACATTTGAATG
	27	T2_h[3](72)h[5](72)	TCGGCCTTGATTCACCAGTCACACTGATAGCC
	28	T2_h[3](104)h[5](104)	TACCGCCAAATCGTCTGAAATGGACCGAACGA
	29	T2_h[3](136)h[5](136)	ATCGGTGCAGGGCGATGGAAATACGGTGAGGC
	30	T2_h[3](168)h[5](168)	CTGGCGAAATTCGCCATTCAGGCATTTTTT
	31	T2_h[3](200)h[5](200)	AAGTTGGGCGCTTCTGGTGCCGGATAATTCGC
	32	T2_h[3](232)h[5](232)	GACGTTGTGGAAGATCGCACTCCATCATCAAC
	33	T2_h[3](264)h[5](264)	TGCATGCCGTTTGAGGGGACGACGCCGTCGGA
	34	T2_h[3](296)h[5](296)	CGGGTACCTAGATGGGCGCATCGTTTGACCGT
	35	T2_h[4](112)h[2](128)	TGACGCTCGCCATTGCAACAGGAACGGGAGCT
	36	T2_h[4](144)h[2](160)	TGTTGGGAGGGCCTCTTCGCTATTCGCTTTCC
T4	37	T2_h[4](208)h[2](224)	TCCGGCACTAACGCCAGGGTTTTCGTGTAAAG
	38	T2_h[4](304)h[3](319)	TTTTTTTTGTTGGTGGAGCTCGTTTTTTTT
	39	T2_h[5](0)h[6](16)	TTTTTGTAAGAATACGCAAATCAACAGTTTTT
	40	T2_h[5](40)h[7](40)	GCTATTAGATCAAACCCTCAATCAAAATATCT
	41	T2_h[5](72)h[7](72)	CTAAAACAAATCTAAAGCATCACCTTAGAGCC
	42	T2_h[5](104)h[7](104)	ACCACCAGGTGCCACGCTGAGAGCATTTAGAA

	43	T2_h[5](168)h[7](168)	AACCAATATTTAAATTGTAAACGGCAATGC
	44	T2_h[5](200)h[7](200)	GTCTGGCCCCCCAAAAACAGGAAGCAAAAGGG
	45	T2_h[5](232)h[7](232)	ATTAAATGATCATATGTACCCCGGTCACCATC
	46	T2_h[5](264)h[7](264)	TTCTCCGTGATGAACGGTAATCGTCTGATAAA
	47	T2_h[5](296)h[7](296)	AATGGGATTTGCCTGAGAGTCTGGTTTGAGAG
	48	T2_h[6](112)h[4](128)	CTGCAACACAGAAGATAAAACAGACTACATTT
T5	49	T2_h[6](144)h[4](160)	TTGTTAAATTTGTTAAATCAGCTCTGCGCAAC
	50	T2_h[6](208)h[4](224)	CAGAAAAGTTCCTGTAGCCAGCTTGCCAGCTT
	51	T2_h[6](304)h[5](319)	TTTTTTTTCAGGTCAAGGTCACTTTTTTTT
	52	T2_h[7](0)h[8](16)	TTTTTTTGAAAGGAATAATTATCATCATTTTT
	53	T2_h[7](40)h[9](40)	TTAGGAGCTTGCGGAACAAAGAAATCATCAAT
	54	T2_h[7](72)h[9](72)	GTCAATAGTTAATTTTAAAAGTTTCTTCTGAA
	55	T2_h[7](104)h[9](104)	GTATTAGAAACTCGTATTAAATCCATATCAAA
	56	T2_h[7](200)h[9](200)	TGAGAAAGAAGCCTCAGAGCATAACTGGAAGT
	57	T2_h[7](232)h[9](232)	AATATGATGCAAGGCAAAGAATTACCAATTCT
	58	T2_h[7](264)h[9](264)	TTAATGCCTAGTAGCATTAACATCCCATTAGA
	59	T2_h[7](296)h[9](296)	ATCTACAAAGCTGAAAAGGTGGCACTGTTTAG
	60	T2_h[8](208)h[6](224)	AAGCAATAGCCGGAGACAGTCAAATTGATAAT
T6	61	T2_h[8](304)h[7](319)	TTTTTTTTGGGCGCGAGGCTATTTTTTTTT
	62	T2_h[9](0)h[10](16)	TTTTTTATTCCTGATTTTAATTGAGAATTTTT
	63	T2_h[9](40)h[11](40)	ATAATCCTACCAGTATAAAGCCAAAATTTAGG
	64	T2_h[9](72)h[11](72)	TAATGGAATGTTTAGTATCATATGAAGAGAAT
	65	T2_h[9](104)h[11](104)	ATTATTTGAACACCGGAATCATAAGTAATTCT
	66	T2_h[9](200)h[11](200)	TTCATTCCTGATAAGAGGTCATTTAAATGCTT
	67	T2_h[9](232)h[11](232)	GCGAACGACAGGATTAGAGAGTACACCATAAA
	68	T2_h[9](264)h[11](264)	TACATTTCCGAACCAGACCGGAAGACTATTAT
	69	T2_h[9](296)h[11](296)	CTATATTTTCAAATATCGCGTTTTCAAAAAGA
	70	T2_h[10](208)h[8](224)	GCTCCTTTATATAACAGTTGATTCGCAAAATT
	71	T2_h[10](304)h[9](319)	TTTTTTTTAAAGACTTCATTTGTTTTTTTT
	72	T2_h[11](0)h[12](16)	TTTTTTCGCCATATTTCTGTCTTTCCTTTTTT
T7	73	T2_h[11](40)h[13](40)	CAGAGGCATTACGAGCATGTAGAAAAGTACCG
	74	T2_h[11](72)h[13](72)	ATAAAGTACCTGAACAAGAAAAATTTTTTATT
	75	T2_h[11](104)h[13](104)	GTCCAGACGAACGCGCCTGTTTATGCCCAATA
	76	T2_h[11](200)h[13](200)	TAAACAGTAGCGAGAGGCTTTTGCACTGGCTC
	77	T2_h[11](232)h[13](232)	TCAAAAATCTCGTTTACCAGACGAAGAAAAAT
	78	T2_h[11](264)h[13](264)	AGTCAGAAGAGGCATAGTAAGAGCGGAACAAC
	79	T2_h[11](296)h[13](296)	TTAAGAGGTAATGCAGATACATAACAGTTGAG
	80	T2_h[12](208)h[10](224)	ACCAAAATTCAGAAAACGAGAATGCTTTAATT
	81	T2_h[12](304)h[11](319)	TTTTTTTTATTCAACAAGCCCGTTTTTTTT
	82	T2_h[13](0)h[14](16)	TTTTTTATCATTCCAATGCACCCAGCTTTTTT
	83	T2_h[13](40)h[15](40)	CACTCATCTTGAAGCCTTAAATCACTAACGAG
	84	T2_h[13](72)h[15](72)	TTCATCGTGTTTTAGCGAACCTCCAGTTACAA
T8	85	T2_h[13](104)h[15](104)	GCAAGCAAAGGCTTATCCGGTATTCCAATCCA
	86	T2_h[13](168)h[15](168)	GAATTACCCGTAACAAAGCTGCTAAAGAAT
	87	T2_h[13](200)h[15](200)	ATTATACCGACAAGAACCGGATATCCCAGCGA
	88	T2_h[13](232)h[15](232)	CTACGTTATAGGCTGGCTGACCTTCAACGGAG
	89	T2_h[13](264)h[15](264)	ATTATTACGGACAGATGAACGGTGTGTGTCGA
	90	T2_h[13](296)h[15](296)	ATTTAGGAATAAGGGAACCGAACTTTAGCCGG
	91	T2_h[14](208)h[12](224)	GTAATCTTAGTCAGGACGTTGGGACGATAAAA
	92	T2_h[14](304)h[13](319)	TTTTTTTTGTCAATCATACCACTTTTTTTT
	93	T2_h[15](0)h[16](16)	TTTTTACAATTTTATCAACCCACAAGATTTTT
	94	T2_h[15](40)h[17](40)	CGTCTTTCAGAGGGTAATTGAGCGAGAAACAA
	95	T2_h[15](72)h[17](72)	AATAAACAGGGAGAATTAACTGAAGAAGCCCT
	96	T2_h[15](104)h[17](104)	AATAAGAAAATAACATAAAAACAGCCGAACAA
T9	97	T2_h[15](136)h[17](136)	AACCAACCCGAAGGCTGAAAATAGACGCAATA
	98	T2_h[15](168)h[17](168)	ACACTAAACATTAAACGGGTAAAAAATCTC
	99	T2_h[15](200)h[17](200)	TTATACCAGAGGACTAAAGACTTTCTTTAATT
	100	T2_h[15](232)h[17](232)	ATTTGTATGGAACGAGGGTAGCAAGAGGTGAA
	101	T2_h[15](264)h[17](264)	AATCCGCGATCGTCACCCTCAGCATAGTTGCG
	102	T2_h[15](296)h[17](296)	AACGAGGCGCTTGCAGGGAGTTAACACGCATA
	103	T2_h[16](112)h[13](136)	ACAGAGAGACGATTTTTTGTTTAAACACCAGAACTATAGAATCA GAGAGTAGTAAATTGGGCTT
	104	T2_h[16](208)h[14](224)	GAGGCTTTAGCGCGAAACAAAGTACATCAAGA
	105	T2_h[16](304)h[15](319)	TTTTTTTTCGCTGAGGCAGACGTTTTTTTT

	106	T2_h[17](0)h[18](16)	TTTTTATTGAGTTAAGAATAAGTTTATTTTTT
	107	T2_h[17](40)h[19](40)	TGAAATAGCAACATATAAAAGAAATGGTTTAC
	108	T2_h[17](72)h[19](72)	TTTTAAGAAGCAAACGTAGAAAATATTCAACC
T10	109	T2_h[17](104)h[19](104)	AGTTACCAATGATTAAGACTCCTTTTGACGGA
	110	T2_h[17](136)h[19](136)	AGAATAATGAATTGCTAACGGAATCACCGTCA
	111	T2_h[17](168)h[19](168)	CAAAAAAAGCGGAGTGAGAATAGTTAGCGG
	112	T2_h[17](200)h[19](200)	GTATCGGTGGGATTTTGCTAAACAAAGTGCCG
	113	T2_h[17](232)h[19](232)	TTTCTTAATTTCCAGACGTTAGTACCCGGAAT
	114	T2_h[17](264)h[19](264)	CCGACAATAGTTAGCGTAACGATCGTTTAGTA
	115	T2_h[17](296)h[19](296)	ACCGATATAACGCCTGTAGCATTCACACTGAG
	116	T2_h[18](112)h[16](128)	GAACTGGCGAAGGAAACCGAGGAACAGCCTTT
	117	T2_h[18](144)h[16](160)	AACTAAAGAATTTTTTCACGTTGAATACGTAA
	118	T2_h[18](176)h[16](192)	CAGTTTCAAGGCTCCAAAAGGAGCTTCATGAG
	119	T2_h[18](208)h[16](224)	TTCTGTATTTATCAGCTTGCTTTCCGGCTACA
	120	T2_h[18](240)h[16](256)	TGTCGTCACAGCTTGATACCGAGCGAAAGA
T11	121	T2_h[18](304)h[17](319)	TTTTTTTTAAACTACATTCGGTTTTTTTTT
	122	T2_h[19](0)h[20](16)	TTTTTTTTGTCACAATAATCAAGTTTGTTTTT
	123	T2_h[19](40)h[21](40)	CAGCGCCAATCGATAGCAGCACCGTCATCGGC
	124	T2_h[19](72)h[21](72)	GATTGAGGTTAGCAAGGCCGGAAAAGCGTTTG
	125	T2_h[19](104)h[21](104)	AATTATTCGCCAGCAAAATCACCACCGGAACC
	126	T2_h[19](136)h[21](136)	CACTCCTCGGCTGAGCGACTTGAGCTCAGAGC
	127	T2_h[19](168)h[21](168)	GGTTTTGCCCTATTATTCTGAAACAGAGCC
	128	T2_h[19](200)h[21](200)	TCGAGAGGTAAACAGTTAATGCCCACCACCAC
	129	T2_h[19](232)h[21](232)	AGGTGTATCGGGGTCAGTGCCTTGGAGGTTGA
	130	T2_h[19](264)h[21](264)	CCGCCACCGATACAGGAGTGTACTATATTCAC
	131	T2_h[19](296)h[21](310)	TTTCGTCACGTTCCAGTAAGCGTCCAGAATGGAAAGCGCAGTCTCT
	132	T2_h[20](16)h[18](32)	AGCGACAGCAATAGAAAATTCATACGCAAAGA
T12	133	T2_h[20](48)h[18](64)	ATGAAACCAAGACAAAAGGGCGACACATACAT
	134	T2_h[20](80)h[18](96)	ATTACCAGAGGGAAGGTAAATAATTACGCA
	135	T2_h[20](112)h[18](128)	GAATTAGAATTAAAGGTGAATTATACCCAAAA
	136	T2_h[20](144)h[18](160)	TATTAAGAAAGAGAAGGATTAGGAAAAGGAAC
	137	T2_h[20](176)h[18](192)	TTTCGGAATCAGTACCAGGCGGATACTTTCAA
	138	T2_h[20](208)h[18](224)	TGCCCGTAGTTGATATAAGTATAGAATGAATT
	139	T2_h[20](240)h[18](256)	GTTTTAACACCGTACTCAGGAGTAAAGTTT
	140	T2_h[20](279)h[18](288)	CTTTTGATCTCAGAATGTACCGTACACAGACA
	141	T2_h[20](304)h[19](319)	TTTTTTTTAATTTACCCAGTACTTTTTTTT
	142	T2_h[21](6)h[20](32)	CTTTAGCGTCAGACTGTAGCGCGTTTTAATCAGT
	143	T2_h[21](40)h[20](64)	ATTTTCGGTCATAGCCCCCTTATTCGTCACCA
	144	T2_h[21](72)h[20](96)	CCATCTTTTCATAATCAAAATCAGTAGCACC
T13	145	T2_h[21](104)h[20](128)	AGAGCCACCACCGGAACCGCCTCCCCATTTGG
	146	T2_h[21](136)h[20](160)	CGCCACCCTCAGAACCGCCACCCTCATGAAAG
	147	T2_h[21](168)h[20](192)	ACCACCCTCAGAGCCGCCACCAGACCTGCCTA
	148	T2_h[21](200)h[20](224)	CAGAGCCGCCGCCAGCATTGACAGAGTAACAG
	149	T2_h[21](232)h[20](256)	GGCAGGTCAGACGATTGGCCTTGGGTAATAA
	150	T2_h[21](268)h[20](288)	AAATAAATCCTCATTAAAGCATACATGG

Table S12. *In this sheet, No.152–No.201 present the modified staple strings with DNA extensions (red) used to attach cholesteryl to v7T2 tiles. In addition, No.201–No.251 show the corresponding truncated sequences without DNA extensions.*

No.	Name	Sequences of modified staple (5'-3')
152	T2_h[40](76)h[0](64)	TAACAGGATTAGCAGAGCGATTTATAATACCACCACACCCGCCGAGCCCCCG
153	T2_h[32](170)h[0](192)	TAACAGGATTAGCAGAGCGATTAATTGCAGAGGCGGTTTGCGTAGTGGACTC
154	T2_h[32](220)h[0](256)	TAACAGGATTAGCAGAGCGAAACATACTGATTGCCCTTCACCGAATAGCC
155	T2_h[32](270)h[0](288)	TAACAGGATTAGCAGAGCGATGTGAAATGTTGCAGCAAGCGGTCCCGAAATCGGCAAAAT
156	T2_h[39](31)h[2](32)	TAACAGGATTAGCAGAGCGAACCCTTCTCAATACTTCTTTGATTTAAAAGAG
157	T2_h[39](74)h[2](64)	TAACAGGATTAGCAGAGCGAATAAAAGGGCCTGAGTAGAAGAACGAAGTGTT
158	T2_h[39](110)h[2](96)	TAACAGGATTAGCAGAGCGAATTGGCAGCTGGTAATATCCAGTTTAGACA
159	T2_h[33](180)h[2](192)	TAACAGGATTAGCAGAGCGAAAAGCGCCAGGGGGATGTGCTGCAAACTCACA
160	T2_h[33](218)h[2](256)	TAACAGGATTAGCAGAGCGAGGCCTCAAAAACGACGGCCAGTTTCCACAC
161	T2_h[33](268)h[2](288)	TAACAGGATTAGCAGAGCGAATCTGCCATGCAGGTCGACTCTAGGTTTCCTG
162	T2_h[38](29)h[4](32)	TAACAGGATTAGCAGAGCGATCAGTTGGTGGCACAGACAATATTGAGATAGA
163	T2_h[38](72)h[4](64)	TAACAGGATTAGCAGAGCGACCTCAAATTCTTTAATGCGCGAACGACCAGTA
164	T2_h[38](108)h[4](96)	TAACAGGATTAGCAGAGCGAAATGAAATCGCCATTAAAAATATTATTTAC
165	T2_h[34](176)h[4](192)	TAACAGGATTAGCAGAGCGAAGCAAATAGGAACGCCATCAAAAAAACCAGGC
166	T2_h[34](216)h[4](256)	TAACAGGATTAGCAGAGCGACATGTCATGAGCGAGTAACAACACAGTATC
167	T2_h[34](268)h[4](288)	TAACAGGATTAGCAGAGCGAAGAGAATCGGGAACAAACGGCGGAAACCGTGC
168	T2_h[37](27)h[6](32)	TAACAGGATTAGCAGAGCGAAGGAGCGGTGAGGAAGGTTATCTAATATCTGG
169	T2_h[37](70)h[6](64)	TAACAGGATTAGCAGAGCGATTATCATTACTAACAACTAATAGATTGCTGAA
170	T2_h[37](106)h[6](96)	TAACAGGATTAGCAGAGCGAAACGTTAATAATACATTTGAGGCAGCAGCA
171	T2_h[35](174)h[6](192)	TAACAGGATTAGCAGAGCGACGGTTGTATGTGTAGGTAAAGATTATTGTATA
172	T2_h[35](214)h[6](256)	TAACAGGATTAGCAGAGCGACATACAGATTCAACCGTTCTAGAAAACTAG
173	T2_h[34](264)h[6](288)	TAACAGGATTAGCAGAGCGAACTAATAGGGAGAGGGTAGCTATTAGCAAACA
174	T2_h[36](25)h[8](32)	TAACAGGATTAGCAGAGCGAAGTAGGGCATCAGATGATGGCAATCCACCAGA
175	T2_h[36](68)h[8](64)	TAACAGGATTAGCAGAGCGAAAATTCTTGATTGTTTGGATTATAGAGTAACA
176	T2_h[36](104)h[8](96)	TAACAGGATTAGCAGAGCGAAAAGCCGGGTTAGAACCTACCTTTGCCCG
177	T2_h[36](172)h[8](192)	TAACAGGATTAGCAGAGCGAGGCTTAGAAACTAAAGTACGGTGTAGCTAAAT
178	T2_h[36](212)h[8](256)	TAACAGGATTAGCAGAGCGAAACAGGTGTAGATTTAGTTTGACAATAAAT
179	T2_h[35](262)h[8](288)	TAACAGGATTAGCAGAGCGACTTCAAAGGCAAATGGTCAATAACTCAATTCT
180	T2_h[35](23)h[10](32)	TAACAGGATTAGCAGAGCGAATAATCGGAACAACGCCAACATGTCGCTCAAC
181	T2_h[35](66)h[10](64)	TAACAGGATTAGCAGAGCGAATCCTAATTTTTCGAGCCAGTAATCGTTATAC
182	T2_h[35](102)h[10](96)	TAACAGGATTAGCAGAGCGAGATAAGTCCGACAAAAGGTAAATTACTAGA
183	T2_h[37](170)h[10](192)	TAACAGGATTAGCAGAGCGATTTTGCCATCATTGAATCCCCCTCTTGCGGAT
184	T2_h[37](210)h[10](256)	TAACAGGATTAGCAGAGCGACATAACCCAGGTCTTTACCCTGCAAACTCC
185	T2_h[36](260)h[10](288)	TAACAGGATTAGCAGAGCGAGGAATTACGCAAAGCGGATTGCATAATTCGAG
186	T2_h[34](21)h[12](32)	TAACAGGATTAGCAGAGCGATGCTATTTGAACGGGTATTAAACCACCAATCA
187	T2_h[34](64)h[12](64)	TAACAGGATTAGCAGAGCGAGGGAGGTTGAGAACAAGCAAGCCGAATATC

		CC
188	T2_h[34](100)h[12](96)	TAACAGGATTAGCAGAGCGAGCGAGGCAGGAATCATTACCGCCAACAATA
189	T2_h[38](168)h[12](192)	TAACAGGATTAGCAGAGCGACAAATCAATTATGCGATTTTAAGAAAAAGAAG
190	T2_h[38](208)h[12](256)	TAACAGGATTAGCAGAGCGAAGGCGCAATAAAACGAACTAACAACACTAT
191	T2_h[37](258)h[12](288)	TAACAGGATTAGCAGAGCGATTGAAAGAAGGTAGAAAGATTCATCGCCAAAA
192	T2_h[33](19)h[14](32)	TAACAGGATTAGCAGAGCGAAGAGAGATCTGAATCTTACCAACGAGATTAGT
193	T2_h[33](62)h[14](64)	TAACAGGATTAGCAGAGCGAACAAAGTCCAGAGCCTAATTTGCCCGACTTGC
194	T2_h[33](98)h[14](96)	TAACAGGATTAGCAGAGCGAATTAGACGCCATATTATTTATCCTAAGAAC
195	T2_h[39](166)h[14](192)	TAACAGGATTAGCAGAGCGAGAAGTTTCACACTCATCTTTGACCTCATTACC
196	T2_h[39](206)h[14](256)	TAACAGGATTAGCAGAGCGACAGCATCCATCGCCTGATAAATTACAGACC
197	T2_h[38](256)h[14](288)	TAACAGGATTAGCAGAGCGATTTGCGGGACCTGCTCCATGTTACGACCAACT
198	T2_h[32](16)h[16](32)	TAACAGGATTAGCAGAGCGACACCACGGCCCAATAATAAGAGCACTAATATC
199	T2_h[32](60)h[16](64)	TAACAGGATTAGCAGAGCGAAAAGGTGGCAATAGCTATCTTACCCACCCTGA
200	T2_h[32](96)h[16](96)	TAACAGGATTAGCAGAGCGAGTATGTTAAAGTAAGCAGATAGGGAAGCGC
201	T2_h[39](254)h[16](288)	TAACAGGATTAGCAGAGCGAGCCCTCATGACAACAACCATCGCCAGGCCGCT
202	T2_h[2](48)h[0](64)	TTTATAATACCACCACACCCGCCGAGCCCCCG
203	T2_h[2](176)h[0](192)	TTAATTGCAGAGGCGGTTTGCGTAGTGGACTC
204	T2_h[2](240)h[0](256)	AACATACTGATTGCCCTTCACCGAATAGCC
205	T2_h[2](276)h[0](288)	TGTGAAATGTTGCAGCAAGCGGTCCCGAAATCGGCAAAAT
206	T2_h[4](16)h[2](32)	ACCCTTCTCAATACTTCTTTGATTTAAAAGAG
207	T2_h[4](32)h[2](64)	ATAAAAGGGCCTGAGTAGAAGAACGAAGTGTT
208	T2_h[4](80)h[2](96)	ATTGGCAGCTGGTAATATCCAGTTTAGACA
209	T2_h[4](176)h[2](192)	AAAGCGCCAGGGGGATGTGCTGCAAACTCACA
210	T2_h[4](240)h[2](256)	GGCCTCAAAAACGACGGCCAGTTTCCACAC
211	T2_h[4](276)h[2](288)	ATCTGCCATGCAGGTCGACTCTAGGTTTCCTG
212	T2_h[6](16)h[4](32)	TCAGTTGGTGGCACAGACAATATTGAGATAGA
213	T2_h[6](32)h[4](64)	CCTCAAATTCTTTAATGCGCGAACGACCAGTA
214	T2_h[6](80)h[4](96)	AATGAAATCGCCATTAAAAATATTATTTAC
215	T2_h[6](176)h[4](192)	AGCAAATAGGAACGCCATCAAAAAAACCAGGC
216	T2_h[6](240)h[4](256)	CATGTCATGAGCGAGTAACAACACAGTATC
217	T2_h[6](276)h[4](288)	AGAGAATCGGGAACAAACGGCGGAAACCGTGC
218	T2_h[8](16)h[6](32)	AGGAGCGGTGAGGAAGGTTATCTAATATCTGG
219	T2_h[8](32)h[6](64)	TTATCATTACTAACAACTAATAGATTGCTGAA
220	T2_h[8](80)h[6](96)	AACGTTAATAATACATTTGAGGCAGCAGCA
221	T2_h[8](176)h[6](192)	CGGTTGTATGTGTAGGTAAAGATTATTGTATA
222	T2_h[8](240)h[6](256)	CATACAGATTCAACCGTTCTAGAAAACTAG
223	T2_h[8](276)h[6](288)	ACTAATAGGGAGAGGGTAGCTATTAGCAAACA
224	T2_h[10](16)h[8](32)	AGTAGGGCATCAGATGATGGCAATCCACCAGA
225	T2_h[10](48)h[8](64)	AAATTCTTGATTGTTTGGATTATAGAGTAACA
226	T2_h[10](80)h[8](96)	AAAAGCCGGGTTAGAACCTACCTTTGCCCG
227	T2_h[10](176)h[8](192)	GGCTTAGAAACTAAAGTACGGTGTAGCTAAAT
228	T2_h[10](240)h[8](256)	AACAGGTGTAGATTTAGTTTGACAATAAAT
229	T2_h[10](276)h[8](288)	CTTCAAAGGCAAATGGTCAATAACTCAATTCT
230	T2_h[12](16)h[10](32)	ATAATCGGAACAACGCCAACATGTCGCTCAAC
231	T2_h[12](48)h[10](64)	ATCCTAATTTTTCGAGCCAGTAATCGTTATAC
232	T2_h[12](80)h[10](96)	GATAAGTCCGACAAAAGGTAAATTACTAGA
233	T2_h[12](176)h[10](192)	TTTTGCCATCATTGAATCCCCCTCTTGCGGAT
234	T2_h[12](240)h[10](256)	CATAACCCAGGTCTTTACCCTGCAAACTCC
235	T2_h[12](276)h[10](288)	GGAATTACGCAAAGCGGATTGCATAATTCGAG
236	T2_h[14](16)h[12](32)	TGCTATTTGAACGGGTATTAAACCACCAATCA
237	T2_h[14](48)h[12](64)	GGGAGGTTGAGAACAAGCAAGCCGAATATCCC
238	T2_h[14](80)h[12](96)	GCGAGGCAGGAATCATTACCGCCAACAATA
239	T2_h[14](176)h[12](192)	CAAATCAATTATGCGATTTTAAGAAAAAGAAG
240	T2_h[14](240)h[12](256)	AGGCGCAATAAAACGAACTAACAACACTAT
241	T2_h[14](276)h[12](288)	TTGAAAGAAGGTAGAAAGATTCATCGCCAAAA

242	T2_h[16](16)h[14](32)	AGAGAGATCTGAATCTTACCAACGAGATTAGT
243	T2_h[16](48)h[14](64)	ACAAAGTCCAGAGCCTAATTTGCCCGACTTGC
244	T2_h[16](80)h[14](96)	ATTAGACGCCATATTATTTATCCTAAGAAC
245	T2_h[16](176)h[14](192)	GAAGTTTCACACTCATCTTTGACCTCATTACC
246	T2_h[16](240)h[14](256)	CAGCATCCATCGCCTGATAAATTACAGACC
247	T2_h[16](276)h[14](288)	TTTGCGGGACCTGCTCCATGTTACGACCAACT
248	T2_h[18](16)h[16](32)	CACCACGGCCCAATAATAAGAGCACTAATATC
249	T2_h[18](48)h[16](64)	AAAGGTGGCAATAGCTATCTTACCCACCCTGA
250	T2_h[18](80)h[16](96)	GTATGTTAAAGTAAGCAGATAGGGAAGCGC
251	T2_h[18](276)h[16](288)	GCCCTCATGACAACAACCATCGCCAGGCCGCT

Table S13. *The detailed sequences for v7T2 nanopore structure and nearby 9 staples.*

Name	Sequences (5'-3')
ST3[76]4[44]	AACTATATGAGAACGCGAGAAAACTTTTTCAATAGCTTATCAGATCAACATCAGTC TGATAAGCTAGAGGCGAATTATTCATTTCAATTTCATTTGAAAATCAA
ST3[65]7[38]	TTGGGTTATATTTTTCCAATACTGCGGAATCGTC
ST3[93]3[92]	GATGCAAACCTAGGTCTGAGAGACGAATTTAATACCGACCGTGTGTCTTCTGACCT AAATTTACCAATCGCAAGACAATAAATGCT
ST1[75]0[62]	AATTTCAATAAATAAGAGAAATTGCGT
ST1[62]7[63]	ATATATTTTAGTTTTTTAGAAGCCTTTAT
ST1[54]4[65]	ATCGCGCAAGATTTTCTAGCTTAGATTAAGACGCTGAGAAGAGTCAATAGTTACCT TTTTAACCTCCGGCTTAGG
ST4[64]3[64]	TGGAAACAGTACATATTACCTTTTTTAA
ST4[43]1[53]	TATATGTGTCCTTGAAAACATAGCGAAGGTTTAACGTCAGAATACCAAGTTACAAA
ST3[35]3[34]	TAACAATTACCTGAGCAAAAGACCTGATTGCTTTGATGAATATACAGTAACAGATT TTCCCTTAGAAAGTGAATATTACATT
ST5[10]4[10]	CGCTATTAATTATACCTTTTACATCGAAACAATAACGGATTCGAGATGATGAAACA ACAAGAAAACAAAATTAAACCTTGCTTCTGTAAATC
ST0[107]5[107]	TTTTGTTTGAATCAAAATCTTTT
Nine staples around nanopore	
T2_h[5](136)h[7](144)_1	GTTAAATTATTCGCAGTCAGTATTAAAAATTTTTAGAACC
T2_h[7](168)h[10](176)_2	CTGAGTAACCAAAAACATTATGAATGTTTTAAATATGCGCTTAATT or CTGAGTAACCAAAAACATTATGAATGTTTTAAATATGCGCTTAATTGCTGAATATA ATGCT
T2_ [9]152[6]160_3	AGCTCAACCCCTGTAATACTTTTGCTCATATATTTTAAATTTAATATT
T2_ [10]112[6]128_4	TAAGAATACACGTAAAACAGAAATTTCAACGCAATTCGACCTTTACAAACAAGGA TAACACCGC
T2_ [11]168[13]168_5	ATAAATATGAGGGGGTAATAGTATCATTGT
T2_ [12]128[12]143_6	TGGAACTGAAACTGG
T2_ [12]112[10]128_7	CAGTTCCAATGTTCAGCTAATGCAGACGACAATAAACAACGCGTTAAA or TGTTCAGCTAATGCAGACGACAATAAACAACGCGTTAAA
T2_ [13]136[15]136_8	GAGATGGTCTTGCCCTGACGAGAACGTCAAAA
T2_ [16]144[12]160_9	TGCCACTATAAAACGAAAGAGGCACATTCAGTGAATAAGGTTAATTTCAACTTTAA AAATGTTTAGACTGGATCCAGTTT or TGCCACTATAAAACGAAAGAGGCACATTCAGTGAATAAGGTTAATTTCAACTTTAA AAATGTTTAGACTGG
T2_cholTAG also called T2-chol	CCTCGCTCTGCTAATCCTGTTA-cholesteryl
input strand	TAGCTTATCAGACTGATGTTGA

Table S14. *The annealing component for v7T2 structure.*

Component	Amount
M13mp18 virus	10μL 100nM
Mixture of 150 staples T1–T13 (no cholesterol binding staples)	7.5μL 666nM
arrangement of 50 cholesterol binding sites	1μL 4.7μM
11 'v7' pore and around 9 staples for v7T2	1μL 5μM
MilliQ water	79μL
2×TMO buffer [40mM Tris-acetate, 1mM EDTA, 20mM MgAc, pH 7.4]	100μL
total Volume	200μL

Note:

The list of T1–T13 for v7T2	original sequence number from **Table S11**.
T1	1 – 12
T2	13 – 24
T3	25 – 36
T4	37 – 48
T5	49 – 60
T6	61 – 72
T7	73 – 84
T8	85 – 96
T9	97 – 108
T10	109 – 120
T11	121 – 132
T12	133 – 144
T13	145 – 150
The list of arrangement of 50 possible cholesteryl binding sites for v7T2	152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201
The list of corresponding truncated strands without DNA extensions. to 50 possible cholesteryl binding sites for v7T2	202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
The list of 11 nanopore sequences for 'v7' and nanopore around 9 staples for v7T2	as shown in Table S13.

Appendix C Materials

C.1 Fluorescent Dye

Excitation and Emission Spectra and Extinction coefficient

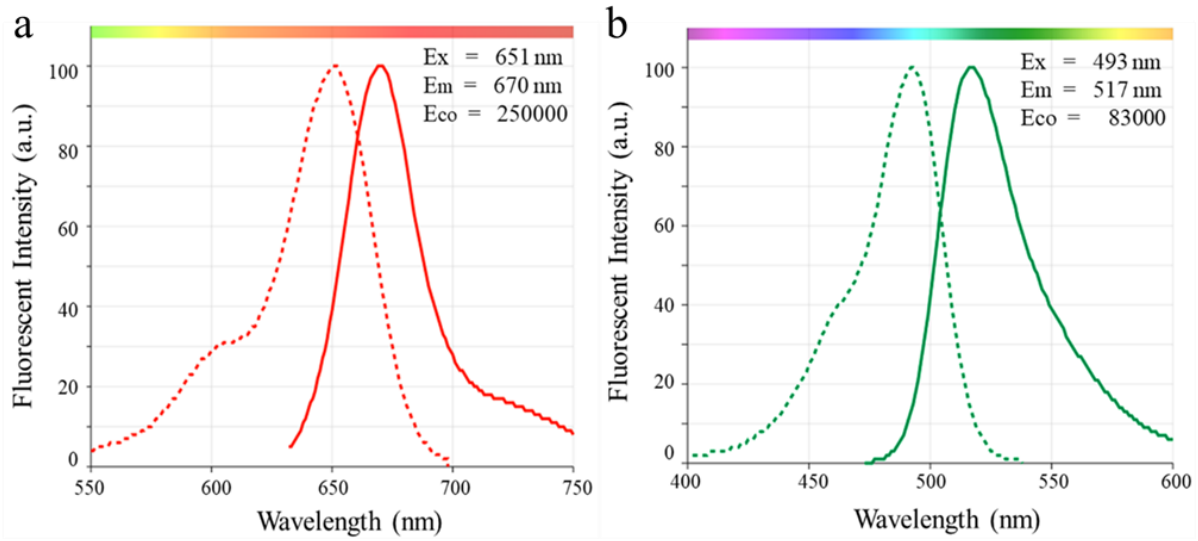


Figure S3. (a) Cy5 (Cyanine-5) excitation and emission spectra. (b) FAM (Carboxyfluorescein) excitation and emission spectra.

Cy5 can be excited using a 640 nm laser paired with a 670/30 nm bandpass filter. FAM can be excited using a 488 nm laser paired with a 530/30 nm bandpass filter. Source: www.aatbio.com

C.2 Supplementary Material Information

Table S15. *List of equipment used in this thesis.*

Equipment and software	Supplier
HEKA Elektronik	HEKA Elektronik, Dr. Schulze GmbH, Wiesenstrasse 71, D-67466 Lambrecht/Pfalz, Germany
PATCHMASTER, multi-channel data acquisition software	HEKA Elektronik
Intan Clamp System: headstages, hardware, Clamp controller, and controller software	Intan technologies
QCM-D sensor, QCM-D sensor crystals (5MHz) reactively sputter-coated with 50nm silicon oxide, plates of 11×11 mm^2 of silicon wafer	Q-sensor

Table S16. *List of reagents used in this thesis.*

Reagent	Supplier
M13mp18 ssDNA (p7249)	N4040S, New England BioLabs Inc., UK M1-11, Tilibit nanosystems, Germany
MilliQ water	Milli-Q, Millipore
Potassium chloride, KCl	Sigma, CAS 7447-40-7, MW 74.55 g/mol
$MgCl_2$	Sigma
magnesium acetate tetrahydrate, $(Mg(C_2H_3O_2)_2 \cdot 4H_2O$, MgAc	Sigma, CAS 16674-78-5, MW 214.45 g/mol
N-(2-Hydroxyethyl)pipe razine-N'-(2-ethanesulfonic acid), $C_9H_{18}N_2O_4S$, HEPES	Sigma, CAS 7365-45-9, MW 238.30 g/mol
poly(ethylene glycol) octyl ether, $CH_3(CH_2)_6CH_2(OCH_2CH_2)_nOH$, OPOE	Sigma, CAS 27252-75-1
1,2-Diphytanoyl-sn-glycero-3-PC, $C_{48}H_{96}NO_8P$, DPhPC	Avanti Polar Lipids, CAS 207131-40-6, MW 846.30 g/mol
1-Palmitoyl-2-oleoyl-sn-glycero-3-PC, $C_{42}H_{82}NO_8P$, POPC	Avanti Polar Lipids, CAS 26853-31-6, MW 760.10 g/mol
1-oleoyl-2-[12-biotinyl(aminododecanoyl)]-sn-glycero-3-phosphocholine, $C_{48}H_{89}N_4O_{10}PS$, POPC-biotin	Avanti Polar Lipids, CAS 2260669-99-4, MW 945.28 g/mol
SYBR Safe stain	S33102, Thermo Fisher Scientific
Silver wire, Ag, 0.125mm and 0.25mm	Sigma, CAS 7440-22-4, MW 107.87 g/mol
Tris base	FSBBP152-5, Fisher Scientific
Tris-HCl powder	Sigma
Boric acid	B7901, Sigma
10× concentrate Tris acetate-EDTA buffer, 10× TAE buffer	Sigma-Aldrich Inc. (1× TAE buffer (pH 8.3) with 40 mM Tris-acetate and 1 mM EDTA)
Agarose powder, $C_{24}H_{38}O_{19}$	Sigma, CAS 9012-36-6, MW 630.50 g/mol
Streptavidin from Streptomyces avidin	affinity purified, CAS 9013-20-1, MW ~60 KDa
Eau de Javel, NaClO, Hypochlorite de sodium (solutions aqueuses)	Laboratoire OXENA (France), CAS 7681-52-9, 9.6% (w/w)
Lyophilized oligonucleotides of HPLC grade	Eurogentec (Belgium)
Millipore membrane filter unit (0.22μm and 0.2μm)	Merck Millipore Ltd. (Germany)
Amicon Centrifugal Filter Units (100 kD MWCO)	Merck Millipore Ltd. (Germany)
Safeseal microcentrifuge tubes with low binding polymer technology (1.7 mL)	Sorenson Bioscience, Inc. (USA)

Note: HEPES has been described as one of the best all-purpose buffers available for biological research. At biological pH, the molecule is zwitterionic, and is effective as a buffer at pH 6.8 to 8.2 (pKa 7.55). It is typically used in cell culture at concentration between 5mM to 30 mM. After the addition of HEPES, the pH is adjusted with NaOH or HCl. HEPES has been used in a wide variety of applications, including tissue culture. It is commonly used to buffer cell culture media in air. HEPES finds its usage in *in vitro* experiments on Mg. At the beginning of our experiments, the buffer condition of 10mM Hepes, pH 6.22 is incorrect. Source from www.chemicalbook.com.