DNA nanotubes and their interaction with membranes: Insights through multiscale molecular dynamics simulations



Vishal Maingi
Department of Biochemistry
University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford.

Exeter College Michaelmas 2016

Acknowledgements

My first and wholesome big thanks to my supervisor *Prof. Mark Sansom* for the excellent supervision he rendered over me. *Dr. Mickaël Lelimousin*, the then Postdoc in our group, whose insights with exhaustive discussions helped me to improve the impact of analyses during first year. *Prof. Stefan Howorka* and *Dr. Jonathan Burns* (the experimental DNA nanotube 'tailors') for various discussions, and great insights about DNA nanotubes. *Prof. Siewert-Jan Marrink* and *Dr. Jaakko Uusitalo* for helpful discussions during my visit to their group to develop DNA nanotube coarse-grain models. *Prof. Syma Khalid* and *James Graham* for initiating working on flexible coarse-grain DNA model. I would also like to acknowledge *Prof. Andrew Turberfield* and *Dr. Philip Fowler* for their helpful suggestions during initial stages of my study.

Also, thanks to *Anna, Antreas, Chen, George, Heidi, Jean, Jemma, Maria, Matthieu, Phill S.,* and *Prafulla, Tom, Tyler* and the *SBCB* group for various generic to specific discussions over three years. Thanks to the *IT team* at Biochemistry for resolving many annoying issues.

I cannot forget the good times I shared with *Amanda, Fiona, Laura, Maria, Naushad, Teresa,* and *Tom*, and the delicious eggless vegetarian bakeries by them.

I am indebted to the *Exeter College* for their support and the funds they provided. Also, thanks to my benefactor late *Shri Krishna Pathak* from Dubai, *Clarendon* fund, *Radhakrishnan* memorial fund, and *Goodger & Schorstein* Research Scholarship. Overall big thanks to the British taxpayers, who remain unseen, for their love and support towards the research activities at Oxford.

Dedicated to my Family and Guru

&

In the memory of Rosalind Franklin

Publications

Publications mentioned in this thesis:

- Gating-like motions and wall porosity in a DNA nanopore scaffold revealed by molecular simulations; *Vishal Maingi, Mickaël Lelimousin, Stefan Howorka, and Mark S. P. Sansom.* ACS Nano 2015, *9*, 11209-11217

- Stability and dynamics of membrane spanning DNA nanopores; *Vishal Maingi, Jonathan R. Burns, Jaakko J. Uusitalo, Stefan Howorka, Siewert J. Marrink, and Mark S. P. Sansom*. Nature Communications (in press).

Abstract

DNA origami offers the possibility of developing novel membrane-spanning pores with potential applications in therapeutics and in nanosensors. In this work multiscale molecular dynamics simulation approaches have been employed to understand the dynamics of a DNA nanotube, and of its interaction with lipid bilayer membranes. All-atom simulation studies performed on the DNA nanotube model allowed exploration of its conformational dynamics, and of its interactions with ions and water molecules. Simulations under different conditions (specifically force fields and temperatures) revealed consistent properties in terms of the pore lumen shape and volume, and the gating-like motions at the mouths of the central pore. Overall the DNA nanotube model has been found to be relatively soft and porous in nature. Using the conformational information obtained from these all-atom simulations, a coarse-grained model of the DNA nanotube was developed in order to study its interactions with lipid bilayer membranes on an extended (microsecond) time scale. A number of different hydrophobic anchors which stabilize the nanotube relative to the hydrophobic core of the bilayer have been explored. Local perturbation of the membrane lipids has been observed. Energetic barriers to membrane insertion and exit for DNA nanotubes have been revealed using steered molecular dynamics approaches. A stable membrane-spanning coarse-grain DNA nanopore model was converted to all-atom resolution and used as the basis of simulation to explore the effect of high salt concentration on the stability and conformational dynamics of the pore. This confirmed that the DNA nanotube was stably embedded in the bilayer, and that ions did not form an alternative permeation pathway between the pore wall and the lipids, in contrast with other recently-reported DNA nanopore designs. Overall, these studies contribute to our understanding of the conformational dynamics and membrane interactions of DNA nanopores, thus providing guidelines to design next generation DNA nanopores rendered with controlled gating properties.

Abbreviations

Å = Angstrom
AMBER = Assisted model building with energy refinement
AT = All-atom
ca. = Circa
CG = Coarse-grain
CoM = Center of mass
DAG = Diacylglycerols
DNA = Deoxyribonucleic acid
dsDNA = double stranded Deoxyribonucleic acid
DNT = DNA nanotube
ENM = Elastic network model
Et = Ethyl
GROMACS = Groningen machine for chemical simulations
HB = Hydrogen bond
i.e.= That is
MD = Molecular dynamics
NAB = Nucleic acid builder
nm = nanometre
PA = Phosphatidic acid
PBC = Periodic boundary conditions
PC = Phosphatidylcholine
PCA = Principal component analysis
PE = Phosphatidylethanolamine
PI = Phosphatidylinositol
PME = Particle mesh Ewald
POPC = 1-Palmitoyl-2-oleoylphosphatidylcholine
PS = Phosphatidylserine
RDF = Radial distribution function
RMSD = Root mean square deviation
RMSF = Root mean square fluctuation
SM = Sphingomyelin
SMD = Steered molecular dynamics
TC-DB = Transport Classification Database
TEG = Triethylene glycol
TM = Transmembrane

Contents

1 Introduction 1
1.1) Biological membranes 1
1.2) Membrane proteins: Focusing on pores and channels of nanotechnological importance 3
1.3) DNA nanotechnology: Origins, Concepts and Advances 16
1.4) DNA nanopores 20
1.4.1) Experimental studies 21
1.4.2) Computational molecular dynamics studies 25

2 Methods 27
2.1) Classical molecular dynamics 27
2.1.1) Bonded potentials 29
2.1.2) Non-bonded potentials 31
2.2) Simulations methodology 33
2.2.1) Periodic boundary conditions 33
2.2.2) Temperature and Pressure control 34
2.2.3) Energy minimisation 35
2.2.4) Multiscale approaches 36
2.3) DNA nanotube modelling 38

3 All-atom simulations of DNA nanotube in an aqueous solution 45
3.1) Introduction 45
3.2) Methods 47
3.2.1) Model building 47
3.2.2) Simulations 49
3.3) Results and discussion 52
3.3.1) Structure and stability of the DNT models 52
3.3.2) Conformational dynamics of DNTs 59
3.3.3) Nanopore radius profiles 65
3.3.4) Ions and water in the DNT nanopore 68
3.4) Conclusions 78

4 Coarse-grained simulations of DNA nanotube interactions with lipid bilayer membranes 82
4.1) Introduction 82
4.2) Methods 83
4.2.1) Model building 83
4.2.2) Simulations 88
4.3) Results and discussion 88
4.3.1) DNA nanopores in a bilayer 88
4.3.2) Local lipid interactions 94
4.3.3) Metastable and stable states 100
4.4) Hydrophobic anchors vs hydrobphobic bands 105
4.5) Conclusions 110

5 All-atom simulations of DNA nanotube interactions with lipid bilayer membranes 111
5.1) Introduction 111
5.2) Methods 112

5.2.1) Modelling *112*
5.2.2) Simulations *113*
5.3) Results and discussion 114
5.3.1) Conformational changes under different salt conditions *114*
5.3.2) Pore characteristics and lipid interactions *117*
5.3.3) Area per lipid and bilayer thickness *124*
5.4) Conclusions 124

6 Summary, conclusions and prospects 126
6.1) Summary 126
6.2) Using improved forcefields 127
6.3) High molar concentrations 129
6.4) Final remarks 130

7 Appendix 132

8 References 135

CHAPTER 1

1 Introduction

1.1) Biological membranes

Living cells, in general, are protected from the external environment by a membranous barrier. Cells are surrounded by a lipid bilayer membrane, where each layer is an assembly of molecules whose polar hydrophilic head groups face towards the aqueous solution, Figure 1.1. The lipid tail part is hydrophobic, and forms the central core of a bilayer. The overall bilayer arrangement of lipids is driven by the hydrophobic effect allowing minimal contact of hydrophobic groups with the aqueous phase whilst keeping maximum hydrogen bonding among the freed water molecules. In terms of thermodynamics the whole process leads to an increased entropy of water molecules, hence making it a spontaneous assembly phenomenon[1,2].

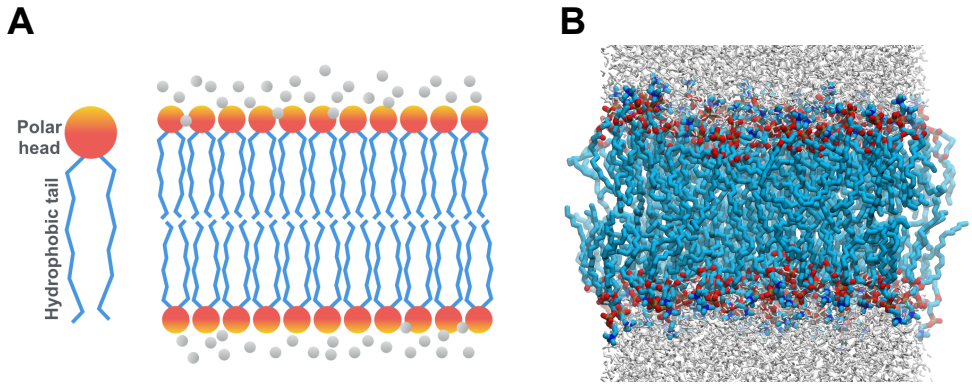

Figure 1.1: **Lipid bilayer membrane**. **(A)** Left panel shows cartoon representation of a typical lipid molecule. Head group as *orange-red* filled circle, and two tails are represented as *blue* lines. Right panel shows the assembly of lipid molecules forming a bilayer membrane with polar head groups facing towards aqueous solution; *white* filled circles. **(B)** Simulated (100 ns) all-atom 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) membrane snapshot. *Blue sticks*; aliphatic chains, *red-darkblue-gold*; head group showing oxygen, nitrogen and phosphorous atoms respectively, *white*; water molecules. Hydrogen atoms of POPC residues are not shown for clarity.

van Meer *et al*.[2,3] review membranes and their functional aspects based on their compositions. Eukaryotic membranes contain glycerophospholipids as major structural lipids which include diacylglycerols (DAG) namely

phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidylserine (PS) and phosphatidic acid (PA). Membrane properties are dependent on the chemical nature and overall geometry of the lipid molecule. For example, the cylindrical molecular shape of PC containing unsaturated (*cis* form) acyl chain makes the membrane fluid at room temperature characterised by low acyl chain order parameter and high translational diffusion coefficient values. While the conical PE molecule, due to a smaller polar head group, induces curvature in the membrane helpful in fusion processes. Another major structural lipid class consist of ceramide backbone sphingolipids called sphingomyelins (SM). Geometrically they are narrow cylinders formed by saturated or *trans*-unsaturated chains, which may introduce a gel phase into the membrane characterised by high acyl chain order parameter and low translational diffusion coefficient values. In such case cholesterol usually helps to keep the fluidic nature of the membrane.[2]

Thus it occurs that depending upon the membrane composition their function is not just limited to act as a protective barrier. They help in fission and fusion processes, protein aggregation, signalling and molecular recognition processes etc. Membranes not only provide a secluded environment for the cell but they also help in sub-compartmentalisation for the functioning of cellular organelles. Figure 1.2 shows lipid composition of typical eukaryotic membranes.

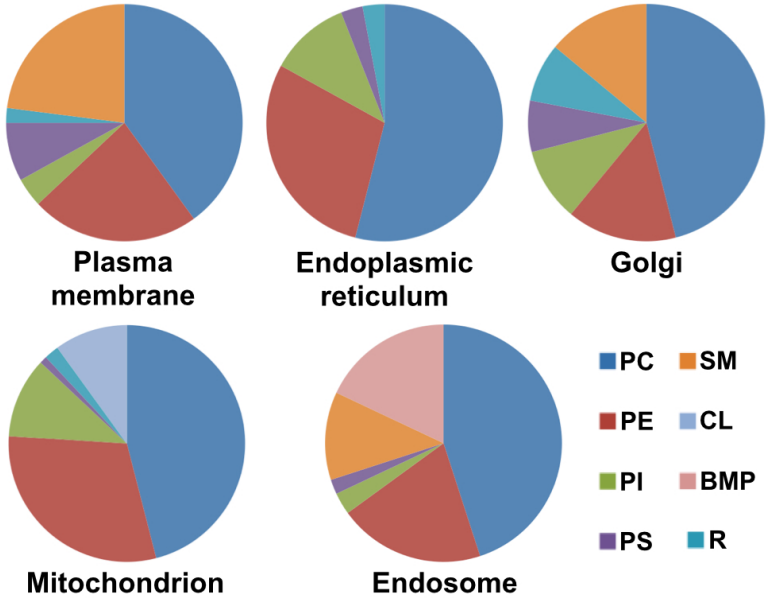

Figure 1.2: **Membrane composition**. Pie plots showing approximate composition (as percentage of total phospholipids) of membranes surrounding different organelles or cell itself. Cholesterol to phospholipids ratios in plasma membrane, endoplasmic reticulum, golgi, mitochondrion and endosomes are 1.0, 0.15, 0.2, 0.1, 0.5 respectively. Approximate composition data taken from van Meer *et al*[2]. *BMP*; bis(monoacylglycero)phosphate, *Chol*; cholesterol, *CL*; cardiololipin, *PC*; phosphatidylcholine, *PE*; phosphatidylethanolamine, *PI*; phosphatidylinositol, *PS*; phosphatidylserine, *R*; remaining lipids, *SM*; sphingomyelin.

1.2) Membrane proteins: Focusing on pores and channels of nanotechnological importance

Membrane proteins are categorised mainly in three classes[1] (Figure 1.3): a) *Integral membrane proteins* embedded in the membrane with their hydrophobic residues (amino-acids) interacting with the hydrophobic core of the membrane (aliphatic chains); *e.g.* G protein coupled receptors, glycophorins, b) *Peripheral membrane proteins* remain at the surface of membrane with favourable electrostatic or hydrogen-bonding interactions; *e.g.* phospholipase C, lipoxygenase, and c) *Lipid-anchored proteins* covalently bonded to the lipid groups with their amino-acids; *e.g.* glycosylphosphatidylinositol proteins[4] (acetylcholinesterase, alkaline phosphodiesterase). Mostly integral membrane proteins are found to be bitopic or polytopic *i.e.* spanning the membrane more than once. The peripheral membrane proteins are monotopic *i.e.* attached only at one side of the membrane, and do not span across the membrane.

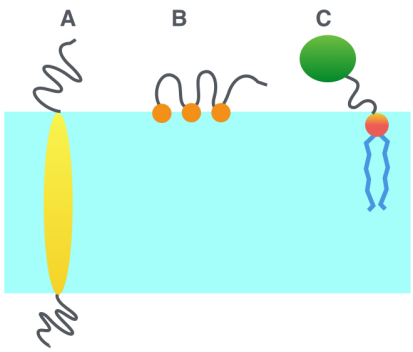

Figure 1.3: **Membrane proteins**. Cartoon representation of three main classes of membrane proteins represented as integral (A), peripheral (B), and lipid-anchored (C) proteins. Figure idea from *en.wikipedia.org/wiki/Membrane_protein.*

Here, I will focus on some pores or channels which are integral membrane proteins, and which have been of interest for nanotechnology applications. The membrane pores existing in nature basically fall under membrane proteins 'Transport Classification' system[5] also available as a database. This system mainly classifies proteins by the way they facilitate transport across the membrane as:

- *Channels or Pores*: Facilitated diffusion performed by α-helical or β-strand membrane-spanning channels (e.g. aquaporins, hemolysin, voltage gated ion channels)
- *Transporters driven by electrochemical potential*: Carrier-mediated chemiosmotic uniport, antiport or symport processes (e.g. glucose, dopamine transporters)
- *Primary active transporters*: An energy source is required for active transport against the concentration gradient (e.g. ABC transporters, Na^+/K^+ ATPase)
- *Group translocators*: Normally involves substrate chemical modification during transportation process (e.g. phosphotransferase system involving sugar phosphorylation with ATP)
- *Electron carriers*: Proteins influence the energetics of the cell involving one or two electron carrier systems (e.g. NADPH oxidases)

An exhaustive and updated list of the classification system with examples can be found online provided by the Transport Classification Database (TC-DB; *http://tcdb.org*).

I will introduce more about Pores and Channels in general, creating a platform for further discussion related to their applications in nanotechnology. The distinction between pores and channels is sometimes unclear. The International Union of Biochemistry and Molecular Biology (IUBMB; http://iubmb.org) define *Pores* in general as protein structures allowing non-selective passage of solutes across the membranes. In contrast, according to IUBMB, *Channels* allow transfer of ions or uncharged molecules while most of them exist in an open or closed state forms. The open state can be triggered by a potential-gradient or due to mechanical stress or by specific binding of a ligand molecule. Some general definitions, thus helpful for the classification, and related examples described below are obtained from TC-DB.

Pore proteins often consist of a transmembrane β-barrel. Porins and α-haemolysin are the common examples in this category (Figure 1.4). *Porin*, as a general terminology, was originally presented to describe pore-forming membrane proteins, which allows free diffusion of molecules across the membranes[6,7]. Porins are mostly found in outer-membranes of mitochondria, plastids and membrane envelopes of Gram-negative bacteria (*e.g.*, *Escherichia coli*). Common examples include OmpC and OmpF porins found in *E. coli*, which exist in the membrane as homo-trimers, and the open barrels are formed by antiparallel 16-β strands. Another similar protein OmpG is monomeric instead of homo-trimer[8]. These porins are usually required for the passage of ions and small sugar molecules across the outer membrane. Another important example α-haemolysin falls in the category of pore-forming toxins also having a characteristic β-barrel. Most of the pore-forming toxins are initially released by bacteria as water-soluble folded monomer units. Lipids/proteins, as receptors, present at the membrane surface facilitate the oligomerisation of these monomers thus forming a pre-pore state. Conformational changes, in the pre-pore state, eventually leads to the formation of membrane pore[9]. Finally, the membrane perforation usually leads to host-cell death due to leakage of the cellular contents; the process ultimately helpful for the growth of bacterial colonies[9]. Haemolysin is released as a monomer by a pathogenic *Staphylococcus aureus* which forms a cytotoxic mushroom shaped homo-heptamer pore in the membrane. The transmembrane stem is formed of antiparallel 14-stranded β-barrel.

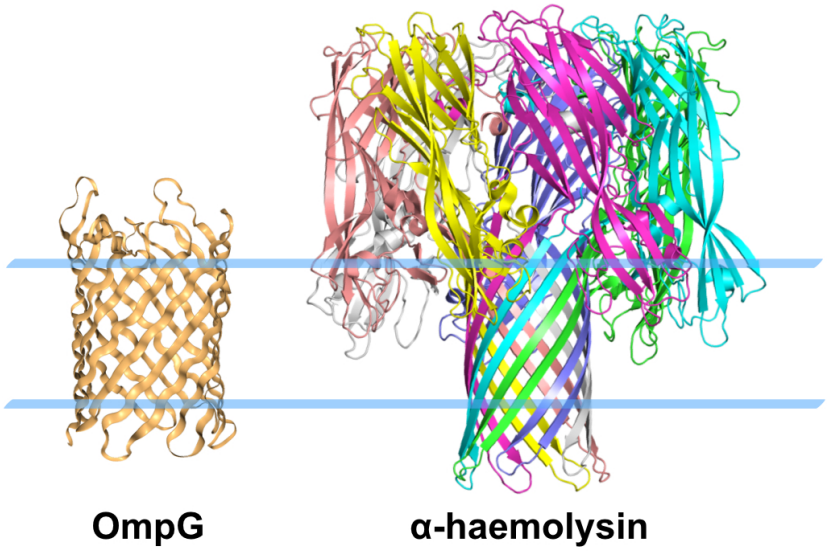

Figure 1.4: **Membrane pores**. Two common membrane β-barrel proteins. An outer membrane protein G; OmpG (from *Escherichia coli*) is shown on the left and on the right shown a heptameric α-haemolysin (from *Staphylococcus aureus*). Transparent blue lines represent approximate position of the membrane. Images taken from http://www.rcsb.org for PDB IDs: 2IWV[10] (left) and 7AHL[11] (right).

Channels are generally formed by cluster of α-helical units. Common examples include potassium channels, ligand-gated ion channels (Figure 1.5). K^+ channels play an essential role in the nervous system electrical signalling processes but have been found both in prokaryotes and eukaryotes[12,13]. Mostly they are formed by four subunits thus tetrameric in nature. There are further three main classes of these channels: voltage-gated (Kv), tandem pore domains (K2P), and inward rectifiers (Kir). KcsA found in *Streptomyces lividans*, is a tetrameric pH sensitive potassium ion selective channel. KcsA are found in the open state at low pH in contrast to, for example, Kv channels, which are activated by membrane depolarisation[14]. K^+ channels in general almost 10^4 times more selective to Na^+ ions[12]. Another class of proteins called ligand-gated ion channels control the ion flux by binding specifically to some ligands (mostly neurotransmitters, *e.g.*, serotonin, acetylcholine, glycine, and γ-amino butyric acid). The structure and mechanism of pentameric ligand-gated ion channels (pLGIC) glycine receptor (GlyR) have been recently explored using cryo-microscopy[15]. GlyR is an urn-shape channel formed by five-subunits which controls the flux of Cl^- ions. The structural, and thus the mechanism-associated, studies of these channels are quite

important to understand neurological anomalies related to hyperekplexia, autism etc.

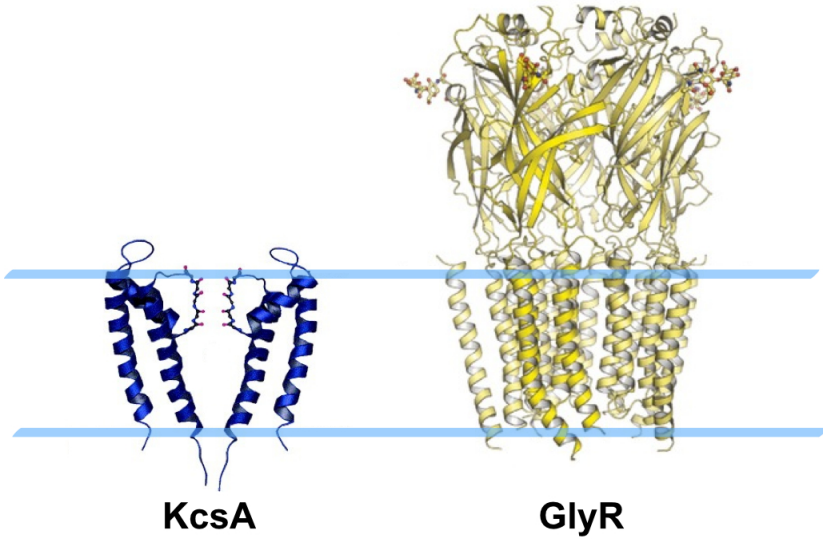

Figure 1.5: **Membrane channels**. Common examples for channels formed by α-helical units. On the left shown two units (of the four) of KcsA channel. Blue ribbons are α-helices and the selectivity filter is represented as ball and stick model. On the right shown a pentameric GlyR cryo-microscopy reconstructed model. Transparent blue lines represent approximate position of the membrane. KcsA image adapted from Sansom *et al.*[14] with permission from Elsevier. GlyR image from Du *et al.*[15], adapted by permission from Macmillan Publishers Ltd: *Nature*, copyright 2015.

Pores and channels of nanotechnological importance. Considering the nanotechnology application aspects of the protein pores/channels, the β-barrel toxins (see above TC-DB; Pore/Channel classification) have been used and engineered (mutation experiments) for various sensing applications[16] (illustrated with examples below); Figure 1.4. In general the conduction state through the channels helps in determining the concentration and the nature of any analytical molecule passing through the pore, which further depends on its binding ability with the amino-acid residues in the lumen[17,18].

In the β-barrel category α-haemolysin (a heptamer; minimal constriction diameter 1.4 nm)[19], OmpG (minimal diameter 0.8 nm in an open conformation[17] *e.g.* in PDBID: 2IWV), and MspA (an octamer; constriction diameter 1.2 nm)[19] proteins have been used quite extensively for single molecule analysis *e.g.* ADP, nucleic

acids[20]. Channels formed by α-helical units have been less explored compared to the β-barrels but few examples will be discussed in the following sections. Finally, the discussion further leads to 'journey to the DNA nanopores'.

β-barrels

Inspired by protein nanopores found naturally in the lipid membranes, it is now possible not only to study and understand the translocation processes through these nanopores but these understandings have also helped to exploit them for various nanotechnology sensing applications discussed below. In general translocation processes through the protein pores are detected based on the resistive-pulse technique (which is historically based on the Coulter counter principles)[21]. In this technique, the conducting state of a protein nanopore is altered when a molecule is partitioned (or simply enters) into the nanopore. Thus, the 'sensing' in these experiments is based on the observed current fluctuation measurements. One of the first study[21] of translocation through a protein pore was reported by Bezrukov *et al*.[22] in which current fluctuations through a single alamethicin channel were induced each time a poly(ethylene glycol) polymer (PEG) translocated through them. This helped them in counting polymers passing through the channel. PEG is a charge neutral molecule thus in their experiment it was a diffusion-driven exchange/translocation through the nanopore. But it is to be noted that even before this some studies were reported in which such polymer were partitioned in other protein pore (e.g. α-haemolysin[23]) to study and measure the pore diameters based on the conductivity of these channels in the presence of different concentrations of non-electrolytes (polyglycols).

After the success of experiments on PEG translocation, interest has been directed towards the characterisation of molecules of biological interest also[21]. Kasianowicz *et al.*[24] reported first study[21] in which polynucleotide sequences were characterised by translocation through α-haemolysin nanopores using resistive-pulse methodology. Since then α-haemolysin has been explored and studied for various sensing purposes mainly due to its stability, reduced spontaneous gating, enough large diameters to accommodate molecules like ssDNA, and commercial availability[20,21,25]. However to optimise the sensing technique, and to obtain more information from the experimental conductance data various modifications on the

nanopore itself or the molecules under question have been suggested (*e.g.* nice review by Ulrich Keyser[21]) mainly to control the rate of translocation. Here I will briefly mention different modifications performed, and purpose for the same, on the α-haemolysin nanopore itself but in no case is an inclusive list.

For sensing applications based on the ionic current fluctuation measurements one would need to slow down the translocation speed in order to differentiate between residues attached in a continuous sequence, *e.g.* nucleic acid bases sequencing. Rincon-Restrepo *et al*.[26] were able to slow down the translocation rate of ssDNA by modifying the lumen lining of α-haemolysin using positively charged amino acid (arginine) residues (Figure 1.6A). They suggested that the reduced translocation rate might be possibly due to DNA phosphodiester interactions with arginine side chains. But this site-directed mutation also reduced or suppressed the ionic current flow through the pore. Thus, they suggested such mutated pores could be used (to slow down and constant rate translocation of nucleic acids) as hybrid-nanopores and mounted on solid-nanopores or other protein nanopores (to measure currents while slowed down nucleic acids are translocating). Similarly for MspA, Butler *et al*.[18] formed two mutant variants in which the negative charge in the constriction of the pore was replaced by neutral or positive charged residues. They also compared the effect of such mutations on the translocation of nucleic acids.

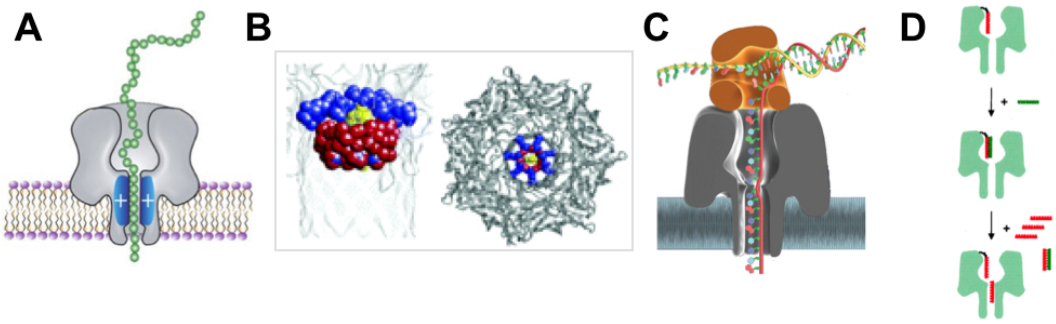

Figure 1.6: **Different functional modifications on α-haemolysin**. **(A)** α-haemolysin nanopore mutated with positive arginine residues shown in blue, and ssDNA shown green beads thread translocating through the pore. **(B)** Mouth of β-barrel in α-haemolysin modified with positive arginine residues (blue) fitted with cyclodextrin molecule (red), while one nucleoside molecule interacting at this interface is shown in yellow (left; side view showing only β-barrel, right; top view). **(C)** phi29 DNA polymerase (orange) is shown mounted on the α-haemolysin pore which controls the translocation of DNA strands. **(D)** α-haemolysin modified with oligonucleotide (red), and sequential events showing binding of complementary strand (green) and release of the same due to competitive binding when excess of red strands are added. Each such event affects the flow of current through the pore. Figure **A** adapted with permission from ref[26], copyright 2011 American Chemical Society; Figure **B** adapted with permission from ref[27], copyright 2006 American Chemical Society; Figure **C** from Schneider *et al.*[28], adapted by permission from Macmillan Publishers Ltd: *Nature Biotechnology*, copyright 2012; Figure **D** from ref[29] Copyright 2011 National Academy of Sciences.

Apart from these experiments, in addition to arginine mutations on the α-haemolysin pore Astier *et al.*[27] non-covalently incorporated an adapter molecule (type of β-cyclodextrin molecule) close to the mouth of pore; Figure 1.6B (one of the references mentioned in the paper leading to how cyclodextrin could be fitted seems to be incorrect but it appeared from the experimental setup that one can simply add such molecule in one of the chambers separated by the membrane). A combination of two positive rings, one ring formed by seven arginine residues and other ring formed by β-dextrin seven amine groups, helped to classify (type or concentration) four different individual nucleoside monophosphates based on different ionic current recordings. Interestingly, nucleoside monophosphate is a product of exonuclease catalytic activity on single stranded nucleic acid polymer

sequence, and it was proposed in the same paper that exonuclease activity can be tuned in with the nanopores. Later, this idea was indeed implemented by Clarke *et al*.[30], and covalently attached β-cylodextrin analogue of α-haemolysin to read out the nucleic acid sequence by tuning the exonulease turnover activity. But ideally they suggested that exonuclease should be in close proximity using covalent attachment or a genetic manipulation) of the pore, and the digested nucleotide products should be fed directly to the nanopore with better control over the enzyme activity. Later two research groups coupled the phi29 DNA polymerase to α-haemolysin (Cherf *et al*.[31]) or MspA (Manrao *et al*.[32]); Figure 1.6C. Main idea was to read the DNA sequence (i.e. translocation of ssDNA chain) as a single file without breaking the DNA strand and with a controlled rate of translocation. Though more details about exact differences (beyond the interest of basic introduction here) in these methodologies can be found in the corresponding references[31,32] but in brief the extension and exonulcease functions of phi29 DNA polymerase helped in the controlled translocation of DNA template[32].

Not only for the sensing applications, nanopores have been employed to study the kinetics of complex formation. Howorka *et al*.[29] studied kinetics of oligonucleotides hybridization (Figure 1.6D) by measuring ionic current fluctuations. They modified *cis* side of α-haemolysin with a covalent attachment of small oligonucleotide. When a complementary oligonucleotide strand was added the ionic current blockade was observed which indirectly described the duplex formation between the strands. In an another experiment Gurnev *et al*.[33] used maltoporin channels (sugar selective bacterial outer membrane trimer channel) as receptors for bacteriophage lambda virus particles to study the interaction kinetics of the complex formation. Nanopores have also been employed to understand protein folding and unfolding kinetics[25]. In general, the folded proteins are unable to enter small pores but if denatured small peptides can enter nanopores. Using simple models obtained by mutating α-haemolysin with different charged amino-acid residues may help better understanding the fundamental process of polypeptide transportation through biological nanopores. For example, Wolfe *et al*.[34] modified the inner lumen lining of α-haemolysin with negatively charged residues to facilitate translocation of cationic polypeptides. This suggested that such studies could be used to understand the free energy

barriers to transport polypeptides through biological nanopores. A nice review by Reiner *et al.*[25] provides some more interesting examples for sensing applications.

α helices

In contrast to robust β-barrels, the channels formed by α helices have not been widely employed[20] to a greater extent for conduction or sensing experiments[22,35,36]. Mayer *et al.*[36] used alamethicin to study the protein-ligand interactions (Figure 1.7A). They covalently attached alamethicin with benzenesulphonamide which acted as a ligand for carbonic-anhydrase. The alamethicin-sulphonamide units could assemble to form membrane channels thus allowing current flow. Addition of carbonic-anhydrase reduced the current flow possibly due to alamethicin-sulphonamide carbonic-anhydrase complex formation, which thus inhibited the channel formation/assembly by alamethicin-sulphonamide α-helices. Further addition of a competitive inhibitor (4-carboxybenzenesulphonamide) released the carbonic-anhydrase which was detected in the form of restored current flow. This study helped them to deduce the dissociation constants of protein-ligand binding.

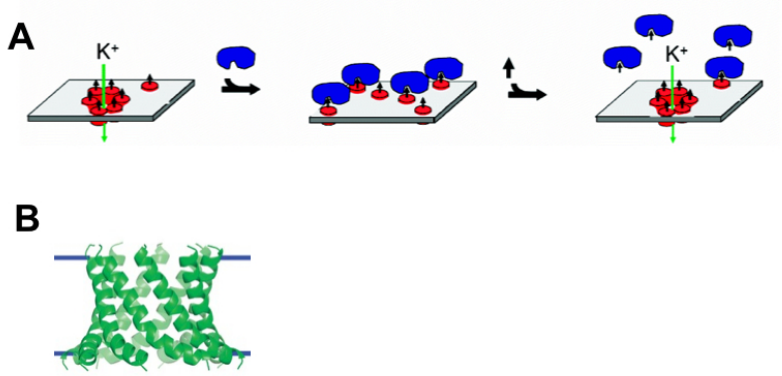

Figure 1.7: **Channels formed by α helices. (A)** Sulphonamide modified (black) alamethicin (red) units (3 to 5) assemble to form a conducting pore in a lipid bilayer (grey surface). The assembly is disrupted with the addition of carbonicanyhdrase (blue) due to specific interaction with sulphonamide moieties. In the last step addition of competitive ligand 4-carboxybenzenesulphonamide restores the conducting state. **(B)** Proposed high-conducting state model of pore formed by 8 synthetic α helical units (green). Blue lines represents membrane. Figure **A** adapted from ref[36], copyright 2008 American Chemical Society; Figure **B** from Mahendran *et al.*[37], adapted by permission from Macmillan Publishers Ltd: *Nature Chemistry*, copyright 2016.

Recent study by Mahendran *et al*.[37] demonstrated and characterised the formation of membrane channel formed, most probably, by eight synthetic α-helical units (Figure 1.7B), which may serve as a new platform for sensing technology. α-helical unit consisted of 35 peptides obtained as consensus sequence from outer membrane D4 subunit of Wza (polysaccharide transporter in gram-negative bacteria).

Journey to DNA nanopores

Another important class for sensing applications are solid-state nanopores. Though, this needs a separate dedicated discussion but I have just briefly introduced them here to fill the gap between the protein-nanopores and very recent class of pores; DNA-nanopores, which are of main interest in my thesis. This will also allude to the fact that despite advancements in the protein-nanopores for nanotechnology applications why we still need to hunt for new nanopores.

Solid-state nanopores are small holes in the artificial membranes (mostly SiN based) created using high-energy focused ion beams which causes atomic level erosions[38]. Depending upon the sensing target applications the pore surface properties can be modified[39,40]. For example, applying lipid coating on the surface with mobile ligands (biotin), Yusko *et al.*[40] controlled the translocation of proteins (streptavidin) through the pores. Still it is difficult to accomplish precision control over the geometry, and the placement of chemical modifications[41] but they can serve as stable platforms to hold other nanopores[42]. On the other hand protein nanopores have stable pore diameters (e.g. α-haemolysin) and one can mutate with different amino-acid residues or perform additional chemical modifications, as illustrated by examples mentioned above, but they require a delicate lipid membrane support. Considering these advantages and limitations of these two different nanopore (protein and solid-state) systems, Hall *et al*.[42] constructed a hybrid nanopore by placing α-haemolysin on a solid-state nanopore. α-haemolysin was hybridised to a long dsDNA thread *via* a covalently attached small ssDNA linker. With the help of DNA thread the complex could be pulled and positioned in a solid-nanopore electrophoretically (voltage driven). It was suggested that these hybrid pores could be used for many sensing applications as discussed above. Bell *et al*.[41] pointed out an important limitation to such nanopores that

protein nanopores have a fixed nanometer range diameters thus limiting wide range of analytes (currently for example *mostly* but not necessarily, as exhaustive examples are not included here, limited to unfolded peptides/proteins and ssDNA), which otherwise could be sensed using nanopores with large diameters. To overcome the later, they designed a new hybrid nanopore possessing merits of protein--solid-state hybrid nanopores. They constructed a DNA origami based nanopore (classic example of DNA nanotechnology; discussed further in Sections 1.3 and 1.4), with central tip opening dimensions of ~7.5x7.5 nm^2 which was then mounted on a solid-state nanopore. Using similar approach in hybrid protein solid-state nanopores, DNA over hangs were attached at the tip of DNA origami nanopore which assisted its mounting over the solid nanopore driven by voltage. Recently, Göpfrich *et al*.[43] inserted a large diameter DNA nanopore (minimal cross-section 6 nm) in a lipid bilayer membrane, which also happens to be the largest synthetic pore with a lipid membrane support. They also review that wide diameter protein pores remain unexplored while at the same it remains a challenge to create synthetic channels with wider diameter range. Large diameter pore can accommodate high molecular-weight analytes, and could possibly be used to understand protein-protein or protein-DNA interactions[43,44].

Apart from fixed diameter (in terms of range) of protein nanopores there are some other complications that may arise also. α-haemolysin is formed by seven subunits (homoheptamer pore), and if one would like to improve the properties of the pore by mutating more than two subunits then one can get a number of permutation combinations for the resultant assembled functional nanopore[17,45] (see Figure 2 in this reference[46]). This feature thus makes it a difficult (though not impossible) process to obtain a desired pore. OmpG, though a monomeric porin has flexible loop regions, which can cause stochastic gating in the nanopore[17,20]. This is where mutation experiments help in engineering a desired nanopore. Using computational driven experimental approaches Chen *et al.*[17] mutated the OmpG bacterial β-barrel protein to reduce the spontaneous gating process to obtain a stable single-molecule detecting sensor pore. It is rather difficult (if not impossible) to control designing myriad shapes from a significantly mutated peptide sequence[47], which may otherwise help in adding new features in the nanopore. Nonetheless, over the decades various groups have attempted to design

proteins *de novo*. In a broad sense, *de novo* protein design aims at creating new desired protein folds using sequences not related to those found in nature, though some groups have used this terminology in different ways as discussed by DeGrado *et al.* [48] The *de novo* design principles and examples are not exhaustively discussed here as it is beyond the interest of major theme of my thesis but a review by Huang *et al*.[49] explain the basic principles, approaches, protein *de novo* structures, and updates in this field. Some *de novo* protein design examples in context of channels/pores are mentioned here. Joh *et al*.[50] designed a membrane spanning four-helix bundle, which acted as an antiport protein for Zn^{+2} (in exchange with protons) as shown with the liposome flux assays. They also determined the stability of this complex using molecular dynamics simulations, and characterised the structure with X-ray studies. Though not membrane pores, Thomson *et al*.[51] designed water soluble barrels formed by α-helices using computational approaches, and also characterised them experimentally with X-ray studies. Also, recently as explained above Mahendran *et al*.[37] formed a membrane spanning α-helices octamer.

The above examples clearly demonstrate that proteins have been exploited, and engineered (few mutations) to achieve the desired features in the nanopore and without doubts the *de novo* designs are progressing, and illustrated quite well[49]. As discussed further in detail, DNA nanotechnology can also be used to design nanostructures achieving precise chemical functionalization (which is similar to introducing mutations in proteins), and additionally shape control. In terms of membrane spanning structures it is to be recalled that, with the help of exposed hydrophobic amino-acid residues, protein nanopores are embedded in the membrane. DNA presents an exposed hydrophilic surface thus not found naturally embedded in the membrane. Motivated by membrane porin proteins, since 2012 a few groups have successfully made functional DNA nanopores embedded in membranes by providing hydrophobic attachments to the DNA surface (discussed ahead).

No one has yet attempted to understand the molecular level energetics involved in DNA nanopore insertion in the membrane. Without providing molecular

understanding, reported experimental studies mention that hydrophobic groups 'help' in the insertion. I wish to use computational simulations to better understand, at a molecular level, the nature of this 'help'.

1.3) DNA nanotechnology: Origins, Concepts and Advances

Amino acid sequences, called polypeptides, fold variously to provide protein structures. Similar to the protein structural hierarchy, nucleic acids can also serve to generate primary, secondary, tertiary or quaternary structures. De-oxy ribonucleic acids (DNA), constituting bases namely adenine (A), guanine (G), cytosine (C) and thymine (T), have been used extensively to form various complex structures (nanostructures) where the secondary structure stability is provided by the Watson-Crick complementary base pairing in an aqueous environment (Figure 1.8). The foundations and recent advancements are discussed below.

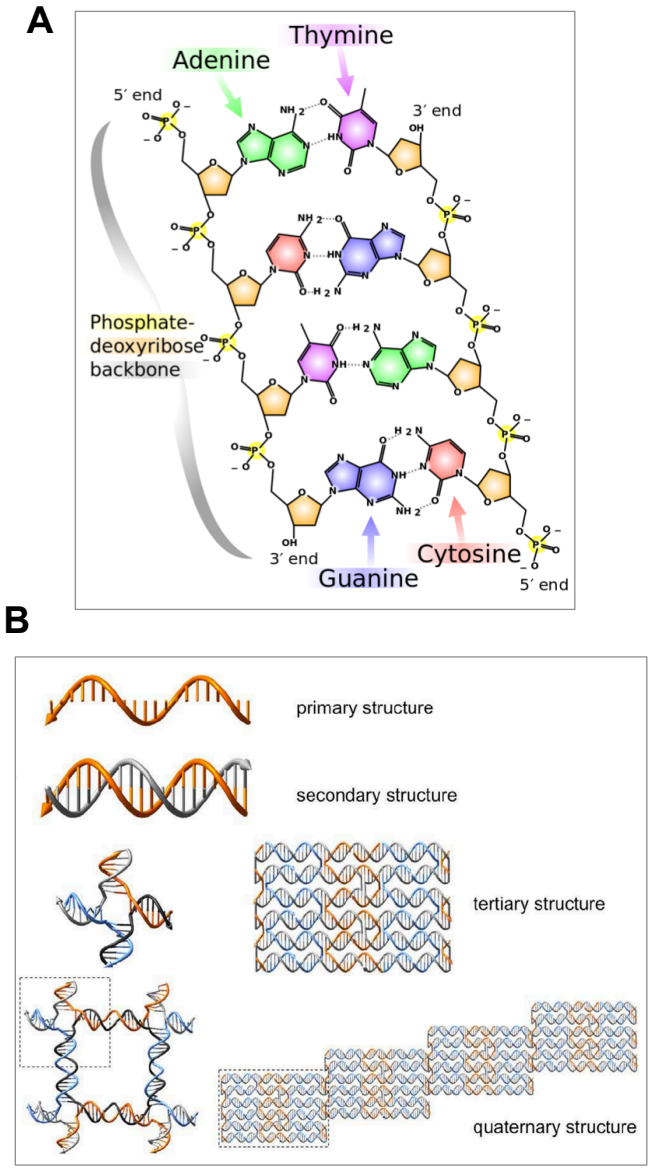

Figure 1.8: **DNA base pairing and structural hierarchy**. **(A)** A-T, G-C base pairs shown along with phosphate backbone, **(B)** Hierarchy seen in DNA from primary to quaternary structures. Figure **A** by Madprime: (downloaded August 2016) *https://commons.wikimedia.org/w/index.php?curid=1848174.* Figure **B** taken from Pfeifer *et al*[52]; with permission John Wiley and Sons.

Nadrian Seeman established the concept[53] of DNA nanotechnology. As the name suggests, DNA nanotechnology employs nucleic acids as building blocks to generate complex nanostructures. Chen and Seeman generated one of the initial three-dimensional DNA nanostructures in the form of a cube[54]. The original methodology developed, often described as the multistrand[55] approach; see Figure 1.9A, uses number of single strands (oligonucleotides), which self-assemble to generate DNA tiles and these tiles themselves further act as building blocks. The

sequence of oligonucleotides are considered in such a way that the possibility of obtaining undesired secondary structures is minimized[52]. This methodology requires stoichiometry control on the participating oligonucleotides, and the experimental procedures are prolonged as each of the building block (DNA tiles) need purification. Also, the myriad of structures that could be generated with this methodology was limited until Paul Rothemund[56] introduced the DNA-origami approach in 2006, also referred as scaffold-based self-assembly[55]. In Rothemund's methodology, see Figure 1.9B, usually a long single strand of DNA (a scaffold strand) is folded with the help of small single strands (staple strands). Depending upon the origami design the scaffold strands run parallel as well without folding[57]. Overall this approach overcomes the disadvantages in multistrand-based procedures. DNA-origami involves dense packing of dsDNA helices thus requiring a high molar concentration of salts to screen negatively charged backbones of helices running parallel to each other (Figure 1.9C). At the current time, this limits their application for therapeutic or physiological purposes.

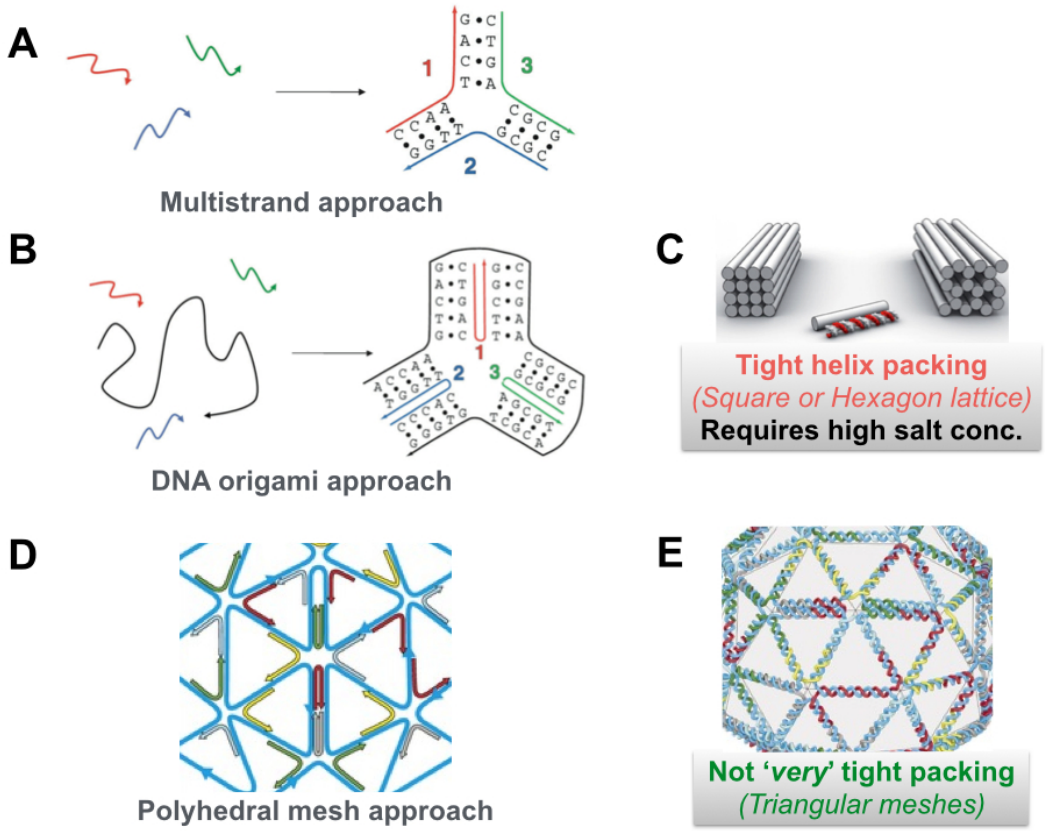

Figure 1.9: **General DNA nanostructures building approaches**. **(A)** A simplified multistrand approach shown using mixture of small strands of DNA. **(B)** A simplified DNA-origami approach using a long scaffold (black) strand folded with the help of small staple strands (red, green, blue). **(C)** A general packing (square or hexagonal lattice) arrangements shown for structures obtained from DNA origami approach **(B)**. Each cylinder represents a dsDNA. **(D)** Triangular meshes can be designed to render almost any surface where each edge of the triangle is a dsDNA. **(E)** Arrangement of DNA helices in polyhedral mesh approach **(D)**. This approach allows less dense DNA nanostructures. Figures **A**, **B** taken from Saccà *et al.*[55] with permission John Wiley and Sons (original image credits from Rothemund *et al.*[58], reprint with permission IEEE 2005). Figure **C** from Castro *et al.*[59], adapted by permission from Macmillan Publishers Ltd: *Nature Methods*, copyright 2011. Figures **D**, **E** taken from Benson *et al.*[60], adapted by permission from Macmillan Publishers Ltd: *Nature*, copyright 2015.

Almost a decade since Rothemunds's work, another important methodology reported by Benson *et al.*[60], see Figure 1.9D, allows generation of very complex shapes. Benson *et al.* described that almost any surface could be rendered in the form of triangular meshes where each edge of the mesh network is defined by a dsDNA helix (Figure 1.9E). It should be noted that the triangular arrangement of dsDNA helices reduce the inter-helix backbone interaction to a greater extent,

requiring relatively lower salt concentrations. I believe this approach is promising in terms of further development of DNA nanotechnology towards physiological and therapeutic applications.

1.4) DNA nanopores

Among the myriad structures reported by various groups, DNA nanotubes (DNT) have been of considerable interest. DNTs are hollow tubes formed using nucleic acids (*e.g.* see Figure 1.10). The first DNT was formed in 1998[61] but remained un-noticed for almost half a decade, and were considered as flat structures. Rothemund *et al*.[62] pointed out this fact in 2004, and characterized DNA nanotubes through experiments. They re-analysed (details not provided in the paper) the AFM images from the previous study by Winfree *et al*.[61], and suggested that 2-D DNA arrays could form tube like structures as well. This motivated them to properly characterise them, and with the help of simple modelling and design principles they studied the important features and modifications required in the DNA tiles that could help (or disrupt) the formation of tube like structures. Consistent characteristics of tube, *e.g.* length and circumference, were analysed using AFM and fluorescence techniques. Interestingly, they have also reported temporal AFM images that shows due to successive scans of the same area (where tube was placed) with AFM tip resulted in opening of the tubes in to sheets which helped them to reveal more structural details of the opened tube (more details can be found in the same paper[62]). It should be noted that tube-like DNA structures were reported earlier by Yan *et al*.[63] but the reasons for their formation and the geometrical description remained less understood, and Liu *et al*.[64] in 2003 used disulphide bonds between the tiles to form the tube like structures. So study by Rothemund *et al*.[62] can be considered as the first report of fully understood and characterised DNA 'only' tube like structures. Since then many potential applications of DNT have been suggested, or speculated, by manipulating the design or nucleic acid modifications. DNTs, once suitably chemically derivatized to aid insertion into lipid bilayers, have been shown to function as nanopores in both artificial and cell membranes[57,65,66], thus offering novel platforms for biomimetic ion transport and for possible therapeutic applications[67].

For the purpose of clarity, in the following discussions, the long and hollow DNA structure is referred as DNA nanotube (DNT) if it exists in an aqueous solution, and as DNA nanopore when embedded in the membrane.

1.4.1) Experimental studies

Nanopores can be formed in biological (i.e. lipid-bilayer) membranes, by a variety of materials, both biological and synthetic[68]. Membrane nanopores have been shown to be formed by carbon nanotubes[69], synthetic molecular assemblies[70], and cyclic peptides[71], in addition to biological pore-forming proteins such as porins[17] and α-hemolysin[11], and bundles of α-helices[72]. Nanopores are of importance not only in terms of enabling exploration of fundamental aspects of solute transport at the nanoscale[73], but also because of their potential applications in biotechnology[74] and biomedicine[75]. For example, they have been used as stochastic sensors[20,76] (called stochastic sensor because the current fluctuations or any observation so made occurs due to reversible random binding events of analytes in the receptor pores), and for the translocation of proteins[40] or nucleic acids[24,77]. In particular, protein nanopores form the basis of new developments in next generation sequencing of genomic DNA[78]. Despite advances in design of novel protein nanopores[51], there is a need to expand the range of materials which may be used to design new membrane-spanning nanopores, allowing controlled pore selectivity and gating. Nucleic acids have recently been shown to constitute suitable building blocks to form novel nanopore scaffolds (see discussed references ahead). Additionally, the scaffold DNA strands allow precise control over the positioning of desired chemical modifications[79]. This feature in particular allows positioning of the membrane anchors and potentially the DNA surface may be modified to act as gating valves. (Further discussion is limited to DNA nanopores with references up to August 2016)

For nanopore applications, DNA nanostructures[80] have been used either in combination with solid-state nanopores[41,81] or these nanostructures acted as nanpores by themselves in the lipid membranes. I will be focusing on nanopores formed by DNA in lipid membranes.

The application of DNA to form stable nanopores in hydrophobic membrane cores requires suitable chemical modification of the DNA surface. Phosphate groups on the DNA backbone provide a negative surface, and overall a DNA helix presents a hydrophilic surface. Ideally, the DNA surface needs to be modified to neutralise the negative charge and to provide hydrophobic anchors supporting interaction with the hydrophobic membrane core.

Langecker *et al.* reported[82] the first membrane embedded DNA nanopore (Figure 1.10A). The pore design consisted of a stem part, which traverses through the membrane while a barrel shape cap sticks at one side of the membrane. In brief, the whole complex was 42 nm in length with pore size of 2 nm, and held in the membrane with the help of 26 cholesterol molecules. Ion conductance, and DNA translocation studies supported the functional and stable nature of the pores. Almost at the same time, Burns *et al.*, working independently, reported[57] later a relatively simpler, and shorter DNA nanopore of length 14 nm having pore size of 2 nm (Figure 1.10B). To render hydrophobicity on the DNA surface they provided 72 ethylphosphorothioate modifications in the middle of nanopore. The thickness of hydrophobic belt was almost same as hydrophobic core of membrane thus providing maximum hydrophobic interaction within the membrane core. Ion conductance studies performed through this pore revealed cation selectivity. Burns *et al*.[65] later reported a similar nanopore but with only two porphyrin rings as membrane anchors (Figure 1.10C). This study suggests that instead of many ethylphosphorothioate modifications only two porphyrin rings can provide sufficient hydrophobicity for membrane retention.

Thereon, various versions of DNA nanopore formed by six helix bundles have been reported, mostly pioneered by Stefan Howorka and Ulrich Keyser (see below).

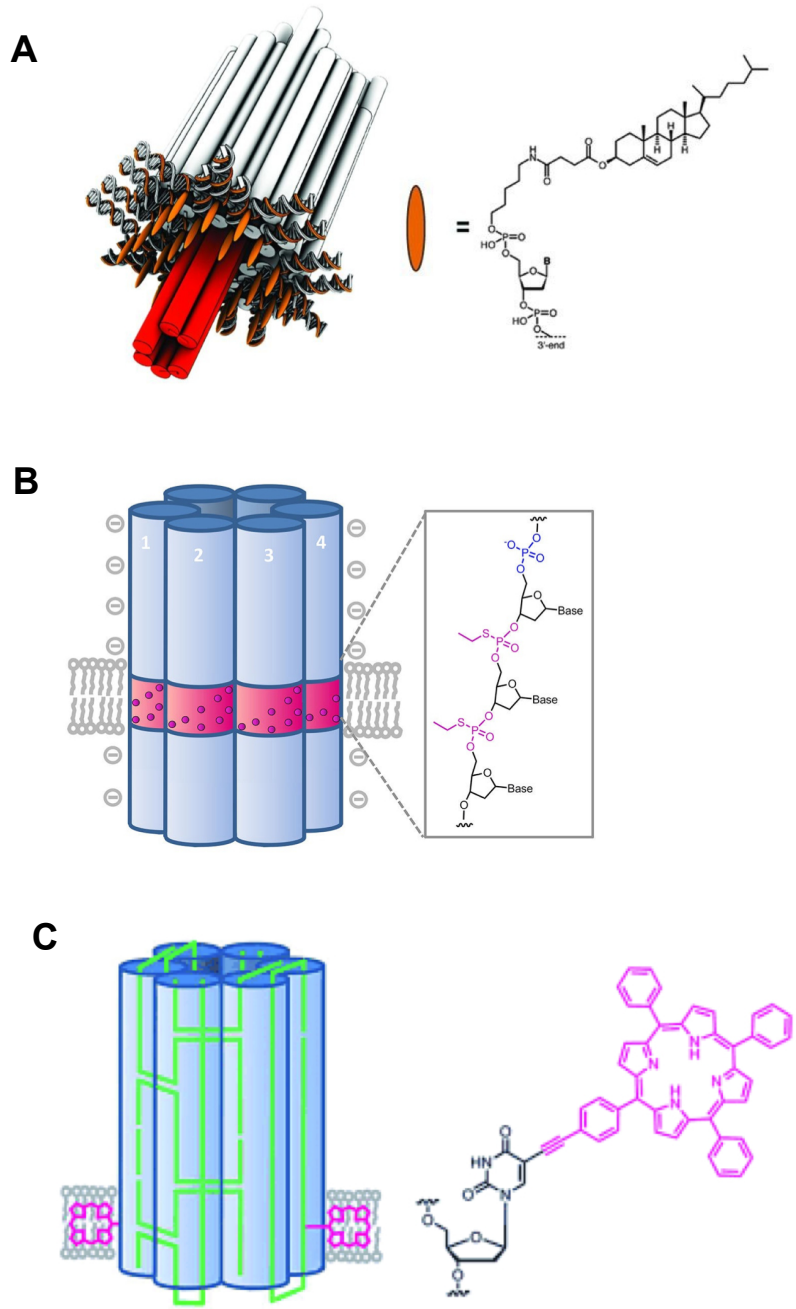

Figure 1.10: **Initial DNA nanopore principle designs**. **(A)** The first designed DNA nanopore consisted of cholesterol anchors (orange ellipsoids; inset on right). The bundle of protruding rods (red) is suggested to be inserted in the membrane while overall cholesterol helps to maintain the interaction of the full structure with the membrane. **(B)** Simple six double helix DNA bundles nanopore design supported by ethylphosophorothioate modifications shown as thick red belts (see right inset). Each cylinder represents a dsDNA though it should be noted that this design consists of folding staple strands (See Chapter 3). **(C)** Another six double helix DNA nanopore design but with only two porphyrin rings (see right inset). Figure **A** from Langecker *et al.*[82], reprinted with permission from AAAS, Figure **B** reprinted with permission from ref[57], copyright 2013 American Chemical Society, and Figure **C** from Burns *et al.*[65], with permission John Wiley and Sons.

Seifert *et al*.[83] studied DNA nanopores with porphyrin anchorages at low and high voltages, and noted higher and lower conducting states, respectively. Such an observation was explained as voltage-induced conformational changes in the DNA nanopore. Over the years, increased understanding of such designs has led to control of functional nature of such nanopores. Burns *et al*.[66] reported new generation cholesterol tagged shorter length, ca. 9 nm, DNA nanopores, and their open and closed conducting states were controlled using a lock and key mechanism. The closed state of the pore featured a 'lock' constituting a DNA strand at the pore entrance forming a duplex, which was also hybridized at the two opposite extended DNA dock sites. The open state was achieved using a strand-displacement approach by adding a 'key' DNA strand, which completes the base-pairs with the 'lock' strand thus finally releasing it free. The simpler design with shorter length DNA nanopore was aimed to avoid structural conformational changes, and current leakage. A four helix bundle DNA channel reported by Göpfrich *et al*.[84] provides a relatively smaller inner lumen diameter compared to the six helix bundle nanopores. Very recently, and interestingly the simplest nucleic acid based channels, by Göpfrich *et al*.[85], exploit just a single DNA duplex embedded in a membrane supported by 6 porphyrin rings (thus as such itself not a 'nanopore' with a lumen). It was suggested that lipid bending towards the DNA surface leads to opening of space around the DNA surface thus leading to conducting states of such channels. Thus, the formation of toroidal lipid nanopore was suggested. Observing this, they also suggested that the total conductance through DNA nanopores results due to ion flow through the lumen as well as through the toroidal lipid pore formed around the DNA nanopore. They also mentioned that previously toroidal lipid pore formed by peripheral membrane truncated α-haemolysin[86] had been reported but in this case the toroidal pore was formed by the integral membrane structure (*i.e.* single dsDNA). It is to be noted that the DNA surface in such case is not completely neutralised, and as discussed in Chapter 4 the negatively charged surface can perturb the lipid bilayer to a more extent compared to the neutralised DNA surface. Apart from this, as mentioned above, one of the recently reported study, Göpfrich *et al*.[43], of a largest membrane embedded (54 nm length and 6 nm minimal cross-section) DNA 'porin' is stabilized by the presence of cholesterol anchors, results in substantive local reorganization of the bilayer and leakage of ions and water around the edges of the

inserted nanopore. Most recently Krishnan *et al*.[44] also created a large DNA nanopore (~4.2x4.2 nm^2 cross-section area). They designed a T-shape DNA nanopore where stem part is formed by a square arrangement of 12 helices attached to a DNA plate. Number of tocopherol moieties were attached to the DNA plate to assist membrane interaction. The large diameter of pore enabled them to translocate dsDNA through the pore.

Also, attempts have been made to exploit such nanopores as anticancer agents. Burns *et al.*[87] used a DNA nanopore with ethylphosphorothioate modifications provided at the terminal end of the nanopore, and incubated them with cervical cancer cells. They noticed pronounced cytotoxic effect with DNA constructs having hydrophobic modifications in contrast to control constructs without such modifications. As such these nanopores were not selective for cancerous cells due to a lack of any targeting moiety but it opens a new field for their applications in therapeutics.

1.4.2) Computational molecular dynamics studies

Computational studies of DNA nanopores (also DNA origami systems) have been sparse compared to experimental reports. Until now (August 2016), to the best of my knowledge there are only three groups globally, including ours, who employ computational molecular simulation studies to understand DNA nanopore or DNT dynamics. Another group employs Monte Carlo and MD techniques to study self-assembly of DNA structures, and also involved in the development of oxDNA coarse-grain forcefield[88].

Yoo *et al*.[89] simulated one of the first complex DNA-origami in the form of an array construct. Their DNA-origami construct consisted of 6x3 honeycomb and 4x4 square arrays of DNA helices. The major aim of the study was to understand mechanical properties of such systems. Joshi *et al*.[90] studied smaller systems constituting 6 and 8 helix DNTs, and found 6 helix bundle nanotubes to be more stable. Later, I published[91] a study on the dynamics of DNT structures in pure aqueous solution, which is discussed in detail in Chapter 3. Immediately after, Yoo *et al*.[92] published simulation studies of a DNA nanopore embedded in a

membrane. They applied transmembrane voltage to understand the conductance properties, and concluded that DNT structure is quite porous (my study suggests the same; see Chapter 3) where a major portion of the current flows through the DNA walls rather than the central lumen. Later, Joshi *et al*.[93] compared 6 helix DNT with and without additional supportive pillars. Using steered MD they found the stretch modulus of the pillar supported DNT was twice as large as the DNT without such support.

One recent study, though not related to DNA nanopores, by Maffeo *et al.*[94] simulated a reported experimental[95] cryo-em DNA-origami 'pointer' structure. They presented a simulation methodology for a huge origami construct, and could achieve stable RMSDs within a reasonable computational time. Interestingly they have used elastic network modelling in all-atom simulations, which is mostly used in coarse-grain modelling. I believe that such simulation methodology could prove to be quite useful for other DNA-origami simulations.

The most recent experimental study, as mentioned above, was supported by computational studies of a very large DNA porin channel[43]. Though this system size is quite huge, thus computationally expensive, I believe this study was not simulated long enough to derive any concrete results using computational approaches. I have explained the reasons; see Chapter 5 (section 5.3.2), where I build up my understanding by simulating relatively simpler but comparable systems of DNA nanotube embedded in a membrane.

In this thesis I will be focusing on one of the experimentally[57] reported and designed DNA nanostructure object, called a DNA nanotube. Simple geometry principles and steps used to construct such structures are explained in Chapter 2 (Methods). In Chapters 3, 4, and 5, I discuss all-atom and coarse-grain simulation results of DNT and their interaction with membranes.

CHAPTER 2

2 Methods

I have used classical molecular dynamics (MD) simulations techniques as a computational 'microscope'[96] to understand molecular interactions at different resolutions, *i.e.* all-atom (AT) and coarse-grain (CG). In particular in this thesis I have used MD to understand DNA nanotube (DNT) and its interaction with the membrane. I discuss here the methods involved in my MD simulations, and the basic geometric principles for DNT modelling.

2.1) Classical molecular dynamics

Molecular dynamics simulations are computational experiments[97] where one prepares a sample (selecting a 'reasonable' model/structure under question), and then the particles constituting the sample are allowed to interact with the evolution of time. In general MD is a deterministic approach[98] where one simulates the system until the system 'properties' do not change with the time, thus called equilibrium dynamics, and finally the equilibrium configurations are used to calculate the average properties. Ideally the interaction of particles should consider all the nuclear elements and electrons, *e.g.* as in quantum mechanical models. Quantum methods are computationally very expensive, and cannot be readily applied to large biomolecular systems. Classical MD is an approach based on empirical force-fields methods where electrons motions are ignored, and only nuclear positions are used to calculate the interactions[98]. The word *classical* comes from the fact that nuclear motions follow classical mechanical laws[97].

The basic MD algorithm based on classical dynamics to calculate equations of motion is shown as a flowchart in Figure 2.1.

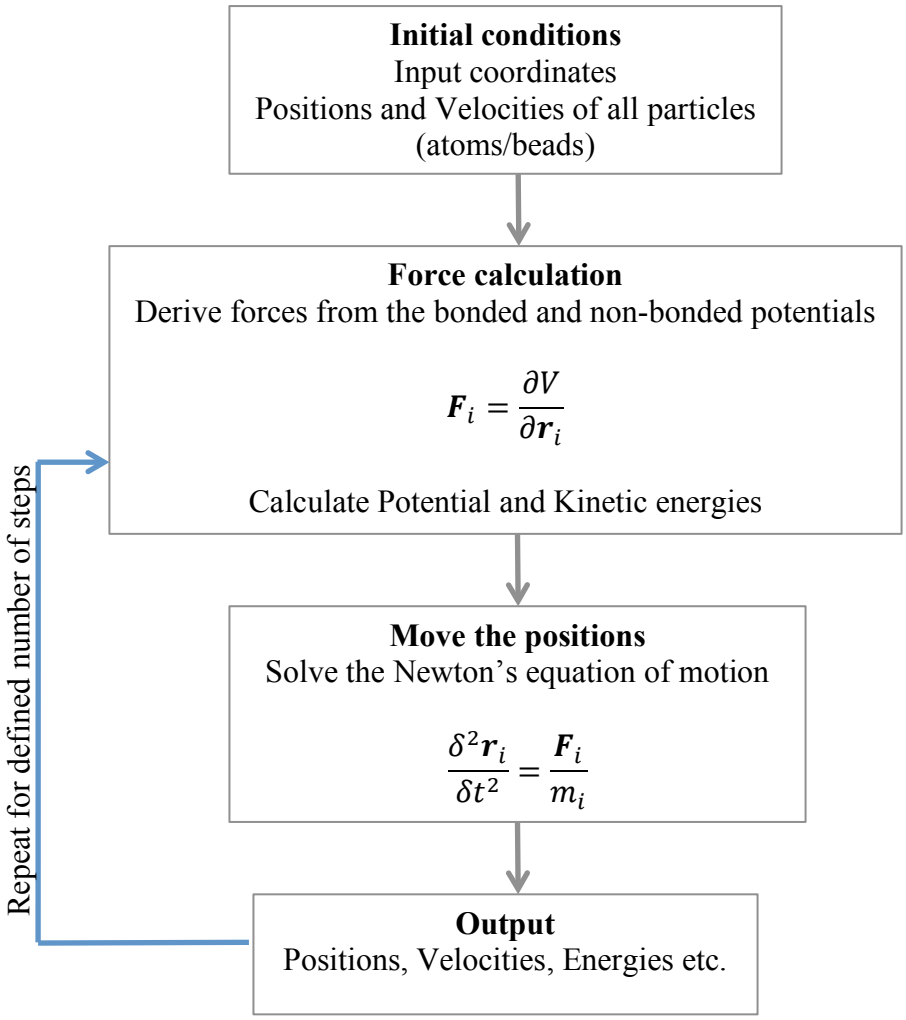

Figure 2.1: A general MD algorithm is shown in the form of flow chart steps based on GROMACS[99] manual 4.6.7.

In general, initial velocities are provided using Maxwell-Boltzmann distribution. Further, the trajectory of motions can be solved using the *finite difference method*. The integration is broken in small steps each separated by a fixed time δt. Force on each particle is calculated as total vector sum of interaction with the other particles. The calculated force is further used to determine the position and velocities at time $t + \delta t$. Again forces are calculated at new positions to determine positions and velocities with a further but constant time step δt. It should be noted that force is assumed to be constant during any δt. This assumption can result in errors, thus choosing δt needs to be optimised as discussed ahead.

In brief, this process can be generalised using the *leap-frog* algorithm implemented in GROMACS[99] (manual 4.6.7):

$$\boldsymbol{v}\left(t+\frac{1}{2}\delta t\right)=\boldsymbol{v}\left(t-\frac{1}{2}\delta t\right)+\frac{\delta t}{m}\boldsymbol{F}(t)$$

$$\boldsymbol{r}(t+\delta t)=\boldsymbol{r}(t)+\delta t\boldsymbol{v}(t+\frac{1}{2}\delta t)$$

Smaller values of δt will lead to fewer integration errors. For the same number of simulation steps, the system will explore only a small phase space with very small δt values, and on the other hand very large δt value can cause particle-particle collision leading to high energy overlaps[98]. In general it is suggested that δt should be a tenth of the period of fastest motion which is usually of the order of a femtoseconds time scale. I have mentioned the chosen time steps in the respective Methods sections of each Chapter.

As shown in the flowchart above (Figure 2.1), the force-fields calculate the interaction of particles as the potential energy between bonded and the non-bonded particles. Thus, the total potential energy (V) of the interacting particles can be written as:

$$V = V_{bonded} + V_{non-bonded}$$

I discuss the individual contributing terms in the following sections.

2.1.1) Bonded potentials

The bonded potential is the summation of four terms defined by 2, 3 or 4 atoms connected by covalent bonds. They are shown in Figure 2.2.

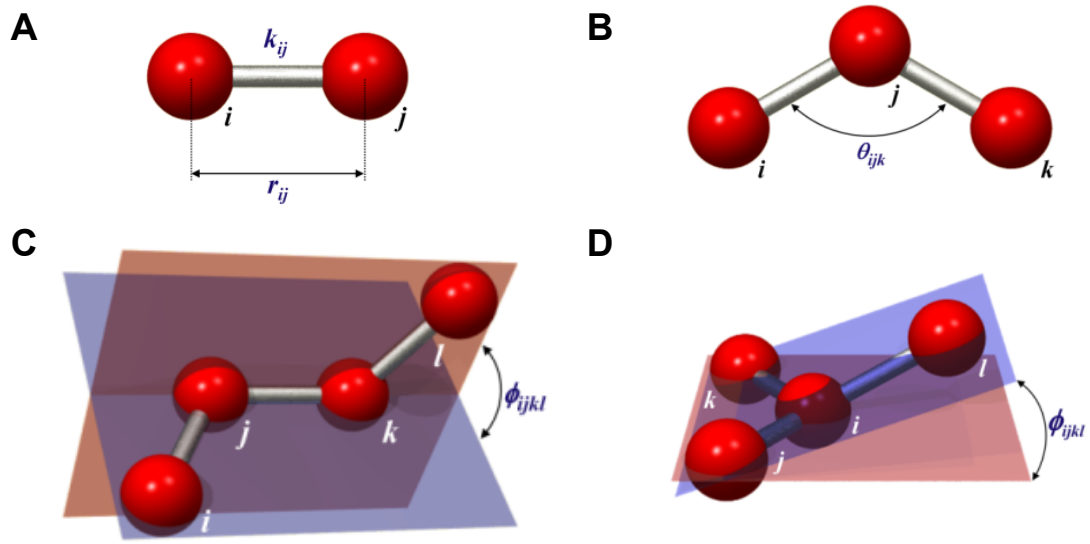

Figure 2.2: **Bonded interactions**. **(A)** Harmonic bond between two covalent connected atoms. **(B)** Angle potential is between three atoms, thus will always be in a plane. **(C)** The proper dihedral defines the angle between two planes defined by four atoms sharing a common 'hinge' edge (*jk* here) along which rotation is allowed. **(D)** Improper dihedral also formed by four atoms as shown which helps keep maintaining the planarity or chirality in selected atoms.
Figures taken from *http://cbio.bmt.tue.nl/pumma/index.php/Theory/Potentials*

In general the total bonded potential can be expressed as:

$$V_{Bonded} = \frac{k_{ij}}{2}\left(r_{ij} - r_{ij}^{0}\right)^2 + \frac{k_{ijk}}{2}\left(\theta_{ijk} - \theta_{ijk}^{0}\right)^2 + k_{ijkl}\left[1 + cos\left(n\phi_{ijk} - \phi_{ijkl}^{0}\right)\right] + \frac{k_{ijkl}}{2}\left(\phi_{ijkl} - \phi_{ijkl}^{0}\right)^2$$

The first and the second term separately define the potential due to bond stretching and angle bending respectively (Figure 2.2 A and B). They are described by Hooke's law, and thus harmonic in nature. r_{ij}^{0} and θ_{ijk}^{0} are 'reference' bond length and angle respectively when other bonding terms are set to zero[98]. Any change from the reference value will lead to a change in potential, which is penalised by the force constants (k_{ij}, k_{ijk}). It is difficult for a system to deform from their reference bond and angle values, which require quite high energy values. The dihedral terms (proper and improper), along with the non-bonding terms, are usually held responsible for any conformational changes. The proper dihedral term is formed by four atoms connected consecutively, and this

periodic function restricts or allows the rotation along the central bond (*kj* in Figure 2.2C). The improper dihedral is also constituted by four atoms but in this case three atoms (*j*, *k*, *l* in Figure 2.2D) are centered around the fourth atom (*i* in Figure 2.2D). This potential function is particularly incorporated to help maintaining the chirality (also prevent 'flipping', *e.g.* in tetrahedral geometry) or planarity in the selected atoms.

In general the reference bond lengths and angles are obtained from microwave spectroscopy or high resolution crystallographic data. The force constants are usually derived from quantum mechanical calculations or vibrational frequencies from infrared data[98,100].

2.1.2) Non-bonded potentials

Non-bonding potentials are contributed by Coulomb and van der Waals interactions.

Electrostatic interaction.

Interaction between any two charged particles q_i and q_j, with free space permittivity ϵ_o, and separated by distance r_{ij} can be defined using Coulombs law as:

$$V_E(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_o r_{ij}}$$

These charges are considered to be restricted to the nuclear centers, and are called partial atomic charges. The distribution of partial charges on each atom in a molecule depends on the electronegativity of each atom[98]. Mostly, for ions the charges are considered to be point charges. It can be seen that Coulomb interaction scales and decays slowly as r^{-1}. Thus long-range electrostatics play an important role.

Considering the expense of computational costs required for any paired interaction calculations in the periodic boundaries (discussed ahead), the MD engines use some cut-off range for short-range interactions but at the same time one cannot ignore the long-range electrostatics (slow decay to zero). Thus, long-range

electrostatics are usually considered in the form of a Particle-Mesh Ewald (PME) approach. The PME is based on Ewald summation, which was developed initially for ionic crystals[101]. In this method the calculated short-range electrostatics converge quickly in real space while the long-range function converges in Fourier space. For the efficiency reasons the Fourier transformations of the charge field density, usually fast Fourier transformation, is performed to evaluate the density fields in a discrete lattice space. Overall this method scales as $N\ln(N)$, which is quite fast compared to N^2 scaling. I have used PME for all-atom MD simulations.

van der Waals interactions.

The van der Waals interaction involves attractive (long-range London dispersive forces) and repulsive (at short distances) forces depending on the distance between the interacting particles. It is usually calculated using a Lennard-Jones potential described as:

$$V_{vdW}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

where ϵ_{ij} is the potential energy well depth (the interaction strength), σ_{ij} is the collision diameter (the separation at which the energy is zero), and r_{ij} is the distance between particles. This potential can be considered to consist of an attractive part, varying as r_{ij}^{-6}, and the repulsive part, which varies as r_{ij}^{-12}. As with the electrostatics the Lennard-Jones potential suffers from a drawback of long-range weak attractive forces. This is a paired interaction thus it becomes computationally quite expensive to consider periodic images. To overcome this, in MD simulations various cut-offs with shift or switch functions are introduced to smoothly decay the function to zero. Cut-off values and the functions depend upon the system, and force-fields. I have mentioned them in the Methods section in each chapter.

2.2) Simulations methodology

2.2.1) Periodic boundary conditions

In MD simulations periodic boundary conditions (PBC) are applied so that any particle can experience forces as if it is in a bulk solution. This helps in avoiding simulating large numbers of particles. In its simplest form it can be considered that in PBC the main unit simulation box is repeated infinitely in all the dimensions. For example, in Figure 2.3 the central unit cell box is surrounded by periodic images ('image' because it is the same copy of main unit cell). If a particle leaves the unit cell from one side (in 2D) or face (in 3D) then it can re-enter the unit cell from the opposite side or face of the box. PBC usually includes a *minimal image convention* according to which, for non-bonding interactions, each atom can interact up to maximum one closest image (repeated units) of the remaining atoms in the system[98].

Depending on the system of interest under study, PBC can suffer from some drawbacks. Some membrane studies involving, for example, undulations[102], and effect of asymmetric distribution of lipids[103] have been found to introduce complications and be difficult to study[104].

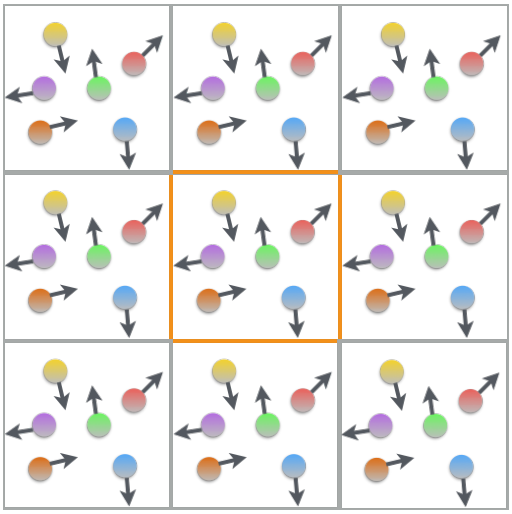

Figure 2.3: **Periodic boundary conditions in 2D**. Central unit cell in *orange* has been replicated which represents a general PBC approach. Each colour bead represents a different species, which can be chemically identical as well.

2.2.2) Temperature and Pressure control

The MD algorithm discussed above leads to conservation of energy (potential and kinetic energy), and thus will produce *NVE* ensembles where *N* number of particles, *V* volume of the system, and *E* energy of the system do not change. For biological systems another form of ensemble *NPT* is quite relevant where pressure *P* and the temperature *T* of the system are kept constant[98]. Various algorithms have been suggested to keep pressure and temperature of the system constant. Here I will discuss briefly the temperature and pressure controls used in my thesis implemented in AMBER[105] or GROMACS[99].

Temperature control.

In general the temperature, *T*, of the system can be related to the kinetic energy, K_{NVT}, as[98]:

$$K_{NVT} = \frac{3}{2} N k_b T$$

where *N* is the number of particles, k_b is the Boltzmann constant.

The Berendsen algorithm[106] is one way to control the temperature in which the system is weakly coupled to the external heat bath at temperature T_0. Any change in the system from this value is corrected slowly as (GROMACS manual 4.6.7):

$$\frac{dT}{dt} = \frac{1}{\tau_T}(T_0 - T)$$

where τ_T is the time constant, and *T* is the temperature. Using this method the fluctuation in kinetic energy is suppressed.

Another method called velocity-rescaling[107] is similar to Berendsen's method but includes an additional stochastic term to correct the kinetic energy distribution. I have also used Langevin dynamics[108], which adds a friction term, and comes with stochastic-temperature coupling. I have conducted my simulations at 300 or 310 K, and mentioned in the Methods section in each chapter.

Pressure control.

Similar to the temperature control, the pressure of the system can be coupled to an external pressure bath. Thus, Berendsen[106] pressure control can be given by:

$$\frac{dP}{dt} = \frac{1}{\tau_P}(P_0 - P)$$

where τ_P is the time constant, and P is the pressure.

In this algorithm the coordinates and box vectors are rescaled after some defined number of steps. Another method called Parrinello-Rahman[109] can also be used to achieve the pressure coupling.

In case of membrane systems the pressure coupling in all dimensions can result in reduced surface area. In such a case semi-isotropic coupling is imposed along the plane of the bilayer (say *xy*), which changes the *x* and *y* box length vectors together while the *z* vector varies independently. In all cases the pressure was maintained at 1 bar.

2.2.3) Energy minimisation

Before any MD is initated the starting configuration should be free from steric clashes, and thus any atomic overlaps should be removed. Energy minimisation is usually done to reach local minima around the given configuration. This is very helpful in case of systems that are inserted in a membrane or solvated in ions and water solution. It is also advantageous in case of systems built using geometric principles, for example in my case where I built initial DNA nanotubes discussed ahead. Energy minimisation can be achieved using steepest descent or conjugate gradient methods. In steepest descent method the geometry (configuration) is minimized in the direction of largest gradient. The forces and potential needs to be calculated at each step, where new position is given by (as implemented in GROMACS[99]):

$$\boldsymbol{r}_{n+1} = \boldsymbol{r}_n + \frac{\boldsymbol{F}_n}{max(|\boldsymbol{F}_n|)} h_n$$

where, $\boldsymbol{r}_n$ and $\boldsymbol{r}_{n+1}$ are old and new position vectors, $\boldsymbol{F}_n$ is the negative gradient of the potential V, h_n is the maximum displacement. In general the algorithm accepts the new position if $V_{n+1} < V_n$, and then h_{n+1} is updated to $1.2h_n$, otherwise position is not accepted and h_n becomes $0.2h_n$. This process is performed for the defined number of steps. In this method the gradients and direction are orthogonal while in conjugate gradient method at each step the gradients are orthogonal while

the direction in next step is *conjugate*. A working example can be found in reference[98].

2.2.4) Multiscale approaches

I have used multiscale simulation approaches at all-atom (AT) and coarse-grain (CG) resolutions. In general AT, as the name suggests, each interacting particle represents an atom itself. In CG, a few atoms are modelled as one bead thus reducing the total number of interacting particles in the system in contrast to AT models. In each chapter I describe simulations conducted at different resolutions.

All-atom forcefield.

For all-atom simulations I used AMBER parmbsc0[110] and CHARMM36[111] forcefields. It is difficult to explain a generalized difference between atomic force fields. Usually the potential function terms are same but the major differences can be seen in the form charge derivation, torsion angles, van der Waals parameters etc.[98] Thus, usually extended simulation studies are performed to compare the performance of forcefields to check the correct reproducibility of some experimental observables[112,113]. Also, it is not advised to introduce missing parameters in one forcefield from the other[104].

Coarse-grain forcefield.

Here I will briefly explain MARTINI CG forcefield for DNA[114] and lipids[115], which I have used exclusively for my CG simulation studies.

For the DNA, MARTINI forcefield maps all-atom DA/DG and DT/DC bases to 7 and 6 CG beads respectively as shown for the DA case in Figure 2.4.

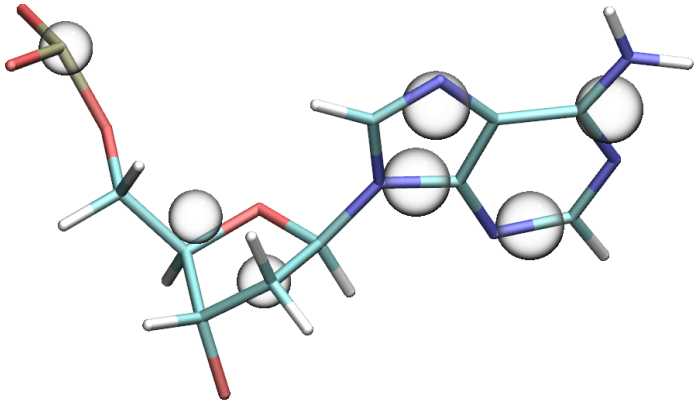

Figure 2.4: **CG DNA**. Representation of DA base in all-atom (connected sticks) and MARTINI CG (beads) resolutions. Note the beads size shown here are not based on their respective σ values. More details about bead types and CG representation for other DNA bases can be found in this reference[114].

Depending on the bead type, each bead can consist of 2 to 4 heavy atoms. In terms of polarity, MARTINI beads are categorised as: polar, intermediate, apolar, and charged particles. These beads have their sublevels based on their hydrogen bonding capacities as donor, acceptor, donor-acceptor or none. It should be noted that in CG the mass of bead does not depend on the constituent atoms. Conventionally for computational efficiency reasons[116] there are only two defined masses in CG (72 or 45 amu.). In general for proteins, lipids and also DNA, the σ (the separation at which the potential is zero) is defined as 0.47 nm for the normal beads and 0.43 nm for 'small' beads. Special 'tiny' beads, with σ = 0.32 nm, are defined for beads involved in stacking of bases. In DNA hydrogen bonding between specific base pairs is quite important, thus in MARTINI special DNA tiny beads were optimised to achieve the same.

For the CG membrane simulations, the latest MARTINI lipid parameters were used[115]. In particular I have used POPC and POPS lipid types, which consist of one bead less than the older MARTINI version. In the recent version one of the beads in one of the aliphatic chains (oleic) has been removed, which was shown to provide overall better membrane properties (thickness, area)[115].

Backmapping.

For multiscale simulations, first I simulated membrane embedded DNA nanotube systems in CG then one of the final configurations was backmapped to all-atom

resolution. The backmapping was done using *backward.py* by Wassenaar *et al*.[117] With additional modifications, their backmapping originates from the methodology by Brocos *et al*.[118] in which, the interpolation of geometry is followed by structural relaxation or energy minimization. For example, in the simplest case of CG bead chain, the first atom is placed on the reference CG particle (with *n* heavy atoms) and the remaining *n*-1 atoms are placed on the trace segment joining two adjacent beads. It is followed by energy minimization process or with small MD runs with defined parameters (quite useful specially in case of system containing water molecules or with membrane systems).

2.3) DNA nanotube modelling

I will discuss my DNA nanotube (DNT) building procedure in Chapter 3, and here I discuss basic DNA folding geometry principles that were used to build the DNT model using my own scripts written in PERL, and integrated with *tleap* module in AMBERTools[105].

I assume B-DNA with 10.5 base pairs per turn for constructing DNT. This provides twist angle ~34.29° (360/10.5) which defines the basic minimal accessible angle for crossing over DNA strands (Figure 2.5A). One can design cross over points, at least theoretically, based on various allowed accessible angles to achieve overall 2D or 3D shape of DNA nanostructure[119]. The 3D structural designs can be based on square lattice or honeycomb hexagonal lattice. My DNT is based on a hexagonal lattice arrangement (Figure 2.5B), thus further discussion and construction is based on this only. A nice review by Kearney *et al.*[119] provides more information about other different geometrical aspects of DNA construction.

In a hexagonal lattice arrangement possible folding can occur at nx7^{th} (where '*n*' is an integer) base pair intervals. The total angle spanned by a continuous 7 bases is 240° (7x34.29°) which is the external angle of a regular hexagon. Thus, this forms the fundamental basis, and the requirement of a hexagonal shape (Figure 2.5B).

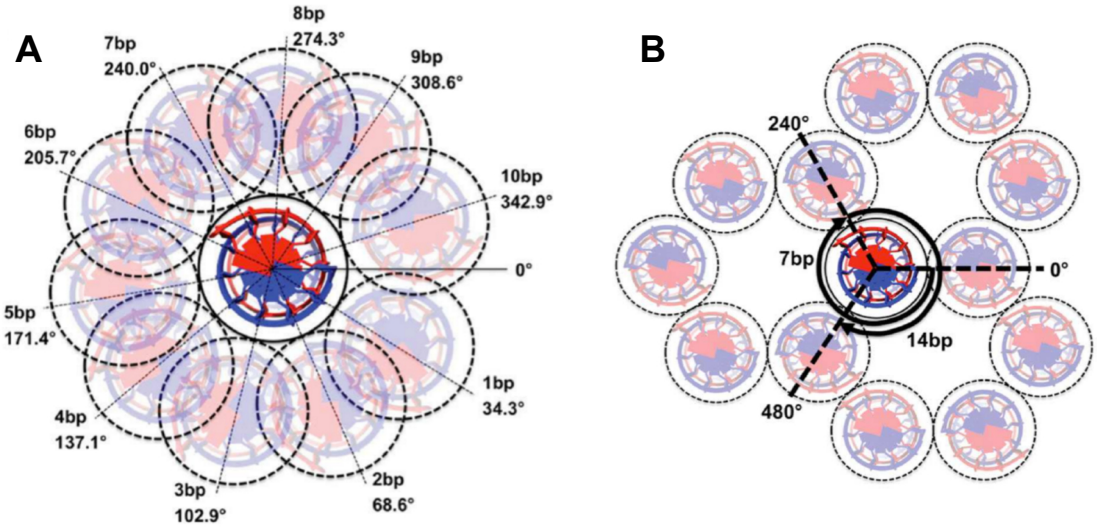

Figure 2.5: **Cross-over accessible angles**. **(A)** Various possible accessible angles shown for DNA cross over points (as seen from top). It is based on the assumption with B-DNA having 10.5 base pairs per turn, which provides a twist angle of 34.29°. **(B)** Honeycomb lattice formed by hexagonal arrangement of six DNA double helices. Note 7 base pairs form an external angle 240° (7x34.29°) of a regular hexagon. The DNT considered in my thesis is based on hexagonal geometry. Figures **A** and **B** from Kearney *et al*.[119], with permission John Wiley and Sons.

I will explain this, using above basic geometric principles, providing direct example from my DNT based on experimental design[57]. Figure 2.6A shows a 2D sketch of 42 base pairs long 'double helix' (only part shown; for full figure see Chapter 3, Figure 3.2) with four cross-over junctions where strands of DNA fold and cross towards the adjacent double helix (shaded region). For the purpose of understanding, and also for the ease of DNT building, the broken strands (highlighted in a dotted box Figure 2.6A) can be considered as a regular single straight strand ignoring crossing over points. It can be seen that the cross-over points are *n*x7 base apart. In this particular example they are at 7th, 14th, 28th and 35th base numbers. If one starts from 5' *cyan* end (base label 1) it will be the O3' of DNA base number 7 that will join P of an adjacent DNA helix base (shaded region). Similarly, P of base number 8 will join with O3' of an adjacent helix (shaded region).

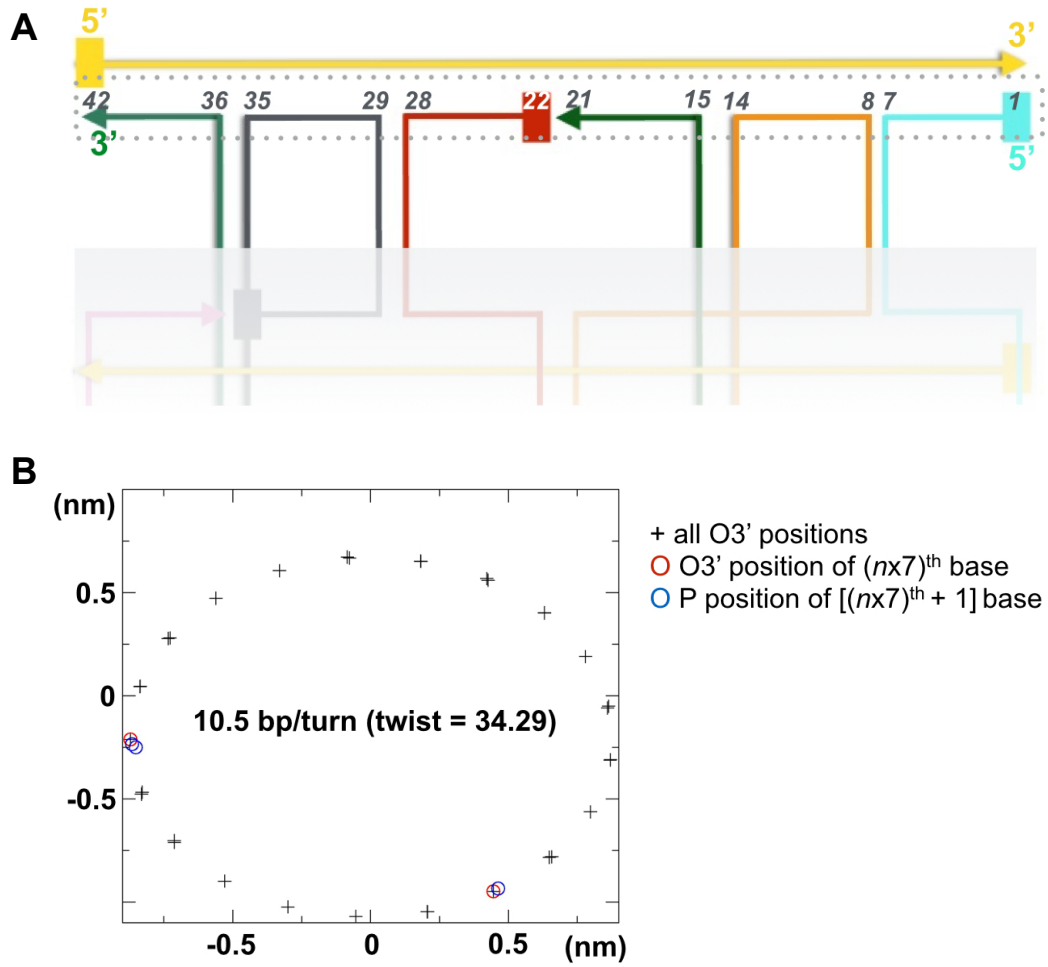

Figure 2.6: **Geometric principles used for DNT building**. **(A)** Part of full 2D sketch of DNT model used in the thesis (for full figure see Chapter 3, Figure 3.2). The *golden* scaffold strand does not fold, while different colours staple strands can fold at nx7th base pair (see also Figure 2.5). Number labels (1,7,8…42) show residue bases forming cross-junction. See main text for details. **(B)** *XY* projected positions of O3' and P of 'assumed straight' staple strand shown in a dotted box in **A**. This shows all the four O3' (*red* circles) and four P (*blue* circles) involved in inter-helix crossover are at 120° (hexagon internal angle), and see Figure 2.8 for clarity. Compare with Figure 2.7.

Also, if one imagines from the top along the *Z*-axis all the O3' atoms (labelled 7, 14, 28, 35) and P atoms (labelled 8, 15, 29, 36) of the connecting residues should be 120° away depending upon where they are positioned (Figure 2.6B). As discussed ahead (Figure 2.8), it should be noted that in Figure 2.6B O3' (*red* circles) in one helix joins the P positioned (*blue* circles) on the adjacent helix (but not P; *blue* circles in the same helix).

To stress the importance of consideration of twist angles, Figure 2.7 shows O3' and P are positioned wrongly if DNA with 10 base pair per turn is considered. Thus building a structure using this geometry will be wrong, as it will introduce extended bond lengths while joining with the residues in an adjacent helix.

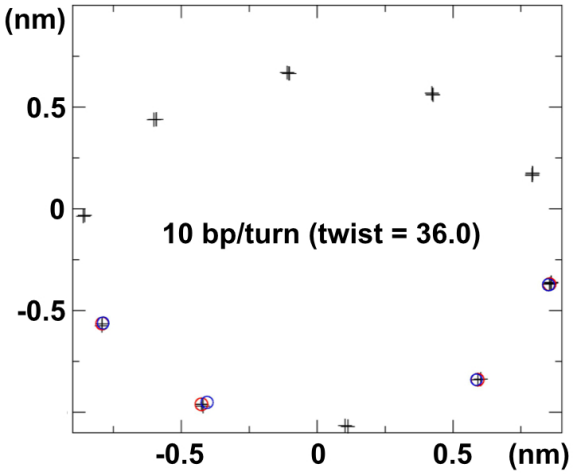

Figure 2.7. **Distribution of connecting residues using '*wrong*' twist**. A similar plot to Figure 2.6B showing the connecting residues (*blue* and *red* circles) are not in proper position, *i.e.* not separated by 120°. This importance can be realised after looking at Figure 2.8. The keys ('+' and circles symbols) are same as in Figure 2.6B. Note: This has been just shown for representational or 'tutorial' purposes to understand the geometric principles, and has not been used to construct the DNT.

Above (Figure 2.6), I have shown just one helix which is part of the six helices bundle DNT hexagon design. Once all the connecting residues are known one can arrange six double helices (more details in Chapter 3 Methods), and join the residues at suitable positions. Figure 2.8 shows six helices arranged in hexagon format where O3' (*red* circles) on one helix are joined to P on the adjacent helix (*blue* circles), thus correctly positioned.

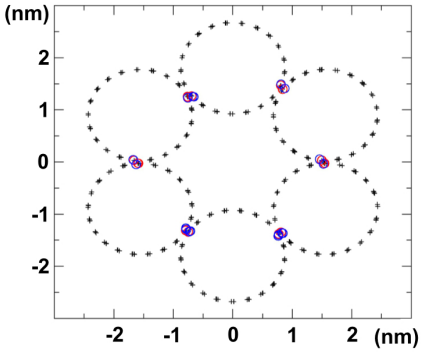

Figure 2.8. **Six helices bundle DNT final design (projected top view of selected O3', P coordinates)**. Each circle showing *XY* projected positions of O3' and P of 'assumed straight' all staple strands (see Figure 3.2). The keys ('+' and circles symbols) are same as in Figure 2.6B. O3' (*red* circles) on one helix are joined to P on adjacent helix (*blue* circles). Note the connecting residue atoms are at their accessible angle of 120°, and also very close to the neighbouring helix connecting residues. This allows reasonable bond lengths when the residues are connected together while building the whole complex.

It is to be noted that DNA-origami design consists of junctions (Holliday-junctions; HJs) where the staple strands fold and crosses over the scaffold strands. It has been found that HJs are polymorphic in nature, and in free solution (under no stress) with ions (Mg^{+2}) they exist in the right-handed stacked conformation. DNA inter-helix electrostatic repulsions results in interaxial helices angle *ca*. 60° as determined by FRET experiments[89,120,121]. Yoo *et al*.[89] observed that in their DNA-origami MD simulations (starting configuration as ideal 'continuous' DNA strands) the HJs exist in the left-handed conformation. It is to be noted that in DNA-origami there is a close packing of duplexes thus HJs under stress are not likely to exist in the same conformation as found in the free solution. While building my DNT models I did not include HJ parameters, *e.g.* inter-helix bending, rise per base etc. But still, it is intriguing that in future one would like to introduce the HJ helical parameters, obtained from the available experimental data, in the DNT starting models for simulations. Without doubt this requires critical thinking and exhaustive literature search. For example, Watson *et al*.[122] critically examine various DNA-only (non-protein bound) HJ structures, and introduced analytical definitions that could help in defining HJ geometry. Apart from the commonly used definitions like (Figure 2 in reference[122]) the interduplex angle (IDA; angle formed by end of one stacked duplex to the center of HJ to the

end of opposite stacked duplex), J_{roll} (angle related to the stacked duplex arms to junction center in the bisecting plane) and J_{slide} (stacked-duplex translation along the respective helical axis towards or away from the junction), they also introduced J_{twist} (angle constituted by two helical axes when projected on the resolving plane of the junction). They further explain that some definitions, *e.g.* IDA, are sensitive to length of the arms, and helix widths. J_{twist} can be equal to IDA only if HJ arms are very long but if calculated over short arms (usually the case as in experiments) both can be quite different, thus stressing the need to be differentiated. These are quite important considerations while defining HJ geometry from crystal *vs.* non-crystallographic data (FRET, simulations etc.). Learning from these definitions one may impose the HJ geometry in the DNT model at suitable cross-over junctions, and it seems to be quite possible (but I have not attempted yet) to implement this as an automated scripted protocol using nucleic acid builder[123] (NAB) and *tleap* available in AMBERTOOLS[105].

Additionally, as explained above in my DNT building method, I have assumed a B-DNA form (helix diameter 2 nm, and 10.5 base pair per turn), which is the most probable (populated) structure in a fully hydrated state[124]. In a low hydration state or in a different solvent (*e.g.* ethanol 85% v/v) the A-DNA form (relatively wider helix diameter ~2.3 nm, and ~11 base pair per turn) dominates over the B form[124–126]. In the Introduction chapter, I discussed DNA nanopores can be embedded in the membrane. Thus, in the embedded transmembrane domain of DNA nanopores (*i.e.* in the hydrophobic membrane core) one can expect to have low water density (also discussed in Chapter 5, the water density is lower in the DNT walls). In terms of experimental set-up one can assume that B-DNA form of DNA-nanopore exists in the solution, and when it gets inserted then the transmembrane domain may exist in the A-DNA form depending upon the extent of loss of water. Banavali *et al.*[125], and Noy *et al.*[126] have studied the free energy landscape of A ←→ B DNA form transitions in the aqueous solutions (both studies in good agreement each other), and additionally in ethanol 85% solution (by the later group). These studies suggest that the barrier between such transitions is quite dependent on the solvent (pure aqueous *vs* ethanol solution). Importantly it is to be noted that, to the best of my knowledge, there is no study reported yet to

determine the free energy landscape of A ←→ B form transitions for a hydrophobic-DNA duplex (*e.g.* ethylphosphorothioate DNA used to build DNA nanopores in reference[57]). Such study conducted on simpler systems might be helpful in determining whether one should consider using mixed form of DNA (B-DNA outermembrane, and A-DNA transmembrane domain of DNA nanopore) while building DNA nanotubes models. Such transition study of using ethylated dsDNA embedded in lipid bilayer can be very useful which might be considered for future studies. Recently, Waters *et al*.[127] provided new metrics A-B index (ABI), based on the phosphate group position and the glycosidic torsion angles, to evaluate continuum A to B transition in atomistic simulations. This might prove to be a useful metric while studying membrane embedded DNA systems to understand basic geometry of DNA helices. I have not used this metric for my simulations.

Also, it will be interesting to see if simulated annealing method could be used to refine the initial built model. In general, in this approach temperature of the system is increased and decreased slowly in steps, and usually one can iterate through number of such cycles. This helps in overcoming a large barrier, at higher temperature, separating different conformations, and when system is cooled down slowly usually one can reach a more stable conformation defined by a force-field in a more 'natural' way. But this approach could be harmful in case of DNA like structures where it will be difficult to recover the broken hydrogen-bonds in a short duration during cooling down phase.

CHAPTER 3

3 All-atom simulations of DNA nanotube in an aqueous solution

The contents of this chapter are taken directly and/or modified from my paper published in ACS Nano *9*, 11209-11217, with permission American Chemical Society Copyright 2015. Additional analyses have also been incorporated.

3.1) Introduction

The protein nanopores have been engineered for various sensing applications as illustrated with examples in Chapter 1. Considering the limitations of existing protein nanopores (mainly the limited range of pore diameters) these examples stressed the importance of requirement for new generation protein nanopores (*e.g.* may be using *de novo*[49] strategies), and also hints at discovering new class of pores. One of the main limitations to have a wide diameter pores was overcome by designing DNA origami-solid state hybrid nanopores[41]. Such DNA pores can be designed as functional biomimetic nanopores, which could possibly be employed for sensing applications. DNA nanopores/nanotubes (DNTs), the one discussed in this chapter and further in this thesis, is a hollow cylindrical nanobarrel formed by six DNA duplexes arranged in a hexagonal lattice[57] (Figure 3.1). It is a self-assembled pore formed by folding 8 staple strands around 6 scaffold strands (which do not fold); see Figure 3.2. The close packing of DNA helices requires higher concentration of salt (usually > 1 M) to overcome the electrostatic repulsions. Experimentally the assembled nanopore was characterised using various techniques. Single peak in UV melting studies of the structure revealed the cooperative nature of the assembled strands otherwise multiple peaks would have been expected. The dimensions, and the monomeric nature were confirmed using dynamic light scattering experiments and atomic force microscopy. DNT was embedded in the membrane, thus forming a DNA nanopore, after suitably modifying with chemical functional groups to provide hydrophobicity at the surface of nanotubes (Figure 3.1). This was established by inserting the pores in the lipid bilayer and measuring the stable ions current flow.

The design and synthesis of DNTs is aided by the principles of DNA self-assembly and DNA origami[41,56,62,82,128,129] which have been discussed in Chapter 1.

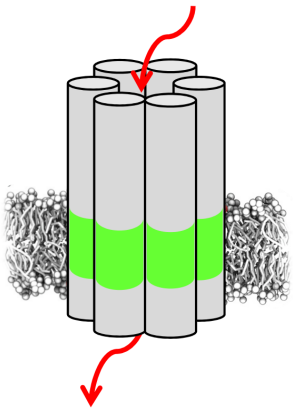

Figure 3.1: **DNA nanopore**. Schematic representation of a 6-helix DNT forming an ion permeable nanopore embedded in a lipid bilayer. DNT helices are represented by grey cylinders with the green regions corresponding to the hydrophobic surface band which stabilizes the nanopore within the hydrophobic core of the membrane. Surrounding lipid bilayer is shown in grey. Figure adapted from Maingi *et al.*[91]

In other experimental studies into their conductance properties established that the DNA pores exhibit two voltage-dependent conductance states[83]. Low transmembrane voltages favour a stable high-conductance level, which corresponds to an unobstructed DNA pore, while higher voltages lead to a low-conductance state possibly caused by electric-field-induced changes in the DNA pore conformation (likely to be flexible and dynamic terminal ends of DNA nanopore). The detailed molecular reason for the two state switching could, however, not be determined due to lack of atomistic insight. An investigation of DNT's conformational dynamics could help settle this question. Furthermore, examining the interactions of DNTs with ions and water will also improve fundamental understanding of the permeation properties and aid the development of novel DNT structures of designed function.

Molecular modelling and simulations can provide important insights into the sub-nanoscale structure and function of nanopores[73]. Molecular dynamics (MD) simulations have also been used to reveal aspects of the mechanical properties of

DNA-origami array structures[89]. MD simulations have also been performed to explore the ionic conductivity and structural deformation of DNA origami nanoplates in an electric field[130]. A recent examination of DNT structures by MD simulations suggested that DNTs formed by six DNA helix bundles may be more stable than those formed by 8-helix bundles[90]. However, when this project was started there was lack of complete understanding of the fundamental conformational dynamics of DNTs and of their interactions with ions and water, which could aid biomimetic design of novel features into such pores. Thus, I have performed all-atom MD simulations of a synthetic 6-helix DNT structure that has previously been shown experimentally to form a scaffold for ion permeable nanopores in membranes (Figure 3.1-3.2)[57].

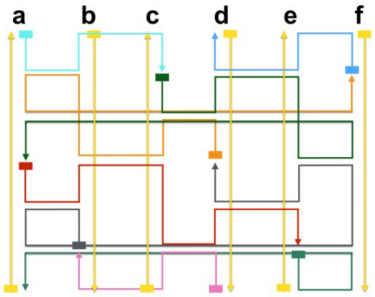

Figure 3.2: **Modelling the 6-helix DNT structure**. 2D representation of the DNT structure displaying the crossover points between the six dsDNA helices. The six scaffold strands are in golden-yellow, whilst the eight staple strands are in other colours. Figure adapted from Maingi *et al.*[91]

3.2) Methods

3.2.1) Model building

The all-atom initial DNT model was generated based on DNA (B-form) geometry. To automate the building procedure I used scripts written in PERL that called the *tleap* module in AMBER[105]. First, I generated six DNA double-helices using the nucleic acid builder (NAB)[123] utility in AMBER and the exact base sequence as used by Burns *et al.*[57] Six DNA double-helices were placed in a hexagonal geometry with their helical axis parallel with respect to each other, and hexagonal edge lengths of 1.8 or 2.0 nm (see Table 3.1 also). Each helix was then rotated with a major angle (between vectors pointing the P atom of joining residue and

along the hexagonal edge originating from centre of DNA double-helix; step 2 in Figure 3.3A) in such a way that bases involved in crossover points were approximately the nearest to each other. To further obtain the closest possible distance between joining residues in adjacent helices, each helix was further rotated (axis rotation) ±2.5° with 1.0° increments. This generated '*N'* number of configurations, and in general it can be written as:

$$N = \left(\frac{\theta}{d\theta}\right)^{h}$$

where 'θ' is the total angle scanned (5.0° in this case), '$d\theta$' is the small increment to complete the scan (1.0° in this case) and 'h' is the number of helices (6 in this case).

Thus, in total I obtained 15625 configurations, and finally the configuration that provided the least average distance and lowest range (calculated over all the joining residues in that configuration) was chosen as the best structure. Lastly, crossover points were created by joining the bonds between the corresponding bases to obtain the initial model as shown in Figure 3.3B. These steps are summarized briefly in the form of a flow chart shown in Figure 3.3A. In the final model the nucleotide sequences of 14 ssDNA and the number of crossover points were same as designed and assembled experimentally[57]. In this study of the DNT *per se* the hydrophobic belt modification was not included. Hydrophobic modifications were done on membrane-embedded DNT models detailed in Chapters 4 and 5.

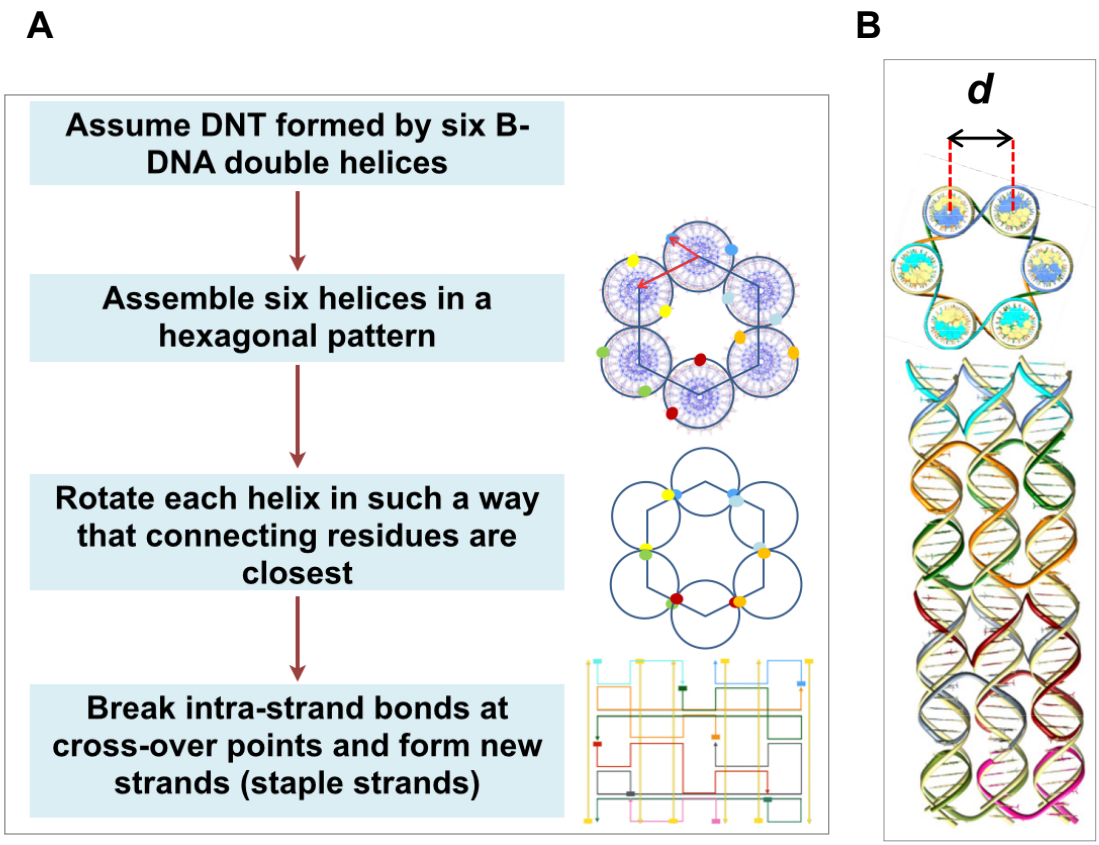

Figure 3.3: **Model building. (A)** Schematic flowchart for the building of the initial all-atom DNT model. **(B)** Initial hexagonal bundle model for DNT. The distance *d* between adjacent helices was either 1.8 or 2.0 nm. The DNA strands are drawn using the same colour scheme as in Figure 3.2. Figures **A** and **B** have been adapted from Maingi *et al*.[91]

3.2.2) Simulations

To check the robustness and dependence of the built DNT model on the chosen forcefields or inter-helix distances, I tested a few independent MD simulation cases. Also, in general, it is usually helpful to compare different forcefields to check, if any, anomalies associated with a particular forcefield for a test system.

The four simulations are summarized in Table 3.1. The AMBER12 CUDA version[131,132] (http://ambermd.org/) was employed for DNT1, DNT2 and DNT3 simulations using the AMBERff99+parmbsc0 force field[110]. The simulation DNT4 was performed using GROMACS 4.6.5[99] (www.gromacs.org) and the CHARMM36 force field[111] in order to compare and assess the results using different force fields. The initial DNT1 (d = 2.0 nm) and DNT2 (d = 1.8 nm) models were solvated with a 1.2 nm buffer of water molecules using the TIP3P

model[133] in a truncated octahedral box with KCl[134] at a concentration of 1 M relevant to experiments[57]. Truncated octahedral box is more 'spherical', and thus requires less number of water molecules in contrast to the cubic box. Additional K^+ ions were added to neutralize the phosphate groups of the DNT backbone.

Model	**Initial Hexagon Length *d*** (nm)	**All-atom Force Field**	**Temperature Control Scheme**	**Simulation Time** (μs)
DNT1	2.0	AMBER parmbsc0	Berendsen	0.17
DNT2	1.8	AMBER parmbsc0	Berendsen	0.165
DNT3	1.8	AMBER parmbsc0	Langevin	0.115
DNT4	2.0	CHARMM36	Langevin	0.5

Table 3.1: Description of the all-atom MD simulations performed to obtain the four DNT models. Table adapted, with different units of measurement, from Maingi *et al*.[91]. The simulation time in DNT4 case was extended (*ca.* three times than other cases) considering one of the questions raised by the reviewers for my ACS Nano paper[91].

The solvated systems were subject to steepest descent followed by conjugate gradient minimization with harmonic constraints applied to the position of the DNT atoms. The energy-minimization was repeated several times, progressively lowering the force constants from ca. $2x10^5$ to $8x10^3$ kJ mol^{-1} nm^{-2} (that is 500 to 20 kcal mol^{-1} $Å^{-2}$ in AMBER units). The systems were then slowly heated while maintaining weak harmonic constraints on the DNT atom positions. At each temperature between 0 K and 300 K MD simulation was performed in the NVT ensemble. The temperatures were kept constant using a Berendsen thermostat[106] and a relaxation time of 0.5 ps. A time step of 2 fs was used to integrate the equations of motion, with the SHAKE algorithm. During this heating process, inter-helix interactions were maintained using the EMAP method[135] which helped to maintain the internal geometry while at the same time allowing inter-helix interactions. This slow relaxation was performed over 1 ns. This strategy successfully removed strong electrostatic repulsions between helices from the initial hexagonal DNT models. The heating was followed by 20 ns of MD simulation in the NVT ensemble at 300 K with weaker harmonic positional

constraints of 4184 kJ mol^{-1} nm^{-2} (that is 10 kcal mol^{-1} $Å^{-2}$ in AMBER units) applied to the DNT atoms. This allowed equilibration of the positions of ions around the DNT surface. The harmonic constraints were then slowly removed by consecutive runs of 1 ns in the NVT ensemble (force constants of 4184, 2092 and then 418.4 kJ mol^{-1} nm^{-2}). Finally the extended, unrestrained MD simulation was performed in the NPT ensemble with a Berendsen thermostat and barostat[106] with heat bath coupling time of 1 ps and pressure relaxation time of 2 ps. A cut-off of 0.9 nm was used to treat short-range van der Waals interactions. The particle mesh Ewald (PME) method[136] was used for electrostatic interactions with a real-space cut-off of 0.9 nm, an order 4 cubic B-spline interpolation and direct space cutoff tolerance set being 10^{-4}.

While I was simulating some DNT models, one study by Yoo *et al.*[89] was published in which Langevin dynamics was used using NAMD software. As this protocol was reported and tested one, I wanted to compare results of my simulation protocols, using Berendsen thermostat, with Langevin thermostat in AMBER and GROMACS MD engines. Though various other suitable combinations of thermostats, barostats, and other parameters (type of ions and their available parameters) could be tested but these were not employed due to limited computational resources. Thus, for DNT3 a slightly different protocol was adopted. The DNT2 structure, obtained after the EMAP step described above, was subjected to the same subsequent NVT and NPT simulations but using a Langevin thermostat with collision frequency 1 ps^{-1}. For the DNT4 simulation, the DNT1 structure obtained after the EMAP step was used but switching to the CHARMM36 forcefield. The DNT4 model was solvated with a 2.4 nm buffer of TIP3P water molecules in a truncated octahedral box. The simulation protocol was the same as for the DNT1 simulation but with a Langevin thermostat. Final snapshots obtained after unrestrained NPT simulations are shown in Figure 3.4 for all the models.

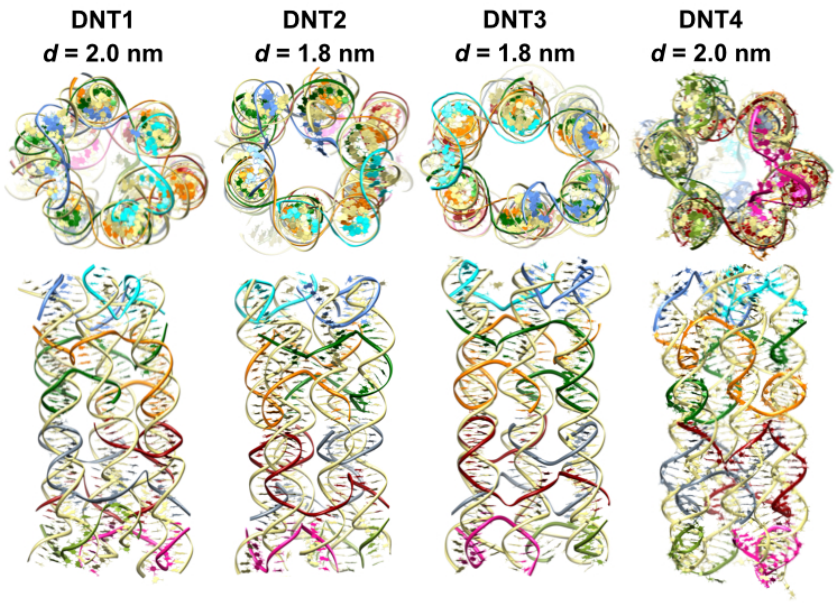

Figure 3.4: Final snapshots showing structures of all the models after unrestrained MD simulations. See Table 3.1 for summary of simulations. The snapshots are adapted from Maingi *et al*.[91] The *d* distances refer to the initial built model interhelix spacing.

3.3) Results and discussion

3.3.1) Structure and stability of the DNT models

The DNT used in my studies features a 6-helix scaffold composed of 14 strands same as in the experimental design[57] (Figure 3.2). Following the principles of DNA origami, six oligonucleotides (Figure 3.2, golden-yellow) run parallel to the duplexes while eight strands form short inter-duplex cross-overs to ensure the cohesion of the nanostructure. The two initially built models (d = 1.8 or 2.0 nm) were used as starting structures for my MD simulations, ranging in duration from ca. 0.15 to 0.5 μs (Table 3.1) of the DNT immersed in a 1 M KCl solution. In these four simulations, I varied the simulation conditions in order to explore the robustness of the conclusions derived from the subsequent analysis.

Analysis, explained further below, of the no-H atom root mean square deviations (RMSDs) suggested that all four simulations equilibrated after 60 ns, to an RMSD relative to their respective average structures of ca. 0.4 nm (Figure 3.5). First, the general procedure followed to calculate the RMSD is mentioned, which is later explained with the DNT1 case as an example. RMSDs were calculated for all the

simulated models with respect to their initial configurations, Figure 3.5A. It can be seen that all models show high RMSD values, which is quite expected, as the initial atomic coordinates were not obtained from X-ray, NMR or any other experimental data. This was also observed for similar structures as reported by Joshi *et al.*[90] and Yoo *et al.*[89]. These RMSD plots helped to identify the part of trajectory suitable to obtain the average structures represented as: <DNT1>, <DNT2>, <DNT3>, and <DNT4>. In order to present the results in a simplified manner, all averaged structures were obtained from the length of trajectory after 60 ns. Furthermore, these averaged coordinates were used as reference structures for their respective trajectories to calculate RMSD as shown in Figure 3.5B. This helped to understand how each model simulation converges to their respective average structures. It is especially useful in my case as I lack any experimental structural details, which could otherwise be used as a reference structure. Such plots are helpful to remove the part of the trajectories which can otherwise lead one to interpret wrongly any analysis meant to be performed on the equilibrated system. To illustrate this procedure the case of DNT1 is presented. Figure 3.5A suggests that DNT1 equilibrates after ca. 0.045 μs, but RMSD calculated from representing average structure (0.045 to 0.17 μs) indicated that an equilibrium is achieved after 0.06 μs (not shown for clarity). Also, in the DNT1 case and other cases, it was found that RMSDs calculated using an average structure obtained from 0.06 to 0.17 μs converged in a similar fashion (after 0.06 μs) as shown in Figure 3.5B. Therefore, I ignored the initial 0.06 μs for all cases to present my results in a simplified manner. Thus, unless stated all further analyses were performed on the length of trajectory after 0.06 μs (60 ns) for all cases; Figure 3.5B.

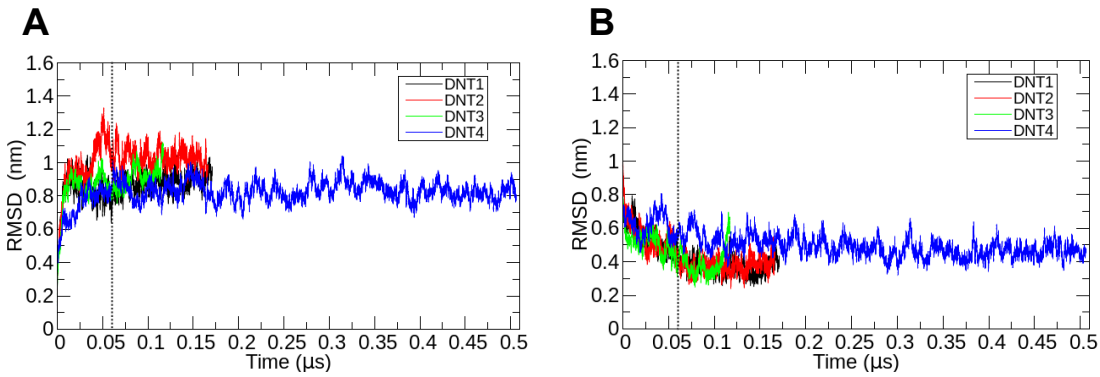

Figure 3.5: **RMSD all-atom plots (hydrogen atoms excluded)**. **(A)** Root mean square deviation (RMSD) of the non-hydrogen atom positions with respect to the starting configuration over the four simulations. An average structure was calculated in each case considering all the frames after 0.06 μs (vertical dotted line). **(B)** RMSD plots for the four models with respect to their average structure. The vertical dotted line was considered as the start of the equilibrated part of the trajectories for further analysis (*i.e.* after 0.06 μs). Figures **A** and **B** adapted from Maingi *et al*.[91]

In order to evaluate how these models compare with respect to each other, RMSD for each simulation was calculated with four different averaged structures. For example, in Figure 3.6B it can be seen that for the DNT1 case the RMSD calculated against <DNT1> has lower values compared to the other cases (DNT2–DNT4). In order to understand the origin of this difference, RMSD was calculated on local parts (G1 to G6) of the structure (Figure 3.6A).

Preliminary analysis showed that different subgroups present in DNT have their own influence on overall RMSDs calculated with respect to average structures. Thus, for each case RMSDs were calculated on different subgroups as: G1, G2-G5 combined and G6 using average structures as reference. For example, in Figure 3.6C, using <DNT1> the RMSD for DNT2(G2-G5) is closer to DNT1(G2-G5), and the major difference in overall DNT2 RMSD relative to DNT1, as observed in Figure 3.6B, arises due to DNT2(G1) and DNT2(G6) subgroups. This was also observed for the DNT3 case with <DNT1> as shown in Figure 3.6D. For the DNT4 case, with the CHARMM36 force field, with <DNT1> significant difference between DNT4(G2-G5) and DNT1(G2-G5) was observed, Figure 3.6E, which may be due to different force fields used in these cases. Remarkably, with the CHARMM36 case also, G1 and G6 show major fluctuations as observed in

DNT2 and DNT3 cases. Similarly, this behaviour was observed for <DNT2> (with DNT1, DNT3, DNT4 simulations), <DNT3> (with DNT1, DNT2, DNT4 simulations), and <DNT4> (with DNT1, DNT2, DNT3 simulations) which are shown in Figures 3.7, 3.8 and 3.9 respectively.

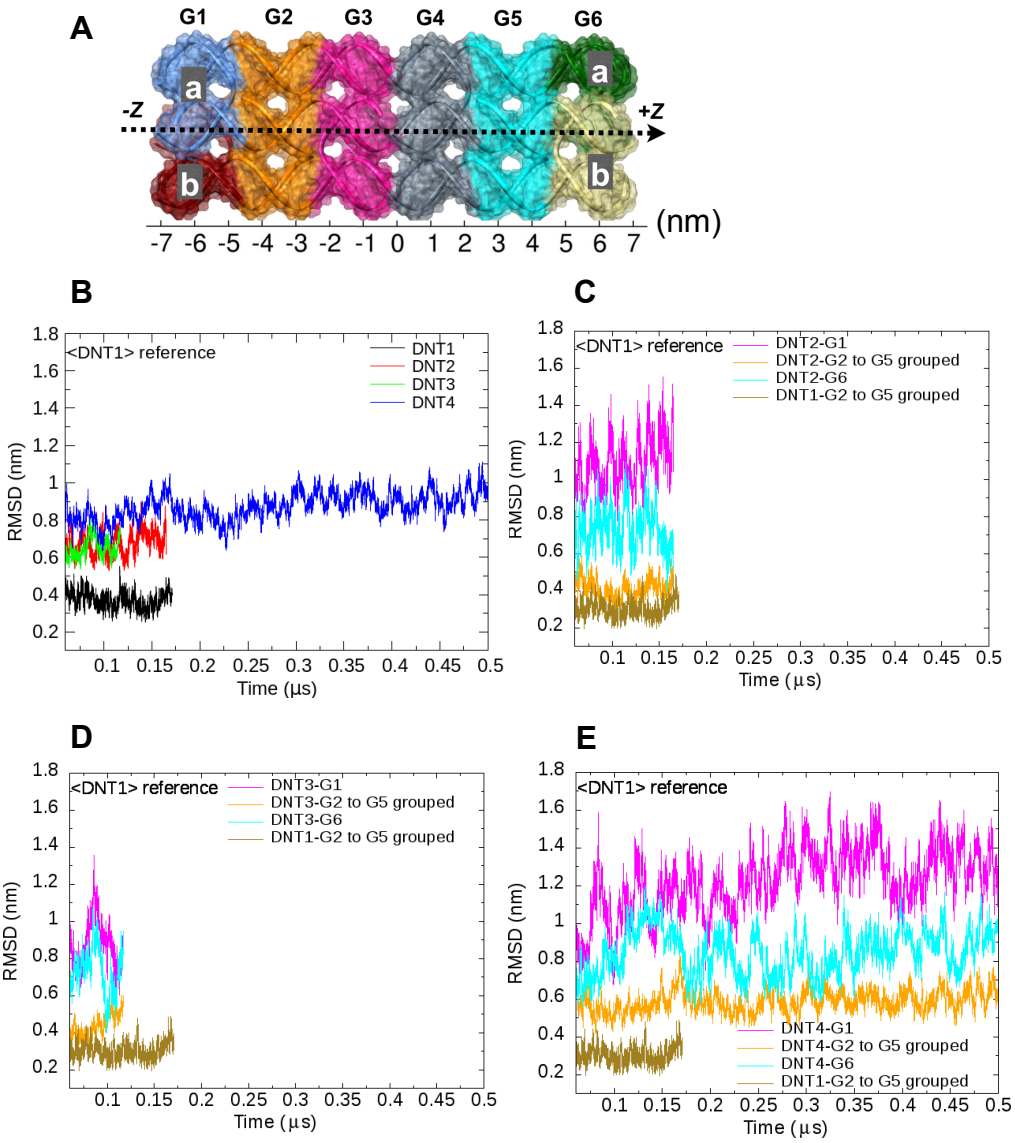

Figure 3.6: **(A)** Definition of the six regions of the DNT, from G1 to G6, with an equal length of seven bases along the pore axis. These regions are based on crossover points where the folding of staple strands occurs. RMSD plots with <DNT1> reference **(B)** for all cases considering full structure, **(C-E)** for all cases considering various subgroups G1, G2-G5 combined and G6. Figure **A** adapted from Maingi *et al.*[91]

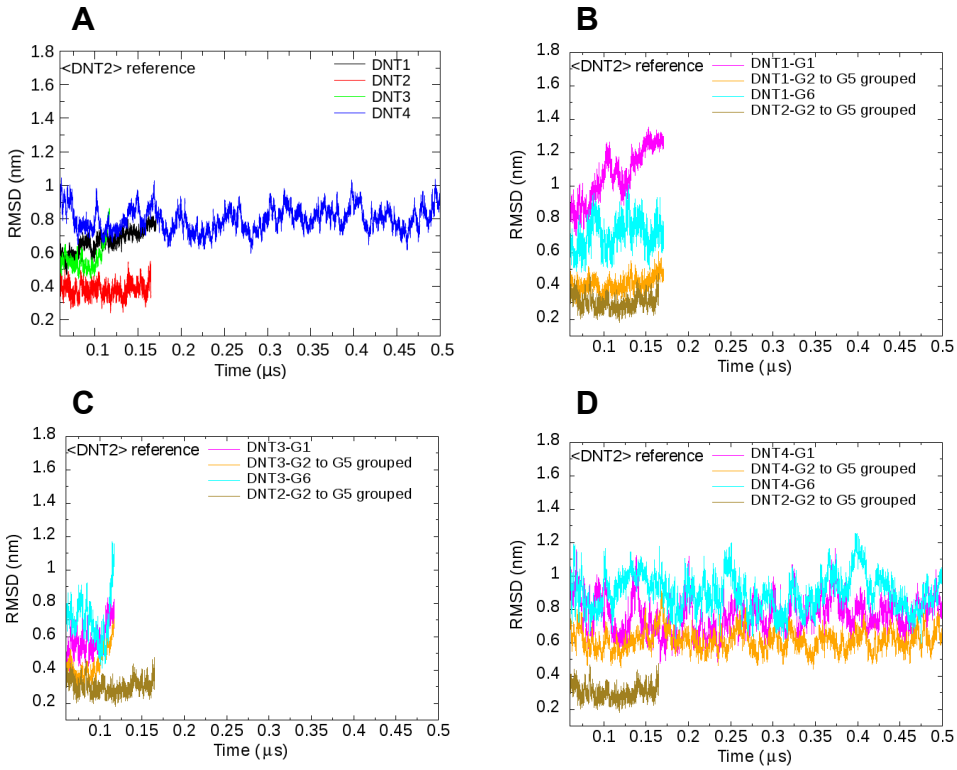

Figure 3.7: RMSD plots with <DNT2> reference **(A)** for all cases considering the full structure, **(B-D)** for all cases considering various subgroups G1, G2-G5 combined and G6.

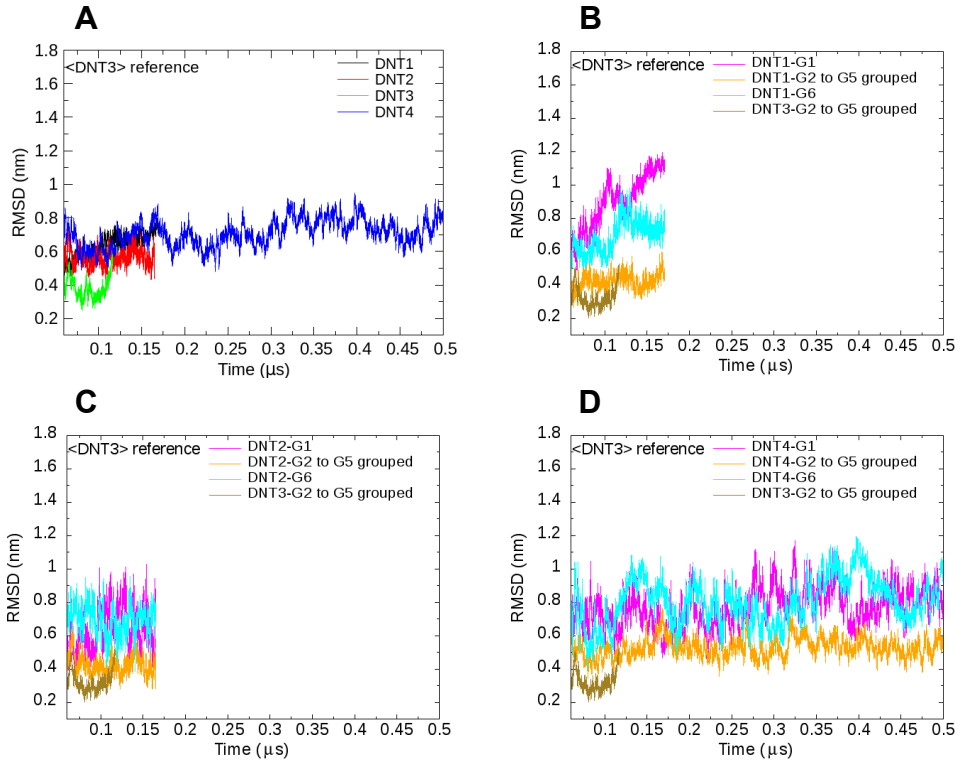

Figure 3.8: RMSD plots with <DNT3> reference **(A)** for all cases considering the full structure, **(B-D)** for all cases considering various subgroups G1, G2-G5 combined and G6.

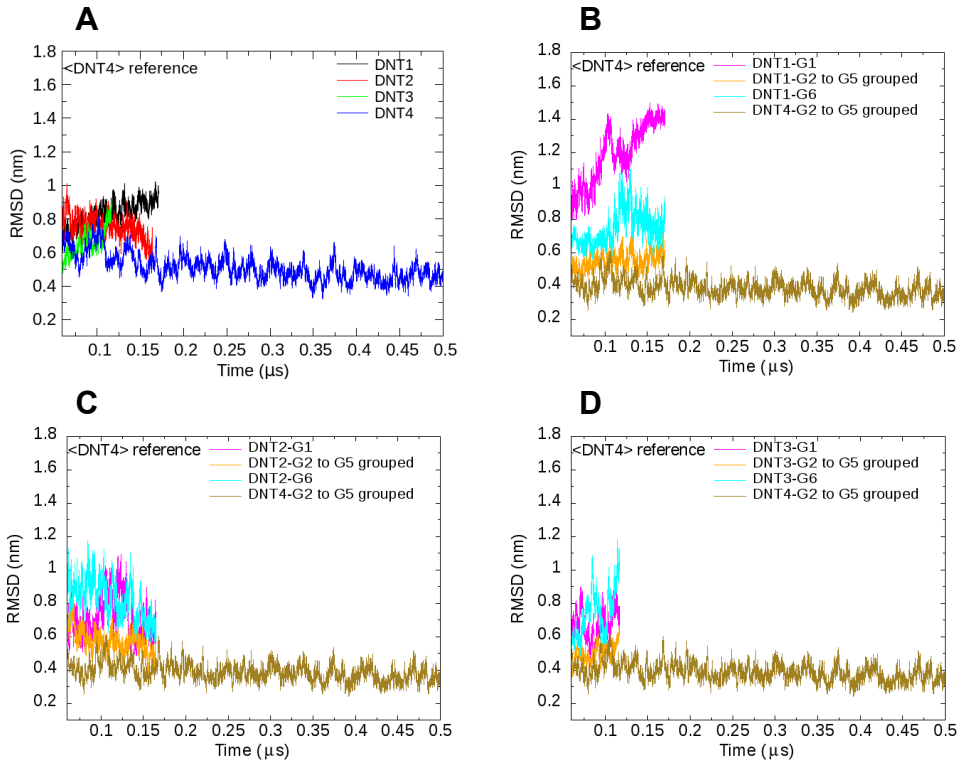

Figure 3.9: RMSD plots with <DNT4> reference **(A)** for all cases considering the full structure, **(B-D)** for all cases considering various subgroups G1, G2-G5 combined and G6.

Apart from RMSD analyses, structural stability of the models was also confirmed by analysis of hydrogen bonds (HBs) between the base pairs within the DNT structure (Figure 3.10). In case of DNT4 it can be seen that number of broken HBs (Figure 3.10A) keeps on increasing with time, and in rest of the cases during initial 0.04 μs there is an increase in fraction of broken HBs which remain stable afterwards. I was interested to explore more about these observations. As in the case of RMSD analyses I calculated the fraction of broken HBs in each subgroup, and also considering residues at the junctions. It can be clearly seen that in all the cases maximum HB disruption occurs in G1 and G6 subgroups (also to some extent in DNT2-G3 and DNT4-G5), and the number of disrupted bonds keeps on fluctuating (similar observation for G1 and G6 in the RMSD fluctuations discussed above). This can be considered as a fraying motion occurring in the DNA due to HB breaking. Thus, the gradual increase in broken HB in DNT4 is mostly contributed by G1 and G6 subgroups (almost goes upto 50% of HB in these subgroups). As other cases, DNT1-3, were not simulated as long as the DNT4 case it is difficult to conclude if there is any anomaly related to different

force-fields (AMBER for DNT1-3, CHARMM for DNT4). Fraction of broken HB in G2-G5 subgroups in all cases remains more stable from the initial stage of simulation. Also, HBs associated with the junction residues remain stable in all the cases.

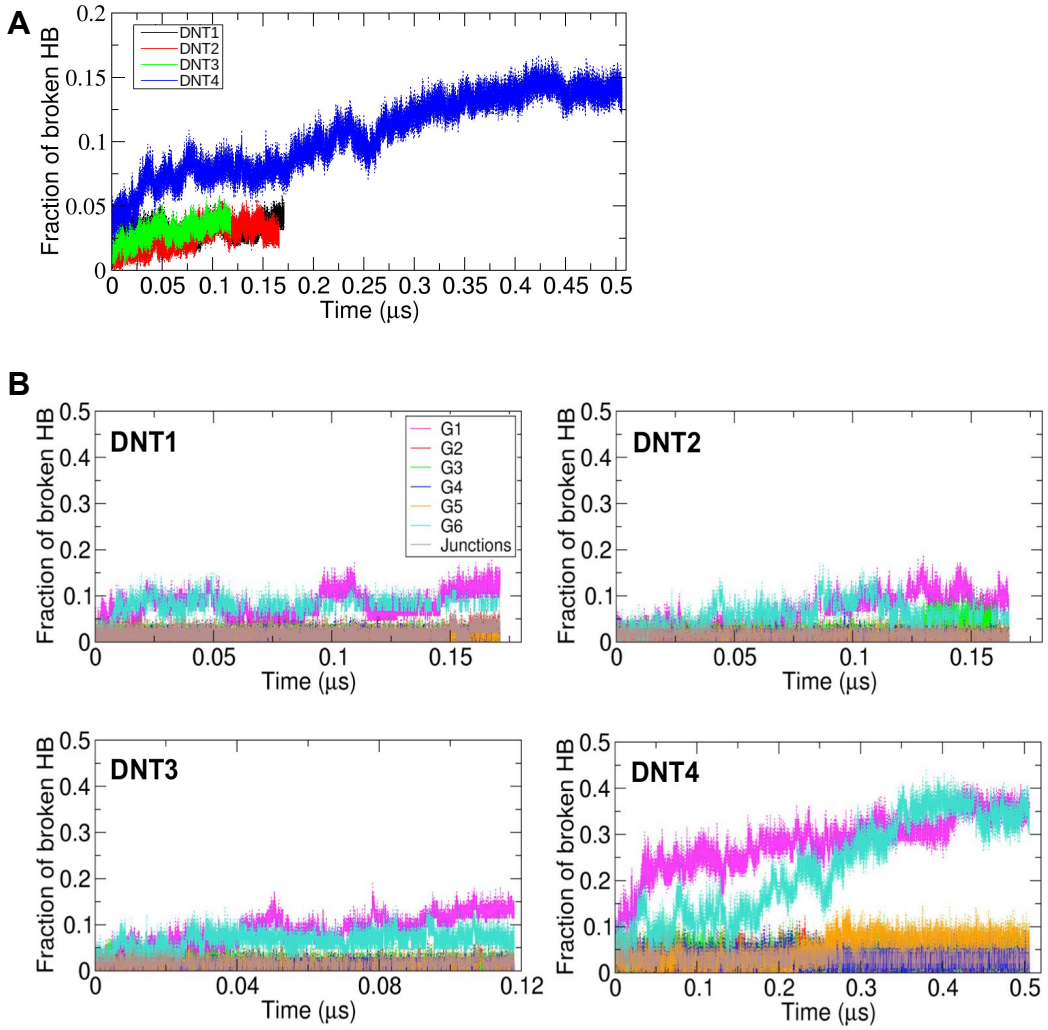

Figure 3.10: **A-T, G-C Hydrogen bond analysis. (A)** Fraction of broken hydrogen bonds (HBs) shown with respect to the time of simulations. The maximum fraction of broken HBs was seen in DNT4 (with the CHARMM36 force field). Nonetheless, in all cases the maximum number of broken HBs is reached during the first 0.04 μs, which was perhaps expected for the relaxation of the theoretical model. Joshi *et al.*[90] and Yoo *et al.*[89] have observed similar behaviour. **(B)** Fraction of broken HBs with respect to time calculated in each subgroup (G1 to G6) and junctions separately; shown for DNT1-4 cases. Note *X*-axis time scale is different in each case. The calculated HBs follow the criteria defined by Olson *et al.*[137] Figure **A** adapted from Maingi *et al.*[91]

I also analysed the significance, if any, of initial inter-helix spacings in our different simulated models. The initial models used in this work were based on a hexagonal arrangement of double helices and the distance between the central axes of any two adjacent helices was chosen to be 2.0 nm (DNT1, DNT4) or 1.8 nm (DNT2, DNT3). Thus, it was of interest to know the change which occurred in six distances during the course of simulation. For this, the distances were calculated between the projected center of masses (CoM) of G2-G5 P atoms in adjacent double helices. The P atoms of G1 and G6 were ignored as their RMSF profiles were found to be highly dynamic (discussed below). In total, six distances were calculated which are represented as **a-b**, **b-c**, **c-d**, **d-e**, **e-f**, **f-a** in Figure 3.2. It can be seen in Table 3.2 that the corresponding distances converge to almost the same lengths regardless of the initial distances (1.8 or 2.0 nm). All distances fall between 2.3 to 2.5 nm and for **e-f** it is around 2.5-2.6 nm but, again, consistent in all the cases. Thus, in all four simulations, the inter-helix spacing converged to $d = 2.4 \pm 0.1$ nm, and the hexagonal shape of the central nanopore was retained roughly.

So, in the subsequent analyses I will focus mainly on simulation DNT1.

Distance	**DNT1** (2.0 nm)	**DNT2** (1.8 nm)	**DNT3** (1.8 nm)	**DNT4** (2.0 nm)
ab	2.36 ± 0.08	2.35 ± 0.08	2.33 ± 0.1	2.37 ± 0.11
bc	2.39 ± 0.07	2.39 ± 0.11	2.42 ± 0.07	2.51 ± 0.1
cd	2.34 ± 0.08	2.38 ± 0.08	2.42 ± 0.08	2.38 ± 0.11
de	2.45 ± 0.07	2.47 ± 0.09	2.45 ± 0.08	2.43 ± 0.09
ef	2.48 ± 0.09	2.54 ± 0.08	2.59 ± 0.11	2.62 ± 0.15
fa	2.34 ± 0.07	2.36 ± 0.07	2.36 ± 0.07	2.27 ± 0.09
Average	2.39 ± 0.05	2.42 ± 0.07	2.43 ± 0.08	2.43 ± 0.01

Table 3.2: Average values and standard deviations of the six adjacent helix-bundle distances, in comparison to the initial hexagonal length (1.8 or 2.0 nm) of the four models. Table adapted, with different units of measurement, from Maingi *et al*.[91]

3.3.2) Conformational dynamics of DNTs

Understanding the conformation and dynamics of DNTs is of interest in the context of designing DNT-based nanopores of defined structural stability or flexibility. To provide an overview of DNT dynamics, I calculated the root mean square fluctuation (RMSFs) of the phosphorous atom positions, relative to the

average structures over the equilibrated part of the simulations (i.e. after 0.06 μs). In all the simulated models I observed low RMSF values in the central section of the DNT (regions G2 to G5) compared to those calculated for the mouth regions, G1 and G6, of the nanopore (Figure 3.11).

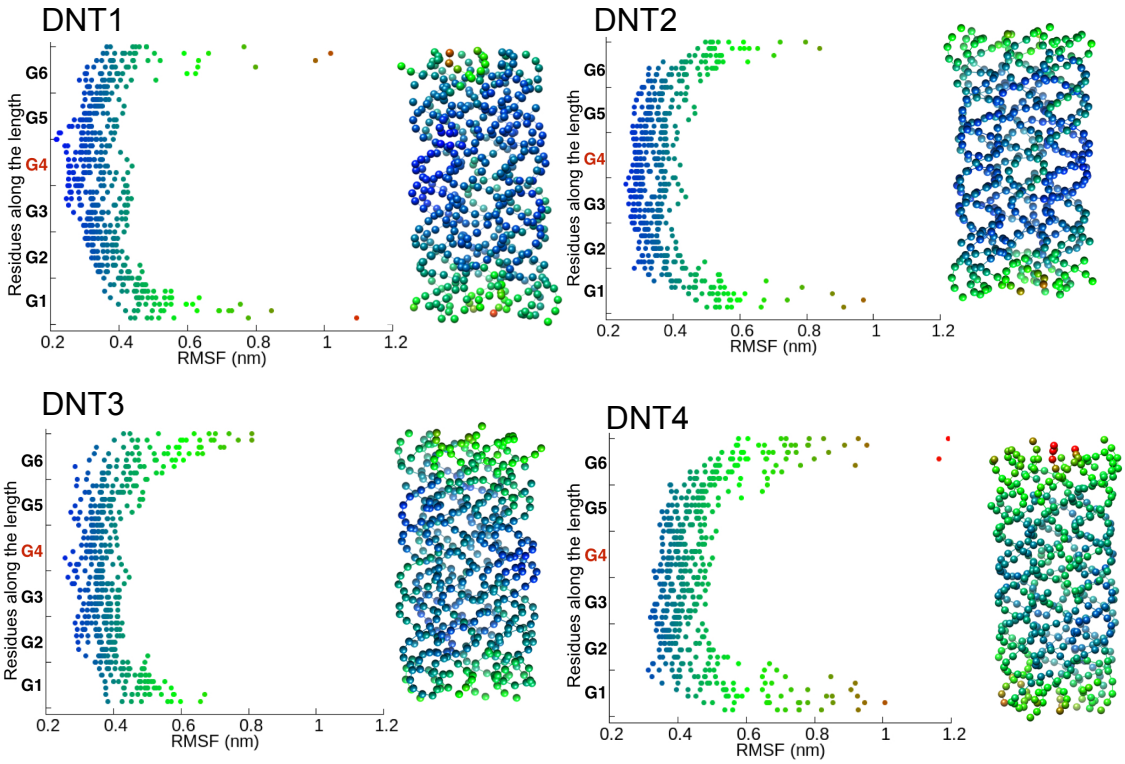

Figure 3.11: **RMSF distribution.** Root mean square fluctuation (RMSF) values of phosphorous atoms along the pore axis of the DNT1, DNT2, DNT3 and DNT4 models. The DNT length consists of 42 bases shown along *Y*-axis as subgroups G1 to G6 (refer Figure 3.6A). Also shown along each RMSF plot (right hand) average structures from each simulation, represented as phosphorous atoms coloured using the same RMSF scale in their corresponding plots. This figure has been modified from Maingi *et al*.[91]

Phosphorous atoms at the junctions, mainly located between G2 and G5, did not show large local motions (Figure 3.12), indicative of overall nanopore stability. In particular the G4 region, where chemical modifications have been made to provide a hydrophobic belt (in the experimental study)[57], showed relatively small fluctuations.

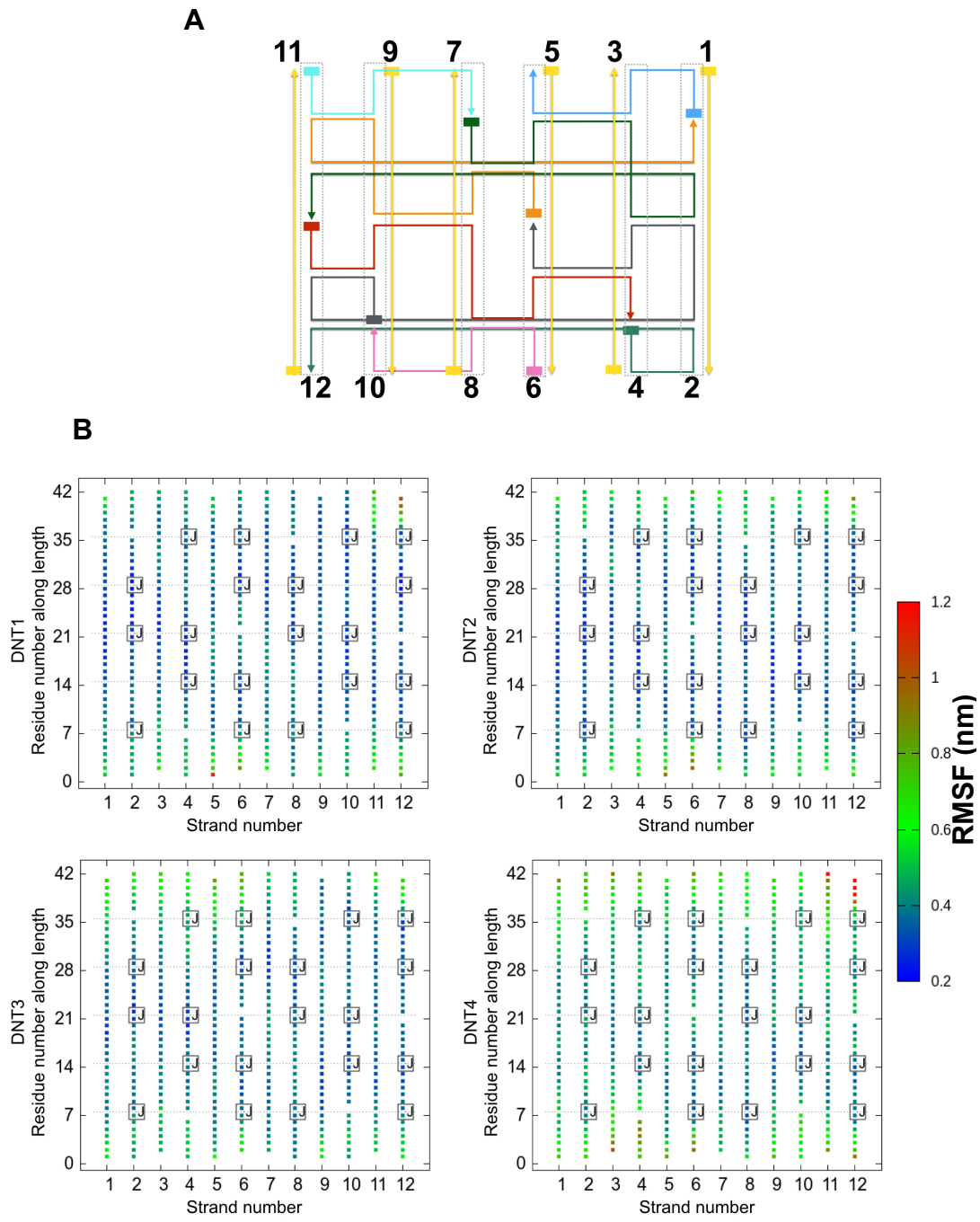

Figure 3.12: **Phosphorous atom RMSF patterns (focus on cross-junctions)**. **(A)** DNT 2D map with odd numbers indicating straight scaffold strands and even number represents assumed single straight strands along which RMSF was calculated. **(B)** RMSF pattern of all the P atoms present in G1 to G6 subgroups. Each colored square represents a phosphorous atom; a missing square means 5' end of staple/scaffold strand; 'J' labels represent cross-over junction. This figure has been adapted from Maingi *et al.*[91] but rescaled and given different units of measurement.

The flexibility observed in groups G1 and G6 is suggestive of motions which might result in possible gating at the mouths of the DNT. To explore this further, I performed principal component analysis (PCA) to identify the major collective motions of the DNT backbone observed during the simulations. The first eigenvector identified through this analysis (Figure 3.13) suggests a 'breathing motion' at the mouth regions of the DNT. Thus, the nanopore in regions G1 and G6 may successively expand and shrink. I identified two subgroups at the DNT termini that lacked staple strands, which could rationalize the motions observed by PCA. These subgroups were defined as G1a and G1b, and as G6a and G6b (Figure 3.14 and also see Figure 3.6A).

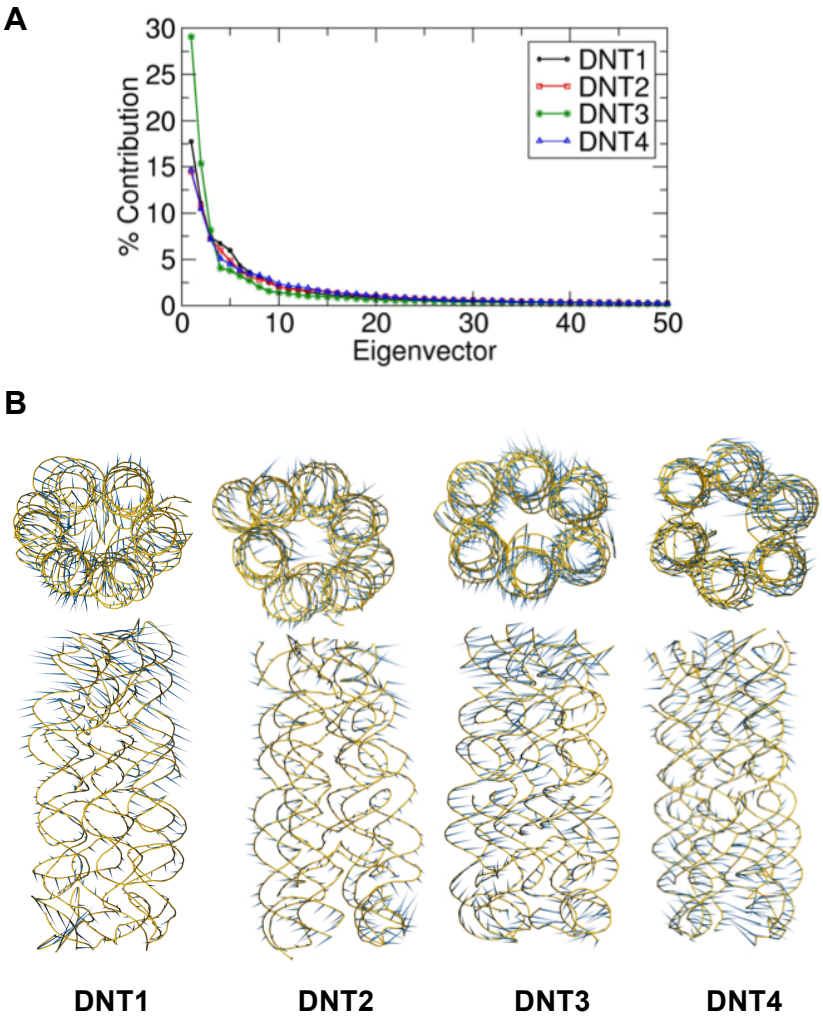

Figure 3.13: **Principal component analysis**. **(A)** Eigenvalue contributions in each model. **(B)** First eigenvector projections for DNT1, DNT2, DNT3 and DNT4 cases. Backbone structures with the P-atom trace are shown in yellow, with blue arrows indicating the eigenvector. Figures **A** and **B** have been adapted and slightly modified from Maingi *et al*.[91]

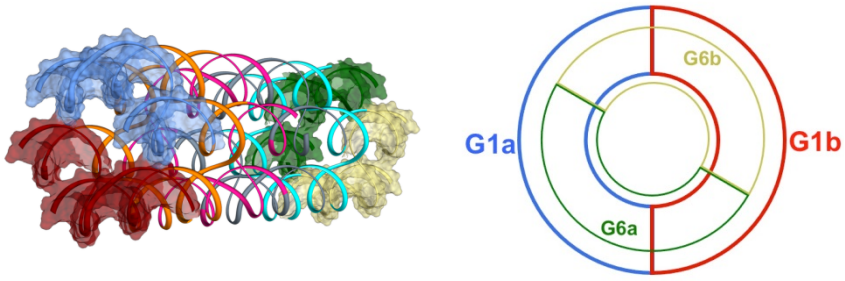

Figure 3.14: **Subgroups chosen for analysis**. A complementary view to Figure 3.6A alongside a simple cartoon with a circle diagram showing the approximate relative positions of the G1a-G1b and G6a-G6b flaps with same colouring scheme for subgroups. Figures **A** and **B** have been adapted and slightly modified from Maingi *et al*.[91]

I also explored possible correlations between motions at the two termini, i.e. between the G1a-G1b and the G6a-G6b distances (Figure 3.15 and 3.16). This analysis did not reveal any significant correlations, suggesting that the motions at the two ends of the nanopore are largely independent of one another.

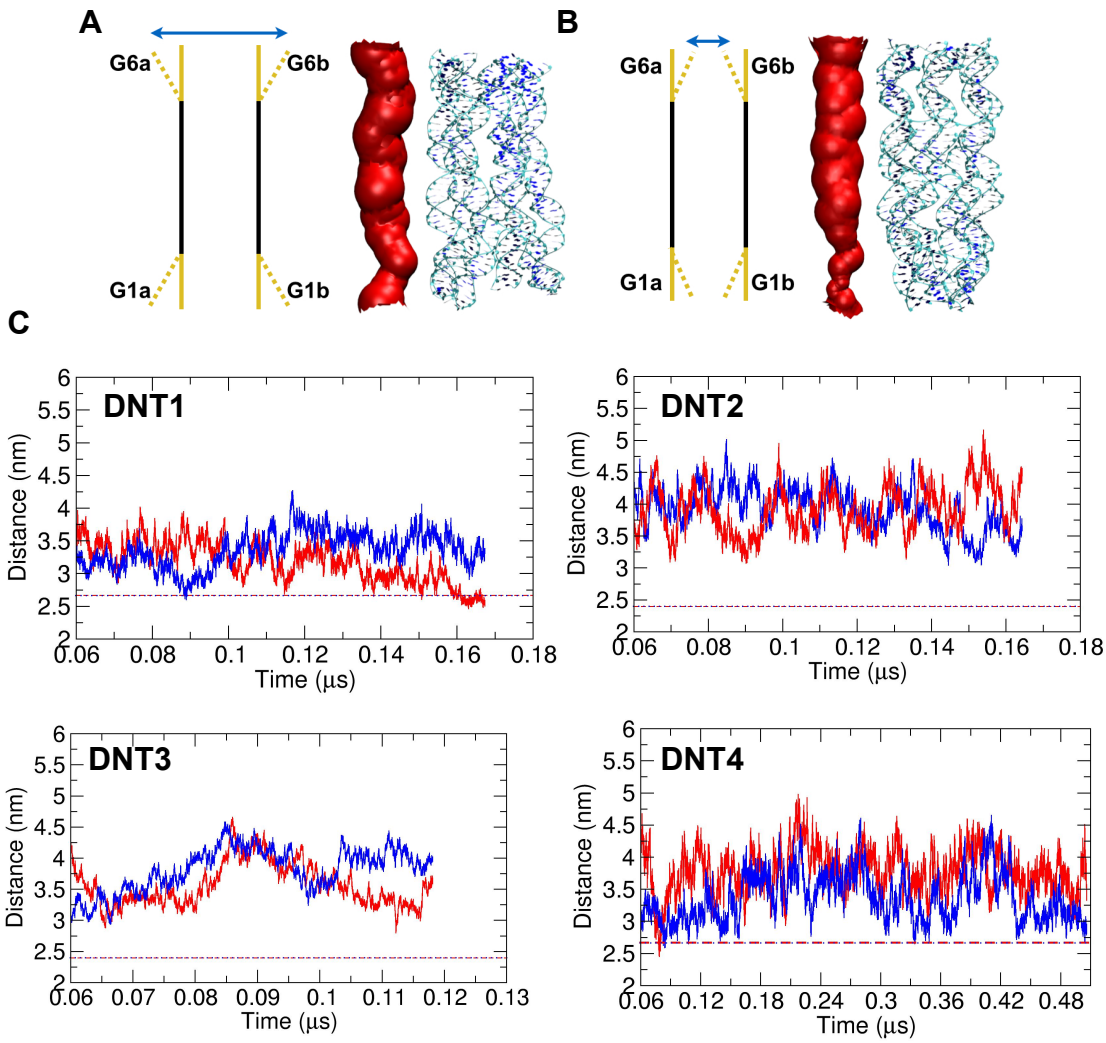

Figure 3.15: **Breathing motions of DNT termini**. **(A, B)** Simple cartoon showing open and closed states, respectively, along with a representative snapshot of the DNT1 structure and a surface view of the lumen represented in red. (**C**) Evolution of the G1a-Gb and the G6a-G6b distances represented in red and blue solid lines, respectively, as a function of the time of the DNT simulations. Distances were calculated between the centers of mass of G1a and G1b at one end and between the centers of mass of G6a and G6b at the other end. The horizontal dotted lines report the initial distances (Note they are on different scales; Time *X*-axis). Figures **A, B** and **C** have been adapted and modified from Maingi *et al.*[91]

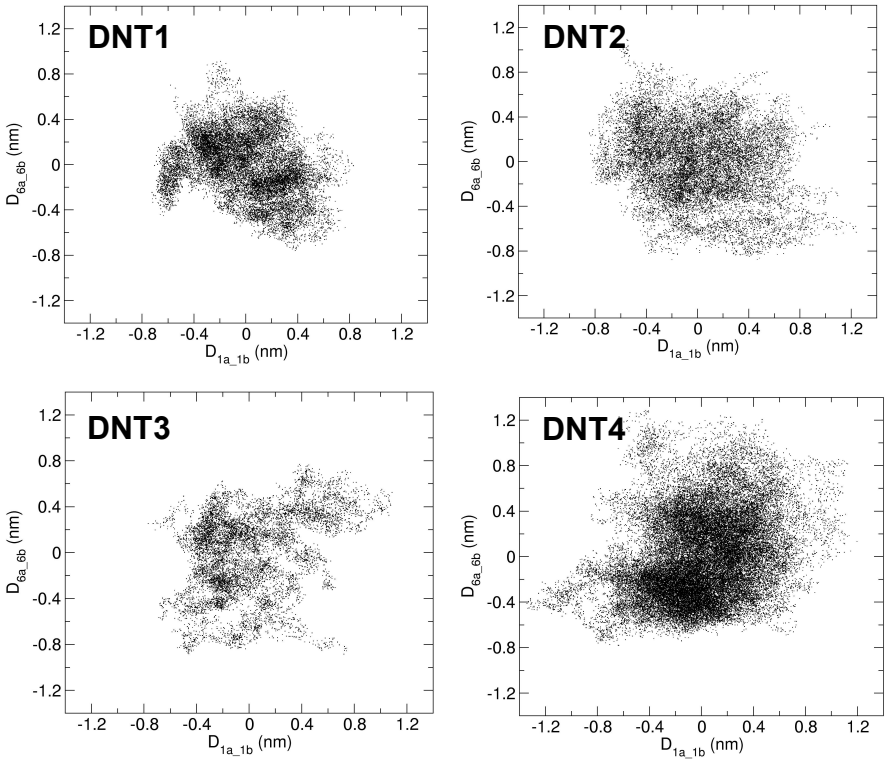

Figure 3.16: **Absence of correlation between the terminal breathing motions**. Scatter plots for DNT cases with *D6a_6b vs. D1a_1b* calculated as D_t - <*D*> for G6 and G1 subgroups where D_t is distance between two flaps in one subgroup at any time *t* (G1a and G1b, and G6a and G6b) and <*D*> is an average over time for the corresponding subgroup. Figure adapted from Maingi *et al.*[91]

3.3.3) Nanopore radius profiles

I examined the radius profile along the length of the nanotube (Figure 3.17) to follow how the DNT structure changes in the simulations as well as to quantify the consequences of the pore dynamics. To calculate the radius profiles I used HOLE[138]. In brief, along each plane, normal to the pore vector, it calculates the largest sphere that can be accommodated without overlapping the van der Waals radii of the atoms. The initial model showed a fairly flat pore radius profile, with an average pore radius of ca. 0.9 nm for DNT1 and DNT4 (d = 2.0 nm); Figure 3.17A, and of ca. 0.7 nm for DNT2 and DNT3 (d = 1.8 nm); Figure 3.17B. In all four simulations the central (regions G2 to G5) pore radius increased to ca. 0.1 nm, consistent with convergence on a equilibrium value of d = 2.4 nm for the inter-helix spacing. In all four simulations a degree of narrowing at one (G1) or other (G6) mouth of the pore was observed, consistent with the suggestions of

gating-like motions at the mouths of the pore discussed above. Thus, in DNT1 and DNT4, both starting from the d = 2.0 nm model, the average mouth radius dropped to ca. 0.6 nm at the G1 end and to ca. 0.8 nm at the G6 end respectively (Figure 3.17A). Similarly, this was observed at the G6 end of DNT2 and DNT3 cases. In order to characterize the gating motions at the pore mouths I also calculated the pore radius profile as a function of simulation time (Figure 3.17C).

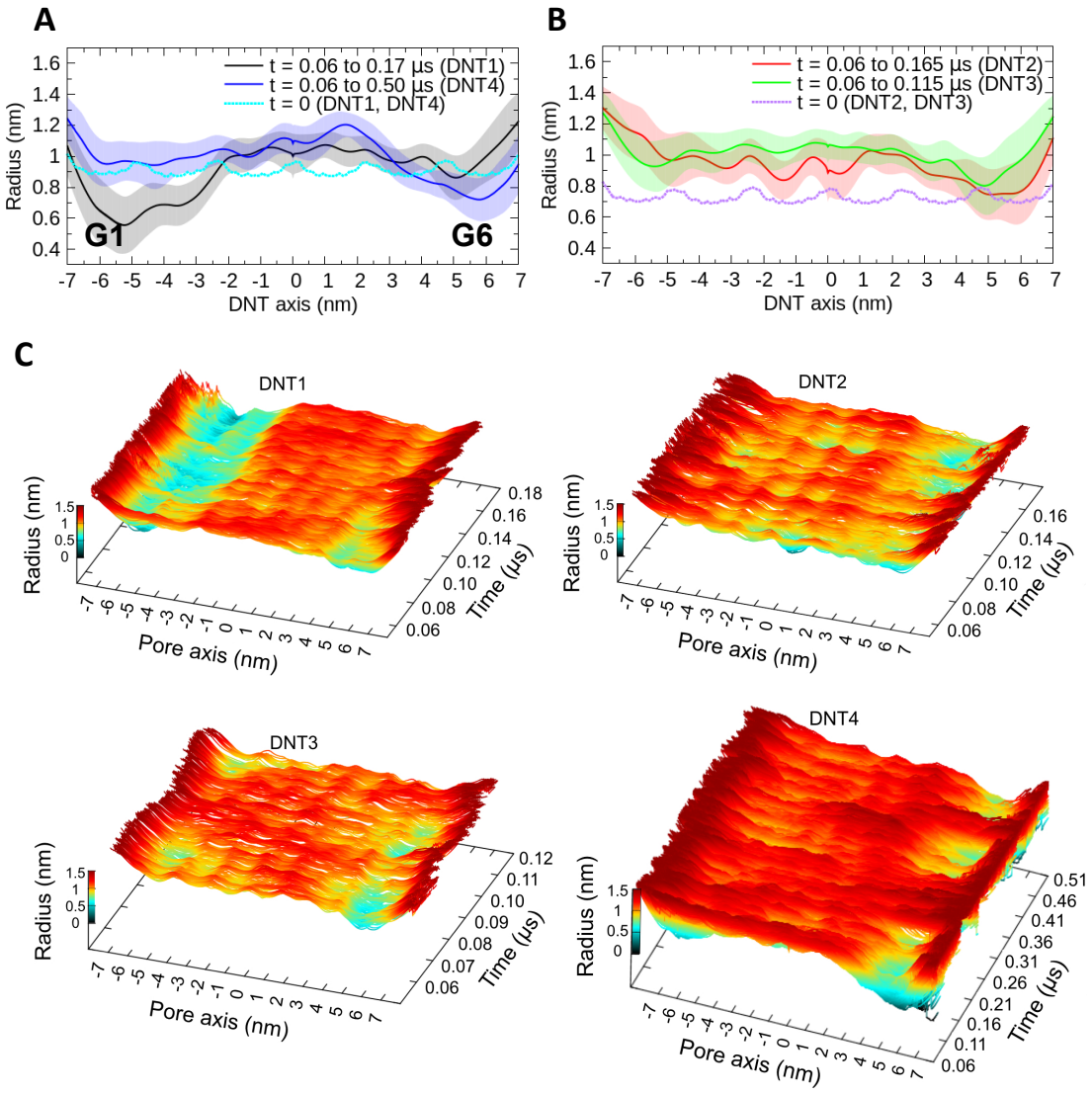

Figure 3.17: **(A, B)** Pore radius HOLE[138] profiles along the principal axis of the nanopore for the DNT1, DNT2, DNT3, and DNT4 simulations. The cyan and purples lines in **A** and **B** respectively shows the radius profile of the initial model of the system at t = 0. In **A**, the black and blue lines, and in **B**, the green and red lines represents the radius profile averaged over the equilibrated section of the simulations (t = 0.06 to 0.17 µs for DNT1 and t = 0.06 to 0.5 µs for DNT4, t =

0.06 to 0.165 μs for DNT2 and t = 0.06 to 0.115 μs for DNT3). The corresponding grey, blue, red and green bands report the standard deviations. **(C)** The pore radius profiles of the DNT1, DNT2, DNT3, and DNT4 simulations are shown as a function of time as a combined heatmap and surface plot. This was obtained by juxtaposing radius profile curves averaged within 100 ps windows. Figures **A**, **B**, **C** have been adapted and modified from Maingi *et al.*[91]

Detailed inspection suggested multiple time-dependent fluctuations in the pore radius. To further analyse such possible gating-like events I analysed the minimum pore radius as a function of time for both the G1 end (-7 to -3 nm) and the G6 end (3 to 7 nm) of the DNT (Figure 3.18).

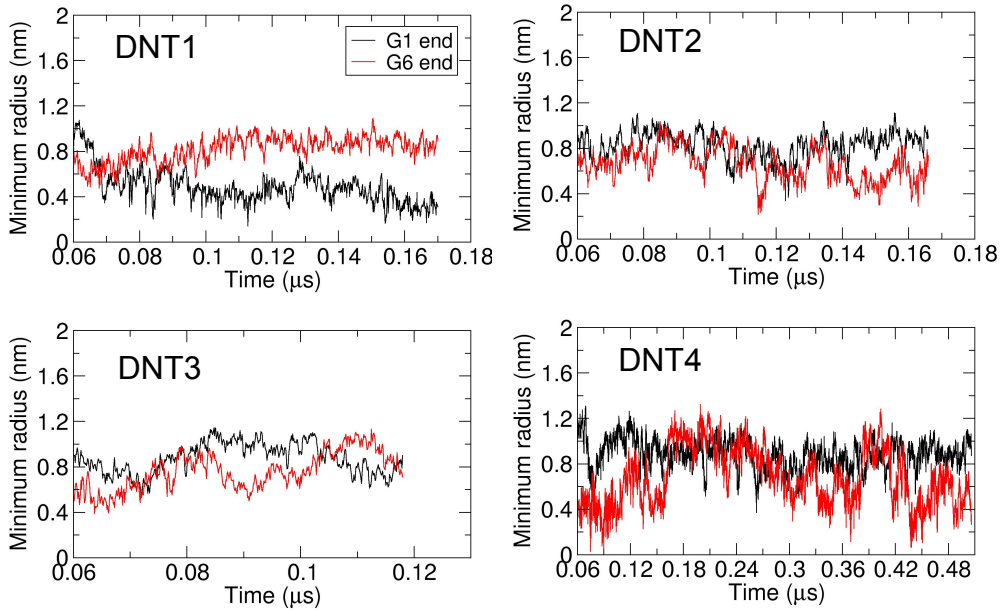

Figure 3.18: **Minimum pore radii at the two ends of the nanopore.** Time evolution of the minimum pore radius as a function of time for the G1 end (-7 to -3 nm) and the G6 end (3 to 7 nm) of the nanopore for the DNT1, DNT2, DNT3 and DNT4 simulations. Figure has been adapted and modified from Maingi *et al.*[91]

The above analysis (Figure 3.18) revealed fluctuations in these minimum radii, which I classified as either O (open-like) or C (closed-like) using a simple 0.6 nm threshold radius (based on consideration of the radii of hydrated ions). This simple analysis yields a signal-like trace of the transitions of the pore between the O and C states (shown in Figure 3.19), revealing multiple transitions during the course of the simulations. For example, from this analysis, for DNT1, I observe a dynamic narrowing/closure of the G6 gate which re-opens at ca. 0.1 μs and remains open

for the rest of the simulation. In contrast, the G1 gate is partly closed at the start and then more fully closes (and remains closed) after ca. 0.1 μs with short transient re-openings from 0.13-0.14 μs. For the longer (0.5 μs, Figure 3.15) DNT4 simulation, I observe multiple brief closures at the G1 mouth, alongside frequent stochastic switching between O and C at the G6 mouth of the nanopore. Analysis of the DNT2 and DNT3 simulations shows similar behaviour.

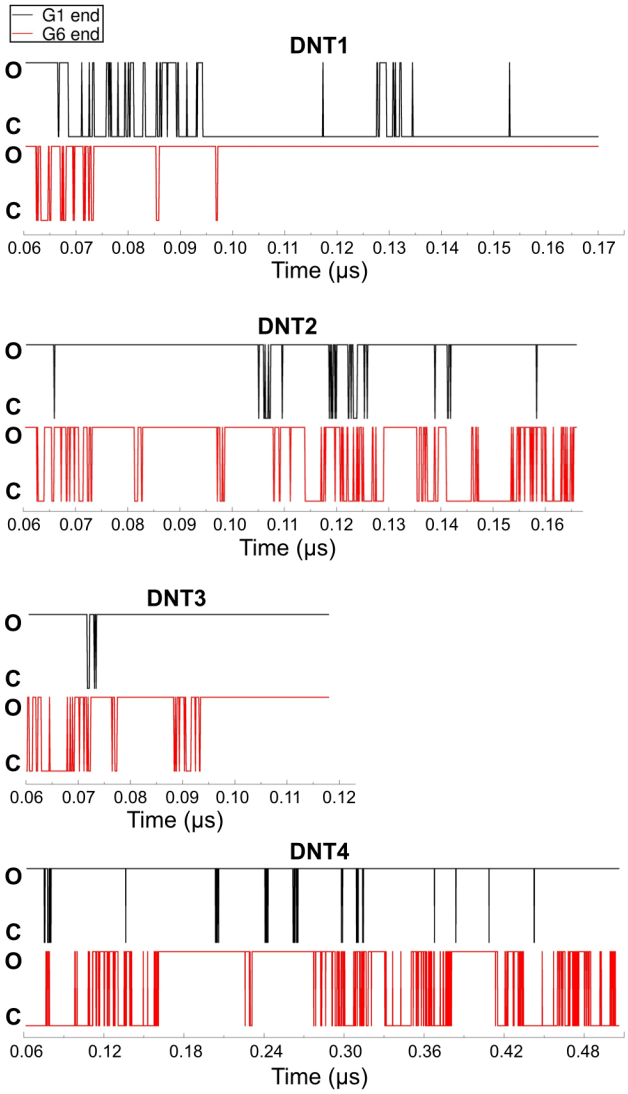

Figure 3.19: **Open or closed states**. Fluctuations in the minimum radii were used to classify the DNT as being either O (open-like) or C (closed-like) state, using a simple 0.6 nm threshold radius, yielding a signal-like trace of the transitions of the nanopore between the O and C states. Figure has been adapted and modified from Maingi *et al.*[91]

3.3.4) Ions and water in the DNT nanopore

The temporal evolution of the lumen dimensions offers an opportunity to investigate the effect on ions and water within the DNT. As discussed above

(Figure 3.17A), in the central lumen region the nanopore radius is *ca*. 1 nm and shows only small fluctuations. Consequently, there are no large temporal variations in the density of water or of ions in the lumen. In contrast, the mouths of the DNT are narrower and undergo dynamic changes over the course of the simulation. I therefore calculated the numbers of both ion species and of water molecules along the pore axis, focussing especially (in simulation DNT1) on the G1 mouth of the nanopore (Figure 3.20A). I observe that as the G1 mouth region narrows (after ca. 0.1 µs), the number of both ions and water decreases in this region (Figure 3.20A, B). This effect is most obvious for the Cl^- ions due to their electrostatic repulsion by the phosphate groups of the DNT backbone.

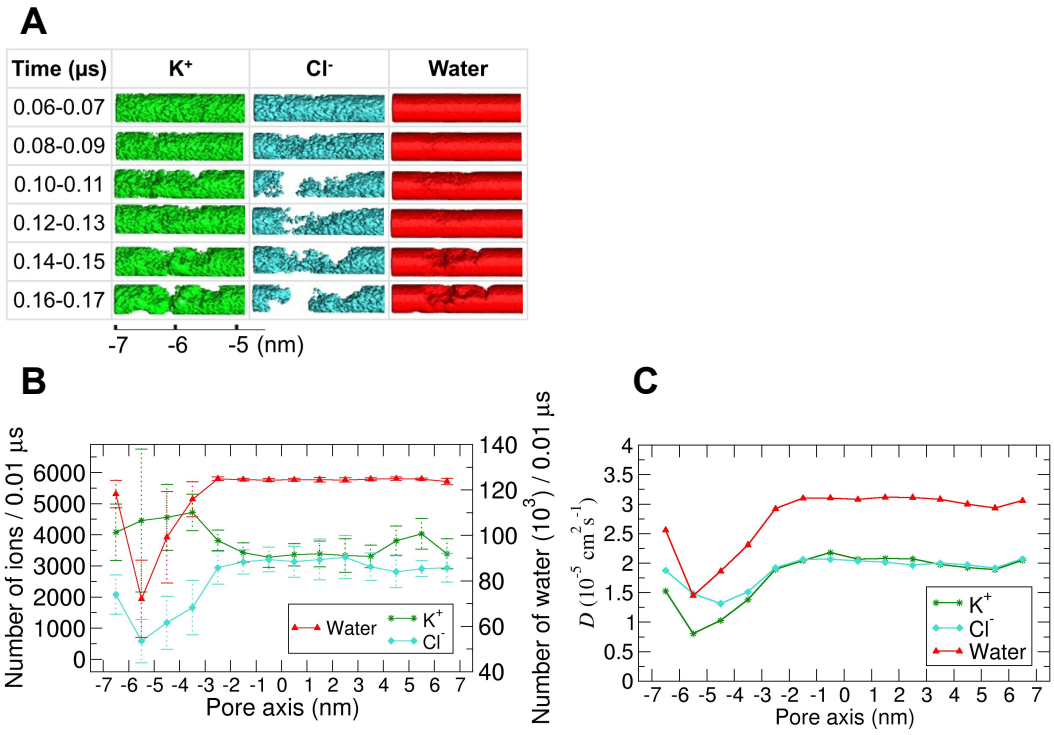

Figure 3.20: **Distribution and diffusion of ions and water along the pore axis**. **(A)** Time-evolution of the distribution of K^+ ions, Cl^- ions and water molecules in the G1-G2 region of the pore (shown for a central cylindrical volume of radius 0.5 nm) for consecutive intervals of the DNT1 simulation. **(B)** Average numbers of ions or waters per 0.01 µs frame along the central pore axis, calculated between 0.06 µs and 0.17 µs of the simulation. **(C)** Diffusion coefficient of K^+ ions and of Cl^- ions and water molecules as a function of position along the pore axis of the DNT1 system calculated between 0.06 µs and 0.17 µs of the simulation. For **B** and **C** cylindrical grids of length 0.1 nm and radius 0.5 nm were constructed along the central pore axis. Figures have been adapted and modified from Maingi *et al.*[91]

I also determined the diffusion coefficients of ions and water molecules in the lumen along the nanopore axis derived from their mean square displacements locally within grid volumes along the DNT; Figure 3.20C. Along the central region of the pore (from G2 to G5) constant diffusion coefficients were observed for water (ca. 3 x 10^{-5} cm^2 s^{-1}) and ions (ca. 2 x 10^{-5} cm^2 s^{-1}). These are just a little lower than the corresponding values in bulk solution (see Figure 3.23D and below). In the constricted G1 mouth these diffusion coefficients were decreased approximately twofold, this effect being most marked for K^+ ions. These findings indicate that even the incomplete 'closure' of the G1 gate at the mouth of the pore significantly slows diffusive motions of both ions and solvent, consistent with a likely effect on ionic conductance. To keep the discussion seamless, details of the diffusion coefficient measurements are provided further at the end of this section (see Note on diffusion analysis) along with other diffusion calculations.

I also compared the volumes of the nanopore for the four simulations. This is consistently ca. 1.1 x 10^2 nm^3 for all four simulations; Figure 3.21 and Table 3.3.

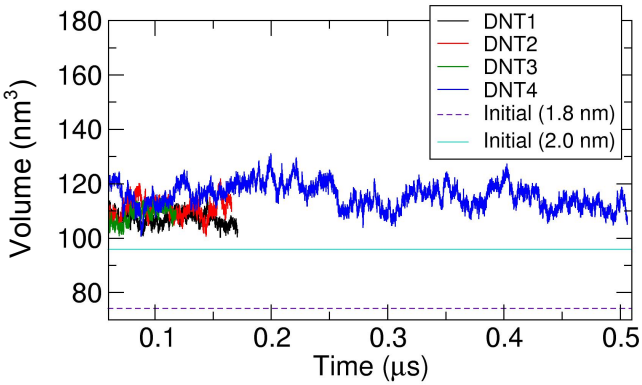

Figure 3.21: **DNT pore volume profiles.** Volumes were calculated using Epock[139]. Horizontal lines show initial volumes for the two initial hexagonal models. In all cases a lumen of constant volume is observed irrespective of the starting model (see also Table 3.3). Figure adapted and modified from Maingi *et al.*[91]

Model	V_{pore} (nm^3)	$[K^+]$ (M)	$[Cl^-]$ (M)
DNT1	107 ± 2.4	1.93 ± 0.14	0.82 ± 0.11
DNT2	110 ± 3.3	1.81 ± 0.15	0.87 ± 0.13
DNT3	108 ± 2.9	1.82 ± 0.16	0.96 ± 0.17
DNT4	115 ± 4.5	1.64 ± 0.15	0.68 ± 0.14

Table 3.3. Average equilibrium (simulation time after 0.06 μs) properties of the lumen region of the full-length DNT models. The ionic concentrations were calculated within a cylinder of radius 1.2 nm passing through the centre of the DNT. Table adapted and modified from Maingi *et al.*[91]

Calculation for the numbers of both ionic species in lumen (Table 3.3 and Figure 3.22) yielded an overall selectivity for cations (K^+) over anions (Cl^-), of ca. 2.4:1, as anticipated given the anionic nature of DNA.

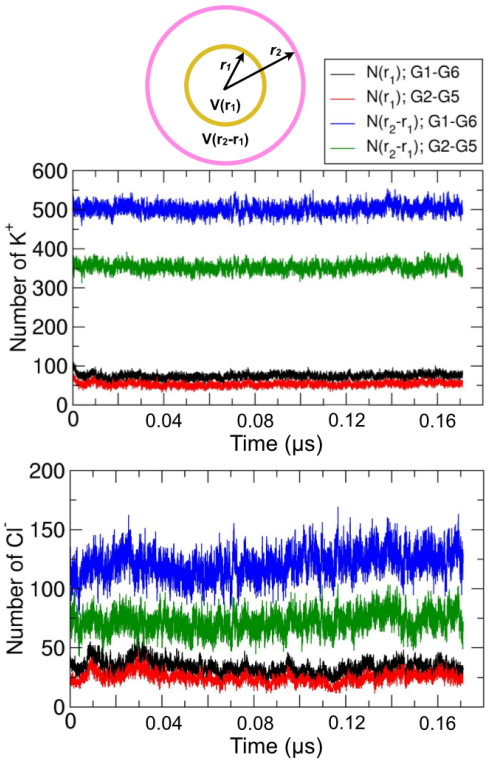

Figure 3.22: **Ion selectivity in DNT1 pore**. Number of ions calculated in cylindrical volumes $V(r_1)$ and $V(r_2 - r_1)$ for lengths G1-G6 and G2-G5. Inner cylinder r_1 = 1.2 nm and outer cylinder r_2 = 3.5 nm. Regardless of the volume chosen the number of K^+ ions is always more than Cl^- ions, and the DNT lumen can be considered as cation selective (see also Table 3.3); as external voltage is not applied in my studies thus the conductance state selectivity under external voltage conditions can not be compared directly. Figure adapted and modified from Maingi *et al.*[91]

Porous Walls.

Visual inspection of the simulations indicated a degree of penetration of both species of ions and of water molecules into the walls of the DNT nanopore. I quantified this by calculation of cylindrical radial distribution functions (similar to grids as shown in Figure 3.23A) of ions, of water molecules, and of the phosphorous atoms of the DNA along the DNT core region G2-G5 (Figure 3.23B, C). The cylindrical radial distribution was achieved by considering cylindrical shell grids originating from the central DNT1 axis up to radius r = 6 nm and length ranging from G2 to G5 regions. The grids were constructed with dr = 0.05 nm. G1 and G6 regions were ignored for this analysis as they were found to be highly dynamic. To calculate densities, the numbers of ions and water molecules were counted at each time step in each cylindrical shell grid and averaged over 10 ns. I have also calculated these densities on randomly chosen 10 ns trajectories (from 0.06 μs to 0.17 μs) and results were similar and consistent quantitatively and qualitatively.

Interestingly, in the very centre of the lumen of the nanopore the number densities of ions and water molecules are very similar to those in the bulk solution, which thus shows only weak cation selectivity. However, adjacent to the walls the cation:anion selectivity rises to >3:1. It can be seen that there are a significant number of K^+ ions in the walls of the nanopore, that the water density falls within the walls by ca. 3x on average, and that there are even some Cl^- ions within the wall region.

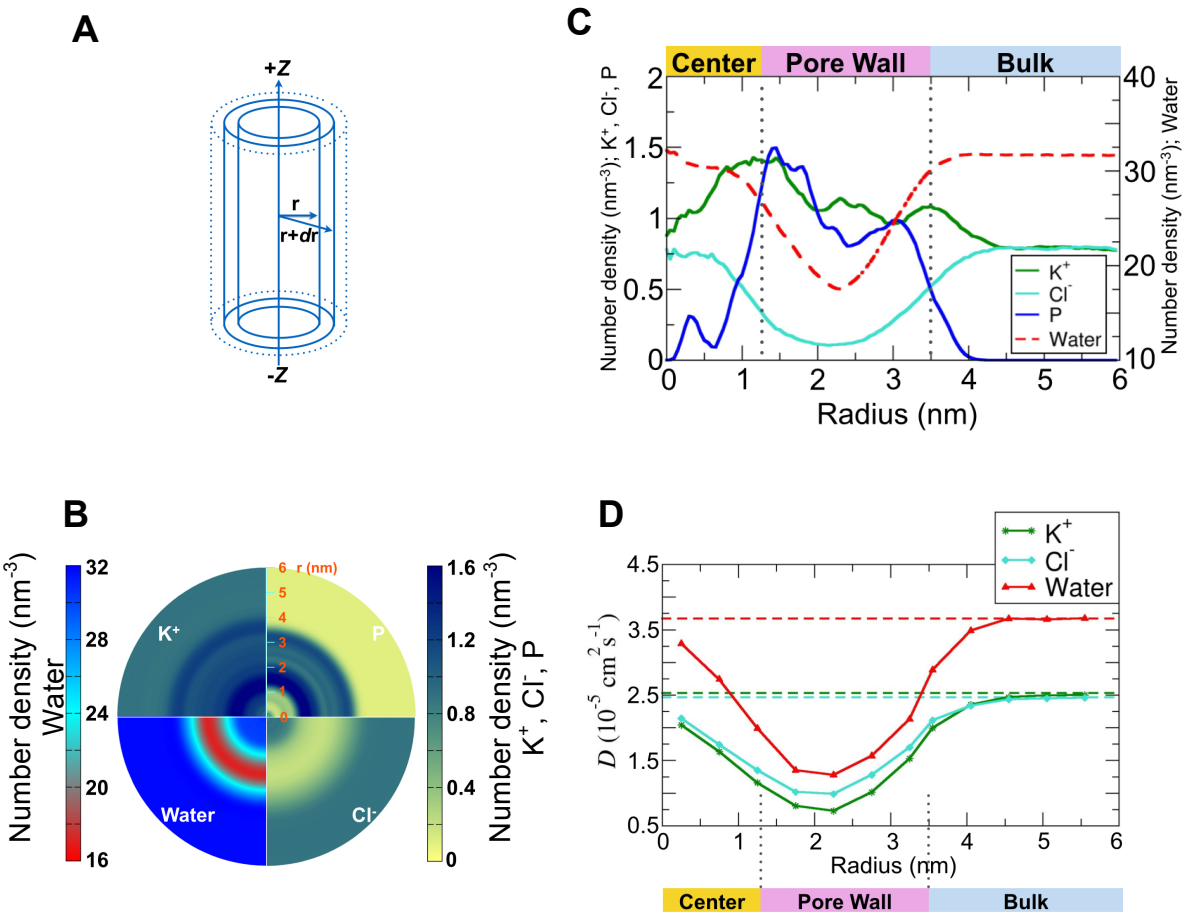

Figure 3.23: **Density and diffusion of ions and water through the DNT walls**. **(A)** Cylindrical shell grids, *r* 0 to 6 nm (with d*r* = 0.05 nm for **B** and **C** and with d*r* = 0.5 nm for **D**) chosen for sampling ions and water molecule along the DNT1 length G2-G5. **(B)** Radial distribution of phosphorous atoms, K^+ ions, Cl^- ions and of water molecule densities. **(C)** Densities of phosphorous atoms, K^+ ions, Cl^- ions and water molecules, as a function of the radial distance from the center of the DNT. The DNT pore wall extends ca. 1.2 to 3.5 nm along the radial axis. **B** and **C** were calculated using cylindrical grid shells of thickness 0.05 nm and length extending over the G2-G5 region of the DNT. **(D)** Diffusion coefficients of K^+ ions, Cl^- ions and water molecules as a function of the radial distance to the center of the DNT. The horizontal dashed lines show values corresponding to bulk conditions. The *D* values were calculated in cylindrical grid shells of thickness 0.5 nm and G2-G5. Figures adapted and modified from Maingi *et al.*[91]

The presence of ions and water within the wall of the DNT nanopore shows the soft nature of this interface. I therefore analysed diffusion coefficients of ions and water as a function of radius (again along the DNT core region G2-G5; Figure 3.23D). Within the wall region, the diffusion coefficients are reduced by ca. 2.5

fold, which is perhaps less than anticipated. This directly reveals the porous nature of the wall of the DNT nanopore.

Returning to visualization of the simulations in this context, detailed examination of simulation snapshots (Figure 3.24A) reveals ‘sideways leaks’ in the wall of the nanopore, which can accommodate ions and water molecules. I therefore analysed the spatial distribution of diffusion coefficients of ions along cylindrical coordinates, Figure 3.24B, and could see a clear (negative) correlation between the local ionic diffusion coefficients and the local density of atoms from the DNA molecules of the DNT (Figure 3.24C). This confirms the sponge-like nature of the walls of the DNT nanopore, with penetration of water and ions into the less densely packed regions of the pore wall. See also the note on diffusion analysis below.

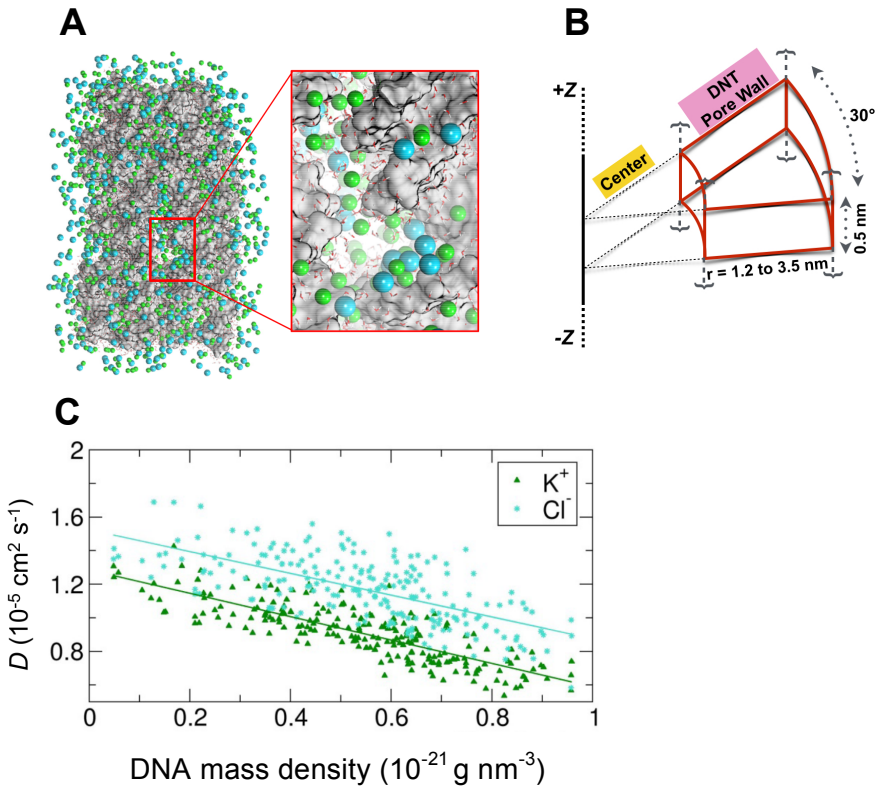

Figure 3.24: Porous nature of the DNT walls. (A) Snapshot of the DNT1 simulation showing water molecules and ions within 1.5 nm of the DNT. The DNT is represented as grey surface, oxygen of water molecules as red points, K^+ ions as green spheres and Cl^- ions as cyan spheres. The right inset shows a magnified leaky region of the DNT wall. **(B)** Representation of one of the 216 grids constructed along the pore wall cylindrical coordinates (dl = 0.5 nm, angle dθ = 30°, and thickness 2.3 nm) used to calculate both the density of DNA atoms and the diffusion coefficient of ionic species. The grids were built in the G2-G5 region pore wall of DNT. **(C)** Scatter plot showing the local diffusion coefficients of K^+ ions (green triangles) and Cl^- ions (cyan asterisks) as a function of the local mass density of DNA, calculated in each bin of the grid constructed along the pore wall in cylindrical coordinates. The green and cyan solid lines represent linear regressions through the data points for K^+ and Cl^-, respectively (with correlation coefficients of -0.70 and -0.65, respectively). Figures adapted and modified from Maingi *et al.*[91]

Note on diffusion analysis: The diffusion coefficients, D, were calculated using Einstein equation (implemented in AMBERTOOLS)[105] on particles (tracked with my PERL script):

$$2nD = \frac{MSD}{t}$$

where *MSD* is mean square displacement, *n* is the dimensions (in my case *n* = 3) and *t* is time. In general the linear relation of MSD with time is observed when *t* goes to infinity. In my case I assume linear relation over small duration of time while averaging over all the particles as my volumes in which diffusion is calculated is quite small and the population of any particles that remains in the box for very long duration would be extremely less. In some cases (grids) I observed that the MSD *vs* time relation is not linear but I used linear regression fitting in all the cases to calculate slopes (Figure 3.25). Above said approximations are not completely correct, and also the linear relation is not always expected on the extended time scales where one can observe anomalous diffusion processes[140]. Various grid sizes and shapes (Figure 3.23A, 3.24B) were used to sample migrating species depending upon the desired analyses.

The diffusion pattern along the central axis of DNT, Figure 3.20C, was calculated in cylindrical grids with radius, r = 0.5 nm and length, dl = 0.1 nm which provided 14 grids in total along the G1 to G6 length.

For the *D* values as shown in Figure 3.23D, I considered cylindrical shell grids up to radius r = 0.6 nm and length ranging from G2 to G5 regions but with dr = 0.5 nm to sample considerable numbers of ions and water molecules. In total, 12 such grids along the central axis were analysed (Figure 3.23A). I obtained *D* as slopes from the calculated MSD values averaged over 0.06 to 0.170 μs. Also, to keep calculations consistent I calculated bulk level *D* in a cylindrical shell of thickness 0.5 nm at r = 7.0 nm shown as horizontal lines in Figure 3.23D.

Finally, for the calculation of *D* values as a function of cylindrical coordinates, a representative sampling grid is shown in Figure 3.24B. In this case the cylindrical grids of height, dl = 0.5 nm, angle dθ = 30°, and thickness 2.3 nm from G2 to G5 region was chosen. The thickness of the pore wall (2.3 nm) was calculated using the peak shoulders of K^+ density represented by vertical lines in the plot shown in Figure 3.23C.

In each grid I calculated the residence time for all the ions and water molecules (oxygen atoms) for MD trajectories from 0.06 to 0.17 μs, and measured their MSD whilst in a particular grid cell. For all diffusion analysis, if the same ion/water molecule after residing in a given grid cell falls out and then returns

back in the same grid cell, after additional one or more step time gap, to reside again for additional time steps in the given grid cell, I considered these events as two different residence times. Thus, I calculated MSD for each residence time, and for each particle (ions and water), separately and then averaged. Also, as the trajectory snapshots for DNT1 (DNT2 and DNT3) case was saved every 2 ps, I assumed that during this time step any ion present inside the grid cell does not fall out and come back in during one event. *D* values were obtained from the slopes of averaged MSDs in each grid cell. In general, the average residence time for particles should be directly proportional to the grid volume and may be affected by the surrounding environment (functional groups). Thus, having enough samples, the time step over which MSD is calculated should be less than the maximum residence time for any particle in a particular grid volume. I found, *e.g.* in the grid cells along the long central axis (r = 0.5 nm, d*l* = 1.0 nm), that the major population of ions/water molecules had a residence time less than 14-16 ps though it spreads up to (maximum residence time) 60 ps. The number of ions/water molecules sampled having residence times of more than 16 ps were far less to obtain the linear MSD relation (beyond 16 ps). Also, in order to assess whether the chosen grid size*, e.g.* (r = 0.5 nm, d*l* = 1.0 nm), and the trajectory time step of 2 ps are satisfactory or not, I sampled water molecules and ions with this same grid dimension approximately 7.0 nm away from DNT and found same bulk *D* values in this grid as measured with other method on larger volumes in bulk. (Each MSD curve has at least 7 points). Figure 3.25 shows various MSD plots.

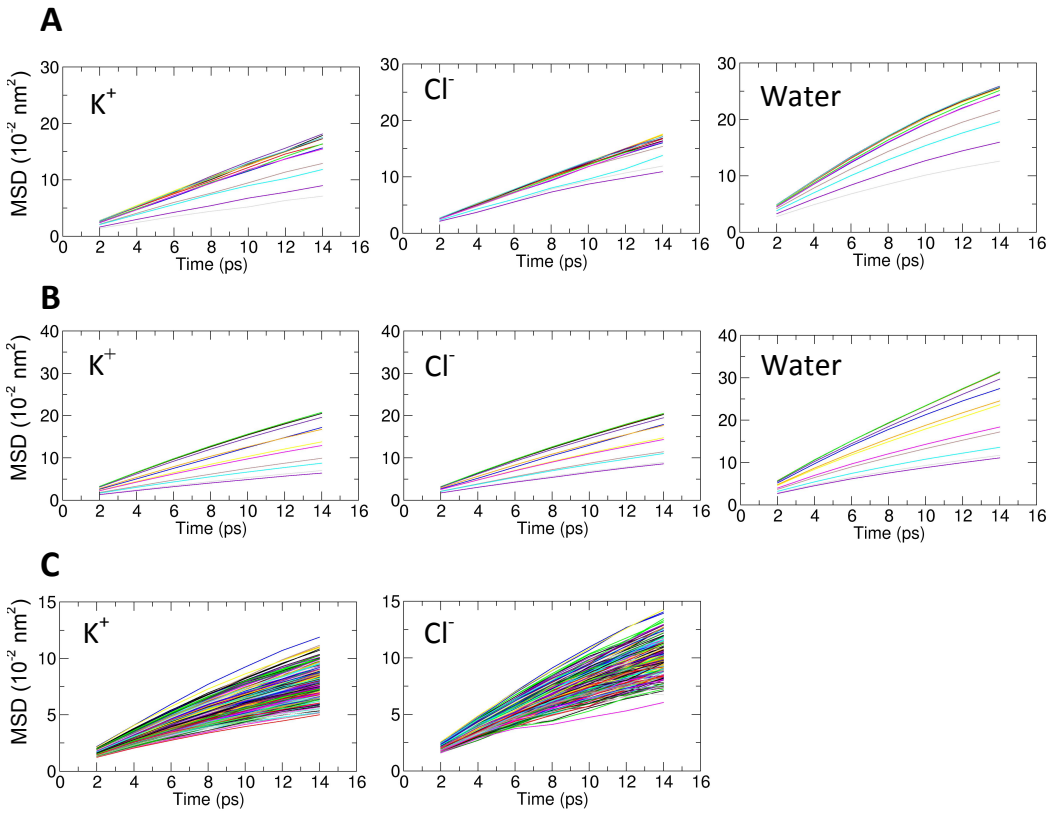

Figure 3.25: **MSD plots. (A)** Corresponding MSD plots for Figure 3.20C with the *D* values calculated in cylindrical grids of length 1.0 nm and radius 0.5 nm. **(B)** Corresponding MSD plots for Figure 3.23D with the *D* values calculated in cylindrical grid shells of thickness 0.5 nm and G2-G5 length. **(C)** Corresponding MSD plots for Figure 3.24C with the *D* values calculated in cylindrical coordinates. Figures adapted from Maingi *et al*.[91]

Specially in case of relatively smaller volume grids, I wished to convince myself whether chosen grid size (r = 0.5 nm, dl = 1.0 nm) and the trajectory time step of 2 ps are satisfactory or not. I sampled water molecules and ions with this same grid shape and dimensions approximately 7.0 nm away from DNT1. I found the same bulk *D* values in this grid as measured with other methods in larger volumes in bulk (cylindrical shell of thickness 0.5 nm at r = 7.0 nm with length G2 to G5).

3.4) Conclusions

In this study I performed all-atom MD simulations of a model DNT structure that has previously been shown to form ion conducting nanopores when embedded within a lipid bilayer membrane[57]. Comparing four simulations, I demonstrated this model is robust and stable on a sub-microsecond timescale in 1 M KCl solution. The four simulations provide a coherent description of the structure and

the conformational dynamics of the DNT. The central lumen does not undergo large fluctuations but the terminal regions of the DNT exhibit dynamic flexibility that provides a gating-like motion at the mouths of the pore on a sub-microsecond timescale. Partial closure of these gate-like regions results in significant reduction in ion and water diffusion at either mouth of the nanopore. Though I have not applied any external voltage but I observed gating-'like' motions in my simulations. Experimentally voltage-induced switching from a low transmembrane-voltage high-conductance level to a high voltage low-conductance state[83] was observed. The molecular nature of the voltage-induced gating is unclear and could include an electric-field-induced fraying of DNA duplexes, flipping of DNA loops at the pore entrances, conformational alterations in the overall duplex structure or the relative position between duplexes, or repositioning of the pore within the bilayer. Whilst my simulations were not performed for a membrane embedded DNT under transmembrane voltage, the intrinsic gating-like behaviour at the pore mouth regions offers a likely mechanism for the experimentally observed conductance switching. In future, this may be confirmed by applying an external voltage using a double-bilayer system in which different ion concentrations in two compartments creates a transmembrane voltage[141] or simply by applying an external electric field as implemented in GROMACS[99].

These fundamental insights into pore dynamics can also aid the design of pores of defined structural and functional properties. For example, the known location for intrinsic gating of DNTs could be exploited to design biomimetic nanodevices that undergo tuneable trigger-dependent conductance changes similar to natural ion channels. By contrast, the knowledge of a stable pore lumen and flexible termini can serve as a structural guide to reengineer DNA nanopores. For sensing applications, the pore should be void of major intrinsic conformational changes which would interfere with the binding or detection of passing analytes. Clearly, the flexible pore termini could be altered to prepare a scaffold of increased stability.

The nanopore is, as anticipated, cation selective, with strong selectivity at the (inner) walls of the nanopore, but little selectivity in the centre of the pore lumen. It should be noted that in my studies I have not applied any voltage across the DNT, and the ion selectivity profile for the conducting state under the applied

voltage across the membrane-embedded DNT can be different, thus should not be compared directly. Perhaps surprisingly, the walls of the DNT structure are shown to be soft and porous, allowing lateral leakage of ions and water molecules. This is in marked contrast to e.g. carbon nanotubes[142] or protein (β-barrel)[73] nanopores, providing an interesting feature which may be exploited in future DNT nanopore designs.

These studies in particular help to probe the diffusive behaviour of ions within the pore formed by a DNA-origami structure. Due to its highly charged nature, DNA interacts strongly with ions. For example, the structure of dsDNA is influenced by ions[143–145]. Within DNA-origami constructs, which are made up of dense assemblies of DNA strands, effects of ionic concentration on DNT structure are expected. For example, changes in ionic concentration have been shown to modify the structure and the behaviour of DNA-plate architectures[130]. It will also be of interest in the future to explore the effects of hydrophobic derivatization of DNTs, as this is required to stabilize such nanopores within a lipid bilayer environment.

To model future DNA nanotubes for MD simulations one can derive a few general features from my results. The terminal motions at the two ends of DNT were observed in all the models but were not consistent in terms of which end shows more fluctuations (G1 *vs* G6). For example, G1 terminal fluctuations were more pronounced in DNT1 model while it was seen at the G6 end for DNT 2, 3 and 4 models. Although, considering the computational resources, I have not performed repeat simulations of the same model but considering results I have obtained so far concludes that these fluctuations depends, *e.g.*, on the force-field used (DNT1 *vs*. DNT4 but at the same time I use different temperature control), different initial inter-helix separation (DNT1 *vs*. DNT2), different temperature controls so used (DNT2 *vs*. DNT3). Though number of combinations could have been tested by varying just one parameter (*e.g.* different temperature controls for a same initial model, and simulated in same MD engine software) but due to limited computational resources these tests were not performed, thus comparative conclusions cannot be drawn properly. But it is quite suggestive (from all the models) that G1 and G6 regions were found be flexible due to absence of some staple-strands. Thus, one can consider designing future DNTs by providing staple-

strands in the terminal ends of the tube. In all the models middle regions (G2-G5) shows consistent behaviour (stable RMSD, low RMSF). Also, the inter-helix separation (for G2-G5) region shows convergence towards a common value of *ca*. 2.4 nm. Thus, future DNT models can be designed using such inter-helix separation instead of 1.8 or 2 nm used in my study. But as discussed in Chapter 2, one will have to consider Holliday-junctions parameters as well thus requiring critical efforts to build future DNT models.

In terms of analysis, tools may be developed to visualise the flow of ions further characterising the porous nature of pores. Earlier Chavent *et al*.[146] visualised lipid diffusion using vector field flow analysis, and as this methodology was developed in our lab there is a possibility to extend this further to visualise the flow of ions as well. Diffusion analysis may also be improved by using more appropriate equations for ions diffusing in relatively smaller volumes (thus short residence times) rather than simpler assumption of using Einstein diffusion equation suitable for bulk volumes. Nonetheless, one should consider performing repeats of simulations using same set of parameters but by generating, for example, different initial velocities. In general, this helps to determine whether the analyses have converged using large set of ensembles.

CHAPTER 4

4 Coarse-grained simulations of DNA nanotube interactions with lipid bilayer membranes

The contents of this chapter are taken directly and/or modified from my paper accepted in Nature Communications (in press). Additional analyses have also been incorporated.

4.1) Introduction

I discussed DNA nanotubes (DNTs) and their dynamics in aqueous solution in Chapter 3. However, a central challenge[47] in the construction of DNTs is how to stably embed a polyanionic DNA assembly within the hydrophobic environment presented by the core of a lipid bilayer[43,57,65,66,82,85,87,147]. This hurdle can be overcome by chemically modifying the outer surface of DNTs with a band of hydrophobic groups to match the nature of the bilayer core. Successfully tested hydrophobic groups to date (till August 2016), as discussed in the introductory Chapter 1, include ethyl modifications[57,87], porphyrin rings[65] and cholesterol[66,82] molecules; Figure 4.1.

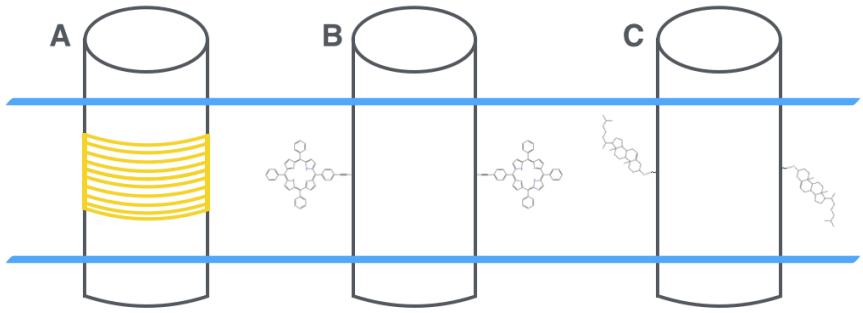

Figure 4.1: **DNT hydrophobic modifications**. **(A)** Ethylphosphorothioate modifications, usually many, span like a belt with a thickness almost matching the membrane hydrophobic core. This involves charge neutralization of all the modified phosphate groups. **(B)**, **(C)** Porphyrin and cholesterol molecules, respectively, attached on the surface of DNTs act as anchors in the membrane core. Usually there are 2-3 modification sites, and phosphate charge neutralization depends on the nature of the chemical linkage. Note that these are schematic representations—the precise positioning and chemical attachment varies depending upon the specific design.

However, the porphyrin or cholesterol modifications do not provide extensive hydrophobic coverage on the DNT surface, and may result in a greater degree of mismatch between the lipid tails, which form the bilayer core, and the DNT surface. Also, in contrast to the ethyl modifications, the number of cholesterols or porphyrins employed is much smaller, and thus the residual negative charge on the DNA surface exposed to the hydrophobic core remains high.

Despite experimental progress in anchoring DNTs into membranes, the molecular structure, dynamics and energetics of the process are largely unknown. In particular, the molecular structure of the interface between the DNA nanotube and the lipid bilayer remains unclear. Similarly, the dynamic structural changes of the inserted pore and the energetics of membrane insertion are unknown. I use molecular dynamics (MD) simulations to explain the interactions of DNA nanopores with lipid bilayers. In this chapter I cover the coarse-grained aspects of MD, with discussion of all-atom simulations covered in Chapter 5.

MD is a powerful tool to explore the interactions of membrane proteins with lipid bilayers[148,149]. MD simulations have also been employed to examine the conformational dynamics and permeation properties of nanopores (e.g.[142,150–153]), including those formed by non-membrane spanning DNTs[90,91,93] and a single bilayer-spanning DNA duplex[85]. Recently, MD simulations have provided an estimate of the substantive free energy barrier to insertion of unmodified DNA into a lipid bilayer, whilst suggesting that the presence of divalent cations may lower the energetic barrier[154]. The use of coarse-grained (CG) MD is well suited to extend simulation timescales,[155] thus enabling exploration of the energetics of protein/lipid interactions[156].

4.2) Methods

4.2.1) Model building

Choosing all-atom structures and cluster analysis.

I selected two all-atom (AT) DNT[91] models as starting points for coarse-grained simulations. One structure was obtained by performing cluster analysis on the

DNT1 model simulation (0.17 μs)[91] described in Chapter 3. Another structure for CG simulations was taken from the initial model AT coordinates of DNT1. Cluster analysis was performed using an algorithm reported by Daura *et al.*[157] implemented in GROMACS[99]. In short, root mean square deviation (RMSD) is calculated among all the structure pairs, and the number of structures is determined around each structure falling within a cut-off value of RMSD (heavy atoms only). This process is iterated over the residual population of structures until it is depleted. I tested cut-offs ranging from 0.2 nm to 0.6 nm. To maintain a sufficient RMSD between the cluster populations, I calculated the RMSD between the two most populated clusters in each cut-off and the percentage of most populated cluster (Figure 4.2). Within the same cut-off, ideally the two most populated clusters should not be either too far separated or close (cut-off 0.2 nm *vs* 0.3 nm) with respect to each other, and the most populated cluster should form at least 50% of the total population of structures. The RMSD between the two most populated clusters does not change much for cut-offs between 0.4-0.6 nm, and with a 0.45 nm cut-off the population is slightly more than 50%. Thus, the central structure obtained with a 0.45 nm cut-off from the most populated cluster (ca. 51%) was chosen for coarse-graining. Also, note that in general AT average structures are usually not ideal candidates for CG models as bonding terms could be far from equilibrated structures thus I have employed cluster analysis to choose a suitable structure.

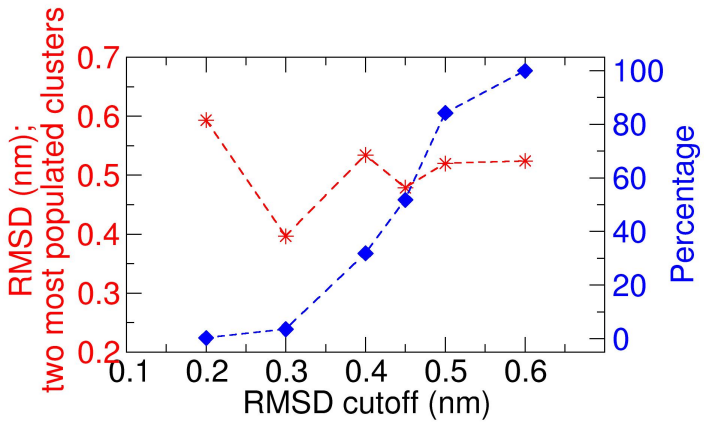

Figure 4.2: **Cluster analysis**. Left *Y*-axis shows the RMSD between the two most populated central structures of clusters for each cut-off RMSD value shown on the *X*-axis. Right *Y*-axis shows the percentage of the most populated cluster for each cut-off value.

Coarse-grained models from all-atom structures.

Representative atomistic structure of the DNTs (from the cluster analysis described above and the initial built model) were converted to a CG MARTINI model using a modified version of the *martinize.py*. This script was kindly modified by Jaakko Uusitalo (University of Groningen) to provide an elastic network model (ENM) to restrain the DNA conformation during the CG simulations. An elastic network was created for each of the six double helices as well as between the staple strands (see Figure 3.2, Chapter 3 for scaffold and staple strands) that form the junctions. A 'stiff' ENM with force constant value of 500 kJ $mol^{-1}nm^{-2}$ was used unless stated otherwise. Initial tests for CG-DNT simulations proved to be unstable, and after removing cross-junction bonding terms, while keeping their elastic-networks, the simulations were found to be stable. This might be due to the fact that MARTINI DNA is designed for straight, and unfolded DNA strands. For 'soft' ENM a lower force constant value of 100 kJ $mol^{-1}nm^{-2}$ was used. The GROMACS files for both the cases are available at *sites.google.com/site/maingivishal/scientific_data*

Hydrophobic particles, mimicking ethyl-phosphorothioate, were attached to selected phosphate beads of the DNT to generate hydrophobic surface bands (Figure 4.3). This approach was used to generate a variety of structures differing in nature of the surface band, which were simulated further, and summarised in Table 1.

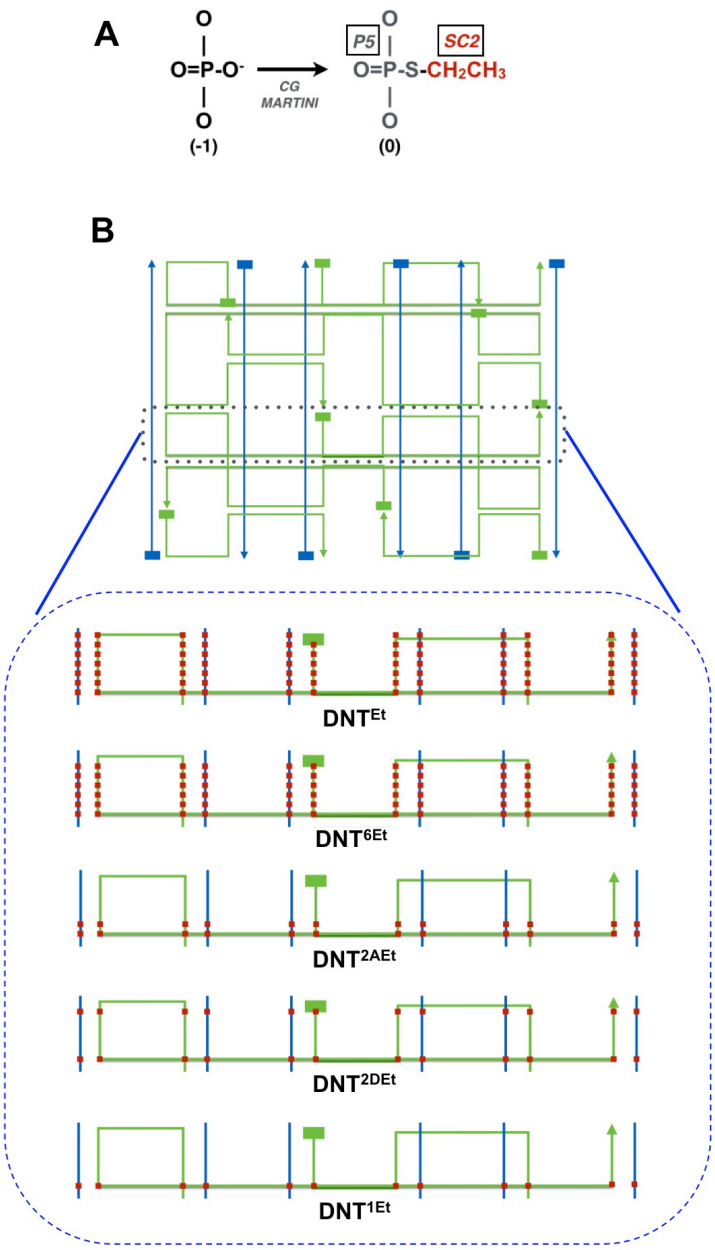

Figure 4.3: **Topology and positioning of hydrophobic particles in the DNA nanotube model**. **(A)** Schematic diagram showing the ethyl modification of the DNA backbone. Boxes indicate the MARTINI particle types used for coarse-graining. **(B)** The upper panel shows the six scaffold strands (blue) and the eight staple strands (green) of the DNA nanotube (design taken from Burns *et al.*[57]). The lower panel shows the positions of the hydrophobic (Et; in red) modifications in the different models presented in Table 4.1 of the main text.

CG simulations	Duration
$DNT^{Et}PC$[1]	6 x 1 μs
$iDNT^{Et}PC$[2]	6 x 1 μs
$iDNT^{Et}PCold$[3]	4 x 1 μs
$DNT^{Et}PCsoft$[4]	6 x 1 μs
$DNT^{Et}PC/PS$[5]	6 x 1 μs
$DNT^{6Et}PC$[6]	6 x 1 μs
$DNT^{1Et}PC$[7]	6 x 1 μs + 2 x 10 μs
$DNT^{2AEt}PC$[8]	3 x 1 μs
$DNT^{2DEt}PC$[9]	3 x 1 μs
5 displaced $DNT^{Et}PC$s (see Figure 4.12)	1 μs each
SMD (push in) $DNT^{Et}PC$ (see Figure 4.13AB)	10 nm/ μs; >2 μs x 3
SMD (push in) DNT PC (see Figure 4.13C)	10 nm/ μs; >2 μs x 3
SMD (push in) $DNT^{Et}PCxy$[10] (see Figure 4.14A)	10 nm/ μs; >2 μs
SMD (push in) DNT PCxy[10] (see Figure 4.14A)	10 nm/ μs; >2 μs
SMD (pull out) $DNT^{Et}PC$ (see Figure 4.15A)	10 nm/ μs; >1.5 μs
SMD (push in) $DNT^{1Et}PC$ (see Figure 4.14B)	10 nm/ μs; >2 μs x 2
SMD (push in) $DNT^{2AEt}PC$ (see Figure 4.14C)	10 nm/ μs; >2 μs x 2

Table 4.1: **Summary of simulations performed**.
[1]This is the main CG simulation - $DNT^{Et}PC$ with 7 hydrophobic bands (see Figure 4.4). All CG simulations were performed using 1.0 M NaCl.
[2]i = initial DNT model used[91] (see Figure 4.5)
[3]old = MARTINI version 2.0 [116] used for lipids (old 5 beads oleoyl tail model); see Figure 4.5
[4]soft = soft ENM used for DNT (see Figure 4.5)
[5]PC/PS = lipid mixture: 80% PC and 20% PS (symmetric); see Figure 4.5
[6]6Et = 6 hydrophobic bands instead of 7 (see Figure 4.5)
[7]1Et = single hydrophobic band (see Figure 4.6)
[8]2AEt = 2 adjacent hydrophobic bands (see Figure 4.6)
[9]2DEt = 2 distal hydrophobic bands (see Figure 4.6)
[10]Positional restraints in the xy-plane used in the SMD to keep the DNT long axis parallel to the bilayer normal (see Figure 4.14A)

4.2.2) Simulations

The CG hydrophobic DNT models were embedded in a phospholipid (POPC) bilayer using the *insane.py* script[115]. In brief, this script helps to generate a lattice of bilayer lipid molecules around the centred molecule, and add solvent molecules above and below the bilayer. The hydrophobic domain was centred in the hydrophobic core of the membrane. No lipids were inserted inside the DNT lumen. CG water particles were also included using *insane*, and the ionic concentration was set to 1 M NaCl. The final dimensions of the simulation box were ca. 18 x 18 x 25 nm^3 (~ 73,000 CG particles).

GROMACS (*www.gromacs.org*) v4.6.5[99] was used for CG simulations with the MARTINI DNA (version 2.1)[114] and lipid (version 2.0 new; 4 bead oleoyl model)[115] forcefield. To remove the initial close contacts, the solvated DNT-membrane system was subjected to steepest descent energy minimization. The minimized system was then equilibrated for 1 ns with a time step of 10 fs. Semi-isotropic pressure coupling at 1 bar was employed using the Berendsen barostat[106] with a 3 ps coupling constant. A velocity rescale thermostat[107] was used to maintain temperature at 300 K with a 1 ps coupling constant. A cut-off of 1.2 nm was used for Coulombic interactions, shifted between 0.0 and 1.2 nm, and van der Waals interactions were shifted between 0.9 and 1.2 nm. The neighbour list was updated after every 10 steps with a cut-off distance set to 1.4 nm. Repeat simulations were run by generating new random seeds for velocities. Production simulations were for up to 1 μs.

4.3) Results and discussion

4.3.1) DNA nanopores in a bilayer

CG MD simulations were performed using the MARTINI forcefield (see Methods above), which has proven suitable to probe interactions of a large variety of biomolecules and nanoparticles with membranes[158]. These simulations yielded a model of the DNA nanopore with a hydrophobic surface band (DNT^{Et} – Figure 4.3B) embedded in a simple model (phospahtidylcholine; PC) lipid bilayer. The resultant DNT-bilayer system, with water and ions (1 M NaCl) on either face of

the membrane, was simulated for 1 μs (6 repeats; see Table 4.1). A snapshot of the final configuration from one of these simulations is shown in Figure 4.4A. The DNT can be seen to span the membrane, with the ethyl groups of the hydrophobic band located within the hydrophobic core of the bilayer. If one examines the bilayer with the DNT removed (Figure 4.4B) it can be seen that there is a degree of local reorganization of the lipids in the vicinity of the inserted nanopore, such that the lipid headgroups are drawn into the bilayer core, so that they may form close contacts to the modified DNA. A more quantitative view may be obtained by examining the bilayer thickness and the distribution of different components as a function of the radial distance from the centre of the nanopore (Figure 4.4C). The thickness profile reveals that membrane thinning occurs around the DNT and persists up to ~2 nm away from its surface. As in my atomistic simulations of a DNT in aqueous solution (Chapter 3), water is present both within the lumen of the nanotube and also (to a lesser extent) within the walls of the nanotube itself. Lipid density can be seen to overlap radially with DNA density, confirming the close interactions of lipid molecules and DNA of the nanopore noted above. The lipid density can also be seen to overlap with the radial density of the ethyl particles of the hydrophobic band (which are present on both the inner and outer surfaces of the nanotube as well as within the walls of the nanotube). Lipid molecules are absent from the lumen of the nanotube and so do not 'clog' the pore.

Initial bilayer thickness analysis performed on different parts of the trajectory suggested that the bilayer thickness profiles are stable after 0.6 μs (especially near the surface of the DNA pore). Thus, subsequently, all the time averaged properties were calculated over the last 0.4 μs of the trajectory.

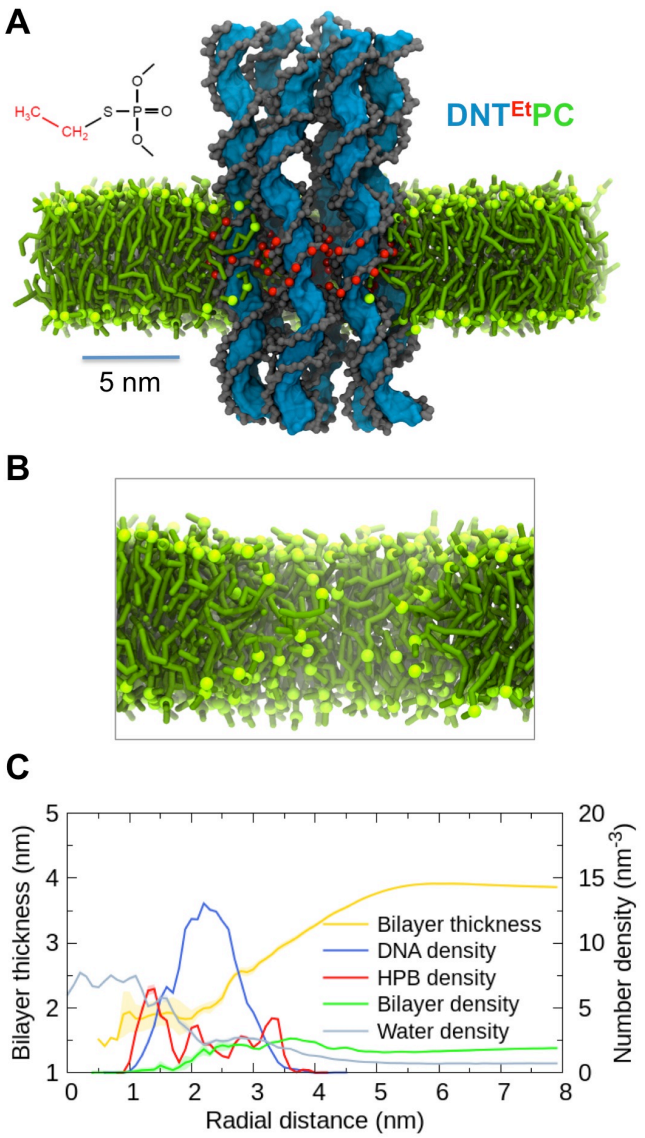

Figure 4.4: **DNA nanopore in a bilayer**. **(A)** Final 1 μs snapshot from a CG simulation (DNT^{Et}PC; see Table 4.1 for details) of a DNA nanopore embedded in a lipid bilayer; water and ions are omitted for clarity. The colour scheme is: blue surface = DNA bases; grey surface = DNA backbone; green = PC tail/phosphate head group; red = hydrophobic (HPB) ethyl modifications. **(B)** A zoomed in view of the same system, showing only the lipids and revealing the local perturbation of the bilayer proximal to the DNA nanopore. **(C)** The bilayer thickness (yellow line; left hand axis) and number densities of the system components (right axis) are shown as a function of radial distance from the centre of the nanopore. Bilayer thickness values were calculated between leaflet phosphates. For number density analysis, cylindrical grid shells (concentric cylindrical shells 0.1 nm thick and 4 nm long, and thus approximating the membrane thickness), were constructed around the central axis of the transmembrane domain of DNT. Thickness and density profiles were averaged over the last 400 ns of each simulation. Results shown were averaged over six different runs. Lipid bilayer and hydrophobic particle (i.e. Et group) densities have been scaled up 5x for clarity.

Simulations were performed to assess the robustness of the results to variations in the structure of the nanotube, the nature of the bilayer, and the simulation conditions employed (Table 4.1 and Figure 4.5). Thus in simulation iDNT^{Et}PC an initial DNT model which had not been subjected to atomistic simulations in aqueous solution prior to bilayer insertion was employed, whereas in DNT^{Et}PCsoft a 'softer' (i.e. lower force constant – see Methods above) ENM was used to maintain the DNA conformation. As might be anticipated, the softer ENM results in some degree of smoothing of the density profiles whilst the iDNT model results in sharpening of the density profile peaks (relative to the DNT^{Et}PC simulations), but the overall distributions of components and the bilayer thickness profiles are not substantially different. Changes in the lipid bilayer composition (the inclusion of 20% of the anionic headgroup lipid PS) or to the lipid parameters (iDNT^{Et}PCold) or to the degree of hydrophobicity of the nanopore (DNT^{6Et}PC) likewise did not result in any substantive changes. Thus it may be concluded that the overall simulation results are robust to minor changes in the details of the simulation model and the parameters.

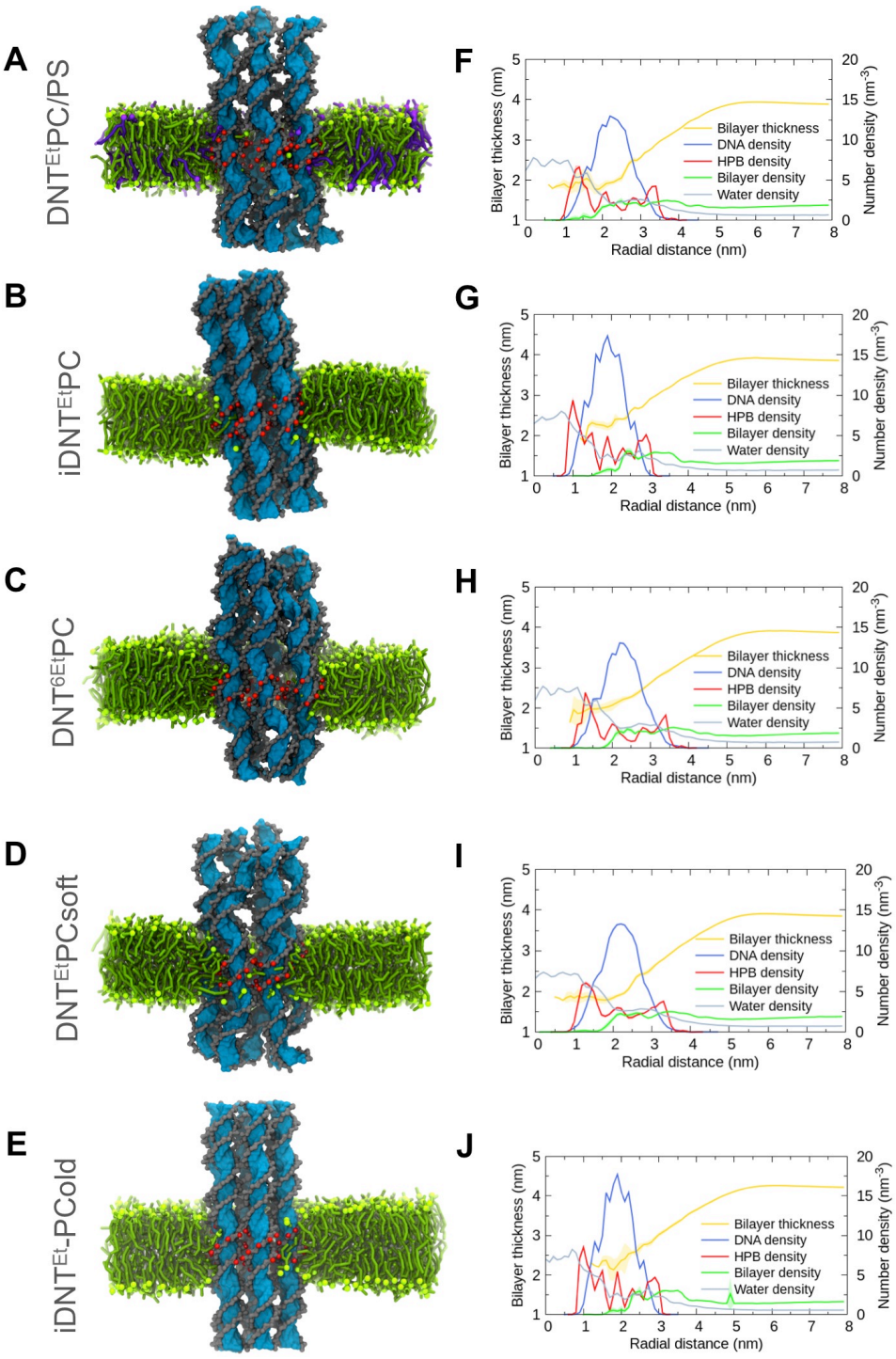

Figure 4.5: **A range of DNT models**. **(A-E)** Final snapshots showing different DNT models embedded in a membrane. See Table 4.1 for summary of simulations performed, and abbreviations used. Red beads, hydrophobic modifications (HPB); blue surface, DNA bases; grey surface, DNA backbone; green sticks/beads, POPC/phosphate head group; purple sticks/beads, POPS/phosphate head groups. **(F-J)** Bilayer thickness profiles around the embedded DNTs along with the radial density distributions of DNA-phosphates, hydrophobic particles, bilayer phosphates, and water particles. Note that the lipid bilayer and hydrophobic particle (i.e. Et group) densities have been scaled up 5x for clarity.

I wanted to determine whether a substantially smaller (functionalized ethyl) hydrophobic band altered the behaviour of the nanopore within the membrane. To this end, I simulated three systems with either: (i) a single ring of ethyl groups around the nanopore (DNT^{1Et}PC) or (ii) with just two such rings (DNT^{2AEt}PC and DNT^{2DEt}PC; see Figure 4.3, Figure 4.6A and Table 4.1 for details). The overall behaviour of each of these nanopores was similar to that of the nanopores with either 6 or 7 rings of ethyl groups (see above). However, analysis of the distribution of tilt angles (between the long axis of the DNT and the bilayer normal; see Figure 4.6B) revealed that nanotubes with fewer rings of ethyl groups 'wobbled' more in the bilayer, as evident from a shift in and broadening of their tilt angle distributions. This is likely to be an important consideration when designing hydrophobically-derivatized DNTs, as a wide range of tilt angles might be anticipated to perturb ion flow through the nanopores. It has been suggested that lipids may be able to occlude nanotubes[159], and this would be expected to be more likely to occur if nanopores were prone to tilting, especially if shorter (relative to the membrane thickness) DNTs were to be explored[66].

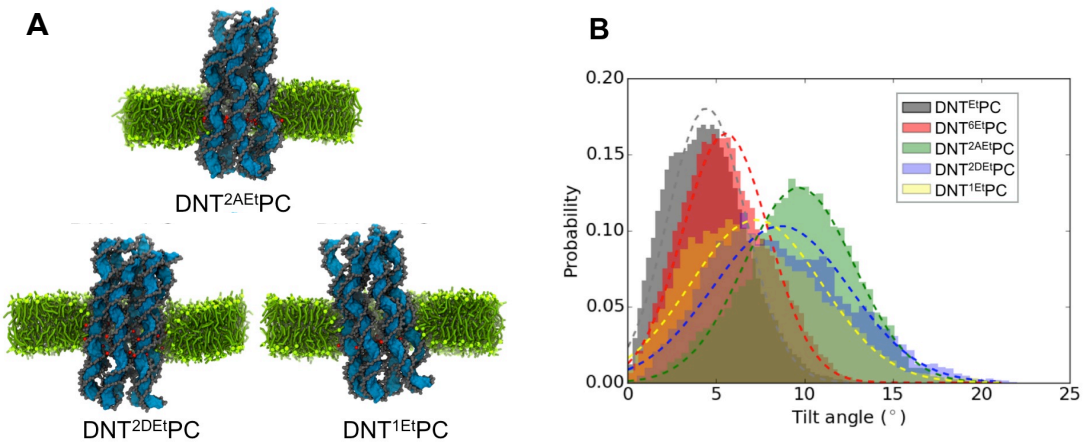

Figure 4.6:**Varying the extent of the hydrophobic anchor surface**. **(A)** Final (1 μs) snapshots are shown for various simulation systems of DNTs having different numbers of ethyl modifications (2x12; adjacent and distal, 1x12; see Figure 4.3). **(B)** DNT tilt angle distributions for the different DNT simulation systems. Tilt was calculated as the angle between the Z-axis (i.e. approximate bilayer normal) and the long axis passing through the center of each DNT.

4.3.2) Local lipid interactions

I explored the interactions of lipids adjacent to the membrane-spanning DNTs (i.e. those next to the TM domain) in more detail. Visual inspection (Figure 4.7A) revealed that a number of lipids penetrated the DNT to the extent that their head groups pointed into the lumen of the nanopore, but did not occlude it. The tails of the lipid molecules in contact with the DNT could be seen to wrap around the DNA double helix, in both the major and (to a lesser extent) the minor grooves (Figure 4.7B). Such locally perturbed lipids remain in close proximity to the DNT along the width of the hydrophobic band.

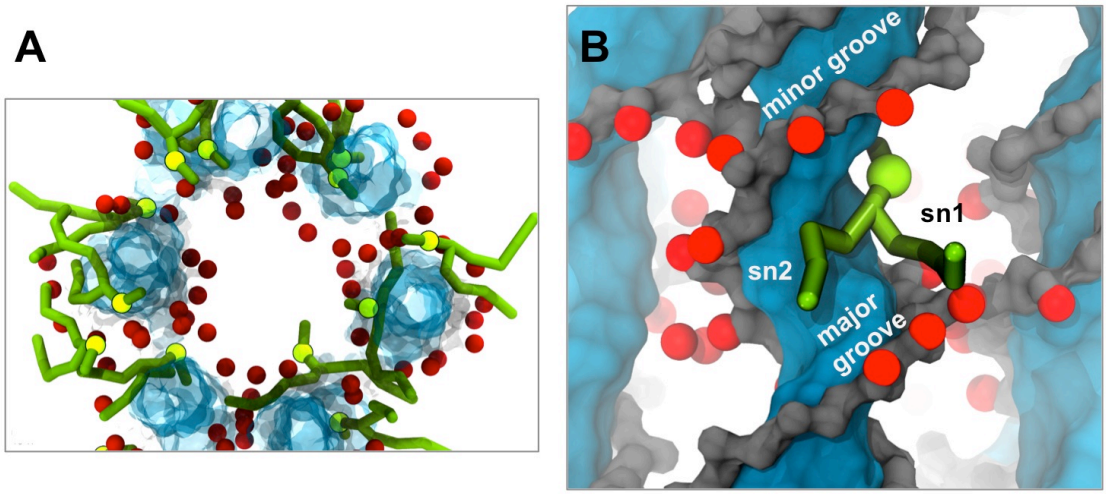

Figure 4.7: **Local lipid interactions**. **(A)** Snapshot from the CG simulation ($DNT^{Et}PC$) showing presence of locally perturbed lipids (green) around the inner surface of the DNT (blue/grey with Et groups in red) lumen. Lipids do not occlude the lumen, but some are exposed on the luminal surface. **(B)** Snapshot of the same system showing a selected lipid, the headgroup and *sn2* acyl chain of which interact with the major groove surface of the DNA.

I examined the pattern of residence times (Figure 4.8) in the minor and major grooves for the different constituent particles of the CG lipid molecules. For both the $DNT^{Et}PC$ and the $iDNT^{Et}PC$ simulations the *sn2* chain interacts with both the major and minor grooves of the DNA more than the *sn1* chain. The choline and phosphate groups (NC3 and PO4 particles) also exhibited preferential interactions with the major groove. This analysis along with the observed local thinning of the bilayer (see above) reveals that the interactions of lipid molecules with the surface of the DNT induce a local perturbation of the lipid bilayer structure.

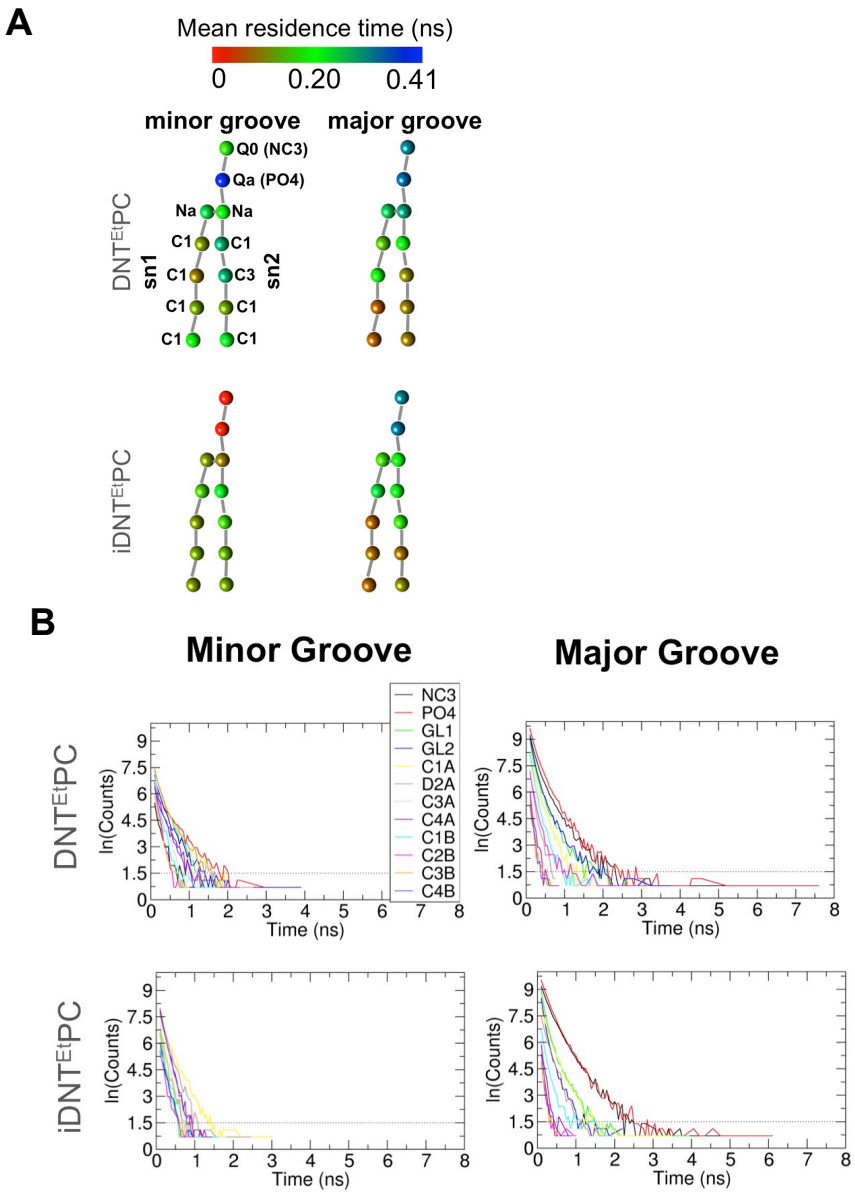

Figure 4.8: **Mean residence times of lipids in the DNA grooves**. **(A)** Patterns of interaction of lipid CG particles (coloured as mean residence times) with minor and major grooves of DNT^{Et} and $iDNT^{Et}$. Perturbed selections were determined based on the PO4 bead (POPC) remaining within *Z* of DNT HPB beads. Then residence contact times, with major or minor grooves, were calculated only for perturbed lipids by using a distance criterion defined as: any perturbed lipid bead (among 12 beads) within 0.5 nm of SC2 (minor groove) or SC3 (major groove) DNA beads, and staying at least for three regular steps; where two adjacent steps differ by 50 ps. Labels along each bead represent bead type, and labels in parenthesis show bead name for CG MARTINI. **(B)** Residence time distributions for lipids in the minor and major grooves of the DNA nanotube. Residence contact times, with grooves, were calculated only for those lipids interacting with the nanotube. Mean residence times were calculated as inverse of slopes obtained from linear regression fitting on the exponential of the form[160] $f = exp\ (-t/\tau)$.

The fact that the DNT surface induces a local perturbation of the lipid bilayer structure is also reflected in the order parameters for the lipid tails in close proximity to the surface of the DNT. POPC in the immediate surroundings of DNT^{Et} (*i.e.* interacting at DNA grooves or near the surface) were found to be highly perturbed as revealed by their density profiles (of PO4 beads) and order parameters (of POPC tails) (Figure 4.9). Lipid arrangement in 3-4.5 nm regions (from DNT center) was relatively more ordered, and reaches the bulk order parameter values found in normal CG POPC membranes[115]. This is reflected in the bilayer thickness profile which increases up to ca. 4.8 nm (Figure 4.4C). Lipid ordering is also quite sensitive to the interacting surfaces. For example, in the DNT PC case, which lacks the hydrophobic band, a negatively charged surface is presented to the hydrophobic core of the membrane. This triggers a hydrophobic mismatch leading to extended lipid order perturbation, up to ca. 4.5 nm, around DNT (Figure 4.10), which is in contrast to DNT^{ET}PC case where lipid perturbation is extended up to 3 nm. Lipid perturbations are not ideal for a membrane embedded nanopore where seamless lipid packing would be preferred to prevent the leakage of current from the outer surface of the nanopore. Such negatively charged surfaces are present in DNA nanopores to which only a few cholesterol or porphyrin moieties are attached[65,66,82]. In contrast to the DNA surfaces, I performed similar analyses on a membrane embedded protein (e.g. BtuB[161]; PDB ID: 2GUF), and the lipids were less perturbed (Figure 4.11).

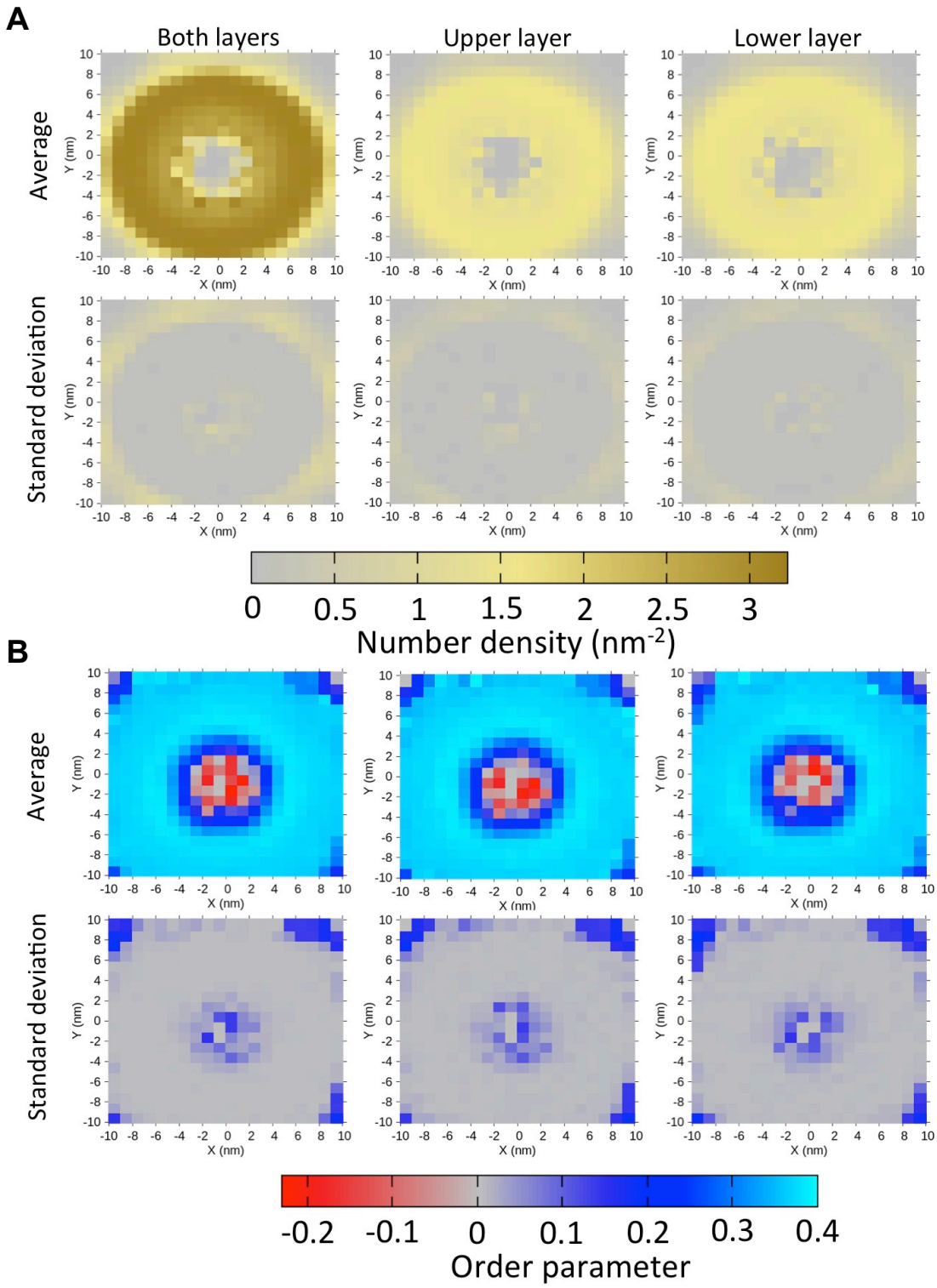

Figure 4.9: Number densities **(A)** and order parameters **(B)** of lipids sampled in square grids (1 nm^2) projected in the *XY* plane and centred around DNT^{Et} at (0, 0). Lipids *XY* positions refer to the lipid headgroup phosphates. Lipid order parameters were calculated for all the lipids in each grid box and averaged over all simulation frames for each run, and then for all six runs. The order parameter was calculated using: $P_2 = 0.5\ (3\ cos^2\langle\theta\rangle - 1)$.

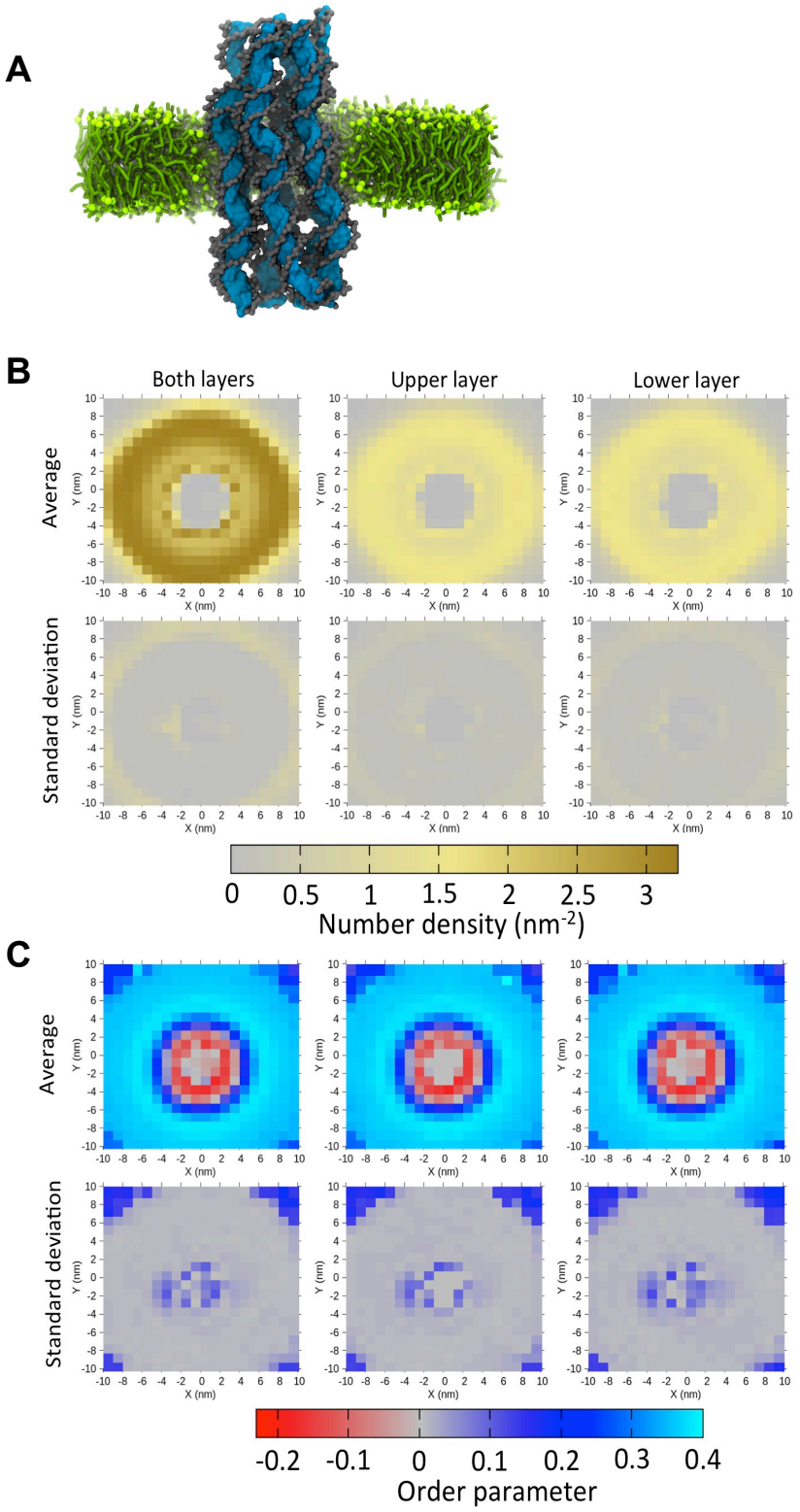

Figure 4.10: **(A)** Final snapshot of a (unmodified) DNT embedded in a POPC membrane. Number densities **(B)** and order parameters **(C)** of lipids sampled in square grids (1 nm^2) projected in the *XY* plane and centred around DNT at (0, 0). Lipids *XY* positions refer to the lipid headgroup phosphates. Lipid order parameters were calculated for all the lipids in each grid box and averaged over all simulation frames for each run, and then for all six runs. The order parameter was calculated using: $P_2 = 0.5\ (3\ cos^2\langle\theta\rangle - 1)$.

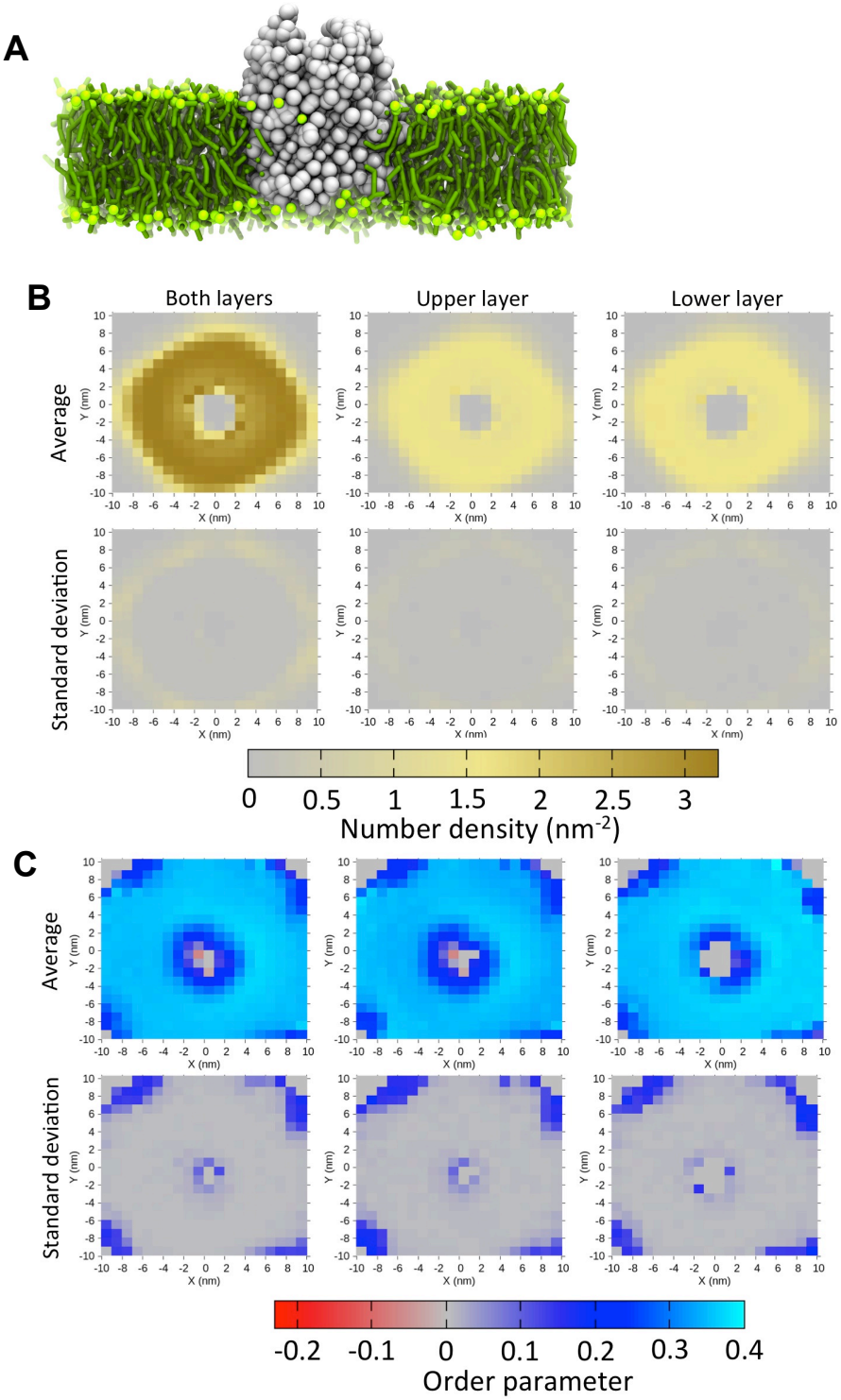

Figure 4.11: **(A)** Final snapshot of BtuB embedded in a POPC membrane. Number densities **(B)** and order parameters **(C)** of lipids sampled in square grids (1 nm^2) projected in the *XY* plane and centred around BtuB at (0, 0). Lipids *XY* positions refer to the lipid headgroup phosphates. Lipid order parameters were calculated for all the lipids in each grid box and averaged over all simulation frames for each run, and then for all six runs. The order parameter was calculated using: $P_2 = 0.5\ (3\ cos^2\langle\theta\rangle - 1)$.

4.3.3) Metastable and stable states

To explore further the (meta) stability of the membrane-embedded DNT I used steered molecular dynamics (SMD)[162] to perturb the system generated at the end of the 1 μs $DNT^{Et}PC$ simulation by displacing the DNT relative to the bilayer (Figure 4.12A). The DNT was displaced vertically to differing extents, ranging from $\Delta z = 0.5$ nm to 4.4 nm (Δz measured as distance between the centers of mass (CoM) of bilayer and hydrophobic band). Each perturbed system was then used as the starting point for a further *unrestrained* 1 μs simulation to examine whether or not the DNT relaxed back to its initial position in the bilayer. Note that for Δz = 0.5 and 1.3 nm, the hydrophobic bands (i.e. the CG ethyl particles) are initially within the bilayer, for $\Delta z = 2.4$ nm roughly half of the ethyl particles are within the bilayer, and for $\Delta z = 3.5$ and 4.4 nm the hydrophobic particles are outside of the bilayer. The behaviour in these five simulations was measured by tracking the CoM of the DNT hydrophobic region relative to that of the bilayer (Figure 4.12BC). In the case of the smaller displacements ($\Delta z = 0.5$, 1.3, and 2.4 nm) the DNT relaxes back to its initial position within ~40 ns. In contrast, for the larger displacements ($\Delta z = 3.5$ and 4.4 nm), which completely removed the hydrophobic band from the bilayer core, the DNT exits from and diffuses away from the bilayer into the bulk solvent. This suggests that there are two (meta)stable states of the system, one with the DNT spanning the membrane and one with the DNT free of the membrane.

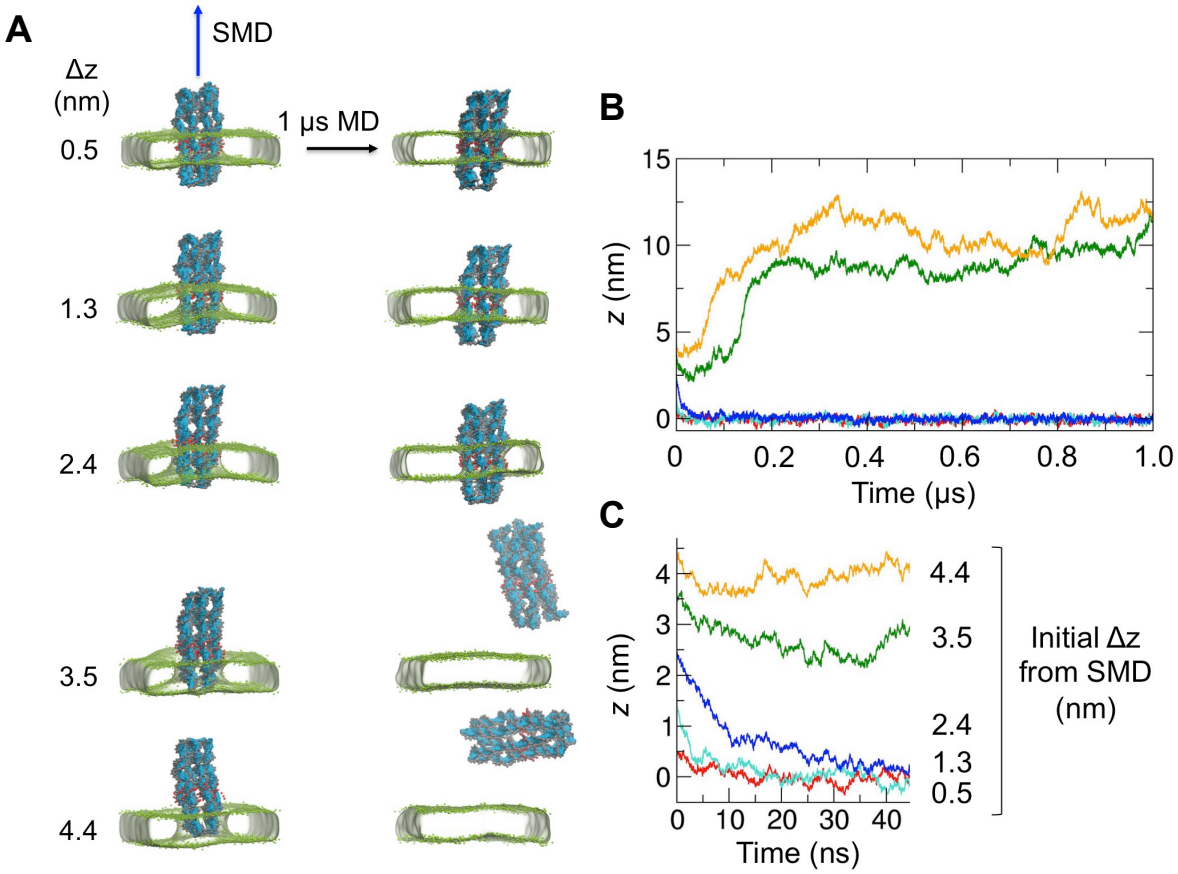

Figure 4.12: **Displacing the DNA nanopore relative to the bilayer**. **(A)** Five displaced states were generated by using steered MD to displace DNT^{Et} by Δz along the bilayer normal relative to its equilibrium position in the membrane. The left hand column shows images of the five initial displaced structures (DNT in blue/grey/red; PC bilayer as a surface defined by the green phosphate particles). Each displaced structure was used as the starting point for a 1 μs equilibrium simulation, yielding the final structure snapshots shown in the right hand column. Note that for Δz = 0.5 and 1.3 nm, the red Et particles are initially within the bilayer, for Δz = 2.4 nm about half of the Et particles are within the bilayer, whereas for Δz = 3.5 and 4.4 nm the particles are outside of the bilayer. **(B)** Position along z (the bilayer normal) of the CoM of the DNT hydrophobic region relative to the CoM of the bilayer as a function of time for the five displaced simulations shown in (**A**). **(C)** An expanded version of the same graph corresponding to the first 45 ns of each simulation. The colour code is: Δz = 0.5 (red), 1.3 (cyan), 2.4 (blue), 3.4 (green) and 4.4 (yellow) nm. SMD was performed by pulling, at the rate of 10 nm/μs, the CoM of the modified groups in the DNT along the *Z*-axis relative to the CoM of membrane. The pulling force constant was set to 1000 $kJmol^{-1}nm^{-2}$.

To explore the nature of the barrier separating these two states I performed further SMD simulations, comparing the DNA nanopore both with ($DNT^{Et}PC$) and without (DNT PC) the hydrophobic band on its surface (Table 4.1). In these simulations the DNT was initially located outside the bilayer, and then was pushed into and across the bilayer at a constant velocity of 10 nm/μs (Figure 4.13A). The DNT experiences some resistance (as evident from the bilayer curvature) as it enters the bilayer, then adopts a transbilayer configuration, and then experiences further resistance as it exits from the other side of the bilayer. The force profile (Figure 4.13B) is remarkably consistent between the 3 repeat simulations and shows a gradual build up of force to ~1000 kJ mol^{-1} nm^{-1} (equivalent to a single molecule force of ~1700 pN) as the DNT is pushed into the membrane. Note that this force is approximately the same whether or not the DNT has a hydrophobic surface band (Figure 4.13BC). It is apparent that in resisting entry to the bilayer the DNT twists away from the bilayer normal. If the DNT is restrained so that its long axis is parallel to the bilayer normal (i.e. pushing it into the bilayer like a nanoneedle[163]; Figure 4.14A) the force profile is similar but the peak force on entry is reduced to ~800 kJ mol^{-1} nm^{-1}). Once inserted into the bilayer the force drops to zero. Then as the DNT is pushed through the bilayer and starts to exit, if there is a hydrophobic band then there is a further barrier. However, if the DNT does not have a hydrophobic band this barrier is absent and the DNT can be pushed right out of the bilayer without any resisting force. If the hydrophobic band is present, the barrier for exit from the bilayer is ~400 kJ mol^{-1} nm^{-1}. I also observe that this barrier is dependent on the thickness of the hydrophobic band (Figure 4.14BC). In general, it can be concluded that for the same DNT with different hydrophobicity levels, the insertion barrier remains about the same, but that the exit barrier, which would determine the stability of the inserted complex, is reduced as the thickness of the hydrophobic band thickness is reduced. The forces experienced on entry into and exit from the bilayer of the DNT are mirrored by local perturbation of the bilayer curvature (Figure 4.13D), the bilayer relaxing back to a plane once the DNT is inserted in a transmembrane orientation, and once again when the DNT has exited.

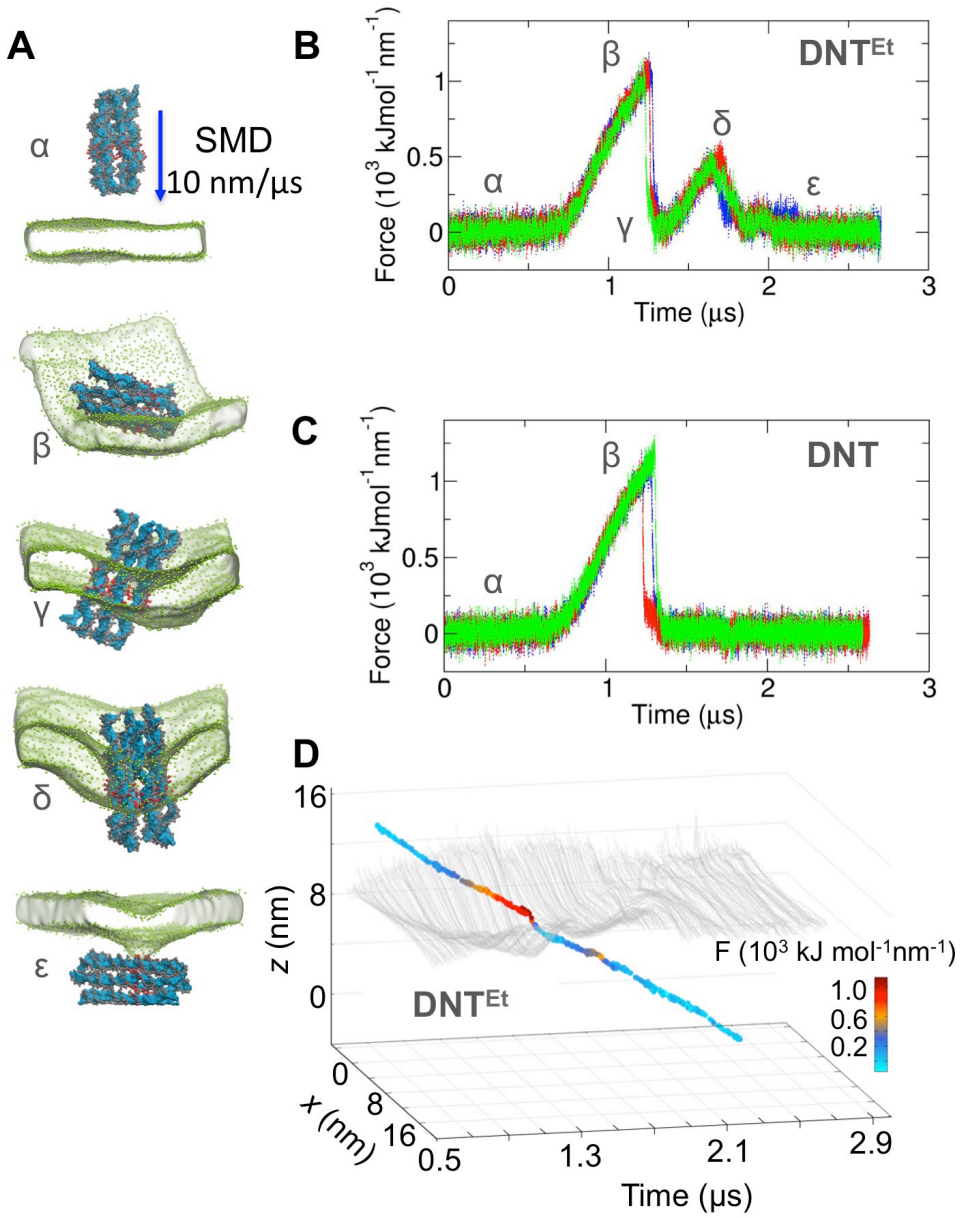

Figure 4.13: **Pushing DNA nanopores through a bilayer**. **(A)** Snapshots α to ε from a steered MD (SMD) simulation in which DNT^{Et} was moved at a rate of 10 nm/μs along the normal to a PC bilayer. In the snapshots shown the nanopore approaches the bilayer (α), interacts with the surface of the bilayer (β), penetrates and spans the bilayer (γ), is pulled out of the bilayer (δ), and finally exits the bilayer (ε). **(B)** A plot of force *vs*. time profile for the SMD simulations of DNT^{Et} illustrated in **(A)**. Three repeat simulations are shown in three different colours. **(C)** A comparable plot for SMD simulations of the unmodified (i.e. no Et groups) DNA nanopore. **(D)** A multi-dimensional plot showing the time evolution, as 4 ns time steps, for the DNT^{Et} simulation in **(A)**, showing membrane surface distortion (grey lines; calculated as the *XZ* projected CoM of the PC beads), the *XZ* projected coordinate of the CoM of the Et modified region of the DNT *vs*. time, with the points coloured by the corresponding force exerted. *SMD details:* A CG POPC bilayer was simulated for 200 ns, and the final snapshot was used as a starting

point for steered MD simulations. A DNT molecule (DNT, DNT^{Et}, DNT^{2AEt} or DNT^{1Et}) was placed above the membrane, with the minimum distance between POPC and DNT set to ~3.2 nm. The distance between the CoM of the membrane and of the hydrophobic region in the DNT was initially ~12 nm. The simulation box size was ~22 x 22 x 34 nm^3 (136,000 CG particles), and the bilayer/DNT system was solvated with water particles and 1 M NaCl. Positional restraints of 1000 $kJmol^{-1}nm^{-2}$ were applied to the DNT for 50 ns to equilibrate the water and ions around DNA and membrane. Positional restraints on the DNT were then removed, and steered MD was performed by pulling, at the rate of 10 nm/μs, the CoM of the modified groups in the DNT (or the same unmodified residues in plain DNT) along the *Z* axis relative to the CoM of membrane. The pulling force constant was set to 1000 $kJmol^{-1}nm^{-2}$. Three repeats for each simulation condition were performed with different random seeds for initial velocities. Using the same initial set ups, additional steered MD simulations were conducted by applying 100 $kJmol^{-1}nm^{-2}$ positional restraints in the *X* and *Y* dimensions on the DNT particles to avoid tumbling of the nanotube relative to the *Z*-axis. The corresponding plots are shown in this figure and Figure 4.14.

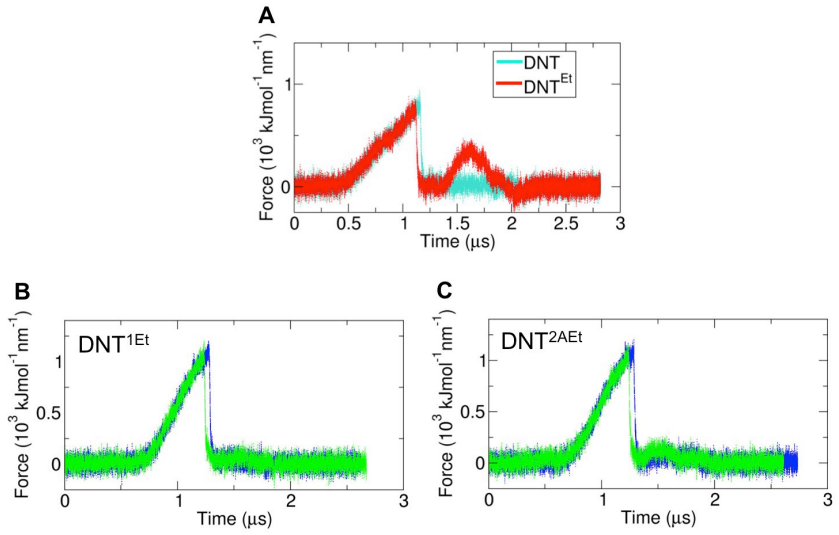

Figure 4.14: **Other steered MD simulations (push in)**. **(A)** Force *vs.* time curves shown for DNT (cyan) and DNT^{Et} (red) steered through a POPC bilayer with positional restraints in *XY* to avoid DNT tumbling. **(B)**, **(C)** Force *vs*. time curves for steered simulations of DNT^{1Et} and DNT^{2AEt}, respectively. Repeat simulations are shown in blue and green. See text for more details.

Furthermore, if one starts with a DNT pre-inserted into a bilayer and SMD is used to pull the DNT out of the bilayer, there is again a barrier of 400 to 500 kJ mol^{-1} nm^{-1} for exit from the bilayer (Figure 4.15).

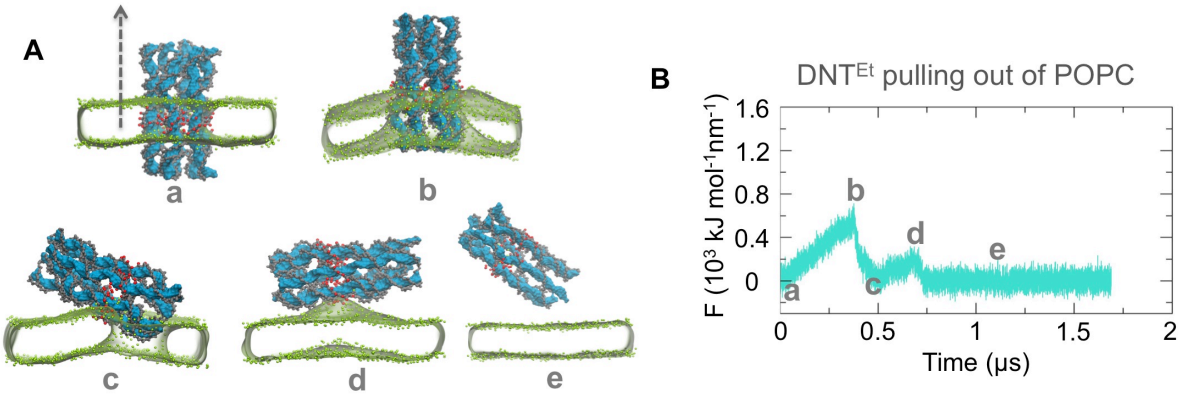

Figure 4.15: Steered MD simulations (pull out). **(A)** Snapshots a-e showing DNT^{Et} steered (i.e. pulled) out from a membrane embedded state at a rate of 10 nm/μs. **(B)** Force *vs*. time profile with the labels a to e corresponding to the snapshots in **(A)**. SMD simulation was performed on initially embedded DNT^{Et} nanopore which was pulled, along the Z-axis, out of the lipid bilayer. The initial configuration was obtained from a 1 μs CG simulation of $DNT^{Et}PC$. The pulling force constant was set to 1000 $kJmol^{-1}nm^{-2}$.

Thus, it may be concluded that the hydrophobic band does not facilitate the entry of the DNT^{Et} into the bilayer, but instead stabilizes the nanopore once in an inserted state to remain within the bilayer. It is interesting to compare the peak force for exit of the DNT^{Et} from the bilayer (~500 kJ mol^{-1} nm^{-1}, equivalent to ~800 pN) with estimates of up to ~300 pN for AFM unfolding and extraction of membrane proteins from lipids bilayers[164,165], whilst noting that there are several orders of magnitude of difference in the pulling rate in SMD and in the AFM experiments.

4.4) Hydrophobic anchors vs hydrobphobic bands

As discussed in Chapter 1, experimentally cholesterol and porphyrin rings can also be attached to DNT. Thus additionally, I have studied DNTs attached with porphyrin and cholesterol molecules anchored in the membrane. The overall DNT design (DNA strands) has been preserved to match the above cases, which is different from relevant experimental studies[65,66]. But the hydrophobic chemical attachments have been followed as in the experimental designs[65,66]. An iDNT model (DNT CG model obtained from initial built all-atom structure) was simulated (with the same parameters as above) in POPC membranes, as summarised in Table 4.2.

CG simulations	Duration
[1]iDNT^{2Chol}PC	6 x 1 μs
iDNT^{3Chol}PC	6 x 1 μs
[2]iDNT^{2TPP}PC	6 x 1 μs

Table 4.2: **Additional hydrophobic DNT cases simulated**.
[1]Chol: Cholesterol molecules (2-3) attached with polyethylene oxide linker.
[2]TPP: Tetraphenylporphyrin (2 modifications total)

I discuss briefly here MARTINI CG bead types used for the TPP and Chol hydrophobic anchors. The bonded parameters and the corresponding GROMACS files are available at *sites.google.com/site/maingivishal/scientific_data*

Cholesterol anchors.

In the experimental studies, Burns *et al*.[66] used cholesterol anchors, connected *via* triethylene glycol (TEG) linker, to stabilise DNA nanopores in the membrane. I have used the same linker and cholesterol molecules (2 or 3 in total) attached symmetrically around the iDNT surface.

I built a simplified computational model for the TEG linker attached to iDNT at one end and to the cholesterol at the other end (Figure 4.16). Cholesterol parameters were taken directly from MARTINI[166] (new version v2.0_CHOL_02). For a polyethylene oxide linker I used parameters from a reported computational study on polymers[167]. For the attachment of linker to cholesterol (hydroxyl end) and iDNT (O3' end) I have used only bond term (i.e. `[ bonds ]`). Figure 4.16 shows the linker with the attachment points along with the MARTINI bead types.

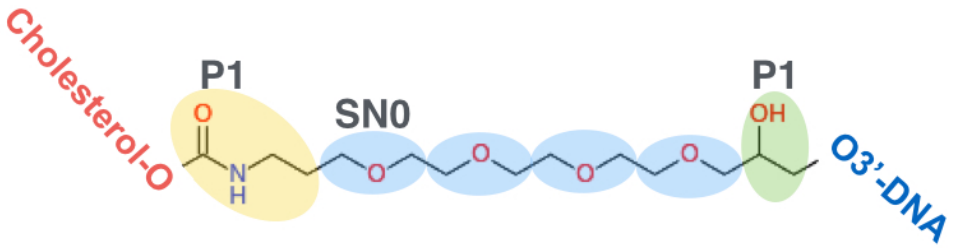

Figure 4.16: **Cholesterol linker**. TEG linker used for connecting cholesterol to the DNA. Transparent filled ovals represent CG particles and lines connect the overlapping atoms. The linker stick diagram was made using *ChemDoodle* free tool available at *https://web.chemdoodle.com/demos/sketcher/*

Tetraphenylporphyrin anchor.

In another experimental study, Burns *et al.*[65] attached two porphyrin rings (see Chapter 1; Figure 1.7) to the deoxyuridine base with an acetylene linker. Since TPP is a planar molecule, I used a simple CG model where the whole molecule was constrained to a planar configuration using only ENM (*i.e.* no angle or dihedral potentials). There was no ENM between TPP and any DNA bead, except for a connecting bond between them. Figure 4.17 shows the overlapped CG model beads on an all-atom TPP structure. The bead types were primarily derived from a Heme-B MARTINI model reported by de Jong *et al*[168]. I attached two TPP anchors on my iDNT design at DT bases at the opposite surface ends (Figure 4.18).

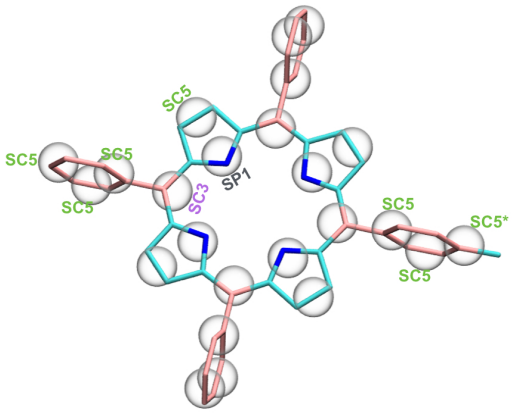

Figure 4.17: **TPP CG model.** Sticks connect the atomic coordinates, and transparent beads show the mapping to representative MARTINI CG particles for TPP, along with labels for bead types. An asterisk (*) labels the bead to which acetylene atoms are mapped, but the distances were adjusted to represent the atomic structure. All beads were connected by ENM only.

The iDNT^{2Chol}PC, iDNT^{3Chol}PC and iDNT^{2TPP}PC complexes were simulated for 1 μs with six repeats. They were all stable in membrane-buried simulations. As for the DNT$^{n\text{-}Et}$ cases, I was interested in calculating their tilt angles to determine the extent to which they 'wobble' in the membrane (Figure 4.18). As was the case for DNT^{Et} and DNT^{6Et}, they don't tilt substantially (compare Figure 4.6B and Figure 4.18B). In terms of tilting angle stability, it appears that the degree of hydrophobicity rendered by multiple Et groups is equivalent to 2 or 3 Chol or TPP anchors. But it should be noted that Chol and TPP anchors do not help in neutralizing the charged DNA surface.

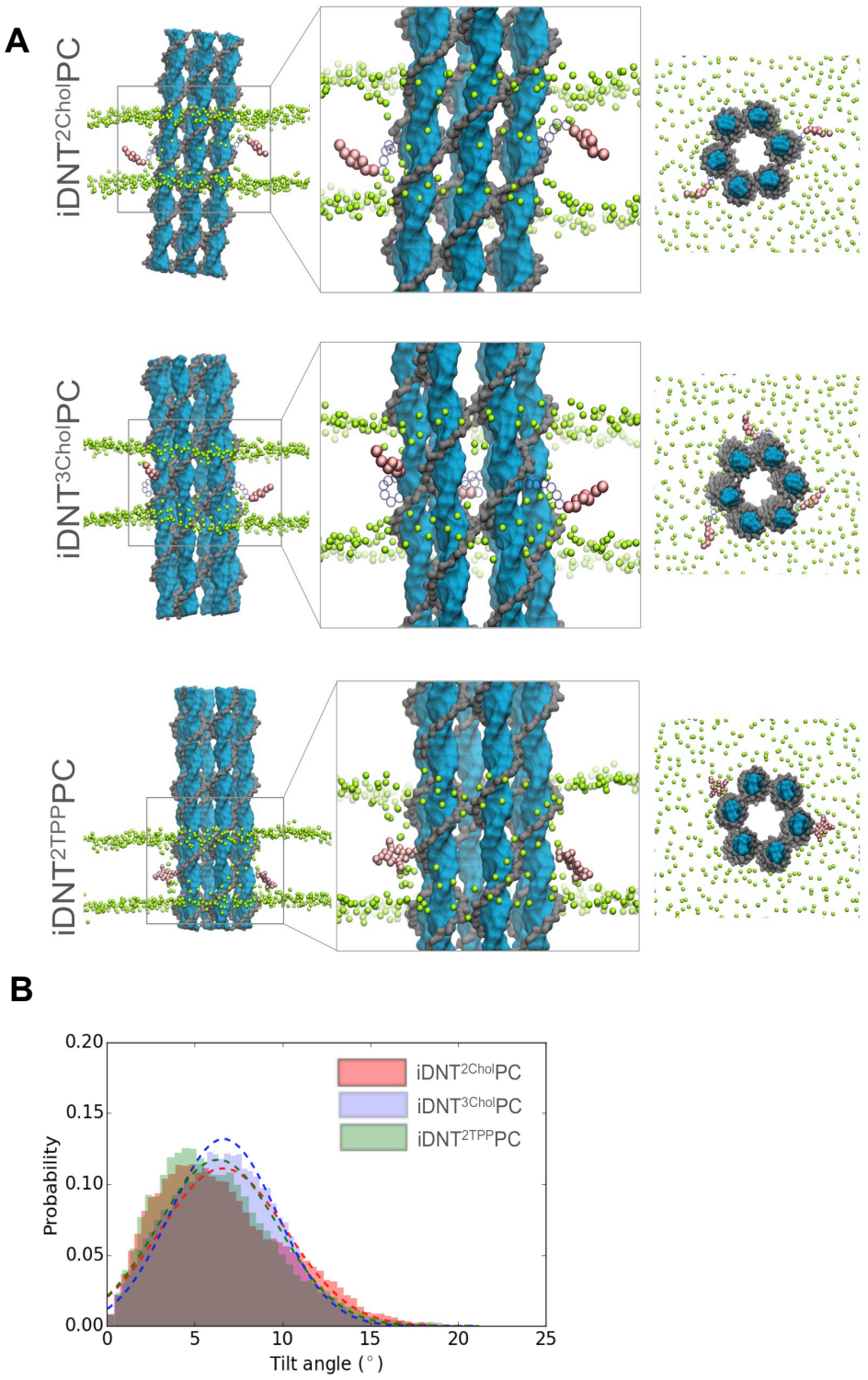

Figure 4.18: **Cholesterol and Porphyrin-decorated DNT simulations**. **(A)** Final 1 μs snapshots showing different DNT models embedded in a membrane (left panel *side* view; middle panel *zoom-in* view; right panel *top* view). See Table 4.2 for summary of simulations performed, and abbreviations used. Blue surface, DNA bases; grey surface, DNA backbone; green, POPC/phosphate head groups; transparent grey beads, linker; pink beads, Cholesterol/TPP. **(B)** DNT tilt angle distributions for the different DNT simulation systems.

4.5) Conclusions

I have performed coarse-grained simulations to explain the origin of the stability of DNA nanopores spanning phospholipid bilayers. Although the 'native' DNT carries a high negative charge density, modification with a central hydrophobic band stabilizes the DNT in a bilayer by providing an energetic barrier to exit of the inserted nanopore. It should be noted that this barrier is obtained from non-equilibrium steered MD simulations so should not be considered as a free energy barrier. Potential of mean force calculations using umbrella sampling techniques could help in revealing the free energy landscape of insertion. This barrier (from steered MD) was found to be dependent on the hydrophobicity rendered on the DNA surface. Instead of hydrophobic bands, where a number of ethyl modifications are required, only a few hydrophobic anchors (porphyrin or cholesterol) can be attached on the DNT surface. My results indicate that in both cases the membrane-buried DNT does not tilt substantially, and remains stably embedded in the membrane. It should be noted that, in contrast to the hydrophobic anchors, introduction of a hydrophobic band could help in reducing overall negative charge on the DNT surface to a large extent. There is a local reorganization of the bilayer structure so as to optimize interactions with the nanopore, by lipids interacting with the surface grooves of the DNA (also observed at all-atom resolution, as discussed in Chapter 5). This contrasts with e.g. a recent study[43] (experimental and computational) of a very large (22 nm) DNA 'porin', insertion of which, though stabilized by the presence of cholesterol anchors, results in substantive local reorganization of the bilayer and 'leakage' of ions and water around the edges of the inserted nanopore.

These studies lay the foundation to understand membrane-buried DNA nanopore systems at an all-atom level. The MARTINI CG DNA model cannot be used to observe any substantial conformational changes, which are quite important and anticipated in DNA-origami systems. Thus, in Chapter 5 I cover DNA nanopore-membrane interactions at atomic resolutions under different ionic strengths.

CHAPTER 5

5 All-atom simulations of DNA nanotube interactions with lipid bilayer membranes

The contents of this chapter are taken directly and/or modified from my paper accepted in Nature Communications (in press). Additional analyses have also been incorporated.

5.1) Introduction

In the case of membrane-buried nanopores it becomes essential to understand the ion and water distribution patterns along the channels revealing their selectivity and gating properties. All-atom (AT) resolution simulations are generally helpful in understanding the dynamics and effects of salt concentrations on the structural properties of systems[169]. The structural dynamics were quite evident in Chapter 3 where I studied AT DNA nanotube (DNT) simulations in an aqueous solution. In Chapter 4, I discussed DNA nanopore-membrane CG simulations using the MARTINI model[114,115]. This model normally involves an elastic network between constituent CG particles, which enables the maintenance of secondary structure in the absence of explicit hydrogen bond representations, but at the cost of structural rigidity. Thus, the CG MARTINI approach is not well suited to study conformational changes[114] that could occur while achieving an equilibrated state, or to study the effect of, for example, different salt concentrations. Also, MARTINI ions are modelled to represent hydrated states and thus their bigger size along with the less flexible DNA model does not enable observation of any possible leakage of ions that could occur through the DNT walls. Nonetheless, CG simulations are quite helpful in the case of membrane embedded structures where long time scales are usually required to assemble lipids properly around the transmembrane (TM) domains.

Starting with an equilibrated CG simulation model of the DNT^{6Et} in a lipid bilayer, I backmapped[117] to a more detailed atomistic representation in order to explore the conformational dynamics of the membrane-buried nanopore with

greater sensitivity. In Chapter 3 for DNT (without hydrophobic modifications) all-atom simulations I used 1M KCl salt concentrations, which was more relevant to the experimental conditions[57]. Further I was interested to see the influence of low salt concentrations, which is more relevant to the physiological saline conditions for therapeutic purposes. Thus, I simulated a membrane inserted nanopore (DNT^{6Et}) with two different ionic strengths: 0.15 M NaCl (corresponding approximately to normal physiological saline, and so facilitating comparison with experiments on nanopores inserted into cell membranes[87]), and 1.0 M NaCl (as in the CG simulations), enabling comparison with *in vitro* experiments on ionic currents through nanopores in artificial membranes[57]. Each simulation, containing ~0.83 x 10^6 atoms, was run for 0.5 μs.

5.2) Methods

5.2.1) Modelling

The final snapshot of the DNT^{6Et}-PC CG simulation was converted to an all-atom configuration using *backward.py*[117]. DNA, POPC, and water were back mapped to all-atom representations (CG ions were mapped as all-atom water), and following this the ions were added to reach 0.15 M or 1.0 M NaCl concentration by randomly removing some water molecules. In order to determine, and validate whether backmapping process using *backward.py* can recover the original model I performed RMSD analysis. All-atom structure obtained from cluster analysis on DNT1 (also used for CG simulations; Chapter 4) was coarse-grained and then again backmapped immediately (*i.e.* no further simulations) to all-atom using *backmap.py*. Corresponding calculated RMSD between two all-atom structures (original from cluster analysis *vs*. immediately backmapped) is 0.25 nm. Thus, *backmap.py* was not able to recover completely the original structure (from cluster analysis) as zero RMSD will be expected in such case. I have also done similar analysis on relatively simple dsDNA crystal structure (PDBID: 1BNA), and found that *backmap.py* could not completely recover the original coordinates in this case as well (RMSD in this case was 0.29 nm). Additionally, I calculated RMSD for backmapped all-atom DNT obtained from coarse-grain DNT-membrane simulated complex (DNT^{6Et}PC; Chapter 4). In this case calculated RMSD between two all-

atom DNT structures (original from cluster analysis *vs*. backmapped CG simulated DNT-membrane complex) is 0.28 nm. Slightly higher RMSD in this case is expected as DNT CG model is not completely 'stiff' as some fluctuations of CG beads is expected to occur during simulations which is reflected in the backmapped structure. As CG simulations use stiff elastic networks to retain the configuration of protein or DNA, so in general one can backmap lipids and waters, and for protein or DNA the original all-atom structure can be fitted (rotational-translational RMSD fitting) on membrane-embedded CG structure. But in my case I have backmapped DNA also.

Ethyl modifications were made to selected residues based on the experimental design[57] positions (Chapter 4, Figure 4.3), using ethyl-phosphates (instead of ethyl-phosphorothioate)[92]. Chemical modifications were done at the 'O1P' atom of DNA bases thus forming all '*S*' chiral phosphate centers. But it is to be noted that experimentally[57] the chemical substitution reaction was not chiral specific. Parameters for the ethyl-phosphate modified nucleotides were derived with CGenFF[170], and other parameters were from CHARMM36[111]. Details regarding charge adjustments done for the modified DNA bases have been provided in the Appendix.

5.2.2) Simulations

Atomistic MD simulations were performed using GROMACS[171] v5.0.2. The initial system configuration was energy minimized using steepest descents. Velocities were generated with 100 ps NVT simulation at 310 K. During this step positional restraints on DNA and lipids were applied in *xyz* with a force constant of 1000 kJ $mol^{-1}nm^{-2}$. A subsequent NPT simulation was performed and positional restraints were removed slowly over ~2 ns. An *unrestrained* NPT simulation of up to 0.5 μs was then performed. Time steps of 1 fs and 2 fs was used in the NVT and NPT simulations, respectively. Bond lengths were constrained using the LINCS algorithm[172]. A velocity rescale thermostat[107] was used to maintain temperature at 310 K with a 0.1 ps coupling time constant. The Parrinello-Rahman algorithm[109] was used to maintain semi-isotropic pressure at 1 bar with a 1 ps time constant. The neighbour list was updated after every 10 steps with a 1.2 nm cut-off. For

Coulombic interactions PME[173] was used with a real space cut-off of 1.2 nm, while the van der Waals interactions were switched to zero between 0.8 and 1.2 nm.

5.3) Results and discussion

5.3.1) Conformational changes under different salt conditions

At 0.15 M NaCl the DNT undergoes a substantial change in conformation from the starting structure (the backmapped structure obtained from coarse-grained simulations DNT^{6Et}-PC case, Chapter 4) compared to the simulation at 1.0 M (Figure 5.1A; also see Figure 5.2). This comparison can be quantified by measuring the root mean square displacement (RMSD) of the DNA heavy (i.e. not H) atoms relative to the starting structure as a function of time (Figure 5.1B). For both the TM and the outside-the-membrane (solvent-exposed) regions of the DNA, as expected, the RMSD shows an initial jump as the structure from the (ENM restrained) CG simulations relaxes. At both ionic strengths the RMSD is then approximately constant for the TM region of the nanopore, indicative of its conformational stability in the membrane environment. The RMSD is higher for the outside solvent-exposed region, suggesting greater structural deformation, especially in the 0.15 M NaCl simulation. The greater conformational stability of the outside region in the presence of 1.0 M NaCl suggests that the increased ionic strength helps to stabilize the densely packed DNA-origami structures. Published computational studies of a comparable nanopore in 1.0 M KCl[92] showed a comparable degree of conformational flexibility, as is also observed for the un-derivatized DNT in the absence of a membrane[91] (discussed in Chapter 3). The lipid organization around the nanopore in the atomistic simulations (Figure 5.1C) remains similar to what was observed at the CG level, with no apparent effect from the salt concentration—in line with the similar RMSD of the TM parts.

For both salt concentrations, RMSD profiles suggest that the DNA transmembrane domain achieves equilibrium after 0.2 μs. Thus, all the time averaged properties reported here have been calculated over the last 300 ns.

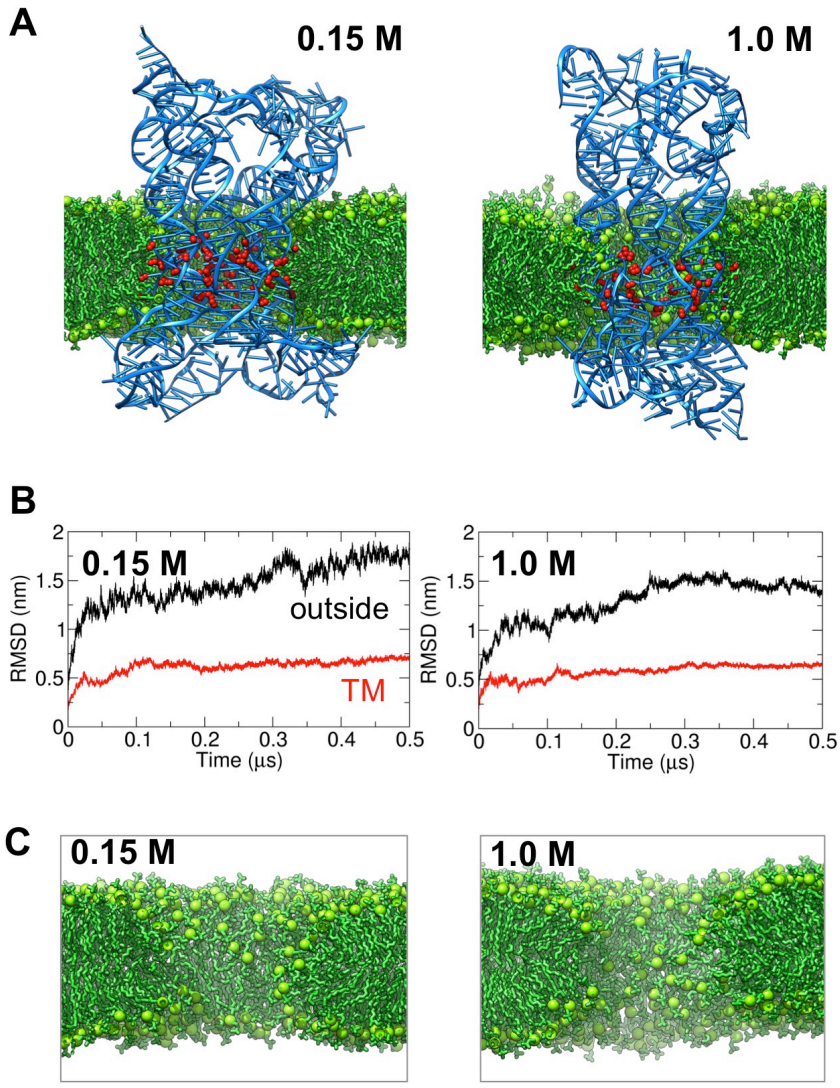

Figure 5.1: **Atomistic simulations of a DNA nanopore embedded in a POPC bilayer.** Snapshots and graphs on the left- and right-hand sides are for simulations in 0.15 M and 1.0 M NaCl, respectively. **(A)** Final (0.5 μs) snapshots for simulation AT-DNT^{6Et} (green sticks and balls = lipid tails and head groups, respectively; blue ribbon ladders = DNA; and red beads = ethyl chains; water and ions omitted for clarity). **(B)** Root mean square displacements (RMSD) for the DNA heavy (i.e. not H) atoms relative to the starting structure as a function of time. RMSDs are shown separately for the transmembrane (TM; red) and outside-the-membrane (black) regions of the DNA. **(C)** Zoomed in views of the same systems, showing only the lipids and revealing the local perturbation of the bilayer by insertion of the DNA nanopore. Compare with Figure 4.4B, Chapter 4.

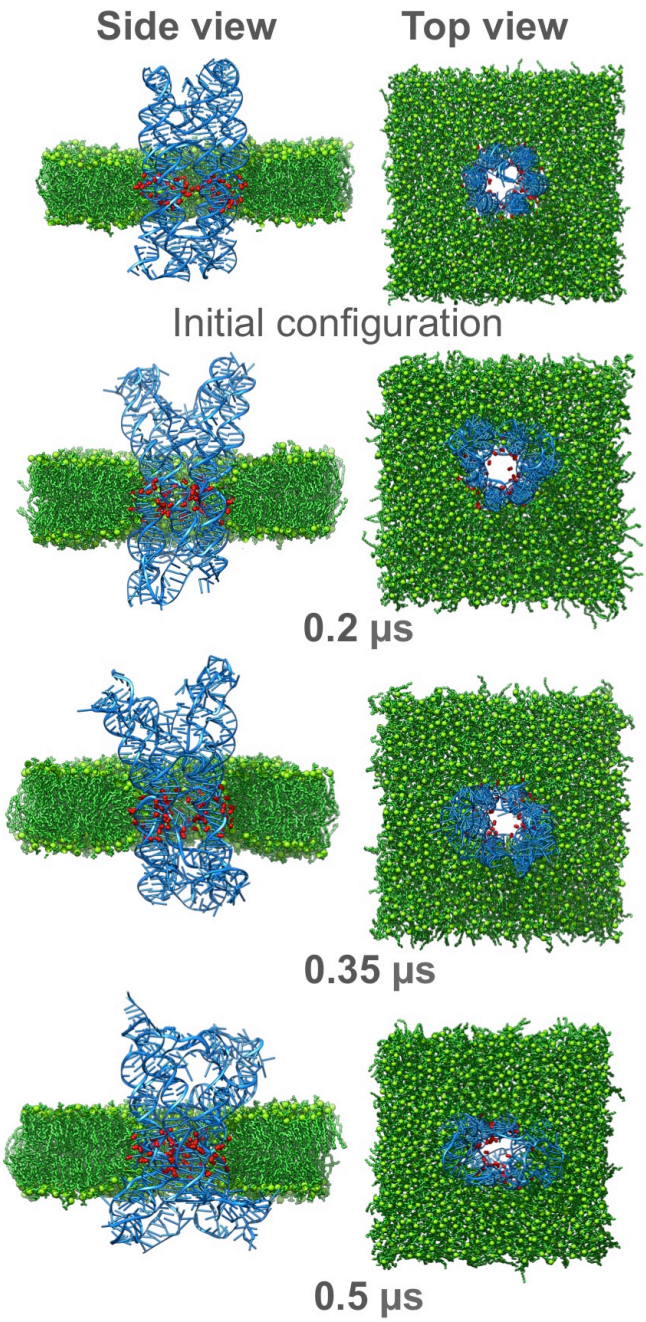

Figure 5.2: **All atom DNT-membrane simulation (0.15 M)**. Successive snapshots are shown for the AT DNT^{6Et}PC simulation. Ethyl groups are shown as *red* spheres, DNA is represented as a *blue* ladder, and POPC molecules are shown in *green*. The simulation was performed in 0.15 M NaCl (waters, ions are not shown for clarity). It can be seen that even though the outermembrane DNA conformation stability is lost progressively but the central lumen pore remains open. My simulation stresses the importance of extended simulation time scales during which such conformational changes could be observed compared to otherwise small duration simulations. Such progressive conformational changes were not observed for 1M case.

5.3.2) Pore characteristics and lipid interactions

In my previous simulations of the un-derivatized DNT in aqueous solution (all-atom simulations discussed in Chapter 3), I observed time dependent fluctuations in the pore radius, which I suggested might provide a model for pore gating[91]. I therefore explored the effect of the bilayer and the ionic strength on fluctuations of the radius of the transmembrane pore by evaluating the pore radius profile as a function of time for each of the simulations (Figure 5.3AB).

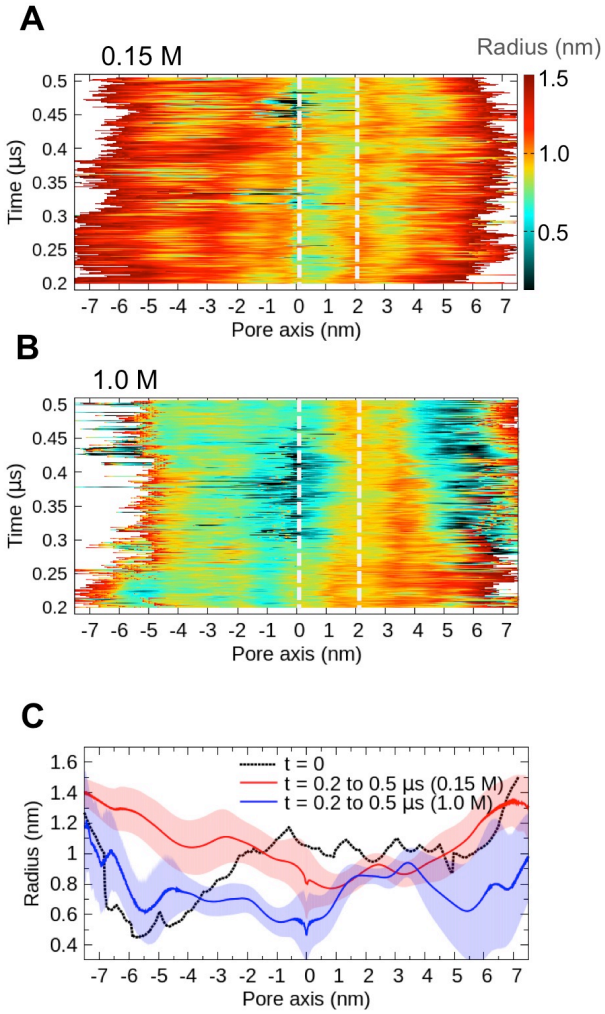

Figure 5.3: **Pore dynamics from the atomistic simulations.** The pore radius profiles (calculated using HOLE[138]) of the ethyl modified hydrophobic DNT are shown as a function of time for the simulations in **(A)** 0.15 M and **(B)** 1.0 M NaCl. Vertical dashed white lines represent the approximate location of the bilayer projected onto the pore axis. This was obtained by juxtaposing radius profiles for each frame after 20 ps. **(C)** Time averaged pore radius profiles for the simulations in (red) 0.15 M and (blue) 1.0 M NaCl. The error bars correspond to one standard deviation of the mean. The pore radius profile of the initial structure is shown in black.

At the lower ionic strength (0.15 M) the radius profile suggests that in the external regions the pore lumen remains open (radius ~1 nm) for most of the simulation time; the TM region is a little narrower at radius ~0.8 nm but also remains open throughout the simulation. At the higher ionic strength (1.0 M) the fluctuations in the pore radius are more pronounced and overall there is a degree of contraction of the pore relative to the starting structure (Figure 5.3C). Thus, the time dependent radius profile reveals opening and closing events ('gating' motions) both close to the 'upper' mouth of the TM pore (i.e. at z ~ 0 nm in Figure 5.3B) and in the 'lower' outside-the-membrane region (z ~ +5 nm). These are similar to the pore radius fluctuations reported in my previous study of DNT simulations in 1.0 M KCl aqueous solution[91] (discussed in Chapter 3). These differences between the behaviour of the pore at the two ionic strengths are reflected in their time-averaged pore radius profiles (Figure 5.3C).

The DNT studied here was obtained from backmapping CG DNT^{6Et} where the DNT originally comes from the cluster analysis of DNT1 simulation (Chapter 3). Thus it is interesting, and more relevant, to compare radius profiles of the membrane embedded DNT1 with the free aqueous solution simulations both conducted at 1 M salt concentrations. The time dependent radius profile of DNT1 (Figure 3.17C) reveals that there is a large degree of opening and closing events at one end of the lumen *i.e.* near G1-G2 regions compared to G5-G6 regions. The central part (G3-G4) of lumen remains open. It is quite evident in the membrane-DNT1; 1 M case, that the central region (G4 part of DNT1), which is embedded in the membrane (and modified as ethyl-phosphates) becomes more constricted at one end (near ~0 nm pore axis; Figure 5.3B) compared to the other transmembrane end (near ~2 nm pore axis; Figure 5.3B). Considering the terminal ends gating-'like' motions, it was observed to a greater extent at the lumen end (near ~5-6 nm pore axis; Figure 5.3B) more near to the membrane surface (G5-G6 in this case compared to G1-G2 end in aqueous solution DNT1 simulations). Thus, the gating like motions observed in my simulations are influenced by the nature of environment (aqueous *vs*. membrane embedded).

In each simulation the time averaged distributions of ions (Figure 5.4A) indicate that the pore is *cation* selective (It should be noted that I have not applied any

voltage across the membrane-embedded DNT, and the ion selectivity profile for the conducting state under the applied voltage should not be compared directly). At 0.15 M the ratio of Na^+ to Cl^- in the lumen of the pore is ~14; at the higher ionic strength it is ~2. Although some hydrophobic ethyl groups are located in the luminal lining of the pore this is not sufficient to cause hydrophobic gating[73] (Figure 5.4B) and so fully hydrated ions and water molecules remain within the lumen of the nanopore throughout the simulations. In summary, in terms of pore properties, at lower ionic strength the DNA nanopore has a high cation selectivity and a higher probability of being open.

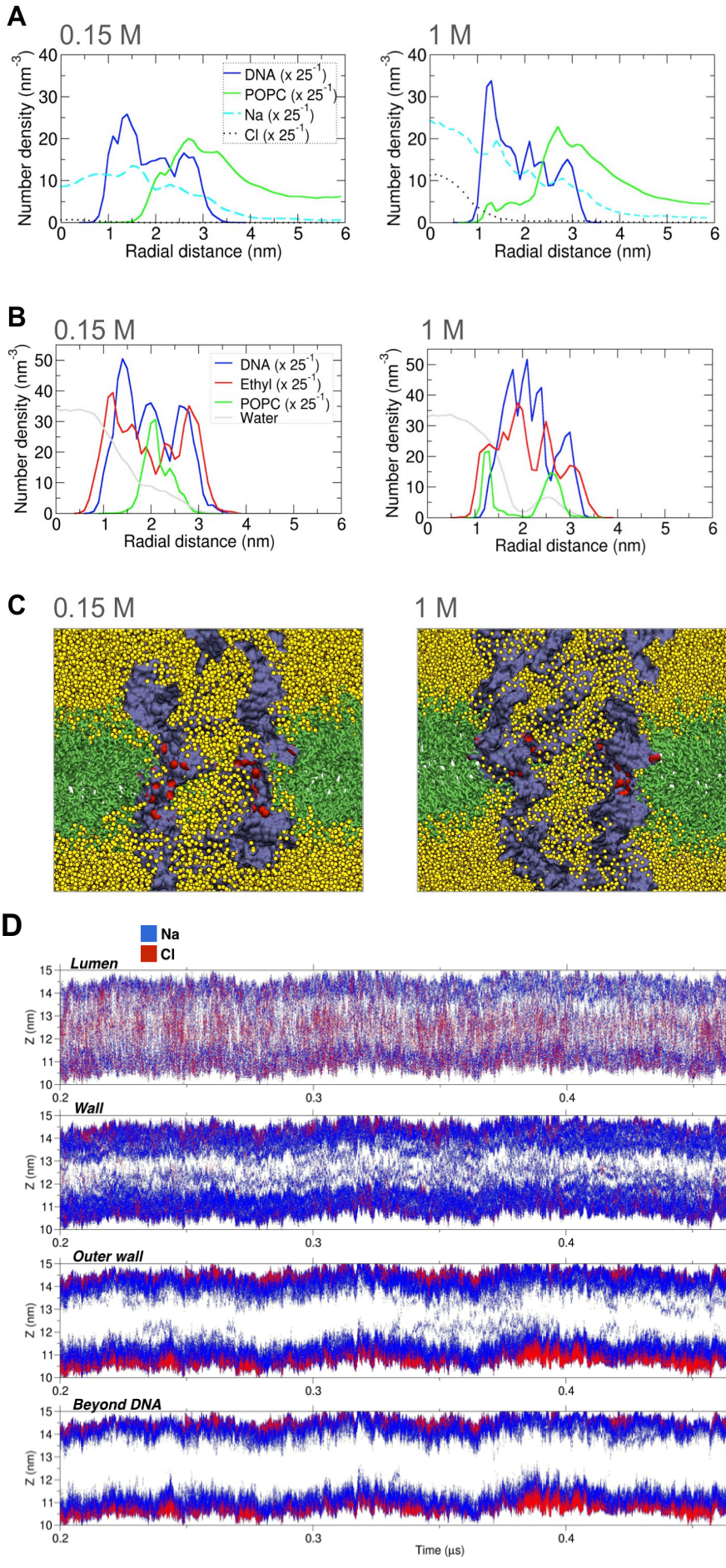
A
0.15 M
1 M
Number density (nm^{-3})
Radial distance (nm)
DNA (x 25^{-1})
POPC (x 25^{-1})
Na (x 25^{-1})
Cl (x 25^{-1})
B
0.15 M
1 M
Number density (nm^{-3})
Radial distance (nm)
DNA (x 25^{-1})
Ethyl (x 25^{-1})
POPC (x 25^{-1})
Water
C
0.15 M
1 M
D
Na
Cl
Lumen
Wall
Outer wall
Beyond DNA
Z (nm)
Time (μs)

Figure 5.4: **Radial density distributions**. **(A)** Radial density distribution of DNA phosphates, lipid phosphates and ions for the AT DNT^{6Et} simulations at 0.15 and 1.0 M NaCl. These distributions correspond to a cylinder of height 4.1 nm which therefore extends from the centre of the bilayer just beyond the headgroup region of the lipids on each side. **(B)** Radial density distribution of DNA phosphates, ethyl group terminal carbons, lipid phosphates and waters for the same two simulations. These distributions correspond to a cylinder of height 0.5 nm which therefore covers the central region of the hydrophobic core of the bilayer. **(C)** Final sliced (along *Z*) snapshots illustrating that the distribution of water in the membrane plane is primarily through the pore lumen, and also the DNA walls but not major leakage outside the walls. *Yellow beads*; water oxygen atoms, *green sticks*; POPC, *blue surface*; DNA; *red beads*; ethyl modifications. **(D)** *Z* distribution/tracking of ions (line plot) with respect to time. Only transmembrane region of 4.1 nm thickness was considered to track the ions. See text for discussion.

Inspection of the distributions of ions and of water around the inserted nanotube (Figure 5.4B) suggests that there is *not* a substantial number of waters trapped between the nanopore and the surrounding lipid molecules (Figure 5.4C). Therefore, there is unlikely to be significant ion 'leakage' possible around the edges of the pore. To explore further I was interested to see the potential leakage of ions by tracking each and every ion for 1M case (as conformation is more stable, and experimentally more relevant in terms of ions concentration used[57]). For this I divided the transmembrane pore in four regions based on DNA number density shoulder peaks (Figure 5.4A; 1M case): *lumen* (0 to < 1.3 nm radial distance), *wall* (1.3 to < 2.9 nm radial distance), *outer wall* (> 2.9 nm radial distance), *beyond DNA* (> 3.4 nm radial distance). In each region, *Z* coordinate value of each ion was extracted, and the sequence of points at each frame (time) can be considered as a line plot of *Z* coordinates (projected on a line). It should be noted that it is difficult to define a 'wall' thus ranges so chosen can lead to different results in the *wall* region (1.3 to < 2.9 nm). It can be clearly seen that ions are mostly found in the *lumen* and in the *wall* regions but far less in the *outer wall* region. In terms of defining the pore to be leaky, the ions can either leak from lumen to the wall or can leak directly while remaining outside the wall region (edges of the pore). I observe that there is no leakage of ions in *beyond DNA* region (except small stochastic burst at 0.38 μs) suggesting that lipid plugs-in the

outer wall region as seen the density profiles also. Thus, the presence of ions in *wall* and *outer wall* establishes the porous nature of the DNT-embedded in a membrane but at the same time it is not leaky as seen in *beyond DNA* regions.

This contrasts with a recently simulated cholesterol supported large conductance DNA 'porin'[43] for which simulations suggest ca. 20% of the ionic current may flow around the *outside* of the channel. Though system size was considerably large and computationally expensive I believe that this system was not simulated long enough to allow lipid molecules assemble and equilibrate around the DNA porin. This could be the reason for leakage of current at the DNA and membrane interface as observed in their computational simulations. My coarse-grained simulation studies suggest that it takes around 0.6 μs for the lipids to equilibrate around the embedded DNA pore. It is to be noted that this is in CG time scale and if one has to initiate simulation at the all-atom resolution then similar equilibrium could be achieved in ~ 4x CG time scale.

Turning to the lipids, the atomistic radial densities (Figure 5.4) are comparable to those from the CG simulations (Figure 4.4C; Chapter 4), and confirm a degree of penetration of the 'walls' of the DNT by lipid molecules (seen as overlap of lipid and DNA radial density). However, the lipids do not occlude the pore lumen. The atomistic simulations also provide an opportunity to examine in more detail the interactions of lipid molecules with the inserted DNT. Visual inspection of the membrane proximal to the nanotube (Figure 5.1C) suggests that the bilayer is locally perturbed in a similar fashion to that observed in the CG simulations (Figure 4.4B; Chapter 4), with some lipid headgroups remaining within the core of the membrane adjacent to the DNT surface.

Examination of radial distribution functions (Figure 5.5) provides clear evidence for favourable interactions of the anionic phosphate DNA backbone and the positively charged choline groups of the zwitterionic lipid headgroups. Thus, local reorganization of the lipids occurs so as to stabilize the DNA nanopore within the bilayer core. More detailed examination of snapshots from the simulation (Figure 5.5D) shows that lipid tails remain within the major and minor grooves of the membrane spanning DNA segments.

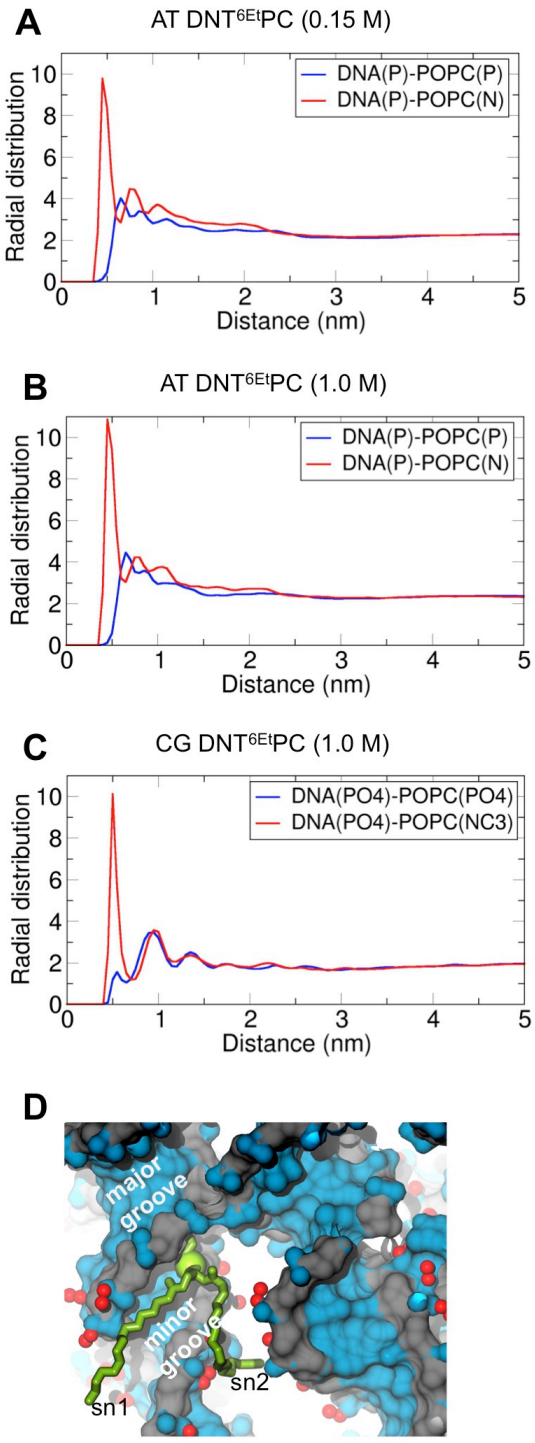

Figure 5.5: **Radial distributions of lipids around the DNA nanopores embedded within the bilayer**. Radial distribution functions (RDFs) for DNA-phosphorus to PC headgroup phosphorus (blue) or, and DNA-phosphorus to PC headgroup choline nitrogen (red) for the AT-DNT^{6Et} simulations in 0.15 M **(A)** and 1.0 M NaCl **(B)**. **(C)** The equivalent RDFs for the CG DNT^{6Et} simulation, for DNA-phosphate particles to PC headgroup phosphates (blue), and DNA-phosphates to PC headgroup choline particles (red). Only negatively charged (i.e. not ethylated) DNA bases near the membrane surface were included in this analysis in both the cases. **(D)** A snapshot showing a lipid molecule in the minor groove of DNT (1.0 M NaCl case).

5.3.3) Area per lipid and bilayer thickness

To check the overall stability of the membranes during the simulations I calculated the area per lipid and bilayer thickness values (Figure 5.6). The effect of high salt concentration (1.0 M NaCl) is clearly observed as a lowered area per lipid and increased membrane thickness.

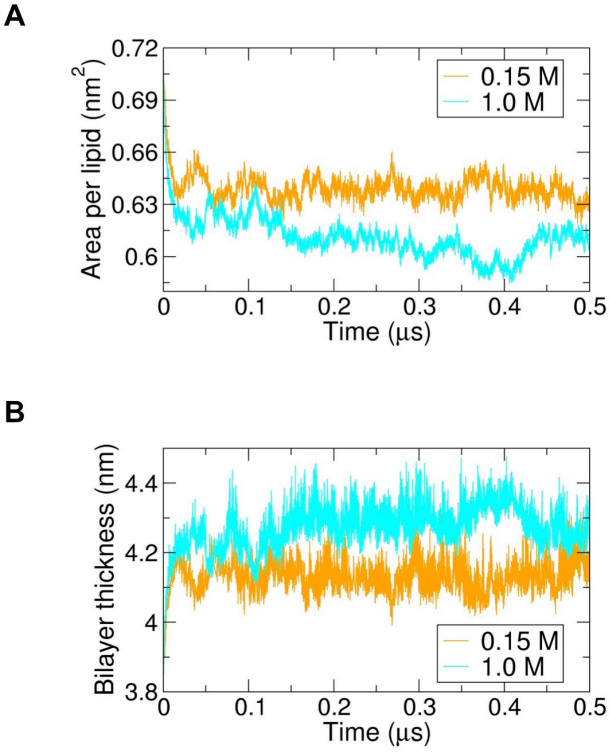

Figure 5.6: POPC bilayer area per lipid and thickness values for DNT^{6Et}PC all-atom simulations at 0.15 M and 1.0 M NaCl. **(A)** Area was averaged for both leaflets separately. **(B)** Bilayer thickness calculated (P-P distance) along the evolution of simulation for both the cases. This analysis was performed using APL@Voro[174].

5.4) Conclusions

The DNA nanopore is dynamic, and responsive to both the bilayer[92] and aqueous environment. The open state of the central pore is stable, against fluctuations (leading to partial pore closure), at low (physiological) ionic strength. Although at low ionic strength (0.15 M) there are substantive fluctuations in the conformation of the DNT regions outside of the bilayer, the interactions with the lipids protect the TM pore against collapse, as evidenced by a stable pore radius profile in this region. In contrast, at higher ionic strength (1.0 M), the fluctuations in the DNT

are lower overall, but more pronounced fluctuations in the pore radius are observed. Furthermore, the ion selectivity of the nanopore (as evidenced by the ratio of cations to anions in the pore lumen) shows a dependence on ionic strength. Also, I noticed that the ions do not leak or cross outside the DNA wall. In a comparable DNT design computational study, Yoo *et al.*[92] also did not observe ion leakage but they additionally applied external voltage. As discussed in Chapter 4, in the reported study where cholesterol anchored DNA 'porin'[43] was studied, ion leakage was observed across the DNA walls.

Over the last decade DNA-origami has provided complex structures, and their wide spectrum of potential applications include biomedicine[67] and therapeutics[175]. For example, DNA-origami nanotubes have been reported to exhibit cytotoxic properties[87], which may be used for targeted anticancer applications, *e.g.*, by attaching cancer cell specific moieties (aptamers[176], folic acid[177] attached to variety of nanoparticles) to the DNA architectures. The insights obtained from all-atom simulations performed here, along with CG simulations, inform design of the next generation of DNTs to selectively permeabilize cell membranes for potential biotherapeutic applications[87]. In particular, the interplay between the role of the lipids and of the electrolyte in stabilizing an open DNA nanopore within a bilayer offers intriguing possibilities for design of novel environment-sensitive gating of DNTs. My all-atom simulation studies at low salt concentration suggest that the DNA nanopore loses conformational stability compared to the high salt concentrations. Low salt concentrations are more relevant and comparable to the physiological conditions. Thus, such DNA nanostructures formed by close packed DNA helices are expected to lose structural rigidity if used for physiological applications. This indicates that an improvement is required in the design of DNA nanostructures for therapeutic purposes, which may be achieved using better positioning the staple strands with reduced length of DNA pores.

CHAPTER 6

6 Summary, conclusions and prospects

I have studied DNA nanotubes (DNT) and their interaction with membranes at different resolutions using computational molecular dynamics techniques. In each chapter I drew some conclusions but this chapter, based on the conclusions, will also focus on some limitations, and thus further improvements that can be considered to study DNTs.

6.1) Summary

In Chapter 3, I discussed all-atom molecular dynamics (MD) simulations of a DNA nanotube (DNT) pore scaffold based on an experimental design[57]. This helped me to gain fundamental insights into its atomic structure, dynamics, and interactions with ions and water. My multiple simulations of models of DNTs that are composed of a six-duplex bundle led to a coherent description. In brief, the central tube lumen adopts a cylindrical shape while the mouth regions at the two DNT openings undergo gating-like motions which provide a possible molecular explanation of a lower conductance state observed in previous experimental study on a membrane-spanning version of the DNT[83]. Similarly, the central nanotube lumen filled with water and ions was characterized by bulk diffusion coefficients. I also observed that the porous nature of the DNT walls allows lateral leakage of ions and water.

Apart from the understandings that could be gained from their aqueous studies, DNTs provide a key example of how DNA nano-assemblies may be inserted into a membrane *via* attachment of hydrophobic anchors. However, the molecular mechanism whereby such anchors stabilize a polyanionic assembly spanning the hydrophobic core of a lipid bilayer remain unclear. In Chapters 4 and 5 I discussed multiscale MD simulations to explore the stability and dynamics of a DNA nanopore inserted into a phospholipid bilayer. Coarse-grained (CG) MD revealed that the inserted nanopore exhibits close interactions of the lipids and the DNA resulting in a degree of local reorganization of the lipid bilayer to

accommodate the nanopore. Steered simulations provided a force profile as the DNT is pushed into and subsequently out of a lipid bilayer by translation along the bilayer normal, revealing both a barrier to entry and a barrier to exit. The exit barrier (using non-equilibrium steered MD) is due to the hydrophobic band on the surface. This barrier should not considered as free energy barrier which may be determined using potential of mean force calculations. Atomistic MD simulations at physiological (0.15 M) and at elevated (1.0 M) salt concentrations revealed that in both cases the nanopore remains stable inserted in the bilayer on a microsecond simulation timescale, with greater flexibility of the extra-membranous DNA region seen at lower ionic strength. Lipids pack closely around the nanotube at both ionic strengths. The pore was found to be cation selective, more so at the lower ionic concentration. Thus, the DNA nanopore exhibits both selectivity and gating properties which are sensitive to the ionic strength of the surrounding aqueous solution.

6.2) Using improved forcefields

Atomistic

Yoo *et al*.[178] suggested some improvements that are required for non-bonding (Lennard-Jones) parameters specifically for amine-phosphate and amine-carboxylate interactions. Using these parameters one can expect to have influence on protein-protein, protein-DNA, protein-membrane, DNA-DNA and DNA-membrane interactions. Recently Cheatham *et al*.[112] exhaustively compared and analysed, on aggregated simulation timescale in milliseconds, mainly the *bsc0* (developed by Pérez *et al*.[110]), *bsc1* (developed by Ivani *et al*.[113]), *OL15* (developed by Zgarbová *et al*.[179]), and *CHARMM36* (developed by Hart *et al*.[111]) forcefields. It should be noted that these forcefields have been improved over the generations[112] for the bonding term (specifically torsion potentials) interactions, and all the references have not been provided here. Cheatham *et al*.[112] suggested the community to use either *bsc1* or *OL15* forcefields for future DNA simulations.

The improved forcefields were published and reported when I had already started my project. I have used *bsc0* (Chapter 3) and *CHARMM36* (Chapters 3 and 5) for

my studies. So for the future studies one can consider using newer and improved forcefields.

Coarse-grain

oxDNA developed by Šulc *et al*.[88] has been used for origami like DNA nanostructures and it seems to be quite promising as one can understand various self-assembly processes[180] within a reasonable computational time. It is to be noted that this forcefield does not consider any explicit waters and ions. In contrast to the atomistic forcefields, oxDNA is not very sensitive to the effect of different molar salt concentrations on the stability of complexes. My major interest was to simulate DNA-membrane interactions (Chapters 4 and 5) so I chose CG MARTINI[114,115], which allows combining lipid and DNA forcefields together which is currently not available with the oxDNA forcefield. Although with this forcefield one can use explicit CG water and ions but one cannot observe any conformational change in the DNA, and the self-assembly processes cannot be studied. Thus, the provided conformation remains as such rigid because of the extensive elastic networks used in the forcefield.

Considering the above stated drawbacks of CG MARTINI, I have conducted preliminary investigations aimed at developing a 'flexible' general CG MARTINI DNA model which ideally would reproduce all-atom forcefield properties. To achieve this I used PyCGTOOL developed by Graham *et al*.[181] (unpublished but available, see reference[181]) at the University of Southampton in Prof. Syma Khalid's lab. This tool uses the all-atom trajectories to provide equilibrium bonding term force constants for the CG beads mapping on all-atom structures defined by the user. The tool performs Boltzmann inversion on the bond distributions obtained from the all-atom trajectory to calculate the equilibrium values and force constants. It should be noted that one has to define not only the CG beads but also has to define the bonds, angles, and dihedrals maps between the beads for which the force constants are calculated. Thus, it is a trial and experience based approach in order to design the best map, and to select optimal CG beads and their types. Nonetheless, Graham *et al*. have successfully tested it for small all-atom organic molecules to obtain coarse-grain parameters. I simulated 20 base pairs AT dsDNA fragment in an aqueous solution (see flow diagram Figure 6.1A). The structure was then coarse-grained with the MARTINI

DNA bead types definition[114], and using PyCGTOOL force constants for the bond maps (bond, angles and dihedrals defined by me) between the beads were derived from the AT trajectory. I attempted a number of different bonded maps to keep the base stacking, and the helical nature intact. I could attain base stacking and helical nature, and flexibility in the structure (see Figure 6.1B) but not with full success. Backmapped structure, all-atoms, shows that the bases are stacked but attain a buckle shaped alignment. Also, the main drawback was that I had to keep the time step for my CG simulation very small from 1 to 2 fs. This time step is comparable to AT simulation thus not advantageous in terms of faster MD calculations. This suggests significant future work is needed to introduce the required flexibility in the current MARTINI DNA model.

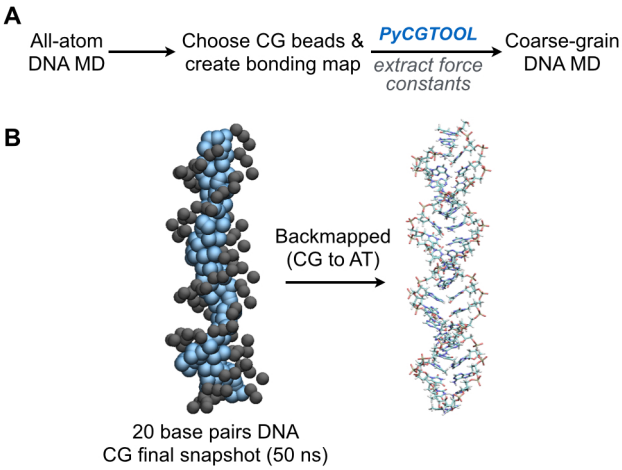

Figure 6.1: **Flexible MARTINI DNA CG model possible approach**. **(A)** Simple flow chart diagram showing the steps to use all-atom trajectory to obtain force constants for the defined CG beads map (see text for details). **(B)** Final 50 ns CG snapshot of 20 base pairs DNA. *Grey beads*; backbone, *blue beads*; base pairs. Also shown on the right side of arrow the backmapped all-atom DNA (*stick* model) obtained from coarse-grained final snapshot on the left. Clearly the base pairs are stacked but align in buckle shape.

6.3) High molar concentrations

In Chapter 5, I showed that DNT systems are not conformationally stable under physiological salt concentrations. Also, understanding from the fact that most of the DNA-origami experiments require high salt concentrations these systems cannot be used directly for therapeutic applications. Though it has been reported that some systems are (or can be made) stable in physiological applications

depending on the shape of the origami[182]. Also, the DNA nanostructures obtained using the polyhedral mesh approach[60] (discussed in Chapter 1) are expected to be more stable as the DNA helices are less densely packed compared to the square or hexagon lattice based DNA nanostructures. Also, depending on the experiments, in some of the MD simulation studies Mg^{+2} ions[89], or combination of both Na^{+} and Mg^{+2} ions[94] have been used to attain DNA nanostructure conformational stability. Nonetheless, one cannot ignore their resistance to nucleases.

6.4) Final remarks

Recent advances in lipid bilayer embedded DNA based nanopores definitely broadens the spectrum of available natural or engineered protein porins, and other nonbiogenic pores. I hope my study will help with the design of DNA nanopores with enhanced stability features for sensing applications possibly with tunable gating properties. Also, insights from my work describe the next generation DNA nanotubes that can be designed to permeabilize cell membranes for therapeutic purposes. Further work on computational studies of DNA nanopores with chemically functionalised internal lumens may help experimentalists to know and understand their functional properties beforehand or better.

Also, it will be interesting to explore other hydrophobic anchors for DNA nanopores. As an original idea, there is the possibility that the DNA can be embedded in the membrane without the need of any hydrophobic anchors. It is to be noted that DNA has a hydrophobic core formed by purine and pyrimidine bases. Thus, a mismatch in any base pair will cause purine or pyrimidine rings to flip outside that can interact with the bilayer membrane hydrophobic core. For example, Figure 6.2 shows a mismatch in the middle of the DNA, which may result in fraying of bases in the middle of DNA, thus allowing flipped-out bases to interact with the membrane.

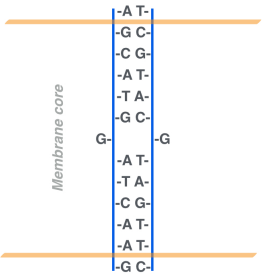

Figure 6.2: Cartoon representation of an idea to embed dsDNA in the membrane without the need of any hydrophobic anchors. *Blue lines*; DNA backbone, *Orange lines*; bilayer membrane boundaries with hydrophobic membrane core. Bases in the middle can be mismatched so that instead of forming hydrogen bonds they can flip outside. Note that I have chosen here two DG bases not only due to mismatch for hydrogen bonding but also due to their size they cannot geometrically fit inside the helix. Similarly it will be good to choose any two purine bases (two DA or DA and DG bases). It should be noted that it is just a representative design and does not mean to be experimentally relevant. For example, this can be considered as one of many such units of staple or scaffold strands in a full design. Also, one can consider using extra sticky-free ends having purine bases.

7 Appendix

Charge adjustments for modified DNA bases

I would like to acknowledge Justin Lemkul for helping me to adjust the charges of modified DNA bases.

Experimentally ethylphosphorothioate DNA bases have been used. Since it involves a C-S-P bond, there are no reliable parameters available for the same[92]. Yoo *et al.*[92] used ethylphosphate (*i.e.* C-O-P) for which reliable parameters were derived from CGenFF[170]. They did not provide details about the parameters and charges used for the modified DNA bases. I used the following protocol and provide here details giving an example of DA base.

To avoid deriving all the parameters for a modified DNA base I derived charges and parameters for a ethlyphosphate cap only. Finally, it needs some charges and parameters adjustments as new atom types and additional bonding terms are introduced. In brief following steps were used for charges adjustment:

1) I designed a ethylphosphate analogue cap shown in Figure A1.

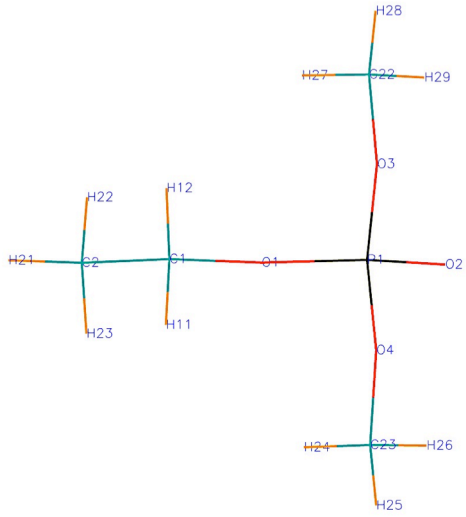

Figure A1: Ethylphosphate cap analogue shown with atom names.

2) Submit this molecule to CGenFF webserver (*cgenff.paramchem.org*).

3) It is to be noted that except for C22 methyl and C23 methyl groups, all charges can be directly used for modified DNA base. Also, C22 and C23 can be considered as C3’ and C5’ respectively found in DA base (Figure A2).

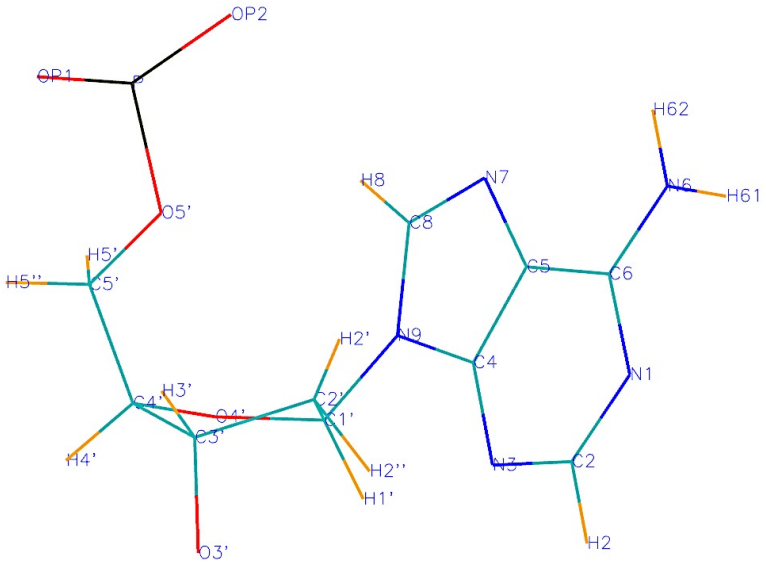

Figure A2: Normal, unmodified, DA molecule with atom names.

4) C3' and H3' have partial charges as:

```
C3'   0.01
H3'   0.09
```

Total charge of this group needs to be 0.1

C22 methyl group has partial charges as:

```
C22   -0.169
H27   0.09
H28   0.09
H29   0.09
```

Total charge of this group needs to be 0.101
Thus new partial charge for C3’ is given by: `0.101-0.09 = 0.011`

5) Similarly, C3' and H3' have partial charges as:

```
C5'    -0.08
H5'1   0.09
H5'2   0.09
```

Total charge of this group needs to be 0.1

C23 methyl group has partial charges as:

```
C23   -0.169
H24   0.09
H25   0.09
H26   0.09
```

Total charge of this group needs to be 0.101
Thus new partial charge for C5’ is given by: `0.101-0.09-0.09 = -0.079`

Bonding terms were taken from CGenFF though it required minor adjustments for the attached ethyl chain depending on the atom types. Final GROMACS format related files are available here *sites.google.com/site/maingivishal/scientific_data*
For an example and a comparison, I have shown in Table A1 a normal DA base and a modified DA base charges.

Normal DA		**Ethylphosphate DA**		
P	1.5000	P	1.9940	CgenFF
O1P	-0.7800	O1P	-0.5340	CgenFF
O2P	-0.7800	O2P	-0.6400	CgenFF
O5'	-0.5700	O5'	-0.5530	CgenFF
C5'	-0.0800	C5'	-0.0790	Adjusted
H5'1	0.0900	H5'1	0.0900	
H5'2	0.0900	H5'2	0.0900	
C4'	0.1600	C4'	0.1600	
H4'	0.0900	H4'	0.0900	
O4'	-0.5000	O4'	-0.5000	
C1'	0.1600	C1'	0.1600	
H1'	0.0900	H1'	0.0900	
N9	-0.0500	N9	-0.0500	
C5	0.2800	C5	0.2800	
N7	-0.7100	N7	-0.7100	
C8	0.3400	C8	0.3400	
H8	0.1200	H8	0.1200	
N1	-0.7400	N1	-0.7400	
C2	0.5000	C2	0.5000	
H2	0.1300	H2	0.1300	
N3	-0.7500	N3	-0.7500	
C4	0.4300	C4	0.4300	
C6	0.4600	C6	0.4600	
N6	-0.7700	N6	-0.7700	
H61	0.3800	H61	0.3800	
H62	0.3800	H62	0.3800	
C2'	-0.1800	C2'	-0.1800	
H2'1	0.0900	H2'1	0.0900	
H2'2	0.0900	H2'2	0.0900	
C3'	0.0100	C3'	0.0110	Adjusted
H3'	0.0900	H3'	0.0900	
O3'	-0.5700	O3'	-0.5530	CgenFF
Sum	-1.0000	CX1 (C1)	-0.0960	CgenFF
		HZ1 (H11)	0.0900	CgenFF
		HZ2 (H12)	0.0900	CgenFF
		CX2 (C2)	-0.2700	CgenFF
		HZ3 (H21)	0.0900	CgenFF
		HZ4 (H22)	0.0900	CgenFF
		HZ5 (H23)	0.0900	CgenFF
		Sum	0.0000	

Table A1: Comparison of partial atomic charges of DNA un-modified DA *vs* ethylphosphate DA. Charges which were taken directly from CGenFF or adjusted (see main text) are labelled accordingly. Rest of the charges are taken from CHARMM36.

8 References

(1) Horton, H. Robert; Moran, Laurence A.; Ochs, Raymond S.; Rawn, J. David; Scrimgeour, K. G. *Principles of Biochemistry*; 2nd ed.

(2) van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane Lipids: Where They Are and How They Behave. *Nat Rev Mol Cell Biol* **2008**, *9*, 112–124.

(3) van Meer, G.; de Kroon, A. I. P. M. Lipid Map of the Mammalian Cell. *J. Cell Sci.* **2010**, *124*, 5–8.

(4) Ikezawa, H. Glycosylphosphatidylinositol (GPI)-Anchored Proteins. *Biol. Pharm. Bull.* **2002**, *25*, 409–417.

(5) Saier, M. H.; Tran, C. V; Barabote, R. D. TCDB: The Transporter Classification Database for Membrane Transport Protein Analyses and Information. *Nucleic Acids Res.* **2006**, *34*, D181–D186.

(6) Jap, B. K.; Walian, P. J. Structure and Functional Mechanism of Porins. *Physiol. Rev.* **1996**, *76*, 1073–1088.

(7) Nakae, T. Identification of the Outer Membrane Protein of E.coli That Produces Transmembrane Channels in Reconstituted Vesicle Membranes. *Biochem. Biophys. Res. Commun.* **1976**, *71*, 877–884.

(8) Conlan, S.; Zhang, Y.; Cheley, S.; Bayley, H. Biochemical and Biophysical Characterization of OmpG: A Monomeric Porin. *Biochemistry* **2000**, *39*, 11845–11854.

(9) Peraro, M. D.; van der Goot, F. G. Pore-Forming Toxins: Ancient, but Never Really out of Fashion. *Nat Rev Micro* **2016**, *14*, 77–92.

(10) Yildiz, Ö.; Vinothkumar, K. R.; Goswami, P.; Kühlbrandt, W. Structure of the Monomeric Outer-membrane Porin OmpG in the Open and Closed Conformation. *EMBO J.* **2006**, *25*, 3702 LP-3713.

(11) Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J. E. Structure of Staphylococcal α-Hemolysin, a Heptameric Transmembrane Pore. *Science (80-.).* **1996**, *274*, 1859–1865.

(12) Doyle, D. A.; Cabral, J. M.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. The Structure of the Potassium Channel: Molecular Basis of K⁺ Conduction and Selectivity. *Science (80-.).* **1998**, *280*, 69 LP-77.

(13) Kuang, Q.; Purhonen, P.; Hebert, H. Structure of Potassium Channels. *Cell. Mol. Life Sci.* **2015**, *72*, 3677–3693.

(14) Sansom, M. S. P.; Shrivastava, I. H.; Bright, J. N.; Tate, J.; Capener, C. E.; Biggin, P. C. Potassium Channels: Structures, Models, Simulations. *Biochim. Biophys. Acta - Biomembr.* **2002**, *1565*, 294–307.

(15) Du, J.; Lu, W.; Wu, S.; Cheng, Y.; Gouaux, E. Glycine Receptor Mechanism Elucidated by Electron Cryo-Microscopy. *Nature* **2015**, *526*, 224–229.

(16) Bayley, H.; Cremer, P. S. Stochastic Sensors Inspired by Biology. *Nature* **2001**, *413*, 226–230.

(17) Chen, M.; Khalid, S.; Sansom, M. S. P.; Bayley, H. Outer Membrane Protein G: Engineering a Quiet Pore for Biosensing. *Proc. Natl. Acad. Sci.* **2008**, *105*, 6272–6277.

(18) Butler, T. Z.; Pavlenok, M.; Derrington, I. M.; Niederweis, M.; Gundlach, J. H. Single-Molecule DNA Detection with an Engineered MspA Protein Nanopore. *Proc. Natl. Acad. Sci.* **2008**, *105*, 20647–20652.

(19) Manara, R. M. A.; Jayne Wallace, E.; Khalid, S. DNA Sequencing with MspA: Molecular Dynamics Simulations Reveal Free-Energy Differences between Sequencing and Non-Sequencing Mutants. *Sci. Rep.* **2015**, *5*, 12783.
(20) Howorka, S.; Siwy, Z. Nanopore Analytics: Sensing of Single Molecules. *Chem. Soc. Rev.* **2009**, *38*, 2360–2384.
(21) Keyser, U. F. Controlling Molecular Transport through Nanopores. *J. R. Soc. Interface* **2011**.
(22) Bezrukov, S. M.; Vodyanoy, I.; Parsegian, V. A. Counting Polymers Moving through a Single Ion Channel. *Nature* **1994**, *370*, 279–281.
(23) Krasilnikov, O. V; Sabirov, R. Z.; Ternovsky, V. I.; Merzliak, P. G.; Muratkhodjaev, J. N. A Simple Method for the Determination of the Pore Radius of Ion Channels in Planar Lipid Bilayer Membranes. *FEMS Microbiol. Immunol.* **1992**, *5*, 93–100.
(24) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. *Proc. Natl. Acad. Sci.* **1996**, *93*, 13770–13773.
(25) Reiner, J. E.; Balijepalli, A.; Robertson, J. W. F.; Campbell, J.; Suehle, J.; Kasianowicz, J. J. Disease Detection and Management via Single Nanopore-Based Sensors. *Chem. Rev.* **2012**, *112*, 6431–6451.
(26) Rincon-Restrepo, M.; Mikhailova, E.; Bayley, H.; Maglia, G. Controlled Translocation of Individual DNA Molecules through Protein Nanopores with Engineered Molecular Brakes. *Nano Lett.* **2011**, *11*, 746–750.
(27) Astier, Y.; Braha, O.; Bayley, H. 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.* **2006**, *128*, 1705–1710.
(28) Schneider, G. F.; Dekker, C. DNA Sequencing with Nanopores. *Nat Biotech* **2012**, *30*, 326–328.
(29) Howorka, S.; Movileanu, L.; Braha, O.; Bayley, H. Kinetics of Duplex Formation for Individual DNA Strands within a Single Protein Nanopore. *Proc. Natl. Acad. Sci.* **2001**, *98*, 12996–13001.
(30) Clarke, J.; Wu, H.-C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H. Continuous Base Identification for Single-Molecule Nanopore DNA Sequencing. *Nat Nano* **2009**, *4*, 265–270.
(31) Cherf, G. M.; Lieberman, K. R.; Rashid, H.; Lam, C. E.; Karplus, K.; Akeson, M. Automated Forward and Reverse Ratcheting of DNA in a Nanopore at 5-A Precision. *Nat Biotech* **2012**, *30*, 344–348.
(32) Manrao, E. A.; Derrington, I. M.; Laszlo, A. H.; Langford, K. W.; Hopper, M. K.; Gillgren, N.; Pavlenok, M.; Niederweis, M.; Gundlach, J. H. Reading DNA at Single-Nucleotide Resolution with a Mutant MspA Nanopore and phi29 DNA Polymerase. *Nat Biotech* **2012**, *30*, 349–353.
(33) Gurnev, P. A.; Oppenheim, A. B.; Winterhalter, M.; Bezrukov, S. M. Docking of a Single Phage Lambda to Its Membrane Receptor Maltoporin as a Time-Resolved Event. *J. Mol. Biol.* **2006**, *359*, 1447–1455.
(34) Wolfe, A. J.; Mohammad; Cheley, S.; Bayley, H.; Movileanu, L. Catalyzing the Translocation of Polypeptides through Attractive Interactions. *J. Am. Chem. Soc.* **2007**, *129*, 14034–14041.
(35) Borisenko, V.; Lougheed, T.; Hesse, J.; Füreder-Kitzmüller, E.; Fertig, N.;

Behrends, J. C.; Woolley, G. A.; Schütz, G. J. Simultaneous Optical and Electrical Recording of Single Gramicidin Channels. *Biophys. J.* **2003**, *84*, 612–622.
(36) Mayer, M.; Semetey, V.; Gitlin, I.; Yang, J.; Whitesides, G. M. Using Ion Channel-Forming Peptides to Quantify Protein−Ligand Interactions. *J. Am. Chem. Soc.* **2008**, *130*, 1453–1465.
(37) Mahendran, K. R.; Niitsu, A.; Kong, L.; Thomson, A. R.; Sessions, R. B.; Woolfson, D. N.; Bayley, H. A Monodisperse Transmembrane α-Helical Peptide Barrel. *Nat Chem* **2016**, *advance on*.
(38) Li, J.; Stein, D.; McMullan, C.; Branton, D.; Aziz, M. J.; Golovchenko, J. A. Ion-Beam Sculpting at Nanometre Length Scales. *Nature* **2001**, *412*, 166–169.
(39) Wanunu, M.; Meller, A. Chemically Modified Solid-State Nanopores. *Nano Lett.* **2007**, *7*, 1580–1585.
(40) Yusko, E. C.; Johnson, J. M.; Majd, S.; Prangkio, P.; Rollings, R. C.; Li, J.; Yang, J.; Mayer, M. Controlling Protein Translocation through Nanopores with Bio-Inspired Fluid Walls. *Nat Nano* **2011**, *6*, 253–260.
(41) Bell, N. A. W.; Engst, C. R.; Ablay, M.; Divitini, G.; Ducati, C.; Liedl, T.; Keyser, U. F. DNA Origami Nanopores. *Nano Lett.* **2012**, *12*, 512–517.
(42) Hall, A. R.; Scott, A.; Rotem, D.; Mehta, K. K.; Bayley, H.; Dekker, C. Hybrid Pore Formation by Directed Insertion of [Alpha]-Haemolysin into Solid-State Nanopores. *Nat Nano* **2010**, *5*, 874–877.
(43) Göpfrich, K.; Li, C.-Y.; Ricci, M.; Bhamidimarri, S. P.; Yoo, J.; Gyenes, B.; Ohmann, A.; Winterhalter, M.; Aksimentiev, A.; Keyser, U. F. Large-Conductance Transmembrane Porin Made from DNA Origami. *ACS Nano* **2016**, *10*, 8207–8214.
(44) Krishnan, S.; Ziegler, D.; Arnaut, V.; Martin, T. G.; Kapsner, K.; Henneberg, K.; Bausch, A. R.; Dietz, H.; Simmel, F. C. Molecular Transport through Large-Diameter DNA Nanopores. *Nat. Commun.* **2016**, *7*, 12787.
(45) Bayley, H. Designed Membrane Channels and Pores. *Curr. Opin. Biotechnol.* **1999**, *10*, 94–103.
(46) Braha, O.; Walker, B.; Cheley, S.; Kasianowicz, J. J.; Song, L.; Gouaux, J. E.; Bayley, H. Designed Protein Pores as Components for Biosensors. *Chem. Biol.* **1997**, *4*, 497–505.
(47) Howorka, S. Changing of the Guard. *Science (80-.).* **2016**, *352*, 890–891.
(48) DeGrado, W. F.; Summa, C. M.; Vincenzo, P.; Flavia, N.; Angela, L. De Novo Design and Structural Characterization of Proteins and Metalloproteins. *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
(49) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature* **2016**, *537*, 320–327.
(50) Joh, N. H.; Wang, T.; Bhate, M. P.; Acharya, R.; Wu, Y.; Grabe, M.; Hong, M.; Grigoryan, G.; DeGrado, W. F. De Novo Design of a Transmembrane Zn²⁺-Transporting Four-Helix Bundle. *Science (80-.).* **2014**, *346*, 1520 LP-1524.
(51) Thomson, A. R.; Wood, C. W.; Burton, A. J.; Bartlett, G. J.; Sessions, R. B.; Brady, R. L.; Woolfson, D. N. Computational Design of Water-Soluble α-Helical Barrels. *Sci.* **2014**, *346*, 485–488.
(52) Pfeifer, W.; Saccà, B. From Nano to Macro through Hierarchical Self-Assembly: The DNA Paradigm. *ChemBioChem* **2016**, *17*, 1063–1080.

(53) Seeman, N. C. Nucleic Acid Junctions and Lattices. *J. Theor. Biol.* **1982**, *99*, 237–247.
(54) Chen, J.; Seeman, N. C. Synthesis from DNA of a Molecule with the Connectivity of a Cube. *Nature* **1991**, *350*, 631–633.
(55) Saccà, B.; Niemeyer, C. M. DNA Origami: The Art of Folding DNA. *Angew. Chemie Int. Ed.* **2012**, *51*, 58–66.
(56) Rothemund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440*, 297–302.
(57) Burns, J. R.; Stulz, E.; Howorka, S. Self-Assembled DNA Nanopores That Span Lipid Bilayers. *Nano Lett.* **2013**, *13*, 2351–2356.
(58) Rothemund, P. W. K. Design of DNA Origami. In *Proceedings of the 2005 IEEE/ACM International Conference on Computer-aided Design*; ICCAD '05; IEEE Computer Society: Washington, DC, USA, 2005; pp. 471–478.
(59) Castro, C. E.; Kilchherr, F.; Kim, D.-N.; Shiao, E. L.; Wauer, T.; Wortmann, P.; Bathe, M.; Dietz, H. A Primer to Scaffolded DNA Origami. *Nat Meth* **2011**, *8*, 221–229.
(60) Benson, E.; Mohammed, A.; Gardell, J.; Masich, S.; Czeizler, E.; Orponen, P.; Hogberg, B. DNA Rendering of Polyhedral Meshes at the Nanoscale. *Nature* **2015**, *523*, 441–444.
(61) Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C. Design and Self-Assembly of Two-Dimensional DNA Crystals. *Nature* **1998**, *394*, 539–544.
(62) Rothemund, P. W. K.; Ekani-Nkodo, A.; Papadakis, N.; Kumar, A.; Fygenson, D. K.; Winfree, E. Design and Characterization of Programmable DNA Nanotubes. *J. Am. Chem. Soc.* **2004**, *126*, 16344–16352.
(63) Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H. DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires. *Science (80-.).* **2003**, *301*, 1882–1884.
(64) Liu, D.; Park, S. H.; Reif, J. H.; LaBean, T. H. DNA Nanotubes Self-Assembled from Triple-Crossover Tiles as Templates for Conductive Nanowires. *Proc. Natl. Acad. Sci. United States Am.* **2004**, *101*, 717–722.
(65) Burns, J. R.; Göpfrich, K.; Wood, J. W.; Thacker, V. V; Stulz, E.; Keyser, U. F.; Howorka, S. Lipid-Bilayer-Spanning DNA Nanopores with a Bifunctional Porphyrin Anchor. *Angew. Chemie Int. Ed.* **2013**, *52*, 12069–12072.
(66) Burns, J. R.; Seifert, A.; Fertig, N.; Howorka, S. A Biomimetic DNA-Based Channel for the Ligand-Controlled Transport of Charged Molecular Cargo across a Biological Membrane. *Nat Nano* **2016**, *11*, 152–156.
(67) Hernández-Ainsa, S.; Keyser, U. F. DNA Origami Nanopores: An Emerging Tool in Biomedicine. *Nanomedicine* **2013**, *8*, 1551–1554.
(68) Tagliazucchi, M.; Szleifer, I. Transport Mechanisms in Nanopores and Nanochannels: Can We Mimic Nature? *Mater. Today*.
(69) Liu, L.; Yang, C.; Zhao, K.; Li, J.; Wu, H.-C. Ultrashort Single-Walled Carbon Nanotubes in a Lipid Bilayer as a New Nanopore Sensor. *Nat Commun* **2013**, *4*.
(70) Chui, J. K. W.; Fyles, T. M. Ionic Conductance of Synthetic Channels: Analysis, Lessons, and Recommendations. *Chem. Soc. Rev.* **2012**, *41*, 148–175.
(71) Ghadiri, M. R.; Granja, J. R.; Buehler, L. K. Artificial Transmembrane Ion Channels from Self-Assembling Peptide Nanotubes. *Nature* **1994**, *369*,

301–304.
(72) Lear, J. D.; Wasserman, Z. R.; DeGrado, W. F. Synthetic Amphiphilic Peptide Models for Protein Ion Channels. *Science (80-.).* **1988**, *240*, 1177–1181.
(73) Trick, J. L.; Aryal, P.; Tucker, S. J.; Sansom, M. S. P. Molecular Simulation Studies of Hydrophobic Gating in Nanopores and Ion Channels. *Biochem. Soc. Trans.* **2015**, *43*, 146–150.
(74) Stoloff, D. H.; Wanunu, M. Recent Trends in Nanopores for Biotechnology. *Curr. Opin. Biotechnol.* **2013**, *24*, 699–704.
(75) Wang, Y.; Zheng, D.; Tan, Q.; Wang, M. X.; Gu, L.-Q. Nanopore-Based Detection of Circulating microRNAs in Lung Cancer Patients. *Nat Nano* **2011**, *6*, 668–674.
(76) Gu, L.-Q.; Braha, O.; Conlan, S.; Cheley, S.; Bayley, H. Stochastic Sensing of Organic Analytes by a Pore-Forming Protein Containing a Molecular Adapter. *Nature* **1999**, *398*, 686–690.
(77) Howorka, S.; Cheley, S.; Bayley, H. Sequence-Specific Detection of Individual DNA Strands Using Engineered Nanopores. *Nat Biotech* **2001**, *19*, 636–639.
(78) Loman, N. J.; Quick, J.; Simpson, J. T. A Complete Bacterial Genome Assembled de Novo Using Only Nanopore Sequencing Data. *Nat Meth* **2015**, *12*, 733–735.
(79) Pinheiro, A. V; Han, D.; Shih, W. M.; Yan, H. Challenges and Opportunities for Structural DNA Nanotechnology. *Nat Nano* **2011**, *6*, 763–772.
(80) Bell, N. A. W.; Keyser, U. F. Nanopores Formed by DNA Origami: A Review. *FEBS Lett.* **2014**, *588*, 3564–3570.
(81) Wei, R.; Martin, T. G.; Rant, U.; Dietz, H. DNA Origami Gatekeepers for Solid-State Nanopores. *Angew. Chemie Int. Ed.* **2012**, *51*, 4864–4867.
(82) Langecker, M.; Arnaut, V.; Martin, T. G.; List, J.; Renner, S.; Mayer, M.; Dietz, H.; Simmel, F. C. Synthetic Lipid Membrane Channels Formed by Designed DNA Nanostructures. *Sci.* **2012**, *338*, 932–936.
(83) Seifert, A.; Göpfrich, K.; Burns, J. R.; Fertig, N.; Keyser, U. F.; Howorka, S. Bilayer-Spanning DNA Nanopores with Voltage-Switching between Open and Closed State. *ACS Nano* **2015**, *9*, 1117–1126.
(84) Göpfrich, K.; Zettl, T.; Meijering, A. E. C.; Hernández-Ainsa, S.; Kocabey, S.; Liedl, T.; Keyser, U. F. DNA-Tile Structures Induce Ionic Currents through Lipid Membranes. *Nano Lett.* **2015**, *15*, 3134–3138.
(85) Göpfrich, K.; Li, C.-Y.; Mames, I.; Bhamidimarri, S. P.; Ricci, M.; Yoo, J.; Mames, A.; Ohmann, A.; Winterhalter, M.; Stulz, E.; *et al.* Ion Channels Made from a Single Membrane-Spanning DNA Duplex. *Nano Lett.* **2016**, *16*, 4665–4669.
(86) Stoddart, D.; Ayub, M.; Höfler, L.; Raychaudhuri, P.; Klingelhoefer, J. W.; Maglia, G.; Heron, A.; Bayley, H. Functional Truncated Membrane Pores. *Proc. Natl. Acad. Sci.* **2014**, *111*, 2425–2430.
(87) Burns, J. R.; Al-Juffali, N.; Janes, S. M.; Howorka, S. Membrane-Spanning DNA Nanopores with Cytotoxic Effect. *Angew. Chemie Int. Ed.* **2014**, *53*, 12466–12470.
(88) Šulc, P.; Romano, F.; Ouldridge, T. E.; Rovigatti, L.; Doye, J. P. K.; Louis, A. A. Sequence-Dependent Thermodynamics of a Coarse-Grained DNA Model. *J. Chem. Phys. 137*, 135101–135114.

(89) Yoo, J.; Aksimentiev, A. In Situ Structure and Dynamics of DNA Origami Determined through Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci.* **2013**, *9*, 20099–20104.
(90) Joshi, H.; Dwaraknath, A.; Maiti, P. K. Structure, Stability and Elasticity of DNA Nanotubes. *Phys. Chem. Chem. Phys.* **2015**, *17*, 1424–1434.
(91) Maingi, V.; Lelimousin, M.; Howorka, S.; Sansom, M. S. P. Gating-like Motions and Wall Porosity in a DNA Nanopore Scaffold Revealed by Molecular Simulations. *ACS Nano* **2015**, *9*, 11209–11217.
(92) Yoo, J.; Aksimentiev, A. Molecular Dynamics of Membrane-Spanning DNA Channels: Conductance Mechanism, Electro-Osmotic Transport, and Mechanical Gating. *J. Phys. Chem. Lett.* **2015**, *6*, 4680–4687.
(93) Joshi, H.; Kaushik, A.; Seeman, N. C.; Maiti, P. K. Nanoscale Structure and Elasticity of Pillared DNA Nanotubes. *ACS Nano* **2016**, *10*, 7780–7791.
(94) Maffeo, C.; Yoo, J.; Aksimentiev, A. De Novo Reconstruction of DNA Origami Structures through Atomistic Molecular Dynamics Simulation. *Nucleic Acids Res.* **2016**, *44*, 3013–3019.
(95) Bai, X.; Martin, T. G.; Scheres, S. H. W.; Dietz, H. Cryo-EM Structure of a 3D DNA-Origami Object. *Proc. Natl. Acad. Sci.* **2012**, *109*, 20012–20017.
(96) Lee, E. H.; Hsin, J.; Sotomayor, M.; Comellas, G.; Schulten, K. Discovery Through the Computational Microscope. *Structure* **2016**, *17*, 1295–1306.
(97) Frenkel, D.; Berend, S. *Understanding Molecular Simulation: From Algorithms to Applications*.
(98) Leach, A. R. *Molecular Modelling: Principles and Applications*.
(99) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
(100) MacKerell, A. D.; Wiorkiewicz-Kuczera, J.; Karplus, M. An All-Atom Empirical Energy Function for the Simulation of Nucleic Acids. *J. Am. Chem. Soc.* **1995**, *117*, 11946–11975.
(101) Ewald, P. P. Die Berechnung Optischer Und Elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369*, 253–287.
(102) Marrink, S. J.; de Vries, A. H.; Tieleman, D. P. Lipids on the Move: Simulations of Membrane Pores, Domains, Stalks and Curves. *Biochim. Biophys. Acta - Biomembr.* **2009**, *1788*, 149–168.
(103) Dolan, E. A.; Venable, R. M.; Pastor, R. W.; Brooks, B. R. Simulations of Membranes and Other Interfacial Systems Using $P2_1$ and Pc Periodic Boundary Conditions. *Biophys. J.* **2016**, *82*, 2317–2325.
(104) Tieleman, D. P. Methods and Parameters for Membrane Simulations. In *Molecular Simulations and Biomembranes : From Biophysics to Function*; Sansom, M. S. P.; Biggin, P. C., Eds.; 2010; pp. 1–25.
(105) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; *et al.* AMBER 12, 2012.
(106) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(107) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101–014107.
(108) Van Gunsteren, W. F.; Berendsen, H. J. C. A Leap-Frog Algorithm for Stochastic Dynamics. *Mol. Simul.* **1988**, *1*, 173–185.

(109) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
(110) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of A/γ Conformers. *Biophys. J.* **2007**, *92*, 3817–3829.
(111) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *J. Chem. Theory Comput.* **2011**, *8*, 348–362.
(112) Galindo-Murillo, R.; Robertson, J. C.; Zgarbová, M.; Šponer, J.; Otyepka, M.; Jurečka, P.; Cheatham, T. E. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, *12*, 4114–4127.
(113) Ivani, I.; Dans, P. D.; Noy, A.; Perez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goni, R.; Balaceanu, A.; *et al.* Parmbsc1: A Refined Force Field for DNA Simulations. *Nat Meth* **2016**, *13*, 55–58.
(114) Uusitalo, J. J.; Ingólfsson, H. I.; Akhshi, P.; Tieleman, D. P.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory Comput.* **2015**, *11*, 3932–3945.
(115) Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational Lipidomics with Insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144–2155.
(116) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
(117) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **2014**, *10*, 676–690.
(118) Brocos, P.; Mendoza-Espinosa, P.; Castillo, R.; Mas-Oliva, J.; Pineiro, A. Multiscale Molecular Dynamics Simulations of Micelles: Coarse-Grain for Self-Assembly and Atomic Resolution for Finer Details. *Soft Matter* **2012**, *8*, 9005–9014.
(119) Kearney, C. J.; Lucas, C. R.; O'Brien, F. J.; Castro, C. E. DNA Origami: Folded DNA-Nanodevices That Can Direct and Interpret Cell Behavior. *Adv. Mater.* **2016**, *28*, 5509–5524.
(120) Hohng, S.; Zhou, R.; Nahas, M. K.; Yu, J.; Schulten, K.; Lilley, D. M. J.; Ha, T. Fluorescence-Force Spectroscopy Maps Two-Dimensional Reaction Landscape of the Holliday Junction. *Science (80-.).* **2007**, *318*, 279 LP-283.
(121) McKinney, S. A.; Declais, A.-C.; Lilley, D. M. J.; Ha, T. Structural Dynamics of Individual Holliday Junctions. *Nat Struct Mol Biol* **2003**, *10*, 93–97.
(122) Watson, J.; Hays, F. A.; Ho, P. S. Definitions and Analysis of DNA Holliday Junction Geometry. *Nucleic Acids Res.* **2004**, *32*, 3017–3027.
(123) Macke, T. J.; Case, D. A. Modeling Unusual Nucleic Acid Structures. In *Molecular Modeling of Nucleic Acids*; ACS Symposium Series; American Chemical Society, 1997; Vol. 682, pp. 24–379.

(124) Ussery, D. W. DNA Structure: A-, B- and Z-DNA Helix Families. In *eLS*; John Wiley & Sons, Ltd, 2001.
(125) Banavali, N. K.; Roux, B. Free Energy Landscape of A-DNA to B-DNA Conversion in Aqueous Solution. *J. Am. Chem. Soc.* **2005**, *127*, 6866–6876.
(126) Noy, A.; Pérez, A.; Laughton, C. A.; Orozco, M. Theoretical Study of Large Conformational Transitions in DNA: The B↔A Conformational Change in Water and Ethanol/water. *Nucleic Acids Res.* **2007**, *35*, 3330–3338.
(127) Waters, J. T.; Lu, X.-J.; Galindo-Murillo, R.; Gumbart, J. C.; Kim, H. D.; Cheatham, T. E.; Harvey, S. C. Transitions of Double-Stranded DNA Between the A- and B-Forms. *J. Phys. Chem. B* **2016**, *120*, 8449–8456.
(128) Mitchell, J. C.; Harris, J. R.; Malo, J.; Bath, J.; Turberfield, A. J. Self-Assembly of Chiral DNA Nanotubes. *J. Am. Chem. Soc.* **2004**, *126*, 16342–16343.
(129) Hariri, A. A.; Hamblin, G. D.; Gidi, Y.; Sleiman, H. F.; Cosa, G. Stepwise Growth of Surface-Grafted DNA Nanotubes Visualized at the Single-Molecule Level. *Nat Chem* **2015**, *7*, 295–300.
(130) Li, C.-Y.; Hemmig, E. A.; Kong, J.; Yoo, J.; Hernández-Ainsa, S.; Keyser, U. F.; Aksimentiev, A. Ionic Conductivity, Structural Deformation, and Programmable Anisotropy of DNA Origami in Electric Field. *ACS Nano* **2015**, *9*, 1420–1433.
(131) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41.
(132) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888.
(133) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*.
(134) Joung, I. S.; Cheatham Thomas E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.
(135) Wu, X.; Subramaniam, S.; Case, D. A.; Wu, K. W.; Brooks, B. R. Targeted Conformational Search with Map-Restrained Self-Guided Langevin Dynamics: Application to Flexible Fitting into Electron Microscopic Density Maps. *J. Struct. Biol.* **2013**, *183*, 429–440.
(136) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*.
(137) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X.-J.; Neidle, S.; Shakked, Z.; *et al.* A Standard Reference Frame for the Description of Nucleic Acid Base-Pair Geometry. *J. Mol. Biol.* **2001**, *313*, 229–237.
(138) Smart, O. S.; Neduvelil, J. G.; Wang, X.; Wallace, B. A.; Sansom, M. S. P. HOLE: A Program for the Analysis of the Pore Dimensions of Ion Channel Structural Models. *J. Mol. Graph.* **1996**, *14*, 354–360.

(139) Laurent, B.; Chavent, M.; Cragnolini, T.; Dahl, A. C. E.; Pasquali, S.; Derreumaux, P.; Sansom, M. S. P.; Baaden, M. Epock: Rapid Analysis of Protein Pocket Dynamics. *Bioinforma.* **2015**, *31*, 1478–1480.
(140) Yamamoto, E.; Kalli, A. C.; Akimoto, T.; Yasuoka, K.; Sansom, M. S. P. Anomalous Dynamics of a Lipid Recognition Protein on a Membrane Surface. *Sci. Rep.* **2015**, *5*, 18245.
(141) Kutzner, C.; Grubmüller, H.; de Groot, B. L.; Zachariae, U. Computational Electrophysiology: The Molecular Dynamics of Ion Channel Permeation and Selectivity in Atomistic Detail. *Biophys. J.* **2011**, *101*, 809–817.
(142) García-Fandiño, R.; Sansom, M. S. P. Designing Biomimetic Pores Based on Carbon Nanotubes. *Proc. Natl. Acad. Sci.* **2012**, *109*, 6939–6944.
(143) Guéroult, M.; Boittin, O.; Mauffret, O.; Etchebest, C.; Hartmann, B. Mg^{2+} in the Major Groove Modulates B-DNA Structure and Dynamics. *PLoS One* **2012**, *7*, e41704.
(144) Egli, M. DNA-Cation Interactions: Quo Vadis? *Chem. Biol.* **2002**, *9*, 277–286.
(145) Várnai, P.; Zakrzewska, K. DNA and Its Counterions: A Molecular Dynamics Study. *Nucleic Acids Res.* **2004**, *32*, 4269–4280.
(146) Chavent, M.; Reddy, T.; Goose, J.; Dahl, A. C. E.; Stone, J. E.; Jobard, B.; Sansom, M. S. P. Methodologies for the Analysis of Instantaneous Lipid Diffusion in Md Simulations of Large Membrane Systems. *Faraday Discuss.* **2014**, *169*, 455–475.
(147) Langecker, M.; Arnaut, V.; List, J.; Simmel, F. C. DNA Nanostructures Interacting with Lipid Bilayer Membranes. *Acc. Chem. Res.* **2014**, *47*, 1807–1815.
(148) Stansfeld, P. J.; Goose, J. E.; Caffrey, M.; Carpenter, E. P.; Parker, J. L.; Newstead, S.; Sansom, M. S. P. MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure* **2016**, *23*, 1350–1361.
(149) Gumbart, J.; Wang, Y.; Aksimentiev, A.; Tajkhorshid, E.; Schulten, K. Molecular Dynamics Simulations of Proteins in Lipid Bilayers. *Curr. Opin. Struct. Biol.* **2005**, *15*, 423–431.
(150) Tang, D.; Yoo, Y.-E.; Kim, D. Molecular Dynamics Simulations on Water Permeation through Hourglass-Shaped Nanopores with Varying Pore Geometry. *Chem. Phys.* **2015**, *453–454*, 13–19.
(151) Beckstein, O.; Sansom, M. S. P. Liquid–vapor Oscillations of Water in Hydrophobic Nanopores. *Proc. Natl. Acad. Sci.* **2003**, *100*, 7063–7068.
(152) Aksimentiev, A.; Schulten, K. Imaging α-Hemolysin with Molecular Dynamics: Ionic Conductance, Osmotic Permeability, and the Electrostatic Potential Map. *Biophys. J.* **2005**, *88*, 3745–3761.
(153) Paulechka, E.; Wassenaar, T. A.; Kroenlein, K.; Kazakov, A.; Smolyanitsky, A. Nucleobase-Functionalized Graphene Nanoribbons for Accurate High-Speed DNA Sequencing. *Nanoscale* **2016**, *8*, 1861–1867.
(154) Antipina, A. Y.; Gurtovenko, A. A. Molecular-Level Insight into the Interactions of DNA with Phospholipid Bilayers: Barriers and Triggers. *RSC Adv.* **2016**, *6*, 36425–36432.
(155) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The Power of Coarse Graining in Biomolecular Simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 225–248.
(156) Hedger, G.; Sansom, M. S. P. Lipid Interaction Sites on Channels,

Transporters and Receptors: Recent Insights from Molecular Dynamics Simulations. *Biochim. Biophys. Acta - Biomembr.* **2016**, *1858*, 2390–2400.
(157) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chemie Int. Ed.* **1999**, *38*, 236–240.
(158) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini Model. *Chem. Soc. Rev.* **2013**, *42*, 6801–6822.
(159) Lopez, C. F.; Nielsen, S. O.; Moore, P. B.; Klein, M. L. Understanding Nature's Design for a Nanosyringe. *Proc. Natl. Acad. Sci. United States Am.* **2004**, *101*, 4431–4434.
(160) García, A. E.; Stiller, L. Computation of the Mean Residence Time of Water in the Hydration Shells of Biomolecules. *J. Comput. Chem.* **1993**, *14*, 1396–1406.
(161) Rassam, P.; Copeland, N. A.; Birkholz, O.; Toth, C.; Chavent, M.; Duncan, A. L.; Cross, S. J.; Housden, N. G.; Kaminska, R.; Seger, U.; *et al.* Supramolecular Assemblies Underpin Turnover of Outer Membrane Proteins in Bacteria. *Nature* **2015**, *523*, 333–336.
(162) Isralewitz, B.; Gao, M.; Schulten, K. Steered Molecular Dynamics and Mechanical Functions of Proteins. *Curr. Opin. Struct. Biol.* **2001**, *11*, 224–230.
(163) Wallace, E. J.; Sansom, M. S. P. Blocking of Carbon Nanotube Based Nanoinjectors by Lipids: A Simulation Study. *Nano Lett.* **2008**, *8*, 2751–2756.
(164) Petrosyan, R.; Bippes, C. A.; Walheim, S.; Harder, D.; Fotiadis, D.; Schimmel, T.; Alsteens, D.; Müller, D. J. Single-Molecule Force Spectroscopy of Membrane Proteins from Membranes Freely Spanning Across Nanoscopic Pores. *Nano Lett.* **2015**, *15*, 3624–3633.
(165) Oesterhelt, F.; Oesterhelt, D.; Pfeiffer, M.; Engel, A.; Gaub, H. E.; Müller, D. J. Unfolding Pathways of Individual Bacteriorhodopsins. *Science (80-.).* **2000**, *288*, 143–146.
(166) Melo, M. N.; Ingólfsson, H. I.; Marrink, S. J. Parameters for Martini Sterols and Hopanoids Based on a Virtual-Site Description. *J. Chem. Phys.* **2015**, *143*, 243152–243159.
(167) Lee, H.; de Vries, A. H.; Marrink, S.-J.; Pastor, R. W. A Coarse-Grained Model for Polyethylene Oxide and Polyethylene Glycol: Conformation and Hydrodynamics. *J. Phys. Chem. B* **2009**, *113*, 13186–13194.
(168) de Jong, D. H.; Liguori, N.; van den Berg, T.; Arnarez, C.; Periole, X.; Marrink, S. J. Atomistic and Coarse Grain Topologies for the Cofactors Associated with the Photosystem II Core Complex. *J. Phys. Chem. B* **2015**, *119*, 7791–7803.
(169) Long, H.; Kudlay, A.; Schatz, G. C. Molecular Dynamics Studies of Ion Distributions for DNA Duplexes and DNA Clusters: Salt Effects and Connection to DNA Melting. *J. Phys. Chem. B* **2006**, *110*, 2918–2926.
(170) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; *et al.* CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671–690.
(171) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput.*

Phys. Commun. **1995**, *91*, 43–56.
(172) Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116–122.
(173) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
(174) Lukat, G.; Krüger, J.; Sommer, B. APL@Voro: A Voronoi-Based Membrane Analysis Tool for GROMACS Trajectories. *J. Chem. Inf. Model.* **2013**, *53*, 2908–2925.
(175) Jiang, Q.; Song, C.; Nangreave, J.; Liu, X.; Lin, L.; Qiu, D.; Wang, Z.-G.; Zou, G.; Liang, X.; Yan, H.; *et al.* DNA Origami as a Carrier for Circumvention of Drug Resistance. *J. Am. Chem. Soc.* **2012**, *134*, 13396–13403.
(176) Wu, X.; Chen, J.; Wu, M.; Zhao, J. X. Aptamers: Active Targeting Ligands for Cancer Diagnosis and Therapy. *Theranostics* **2015**, *5*, 322–344.
(177) Zwicke, G. L.; Mansoori, G. A.; Jeffery, C. J. Utilizing the Folate Receptor for Active Targeting of Cancer Nanotherapeutics. *Nano Rev.* **2012**, *3*, 10.3402/nano.v3i0.18496.
(178) Yoo, J.; Aksimentiev, A. Improved Parameterization of Amine–Carboxylate and Amine–Phosphate Interactions for Molecular Dynamics Simulations Using the CHARMM and AMBER Force Fields. *J. Chem. Theory Comput.* **2016**, *12*, 430–443.
(179) Zgarbová, M.; Šponer, J.; Otyepka, M.; Cheatham, T. E.; Galindo-Murillo, R.; Jurečka, P. Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *J. Chem. Theory Comput.* **2015**, *11*, 5723–5736.
(180) Snodin, B. E. K.; Romano, F.; Rovigatti, L.; Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K. Direct Simulation of the Self-Assembly of a Small DNA Origami. *ACS Nano* **2016**, *10*, 1724–1737.
(181) Graham, J. PyCGTool https://github.com/jag1g13/pycgtool.
(182) Zhang, Q.; Jiang, Q.; Li, N.; Dai, L.; Liu, Q.; Song, L.; Wang, J.; Li, Y.; Tian, J.; Ding, B.; *et al.* DNA Origami as an In Vivo Drug Delivery Vehicle for Cancer Therapy. *ACS Nano* **2014**, *8*, 6633–6643.