



Publicly Accessible Penn Dissertations

Summer 7-2-2009

Probing the Structure and Function of Biopolymer-Carbon Nanotube Hybrids with Molecular Dynamics

Robert R. Johnson

University of Pennsylvania, robertjo@physics.upenn.edu

Follow this and additional works at: <http://repository.upenn.edu/edissertations>

 Part of the [Biological and Chemical Physics Commons](#), and the [Physical Chemistry Commons](#)

Recommended Citation

Johnson, Robert R., "Probing the Structure and Function of Biopolymer-Carbon Nanotube Hybrids with Molecular Dynamics" (2009). *Publicly Accessible Penn Dissertations*. 5.
<http://repository.upenn.edu/edissertations/5>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/edissertations/5>
For more information, please contact libraryrepository@pobox.upenn.edu.

Probing the Structure and Function of Biopolymer-Carbon Nanotube Hybrids with Molecular Dynamics

Abstract

Nanoscience deals with the characterization and manipulation of matter on the atomic/molecular size scale in order to deepen our understanding of condensed matter and develop revolutionary technology. Meeting the demands of the rapidly advancing nanotechnological frontier requires novel, multifunctional nanoscale materials. Among the most promising nanomaterials to fulfill this need are biopolymer-carbon nanotube hybrids (Bio-CNT). Bio-CNT consists of a single-walled carbon nanotube (CNT) coated with a self-assembled layer of biopolymers such as DNA or protein. Experiments have demonstrated that these nanomaterials possess a wide range of technologically useful properties with applications in nanoelectronics, medicine, homeland security, environmental safety and microbiology. However, a fundamental understanding of the self-assembly mechanics, structure and energetics of Bio-CNT is lacking. The objective of this thesis is to address this deficiency through molecular dynamics (MD) simulation, which provides an atomic-scale window into the behavior of this unique nanomaterial. MD shows that Bio-CNT composed of single-stranded DNA (ssDNA) self-assembles via the formation of high affinity contacts between DNA bases and the CNT sidewall. Calculation of the base-CNT binding free energy by thermodynamic integration reveals that these contacts result from the attractive pi-pi stacking interaction. Binding affinities follow the trend $G > A > T > C$. MD reveals that long ssDNA sequences are driven into a helical wrapping about CNT with a sub-10 nm pitch by electrostatic and torsional interactions in the backbone. A large-scale replica exchange molecular dynamics simulation reveals that ssDNA-CNT hybrids are disordered. At room temperature, ssDNA can reside in several low-energy conformations that contain a sequence-specific arrangement of bases detached from CNT surface. MD demonstrates that protein-CNT hybrids composed of the Coxsackie-adenovirus receptor are biologically active and function as a nanobiosensor with specific recognition of Knob proteins from the adenovirus capsid. Simulation also shows that the rigid CNT damps structural fluctuations in bound proteins, which may have important ramifications for biosensing devices composed of protein-CNT hybrids. These results expand current knowledge of Bio-CNT and demonstrate the effectiveness of MD for investigations of nano-biomolecular systems.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Physics & Astronomy

First Advisor

A.T. Charlie Johnson

Second Advisor

Michael L. Klein

Keywords

carbon nanotube, dna, protein, nanoscience, molecular dynamics, nano bio

Subject Categories

Biological and Chemical Physics | Physical Chemistry

PROBING THE STRUCTURE AND FUNCTION OF
BIOPOLYMER-CARBON NANOTUBE HYBRIDS
WITH MOLECULAR DYNAMICS

Robert R. Johnson

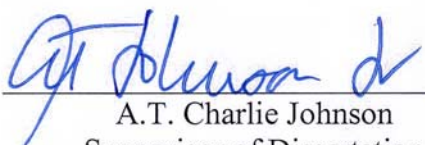
A DISSERTATION

in


PHYSICS AND ASTRONOMY

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

2009



A.T. Charlie Johnson
Supervisor of Dissertation



Ravi K. Sheth
Graduate Group Chairperson

Dissertation committee

Andrea Liu
Philip Nelson
Eugene W. Beier

Probing the Structure and Function of Biopolymer-Carbon Nanotube Hybrids with
Molecular Dynamics

© Copyright 2009 by Robert R. Johnson.

All rights reserved.

For my parents

Acknowledgements

I'd like to thank Dr. Charlie Johnson for serving as my advisor and providing me with five years of interesting research projects, scientific knowledge and useful advice. Thanks also to Dr. Michael Klein for providing helpful scientific insights and discussion in addition to a great workspace at the Center for Molecular Modeling (CMM) with access to world class computational resources. I'd like to thank the entire CMM for their openness to offer suggestions and answer questions over the years. I benefited immensely from my collective interactions with group members. I'm very thankful for Dr. Katrin Spiegel for many useful discussions about DNA, Dr. Matteo Dal Peraro for his mentorship, help setting up simulations and generating ideas for research projects, Dr. Jérôme Hénin and Dr. David Mobley for their valuable advice regarding free energy calculations and Dr. Marco De Vivo for his thoughtful commentary on my first few publications. Very special thanks to Dr. Axel Kohlmeyer who patiently and enthusiastically helped with the daily technical problems of installing and compiling programs, getting jobs running on supercomputers and writing data analysis scripts. He was never too busy for discussion and I'm extremely grateful for his help and guidance. I'd also like to thank the entire Johnson group and especially Dr. Sam Khamis for many useful interactions with the experimentalists who had daily hands-on access to the real nanoscale materials that I have modeled. Thanks to all of the great undergraduate students that I mentored on summer research projects over the years. I had a great time working with all of you. Thanks to my parents who have provided me with a lifetime of love and support. It is because of you that I have been able to complete this thesis. Finally, I give great thanks to God who makes all things possible in my life.

ABSTRACT

PROBING THE STRUCTURE AND FUNCTION OF BIOPOLYMER-CARBON NANOTUBE HYBRIDS WITH MOLECULAR DYNAMICS

Robert R. Johnson

A.T. Charlie Johnson and Michael L. Klein

Nanoscience deals with the characterization and manipulation of matter on the atomic/molecular size scale in order to deepen our understanding of condensed matter and develop revolutionary technology. Meeting the demands of the rapidly advancing nanotechnological frontier requires novel, multifunctional nanoscale materials. Among the most promising nanomaterials to fulfill this need are biopolymer-carbon nanotube hybrids (Bio-CNT). Bio-CNT consists of a single-walled carbon nanotube (CNT) coated with a self-assembled layer of biopolymers such as DNA or protein. Experiments have demonstrated that these nanomaterials possess a wide range of technologically useful properties with applications in nanoelectronics, medicine, homeland security, environmental safety and microbiology. However, a fundamental understanding of the self-assembly mechanics, structure and energetics of Bio-CNT is lacking. The objective of this thesis is to address this deficiency through molecular dynamics (MD) simulation, which provides an atomic-scale window into the behavior of this unique nanomaterial. MD shows that Bio-CNT composed of single-stranded DNA (ssDNA) self-assembles via the formation of high affinity contacts between DNA bases and the CNT sidewall. Calculation of the base-CNT binding free energy by thermodynamic integration reveals that these contacts result from the attractive π - π stacking interaction. Binding affinities follow the trend $G > A > T > C$. MD reveals that long ssDNA sequences are driven into a

helical wrapping about CNT with a sub-10 nm pitch by electrostatic and torsional interactions in the backbone. A large-scale replica exchange molecular dynamics simulation reveals that ssDNA-CNT hybrids are disordered. At room temperature, ssDNA can reside in several low-energy conformations that contain a sequence-specific arrangement of bases detached from CNT surface. MD demonstrates that protein-CNT hybrids composed of the Coxsackie-adenovirus receptor are biologically active and function as a nanobiosensor with specific recognition of Knob proteins from the adenovirus capsid. Simulation also shows that the rigid CNT damps structural fluctuations in bound proteins, which may have important ramifications for biosensing devices composed of protein-CNT hybrids. These results expand current knowledge of Bio-CNT and demonstrate the effectiveness of MD for investigations of nano-biomolecular systems.

Contents

Overview	1
1.1 Nanoscience	1
1.2 This Thesis	7
Biopolymer-Carbon Nanotube Hybrids	9
2.1 Carbon Nanotubes	9
2.1.1 Structure.....	12
2.1.2 Electronic Properties	14
2.1.3 Challenges in Carbon Nanotube Synthesis and Purification.....	16
2.2 Biopolymers	17
2.2.1 DNA.....	18
2.2.2 Proteins	20
2.3 Biopolymer-Carbon Nanotube Hybrids	24
2.3.1 DNA-Carbon Nanotube Hybrids.....	24
2.3.2 Protein-Carbon Nanotube Hybrids.....	28
2.4 Computer Simulations in Nanoscience	30
Molecular Dynamics	34
3.1 Methodology	35
3.1.1 Initialization of the System.....	36
3.1.2 Force Calculation	36
3.1.3 Temperature and Pressure Regulation.....	39

3.2 Atomistic Force Field	40
3.2.1 Electrostatic Interaction	40
3.2.2 van der Waals Interaction	41
3.2.3 Bond Stretching	42
3.2.4 Angle Bending	43
3.2.5 Bond Torsion	43
3.2.6 Parameterization	44
3.3 Thermodynamic Integration	45
3.4 Replica Exchange Molecular Dynamics	50
3.5 Simulation Protocol	53
Structure of DNA-Carbon Nanotube Hybrids	55
4.1 DNA-Carbon Nanotube Hybrid Self-Assembly	56
4.1.1 Simulation Details	57
4.1.2 Results and Discussion	57
4.2 The Nature of DNA-Carbon Nanotube Interactions	61
4.2.1 Simulation Details	62
4.2.2 Results and Discussion	66
4.3 Stability of Proposed Structures	73
4.3.1 Stability of Poly GT Dimers Adsorbed to Carbon Nanotubes	74
4.3.2 Stability of Helically Wrapped DNA About Carbon Nanotubes	77
4.3.3 Mechanism for Helical Wrapping of DNA about Carbon Nanotubes	79
4.4 Free Energy Landscape of a DNA-Carbon Nanotube Hybrid	85

4.4.1 Simulation Details	86
4.4.2 Results and Discussion.....	88
4.5 Summary	97
Structure and Function of a Nanobiosensor:	
A Carbon Nanotube Functionalized with the Coxsackie-	
Adenovirus Receptor	100
5.1 Simulation Details	102
5.2 Protein Simulations Under Native Conditions.....	104
5.3 Covalent Attachment of the Coxsackie-Adenovirus Receptor ...	106
5.4 Structure and Function of a Nanobiosensor.....	108
5.5 Discussion.....	112
5.6 Summary	113
Summary and Conclusion	114
6.1 Future Research	117
Appendix A: Nanotube Builder 1.0 TCL Script	121

List of Tables

Table 3.1: MD algorithm.	35
Table 4.1: Base-CNT binding free energy	67
Table 5.1: Summary of CAR-CNT simulations performed.	103

List of Figures

Figure 1.1: Two allotropes of carbon: graphite and diamond.....	6
Figure 2.1: A single-walled carbon nanotube.....	10
Figure 2.2: A multi-walled carbon nanotube.....	11
Figure 2.3: Rolling a graphene sheet into CNT.....	12
Figure 2.4: Various atomic structures of CNT.....	13
Figure 2.5: Band structure for graphene π electrons.....	15
Figure 2.6: Chemical structure of a DNA strand and A-T and G-C base pairs.....	18
Figure 2.7: Forms of dsDNA (A, B, Z) and ssDNA (hairpin, random coil).....	20
Figure 2.8: Chemical structure of an amino acid and a protein.....	21
Figure 2.9: Structure of an α -helix.....	22
Figure 2.10: Backbone and hydrogen bond pattern a β -sheet.....	22
Figure 2.11: Protein visualization styles.....	23
Figure 2.12: DNA-assisted dispersion and separation of CNTs in aqueous solution.....	25
Figure 2.13: AFM image of CNT wrapped with (GT) ₃₀	26
Figure 2.14: CNT-FET functionalized with ssDNA for chemical sensing.....	27
Figure 2.15: CNT-FET functionalized with receptor proteins for biological sensing.....	29
Figure 2.16: Diimide-activated amidation.....	30
Figure 2.17: Time and length scales accessible to different simulation techniques.....	31
Figure 3.1: Schematic of periodic boundary conditions.....	37
Figure 3.2: The van der Waals interaction modeled by the Lennard-Jones potential.....	42
Figure 3.3: Stretching of a chemical bond.....	42
Figure 3.4: Bending of an angle in a molecule.....	43

Figure 3.5: Torsion angle.....	44
Figure 3.6: Thermodynamic cycle to compute protein-ligand binding free energy	47
Figure 3.7: Soft-core potential for van der Waals interactions for various λ values.....	48
Figure 3.8: Instantaneous temperature of one replica during an REMD simulation	51
Figure 3.9: Energy distribution of each replica in an REMD simulation	52
Figure 4.1: Self-assembly of DNA-CNT in aqueous solution.....	58
Figure 4.2: Number of adsorbed bases N_{adsorb} during DNA-CNT self-assembly	59
Figure 4.3: Metastable kinked ssDNA structure.....	60
Figure 4.4: Thermodynamic cycle to calculate the base-CNT binding free energy	62
Figure 4.5: $\langle \partial H / \partial \lambda \rangle$ curves for annihilating adenine	67
Figure 4.6: Change in various interaction energies during base-CNT binding	68
Figure 4.7: Water density around adenine and CNT in bound and unbound states	69
Figure 4.8: Water-adenine hydrogen bonds.....	70
Figure 4.9: Proposed dimeric structure of poly GT around CNT.....	74
Figure 4.10: Energy minimization of a (GT) ₂ dimer on CNT	75
Figure 4.11: G-T Wobble base pair	76
Figure 4.12: (GT) ₃₀ on a (11,0) CNT with regular 18 nm pitch	78
Figure 4.13: Different sugar-phosphate backbone orientations with respect to CNT	78
Figure 4.14: Simulation of (GT) ₂₀ about CNT	80
Figure 4.15: Structural and energetic changes for (GT) ₂₀ about CNT.....	82
Figure 4.16: Desorbed bases within adsorbed ssDNA	84
Figure 4.17: Free energy landscape of (GT) ₇ -CNT hybrid at room temperature	89
Figure 4.18: Comparison of free energy landscapes from MD and REMD	90

Figure 4.19: Probability distribution of number of adsorbed bases in (GT) ₇	91
Figure 4.20: Free energy difference for adsorbed and desorbed states for (GT) ₇ bases..	93
Figure 4.21: Probability density of nucleotide orientation on CNT surface.....	94
Figure 4.22: Intrastrand hydrogen bonds for adsorbed ssDNA.....	95
Figure 4.23: (GT) ₇ free energy landscape at various temperatures	96
Figure 4.24: Loop conformation at 5' end of helical conformations in 40-60mers.....	97
Figure 5.1: Adenovirus binding to the coxsackie-adenovirus receptor D1 domain	101
Figure 5.2: Visualization of CAR showing its secondary structural elements	104
Figure 5.3: Main contacts in CAR-Knob binding site.....	106
Figure 5.4: Potential CAR-CNT binding sties.....	107
Figure 5.5: Structural changes in CAR that occur after CNT binding	110
Figure 5.6: Average heights of bound CAR-Knob from CNT	111

Chapter 1

Overview

One of the most important scientific discoveries of all time is that matter is not smooth and continuous, but is granular and composed of atoms. As it turns out, atoms have a substructure of their own and consist of a cloud of electrons surrounding a nucleus. Furthermore, the nucleus is subdivided into protons and neutrons and these objects can be further broken down into quarks. Atoms and their constituent particles have been heavily studied for a century. While there remain many unanswered questions about the fundamental nature of these individual objects, they tend to behave in a comparatively simple way. The richness of matter begins to become evident when looking at collections of atoms. Multiple atoms interact with one another and can spontaneously self-arrange into complex geometries by forming chemical bonds. These geometries can be periodic and extended, such as in a crystal, or finite and anisotropic, such as in a molecule. The behavior of these atomic assemblies, which varies according to the assembly's elemental composition and structure, is ultimately responsible for all material properties. Thus, much can be gained by studying how atoms and molecules interact, assemble and behave.

1.1 Nanoscience

Nanoscience deals with the manipulation and characterization of matter on the atomic and molecular size scale. The physical dimensions of the materials encountered at this scale typically range from one to several hundred nanometers. The prefix “nano” stands

for 10^{-9} , or one billionth and thus, a nanometer (nm) is one billionth of a meter. To grasp a sense of this size, it would require approximately 100,000 nanometer-sized objects placed side by side to equal the width of a single piece of paper. At smaller and smaller length scales, new forces and effects become increasingly important and cause nanoscale objects to behave differently than human (macro) scale objects. The unique characteristics of the nanoscale pave the way for revolutionary technological applications not possible with conventional macroscale materials. The behavior of atoms, molecules and other nanoscale objects differs from macroscale objects for several reasons:

1. Nanoscale objects are subject to different laws of physics

The classical laws of physics (Newton's Laws) break down at the nanoscale and are replaced by quantum mechanics which provide the correct description of atomic behavior. Particles subject to quantum mechanics behave in a fundamentally different way than classical particles.

2. Nanoscale objects are subject to different forces of nature

The strength of the various forces of nature depends on scale. On astronomical scales, gravity is the most important force. At the nanoscale, however, gravity is weak and the electromagnetic force dominates. This force can be either attractive or repulsive and lies at the root of all nanoscale phenomena including chemical bonding, electronic transport and molecular self-assembly.

3. Nanoscale objects are subject to thermal motion

At finite temperature, atoms are in constant motion due to random collisions with other atoms. These thermal fluctuations result in the passive transport of atoms

and molecules (diffusion) throughout a fluid and induce spontaneous structural changes in nanoscale objects.

4. Nanoscale objects have a high surface area to volume ratio

Most atoms in a macroscale object are located within the interior. However, because of their small size, nanoscale objects contain a significant fraction of their atoms at the surface. This enables nanoscale objects to interact differently with the environment and have a greater chemical reactivity.

These factors all contribute to give rise to properties at the nanoscale that differ from macroscale objects. For example, as one goes to the nanoscale copper goes from reddish-brown in color to transparent; gold goes from a solid at room temperature to a liquid; silicon goes from an insulator to a conductor.

Because all materials are composed of atoms, nanoscience is highly interdisciplinary and draws from the traditional natural sciences of physics, chemistry, biology and materials science. Additionally, this field not only seeks to understand the behavior of the nanoscale, but to learn how to manipulate molecular sized objects for the development of new (nano)technology. Thus, nanoscience contains a strong engineering presence as well. Nanoscience, as it aspires to understand and control the elementary building blocks of matter, holds the potential to revolutionize the way machines are built. Many traditional fabrication techniques rely on top-down processes where external tools are used to fashion materials into machine parts of the desired shape and size. These parts are then manually combined into a useful device. Much of nanotechnology, on the other hand, seeks bottom-up processes that employ the unique ability of specially designed molecular substances to self-assemble and self-organize into useful devices and/or structures with

the desired properties and function. The latter method has the potential to produce machinery in parallel and therefore be much cheaper and scalable than top-down processes.

Another central goal of the field is to take advantage of the unique properties of a host of different nanoscale materials in order to develop revolutionary nanotechnologies. Among the most promising materials being employed to fulfill this objective are carbon nanotubes (CNT), due to their unprecedented mechanical, electrical and optical properties, and biopolymers such as DNA and proteins, due to their biological significance and molecular recognition capabilities. CNTs are cylindrical sheets of carbon atoms with diameters of ~ 1 nm and have many potential applications as miniaturized electronics. They have already been fashioned into nanoscale transistors,¹ sensors² and memory devices.³ DNA is complex biological heteropolymer that displays self-recognition in the hybridization of the DNA double helix and can be engineered through “directed evolution” for recognition of other molecular species.⁴ These properties can be utilized to design complex three-dimensional (3D) molecular structures that self-assemble⁵ and direct the construction of nanodevices.⁶⁻⁸ Proteins are ubiquitous biological macromolecules that are involved in virtually every cellular process. One of the holy grails of protein nanoscience is the synthesis of engineered proteins that carry out customized functions valuable to technology, medicine and research.⁹⁻¹¹ While this is a challenging goal, there has been limited success in applying protein design algorithms to stabilize proteins,¹² develop new enzymes¹³ and structures¹⁴ and solubilize membrane proteins.¹⁵

Amazingly, even though CNTs and biopolymers have never encountered one another in nature, they are chemically compatible and are easily combined to form hybrid inorganic/organic nanomaterials with a set of unique properties that derive from the two components. These biopolymer-carbon nanotube hybrids (Bio-CNT) are multifunctional materials that are truly greater than the sum of their parts and hold promise to drive advances in nanoelectronics, molecular biology, environmental safety, medicine and homeland security. Experiments have already demonstrated the value of Bio-CNTs in biomedical devices,¹⁶ CNT solubilization,¹⁷⁻¹⁹ cellular delivery of biomolecules²⁰ and label-free chemical²¹ and biological²²⁻²⁴ sensing devices. Despite the importance of Bio-CNT, a fundamental understanding of its self-assembly mechanisms, physical properties and molecular interactions is lacking.

This information can be gained by resolving the 3D atomic structure of Bio-CNT, as all properties (chemical and physical) of a nanoscale object are determined by its structure. In other words, “form determines function”. This adage is best exemplified by the two allotropes of carbon: graphite and diamond. Each of these materials is composed entirely of carbon atoms. The only difference between the two is the 3D arrangement of the atoms (Figure 1.1). Graphite consists of a series of two-dimensional hexagonal sheets stacked on top of one another. The interlayer spacing is 0.34 nm and the carbon-carbon bond length is 0.142 nm. Carbon atoms in diamond reside in a face centered cubic lattice where the carbon-carbon spacing is 0.154 nm. Owing to these structural differences, the properties of graphite and diamond are polar opposites. Graphite is soft and brittle which makes it a good lubricant, whereas diamond is one of the hardest known materials and is

used in cutting tools and abrasives. Graphite is opaque while diamond is transparent. And finally, graphite is a good conductor of electricity while diamond is an insulator.

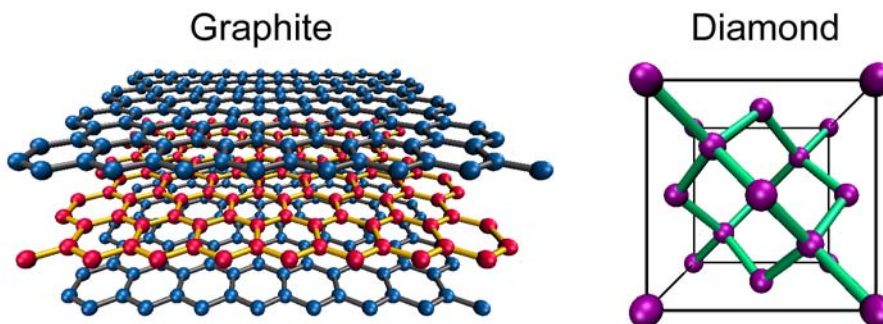


Figure 1.1: Two allotropes of carbon: graphite and diamond. Their differing atomic structures result in their drastically different properties.

The intimate relation of structure and function is rampant in biology as well. The particular tasks carried out by biological molecules such as proteins, RNA and DNA are determined by their 3D shapes. There exists an entire branch of science known as structural biology that is dedicated to archiving, classifying, understanding and predicting the structures of biomolecules.

Among the experimental tools available to study the structure of nanoscale objects are atomic force microscopy (AFM), scanning electron microscopy (SEM), transmission electron microscopy (TEM), scanning tunneling microscopy (STM), nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography. Under the right conditions, atomic-scale resolutions can be achieved. However, due to Bio-CNT's heterogeneity and complexity, obtaining resolutions less than a few nanometers with these techniques has been difficult. Computer simulation is another valuable tool that enables study of the structure and function of nanoscale objects. Unlike experimental methods, computation can probe arbitrarily small length scales, albeit in a virtual world. Molecular dynamics

(MD) simulation is a computational technique that is especially useful for studying biophysical systems such as Bio-CNT. MD computes the trajectories of complex systems of interacting particles and therefore, provides a detailed microscopic view into the dynamical behavior of nanoscale systems. MD is advantageous because it can simulate real experimental conditions such as temperature, pressure and the aqueous environment.

1.2 This Thesis

A deeper understanding of Bio-CNT lies at a powerful advancing frontier of fundamental research in nanoscience and will enable numerous proposed applications of this hybrid organic/inorganic nanomaterial. In order to meet this need, this thesis presents the results of a series of MD simulations that were used to investigate the structure, dynamics and energetics of Bio-CNT.

Chapter 2 provides introductory and background information about CNT, biopolymers and hybrids composed of these two macromolecules. This chapter also presents an overview of the various computational methods useful for studying nanoscale systems. Chapter 3 contains an outline of the computational tools employed in this thesis: molecular dynamics (MD), thermodynamic integration (TI) and replica exchange molecular dynamics (REMD).

Chapter 4 presents a computational study of DNA-carbon nanotube hybrids (DNA-CNT). MD shows that this material spontaneously self-assembles in aqueous solution due to the formation of high affinity contacts between DNA bases and the CNT sidewall.^{25, 26} Free energy calculations using TI demonstrate that these contacts are due to π - π stacking, an interaction due to van der Waals forces between aromatic species. MD was also used

to test the stability of several proposed DNA wrapping conformations that were based on experimental observations. The simulations show that long DNA sequences prefer helical wrapping about CNT.²⁵ The wrapping is driven by an interplay of electrostatic and torsional interactions within the DNA backbone. Additionally, MD shows that multiple poly GT DNA sequences prefer to adsorb to CNT separately and not in a dimer configuration that was initially proposed by experimentalists. REMD was employed to compute the free energy surface and study the full ensemble of DNA conformations in a DNA-CNT composed of a short 14-base long oligonucleotide.²⁷ At low temperature, the free energy surface is rugged with six energy minima. At room temperature, each energy minimum is significantly populated, indicating that DNA-CNT contains significant structural disorder with contributions from multiple DNA configurations. Additionally, thermal fluctuations and steric limitations limit base-CNT binding; at room temperature, the hybrid contains multiple unbound bases.

Chapter 5 presents a computational study of a nanobiosensor consisting of CNT covalently attached to the coxsackie-adenovirus receptor (CAR).²⁸ The results show that CAR retains its biologically active form even when bound to CNT and remains capable of specifically binding Knob proteins from the adenovirus capsid. In this study, we show that usefulness of MD in the design and understanding of nanobiosensing devices.

Chapter 6 concludes this thesis with a summary of the work, presented results and offers an outlook for future study.

Chapter 2

Biopolymer-Carbon Nanotube Hybrids

Nanotechnology strives to employ the unique properties of nanoscale materials for the development of revolutionary technologies. Some of the materials with the highest potential to fulfill this goal are carbon nanotubes, due to their extraordinary mechanical, electrical and optical properties, and biopolymers, due to their biological significance and molecular recognition capabilities. Years of careful research have provided a comprehensive understanding of these materials alone. Recent work has focused on combining the two in order to incorporate the useful properties of each into a single, multifunctional biopolymer-carbon nanotube hybrid (Bio-CNT). As research in nanoscience and especially in Bio-CNTs proceeds, reliable methods for the characterization of molecular sized objects are needed. Computer simulation is an invaluable tool that provides a detailed, atomic-scale probe of the structure and nature of nanoscale systems, but has been underexploited for study of Bio-CNT. This chapter provides an introduction to carbon nanotubes, biopolymers, Bio-CNT and how these nanomaterials can be understood through computation.

2.1 Carbon Nanotubes

Carbon atoms possess a versatile electronic structure that can reside in a number of hybridization states such as sp , sp^2 and sp^3 .²⁹ This enables carbon atoms to form a variety of covalent bonds (single, double and triple bonds) which results in a multitude of

carbonaceous materials with diverse physical properties. These materials include diamond, graphite, amorphous carbon, buckyballs and carbon nanotubes.

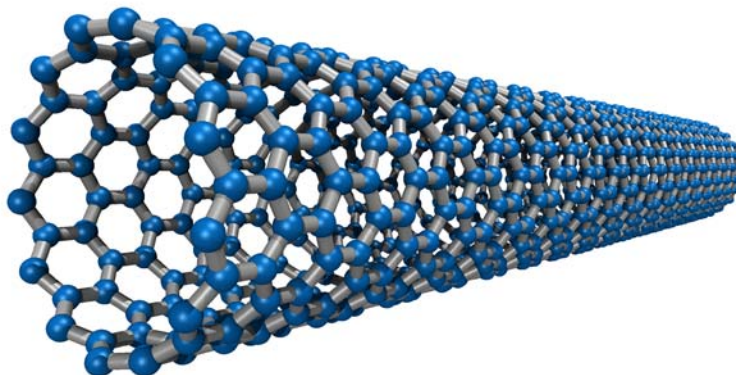


Figure 2.1: A single-walled carbon nanotube. Carbon atoms and the covalent bonds between them are shown in blue and gray, respectively.

Carbon nanotubes (CNT) are cylindrical tubes of carbon atoms with diameters of about 1 nm (Figure 2.1). Since their discovery in transmission electron microscope images in 1991,³⁰ the physical properties of CNTs have been extensively studied. CNTs can exist in single- (Figure 2.1) or multi-walled form (Figure 2.2). Despite their small diameters, CNTs vary in length from ~ 10 nm all the way up to ~ 1 cm.³¹ CNTs are exceptionally strong and stiff; they have the largest tensile strength and elastic modulus of any known material.³²⁻³⁴ Because they consist of hollow tubes of relatively low mass carbon atoms, CNTs are extremely lightweight. These properties make CNTs ideal for use in lightweight, high strength composite materials.^{35, 36} In fact, Easton Sports has incorporated CNTs into composite hockey sticks, baseball bats and bike frames.³⁷ CNTs have extraordinary electronic properties and can act as semiconductors or metals depending on their underlying atomic structure (see Sections 2.1.1 and 2.1.2). Remarkably, the carrier mobility of semiconducting CNTs and the electrical conductivity of metallic CNTs are among the largest of known materials.²⁹ Owing to their small size

and high aspect ratio, CNTs are excellent field emitters and suitable for flat panel displays³⁸ and X-ray sources.³⁹

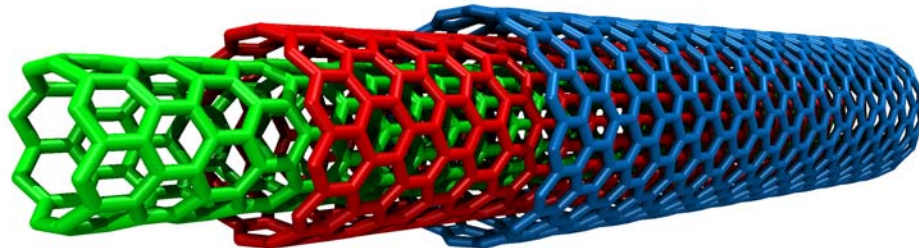


Figure 2.2: A multi-walled carbon nanotube.

One of the most compelling applications of CNTs lies in nanoscale molecular sensing. The electronic transport properties of semiconducting CNTs can be easily controlled with external electric fields.¹ Additionally, because these nanomaterials consist of single tubular sheets, all carbon atoms and electron states associated with conduction reside on the surface and thus, the electronic properties of semiconducting CNTs are highly susceptible to minute changes in the local electrostatic environment. Because of this property, CNT field-effect transistors (FET) have been employed to detect low concentrations of a number of gaseous chemicals including NH_3 , NO_2 , H_2 and O_2 .^{2, 40, 41} Each of these molecules produces large changes in the electrical resistance and current-gate voltage characteristic of the CNT device. Additionally, controlling the chemical affinity (i.e. the magnitude of electrical response produced by a given chemical) of such devices has been demonstrated by functionalizing CNT-FETs with biopolymers (see Section 2.3).^{21, 22} These results provide support that label-free, electronic detection of small concentrations of molecular analytes can be achieved with CNT devices.

Below are further details about the structure and electronic properties of CNTs. Additional information can be found in several textbooks.^{29, 42}

2.1.1 Structure

A CNT can be thought of as a single sheet of graphite (graphene) rolled up into a seamless molecular cylinder (Figure 2.3). \vec{C} is the chiral vector and indicates the direction of the rolling.

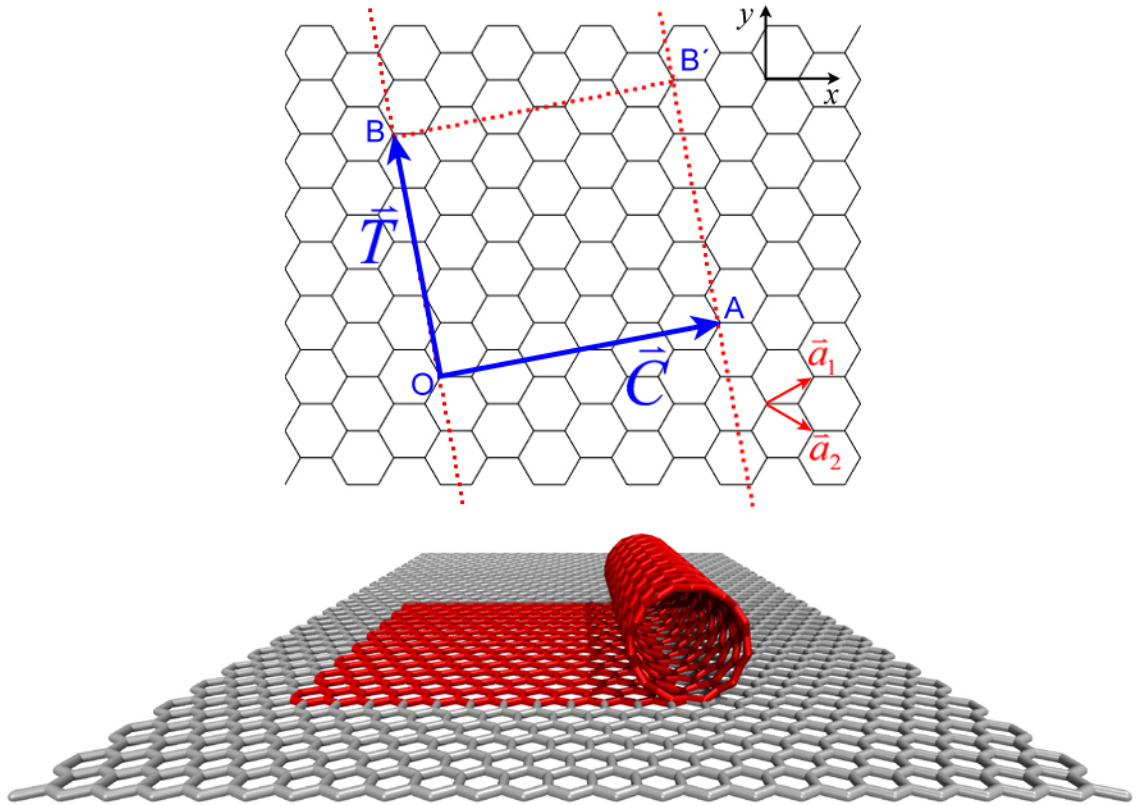


Figure 2.3: Rolling a graphene sheet into CNT. The CNT is constructed by connecting OB and AB' . \vec{C} is the chiral vector whose length equals the CNT circumference. \vec{a}_1 and \vec{a}_2 are the basis vectors for the graphene honeycomb lattice. \vec{T} is the translation vector whose length equals the unit cell length of the resulting CNT.

\vec{C} extends from one carbon atom to a crystallographically equivalent atom on the graphene lattice. Thus, \vec{C} can be written as a linear combination of the lattice basis vectors:

$$\vec{C} = n\vec{a}_1 + m\vec{a}_2 \quad (2.1)$$

Here, n and m are positive integers known as the chiral indices. All physical properties of a given CNT ultimately depend on these two numbers. CNTs of different atomic structure result for various choices of n and m (Figure 2.4). The length of \vec{C} equals the circumference of the resulting CNT. Thus, the CNT diameter d is given by

$$d = \frac{|\vec{C}|}{\pi} = a\sqrt{n^2 + nm + m^2} \quad (2.2)$$

Here, $a = 0.249$ nm is the lattice constant of the graphene honeycomb lattice.

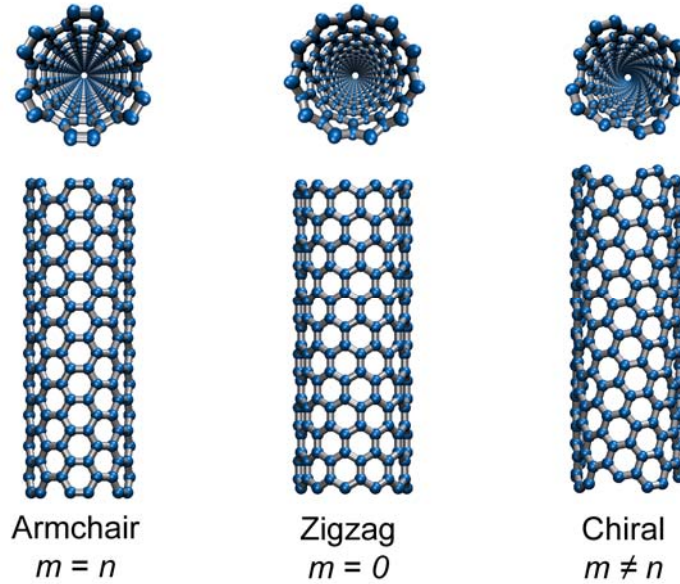


Figure 2.4: Various atomic structures of CNT. High symmetry armchair and zigzag CNTs occur for $m = n$ and $m = 0$, respectively.

The translation vector \vec{T} connects two equivalent carbon atoms along the CNT axis and is given by

$$\vec{T} = \frac{(2m+n)\vec{a}_1 - (2n+m)\vec{a}_2}{\text{gcd}(2n+m, 2m+n)} \quad (2.3)$$

where gcd is the greatest common divisor function. The length of \vec{T} gives the length of the CNT unit cell and can be written

$$|\bar{T}| = \frac{\sqrt{3}d}{\text{gcd}(2n+m, 2m+n)} \quad (2.4)$$

These definitions have been implemented in Nanotube Builder 1.0, a home-written plugin we developed for Visual Molecular Dynamics (VMD)⁴³ that generates the 3D of CNT of arbitrary length and chirality.

2.1.2 Electronic Properties

Electrons in a crystal are only able to occupy states within an allowed band of energy levels. The collection of allowed levels, known as the band structure, is specified by a dispersion relation that gives the energy E of an electron propagating with wave number \bar{k} .⁴⁴ An approximate band structure of a CNT can be derived from that of graphene. Each carbon atom has four valence electrons. In graphene (and CNT), three of them hybridize in the sp^2 state to form σ bonds with neighboring atoms. The fourth electron lies perpendicular to the carbon surface in a $2p_z$ orbital and is involved in π bonding. These π electrons are delocalized throughout the lattice and are most responsible for the material's electronic transport. The energy dispersion relation⁴⁴ $E(k_x, k_y)$ for π electrons in graphene can be obtained via the tight binding approximation^{29, 44} and is given by⁴⁵

$$E(k_x, k_y) = \pm \gamma_0 \sqrt{1 + 4 \cos\left(\frac{3k_x a}{2}\right) \cos\left(\frac{\sqrt{3}k_y a}{2}\right) + 4 \cos^2\left(\frac{\sqrt{3}k_y a}{2}\right)} \quad (2.5)$$

Here, $\gamma_0 = 2.5$ eV is the nearest-neighbor overlap integral.²⁹ This relation is plotted in Figure 2.5. The graphene valence and conduction bands touch at exactly six points. Moreover, because the graphene honeycomb lattice is composed of a two atom basis, each unit cell contributes two π electrons and the valence band is completely filled. Thus,

the Fermi level resides at the interface of the valence and conduction bands making graphene a zero band-gap semiconductor.

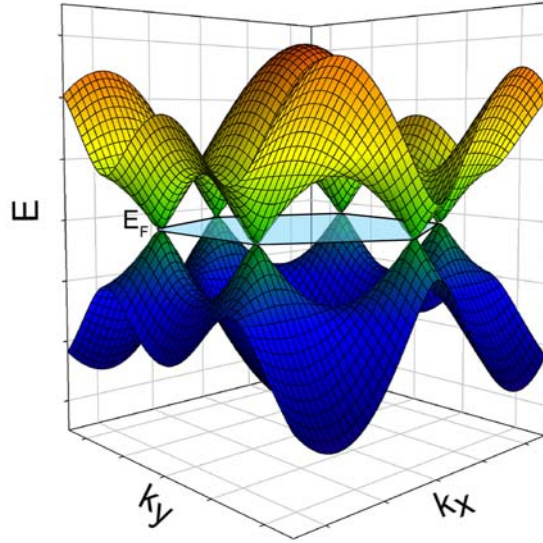


Figure 2.5: Band structure for graphene π electrons. The Fermi level (blue hexagon) E_F is located where the valence (bottom surface) and conduction (top surface) bands touch.

After graphene has been rolled up into CNT, electron momenta in the circumferential direction must fit an integer number of wavelengths within the circumference. Thus, momentum will be quantized in this direction and will be subject to the condition

$$\vec{C} \cdot \vec{k} = 2\pi p \quad (2.6)$$

for any integer p . Imposing this condition on the graphene dispersion relation (Equation 2.5) produces an energy gap E_g for all CNTs with $(n-m) \bmod 3 \neq 0$. Thus, metallic CNTs result for $(n-m) \bmod 3 = 0$ and semiconducting CNTs result if otherwise. It follows that all armchair CNTs are metallic while most zigzag and chiral CNTs are semiconducting. For a random sample of CNTs, two-thirds will be semiconducting while one-third will be metallic. The band-gap is related to the diameter via

$$E_g = \frac{\gamma_0 a}{d\sqrt{3}} \quad (2.7)$$

2.1.3 Challenges in Carbon Nanotube Synthesis and Purification

CNTs can be synthesized by arc discharge⁴⁶, where a plasma is generated between graphite rods, laser ablation,⁴⁷ where a high-intensity laser is focused on graphite or chemical vapor deposition,⁴⁸ where a carbon-containing gas is heated in a furnace in the presence of a catalyst. These methods can result in large amounts of impurities which must be subsequently removed using purification processes.^{49, 50} Additionally, these synthesis methods typically produce a random assortment of metallic and semiconducting CNTs of various lengths and diameters. These limitations present significant challenges for applications that require high-purity, individually addressable CNTs of similar size and electronic character. In order to meet these requirements, the science community has sought after cost-effective methods for the synthesis of monodisperse CNT samples.⁵¹

Dispersing CNTs in solvent is favorable as it enables access to a number of solution based sorting/separation strategies.⁵²⁻⁵⁵ However, owing to their hydrophobic nature and mutual attraction via van der Waals forces, CNTs tend to aggregate in bundles in a solution environment. It was demonstrated that CNTs could be solubilized in water by covalent modification of the sidewall with polar groups.^{56, 57} However, these schemes have the disadvantage of disrupting the delocalized π electron network that is responsible for the extraordinary properties of pristine CNTs.⁵⁸ Thus, focus has shifted towards CNT solubilization using surfactants which can attach without the formation of chemical bonds. Among the successful surfactants are sodium dodecylbenzene sulfonate,^{59, 60} sodium cholate and sodium dodecyl sulfate.⁶¹ More recently, effective CNT solubilization and even separation according to electronic character has been achieved using biopolymers.¹⁷⁻¹⁹ This is described in more detail in Section 2.3.

2.2 Biopolymers

Polymers are large chain-like molecules constructed of repeating chemical units (monomers) connected by covalent bonds. Biopolymers are a class of polymers that are synthesized by the molecular machinery of living organisms. These include polysaccharides, proteins, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA).⁶² Biopolymers tend to have complex, well-defined 3D structures that enable them to carry out specific cellular functions including molecular recognition. Molecular recognition is the selective binding of two complementary molecules via non-covalent interactions such as hydrogen bonding, van der Waals or electrostatic forces, metal coordination or hydrophobic effects.⁶³ Additionally, biopolymers synthesized *in vivo* and *in vitro* are monodisperse (i.e. all composed of exactly the same number of monomers). Synthetic polymers, on the other hand, tend to be polydisperse (i.e. composed of a variable number of monomers) and have more disordered structures and nonspecific molecular interactions.

The structure of a given biopolymer can be studied on multiple levels. The primary structure specifies the particular sequence of monomeric units that make up the biopolymer. The secondary structure is the local structural motif of segments of the biopolymer. A biopolymer's tertiary structure refers to its global 3D structure. The quaternary structure indicates the arrangement of multiple, interacting biopolymers.

Below are additional details about two of the most important biopolymers: DNA and proteins. Further details can be found in several textbooks.^{62, 64, 65}

2.2.1 DNA

DNA is a complex biological heteropolymer that encodes genetic information in living organisms.⁶⁵ A DNA strand, whose chemical structure is shown in Figure 2.6, consists of a sequence of nucleotides. A nucleotide is composed of a phosphate, a sugar known as deoxyribose and one of four bases: adenine (A), cytosine (C), guanine (G) or thymine (T). A strand of DNA is constructed by connecting the phosphates and sugars of adjacent nucleotides with a covalent bond. This chain of sugar-phosphate linkages is referred to as the backbone. The sugar groups are chiral, meaning their mirror images are not identical. It follows that the DNA backbone is also chiral and has asymmetric termini called the 5' and 3' ends. Typically, DNA sequences are read from the 5' to the 3' end. At neutral pH, phosphates will be deprotonated and the DNA backbone will carry a negative charge $Q = -e(N - 1)$, where N is the number of bases in the sequence and e is the elementary charge. This makes DNA highly soluble in aqueous solution.

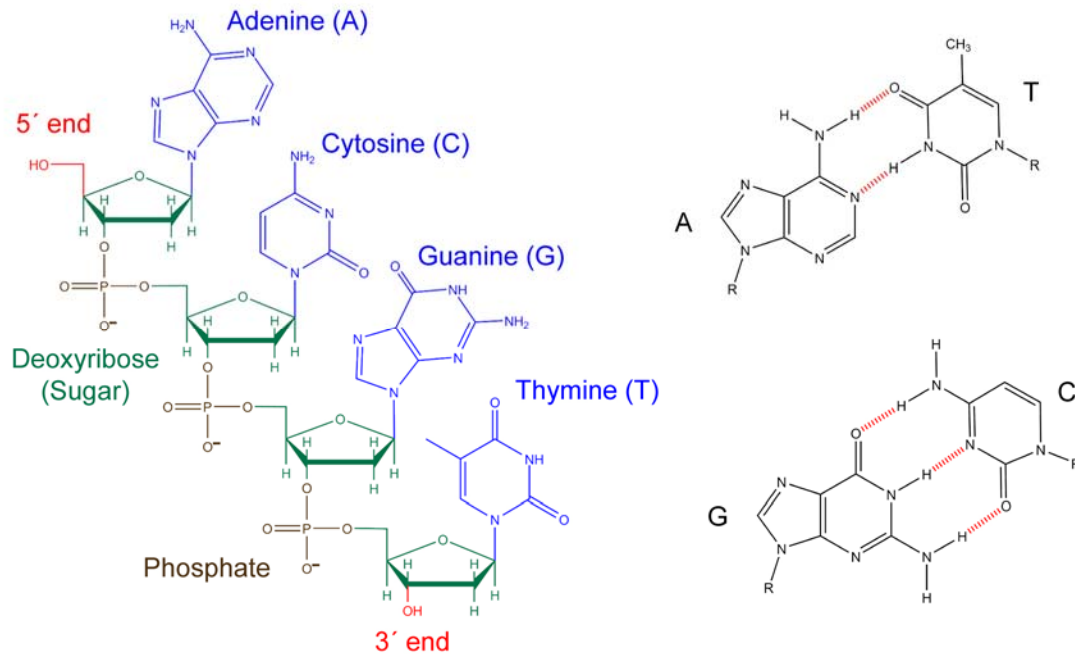


Figure 2.6: (Left) Chemical structure of a DNA strand. (Right) A-T and G-C base pairs.

DNA bases are complementary and can form pairs held together by specific hydrogen bond interactions. In Watson-Crick base pairing, the most common base pairing scheme, A pairs with T and G pairs with C (Figure 2.6). The G-C base pair, with its three hydrogen bonds is more stable than the A-T base pair, which has only two hydrogen bonds. While non-Watson-Crick schemes such as wobble^{66, 67} and Hoogsteen⁶⁵ base pairing exist, they are generally less favorable and not as commonly found.

DNA can exist in single- (ssDNA) or double-stranded (dsDNA) form (Figure 2.7). dsDNA is composed of two antiparallel, complementary ssDNA strands arranged in a double-helix and held together by base pairs. In this helical form, adjacent bases stack on top of each other in a spiral staircase fashion. These stacking interactions provide significant rigidity in dsDNA and are even more important for stabilizing the helical structure than base pairing. dsDNA can assume a variety of helical forms that depend on a complex interplay of effects including sequence, hydration level, type and concentration of dissolved salts, chemical modifications of bases and the presence of polyamines. Several forms of dsDNA are shown in Figure 2.7. In the cell nucleus, dsDNA is well-hydrated and normally assumes a right-handed double-helical B DNA form. Under (non-physiological) dehydrated conditions, the A DNA form occurs. A DNA is also right-handed, but, compared to the B form, is more compact with bases tilted with respect to the helical axis. Under certain circumstances in the cell, bases can become modified with methyl groups to produce left-handed Z DNA. ssDNA, on the other hand, generally has a less ordered structure than dsDNA. For short ssDNA sequences, equilibrium exists between a helical form stabilized by base-base stacking and a random

coil. When two segments of ssDNA are complementary, a folded hairpin structure can form (Figure 2.7).

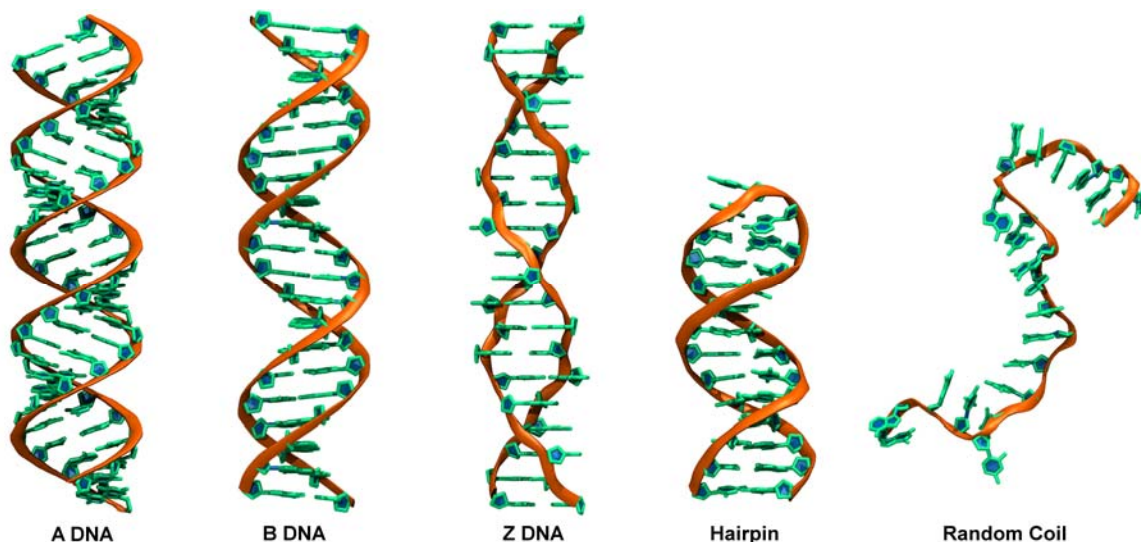


Figure 2.7: Forms of dsDNA (A, B, Z) and ssDNA (hairpin, random coil). The backbone is represented by an orange ribbon. The base and sugar groups are shown in green.

2.2.2 Proteins

Proteins are ubiquitous biopolymers composed of sequences of amino acids that participate in virtually every cellular process.⁶⁴ An amino acid consists of a central carbon atom named C_{α} connected to a hydrogen atom, an amino group (NH_2), a carboxylic acid group ($COOH$) and a side chain (Figure 2.8). Each amino acid has a unique side chain that is either charged (positive or negative), polar or hydrophobic. There are twenty naturally occurring side chains specified by the genetic code. Proteins are assembled by linking the carboxyl and amino groups of adjacent amino acids with a peptide bond (Figure 2.8). This series of peptide bonds composes the protein backbone and serves as a scaffold for the various side chains. Individual amino acids in a protein are referred to as residues. The two protein ends are called the N-terminus and C-

terminus, referring to the end terminated by the amino group and carboxylic acid group, respectively.

In aqueous solution, a protein folds into a complicated 3D structure that depends on the amino acid sequence. Folding buries hydrophobic residues in the interior and exposes charged and polar residues to solvent to produce a highly favorable protein conformation. The function of a protein is specified by its structure. Because of the considerable number of amino acids, proteins come in a vast number of shapes and sizes and carry out a myriad of biological functions. Proteins are vital for metabolism, cell structure and stability, the cell life cycle, immune response, signaling and molecular transport within the cell.

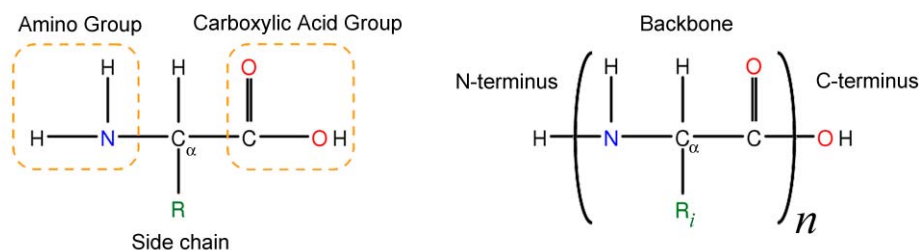


Figure 2.8: Chemical structure of an amino acid (left) and a protein (right). A protein of n amino acids consists of a backbone with a series of side chains projecting outwards.

The creation of regular secondary structural elements enables a protein to assume structures with a hydrophobic core and a hydrophilic surface. The two main types of secondary structure are the α -helix and β -sheet. α -helices are backbone segments assuming a right-handed helix stabilized by hydrogen bonds between the carboxyl carbon of residue i and the amino nitrogen of residue $i + 4$ (Figure 2.9). β -sheets result when two parallel or antiparallel backbone segments adopt extended, linear conformations held together by a distinct hydrogen bond pattern (Figure 2.10). These secondary structural elements are connected by loop regions that reside on the protein surface and are rich in

polar and charged residues. Loops tend to be more disordered and flexible than α -helices and β -sheets (Figure 2.11).

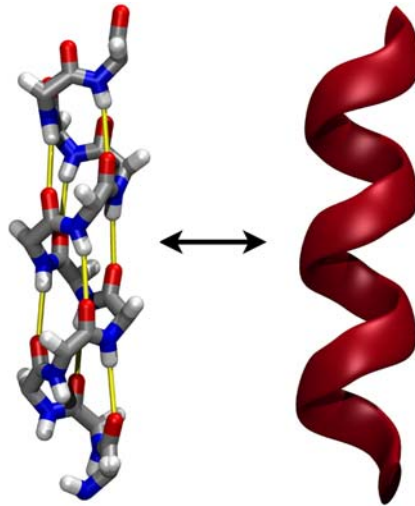


Figure 2.9: (Left) An α -helix consists of hydrogen bonds (yellow) along the backbone between residue i and $i + 4$. Side chains are not shown. (Right) A ribbons representation of an α -helix.

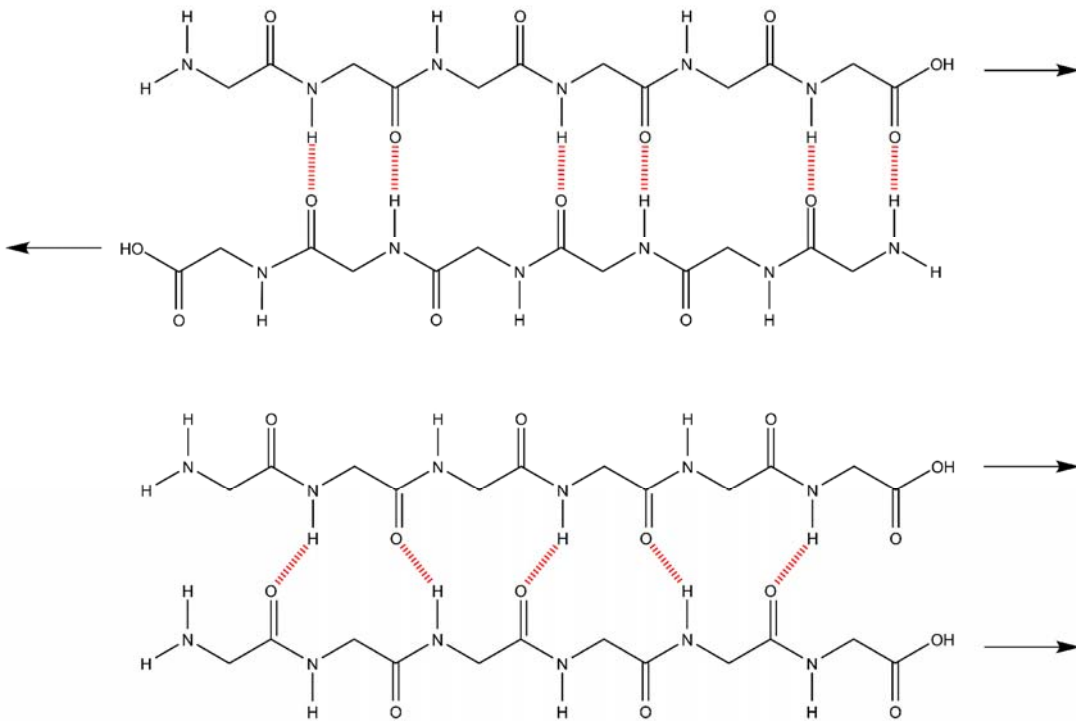


Figure 2.10: Backbone and hydrogen bond pattern (red) of an antiparallel (top) and parallel (bottom) β -sheet. Arrows indicate the direction from the N-terminus to the C-terminus. For clarity, side chains are not shown.

Because proteins assume complicated 3D shapes, they are commonly visualized with drawing methods that emphasize important structural elements. Common styles include the bond, space filling, ribbons and surface visualizations (Figure 2.11).

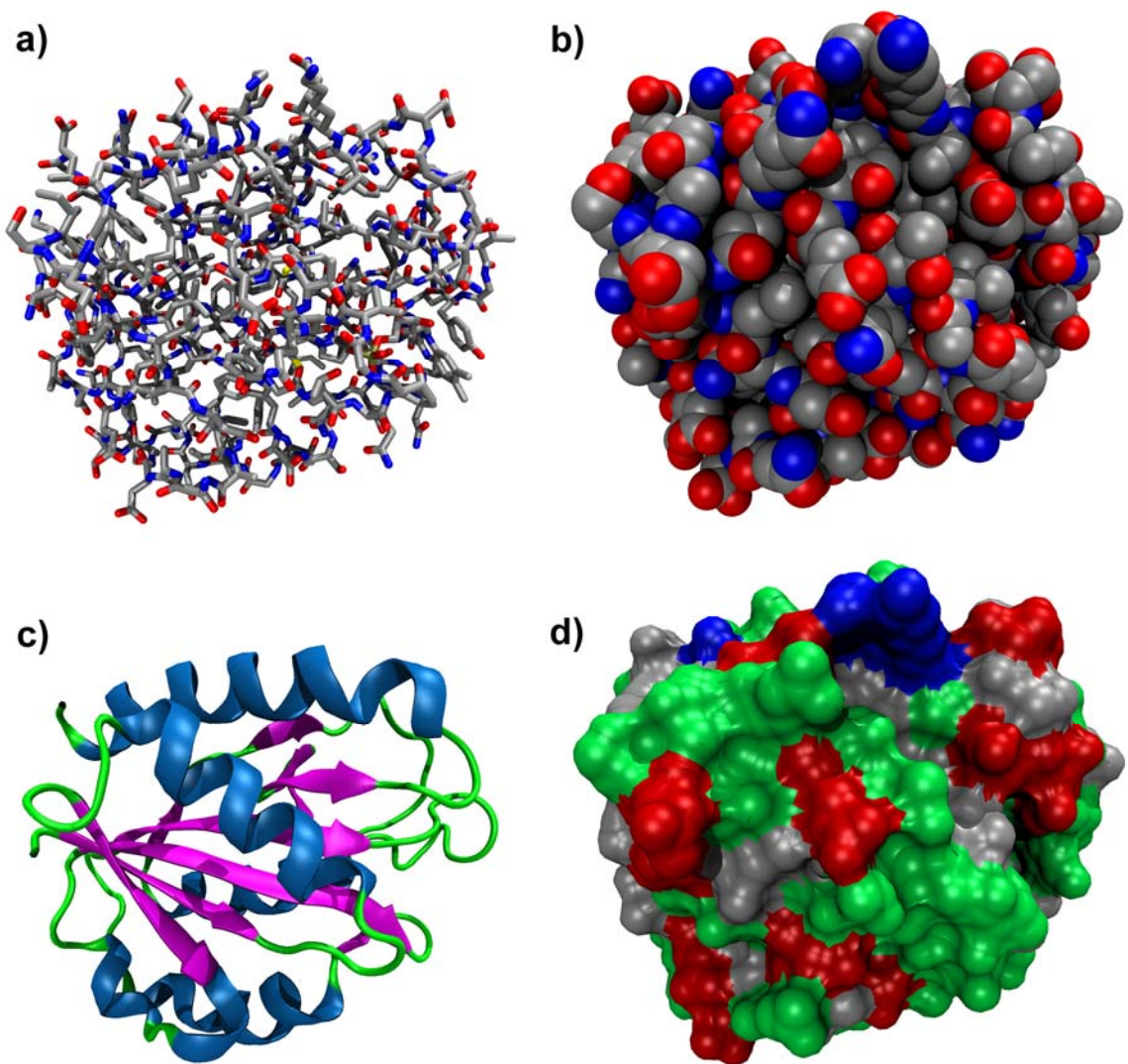


Figure 2.11: Protein visualization styles. (a) Bond and (b) space filling representation showing the atomic detail. Carbon, oxygen, nitrogen and sulfur atoms are colored gray, red, blue and yellow, respectively. Hydrogen atoms are not shown. (c) Ribbons representation showing the secondary structure. α -helices, β -sheets, and loops are colored blue, magenta and green, respectively. (d) Surface representation showing the 3D shape. Basic (positively charged), acidic (negatively charged), polar and hydrophobic regions are colored blue, red, green and gray, respectively.

2.3 Biopolymer-Carbon Nanotube Hybrids

As research in nanotechnology proceeds, the development of versatile classes of molecular materials is needed to fulfill the various technological needs for novel applications. While traditional nanomaterials such as nanocrystals, nanowires, fullerenes and synthetic or biological polymers are intrinsically interesting and technologically useful, they tend to have a predefined set of properties that limits their possible applications. Merging two or more of these nanomaterials into a hybrid has great potential to initiate further advancements in nanotechnology. Not only do hybrids encompass the strengths of each material, they also tend to be “greater than the sum of their individual parts.” That is, in a hybrid, nanomaterials reside in arrangements that do not occur naturally and can result in materials with new and unexpected properties.

A class of hybrid materials composed of biopolymers and CNTs show much promise for a wide range of nanotechnological applications. The structure, function and molecular recognition capabilities of biopolymers combined with the robust mechanical, electrical and optical properties of CNT make biopolymer-carbon nanotube hybrids (Bio-CNT) an intriguing material that may revolutionize many areas of technology. Below, two types of Bio-CNT based on ssDNA and proteins are introduced.

2.3.1 DNA-Carbon Nanotube Hybrids

It was first discovered that ssDNA and CNT spontaneously bind and form stable DNA-carbon nanotube hybrids (DNA-CNT) during CNT solubilization experiments. These experiments demonstrated that, with mild sonication, ssDNA could disperse CNT bundles and solubilize individual CNTs in aqueous solution.¹⁷ Moreover, the resulting

suspension of DNA-CNTs could then be separated according to the electronic character of the underlying CNT using anion exchange chromatography⁶⁸, a process that enables the separation of ions according to their net charge (Figure 2.12).⁶⁹ Because of the charged phosphate backbone, DNA-CNT carries an overall charge. However, due to differences in the dielectric properties (i.e. the way charges are screened) of metals and semiconductors⁷⁰, the effective charge of the hybrid will depend on the electronic character of the CNT. These ssDNA mediated charge differences enable a successful separation of metallic and semiconducting CNTs. This property of DNA-CNT alone holds promise to facilitate developments of CNT based nanotechnology in a very fundamental way by providing a route to monodisperse CNT samples.⁵¹

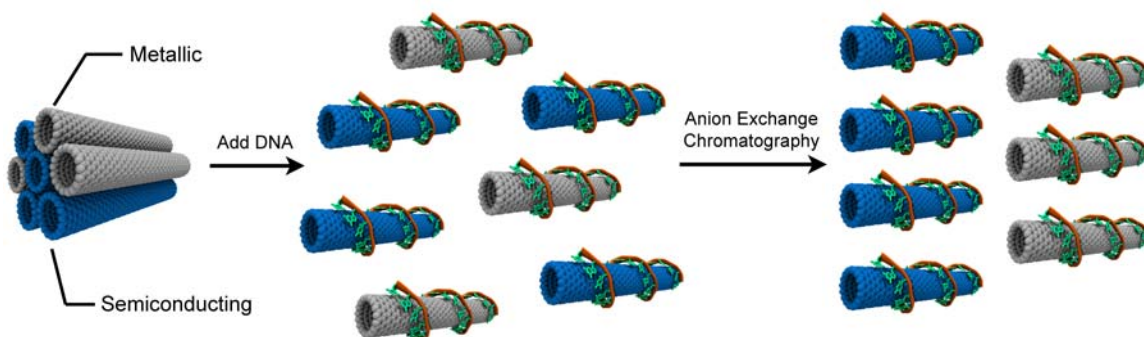


Figure 2.12: DNA-assisted dispersion and separation of CNTs in aqueous solution. ssDNA disperses CNT bundles and results in monodisperse water soluble DNA-CNTs. DNA-CNT can then be separated according to the electronic properties of CNT using anion exchange chromatography.

The solubilization and separation capabilities of ssDNA depend on its sequence; poly T (sequences of repeating thymines) has the highest dispersion efficiency (though arbitrary sequences have comparable performance) while poly GT (sequences of guanine-thymine repeats) provides the best separation by far.^{17, 68} Initial atomic force microscope (AFM) images taken by Zheng *et al.* of DNA-CNT based on poly GT sequences

displayed periodic bands of high and low regions on the surface of the hybrid with a uniform 18 nm spacing (Figure 2.13). This differed from other sequences that showed little or no structural regularity. Because multiple GT-rich sequences are known to hybridize via non-Watson-Crick base pairing interactions⁷¹, Zheng *et al.* hypothesized that poly GT binds to CNT in a duplex configuration that results in a more regular ssDNA conformation. This would yield a more uniform DNA-CNT charge density that would enable improved separation.⁶⁸ However, questions surrounding the nature of sequence dependent DNA-CNT structure have yet to be firmly resolved. AFM measurements by other groups reveal a similar band pattern on the surface of DNA-CNT that is independent of sequence.⁷² Additionally, our computational work (see Sections 4.3 and 4.4 for more details) rules out the possibility of a poly GT duplex adsorbed to CNT and shows that sequence has no discernable effect on global DNA-CNT structure.²⁵

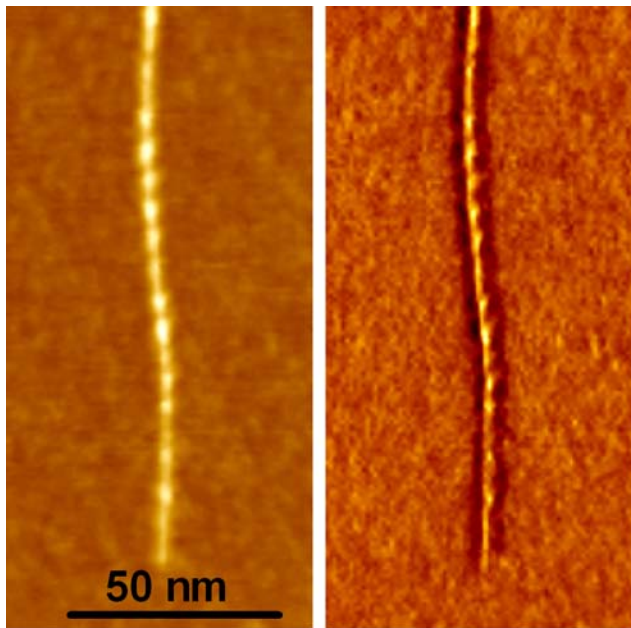


Figure 2.13: AFM image courtesy of Zheng *et al.*⁶⁸ of CNT wrapped with $(GT)_{30}$. The periodic bands are interpreted as ssDNA wrapping helically about CNT with a pitch commensurate with the band spacing.

ssDNA has also been used to improve and expand the chemical sensing capabilities of CNT-FETs. Detectable changes in the electronic properties of conventional CNT-FETs only occur for a limited number of gaseous chemicals. Coating these devices with a nanoscale layer of ssDNA drastically increases sensitivity and enables recognition of an expanded library of molecules. A schematic of these DNA-functionalized CNT-FETs is shown in Figure 2.14.

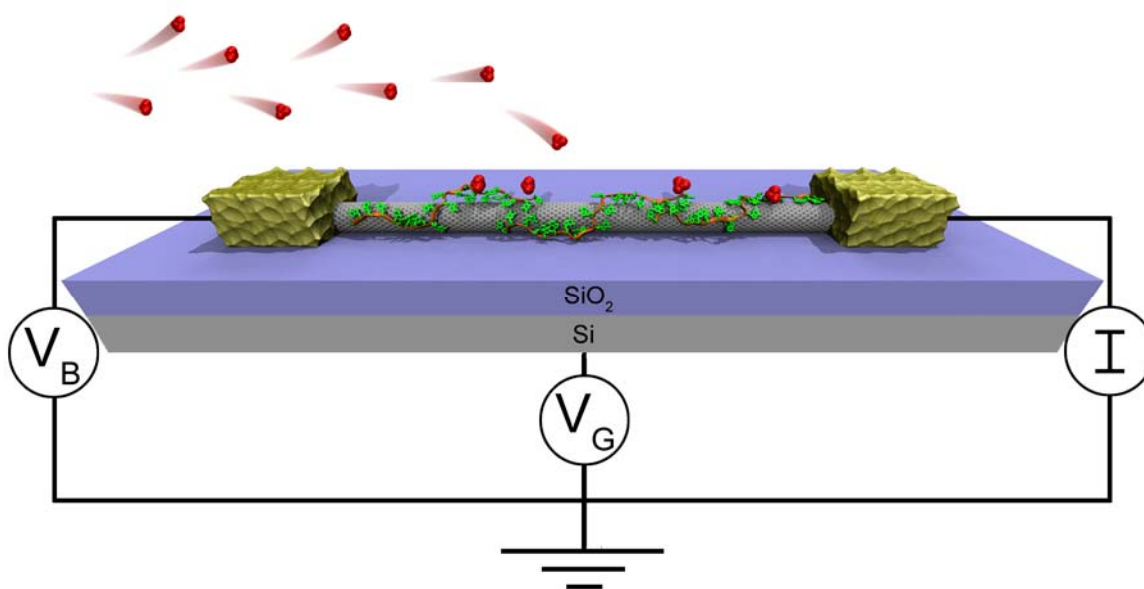


Figure 2.14: CNT-FET functionalized with ssDNA (orange ribbon backbone and green bases). Gaseous molecules (red) flowing in the vicinity of the device interact with the ssDNA-coated CNT and produce changes in the electrical resistivity and current-gate voltage (I - V_G) characteristic.

Interestingly, the sensitivity of these devices to any particular analyte varies with ssDNA sequence. This feature makes these devices ideal for electronic nose applications. An array of these devices each coated with a different ssDNA sequence would produce a multitude of different signals upon exposure to complex mixtures of gases. These signals could then be fed to a neural network programmed to identify the contents of the sample.

Such technology would have far-reaching impact on homeland security, disease diagnosis and environmental safety.

Another important application of DNA-CNT is for the label-free detection of DNA-hybridization. The ability to detect the hybridization of two complementary sequences of DNA has many important applications in microbiology,⁷³⁻⁷⁵ environmental science^{76, 77} and medicine.^{78, 79} It has been shown that DNA-CNT produces an electronic⁸⁰ or optical²⁴ response when a complementary DNA sequence hybridizes with the one bound to CNT.

Because of the importance of DNA-CNT for advancements in nanotechnology, there have been several studies aimed at understanding the structure, interactions and self-assembly of these hybrids. Molecular mechanics calculations using energy minimization principles were employed to locate low energy ssDNA conformations about CNT. These computations showed that ssDNA can reside in a helical wrapping with its bases lying flat (stacked) on top of CNT.¹⁷ *Ab initio* computations have also shown that DNA bases prefer a stacked geometry when bound to CNT.^{81, 82} Experiments have corroborated these results.^{68, 72, 83, 84} However, a truly dynamical understanding of the structure and function of DNA-CNT has been lacking. The computational work presented in the remainder of this thesis has provided an expanded and more complete understanding of this fascinating hybrid material.

2.3.2 Protein-Carbon Nanotube Hybrids

Many naturally occurring proteins will spontaneously adsorb to CNT via hydrophobic interactions⁸⁵ and result in water soluble protein-CNT hybrids.^{18, 19} It is also possible to design synthetic proteins that selectively bind to CNT.⁸⁶ For some proteins, binding can

even be controlled by adjusting the pH of the solution.⁸⁷ Thus, proteins are another promising candidate for controllable solubilization and separation of CNTs in aqueous solution. Protein-CNT hybrids also have great potential for therapeutic applications. It has been shown that CNTs can penetrate cell membranes and transport molecular cargo, including proteins, into the cell.⁸⁸⁻⁹⁰ Once inside the cell these proteins remain capable of performing biological functions even when attached to CNT, making protein-CNT hybrids relevant for drug delivery and gene therapy purposes.⁸⁹ Protein-CNT complexes composed of lysozyme are also of special biomedical interest. Lysozyme⁹¹, a powerful antibacterial protein, will spontaneously bind and solubilize CNT. The resulting lysozyme-CNT hybrids combine the mechanical robustness of CNTs with the enzymatic activity of lysozyme and have use as an antimicrobial coating.¹⁶

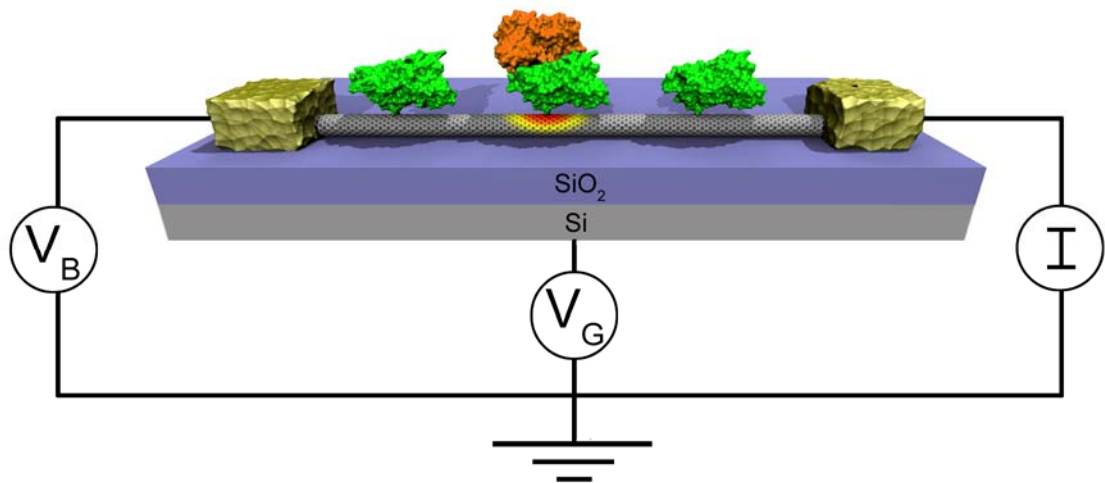


Figure 2.15: Nanoscale biological sensor constructed of a CNT-FET functionalized with receptor proteins (green). Ligands (orange) bind to the receptor and alter the electronic characteristics of CNT.

Another exciting use of protein-CNT hybrids is as nanoscale biological sensors.^{22, 23} These devices typically consist of a CNT-FET functionalized with receptor proteins. The binding of a ligand (which could be a complementary protein, a hormone, a drug or a

toxin) to the receptor modifies the electronic transport properties of the CNT-FET. Thus, the presence of ligands can be detected using a label-free scheme with electronic readout. A schematic of these devices is shown in Figure 2.15.

It is also possible to attach proteins to CNTs by chemical means. Covalent attachment is oftentimes preferable, as it provides a universal method to immobilize the protein on CNT. Diimide-activated amidation is the chemical process that is widely used for this purpose (Figure 2.16).⁹² In this process, carboxylic acid defects are introduced on the surface of CNT via treatment with nitric or sulfuric acid. Stable esters are then created at these defect sites with *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDAC) and *N*-hydroxysuccinimide (NHS). Amine groups located on protein side chains will react with these esters resulting in a chemical bond between the protein and CNT.

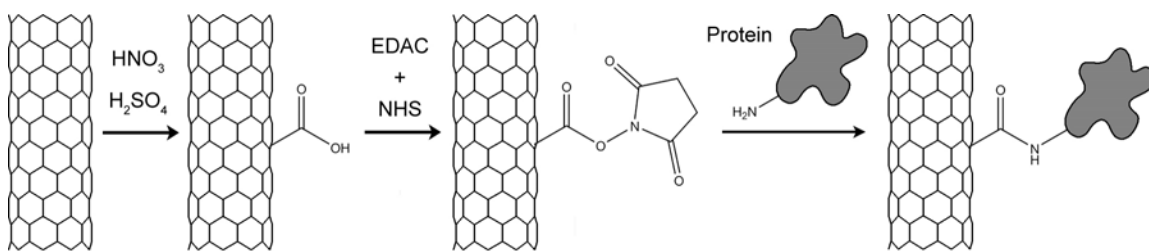


Figure 2.16: Diimide-activated amidation used for covalently attaching proteins to CNT.

2.4 Computer Simulations in Nanoscience

Computer simulation is an excellent tool in nanoscience that complements experimental study, but has been underexploited in Bio-CNT research. There exists a spectrum of computational techniques that enable investigation of nanoscale systems at various levels of precision (Figure 2.17). No method by itself can provide complete information about a system. Therefore, picking the right methodology is of paramount importance in order to answer the most important questions for a given system.

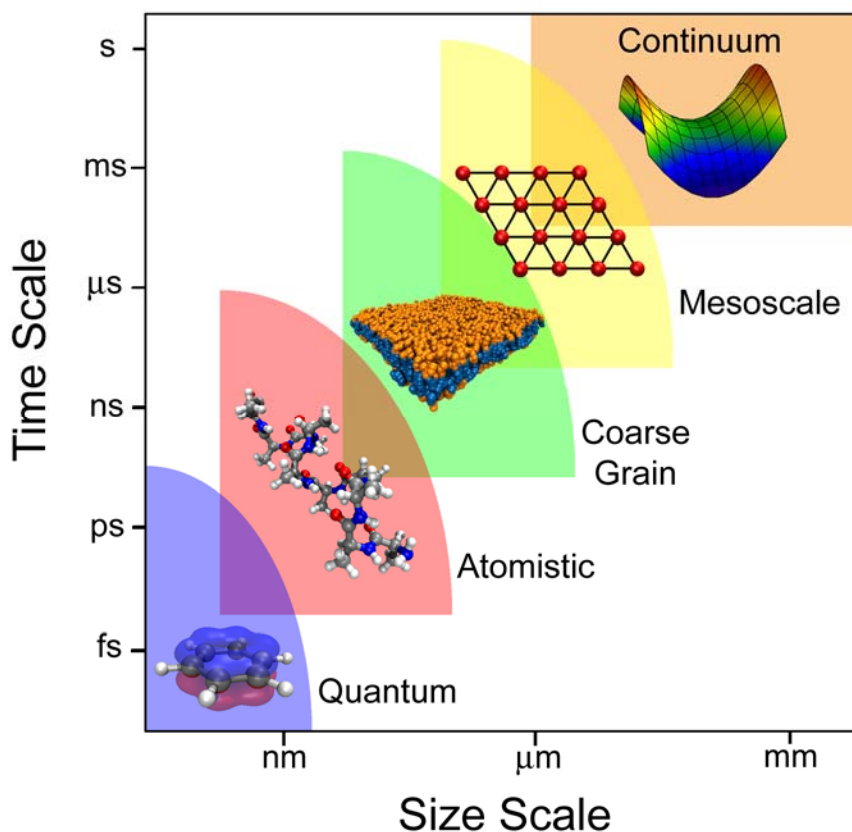


Figure 2.17: Time and length scales accessible to different simulation techniques.

Ab initio (meaning “from first principles”) methods, such as Hartree-Fock⁹³ and density functional theory,⁹⁴ employ quantum mechanics to study the nature of condensed matter at the electronic level. These methods rely very little on fits to phenomenological models or experimental data and therefore, are typically extremely accurate. However, they have the drawback of being computationally intensive and can only be routinely used to study small systems (~100 atoms) over short time scales (~10 ps).

Atomistic methods, such as molecular dynamics and Monte Carlo, simulate a system of atoms by utilizing models derived from a combination of *ab initio* computations and experimental data.⁹⁵ Atomistic simulations assume particle motion is governed by classical mechanics. While this method contains many approximations, it provides a good

compromise between computational efficiency and overall accuracy. System sizes and time scales that can be comfortably handled by this method are 10^3 – 10^6 atoms and 1–1000 ns, respectively.

Coarse grain methods use models similar to those in atomistic methods except that, in coarse grain simulations, each particle represents a collection of atoms. Typically each coarse grain particle represents three to four heavy (non-hydrogen) atoms and thus, detailed atomic-scale motion (e.g. bond vibration), structure and interactions (e.g. hydrogen bonding) are averaged out. This considerably reduces the complexity of the system and enables simulations of large systems (~ 100 nm) for long time scales (~ 1 μ s) while retaining the granularity of nanoscale systems.

Mesoscale methods are particle based methods that enable study of mesoscopic (~ 1 μ m) systems. In this method, each particle represents tens to even thousands of atoms. Solvent effects are usually handled in an implicit way by applying dissipative (viscous) and random (due to Brownian motion) force to each particle. Mesoscale simulation techniques include dissipative particle dynamics and Lattice Boltzmann methods.⁹⁵

Continuum methods describe a system by partial differential equations and thus, are useful when the discreteness of matter can be approximated by a continuum. These methods oftentimes employ finite element techniques to obtain approximate solutions to the differential equations.⁹⁶

Atomistic simulation, specifically molecular dynamics, is the most desirable for Bio-CNT. This hybrid nanomaterial is constructed of biopolymers that contain intricate local geometries and highly specific short-range (hydrogen bonding) interactions. Thus, retaining atomic detail of Bio-CNT is vital for a complete understanding of its structure

and function. An introduction to the molecular dynamics techniques used in this thesis is given in the next chapter.

Chapter 3

Molecular Dynamics

Atomistic molecular dynamics simulation, hereafter referred to as molecular dynamics (MD), is an extremely versatile and powerful method for computing the equilibrium and dynamical properties of a system of N atoms in a volume V .^{95, 97} This method provides a window into the microscopic behavior of a nanoscale system by computing the trajectory of each atom over a specified period of time. There exist a number of freely available MD software packages accessible on the Internet that are compatible with many operating systems and parallel computing architectures.

In MD, atoms behave according to the laws of classical mechanics. In most cases, this is an excellent approximation that only breaks down when considering the motion of light atoms/molecules (H_2 , He, D_2) or the statistical mechanics of high frequency vibrations.⁹⁵ MD also employs the Born-Oppenheimer approximation,⁹⁸ which assumes that electrons always reside in their ground states. This typically means that MD can not be used to study excited states, electron transfer or chemical reactions. Additionally, atomic interactions are specified by potential energy functions derived from *ab initio* calculations and experimental data (see Section 3.2). Despite these limitations, MD remains advantageous over other methods because it can simulate large ($\sim 10^6$ atoms) systems over considerable timescales ($\sim 1 \mu s$) in atomic detail, thus enabling simulation of biologically relevant systems.⁹⁹ This chapter contains a brief summary of the methods and algorithms employed in MD as well as an overview of two special MD techniques used in this thesis: thermodynamic integration and replica exchange molecular dynamics.

3.1 Methodology

A schematic of a general MD simulation is given in Table 3.1.

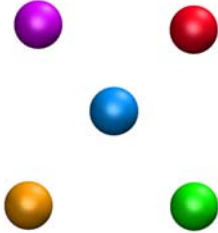
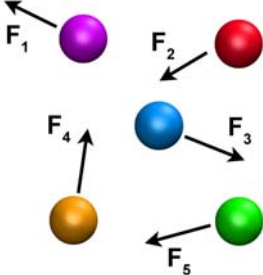
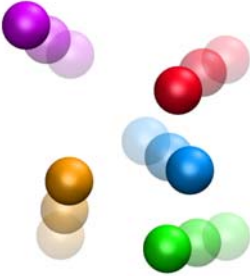


<p>1. Initialize system</p> <p>Specify potential $U(\vec{r}_1, \dots, \vec{r}_N)$ for atomic interactions (force field) as function of positions of N atoms</p> <p>Specify initial positions and velocities</p> <p>Specify run parameters (volume, temperature, pressure, etc.)</p>	
<p>2. Calculate forces, temperature and pressure</p> <p>Net force on atom i given by:</p> $\vec{F}_i = \sum_j \vec{F}_{ij} = -\vec{\nabla}_i U$	
<p>3. Update configuration</p> <p>Determine new atomic positions by solving Newton's equations of motion for each atom:</p> $m_i \frac{d^2 \vec{r}_i}{dt^2} = \sum_i \vec{F}_i$ <p>Regulate temperature and pressure.</p>	
<p>4. Output</p> <p>Write positions, velocities, energies, temperature, pressure, etc.</p>	
<p>5. Iterate steps 2-4 for desired number of steps</p>	

Table 3.1: MD algorithm.

3.1.1 Initialization of the System

Coordinates for proteins and DNA can be obtained from experimentally resolved crystal structures published on the Protein Data Bank (www.pdb.org) or the Nucleic Acid Data Bank (ndbserver.rutgers.edu). Because of its structural regularity, coordinates of double-stranded DNA can also be generated in a variety of helical forms using programs such as AMBER (Assisted Model Building with Energy Refinement).¹⁰⁰ Solutes are then placed in an empty simulation box of appropriate dimension. The box should be large enough so that periodic images of the solute lie outside the cutoff distance (see Section 3.1.2). The box is then solvated by a random placement of solvent molecules using programs such as AMBER¹⁰⁰ or Visual Molecular Dynamics (VMD).⁴³ This step usually generates a number of close (unphysical) contacts between solute and solvent, resulting in large repulsive forces. Thus, to generate more realistic initial coordinates and ensure the stability of the subsequent MD simulation, systems must undergo geometry optimization using algorithms such as steepest descents¹⁰¹ that translate the system to the closest local energy minimum.

3.1.2 Force Calculation

Summation of forces on each particle is the most costly portion of an MD simulation. Because the forces are pairwise additive, for a system of N particles, $N(N-1)/2$ pair distances must be evaluated. In order to simulate bulk phases (e.g. solute in a macroscopic volume of solvent), periodic boundary conditions (PBC) are applied. That is, the volume containing N atoms is treated as the primitive cell in an infinite periodic lattice of identical cells (Figure 3.1). While these conditions are unphysical, they only

present small errors for large systems. With PBC, an atom i interacts with all other atoms in this infinite crystal lattice including its own periodic image. Obviously, summing an infinite number of forces is impossible. Fortunately, atomic interactions are typically short-ranged and the interaction of two atoms separated by a large enough distance r_c (the cutoff distance) can be ignored. Using cutoffs, atom i only interacts with atoms, including those in the next periodic box, that are within r_c (Figure 3.1).^{95, 97} The simulation box is chosen such that the shortest dimension is larger than twice r_c , which ensures that atoms do not interact with their periodic images.

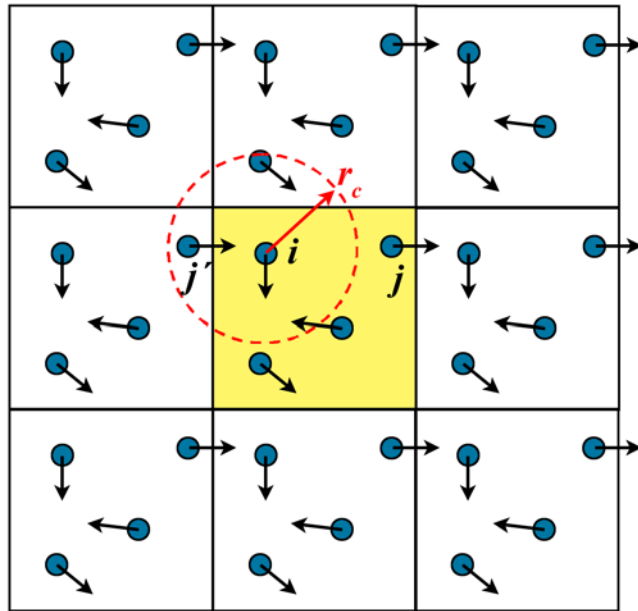


Figure 3.1: Schematic of periodic boundary conditions. The primitive cell is shown in yellow. Atom i interacts only with those atoms within a cutoff distance r_c .

While conceptually, simulations using PBC contain an infinite number of atoms, in practice one only has to compute trajectories for the original N atoms of the primitive cell. To illustrate this, consider the motion of atom j (Figure 3.1). As j diffuses into the periodic box on the right, its image j' from the left periodic box enters the primitive cell. Therefore, given the trajectory of an atom from the primitive cell, trajectories of all its

images are determined. Thus, PBC enables efficient simulation of atoms that experience forces and dynamics consistent with a homogeneous bulk system.

Electrostatic interactions are long-ranged and can not be simply cut off.^{95, 102} There exist a number of special algorithms for handling electrostatic interactions within the PBC framework. These algorithms include Ewald and Particle Mesh Ewald methods.^{95,}

102

3.1.3 Integration of the Equation's of Motion

When the forces acting on each atom are known, the trajectory can be calculated via Newton's 2nd Law:

$$m_i \frac{d^2 \bar{r}_i}{dt^2} = \sum_i \bar{F}_i \quad (3.1)$$

There exist a number of algorithms to numerically solve this equation at discrete time steps Δt . A simple and effective choice that provides stable solutions for a large number of atoms is the Verlet algorithm.⁹⁵ Deriving this algorithm is simple. First, a Taylor expansion of an atom's position $\bar{r}(t)$ is taken around time t :

$$\begin{aligned} \bar{r}(t + \Delta t) &= \bar{r}(t) + \bar{v}(t)\Delta t + \frac{\bar{F}(t)}{2m}\Delta t^2 + \frac{\ddot{\bar{r}}}{6}\Delta t^3 + O(\Delta t^4) \\ \bar{r}(t - \Delta t) &= \bar{r}(t) - \bar{v}(t)\Delta t + \frac{\bar{F}(t)}{2m}\Delta t^2 - \frac{\ddot{\bar{r}}}{6}\Delta t^3 + O(\Delta t^4) \end{aligned} \quad (3.2)$$

Summing these two equations gives,

$$\bar{r}(t + \Delta t) + \bar{r}(t - \Delta t) = 2\bar{r}(t) + \frac{\bar{F}(t)}{m}\Delta t^2 + O(\Delta t^4) \quad (3.3)$$

Rearranging, and dropping the highest order terms in Δt gives the Verlet algorithm:

$$\vec{r}(t + \Delta t) \approx 2\vec{r}(t) - \vec{r}(t - \Delta t) + \frac{\vec{F}(t)}{m} \Delta t^2 \quad (3.4)$$

Thus, given the position \vec{r} of an atom at times t and $t - \Delta t$ and the total force \vec{F} at time t , the position at a future time $\vec{r}(t + \Delta t)$ can be computed. This algorithm contains an error of order Δt^4 . To ensure accuracy, the time step Δt must be significantly larger than the oscillation period τ of the system's highest frequency vibrational motion. Typically, $\Delta t = 1.0 - 2.0$ fs.

3.1.3 Temperature and Pressure Regulation

Because atomic interactions are described by conservative forces (see Section 3.2), the total energy E is a constant of motion. However, systems in thermal equilibrium under ambient conditions experience a constant temperature T and pressure P . Thus, simulations are performed in the NPT -ensemble (constant number of particles, temperature and pressure). To do this, thermostats and barostats are employed to regulate T and P , respectively. The microscopic definitions of T and P are given by⁹⁵

$$T = \frac{2}{3Nk_B} \langle K \rangle \quad (3.5)$$

$$P = \frac{1}{V} \left[Nk_B T + \frac{1}{3} \left\langle \sum_{i>j} \vec{f}(r_{ij}) \cdot \vec{r}_{ij} \right\rangle \right]$$

Here, $\langle K \rangle$ is the average kinetic energy of the system and $\vec{f}(r_{ij})$ is the force between atoms i and j at a distance r_{ij} . Thus, T and P are related to the atom velocities (through $\langle K \rangle$) and system volume V , respectively. Thus, thermostats and barostats rescale the atom velocities and box volume during a simulation to maintain constant T and P . A detailed description of the algorithms that perform these functions is given elsewhere.⁹⁵

3.2 Atomistic Force Field

Accurate models for atomic forces are crucial for a realistic simulation of condensed matter. Fundamentally, atomic interactions are solely electrostatic in nature and include electron-electron, electron-nuclei and nuclei-nuclei interactions.⁹³ However, within classical simulations, the electronic degrees of freedom are averaged out and interactions are thus described in a phenomenological way using effective potentials U . These potentials provide an atomistic force field $\vec{F} = -\vec{\nabla}U$ that enters Equation 3.1 and enables computation of atomic trajectories. Interactions occurring between atoms that share a chemical bond (bonded interactions) and those that do not (non-bonded interactions) are considered separately. The latter include the electrostatic and van der Waals interactions and take the form of pairwise additive (two-body) potentials that only depend on the separation r of the atoms. This is an approximation that neglects many-body effects such as electronic polarization. For many cases, this approximation is satisfactory. However, there is considerable effort to improve existing atomistic force fields by explicitly including these effects.¹⁰³ Bonded interactions include effective two-, three- and four-body terms that describe bond stretching, angle bending and bond torsion, respectively. The potentials for each of these interactions are described below.

3.2.1 Electrostatic Interaction

Every atom possesses a partial charge q which corresponds to its local electron density. The electrostatic interaction⁷⁰ between two charges q_i and q_j separated by a distance r_{ij} is given by the Coulomb potential:

$$U_{\text{ele}} = \frac{1}{4\pi\epsilon} \frac{q_i q_j}{r_{ij}} \quad (3.6)$$

Here, ϵ is the permittivity of the medium surrounding the charges. For simulations with explicit solvent (i.e. water), $\epsilon = \epsilon_0$, the permittivity of free space.

3.2.2 van der Waals Interaction

All non-bonded atom pairs experience a van der Waals interaction¹⁰⁴ which includes a short-range repulsion due to the overlap of atomic electron clouds (Pauli exclusion) and a longer-range attraction due to induced dipole–induced dipole interactions (London dispersion forces). The van der Waals interaction between two identical atom types separated by a distance r is described by the Lennard-Jones potential:¹⁰⁴

$$U_{\text{vdw}} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (3.7)$$

U_{vdw} is plotted in Figure 3.2. Here, ϵ is the strength of the attraction and σ is the effective atom diameter. For interactions between two different atom types i and j , these parameters are given by¹⁰⁵

$$\begin{aligned} \sigma &= \frac{1}{2}(\sigma_i + \sigma_j) \\ \epsilon &= \sqrt{\epsilon_i \epsilon_j} \end{aligned} \quad (3.8)$$

With this model, two interacting atoms will have a minimum potential energy of $-\epsilon$ when separated by $2^{1/6} \sigma \approx \sigma$. Typically, $\sigma \sim 0.3$ nm and thus, the van der Waals interaction is extremely weak for separations greater than ~ 1 nm.

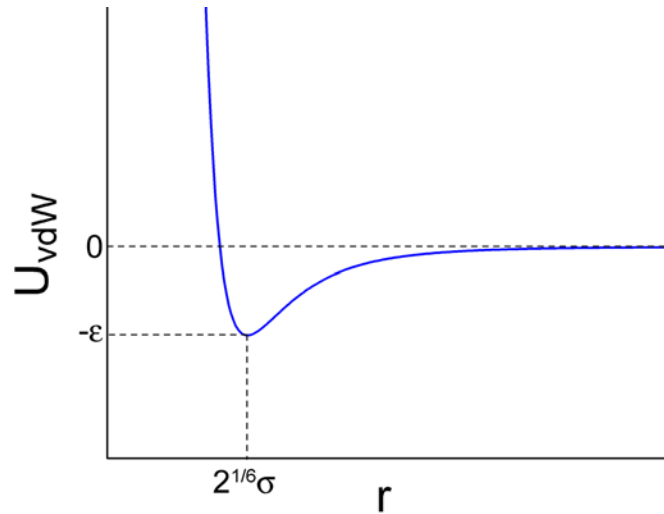


Figure 3.2: The van der Waals interaction modeled by the Lennard-Jones potential.

3.2.3 Bond Stretching

Atoms that share a chemical bond undergo small amplitude vibrations about an equilibrium bond length r_0 (Figure 3.3). Thus, bond stretching can be accurately modeled with a harmonic potential:

$$U_{\text{bond}} = \frac{1}{2}k_r(r - r_0)^2 \quad (3.9)$$

Here, k_r and r are the force constant that describes the stiffness of the stretching and the instantaneous bond length, respectively. Because this potential contains only a single minimum at r_0 , this potential can not be used to describe bond breaking/forming and molecules will maintain their initial connectivity throughout the simulation.

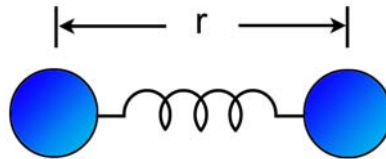


Figure 3.3: Stretching of a chemical bond about an equilibrium bond length r_0 .

3.2.4 Angle Bending

The angles between two chemical bonds in a molecule undergo low amplitude bending about an equilibrium angle θ_0 (Figure 3.4) and are also modeled with a harmonic potential:

$$U_{\text{angle}} = \frac{1}{2}k_{\theta}(\theta - \theta_0)^2 \quad (3.10)$$

Here, k_{θ} and θ are the force constant and instantaneous angle, respectively.

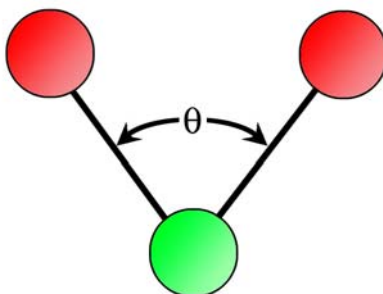


Figure 3.4: Bending of an angle in a molecule about its equilibrium value θ_0 .

3.2.5 Bond Torsion

Four consecutive (bonded) atoms in a molecule form planes ijk and jkl (Figure 3.5). The dihedral angle ϕ is defined as the angle between these planes. In a molecule, atoms i , j , k and l have a set of preferred (minimum energy) dihedral angles separated by energy barriers. The potential that describes these barriers is given by:

$$U_{\text{torsion}} = \frac{k_{\phi}}{2}(1 + \cos(n\phi - \phi_0)) \quad (3.11)$$

Here, k_{ϕ} , n are the barrier height, number of energy minima encountered as the angle is rotated from 0° to 360° and a dihedral angle. ϕ_0 is 0 or π depending on whether the *trans* or *cis* conformation of the four atoms is more favorable.

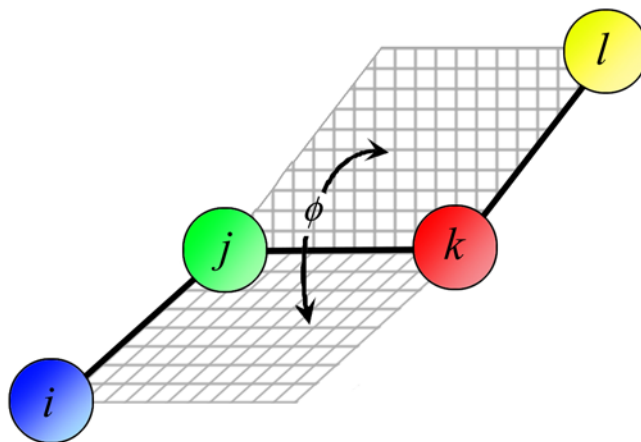


Figure 3.5: Dihedral angle ϕ for atoms i, j, k and l .

3.2.6 Parameterization

Each of the potentials above contains a number of parameters that are specific to the particular atoms involved in the interaction. These parameters must be carefully chosen in order to build a model system that accurately reproduces experimental results. Two of the most popular parameterizations for biological systems are the AMBER¹⁰⁵ and CHARMM¹⁰⁶ force fields. Developers of both these force fields have followed a parameterization strategy that involves fitting the parameters to experimental data and *ab initio* calculations. Additionally, both of these force fields have been parameterized for the TIP3P water model.¹⁰⁷ For more details on how these force fields were developed, please see the references.^{105, 106} Classical force fields have been widely used for biological simulations and have achieved a great deal of success.^{99, 108, 109} However, there are also limitations and care must be taken to ensure reliability of the simulation results.^{103, 109}

3.3 Thermodynamic Integration

The natural tendency of a thermodynamic system is to achieve a minimum of free energy.¹¹⁰ Thus, much information about the nature of a system can be gained by measuring the free energy change ΔF that occurs as the system undergoes a reversible transformation from an initial state to a final state. ΔF is the work exchanged with the system's surroundings during such a transformation at constant volume and is defined as

$$\Delta F = \Delta E - T\Delta S \quad (3.1)$$

Here, ΔE is the change in total energy, T is the temperature and ΔS is the change in entropy. Favorable (spontaneous) transformations release energy to the environment and are characterized by $\Delta F < 0$. Transformations with $\Delta F > 0$ consume energy from the environment and are thus unfavorable (nonspontaneous).

Within statistical mechanics, the free energy F is related to the partition function Z by

$$F = -k_{\text{B}}T \ln Z \quad (3.2)$$

The partition function is obtained by calculating a multidimensional integral over the entire phase space accessible to the system:

$$Z = \frac{1}{N!h^{3N}} \int e^{-\beta H} d^{3N} p d^{3N} q \quad (3.3)$$

Z is not a canonical average and thus, can not be directly measured in an MD simulation. However, there exist a number of methods that enable computation of ΔF between two well-defined states with MD.

Thermodynamic integration (TI) a technique used within the alchemical free energy framework for calculating the free energy difference ΔF between two states A and B .^{95,}

^{111, 112} In TI, a continuous set of unphysical states intermediate between A and B are

generated by introducing a coupling parameter λ into the Hamiltonian $H = H(\lambda)$. Here, $\lambda = 0$ and $\lambda = 1$ refer to state A and B , respectively: $H(0) = H_A$ and $H(1) = H_B$. Unphysical intermediates occur for λ values between 0 and 1. Substituting $H(\lambda)$ into Equation 3.3 and differentiating with respect to λ gives

$$\frac{dF}{d\lambda} = \frac{\int \frac{\partial H}{\partial \lambda} e^{-\beta H(\lambda)} d^{3N} p d^{3N} q}{\int e^{-\beta H(\lambda)} d^{3N} p d^{3N} q} = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle \quad (3.4)$$

The free energy difference of systems A and B can then be obtained through integration:

$$F_B - F_A = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle d\lambda \quad (3.5)$$

TI is advantageous because 1) this method enables accurate computation of large free energy differences and 2) the quantity $\langle \partial H / \partial \lambda \rangle$ can be computed with MD. In practice, $\langle \partial H / \partial \lambda \rangle$ is obtained at a discrete set of λ values and the integral is evaluated numerically. With the formalism presented here, exact computation of ΔF can be achieved. However, computed in this way ΔF is a statistical quantity and is subject to errors. When carrying out TI, one must employ adequate MD simulation timescales to ensure proper convergence of $\langle \partial H / \partial \lambda \rangle$ at each λ value. This convergence can be assessed using one of several error analysis methods.¹¹²⁻¹¹⁴

One application of this method that is of particular importance in nanoscience and biochemistry is for the computation of free energy differences between the bound and unbound states of two interacting molecular species. The binding free energy ΔF_{bind} is defined as

$$\Delta F_{\text{bind}} = F_{\text{bound}} - F_{\text{unbound}} \quad (3.6)$$

ΔF_{bind} is a measure of the attraction between two interacting species and their preference for binding. Computing ΔF_{bind} for a ligand binding to a receptor protein, for example, is carried out with TI by computing free energy differences along an appropriate thermodynamic cycle (Figure 3.6). With this cycle ΔF_{bind} is determined by calculating $\Delta F_{\text{ann}}^{\text{unb}}$ and $\Delta F_{\text{cre}}^{\text{b}}$, the free energy change for annihilating (ann) and creating (cre) the ligand in the unbound (unb) and bound (b) states, respectively. Because Figure 3.6 makes up a closed cycle, the sum of free energy changes over the cycle is zero. Thus, ΔF_{bind} is then given by

$$\Delta F_{\text{bind}} = \Delta F_{\text{ann}}^{\text{unb}} + \Delta F_{\text{cre}}^{\text{b}} \quad (3.7)$$

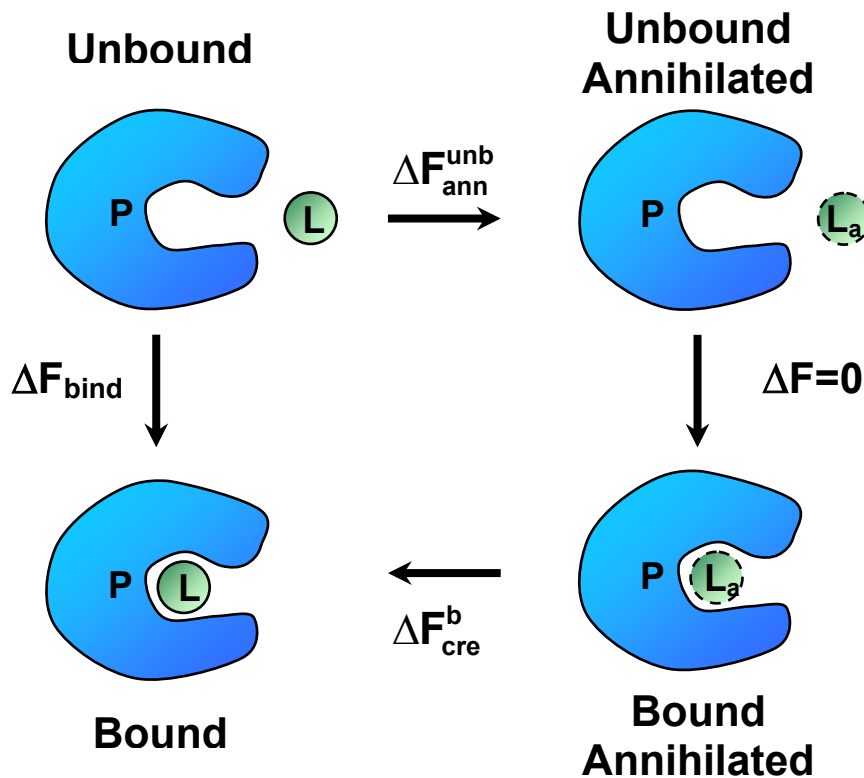


Figure 3.6: Thermodynamic cycle used to compute the binding free energy ΔF_{bind} of a ligand L to a receptor protein P. L_a is an annihilated (non-interacting) ligand that experiences no non-bonded interaction with itself or the environment.

To annihilate (or create) the ligand, λ is introduced into its atomic charges $q(\lambda)$ and van der Waals (vdW) interaction potential $U_{\text{vdW}}(\lambda)$. Varying λ from 0 to 1 (or 1 to 0), effectively turns these interactions off (on). That is,

$$q(\lambda) = (1 - \lambda)q$$

$$U_{\text{vdW}}(\lambda) = 4\epsilon(1 - \lambda) \left[\frac{\sigma^{12}}{(\alpha\sigma^6\lambda^p + r^6)^2} - \frac{\sigma^6}{(\alpha\sigma^6\lambda^p + r^6)} \right] \quad (3.8)$$

The singularity in the r^{-12} portion of the Lennard-Jones potential results in slow convergence of $\langle \partial H / \partial \lambda \rangle$ when perturbing vdW interactions. This problem can be alleviated by removing the singularity with the “soft-core” form of $U_{\text{vdW}}(\lambda)$ given in Equation 3.8.^{112, 115, 116} A plot of $U_{\text{vdW}}(\lambda)$ for recommended parameter values of $\alpha = 0.5$ and $p = 1$ is shown in Figure 3.7.¹¹⁶ For more details on how λ can be introduced into atomic interaction potentials, see the GROMACS Manual.¹⁰¹

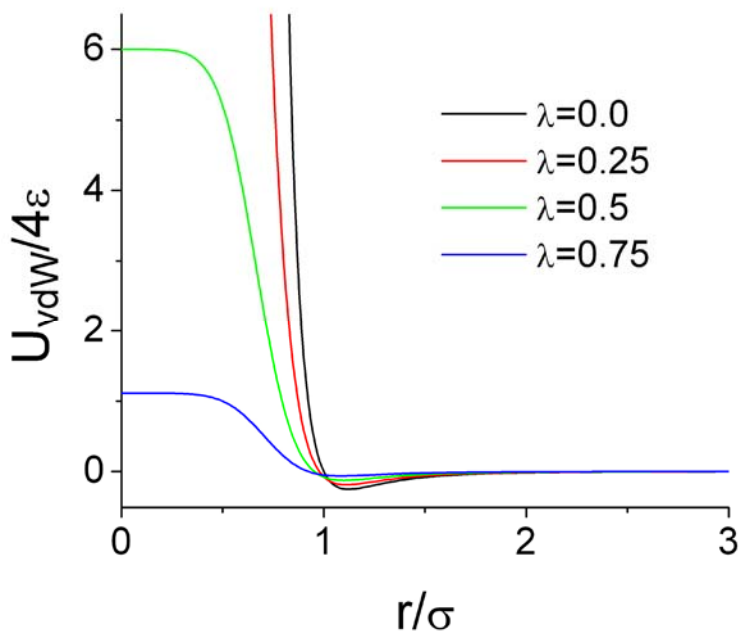


Figure 3.7: Soft-core potential $U_{\text{vdW}}(\lambda)$ for van der Waals interactions for various λ values with $\alpha = 0.5$ and $p = 1$.

Convergence of $\langle \partial H / \partial \lambda \rangle$ during annihilation or creation of a ligand is also improved by perturbing charges first, followed by vdW interactions.^{112, 116} Thus, $\Delta F_{\text{ann/cre}} = \Delta F_{\text{ele}} + \Delta F_{\text{vdw}}$ where ΔF_{ele} and ΔF_{vdw} are the free energy changes of perturbing charges and vdW interactions, respectively. Typically, when perturbing charges, $\langle \partial H / \partial \lambda \rangle$ is a smoothly varying function and only a few (3-5) λ values are required to achieve an accurate value of ΔF_{ele} . However, as vdW interactions are perturbed, solvent molecules can penetrate ligand atoms which results in a non-trivial $\langle \partial H / \partial \lambda \rangle$ curve. Thus, perturbing vdW interactions over many (10-20) λ values is necessary to accurately calculate ΔF_{vdw} .

In the bound state, the ligand is held in place by a combination of electrostatic and vdW interactions with the receptor. Perturbing these interactions (bottom branch of Figure 3.1) enables the ligand to escape the binding pocket and explore the entire simulation box. This causes $\langle \partial H / \partial \lambda \rangle$ to converge very slowly during this transformation. To alleviate this problem, restraints are applied to the bound ligand to reduce the space that it can sample during this step. This drastically improves convergence of $\Delta F_{\text{cre}}^{\text{b}}$. However, the free energy contribution of these restraints ΔF_{res} must then be accounted for to ensure that the free energy difference between the bound annihilated and unbound annihilated systems is zero (right branch of Figure 3.1). ΔF_{res} will depend on the nature of the restraints. A detailed overview of how to choose and calculate ΔF_{res} is given elsewhere.¹¹⁷⁻¹²⁰ For other practical information regarding alchemical free energy these calculations, see www.alchemistry.org.

3.4 Replica Exchange Molecular Dynamics

Flexible polymers such as single-stranded DNA have many degrees of freedom and a rugged potential energy landscape. Because of this, a phenomenon known as kinetic trapping can occur where polymers become trapped in metastable configurations associated with local energy minima that persist over typical MD timescales (10–100 ns). Thus, complete sampling of configuration space and determination of accurate equilibrium properties of polymeric systems with conventional MD is usually not possible. Replica exchange molecular dynamics (REMD) overcomes this limitation by performing multiple MD simulations (replicas) in parallel at multiple temperatures. REMD has enabled explorations of the folding pathways of proteins^{121, 122} and oligonucleotides¹²³ and produced structures in agreement with experiment.

There are several detailed accounts of REMD applied to biomolecular systems.^{122, 124,}
¹²⁵ In short, REMD constructs M independent copies (replicas) of a system, each at a different temperature T_m ($m = 0, \dots, M-1$). The trajectory of each replica is then computed concurrently using standard MD. After a fixed number of MD steps, adjacent replicas (replicas i and $i+1$) attempt to exchange temperatures. The probability of a successful exchange is given by the distribution

$$P_{\text{exchange}} = \begin{cases} 1 & , \text{ for } U_{i+1} < U_i \\ \exp[(\beta_i - \beta_{i+1})(U_i - U_{i+1})] & , \text{ for } U_{i+1} > U_i \end{cases} \quad (3.1)$$

Here, β_i and U_i are the inverse temperature and instantaneous total potential energy of replica i , respectively. The time between consecutive exchange attempts τ_{exchange} must be larger than the system's thermal relaxation time τ_{relax} in order to ensure that all replicas are properly thermalized. High temperature replicas enable the system to overcome

energy barriers and explore configurations that would not be sampled at low temperature due to kinetic trapping. The temperature exchanges and form of P_{exchange} effectively filter energetically favorable configurations for each temperature according to the Boltzmann distribution. REMD drastically enhances sampling of configuration space and provides a means for obtaining accurate equilibrium properties of systems described by rugged energy landscapes. Figure 3.8 shows the instantaneous temperature of one replica over the course of an REMD simulation.

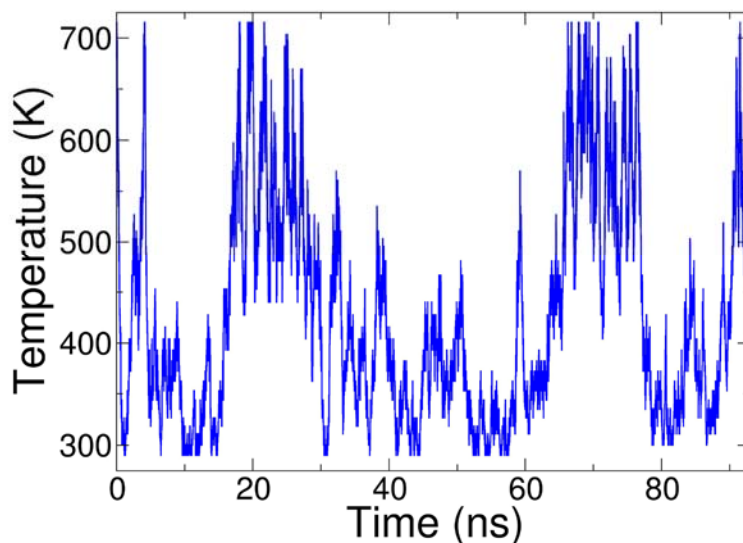


Figure 3.8: Instantaneous temperature of one replica during an REMD simulation with temperatures ranging from 290 K to 715 K. The replica visits many temperatures over the course of the simulation.

Determining an adequate temperature tiling $\{T\} = T_0, T_1, \dots, T_{M-1}$ of the replicas is crucial for a successful REMD simulation. To take advantage of the enhanced sampling potential of REMD, a subset of $\{T\}$ must include high enough temperatures at which kinetic trapping effects are minimized. Additionally, $\{T\}$ is typically chosen such that $P_{\text{exchange}} \approx 20 - 30\%$ for all replicas, thus allowing the system to perform a random walk in

temperature space. The temperature spacing $\Delta T_i = T_{i+1} - T_i$ between adjacent replica i and $i+1$ should not be constant. Because the magnitude of thermal fluctuations grows with increasing temperature, ΔT_i should be larger at high temperature in order to maintain a constant P_{exchange} . Figure 3.9 shows the energy distributions in an REMD simulation involving 64 replica. Because of the appropriately chosen $\{T\}$, the distributions contain significant overlap and enable a constant P_{exchange} .

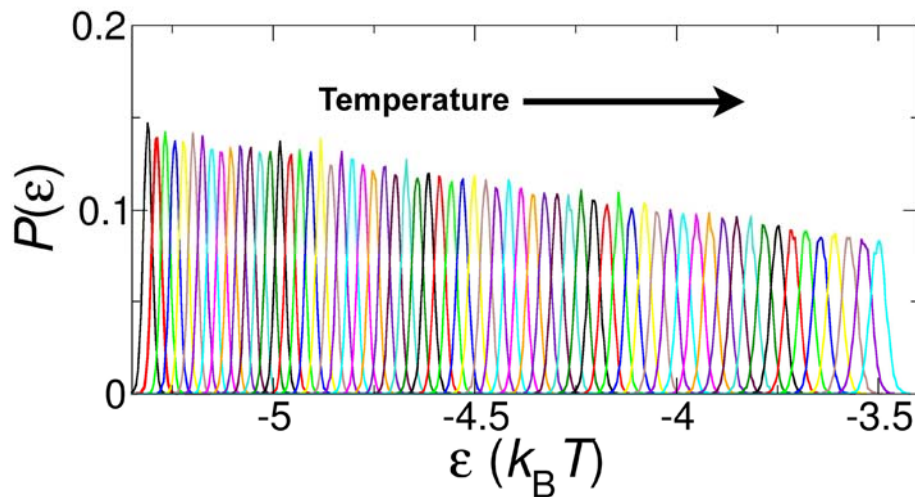


Figure 3.9: Energy distribution of each replica in an REMD simulation. The distributions are broadened and shifted to higher energies with increasing temperature. This is indicative of the larger thermal fluctuations the system experiences at high temperature. $\{T\}$ was chosen such that the temperature of the m -th replica is given by $T_m = 125.7\text{K} + 164.3\text{K} e^{m/49.3}$ with $m = 0, 1, \dots, 63$.

3.5 Simulation Protocol

DNA structures were built with the Amber 7.0 Nucleon tool, which enables the user to generate 3D coordinates of nucleic acids in various helical forms.¹⁰⁰ CNT structures were generated using Nanotube Builder 1.0, a home-written C program we developed that constructs CNT coordinates of arbitrary length and chirality. This code has been implemented as a TCL plugin in version 1.8.7 of the molecular visualization program Visual Molecular Dynamics (VMD) (www.ks.uiuc.edu/Research/vmd).⁴³ Systems were set up within the Amber 7.0¹⁰⁰ xLeap toolkit and VMD.⁴³ These programs provide a graphical interface where the user can generate coordinates and topologies for molecular systems. All simulations were performed with the GROMACS¹²⁶ MD package. This package was advantageous because it is a flexible, efficient, parallelized code that supports a multitude of MD methodologies and contains numerous tools for setting up and analyzing simulations. Unless stated otherwise, simulations were run across 4–16 CPUs on a local computing cluster. Energy minimization using the steepest descent algorithm with a step size of 0.1 nm was employed to optimize the initial coordinates of each system prior to the MD simulation. Trajectories were computed using the leap-frog algorithm⁹⁵ with a 1.5 fs time step in the *NPT*-ensemble under ambient conditions (1 atm, 300 K unless stated otherwise) using the Parrinello-Rahman barostat¹²⁷ and Berendsen thermostat.¹²⁸ Time constants of 0.1 ps were used for both these temperature and pressure coupling schemes. Electrostatic interactions were calculated with the particle mesh Ewald method¹⁰² using a 0.12 nm grid spacing and fourth order interpolation. 3D periodic boundary conditions were applied in all cases. A cutoff of 1.2 nm was used for van der

Waals interactions. All bonds containing hydrogen atoms were constrained with the LINCS algorithm.¹²⁹

All simulations reported here included biopolymers (ssDNA or proteins) interacting with an infinite CNT in aqueous solution. To simulate an infinite CNT, a segment with a length commensurate with the L_z box dimension was aligned along the z-axis with the terminal carbon atoms sharing a chemical bond. Based on previous work,¹³⁰⁻¹³² CNT atoms were modeled as uncharged Lennard-Jones particles using sp^2 carbon parameters from the AMBER99 force field.¹³³ For simplicity, the positions of all CNT atoms were constrained with a harmonic potential with a force constant of 240 kcal/mol·nm². The AMBER99 force field was also used to model biopolymers. For simulations involving ssDNA, Na⁺ counterions were included to exactly neutralize the negatively charged sugar-phosphate backbone unless stated otherwise. For simulations containing proteins, Na⁺ and Cl⁻ counterions were included to reproduce a physiological salt concentration of 100 mM in addition to those necessary to neutralize the protein. All simulations were performed with explicit solvent using the TIP3P water model.¹⁰⁷ At all times the interior of CNT remained hollow and devoid of water molecules, which is reasonable for pristine, unoxidized CNTs in aqueous solution. Analysis and visualization of MD trajectories was performed with home-written TCL scripts executed within VMD.⁴³

Chapter 4

Structure of DNA-Carbon Nanotube Hybrids

The results presented in this chapter have appeared in the following publications:

- R. R. Johnson, A. Kohlmeyer, A.T. C. Johnson, M. L. Klein. *Nano Letters*, **9**, 537-541 (2009).
- R. R. Johnson, A.T. C. Johnson, M. L. Klein. *Nano Letters*, **8**, 69-75 (2008).
- A.T.C. Johnson, C. Staii, M. Chen, S. Khamis, R. R. Johnson, M. L. Klein, A. Gelperin. *Semiconductor Science and Technology*, **21**, S17-S21 (2006).

It turns out that ssDNA and CNT, two macromolecules that have never encountered each other in nature, are chemically compatible and can be easily combined to form DNA-carbon nanotube hybrids (DNA-CNT). These unique nanomaterials possess a remarkable set of technologically useful properties with potential applications in CNT solubilization¹⁷ and separation,⁶⁸ chemical²¹ and biological²⁴ sensing and ultrafast DNA sequencing.⁸¹ The importance of this composite nanomaterial motivates a quantitative understanding of its interactions, structure, and physical properties. Experimental methods for studying DNA-CNT have included atomic force microscopy^{68, 72, 134}, scanning tunneling microscopy,¹³⁵ circular dichroism⁸³ and optical absorption.⁸⁴ While these techniques have provided useful insights into the physical properties of DNA-CNT, they leave a lot to be desired. Such techniques have resolutions limited to several nanometers or probe the properties of bulk solutions of DNA-CNT. Because of the small size of this nanomaterial, structural information at atomic-scale resolution is needed to

fully understand and realize the potential of DNA-CNT. At the moment, a truly atomistic picture of this hybrid material is lacking.

MD simulations are well suited for gaining detailed information of biomolecular complexes, such as DNA-CNT, with atomic resolution. Thus, to expand our knowledge of this hybrid nanostructure, we conducted a series of MD simulations to study the self-assembly, dynamics, energetics and structure of DNA-CNT. The work presented here has developed a deeper understanding of this organic/inorganic hybrid nanostructure and will facilitate numerous proposed applications of DNA-CNT.^{17, 21, 24, 68}

4.1 DNA-Carbon Nanotube Hybrid Self-Assembly

Zheng *et al.* demonstrated that ssDNA and CNT will spontaneously self-assemble into water soluble hybrids in aqueous solution.¹⁷ Several prior computational studies have attempted to shed light on the nature of this self-assembly process. Molecular mechanics calculations have been employed to locate low energy conformations of ssDNA adsorbed onto CNT.^{17, 81} These calculations showed that energetically favorable configurations result when ssDNA wraps helically¹⁷ about CNT with its bases stacked to the CNT surface.^{17, 81} However, such calculations were based on energy minimization principles and did not provide any information about the dynamics and conformational changes that occur during DNA-CNT self-assembly. Additionally, the prior studies neglected the effects of the solvent and temperature. Thus, in order to obtain a microscopic dynamical understanding of DNA-CNT self-assembly, we performed a series of simulations of ssDNA adsorbing to a CNT in aqueous solution under ambient conditions.

4.1.1 Simulation Details

Our system consisted of a random 14-base (ATCGATACGTGACT) oligonucleotide initially separated from a (11,0) CNT by about 1.5 nm (Figure 4.1a). ssDNA was initialized in a helical stacked conformation (obtained from one strand of B form dsDNA⁶⁵), which is a reasonable structure for short oligonucleotides in aqueous solution.⁶⁵ A simulation box size of $5 \times 5 \times 8.083 \text{ nm}^3$ was used. The system was then allowed to equilibrate over the course of MD at a constant temperature of 330 K for 21 ns. The slightly elevated temperature was employed in this particular simulation to effectively accelerate the adsorption kinetics so that the entire process would occur within typical timescales accessible with MD ($\sim 10 \text{ ns}$). By examining the results of approximately twenty additional simulations conducted at 300 K and involving five different oligonucleotide sequences, we have verified that this temperature does not affect the mechanics involved in DNA-CNT self-assembly. Thus, the results of this simulation are relevant at room temperature as well.

4.1.2 Results and Discussion

Within the first 500 ps, several ssDNA segments make contact to CNT. These segments undergo a conformational change where bases rotate by 90° relative to the sugar-phosphate backbone, thus becoming unstacked from their neighbors. This enables individual bases to adsorb (stack) to the CNT surface at a distance similar to that found for neighboring planes in graphite ($\sim 0.34 \text{ nm}$). These bases are held tightly against the CNT sidewall and anchor ssDNA in the radial direction. However, the oligonucleotide freely diffuses along the CNT axial and circumferential directions. Within 5.5 ns, the

entire ssDNA backbone is drawn close to CNT which permits additional bases to bind to the sidewall (Figure 4.1b and Figure 4.2). Over the next 16 ns, many of the remaining, unbound bases adsorb and, even more remarkably, ssDNA spontaneously wraps around CNT into a left-handed helix (Figure 4.1c).

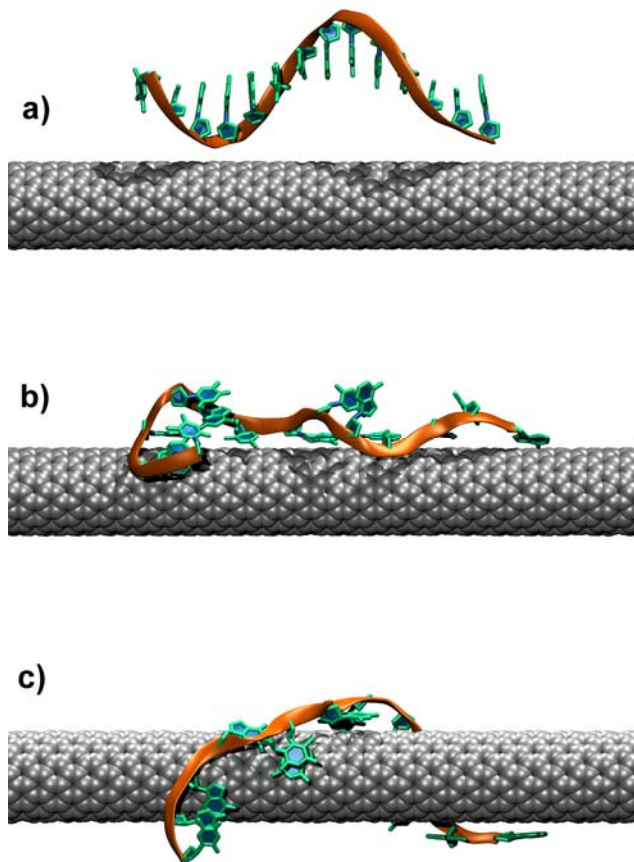


Figure 4.1: Self-assembly of DNA-CNT in aqueous solution. Water molecules and Na^+ counterions have been removed for visualization purposes. (a) Initial configuration. (b) Configuration after 5.5 ns. (c) Final configuration after 21 ns.

The self-assembly of DNA-CNT is driven by strong, attractive interactions between the faces of the bases and the CNT sidewall (Figure 4.2). These interactions, which result in bases stacking to the CNT surface, occur for all four DNA bases (A, C, G and T) on CNT of arbitrary diameter and chirality. These results agree with computations^{17, 81, 82, 136-}

¹³⁸ and experiments.⁸⁴ The nature of base-CNT binding (as discussed in more detail in Section 4.2) is due to π - π stacking:¹³⁹ an attractive interaction due to van der Waals forces between aromatic moieties. Because all DNA bases can bind to CNT of arbitrary diameter and chirality, ssDNA of any sequence should readily form DNA-CNTs. We have verified this by obtaining similar results with another random sequence as well as homopolymers (each 21 nucleotides in length) poly A, poly C, poly G and poly T adsorbing to several other CNTs of differing diameter and chirality.

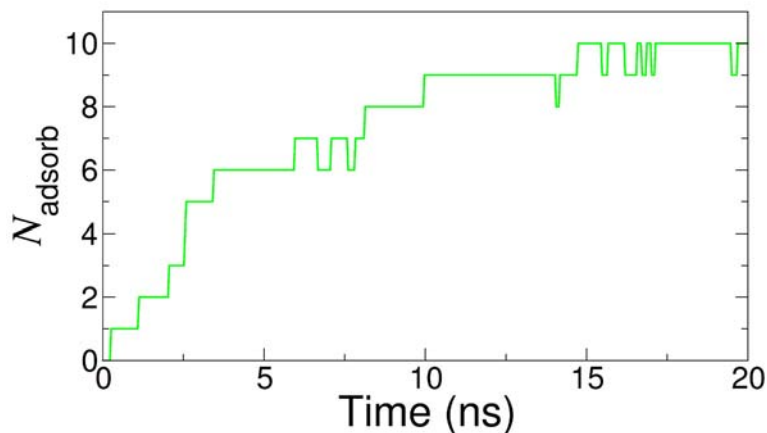


Figure 4.2: Number of adsorbed bases N_{adsorb} during DNA-CNT self-assembly. Self-assembly proceeds via formation of high affinity contacts between bases and CNT.

Perhaps the most striking result in this simulation is that ssDNA spontaneously wraps into a compact left-handed helix about CNT circumference. In other trials with different initial configurations of the system, we observed formation of right-handed helices, loops and disordered, kinked structures. These kinked structures occur quite frequently in such DNA-CNT self-assembly simulations (Figure 4.3). However, they represent high energy (metastable) ssDNA configurations and can be annealed to more favorable conformations (Figure 4.3). The sensitivity of the final ssDNA conformations to the initial conditions and the persistence of metastable kinked structures indicates that conventional MD is

insufficient to fully sample DNA-CNT configuration space and achieve accurate equilibrium information over typical simulation timescales (10-100 ns). ssDNA is a flexible polymer with many degrees of freedom and a rugged potential energy landscape containing many local minima. Because of this, such systems can be difficult to sample at low temperature ($T \sim 300$ K) with conventional MD. These sampling limitations in MD are widely appreciated and there exist a variety of techniques such as replica exchange molecular dynamics (Section 3.4) that can overcome this.¹²⁴ We have applied this method to study the entire ensemble of low energy ssDNA conformations about CNT. A discussion of these results is presented in Section 4.4.

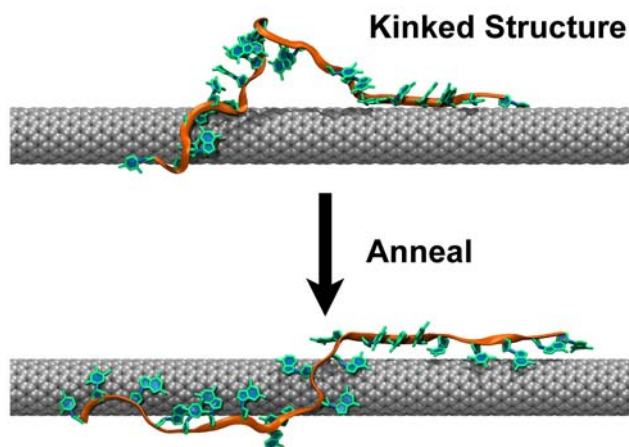


Figure 4.3: (Top) Metastable kinked ssDNA structure which is the result of kinetic trapping. (Bottom) The kink can be annealed into a more energetically favorable configuration.

4.2 The Nature of DNA-Carbon Nanotube Interactions

Our previous computations have shown that DNA-CNT spontaneously self-assembles in aqueous solution via high affinity contacts between DNA bases and CNT sidewall. These contacts consist of bases binding and lying flat (stacking) on CNT surface. The importance of base-CNT stacking on the stability of DNA-CNT has motivated several computational studies of the nature of such interactions. Both MD and *ab initio* methods have been employed to estimate the base-CNT interaction energy in vacuum for the DNA bases adenine (A), cytosine (C), guanine (G) and thymine (T).^{25, 81, 82, 137} Except for one,⁸¹ these studies predict substantial energy values of ~ 10 kcal/mol ($\sim 17 k_B T$) that follow the trend $G > A > T > C$.^{25, 82, 137} While the interaction energy is useful for gaining information about base-CNT adhesion strength, it does not provide a complete understanding of all the relevant effects involved in DNA-CNT. Because DNA-CNT is usually assembled and employed in aqueous solution at finite temperature, it is expected that base-CNT interactions includes a complex interplay of enthalpic, entropic, and solvent mediated effects. Simply calculating the base-CNT interaction energy does not provide any insight into the importance of these effects. The relevant quantity that captures all of these effects is the base-CNT binding free energy ΔF_{bind}^F in aqueous solution. Here, we have employed alchemical free energy methods and thermodynamic integration (TI) to compute ΔF_{bind}^F for all four bases. We find that the binding free energies follow the trend: $\Delta F_{\text{bind}}^G < \Delta F_{\text{bind}}^A < \Delta F_{\text{bind}}^T < \Delta F_{\text{bind}}^C$. The binding is largely governed by base-CNT van der Waals (vdW) forces with solvation and entropic effects playing a relatively minor role.

4.2.1 Simulation Details

TI is employed to compute the base-CNT binding free energy ΔF_{bind} for all four bases in aqueous solution. ΔF_{bind} is defined as the free energy difference between the bound and unbound states of the base-CNT system. This was carried out by calculating free energy differences along the thermodynamic cycle shown in Figure 4.4. ΔF_{bind} is obtained by computing the free energy required to annihilate (ann) and create (cre) the DNA base in the unbound $\Delta F_{\text{ann}}^{\text{unb}}$ and bound $\Delta F_{\text{cre}}^{\text{b}}$ state, respectively and is given by

$$\Delta F_{\text{bind}} = \Delta F_{\text{ann}}^{\text{unb}} - \Delta F_{\text{ann}}^{\text{b}} \quad (4.1)$$

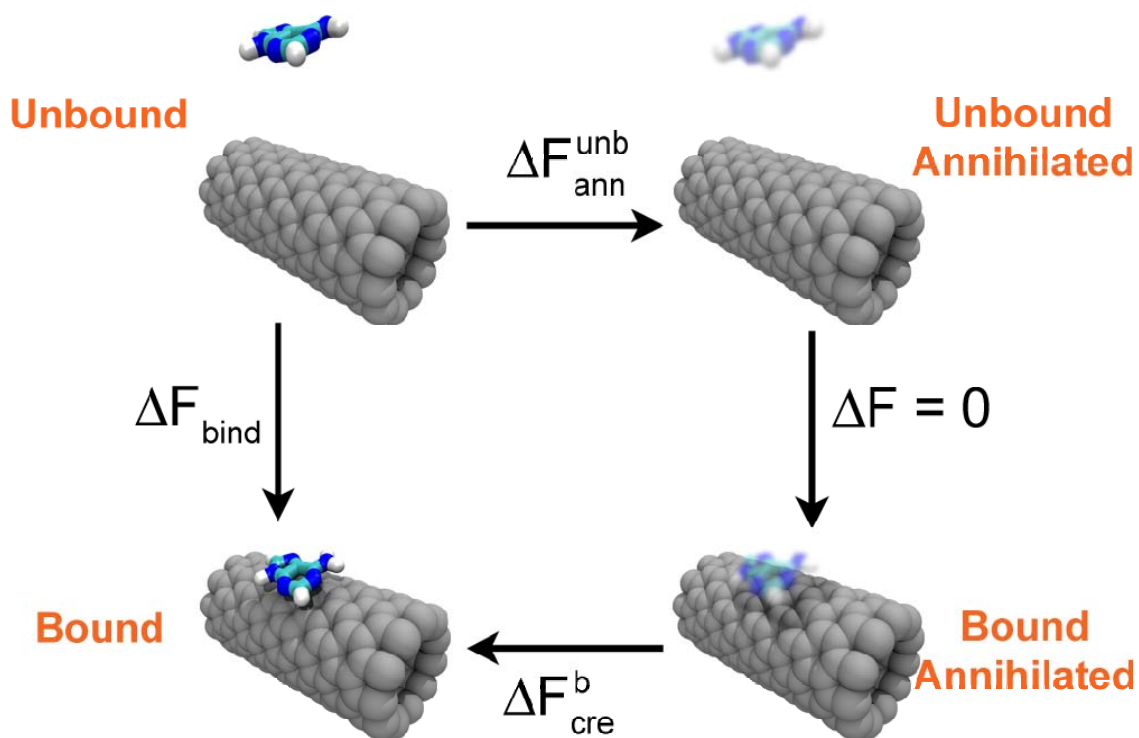


Figure 4.4: Schematic of the thermodynamic cycle employed to calculate the base-CNT binding free energy. The binding free energy is defined as $\Delta F_{\text{bind}} \equiv F_{\text{bound}} - F_{\text{unbound}}$. $\Delta F_{\text{cre}}^{\text{b}}$ and $\Delta F_{\text{ann}}^{\text{unb}}$ are the free energy required to create and annihilate base interactions in the bound and unbound states, respectively. With this scheme $\Delta F_{\text{bind}} = \Delta F_{\text{ann}}^{\text{unb}} + \Delta F_{\text{cre}}^{\text{b}}$.

The annihilated system consists of a non-interacting DNA base, i.e. the base has atomic charges and vdW interaction potential set to zero and experiences no interaction with itself, water or CNT. Bound systems consisted of a single base adsorbed to an infinite (11,0) CNT in aqueous solution in a $3.4 \times 3.4 \times 2.127 \text{ nm}^3$ box. Unbound systems consisted of a single base in aqueous solution alone in a $3.0 \times 3.0 \times 3.0 \text{ nm}^3$ box. Previous *ab initio* calculations have shown that DNA base-CNT interactions are dominated by vdW forces; polarization and charge transfer (chemisorption) effects are negligible.^{82, 137, 138, 140} Thus, the classical atomistic models employed here included the relevant physics of DNA base-CNT interactions and should enable a reasonable estimate for ΔF_{bind} . The π - π stacking interaction plays a critical role in base-CNT binding (see below). In the AMBER99 force field, stacking interactions among aromatic species are parameterized within the vdW parameters of each atom type. Specific electrostatic interactions among π electrons are thus included in an average way. A discussion on modeling π - π stacking interactions is reported elsewhere.¹³⁹ Technically, the AMBER99 force field contains charge and bonded interaction parameters for nucleotides. That is, bases connected to a sugar and phosphate group. To model individual DNA bases, we simply removed these groups from the model and capped atom N₉ and N₁ with hydrogen for purines and pyrimidines, respectively. The charge on the hydrogen was then adjusted to maintain electrical neutrality of the single base. Prior to the free energy calculation, each system was equilibrated at 300 K¹²⁸ and 1 atm¹²⁷ to ensure a proper solvent density.

Computing binding free energies involved calculating the free energy required to annihilate and create a base in the unbound $\Delta F_{\text{ann}}^{\text{unb}}$ and bound $\Delta F_{\text{cre}}^{\text{b}}$ states, respectively.

Following previous work,^{116, 117} these free energy differences were computed using a two step process:

- 1) Electrostatic interactions were first turned off/on by annihilating/creating atomic charges using λ values of 0.0, 0.25, 0.50, 0.75, and 1.0.
- 2) vdW interactions were then annihilated/created using λ values of 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. Soft-core Lennard-Jones potentials were employed with soft-core parameter $\alpha = 0.5$ and λ power $p = 1$.^{114, 126}

In addition to eliminating sampling problems at $\lambda = 0$ and $\lambda = 1$, this method also enabled the estimation of the electrostatic ΔF_{ele} and vdW ΔF_{vdw} contributions to ΔF_{bind} . At each λ value, $\langle \partial H / \partial \lambda \rangle$ was extracted from the last 14 ns of a 15 ns trajectory. Error in each $\langle \partial H / \partial \lambda \rangle$ value was estimated using the block averaging method.¹¹³ Each trajectory evolved at constant temperature¹²⁸ (300 K) and volume. To eliminate sampling problems associated with annihilating vdW interactions, harmonic position restraints with force constant of 240 kcal/mol·nm² were applied to DNA bases in the bound state. The entropic ΔS_{bind} and energetic ΔE_{bind} contributions to ΔF_{bind} were obtained by repeating these free energy calculations at 320 K and 280 K. ΔS_{bind} and ΔE_{bind} , formally derivatives of ΔF_{bind} , were calculated as finite differences:¹¹¹

$$\Delta S(T) = -\frac{\Delta F(T + \Delta T) - \Delta F(T - \Delta T)}{2\Delta T} \quad (4.2)$$

$$\Delta E = \Delta F + T\Delta S$$

ΔF_{bind} depends on the concentrations of both the bases and CNTs in aqueous solution. The concentration dependence enters in the free energy contribution of the position restraints. Typically, one is interested in obtaining the free energy under standard state

conditions. That is, at a temperature, pressure and concentration of 300 K, 1 atm and 1 M respectively. To obtain the standard state value of ΔF_{bind} , we compute the free energy contribution of the position restraints ΔF_{res} as

$$\Delta F_{\text{res}} = k_B T \ln \left(\frac{V_{\text{std}}}{V_{\text{bound}}} \right) \quad (4.3)$$

Here, $V_{\text{std}} = 1660 \text{ \AA}^3$ is the molar standard state volume occupied by a single base in the unbound state. $V_{\text{bound}} = \pi(r_{\text{max}}^2 - r_{\text{min}}^2)L_{\text{CNT}}$ is the volume accessible to the base in the bound state. Here, $L_{\text{CNT}} = 2.127 \text{ nm}$ and r_{max} and r_{min} are given by

$$\begin{aligned} r_{\text{max}} &= \langle r \rangle + 2\sigma_r \\ r_{\text{min}} &= \langle r \rangle - 2\sigma_r \end{aligned} \quad (4.4)$$

Here, $\langle r \rangle$ and σ_r is the average and standard deviation of the radial position of the center of the base. r_{max} and r_{min} represent the maximum and minimum radial distances from CNT center that the base can access in the bound state. With these definitions, the base-CNT binding free energy corresponds to an aqueous DNA base concentration of 1 M that bind to CNT at a linear density of $1/2.127 \text{ nm}^{-1}$.

The form of ΔF_{res} assumes that the base behaves as an ideal gas that can explore a volume V_{bound} in the bound state. This is an approximation that neglects energetic contributions of the position restraints. However, this approximation is reasonable because of the nature of the base-CNT bound state. The CNT surface is quite smooth and the DNA base can slide laterally along its surface with ease; the energy barriers for lateral motion along the CNT are less than $k_B T$. Additionally, in the bound state, the DNA lies flat and firmly attached to CNT surface and experiences only minor fluctuations in its

radial position. To summarize, the definitions above assume that, in the bound state, the DNA base experiences a potential that is flat in the axial and circumferential directions, and has the form of an infinite well with walls at r_{\max} and r_{\min} in the radial direction.

4.2.2 Results and Discussion

The $\langle \partial H / \partial \lambda \rangle$ curves for perturbing (creating/annihilating) the bases were smooth, well-behaved functions which indicates that the spacing of λ values is adequate (Figure 4.5). The average error in each $\langle \partial H / \partial \lambda \rangle$ value is only 1%, which shows that our results are well converged. Table 4.1 gives ΔF_{bind} along with ΔF_{ele} , ΔF_{vdw} and ΔF_{res} , the electrostatic, van der Waals and position restraint contributions, respectively. G has the lowest (most negative) ΔF_{bind} and thus, has the highest affinity for CNT. Alternatively, C has the highest ΔF_{bind} and lowest binding affinity. ΔF_{bind} for the four bases follows the trend

$$\Delta F_{\text{bind}}^{\text{G}} < \Delta F_{\text{bind}}^{\text{A}} < \Delta F_{\text{bind}}^{\text{T}} < \Delta F_{\text{bind}}^{\text{C}} \quad (4.5)$$

This trend agrees with measurements of the adsorption isotherms of bases on graphite.¹⁴¹ The magnitude of all four ΔF_{bind} values are on the order of 10 kcal/mol ($17 k_{\text{B}}T$), indicating strong base-CNT binding. ΔF_{vdw} accounts for almost the entirety of ΔF_{bind} ; there are only minor contributions from ΔF_{ele} and ΔF_{res} (Table 4.1). This indicates that vdW interactions play the most dominant role in base-CNT binding. The values for ΔF_{bind} calculated at 280 K and 320 K were approximately the same as those computed at

300 K. Thus, ΔF_{bind} has very little temperature dependence and $T\Delta S_{\text{bind}} \ll \Delta E_{\text{bind}}$.

Therefore, base-CNT binding is governed by energetic effects and $\Delta F_{\text{bind}} \approx \Delta E_{\text{bind}}$.

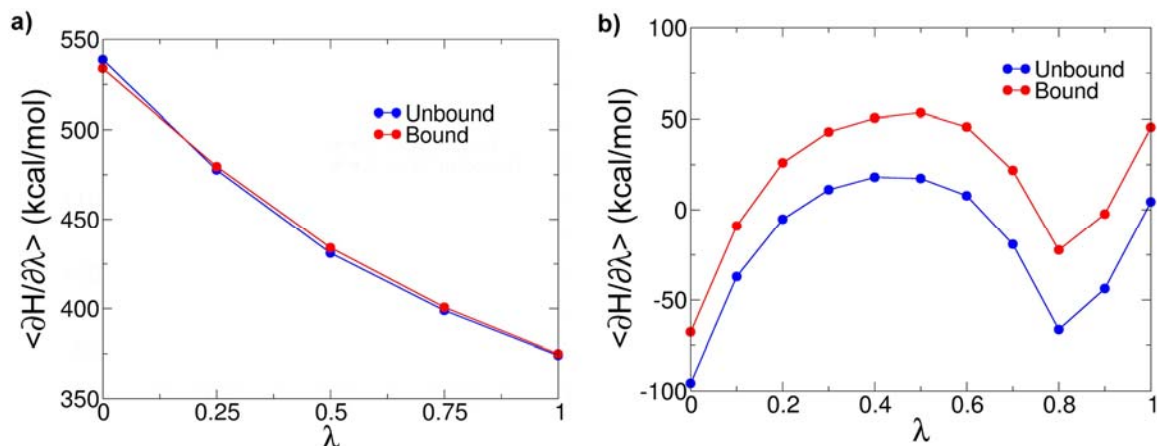


Figure 4.5: $\langle \partial H / \partial \lambda \rangle$ curves for annihilating A. Error bars are smaller than the data point symbols. Similar plots are obtained for C, G and T (data not shown).

Base	ΔF_{ele}	ΔF_{vdw}	ΔF_{res}	ΔF_{bind}
A	-0.3 ± 0.1	-8.6 ± 0.3	0.5 ± 0.1	-8.4 ± 0.3
C	-0.3 ± 0.1	-6.7 ± 0.4	0.3 ± 0.1	-6.7 ± 0.4
G	-0.8 ± 0.1	-10.0 ± 0.3	0.5 ± 0.1	-10.3 ± 0.4
T	-0.1 ± 0.1	-8.1 ± 0.3	0.4 ± 0.1	-7.7 ± 0.4

Table 4.1: Base-CNT binding free energy ΔF_{bind} . ΔF_{ele} , ΔF_{vdw} and ΔF_{res} are the electrostatic, vdW and position restraint contributions to ΔF_{bind} , respectively. Here, $\Delta F_{\text{bind}} = \Delta F_{\text{ele}} + \Delta F_{\text{vdw}} + \Delta F_{\text{res}}$. All values in kcal/mol.

ΔE_{bind} can be expanded as a summation of pairwise additive terms for the interactions between the three components of the system: base (B), carbon nanotube (CNT) and water (W). Thus,

$$\Delta E_{\text{bind}} = \Delta E_{\text{B-B}} + \Delta E_{\text{B-CNT}} + \Delta E_{\text{B-W}} + \Delta E_{\text{CNT-CNT}} + \Delta E_{\text{CNT-W}} + \Delta E_{\text{W-W}} \quad (4.6)$$

These contributions are summarized graphically in Figure 4.6. Because both the bases and CNT are rigid, they do not undergo any structural rearrangement upon binding and

$\Delta E_{B-B} = \Delta E_{CNT-CNT} = 0$. Below we examine the remaining contributions to ΔE_{bind} .

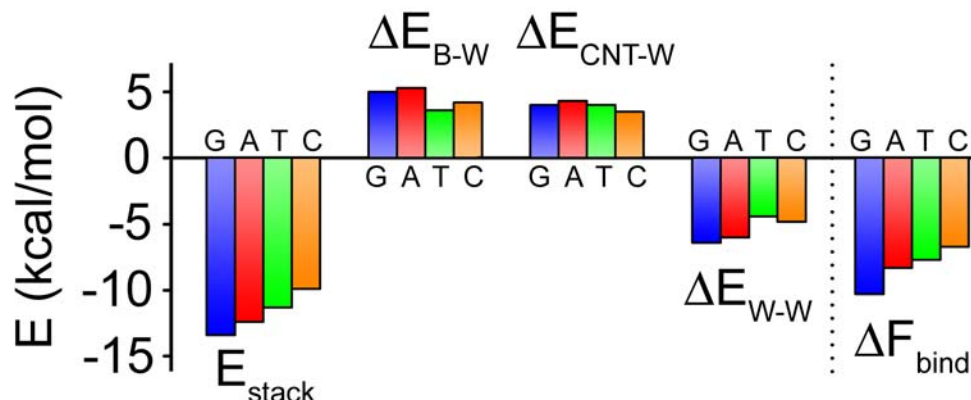


Figure 4.6: Average base-CNT π - π stacking (vdW) energy E_{stack} , difference in base-water energy ΔE_{B-W} , CNT-water energy ΔE_{CNT-W} and water-water energy ΔE_{W-W} for bound and unbound states. The sum of these four terms is approximately equal to ΔF_{bind} .

In the bound state, strong vdW forces occur between the base and CNT. This interaction is expected among planar, aromatic molecules and is referred to as the π - π stacking interaction. CNT atoms are uncharged and experience no electrostatic interactions with the environment. Therefore, $\Delta E_{B-CNT} = E_{stack}$, where E_{stack} is the base-CNT π - π stacking (vdW) energy. The strength of this interaction varies among the four bases (Figure 4.6) and follows the trend:

$$E_{stack}^G < E_{stack}^A < E_{stack}^T < E_{stack}^C \quad (4.7)$$

This trend is understandable from geometric considerations. In the bound state, all atoms of the planar bases are in contact with CNT. E_{stack} scales with surface area overlap of the interactants; the purines, G and A, which contain two aromatic rings have a stronger interaction with CNT than pyrimidines, C and T, which contain only a single ring.

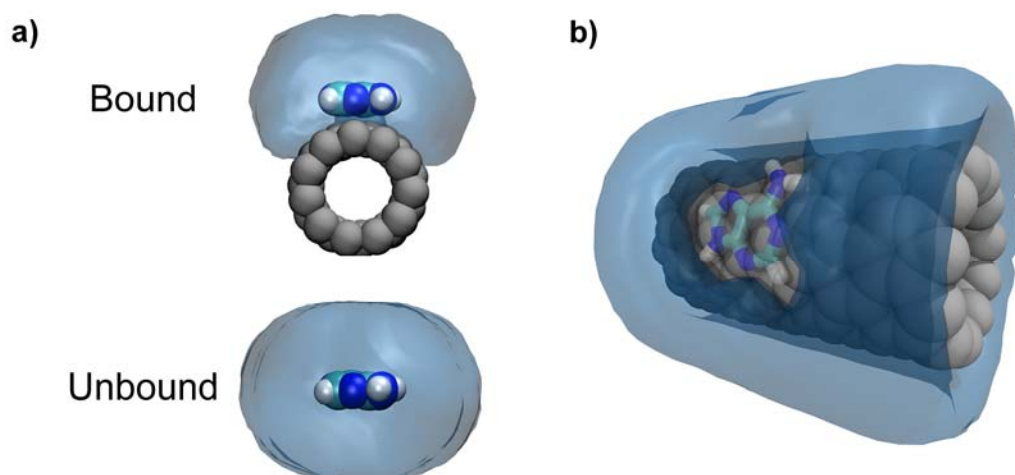


Figure 4.7: (a) Water density around bound and unbound adenine. In the unbound state, water forms a 3D solvation shell around the base. In the bound state, solvent density on one face is reduced to the excluded volume effects of CNT. (b) Water density around CNT showing water excluded from the CNT surface due to the presence of the bound base.

Unlike, $\Delta E_{\text{B-CNT}}$ which provides a comparatively large, negative contribution to ΔF_{bind} , $\Delta E_{\text{B-W}}$ and $\Delta E_{\text{CNT-W}}$ are positive and represent desolvation penalties to base-CNT binding (Figure 4.6). In the unbound state, both the DNA base and CNT are completely solvated by water molecules. However, after binding, water is excluded from the base-CNT interface (Figure 4.7). The reduced solvation diminishes the number of attractive interactions with water molecules, thereby increasing the total CNT-water and base-water vdW energy. Interestingly, even though water molecules are completely excluded from one face of the bound base, no base-water hydrogen bonds are broken upon binding. Because of the base's planar structure, polar groups on the periphery remain well exposed to solvent. Additionally, since the bound base rests on the convex (outer) CNT surface, water molecules are able to solvate the base-CNT interface from underneath (Figure 4.7a). These two properties enable bases to maintain their local

solvation structure (i.e. the first solvation shell) and remain capable of hydrogen bonding with water molecules even in the bound state (Figure 4.8).

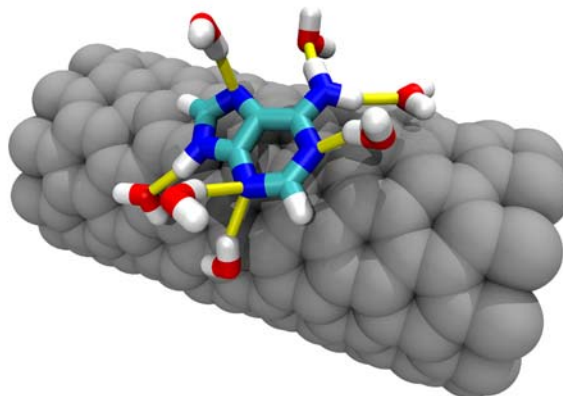


Figure 4.8: Water remains capable of forming hydrogen bonds (yellow) with bound bases (adenine shown here). Hydrogen bonding is not impeded by CNT.

The change in water-water interaction energy upon base-CNT binding $\Delta E_{\text{w-w}}$ is difficult to converge because it involves subtracting large numbers that differ by only a small amount. Therefore, we estimate this quantity by subtracting the other energy values from ΔF_{bind} . Thus,

$$\Delta E_{\text{w-w}} \approx \Delta F_{\text{bind}} - E_{\text{stack}} - \Delta E_{\text{B-W}} - \Delta E_{\text{CNT-W}} \quad (4.8)$$

$\Delta E_{\text{w-w}}$ is negative for all bases and partially compensates for the energetic penalties of desolvating the base and CNT (Figure 4.6). In the unbound state, water completely solvates the hydrophobic CNT surface resulting in unfavorable dangling hydrogen bonds at the water-CNT interface.¹⁴² The presence of a bound base reduces the amount of hydrophobic CNT surface area accessible to solvent. This decreases the amount of dangling hydrogen bonds and results in more favorable water-water interactions in the bound state.

While the work presented here analyzed the binding of DNA bases to the outside of one specific type of CNT, the results are easily generalized for other CNT chiralities and binding geometries. The π - π stacking energy becomes stronger as more base atoms come into contact with CNT. Because the base has a planar structure, base-CNT contact will be maximized on a CNT of zero curvature, i.e. on the graphene surface. As CNT curvature increases, base-CNT contacts are lost and the π - π stacking interaction is weakened. Solvent effects follow a differing trend; the base is most accessible to water molecules (and thus, base-water hydrogen bonds) when bound to the outer wall of CNTs of high curvature. Solvent accessibility steadily diminishes with decreasing CNT curvature. Additionally, because of the concave surface geometry, base-water interactions would be minimized for bases bound to the CNT inner wall. Thus, considering π - π stacking and solvent accessibility together, base-CNT binding should be greatest on a graphene sheet and weakest on the inner CNT surface.

The calculations presented here involve a zigzag CNT whereas, under typical experimental conditions, ssDNA encounters CNTs of widely distributed chirality. However, our calculations suggest that base-CNT binding is not affected by CNT chirality. With the classical models employed here, the energy barriers for lateral motion of bases across CNT surface are less than $k_B T$. For example, these barriers are $0.4k_B T$ and $0.6k_B T$ for armchair and zigzag CNTs, respectively. This indicates that at room temperature, the CNT presents small barriers for lateral movement of bound bases, in agreement with molecular mechanics calculations of adenine on graphite.¹⁴³ Thus, the orientation of the base is not sensitive to the underlying CNT atomic structure. This is consistent with recent measurements that suggest that positional fluctuations in

microbeads tethered to ssDNA-coated CNTs may be due to motion of ssDNA on the CNT surface.¹⁴⁴ However, *ab initio* calculations of bases on graphene^{137, 138} or CNT⁸¹ predict a rougher energy corrugation of $1.2k_{\text{B}}T-4.0k_{\text{B}}T$. These larger values may be due explicit electron-electron interactions that are underestimated using classical potentials. However, they also may be due to artifacts of the density functionals employed in these computations.¹⁴⁵

4.3 Stability of Proposed Structures

Atomic force microscopy (AFM) has been the most popular experimental tool to study DNA-CNT structure. AFM images of DNA-CNT based on the poly GT (sequences of repeating guanine and thymine nucleotides) sequence and formed using sonication are reported to show alternating bands of high and low regions on the surface of the hybrid with a uniform spacing of 18 nm.⁶⁸ These features, along with the observed effectiveness of poly GT sequences in ssDNA mediated CNT sorting, led Zheng *et al.* to propose that poly GT forms a homodimeric structure held together by an exotic, non-Watson-Crick hydrogen bond network and wrap in a helical fashion about CNTs with a pitch commensurate with this 18 nm spacing.⁶⁸ Their proposed ssDNA wrapping conformation and base pairing arrangement is shown in Figure 4.9. Later AFM measurements of DNA-CNT have reported similar band features but with a smaller spacing of about 14 nm that was independent of sequence.⁷² A recent scanning tunneling microscopy (STM) study suggests an even smaller band spacing of 3.3 nm.¹³⁵ Measurements of the circular dichroism of DNA-CNT also suggest helical wrapping.⁸³ However, AFM measurements on DNA-CNT formed without sonication reveal that adsorbed ssDNA forms a thicker, presumably disordered, layer on CNTs that is featureless in AFM.²¹ To make sense of the disparate set of experimental data and test the validity of these interpretations, we performed a series of MD simulations that investigate the stability of these proposed structures and others related to them.

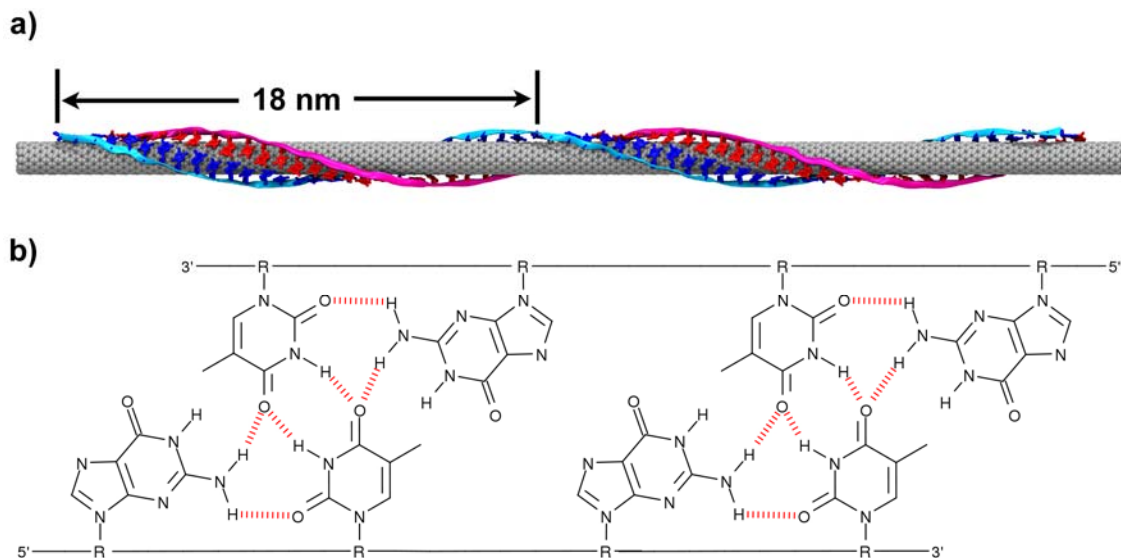


Figure 4.9: (a) Proposed dimeric structure of poly GT ((GT)₃₀ shown here) around CNT. (b) Proposed non-Watson-Crick hydrogen bond scheme for poly GT sequences adsorbed to CNT.

4.3.1 Stability of Poly GT Dimers Adsorbed to Carbon Nanotubes

First, we investigated the structural stability of the poly GT dimer. Two (GT)₂ oligonucleotides were constructed in linear conformations and placed adjacent to each other on a (11,0) CNT in a $3.5 \times 3.5 \times 3.83 \text{ nm}^3$ box (Figure 4.10a). In order to facilitate dimerization, the ssDNA backbones were placed antiparallel and their bases oriented to mimic the proposed hydrogen bond pattern (Figure 4.10b). However, we found that this arrangement induces high stress within the ssDNA sugar residues and glycosidic bonds. Performing energy optimization on this structure alleviated this stress, but significantly altered the geometry and led to a configuration that was incompatible with the original base pairing scheme (Figure 4.10a). We therefore conclude that the proposed hydrogen bond pattern is energetically unfavorable and structurally unstable.

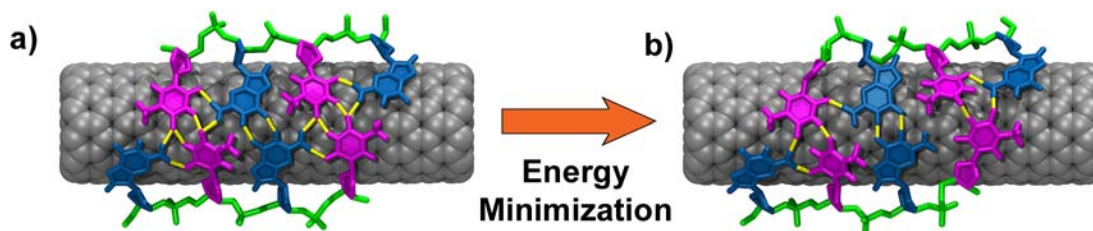


Figure 4.10: (a) Initial (GT)₂ dimer following the proposed base pair scheme. Guanine and thymine are shown in blue and magenta, respectively. Hydrogen bonds are shown in yellow. (b) (GT)₂ dimer after energy minimization. Several of the hydrogen bonds have been broken.

Nevertheless, the optimized structure retains some of its original hydrogen bonds (Figure 4.10b). To explore the possibility that these remaining bonds could stabilize a dimer, we performed a 5 ns simulation starting from this energy minimized configuration. The simulation results in the two strands gently separating and diffusing away from one another; no stable dimer is formed. Apparently, any attraction between the oligonucleotides is nonspecific and insufficient to survive thermal fluctuations at 300 K. It is reasonable to imagine that longer oligonucleotides containing many more inter-ssDNA hydrogen bonds would be more prone to dimerize than the short (GT)₂ sequences used here. However, we carried out the identical simulation using (GT)₂₀, a sequence ten times longer than the previous one, and obtained qualitatively similar results. The initial (GT)₂₀ dimer contained a total of 73 interstrand hydrogen bonds. However, after 5 ns of MD, only 12 remained. These residual hydrogen bonds occurred between random bases and were not arranged in any specific fashion and did not impose any overall structure on the adsorbed oligonucleotides.

We also considered dimer structures based on “wobble base pairing”, which is known to initiate formation of double helices of poly GT sequences at low temperature.^{66, 67} However, this scheme is also highly unlikely to occur among adsorbed ssDNA for the

following reason. In wobble base pairs, one hydrogen bond exists between atoms O₆ of guanine and N₃ of thymine and another between atoms N₁ of guanine and O₂ of thymine (Figure 4.11a). However, when thymine is adsorbed to CNT, the O₂ and N₃ atoms are oriented towards the interior of the oligonucleotide which restricts their ability to hydrogen bond to an opposing guanine (Figure 4.11b).

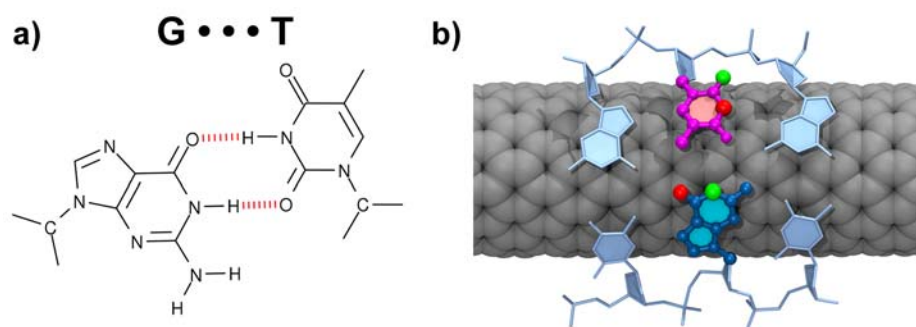


Figure 4.11: (a) G-T Wobble base pair. (b) Two poly GT strands adsorbed to SWCN. Guanine and thymine are shown in blue and magenta, respectively. Atom pairs that would normally share hydrogen bonds in a wobble base pair are colored red and green, respectively. The geometry assumed by adsorbed bases makes them incompatible with wobble base pairing.

While not an exhaustive search of all possible base pairing schemes, the preceding results are evidence against the formation of stable poly GT dimers adsorbed to CNTs. This conclusion is further corroborated by established facts about double stranded DNA (dsDNA). In dsDNA, base pairs reside in a planar geometry which maximizes the strength of the interstrand hydrogen bonds.⁶⁵ However, the adsorbed poly GT nucleotides reside on the curved CNT surface and are unable to form planar base pairs. This geometry significantly reduces the cohesive strength of interstrand hydrogen bonds. Additionally, in dsDNA, adjacent bases are stacked on top of one another in a spiral staircase fashion.⁶⁵ Because of this stacked geometry, bases are freely able to hydrogen

bond to their counterparts without steric hindrance. These stacking interactions also provide rigidity in dsDNA and are as important as base pair interactions in stabilizing the double helix.⁶⁵ However, in the adsorbed state, adjacent poly GT bases lie roughly in the same plane and must compete for space on the CNT surface. Thus, adsorbed bases are subjected to steric restrictions that impede their ability to form base pairs. Also, because of the lack of intrastrand stacking interactions, adsorbed poly GT adopts a more disordered structure compared to dsDNA. Consequently, interstrand interactions among the adsorbed oligonucleotides tend to be more random and nonspecific.

4.3.2 Stability of Helically Wrapped DNA About Carbon Nanotubes

Since we find that dimer formation is improbable, we performed several simulations of a single adsorbed poly GT oligonucleotide in order to determine the stability of helically wrapped poly GT. We constructed a (11, 0) CNT initially wrapped with a sixty base long poly GT ((GT)₃₀) sequence adopting an 18 nm pitch helix (Figure 4.12). This oligonucleotide is identical to that used by Zheng *et al.* in their original AFM experiments.⁶⁸ Due to the chiral ssDNA backbone, we identified four distinct conformations consistent with helical wrapping. These structures differ in helical handedness (*left* or *right*) or orientation of the sugar-phosphate backbone with respect to CNT. For clarity, the backbone orientation is defined by the position of the O_{4'} atom of the sugar group, which can point radially *inward* or *outward* from CNT (Figure 4.13). Thus, the four initial structures are *LH-inward*, *LH-outward*, *RH-inward* and *RH-outward*, where *LH* and *RH* stand for left-handed and right-handed, respectively. To construct each helix, the orientation of (GT)₁, a two base long poly GT sequence, was

optimized on the surface of CNT in both the *inward* and *outward* configuration. Then a series of translations and rotations about the CNT axis were applied to the bases with coordinates being saved after each step. These transformations caused the bases to trace out a left- or right-handed helix of 18 nm pitch with 26 bases per turn. This value was determined to minimize both steric repulsion between adjacent bases and bond stretching of the backbone under the constraint of an 18 nm pitch. The coordinates were then concatenated to produce (GT)₃₀ in each of the four helical conformations described above. Each structure was relaxed over the course of an 80–100 ns MD simulation in aqueous solution in a $3.5 \times 3.5 \times 39.9 \text{ nm}^3$ box.

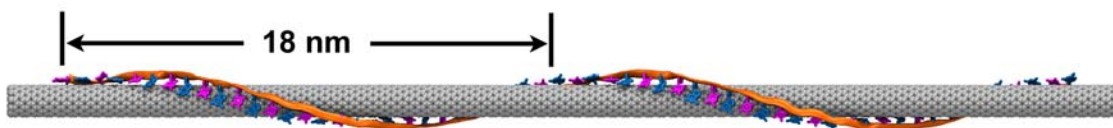


Figure 4.12: (GT)₃₀ on a (11,0) CNT with regular 18 nm pitch.

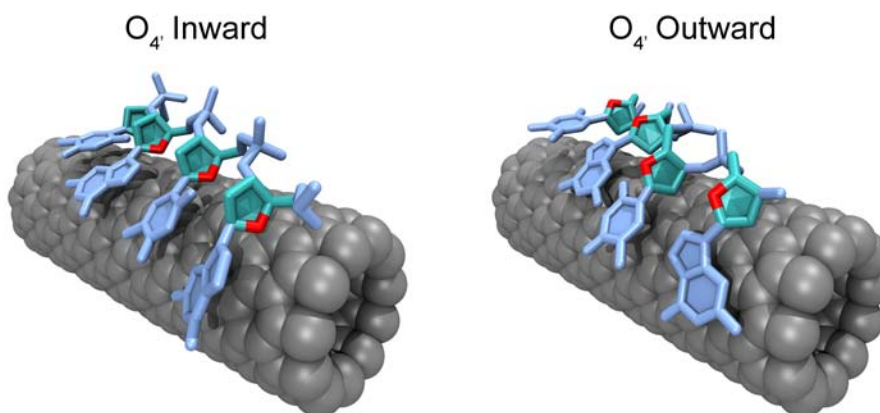


Figure 4.13: Different sugar-phosphate backbone orientations with respect to CNT. O₄ (red) points either radially inward or outward from CNT.

In each system, (GT)₃₀ retains a helical wrapping about CNT but undergoes an overall reduction in pitch. For *RH-outward* and *LH-inward*, the oligonucleotide actually unwraps into a linear configuration before rewrapping into a helix of opposite initial handedness. The final pitch values for the four systems range from 2 nm for *LH-outward* to 8 nm for *RH-inward*. The decrease in pitch is accompanied by an increase in the number of helical turns about CNT circumference. These results indicate that helical wrapping is a viable ssDNA structure about CNT. However, pitch values over about 10 nm are unfavorable.

Strikingly, the *inward* and *outward* ssDNA conformations are energetically distinguishable; the *inward* structure has the lower potential energy by 2.5 kcal/mol-base. This energetic difference is largely due to the ssDNA-water and ssDNA-CNT interactions. Compared to the *outward* conformation, adjacent bases residing in the *inward* conformation tend to be more spatially separated and more heavily solvated by water. As a result, each thymine has approximately 0.67 more hydrogen bonds with water for the *inward* conformation. Additionally, the *inward* ssDNA backbone is about 0.8 Å closer to CNT which results in a more favorable ssDNA-CNT van der Waals interaction.

4.3.3 Mechanism for Helical Wrapping of DNA about Carbon Nanotubes

To investigate the mechanism responsible for ssDNA helical wrapping, we carried out simulations of two systems, S_1 and S_2 , where a (GT)₂₀ oligonucleotide initially adopts a linear, ahelical conformation on top of a (11,0) CNT in a $3.5 \times 3.5 \times 27.7 \text{ nm}^3$ box (Figure 4.14a). In each system, ssDNA was initialized in the *inward* configuration, but with a different set ($\alpha, \beta, \gamma, \delta, \epsilon, \zeta$) of backbone torsion angles. The initial average torsional angles were (260°, 162°, 59°, 140°, 159°, 135°) and (194°, 116°, 62°, 104°,

143°, 216°) for S_1 and S_2 , respectively. The initial configurations were constructed using similar methods to those employed above for construction of the (GT)₃₀ helices. However, since linear conformations were desired in this case, only translations (no rotation) along the CNT axis were applied to the bound (GT)₁ sequence. Each system was allowed to relax in aqueous solution for 7 ns.

Remarkably, the linear oligonucleotide in S_1 and S_2 spontaneously winds around the CNT into a right- and left-handed helix of 2–3 nm pitch, respectively. In both systems, helical wrapping does not occur uniformly over the entire length of the oligonucleotide. Instead, the 5' end remains essentially stationary (apart from thermal motion) while the 3' end rapidly encircles the CNT circumference (Figure 4.14b). As a result, additional helical turns are generated at the 3' end, which then propagate towards the 5' end. The winding continues until ssDNA forms a compact helix about CNT (Figure 4.14c).

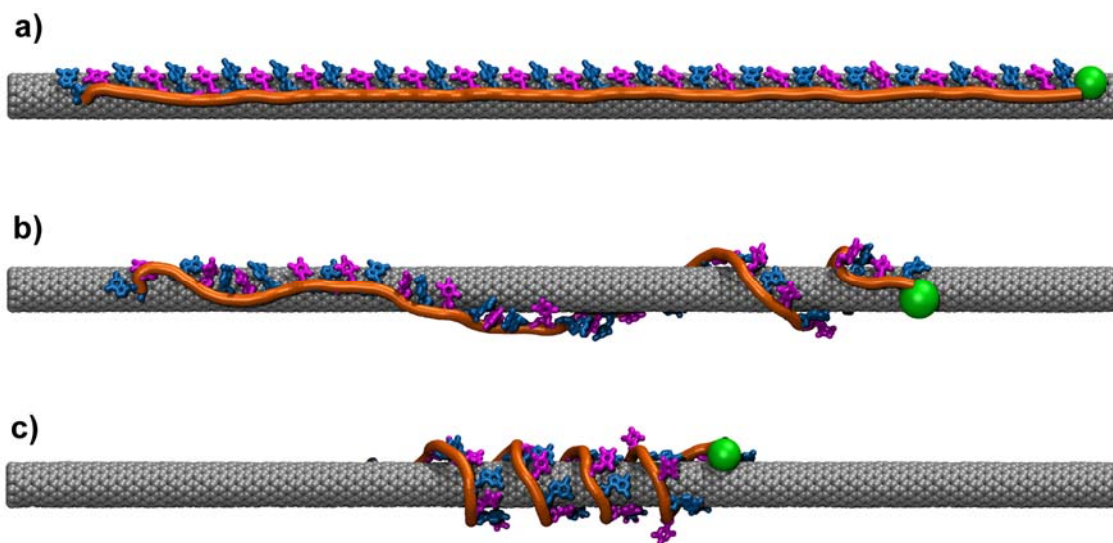


Figure 4.14: Simulation of S_1 displaying right-handed helical wrapping of (GT)₂₀ about CNT. (a) Initial configuration. (b) Configuration after 2.2 ns. (c) Configuration after 7 ns. The green sphere marks the ssDNA 3' end. Similar results occur for S_2 , but with left-handed helical wrapping.

Analysis of structural and energetic changes that occur in the two systems reveals that electrostatic interactions within the ssDNA backbone are responsible for wrapping the initially linear oligonucleotide into a helical structure. As the helix forms, the average distance between adjacent phosphates d_{P-P} steadily increases, thereby relieving electrostatic repulsion E_{P-P} within the backbone (Figure 4.15). These structural changes proceed via a rearrangement of the backbone torsional angles (Figure 4.15). It is apparent that the differing initial sets of torsional angles in S_1 and S_2 enable ssDNA to evolve along two drastically different pathways that ultimately lead to a right- and left-handed structure, respectively. The importance of backbone torsion in helix formation is most likely why the wrapping occurs asymmetrically with turns being generated at the 3' end. The chiral sugar-phosphate backbone results in chiral forces along the oligonucleotide which may facilitate deformations along preferred directions.

It should be pointed out that the preceding observations arise from the artificial initial condition of linear ssDNA and that it is possible that other non-helical conformations may result from a more random initial configuration. However, these results provide several interesting insights about the mechanics of oligonucleotides adsorbed to CNT. Firstly, the sugar-phosphate backbone contains intrinsic curvature and prefers a helical wrapping to a linear structure. Secondly, while π - π stacking is the main driving force for DNA-CNT self-assembly, the backbone dictates the overall ssDNA conformation. Because the sugar-phosphate backbone is not specific to base sequence, general ssDNA sequences are thus expected to wrap CNT in a similar manner to that observed here. We have verified this by obtaining similar results with a random 40-base long sequence. There is also reason to believe that the 3' to 5' wrapping may be a general feature of

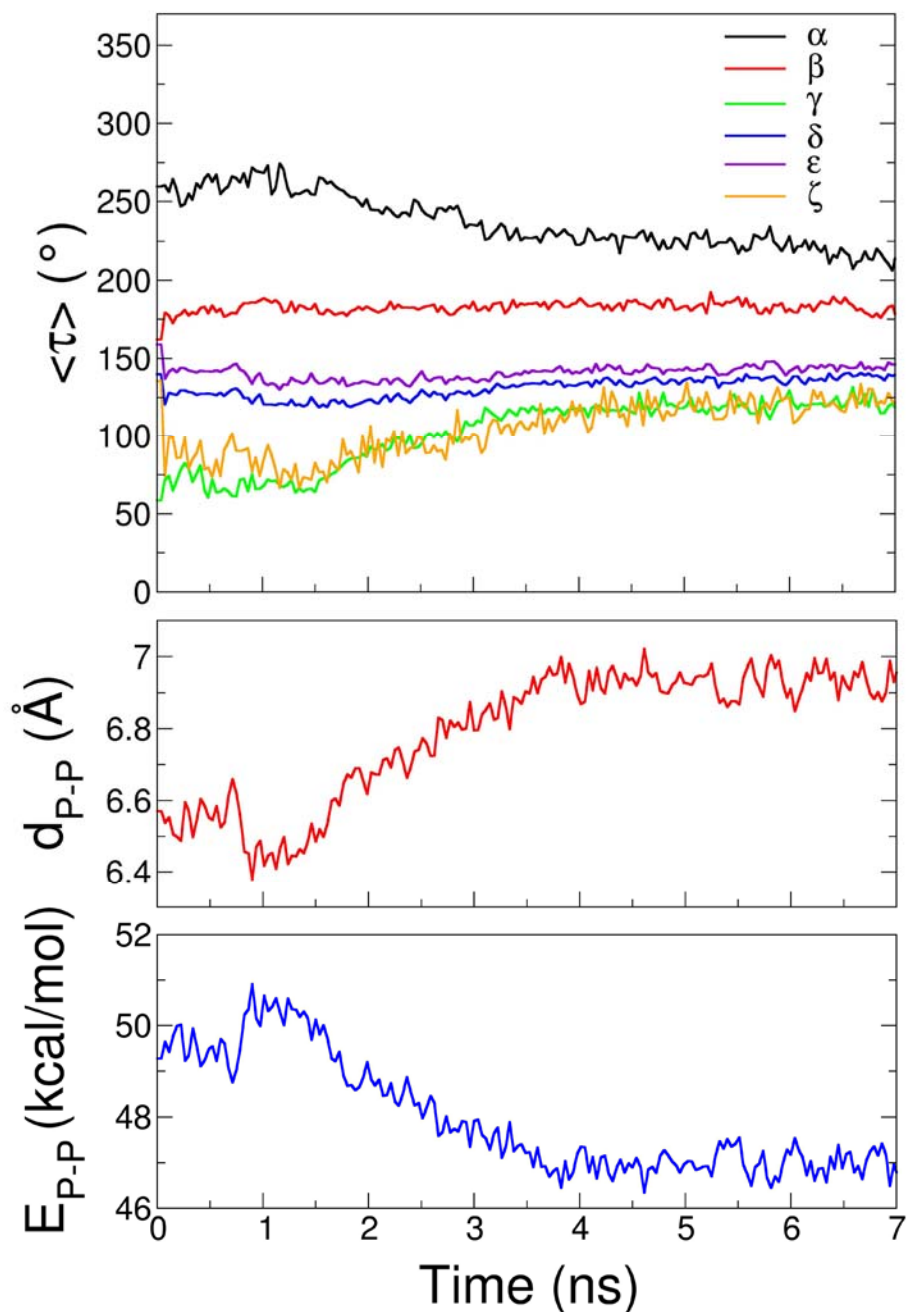


Figure 4.15: Structural and energetic changes for S_1 . The average distance (d_{P-P}) between adjacent phosphorous atoms (top) increases as the helix forms, thereby reducing the electrostatic repulsion (E_{P-P}) between them (middle). These changes proceed via a rearrangement of the average torsional angles ($\langle \tau \rangle$) in the ssDNA backbone (bottom). Qualitatively similar data (not shown) occurs for S_2 .

DNA-CNT as it has been observed in other systems with differing initial conditions. For example, in the previously described simulations involving the relaxation of the (GT)₃₀ 18 nm pitch helices, the overall reduction of pitch and increase in the number of helical turns also proceeded via a 3' to 5' wrapping.

It is widely appreciated that electrostatics play a vital role in DNA deformation^{65, 146} and in the polymorphism of the double helix.⁶⁵ For example, the form⁶⁵ and mechanical properties¹⁴⁷ of the double helix are extremely sensitive to the salt concentration of the solvent. Counterions alter the dielectric properties of the solvent and screen electrostatic repulsion between phosphates. We have observed salt concentration dependent effects in our simulations as well. Adding 0.5 M NaCl to S_1 and S_2 effectively neutralizes the negatively charged backbone and quenches helical wrapping. This justifies the conclusion that helix formation is due to electrostatic repulsion within the backbone.

At low salt concentration, these MD simulations show that adsorbed ssDNA is found to prefer a compact helical wrapping with a pitch less than 10 nm, while at high salt concentration, helical wrapping is suppressed entirely. It remains an open question whether a more elongated helical structure is preferred at intermediate salt concentrations. Experiments suggest a variety of pitch values ranging from 3.3 – 18 nm.^{68, 72, 135} The DNA-CNT in these measurements are dried in air prior to imaging, thus altering the dielectric environment around the hybrids from that in aqueous solution. This may result in pitch values that differ from those found in a fully hydrated DNA-CNT.

Although the resulting helical structures observed here exhibit global order in the form of helical wrapping, a considerable amount of local disorder is present throughout adsorbed ssDNA. While the oligonucleotide is initialized with all bases stacked to CNT,

thermal fluctuations and steric limitations cause many bases to desorb throughout the simulation (Figure 4.16). While this could be due to an energetically unfavorable starting configuration with adjacent bases placed too close together, we find that they are actually a characteristic structural feature of DNA-CNT based on a more detailed analysis of desorbed bases presented in Section 4.4.2. Disorder is also present at both ends of ssDNA. Because the ends reside outside of the interior “bulk” region of the ssDNA backbone, approximately the first and last three bases tend to self-arrange on CNT surface in a nonspecific, ahelical way.

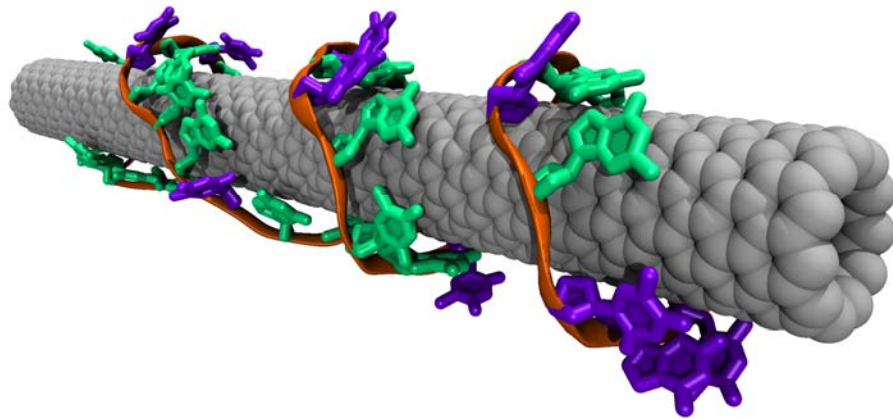


Figure 4.16: Desorbed bases (purple) introduce significant disorder within adsorbed ssDNA. Thermal fluctuations and steric limitations cause many of the initially adsorbed bases to detach from CNT surface.

4.4 Free Energy Landscape of a DNA-Carbon Nanotube Hybrid

The MD simulations in previous Sections show that DNA-CNT self-assembles via π - π stacking, with ssDNA adopting various energetically and structurally distinct wrapping conformations about CNT.²⁵ However, final ssDNA configurations obtained in such simulations are highly sensitive to the system's initial conditions. This indicates that such simulations do not fully sample the entirety of ssDNA's configuration space over the simulation timescale. Oligonucleotides, a flexible biopolymer, have many degrees of freedom and a rugged potential energy landscape. They can become trapped in metastable configurations associated with local energy minima that persist over typical MD timescales (10–100 ns). This severely limits conformational sampling and prohibits accurate determination of equilibrium properties using conventional MD techniques. Replica exchange molecular dynamics¹²⁴ (REMD) overcomes this limitation through the use of multiple MD simulations (replicas) in parallel at multiple temperatures. High temperature replicas enable rapid barrier crossing and sample additional, possibly rare, configurations not easily observable in conventional room temperature MD simulations. We have employed REMD in order to fully characterize the equilibrium structure of a DNA-CNT. We calculate the free energy landscape and find minima corresponding to six distinct conformations, with a non-helical loop structure as the global minimum. The hybrid contains significant structural disorder, with desorbed bases as an important structural feature. These results expand our understanding of DNA-CNT and indicate the relevance of REMD for explorations of the physical properties of organic-inorganic multifunctional nanomaterials.

4.4.1 Simulation Details

REMD was performed to explore the conformations of a (GT)₇ oligonucleotide adsorbed to an infinite (11,0) CNT in aqueous solution. The initial configuration of the (GT)₇-CNT hybrid was taken from the final frame of a separate MD simulation of the adsorption of these materials following the methods of Section 4.1. The hybrid was solvated with 2559 water molecules in a box of dimensions $L_x = L_y = 3.5$ nm and $L_z = 8.083$ nm. Prior to REMD, a 1 ns MD simulation was performed at constant temperature¹²⁸ (300 K) and pressure¹²⁷ (1 atm) in order to equilibrate the box dimensions and ensure a proper solvent density.

Determining an adequate temperature tiling $\{T\} = T_0, T_1, \dots, T_{M-1}$ of the replicas is crucial for a successful REMD simulation (Section 3.4). To acquire $\{T\}$, we used the following method. First, we varied the temperature in a series of MD simulations to determine the temperature at which kinetic trapping of ssDNA was minimized. For this system, because of the strong ssDNA-CNT cohesion, temperatures exceeding 600 K were necessary to remove kinetic trapping effects and generate large conformational changes over sub-nanosecond timescales. Secondly, two series of REMD simulations, one at low temperature ($T \sim 300$ K relevant to ambient conditions) and the other at high temperature ($T \sim 700$ K where kinetic trapping effects are negligible), were performed. In each simulation, two replica (replica 0 and 1) were employed with temperatures given by

$$\begin{aligned} T_0^{\text{low}} &= 300 \text{ K}, & T_1^{\text{low}} &= T_0^{\text{low}} + \Delta T^{\text{low}} \\ T_0^{\text{high}} &= 700 \text{ K}, & T_1^{\text{high}} &= T_0^{\text{high}} + \Delta T^{\text{high}} \end{aligned} \quad (4.3)$$

The temperature spacings ΔT^{low} and ΔT^{high} were then adjusted for the two simulations until an acceptance ratio of 20% occurred between replica 0 and 1. This provided the

desired temperature spacing at low and high temperature. Using the values of ΔT^{low} and ΔT^{high} , the temperature spacing at any temperature $\Delta T(T)$ could be fit to a linear function given by

$$\Delta T(T) = -2.58\text{K} + 0.0205T \quad (4.2)$$

This function was used to determine the proper temperature tiling and total number of replicas for our system.

For our REMD simulation of a (GT)₇-CNT hybrid, sixty four replicas were employed with T_m , the temperature of replica m , spanning from 290 K to 715 K. Following from Equation 4.2, the temperature of the m -th replica was given by

$$T_m = 125.7\text{K} + 164.3\text{K} e^{m/49.3} \quad (4.3)$$

This temperature distribution resulted in an acceptance ratio of $21\% \pm 1\%$ between all replicas. Each replica was run at constant temperature¹²⁸ and volume. Exchanges between adjacent replicas were attempted every $\tau_{\text{exchange}} = 0.6$ ps. (GT)₇ coordinates were saved every 0.6 ps. To enhance sampling and reduce any artifacts introduced by our choice of initial conditions, we performed two REMD simulations. The first was run for 55.5 ns using the starting configuration described previously. We then ran an additional 38.3 ns starting from the most probable conformation (global free energy minimum) of the first simulation. The two simulations combined provided a 94 ns trajectory (6 μs total integration time) yielding almost 10 million ssDNA configurations over the entire temperature range. These computations were run across 2048 CPUs (32 CPUs per replica) of the IBM Blue Gene/L at the San Diego Supercomputer Center. Recently, another group performed an independent 5 ns REMD simulation of ssDNA 10mers in complex with CNT using 12 replicas spanning 298 K to 342 K.¹⁴⁸ While many of their

findings agree with the results presented here, there are a few marked differences that are discussed below.

4.4.2 Results and Discussion

Free energy landscapes of the (GT)₇-CNT hybrid reveal the preferred oligonucleotide conformations at each temperature. The free energy is calculated as a function of two order parameters x and y :

$$\frac{\Delta F(x, y)}{k_B T} = -\ln\left(\frac{P(x, y)}{P_{\max}}\right) \quad (4.4)$$

Here, x and y are the minimum root-mean-square deviations (RMSD) of the (GT)₇ phosphorous atoms from an ideal left-handed helix and linear conformation, respectively. Each (GT)₇ conformation underwent a rigid body translation and rotation about the CNT axis to minimize the RMSD. $P(x, y)$ is the probability distribution obtained from the REMD data, and P_{\max} is the maximum of $P(x, y)$.

At room temperature, the free energy landscape is rugged and reveals that six distinct (GT)₇ conformations characterize the hybrid's structure (Figure 4.17). The global free energy minimum, and thus most probable oligonucleotide configuration, is the loop. This conformation resembles a letter U draped laterally across CNT and has an average end-to-end length of 3.3 nm. The local minima at (13.5 Å, 11 Å) and (17 Å, 24 Å) (Figure 4.17) are similar structures with altered size that we refer to as the compact and elongated loop, with average end-to-end lengths of 2.2 nm and 4.5 nm, respectively. In agreement with our previous MD simulations, the hybrid contains helical oligonucleotide conformations. The minima at (15.5 Å, 22.5 Å) and (11 Å, 22.5 Å) correspond to distorted right- and left-handed helices. These structures have a pitch of 2.5 nm and

deviate from ideal helices because the 5' end, the 3' end or both bend opposite to the helical turns. Ideal left-handed helices are present in the minimum at (5.5 Å, 17 Å); their backbone is more elongated than the distorted helices (3.2 nm average pitch), and it traces out a well-defined helix with only slight deformations and no irregular bends. Interestingly, an ideal right-handed helix is not encountered, suggesting that (GT)₇ assumes a preferred handedness among its helically wrapped conformations. In all cases, the O_{4'} atom of sugar residues prefers to point towards CNT in agreement with previous results that show that this is the preferable sugar-phosphate backbone configuration.

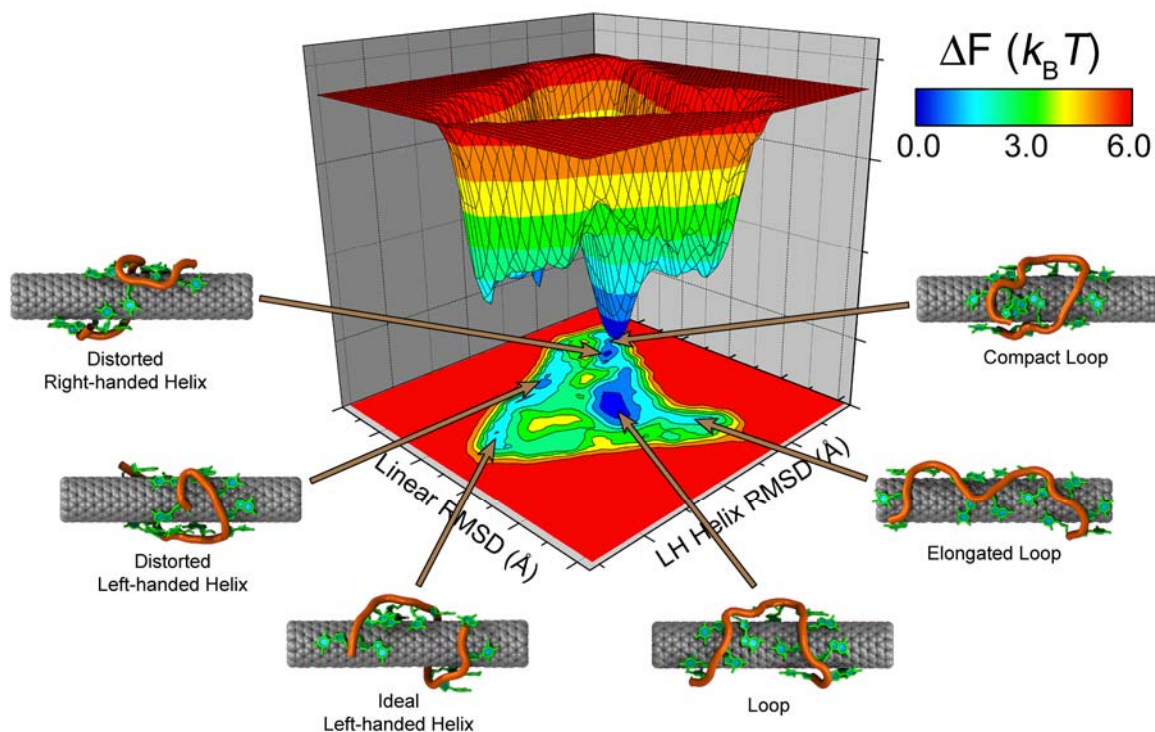


Figure 4.17: Free energy landscape of (GT)₇-CNT hybrid at room temperature. Representative conformations for each local minimum are displayed. The sugar-phosphate backbone is depicted in orange and bases are shown in green.

The REMD study by Martin *et. al* also found that several oligonucleotide conformations contribute to DNA-CNT structure.¹⁴⁸ However, many configurations

found in their work disagree with those observed here. While several of their structures resemble loops, they also found ssDNA conformations containing kinked regions detached from CNT surface that resemble the metastable configurations found in our self-assembly simulations (Section 4.1). We have found that kinks represent high energy ssDNA configurations and can be annealed to more favorable conformations. These discrepancies may be due to our use of a more stringent REMD protocol that included higher temperature replica, which accessed longer timescales and sampled a more extensive region of DNA-CNT configuration space. Additionally, the enhanced sampling performance of our REMD simulation is evident when comparing the (GT)₇-CNT hybrid free energy landscape of Figure 4.17 with that obtained with a conventional 55 ns MD simulation (Figure 4.18). Even with a relatively long MD simulation, kinetic trapping limits sampling of the oligonucleotide's configuration space. As a result, MD visits only a single (GT)₇ conformation as opposed to the entire ensemble sampled with REMD.

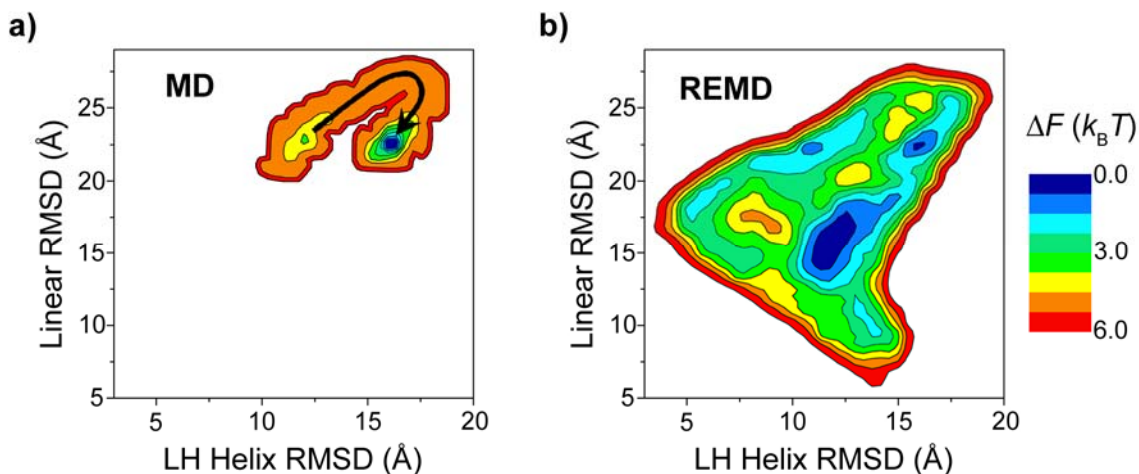


Figure 4.18: (GT)₇-CNT free energy landscape for (a) 55 ns regular MD and (b) 55 ns REMD simulation. The black arrow indicates the evolution of the oligonucleotide into a local free energy minimum. Configurational sampling is drastically improved with REMD.

Each low energy ssDNA conformation contains many bases adsorbed (stacked) to CNT surface. However, base adsorption is limited by steric effects. Unlike the vertical stacking in double stranded DNA (dsDNA), adjacent adsorbed bases lie roughly coplanar and compete for binding space on CNT surface. For many conformations, these spatial limitations result in desorption of bases. Desorbed bases can be stabilized by stacking interactions with neighboring bases (Figure 4.19a), similar to what occurs in dsDNA where adjacent bases stack in a spiral staircase fashion. The global minimum loop conformation maximizes the number of adsorbed bases (Figure 4.19b). Of the sampled loops, the vast majority (over 76%) have all 14 bases adsorbed. This differs from non-loops (i.e. conformations outside the global minimum) where (GT)₇ generally contains at least one desorbed base. Base adsorption results from strong van der Waals forces (π - π stacking interactions) and leads to a more energetically favorable (GT)₇-CNT configuration. The strong preference for loop conformations indicates that the free energy of the hybrid is minimized by maximizing base-CNT stacking. This phenomenon is likely

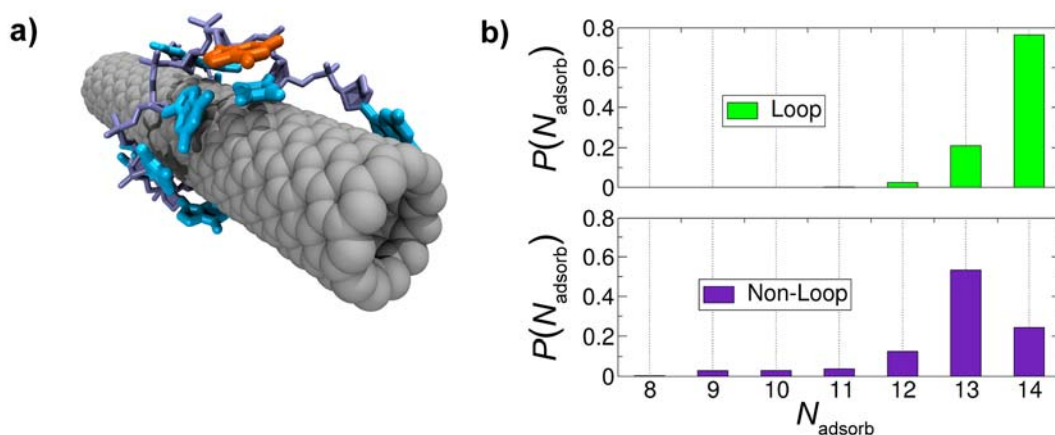


Figure 4.19: (a) Desorbed base (orange) stacks on top of a neighbor. Adsorbed bases are shown in cyan. (b) Probability distribution of number of adsorbed bases $P(N_{\text{adsorb}})$ for loop (green) and non-loop (purple) conformations.

sequence independent for ssDNA of this length since all nucleotides can bind to CNT in a similar fashion.

Each base experiences unique steric limitations that depend on its identity and that of its neighbors, leading to a distinct free energy difference ΔF_{adsorb} between the adsorbed (stacked to CNT) and desorbed (detached from CNT surface and stacked to neighboring base) states for each base (Figure 4.20). Here, ΔF_{adsorb} is given by

$$\Delta F_{\text{adsorb}} = -k_{\text{B}}T \ln\left(\frac{1 - P_{\text{desorb}}}{P_{\text{desorb}}}\right) \quad (4.6)$$

Here, a base is considered desorbed if the distance between its center and CNT axis is larger than 8.5 Å. This distance value provides a good estimate of when a base unstacks from CNT. It is evident that guanine has the greater preference for the adsorbed state with $\Delta F_{\text{adsorb}}/k_{\text{B}}T$ ranging from -7.4 for G₁₃ to -3.1 for G₉. In contrast, thymine has a lower preference for adsorption with $\Delta F_{\text{adsorb}}/k_{\text{B}}T$ ranging from -5.7 for T₂ to -1.5 for T₆. These results are consistent with the calculations of Section 4.2 that show that CNT binding for guanine is stronger than for thymine. The magnitude of ΔF_{adsorb} is $\sim k_{\text{B}}T$ for several bases, indicating that thermal fluctuations are sufficient to cause these weakly bound bases to unstack from CNT. These results indicate that desorbed bases are a significant structural feature of DNA-CNT at room temperature. Additionally, because ΔF_{adsorb} varies according to the identity and position of the base on the oligonucleotide backbone, DNA-CNT will contain a sequence specific arrangement of desorbed bases. This has important ramifications for ssDNA-functionalized CNT-FET chemical sensors that display sequence specific sensing responses to gaseous analytes.²¹ Desorbed bases reside the furthest from CNT surface and can readily attract nearby molecules as opposed to

adsorbed bases that may be unable to interact strongly with the environment due to their proximity to CNT. Thus, the sequence specific chemical detection capabilities of these devices may result from interactions between desorbed bases and gaseous analytes.

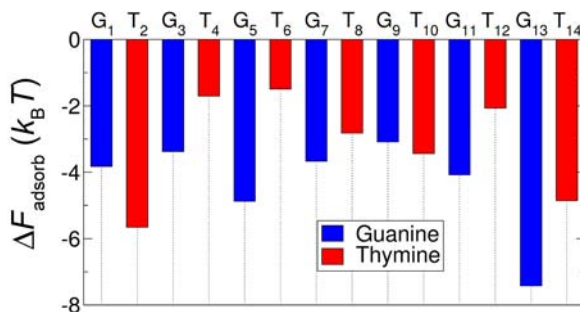


Figure 4.20: Difference in free energy between adsorbed and desorbed states ΔF_{adsorb} for each $(GT)_7$ base. Letters and subscripts on the abscissa refer to the identity and position of each base on the backbone

Because of the small (less than $k_B T$) energy corrugation of the CNT surface, adsorbed nucleotides display considerable configurational disorder. While the collective average orientation of the bases' dipole moments align with CNT axis, which is in agreement with the prior REMD study¹⁴⁸ and optical absorption experiments,⁸⁴ individual bases adopt a multitude of orientations that reflect the local backbone geometry as opposed to nucleotide-specific CNT registration (Figure 4.21). Adsorbed nucleotides are thus free to slide and rotate, indicating that the overall ssDNA configuration is governed by the backbone rather than the nucleotide sequence. We have verified this by performing an additional 55.5 ns REMD simulation of DNA-CNT using a 14 base *random sequence* (ATCGATACGTGACT) and obtained a free energy landscape qualitatively similar to the $(GT)_7$ system. This finding is consistent with AFM measurements of DNA-CNT that show little variation with sequence.⁷² In contrast, *ab initio* computations suggest a more rugged CNT surface with barriers of several $k_B T$.^{81, 137} Increased corrugation could lead to

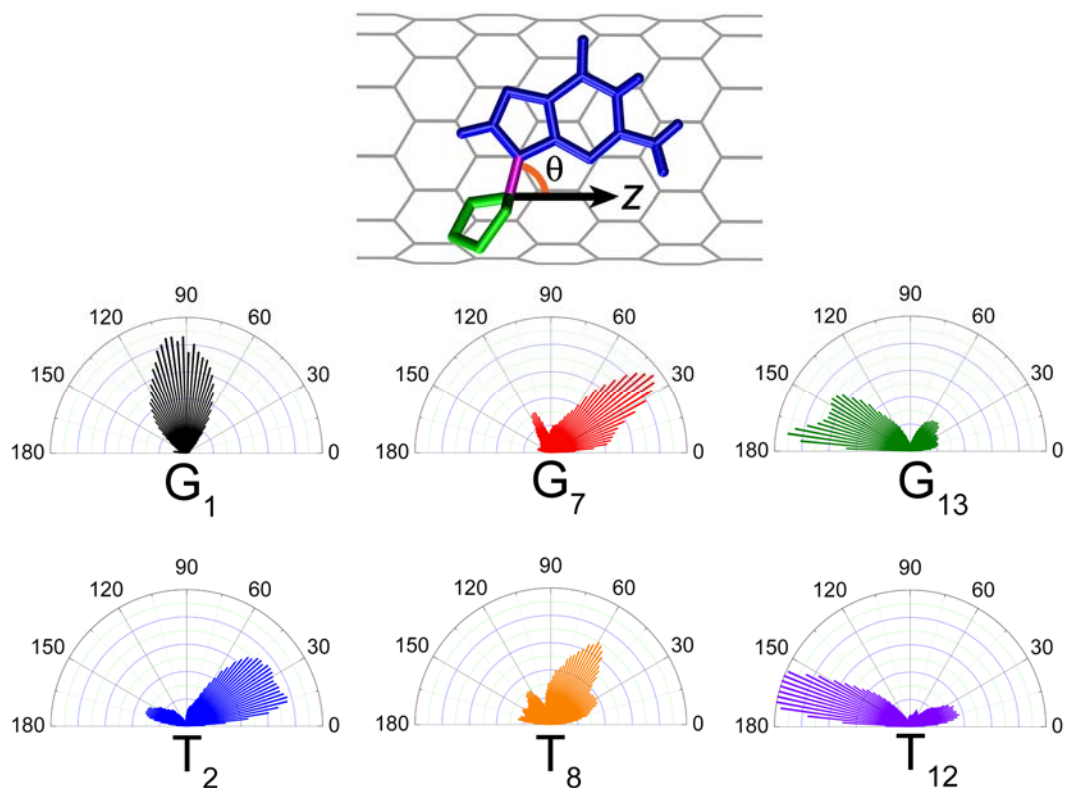


Figure 4.21: Probability density of nucleotide orientation on CNT surface. Orientation is measured by the angle between the glycosidic bond (magenta) and the z-axis (CNT axis). The frequency of each angle is denoted by the length of the line in each polar plot. Each base (several shown below) has a preference for a certain range of orientations

preferred nucleotide orientations, evidence for which has been seen in optical absorption measurements of DNA-CNT,⁸⁴ and sequence dependent DNA-CNT structure observed in other AFM experiments.^{68, 134} Further work is necessary to illuminate these discrepancies and elucidate the effects of sequence on DNA-CNT structure.

Adsorbed bases in DNA-CNT lie flat on the CNT surface and can potentially interact via hydrogen bonding. Base-base, base-backbone and backbone-backbone hydrogen bonds are all found to be present within the (GT)₇-CNT hybrid. A hydrogen bond at the 3' end connects atom O_{3'} of T₁₄ and O_{2P} of G₁₃ and is present in 67% of the conformations. This bond causes T₁₄ to bend towards the backbone (Figure 4.22a). A similar hydrogen bond (not shown) is formed at the 5' end between atom O_{5'} of G₁ and

O_{2P} of G_3 , although this is found for only 39% of the conformations. On average, $(GT)_7$ contains 3.5 intrastrand hydrogen bonds. However, other than those at the 5' and 3' end, intrastrand hydrogen bonds tend to be short lived and occur between random atoms. A few highly ordered conformations are observed that contain upwards of twelve internal hydrogen bonds (Figure 4.22b). However, such conformations are extremely rare. Thus, we conclude that intrastrand hydrogen bonding, similar to interstrand hydrogen bonding, is irrelevant in DNA-CNT.

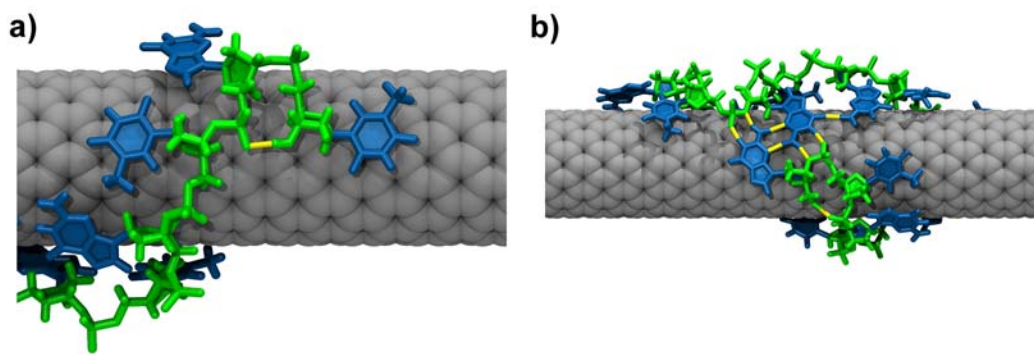


Figure 4.22: (a) Hydrogen bond (yellow) at 3' end resulting in bending of the terminal nucleotide. A similar configuration occurs at the 5' end. (b) Rare, highly ordered conformation containing twelve internal hydrogen bonds (only eight are shown).

The $(GT)_7$ -CNT hybrid free energy landscape varies somewhat with temperature (Figure 4.23). At low temperature ($T \sim 300$ K), the landscape is rugged with many sharp local minima. Several distinct low energy conformations contribute to the hybrid's structure. At higher temperatures ($T \sim 350$ K), the landscape is smoother and dominated by two broad energy minima: a deep, global minimum containing loop conformations and a shallow, local minimum containing left-handed helices. The broad energy minima contain a continuum of isoenergetic conformations, indicating that the hybrid's structural disorder increases with temperature. However, the locations of the free energy minima do

not vary with temperature, indicating that the same ssDNA conformations contribute at all temperatures.

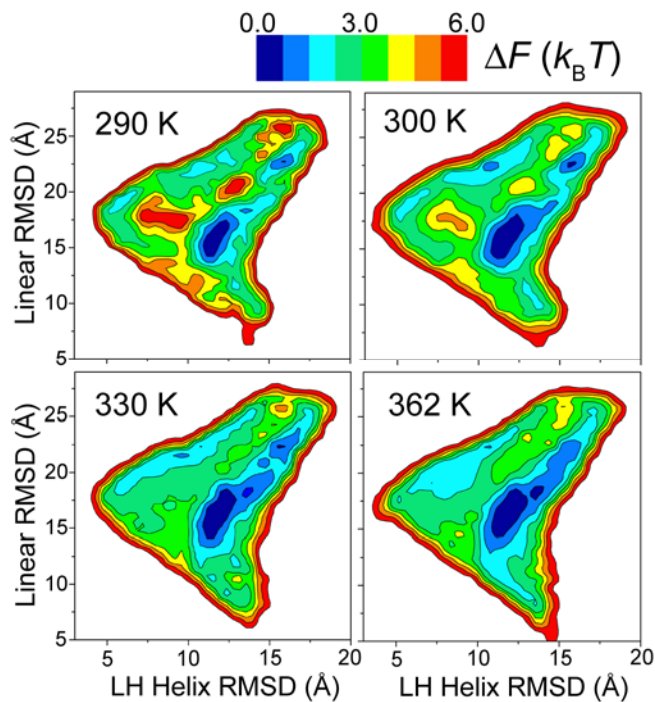


Figure 4.23: The (GT)₇ free energy landscape at various temperatures.

REMD shows that a non-helical loop conformation is the global free energy minimum of the (GT)₇-CNT hybrid. Our previous MD simulations involving the longer 40-60mer oligonucleotides indicate a strong preference for helically-wrapped DNA-CNT. However, these longer sequences may behave qualitatively different from the short sequence employed here. We find that approximately the first and last three nucleotides in adsorbed ssDNA exhibit a more random configuration than those in the interior, “bulk” region of the oligonucleotide. Due to steric interactions of adjacent bases, terminal nucleotides tend to splay out on the CNT surface. As the length of the oligonucleotide increases, end effects will become less important and helical conformations may be a

natural extension of the loop structure. Interestingly, we observe loop-like structures at the 5' end of 40-60mer helices in the prior MD simulations (Figure 4.24).

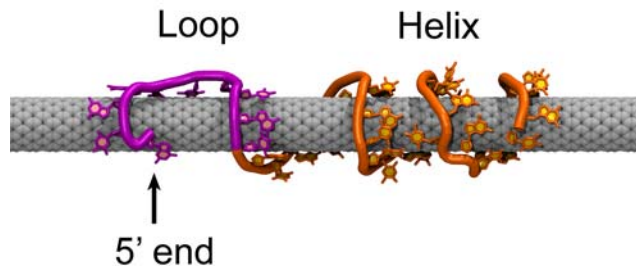


Figure 4.24: Loop is present at the 5' end of helical conformations in longer 40-60mer oligonucleotides. Conformation for (GT)₂₀ shown here.

4.5 Summary

In this chapter, we presented the results of a series of atomistic MD simulations that explored the self-assembly mechanisms, energetic properties and structure of DNA-CNT. The simulations show that ssDNA of any sequence and CNT of arbitrary chirality will spontaneously self-assemble into DNA-CNTs. Self-assembly is driven by the attractive π - π stacking interaction between DNA bases and CNT sidewall; solvent and entropic effects play only a minor role. The strength of this interaction varies for each base and results in CNT binding affinities that follow the trend $G > A > T > C$. However, base-CNT binding is strong with binding free energy magnitudes ranging from 10.3 kcal/mol for G to 6.7 kcal/mol for C.

The flexibility of ssDNA enables a wide range of wrapping conformations about CNTs including right- and left-handed helices and U-shaped loops. At low salt concentration, short ssDNA sequences (less than about twenty bases) prefer to reside in loop conformations that maximize base-CNT stacking, while longer sequences show a greater preference for a 2–8 nm pitch helical wrapping. Helix formation is found to

derive from electrostatic and torsional interactions within the sugar-phosphate backbone, which results in ssDNA wrapping about CNT from the 3' end to the 5' end. Thermal fluctuations and steric limitations introduce structural disorder in these ssDNA configurations. These effects result in a sequence specific arrangement of desorbed bases in DNA-CNT. They also produce disorder in the sugar-phosphate backbone at the 3' and 5' ends which results in the qualitatively different wrapping conformations for short and long ssDNA sequences.

Our MD results show that, while base-CNT stacking provides the main cohesive forces within DNA-CNT, the overall ssDNA conformation is controlled by interactions within the backbone. Effects such as temperature, nucleotide specific CNT registration or hydrogen bonding within a single strand or between multiple strands of adsorbed DNA play only a minor role in DNA-CNT structure. These results rule out the possibility that ssDNA adsorbs to CNT as a dimer. They also show that the overall ssDNA wrapping configuration is not dependent on sequence.

Many experiments have been performed in attempt to measure the helical pitch value and how it is affected by ssDNA sequence, but have produced inconsistent results. Experiments have resulted in pitch values ranging from 3 nm to 18 nm.^{68, 72, 135} Additionally, some experiments have shown that these values vary strongly with sequence, whereas others have not. This information is important for obtaining a full understanding of sequence dependent effects in CNT separation⁶⁸ and in ssDNA-functionalized CNT-FET chemical sensors.²¹ Further experimental work is necessary to illuminate these discrepancies and facilitate a comparison with our computational results.

While it is commonplace to visualize DNA as a helical structure, it is important to note that the conformations found here would not result without the presence of CNT. Upon adsorption, CNT provides a cylindrical template for helical wrapping. The resulting ssDNA conformations in DNA-CNT are drastically different than those found in double stranded DNA or even in ssDNA in aqueous solution. Thus, as research proceeds in combining inorganic nanostructures (e.g. carbon nanotubes, graphene, nanowires, nanocrystals) with organic macromolecules (e.g. proteins, DNA, synthetic polymers), objects that have never been in contact in nature, there is a great possibility of discovering composite materials, such as DNA-CNTs, that possess brand new structural and physical properties.

Chapter 5

Structure and Function of a Nanobiosensor:

A Carbon Nanotube Functionalized with the Coxsackie-Adenovirus Receptor

The results presented in this chapter have appeared in the publication:

- R. R. Johnson, B. J. Rego, A.T. C. Johnson, M. L. Klein. *Journal of Physical Chemistry B*. **34** (2009).

Adenoviruses are a family of viruses that can result in a number of upper respiratory infections or gastroenteritis.¹⁴⁹ The adenovirus consists of a double-stranded DNA genome enclosed in an icosahedral capsid. Emanating from the capsid are spike-like fibers that terminate in globular Knob domains. Adenovirus infection is initiated when Knob binds to the Coxsackie-adenovirus receptor D1 domain (hereafter be referred to as CAR), the extracellular portion of a receptor protein located in the membrane of the host cell (Figure 5.1). Each Knob is composed of three identical, self-complementary proteins arranged in a triskelion.¹⁵⁰ CAR and Knob form a high-affinity complex that is largely stabilized by hydrogen bonds and salt bridges between a Knob monomer and one face of CAR, as determined by crystal structure.¹⁵¹

Recently, highly specific and sensitive biosensors for label-free, electronic detection of the adenovirus have been developed by functionalizing CNT FETs with CAR.²² Fabrication of the device included covalent attachment of CAR to the sidewall of oxidized CNT FETs via diimide activated amidation.⁹² Introducing Knob proteins to these devices resulted in a drastic shift in the on-state current and threshold voltage of the

CNT FET current-gate voltage characteristic. No detectable shifts occurred when other proteins were introduced, indicating that these devices are specific to the adenovirus Knob only.²²

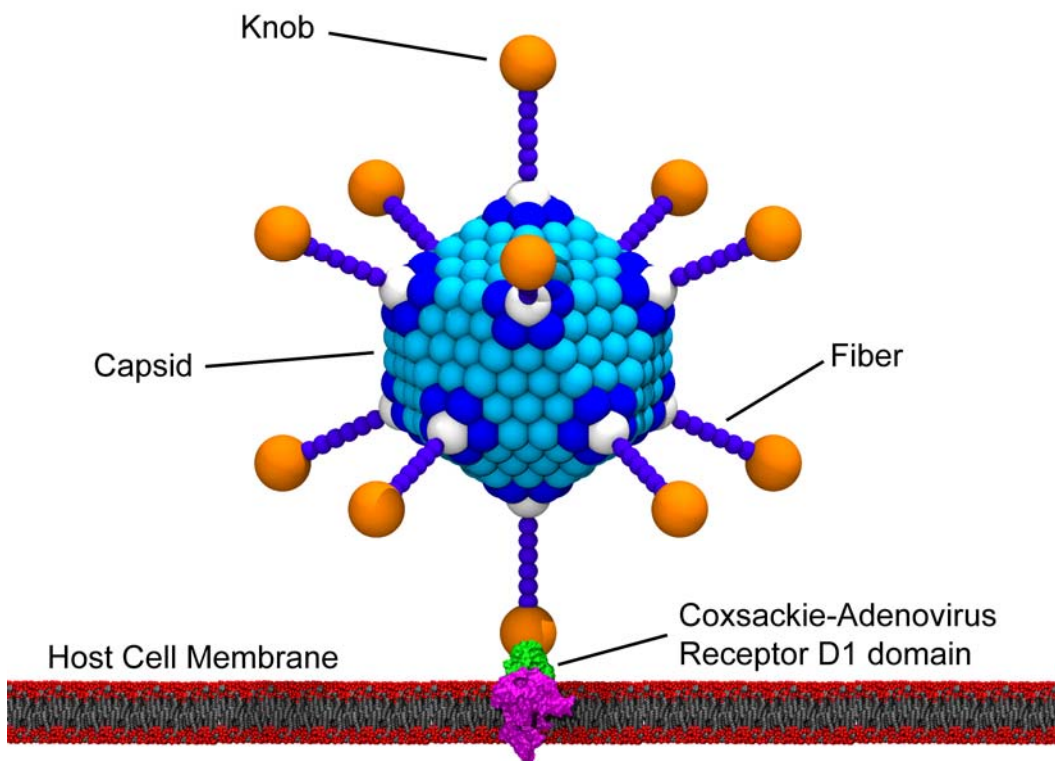


Figure 5.1: Adenovirus binding to the coxsackie-adenovirus D1 domain (CAR). Knob domains are represented as orange spheres.

An axiom of biology is that the three-dimensional structure of a protein determines its *in vivo* functionality.⁶⁴ Experiments with these CAR-CNT biosensors suggest that CAR retains its biologically active form even when chemisorbed to CNT; i.e., the bound CAR remains capable of specifically recognizing Knob. Curiously, however, atomic force microscope (AFM) measurements suggest that the physical dimensions of CAR and CAR-Knob complex (hereafter denoted CAR-Knob) when bound to CNT are significantly smaller than the sizes of these proteins in their native states.²² One

explanation for this observation is that the conformations of these proteins are substantially altered in the presence of CNT. Another possibility is that pressure from the AFM tip produced distortions in the CAR and CAR-Knob structures, thereby temporarily reducing their apparent sizes.

Even though experiments have demonstrated the feasibility of these nano-bio devices to detect adenovirus proteins, the extent of the structural rearrangements, dynamics and specificity of the bound proteins has not been studied in detail. Herein we present results of all-atom classical MD simulations to understand the structure and stability of CAR and CAR-Knob in both aqueous solution alone and when bound to CNT. We find minor structural differences between CAR structures in the bound and unbound state. These differences are due to the suppression of structural fluctuations in the bound CAR by the rigid CNT. Nevertheless, these differences do not significantly alter CAR's ability to bind Knob, suggesting that CAR retains its biologically active form when attached to CNT. The work presented here exemplifies a computational route to gain microscopic understanding of CNT-based biosensing devices.

5.1 Simulation Details

In order to compare CAR and CAR-Knob in the bound (attached to CNT) and unbound (alone in aqueous solution) states, we performed simulations of four systems: CAR_{cnt} , CAR_{nat} , $CAR-Knob_{cnt}$, and $CAR-Knob_{nat}$, as summarized in Table 5.1. Here, the subscripts indicate whether the proteins are simulated in their native conditions (nat), i.e. in aqueous solution alone, or bound to CNT (cnt). Because most CAR-Knob contact occurs along a face of a single Knob monomer, the initial CAR-Knob configuration

employed in our simulations was obtained from crystal structure 1KAC of the Protein Data Bank (PDB).¹⁵¹ This structure contained the coordinates of CAR bound to a Knob monomer. Initial structures for simulations involving CAR only were obtained by removing Knob coordinates from this crystal structure. This is justified because CAR undergoes only minor rearrangements upon binding Knob.¹⁵⁰ To construct CAR-Knob_{cnt}, the crystal structure was aligned to the equilibrated CAR configuration obtained from the CAR_{cnt} simulation. In each system, the structural properties of the proteins were studied at equilibrium. Equilibrium was considered to be reached when the average root mean square deviation (RMSD) of C_α atoms from their initial positions no longer varied with time. We define the best representative protein structure (denoted by curly braces) of a given trajectory as the instantaneous structure with the smallest RMSD from the average. For example, {CAR_{cnt}} is the best representative CAR structure in the bound state. When referring to individual protein residues, we use the standard three letter code for the amino acid followed by its residue number in PDB entry 1KAC.

System	Description	Time (ns)	Box Dimensions (nm ³)	Number of atoms
CAR _{nat}	CAR in native conditions	11.25	6.9×6.9×6.9	10586
CAR _{cnt}	CAR covalently attached to CNT in aqueous solution	11.25	6.1×6.1×8.1	8200
CAR-Knob _{nat}	CAR-Knob in native conditions	10.80	8.8×8.8×8.8	21875
CAR-Knob _{cnt}	CAR-Knob attached to CNT in aqueous solution	10.14	8.5×8.5×8.1	40338

Table 5.1: Summary of simulations performed.

5.2 Protein Simulations Under Native Conditions

In the CAR_{nat} simulation, initially the protein resides in a β -sandwich fold with β -sheets as its primary secondary structural element (Figure 5.2). This structure slowly relaxes over the course of an 11.25 ns trajectory. The β -sheets are the most rigid structural elements of the protein and quickly equilibrate in the first 3.75 ns. The loop regions, on the other hand, are more flexible and require 8.25 ns or more to stabilize. Despite the small structural relaxation, CAR maintains its original folded conformation throughout the simulation. The final RMSD from the crystal structure for C_{α} atoms in β -sheets is only 1.0 Å. Thus, over the simulation timescale, the force fields employed here provide a stable CAR structure that is almost identical to that of PDB 1KAC. Owing in part to a disulfide bridge between Cys43 and Cys122 that connect the bottom and top β -sheet layers (Figure 5.1), the β -sheet C_{α} atoms are held relatively rigid and undergo root mean square fluctuations (RMSF) less than 0.9 Å. However, loop regions are quite mobile with RMSF values up to 2.3 Å. The CAR C-terminus experiences the largest fluctuations with RMSF values greater than 8 Å. These results are in good agreement with NMR studies of CAR's structure and dynamic properties.¹⁵⁰

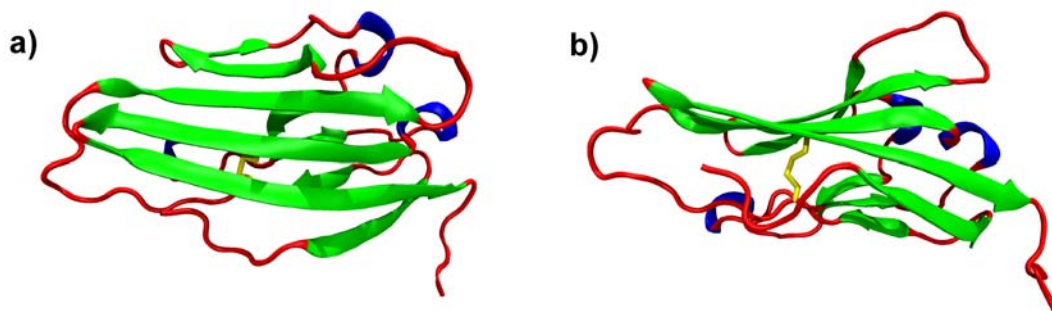


Figure 5.2: Visualization of CAR showing its secondary structural elements: β -sheets (green), 3_{10} helices (blue) and loops (red). A disulfide bridge (yellow) connects the top and bottom β -sheet layers (a) Top view. (b) Side view.

The CAR-Knob_{nat} simulation provides information about the stability and dynamic properties of the protein complex in its native state. In this simulation, CAR-Knob remains in a stable complex with only minor structural rearrangements in the flexible loop regions during equilibration. The RMSD from the crystal structure for β -sheet C_{α} atoms is only 1.4 Å. While many residues at the CAR-Knob interface fluctuate in and out of contact, there are several contacts that are well-preserved throughout the trajectory and compose the main anchor points for the complex (Figure 5.3). The salt bridges between Asp415 and Lys123, Asp415 and Lys125, and Lys429 and Glu58 are maintained throughout the entire trajectory. The hydrogen bonds between Leu426 O and Tyr85 OH and between Lys451 N ζ and Asp83 O are also constant during the simulation. Additionally, there are several transient contacts that may also be important for CAR-Knob stability. Gln487 is in close contact with two polar residues on CAR: Tyr82 and Ser77. These three residues fluctuate continuously throughout the simulation and form temporary hydrogen bonds between Gln487 N ϵ 2 and Tyr82 OH or Gln487 N ϵ 2 and Ser77 O γ . Ser497 changes rotamer conformation to form a short-lived hydrogen bond between Ser497 O γ and Pro128 O. Finally, Glu425 and Lys80 side chains can change conformation to form a salt bridge that is present during 20% of the trajectory. Overall, these results are in good agreement with analysis of the CAR-Knob crystal structure that suggests that the residues on the Knob AB loop region provide the most important contacts in the complex.¹⁵¹

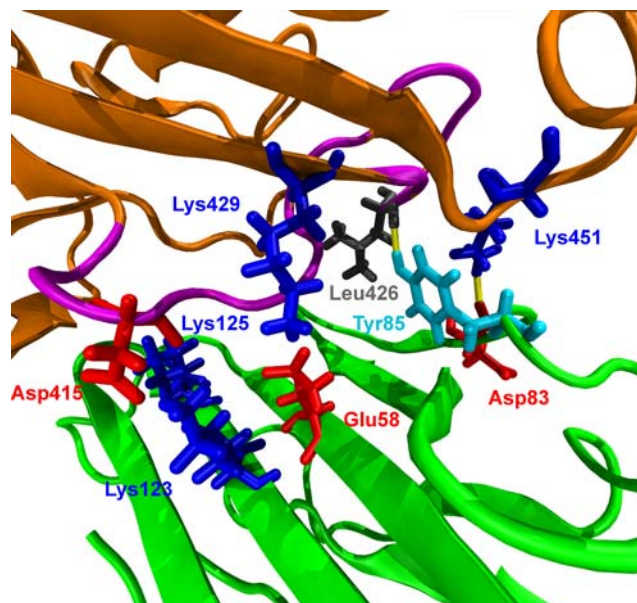


Figure 5.3: Main contacts in CAR-Knob binding site. CAR and Knob are shown in green and orange, respectively. Blue, red, gray and cyan refer to positively charged, negatively charged, hydrophobic and polar residues, respectively. There are three salt bridges: Asp415-Lys123, Asp415-Lys125 and Lys429-Glu58. There are hydrogen bonds (yellow) between Tyr85 and Leu426 backbone and between Lys451 and Asp83 backbone. The Knob AB loop is colored magenta.

5.3 Covalent Attachment of the Coxsackie-Adenovirus Receptor

There were several options for attaching CAR to CNT in a manner that reproduced the experimental conditions. Diimide-activated amidation results in a covalent link between carboxylic acid defects on CNT and amine groups on the protein.⁹² Amine groups are located on the side chains of amino acids Arg, Asn, Gln and Lys and at the N-terminus of the protein backbone. However, at neutral pH values, Lys side chains and the N-terminus are protonated and can not participate in this chemical reaction. Thus, only Arg, Asn and Gln side chains provide viable protein-CNT binding sites. CAR has one Arg, five Asn and five Gln for a total of eleven potential CAR-CNT binding sites (Figure 5.4). However, not all binding sites are equally accessible to carboxylic acid defects on

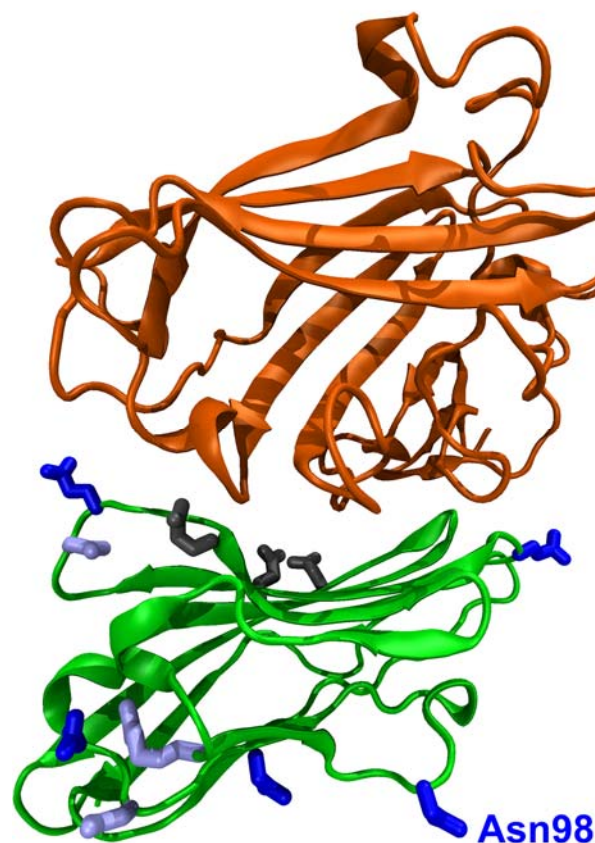


Figure 5.4: CAR (green) can covalently bind to CNT (not shown) via amine groups on Arg, Asn and Gln side chains. Side chains that are oriented towards the CAR interior and are most likely not involved in CAR-CNT binding are shown in light blue. Side chains located on or near the CAR receptor region are shown in black. Side chains on the CAR periphery that are oriented towards the solvent and represent the best CAR-CNT binding sites are shown in blue. Knob is shown in orange.

CNT. For example, the side chains of Asn66, Arg92 and Asn111 (Figure 5.4, light blue residues) are oriented towards the protein interior and are most likely unable to covalently attach to CNT without CAR undergoing major conformational changes. Additionally, other potential CAR-CNT binding sites, such as Gln71, Gln121 and Gln132, are located on or near the CAR receptor region (Figure 5.4, black residues). These sites would orient CAR such that Knob binding would be blocked by CNT. Therefore, we consider the “good” binding sites to be Gln52, Gln67, Asn98, Asn108,

Gln113, (Figure 5.4, blue residues). The side chains of these residues are oriented towards the solvent and are easily accessible to CNT carboxylic acid defects. For our study, we chose to attach CAR to CNT via Asn98. This particular residue was well-exposed and located in the middle of a flexible loop region which can undergo conformational changes without significantly affecting CAR's structure.

5.4 Structure and Function of a Nanobiosensor

Figure 5.5a shows the starting configuration in the CAR_{cnt} simulation. Initially, CAR only contacts CNT through Asn98 which is chemisorbed to a carboxylic acid defect site. However, drawn by van der Waals forces, CAR rotates about Asn98 and physisorbs to CNT over the first 1.5 ns of the trajectory (Figure 5.5b). After this initial rigid body rotation, CAR quickly reaches structural equilibrium: loops equilibrate in 3.5 ns while β -sheets achieve a steady state almost instantaneously. This is significantly faster than the relaxation timescales in CAR_{nat} where a trajectory in excess of 8 ns is required to equilibrate the protein structure. A convenient way to quantify the structural differences in CAR_{cnt} and CAR_{nat} is to compare the best representative protein structures in these two simulations. Here, we assume that $\{CAR_{\text{nat}}\}$ represents the true native state of the protein and, thus, is used as the reference system. The RMSD of C_{α} atoms in $\{CAR_{\text{cnt}}\}$ from those in $\{CAR_{\text{nat}}\}$ is only 2.0 Å, indicating that these two CAR structures are approximately identical. Analyzing the displacement $|\Delta\vec{r}|$ of individual C_{α} atoms shows that most of the structural differences occur in the loop regions (Figure 5.5c). The displacement $|\Delta\vec{r}|$ is given by,

$$|\Delta\bar{r}| = |\{\bar{r}_{\text{cnt}}\} - \{\bar{r}_{\text{nat}}\}| \quad (5.1)$$

Here, $\{\bar{r}_{\text{cnt}}\}$ and $\{\bar{r}_{\text{nat}}\}$ are the positions of C_{α} atoms in $\{\text{CAR}_{\text{cnt}}\}$ and $\{\text{CAR}_{\text{nat}}\}$, respectively. These structural differences can be understood from analysis of the protein's dynamic properties. The RMSF of C_{α} atoms are drastically reduced in the bound state (Figure 5.5d). The most notable changes are observed in the loop regions and C-terminus. While these regions are highly mobile and undergo large fluctuations in CAR_{nat} , such motion is arrested upon CNT binding. The structural differences between the bound and unbound states are not so much due to CNT induced deformation. Rather, CNT acts as a rigid platform that damps relative motion and restricts the configuration space of the bound CAR.

The results of the $\text{CAR-Knob}_{\text{cnt}}$ simulation are qualitatively similar to those in the CAR_{cnt} simulation. Namely, the bound protein structures equilibrate more quickly due to the damping of structural fluctuations caused by CNT. Because of this, again most structural differences between the bound and unbound CAR-Knob occur in the loops. Despite these differences, the RMSD of C_{α} atoms in $\{\text{CAR-Knob}_{\text{cnt}}\}$ from $\{\text{CAR-Knob}_{\text{nat}}\}$ is only 1.6 Å, demonstrating again that CNT binding does not strongly perturb the proteins from their native states. The increased rigidity of $\text{CAR-Knob}_{\text{cnt}}$ seems to have little effect on the binding properties of the complex. The interfacial salt bridges and hydrogen bonds still remain permanent fixtures for the bound CAR-Knob (Figure 5.3). However, because of the decreased mobility, the transient salt bridge between Glu425 and Lys80 and hydrogen bonds among Gln487, Tyr82 and Ser77 present in $\text{CAR-Knob}_{\text{nat}}$ are no longer present in $\text{CAR-Knob}_{\text{cnt}}$. The loss of these transient contacts may reduce the

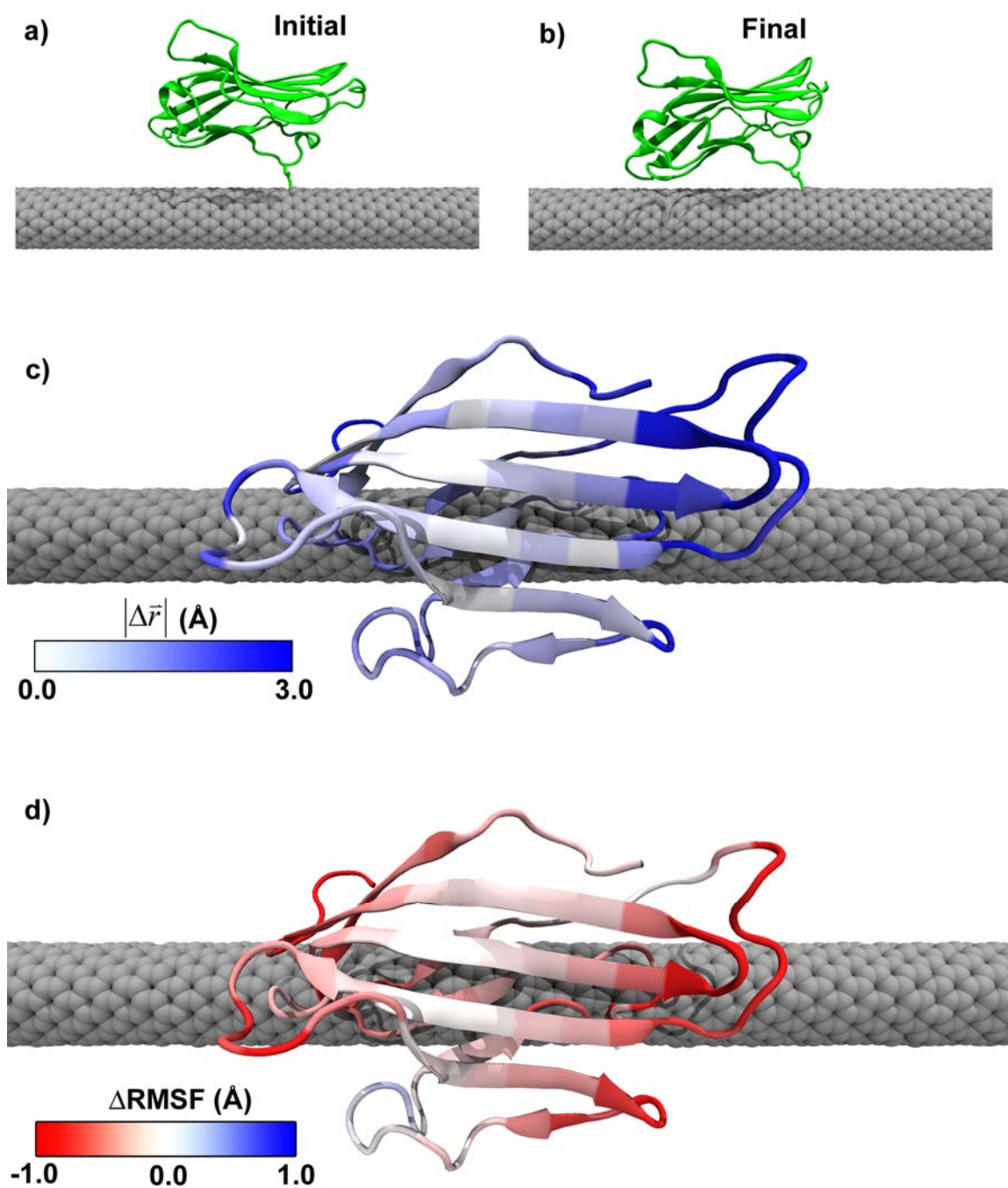


Figure 5.5: (a) Initial CAR structure chemisorbed to CNT via Asn98. (b) Final CAR structure after 11.25 ns MD that includes physisorption to CNT. (c) Bound CAR. Color scale shows the displacement of C_{α} atoms in $\{CAR_{\text{cnt}}\}$ from those in $\{CAR_{\text{nat}}\}$. Most structural differences occur in the flexible loop regions leaving the β -sheets unaffected. (d) Difference in RMSF ($\Delta\text{RMSF} = \text{RMSF}_{\text{cnt}} - \text{RMSF}_{\text{nat}}$) of C_{α} atoms between CAR_{cnt} and CAR_{nat} . Red (blue) indicates regions that are more (less) rigid than those in CAR_{nat} .

magnitude of the binding free energy of the complex, but has no detectable deleterious effect on the stability of CAR-Knob.

Because of the small structural rearrangement, the bound protein complexes retain their native sizes. Thus, the average heights from the CNT surface of CAR and CAR-Knob are 3.2 nm and 6.4 nm, respectively (Figure 5.6), much larger than values of 0.5 nm and 2.5 nm obtained from AFM measurements.²² It seems unlikely that CAR could maintain its biological activity if its height decreases by over 80% when bound to the CNT. The simulation results presented here and the observation that CAR retains its molecular recognition functionality when bound to CNT²² indicates that pressure-induced distortions from the AFM tip is the most likely cause for the apparent reduced dimensions of the bound protein complexes.

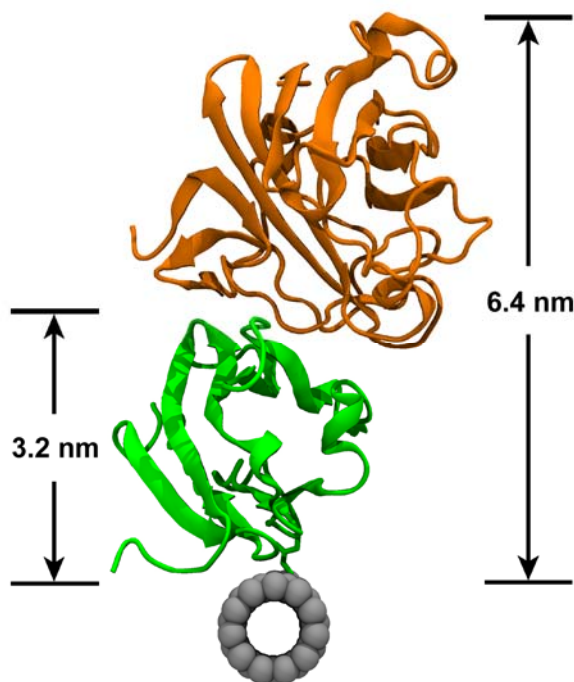


Figure 5.6: Average heights of bound CAR-Knob from CNT surface over 10.14 ns MD.

5.5 Discussion

There is evidence that the degree to which CNT binding affects a protein's structure and function largely depends on the particular protein in question; certain proteins can denature and lose biological activity when in contact with CNT while others do not.¹⁵² Thus, while our simulations show that CAR is not significantly perturbed by CNT, this result is most likely not the case for general proteins. However, we can draw the following general conclusion on the nature of protein-CNT interactions from this study. Because the reduction of structural fluctuations in CAR is a direct consequence of the mechanical properties of CNT, similar damping effects should be expected for general proteins as well. These effects may be manifested in receptor proteins that require large conformational changes to accommodate protein binding. In such systems where structural flexibility is required, protein binding may be slowed or completely abrogated when the receptor is attached to CNT. There has also been significant development of similar biosensing devices using silicon nanowires in place of CNT.¹⁵³ Such structural damping effects would also be expected to occur for proteins in contact with silicon nanowires or any other solid nanostructure. Thus, solid state biosensing devices, similar to the one modeled here, may be most compatible with biomolecules that are relatively rigid and do not require large structural rearrangement.

Performing simulations analogous to those reported here may provide a computational means of identifying suitable proteins for use in similar nano-biosensing devices. For some proteins, amine containing residues may not be conveniently located on the surface and can not readily bind to CNT carboxylic acid defects. MD can be employed to examine the protein conformational changes required to accommodate CNT

binding to such inaccessible amine groups and predict how such structural changes ultimately affect protein functionality. Alternatively, MD could be used to rationalize mutations to generate CNT compatible proteins that contain easily accessible protein-CNT binding sites. Finally, MD enables one to assess the functionality of such nanobiosensors through inspection of the structural stability and mobility of proteins bound to CNT.

5.6 Summary

MD was employed to gain microscopic understanding of the structural and dynamic properties of CNT-based biosensors used for detecting the binding of the adenovirus Knob domain to its cellular receptor CAR. We find that structural fluctuations in CAR are significantly damped when the protein is bound to CNT. However, despite the changes in the internal dynamics of CAR, its overall structure experiences only minor deformation from its native structure. Thus, AFM measurements showing a drastically reduced CAR height when attached to CNT was most likely due to artifacts of the experiment. The nature of CAR-Knob binding when attached to CNT is largely identical to that under native conditions. These findings are consistent with experiments that demonstrate that CAR retains its biological activity when attached to CNT and is a suitable protein for CNT-based biosensing applications. The work presented here demonstrates the value of MD simulations in aiding and understanding nanobiosensor design and function.

Chapter 6

Summary and Conclusion

Learning how to navigate and utilize the interface between nanotechnology and biology is one of the most important challenges facing nanoscience today. Connecting the mechanical and electrical robustness of solid state nanomaterials with the functionality of cellular molecular machinery promises to yield far reaching societal and technological impacts. Biopolymer-carbon nanotube hybrids (Bio-CNT) are exciting nanoscale materials of current interest that lie at this nano-bio interface. Bio-CNTs combine single-walled carbon nanotubes (CNT), one of the most heralded inorganic nanomaterials, with the ubiquitous, essential biological macromolecules DNA or proteins. Even though these objects have never encountered one another in nature, they readily self-assemble into a hybrid that contains a unique set of properties with the possibility to transfigure current technology. To realize this potential, a detailed understanding of the mechanics of Bio-CNT is needed. Up until now, atomic force microscopy (AFM) has been the most popular (experimental) tool to study the nature of this nanomaterial. These measurements have produced a contrasting collection of results that have arguably produced more questions than answers.^{21, 22, 68, 72, 134} Several AFM images of ssDNA-carbon nanotube hybrids (DNA-CNT) have suggested adsorbed ssDNA assumes a helical wrapping of various pitch that may^{68, 134} or may not⁷² be sequence specific. However, others have indicated amorphous ssDNA coating of CNT.²¹ Additionally, AFM studies of protein-CNT hybrids composed of the Coxsackie-adenovirus receptor (CAR) measure a CAR height that is 80% smaller than the expected value. This is especially curious considering that CAR

seemingly retains its biological activity when bound to CNT.²² The time is ripe for the use of additional methods in attempt to verify and organize the existing data set on Bio-CNT.

This thesis has presented the results of a series of molecular dynamics (MD) simulations that studied the self-assembly mechanisms, dynamics, energetics, structure and function of biopolymer-carbon nanotube hybrids (Bio-CNT). These computations provided an atomic-resolution glimpse into the nature of Bio-CNT and complement the prior experimental work. In contrast to experiment, computation gives the researcher complete control over the microscopic character of his or her nanoscale system. In an MD simulation, the researcher can precisely alter the hydration level, salt concentration, temperature, pressure and even the strength of specific (electrostatic, van der Waals, bond stiffness, etc.) interactions between atoms. Because nanoscience deals with imperceptibly small objects, much experimental effort goes into purifying and identifying the nanoscale materials within their samples before measurements can even be performed. These difficulties make a precise tool such as MD all the more valuable to this field. However, until now, MD has been underutilized in Bio-CNT research. The thesis can be divided into two parts. The first part offers the results of a series of MD simulations of DNA-CNT. The second part presents a detailed study of a potential nanobiosensor composed of the Coxsackie-adenovirus receptor attached to CNT.

The MD simulations presented here show that, over timescales of a few nanoseconds, CNT induces a conformational change in hydrated ssDNA whereby bases rotate about the backbone and stack on the CNT sidewall. Thus, DNA-CNT self-assembly is driven by attractive interactions between DNA bases and the CNT sidewall, in agreement with

experiment⁸⁴ and *ab initio*⁸¹ and molecular mechanics¹⁷ calculations. These interactions exist for all four bases and thus, arbitrary ssDNA sequences will readily form DNA-CNT as verified by the simulations here. This agrees with experiments that show all DNA sequences can effectively disperse CNT bundles in solution with similar performance.¹⁷

Using thermodynamic integration (TI), the nature of base-CNT interactions is shown to result primarily from the π - π stacking interaction with almost negligible contributions from entropic and solvent mediated effects. The free energy difference between the bound and unbound state of the base-CNT system ΔF_{bind} follows the trend $\Delta F_{\text{bind}}^{\text{G}} < \Delta F_{\text{bind}}^{\text{A}} < \Delta F_{\text{bind}}^{\text{T}} < \Delta F_{\text{bind}}^{\text{C}}$ which agrees with the adsorption isotherms of bases on graphite.¹⁴¹ The strength of base-CNT binding is quite strong with $|\Delta F_{\text{bind}}| \sim 17 k_{\text{B}} T$ for each base.

MD simulations that tested the feasibility of several proposed structures for poly GT-CNT hybrids ruled out the possibility of poly GT dimers on the surface of CNT. Additionally, they showed that helical wrapping about CNT minimizes electrostatic and torsional backbone interactions and thus, is a highly favorable conformation for long ssDNA strands. This is consistent with interpretations of several AFM images. However, MD predicts a pitch less than 10 nm which contrasts with several of the AFM studies, but agrees with a recent STM measurement.¹³⁵

A study of the full ensemble of configurations for a short oligonucleotide adsorbed to CNT with replica exchange molecular dynamics (REMD) shows that DNA-CNT is disordered at room temperature. ssDNA bound to CNT contains a sequence-dependent arrangement of desorbed bases due to a combination of steric limitations and thermal fluctuations. This study also shows that temperature, intrastrand hydrogen bonding and

sequence have no effect on DNA-CNT structure. Thus, while self-assembly is driven by base-CNT interactions, the overall ssDNA conformation is determined by the sequence-nonspecific backbone.

MD shows that CAR maintains its native conformation and thus, biological functionality even when attached to CNT. This implies that pressure from the AFM tip was the cause for the apparent reduced size of CAR when attached to CNT-FETs. Moreover, the simulations show that the rigid CNT damps structural fluctuations in bound proteins. Thus, biological sensors based on proteins bound to solid state nanostructures, such as CNTs or silicon nanowires, may yield the highest performance for fairly stiff receptor proteins that do not require large conformational changes to accommodate ligand binding.

6.1 Future Research

This thesis has attempted to provide a deeper understanding of the fundamental nature of Bio-CNT. However, Bio-CNT research is still in its infancy and there exist many possibilities for future computational projects.

Now that this thesis has provided insights into the structural properties of DNA-CNT structure, further study on the functionality of this nanomaterial may proceed. One of the most interesting applications of this material is as a chemical sensor (see Section 2.3). DNA-functionalized CNT-FETs display a characteristic electronic response to a variety of gaseous analyte that varies according to the sequence of adsorbed ssDNA. At the present moment, neither the mechanism for this response nor the role of ssDNA in device performance is understood.

Adsorbates can affect the electronic properties of CNT-FETs in a number of ways. The first is through the direct transfer of charge from the molecule to the CNT. The second is by altering the properties of the CNT-metal contact, thereby affecting the ability of charge carriers from the leads to enter or leave the CNT. The third is through chemical gating. In this mechanism, polar adsorbates produce a net electrostatic potential at the CNT surface and thereby produce effects similar to applying a voltage V_G to the FET back-gate. The presence of ssDNA could enhance any these effects by increasing the analyte-CNT binding affinity. Before ssDNA functionalization, the device consists of a bare CNT that is hydrophobic and presumably does not interact strongly with charged or polar molecular species. However, ssDNA is anionic and exhibits strong electrostatic interactions with such molecules. Thus, the overall concentration of molecular analytes in the vicinity of the device may be dramatically increased after ssDNA application. Additionally, the REMD simulation presented in Section 4.4 shows that DNA-CNT contains a sequence-specific arrangement of desorbed bases. These desorbed bases reside the furthest from the CNT surface and could potentially experience stronger interactions with analytes than those that remain firmly attached to CNT. If such interactions are base-specific, this may explain the observed sequence-dependent chemical sensing capabilities of the device.

The prior hypotheses could be tested and/or modified by performing two MD simulations of a small segment of a CNT resting on top of a silicon dioxide substrate. Under (experimental) ambient conditions, a nanoscale film of water will be present on the substrate and thus, should be included in the simulation.¹⁵⁴ In one simulation, the CNT should remain bare while in the other, an ssDNA coating should be applied. Various

analyte molecules could then be introduced into the system. The simulation could help determine how strongly the molecules interact with CNT with and without ssDNA functionalization. Moreover, one could shed light on the correlation between desorbed bases and the sensing response of the device by calculating distributions of the average position of the analytes relative to desorbed bases. The effects of chemical gating could also be estimated from the resulting MD trajectories. One could compute the average change in the net electrostatic potential ΔV_{CNT} at the CNT surface upon exposure to analytes and see if it is consistent with the electronic transport changes that occur in the experiments. Because undoped CNT-FETs exhibit p-type conduction at low bias voltages,¹⁵⁵ $\Delta V_{\text{CNT}} > 0$ would deplete hole density and conduction would be expected to decrease upon analyte exposure, whereas $\Delta V_{\text{CNT}} < 0$ would increase hole density and enhance current in the device.

The sensing mechanism of biosensors consisting of protein-functionalized CNT-FETs (Section 2.3) is also not fully understood. The previously described electrostatics calculations could be useful to determine the importance of chemical gating in the function of these devices as well.

MD could also be used to test the stability and structure of a series of small, synthetic proteins that have been designed to selectively bind and solubilize CNT.¹⁵⁶ Almost all of these proteins were rich in tryptophan residues, for reasons that remain unclear. While the structure of these proteins is not known for certain, there is indication that they assume an α -helix at least when bound to CNT (measurements suggest the proteins assume a random coil when in solution alone). More interestingly, there is experimental evidence that amphiphilic proteins are induced into α -helices by CNT.¹⁵⁷ Because of the small size

and simple nature of these synthetic proteins, MD has the potential to provide much supplementary information about this system. A series of simulations could be conducted that monitor the stability of α -helical structures of these proteins when bound to CNT and compare these results to the structures when in solution alone. With this method it may also be possible to reveal the importance of tryptophan residues for CNT binding.

It is also of interest to understand how certain proteins change conformation when adsorbing to CNT. Just as with DNA-CNT, REMD is the best method to study the entirety of possible protein-CNT structures. However, because this method runs multiple MD simulations in parallel, REMD would be computationally unfeasible for all but the smallest proteins. However, it would be possible to perform REMD on small proteins that are known to adopt a regular α -helix or β -sheet in order to rationalize how these secondary structural elements may be affected in larger proteins upon CNT binding.

There remain many unanswered questions about the fundamental properties of Bio-CNT and thus, a vast potential for further computational study. Bio-CNT's novelty and versatility ensures that research with this hybrid nanomaterial will continue to provide fruitful ground for interesting computational and experimental studies alike.

Appendix A: Nanotube Builder 1.0 TCL Script

```
#Generate a single wall carbon nanotube

#Author: Robert R. Johnson (robertjo@physics.upenn.edu)

package provide nanotube 1.0

namespace eval ::Nanotube:: {
    variable w

    variable l
    variable n
    variable m
}

proc nanotube { args } {return [eval ::Nanotube::nanotube_core $args] }

#Instructions on how to use plugin
proc ::Nanotube::nanotube_usage { } {
    puts "Usage: nanotube -l <length> -n <n> -m <m>"
    puts "  <length> is length of nanotube in nanometers"
    puts "  <n> and <m> are the chiral indices of the nanotube"
    error ""
}

#Process to generate nanotube coordinates
proc ::Nanotube::nanotube_core { args } {
    # Check if proper number of arguments was given
    set n_args [llength $args]
    if { [expr fmod($n_args,2)] } { nanotube_usage }
    if { $n_args < 6 && $n_args > 8 } { nanotube_usage }

    for { set i 0} {$i < $n_args} {incr i 2} {
        set key [lindex $args $i]
        set val [lindex $args [expr $i + 1]]
        set cmdline($key) $val
    }

    # Check if mandatory options are defined
    if { ![info exists cmdline(-l)] \
        || ![info exists cmdline(-n)] \
        || ![info exists cmdline(-m)] } {
        nanotube_usage
    }

    #Set nanotube parameters
    set length $cmdline(-l)
    set n $cmdline(-n)
    set m $cmdline(-m)
    set a 1.418
    set pi 3.14159265358979323846
}
```

```

#Check that input is reasonable
if { $n < 0 || $m < 0 || int($n) != $n || int($m) != $m } {error "n
and m must be positive integers"}
if {$m==0 && $n==0} {error "n and m can not both be zero"}
if {$length <= 0} {error "Nanotube length must be a positive value"}

#Calculate greatest common divisor d_R
set num1 [expr 2*$m + $n]
set num2 [expr 2*$n + $m]
while { $num1 != $num2 } {

    if { $num1 > $num2 } {
        set num1 [expr $num1 - $num2]
    } else {
        set num2 [expr $num2 - $num1]
    }

}
set d_R $num1

#Compute geometric properties
set C [expr $a*sqrt(3*($n*$n + $m*$n + $m*$m))]
set R [expr $C/(2*$pi)]
set L_cell [expr sqrt(3)*$C/$d_R]

#Number of unit cells
set N_cell [expr ceil($length*10/$L_cell)]

#Index min/max
set pmin 0
set pmax [expr int(ceil($n + ($n + 2*$m)/$d_R))]
set qmin [expr int(floor(-(2*$n + $m)/$d_R))]
set qmax $m
set i 0

#Generate unit cell coordinates
for {set q $qmin} {$q <= $qmax} {incr q} {
    for {set p $pmin} {$p <= $pmax} {incr p} {

        #First basis atom
        set xprime1 [expr 3*$a*$a*($p*(2*$n + $m) + $q*($n +
2*$m))/(2*$C)]
        set yprime1 [expr 3*sqrt(3)*$a*$a*($p*$m - $q*$n)/(2*$C)]

        #Second basis atom
        set xprime2 [expr $xprime1 + 3*$a*$a*($n + $m)/(2*$C)]
        set yprime2 [expr $yprime1 - $a*$a*sqrt(3)*($n - $m)/(2*$C)]

        set phi1 [expr $xprime1/$R]
        set phi2 [expr $xprime2/$R]

        set x1 [expr $R*cos($phi1)]
        set x2 [expr $R*cos($phi2)]
        set y1 [expr $R*sin($phi1)]
        set y2 [expr $R*sin($phi2)]
        set z1 $yprime1

```

```

set z2 $yprime2

#Store coordinates of unit cell in an array
# 0 <= xprime1 < C and 0 <= yprime1 < L_cell
if {0 <= $xprime1 \
&& $p*(2*$n + $m) + $q*($n + 2*$m) < 2*($n*$n + $n*$m + $m*$m) \
&& 0 <= $yprime1 \
&& $d_R*($p*$m - $q*$n) < 2*($n*$n + $n*$m + $m*$m) } {
  set coord1($i,0) $x1
  set coord1($i,1) $y1
  set coord1($i,2) $z1

  set coord2($i,0) $x2
  set coord2($i,1) $y2
  set coord2($i,2) $z2

  incr i
}
}
}

set num_atom $i

set k 0

#Open file for writing
set file [open temp_nanotube.pdb "w"]

#Generate PDB header
puts $file [format "CRYST1 100.000 100.000 %7.3f 90.00 90.00
90.00 P 1" [expr $N_cell*$L_cell]]

#Generate nanotube
for {set j 0} { $j < $N_cell } {incr j} {
  for {set i 0} { $i < $num_atom } {incr i} {

    set x1 $coord1($i,0)
    set y1 $coord1($i,1)
    set z1 [expr $coord1($i,2) + $j*$L_cell]

    set x2 $coord2($i,0)
    set y2 $coord2($i,1)
    set z2 [expr $coord2($i,2) + $j*$L_cell]

    puts $file [format "%-6s%5d %-4s
%3s%5d%12.3f%8.3f%8.3f%6.2f%6.2f" "ATOM" [incr k] "C" "CNT" $k $x1 $y1
$z1 1.00 0.00]
    puts $file [format "%-6s%5d %-4s
%3s%5d%12.3f%8.3f%8.3f%6.2f%6.2f" "ATOM" [incr k] "C" "CNT" $k $x2 $y2
$z2 1.00 0.00]
  }
}

#Close file
close $file

#Load molecule

```

```

mol new temp_nanotube.pdb
mol rename top Nanotube

#Delete file
file delete temp_nanotube.pdb
}

#GUI
proc ::Nanotube::nanotube_gui {} {
    variable w
    variable l
    variable n
    variable m

    if { [wininfo exists .nanotube] } {
        wm deiconify $w
        return
    }

    set w [toplevel ".nanotube"]
    wm title $w "Nanotube"
    wm resizable $w yes yes
    set row 0

    set ::Nanotube::l 5
    set ::Nanotube::n 5
    set ::Nanotube::m 10

    #Add a menubar
    frame $w.menubar -relief raised -bd 2 -padx 10
    grid $w.menubar -padx 1 -column 0 -columnspan 4 -row $row -sticky
ew
    menubutton $w.menubar.help -text "Help" -underline 0 \
    -menu $w.menubar.help.menu
    $w.menubar.help config -width 5
    pack $w.menubar.help -side right
    menu $w.menubar.help.menu -tearoff no
    $w.menubar.help.menu add command -label "About" \
    -command {tk_messageBox -type ok -title "About Nanotube Builder" \
    -message "Nanotube building tool."}
    $w.menubar.help.menu add command -label "Help..." \
    -command "vmd_open_url [string trimright [vmdinfo www]
/]plugins/nanotube"
    incr row

    grid [label $w.nlabel -text "Nanotube chiral index n: "] \
    -row $row -column 0 -columnspan 3 -sticky w
    grid [entry $w.n -width 7 -textvariable ::Nanotube::n] -row $row -
column 3 -columnspan 1 -sticky ew
    incr row

    grid [label $w.mlabel -text "Nanotube chiral index m: "] \
    -row $row -column 0 -columnspan 3 -sticky w
    grid [entry $w.m -width 7 -textvariable ::Nanotube::m] -row $row -
column 3 -columnspan 1 -sticky ew
    incr row

```

```

grid [label $w.llabel -text "Nanotube length (nm): "]\
  -row $row -column 0 -columnspan 3 -sticky w
grid [entry $w.l -width 7 -textvariable ::Nanotube::l] -row $row -
column 3 -columnspan 1 -sticky ew
incr row

grid [button $w.gobutton -text "Generate Nanotube" \
  -command [namespace code {
    puts "nanotube_core -l $l -n $n -m $m"
    nanotube_core -l "$l" -n "$n" -m "$m"
  } ]] -row $row -column 0 -columnspan 4 -sticky nsew
}

proc nanotube_tk {} {
  ::Nanotube::nanotube_gui
  return $::Nanotube::w
}

```

Bibliography

1. Tans, S. J.; Verschueren, A. R. M.; Dekker, C. *Nature* **1998**, 393, (6680), 49-52.
2. Kong, J.; Franklin, N. R.; Zhou, C. W.; Chapline, M. G.; Peng, S.; Cho, K. J.; Dai, H. J. *Science* **2000**, 287, (5453), 622-625.
3. Fuhrer, M. S.; Kim, B. M.; Durkop, T.; Brintlinger, T. *Nano Lett.* **2002**, 2, (7), 755-759.
4. Breaker, R. R. *Nature* **2004**, 432, (7019), 838-845.
5. Rothmund, P. W. K. *Nature* **2006**, 440, (7082), 297-302.
6. Keren, K.; Berman, R. S.; Buchstab, E.; Sivan, U.; Braun, E. *Science* **2003**, 302, (5649), 1380-1382.
7. Seeman, N. C. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, 27, 225-248.
8. Zheng, J.; Constantinou, P. E.; Micheel, C.; Alivisatos, A. P.; Kiehl, R. A.; Seeman, N. C. *Nano Lett.* **2006**, 6, (7), 1502-1504.
9. Jackel, C.; Kast, P.; Hilvert, D. *Annual Review of Biophysics* **2008**, 37, (1), 153-173.
10. Butterfoss, G. L.; Kuhlman, B. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, 35, (1), 49-65.
11. Kang, S.-g.; Saven, J. G. *Curr. Opin. Chem. Biol.* **2007**, 11, (3), 329-334.
12. Malakauskas, S. M.; Mayo, S. L. *Nat Struct Mol Biol* **1998**, 5, (6), 470-475.
13. Chevalier, B. S.; Kortemme, T.; Chadsey, M. S.; Baker, D.; Monnat, R. J.; Stoddard, B. L. **2002**, 10, (4), 895-905.
14. Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. *Science* **2003**, 302, (5649), 1364-1368.

15. Slovic, A. M.; Kono, H.; Lear, J. D.; Saven, J. G.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, (7), 1828-1833.
16. Nepal, D.; Balasubramanian, S.; Simonian, A. L.; Davis, V. A. *Nano Lett.* **2008**, 8, (7), 1896-1901.
17. Zheng, M.; Jagota, A.; Semke, E. D.; Diner, B. A.; McLean, R. S.; Lustig, S. R.; Richardson, R. E.; Tassi, N. G. *Nat. Mater.* **2003**, 2, (5), 338-342.
18. Karajanagi, S. S.; Yang, H. C.; Asuri, P.; Sellitto, E.; Dordick, J. S.; Kane, R. S. *Langmuir* **2006**, 22, (4), 1392-1395.
19. Nepal, D.; Geckeler, K. E. *Small* **2007**, 3, (7), 1259-1265.
20. Kam, N. W. S.; Liu, Z. A.; Dai, H. J. *Angew. Chem., Int. Ed.* **2006**, 45, (4), 577-581.
21. Staii, C.; Chen, M.; Gelperin, A.; Johnson, A. T. *Nano Lett.* **2005**, 5, (9), 1774-1778.
22. Zhang, Y. B.; Kanungo, M.; Ho, A. J.; Freimuth, P.; van der Lelie, D.; Chen, M.; Khamis, S. M.; Datta, S. S.; Johnson, A. T. C.; Misewich, J. A.; Wong, S. S. *Nano Lett.* **2007**, 7, (10), 3086-3091.
23. Chen, R. J.; Bangsaruntip, S.; Drouvalakis, K. A.; Wong Shi Kam, N.; Shim, M.; Li, Y.; Kim, W.; Utz, P. J.; Dai, H. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, (9), 4984-4989.
24. Jeng, E. S.; Moll, A. E.; Roy, A. C.; Gastala, J. B.; Strano, M. S. *Nano Lett.* **2006**, 6, (3), 371-375.
25. Johnson, R. R.; Johnson, A. T. C.; Klein, M. L. *Nano Lett.* **2008**, 8, (1), 69-75.

26. Johnson, A. T. C.; Staii, C.; Chen, M.; Khamis, S.; Johnson, R.; Klein, M. L.; Gelperin, A. *Semiconductor Science and Technology* **2006**, 21, (11), S17-S21.
27. Johnson, R. R.; Kohlmeyer, A.; Johnson, A. T. C.; Klein, M. L. *Nano Lett.* **2009**, 9, (2), 537-541.
28. Johnson, R. R.; Rego, B. J.; Johnson, A. T. C.; Klein, M. L. *J. Phys. Chem. B* **2009**, 34, (0).
29. Saito, R.; Dresselhaus, G.; Dresselhaus, M. S., *Physical properties of carbon nanotubes*. Imperial College Press: London, 1998.
30. Iijima, S. *Nature* **1991**, 354, (6348), 56-58.
31. Hong, B. H.; Lee, J. Y.; Beetz, T.; Zhu, Y. M.; Kim, P.; Kim, K. S. *J. Am. Chem. Soc.* **2005**, 127, (44), 15336-15337.
32. Li, F.; Cheng, H. M.; Bai, S.; Su, G.; Dresselhaus, M. S. *Appl. Phys. Lett.* **2000**, 77, (20), 3161-3163.
33. Yu, M. F.; Files, B. S.; Arepalli, S.; Ruoff, R. S. *Phys. Rev. Lett.* **2000**, 84, (24), 5552-5555.
34. Demczyk, B. G.; Wang, Y. M.; Cumings, J.; Hetman, M.; Han, W.; Zettl, A.; Ritchie, R. O. *Materials Science and Engineering a-Structural Materials Properties Microstructure and Processing* **2002**, 334, (1-2), 173-178.
35. Peigney, A. *Nat. Mater.* **2003**, 2, (1), 15-16.
36. Esawi, A. M. K.; Farag, M. M. *Materials & Design* **2007**, 28, (9), 2394-2401.
37. Easton Sports. www.eastonsports.com
38. de Heer, W. A.; Châtelain, A.; Ugarte, D. *Science* **1995**, 270, (5239), 1179-1180.

39. Cao, G.; Lee, Y. Z.; Peng, R.; Liu, Z.; Rajaram, R.; Calderon-Colon, X.; An, L.; Wang, P.; Phan, T.; Sultana, S.; Lalush, D. S.; Lu, J. P.; Zhou, O. *Phys Med Biol* **2009**, *54*, (8), 2323-2340.
40. Kong, J.; Chapline, M. G.; Dai, H. J. *Advanced Materials* **2001**, *13*, (18), 1384-1386.
41. Collins, P. G.; Bradley, K.; Ishigami, M.; Zettl, A. *Science* **2000**, *287*, (5459), 1801-1804.
42. Harris, P. J. F., *Carbon Nanotube Science: Synthesis, Properties and Applications*. 1st ed.; Cambridge University Press: New York, 2009.
43. Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, (1), 33-38.
44. Ashcroft, N. W.; Mermin, N. D., *Solid state physics*. Holt: New York,, 1976; p xxi, 826 p.
45. Wallace, P. R. *Physical Review* **1947**, *71*, (9), 622-634.
46. Ebbesen, T. W.; Ajayan, P. M. *Nature* **1992**, *358*, (6383), 220-222.
47. Guo, T.; Nikolaev, P.; Thess, A.; Colbert, D. T.; Smalley, R. E. *Chem. Phys. Lett.* **1995**, *243*, (1-2), 49-54.
48. Endo, M.; Takeuchi, K.; Igarashi, S.; Kobori, K.; Shiraishi, M.; Kroto, H. W. *J. Phys. Chem. Solids* **1993**, *54*, (12), 1841-1848.
49. Terrones, M. *Annual Review of Materials Research* **2003**, *33*, 419-501.
50. Baughman, R. H.; Zakhidov, A. A.; de Heer, W. A. *Science* **2002**, *297*, (5582), 787-792.
51. Hersam, M. C. *Nat Nano* **2008**, *3*, (7), 387-394.

52. Liu, J.; Rinzler, A. G.; Dai, H.; Hafner, J. H.; Bradley, R. K.; Boul, P. J.; Lu, A.; Iverson, T.; Shelimov, K.; Huffman, C. B.; Rodriguez-Macias, F.; Shon, Y.-S.; Lee, T. R.; Colbert, D. T.; Smalley, R. E. *Science* **1998**, 280, (5367), 1253.
53. Doorn, S. K.; Fields, R. E.; Hu, H.; Hamon, M. A.; Haddon, R. C.; Selegue, J. P.; Majidi, V. *J. Am. Chem. Soc.* **2002**, 124, (12), 3169-3174.
54. Duesberg, G. S.; Burghard, M.; Muster, J.; Philipp, G.; Roth, S. *Chemical Communications* **1998**, (3), 435-436.
55. Farkas, E.; Elizabeth Anderson, M.; Chen, Z.; Rinzler, A. G. *Chem. Phys. Lett.* **2002**, 363, (1-2), 111-116.
56. Pompeo, F.; Resasco, D. E. *Nano Lett.* **2002**, 2, (4), 369-373.
57. Georgakilas, V.; Tagmatarchis, N.; Pantarotto, D.; Bianco, A.; Briand, J. P.; Prato, M. *Chemical Communications* **2002**, (24), 3050-3051.
58. Bahr, J. L.; Yang, J.; Kosynkin, D. V.; Bronikowski, M. J.; Smalley, R. E.; Tour, J. M. *J. Am. Chem. Soc.* **2001**, 123, (27), 6536-6542.
59. Islam, M. F.; Rojas, E.; Bergey, D. M.; Johnson, A. T.; Yodh, A. G. *Nano Lett.* **2003**, 3, (2), 269-273.
60. Matarredona, O.; Rhoads, H.; Li, Z.; Harwell, J. H.; Balzano, L.; Resasco, D. E. *J. Phys. Chem. B* **2003**, 107, (48), 13357-13367.
61. McDonald, T. J.; Engtrakul, C.; Jones, M.; Rumbles, G.; Heben, M. J. *J. Phys. Chem. B* **2006**, 110, (50), 25339-25346.
62. Lehninger, A. L.; Cox, M. M.; Nelson, D. L., *Principles of biochemistry*. 2nd ed.; Worth Publishers: New York, N.Y., 1993; p 1 v. (various pagings).
63. Gellman, S. H. *Chemical Reviews* **1997**, 97, (5), 1231-1232.

64. Branden, C.; Tooze, J., *Introduction to protein structure*. 2nd ed.; Garland Pub.: New York, 1999.
65. Saenger, W., *Principles of nucleic acid structure*. Springer-Verlag: New York, 1984.
66. Gray, D. M.; Ratliff, R. L. *Biopolymers* **1977**, 16, (6), 1331-1342.
67. Early, T. A.; Olmsted, J.; Kearns, D. R.; Lezius, A. G. *Nucleic Acids Research* **1978**, 5, (6), 1955-1970.
68. Zheng, M.; Jagota, A.; Strano, M. S.; Santos, A. P.; Barone, P.; Chou, S. G.; Diner, B. A.; Dresselhaus, M. S.; McLean, R. S.; Onoa, G. B.; Samsonidze, G. G.; Semke, E. D.; Usrey, M.; Walls, D. J. *Science* **2003**, 302, (5650), 1545-1548.
69. Fritz, J. S.; Gjerde, D. T.; Gjerde, D. T., *Ion chromatography*. 3rd, completely rev. and enl. ed.; Wiley-VCH: Weinheim ; New York, 2000; p xiii, 254 p.
70. Griffiths, D. J., *Introduction to electrodynamics*. 3rd ed.; Prentice Hall: Upper Saddle River, N.J., 1999; p xv, 576 p.
71. Marsh, T. C.; Vesenka, J.; Henderson, E. *Nucleic Acids Research* **1995**, 23, (4), 696-700.
72. Campbell, J. F.; Tessmer, I.; Thorp, H. H.; Erie, D. A. *J. Am. Chem. Soc.* **2008**.
73. Pease, A. C.; Solas, D.; Sullivan, E. J.; Cronin, M. T.; Holmes, C. P.; Fodor, S. P. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, 91, (11), 5022-5026.
74. Heller, M. J. *Annu. Rev. Biomed. Eng.* **2002**, 4, (1), 129.
75. Wei, Y.; Lee, J.-M.; Richmond, C.; Blattner, F. R.; Rafalski, J. A.; LaRossa, R. A. *J. Bacteriol.* **2001**, 183, (2), 545-556.

76. Louws, F. J.; Fulbright, D. W.; Stephens, C. T.; de Bruijn, F. J. *Appl. Environ. Microbiol.* **1994**, 60, (7), 2286-2295.
77. Juretschko, S.; Timmermann, G.; Schmid, M.; Schleifer, K.-H.; Pommerening-Roser, A.; Koops, H.-P.; Wagner, M. *Appl. Environ. Microbiol.* **1998**, 64, (8), 3042-3051.
78. Nolte, F. S.; Metchock, B.; McGowan, J. E., Jr.; Edwards, A.; Okwumabua, O.; Thurmond, C.; Mitchell, P. S.; Plikaytis, B.; Shinnick, T. *J. Clin. Microbiol.* **1993**, 31, (7), 1777-1782.
79. Hamilton-Dutoit, S. J.; Pallesen, G.; Franzmann, M. B.; Karkov, J.; Black, F.; Skinhoj, P.; Pedersen, C. *Am J Pathol* **1991**, 138, (1), 149-163.
80. Star, A.; Tu, E.; Niemann, J.; Gabriel, J. C. P.; Joiner, C. S.; Valcke, C. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, 103, (4), 921-926.
81. Meng, S.; Maragakis, P.; Papaloukas, C.; Kaxiras, E. *Nano Lett.* **2007**, 7, (1), 45-50.
82. Gowtham, S.; Scheicher, R. H.; Pandey, R.; Karna, S. P.; Ahuja, R. *Nanotechnology* **2008**, 19, (12), 125701.
83. Dukovic, G.; Balaz, M.; Doak, P.; Berova, N. D.; Zheng, M.; McLean, R. S.; Brus, L. E. *J. Am. Chem. Soc.* **2006**, 128, (28), 9004-9005.
84. Hughes, M. E.; Brandin, E.; Golovchenko, J. A. *Nano Lett.* **2007**, 7, (5), 1191-1194.
85. Shim, M.; Kam, N. W. S.; Chen, R. J.; Li, Y. M.; Dai, H. J. *Nano Lett.* **2002**, 2, (4), 285-288.

86. Wang, S.; Humphreys, E. S.; Chung, S.-Y.; Delduco, D. F.; Lustig, S. R.; Wang, H.; Parker, K. N.; Rizzo, N. W.; Subramoney, S.; Chiang, Y.-M.; Jagota, A. *Nat Mater* **2003**, 2, (3), 196-200.
87. Nepal, D.; Geckeler, K. E. *Small* **2006**, 2, (3), 406-412.
88. Shi Kam, N. W.; Jessop, T. C.; Wender, P. A.; Dai, H. *J. Am. Chem. Soc.* **2004**, 126, (22), 6850-6851.
89. Kam, N. W. S.; Dai, H. *J. Am. Chem. Soc.* **2005**, 127, (16), 6021-6026.
90. Cherukuri, P.; Bachilo, S. M.; Litovsky, S. H.; Weisman, R. B. *J. Am. Chem. Soc.* **2004**, 126, (48), 15638-15639.
91. Jollès, P.; Jollès, J. *Mol. Cell. Biochem.* **1984**, 63, (2), 165-189.
92. Jiang, K. Y.; Schadler, L. S.; Siegel, R. W.; Zhang, X. J.; Zhang, H. F.; Terrones, M. *J. Mater. Chem.* **2004**, 14, (1), 37-39.
93. Szabo, A.; Ostlund, N. S., *Modern quantum chemistry : introduction to advanced electronic structure theory*. 1st. ed.; McGraw-Hill: New York, 1989; p xiv, 466 p.
94. Parr, R. G.; Yang, W., *Density-functional theory of atoms and molecules*. Clarendon Press ;
Oxford University Press: Oxford [England]
New York, 1989; p ix, 333 p.
95. Frenkel, D.; Smit, B., *Understanding molecular simulation : from algorithms to applications*. 2nd ed.; Academic Press: San Diego, Calif., 2002; p xxii, 638 p.
96. Becker, A. A., *An introductory guide to finite element analysis*. ASME Press: New York, 2004; p xiv, 171 p.
97. Allen, M. P.; Tildesley, D. J., *Computer simulation of liquids*. Clarendon Press ;

Oxford University Press: Oxford [England]

New York, 1989; p xix, 385 p.

98. Sakurai, J. J.; Tuan, S. F., *Modern quantum mechanics*. Rev. ed.; Addison-Wesley Pub. Co.: Reading, Mass., 1994; p x, 500 p.

99. Karplus, M.; McCammon, J. A. *Nat Struct Biol* **2002**, 9, (9), 646-652.

100. D.A. Case, D. A. P., J.W. Caldwell, T.E. Cheatham III, J. Wang, W.S. Ross, C.L.; Simmerling, T. A. D., K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley,; V. Tsui, H. G., R.J. Radmer, Y. Duan, J. Pitner, I. Massova, G.L. Seibel, U.C.; Singh, P. K. W. a. P. A. K. **2002**.

101. D. van der Spoel, E. L., B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff,; D. P. Tieleman, A. L. T. M. S., K. A. Feenstra, R. van Drunen and H. J. C.; Berendsen, *Gromacs User Manual version 3.3*. www.gromacs.org; 2005.

102. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, 98, (12), 10089-10092.

103. Ponder, J. W.; Case, D. A. *Protein Simulations* **2003**, 66, 27-+.

104. Kittel, C.; McEuen, P., *Introduction to solid state physics*. 8th ed.; J. Wiley: Hoboken, NJ, 2005; p xix, 680 p.

105. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, 117, (19), 5179-5197.

106. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.;

- Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, 102, (18), 3586-3616.
107. Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, 103, (2), 335-340.
108. Cheatham, T. E.; Kollman, P. A. *Annu. Rev. Phys. Chem.* **2000**, 51, 435-471.
109. Mackerell, A. D. *J. Comput. Chem.* **2004**, 25, (13), 1584-1604.
110. Reif, F., *Fundamentals of statistical and thermal physics*. McGraw Hill: New York,, 1965; p x, 651 p.
111. Chipot, C.; Pohorille, A., *Free energy calculations: theory and applications in chemistry and biology*. Springer: Berlin ; New York, 2007.
112. Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Wheeler, D. C. S. a. R., Chapter 4 Alchemical Free Energy Calculations: Ready for Prime Time? In *Annual Reports in Computational Chemistry*, Elsevier: 2007; Vol. Volume 3, pp 41-59.
113. Hess, B. *J. Chem. Phys.* **2002**, 116, (1), 209-217.
114. Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, 125, (8), -.
115. Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, 222, (6), 529-539.
116. Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, 122, (13), 134508.
117. Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. *J. Mol. Biol.* **2007**, 371, (4), 1118-1134.
118. Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. B* **2003**, 107, (35), 9535-9551.
119. Hermans, J.; Wang, L. *J. Am. Chem. Soc.* **1997**, 119, (11), 2707-2714.
120. Deng, Y. Q.; Roux, B. *J. Phys. Chem. B* **2009**, 113, (8), 2234-2246.

121. Garcia, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, (24), 13898-13903.
122. Garcia, A. E.; Sanbonmatsu, K. Y. *Proteins: Struct., Funct., Genet.* **2001**, 42, (3), 345-354.
123. Kannan, S.; Zacharias, M. *Biophys. J.* **2007**, 93, (9), 3218-3228.
124. Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, 60, (2), 96-123.
125. Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, 314, (1-2), 141-151.
126. Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, 7, (8), 306-317.
127. Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, 52, (12), 7182-7190.
128. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, 81, (8), 3684-3690.
129. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, 18, (12), 1463-1472.
130. Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. *Nature* **2001**, 414, (6860), 188-190.
131. Gao, H. J.; Kong, Y.; Cui, D. X.; Ozkan, C. S. *Nano Lett.* **2003**, 3, (4), 471-473.
132. Zhao, X. C.; Johnson, J. K. *J. Am. Chem. Soc.* **2007**, 129, (34), 10438-10445.
133. Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, 21, (12), 1049-1074.
134. Jin, H.; Jeng, E. S.; Heller, D. A.; Jena, P. V.; Kirmse, R.; Langowski, J.; Strano, M. S. *Macromolecules* **2007**, 40, (18), 6731-6739.
135. Yarotski, D. A.; Kilina, S. V.; Talin, A. A.; Tretiak, S.; Prezhdo, O. V.; Balatsky, A. V.; Taylor, A. J. *Nano Lett.* **2009**, 9, (1), 12-17.

136. Manohar, S.; Tang, T.; Jagota, A. *J. Phys. Chem. C* **2007**, 111, (48), 17835-17845.
137. Gowtham, S.; Scheicher, R. H.; Ahuja, R.; Pandey, R.; Karna, S. P. *Phys. Rev. B* **2007**, 76 033401-033403.
138. Ortmann, F.; Schmidt, W. G.; Bechstedt, F. *Phys. Rev. Lett.* **2005**, 95, (18), 186101.
139. Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, 112, (14), 5525-5534.
140. Shtogun, Y. V.; Woods, L. M.; Dovbeshko, G. I. *J. Phys. Chem. C* **2007**, 111, (49), 18174-18181.
141. Sowerby, S. J.; Cohn, C. A.; Heckl, W. M.; Holm, N. G. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, (3), 820-822.
142. Gordillo, M. C.; Marti, J. *Phys. Rev. B* **2008**, 78, (7), 075432.
143. Edelwirth, M.; Freund, J.; Sowerby, S. J.; Heckl, W. M. *Surf. Sci.* **1998**, 417, (2-3), 201-209.
144. Zhu, J.; Brink, M.; McEuen, P. L. *Nano Lett.* **2008**, 8, (8), 2399-2404.
145. Scheicher, R. H. *Private discussion*
146. Williams, L. D. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, 29, 497-521.
147. Baumann, C. G.; Smith, S. B.; Bloomfield, V. A.; Bustamante, C. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, 94, (12), 6185-6190.
148. Martin, W.; Zhu, W.; Krilov, G. *J. Phys. Chem. B* **2008**.
149. Fields, B. N.; Knipe, D. M.; Howley, P. M., *Fields' virology*. 5th ed.; Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia, 2007.

150. Jiang, S. K.; Jacobs, A.; Laue, T. M.; Caffrey, M. *Biochemistry* **2004**, 43, (7), 1847-1853.
151. Bewley, M. C.; Springer, K.; Zhang, Y. B.; Freimuth, P.; Flanagan, J. M. *Science* **1999**, 286, (5444), 1579-1583.
152. Karajanagi, S. S.; Vertegel, A. A.; Kane, R. S.; Dordick, J. S. *Langmuir* **2004**, 20, (26), 11594-11599.
153. Cui, Y.; Wei, Q. Q.; Park, H. K.; Lieber, C. M. *Science* **2001**, 293, (5533), 1289-1292.
154. Asay, D. B.; Kim, S. H. *J. Phys. Chem. B* **2005**, 109, (35), 16760-16763.
155. Sze, S. M.; Ng, K. K., *Physics of semiconductor devices*. 3rd ed.; Wiley-Interscience: Hoboken, NJ, 2007; p x, 815 p.
156. Zheng, L. F.; Jain, D.; Burke, P. *J. Phys. Chem. C* **2009**, 113, (10), 3978-3985.
157. Dieckmann, G. R.; Dalton, A. B.; Johnson, P. A.; Razal, J.; Chen, J.; Giordano, G. M.; Munoz, E.; Musselman, I. H.; Baughman, R. H.; Draper, R. K. *J. Am. Chem. Soc.* **2003**, 125, (7), 1770-1777.