2016

Sequence Dependent Interactions and Recognition between DNA and Single-Walled Carbon Nanotubes

Akshaya Shankar
Lehigh University

Sequence Dependent Interactions and Recognition between DNA and Single-Walled Carbon Nanotubes

By

Akshaya Shankar

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy of the Degree of

Doctor of Philosophy

in

Chemical Engineering

Lehigh University

May 2016

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Date

_____

Prof. Anand Jagota

(Dissertation Director)

_____

Accepted Date

Committee Members:

_____

Prof. Anand Jagota

(Committee Chair)

_____

Prof. Jeetain Mittal

_____

Prof. Mark A. Snyder

_____

Prof. Dmitri Vezenov

_____

Prof. Vyacheslav V. Rotkin

Acknowledgements

The journey towards completing my PhD has been an amazing experience and has served to change and teach me in many ways. A lot of people have contributed to this experience and I wish to thank them for everything. I would like thank my parents for their unconditional support and making a lot of sacrifices in order to enable me to reach where I am. I would like to dedicate this work to them. I would like to thank my advisor Prof. Anand Jagota for being such a wonderful and patient advisor, for supporting, encouraging and guiding me through the difficult patches. Thank you for being such a role model and always motivating and pushing me throughout these years, helping me become a better researcher. I would also like to thank Prof. Jeetain Mittal, for patiently supporting and encouraging me during my first steps into the world of research and teaching me so patiently. I am really grateful for all that I have learned from you and honored to have been mentored by you. I would also like to thank Prof. Dmitri Vezenov, Prof. Slava Rotkin and Prof. Mark Snyder for all the advice, suggestions and help over the years. Your valuable feedback has helped me tremendously in improving my work. I would also like to thank all the professors at Lehigh who have taught me so much through coursework and beyond. Thank you to all the professors who I have assisted in teaching. A special mention for Prof. Lori Herz and Dr. Logan MacDonald for patiently introducing me to the world of bioengineering laboratory.

I would like to thank Dr. Daniel Roxbury for being the best teacher I could wish for, patiently helping me and going out of your way to help me. I would also like to especially thank Sara Iliafar Cook, Apratim Bhattacharya, Yoona Yang, Minseok Song, Nichole Nadermann, Ying Bai, Walter Jokiel, Nicole Fortoul, Dadhichi Paretkar, Snow Xu, Pankaj Singh, Zhenping He, Jonathan Dillen and Nichole Asermely for being such wonderful friends and colleagues! You have helped and taught me

countless times over the years and I am really grateful for your support and companionship. I wish you all the best in your future careers and lives. Thank you so much Coco Xue for your invaluable and patient help with the experiments. I would like to thank my collaborators at NIST, Dr. Ming Zheng, Dr. Constantine Khripin and a very special thanks to Dr. Ao Geyou for teaching me so much about aqueous two phase separation and allowing me the opportunity to have such a wonderful experience at NIST. I would like to thank all the staff at the Chemical Engineering and Bioengineering departments, especially Barbara, Tracey, Janine and Cindy for going out of your way to help me countless times at Lehigh.

My experience at Lehigh is what it is due to all my friends who have made it so enjoyable and memorable. My lovely housemates and closest friends, Preeti, Sonam and Fon, you have been my biggest pillars of strength and I cannot thank you enough. Mettu, Dhruthi, Rikhi, Abhishek, AK, Anshu, Chinmoy, Nipun, Midhun, Tharanga, Erg, Aditi, Nishant and everyone from India Club, you have made my life at Lehigh so memorable and enjoyable and I thank you for everything.

My biggest thanks are to my family, my parents, my sister Nithya who patiently stood by me and supported me during the thick and thin even when I did not have enough time for them., my grandparents, uncles, aunts and cousins, you have made me what I am today and I love you very much. Finally I would like to thank my husband and best friend, Anand Srinivas Guruswamy, without whose love and support this would not be possible. Thank you for bringing such a lovely family, Ma, Pa, Aditya and Amrita into my life and enriching it.

Table of Contents

List of Tables

List of Figures

Abstract

Since DNA-SWCNT hybrids have a number of potential biomedical applications such as molecular sensing, drug delivery and cell imaging, it is essential to characterize them and to understand their structure and properties. Certain single stranded DNA (ssDNA) sequences are known to recognize their partner single wall carbon nanotube (SWCNT). We report here the activation energies for removal of several ssDNA sequences from a few SWCNT species by a surfactant molecule. We found that DNA sequences systematically have higher activation energy of dissociation from their carbon-nanotube recognition partner than on non-partner species. Since the difference in binding affinity and difference in partitioning can depend on DNA structure on the single walled carbon nanotube (SWCNT), we studied the partitioning of the various DNA sequences in an aqueous two phase system. We found that for two sequences of same length, $(CCA)_{10}$ on (6,5) SWCNT requires much higher amount of modulant to be moved from the relatively hydrophilic phase to the more hydrophilic phase as compared to $(GT)_{15}$ on (6,5), suggesting that the solvation energy depends greatly on the DNA sequence. We also found that various sequences with the same length but different repeating units of two bases exhibit different hydration energies on the same SWCNT (6,5).

Unlike the majority of DNA structures in bulk that are stabilized by canonical Watson-Crick pairing between Ade-Thy and Gua-Cyt, those adsorbed on surfaces are often stabilized by non-canonical base pairing, quartet formation, and base-surface stacking. All-atom molecular simulations of DNA bases in two cases - in bulk water and strongly adsorbed on a graphite surface – are conducted to study the relative strengths of stacking

and hydrogen bond interactions for each of the ten possible combinations of base pairs. We find that stacking interactions exert the dominant influence on the stability of DNA base pairs in bulk water in the order, Gua-Gua > Ade-Gua > Ade-Ade > Gua-Thy > Gua-Cyt > Ade-Thy > Ade-Cyt > Thy-Thy > Cyt-Thy > Cyt-Cyt. On the other hand, mutual interactions of surface adsorbed base pairs are stabilized mostly by hydrogen bonding interactions in the order, Gua-Cyt > Ade-Gua > Ade-Thy > Ade-Ade > Cyt-Thy > Gua-Gua > Cyt-Cyt > Ade-Cyt > Thy-Thy > Gua-Thy. Interestingly, several non-Watson-Crick base pairings, that are commonly ignored, have similar stabilization free energies due to inter-base hydrogen bonding as Watson-Crick pairs. This clearly highlights the importance of non-Watson-Crick base pairing in the development of secondary structures of oligonucleotides near surfaces.

Hybrids of single stranded DNA and single walled carbon nanotubes have proven very successful in separating various chiralities and, very recently, enantiomers of carbon nanotubes using aqueous two-phase separation. This technique sorts objects based on small differences in hydration energy, which is related to corresponding (small) differences in structure. Separation by handedness requires that a given ssDNA sequence adopt different structures on the two SWCNT enantiomers. Here we study the physical basis of such selectivity using a coarse grained model to compute the energetics of ssDNA wrapped around an SWCNT. Our model suggests that difference by handedness of the SWCNT requires spontaneous twist of the ssDNA backbone. We also show that differences depend sensitively on the choice of DNA sequence.

Chapter 1 Introduction to carbon nanotube – DNA hybrids and their characterization methods

1.1. **Carbon nanotubes: chiralities, enantiomers, properties and applications**

Single walled carbon nanotubes (SWCNTs) are low dimensional tubular structures of a single layer of sp^2 hybridized carbon atoms bonded in a hexagonal lattice except at their ends.[1] They are often visualized as seamlessly rolled up graphene.[2] As shown in Figure 1.1. (b), a carbon atom on the graphene sheet is chosen as origin and the chiral vector C_h is drawn from the origin atom to another atom on the same graphene sheet. When n and m are two integers and a_1 and a_2 are the unit cell vectors of the two-dimensional lattice formed by the graphene sheet, $C_h = na_1 + ma_2$. The chirality of the carbon nanotube so formed is represented as (n,m), where the direction of the nanotube axis is perpendicular to the chiral vector C_h. Single walled carbon nanotubes can be classified as chiral and achiral SWCNTs. This has a direct effect on their properties such as being metallic and semiconducting.[3] SWCNTs are found to be metallic if |n-m| is a multiple of 3 and semiconducting if not. Some exceptions to this rule exist, such as (5,0) which is metallic instead of semiconducting, as curvature effects on the electrical properties are significant in case of small diameter nanotubes. Like other chiral molecules, chiral SWCNTs can exist as right handed and left handed enantiomers.[4] Popular SWCNT production techniques such as electric arc discharge, laser ablation, gas phase catalytic growth from carbon monoxide and

chemical vapor deposition (CVD) from hydrocarbons[5] result in a mixture of different chiralities, which consequently have different electronic structure.

Due to their extraordinary mechanical, electronic and optical properties[6,7], SWCNTs have found a number of applications in the recent times. Since they have high carrier mobility, they have been used in thin film field effect transistors (FET)[8,9]. This combined with their high optical transparency and high chemical stability makes them useful for organic photovoltaics[10]. Additionally, this high chemical stability along with the high aspect ratio which results in all the atoms being exposed to the environment has led to applications as effective catalyst supports.[11,12] SWCNTs have very small, nanometer-scale, dimensions and exhibit band-gap fluorescence which is highly sensitive to the surrounding environment. This makes them perfect candidates for various types of biosensors[13–15]. They also become more bio-compatible when coated with biological molecules and can cross mammalian cell membranes, making them promising for targeted drug delivery[16,17].

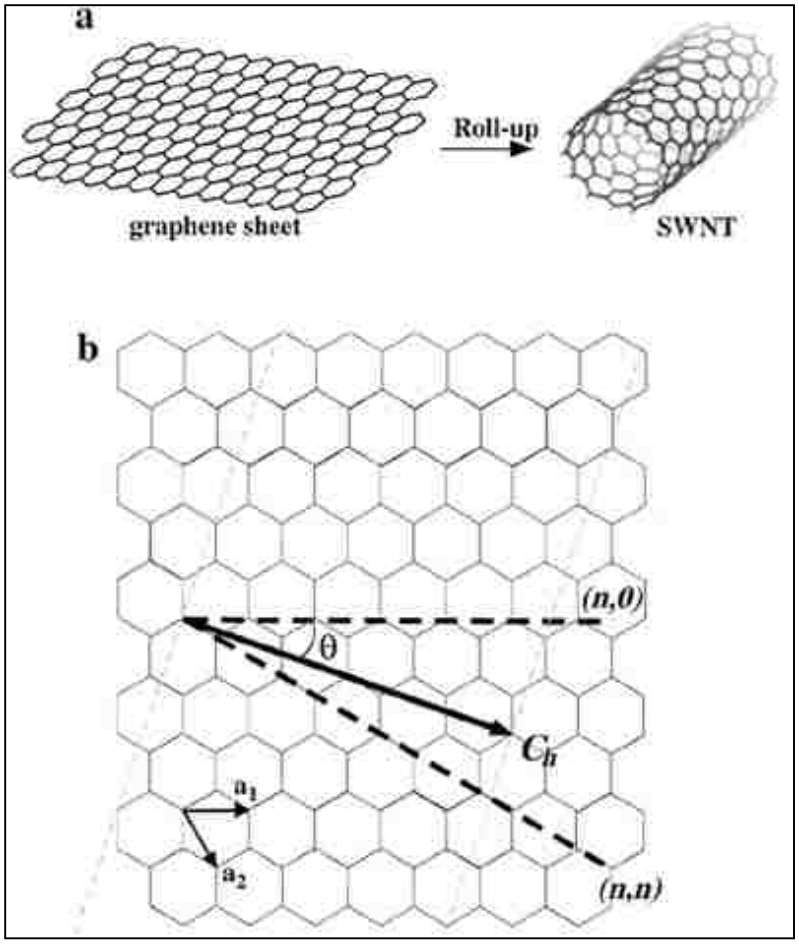


Figure 1.1. Single walled Carbon nanotube (SWCNT) chirality schematic

(a) Schematic of a portion of a graphene sheet rolled up to form an SWCNT. (b) 2D graphene sheet illustrating lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$, and the roll-up vector $C_h = n\mathbf{a}_1 + m\mathbf{a}_2$. The achiral, limiting cases of (n, 0) and (n, n) armchair are indicated with thick**,** dashed lines, and the chiral θ angle is measured from the zigzag direction. The light, dashed parallel lines define the unrolled, infinite SWCNT. The diagram has been constructed for (n, m) = (4, 2). Figure 1.1.(a) and Figure 1.1.(b) were published by Odom et al.[2] and proper copyright permission was obtained from John Wiley and Sons prior to submittal of this document.

1.2. **Single stranded deoxyribonucleic acid (DNA)**

Deoxyribonucleic acid (DNA) is one the most important biological molecules which is usually present in nature in the double stranded form. Single stranded DNA (ssDNA) is a biopolymer whose monomeric unit is a nucleotide composed of three units: deoxyribose sugar, negatively charged phosphate group and a nucleobase. There are four nucleobases: Adenine (Ade or A) and Guanine (Gua or G), which are purines with two aromatic rings each, and Thymine (Thy or T) and Cytosine (Cyt or C), which are pyrimidines with one aromatic ring each. The free phosphate end of the strand is called the 5' end and the free hydroxyl group end is called the 3' end. Every nucleotide has a hydrophilic part, i.e., the phosphate group and a hydrophobic part, i.e., the aromatic nucleobase. The single stranded DNA is thus amphiphilic and has surfactant like properties.

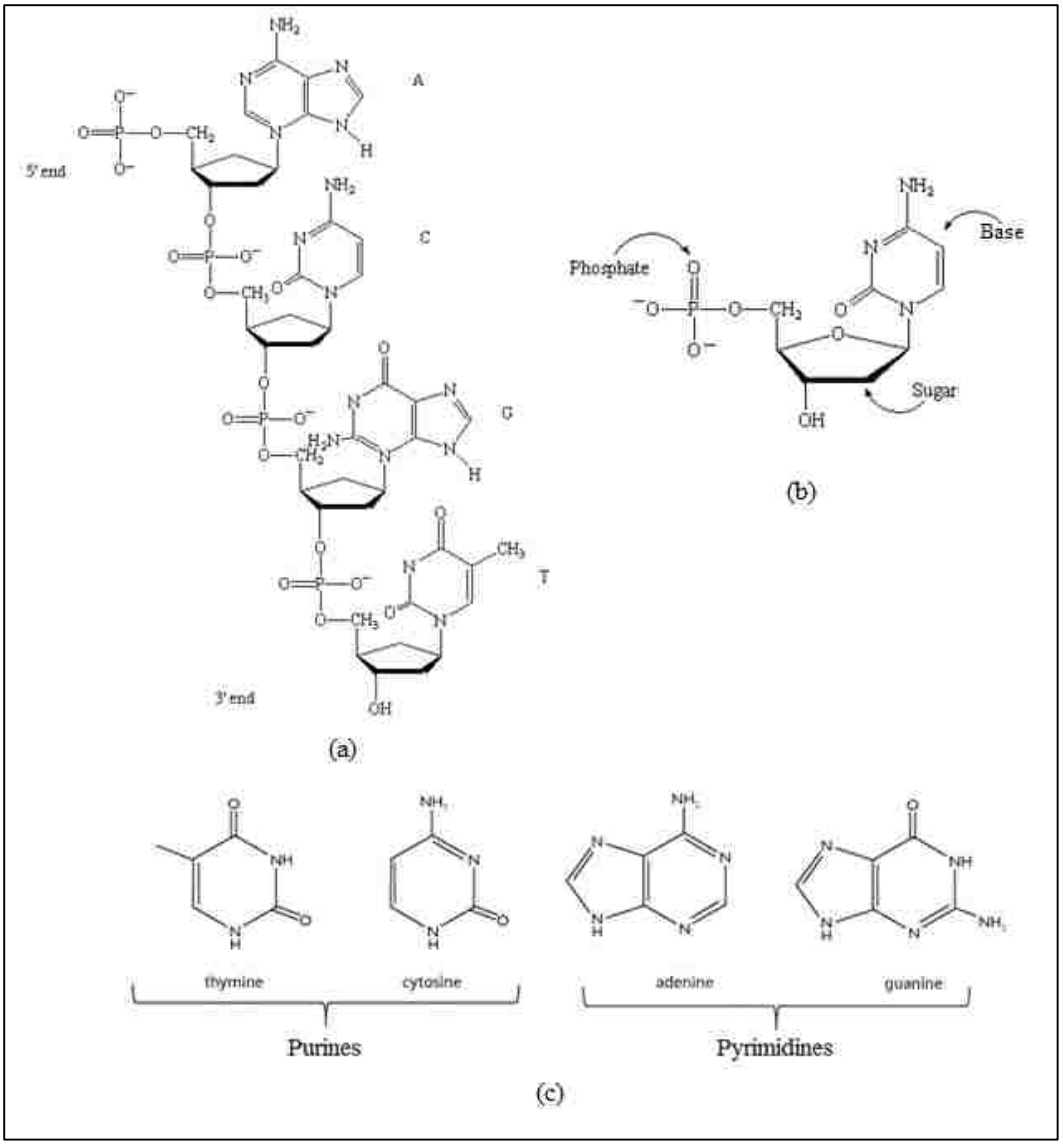


Figure 1.2 Single stranded DNA structure

(a) Single stranded DNA sequence (5'-ACGT-3')[18], (b) Cytosine nucleotide[18], (c) DNA nucleobases: Thymine (Thy or T) and Cytosine (Cyt or C) are Purines, Adenine (Ade or A) and Guanine (Gua or G) are Pyrimidines[19,20]

1.3. **Single walled nanotubes – single stranded DNA hybrids**

Single walled carbon nanotubes have surfaces which are hydrophobic and hence they aggregate into clumps when placed in an aqueous solution. Such poor dispersions are usually not desirable for further processing for most applications. In order to produce uniform solutions of singly dispersed nanotubes, SWCNTs are wrapped with amphiphilic molecules such as surfactants, single stranded DNA and proteins. The hydrophobic parts of the amphiphilic molecule interact with the hydrophobic surface of the SWCNT and the hydrophilic regions render the resulting hybrid water-soluble. In case of single-stranded DNA, the hydrophobic aromatic rings in the bases are thought attach non-covalently on the SWCNT surface via π-stacking. Each of the resulting ssDNA-SWCNT hybrids is negatively charged.

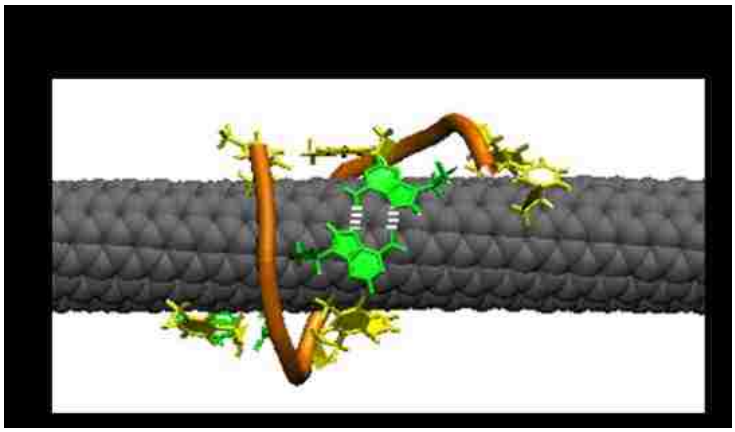


Figure 1.3 Representation of single stranded DNA-single walled carbon nanotube hybrid. $(TAT)_4$ on (6,5) SWCNT

1.4. **Characterization techniques**

Since DNA-SWCNT hybrids have a number of potential applications as discussed above, it is essential to characterize them and to understand their structure and properties. The structure of DNA on SWCNT has been probed in many ways including molecular dynamics (MD) simulations[21–24], measuring activation energy of displacement of DNA by a surfactant[25,26], Atomic Force Microscope (AFM) studies[27–29] and aqueous two phase separations[30]. Other characterization techniques include capillary electrophoresis, and optical spectroscopic techniques such as UV-Vis-NIR, Photoluminescence and Raman spectroscopy. Single stranded DNA MD simulations have given some idea of how the single stranded DNA adsorbs onto the SWCNT surface. Tools like surfactant exchange studies, aqueous two phase studies and AFM studies have helped estimate the total binding free energy, activation barriers and small differences in solvation energy for different DNA sequences on different SWCNT chiralities. Some of these techniques are described as follows:

1.4.1. Optical spectroscopy

Single-walled carbon nanotubes (SWCNTs) have characteristic optical transition energies associated with their (n,m) chirality.[31] They also exhibit solvatochromic shift where these transition energies depend on the surrounding solvents and adsorbed molecules.[32] For example, when DNA strands adsorbed onto SWCNTs are replaced by surfactant molecules such as sodium dodecyl benzene sulfonate (SDBS), a blueshift is observed.[26] This can be

tracked using optical measurement techniques such as absorbance and fluorescence spectroscopy.

1.4.2. Hydration energy using aqueous two phase system

Aqueous two phase systems have been used to sort SWCNTs based on chirality and enantiomers.[30] It is proposed that the DNA-SWCNT hybrids partition in the aqueous two phase because of sensitive dependence of the free energy of hydration on the spatial distribution of hydrophilic groups in the DNA-SWCNT hybrid. Hence it may be possible to study or rank hydration or solvation free energy of these hybrids using such systems.

1.4.3. Molecular Dynamics simulations (all atom)

Most of the DNA sequences studied in the DNA-CNT hybrids are relatively short (from 6mers to 100mers), so the system size is quite small and can be simulated using all atom molecular models effectively. The number of types of interactions present in such systems are also sufficiently limited for the available force fields to describe these interactions. A number of molecular dynamics studies have been conducted on the DNA-SWCNT hybrid previously in order to study the structure of DNA near SWCNTs.[21,22,24,33,34]

1.4.4. Atomic Force Microscopy (AFM)

Atomic Force Microscopy (AFM) is one of the most important single molecule experimental techniques to study and quantify interactions between biomolecules and material surfaces. It can be used for both imaging areas upto 100 x 100 μm^2 in a line by

line fashion with sub-nanometer lateral resolution and subatomic (<1A^{o}) vertical resolutions. Single walled carbon nanotubes range from 0.4 nm to 2 nm in diameter and a few hundred nm to several millimeters in length. Hence AFM is ideal for imaging SWCNT coated with different biomolecules. The most important use of AFM in quantifying SWCNT-DNA interactions is due to its Pico newton force sensitivity. Peeling experiments can hence be conducted using AFM, measuring the intra and intermolecular forces which separate surfaces at the single molecular level, enabling quantification of DNA base-SWCNT interactions.

1.4.5. X-ray Photoelectron Spectroscopy (XPS)

X-ray photoelectron spectroscopy (XPS) is a technique that measures the elemental composition along with their chemical and electronic states within 2 to 10 nm of a surface with low detection limits of ~ - 0.1 at%.[35] This makes it suitable to study the DNA bound to the surface of the SWCNT.

1.5. **Outline of Thesis**

It is empirically known from experiments reported by Tu et al. that certain special DNA sequences called ‘recognition’ sequences are able to selectively separate certain SWCNT chiralities via ion exchange chromatography.[36] Roxbury et al. showed that even the slightest change to the recognition sequence significantly affects its binding affinity to the partner SWCNT chirality.[26] It is hypothesized that the separation of SWCNTs via

recognition sequences is related to their binding affinity. The DNA assisted separation technique is carried out on a mixture of SWCNT chiralities. In *Chapter 2*, we studied the binding affinity of various recognition sequences with their partner and non-partner SWCNT chiralities to find if separation due to recognition ability is correlated to its binding affinity. We also studied the effect of length of the single stranded DNA with the binding affinity for a small range close to a recognition sequence.

It is hypothesized that the DNA-SWCNT hybrids partition in the aqueous two phase system because of sensitive dependence of the free energy of hydration on the spatial distribution of hydrophilic groups in the DNA-SWCNT hybrid.[30] Since the differences in SWCNT-DNA hybrids are attributed to the differences in the structure of DNA on various SWCNTs, the hydration energy of these hybrids should also be slightly different. In *Chapter 3*, we show some preliminary studies of the hydration energy of various recognition sequences and if this is dependent on the length or composition of the DNA sequences.

Several molecular dynamics simulation studies have shown that the structure of single stranded DNA on the SWCNTs is stabilized by hydrogen bonds. It is seen that some of these hydrogen bonds occur between non-Watson-Crick pairs such as Adenine-Adenine. This is very different from double stranded DNA where the hydrogen bonds between the two single stranded DNA are Watson-Crick pairs only (Adenine-Thymine and Guanine-Cytosine). This raises the question of whether the interactions that give rise to DNA

structure in aqueous bulk phase are different from those in case of DNA structures near surfaces such as SWCNTs and graphene. In *Chapter 4*, molecular dynamics simulations were conducted for all possible DNA base pairs including all non-Watson-Crick pairs in aqueous bulk phase and near graphite surface. The stability of these interactions as represented by binding free energy for each of the ten possible pairs was then ranked.

Recently various SWCNT enantiomers have been successfully separated using the aqueous two phase system. Separation by handedness requires that a given single stranded DNA sequence adopt different structures on the two SWCNT enantiomers. In *Chapter 5*, we study the physical basis of such selectivity using a coarse grained model to compute the energetics of single stranded DNA wrapped around an SWCNT.

It has been shown experimentally that various DNA sequences which are closely related to the recognition sequence i.e. $(TAT)_4$ show very different binding affinity to the partner SWCNT (6,5). It is proposed that this is because of subtle differences in the structure of these sequences on the same SWCNT. Certain DNA sequences have also been successfully used to separate both enantiomers of (6,5) SWCNT. In *Chapter 6*, molecular dynamics simulations of these DNA-SWCNT combinations were conducted and analyzed.

In *Chapter 7*, the major findings of this thesis are summarized and some possibilities for future work are discussed.

Chapter 2 Sequence and chirality dependence of binding between DNA and Carbon Nanotubes

Certain single stranded DNA (ssDNA) sequences are known to recognize their partner single wall carbon nanotube (CNT). Here we report on activation energies for removal of several ssDNA sequences from a few SWCNT species by a surfactant molecule. We find that DNA sequences systematically have higher activation energy on their carbon-nanotube recognition partner than on non-partner species. For example, the DNA sequence $(TAT)_4$ has much lower activation energy on the (9,1) SWCNT than on its partner (6,5) SWCNT whereas the DNA sequence $(CCA)_{10}$ binds strongly to its partner (9,1) SWCNT compared to (6,5). The (6,5) and (9,1) SWCNTs have the same diameter but different electronic properties, suggesting that activation energy difference can detect differences in the arrangement of carbon atoms of the underlying SWCNT. The activation energies of increasing lengths of closely related sequences from the 11mer $(TAT)_3TA$ to the 21mer $(TAT)_7$ on three different SWCNT species (9,1), (6,5), and (8,3) were measured. For the shorter sequences, the activation energy on the SWCNT varies periodically with sequence.

The work described in this chapter was published in "Shankar, A., Mittal, J., & Jagota, A. (2014). Binding between DNA and carbon nanotubes strongly depends upon sequence and chirality. *Langmuir*, *30*(11), 3176–83"

2.1. **Introduction**

Single stranded DNA conjugated with single-walled carbon nanotubes (CNT) allows the latter to be dispersed readily in water. Certain special short DNA sequences have recognition ability towards partner species of single-walled carbon nanotubes.[36] This ability has been useful in sorting these species from a mixture of carbon nanotubes. There is also considerable interest in the hybrids of carbon nanotubes and DNA due to their potential in biomedical applications such as targeted cellular drug and siRNA delivery[37,38], sensing[39–41], and *in vivo* imaging[42,43]

In order to understand DNA-CNT interactions and to predict their behavior in such applications, it is important to study the relationship between DNA sequence, the SWCNT species, and the binding affinity between these two. The adsorption and assembly of DNA bases at a liquid-solid interface has been well studied.[44–48] It is known that the binding strength of DNA bases on the graphite surface in the presence of water follows the order guanine > adenine > thymine > cytosine.[46] Direct measurements of the binding strength of DNA oligomers on graphite surfaces using single molecule force spectroscopy[27,29,49] reported binding free energy in the range 7-11 k_BT decreasing in the order thymine > adenine > guanine > cytosine. A recent single molecule force spectroscopy study reported the free energy of binding of DNA to carbon nanotube surface in the range of 17-38 k_BT per nucleotide decreasing in the order adenine > guanine > thymine > cytosine, which is much larger than on graphite surface.[28] Studies of similar systems have also been conducted using molecular dynamics simulation, where it has been found that short DNA

homopolymers also follow the same order of adhesion strength on graphite surface as that measured experimentally for single bases.[50] Short DNA-CNT hybrids are known to dissociate thermally in the order: guanine > cytosine > adenine > thymine and it is reported that the binding free energies of the ssDNA to SWCNTs increased monotonically with increasing DNA sequence length.[51]The structures formed by the DNA strands on SWCNT surfaces and the factors contributing to their stability have been extensively studied using various experimental techniques[52,53] as well as molecular dynamics simulations.[21–24,33,34,54,55]The order of strength of the two interactions which contribute to the stability of DNA structure on graphite, i.e., stacking and hydrogen bonding for all the ten possible DNA base pairs, have been studied and it has been found that non-Watson Crick hydrogen bonding plays an important role in stabilizing DNA structures near surfaces like graphite and SWCNT.[22,56]Recently, an experimental study was conducted on the equilibrium thermodynamics of different polycytosine sequences and a surfactant, sodium cholate, on different SWCNT species.[57]

It has been found that the recognition ability of certain single stranded DNA sequences for their partner SWCNT correlates strongly with binding affinity.[26] Roxbury et al.[26]measured binding affinity between (6,5) SWCNTs, its recognition sequence $(TAT)_4$, and closely related DNA sequences $(TAT)_3T$, $(TAT)_3TA$, $(TAT)_4T$ and $(TAT)_4TA$ by measuring the kinetics (and hence the activation energy) of DNA displacement by a surfactant molecule. It was shown that addition or subtraction of just one base from the recognition sequence can enhance the kinetics of DNA displacement by some 20-fold.[26]The study by Roxbury et al.[26]showed that activation energy for a particular SWCNT

is strongly DNA sequence-specific in that small changes to the recognition DNA sequence result in large changes in activation energy. However, separation in practice is performed with a single DNA sequence dispersing a mixture of SWCNTs. The question of the difference in activation energy for a given sequence and different SWCNTs remains unexplored; addressing it is the main objective of this work. Some questions that remain unanswered are: will other recognition sequences also have similarly high activation energy with their partner SWCNTs compared to their compositional cousins? What is the typical difference in activation energy between recognition and non-recognition partners?

In order to answer these questions, we performed a study of the activation energies of several DNA sequences and SWCNTs. We used a surfactant molecule, sodium dodecyl benzene sulfonate (SDBS), in sufficient concentration to displace DNA off the surface of SWCNTs. By monitoring the kinetics of this process, we obtained activation energy as a measure of binding affinity.

2.2. **Experimental Methods:**

2.2.1. Sample Preparation

Single stranded DNA was obtained from Integrated DNA Technologies (IDT). All the chemicals apart from the SWCNT and DNA were procured from Sigma Aldrich. The DNA and the Hipco SWCNT (NanoIntegris) in a 2 mg/1 mg weight ratio were sonicated using a Branson model 150 probe ultrasonicator for 90 minutes at 8W output power in a 10 mM 7.1 pH phosphate buffer. The resultant dispersion was centrifuged using an Eppendorf microcentrifuge at 13.1 K rpm for 90 minutes to precipitate impurities such as

catalysts and undispersed SWCNTs, and the supernatant was separated out. This supernatant was then passed through a 100kDa Millipore microcentrifuge filter and redispersed in the phosphate buffer repeatedly (three times) to remove the excess DNA. All the dispersions were then diluted such that the absorbance at 990 nm was ~ 0.5 in order to standardize all the dispersions to the same concentration. It may be noted here that in the previous work by Roxbury et al.[26] the DNA-CNT dispersions were passed through size exclusion chromatography (SEC) to remove excess DNA and to sort SWCNTs according to length. However, the samples studied in this work were not passed through SEC for two reasons. Firstly, the SEC step would not be done normally during the process of DNA-assisted separation of SWCNTs. Secondly, this step caused a major loss of sample for the relatively weakly bound sequences. A solution of 0.2 wt % sodium dodecyl benzene sulfonate (SDBS) was prepared in the phosphate buffer solution. 80 μl of the SDBS solution was pre-heated in a quartz cuvette to the required temperature using a Peltier temperature control device. An equal volume (i.e., 80 μl) of the DNA-CNT dispersion was added to the cuvette and mixed using the pipette. Thus, the SDBS concentration in the reaction mixture was 0.1wt %.

2.2.2. Absorbance Spectroscopy

The Hipco SWCNT mixture contains various SWCNT species such as (6,5), (7,5), (7,6), (8,3), (9,1) etc. In this work, due to the range limitations of our instrument, we focus attention on (9,1), (8,3), and (6,5) species. To study the relationship between sequence/CNT match and activation energy, Roxbury et al.[26]developed a surfactant-

exchange technique in which the progress of the replacement of ssDNA from the surface of SWCNTs by surfactant molecules could be monitored by absorbance spectroscopy (UV-Vis spectrophotometer: Varian Cary 50).

It is known that a change in the environment of the carbon nanotube causes a peak shift in the optical absorbance and fluorescence (solvatochromic effect[32,58]). Figure 2.1(a) shows two absorbance spectra for a mixture of SWCNTs. The initial one corresponds to coating by the ssDNA sequence $(TAT)_4$; the final one is after ssDNA has been replaced by the SDBS surfactant molecule. The wavelengths assigned to the various SWCNT species correspond to the peak assignment by Tu et al.[36]

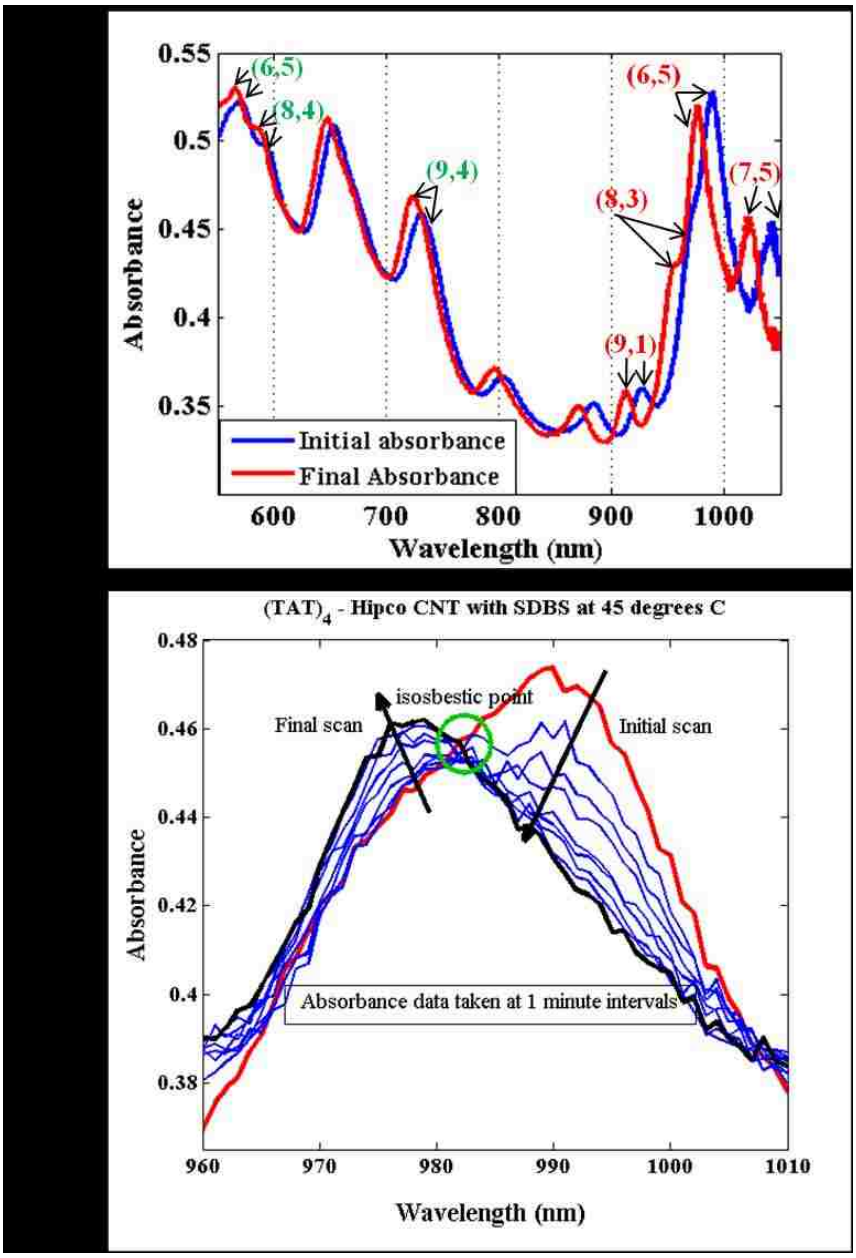


Figure 2.1. Time resolved absorbance spectra during surfactant exchange on DNA wrapped SWCNTs

(a) Absorbance peak shift for the various nanotube SWCNT species before and after the exchange reaction of DNA with SDBS molecules. (b) Time resolved absorbance spectra show an isosbestic point where the absorbance intensity does not change with time.[26]

Figure 2.1. (b) shows the time evolution of the absorbance spectrum in the vicinity of the (6,5) absorbance peak as the DNA coating on the SWCNT surface is replaced by SDBS. A characteristic feature is the existence of an isosbestic point indicated by the green circle in Figure 2.2., where the absorbance does not change during this process. At this wavelength, the absorbance intensity is same for DNA and SDBS coated species. The fact that it does not change with time, along with other supporting evidence discussed by Roxbury et al.[26], indicates that the absorbance at a particular wavelength can be represented as a linear combination of each of two pure species, i.e., a particular SWCNT exists with complete coverage either solely by DNA or by SDBS. Intermediate stages where both DNA and SDBS are on the same SWCNT surface are very short-lived.

2.2.3. Fluorescence Spectroscopy

To confirm independently that there is indeed a shift in the peaks for each of the SWCNT species as the DNA adsorbed onto it is replaced by the SDBS, two dimensional fluorescence maps of the dispersion were measured (Fluorolog-3; Horiba-Jobin Yvon) before and after the SDBS exchange with the DNA on the nanotube surface. (The absorbance spectra were measured during the exchange process.) We used 500−800 nm excitation with increments of 3 nm and slit width of 8 nm and 900−1200 emission with increments of 3 nm and slit width of 8 nm. Dark count correction factors were applied with a 1.0 s time of integration. As shown in appendix Figure 2.8. , the shift in the absorbance peaks due to the change in the coating of the SWCNT is also seen in the photoluminescence

spectra. In appendix we also present table 2.3. (a) and (b) that compares peak values as measured by absorbance and fluorescence spectroscopy.

2.2.4. Analysis of time-resolved absorbance spectra

Figure 2.2 draws schematically the characteristic free-energy landscape followed when the initial reactants – DNA-covered SWCNTs – are converted to the product, SDBS-covered SWCNTs and free DNA, in the presence of solvent and excess surfactant. The reaction runs irreversibly from reactants to products and has an activation free energy barrier, $\Delta G^{\ddagger}$. For a given type of SWCNT, the product free energy is presumably the same, regardless of the type of DNA sequence that coats the SWCNT initially. Therefore, the activation energy barrier, which we infer by measuring the kinetics of DNA removal, reflects differences in how different DNA sequences bind to the same SWCNT. This also assumes that SDBS-DNA interactions are either too weak to matter or do not depend on DNA sequence. Our choice of SDBS as the surfactant is based on surface tension measurements suggesting that the interaction between SDBS and DNA is quite weak.[26] Differences between sequences could arise either because the free energy of the initial state is different or because the energy of the transition state is different. Underlying differences in the activation barrier manifest as differences in the kinetics of DNA removal from the SWCNT surface, which is the quantity we measure.

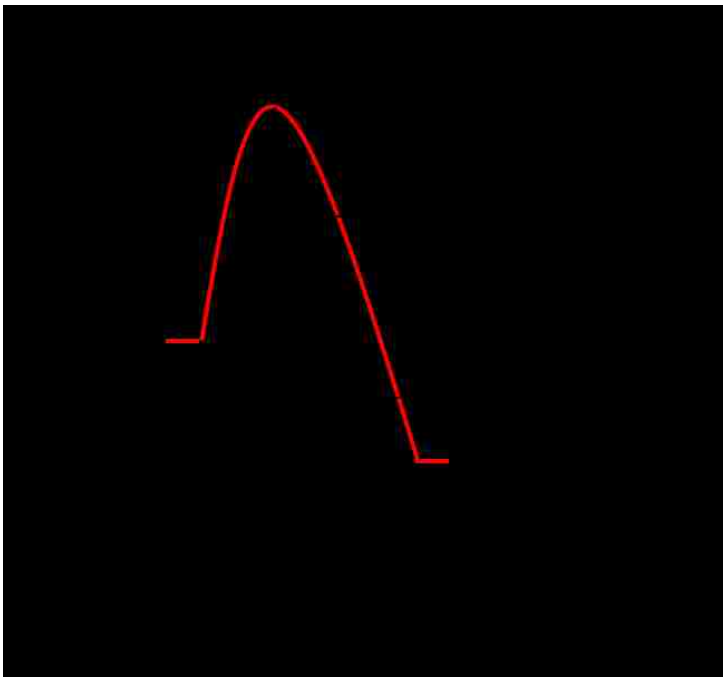


Figure 2.2 Free energy change during conversion of DNA-CNT to SDBS-CNT

For samples where the absorbance spectra are dominated by a single SWCNT species, the kinetics of the exchange reaction can be followed by tracking the absorbance at a single wavelength as a function of time.[26]However, for a mixture of SWCNT species, the absorbance lines overlap, so that absorbance change at a given wavelength depends on the kinetics of surfactant exchange from several different SWCNTs. In our typical samples, this means that the larger peaks such as the (6, 5) peak strongly influence weaker peaks like (8, 3) and (9, 1). We therefore developed a procedure to decompose the absorbance spectra into contributions from individual species.

The details of the procedure can be found in appendix, section S2. In essence, it consists of the following:

a) Construction of fits to absorbance spectra of pure species, i.e., either purely DNA or purely SDBS coated SWCNTs, by an asymmetric Lorentzian line shape.[36] Let these fits be given by $g^{DNA}{}_{(m,n)}(\lambda)$, and $g^{SDBS}{}_{(m,n)}(\lambda)$, two functions for each SWCNT type.

b) Decomposition of absorbance spectra for samples with only DNA (Figure 2.3.) or only SDBS coated SWCNTs (appendix, Fig 2.10.).

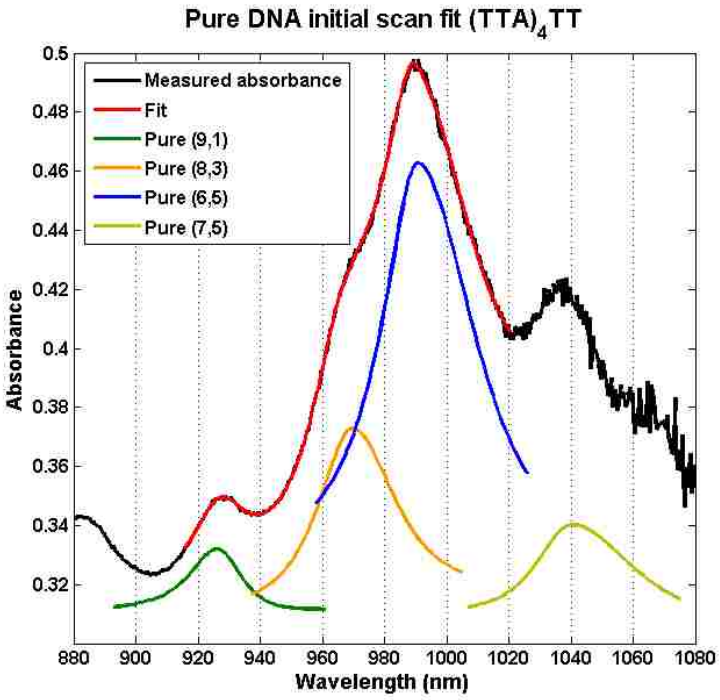


Figure 2.3 Fitted data for purely DNA-coated SWCNTs before surfactant reaction

Decomposition of measured absorbance of purely DNA-coated SWCNTs into contributions from four SWCNT species for $(TTA)_4TT$ – SWCNT dispersion. This is the initial mixture before the exchange reaction. Note that the although the influence of the (7,5) peak is included in the fit, the raw data is fitted only upto 1020 nm because for higher wavelengths the spectrum is influenced by tails of peaks at higher wavelengths, outside the range of the instrument used.

c) Fit to spectra obtained during intermediate stages of transformation from DNA to SDBS coating as linear combinations of contributions from constituents of the pure DNA and SDBS-coated spectra (Appendix, Fig 2.4.).

d) Let $D_{(m,n)}$ and $S_{(m,n)}$ be the coefficients corresponding to contributions from DNA and SDBS coated SWCNTs. Then, the value of $D_{(m,n)}$ decreases and that of $S_{(m,n)}$ increases from their initial to final values (Figure 2.4.). At higher temperatures we observe that the decay can be represented well by a single exponential (Figure 2.4.a) but at lower temperatures there is clear evidence of two processes, each with its own characteristic decay time. We therefore fit the time-evolution of $S_{(m,n)}$with a first order rate equation with two exponential terms:

$$S_{m,n} = a_1 * e^{-k_1 * t} + a_2 * e^{-k_2 * t} + S_{final} \quad (2.1)$$

Where k_1 and k_2 are rate constants in sec^{-1} and *t* is time in seconds.

Let the first, faster, reaction step be represented by a_1 and k_1, and the second slower reaction step be represented by a_2 and k_2. At higher temperatures, both reactions are completed within the first few minutes, and we can determine S_{final} for a particular pair of DNA sequence and SWCNT. This value of S_{final} is found for all the temperatures for which the reaction is complete for each DNA sequence and SWCNT pair. For example in figure 2.4. (a), we see that at 45 degrees C, for $(TAT)_4$ on (6,5) SWCNT, the value of $S_{m.n}$ becomes constant and equal to about 4.45 once both reaction steps are completed. Then the mean and standard deviation is calculated for S_{final}. So we can write equation (2.1) for $(TAT)_4$ on (6,5) as

$$S_{6,5} = a_1 * e^{-k_1 * t} + a_2 * e^{-k_2 * t} + S_{final}^{6,5} \quad (2.2)$$

We then fit the five parameter equation (2.1) at various temperatures for $(TAT)_4$ – (6,5) to the data on the SDBS coefficient with time, using the curvefit toolbox available in Matlab®. For the value of S_{final} we used the mean as guess value and twice the standard deviation on each side of the mean as lower and upper bounds for S_{final}. The values for k_1 and k_2 are constrained to take positive values. The values for a_1 and a_2 are constrained to take negative values in case of the SDBS coefficients and positive values in case of DNA coefficients. An example is given in Figure 2.4. (b) for $(TAT)_4$ – (6,5) SWCNT at 35 degrees C. Another example is given in Appendix Figure 2.5. (a) and (b) for the DNA coefficient.

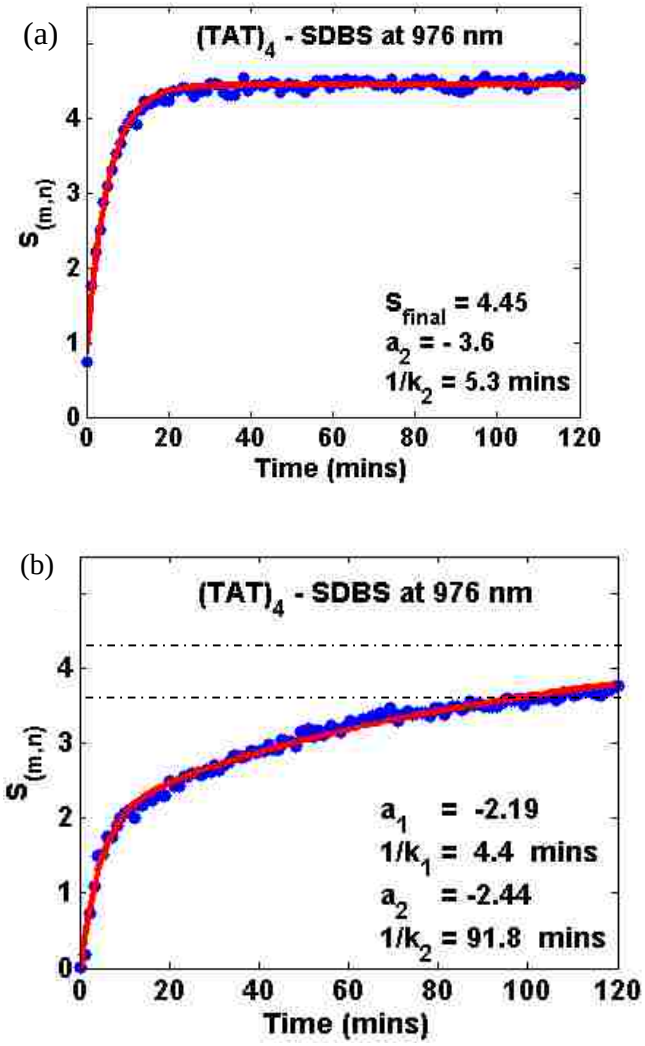


Figure 2.4 Fit to obtain decay time from surfactant exchange experiment

(a) Fit to obtain S_{final} for $(TAT)_4$ – (6,5) SWCNT at 45 degrees C where the reaction is completed in the experimental time.
(b) Fit to second order rate equation for the SDBS coefficient at 35 degrees C, giving the decay times and hence the rate constants k_1 and k_2

We conducted the surfactant exchange reactions at various temperatures to obtain rate constants as a function of temperature. Our observation of a fast and a slow process is consistent with the observations made Roxbury et al.[26]. There we interpreted the fast process as being due to quick transformation of those DNA-coated SWCNTs (to SDBS

coating) that have pre-existing defects. The slow process was interpreted as representing the rate of nucleation of new defects. We adopt the same interpretation here. The two-step process used to represent the kinetics provided reliable values for k_2 as a function of temperature. However, in many cases it did not provide reliable values for k_1 because the first step is often very rapid (e.g., Figure 2.4.a).

We therefore focus our attention on the second, slower process and, henceforth, we will refer to the rate constant of the second step 'k_2' as 'k'. In appendix we present data on the $(CCA)_{10}$ sequence to show that the fast and the slow process are correlated, supporting our focus on only k_2 as a measure of binding affinity. We plot ln (k/T) versus (1/T). According to the Eyring's activated rate theory for a desorption process, [59,60]

$$\ln\left(\frac{k}{k_B \mathrm{T}/h}\right) = \ln(\omega) - \frac{\Delta \mathrm{H}^{\ddagger}}{\mathrm{k_B T}} + \frac{\Delta \mathrm{S}^{\ddagger}}{\mathrm{k_B}} \tag{2.3}$$

Where k is rate constant in sec^{-1}, T is absolute temperature, $\Delta H^{\ddagger}$ is activation enthalpy, $\Delta S^{\ddagger}$ is activation entropy, h is Planck's constant, and k_B is the Boltzmann constant.

The constant ω is the number of activated complexes per SWCNT. We explicitly separate it because the rate we measure is of transforming a SWCNT from DNA to SDBS coverage, whereas the activation enthalpy and entropy are for a single activated complex. Since the entire SWCNT transforms from DNA to SDBS coverage rapidly when a single defect on its side wall is nucleated, the measured rate k is the product of the rate of defect nucleation

multiplied by the number of defects per SWCNT. The activation enthalpy can be obtained from the slope of ln (*k*/T) versus 1/T in the Eyring plot.

2.3. **Results and Discussion**

2.3.1. Recognition sequences on various SWCNTs

In order to understand how the activation energy differs for the same recognition DNA sequence on its partner SWCNT and non-partner SWCNT species, we studied the recognition sequence for the (6,5) SWCNT, i.e., $(TAT)_4$, for the (9,1) SWCNT, i.e., $(CCA)_{10}$, and for the (8,3) SWCNT, i.e., $(TTA)_4TT$ on each of the three SWCNT species (6,5), (9,1) and (8,3).

As an example, in Figure 2.5. , the time resolved absorbance spectra show that for $(TAT)_4$ covered SWCNT, at 15 degrees C, the (9,1) peak shifts considerably within the first few minutes, whereas the (6,5) and (7,5) peaks do not show much shift even after two hours. The lower the rate constant of the exchange reaction of a DNA sequence with SDBS for the surface of a particular SWCNT species, the higher is the activation energy. That means $(TAT)_4$ is significantly more weakly bound to (9,1) compared to (6,5) and (7,5) SWCNTs. More quantitative information is obtained by examining how the relative populations of SDBS and DNA-coated SWCNTs change with time.

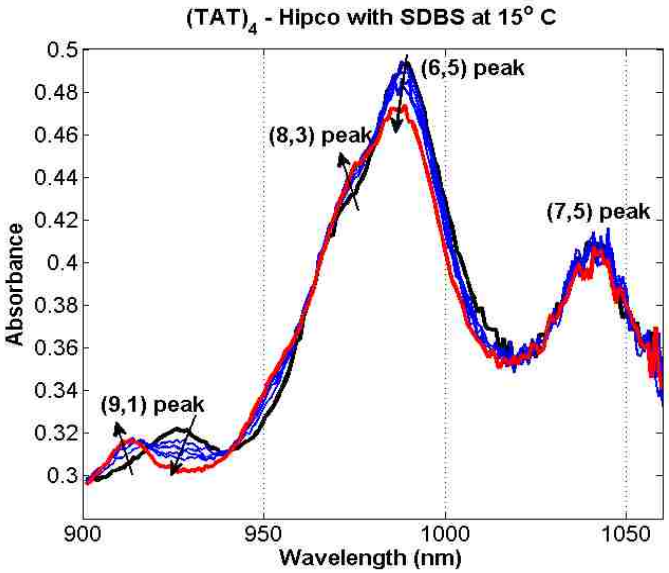


Figure 2.5 Time resolved absorbance spectra showing difference in rate of peak shift at different wavelengths

Time resolved absorbance spectra of $(TAT)_4$ – dispersed SWCNTs displaced by SDBS at 15 degrees C, showing that the (9,1) SWCNT peak shifts within the first several minutes, but the (6,5) and (7,5) peaks do not shift much even after two hours. The black line shows the initial spectrum, successive scans taken every 5 minutes are shown for the first twenty minutes (blue lines), and the red line shows the final scan after two hours.

We find that each of the three DNA sequences binds strongest to its recognition partner SWCNT. This difference is largest and clearest for $(TAT)_4$, and present but relatively weak for the other two DNA sequences (Figure 2.6). Figure 2.6. (a) shows the Eyring plot for $(TAT)_4$-dispersed SWCNTs (left panel) and the average ratio of rate constant normalized by its value for the $(TAT)_4$-(6,5) recognition pair (right panel). Consistent with the data

presented in Figure 2.5, Figure 2.6. (a) shows that the rate of $(TAT)_4$ removal from its partner SWCNT (6,5) is significantly lower than from (8,3) and especially (9,1) SWCNTs. That is, the affinity of $(TAT)_4$ is significantly greater to (6,5) than to (8,3) or (9,1). The large difference in the rate constant of removal of $(TAT)_4$ from (6,5) and (9,1) is striking, given that they have exactly same diameter (but differ in their chirality and electronic properties).

Figure 2.6. (b) and (c) show the Eyring plots and relative rate constant ratios for $(CCA)_{10}$and $(TTA)_4TT$ – dispersed SWCNTs, respectively. We find that $(CCA)_{10}$ binds more strongly to its recognition partner SWCNT (9,1) than to the other two non-partner SWCNTs. However, the differences are relatively small compared to the $(TAT)_4$ case. Figure 2.6. (c) shows that $(TTA)_4TT$ binds very slightly more strongly to its partner SWCNT (8,3) as compared to the other three SWCNT species. The order of binding affinity is not discernable for the non-partner SWCNT species in this case.

The data in Figure 2.6. suggest that for a given DNA sequence, the slope, which represents activation enthalpy, does not change with SWCNT type. (See also appendix Figure 2.13. in which we fit independent lines to data for each DNA-CNT combination.) The data also suggest that slopes and therefore enthalpies change when one changes the type of DNA sequence. We test this hypothesis by fitting three straight lines to each of the data sets in Figure 2.6. For a given DNA sequence, the lines are constrained to have the same slope and three different intercepts, and these four parameters are determined by least-squares fitting. This procedure yields a value for activation enthalpy of each DNA sequence along with a confidence interval for it estimate.

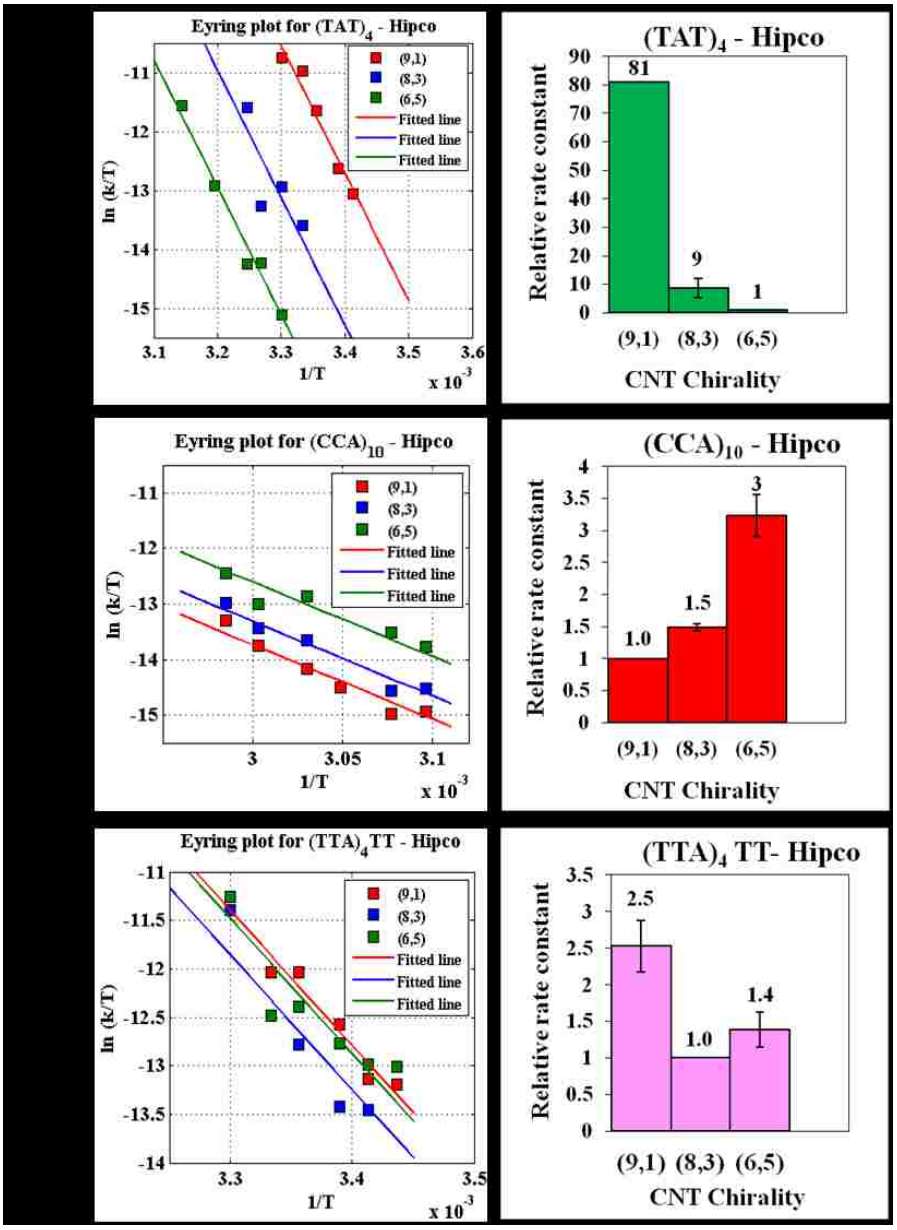


Figure 2.6 Eyring plots and relative rate constants for recognition sequences on different SWCNTs

(left) Eyring plot for surfactant exchange with (a) $(TAT)_4$ wrapped SWCNTs (b) $(CCA)_{10}$ wrapped SWCNTs, and (c) $(TTA)_4TT$ wrapped SWCNTs, (right) Relative rate constant (ratio with respect to rate constant of the recognition DNA sequence – partner SWCNT species at the same temperature). The lower the rate constant, the higher is the activation energy of the DNA sequence to the SWCNT species. Here the error bars are the standard errors. Note that each sequence binds strongest to its recognition partner SWCNT.

Table 2.1 contains values of the activation enthalpy $\Delta H^{\ddagger}$, obtained from the slope of the linear fits in Figure 2.6. With 90% confidence we find that the activation enthalpy of $(TAT)_{4,}$estimated to be 72 ± 11.82 k_BT, is significantly different from the activation enthalpy of $(CCA)_{10}$ and $(TAT)_4TT$;the latter two are not statistically distinguishable from each other. Also, the enthalpy value for the (TAT) family reported by Roxbury et al.[26] of 29 ± 7.47 k_BT is significantly lower than what we found here. The samples used in that previous study were first subjected to size exclusion chromatography (SEC). The observation of sample loss for weakly bound DNA-CNT combinations and the significant difference in activation enthalpies both indicate that the DNA-CNT hybrid, at least for these relatively short oligomers, is quite plastic. That is, SEC very likely alters the structure and concentration of DNA adsorbed on the SWCNT, perhaps by stripping away loosely bound DNA strands and creating a population of defects not present in the original sample.

Table 2.1 Activation enthalpies obtained from Eyring plots (Figure 2.6)

Sequence	**$\Delta H^{\ddagger}/k_B$ T (300K) with 90 % CI**
$(TAT)_4$ ((6,5) recognition sequence)	72 ± 11.82
$(CCA)_{10}$ ((9,1) recognition sequence)	44 ± 5.98
$(TTA)_4TT$ ((8,3) recognition sequence)	46 ± 9.67

For a given DNA sequence taking $\Delta H^{\ddagger}$ to be the same for different SWCNTs, for two different SWCNT species we can write,

$$\ln\left(\frac{k_1}{\frac{k_B \mathrm{T}}{h}}\right) = \ln(\omega_1) - \frac{\Delta H^{\ddagger}}{k_B T} + \frac{\Delta S_1^{\ddagger}}{k_B} \tag{2.4}$$

$$\ln\left(\frac{k_2}{\frac{k_B \mathrm{T}}{h}}\right) = \ln(\omega_2) - \frac{\Delta H^{\ddagger}}{k_B T} + \frac{\Delta S_2^{\ddagger}}{k_B} \tag{2.5}$$

Or,

$$k_B \mathrm{T} \ln\left(\frac{k_1}{k_2}\right) = k_B \mathrm{T} \ln\left(\frac{\omega_1}{\omega_2}\right) + \mathrm{T}\left(\Delta S_1^{\ddagger} - \Delta S_2^{\ddagger}\right) \tag{2.6}$$

If we interpret the left hand side of equation (2.6) as a difference in free energy, ΔΔG, we see that within one DNA type, this difference between different SWCNTs is due to difference either in defect density, $k_B \mathrm{T} \ln\left(\frac{\omega_1}{\omega_2}\right)$, or activation entropy, $\mathrm{T}\left(\Delta S_1^{\ddagger} - \Delta S_2^{\ddagger}\right)$.

Table 2.2 provides values of $\Delta\Delta G$ for the three different DNA sequences obtained by averaging over different temperatures.

Table 2.2 $\Delta\Delta G/ k_BT$ (at 300K, subtracted from the recognition DNA sequence – partner SWCNT) *

	(6,5)	**(9,1)**	**(8,3)**
$(TAT)_4$	Reference →	4.39	1.93 ± 0.51
$(CCA)_{10}$	1.15 ± 0.11	← Reference →	0.40 ± 0.03
$(TTA)_4TT$	0.25 ± 0.17	0.89 ± 0.13	← Reference →

* In the case of (9,1) versus (6,5) comparison, since we can calculate the ratio at only one temperature, we do not provide an estimate of the standard error.

Table 2.2 shows that the difference in activation free energy between weakly and strongly binding sequences is in the range 0.25 to 4.39 k_BT. A DNA sequence with its partner recognition SWCNT has the strongest binding for all three sequences and chiralities. For a given DNA sequence, based on the fact that activation enthalpy does not vary with SWCNT type, if we make the additional conjecture that the activation entropy is coupled to the activation enthalpy, it would also not vary with SWCNT type. Then, it would follow that for a given DNA sequence variation in activation energy with different SWCNTs arises due to differences in the number of sites per SWCNT at which a transition state can occur.

Removal of DNA requires its bases to be sufficiently displaced for the insertion of water molecules between them and the SWCNT surface. This event, which presumably

corresponds to the transition state, needs to occur before the surfactant molecule can replace DNA on the SWCNT surface. If this is the case, we can compare the value of $\Delta H^{\ddagger}$ to the free energy of removal of a base from the SWCNT surface. The free energy of a DNA base removed from a graphite surface as measured by single molecule force spectroscopy is in the range 7-11 k_BT.[27] Comparing to values of $\Delta H^{\ddagger}$ given in Table 2.1., suggests that a cluster of about 4-6 bases are involved in the activated complex.

2.3.2. Relative activation energy of closely related DNA sequences

The results of the previous section establish that DNA sequences bind more strongly on their partner SWCNT species compared to the non-partner species. Roxbury et al.[26] have shown that the dependence of the activation energy of the DNA on the SWCNT surface with the length of the DNA strand is not monotonic for the smaller sequences (10 to 14mers). They also showed that for longer sequences (30mers), the binding is significantly stronger and not as strongly dependent on sequence. Hence, we can say that the recognition ability of DNA sequences for SWCNT species is not due to the absolute value of activation energy. It comes, rather, from the difference in the binding between compositionally similar sequences and SWCNTs.

To study the transition from weak but discriminative binding of shorter sequences to the strong but non-discriminative binding of longer sequences, we studied eleven different DNA sequences in the$(TAT)_n$ family, ranging from the 11mer $(TAT)_3TA$ to the

21mer $(TAT)_7$, on three different SWCNT species. The SDBS exchange reaction is carried out at two temperatures, 30^0C and 40^0C. Figure 2.7 plots time constants for the three SWCNT types as a function of sequence length. (All the time constants mentioned in this section are relative ratios to that of $(TAT)_4$ on its partner SWCNT (6,5) at the same temperature.) Two features of the behavior of time constant with sequence length are evident. Firstly, the relative time constants vary periodically with sequence length for the short sequences with a period of two or three bases. Secondly, imposed on this periodicity is an overall increase of time constant with sequence length. It is apparent that this family of sequences has comparatively higher activation energy to the SWCNT (6,5) than to the other two SWCNT species. However, for shorter strands there is strong sequence-dependence. For example, whereas $(TAT)_4$ binds stronger to (6,5) than to either (9,1) or (8,3), its immediate compositional neighbors $(TAT)_4T$ and $(TAT)_3TA$bind relatively weakly to all three SWCNTs.

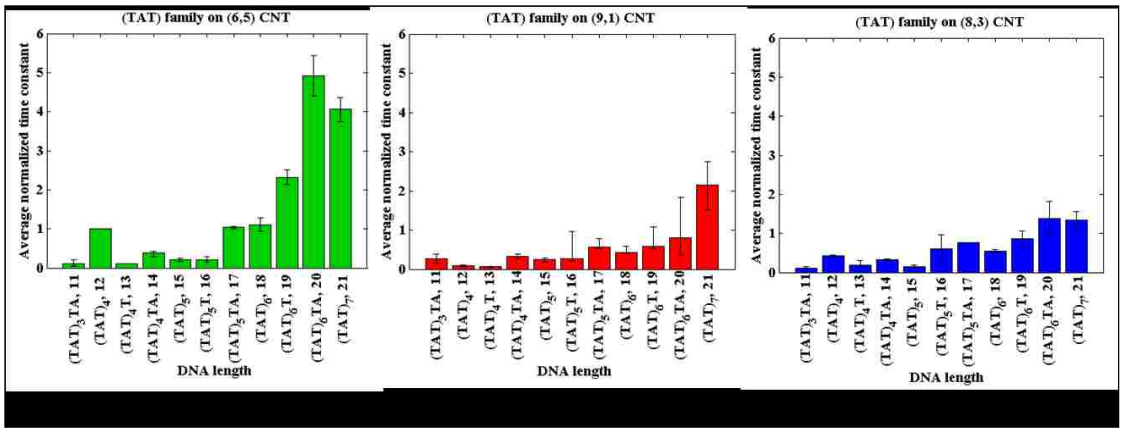


Figure 2.7 Time constants for removal of closely related DNA sequences as function of DNA length

Time constants for eleven closely related DNA sequences on (a) (6,5), (b) (9,1),and (c) (8,3) SWCNT species relative to the time constant for $(TAT)_4$ on its partner SWCNT (6,5) at the same temperature. The error bars are the minimum and maximum values of time constants from various experimental measurements.

The phenomenon of increase of activation energy with the length of DNA sequence is consistent with the model proposed by Roxbury et al.[26] for the surfactant exchange process. As they showed, DNA removal proceeds by the nucleation of a defect on the SWCNT outer surface followed by rapid substitution of the DNA by SDBS. The nucleated defect corresponds to the activated state, and this evidently involves fewer DNA bases than are in each entire chain. (For example, we know that the free energy for removal of one DNA base from graphite surface is 7-11 k_BT.[27] Assuming the value is 10 k_BT for a single base and multiplying by number of the bases in the DNA strand, the activation enthalpy for removal of an entire 30-mer would need to be > 300 k_BT.)That is much higher than the

activation enthalpy values we found. So it is reasonable to suppose the activated state corresponds to a defect near the ends of the DNA. Roxbury et al.[26] showed that this presumed defect nucleation occurs homogeneously on the SWCNT surface and not just on the ends of the SWCNT. This starts the process of replacing the DNA with SDBS. Now if we assume that the defect nucleation happens generally at the ends of the DNA strands, then shorter DNA strands provide more sites and hence greater probability for defect nucleation on any given SWCNT than do the longer strands. Hence, shorter sequences in general will have a higher rate constant, even if the activated state is the same. It is possible that exceptions to this rule, for example $(TAT)_4$ on (6,5) achieve higher strength by a mechanism such as assembly into a structure that resembles a longer, effectively ligated, DNA chain, with fewer associated defects per unit length of the SWCNT.

2.4. **Conclusions**

We have studied the activation energy for removal by as surfactant molecule of three DNA sequences ($(CCA)_{10}$; $(TTA)_4TT$; $(TAT)_4$) on three different SWCNT species, ((9,1);(8,3);(6,5)). Of the nine combinations, three are recognition pairs (($(CCA)_{10}$-9,1);($(TTA)_4$TT-8,3);($(TAT)_4$-6,5)) as determined by their ability to sort a particular SWCNT from a mixture. We found that recognition sequences have significantly higher activation energy on their partner SWCNT than on the non-partner SWCNT species. Specifically, $(TAT)_4$ has much lower activation energy on the (9,1) SWCNT than on its recognition partner SWCNT (6,5). Conversely, we also found that $(CCA)_{10}$ has higher

activation energy for its partner SWCNT (9,1) than for (6,5). The (6,5) and (9,1) SWCNTs have the same diameter but different chirality and electronic properties. We calculated the lateral energy barrier fraction for a DNA base on a graphene surface (neglecting curvature effects for SWCNT) to be about 0.01 k_BT per base (see appendix Figure 7), which is small. We therefore conclude that the lateral energy barriers are not sufficient for the DNA strand to be cognizant of the chirality of the SWCNT based purely on the arrangement of the C atoms. But it must be noted that recent experiments by Geyou et al. have shown that SWCNT enantiomers wrapped with the same DNA sequence can be separated, hinting that certain DNA sequences do have the ability to differentiate between the handedness of the SWCNT. Thus the difference of activation energy of $(TAT)_4$ to (6,5) and (9,1) can be attributed to difference in SWCNT electronic properties (which are typically not represented in force-field-based molecular simulations).

We also studied increasing lengths of sequences in the $(TAT)_n$ family from the 11mer $(TAT)_3TA$ to the 21mer $(TAT)_7$ on three different SWCNT species. We found that for these sequences, the activation energy varies periodically with DNA sequence length. That is, certain sequences have higher activation energy than their closely related neighbor sequences obtained by adding or removing just one base from the end. This periodic variation is superimposed on an overall trend of increasing activation energy with increasing sequence length. We have interpreted this to mean that DNA replacement by surfactant molecules occurs by the formation of an activated complex at the ends of the DNA sequence.

Overall, experimental data on the nine combinations of DNA sequences and SWCNTs showed that for a given DNA sequence, the activation energies do not change with SWCNT type. The difference in activation energy for a given sequence and different SWCNTs can be attributed to some combination of the number of defects per SWCNT and activation entropy of the activated state. Through molecular simulation, we have previously shown that a given DNA strand on different SWCNTs, as well as different DNA strands on the same SWCNT, form different ordered structures [26-28], showing that a structural basis can exist for differences in activation energy.

2.5. **Acknowledgement**

We acknowledge the assistance of Miss Ke Xue of Lehigh University in conducting the experiments and several useful discussions with Drs. Ming Zheng, Jeffrey Fagan, and Constantine Khripin. This work was supported by the National Science Foundation through grant CMMI-1014960.

2.6. **Appendix**

2.6.1. Photoluminescence spectra of DNA and SDBS-coated SWCNT dispersions

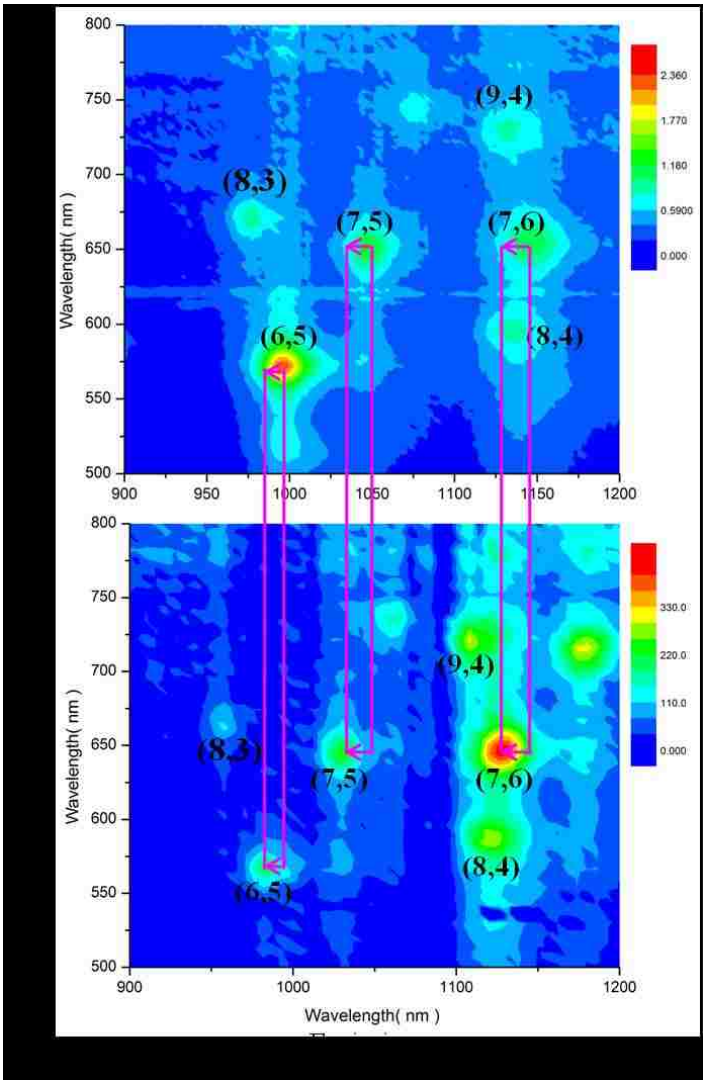


Figure 2.8 Photoluminescence spectra confirming solvatochromic shifts upon exchange of ssDNA by SDBS.

(Top) Fluorescence map for DNA coated SWCNT (before surfactant exchange) (Bottom) Fluorescence map for SDBS coated SWCNT (after surfactant exchange). These measurements confirm the solvatochromic shift in absorbance that we used to monitor kinetics.

Table 2.3 Peak positions for (a) absorbance spectroscopy and (b) fluorescence spectroscopy

(a)

CNT Chirality (n,m)	Pure DNA - SWCNT	Pure SDBS-CNT
	Wavelength (nm)	Wavelength (nm)
(9,1)	926	912
(6,5)	990	976
(8,3)	968	952
(7,5)	1042	1024

(b)

CNT Chirality (n,m)	Pure DNA - SWCNT	Pure SDBS-CNT
	Wavelength (nm)	Wavelength (nm)
(9,1)	930	918
(6,5)	990	984
(8,3)	972	960
(7,5)	1044	1029

2.6.2. Methods for decomposition of absorbance spectra

a) Fit for pure species:

First we fitted the initial (pure DNA-CNT) reaction mixture and the final (pure SDBS-CNT) reactant mixture as a sum of the individual absorption spectra, $g_{(m,n)}(x)$, of purified chiralities (in our case the four chiralities (9,1), (8,3), (6,5) and (7,5)).

$$g_{mixture}(x) = e + \sum_{(m,n)} d_{(m,n)}\left[g_{(m,n)}(x) - c\right] \tag{2.7}$$

Here, $g_{mixture}(x)$ is the function representing the spectrum of the mixture, and $d_{(m,n)}$ are coefficients determined by the fitting procedure. The function $g_{(m,n)}(x)$ used was:

$$g_{(m,n)}(x) = \begin{cases} c + a/\left(1 + \left((x-p)/b_l\right)^2\right)^{\alpha_l} & x \le p \\ c + s + (a-s)/\left(1 + \left((x-p)/b_r\right)^2\right)^{\alpha_r} & x \ge p \end{cases} \tag{2.8}$$

It is continuous and smooth at the point of joining the left and right hand portions. The parameters have the following interpretation: *c* represents a baseline; *a* represents the amplitude; *s* represents the difference in baseline on the left and right hand sided; *p* represents the peak wavelength; b_l and b_r represent spread on the left and right sides of the peak, respectively, and α_l and α_r are exponents that also allow asymmetry. Note that by subtracting 'c' from the pure-species fitting function, we absorb it into a single fitting baseline parameter, 'e'.

In order to get good fits, we used previously known values for c, a, s, α_l, α_r for the four chiralities from Tu et al. We fitted the coefficient $d_{(m,n)}$, p, b_l and b_r for the pure DNA- SWCNT fits for each of the four chiralities. As an example, the fit for the pure $(TTA)_4TT$– SWCNT is shown in Figure 2.9.

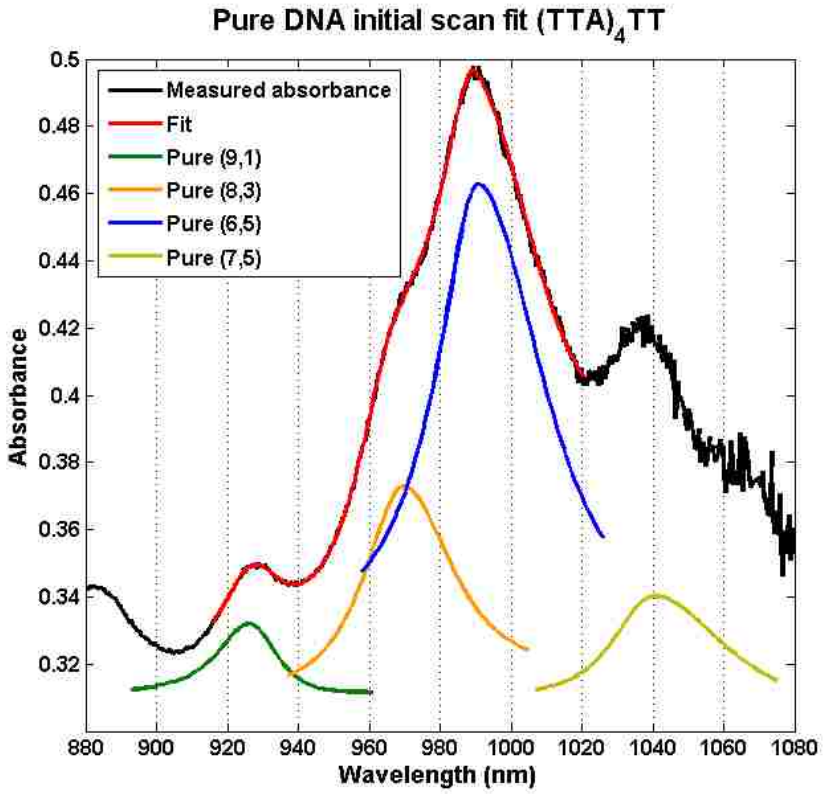


Figure 2.9 Decomposition of measured absorbance spectra of initial state

Decomposition of measured absorbance into contributions from four SWCNT species for $(TTA)_4TT$– SWCNT dispersion which is the initial mixture before exchange reaction

For the pure SDBS-CNT fits, since the individual E_{11} absorption spectra of individual purified SDBS-coated chiralities were not available, we assumed that the basic shape of

the peak of a particular chirality does not change and hence parameters like c,s, α_l, α_r could be kept same as that of the DNA-coated species of the same chirality. We fitted the coefficient $d_{(m,n)}$, p, b_l, b_r, a (to account for possible change in height) for the pure SDBS-CNT fits for each of the four chiralities. The fit for the pure SDBS – SWCNT is shown in Figure 2.10.

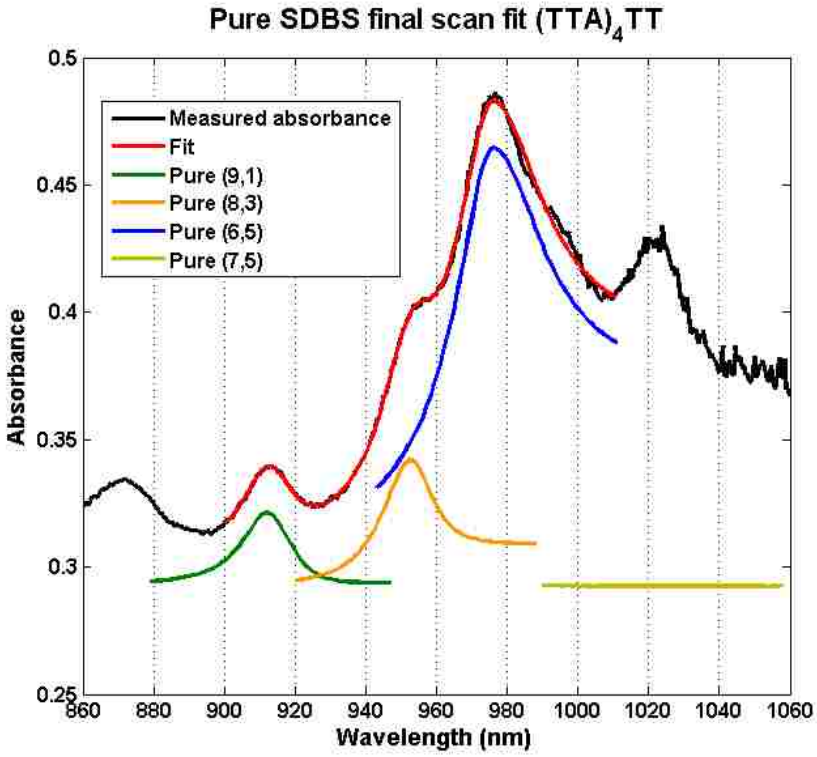


Figure 2.10 Decomposition of measured absorbance spectra of final state

Decomposition of measured absorbance into contributions from four SWCNT species for SDBS– SWCNT dispersion which is the final mixture after exchange reaction

b) Fit for intermediate species:

Now, to fit the intermediate data, which is a mixture of DNA-coated and SDBS-coated SWCNTs, we assumed that all the intermediate absorbance spectra can be considered a linear combination of the pure DNA-coated and SDBS-coated SWCNTs for each chirality. So for any wavelength, *x*, evolution of its absorbance is a sum of contributions from all 8 species:

$$g_{mixture}(x) = e + \sum_{(m,n)} D_{(m,n)} d_{(m,n)} \left[g_{(m,n)}(x) - c \right] + \sum_{(m,n)} S_{(m,n)} d_{(m,n)} \left[g_{(m,n)}(x) - c \right]$$

Where $D_{(m,n)}$ are all the four coefficients for the DNA coated chiralities and $S_{(m,n)}$ all the four coefficients for the SDBS coated chiralities. As the surfactant exchange takes place, $D_{(m,n)}$ all decrease and $S_{(m,n)}$ all increase with time till the reaction is complete. Because for the (7,5) SWCNTs, there will be some influence due to adjacent peaks at higher wavelengths not measured in our experiments, we did not include it in the range of wavelengths fitted. Figure 2.11 shows an example of a fit to a spectrum intermediate between purely DNA and SDBS coated SWCNTs. Figure 2.12 shows two examples of the decay in $D_{(m,n)}$ with time. At the lower temperature, it is evident that there are two separate processes each with its characteristic decay time. At the higher temperature, the first process proceeds too rapidly to be captured and the data fit a single exponential decay.

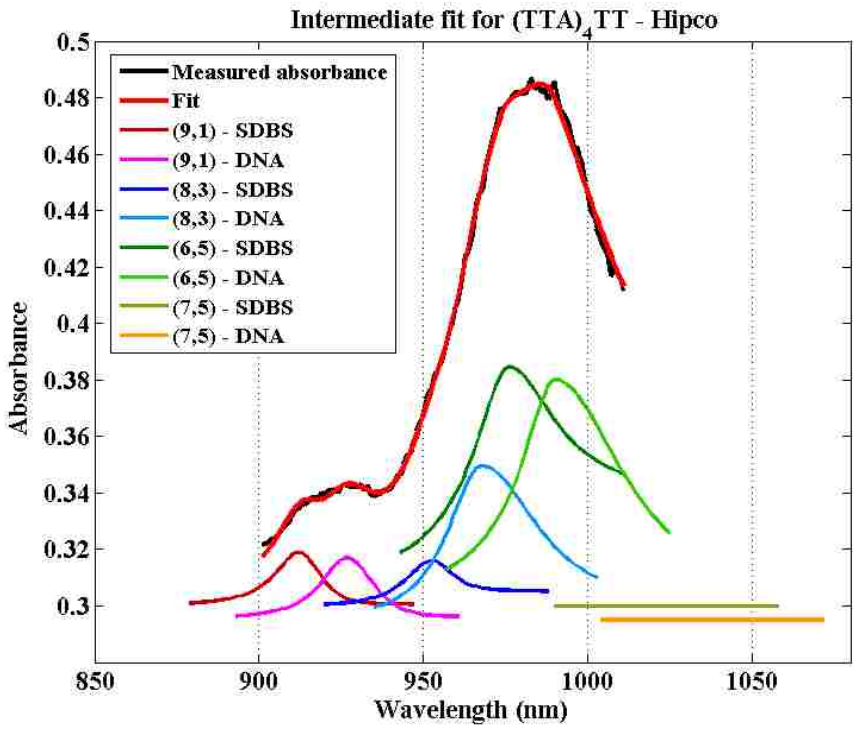
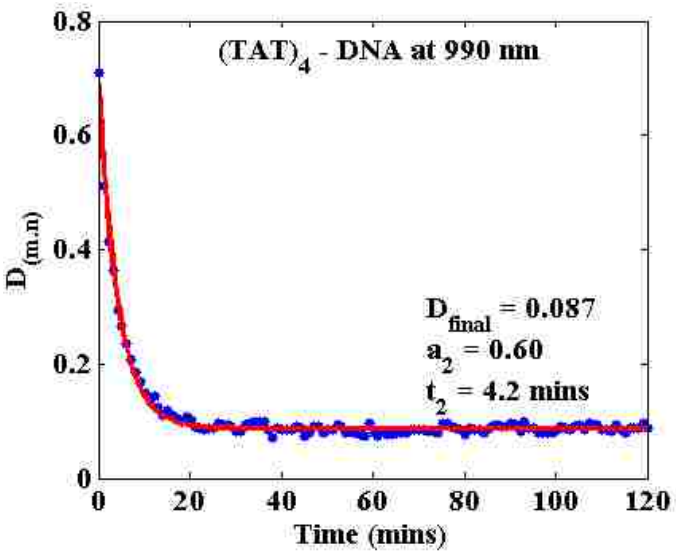
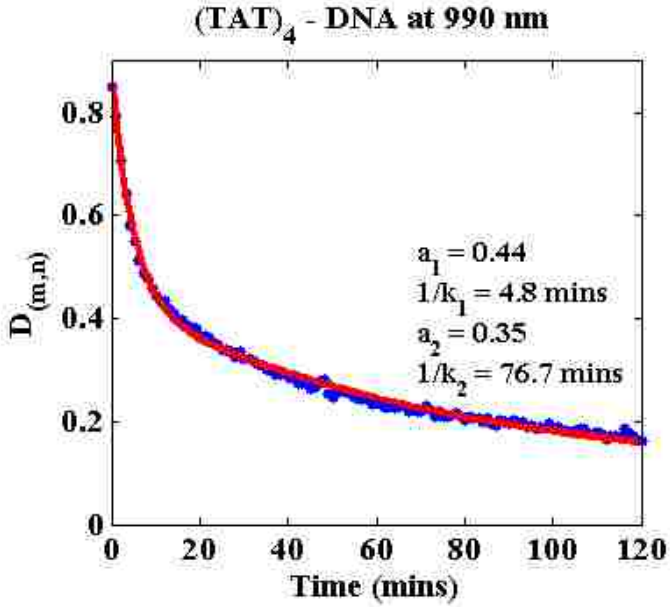


Figure 2.11 Intermediate Fit for $(TTA)_4TT$ – Hipco at 27 degrees C



(a)



(b)

Figure 2.12 Fit to obtain rate constant of DNA removal from SWCNT surface

(a) At 45 degrees C the decay in $D_{(m,n)}$for $(TAT)_4$ – (6,5) SWCNT can be fitted well by a single exponential – the faster first step is too rapid to be captured in the measurement. The fit provides D_{final} at 45 degrees C as the reaction is completed in the experimental time.
(b) Fit to second order rate equation for the DNA coefficient at 35 degrees C, yields the decay times and hence the rate constants k_1 and k_2.

Correlation between the two kinetic constants k_1 and k_2:

To consider if and how k_1 and k_2 are related, consider equation 2 that connects the rate constant with the activation enthalpy and activation entropy

$$\ln\left(\frac{k}{k_B\mathrm{T}/h}\right) = \ln(\omega) - \frac{\Delta\mathrm{H}^{\ddagger}}{\mathrm{k_B T}} + \frac{\Delta\mathrm{S}^{\ddagger}}{\mathrm{k_B}}$$

Now we can rewrite this as

$$\frac{k}{\frac{k_B\mathrm{T}}{h}} = \omega * \exp\left(\frac{-\Delta\mathrm{H}^{\ddagger}}{\mathrm{k_B T}} + \frac{\Delta\mathrm{S}^{\ddagger}}{\mathrm{k_B}}\right)$$

Applying this to the two rate processes,

$$\frac{k_1}{\frac{k_B\mathrm{T}}{h}} = \omega_1 * \exp\left(\frac{-\Delta\mathrm{H}_1^{\ddagger}}{\mathrm{k_B T}} + \frac{\Delta\mathrm{S}_1^{\ddagger}}{\mathrm{k_B}}\right)$$

$$\frac{k_2}{\frac{k_B\mathrm{T}}{h}} = \omega_2 * \exp\left(\frac{-\Delta\mathrm{H}_2^{\ddagger}}{\mathrm{k_B T}} + \frac{\Delta\mathrm{S}_2^{\ddagger}}{\mathrm{k_B}}\right)$$

Hence,

$$\frac{k_1}{\mathrm{k_2}} = \frac{\omega_1}{\omega_2} * \exp\left(\frac{-\Delta\mathrm{H}_1^{\ddagger}}{\mathrm{k_B T}} + \frac{\Delta\mathrm{S}_1^{\ddagger}}{\mathrm{k_B}} + \frac{\Delta\mathrm{H}_2^{\ddagger}}{\mathrm{k_B T}} - \frac{\Delta\mathrm{S}_2^{\ddagger}}{\mathrm{k_B}}\right)$$

$$\ln\left(\frac{k_1}{\mathrm{k_2}}\right) = \ln\left(\frac{\omega_1}{\omega_2}\right) + \left(\frac{-\Delta\mathrm{H}_1^{\ddagger}}{\mathrm{k_B T}} + \frac{\Delta\mathrm{H}_2^{\ddagger}}{\mathrm{k_B T}} + \frac{\Delta\mathrm{S}_1^{\ddagger}}{\mathrm{k_B}} - \frac{\Delta\mathrm{S}_2^{\ddagger}}{\mathrm{k_B}}\right)$$

$$\ln\left(\frac{k_1}{\mathrm{k_2}}\right) = \left(\frac{-\Delta\mathrm{H}_1^{\ddagger} + \Delta\mathrm{H}_2^{\ddagger}}{\mathrm{k_B T}}\right) + \left(\ln\left(\frac{\omega_1}{\omega_2}\right) + \frac{\Delta\mathrm{S}_1^{\ddagger} - \Delta\mathrm{S}_2^{\ddagger}}{\mathrm{k_B}}\right)$$

Therefore a plot of ln (k_1/k_2) versus 1/T should be linear, with the slope giving $\frac{-\Delta H_1^{\ddagger}+\Delta H_2^{\ddagger}}{k_B}$ and the intercept giving $\ln\left(\frac{\omega_1}{\omega_2}\right)+\frac{\Delta S_1^{\ddagger}-\Delta S_2^{\ddagger}}{k_B}$. For the $(CCA)_{10}$ sequence we have sufficient data for both k_1 and k_2 to make this plot, which is shown below. Two conclusions can be drawn. Firstly, the plot of ln (k_1/k_2) versus 1/T yields a positive slope, showing that $\Delta H_2^{\ddagger} > \Delta H_1^{\ddagger}$. Secondly, the slopes are similar for different DNA-CNT combinations, showing that $\Delta H_2^{\ddagger} - \Delta H_1^{\ddagger}$ does not change, i.e., change in activation energy the fast process is correlated with change in activation energy for the slow process.

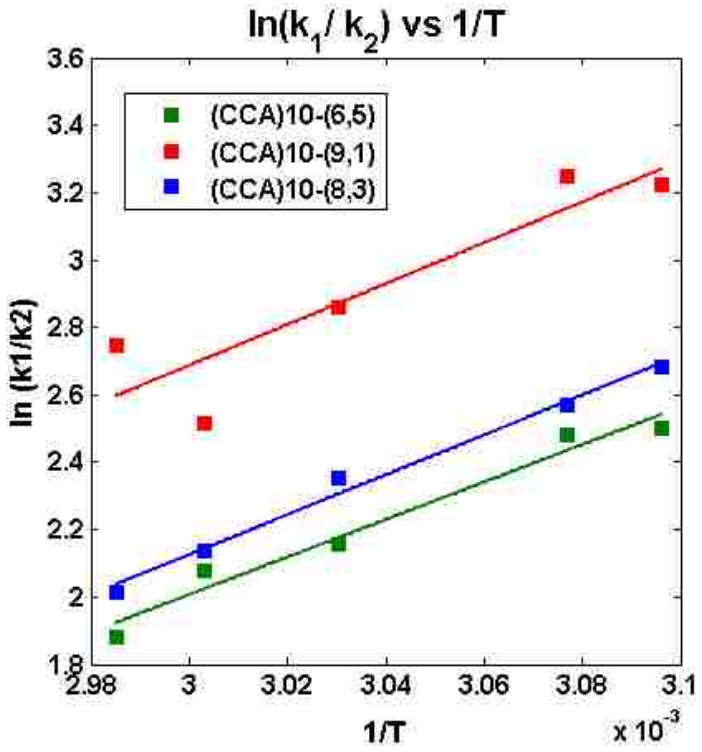


Figure 2.13 Plot of ln (k_1/k_2) versus 1/T shows a linear relation for each of the DNA – SWCNT combinations with similar slopes

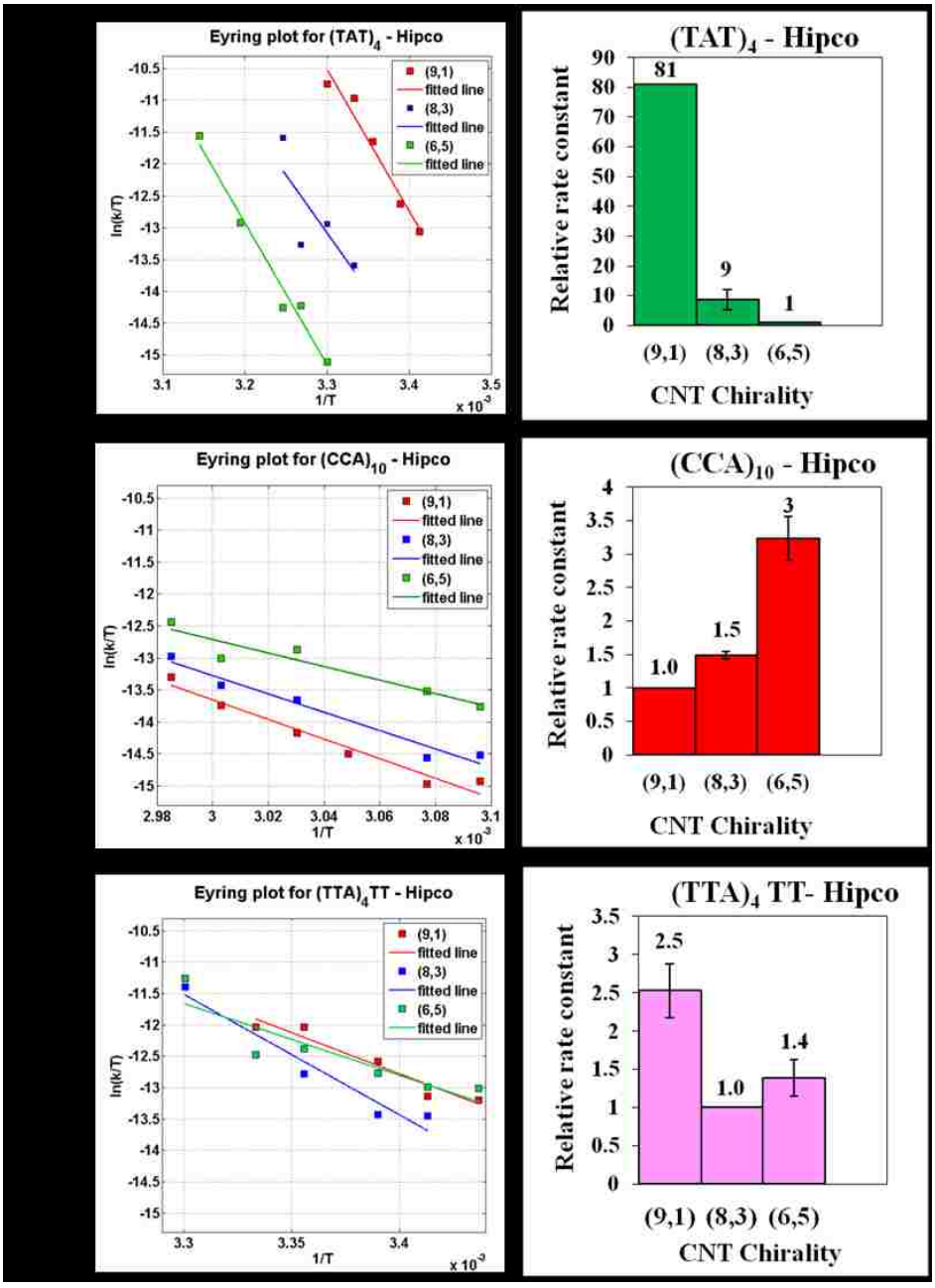


Figure 2.14 Eyring plots and relative rate constant for recognition sequences

(left) Eyring plots for surfactant exchange with (a) $(TAT)_4$ wrapped SWCNTs, (b) $(CCA)_{10}$ wrapped SWCNTs, (c) $(TTA)_4TT$ wrapped SWCNTs,

(right) Relative rate constant (ratio with respect to rate constant of the recognition DNA sequence – partner SWCNT species at the same temperature) The lower the rate constant, the higher is the activation energy of the DNA sequence to the SWCNT species. Here the error bars are the standard errors.

2.6.3. Corrugated Energy Profile above a Graphene Sheet

The difference of binding of a DNA sequence to different SWCNTs can depend on two factors: the difference in arrangement of the carbon atoms and the difference in the optical properties of the different SWCNT chiralities. In order to understand the undulating energy profile over a SWCNT surface, we simplify it by considering it as a flat graphene surface and ignoring the curvature effects. The carbon atoms are arranged in a in a graphene lattice as in SI Figure 12 (a). We took the distance between carbon atoms to be 1.42 A. A probe Carbon atom interacts with the graphene lattice via an L-J potential:

$$\frac{V}{4\varepsilon} = \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}$$

where *r* is the distance between the probe atom and a graphene atom. We take σ=3.39 A. The total potential energy of interaction is a sum of this interaction for each pair of (probe atom, graphene atom). SI Figure 12 (b) shows a contour plot of potential energy in units of εas a function of *x* and *z*. The contour is for a value of *y* depicted by the dashed green line in the SI Figure 12 (a).

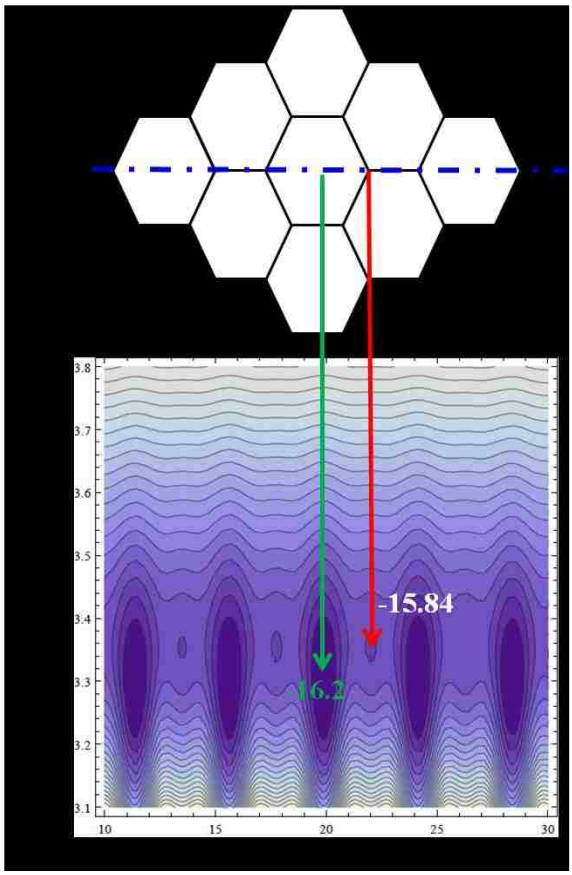


Figure 2.15 Potential on corrugated graphene surface

So difference between maximum and minimum (V/4ε) = 16.2 – 15.84 = 0.36

But ε = 0.339 kJ/mol

Hence potential energy barrier fraction (V/4ε) = 0.36/ ((16.2+15.84)/2) = 0.36/16.02 = 0.022472

Potential energy barrier fraction = 0.022472 * 4 * 0.339 = 0.030472 kJ/mol = 0.0123 k_BT

Chapter 3 Relative hydration of DNA- single walled carbon nanotube hybrids using aqueous two phase system

Since the difference in binding affinity and difference in partitioning can depend on DNA structure on the single walled carbon nanotube (SWCNT), we studied the partitioning of the various DNA sequences in an aqueous two phase system. DNA-SWCNT partitioning in the aqueous two phase is because of sensitive dependence of the free energy of hydration on the spatial distribution of hydrophilic groups in the DNA-SWCNT hybrid. In this way, the aqueous two phase process is at the same time a technique for separation and a method by which to evaluate and rank hydration or solvation free energy. We found that $(CCA)_{10}$ *on (6,5) SWCNT requires much higher amount of modulant to be moved from the relatively hydrophilic phase to the more hydrophobic phase as compared to* $(GT)_{15}$ *on (6,5) even though they are both 30mers, suggesting that the solvation energy depends greatly on the DNA sequence. We also found that various sequences with same length but different repeating units of two bases exhibit different hydration energies on the same SWCNT (6,5).*

3.1. **Introduction:**

Solute distribution in a two phase system depends on the relative solvation energy of the solute in the two phases, which in turn depends on the exact structure of the solute surface exposed to the phases. The single walled carbon nanotube (SWCNT) – single stranded DNA hybrid is essentially an amphiphilic system, with the DNA backbone being the hydrophilic area and the DNA bases as well as the SWCNT surface being

hydrophobic regions. Various simulation studies have suggested that the DNA bases adsorb onto the SWCNT surface while the backbone is away from the SWCNT surface and faces the surrounding aqueous phase.[21,22,26] This suggests that the hydrophobicity of the hybrid surface greatly depends on the exact coverage of the SWCNT by the DNA, which in turn depends on the exact structure of the DNA strand on the SWCNT surface. A polymer aqueous two-phase system consists of two separate but permeable aqueous phases which vary only slightly in their physical properties due to the difference in the polymer concentration in the two phases. The aqueous two phase system has been widely used for separation of biomolecules as the phases do not denature the biomolecules, the interfacial stress is much lower than in case of water – organic solvent system, leading to lesser chances of modification of the solute structure as it passes through the interface, and the small difference in hydration energy in these systems is ideal for separating solutes with very small structural differences.[61,62]

Recently, Khripin et al.[63] and Fagan et al.[64] showed that surfactant coated SWCNTs could be separated very effectively into various chiralities using a polymer aqueous phase system.[63,64] Further work by Geyou et al.[30] have shown that partitioning of DNA−SWCNT hybrids in a given polymer two-phase system is strongly sequence-dependent and can be further modulated by salt and polymer additives. SWCNT partition in the aqueous two phase system is determined by the SWCNT solvation energy difference between the two phases. Hence it is proposed that the DNA-SWCNT partitioning in the aqueous two phase is because of sensitive dependence of the free energy of hydration on the spatial distribution of hydrophilic groups in the DNA-SWCNT

hybrid. In this way, the aqueous two phase process is at the same time a technique for separation and a method by which to evaluate and rank hydration or solvation free energy.

3.2. **Methods:**

Since the difference in binding affinity and difference in partitioning can depend on DNA structure on the SWCNT, we studied the partitioning of the various DNA sequences in PEG-Dextran aqueous two phase system. We used polyvinyl pyrrolidone (PVP) as a modulant to move the DNA-SWCNT hybrids to top phase. The amount of PVP required to move the hybrid from relatively hydrophilic bottom phase to relatively hydrophobic top phase gives a measure of the hydrophobicity of the hybrid.

For dispersion of SWCNTs with a given DNA sequence, a total volume of 1 mL of the DNA and SWCNT mixture in phosphate buffer was sonicated in an ice bath for 90 minutes at a power level of 8 W. The SWCNT/DNA mass ratio was 1:1.5. After centrifugation at 16100g for 90 min, the supernatant of the DNA-SWCNT dispersion was collected. The dispersion was then passed through an Amicon 100kDa filter and resuspended in DI water three times in order to remove free DNA and the phosphate salts. The concentration of the dispersion was adjusted such that at 20 times dilution, the absorbance at 990 nm was ~0.5.

The aqueous two phase system consisted of 5 % (w/w) PEG (6 kDa) and 10 % (w/w) Dextran (70 kDa) in DI water. The total volume of the aqueous two phase system including the DNA-SWCNT dispersion and PVP solution was fixed at 500 uL. The volume of

dispersion added was fixed at 25 uL. Partitioning of SWCNTs in the aqueous two phase was obtained by vortex mixing of the mixture of PEG solution, Dextran solution, PVP solution and DNA-SWCNT dispersion in a microcentrifuge tube for 1 min followed by centrifugation at 16100g for 2 min. Figure 3.1. shows qualitatively the use of PVP to adjust the phase in which the DNA/SWCNT hybrids reside. With increasing PVP concentration the hybrids move from the bottom to the top phase.

Figure 3.1 With the addition of PVP, the DNA-CNT hybrid moves from being mostly in the bottom phase to mostly in the top phase

Absorbance measurements were performed on a Varian Cary 50 spectrophotometer over the wavelength range of 200−1100 nm using a 10 mm path length quartz microcuvette. Fractions of the top and bottom phases were collected using a pipette and diluted for

absorbance measurements. Blank top or bottom phases of aqueous two phase systems without SWCNTs were collected and diluted in the same way as the corresponding SWCNT fractions for baseline measurements.

Figure 3.2 shows typical absorbance spectra of the bottom phase demonstrating the transfer of material from the lower to upper phase. We can also observe that the transfer is not uniform across all species. For example, we observe the "hump" corresponding to (8,3) vanishes, i.e., with increasing PVP concentration it appears to move to the top phase earlier than (6,5).

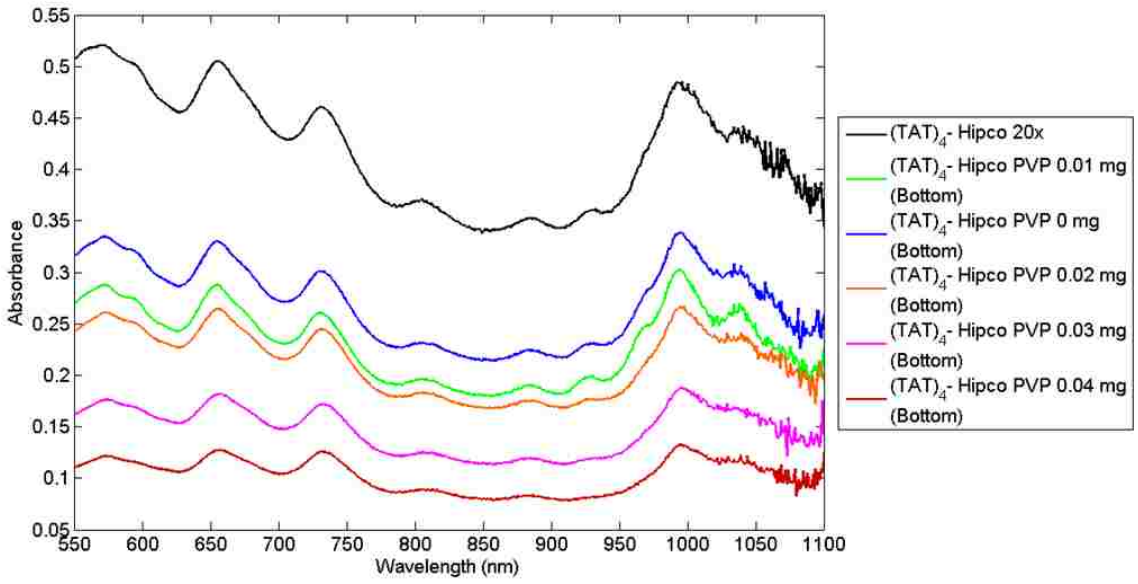


Figure 3.2 Absorbance spectra of the bottom phase showing decrease in DNA-CNT presence in bottom phase with addition of PVP modulant

For quantitative analysis we decompose the absorbance spectra using methods developed earlier for analysis of surfactant exchange experiments.[25] Figure 3.3 shows a typical decomposition of the absorbance spectra into contributions from (9,1), (8,3), and (6,5) SWCNTs. By using this decomposition, we can separately track the change in absorbance intensity of each SWCNT chirality.

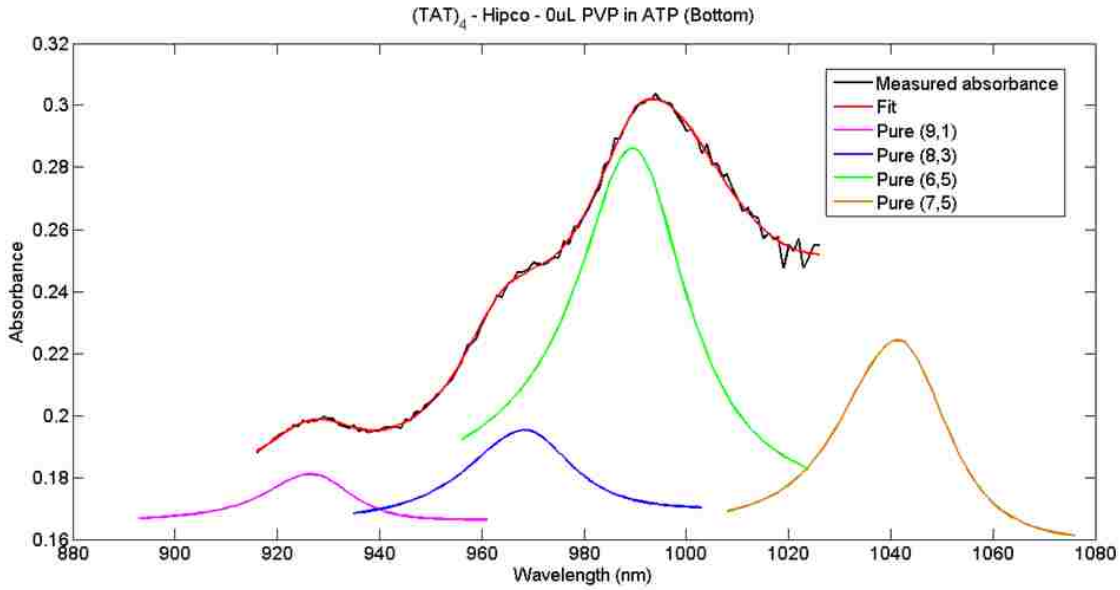


Figure 3.3 Fitted spectra showing contribution of different SWCNTs to the measured spectra

The primary measurement is the ratio of concentrations in the two phases, the partition coefficient

$$K = \frac{C_b}{C_t} \qquad (3.1)$$

Note that we are actually measuring only concentration in the bottom layer and inferring the one in the top layer by mass balance. This is to avoid error due to loss of some

hybrids in the interface. Assuming dilute conditions, the partition coefficient is related to the difference in free energy for insertion of a single hybrid in either phase,

$$K = \frac{c_b}{c_t} = \exp(-\frac{\mu_b^o - \mu_t^o}{kT}) \qquad (3.2)$$

Based on previous experience in aqueous two phase, we make the assumption for the sake of analysis that modulants (PEG, Dextran, and PVP) do not interact directly with the solute (DNA/SWCNT hybrid), but only indirectly affect the hydration free energy. Any given DNA-SWCNT hybrid can be identified by the SWCNT chiral indices (n,m) and the DNA sequence 'd'. 'Recognition' DNA sequence and SWCNT pairs were obtained from data reported by Tu et al. [36]

3.3. **Results and discussion**

3.3.1. Recognition sequences

We first studied the (6,5) recognition sequence: $(TAT)_4$, the (9,1) recognition sequence: $(CCA)_{10}$ and a control sequence: $(GT)_{15}$. Initially most of the DNA-CNTs are present in the bottom phase only. We can move the DNA-CNTs gradually to the top phase by adding increasing amounts of PVP to the two phase system. Figure 3.4. shows the distribution coefficient $c_b/(c_t + c_b)$ for these three sequences in hybrid with (6,5). We found that there is considerable difference in the amount of PVP required to move all the DNA-CNTs from the bottom phase to top phase for the different sequences.

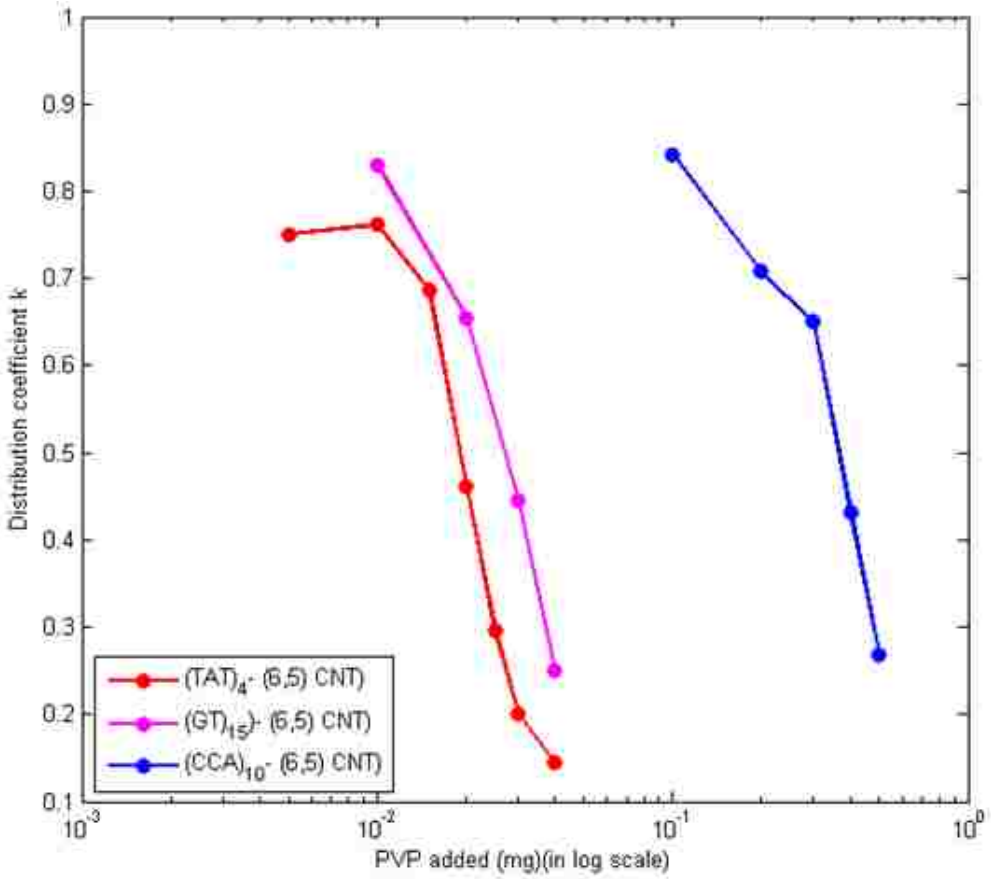


Figure 3.4 Distribution coefficient $c_b/(c_t + c_b)$ for three sequences.

We see that $(GT)_{15}$ and $(CCA)_{10}$ behave very differently even though the lengths are same (30mers). This clearly confirms that the solvation free energy depends strongly on the DNA sequence. From previous experiments described in chapter 2, we know that the rate at which a surfactant removes $(CCA)_{10}$ is about ten times slower than removal of $(TAT)_4$.[25] This appears to correlate with the fact that it take considerably larger amounts of PVP to move the $(CCA)_{10}$/(6,5) hybrid to the top phase, i.e., it has a more negative solvation free energy than $(TAT)_4$/(6,5). We also know from previous experiments that $(GT)_{15}$/(6,5) is removed about 7 times slower than $(TAT)_4$/(6,5).[26] But here, $(GT)_{15}$/(6,5)

is only slightly more hydrophilic than the $(TAT)_4$/(6,5). This suggests that even though the order of hydrophilicity is same as order of binding activation energy, the two properties may not be directly correlated necessarily.

3.3.2. Two repeat 30mers

To study more systematically the effect of base composition in the DNA sequence on the hydrophobicity as measured by aqueous two phase, we looked at all the sixteen combinations of the four DNA bases taken two at a time and repeated fifteen times (i.e. 30mer length).

$$\begin{bmatrix} AA & AT & AG & AC \\ TA & TT & TG & TC \\ GA & GT & GG & GC \\ CA & CT & CG & CC \end{bmatrix} \tag{3.3}$$

Figure 3.5 shows the data on nine of these sequences. One can see that the order of the bases in the repeating unit is not too important, e.g., $(TA)_{15}$ is approximately equivalent to $(AT)_{15}$. However, the composition is very important; see the stark difference between $(GT)_{15}$ and $(AC)_{15}$.

From the data shown in figure 3.5, we can say qualitatively that on (6,5) SWCNT, the 30 mers with repeating units are ranked (GT) < (AT) < (GC) < (CT) < (AG) < (AC) in terms of hydration energy (relative hydrophilicity). However, in order to rank them quantitatively, a theoretical model needs to be proposed by which the amount of modulant required to move the DNA-SWCNT hybrid from one phase to another can be converted into absolute or relative free energy of hydration.

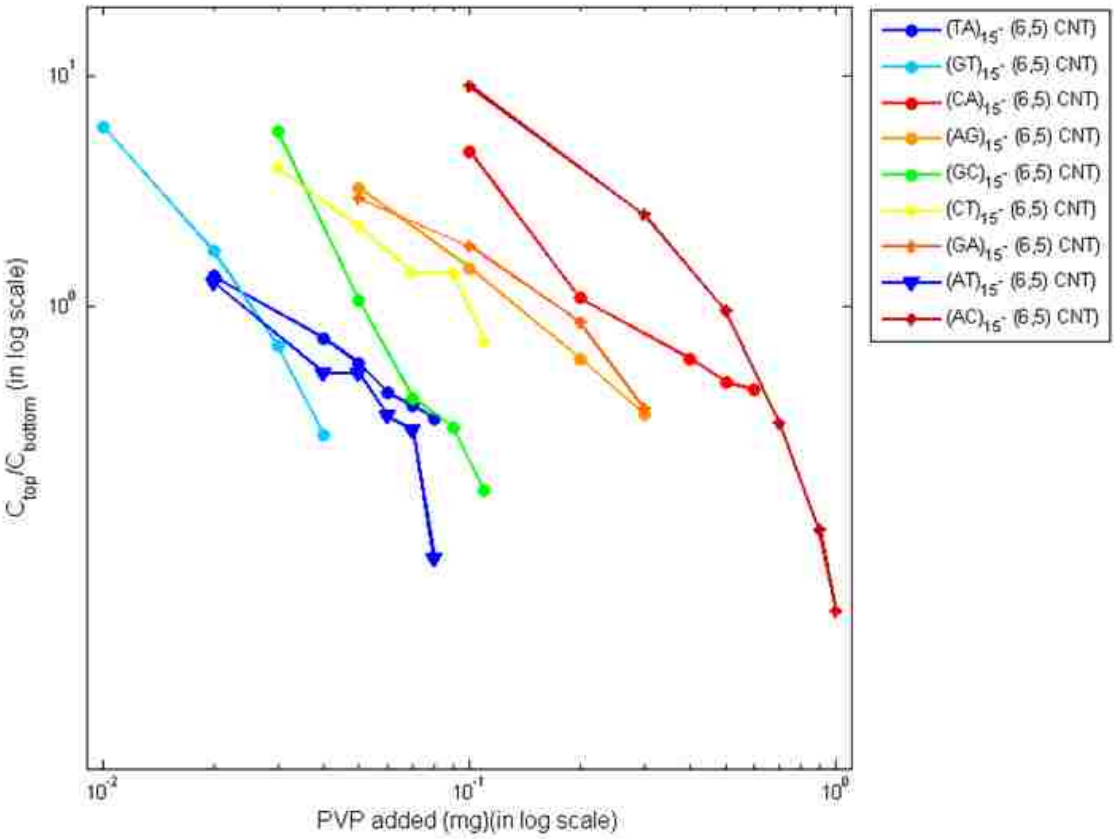


Figure 3.5 Partition coefficient C_{top} / C_{bottom} (as a function of PVP concentration for a number of sequences paired with the (6,5) SWCNT.

3.4. **Conclusion and future work**

We interpreted the increasing amount of modulant required to push a DNA-SWCNT hybrid from the relatively more hydrophilic phase to the more hydrophobic phase in the aqueous two phase system as a measure of relatively higher hydrophilicity of the surface of the hybrid.

We found that $(CCA)_{10}$ on (6,5) SWCNT requires much higher amount of modulant to be moved from the relatively hydrophilic phase to the more hydrophilic phase as compared

to $(GT)_{15}$ on (6,5) even though they are both 30-mers. This suggests that the solvation energy depends greatly on the DNA sequence composition. Studying the repeating units of two of the four DNA bases taken at a time for the same length i.e. 30mer, we find that the solvation energy depends greatly on the bases constituting the repeating unit. The order of the base in the repeating unit is not so important.

These preliminary findings are very promising and suggest that the aqueous two phase system can be used as an effective method to evaluate and possibly rank the solvation free energy of various DNA sequence - SWCNT chirality combinations. Various sequences in the four dimensional sequence space from say $(AT)_{15}$ to $(AC)_{15}$ can be studied systematically. Additionally, the four 30mer homopolymers can also be compared to the sequences studied in this work. Additionally, similar studies can be conducted on various SWCNT chiralities and enantiomers also. In order to rank the hydration energy of these hybrids quantitatively, a theoretical model needs to be proposed by which the amount of modulant required to move the DNA-SWCNT hybrid from one phase to another can be converted into absolute or relative free energy of hydration.

3.5. **Acknowledgement**

This work was done in collaboration with Dr. Ming Zheng (Materials Science and Engineering Division, National Institute of Standards and Technology). I would like to thank Dr. Ao Geyou (Materials Science and Engineering Division, National Institute of Standards and Technology) and Dr. Constantine Khripin for the aqueous two phase separation techniques.

Chapter 4 Stabilization of DNA base dimers near graphite surfaces by hydrogen bonding interactions including non-Watson–Crick pairing

Single and double stranded DNA are increasingly being paired with surfaces and nanoparticles for numerous applications such as sensing, imaging, and drug delivery. Unlike the majority of DNA structures in bulk that are stabilized by canonical Watson-Crick pairing between Ade-Thy and Gua-Cyt, those adsorbed on surfaces are often stabilized by non-canonical base pairing, quartet formation, and base-surface stacking. Not much is known about these kinds of interactions. To build an understanding of the role of non-Watson-Crick pairing on DNA behavior near surfaces, one requires basic information on DNA base pair stacking and hydrogen bonding interactions. All-atom molecular simulations of DNA bases in two cases - in bulk water and strongly adsorbed on a graphite surface, are conducted to study the relative strengths of stacking and hydrogen bond interactions for each of the ten possible combinations of base pairs. The key information obtained from these simulations is the free energy as a function of distance between two bases in a pair. We find that stacking interactions exert the dominant influence on the stability of DNA base pairs in bulk water as expected. The strength of stability for these stacking interactions is found to decrease in the order, Gua-Gua > Ade-Gua > Ade-Ade > Gua-Thy > Gua-Cyt > Ade-Thy > Ade-Cyt > Thy-Thy > Cyt-Thy > Cyt-Cyt. On the other hand, mutual interactions of surface adsorbed base

pairs are stabilized mostly by hydrogen bonding interactions in the order, Gua-Cyt > Ade-Gua > Ade-Thy > Ade-Ade > Cyt-Thy > Gua-Gua > Cyt-Cyt > Ade-Cyt > Thy-Thy > Gua-Thy. Interestingly, several non-Watson-Crick base pairings, that are commonly ignored, have similar stabilization free energies due to inter-base hydrogen bonding as Watson-Crick pairs. This clearly highlights the importance of non-Watson-Crick base pairing in the development of secondary structures of oligonucleotides near surfaces.

The work described in this chapter was published in Shankar, A., Jagota, A., & Mittal, J. (2012). "DNA base dimers are stabilized by hydrogen-bonding interactions including non-Watson-Crick pairing near graphite surfaces." *J. Phys. Chem. B,* 116(40), 12088–94.

4.1. **Introduction**

Apart from being one of the important macromolecules essential for most forms of life, DNA is now the focus of a wide range of technological applications such as DNA-based nanomaterials,[65–67] recognition and separation of carbon nanotubes of various chiralities,[36] biosensors and microarrays,[68] gene delivery,[69] and so on. Interactions between the DNA bases, such as stacking and hydrogen bonding determine the structure of the DNA strands, which in turn determine their function. Various structures of DNA such as the B,[70] A,[71] and Z[72] forms of the double stranded DNA and aptamers[73,74] have been studied in detail. In nearly all cases, double-stranded DNA structure is governed by Watson-Crick (WC) base pairing rules. But when these DNA strands are strongly adsorbed on a surface, some of these conventional DNA structures may not be possible any more. Different structures may be formed in these cases which may depend on non-Watson Crick pairing[24] and surface adsorption effects.[49]

Hydrogen bonding and stacking are the noncovalent interactions which stabilize the DNA double helix.[20,75] Various experimental and computational studies have been carried out in the past on the relative strengths of stacking and hydrogen bonding interactions between DNA bases. Sínanoglu and Abdulnur reported some of the first theoretical work on the effect of solvent on base interactions and suggested that stacking interactions were favored by DNA bases in water due to the large surface enthalpy of water.[76] NMR studies by Mitchell and Sigel on the stacking abilities of various purine and pyrimidine derivatives

such as adenosine, ATP, $Mg(ATP)^{2-}$, uridine and UTP suggested that purines have higher stacking abilities than pyrimidines.[77] Martel's comparison of 6-methylpurine solutions with that of its crystalline precipitates using neutron diffraction suggested base stacking interactions at a separation distance of ~ 0.34 nm between the adjacent bases.[78] Hunter reported a computational model studying the DNA double helix in water which can predict sequence dependent structure and properties of the double helical DNA in water.[79] Danilov and Tolokh conducted Monte Carlo simulations of uracil and thymine dimers in water, finding that stacking interactions were favored over hydrogen bonding interactions in water due to favorable water structure around the stacked configuration.[80] Pohorille and Pratt carried out simulations of self-associated adenine, uracil, guanine, cytosine and adenine-uracil, guanine-cytosine in carbon tetrachloride solvent and self-associated adenine and uracil in water.[81] They reported that stacking interactions were favored in water and hydrogen bonded complexes in nonpolar solvents like carbon tetrachloride. Cieplak and Kollman conducted computational studies on adenine-thymine and guanine-cytosine pairs and found that while stacking was preferred in water, hydrogen bonded complexes were preferred in the gas phase.[82] They also observed that hydrogen bonding interactions were mainly stabilized by electrostatic contribution to the free energies while stacking interactions were mainly stabilized by van der Waals interactions. However they reported that the stacking ability of guanine-cytosine pair was higher than that of adenine-thymine and adenine-adenine which is not consistent with the experimental observations.[83] Dang and Kollman reported potential of mean force (PMF) curves for adenine-thymine pair in

water and gas phase and found free energy of stacking interactions for the pair which is reasonably close to the experimentally obtained value.[84]

Guckian *et al.*[75] have carried out thermodynamic measurements to study the relative stacking ability of the bases. The experiments were based on the 'dangling end' effect which occurs when a single unpaired base is added to the end of a duplex, stabilizing the helix by stacking on it. They have also correlated the stacking ability of the four bases to their physical properties such as polarizibility, dipole moment, hydrophobicity, and surface area. They reported the relative stacking ability to be adenine > guanine ≥ thymine ≈ cytosine. Ke *et al.* have reported measurement of stacking interaction of single stranded poly-adenine and poly-thymine using atomic force spectroscopy.[85] They have studied the effect of base stacking interactions on the molecular elasticity of the two ssDNA sequences. Manohar et al.[29] and Iliafar *et al.*[27,28] conducted measurements of the binding strength of DNA homopolymers by atomic force microscopy-based single molecule force spectroscopy. Binding strengths of various DNA sequences on single walled carbon nanotube of (6,5) chirality were measured by Roxbury *et al.*[26] by studying the kinetics of DNA replacement by surfactant molecules.

Very recently Spiwok et al. showed through their simulations that pairing formed by methyl-adenine and methyl-thymine favored hydrogen bonding interactions when adsorbed onto a graphene surface.[86] Another very recent study by Linak *et al.* demonstrated the increase in accuracy of coarse grained models for DNA when Hoogsteen pairings are also taken into account along with Watson Crick pairings.[87] The efficacy of this model was demonstrated by studying various DNA related phenomena such as melting of a DNA

hairpin, folding of thrombin aptamer containing guanine quartets and strand invasion during triplex formation. However, despite the importance of base stacking and base-pairing in bulk and for surface adsorbed DNA, the associated free energies for the various possible combinations are not easily available from experiments or all-atom simulations.

Some studies on adsorption of single stranded DNA on single walled carbon nanotubes have already been conducted by our group.[22–24] Now, we take a step back and study the behavior of free deoxynucleosides in the presence of a surface. In this paper, we have performed comprehensive all-atom simulations utilizing an explicit solvent nucleic acid model (CHARMM27) along with enhanced sampling techniques (umbrella sampling,[88] Hamiltonian exchange[89]) to obtain base pair free energy as a function of distance between the bases in bulk water and near a graphite surface. We find that all base pairs in bulk water are stabilized due to stacking interactions and the free energy of stabilization is of the order of 3 to 5 kcal/mol. By contrast, base pairs adsorbed on a graphite surface are stabilized by hydrogen bonding between bases (and sometimes base and sugar) and the free energy of stabilization ranges from 1 to 3.5 kcal/mol.

4.2. **Models and Simulation Methods**

As shown in the first column of Table 4. 1, there are ten possible dimer combinations for the four DNA bases. We obtain the all atom co-ordinates of the ten possible deoxynucleoside (only the nitrogenous base attached to a deoxyribose sugar) pairs of DNA. We will hereafter refer to the four deoxynucleosides i.e. Deoxyadenosine,

Deoxyguanosine, Deoxythymidine, and Deoxycytidine as Ade, Gua, Thy, and Cyt respectively. We use the GROMACS 4.5.3 molecular dynamics (MD) package[90,91] to simulate these base pairs in a 4.0 × 4.0 × 4.0 nm^3 octahedral explicit water box with about 1600 TIP3P[92] water molecules, with a total system size of about 5000 atoms in the bulk water case. For the case where the bases are adsorbed on a graphite surface, we simulate the base pairs in a 3.68 × 3.83 × 6.0 nm^3 rectangular explicit water box with 2200 TIP3P water molecules, with a total system size of about 8500 atoms (see Figure 4.1).

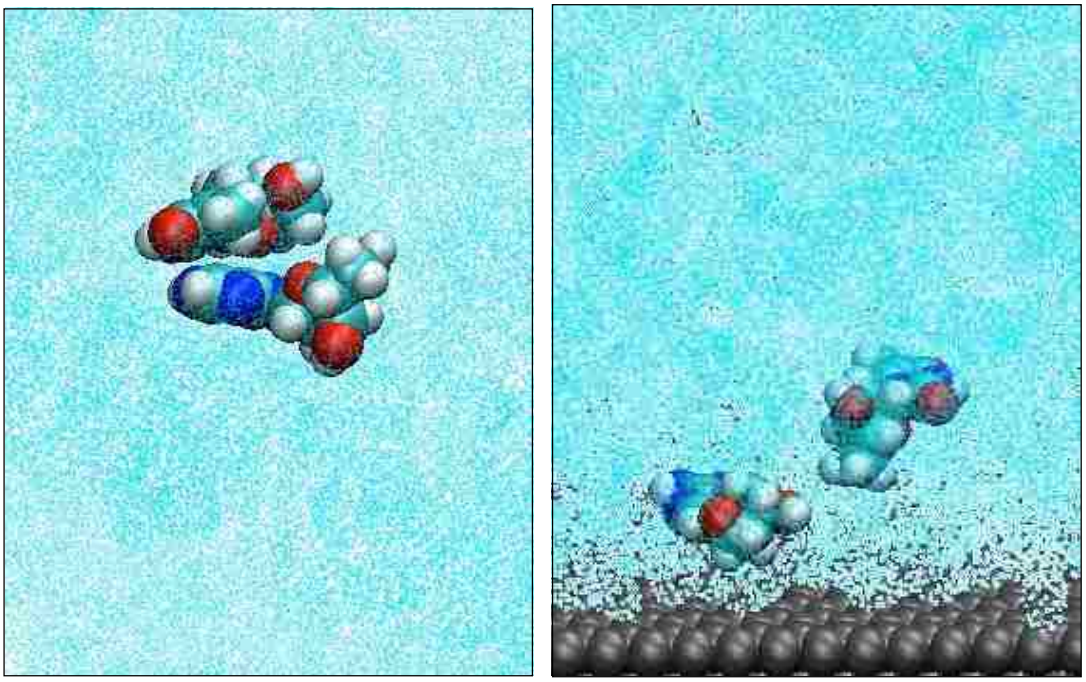


Figure 4.1 Initial configurations for molecular simulations of DNA bases in (left) bulk explicit water box and (right) adsorbed on graphite surface in an explicit water box.

All the simulations are conducted using the all atom CHARMM27 force field [93,94] for nucleic acids for a time period of 27 nanoseconds, using a time step of 2 femtoseconds. Electrostatics interactions are treated with the particle mesh Ewald (PME) method[95] with a cutoff of 0.9 nm. Van der Waals interactions are calculated with a cutoff of 1.2 nm. The simulations are carried out at temperature 300 K and pressure 1 atm. The simulation snapshots are saved every 2 picoseconds in a trajectory. We use Parrinello-Rahman isotropic pressure coupling barostat for the bulk water case and Parrinello-Rahman semi-isotropic pressure coupling barostat[96] in case of surface adsorbed bases with a time constant of 0.5 picoseconds. The graphite atoms are constrained to a fixed position throughout the simulations.

In order to sample the regions which would otherwise be inaccessible due to high free energy barrier along the center of mass distance between the two base nitrogenous rings, we use umbrella sampling method.[88] In this method, one applies a biasing harmonic potential so that regions with low probabilities are also sampled. In case of bases in bulk water, we applied an umbrella bias with a force constant of 1000 kJ/(nm^2-mole) to constrain the center of mass distance between the two base nitrogenous rings at twenty different values ranging from 0.25 nm to 1.5 nm. In case of bases adsorbed on the graphite surface, application of only umbrella bias did not ensure sufficient sampling of all possible pair structures and reproducible results could not be obtained. Specifically, the calculated free energy curve depended on the initial configuration of the bases. This is due to the fact that the hydrogen bonds between bases once formed did not break and re-form to allow

sampling of all possible structures within the given distance constrained by umbrella bias. Hence, umbrella sampling with Hamiltonian exchange[89] is carried out using the Plumed[97] plugin along with the GROMACS molecular dynamics package (see appendix Figure S1 and S2). An umbrella bias with a force constant of 2000 kJ/(nm^2-mol) is applied to restrain the center of mass distance between the two base nitrogenous rings at sixteen different values ranging from 0.35 nm to 1.5 nm. We obtain the potential of mean force (PMF) plots, showing the free energy $G(r)$ at various center of mass distances r between the base nitrogenous rings, for all the ten base pairs using the weighted histogram analysis method (WHAM).[98] The first seven nanoseconds are discarded as equilibration time. Error bars are obtained by averaging over two halves of the data from the remaining 20 nanoseconds. We also obtain the typical base pair structures at the free energy minimum by using a clustering technique[99] based on structural similarity. In this technique, we compare the structures of the base pairs in the trajectories and group them into clusters with a root mean square (RMSD) cutoff distance of 0.15 nm. Thus, we can identify the structures the base pairs are in for most of the simulation time. The clusters are visualized using the visual molecular dynamics (VMD) package.[100]

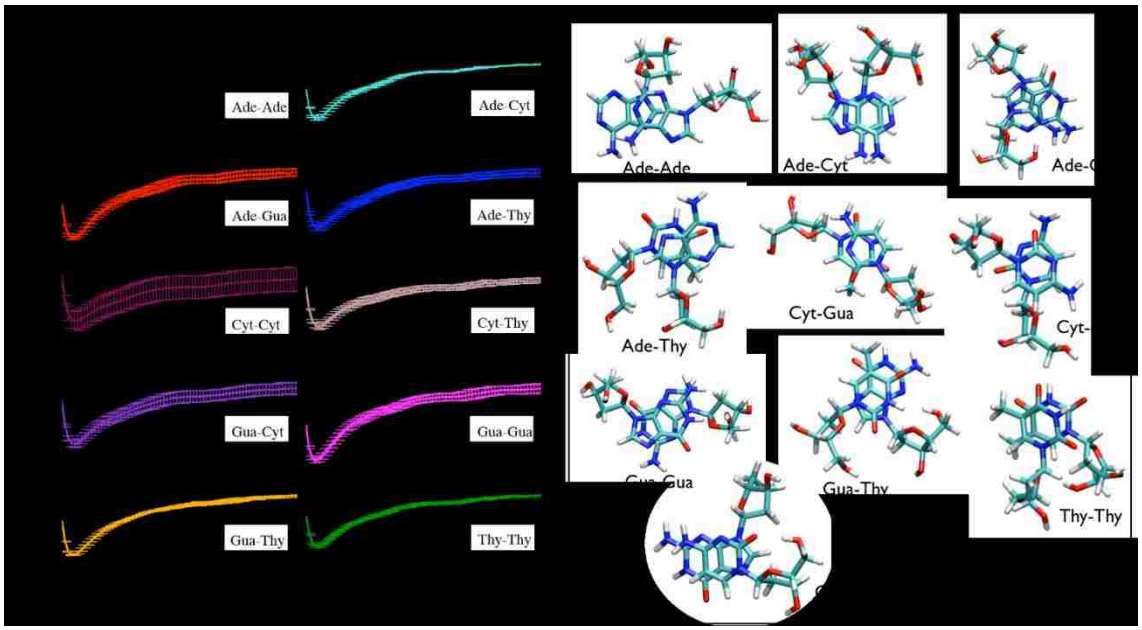


Figure 4.2 PMF and typical structures of base pairs in bulk water

(Left) PMF plots showing free energy G (kcal/mol) at various distances between bases in bulk water. (Right) Typical clusters of the base pairs in bulk water at the free energy minima show only stacking interactions

Table 4.1 Minimum in free energy and the location of minimum for stacking and Hydrogen bonding (HB) interactions

Base pair	G^{stack}_{bulk} (kcal/mol)	r^{stack} (nm)	$G^{HB}_{graphite}$ (kcal/mol)	r^{HB} (nm)
Gua-Gua	-5.25 ± 0.18	0.36 ± 0.01	-1.92 ± 0.00	0.70 ± 0.01
Ade-Gua	-4.80 ± 0.11	0.36 ± 0.01	-3.08 ± 0.00	0.69 ± 0.01
Ade-Ade	-4.60 ± 0.13	0.36 ± 0.01	-2.18 ± 0.00	0.64 ± 0.01
Gua-Thy	-4.26 ± 0.04	0.36 ± 0.01	-0.79 ± 0.10	0.61 ± 0.01
Ade-Thy	-4.05 ± 0.20	0.36 ± 0.01	-2.82 ± 0.00	0.60 ± 0.01
Gua-Cyt	-4.04 ± 0.26	0.38 ± 0.01	-3.49 ± 0.01	0.55 ± 0.01
Ade-Cyt	-4.02 ± 0.09	0.36 ± 0.01	-1.40 ± 0.18	0.61 ± 0.01
Thy-Thy	-3.65 ± 0.07	0.37 ± 0.01	-1.36 ± 0.03	0.62 ± 0.01
Cyt-Thy	-3.37 ± 0.21	0.38 ± 0.01	-2.09 ± 0.15	0.52 ± 0.01
Cyt-Cyt	-2.97 ± 0.70	0.37 ± 0.01	-1.77 ± 0.03	0.53 ± 0.01

4.3. **Results and Discussion**

4.3.1. Potential of Mean Force for DNA base pairs in bulk

First we look at the base pair stability in bulk water in the absence of any surface. The potentials of mean force (PMF) curves show the relative free energies at different center of mass distances between the two base nitrogenous rings. Hence, the minimum free energy, associated with contact between DNA base pairs, is representative of their relative stability.

Figure 4.2 (left panel) shows the presence of a dominant free energy minimum at contact for all the ten base pairs. This minimum is located at a center of mass distance of ≈ 0.37 nm between the two base nitrogenous rings. This suggests that the base pairs in bulk liquid are stabilized due to stacking interactions as the equilibrium spacing for base stacking is roughly the same.[78] In order to obtain and study typical base pair configurations, we use a clustering technique based on structural similarities and visualize them using VMD.[100] The analysis confirms that interactions between base pairs in bulk water are mainly stabilized by stacking interactions between them, as shown in Figure 4.2 (right panel).

In Table 4.1, the second and third columns indicate the stacking free energy (minimum in the PMF curve) for a given base pair and the distance between the centers of mass of the bases' nitrogenous rings where this free energy minimum occurs, respectively. These free energy values represent stacking interactions alone as checked from the simulation trajectory and do not have any contribution from hydrogen bonding interactions. It can be seen that the order of stacking free energies for the various base pairs in bulk water found in our simulations tally qualitatively with the known data obtained from previous experimental and theoretical work[101] i.e. Purine - Purine > Purine - Pyrimidine > Pyrimidine - Pyrimidine. But the quantitative values vary from the data obtained by the previous study which used implicit solvent environment (see appendix Figure S3). As summarized by Friedman and Honig, a wide range of free energy values have been obtained for the same base pair by experimental studies.[101] For example, the value for stacking free energy for Ade-Ade has been reported as 1.21 kcal/mol by Mitchell and Sigel[77] and as 5.73

kcal/mol by Morcillo *et al.*[102] The value we obtained from our simulations for Ade-Ade stacking free energy i.e. 4.6 kcal/mol falls within this range. The nitrogenous aromatic rings of the bases have two distinct faces. For the sake of clarity we follow a nomenclature in which the two faces of the planar aromatic rings of each nitrogenous base are assigned the name 'a' or 'b' as explained in the Appendix (Figure S4). Since we apply umbrella bias to constrain the center of mass distance between the two base nitrogenous rings only, the two bases are theoretically allowed to sample both orientations with respect to each other. But we find that once the bases enter a stable configuration, they remain in the same configuration throughout the simulation as the energy penalty for the bases moving away from each other and flipping into the other configuration is high. We can allow the bases to change configurations during the simulation by using umbrella sampling with Hamiltonian exchange. We find that the bases have almost equal probability of being in each of the two possible configurations (see Appendix Figure S5, suggesting that the bases in bulk water do not have a preference for any one of the two possible configurations). Also, the PMF we obtain by applying umbrella sampling and umbrella sampling with Hamiltonian exchange are nearly identical (see Appendix Figure S6).

4.3.2. Configurations of DNA bases adsorbed on graphite surface

We now look at the base pair interactions between the DNA bases adsorbed on to a graphite surface. The base pairs can be adsorbed on the surface in three possible configurations in case of the four homogeneous base pairs (Ade-Ade, Cyt-Cyt, Gua-Gua, and Thy-Thy) and

four possible configurations in case of the remaining heterogeneous base pairs. This can be visualized by comparing the planar aromatic rings of the nitrogenous bases with coins with two distinct faces. Two different coins with two distinct faces can be flipped and placed on a surface in four ways. Two identical coins can be placed in three different ways. For the sake of clarity we follow a nomenclature in which the two faces of the planar aromatic rings of each nitrogenous base are assigned the name 'a' or 'b' as explained in the Appendix (Figure S4). To account for the effect of the orientation of the aromatic ring with respect to the substrate on which it is adsorbed, we simulated Ade-Ade, Ade-Thy and Gua-Cyt base pairs in all their possible configurations. The resulting potential of mean force plots are shown in Figure 4.3 (left panel). In case of Ade-Ade and Ade-Thy, all PMF curves show similar free energy minima at approximately the same distances. Hence, we can make the assumption that the free energies for the different configurations of a particular base pair are not very different. The one exception to this is Gua-Cyt, as they can form three hydrogen bonds in the a-b and b-a configurations, but only two hydrogen bonds in the a-a and b-b configurations (See Figure 4.3). So we study Gua-Cyt pair in the a-b configuration. For all other base pairs, we study the a-a configuration. Figure 4.3 also shows the typical base pair structures using a clustering techniques based on structural similarity, as done in the previous case. As shown in Figure 4.3 (right panel), the Ade-Thy pair is found to exhibit five structures,[14] Watson-Crick pairing, reverse Watson-Crick pairing, Hoogsteen pairing, reverse Hoogsteen pairing and another wobble base pair structure. At least three of these structures are observed for each of these four possible starting configurations for Ade-Thy. In case of Ade-Ade, we observe two hydrogen bonds between the two bases for all the

three starting configurations. In case of Gua-Cyt, we observe three hydrogen bonds between the bases in the a-b and b-a configurations and only two hydrogen bonds in the other two configurations.

4.3.3. Potential of Mean Force for DNA bases adsorbed on graphite surface

Next we obtain the PMF plots for the ten pairs of DNA bases adsorbed on a graphite surface. It can be observed in Figure 4.4. (left panel) that most base pairs have primary as well as secondary free energy wells. This is due to the fact that the base pairs can form hydrogen bonds at lower distances to form various pair structures and intermolecular sugar-base hydrogen bonds can form at higher distances (see Appendix Figure S7). Using a clustering technique based on structural similarity, with RMSD cutoff of 0.15 nm, we obtain the typical base pair structures stabilizing the system at the distance *r* where we observe a primary free energy minimum.

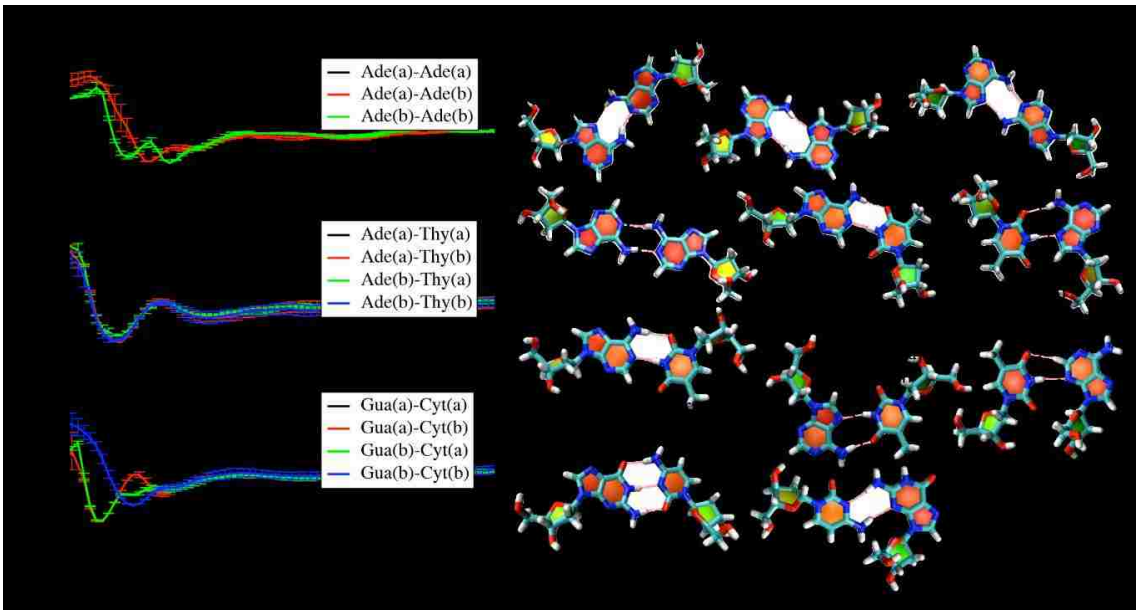


Figure 4.3 PMF and typical structures of base pair configurations on graphite surface

(Left) PMF plot showing free energy minima for (a) the three possible configurations of Ade-Ade, (b) the four possible configurations of Ade-Thy and (c) the four possible configurations of Gua-Cyt. (Right) The typical structures of different configurations of the base pairs Ade-Ade, Ade-Thy and Gua-Cyt adsorbed on graphite surface at free energy minima showing hydrogen bond interactions.

It is observed that these base pairs, when adsorbed on the surface, form hydrogen bond interactions (Figure 4 (right panel)). This is consistent with the recent findings by Spiwok *et al.*,[86] where they reported that hydrogen bonding interactions are more stable arrangements for a base pair when adsorbed onto a surface. It must be noted here that when bases are adsorbed onto the graphite surface, the base - surface π stacking is very important and a dominant force.[29,103] However, base adsorption on a graphite surface is so strong that base pairing strength in this case depend on differences in interactions mainly hydrogen bonding. Base stacking is not favored here and the aromatic rings of adsorbed bases lie flat on the graphite surface. It may be useful to contrast this with

previous work by Wenping[104] where the bases stack with each other such that the plane of the base nitrogenous rings are perpendicular to the surface of a highly charged surface of a carbon nanotube. Similar configurations have been reported in another previous study conducted by Akca *et al.*[105] where they found that short strands of single-stranded DNA assemble into two distinct patterns, small spherical particles and elongated networks on the graphite surface. They presented the argument that the observed assembly behavior was caused by a crossover in the competition between base-base π stacking and base-graphene π stacking. But in the latter study, the structures were formed in the presence of a backbone. In our studies, we study free DNA bases in the absence of a backbone.

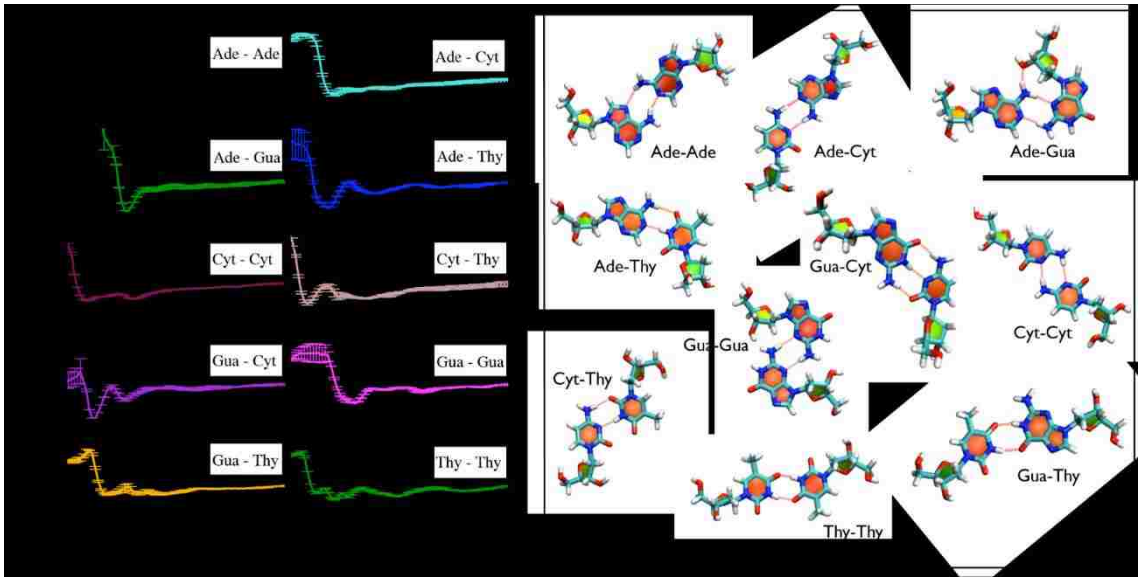


Figure 4.4 PMF and typical structures of base pairs on graphite surface

(Left) PMF plots showing free energy G (kcal/mol) at various distances between bases when adsorbed on graphite surface.
(Right) Typical structures of various base pairs adsorbed on a graphite surface (not shown) at free energy minimum showing hydrogen bond interactions by red dashed lines.

In Table 4.1, the fourth and fifth columns indicate the minimum value for free energy of hydrogen bonding for the corresponding base pairs and the distance between the centers of mass of the bases where this minimum free energy occurs, respectively. These values represent in-plane base-base hydrogen bonding interactions only and do not involve base stacking. The data in Table 4.1 show that the two Watson-Crick base pairs, Ade-Thy and Gua-Cyt, are the most stable. But it is seen that Ade-Ade base pair interaction is also significant when compared to that of the Watson-Crick pairs. Indeed, free energies for several other non-Watson-Crick base pairs are significant in value compared to the Watson-Crick base-pairs. It is observed that the relatively high stability of Ade-Gua is due to the formation of intermolecular base-sugar hydrogen bond. We notice that free energy minimum is at larger distances ≈ 0.7 nm for the larger base pairs, i.e. Ade-Gua, Gua-Gua, and Ade-Ade, and at smaller distances ≈ 0.50 nm for smaller base pairs such as Cyt-Cyt, Cyt-Thy, and Thy-Thy. It is also noted that the hydrogen bonding free energy of Gua-Cyt is not 1.5 times that of Ade-Thy. Hence, for an accurate coarse grained model for DNA near surfaces, the non-Watson-Crick base pair interactions cannot be ignored. For example Linak *et al.*[87] have shown in their recent studies the importance of Hoogsteen pairing in building a more accurate coarse grained model for DNA. Various experimental measurements of Watson-Crick hydrogen bond strength have been conducted and reported in the literature.[106–109] Most recently, Kool and co-workers have estimated these values to be between 0.7 to 1.6 kcal/mol per hydrogen bond .[110] All the base pairs except Gua-Cyt in one particular configuration form one or two hydrogen bonds. For two hydrogen bonds, the range of values will be 1.4 to 3.2 kcal/mol without taking into account the reduction in free

energy due to co-operative effect for multiple hydrogen bonds. The hydrogen bond free energies we obtain range from 1.36 to 3.08 kcal/mol (excluding Gua-Cyt which forms three hydrogen bonds) approximately fall within the expected range as mentioned above. Similarly, the expected range for three hydrogen bonds without taking into account the reduction in free energy due to co-operative effect for multiple hydrogen bonds is 2.1 to 4.8 kcal/mol. The value we obtain for Gua-Cyt is 3.49 kcal/mol, which falls within this range.

One should keep in mind that our simulations include DNA dimer interactions only. Other interactions involving more than two bases, such as guanine quartets,[111] adenine-thymine quartets[47] will be conducted in future studies. We also note that one may expect to find differences in the interactions between base pairs in the absence of a backbone (as studied here) and base pairs as part of a long DNA strand. In the former case, all of the hydrophobic base rings are exposed and can interact with the hydrophobic surface with little hindrance. In the latter case, the hydrophobic base rings in the DNA strand will interact with the hydrophobic surface in the presence of competing backbone-solvent and backbone-surface interactions. We expect that the results presented in this manuscript, quantifying base-base interaction parameters for isolated base pairs, will allow a systematic study of such questions in future.[22,112]

4.4. **Concluding Remarks**

We find that base pairs in bulk water are stabilized by stacking interactions as expected. The order of stability in bulk liquid is found to be (Gua-Gua > Ade-Gua > Ade-Ade > Gua-Thy > Gua-Cyt > Ade-Thy > Ade-Cyt > Thy-Thy > Cyt-Thy > Cyt-Cyt). Hydrogen bonding interactions are not observed in any of these cases, indicating that stacking interactions are much stronger than hydrogen bonding interactions under these conditions. In case of base pairs adsorbed on a graphite surface (strong base-surface stacking due to π–π interactions), because the propensity of bases to stack is satisfied by their adsorption onto the surface, dimer stabilization is dominated by hydrogen bonding interactions. The order of strength for hydrogen bonding interactions is (Gua-Cyt > Ade-Gua[1] > Ade-Thy > Ade-Ade > Cyt-Thy > Gua-Gua > Cyt-Cyt > Ade-Cyt > Thy-Thy > Gua-Thy).

It is found that the base pair Ade-Ade has comparable stability relative to the Watson-Crick pairs Ade-Thy and Gua-Cyt. The other non-Watson-Crick pairs too have free energies which are comparable to that of Watson Crick pairing free energy.

Various all atom force-field models for nucleic acids have been able to successfully predict the structural properties of DNA. [113–117] However, most of the interesting biological and physical phenomena involving DNA have characteristic time and length scales which demand currently infeasible computational requirements for long base sequences.[118] Hence, there is a need for an accurate coarse grained model to enable the study of these DNA

[1] The high stability of this pair is due to the formation of intermolecular base-sugar hydrogen bonding interactions during the course of our simulations

related phenomena. There have been various coarse grained models proposed for single stranded and double stranded DNA sequences.[118–123] But none of the existing coarse grained models (except the one by Linak *et al.*[87] that includes Hoogsteen base pairing) account for the possibility of non-Watson-Crick base pairings which are very important for accurate modeling of single stranded DNA adsorbed onto surfaces. Thus, these models are incapable of capturing several phenomena involving DNA sequences in the presence of surfaces. The essential information needed for building such a model is the relative strengths of stacking and hydrogen bonding interactions for each of the ten possible pairs of DNA bases which we obtain here from all-atom simulations.

4.5. **Acknowledgements**

This work was supported by the National Science Foundation through grant CMMI-1014960. This study utilized the high-performance computational capabilities of Ranger cluster at Texas Advanced Computing Center through TeraGrid resources provided by National Science Foundation. We would like to thank Alan Grossfield for making available the WHAM code on the website. AS acknowledges useful discussions with Dr. Daniel Roxbury and Mr. Apratim Bhattacharya of Lehigh University.

4.6. **Appendix**

S1. Sampling efficiency for DNA bases adsorbed on graphite

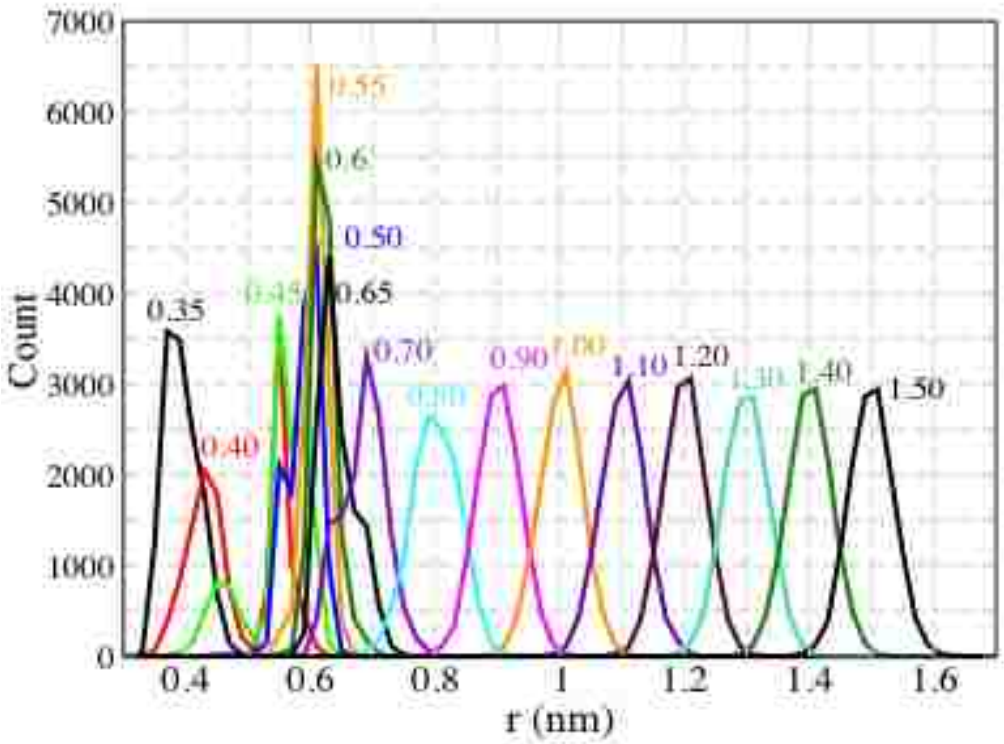


Figure 4.5 Sampling histogram for Gua(a) - Cyt(a) adsorbed on graphite surface using umbrella sampling with Hamiltonian exchange

The simulations are carried out using umbrella sampling with Hamiltonian exchange. An umbrella bias of 2000 kJ/(nm^2/mol) is applied to restrain the distance between the centers of mass of the two DNA bases at 16 different values ranging from 0.35 to 1.5 nm. The histogram depicting the sampling shows that sampling is quite good in the region 0.60 nm to 1.50 nm. There is sufficient overlap between the replicas to enable the extraction of potential of mean force (PMF) curve from this data using the WHAM algorithm.[98]

S2. Drawback of simulations using umbrella sampling for DNA bases adsorbed on graphite surface

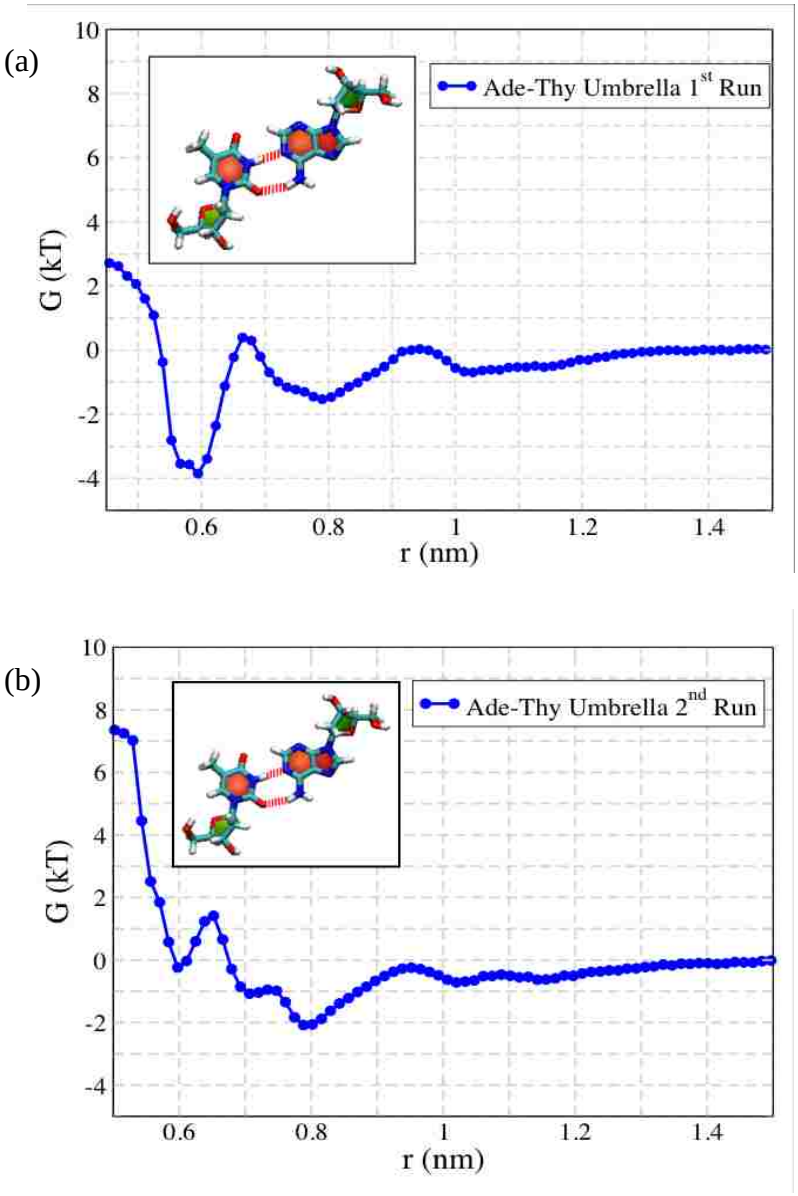


Figure 4.6 PMF and typical configuration of Ade-Thy obtained using umbrella sampling from (a) first run and (b) second run

When we simulate the DNA bases Ade-Thy adsorbed on graphite surface using umbrella sampling two times, we observe that the results vary each time. During the first run (see Figure S2 (a)), the base pairs remain in the reverse Watson-Crick configuration for most of the simulation time. In case of the second run (see Figure S2 (b)), the bases remain in the Hoogsteen configuration for most of the simulation time. This happens because without Hamiltonian exchange, the hydrogen bonds are not broken and re-formed. Hence we are unable to sample all the possible configurations in a reasonable amount of simulation time. In order to overcome this problem, we use umbrella sampling with Hamiltonian exchange, allowing us to sample the various possible configurations of the bases.

S3. Qualitative comparison of stacking free energy values

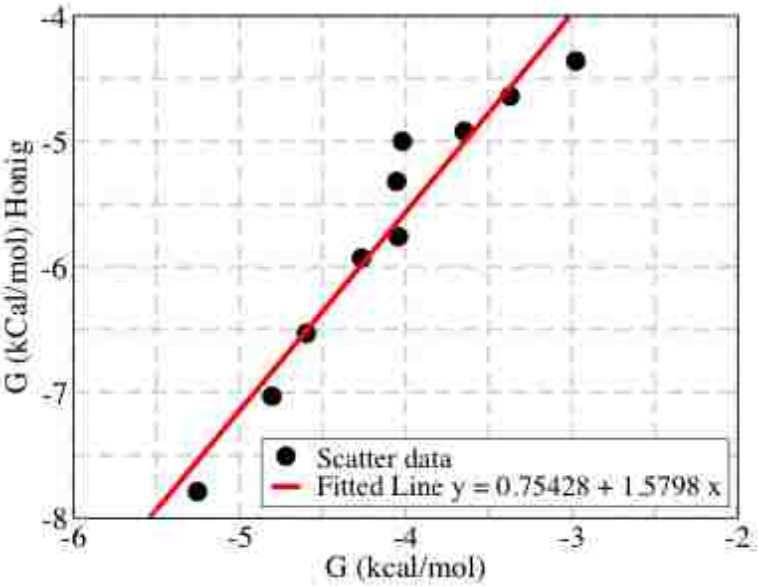


Figure 4.7 Scatter plot comparing stacking free energy values obtained by us and by Friedman & Honig

The scatter plot shows that qualitatively the values for free energy of stacking of the base pairs obtained by us and from previous studies reported Friedman and Honig[101] are similar i.e. the stacking order is same. We obtain new quantitative values using explicit solvent and assign ranking to the base pairs according to their stacking free energy.

S4. Nomenclature for various starting configurations of DNA bases adsorbed on graphite surface

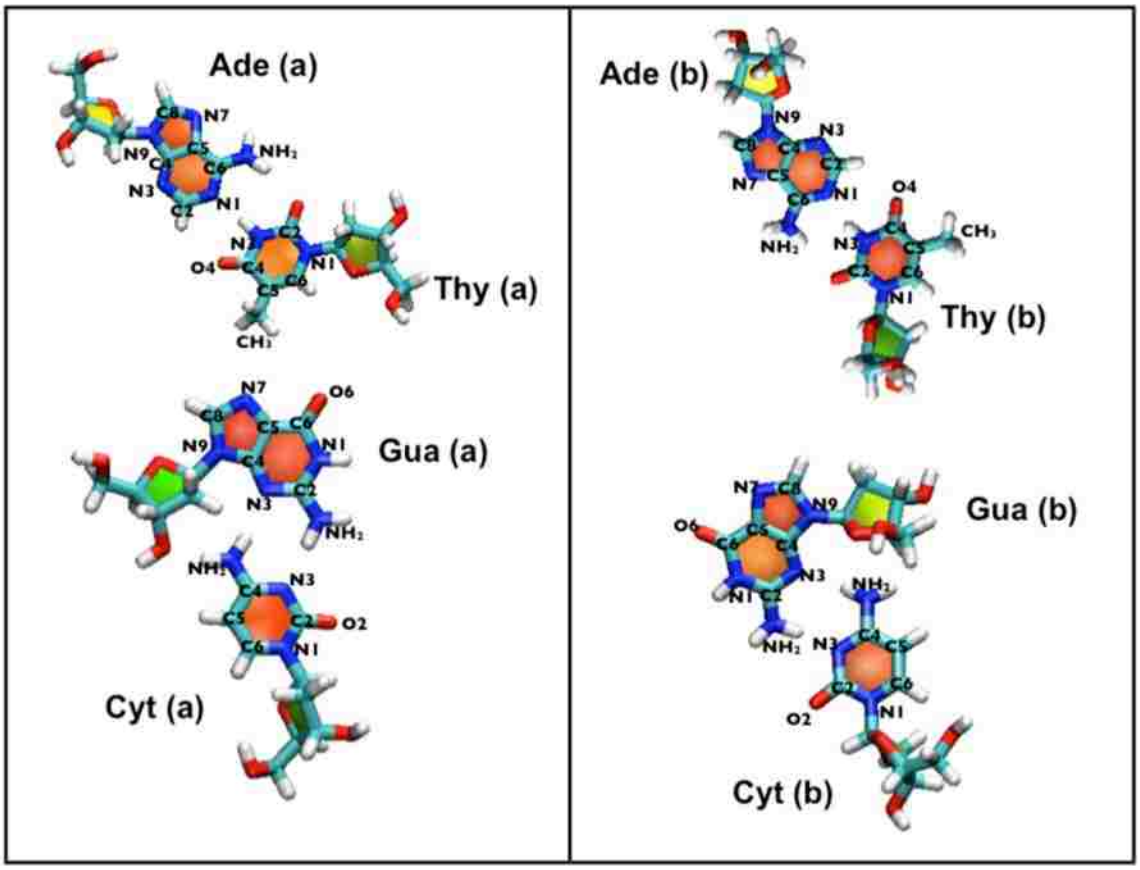


Figure 4.8 Nomenclature for the two possible configurations of each DNA base (Ade, Thy, Gua and Cyt) adsorbed on a graphite surface

Here, we are looking at the plane of the nitrogenous rings of the bases from above and the graphite sheet is below this plane.

For purines (Ade and Gua): When we move from the sugar – N9 bond towards the right (anti-clockwise from N9), if the next atom is C4, then the configuration is designated as '**a**', else as '**b**'.

For pyrimidine (Thy and Cyt): When we move from the sugar – N1 bond towards the right (anti-clockwise from N1), if the next atom is C2, then the configuration is designated as '**a**', else as '**b**'.

S5. Comparison of time spent by DNA bases in bulk water in each of the two configurations a-a and a-b in simulations with umbrella sampling and umbrella sampling with Hamiltonian exchange

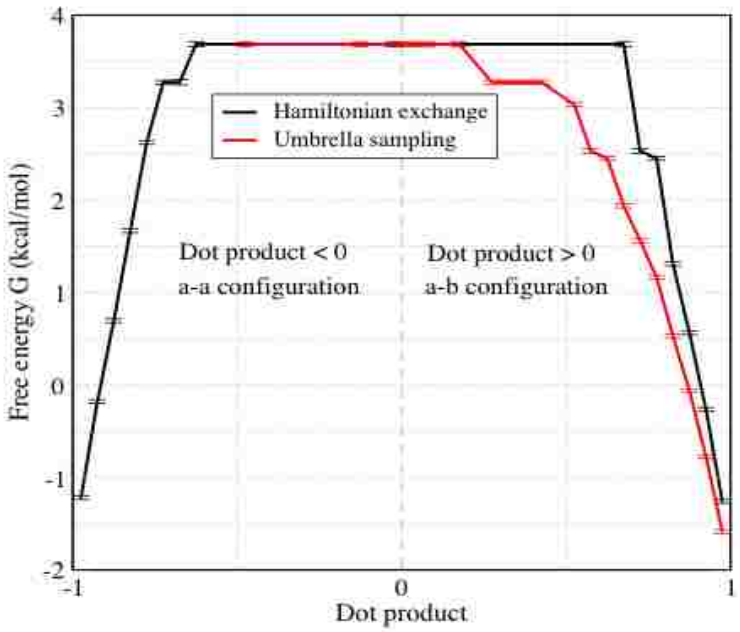


Figure 4.9 PMF with respect to dot products of the unit normals drawn to the planes of the nitrogenous aromatic rings of Ade and Thy in bulk water

In simulations with umbrella sampling, the bases remain in a-b configuration throughout the simulation. This is because the applied umbrella bias results in high energy penalty for the bases to move away from each other and flip into the other configuration.

In simulations with umbrella sampling with Hamiltonian exchange, the bases can sample both configurations. The bases spend almost equal time in both the configurations,

suggesting that the bases do not have a preference for any one of the two possible configurations.

S6. Comparison of the PMF plots obtained from simulations with umbrella sampling and umbrella sampling with Hamiltonian exchange

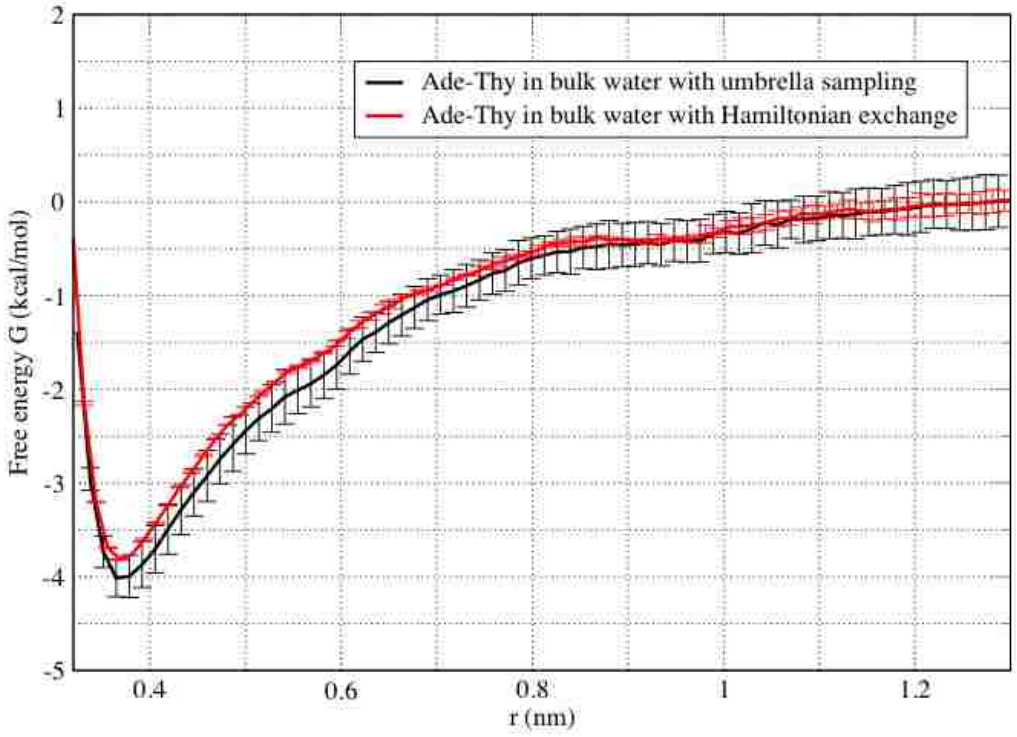


Figure 4.10 PMF plots for Ade-Thy in bulk water obtained from simulations with umbrella sampling and with umbrella sampling with Hamiltonian exchange are nearly identical.

S7. Secondary minima due to hydrogen bonding interactions between DNA base and sugar

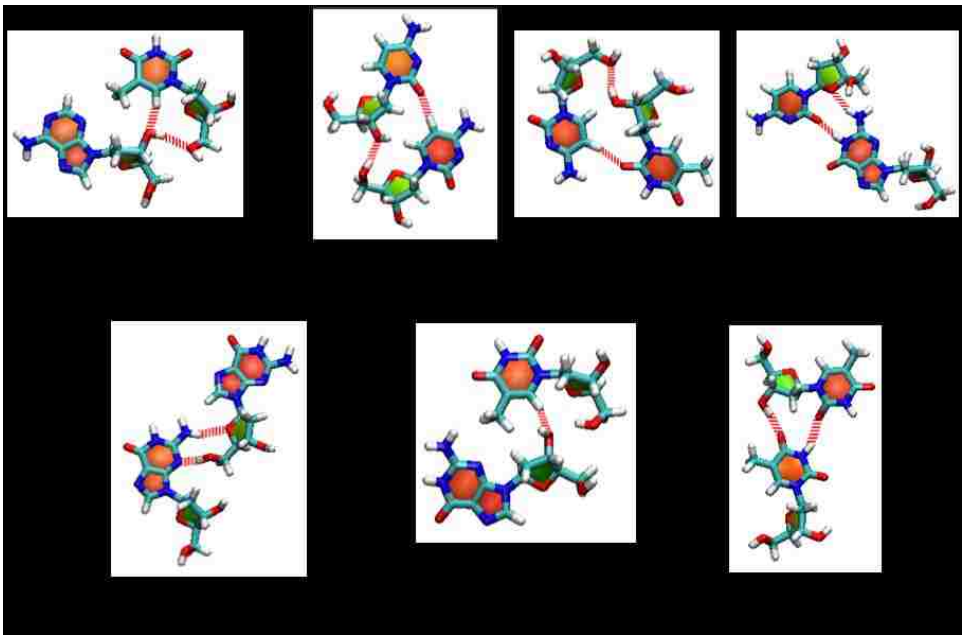


Figure 4.11 Hydrogen bonding interactions involving both base and sugar, causing secondary free energy minima for bases adsorbed on graphite surface at distances higher than where the primary minima occur

Chapter 5 Energetic Basis of Single Wall Carbon Nanotube Enantiomer Recognition by Single Stranded DNA

Hybrids of single stranded DNA and single walled carbon nanotubes have proven very successful in separating various chiralities and, very recently, enantiomers of carbon nanotubes using aqueous two-phase separation. This technique sorts objects based on small differences in hydration energy, which is related to corresponding (small) differences in structure. Separation by handedness requires that a given ssDNA sequence adopt different structures on the two SWCNT enantiomers. Here we study the physical basis of such selectivity using a coarse grained model to compute the energetics of ssDNA wrapped around an SWCNT. Our model suggests that difference by handedness of the SWCNT requires spontaneous twist of the ssDNA backbone. We also show that differences depend sensitively on the choice of DNA sequence.

5.1. Introduction

Single walled carbon nanotubes (SWCNTs) are low dimensional tubular structures of a single layer of sp^2 hybridized carbon atoms bonded in a hexagonal lattice except at their ends.[1] They can be classified as chiral and achiral SWCNTs, which have a direct effect on their properties such as being metallic and semiconducting.[3] Like other chiral molecules, chiral SWCNTs can exist as right handed and left handed enantiomers.

Due to their extraordinary mechanical, electronic and optical properties[6,7], SWCNTs have found a number of applications in the recent times such as thin film transistors[8,9], organic photovoltaics[10], various types of biosensors[13–15] and targeted drug delivery[16,17]. Many potential biomedical and sensing applications require separation of nanotubes according to chirality and handedness such as optical transition based applications[124,125], molecular sensing of amino acids[126] and chirality based effects.[127,128] Because of this, it is important to develop separation techniques for sorting nanotubes based on their chirality and handedness. A number of techniques have been developed, including density gradient centrifugation, DNA-assisted separation[129] and aqueous two phase separation[30,63,64]. However, most of these techniques can only separate the nanotubes according to their chiralities and not handedness. Some previous work has been done in which SWCNT chiral enantiomers have successfully been separated using density gradient ultracentrifugation[4,130] and block co-polymers[131] .

The structure of DNA on SWCNT has been probed in many ways including molecular dynamics (MD) simulations[21–24], measuring activation energy of displacement of DNA by

a surfactant[25,26], AFM studies[27–29] and aqueous two phase separations[30]. MD simulations have given some idea of how DNA adsorbs onto the SWCNT surface. Surfactant exchange studies, aqueous two phase studies and AFM studies have given an idea of the total binding free energy, activation barriers and small differences in solvation energy for different DNA sequences on different SWCNT chiralities.

The DNA assisted methods of separation usually exploit the differences in the DNA structure on various SWCNT types including chiralities and enantiomers. The purpose of this paper is to develop a physical model for the energetics behind DNA based enantiomer separation. From all the previous work, we hypothesize that separation relies on small differences in solvation free energy that are reflected in small differences in structure, for which we have as proxy small differences in binding energy of the hybrid. Accounting for the various contributions to this energy, such as adhesion between bases and the SWCNT, hydrogen bonding between bases, bending and torsion of ssDNA, we ask what it is that allows enantiomeric recognition to happen.

5.2. **Methods**

5.2.1. Representation and Structure of the Hybrid

We studied three different SWCNT types, (10,0), (6,5) and (5,6) The coordinates for the carbon atoms in each SWCNT of length ~ 40 nm were generated using the "Nanotube builder" in the VMD package.[100] We used a coarse grained model in which each 'mer' in ssDNA is represented by two-beads. This is the simplest coarse-grained model that allows us to incorporate base-specificity and surface adsorbed structures with bases that alternate

on each side.[121] One bead represents the backbone (consisting of the sugar and phosphate) and the second bead is for the base. The co-ordinates for each bead were generated using code written in MATLAB®. While building this model, we made some assumptions based on data acquired using all atom molecular dynamics simulations done previously such as: each DNA strand backbone adopts a strictly helical configuration and the DNA bases alternate on each side of the backbone.[21–24,132]

The SWCNT is placed along the Z-axis and the DNA backbone beads are arranged along a helical path which is co-axial with the SWCNT. The bond between the backbone bead and the base bead is perpendicular to the local tangent to the backbone, and alternates on either side of it.[22] Given a starting point of the helical backbone, all bead positions are therefore specified once the helical angle or pitch is given. In case of a single DNA strand on the SWCNT, the position of the first backbone bead is sampled over the nanotube surface in the angular and axial direction. A representative figure of one DNA strand on nanotube model is shown in figure 5.1, with the carbon nanotube atoms in black, DNA backbone beads in blue, and DNA base beads in pink. The pitch angle 'θ' is as defined in the figure 5.1 inset, where 'a' is the helix diameter and $2\pi c$ is the pitch of the helix. When the helix is left handed, 'c' and 'θ' are both negative and for right handed helix, they are positive. When 'θ' is $\pm \pi/2$, the helix simply becomes a circle and when it is zero, the helix is a straight line parallel to the helix axis (in this case the z-axis).

Table 5.1 provides values of the various parameters used to construct the model.

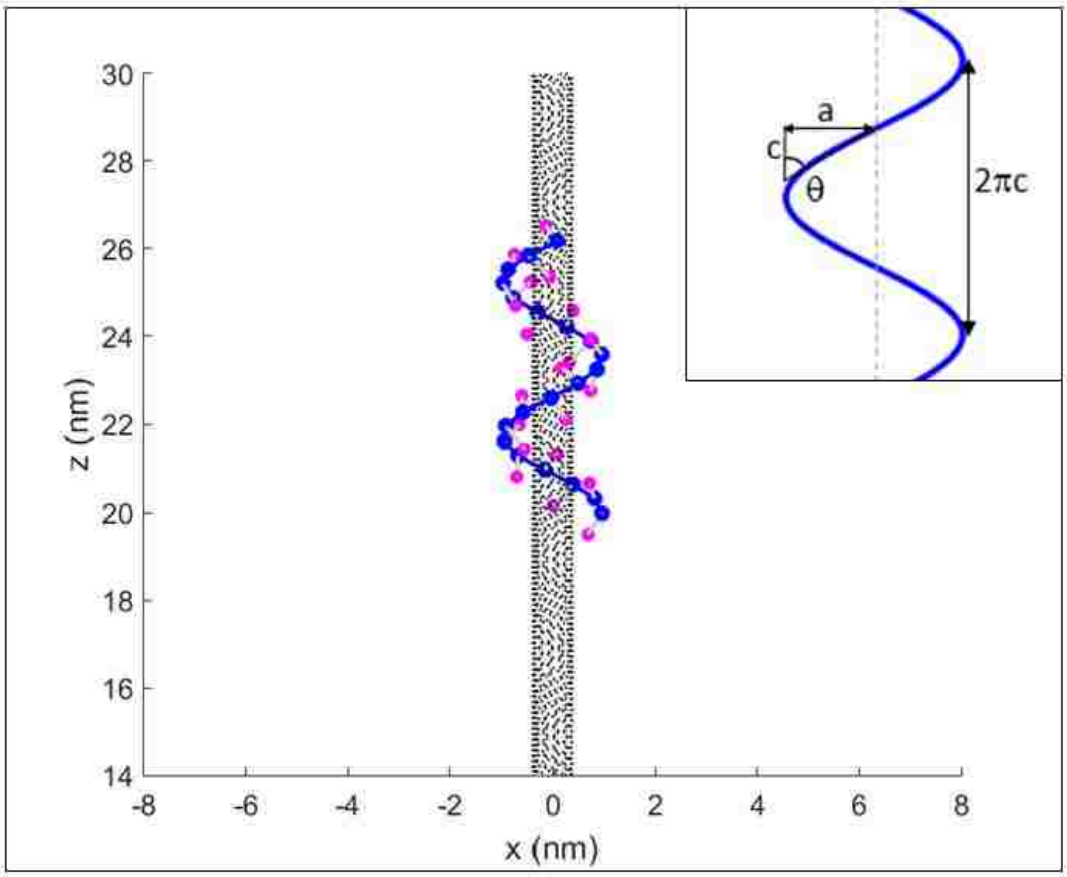


Figure 5.1 DNA-SWCNT coarse grained model

Helical conformation of an $(AT)_{10}$ ssDNA sequence wrapped helically around a (6,5) SWCNT. SWCNT carbon atoms are in black, the DNA backbone beads (phosphate + sugar) are in blue, the blue line shows the helical backbone, and the DNA base beads are in pink.

Table 5.1 Parameters for DNA-CNT model co-ordinates

	Value used	Source
Radius of SWCNT i) (6,5) ii) (5,6) iii) (10,0)	i) 0.373 nm ii) 0.373 nm iii) 0.392 nm	CNT Diameter = $(n^2 + m^2 + nm)^{1/2} * 0.0783$ nm[133]
Distance between SWCNT surface and ssDNA backbone (d_{cp})	0.592 nm	Ref. [22]
Distance between SWCNT surface and base (d_{cb})	0.365 nm	Ref. [22]
Radius of helix i) (6,5) ii) (5,6) iii) (10,0)	i) 0.965 nm ii) 0.965 nm iii) 0.984 nm	CNT radius + distance from SWCNT to backbone

Distance of base from SWCNT axis i) (6,5) ii) (5,6) iii) (10,0)	i) 0.738 nm ii) 0.738 nm iii) 0.757 nm	CNT radius + Distance from SWCNT to base
Distance from backbone to base (Adenine) ($d_{pb\text{-}Ade}$)	0.64 - 0.8 nm	Ref. [118]
Distance from backbone to base (Thymine) ($d_{pb\text{-}Thy}$)	0.49 - 0.6 nm	Ref. [118]
Distance between two consecutive backbone beads (d_{pp})	0.65 nm	Ref. [22]

5.2.2. Energy Potentials

Our goal is to calculate the energy of a DNA-SWCNT hybrid as a function of helical pitch and number of strands. We treat the SWCNT as rigid and so its carbon atoms are fixed and there is no need to specify or compute their internal interactions. The variable contributions that are summed to obtain the total energy are

- Adhesive interaction between DNA bases and SWCNT atoms,
- Hydrogen bonding between DNA bases,
- Bending and torsional energy of the DNA backbone
- Base-backbone repulsion
- Exclusion energy to prevent overlap of bead volume

It is assumed that electrostatic interactions are subsumed into the bending and torsional energies.

5.2.2.1. Adhesion energy between bases and SWCNT

We represent the adhesive interaction between the base beads and the SWCNT by a 12-6 Lennard-Jones potential

$$E_{adhe} = \sum 4 \in_{adhe} \left[\left(\frac{\sigma_{adhe}}{r_{cb}} \right)^{12} - \left(\frac{\sigma_{adhe}}{r_{cb}} \right)^{6} \right] \quad (5.1)$$

Where r_{cb} is the distance between the base bead and a SWCNT carbon atom. The values of σ_{adhe} and ε_{adhe} for each type of base are obtained by calibrating equation (5.1) against

experimental data reported by Iliafar et al. for adhesion energy per base for each DNA base on graphite [27] and SWCNTs[28]. We set σ_{adhe} as 0.34 nm and found the values of ε_{adhe} for which E_{adhe} equaled the experimentally obtained values for the particular base and substrate in the above mentioned works on graphite and SWCNT. These values of ε_{adhe} and σ_{adhe} were then used in equation (5.1) to obtain adhesion energies between base beads and SWCNT carbon atoms. The values we found and used in our model are ε_{adhe} = 0.636 and 0.726 kT for A and T on graphite, respectively to get E_{adhe} ~ 9.9 kT per Ade on graphite and E_{adhe} ~ 11.3 kT per Thy on graphite. Based on measurements of adhesion energy per base on SWCNTs, we obtain ε_{adhe} = 3.346 kT and 2.039 kT for A and T, respectively. Note that there is a significant and surprising increase in binding energy between ssDNA and SWCNTs compared to flat graphite. Unless otherwise stated, we have used the measured value for binding energy for ssDNA strands against surface deposited SWCNTs.

Note that the ssDNA base-SWCNT interaction is sensitive to chirality in that a given DNA conformation will have different energies depending on the chirality of the SWCNT (e.g., (6,5) vs (5,6)) because it 'sees' a different arrangement of atoms in the two cases. However, this interaction has the following symmetry. If we take a certain ssDNA-SWCNT hybrid, say right-handed $(TAT)_4$ on (6,5), then it will have the same energy as its mirror image, i.e., a left-handed $(TAT)_4$ conformation on (5,6).

5.2.2.2. Bending energy of backbone

We assign a resistance to bending that is quadratic in the bending angle, and allow the possibility of spontaneous bending, i.e., that the ssDNA is naturally bent.

$$E_{bending} = \frac{1}{2} k_b (\theta_{bend} - \theta_0)^2 \quad (5.2)$$

The bending stiffness includes intrinsic bending resistance and electrostatic self repulsion. To establish the order of magnitude value of k_b we rely on experimental measurements that show persistence length of ssDNA in solution to be about 8 Angstroms, although it can be quite a bit larger at lower salt concentration.[134] For this reason we can treat it as a lower limit. We first tried to get a reasonable value for k_b (see appendix S1) and estimate $k_b \sim 4kT$.

θ_{bend} is the angle formed by three consecutive backbone beads. So for an ssDNA with 'n' bases, there will be (n-2) bending angles, which will be identical for a given helix pitch.

DNA is known to exist in nature in the double helical structure, usually as B-DNA. In this structure, each of the two single strands has a curvature and a torsion, giving rise to the idea that each strand DNA ought to have a spontaneous bending angle and torsional angle. That is, we assume that the double stranded B-DNA conformation is the unstressed, relaxed state. Assuming that the spontaneous bending angle comes from the double stranded B-DNA found in nature, we find a value for θ_0 to be 30 degrees. (See appendix S2)

Figure 5.2 shows how bending angle θ_{bend} (in blue) changes as a function of the helix pitch angle. Note that it is symmetric with respect to handedness (negative pitch corresponds to left-handed helices; positive to right-handed ones.) Figure 5.3 shows the corresponding bending energy as described by equation (5.2) as a function of the helix pitch. Note that the spontaneous curvature makes the minimum in free energy at some bent state, but this is symmetric with respect to handedness. That is, a right handed and a corresponding left handed helix have the same bending energy.

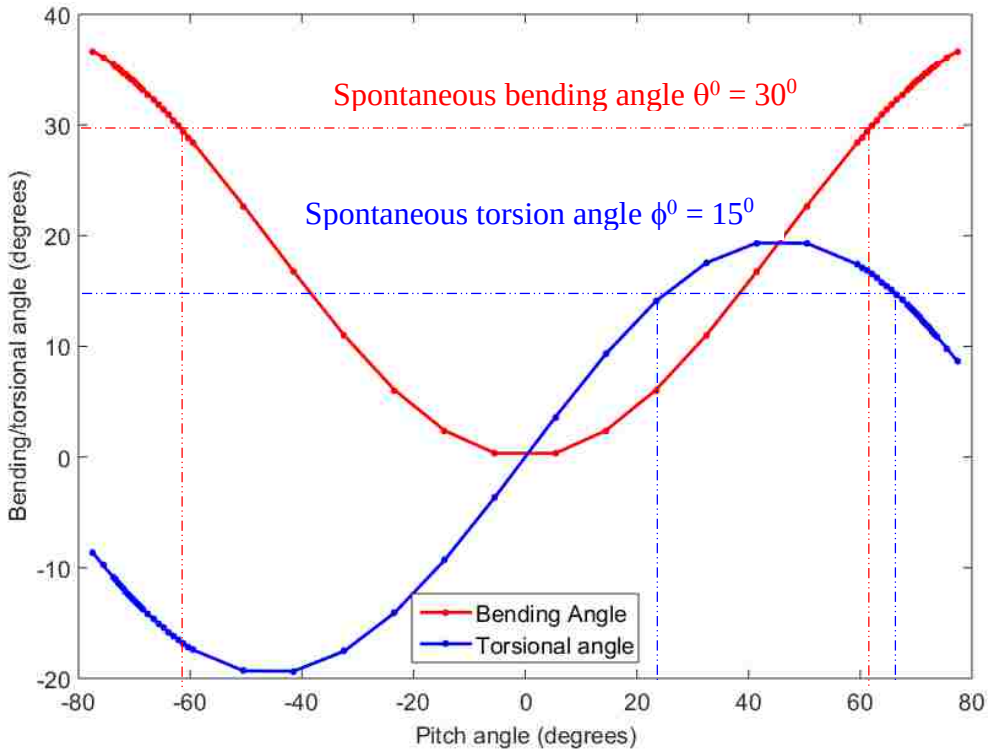


Figure 5.2 Bending angle θ_{bend} and torsional angle ϕ_{ang} as a function of helical pitch angle (all the angles are shown in degrees but used in radians in equations (5.2) and (5.4))

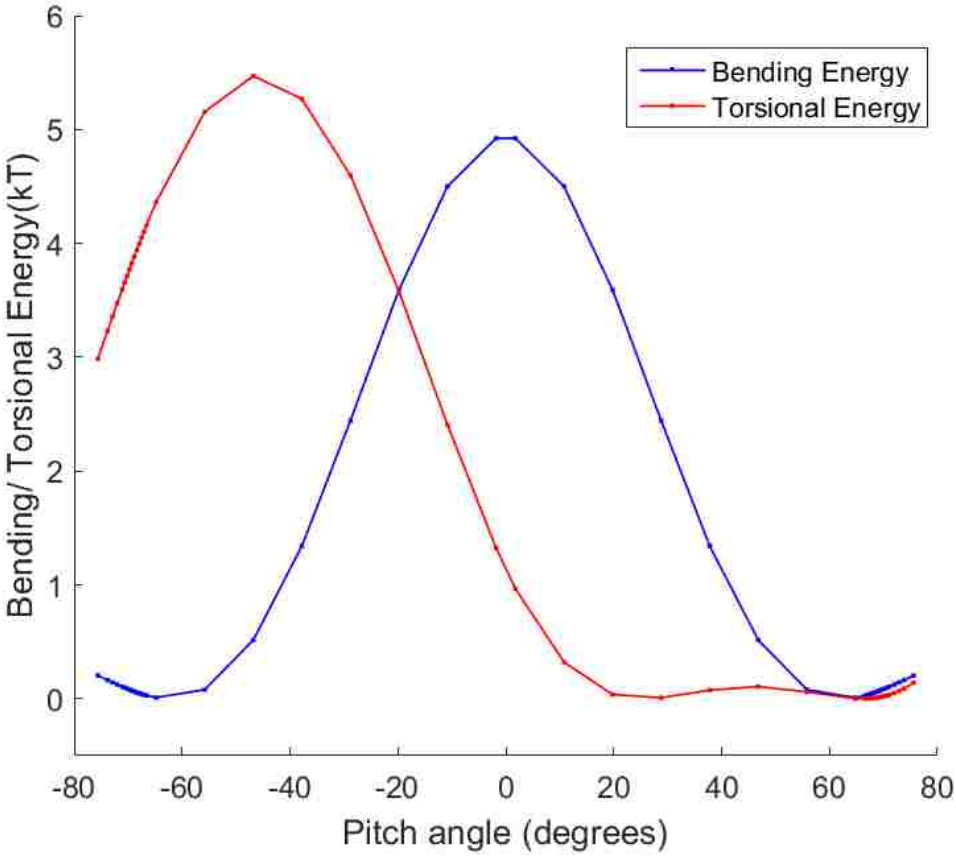


Figure 5.3 Bending and torsional energy as a function of pitch angle

In blue: Bending Energy $E_{bending}$ (kT) versus helix pitch angle (degrees) for $\theta_0 = 30$ degrees,
In red: Torsional Energy $E_{torsion}$ (kT) versus helix pitch angle (degrees) for $\varphi_{ang} = 15$ degrees

5.2.2.3. Torsional energy

Similarly, we assign to the single-stranded DNA backbone a spontaneous torsion corresponding to its native double-stranded B-DNA structure. There have been various studies reporting different single stranded DNA sequences showing signals when subjected to circular dichroism, consistent with spontaneous torsion.[135] The torsional angle ϕ_{ang} is

formed by four consecutive backbone beads. So, for an ssDNA with 'n' bases, there will be (n-3) torsional angles, which will be identical for a given helix pitch. We assign a resistance to torsion that is quadratic in the torsional angle ϕ_{ang} and also sensitive to the handedness as torsional angle varies with handedness, and allow the possibility of spontaneous torsion, i.e., that the ssDNA is naturally twisted.

$$E_{torsion} = \frac{1}{2} k_t \sin^2(\phi_{ang} - \phi_0) \tag{5.3}$$

Even though the exact value of the spontaneous torsional angle ϕ_0 may vary, depending on the exact DNA sequence, we simplified the model by assuming a single value for ϕ_0 as 15 degrees. (See appendix S2)

We need to make a reasonable assumption for the value of k_t in equation (5.3). From literature, we know that k_t is roughly half of k_b for double stranded DNA.[136] Now to get upper and lower estimates of k_t, we can assume single stranded DNA to be various shapes. The bending rigidity k_b is proportional to the area moment of inertia while the torsional rigidity k_t is proportional to the moment of inertia about an axis.[137] If we assume the single stranded DNA to be a circular rod, then k_t will be equal to 0.5 times k_b (see appendix S3). If single stranded DNA is assumed to be a thin rectangular block, k_t is about 60 times the value of k_b (see appendix S3). Thus, we estimate the value of k_t based on value of k_b and the ratio k_t/k_b assumed.

Figure 5.2 shows how torsional angle ϕ_{ang} (in red) changes as a function of the helix pitch. Figure 5.3 shows the torsional energy (in red) as described in equation (5.3) as a function of the helix pitch. Note that this torsional energy with spontaneous torsion introduces a chirality-dependence. Right handed helices have a much lower torsional energy than their corresponding left handed mirror images.

5.2.2.4. Hydrogen bonding between bases

Previous molecular dynamics simulation studies have shown that non-Watson-Crick base pairing can be quite significant in DNA structures near nanotube and graphite surfaces. [21–24,56] Hence, we assume that hydrogen bonding can occur between any two bases (including non-Watson-Crick pairs) if they are separated by at least two bases, i.e., the base n can have hydrogen bonding with all others on the same DNA strand except bases (n ± 1) and (n ± 2) bases.

We represent the hydrogen bonding energy between the base beads by a 12-6 Lennard Jones potential, multiplied by a factor that is a Gaussian with argument $(\alpha_{\mathrm{HB}} - \theta_{HB})^2$. This allows us to capture the strong directionality of the hydrogen bonding interaction which decays rapidly if the angle α_{HB} differs from θ_{HB}.

$$E = \sum 4 \in_{HB} \left[\left(\frac{\sigma_{HB}}{r_{bb}}\right)^{12} - \left(\frac{\sigma_{HB}}{r_{bb}}\right)^{6}\right] * \exp\left(-\frac{(\alpha_{\mathrm{HB}} - \theta_{HB})^2}{2\, s_{HB}^2}\right) \qquad (5.4)$$

Where r_{bb} = distance between the base beads

α_{HB} is the angle formed by backbone bead, base and another base.

θ_{HB} is the cutoff angle for hydrogen bonding which is assumed to be 0 degrees.

s_{HB} is the deviation allowed for the angle which we assume to be 30 degrees.

We fitted the Lennard Jones 12-6 potentials to the Potential of Mean Force (PMF), which is free energy as a function of distance between the two bodies, obtained from Molecular Dynamics simulations for various DNA bases on graphite previously[56]. Thus we can obtain the values for ε_{HB} and σ_{HB}.

ε_{HB} for A-T = 3.70, ε_{HB} for A-A = 2.70, ε_{HB} for T-T = 1.37 (kT)

σ_{HB} for A-T = 0.53 nm, σ_{HB} for A-A = 0.59 nm , σ_{HB} for T-T = 0.77 nm

In case we have Gua and Cyt also in the DNA sequence, we will have to account for the remaining possible 7 hydrogen bonding base pairs also.

5.2.2.5. Repulsion between base and backbone of adjacent strand:

In order to preclude the possibility of a DNA base crossing over an adjacent strand to interact with a base on non-adjacent strand, we introduced an energy penalty: if the base gets closer than a distance equal to the distance between consecutive backbone beads (d_{pp}) from a backbone bead on an adjacent strand, a hard wall energy potential is introduced.

Table 5.2 Parameters for energy calculations (force field)

Parameters	Values	Source
Adhesion energy between bases and graphite	σ_{adhe} = 0.34 nm ε_{adhe} (Ade) = 0.636 kT ε_{adhe} (Thy) = 0.726 Kt	Calibrated using experimental value for adhesion energy[27]
Adhesion energy between bases and SWCNT	σ_{adhe} = 0.34 nm ε_{adhe} (Ade) = 3.346 kT ε_{adhe} (Thy) = 2.039 Kt	Calibrated using experimental value for adhesion energy[28]
Bending energy	k_b = 4kT θ_0 = 30 degrees	Appendix S2 and s3
Torsional energy	k_t = 0.5 to 60 times kb ϕ_0 = 15 degrees	Appendix S2 and S3
Hydrogen bonding energy	ε_{HB} = 2.7 kT, (A-A) σ_{HB} = 0.59 nm (A-A) ε_{HB} = 3.7 kT, (A-T) σ_{HB} = 0.53 nm (A-T) ε_{HB} = 1.37 kT, (T-T) σ_{HB} = 0.77 nm (T-T)	[56]

5.3. **Results and discussion:**

As explained previously, when the helix is left handed, pitch angle 'θ' is negative and for right handed helix, it is positive. When 'θ' is $\pm \pi/2$, the helix simply becomes a circle and when it is zero, the helix is a straight line parallel to the helix axis (in this case the z-axis).

Figure 5.4 shows a typical plot of energy per unit length as a function of helical pitch angle for the DNA wrap, for one strand of $(AT)_{10}$ on (6,5) SWCNT. The plot shows two energy minima for one strand of DNA on the SWCNT surface, where one minimum AR corresponds to the right handed DNA helix (positive pitch angle) and the other AL corresponds to the left handed DNA helix (negative pitch angle). Whichever of the two minima is lower corresponds to the more stable handedness of the DNA helix for that particular case.

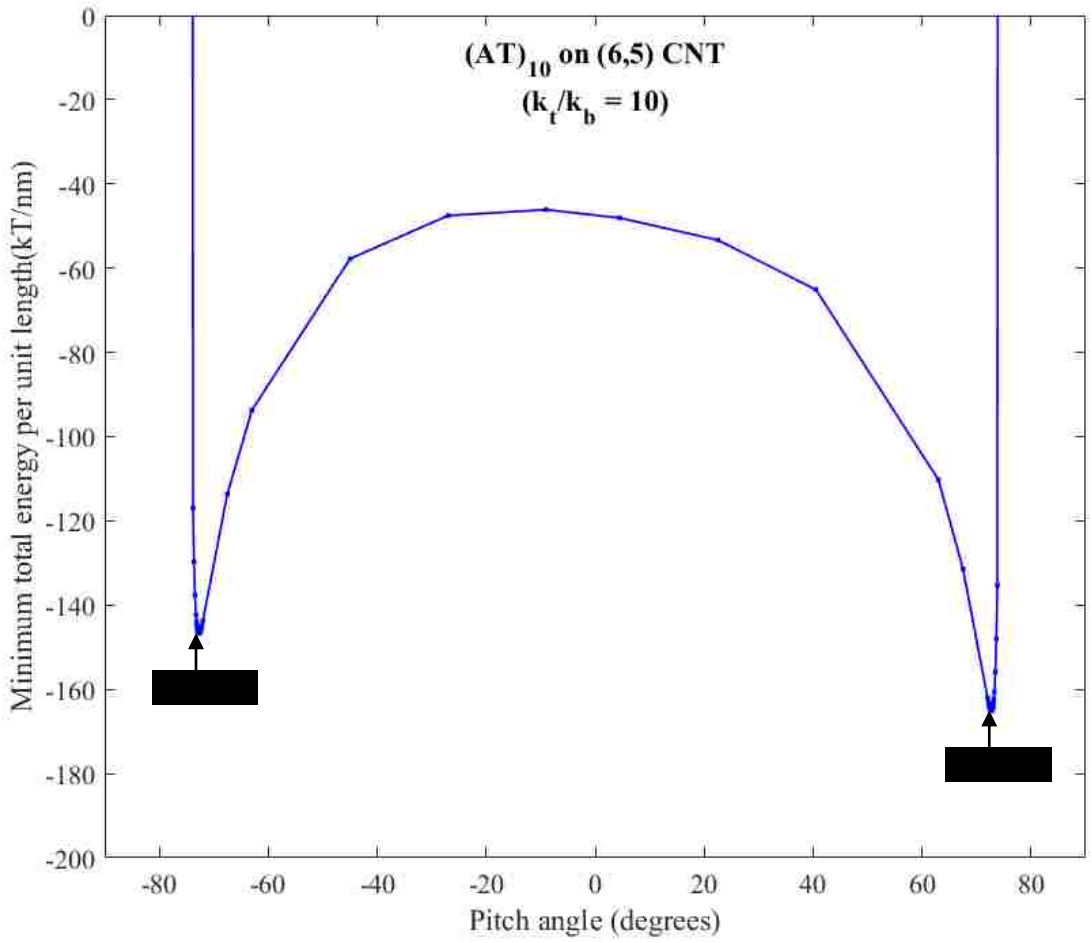


Figure 5.4: Minimum total average energy points corresponding to one DNA strand around nanotube, corresponding to left handed and right handed configurations.

5.3.1. Recognition of handedness requires spontaneous torsion

In order to explore systematically which contributions to the energy matter most for recognition of handedness, we calculated the total energy per unit length for a single strand of $(AT)_{10}$ on various achiral and chiral SWCNTs while considering the different contributions to the energy as shown in Table 5.3.

Table 5.3 Handedness preference for one DNA strand on various SWCNTs

CNT: Carbon nanotube chirality
DNA: DNA sequence
Adhesion: Adhesion between nanotube and DNA bases
Actual energy values = reference value (*) + actual number in the table

	CNT	DNA	Adhesion	Hydrogen Bonding	Bending	Torsion	Exclusion	AL (kT/nm)	AR (kT/nm)
1	(10,0)	$(AT)_{10}$	✓	✓	×	×	✓	-168.20 ± 0.01	-168.22 ± 0.01
2	(6,5)	$(AT)_{10}$	✓	✓	×	×	✓	* ± 0.006	0.1 ± 0.015
3	(5,6)	$(AT)_{10}$	✓	✓	×	×	✓	0.09 ± 0.016	0 ± 0.005
4	(6,5)	$(AT)_{10}$	✓	✓	✓	×	✓	0.08 ± 0.006	0.14 ± 0.014
5	(5,6)	$(AT)_{10}$	✓	✓	✓	×	✓	0.14 ± 0.014	0.08 ± 0.005
6	(6,5)	$(AT)_{10}$	✓	✓	✓	✓	✓	1.92 ± 0.006	0.18 ± 0.015
7	(5,6)	$(AT)_{10}$	✓	✓	✓	✓	✓	2.01 ± 0.014	0.11 ± 0.006

* -165.17 kT/nm is the reference value here

All the values except the first row are shown with respect to a fixed value indicated by an asterisk, as the adhesion energy is about two orders of magnitude larger than other energy contributions and we wish to focus on the relatively smaller differences in energies that account for recognition. First we consider the DNA sequence $(AT)_{10}$ wrapped on an achiral nanotube (10,0) in row 1 of Table 5.3. (10,0) is an achiral SWCNT which is very close to (6,5) and (5,6) in diameter. We consider the energy contributions from adhesion, hydrogen bonding and exclusion energy. From the values AL and AR in row 1, we see that the total energy as a function of helix pitch is perfectly symmetric with respect to pitch angle; there is equal preference for both handedness, (Note that in case of (10,0) in Table 5.3, the energy value is absolute and not relative to the asterisk.)

Next we considered $(AT)_{10}$ wrapped on chiral nanotubes enantiomers (6,5) and (5,6). We take into account the energy contributions from adhesion, hydrogen bonding and exclusion energy. As seen in row 2, column AL and AR, the DNA is slightly more stable as a left handed helix (AL) on (6,5) compared to the right handed helix (AR) on (6,5). The situation is reversed in case of (5,6) , as seen in row 3, where the right handed is slightly more stable. We can also see that a left-handed helix on (6,5) is exactly as stable as a right-handed helix on (5,6) and vice versa. It means that there is no basis for separation of (6,5) and (5,6) if we only consider the energy contributions from adhesion, hydrogen bonding and exclusion energy. This is because the lack of symmetry comes only from registration of base onto the SWCNT carbon atoms that have a chiral arrangement

We introduced another energy contribution apart from the ones mentioned in the previous case which is bending energy with a spontaneous bending angle of 30 degrees. As seen in row 4, the left handed helix on (6,5) is still slightly more stable as compared to the right handed helix. The reverse is still true for (5,6) and the opposite handed helices on the two SWCNT enantiomers still are nearly equally stable. This shows that bending energy cannot explain the possibility of separation of two SWCNT enantiomers wrapped by the same DNA sequence. Figure 5.3 shows that bending energy is actually symmetric around zero pitch. Therefore, if one starts with a mixture of enantiomers, there is again no basis for separation of chiral nanotube enantiomers.

In rows 7 and 8, we also considered the contribution of torsional energy with a spontaneous torsional angle of 15 degrees, in addition to the other four energy contributions. We used a ratio of constants for torsional and bending energies (k_t/k_b) as 1. Now we see that the right handed helix (AR) is more preferable for both the SWCNT enantiomers (6,5) and (5,6). However, the right handed helix is very marginally more stable on (5,6) as compared to (6,5). This energy difference is actually not so small considering that it is in the units kT/ nm and the average length of SWCNTs used in separation experiments vary between 200 to 700 nm. It is sufficient for separation to be possible in very sensitive separation techniques such as the aqueous two phase system. The crucial contributor that allows for difference in energy minimum is torsional energy, specifically, when one invokes a spontaneous torsion angle.

5.3.2. Recognition of handedness depends on DNA sequence, geometry and SWCNT chirality

The model used here to predict the total energy and hence the stability of the DNA-SWCNT hybrid relies on various parameters which are not very accurately known. This includes exact adhesion energy between the various DNA bases and SWCNT carbon atoms, the values for torsional rigidity k_t and distance between the backbone and base bead. The total energy of the system also depends on the DNA sequence composition and the SWCNT chirality and enantiomer.

In table 5.4, we consider a DNA strand wrapped around the SWCNT for eight different cases. From the previous section, we know that torsional energy is critical in enabling separation of DNA wrapped SWCNT enantiomers. We explore the effects of altering the value of torsional rigidity k_t as mentioned in equation 5.3. As explained in the appendix 5.6.3, k_t can be estimated to be anywhere between 1 to 60 times the value of the bending rigidity k_b. We look at one strand of $(AT)_{10}$ wrapped around (6,5) for three cases where k_t is 1, 10 and 20 times the values of k_b (see row 1,2 and 4 of table 5.4). For a low value of k_t, energy difference between the left handed and the right handed configurations is ~ 1.76 kT/nm, with the right handed one being preferred. On increasing k_t to 10 times k_b, this preference for the right handed configuration goes up to 18.03 kT/nm.

In all the data presented previously, we used the values of distance from backbone to base beads from a three bead model by Knotts et al.[118] Since we have a two bead model here, the distance is somewhat underestimated. In order to understand the effect of base-

backbone distance, we increase these values from 0.68 to 0.8 nm in case of Ade and 0.49 to 0.6 nm in case of Thy (see row 3 of table 5.4). In this case, the right handed configuration of $(AT)_{10}$ around the (6,5) SWCNT is the still more preferred one when the k_t is 10 times k_b. On further increasing k_t to 20 times k_b (which is still well within the range of values of k_t possible), we see in row 4 that the preference for the right configuration over the left has gone up to 35.19 kT/nm. This shows that the degree of preference for one handedness over the other depends on what value of torsional constant k_t is used as well as the geometrical values used.

In rows 4 and 5, the same set of parameters were used for multiple strands of $(AT)_{10}$ on (6,5) and (5,6) enantiomers. Both showed preference for the right handed helix. But there is an energy difference of 0.02 kT/nm between the right handed configuration on (6,5) and (5,6). This energy difference is actually not so small considering that it is in the units kT/ nm and the average length of SWCNTs used in separation experiments vary between 200 to 700 nm. This shows the basis of separability of the two enantiomers wrapped with the same DNA sequence. For this energy difference, the enantiomers can only be separated using a technique sensitive to small energy differences such as aqueous two phase system.[30]

For the same set of parameters, we changed the DNA sequence from $(AT)_{10}$ on (6,5) to A_{20} on (6,5) and (5,6) (see row 5 and 6). In case of A_{20}, which is also a 20mer like $(AT)_{10}$, the most preferred configuration is still the right handed single helix just as in case of of $(AT)_{10}$, but the energy difference between the left and right handed configurations changes marginally. Thus, for a given set of parameters, the preference for a particular

configuration depends on the exact DNA sequence. It is also seen that the right handed configurations of A_{20} on the two SWCNT enantiomers (6,5) and (5,6) have the same energy, suggesting that it is not necessarily possible to separate the SWCNT enantiomers using any DNA sequence. This is supported by previous work done by Geyou et al.[30] where different DNA sequences are used to separate different SWCNT chiralities and enantiomers.

For the data shown in table 5.3 and table 5.4 rows 1 to 7, we have used adhesion energy data from the peeling energy data of single stranded DNA from carbon nanotube surfaces by Iliafar et al.[28] which is considerably higher than on graphite substrate[27]. If we use the adhesion energy data from graphite instead, the adhesion energy becomes lower but the same preference for the right handed configuration is still exhibited (see row 8 of table 5.4). Thus our results are quite robust and do not depend on whether graphite or SWCNT substrate is used to calibrate the adhesion energy parameters.

Table 5.4 Energy differences for different DNA-CNT combinations and varying parameters

[Actual energy value = reference value (*) + actual number in the table]

	CNT	DNA	k_t/k_b	Substrate used for adhesion calibration	Base - Backbone Distance (Ade/Thy)	AL (kT/nm)	AR (kT/nm)
1	(6,5)	$(AT)_{10}$	1	CNT	0.64 / 0.49 nm	* (-163.18) ± 0.010	-1.76 ± 0.027
2	(6,5)	$(AT)_{10}$	10	CNT	0.64 / 0.49 nm	* (-146.60) ± 0.010	-18.03 ± 0.027
3	(6,5)	$(AT)_{10}$	10	CNT	0.8 / 0.6 nm	* (-126.23) ± 0.008	-17.59 ± 0.019
4	(6,5)	$(AT)_{10}$	20	CNT	0.8 / 0.6 nm	* (-108.53) ± 0.004	-35.19 ± 0.009
5	(5,6)	$(AT)_{10}$	20	CNT	0.8 / 0.6 nm	* (-108.52) ± 0.009	-35.21 ± 0.004

6	(6,5)	$(A)_{20}$	20	CNT	0.8 nm	* (-141.54) ± 0.007	-35.07 ± 0.019
7	(5,6)	$(A)_{20}$	20	CNT	0.8 nm	* (-141.54) ± 0.019	-35.07 ± 0.007
8	(6,5)	$(AT)_{10}$	1	Graphite	0.64/0.49 nm	* (-39.66) ± 0.002	-1.78 ± 0.007

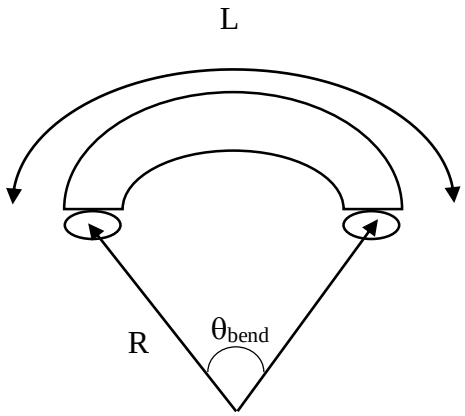
5.4. **Conclusion**

From previous surfactant experiments[25], we know that certain recognition DNA sequences have different activation energy of binding to their partner chiralities.[25,26] The same recognition sequences are also useful for separation of nanotube chiralities using the aqueous two phase technique[30], hinting that the hydrophobicity of these pairs are different which in turn indicates some difference in the coverage of the hydrophobic nanotube surface by the DNA strand. All this points to a conclusion that there is a difference in the structure of DNA on different SWCNT chiralities and enantiomers. This difference is only enough to cause a small difference in the energies of the system. But this small energy

difference is sufficient to permit separation of SWCNT enantiomers using DNA-assisted aqueous two phase separation which is very sensitive to very small energy differences. We also see that torsional energy is the only energy contribution which causes a difference in the stability of the opposite handed DNA helices on the nanotube enantiomers which enable the separation of SWCNT enantiomers using DNA assisted techniques.

5.5. **Appendix**

5.5.1. Estimating value of bending rigidity k_b for calculating bending energy



Consider a bent rod of length L. When bent, the radius of curvature is R, and angle is θ_{bend}.

Let k_f be the flexural rigidity and k_b the bending rigidity, defined by

$$E_{bend} = \frac{1}{2} * \frac{k_f}{L}\theta^2 = \frac{1}{2} * k_b\theta^2$$

Hence, $k_b = k_f/L$

By definition, when L = persistence length l_p, is

$$l_p = \frac{k_f}{k_B T} = \frac{k_b\, L}{k_B T}$$

Hence, $k_b = \frac{l_p}{L} * k_B T$

We assumed that persistence length l_p is about 5 bases. Hence $l_p = 4 * d_{pp.}$

L is length of bending angle ~ $d_{pp.}$

Hence, $k_b \sim l_p/L$ kT = 4 d_{pp} / d_{pp} = 4 kT

5.5.2. Estimate spontaneous bending angle and spontaneous torsional angle

Assuming that the spontaneous bending and torsion angles correspond to the double stranded B-DNA structure, we calculate a value for θ_0 as follows.

As seen in figure 5.1, 'a' is the helix diameter and $2\pi c$ is the pitch of the helix.

Now in B-DNA, pitch $2\pi c$ is 3.4 nm so c = pitch/(2π) = 0.54 nm [20]

's' is helical path.

Curvature κ of the helix is defined as follows:

$$\kappa = \frac{a}{a^2+c^2} = \frac{0.9989}{0.9989^2+0.54^2} = 0.7747\ nm^{-1} \quad (5.5)$$

$$\theta_0 = \kappa\, \Delta s = 0.7747\ nm^{-1} * d_{pp} = 0.7747\ nm^{-1} * 0.65\ nm$$

$$= 0.503555\ radians \sim 30\ degrees$$

Hence we use θ_0 = 30 degrees (note that θ_0 is in radians in the equation 5.2)

Just like θ_0, we decided to estimate ϕ_0 based on double stranded B-DNA.

Torsion τ of the helix is defined as follows:

$$\tau = \frac{c}{a^2+c^2} = \frac{0.54}{0.9989^2+0.54^2} = 0.4188\ nm^{-1} \quad (5.6)$$

$$\phi_0 = \tau \Delta s = 0.4188\ nm^{-1} * d_{pp} = 0.4188\ nm^{-1} * 0.65\ nm =$$

$$0.27222\ radians \sim 15\ degrees$$

Hence we use ϕ_0 = 15 degrees. (Note that ϕ_0 is in radians in the equation 5.3)

5.5.3. Estimate value of torsional rigidity k_t for calculating torsional energy

The bending rigidity k_b is proportional to the area moment of inertia I_{AA} while the torsional rigidity k_t is proportional to the moment of inertia about an axis I_{BB}.

We can get the upper and lower bounds for the value of k_t by making a number of assumptions.

a) Round cylinder

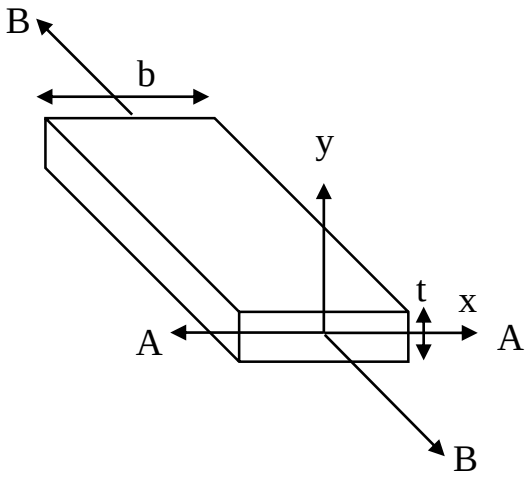
Assuming that the single stranded DNA is a round cylinder with radius r.

$$I_{AA} = \int y^2 dA = \frac{1}{2}\int (x^2 + y^2) dA = \frac{1}{2}\int r^2 dA = \frac{1}{2}\int r^2 . 2\pi r \, dr = \frac{1}{2}\int 2\pi \, r^3 dr = \frac{\pi r^4}{4}$$

$$I_{BB} = \int y^2 dA = \int r^2 . 2\pi r \, dr = \int 2\pi r^3 \, dr = \frac{\pi r^4}{2}$$

Hence $k_t / k_b = I_{BB} / I_{AA} = 2$

b) Thin rectagular block (ribbon)



Assuming that the single stranded DNA is a thin rectangular block with breadth b, thickness t

$$I_{AA} = \int_{-\frac{t}{2}}^{\frac{t}{2}} y^2 b \, dy = \frac{2b}{3} * \frac{t^3}{8} = \frac{bt^3}{12}$$

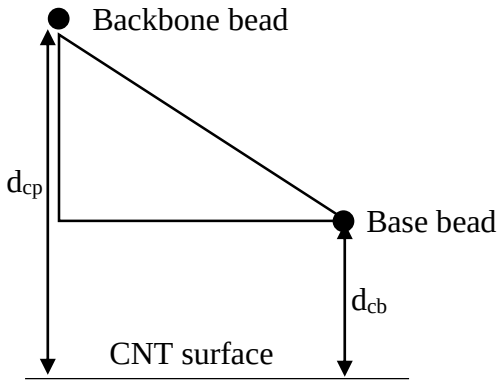
$$I_{BB} = \int_{-\frac{b}{2}}^{\frac{b}{2}} x^2 t \, dx = \frac{b^3 t}{12}$$

$$\frac{I_{BB}}{I_{AA}} = \left(\frac{b}{t}\right)^2$$

In our model, for (6,5) and (5,6) SWCNT, as per Table 5.1, distance between backbone and surface of nanotube = 0.592 nm and distance between base and surface of nanotube = 0.365 nm

$d_{pb\text{-}Ade}$ = 0.64 nm and $d_{pb\text{-}Thy}$ = 0.49 nm

For maximum hydrogen bonding energy, σ_{HB} for A-T = 0.53 nm, σ_{HB} for A-A = 0.59 nm , σ_{HB} for T-T = 0.77 nm



Assuming distance between two bases ~ σ_{HB}

$$b = 2 * \sqrt{d_{pb}^2 - \left(d_{cp} - d_{cb}\right)^2} + \sigma_{HB} = 2 * 0.6 + 0.59\ (for\ (Ade)_n) = 1.78\ nm$$

$$t = (d_{cp} - d_{cb}) = 0.227\ nm$$

$$\frac{k_t}{k_b} = \frac{I_{BB}}{I_{AA}} = \left(\frac{1.78}{0.227}\right)^2 = 61$$

So we can define k_t as 0.5 to 60 times the value of k_b. Thus, we estimate the value of k_t based on value of k_b and the ratio k_t / k_b assumed.

Chapter 6 Molecular dynamics simulations of closely related DNA sequences on nanotube enantiomers

6.1. **Introduction**

Single stranded DNA wrapped on single walled carbon nanotubes (SWCNTs) have been studied for various applications ranging from sorting of SWCNTs according to chirality[36] to molecular sensing[40,138] to cell imaging[42] and drug delivery. [17,38] A number of molecular dynamics studies have been conducted on the DNA-SWCNT hybrid previously in order to study the structure of DNA near SWCNTs.[21–24,33,34,132] As most of the DNA sequences used in these DNA-CNT hybrids are relatively short (from 6mers to 100mers), the system size is quite small and can be simulated using all atom models effectively. The number of types of interactions present in such systems are also sufficiently limited for the available force fields to describe these interactions. Roxbury et al. have reported a number of motifs formed by the DNA when wrapped around the nanotube surface , such as self-stitched structures[22], β-barrels[23], clasp structures[24] etc.

Certain special DNA sequences called 'recognition sequences' have been empirically found to selectively bind to certain SWCNT chiralities and enable their separation.[36] A number of studies of SWCNT-DNA hybrids have shown that many of their properties are highly sequence and chirality specific. For example, the recognition DNA sequence $(TAT)_4$ has been known to bind far more strongly to its partner SWCNT chirality (6,5) far more strongly as compared to very closely related sequences such as $(TAT)_3TA$ and $(TAT)_4T$.[26] It is also known that $(TAT)_4$ has higher affinity for its partner SWCNT

chirality (6,5) than for another SWCNT of exactly the same diameter i.e. (9,1).[25] More recently, recognition sequences have been used to separate the two enantiomers of (6,5) using aqueous two phase system. It is hypothesized that the differences in the properties of these DNA-SWCNT hybrids arise due to differences in the structure of the DNA on the SWCNT which may be both sequence and chirality specific. In order to develop a structural basis for experimental findings, we employed molecular dynamics to investigate a few of the ssDNA/CNT combinations.

6.2. **Models and simulation methods**

All atom Replica Exchange Molecular Dynamics (REMD)[89] simulations were performed to study the various structures that can be formed by single or multiple strands of single stranded DNA when placed close to a single walled carbon nanotube (SWCNT) surface. We studied closely related DNA strands from the (TAT) family as well as a palindromic sequence $TTA(TAT)_2ATT$ on (6,5) chirality SWCNT, its enantiomer (5,6) and another SWCNT of exactly the same size (9,1). The length of the (6,5) and (5,6) SWCNT were 7.97 nm and that of (9,1) was 8.161 nm. The diameter of all three SWCNTs was 0.746 nm. The end carbon atoms were covalently bonded to adjacent image carbons to mimic an infinitely long SWCNT. The nanotube co-ordinates were obtained using VMD Nanotube builder.[100]

The REMD simulations were performed using GROMACS 4.5.3 simulation package[90,91,139] with the CHARMM27[140] force field. The DNA-CNT hybrids were solvated

in a water box of size 7.97 nm x 3.46 nm x 3.46 nm in case of (6,5) and (5,6) and 8.161 nm x 3.46 nm x 3.46 nm in case of (9,1) with ~2,500 TIP3P model water molecules and the appropriate number of sodium counter ions to balance the negative phosphate charges. Total system size is about 10,000 atoms. Periodic boundary conditions were applied in all directions with long range electrostatics calculated using the particle mesh Ewald (PME) method. All structures were subjected to 100 ps heating (NVT) to attain a 300 K starting temperature. Forty replicas of each configuration were created for REMD NVT simulation, having temperatures ranging from 296 to 587 K. Replica temperatures were chosen such that exchange acceptance ratios remained around 10%, with an exchange time of 2 ps. Each of the one strand cases was run for 200 ns, for a total computation time of 40 × 200 = 8 μs. The first 50 ns were discarded as equilibration time. Each of the three strand cases were run for 400 ns, for a total computation time of 40 × 400 =16μs.The first 100 ns were discarded as equilibration time. The time step of the simulation was 2 fs. The trajectories were saved at every 10 ps, yielding a total of 15 000 snapshots and 30 000 snapshots for one strand and three strand cases of the DNA on SWCNTs, respectively, for production analysis.

6.3. **Results and Discussion**

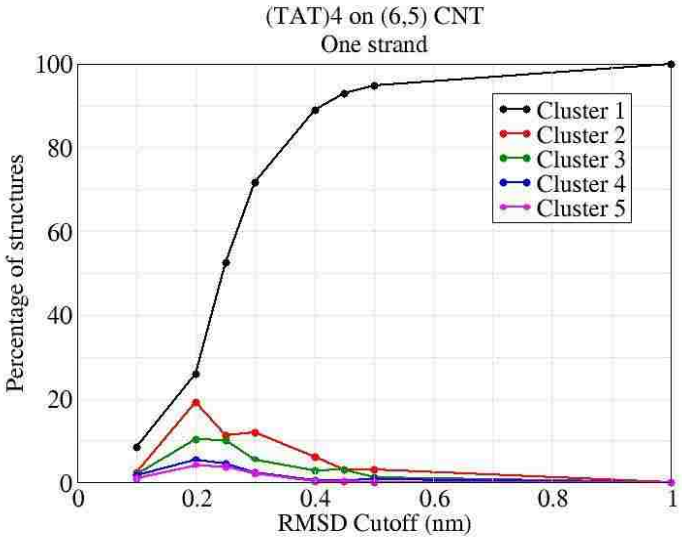
6.3.1. $(TAT)_4$ on (6,5) and (9,1) SWCNTs

a) Structure based clustering

We compared the structures formed by $(TAT)_4$ on (6,5) and (9,1) SWCNTs. The simulated equilibrium trajectories are clustered based upon the backbone atom positions of all available DNA residues, allowing one to extract the dominant structure for each stable configuration. Different cutoff sizes ranging from 0.1 to 1 nm for the one strand case and 0.1 to 2.5 nm for the three strand case are used for clustering the structures with respect to the backbone. In Figure 6.1 (a) and 1 (b) for three strands, the percentage of the structures in the majority cluster for the same cutoff for the two different SWCNT species are shown.

We first look at one strand of $(TAT)_4$ adsorbed onto the surface of the (6,5) and (9,1) SWCNTs. From Figure 6.2, we see that there is no significant difference in the typical structures of one strand of $(TAT)_4$ adsorbed onto the surface of the (6,5) and (9,1) SWCNTs.

(a)



(b)

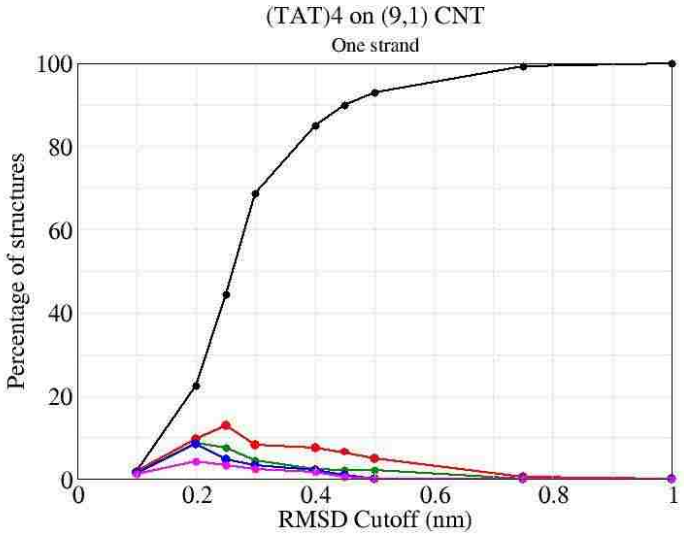


Figure 6.1 Structure based clustering for various cut off distances for $(TAT)_4$ on (6,5) and (9,1) SWCNTs

(a) Percentage of structures in first five majority clusters for $(TAT)_4$ on (6,5) chirality for various cut off distances with respect to the DNA backbone
(b) Percentage of structures in first five majority clusters for $(TAT)_4$ on (9,1) chirality for various cut off distances with respect to the DNA backbone

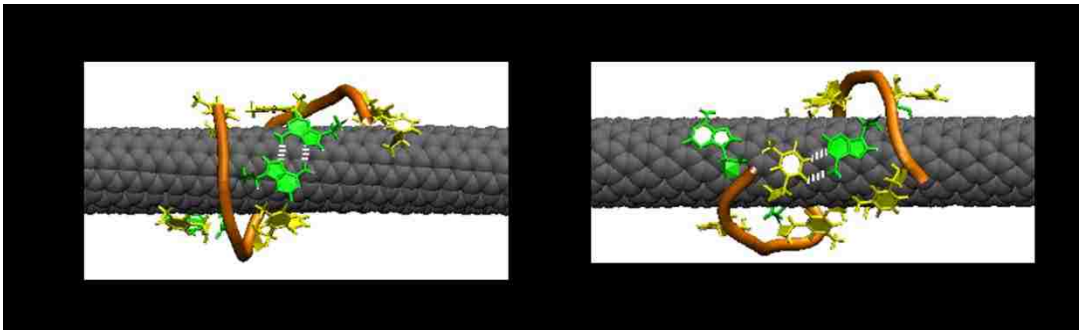


Figure 6.2 Typical structures of one strand of $(TAT)_4$ on (6,5) and (9,1) SWCNTs

Typical structures of one strand of $(TAT)_4$ on (6,5) and (9,1) SWCNTs in first two majority clusters based on the DNA backbone, using rmsd cutoff distance 0.25 nm

b) Hydrogen Bonding

It is known that hydrogen bonding between the DNA bases is a very important factor in the stabilization of DNA structures on the SWCNT surface.[22,24] Hence, we calculated the average number of hydrogen bonds formed during the simulation runs. The total number of hydrogen bonds are normalized by dividing by the number of bases in each strand, i.e., 12 in case of one strand and 36 in case of three strands each of $(TAT)_4$. The error bars are calculated by dividing the analyzed trajectory into ten equal parts, for each one of which we calculate the quantity of interest, using which we obtain the mean and standard deviation.

As shown in Figure 6.3 (a), we see that there is no significant difference in number of hydrogen bonds for one strand case of $(TAT)_4$ on (6,5) and on (9,1). But there seems to be a slightly higher number of hydrogen bonds present in case of $(TAT)_4$ on (9,1) than on (6,5) SWCNT. We further classified the total number of hydrogen bonds depending on the

types of bases involved in the hydrogen bonding. Since there are only two types of DNA bases, adenine and thymine present in the DNA sequence $(TAT)_4$, there are only three possible combinations i.e. adenine-adenine, adenine-thymine, and thymine-thymine. The numbers of hydrogen bonds between various types of bases are divided by the total number of hydrogen bonds present at that time instant to normalize them. We can compare these ratios with the probability of formation of hydrogen bonds if the bases were picked independently. For example, in case of $(TAT)_4$, combinations would appear with probability 64/144, 64/144 and16/144 for TT, AT, and AA respectively. We see that from Figure 6.3 (b) and (c) that there is significant deviation from these calculated probabilities. The majority of the hydrogen bonds stabilizing the DNA structure are between the Watson-Crick pair adenine and thymine. But there are also a significant number of non-Watson Crick pair hydrogen bonds between adenine and adenine, and thymine and thymine. But it is known from our experimental data in this paper that $(TAT)_4$ binds far more strongly to (6,5) than to (9,1). Hence it is possible that this difference in strength is due to difference in the DNA structure which would be visible as difference in number of hydrogen bonds and difference in the relative number of hydrogen bonds between various types of DNA bases.

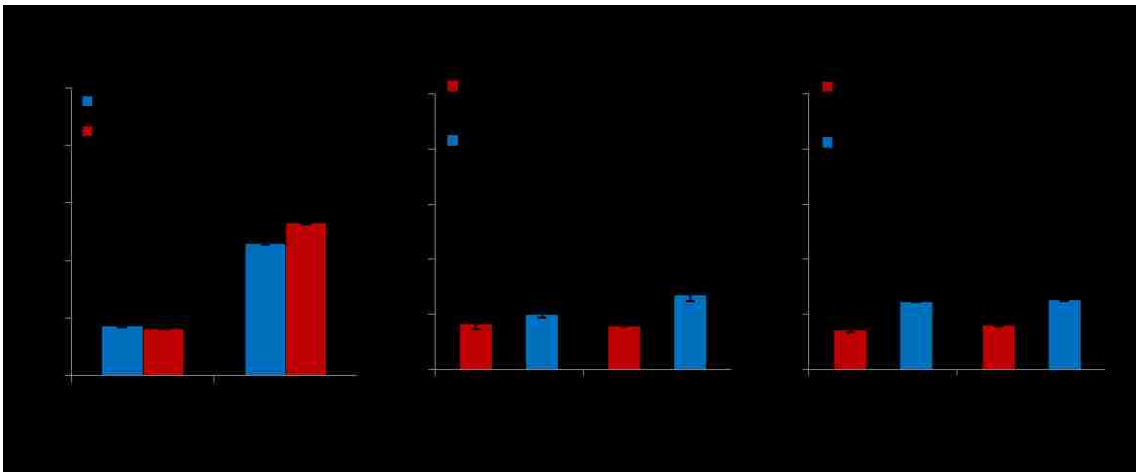


Figure 6.3 Normalized average number of hydrogen bonds formed for $(TAT)_4$ on (6,5) and (9,1)

Normalized average number of hydrogen bonds formed (a) in total, (b) between various kinds of bases for one DNA strand on SWCNT and, (c) between various kinds of bases for three DNA strands on SWCNT

c) Solvent accessible surface area of the SWCNT

According to Roxbury et al.[26] the replacement of the DNA on the SWCNT surface by the SDBS requires creation of a defect on the DNA covered SWCNT surface. Hence the coverage of the SWCNT surface by the DNA is an important parameter in explaining the binding strength of DNA onto the SWCNT surface.

For the two cases i.e. three strands of $(TAT)_4$ on (6,5) and $(TAT)_4$ on (9,1), the amount of SWCNT surface area excluded from water by the DNA is calculated (by placing a sphere the size of a water molecule near the surface of the SWCNT).

In Figure 6.4, it is seen that there is a difference in the solvent accessible surface area for both cases. Here the error bars are again calculated by dividing the analyzed trajectory into ten equal parts and calculating the standard deviation. For the one strand cases, the solvent

accessible surface area obviously is higher than that for the three strands as one DNA strand has lesser coverage on the SWCNT surface than three strands. The difference in the solvent accessible surface area on the two SWCNTs means that the DNA structures are not equally well-packed on both SWCNT species, again suggesting that there is a difference in the DNA structures on the two SWCNT surfaces.

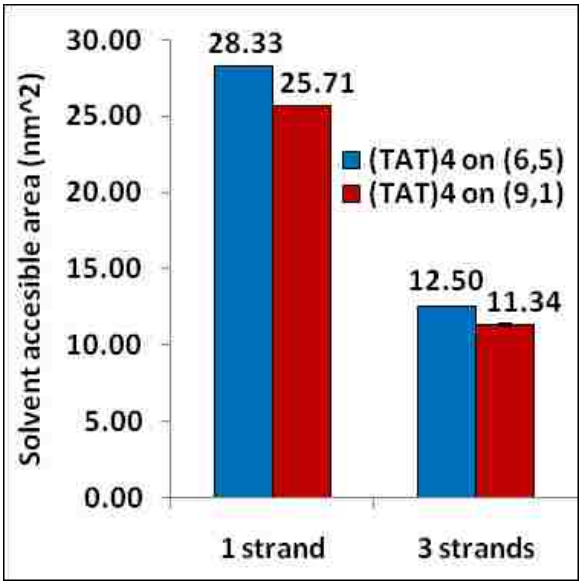


Figure 6.4 Average solvent accessible surface area for the SWCNT surface for $(TAT)_4$ wrapped (6,5) and (9,1) SWCNT species

6.3.2. (TAT) family on the (6,5) SWCNT

Similarly, we suppose that the difference in binding strength for different lengths of DNA sequences is because of difference in the DNA structures formed. In order to probe

the validity of this statement, we compare the structures predicted by molecular simulation formed by 5 different lengths ranging from 10 to 14mer DNA sequences belonging to the (TAT) family on the same SWCNT type i.e. (6,5).

a) Structure based clustering

We compared the structures formed by one strand placed on the (6,5) SWCNTs. The simulated equilibrium trajectories are clustered based upon the backbone atom positions of all available DNA residues, allowing one to extract the dominant structure for each stable configuration. A cutoff size of 0.25 nm was used for clustering the structures with respect to the backbone. In Figure 6.5, the percentage of the structures in the majority cluster for the same cutoff for the different closely related DNA sequences on the same SWCNT (6,5) are shown. The recognition sequence for (6,5) which is $(TAT)_4$ shows a much larger percentage of structures in the majority cluster, indicating that the $(TAT)_4$ on (6,5) is more ordered than the other DNA sequences.

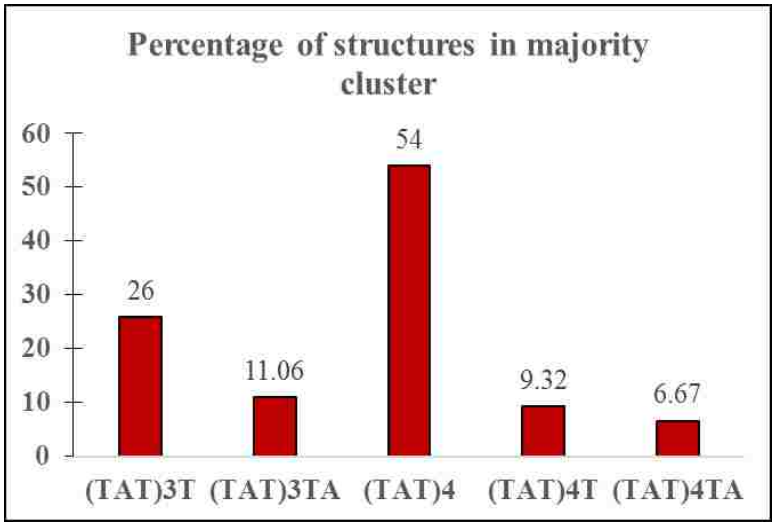


Figure 6.5 Percentage of structures in the majority clusters for different closely related DNA sequences on (6,5) chirality for 0.25 nm rmsd cut off distance with respect to the DNA backbone

We calculated the average number of hydrogen bonds formed during the simulation runs. The total number of hydrogen bonds are normalized by dividing by the number of bases in each strand i.e. 30, 33, 36, 39 and 42 in case of three strands each of $(TAT)_3T$, $(TAT)_3TA$, $(TAT)_4$, $(TAT)_4T$ and $(TAT)_4TA$ respectively. The error bars are calculated by dividing the analyzed trajectory into ten equal parts, for each one of which we calculate the quantity of interest, using which we obtain the mean and standard deviation.

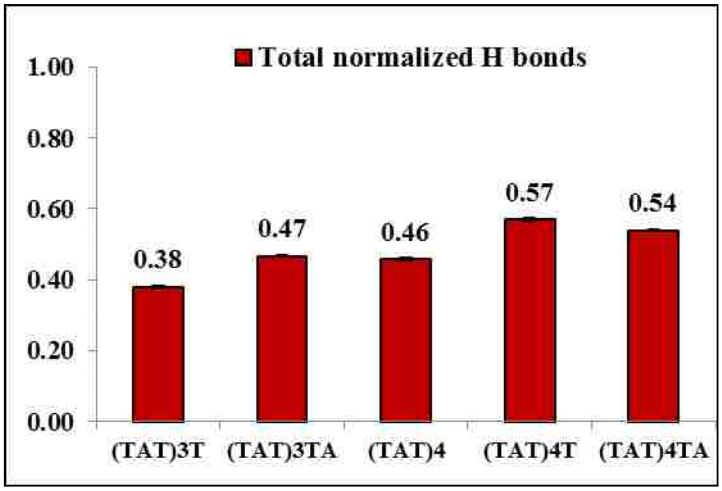
As shown in Figure 6.6 (a), we see that there is very slight difference in number of hydrogen bonds for three strand cases of $(TAT)_3T$, $(TAT)_3TA$, $(TAT)_4$, $(TAT)_4T$ and $(TAT)_4TA$ on (6,5) SWCNT. As done previously, we also further classified the total number of hydrogen bonds depending on the types of bases involved in the hydrogen bonding. The numbers of hydrogen bonds between various types of bases are divided by

the total number of hydrogen bonds present at that time instant to normalize them. We can compare these ratios with the probability of formation of hydrogen bonds if the bases were picked independently. For example, in case of $(TAT)_3T$, combinations would appear with probability 49/100, 42/100 and 9/100 for TT, AT, and AA respectively.

We see that from Figure 6.6 (b) that there is significant deviation from these calculated probabilities in some cases. The majority of the hydrogen bonds stabilizing the DNA structure are between the Watson-Crick pair adenine and thymine. But there are also a significant number of non-Watson Crick pair hydrogen bonds between adenine and adenine, and thymine and thymine. Hence it is possible that the difference in strength is due to difference in the DNA structure which would be visible as difference in the relative number of hydrogen bonds between various types of DNA bases.

b) Hydrogen Bonding

(a)



(b)

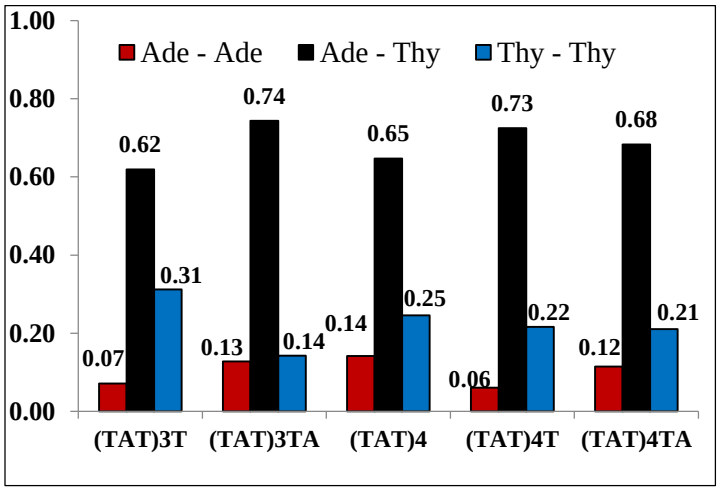


Figure 6.6 Normalized average number of hydrogen bonds formed

(a) in total, and (b) between various kinds of bases for three DNA strands on SWCNT

c) Self-stitching

Hydrogen bond self-stitching has been thought to be a significant factor in stabilizing single strands of DNA on the nanotube surface.[22] A 'stitch' is said to be present between two bases if distance between base centroids is less than 9 Angstroms and the bases are more than 7 bases away from each other. Table 6.1. shows the percentage of the simulation time in which the stitches are present and also the number of stitches present. It is seen that in case of the recognition sequence $(TAT)_4$, at least one stitch is present for the highest percentage of the simulation time as compared to very closely related sequences on the same SWCNT. This hints that the structures formed by the recognition sequence are stabilized by significantly more 'self-stitch' hydrogen bonds than their closely relates sequences.

Table 6.1 Percentage of the simulation time in which one or more stitches are present

Sequence	CNT	Percentage of time with:				
		1 stitch	2 stitches	3 stitches	4 stitches	At least 1 stitch
$(TAT)_3T$	(6,5)	5.79 %	0 %	0 %	0 %	5.79 %
$(TAT)_3TA$	(6,5)	11.18 %	4.79 %	2.79 %	0 %	18.76%
$(TAT)_4$	(6,5)	39.32 %	16.17 %	2.20 %	0 %	57.68 %
$(TAT)_4T$	(6,5)	19.16 %	9.18 %	3.99 %	0 %	32.34 %
$(TAT)_4TA$	(6,5)	11.58 %	8.78 %	6.58 %	0.6 %	27.74

d) Helicity

To quantify the DNA helical handedness, phosphorus atom positions were identified and monitored. For example, $(TAT)_4$ has 11 P atoms i.e. 10 adjacent P pairs. Local helical angles determined for each P pair (10 local helical angles per structure).

While traversing the SWCNT axis (with increasing axial coordinate), phosphorus pairs determined to be locally right-handed were assigned a "1", else assigned a "0" for a locally left-handed helix. Because of intrinsic DNA-backbone flexibility, a DNA strand with a visibly right-handed overall helical conformation may have a few local left-handed helical angles

If 8 out of 10 local helical angles were assigned "1", the structure was classified as right-handed and vice versa for left-handed. Furthermore, if a structure contained 5 consecutive local helical angles with the designation "1", without previous right-handed classification, it was also delineated as such. The same applies for left-handed if there were 5 in a row with designation "0". Structures not designated either right- or left-handed were termed "unclear" and generally formed a loop structure.

Table 6.2 shows that a single strand of recognition sequence $(TAT)_4$ on (6,5) shows a right-handed structure for majority of the time as compared to the closely related sequences which are left-handed for a majority of the simulation time.

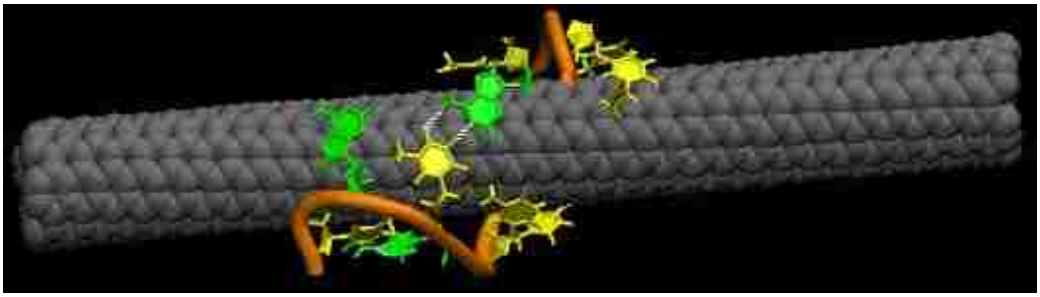
Table 6.2 DNA handedness as percentage of simulation time for one DNA strand on SWCNT

DNA sequence	SWCNT	Left-handed	Right-handed	Unclear
$(TAT)_3T$	(6,5)	85.4 %	0 %	14.6 %
$(TAT)_3TA$	(6,5)	79.5 %	0 %	20.5 %
$(TAT)_4$	(6,5)	0 %	96.1 %	3.9 %
$(TAT)_4T$	(6,5)	79.5 %	0 %	31.7 %
$(TAT)_4TA$	(6,5)	63.6 %	0 %	35.8 %

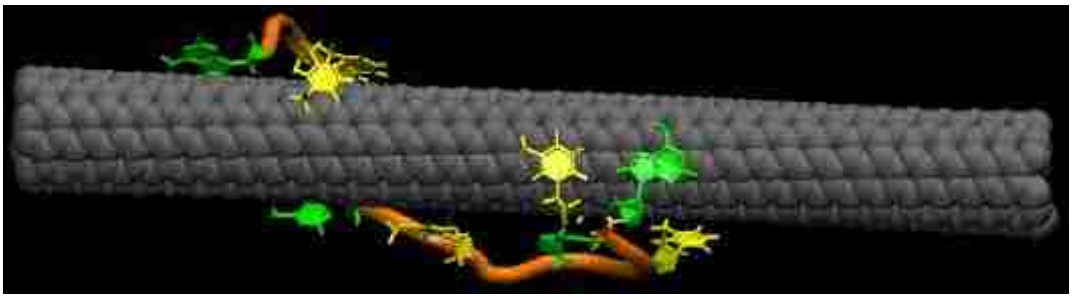
6.3.3. $(TAT)_4$ on (6,5) and (5,6) SWCNTs

a) Structure based clustering

We compared the structures formed by one strand of $(TAT)_4$ placed on the (6,5) and (5,6) SWCNTs. The simulated equilibrium trajectories are clustered based upon the backbone atom positions of all available DNA residues, allowing one to extract the dominant structure for each stable configuration. A cutoff size of 0.25 nm was used for clustering the structures with respect to the backbone. Figure 6.8 shows that the recognition sequence for (6,5) which is $(TAT)_4$ shows a much larger percentage of structures in the majority cluster, indicating that the $(TAT)_4$ on (6,5) is more ordered than on its enantiomer (5,6) SWCNT.



(a)



(b)

Figure 6.7 One strand of $(TAT)_4$ on (6,5) and (5,6) SWCNT

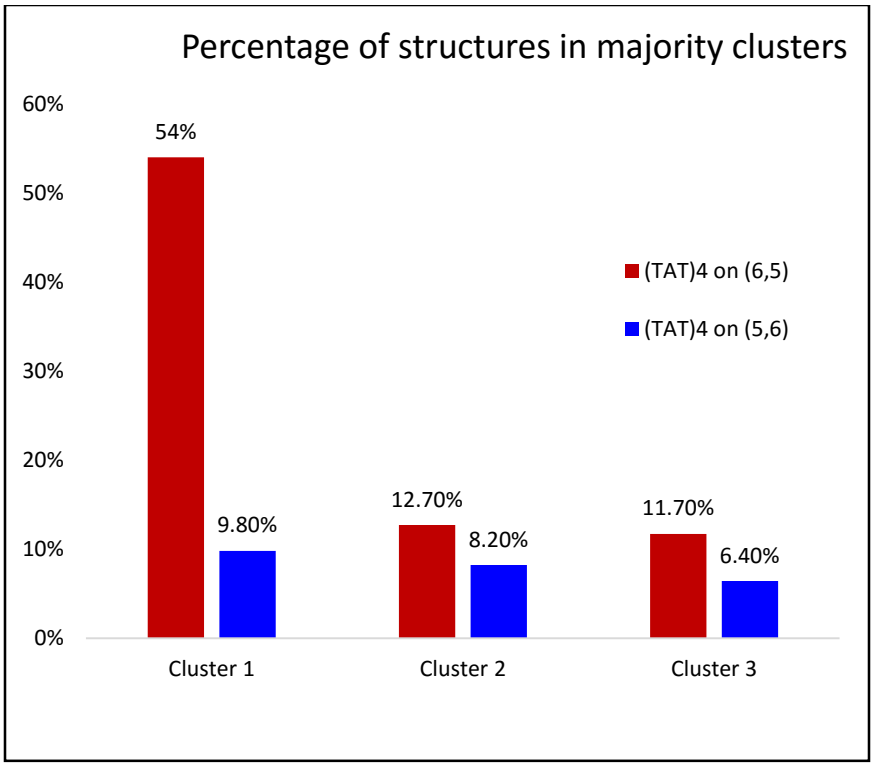


Figure 6.8 Percentage of structures in the majority clusters for different closely related DNA sequences on (6,5) chirality
(for 0.25 nm rmsd cut off distance with respect to the DNA backbone)

b) Helicity

Table 6.3 shows that a single strand of recognition sequence $(TAT)_4$ on (6,5) shows a right-handed structure for majority of the time as compared to $(TAT)_4$ on its enantiomer SWCNT (5,6) which are left-handed for a majority of the simulation time. This indicates clear structural difference for the same DNA sequence on enantiomer SWCNTs

Table 6.3 DNA handedness as percentage of simulation time for one DNA strand on SWCNT

	$(TAT)_4$ on (6,5)	**$(TAT)_4$ on (5,6)**
Left handed	0 %	68.6 %
Right handed	96.1 %	0 %
Unclear	3.9 %	31.4 %

c) Self-stitching

Table 6.4. shows the percentage of the simulation time in which the stitches are present and also the number of stitches present. It is seen that in case of the recognition sequence $(TAT)_4$, on (6,5) SWCNT at least one stitch is present for a higher percentage of the simulation time as compared to $(TAT)_4$, on (5,6) SWCNT. This hints that the structures formed by the recognition sequence on one enantiomer are stabilized by significantly more 'self-stitch' hydrogen bonds than on the other SWCNT enantiomer.

Table 6.4 Percentage of the simulation time in which one or more stitches are present

Sequence	CNT	Percentage of time with:				
		1 stitch	**2 stitches**	**3 stitches**	**4 stitches**	**At least 1 stitch**
$(TAT)_4$	(6,5)	39.32 %	16.17 %	2.20 %	0 %	57.68 %
$(TAT)_4$	(5,6)	12.37 %	0 %	0 %	0 %	12.37%

6.3.4. $TTA(TAT)_2ATT$ on (6,5) and (5,6) SWCNTs

The palindromic sequence $TTA(TAT)_2ATT$ is known to have the ability to bind to both (6,5) and (5,6) and separate in an aqueous two phase system. We hypothesize that this difference is due to difference in the hydration energy of the two hybrids which in turn is related to difference in the structure of this DNA strand on the two SWCNT enantiomers. Figure 6.9 and table 6.5 both indicate that there is very little difference observed for the palindromic sequence $TTA(TAT)_2ATT$ on the two SWCNT enantiomer sequences (6,5) and (5,6). It is possible that the force field used for these simulation cannot model small differences.

a) Structure based clustering

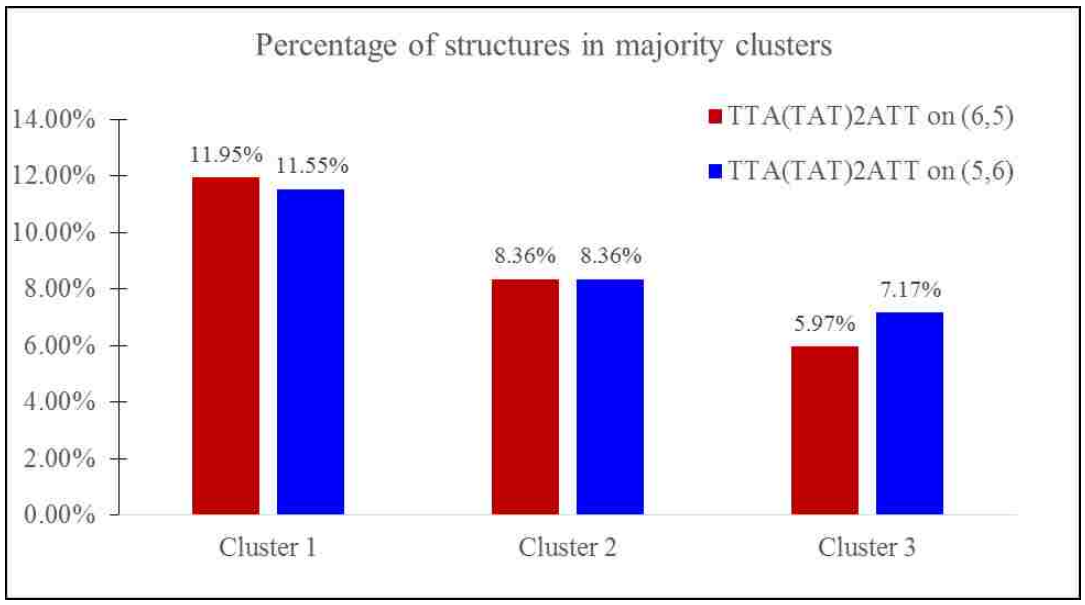


Figure 6.9 Percentage of structures in the majority clusters for different closely related DNA sequences on (6,5) chirality **(**for 0.25 nm rmsd cut off distance with respect to the DNA backbone)

b) Helicity

Table 6.5 DNA handedness as percentage of simulation time for one DNA strand of palindromic sequence on SWCNT

	One strand of $TTA(TAT)_2ATT$ on (6,5)	One strand of $TTA(TAT)_2ATT$ on (5,6)
Left handed	68.26 %	71.46 %
Right handed	0 %	1.00 %
Unclear	31.74 %	27.54 %

6.4. **Conclusion**

The simulations showed significant differences between the structures of the recognition sequence $(TAT)_4$ on its partner chirality (6,5) and closely related sequences. However, on conducting a repeat simulation of the same hybrid, we found that the handedness had reversed from right-handed to left-handed like the remaining closely related sequences. It is possible that the simulation has not converged sufficiently and depends on the initial structure. Further investigation is required before any conclusions are drawn from these simulations.

Chapter 7 Summary and future work

7.1. **Sequence and chirality dependence of binding**

We reported the activation energies for removal of several ssDNA sequences from a few SWCNT species by a surfactant molecule. We found that DNA sequences systematically have higher activation energy on their carbon-nanotube recognition partner than on non-partner species. For example, the DNA sequence $(TAT)_4$ has much lower activation energy on the (9,1) SWCNT than on its partner (6,5) SWCNT whereas the DNA sequence $(CCA)_{10}$ binds strongly to its partner (9,1) SWCNT compared to (6,5). The (6,5) and (9,1) SWCNTs have the same diameter but different electronic properties, suggesting that activation energy difference is related to electronic properties. But it must be noted that recent experiments by Geyou et al. have shown that SWCNT enantiomers wrapped with the same DNA sequence can be separated, hinting that certain DNA sequences do have the ability to differentiate between the handedness of the SWCNT. The activation energies of increasing lengths of closely related sequences from the 11mer $(TAT)_3TA$ to the 21mer $(TAT)_7$ on three different SWCNT species (9,1), (6,5), and (8,3) were measured. For the shorter sequences, the activation energy on the SWCNT varies periodically with sequence.

7.2. **Relative hydration of DNA- single walled carbon nanotube hybrids using aqueous two phase system**

Since the difference in binding affinity and difference in partitioning can depend on DNA structure on the single walled carbon nanotube (SWCNT), we studied the partitioning of the various DNA sequences in an aqueous two phase system. DNA-SWCNT partitioning in the aqueous two phase is because of sensitive dependence of the free energy of hydration on the spatial distribution of hydrophilic groups in the DNA-SWCNT hybrid. In this way, the aqueous two phase process is at the same time a technique for separation and a method by which to evaluate and rank hydration or solvation free energy. We found that $(CCA)_{10}$ on (6,5) SWCNT requires much higher amount of modulant to be moved from the relatively hydrophilic phase to the more hydrophilic phase as compared to $(GT)_{15}$ on (6,5) even though they are both 30mers, suggesting that the solvation energy depends greatly on the DNA sequence composition. We also found that various sequences with same length but different repeating units of two bases exhibit different hydration energies on the same SWCNT (6,5).

7.3. **DNA base dimers are stabilized by hydrogen bonding interactions including non-Watson-Crick pairing near graphite surfaces**

We find that base pairs in bulk water are stabilized by stacking interactions as expected. The order of stability in bulk liquid is found to be (Gua-Gua > Ade-Gua > Ade-Ade >

Gua-Thy > Gua-Cyt > Ade-Thy > Ade-Cyt > Thy-Thy > Cyt-Thy > Cyt-Cyt). Hydrogen bonding interactions are not observed in any of these cases, indicating that stacking interactions are much stronger than hydrogen bonding interactions under these conditions. In case of base pairs adsorbed on a graphite surface (strong base-surface stacking due to π–π interactions), because the propensity of bases to stack is satisfied by their adsorption onto the surface, dimer stabilization is dominated by hydrogen bonding interactions. The order of strength for hydrogen bonding interactions is (Gua-Cyt > Ade-Gua > Ade-Thy > Ade-Ade > Cyt-Thy > Gua-Gua > Cyt-Cyt > Ade-Cyt > Thy-Thy > Gua-Thy). Several non-Watson-Crick base pairings, that are commonly ignored, have similar stabilization free energies due to inter-base hydrogen bonding as Watson-Crick pairs. This clearly highlights the importance of non-Watson-Crick base pairing in the development of secondary structures of oligonucleotides near surfaces.

7.4. Energetic Basis of Single Wall Carbon Nanotube Enantiomer Recognition by Single Stranded DNA

DNA assisted separation by handedness of SWCNTs requires that a given single stranded DNA sequence adopt different structures on the two SWCNT enantiomers. We studied the physical basis of such selectivity using a coarse grained model to compute the energetics of ssDNA wrapped around an SWCNT. Our model suggests that difference by handedness of the SWCNT requires spontaneous twist of the ssDNA backbone. We also show that differences depend sensitively on the choice of DNA sequence.

7.5. Molecular Dynamics simulations of closely related DNA sequences on closely related carbon nanotubes

In order to develop a structural basis for previous experimental findings which found differences in SWCNT-DNA hybrids based on both sequence and chirality, we employed molecular dynamics to investigate a few of the ssDNA/CNT combinations. We were able to find differences but since the repeat simulations did not show the same result, it is not possible to draw any conclusions from this work. Further investigation is required for this.

7.6. Future work

In chapter 2, we studied binding affinity for recognition sequences on their partner and non-partner chiralities. It may be useful to study a larger library of DNA sequences and SWCNT chiralities to better understand the correlation between recognition and binding affinity. Another important study that could be carries out is equilibrium state exchange between DNA wrapped SWCNTs and surfactant wrapped SWCNTs. This can yield equilibrium free energy of displacement of the DNA from the SWCNT surface. Analysis of molecular dynamics simulations may help understand why the binding affinity as a function of length of DNA sequence is periodic with period of two bases.

In chapter 4, the preliminary findings are very promising and suggest that the aqueous two phase system can be used as an effective method to evaluate and possibly rank the solvation free energy of various DNA sequence - SWCNT chirality combinations. Various sequences in the four dimensional sequence space from say $(AT)_{15}$ to $(AC)_{15}$ can be studied systematically. Additionally, the four 30mer homopolymers can also be compared to the sequences studied in this work. Additionally, similar studies can be conducted on various SWCNT chiralities and enantiomers also.

The model developed in chapter 5 may be further expanded to Monte-Carlo or molecular dynamics simulations to sample the possible configurations of the DNA-SWCNT hybrid more effectively.

References

1. Iijima, S. & Ichihashi, T. Single-shell carbon nanotubes of 1-nm diameter. *Nature* **363,** 603–605 (1993).
2. Teri Wang Odom, Jin-Lin Huang, and C. M. L. Single-Walled Carbon Nanotubes From Fundamental Studies to New Device Concepts. (2002).
3. Baughman, R. H., Zakhidov, A. A. & de Heer, W. A. Carbon nanotubes--the route toward applications. *Science (80-.).* **297,** 787–792 (2002).
4. Green, A. A., Duch, M. C. & Hersam, M. C. Isolation of Single-Walled Carbon Nanotube Enantiomers by Density Differentiation. *Nano Res* **2,** 69–77 (2009).
5. Thostenson, E. T., Ren, Z. & Chou, T.-W. Advances in the science and technology of carbon nanotubes and their composites: a review. *Compos. Sci. Technol.* **61,** 1899–1912 (2001).
6. Ajayan, P. M. Nanotubes from Carbon. *Chem. Rev.* **99,** 14 (1999).
7. Sinha, N. & Yeow, J. T. W. Carbon nanotubes for biomedical applications. *IEEE Trans. Nanobioscience* **4,** 180–195 (2005).
8. Javey, A. *et al.* Carbon nanotube field-effect transistors with integrated ohmic contacts and high-k gate dielectrics. *Nano Lett.* **4,** 447–450 (2004).
9. Hu, L., Hecht, D. S. & Grüner, G. Carbon nanotube thin films: Fabrication, properties, and applications. *Chem. Rev.* **110,** 5790–5844 (2010).
10. O'Regan, B. & Grätzel, M. A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO2 films. *Nature* **353,** 737–740 (1991).
11. Gong, K., Du, F., Xia, Z., Durstock, M. & Dai, L. Nitrogen-doped carbon

nanotube arrays with high electrocatalytic activity for oxygen reduction. *Science (80-.).* **323,** 760–4 (2009).

12. Wang, X., Li, W., Chen, Z., Waje, M. & Yan, Y. Durability investigation of carbon nanotube as catalyst support for proton exchange membrane fuel cell. *J. Power Sources* **158,** 154–159 (2006).
13. Guiseppi-Elie, A., Lei, C. & Baughman, R. H. Direct electron transfer of glucose oxidase on carbon nanotubes. *Nanotechnology* **13,** 559–564 (2002).
14. Gooding, J. J. *et al.* Protein electrochemistry using aligned carbon nanotube arrays. *J. Am. Chem. Soc.* **125,** 9006–9007 (2003).
15. Goldsmith, B. R. Conductance-Controlled Point. **77,** 77–82 (2007).
16. Bhirde, A. A. *et al.* Targeted killing of cancer cells in vivo and in vitro with EGF-directed carbon nanotube-based drug delivery. *ACS Nano* **3,** 307–16 (2009).
17. Kavitha, T., Abdi, S. I. H. & Park, S.-Y. pH-sensitive nanocargo based on smart polymer functionalized graphene oxide for site-specific drug delivery. *Phys. Chem. Chem. Phys.* **15,** 5176–85 (2013).
18. Nucleic Acids. at <http://chemed.chem.purdue.edu/genchem/topicreview/bp/1biochem/nucleic8.html>
19. NUCLEOTIDES AND THE DOUBLE HELIX. at <http://cyberbridge.mcb.harvard.edu/dna_1.html>
20. Saenger, W. *Principles of Nucleic Acid Strucure*. (Springer-Verlag, 1983).
21. Roxbury, D., Jagota, A. & Mittal, J. Structural Characteristics of Oligomeric DNA

Strands Adsorbed onto Single-Walled Carbon Nanotubes. *J. Phys. Chem. B* **117,** 132–140 (2013).

22. Roxbury, D., Jagota, A. & Mittal, J. Sequence-specific self-stitching motif of short single-stranded DNA on a single-walled carbon nanotube. *J. Am. Chem. Soc.* **133,** 13545–13550 (2011).

23. Roxbury, D., Manohar, S. & Jagota, A. Molecular Simulation of DNA B-Sheet and B-Barrel Structures on Graphite and Carbon Nanotubes. *J. Phys. Chem. C* **114,** 13267–13276 (2010).

24. Roxbury, D., Mittal, J. & Jagota, A. Molecular-basis of single-walled carbon nanotube recognition by single-stranded DNA. *Nano Lett.* **12,** 1464–1469 (2012).

25. Shankar, A., Mittal, J. & Jagota, A. Binding between DNA and carbon nanotubes strongly depends upon sequence and chirality. *Langmuir* **30,** 3176–83 (2014).

26. Roxbury, D., Tu, X., Zheng, M. & Jagota, A. Recognition ability of DNA for carbon nanotubes correlates with their binding affinity. *Langmuir* **27,** 8282–93 (2011).

27. Iliafar, S., Wagner, K., Manohar, S., Jagota, A. & Vezenov, D. Quantifying Interactions between DNA Oligomers and a Graphite Surface Using Single Molecule Force Spectroscopy. *J. Phys. Chem. C* **116,** 13896–13903 (2012).

28. Iliafar, S., Mittal, J., Vezenov, D. & Jagota, A. Interaction of single-stranded DNA with curved carbon nanotube is much stronger than with flat graphite. *J. Am. Chem. Soc.* **136,** 12947–57 (2014).

29. Manohar, S. *et al.* Peeling Single-Stranded DNA from Graphite Surface to

Determine Oligonucleotide Binding Energy by Force Spectroscopy. *Nano Lett.* **8,** 4365–4372 (2008).

30. Ao, G., Khripin, C. Y. & Zheng, M. DNA-controlled partition of carbon nanotubes in polymer aqueous two-phase systems. *J. Am. Chem. Soc.* **136,** 10383–10392 (2014).

31. Bachilo, S. M. *et al.* Structure-assigned optical spectra of single-walled carbon nanotubes. *Science (80-.).* **298,** 2361–6 (2002).

32. Choi, J. H. & Strano, M. S. Solvatochromism in single-walled carbon nanotubes. *Appl. Phys. Lett.* **90,** 223114 (2007).

33. Johnson, R. R., Kohlmeyer, A., Johnson, a T. C. & Klein, M. L. Free energy landscape of a DNA-carbon nanotube hybrid using replica exchange molecular dynamics. *Nano Lett.* **9,** 537–41 (2009).

34. Johnson, R. R., Johnson, a T. C. & Klein, M. L. Probing the structure of DNA-carbon nanotube hybrids with molecular dynamics. *Nano Lett.* **8,** 69–75 (2008).

35. Castner, D. G. & Ratner, B. D. Biomedical surface science: Foundations to frontiers. *Surf. Sci.* **500,** 28–60 (2002).

36. Tu, X., Manohar, S., Jagota, A. & Zheng, M. DNA sequence motifs for structure-specific recognition and separation of carbon nanotubes. *Nature* **460,** 250–3 (2009).

37. Kam, N. W. S., Liu, Z. & Dai, H. Functionalization of carbon nanotubes via cleavable disulfide bonds for efficient intracellular delivery of siRNA and potent gene silencing. *J. Am. Chem. Soc.* **127,** 12492–3 (2005).

38. Liu, Z. *et al.* Drug delivery with carbon nanotubes for in vivo cancer treatment. *Cancer Res.* **68,** 6652–60 (2008).

39. Staii, C., Johnson, A. T., Chen, M. & Gelperin, A. DNA-decorated carbon nanotubes for chemical sensing. *Nano Lett.* **5,** 1774–8 (2005).

40. Heller, D. A. *et al.* Multimodal optical sensing and analyte specificity using single-walled carbon nanotubes. **4,** (2009).

41. Cha, T.-G. *et al.* Optical nanosensor architecture for cell-signaling molecules using DNA aptamer-coated carbon nanotubes. *ACS Nano* **5,** 4236–44 (2011).

42. Welsher, K., Liu, Z., Daranciang, D. & Dai, H. Selective probing and imaging of cells with single walled carbon nanotubes as near-infrared fluorescent molecules. *Nano Lett.* **8,** 586–90 (2008).

43. De la Zerda, A. *et al.* Carbon nanotubes as photoacoustic molecular imaging agents in living mice. *Nat. Nanotechnol.* **3,** 557–62 (2008).

44. Tao, N. J. & Shi, Z. Monolayer Guanine and Adenine on Graphite in NaCl Solution: A Comparative STM and AFM Study. *J. Phys. Chem.* **98,** 1464–1471 (1994).

45. Sowerby, S. J., Edelwirth, M. & Heckl, W. M. Self-Assembly at the Prebiotic Solid-Liquid Interface : Structures of Self-Assembled Monolayers of Adenine and Guanine Bases Formed on Inorganic Surfaces. **5647,** 5914–5922 (1998).

46. Sowerby, S. J., Cohn, C. a, Heckl, W. M. & Holm, N. G. Differential adsorption of nucleic acid bases: Relevance to the origin of life. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 820–2 (2001).

47. Mamdouh, W., Dong, M., Xu, S., Rauls, E. & Besenbacher, F. Supramolecular nanopatterns self-assembled by adenine-thymine quartets at the liquid/solid interface. *J. Am. Chem. Soc.* **128,** 13305–11 (2006).

48. Mamdouh, W., Kelly, R. E. a, Dong, M., Kantorovich, L. N. & Besenbacher, F. Two-dimensional supramolecular nanopatterns formed by the coadsorption of guanine and uracil at the liquid/solid interface. *J. Am. Chem. Soc.* **130,** 695–702 (2008).

49. Manohar, S., Tang, T. & Jagota, A. Structure of Homopolymer DNA-CNT Hybrids. *J. Phys. Chem. C* **111,** 17835–17845 (2007).

50. Shi, X., Kong, Y., Zhao, Y. & Gao, H. Molecular dynamics simulation of peeling a DNA molecule on substrate. *Acta Mech. Sin.* **21,** 249–256 (2005).

51. Albertorio, F., Hughes, M. E., Golovchenko, J. A. & Branton, D. Base dependent DNA-carbon nanotube interactions: activation enthalpies and assembly-disassembly control. *Nanotechnology* **20,** 395101 (2009).

52. Cathcart, H. *et al.* Ordered DNA wrapping switches on luminescence in single-walled nanotube dispersions. *J. Am. Chem. Soc.* **130,** 12734–44 (2008).

53. Coleman, J. N. Liquid-Phase Exfoliation of Nanotubes and Graphene. *Adv. Funct. Mater.* **19,** 3680–3695 (2009).

54. Neihsial, S., Periyasamy, G., Samanta, P. K. & Pati, S. K. Understanding the binding mechanism of various chiral SWCNTs and ssDNA: a computational study. *J. Phys. Chem. B* **116,** 14754–9 (2012).

55. Martin, W., Zhu, W. & Krilov, G. Simulation study of noncovalent hybridization

of carbon nanotubes by single-stranded DNA in water. *J. Phys. Chem. B* **112,** 16076–89 (2008).

56. Shankar, A., Jagota, A. & Mittal, J. DNA base dimers are stabilized by hydrogen-bonding interactions including non-Watson-Crick pairing near graphite surfaces. *J. Phys. Chem. B* **116,** 12088–94 (2012).

57. Kato, Y., Inoue, A., Niidome, Y. & Nakashima, N. Thermodynamics on soluble carbon nanotubes: how do DNA molecules replace surfactants on carbon nanotubes? *Sci. Rep.* **2,** 733 (2012).

58. Suppan, P. Invited review solvatochromic shifts: The influence of the medium on the energy of electronic states. *J. Photochem. Photobiol. A* **50,** 293 – 330 (1990).

59. Eyring, H. The activated complex and the absolute rate of chemical reactions. *Chem. Rev.* **17,** 65–77 (1935).

60. Atkins, P. & de Paula, J. *Physical Chemistry*. (Oxford University Press, 2010).

61. Zaslavsky, B. Y. *Aqueous Two-Phase Partitioning: Physical Chemistry and Bioanalytical Applications*. (Marcel Dekker Inc., 1994).

62. Albertsson, P.-A. *Partition of Cell Particles and Macromoleucles*. (Wiley-Interscience, 1971).

63. Khripin, C. Y., Fagan, J. a & Zheng, M. Spontaneous partition of carbon nanotubes in polymer-modified aqueous phases. *J. Am. Chem. Soc.* **135,** 6822–5 (2013).

64. Fagan, J. a *et al.* Isolation of Specific Small-Diameter Single-Wall Carbon Nanotube Species via Aqueous Two-Phase Extraction. *Adv. Mater.* (2014).

doi:10.1002/adma.201304873

65. Seeman, N. C. Nanomaterials based on DNA. *Annu. Rev. Biochem.* **79,** 65–87 (2010).

66. Nykypanchuk, D., Maye, M. M., van der Lelie, D. & Gang, O. DNA-guided crystallization of colloidal nanoparticles. *Nature* **451,** 549–52 (2008).

67. Park, S. Y. *et al.* DNA-programmable nanoparticle crystallization. *Nature* **451,** 553–6 (2008).

68. Sassolas, A., Leca-Bouvier, B. D. & Blum, L. J. DNA biosensors and microarrays. *Chem. Rev.* **108,** 109–39 (2008).

69. Prato, M., Kostarelos, K. & Bianco, A. Functionalized carbon nanotubes in drug design and discovery. *Acc. Chem. Res.* **41,** 60–8 (2008).

70. Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids. *Nature* **171,** 737 – 738 (1953).

71. Franklin, R. E. & Gosling, R. G. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* **172,** 156–157 (1953).

72. Wang, A. H. *et al.* Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **282,** 680–686 (1979).

73. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346,** 818–822 (1990).

74. Mao, X., Marky, L. A. & Gmeiner, W. H. NMR structure of the thrombin-binding DNA aptamer stabilized by Sr2+. *J Bio Mol Struc Dynam* **22,** 25–33 (2004).

75. Guckian, K. M. *et al.* Factors Contributing to Aromatic Stacking in Water:

Evaluation in the Context of DNA. *J. Am. Chem. Soc.* **122,** 2213–2222 (2000).

76. Sínanoğlu, O. & Abdulnur, S. HYDROPHOBIC STACKING OF BASES AND THE SOLVENT DENATURATION OF DNA. *Photochem. Photobiol.* **3,** 333–342 (1964).

77. Mitchell, P. R. & Sigel, H. A proton nuclear-magnetic-resonance study of self-stacking in purine and pyrimidine nucleosides and nucleotides. *Eur. J. Biochem.* **88,** 149–54 (1978).

78. Martel, P. Base crystallization and base stacking in water. *Eur. J. Biochem.* **96,** 213–9 (1979).

79. Hunter, C. A. Sequence-dependent DNA structure. The role of base stacking interactions. *Journal of Molecular Biology* **230,** 1025–1025 (1993).

80. Danilov, V. I. & Tolokh, I. S. Nature of the stacking of nucleic acid bases in water: a Monte Carlo simulation. *J Bio Mol Struc Dynam* **2,** 119 (1984).

81. Pohorille, A., Pratt, L. R., Burt, S. K. & MacElroy, R. D. Solution influence on biomolecular equilibria: nucleic acid base associations. *J Bio Mol Struc Dynam* **1,** 1257–80 (1984).

82. Piotr Cieplak, J. Am. Chem. SOCPeter A. Kollman, Cieplak, P. & Kollman, P. A. Calculation of the Free Energy of Association of Nucleic Acid Bases in Vacuo and Water Solution. *J. Am. Chem. Soc.* **110,** 3734–3739 (1988).

83. Nakano, N. I. & Igarashi, S. J. Molecular Interactions of Pyrimidines, Purines, and Some Other Heteroaromatic Compounds in Aqueous Media. *Biochemistry* **9,** 577 – 583 (1970).

84. Dang, L. X. & Kollman, P. a. Molecular dynamics simulations study of the free energy of association of 9-methyladenine and 1-methylthymine bases in water. *J. Am. Chem. Soc.* **112,** 503–507 (1990).

85. Ke, C., Humeniuk, M., S-Gracz, H. & Marszalek, P. E. Direct Measurements of Base Stacking Interactions in DNA by Single-Molecule Atomic-Force Spectroscopy. *Phys. Rev. Lett.* **99,** 1–4 (2007).

86. Spiwok, V., Hobza, P. & Řezáč, J. Free-Energy Simulations of Hydrogen Bonding versus Stacking of Nucleobases on a Graphene Surface. *J. Phys. Chem. C* **115,** 19455–19462 (2011).

87. Linak, M. C., Tourdot, R. & Dorfman, K. D. Moving beyond Watson–Crick models of coarse grained DNA dynamics. *J. Chem. Phys.* **135,** 205102 (2011).

88. Torrie, G. & Valleau, J. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **23,** 187–199 (1977).

89. Sugita, Y., Kitao, A., Okamoto, Y. & Introduction, I. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* **113,** 6042–6051 (2000).

90. Lindahl, E. & Hess, B. GROMACS 3.0 : a package for molecular simulation and trajectory analysis. *J. Mol. Model.* 306–317 (2001). doi:10.1007/s008940100045

91. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4,** 435–447 (2008).

92. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L.

Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926 – 935 (1983).

93. Foloppe, N. & Mackerell, A. D. All-Atom Empirical Force Field for Nucleic Acids : I . Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **21,** 86–104 (2000).

94. MacKerell, A. D. & Banavali, N. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J. Comput. Chem.* **21,** 105–120 (2000).

95. York, D. M., Darden, T. a. & Pedersen, L. G. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. *J. Chem. Phys.* **99,** 8345 – 8348 (1993).

96. Parrinello, M., Rahman, A. & Introduction, I. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52,** 7182–7190 (1981).

97. Bonomi, M. *et al.* PLUMED: A portable plugin for free-energy calculations with molecular dynamics☆. *Comput. Phys. Commun.* **180,** 1961–1972 (2009).

98. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13,** 1011–1021 (1992).

99. Shao, J., Tanner, S. W., Thompson, N. & Cheatham, T. E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theory Comput.* **3,** 2312–2334 (2007).

100. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J.*

Mol. Graph. **14,** 33–8, 27–8 (1996).

101. Friedman, R. & Honig, B. A free energy analysis of nucleic acid base stacking in aqueous solution. *Biophys. J.* **69,** 1528–35 (1995).

102. Morcillo, J., Gallego, E. & Peral, F. A critical study of the application of ultraviolet spectroscopy to the self-association of adenine, adenosine and 5-Amp in aqueous solution. *J. Mol. Struct.* **157,** 353–369 (1987).

103. Zhao, X. Self-Assembly of DNA Segments on Graphene and Carbon Nanotube Arrays in Aqueous Solution: A Molecular Simulation Study. *J. Phys. Chem. C* **115,** 6181–6189 (2011).

104. Lv, W. The adsorption of DNA bases on neutral and charged (8, 8) carbon-nanotubes. *Chem. Phys. Lett.* **514,** 311–316 (2011).

105. Akca, S., Foroughi, A., Frochtzwajg, D. & Postma, H. W. C. Competing interactions in DNA assembly on graphene. *PLoS One* **6,** e18442 (2011).

106. Sugimoto, N., Kierzek, R. & Turner, D. H. Sequence dependence for the energetics of dangling ends and terminal base pairs in ribonucleic acid. *Biochemistry* **26,** 4554–8 (1987).

107. Turner, D. H., Sugimoto, N., Kierzek, R. & Dreikert, S. D. Free Energy Increments for Hydrogen Bonds in Nucleic Acid Base Pairs. *J. Am. Chem. Soc.* **109,** 3783–3785 (1987).

108. Guckian, K. M. *et al.* Experimental Measurement of Aromatic Stacking Affinities in the Context of Duplex DNA. *J. Am. Chem. Soc.* **118,** 8182–8183 (1996).

109. SantaLucia, J., Kierzek, R. & Turner, D. H. Context dependence of hydrogen bond

free energy revealed by substitutions in an RNA hairpin. *Science (80-.).* **256,** 217–9 (1992).

110. Kool, E. Hydrogen bonding, base stacking and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.* **30,** 1–22 (2001).

111. Williamson, J. R., Raghuraman, M. K. & Cech, T. R. Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell* **59,** 871–880 (1989).

112. Zhao, X. & Johnson, J. K. Simulation of adsorption of DNA on carbon nanotubes. *J. Am. Chem. Soc.* **129,** 10438–45 (2007).

113. MacKerell, A. D., Wiorkiewicz-Kuczera, J. & Karplus, M. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* **117,** 11946–11975 (1995).

114. Langley, D. R. Molecular Dynamic Simulations of Environment and Sequence Dependent DNA Conformations: The Development of the BMS Nucleic Acid Force Field and Comparison With Experimental Results. *J Bio Mol Struc Dynam* **16,** 487 (1998).

115. Orozco, M., Perez, A., Noy, A. & Luque, F. J. Theoretical methods for the simulation of nucleic acids. *Chem. Soc. Rev.* **32,** 350 – 364 (2003).

116. Cheatham, T. E. & Kollman, P. a. Insight into the stabilization of A-DNA by specific ion association: spontaneous B-DNA to A-DNA transitions observed in molecular dynamics simulations of d[ACCCGCGGGT]2 in the presence of hexaamminecobalt(III). *Structure* **5,** 1297–311 (1997).

117. Cieplak, P., Cheatham, T. E. & Kollman, P. a. Molecular Dynamics Simulations

Find That 3‘ Phosphoramidate Modified DNA Duplexes Undergo a B to A Transition and Normal DNA Duplexes an A to B Transition. *J. Am. Chem. Soc.* **119,** 6722–6730 (1997).

118. Knotts, T. a, Rathore, N., Schwartz, D. C. & de Pablo, J. J. A coarse grain model for DNA. *J. Chem. Phys.* **126,** 084901 (2007).

119. Ouldridge, T. E., Louis, A. a. & Doye, J. P. K. DNA Nanotweezers Studied with a Coarse-Grained Model of DNA. *Phys. Rev. Lett.* **104,** 1–4 (2010).

120. Morriss-Andrews, A., Rottler, J. & Plotkin, S. S. A systematically coarse-grained model for DNA and its predictions for persistence length, stacking, twist, and chirality. *J. Chem. Phys.* **132,** 035105 (2010).

121. Kenward, M. & Dorfman, K. D. Brownian dynamics simulations of single-stranded DNA hairpins. *J. Chem. Phys.* **130,** 095101 (2009).

122. Zhang, F. & Collins, M. Model simulations of DNA dynamics. *Phys. Rev. E* **7,** 915–4224 (1995).

123. Drukker, K., Wu, G. & Schatz, G. C. Model simulations of DNA denaturation dynamics. *J. Chem. Phys.* **114,** 579 – 590 (2001).

124. Samsonidze, G. G. *et al.* Interband optical transitions in left- and right-handed single-wall carbon nanotubes. *Phys. Rev. B* **69,** 205402 (2004).

125. Tasaki, S. & Yamabe, T. π-band contribution to the optical properties of carbon nanotubes: Effects of chirality. *Phys. Rev. B* **57,** 9301–9318 (1998).

126. Vardanega, D., Picaud, F. & Girardet, C. Chiral response of single walled carbon nanotube based sensors to adsorption of amino acids: a theoretical model. *J. Chem.*

Phys. **127,** 194702 (2007).

127. Strano, M. S. Carbon nanotubes: sorting out left from right. *Nat. Nanotechnol.* **2,** 340–341 (2007).

128. Ivchenko, E. L. & Spivak, B. Chirality effects in carbon nanotubes. *Phys. Rev. B* **66,** 23 (2002).

129. Zheng, M. *et al.* DNA-assisted dispersion and separation of carbon nanotubes. *Nat. Mater.* **2,** 338–342 (2003).

130. Ghosh, S., Bachilo, S. M. & Weisman, R. B. Advanced sorting of single-walled carbon nanotubes by nonlinear density-gradient ultracentrifugation. (2010). doi:10.1038/NNANO.2010.68

131. Akazaki, K., Toshimitsu, F., Ozawa, H., Fujigaya, T. & Nakashima, N. Recognition and One-Pot Extraction of Right-and Left-Handed Semiconducting Single-Walled Carbon Nanotube Enantiomers Using Fluorene-Binaphthol Chiral Copolymers. doi:10.1021/ja304244g

132. Roxbury, D. Sequence Dependent Interactions Between DNA and Single-Walled Carbon Nanotubes. (Lehigh University, 2012). at <http://gradworks.umi.com/35/10/3510120.html>

133. Dresselhaus, M. S., Dresselhaus, G. & Saito, R. Physics of carbon nanotubes. *Carbon N. Y.* **33,** 883–891 (1995).

134. Murphy, M. C., Rasnik, I., Cheng, W., Lohman, T. M. & Ha, T. Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys. J.* **86,** 2530–7 (2004).

135. Kypr, J., Kejnovska, I., Renciuk, D. & Vorlickova, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37,** 1713–1725 (2009).

136. Moroz, J. D. & Nelson, P. Torsional directed walks, entropic elasticity, and DNA twist stiffness. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 14418–14422 (1997).

137. Salencon, J. *Handbook of continuum mechanics*. (Springer-Verlag).

138. Zhang, J. *et al.* Single Molecule Detection of Nitric Oxide Enabled by d(AT)(15) DNA Adsorbed to Near Infrared Fluorescent Single-Walled Carbon Nanotubes. *J. Am. Chem. Soc.* 567–581 (2010). doi:10.1021/ja1084942

139. Berendsen, H. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91,** 43–56 (1995).

140. MacKerell, A. D., Banavali, N. & Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **56,** 257–65 (2001).

Vita

Akshaya Shankar was born in Mumbai, India on July 20 1988 to Ganesan Sankar and Jayshree Shankar. She received her Bachelors in Chemical Engineering from Institute of Chemical Technology (formerly UDCT), Mumbai, India in May 2010. She did an internship at National Peroxode Ltd in Kalyan, India in Summer 2008. Since August 2010, she has been attending Lehigh University to earn her PhD in Chemical and Biomolecular Engineering and working with her advisor Prof. Anand Jagota.