COMPUTATIONAL TECHNIQUES FOR SIMULATING THE INTERACTIONS BETWEEN PEPTIDES AND CARBON NANOTUBES

CHENG YUAN

NATIONAL UNIVERSITY OF SINGAPORE

2007

COMPUTATIONAL TECHNIQUES FOR SIMULATING THE INTERACTIONS BETWEEN PEPTIDES AND CARBON NANOTUBES

CHENG YUAN

(B. S., FUDAN UNIVERSITY, CHINA)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MACHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2007

Acknowledgements

I would like to express my deepest gratitude and appreciation to my supervisor, Professor **Liu Gui-Rong** for his dedicated support, invaluable guidance, and continuous encouragement in the duration of the study. His influence on me is far beyond this thesis and will benefit me in my future research work. I am much grateful to my co-supervisor, Dr. **Lu Chun**, for his inspirational help and valuable guidance in my research wrok. I would also like to thank Mr. **Li Zi-rui** and Dr. **Mi Dong** for their helpful discussion, suggestion, recommendations and valuable perspectives.

To my friends and colleagues in the ACES research center, Ms **Zhang Ying-Yan**, Dr. **Zhang Gui-Yong**, Dr. **Dai Ke-Yang**, Dr. **Li Wei**, Dr. **Deng Bin**, Mr. **Zhou Cheng-En**, Dr. **Zhao Xin**, Mr. **Kee Buck Tong Bernard**, Mr. **Zhang Jian**, Mr. **Song Cheng-Xiang**, Mr. **Khin Zaw**, Mr. **Luo Rongmo**, I would like to thank them for their friendship and help.

To my family, I appreciate their love, encouragement andsupport. Especially to my husband, Mr. **Li Ang**, it is impossible for me to finish this work without his support and encouragement.

I am grateful to the National University of Singapore for granting me the research scholarship which makes my study in NUS possible. Many thanks are conveyed to Center for Advanced Computations in Engineering Science (ACES) and Department of Mechanical Engineering, for their material support to every aspect of this work.

Table of Contents

Summary

The exceptional properties of carbon nanotubes (CNTs) facilitate their wide application in a number of fields in physics, chemistry, and biomedicine. Although the marvellous properties of CNTs have triggered great interest of researchers to explore potential applications of CNTs, the mechanism of CNTs interacting with biomolecules still remains unclear.

This thesis focuses on investigation of interaction mechanism between peptides and CNTs based on different levels of molecular description. Computational strategies adopting either all-atom model or coarse-grained model are implemented. The major works reported in this thesis are listed as follows

1) An all-atom model is developed to study self-insertion behaviors of different peptides into SWCNTs in explicit water environment using molecular dynamics (MD) simulation. The conformational changes of the peptide and energetics of the interaction are traced. Variations in affinity of different peptides for single-walled carbon nanotubes (SWCNTs) are also observed.

2) The Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) method is extended to evaluate the free energy of peptides interacting with CNTs. The relative binding affinities are compared with the experimental results to validate the

model. The physical mechanism involved in this process is then studied in detail. Other effects that may influence peptide-CNT interaction are also investigated.

3) In order to obtain a general view of different binding affinity of hydrophobic and hydrophilic amino acids for the CNTs, binding free energy between each amino acid and the same CNT is estimated individually based on the all-atom model. The relative binding affinities of amino acids from the hydrophobic and hydrophilic groups are compared.

4) A coarse-grained hydrophobic-polar (HP) lattice model is developed performing MC simulation to observe the macroscopic properties of the adsorption of peptides onto CNT surfaces. The preliminary energy parameters are developed according to experimental observations and numerical results from the all-atom model. The thermodynamic quantities and conformational characteristics of peptides are also clarified.

Through these studies I am not only able to explore the detailed conformational properties and energetics of peptides interacting with CNTs, but also the peptide-CNT interaction mechanism from both microscopic and macroscopic views. The results obtained through this study provide valuable information on the potential applications of CNTs in the field of drug delivery, drug design and protein control.

Nomenclature

a	particle's acceleration
A	Helmholtz free energy; accessible surface area
A_{ij}	the area of sphere i buried inside sphere j
$\langle A \rangle$	ensemble average value of property A
c_1	the constant representing the initial velocity of the particle
c_2	the constant representing the initial position of the particle
d	the center of mass distance between the peptide and the nanotube at instant simulation time
d_0	the initial center of mass distance between the peptide and the nanotube
E_{des}	desired energy
E_{MM}	gas-phased molecular mechanics energy
E_{internal}	the internal energy
E_{ele}	the electrostatic energy
E_{vdw}	the van der Waals interaction energy

E_i	potential energy of HP lattice model at certain state
E_{ele_total}	the energy sum of E_{ele} and G_{pol}
E_{MM}	gas-phased molecular mechanics energy
$\mathbf{f}_i^{tot}$	the sum of inter-molecular forces and external forces
$\mathbf{F}$	the force acting on the particle
G_{cnt}^s, $G_{peptide}^s$, $G_{complex}^s$	free energy of the peptide, the carbon nanotube, and the peptide-nanotube complex solvated in water, respectively
ΔG	interaction free energy
G_{sol}	the solvation free energy
G_{nonpol}	polar energy
G_{pol}	nonpolar energy
h	Planck's constant
$H(\mathbf{\Gamma})$	Hamiltonian
k_θ	force constant for angle bending potential
k_r	force constant for bond stretching potential
k_B	Boltzmann constant
K	kinetic energy
K_{MU}	equilibrium constant
l	bond length

l_0	the reference bond length
m	particle's mass
N_k	the total number of contact pairs in a certain HP lattice structure
N	number of particles
p_{mn}	the trial move
$\mathbf{p}$	momentum
P	pressure
q_i	atomic partial charges
$Q_{NVE}, Q_{NVT}, Q_{\mu VT}, Q_{NPT}$	the partition function for different ensembles
r_{ij}	distance between atoms
$\mathbf{r}$	displacement of the particle
R_i	atomic radius
S_i	the surface area of the isolated sphere
S	entropy
t	time
T^*	the dimensionless temperature
T	temperature
$U(\mathbf{r})$	potential energy
$\mathbf{v}$	particle's velocity
V	box of volume
V_n	the rotational barrier height

w_i	occurrence of each amino acid in proteins
W	internal virial
Z	partition function of the system
α_i	effective Born radius
α_{mn}	stochastic matrix element between state m and state n
β	inverse temperature $(1/k_BT)$
$\mathbf{\Gamma}$	the set of particle positions and momenta
$\gamma_S(A_i, A_j)$	interaction potential energy between the two amino acids residues for HP lattice model
ε_0	solvent dielectric constant
ε	well depth; intermolecular potential energy parameter
σ	the collision diameter
θ	bond angle
θ_0	reference bond angle
μ	the chemical potential of the simulated system
κ	Debye-Huckel screening parameter
η_i	intrinsic radius of atom i
ρ_m	the probability that the system is in state m
ξ	random number (usually in range 0 to 1)
Ψ_{ens}	the characteristic thermodynamic function

ω	torsion angle
$\Omega(E_i)$	the density of states with energy E_i
π_{mn}	the probability of moving from m to state n

List of tables

List of figures

Chapter1

Introduction

1.1 Background information for Carbon nanotubes (CNTs) and peptides

1.1.1 General overview of CNTs

Carbon nanotubes (CNTs) are hollow cylindrical tubes consisting of webs of carbon atoms. Since their discovery in 1991 (Iijima, 1991), CNTs have stimulated ever-broader research activities in science and engineering devoted to production and application of various CNTs. The outstanding properties of CNTs such as high mechanical strength and remarkable electronic structure make CNTs special in applications in a vast variety of fields. A number of excellent reviews on general properties of CNTs are available (Harris et al., 1999; Dresselhaus et al., 1996; Dresselhaus et al., 2001), here I make this effort with emphasis on the applications of CNTs in biomedical areas.

1.1.1.1 Molecular structure of CNTs

CNTs are normally classified into two categories: single-walled carbon nanotubes (SWCNTs) and multi-walled carbon nanotubes (MWCNTs). SWCNTs are made from a graphite sheet rolled into a cylinder, while MWCNTs are composed of multiple concentric graphite cylinders, as illustrated in Figure 1.1. Compared with

MWCNTs, SWCNTs are more expensive and difficult to manufacture and clean, but they have been of great interest to researchers owing to their specific electronic, mechanical, and gas adsorption properties (Ebbesen et al., 1997).

CNTs can be considered as rolled-up graphite sheets. When carbon atoms geometrically combine together to form graphite, sp^2 hybridization occurs (Brown et al., 1999). Different types of CNTs can be characterized by a linear combination of base vectors $\mathbf{a}$ and $\mathbf{b}$ of the hexagon, or $\mathbf{r} = n\mathbf{a} + m\mathbf{b}$, where n and m are integers of the vector equation (Thostensona et al., 2001; Qian et al., 2002) as shown in Figure 1.2. The values of n and m uniquely determine the chirality, or twist style of the nanotube. Three major categories of CNTs can be defined based on the value of n and m. If $n = m$, the CNT is armchair, if $n = 0$ or $m = 0$, the CNT is classified as zigzag. When $n \neq m$, the CNT is generally chiral. The chirality affects the conductance, the density, the lattice structure, and therefore affects other properties of the nanotube. A SWCNT is considered metallic if the value $n - m$ is divisible by three. Otherwise, the nanotube is semiconducting. Consequently, when tubes are formed with random values of n and m, it is expected that two-thirds of nanotubes would be semi-conducting, while the other third would be metallic, which happens to be the case. Representative configurations of the three types of CNTs are illustrated in Figure 1.3.

Given the chiral vector (n, m), the diameter d and the chiral angle θ of a carbon nanotube can be determined as

$$\begin{cases} d = (n^2 + m^2 + nm)^{1/2} * 0.0783 \text{ nm} \\ \theta = \sin^{-1}\left[\dfrac{\sqrt{3}m}{2(n^2 + nm + m^2)}\right] \end{cases} \quad (1.1)$$

1.1.1.2 Properties of CNTs and their applications

Many efforts have been made in order to investigate the mechanical properties of CNTs. For example, they were found to be bent mechanically by mechanical milling or embedding in a polymeric resin (Ajayan et al., 1994; Iijima et al., 1996; Chopra et al., 1995; Ruoff et al., 1995). This flexibility property was also predicted through theoretical calculations (Overney et al, 1993; Robertson et al., 1992; Tersoff, 1992). Treacy et al. (1996) first investigated the elastic modulus of isolated multi-walled nanotubes by measuring the amplitude of their intrinsic thermal vibration through the transmission electron microscope (TEM). Direct measurement of the stiffness and strength of individual, structurally isolated multi-wall CNTs has also been performed with an atomic-force microscope (AFM) (Wong et al., 1997). High Young's modulus of CNTs was observed through these measurements. This high Young's modulus implies that CNTs are very strong material. On the other hand, the mechanical properties of composite materials containing CNTs are expected to be greatly enhanced, although those materials will not be as robust as individual nanotubes.

CNTs also possess unique electrical properties. These properties are sensitive to the orientation of the hexagonal graphite lattice because it determines the density of electron states at the Fermi level (Gao et al., 2004). Hamada et al. (1992) found theoretically that all the armchair nanotubes are electronic conductors, while zig-zag nanotubes are

semiconductors except for those $n-m$ is divisible by three. For CNTs whose radius is greater than 1nm, this simple model works remarkably well. In those cases that the radius of CNTs are smaller, however, the atomic arrangement of CNTs is highly curved and this simple rule is no longer valid owing to strong mixing between the in-plane and out-of-plane electronic orbitals. Therefore first-principles calculations are needed to adequately describe the electronic properties of very small diameter CNT systems (Blase et al., 1994). Furthermore, SWCNTs tend to self-assemble into bundles. The internal interactions of the tube may introduce small pseudogaps in bundles of nominally metallic nanotubes (Delaney et al., 1998; Kwon and Tomanek, 1998).

The exceptional mechanical and electrical properties of CNTs facilitate their wide application in a number of fields in physics, chemistry, and material science including biosensors (Balavoine et al., 1999), atomic force microscopy (AFM) (Jarvis et al., 2000; Li et al., 1999) and fuel storage (Lee et al., 2000; Wang and Johnson, 1999). Their outstanding mechanical properties suggest that they could act as unique force transducers to the molecular world. The inversed electromechanical effect of CNTs enables the application of CNTs in nanomechanical applications, such as tweezers (Poncharal et al., 1999) and actuators (Baughman et al., 1999). The coulomb blockade was detected in transport measurements (Tans et al., 1997; Bockrath et al., 1998), which implies that the nanotubes are suitable building blocks of single-electron transistors. Recently some functional structures based on CNTs have also been fabricated, including nanotube transistor (Tans et al., 1998), nano-diode (Antonov and Johnson, 1999), and color display (Choi et al., 1999). Nanotubes can also form novel nanostructures which

may have potential application in nanoelectronics and nanophotonics, e.g., molecular junctions by jointing CNTs (Andriotis et al., 2000; Terrones et al., 2002; Srivastava et al., 2003), organized assembly of CNTs (Wei et al., 2002), and nano-films (Shimoda et al., 2002) composed of aligned uniform nanotubes, are to be manufactured in industry.

CNTs also show great potential for biomedical applications owing to their high strength and biocompatibility. For example, recent demonstration of CNT artificial muscle implied a dramatic increase in work density output and force generation over known technologies, along with the ability to operate at low voltage (Baughman et al., 1999). CNTs can also be utilized in gene and drug delivery. For example, they could be implanted at the sites where a drug is needed without trauma, and slowly release a drug effectively over a period of time (Harutyunyan et al., 2002). It is also promising in applying CNTs in the area of cellular experiments, where CNTs can be utilized as nanopipettes for the distribution of extremely small volumes of liquid or gas into living cells or onto surfaces. It is also conceivable that they could serve as a medium for implantation of diagnostic devices.

1.1.2 Proteins and peptides

Proteins are building blocks of a living cell, and they participate in essentially all cellular processes. One of the major functions of proteins is enzymatic catalysis of chemical conversions inside and around the cell. In addition, regulatory proteins control gene expression, and receptor proteins (which locate in the lipid membrane) accept intercellular signals that are often transmitted by hormones, which are proteins

as well. Structural proteins form microfilaments and microtubules, as well as fibrils, hair, silk and other protective coverings. These proteins reinforce membranes and maintain the structure of cells and tissues. Transfer proteins transfer other molecules. Some proteins provide the human body with entire bioenergetics, for example, light absorption, respiration, ATP production, *etc*.

Proteins are polymers built of amino acids arranged in a linear chain and joined together by peptide bonds between the carboxyl and amino groups of adjacent amino acid residues. An $\alpha-$ amino acid consists of a central carbon atom, called the α carbon, lined to an amino group, a carboxylic acid group, a hydrogen atom, and a distinctive R group. The R group is often referred to as the side chain. There are twenty kinds of amino acids, classified according to their side chains. The detailed structures for the individual amino acids can be found in references (e.g., Berg et al., 2002). The twenty types of side chains vary in size, shape, charge, hydrogen-bonding capacity, hydrophobic character, and chemical reactivity. All the proteins in all species are constructed from the same set of twenty amino acids. Owing to the diversity and versatility of these twenty building blocks, proteins are able to perform a wide range of functions.

Amino acids are often designated by a three-letter abbreviation or a one-letter symbol (Table 1.1). Their essential properties such as the occurrence in proteins and the hydrophobicity scale of each amino acid are also listed. Hydrophilic molecules are in favor of interacting with water while hydrophobic ones tend to be nonpolar and thus prefer other neutral molecules and nonpolar solvent. The value of hydrophobicity is

listed according to K-D method (Kyte and Doolittle, 1982), in which each amino acid has been assigned a value reflecting its relative hydrophilicity and hydrophobicity. A positive hydrophobicity value indicates that the amino acid is hydrophobic, and the negative value implies the hydrophilic property of the amino acid. The higher the hydrophobicity values, the more hydrophobic the amino acid is.

Protein structures can be described at four levels. The primary structure refers to the amino acid sequence. A series of amino acids joined by peptide bonds form a polypeptide chain, and each amino acid unit in a polypeptide is called a residue. The polymer chain consists of a chemically regular backbone called main chain and various side chains (R_1, R_2, …, R_M). The number M of residues in one protein could range from a few dozens to many thousands. This number is gene-encoded, and so are the positions of these amino acids in the protein chain. Most natural polypeptide chains contain between 50 and 2000 animo acid residues and are usually referred to as proteins. Polypeptides made of small number of amino acids are called oligopeptides or simply peptides.

Secondary structure refers to the conformation of the local regions of the polypeptide chain. Polypeptide chains can fold into regular structures such as the alpha helix, the beta sheet, and turns and loops. Although the turn or loop structures are not periodic, they are well defined and contribute together with alpha helices and beta sheets to form the final protein structure.

Tertiary structure describes the overall folding of the polypeptide chain. Finally, quaternary structure refers to the specific association of multiple polypeptide chains to

form multisubunit complexes. A knowledge of the 3D structure of a protein is essential to understanding its function.

1.2 Functionalization of CNTs with Biomolecules

1.2.1 Experimental approaches

Although there is growing interest in exploring the application of CNTs in novel fields, CNTs are extremely hydrophobic and form insoluble aggregates in solvent, which makes them difficult to assemble into applicable structures. The solubilization of SWCNTs has been a research goal for the past few years, and study on solution-phase handling would be very useful for many of the CNT applications. Ausman et al. investigated the room-temperature solubility of SWCNTs in a variety of solvents (Ausman et al., 2000). It was found that a class of non-hydrogen-bonding Lewis bases could lead to better solubility, but this was only a possible way that can provide better solvents capable of solvating pristine tubes. The problem that SWCNTs are insoluble in all solvents is still difficult to overcome.

In order to make CNTs soluble, as well as to facilitate the possible applications of CNTs in various areas, many experimental efforts have been made, either through covalent or noncovalent interactions between biomaterials and CNTs to explore the biological applications of CNTs.

For example, nanotubes could be solubilized well by functionalizing the end-caps with long aliphatic amines (Chen et al., 1998). Furthermore, it has been

reported that SWCNTs have been solubilized by functionalizing their sidewalls with fluorine (Mickelson et al., 1999) and with alkanesn (Boul et al., 1999).

In addition, since the electronic properties of CNTs are sensitive to surface charge transfer and changes in the surrounding electrostatic environment, it is expected that functionalization of CNTs by attaching various functional groups or molecules to its outer surface could be applied to controllably modify the intrinsic chemical and physical properties for specific chemical and biomedical applications (Zhao et al., 2002; Hirsch et al., 2002; Wong et al., 1998; Erlanger et al., 2001; Azamian et al., 2002; Nguyen et al., 2002; Williams et al., 2002; Pantarotto et al., 2003). Among them Wong et al. reported the modification of MWCNTs through amide bond. The amide bond formed between amine and carboxy functional groups bonded to the open ends of MWCNTs. The modified complex could be applied as AFM tips, so that the binding force between single protein-ligand pairs can be measured.

However, for those applications requiring high conductivity properties of CNTs, the modification through noncovalent bond is more attractive. Another strategy that scientists are eager to explore is to attach organic molecules to these tubular nanostructures in a noncovalent way in order to preserve the nanotubes' π networks- and thus their electronic characteristics. Scientists can manipulate nanotubes into ordered array without destroying their instinct structure through noncovalent modification approaches.

Chen et al. (2001) explored π-stacking interactions between the CNT and a molecule containing a planar pyreny group through noncovalent contact. The pyreny group irreversibly absorbed to the surface of a SWCNT driven by π-stacking forces. The molecule's tail was tipped with a succinimidyl ester group. While an amine group attacked the ester function, the ester group could be substituted and an amide bond forms. This strategy may be very useful not only for immobilizing proteins or DNA, but also for solubilizing CNTs. There are also some studies referring to the noncovalent interactions between peptides and CNTs. Diechmann et al. (2003) designed an amphiphilic α-helical peptide not only to coat and solubilize CNTs, but also to control the assembly of the peptide-coated nanotubes into macromolecular structures through peptide-peptide interactions. The phage display method was used to identify peptides with selective affinity for CNTs (Wang et al., 2003). It was found that CNTs have strong affinity for peptide sequences rich in His and Trp. Several of the binding peptides had a hydrophobic structure of symmetric detergents.

In addition to binding and attaching of functional groups to the outer surface of the CNTs, the hollow interior of CNTs can also be filled with smaller nanoparticles and molecules. For example, gas molecules, C_{60} and metallofullerences could be encapsulated into the inner space of CNTs to functionalize them (Gogotsi et al., 2001; Hirahara et al., 2000; Smith et al., 1998). Ito et al. reported observation of DNA transport through a SWCNT channel by fluorescence microscopy (Ito et al., 2003).

1.2.2 Simulation approaches

A computational simulation allows researchers to gain insight into the processing of materials and propose new directions for design without expensive and time consuming experimentation in a laboratory.

There have been only a few studies exploring the biomolecules-CNT interactions through computational methods. Hummer et al. showed that SWCNTs could act as a hydrophobic channel for conduction of water molecules (Hummer et al., 2001) molecular dynamics (MD) simulation. Gao and his colleagues simulated spontaneous insertion of DNA oligonucleotides into SWCNTs in water solvent environment (Gao et al., 2003). More recently the electrophoretic transport of single-stranded RNA molecules through SWCNT membranes was investigated using MD simulations (Yeh and Hummer, 2004). The numerical simulation results revealed that the translocation kinetics of RNA through the nanotube membranes was sequence-dependent. These works inspired us to further explore the problem using computational approaches.

1.3 Molecular simulation models based on different levels of description

While the previous works have provided us with hints on possible applications of CNTs, further research is needed to clarify the mechanism of interactions between biomolecules and CNTs. Therefore a systematic study based on different levels of description for modeling of peptide-CNT interaction is particularly essential. The all-atom simulations allow us to follow the delicate interplay of various chemical interactions leading to the formation of native or the equilibrium states with microscopic detail. On the other hand, simplified coarse-grained models are very

useful and efficient to gain insights into the general thermodynamic and kinetic features of the folding process.

In this work different computational strategies based on the used of either all-atom or coarse-grained descriptions are discussed. These levels of description for a given system order themselves in terms of the amount of information captured by the relevant variables. Each level of description is characterized by a set of relevant variables that specify the state of the system at that level. Less detailed levels (coarser levels) have a smaller number of variables and capture less information than the all-atom level.

1.3.1 The atomic model

In atomic-level models, all the atoms can be explicitly simulated. Within a classical perspective, the appropriate tool to capture the detailed dynamical and thermodynamical aspects is constituted by simulations based on all-atom potentials. Although the time scale that could be handled by this model is limited by its large computational cost, it has proven useful in several important contexts. Examples involve tracing the detailed characterization of complete pathways, exploring the interactions between ligands and receptors, and design of possible drugs capable of interacting with specific mutants.

Molecular simulations based on both MD and Monte Carlo (MC) approaches using all-atom force fields are frequently used. Among them MD simulations are particularly appealing in this field, since MD simulation method represents the only

computational method that can provide a time-dependent analysis of a system in molecular biology. Therefore, a complete description of the folding mechanism of a protein can be gained.

In MD simulations, successive configurations of the system are generated by integrating Newton's law of motion. The result is a trajectory that specifies how the positions and velocities of the particles in the system change with time. There are three essential components for a MD program: a model describing the interactions between system constituents (electrons, atoms/molecules, etc.); an integrator that propagates particle positions and velocities from simulation time t to δt (The equations of motions are usually integrated using a finite difference method)； and a statistical ensemble where thermodynamic quantities such as temperature, pressure, or the number of particles are controlled.

At the most basic level of model building, quantum mechanics (QM)-based ab initio MD method evaluates the interatomic forces from the electronic structure calculations during the process of simulation. The typical length and timescales are of the order of angstroms (Å) and picoseconds. Nevertheless, as the advent of more powerful, massively parallel computers, coupled with spectacular advances in theoretical framework of method (Carloni et al., 2002), enables the modeling and simulations of novel materials based on electronic level. For example, the electronic structure of DNA molecules (de Pablo et al., 2000; Gervasio et al., 2002) and reaction mechanism of enzymes (Carloni et al., 2002) were clarified. Classical MD models interatomic interactions via empirical molecular force fields (Stutman, 2002), where

the electronic distributions are estimated either by putting fixed partial charges on interaction sites or by adding an approximate model for polarization effects. The accessible length and time scale are in order of tens of nanometers and nanoseconds.

Classical MD simulations are applied in a wide range of applications. They are often used to study the thermodynamic properties of gas, liquid, solid, phase transitions, as well as motions of bio-molecular systems (Kaplus, 1990; van Gunsteren, 1994), including structural dynamics of biomolecules, protein/DNA interaction, and the effect of solvent. Owing to the large area of applicability, simulation packages for MD were developed by a number of research groups, such as Amber (Cornell et al., 1995), Charmm (Brooks et al., 1983), NAMD (Kale et al., 1999), and Gromacs (Berendsen et al., 1995).

Estimating free energy through MD simulation method has been a great challenge for scientists. Free energy is the most important general concept in physical chemistry. The free energies of molecular systems describe their tendencies to associate and react. Thus, being able to predict this quantity using molecular theory would be essential for us to understand the mechanism of physical and chemical phenomenon.

Among the interactions between molecules, the ability to predict the strength of noncovalent binding between molecules has been a longstanding goal in computational chemistry. Gaining into the energetics of binding is a problem that is extremely difficult to solve using conventional computational free energy techniques. During the past 10-15 years, many efforts have been made by researchers to address the question

of ligand-receptor binding (Kollman 1993; Lamb and Jorgensen 1997; Bohm and Stahl 1999). The most commonly-used methods include the free energy perturbation (FEP) theory, the linear interaction energy (LIE) and the Generalized Born Surface Area (GBSA) methods.

Based on an all-atom model, FEP theory (Beveridge and DiCapua 1989; Jorgensen 1989; Kollman 1993) combined with conformational sampling by MD or MC simulations provide a rigorous way of calculating free energies upon modifactions of a ligand or a receptor. Most FEP calculations take advantage of a thermodynamic perturbation cycle and modifications of the ligand or receptor are achieved through nonphysical transformation process. As a result of sampling and convergence， problems related to large perturbation FEP calculations are in most cases limited to the evaluation of relative binding free energies for compounds of similar chemical structure. Even calculations of relative binding free energy may pose a major problem if a lot of modifications are required to bring the system from one state to another.

The LIE method, first proposed by Åqvist et al. (1994), was based on the electrostatic linear response approximation and an empirical estimate of the nonpolar binding contribution. This method is an alternative to FEP method. In contrast to FEP calculations, the LIE method requires only simulations of the corners of the thermodynamic perturbation cycle. However, explicit solvent is used and relatively long computational time is required by these approaches.

The analytic Generalized Born (GB) model efficiently describes electrostatics of molecules in water environment. It treats the solvent implicitly as continuum with the

dielectric properties of water, and includes the charge screening effects of salt. The nonpolar free energy is estimated proportional the surface areas (SA) to represent the cavity and van der Waals contributions to solvation. The surface area is commonly calculated using the Linear Combinations of Pairwise Overlaps (LCPO) model (Weister et al., 1999).

There are several advantages for using GB models. For example, the computational cost of using the GB model in MD simulation is generally significantly smaller than the cost of simulations with explicit water. The model describes instantaneous solvent dielectric response which eliminates the need for length equilibration of water necessary in explicit water simulations. The GB model assumes that the systems are solvated in an infinite volume of solvent, therefore avoiding possible artifacts of replica interactions in periodic system treatments to speed-up explicit water calculations. Since the solvent degrees of freedom are taken into account implicitly, estimating energies of solvated structures is much more straightforward than with explicit water models.

The GBSA continuum solvent model is generally combined with the molecular mechanics (MM) of the molecules to describe solvation free energies. Calculations of binding free energy using MM-GBSA method only takes into account the physical states at both end points of binding reaction and therefore there is no need to devote computer time on intermediate states. The method has been applied to compare relative stabilities of different conformations of nucleic acids (Srinivasan et al., 1998), to identify correctly folded proteins (Lee et al., 2001) , and to estimate binding

affinities of small molecules or ligands binding to proteins (Kuhn and Kollman, 2000; Lee and Kollman, 2000; Wang et al., 2001). The method can also be utilized to predict the effects of amino acid mutations on binding affinities (Wang and Kollman, 2000), and could be extended to study the interaction free energies between CNTs and biomolecules.

1.3.2 The coarse-grained hydrophobic-polar (HP) lattice model

While computational simulation is a powerful tool which permits us to observe, examine and manipulate the smallest detail in many ways beyond the access of experiment, computer equipment is also a limited resource. Although MD simulation of the all-atom models can provide us with great insight into the peptide-CNT interaction mechanism, it is currently only suitable for simulating short peptides in a relatively short time scale (typically nanoseconds). Such an approach is not applicable to the study of the whole protein folding process which is typically in order of microseconds to seconds. Therefore it is also necessary to develop simulation models which are able to capture the essential features of the materials with the minimum of computational units and computational time. The atomic model and the coarse-grained model can serve as complements for each other.

Molecular systems can be modeled at different levels of spatial resolution. The process of representing a system with fewer degrees of freedom than those actually present in the system is called coarse-graining. By coarse-graining I am not only reducing computational units in the system, but also on the possibility of exploring a

much larger time span. The validity of the coarse-gained models is inferred by confronting its predictions with experiments. Different from the classical atomic level representations of biomolecules, these coarse-grained models and their correspondingly simplified force fields consist of beads representing groups of atoms, monomers, or even several monomer units. The beads interact with each other through effective interaction functions that take into account the response of the omitted degrees of freedom effectively in an average way. They have proven to supply accurate thermodynamic descriptions of partitioning in homogeneous systems (Baron et al., 2007).

In recent years there has been an emerging interest in the development of simple coarse-grained models for a variety of polymers (Baschnagel et al., 2000; MQller-Plathe, 2002; Kremer, 2003), lipids and surfactants (Marrink et al., 2004; Smit et al., 1990; Goetz and Lipowsky, 1998), and proteins (Tozzini, 2005; Shih et al., 2006; Bond et al., 2006). These studies focused on computer simulations of longer time and larger length scales at the expense of lower resolution of structural and dynamical properties.

Among the many coarse-grained models, the HP lattice model is one of the most widely adopted one and has been shown successful in clarifying protein folding mechanism. First proposed by Dill and Lau (Lau and Dill, 1989), the HP model is based on the assumption that the hydrophobic interaction is the dominant force in protein folding. Each residue in the protein sequence is represented by either of the two types, hydrophobic (H) or polar (P). A protein is modeled as a specific sequence

of H and P monomers. The structure of a protein is modeled as linked beads placed on the nodes of a Cartesian coordinate grid. Each protein chain contains N residues, connected through $N-1$ links. All connecting vectors move parallel to either the x- or the y-axis with a self-avoiding configuration. The lattice spaces which are not occupied by the amino acid residues are assumed to be solvent units. Interaction energies are calculated between monomers that are adjacent in space but not covalently lined.

The HP lattice model is commonly implemented using MC algorithm, which is a stochastic algorithm that depends on probabilities. Despite its simplicity, the HP lattice model using MC simulation method is capable of capturing most essential mechanism of protein folding such as hydrophobic effect, multi-stage folding kinetics and molten globule states (Lau and Dill, 1989). Most importantly, the model captures the main physics of protein folding—the hydrophobic interactions, conformational freedom of the chain, and the steric restrictions imposed by excluded volume (Miller and Dill, 2006). Furthermore, the full conformational spaces can be enumerated exhaustively and insights into the nature of free energy landscapes can be obtained. For short chains, all the possible conformations can be enumerated and their energies are evaluated to find the global energetic minimum for each HP sequence. Therefore the model has been successfully used to explore kinetics and thermodynamics of protein folding in bulk solvent (Lau and Dill, 1989; Miller and Dill, 2006; Camacho and Thirumalai, 1993; Chan and Dill; 1993**;** Li et al., 1996; Kumar et al., 2005; Chikenji et al., 2006) or adsorption of protein to various surfaces (Zhdanov, and Kasemo, 1998;

Anderson et al., 2000; Castells et al., 2002; Liu and Haynes, 2004). For example, Zhdanov and Kasemo investigated the relative rates of denaturation of HP chains in the presence of a solid-liquid interface. Their results revealed that at relatively high temperature, unfolding of the chains follows a first-order rate equation. This phenomenon is similar to what was observed without a surface present. At lower temperatures, however, the denaturation pathway at the surface differed in metastable states were formed. Castells et al. used a MC algorithm to demonstrate that mesoscopic protein chains attach to a surface in an unfolded state, and that the degree of unfolding is correlated to the degree of attraction of the residue to the surface.

In this thesis the HP lattice model using MC simulation method is extended to study peptide-CNT interactions.

1.4 Objectives and significance of this study

In spite of the fact that the marvellous properties of CNTs have triggered great interest of researchers to explore wide applications of CNTs, the mechanism of CNTs interacting with biomolecules still remains unclear. Furthermore, there have been only a few studies focusing on interactions between peptides and CNTs, and therefore a systematically research in this area based on computational modelling is particularly novel and will significantly benefit the application of CNTs. Despite the challenge, this thesis focuses on developing both all-atom models and coarse-grained models to investigate the peptide-CNT interaction mechanism. Major works reported in this thesis are as follows.

1) An all-atom model is developed to study self-insertion behaviors of different peptides into SWCNTs in explicit water environment using MD simulation. The conformational change of the peptide and energetics of the interaction are traced. Variations in affinity of different peptides for SWCNTs are also observed.

2) The MM-GBSA method is extended to evaluate the free energy of peptides interacting with CNTs. The relative binding affinities are compared with the experimental results to validate the model. The physical mechanism involved in this process is then studied in detail. Other effects that may influence peptide-CNT interaction are also investigated.

3) In order to obtain a general view of different binding affinity of hydrophobic and hydrophilic amino acids for the CNTs, binding free energy between each amino acid and the same CNT is estimated individually based on the all-atom model. The relative binding affinities of amino acids from the hydrophobic and hydrophilic groups are compared.

4) A coarse-grained HP lattice model is developed performing MC simulation to observe the macroscopic properties of the adsorption of peptides onto CNTs. The preliminary energy parameters are developed according to experimental observations and numerical results from the all-atom model. The thermodynamic quantities and conformational characteristics of peptides are also clarified.

Through these studies I am not only able to explore the detailed conformational properties and energetics of peptides interacting with CNTs, but also the peptide-CNT interaction mechanism from both microscopic and macroscopic views. The approach

used in this study would greatly benefit our understanding of the mechanisms of the interactions between CNTs and bio-materials, and hence promote the design of novel nano-devices.

These works will be thoroughly discussed in the following chapters.

1.5 Main contribution of the thesis

MD simulation method is used to provide a microscopic view of peptide-CNT interaction based on the atomic level. Self-insertion of peptides into SWCNTs and binding of peptides to the outer surface of SWCNTs are simulated. The energetics of interaction, as well as the conformational change of peptides is also discussed. Hydrophobicities of the peptides have high correlations with the peptides' propensities for the CNT.

Free energies of peptides binding to SWCNTs and peptides encapsulated into SWCNTs are estimated based on MM-GBSA method. The computational results are qualitatively comparable to binding affinities observed in experiments. It is observed that the van der Waals interaction plays a dominant role in peptide-SWCNT interaction.

In order to complement the all-atom MD simulation studies, a 2D HP lattice model is developed to simulate the interactions between peptides and CNTs. The parameters are qualitatively developed based on both experimental studies and MD simulation results. Both the internal energy and conformational entropy contribute to

the Helmholtz free energy. These results promise potential application of CNTs in drug delivery, drug design and protein control.

1.6 Organization of the thesis

The thesis consists of five chapters and the contents of each chapter are as follows.

In chapter 1, an overview of CNTs and peptides is presented. A literature review of bio-molecule functionalized CNT for various purposes of applications is carried out, which includes both experimental and theoretical approaches. Furthermore, the all-atom model and the coarse-grained model and the involved simulation algorithm for investigation of peptide-CNT interaction are introduced.

In Chapter 2, MD simulation method is used to provide a microscopic view of peptide-CNT interaction based on the atomic level. Intensive numerical simulation has been carried out for a large number of different peptides. Self-insertion of peptides into SWCNTs and binding of peptides to the outer surface of SWCNTs are simulated. The energetics of interaction, as well as the conformational change of peptides is discussed.

In Chapter 3, free energies of peptides binding to SWCNTs and peptides encapsulated into SWCNTs are estimated based on MM-GBSA method. The computational results are qualitatively compared with experimental results. It is observed that the van der Waals interaction plays a dominant role in peptide-SWCNT interaction. The properties of peptides that may have influence on their interaction

such as peptide length, hydrophobicities, aromatic rings are discussed. In discussion section, further simulations are performed to estimate binding free energy of each amino acid binding to the same SWCNT, and average binding affinity of hydrophobic and hydrophilic amino acids are obtained and compared.

In Chapter 4, a coarse-grained HP lattice model is developed to simulate peptides interacting with CNTs. The parameters are qualitatively developed based on experimental studies and MD simulation results. The exact enumeration based on this coarse-grained model enables us to look into thermodynamics of the interaction process. The impact of internal energy and entropy are discussed and conformational change of peptides interacting with CNT surface is traced.

The final Chapter 5 presents concluding remarks and recommendation for future research work.

Table 1.1 Abbreviations for amino acids, hydrophobicity (by K-D method) and the occurrence of the amino acids in proteins.

Amino acid	Three-letter abbreviation	One-letter abbreviation	Hydrophobicity	Occurrence (%)
Isoleucine	Ile	I	4.5	5.3
Valine	Val	V	4.2	6.6
Leucine	Leu	L	3.8	9.1
Phenylalanine	Phe	F	2.8	3.9
Cysteine	Cys	C	2.5	1.9
Methionine	Met	M	1.9	2.3
Alanine	Ala	A	1.8	7.8
Glycine	Gly	G	-0.4	7.2
Threonine	Thr	T	-0.7	5.9
Serine	Ser	S	-0.8	6.8
Tryptophan	Trp	W	-0.9	1.4
Tyrosine	Tyr	Y	-1.3	3.2
Proline	Pro	P	-1.6	5.2
Histidine	His	H	-3.2	2.3
Glutamic acid	Glu	E	-3.5	6.3
Asparagine	Asn	N	-3.5	4.3
Glutamine	Gln	Q	-3.5	4.2
Aspartic acid	Asp	D	-3.5	5.3
Lysine	Lys	K	-3.9	5.9
Arginine	Arg	R	-4.0	5.1

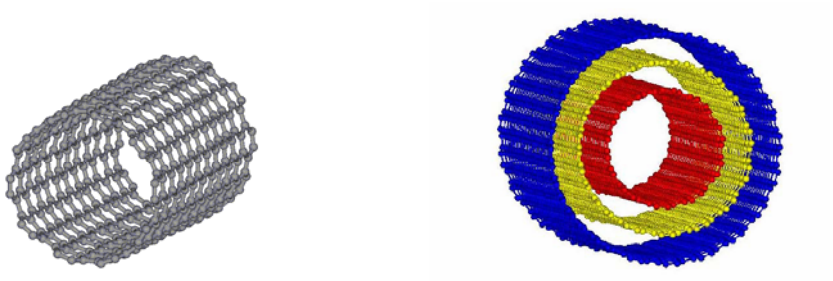


Figure 1.1 Structure of single-walled carbon nanotubes (SWCNT) and multi-walled carbon nanotubes (MWCNT).

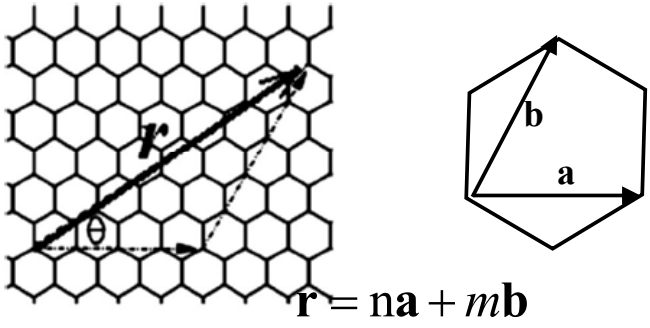


Figure 1.2 Definition of roll-up vector as linear combinations of base vectors *a* and *b* .

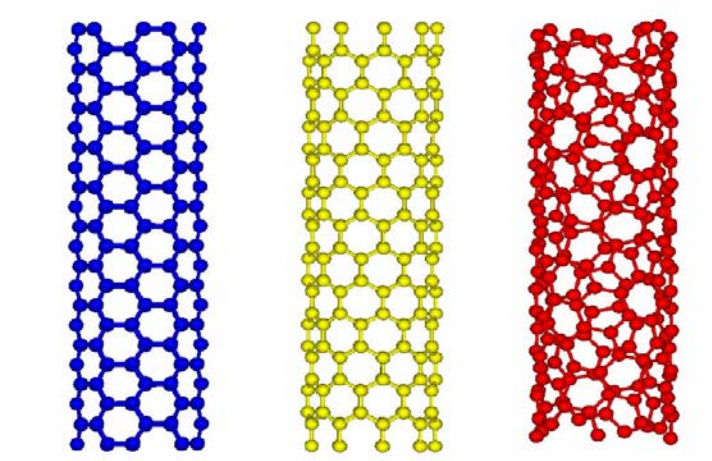


Figure 1.3 Three types of CNTs: armchair, zigzag, and chiral nanotubes.

Chapter 2

Molecular dynamics (MD) simulation based on the all-atom model

Molecular simulation based on the all-atom model has provided us with great insights into various biophysical phenomena. In this chapter, self-insertion behaviours of different peptides into SWCNTs in explicit water solvent are studied using MD simulation method. Section 2.1 introduces the molecular mechanics and force field for modelling of peptide-SWCNT potential function, and the algorithms for simulating the motions of the systems. Intensive numerical simulation is carried out for a large number of different peptides. In Section 2.2, the conformational change of the peptides and the energetics of peptide-SWCNT are analyzed. In addition, discussions on other factors that may influence the encapsulation mechanism are also included. Finally, the remarks are given in Section 2.3. In the following sections of this chapter, the word "propensity" will be used to represent the propensity of self-insertion of a peptide into a nanotube.

2.1 Modeling and simulation methods

2.1.1 Molecular Mechanics and empirical force fields for molecular simulation

In order to model the peptide-SWCNT interaction at the all-atom level, the analytical potential energy functions within the frame of classical molecular mechanics is used. The applications of molecular mechanics have employed MC methods, energy minimization, and MD methods to carry out the analytical potential energy calculations (Leach, 2001). Such methods have been widely used in computational biochemistry and biophysics, including the intrinsic strain of organic molecules, structures and dynamics of simple and complex liquids, thermodynamics of ligand binding to proteins, and conformational transitions in nucleic acids. Through these methods I am able to gain insights into a lot of biophysical phenomenon, especially for those are not accessible to experimental observations. Given the significance of the molecular mechanics model, many efforts have been made to consider both of the functional form and the parameters that is needed to apply the potential energy functions (or force fields).

The basis for molecular modeling is to construct the energy potential between atoms and molecules, and to develop parameters comprising the force field. The force fields used in molecular modeling are primarily designed to reproduce molecular structural properties, but they can also be used to predict other properties, such as molecular spectra. Generally a force field is designed to predict certain properties and the empirical parameters will be developed accordingly. Here the potential function

and the parameters are constructed according to the Amber99 force field (Wang et al., 2000).

In a molecular system, atoms are not rigidly positioned in a system, and both external and internal forces can induce atomic motions. Some motions may even have chemical effect. The effects of motions are described by energy because energy demonstrates the measurement of the ability of work. There are two types of energies involved in molecular modeling: the first one is kinetic energy, or motional energy, which is related to the speed and mass of a moving object. The higher the speed and the heavier the object is, the bigger work it can do. The second one is potential energy, or positional energy. That is, the value of potential energy is determined by inter-molecular, intra-molecular, and environmental forces. Objects generally move from higher potential energy place to lower potential energy place. A molecule changes from higher potential energy form to lower potential energy form and the configuration of the system is optimized.

In our model, the total potential energy between atoms is usually calculated as:

Total Potential Energy = Bond Stretching Energy + Angle Bending Energy

+ Torsional Energy + Non-Bonded Interaction Energy

Other forms of potential energy may also be added to the total potential energy function, according to different formula of force fields. The forces contributing to the whole simulated systems in this study are calculated as below.

Bond stretching

There are several models for estimation of bond stretching energy, and the most widely used approach is adopted, namely the Hooke's law formula. In this formula the energy varies with the square of the displacement from the reference bond length l_0

$$U(l) = \frac{k_r}{2}(l - l_0)^2 \tag{2.1}$$

Where the parameter l_0 is the reference bond length, or commonly called the equilibrium bond length (Figure 2.1). This parameter is the length that the chemical bond adopts when all the other terms in the force field are set to zero. The bond stretching forces between atoms are very strong and considerable energy is required to make a bond deviate significantly from its equilibrium length. This is reflected in the magnitude of the force constants k_r for bond stretching.

Angle bending

The contribution of angle bending energy is also described as the deviation of angles from their reference values using a Hooke's law:

$$U(\theta) = \frac{k_\theta}{2}(\theta - \theta_0)^2 \tag{2.2}$$

The contribution of each angle is characterized by a force constant k_θ and a reference value θ_0, as shown in Figure 2.1. Compared with the bond stretching energy, less energy is required to distort the angle away from equilibrium and the force constants for angle bending are proportionately smaller.

Torsional terms

For bond-stretching and angle-bending terms, substantial energies are required to cause significant deformations from their reference values. Therefore they are often referred to as 'hard' degrees of freedom. Most of the variation in structure and relative

energies is due to the complex interplay between the torsional and non-bonded contributions. The torsional potential represents the effect of rotation around a bond and as such it is a function of the dihedral angles. The torsional potential is expressed as

$$U(\omega) = \sum_{n=0}^{N} \frac{V_n}{2}[1 + \cos(n\omega - \gamma)]^2 \tag{2.3}$$

where ω is the torsion angle, V_n is often referred to as the rotational 'barrier' height, which is a qualitative indication of the relative energy barriers to rotation, and n denotes the periodicity.

Non-bonded interaction

The non-bonded interaction between independent molecules and atoms also plays an important role in determining the structure of individual molecules. The non-bonded interactions do not depend upon a specific bonding relationship between atoms. As shown in Figure 2.1, the atoms interact through space and the energy potential is usually modeled as a function of some inverse power of the distance. The non-bonded energy terms in a force field are usually considered in two forms, the electrostatic interactions and the van der Waals interactions. The electrostatic properties are calculated as a sum of interactions between pairs of point charges, using Coulomb's law. The energy function for the electrostatic interaction is written as:

$$U(r) = \sum_{i<j}^{atoms} \frac{q_i q_j}{\varepsilon_0 r_{ij}} \tag{2.4}$$

where q_i and q_j are atomic partial charges, ε_0 is the solvent dielectric constant.

The van der Waals interaction consists of both attractive and repulsive forces between atoms. The attractive contribution is due to dispersive forces. The short-range repulsive forces are often referred to as exchange forces, or overlap forces. This is to prevent the atoms from occupying the same region of space. The van der Waals potential function is usually modeled as Lennard-Jones 12-6 function.

$$U(r) = 4\varepsilon\left[\sum_{i<j}^{atoms}\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{2.5}$$

There are two adjustable parameters in the Lennard-Jones 12-6 potential energy function: the collision diameter σ and the well depth ε.

Therefore the total potential energy function is defined as:

$$U(\mathbf{r}) = \sum_{bonds}\frac{k_r}{2}(l-l_0)^2 + \sum_{angles}\frac{k_\theta}{2}(\theta-\theta_0)^2 + \sum_{dihidrals}\frac{V_n}{2}[1+\cos(n\omega-\gamma)]^2 + 4\varepsilon\left[\sum_{i<j}^{atoms}\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] + \sum_{i<j}^{atoms}\frac{q_i q_j}{\varepsilon_0 r_{ij}} \tag{2.6}$$

Energy parameters for all the atoms are constructed based on force field Amber99, which is suitable for general organic systems and drug-protein interaction simulation (Wang and Kollman, 2000). Carbon atoms on nanotubes are modeled as uncharged Lennard-Jones particles with a cross-section of $\sigma_{cc} = 3.400\text{Å}$ and a well depth of $\varepsilon_{cc} = 0.086\text{kcal mol}^{-1}$. Carbon-carbon bond length of 1.4 Å and angles of 120° are maintained by harmonic potentials with spring constants of 938 $\text{kcal mol}^{-1}\text{Å}^{-2}$ and 126 kcal mol^{-1} rad^{-2}, corresponding to parameters of sp2 carbons in the AMBER99 force field.

Van der Waals parameters between different types of atoms are calculated from the parameters of the pure atoms using combination rules (Cornell et al., 1995). The

cross-term non-bonded Lennar-Jones potential between different types of atoms could be obtained through certain combination rules, which is calculated as

$$\sigma_{\alpha\beta} = \frac{\sigma_{\alpha\alpha} + \sigma_{\beta\beta}}{2}, \qquad \varepsilon_{\alpha\beta} = \sqrt{\varepsilon_{\alpha\alpha}\varepsilon_{\beta\beta}} \tag{2.7}$$

2.1.2 The criteria of peptide selection

The selection of peptides is based on their specific biochemical properties such as hydrophobicity or their potential of being used as drugs. As shown in Table 2.1, some peptides are commonly used as therapeutic agents or extracted from disease-related proteins, others are designed to facilitate the identification of factors that may have influence on peptides' insertion into CNTs. Specifically, oxytocin (Pep3) is a commonly used drug peptide, Angiotensin II (Pep6) is one of the famous peptide hormones (Spyroulias et al., 2003); Pep7 and pep13 are extracted from N-terminal domain of mammalian PrPC (Chen and Prusiner, 1998) and yeast protein Sup35 (Gsponer et al., 2003). In this thesis I refer to the peptides by their assigned names in the table instead of presenting their sequences. The properties of these peptides, including the hydropathy distributions of these peptides, and the number of residues containing aromatic rings on the peptides are also listed in Table 2.2. Each amino acid on the peptide is indicated as either ‘H’ (hydrophobic) or ‘P’ (polar), according to K-D method (Kyte and Doolittle, 1982).

2.1.3 Generation of initial structures

Initial structures for simulation are generated using LEAP module in Amber 7

package (Case et al., 2002). Each simulated system consists of one SWCNT and one peptide solvated in water solvent. Initial structures of oxytocin and Angiotensin II are obtained from Protein Data Bank (PDB entry 1NPO and entry 1N9V, respectively). Other peptides adopt stretched initial conformations. The (12, 12) (diameter 16.1 Å) or (14, 14) (diameter 18.9 Å) types of SWCNTs are constructed through folding a graphite sheet of carbon rings to cylinder. Initially, the peptide is aligned along the nanotube axis with an appropriate initial distance between them, as illustrated in Table 2.1. The peptide-SWCNT complex is then surrounded by TIP3P water molecules (Jorgensen et al., 1983).

Periodic boundary conditions are applied to the system. The periodic boundary conditions enable a simulation to be performed using a relatively small number of particles, in such a way that the particles experience forces as if they were in bulk fluid. The particles being simulated are enclosed in a box which is then replicated in all three dimensions to give a periodic array. During the simulation only one of the particles is represented, but the effects are reproduced over all the images. Each particle interacts not only with the other particles but also with its own images in neighbouring boxes. The particles that leave one side of the box re-enter from the opposite side as their image. In this way the total number of particles in the central box remains constant.

2.1.4 Energy Minimization

Energy minimization algorithm is implemented on the generated initial structures before MD simulation is run. This method is widely used in molecular modeling and is

an integral part of techniques such as conformational search procedures, and can also be used to prepare a system for other types of calculation. For example, energy minimization is often used prior to a MD simulation in order to relieve any unfavorable or overlap conformations

In molecular modeling I am especially interested in minimization points on the energy surface. Minimum energy arrangements of the atoms correspond to stable states of the system; any movement away from the minimum energy configuration will lead to a higher energy. There may be a very large number of minima on the energy surface. The energy minimum corresponding to the very lowest energy is known as the global energy minimum. In order to identify those geometries of the system that correspond to minimum points on the energy surface, a minimization algorithm is used. I am also interested to know how the system changes from one minimum energy structure to another. The highest point on the pathway between two minima is of particular interest and is know as the saddle point, with the configuration of the atoms being the transition structure. Both energy minima and saddle points are stationary points on the energy surface, where the first derivation of the energy function is zero regarding all the coordinates.

2.1.4.1 Statement for the energy minimization problem

The minimization problem can be stated as: given a function f which depends on one or more independent variables $x_1, x_2, x_3, ..., x_i$, and the purpose is to find the values of those variables where f has a minimum value. At a minimum point the

first derivative of the function with respect to each of the variables is zero and the second derivatives are all positive, and the equation is written as

$$\frac{\partial f}{\partial x_i} = 0; \qquad \frac{\partial^2 f}{\partial x_i^2} > 0 \tag{2.8}$$

In this study, molecular mechanics minimizations are performed in Cartesian coordinates, where the energy is a function of 3N variables. For analytical functions, the minimum of a function can be found using standard calculus methods. However, this is not generally possible for molecular systems because of the complicated way in which the energy varies with the coordination. Therefore, minima are generally calculated and located using numerical methods, which gradually change the coordinates to produce configurations with lower and lower energies until the minimum is reached.

Minimization algorithms can be classified into two groups: those which use derivatives of the energy with respect to the coordinates and those which do not. Derivatives have the advantage that they provide information about the shape of the energy surface, and if used properly, they can significantly enhance the efficiency with which the minimum is located. The ideal minimization algorithm is the one that provides the answer as quickly as possible, using the least amount of memory, and yields the least standard errors. Most minimization algorithms can only go downhill on the energy surface and therefore they can only locate the minimum that is nearest (in a downhill sense) to the starting point. Although some specialized minimization methods can make uphill moves to seek out minima lower in energy than the nearest one, but no algorithm has yet demonstrated to be capable of locating the global energy minimum

from an arbitrary starting position. The shape of the energy surface may be important if one wishes to calculate the relative populations of the various minimum energy structures.

2.1.4.2 Derivative Minimization methods

There are non-derivative minimization methods and derivative minimization methods, here only the later one is performed. In order to use a derivative minimization method the derivatives of the energy with regard to the variables should be calculated. Derivatives provide very useful information in energy minimization. The direction of the first derivative of the energy (the gradient) indicates where the minimum lies, and the magnitude of the gradient implies the steepness of the local slope. The energy of the system can be lowered by moving each atom in response to the force acting on it; the force is equal to minus the gradient. Second derivatives indicate the curvature of the energy function, and the information that can be used to predict the place where the function will change direction (i.e. pass through a minimum or some other stationary point).

In order to calculate the derivative, the energy function is first written as a Taylor series expansion about the point x_k

$$f(x) = f(x_k) + (x - x_k) f'(x_k) + (x - x_{k)})^2 f''(x_k)/2 + \ldots \quad (2.9)$$

For a multidimensional function, the variable x is replaced by the vector $\mathbf{x}$ and matrices are used for the various derivatives. The derivative methods can be classified according to the highest-order derivative used. For example, first-order methods use

the first derivatives whereas second-order methods use both first and second derivatives. Here derivation methods based on the first-order minimization methods are used. The two most frequently used methods in molecular simulations, the steepest descents and the conjugate gradient method are sequentially used on the input initial structure. Both algorithms gradually change the coordinates of the atoms as they move the system closer and closer to the minimum energy point.

1) The steepest Descents Method

The steepest descents method moves in the direction parallel to the net force, which corresponds to walking straight downhill in the direction of the gradient. For 3N Cartesian coordinates this direction is most conveniently represented by a 3N-dimensional unit vector, $\mathbf{s}_k$. Thus

$$\mathbf{s}_k = -\mathbf{g}_k / \left|\mathbf{g}_k\right| \tag{2.10}$$

where $\mathbf{g}_k$ is the first derivation of the potential energy. After determining the direction along which to move, the minimum point can choose to locate by performing a line search or take a step of arbitrary size along the direction of the force.

2) Conjugate Gradients Minimization

The conjugate gradients method produces a set of directions and does not show the oscillatory behaviors of the steepest descents method in narrow energy valleys. In the steepest descents method both the gradients and the direction of successive steps are orthogonal. In conjugate gradients, however, the gradients at each point are orthogonal but the directions are conjugate. A set of conjugate directions has the property that for a quadratic function of M variables, the minimum will be reached within M steps. The

conjugate gradients method moves in a direction $\mathbf{v}_k$ from point $\mathbf{x}_k$, where $\mathbf{v}_k$ is computed from the gradient at the point and the previous direction vector $\mathbf{v}_{k-1}$:

$$\mathbf{v}_k = -\mathbf{g}_k + \gamma_k \mathbf{v}_{k-1} \tag{2.11}$$

where γ_k is a scalar constant calculated by

$$\gamma_k = \frac{\mathbf{g}_k \cdot \mathbf{g}_k}{\mathbf{g}_{k-1} \cdot \mathbf{g}_{k-1}} \tag{2.12}$$

In the conjugate gradients method all of the directions and gradients satisfy the following relationships:

$$\mathbf{g}_i \cdot \mathbf{g}_j = 0 \tag{2.13}$$

$$\mathbf{v}_i \cdot U_{ij}'' \cdot \mathbf{v}_j = 0 \tag{2.14}$$

$$\mathbf{g}_i \cdot \mathbf{v}_j = 0 \tag{2.15}$$

It is noted that Equation (2.11) can only be implemented from the second step onwards and therefore the first step in the conjugate gradients method is the same as the steepest descents. The one-dimensional minimum in each direction is located to ensure that each gradient is orthogonal to all previous gradients and that each direction is conjugated to all previous directions.

The steepest descent method is performed prior to conjugate gradient, as the starting structure is in some distance away from the minimum. However, conjugate gradients is much better once the initial strain has been removed by steepest descent method (Leach, 2001). Combining both of the algorithms is efficient for obtaining the energy-optimized structure of the system.

2.1.5 Integration of the motions of particles using finite difference method

Although energy minimization is able to lead the system to the optimized energy states, the system may be trapped in local minimum if only minimization method is used. Therefore MD simulation is used after energy minimization to simulate the system into equilibrium state. MD simulations are applied in a wide range of applications. They can be used to study the thermodynamic properties of gas, liquid, solid, phase transitions, as well as internal motions of macromolecules (e.g., polymers, DNA, proteins). On the other hand, hydrodynamical fluid flow, plasma and electrons, and transport phenomena could also been simulated by MD calculations.

A common feature of all MD algorithms is that the positions of molecules are evolved with time by integrating the Newton's equations of motion. Macroscopic quantities are extracted from the microscopic trajectories of particles. In principle, these equations can be solved by using any standard finite difference algorithm. However, in practice, the evaluation of force is the most time consuming step in a molecular simulation, and in MD simulation it is required that the algorithm should be efficient to handle large systems. Therefore many commonly used algorithms such as the conventional Runge-Kutta method are computationally too expensive because they require multiple force evaluations. The other requirement is that the integration algorithm must be well behaved for the force encountered in MD, which implies that the order of the algorithm defined as the highest order of the time step for the equation of the coordinates should be at least two (Sadus, 1999). In order to meet these requirements some finite difference algorithms have been developed and specifically tuned for applications to MD (Dahlquist and Bjork, 1974).

Finite difference algorithms used in molecular simulation can be classified as either predictor or predictor-corrector methods. In the predictor methods, the molecular coordinates are updated from quantities that are either calculated in the current step or that are from the previous steps. The Verlet (1967) algorithm and its subsequent modifications are typically widely used predictor algorithms. In contrast, predictor-corrector algorithms involve predicting new molecular coordinates. The predicted coordinates are used to calculate the value of related functions and to correct the initial prediction. The Gear (1971) algorithm is also one of the most widely used predictor-corrector algorithm for MD calculations. The detailed discussion on various integration schemes have been discussed in a number of references (Berendsen and van Gunsteren, 1986; Allen and Tildesley, 1987; Hockney and Eastwood, 1988; Rapaport, 1995; Gubbins and Quirke, 1996; Frenkel and Smit, 1996), while leap-frog Verlet algorithm is used for molecular simulations in this study.

The problem posed specifically by MD is formulated as follows. The force acted on a particle during displacement is expressed as Newton's second law of motion

$$\mathbf{F}_i = m_i \mathbf{a}_i \tag{2.16}$$

Where $\mathbf{F}$ is the force acting on the particle, and m and $\mathbf{a}$ are the particle's mass and acceleration, respectively. The equation can be written as the second derivative of displacement $\mathbf{r}$ with respect to time t.

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} \tag{2.17}$$

The dynamic behavior of the whole system can be obtained by solving Equation (2.17). Integrate the equation with respect to time for a small interval yields:

$$\frac{d\mathbf{r}_i}{dt} = \left(\frac{\mathbf{F}_i}{m_i} \right) t + c_i \tag{2.18}$$

At the time $t = 0$, the velocity is given by the constant c_1 which represents the initial velocity $\mathbf{v}_i$. Therefore the expression of the velocity at time t is shown as:

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{a}_i t + \mathbf{v}_i \tag{2.19}$$

The equation can be integrated and and the displacement is written as

$$\mathbf{r}_i = \mathbf{r}_i t + \frac{\mathbf{a}_i t^2}{2} + c_2 \tag{2.20}$$

Where the constant c_2 is the initial position of the particle. Equation (2.20) allows us to calculate the particle's displacement from the initial velocity and acceleration.

As illustrated in Equation (2.20), the coordinates of the particle evolves with time. Any changes in particle displacement will also affect the strength of forces imposed between the particles, as well as the velocities, and accelerations between particles. Consequently, a general method is required to estimate all of the dynamic properties of the system. Different integrators have their specific advantages and shortcomings, and they are selected according to different requirements in various simulation systems. For example, the Gear algorithms are slower than the Verlet algorithms but the difference is not significant. The Verlet algorithms are easier to implement than the Gear algorithms, which is an advance. The Verlet algorithm belongs to the symplectic class of integrators that preserves a canonical structure. It was found that symplectic algorithms are likely to be the preferred choice for the long-time integration of Hamiltonian dynamic systems (Leimkuhler and Skeel, 1994), and therefore are suitable for carrying out simulations of the interactions between peptides and CNTs.

The Verlet Algorithm

Verlet proposed that a simple predictor method can be successfully used in MD simulations (Verlet, 1976). Based on the original Verlet algorithm, many subsequent modifications and improvements have been made on the algorithm. These 'Verlet family' of algorithms are now widely used in MD simulations. In this section, the implementation of Verlet algorithm and the Leap-frog Verlet algorithm are shown.

1) The original Verlet Algorithm

The starting point for the Verlet algorithm is a Taylor series expansion, which is in common with the corrector-predictor algorithm. I start by considering the Taylor series expansions about $\mathbf{r}(t)$.

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \frac{d\mathbf{r}}{dt}\Delta t + \frac{1}{2!}\frac{d^2\mathbf{r}}{dt^2}\Delta t^2 .... \tag{2.21}$$

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \frac{d\mathbf{r}}{dt}\Delta t + \frac{1}{2!}\frac{d^2\mathbf{r}}{dt^2}\Delta t^2 .... \tag{2.22}$$

By adding the two equations, I have

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \frac{d^2\mathbf{r}}{dt^2}\Delta t^2 .... \tag{2.23}$$

Equation (2.23) is known as Verlet's algorithm. Such algorithm enables us to calculate the position of the molecules without calculating their velocities. However, the velocities could also be derived in order to determine the system kinetic energy,

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)}{2\Delta t} \tag{2.24}$$

In the Verlet algorithms, the velocity calculations are optional because they do not actually contribute to the estimation of the atomic coordinates. However, in order to determine the total linear momentum and kinetic energy, the velocity calculations

should be included. The advantage of Verlet method lies in that it's compact, easy to implement, time reversible, and the energy is conservable for long time steps.

2) Leap-Frog Verlet Algorithm

The 'leap-frog' algorithm proposed by Hockney (1970) calculates the velocity at half intervals.

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \Delta t \mathbf{v}(t+\Delta t/2) \tag{2.25}$$

$$\mathbf{v}(t+\Delta t/2) = \mathbf{v}(t-\Delta t/2) + \Delta t \mathbf{a}(t) \tag{2.26}$$

$$\mathbf{v}(t) = \frac{\mathbf{v}(t+\Delta t/2) - \mathbf{v}(t-\Delta t/2)}{2} \tag{2.27}$$

In the original Verlet method, the next position is determined both by the current position and the velocity at the next half-time interval. In leap-frog algorithm, however, the velocity at the next half-time interval is determined by the current acceleration and the velocity at the previous half-time intervals. The current velocity is calculated as the average value of the velocities at the next and the previous half-time intervals.

The major advantage of the leap-frog algorithm is that numerical imprecision is reduced because it is based on differences between smaller quantities. However, the calculation of velocities relies on velocity averaging at different time intervals and the algorithm is not self starting. I need to calculate the velocity at $t = -\Delta t/2$ initially. This difficulty can be overcome by using reverse Euler estimates of velocities.

$$\mathbf{v}(-\Delta t/2) = \mathbf{v}(0) - \Delta t \mathbf{a}(0)/2 \tag{2.1.30.}$$

In common with the original Verlet algorithm, the calculation of velocities lags one time interval behind the calculation of positions. The time step should be properly

chosen to guarantee the stability of the integrator. In our simulation Δt is chosen in the length scale of femtoseconds (fs).

2.1.6 Statistical mechanics ensembles

The target of the statistical mechanics is to convert the information generated at the microscopic level (including atomic and molecular positions, velocities, *etc.*) into macroscopic terms. The thermodynamic variables of a system include number of particles (N), temperature (T), pressure (P), internal energy (E), *etc.* The system is often analyzed with a few variables maintained at constant values. The choice of the ensemble determines which thermodynamic quantities can be evaluated and it also governs the overall simulation algorithm (Hill 1956; McQuarrie 1976; Chandler, 1987). Algorithms for implementation of different types of ensembles are discussed as below.

2.1.6.1 Implementation of statistical ensembles

There are generally four types of ensembles in common use, including the microcanonical, or constant-NVE, the canonical, or constant- NVT ensemble, the isothermal-isobaric constant-NPT ensemble, and the grand canonical constant- μVT ensemble. For each ensemble, the aforementioned thermodynamic quantities are specified or kept constant. Other thermodynamic variables could be determined by ensemble averaging. The fluctuations are evaluated by calculating the deviations of the thermodynamic quantities from the average value at any particular states.

The microcanonical ensemble, or constant-NVE ensemble is the default ensemble for MD calculations because Newton's equations of motion conserve energy. Therefore a microcanonical ensemble is generated automatically by solving Newton's equation of motion.

The probability density for the microcanonical ensemble is proportional to

$$\delta(H(\mathbf{\Gamma}) - E_{des}) \tag{2.28}$$

Where $\mathbf{\Gamma}$ denotes the set of particle positions and momenta (or quantum numbers), and $H(\mathbf{\Gamma})$ is the Hamiltonian. The delta function selects out those states of an N-particle system in a box of volume V that have the desired energy E_{des}.

The microcanonical partition function is written as

$$Q_{NVE} = \sum_{\mathbf{\Gamma}} \delta(H(\mathbf{\Gamma}) - E_{des}) \tag{2.29}$$

For an quasi-classical expression for Q_{NVE}, in our atomic systems, the quantity is calculated from a factor of $N!$

$$Q_{NVE} = \frac{1}{N!}\frac{1}{h^{3N}} \int d\mathbf{r}d\mathbf{p}\,\delta(H(\mathbf{r},\mathbf{p}) - E_{des}). \tag{2.30}$$

Where $\int d\mathbf{r}d\mathbf{p}$ stands for integration over all $6N$ phase space coordinates, and h is the Planck's constant. The appropreiate thermodynamic potential is the negative of the entropy

$$-S/k_B = -\ln Q_{NVE} \tag{2.31}$$

The density for the canonical ensemble is proportional to

$$\exp(-H(\mathbf{\Gamma})/k_B T) \tag{2.32}$$

And the partition function is expressed as

$$Q_{NVT} = \sum_{\Gamma} \exp(-H(\mathbf{\Gamma})/k_B T) \tag{2.33}$$

For our atomic system, the function can also be expressed in quasi-classical form

$$Q_{NVT} = \frac{1}{N!}\frac{1}{h^{3N}}\int d\mathbf{r}d\mathbf{p}\exp(-H(\mathbf{r},\mathbf{p})/k_BT). \tag{2.34}$$

The appropriate thermodynamic function is the Helmholtz free energy A

$$A/k_BT = \ln Q_{NVT} \tag{2.35}$$

The probability density for the isothermal-isobaric (NPT) ensemble is proportional to

$$\exp(-(H+PV)/k_BT) \tag{2.36}$$

The ensemble averaged term of $<H>+P<V>$ gives the thermodynamic enthalpy. The appropriate partition function is

$$Q_{NPT} = \sum_{\Gamma}\sum_{V}\exp(-(H+PV)/k_BT) = \sum_{V}\exp(-PV/k_BT)Q_{NVT} \tag{2.37}$$

The equation could be written in quasi-classical form for the atomic system:

$$Q_{NPT} = \frac{1}{N!}\frac{1}{h^{3N}}\frac{1}{V_0}\int dV\int d\mathbf{r}d\mathbf{p}\exp(-(H+PV)/k_BT) \tag{2.38}$$

The corresponding thermodynamic function is the Gibbs free energy G

$$G/k_BT = -\ln Q_{NPT} \tag{2.39}$$

In order to generate state points in the constant-NPT ensemble, the changes in the sample volume must be provided along with energy.

The density function for the grand canonical ensemble is proportional to

$$\exp(-(H-\mu N)/k_BT) \tag{2.40}$$

Where μ is defined as the chemical potential of the simulated system. In the equation number of particles N, along with the coordinates and momenta of those particles are all variables. The grand canonical partition function is

$$Q_{\mu VT} = \sum_{\Gamma}\sum_{N} \exp(-(H-\mu N)/k_B T) = \sum_{N} \exp(\mu N / k_B T) Q_{NVT} \tag{2.41}$$

The equation could be written in quasi-classical form for an atomic system

$$Q_{\mu VT} = \sum_{N} \frac{1}{N!}\frac{1}{h^{3N}} \exp(\mu N / k_B T) \int d\mathbf{r} d\mathbf{p} \exp(-PV / k_B T) \tag{2.42}$$

and the corresponding thermodynamic function is $-PV/k_BT$:

$$-PV/k_BT = -\ln Q_{\mu VT} \tag{2.43}$$

Addition and removal of particles are allowed for the grand ensemble. The constant NVT and the constant NPT ensembles are most frequently used in this work, and quantities are transferable between different ensembles.

2.1.6.2 Thermodynamic average

The ensembles enable us to calculate the basic thermodynamic properties of the model system provided a suitable phase function is identified. In this section I derive the equations for calculation of thermodynamic quantities regardless of subscripts which identify particular ensembles. The functions are often derivatives of one of the characteristic thermodynamic functions Ψ_{ens} .

In a model system, the total energy of the system is calculated as the sum of kinetic energy and the potential energy

$$E_{tot} = < K > + < U > \tag{2.44}$$

The kinetic energy is a sum of contributions from all the individual particle momenta, while the potential energy is calculated by summing over all pairs, triplets etc. of particles, corresponding to the potential energy function.

The temperature and pressure are evaluated using the viral theorem, which is written in the form of generalized equipartition (Munster, 1969):

$$< p_k \partial H / \partial p_k >= k_B T \tag{2.45}$$

$$< q_k \partial H / \partial q_k >= k_B T \tag{2.46}$$

for any generalized coordinate q_k or momentum p_k. For an atomic system, 3N terms of the form $p_{i\alpha}^2 / m_i$ (α rages from 1 to 3) may be summed up and Equation (2.45) is expressed as

$$< \sum_{i=1}^{N} |\mathbf{p}_i|^2 / m_i >= 2 < K >= 3Nk_B T \tag{2.47}$$

This equation is the familiar equipartition principle: an average energy of $k_B T / 2$ counts for each degree of freedom. An instantaneous kinetic temperature function could be defined as

$$T = 2K / 3Nk_B = \frac{1}{3Nk_B} \sum_{i=1}^{N} |\mathbf{p}_i|^2 / m_i \tag{2.48}$$

The pressure is evaluated via Equation (2.46). Each coordinate derivative in Equation (2.46) is the negative of a component of the force $\mathbf{f}_i$ on some molecule i, and by summing over N molecules, I have:

$$-\frac{1}{3}\left\langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \nabla_{\mathbf{r}_i} V \right\rangle = \frac{1}{3}\left\langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{f}_i^{tot} \right\rangle = -Nk_B T \tag{2.49}$$

where $\mathbf{f}_i^{tot}$ denotes the sum of inter-molecular forces and external forces. The latter re related to external pressure, and can be considered as the effect of the container walls on the system:

$$\frac{1}{3}\left\langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{f}_i^{ext} \right\rangle = -PV \tag{2.50}$$

If the 'internal virial' W is defined

$$-\frac{1}{3}\sum_{i=1}^{N}\mathbf{r}_i\cdot\nabla_{\mathbf{r}_i}V=\frac{1}{3}\sum_{i=1}^{N}\mathbf{r}_i\cdot\mathbf{f}_i^{tot}=W \tag{2.51}$$

Regarding the inter-molecular forces, I have

$$PV=Nk_BT+\langle W\rangle \tag{2.52}$$

The pressure is evaluated as the ensemble averaged value of the instantaneous pressure function (Cheung, 1977)

$$P=\rho k_BT+W/V=P^{it}+P^{ex} \tag{2.53}$$

where ρ is the number density.

2.1.7 Implementation details

In this work, MD simulations are performed using the program based on the molecular simulation package Amber 7 (Case et al., 2002). Force field Amber 99, is used for modeling the peptide-SWCNT interaction system.

The procedure of implementing simulations of peptide insertion into SWCNT interaction is shown in Figure 2.2. After the simulation system is selected and initial structure is generated, potential energy minimization is performed on each of the initial systems, the steepest descent algorithm is used for the first 10 cycles and conjugate gradient minimization is run for the following 19990 cycles. Then MD simulations are implemented on the energy-minimized systems, and the structure is saved every 5 ps. Constant-NVT ensemble is run first for 100 ps to target the temperature of the system from initial value of 0 K to 300 K. Subsequently the simulation is performed at constant pressure (1 bar), and constant temperature (300 K) for 1900 ps (Berendsen et

al., 1984). Newton's equations of motion are integrated with a step size of 1 fs, with all non-bonded cutoff distance of 8 Å was used. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm and a relative tolerance of 0.00001 (Rychaert et al., 1977) The partical-mesh Ewald method (PME) algorithm with cubic-spline interpolation (1 Å grid width) was applied to calculate the full electrostatic energy of the simulated system (Darden et al., 1993).

2.2 Results and Discussion

2.2.1 Diverse propensities

The stability of the conformational change of the peptides upon interacting with SWCNTs is tested through analysis of the root meant square deviations (RMSDs) of the backbone atoms on peptides. The RMSD trajectory of backbone atoms of one representative peptide during the 2ns of simulation time is shown in Figure 2.3. It is observed that over the simulation time, the RMSDs are stable with no unreasonable oscillations.

Analyzing the simulation results, it is shown that some peptides are able to insert into SWCNTs while others are not. As an example of peptides that have strong affinities for SWCNTs, Figure 2.4 (a)-(f) show the snapshots of structures of oxcytocin (pep3)-SWCNT system at different simulation time. Water molecules are stripped for a clearer visualization purpose. It is observed that the first residue of oxcytocin begins to enter the nanotubes at about 50 ps, and it has been encapsulated in the nanotube completely by the time up to 500 ps. Afterwards the peptide is trapped in the tube and

does not escape, with only slight fluctuations in its relative position to the nanotube.

Significant differences in propensities are also observed. Five among the 17 simulated peptides, pep1 through pep5, could quickly self-insert into the SWCNTs. Pep12 through pep17 fail to insert into the SWCNTs, they either move away or approach the outer surface of the nanotube. Representative conformation of pep13-SWCNT interaction configuration at simulation time of 2ns is shown in Figure 2.5. The remaining six peptides, pep6 through pep11 are able to enter or partly enter SWCNTs, but the insertion processes are very slow. Based on the above observation, these simulated peptides can be roughly classified into three classes as shown in Table 2.3: (i) Peptides that have strong affinities for SWCNTs, which could insert into SWCNTs within 800 ps simulation time (pep1 through pep5); (ii) Pep6 through pep11, which could completely or partly enter SWCNTs within 2 ns simulation time, but the insertion progresses are greatly hindered compared to those in class (i) (encapsulated completely at least after 1.2 ns); (iii) Peptides that move away from the SWCNTs (pep12 through pep17).

Different behaviors of peptides interacting with SWCNT could also be illustrated by tracing the relative distance between the two objects at instant simulation time. Figure 2.6 shows the normalized center of mass (COM) distances as the function of simulation time of three representative peptides from the respective three categories. It is apparent that the peptide from the first class shows the sharpest reduction of COM distance to the CNT, while the one from the third class indicates the opposite tendency. The peptide from the second class takes the intermediate.

2.2.2 Energetics of peptide-CNT interaction

In order to examine the convergences and stabilities of MD simulations, the energetic trajectories of the simulation are also traced. The total potential energy trajectories and its components are analyzed with reference to simulation time. The data demonstrates that the energies converge with only small fluctuations.

The energetics of the process of peptides insertion into SWCNTs is further analyzed. For peptides that could insert into the nanotube, the potential energy of the whole system decreases with the reduction of COM distance between the peptide and the nanotube, as shown in Figure 2.7(a). The system potential energy reaches its minimum when the mass center of the peptide is close to the centre of the SWCNT, and subsequently remains at the minimum value. In contrast, in the case that the peptide could not insert into nanotube, the potential energy is not dependent on the distance between the peptide and the nanotube. For instance, the COM distance between pep13 and the nanotube never reaches low values (Figure 2.8(a)). The peptide tends to interact with the water molecules more than the nanotube.

Figure 2.7 (b) demonstrates the relationship between the energy sum of van der Waals and electrostatic interactions (non-bonded interactions) and the COM distance. Comparing Figure 2.7 (a) with Figure 2.7 (b), it is found that the energy sum is dominant in driving the peptide into the potential well. Contributions of other interaction energies are shown in Figure 2.7 (c), the difference between potential energy and non-bonded interactions energy is not dependent on the COM distance

between the two objects in our simulation system. It is found that the electrostatic interactions and the van der Waals interactions among the peptide, the CNT and water play an important role in determining whether peptides could be encapsulated into SWCNTs. Although pep13 is not able to self-insert into the SWCNT, the consistent energetic contributions are also observed as shown in Figure 2.8 (a)-(c).

In order to further examine the significance of van der Waals interaction on the insertion of the peptide into SWCNT, I selected three systems in which the peptides could insert into the nanotubes with original van der Waals parameters, cut the cross-section parameter σ_{cc} by half (so that the van der Waals attraction is artificially reduced), and repeated the simulation for 2ns with the same initial conformations. Simulation results indicate that these peptides are not able to enter the nanotube with modified parameters. For the three selected systems with normal and modified van der Waals parameters, Figure 2.9 compares the normalized COM distances between the peptide and nanotube as the function of simulation time. It is shown that the COM distances decrease rapidly within first 800ps for the normal van der Waals parameters cases; while the COM distances remained roughly unchanged with time for the systems with modified van der Waals parameters. This indicates that with the changed van der Waals interactions, the peptide only make random motion around the nanotube, and hence the insertion progress of peptide into nanotube is greatly hindered. These results imply that van der Waals interactions play an important role in the self-insertion of peptide into SWCNT.

For pep13 and pep14 (from the third class), I have also simulated the system with

initial structure of peptide being positioned within the inner space of SWCNT, then observe its conformational change over 1 ns of simulation time. It is found that both peptides remain inside the SWCNT. It is speculated that if initially positioned inside the CNT, the peptide would be trapped by the potential well of the CNT. The peptide does not necessarily move out spontaneously due to the energy barrier, even for peptides exhibiting low propensities.

2.2.3 Impacts of CNT size

Other factors that may have effects on the propensities involve diameters of SWCNTs, lengths of SWCNTs, etc.

Taking into consideration that both the steric hindrance effect of CNTs and the van der Waals interactions between CNTs and peptides depend on the sizes of CNTs, the insertion process of peptides should be CNT-dimension dependent.

As shown in Figure 2.10, if the diameter of the nanotube is too small, some bulky residues (for example, those have aromatic rings) may be stucked at the entrance of the SWCNT even though the insertion process starts normally. Therefore peptides would not be able to fully enter the nanotube. By contrast, if the diameter of the nanotube is large enough (Figure 2.4), the same peptide can be encapsulated into the nanotube completely. As an example, the critical diameter for a SWCNT to intake pep3 is approximately 18.9 Å .

In order to clarify the effect of the SWCNT length on the propensity, two systems are tested, each composes of the same peptides but CNTs with different

lengths. In both cases, peptides could self-insert into the CNTs, and their normalized COM distances are compared in Figure 2.11. It is observed that the COM distance between the longer CNT and the peptide does not decrease after the peptide is fully encapsulated into the CNT. This may be due to that for such cases, there exist a potential well with broadened at bottom, the system could stay anywhere at the flat bottom of the potential well, and thus the peptide does not necessarily reach the central of the CNT. Therefore lengths of SWCNTs are not critical factors as long as the tubes are long enough to encapsulate the peptides.

2.2.4 Correlations between hydrophobicities and propensities

It is well-established that hydrophobic interactions play an essential role on the interactions between CNTs and peptides. In our study, hydrophobic interactions are not calculated explicitly because the macroscopic statistical properties such as averaged free energy and the entropy are not easily accessible in explicit solvent for MD system. Instead, I use K-D method (Kyte and Doolittle, 1982) to assign each amino acid residue a hydrophobicity value. The average hydrophobicity for each peptide corresponding to Table 2.1 is then calculated and plotted as bars in Figure 2.12. Comparing their average hydrophobicity values, it is found that peptides consisting of more hydrophobic amino acid residues tend to enter the SWCNTs more easily than those consisting of more hydrophilic polar residues. Such result is not surprising from the viewpoint of thermodynamics: hydrophobic solutes such as the SWCNT and hydrophobic peptide tend to aggregate to effectively reduce the hydrophobic surface

exposed to polar solvent and consequently to reduce the overall free energy of the system. It is also observed that pep8 through pep17 have comparable average hydrophobicity value, while they are differentiated in their affinities for SWCNTs as mentioned above. Other characteristics of amino acid residues such as sizes, charges and aromaticities may have impact on the interactions between SWCNTs and peptides.

2.3 Remarks

In this chapter, self-insertion behaviors of peptides into SWCNTs are studied in explicit solvent environment. Intensive numerical studies show that some peptides are capable of self-inserting into SWCNTs while others are not. Energetic analysis indicates that the electrostatic interactions, combined with the van der Waals interactions, contribute most to the insertion of peptides into SWCNTs. Other properties of peptides are also tested for their impact on the interaction. For example, sizes of SWCNTs should be chosen properly corresponding to dimensions of peptides to facilitate the insertion. From a macroscopic view, the hydrophobicity of a peptide correlates with the propensity.

Table 2.1 The list of the simulated peptides, type of SWCNTs, number of surrounding water molecules as well as the initial distance between the most adjacent two atoms of the peptide and the SWCNT along the nanotube axis.

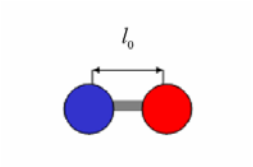
Assigned peptide name	Peptides sequences	Type of SWCNTs	Number of surrounding Water Molecules	Initial distances (Å)
pep1	VEAGG	(12, 12)	3172	2.5
pep2	AAFEL	(12, 12)	3072	4.1
pep3	GLPCNQIYC	(14, 14)	6373	2.7
pep4	LLLLLLLL	(12, 12)	4174	4.0
pep5	FLIGI	(12, 12)	3201	2.4
pep6	DRVYIHPF	(12, 12)	4985	5.6
pep7	PHGGGWGQ	(12, 12)	3750	4.0
pep8	EEEEEEEE	(12, 12)	4484	3.2
pep9	QQQQQQQQ	(12, 12)	4468	2.5
pep10	KKKKKKKK	(12, 12)	4628	3.2
pep11	SQNGNRE	(12, 12)	4058	3.5
pep12	DNNNRTEE	(12, 12)	3052	4.0
pep13	GNNQQNY	(12, 12)	4539	3.9
pep14	DDDDDDDD	(12, 12)	4488	3.2
pep15	DKNNRQE	(12, 12)	3993	4.4
pep16	RRRRRRRR	(12, 12)	4808	3.2
pep17	NNNNNNNN	(12, 12)	4451	4.0

Table 2.2 The properties of simulated peptides. For hydropathy distributions, each amino acid on the peptide is indicated as either 'H' (hydrophobic) or 'P' (polar), according to K-D method.

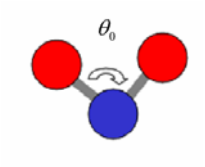
Assigned peptide name	hydropathy distributions	Number of residues containing aromatic rings
pep1	HPHPP	0
pep2	HHHPH	1
pep3	PHPHPPHPH	1
pep4	HHHHHHHH	0
pep5	HHHPH	1
pep6	PPHPHPPH	2
pep7	PPPPPPPP	0
pep8	PPPPPPPP	0
pep9	PPPPPPPP	0
pep10	PPPPPPPP	0
pep11	PPPPPPP	0
pep12	PPPPPPPP	0
pep13	PPPPPPP	1
pep14	PPPPPPPP	0
pep15	PPPPPPP	0
pep16	PPPPPPPP	0
pep17	PPPPPPPP	0

Table 2.3 The list of the simulated peptides classified into three classed based on the insertion behaviors.

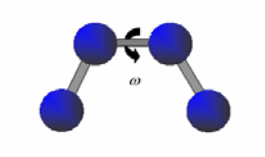
peptide name	Class
pep1-pep5	First class
pep6-pep11	Second class
pep12-pep17	Third class



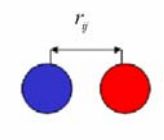
Bond stretching



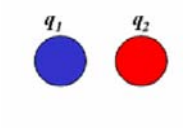
Angle bending



Torsional



Electrostatic interaction



Van der Waals interaction

Figure 2.1 Illustrations for potential energies between particles.

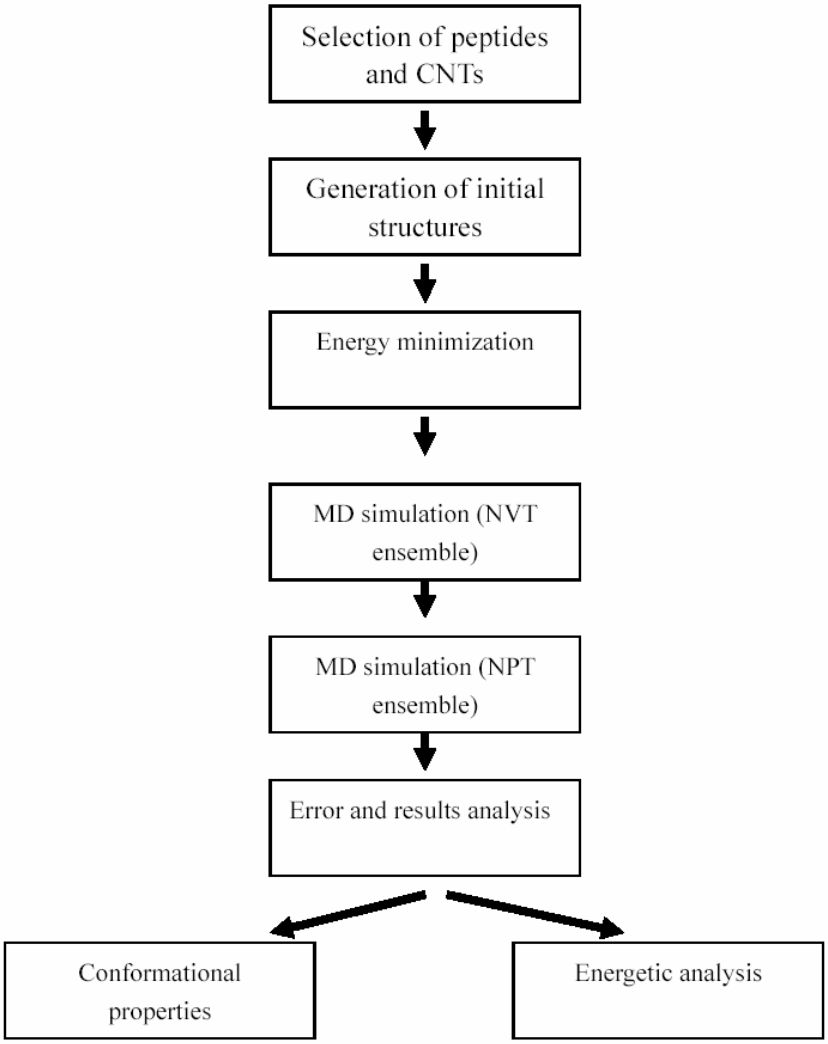


Figure 2.2 Strategies for implementation of simulation procedure.

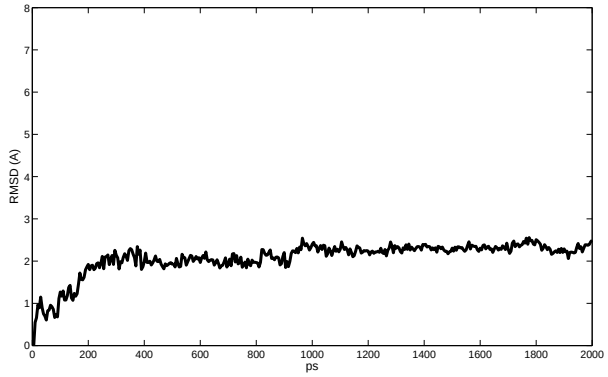


Figure 2.3 The RMSDs for the backbone atoms on pep3 against the simulation time.

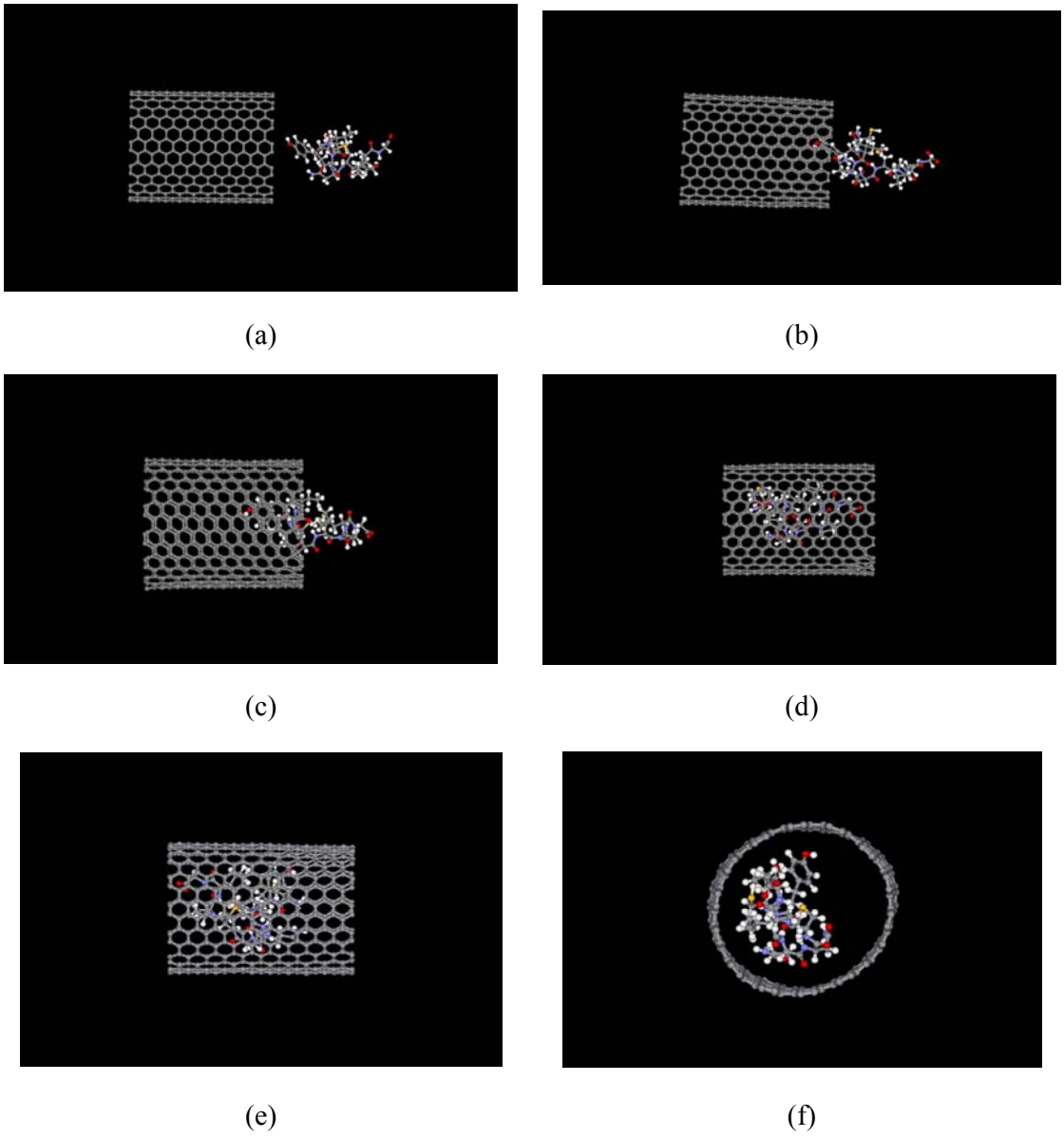
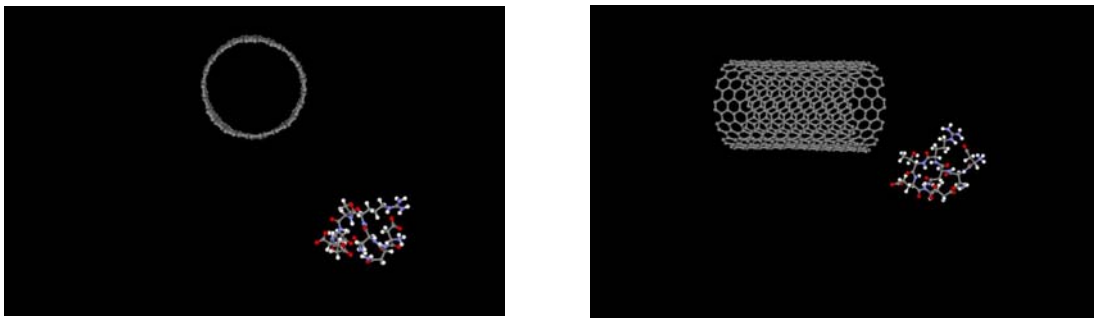


Figure 2.4 The snapshots of the conformation of oxytocin (pep3) insertion into SWCNT at different simulation time: (a) initial structure, (b) 50ps, (c) 100ps, (d) 500ps, (e) 2ns. (f) shows the final structure (2ns) viewed along the axis of nanotube. The images are created with DS ViewerPro 5.0 software (Accelrys Inc., San Diego, CA)



(a) (b)

Figure 2.5 The snapshots of the final structure of pep13 interacting with SWCNT at simulation time of 2ns. The images are created with DS ViewerPro 5.0 software (Accelrys Inc., San Diego, CA)

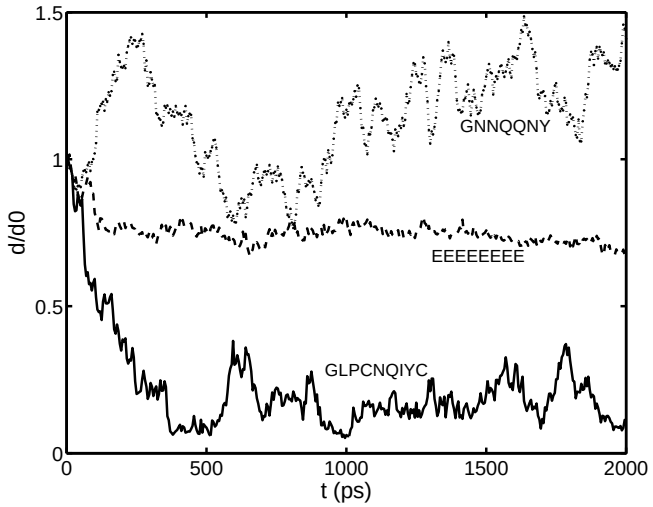


Figure 2.6 Normalized Center of Mass (COM) distances between the peptide and SWCNT as the function of MD simulation time. d_0 is the initial COM distance between the peptide and the SWCNT, and d is the distance at the Corresponding simulation time.

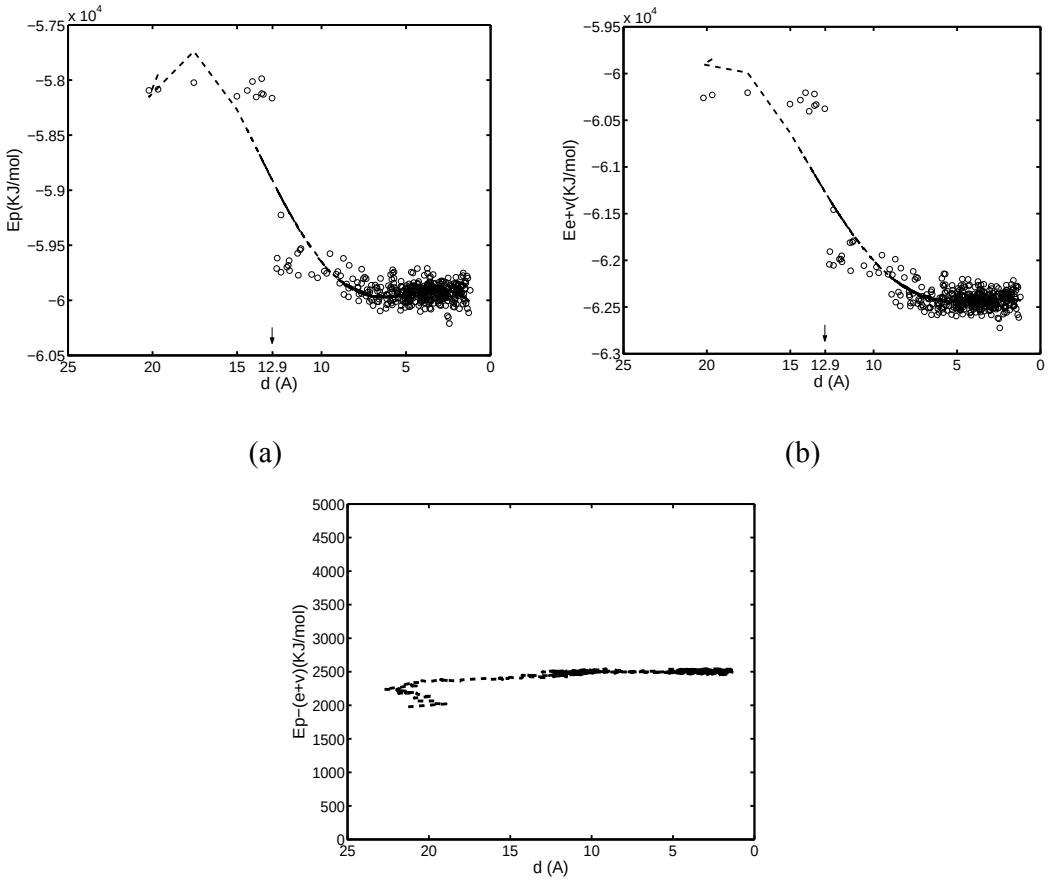
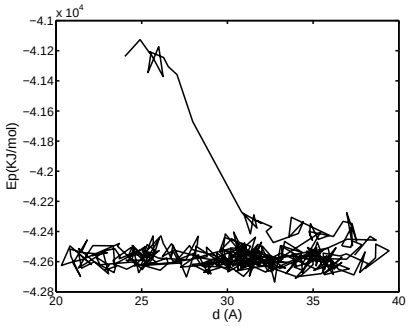
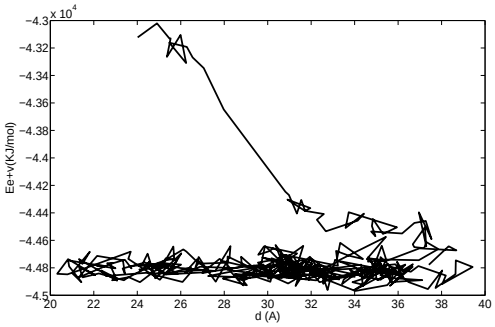


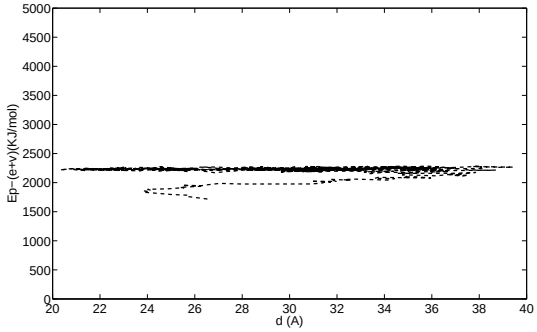
Figure 2.7 (a) Potential energy of the simulated oxytocin (pep3)-SWCNT system as the function of COM distance between SWCNT and pep3. (b) Energy sum of the van der Waals energy and the electrostatic energy (non-bonded interaction energy) as the function of COM for pep3-SWCNT system. (c) The difference between potential energy and non-bonded interaction energy as the function of COM distance between pep3 and SWCNT. The half length of the nanotube is 12.9 Å .



(a)



(b)



(c)

Figure 2.8 (a) Potential energy of the pep13-SWCNT system as the function of COM distance of SWCNT and pep13. (b) Energy sum of the van der Waals energy and the electrostatic energy (non-bonded interaction energy) as the function of COM for pep13-SWCNT system. (c) The difference between potential energy and non-bonded interaction energy as the function of COM distance between pep13 and SWCNT. The half length of the nanotube is 14.6 Å .

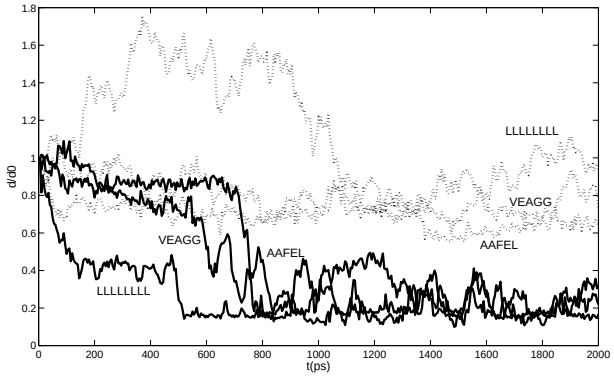


Figure 2.9 Normalized COM distances between the peptide and nanotube as the function of simulation time. Solid lines represent the cases with normal van der Waals parameters, dash lines are for the cases with the modified van der Waals parameters.

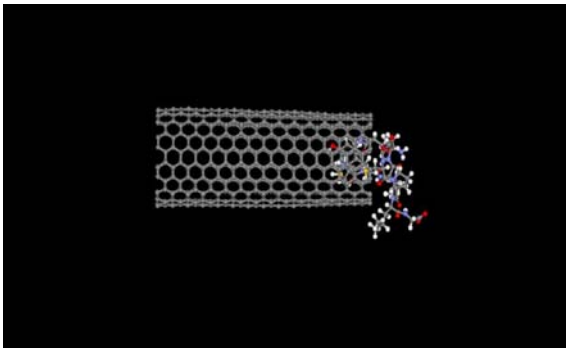


Figure 2.10 Snapshots of conformation of oxcytocin and (12, 12) type SWCNT at simulation time of 2ns. The diameter of the nanotube is 16.1 Å , smaller than that of (14,14) in Figure 2.4.

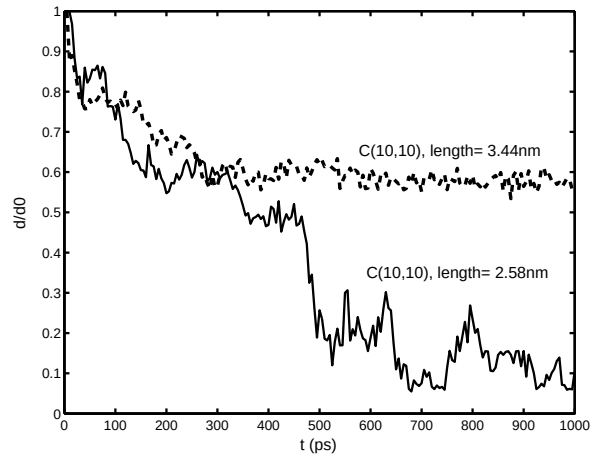


Figure 2.11 Normalized Center of Mass (COM) distances between the peptide and SWCNT as the function of MD simulation time for the same peptide inserting into SWCNTs of different length.

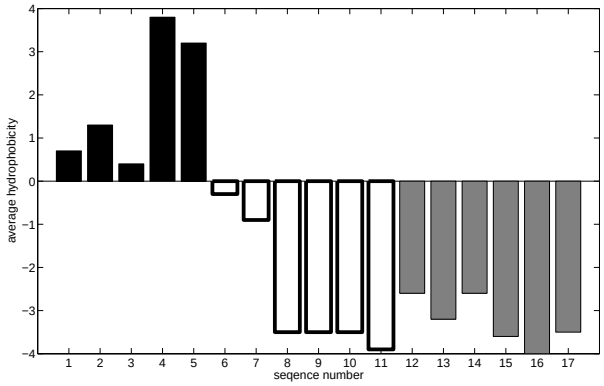


Figure 2.12 Average hydrophobicity for simulated peptides. Higher values of the average hydrophobicity imply that the peptides are more hydrophobic. Sequence numbers of peptides are in accordance as listed in Table 1. Pep1 through pep5 rapidly insert into the SWCNTs, pep6 through pep11 partially insert into SWCNTs or insert completely with slow speed, pep12 through pep17 fail to insert into SWCNTs.

Chapter 3

Estimation of interaction free energy

The capability of evaluating the interaction free energy is essential for us to understand the mechanism of physical and chemical reactions. The MM-GBSA method has been successfully applied in many studies to predict the interaction free energies between biomolecules.

In this chapter, I extend the MM-GBSA method to estimate the binding free energy of peptides onto the SWCNTs. The organization of this chapter is listed as follows. Section 3.1 presents the method of using continuum water medium solvent for calculating energetic contributions. The binding free energy model used in this work takes into account contributions of both the solute and the solvent and therefore the paths though the intermediate states are not traced. In section 3.2, the change in free energies upon binding are compared with binding affinities reported from experiments. Furthermore, the energetic contributions are analyzed. Our results show that the five peptides tested have diverse affinities for CNTs. The van der Waals interaction is the most significant contributor. The interactions between aromatic rings have also been explored.

Section 3.3 gives further discussions arising from the previous study. As the computational model is qualitatively validated by comparing with the experimental results, the interaction free energy between encapsulated peptides and the SWCNT is

further investigated. The consistent finding that the van der Waals interaction dominates the interaction is drawn.

In order to understand the various interaction affinities between individual amino acid and the same CNT, the same simulation strategy is applied on the free energy estimation of each of the amino acid-SWCNT interaction. A general comparison of the relative binding affinity between hydrophobic and hydrophilic groups of amino acids for the same SWCNT is carried out.

Section 3.4 gives the remarks of this chapter.

3.1 Methods

There are two major steps involved to estimate the binding free energy of each peptide to CNTs. Firstly, MD simulations in explicit solvent to obtain the equilibrium structure of the peptide, the CNT, and the peptide-CNT complex are performed separately. Secondly the water molecules are removed from the equilibrated structure and the energy calculation is performed in implicit solvent using the MM-GBSA method. The peptide-CNT binding free energy is then obtained as the energy difference between the complex and the individual systems.

3.1.1 Generation of initial structures

To evaluate the free energies of peptides binding to SWCNTs, I adopt five peptide sequences, where the experimental relative binding affinities of these peptides for CNTs are available (Wang et al., 2003). Sequences of the peptides and their

average hydrophobicities calculated by the K-D method (Kyte and Doolittle, 1982) are listed in Table 3.1 for reference. A positive hydrophobicity value indicates that the peptide is hydrophobic and the negative value corresponds to hydrophilic peptides. Throughout the rest of the thesis, I refer to the peptides by their assigned name in Table 3.1, instead of listing the whole lengthy residue chains. The molecular properties of these peptides including the hydropathy distribution and the number of residues bearing aromatic rings are further illustrated in Table 3.2.

In order to estimate the free energy change upon binding, for each system estimated, MD simulation experiment is carried out respectively for the complex of the peptide and the SWCNT solvated in water, the peptide in water, and the SWCNT in water. The Amber99 force field is used for building amino acid residues.

Initial structure of a SWCNT is constructed as a hollow cylinder rolled up from a graphite sheet. A (6, 6) type SWCNT with diameter of 8.1 Å and length of 25.8 Å is used. Parameters of carbon atoms on SWCNTs are adopted the same as illustrated in Chapter 2.

Initially, the peptide is constructed as a fully extended structure. Each complex of peptide-SWCNT contains one SWCNT and one peptide. The peptide is positioned approximately parallel to the SWCNT and parts of them contact directly. Subsequently the complex of peptide-SWCNT is surrounded by a layer of at least 10 Å of TIP3P water molecules. Water molecules are not accessible to the contact regions of peptide-SWCNT complex. Periodic boundary conditions are applied throughout the simulation.

3.1.2 MD simulation in explicit solvent

The procedure of the MD simulation in explicit solvent is roughly identical to those stated in Chapter 2, while a brief description is given below. Firstly, for each initial structure of the complex, peptide or SWCNT solvated in explicit water molecules, energy minimization is performed to avoid steric clashes. The steepest descent method of minimization is used for the first 10 cycles and conjugate gradient minimization is run for the following 19990 cycles. MD simulation of constant volume and constant temperature (NVT) ensemble is then run for 100 ps to raise the temperature from initial value of 0K to 300K. Subsequently the structure is simulated for 1 ns under the conditions of constant pressure and constant temperature (NPT) ensembles.

3.1.3 Calculations of energy contributions

In MM-GBSA model, the molecular mechanical (MM) energies of the molecules are combined with the GBSA continuum solvent model to describe total free energies.

3.1.3.1 Implementation of the GB model

To calculate the solvent induced free energies of peptide-SWCNT interaction, the GBSA method is implemented. Both the polar and nonpolar terms are taken into consideration. The polar term is calculated by implementing a pairwise descreening

approximation. This particular form includes a Debye-Huckel term to account for salt effects at low salt concentrations (Srinivasan et al., 1999):

$$G_{pol} = -\frac{1}{2}\left(1 - \frac{e^{-\kappa f^{gb}}}{\varepsilon_0}\right)\sum_{ij}^{atoms}\frac{q_i q_j}{f^{gb}} \tag{3.1}$$

Where q_i and q_j are atomic partial charges, ε_0 is the solvent dielectric constant, κ is the Debye-Huckel screening parameter, and the double sum runs over all pairs of atoms. f^{gb} is a function that interpolates between the effective Born radius α_i, when the distance r_{ij} between atoms is short, and the function is written as:

$$f^{gb} = \left[r_{ij}^2 + \alpha_i \alpha_j \exp(-r_{ij}^2 / 4\alpha_i \alpha_j)\right]^{1/2} \tag{3.2}$$

In the equation the effective Born radius α_i demonstrates how deeply buried a charge is in the low-dielectric medium, for example a protein or nucleic acid. The variable α_i depends not only on the intrinsic radius η_i of atom i, but also on the relative positions and intrinsic radii of all the other atoms in the molecules in the system:

$$\alpha_i^{-1} = \eta_i^{-1} - \sum_{j \neq i} g(\mathbf{r}, \eta) \tag{3.3}$$

These general characteristics can be applied to several pairwise analytical approximations of the Born radii (Bashford et al., 2000; Hawkins et al., 1995). In particular, the equation for $g(\mathbf{r},\ \eta)$ has been provided by Hawkins et al. (Hawkins et al., 1995).

The intrinsic radius η_i is estimated from the atomic radius R_i by Equation (3.4). The overlap factors S_i approximately account for simultaneous overlap of three or more atoms. An overall offset is included to adjust the magnitudes of solvation

energies to match those from Poisson-Boltzmann calculation, therefore η_i is calculated by

$$\eta_i = S_i(R_i + b_{offset}) \tag{3.4}$$

The parameters are referred to the work of Tsui and Case (Tsui and Case, 2001).

Although the GB model is an approximation to the physical situation with actual solvent molecules, eliminating the solvent molecules can greatly decrease the number of atoms in the system. This generally results in speed ups in the simulation, which is particularly significant as the number of water molecules required in the explicit solvent simulation is large.

Due to the lack of a predefined solvent box, GB simulations are independent of the shape of the system. Another advantage of continuum models lies in the implicit averaging over the solvent degrees of freedom, so that this averaging does not need to be done by sampling the solute configurations explicitly. This property greatly simplifies calculations of thermodynamic quantities, and is very efficient for examining binding free energies between peptides and CNTs.

The GBSA model

While the GB model is used to compute the electrostatic solvent polarization energy, the total solvation free energy also includes nonpolar contributions. The nonpolar free energy arises from solvent–solvent cavity terms and attractive van der Waals solvent–solute interaction terms (Tomasi and Persico, 1994). These two terms are approximately proportional to the solvent accessible surface area (SA) of the solute molecule, with a surface tension term used as the linear scaling factor. The two energy

terms, polar (G_{pol}) and nonpolar (G_{nonpol}) energy together make up the GBSA model (Qiu et al., 1997).

In order to obtain solvent accessible surface areas from solute configurations, the linear combination of pairwise orbitals (LCPO) approach (Weiser et al., 1999) is implemented. This is a fast analytical approximation for computing exposed areas of atoms in molecules. This method makes use of linear combinations of terms composed from pairwise overlaps of hard spheres. First derivatives with respect to atomic coordinates are calculated to obtain the forces arising from this term.

When two hard spheres, i and j show an overlap conformation, then A_i, the accessible surface area of sphere i, is given by

$$A_i = S_i - A_{ij} \tag{3.5}$$

Where S_i is the surface area of the isolated sphere,

$$S_i = 4\pi r_i^2 \tag{3.6}$$

In which r_i is the radius of the sphere.

A_{ij} is the area of sphere i buried inside sphere j, or the overlap area of sphere i with j, which can be expressed as the function of the internuclear distance d_{ij} and the sphere's radii r_i and r_j

$$A_{ij} = 2\pi r_i \left(r_i - \frac{d_{ij}}{2} - \frac{r_i^2 - r_j^2}{2d_{ij}} \right) \tag{3.7}$$

The main equation in the LCPO method takes the following form to compute the accessible surface area of atom i:

$$A_i = P_1 S_1 + P_2 \sum_{j \in N(i)} A_{ij} + P_3 \sum_{\substack{j,k \in N(i) \\ k \in N(j) \\ k \neq j}} A_{jk} + P_4 \sum_{j \in N(i)} A_{ij} \left(\sum_{\substack{j,k \in N(i) \\ k \in N(j) \\ k \neq j}} A_{jk} \right) \tag{3.8}$$

Where $N(i)$ denotes the neighbor list of i, or the list of spheres that overlap with sphere i. The second term in the equation involves the sum of pairwise overlaps if sphere i with its neighbors. The third term is the sum of overlaps of the neighbors of i with each other. The fourth term is a further correction for multiple overlaps in the system. $P_1 - P_4$ are parameters estimated through multiple linear regression.

Contributions of the nonpolar term are then calculated proportional to the solvent accessible surface (A) using Equation (3.9). Here the parameter $\sigma = 0.005 kcal/mol * \text{Å}^2$ (Sitkoff et al., 1994).

$$G_{nonpol} = \sigma A \tag{3.9}$$

3.1.3.2 Evaluation of binding free energy from its components

In order to implement the free energy calculation in continuum models, the final PDB structures of peptide-SWCNT complex, the SWCNT, and unbound peptides structures obtained from explicit solvent MD simulation are adopted. After the removal of water molecules, 50ps of MD simulations are carried out to arrive at the equilibrium state and another 50ps for data collections. All the energy components are sampled by averaging the results over the final 50ps.

As shown in Equation (3.10), the binding free energy between a peptide and a SWCNT is estimated as the difference between the free energies of the complex in

water solvent ($G^s_{complex}$), and that of the sum of the SWCNT (G^s_{cnt}), the peptide ($G^s_{peptide}$) solvated in water, respectively.

$$\Delta G = G^s_{complex} - (G^s_{cnt} + G^s_{peptide}) \tag{3.10}$$

The strategy is also shown in Figure 3.1.

For each system representing the solvated molecules and the surrounding solvent, the free energy is calculated from the solute's gas-phased molecular mechanics energy E_{MM} , and the solvation-induced free energy G_{sol} , which is expressed as

$$G^s = E_{MM} + G_{sol} \tag{3.11}$$

According to molecular mechanics theory, E_{MM} could be calculated as

$$E_{MM} = E_{\text{internal}} + E_{vdw} + E_{ele} \tag{3.12}$$

Which means that E_{MM} is composed of the internal energy (E_{internal}), the van der Waals interaction energy (E_{vdw}), and the electrostatic energy (E_{ele}).

The internal energy includes the bond stretching, the angle bending and the torsion energy, which can be further expressed as

$$E_{\text{internal}} = E_{bond} + E_{angle} + E_{torsion} \tag{3.13}$$

The contribution of the solvation free energy, G_{sol} , includes both the polar and nonpolar terms:

$$G_{sol} = G_{pol} + G_{nonpol} \tag{3.14}$$

The polar (G_{pol}) and nonpolar (G_{nonpol}) energy contribution to the solvation free energy is estimated using GBSA (Onufriev et al., 2000; Tsui and Case, 2001) method as stated below.

Note that the energies and free energy contributions are all state functions. Therefore they can be calculated at different stages of the interaction systems regardless of their paths of evolutions to these states.

In order to obtain a clear picture of the energy contributions, each term of ΔE and ΔG is also listed, which is calculated as the difference of this value between two states before and after binding.

$$\Delta E = E(complex) - (E(peptide) + E(nanotube)) \tag{3.15}$$

The binding free energy can be finally calculated using

$$\Delta G = \Delta E_{MM} + \Delta G_{sol} = \Delta E_{\text{internal}} + \Delta E_{vdw} + \Delta E_{ele} + \Delta G_{pol} + \Delta G_{nonpol} \tag{3.16}$$

3.2 Results

3.2.1 Peptides display diverse propensities

In the five peptide-SWCNT complex systems, the conformations of peptides change to favor interactions with SWCNTS. Some peptides wrap around the nanotubes completely while others partly contact the surface of nanotubes. As shown in Figure 3.2, pep22 (Figure 3.2 (a)) does not wrap the SWCNT completely, but it interacts more with water instead. Pep20 binds tightly onto the SWCNT surface (Figure 3.2 (d)), which means that a larger part of the peptide interacts with the SWCNT. Apparently the contact area of pep22 with the SWCNT surface is much smaller than that of pep20. Behaviors of unbound peptides solvated in water are also simulated and analyzed. Pep22 is folded driven by the clustering in its end groups. Pep20 shows a high tendency of clustering in aromatic rings. This structure implies a favorable

conformation of pep20 interacting with SWCNT, particularly for $\pi - \pi$ stacking of aromatic rings on the peptide and SWCNT surfaces (Wang et al., 2003).

3.2.2 Error analysis of the systems in explicit solvent

To examine the convergences and stabilities of MD simulations, the energetic trajectories and structural changes of the systems are traced. The potential energy trajectories are analyzed with reference to simulation time. The mean value of the potential energies and their standard deviations during the last 500ps are provided in Table 3.3 to verify the thermal stability of the simulations. The data demonstrates that the energies converge with only small fluctuations.

The stability of the simulation is further studied through analysis of the root meant square deviations (RMSDs) of the backbone atoms on peptides, both in bound and unbound states. One randomly selected RMSD trajectory of backbone atoms of pep20 in the two states is provided in Figure 3.3. The RMSDs are stable with no unreasonable oscillations.

3.2.3 Free energy calculations and energetic analysis

Binding affinities between CNTs and peptides are sensitive to amino acid sequences, implying the possibility for design of nanotubes-based probes. In order to understand interactions between these two kinds of materials, analyzing the binding

free energies between CNTs and different peptides sequence is an efficient and reliable approach.

For the five calculated systems, the mean values of the absolute energy contributions and their standard deviations estimated from the 50ps data-collection period are provided in Table 3.4. Overall the energies maintain constant and fluctuate within standard errors, except that E_{ele} and G_{pol} appear to have relatively larger fluctuation. However, the energy sum of these two terms, E_{ele_total} converges, and the errors are canceled by each other. Furthermore, one prolonged run for pep22 has been performed using the same method for 1ns for testing, 500ps for equilibrium and 500ps for data collection. However, the longer period of simulation doesn't add to the convergence of the energy contributions.

Based on the results of MD simulations and the free energy calculations, the binding free energies ΔG and energy contributions are shown in Table 3.4. In this table, ΔG qualitatively correlates with the binding affinities between peptides and SWCNTs. The greater the free energy changes between the two states before and after binding, or the lower the value of ΔG, the stronger is the binding affinities. The free energy implies that the peptide and the SWCNT should overcome certain value of energy barrier to dissociate once they bind to each other.

Figure 3.4 lists the scaled experimental results of peptides' binding affinities to CNTs (Wang et al., 2003) and our calculated binding free energies. In the experimental study, values of plaque-forming units correlate with binding affinities. I qualitatively compare our estimated free energies with that of the plaque-forming units. Among the

five peptides, absolute free-energy value of pep20 binding to CNT is the highest, corresponding to the strongest binding affinity observed in experiments. Furthermore, in agreement with experimental results, with only one mutation from Trp to Ser at the sixth position of peptide sequences, from pep18 to pep19, and pep20 to pep21, both mutations show apparent influence on decreasing of the binding affinities.

The relative energy values are in good agreement with experiments, except for the calculation of pep22. From experiment results, pep22 has much weaker affinity for SWCNTs than pep18 and pep20, but slightly stronger than pep19 and pep21. In our simulation, binding free energy of pep22 to the SWCNT ranks the weakest among the five.

Although the calculated free energies qualitatively reflect experimental observations, the energy contributions and their roles in binding of peptides to SWCNTs remain unclear. I observe that the contributions of the internal energies are quite small, and hence non-bonded interactions play a dominant role.

It is also noted that the electrostatic interaction energy is balanced to some degree by the polar solvation energy. The sum of the two terms could be observed by E_{ele_total}. The polar contribution of the solvation screen much of the electrostatic interaction in gas-phase, so that the value of ΔE_{ele_total} is much smaller than that of ΔE_{ele} or ΔG_{pol} respectively. On the other hand, the contribution of the nonpolar solvation energy to the total binding energy is almost negligible. Therefore it seems that the van der Waals interaction is the driving force for the binding process. For example, both pep18 and pep20 have stronger van der Waals interactions, and they have stronger

binding affinities. Table 3.6 shows the binding free energy differences between pep18 and pep19, pep20 and pep21. There is only one mutation from Trp to Ser at the sixth position of the peptide sequences for both pairs. This results in an unfavorable loss in binding free energies. Comparing interaction energies of pep18-SWCNT with that of peptide 19-SWCNT complex, there's a loss of 9.47 kcal/mol in the van der Waals interaction and 4.93 kcal/mol in ΔE_{ele_total}, the sum of ΔE_{ele} and ΔG_{pol}. The mutation from peptide 20 to peptide 21 leads to a loss of 10.24 kcal/mol in the van der Waals interaction and 2.43 kcal/mol in ΔE_{ele_toltal}. In both cases, although ΔE_{ele} and ΔG_{pol} have different trends of fluctuation, their summation shows a slight loss after the mutation. There is a stronger contribution of the van der Waals interaction before the mutation from Trp to Ser. The major loss in binding free energy is due to the loss of this interaction.

3.2.4 The effect of aromatic rings

The interactions between aromatic rings appear to be essential for the binding of peptides to SWCNTs. It has been observed experimentally that aromatic rings affected ligands or peptides' affinities for CNTs significantly (Chen et al., 2001; Wang et al., 2003).

Similar results through calculations of the five peptides' affinities for SWCNTs are also obtained. As discussed in the previous section, pep19 has only one mutation from pep18 at the sixth position, from Trp to Ser, and the same mutation is from pep20 to pep21. It is found that significant propensity changes result from this crucial

mutation. The snapshots of pep20-SWCNT complex conformation over the course of the MD simulation are also recorded. The orientation of planar of aromatic ring on Trp on the sixth position of pep20 approximately parallels the CNT surface, with an offset from the rings in CNT surface. This is also called an “offset stacked” interaction configuration (Waters, 2002). The configurational trace reveals that a stable structure is formed due to the aromatic affinities within the peptide-CNT complex. Aromatic rings on other positions of the peptide may interact with the CNT surface with edge-to-face or other contact configurations.

Based on the results that the van der Waals interaction dominates the binding of peptides to SWCNTs, it is expected that the stacked structure leads to stronger interaction energy. In order to further clarify the interaction mechanisms between aromatic rings and SWCNTs, a complex comprised of only one amino acid of Trp and one SWCNT is simulated for energy calculation. The complex in explicit water solvent is simulated first to obtain the equilibrium structure. The final equilibrium structure shows that the aromatic ring on the residue Trp also presents offset stacked configuration towards rings on the SWCNT. Afterwards the complex structure is adopted for potential energy calculation in vacuum. The Trp residue is located at different positions along the length of the SWCNT, each separated by approximately the length of the radius of one aromatic ring, with the trajectory parallel to the CNT axis. On the other hand, the residue is positioned in a greater distance from the CNT surface, which is approximately the radius of an aromatic ring. Illustration of the interaction conformation is provided in Figure 3.5. Potential energies of the system

with Trp in these different positions are calculated repeatedly with restrained structures. Analysis of the results indicates that the potential energy fluctuation along the length of CNT surface is quite small. The energy difference is within the range of 1 kcal/mol. However, potential energy change depending on the distance between the ring and CNT surface is much more substantial, the energy increases by 8.3 kcal/mol as the distance is enlarged approximately the radius of a ring. Despite the variations of the potential energy at different locations on the CNT surface, the distance between aromatic rings on the amino acid residue and the CNT surface is more crucial, at least within certain range. Therefore the stacked conformation is expected to possess stronger van der Waals interaction, which leads to an optimized lower free energy.

3.3 Discussions

3.3.1 Functionalizing CNTs with peptides

One of the main challenges in applications of CNTs is the dispersion of nanotubes in solution and control of their assembly in solvent. Coating the CNTs with peptides could enable peptides to interact noncovalently with CNTs and therefore is one approach to modify the solubility of SWCNTs. Average hydrophobicities of the five peptides investigated in this chapter are all slightly below zero, which indicates that on the whole, these peptides are prone to be hydrophilic and are in favor of interacting with water. When hydrophilic peptides are used to coat the CNTs, the tubes will be more soluble in water instead of being highly hydrophobic.

3.3.2 Calculations of the entropic term

In this study, energy contributions are summed unweighted to calculate the change of free energies, and the entropic change of solute is not taken into consideration. The entropy calculation is one of the greatest challenges in MD simulation, accurate and complete estimation for entropy through MD calculation is still under exploration. The entropy contribution can be estimated by performing normal mode analysis on the three species, for example the vibrational, transitional, and torsional terms of the entropy of the solute to the binding free energies (Massova et al., 1999), which were options for estimation of the entropy. However, in practice entropy contributions can be neglected if only a comparison of states of similar entropy is expected, such as two ligands binding to the same protein, or peptides in the comparable length interacting with the same SWCNT. The reason for this is that normal mode analysis calculations are computationally expensive and tend to have a large margin of error that introduces significant uncertainty in the result (Case et al., 2002).

3.3.3 Calculations of free energy of peptides encapsulated into SWCNTs

3.3.3.1 Implementation details

Based on the procedure of evaluating free energies of peptides binding to SWCNTs, the free energies of peptides encapsulated into SWCNTs are further discussed. Three representative peptides, named according to Chapter 2 as pep3 (GLPCNQIYC), pep4 (LLLLLLLL), and pep4a (LLLL) are selected for investigation

of the interaction free energy of the peptide encapsulated into the SWCNT. As illustrated in Chapter 2, peptide 3 is a commonly used drug peptide named oxytocin, pep4 and pep4a are peptides composed of the same hydrophobic amino acid residue Leu but with different length. All the three peptides are able to spontaneously insert into the SWCNT. The purpose of investigating pep4 and pep4a is to clarify the significance of peptide length on affinity of the peptide for SWCNTs.

The procedure of implementing free energy calculation based on MM-GBSA method follows that introduced in Section 3.1. In order to estimate the interaction free energy, for each estimated system, MD simulation experiment is carried out respectively for the complex of the peptide and the SWCNT solvated in water, the peptide in water, and the SWCNT in water. The initial structure of pep3 is extracted from Protein Data Bank with PDB entry 1NPO. Initial structures of pep4 and pep4a are generated using LEAP module provided by Amber 7 suit of programs in their stretched structure. A type (14, 14) SWCNT is tested for pep3 and type (12, 12) SWCNT for pep4 and pep4a respectively to encapsulate peptides. In each peptide-SWCNT system, the peptide is initially positioned along the axis of the SWCNT with a small distance between them (approximately 2 nm). All the subsystems are surrounded by a 10 Å layer of TIP3P water molecules. Periodic boundary conditions are applied.

MD simulation is carried out on the generated structures solvated in explicit solvent. After energy minimization, MD simulation of constant volume and constant temperature (NVT) ensemble is then run for 100 ps to raise the temperature from

initial value of 0K to 300K. Subsequently the structure is simulated for 1900 ps under the conditions of constant pressure and constant temperature (NPT) ensembles. A time step of 1 fs is used to integrate the Newton's equation of motion and the coordinates of the structures are saved every 5ps. The free energy and the energy contributions are calculated according to MM-GBSA model, with 50ps for equilibrium and 50ps for data collection.

3.3.3.2 Results

Based on the two-state theory, the conformations of the peptide in solvent and in peptide-SWCNT complex are compared. Figure 3.6 illustrates the conformations of the peptide in solvent and pep3-SWCNT complex at the two end states, respectively. It is shown that the peptide exhibits a compact structure in bulk solvent (Figure 3.6 (a)). However, as the peptide is encapsulated into the SWCNT, stretched conformations are observed. As shown in Figure 3.6 (c), the side view of the peptide-SWCNT complex indicates an adsorption of the peptide onto the inner surface of the SWCNT. It is inferred that the existence of the CNT would influence the conformation of the peptide.

The mean values of the absolute energy contributions and the standard deviations estimated based on the 50ps data collection period are listed in Table 3.7. Overall the energies converge and fluctuate within standard errors. As shown in Table 3.8, the contributions of the internal energies are quite small, and hence non-bonded interactions play a dominant role. The nonpolar solvation energy $G_{nonpolar}$ is

proportional to the change of the surface accessible area, and overall its contribution to the total binding energy is small. The van der Waals interaction has a significant contribution to the binding, which is consistent with the case of peptides binding to SWCNTs.

The effect of the peptide length is also explored. If the energetic contributions of pep4a is compared to that of the pep4, it is found that the binding free energy of pep4a to the SWCNT is much higher than that of pep4, indicating a relatively weaker binding affinity. In addition, the major loss is the van der Waals interaction, which supports the point that the van der Waals interaction dominates the binding process. For example, ΔE_{vdw} is equal to -49.6 $kcal/mol$ for pep4a, which contains four residues of Leu, while the value of ΔE_{vdw} decrease to -106.5 $kcal/mol$ as the number of Leu increases to eight for pep4. A longer peptide leads to a stronger van der Waals interaction between the peptide and the SWCNT.

3.3.4 The influence of hydrophobicities of amino acids

In order to clarify the affinity of different amino acids for SWCNTs, intensive simulations have been carried out on twenty amino acids individually based on MM-GBSA method. Through this a general picture of relative binding strength of hydrophobic and hydrophilic amino acids to SWCNT can also be obtained.

The peptides are replaced with single amino acid, and approximately the same implementation method and procedure is used. After the simulation of the amino acid, the SWCNT, and the amino acid-SWCNT interaction in the explicit solvent,

respectively, the MM-GBSA method is performed to calculate the binding free energy for 10 ps of equilibration and 20 ps of data collection period. The binding free energy is then obtained as the energy difference between the complex and the individual systems.

$$\Delta G = G^s_{complex} - (G^s_{cnt} + G^s_{aa}) \tag{3.17}$$

The stability of both energetic and structural trajectories throughout the MD simulation in explicit solvent is analyzed before the free energy is evaluated. In the stage of free energy calculation using MM-GBSA method, the average free energy and their standard deviations during the data collection period are provided in Table 3.9. The standard errors are also analyzed to trace the stability of energy perturbation.

The calculation results indicate that the binding free energies vary with different amino acids. In order to qualitatively derive the interaction between hydrophobic and hydrophilic the occurrence of each amino acid in proteins w_i (Doolittle, 1989) and their estimated free energy ΔG_i are utilized to evaluate the average binding free energy for both the hydrophobic group and the hydrophilic group. w_i is listed in Table 1.1.

According to the K-D hydropathy scale (Kyte and Doolittle, 1982), each amino acid has been assigned a value indicating its relative hydrophilicity and hydrophobicity. The twenty amino acids are classified into two groups based on this scale. The amino acids with positive values (including Ile, Val, Leu, Phe, Cys, Met, Ala) are considered hydrophobic (H) and the others with negative values are classified

into hydrophilic (P) groups. For hydrophobic group, the average binding free energy is $\overline{\Delta G_H} = \sum_1^{n_h} \Delta G_i w_i / \sum_1^{n_h} w_i = -5.63 kcal/mol$, and $\overline{\Delta G_P} = \sum_1^{n_p} \Delta G_i w_i / \sum_1^{n_p} w_i = -3.97 kcal/mol$ for hydrophilic group. Where $n_h = 7$ and $n_p = 13$, and values of w_i are listed in Table 1.1. These results imply that the both hydrophobic and hydrophilic amino acids have affinities for CNTs and generally the binding affinity of hydrophobic amino acid is stronger. In addition, experimental studies have shown that both hydrophobic and hydrophilic peptides may spontaneously bind to CNTs. On the other hand, it is also observed that hydrophobic peptides indeed have stronger affinities for CNTs than hydrophilic ones (Wang et al., 2003). Therefore our simulation results also agree with that of experiments.

3.3.5 Impact of the aromatic ring

It is also inferred from Table 3.9 that amino acids bearing aromatic rings have relatively stronger binding free energies for SWCNT, for example Trp and Phe. Consistently, it is observed that for pep18 through pep22, those with lower values of average hydrophobicities may show high affinities for SWCNTs, when they contain aromatic rings. For a peptide, hydrophobicity properties on its own can not determine its affinity for SWCNT. Aromatic rings also make contributions here, concluded both from energetic analysis and conformational observation. The stacked structure of aromatic rings may decrease the hydrophobic surface exposed to solvent. The van der Waals interaction is also stronger for such structure.

3.4 Remarks

In this Chapter, free energies of peptides interacting with SWCNTs are calculated based on combined simulation methods of MD and continuum solvation model. The interaction free energies take into accounts both contributions of the solute and solvent. Both binding of peptides onto the outersurface of the SWCNT and encapsulation of peptides into the SWCNT are studied. The calculation results of binding free energies are proven to be satisfactory compared with experimental results of binding affinities of different peptides for SWCNTs.

The energy contributions are also analyzed, it is found that noncovalent bond interactions dominate this binding process, among which the van der Waals interaction appears to be the most significant contributor. Aromatic rings on peptides have strong affinities for CNT surface, which is also driven by the van der Waals interaction.

Further discussion implies that generally hydrophobic amino acids possess stronger affinity for SWCNTs than hydrophilic ones, while other effects such as aromatic rings on the amino acid also have impacts on the interaction.

Despite the difficulties on estimation of some terms contributing to the total free energy, such as entropic change of the solute, the method I adopt is a good estimation of the binding affinities between peptides and CNTs. The method is also applicable for calculation of the interaction energies between other non-biological materials and biological materials.

Table 3.1 Sequences of five 12-residue peptides, as well as their average hydrophobicity. The hydrophobicity values of amino acid residues are calculated using the K-D method.

Assigned peptide name	Peptide sequence	Average hydrophobicity
pep18	HWKHPWGAWDTL	-1.067
pep19	HWKHPSGAWDTL	-1.058
pep20	HWSAWWIRSNQS	-1.083
pep21	HWSAWSIRSNQS	-1.075
pep22	LPPSNASVADYS	-0.192

Table 3.2 The properties of simulated peptides. For hydropathy distributions, each amino acid on the peptide is indicated as either ‘H’ (hydrophobic) or ‘P’ (polar), according to K-D method.

Assigned peptide name	hydropathy distributions	Number of residues containing aromatic rings
pep18	PPPPPPPHPPPH	3
pep19	PPPPPPPHPPPH	2
pep20	PPPHPPHPPPPP	3
pep21	PPPHPPHPPPPP	2
pep22	HPPPPHPHHPPP	1

Table 3.3 The average values of potential energies and their standard deviations over the last 500ps for simulated systems solvated in explicit TIP3P water molecules.

Potential energies of the systems (kcal/mol)	Mean	Std dev
Pep18	-31480.0547	50.2578
Pep8-SWCNT	-31424.5957	54.0284
Pep19	-25984.2227	52.9747
Pep19-SWCNT	-29909.0117	50.4474
Pep20	-24530.0176	37.2812
Pep20-SWCNT	-26704.5547	48.2280
Pep21	-23029.5352	41.4693
Pep21-SWCNT	-28454.0020	52.7485
Pep22	-19882.0176	41.2324
Pep22-SWCNT	-41072.1289	61.7204
SWCNT	-11439.8437	30.5320

Table 3.4 (a)-(e) The energy contributions of the five peptides binding to SWCNTs, and the standard deviations of the energy terms.

Contributions (kcal/mol)	Pep18-SWCNT complex		Pep18		SWCNT	
	mean	Std dev	mean	Std dev	mean	Std dev
E_{internal}	1638.3746	12.2537	276.6918	8.2177	1367.0123	9.0915
E_{ele}	-12.1162	60.7432	7.3884	66.93	0	0
E_{vdw}	242.9762	7.5688	-23.9423	4.6645	309.2535	4.8121
E_{vac}	1869.2349	62.0499	260.1379	67.0886	1676.2667	8.3616
G_{pol}	-394.1432	60.7556	-425.8539	65.7464	0	0
G_{nonpol}	9.4334	0.2452	6.4378	0.284	3.8638	0.0123
G_{sol}	-384.7097	60.7575	-419.4159	65.6952	3.8638	0.0123
E_{ele_total}	-406.2593	4.2575	-418.4655	4.4762	0	0
G	1484.525	11.2766	-159.2781	7.1467	1680.1296	8.3621

(a)

Contributions (kcal/mol)	Pep19-SWCNT complex		Pep19		SWCNT	
	mean	Std dev	mean	Std dev	mean	Std dev
E_{internal}	1631.5192	12.0961	272.147	8.1537	1367.0123	9.0915
E_{ele}	142.6086	28.6066	170.8088	76.1325	0	0
E_{vdw}	250.7564	4.3113	-25.6338	4.6942	309.2535	4.8121
E_{vac}	2024.8838	31.429	417.322	77.4235	1676.2667	8.3616
G_{pol}	-557.3611	28.412	-602.6929	75.0763	0	0
G_{nonpol}	9.9994	0.0369	5.7827	0.3024	3.8638	0.0123
G_{sol}	-547.3617	28.4117	-596.9105	74.9688	3.8638	0.0123
E_{ele_total}	-414.7525	2.2455	-431.8842	4.6101	0	0
G	1477.5217	11.5573	-179.5883	7.5352	1680.1296	8.3621

(b)

Contributions (kcal/mol)	Pep20-SWCNT complex		Pep20		SWCNT	
	mean	Std dev	mean	Std dev	mean	Std dev
E_{internal}	1638.374	12.8965	270.0867	10.5467	1367.0123	9.0915
E_{ele}	-345.1888	76.0889	-246.2108	91.7371	0	0
E_{vdw}	244.1539	6.3216	-23.1936	5.4034	309.2535	4.8121
E_{vac}	1537.3383	75.3554	0.6823	91.2897	1676.2667	8.3616
G_{pol}	-319.562	74.7975	-417.1699	89.6081	0	0
G_{nonpol}	10.1645	0.1696	6.9329	0.3363	3.8638	0.0123
G_{sol}	-309.3974	74.8006	-410.2368	89.6046	3.8638	0.0123
E_{ele_total}	-664.7509	5.5092	-663.3806	6.1703	0	0
G	1227.9415	11.772	-409.5546	7.8099	1680.1296	8.3621

(c)

Contributions (kcal/mol)	Pep21-SWCNT complex		Pep21		SWCNT	
	mean	Std dev	mean	Std dev	mean	Std dev
E_{internal}	1624.3468	13.0358	250.7221	8.4374	1367.0123	9.0915
E_{ele}	-409.3298	86.1343	-526.4892	96.0464	0	0
E_{vdw}	248.9652	6.7611	-28.4102	5.2523	309.2535	4.8121
E_{vac}	1463.9817	88.3124	-304.1774	96.097	1676.2667	8.3616
G_{pol}	-259.5987	85.1361	-144.1688	93.6886	0	0
G_{nonpol}	9.2655	0.1538	5.8317	0.2158	3.8638	0.0123
G_{sol}	-250.3332	85.1506	-138.3371	93.6964	3.8638	0.0123
E_{ele_total}	-668.9283	5.2098	-670.658	5.8716	0	0
G	1213.6492	11.9748	-442.5143	7.44	1680.1296	8.3621

(d)

Contributions (kcal/mol)	Pep22-SWCNT complex		Pep22		SWCNT	
	mean	Std dev	mean	Std dev	mean	Std dev
E_{internal}	1582.9836	12.3242	219.3674	7.7071	1367.0123	9.0915
E_{ele}	-198.8516	64.7461	-293.2146	72.3533	0	0
E_{vdw}	267.3004	7.1236	-7.6251	4.8981	309.2535	4.8121
E_{vac}	1651.432	65.2193	-81.4723	72.9131	1676.2667	8.3616
G_{pol}	-268.2461	63.3083	-194.4066	71.0038	0	0
G_{nonpol}	9.0397	0.15	6.0703	0.1855	3.8638	0.0123
G_{sol}	-259.2063	63.2931	-188.3364	70.9858	3.8638	0.0123
E_{ele_total}	-467.0977	4.1893	-4 87.621	4.0205	0	0
G	1392.226	10.6293	-269.8085	6.3475	1680.1296	8.3621

(e)

Table 3.5 The comparison of energy contributions of peptides binding to SWCNTs.

Contributions(kcal/mol)	pep18	pep19	pep20	pep21	pep22
$\Delta E_{\text{internal}}$	-5.3295	-7.6401	1.275	6.6124	-3.3961
ΔE_{ele}	-19.5046	-28.2002	-98.978	117.1594	94.363
ΔE_{vdw}	-42.335	-32.8633	-41.906	-31.8781	-34.328
ΔE_{vac}	-67.1697	-68.7049	-139.611	91.8924	56.6376
ΔG_{pol}	31.7107	45.3318	97.6079	-115.43	-73.8395
ΔG_{nonpol}	-0.8682	0.3529	-0.6322	-0.43	-0.8944
ΔG_{sol}	30.8424	45.685	96.9756	-115.86	-74.7337
ΔE_{ele_total}	12.2062	17.1317	-1.3703	1.7297	20.5233
ΔG	-36.3265	-23.0196	-42.6335	-23.9661	-18.0951

Table 3.6 Relative binding free energies between pep18 and pep19, and pep20 and pep21. $\Delta\Delta G$ of pep18-pep19 is calculated as $\Delta\Delta G = \Delta G(pep18 - SWCNT) - \Delta G(pep19 - SWCNT)$, and the same with other energy contributions and that of pep20-pep21.

Contributions(kcal/mol)	pep18-pep19	pep20-pep21
$\Delta\Delta E_{\text{internal}}$	2.3106	-5.2929
$\Delta\Delta E_{ele}$	8.6956	-206.4158
$\Delta\Delta E_{vdw}$	-9.4717	-10.2413
$\Delta\Delta E_{vac}$	1.5352	-221.95
$\Delta\Delta G_{pol}$	-13.6211	203.9839
$\Delta\Delta G_{nonpol}$	-1.2211	-0.2019
$\Delta\Delta G_{sol}$	-14.8426	203.782
$\Delta\Delta E_{ele\_toltal}$	-4.9255	-2.4319
$\Delta\Delta G$	-13.3069	-18.168

Table 3.7 (a)-(c) The energy contributions of the three peptides inserting into to SWCNTs, and their standard deviations of the energy terms, respectively.

Contributions (kcal/mol)	Pep3-SWCNT complex		Pep3		SWCNT	
	mean	Std dev	mean	Std dev	mean	Std dev
E_{internal}	1191.5861	15.3248	169.0035	6.5918	1019.4637	12.4125
E_{ele}	-204.6609	51.6702	-148.585	67.1861	0	0
E_{vdw}	495.4647	7.6932	-3.8899	3.9863	583.4814	4.3773
E_{vac}	1482.3896	53.4126	16.5285	67.8999	1602.9449	11.6556
G_{pol}	-202.7411	51.0705	-276.8427	67.0532	0	0
G_{nonpol}	10.8287	0.1969	5.4019	0.1976	7.9816	0.026
G_{sol}	-191.9124	51.0292	-271.4408	66.9631	7.9816	0.026
E_{ele_total}	-407.4019	3.3831	-425.4276	3.627	0	0
G	1290.4775	11.2524	-254.9122	7.8562	1610.9267	11.2014

(a)

Contributions (kcal/mol)	Pep4a-SWCNT complex		Pep4a		SWCNT	
	mean	Std dev	mean	Std dev	mean	Std dev
E_{internal}	1187.9714	12.5567	78.6398	5.3064	1110.8688	11.0425
E_{ele}	-179.9952	26.3942	-104.3704	25.3194	0	0
E_{vdw}	509.0128	7.0548	5.1788	2.5071	553.4804	4.4837
E_{vac}	1516.9889	32.071	-20.5519	26.0232	1664.3497	10.6876
G_{pol}	-66.8007	25.8641	-149.3291	25.3112	0	0
G_{nonpol}	8.5188	0.1875	3.4415	0.0816	7.0454	0.0201
G_{sol}	-58.2819	25.8924	-145.8875	25.294	7.0454	0.0201
E_{ele_total}	-246.7959	3.3896	-253.6995	1.8106	0	0
G	1458.7071	105946	-166.4394	6.4258	1671.3946	10.5351

(b)

Contributions (kcal/mol)	Pep4-SWCNT complex		Pep4		SWCNT	
	mean	Std dev	mean	Std dev	mean	Std dev
E_{internal}	1286.4398	15.0647	162.9834	6.4882	1110.8688	11.0425
E_{ele}	-227.7805	37.5617	-143.454	58.2402	0	0
E_{vdw}	450.0414	9.4405	3.0397	3.5016	553.4804	4.4837
E_{vac}	1508.7008	39.9636	22.5692	58.5838	1664.3497	10.6876
G_{pol}	-132.3807	36.4556	-239.3704	57.5327	0	0
G_{nonpol}	10.0537	0.1912	5.8221	0.1372	7.0454	0.0201
G_{sol}	-122.327	36.4765	-233.5482	57.5546	7.0454	0.0201
E_{ele_total}	-360.1612	3.2885	-382.8243	3.0567	0	0
G	1386.3737	10.3652	-210.9792	7.5981	1671.3946	10.5351

(c)

Table 3.8 The comparison of energy contributions of peptides inserting into SWCNTs.

Contributions(kcal/mol)	pep3	pep4a	pep4
$\Delta E_{\text{internal}}$	3.1189	-1.5372	12.5876
ΔE_{ele}	-56.0759	-75.6248	-84.3265
ΔE_{vdw}	-84.1268	-49.6464	-106.4787
ΔE_{vac}	-137.0838	-126.8089	-178.2181
ΔG_{pol}	74.1016	82.5284	106.9897
ΔG_{nonpol}	-2.5548	-1.9681	-2.8138
ΔG_{sol}	71.5468	80.5602	104.1758
ΔE_{ele_total}	18.0257	6.9036	22.6631
ΔG	-65.537	-46.2481	-74.0417

Table 3.9 Binding free energies and the standard deviations estimated using MM-GBSA method. The energy unit in this table is kcal/mol. The free energy of the SWCNT for all the twenty systems is $G^s_{cnt} = 1653.5895$ kcal/mol and the standard deviation is 8.7468 kcal/mol. The binding free energy is estimated as $\Delta G = G^s_{complex} - (G^s_{cnt} + G^s_{aa})$. The lower value of ΔG correlates to a stronger binding affinity.

Amino acid	The amino acid-SWCNT		Amino Acid		ΔG
	mean	Std dev	mean	Std dev	
Ile	1584.591	8.2012	-62.4168	1.7384	-6.5813
Val	1559.739	9.7484	-89.0848	1.7052	-4.7658
Leu	1565.049	10.427	-81.4391	1.8871	-7.1018
Phe	1598.538	8.5823	-48.4989	1.7332	-6.5523
Cvs	1596.942	9.0752	-53.5529	1.429	-3.0945
Met	1586.07	9.1196	-59.2105	1.9881	-8.3093
Ala	1596.049	8.4567	-54.1631	1.7447	-3.3771
Gly	1573.276	10.1111	-76.9428	1.0023	-3.3711
Thr	1552.899	9.1481	-97.6639	2.1033	-3.0267
Ser	1570.769	8.7619	-80.4642	1.4295	-2.3565
Trp	1601.97	9.2565	-43.6792	2.9699	-7.9406
Tyr	1571.817	9.3792	-75.6519	1.8876	-6.1211
Pro	1594.68	8.6813	-55.662	1.8604	-3.2473
His	1605.269	9.9353	-42.3449	2.2036	-5.9758
Glu	1623.309	8.8987	-25.4096	1.9013	-4.8712
Asn	1515.04	7.7544	-133.307	1.5842	-5.243
Gln	1536.462	8.6544	-114.396	1.9972	-2.7319
Asp	1613.098	10.0269	-34.2139	1.5893	-6.2779
Lys	1570.113	10.2352	-81.9991	2.2068	-1.4772
Arg	1400.151	8.9225	-248.73	2.5351	-4.7084

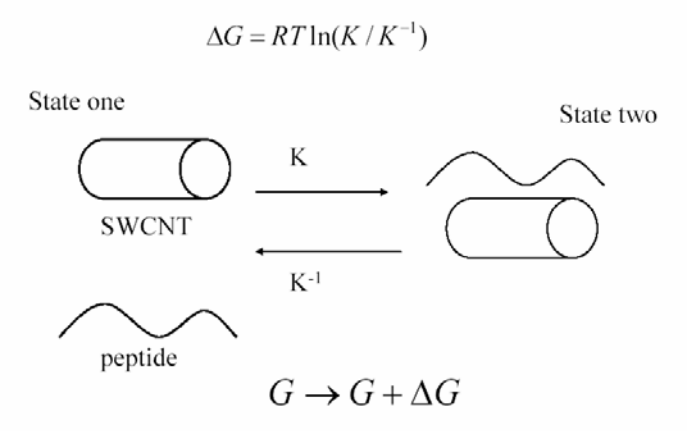
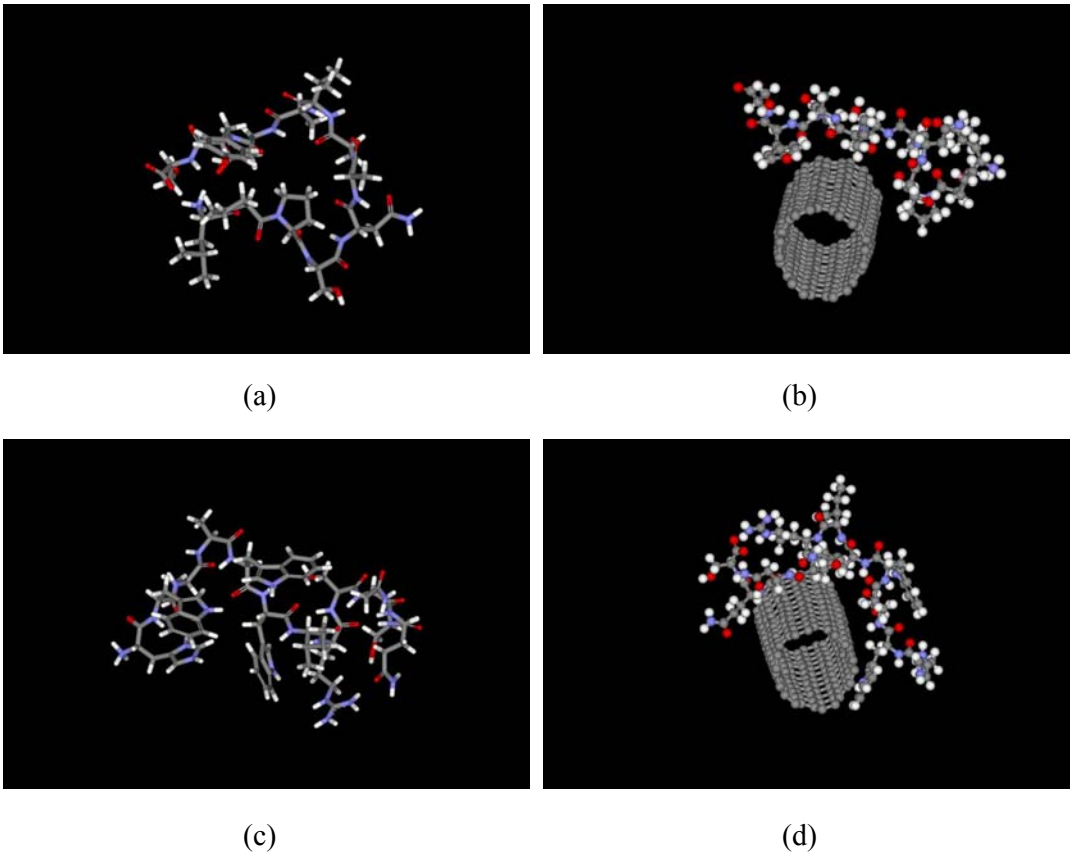


Figure 3.1 The strategy of estimating interaction free energy between two states.



(a) (b)

(c) (d)

Figure 3.2 Snapshots of final structures of peptides and peptide-SWCNT complex in water solvent. (a) pep22 (b) pep22-SWCNT complex (c) pep20 (d) pep20-SWCNT complex. The images are created with DS ViewerPro 5.0 software (Accelrys Inc., San Diego, CA)

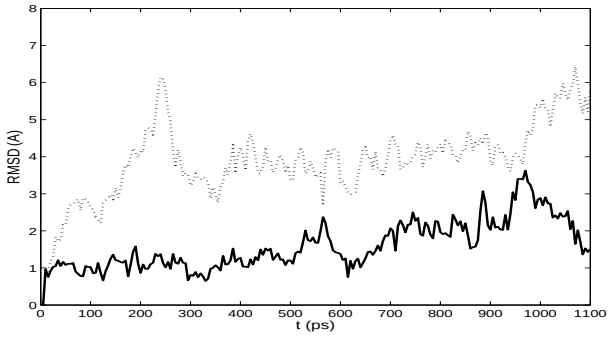


Figure 3.3 The RMSDs for the backbone atoms on pep20. The dotted lines represent the unbound peptide and the solid lines represent the peptide in the complex.

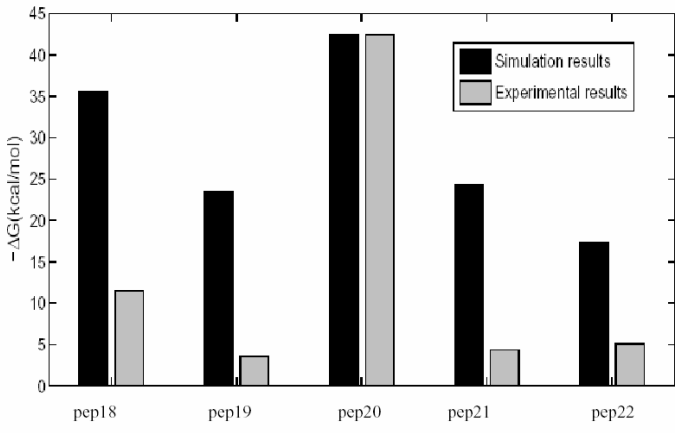


Figure 3.4 The comparison of binding free energies with experimental results. The binding free energies are drawn as their absolute values (kcal/mol). The plaque-forming units from experimental results are scaled linearly in relation to the absolute values of the binding free energy of pep20. Larger ΔG and plaque-forming unit values correspond to higher binding affinities.

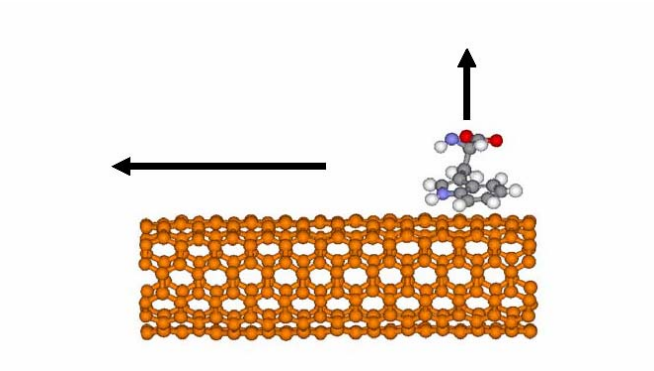
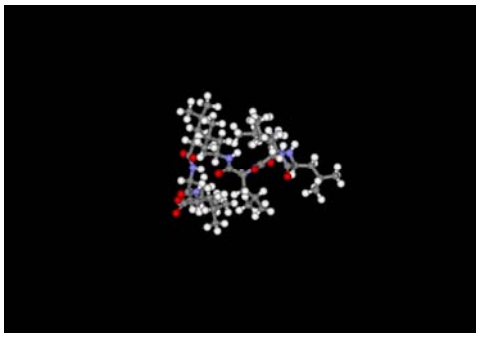
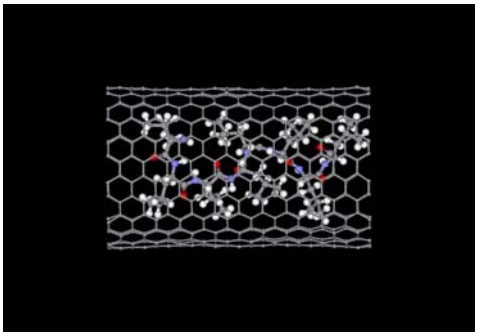
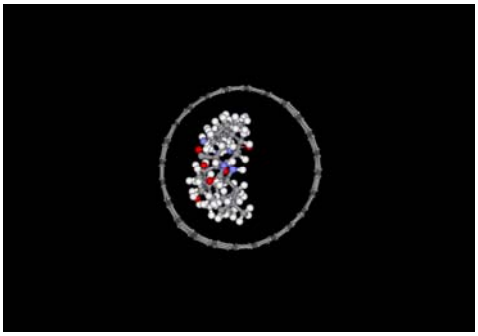


Figure 3.5 The scheme for calculating energy potential of residue Trp on the surface of a SWCNT. The residue containing an aromatic ring is moved along two directions for positioning and energy calculations.



(a)





(b) (c)

Figure 3.6 Snapshots of final structures of peptides and peptide-SWCNT complex in water solvent. (a) pep4 (b) pep4-SWCNT complex (c) side view of pep4-SWCNT complex. The images were created with DS ViewerPro 5.0 software (Accelrys Inc., San Diego, CA)

Chapter 4

Thermodynamic studies based on a hydrophobic-polar (HP) lattice model

Computational simulation has been critical in revealing the essential chemical and physical phenomena that are not accessible to experiments. As shown in Chapter 2 and Chapter 3, I have studied the self-insertion of peptides into single-walled carbon nanotubes (SWCNTs) and binding of peptides to the outer-surface of the SWCNTs based on the all-atom models. The energetic and conformational analysis suggests that the hydrophobicities of peptides correlate to their affinities for CNTs.

While MD simulation of the all-atom models have enabled us to gain great insight into the peptide-CNT interaction mechanism, the time scale it can handle is far less than that of the coarse-grained model. The typical accessible time scale for the all-atom models is up to nanoseconds, while the coarse-grained model can simulate the reaction process with the time scale from microseconds to seconds. In order to obtain a clearer picture of the entire interaction process between peptides and CNTs, investigating coarse-grained model seems also essential. The advantage of the HP lattice model is that it is capable of capturing most essential mechanism of protein folding such as hydrophobic effect and multi-stage folding kinetics (Lau and Dill, 1989). Furthermore, the full conformational spaces can be enumerated exhaustively

and insights into the nature of free energy landscapes can be obtained. Therefore the model has been successfully applied to explore kinetics and thermodynamics of protein folding in bulk solvent and adsorption onto various surfaces.

In this chapter, I extend the previous work to study the mechanism of peptides interacting with CNTs using HP lattice model and MC simulation method. In Section 4.1, the simulation models and methods are presented. Exact enumeration of all the possible conformations of peptides has been carried out on 2D lattices with the chain length of 16 residues. In order to simulate the peptide-CNT interactions, a model wall with an energetically favorable potential to peptide beads is introduced. In Section 4.2, thermodynamics and conformational changes of peptides folding in bulk water and interacting with CNTs are discussed. The hydrophobicity of the CNT is incorporated into the existing HP lattice paradigm and the interaction parameters between model chain residues and the CNT monomers can be qualitatively determined based on experimental data and simulation results from the all-atom model. Section 4.3 discusses the comparison between MD and MC algorithms, and Section 4.4 gives the remarks of the chapter.

4.1 HP lattice model using Monte Carlo (MC) simulation methods

4.1.1 2D HP lattice model for modeling peptide-CNT interactions

The HP model was first proposed by Dill and Lau (Lau and Dill, 1989). The assumption of the model is that the hydrophobic interaction is the dominant force in

protein folding. Each residue in the protein sequence is represented by either of the two types, hydrophobic (H) or polar (P).

The simplified HP lattice model has been found useful for modeling protein properties because it shows several protein-like features, including cooperative collapse, native structures having a nonpolar core and definable secondary structures, multi-stage folding kinetics, and molten globule states. Most importantly, the model captures the main physics of protein folding—the hydrophobic interactions, conformational freedom of the chain, and the steric restrictions imposed by excluded volume (Miller and Dill, 2006). For short chains, all the possible conformations can be enumerated and their energies are evaluated to find the global energetic minimum for each HP sequence. The thermodynamic properties of the system could be estimated through statistical mechanics. Therefore the model has successfully provided insight into the kinetics and thermodynamics for protein folding in bulk solvent or adsorption of protein to different surfaces.

In HP lattice model, chains are configured as self-avoiding walks on 2D square lattices or 3D simple cubic lattices. While the HP model is most intuitively defined in 3D to match the physical phenomenon, a 2D model is actually more suitable for currently computationally feasible sizes. For example, in order to represent the appropriate surface-interior ratios of protein molecule, 16-20 monomers in two dimensions can reproduce a chain of 154 monomers in three dimensions (Chan and Dill, 1993). The latter case is obviously beyond the scope of exact enumeration. A chain with 27 monomers is a feasible size in three dimensions, but unfortunately

there's only one interior residue in a maximally compact conformation. Therefore a 2D lattice model is a feasible model allowing exact enumeration of all the possible configurations. In some recent work, the researchers also employed the 2D 16-monermer of HP lattice model and revealed important principles on protein folding (Kumar et al., 2005; Chikenji et al., 2006).

Therefore a 2D short chain is a realistic choice considering the number of monomers can be accurately enumerated. It was suggested that the ideal number of monomers for a 2D lattice model is about 16-20 (Chan and Dill, physics today, 1993). In order to obtain an accurate enumerate of all the possible conformations that the peptide undergoes, and to test the performance of the model using different sets of parameters, a 16-monomer system is a reasonable length for obtaining targeted results and within affordable computational time. It was also established that even the 2D 16-monermer of HP lattice model can be used to reveal important principles on protein folding (Kumar et al., 2005; Chikenji et al., 2006). Therefore I simply apply 2D 16-monomer model to represent peptides in this study.

For our 2D lattice model, the structure of a protein is modeled as linked beads. Each protein chain contains $N = 16$ residues, connected through $N - 1$ links. The protein chain is placed on the nodes of a Cartesian coordinate grid. All connecting vectors move parallel to either the x- or the y-axis with a self-avoiding configuration. The lattice spaces which are not occupied by the amino acid residues are assumed to be solvent units.

The presence of a CNT surface is introduced into the simulation by modeling a straight wall on the lattice space. Each bead along the wall represents a segment of the CNT. I refer to such beads as type "C". The wall is regarded as a rigid boundary, where no residue can go through it.

The potential energy of a lattice chain containing N monomers is defined as the sum of the interaction energies between all the contact pairs of the system. In the case of a peptide chain solvated in bulk water, the two elements which are adjacent in coordinate but not directly connected are defined as one contact pair. If both elements are hydrophobic, it's called a native contact. For each contact pair (i, j), a corresponding variable $\gamma_S(A_i, A_j)$ is counted to the interaction potential energy between the two amino acids residues. The parameters $\gamma_S(H,H)$, $\gamma_S(H,P)$, and $\gamma_S(P,P)$ are determined according to the relative affinities between these element types. The energy parameters $\gamma_S(H,H) = -2\varepsilon$, $\gamma_S(H,P) = -\varepsilon$, and $\gamma_S(P,P) = 0$ are adopted in this model. In which $\varepsilon > 0$ and 2ε is the free energy required to unfold one "H-H" contact. The advantage of this set of parameter is that compact conformations of peptides in water can be obtained (Li et al., 1996).

The interaction energies between the CNT and the peptide residues are only counted if the residues are in direct contact with the monomers of type "C". In other cases, the interaction energy is zero. The energy parameters $\gamma_S(H,C)$ and $\gamma_S(P,C)$ are discussed in section 4.2.2.1 based on experimental and MD simulation results. Therefore the total energy (E) of a system is calculated as $E = \sum_{(i,j)}^{N_K} \gamma_S(A_i, A_j)$, where

N_k denotes the total number of contact pairs in a certain structure, including those between pairs of residues in the peptide and also the residues and the CNT.

4.1.2 MC simulation of peptide-CNT interactions

The MC method is a stochastic algorithm that depends on probabilities. The general procedure of performing MC method is achieved by: (a) generating a random trial configuration; (b) evaluating an 'acceptance criterion' by calculating the change in energy and other properties in the trial configuration; (c) comparing the acceptance criterion to a random number and either accepting or rejecting the trial configuration. Note that all the generated states will make a significant contribution to the configurational properties of the system. Therefore it is essential to sample those states that make the most significant contributions in order to accurately determine the properties of the system within the finite simulation time. A Markov chain is often generated to evaluate these important states.

A Markov chain is a sequence of trials in which the outcome of successive states only depends on the immediate predecessor. In a Markov chain, a new state will only be accepted if it is more favorable than the existing state. Specifically, the Metropolis algorithm (Metropolis, 1953) is implemented and ensemble sampling is carried out to calculate thermodynamic quantities. During the molecular simulation, the more favorable state usually correlates to the new trial state that is lower in energy.

4.1.2.1 Random number generators

The random number generator is essential to every MC simulation program, and is accessed a very large number of times. The random number generator is used not only to generate new configurations but also to decide whether a given move should be accepted or not. Actually, the numbers produced by a random number generator are not in fact truly random. The same sequence of numbers should always be generated when the program is run with the same initial conditions. The sequences of numbers are therefore often referred to as 'pseudo-random' numbers as they possess the statistical properties of true sequences of random numbers of a different 'seed' is provided. In this way, several independent runs can be carried out using different seeds.

The random numbers produced by a random number generator should satisfy certain criteria. A good random number generator should yield a random distribution of points while being computationally fast. The linear congruential method is used in this study for generating random numbers. Each number in the sequence is generated by taking the previous number, multiplying by a constant a, adding a second constant b, and taking the remainder when divided by a third constant (the modulus, m). The initial value is the seed, supplied by the user. The algorithm is written as:

$$\xi[1] = seed \tag{4.1}$$

$$\xi[i] = MOD\left\{\left(\xi[i-1]\times a + b\right), m\right\} \tag{4.2}$$

Where the MOD function returns the remainder when the first argument is divided by the second.

4.1.2.2 Implementation of the Metropolis algorithm

During the dynamic MC simulation of the lattice model, the peptide undergoes three types of allowable Verdier–Stockmeyer moves (crankshaft, flip, and turn) (Verdir and Stockmayer, 1962), as illustrated in Figure 4.1. The MC simulation of the lattice structures obeys the Metropolis algorithm (Metropolis, 1953). The Metropolis algorithm generates a Markov chain of states, suppose that the system is in a state m. It is assumed that the probability of moving to state n is π_{mn}. The various π_{mn} can be considered to constitute an $N \times N$ matrix $\boldsymbol{\pi}$, called the transition matrix, where N is the number of possible states. Each row of the transition matrix sums to 1. Therefore the probability that the system lies in a particular state is represented by a probability vector $\boldsymbol{\rho}$:

$$\boldsymbol{\rho} = (\rho_1, \rho_2, ..., \rho_m, \rho_n ..., \rho_N) \tag{4.3}$$

Here ρ_1 is the probability that the system is in state 1 and ρ_m represents the probability that the system is in state m. If $\rho(1)$ represents the initial configuration which is randomly chosen, the probability of the second state is given by equation:

$$\boldsymbol{\rho}(2) = \boldsymbol{\rho}(1)\pi \tag{4.4}$$

The probability for the system in the third state is:

$$\boldsymbol{\rho}(3) = \boldsymbol{\rho}(2)\pi = \boldsymbol{\rho}(1)\pi\pi \tag{4.5}$$

The equilibrium distribution of the system can be determined by considering the result of applying the transition matrix an infinite number of times. This limiting distribution of the Markov chain is given by $\boldsymbol{\rho}_{\text{limit}} = \lim_{N\to\infty} \boldsymbol{\rho}(1)\pi^N$.

Closely related to the transition matrix $\boldsymbol{\pi}$ is the stochastic matrix, whose elements are labeled α_{mn}. This matrix shows the probability of choosing the two states m and n between which the move is to be made. It is also referred to as the underlying matrix of the Markov chain. If the probability of accepting a trial move from m to n is p_{mn}, then the probability of making a transition from m to n (π_{mn}) is calculated by multiplying the probability of choosing states m and n (α_{mn}) by the probability of accepting the trial move (p_{mn}):

$$\pi_{mn} = \alpha_{mn} p_{mn} \tag{4.6}$$

It is assumed that the stochastic matrix $\boldsymbol{\alpha}$ is symmetrical. If the probability of state n is greater than that of state m in the limiting distribution, for example, if the Boltzmann factor of n is greater than that of state m because the potential energy of state n is lower than the energy of state m, then the transition matrix element π_{mn} for progressing from m to n equals the probability of selecting the two states in the first place. If the Boltzmann weight of the state n is less than that of state m, then the probability of permitting the transition is obtained by multiplying the stochastic matrix element α_{mn} by the ratio of the probabilities of the state n to the previous state m. The algorithm is expressed as:

$$\pi_{mn} = \alpha_{mn} \qquad (\rho_n \geq \rho_m) \tag{4.7}$$

$$\pi_{mn} = \alpha_{mn}(\rho_n / \rho_m) \qquad (\rho_n < \rho_m) \tag{4.8}$$

These two conditions only apply when the states m and n are different. If m and n are the same state, then the transition matrix element is calculated from the truth that the rows of the stochastic matrix sum up to 1:

$$\pi_{mm} = 1 - \sum_{m \neq n} \pi_{mn} \tag{4.9}$$

In order to apply the Metropolis algorithm to simulate the dynamics of lattice peptide, at each MC cycle, a single movement is randomly selected for the peptide, and the energy of the newly generated system is evaluated. According to the Metropolis algorithm, if the new system energy is found to be equal or less than that of the original conformation, the configuration is updated and a new cycle starts. If the new conformation results in an increase in system energy, however, the move with a probability is chosen according to Equation (4.6). This is achieved by comparing the Boltzmann factor $P = \exp(-\Delta E(\mathbf{r}^N) / k_B T)$ (where $\Delta E(\mathbf{r}^N) = E_n - E_m$) to a random number r between 0 and 1. The quantity k_B is the Boltzmann constant, taken as the dimensionless unit with $k_B = 1$, and T is the temperature. $k_B T$ has the same dimension as ε. If $P > r$, the new configuration is accepted; otherwise, the previous configuration is retained.

In this way, if the energy of new state E_n is very close to that of the old state E_m , then the Boltzmann factor of their energy difference is very close to 1, and the move is likely to be accepted. However, if the energy difference is very large, then the Boltzmann factor tends to be zero and the move is unlikely to be accepted.

The Metropolis algorithm is derived based on the condition of microscopic reversibility, which indicates that at equilibrium, the transition between two states occurs at the same rate. The rate of transition from a state m to another state n equals the product of the population (ρ_m) and the appropriate element of the transition matrix π_{mn}. Therefore at equilibrium I have

$$\pi_{mn} / \pi_{nm} = \exp(-\Delta E(\mathbf{r}^N) / k_B T) \tag{4.10}$$

4.1.3 Molecular Simulation of Ensembles

The original MC method sampled the canonical (NVT) ensemble, and it could also be extended to other ensembles by using importance sampling for Markov chain process. Here the canonical ensemble is sampled to calculate the thermodynamic quantities from the calculated microscopic properties.

The average of any thermodynamic property $\left\langle A(\mathbf{r}^N) \right\rangle$ can be calculated by evaluating the following multidimensional integral over the degrees of freedom on the N particles in the system

$$\left\langle A(\mathbf{r}^N) \right\rangle = \int A(\mathbf{r}^N) \rho(\mathbf{r}^N) dr^N \tag{4.11}$$

where $\rho(\mathbf{r}^N)$ is the probability of obtaining configuration $\mathbf{r}^N$ which depends on the potential energy (E) of the system.

$$\left\langle \rho(\mathbf{r}^N) \right\rangle = \frac{\exp[-\beta E(\mathbf{r}^N)]}{\int \exp[-\beta E(\mathbf{r}^N)] d\mathbf{r}^N} \tag{4.12}$$

Where $\beta = 1/(k_B T)$. These integrals can not be evaluated analytically, and therefore the MC solution is utilized to generate a large number of trial configurations

$\mathbf{r}^N$ and replace the integrations by summations over a finite number of configurations. If the configurations of the system are chosen randomly, Equation (4.11) is expressed as

$$< A >= \frac{\sum_{i=1}^{N_{trial}} A_i(\mathbf{r}^N) \exp[-\beta E_i(\mathbf{r}^N)]}{\sum_{i=1}^{N_{trial}} \exp[-\beta E_i(\mathbf{r}^N)]} \tag{4.13}$$

However, this simple approach is not feasible in practice because random sampling may yield many configurations having a very small Boltzmann factor. Such configurations make very little contribution to the average. Therefore a large number of configurations should be generated and evaluated to obtain the accurate answer.

The Metropolis sampling biases the generation of configurations towards those that make the most significant contribution to the integral, as introduced in the previous section. The limitations of random sampling can be avoided by generating configurations that make a large contribution to the right hand side of Equation (4.13).

The MC method can also be extended to other ensembles because importance sampling can be used for any Markov chain process. The canonical average of any function of particle coordinates <A> can be obtained for Equation (4.13). In principle, <A> could be calculated through generating a large number of configurations (M) of particles, which is estimated by replacing the integrals with finite sums.

$$< A >= \frac{\sum_{i=1}^{M} A(i) \exp[-\beta E(i)]}{\sum_{i=1}^{M} \exp[-\beta E(i)]} \tag{4.14}$$

It is noted that the system should come to equilibrium state before contributions to the ensemble average are accumulated. The equilibration period is important for obtaining reliable ensemble averages because the initial state my not accurately reflect the system configuration or properties at equilibrium. Therefore the equilibration period is used before calculation of ensemble averaged properties.

In MC simulations, particle momenta are not involved, which is different from MD simulation method. It is unnecessary to know the particle momenta to calculate thermodynamic properties of the system because the momenta contribute exclusively to the ideal gas term. Deviations from ideal gas behavior are caused by interactions between particles which can be calculated from a potential energy function. The potential energy depends only on the positions of atoms or particles but not their momenta.

4.1.4 Calculations of thermodynamics for peptide-CNT binding process

The thermodynamic quantities of the model peptides binding to CNTs are thus calculated based on canonical ensemble. The partition function of the system, Z, is expressed by means of sampling all the possible configurations of the peptide chain in each energy state,

$$Z = \sum_{x} \exp[-E(x)/k_B T] = \sum_{E_i} \Omega(E_i) \exp(-E_i/k_B T), \tag{4.15}$$

where $\Omega(E_i)$ is the density of states with energy E_i. In this case, E_i is determined by both the "H-H" and "H-P" contact interactions within the lattice chain

and the neighboring "H-C" and "P-C" interactions between the chain residues and the "C" type monomers.

The probability (ρ_M) that the system is in its native (lowest-accessible energy) state E_M is expressed as

$$\rho_M(T) = \frac{\exp(-E_M / k_B T)\Omega(E_M)}{Z}. \tag{4.16}$$

It should be noticed that, the lowest-accessible energy state of the system may be degenerated. The peptide may adopt several different conformations in the lowest-accessible energy state. The probability that the system is in the denatured (non-lowest energy) states is

$$\rho_U = 1 - \rho_M. \tag{4.17}$$

For a system at equilibrium, the probabilities ρ_M and ρ_U are connected by the equilibrium constant K_{MU},

$$K_{MU} = \frac{[U]}{[M]} = \frac{\rho_U}{\rho_M}. \tag{4.18}$$

where $[M]$ and $[U]$ are the concentrations of the systems, which correspond to the lowest and non-lowest energy states respectively. Furthermore, the free energy change can be calculated as

$$\Delta G_{MU} = G_U - G_M = -k_B T \ln K_{MU} \tag{4.19}$$

The internal energy U is calculated as the ensemble averaged energy E,

$$U = \langle E \rangle = \sum_i \rho_i(E_i) E_i, \tag{4.20}$$

where $\rho_i(E)$ is the probability that the energy of the system is E_i. Different energy states are enumerated and the corresponding probabilities are calculated through the partition function.

$$\rho_i(E_i) = \frac{\exp(-E_i / k_B T)\Omega(E_i)}{Z} \tag{4.21}$$

The Helmholtz free energy of the system, A, could be calculated as

$$A = \langle A \rangle = -k_B T \ln Z \tag{4.22}$$

Since the CNT surface is stationary, the conformational entropy of the model peptide is the entropy of the whole system, which is determined through standard thermodynamic equation,

$$S = \frac{U - A}{T} \tag{4.23}$$

4.2 Results

4.2.1 Thermal unfolding of model peptide

In this section, the thermal unfolding behavior of a randomly selected 16-monomer model peptide in sequence HPPHHHPPHPPHHPPH is investigated. The model peptide (referred to as peptide I) contains eight hydrophobic residues and can fold into a unique native conformation with seven hydrophobic contacts, as shown in Figure 4.2.

During the MC running, the conformation of the peptide undergoes Verdier-Stockmeyer moves. Figure 4.3 reports simulation results of $5*10^6$ iterations

for the ensemble averaged number of native contacts, $< Q >$, in bulk water as the function of the dimensionless temperature, T^*. The temperature T should be in unit of ε / k_B, here I refer to the dimensionless $T^* = T /(\varepsilon / k_B)$ to avoid lengthy presentation. It is observed from the figure that at low temperature, the peptide tends to be fully folded. The average number of hydrophobic contact pairs lies close to seven, which indicates that the peptide folds into its native state. As the temperature is increased, there are fewer native contacts, and the peptide is denatured from its native structure.

The thermal unfolding of peptide I at different temperatures is simulated over $5*10^6$ iterations. Here the dimensionless temperature T^* is referred, which is in unit of ε / k_B. It is observed that at low temperature, the most probable state is the native state, while at high temperature, the denatured conformations are more populated. The denaturation in terms of the unfolding free energy change ΔG_{MU} can be described. For example, as illustrated in Table I, at the temperature $T^* = 0.2$, the corresponding free energy change $\Delta G_{MU} = G_U - G_M >> 0$, indicating a stable native structure. On the other hand, at the temperature $T^* = 1.6$, $\Delta G_{MU} < 0$, which shows that at higher temperature, the probability of finding the model peptide in the native state decreases drastically, and hence a denatured structure is preferred by the model peptide.

In this simulation, the temperature acts as a thermodynamic property which controls the folding behavior of the model peptide chain. The higher temperature induces the unfolding and therefore influences other macroscopic properties of the

peptide. It is assumed that the Gibbs free energy difference between the native and unfolded states is zero at the mid transition temperature. For peptide I, at the temperature $T^* \approx 0.7$, $\rho_M \approx 0.5$, and $\Delta G_{MU} = 0$, a phase transition state where the peptide exists both in folded and unfolded states is observed. This temperature is often defined as the protein folding temperature (Camacho and Thirumalai, 1993). I have tested the dynamic process of peptide folding in bulk solvent and binding to CNTs at different temperatures, including those below and above the folding temperature. Other values of temperatures that may go to extremes are not considered here.

These observations imply that the lattice model is able to reproduce the essential properties of protein folding in bulk water.

4.2.2 Thermodynamics of peptides interacting with CNTs

4.2.2.1 The selection criteria for the interaction energy parameters and the analysis of thermodynamic quantities

In Chapter 3, I have estimated the average binding free energy of hydrophobic and hydrophilic groups of amino acids to the same CNT based on the atomic model. For hydrophobic group, the average binding free energy is $\overline{\Delta G}_H = -5.63 kcal/mol$, and that of the polar group is $\overline{\Delta G}_P = -3.97 kcal/mol$.

These results imply that the both hydrophobic and hydrophilic amino acids have affinities for CNTs and generally the binding of hydrophobic amino acid is stronger. In addition, experimental studies have shown that both hydrophobic and hydrophilic

peptides may spontaneously bind to CNTs. On the other hand, it is also observed that hydrophobic peptides indeed have stronger affinities for CNTs than hydrophilic ones (Wang et al., 2003).

Based on the experimental observations and the energetic analysis of the atomic model, I qualitatively develop the interaction energy parameters for HP lattice model according to the following assumptions: the binding affinity of CNTs for hydrophobic residues should be stronger than hydrophilic ones, which implies $\gamma_S(H,C) < \gamma_S(P,C)$. In addition, both hydrophobic and hydrophilic amino acid residues have affinities for CNTs, which implies that $\gamma_S(P,C) < 0$ and $\gamma_S(H,C) < 0$.

As to the relative energy parameters between "H-C" and "H-H" types of elements, the previous one is assumed to be lower. Such parameter is modeled with respect to the phenomenon that the CNT is extremely hydrophobic, which can form stable structure with hydrophobic residues in solvent (Dieckmann et al., 2003).

Referring to the relative values of $\gamma_S(H,H) = -2\varepsilon$, $\gamma_S(H,P) = -\varepsilon$ and $\gamma_S(P,P) = 0$, several parameter sets may be acceptable if they satisfy the criteria. I repeat the simulations of representative peptide binding to the CNT using the same setups and experimental conditions except for the interaction parameters. Although the average system energy and the accessible structures of the peptide are parameter-dependent, the binding mechanisms are qualitatively the same for these acceptable parameter sets. The thermodynamic quantities of peptide I binding to the

CNT using preliminary parameters at representative temperatures over $5*10^6$ MC runs are listed in Table 4.2.

As shown in Table 4.2, the process of peptides interacting with the CNT surface is a balance between an energetic gain and a conformational entropy loss. Both the internal energy and the peptide conformational entropy contribute to the Helmholtz energy. The internal energy is composed of two terms: the intra-molecular interaction within the peptide, and the peptide-CNT interaction. The selection of the parameters affects not only the absolute value of internal energy, but also the balance between the internal energy and the entropy. For example, at $T^* = 1.6$, the contribution of the internal energy U is -39.4 ε , when $\gamma_S(H,C) = -4\ \varepsilon$, $\gamma_S(P,C) = -3\varepsilon$. As the parameters are set as $\gamma_S(H,C) = -5\ \varepsilon$, $\gamma_S(P,C) = -4\ \varepsilon$, U decreases to -51.7 ε. The ratio of U to the Helmholtz free energy also enlarges.

Such observation is reasonable because if both “H” and “P” types of elements have strong affinity for the CNT, the intra-molecular interaction, which is stabilizing the compact structure of the peptide, becomes relatively weak. The strong binding affinity of the peptide to the CNT induces the conformational change of the peptide. Both the internal energy and the entropy would then be affected by this strength of affinity.

In addition to those parameters listed in Table 4.2, I have also taken into account other values of $\gamma_S(H,C)$ and $\gamma_S(P,C)$. If the interaction between “H-C” and “P-C” types of monomers is considered weak, for example $\gamma_S(H,C) = -2\ \varepsilon$, $\gamma_S(P,C) = -\varepsilon$, or $\gamma_S(H,C) = -\varepsilon$, $\gamma_S(P,C) = 0$, the entire peptide weakly binds to

the surface of the CNT. Such an observation is not consistent with the experimental and MD simulation results. Therefore these data are not listed for discussion. It is inferred that choosing $\gamma_S(H,C) < \gamma_S(H,H)$ seems also essential.

4.2.2.2 Conformational changes of peptide chain binding to CNT surface

Based on the developed parameters, the conformational changes of peptides upon binding are simulated and captured. In order to make the discussion more general, the second model peptide, namely peptide II is introduced. Compared with peptide I, peptide II possesses more hydrophilic monomers, and has no unique native structure.

At the start of each simulation, the model peptide is positioned partly in contact with the model CNT in their compact conformation, as shown in Figure 4.4. The vertical line beside the left-hand of the model peptide represents the CNT surface. A typical binding process is discussed on peptide I with preliminary parameters of $\gamma_S(H,C) = -5\ \varepsilon$, $\gamma_S(P,C) = -4\varepsilon$. Since the local energy barriers can be against the attempts that the chain unfolds into conformational trajectories leading to the lowest-accessible energy, the model chain binds reversibly at energies well above the energy minimum in the early stage of the binding process. Representative conformations of the two peptides shortly after their binding process are shown in Figure 4.5.

In the later stage of the adsorption, the chain is irreversibly bound to the surface, at least within the duration of the simulation. As shown in Figure 4.6, it is observed

that the peptide unfolds at the interface, and the adsorption is essentially irreversible. Some native contacts may also be retained due to the internal attractions between peptide residues. The similar configuration at equilibrium was also observed through MD simulations of atomic model. The corresponding system potential energy decreases to a lower average value, and fluctuates between low energy states.

The conformational change of peptide I is further illustrated in Figure 4.7. The number of monomers in the first and the fourth layers adjacent to CNT surface is traced against the MC simulation cycles. In the first layer, the number of monomers on the peptide increases most drastically to show the adsorption of the peptide. On the other hand, in the layers which are further away from the wall, for example, the fourth one, the number of monomers drops, and the denaturation of the peptide is observed. It is also observed that the average number of hydrophobic contacts for peptide I in bulk water is more than that of the peptide binding to the CNT surface at the same temperature, indicating a change of peptide properties upon binding.

Temperature is a crucial factor that affects the behavior of peptides binding to CNTs. For example, at $T^* = 0.8$, peptide I is trapped at its local minimum over the course of MC simulation. The monomer numbers in the first and the fourth layers do not change significantly at $T^* = 0.8$, showing that the peptide maintains its own conformation at low temperature, and the binding is reversible. On the other hand, at $T^* = 1.6$, the peptide is found denatured and bound to the CNT irreversibly. Further investigations show that the lowest-accessible energy state for peptide I is $E_M = -32.0\varepsilon$ at $T^* = 0.8$, but $E_M = -53.0\varepsilon$ at $T^* = 1.6$. Increasing the temperature

may facilitate the adoption of previously inaccessible low energy states of peptide-CNT system.

4.3 Discussions on comparison of MD and MC methods

While both the all-atom model and the coarse-grained model are developed to investigate the peptide-CNT interaction, different simulation techniques have also been adopted to implement the simulation. Among which both the MD and MC simulation methods are frequently used whereas they differ in a variety of ways. The most obvious difference is that MD provides information about the time dependence of the properties of the system whereas there is no temporal relationship between successive MC configurations. In a MC simulation the outcome of each trial move depends only upon its previous state, whereas in MD simulations it is possible to predict the configuration of the system at any time in the future or at any time in the past.

The two methods also differ in their ability to explore phase space. A MC simulation often gives much more rapid convergence of the calculated thermodynamic properties of a simple molecular liquid. But it may explore the phase space of large molecules very slowly due to the need for small steps unless special techniques such as the configurational bias MC method are used. MD may fail to cross the barriers between the conformations sufficiently often to ensure that each conformation is sampled and it can be very useful for exploration of the local phase space, whereas the MC method may be more effective for conformational changes, which jump to a completely different area of phase space.

MD has a kinetic energy contribution to the total energy whereas in a MC simulation the total energy is determined from the ensembles, originally canonical or constant-NVT ensembles. MD simulation is traditionally performed under conditions of the microcanonical or constant NVE ensemble. Both the MD and MC techniques can be modified to sample from other ensembles. As the two techniques in some ways complement each other in their ability to explore phase space, they can also be combined to facilitate the exploration of physical or chemical phenomena.

4.4 Remarks

In this work, the coarse-grained HP lattice model is employed to study the interaction between peptides and CNTs. This simple model is capable of reconstructing a wide variety of complex characteristics of the target problem. Using the dynamic MC simulations, this HP lattice model can reproduce the dynamic processes of peptides folding in bulk water, and peptides binding to the CNT surface.

The interaction parameters between the CNT monomers and two types of residues in our model are developed based on the related experimental data and MD simulation results. To the best of my knowledge, this is the first time that such a set of preliminary parameters are developed to study peptides' binding to the CNT. It is found that both hydrophobic and hydrophilic residues have affinity for CNTs. The simulation results imply that the suggested parameters here are acceptable and can qualitatively reflect the mechanisms of the binding process.

Since all the possible configurations of the system are enumerated, thermodynamics and conformational change of peptides binding to CNTs can be explored. The analysis of the thermodynamic quantities suggests that both the internal energy and the peptide conformational entropy contribute to the binding process. Upon binding to the CNT, peptides generally unfold into their denatured states to reach the low energy states of the system. In order to access the low energy levels, the peptide has to escape from local energy minima and the average number of native contacts may decrease. Temperature has a significant influence on the interaction affinity. The model developed in this work is validated by qualitatively agreeing with experimental observations and simulation results of high-resolution model, and is able to provide further insight that is crucial for designing nanotube-based devices and drug delivery systems.

Table 4.1 Thermodynamic quantities of sequence I in bulk water at different temperatures. In the table T^* is the dimensionless temperature, U is the internal energy, ΔG_{MU} is the standard free energy change, S is the conformational entropy of the peptide, A is the Helmholtz free energy, ρ_M is the probability that the system lies in the lowest-accessible energy of the system. The energy unit is ε.

T^*	$U(\varepsilon)$	$\Delta G_{MU}(\varepsilon)$	$T*S(\varepsilon)$	$A(\varepsilon)$	ρ_M
0.2	-14.0	$+\infty$	0.0	-14.0	1.00
0.8	-12.2	-0.3	3.9	-16.1	0.40
1.6	-6.2	-5.1	12.9	-19.1	0.04

Table 4.2 Thermodynamic properties of sequence I binding to the CNT using different parameters at representative temperatures. In the table E_M is the lowest-accessible potential energy. Other quantity units can be referred to Table 4.1.

$\gamma_S(H,C)$ $\gamma_S(P,C)$ (ε)	T^*	E_M (ε)	U (ε)	ΔG_{MU} (ε)	$T*S$ (ε)	A (ε)	ρ_M
-4,-3	1.6	-42.0	-39.4	0.3	3.6	-43.0	0.54
-5,- 4	1.6	-53.0	-51.7	1.5	1.8	-53.5	0.72
-4,-3	0.8	-28.0	-27.9	3.1	0.1	-28.0	0.98
-5,- 4	0.8	-32.0	-31.9	3.1	0.1	-32.0	0.98

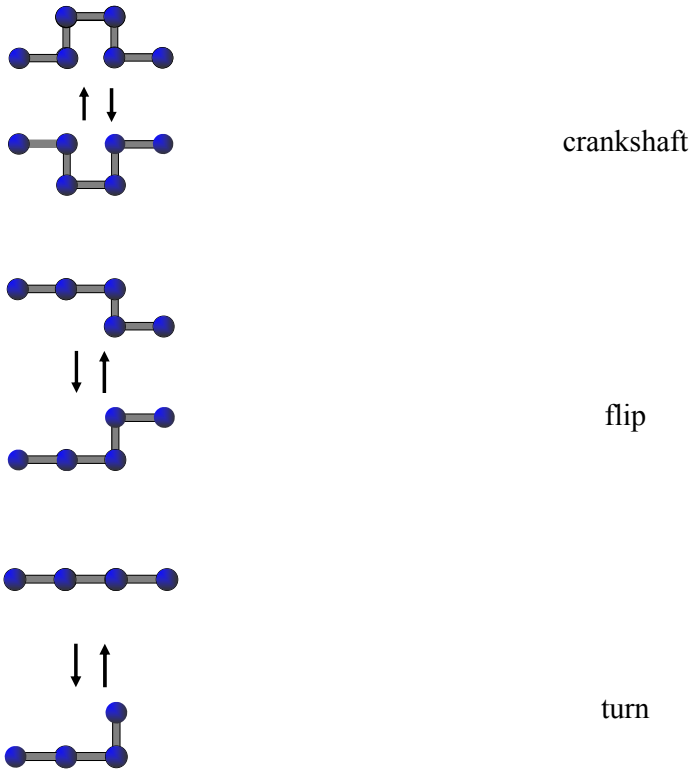


Figure 4.1 The Verdier–Stockmeyer moves allowed for peptide conformational transition.

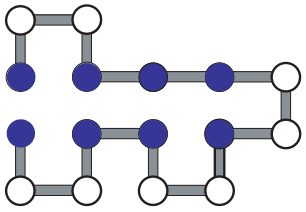


Figure 4.2 The initial conformation of model peptide I. The filled cycles represent hydrophobic elements, while the unfilled ones represent polar elements.

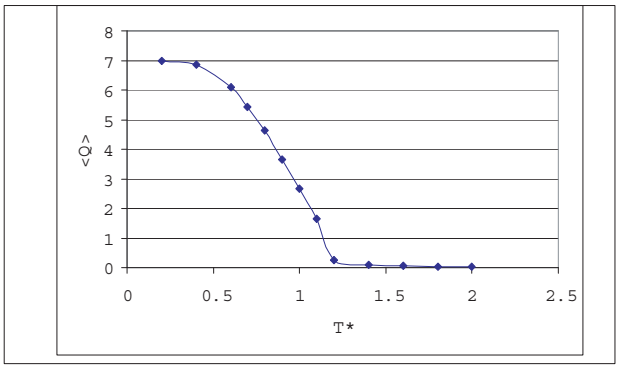


Figure 4.3 The ensemble averaged number of hydrophobic contacts $< Q >$, as the function of the dimensionless temperature, T^*.

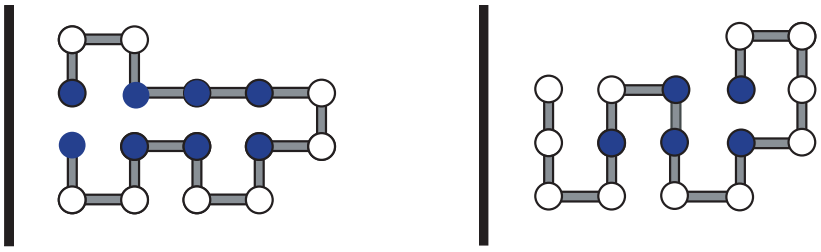


Figure 4.4 Initial structures of peptide sequence I (left) and sequence II (right) interacting with model CNT surface. Peptide sequence I has eight hydrophobic residues and sequence II possesses five. The filled cycles represent hydrophobic elements while unfilled ones represent the polar elements.

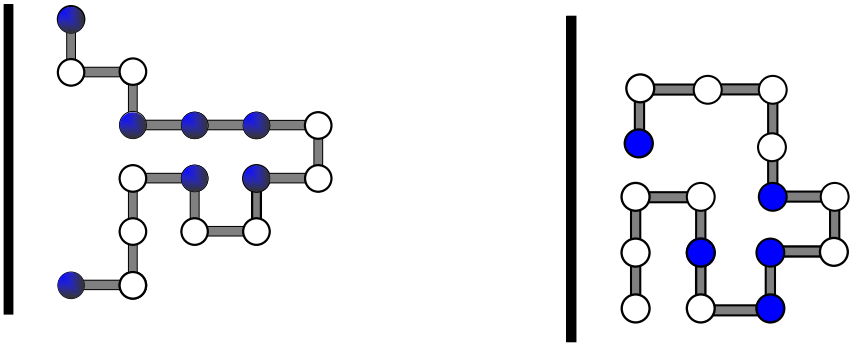


Figure 4.5 The representative conformations of sequence I (left) and sequence II (right) shortly after their binding to the CNT surface. The peptide-CNT interaction energy parameters are $\gamma_S(H,C) = -5\ \varepsilon$, $\gamma_S(P,C) = -4\varepsilon$.

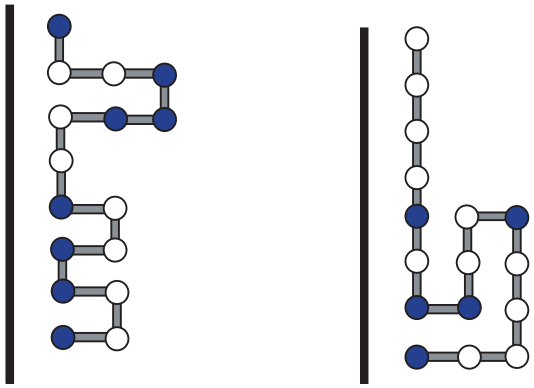


Figure 4.6 Representative conformations of sequence I (left) and sequence II (right) binding to CNT surface at $T^* = 1.6$ at $T^* = 1.6$. The peptide-CNT interaction energy parameters are $\gamma_S(H,C) = -5\ \varepsilon$, $\gamma_S(P,C) = -4\varepsilon$.

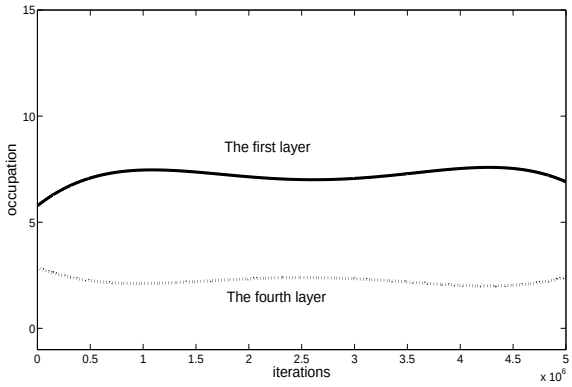


Figure 4.7 Illustrations of the averaged number of monomers in the first and the fourth layers adjacent to CNT surface against the MC cycles for peptide I at $T^* = 1.6$ (fitted using fourth order polynomials).

Chapter 5

Conclusions and Future work

5.1 The major conclusion from the study

This study has focused on investigating interaction mechanism between peptides and CNTs based on different levels of molecular description. Different computational strategies adopting either all-atom model or coarse-grained model are implemented. Through the studies, the following conclusions are drawn:

1) MD simulation of self-insertion of peptides into SWCNTs is carried out based on molecular mechanics of atomic model. The energetic analysis on the total potential energy of the system indicates that the non-bonded interaction (including the van der Waals interaction and the electrostatic interaction) dominates the process. It is also found that some of the simulated peptides are capable of self-inserting into SWCNTs while others are not. The propensities of self-insertion of peptides into SWCNTs correlate with their average hydrophobicity. Furthermore, sizes of SWCNTs should be chosen properly corresponding to dimensions of peptides to facilitate the insertion.

2) Free energies of peptides binding to SWCNTs are calculated using combined simulation methods of MD and MM-GBSA model. The calculation results of binding free energies qualitatively agree with experimental results of binding affinities of different peptides for SWCNTs. From a microscopic view, the van der Waals interaction plays a key role in driving the peptides onto the surface of CNTs. Further

simulation is carried out to estimate the interaction free energy between encapsulated peptides and SWCNTs, and consistent results on the energetic contributions are obtained.

3) In order to understand the affinity of different amino acids for SWCNTs, intensive molecular simulations have been carried out between individual amino acids and the same SWCNT using MM-GBSA method. Results show that the binding free energies vary with different amino acids. Generally hydrophobic amino acids possess stronger binding affinity for CNTs than the hydrophilic ones, and the average binding free energy from each group is evaluated.

While the hydrophobicity of peptides correlates with strength of peptide-CNT affinity, other properties of peptides, for example the aromatic rings, the peptide length and the CNT size also have impact on their interaction free energy.

4) The coarse-grained HP lattice model is employed to study the thermodynamics of the interaction between peptides and CNTs. The simplified model is capable of reconstructing a wide variety of complex characteristics of the target problem. The HP lattice model using MC simulation method can reproduce the dynamic processes of peptides folding in bulk water, as well as peptides binding to the CNT surface.

The interaction parameters between the CNT monomers and ‘H’ or ‘P’ types of residues in our model are developed based on the relevant experimental data and MD simulation results. A set of preliminary parameters are developed to study peptides’ binding to the CNT based on the observed affinity of hydrophobic and hydrophilic

amino acids for CNTs. The simulation results imply that the suggested parameters here can qualitatively reflect the mechanisms of the binding process.

5) Thermodynamics and conformational change of peptides binding to CNTs are also explored since all the possible configurations of the system can be enumerated by the coarse-grained HP lattice model. The analysis of the thermodynamic quantities suggests that both the internal energy and the peptide conformational entropy contribute to the binding process. Upon binding to the CNT, peptides generally unfold into their denatured states to reach the low energy states of the system. In order to access the low energy levels, the peptide has to escape from local energy minima and the average number of native contacts may decrease. Temperature has a significant influence on the conformational evolution of the peptide.

5.2 Recommendations for future research work

1) Despite the efforts we have made, there are still many open questions in application of biomolecule-CNT systems. For example, there have been limited studies on the use of CNTs to deliver drugs to live animals. More investigations should be performed to study the cellular uptake mechanism of CNTs. On the other hand, more work should be carried out to confirm that drugs or other molecular cargo can be successfully loaded onto or into the nanotube, and subsequently triggered to unload at some designated therapeutic site.

2) Given that the computational resources are more powerful, larger systems of simulations could be performed based on all-atom models. For example, the full sized

proteins consisting of hundreds of amino acids interacting with SWCNTs, or MWCNTs can be studied. Therefore our understanding of kinetics and energetics of protein-CNT interaction can be enhanced, and properties of the novel materials such as their biocompatibility can be clarified.

3) Although the MM-GBSA method is efficient in evaluating the interaction free energy based on the two-state theory, the paths through the end states are ignored. It is also recommended to develop the appropriate path-dependent methods for free energy estimation, again, on condition that the computer resource should be powerful enough to handle the simulation within reasonable simulation time.

4) Based on the coarse-grained model, further studies can be carried out to understand the free energy landscape of peptide-CNT. For example different sequences of peptides, as well as peptides composed of different number of monomers can be tested to further investigate the peptide-CNT interactions.

5) The computational models developed in this thesis can also be extended to investigate the interaction between other species of bio-materials. For example, they could be applied in the investigation of the interactions between proteins and biopolymers, or drug-resistant mutation systems, *etc*.

6) Both the all-atom model and the coarse-grained model have some limitations. For example, it is not easy to enumerate all the orientations of the peptide-CNT interaction. As a recommendation for future work, a continuum model could be developed to provide further insights into the problem, and therefore a sound theory could be established.

References

Ajayan PM, Stephan O, Colliex C and Trauth D (1994), Aligned carbon nanotube arrays formed by cutting a polymer resinnanotube composite. Science, 265, 1212-1214.

Allen MP and Tildesley DJ (1987), Computer Simulation of Liquids, Clarendon Press, Oxford.

Anderson RE, Pande VS and Radke CJ (2000), Dynamic lattice Monte Carlo simulation of a model protein at an oil/water interface. J. Chem. Phys., 112, 9167-9185.

Andriotis A, Menon M and Froudakis G (2000), Various bonding configurations of transition-metal atoms on carbon nanotubes: their effect on contact resistance. Appl. Phys. Lett., 76, 3890-3892.

Antonov RD and Johnson AT (1999), Subband population in a single-wall carbon nanotubes diode. Phys. Rev. Lett., 83, 3274-3276.

Åqvist J, Medina C and Samuelsson JE (1994), A new method for predicting binding affinity in computer-aided drug design. Protein Eng., 7, 385–391.

Ausman KD, Piner R, Lourie O and Ruoff, RS (2000), Controlled sliding and pullout of nested shells in individual multiwalled carbon nanotubes. J. Phys. Chem B, 104, 8911-8915.

Azamian BR, Davis JJ, Coleman KS, Bagshaw CB and Green MLH (2002), Bioelectrochemical single-walled carbon nanotubes. J. Am. Chem. Soc., 124, 12664-12665.

Balavoine F, Schultz P, Richard C, Mallouh V, Ebbesen TW and Mioskowski C (1999), Helical crystallization of proteins on carbon nanotubes: A first step

towards the development of new biosensors. Angew. Chem., 38(13/14), 1912-1915.

Baron R, Trzesniak D, de Vries AH, Elsener A, Marrink SJ and van Gunsteren WF (2007), Comparison of Thermodynamic Properties of Coarse-Grained and Atomic-Level Simulation Models, Chem Phys Chem, 8, 452 – 461.

Baschnagel J, Binder K, Doruker P, Gusev AA, Hahn O, Kremer K, Mattice WL, MQller-Plathe F, Murat M, Paul W, Santos S, Suter UW and Tries V (2000), Viscoelasticity, Atomistic Models, Statistical Chemistry, Vol 152, Advances in Polymer Sciences, Springer, Heidelberg, 41–165.

Bashford D and Case DA (2000), Generalized Born models of macromolecular solvation effects. Annu Rev Phys Chem, 51, 129–152.

Baughman RH, Cui CX, Zakhidov AA, Iqbal Z, Barisci JN, Spinks GM, Wallace GG, Mazzoldi A, De Rossi D, Rinzler AG, Jaschinski O, Roth S and Kertesz M (1999), Carbon nanotube actuators. Science, 284(5418), 1340–1344.

Berendsen HJ C, van Gunsteren WF, In Ciccotti G and Hoover WG (1986), Molecular-Dynamics Simulation of Statistical-Mechanical Systems North-Holland, Amsterdam.

Berendsen HJC, Postma JPM, van Gunsteren WF, Nola AD and Haak JR (1984), Molecular dynamics with coupling to an external bath. J. Chem. Phys., 81, 3684–3690.

Berendsen HJC, Van der Spoel D and van Drunen R (1995), GROMACS: A message-passing Parallel Molecular Dynamics Implementation. Comp. Phys. Commun., 91, 43-56.

Berg JM, Tymoczko JL, and Stryer L (2002), Biochemistry. 5th ed. W.H. Freeman, New York.

Beveridge DL and DiCapua FM (1989), Free energy via molecular simulation: Applications to chemical and biomolecular systems. Annu. Rev. Biophys. Biophys. Chem., 18, 431–492.

Blase X, Benedict LX, Shirley EL and Louie SG (1994), Hybridization effects and metallicity in small radius carbon nanotubes. Phys. Rev. Lett., 72, 1878-1881.

Bockrath M, Cobden DH, McEuen PL, Chopra NG and Zettl A (1997), single-electron transport in ropes of carbon nanotubes. Science, 275, 1922-1925.

Böhm HJ and Stahl M (1999), Rapid empirical scoring functions in virtual screening applications. Med. Chem. Res., 9, 445–462.

Bond PJ and Sansom MSP(2006), Insertion and assembly of membrane proteins via simulation. J. Am. Chem. Soc. 128, 2697-2704.

Boul PJ, Liu J, Mickelson ET, Huffman CB, Ericson LM, Chiang IW, Smith KA, Colbert DT, Hauge RH, Margrave JL and Smalley RE (1999), Reversible sidewall functionalization of buckytubes. Chemical Physics Letters, 310, 367 – 372.

Brooks BR, Bruiccoleri RE, Olafson BD, States DJ, Swaminathan S, and Karplus M. (1983), CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem., 4, 187-217.

Brown TLL, Bursten BE, and Lemay HE (1999), Chemistry: The Central Science, 8th ed, Prentice Hall PTR.

Camacho CJ and Thirumalai D (1993), Kinetics and thermodynamics of folding in model proteins. Proc. Natl. Acad. Sci. USA, 90, 6369-6372.

Carloni P, Rothlisberger U and Parrinello M (2002), The role and perspective of ab initio molecular dynamics in the study of biological systems, Acc. Chem. Res, 35, 455 – 464.

Carloni P, Rothlisberger U, Parrinello M (2002), The role and perspective of ab initio molecular dynamics in the study of biological systems. Acc. Chem. Res. 35(6), 455–464.

Case DA (1994), Normal Mode Analysis of protein dynamics. Current Opinion in Structure Biology, 4, 285-290.

Case DA, Pearlman DA, Caldwell JW, Cheatham TE, Wang J, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Tsui V, Gohlke H, Radmer RJ, Duan Y, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, and Kollman PA (2002), Assisted Model Building with Energy Refinement 7 (AMBER 7). University of California, San Francisco, CA.

Castells V, Yang S and Van Tassel PR (2002), Surface-induced conformational changes in lattice model proteins by Monte Carlo simulation. Phys. Rev. E, 65, 31912-31918.

Chan HS and Dill KA (1993), The protein folding problem. Phys. Today, 46, 24-32.

Chandler D (1987), Introduction to Modern Statistical Mechanics. Oxford University Press, New York.

Chen J, Hamon MA, Hu H, Chen Y, Rao AM, et al. (1998), Solution properties of single-walled carbon nanotubes. Science, 282, 95-98.

Chen RJ, Zhan YG, Wang DW and Dai H (2001), Noncovalent Sidewall Functionalization of Single-Walled Carbon Nanotubes for Protein Immobilization. J. Am. Chem. Soc., 123, 3838-3839.

Cheung, PSY (1977), On the calculation of specific heat, thermal pressure coefficients, and compressibilities in molecular dynamics simulations. Mol. Phys. 33, 519-526.

Chikenji G, Fujitsuka Y, and Takada S (2006), Shaping up the protein folding funnel by local interaction: Lesson from a structure prediction study. Proc. Natl. Acad. Sci. USA, 103, 3141-3146.

Choi WB, Chung DS, Kang JH, Kim HY, Jin YW, et al. (1999), Fully sealed, highbrightness carbon-nanotube field-emission display. Appl. Phys. Lett. 75, 3129-3131.

Chopra NG, Benedict LX, Crespi VH, Cohen ML, Louie SG and Zettl A (1995), Fully collapsed carbon nanotubes. Nature, 377, 135-137.

Cohen FE and Prusiner SB (1998), Pathologic conformations of prion proteins. Annu. Rev. Biochem., 67, 793-819.

Colombo G and Micheletti C (2006), Protein folding simulations: combining coarse-grained models and all-atom molecular dynamics, Theor Chem Acc, 116, 75–86.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Ferguson DC, Spellmeyer DC, Fox T, Caldwell JW, and Kollman PA (1995), A second

generation force field for the simulation of proteins, nucleic acids, and organic molecules. J. Am. Chem. Soc., 117, 5179-5197.

Dahlquist G and Bjork A (1974), Numerical Methods, Prentice-Hall, Englewood Cliffs, NJ.

Darden T, York D, and Pedersen L (1993), Particle mesh Ewald: an N_log (N) method for Ewald sums in large systems. J. Chem. Phys., 98, 10089-10092.

de Pablo P, Moreno-Herrero F, Colchero J, G´omez-Herrero J, Herrero P, et al. (2000), Absence of dc-conductivity in λ-DNA. Phys. Rev. Lett., 85, 4992-4995.

Delaney P, Choi HJ, Ihm J, Louie SG and Cohen ML (1998), Broken symmetry and pseudogaps in ropes of carbon nanotubes. Nature, 391, 466-468.

Dieckmann GR, Dalton AB, Johnson PA, Razal J, Chen J, Giordano G M, Munoz E, Musselman IH, Baughman RH and Draper RK (2003), J. Am. Chem. Soc. 125, 1770-1777.

Doolittle RF (1989), Prediction of protein structure and the Principles of Protein Conformation (Fasman, G. D., ed) Plenum Press, NY, 599-623.

Dresselhaus MS and Avouris P (2001), Introduction to carbon materials research, Carbon Nanotubes. 80, 1–9.

Dresselhaus MS, Dresselhaus G, and Eklund PC (1996), Science of Fullerenes and Carbon Nanotubes, Academic Press, San Diego.

Ebbesen TW (1997), Carbon Nanotubes: Preparation and properties, CRC Press, Baca Raton, FL.

Erlanger F, Chen B, Zhu M, and Brus L (2001), Binding of an anti-fullerene IgG monoclonal antibody to single-wall carbon nanotubes. Nano Lett., 1, 465-467.

Frenkel D and Smit B (1996), Understanding Molecular Simulation. From Algorithms to Applications. Academic Press, San Diego.

Gao H, Kong Y, Cui D, and Ozkan CS (2003), Spontaneous insertion of DNA Oligonucleotides into carbon nanotubes. Nano Lett., 3, 471-473.

Gao HJ, and Kong Y (2004), Simulation of DNA-Nanotube Interactions, Annu. Rev. Mater. Res., 34, 123-150.

Gear CW (1971), Numerical Initial Value Problems in Ordinary Differential Equations, Prentice-Hall, Englewood Cliffs, NJ.

Georagakilas V, Kordatos K, Prato M, Guldi DM, Holzinger M, and Hirsch AJ (2002), Organic functionalization of carbon carbon nanotubes. J. Am. Chem. Soc., 124, 760–761.

Georagakilas V, Tagmatarchis N, Pantarotto D, Bianco A, Briand JP, and Prato M (2002), Amino acid functionalisation of water soluble carbon nanotubes. Chem Commun., 24, 3050-3051.

Gervasio FL, Carloni P, Parrinello M (2002), Electronic structure of wet DNA. Phys. Rev. Lett., 89, 108102-108105.

Goetz R, Lipowsky R (1998), Computer simulations of bilayer membranes: Self-assembly and interfacial tension. J. Chem. Phys., 108, 7397-7409.

Gogotsi Y, Libera JA, Guvenc-Yazicioglu A, and Megaridis CM (2001), In-situ multiphase fluid experiments in hydrothermal carbon nanotubes, Appl. Phys. Lett., 79, 1021-1023.

Gsponer J, Haberthur U and Caisch A (2003), The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. Proc. Natl. Acad. Sci. USA, 100, 5154- 5159.

Gubbins KE and Quirke N (1996), Molecular Simulation and Industrial Applications. Methods, Examples and Prospects, Gordon and Breach, Amsterdam.

Hamada N, Sawada S, Oshiyama A (1992), New one-dimensional conductors: graphitic microtubules. Phys. Rev. Lett., 68, 1579-1581.

Harris PJF (1999), Carbon Nanotube and Related Structures: New Materials for the 21st Century, Cambridge University Press, Cambridge, UK.

Harutyunyan AR, Pradhan1 BK, Sumanasekera1 GU, Korobko EY, and Kuznetsov AA (2002), Carbon nanotubes for medical applications European Cells and Materials., 3 Suppl. 2, 84-87.

Hawkins GD, Cramer CJ, Truhlar DG (1995), Pairwise solute descreening of solute charges from a dielectric medium. Chem Phys Lett., 246, 122–129.

Hill TL (1956), Statistical Mechanics. McGraw-Hill, New York.

Hirahara K, Suenaga K, Bandow S, Kato H, Okazaki T, Shinohara H, and Iijima S (2000), One-Dimensional Metallofullerene Crystal Generated Inside Single-Walled Carbon Nanotubes. Phys. Rev. Lett., 85, 5384-5387.

Hirsch A (2002), Functionalization of single-walled carbon nanotubes. Angew. Chem., Int. Ed. Engl, 41, 1853-1859.

Hockney RW and Eastwood JW (1988), Computer simulation using Particles, Adam Hilger, Bristol.

Hummer G, Rasalah JC, Noworyta JP (2001), Water conduction through the hydrophobic channel of a carbon nanotubes, Nature, 414, 188-190.

Hunter CA and Sanders JKM (1990), The nature of $\pi - \pi$ interactions. J. Am. Chem. Soc., 112, 5525-5534.

Iijima S (1991), Helical microtubules of graphitic carbon. Nature, 354, 56-58.

Iijima S, Brabec C, Maiti A, Bernholc J (1996), Structural flexibility of carbon nanotubes. J. Chem. Phys. 104, 2089-2092.

Ito T, Sun L and Crooks RM (2003), Observation of DNA transport through a single carbon nanotube channel using fluorescence microscopy, Chem. Commum., 7, 1482-1483.

Jarvis SP, Uchihashi T, Ishida T, Tokumoto H and Nakayama Y (2000), Local solvation shell measurement in water using a carbon nanotube probe, J. Phys. Chem. B, 104(26), 6091-6094.

Jorgensen WJ, Chandreskhar J, Madura J, Imprey R and Klein M (1983), Comparison of simple potential functions for simulating water, J. Chem. Phys., 79, 926–935.

Jorgensen WL (1989), Free-energy calculations: A breakthrough for modeling organic-chemistry in solution. Acc. Chem. Re., 22, 184–189.

Kale L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, et al. (1999), NAMD2, greater scalability for parallel molecular dynamics. J. Comp. Phys., 151, 283- 312.

Kaplus M (1990), Molecular dynamics simulations in biology. Nature, 347, 631-639.

Kollman PA (1993), Free energy calculations: Applications to chemical and biochemical phenomena. Chem. Rev., 93, 2395–2417.

Kremer K (2003), Computer simulations for macromolecular science. Macromol. Chem. Phys., 204, 257-264.

Kuhn B, Baker D, Kollman PA (2000), A ligand is predicted to bind better to avidin than biotin: Insights from computational fluorine scanning. J. Am. Chem. Soc., 122, 3909-3916.

Kumar S, Giri D, Bhattacharjee SM (2005), Force-induced triple point for interacting polymers. Phys. Rev. E, 71, 051804-051808.

Kwon YK, Tomanek D (1998), Electronic and structural properties of multiwall carbon nanotubes. Phys. Rev. B, 58, R16001−R16004.

Kyte J and Doolittle RF (1982), A simple method for displaying the hydrophobic character of a protein, J. Mol. Biol., 157, 105-132.

Lamb ML and Jorgensen, WL (1997), Computational approaches to molecular recognition. Curr. Opin. Chem. Biol., 1, 449–457.

Lau KF, and Dill KA (1989) Macromolecules, A lattice statistical mechanics model of the conformational sequence spaces of proteins. 22, 3986-3997.

Leach AR (2001), Molecular Modelling: principles and applications, Published by Pearson Education EMA.

Lee MR, Baker D, Kollman PA (2001), 2.1 and 1.8 Å average C_α RMSD structure predictions on two small proteins, HP-36 and s15. J. Am. Chem. Soc, 123, 1040-1046.

Lee SM, Lee YH (2000), Hydrogen Storage In Single-Walled Carbon Nanotubes. Appl. Phys. Lett., 76 (20), 2877-2879.

Lee T, Kollman PA (2000), Theoretical studies suggest a new antifolate as a more potent inhibitor of thymidylate. Synthase. J Am chem. Soc., 122, 4385-4393.

Leimkuhler BJ and Skeel RD (1994), Symplectic numerical integrators in constrained Hamiltonian systems. J. Comput. Phys., 112, 117-125.

Li H, Helling R, Tang C, Wingreen NS (1996), Emergence of preferred structures in a simple-model of protein-folding. Science, 273, 666-669.

Li J, Cassell AM, and Dai H (1999), The carbon nanotube as AFM tips: Measuring DNA Molecules at the Liquid/Solid Interface, Surf. Interface Anal, 28, 8-11.

Liu SM and Haynes CA (2004), Mesoscopic analysis of conformational and entropic contributions to nonspecific adsorption of HP copolymer chains using dynamic Monte Carlo simulations, J. Colloid Interface Sci. 275, 458-469.

Marrink SJ, de Vries AH, Mark AE (2004), Coarse grained model for semiquantitative lipid simulations, J. Phys. Chem. B, 108, 750-760.

Massova I and Kollman PA (1999), Computational Alanine Scanning To Probe Protein-Protein Interactions: A Novel Approach To Evaluate Binding Free Energies, J. Am. Chem. Soc, 121, 8133-8142.

McQuarrie, D A (1976), Statistical Mechanics. Harper and Row, New York.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, and Teller E (1953), Equation of state calculations by fast computing machines. J. Chem. Phys., 21, 1087-1092.

Mickelson ET, Chiang IW, Zimmerman JL, Boul PJ, Lozano J, Smalley RE, Hauge RH, Margrave JL (1999), J. Phys. Chem. B, 103, 4318-4322.

Miller DW and Dill KA (2006), Ligand binding to proteins: the binding landscape model, Protein. Sci., 6, 2166-2179.

Mueller-Plathe F (2002), Coarse-graining in Polymer Simulation: From the Atomistic to the Mesoscopic Scale and Back, Chem Phys Chem, 3, 754-769.

Munster (1969), A. Statistical thermodynamics. Springer, Academic Press.Berlin, New York.

Nguyen CV, Delzeit L, Cassel AM, Li J, Han J, and Meyyappan M (2002), Preparation of nucleic acid functionalized carbon nanotube arrays. Nano Lett, 2, 1079-1081.

Onufriev A, Bashford D, and Case DA (2000), Modification of the Generalized Born Model Suitable for Macromolecules. J. Phys. Chem. B, 104, 3712-3720.

Ordej´on P (2000). Linear scaling ab initio calculations in nanoscale materials with SIESTA. Phys. Stat. Sol. B, 217, 335-356.

Overney G, Zhong W, Tomanek D (1993), Structural rigidity and low frequency vibrational modes of long carbon tubules. Z. Phys. D, 27, 93-96.

Pantarotto D, Partidos CD, Graff R, Howbeke J, Briand JP, Prato M, and Bianco A (2003), Synthesis, structural characterization and immunological properties of carbon nanotubes functionalized with peptides. J. Am. Chem. Soc., 125, 6160-6164.

Pantarotto D, Partidos CD, Hoebeke J, Brown F, Kramer E, Briand JP, Muller S, Prato M, and Bianco A (2003), Immunization with peptide-functionalized carbon nanotubes enhances virus-specific neutralizing antibody responses. Chem Biol. 10(10), 961-966.

Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, et al. (1995), AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comp. Phys. Commun. 91, 1-41.

Poncharal P, Wang ZL, Ugarte D, de Heer WA (1999), Electrostatic deflections and electromechanical resonances of carbon nanotubes. Science, 283, 1513-1516.

Qian D, Wagner GJ, and Liu WK, Yu MF, and Ruoff RS (2002), Mechanics of carbon nanotubes, Appl Mech Rev, 55, 495-533.

Qiu D, Shenkin PS, Hollinger FP, Still WC (1997), The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate born radii. J Phys Chem A, 101, 3005–3014.

Rapaport, DC (1995), The art of Molecular Dynamics Simulation, Cambridge University Press, Cambridge.

Robertson DH, Brenner DW, Mintmire JW (1992), Energetics of nanoscale graphitic tubules. Phys. Rev. B, 45:12592-12595.

Ruoff RS, Lorents DC (1995), Mechanical and thermal properties of carbon nanotubes. Carbon, 33, 925- 930.

Rychaert JP, Ciccotti G, Berendsen HJC (1977), Numerical integration of the cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes, J. Computat. Phys. 23327-341.

Sadus RJ (1999), Molecular Simulation of Fluids, Theory, Algorithms and Object-orientation, Amsterdam: Elsevier,.

Shih AY (2006), Arkhipov A, Freddolino PL, and Schulten K, J. Phys. Chem. B, 110, 3674-3684..

Shimoda H, Oh SJ, Geng HZ,Walker RJ, Zhang XB, et al. (2002), Self-assembly of carbon nanotubes. Adv. Mater., 14, 899-901.

Sitkoff D, Sharp KA, Honig B (1994), Accurate calculation of hydration free energies using macroscopic solvent models. J. Phys. Chem., 98, 1978-1988.

Smit B, Hilbers PAJ, Esselink K, Rupert LAM, van Os NM, Schlijper AG (1990), Nature, 348, 624-625.

Smith BW, Monthioux M, Luzzi DE (1998), Encapsulated C_{60} in carbon nanotubes, Nature, 396, 323-323.

Spyroulias GA, Nikolakopoulou P, Tzakos A, Gerothanassis IP, Magafa V, Manessi-Zoupa E and Cordopatis P (2003), Comparison of the solution structures of angiotensin I & II: Implication for structure-function relationship. Eur. J. Biochem., 270, 2163-2173.

Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA (1998), Continuum solvent studies of the stability of dna, rna, and phosphoramidate-dna helices. J Am Chem Soc, 120, 9401- 9409.

Srinivasan J, Trevathan MW, Beroza P, Case DA (1999), Molecular dynamics and NMR spin relaxation in proteins. Theor Chem Acc, 101, 426–434.

Srivastava D, Menon M, Ajayan PM (2003), Branched carbon nanotube junctions predicted by computational nanotechnology and fabricated through nanowelding. J. Nanopart. Res., 5, 395-398.

Stutman G (2002) Classical molecular dynamics. In Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms, 211–254.

Tans SJ, Devoret MH, Dai H, Thess A, Smalley RS, et al. (1997), Individual

singlewall carbon nanotubes as quantum wires. Nature, 386, 474-476.

Tans SJ, Verschueren ARM and Dekker C (1998), Room temperature transistor based on a single carbon nanotube. Nature, 393, 49-52.

Terrones M, Banhart F, Grobert N, Charlier JC, Terrones H, et al. (2002), Molecular junctions by joining single-walled carbon nanotubes. Phys. Rev. Lett., 89, 075505.

Tersoff J (1992), Energies of fullerenes. Phys. Rev. B, 46, 15546-15549

Thostenson ET, Ren Z and Chou TW (2001), Advances in the science and technology of carbon nanotubes and their composites: a review, Composites Science and Technology, 61, 1899–1912.

Tomasi J and Persico M (1994), Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. Chem Rev, 94, 2027– 2094.

Tozzini V (2005), Coarse-grained models for proteins, Curr. Opin. Struct. Biol., 15, 144-150.

Treacy MMJ, Ebbesen TW, and Gibson JM (1996), Exceptionally high Young's modulus observed for individual carbon nanotubes. Nature 381, 678-680.

Tsui V and Case DA (2001), Theory and applications of the generalized Born solvation model in macromolecular simulation. Biopolymers, Nucl. Acid. Sci., 56, 275-291.

van Gunsteren WF, Luque FJ, Timms D, and Torda AE (1994), Molecular mechanics in biology: from structure to function, taking account of solvation. Annu. Rev. Biophys. Biomol. Struct., 23, 847-863.

Vegt and Yu HB (2006), Biomolecular modeling: goals, problems, perspectives. Angew. Chem. Int. Ed., 45, 4064 – 4092.

Verdir PH and Stockmayer WH (1962), Monte Carlo calculations on the dynamics of polymers in dilute solution. J. Chem. Phys., 36, 227-235.

Verlet L (1967) Computer "experiments" on classical fuids. I. Thermodynamical properties of Lennard-Jones molecules. Phys. Rev., 159, 98-103.

Wang J, Cieplak P and Kollman PA (2000), How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules. J. Comput. Chem., 211049-211074.

Wang J, Morin P, Wang W, and Kollman PA (2001), Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J. Am. Chem. Soc., 123, 5221- 5230.

Wang Q, and Johnson JKJ (1999), Optimization of carbon nanotubes for hydrogen adsorption Phys. Chem. B, 103, 4809-4813.

Wang S, Humphreys ES, Chung SY, Delduco DF, Lustig SR, Wang H, Parker KN, Rizzo NW, Subramoney S, Chiang YM, and Jagota A (2003), Peptides with selective affinity for carbon nanotubes, Nat. Mater., 2(3), 196-200.

Wang W, and Kollman PA (2000), calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. J. Mol. Biol., 303, 567-582.

Waters ML (2002), Aromatic interactions in model systems, current Opinion in Chemical Biology, 6, 736-741.

Wei BQ, Vajtai R, Jung Y, Ward J, Zhang R, et al. (2002), Microfabrication technology: organized assembly of carbon nanotubes. Nature, 416, 495-496.

Weiser J, Shenkin PS, and Still WC (1999), Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). J. Computat. Chem., 20, 217-230.

Williams KA, Veenhuizen PTM, de la Torre BG, Eritjia R, and Dekker C (2002), Carbon nantotubes with DNA recognition. Nature, 420, 761-761.

Wong EW, Sheehan PE, and Lieber CM (1997), Nanobeam mechanics: elasticity, strength and toughness of nanorods and nanotubes. Science, 277, 1971-1975.

Wong SS, Joselevich E, Woolley AT, Cheung CL, and Lieber CM (1998), Covalently functionalized nanotubes as nanometersized functionalized nanotubes as nanometersized. Nature, 394, 52-55.

Yeh IC and Hummer G (2004), Nucleic acid transport through carbon nanotube membranes. PNAS, 101, 12177–12182.

Zhao W, Song C, and Pehrsson PE (2002), Water-soluble and optically pH-sensitive single-walled carbon nanotubes from surface modification. J. Am. Chem. Soc.124, 12418-12419.

Zhdanov VP, and Kasemo B (1998), Monte Carlo simulation of denaturation of adsorbed proteins. Proteins: Structure, function, and Genetics. 30, 168-176.

Zoetea V, Michielinb O, and Karplus M (2003), Protein-ligand binding free energy estimation using molecular mechanics and continuum electrostatics. Application to HIV-1 protease inhibitors. Journal of Computer-Aided Molecular Design, 17, 861–880.

Publications arising from thesis

Journal papers

1. Cheng Y, Liu GR, Li ZR, Lu C and Mi D (2008), A thermodynamic study of peptides binding to carbon nanotubes based on a hydrophobic-polar lattice model using Monte Carlo Simulations, Journal of Physics D: Applied Physics, in press.

2. Cheng Y, Liu GR, Li ZR and Lu C (2006), Computational Analysis of Binding Free Energies between Peptides and Single-Walled Carbon Nanotubes, Physica A, 367, 293–304.

3. Liu GR, Cheng Y, Mi D and Li Z R (2005), A study on self-insertion of peptides into single-walled carbon nanotubes based on molecular dynamics simulation. International Journal of Modern Physics C., 16, (8), 1239-1250.

4. Li ZR, Liu GR, Cheng Y (2005), Thermodynamic analysis of protein sequence-structure relationships in monomer and dimer forms, Physica A, 354, 381-392.

Conference papers

1. Cheng Y, Liu GR, Li ZR, Lu C (2007), Analysis of thermodynamics of Peptides encapsulated into Single-Walled Carbon Nanotubes based on an Atomistic Model, Challenges in Computer Simulations (CCS2007), Singapore.

2. Cheng Y, Liu GR, Li ZR, Lu C (2006), Molecular Dynamics Simulations of Interactions between Peptides and Single-Walled Carbon Nanotubes, 7th World Congress on Computational Mechanics, Los Angles, the USA.

3. Cheng Y, Liu GR, Mi D, Li ZR (2004), Molecular dynamics simulation of peptides insertion into single-walled carbon nanotubes, The International Conference on Computational Methods, Singapore.

4. Li ZR, Liu GR, Cheng Y (2004), Designability of proteins and stability analysis upon dimerization using 2D lattice model, The International Conference on Computational Methods, Singapore.