

**Molecular Dynamics Simulation of Linear
Perfluorocarbon and Hydrocarbon Liquid-Vapor
Interfaces**

by

Jennifer Tsengjian Chin

B.A., Chemistry

Barnard College, Columbia University (1990)

M.S.C.E.P., Chemical Engineering

Massachusetts Institute of Technology (1995)

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN CHEMICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1999

© Massachusetts Institute of Technology 1999. All rights reserved.

Author

Department of Chemical Engineering

May 25, 1999

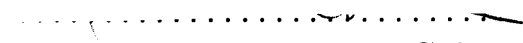
Certified by



Jonathan G. Harris

Thesis Supervisor

Accepted by



Robert E. Cohen

St. Laurent Professor of Chemical Engineering

Chairman, Committee for Graduate Students

Molecular Dynamics Simulation of Linear Perfluorocarbon and Hydrocarbon Liquid-Vapor Interfaces

by

Jennifer Tsengjian Chin

Submitted to the Department of Chemical Engineering
on May 25, 1999, in partial fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY IN CHEMICAL ENGINEERING

Abstract

This thesis aims to use molecular dynamics simulation techniques and perturbation theory to reproduce bulk and liquid-vapor interfacial properties of two classes of technologically important compounds – linear perfluoroalkanes and linear alkanes; to understand the underlying physical factors that contribute to the experimentally observed significant difference in surface tension values between perfluoroalkanes and their corresponding alkanes; and to explain why molecular dynamics simulation studies consistently over-estimate the surface tension values for both linear perfluoroalkanes and linear alkanes.

A new united-atom Lennard-Jones forcefield was developed and validated through simulating bulk properties of liquid perfluoroalkanes. This new forcefield was subsequently applied to liquid-vapor interface simulations of perfluoroalkanes, and produced significantly improved bulk and interfacial properties. In addition, the effects of longer cut-off radius in simulation, different torsional energy barrier, and wider bond angle were tested. Overall, the new Lennard-Jones forcefield produced the most significant improvement whereas the other factors evaluated displayed much smaller effects to surface tension prediction. Although the new Lennard-Jones forcefield still over-predicted the experimental surface tension values, when compared with liquid-vapor interface simulation results for alkanes, it predicted beyond simulation uncertainty, that perfluoroalkanes have lower surface tension values than their alkane counterparts.

To further explore the cause of surface tension over-estimation and provide an estimate for the contribution from three-body interactions, perturbation theory was applied to alkane and perfluoroalkane liquid-vapor interfaces. The results of the molecular dynamics simulation of liquid-vapor interfaces were used as the reference system. The inclusion of the three-body contribution significantly lowered the predicted surface tension values for both linear perfluoroalkanes and linear alkanes, and brought the predicted surface tension values to good agreement with experimental values. This work suggests that three-body effect is non-negligible for interfacial properties such as surface tension, and must be included explicitly by simulation or

by theory in order to more accurately predict liquid-vapor interfacial properties.

Thesis Supervisor: Jonathan G. Harris

Title:

Acknowledgments

I would like to express my gratitude to my thesis advisor Prof. Jonathan G. Harris for his mentorship, and to members of my thesis committee – Profs. Daniel Blankschtein, Robert Brown, Sow-Hsin Chen, and Gregory Rutledge for lighting the way. I am very grateful for the support, guidance, and friendship of my Chemical Engineering Practice School Directors: Drs. Colin Walden, Friedrich K. von Gottberg (Natick Station), and Barry S. Johnston (Midland Station). I also wish to thank Profs. Alexander M. Klibanov and Robert S. Langer for their guidance on the project of protein processing in organic media.

I want to take this opportunity to thank my officemates: Sandeep A. Patel (my “brother” the “Deepman”) for sharing the ups and downs of research, and many “deep” general discussions of the world; Jason Cline for all the help he provided on computer networking; and Matthew Reagan and Randy Weinstein for their friendship.

I would like to acknowledge the Office of Naval Research for funding part of this project.

I am forever indebted to my family. My grandparents in Shanghai, with whom I spent the first seventeen years of my life, instilled into me the belief that education is the most valuable personal possession in my life. My dad, who passed away when I was fourteen, left me with so many precious memories to cherish and draw strength from. My mom, who made countless sacrifices for me, always had confidence in me even when I had doubts about myself. My husband and best friend, Jianxin, gave me love, support, encouragement, and understanding when I needed most.

To My Mother Nai-Xun

Contents

1	Introduction	12
1.1	Motivation	12
1.2	Statement of Objectives	17
2	Molecular Dynamics Simulation – Fundamentals and Methods	19
2.1	Foundation of Molecular Dynamics	19
2.1.1	Newtonian Equation of Motion	20
2.1.2	Hamiltonian Equation of Motion	20
2.1.3	Lagrangian Equation of Motion	21
2.2	Integration	23
2.2.1	Verlet Algorithm	23
2.2.2	SHAKE and RATTLE	24
2.3	Simulation Techniques	25
2.3.1	Simulation Cell and Periodic Boundary Conditions	26
2.3.2	Cutoff Distance and Mean-Field Correction	27
2.3.3	Temperature Calculation and Control	27
2.3.4	Pressure Calculation and Control	31
2.4	Forcefields	33
2.4.1	Fundamentals of Forcefields	33
2.4.2	Various Types of Forcefield	34
3	Previous Work	39
3.1	Experimental	39

3.1.1	Melting and Boiling Points	39
3.1.2	Densities	42
3.1.3	Heats of Vaporization	42
3.1.4	Polarizability	42
3.1.5	Molecular Structures	42
3.1.6	Surface Measurements	49
3.2	Liquid-Vapor Interface Simulation	49
3.2.1	Rare Gases – Argon/Krypton/Xenon	52
3.2.2	Water	53
3.2.3	Metal	54
3.2.4	Alkane Oligomers	54
3.2.5	Perfluorohydrocarbons and Semifluorinated Alkane Diblocks	54
3.3	Review of Relevant Forcefields	55
3.3.1	Forcefields for Hydrocarbons	55
3.3.2	Forcefields for Perfluorinated Carbons	57
4	New Lennard-Jones United-Atom Intermolecular Forcefields	60
4.1	Rational and Motivation	60
4.2	Simulation Details	63
4.3	Results and Discussion	68
4.3.1	Parameterization	68
4.3.2	Validation	81
4.4	Conclusion	82
5	Liquid-Vapor Interface Simulation	83
5.1	Objectives	83
5.2	Simulation Details	83
5.3	Property Calculation	88
5.3.1	Density Profile and Width of Interface	88
5.3.2	Surface Tension	88
5.4	Simulation Results	90

5.4.1	Density Profiles	90
5.4.2	Effect of Various Factors on Surface Tension	99
5.4.3	Chain Orientation	100
5.5	Conclusion	103
6	Perturbation Theory and Three-Body Interaction	105
6.1	Introduction	105
6.2	Three-Body Interaction	106
6.3	Perturbation Theory	108
6.4	Estimate Three-body Contribution to Surface Tension with Perturbation Theory: Calculation and Results	109
6.4.1	Estimating Three-Body Coefficient	110
6.4.2	Calculate Transverse Pair Correlation Function	112
6.4.3	Density Profile	113
6.4.4	Three-Body Contribution to Surface Tension	113
6.5	Conclusion	117
7	Summary and Suggested Future Work	118
A	Surface Tension and Its Long-Range Correction	120
B	Published Paper on Protein Solubility in Organic Solvents	123

List of Figures

1-1	Sketch of Liquid-Vapor Interface. (A) a molecule in bulk liquid phase; and (B) a molecule near/at liquid-vapor interface.	13
1-2	Interplay of Experiment, Simulation, and Theory	15
2-1	A two-dimensional periodic system	26
2-2	Graphic illustration of intra- and inter- molecular interactions (Adapted from Discover User Guide Version 2.9/3.1)	36
3-1	Structure of alkanes and perfluoroalkanes	48
4-1	Torsional Forcefields, by Dixon and by Shin <i>et al.</i>	64
4-2	Lennard-Jones Forcefield, Hariharan <i>et al.</i>	70
4-3	Newly Developed Lennard-Jones Forcefield	71
4-4	Lennard-Jones Forcefield for $CF_2 - CF_2$	72
4-5	Lennard-Jones Forcefield for $CF_3 - CF_3$	73
4-6	Error (E) Contour at C ($CF_2 - CF_2$) = 0.0681 ($nm^{12}K$) and C ($CF_3 - CF_3$) = 0.0216 ($nm^{12}K$)	74
4-7	Error (E) Contour at C ($CF_2 - CF_2$) = 0.0681 ($nm^{12}K$) and D ($CF_2 - CF_2$) = 2.743 (nm^6K)	75
4-8	Error (E) Contour at C ($CF_2 - CF_2$) = 0.0681 ($nm^{12}K$) and D ($CF_3 - CF_3$) = 3.421 (nm^6K)	76
4-9	Error (E) Contour at C ($CF_3 - CF_3$) = 0.0216 ($nm^{12}K$) and D ($CF_2 - CF_2$) = 2.743 (nm^6K)	77

4-10 Error (E) Contour at C ($CF_3 - CF_3$) = 0.0216 ($nm^{12}K$) and D ($CF_3 - CF_3$) = 3.421 (nm^6K)	78
4-11 Error (E) Contour at D ($CF_2 - CF_2$) = 2.743 (nm^6K) and D ($CF_3 - CF_3$) = 3.421 (nm^6K)	79
5-1 Sketch of Liquid-Vapor Interface Simulation Geometry	84
5-2 Scaled Number Density Profile, Simulation Set I	90
5-3 Scaled Number Density Profile, Simulation Set II	91
5-4 Scaled Number Density Profile, Simulation Set III	92
5-5 Scaled Number Density Profile, Simulation Set IV	93
5-6 Scaled Number Density Profile, Simulation Set V	94
5-7 Scaled Number Density Profile, Simulation Set VI	95
5-8 Scaled Number Density Profile, Simulation Set VII	96
5-9 Liquid-Vapor Interface, n-Perfluorodecane at 323 K, Top view, Interface is in the plane of the page (z axis perpendicular to the plane of the page), $L_x=L_y=5.0nm$, Periodic boundary condition in all directions, End groups labeled in red, Center groups labeled in green.	97
5-10 Liquid-Vapor Interface, n-Perfluorodecane at 323 K, Side view, Interface is perpendicular to the plane of the page and at the top and bottom of the page, $L_x=L_y=5.0nm$, Periodic boundary condition in all directions, End groups labeled in red; Center groups labeled in green.	98
5-11 Chain Orientation Parameters, n-perfluorodecane at 323 K	101
5-12 Chain Orientation Parameters, n-perfluorodecane at 400 K	102
6-1 Non-additivity in Long-range Dispersion Energy. (I)Case where correlation between A and B is enhanced by C; (II)Case where correlation between A and B is reduced by C, angles and separations in three-body system indicated	107
6-2 Transverse Pair Correlation Function, n-perfluorodecane, 400K	114
6-3 Transverse Pair Correlation Function, n-perfluorodecane, 323K	115
6-4 Transverse Pair Correlation Function, n-decane, 400K	116

List of Tables

3.1	Boiling Point and Melting Point of Linear Alkanes	40
3.2	Boiling Point and Melting Point of Linear Perfluoroalkanes	41
3.3	Densities of Alkanes.	43
3.4	Densities of Perfluoroalkanes, $t = \text{temp. in } (^{\circ}C), T = \text{temp. in } K$. .	44
3.5	Heats of Vaporization of Linear Alkanes	45
3.6	Heats of Vaporization of Linear Perfluoroalkanes	46
3.7	Polarizability of Linear Alkanes	47
3.8	Polarizability of Linear Perfluoroalkanes	47
3.9	Experimental Surface Tension Value of Hydrocarbons	50
3.10	Experimental Surface Tension Value of Perfluorocarbons	51
4.1	Lennard-Jones Parameters, by Hariharan <i>et al.</i>	61
4.2	Experimental Bulk Properties	62
4.3	Parameters of Torsional Forcefields	65
4.4	New Lennard-Jones Parameters	69
4.5	Simulated Bulk Properties, with Different Lennard-Jones Forcefields .	80
4.6	Forcefield Validation: Experimental and Simulated Bulk Properties .	81
5.1	Liquid-Vapor Simulation Conditions Summary	88
5.2	Liquid-Vapor Interface Simulation Results	99
6.1	Polarizability of Linear Perfluoroalkanes	111
6.2	Three-Body Coefficients	112
6.3	Three-Body Contribution to Surface Tension	117

Chapter 1

Introduction

1.1 Motivation

The physical properties of inhomogeneous systems, such as liquid-vapor, liquid-liquid, and liquid-solid interfaces, are very different from those of bulk phases. For example, at a liquid-vapor interface, as Figure 1-1 illustrates, every molecule in the liquid experiences strong attractive intermolecular forces from its neighboring molecules. When a molecule is in the bulk liquid (molecule A), far away from the interface, its surrounding is isotropic and it has equal interactions in all directions. Therefore, there is no net force acting on the molecule in the bulk liquid. On the contrary, when a liquid molecule (molecule B) is near or at the liquid-vapor interface, it experiences a much smaller attractive force from the molecules on the low density side of the interface, hence there is a net inward (i.e., toward the liquid side) force on molecule B. In order to bring molecules to the liquid-vapor interface and increase the interfacial area, work must be done against this net inward force at the interface, hence the liquid-vapor interface has higher free energy than the bulk liquid.

The study of interfacial behavior has long been an area of active research. From an applicational viewpoint, the properties at the interfaces are crucial in understanding many industrial processes such as wetting, lubrication, and coating. Properties of materials at interfaces are governed by many different factors, such as chemical and physical composition of the materials, temperature, orientation of materials at the

Vapor

Liquid

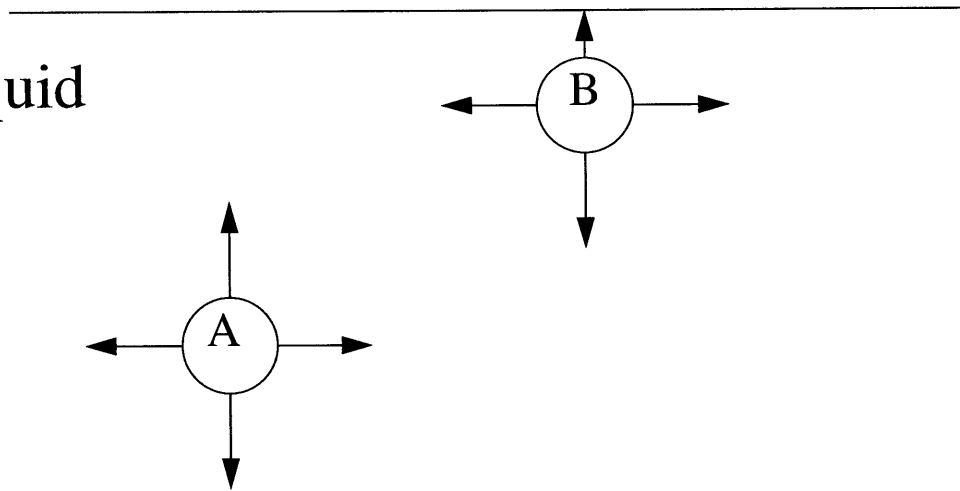
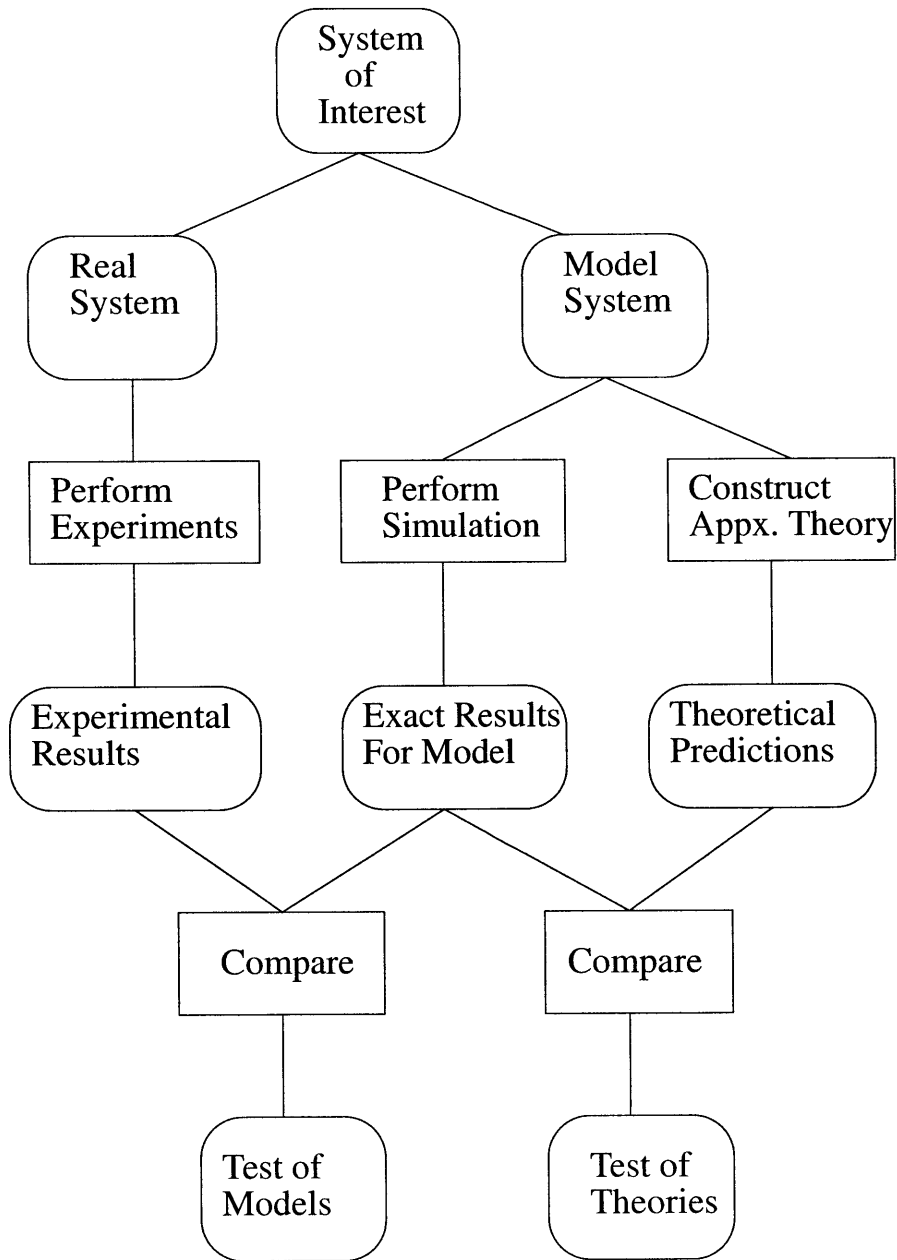


Figure 1-1: Sketch of Liquid-Vapor Interface. (A) a molecule in bulk liquid phase; and (B) a molecule near/at liquid-vapor interface.

interface, etc. To gain knowledge of equilibrium and non-equilibrium properties of these technologically important interfaces, scientists must explore these systems at a molecular and microscopic level.

Since even modern experimental techniques have a limited ability to resolve the many fine details of interfaces, researchers must resort to theoretical calculations or computer simulations to understand interfacial structures and interpret experimental observations. As Figure 1-2 illustrates, experiments, simulations, and theories are closely tied in the study of a physical system. Given a system of interest, researchers can perform real experiments and obtain experimental measurements or they can develop models to mimic the system. Given a specific model, scientists can make mathematical approximations and develop analytical theories, or carry out simulations. On one hand, since results from simulations are exact for a specific model, they can be compared to results from theory to check the accuracy and validity of the approximations made in the theory. This approach is particularly useful for testing the theories for idealized systems such as the hard spheres. On the other hand, comparison between experiment and simulation results can test the accuracy of a given model. An accurate model can in turn provide valuable insight for future experiments. With a model that is capable of reproducing properties of a class of compounds, scientists can predict properties of new materials before actual synthesis or simulate conditions which are hard to achieve experimentally. Hence, simulation serves as a valuable link between an actual experiment and a theory.

Theoretical studies of liquid-vapor interfaces of oligomers and small molecules include van der Waals and Cahn-Hilliard theories, density functional theories based on the gradient expansion, lattice-based mean-field theories, and integral equation theories. The van der Waals^[1] and Cahn-Hilliard^[2] approaches have been used for the liquid-vapor interfaces of atomic fluids and non-polar molecules. Density functional theories have been applied to atomic fluid liquid-vapor interfaces,^[3] liquid-solid interfaces,^[4, 5, 6] wetting and atomic fluid in confined geometries.^[7] Mean-field theory has been applied to polymer free surfaces and polymer-solid interfaces.^[8, 9, 10] Integrated equation theories have been utilized in the study of atomic liquid-solid



Adapted from “Computer Simulation of Liquids”, by Allen & Tildesley

Figure 1-2: Interplay of Experiment, Simulation, and Theory

interfaces.^[7, 11, 12]

With the arrival of ever faster computers, molecular simulation are becoming more feasible and powerful. Molecular simulations such as Monte Carlo (MC) and molecular dynamics (MD) have been successfully applied to many areas, such as the liquid-vapor interfaces of simple atomic systems,^[13, 14, 15, 16, 17, 18] liquid alkali metals and their alloys,^[19, 20] mixtures of Lennard-Jones molecules,^[16, 15, 21] water,^[22] and methanol.^[23] Other applications of molecular simulation to interfacial phenomena include cavitation in octane,^[24] benzene-water interface,^[25] hexane water interface,^[26] hydrocarbons between two impenetrable hard smooth walls,^[27] and hexadecane between two nickel plates.^[28]

Many surfaces of technological and biological importance involve oligomers or polymers. Two classes of compounds, namely, hydrocarbons and perfluorocarbons, have long been subjects of interest because they have a wide range of important applications. Hydrocarbons are the building blocks of the entire petrochemical industry and are vital to our daily life. Perfluorocarbons have been used extensively in retinal detachment and other eye surgeries because of their low surface tensions, high specific gravities and optical clarities.^[29, 30, 31] Short chain perfluorocarbons such as C_3F_8 and C_2F_6 are used as cleaning agents in chemical vapor deposition (CVD) processes.^[32] Perfluorocarbons are being explored as more environmentally friendly alternatives of chlorofluorocarbons (CFC's).^[33] Perfluorocarbons have also been used as artificial blood substitutes because of their ability to dissolve significant amount of oxygen.^[34, 35, 36, 37, 38] In addition, hydrocarbon and perfluorocarbon oligomers also provide good models for industrially important polymers such as polyethylene (PE) and polytetrafluoroethylene (PTFE).

Molecular dynamics simulations and theoretical studies of alkanes (of lengths 10 and 20 carbons),^[39] perfluoroalkanes, and semi-fluorinated alkane diblocks ($CH_3(CH_2)_4(CF_2)_4CF_3$ and $CH_3(CH_2)_9(CF_2)_9CF_3$),^[40] have revealed many features of surface structures. While the simulated surface tensions increase with increasing chain lengths and decreasing temperatures as the experimental data indicate, in all cases tested, the simulations severely overestimate the surface tensions. Fur-

thermore, these simulations are not capable of predicting that perfluorinated alkanes have much lower surface tensions than the corresponding alkanes.^[40] The above studies suggest that a deeper understanding at the microscopic level and better forcefields are necessary to accurately predict surface properties.

1.2 Statement of Objectives

In this study, I aim to:

- (1) Develop and validate new united-atom intermolecular forcefields for linear perfluorocarbons through simulating bulk properties;
- (2) Apply the newly developed forcefields to simulation of liquid-vapor interfaces of linear perfluorocarbons and compare simulated properties, such as surface tension values, with experimental data;
- (3) Investigate the effects of other factors: longer cut-off distance, wider C-C bond angle, and different torsion energy barriers, for simulation of linear perfluorocarbon interfaces; and
- (4) Apply perturbation theory to liquid-vapor interface of linear alkanes and perfluoroalkanes, and estimate the contribution of three-body interactions to surface tension prediction.

The remainder of this thesis is organized to address these objectives in order. Chapter 2 contains a brief overview of the fundamentals and techniques of molecular dynamics simulation used through out this thesis. Chapter 3 reviews the relevant experimental, theoretical, and simulation work. Chapter 4 covers the determination of the Lennard-Jones forcefield parameters for linear perfluoroalkanes in the bulk liquid state. Chapter 5 displays the results of applying the new Lennard-Jones forcefield to liquid-vapor interfaces of linear perfluoroalkanes, and also shows the effects of bond-angle, torsion energy, and cut-off distance on liquid-vapor interfacial properties. Chapter 6 contains an estimate of the three-body contribution to surface tension

values using perturbation theory. Lastly, Chapter 7 summarizes the major findings of this study and suggests directions for future studies.

Chapter 2

Molecular Dynamics Simulation – Fundamentals and Methods

2.1 Foundation of Molecular Dynamics

In a system of N particles, at an instant time t , the positions and momenta of the N particles form a $6N$ -dimensional hyperspace. This hyperspace, called phase space, is composed of two parts: (1) a $3N$ -dimensional configuration space, in which the coordinate axes are the components of the position vectors $\mathbf{r}_i(t)$ and (2) a $3N$ -dimensional momentum space, in which the coordinate axes are the components of the momentum vectors $\mathbf{p}_i(t)$. The positions and momenta of the entire N -particle system at any instance are represented by one point in this space. As the positions and the momenta of the N particles change with time, the point moves, describing a trajectory in phase space.

The essence of molecular dynamics is the realization that, given the forces of interaction and initial conditions, this trajectory in the phase space can be computed by numerically solving the Newtonian classical equation of motion, or equivalently, Hamiltonian equation of motion, or more generalized, Lagrangian equation of motion for systems with internal constraints.

2.1.1 Newtonian Equation of Motion

In the Newtonian view, motion is a response to an applied force. The translational motion of a spherical particle and the force externally applied to the i th particle are explicitly related through Newton's equation of motion:

$$\mathbf{F}_i = m \frac{d^2 \mathbf{r}_i}{dt^2} \quad (2.1)$$

where m is the mass of the particle and \mathbf{r}_i is a position vector. For N spherical particles, it represents $3N$ second-order, ordinary differential equations of motion.

2.1.2 Hamiltonian Equation of Motion

Contrast to the Newtonian view, Hamiltonian dynamics does not express the applied force explicitly. Instead, motion occurs in such a way to preserve some function of positions and velocities, called the Hamiltonian H , whose value is constant,

$$H(\mathbf{r}^N, \mathbf{p}^N) = \text{constant} \quad (2.2)$$

where \mathbf{p}_i is the momentum of the i th particle, defined in terms of its velocity by

$$\mathbf{p}_i = m \dot{\mathbf{r}}_i \quad (2.3)$$

Since the Hamiltonian is constant and has no explicit time dependence, one obtains, by taking the total time derivative of Eq. (2.2)

$$\frac{dH}{dt} = \sum_i \frac{\delta H}{\delta \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i + \sum_i \frac{\delta H}{\delta \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i = 0 \quad (2.4)$$

For an isolated system, the total energy E is conserved and equals the Hamiltonian. E is the sum of kinetic energy K and potential energy U .

$$H(\mathbf{r}^N, \mathbf{p}^N) = E = K + U$$

$$= \frac{1}{2m} \sum_i \mathbf{p}_i^2 + U(\mathbf{r}^N) \quad (2.5)$$

Taking the total time derivative of Eq. (2.5), one obtains

$$\frac{dH}{dt} = \frac{1}{m} \sum_i \mathbf{p}_i \cdot \dot{\mathbf{p}}_i + \sum_i \frac{\delta U}{\delta \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i = 0 \quad (2.6)$$

Comparing Eq. (2.4) and Eq. (2.6), one obtains that for each molecule i ,

$$\frac{\delta H}{\delta \mathbf{p}_i} = \frac{\mathbf{p}_i}{m} = \dot{\mathbf{r}}_i \quad (2.7)$$

Combining Eq. (2.4) and Eq. (2.7), and satisfying the condition that all velocities are independent of each other, one obtains,

$$\frac{\delta H}{\delta \mathbf{r}_i} = -\dot{\mathbf{p}}_i \quad (2.8)$$

Eq. (2.7) and Eq. (2.8) are the Hamiltonian equations of motion for an isolated system. For a system of N spherical particles, they represent a set of $6N$ first-order differential equations and are equivalent to Newton's $3N$ second-order equations. In the cases where the system can exchange energy with its surrounding, H no longer equals the system's total energy E , but instead contains extra terms to account for the energy exchanges. H is still conserved, but E is not.

2.1.3 Lagrangian Equation of Motion

Lagrangian dynamics is the most general form of equation of motion and covers all previous versions of equations of motion.^[41, 42, 43] In cases when the systems have internal constraints (e.g., rigid bond) which give extra terms in the form of internal forces, Lagrangian formulation solves dynamics problems in the most economical coordinates by selecting the coordinates that do not violate the physical constraints of the systems. Using Newtonian dynamics, one can include the applied \mathbf{F}_i and constraint force \mathbf{C}_i , and obtain:

$$\sum_{i=1}^N \mathbf{F}_i \cdot \delta \mathbf{r}_i + \sum_{i=1}^N \mathbf{C}_i \cdot \delta \mathbf{r}_i - \sum_{i=1}^N m_i \mathbf{a}_i \cdot \delta \mathbf{r}_i = 0 \quad (2.9)$$

Since the work of the constraint force caused by displacement is zero, Eq. (2.9) reduces to

$$\sum_{i=1}^N \mathbf{F}_i \cdot \delta \mathbf{r}_i - \sum_{i=1}^N m_i \mathbf{a}_i \cdot \delta \mathbf{r}_i = 0 \quad (2.10)$$

The old coordinates are transformed into a set of new independent generalized coordinates $q_j (j = 1, 2, \dots, M)$, where

$$\begin{aligned} \mathbf{r}_i &= \mathbf{r}_i(q_1, q_2, q_3, \dots, q_M, t) \\ i &= 1, 2, 3, \dots, N \end{aligned} \quad (2.11)$$

The generalize force Q_j is defined to be

$$Q_j \equiv \sum_{i=1}^N \mathbf{F}_i \cdot \frac{\delta \mathbf{r}_i}{\delta q_j} \quad (2.12)$$

Combining Equations (2.10)(2.11)(2.12), we obtain,

$$\sum_{j=1}^M Q_j \delta q_j = \sum_{j=1}^M \left(\sum_{i=1}^N m_i \mathbf{a}_i \cdot \frac{\delta \mathbf{r}_i}{\delta q_j} \right) \delta q_j \quad (2.13)$$

Utilizing the definition of kinetic energy K , in terms of velocity, and the fact that all the generalized coordinates are independent, after much mathematical manipulation,^[41] one obtains the Lagrange's equation of motion:

$$\frac{d}{dt} \left(\frac{\delta L}{\delta \dot{q}_j} \right) - \frac{\delta L}{\delta q_j} = 0 \quad (j = 1, 2, \dots, M) \quad (2.14)$$

where the lagrangian L is defined as

$$L = K - U \quad (2.15)$$

The number of equations of motion are equal to the number of degrees of freedom

M of the system, which is the minimum number of independent coordinates needed to completely describe the position of the system. Lagrange's equations of motion are simpler to use than Newton's equations because there are fewer unknowns to solve for, and fewer equations of motion to deal with.

2.2 Integration

From a mathematical point of view, molecular dynamics is an initial value problem. The classical tool for solving the initial value problems are finite-difference methods. The following subsections contain a summary of the numerical integration scheme used in this study.

2.2.1 Verlet Algorithm

Original Verlet Algorithm

Verlet algorithm is the most widely used finite difference integration method for molecular dynamics. It is a direct solution of the second-order equations (Eq. (2.1)). It results from a combination of two Taylor expansions:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \frac{d\mathbf{r}(t)}{dt}\Delta t + \frac{1}{2}\frac{d^2\mathbf{r}(t)}{dt^2}\Delta t^2 + \frac{1}{3!}\frac{d^3\mathbf{r}(t)}{dt^3}\Delta t^3 + O(\Delta t^4) \quad (2.16)$$

and

$$\mathbf{r}(t - \Delta t) = \mathbf{r}(t) - \frac{d\mathbf{r}(t)}{dt}\Delta t + \frac{1}{2}\frac{d^2\mathbf{r}(t)}{dt^2}\Delta t^2 - \frac{1}{3!}\frac{d^3\mathbf{r}(t)}{dt^3}\Delta t^3 + O(\Delta t^4) \quad (2.17)$$

Adding Eq. (2.16) and Eq. (2.17) eliminates all odd-order terms and yields the Verlet algorithm for positions:

$$\mathbf{r}(t + \Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \frac{d^2\mathbf{r}(t)}{dt^2}\Delta t^2 + O(\Delta t^4) \quad (2.18)$$

Verlet algorithm utilizes the position $\mathbf{r}(t)$, acceleration $\frac{d^2\mathbf{r}(t)}{dt^2}$, and the position

$\mathbf{r}(t-\Delta t)$ from the previous step. The local truncation error is on the order of $(\Delta t)^4$.

Velocity is not needed for computing the trajectories but is useful for estimating the kinetic energy of the system. Velocity can be estimated as:

$$\mathbf{v}(t) \approx \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t)}{2\Delta t} \quad (2.19)$$

The Verlet algorithm has been shown to have excellent stability for relatively large time steps.^[44, 45, 46] However, it suffers several deficiencies. First, it is not self-starting. It estimates $\mathbf{r}(t + \Delta t)$ from the current position $\mathbf{r}(t)$ and the previous position $\mathbf{r}(t - \Delta t)$. To begin a calculation, special technique such as the backward Euler method must be used to get $\mathbf{r}(-t)$. Second deficiency of the Verlet algorithm is that, conflicting with the view that phase-space trajectory depends equally on position $\mathbf{r}(t)$ and velocity \mathbf{v} , it purely rely on positions. Velocities are not explicitly included in the integration and hence this method requires extra computation and storage effort. A modified version, namely the velocity version of Verlet algorithm, averts these two drawbacks.

Velocity Version of the Verlet Algorithm

Swope, Andersen, Berens, and Wilson^[47] proposed the velocity version of Verlet algorithm which takes the form

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t\mathbf{v}(t) + \frac{1}{2}\Delta t^2\frac{d^2\mathbf{r}(t)}{dt^2} \quad (2.20)$$

and

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{1}{2}\Delta t\left[\frac{d^2\mathbf{r}(t)}{dt^2} + \frac{d^2\mathbf{r}(t + \Delta t)}{d(t + \Delta t)^2}\right] \quad (2.21)$$

The velocity version of Verlet is simple, compact, stable, and more commonly used compared to the original Verlet algorithm.

2.2.2 SHAKE and RATTLE

For polyatomic systems where there are internal constraints, the regular Verlet in-

tegration method is not sufficient. As mentioned in the earlier section on Lagrangian dynamics, a set of independent generalized coordinates must be constructed to obey the constraint-free equations of motion. SHAKE^[48, 49, 50] and RATTLE^[51] are two methods that deal with Verlet integration with internal constraints. SHAKE corresponds to the original position-oriented Verlet formulation, and RATTLE corresponds to the velocity version of the Verlet algorithm.

SHAKE is a procedure that approaches internal constraints by going through the constraints one by one, cyclically, adjusting the coordinates to satisfy each constraints in turn. The procedure is repeated until all constraints are satisfied to within a specified tolerance level.

RATTLE is a modification of SHAKE, based on the velocity version of Verlet algorithm^[51]. It calculates the positions and velocities at the next timestep from the positions and velocities at the present timestep, without requiring information about the earlier history. Like SHAKE, it retains the simplicity of using Cartesian coordinates for each of the atoms to describe the configuration of a molecules with internal constraints. It guarantees that the coordinates and velocities of the atoms within a molecule satisfy the internal constraints at each timestep.

RATTLE has two advantages over SHAKE: (1) on computers of fixed precision, it is of higher precision than SHAKE; (2) since RATTLE deals directly with the velocities, it is easier to modify RATTLE for use with constant temperature and constant pressure molecular dynamics methods and with the non-equilibrium molecular dynamics methods that make use of rescaling of the atomic velocities. RATTLE is used to conduct integrations in this study.

2.3 Simulation Techniques

Some useful techniques frequently employed in molecular dynamics simulations are introduced here. For deeper understanding of these techniques, one should refer to the references listed.^[48, 43]

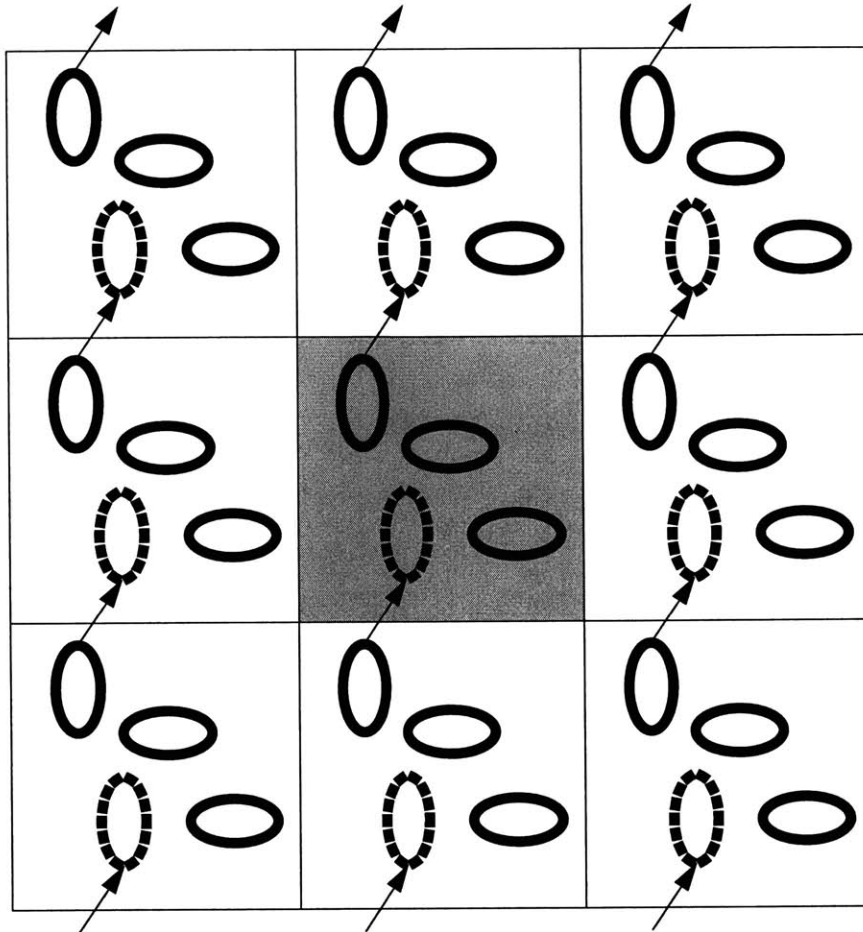


Figure 2-1: A two-dimensional periodic system

2.3.1 Simulation Cell and Periodic Boundary Conditions

To implement MD, one must first define a simulation cell containing N (N usually on the order of 100 – 10,000) particles, and specify their interactions. A simulation cell of this size is usually dominated by surface effect. When simulating the behavior of bulk liquids, to overcome the effect of this artificial cell surface, the cell is considered to be surrounded by replicas of itself. [52] In order to conserve the number density in the central cell, periodic boundary conditions are needed. That is, when a particle leaves the central cell and enters one of the surrounding replicas, its image enters the central cell from the opposite surrounding replica. Figure 2-1 illustrates a two-dimensional periodic system. Particles can enter and leave each box across each of the four sides in the 2D-example. In a three-dimensional case, particles would be free

to cross any of the six neighboring faces.

The limitation of the periodic boundary condition is that it will neglect any density waves with wavelengths longer than the side length of the simulation cell. For liquid systems far from the critical point and for which the interactions are short-ranged, periodic boundary conditions are found to be very useful.

2.3.2 Cutoff Distance and Mean-Field Correction

In principle, to calculate the force acting on a specified particle, one needs to account for interactions between this particle and all other particles in the simulation box. In reality, the concept of “cut-off distance” R_c is introduced to decrease the required computational effort, as the potential and force to be calculated are mainly contributed by particles within a close distance. Only interactions within a sphere of radius R_c from each particle are included. The radial distribution function of regions beyond the cut-off is set to unity and mean-field theory is often used to provide corrections to properties from interactions beyond the cut-off distance. In cases where the interaction beyond R_c is significant, double cut-off and multiple timesteps can be used in MD simulations. Two cut-off distances – an inner and an outer one – are assigned. The inner cut-off distance defines a spherical region where interactions with the center atom are treated as usual and updated every timestep. The outer cut-off distance defines a shell outside the spherical region. The interactions with the center atom in the outer shell are updated much less frequently. The underlying assumption for this techniques is that contribution from outer shell varies at a much slower rate. This double cut-off/multiple timestep technique allows gaining computational accuracy at a minimal increase of cost.

2.3.3 Temperature Calculation and Control

Temperature is a measurable macroscopic property of the system, and can be calculated from microscopic details of the system. In addition, in many cases, to mimic experimental conditions, one must be able to maintain the temperature of the

system constant during the MD simulation.

Temperature Calculation

Temperature is related to the average kinetic energy of the system through the equipartition principle which states that every degree of freedom has an average energy of $k_B T/2$ associated with it. Hence one obtains:

$$\left\langle \sum_i^N \frac{\mathbf{p}_i^2}{2m} \right\rangle = \left\langle \sum_i^N \frac{m\mathbf{v}_i^2}{2} \right\rangle = \langle K \rangle = \frac{f k_B T}{2} \quad (2.22)$$

and

$$T = \langle T_{inst} \rangle = \left\langle \frac{2}{(3N - N_c)k_B T} \sum_i^N \frac{m_i \mathbf{v}_i^2}{2} \right\rangle \quad (2.23)$$

where f is the degrees of freedom, N is the number of the particles, N_c is the number of constraints on the ensemble, m_i is the mass of the i th particle, T_{inst} is the instantaneous kinetic temperature, T is the temperature, and k_B is the Boltzmann's constant.

Temperature and the distribution of velocities in a system are related through the Maxwell-Boltzmann expression:

$$f(v)dv = \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-\frac{mv^2}{2k_B T}} 4\pi v^2 dv \quad (2.24)$$

which calculates the probability $f(v)$ that a particle of mass m has a velocity of v when it is at temperature T .

Maintaining Constant Temperature

Even if the initial velocities are generated according to the Maxwell-Boltzmann distribution at the desired temperature, the velocity distribution will not remain constant as the simulation continues. To maintain the correct temperature, the computed velocities needed to be adjusted. Besides getting the temperature to the right target, the temperature-control mechanism should also produce the correct statistical ensem-

bles. Several methods for temperature control have been developed. They are (1) stochastic method, (2) extended system method, (3) direct velocity scaling method, and (4) Berendsen method.^[53]

Stochastic Method

A system corresponding to the canonical ensemble is one that involves interactions between the particles of the system and the particles of a heat bath at a specified temperature. Exchange of energy occurs across the system boundaries. At intervals, the velocity of a randomly chosen particle is reassigned with a value selected according to the Maxwell-Boltzmann distribution. This process corresponds to a system particle colliding with a heat bath particle. When such a collision takes place, the system jumps from one constant energy surface onto a different constant energy surface. In this fashion, the system samples all the important regions of the phase space and generate an irreducible Markov chain corresponds to that of the canonical ensemble.^[54]

If the collisions take place very frequently, it will slow down the speed at which the particles in the system explore the configuration space. If the collisions occur too infrequently, the canonical distribution of energy will be sampled too slowly. For a system to mimic a volume element in real liquid in thermal contact with the heat bath, a collision rate

$$R_{particle\ collision} = \frac{\lambda^T}{\rho^{1/3} N^{2/3}} \quad (2.25)$$

is suggested by Andersen^[54] where λ_T is the thermal conductivity, N is the number of particles, and ρ is the liquid density.

Instead of changing the velocity of the particles one at a time as described above, massive stochastic collision method assigns the velocities of all the particles at the same time at a much less frequency at equally-spaced time intervals.

Extended System Method

Another way to describe the dynamics of a system in contact with a heat bath is to add an extra degree of freedom to represent the heat bath and carry out a simulation of this “extended system”. The heat bath has a “thermal inertia” and energy is allowed to flow between the bath and the system.^[55] The extra degree of

freedom is denoted s and it has a conjugate momentum p_s . The real particle velocity is

$$\mathbf{v} = s\dot{\mathbf{r}} = \frac{\dot{\mathbf{p}}}{\mathbf{m}}s \quad (2.26)$$

The extra potential energy associated with s is

$$U_s = (f + 1)k_B T \ln s \quad (2.27)$$

where f is the number of degrees of freedom and T is the specified temperature.

The kinetic energy associated with s is

$$K_s = \frac{1}{2}Q\dot{s}^2 = \frac{p_s^2}{2}Q \quad (2.28)$$

where Q is the thermal inertia parameter in units of $(energy)(time)^2$ and controls the rate of temperature fluctuations.

The extended system Hamiltonian

$$H_s = K + K_s + U + U_s \quad (2.29)$$

is conserved and the extended system density function

$$\rho_{NVE_s}(\mathbf{r}, \mathbf{p}, s, p_s) = \frac{\delta(H_s - E_s)}{\int d\mathbf{r}d\mathbf{p}dSdp_s\delta(H_s - E_s)} \quad (2.30)$$

Integration over s and p_s leads to a canonical distribution of the variables \mathbf{r} and \mathbf{p}/s .^[55]

The parameter Q is often chosen by trial and error. If Q is too high, the energy flow between the system and the heat bath is slow, If Q is too low, there exists long-lived, weakly damped oscillation of the energy, resulting in poor equilibration.^[55]

Direct Velocity Scaling Method

This method involves rescaling the velocities of each particle at each timestep by a

factor of $(T_{target}/T_{current})^{1/2}$ where T_{target} is the desired thermodynamic temperature and $T_{current}$ is the current kinetic temperature. Even though this method transfer energy to/from the system very efficiently, ultimately the speed of this method depends on the potential energy expression, the parameters, the nature of the coupling between the vibrational, rotational, and translational modes, and the system sizes, because the fundamental limitation to achieving equilibrium is how rapidly energy can be transferred to/from/among the various internal degrees of freedoms of the molecules.

Berendsen Method of Temperature Coupling

Berendsen method is a refined approach to velocity rescaling.^[56] Each velocity is multiplied by a factor χ at each time step Δt

$$\chi = \left(1 + \frac{\Delta t}{\tau} \left(\frac{T_{target}}{T_{current}} - 1\right)\right)^{1/2} \quad (2.31)$$

where $T_{current}$ is the current kinetic temperature, T_{target} is the desired thermodynamic temperature, and τ is a preset time constant. This method forces the system towards the desired temperature at a rate determined by τ , while only slightly perturbing the forces on each molecule.

2.3.4 Pressure Calculation and Control

Pressure is another thermodynamic property that we would like to calculate and control.

Pressure Calculation

Pressure is a tensor:

$$\underline{\underline{\mathbf{P}}} = \begin{bmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{yx} & P_{yy} & P_{yz} \\ P_{zx} & P_{zy} & P_{zz} \end{bmatrix} \quad (2.32)$$

Each element of the pressure tensor is the force acting on the surface of an infinitesimal cubic volume that has edges parallel to the x, y, and z axes. The first subscript

denotes the direction of the normal to the plane on which the force acts, and the second subscript denotes that direction of the force.

Pressure is contributed by two components: (1) the momentum carried by the particles as they cross the surface area and (2) the momentum transferred as a result of forces between interacting particles that lie on different sides of the surface. Hence, $\underline{\underline{\mathbf{P}}}$ can be expressed as

$$\underline{\underline{\mathbf{P}}} = \frac{1}{V} \left[\sum_{i=1}^N m_i \mathbf{v}_i \cdot \mathbf{v}_i^T + \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i^T \right] \quad (2.33)$$

where

$$\sum_{i=1}^N m_i \mathbf{v}_i \cdot \mathbf{v}_i^T = \begin{bmatrix} \sum_i m_i v_{ix} \cdot v_{ix} & \sum_i m_i v_{ix} \cdot v_{iy} & \sum_i m_i v_{ix} \cdot v_{iz} \\ \sum_i m_i v_{iy} \cdot v_{ix} & \sum_i m_i v_{iy} \cdot v_{iy} & \sum_i m_i v_{iy} \cdot v_{iz} \\ \sum_i m_i v_{iz} \cdot v_{ix} & \sum_i m_i v_{iz} \cdot v_{iy} & \sum_i m_i v_{iz} \cdot v_{iz} \end{bmatrix} \quad (2.34)$$

and

$$\sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i^T = \begin{bmatrix} \sum_i r_{ix} \cdot f_{ix} & \sum_i r_{ix} \cdot f_{iy} & \sum_i r_{ix} \cdot f_{iz} \\ \sum_i r_{iy} \cdot f_{ix} & \sum_i r_{iy} \cdot f_{iy} & \sum_i r_{iy} \cdot f_{iz} \\ \sum_i r_{iz} \cdot f_{ix} & \sum_i r_{iz} \cdot f_{iy} & \sum_i r_{iz} \cdot f_{iz} \end{bmatrix} \quad (2.35)$$

where $r_{i\alpha}$, $v_{i\alpha}$, and $f_{i\alpha}$ indicate the α ($\alpha = x, y, \text{ or } z$) components of the position, velocity, and force vector of the i th particle, respectively. In an isotropic situation, the pressure tensor is diagonal, and the instantaneous hydrostatic pressure is calculated as

$$P = \frac{1}{3} (P_{xx} + P_{yy} + P_{zz}). \quad (2.36)$$

Maintaining Constant Pressure

Pressure can be adjusted by changing the volume of the simulation box. Two common methods for manipulating pressure while maintaining the shape of the simulation box have been developed: (1) Berendsen method and (2) extended system method (Andersen method).

Berendsen Method

Berendsen method^[57] couples the system to a “pressure bath” to maintain the

pressure at a target P_{target} . The strength of the coupling is determined by the isothermal compressibility of the system β_T and a relaxation time constant τ . At each timestep ΔT , the x , y , and z coordinates of each particle are scaled by the factor χ

$$\chi = [1 - \beta_T \frac{\Delta t}{\tau} (P_{current} - P_{target})]^{1/3} \quad (2.37)$$

where $P_{current}$ is the instantaneous pressure.

Extended System Method

Andersen^[54] proposed a method to treat the volume of the simulation unit as a dynamic variable in an extended system. This mimics the action of a piston of mass M on a real system. The Lagrangian of the extended system is modified so that it contains a kinetic energy term associated with the piston and a potential term PV where P is the desired pressure.

The piston mass M is an adjustable parameter in Andersen's method. A low M will result in rapid box size oscillations, which are not damped very efficiently by the random motions of the particles. A large mass will give rise to slow exploration of volume space. To mimic event in a small volume of real liquid, Andersen^[54] recommends the time scales for the box-volume fluctuations should be roughly the same as the time scale for a sound wave crossing the simulation box.

2.4 Forcefields

This section provide an overview of the fundamental of forcefields, followed by an introduction to various types of forcefield related to chain molecules.

2.4.1 Fundamentals of Forcefields

The Schrödinger equation,

$$\hat{H}\Psi(\mathbf{R}, \mathbf{r}) = \mathbf{E}\Psi(\mathbf{R}, \mathbf{r}) \quad , \quad (2.38)$$

where \hat{H} is the Hamiltonian operator, Ψ is the wave function, E is the total energy, \mathbf{R} is the vector containing the $3N$ coordinates of the nuclei, and \mathbf{r} is the vector of the electrons' coordinates, is a complete mathematical description of a molecule. Since electrons are several thousand times lighter than the nuclei and move much faster than the nuclei, the Born-Oppenheimer approximation can be used to decouple the motion of the electrons from that of the nuclei. Two separate equations are thus derived from Eq. (2.38):

$$\hat{H}_{el}\psi(\mathbf{r}; \mathbf{R}) = \mathbf{E}\psi(\mathbf{r}; \mathbf{R}) \quad (2.39)$$

and

$$\hat{H}_{nucl}\Phi(\mathbf{R}) = \mathbf{E}\Phi(\mathbf{R}) \quad , \quad (2.40)$$

where ψ is the electronic wave function and Φ is the nuclear wave function. ψ only parametrically depends on the nuclear positions.^[58, 59]

$$\hat{H}_{nucl} = \sum_{i=1}^N \frac{|\hat{P}_i|^2}{2m_i} + E_0(\mathbf{R}) \quad , \quad (2.41)$$

where \hat{P}_i and m_i are the momentum operator and the mass of the i th nucleus, respectively.

Eq. (2.39) describes the motion of the electrons only and Eq. (2.40) describes the motion of the nuclei only. $E_0(\mathbf{R})$ is the potential energy surface and is only a function of the position of the nuclei.

In principle, Eq. (2.39) and Eq. (2.40) can be solved for \mathbf{E} and \mathbf{R} . However, this process is often extremely demanding mathematically, hence further approximations are often made. An empirical fit to the potential energy surface, called a forcefield or potential, is usually used instead of solving Eq. (2.39), and Newton's equation of motion is used instead of Eq. (2.40) based on the fact that non-classical effects are extremely small for the relatively heavy nuclei.

2.4.2 Various Types of Forcefield

The forcefields used for describing chain molecules are usually of the following

general form:

$$\begin{aligned}
 U_{total\ forcefield} &= U_{bond-stretching} + U_{bond-angle-bending} + U_{torsional-angle\ rotation} \\
 &+ U_{out-of-plane\ interaction} + U_{off-diagonal\ coupling} + U_{van\ der\ Waals} \\
 &+ U_{electrostatic} + U_{hydrogen\ bond} \quad .
 \end{aligned}
 \tag{2.42}$$

As Figure 2-2 illustrates, the total forcefield includes a combination of internal coordinates to describe the bond part (term (1) bond-stretching; (2) bond-angle-bending; (3) torsional-angle rotation; (4) out-of-plane interaction; (5)–(11) off-diagonal coupling) in Figure 2-2) of the potential surface, and external coordinates to describe the van der Waals, hydrogen bonding, and electrostatic interactions (term (12) in Figure 2-2) between atoms. Depending on the property one is interested in studying, not all terms are required for every simulation. For instance, the bond-stretching term is often omitted to reduce computational effort since it is not essential in reproducing thermodynamic properties. The off-diagonal terms that represent the couplings between deformations of internal coordinates are only necessary for accurately reproducing experimental vibrational frequencies, which are the dynamic properties of a molecule.^[58]

The bond-stretching and bond-bending terms are usually quadratic and based on considering the bonds and angles as Hookean springs. They have the forms:

$$U_{bond-stretching} = \sum_{bonds} K_r (r - r_{eq})^2 \quad ,
 \tag{2.43}$$

and

$$U_{bond-angle-bending} = \sum_{bond-angles} K_\theta (\theta - \theta_{eq})^2 \quad .
 \tag{2.44}$$

These quadratic expressions have satisfactorily described the structures and energies for relatively unstrained proteins and nucleic acids.^[60]

The torsional-angle-rotation term can often be represented as a Fourier series expansion and can provide a relatively accurate representation of the conformational

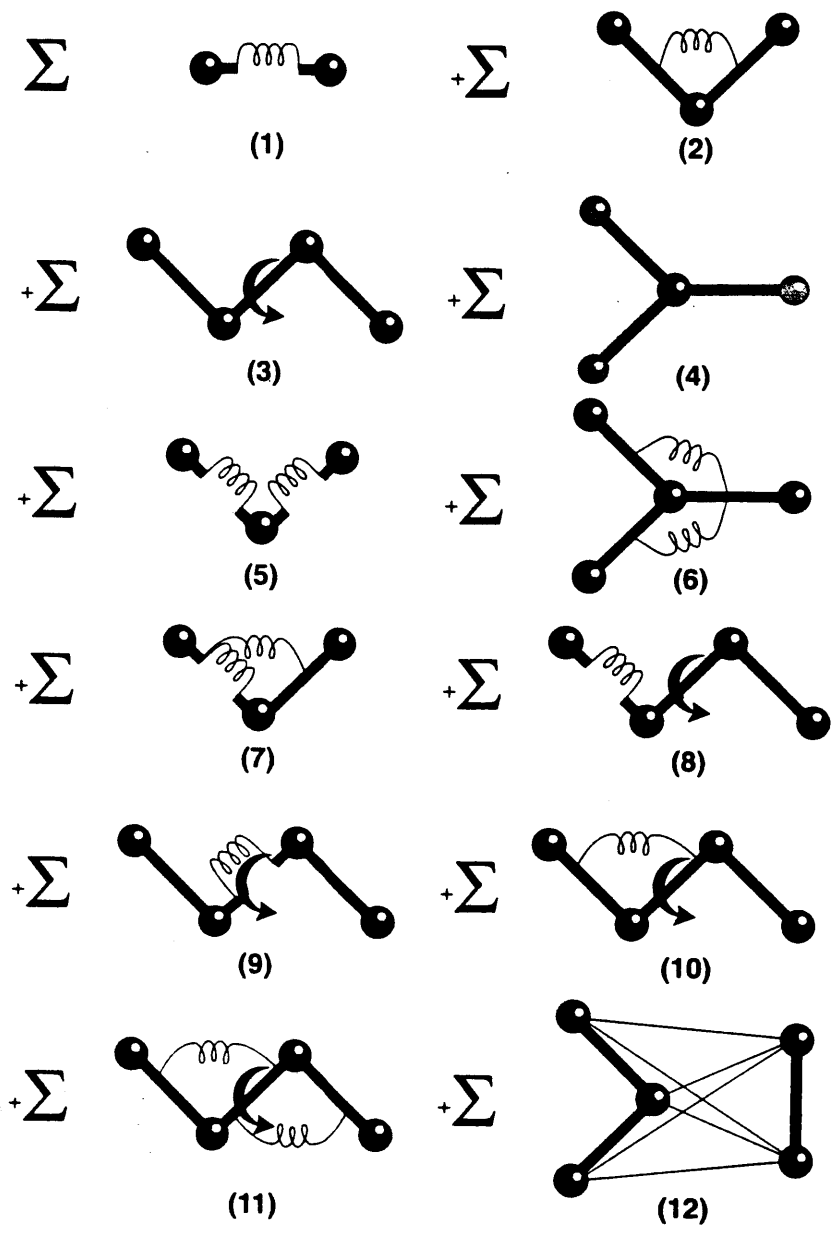


Figure 2-2: Graphic illustration of intra- and inter- molecular interactions (Adapted from Discover User Guide Version 2.9/3.1)

isomers of simple and complex molecules:^[60]

$$U_{\text{torsional angle rotation}} = \sum_{\text{dihedral angles}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \quad . \quad (2.45)$$

The electrostatic terms obey the classical Coulombic point-charge interaction

$$U_{\text{electrostatic}} = \sum_{i < j} \frac{q_i q_j}{\epsilon r_{ij}} \quad . \quad (2.46)$$

The van der Waals dispersion is typically represented by a Lennard-Jones 12-6 forcefield

$$U_{\text{van der Waals}} = \sum_{i < j} \left[\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^6} \right] \quad . \quad (2.47)$$

where r_{ij} is the interparticle distance.

The effect of hydrogen bonding is inherent in the electrostatic and van der Waals interactions, but some researchers often include an explicit hydrogen bonding energy^[61]

$$U_{\text{hydrogen bond}} = \sum_{\text{H-bonds}, i < j} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{10}} \right] \quad (2.48)$$

in hopes that it will better represent the real systems.

In order to greatly reduce the computational effort, united-atom (UA) model is often used instead of explicit-atom (EA) model. In a UA model, a group of atoms (e.g., CH_3 , CH_2 , CF_2 , or CF_3) is considered as a condensed “united-atom”. The parameters of the united-atom are determined from fitting to thermodynamics properties. This grouping significantly cuts down the number of “atoms” in the simulation and makes computation less demanding. The UA approach allows the simulation of larger molecules while preserving the main features of the molecules such as the torsional angles and the backbone bond angles.

The functional forms of forcefields may vary from the ones listed above. In the next chapter, there will be a review of the forcefields relevant to hydrocarbons and perfluorocarbons. In summary, the role of a forcefield is to describe an entire class of molecules with reasonable accuracy. A forcefield is parameterized by an empirical

fit to experimental data for a limited set of molecules and then used to calculate properties for a larger set of related molecules and structures.

Chapter 3

Previous Work

In this chapter, experimentally measured bulk and surface, structural and thermodynamical properties of alkanes, perfluoroalkanes, PE, and PTFE in literature will be reviewed first. Next, previous simulations of bulk and liquid-vapor interfaces will be discussed. Finally, a review of forcefields related to hydrocarbon and perfluorocarbon is included.

3.1 Experimental

There are numerous experimental data available for alkanes and polyethylenes. Perfluoroalkanes and polytetrafluoroethylene, are also relatively well-studied experimentally, although not as well as their hydrocarbon counterparts. The experimental data compiled in this section will serve three purposes for this study: (1) to establish conditions under which simulations will be conducted, (2) to provide basis for developing new forcefields, and (3) to be compared with simulated properties and evaluate the quality of the models used in the simulations.

3.1.1 Melting and Boiling Points

Table 3.1^[62] and Table 3.2^[63, 62] list the molecular weight, melting point, and normal boiling point values of linear alkanes and linear perfluoroalkanes, respectively,

Table 3.1: Boiling Point and Melting Point of Linear Alkanes

Compound	MW (<i>g/mol</i>)	Melting Point (<i>K</i>)	Boiling Point (<i>K</i>)
Methane	16.04	89	112
Ethane	30.07	101	185
Propane	44.10	83	231
Butane	58.12	138	273
Pentane	72.15	142	309
Hexane	86.18	179	342
Heptane	100.20	182	371
Octane	114.23	216	399
Nonane	128.26	220	424
Decane	142.28	242	447

Table 3.2: Boiling Point and Melting Point of Linear Perfluoroalkanes

Compound	MW (<i>g/mol</i>)	Melting Point (<i>K</i>)	Boiling Point (<i>K</i>)
Perfluoromethane	88.00	89	145
Perfluoroethane	138.01	167	195
Perfluoropropane	188.02	90	235
Perfluorobutane	238.03	188	272
Perfluoropentane	288.04	148	302
Perfluorohexane	338.04	187	330
Perfluoroheptane	388.05	193	355
Perfluorooctane	438.06	208	377
Perfluorononane	488.07	257	398
Perfluorodecane	538.07	309	417
Perfluoroundecane	588.08	330	434
Perfluorododecane	638.09	348	451
Perfluorohexdecane	838.12	398	505

in the order of increasing chain lengths. Because of the high atomic weight of the F atoms, molecular weight of perfluoroalkanes are significantly larger than their alkane counterparts. Despite this difference in molecular weight, the boiling points of perfluoroalkanes are surprisingly similar to those of the corresponding alkanes. This phenomenon suggests that the intermolecular forces between perfluoroalkane molecules are relatively weak. Temperatures between the melting points and boiling points of the compounds will be chosen for all the simulations conducted hereafter.

3.1.2 Densities

Densities of liquid alkanes and perfluoroalkanes are listed in Table 3.3 and Table 3.4^[63, 64]. Due to the high molecular weight of F atom, densities of liquid perfluoroalkanes are much higher than their alkane counterparts. For both alkanes and perfluoroalkanes, density decreases as temperature increases. These density data will play an important role in the parameterization and validation of the forcefield in the chapters to follow.

3.1.3 Heats of Vaporization

Heats of vaporization for alkanes and perfluoroalkanes are listed in Table 3.5 and Table 3.6, respectively.^[65, 66] These heats of vaporization data will play important roles in the parameterization and validation of the forcefield in the chapters to follow.

3.1.4 Polarizability

The polarizability of some linear alkanes and linear perfluoroalkanes are available.^[67, 66]

The values listed in Table 3.7 and Table 3.8 will be used in Chapter 6 for calculating the three-body interaction coefficients.

3.1.5 Molecular Structures

Crystallographic studies of *n*-perfluoroalkanes and polytetrafluoroethylene have revealed that they depart from structures of *n*-alkanes and polyethylene.^[68] As Fig-

Table 3.3: Densities of Alkanes.

Compound	$D_4^{20}(g/ml)$	$\frac{dD}{dt}(g/(ml \cdot ^\circ C))$	t Range ($^\circ C$)
$n - C_5H_{12}$	0.62638	-0.0009686	10 – 30
$n - C_6H_{14}$	0.65942	$-0.0008790(1+0.0000764t)$	-90 – 50
$n - C_7H_{16}$	0.68375	$-0.0008411(1+0.0009019t)$	-90 – 90
$n - C_8H_{18}$	0.70283	$-0.0008096(1+0.000857t)$ -0.0008268	-50 – 110 10 – 30
$n - C_9H_{20}$	0.71790	$-0.0007847(1+0.000668t)$ -0.0007952	-50 – 150 10 – 30
$n - C_{10}H_{22}$	0.72985	$-0.0007675(1+0.00511t)$	-30 – 170
$n - C_{11}H_{24}$	0.7404	$-0.0007061(1+0.000617t)$ -0.0007150	-10 – 190 10 – 30
$n - C_{12}H_{26}$	0.7493	$-0.0006942(1+0.000617t)$ -0.0007023	-10 – 190 10 – 30
$n - C_{13}H_{28}$	0.7568	-0.0007104	0 – 99
$n - C_{14}H_{30}$	0.7636	-0.0007069	0 – 100
$n - C_{15}H_{32}$	0.7688	-0.0006962	15 – 100
$n - C_{16}H_{34}$	0.7739	-0.0006932	-10 – 190
$n - C_{20}H_{42}$	0.7777 at $36.4^\circ C$	-0.0006759	36.4 – 100

Table 3.4: Densities of Perfluoroalkanes, t = temp. in ($^{\circ}C$), T = temp. in K

Compound	Density (g/ml)	Temp. Range
$n - C_4F_{10}$	$1.6484 + 3.18(10^{-3})(259.88 - T) - 0.0003(10^{-6})(259.88 - T)^2$	
$n - C_5F_{12}$	$1.6195 + 0.003375(293.16 - T) - 6.374(10^{-6})(293.16 - T)^2$	258 - 293 (K)
$n - C_6F_{14}$	$1.7450 - 2.847(10^{-3})t - 3.324(10^{-6})t^2$	15 - 45 ($^{\circ}C$)
$n - C_7F_{16}$	$1.801 - 0.00240t$	0 - 60 ($^{\circ}C$)
$n - C_8F_{18}$	$1.839 - 0.00251t$	0 - 60 ($^{\circ}C$)
$n - C_9F_{20}$	$1.860 - 0.00248t$	0 - 60 ($^{\circ}C$)
$n - C_{10}F_{22}$	$1.873 - 0.00230t$	35 - 60 ($^{\circ}C$)
$n - C_{11}F_{24}$	$1.919 - 0.00248t$	60 - 130 ($^{\circ}C$)
$n - C_{12}F_{26}$	$1.961 - 0.00256t$	80 - 175 ($^{\circ}C$)

Table 3.5: Heats of Vaporization of Linear Alkanes

Compound	temp (K)	P (mmHg)	ΔH_{vap} (kcal/mol)
$n - C_5H_{12}$	298	–	6.316
	309	760	6.16
$n - C_6H_{14}$	298	–	7.54
	342	760	6.896
$n - C_7H_{16}$	298	–	8.735
	371	760	7.575
$n - C_8H_{18}$	298	–	9.915
	399	760	8.214
$n - C_9H_{20}$	298	–	11.099
	424	760	8.82
$n - C_{10}H_{22}$	298	–	12.276
	447	760	9.39
$n - C_{11}H_{24}$	298	–	13.464
	469	760	9.92
$n - C_{12}H_{26}$	298	–	14.647
	489	760	10.44
$n - C_{13}H_{28}$	298	–	15.83
	509	760	10.9
$n - C_{14}H_{30}$	298	–	17.01
	527	760	11.4
$n - C_{15}H_{32}$	298	–	18.20
	544	760	11.8
$n - C_{16}H_{34}$	298	–	19.38
	560	760	12.3
$n - C_{17}H_{36}$	298	–	20.6
	576	760	12.7

Table 3.6: Heats of Vaporization of Linear Perfluoroalkanes

Compound	temp (K)	P (mmHg)	ΔH_{vap} (kcal/mol)
CF_4	145.1	760	3.0
C_2F_6	194.9	760	3.9
C_3F_8	236.7	760	4.69
$n - C_4F_{10}$	233.2	124	6.1
	253.2	349	5.75
	271.2	760	5.48
$n - C_5F_{12}$	302.4	760	6.25
$n - C_6F_{14}$	293.2	176	7.64
	330.3	760	6.8
$n - C_7F_{16}$	298.2	77	8.69
	355.66	760	7.54

Table 3.7: Polarizability of Linear Alkanes

Compound	Polarizability (10^{-24}cm^3)
C_2H_6	4.47
	4.43
C_3H_8	6.29
	6.37
$n - C_4H_{10}$	8.20
$n - C_5H_{12}$	9.99
$n - C_6H_{14}$	11.9
$n - C_7H_{16}$	13.7
$n - C_8H_{18}$	15.9

Table 3.8: Polarizability of Linear Perfluoroalkanes

Compound	Polarizability (10^{-24}cm^3)
C_2F_6	6.82
	6.03
$n - C_5F_{12}$	13.53
$n - C_6F_{14}$	16.22

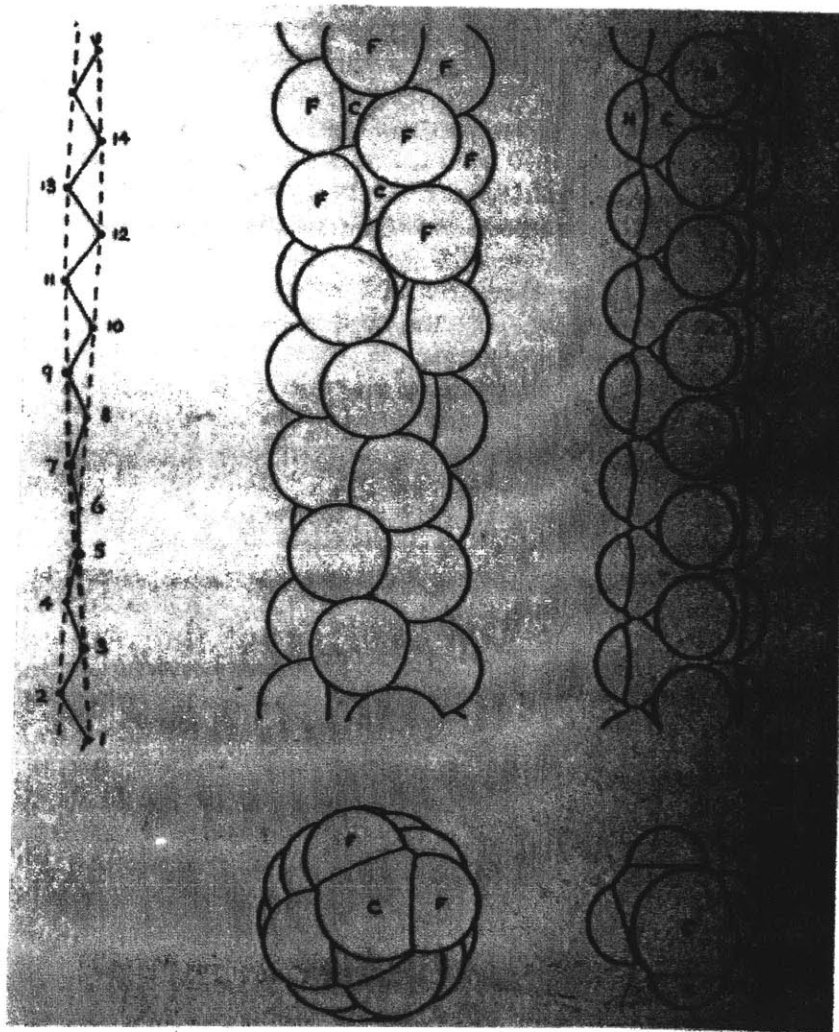


Figure 3-1: Structure of alkanes and perfluoroalkanes

ure 3-1 demonstrates, perfluorocarbon molecules form helical chains with 13-C-atom repeating units, while corresponding hydrocarbon molecules adopt a planar zig-zag carbon chain conformation. This difference is due to the fact that the fluorine atoms on the perfluorinated molecule have larger van der Waals radii than those of the hydrogen atoms on the alkanes. Since the van der Waals radii for the hydrogen atoms are only 1.1-1.2 Å, *n*-alkanes and PE are not crowded in planar zig-zag conformation. Perfluoroalkanes and PTFE, on the other hand, containing many fluorine atoms of radius 1.35 Å would be overcrowded in the same conformation. In order to increase the distance between the neighboring fluorine atoms, perfluoroalkanes and PTFE form a helical structure by rotating around the chain-bonds and by widening the C-C-C bond-angles from 109° to 116°. The backbone dihedral angles for the *trans* conformation in perfluoroalkanes and PTFE are displaced by about 17° from a true *trans*. These differences in structure will be accounted for through adjusting the forms and parameters of the torsional and bond-angle stretching forcefields.

3.1.6 Surface Measurements

Surface tensions of linear alkanes are quite well documented.^[39, 69] Surface tensions for linear perfluorinated alkanes are reported by Sauer and Dee.^[70, 71] The same authors also reported the surface tension values for PE, PTFE, and diblock co-polymers of the two. Table 3.9 and Table 3.10 shows the experimental surface tensions of alkanes and perfluoroalkanes, respectively. Surface tension is a very strong function of the temperature and decreases with increasing temperature. Surface tension value increases as the chain length increases for a homologous series. These experimental surface tension values will provide feedback on whether our forcefield and/or simulation methodology are appropriate for capturing the characteristics of the liquid-vapor interface.

3.2 Liquid-Vapor Interface Simulation

Computer simulation has been applied to a number of different types of liquid-

Table 3.9: Experimental Surface Tension Value of Hydrocarbons

Num. C	Compound	Temperature (K)	Surface Tension (dyne/cm)
5	pentane	298	15.49
6	hexane	298	17.89
7	heptane	298	19.65
8	octane	298	21.14
9	nonane	298	22.38
10	decane	298	23.37
		318	21.53
11	undecane	298	24.21
		343	20.15
12	dodecane	298	24.91
		387	17.08
13	tridecane	298	25.55
14	tetradecane	298	26.13
15	pentadecane	298	26.64
16	hexadecane	298	27.05

Table 3.10: Experimental Surface Tension Value of Perfluorocarbons

Num. C	Compound	Temperature (K)	Surface Tension (dyne/cm)
5	perfluoropentane	298	9.42
6	perfluorohexane	298	11.44
7	perfluoroheptane	298	12.78
8	perfluorooctane	298	13.7
9	perfluorononane	298	14.4
10	perfluorodecane	318	13.5
		323	13.2
		400	7.9
11	perfluoroundecane	343	12.7
12	perfluorododecane	387	10.6

vapor interfaces such the rare gases, metals, and oligomers.

3.2.1 Rare Gases – Argon/Krypton/Xenon

Lee, Barker and Pound calculated the surface tension of liquid argon using the Monte Carlo method with three different forcefields: (1) the Lennard-Jones forcefield, (2) the two-body Barker forcefield, and (3) the Barker two-body forcefield plus the three-body Axilrod-Teller contribution.^[13, 18] At all five temperatures tested (84-144 K), both the Lennard-Jones and two-body Barker forcefields significantly overestimated the surface tension by at least 20% of the experimental values. The Barker two-body forcefield in combination with the Axilrod-Teller three-body contribution (calculated using perturbation theory), on the other hand, showed significant better agreements with experimental values. In the low temperature range (84-102 K), the simulated surface tensions showed only a 1-4% deviation from the experimental values. Near the critical temperature of argon (150.7 K), larger discrepancies were displayed since the higher-order many-body terms were not included. However, the forcefield with the Axilrod-Teller contribution still performed much better than the other two near the critical temperature. This study, for the first time, indicated the importance of the three-body interactions for accurately predicting surface tension values.

Encouraged by the abovementioned findings, Barker *et al.*^[72, 73] conducted more similar Monte Carlo calculation of surface tension of liquid argon, krypton, and xenon in all temperature range. Reliable pair potential functions were used and calculations were made both with and without the Axilrod-Teller-Muto (ATM) three-body interaction. Again, the surface tensions calculated with pair potentials alone were higher than the experimental values by 19% to 35%. The surface tensions calculated with the ATM three-body interaction included, on contrary, agreed within 2.2% of the experimental values for all three rare gas liquids.

These studies demonstrated that a model with accurate pair potential plus three-body interaction is an excellent model for predicting the liquid-vapor surface properties of the rare gases.

3.2.2 Water

Besides its well-known characteristics of strong hydrogen bonding, high heat capacity and high dielectric constant, water also has a very high surface tension. The experimental values of water surface tension are 72 dyne/cm at 298 K and 68 dyne/cm at 325 K . Numerous attempts have been made to reproduce these experimental data with computer simulation, and the results vary widely.

Using MC technique, Borstnik first performed a simulation of a water surface for 64 water molecules^[74] interacting via the Matsuoka-Clementi-Yoshimine (MCY) forcefield.^[75, 76] The result indicated that the molecules at the surface preferred to orient in a way such that one hydrogen atom pointed towards the vapor phase. With a system of 256 molecules interacting with the ST2 forcefield,^[77] Lee and Scott reported an MC water surface tension of $97 \pm 6 \text{ dyne/cm}$ at 298 K .^[78] Subsequently, Jorgensen *et al.* simulated a system of 342 water molecules interacting via the TIP4P forcefield^[79] with MD technique, and predicted a surface tension of $132 \pm 46 \text{ dyne/cm}$ at 325 K .^[80] An MD simulation by Matsumoto *et al.* of a system of 1000 water molecules with the Carravetta-Clementi (CC) forcefield^[81] predicted a surface tension value of $30.5 \pm 2.7 \text{ dyne/cm}$ at 300 K .^[82] More recently, a system of 490 molecules was simulated by MC using the MCY forcefield. The surface tension calculated at 298 K is $23.7 \pm 3.4 \text{ dyne/cm}$. This simulation predicted the interfacial 10-90 thickness to be 4.70 \AA , with the dipoles of the water molecules near the liquid phase pointing slightly towards the liquid phase and those near the gas phase pointing towards the gas phase. The smaller root-mean-squared fluctuations of the dipole directions indicated that interfacial water molecules were more restricted in their rotation. The temperature derivative of the surface tension for the MCY water was found to be $-0.32 \text{ dyne}/(\text{cm}^2 \text{ K})$.^[76]

The wide range of simulated surface tension typifies the accuracy often achieved by simulation of liquid-vapor interface.

3.2.3 Metal

The Embedded Atom Method (EAM) was used to compute the surface enthalpy of liquid metals such as Cu, Au, Ag, Ni, and Pt.^[83] The results were 50% to 80% of the expected experimental values. These results followed the same trend as those of the solid-vapor metal interfaces, which were lower than experimental value by about 25%.^[84] Holtzman *et al.* suggested that it might be due to the fact these EAM calculations did not include directional, three-body terms.^[83]

3.2.4 Alkane Oligomers

Harris^[39] performed molecular dynamics simulations for the liquid-vapor interfaces of linear decane($C_{10}H_{22}$) and eicosane($C_{20}H_{42}$) with united Lennard-Jones intermolecular forcefields, fixed bond-length, a flexible bond-angle forcefield, and a torsional forcefield. Even though the total mass density profile is monotonic, as that of simple atomic fluids, these long linear molecules displayed a rich equilibrium local structures. The density profiles of the chain centers of mass and central segments were strongly peaked. In order to maximize entropy, the outer edge of the surface was dominated by chain ends. At a fixed temperature, the interface width as defined by the total mass density profile decreased with increasing chain length. The chains were flattened in the outer regions of the surface but slightly elongated below in the region corresponding to the peak in the center of mass density profile. The calculated surface tensions and liquid-vapor co-existence densities showed the correct trends with variation in chain length and temperature. The bulk liquid densities were very close to experimental values. However, the absolute values of surface tensions were significantly higher than experimental data.^[39]

3.2.5 Perfluorohydrocarbons and Semifluorinated Alkane Diblocks

Hariharan *et al.*^[40] carried out molecular dynamics simulations of the liquid-vapor interface of short fluorocarbon chains and short fluorocarbon-hydrocarbon diblock

chains, with united Lennard-Jones intermolecular forcefields, fixed bond-length, flexible bond-angle, and a realistic torsional forcefield.^[85] The perfluoromethyl segments had larger Lennard-Jones diameters and shallower well-depths than those of CH_2 and CH_3 . At high temperatures, the density profiles of the co-polymer melt decayed monotonically from the bulk liquid density to the melts. This study also showed that fluorocarbon segments in the diblocks had a tendency to segregate at the free surface when the chain length was increased or the temperature was decreased. The constraint of connectivity between the two diblocks resulted in oscillatory density profiles and a rich structure. This model predicted that a co-polymer could have a lower surface tension than either homopolymer of similar length. It again overestimated the surface tensions of the fluorocarbon and fluorocarbon-hydrocarbon diblocks, and failed to predict the lower surface tension of perfluorodecane compared that of decane.^[39, 40]

In the studies of linear alkane, perfluorinated alkane, and diblock liquid-vapor interfaces,^[39, 40] the long-range correction contributions to surface tension^[39] were quite significant, about 50% of the total values.

3.3 Review of Relevant Forcefields

3.3.1 Forcefields for Hydrocarbons

At present, there are many forcefields available for hydrocarbons. These include both explicit-atom models and more often, united-atom models. In a united-atom model, a group of atoms, for example, CH_3 , CH_2 are condensed into a “united atom”. This simplification greatly cuts down the number of “atoms” in the simulation, and allows efficient handling of larger molecules while preserving the main features of the system such as the torsional angles. The following is a list of some of the forcefields frequently used. Some of these forcefields are developed explicitly for hydrocarbons, and others cover a much wider variety of molecules.

Lifson & Warshel^[86] developed a EA forcefield with explicit hydrogens based on

equilibrium conformation, vibrational enthalpies for alkane crystals.

Kollman and co-workers^[60, 87] developed both EA and UA forcefields, which are usually known as the AMBER forcefields, based on structure, energies and vibrational frequencies of model protein and DNA crystals. These forcefields include the standard bond stretching, bond angle, torsional angle, 12-6 Lennard-Jones, 12-10 hydrogen bond and Coulombic terms. These forcefields have been used extensively for hydrocarbons and polymers.

Allinger and co-workers developed MM2/MM3 forcefields for several classes of molecules.^[88, 89] The main features of these potentials are: (1) They give more accurate representation of intramolecular energies. In addition to the usual terms, they also include the coupling terms between different modes. (2) Instead of the usual Coulombic term, they model the electrostatic energy via dipole-dipole, dipole-charge, charge-charge interaction.^[88]

Brooks and co-workers developed a united-atom forcefield called CHARMM for macromolecules. It contains an improper torsion term which is used to maintain the planarity about certain planar atoms and chirality about a tetrahedral extended heavy atoms.^[90]

Instead of fitting to experimental data from crystals, Jorgensen^[79, 91, 92] argued that intermolecular forcefields developed directly from liquid properties will be better suited for predicting liquid properties. He developed optimized intermolecular potential functions for liquid (OPLS)^[79, 91, 92, 93, 94] for a variety of molecules, including hydrocarbons, amides and peptides. The OPLS forcefields use fixed bond length, fixed bond angle, and torsional terms for intramolecular energies. They contain 12-6 Lennard-Jones, and Coulombic terms for intermolecular energies. Based on Monte Carlo simulations for 15 hydrocarbon liquids (linear and branched), twelve constituent groups were identified and their Lennard-Jones parameters optimized. The energies and densities of the liquids are within 2% of the experimentally measured values.^[91]

Toxvaerd^[95, 96] reported that molecular dynamics calculations of the pressure in fluids of propane, pentane, and decane scaled incorrectly with respect to temperature and density if the methylene and methyl groups were treated as isotropical interaction

units. An anisotropic united-atom (AUA) forcefield for alkanes was proposed and proven to overcome this shortcoming. Unlike the traditional UA forcefield, AUA forcefield model recognizes the different conformations of the H atoms on the C atom and chooses the forcefield accordingly. The anisotropy is achieved by taking the geometrical mean position of the position for the valence electrons as the origin for intermolecular interaction of the methylene and methyl units. The AUA forcefield effectively incorporates some molecular details contained in explicit-atom model, yet maintains the computational simplicity of the UA forcefield.

In addition, a totally new approach to develop forcefields has been taken by Hagler and co-workers[97]. Realizing the need to predict properties of classes of compounds for which experimental data are scarce, they derived a quantum mechanical forcefield (QMFF) by fitting to *ab initio* potential surfaces. In an explicit atom model for hydrocarbon, with same set of parameters, this forcefield reasonably well predicted properties for both unstrained and highly strained hydrocarbons.[97]

3.3.2 Forcefields for Perfluorinated Carbons

Contrast to the abundance of forcefields for hydrocarbon and PE, there are only a few forcefields available for perfluorocarbons or perfluorocarbon-related materials. Much of the forcefield development has been focused on the structure of the perfluorocarbon chains, particularly, the dihedral angle and the torsional energy barriers.

Shin *et al.*[98, 85] used a united-atom model to simulate water supported monolayers of partially fluorinated amphiphiles such as $CF_3(CF_2)_9CH_2COOH$, $CF_3(CF_2)_6CH_2(CF_2)_3COOH$ and $CF_3(CF_2)_6(CH_2)_4COOH$. The Lennard-Jones core size parameters σ for the $CF_2 - CF_2$ and $CF_3 - CF_3$ interactions were assumed to be the same and determined from the known lattice spacing of the hexagonal structure of perfluoroalkanes. The Lennard-Jones well-depth parameters ϵ were derived from the well-depth parameters for $CH_2 - CH_2$ and $CH_3 - CH_3$ using CF_x/CH_x (x=2, 3) polarizability scaling. In addition to the usual fixed bond length, flexible bond angle, and Lennard-Jones intermolecular interaction approach, to account for the helical nature of the fluorinated chains, an $Aexp(-B\phi^2)$ term was included in the torsional

potential so that instead of reaching minimum at true *trans*- ($\phi = 0$), the torsional potential has a global minimum at $\phi = \phi_0 \neq 0$. The inclusion of this term improved the prediction of collective tilted angle for these amphiphiles^[85] when compared to experimental values obtained through grazing incidence X-ray diffraction.^[99, 85] However, when the exact model was applied to the calculation of the liquid-vapor phase equilibrium by Siepmann *et al.*,^[100] the liquid density was found to be severely underestimated. In addition, this model also under-estimated the critical densities of perfluoropentane and perfluorooctane, and failed to predict critical temperatures that scale correctly with chain lengths^[100]

Hariharan *et al.*^[40] determined a new set of Lennard-Jones parameters that describe both the intramolecular interactions between $CF_x - CF_x$ ($x = 2$ or 3) that were separated by at least three other CF_x along the chain and the intermolecular interactions between any pair of CF_x belonging to different molecules. σ parameters were again kept the same for $CF_2 - CF_2$ and $CF_3 - CF_3$. σ and ϵ were optimized to predict the density and enthalpy of vaporization of liquid perfluorohexane at 1 *atm* and 293.2 *K*. The development of this forcefield will be discussed in more detail in the next chapter.

Sets of EA van der Waals parameters for Lennard-Jones (6-12) and 6-9 forcefields and equilibrium parameters for perfluoroalkanes and PTFE were derived by Holt *et al.* from MOPAC AM1 semiempirical calculation results (specifically, heats of formation, charges and geometries) for the model molecule perfluorohexadecane.^[101] Reasonable intramolecular geometry and intermolecular packing arrangements were obtained with all parameter sets developed.

Dixon^[102] calculated the potential energy surface for torsion about the central C-C bond in *n*-perfluorobutane using *ab initio* molecular orbital theory. The geometries of the symmetrical minima (t_{\pm} ($\Phi = \pm 15.4^\circ$), g_{\pm} ($\Phi = \pm 116.7^\circ$)) and the three transition states separating these minima were obtained and optimized with the $DZ + D_C$ basis set, which had been shown to give good geometry predictions for a wide range of fluorocarbons. Perfluorobutane energy calculations were also conducted with a full polarized double- ξ basis set $DZ + P$. Correlation corrections to the final energies

were done at the MP-2 level for all the valence electrons. The global minimum was found to be at t_{\pm} , with the g_{\pm} states 6.2 kJ/mol higher in energy on the electronic energy surface. The true-trans state (i.e. $\Phi = 0$) was 1.6 kJ/mol higher in energy than the t_{\pm} states. The transition barriers from t_{\pm} to g_{\pm} were 10.1 kJ/mol and those from g_{\pm} to g_{\mp} were 33.5 kJ/mol .

Later, Smith, Jaffe and Yoon^[103] established a new rotational isomeric model for PTFE based on *ab initio* electronic structure calculations on perfluorobutane, perfluoropentane, and perfluorohexane. 4-31G and D95+* basis sets were used, each at two different levels: SCF and MP2. Unlike the earlier Dixon model which had only four minima, this model contained six minima and fully took into account the strong coupling of conformations for adjacent bond pairs. The conformational energy contours confirmed the t_{\pm} states at $\Phi = \pm 17^{\circ}$, but also showed that the gauche states further splitted at $\pm 124^{\circ}$ and $\pm 84^{\circ}$ to relieve steric crowding. This model, without adjustment of geometric or energy parameters estimated from *ab initio* electronic structure calculations on small molecules, predicted the characteristic ratio of PTFE at 600 K that was consistent with recent experiments.

Chapter 4

New Lennard-Jones United-Atom Intermolecular Forcefields

4.1 Rational and Motivation

A good intermolecular forcefield is pivotal to the accurate prediction of both bulk and interfacial properties. Previously, Hariharan *et al.*^[40] obtained the Lennard-Jones parameters σ and ϵ for CF_2 and CF_3 by optimizing the parameters to predict the experimental density and enthalpy of vaporization of liquid perfluorohexane at 1 *atm* and 293.2 *K*. *NPT* MD simulation was conducted with 50 *n*-perfluorohexane molecules. During the optimization, corrections for dispersion interaction over distances beyond the cut-off distance (2.5σ) were included. The σ values for CF_2 and CF_3 were kept the same, and geometric mean combining rule were used to obtain the ϵ and σ between unlike monomers. The optimized set of Lennard-Jones parameters were listed in Table 4.1. The authors reported that the simulated and the experimental molar volumes agree to within 0.7%, while the calculated and experimental enthalpies agree to within 3%. When these parameters were applied to a liquid-vapor interface of *n*-perfluorodecane, the predicted bulk density was higher than experimental value by about 2%.

Upon careful examination of their results, we discovered a discrepancy between the experimental molar volume of *n*-perfluorohexane cited by Hariharan *et al.* and the val-

Table 4.1: Lennard-Jones Parameters, by Hariharan *et al.*

Parameter	Value
$\epsilon_{CF_2-CF_2}$	46.58 K
$\epsilon_{CF_3-CF_3}$	66.234 K
$\sigma_{CF_2-CF_2}$	4.672 Å
$\sigma_{CF_3-CF_3}$	4.672 Å

ues reported in other references.^[67] The correct molar volume of *n*-perfluorohexane is 198.9 cm^3 , not 190.0 cm^3 as reported by Hariharan *et al.* The parameters listed in Table 4.1 would actually under-predict molar volume and over-predict bulk density by about 5%. Furthermore, as shown in the sections to follow, these parameters would lead to very significant errors in predicting bulk properties of other linear perfluoroalkanes.

To overcome the aforementioned deficiency of the Hariharan forcefield, we decided to first develop a new intermolecular forcefield. Three criteria are considered important to the design of this new forcefield. The first is that the forcefield accurately predicts the thermodynamic properties of the bulk liquids over as wide a range of conditions as possible. The second criterion is that the forcefield is transferable to a wide variety of compounds which have similar functional units. The third criterion is the simplicity and computational tractability of the forcefield. Based on these three criteria, we used the united-atom treatment of the CF_2 and CF_3 units because it would require far less computational effort than forcefields with explicit-atoms. We also kept the 12-6 form of Lennard-Jones forcefield. The forcefield parameters would be determined based on the experimental densities (equivalently, molar volumes) and heats of vaporizations of *n*-perfluorobutane and *n*-perfluoroheptane at several differ-

Table 4.2: Experimental Bulk Properties

Molecule Type	$P(mmHg)$	$T(K)$	$\rho_{exp}(g/ml)$	$\Delta H_{exp}^{vap}(kJ/mol)$
<i>n</i> -perfluorobutane	124	233	1.723	25.522
<i>n</i> -perfluorobutane	349	253	1.659	24.058
<i>n</i> -perfluoroheptane	77	298	1.718	36.342
<i>n</i> -perfluoroheptane	760	355	1.551	31.56

ent temperature and pressure conditions (Table 4.2). Since the density values at the pressures listed are not available, the experimental densities at 1 *atm* are used as approximations. Because liquid density is not a strong function of pressure, the error caused by this approximation should be very small, much smaller than the uncertainties of the simulations. For example, for *n*-perfluorohexane at 288 *K*, the densities at equilibrium vapor pressure (139 *mmHg*) and air saturated at 1 *atm* are 1.7016 *g/ml* and 1.69943 *g/ml*, respectively, a difference of less than 0.2%.^[66]

There are basically two stages in constructing an intermolecular forcefield: parameterization and validation. Parameterization involves choosing a set of experimental data, surveying the space of reasonable parameter values, conducting simulations, predicting macroscopic properties and comparing them with the experimental data set. A set of parameters that gives the least error between the predicted and the experimental properties will be chosen. Once this set of parameters is chosen, it will be validated. Validation involves applying the chosen set of parameters to simulate the properties of material and/or conditions not already used in the parameterization process and compare them with the experimental values. Through the combination of parameterization and validation processes, we can improve and fine-tune the forcefield in a systematic way.

4.2 Simulation Details

Perfluorinated methylene (CF_2) and perfluorinated methyl (CF_3) groups were modeled as united-atoms. The intramolecular interactions included bond length stretching, bond angle bending, torsional rotation, and intramolecular van der Waals forces. Intermolecular interactions were van der Waals forces between united-atoms on different molecules.

The bond stretching forcefield took the form:

$$U_{bond-stretching} = \sum_{bonds} K_r (r - r_{eq})^2 \quad , \quad (4.1)$$

where $K_r = 217567 \text{ KJ/nm}^2/\text{mol}$ and $r_{eq} = 0.153 \text{ nm}$.

The bond angle bending forcefield was:

$$U_{bond\ angle\ bending} = \sum_{bond\ angles} K_\theta (\theta - \theta_{eq})^2 \quad (4.2)$$

where the parameters $k_\theta = 527.184 \text{ kJ/mol/rad}^2$ and $\theta_{eq} = 112.4^\circ$, determined by Kollman and co-workers.^[60]

The torsional angle forcefield which described the interactions between any four successive united-atom along the chain, took the form:

$$U_{torsion} = \sum_{i=0}^5 C_i \cos^i \Phi + [a \exp(-b\Phi^2)] \quad (4.3)$$

where Φ , dihedral angle, was defined according to the convention that the minimum energy *trans* configuration was at $\Phi = 0$ when $a = b = 0$. The exponential terms were included to account for the helical superstructure of perfluoroalkanes as observed by crystallographic studies^[68]. The parameters ($C_0 - C_5$, a , and b) were obtained by fitting to the results of the quantum mechanical calculations by Dixon^[102]. The Dixon torsional forcefield had lower energy barriers from those of torsional forcefield developed by Shin *et al.*^[98] as previously used by Hariharan.^[40] The Dixon torsional forcefield was used for all the simulations conducted in this chapter. The parameters

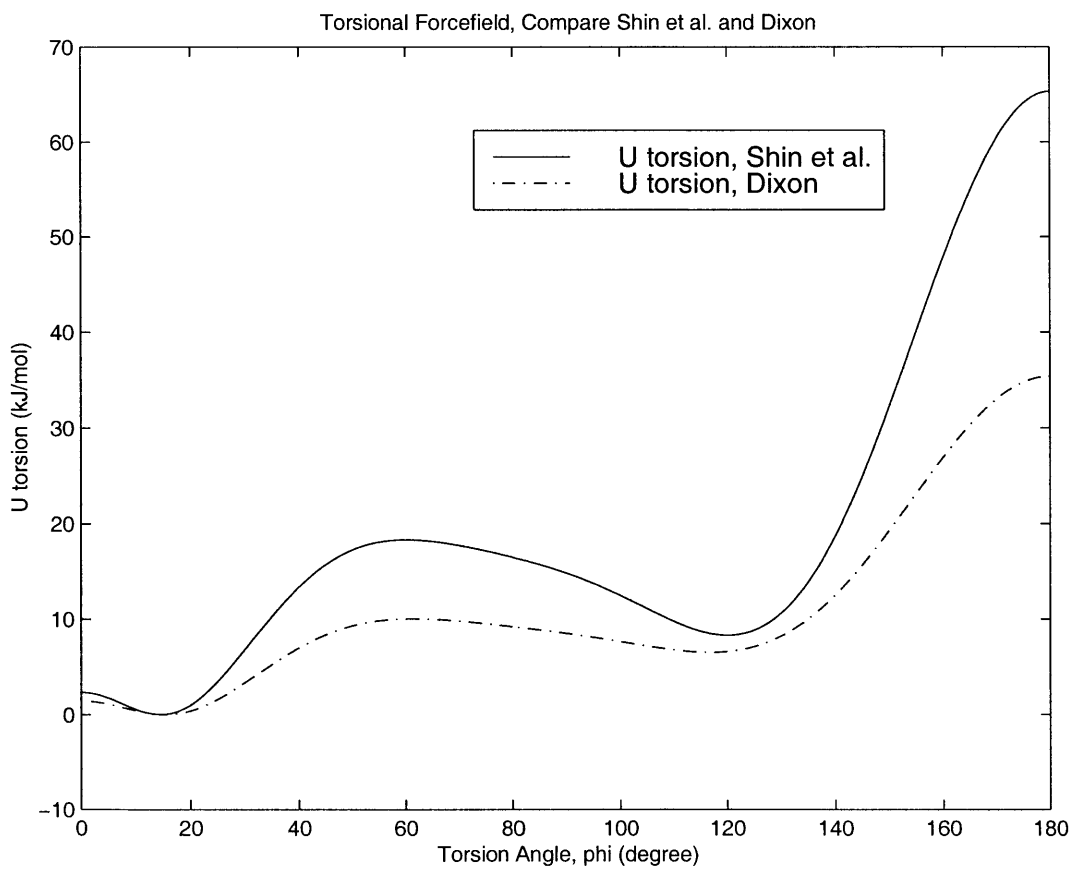


Figure 4-1: Torsional Forcefields, by Dixon and by Shin *et al.*

Table 4.3: Parameters of Torsional Forcefields

Parameter	by Dixon	by Shin <i>et al.</i>
C_0	8.51331	14.8270
C_1	4.40983	11.2778
C_2	-2.85735	-12.0000
C_3	3.29895	9.7778
C_4	8.47115	24.0000
C_5	-28.9601	-59.5556
a	8.54774	14.0000
b	11.4505	12.7176

of the two torsional forcefields were listed in Table 4.3 and the forcefields were plotted in Figure 4-1.

Van der Waals interaction between united-atoms that were separated by at least three carbon atoms along the chain within the same molecule, and that between united-atoms in different molecules were described by Lennard-Jones forcefields:

$$U_{L-J} = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right] \quad (4.4)$$

where σ was the Lennard-Jones core diameter, ϵ was the forcefield well-depth, and r was the interatomic distance. Interaction parameters for united-atoms of different types were taken as the geometric means of the corresponding interaction parameters for the two united-atoms:

$$\sigma_{CF_2-CF_3} = (\sigma_{CF_2-CF_2}\sigma_{CF_3-CF_3})^{1/2} \quad (4.5)$$

$$\epsilon_{CF_2-CF_3} = (\epsilon_{CF_2-CF_2}\epsilon_{CF_3-CF_3})^{1/2} \quad (4.6)$$

By defining

$$C = 4\epsilon\sigma^{12} \quad (4.7)$$

and

$$D = 4\epsilon\sigma^6 \quad (4.8)$$

the Lennard-Jones forcefield can be equivalently written as

$$U_{L-J} = \frac{C}{r^{12}} - \frac{D}{r^6} \quad (4.9)$$

and the parameters to be determined are $C_{CF_2-CF_2}$, $C_{CF_3-CF_3}$, $D_{CF_2-CF_2}$, and $D_{CF_3-CF_3}$.

Simulations were conducted in *NPT* ensembles. Cubic boxes with periodic boundary conditions in all three directions were used. For each simulation, 100 linear perfluoroalkane molecules of interest were included in the simulation box. The initial configurations were generated using the Discover Biosym software. To eliminate initial overlap of molecules, σ values of the united-atoms were initially set to close to zero and gradually increased to the desired values.

For each set of parameters tested, there were two stages in the simulation: equilibration stage followed by data collection stage. To make certain the system had been well equilibrated, two checking mechanisms were used. First, various thermodynamic quantities such as energy, temperature, and volume were plotted against time. When equilibrium had been achieved, these quantities fluctuated around their averages, which remained constant over time. Second, calculations were started with different initial configurations and different initial velocities. Convergence to similar configurations and properties from different initial values again indicated equilibrium had been reached.

During the equilibration phase, to maintain constant temperature, Berendsen's thermostat, with a time constant of 2.5-ps was used in addition to massive stochastic collisions in which velocities were randomly reassigned from the Maxwell Boltzmann distribution every 7 ps. During the sampling phase, only massive stochastic collisions

of 7 ps frequency were employed to maintain a constant temperature. For *NPT*, Andersen’s extended system method was used to maintain a constant pressure. A piston weight of 25 ($kJps^2$) was used for all subsequent *NPT* simulations. Each system was equilibrated for a 70-140 ps period and samples were taken over the subsequent 70-140 ps.

The equation of motion were integrated using RATTLE with a timestep of 0.007 ps. The 0.007-ps timestep was chosen because at this timestep, no significant temperature of energy drift occurred over a 70-ps microcanonical (*NVE*) simulation of *n*-perfluorohexane at $T = 293\text{ K}$ and the root-mean-square fluctuations in the total energy were less than 10% of the fluctuations in the kinetic energy.

As discussed earlier in this chapter, the main objective of this chapter was to obtain a new set of Lennard-Jones parameters ($\epsilon_{CF_2-CF_2}$, $\epsilon_{CF_3-CF_3}$, $\sigma_{CF_2-CF_2}$ and $\sigma_{CF_3-CF_3}$, or equivalently, $C_{CF_2-CF_2}$, $C_{CF_3-CF_3}$, $D_{CF_2-CF_2}$ and $D_{CF_3-CF_3}$) that were capable of covering a wider variety of *n*-perfluoroalkanes under various temperature and pressure conditions. To achieve this objective, different trial Lennard-Jones parameter sets were tested in *NPT* ensembles to simulate macroscopic bulk liquid properties for linear perfluoroalkanes. The simulated heats of vaporization and molar-volumes were compared to the experimental values. The consistency or lack thereof between the experimental data and simulation data indicated whether the trial Lennard-Jones parameters were proper and pointed out directions for adjusting the parameters to better match the two data sets.

The block averages of macroscopic properties were taken by grouping data in blocks large enough so the error estimates do not increase significantly with the blocking factor, and 95% confidence intervals were calculated.

The heats of vaporization were calculated using

$$\Delta H_{vap} = U_{intra}(g) - (U_{inter}(l) + U_{intra}(l)) + RT \quad (4.10)$$

where $U_{intra}(g)$, $U_{inter}(l)$, and $U_{intra}(l)$ were molar intramolecular internal energy in gas, molar intermolecular internal energy in liquid, and molar intramolecular internal

energy in liquid, respectively. R was the gas constant. The chief approximation in Equation 4.10 was that the sum of the kinetic and vibrational energies were the same for the gas and liquid. The results of Monte Carlo simulation for a wide variety of hydrocarbon molecules performed by Jorgensen *et al.* [79] suggested

$$U_{intra}(g) \approx U_{intra}(l) \quad (4.11)$$

and we further approximated Equation 4.10 with

$$\Delta H_{vap} = -U_{inter}(l) + RT \quad (4.12)$$

4.3 Results and Discussion

4.3.1 Parameterization

To standardize the comparison between the errors associated with different properties, a positive function E is defined as

$$E = \sum_M (e^2) \quad (4.13)$$

where

$$e = \frac{\text{predicted value} - \text{experiment value}}{\text{experiment value}} \quad (4.14)$$

and

$$M = \text{number of properties predicted} \quad (4.15)$$

Since our parameterization process involves 4 heat of vaporization values and 4 molar volume values, M is equal to 8. The best set of parameters is the one that yields a value of error E closest to 0.

After exploring a wide range of reasonable Lennard-Jones parameters, a set of new Lennard-Jones parameters were obtained and listed in Table 4.4. Figure 4-2 shows the Lennard-Jones forcefield used in the study by Hariharan *et al.* Figure 4-3 shows the

Table 4.4: New Lennard-Jones Parameters

Parameter	Hariharan	New
$\epsilon_{CF_2-CF_2}$	46.58 <i>K</i>	27.7 <i>K</i>
$\epsilon_{CF_3-CF_3}$	66.234 <i>K</i>	135.3 <i>K</i>
$\sigma_{CF_2-CF_2}$	4.672 <i>A</i>	5.4 <i>A</i>
$\sigma_{CF_3-CF_3}$	4.672 <i>A</i>	4.3 <i>A</i>
$C_{CF_2-CF_2}$	0.0202 $nm^{12}K$	0.0681 $nm^{12}K$
$C_{CF_3-CF_3}$	0.0287 $nm^{12}K$	0.0216 $nm^{12}K$
$D_{CF_2-CF_2}$	1.938 nm^6K	2.747 nm^6K
$D_{CF_3-CF_3}$	2.755 nm^6K	3.421 nm^6K

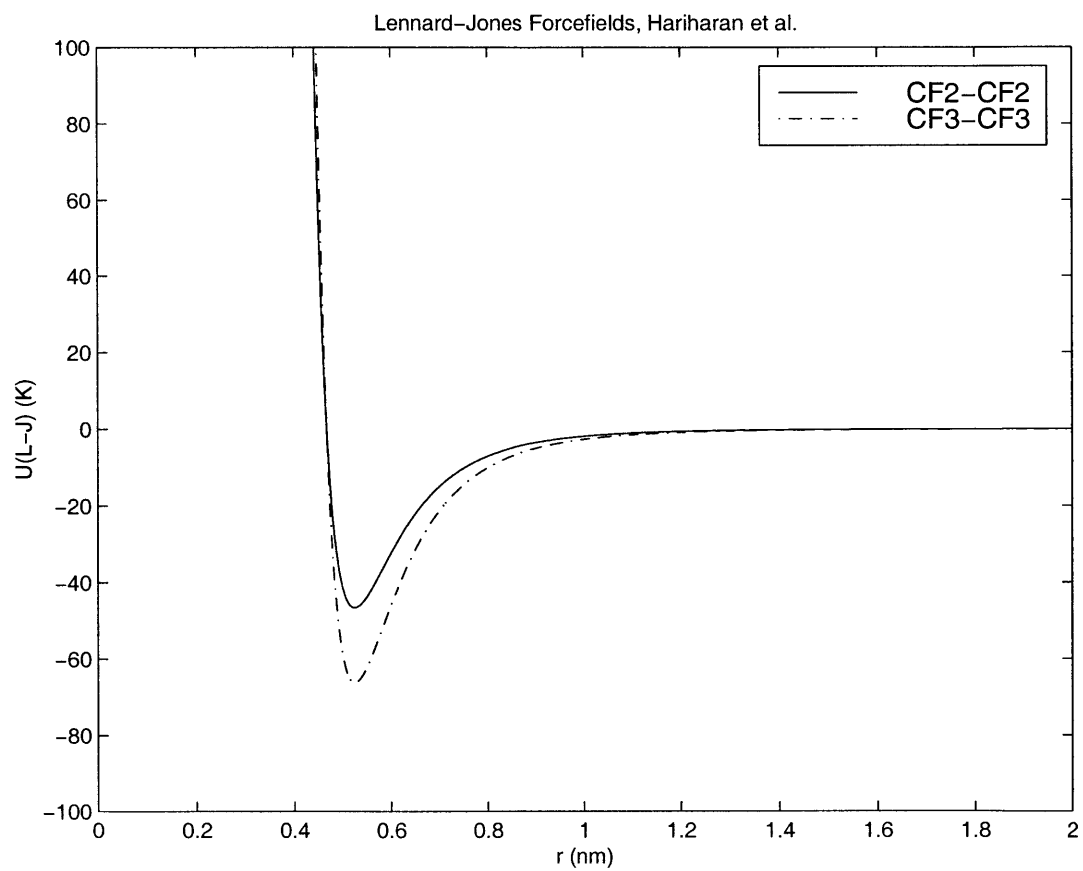


Figure 4-2: Lennard-Jones Forcefield, Hariharan *et al.*

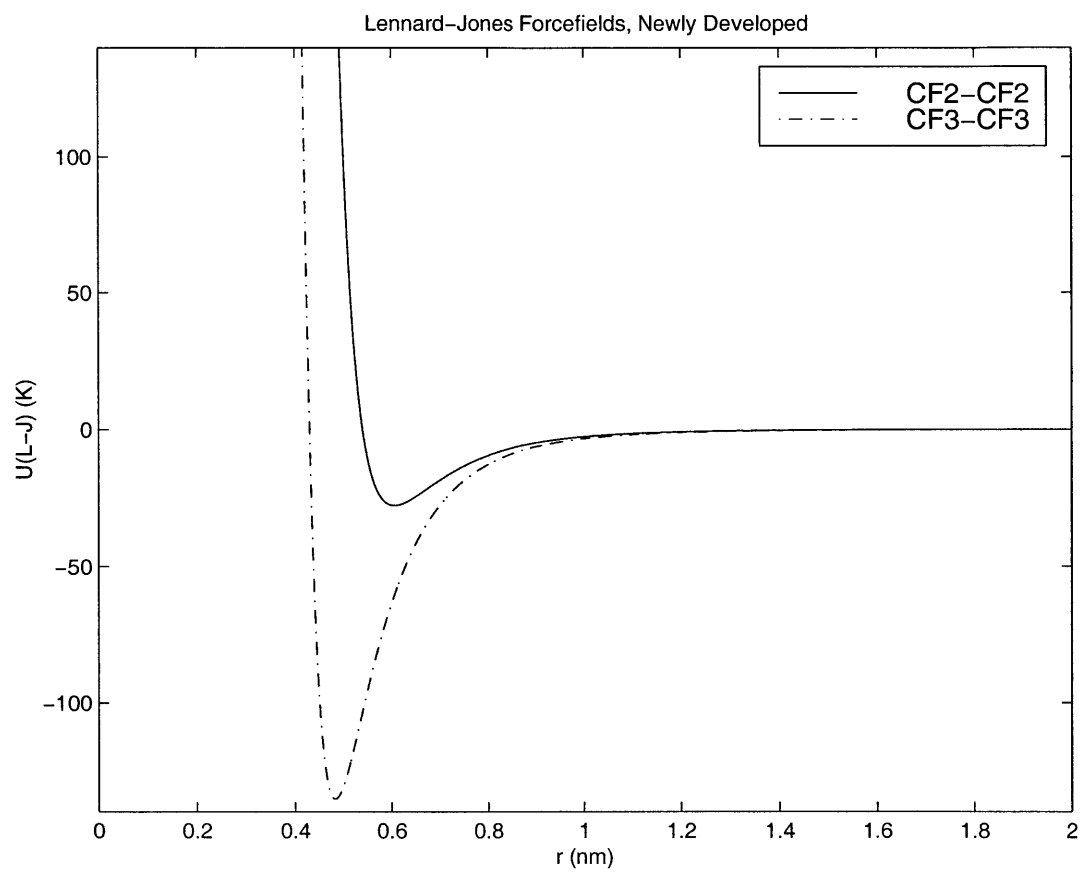


Figure 4-3: Newly Developed Lennard-Jones Forcefield

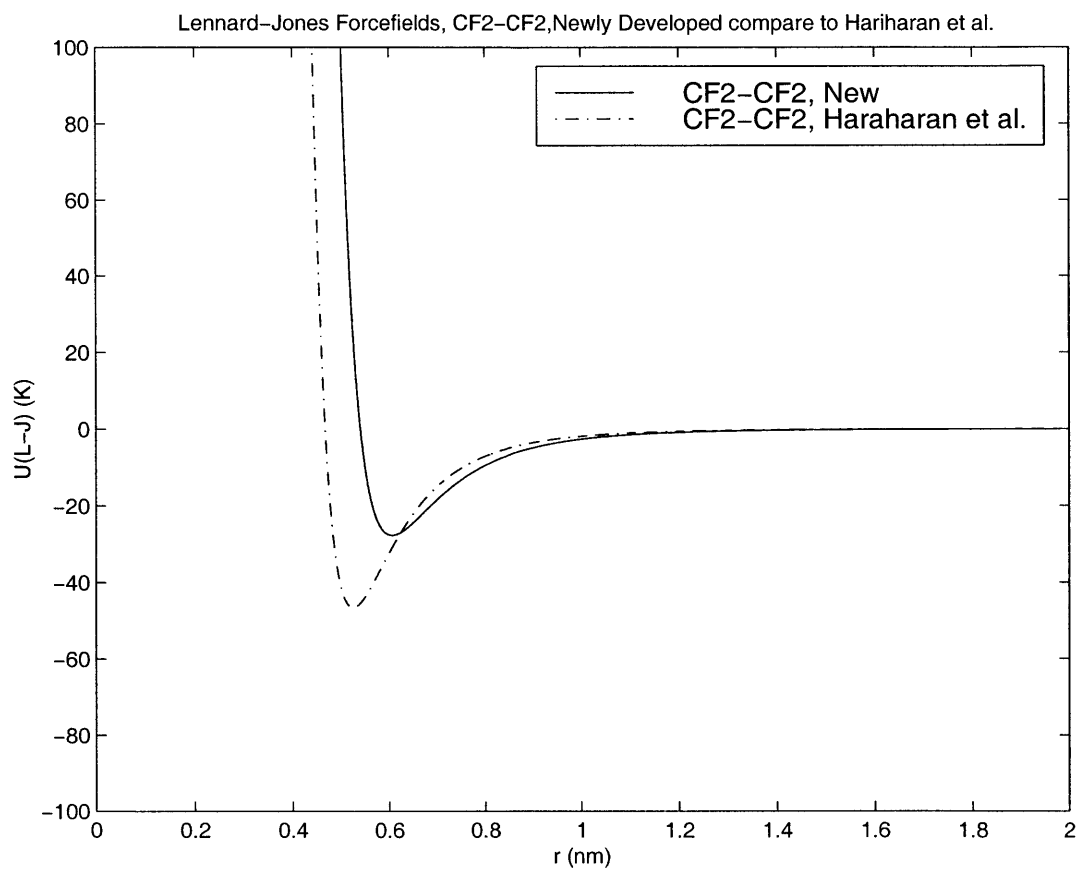


Figure 4-4: Lennard-Jones Forcefield for $CF_2 - CF_2$

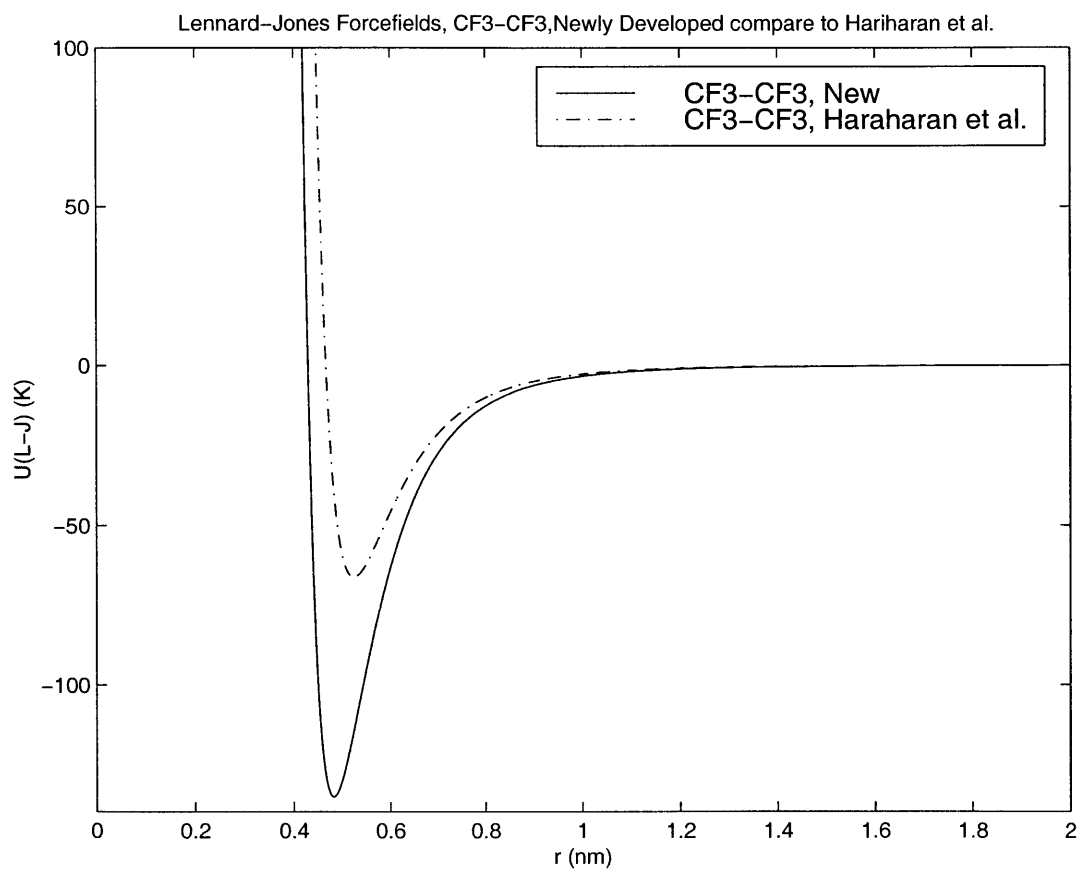


Figure 4-5: Lennard-Jones Forcefield for $CF_3 - CF_3$

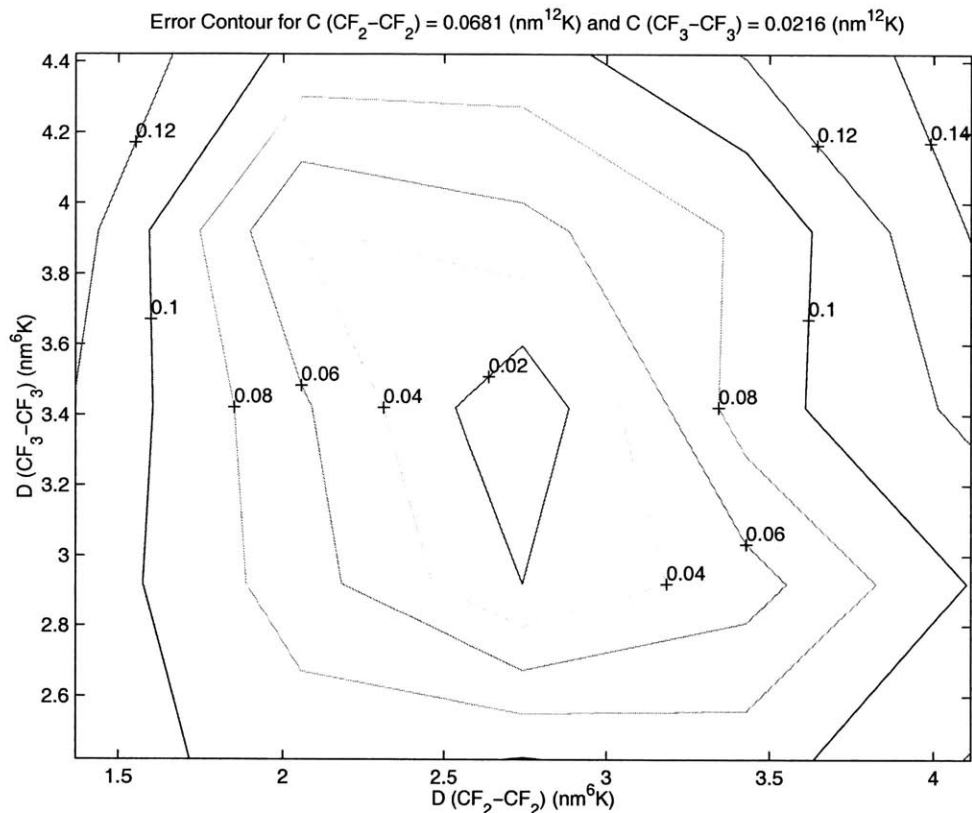


Figure 4-6: Error (E) Contour at $C (CF_2-CF_2) = 0.0681 (nm^{12}K)$ and $C (CF_3-CF_3) = 0.0216 (nm^{12}K)$

newly-developed Lennard-Jones forcefield. Figure 4-4 shows the comparison between the new forcefield and Hariharan *et al.* forcefield for CF_2-CF_2 interactions. Figure 4-5 shows the comparison between the new forcefield and Hariharan *et al.* forcefield for CF_3-CF_3 interactions.

Figure 4-6 – Figure 4-11 are the contour diagrams of error E as a function of two of the four Lennard-Jones parameters while the other two are kept constant.

It is interesting to note that the σ value of CF_2 is larger than that of CF_3 . In previous united-atom models,^[98, 85, 40] the two were the same. However, we found that if the same values of σ are kept for both CF_2 and CF_3 , it is not possible to achieve a reasonably good fit to the densities and enthalpies of vaporization for both *n*-perfluorobutane and *n*-perfluoroheptane. A qualitative understanding of the larger effective sizes of the CF_2 can be achieved by considering the directions of interactions involving CF_3 and CF_2 groups. For another united atom to approach CF_3 or CF_2 ,

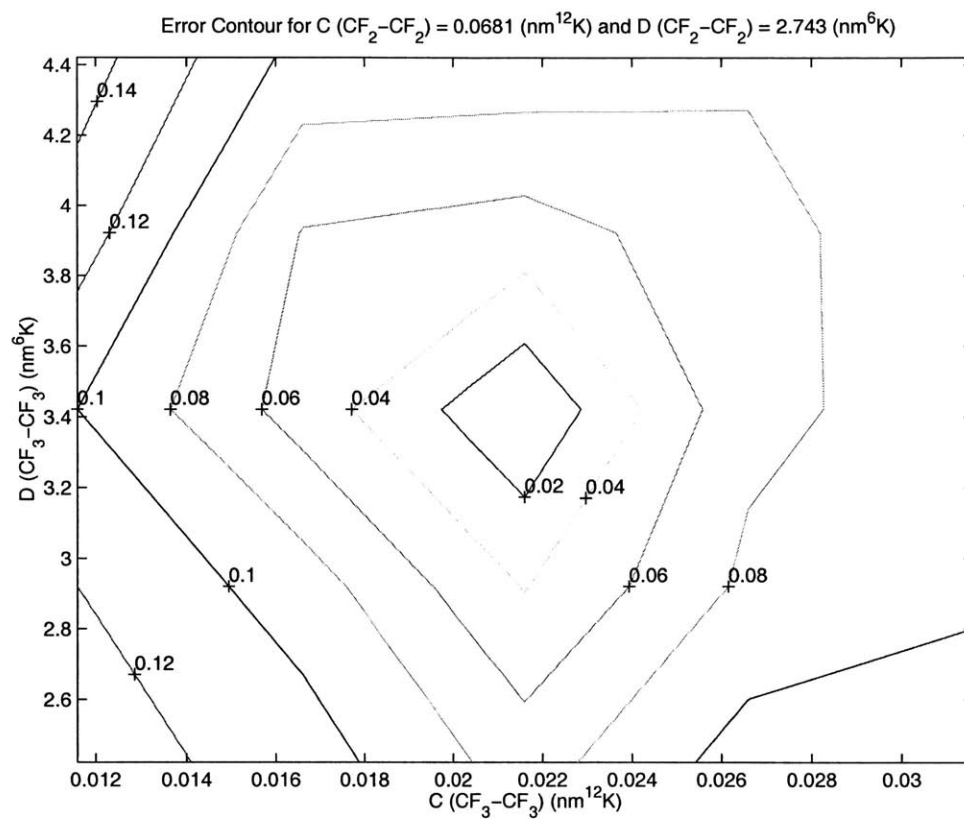


Figure 4-7: Error (E) Contour at C (CF_2-CF_2) = 0.0681 ($nm^{12}K$) and D (CF_2-CF_2) = 2.743 (nm^6K)

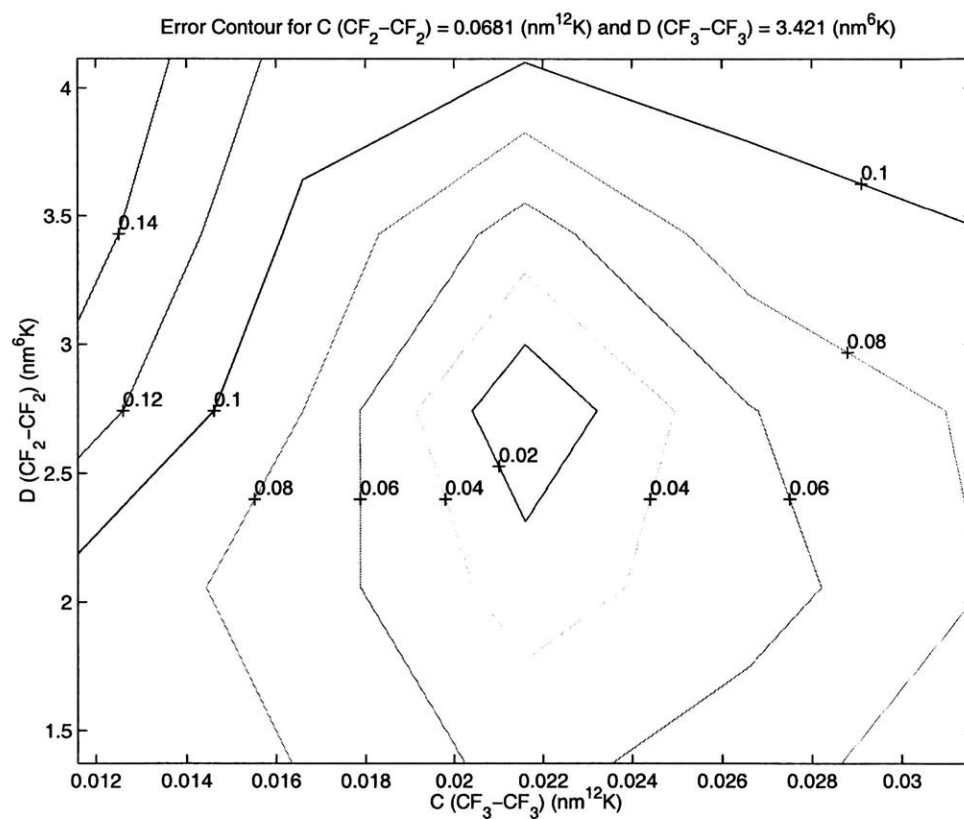


Figure 4-8: Error (E) Contour at C (CF_2-CF_2) = 0.0681 ($nm^{12}K$) and D (CF_3-CF_3) = 3.421 (nm^6K)

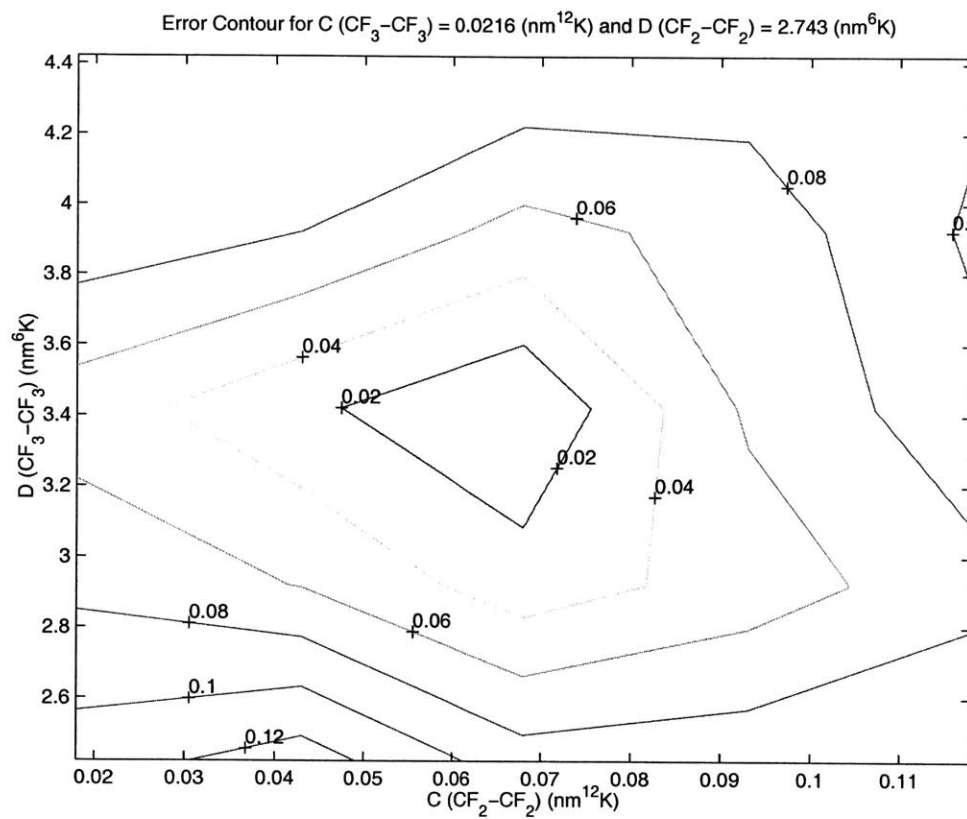


Figure 4-9: Error (E) Contour at $C (CF_3-CF_3) = 0.0216 (nm^{12}K)$ and $D (CF_2-CF_2) = 2.743 (nm^6K)$

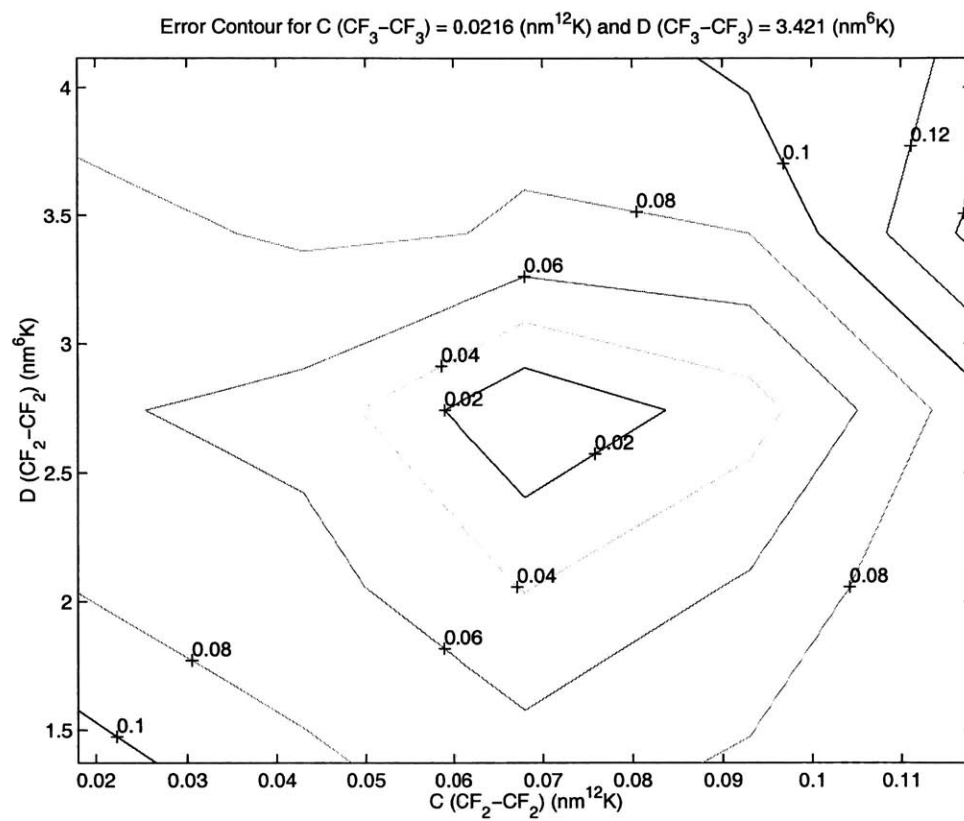


Figure 4-10: Error (E) Contour at $C (CF_3 - CF_3) = 0.0216 (nm^{12}K)$ and $D (CF_3 - CF_3) = 3.421 (nm^6K)$

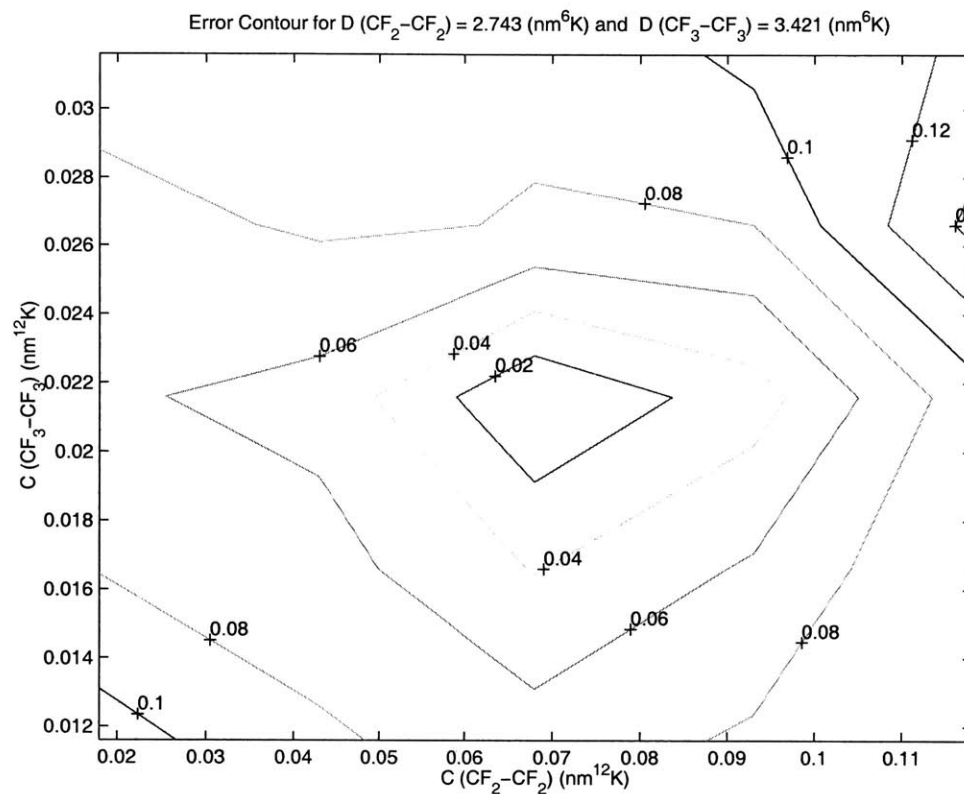


Figure 4-11: Error (E) Contour at $D(CF_2-CF_2) = 2.743 (nm^6K)$ and $D(CF_3-CF_3) = 3.421 (nm^6K)$

Table 4.5: Simulated Bulk Properties, with Different Lennard-Jones Forcefields

Molecule/Condition	Property	Exp.	New	Hariharan
<i>n</i> -perfluorobutane	ΔH_{vap} (kJ/mol)	25.522	25.2±0.4	19.3±0.4
233 K, 124 mmHg	Mol. Vol. (cm ³ /mol)	138.149	137.9±0.5	142.6±0.9
<i>n</i> -perfluorobutane	ΔH_{vap} (kJ/mol)	24.058	24.5±0.4	18.7±0.4
253 K, 349 mmHg	Mol. Vol. (cm ³ /mol)	143.478	141.9±0.7	148.3±0.7
<i>n</i> -perfluoroheptane	ΔH_{vap} (kJ/mol)	36.342	35.6±0.5	34.6±0.5
298 K, 77 mmHg	Mol. Vol. (cm ³ /mol)	225.885	229.5±1.1	208.2±1.3
<i>n</i> -perfluoroheptane	ΔH_{vap} (kJ/mol)	31.56	32.1±0.4	32.3±0.6
355 K, 760 mmHg	Mol. Vol. (cm ³ /mol)	250.206	251.2±1.5	221.0±1.9

the contact distance is primarily determined by the fluorine atoms on the CF_3 group. The fluorine atoms on the CF_3 end are less sterically hindered than the fluorine atoms on the CF_2 group which are in the middle of the chain. The F atom on CF_3 can rotate away from the approaching united atom, allowing it to pack more closely to the CF_3 group. In addition, the united atoms bonded to CF_2 group serve to direct the approach along the C-F bond directions, causing the van der Waals force of the F to be felt sooner than for close approach to a CF_3 group. Therefore, the effective size of the CF_2 group is larger than that of CF_3 . Similar results were observed for linear alkanes.^[104, 105]

As Table 4.5 shows, the new Lennard-Jones parameters significantly improved the prediction of bulk properties in all cases. The 95% confidence interval of the simulation results are also included. With the parameters used by Hariharan *et al.*, the relative errors in ΔH_{vap} can be as high as 20% of the experimental values and

Table 4.6: Forcefield Validation: Experimental and Simulated Bulk Properties

Molecule/Condition	Property	Experiment	Predicted
<i>n</i> -perfluoropentane 302.4 K, 760 mmHg	ΔH_{vap} (kJ/mol)	26.2	26.5±0.4
	Mol. Vol. (cm ³ /mol)	181.4	179.5±0.6
<i>n</i> -perfluorohexane 330.3 K, 760 mmHg	ΔH_{vap} (kJ/mol)	28.5	28.0±0.5
	Mol. Vol. (cm ³ /mol)	215.0	217.9±0.9
<i>n</i> -perfluorodecane 333 K, 760 mmHg	Mol. Vol. (cm ³ /mol)	310.1	313.6±1.7

the relative errors in molar volume can be as high as 10% of the experimental values. With the new parameters, the relative errors in ΔH_{vap} and molar volume are within 2% of the experimental values for all cases.

4.3.2 Validation

Next, we apply the forcefield obtained above to simulate bulk properties of linear perfluoroalkanes not previously parameterized against. Ideally, we would like to predict molar volumes and heats of vaporization for chains longer than 7 carbons. However, we are unable to obtain any experimental heat of vaporization value for chains longer than perfluoroheptane. So we validated our new forcefield using the molar volume and heat of vaporization values of perfluoropentane and perfluorohexane, and the molar volume of perfluorodecane. As Table 4.6 shows, the new Lennard-Jones forcefield predicted molar volume and heat of vaporization values that are in excellent agreement with the experimental values, and was proven valid.

4.4 Conclusion

A new set of Lennard-Jones parameters for perfluoroalkane is obtained and validated through predicting bulk liquid properties such as molar volumes and heats of vaporization for a variety of linear perfluoroalkanes and at a variety of temperatures and pressures. Compare to previous sets of Lennard-Jones parameters, this new set has larger σ for perfluoromethylene group than perfluoromethyl group. The energy well for perfluoromethyl group is much deeper than that of Hariharan *et al.* The energy well for perfluoromethylene is much shallower than that of Hariharan *et al.* Compare to the Lennard-Jones forcefields previously used by Hariharan *et al.*, this new set significantly improved the prediction of bulk liquid properties for linear perfluoroalkanes.

Chapter 5

Liquid-Vapor Interface Simulation

5.1 Objectives

The main objective of the work presented in this chapter is to test the performance of the Lennard-Jones united-atom forcefield developed in the preceding chapter in simulating linear perfluoroalkane liquid-vapor interfaces. Other issues such as the effect of longer cut-off radius, effect of wider bond angle, and effect of different torsion forcefields is addressed in this chapter as well.

5.2 Simulation Details

The simulation system was an NVT ensemble in a rectangular box of dimension L_x , L_y , and L_z , as illustrated in Figure 5-1, with periodic boundary conditions in all directions. $L_x = L_y = 6.093nm$ for all simulation sets, except Sets VI and VII, where $L_x = L_y = 5.0nm$. L_z was set to about 4-times the value of L_x and L_y to ensure the equilibrium configuration had two independent free surfaces on either side of the bulk liquid phase, and particles in the liquid phase did not interact with each other across the vapor phase.

For each liquid-vapor interfacial simulation, 400 n -perfluorodecane or n -decane molecules were used. The initial configuration were generated using Discover software. To eliminate overlap of molecules, the σ values of the united atoms were set

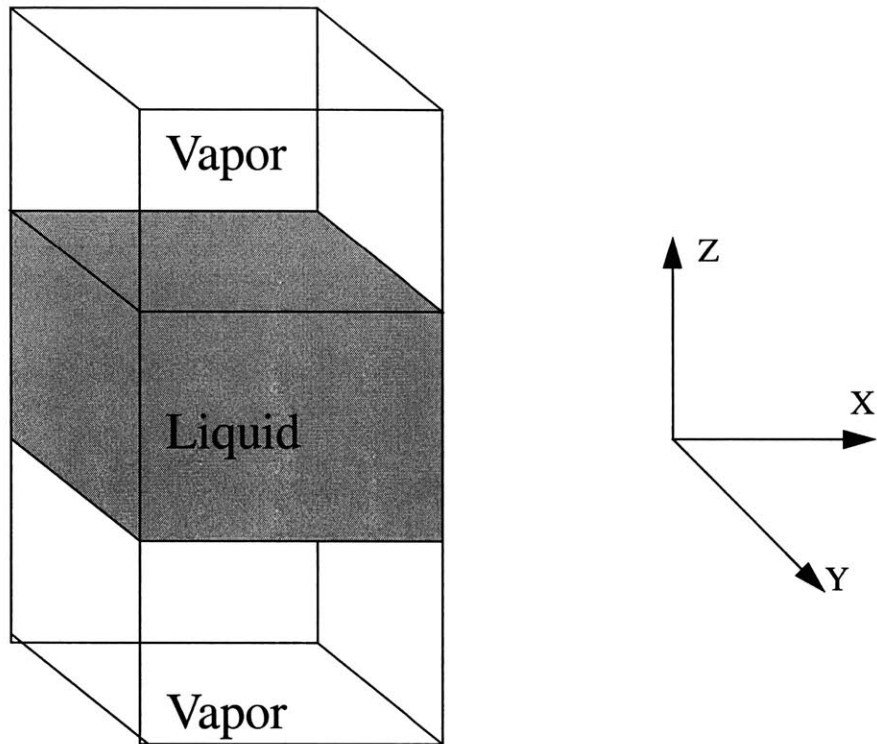


Figure 5-1: Sketch of Liquid-Vapor Interface Simulation Geometry

to close to zero and gradually increased to the desired values. The simulations were started by placing the chains in a rectangular box of the volume corresponded to the experimental bulk density at 1 *atm* and the temperature value of interest. Then two interfaces were formed by moving the boundaries in the z-direction symmetrically to L_z . So the simulation system represented a liquid-vapor co-existence at pressure slightly lower than 1 *atm*.

The equations of motion were integrated using RATTLE with a timestep of 0.007 *ps*. At this timestep, no significant temperature or energy drift occurred over a 280-*ps* microcanonical (*NVE*) simulation of *n*-perfluorodecane at $T = 400$ *K*. The root-mean-square fluctuations in the total energy were 9% of the fluctuations in the kinetic energy during the simulation.

Equilibration was ensured with two checking mechanisms. First, various thermodynamic quantities such as energy, temperature, and pressure were plotted versus time. When equilibrium had been reached, these quantities fluctuated around their averages, which remain constant over time. Second, simulations were started with different initial conformations and initial velocities. Convergence to similar values of properties from different initial conditions again verified that equilibrium has been reached.

During the process of equilibration, constant temperature was maintained by employing both Berendsen's thermostat (with a time constant of 2.5 *ps*) and massive stochastic collisions in which velocities were randomly reassigned from the Maxwell Boltzmann distribution every 7 *ps* (1000 timesteps). After equilibrium had been reached, during the subsequent sampling phase, constant temperature was maintained by using only massive stochastic collision in which velocities were randomly reassigned from the Maxwell Boltzmann distribution every 7 *ps* (1000 timesteps).

The sampling phases were at least 700 *ps* (1×10^5 timesteps), typically more than 1400 *ps* (2×10^5 timesteps). Atomic and molecular virials, atomic and molecular temperatures, and energetic quantities were saved every 0.14 *ps* (20 timesteps). Configurations were saved every 0.70 *ps* (100 timesteps) for analysis of density and structural properties.

Perfluorinated methylene (CF_2) and perfluorinated methyl (CF_3) groups were modeled as united-atoms. The intramolecular interactions included bond length stretching, bond angle bending, torsional rotation, and intramolecular van der Waals forces. Intermolecular interactions were van der Waals forces between united-atoms on different molecules.

The bond stretching forcefield took the form:

$$U_{bond-stretching} = \sum_{bonds} K_r (r - r_{eq})^2 \quad , \quad (5.1)$$

where $K_r = 217567 \text{ KJ/nm}^2/\text{mol}$ and $r_{eq} = 0.153 \text{ nm}$.

The bond angle bending forcefield is:

$$U_{bond\ angle\ bending} = \sum_{bond\ angles} K_\theta (\theta - \theta_{eq})^2 \quad (5.2)$$

where the parameters $k_\theta = 527.184 \text{ kJ/mol/rad}^2$ and $\theta_{eq} = 112.4^0$, determined by Kollman and co-workers.^[60] In addition, since X-ray crystallography had suggested a bond angle value of 116^0 , a simulation was conducted using $\theta_{eq} = 116^0$.

The torsional angle forcefield which described the interactions between any four successive united-atom along the chain, took the form:

$$U_{torsion} = \sum_{i=0}^5 C_i \cos^i \Phi + [a \exp(-b\Phi^2)] \quad (5.3)$$

where Φ , dihedral angle, was defined according to the convention that the minimum energy *trans* configuration was at $\Phi = 0$ when $a = b = 0$.

To evaluate the effect of different torsional potential barrier, two different of sets of parameters were used, one corresponded to the Dixon torsional potential, the other corresponded to the Shin *et al.* torsional potential. These parameters are listed in the preceding chapter. Dixon torsion potential has lower energy barriers than those of Shin *et al.*

Van der Waals interaction between united-atoms that were separated by at least three carbon atoms along the chain within the same molecule, and that between

united-atoms in different molecules were described by Lennard-Jones forcefields:

$$U_{L-J} = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right] \quad (5.4)$$

where σ was the Lennard-Jones core diameter, ϵ was the forcefield well-depth, and r was the interatomic distance. Interaction parameters for united-atoms of different types were taken as the geometric means of the corresponding interaction parameters for the two united-atoms:

$$\sigma_{CF_2-CF_3} = (\sigma_{CF_2-CF_2}\sigma_{CF_3-CF_3})^{1/2} \quad (5.5)$$

$$\epsilon_{CF_2-CF_3} = (\epsilon_{CF_2-CF_2}\epsilon_{CF_3-CF_3})^{1/2} \quad (5.6)$$

The Lennard-Jones parameters developed in the preceding chapter was tested in the liquid-vapor interface system. As a comparison, simulation was also run with the parameters developed by Hariharan *et al.*

The effect of longer cut-off distance is explored by conducting liquid-vapor simulation at two difference cut-off distances: $2.5\sigma_{CF_2-CF_2}$ and $5.0\sigma_{CF_2-CF_2}$.

Simulations for n-perfluorodecane were conducted at two different temperatures: 400 K and 323 K. To ensure the correct implementation of computer simulation, a MD run of the *n*-decane liquid-vapor interface was conducted under the exact same condition as reported by Harris.^[39]

Table 5.1 summaries the different sets of liquid-vapor MD simulations conducted. Comparing sets I and II can demonstrate the effect of different Lennard-Jones force-field. Comparing sets II and III can demonstrate the effect of different torsional barriers. Comparing sets I and IV can demonstrate the effect of bond angles. Comparing sets I and V can demonstrate the effect of cut-off distances. Comparing sets I and VI can demonstrate the effect of temperature.

Table 5.1: Liquid-Vapor Simulation Conditions Summary

set	Molecule	Temp (K)	L-J	Torsion	θ_{eq} (<i>degree</i>)	$R_{cut-off}$ (<i>nm</i>)
I	$C_{10}F_{22}$	400	New	Dixon	112.4	1.35
II	$C_{10}F_{22}$	400	Hariharan	Dixon	112.4	1.35
III	$C_{10}F_{22}$	400	Hariharan	Shin	112.4	1.35
IV	$C_{10}F_{22}$	400	New	Dixon	116.0	1.35
V	$C_{10}F_{22}$	400	New	Dixon	112.4	2.7
VI	$C_{10}F_{22}$	323	New	Dixon	112.4	1.35
VII	$C_{10}H_{22}$	400	OPLS	OPLS	112.4	0.976

5.3 Property Calculation

5.3.1 Density Profile and Width of Interface

The density profiles were generated by averaging the configurations accumulated during the sampling phase. The width of each slice of the density histogram was $0.1nm$. Total density density profiles, end group density profiles, and middle group density profiles were generated for each set of simulations. As measures of thickness of the liquid-vapor interfaces, 10-90 widths (W_{10-90}), defined as the distance over which the total density decreased from 90% of bulk liquid density value to 10% of the bulk liquid density value, were also recorded.

5.3.2 Surface Tension

Mansfield and Theodorou^[106, 107] identified the internal energy component of

the surface tension, γ^U , as

$$\gamma^U = \left. \frac{\partial U}{\partial A} \right|_{T,V,N} = \frac{\Delta U}{\Delta A} \Big|_{T,V,N} \quad , \quad (5.7)$$

where U is the internal energy and A is the surface area. ΔU and ΔA represent a finite small change in the internal energy and surface area, respectively.

Harris^[39] extended the method of calculating surface tensions from molecular dynamics trajectories to molecular fluids. We use this method to computer surface tension from the molecular virial:

$$\gamma = \frac{1}{2A} \langle V_{zz} - V_{xx} - V_{yy} \rangle \quad , \quad (5.8)$$

where A is the total surface area, $2L_x L_y$. $V_{\alpha\alpha}$, the molecular virial tensor, is defined as

$$V_{\alpha\alpha} = \sum_{a < b} F_{ab}^{\alpha} R_{ab}^{\alpha} \quad . \quad (5.9)$$

The sum is taken over all molecules a and b. R_{ab}^{α} is the α component in direction $\alpha = (x, y, or z)$ of the vector between the centers of mass of molecules a and b. F_{ab}^{α} is the total force between the molecules.

Harris calculated surface tensions of chain molecules such as decane ($C_{10}H_{22}$) and eicosane ($C_{20}H_{42}$), and included the long-range correction factor for interactions beyond the cutoff-distance by using a mean-field approximation:^[39]

$$\gamma_{Long-Range} = \frac{\pi}{4} \sum_{\alpha, \beta=1}^m \int_{-\frac{L}{2}}^{\frac{L}{2}} dz_{\alpha} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz_{\beta} \int_{p_{min}}^{\infty} dp p \rho(z_{\alpha}) \rho(z_{\beta}) r \phi'(r) \left(1 - \frac{3(z_{\alpha} - z_{\beta})^2}{r^2}\right) \quad (5.10)$$

where R_c is the cut-off radius, α, β represent united atom units of different type, $p_{min} = (\max[0, R_c^2 - (z_{\alpha} - z_{\beta})^2])^{\frac{1}{2}}$, and $r^2 = p^2 + (z_{\alpha} - z_{\beta})^2$.

A detailed derivation of the long-range correction to surface tension is included in the appendix.

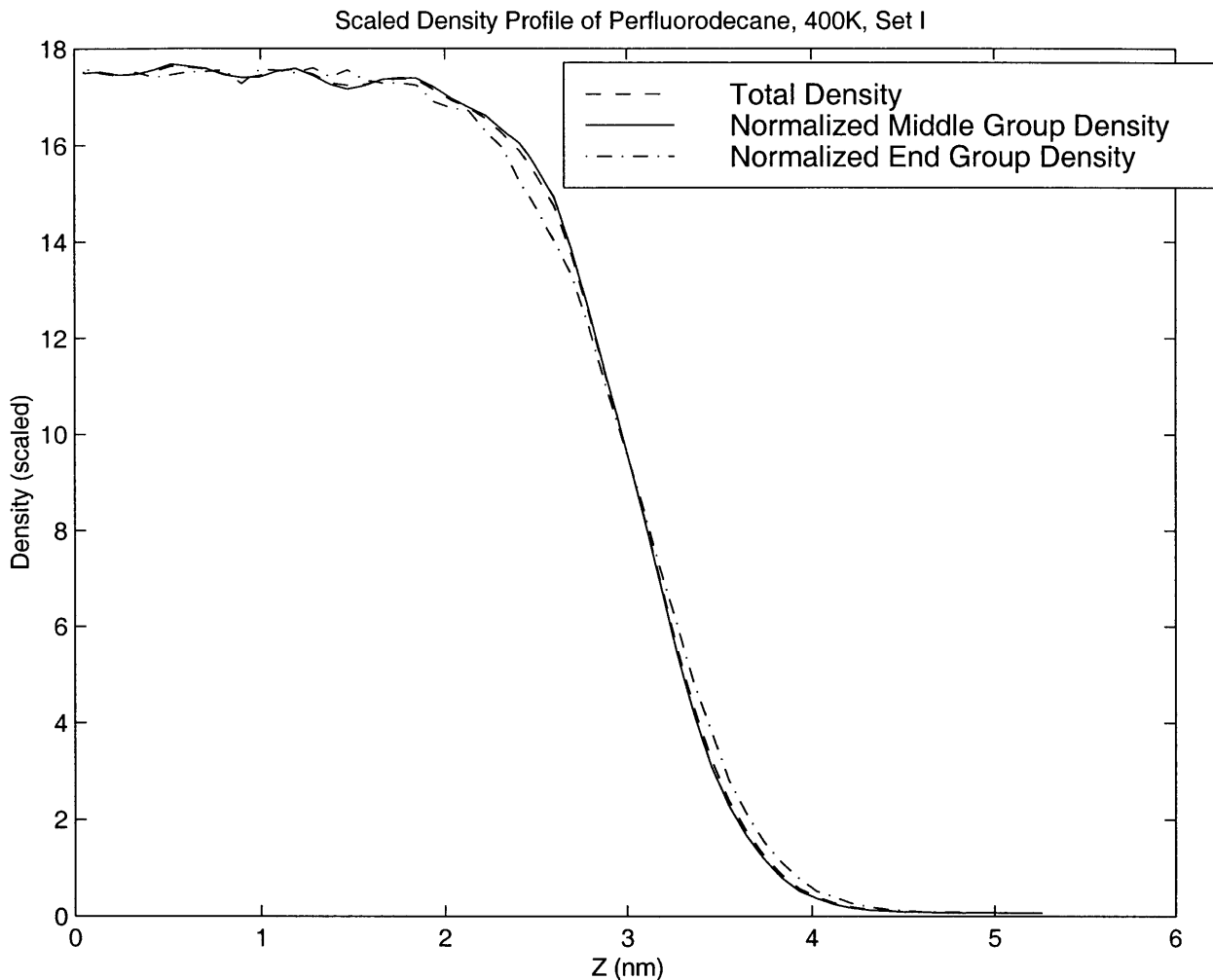


Figure 5-2: Scaled Number Density Profile, Simulation Set I

5.4 Simulation Results

5.4.1 Density Profiles

Figure 5-2, Figure 5-3, Figure 5-4, Figure 5-5, Figure 5-6, Figure 5-7, and Figure 5-8 are the total united-atom number density, and the scaled end-group and middle-group density profiles for simulation sets I–VII, respectively, averaged over the two symmetric interfaces. In all these sets, the density profiles can be divided into three distinct regions: (1) a bulk region in the middle of the simulation box ($z=0$), (2) an inner interfacial region (closer to the liquid phase), and (3) an outer interfacial region (closer to the vapor phase). In the bulk region, the total number density, the

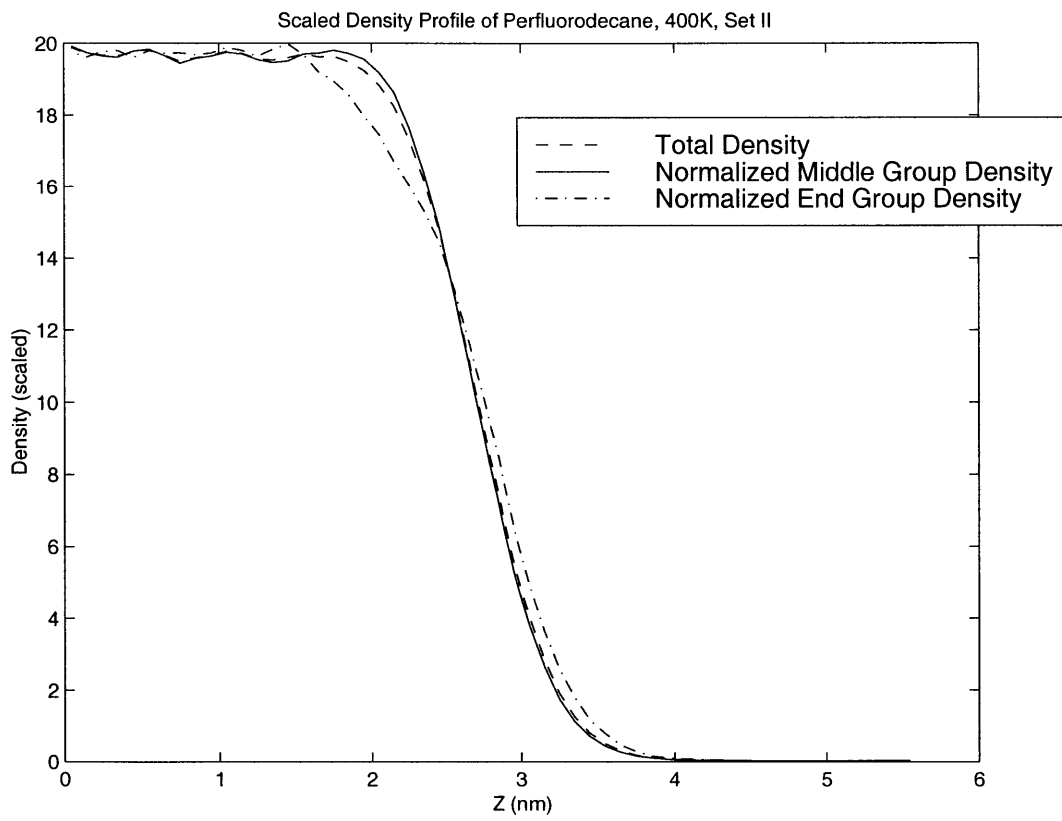


Figure 5-3: Scaled Number Density Profile, Simulation Set II

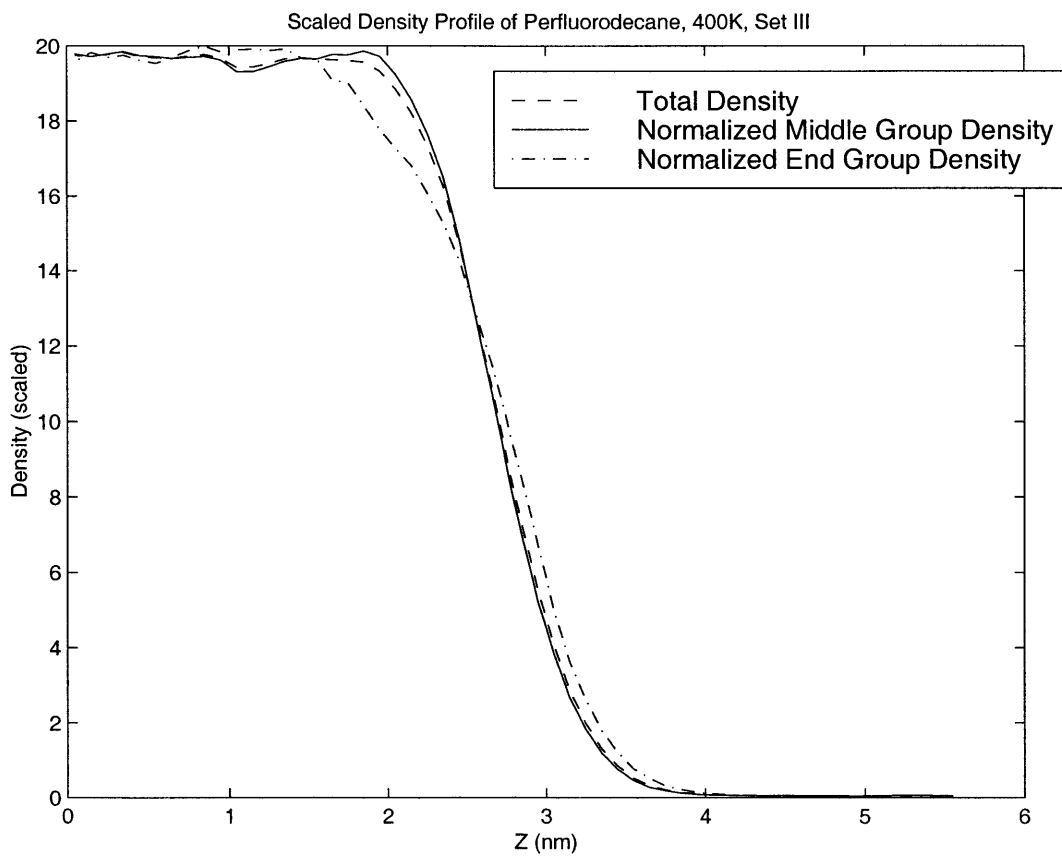


Figure 5-4: Scaled Number Density Profile, Simulation Set III

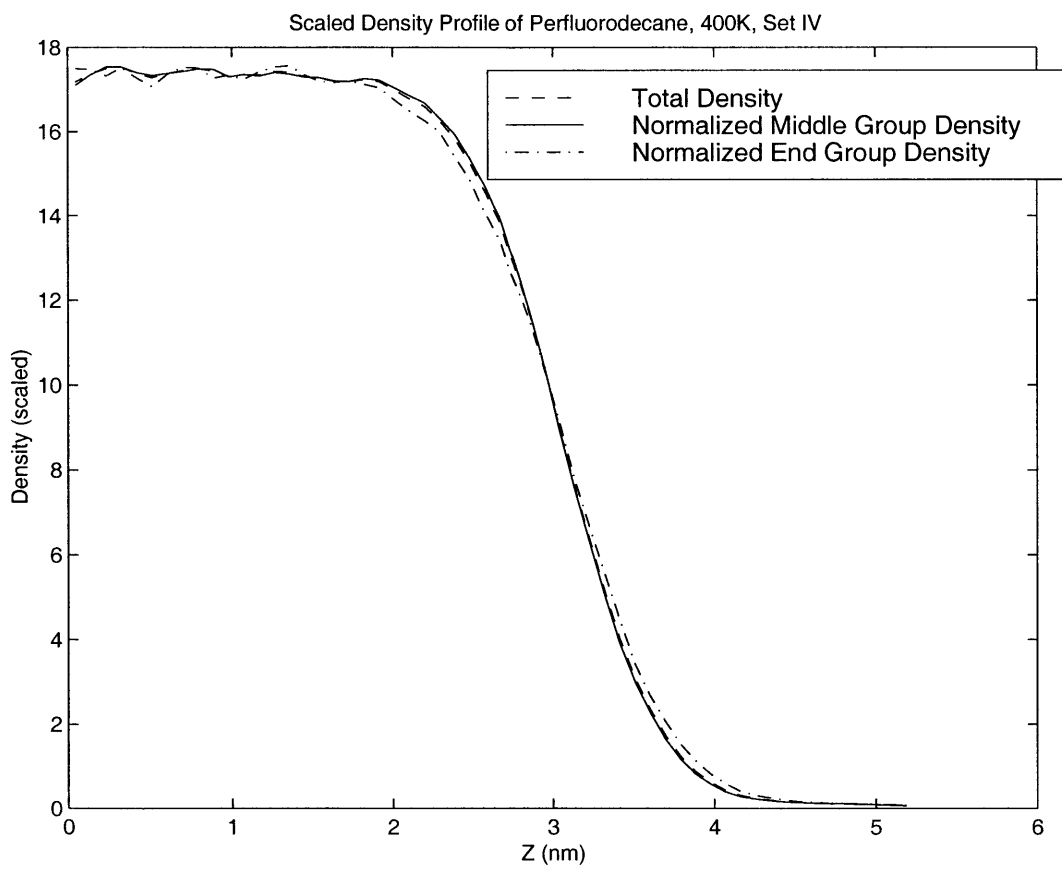


Figure 5-5: Scaled Number Density Profile, Simulation Set IV

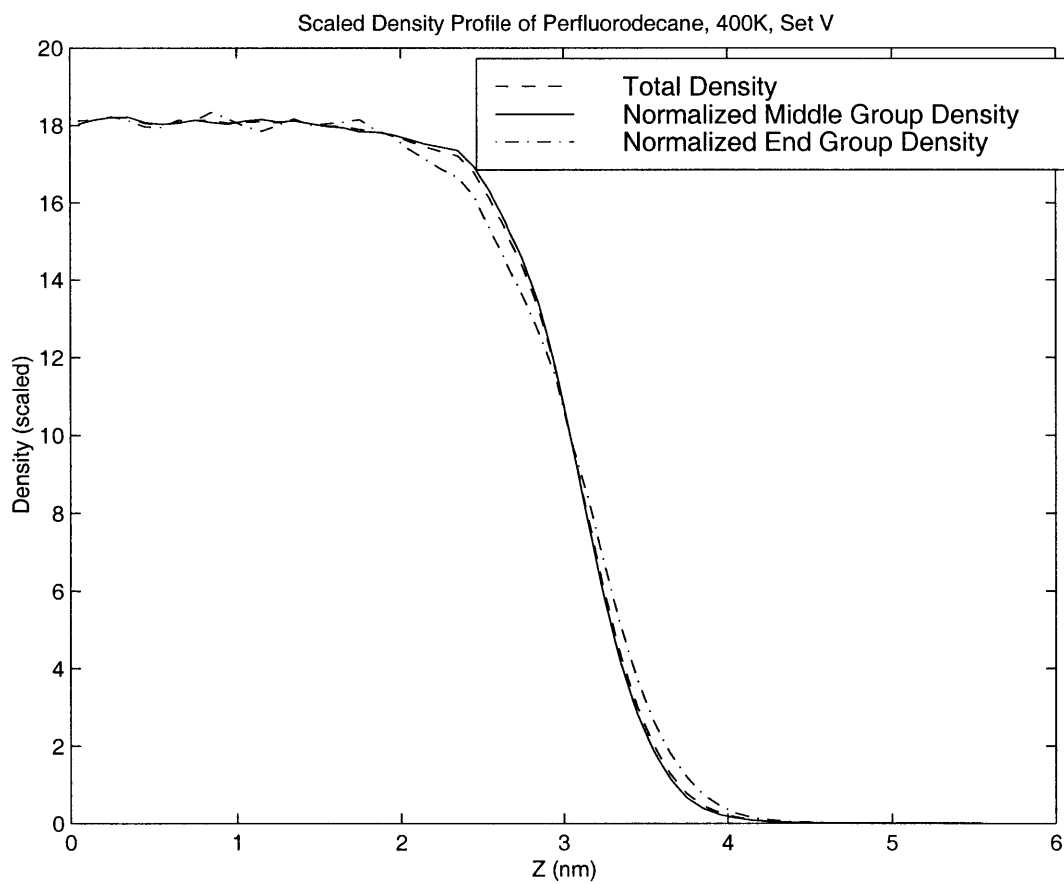


Figure 5-6: Scaled Number Density Profile, Simulation Set V

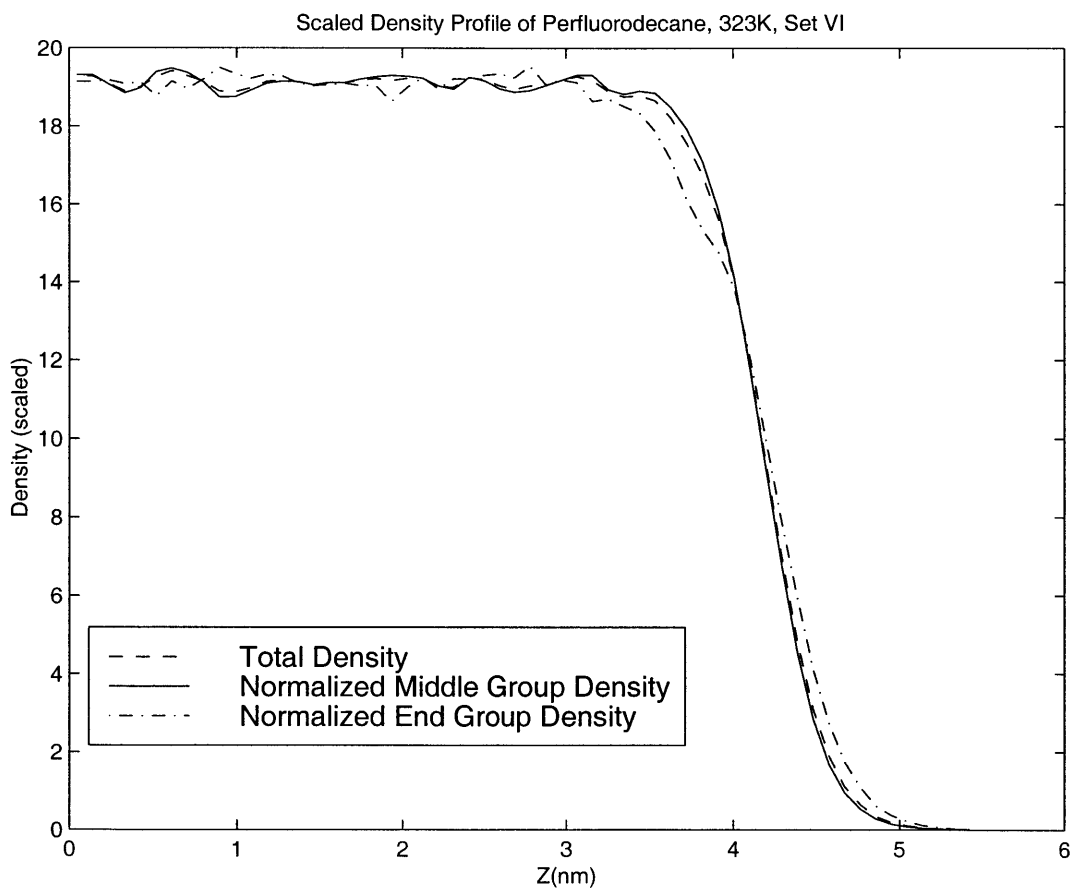


Figure 5-7: Scaled Number Density Profile, Simulation Set VI

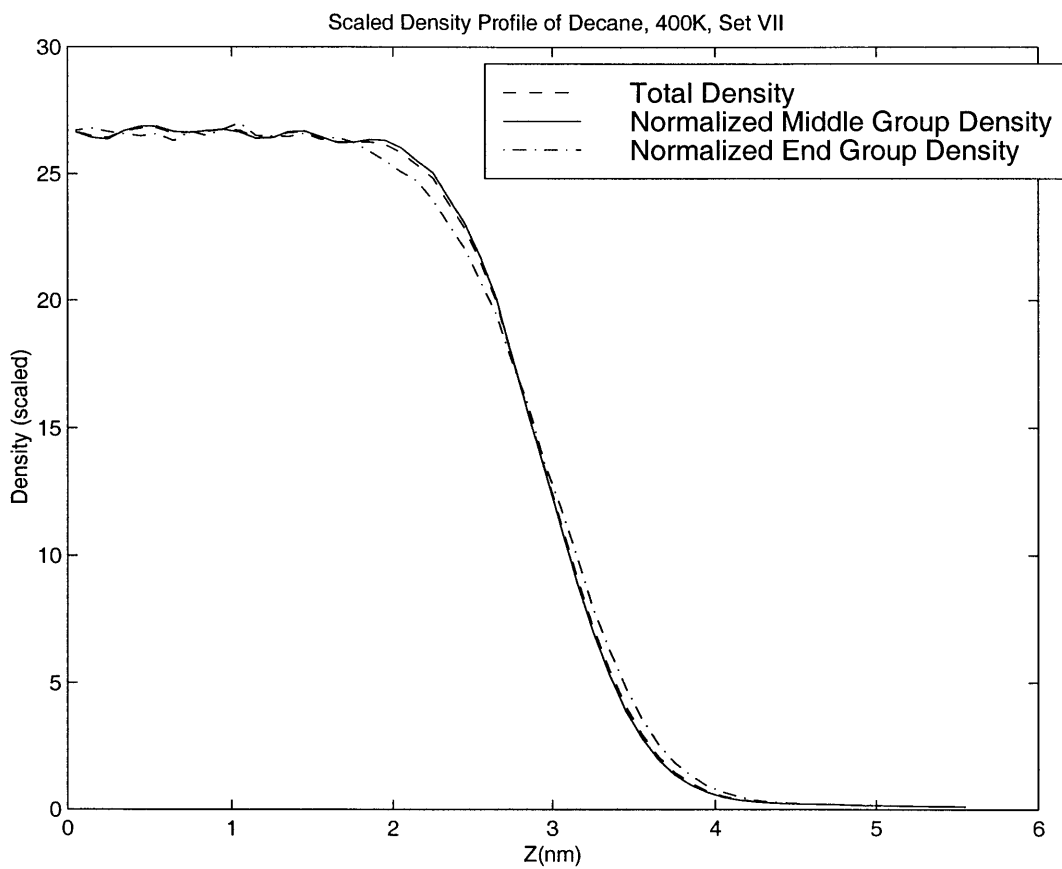


Figure 5-8: Scaled Number Density Profile, Simulation Set VII

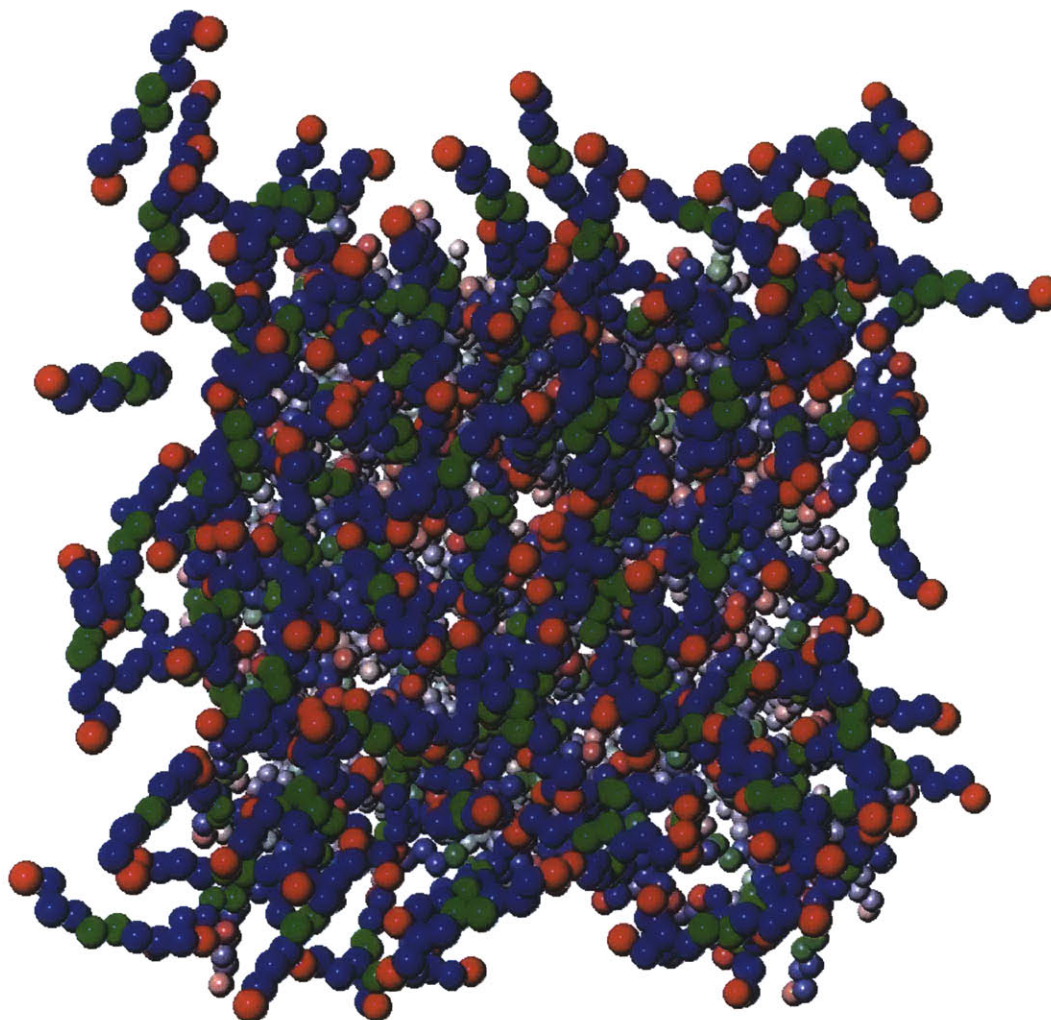


Figure 5-9: Liquid-Vapor Interface, n-Perfluorodecane at 323 K, Top view, Interface is in the plane of the page (z axis perpendicular to the plane of the page), $L_x=L_y=5.0\text{nm}$, Periodic boundary condition in all directions, End groups labeled in red, Center groups labeled in green.

scaled end and middle group number density are almost indistinguishable. In the outer interfacial region, there is an excess of the end groups. This is due to the gain of entropy by placing end groups at the interface. In the inner interfacial region, there is an excess of middle segments. This is explained by the chain-connectivity of the middle and the end groups.

Figure 5-9, Figure 5-10 are snapshots of the liquid-vapor interface for *n*-perfluorodecane at 323 K.

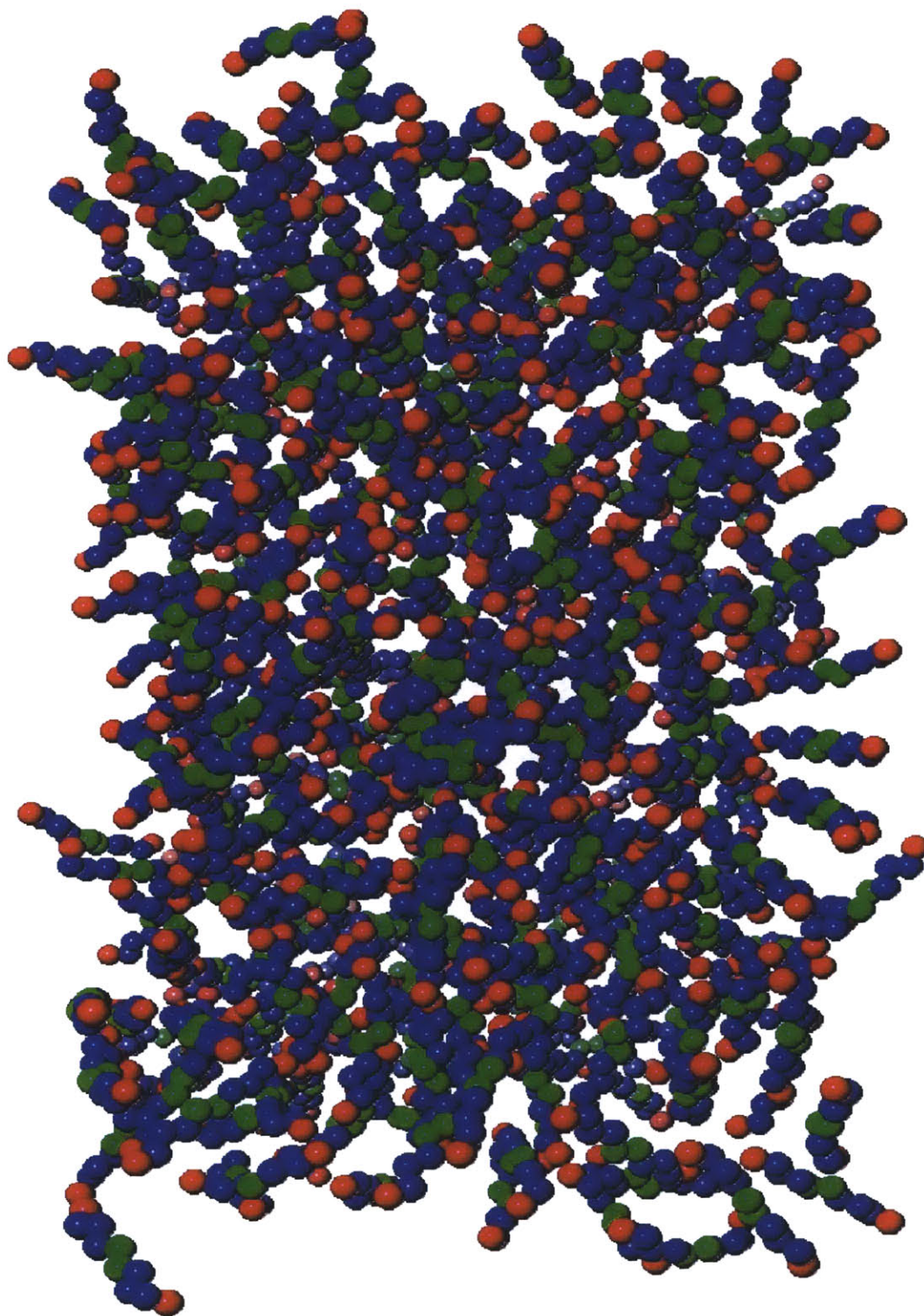


Figure 5-10: Liquid-Vapor Interface, n-Perfluorodecane at 323 K, Side view, Interface is perpendicular to the plane of the page and at the top and bottom of the page, $L_x=L_y=5.0\text{nm}$, Periodic boundary condition in all directions, End groups labeled in red; Center groups labeled in green.

Table 5.2: Liquid-Vapor Interface Simulation Results

Set	ρ_{bulk}^{exp} C/nm^3	ρ_{bulk}^{sim} C/nm^3	γ_{exp} $dyne/cm$	γ_{tot}^{sim} $dyne/cm$	W_{10-90} nm
I	17.7	17.5±0.1	7.9	12.5±0.9	1.21
II	17.7	19.7±0.1	7.9	17.0±0.8	0.95
III	17.7	19.7±0.2	7.9	17.3±0.7	1.05
IV	17.7	17.4±0.1	7.9	12.6±0.8	1.28
V	17.7	18.1±0.1	7.9	13.5±0.9	1.10
VI	19.7	19.1±0.2	13.2	21.1±1.1	0.80
VII	27.1	26.7±0.2	14.0	17.2±0.8	1.25

5.4.2 Effect of Various Factors on Surface Tension

Table 5.2 is a summary of the simulation predicted bulk liquid densities, surface tension, and interface width, along with the corresponding experimental data.

Comparison between data sets I and II demonstrated that our new Lennard-Jones forcefield performed significantly better than previous Lennard-Jones forcefields in predicting both bulk liquid density and the surface tension values. Set I predicted density of bulk liquid perfluoroalkanes density at 400 K to within 1.2 % of the experimental value, whereas Set II over-estimated the bulk liquid density by as much as 11.3%. Set II over-estimated the surface tension by 115%, whereas Set I reduces the discrepancy by about half. Set I also predicts a thicker interface than Set II.

Comparison between data sets II and III shows little difference in density and surface tension prediction, indicating modifying torsional energy barrier has little effect on obtaining correct density and surface tension values.

Comparison between data sets I and IV shows little difference in density and surface tension prediction, indicating fine-tuning C-C bond angle has little effect on

obtaining correct density and surface tension values.

Comparison between data sets I and V demonstrates increasing cut-off radius in simulation will increase the both the bulk liquid density and surface tension slightly. Hence, despite the significant contribution to surface tension from interactions beyond the 2.5σ (1.35 nm) cut-off radius, it is unlikely that using longer cut-off distance instead of mean-field correction will lead better prediction of the surface tension values.

Data set VI shows the new Lennard-Jones forcefield predicts the bulk liquid density of perfluorodecane at a lower temperature with reasonable accuracy (within 3%). This confirms that the new forcefield is transferable over a range of compounds and a range of temperature. Comparison between data set I and VI demonstrates the new Lennard-Jones forcefield correctly predicted the trends that bulk density and surface tension increases as temperature decreases. The predicted interface width is significantly lower at lower temperature.

Data set I and VII show that using our newly developed Lennard-Jones forcefield, along with OPLS forcefield for hydrocarbon, simulation can predict, beyond simulation uncertainty, that perfluoroalkanes have lower surface tension than their alkane counterparts under the same condition.

5.4.3 Chain Orientation

To investigate the effect of the liquid-vapor interface on the orientation of the perfluoroalkane chain molecules, an orientation parameter P is defined as

$$P(z) = \frac{1}{2} \langle 3\cos^2\theta - 1 \rangle \quad (5.11)$$

where θ is the angle between the interface normal and the vector connecting a pair of united-atoms that are two units apart in the chain (i.e., united-atoms k and $k + 2$), and the average, $\langle \rangle$, is taken over all timesteps and vectors within a specified slice over a region in the z -direction. A vector connecting united-atoms k and $k + 2$ is assigned the z -value of the midpoint between the two united-atoms. According to

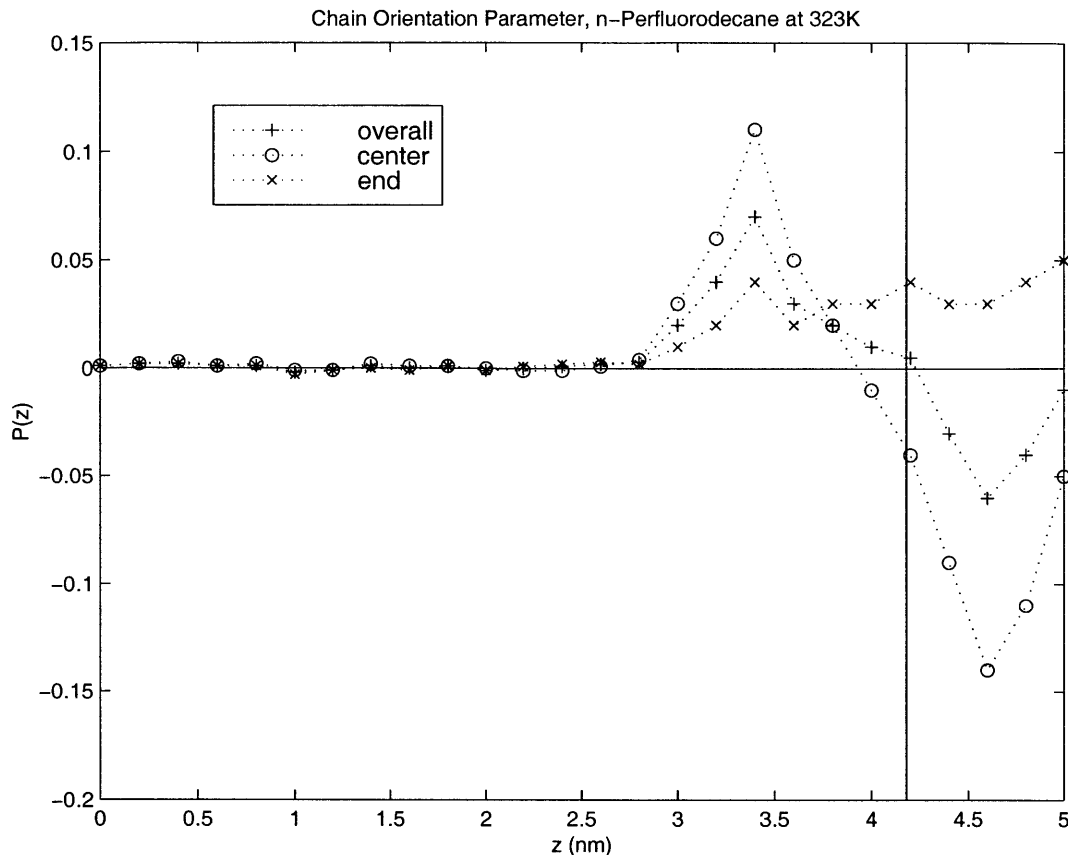


Figure 5-11: Chain Orientation Parameters, n-perfluorodecane at 323 K

this definition, $P = 0$ corresponds to completely random orientation of the molecules in the region; $P = -0.5$ corresponds to the case where all vectors are parallel to the interface; and $P = 1$ corresponds to the case where all vectors are perpendicular to the interface. $P(z)$ values can be calculated for all vectors on the chain molecules, or for vectors of selected segment of the chain.

Figure 5-11 and Figure 5-12 depict the chain orientation parameter $P(z)$ for *n*-perfluorodecane at 323 K (simulation set VI) and 400 K (simulation set I), respectively. In each figure, there are three different sets of $P(z)$ values computed from: (1) all vectors, (2) the two vectors at the end of the chain (i.e., 1-3 and 8-10), and (3) the two vectors at the center of the chain (i.e., 4-6 and 5-7). The vertical line indicate where the Gibbs equal-molar dividing surface is and the horizontal line at $P(z) = 0$ corresponds to randomly oriented chains. As Figure 5-11 and Figure 5-12 show that the chain orientation behavior at the interface can be divided into two regions. In the

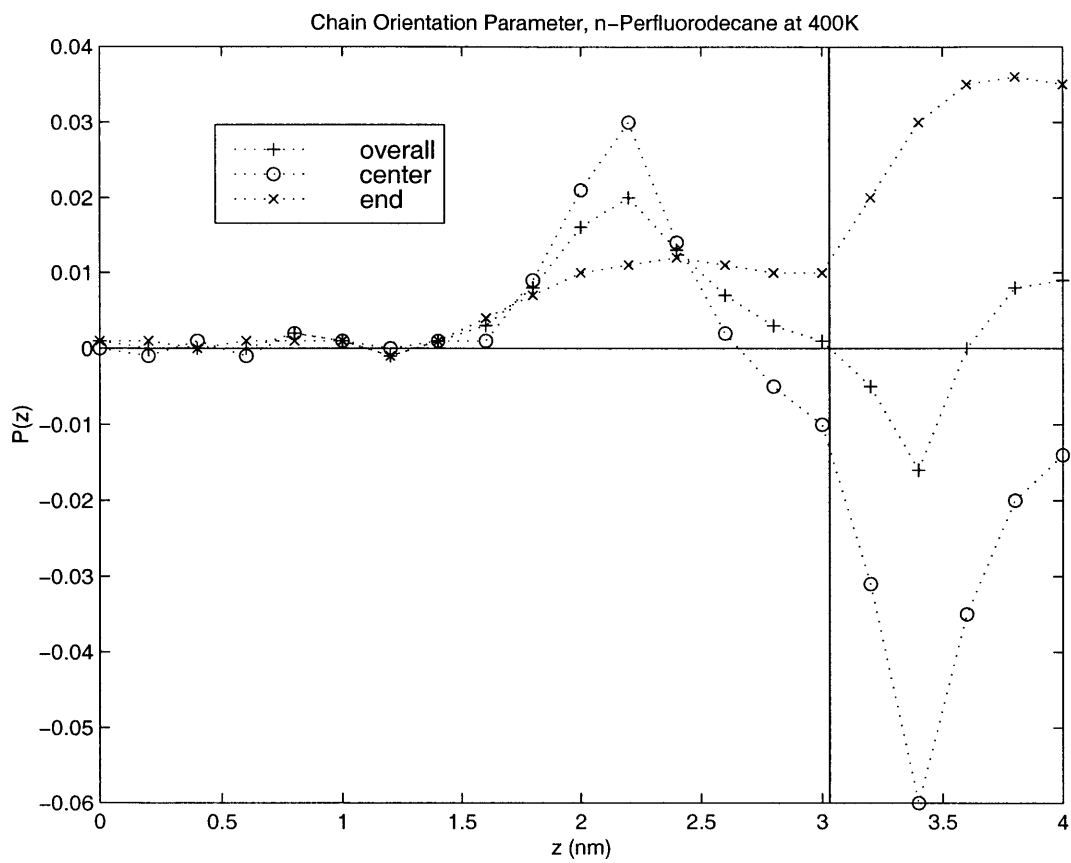


Figure 5-12: Chain Orientation Parameters, n-perfluorodecane at 400 K

region closer to the vapor phase, $P(z)$ is less than 0, suggesting that the molecules prefer to align parallel to the interface. In the region closer to the liquid phase, $P(z)$ is positive, indicating that the molecules prefer to orient slightly perpendicular to the interface. This effect is stronger for *n*-perfluorodecane at 323 *K* and weaker at 400 *K*.

When the $P(z)$ values for chain ends and chain centers, the difference is even more pronounced. In the vapor-side of the interface, the chain ends are oriented perpendicular to the interface, while the chain centers are parallel to it. In this region, most of the chain ends are connected to a chain centered more deeply into the liquid. For the chain ends to stick out to the vapor phase, the bonds must be perpendicular to the interface. On the other hand, if a vector in the middle of the chain were perpendicular to the interface, some portion of the molecule would have to stick out to the low density region and thus have a higher energy. These chain orientation behaviors are very similar to those of *n*-decane as previously reported by Harris.^[39]

5.5 Conclusion

In the simulation of liquid-vapor interface of perfluorodecane, the Lennard-Jones forcefield developed through fitting to bulk liquid properties of perfluoroalkanes performed significantly better than the previous Lennard-Jones forcefield. This new forcefield accurately predicted the density of perfluorodecane in the bulk region at two different temperatures within 1-3% of the experimental values. It also correctly predicted the trend that bulk liquid density and surface tension increases as temperature decreases. Even though the new Lennard-Jones forcefield still over-predicted the experimental surface tension values, compared to liquid-vapor simulation for alkanes with OPLS, it can predict, beyond simulation uncertainty, that perfluoroalkanes have lower surface tension than their alkane counterparts.

The effects of different torsion energy barrier, wider C-C bond angle, and longer cut-off radius were also examined. These factors have relatively insignificant effect

on the bulk liquid density and the surface tension prediction.

Chapter 6

Perturbation Theory and Three-Body Interaction

6.1 Introduction

In the molecular dynamics simulations of bulk and liquid-vapor interfaces performed in the preceding chapters, the intermolecular interactions were all characterized as Lennard-Jones 12-6 interactions between pairs of united-atoms. However, strictly speaking, the total energy of three or more particles interacting simultaneously is not exactly equal to the sum of all the isolated pair-wise interaction energies.^[108, 109, 13] The presence of the other nearby particles perturbs the pair-wise interactions and the total energy of a group of N particles should be expressed more accurately as:

$$U_N = \sum_{j>i} U_{ij} + \sum_{k>j} \sum_{j>i} U_{ijk} + \dots \quad (6.1)$$

While the first term, the sum of pair energies, is dominant and contributes to most of the total energy, in a dense system such as liquid or solid, the contributions from many-body interactions (e.g., three-body interaction – the second term in Eq. (6.1)) may not be negligible, and can account up to 10% of the total intermolecular energy.

Despite the significant contribution of the many-body interactions, many-body

terms are rarely explicitly included in computer simulations due to the high-cost associated with summation over triplets and quadruplets, etc. Instead, the effect of many-body interactions are often included by defining an “effective” two-body interaction term. This effective two-body interaction, unlike the “isolated” two-body interaction, may be a function of density and temperature. When the effective two-body interactions are used in simulations of bulk liquid, the results are quite satisfactory. However, at a liquid-vapor interface where density profile is rapidly changing, the effective two-body interaction approach may no longer be accurate and explicit three-body interaction terms may be needed. Lee, Baker, and Pounds included the three-body interaction in the liquid-vapor interface of argon using perturbation theory and demonstrated (1) three-body interaction contribute negatively to the value of surface tension for argon; (2) explicit inclusion of three-body interaction significantly improved the surface tension of argon over a range of temperatures. Subsequently, similar results were obtained for Kr and Xe.^[73, 72]

In this chapter, perturbation theory will be applied to estimate the effect of three-body interaction in the alkane and perfluoroalkane liquid-vapor systems, and investigate whether the inclusion of three-body interaction can account for the over-estimation of the surface tension values. The result of the Lennard-Jones molecular dynamics simulation described in the previous chapter will be used as the reference system, and the three-body interaction will be treated as the perturbation under the assumption that the three-body interaction has little effect on the structure of the system.

6.2 Three-Body Interaction

In the non-polar cases, the major contribution to the non-additivity arises from the long-range dispersion energy. Dispersion forces arise from the electronic motions getting into phase. Lennard-Jones interaction is a result of two-body dispersion interaction. As Figure 6-1 demonstrated, if a third molecule C approaches a pair (A-B) linearly, correlation of the electrons in molecules A and B can be enhanced by both

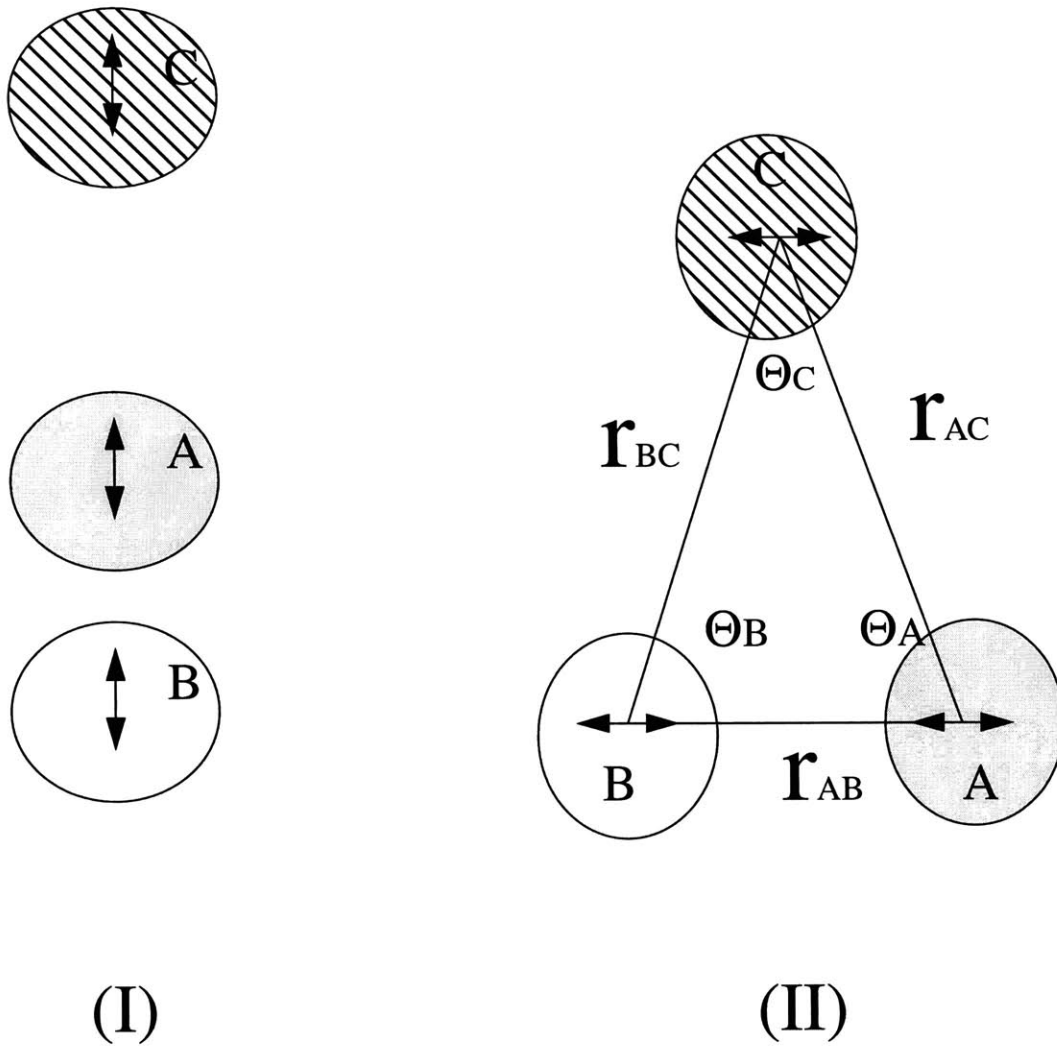


Figure 6-1: Non-additivity in Long-range Dispersion Energy. (I) Case where correlation between A and B is enhanced by C; (II) Case where correlation between A and B is reduced by C, angles and separations in three-body system indicated

correlating with C. Thus the dispersion forces are increased for this geometry. On the other hand, if C forms an approximately equilateral triangle with A and B, the coupling of electronic motion between A and B is reduced when both try to interact with C. This leads to a weakening of the dispersion interaction.^[110]

The leading three-body dispersion correction term arising from this behavior was first evaluated by Axilrod and Teller and is called the triple-dipole contribution,

$$U_{(DDD)_3} = \frac{\nu_{ABC}(1 + 3\cos\theta_A\cos\theta_B\cos\theta_C)}{r_{AB}^3 r_{BC}^3 r_{AC}^3} \quad (6.2)$$

where the internal angles θ 's and interatomic distances r 's are defined in part (II) of Figure 6-1, and ν is a coefficient which can be estimated from the polarizability and the leading coefficient of dispersion energy.

The sign of the Axilrod-Teller triple-dipole energy depends on the internal angles of this triangle. For acute triangles, the triple-dipole energy is always positive, and for most obtuse triangles it is negative. This explains why near-linear arrays of molecules are stabilized by this effect, while most triangular arrangements are destabilized. For most molecular configurations appropriate to solids and liquids the net energy from the triple-dipole correction is positive.^[109]

Other three body terms are $(DDQ)_3$, $(DQQ)_3$, $(QQQ)_3$, and $(DDO)_3$, where D, Q, and O represent dipole, quadrupole, and octopole contributions, respectively. These terms are less important compared to the Axilrod-Teller term, because instead of being proportional to r^{-9} , they are proportional to the r^{-11} , r^{-13} , r^{-15} and even higher orders, and hence fall off very rapidly with distances.

6.3 Perturbation Theory

Perturbation theory is originally developed based on the assumption that the structure of a liquid is primarily determined by the short-range repulsive part of the pair potential and that the relatively longer-range attractive part of the the potential provides a net force that gives a somewhat uniform attractive potential. This assumption has been proven to be quite accurate. For instance, X-ray scattering experiments

and molecular dynamics studies have both demonstrated that for simple monoatomic systems, except for the discontinuity at $r = \sigma$, the radial distribution function of a real fluid is very similar to that of a hard sphere fluid, which has no attractive part in the potential.

Much of the work on perturbation theory used the hard sphere potential as the reference potential since the hard sphere system is very well known, from both computer simulation studies and from the Percus-Yevick equation. However, the choice of reference system is not limited to the hard-sphere system. The use of more realistic reference system can improve the accuracy of the calculations.

6.4 Estimate Three-body Contribution to Surface Tension with Perturbation Theory: Calculation and Results

Toxvaerd^[111] first extended the Barker-Henderson perturbation theory^[112] to non-uniform systems. Lee, Barker and Pounds first derived the expression for surface tension with the three-body interaction included.^[13]

Using hard-sphere as the reference system and placing Gibb's equimolar dividing surface at $z = 0$, Lee *et al.* obtained an expression for surface tension γ :

$$\gamma = \lim_{h \rightarrow \infty} \left[\int_{-h}^h \rho(z_j) f(z_j) dz_j - h(\rho(-h)f(-h) + \rho(h)f(h)) \right] \quad (6.3)$$

where

$$f = f_0 + f_1 + f_2 + f_3 \quad (6.4)$$

$$f_1(z_j) = \frac{1}{2} \int_V u_{2b}(r) \rho(z_j + z) g_2^{(0)}(r, \rho(z_j)) d\mathbf{r} \quad (6.5)$$

$$f_2(z_j) = -\frac{1}{4} \int_V u_{2b}^2(r) \rho(z_j + z) g_2^{(0)}(r, \rho(z_j)) \left(\frac{\partial \rho(z_j + z)}{\partial p} \right)_0 d\mathbf{r} \quad (6.6)$$

$$f_3(z_j) = \frac{1}{6} \int_V \int_V u_{3b}(r_1, r_2, r_3) g_3^{(0)}(\mathbf{r}_2, \mathbf{r}_3, \rho(z_j + z_2), \rho(z_j + z_3)) \cdot \rho(z_j + z_2) \rho(z_j + z_3) d\mathbf{r}_2 d\mathbf{r}_3 \quad (6.7)$$

$f_1(z_j)$ and $f_2(z_j)$ are the perturbation terms resulted from the two-body interactions. $f_3(z_j)$ is the perturbation term that arises from the three-body interaction.

Since we are using the result of the MD simulation as our reference system, $f_3(z_j)$ is the only perturbation term in our system. If we make the assumption that three-body interaction has little effect on the structure of the liquid-vapor interface, we can then use the radial distribution function calculated from the MD results to calculate the three-body contribution to surface tension. For simplicity, we also adopt the usual superposition approximation:

$$g_3^{(0)}(\mathbf{r}_2, \mathbf{r}_3, \rho(z_j + z_2), \rho(z_j + z_3)) = g_2^{(0)}(r_1, \rho(z_j)) \cdot g_2^{(0)}(r_2, \rho(z_j)) \cdot g_2^{(0)}(r_3, \rho(z_j)) \quad (6.8)$$

Once we know the density, pair correlation function, and the three-body coefficients, we can use Eq. (6.3) and Eq. (6.7) to calculate the three-body contribution to surface tension.

6.4.1 Estimating Three-Body Coefficient

For cases where the Axilrod-Teller three-body coefficients are not readily available, they can be estimated with^[109]

$$\nu_{3b} \cong -\frac{3}{4} \alpha C_6 \quad (6.9)$$

where α is the polarizability of the atom and C_6 is the leading coefficient in the dispersion energy.

The polarizability of the homologous series of liner alkanes and linear perfluoroalkanes are listed in Table 3.7 and Table 3.8. Since polarizability is a group-additive property,^[109, 66] α values for united-atoms CH_2 , CH_3 , CF_2 and CF_3 can be calcu-

Table 6.1: Polarizability of Linear Perfluoroalkanes

Compound	α (10^{-24}cm^3)
CH_2	1.88
CH_3	2.21
CF_2	2.39
CF_3	3.27

lated and are listed in Table 6.1. C_6 values for CH_2 and CH_3 are estimated with the coefficient of the OPLS r^{-6} term.^[79, 91, 92] The C_6 values for CF_2 and CF_3 are estimated with the coefficient of the r^{-6} term in our new Lennard-Jones forcefield.

Table 6.2 list the values of three-body coefficients for different united-atoms. To further simplify the perturbation theory calculation, instead of accounting the different triplet combination separately, a “blended” coefficient is used:

For decane,

$$\frac{\text{Number of } CH_3}{\text{Number of } CH_2} = \frac{1}{4} \quad (6.10)$$

so

$$\nu_{CH_x-CH_x-CH_x} = (\nu_{CH_3-CH_3-CH_3})^{1/5} (\nu_{CH_2-CH_2-CH_2})^{4/5} \quad (6.11)$$

$$= 1.85 \times 10^{-82} \text{erg} \cdot \text{cm}^9 \quad (6.12)$$

and similarly, for perfluorodecane,

$$\frac{\text{Number of } CF_3}{\text{Number of } CF_2} = \frac{1}{4} \quad (6.13)$$

Table 6.2: Three-Body Coefficients

3-Body Coeff.	Value
$\nu_{CH_3-CH_3-CH_3}$	$2.85x10^{-82}erg \cdot cm^9$
$\nu_{CH_2-CH_2-CH_2}$	$1.66x10^{-82}erg \cdot cm^9$
$\nu_{CF_3-CF_3-CF_3}$	$11.58x10^{-82}erg \cdot cm^9$
$\nu_{CF_2-CF_2-CF_2}$	$6.80x10^{-82}erg \cdot cm^9$

so

$$\nu_{CF_x-CF_x-CF_x} = (\nu_{CF_3-CF_3-CF_3})^{1/5}(\nu_{CF_2-CF_2-CF_2})^{4/5} \quad (6.14)$$

$$= 7.56x10^{-82}erg \cdot cm^9 \quad (6.15)$$

6.4.2 Calculate Transverse Pair Correlation Function

The radial distribution function $g(r)$ displays the static structure of matter, It is a measure of how particles organizes themselves around one another. $\rho g(r)d\mathbf{r}$ is the non-normalized probability of finding a second particle at distance $d\mathbf{r}$ given that there is a particle at the origin of \mathbf{r} . It is defined as the probability of finding a pair of particles at a distance r apart, relative to the probability expected for a completely random distribution at the same density.

Transverse (parallel to the interface) pair correlation functions are calculated according to the methods described by Harris, Gryko and Rice.^[20] For each thin slab sliced parallel to the interface, the pair correlation function is computed from a histogram of the separations between all pairs of particles within the slab. The expression

for the pair correlation function is:

$$g(r) = \frac{2N(r, \Delta r)}{V_{int} N_T^2} V_s \quad (6.16)$$

where $N(r, \Delta r)$ is the average number of pairs of particles within the section whose separations are between r and $r + \Delta r$; V_{int} is the average volume of the intersection of the slab with the space between the spheres of radius r and $r + \Delta r$; V_s is the total volume of the thin slab; and N_T is the average number of particles within the slab. Intervals of constant $\Delta(r^2)$ are used so that the average statistical fluctuation in $g(r)$ at each point is roughly independent of r . Only pairs that interact via Lennard-Jones interactions in MD are counted.

Figure 6-2, Figure 6-3 and Figure 6-4 show the transverse pair correlation functions for perfluorodecane at 400 K, 323K, and decane at 400 K, respectively. Moving from bulk phase to the interface, the overall upward shift in the transverse pair correlation functions is due to the sharp decrease in density within the finite width of slab. More accurate transverse $g(r)$ can be obtained by taking much thinner slices and averaging over a greater number of configurations.

6.4.3 Density Profile

Gibb's equimolar density profiles are obtained from the density profiles shown in the last chapter by shifting the location of $z = 0$ to the Gibb's equimolar dividing line. The density profiles are then fitted to hyperbolic functions.

6.4.4 Three-Body Contribution to Surface Tension

Table 6.3 shows the three-body contribution to surface tension. The \pm values account for the error associated with sectioning $g(r)$ differently and with integration.

These results demonstrate that the Axilrod-Teller three-body potential have a significant, negative contribution to the surface tension values of linear alkanes and perfluoroalkanes. By combining MD surface tension prediction and the three-body contribution to surface tension, we have improved the prediction of surface tension.

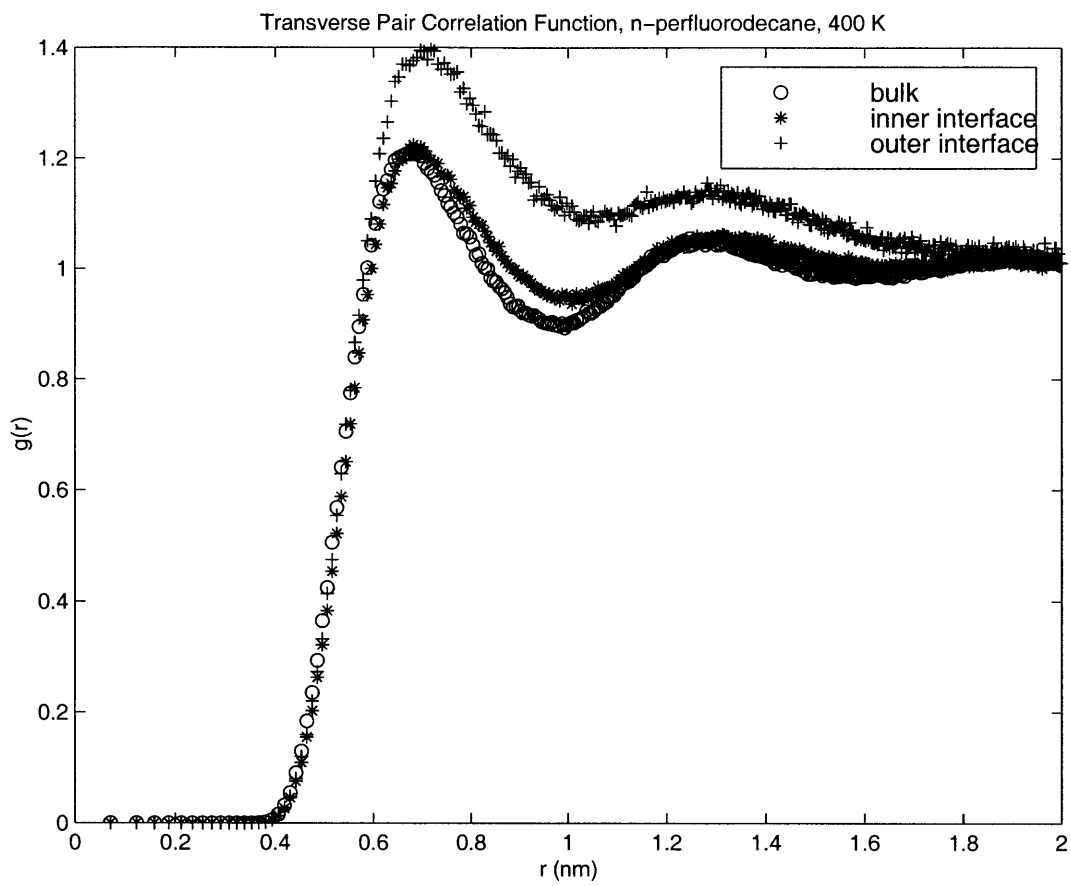


Figure 6-2: Transverse Pair Correlation Function, n-perfluorodecane, 400K

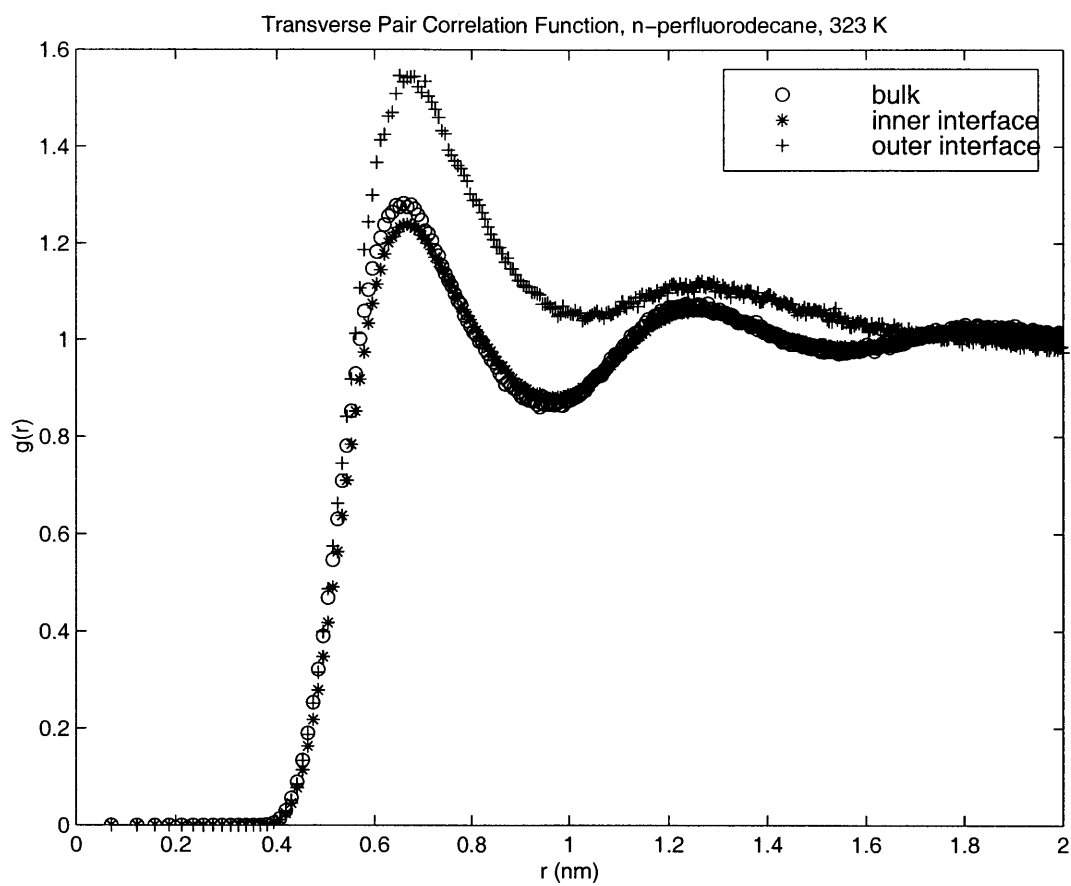


Figure 6-3: Transverse Pair Correlation Function, n-perfluorodecane, 323K

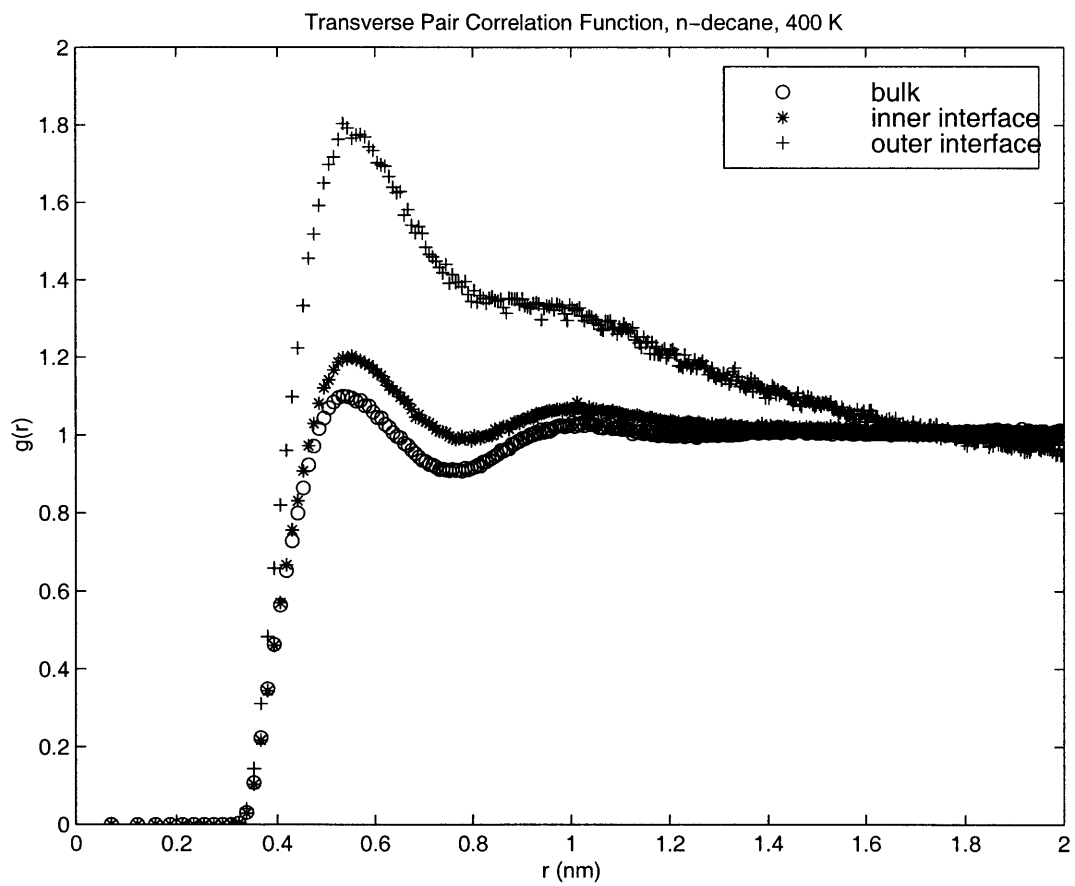


Figure 6-4: Transverse Pair Correlation Function, n-decane, 400K

Table 6.3: Three-Body Contribution to Surface Tension

Molecule	Temp. (K)	γ_{exp} (dyne/cm)	γ_{MD+L} (dyne/cm)	γ_{3b} (dyne/cm)	$\gamma_{MD+L+3b}$ (dyne/cm)
<i>n</i> -Perfluorodecane	400	7.9	12.5 ± 1.0	-3.9 ± 1.0	8.6 ± 2.0
<i>n</i> -Perfluorodecane	323	13.5	21.1 ± 1.1	-5.0 ± 1.0	15.1 ± 2.1
<i>n</i> -Decane	400	14.0	17.2 ± 1.0	-3.5 ± 1.0	13.7 ± 2.0

Our results also show that the three-body contribution is more significant at lower temperature.

6.5 Conclusion

In this work, the contributions to surface tension by Axilrod-Teller three-body interaction for chain molecules such as alkanes and perfluoroalkanes are estimated using perturbation theory. This is the first time three-body interactions are accounted for explicitly in the oligomer liquid-vapor interface. Despite the many approximations made due to computation efficiency or lack of more accurate values, this work has shown that for oligomers such as decane and perfluorodecane, the Axilrod-Teller three-body potential has a significant and negative contribution to the surface tension values. The effect of the three-body interaction at the liquid-vapor interface maybe not be sufficiently included through the use of an “effective” two-body potential. Our results also suggests that the three-body contribution is more significant at lower temperature. By combining the results of molecular dynamics simulation and the perturbation theory, the predicted surface tension values for decane and perfluorodecane are in very good agreement with the experimental values.

Chapter 7

Summary and Suggested Future Work

In this work, a new set of united-atom Lennard-Jones forcefield parameters was developed from simulation of bulk liquid perfluoroalkanes. Compare to the Lennard-Jones forcefield used in previous studies, the new Lennard-Jones forcefield has the following features: (1) larger σ for perfluoromethylene group than perfluoromethyl group; (2) the energy well for perfluoromethyl group is much deeper than that of Hariharan *et al*; and (3) the energy well for perfluoromethylene is much shallower than that of Hariharan *et al*. This new forcefield significantly improved the simulation of bulk properties such as molar volume and heat of vaporization for the homologous series of linear perfluoroalkanes.

The new forcefield was used to simulate the liquid-vapor interface of perfluorodecane at two different temperatures, and resulted significantly improved prediction of surface tension and density values. It accurately predicted the density of perfluorodecane in the bulk region to within 1-3% of the experimental values. It also correctly predicted the trend that bulk liquid density and surface tension increases as temperature decreases. Even though the new Lennard-Jones forcefield still over-predicted the experimental surface tension value, used along with OPLS for alkanes, it can predict, beyond simulation uncertainty, that perfluoroalkanes have lower surface tension than their alkane counterparts.

Despite the significant improvement in bulk liquid density and surface tension using the new Lennard-Jones forcefield for linear perfluoroalkanes, the predicted surface tension values are still significantly higher than the experimentally measured values. The same trend is observed for linear alkanes as well. Using perturbation theory and the results of MD liquid-vapor interface simulation as the reference system, we estimated the contribution to surface tension from the three-body interactions. It has been shown that the three-body interactions have a negative contribution to surface tension value. The explicit inclusion of the three-body contribution brought the calculated surface tension to excellent agreement with experiment.

To more accurately account for the three-body interaction, one possible direction to explore in the future is to include an explicit three-body term during the MD simulation. To avoid the time-consuming summation over triplet during the simulation process, one can use the novel method suggested by Barker.^[73] A C_9/r^9 term can be included in the intermolecular/intramolecular forcefield during the MD simulation. The magnitude of C_9 can be evaluated from the averages of the three-body forcefield and of the sum over pair of $1/r^9$ for a relatively small number of slab of configurations generated using the pair forcefield alone. This procedure can be repeated until self-consistency is achieved. The difference between U_{ABC} and C_9/r^9 will be relatively small and can be treated as a very minor perturbation.

Another possible future direction is to apply the new Lennard-Jones forcefield to branched perfluoroalkanes, polytetrafluoroethylene, and diblocks of polytetrafluoroethylene and polyethylene. One can test whether the new forcefield correctly predicts the bulk and surface properties of these compounds. As another more stringent test for the new forcefield, one can examine the behavior of linear hydrocarbons and perfluorocarbons on PE and PTFE coated surfaces.

Furthermore, neutron scattering can be performed on selectively labeled alkanes and perfluoroalkanes. The neutron scattering experiments can measure radial distribution functions, interfacial width, and interfacial chain orientation. Results of these experiment can be compare to our simulated structural properties and provide further feedback.

Appendix A

Surface Tension and Its Long-Range Correction

The following is the derivation of the surface tension calculation from molecular virial and the long-range correction expression given by Harris^[39].

For the rectangular liquid-vapor interfacial system described in Figure 5-1 of dimensions $L_x L_y L_z$ where $L_x = L_y = L_1$ and total surface area $A = L_1^2$. The surface tension is

$$\frac{-kT}{Z} \left(\frac{\delta Z}{\delta A} \right)_{TVN} = \frac{-kT}{Z} \left[\frac{1}{4L_1} \left(\frac{\delta Z}{\delta L_x} \right)_{L_y L_z} + \frac{1}{4L_1} \left(\frac{\delta Z}{\delta L_y} \right)_{L_x L_z} - \frac{L_z}{2L_1^2} \left(\frac{\delta Z}{\delta L_z} \right)_{L_x L_y} \right] \quad (\text{A.1})$$

where

$$Z = \int_{L_x L_y L_z} \exp[-U(\mathbf{r}^{3Nm})/kT] d\mathbf{r}^{3Nm} \quad (\text{A.2})$$

is the configurational integral for a system with N molecules containing m atoms each.

Equation (A.1) describes the effects of a change in area by extending the lengths in the x and y directions while contracting the box in the z direction to maintain a constant volume. To evaluate the derivatives in Eq. (A.1), we must transform the integral into a form where the limits of integration are fixed and the system dimensions appear in the integrand.

To obtain the molecular virial, the center of mass coordinates are rescaled by the transformation

$$\begin{aligned} r_{aj}^\alpha &= L_\alpha X_a^\alpha + \sigma_{aj}^\alpha \\ 1 \leq j \leq m \end{aligned} \quad (\text{A.3})$$

where the $L_\alpha X_a^\alpha$ are the α components of the center of mass and the σ_{aj}^α are the displacements from the molecular center of mass of atom j on the molecule a in the α direction.

In the transformed coordinate system, the system dimensions only appear in the integrand – in the potential energy and the Jacobian determinant. The limits of integration of the σ_{aj}^α are irrelevant and can be set to $\pm\infty$ because the intramolecular interactions insure that the Boltzmann factor vanishes when $|\sigma_{aj}^\alpha|$ get large. When the partition function is differentiated, only derivatives of the intermolecular interactions appear, because the intramolecular interactions are functions of the σ_{aj}^α , which are not scaled by the box dimension in Eq. (A.3). Thus,

$$\begin{aligned} \frac{dZ}{dL_\alpha} &= \int_{X_a^\alpha=0}^1 d\mathbf{X}^{3N} d\sigma^{3N(m-1)} \frac{d|J|}{dL_\alpha} \exp[-U(\mathbf{r}^{3Nm})/kT] \\ &- \frac{1}{kTL_\alpha} \int_{X_a^\alpha=0}^1 d\mathbf{X}^{3N} d\sigma^{3N(m-1)} |J| \exp[-U(\mathbf{r}^{3Nm})/kT] \sum_{a<b} \sum_{kl} \frac{R_{ab}^\alpha \mathbf{r}_{abkl}^\alpha}{|\mathbf{r}_{abkl}|} \phi'(|\mathbf{r}_{abkl}|) \end{aligned} \quad (\text{A.4})$$

where $|J|$ is the Jacobian determinant of the transformation, $R_{ab}^\alpha \mathbf{r} = L_\alpha X_{ab}^\alpha$ is the α component of the vector between the centers of mass of molecules a and b . \mathbf{r}_{abkl} is the vector between atoms k on molecule i and l on molecule j , and ϕ' is the derivative of the Lennard-Jones potential with respect to distance.

Substitue Eq. (A.4) into Eq. (A.1), we obtain the expression of surface tension in terms of molecular virials:

$$\gamma = \frac{1}{2A} \langle V_{zz} - V_{xx} - V_{yy} \rangle \quad , \quad (\text{A.5})$$

where A is the total surface area. $V_{\alpha\alpha}$, the molecular virial tensor, is defined as

$$V_{\alpha\alpha} = \sum_{a < b} F_{ab}^{\alpha} R_{ab}^{\alpha} \quad . \quad (\text{A.6})$$

The sum is taken over all molecules a and b. R_{ab}^{α} is the α component in direction $\alpha = (x, y, \text{or } z)$ of the vector between the centers of mass of molecules a and b. F_{ab}^{α} is the total force between the molecules:

$$F_{ab}^{\alpha} = - \sum_{k,j=1}^m \frac{r_{abkl}^{\alpha}}{r_{abkl}^{\alpha}} \phi'(|\mathbf{r}_{abkl}|) \quad (\text{A.7})$$

Appendix B

Published Paper on Protein Solubility in Organic Solvents

Communication to the Editor

On Protein Solubility in Organic Solvents

Jennifer T. Chin,¹ Sarah L. Wheeler,² and Alexander M. Klibanov^{2*}

¹Department of Chemical Engineering and ²Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received September 21, 1993/Accepted January 7, 1994

Solubility of a model protein, hen egg-white lysozyme, was investigated in a wide range of neat nonaqueous solvents and binary mixtures thereof. All solvents that are protic, very hydrophilic, and polar readily dissolve more than 10 mg/mL of lysozyme (lyophilized from aqueous solution of pH 6.0). Only a marginal correlation was found between the lysozyme solubility in a nonaqueous solvent and the latter's dielectric constant or Hildebrand solubility parameter, and no correlation was observed with the dipole moment. Lysozyme dissolved in dimethyl sulfoxide (DMSO) could be precipitated by adding protein nondissolving co-solvents, although the enzyme had a tendency to form supersaturated solutions in such mixtures. The solubility of lysozyme, both in an individual solvent (1,5-pentanediol) and in binary solvent mixtures (DMSO/acetonitrile), markedly increased when the pH of the enzyme aqueous solution prior to lyophilization was moved away from the protein's isoelectric point. © 1994 John Wiley & Sons, Inc. Key words: lysozyme • nonaqueous solvents • protein solubility • binary solvent mixtures • lyophilized proteins

INTRODUCTION

A surging interest in the biotechnological potential of enzymes suspended in neat organic solvents^{5,6,12,13,23} rekindles the attention to the issue of protein solubility in such media. It has been known since Singer's classical studies²² that although typical, hydrophilic proteins are insoluble in nearly all organic solvents, a few nonaqueous solvents, in particular, dimethyl sulfoxide (DMSO), ethylene glycol, and formamide, as well as some halogenated alcohols,¹⁰ can dissolve significant concentrations of common proteins. It is still unclear, however, what makes these protein-dissolving solvents so special, i.e., which of their physicochemical characteristics enable them to dissolve proteins. In addition to its relevance to nonaqueous enzymology,¹² this question is also important for recently proposed^{3,4} downstream protein processing in such protein-dissolving organic solvents.

In the present study, as a step toward answering the foregoing question, we have systematically and quantitatively examined protein solubility in a wide range of neat nonaqueous solvents and their binary mixtures. A counterintuitive dependence of the solubility of a protein on its charge has been observed, and a number of new

highly protein-dissolving organic solvents has been uncovered.

MATERIALS AND METHODS

Hen egg-white lysozyme (EC 3.2.1.17) was purchased from Sigma Chemical Co., and had a specific activity of 48,000 units/mg solid. The enzyme was dissolved in deionized water at 4 mg/mL and then dialyzed against 16 volumes of deionized water at 4°C for 16 h with three changes of water in the interim. The pH of the resultant enzyme solution was adjusted to a desired value (usually pH 6.0) with aqueous NaOH or HCl. The solution was subsequently lyophilized for 48 h, and the solid enzyme was stored in a desiccator at 4°C.

All nonaqueous solvents used in this work were obtained from Aldrich Chemical Co., except for glycerol which was from Mallinckrodt Specialty Chemicals Co. All of them were 99% pure or better, with the exception of 1,3-propanediol, 1,5-pentanediol, and hexanol which had the purities of 98%, 96%, and 98%, respectively. The solvents were used without further purification or drying. All other chemicals employed herein were purchased from commercial suppliers and were of the highest purity available.

The dissolved protein concentration in all organic solvents was determined using the Lowry¹⁶ assay. *Water-miscible* solvents were diluted more than ten fold in the course of the assay; in each case, it was established separately that the residual solvent did not affect the assay. For *water-immiscible* solvents, the protein was first extracted with an equal volume of 10 mM aqueous phosphate buffer, pH 7.0 (repeated extraction yielded no additional protein, indicating that all the protein was extracted the first time), followed by measuring the protein concentration in the aqueous extract.

Unless stated otherwise, lysozyme solubility in organic solvents was measured by placing the solid protein (typically 40 mg) into a 5-mL screw-cap scintillation vial, followed by addition of 2 mL of the solvent. Before closing, the vial was sealed with aluminum foil and Teflon tape. The resultant suspension was shaken at 30°C and 300 rpm for a specified period of time and then centrifuged at 30,000 g for 30 min at 30°C. The protein content of the supernatant was determined as described above.

* To whom all correspondence should be addressed.

The water content of lyophilized lysozyme was measured using the Karl Fischer titration.¹⁵

RESULTS AND DISCUSSION

As a model for our investigation, we selected hen egg-white lysozyme as it is a typical and well-studied protein.⁸ Lysozyme was dissolved in deionized water, desalted by extensive dialysis against fresh deionized water, and then lyophilized from a pH 6.0 aqueous solution. The water content of the resultant amorphous powder was found to be $8.5 \pm 0.1\%$.

First, we attempted to determine quantitatively the solubility of this lysozyme sample in three classical^{4,22} protein-dissolving nonaqueous solvents—dimethyl sulfoxide (DMSO), ethylene glycol, and formamide. The lyophilized enzyme was added to each of these solvents, followed by shaking at 30°C and centrifugation, as described in Materials and Methods. In all three cases, the protein solubility was found to be so high that it could not be measured precisely, because at 50 mg/mL the lysozyme solution became so viscous that it behaved like a transparent gel.

Because it was experimentally impossible to determine the solubility of lysozyme in the aforementioned three neat organic solvents and because such solubility values would be required for quantitative correlations, we decided to employ the following approach. Because lysozyme is insoluble in most common solvents, the binary mixtures thereof with, say, DMSO should dissolve less protein than DMSO itself.⁴ Consequently, we examined the solubility of lysozyme in various binary mixtures of DMSO with different co-solvents—*n*-octanol, *tert*-amyl alcohol, *tert*-butanol, *n*-hexanol, acetonitrile, nitrobenzene, methylene chloride, and *N,N*-dimethylformamide. (Note that these solvents are infinitely miscible with DMSO,⁷ and thus all the binary mixtures are monophasic.) The data obtained are depicted in Figure 1. It is seen that for each of the co-solvents the lysozyme solubility drops as the co-solvent concentration is raised. Furthermore, for each co-solvent the protein solubility can be varied from zero to more than 10 mg/mL simply by adjusting the ratio of the co-solvent to DMSO.

For most of the data points in Figure 1, at least a half-day shaking at 30°C was required to reach the constant concentration of lysozyme in the binary mixture. We found that a much faster route to the same DMSO/co-solvent/lysozyme system was to first dissolve the protein in pure DMSO and then add a co-solvent, thereby bringing about precipitation of the excess protein. This approach (henceforth referred to as the *precipitation method*), in which a constant concentration of the dissolved protein was attained in less than 2 h, was explored further.

We dissolved lysozyme lyophilized from pH 6.0 in neat DMSO, and then investigated the effect of the increasing fraction of primary straight-chain alcohols on this solution. For ethanol and butanol, the protein was still soluble at (the

arbitrarily chosen concentration of) 5 mg/mL even at 60% (v/v) alcohol. For higher alcohols, lysozyme precipitation from its 5 mg/mL solution in the binary mixture was observed at some 40% to 50% (v/v) co-solvents. For one such typical alcohol, heptanol, the lysozyme solubility, obtained by the precipitation method, was examined as a function of heptanol concentration.

The middle column in Table I shows the protein solubilities obtained for 40% to 90% (v/v) heptanol content in its mixtures with DMSO. Independently, the solubilities of lysozyme were also measured by *directly solubilizing the protein in the same binary solvent mixtures*. Comparison of the solubilities obtained by the two methods (the last two columns in Table I) reveals that, despite similar trends, there is a large discrepancy between them. In all instances, the precipitation method resulted in much higher solubility values: e.g., for 70% heptanol, the lysozyme solubility obtained by the precipitation method exceeded that obtained via direct dissolution by 20-fold. Similar differences between the results of the two methods, although of a lesser magnitude, were observed for hexanol and octanol, as well as for such other co-solvents as butyl acetate, acetonitrile, and *N,N*-dimethylformamide.

Which method gives the right results? We hypothesized that the precipitation method yields relatively stable supersaturated lysozyme solution, thus not providing the true solubility values. This hypothesis was verified in the following experiment. To a lysozyme solution in DMSO, 70% (v/v) heptanol was added to bring the protein concen-

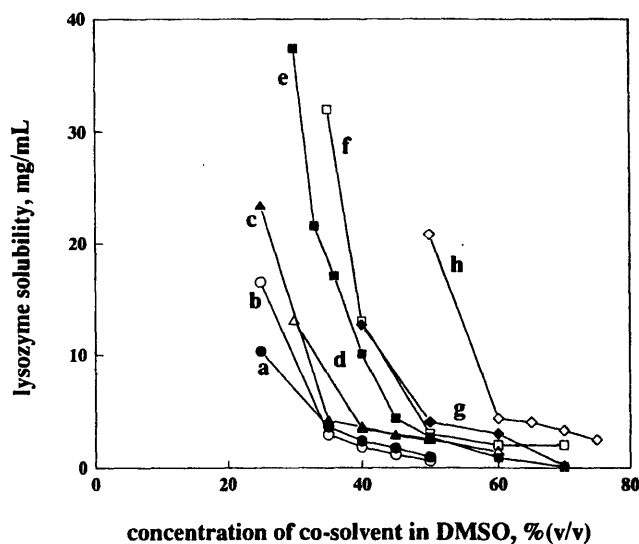


Figure 1. Solubility of lysozyme, lyophilized from aqueous solution of pH 6.0, in various binary mixtures of dimethyl sulfoxide (DMSO) with octanol (a), *tert*-amyl alcohol (b), *tert*-butanol (c), hexanol (d), acetonitrile (e), nitrobenzene (f), methylene chloride (g), and *N,N*-dimethylformamide (h). The solubility was determined by placing 50 mg/mL of the lyophilized protein powder in the appropriate solvent mixtures (all monophasic), shaking at 30°C, subsequent removal of the undissolved protein by centrifugation, and protein determination in the supernatant; for other conditions and procedures, see Materials and Methods.

Table I. Solubility of lysozyme in various mixtures of DMSO and *n*-heptanol obtained by two independent methods—precipitation from DMSO solutions by heptanol and direct dissolution in binary mixtures.^a

Concentration of heptanol, %(v/v)	Concentration of dissolved lysozyme, mg/mL	
	By precipitation	By direct dissolution
40	3.8	2.3
50	2.6	0.9
60	2.1	0.5
70	2.0	0.1
80	0.4	0.1
90	0.0	0.0

^aLysozyme lyophilized from aqueous solution of pH 6.0 was dissolved in DMSO/heptanol binary mixtures at 30°C by one of the two methods. In the precipitation method, to a protein solution in DMSO a desired concentration of heptanol (left column) was added to give 5 mg/mL lysozyme. The resultant suspension was shaken for 2 h (which was shown to be sufficient to attain the constant concentration of the dissolved protein in the supernatant), followed by centrifugation and protein assay of the supernatant. In the direct dissolution method, 5 mg/mL lysozyme was placed in a desired DMSO/heptanol binary mixture, followed by shaking and protein assay as described in Materials and Methods.

tration to 5 mg/mL. Some protein precipitated leaving a 2 mg/mL concentration in the supernatant (Table I). When the heptanol content was raised further to 90% (v/v), the lysozyme concentration in the supernatant dropped to zero (Table I). Finally, some DMSO was added to the suspension to lower the heptanol content back to 70%, and the suspension was shaken at 30°C for 12 h, followed by centrifugation and protein assay of the supernatant. The precipitation method would predict the lysozyme concentration in the supernatant of 2 mg/mL, whereas the direct dissolution method would predict 0.1 mg/mL (Table I). In fact, the latter value was obtained. This finding supports our supersaturation hypothesis and seems to invalidate the precipitation method. Therefore, the direct dissolution method was employed in all subsequent experiments.

The next step was to examine the solubility of lysozyme as a function of the nature of the solvent. To this end, the protein solubility was measured in 34 nonaqueous solvents covering a wide range of physicochemical properties. The data obtained, presented in Table II, are conducive to several insightful conclusions. As many as 14 solvents afford a very high lysozyme solubility—greater than 10 mg/mL (in fact, for 10 of them, the solubility exceeded 20 mg/mL). Almost all of these protein-dissolving solvents are protic, suggesting the importance of the solvent's propensity to form hydrogen bonds. In 16 of the solvents tested (the lower part of the table), the lysozyme concentration was below the sensitivity limit of our measurement (0.05 mg/mL), with the remaining four solvents providing the intermediate solubility values.

We attempted to correlate the protein solubility with some commonly used basic solvent characteristics. Those that are available for most of the solvents tested (Table II)

were found to fall into three distinct groups. The first group consists of the solvent properties with which lysozyme solubility showed a good correlation with very few exceptions. These include solvent hydrophobicity (as reflected by its $\log P$ value, where P is the partition coefficient of the solvent between *n*-octanol and water¹⁹) and the empirical solvent polarity parameter¹⁸ E_T^N . With respect to the former characteristic, protein-dissolving solvents were hydrophilic (negative $\log P$) and nondissolving ones hydrophobic. Notable exceptions were phenol which, despite its hydrophobicity, dissolved >10 mg/mL lysozyme and dioxane which, its hydrophobicity notwithstanding, dissolved less than 0.05 mg/mL lysozyme. Likewise, protein-dissolving solvents were quite polar ($E_T^N > 0.6$), whereas nondissolving ones much less so. Exceptions here were DMSO which is an excellent solvent for lysozyme but has $E_T^N = 0.444$ and ethanol which dissolves less than 0.05 mg/mL lysozyme but has $E_T^N = 0.654$.

Lysozyme's solubility displayed only a marginal correlation with solvent characteristics from the second group, namely, the dielectric constant (ϵ) and the Hildebrand solubility parameter (δ). Although in general protein-dissolving solvents tend to have higher ϵ and δ values, there are numerous exceptions (Table II). Finally, as seen in Table II, virtually no correlation was observed between lysozyme solubility and the dipole moment of the solvent (the third group of characteristics).

One can conclude from the results in Table II that there is no single solvent characteristic that can serve as an unmistakable predictor of the solvents's protein-dissolving ability. Nevertheless, all solvents that are protic, very hydrophilic, and polar dissolve at least 10 mg/mL of lysozyme.

Half of the strongly protein-dissolving solvents in Table II are diols or polyols. Therefore, to test further some of the foregoing trends we examined in more detail lysozyme's solubility in several homologous members of this conspicuous class of solvents. It is seen in Table III that the solubility of lysozyme declines in the series of 1,*n*-terminal diols (where *n* is the number of methylene groups) as *n* (and hence, hydrophobicity) increases. Unfortunately, again, in most cases the exact solubility value could not be determined because the solution becomes exceedingly viscous.

Comparison of our lysozyme solubility data in nonaqueous solvents with those from the literature¹⁷ reveals some major differences. For example, we could dissolve more than 50 mg/mL lysozyme in ethylene glycol and formamide, whereas Rees and Singer¹⁷ reported less than 1 mg/mL. We ascribe this discrepancy to the fact that, although we lyophilized lysozyme from the carefully controlled pH 6.0, the previous workers¹⁷ used "commercial protein samples," i.e., it was not known from what pH they had been prepared. Because proteins have a "pH-memory" in nonaqueous solvents,¹² this effect could be important.

To explore this hypothesis, we lyophilized lysozyme from aqueous solutions of two additional pH values— one closer to the protein's isoelectric point of 11,⁸ pH 10.0, and

Table II. Solubility of lysozyme in various neat organic solvents and physicochemical parameters thereof.

Solvent	Solubility (mg/mL) ^a	log <i>P</i> ^b	ε ^c	μ (D) ^d	<i>E</i> _T ^N ^e	δ (MPa ^{1/2}) ^f
Glycerol	>10	-3.0	42.5	g	0.812	34.9
Triethylene glycol	>10	-2.8	23.7	3.0	0.704	24.7
Diethylene glycol	>10	-2.3	31.7	2.3	0.713	27.4
Formamide	>10	-2.1	111	3.37	0.799	39.3
2,2,2-Trifluoroethanol	>10	-2.1	g	g	0.898	g
Ethylene glycol	>10	-1.8	37.7	2.28	0.790	32.9
1,2-Propanediol	>10	-1.4	32	2.3	0.722	28.0
1,3-Propanediol	>10	-1.3	35.0	2.5	0.747	g
DMSO	>10	-1.3	46.5	4.0	0.444	g
<i>N</i> -methylformamide	>10	-1.3	182	3.87	0.722	32.9
<i>N</i> -methylacetamide	>10	-1.1	191	4.26	0.657	29.9
1,4-Butanediol	>10	-0.8	30.2	2.58	0.704	24.7
Methanol	>10	-0.7	32.7	2.87	0.762	29.6
Phenol	>10	1.5	9.78	1.45	0.948	24.1
1,5-Pentanediol	4.8	-0.3	g	g	g	23.5
<i>N,N</i> -dimethylformamide	0.30	-1.0	36.7	3.86	0.404	24.8
Ethanol	0.24	-0.24	24.6	1.74	0.654	26.3
Pyridine	0.16	0.71	12.9	2.37	0.302	21.9
1,4-Dioxane	<0.05	-1.1	2.20	0.45	0.164	20.5
Acetonitrile	<0.05	-0.3	35.9	3.4	0.460	24.4
Ethyl acetate	<0.05	0.68	6.02	1.83	0.228	18.4
<i>tert</i> -Butanol	<0.05	0.8	12.5	1.69	0.389	21.7
Cyclohexanone	<0.05	0.96	18.3	3.01	0.281	20.0
Methylene chloride	<0.05	1.2	8.9	1.62	g	20.1
<i>tert</i> -Amyl alcohol	<0.05	1.3	5.8	1.7	g	g
<i>tert</i> -Butyl methyl ether	<0.05	1.4	4.5	1.23	0.148	g
Butyl acetate	<0.05	1.7	5.01	1.84	g	g
Hexanol	<0.05	1.8	12.1	1.65	0.559	21.9
Nitrobenzene	<0.05	1.8	34.8	4.03	0.324	21.4
Toluene	<0.05	2.5	2.38	0.30	0.099	18.2
Octanol	<0.05	2.9	9.4	1.69	0.543	21.1
Octane	<0.05	4.5	1.95	0.0	g	15.6
Propylene carbonate	<0.05	g	64.9	4.95	g	27.2
Tetramethylene sulfone	<0.05	g	43.3	4.80	0.410	27.4

^aThe solubility was measured by placing 10 mg/mL lysozyme, lyophilized from pH 6.0, in a solvent and then shaking the resultant mixture at 30°C (except at 45°C in the case of phenol because of its mp of 41° to 42°C) for 12 h (24 h for 1,5-pentanediol), followed by centrifugation and protein assay of the supernatant, as described in Materials and Methods. The solubility of >10 mg/mL means that this concentration of the enzyme resulted in a clear solution; <0.05 mg/mL (the sensitivity limit of our assay) means that no dissolved lysozyme was detected in the supernatant.

^bThe measure of solvent's hydrophobicity where *P* is the partition coefficient for the solvent between octanol and water. The values provided were taken from, or calculated on the basis of, Rekker.¹⁹

^cDielectric constant ε values were taken from the monographs by Reichardt¹⁸ and Riddick and Bunger.²⁰

^dDipole moment μ values (in debyes) were taken from the two sources mentioned in footnote c.

^eThe empirical solvent polarity parameter *E*_T^N values (normalized) were taken from Reichardt's monograph¹⁸ (pp. 365–370 and 408–410).

^fThe Hildebrand solubility parameter δ values (in megapascals) were taken from Barton's monograph.¹ When several values were listed there, their average is given in the table.

^gValues are not available.

another one farther from it, pH 2.0, than pH 6.0. The resultant samples, as well as the pH 6 lysozyme, were then dissolved in various binary mixtures of DMSO and acetonitrile presented in Figure 1. The data obtained, depicted in Figure 2A, are quite striking. There is indeed a marked dependence of the lysozyme solubility in the binary solvent mixture on the pH from which the protein was lyophilized. In particular, the farther away that pH is from the isoelectric point, the greater the protein solubility. For example, the solubility of lysozyme lyophilized from pH 2.0 in 56% acetonitrile/44% DMSO was 44 mg/mL, whereas for the pH 10.0 lysozyme the solubility in the same binary mixture was at least 1000 times lower.

The same trend was observed with the solubility of lysozyme in an individual nonaqueous solvent, 1,5-pentanediol (the only solvent from Table II which afforded a high, and yet measurable exactly, protein concentration). One can see in Figure 2B that the protein solubility in this solvent gradually declines to zero as the pH of the aqueous solution from which lysozyme has been lyophilized is increased from 4 to 11.

The data in Figure 2 clearly demonstrate that the pH of protein aqueous solution prior to lyophilization defines the subsequent solubility in nonaqueous solvents. In addition, protein solubility in such solvents, as in water, is the lowest when the macromolecule is near its isoelectric point.

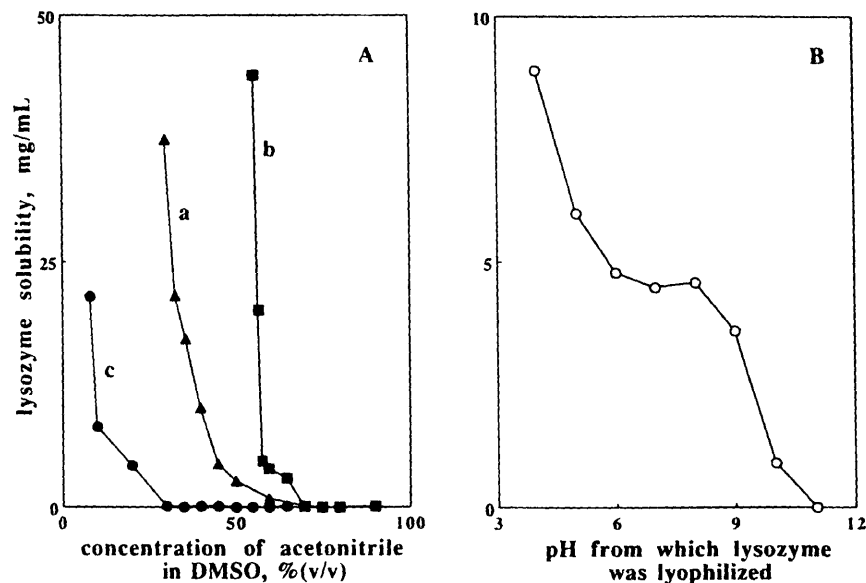


Figure 2. Dependence of the solubility of lysozyme in various DMSO/acetonitrile binary mixtures (A) and in neat 1,5-pentanediol (B) on the pH of the enzyme aqueous solution prior to lyophilization. For A, pH values were 6.0 (a), 2.0 (b), or 10.0 (c). Other conditions and procedures are the same as in Figure 1.

Table III. Solubility of lysozyme in neat homologous diols.^a

Solvent	Solubility (mg/mL)	Comments
Ethylene glycol	>50	Clear gel at 50 mg/mL
1,3-Propanediol	>50	Clear gel at 50 mg/mL
1,4-Butanediol	>40, <50	Clear gel at 40 mg/mL, turbid gel at 50 mg/mL
1,5-Pentanediol	4.8	Clear solution

^aThe solubilities of lysozyme (lyophilized from pH 6.0) were measured at 30°C as outlined in footnote a to Table II, except that 50 mg/mL protein was employed for ethylene glycol and 1,3-propanediol and both 50 and 40 mg/mL protein concentrations were employed for 1,4-butanediol.

Although this phenomenon can be readily rationalized for water, it is surprising for nonaqueous solvents, in which one would expect the energetics of charge-solvent (e.g., 1,5-pentanediol, see Fig. 2B) interactions to be unfavorable. Its explanation is largely dependent on the protein conformation in dissolving solvents. Whereas *nondissolving* organic solvents apparently have little effect on suspended lyophilized enzymes,² *protein-dissolving* solvents severely disrupt the tertiary structure.²² However, there is a considerable controversy with respect to the fate of the secondary structure. For example, it was reported²¹ that the helical content of bovine pancreatic ribonuclease dissolved in ethylene glycol is virtually unchanged compared with water, whereas subsequent studies by other authors^{9,14} concluded that proteins dissolved in DMSO and formamide were nearly devoid of secondary structure. Work is currently in progress to clarify this and related issues.

This research was supported by the NSF Biotechnology Process Engineering Center at the Massachusetts Institute of Technology.

References

- Barton, A. F. M. 1983. CRC handbook of solubility parameters and other cohesion parameters, pp. 43, 48–56, and 94–109. CRC Press, Boca Raton, FL.
- Burke, P. A., Griffin, R. G., Klibanov, A. M. 1992. Solid-state NMR assessment of enzyme active center structure under nonaqueous conditions. *J. Biol. Chem.* **267**: 20057–20064.
- Chang, N., Klibanov, A. M. 1992. Protein chromatography in neat organic solvents. *Biotechnol. Bioeng.* **39**: 575–578.
- Chang, N., Hen, S. J., Klibanov, A. M. 1991. Protein separation and purification in neat dimethyl sulfoxide. *Biochem. Biophys. Res. Comm.* **176**: 1462–1468.
- Dordick, J. S. 1989. Enzymatic catalysis in monophasic organic solvents. *Enzyme Microb. Technol.* **11**: 194–211.
- Faber, K., Riva, S. 1992. Enzyme-catalyzed irreversible acyl transfer. *Synthesis* 895–910.
- Godfrey, N. B. 1972. Solvent selection via miscibility number. *Chemtech* **2**: 359–363.
- Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., Rupley, J. A. 1972. Vertebrate lysozymes, pp. 665–868. In: P. D. Boyer (ed.), *The enzymes*, vol. 7. Academic Press, New York.
- Jackson, M., Mantsch, H. M. 1991. Beware of proteins in DMSO. *Biochim. Biophys. Acta* **1078**: 2231–2235.
- Jackson, M., Mantsch, H. H. 1992. Halogenated alcohols as solvents for proteins: FTIR spectroscopic studies. *Biochim. Biophys. Acta* **1118**: 139–143.
- Klibanov, A. M. 1986. Enzymes that work in organic solvents. *Chemtech* **16**: 354–359.
- Klibanov, A. M. 1989. Enzymatic catalysis in anhydrous organic solvents. *Trends Biochem. Sci.* **14**: 141–144.
- Klibanov, A. M. 1990. Asymmetric transformations catalyzed by enzymes in organic solvents. *Acc. Chem. Res.* **23**: 114–120.
- Klyosov, A. A., Van Viet, N., Berezin, I. V. 1975. The reactions of α -chymotrypsin and related proteins with ester substrates in nonaqueous solvents. *Eur. J. Biochem.* **59**: 3–7.
- Laitinen, H. A., Harris, W. E. 1975. *Chemical analysis*. 2nd edition. McGraw-Hill, New York.

16. Lowry, O.H., Rosenbrough, N.J., Farr, A.L., Randall, R.J. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**: 265–275.
17. Rees, E.D., Singer, S.J. 1956. A preliminary study of the properties of proteins in some nonaqueous solvents. *Arch. Biochem. Biophys.* **63**: 144–159.
18. Reichardt, C. 1988. Solvents and solvent effects in organic chemistry. 2nd edition. VCH, Weinheim.
19. Rekker, R.F. 1977. The hydrophobic fragmental constant. Elsevier, New York.
20. Riddick, J.A., Bunger, W.B. 1970. Organic solvents. Physical properties and methods of purification. 3rd edition. Wiley, New York.
21. Sage, H.J., Singer, S.J. 1962. The properties of bovine pancreatic ribonuclease in ethylene glycol solution. *Biochemistry* **1**: 305–317.
22. Singer, S.J. 1962. The properties of proteins in non-aqueous solvents. *Adv. Protein Chem.* **17**: 1–68.
23. Zaks, A., Russell, A.J. 1988. Enzymes in organic solvents: properties and applications. *J. Biotechnol.* **8**: 259–270.

Bibliography

- [1] J.S. Rowlinson and B. Widom. *Molecular Theory of Capillarity*. Clarendon Press, (1982).
- [2] C.I. Poser and I.C.J. Sanchez. *J. Colloid Interface Sci.*, **69**:539–548, (1979).
- [3] R. Evans. *Advances in Physics*, **28(2)**:143–200, (1979).
- [4] R. Evans. *J. Phys.: Condens. Matter*, **2**:8989–9007, (1990).
- [5] E. Velasco and P. Tarazona. *Phys. Rev. A*, **42**:7340–7346, (1990).
- [6] E. Velasco and P. Tarazona. *J. Chem. Phys.*, **91**:7916–7924, (1989).
- [7] T.K. Vanderlick, L.E. Scriven, and Davis H.T. *J. Chem. Phys.*, **90**:2422–2436, (1989).
- [8] D.N. Theodorou. *Macromolecules*, **21**:1391–1400, (1987).
- [9] D.N. Theodorou. *Macromolecules*, **22**:4578–4589, (1989).
- [10] D.N. Theodorou. *Macromolecules*, **21**:1400–1410, (1988).
- [11] D. Henderson, F.F. Abraham, and J.A. Barker. *Mol. Phys.*, **31**:1291–1295, (1976).
- [12] S. Sarman. *J. Chem. Phys.*, **92**:4447–4455, (1990).
- [13] J.K. Lee, J.A. Barker, and G.M. Pound. *Journal of Chemical Physics*, **60(5)**:1976–1980, (1974).

- [14] M.H. Kalos, J.K. Percus, and M. Rao. *J. Stat. Phys.*, **17**:111, (1977).
- [15] D.J. Lee, M.M. Telo da Gama, and K.E. Gubbins. *Mol. Phys.*, **53**:1113–1130, (1984).
- [16] E. Salomons and M.J. Mareschal. *J. Phys.: Condens. Matter*, **3**:3645–3661, (1991).
- [17] J.A. Holcomb, C.D. Zollweg. *J. of Physical Chemistry*, **97**:4797–4805, (1993).
- [18] J.A. Barker, R.A. Fisher, and R.O. Watts. *Molecular Physics*, **21(4)**:657–673, (1971).
- [19] M.P. D'Evelyn and S.A. Rice. *J. of Chem. Phys.*, **32**:2909, (1985).
- [20] J.G. Harris, J. Gryko, and S.A. Rice. *J. of Chem. Phys.*, **87**:3069–3081, (1987).
- [21] G.G. Chapela, G. Saville, S.M. Thompson, and J.S Rowlinson. *Chem. Soc. Faraday Trans. 2*, **73**:1133, (1977).
- [22] R.M. Townsend, J. Gryko, and S.A. Rice. *J. Chem. Phys.*, **82**:4391, (1985).
- [23] M. Matsumoto and Y. Kataoka. *Journal of Chemical Physics*, **90(4)**:2398–2407, (1989).
- [24] T.A. Weber and E. Helfand. *J. Chem. Phys.*, **72**:4014–4018, (1980).
- [25] P.J. Linse. *J. Chem. Phys.*, **86**:4177–4187, (1987).
- [26] I.L. Carpenter and W.J. Hehre. *J. Chem. Phys.*, **94**:531–536, (1990).
- [27] M. Vacatello, D.Y. Yoon, and B.C. Laskowski. *Journal of Chemical Physics*, **93(1)**:779–786, (1990).
- [28] U. Landman, W.D. Luedtke, J. Ouyang, and T.K Xia. *Japanese Journal of Applied Physics*, **32**:1444–62, (1993).
- [29] C. Morel, S. Doan, O. Rivoal, B. Ameline, C. Monin, and P. Larricart. *JOURNAL FRANCAIS D OPHTALMOLOGIE*, **25(4)**:537–578, (1961).

- [30] S. Chang, E. Ozmert, and N.J. Zimmerman. *American Journal of Ophthalmology*, **106**:668–674, (1988).
- [31] R.G. Devenyi. *Canadian Journal of Ophthalmology*, **33(7)**:385–386, (1998).
- [32] 3M Specialty Fluids. 3M Webpage:
<http://www.mmm.com/market/industrial/fluids/pfg3218.html>.
- [33] A.R. Ravishankara and S. Solomon. *Science*, **259**:194, (1993).
- [34] T. Frietsch, C. Lenz, and K.F. Waschke. *EUROPEAN JOURNAL OF ANAESTHESIOLOGY*, **15(5)**:571–584, (1998).
- [35] H.A. Sloviter. *Fed Proc (EUV)*, **34 (6)**:1484–1487, (1975).
- [36] J.G. Riess and M. Le Blanc. *Angew Chem Int Ed Engl*, **17(9)**:621–634, (1978).
- [37] R.W. Colman, L.K. Chang, B. Mukherji, and H.A. Sloviter. *Lab Clin Med (IVR)*, **95(4)**:553–562, (1980).
- [38] S.M. Panchenko and N.I. Afonin. *Gematol Transfuziol*, **33(5)**:29–31, (1988).
- [39] J.G. Harris. *J. of Phys. Chem.*, **96**:5077–5086, (1992).
- [40] A. Hariharan and J.G. Harris. *J. Chem. Phys.*, **101**, (1994).
- [41] B.J. Torby. *Advanced Dynamics For Engineers*. Holt, Rinehart and Winston, (1984).
- [42] D.C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 1995.
- [43] J.M. Haile. *Molecular Dynamics Simulation*. John Wiley and Sons, Inc., (1997).
- [44] L. Verlet. *Phys. Rev.*, **159**:98–103, (1967).
- [45] D. Findcham and D.M. Heyes. *CCP5 Quarterly*, **6**:4–10, (1982).
- [46] D.M. Heyes and K. Singer. *CCP5 Quarterly*, **6**:11–23, (1982).

- [47] W.C. Swope, H.C. Anderson, P.H. Berens, and K.R. Wilson. *J. Chem. Phys.*, **76**:637–649, (1982).
- [48] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, (1987).
- [49] J.P. Ryckaert, G. Ciccotti, and H.J.C. Berendsen. *J. Comput. Phys.*, **23**:327–341, (1977).
- [50] G. Ciccotti and J.P. Ryckaert. *Comput. Phys. Rep.*, **4**:345–392, (1986).
- [51] H.C. Andersen. *Journal of Computational Physics*, **52**:24–34, (1983).
- [52] M. Born and Von Karman Th. *Physik. Z.*, **13**:297–309, (1912).
- [53] H.C. Andersen, M.P. Allen, A. Bellemans, J. Board, J.H.R. Clarke, M. Ferrario, J.M. Haile, S. Nose, J.V. Opheusden, and J.P. Ryckaert. *Rapport d'activite scientifique du CECAM*, pages 82–115, (1984).
- [54] H.C. Andersen. *J. Chem. Phys.*, **72**:2384–2393, (1980).
- [55] S. Nose. *Mol. Phys.*, **52**:255–268, (1984).
- [56] H.J.C. Berendsen and W.F. Van Gunsteren. *Molecular liquids, dynamic and interaction, NATO ASI series*. Reidel, New York, (1984).
- [57] H.J.C. Berendsen. *Rapport d'activite scientifique du CECAM*, pages 166–175, (1984).
- [58] Biosym Technologies. *Discover User Guide, Version 2.9/3.1*. (1993).
- [59] I.N. Levine. *Quantum Chemistry*. Prentice Hall, 4 edition, 1991.
- [60] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Jr. Profeta, and P. Weiner. *Journal of American Chemical Society*, **106**:765–784, (1984).

- [61] A. Hagler, E. Euler, and S. Lifson. *Journal of American Chemical Society*, **96**:5319, (1974).
- [62] R.E. Banks. *Fluorocarbons and Their Derivatives*. Macdonald Technical & Scientific, London, (1970).
- [63] R.N. Haszeldine and F. Smith. *J. Chem. Soc.*, pages 603–608, (1951).
- [64] G. Egloff. *Physical Constants of Hydrocarbons, Vol. 1*. Reinhold Publishing Corp., (1939).
- [65] *Selected Values of Physical and Thermodynamic Properties of Hydrocarbons and Related Compounds, American Petroleum Institute Research Project 44*. Carnegie Press, Pittsburgh, (1953).
- [66] J.H. Simons. *Fluororine Chemistry, Vol. V*. Academic Press, (1964).
- [67] *Handbook of Chemistry and Physics*. CRC Press, 66th Edition, (1986).
- [68] C.W. Bunn and E.R. Howells. *Nature*, **174**:549–551, (1954).
- [69] J.J. Jasper. *Journal of Physical and Chemical Reference Data*, **1**:841–1007, (1972).
- [70] G.T. Dee and B.B. Sauer. *Journal of Colloid and Interface Science*, **152(1)**:85–103, (1992).
- [71] B.B. Sauer and G.T. Dee. *Journal of Colloid and Interface Science, in press*, (1994).
- [72] J. Miyazaki, J.A. Barker, and G.M. Pound. *Journal of Chemical Physics*, **64(8)**:3364–3369, (1976).
- [73] J.A. Barker. *Molecular Physics.*, **80**:815–820, (1993).
- [74] B. Borstnik, D. Janezic, and A. Azman. *Acta Physica Academiae Scientiarum Hungaricae*, **4**:297, (1980).

- [75] O. Matsuoka, E. Clementi, and M. Yoshimine. *Journal of Chemical Physics*, **64**:1351, (1974).
- [76] G.C. Lie, E. Clementi, and M. Yoshimine. *Journal of Chemical Physics*, **64**:2195, (1975).
- [77] F.H. Stillinger and A. Rahman. *Journal of Chemical Physics*, **60**:1545, (1974).
- [78] C.Y. Lee and H.L. Scott. *Journal of Chemical Physics*, **73**:5351, (1980).
- [79] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Imprey, and M.L. Klein. *Journal of Chemical Physics*, **97**:926, (1983).
- [80] M.A. Wilson, A. Phohorille, and L.R. Pratt. *Journal of Physical Chemistry*, **91**:4873, (1987).
- [81] V. Carravetta and E. Clementi. *Journal of Chemical Physics*, **81**:2464, (1984).
- [82] M. Matsumoto and Y. Kataoka. *Journal of Chemical Physics*, **88**(5):3233–3245, (1988).
- [83] L.M. Holtzman, J.B. Adams, S.M. Foiles, and W.N.G. Hitchon. *Journal of Materials Research*, **6**(2):298–302, (1991).
- [84] M.S. Daw and M.I. Baskes. *Physical Review B*, **29**(12):6443–6453, (1984).
- [85] S. Shin, N. Collazo, and S.A. Rice. *Journal of Chemical Physics*, **98**(4):3469–3474, (1993).
- [86] S. Lifson and A. Warshel. *Journal of Chemical Physics*, **49**(11):5116–5129, (1968).
- [87] S.J. Weiner, P.A. Kollman, D.T. Nguyen, and D.A. Case. *Journal of Computational Chemistry*, **7**(2):230–252, (1986).
- [88] N.L. Allinger, Y.H. Yuh, and J.-H. Li. *Journal of the American Chemical Society*, **111**(23):8551–8565, (1989).

- [89] N.L. Allinger, Z.-Q. Zhu, and K. Chen. *Journal of the American Chemical Society*, **114**:6120–6133, (1992).
- [90] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus. *Journal of Computational Chemistry*, **4(2)**:187–217, (1983).
- [91] W.L. Jorgensen, Jeffery D. Madura, and Carol J. Swenson. *Journal of the American Chemical Society*, **106(22)**:6638–6646, (1984).
- [92] W.L. Jorgensen and Carol J. Swenson. *Journal of the American Chemical Society*, **107**:569–578, (1985).
- [93] J.M. Briggs, T.B. Nguyen, and W.L. Jorgenson. *Journal of Computational Chemistry*, **11(8)**:958–971, (1990).
- [94] J.M. Briggs, T.B. Nguyen, and W.L. Jorgenson. *Journal of Physical Chemistry*, **95**:3315–3322, (1991).
- [95] S. Toxvaerd. *Journal of Chemical Physics*, **93(6)**:4290–4295, (1990).
- [96] P. Padilla and S. Toxvaerd. *Journal of Chemical Physics*, **95(1)**:509–519, (1991).
- [97] J.R. Maple, M.-J. Hwang, T.P. Stochfisch, U. Dinur, C.S. Waldman, M. Ewig, and A.T. Hagler. *Journal of Computational Chemistry*, **15(2)**:162–182, (1994).
- [98] S. Shin, N. Collazo, and S.A. Rice. *Journal of Chemical Physics*, **96(2)**:1352–1366, (1992).
- [99] S.W. Barton, A. Goudot, O. Boulassa, F. Rondelez, B. Lin, F. Novak, A. Acero, and S.A. Rice. *Journal of Chemical Physics.*, **96(2)**:1343–1351, (1992).
- [100] J.I. Siepmann, S. Karaborni, B. Smit, and M.L. Klein. *AIChE Spring Meeting*, page paper 94c, (1994).
- [101] D.B. Holt, B.L. Farmer, K.S. Macturk, and R.K. Eby. *Polymer*, **37**:1847–1855, (1996).

- [102] D. Dixon. *J. Phys. Chem.*, **96**:3698, (1992).
- [103] G.D. Smith, R.L. Jaffe, and D.Y. Yoon. *Macromolecules*, **27**:3166–3173, (1994).
- [104] L.G. Dunfield, A. Burgess, and H.A. Scheraga. *J. Chem. Phys.*, **82**:2609, (1978).
- [105] W.L. Jorgensen. *Journal of the American Chemical Society*, **103**:335, (1981).
- [106] K.F. Mansfield and D.N. Theodorou. *Macromolecules*, **23**:4430–4445, (1990).
- [107] K.F. Mansfield and D.N. Theodorou. *Macromolecules*, **24**:6283–6294, (1991).
- [108] B.M. Axilrod and Teller. *J. Chem. Phys.*, **11**:299–300, (1943).
- [109] G.C. Maitland, M. Rigby, E.B. Smith, and W.A. Wakeham. *Intermolecular Forces, Their Origin and Determination*. Clarendon Press, (1981).
- [110] M. Rigby, E.B. Smith, W.A. Wakeham, and G.C. Maitland. *The Forces Between Molecules*. Clarendon Press, (1986).
- [111] S. Toxvaerd. *J. Chem. Phys.*, **55**(7):3116–3120, (1971).
- [112] J.A. Barker and D. Henderson. *J. Chem. Phys.*, **47**:4714, (1967).