Liquid-Vapor and Liquid-Liquid Interfaces of Small Molecule Esters and Ester Oligomers: A Molecular Dynamics Simulations Study

by

Sandeep A. Patel

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1999

Author .....
Department of Chemical Engineering
March 25, 1999

Certified by .....
Jonathan G. Harris
Visiting Professor
Thesis Supervisor

Certified by .....
Linda Griffith
Associate Professor
Thesis Supervisor

Accepted by .....
Robert E. Cohen
St. Laurent Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Liquid-Vapor and Liquid-Liquid Interfaces of Small Molecule Esters and Ester Oligomers: A Molecular Dynamics Simulations Study

by

Sandeep A. Patel

Submitted to the Department of Chemical Engineering
on May 25, 1999, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The interfaces between polymeric biomaterials and physiological environments mediate numerous phenomena such as protein adsorption, cell adhesion/clotting, foreign body response, and effects of the medium on the material such as degradation. Furthermore, it is the molecular-level structure of the material on which the responses are predicated. For this reason, a molecular-level description of the structure of polymeric materials and the response of this structure to external contacting media is relevant in the context of predicting the material performance.

One subclass of polymeric biomaterials which has exhibited great potential has been that of bioresorbable polymers of which the polyesters of alpha-hydroxy acids, polylactic (PLA) and polyglycolic (PGA) have found a secure niche as suture materials, matrices for drug delivery systems (in conjuction with derivatives), and as matrices for cell transplantation/tissue engineering applications. The present work undertakes a study of the liquid-vapor and liquid-liquid interfaces of these materials, the latter involving contact with water, via molecular dynamics simulations. Towards understanding the interfaces of these polymers, and in particular, the interfacial behavior of the carbonyl functionality defining this class of materials, we first investigate the liquid-vapor interface, or *free surface*, of small-molecule esters, namely methyl acetate. Liquid free surfaces have been studied extensively via theory, simulation, and experiment; they demonstrate unique properties, and, in this respect, are interesting 'states' of matter in and of themselves.

The molecular dynamics method is applied to study the liquid-vapor interface of methyl acetate. The OPLS (Optimized Potentials for Liquids Simulation) interaction model is applied. The model is augmented with bond angle bending and out-of-plane angle bending potentials parameterized as part of this work. It is one of a few forcefields capturing the bulk liquid energetics (enthalpy of vaporization to the ideal gas state) and intramolecular physics (conformational energetics). The system establishes a stable interface. The bulk liquid is structureless as observed in earlier

computations. The interface is monotonic, with a 10-90 thickness of 0.6 nanometers. The carbonyl methyl group orients towards the vapor phase in order to minimize energetic losses. This orientation is confirmed by longitudinal profiles of the order parameters P_1 and P_2. These indicate that the carbonyl-carbon/carbonyl-oxygen bond vector aligns antiparallel to the interface normal in the interfacial region with the oxygen adopting an orientation away from the vapor phase. For methyl acetate, a liquid not known to associate in the bulk, the interface induces a strong orientational order; effectively, the ester behaves almost as a 'surfactant'. Atomic and molecular surface tensions of 22.70 and 21.06 dyne/cm, respectively, are obtained; these values are within 15 percent of the experimental measurements (the atomic virial calculation is within 10 percent). This inadequacy of the parameter set is not too surprising as interfacial thermodynamic properties for liquids have traditionally proved to be intractable via molecular simulations; nevertheless, this result gives a base parameter set to which refinements can be made.

We have performed molecular dynamics simulations of the methyl acetate/water system using the OPLS potential model for the ester and the SPC model for water. Starting from a two-phase configuration, the system shows no *mixing* of the two species (the acetate and water are not present uniformly throughout the simulation cell); we thus report that the interaction is incapable of reproducing the experimental solubility at 298 K. An interesting result of this simulation, nevertheless, is the reversal of the orientation of methyl acetate molecules in the vicinity of the water, the 'interfacial' region, relative to the free surface case. This reorientation suggests of some association between ester and water, and this is corroborated by analysis for hydrogen bonding interactions between the species. Radial distribution functions for the carbonyl oxygen - water hydrogen show a first peak around 1.9 angstroms; the angle made by the triplet of the carbonyl oxygen, water hydrogen, and water oxygen averages to 160 degrees. This geometry is remarkably close to the geometrical results of density functional calculations on methyl acetate - water bimolecular complexes which yield hydrogen-bond distances of 1.938 and 1.916 angstroms and corresponding angles of 170.8 and 164.5 degrees.

Prior to investigating the liquid-vapor interface of an oligomer model of polylactic acid (PLA), the need to fit an internal torsion potential arose as a united-atom parameter set for this degree of freedom was not available in the literature. In fact it is quite surprising that a full model describing the internals of such a simple system as studied here, a simple linear aliphatic polyester, is not found in the literature. With this being the case, the following approach is implemented. An *ab initio* torsion angle profile is generated for a model molecule possessing the pertinent internal; several levels of Hartree-Fock theory including corrections for electron correlation are applied. Next, the CFF91 all-atom potential (which itself did not have the parameterization for the needd torsion) was used to generate the same torsion profile; in this case, the contribution of the required torsion is missing. Thus, the CFF91 profile is subtracted from the total *ab initio* profile, leaving the residual torsion profile representing the missing component. A standard torsion potential model was fitted to this profile. This potential was then added to the CFF91 model, which was used to generate distributions of all internals. United-atom potentials were then fitted to these

distributions, giving the final set of internals which could be used in the subsequent oligomer simulations.

We next begin a study of the liquid-vapor interface of the oligomer model of PLA. Center-of-mass and atom density profiles show an interfacial structure distinct from that of the bulk. We observe a segregation of chain end and middle segments as reported in earlier MD simulations of chain molecules and predicted by mean-field theories. With respect to the nature of the orientations of the carbonyl functionality, two distinct behaviors arise. The end carbonyl vectors tend to orient with the oxygen atoms toward the vapor, with the internal groups orienting essentially randomly. Although the middle group orientation is consistent with the trends observed for the methyl acetate free surface, the end carbonyl groups tend to orient in reverse of what the earlier simulations indicate. Although this appears counter-intuitive, energetic arguments suggest that this molecular conformation coupled with the individual bond vector orientations is in response to the drive for free energy minimization. The energetic loss of having two carbonyls in a high energy state is recovered with the screening of the middle carbonyl groups away from the vapor phase, more in proximity to the bulk. Again, chain connectivity now plays a big role in dictating the molecular structure relative to the case of the free surface of the small molecule, methyl acetate. From a biomaterials perspective, this result indicates that the free surface of the PLA oligomer model is polar and hydrophilic, thus able to facilitate adsorption of components of the physiological millieu such as water and importantly, proteins (which in a global sense can be viewed as large micellar entities).

Thesis Supervisor: Jonathan G. Harris
Title: Visiting Professor

Thesis Supervisor: Linda Griffith
Title: Associate Professor

Acknowledgments

As I sit here trying to write this section of my thesis, I realize what a smart thing it was to decide to write it last, for it is most pleasureable to think of the many great people whom I've had the greatest luck and privilege to meet, work, and associate with during my stay at M.I.T.

First, I would like to express my deepest gratitude to Dr. Jonathan G. Harris for sparking my interest in statistical mechanics and its application via computer simulations. In his *inimitable* style, he opened my eyes to matters ranging from the highly abstract world of electrons to the pragmatic issues associated with the world of high-stakes finance. A true academic in every sense of the word. I would like to extend a special thanks to Professor Gregory Rutledge for his interest in this work and for all his assistance, particularly in allowing me to participate with his group meetings and in providing valuable feedback concerning the content of this thesis. Above all, I acknowledge his always-encouraging words, *constructive* criticism (i.e. after committee meetings), and willingness to share his knowledge, insight, and, something which he does not have too much of, *time.* I extend my gratitude to Professor Linda Griffith, my co-advisor for her help and support during this thesis, and allowing me to attempt to study theoretically the systems which she attends to in *reality.* I also thank the members of my committee, Professors Sidney Yip, Sow-Hsin Chen, and Edward Merrill for their time and input, and support in this work. Last but not least, *many, many, many* thanks to Elaine Aufiero and Janet Fischer for their much-needed help during the final stages in preparation of this thesis (and their timely advice on all matters).

Although there have been many whose paths I've crossed along this journey over the last seven years, I must point out several individuals who really made my life at M.I.T. a true *learning* experience . Through the whole ordeal were two guys who have hung with me and to whom I owe a lot— Howard Covert and Talid Sinno. Thanks for always being there (during the moves from apartment to apartment), for listening and advising (on matters related and *unrelated* to science), for the cold, wet, and cold and wet runs (Talid–sorry I never repaid my debt on the bike), for the late night chats (Howard, go home now, normal people are just waking up), and for the great times we had enjoying the Boston nightlife. I am fortunate to have had you two to enrich my life. I can't forget the rest of 'the boyz', John Tolsma and Shahin Ali—thanks for all the good times (and there were probably too many for a normal M.I.T. graduate student). Shahin, it was always interesting to be around you when there was *work* to be done. John, thanks for enjoying the dance floor as much as I did.

When it comes to the daily grind, I must thank my officemates and colleagues Jennifer Chin and Matthew Reagan, for putting up with my *many* eccentricities (running clothes drying all over the office, biological experiments on my desk and elsewhere, *my cluttered desk*!!) I do appreciate all that you two tolerated–that in itself is probably worth some advanced degree! But most of all, thanks for all the information and ideas you exchanged with me on a daily basis. I also must acknowledge Jason Cline for his mastery of our computational facility (as well as many other neat things!) and

its constant reliable performance. The latter part of this work would not have been had it not been for his expertise and patience.

Thanks to all my fellow MIT colleagues–Mary Jane (Moj) O'Rourke (and Tim), Nina Shapley, Shawn Mowry, Fred Colhoun, Dr. Leo Lue (long-live rustcrumb—at least in our minds and hearts if nowhere else), and all the many others who have been a part of my daily life at MIT.

Contrary to popular belief, one does have a life outside of M.I.T., and it is in this domain that I wish to acknowledge several individuals. I must extend my warmest gratitude to Dr. Arvind Hariharan for just giving a damn, for keeping in touch despite the distance separating us, and for the always-stimulating discussions on technical and not-so-technical topics.

I also would like to thank the members of the Greater Boston Track Club for their comaraderie and friendship over the past two years. Thanks to Dr. Doug Burdi for introducing me to the 'track' and always having an encouraging word or two during the initiation. Many thanks to Alex Caracuzzo and Bill Turner for their congeniality and down-to-earth style. Many thanks to Dr. Bruce Bond for sharing his wisdom based on over twenty years of running. And finally, I cannot forget our beloved coach Tom Derderian, a running legend, for his humourous approach to life and coaching, and his sadistic workouts. Thanks Tom for helping me become a *somewhat faster* runner than I was when I joined!

While on the topic of running, I must mention Dr. Henning Richter, with whom I forged a friendship through running. Thanks, Henning, for our many 'heartbreaks' on Heartbreak Hill, for our many-times sadistic hill workouts in Beacon Hill or on Summit Avenue because 'we must'. Thanks for sharing your wisdom on just about everything from running to the stock market to science to human nature. You truly do justice to the name of an academic. I am indebted to you for always lending an ear to my thoughts, complaints (especially during the latter stages of my graduate student life), and ideas. And above all, thanks for sharing your friendship.

Before I close, I must mention one person who has unconditionally supported me during probably the worst time of my graduate student life–the writing phase. My most sincere gratitude and love I extend to Ratna Desai, for putting up with my violent mood swings, my attacks of depression, my spending late nights writing (even weekends) ; thank you–it's as simple as that.

Finally I would like to acknowledge the support of my family and long-time friends during this part of my life. It was always great to hear from you and live vicariously through you at times.

In closing, I would like to thank my Mother for instilling in me an intense work ethic as well as passing on her *intelligence* gene (of which I have made very little use in my life—sorry Mom). I cannot express enough gratitude to my Mother for her sacrifices which, I know without a doubt, have afforded me the luxury to pursue this research, and just about all that I have done in my life up to this point. Had it been another time, another place, another set of circumstances, I know that today there would exist a document of this type (although of orders of magnitude higher quality) with her as author. And so, I dedicate this thesis to my Mother.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Motivation

1.1 Introduction

Numerous biomaterials, such as polymers, metals, ceramics, reconstituted or specially treated natural tissues, and composites of these, are used clinically as elements of implants or devices for diagnosis and therapy. Synthetic polymers constitute the most diverse class of biomaterials, mainly because they provide a wide variety of compositions, properties, and forms as well as the ability to be fabricated into numerous geometries [1]. The applications of polymeric biomaterials as a class range over an extensive gamut. Table 1.1 lists some of the major applications of polymeric biomaterials within the last half century.

One subclass of polymeric biomaterials which has exhibited great potential has been that of bioresorbable polymers. Bioresorbable materials are known for their ap-

Table 1.1: Selected Biomedical Applications of Polymeric Materials

Medical Applications of Polymeric Materials
Vascular Grafts — Arterial Repair
Internal Supports — Bone Replacement, Joints, Pins, Screws
Components of Artificial Hearts — Valves
Drug Delivery Systems
Cell Transplantation — Scaffolds, Matrices
Tissue Engineering — Scaffolds, Matrices

plication as suture materials (DEXON, Medifit, VICRYL, a copolymer of glycolic and lactic acid), internal fixation devices for bone fracture repair (pins, screws, staples, high strength orthopedic implants), reconstructive substitutes, intraluminal grafts (devices implanted into coronary arteries following procedures such as balloon angioplasty in order to prevent the collapse and reblocking of these blood vessels), temporary vascular grafts, artificial skin, sheets for preventing adhesion, matrices for drug delivery systems, and as matrices for cell support in tissue engineering applications [2]. Among the bioresorbable polymers, the most successful have been polyesters of alpha-hydroxy acids, polylactic acid (PLA) and polyglycolic acid (PGA). These polymers enjoy a long clinical history, elicit minimal tissue response, and through a variety of compositions as copolymers, provide a wide range of properties and degradation behavior (the degradation behavior referring to the kinetics of degradation which can be customized for a particular application by variation of composition or crystallinity, for example) [3].

Structurally, PLA and PGA are represented as [4]

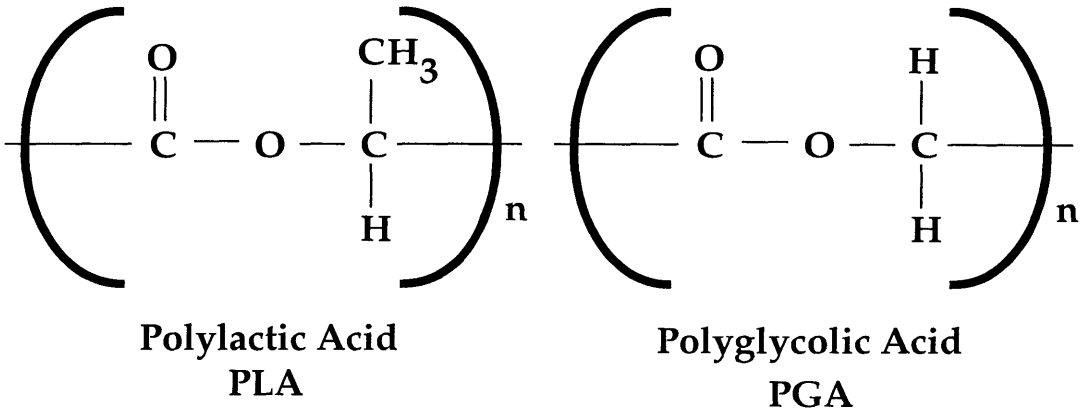


Figure 1-1: PLA and PGA Repeat Units

The ends may be capped with hydroxyl groups. Lactic acid is a chiral compound, and thus the homopolymer can exist in two enantiomeric forms–the L and D forms. The two enantiomers have equivalent intrinsic chemical properties but different structural and optical properties [5]. The synthesis of PGA and PLA is possible

through simple polycondensation reactions of the parent glycolic and lactic acids in the presence of Sb2o3; the product polymer, however, has low molecular weight and sub-optimal properties [4]. High molecular weight polymers are prepared via cationic ring-opening polymerization of the parent lactide or glycolide (the cyclic dimeric esters) in the presence of a long-chain alcohol, such as lauryl alcohol, using a metal catalyst such as antimony, zinc, lead, or preferably tin as shown in Figure 1-2 [4, 6]:

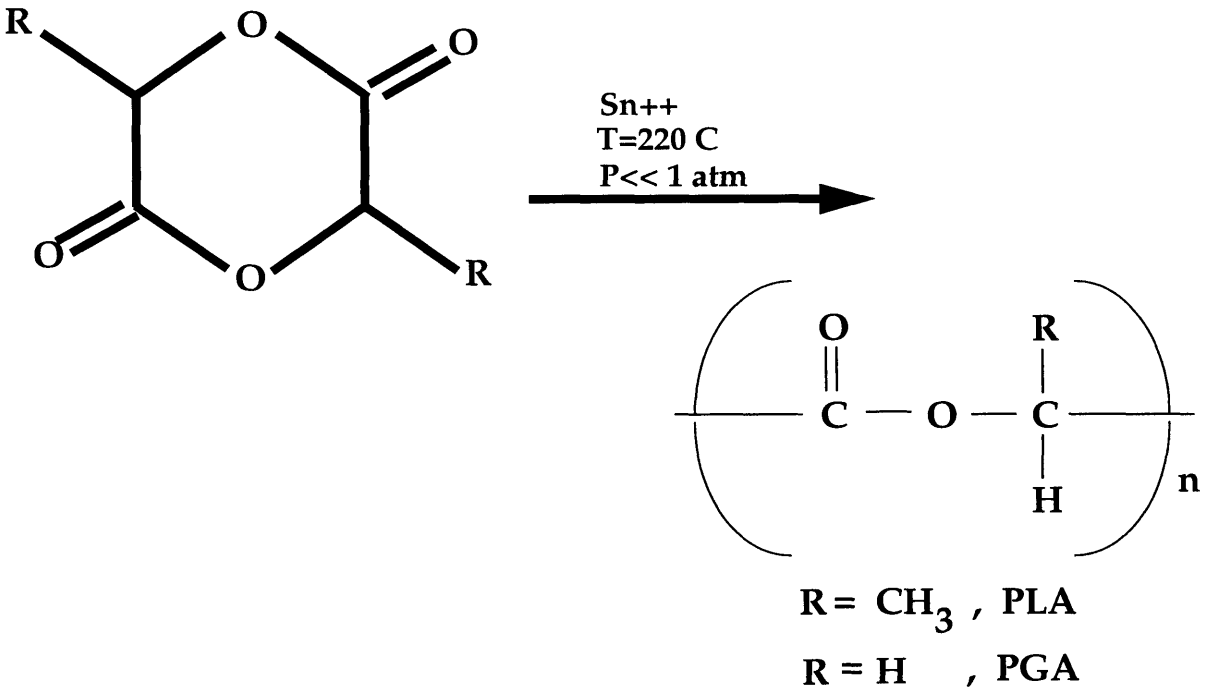


Figure 1-2: Polymerization of Glycolid/Lactide to the Polyester

Following the reaction, residual monomers are removed via precipitation of the polymer from methylene chloride solution with methanol [6]. This of course introduces the need for rigorous separation and purification procedures as the final product will be introduced into a physiological environment.

Tables 1.2 and 1.3 show the physico-chemical and mechanical properties of PGA (MW = 50000) and PLA of varying stereoregularity and molecular weight [2]. Both PGA and the L stereoregular PLA are semi-crystalline, fiber- and film-forming polymers. Due to its relatively high crystallinity (ranging from 46-52Copolymers of PLA

and PGA can be synthesized from appropriate mixtures of the parent glycolides. Morphologically, the copolymers are less crystalline than the homopolymers, and for the GA/LA ranges between 25-70 mole percent of GA monomer, the copolymers are amorphous. If the D,L-LA stereoisomer is used, the amorphous region is extended to 0-70 mole percent GA as shown in Figure 1 [7]. Note that all compositions can be make amorphous through melting and rapid quenching.

Table 1.2: PGA, PLA: Physical and Thermal Properties

Polymer	Molecular Weight	T_g (Celsius)	T_{melt} (Celsius)	T_{decomp} (Celsius)	$H_{formation}$ (J/g)	Crystallinity (Percent)
PGA	50000	35	210	254	71	52
L-PLA	50000	54	170	242	41	30
L-PLA	100000	58	159	235	20	15
L-PLA	300000	59	178	255	39	29
D,L-PLA	21000	50	—	255	amorphous	—
D,L-PLA	107000	51	—	254	amorphous	—
D,L-PLA	550000	53	—	255	amorphous	—

Table 1.3: PGA, PLA: Mechanical Properties

Polymer	Molecular Weight	Tensile Strength (MPa)	Tensile Modulus (MPa)	Flexural Modulus (MPa)	Elongation Yield (Percent)	Elongation Break (Percent)
PGA	50000	—	—	—	—	—
L-PLA	50000	28	1200	1400	3.7	6
L-PLA	100000	50	2700	3000	2.6	3.3
L-PLA	300000	48	3000	3250	1.8	2
D,L-PLA	21000	—	—	—	—	—
D,L-PLA	107000	29	1900	1950	4	6
D,L-PLA	550000	35	2400	2350	3.5	5

The success of the polyesters PLA and PGA as bioresorbable polymers is due to the fact that the degradation products are normal body metabolites which can be eliminated through the Krebs Cycle [3]. This obviates the need for invasive procedures to remove foreign materials introduced into the body, for example. The degradation mechanism of aliphatic polyesters has been studied by many researchers, and at present, the degradation is regarded as depending on chemical hydrolysis of the ester linkage (Figure 1-3); nevertheless, in vivo the contribution of enzymatic

action has been considered, and the question of the effects of enzymes on degradation is not fully answered [5]. Furthermore, it is maintained that the hydrolysis occurs preferentially in the amorphous regions of the polymers into which water is able to penetrate more rapidly as compared to the crystalline regions. Following degradation of the amorphous regions, the hydrolysis proceeds to the crystalline areas [5]. Various groups have reported studies on the kinetics, or more specifically, time dependence of bulk degradation and the factors which ap pear to influence this rate. The major factors affecting degradation rates of the homopolymers and copolymers are polymer morphology (crystalline versus amorphous), copolymer composition, hydrophillic/hydrophobic characteristics, and molecular weight [6].

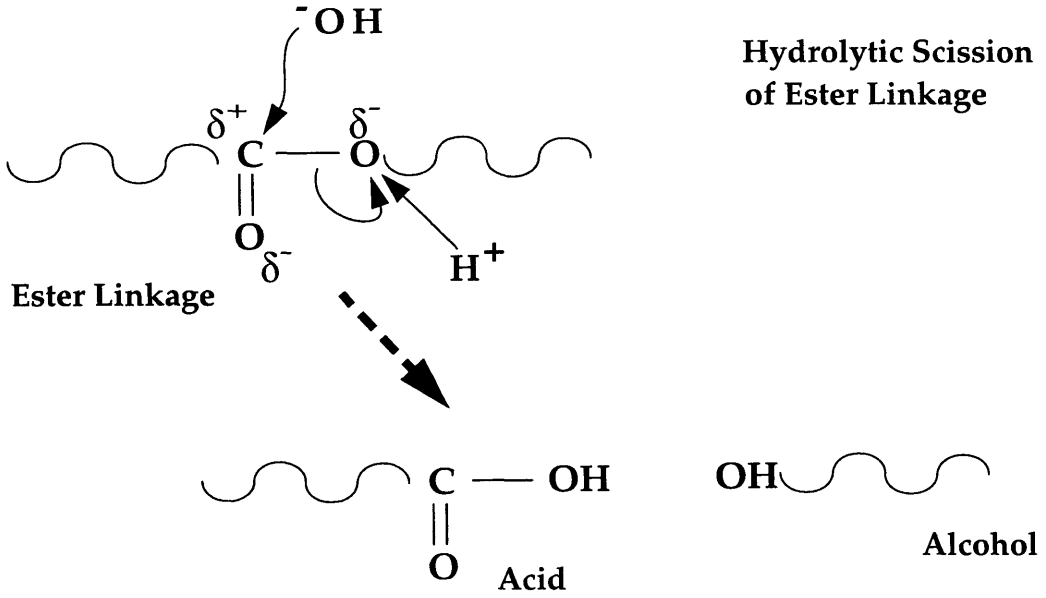


Figure 1-3: Hydrolytic Degradation of Ester Linkage: Nucleophilic Attack by Hydroxyl Group

Both bulk and surface structure play key roles in the properties and behavior of the final device in physiological environments. Factors such as polymer morphology, homogeneity, composition, and molecular structure are invariably interlinked and significantly influence polymer characteristics and performance. Moreover, introducing a polymeric material into a physiological environment creates a dynamic interface at which important phenomena appear [8]. In the presence of a physiological millieu, the interface will accommodate adsorption of native proteins followed possibly by

cell adhesion, organization, and proliferation. The philosophy behind tissue engineering/cell transplantation is based on these fundamental mechanisms. The nature of the interface is in large part influenced by the surface characteristics of the polymer. Surface mobility (ability of chains to reorient/reconfigure upon change in environment), surface tension, surface free energy, the hydrophillic or hydrophobic nature of the polymer surface, and surface structure play an important role in defining the in vivo performance of the biopolymer. It is without question that a knowledge of the structure and thermodynamic factors influencing the bulk and surface is of paramount importance to understand ing the properties and performance of these polymers; furthermore, from a design and development perspective, knowledge of the dependence of performance on structure would facilitate synthesis of novel polymers with optimum performance properties.

Computer simulations of polymeric materials have become helpful tools for investigation of polymer macroscopic structure and properties via application of statistical mechanical methods [9, 10, 11, 12]. With information on the fundamental interactions of a many particle system, computer simulations provide exact solutions to the problem of the statistical physics of many-body systems. They are a means of validating theories and models of matter, the former via comparison of theoretically obtained results to simulation results, and the latter through comparison of simulation results to experimental data [13].

1.2 Background and Rationale

1.2.1 Biopolymers: General Considerations on Biocompatibility

Researchers are considering the use of synthetic biodegradable polymers as scaffolding, or matrix support structures, for cell implantation. Viable cells would be loaded onto a polymer matrix which would then be surgically implanted into the body. Growth of vasculature into the implant and concomitant degradation of the polymer matrix

would follow. In a successful attempt, the implanted cells will continue to proliferate and differentiate into a properly functioning substitute of the original tissue [14]. The behavior of cells on the surface of the polymer is crucial. Cells must adhere to the surface (or in the looses sense be anchored to the matrix surface) and simultaneously maintain proper geometry and tissue-specific functions. Several studies have indicated that normal cell geometry is influential to cell growth and function. Variations in cell shape, it is contended, translate to alterations in cytoskeletal architecture which in turn can affect gene expression [15]. One author has shown that differences in cell shape, induced as a results of varying the density of extracellular matrix proteins on the cell attachment substrate, led to changes in cell function. Surface structure of biomaterials, for cell-transplantation technology, appears to be of prime importance [15].

Once a foreign material is placed in a physiological environment, a major concern is that of biocompatibility, that the implant not evoke any unwanted reactions and that the physiological environment not damage the implant. Reactions to foreign materials generally involve aggregation of multifarious cells on the material surface. A layer of adsorbed proteins such as fibrinogen, fibronectin, and von Willebrand factor mediates the adhesion of cells to the surface [16]. Protein adsorption is widely held as the first readily observable event at the interface between a foreign material and physiological fluid [17].

Protein adsorption is an inevitable event for systems in which protein containing fluids are in contact with surfaces. The composition of this layer of surface-adsorbed proteins depends upon the material's surface properties, the nature of the proteins, their concentration, and their relative affinities for the surface [8]. Also, as there exists a complex mixture of proteins in blood plasma and intercellular fluids, adsorption is a competitive kinetic process. Knowledge of parameters such as adsorption coefficients for different proteins would prove useful in design purposes. Researchers have elucidated several factors which influence protein adsorptino. These include electrostatic interactions, isoelectric point, pH, surface charge, coadsorption of low molecular weight species (ions and molecules), solute-solvent interactions, strength of

bonding to surface functional groups, the chemistry, topology, and morphology of the surface, and the presence of other proteins in the ambient fluid with which a given protein my interact, both away from the surface and once adsorbed on to the surface [8]. The attachment of cells, such as bacterial cells, to surfaces of biomaterials is also important because of the possible means of infection it provides [18]. Variables such as critical surface tension, surface free energy, charge, and hydrophobicity play important roles in this phenomenon.

1.2.2 Characteristics of Poly-Lactic Acid, Poly-glycolic Acid, and Copolymers

The biodegradability of PLA, PGA, and their copolymers makes them attractive for biomedical applications. Thus, the degradation process is an important characteristic of these polymers, and the effect of morphology and copolymer composition, molecular weight, macroscopic shape, macromolecular orientation, and extent of blockiness of copolymers, on degradation rate is an important consideration. The influence of the former two on degradation rate has been documented in the literature. Miller *et al* [19] determined the differences in degradation rates between the pure homopolymers and copolymers of varying monomer compositions. They presented graphically the approximate half-life for different copolymer ratios. The equi-compositional copolymer yielded the highest degradation rate, with a half-life of about one week. Furthermore, through differential isotopic labeling of the glycolic and lactic units, they indicated that the copolymer degrades as a homogeneous copolymer rather t han as a mixture of two polymers degrading at different rates [19].

Gilding and Reed [7] demonstrated that crystallinities of the copolymers vary with varying monomer ratios. Consequently, one may expect correlation between degradation rate and morphology as well as with copolymer composition. The crystallinity variation with monomer ratio has been attributed to the difference in crystallinities of the homopolymers, PGA being approximately 46-52 percent and PLA about 37percent crystalline. A further influence on degradation rate in copolymers suggested by

Dunn *et al* [3] is the extent of blockiness ('segmentedness') of the copolymer, a consideration arising from the estimation of reactivity ratios of the monomers by Gilding and Reed. At a reaction temperature of 200 C, the reactivity of the glycolide and lactide monomer units was determined as Rg = 2.8 and Rl = 0.2. Based on these values, it was concluded that a chain with a terminal glycolide unit has a three to one preference for adding glycolide [7]. This behavior leads, effectively, to bloc ks of glycolide separated by single lactide units. Because the two polymers demonstrate significantly different properties, the extent of blockiness, and its control, is of interest in synthesizing these polymers.

1.2.3 Polymer Surface Reconfiguration and Surface Chain Mobility

An important aspect of the practical application of bioresorbable polymers is that the ultimate physiological environment in which the polymers are used is dramatically different than that under which they are fabricated, processed, and characterized. Thus, by implanting these polymers into the body, one is significantly modifying their environment. In recent years, several authors have proposed that, contrary to the conventional assumption that polymer surfaces are rigid, unperturbable planes on which substances may adsorb and chemical reactions take place, glassy polymer surfaces are highly mobile and that the surface configuration is sensitive to the immediate surroundings [20]. That is , the conformational state of macromolecules at an interface responds to a variation in the contacting medium. Surface configuration is the specific spatial arrangement of chemical functionalities of a macromolecule at the polymer-fluid interface [21].

One may question at this point why surface dynamics would exist in a *glassy* state in which large-scale molecular motion is arrested. The answer lies in the question–that is, below the glass transition temperature of glass-forming polymers, only large-scale motions are quenched, whereas surface reconfiguration is held to occur via short-ranged motions such as rotations around bounds (or through concerted motion in

different chains over very small length scales). Moreover, as large segmental motions are certainly arrested, there still exist molecular modes of motion below the glass transition. The transitions associated with these modes are labeled as α, β, with the glass transition being the former and subsequent labels pertaining to transitions at lower and lower temperatures [20]. Furthermore, when the molecular interactions experienced by the chains at the surface change, surface chain segments experience varying degrees of motion.

The phenomenon of surface reconfiguration has been documented in the literature for several systems. Holly and Refojo [22] investigated the water wettability of PHEMA hydrogels of varying water content. They performed contact angle measurements using the sessile droplet and captive bubble techniques. Their results indicated an unexpectedly high advancing contact angle (thus low wettability), and a large difference between the advancing and receding angles, a phenomenon known as contact angle hysteresis. The authors offered the following interpretation. When polymer chain segments orient themselves so as to expose the hydrophobic moieties to the air and bury the hydrophilic elements in the aqueous phase within the gel. The gel surface thus exhibits hydrophobic character via a high advancing contact angle. Once the hydrophobic gel surface is immersed in water, the polymer segments reorient so as to expose more of the hydrophilic moieties, thus lowering the surface tension and equivalently the contact angle. Furthermore, results based on ESCA and Auger spectroscopy indicate a surface reconfiguration of polymers below the glass transition [23]. Also, Tezuka *et al* [24] studied contact angles via the captive bubble technique on surfaces of block and graft copolymers of polyurethane and polysiloxane. The authors observed that a change of contacting medium from a dry to an aqueous phase induces a dynamic surface reconfiguration causing a polysiloxane segment dominated surface to become polyurethane enriched. This is intuitive if one considers that the polyurethane component, containing polar ester linkages, is more hydrophillic than the siloxane block, which includes significantly more methyl groups. Hence, overall, the driving force for the surface rearrangement process is held to be the minimization of interfacial free energy in response to a change in contacting medium. Further-

more, the authors showed variation of contact angles and contact angle hysteresis with varying blockiness of the copolymer–specifically, with varying siloxane segment lengths.

An interesting phenomenon associated with the structural rearrangement of the polymer is the effect of the polymer structure on the structure of the contacting medium which in this specific case will be water. The properties of liquid water in the vicinity of solid interfaces have been observed to differ from those in the bulk phases. This line of reasoning is extended to mean that the *structure* of water near an interface will differ from its structure in the bulk; this 'modified' water, referred to as *vicinal water* [25], has been studied experimentally and theoretically, and it has been ascribed physical properties such as density, viscosity, heat capacity, and ion selectivity which differ from the analogs in the bulk. Furthermore, there is evidence to suggest that the vicinal structure decays exponentially over a range of 3-6 nanometers from an interface [26]. This is a rather long-ranged ordering phenomenon and has been shown to significantly influence biological systems. In the context of the present work, it would be interesting to determine what the vicinal structure of water is as captured from a simulation model of the esters and water.

1.3 Objectives

Based on the above considerations of the importance of the interface within the context of biomedical applications, this thesis will attempt to investigate the nature of the surface (both free and that in contact with an aqueous medium) of an oligomeric model of polylactic acid using the molecular dynamics method. More specifically, this work will focus primarily on the molecular-level structure of the ester interface in terms of the orientation of particular functional groups along the backbone of the molecule. As discussed above, the nature of the interface, specifically in terms of its chemical composition (e.g., what chemical species does the interface present to the contacting medium) which in turn is directly related to the orientational structure of the constituent molecules. For example, as discussed earlier, the site of hydrolytic

scission of the ester linkage is the carbonyl carbon, and for this reason, it is logical to discern the behavior of this functionality. Moreover, of profound interest is the response of this orientation to a change in the contact medium, as well as any effects, structural or otherwise, induced in the contacting medium, the latter representative of phenomena such as those associated with vicinal water. Furthermore, the conformational distributions of bond, torsion, and out-of-plane angles will be studied in order to derive a clearer picture of the molecular geometry in the interface relative to the bulk. The fluid structure will also be studied by noting the behavior of pair correlation functions in the bulk and interfacial regions. As for thermodynamic properties, the surface tension of the liquid-vapor and interfacial free energy of the liquid-liquid interface will be computed, although quantitative agreement with experimental data is not expected based on the large inherent errors associated with the computations and the anticipated deficiencies (with respect to predictive capabilities of interfacial properties) of the potential models used. It is emphasized here that the primary objective of this work is not to develop a potential model *de novo* for the materials studied; rather, one would like to apply an existing model to a novel system, in this case the ester- and ester-water interfaces. This is not to say that modifications or additions to existing potentials will not be considered; on the contrary, when needed, selected models will be modified to conform to experimental observations. Thus, issues concerning selection and validation of suitable models will be initially addressed in this work.

Chapter 2

Molecular Dynamics: Theory, Methodology, Application

2.1 Fundamentals

The molecular dynamics method (MD) is one approach to probing the phase-space of a many-body system by exploiting Newton's physics, although the implementations are invariably formulated in a Hamiltonian sense (more on this below). By generating phase-space trajectories of a defined system, a set of macroscopic properties is derived from the trajectories as raw data [27]. The phase-space trajectory is computed using classical mechanics and analyzed via concepts of kinetic theory, statistical mechanics, and sampling theory. Along the way, classical nonlinear dynamical theory must be invoked to address issues of ergodicity, and the testing of the trajectory is left for conservation principles. The following sections briefly address these topics, and the reader is referred to the relevant literature for more detailed discussions.

2.1.1 Classical Mechanics: Newtonian and Hamiltonian Dynamics

Fundamentally, MD is numerical integration of Newton's equations of motion; however, as the next section will describe, based on the concept of a 'constant of the

motion', the Hamiltonian formulation is more illustrative. Thus, descriptions of both perspectives are given. In Newtonian dynamics, the force on a particle 'i' is related to its translational motion – namely the acceleration, via Newton's' Second Law:

$$\mathbf{F}_i = m\ddot{\mathbf{r}}_i \tag{2.1}$$

where m is the constant mass of the particle and the acceleration is given as:

$$\ddot{\mathbf{r}}_i = \frac{d^2\mathbf{r}_i}{dt^2} \tag{2.2}$$

in which $\mathbf{r}_i$ is the vector giving the location of particle 'i' relative to a laboratory fixed reference frame [27]. For N particles, Newton's second law gives 3N second-order, ordinary differential equations of motion. Finally, Newton's third law gives the relation between the forces between a pair of particles in isolated system; use of this concept is central to the computational approach to MD as will be discussed below. For the present, the third law is:

$$\mathbf{F}_i = -\mathbf{F}_j \tag{2.3}$$

where 'i' and 'j' are particles in an isolated system (no external forces acting on system constituents) [27].

Now, since the functional form of Newton's second law is time-independent, e.g. the form is invariant under time translations, there is expected to be a function of positions and velocities (or momenta) which is conserved [28]. This function is called the Hamiltonian, which is taken to be a function of the particle positions and *momenta* , $\mathbf{p}_i$

$$\mathbf{p}_i = m\dot{\mathbf{r}}_i \tag{2.4}$$

For isolated systems, the Hamiltonian is identically the total system energy, kinetic plus potential; thus, the Hamiltonian becomes more specifically:

$$H(\mathbf{r}^N, \mathbf{p}^N) = \frac{1}{2m}\sum_i \mathbf{p}_i^2 + U(\mathbf{r}^N) = E_{tot} \tag{2.5}$$

To obtain the equations of motion, one first takes the total time derivative of the general Hamiltonian:

$$\frac{dH}{dt} = \sum_i \frac{\partial H}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i + \sum_i \frac{\partial H}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i + \frac{\partial H}{\partial t} \tag{2.6}$$

With no explicit time dependence for H , the last term is identically zero, and the total time derivative vanishes giving the general result:

$$\frac{dH}{dt} = \sum_i \frac{\partial H}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i + \sum_i \frac{\partial H}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i = 0 \tag{2.7}$$

Taking the total time derivative of the isolated-system Hamiltonian gives:

$$\frac{dH}{dt} = \frac{1}{m}\sum_i \mathbf{p}_i \cdot \dot{\mathbf{p}}_i + \sum_i \frac{\partial U}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i = 0 \tag{2.8}$$

Comparing Equation 2.7 and Equation 2.8 gives :

$$\frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m} = \dot{\mathbf{r}}_i \tag{2.9}$$

and

$$\frac{\partial H}{\partial \mathbf{r}_i} = \frac{\partial U}{\partial \mathbf{r}_i} \tag{2.10}$$

The partial derivative of H with respect to the particle positions is obtained by substituting Equation 2.9 into Equation 2.7 which gives:

$$\sum_i \dot{\mathbf{r}}_i \cdot \dot{\mathbf{p}}_i + \sum_i \frac{\partial H}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i = 0 \tag{2.11}$$

$$\sum_i (\dot{\mathbf{p}}_i + \frac{\partial H}{\partial \mathbf{r}_i}) \cdot \dot{\mathbf{r}}_i = 0 \tag{2.12}$$

With all velocities independent of one another, each term in parentheses vanishes,

giving for each particle:

$$\frac{\partial H}{\partial \mathbf{r}_i} = -\dot{\mathbf{p}}_i \tag{2.13}$$

Thus, Equation 2.9 and Equation 2.13 are the Hamiltonian equations of motion, and for a system of N particles, give 6N first-order differential equations which are equivalent to the 3N second order Newtonian equations [27]. Thus, both descriptions give the same dynamics albeit in different ways. The Newtonian perspective gives rise to motion as a result of applied forces, while the Hamiltonian description does not include forces explicitly; motion occurs in a way to conserve the Hamiltonian. Note in passing the assumptions leading to the Hamiltonian results. Firstly, for the present derivation, an isolated system was employed, although more useful systems would allow energy and mass transfer across boundaries. Even still, the Hamiltonian, which would then include extra terms to describe these interactions, would be a conserved quantity. It would not necessarily equal the total system energy, which would not necessarily be a constant as a result of the interactions. This is an important point to note as extended Hamiltonian formulations of MD are used to construct simulations which generate phase-space trajectories in various ensembles (NPT, NVT, NPH, etc). These will be addressed below. Secondly, the relation between velocities and momenta is given by a relation such as Equation 2.9. Finally, there is no explicit time-dependence attributed to the Hamiltonian itself [27].

2.1.2 Sampling of Phase Space: Ergodicity

The previous sections described the machinery which allows one to compute a phase space trajectory for an N-body system in general terms. It should be kept in mind that although a system has an inherent total phase space available to it, external conditions imposed on the system (e.g.. isolation, constant temperature or pressure) limit the volume of this space which is available to any given phase point. Thus, it is this restriction of phase space which is the mechanism by which external constraints are related to specific values of macroscopic state variables [27]. Now, the phase-space

trajectory is used to obtain macroscopic observables by defining the experimentally observable property A_{obs} in terms of the instantaneous value dependent on the phase point of the trajectory (the set of momenta and positions) $A(\Gamma(t))$ as:

$$A_{obs} = < A >_{time} = < A(\Gamma(t)) >_{time} = lim_{t_{obs}\to\infty} \frac{1}{t_{obs}} \int_o^{t_{obs}} A(\Gamma(t))dt \qquad (2.14)$$

where t_{obs} is a sufficiently long time interval. Since the upper limit of integration cannot be extended to infinite time, a long finite time is used for the averaging, and the time average is written as

$$A_{obs} = < A >_{time} = \frac{1}{\tau_{obs}} \sum_{\tau=1}^{\tau_{obs}} A(\Gamma(\tau)) \qquad (2.15)$$

where τ_{obs} is a large, finite number of time steps used for the numerical integration of the equations of motion.

The question now is whether the system trajectory will traverse a sufficient region of the available phase space over which there is a significant probability density to give satisfactory time averages within a tractable computational time. Thus, the problem is now whether the system trajectory samples sufficiently the relevant volume of phase space.

This is relevant to the notion of system equilibration since the system must evolve from a non-equilibrium to an equilibrium state as it moves along a trajectory. Depending on the initial conditions and the *nature* of the trajectory, the final equilibrium state may or may not be reached; it is this equilibrium state over which the time averages must be taken. Thus, the concepts of nonlinear dynamical theory must be applied to describe the trajectory, and ultimately, place necessary criteria on it to substantiate the MD approach.

Without delving into the details of dynamical theory, suffice it to say that the standard assumption made about the nature of MD trajectories is that they are ergodic (or undergo ergodic motion). That is, over a sufficiently long time, each phase point passes through all locations of the constant-Hamiltonian surface. Although

this is not a rigorous definition of ergodicity, and in fact, leads to an inconsistency in the classification of real dynamical systems [27], the modified use of the *ergodic hypothesis* is that the time average introduced above is equivalent to an ensemble average, the ensemble defined as the collection of identical systems (identical in terms of the external constraints imposed on each) each representing a distinct microscopic state in phase space. Now, instead of following a single point as it moves along the ergodic trajectory (visiting all configurations of phase space), the ergodic hypothesis allows one to look at a single 'snapshot' of all the members of the ensemble (each representing a distinct point in phase space), and taking the macroscopic observable property as an average over the instantaneous value of each member. Statistical mechanics gives observables as ensemble averages of instantaneous properties [13].

Thus, the macroscopic observable A_{obs} can be set equivalent to the ensemble average

$$A_{obs} = < A >_{ens} = < A|\rho_{ens} > = \sum_{\Gamma} A(\Gamma)\rho_{ens}(\Gamma) \tag{2.16}$$

where ρ_{ens} is the phase-space probability distribution for the given ensemble, the ensemble being determined by the external constraints on the system. Along an MD trajectory, each configuration generated via numerical integration of the equations of motion can be considered a 'distinct' state, or member of the ensemble, captured at a given time (with the initial time being equal for all configurations) , thus allowing for averaging static properties such as temperature, pressure , structure factors [13].

It should be noted that one can question the validity of the assumption that a finite interval time average is equal to the infinite time average, an equivalence which is central to the MD method and introduced above. Pure statistical mechanics avoids the calculation of time averages by resorting to the ensemble picture; but MD attacks the problem directly by computing a time average for a finite system over a finite time. By the Birkhoff-von Neumann theorems [29, 30, 31, 32], the assumption is valid under circumstances where a limiting average value $< A >$ exists for all trajectories and that for any trajectory, the average $< A >$ is independent of the initial conditions

[27]. This is always taken to be the case for MD trajectories.

Finally, it is mentioned in passing that the stability of the trajectories also is important in terms of the ability of the system to move from non-equilibrium to equilibrium states. Note that an ergodic trajectory can be stable with the resultant trajectory of a perturbation remaining close to and in close correlation with the parent trajectory. Thus, ergodicity does not tell whether a system starting from an arbitrary non-equilibrium state will irreversibility tend towards equilibrium states. For this evolution to occur, the trajectory must be unstable to perturbations. The reader is referred to the literature for details into the stability analysis of various types of trajectories, but it is mentioned here that for system evolution towards equilibrium, all of phase-space must be accessible, and the trajectory motion must be mixing so as to allow divergence of a perturbed trajectory from the unperturbed parent trajectory [27].

For details on statistical thermodynamic properties and ensemble theory, the reader is referred to the plethora of treatises available in the literature.

2.2 Integration Schemes

2.2.1 Introduction: Finite Difference Methods

The set of 3N ordinary second-order differential equations representing the equations of motion in MD are invariably solved via finite difference methods. The general idea is that given particle positions, velocities, and other dynamic information at a particular time t, the same values are determined for a later time, $t + \delta t$ to a desired level of accuracy (the desired level of accuracy is an ambiguous term as there will ultimately be a level of inaccuracy which must be accepted due to machine round-off and algorithmic error effects; these will be addressed below). The difference equations are solved step-by-step using a timestep that is significantly smaller than some characteristic time-scale of the system such as the time required for the molecule to move a distance equal to one molecular length. The time step is also determined by the

numerical algorithm chosen via requirements on the stability of the integrator. There are a host of classes of methods such as Runge-Kutta, predictor-corrector, leapfrog methods for solving finite difference equations. Selection from among these is facilitated by examining some desirable traits for a successful simulation algorithm. First, the algorithm should be fast and require little memory. Second, a large timestep should be applicable without incurring too large an error. Third, it should closely describe the classical trajectory. Fourth, it should be time-reversible and satisfy the conservation laws for energy (more rigorously, the system Hamiltonian) and momentum. Finally, it should be simple in form and straightforward to program [13].

For MD, some of these requirements are more severe than others. In terms of the speed requirement (1), as the force/energy computations are the most cpu-intensive, the absolute speed of the integration algorithm is not of too much concern. As for the last requirement, simplicity is advantageous in terms of the amount of data required to be stored, such as coordinates, velocities, and perhaps accelerations (depending on the algorithm). There is little value in programming complicated integration algorithms for an MD application when the overwhelming time is spent in the force/energy calculations. Thus, a simple method is much preferred over a more cumbersome one [13].

The major issues concerning the integration algorithm of choice are its ability to reproduce faithfully the classical trajectory associated with the initial conditions of the simulation, its stability, and the size of the time-step allowed by the algorithm. These issues all are tied together by the concept of truncation and round-off errors encountered during an MD run [27].

Truncation errors refer to the accuracy with which a finite difference method approximates the true solution to a differential equation. For example, when the finite difference is represented by an Taylor expansion (as is done in the more widely used methods), the *local* truncation error is taken to be the order of the first non-zero term neglected in the expansion. Thus, for a Taylor expansion such as [27]

$$x(t+\Delta t) = x(t) + \frac{dx}{dt}\Delta t + \frac{1}{2}\frac{d^2x(t)}{dt^2}\Delta t^2 + \frac{1}{3!}\frac{d^3x(t)}{dt^3}\Delta t^3 + \cdots \tag{2.17}$$

which is truncated as follows,

$$x(t+\Delta t) = x(t) + \frac{dx}{dt}\Delta t + O(\Delta t^2) \tag{2.18}$$

the local truncation error is second-order in the timestep (as shown).

In general, an algorithm whose truncation error goes as $(\Delta t)^{n+1}$ is considered an n'th-order method. Since the magnitude of Δt is standard MD units is such that $\Delta t < 1$, for a given timestep, higher order methods have a smaller local truncation error than lower order methods. Note that truncation error is an inherent entity; it is present in every approximate finite difference algorithm regardless of the manner in which the computation is carried out, whether on a computer, calculator, or paper [27].

A further concept to reckon with is that of *global* error. Local error is that incurred during one time step of the calculation, while global error is that accumulated over the total duration of the simulation run. Local error is used to define the order of a method, but it is the global error which (being larger) is more important to acknowledge when considering an algorithm. The global error is generally an order of magnitude less than the local error as can be shown as follows. For an n-th order algorithm with local truncation error [27]

$$lte = kx^{n+1}\Delta t^{n+1} \tag{2.19}$$

where k is a constant and x^{n+1} equals the (n+1)'th derivative of x. Over M integration steps, the global truncation error is taken as the sum of the individual local errors:

$$gte = k\sum_{i=1}^{M} x_i^{n+1}\Delta t^{n+1} = k\Delta t^{n+1}\sum_{i=1}^{M} x_i^{n+1} \tag{2.20}$$

Although the derivatives in the sum will vary from iteration to iteration, applying a mean-value theorem, the M-term sum can be replaced by M times the average to give

$$gte = k\Delta t^{n+1} M < x^{n+1} > \tag{2.21}$$

The number of timesteps M is related to the timestep by $M = t/\Delta t$, so the global error becomes

$$gte = k\Delta t^{n+1} \frac{t}{\Delta t} < x^{n+1} > = k\Delta t^n t < x^{n+1} > \tag{2.22}$$

showing that global truncation error does indeed vary as $(\Delta t)^n$ while local error goes as $(\Delta t)^{n+1}$ [27].

Along with the truncation errors introduced by the approximations in the finite difference method, MD suffers from round-off errors resulting from the manner in which the calculations are performed. Specifically, round-off errors is influenced by the number of significant figures maintained at each step, the order of specific calculations, and by approximations made in calculating functions such as exponentials or square roots (for example, using tabulated values for a square-root or exponential function in the interest of time-saving) [27].

Both global truncation and round-off error depend on the timestep Δt chosen for the integration, and so for a chosen algorithm, one must manipulate the magnitude of Δt in order to achieve acceptable error. Complicating the matter is the fact that global truncation error (gte) and global round-off error (gre) vary oppositely to changes in timestep. Global truncation error decreases with decreasing time step as discussed above. Global round-off error decreases with increasing timestep (for a given total simulation time). This is because global round-off error depends on the total number of calculations, and for a given total time, increasing the time step means decreasing the number of times a calculation is performed (and thus a round-off error is incurred); thus, globally, the round-off error decreases with increasing time step [27].

In application, a time step that yields the lowest total error is usually too small

to be useful, giving rise to too much time required for a simulation. Consequently, a larger time step is used with the acceptance of a level of error within desired bounds, the value being chosen via test calculations. The measure of the acceptability of the time step is now tied to another concept, the stability of the integrator.

The stability properties of a method describe the manner in which the algorithm *propagates* truncation and round-off errors iteration to iteration. An unstable algorithm amplifies errors step to step, whereas a stable algorithm does not. Realistically, and practically, the algorithms used for MD simulations are invariably *conditionally stable* – stable for certain time steps, but becoming unstable at a certain 'critical' value (thus leading to perhaps numerical overflow in the computation) [27].

For linear ordinary differential equations, one can perform analytic stability analysis; however, for the non-linear equations encountered in MD, analytic stability analysis cannot be used [27], and so one can resort to approximate methods such as linearizing the differential equations, or applying Lyapunov methods [27]. As these are generally clumsy when applied to the equations of motion used in MD, a numerical analysis using various values of Δt is used. Specifically, a series of short test runs (all at constant total simulation time) are performed with varying timesteps; the limiting value of Δt beyond which the algorithm becomes unstable is identified, and a timestep within this limit is chosen such that the integration is stable and allows conservation of the relevant Hamiltonian [27].

2.2.2 Integration Algorithms

There exist a variety of numerical methods for integrating the equations of motion if the only concern is to arrive at a solution in and of itself. However, with MD, one is concerned with computational economy, an algorithm which allows a large time step so as to probe more of phase-space in a shorter time. On this basis, many schemes, although stable and robust, can be neglected. This is due to the fact that algorithms such as Runge-Kutta or adaptive timestep methods, which have good stability characteristics, require multiple force calculations per interaction per particle. As the force/energy calculation is the heaviest load in the MD prescription, these methods

are overall less desirable, and in general, any integration method needing more than one force calculation per time step is wasteful unless it offers a commensurate increase in the timestep allowable while maintaining the same accuracy [33].

Based on the above considerations, two classes of method have attained widespread use, a low-order leapfrog technique and predictor-corrector approaches. As the MD work presented in this work exclusively uses the velocity Verlet algorithm (with the RATTLE formalism) and variants thereof, the next section describes this algorithm, an example of leapfrog type methods, and the reader is referred to the literature for discussions of other methods [27, 33].

The Verlet Algorithm

The simplest widely used finite difference method is the Verlet algorithm [27]. It is derived simply from the combination of two Taylor series for the position, one for t to $t + \Delta t$, and the second for t to $t - \Delta t$. The first expansion is:

$$x(t + \Delta t) = x(t) + \frac{dx}{dt}\Delta t + \frac{1}{2}\frac{d^2x(t)}{dt^2}\Delta t^2 + \frac{1}{3!}\frac{d^3x(t)}{dt^3}\Delta t^3 + O(\Delta t^4) \qquad (2.23)$$

and the second for the reverse evolution is:

$$x(t - \Delta t) = x(t) - \frac{dx}{dt}\Delta t + \frac{1}{2}\frac{d^2x(t)}{dt^2}\Delta t^2 - \frac{1}{3!}\frac{d^3x(t)}{dt^3}\Delta t^3 + O(\Delta t^4) \qquad (2.24)$$

Adding the above two equations eliminates the odd-order terms giving,

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + \frac{d^2x(t)}{dt^2}\Delta t^2 + O(\Delta t^4) \qquad (2.25)$$

This is the position Verlet algorithm with a local truncation error which goes as $(\Delta t)^4$ and hence a global error of third order in the timestep [27]. As the velocities are not present in this prescription, various schemes to estimate the velocity have been applied. For instance, the velocity at the half-step is estimated as:

$$v(t + \Delta t/2) = \frac{1}{\Delta t}[x(t + \Delta t) - x(t)] \tag{2.26}$$

or first-order central difference estimator used by Verlet originally [27],

$$v(t) = \frac{1}{2\Delta t}[x(t + \Delta t) - x(t - \Delta t)] \tag{2.27}$$

The Verlet algorithm as described here has several drawbacks [34]. The velocities are not explicitly included in the integration scheme and thus required is extra effort and/or storage to compute them (as mentioned with the approximating schemes above). Starting the algorithm is difficult since the values of the positions are required from the previous steps as well as the current; initial conditions [x(0),v(0)] are not sufficient to begin a calculation, and invariably something special must be done at the beginning of the simulation. Furthermore, because of this dependence on the previous time step, it is difficult to restart the simulation with a different time step. Finally, precision is difficult to maintain since quantities of differing orders of magnitude are added at each step (Andersen). Consider that in Equation 2.25, quantities of order $(\Delta t)^0$ and $(\Delta t)^2$ are added, the latter involving the force. With the timestep being less than unity (in the standard units), this means that only a few significant figures of the computed force are used with a resultant inevitable deterioration in precision. Note also that it is difficult to implement the extended Hamiltonian methods for constant temperature and pressure, velocity scaling methods for constant temperature MD, and some non-equilibrium MD schemes which require velocities at each timestep [34].

In response to these disadvantages of the position Verlet algorithm, the velocity Verlet scheme is popularly implemented [35]. The velocity Verlet equations are:

$$x(t + \Delta t) = x(t) + \frac{dx}{dt}\Delta t + \frac{1}{2}\frac{d^2x(t)}{dt^2}\Delta t^2 \tag{2.28}$$

$$\dot{x}(t + \Delta t) = \dot{x}(t) + \frac{\Delta t}{2}(\frac{d^2x(t)}{dt^2} + \frac{d^2x(t + \Delta t)}{dt^2}) \tag{2.29}$$

The local error is third order in timestep [34]. The global error is of order Δt^2,

the same as for the position Verlet algorithm to which it is equivalent. With the velocity formulation, one can start with the positions and velocities at time t and compute trajectories forward with no difficulties. This scheme explicitly includes velocities making extended Hamiltonian and velocity rescaling algorithms more easily implementable. Precision is increased on a given machine since terms of order Δt and Δt^2 are added (for the positions), and Δt^0 and Δt^1 for the velocities [34].

Algorithms for Molecular Systems with Internal Constraints: SHAKE and RATTLE

The traditional Verlet algorithms are not applicable to systems with internal constraints such as bond-lengths or bond-angles (these are further classified as holonomic constraints [28]). For such systems (and for the molecular models used in this work), algorithms derived from the Verlet schemes have been developed. The two commonly used methods are SHAKE [36] and RATTLE [34], the former being analogous to the position Verlet algorithm for systems with constraints, and the latter improving on the deficiencies of the original. The following is a brief discussion of SHAKE and RATTLE without delving into the rigorous mathematics of the methods. The reader is referred to the literature for derivations and further discussion.

The SHAKE algorithm, as mentioned, was introduced by Ryckaert *etal* to treat molecular systems with internal constraints while maintaining the dynamical description in Cartesian coordinates. The forces are given as [36]:

$$\frac{d\dot{x}(t)}{dt} = f[x(t)] + g[x(t), \dot{x}(t)] \tag{2.30}$$

where f contains all the forces arising from intermolecular and intramolecular n-body interactions, and g contains the constraint forces. The Verlet scheme for this differential equation is:

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + \Delta t^2 (f[x(t)] + g[x(t), \dot{x}(t)]) \tag{2.31}$$

Since the intramolecular constraints would eventually be violated due to the in-

exactness of the algorithm, the authors stipulated the use of a g_S function (that is an approximation to the exact g) requiring that the coordinates at the next time step exactly (or to within a preset tolerance) meet the constraints. The requirement is satisfied by the proper choice of time-dependent Lagrange multipliers which are determined iteratively so as to ensure the new configuration satisfies the constraints. The SHAKE algorithm is thus [36]:

$$x(t+\Delta t) = 2x(t) - x(t-\Delta t) + \Delta t^2(f[x(t)] + g_S[x(t), \dot{x}(t)] \tag{2.32}$$

where g_S is the SHAKE approximation for the constraint force.

SHAKE suffers from the same maladies as the position Verlet algorithm. A 'velocity' form of the SHAKE equations such as:

$$x(t+\Delta t) = x(t) + \dot{x}(t)\Delta t + \frac{\Delta t^2}{2}(f[x(t)] + g[x(t), \dot{x}(t)]) \tag{2.33}$$

$$\dot{x}(t+\Delta t) = \dot{x}(t) + \frac{\Delta t}{2}(f[x(t)] + g[x(t), \dot{x}(t)] + f[x(t+\Delta t)] + g[x(t+\Delta t), \dot{x}(t+\Delta t)]) \tag{2.34}$$

is inconsistent as an iterative procedure since the above two equations suggest that one needs to know $\dot{x}(t+\Delta t)$ before calculating $g(t)$, but needs $g(t+\Delta t)$ in order to get $\dot{x}(t+\Delta t)$ [34]. RATTLE overcomes this obstacle by proposing two different approximations to the exact g, one for the position equation and the second for the velocity equation [34].

The position is given as [34]:

$$x(t+\Delta t) = x(t) + \dot{x}(t)\Delta t + \frac{\Delta t^2}{2}(f[x(t)] + g_{RR}(t)) \tag{2.35}$$

with the $g_{RR}(t)$ being selected to make sure the position constraints are met

The velocity is computed with [34]:

$$\dot{x}(t+\Delta t) = \dot{x}(t) + \frac{\Delta t}{2}(f[x(t)] + g_{RR}(t) + f[x(t+\Delta t)] + g_{RV}(t)) \qquad (2.36)$$

where the forces at the next time step are computed using the new positions, and the g_{RV} is another approximate constraint force function which makes sure that the new velocities satisfy the time derivatives of the constraints exactly.

Thus, RATTLE can be used to integrate the equations of motion for systems with molecular internal constraints. It computes both positions and velocities one step forward with a knowledge of the positions and velocities at the current time step only, without having to resort to earlier data. The coordinates satisfy the constraints at *each* time step, the velocities also meet constraint conditions at every step, and the precision is comparable to the velocity Verlet algorithm. The algorithm incurs second-order global error and is easily applicable to extended Hamiltonian methods. Although RATTLE requires two force calculations per iteration compared to SHAKE's one, since the bulk of the cpu time is spent on the force/energy calculations, the difference is virtually negligible [34]. The RATTLE integrator is used for the current research purposes.

2.3 Extended Hamiltonian Methods

In the following sections, brief descriptions of the approaches used to probe dynamics in the canonical and constant temperature-constant pressure ensembles are given. A further review for constant temperature and pressure methods is given in the reference by Andersen *et al* [37].

2.3.1 Constant Temperature Methods

Stochastic Methods

From a statistical mechanical point of view, the physical system corresponding to the canonical ensemble is one in thermal contact with a heat bath/reservoir allow-

ing random interactions of the system particles with particles of the bath. In this manner, there is an interchange of energy across system boundaries. In Andersen's method [38], the velocity of a randomly selected particle is replaced with a value from the appropriate Maxwell-Boltzmann distribution; this corresponds to a collision event and subsequent energy transfer between the thermal bath particle. The system under study continues to move along a constant energy surface until the next 'collision' at which point it jumps to another constant-energy hypersurface. This allows the system to sample regions of phase-space in a Markov fashion, and the limiting probability distribution of the trajectory is the canonical phase-space density. Hence, the random velocity changes introduce the stochastic element which effectively facilitates the sampling of phase-space (this is in much the same philosophy of hybrid Monte Carlo techniques used in polymer dynamics simulations – see Chapter 3 for a discussion of Hybrid Monte Carlo as applied to the current research work). The collision frequency is determined by selecting time between collisions from a Poisson distribution centered about a desired mean collision time. As for the frequency at which collisions are made, if collisions occur infrequently, kinetic energy fluctuations will occur slowly and in much the same manner as conventional MD. With frequent collisions, the energy fluctuates dramatically, and the systematic fluctuations are overpowered by those associated with the underlying collisions. To simulate a volume element in a real fluid in thermal contact with some sort of heat bath, Andersen proposes a collision rate [38]

$$f_{collision} = \frac{\lambda_T}{\rho^{\frac{1}{3}} N^{\frac{2}{3}}} \tag{2.37}$$

where λ_T is the thermal conductivity.

Instead of altering the velocity of a single particle, 'massive' stochastic collisions can be used to change the velocities of all particles at once; the collisions occur less frequently and at equally spaced intervals [39].

Heyes [40] proposes a scheme in which all particle velocities are rescaled and an acceptance/rejection criterion is applied to probe phase-space.

Note here that the above methods do not begin with a well-defined Hamiltonian and thus are not as rigorous as extended Hamiltonian methods in averaging for equilibrium properties.

Extended Hamiltonian Methods

Extended system methods treat constant temperature simulations by introducing an extra degree of freedom representing the external heat reservoir to the Hamiltonian and carrying out a simulation of this extended system [13]. The extended Hamiltonian is conserved (the extended system is essentially an isolated system), and the dynamics of the original physical system probe the canonical phase-space. The extra degree of freedom is denoted by s and has a conjugate momentum p_s. The real particle velocities are scaled by s as

$$\mathbf{v} = s\dot{\mathbf{r}} = s(\mathbf{p}/m) \tag{2.38}$$

The potential energy associated with s is

$$U_s = (f+1)k_B T ln s \tag{2.39}$$

where f is the number of degrees of freedom of the actual system and T is the desired temperature. This choice of potential energy allows the canonical ensemble averages to be recovered [41]. The kinetic energy associated with s is taken to be:

$$KE_s = \frac{1}{2}Q\dot{s}^2 = \frac{p_s^2}{2Q} \tag{2.40}$$

in which Q is a thermal inertia parameter with units of $(energy)(time)^2$. The extended Hamiltonian is:

$$H_s = KE + KE_s + U + U_s \tag{2.41}$$

with the extended system density function begin microcanonical (as alluded to earlier)

$$\rho_{NVE_s}(\mathbf{r}, \mathbf{p}, s, p_s) = \frac{\delta(H_s - E_s)}{\int d\mathbf{r}d\mathbf{p}dsdp_s\delta(H_s - E_s)} \tag{2.42}$$

As Nose shows, by eliminating the variables associated with s, one recovers the canonical distribution function. The conservation of H_s acts as a check on the stability and accuracy of the algorithm [41].

The thermal inertia parameter, Q is chosen by trial and error so as to achieve equilibration rapidly without entailing long-lived correlations in energy [41].

The detailed dynamics of the extended system dependes on the value of Q, the thermal inertia parameter, selected. With small Q values, the s degree of freedom decouples from the real system, while large Q values result in inefficient sampling of phase space. The optimum scenario is one in which the time scales for fluctuations of the physical system and the s variable are of the same order of magnitude. The frequency for the s oscillation is determined from its equation of motion (generated from the extended system Lagrangian) [41]:

$$Q\ddot{s} = \sum_i \frac{\mathbf{p}_i^2}{m_i s^3 - \frac{f+1}{s} k_b T_{eq}} \tag{2.43}$$

where f is the number of degrees of freedom of the real system (excluding the s degree of freedom). Assuming the extended system to be in equilibrium, and that s fluctuates around an average value, $< s >$, as $s = < s > + \delta s$ [41]. Combining this relation for s with Equation 2.43 gives

$$Q(\delta\ddot{s}) = f k_b T_{eq} \left(\frac{< s >^2}{s^3} - \frac{1}{s} \right) = -\frac{2 f k_b T_{eq}}{< s >^2} \delta s \tag{2.44}$$

which is the equation of a harmonic oscillator with frequency [41]

$$\omega = \left(\frac{2 f k_b T_{eq}}{Q < s >^2} \right)^{1/2} \tag{2.45}$$

and period,

$$t_o = \frac{2\pi}{\omega} = 2\pi \left(\frac{Q < s >^2}{2 f k_b T_{eq}} \right)^{1/2} \tag{2.46}$$

The Q value can be chosen such that the oscillation frequency is the same order of magnitude as the second moment of the frequency spectrum of the velocity autocorrelation function of the real system. This time scale approximates the time it would take a sound wave to travel the nearest neighbor distance in the fluid.

It is noted in passing here that other methods such as those of Berendsen et al, Hoover, and Langevin methods [42, 43] exist and are popular for equilibration purposes. The reader is referred to the literature for further details.

2.3.2 Constant Pressure Methods

Extended Hamiltonian Methods

A popular extended system method for constant pressure MD is that proposed by Andersen [38] in which the actual system is coupled to an external variable, V representing the volume of the simulation cell. The coupling is considered akin to the action of a piston of 'mass' W on a container of the fluid, the volume V giving the coordinate of the piston. Note that this picture is not totally true, as the 'piston' is not of the usual cylindrical type which contracts or expands the system along one dimension, but rather, changes in V cause an *isotropic* change in the system volume (contraction/expansion) [38]. The piston degree of freedom is given a kinetic energy,

$$KE_V = \frac{1}{2}W\dot{V}^2 \tag{2.47}$$

and a potential energy,

$$U_V = PV \tag{2.48}$$

with P the desired pressure [38].

The constant pressure and constant temperature extended system methods are generally combined to generate trajectories sampling the constant NPT phase-space. The extended system variables, also called virtual variables, are related to the real system variables $(\mathbf{q}_i', \mathbf{p}_i', s, V, t')$ by scaling of the coordinates by $V^{1/3}$ and scaling of

the time by s as follows (note that the values of the coordinates, $\mathbf{q}_i$ are limited to [0,1]),

$$\mathbf{q}_i^{'} = V^{1/3}\mathbf{q}_i \tag{2.49}$$

$$\mathbf{p}_i^{'} = \mathbf{P}_i/V^{1/3}s \tag{2.50}$$

$$t^{'} = \int^t \frac{dt}{s} \tag{2.51}$$

The extended system Hamiltonian is

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m_i V^{2/3}s^2} + \phi(V^{1/3}\mathbf{q}) + \frac{p_s^2}{2Q} + gkTlns + \frac{p_V^2}{2W} + P_{ex}V \tag{2.52}$$

The equations of motion are given elsewhere [44]. Importantly, the averages of functions of $\mathbf{p}^{'}, \mathbf{q}^{'}, and V$ over the extended system trajectory are equivalent to equilibrium properties of the constant NPT ensemble.

The choice of the pressure inertial parameter, W, is again governed by a balance of the fluctuations of the real system and the pressure inertial degree of freedom. A low value for W will lead to reapid box size oscillations that are not damped efficiently by particle motions. A large value explores volume-space sluggishly, and an infinite value retrieves conventional MD. To simulate a small volume element of fluid, again, the time scale for box-size fluctuations should be close to the time needed for a sound wave to traverse the simulation box length [38].

Again, the attractiveness of this method is that one has a well-defined Hamiltonian function which is a constant of the motion over the microcanonical extended system; this property is useful for checking the stability and accuracy of the algorithm as well as computing equilibrium NPT properties.

2.4 Molecular Dynamics: Further Issues

The previous sections have introduced the reader to some of the fundamental concepts associated with the MD method. From an implementation perspective, there are a host of so-called 'tricks of the trade' that are applied within specific MD applications. These techniques are used to expedite the computations and/or, address some of the obstacles encountered during implementation. Ideas such as periodic boundary conditions, minimum image convention, neighbor lists and linked cells, lookup tables, shifted-force potentials, and techniques for long-range interactions (just to mention a few) are thoroughly reviewed in the extensive literature describing the MD technique and will not be addressed here. Furthermore, aspects of parallel processing and the implementation of MD within this framework are not discussed here. Parallel methods were not adopted for the current research although it is strongly felt that such methods would certainly have benefitted the computations. Larger system sizes would be possible, a strong advantage for interfacial simulations which suffer from low statistical precision (or require very long simulation trajectories for proper statistics). This is naturally a result of the fact that only particles at the interface contribute to its properties, and for a given system, the number of particles at the interface is much less than in the bulk phases.

Before concluding this chapter with a discussion on forcefields, we mention that MD is completely deterministic approach to generating phase-space trajectories. There is a temporal correlation between particle positions, and this is the mechanism which yields a *dynamic* picture of the system. On the other end of the spectrum are Monte Carlo techniques which generate trajectories via stochastic events; each phase point is then determined by the phase point before it (on any given trajectory), and the whole process is termed Markovian. Both the totally deterministic and totally stochastic approaches have their benefits and drawbacks, and of course, there are hybrid methods which try to exploit the advantages of both extremes.

Molecular dynamics suffers from the limitations of computational hardware, namely, speed and storage. Thus, studies are generally performed on system sizes ranging in

the 10^2 to 10^3 particle number range, although Swope *et al* report a simulation study for 10^6 particles. Because of speed limitations, short-lived phenomena are generally targeted for study; these include events on the timescale of 100 to 1000 picoseconds. The characteristic relaxation for relevant phenomena must be small enough so that several relaxation times are sampled over the total simulation duration. Note that this limitation, however, MD is suitable for studying fast events which may not be accessible experimentally, and allows a test of theories for these types of phenomena.

2.5 Interaction Potentials

To understand interaction potentials and their place within the MD approach, one must begin by examining the non-relativistic time-independent Schrodinger, equation:

$$H\Psi(R,r) = E\Psi(R,r) \tag{2.53}$$

where H is the quantum mechanical Hamiltonian operator, Ψ is a stationary wavefunction, and E is the energy eigenvalue associated with the system state Ψ . Note that rigorously, the complete description of a molecular system would be given by a relativistic, quantum mechanical model of a system of particles, albeit the solution of such a problem would be a formidable task [45]. Unfortunately, as there currently exists no full relativistic quantum mechanical theory to suitably describe molecular systems, the starting point is taken as above; in the final analysis, classical methods such as MD and molecular mechanics incorporate quantum and relativistic effects via the empirical data on which they are in part based. Referring back to Equation 2.53, the wavefunction Ψ depends on the nuclear and electronic coordinates. As this equation is essentially unsolvable in a practical and exact sense for any but the smallest systems, approximations are invoked to allow a tractable solution. The classic approximation in this case is that of Born-Oppenheimer [46] which exploits the disparity in the nuclear and electronic masses, and subsequent differences in the time scales of their motions, to decouple the nuclear and electronic motions. This leads to two separate equations describing the dynamics of the electrons and the nuclei. The

electronic motion is given by [47]

$$H_{elec}\psi_{elec}(r:R) = E_{elec}\psi_{elec}(r:R) \tag{2.54}$$

where the electronic wavefunction now depends explicitly on the electron coordinates and parametrically on the nuclear coordinates. This is akin to solving for the dynamics of electrons within the field of stationary, heavy nuclei (the Born-Oppenheimer approximation). Importantly, the energy defined here is a function of the nuclear coordinates only, and is termed the potential energy surface [47].

The nuclear motion is then given by [47]

$$H_{nuc}\Phi(R) = E_{nuc}\Phi(R) \tag{2.55}$$

See Appendix A for a more detailed discussion of the quantum mechanical aspects and definitions of the operators introduced here (also given is a brief introduction to applied quantum mechanical methods used in the work presented in Chapter 6).

Ab initio quantum mechanics deals with the solution of the electronic equation, the computation the electronic energy as a function of nuclear coordinates. Also available for this goal are semi-empirical methods which approximate the integrals generated within the mathematical framework of *ab initio* schemes by empirically fit functions [48].

Now, to obtain a dynamical picture of the molecular system, one can solve for the potential energy surface, Equation 2.54, and then solve the nuclear equation for the dynamics of the nuclei on this surface. But since the computational cost to derive the potential energy surface is immense (especially for practical systems), some empirical potential is fit to the potential energy surface (and/or in part to experimental data) and instead of solving the quantum dynamical equation for the nuclear dynamics, one takes a classical approach to the nuclear motion, thus arriving at molecular dynamics.

Newton's equations of motion,

$$-\frac{dU}{dR} = m\frac{d^2R}{dt^2} \tag{2.56}$$

are solved for the nuclear dynamics. Note here that central to the method is the empirical (or semi-empirical) function, U, fit to the underlying (true) potential energy surface. The function U is the forcefield (or interaction potential, or potential if encountered in the MD literature). It defines the coordinates in which the dynamics is framed, gives the functional form of the equations describing the physical interactions within the system, and proposes a set of parameters for the functions which yield the best fit of the forcefield to the potential energy surface [48].

Concerning the functional form of the interaction potential, several comments must be made. For a general N-particle system, the potential energy is a complicated entity composed of a progression of many-body interactions as [13]

$$U = \sum_i u_1(\mathbf{r}_i) + \sum_i \sum_{j>i} u_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \sum_{j>i} \sum_{k>j>i} u_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \cdots \tag{2.57}$$

The first term represents the interaction of each particle with an external field. The second term is the two-body interaction, or the pair potential. This turns out to be the most important contributor in terms of computational implementation. An example of this type of interaction is the Lennard-Jones pair potential which realistically describe the repulsive, close-range forces and the longer ranged dispersion/van der Waals interactions [13]. The attractive long range tail arises from the electronic correlation effects, while the steep short-ranged wall is due to overlapping of electron clouds. The third term is the three-body, or triplet interaction term. The three-body energy is documented to account for up to 10 percent of the potential energy (in lattice and liquid systems), and so these contributions are fairly significant. Higher order terms are generally taken to be negligible, although in long-chain molecular systems, intramolecular interactions such as torsions and improper torsions (out-of-plane angles) can be considered to be a special type of four-body interaction.

Despite the acknowledged contribution of three-body effects in simulations, these are not universally included in computations because of the high cost incurred by the sum over triplets. In short, the trade off between including these contributions

and computational time is one that is weighed case by case. It can be said that the two-body interaction can then be considered an 'effective' potential, effective in the sense that it includes the three-body influence [13]:

$$U = \sum_i u_1(\mathbf{r}_i) + \sum_i \sum_{j>i} u_2^{eff}(\mathbf{r}_i, \mathbf{r}_j) \tag{2.58}$$

It should be noted that as a consequence of including higher order effects in the effective pair potential, there may be incurred a density and temperature dependence, whereas the inherent two-body interaction does not depend on thermodynamic state [13].

Molecular forcefields include two types of coordinates to describe the potential energy surface; the first are internal coordinates such as bond distances, angles, and torsions and the second are non-bonded distances defining the dispersive and electrostatic interactions. The functional forms for each type of interaction run over a gamut depending on the philosophy espoused by the developer of the forcefield (just consider the many commercial forcefields available and focus on the diversity in the functional forms of the internal coordinates used). These can be simple quadratic forms to higher order polynomials to Morse functions, Fourier expansions, and the above-mentioned Lennard-Jones pair potential. The forcefield attempts to describe an entire class of compounds with the same functional form with a high degree of accuracy in terms of reproducibility of experimental and theoretical observations.

This section concludes with a final note concerning forcefields. The forcefield is an attempt to understand fundamental quantum mechanical physics in a classical sense as this is computationally tractable (again, this tractability is a relative concept). One can question whether the MD approach with its classical description is reasonable to apply in describing what are fundamentally quantum effects. For example, consider the classical and quantum harmonic oscillator probability density functions– for the quantum oscillator, there is a finite probability of finding the particle outside the region defined by the conservation of its total energy, a phenomenon known as tunneling. This obviously cannot be captured by any classical approach. However,

the numerous experimental properties such as vibrational frequencies, vaporization energies, and crystal structures contain the quantum effects, and as the forcefield is fitted to these quantities, it in effect has incorporated within it the fundamental quantum nature of all systems.

2.6 Application of MD to Investigations of Interfacial Phenomena: Liquid-Vapor and Liquid-Liquid Interfaces

Complementary to the host of experimental methods (see Appendix B) currently applied to the study of interfacial systems, computational approaches to the study of interfacial systems have established themselves mainly due to the phenomenal advances in computing resources (hardware and algorithms) over the last decade [49]. Computer simulations have been used to probe the molecular-level physics of bulk liquids, and liquid-vapor, liquid-liquid, and liquid-solid interfaces. One of the earliest applications of computer simulations was the Monte Carlo study of the benzene-water interface by Linse [50]. This work demonstrated the feasibility of simulating the liquid-liquid interface in much the same manner as the liquid-vapor had been studied up to that time; the step forward was in the use of periodic boundary conditions in all three dimensions. Furthermore, this work showed that the model potentials used for simulations could indeed reflect the experimental data, particularly with respect to preserving a two-phase system (in effect, in agreement with observed mutual solubilities). Finally, Linse showed that with an atomistically detailed model, one could observe preferential molecular orientation at the interface. In this regard, Linse reports that the water structure in the interfacial region is distinct from that in the bulk; the nonpolar benzene phase induces water dipoles to align parallel to the interface; this structural arrangement is limited to within the interfacial region. This orientational ordering is much the same as that observed by Gubbins and Thompson [51] for molecular dipoles (e.g., chlorine). Meyer *et al* [52] explored reported the effect

of interaction strength between molecules on liquid-liquid miscibility; furthermore, they also observed that diffusion in the interface is anisotropic.

The study of small-molecule liquid-liquid and liquid-vapor interfacial systems has since led to a large body of literature on various aspects of these types of systems. The liquid-vapor interface of water has been the subject of numerous studies probing thermodynamic and structural aspects of water at this interface; furthermore, these studies use a wide range of potential models for water. Townsend and Rice [53] report on molecular dynamics simulations of the water liquid-vapor interface using the ST-2 and Lemberg-Stillinger-Rahman potentials. The authors compute density profiles and an liquid-vapor interfacial "10-90" thickness of 3.45 angstroms in good agreement with an experimental value of 3.30 from X-ray reflectance measurements. They also report on the water orientation at the interface, claiming that the water molecules prefer to lie with the HOH bisector in the interfacial plane with one OH bond jutting out into the vapor phase. Furthermore, the authors report that the bulk structure (from pair correlation functions computed as a function of position from the interface) varies with position, transitioning from the bulk tetrahedral configuration to a dimerized state in the interface.

Taylor *et al* [54] have performed MD simulations of the liquid-vapor interface of SPC/E water reporting a bimodal distribution of water orientation at the interface as well as an increased self-diffusion constant relative to the bulk.

Matsumoto and Kataoka [55] also have reported on MD simulations of the water liquid-vapor interface. Their simulations use the Carravetta-Clementi potential. The authors find orientational structuring of water near the interface, with the vapor side being dominated with water molecules projecting one hydrogen atom toward the vapor phase, and the liquid side consisting of molecules close to parallel to the interface with the hydrogens slightly directed towards the bulk liquid. These results are in much the same spirit of the Taylor *et al* results.

The same authors also present an MD study of the liquid-vapor interface of methanol which indicates orientational ordering in the interface with the methanol methyl group projecting towards the vapor phase, thus allowing for more energetically

stabilizing hydrogen-bonding in the vicinity of the interface [56].

Zhang *et al* [57] discuss their MD simulations of the octane-water liquid-liquid interface. The authors report surface tension values for the octane-water, octane-vacuum, and water-vacuum interfaces calculated with the CHARMM PARM22b4b parameter set to be in good to excellent agreement with experimental data. They also report of water ordering in the interfacial region.

van Buuren *et al* [58] report a molecular dynamics of the decane-water interface. They study the sensitivity of the surface properties to the van der Waals parameters, noting the change in the sharpness of the interface with changing interaction strengths. Furthermore, the authors report an orientational preference for water in the interface, while the decane molecules orient more parallel with the surface.

Carpenter and Hehre [59] performed molecular dynamics simulations of the hexane-water interface, reporting much the same flavor of information as mentioned above.

Finally, it is mentioned that more complex systems at interfaces have also been probed with computational techniques. Harris investigated the free surface of alkane oligomers using chemically realistic models [60]. Tarek *et al* [61] studied tetradecyltrimethylammonium bromide monolayers at an air/water interface using MD. Others have offered simulations on hexadecane between two nickel plates [62], hydrocarbons between impenetrable hard walls [63], and an energy minimized amorphous polypropylene glass at the free surface and at the polymer-graphite interface [64].

Chapter 3

Forcefield Selection and Validation

3.1 Introduction

As our research approach involves the simulation of lactic and glycolic acid polyesters via molecular dynamics, the initial task was to select and validate a suitable forcefield/potential model for the systems of interest. Moreover, as structural properties such as molecular conformations,orientations of specific groups (e.g. carbonyl oxygens, terminal methyl groups), etc. are of prime interest, the initial requisite criteria for the potential were of a geometrical nature. More specifically, the potential model was required to reproduce experimental and theoretical geometries and energy differences between conformational states. Concerning thermodynamic properties, enthalpies of vaporization were computed using different potential models. These computations also afforded an opportunity to investigate two approaches for handling long- range electrostatic interactions–the Ewald summation and Reaction Field methods.

3.1.1 Forcefields

As alluded to earlier, the potential or forcefield model to be input to a molecular dynamics simulation lies at the heart of the approach. Numerous factors affect the qualitative and quantitative accuracy of a simulation (for example, system size, treat-

ment of long-range forces which may lead to long-range structural artifacts, etc.). Unfortunately, despite any pains to minimize deficiencies with respect to these factors, simulation accuracy will suffer due to insufficiencies in the underlying forcefield [65]. With this in mind, some effort was spent to determine a suitable forcefield model for small molecule esters which could then be extended to longer chain entities such as oligomers.

It is pointed out that attention has been limited to forcefields focused more on biological applications. There are a host of forcefields dedicated to studying a myriad of systems ranging from zeolite structures, metallic liquids, semi-conductor crystals (and other non-organic crystalline systems), to the simple yet surprisingly complex liquid, water! The potentials for water represent a class of forcefields in themselves and will be discussed further below.

Several general observations concerning the nature of forcefields are worth noting. Fundamentally, the forcefield model is an attempt to model the potential energy surface of an arbitrary arrangement of atoms or molecules (in the case where there is unambiguously definable molecular entity). The physics it tries to capture is the motion of heavy nuclei in the electronic field of the much lighter and faster electrons (within the Born-Oppenheimer approximation). As traditional molecular dynamics implementations are of a classical nature, quantum effects are not explicitly addressed. Thus, the dynamics of the system is governed by classical Newtonian dynamics, and classical statistical mechanics provides the connection from the microscopic system properties to macroscopic observables. Note that the classical approximation is justified as long as the thermal de Broglie wavelength,

$$\Lambda = \frac{h}{(2\pi mkT)^{1/2}} \tag{3.1}$$

is much less than some characteristic spacing, a, between system entities [66]. For instance, in the case of argon near its triple point, Λ/a is roughly 0.08 and quantum effects are negligible. In lighter liquids such as neon and helium, with ratios of 0.26 and unity, quantum effects are dominant and warrant care [66]. For the present work,

the thermal de Broglie wavelength is small enough to validate a classical description of our system. As a simple demonstration, the wavelength can be calculated as follows, with the shown typical values for relevant parameters. As a stringent test, the lightest mass, that of hydrogen, is used; the volume is taken to be a that which would be used for a typical small molecule bulk molecular dynamics simulation :

$$\frac{\Lambda}{V^{1/3}} << 1 \tag{3.2}$$

$$\Lambda = \frac{h}{(2\pi mkT)^{1/2}} \tag{3.3}$$

$$h = Planck'sConstant\,,\ m_{hydrogen} = 1.00797amu\ ,\ T = 298K \tag{3.4}$$

$$V = 30nm^3 \tag{3.5}$$

Thus,

$$\frac{\Lambda}{L} = 0.046 \tag{3.6}$$

which is sufficiently small in the limiting case for the current purposes. The classical approximation (high temperature, high mass limit) is safely applied.

In terms of derivation, first-generation forcefields were all established through fitting parametrical forms to properties obtained experimentally. These properties include thermodynamic properties (PVT equations of state, enthalpies of vaporization and sublimation, etc..), vibrational frequencies, gas-phase molecular structures, and crystal structures [65]. This reliance on experimental data has plagued the art of forcefield development since in instances where there is a lack of such experimental data, the parameterizing and testing of accurate potential energy functions become impossible [65]. In response to this situation, newer, Class II forcefields have been introduced. These potentials attempt to generate observables to which functional parameters are fit through ab initio calculations of equilibrium and non-equilibrium

energies, forces (first derivatives of the potential surface), and force constants (second-derivatives). From this approach to forcefield development, one also gains insight to the $functional form$ of the various contributions to the potential energy surface [65].

3.1.2 Forcefields Specific for Molecules Containing the Ester Functionality

At the start of this research, very little work on forcefields specific to simple ester molecules or aliphatic polyesters had been published. The major forcefields proposed for esters include the OPLS potentials by Jorgensen and coworkers (based on MC simulations of acetic acid and methyl acetate) [67], a forcefield for conformational energy calculations on aliphatic polyesters parameterized by Boyd and Smith [68], the CFF91 class II forcefield [65, 69] and the CVFF [70, 71, 72, 73] forcefield for organic systems, and the traditional protein and nucleic acid forcefields including CHARMM [74], AMBER [75, 76], and MM3 [77, 78, 79]. For the present study, the OPLS, CFF91, AMBER, CVFF forcefields were considered initially, although a host of other models were available. These four are chosen since they collectively represent the major types of forcefields used in biomolecular/organic materials simulations (the nature of these types will become apparent in the discussions of the specific forcefields presented in the next sections); furthermore, on a more simplistic level, these particular forcefields were readily available. For comparisons, the test molecule was chosen to be methyl acetate as experimental data on gas-phase molecular structures and energetic properties was available readily. It is emphasized here that the forcefield selected was required to be transferable, and in this respect, highly flexible for simulating ester type compounds. Furthermore, the necessary physics should be included in the model without necessarily encumbering the description of the energy surface. Finally, the forcefield should reproduce experimental and theoretical gas-phase geometries and energy differences between conformational states as well as enthalpies of vaporization. The fidelity to experimental conformational energetics results is emphasized as it is the goal of this work to probe the transitions of long-chain molecules

between conformational states in response to external environments. In particular, the transitions due to rotations of the carbonyl oxygen atom are of interest as this is a relevant site in the hydrolytic degradation of the polymers in vivo.

3.1.3 Optimized Potentials for Liquids Simulation (OPLS)

Jorgensen and coworkers have developed an extensive parameter set for molecular simulations of a broad spectrum of organic liquids [67, 80, 81, 82, 83]. In considering this work, emphasis is placed on the fact that the potentials developed in this lab are for liquids specifically. Up until the time of Jorgensen's work, potentials were parameterized and tested with crystal data on hydrocarbons and proteins and their nucleic acids. These were then applied to liquids as well as solids, an extension whose validity is often quite tenuous. Thus, the work of Jorgensen's groups has been to develop potentials that reproduce experimental data on liquids (solutions or organic and biochemical substrates) and which are transferable to molecules outside the sample used for fitting [80].

Within the OPLS framework, molecules are represented by interaction sites located on the nuclei. The interaction energy between two molecular entities is determined by Coulombic and Lennard-Jones interactions between all intermolecular site-pairs as [80]:

$$U(r_{ij}) = \sum_{i}^{n_a} \sum_{j}^{n_b} \frac{q_i q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}} \tag{3.7}$$

where the sums are over the atoms on different molecules. The first term represents the Coulombic electrostatic interaction, and the second and third terms account for the intermolecular dispersion (van der Waals) and repulsion interactions.

A geometric mean (Lorentz-Berthelot) combining rule is used to obtain the parameters A and C [80]:

$$A_{ij} = \sqrt{A_{ii} A_{jj}} \tag{3.8}$$

$$C_{ij} = \sqrt{C_{ii}C_{jj}} \tag{3.9}$$

Alternatively, the A and C parameters can be expressed in terms of Lennard-Jones ϵ and σ as:

$$A_{ii} = 4\epsilon_i\sigma_i^{12} \tag{3.10}$$

$$C_{ii} = 4\epsilon_i\sigma_i^{6} \tag{3.11}$$

Hydrogens on carbon are implicit, while those on heteroatoms (atoms other than carbon) are explicitly considered [80]. This is the united-atom approximation which has been found to be sufficient in the study of a range of small molecules, bio-organic compounds, and polymers. In general, the united-atom models are parameterized ("calibrated") to fit the thermodynamic properties of the material under study, and in particular, it is the Lennard-Jones type interaction parameters σ and ϵ that are optimized to obtain satisfactory fits to experimental P-V-T data, for instance. Though it is an attractively simple model, the united-atom approach does have several drawbacks as noted in the literature [84, 85, 86]. Toxvaerd reports that the P-V-T behavior of united-atom models of n-alkanes (specifically, propane, pentane, and decane) does not reproduce the experimental isotherms; the molecular dynamics pressures for the series of alkanes scale systematically differently than the experments, with the MD pressure for propane being higher than experiment, and that for the pentane and decane being much lower [85]. The agreement with experiment for all test compounds can be improved by adjustment of the Lennard-Jones (or other relevant interparticle interaction parameters) parameters for each liquid; however, this contradicts the experimental observation that at room temperature, the molecular volume of n-alkanes increases proportionally with the chain length–a result which indicates that the mean volume of a methylene group in an alkane fluid is independent of the chain length [85]. This implies that if a united-atom model were correct, the value of the σ parameter would be consistently the same for all n-alkanes. The fact that the above

mentioned isotherms scale differently and can be make to agree with experiment by using different parameter values, argues Toxvaerd, seems to indicate that deficiency of the united-atom model [85]. Note that this is a deficiency with respect to the bulk thermodynamic properties of the materials studied, and should not be generalized to include all properties such as structural and dynamic. Toxvaerd claims that the united-atom model oversimplifies the underlying details of molecular structure which add anisotropy to the atomistic picture of liquids; the argument is that the full-atomic model influences the pressure through the anisotropy, or orientational dependence, of the site-site interactions–that is, the net interaction between two methylene units becomes orientation-dependent when one goes from a united-atom model to an explicit-atom model [85]. Thus, since a full model would be computationally expensive, he proposes an anisotropic united-atom model as a better approximation to the actual molecular physics [85].

Although the concept of an anisotropic united-atom model does have merit in the molecular physics it attempts to describe, and, as Toxvaerd shows, can faithfully reproduce the equation of state for linear alkanes, the AUA model has not been widely adopted, particularly for simulation work of polymeric materials. Boyd and coworkers have extensively used and studied the application of AUA potentials to simulations of polymers such as polyethylene and polystyrene with respect to reproducibility of P-V-T behavior, molecular packing in crystalline environments, and small-molecule diffusion within such systems [86, 87, 88]. Boyd and Pant report an adjusted set of AUA potential parameters which result in P-V isotherms and enthalpies of vaporization (as a function of temperature) for polyethylene that are in agreement with experimental data; the parameter set is an improvement over the initial Toxvaerd set [86]. As satisfying as the performance of the AUA model is with respect to the properties of polyethylene, there is no definitive evidence that this is universal. In a study of small-molecule diffusion in polystyrene, Han and Boyd report that for the aromatic CH groups used in the potential model, *united atom* model better reproduced the cohesive energy density and temperature/specific volume curves; more specifically, smaller values of the AUA offset distance resulted in better agreement of

the V-T curve and cohesive density. Thus, for this work, the authors adopted a zero offset, effectively giving a united-atom structure for these groups [87]. This certainly seems to indicate that the AUA formalism may not be universally applicable.

After considering the previous thoughts, one must keep in mind that no matter what pains are taken to reproduce certain experimental observations, there will be other data that will not be captured exactly by the forcefield model (to be more precise, one can say that not all properties will be faithfully reproduced to within some preset tolerance for error). One must decide what propoerties are of interest and proceed with the model which will afford the most efficient path to studying them. In the present case, the specific orientations of certain functional groups along a chain backbone are of primary importance; furthermore, the response of this orientation to changing interfacial environments is to be investigated. Thus, we opt to accept the cited deficiencies of this approach and employ it to the studies of orientations at interfaces. Moreover, the primary goal of this work is not to create a potential model *de novo*, and so again, the most expedient approach, to modify existing models, is taken. Finally, one can note the broad range of commercial and free forcefields available that employ a united-atom approach with satisfactory results. Thus, as a matter of simplicity and practicality, the nuclear-centered united-atom approximation is retained in the OPLS as well as the current research [80].

For intramolecular potentials, the OPLS models only four-body interactions by a Fourier series in the dihedral angle, ϕ as:

$$U_{torsion}(\phi) = \frac{1}{2}V_1(1 + \cos\phi) + \frac{1}{2}V_2(1 - \cos 2\phi) + \frac{1}{2}V_3(1 + \cos 3\phi) \qquad (3.12)$$

No other intramolecular interactions are considered in the original OPLS forcefield, and as will be discussed below, this necessitates having to determine parameters for intramolecular motions to perform molecular dynamics simulations.

Since the parameterization of ester compounds was done after a bulk of data on parameters for related compounds such as ethers and amides [80, 81, 82] had been

completed, the Lennard-Jones parameters for the hydroxy and alkoxy oxygens were adopted from ether oxygens, and the parameters for the carbonyl and methyl groups were transferred from those for amides. Initial charges were computed via Mulliken population analyses of 6-31G(d) wave functions for the molecule in its optimized experimental geometry [67, 80, 81, 82, 83]. Charges were adjusted to yield reasonable agreement to experimental dipole moments. The final charge values give a dipole moment of $1.44D$ for the Z conformer of methyl acetate; the experimental value is $1.72D$ [67].

3.1.4 AMBER

The AMBER forcefield was developed specifically for molecular mechanical/dynamical simulations of nucleic acids and proteins [75] following the cartesian-coordinate energy refinement approach of Lifson and Warshel [89]. This is the second generation forcefield developed by Lifson and coworkers. A brief description of the forcefield follows.

The total interaction energy (including inter- and intramolecular contributions) is [75]:

$$U_{total} = \sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_o)^2 + \sum_{dihedrals} \sum_{n=1}^{3} \frac{V_n}{2}[1 + cos(n\phi - \phi_o)] + \sum_{i<j}[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}}] + \sum_{H-bonds}[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}}] \quad (3.13)$$

Bond stretching and bond-angle bending functions are quadratic, allowing sufficient description of the energy surface of slightly strained proteins and nucleic acids. The third term is a Fourier series representation of the torsional energy. A 6-12 function is used for the non-bond dispersion/repulsion interaction. Electrostatic interactions derive from an atom-centered monopole model, with point charges derived from fits to quantum mechanical electrostatic potential calculations. Note that a *distance dependent dielectric* is used for purposes of including solvation effects when

water is not explicitly modelled in the simulation. This is to incorporate the polarization effect in attractive interactions (e.g., interactions at shorter separations are weighted more heavily than those at larger separations). This does not present a difficulty when including water explicitly as a constant dielectric is allowable (and more appropriate) and consistent with the same set of charges. A united-atom representation for $-CH, -CH_2, and - CH_3$ units is invoked for computational efficiency (although the united-atom model yields Φ, Ψ maps of alanyl and glycyl dipeptides similar to those from an all-atom model) [75].

The approach to developing the parameters for this potential model is given here in brief. The reader is referred to the original paper and references therein for further details [75]. The fitting procedure is begun by obtaining equilibrium bond lengths and bond angles from experimental measurements including microwave spectroscopy and neutron diffraction and from earlier molecular mechanics calculation. Torsional constants were obtained from microwave spectroscopy, NMR spectroscopic measurements, and molecular mechanics computations. Non-bonded parameters were initially taken from prior forcefield parameter sets and crystal packing calculations. A partial atomic charge model was fitted to quantum mechanical electrostatic potentials. These parameters are then optimized via molecular mechanics performed on model compounds to reproduce structures, energies, and vibrational frequencies [75].

3.1.5 CVFF: Consistent Valence Forcefield

The CVFF forcefield is an intermediate type (diagonal plus some off-diagonal elements) forcefield developed by Dauber and Osguthorpe [73]. It is a generalized valence forcefield with parameters for amino acids, water, and other functional groups.

The functional forms for the diagonal terms are:

Bond angle bending, represented via a Morse potential:

$$U_b = \sum_b D_b[1 - e^{-\alpha(b-b_o)^2}] \tag{3.15}$$

Bond angle bending:

$$E_\theta = H_\theta(\theta - \theta_o)^2 \tag{3.16}$$

The potential energy function for dihedrals, ϕ is given by:

$$E_\phi = H_\phi(1 + s\cos n\phi) \tag{3.17}$$

The out-of-plane bending is:

$$E_\chi = \sum_\chi H_\chi \chi^2 \tag{3.18}$$

The off-diagonal elements include bond-bond, angle-angle, bond-angle, torsion-angle-angle, and out-of-plane/out-of-plane coupling:

Bond/bond coupling goes as:

$$E_{bb'} = K_{bb'}(b - b_o)(b' - b'_o) \tag{3.19}$$

Bond/angle coupling is:

$$E_{b\theta} = K_{b\theta}(b - b_o)(\theta - \theta_o) \tag{3.20}$$

Angle/angle cross terms are of the form:

$$E_{\theta\theta'} = K_{\theta\theta'}(\theta - \theta_o)(\theta' - \theta'_o) \tag{3.21}$$

Angle/angle/torsion coupling is:

$$E_{\theta\theta'\phi} = K_{\theta\theta'\phi}(\theta - \theta_o)(\theta' - \theta'_o)\cos\phi \tag{3.22}$$

Coupling of out-of-plane angles is:

$$E_{\chi,\chi'} = \sum_\chi \sum_{\chi'}{}' F_{\chi,\chi'}\chi\chi' \tag{3.23}$$

The non-bond dispersion and electrostatic terms are:

$$E_{dispersion} = \epsilon[(\frac{r^*}{r})^{12} - 2(\frac{r^*}{r})^6] \tag{3.24}$$

$$E_{elec} = \sum \frac{q_i q_j}{\epsilon r_{ij}} \tag{3.25}$$

3.1.6 CFF91: Class II Consistent Forcefield from BIOSYM Technologies

The CFF91 forcefield is an all-atom, class II forcefield developed primarily from ab initio data on the quantum mechanical potential energy surface of a sample of model compounds [65, 69]. The philosophy of this and class II forcefields in general is to exploit the information available from ab inito methods to complement available experimental data with the aim of parameterizing quantitative, transferrable empirical energy functions as well as determining the optimum functional forms for these potentials. Quantum mechanical 'observables' are the computed equilibrium energies, first energy derivatives, and second energy derivatives (force constants). The energy derivatives are determined from quantum mechanical calculations on molecules distorted from equilibrium by moving atoms along the normal modes. The forcefield is an all-atom model with an extensive list of atom types (atomic environments). The functional form includes several contributions which are given here individually [65, 69].

The diagonal contributions for bond stretching are given as:

$$E_b = K_{2,b}(b - b_o)^2 + K_{3,b}(b - b_o)^3 + K_{4,b}(b - b_o)^4 \tag{3.26}$$

where b is the bond length coordinate and b_o the equilibrium value.

The diagonal contributions for bond-angle bending are:

$$E_\theta = K_{2,\theta}(\theta - \theta o)^2 + K_{3,\theta}(\theta - \theta_o)^3 + K_{4,\theta}(\theta - \theta_o)^4 \tag{3.27}$$

The potential energy function for dihedrals, ϕ is given by:

$$E_{\phi} = K_{1,\phi}(1 - \cos\phi) + K_{2,\phi}(1 - \cos 2\phi) + K_{3,\phi}(1 - \cos 3\phi) \tag{3.28}$$

The non-bond interactions between atoms i and j (separated by at least three bonds within a molecule) at a separation of r are:

$$E_{nb} = (\frac{q_i q_j}{r}) + \epsilon[2(\frac{r^*}{r})^9 - 3(\frac{r^*}{r})^6] \tag{3.29}$$

The combination rules for relating the individual atomic parameters to the pair-wise potential are [65, 69, 90]:

$$r^* = [\frac{(r_i^*)^6 + (r_j^*)^6}{2}]^{\frac{1}{6}} \tag{3.30}$$

$$\epsilon = \frac{2(\epsilon_i \epsilon_j)^{1/2}(r_i^* r_j^*)^3}{[(r_i^*)^6 + (r_j^*)^6]} \tag{3.31}$$

The next series of functional forms represent the off-diagonal contributions to the intramolecular potential energy.

Bond/bond coupling goes as:

$$E_{bb'} = K_{bb'}(b - b_o)(b' - b_o') \tag{3.32}$$

Bond/angle coupling is:

$$E_{b\theta} = K_{b\theta}(b - b_o)(\theta - \theta_o) \tag{3.33}$$

Angle/angle cross terms are of the form:

$$E_{\theta\theta'} = K_{\theta\theta'}(\theta - \theta_o)(\theta' - \theta_o') \tag{3.34}$$

Angle/angle/torsion coupling is:

$$E_{\theta\theta'\phi} = K_{\theta\theta'\phi}(\theta - \theta_o)(\theta' - \theta_o')\cos\phi \tag{3.35}$$

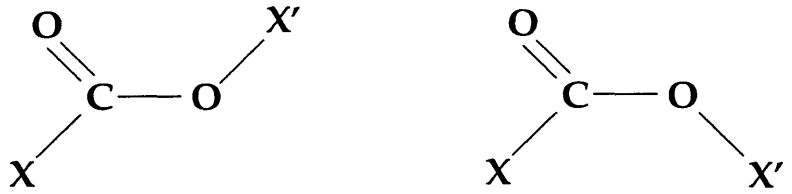
Finally, bond/torsion and angle/torsion coupling are given by:

$$E_{\phi b} = (b - b_o)[K_{1,\phi,b}\cos\phi + K_{2,\phi,b}\cos 2\phi + K_{3,\phi,b}\cos 3\phi] \tag{3.36}$$

$$E_{\phi\theta} = (\theta - \theta_o)[K_{1,\phi,\theta}\cos\phi + K_{2,\phi,\theta}\cos 2\phi + K_{3,\phi,\theta}\cos 3\phi] \tag{3.37}$$

3.1.7 Conformational Energetics Test of Ester Potentials

As the focus of this work is to study the torsional changes in the backbone structure of long-chain aliphatic esters, the forcefield selected for use in molecular dynamics simulations had to describe the rotation about the Carbon-Oxygen bond adjacent to the carbonyl group faithfully. As a test of this requirement, the dihedral profiles about C-O bond were calculated for the selected forcefields, the resulting data being compared to experimental and computational/theoretical data on the behavior of this torsion. Molecules of the group $X - C(O) - O - X'$ where X, $X' = H$, CH_3 can exist in two conformational states, namely, the syn periplanar and the anti periplanar, Figure 3-1 [91, 92, 93, 94].



syn periplanar **anti periplanar**

Figure 3-1: Syn and Anti-Periplanar Configurations of Methyl Acetate

The molecules are almost exclusively found in the syn conformation [94, 91, 93] where the $O - X'$ bond vector is periplanar to the $C = O$ bond of the carbonyl functionality; this gives essentially a *planar* heavy-atom skeleton in the syn conformation [91, 92, 93, 94]. This planarity will be important in interpreting the molecular-level

picture of ester interfaces to be presented later. In the case where the substituent is a methyl group, the preferred conformation of the methyl of the $O - CH_3$ group has the hydrogens staggered with respect to the carbonyl $C = O$ bond [95]. Moreover, the preferred syn conformation is rationalized through resonance arguments; with the methyl group syn to the carbonyl oxygen, there is charge transfer from teh ether oxygen to the carbonyl oxygen, thus leading to the stability of this conformer Figure 3-2 [95].

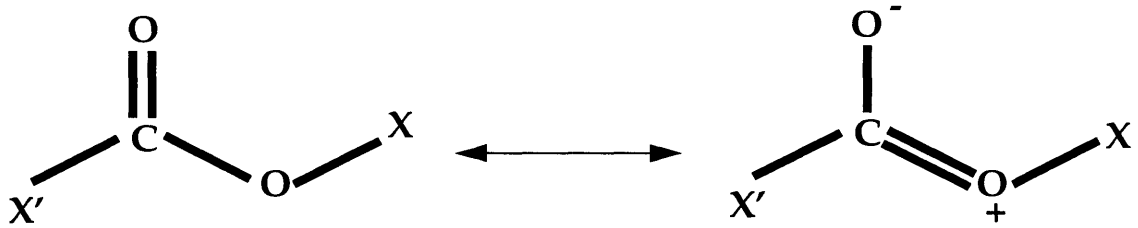


Figure 3-2: Resonance Structures for the Ester Carbonyl Functionality

Furthermore, from infrared intensities and *ab initio* calculations, there is reported a hyperconjugation effect in the acetyl group; negative electronic charge is pumped from the methyl group towards the π system of the adjacent carbonyl group [96]. Low-temperature matrix infrared spectroscopic measurements indicate that the energy difference between the anti and syn conformers is about $8.5 \pm 1.0\, kcal\, mol^{-1}$ [91]. This is for methyl acetate specifically. A barrier height for crossing from one state to another was not directly measured for methyl acetate, but from the value for methyl formate, it is estimated to be in the range from 10-15 $kcal\, mol^{-1}$ [91, 95]. As for computed values for the conformational energetics, the most sophisticated calculations are the ab initio simulations of methyl acetate using Hartree-Fock theory with correction for electron correlation by Wiberg *et al* [95]. The authors find the energy difference between the E (anti periplanar) and Z (syn periplanar) conformers to be 8.6 $kcal\, mol^{-1}$ which is in excellent agreement with the matrix IR data. The authors report geometry optimizations at the Hartree-Fock 6-31G* level followed by electron correlation correction via Moller-Plesset perturbation theory through third

order (MP3) [95].

Furthermore, the experimental and theoretical data seem to indicate that the torsional profile is twofold, with only two minima and a large barrier to rotation. This is to be kept in mind when comparing various forcefields.

Figure 3-3 shows graphically the torsional profile about the central axis of the methyl acetate molecule (torsion about the $C_{sp2} - O$ bond). Also shown is a curve representing semi-empirical calculations performed as part of this research. These calculations were performed with the AM1 model Hamiltonian for MNDO theory. It is evident from this plot and Table 3.1 that the CFF91 potential faithfully reproduces experimental and calculated profiles. Note that the OPLS potential, although fitted to ester simulations, uses a somewhat smaller value for the cis-trans energy difference (ΔE=5 kcal/mole) as well as a lower value for the barrier heights (8.9kcal/mole), for which the experimental values are 8.5 kcal/mole and 13.0 kcal/mole, respectively [67]. However, the barrier height is still large enough to remain valid at the temperatures we are interested in; that is, the energetic barrier to crossing from one conformer to the other is greater than the thermal energy, kT. The MM3 value is also within acceptable bounds with respect to the experimental data, but in essence, the MM3 model is a class II forcefield much like the CFF91. A comparison of the absolute values would lead one to accept the CFF91 forcefield over the MM3. For this reason, the MM3 forcefield was discarded from further consideration as the CFF91 is an equivalent (if not better) forcefield of the same genre. The AMBER forcefield is clearly unacceptable, as not only does it err on the prediction of conformational energy differences, it misrepresents the form of the torsional energy surface.

Based on the above considerations, the OPLS and CFF91 forcefield models for methyl acetate were selected to represent the ester functionality for the present work. As the next step, the ability of these models to reproduce bulk liquid energetics was studied as a prelude to applying a forcefield model to the study of the ester interface.

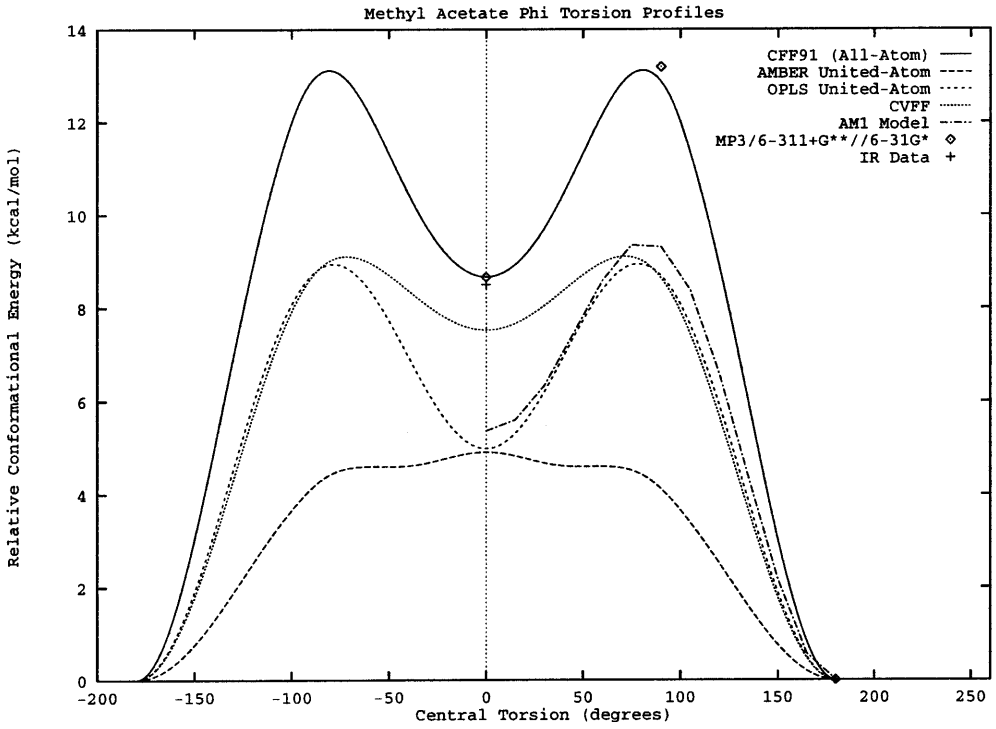


Figure 3-3: Methyl Acetate Phi Torsion Profile: Forcefield Dependence

Table 3.1: Methyl Acetate: Cis-Trans Energy Differences and Torsional Barriers

$\Delta E_{cis,trans}$ (Kcal/mole)	Source
5	AMBER Forcefield (DISCOVER) constant dielectric, $\mu = 1$ no 1-4 non-bonded interaction scaling
2.6	AMBER Forcefield (DISCOVER) distance dependent dielectric, $\mu = 4 * r$ 1-4 non-bonded interaction scaling by 0.5
5.31	MOPAC (AM1 Hamiltonian)
7.21	SCF-Hartree Fock (TURBOMOLE) STO-3G basis set
8.5±1.0	Blom and Gunthard matrix IR spectra
8.71	Allinger MM3
7.8	Allinger MM2
8.6	Wiberg and Laidig SCF-Hartree Fock MP3/6-311+G**//6-31G*
8.67	CFF91 Forcefield (DISCOVER)
4.98	OPLS

3.2 Bulk Liquid Energetics I. Computational Methods

Along with structural and geometrical properties, thermodynamic observables are also of interest. A fundamental measure of the soundness of a potential is its ability to reproduce experimental thermophysical properties, although in practicality, researchers have employed other tests such as reproducibility of vibrational frequencies, crystal structures, zero-point lattice energies, phonon frequencies, and dispersion curves (crystals). In the final analysis, as there are a myriad number of observable properties which can be used to develop a potential model, one usually chooses a subset typical of the properties of interest for a given study. From this perspective, no one potential model can give fully accurate reproduction of all properties for a given class of materials. One hopes at best that relevant observables are reproduced to within acceptable error. With this in mind, in the present work, the enthalpy of vaporization to the ideal gas is calculated as a test of the validity of the potential. The enthalpy of vaporization is directly connected to the internal energy of the liquid, which in turn is reflective of the energetic interaction between the constituent molecular entities. It has been reported that details of structural and dynamic properties of chain molecules at interfaces depend on accurate descriptions of the polymer-polymer interaction; thus, capturing the nature of the intermolecular interaction as accurately as possible is of prime importance [97]. For the purpose of validating a forcefield, bulk simulations for the flexible OPLS model and the CFF91 models were performed.

Enthalpies of vaporization to the ideal gas were computed from the following equation [67]:

$$H = E_{intra}(gas) - [E_{intra}(l) + E_{inter}(l)] + RT \tag{3.38}$$

$E_{intra}(gas)$ is the gas-phase intramolecular potential energy, $E_{intra}(l)$ the liquid-phase intramolecular potential energy,$E_{inter}(l)$ the liquid-phase inter- molecular potential energy, and RT is the standard ideal gas enthalpy. This view considers the kinetic and zero-point vibrational eneries of the liquid and vapor to be equivalent [80]. The

gas-phase quantities were computed using a Hybrid Monte Carlo algorithm. This involved running a single-molecule molecular dynamics sequence for a large number of Monte Carlo steps. At the beginning of each Monte Carlo cycle, the particle velocities are re-selected from a Boltzmann distribution of velocities at the desired temperature (in the present simulations, T=298.0K). A single- molecule simulation is justified in this case as the vapor is essentially at low pressure, low density (physically). Thus, in the vapor phase, any given molecule 'sees' no other molecule and the only interactions are intramolecular.

At this point note that the CFF91 potential is an all-atom model, while the OPLS model specific for methyl acetate is a united-atom model. This is an important distinction in terms of the computational tractability of simulations, particularly when system size increases as with polymer simulations. To perform actual test simulations using the CFF91 model, a multiple time step algorithm was implemented. This is briefly described in the next section. Next, a brief discussion of the Hybrid Monte Carlo algorithm follows. This technique is applied in the present work as a tool in parameter fitting for internals distributions, and more relevant to the calculation of enthalpy of vaporization calculations, for determining single-molecule energetic properties. The ensuing section describes the algorithms used to compute the electrostatic interactions, a need arising from the partial charges assigned as part of the forcefield model. Finally, we describe a procedure for parameterizing an intramolecular bond angle and out- of-plane potential function for the OPLS model. This is necessary since the original OPLS model is rigid, constraining bond lengths and bond angles. As we are dealing with heavy-atom motions (soft modes) characterized by force constants lower than that for C-H stretching (low frequency modes), we cannot represent these motions as rigid (in a time-averaged sense) over the time-scales we probe during an MD simulation. Thus, before presenting bulk liquid simulation details and results, the parameterization of bond-angle and out-of-plane potentials to be merged with the OPLS parameter set is presented.

3.2.1 Reversible Multiple Time Step Molecular Dynamics (r-Respa)

As mentioned above, the CFF91 potential is an all-atom forcefield which models all atomic sites including light atoms such as hydrogen. The low mass of hydrogens leads to difficulties as the dynamics associated with these light particles is much more rapid compared to the dynamics of heavier atoms. More specifically, as the motions involving hydrogens (or light particles in general) are of high frequency, the timestep used in the numerical integration of the equations of motion must be small enough so as to capture the full dynamics (e.g., one must sample sufficiently within the period of the motion). Furthermore, in general, intramolecular dynamics are inherently of differing time-scales. For example, bondstretching is more rapid than bond-angle bending which in turn is more rapid than dihedral motion, which ultimately occurs on a shorter time scale compared to overall molecular translation and rotation [98]. The underlying problem this discussion suggests is that in standard molecular dynamics simulations of systems with dynamical modes of widely differing time-scales, the fast (hard) modes limit the magnitude of the time step that can be used for the numerical integration of the equations of motion, or in other words, propagating the trajectory through phase space. This can be computationally costly particularly when the slow (soft) modes are of primary interest.

Several approaches to circumvent this obstacle are widely used [98]. Constraints are applied to bond-stretching and bond angle bending modes so that a large timestep can be used for integrating the equations of motion for the remaining degrees of freedom. However, bond angle constraints are difficult to apply for macromolecular systems (small molecules are sufficiently treated with bond-angle and bond-length constraints) [98]. A second method reduces the system to internal coordinate space and allows only the dihedral modes to propagate [98]. A further 'trick' is to increase the mass of light atoms (hydrogen) to allow a larger time step. This, of course, will perturb the dynamics away from the true dynamics, but in cases where dynamical information is not of primary interest (diffusion constants), this approach is adequate

[98].

The approach applied in the present case is a multiple-time step technique developed by Berne and coworkers termed r-RESPA (reversible Reference System Propagator Algorithm); this is a reversible integrator method based upon the multiple-scale algorithms developed by Berne and coworkers [99, 100, 101, 102, 103, 104]. This approach includes all degrees of freedom, but uses different time steps to propagate the differing modes inherent in the system. It is based on Trotter factorization of the system Liouvillean and a carefully chosen separation of the forces [99]. The essence of the method is to integrate the faster motions more frequently than slower ones. At regular intervals, the slower degrees of freedom are updated. This approach yields a time reversible propagator, a requirement for the present work as this molecular dynamics scheme will be incorporated into a Hybrid Monte Carlo algorithm which requires strict reversibility in the dynamics to meet the requirement of detailed balance (other multiple timestep algorithms are available, but suffer from irreversibility [100, 101, 102]). Furthermore, this approach is applicable to extended hamiltonians such as Nose thermostats, thus allowing molecular dynamics in the canonical ensemble [99].

In the following subsection is presented in brief the formalism of a reversible multiple timestep propagator after which the current implementation is described.

Reversible Multiple Time Step Propagator (r-RESPA) Formalism

As molecular dynamics is fundamentally probing the underlying phase (configuration) space of a system, the starting point for generating a propagator is the classical Liouville operator. The Liouville operator for a system of N degrees of freedom in Cartesian coordinates is [99, 98]:

$$iL = \sum_{i=1}^{N} (\dot{x}_i \frac{\partial}{\partial x_i} + F_i \frac{\partial}{\partial p_i}) = [\cdots, H] \tag{3.39}$$

where x_i and p_i are the position and conjugate momentum for degree of freedom i, $\dot{x}_i$ is the time derivative of x_i, F_i is the force on degree of freedom i, and the

brackets denote the Poisson bracket (not the quantum mechanical commutator!) of the system. The state of the system is represented as $\Gamma[x_i, p_i]$ in terms of degree of freedom coordinates and momenta. Note that the Liouville operator is a linear Hermitian operator on the space of square integrable functions of Γ. The classical propagator is then [99, 98]

$$U(\Delta t) = e^{iL\Delta t} \tag{3.40}$$

so that starting with a system at state $\Gamma(0)$ at time zero, the state at a time interval Δt later is:

$$\Gamma(\Delta t) = U(\Delta t)\Gamma(0) \tag{3.41}$$

Note here that since the classical Liouville operator is self-adjoint, its propagator is unitary:

$$U(-\Delta t) = U^{-1}(\Delta t) \tag{3.42}$$

and the time development of the system throughout phase space is reversible [99, 98].

Also note here that the propagator can be expressed as a series in $iL\Delta t$, a representation which is useful to the application of Trotter's theorem [105] later [99, 98]:

$$U(\Delta t) = \sum_{n=0}^{\infty} \frac{(\Delta t)^n}{n!}(iL)^n \tag{3.43}$$

Now, the Liouville operator is decomposed (the manner of decomposition is irrelevant for the present purposes; it will be discussed further below) into two parts as:

$$iL = iL_1 + iL_2 \tag{3.44}$$

Applying Trotter factorization:

$$e^{i(L_1+L_2)} = [\, e^{i(L_1+L_2)\Delta t/P}\,]^P \tag{3.45}$$

$$= [\, e^{iL_1(\Delta t/2)}\, e^{iL_2\Delta t}\, e^{iL_1(\Delta t/2)}]^P \tag{3.46}$$

$$+O((\Delta t)^3) \tag{3.47}$$

Here the error is indicated as third order in the timestep.

The second term $e^{iL_2\Delta t}$ can be further factorized to yield:

$$e^{iL_2\Delta t} = (\, e^{iL_2\Delta\tau}\,)^n + O((\Delta t)^3) \tag{3.48}$$

where $\Delta\tau = \Delta t/n$. The smaller timestep $\Delta\tau$ and the integer value n are chosen so as to guarantee a stable dynamics for the system. From the above progression, the concept of multiple timesteps is quite evident and appears as a natural consequence of the factorization of the propagator following the decomposition of the Liouville operator. Thus, in a system composed of several time scales, a proper splitting of the Liouvillean will allow one to effectively sample the modes of interest by separating them out. One must now consider the decomposition of the Liouville operator. However, the above derivation is first generalized for a splitting of the operator into more than two components.

If the system Liouvillian is a sum of individual operators of the form

$$iL = \sum_{k=1}^{n} iL_k \tag{3.49}$$

the propagator is expressed as:

$$e^{iL\Delta t} = [[\prod_{k=1}^{n-1} U_k(\frac{\Delta t}{2}]U_n(\Delta t)[\prod_{k=1}^{n} U_{n-k}(\frac{\Delta t}{2})]]^P + O(\frac{(\Delta t)^3}{P^2}) \tag{3.50}$$

where $U_k(h) = e^{iL_k h}$. This is a continuous representation as the propagator is applied P times. One defines a discrete time propagator as:

$$G(\Delta t) = [\prod_{k=1}^{n-1} U_k(\frac{\Delta t}{2}]U_n(\Delta t)[\prod_{k=1}^{n} U_{n-k}(\frac{\Delta t}{2}] \tag{3.51}$$

This leads to a time-reversible propagator given that each L_k is Hermitian so that for each $U_k(\Delta t)$

$$U_k^{\dagger}(\Delta t) = U_k^{-1}(\Delta t) = U_k(-\Delta t) \tag{3.52}$$

Thus, $G(\Delta t)\, G(-\Delta t) = 1$ and $G(\Delta t)$ gives rise to reversible dynamics [99] .

In applying the r-RESPA algorithm to complex flexible molecules, it is not convenient to separate the fast and slow (hard and soft) degrees of freedom as they are defined in terms of internal coordinates (bond length, bond angles) whereas the dynamics is computed exclusively in Cartesian coordinate space. Consequently, the natural decomposition scheme is to separate the Liouvillian according to type of force, specifically, hard forces, F_h, and soft forces, F_s [98]. For the present work, the hard forces are considered to be bond-stretching, bond-angle bending, and out-of-plane angle bending forces. The remaining forces – dispersion forces, electrostatic forces, and dihedral forces – are lumped as soft forces. The next section describes the present implementation of r-RESPA.

Reversible Multiple Timestep Algorithm for Nose Dynamics: Application to CFF91 Forcefield Model

As the bulk simulations for the methyl acetate liquid energetics are performed at constant temperature and density (canonical ensemble MD), a r-RESPA algorithm for Nose dynamics is developed. The Nose equations of motion are [99]:

$$\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{ms^2} \tag{3.53}$$

$$\dot{p} = -\frac{\partial H}{\partial x} = -\frac{\partial V}{\partial x} \tag{3.54}$$

$$\dot{s} = \frac{\partial H}{\partial p_s} = \frac{p_s}{Q} \tag{3.55}$$

$$\dot{p_s} = -\frac{\partial H}{\partial s} = \frac{1}{s}[\sum \frac{p^2}{ms^2} - (3N+1)kT] \tag{3.56}$$

These equations are generated from the Hamiltonian [99]

$$H = \sum \frac{p^2}{ms^2} + (3N+1)kT \ln s + V(x) + \frac{p_s^2}{2Q} \tag{3.57}$$

Hoover introduced a time rescaling $dt \rightarrow dt/s$ and the change of variables $p \rightarrow p/s$, $p_s \rightarrow p_s/s$, along with the dynamical variable $\eta = \ln s$ which gives the Nose-Hoover equations: [99]

$$\dot{x} = \frac{p}{m} \tag{3.58}$$

$$\dot{p} = -\frac{\partial V}{\partial x} - \frac{p_\eta}{Q} p \tag{3.59}$$

$$\dot{\eta} = \frac{p_\eta}{Q} \tag{3.60}$$

$$\dot{p_\eta} = \sum \frac{p^2}{m} - 3NkT \tag{3.61}$$

To derive a reversible propagator, one turns to the following Liouvillian: [99]

$$iL = \dot{x}\frac{\partial}{\partial x} + F_h(x)\frac{\partial}{\partial p} + F_s(x)\frac{\partial}{\partial p} - \dot{\eta}p\frac{\partial}{\partial p} + \dot{\eta}\frac{\partial}{\partial \eta} + F_\eta(p)\frac{\partial}{\partial p_\eta} \tag{3.62}$$

where

$$F_\eta(p) = \sum \frac{p^2}{m} - 3NkT \tag{3.63}$$

and equivalently

$$F_\eta(\dot{x}) = \sum m\dot{x}^2 - 3NkT \tag{3.64}$$

Note that the total force on the η coordinate is a function of the momenta of the remaining degrees of freedom.

For the present case, the Liouville operator is separated as :

$$iL = iL_1 + iL_2 \tag{3.65}$$

where

$$iL_1 = \dot{\eta}\frac{\partial}{\partial \eta} - \dot{\eta}p\frac{\partial}{\partial p} + F_\eta(p)\frac{\partial}{\partial p_\eta} + F_s(x)\frac{\partial}{\partial p} \tag{3.66}$$

and

$$iL_2 = \dot{x}\frac{\partial}{\partial x} + F_h(x)\frac{\partial}{\partial p} \tag{3.67}$$

Thus, the discrete time propagator, $G(\Delta t)$ becomes: [99]

$$G(\Delta t) = e^{\frac{\Delta t}{2}F_\eta(p)\frac{\partial}{\partial p_\eta}}\, e^{\Delta t/4F_s(x)\frac{\partial}{\partial p}}\, e^{\frac{-\Delta t}{2}\dot{\eta}p\frac{\partial}{\partial p}}\, e^{\Delta t/4F_s(x)\frac{\partial}{\partial p}}\, e^{\frac{\Delta t}{2}\dot{\eta}\frac{\partial}{\partial \eta}}\, e^{\Delta t iL_2} \tag{3.68}$$

$$\times\, e^{\frac{\Delta t}{2}\dot{\eta}\frac{\partial}{\partial \eta}}\, e^{\Delta t/4F_s(x)\frac{\partial}{\partial p}}\, e^{\frac{-\Delta t}{2}\dot{\eta}p\frac{\partial}{\partial p}}\, e^{\Delta t/4F_s(x)\frac{\partial}{\partial p}}\, e^{\frac{\Delta t}{2}F_\eta(p)\frac{\partial}{\partial p_\eta}} \tag{3.69}$$

We can further decompose the second operator, iL_2 to give two further operators:

$$iL_{2,a} = F_h(x)\frac{\partial}{\partial p} \tag{3.70}$$

$$iL_{2,b} = \dot{x}\frac{\partial}{\partial x} \tag{3.71}$$

With the above two operators, an intermediate discrete time propagator can be defined as:

$$G_{int}(\Delta t) = [e^{iL_{2,a}\delta t/2}\, e^{iL_{2,b}\delta t}\, e^{iL_{2,a}\delta t/2}]^n \tag{3.72}$$

where $\delta t = \Delta t/n$.

The significance of the intermediate discrete time propagator is that it represents the standard velocity Verlet integrator used widely in conventional molecular dynamics simulations. [99]. Thus, embedded in the total propagator for Nose dynamics is the Verlet algorithm. Furthermore, it is the hard forces which contribute to the velocity Verlet integration. The hard forces serve as generating a reference system to which the soft forces are a correction updated every n time steps of the integration for the reference system. Applying the total discrete time propagator to an initial state $[x(0), p(0), \eta(0), p_\eta(0)]$ yields reversible Nose dynamics for constant temperature molecular dynamics.

Performance of r-RESPA: Integrator Stability and Speedup

In the present implementation of r-RESPA, the hard forces are considered to arise from bond stretching, bond-angle bending, and out-of-plane (improper torsion) motion. This is in keeping with reported approaches in the literature as well. Also, based on the time periods of the relevant motions, it is evident that the faster modes are generally the intramolecular modes excluding the torsions. The stability of the integrator is measured by the fluctuation [98]:

$$\Delta E^{rms} = \frac{\Delta E_{total}}{\Delta E_{kinetic}} \tag{3.73}$$

where ΔE_{total} is the root-mean-square (rms) fluctuation of the total energy (this is the extended Hamiltonian in cases where baths are used) and $\Delta E_{kinetic}$ is the rms fluctuation of the kinetic energy. The criterion for a stable integrator is that $\Delta E^{rms} \leq 0.1$, although some would claim to use a looser bound [98].

To study the stability of the algorithm and select a working timestep for the bulk methyl acetate simulations using the all-atom CFF91 forcefield, short simulations of a

bulk system of methyl acetate were performed. Canonical ensemble (NVT) molecular dynamics using a Nose thermostat at the experimental density of $0.927 gr/cm^3$ were performed in a cubic simulation box with periodic boundary conditions and minimum image convention for non-bond interactions. The system consists of 216 molecules in a box of dimensions $L_x = L_y = L_z = 3.06nm$. The non-bond interactions, dispersion and electrostatic, are computed out to a spherical cutoff value of 1.2 nanometers. No switching or smoothing function for the dispersion interactions was applied. Electrostatic interactions were computed with a reaction field method to be discussed below.

Figure 3-4 shows a plot of the ΔE^{rms} as a function of the larger timestep, Δt, for a series of n values (n essentially giving the size of the smaller timestep). Also shown is a curve representing conventional molecular dynamics. From this data it is evident that the multiple time step algorithm allows a definitively larger timestep for a given integrator stability as demonstrated by the value of ΔE^{rms}. More specifically, for a ΔE^{rms} within 10 percent, the time step increases from 0.001 to roughly 0.004, a fourfold speedup. Furthermore, the effect of the smaller step size on stability reflects the same behavior, as the stability for the larger n values is greater than for lesser values. Also, the data seems to indicate no great difference between $n = 5$ and $n = 10$ for the smaller time steps, with a slight deviation at larger values. This is all in keeping with the fact that a 1 fs timestep is the limiting case for the dynamics for the hard modes as evidenced by the conventional MD runs.

Finally, Figure 3-5 shows the time requirements for a one picosecond simulation run using a series of large time steps with various n values. Again, for acceptable integrator stability at large time steps, the r-RESPA algorithm yields appreciable speedup over conventional MD.

One may argue that simulations with constraints are just as reasonable for MD simulations. This is true; however, for forcefields not parameterized with constraints, it is necessary to account explicitly for the faster motions. Furthermore, it may be that the fast modes are of primary interest to the scientist in terms of validating a forcefield through frequency calculations. In this case, the utility of multiple time

step methods is unquestionable.

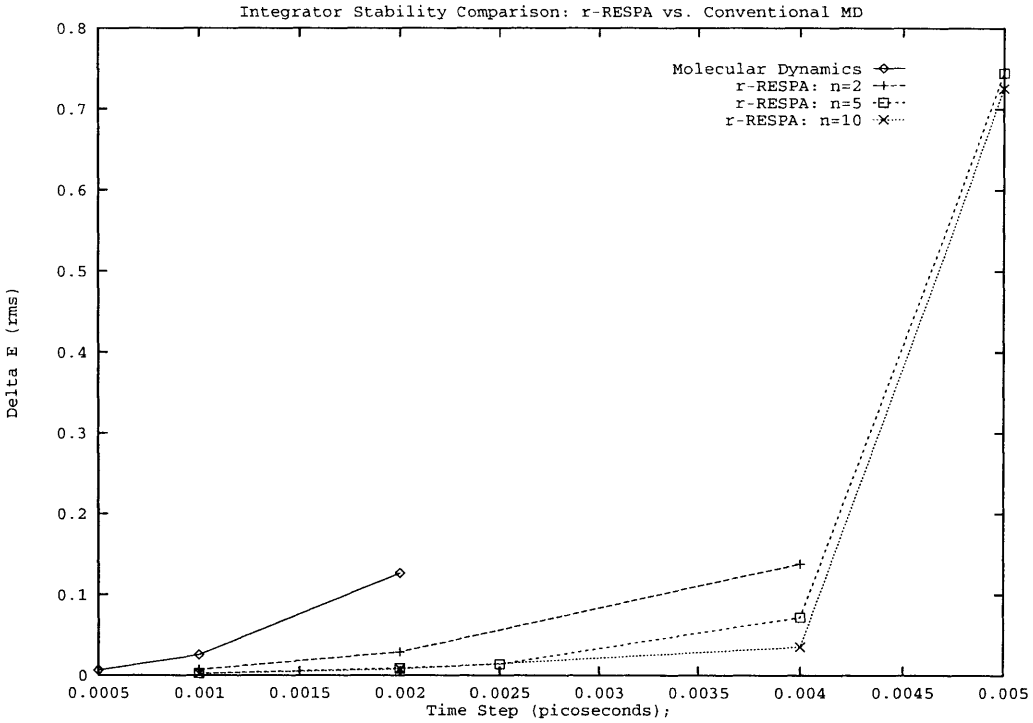


Figure 3-4: Stability Comparison: r-RESPA versus Conventional Molecular Dynamics

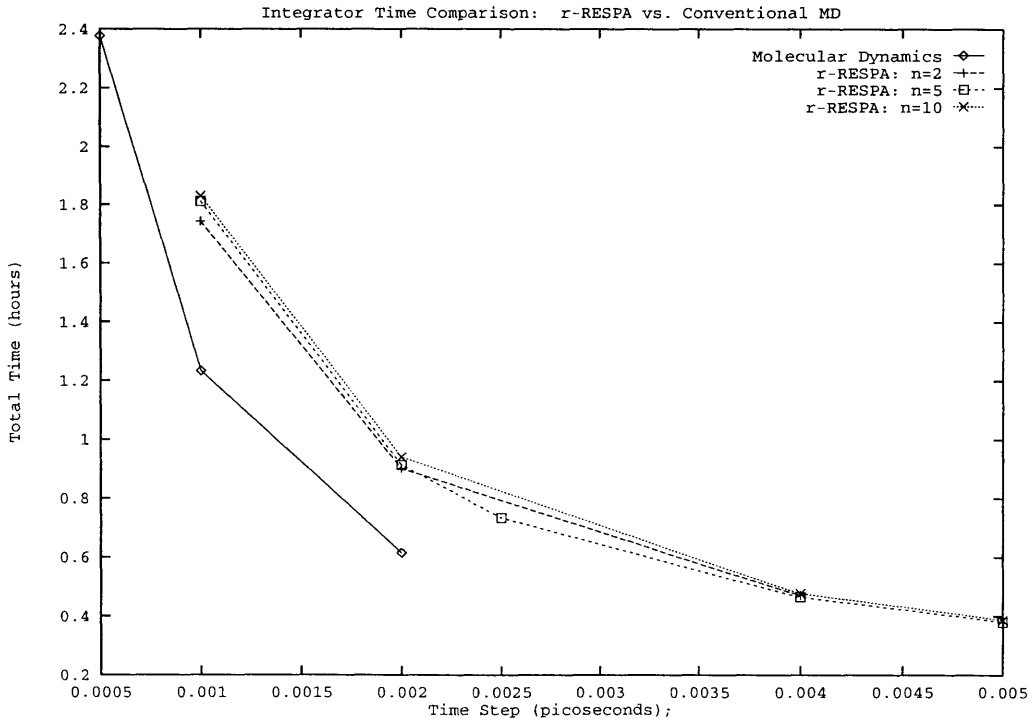


Figure 3-5: Time Comparison: r-RESPA versus Conventional Molecular Dynamics

3.2.2 Hybrid Monte Carlo: Formalism and Implementation

The Hybrid Monte Carlo (HMC) method is used for determining the single-molecule energetic components. In this section, a brief discussion of the formalism and current implementation of HMC is given. For details not given here, the reader is referred to the literature cited [106, 107, 108]. The method originally was applied to quantum chromodynamics simulations of the behavior of quarks in nucleons [106]. Its fundamental use is in effectively sampling the phase space representing a many-body system. The *raison d'etre* for the method is to combine the best features of Monte Carlo and Molecular Dynamics simulations to arrive at an algorithm which samples more of phase space efficiently; that is, the algorithm allows the system to escape limit cycles and attractor points encountered via MD sampling by the action of the random walk nature of the MC rejection/acceptance test [106].

To discuss the theoretical foundations of HMC, one begins with the Hamiltonian of a N-particle system with interaction potential V (V is a function of position) [106]:

$$H(x,p) = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_{i}^{N} V(x_i) \tag{3.74}$$

where m_i, x_i, and p_i are the mass, position, and momentum of particle i. The equations of motion derived from the Hamiltonian are:

$$\dot{x}_i = \frac{p_i}{m_i} \tag{3.75}$$

$$\dot{p}_i = -\frac{\partial V(x_i)}{\partial x_i} \tag{3.76}$$

Standard molecular dynamics simulation involves integration of the equations of motion with a suitable integrator with initial conditions on positions and momenta (velocities) given as input. This integration generates the phase space trajectory; note, however, that the trajectory is dependent on the initial configuration (initial conditions) and the ergodicity of the trajectory will be compromised if it so happens that the initial phase point is near a local minimum , limit cycle, or attractor [106].

The Monte Carlo algorithm generates phase space trajectories which are random walks. The probability of a phase point (any given set of positions and momenta) is the Boltzmann probability [106]:

$$P(x,p)\,dx\,dp = \frac{1}{Z}\,dx\,dp\,e^{-\beta H(x,p)} \tag{3.77}$$

where $\beta = 1/kT$ and the convention adopted is that $dx = \prod_{i=1}^{N} dx_i$. Starting from an initial configuration, a Markov chain of states is generated through random changes in the positions and/or momenta; the probability of accepting a new state (configuration) is given by the Metropolis probability [106]:

$$P_{acc} = min(1, e^{-\beta\Delta H}) \tag{3.78}$$

where

$$\Delta H = H(x^{'}p^{'}) - H(x,p) \tag{3.79}$$

The acceptance probability, P_{acc} , satisfies detailed balance which allows the algorithm to generate states with the desired equilibrium probability distribution. In theory, the number of coordinates or momenta that are changed simultaneously is irrelevant; also the order in which the changes are made is unimportant. However, in practice, the convergence (the rate at which the configurations decorrelate) and the acceptance rate rely heavily on the above parameters. If changes are made to all particles, the new configurations will have a low probability due to a large unfavorable energy change (thus low acceptance rate). If many small changes are made, the acceptance rate will increase, but the decorrelation will take longer and phase space will be sampled much more slowly [106].

The Hybrid Monte Carlo method combines the advantages of both MD (fast phase space sampling) and MC (sampling of a larger portion of phase space through escaping from local equilibrium points). The algorithm is in essence Monte Carlo with MD 'sandwiched' in between. A single HMC step involves first reassigning the velocities of all particles with values sampled from the relevant equilibrium distribution $P(p_i)$ =>

$exp(-\beta p_i^2/2m_i)$. Next, a reversible integrator is used to propagate the system (MD). Finally, after L steps of MD, the Metropolis acceptance criteria is applied and the final state is rejected or accepted. These steps are then repeated for the necessary number of MC steps.

Note that it is essential to have a reversible integrator to meet the detailed balance requirement. Figure 3-6 shows a schematic of the HMC algorithm [106].

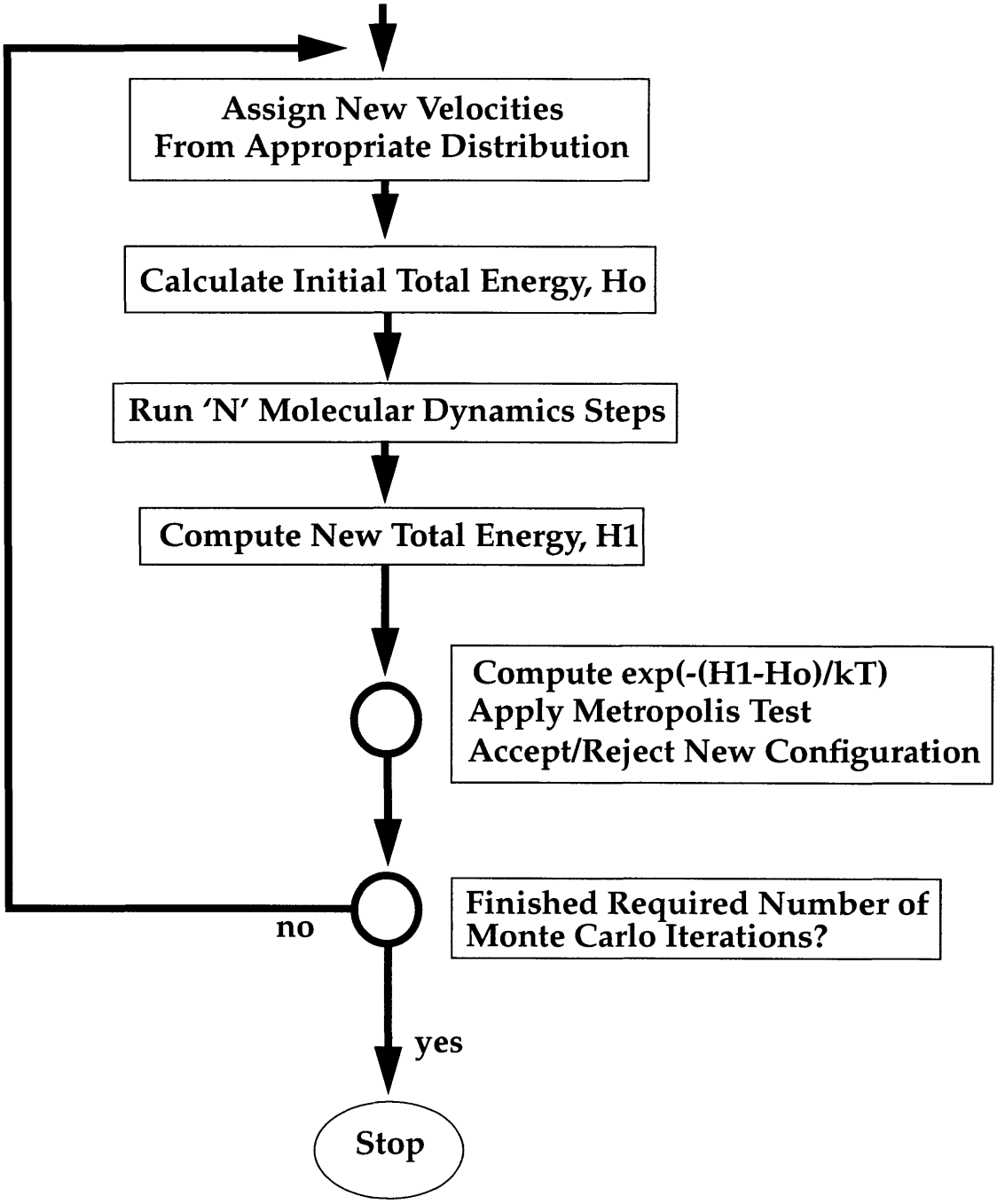


Figure 3-6: Hybrid Monte Carlo Algorithm

3.2.3 Long-Range Interactions: Electrostatic Energies and Forces in Molecular Simulation

Ewald Summation

Polar systems and systems with ionic species pose a difficult problem for molecular dynamics simulations. The electrostatic Hamiltonian,

$$H = \frac{1}{2}\sum_{\mathbf{n}}\left[\sum_{i=1}^{N}\sum_{j=1}^{N}\phi(r_{ij} + L\mathbf{n})\right] \tag{3.80}$$

represents a sum of the Coulomb potential over all the simple cubic lattice points with integer coordinates $\mathbf{n} = (l, m, n)$, th e $\mathbf{n} = \mathbf{0}$ terms omitted [109]. As the Coulomb potential is long-ranged and fails to satisfy

$$|\phi(r)| \leq A|r^{-3}| \tag{3.81}$$

the lattice sum in the Hamiltonian is not absolutely convergent, but rather conditionally convergent, depending on the nature of the summation [109]. The difficulty thus is that to achieve convergence, and thus an accurate accounting of all the interactions, one must include a large number of terms in the summation. From a simulation standpoint, this is costly as the pairwise interactions and forces consume the most time during the dynamics. Over the decades, researchers have developed approaches to circumvent this difficulty. The Ewald summation technique, first developed for simulations of crystalline solids, is widely used in condensed matter simulations ranging from crystals to proteins and organic solutions [13].

The Ewald summation technique essentially decomposes the Hamiltonian into short and long range components which are summed in real and reciprocal (Fourier) space, respectively. An associated convergence factor converts the conditionally convergent summation to an absolutely convergent one [109]. For a given system of charges, a screening charge distribution is introduced about each initial point charge. To retrieve the original system, a canceling charge distribution is also introduced. The interactions between the initial and screening charge distributions are summed in real

space (explicitly), and the interactions of the canceling distribution are summed in fourier space. One can with proper selection of parameters, restrict the real space summation to the original simulation box, and carry out the Fourier space summation over a minimal number of wavevectors(reciprocal lattice vectors) [13]. The Coulomb potential representing the Ewald summation approach is [13]:

$$
\begin{aligned}
U^{Coulomb} &= \frac{1}{L}\sum_{i,j}^{N}\sum_{a,b} q_i^a q_j^b \sum_{\mathbf{n}} \left(\frac{erfc(\alpha \left| \frac{\mathbf{r}_{ij}}{|\mathbf{L}|} + \mathbf{n} \right|)}{\left| \frac{\mathbf{r}_{ij}}{|\mathbf{L}|} + \mathbf{n} \right|} \right) \\
&+ \frac{1}{2L}\sum_{i=1}^{N}\sum_{a} (q_i^a)^2 \sum_{\mathbf{n}\neq\mathbf{0}} \left(\frac{erfc(\alpha|\mathbf{n}|)}{|\mathbf{n}|} \right) \\
&+ \frac{1}{2L}\sum_{i=1}^{N}\sum_{a\neq b} q_i^a q_i^b \sum_{\mathbf{n}} \left(\frac{erfc(\alpha \left| \frac{\mathbf{r}_{ij}}{|\mathbf{L}|} + \mathbf{n} \right|)}{\left| \frac{\mathbf{r}_{ij}}{|\mathbf{L}|} + \mathbf{n} \right|} - \frac{1}{\left| \frac{\mathbf{r}_{ii}^{ab}}{|\mathbf{L}|} \right|} \right) \\
&+ \frac{1}{2L}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{a,b} q_i^a q_j^b \sum_{\mathbf{n}\neq\mathbf{0}} \frac{1}{\pi|\mathbf{n}|^2} exp\left(\frac{-\pi^2|\mathbf{n}|^2}{\alpha^2} + 2\pi i \mathbf{n}\cdot\frac{\mathbf{r}_{ij}^{ab}}{|\mathbf{L}|} \right) \\
&- \frac{\alpha}{\sqrt{\pi L}}\sum_{i=1}^{N}\sum_{a} (q_i^a)^2 + \frac{2\pi}{3L^3}\sum_{x} N_x \mu_x^2
\end{aligned}
\tag{3.82}
$$

The parameter α results from the conversion to an absolutely convergent sum and controls the relative rate at which the two sums converge [13]. Increasing its value causes the direct space summation to converge more rapidly at the expense of a slower convergence of the reciprocal sum. Furthermore, if one considers the screening charge cloud invoked about each initial point charge, the α parameter represents the width, or diffuseness, of the cloud [13]. With a large enough value, the only term which contributes to the direct space term is the n=0, or the conventional minimum image representation. Thus, in practice, α is chosen large enough so that the direct space sum is computed with the minimum image convention, and the summation in reciprocal space is truncated at an appropriate number of k-vectors at which the summation converges to within a present tolerance. Rigorously , the full Hamiltonian is independent of the value of α [13].

Reaction Field

The reaction field technique is a cutoff-based approach to evaluating long-range electrostatic interactions. The reaction field approach combines a cutoff methodology with a 'correction' due to material outside of the cutoff sphere via a continuum dielectric model of the substance. Within the cutoff sphere, interactions between charged sites are explicitly calculated. As there will develop a net dipole moment due to the arrangement of charges, there will be an interaction of this cavity dipole with the surrounding dielectric medium which in turn will induce a reaction field acting on the charges in the sphere [110]. Note that the original formulation of this method was developed by Onsager for point dipoles, but one can show that it is equally applicable to an arrangement of point charges [111]. Also note that the reaction field approach yields a Coulombic interaction for which both the potential and force are continuous at the cutoff separation. With this said, we model the Coulombic interaction via reaction field as follows:

$$U_{ij}^{coulomb}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_o}\left[\frac{1}{r_{ij}} + \frac{r_{ij}^2}{2R_c^3} - \frac{3}{2R_c}\right], \qquad r_{ij} \leq R_c \tag{3.83}$$

$$U_{ij}^{coulomb}(r_{ij}) = 0, \qquad r_{ij} > R_c \tag{3.84}$$

The cutoff radius, R_c is taken to be 1.2 nanometers.

One can derive the reaction field based on Gauss' Law applied to a distribution of charges, each surrounded by a sphere of uniform charge density which acts to neutralize the point charge. Refer to Figure 3-7 for the system geometry relevant for the following discussion.

Gauss' law for a charge density, $\rho(\mathbf{x})$ enclosed within a surface of arbitrary geom-

etry is:

$$\oint_S \mathbf{E} \cdot d\mathbf{a} = 4\pi \int_V \rho(\mathbf{x}) d^3x \tag{3.85}$$

where V is the volume enclosed by the surface, S.

The left-hand side represents the electric flux and the right-hand side integral is the TOTAL charge enclosed within the closed surface. The uniform charge density is:

$$\rho_o = \frac{-q}{\frac{4}{3}\pi r_c^3} \tag{3.86}$$

Consider two cases for computation of the electric field: 1. a Gaussian surface within the cutoff, r_c and 2. a Gaussian surface outside the cutoff. For the present analysis, the Gaussian surface is naturally a sphere.

For the region $r > r_c$, the enclosed charge is zero and there is no need for further consideration.

For the region $r < r_c$, the enclosed charge becomes a function of position as follows:

$$\begin{aligned} Q_{enc} &= q - \int_o^r \rho_o \, d\tau && (3.87)\\ &= q - \frac{3q}{4\pi} \int_0^r \frac{1}{r_c^3} 4\pi r^2 dr && (3.88)\\ &= q - \frac{3q}{r_c^3} \int_0^r r^2 dr && (3.89)\\ &= q - \frac{q}{r_c^3} r^3 && (3.90) \end{aligned}$$

Thus, Gauss' theorem becomes:

$$\oint \bar{E} \cdot d\bar{a} = \frac{1}{\epsilon_o} [q - \frac{q}{r_c^3} r^3] \tag{3.91}$$

Invoking spherical symmetry, the electric field is taken as always directed radially, and thus, along the normal to any differential surface area element; further, taking the electric field to be constant at a given radial position, the resulting simplification

is:

$$\oint E da = \frac{1}{\epsilon_o}[q - \frac{q}{r_c^3}r^3] \tag{3.92}$$

$$4\pi E r^2 = \frac{1}{\epsilon_o}[q - \frac{q}{r_c^3}r^3] \tag{3.93}$$

Thus,

$$E = \frac{1}{4\pi\epsilon_o}[\frac{q}{r^2} - \frac{q}{r_c^3}r] \tag{3.94}$$

Now, from elementary electrostatics, the force on a charge, j, is derived from the electric field, E, as:

$$F_j = q_j E \tag{3.95}$$

$$= \frac{1}{4\pi\epsilon_o}\left[\frac{q_i q_j}{r^2} - \frac{q_i q_j}{r_c^3}r\right] \tag{3.96}$$

With the force defined by the above equation, the electrostatic potential is easily obtained as:

$$F = -\frac{dU}{dr} \tag{3.97}$$

$$U(r) = -\int F dr \tag{3.98}$$

Integrating with the shown boundary conditions:

$$U(r) = \frac{-1}{4\pi\epsilon_o}\int[\frac{q_i q_j}{r^2} - \frac{q_i q_j r}{r_c^3}]dr + C \tag{3.99}$$

$$U = 0, r = r_c \tag{3.100}$$

yields the site-site reaction field equation as given above:

$$U_{ij}^{coulomb}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_o}\left[\frac{1}{r_{ij}} + \frac{r_{ij}^2}{2R_c^3} - \frac{3}{2R_c}\right], \qquad r_{ij} \leq R_c \tag{3.101}$$

$$U_{ij}^{coulomb}(r_{ij}) = 0, \qquad r_{ij} > R_c \tag{3.102}$$

In terms of dipole-dipole interactions, the reaction field potential corresponds to the modified dipole-dipole tensor [111]

$$T_{\alpha\beta}^{RF}(r) = T_{\alpha\beta}(r) + \frac{2(\epsilon_{RF}-1)}{2\epsilon_{RF}+1}\frac{\delta_{\alpha\beta}}{r_c^3} \qquad r < r_c \tag{3.103}$$

where

$$T_{\alpha\beta}(r) = \nabla_\alpha \nabla_\beta \frac{1}{r} \tag{3.104}$$

Taking an infinite dielectric constant, $\epsilon_{RF} = \infty$, gives:

$$T_{\alpha\beta}^{RF}(r) = T_{\alpha\beta}(r) + \frac{\delta_{\alpha\beta}}{r_c^3} \qquad r < r_c \tag{3.105}$$

Note that the reaction field method gives energies and forces which are continuous at the cutoff distance, r_c. Furthermore, the site-site formalism presented here is equivalent to the reaction field applied from a molecular perspective. The reader is referred to the literature for further discussion of this equivalence [111].

Finally, we note here that polarization effects are not considered explicitly in this work, as the models used for ester, water, and oligomers (in subsequent chapters) neglect direct treatment of molecular or atomic polarizability.

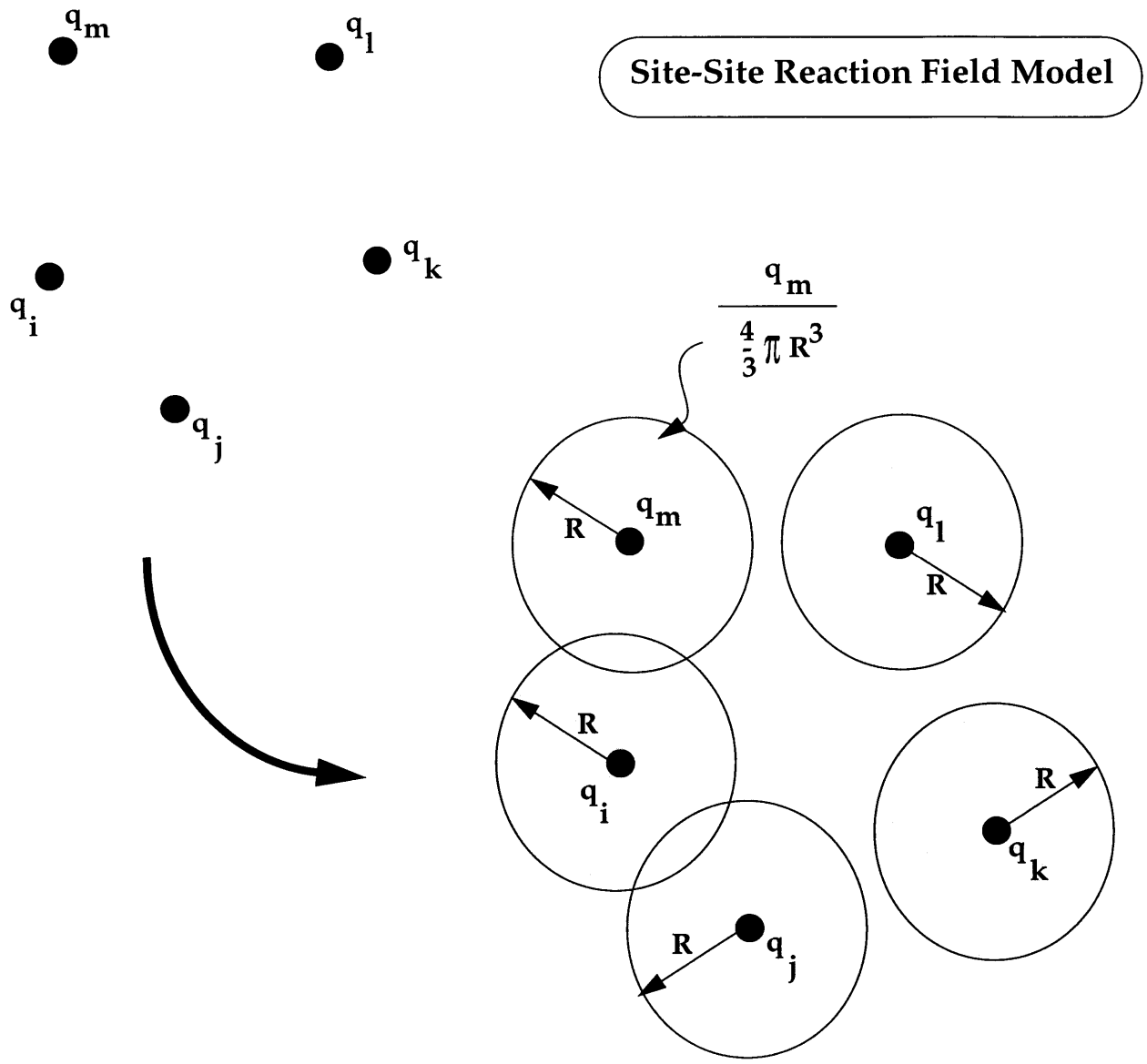


Figure 3-7: Site-Site Reaction Field Geometry

3.2.4 Bond Angle and Out-of-Plane Potentials

As mentioned above, molecular dynamics of the OPLS model requires explicit modeling of internal degrees of freedom. Specifically, to the original OPLS model need to be added bond angle and out-of-plane components. This is due to the fact that the intramolecular modes represented by these motions are observable on the time-scales probed by the MD technique. For example, consider the force constant for a typical $CH_2 - CH_2 - CH_2$ bond angle bending mode in a harmonic approximation:

$$U(\theta) = \frac{1}{2}K_\theta(\theta - \theta_o)^2 \tag{3.106}$$

$$k_{\theta, CH_2-CH_2-CH_2} = 527.2\frac{kJ}{mol\,rad^2} \tag{3.107}$$

Within the harmonic approximation, the fundamental frequency for an internal mode is given as:

$$\omega = \sqrt{\frac{k}{\mu}} \tag{3.108}$$

where k is an effective force constant and μ is an associated reduced mass for the mode of interest.

For a configuration of L particles, the reduced mass is defined by [66]:

$$\mu = \frac{\prod_i^L m_i}{\sum_i^L m_i} \tag{3.109}$$

One can compare the frequencies for bond-stretching and bond-angle bending modes for frequently-occurring bonds and angles to get a feel for the time scales involved.

Applying a harmonic approximation for small angle changes away from equilibrium gives (united-atom CH_2 units):

$$\mu = \frac{14^3}{14 + 14 + 14} \tag{3.110}$$

$$\omega_\theta = 6\,X\,10^{13} sec^{-1} = O(10^{13}) sec^{-1} \tag{3.111}$$

$$f = \frac{\omega}{2\pi} = 9\,X\,10^{12} sec^{-1} = O(10^{12}) sec^{-1} \tag{3.112}$$

Thus,

$$\frac{1}{f} = O(10^2) f sec \tag{3.113}$$

The bond stretching modes are an order of magnitude faster (sometimes regarded as hard modes) and are often constrained; the approach taken in the present work is to use bond length constraints. This also allows for a larger timestep (as discussed in reference to the multiple timestep methods above).

Thus, this section briefly discusses the procedure for obtaining flexible bond angle and out-of-plane potentials for OPLS methyl acetate.

In the interest of simplicity, a quadratic bond-angle and out-of-plane potential are desired. However, the preference for this form goes deeper than computational simplicity. Despite the need to account for bond-angle bending anharmonicity in highly strained molecules such as those containing three- and four-membered rings (cyclopropane, cyclobutane), the anharmonic contribution for linear systems is small and adds little physics to the picture. Furthermore, studies comparing the fits of various functional forms for empirical forcefields to *ab initio* calculations of equilibrium energies and first and second energy derivatives indicate that inclusion of angle-bending anharmonicity does not contribute to reproducing the quantum energy surface (Maple, Hwang, et al). Thus, in the present work, a quadratic bond-angle bending functional form is retained. Table 3.2 shows the angle and out-of-plane (Wilson definition, [112]) types required to describe the internals for a united-atom OPLS methyl acetate . There are four bond angles and a single out-of-plane mode (Figure 3-8).

After a search for the required angle potentials in existing all- and united-atom forcefields (e.g. CHARMm, AMBER,MM3), which led to still missing potentials,

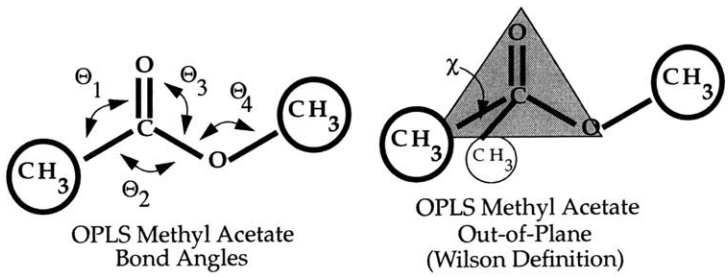


Figure 3-8: Definition of the Bond Angles and Out-of-Plane Angle for OPLS Methyl Acetate

it was decided to fit functions to the bond-angle and out-of-plane potentials of the CFF91 potential. Furthermore, the parameterization was oriented towards reproducing probability distributions of various angles, not the potentials directly. This approach is rigorous as it allows the incorporation of the effects of other internal modes (coupling effects) without having to perform a direct parameterization of all internals.

Parameterization of Bond-Angle and Out-of-Plane Functions: Procedure

This section describes the procedure followed to obtain a set of force constants for the OPLS bond-angle and out-of-plane angle bending modes.

First, a single-molecule Hybrid Monte Carlo simulation using the CFF91 all-atom forcefield was performed to generate a sample of bond and out-of-plane angles. Next the probability distributions for each of the internal angles was computed by normalizing a histogram of number of angles of a particular value, the histogram being calculated from all configurations.

Next, these distributions were used to determine force constants for the internal modes.

The functional form for the angle bending is:

$$U_\theta = \frac{1}{2} K_\theta (\theta - \theta_o)^2 \tag{3.114}$$

and that for the Wilson out-of-plane motion [112]:

$$U_\chi = K_\chi \chi^2 \tag{3.115}$$

The potentials were parameterized in the variables K_θ and K_χ by performing a least-squares fit of the Boltzmann distribution of angles:

$$P(\theta) = \frac{1}{N_\theta} e^{[(-\frac{1}{2}K_\theta(\theta-\theta_o)^2)/RT]} \tag{3.116}$$

$$P(\chi) = \frac{1}{N_\chi} e^{[(-\frac{1}{2}K_\chi(\chi^2)/RT]} \tag{3.117}$$

where the N_θ and N_χ are normalizing factors:

$$N_\theta = \sum_{i=1}^{M_\theta} (e^{[(-\frac{1}{2}K_\theta(\theta_i-\theta_o)^2)/RT]}) \tag{3.118}$$

$$N_\chi = \sum_{i=1}^{M_\chi} (e^{[(-\frac{1}{2}K_\chi(\chi^2)/RT]}) \tag{3.119}$$

The fitting involved:

1. Picking a value of K

2. Calculating P(θ) or P(χ)

3. Calculating sum of squares (of deviations)

4. Minimizing sum of squares with respect to K

Initial values of K were selected based on typical values of similar bond angles found in the literature.

Figures 3-9 and 3-10 show curves of the sum-of-squares error versus force constant for the relevant internals.

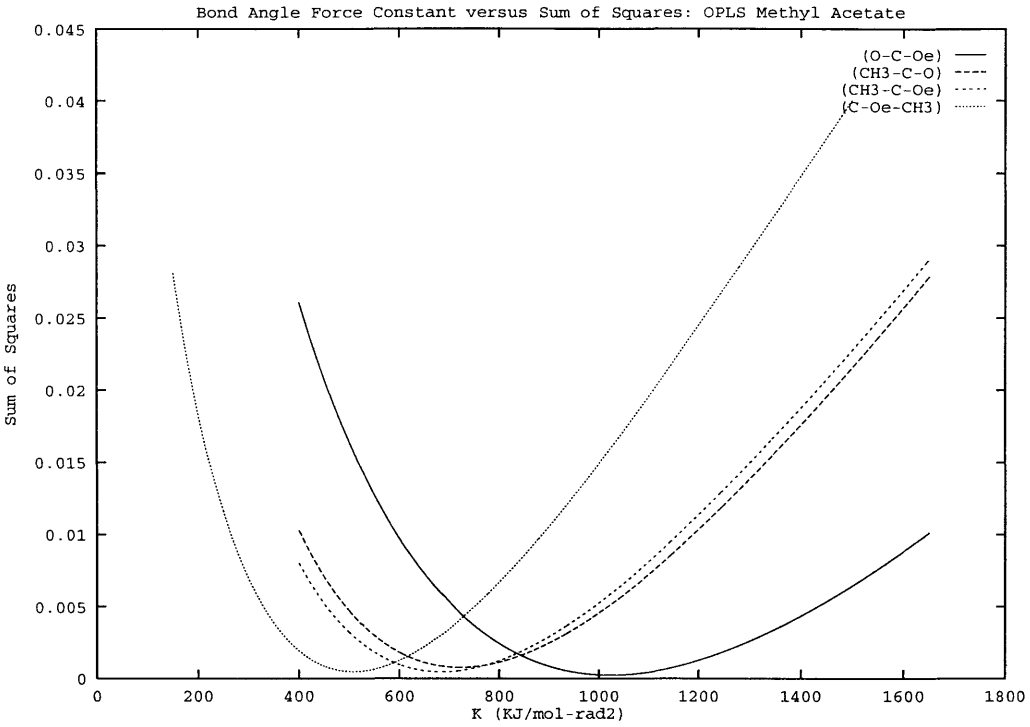


Figure 3-9: Sum of Squares versus Bond Angle Bending Force Constant

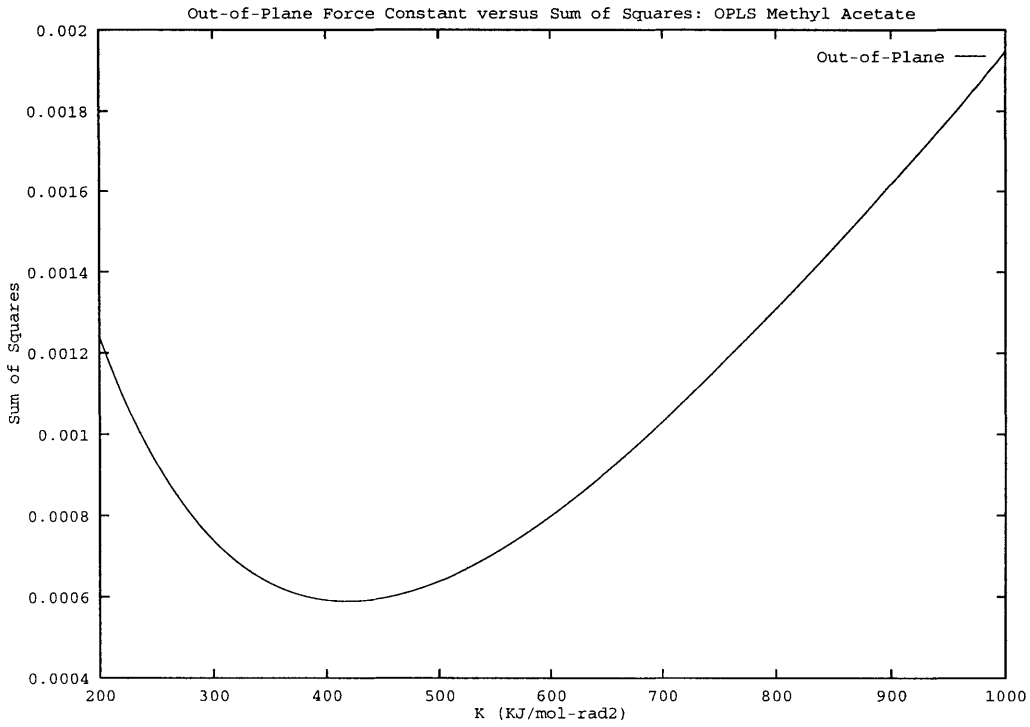


Figure 3-10: Sum of Squares versus Out-of-Plane Force Constant

Bond Angle and Out-of-Plane Parameterization: Results

Figures 3-11, 3-13, 3-15, 3-17, and 3-19 show the CFF91 all-atom bond angle and out-of-plane angle distributions computed from a series of single molecule Hybrid Monte Carlo simulations runs (solid line) along with the fitted distributions (points). The abscissa show the relevant internal (bond or out-of-plane angle) in units of radians, and the ordinate values are the probabilities. The distributions are well-behaved, with well-defined maxima near the most probable values of the internal. Figures 3-9 and 3-10 show the results from a simple least-squares fitting of the distributions of Figures 3-11, 3-13, 3-15, 3-17, and 3-19 with the above described functional forms. The plots show the values of the residual sum of squares as a function of the force constant K_θ and K_χ. From this data, the minimum of the curve was taken as the optimum value for the force constant. Using these values, a second set of single-molecule Hybrid Monte Carlo simulations (using the revised OPLS united-atom forcefield model) were performed and the results of these experiments are shown in Figures 3-12, 3-14, 3-16, 3-18, and 3-20. These plots show the angle distributions obtained with the CFF91 all-atom forcefield (solid line) and with the united-atom OPLS forcefield with angle potentials (points), parameterized in this work. The fits are reasonable within the error of the computations. Table 3.2 gives the values for the relevant internals along with some literature values for those angles on which there is published data.

Table 3.2: Angle Parameters for OPLS Methyl Acetate United-Atom Model

Angle Type	Equilibrium Angle (Degrees)	$K_{\theta,\chi}$ $(\frac{kJ}{mol\,rad^2})$
$CH_3 - C - O$	124.037	721.0
$CH_3 - C - Oe$	110.514	684.25
$O - C - Oe$	125.243	1023.0
$C - Oe - CH_3$	116.812	508.25
$CH_3 - C - Oe - O$ (Out-of-Plane Angle)	0.0	418.3

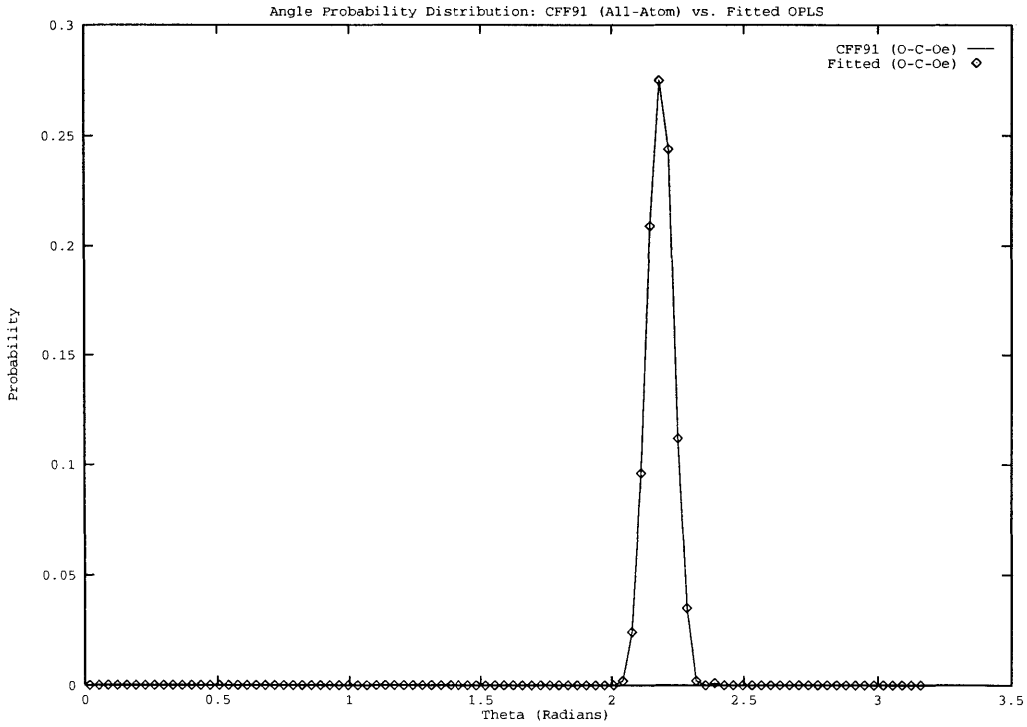


Figure 3-11: Angle 213 Fitted: This is the all-atom distribution (points) to which a united-atom potential form is fitted (line)

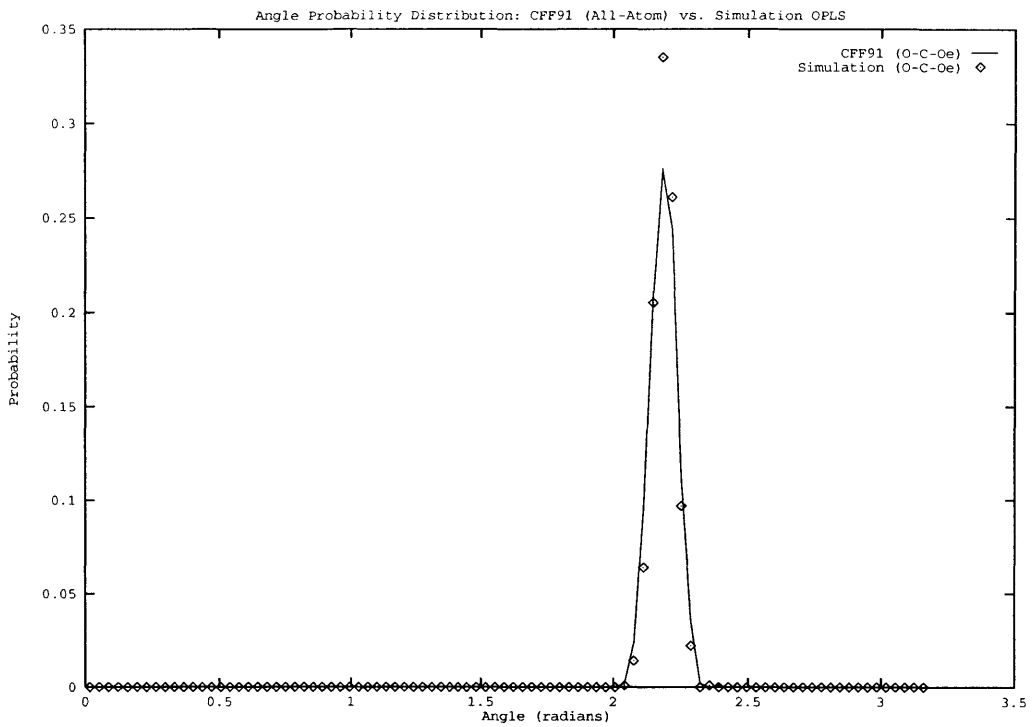


Figure 3-12: Angle 213 Simulation: This is the distribution of bond angle values computed from the results of a simulation using the parameters obtained via fitting to the all-atom distribution shown in the previous Figure

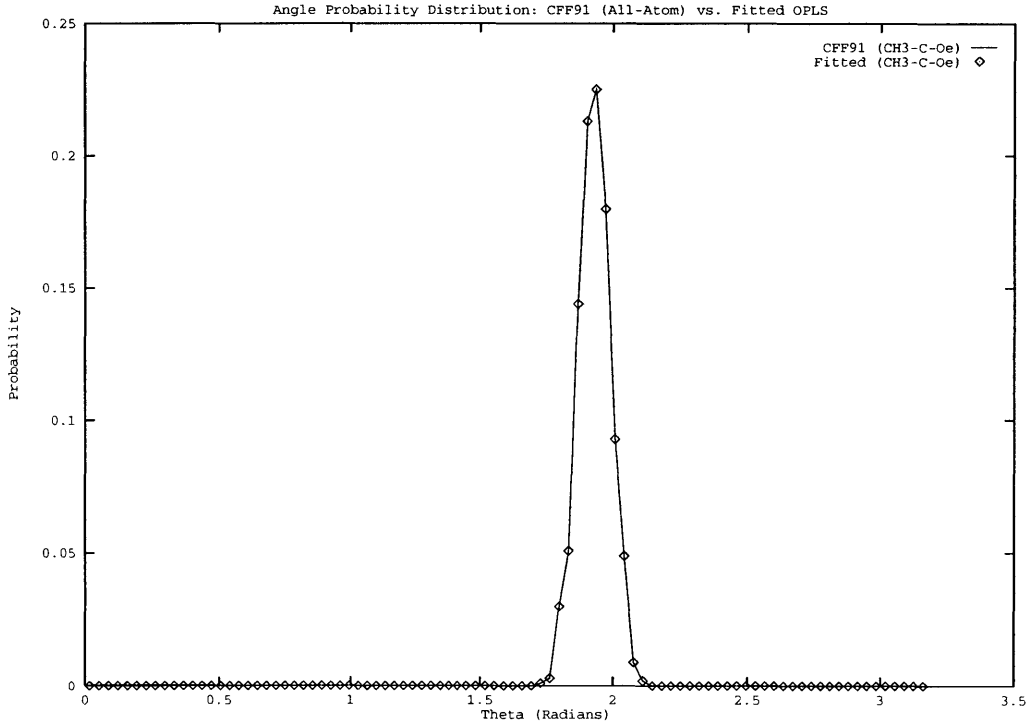


Figure 3-13: Angle 413 Fitted (see caption for Angle 213 fitted for details)

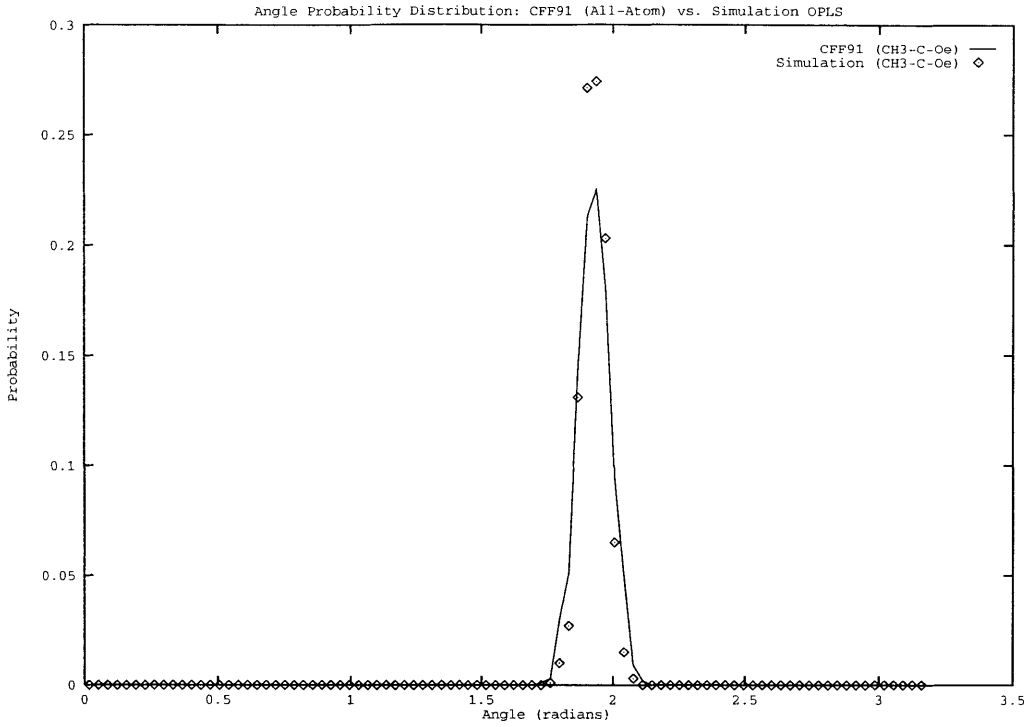


Figure 3-14: Angle 413 Simulation (see caption for Angle 213 simulation for details)

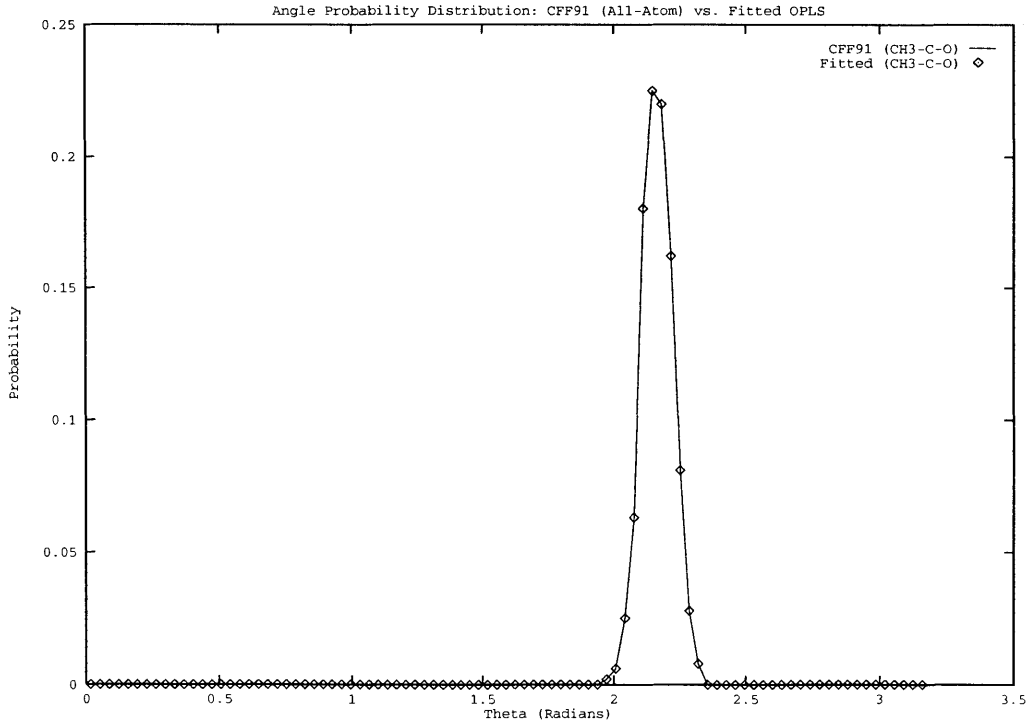


Figure 3-15: Angle 412 Fitted (see caption for Angle 213 fitted for details)

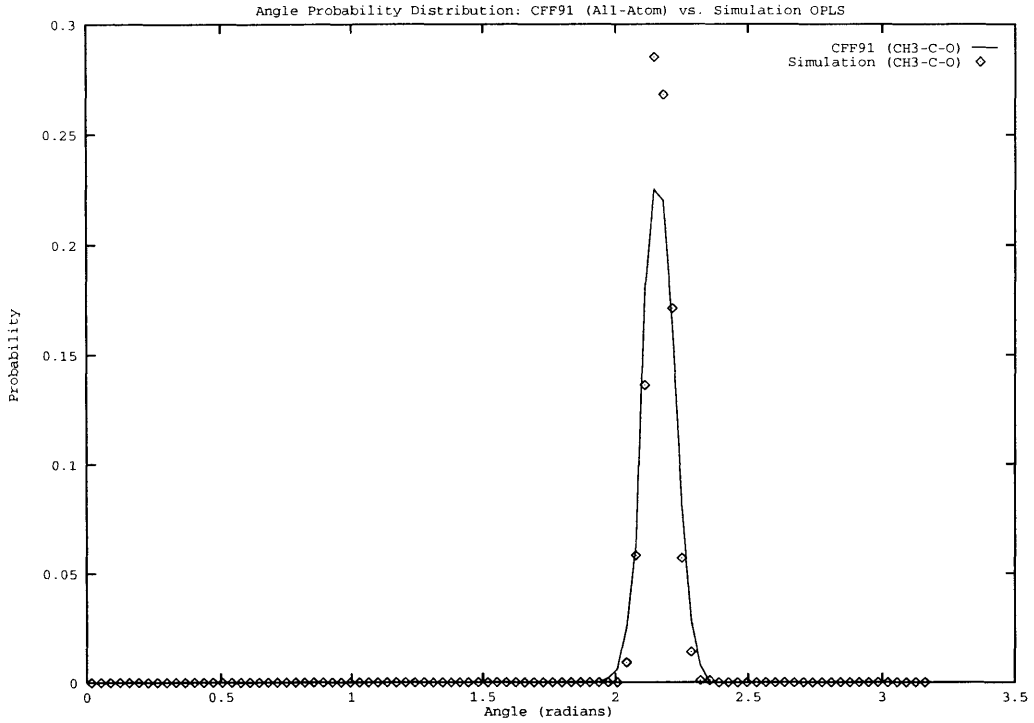


Figure 3-16: Angle 412 Simulation (see caption for Angle 213 simulation for details)

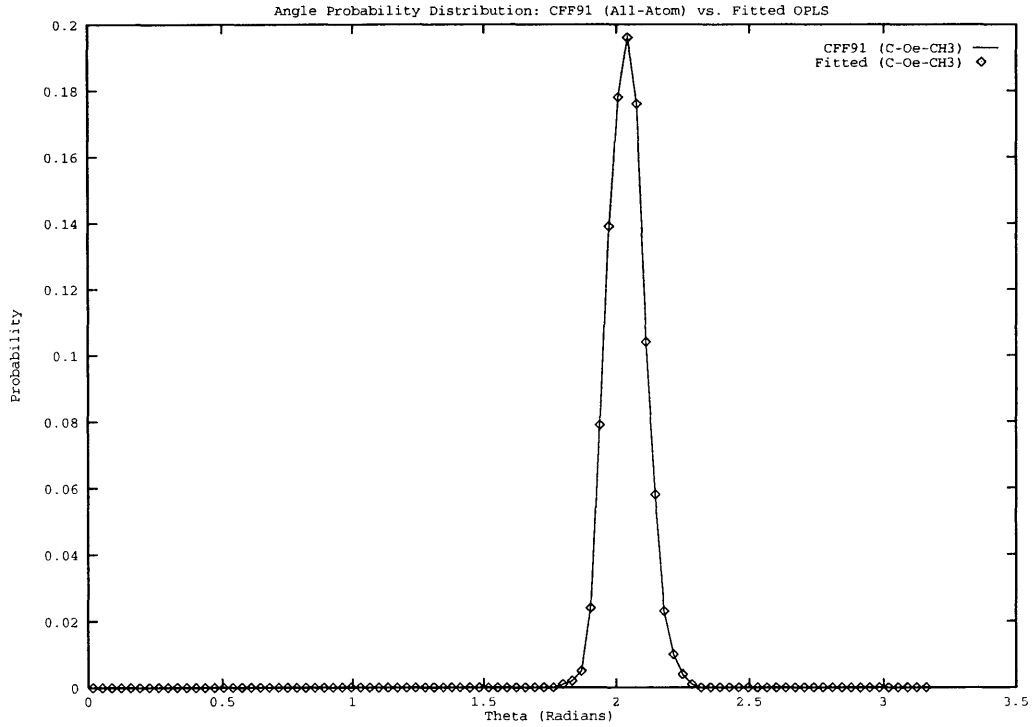


Figure 3-17: Angle 138 Fitted (see caption for Angle 213 fitted for details)

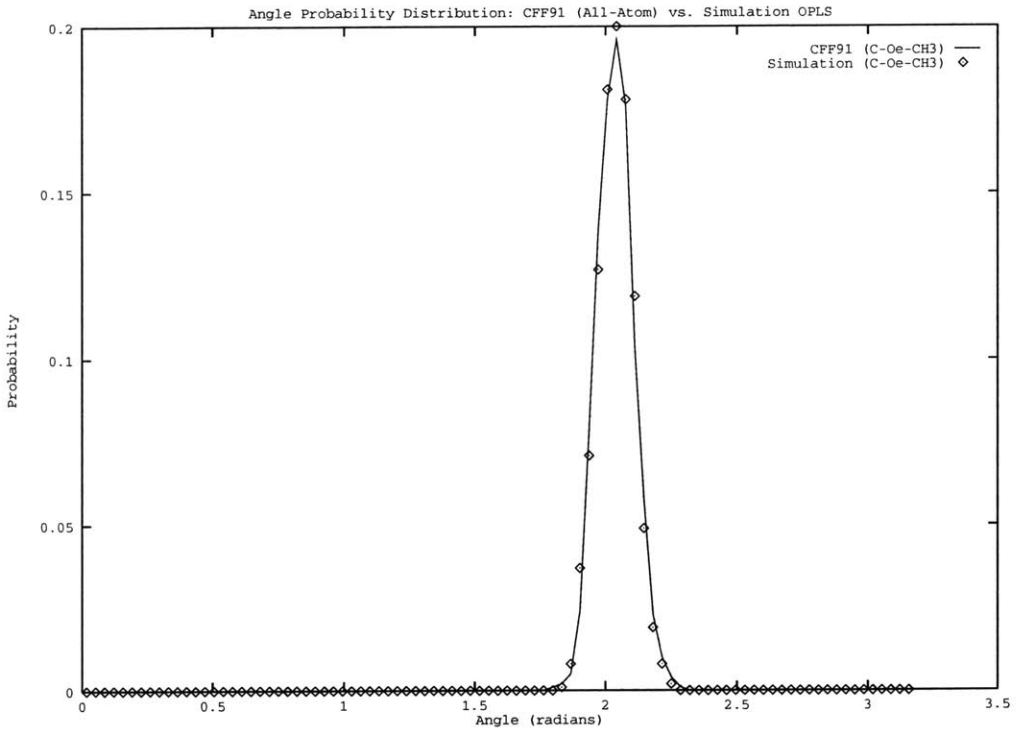


Figure 3-18: Angle 138 Simulation (see caption for Angle 213 simulation for details)

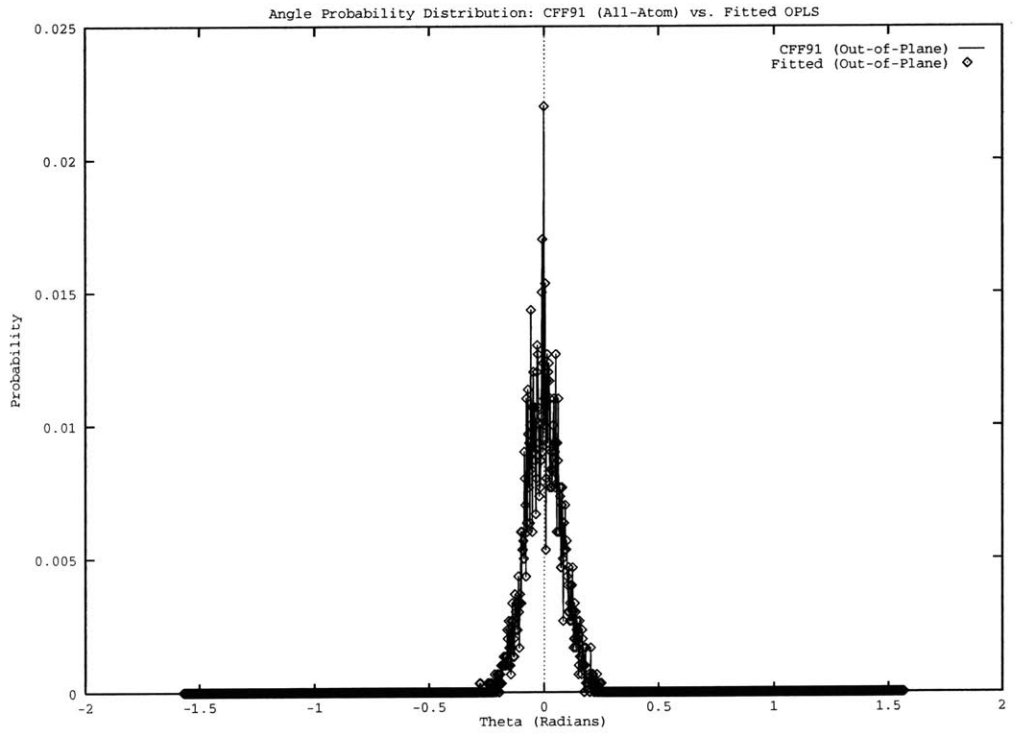


Figure 3-19: Out-of-Plane Angle 4132 Fitted (see caption for Angle 213 fitted for details)

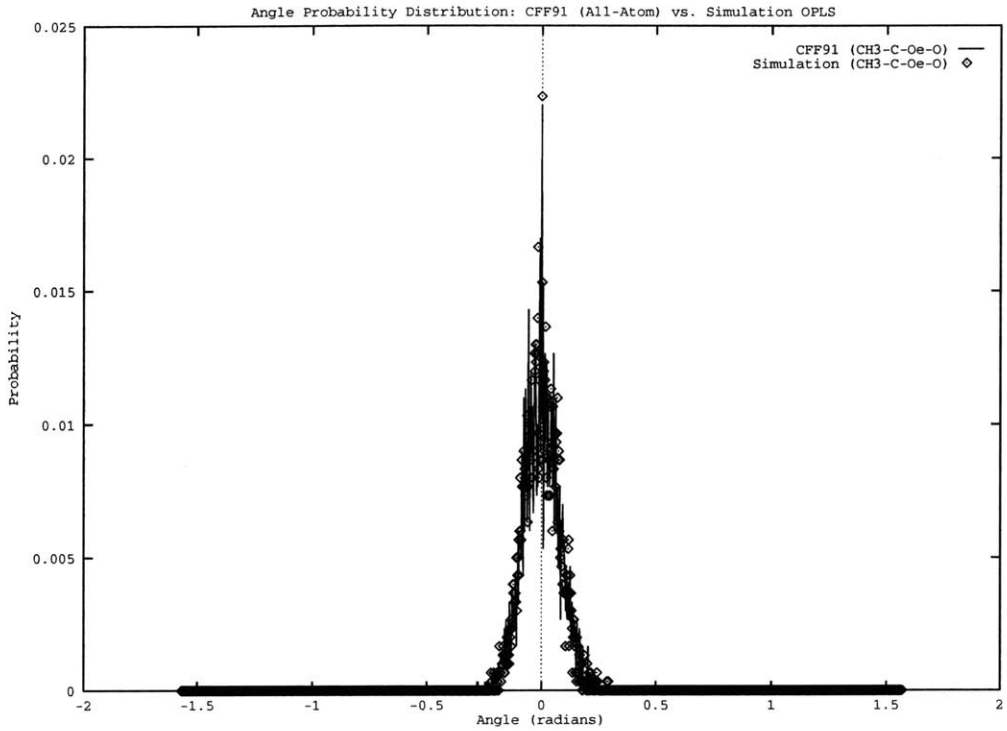


Figure 3-20: Bond Angle 4132 Simulation (see caption for Angle 213 simulation for details)

3.3 Gas Phase Simulations: Single Molecule Hybrid Monte Carlo

To compute the values for the vapor-phase intramolecular energy to be used in Equation 3.38, single molecule Hybrid Monte Carlo simulations were performed. Four sets of simulations were required to study the OPLS and CFF91 parameter sets using the reaction field and Ewald summation techniques. For the OPLS potential, standard molecular dynamics was performed for each Monte Carlo step; for the CFF91 (all-atom) forcefield, the r-RESPA algorithm was implemented for the MD portion of the hybrid technique. For the OPLS runs, a timestep of 0.004 picoseconds was used; for the CFF91 potential with r-RESPA, a large timestep of 0.0025 picoseconds and a small step of 0.0005 picoseconds with n=5 was used. For the Hybrid Monte Carlo algorithm, 100,000 MC steps were performed; for each MC step, two hundred MD steps were performed. A single molecule was place within a cubic box of dimensions of 10.0 nanometers. For the Ewald summation, the direct space sum is truncated to the minimum image; the α parameter is 5.0 and the reciprocal space summation is over 1147 vectors.

3.4 Bulk Liquid Energetics II.: Simulation Details and Results

3.4.1 Bulk Simulations with Reaction Field Electrostatics

OPLS Methyl Acetate

The bulk liquid simulations for the OPLS methyl acetate model with reaction field electrostatics are described in this section. The simulation cell is a cubic box of dimensions $L_x = L_y = L_z = 3.060nm$ representing a density of $0.927gr/cm^3$ which corresponds to the experimental density at 298K (216 OPLS molecules). A timestep of 0.004 picoseconds allows for sufficient integration stability with a ΔE^{rms} of three

percent over 20,000 simulation steps in a time of 5.2 hours (which yields about 0.94 seconds per time step). The r-RESPA algorithm was not used for the OPLS model simulations since there are no bond-stretching modes due to the bond-length constraints used; thus, a fairly large timestep is employable without any further consideration. A Nose-Andersen thermostat (as opposed to the Nose-Hoover used in the r-RESPA simulations using the CFF91 potential) is used to maintain temperature at 298K for canonical ensemble (NVT) simulations; a time constant of 5.5 is used. Data for statistical analysis is collected every 20 simulation steps (every 0.08 picoseconds). A spherical cutoff is applied to both dispersion and electrostatic interactions; a value of 1.2 nm is used. A lookup table which holds the non-bond energies as a function of r^2 as well as an atomic neighbor list are used to expedite the force/energy calculations. Forces are computed from the tabulated energies via interpolation using the Newton-Gregory forward difference method [13]. Long-range corrections for pressure and energy are computed from standard analytic expressions [13],

$$U_{long-range} = 2\pi N\rho \int_{r_c}^{\infty} r^2 u(r)\, dr \tag{3.120}$$

$$(PV)_{long-range} = -\frac{2}{3}\pi n\rho \int_{r_c}^{\infty} r^2 w(r)\, dr \tag{3.121}$$

These equations assume that in the region beyond the spherical cutoff, the pair correlation is unity. The long-range corrections given here are applied only for the non-bond dispersion contributions to the energy and pressure (long-range being outside the cutoff radius of 1.2 nm). The electrostatic contributions are effectively accounted for by the reaction field approach. The need for the above two equations to be applied in the case of the dispersion interactions is that these components do not go to zero at the cutoff (and remain so beyond it) as in the case of the reaction field electrostatics.

The parameters used for the simulation are given in Table 3.3.

Table 3.3: OPLS Non-Bond Parameters: Methyl Acetate United-Atom Model

Atom Type	σ (nm)	ϵ ($\frac{kJ}{mol}$)	Mass (amu)	Charge (e)
Carbonyl Carbon	0.375	1.75728	12.01115	0.550000
Carbonyl Oxygen	0.296	3.51456	15.9994	-0.450000
Ether Oxyen	0.30	2.84512	15.9994	-0.40
Carbonyl Methyl	0.391	2.67776	15.03506	0.050
Ether Methyl	0.380	2.84512	15.03506	0.250

CFF91 Methyl Acetate

This section describes the bulk liquid simulations for the CFF91 methyl acetate model with reaction field electrostatics. The simulation cell is a cubic box of dimensions $L_x = L_y = L_z = 3.060389593nm$ representing a density of $0.927gr/cm^3$ w hich corresponds to the experimental density at 298K (216 OPLS molecules). The r-RESPA multiple timestep algorithm for Nose dynamics is used to carry out the integration. A large timestep of 0.0025 picoseconds and a small step of 0.0005 picoseconds with n=5 is used. This choice of parameter values leads to stable integration with a ΔE^{rms} of 1.7 percent over 10000 simulation steps in a time of 22 hours (which yields about 7.9 seconds per time step). Data for statistical analysis is collected every 20 simulation steps (every 0.05 picoseconds). Further details are as for the OPLS simulations discussed in the previous section. Table 3.4 gives the non-bond parameters for the CFF91 methyl acetate model.

Table 3.4: CFF91 Non-Bond Parameters: Methyl Acetate All-Atom Model

Atom Type	σ (nm)	ϵ ($\frac{kJ}{mol}$)	Mass (amu)	Charge (e)
Carbonyl Carbon	0.3308	0.502079	12.01115	0.3994
Carbonyl Oxygen	0.3535	1.117127	15.9994	-0.3964
Ether Oxyen	0.3535	1.004159	15.9994	-0.1163
Carbonyl Methyl Carbon	0.401	0.225936	12.011150	-0.159
Ether Methyl Carbon	0.401	0.225936	12.01115	-0.0457
Hydrogen	0.2995	0.08368	1.00797	0.053

3.4.2 Bulk Simulations with Ewald Summation Electrostatics

OPLS Methyl Acetate

For the bulk simulations using the flexible OPLS model, the major simulation details are the same as for the bulk reaction field computations. The major differences are outlined. The timestep used for these runs is 0.002 picoseconds with the same sampling frequency for statistics. The integration stability over 5000 simulation steps is rather good, with a ΔE^{rms} of 1 percent. For the Ewald summation, the direct space sum is truncated to the minimum image; the α parameter is 5.0 and the reciprocal space summation is over 1147 vectors.

CFF91 Methyl Acetate

For this set of simulations, the α parameter was again set as 5.0 with the number of k-vectors reduced to 501 in the interest of time. The same r-RESPA parameters are used. A simulation of 20,000 MD timesteps is performed. The total time is roughly 3 days.

3.4.3 Results of Bulk Liquid Energetics Experiments

Table 3.6 shows the component energies computed from the bulk liquid and single molecule Hybrid Monte Carlo simulations using the CFF91 forcefield; Table 3.7 shows the analogous data for the OPLS model. The component energies include the non-bond intermolecular energy (Van der Waals, dispersion), the intramolecular non-bond energy, the angle-bending energy, out-of-plane angle bending energy, torsion energy, bond-stretching energy, and the total electrostatic energy.

Table 3.5 presents the overall bulk liquid energetics results from the present calculations. Values of total bulk liquid potential energy, gas-phase intra molecular energy, computed enthalpies, and deviations from experimental data are shown. Simulations were run with the CFF91 and OPLS potentials as well as with varying methods of

computing electrostatic interactions (as mentioned above). Two sets of data are included to show the effect of including long-range corrections to non-bond energies and pressures.

From the results shown in Table 3.5, it is evident that both the OPLS and CFF91 potentials provide reliable enthalpies of vaporization when the electrostatic interaction is handled exactly via Ewald summation. With the Ewald method, the CFF91 model yields an enthalpy change within 2% of the experimental value. This high fidelity to experiment comes for a significant price, as the computational expense to achieve such accuracy is large; for the methyl acetate system, a 5ps simulation for 216 molecules requires approximately 16 hours of real time, as opposed to 2.4 hours for a reaction field simulation of equal length. For dense, non-ordered macromolecular systems, the Ewald method would not be practical for obtaining meaningful statistics in respectable times.

An interesting point to note is the consistency of the calculated enthalpies using the OPLS model with reaction field and Ewald summation. As the Ewald result can be considered an exact result for the electrostatic model (i.e., charges for each atom), we see that the reaction field performs fairly well in faithfully representing the electrostatics of the system. The agreement between the reaction field and Ewald methods has also been observed in simulations of spc (non-polarizable) water [113] and a three-site CO2 potential [114]. One can note that in all these cases, there are no INTRAmolecular interactions of non-bonded nature, so we are led to believe that there may be some inconsistency of the reaction field applied to systems with intramolecular non-bond (both van der waals and electrostatic) interactions. More specifically, the dependence of the reaction field energy on the choice of cutoff is of concern since depending on the system and molecular geometry, the nature of intramolecular contributions to the total electrostatic energy will change with cutoff; the net charge within the cutoff radius is not necessarily zero for any given charge at any given time. This is obviously true since the cutoff is simply a site-based spherical region; there can by any number of particles (consistent with the system density of course) within this region, and since no strict constraints on charge neutrality within

the cutoff region are imposed, the net charge in the cutoff sphere does not have to vanish. Hoever, the net system charge is zero as is the not dipole since we are dealing with a bulk system. Furthermore, Table 3.6 shows that for the CFF91 forcefield calculations, the largest discrepancy in the component energies is with respect to the value of the electrostatic contribution; for the OPLS model, there is no significant difference in component energies and hence both RF and Ewald summation yield small errors. Further note that the agreement between the OPLS simulations and experiment is not a big surprise since the parameterization of the non-bond parameters involved fitting to the experimental enthalpy of vaporization to the ideal gas. Also from Table 3.6 and 3.7 and note that the intramolecular energetics do not change from the liquid to the vapor phase for both forcefield models.

Table 3.5 also shows a significant contribution from the long-range corrections to intermolecular energy, lowering the deviation from experimental for both models. This merely indicates the need to account for the long-range interactions neglected by a cutoff treatment of the non-bond interactions. Unfortunately, the pressures predicted for the models are two orders of magnitude greater than experimental values.

Based on the results of the liquid simulations, the OPLS parameter set is selected as the forcefield model to study aspects of interfaces of ester materials.

Table 3.5: Bulk Liquid Energetics

Potential	Density gr/cm^3	P_{calc} (bar)	$< U_{total} >$ (kJ/mole)	$< U_{intra}(gas) >$ (kJ/mole)	ΔH_v^o (kJ/mole)	Error (percent)
LR^*						
CFF91(RF)	0.927	257.23	15.13	50.28	37.63	15.88
CFF91(EW)	0.927	123.55	4.35	35.00	33.13	2.04
OPLS(RF)	0.927	344.16	-23.82	6.30	32.60	0.39
OPLS(EW)	0.927	320.31	-24.09	6.18	32.74	0.84
OPLS(Monte Carlo)	0.905	1.0	-29.46	1.30	32.26	2.4
NLR^{**}						
CFF91(RF)	0.927	631.43	16.50	50.28	36.26	11.66
CFF91(EW)	0.927	464.60	5.72	35.00	31.76	2.18
OPLS(RF)	0.927	631.13	-22.67	6.30	31.45	3.15
OPLS(EW)	0.927	607.28	-22.94	6.18	31.60	2.69
Temp. = 298K $\Delta H_v^o(expt.) = 32.47 kJ/mole$						

LR^* Long-Range Correction to Intermolecular Energy Included
NLR^{**} Long-Range Correction to Intermolecular Energy Excluded

Table 3.6: CFF91 Component Energies

Type	$< U_{inter} >$ (kJ/mole)	$< U_{intra} >$ (kJ/mole)	$< U_{oop} >$ (kJ/mole)	$< U_{angle} >$ (kJ/mole)	$< U_{tor} >$ (kJ/mole)	$< U_{elec} >$ (kJ/mole)
Bulk: RF	-32.401	13.011	1.323	38.031	2.632	-21.170
HyMC: RF	0.0	12.717	1.50	37.194	3.043	-17.989
Bulk: EW	-32.408	13.009	1.333	38.051	2.608	-31.961
HyMC: EW	0.0	12.345	0.946	35.318	2.819	-28.617

Table 3.7: OPLS Component Energies

Type	$< U_{inter} >$ (kJ/mole)	$< U_{intra} >$ (kJ/mole)	$< U_{oop} >$ (kJ/mole)	$< U_{angle} >$ (kJ/mole)	$< U_{tor} >$ (kJ/mole)	$< U_{elec} >$ (kJ/mole)
Bulk: RF	-26.747	0.0	1.308	3.860	1.460	-3.697
HyMC: RF	0.0	0.0	1.205	3.764	1.332	0.0
Bulk: EW	-26.791	0.0	1.266	3.724	1.452	-3.737
HyMC: EW	0.0	0.0	1.150	3.694	1.337	0.00

Chapter 4

Methyl Acetate Liquid-Vapor Interface

4.1 Introduction and Objectives

This chapter describes the results of simulations of the methyl acetate liquid-vapor interface. The liquid-vapor interface of small molecules and atomic fluids has been studied in great detail via theory (such as van der Walls and Cahn-Hilliard theories), density functional theories, lattice-based mean-field theories, integral equation theories, and molecular simulations. The molecular dynamics method has been applied to the direct simulation of coexisting phases. The approach has been applied,for example, to atomic liquids [115, 116], alkane oligomers [60], methanol [56], water [55], and water-methanol mixture [117].

Esters are an important chemical species as they are used industrially for a range of applications from fragrances/flavorings to the manufacture of nitrocellulose, acetyl cellulose, and a variety of resins and oils [118]. Furthermore, aspects of the liquid-vapor interface are relevant to the understanding of the use of esters in these applications. From a fundamental, physical viewpoint, the study of the liquid-vapor interface of small-molecule esters is relevant and important to the process of understanding the nature and physics (structure, thermodynamics, dynamics) of more complicated ester-group containing molecules and systems such as oligomers, poly-

mers, surfactants and other complex fluids. In the simplest sense, there is only the single monomeric (if one is to think of the small molecule as being some elemental unit of a longer chain entity) species which captures the important structural and energetic components; there are no issues of chain connectivity, excluded volume effects, intramolecular segmental correlations (in terms of conformations) to couple with more fundamental physical mechanisms such as hydrogen-bonding which give rise to the observed microscopic and macroscopic materials properties. Moreover, the methyl acetate molecule by itself possesses a variety of physical elements. For instance, there is the juxtaposition of what can be considered hydrophobic and hydrophillic groups (the carbonyl oxygen and the methyl groups), the presence of a permanent dipole moment, and a torsional degree of freedom (within the OPLS model) which allows for molecular conformational changes which translates into the shifting of internal components relative to one another in response to internal as well as external stimuli.

Now, one might ask why study the liquid-vapor interface (or equivalently, the free surface) of methyl acetate; moreover, one may go so far as to ask why study the interface of a liquid at all. Free surfaces have been extensively studied theoretically (as mentioned in the opening of this Chapter) and experimentally. Interesting phenomena occur only at the interfaces between the coexisting phases. These phenomena are sometimes in contrast to the nature of the bulk fluid. For instance, strongly associating liquids such as water, methanol, and higher n-alcohols seem to show a strong orientational preference at the liquid free surface. This has been observed via molecular dynamics simulations [55, 56] and non-linear surface vibrational spectroscopic methods (sum frequency and second harmonic generation) [119, 120, 121]. The mechanism for this structuring is the maximization of favorable energetic interactions, such as *hydrogen bonds*, which are characteristic of the bulk liquid. In essence, the liquid 'order' is maintained right through to the surface. In the case of methanol, this leads to the hydroxyl group orienting towards the bulk liquid, with the 'hydrophobic' methyl group forced into the vapor. Much the same occurs with the normal alcohols up to C_8. Furthermore, Good [122] in 1957 published a set of data indicating anamolously lower surface entropies of hydrogen-bonding liquids as compared to non-

polar and polar, but non-hydrogen bonding fluids. This trend was associated to the surface orientation induced by the maximization of hydrogen bonding interactions at the surface. Fundamentally, the total free energy is minimized by the energetic gains of strong association which dominate the entropic loss of orientation. On more than one occasion, the behavior of these ordinarily simple liquids has been likened to that of surfactants (surface active agents).

In the case of methyl acetate, there is no evidence (experimental or otherwise) of strong association in the bulk. In fact, methyl acetate would be considered a *normal* liquid in Good's jargon [122] (although, we note, Good does not specifically mention the class of esters in his study). However, one can draw the analogy to a 'surfactant' type molecule based on the juxtaposition of quite differing substituents. The polar carbonyl oxygen shares a proximal location to what can be considered a 'hydrophobic' methyl group (the alkoxy methyl group). Thus, it would be interesting to determine the behavior of this low molecular weight liquid at the free surface. This will allow one to examine the quality of the potential model employed (when compared with the appropriate experiments), as well as provide insight as to the behavior of the carbonyl functionality at the free surface. This information will be relevant to the interpretation of the results of the oligomer free surface to be presented in Chapter 7.

In this chapter, the liquid-vapor interface of methyl acetate is studied via direct molecular dynamics simulations. It is the aim of this work to investigate the structure of the interface in terms of the molecular geometry (bulk versus interface) and thermodynamics. This will allow a test of the OPLS potential functions with respect to its ability to predict surface tension; Chapter 3 details the soundness of this parameter set in sufficiently reproducing bulk liquid energetics, specifically, the heat of vaporization. As it is widely accepted that prediction of surface tension with current forcefields and techniques is prone to significant error, it is worthwhile to investigate this phenomena with respect to the OPLS bulk parameters. Furthermore, molecular orientation at the interface is targeted since this will provide insight on the structure of the free ester interface in terms of what functionalities are present where spatially;

this is directly related to the importance of the interface in biomedical applications. The concept of unsymmetrical molecules orienting preferentially in the interfacial region is well-accepted in the surface science community [123]. The phenomenon is held to be a result of the tendency to reduce total free energy by increasing the number of favorable interactions. It has also been put forth that molecular orientation is a result of the "principle of least abrupt change" in force fields; that is, molecules near and within the interface orient in such a way so as to provide the most gradual transition from the bulk environment of one phase (fluid) to that of the coexisting one [123].

Section 4.2 gives a description of the simulations and Section 4.3 gives the results of the simulations in terms of density profiles, bulk and interface radial distribution functions, orientational order parameters describing the orientations of molecular bond vectors relative to the interface normal, and surface tensions calculated from the atomic and molecular virials as well as from the Irving-Kirkwood definition of the pressure tensor at a point.

4.2 System Description and Methods

The simulation cell consists of 648 OPLS united atom methyl acetate molecules. The simulation geometry was established by placing three equilibrated bulk slabs of 216 molecules adjacent to each other. The configurations are from the bulk reaction field runs for the energetics study of the previous chapter. The slabs are equilibrated at the experimental density of 0.927 gr/cm^3 at temperature of 298K. The interface is generated by imposing an asymmetry in the box dimensions; the coordinate normal to the planar interface is made roughly five times larger than the transverse dimensions without rescaling particle coordinates. This ensures that a bulk region is established in the center of the cell as well as preventing interactions between particles at the two liquid-vapor interfaces. Periodic boundary conditions are applied in *all* Cartesian dimensions; this is akin to having a reflecting wall normal to the z-direction, except that here, molecules cross the wall and appear on the other equivalent side of the liquid slab. The simulation box dimensions are $L_x = L_y = 3.08483nm$ by $L_z = 15.0nm$.

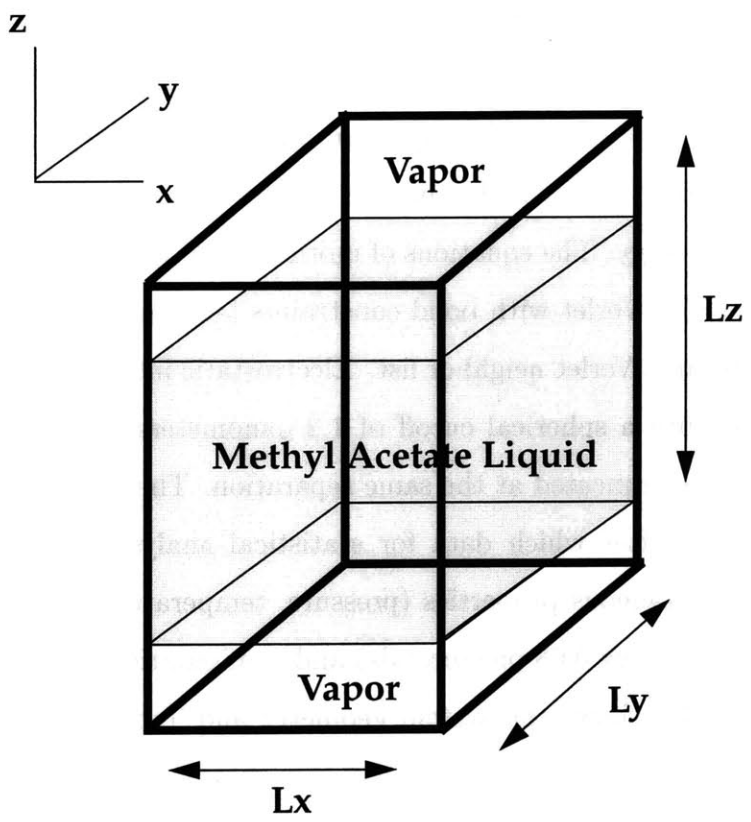
This geometry along with a cutoff of 1.2 nm makes sure that the two interfaces created do not interact with one another. NVT simulations were performed with Nose-Andersen thermostat with time constant of 5.5 picoseconds. A time step of 4 femtoseconds was used, with the fluctuation in total energy being within 10 percent that of the kinetic energy. The equations of motion are integrated using the RATTLE algorithm for Velocity Verlet with bond constraints [34]. Intermolecular interactions were computed using a Verlet neighbor list. Electrostatic interactions were computed via reaction field with a spherical cutoff of 1.2 nanometers. Non-bond interactions were also spherically truncated at the same separation. The system was equilibrated for 440 picoseconds after which data for statistical analysis was collected for 1.2 nanoseconds. Instantaneous properties (pressure, temperature, energies) were saved every 20 time steps (every 0.08 picoseconds) and configurations were saved every 500 timesteps. Figure 4-1 shows the system geometry and Table 4.1 gives the relevant parameters.

Table 4.1: OPLS Non-Bond Parameters: Methyl Acetate United-Atom Model

Atom Type	σ (nm)	ϵ ($\frac{kJ}{mol}$)	Mass (amu)	Charge (e)
Carbonyl Carbon	0.375	1.75728	12.01115	0.550000
Carbonyl Oxygen	0.296	3.51456	15.9994	-0.450000
Ether Oxygen	0.30	2.84512	15.9994	-0.40
Carbonyl Methyl	0.391	2.67776	15.03506	0.050
Ether Methyl	0.380	2.84512	15.03506	0.250

4.3 Results

In this section, the structural and thermodynamic results of the OPLS flexible methyl acetate model liquid-vapor interface simulations are presented. For structural analysis, various density profiles including total molecular number and component number density profiles are discussed. All profiles are computed with respect to the center of mass of the system, and the symmetry of the system is exploited for averaging



Lx = Ly = 3.08483 nm
Lz = 15.0 nm

Figure 4-1: Liquid Vapor Interface Simulation Cell Geometry

purposes; thus, only the symmetrized profiles are shown. Also presented are the order parameters P1 and P2 (first and second Legendre coefficients) representing the orientation of various bond vectors relative to the interface normal. Again, these are computed with respect to the system center of mass and the symmetrized profiles are shown. Finally, radial distribution functions for the bulk and interfacial regions are computed. We note that polarization effects are not considered in the present work since the models used for the ester (and water in subsequent chapters) neglect molecular polarizability.

4.3.1 Density Profiles

The symmetrized total and component number density profiles are shown in Figure 4-3 through Figure 4-5 . Figure 4-3 shows the total molecular number density. A sharp interface is evident; the bulk density drops to essentially zero within one nanometer. This is typical of small molecule liquid-vapor interfaces at temperatures far from the critical point. More importantly, this is a substantial result in that a stable heterogeneous system is established (at the temperature selected) without any need to modify the intermolecular interaction component of the OPLS potential model. We observe no significant deterioration of the system structure as a function of time. We confirm this by monitoring the positions of the Gibbs dividing surfaces during the simulation. Arising from the symmetry of the simulation cell are two dividing surfaces. The z-coordinate of the Gibbs dividing plane is computed for each configuration as,

$$Z_G = Z_{com} \pm \frac{N_{\pm}}{L_x L_y \rho^{bulk}} \tag{4.1}$$

where Z_G is the Gibbs dividing surface, Z_{com} is the center of mass of the total system (all molecules), N_+ represents the number of molecules whose centers of mass are located at higher z-coordinate values than the system center of mass (N_- is analogous), L_x and L_y are the in-plane dimensions, and ρ^{bulk} is the bulk density. This equation derives directly from the definition of the Gibbs dividing surface as being the location which gives no surface excess of material; that is, for a given density profile, the areas between the actual density profile and the constant bulk values on either side of the dividing surface are equal. That is, we have the equality

$$\int_{Z_{com}}^{Z_G} \left[\rho^{bulk} - \rho(z)\right] dz = \int_{Z_G}^{\infty} \left[\rho(z) - \rho^{vapor}\right] dz \tag{4.2}$$

where $\rho^{bulk}$is the bulk liquid denstiy and ρ^{vapor} is the bulk vapor density. Taking the bulk liquid density to be a constant, ρ^{bulk}, and the bulk vapor density for the present simulation configuration to be zero, $\rho^{vapor} = 0$, the equality becomes,

$$\rho^{bulk}\,(Z_G - Z_{com}) \;-\; \int_{Z_{com}}^{Z_G} \rho(z)dz \;=\; \int_{Z_G}^{\infty} \rho(z)dz \tag{4.3}$$

Rearranging so as to place both integrals on the same side of the equality, we get

$$\rho^{bulk}\,(Z_G - Z_{com}) \;=\; \int_{Z_{com}}^{Z_G} \rho(z)dz \;+\; \int_{Z_G}^{\infty} \rho(z)dz \tag{4.4}$$

The two integrals can be combined as

$$\int_{Z_{com}}^{Z_G} \rho(z)dz \;+\; \int_{Z_G}^{\infty} \rho(z)dz \;=\; \int_{Z_{com}}^{\infty} \rho(z)dz \;=\; N^{+} \tag{4.5}$$

We have introduced the value N^{+} which is used in Equation 4.1 which represents the density profile averaged number of molecules per unit area. We can now rearrange and obtain straightforwardly Equation 4.1.

Figure 4-2 shows the time profiles for the last 800 picoseconds of the simulation. The surface for z-coordinate values less than zero has been shifted up by 9.0 nanometers for clarity. The profiles demonstrate that the positions of the dividing surface fluctuate about a mean value (and hence signify the stability of the interface); the fluctuation clearly indicates the density fluctuations at the interface. The average bulk density is 7.109 nm^{-3} which is 0.874 gr/cm^3 (0.0026), a value approximately 5.7 percent too low compared to the experimental value at the same temperature; however, this is within the general order of magnitude error found for bulk densities in interfacial simulations (consider the 12 percent error obtained in bulk densities of methanol by Matsumoto and Kataoka [56]); furthermore, this drop in density results from imposing a finite cutoff on the distance over which intermolecular interactions are accounted, as well as the increase in the longitudinal simulation cell dimension which allows the system to expand in this dimension. The interface is reached within 4.5 nanometers from the bulk center and extends out to about 5.5 nanometers. The 10-90 thickness is computed to be 0.6 nanometers (6 angstroms). This is the thickness over which the ester density changes from 90 percent to 10 percent of its bulk liquid value (here taken as the average value of 7.109 nm^{-3}). The 10-90 thickness corresponds to roughly 1.2 molecular diameters (the molecular diameter is taken to

be 0.4936 nm [124]); this is in keeping with the interfacial thicknesses observed for the free surfaces of water and methanol at approximately equivalent absolute temperatures [53, 55, 56, 125, 126, 54]. When compared on the basis of reduced temperature, T/T_c, where T_c is the critical temperature, we also find that the computed thickness is consistent with previous simulations of molecular systems. We note, however, that the model fluids simulated are just that–they are approximate models. Due to the fact that not all the models have been characterized with respect to the associated critical points, we have used experimental values and accept the uncertainty in so doing.

Comparison of the 10-90 thickness to experimental data from ellipsometry and X-ray reflectivity is often difficult as simulations systematically underestimate the 10-90 thickness due to the lack of capillary waves being included in the simulation. Capillary waves are inherent in experimental setups and in theory will broaden the interface, giving higher interfacial thickness measurements [54]. Figure 4-4 shows a component number density profile. Here there is no strong structuring at the interface per se. There is a slight difference in the regions where certain functionalites are present. In the bulk region, there is obviously no preferential spatial separation of the various groups. Moving out towards the interfacial region, there appears to be a slightly higher tendency for the more electronegative carbonyl and ether oxygens to segregate to this region, while in the outermost regions of the interface, the methyl group density is only very slightly enriched. Figure 4-5 shows a magnified region of the component density profile to more clearly demonstrate the above observations. Finally, the total molecular number density profile over various time intervals during the production run are presented in Figure 4-6. These show a quite stable behavior over the time of the simulation. Also, Figure 4-7 shows a time profile of the bulk density (with the time average value shown as a straight line).

Finally note that the interface shown here is more likely representative of the inherent/intrinsic interface due to suppression of long wavelength capillary waves in MD simulations. Capillary waves, originating from the thermal motions of particles, tend to widen (or analogously, 'roughen') the interface (both experimentally and

via simulation), and so one can expect that the true width is somewhat wider than computed [50]. Continuum theory of interfaces indicate two characteristic interfacial widths, 1). the intrinsic width over which the density at a planar interface changes from one bulk value to the coexisting value, and 2). the amplitude of capillary waves on the inherent planar surface; thus, capillary waves can be thought of as being superimposed upon the inherent interface which has its own thickness [127]. Capillary wave broadening is more pronounced at liquid-liquid rather than liquid-vapor interfaces since in general the mass density and surface tension are smaller in the liquid-liquid case [50]. Since the computational limitations restrict the number of particles that are included in the simulation system, and thus the size of the interface indirectly, for an interface of dimensions L x L, only the short wavelength modes will be visible via molecular dynamics simulation. That is, only those capillary waves with wavelength longer than order L will not be described via molecular simulation for the system sizes commonly used [50]. Furthermore, the major contribution of capillary waves to interfacial broadening is from small k waves, or those with large wavelengths [128]. For example, for a system with $L = 1cm$, a surface tension of $\gamma = 50dyne/cm$, $T = 300K$, and $l_m = 5 \times 10^{-8}cm$, typically, 75 percent of the contribution to the capillary amplitude is from waves of length greater than 100 times l_m, the minimum wavelength definable for the interface (which for the present discussion is taken to be a characteristic molecular diameter) [128]. In the final analysis, the simulated system can be viewed as a locally 'flat' area of a long-wavelength, periodic structure (the dimensions of the simulated system are such that in effect, the radius of curvature of the wave structure is immensely large) [61]. The section 'Capillary Wave Theory and Fluid Interfacial Broadening' goes into a little more theoretical detail concerning the capillary wave picture of interfaces.

At this time, we have not found any experimental data on the methyl acetate free surface thickness; however, based on literature reports of computed interfacial thicknesses, we believe that the present computations will fare no better in reproducing experimental observations [61, 50, 126, 54, 56]. One can estimate the capillary wave broadening of the interface by the mean square amplitude of capillary waves given by

[57]

$$< (\delta\zeta)^2 >= \frac{k_B T}{2\pi\gamma} ln(\frac{L}{l}) \tag{4.6}$$

where γ is the surface tension, L is the upper wavelength limit determined by the size of the surface, and l is the lower limit which is taken to be of the order of a molecular diameter. The surface tension is computed from the molecular and atomic virials as well as from the Irving-Kirkwood formula as will be discussed below.

For the present case, $L = 3.085nm$, l, an estimate of the ester molecular diameter, is taken to be 3.8 angstroms (0.38 nm) (calculated from the geometry with all internals at equilibrium values), $T = 298K$, and the surface tension, γ is that calculated from the molecular virial route, 21.061 dyne/cm. This gives for a rough estimate of the interface broadening:

$$< (\delta\zeta)^2 >= \frac{298.0\, k_B}{2\pi\,(21.061\, dyne/cm)} ln\frac{3.085}{0.38} \tag{4.7}$$

$$< (\delta\zeta) >= 2.5\, angstroms \tag{4.8}$$

This value is within the range of typical values reported in the literature for capillary wave amplitudes of small molecules and oligomeric species [60, 50, 57]. An interesting note is that continuum theory gives an expression for l as [129],

$$l^2 = \frac{\gamma}{(\rho_1 - \rho_2)g} \tag{4.9}$$

where the ρ_i are the densities of the coexisting liquids, and g is the gravitational constant. In the limit of vanishing gravity, we see that the capillary wave amplitude diverges, suggesting that no interface exists in this situation [129].

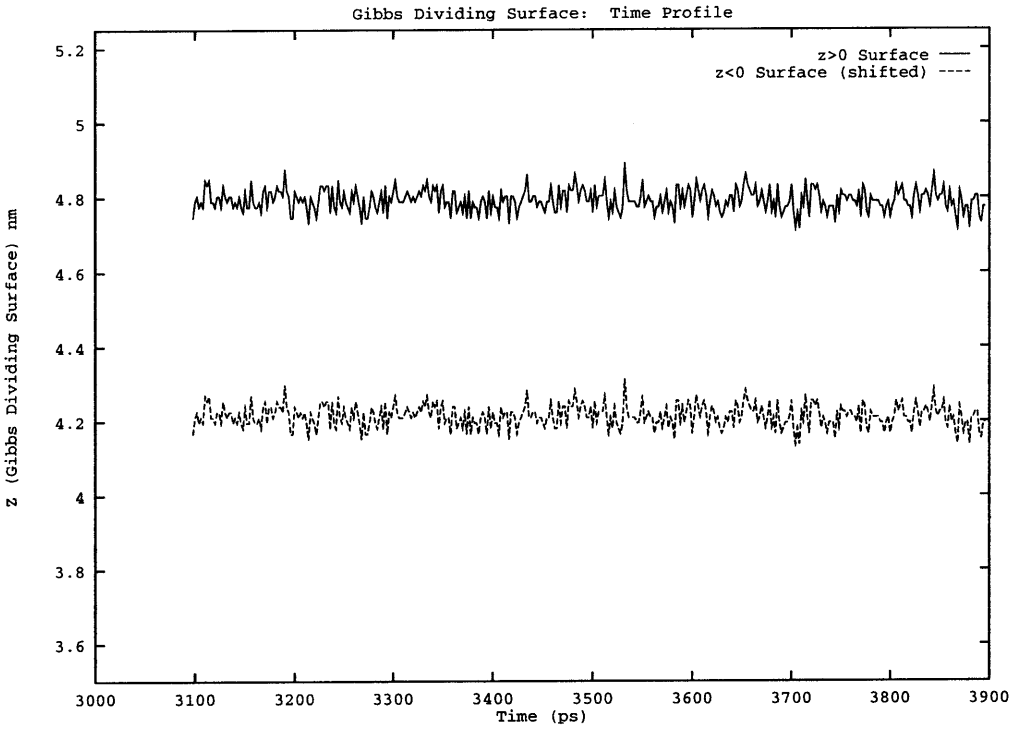


Figure 4-2: Time profile of the positions of the Gibbs dividing surfaces. The Z¡0 profile is shifted up by 9 nm for clarity.

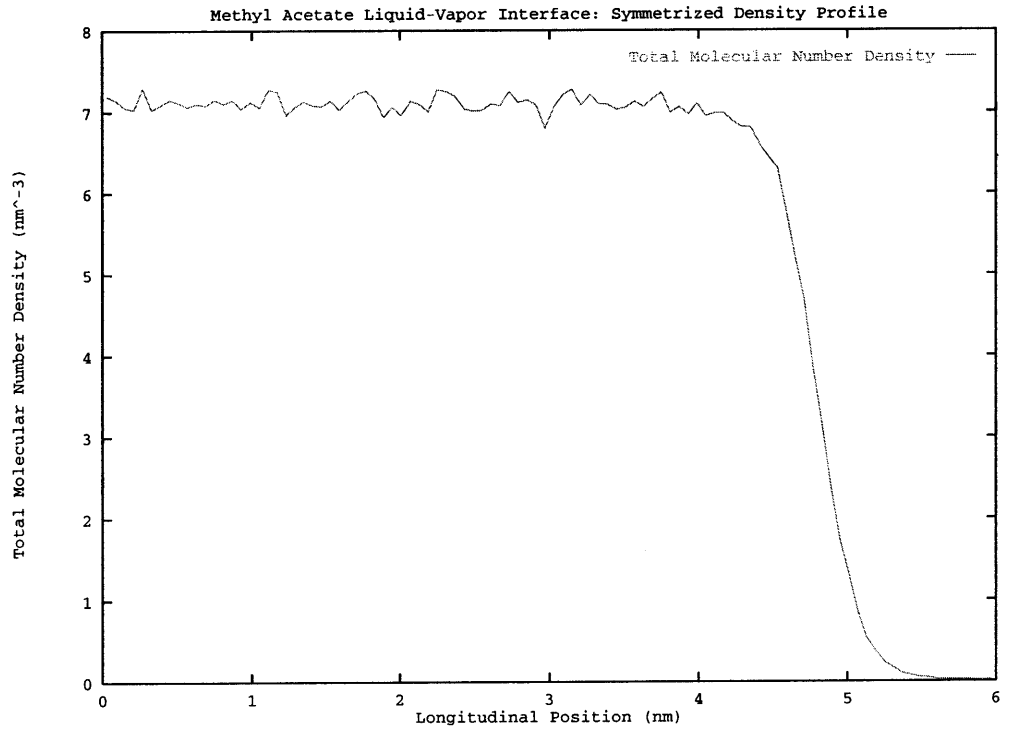


Figure 4-3: Methyl Acetate L-V Interface: Symmetrized Molecular Density Profile

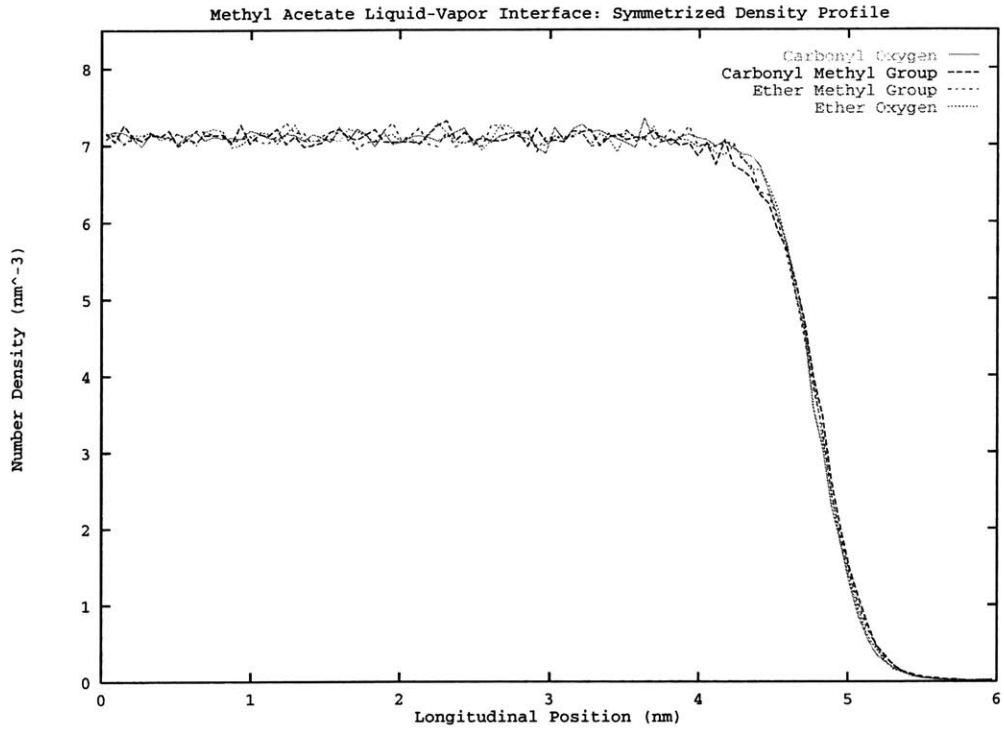


Figure 4-4: Methyl Acetate L-V Interface: Symmetrized Atom Density Profiles

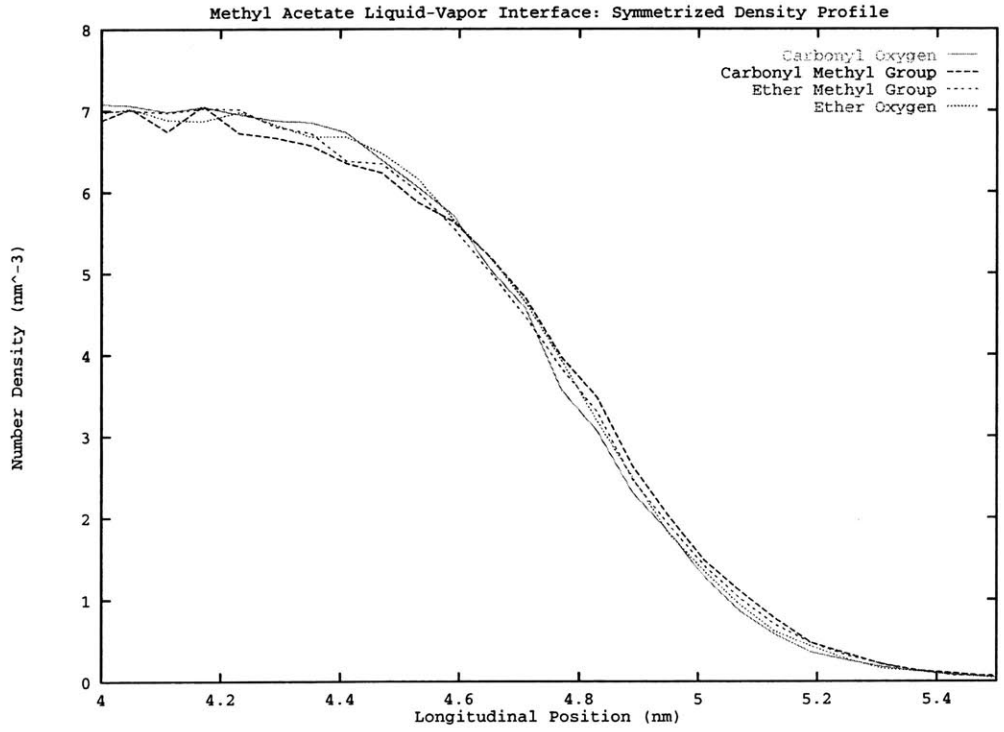


Figure 4-5: Methyl Acetate L-V Interface: Symmetrized Atom Density Profiles

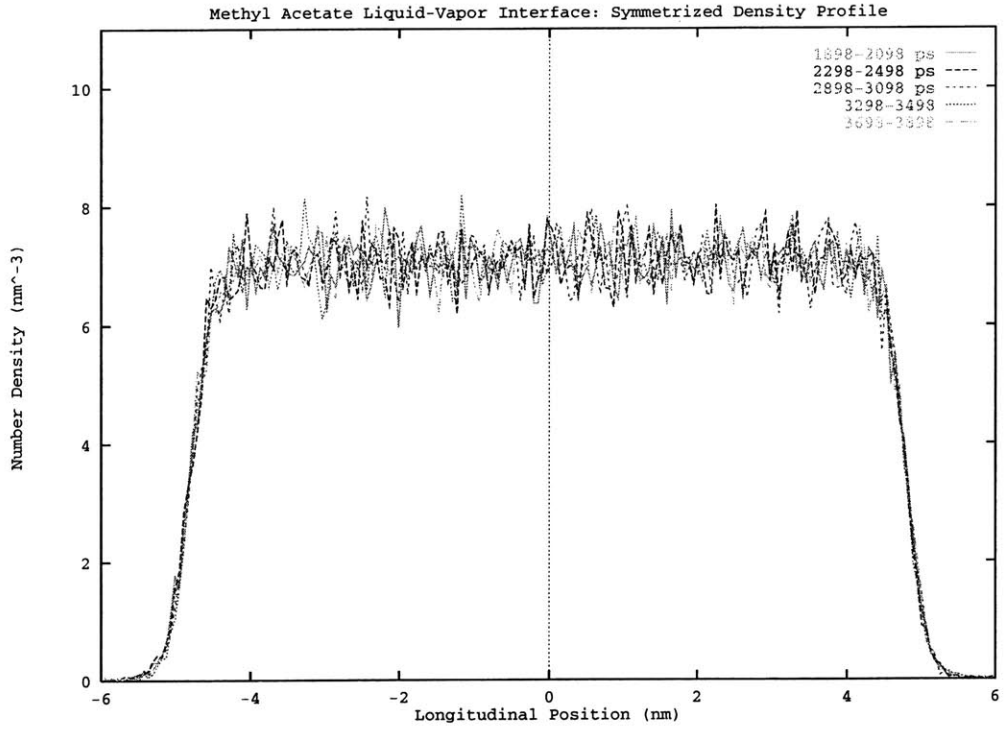


Figure 4-6: Unsymmetrized molecular density profiles. These are averaged over several blocks to show the constancy of the density profile with time.

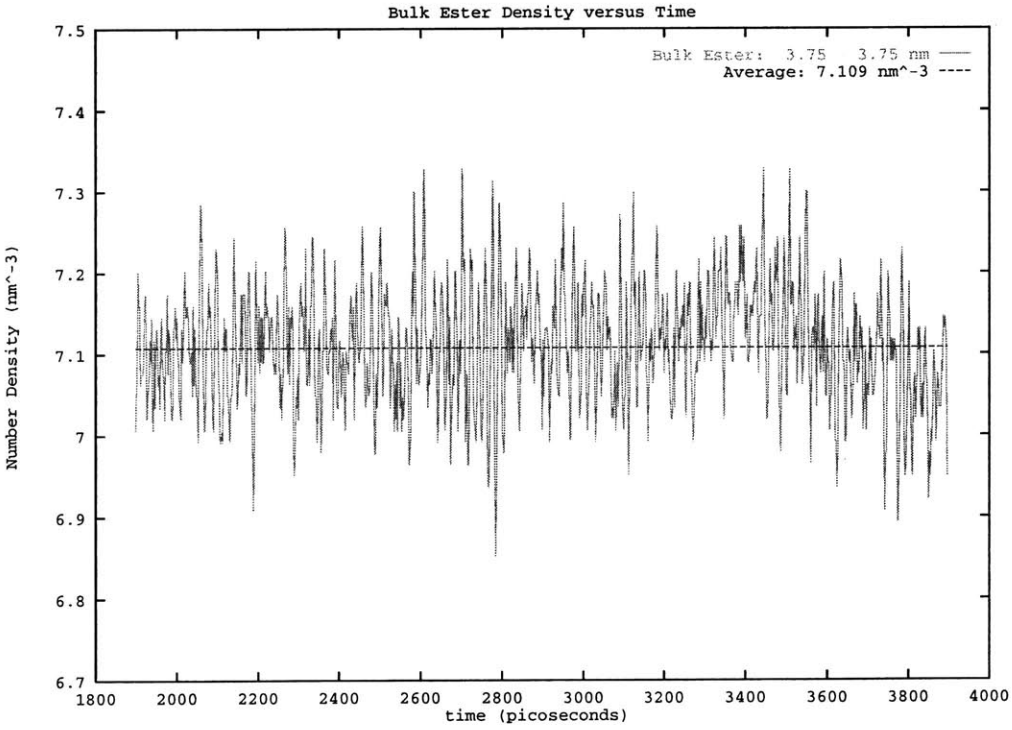


Figure 4-7: Bulk ester density as a function of time. The bulk region is taken to be the middle 7 nm of the simulation cell

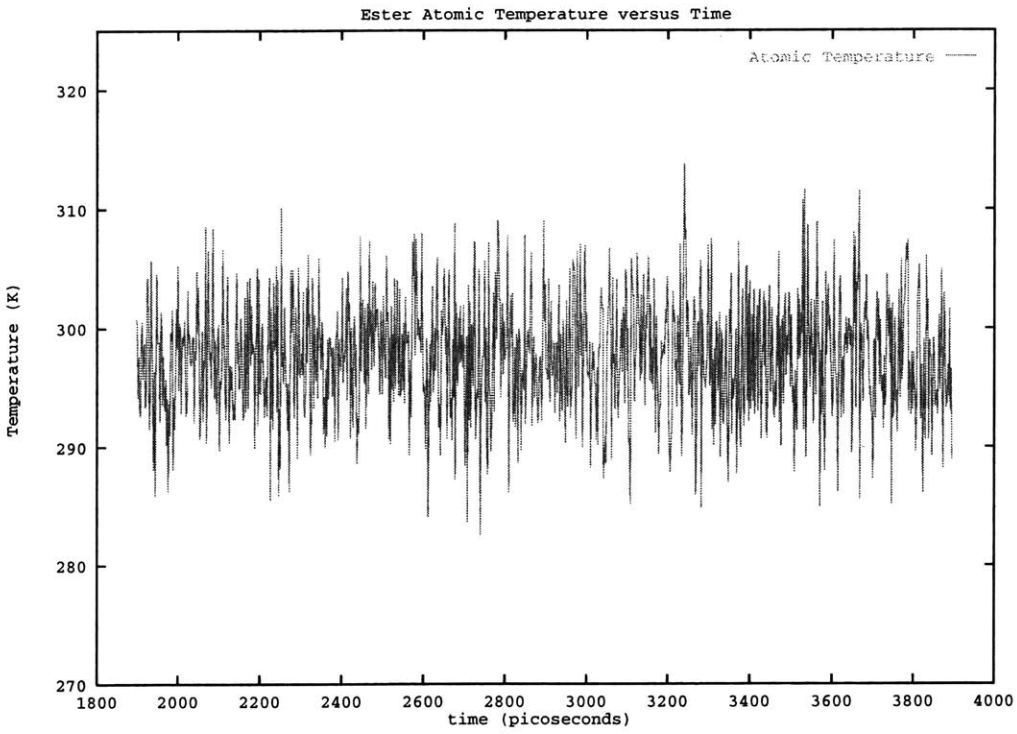


Figure 4-8: Time profile of the ester atomic temperature. The average is computed over all atoms over all configurations.

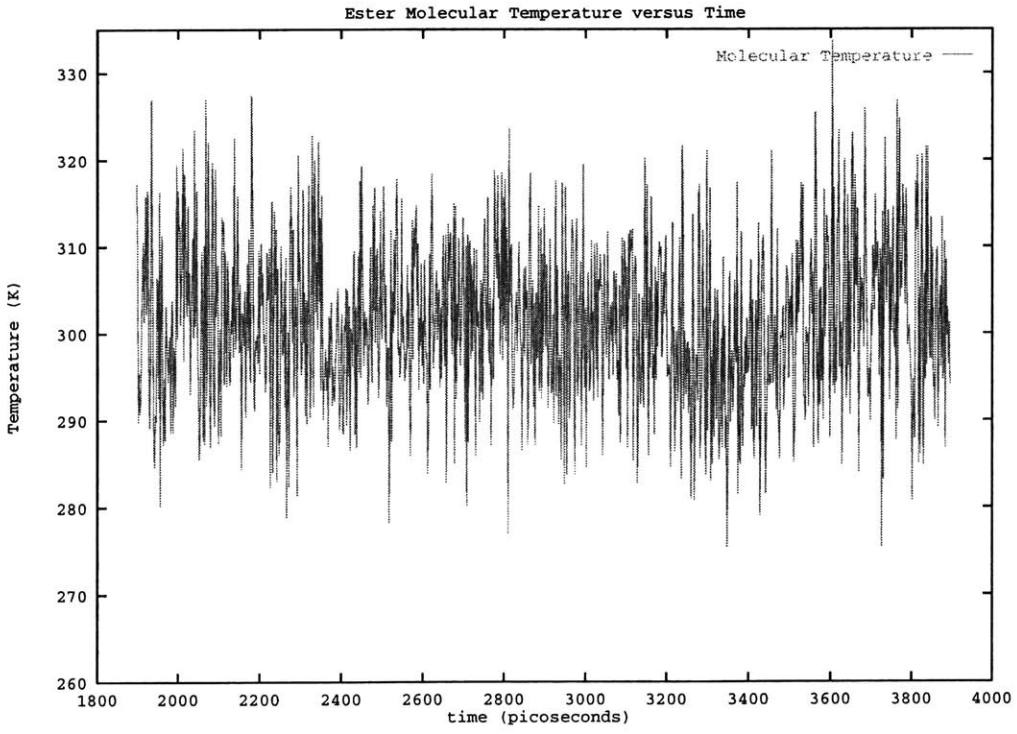


Figure 4-9: Time profile of the ester molecular temperature. The average is computed over all molecules over all configurations.

Capillary Wave Theory and Fluid Interfacial Broadening

This section provides a brief overview of capillary wave theory as applied to fluid-fluid interfaces of molecular fluids; the discussion closely follows the work of Davis [128] which in turn follows the original work of Buff, Lovett, and Stillinger [130]. The notion of capillary waves is invoked to ascribe a density fluctuation to an interface since local fluctuations of density occur at all points of a macroscopic, homogeneous fluid [129]. A simple and natural approach to impose density fluctuations at an interface is to introduce the description of the interface in terms of a spectrum of capillary waves superimposed on a bare, or *intrinsic*, density profile [129]; this will become apparent in the following discussion.

Consider an equimolar dividing surface in the x-y plance, $\zeta(x,y)$, between fluids of density ρ_1 and ρ_2; ζ is a measure of the vertical distance of the interface from the x-y plane which is taken to be the interfacial plane. Now, taking the interfacial tension between the two fluids to be γ_o and considering an external potential (gravitational) of $v(z) = mgz$ which acts to maintain the planarity of the interface [129], the reversible work to create the dividing surface $\zeta(x,y)$ is [128]

$$W_{reversible} = (\rho_1 - \rho_2) \int_{\mathcal{A}} \int \left[\int^{\zeta(x,y)} mgz dz \right] dx\, dy + \gamma_o \int_{\mathcal{A}} \int (1 + \zeta_x^2 + \zeta_y^2)^{1/2} dx\, dy \tag{4.10}$$

where A is the interfacial plane area, and ζ_x and ζy are the x- and y-derivatives of $\zeta(x,y)$, respectively. The gravitational potential energy term in Equation 4.10 is calculated relative to that of a flat (planar) interface at $z = 0$. Furthermore, the system is restricted to be square in the interfacial plane with area, $\mathcal{A} = L^2$, with L being the length of a side of the square [128].

The dividing surface is now represented as a combination of decoupled harmonic surface waves as [128],

$$\zeta(x,y) = \sum_{\mathbf{k}} A(k) e^{i\mathbf{k}\cdot\mathbf{s}} \tag{4.11}$$

$$\mathbf{s} = x\hat{\mathbf{i}} + \hat{\mathbf{j}} \tag{4.12}$$

Note that the sum here is infinite, but will later be bounded based on physical arguments.

Now, substituting Equation 4.11 into Equation 4.10 and expanding $(1 + \zeta_x^2 + \zeta_y^2)^{1/2}$ keeping terms up to second order in ζ_x and ζ_y only (this is because the fluctuations are anticipated to be small in amplitude, or of long wavelength) gives [129, 128],

$$W_{reversible} = \gamma_o \mathcal{A} + \frac{1}{2} \sum_{\mathbf{k},\mathbf{k}'} A(\mathbf{k}) A(\mathbf{k}') \int_0^L \int_0^L e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{s}} \left[(\rho_1 - \rho_2) g - \gamma_o (k_x k_x' + k_y k_y') \right] dx\, dy \tag{4.13}$$

The allowed values of $\mathbf{k}$ are obtained from the periodic boundary conditions imposed in the planar dimensions,

$$\zeta(0, y) = \zeta(L, y) \tag{4.14}$$

$$\zeta(x, 0) = \zeta(x, L) \tag{4.15}$$

The resulting vectors are,

$$\mathbf{k} = \frac{2\pi}{L}(n_x \hat{i} + n_y \hat{j}) \qquad n_x, n_y = 0, \pm 1, \pm 2, \cdots \tag{4.16}$$

Substituting the allowable vectors into Equation 4.13, and noting that only terms with $\mathbf{k} = -\mathbf{k}'$ contribute to the integral over the area gives for the reversible work [128],

$$W_{reversible} = \gamma_o \mathcal{A} + \sum_{\mathbf{k}} |A(\mathbf{k})|^2 \left[\frac{1}{2}(\rho_1 - \rho_2) g + \frac{1}{2} \gamma_o k^2 \right] \mathcal{A} \tag{4.17}$$

For a set of amplitudes $A(\mathbf{k})$, the average value of ζ, $\bar{\zeta}$ is,

$$\bar{\zeta} = \frac{1}{\mathcal{A}} \int_0^L \int_0^L \zeta dx dy = A(0) \tag{4.18}$$

and the mean square fluctuation is,

$$\bar{\zeta^2} = \frac{1}{\mathcal{A}} \int_0^L \int_0^L \zeta^2 dx dy = \sum_{\mathbf{k}} |A(\mathbf{k})|^2 \tag{4.19}$$

Note that as ζ is the fluctuation of a planar dividing surface about $z = 0$, the average value of ζ must vanish, thus implying $A(\mathbf{k} = 0) = 0$; consequently, the sum in Equation 4.17 is taken over non-zero vectors [128].

The probability for an occurrence of a given set of amplitudes $A(\mathbf{k})$ through thermal fluctuations is given by the Boltzmann factor $e^{-\beta W_{reversible}}$, where $\beta = \frac{1}{k_b T}$ (and k_b is Boltzmann's constant), and T is the absolute temperature. The thermal (canonical) average of the mean square amplitude computed by averaging over all sets of amplitudes, $A(\mathbf{k})$, gives the mean-square capillary wave dispersion [128],

$$\sigma_\zeta^2 \equiv \left\langle (\bar{\zeta^2} - \left\langle \bar{\zeta} \right\rangle^2) \right\rangle = \sum_{\mathbf{k}>0} \left(\beta L^2 \left[\frac{1}{2}(\rho_1 - \rho_2)g + \frac{1}{2}\gamma_o k^2 \right] \right)^{-1} \tag{4.20}$$

The sum over $\mathbf{k}$ is transformed to an integral using the relation $1 \equiv \delta n_x \delta n_y = (L/2\pi)^2 \delta k_x \delta k_y$ to obtain

$$\sigma_\zeta^2 \simeq \frac{1}{(2\pi)^2 \gamma_o \beta} \int \int (\zeta^2 + k^2)^{-1} d^2 k \tag{4.21}$$

$$\sigma_\zeta^2 = \frac{1}{4\pi\gamma_o\beta} ln \left(\frac{\xi^2 + (2\pi/l_m)^2}{\xi^2 + (2\pi/L)^2} \right) \tag{4.22}$$

where

$$\xi^2 \equiv \frac{(\rho_1 - \rho_2 g}{2\gamma_o} \tag{4.23}$$

Note that the sum in Equation 4.20 is bounded by a minimum and maximum value for the k-vectors. The sum begins at $(2\pi/L)$, the lowest value of k. The maximum value is determined by the physical dimensions of the system; that is, for

a wave to be defined in a molecular system, the upper bound for k is $2\pi/l_m$, where l_m is a minimum wavelength definable for the interface. Although in the original proposition of capillary wave theory, Buff *et al* took the value of l_m to be the mean thickness $< \zeta^2 >^{1/2}>$, an equivalent choice is a length of the order of hte diameter of a molecule, and thus on the order of the mean nearest neighbor separation in the liquid [128].

One can see from Equation 4.22 that in the thermodynamic limit, $L \to \infty$, the mean square capillary amplitude is [129],

$$\sigma^2 = \frac{1}{4\pi\gamma_o}\beta ln\left(1 + (2\pi/l_m)\frac{2\gamma_o}{(\rho_1 - \rho_2)g}\right) \tag{4.24}$$

As the gravitational potential diminishes to zero, the capillary wave amplitude diverges (in the thermodynamic limit) as $(-ln\, g)^{1/2}$ [129].

For a system of finite size (typically which is what one is limited to in a simulation), the maximum value of the capillary amplitude is [128]

$$\sigma_{max} = \left[\frac{1}{2\pi\gamma_o\beta} ln\left(\frac{L}{l_m}\right)\right]^{1/2} \tag{4.25}$$

This is the zero gravity value of the amplitude for a finite sized system; notice that in this case, the finiteness of the system suppresses the divergence as zero gravity is approached [128].

4.3.2 Bulk vs. Interfacial Structure I. Radial Distribution Functions

The pair correlation function for a slice is calculated from the histogram of the separations between all pairs of particle types within the slice. The normalized pair correlation function is given by [131] :

$$g(R) = \frac{2N(R, \Delta R)}{V_{int} N_T^2} V_s \tag{4.26}$$

where $N(R, \Delta R)$ is the average number of pairs of particles within the section with

separation between R and $R + \Delta R$, V_{int} is the average volume of the intersection of the section with the space between spheres of radii R and $R + \Delta R$, V_s is the total slice volume, and N_T is the average number of particles in the slice.

Figure 4-10 through Figure 4-14 show the radial distribution functions computed for various site-type pairs, these specifically being the carbonyl oxygen to carbonyl oxygen rdf (O-O), the ether oxygen to ether oxygen rdf(Oe-Oe), the carbonyl oxygen to ether oxygen(Oe-O) rdf, the carbonyl oxygen to carbonyl methyl group rdf(O-MeE), and the ether oxygen to carbonyl methyl group rdf (Oe-MeE). The radial distribution profiles are for the central bulk region defined as that in which $-3.00nm \leq R_z \leq 3.00nm$. For the interfacial region, $4.4nm \leq R_z \leq 5.3nm$ and likewise for the mirror side of the simulation cell; the symmetry of the system is taken advantage of for averaging purposes, and the profiles for the interface are averaged over the symmetric halves of the cell. The profiles are computed in 100 bins out to a separation of 1.2 nm (which was the spherical cutoff used for non-bond interactions). Averages are from the last 1.2 nanoseconds of the simulation (production phase).

All the bulk radial distribution functions show relatively structureless homogeneous fluids, a result which is in keeping with the rdf's calculated by Jorgensen *etal* based on NPT Monte Carlo computations of bulk methyl acetate at 298K and 1 atmosphere [67]. Although the density reported by Jorgensen *et al* is 0.905 gr/cm^3 compared to the 0.874 gr/cm^3 value obtained for the bulk ester region, the present simulation does not show any fundamental difference in the structure of the bulk fluid. One can note that based on the first peak in the functions for the carbonyl methyl group, there is a tendency for the more electronegative carbonyl oxygen to attract the carbonyl methyl group more so than the ether oxygen does.

Finally, comparing the bulk and interfacial pair distribution functions, it is evident that the liquid structure is not lost as one moves from the bulk to interface. The peaks in general are higher due to the lower density of the interface.

At this point, some comments are made about the behavior of the radial distribution functions as the interface is approached from the bulk region. Figures 4-10 through 4-14 show an increasing first peak in pair correlation functions with de-

creasing density moving from the bulk to the interface (note that for such a rapidly decreasing density profile, one must keep in mind the nominal density of the region being considered).

This same effect is seen for two water models, the ST2 and LSR (Lemberg-Stillinger- Rahman) potentials, applied to the study of clusters of water molecules by Townsend and Rice [53]. The auth ors present oxygen-oxygen, oxygen-hydrogen, and hydrogen-hydrogen radial distribution functions for shells encompass ing a bulk region and various positions within the inhomogeneous region (the interfacial region characterized by monotically decreasing density profiles). For both po tentials studied by the authors, the oxygen-oxygen correlations became stronger as the density decreased, with a threshold for increasing amplitude of the first peak of the oxygen-oxygen distribution being about ninety percent of the bulk density. Furthermore, the second peak of the oxygen-oxygen distribution is independent of the density, showing a slight enhancement for densities less than 25 percent of the bulk. The oxygen-hydrogen and hydrogen-hydrogen functions are shown to parallel the oxygen-oxygen functions. Based on these results, the authors claim that the behavior of the distribution functions suggest that with a breakup of the hydrogen-bonded structure of the bulk liquid in the lower density interfacial region, there is precipitated dimer formation; it is this dimerized structure which yields the stark enhancement in the first peaks of radial distribution functions.

For methyl acetate, the radial distribution functions presented seem to indicate a similar phenomenon; whether this is a result of 'dimerization' or some higher-order 'clustering' is not clear from the data shown. Based on Monte Carlo simulations of bulk methyl acetate [67], one does not observe strong structure in the bulk liquid (as demonstrated by the relatively featureless rdf's shown for the bulk region). Furthermore, the Monte Carlo simulations did not reveal any strong interaction-induced structure as in the case of acetic acid, for which the bulk fluid structure incorporated strong hydrogen-bonding interactions. At this point, we suggest the possibility of some type of clustering occuring at the liquid-vapor interface. From an energetic argument, this may be the result of two (or more) ester molecules, within the low-

density region of the interface, coming together and proceeding along a trajectory for some characteristic time after which the 'cluster' breaks up. With the current potential model, it may be that the energetic gains compensate for the overall entropic loss of forming the cluster. To more accurately analyze the results of the pair correlation function calculations, we see the need to determine first, the nature, if any, of the clustering at the interface, and second, the dynamical aspect of the association in terms of the mean time over which the cluster remains a stable entity.

One caveat to this analysis is the fact that there is an inherent limitation in the nature of the calculation of the *interfacial* radial distribution functions. One can note that the density over a small section within the monotonically decaying section of the density profile is dramatically changing. Now, the calculation of the rdf for this region requires normalizing the histogram of pair separations; this is done by employing a *uniform density* for this region. In so doing, the smaller pair separations are weighted more, and thus, one observes an enhancement in the first peak of the rdf in the interfacial region relative to the bulk. This effect becomes more pronounced as the degree of inhomogeneity within the region becomes more severe, since this causes the uniform density used for normalization to decrease substantially. This is a topic for further investigation and will not be considered further in this work; it would be interesting to reinvestigate the phenomenon of water clustering at the free surface (based on the previous studies which seem to show very similar results in terms of the rdf behavior) with the intent to consider if this effect is indeed 'real'.

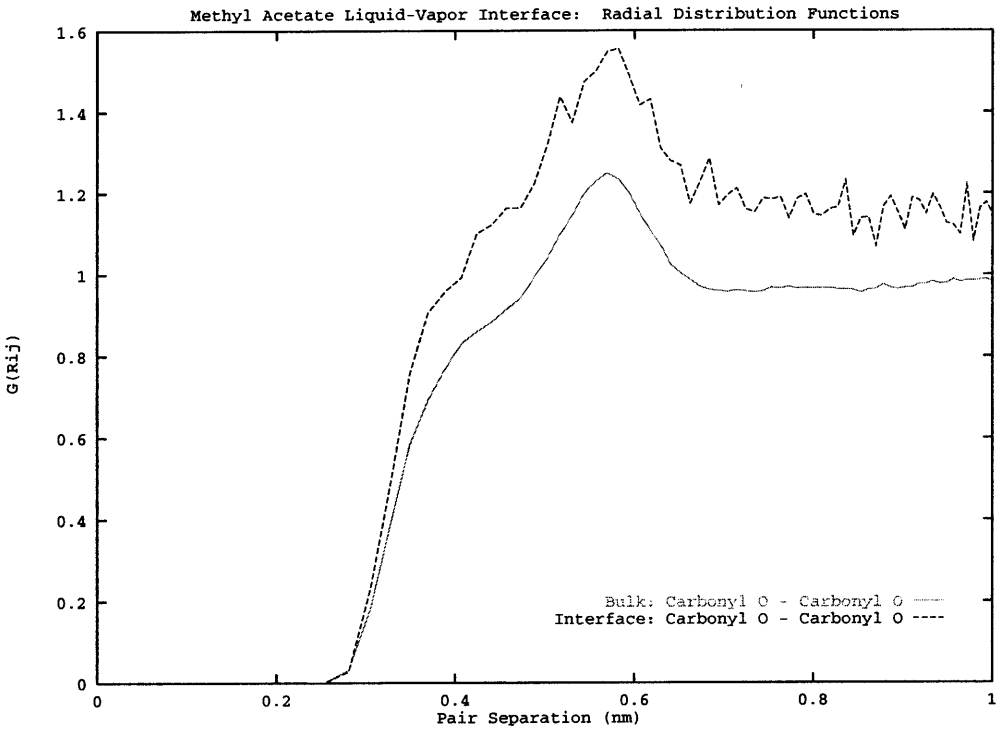


Figure 4-10: Methyl acetate carbonyl oxygen / carbonyl oxygen radial distribution functions. The bulk region is taken to be the central 6.0 nm of the simulation cell. The interfacial distributions are computed for a region from 4.4 to 5.3 from the center of mass of the simulation box (in both the positive and negative directions); the two profiles are then averaged.

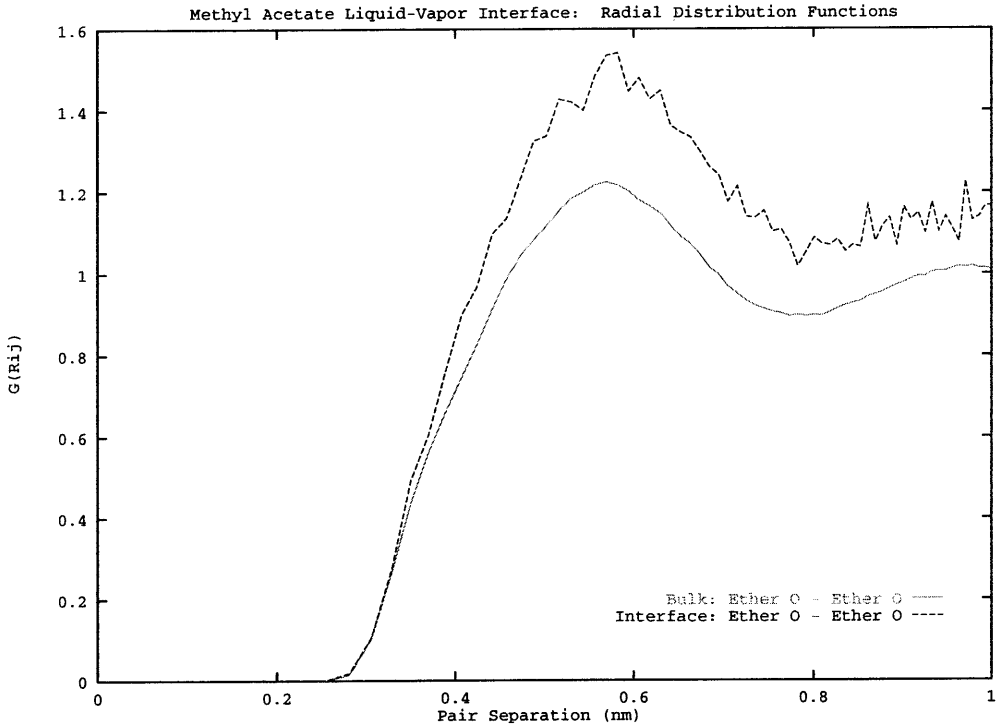


Figure 4-11: Methyl acetate ether oxygen / ether oxygen radial distribution functions (see carbonyl oxygen / carbonyl ogygen rdf for details)

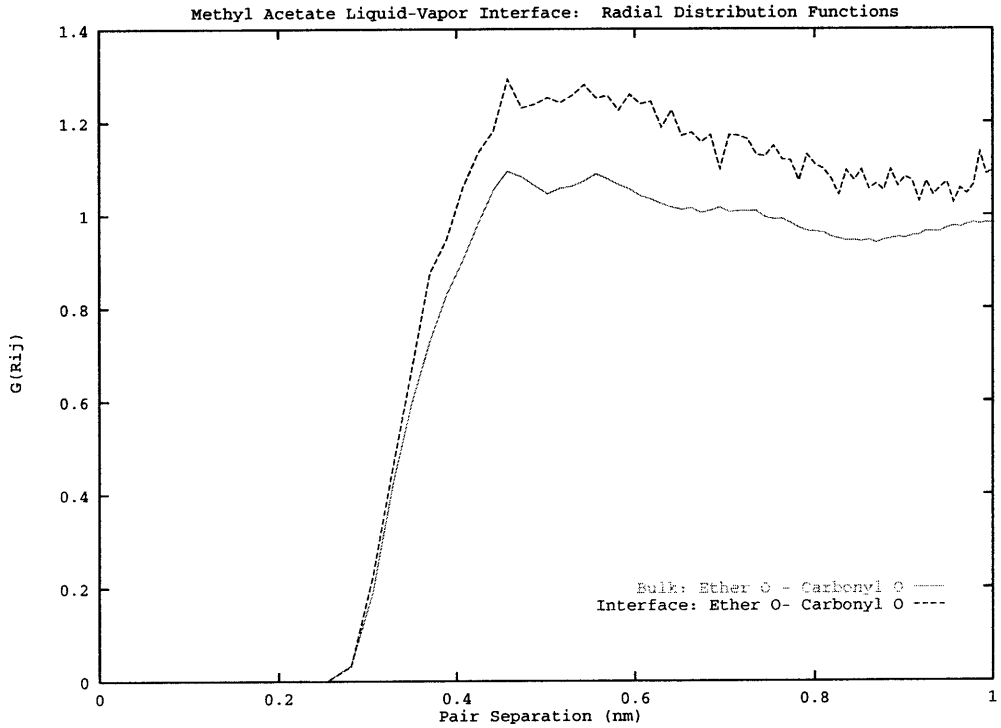


Figure 4-12: Methyl acetate ether oxygen / carbonyl oxygen radial distribution functions (see carbonyl oxygen / carbonyl oxygen rdf for details

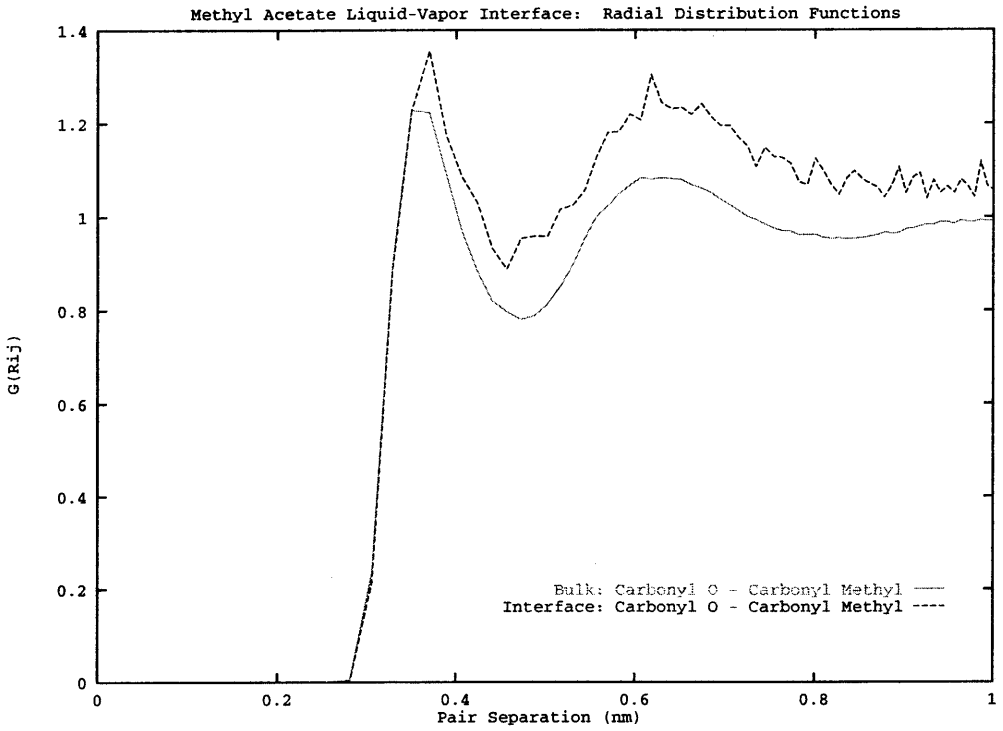


Figure 4-13: Methyl acetate carbonyl oxygen / methoxy methyl radial distribution functions (see carbonyl oxygen / carbonyl oxygen rdf for details)

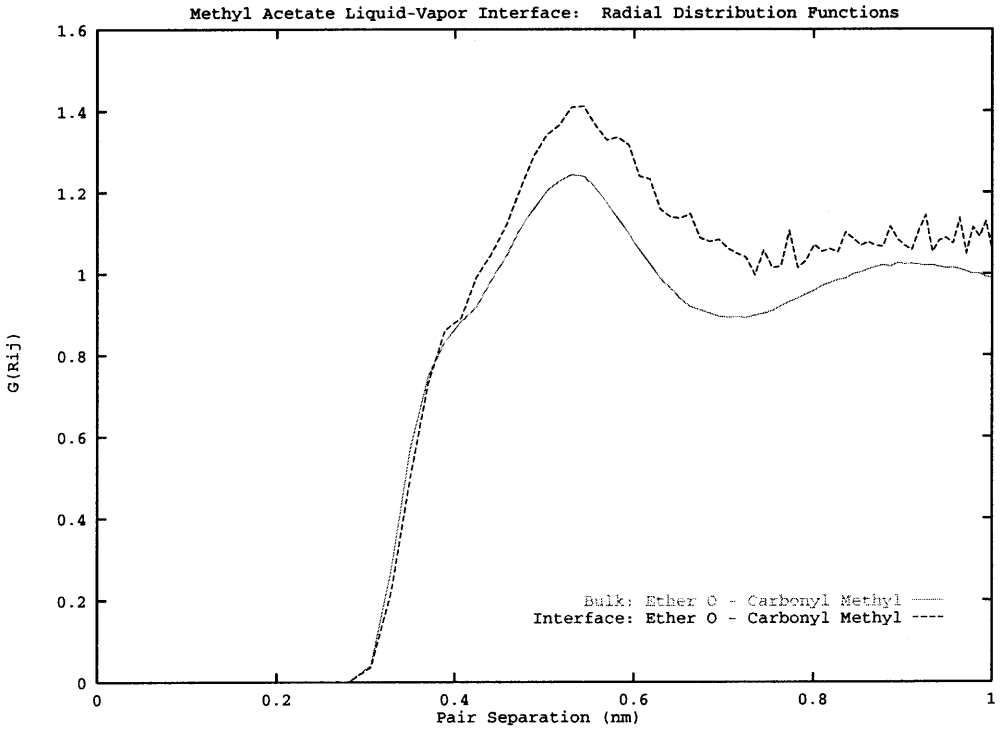


Figure 4-14: Methyl acetate ether oxygen / methoxy methyl radial distribution functions (see carbonyl oxygen / carbonyl oxygen rdf for details)

4.3.3 Bulk vs. Interfacial Structure II. Torsion and Out-of-Plane Distributions

Figure 4-16 shows the distribution of the out-of-plane angle averaged over all molecules in the bulk region as defined above. Also, shown is the distribution for the interfacial region. As expected, the distributions are sharply peaked around zero, signifying a flat trigonal center at the carbonyl carbon atom. Furthermore, the distribution of the central torsion angle is seen to be peaked again at zero degrees in Figure 4-15; this coupled with the out-of-plane angle distribution confirms the average planar molecular structure of the molecule. This is useful in interpreting the orientational order parameters in terms of the spatial relation of the various bond vectors and how this relation is part of the structure shown by the order parameters.

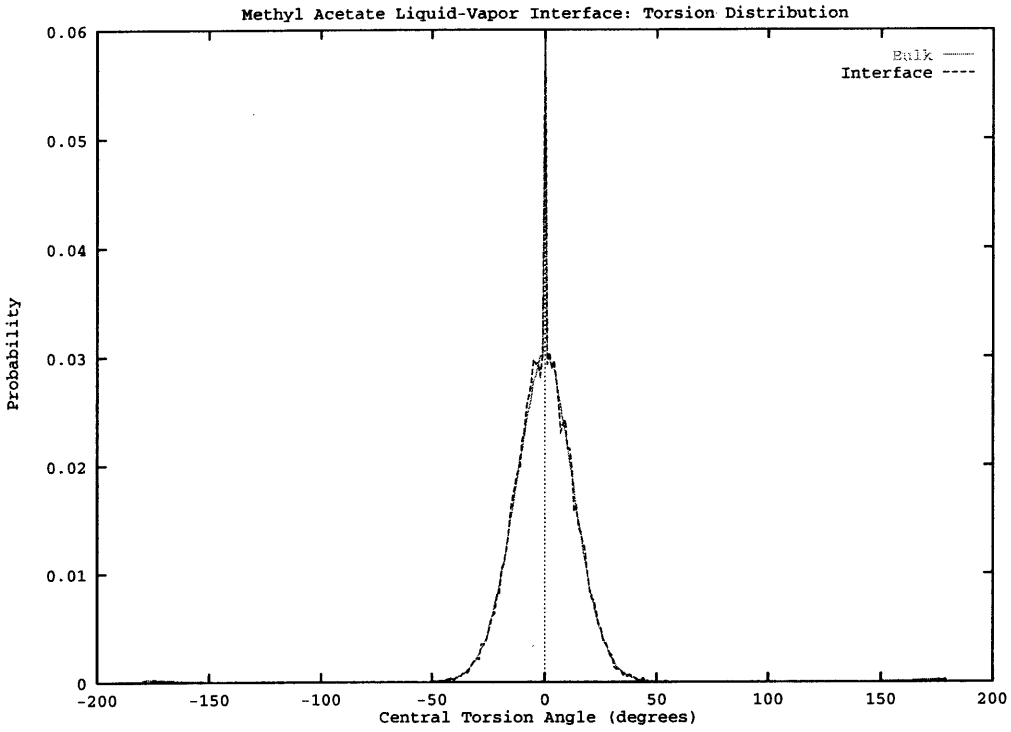


Figure 4-15: Torsion angle distribution for the bulk and liquid-vapor interface.

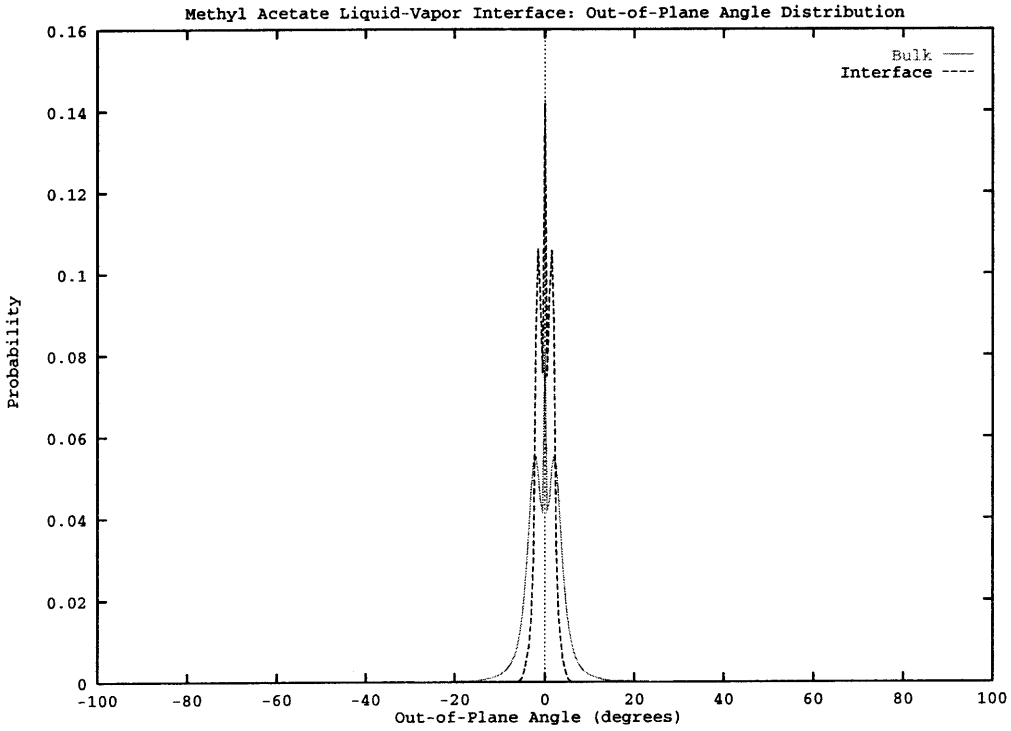


Figure 4-16: Out-of-plane angle distribution for the bulk and liquid-vapor interface.

4.3.4 Bulk vs. Interfacial Structure III. Orientational Order Parameters

The effect of the interface on the orientations of various bond vectors is studied through orientational order parameters (first and second Legendre coefficients) defined as [60, 50]:

$$P_1(z) = < \cos\theta > \tag{4.27}$$

and

$$P_2(z) = \frac{1}{2} < 3(\cos^2\theta) - 1 > \tag{4.28}$$

where θ is defined as the angle between the interface normal (along the positive and negative z-axes) and a bond vector associated with a specific functional group. For the present case, the carbonyl carbon - carbonyl oxygen bond vector, r_{CO}, the carbonyl carbon - methyl bond vector, r_{C-CH3}, and the e ther oxygen - methyl group bond vector, r_{O-CH3}, are considered for methyl acetate. The average is over all vectors and time steps within a thin slab perpendicular to the interface normal. If the vectors in a region are completely randomly oriented, $P_2 = 0$. If there is a prevalent orientation parallel to the interface normal, $P_2 = 1.0$; with an orientation perpendicular to the normal, the order parameter goes as -0.5. Note that the bond vector is defined as $r_{ij} = r_i - r_j$. Thus, for methyl acetate, the relevant bond vectors are such that the r_{CO} vector points from the carbonyl oxygen to the carbonyl carbon, the r_{O-CH3} vector is directed from the methoxy methyl group to the ether oxygen , and the r_{C-CH3} vector points from the carbonyl methyl group to the carbonyl carbon.

Figure 4-17 shows symmetrized P1 profiles for the noted bond vectors. In the bulk region, there is virtually no preferential orientation of the bond vectors. In theory, this should be zero, but the fluctuation is a result of the finite sample size. Note that in the bulk region, the oscillatory behavior of the carbonyl methyl group - carbonyl carbon bond vector is out of phase relative to the other two profiles. This is in

keeping with the observation that the predominant (essentially the only) conformer present in the bulk neat liquid is the syn periplanar form in which the carbonyl methyl group is trans to the ether methyl group. With this molecular geometry, one anticipates the observed behavior. Interesting is the dramatic divergence in the profiles once the interfacial region is encountered beginning around 4 nanometers out from the bulk center. Thus, we see that the anisotropic field generated simply by the heterogeneity of the density gives rise to an order of magnitude change in the orientational preference. The P1 profile shows that in the interface both the C-O and Oe-CH3 bond vectors tend to point inwards to the bulk, while the carbonyl methyl group juts towards the vapor phase. This tendency is acceptable from an energetic perspective as the electrostatic interaction loss is minimized as only the less electropositive carbonyl methyl group (0.05e as opposed to the ether methyl group with charge of 0.25e) is forced to give up interactions. This is also consistent with the results of MD simulations of the methanol liquid-vapor interface presented by Matsumoto and Kataoka [56]. In their study to address the anamolous behavior of strongly hydrogen-bonding molecules at the liquid-vapor interface, the authors report the evolution of a strong tendency of the methanol molecule to show orientational ordering at the interface, even more so than water. Based on distributions of internal angles, the authors conclude that a methanol molecule near the interface prefers to project the methyl group towards the vapor phase. The physical reasoning for this phenomenon is that the methyl group, which is essentially a hydrophobic element unable to participate in hydrogen bonding, is pushed towards the vapor phase so that energetically, the liquid phase is stabilized through more opportunities for hydrogen bonding; the authors thus claim that in this sense, methanol can be regarded as one of the simplest representations of surfactant molecules—although this is by no way a rigorous use of the definition of surfactants, the manifest phenomenon is the same. Now, Matsumoto and Kataoka use Jorgensen's TIPS united atom model for methanol in which the methyl group is attributed a charge of $0.285e$ (with e being the elementary charge); for the methyl acetate OPLS model, the carbonyl methyl group, as noted, has a partial charge of $0.05e$. Thus, although the ester system is not a hydrogen bonded

system, the fundamental mechanism for the orientation, which is the electrostatic interaction which contributes more to the energetic stabilization of the liquid, is still at work. Furthermore, we note that the authors of the previous study use Ewald summation for tracking the electrostatic energy and forces; the present study uses a much simpler (both conceptually and programming-wise) reaction field. In effect, both approaches seem to indicate the same phenomena at the interface. This is not altogether unexpected since the orientation is a local event, basically confined to the inhomogeneous region. Thus, we expect that both approaches should yield the same physics, as they apparently do.

Figure 4-18 shows the P2 profiles for the relevant bond vectors. This figure shows that the distributions of the cosine of the angle are diffuse about the average value indicated by P1. There is a slightly favored orientation along the interface normal, but within the precision of the simulation, this is a small effect. In the very outer regions of the interface, the fluctuations for the individual bins become very large as the number of molecules present in the vapor is insignificant to accumulate sufficient statistics over the simulation time covered in the present work.

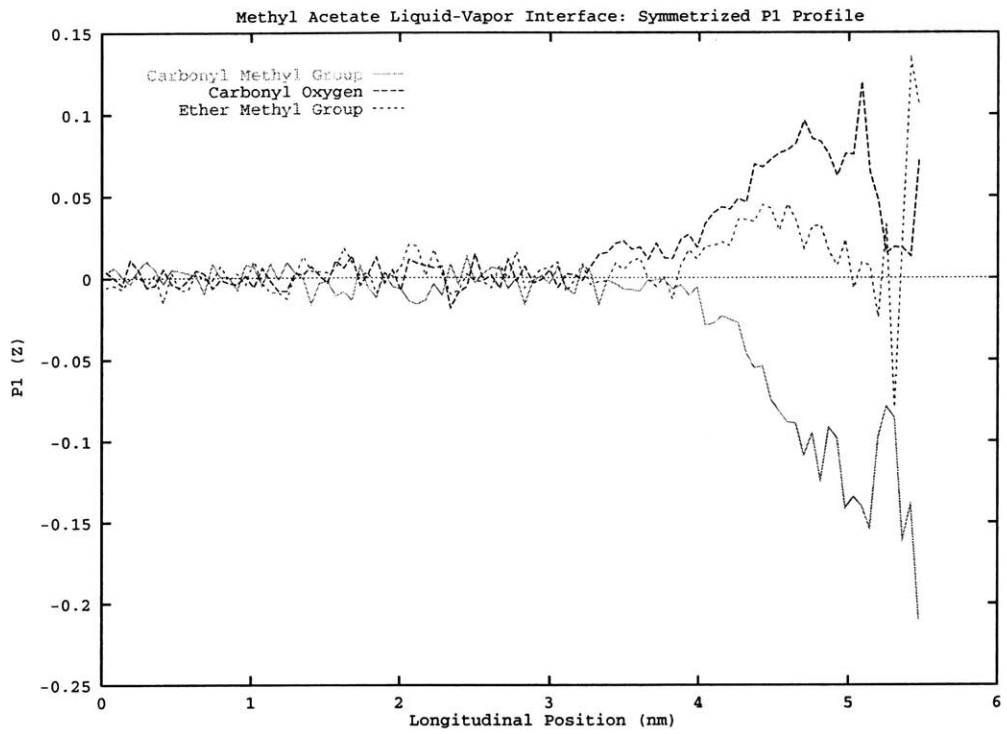


Figure 4-17: Symmetrized P1 profiles for the indicated bond vectors of methyl acetate. The longitudinal position is relative to the center of mass of the simulation cell.

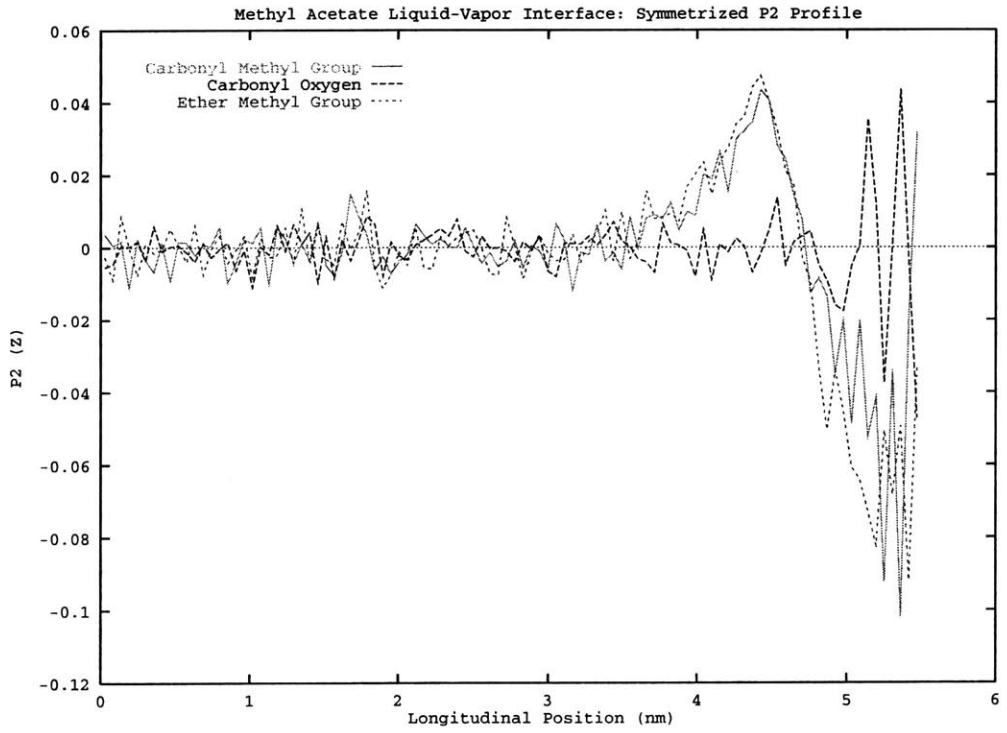


Figure 4-18: Symmetrized P2 profiles for the indicated bond vectors of methyl acetate. The longitudinal position is relative to the center of mass of the simulation cell.

4.4 Surface Tension

In order to evaluate the accuracy of the OPLS ester potential parameters for interfacial thermodynamics, the molecular virial is computed to obtain the molecular surface tension via [60, 126]:

$$\gamma = \frac{1}{2A}(2V_{zz} - V_{xx} - V_{yy}) \tag{4.29}$$

where $V_{\alpha\alpha}$ is the diagonal molecular virial component and $A = 2L_xL_y$ is the total surface area of the slab. The molecular virial is:

$$V_{\alpha\alpha} = \sum_{a<b} F_{ab}^{\alpha} R_{ab}^{\alpha} = \sum_{a} F_{a}^{\alpha} R_{a}^{\alpha} \tag{4.30}$$

where the sum is over all molecules a and b and R_{ab}^{α} and F_{ab}^{α} are the $\alpha = x, y, orz$ component of the molecular center of mass separation and total force,

$$F_{ab}^{\alpha} = -\sum_{k,l=1}^{m} \frac{r_{abkl}^{\alpha}}{|r_{abkl}|}\phi'(r_{abkl}) \tag{4.31}$$

An alternate and theoretically equivalent route to the surface tension is the atomic virial [36],

$$V_{\alpha\alpha} = \sum_{a}\sum_{k} f_{ak}^{\alpha} r_{ak}^{\alpha} \tag{4.32}$$

where the f's and the r's now indicate the forces acting upon, and the positions of, the individual atoms.

The atomic virial includes contributions from intramolecular interactions such as constraint forces, bond angle forces, torsional forces, intramolecular Lennard-Jones type forces, and bond-stretching forces (although not in the present case). Although the total contribution from these sources is in theory zero, the contributions to the individual components of the pressure tensor do not necessarily cancel, thus giving non-zero values for these components. Note that although the molecular and atomic virials should result in equivalent ensemble average values of the surface tension, the

instantaneous values and fluctuations can differ dramatically [36]. The atomic and molecular virials do yield different values for the individual $V_{\alpha\alpha}$, but the atomic and molecular kinetic contributions to the total pressure tensor components cancel this difference.

Furthermore, the differential between the normal and tangential components of the molecular pressure tensor is computed as a function of longitudinal position. This is done as a check on equilibration as the normal and tangential components should be equivalent away from the interfaces and the normal component should be constant throughout the system. Furthermore, analyzing the components of the pressure tensor acts a check on the 'appropriateness' of the potential parameters. If the potential model is correct at the state point under study, the average normal component of the pressure tensor (normal to the interface) should essentially be atmospheric pressure (and constant throughout the longitudinal profile), and the tangential (in-plane) components give rise to the surface tension which can be compared to experimental data. The off-diagonal elements of the pressure tensor should be zero (or 'small'). It is im portant to note that simulation surface tensions for systems containing charges or hydrogen-bond interactions are plagued by large errors; surface tension estimation even for simple alkane oligomers is subject to large inaccuracies. Thus one does not anticipate great accuracy in the surface tension computed. However, it is reiterated that analyzing the pressure tensor components does have merit as a check on equilibration and consistency of potential model [61].

The profile of the normal-tangential pressure differential is computed as:

$$\gamma(z) = \frac{1}{2}\int_{-\infty}^{z} dz'(P_N(z) - P_T(z)) \tag{4.33}$$

where $P_N(z)$ and $P_T(z)$ are the normal and tangential components of the stress across a plane parallel to the interface at the longitudinal position z. $\gamma(\infty)$ is the surface tension.

The normal and tangential components are computed via the Irving-Kirkwood definition of the pressure tensor [60, 132, 126]

$$P_T(z) = < \rho(z) > kT - \frac{1}{A} < \sum_{j>i}^{N} \sum_{a}^{m} \sum_{b}^{m} \frac{(x_{ij}x_{iajb} + y_{ij}y_{iajb})}{2r_{iajb}} \frac{dU(r_{iajb})}{dr_{iajb}} \frac{1}{|z_{ij}|} \theta(\frac{z - z_i}{z_{ij}}) \theta(\frac{z_j - z}{z_{ij}}) > \tag{4.34}$$

and

$$P_N(z) = < \rho(z) > kT - \frac{1}{A} < \sum_{j>i}^{N} \sum_{a}^{m} \sum_{b}^{m} \frac{z_{ij}z_{iajb}}{r_{iajb}} \frac{dU(r_{iajb})}{dr_{iajb}} \frac{1}{|z_{ij}|} \theta(\frac{z - z_i}{z_{ij}}) \theta(\frac{z_j - z}{z_{ij}}) > \tag{4.35}$$

In these equations, $\rho(z)$ is the total molecular density at z (its average over all configurations is used), k is Boltzmann's constant, and T is the absolute temperature. $\theta(q)$ is the Heavyside step function which is zero for arguments less than zero and equal to unity for arguments greater than or equal to zero. For clarification, the x_{ij} , y_{ij} , and z_{ij} are the molecular center of mass separation components, while the x_{iajb} , y_{iajb} , and z_{iajb} are the site-site separation components. The sums are over all N molecule pairs and over the m sites on each molecule (here, all molecules are taken to be equivalent).

The molecular virial route to the surface tension yields a value of 21.06 (0.87) dyne/cm which is about 17 percent lower than the experimental value of 24.73 dyne/cm at 298 K. The atomic virial calculation yields a value of 22.70 (1.09) dyne/cm. Both computed values exclude long-range corrections beyond the cutoff for Lennard-Jones interactions. Although there is a slight difference between the two values, note that both yield a consistent measure of the standard deviation of the virial (pressure) which is related to the compressibility. Furthermore, the error estimates imply that the distribution of atomic surface tension values is more diffuse relative to the molecular distribution for a given sampling frequency. This trend is also observed in the computation of surface tensions for eicosane at 400 K [60]. As to the accuracy of the potential in predicting surface tensions, it is noted that although the uncorrected surface tension values underestimate the experimental data, trends indicate that the corrections simply for the Lennard-Jones contributions can amount to almost half of

the total value [60], and thus lead to an overestimation of the surface tension. Surely for the present case, a correction of such magnitude will not prove any better.

Figure 4-20 shows the longitudinal profile of the differential of the normal and tangential components of the pressure tensor. The pressure components are in program units of $kJ/molcm^3$. There are two clearly defined interfacial regions giving roughly the same contribution to the stress differential. Although the bulk region does not show that the differential identically vanishes in the average sense, we accept this to be a limitation of the statistics accumulated during the simulations. The sampling frequency for configurations (every 500 time steps) is too low to provide any better precision thus resulting in the deviations from zero average pressure differential in the bulk.

Figure 4-19 shows the integrated profile of the differential of the normal and tangential components of the pressure tensor. Although there is a non-uniformity in the bulk region, the two interfaces essentially contribute equally to the surface tension. Also, the surface tension obtained from this curve is $22.852 dyne/cm$ which is well within the error bounds of the surface tension reported above calculated from the atomic virial. The level of precision indicated by this curve is consistent with the results of the virial calculations of the surface tension presented above.

4.5 Conclusions

We have performed direct molecular dynamics simulations of the methyl acetate liquid-vapor interface using the OPLS forcefield augmented with bond angle and out-of-plane potentials. Our computations indicate the evolution of a sharp, stable interface with a 10-90 thickness of 6 angstroms. This is characteristic of the interfaces of small molecules studied by simulation techinques, although somewhat larger as compared to hydrogen-bonding systems. In terms of liquid structure, we observe that the bulk fluid is rather structureless (as has been reported earlier based on Monte Carlo simulations) and retains the same structure within the interfacial region. Molecular and atomic species density profiles indicate a layering effect in

the interfacial region, with the carbonyl oxygen atoms being enriched in a layer just beneath the outermost surface regions, e.g., that closest to the vapor phase. Consistent with this picture, we observe that the P_1 orientational order parameter profile indicates that the carbonyl carbon/carbonyl oxygen bond vector aligns antiparallel to the interface normal within the interfacial region. The oxygen tends to orient toward the bulk liquid phase while the less electronegative methyl groups are exposed to the vapor; this is understandable if one considers the energetic interactions retained for the oxygen by taking such a spatial orientation. An interesting of this work will be to observe the response of the interfacial structure to a change in the contacting medium. We plan to present simulations of the methyl acetate/water liquid-liquid interface in the future. We expect to find that the orientations of the carbonyl bond vector will change due to the enhanced opportunities for hydrogen bonding due to the presence of water. This phenomenon has interesting implications with respect to biomaterials, and in particular, those which are used in applications requiring contacting with physiological fluids containing a vast variety of proteins, small-molecule solutes, and the like. Finally, our work shows that the OPLS non-bond parameters cannot predict liquid-vapor surface tension to any better than within 10 percent of the experimental measurements, and we believe that this will hold true for the whole class of simple esters.

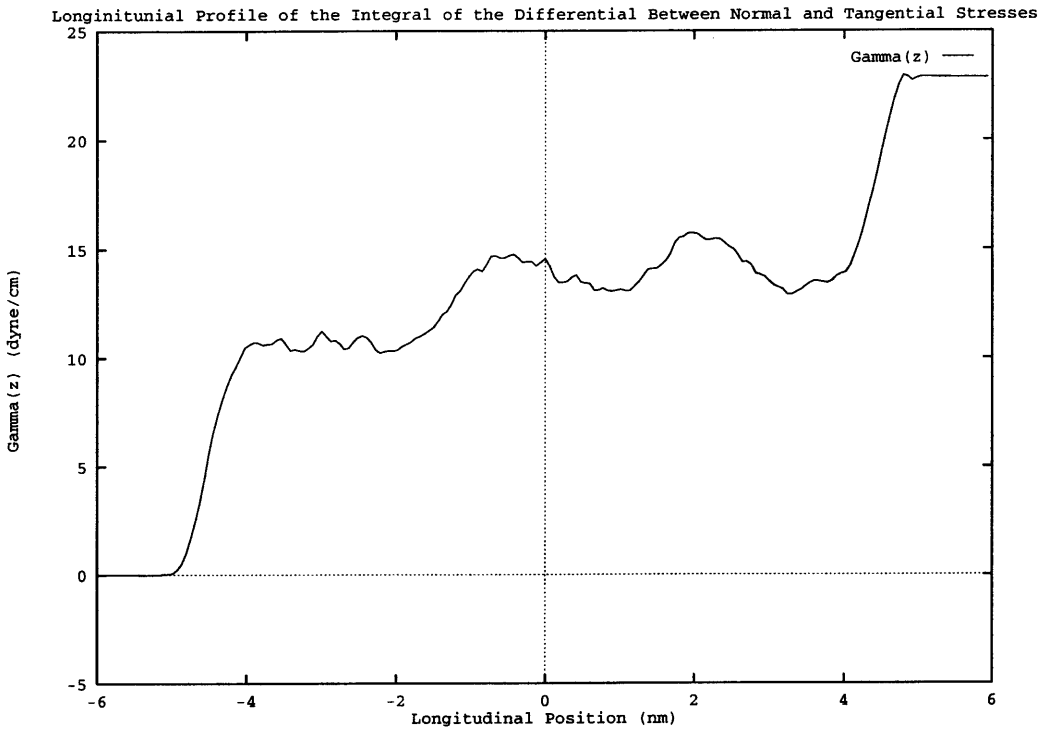


Figure 4-19: The difference between the normal and tangential components of the pressure tensor.

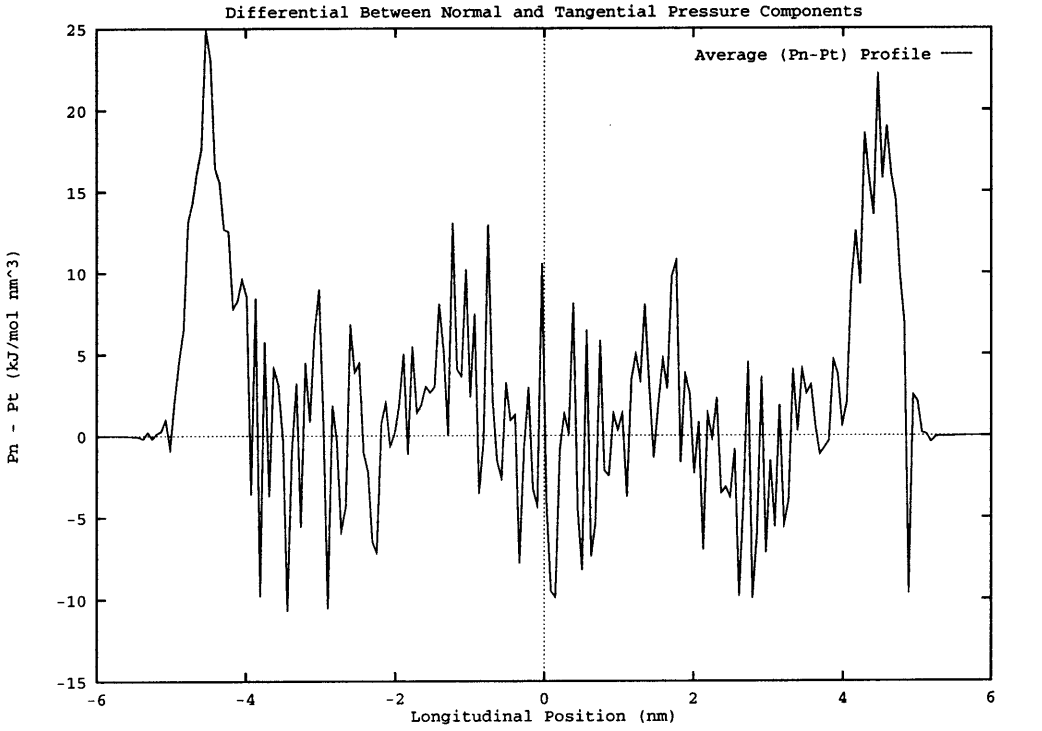


Figure 4-20: The integral of the difference between the normal and tangential components of the pressure tensor.

Chapter 5

Methyl Acetate - Water: Assesment of OPLS-SPC Model

5.1 Introduction

Ultimately this study is interested in probing the oligomer-water interface; thus, with the small-molecule model at hand, it is worth delving into the equilibrium properties of the liquid-liquid interface for small molecules, in this case methyl acetate and water. In a general sense, the study of properties of liquid-liquid interfaces is of great interest in numerous ares of science, engineering, and medicine. Apart from the physiological effects of the interface as discussed in the introduction (from the perspective of long-chain molecules), interfaces of small molecules are important in cases where the transport mechanisms of materials across interfaces directly depends on structural, thermodynamic, and/or dynamic properties of the interface [49].

Despite their fundamental importance, liquid-liquid interfaces have proven difficult to study on a molecular level via both theoretical and experimental approaches [49]. Traditionally, experiments on liquid-liquid interfaces dealt with interfacial thermodynamical aspects such as surface tension or surface potential [121]. These types of experiments can only provide limited information on the orientations of molecules at an interface [121]. Furthermore, typical surface analysis techniques used to probe solid surfaces in ultrahigh vacuum are intractable for liquid systems [120]. From a theoretical standpoint, liquid-liquid interfaces have been treated with continuum approaches, thus offering very little insight to the molecular level detail [49]. Recent advances in experimental and computational technologies such as second harmonic generation (SHG) or sum frequency generation (SFG) [120, 121, 133] and multiprocessor, parallel architectures and codes, respectively, have allowed access to more detailed information about the liquid-liquid interface [49]. Note, however, that even SFG/SHG techniques are not straightforwardly applicable to the study of *liquid-liquid* interfaces primarily due to, for example, IR signal loss in one liquid-phase before the interface is reached [120, 121]. Computationally, simulations on the molecular-level have greatly elucidated some of the fundamental aspects of liquid-liquid, liquid-vapor, and liquid-solid interfacial thermodynamics, structure and dynamics. Early work on computer simulation of the liquid-liquid interface is that of Linse who reported on

the structure and thermodynamics of the benzene-water interface via Monte Carlo methods [50]. Linse reports the formation of a stable interface as well as of a preferred orientation of water molecules near the interface; furthermore, he indicates the reinforced hydrogen bonding at the interface. Gao and Jorgensen [134] report Monte Carlo simulations of the 1-hexanol, water interface in monolayer, bilayer, and double bilayer configurations; they conclude that the water penetrates very little into the amphiphilic aggregates. Molecular dynamics studies on the hexane-water interface [45], water-1,2-dichloroethane interface [135], and decane-water interface [58] hav e recently shed light on the structure and thermodynamics of pure liquid-liquid interfaces. Going beyond the study of the interface itself, phenomena occurring at/across the interface. For example, water transport across a lipid membrane was studied by Berendsen and coworkers [136]; Hayoun *et al* [137] studied transfer of a solute across a liquid-liquid interface; and Benjamin *et al* [138] studied the mechanism of ion transport across the water/dichloroethane liquid-liquid interface.

This Chapter presents work probing the methyl acetate - water liquid-liquid interface, with particular interest in the structure of the interface in terms of ester and water orientations. Furthermore, we are interested in comparing the structure of the liquid-liquid interface to the liquid-vapor interface for both species–that is, how does the molecular orientation of a water molecule change on going from its own vapor as contact medium to a second liquid ester phase as contact medium? Furthermore, we are concerned with hydrogen bonding effects which will tend to reinforce any preferential orientation at the interface.

One aspect of our understanding of the molecular interactions between the two species will be the quality of the OPLS-SPC interaction model. That is, does the interaction model we use, based on standard combination rules for forcefields (which will be discussed below), accurately describe the two-phase behavior of the ester-water system? *A priori* it is difficult to judge the accuracy of the potential, particularly as the two models are generated based on optimizations to individual species properties, and more specifically, *bulk* properties. If anything, one expects the potential model to fail; however, this pre-judgement is applicable to any available (commercial

or academic) potential model; the question is to what extent does any model fail. Thus, we proceed with the investigation of the ester-water system with this in mind. Furthermore, as it is not the ultimate goal of this work to refine the potentials for the low molecular weight systems, we will note the inadequacies and leave it to future researchers to develop accurate quantitative models for this system. The next section presents the details of the simulations, and the following sections describe results.

5.2 Simulation Details and Methodology

This section discusses the details of the simulations performed for the methyl acetate-water liquid-liquid interface. More specifically, the system and the various techniques implemented are given. The system geometry is shown in Figure 5-1. Here we have placed a bulk methyl acetate layer in between two slabs of SPC water. The methyl acetate layer is an equilibrated system from a simulation of the bulk ester. The water layers are from a simulation of bulk SPC water well after equilibration has been reached. The SPC configurations are for T=298K and $\rho = 0.97 gr/cm^3$. For SPC water, this corresponds to a system at pressure of 1 bar. This discrepancy in the equilibrium SPC water densities is well documented in the literature [113]. We feel however, that this is not a significant discrepancy when looking at the structural nature of the interface. We anticipate that any thermodynamic properties of the interface will suffer due to this flaw in the potential. The same argument holds for the deficiencies in the ester potential.

The overall system contains 648 methyl acetate molecules and 1999 SPC water molecules. As Figure 5-1 shows, the system is rectangular in geometry. The ester layer is 'sandwiched' by the two water layers. Beyond the water layers, a region of space is allowed so that the water phase can evolve a 'vapor' region associated with it. This in effect allows a water/water vapor interface. This is quite useful as it allows a simultaneous comparison of the nature of the water at the water-water vapor interface and at the water-ester interface. One would anticipate a disparity in the structure and thermodynamics of the two interfaces a priori. We note here that the

dimensions of the simulation cell are $L_x = L_y = 3.084829966nm$ and $L_z = 23.0nm$. The lengths of the simulation cell are maintained constant for the duration of the runs. The longitudinal length, L_z was chosen to be larger than the lateral lengths so as to create the water vapor region as well as to prevent interaction between the two water-water vapor interfaces. Furthermore, as the width of the interface was not known prior to the simulation, it was necessary to select a dimension which would allow for a bulk ester region. Finally, we introduce periodic boundary conditions in all three Cartesian coordinates. This approach is standard fare for NVT simulations of liquid-liquid, liquid-vapor interfaces. Note that the literature tends to favor the choice of periodic boundary conditions for all cartesian dimensions for studying liquid interfaces, particularly if the surface tension is desired [139]. Unfortunately, when electrostatics are concerned, the choice of which boundary conditions are most effective is not well-understood.

Finally, the model parameters for OPLS methyl acetate are not given here; the reader is referred to Chapter 4 which gives the detailed information. The water model adopted for this work, as just mentioned, is the Simple Point Charge (SPC) model of Berendsen *et al* which consists of a Lennard-Jones center located on the oxygen atom, and three point charges, one located on each of the three atoms. The oxygen is of charge -0.82e, and the hydrogens each carry +0.41e. The oxygen-hydrogen bond length is constrained to 0.1 nanometers, and the lone bond angle is constrained to 109.47 degrees. The Lennard-Jones well-depth for the oxygen center is $\epsilon_o = 0.648kJ/mole$, and the core diameter is $\sigma = 0.3166nm$. The SPC water hydrogen atoms interact with other system species via Coulombic interactions only. We note that the present water potential does not explicitly incorporate hydrogen bonding interactions—any indications of a hydrogen bonded structure resulting from simulations is attributed to the physical dispersion and electrostatic interactions included. We feel that this is a more rigorous approach since this allows the system to evolve a 'hydrogen-bonded' structure without any prior bias (as introduced by a 'hydrogen-bonding' potential). Finally, the SPC model neglects dissociation effects; thus, issues associated with the water ions present (such as acid catalysis of

de-esterification) cannot be addressed directly. This is an accepted limitation of the current models used.

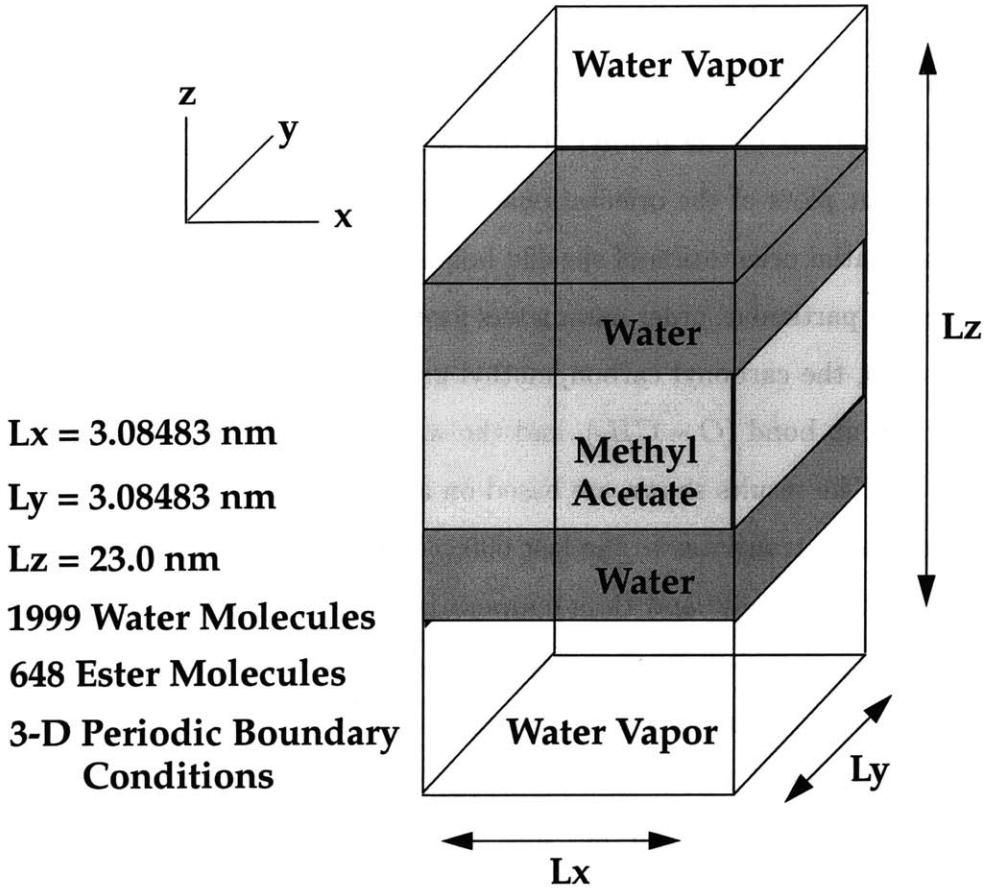


Figure 5-1: Methyl Acetate / Water Simulation Cell

5.3 Results: Methyl Acetate-Water Liquid-Liquid Interface

In this section, we present structural and thermodynamic results from the molecular dynamics simulations of the methyl acetate-water interface. In terms of structural data, we present plots of the orientational order parameters,$P_1(z)$ and $P_2(z)$, which describe the spatial orientation of specific bonds (bond vectors) relative to the interface normal. In particular, order parameters for the carbonyl carbon/carbonyl oxygen bond ($C = O$), the carbonyl carbon/methyl group bond ($C - CH_3$), the ester oxygen/methyl group bond ($O - CH_3$), and the water molecular dipole moment vector are presented. The results shown are based on averages over configurations following equilibration. This translates to the last 600 picoseconds of the simulation as shown in the density-time profile and time-temperature plots. The behavior of the $P_1(z)$ and $P_2(z)$ order parameters for the methyl acetate in the case of the liquid-vapor and the liquid-liquid interfaces show dramatically different behavior, thus providing some insight into the effect of the external environment on the nature of the interface. Also presented are radial distribution functions for the bulk and interfacial regions of the ester. This should act as a check for the bulk since the radial distribution function for this region should be similar to that of the bulk liquid (in this case there will be a slight difference between the bulk rdf's and those computed from NVT simulations of the bulk at the experimental density due to the lower density of the bulk region derived from the interfacial simulations). Furthermore, the distributions of various internal degrees of freedom including the central torsion angles and the out-of-plane angle are presented for the bulk and interfacial regions. These serve to indicate the nature of the molecular geometry which when coupled with the order parameters for the individual bond components, allows a detailed picture of the nature of the interfacial structure. Specifically, the atomistic description of the system allows a view of the molecular orientation and the effect of the anisotropic environment on it.

In terms of thermodynamic results, we present density profiles for total molecular number density and individual molecular component (e.g. atom or group in

an united-atom sense) densities. We also present longitudinal profiles of the excess chemical potential of water in methyl acetate as a check on equilibration. Furthermore, the differential between the normal and tangential components of the molecular pressure tensor is computed as a function of longitudinal position. This is done as a check on equilibration as the normal and tangential components should be equivalent away from the interfaces and the normal component should be constant throughout the system. Furthermore, analyzing the components of the pressure tensor acts a check on the 'appropriateness' of the potential parameters. If the potential model is correct at the state point under study, the average normal component of the pressure tensor (normal to the interface) should essentially be atmospheric pressure (and constant throughout the longitudinal profile), and the tangential (in-plane) components should give rise to the surface tension which can be compared to experimental data. The off diagonal elements of the pressure tensor should be zero (or 'small').

5.3.1 Density Profiles

Figure 5-2 shows the symmetrized total molecular number densities ($molecules/nm^3$) with respect to the system center of mass for methyl acetate and water. These profiles are computed by averaging over blocks of 100 configurations, with a configuration sampling frequency of every 0.4 picoseconds (every 100 configurations). The abscissa represents the longitudinal position along the interface normal, with the zero being in the bulk ester. The total molecular density profiles are typical of small-molecule interfaces between *immiscible* fluids such as hexane-water or benzene-water reported in the literature. This observation is the first indication of the inadequacy of the current combination of the OPLS-SPC interaction models to describe the methyl acetate - water system. Experimentally, the two liquids are soluble, so one should expect a binary system at 298 K. For the methyl acetate - water system, there are published solubility data for a range of temperatures at atmospheric pressure. Tables reftbl:esterinwater and 5.2 show some characteristic values of the solubilities of

the ester in water (and water in ester) at several temperatures; also shown are the mole fraction compositions [?].

Table 5.1: Solubilites of Methyl Acetate (1) in Water (2)

Temperature (K)	gr(1)/100 gr soln.	100 X Mole Fraction (1)
278	23.7 ± 1.1	7.02
283	23.5 ± 1.1	6.95
293	23.9 ± 0.6	7.10
298	23.1 ± 0.2	6.81
303	22.9 ± 0.4	6.74
313	22.7 ± 0.8	6.67
323	22.9 ± 0.8	6.74

Table 5.2: Solubilities of Water (2) in Methyl Acetate (1)

Temperature (K)	gr(2)/100 gr soln.	Mole Fraction (2)
273	6.4	0.219
283	6.8	0.231
293	7.6 ± 0.3	0.253
298	8.1 ± 0.1	0.266
303	8.5 ± 0.2	0.276
313	9.7 ± 0.8	0.306
323	11.1± 1.4	0.339

At 298 K, both the water-rich and ester-rich phases show appreciable solubilities of the other species; thus, if the current combination of interaction parameters is to any extent correct, one should observe a 'mixing' over time (as we have started from a configuration with three equilibrated slabs of the pure liquids). However, we do not observe any mixing with time apart from the intermingling of the two species at the pseudo-interface formed due to the inadequacy of the potential to describe the solubility (this will be addressed in more detail further below). This is demonstrated by computing the Gibbs dividing surface for the individual liquids as a function of time. The same relations are used as in Chapter 4 for the analysis of the methyl acetate free surface stability. Figure 5-3 and Figure 5-4 show the time profiles over the last 800 picoseconds of the simulation of the ester and water

Gibbs dividing surface locations. The plots indicate the expected thermal fluctuations (capillary fluctuations) leading to surface roughness. However, there is no systematic drift of either dividing plane over a rather lengthy simulation time which is much longer than the time required for diffusion of either species over a length of several molecular diameters. In fact, the density profile is much like that of an immiscible system [50, 59]. We can further compute the bulk concentrations of water in ester and ester in water as a function of time. Figure 5-5 shows the ester mole fraction in the bulk SPC water phase (the middle 10 angstroms), and Figure 5-6 shows the water mole fraction in the bulk ester phase (the middle 40 angstroms). For the ester in water case, there is essentially no significant solubility of the ester in the water. There appears to be an interval of time over which there is some ester present, but this quickly diminishes and we observe no appreciable solubility. For the water in ester case, there appears to be some presence of water molecules in the bulk ester, but even at the maximum composition, the water is two orders of magnitude less concentrated than that predicted by experiment. In the experimental system, the solubilities correspond to having 3 molecules of ester for every one molecule of water in the ester-rich phase, and 14 molecules of water for every one molecule of ester in the water-rich phase.

Finally, we note that density profiles of both the ester and water taken over intervals show no consistent mixing of the two phases. So we are confident that the system has equilibrated with respect to compositions. The final verdict is that the OPLS-SPC interaction model we have used is deficient in terms of modeling the two-phase mixture system. However, at this point, we accept this fallacy.

The interface is sharp with a 10-90 width of 0.745 nm (the average interface is thus located within roughly a 1 nm region). Since the width of the interface between two *immiscible* liquids is difficult to unambiguously specify, the 10-90 width is taken to be the distance over which the SPC water density falls from 90 percent to 10 percent of its bulk value. This is also done as it is the water density profile in the interfacial region that encompasses the larger width. The bulk ester density is 0.871 gr/cm^3 which is lower than the expected value for the OPLS parameter set at 298K

and lower than the experimental density by about 6 percent. The 10-90 width for the water liquid-vapor interface is 0.465 nm. This smaller width is generally regarded as a consequence of the hydrogen bonding which tends to compact the interface in order to minimize energetic losses [56]. Figure 5-7 shows the symmetrized ester atomic component densities plotted as a function of longitudinal position, with the center of the bulk at zero longitudinal position. The curves represent the carbonyl oxygen, the carbonyl methyl group, and the ether (or methoxy) methyl group density profiles. This data shows a structuring effect at the interface as there develops a spatial separation of the various components. The carbonyl oxygen is suppressed between 4 and 4.5 nanometers out from the bulk, but is dominant in the outer layer of the interface. Furthermore, the methoxy methyl group appears to be in closer proximity to the carbonyl oxygen than the carbonyl methyl group at it density hovers in between the other two components. This is consistent with the fact that due to the torsional potential favoring the syn periplanar conformation (the Z conformer of methyl acetate is the dominant species in both the liquid and vapor [67]) the methoxy methyl is closer to the carbonyl oxygen. Further evidence in favor of the molecular planarity is given below via out-of-plane angle and torsion angle distributions. It is interesting that such a significant structuring is observed in the case of such a small molecule system, and indicates that the potential model used does allow for a physical mechanism to manifest itself at the interface. In this case, the major contributor to the physics and structure at the interface is the electrostatic component of the interaction.

Finally, Figure 5-8 and Figure 5-9 show the time profiles of the bulk ester and water densities. These show that two stable bulk regions are established and maintained over the time period of the simulation. The average bulk methyl acetate density is 0.871 gr/cm^3, roughly 6.02 percent lower than the experimental density at 298K (0.927 gr/cm^3). This is within the error quite often reported in the literature for bulk densities computed from interfacial simulations. This also is in keeping with the slightly negative system pressure computed for the current simulations. The average bulk water density is 1.078 gr/cm^3 for the SPC water model with deuterium atoms

replacing the hydrogens (as done in this work); for reference, this density corresponds to 0.97 gr/cm^3 for SPC water with hydrogens which is in keeping with the PVT behavior of SPC water at 298K (the discrepancy between experimental and model densities is well documented in the literature).

Figure 5-11, Figure 5-12, Figure 5-13, and Figure 5-14 show the atomic and molecular temperature time profiles for the ester and water in order to ascertain that one component is not 'hotter' than the other as may happen if the system has not equilibrated properly.

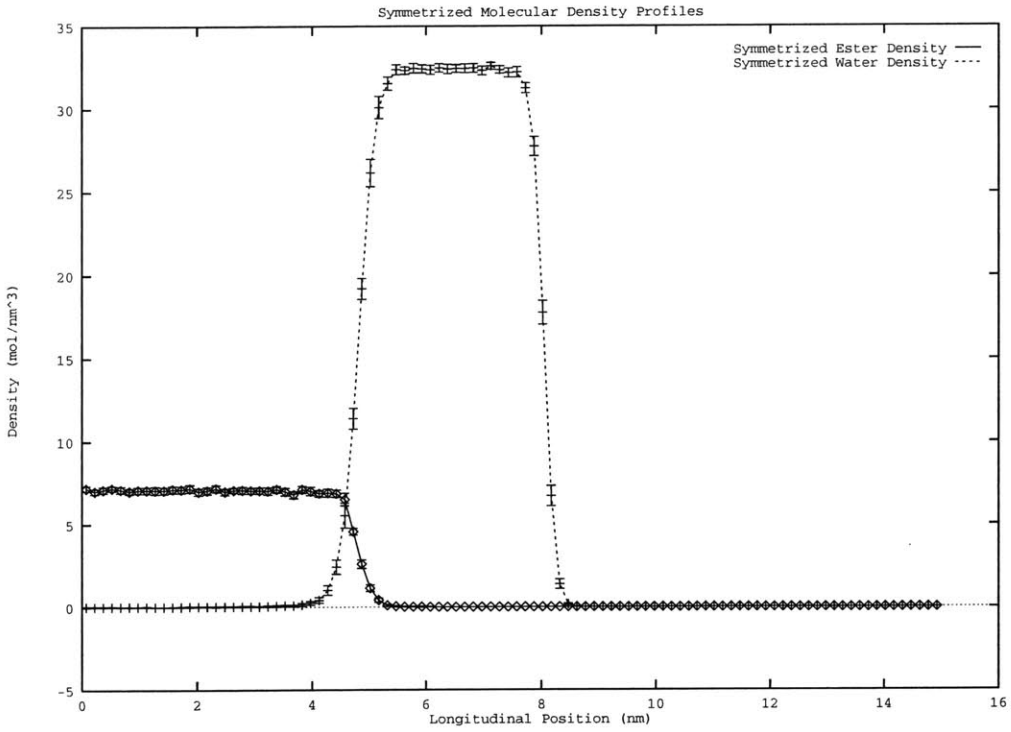


Figure 5-2: Symmetrized molecular density profiles of water and methyl acetate. Longitudinal position is relative to the center of mass of the simulation cell.

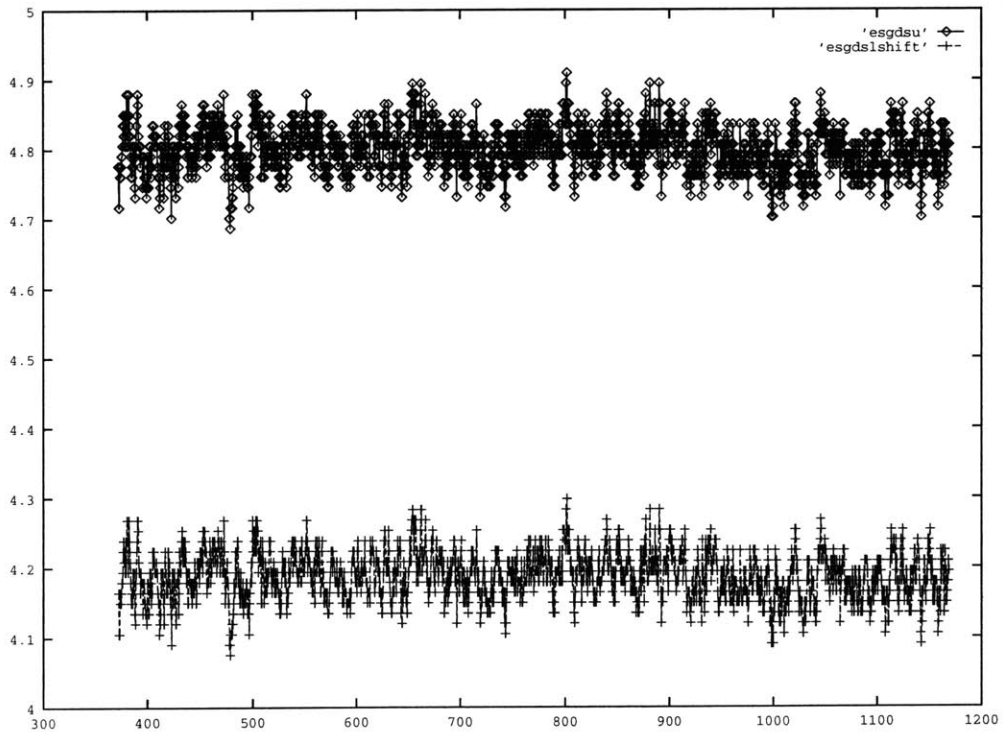


Figure 5-3: Time profile of the location of the Gibbs dividing surface: Methyl Acetate

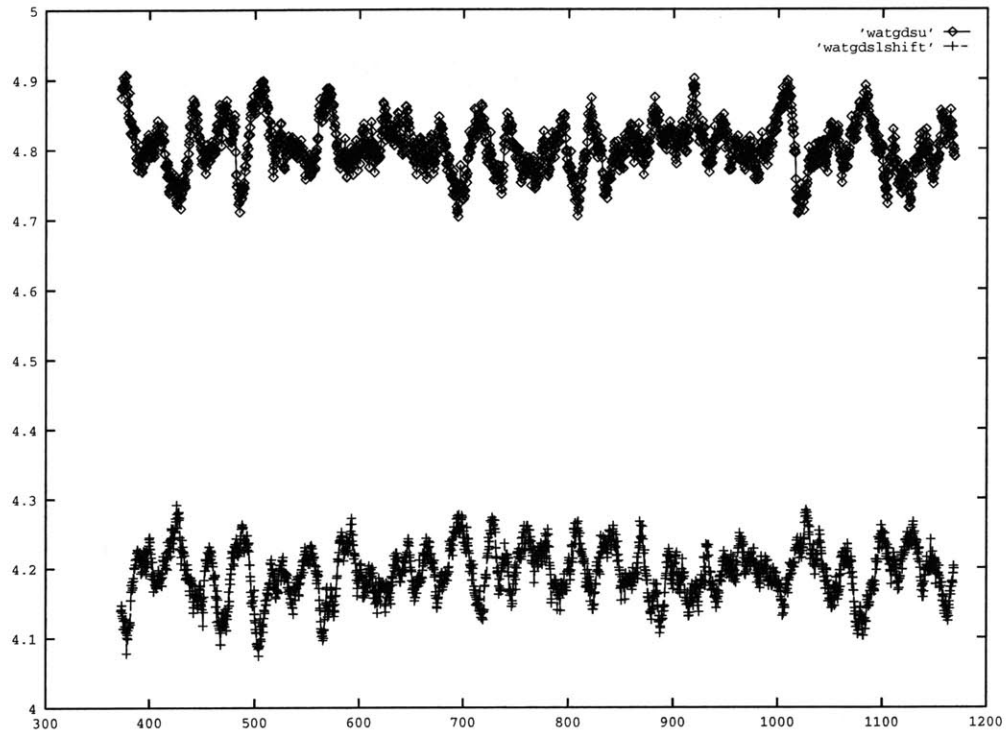


Figure 5-4: Time profile of the location of the Gibbs dividing surface: Water

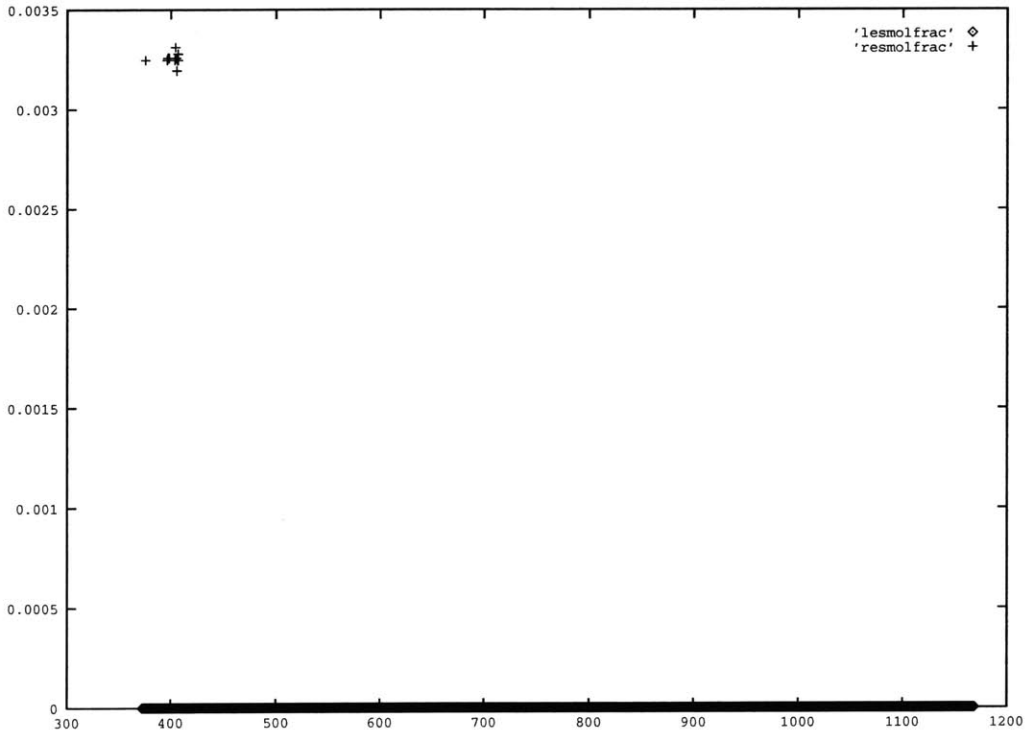


Figure 5-5: Time profile of the mole fraction of ester in bulk water

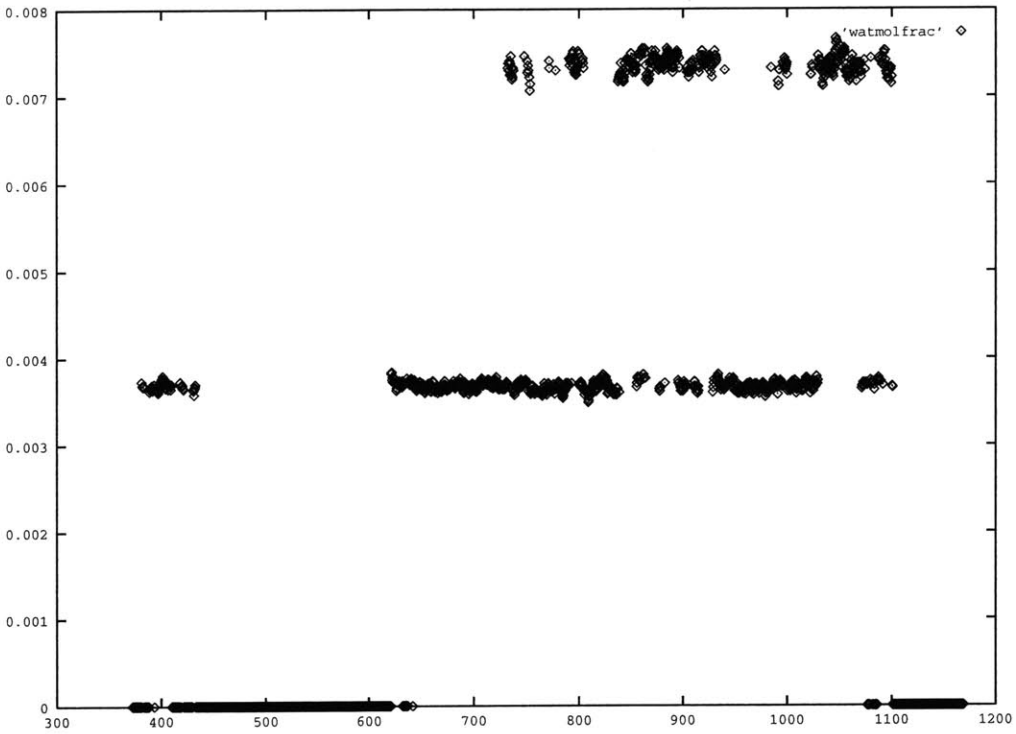


Figure 5-6: Time profile of the mole fraction of water in bulk ester

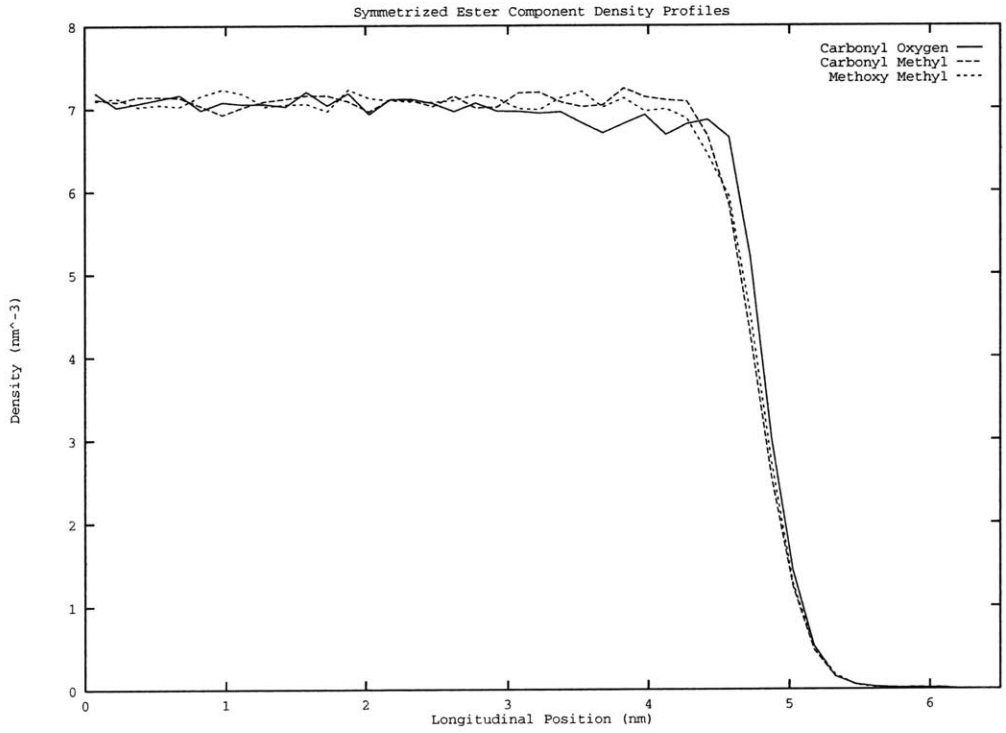


Figure 5-7: Symmetrized ester atom density profiles.

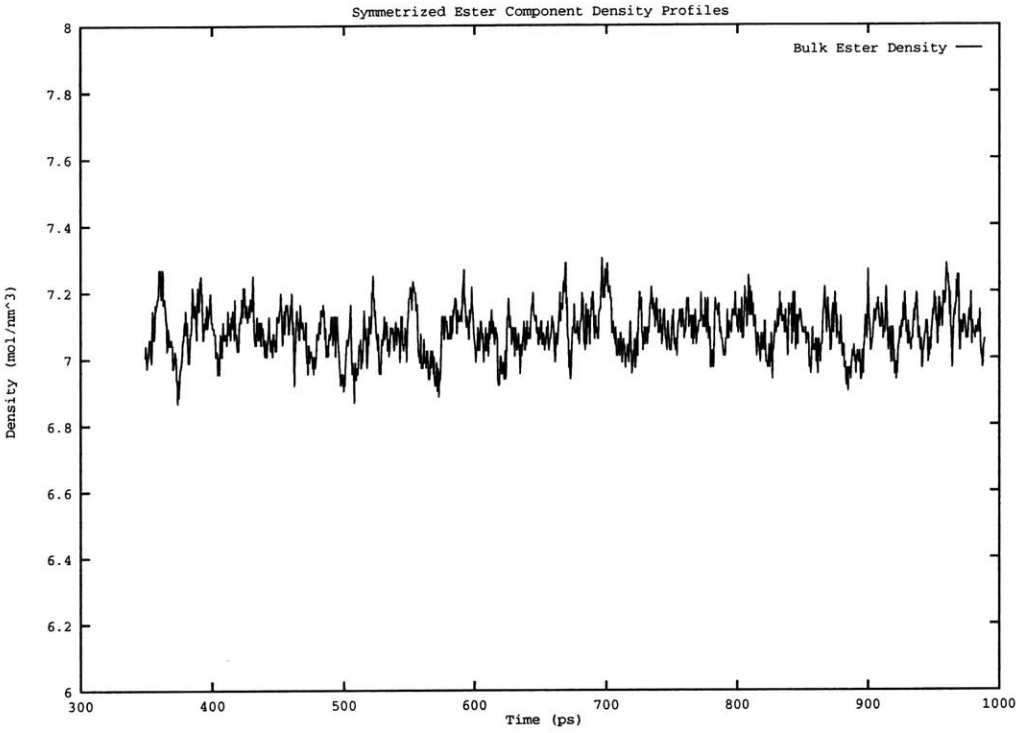


Figure 5-8: Methyl acetate bulk density as a function of time.

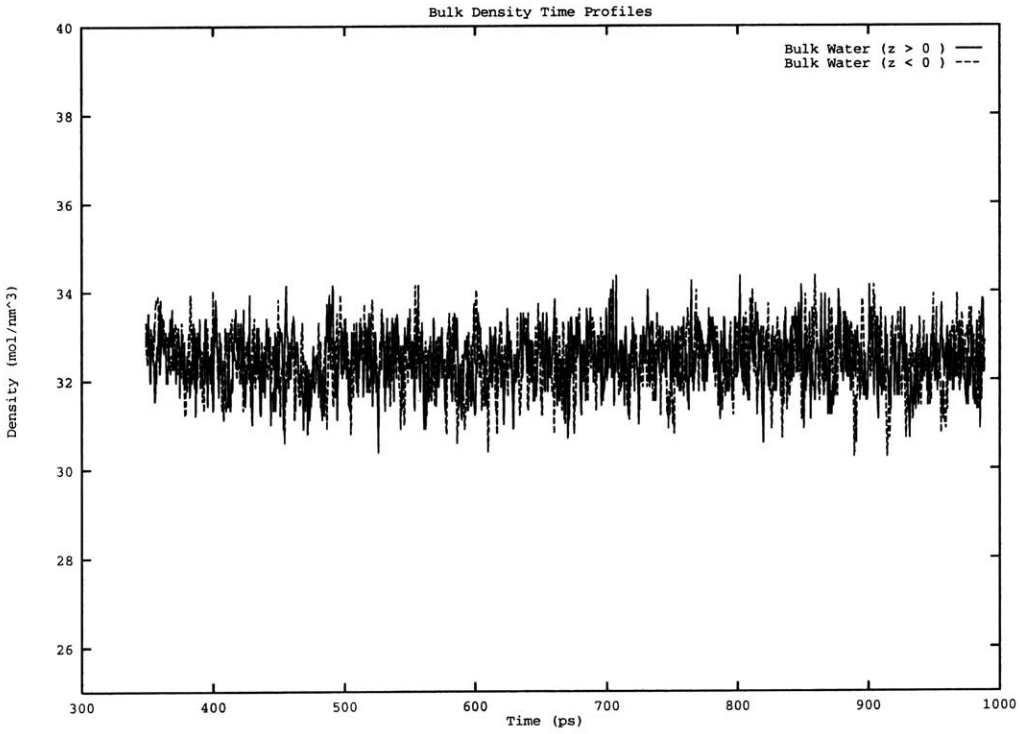


Figure 5-9: Water bulk density as a function of time.

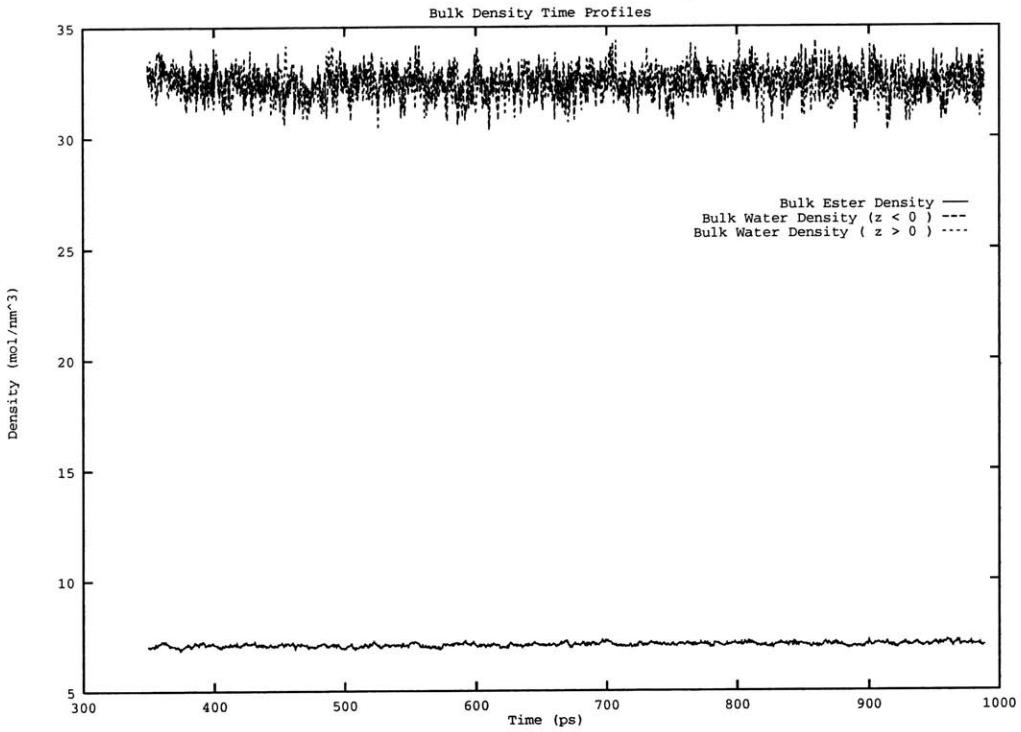


Figure 5-10:

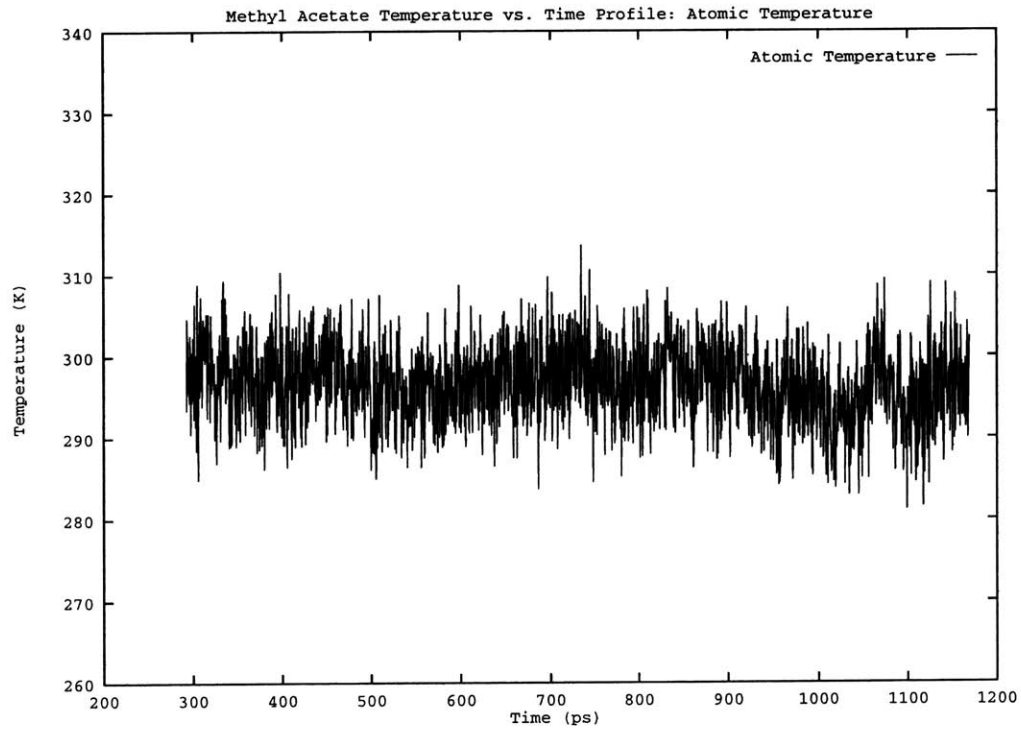


Figure 5-11: Methyl acetate atomic temperature time profile.

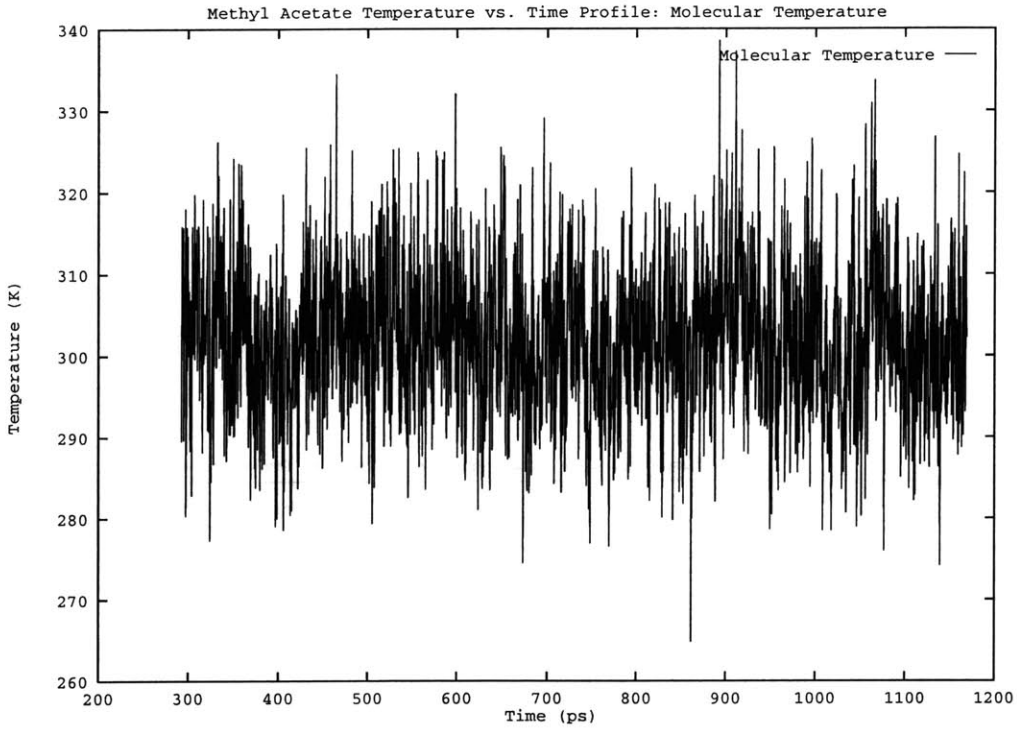


Figure 5-12: Methyl acetate molecular temperature time profile.

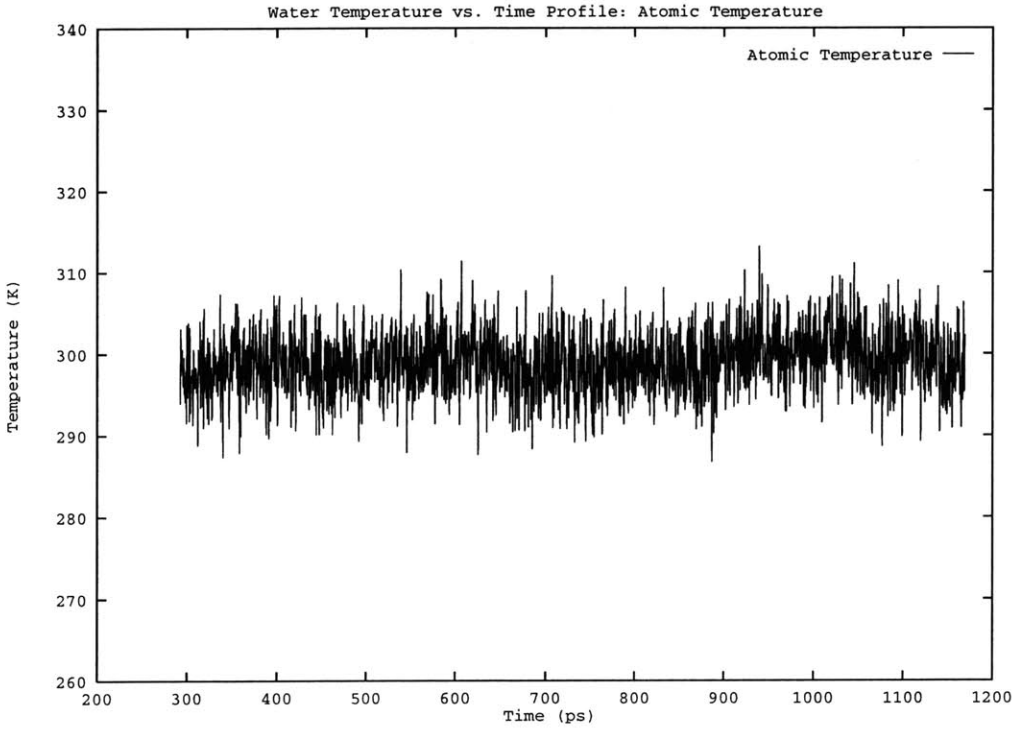


Figure 5-13: Water atomic temperature time profile.

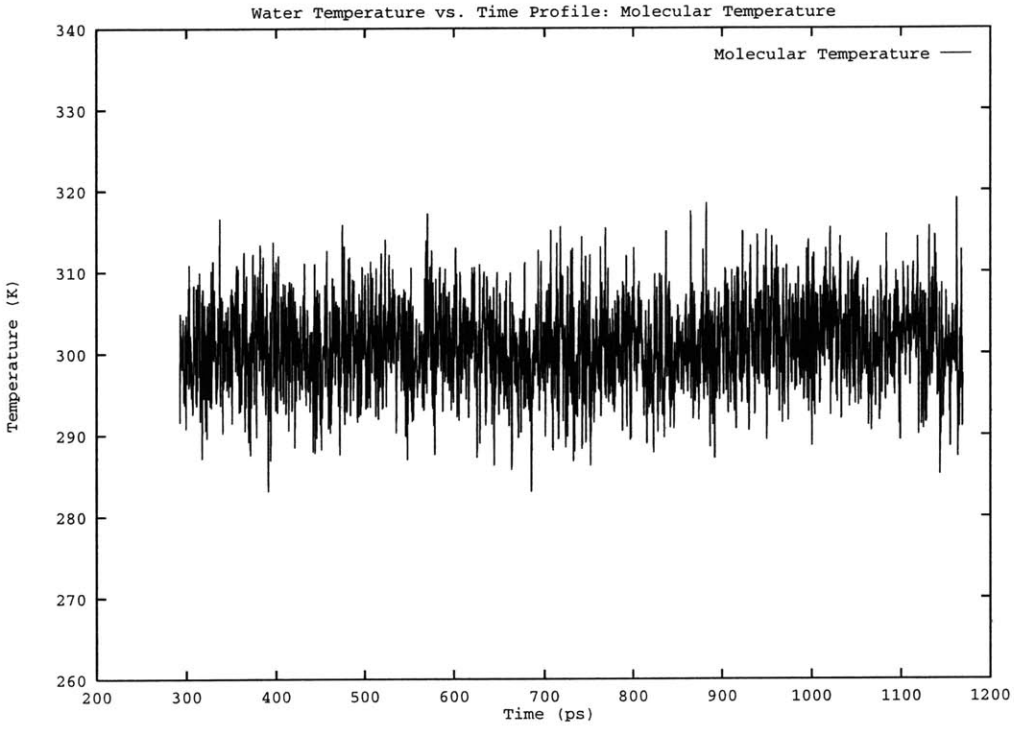


Figure 5-14: Water molecular temperature time profile.

5.3.2 Effect of Interface on Molecular Orientation: Orientational Order Parameters

Despite the fact that the Lorentz-Berthelot mixing rules we have applied to derive the ester-water interaction from the pure component parameters fail to capture the equilibrium solubility properties of the ester-water binary, out of curiosity, we proceed to consider some *interfacial* aspects of the system—that is, we will concern ourselves with the structural aspects of the ester and water in the vicinity of the region where the two species come together. This is motivated by the intuition that possibly the change in contact medium may elicit a response of the ester orientation; one may say that, in an exaggerated sense, we are substituting the ester vapor phase with a dense vapor with high moisture content.

The effect of the interface on the orientations of various bond vector is studied through orientational order parameters (first and second Legendre coefficients) defined as [60, 58]:

$$P_1(z) = < \cos\theta > (z) \tag{5.1}$$

and

$$P_2(z) = \frac{1}{2} < 3(\cos^2\theta) - 1 > (z) \tag{5.2}$$

where θ is defined as the angle between the interface normal (along the positive and negative z-axes) and a bond vector associated with a specific functional group. For the present case, the carbonyl carbon - carbonyl oxygen bond vector, r_{CO}, the carbonyl carbon - methyl bond vector, r_{C-CH3}, and the ether oxygen - methyl group bond vector, r_{O-CH3}, are considered for methyl acetate. For the SPC water, the vector directed from the bisection point of the line joining the two hydrogen atoms to the oxygen atom (a vector antiparallel to the molecular dipole moment vector) is the relevant vector. The average is a over all vectors and time steps within a thin slab perpendicular to the interface normal. If the vectors in a region are completely

randomly oriented, $P_2 = 0$. If there is a prevalent orientation parallel to the interface normal, $P_2 = 1.0$; with an orientation perpendicular to the normal, the order parameter goes as -0.5. Note that the bond vector is defined as $r_{ij} = r_i - r_j$. Thus, for methyl acetate, the relevant bond vectors are such that the r_{CO} vector points from the carbonyl oxygen to the carbonyl carbon, the r_{O-CH3} vector is directed from the methoxy methyl group to the ether oxygen, and the r_{C-CH3} vector points from the carbonyl methyl group to the carbonyl carbon.

Figure 5-19 shows theP_1 order parameter as a function of longitudinal position along the interface normal starting from the bulk ester liquid and moving outwards to the ester-water interface. The shown profile is symmetrized with respect to the center of mass of the total system. The P_1parameter represents the average value of the cosine of the angle between the interface normal and the bond vector, and as Figures 5-19 and 5-20 show, even with this small-molecule system, the interface displays a significant structuring, the nature of which is dictated by the surrounding environment. In the bulk ester liquid region, all the bond vectors more or less show no preferential orientation and the local environment is isotropic. As the interfacial region is approached, the loss of isotropy induces a marked divergence in the order parameters. The carbonyl carbon - carbonyl oxygen bond vector preferentially directs towards the bulk ester, with the oxygen extending towards the water layer. Commensurately, the methoxy methyl - ether oxygen vector orients in much the same way as the former. However, due to the molecular geometry, the average angle is smaller, hence giving a slightly more positive value for P_1. Finally, the carbonyl methyl group - carbonyl carbon vector shows the opposite behavior, directing towards the water layer, with an average angle close to that made by the ether methyl group. Again this is in keeping with the molecular geometry.

Much has been said of the molecular geometry, and at this point, some comments are made as to what that geometry is. This knowledge coupled with the P_1 and P_2 profiles gives a very nice picture of the average orientation of the ester interface. Firstly, the individual molecular geometry is essentially planar as evidenced by the out-of-plane and central torsion angle profiles shown in Figure 5-17 and Figure 5-

18. The out-of-plane angle distribution is sharply peaked at zero degrees for both the molecules considered in the bulk region and those in the total interfacial region. Thus, the carbonyl carbon is at the center of a planar arrangement (this is in keeping with the sp2 nature of this atomic environment). The planarity of the molecule as a whole is realized once the central torsion angle distributions for bulk and interface are considered. The dihedral under consideration is that giving the angle between the planes containing the carbonyl carbon - carbonyl oxygen bond vector and the ether oxygen - methoxy methyl group bond vector. Figure 5-17 shows a sharply peaked distribution around the syn conformation ($\phi = 0.0$). Thus, on average, the molecule is in a planar state.

Next, the P_2 order parameter is considered as it gives information on the width of the distribution of the angle cosine about the average value given by P_1. P_2 gives an indication of the degree to which a certain vector is aligned with the interface normal. For instance, the difference between an angle of 10 degrees and an angle of 50 degrees between the normal is picked up in the P_2 profile. Again, a P_2 value of zero signifies completely randomly oriented vectors. Values towards -0.5 indicate a more planar orientation, while positive values indicate vector orientation more parallel to the normal. Figure 5-20 shows the symmetrized P_2 profiles for the aforementioned ester bond vectors. The carbonyl carbon - carbonyl oxygen profile indicates a significant orientation along the interface normal quite distinct from the randomness of the bulk. Note that the minor fluctuations in the bulk (deviations from random orientation) are an artifact of the finite sample size; fortunately, these deviations are much smaller than those in the interface (this phenomena is also reported in the literature [60]). In the outer interface, there is a slight tendency for a parallel orientation. This results from the increased water density in the outer region, which lessens the need for the carbonyl oxygen to be normal to the interface. Concurrently, the remaining bond vectors follow suit and also tend towards a more planar orientation in the outer interface.

As alluded to above, the orientational changes on going from bulk to interface are induced by the environment, in this case, the water layer adjacent to the ester.

Reciprocally, the ester has an influence on the structure of the water in the interfacial region which is detected through the P_1 and P_2 profiles for the water dipole moment vector. Figure 5-21 shows the symmetrized P_1 profile for the water plotted as a function of the longitudinal position. In the vicinity of the ester-water interface between 4 and 5 nanometers from the center of mass, P_1 indicates that the preferred orientation of the water molecule is such that the hydrogens are directed towards the ester layer, and more specifically, attracted to the more electronegative field generated by the higher number density of the carbonyl oxygen relative to the water oxygen. This disproportion in densities in the interfacial region is shown by the density plots. Furthermore, the orientational order becomes less pronounced moving towards the outer interfacial layer between 4.5 and 5.0 nanometers, finally become random in the bulk of the water film. There is a decreasing trend in the P_1 profile and this suggests a long-range correlation in the orientation; this is most probably due to the finite thickness of the water layer, which probably should have been made larger in order to allow for damping of this correlation. An alternative explanation which is motivated by experimental observations is that water near interfaces is known to possess differing structural and thermodynamic properties than the pure bulk liquid [26, 25]. A vicinal structure is observed to extend 3 to 6 nanometers from the interface [26]; this is the same spatial range over which we observe the decreasing trend in the water P_1 profile. This seems to suggest that the SPC model does capture the structure of vicinal water and that the spatial orientational correlation we observe via the water P_1 profile may not be an artifact of insufficient equilibration. A definitive answer to this question would most probably result from a simulation with a thicker water layer. A very interesting phenomenon is the reversal of the preferred orientation at the water liquid-vapor interface beyond 8.0 nanometers out. The P_1 behavior indicates that now the dipole vector is slightly protruding towards the vapor region. The P_2 profile for water shown in Figure 5-22 shows that in the inner ester-water interface (closer to the ester layer), the orientation of the water dipole moment vector is perpendicular to the interface and becomes more planar traversing through to the outer interface. The isotropic bulk gives the random orientation expected. At the liquid-vapor interface,

there is a marked difference in P_2, as it shows a planar orientation. Thus, at the liquid-vapor interface, the water is slightly displaced from planarity, with at least one hydrogen jutting into the vapor phase, allowing the dipole moment vector to have a non-zero average value (as demonstrated by the P_1 profile).

We can further analyze the orientation of water molecules at the liquid-vapor and liquid-liquid interfaces by computing distributions of the angles made by 1). the water dipole moment vector with the interface normal and 2). by the O-H bond vector and the interface normal. More specifically, one can look at the distributions of the cosines of these angles. First, consider the water liquid-vapor interface. Figure 5-23 and 5-24 show the distribution of the cosine of the angle between the interface normal and the water dipole moment vector for three regions. The vapor side corresponds to the outermost region of the interface, here taken to be between 8.1 and 8.5 nm (in the symmetrized sense). The region represented by the curve labelled with delta = 0.2 nm spans the width from 7.9 nm to 8.1 nm; the remaining region spans 7.8 nm to 7.9 nm. Distributions were computed simply by normalizing histograms of the number of angles falling within evenly spaced intervals between -1 and 1 values of the cosine of the angle. From Figure 5-23, we see that in the outermost, 'vapor' region of the interface (8.1-8.5 nm), the dipole vector juts out towards the vapor phase with the peak in the cosine distribution being at about 0.4; this corresponds to an angle of 66 degrees from the interface normal. Moving in to the liquid phase, Figure 5-24 shows that the distribution shifts to lower values of the cosine, indicating that the molecules in this region are preferentially lying in the plane of the interface. This behavior is analogous to that described by Taylor *et al* and Matsumoto and coworkers [54, 55] as will be discussed below.

Next, one can consider the distribution of the angle, ϕ, between the O-H bond vector and the interface normal to get a more detailed picture of the molecular orientation. Figure 5-25 shows the distribution of ϕ in the outer interfacial region. The O-H bond vector is taken as that from the oxygen to the hydrogen atom. This profile unequivocally shows a bimodal distribution of the O-H bond vector orientations, with one orientation having the vector directed toward the bulk liquid phase (distribution

peak at around $cos(\phi) = -0.4$) and the other with the vector pointing out into the vapor phase (peak at $cos(\phi) = 1.0$). Figure 5-26 shows the distributions of ϕ moving into regions toward the bulk; the delta = 0.2 nm curve shows the distribution for molecules located in a 0.2 nm slice away from the interface, and the delta = 0.3 nm curve shows the same for molecules in an adjacent one nanometer slice further into the bulk region. In the layer directly next to the outer interface, the distribution indicates a planar orientation with the hydrogens lying in the plane of the interface. In the next inner region, the distribution appears to split again, the water molecule adopting an orientation that is slightly askew from the planar configuration. This appears to indicate that one O-H vector aligns in the bulk and the other protudes towards the upper layer; this is in essence a mirror image of the distribution of the outer interface (the distribution being inverted through the origin, for instance). These curves indicate that there is an alternating structure in the interfacial region and its vicinity. This strucutre oscillates between layers of planar water molecules and molecules whose dipole lies in the interfacial plane, but with the hydrogen atoms displaced from the in-plane configuration. This structuring is shown in Figure 5-15. An interesting note is that although similar distributions for the SPC/E potential model have been published by Taylor *et al* [54], it seems that the interpretation differs. The authors claim that in the liquid side of the interface, the hydrogens lie in the plane of the interface or *are directed into the bulk with an angle of* 94 *degrees relative to the interface normal.* This interpretation of the molecular orientation does not seem consistent with the data presented.

Now, the same analysis of the distributions of the water dipole vector and O-H bond vectors is presented for the liquid-liquid interface. For this analysis, the initial task was to determine how many distinct regions are definable at this interface. This was rather straightforwardly done by taking 0.2 nm slices and computing the distributions for each slice beginning at 4.0 nm from the center of mass and moving out towards the bulk water phase. The results of these calculations essentially reveal three zones in which one sees a characteristic structure. These are termed for the present analysis the outer, inner, and bulk regions. The outer region extends from

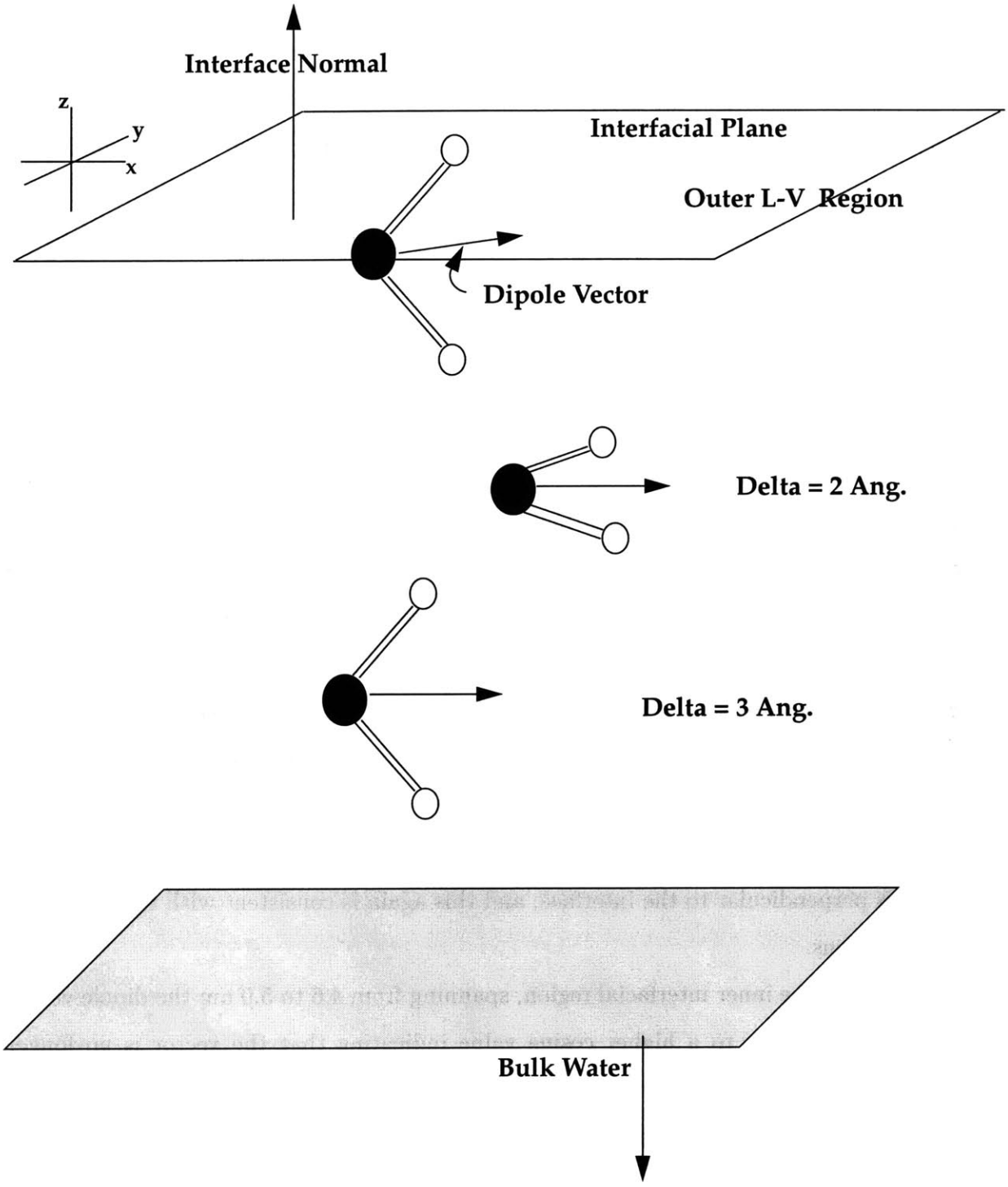


Figure 5-15: Schematic description of water molecular orientation at the water liquid-vapor interface.

4.0 to 4.6 nm (it is referred to as the 'outer' region based on the perspective from the bulk water phase). The next region, the inner region, spans from 4.6 to 5.0 nm, and the final bulk region is taken here to be from 6.0 to 7.5 nm, although this could be extended to lower values. Figures 5-28 and 5-27 shows the profiles computed over 600 configurations for slices of 0.2 nm thickness starting from 4.0 nm. These demonstrate the regions just described.

Figure 5-30 and Figure 5-29 show the average dipole vector and O-H bond vector orientation distributions for the regions delineated in the figures. In the outer region, the water dipole vector highly prefers to orient along the interface normal in a configuration which allows the hydrogens to be protruding towards the ester phase. This is sensible from the perspective that this configuration allows for a sufficient extent of hydrogen bonding interaction between the water hydrogens and the ester carbonyl oxyen. The ester-water hydrogen bonding is indicated by the hydrogen-oxygen radial distribution functions to be presented further below. Furthermore, the orientation of the dipole vector along the interface normal is also suggested by the P_1 and P_2 profiles shown in Figure 5-21 and Figure 5-22. In these figures, within the region from 4.0 to 4.6 nm, the P_1 profile is shows a marked orientation perpendicular to the interface (the P1 profile is positive here since this curve is computed for the vector antiparallel to the dipole vector). The average negative value of P1 corresponds to the distribution being peaked at -1. Also, the P2 profile indicates that the orientation is perpendicular to the interface, and this again is consistent with the computed distributons.

Now, in the inner interfacial region, spanning from 4.6 to 5.0 nm the dipole vector distribution shifts to a higher cosine value indicating that the vector is no longer directed identically along the interface normal, but adopts an orientation slightly skewed from the vertical with the dipole vector jutting towards the ester phase (but not as dramatically as in the outer region). This behavior corresponds to the behavior of the P1 and P2 curves in this region as well. The P1 profile increases moving in towards the bulk water (again, remember, to compare the distribution and P1 profile curves, the P1 curve should be flipped); this corresponds to the shifting to

higher values of the cosine distribution. Furthermore, the P2 profile dips slightly below zero to about -0.5; this suggests that the orientation in this region is more planar than in the outer region and this again is consistent with the behavior of the distribution. The O-H bond vector distribution for the inner region shows a slightly bimodal distribution. One hydrogen is essentially jutting towards the outer interfacial region (which should have an excess of oxygen atoms for hydrogen bonding), and the other is in the plane of the interface, an orientation which allows the latter hydrogen and the oxygen to also take part in hydrogen bonding interactions with the bulk.

Finally, in the bulk, the orientations are essentially random.

Figure 5-16 shows the possible configurations suggested by the above data.

We note here similar results reported in the literature for simulations of water at different interfaces and described with varying potential models. Per Linse [50] reports in an early work of the preferential orientation of water molecules at the benzene-water interface. Using P_1 and P_2 order parameters, the first and second Legendre polynomials of the projection of the water dipole vector along the interface normal, Linse reports that at the water-benzene interface, the water dipole moment vector prefers an orientation lying in the plane of the interface; this is based on a the P_2 value being negative in the interfacial regions. This is similar to water at the liquid-vapor interface since there is an energetic penalty in addition to the entropic loss of ordering, due to the loss of hydrogen bonding if the dipole vector were to deviate too much from planarity. Thus, the current results appear to corroborate Linse's interpretation of the preferential structuring of water in the interfacial region. Studying the water liquid-vapor interface using the Lemberg-Stillinger-Rahman (LSR) potential, Townsend and Rice [53] reported that the bisector of the HOH angle tends to lie in the plane of the interface; however, since specification of only the bisector-normal angle is not sufficient to unambiguously define the molecular orientation, the authors calculate the angle between the molecular plane and a specified fixed plane to pinpoint the orientation of the oxygen-hydrogen bonds. From their analysis, they report that although the molecular dipole vector lies in the interfacial plane, the molecule orients such that one O-H bond vector lies outside the interface, and the other within. Similar behavior

was reported by Matsumoto and Kataoka using the Carravetta-Clementi potential [55]. These authors, in an earlier study, proposed two typical molecular orientations for the water–one corresponding to the inner interface and a second pertaining to the outer interface. I n the vapor (outer) region of the interface, the water molecule orients with one O-H bond projecting out of the interface and the other into the bulk. In the liquid side, the molecule adopts a planar orientation with both hydrogens in the plane of , or slightly projecting inwards from, the interface. Zhang *et al* [57] report the bimodal distribution of water orientation in simulations of water-octane interfaces. Taylor, Dang, Garrett [54] also report the bimodal water structuring at the liquid-vapor interface. They claim that on the vapor side of the interface, the most probable orientation of that with the water dipole directed out of the liquid at an angle of 74 degrees relative to the interface normal; on the liquid side the water molecules are oriented such that the dipoles are in the plane of the interface. They also show that on the vapor side, one of the O-H bonds is protruding into the vapor phase, while in the next layer closer to the liquid, both bonds lie in the interfacial plane. The present work does indicate the more planar orientation of water at the liquid-vapor interface as evidenced by the P_2 profile. The negative P_1 profile at the liquid-vapor interface simply means that there is a tendency for the molecular dipole moment vector to be point ing out into the vapor (this is consistent with the report of Talyor *et al*). Considering the present results closely, the liquid-vapor interface does indicate a two-region splitting for the water. Consider the the outer liquid-vapor interface. In the outermost region, the P_1 deviates most from zero, indicating the dipole vector jutting in to the vapor phase. The P_2 shows that this vector is planar, but, as one moves into the inner region, the P_2 becomes more negative, signifying that the vector is progressing towards a more planar orientation; commensurately, the P_1 is increasing, signifying that the vector is moving towards a planar orientation before randomizing in the bulk.

Thus, in considering the results of the water orientation at the liquid-liquid and liquid-vapor interfaces, we are confident of the orientation we see based on the reported literature. Furthermore, what is more interesting is the fact that such local

behavior is observed irrespective of the type of method used to account for electrostatic interactions which are fundamentally at the root of the observed phenomena. For instance, Zhang *et al* [57] use simple coulomb potential shifted to zero at a spherical cutoff value of 12 angstroms. Matsumoto and Kataoka [55] use Ewald summation. The current work uses reaction field.

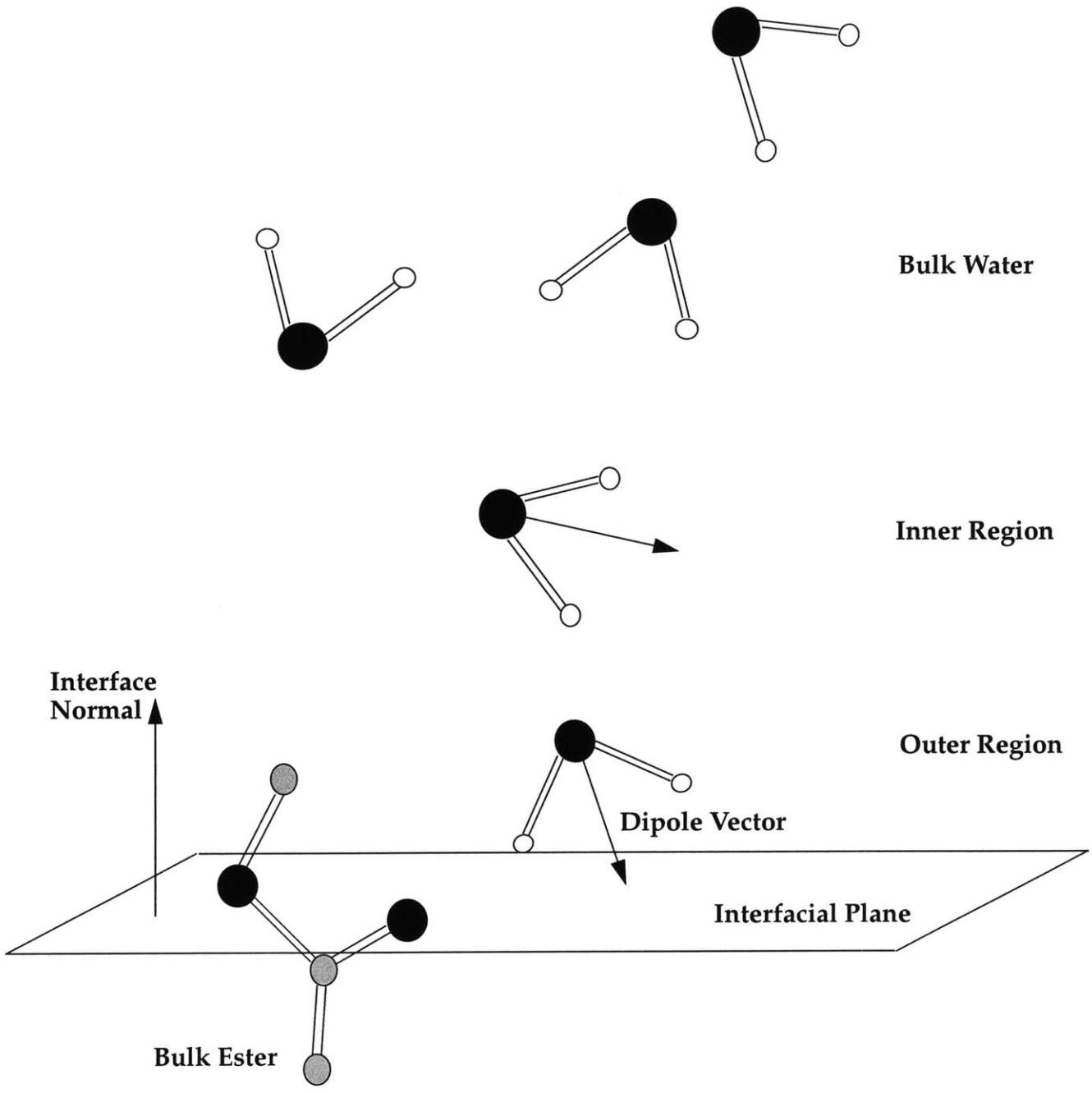


Figure 5-16: Schematic description of the water molecular orientation at the ester-water interface.

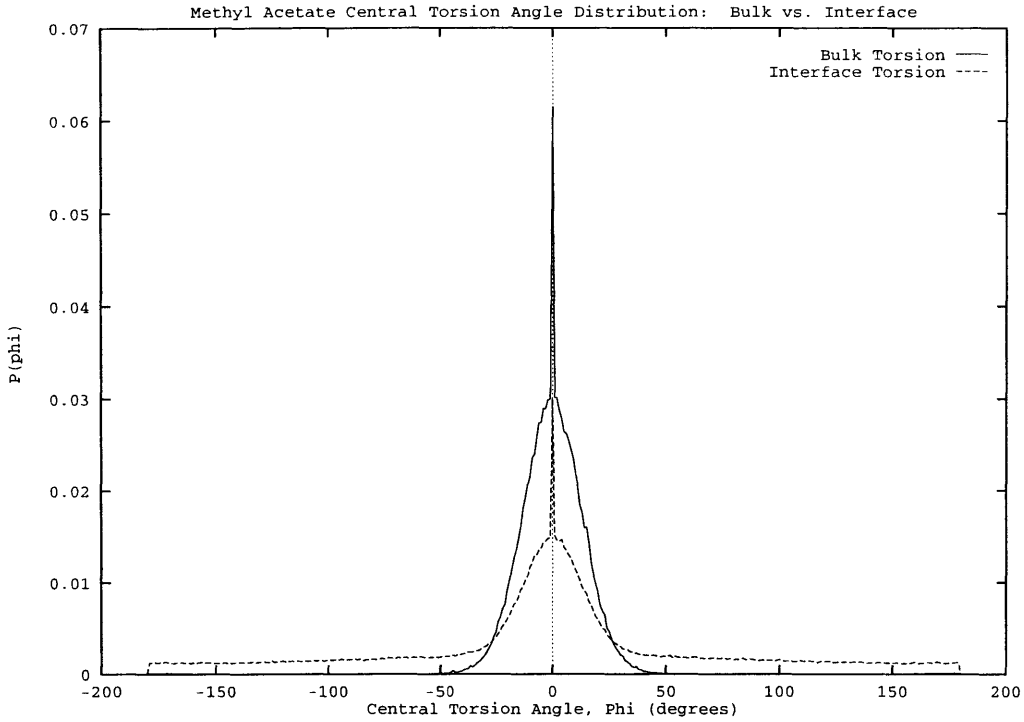


Figure 5-17: Methyl acetate torsion angle distributions for the bulk and interfacial regions

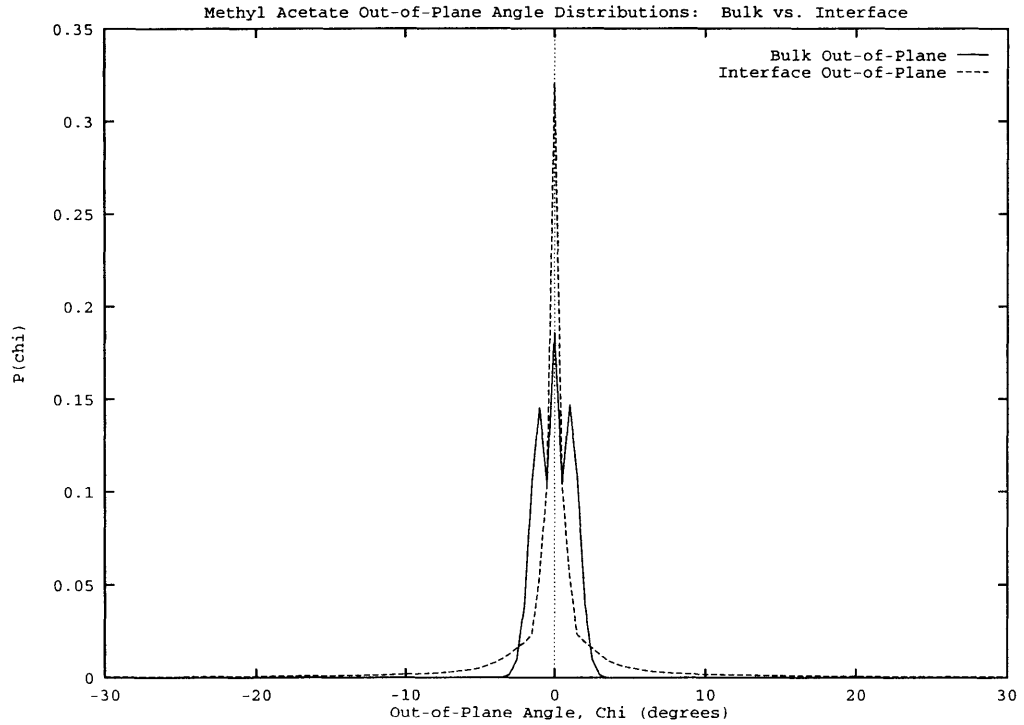


Figure 5-18: Methyl acetate out-of-plane angle distributions for the bulk and interfacial regions

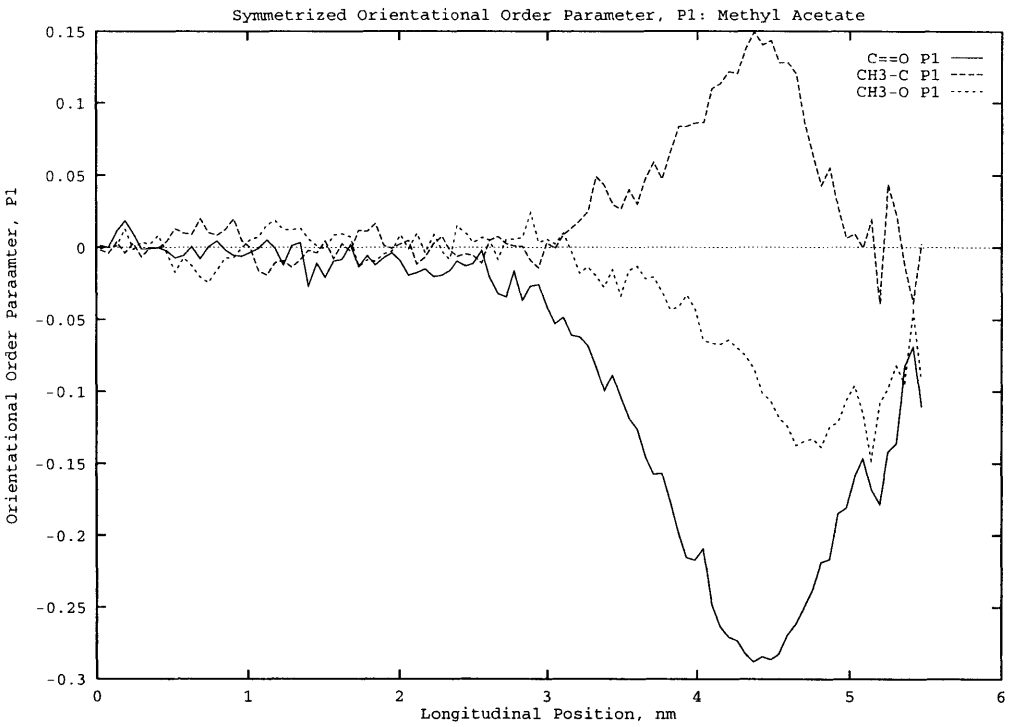


Figure 5-19: Symmetrized P1 profile for methyl acetate. The longitudinal position is relative to the total simulation cell center of mass.

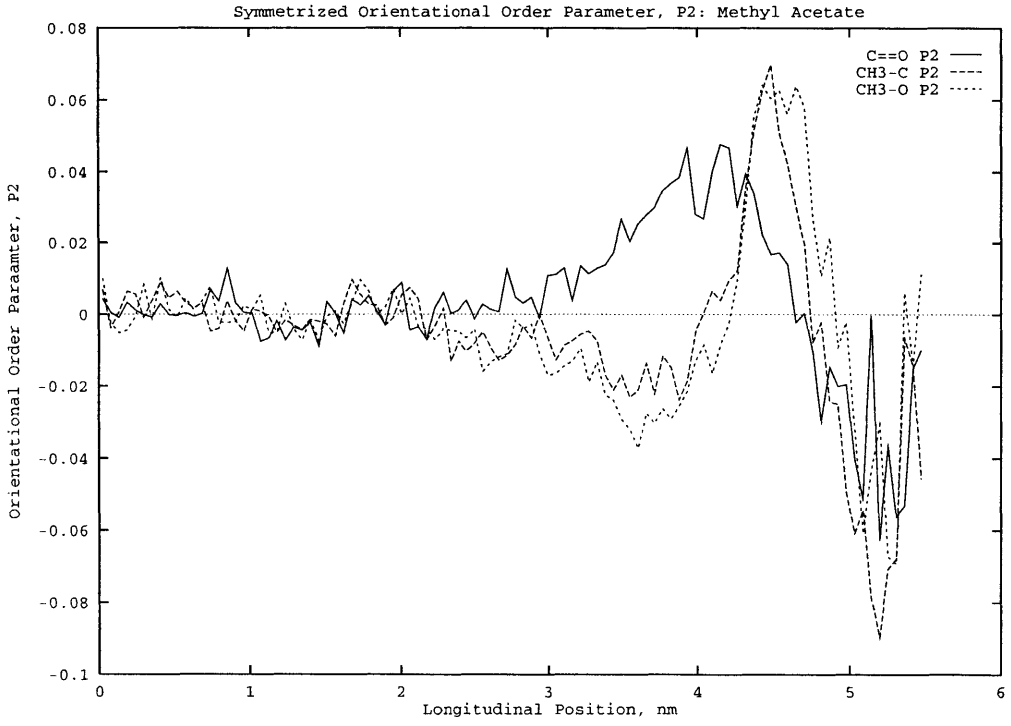


Figure 5-20: Symmetrized P2 profile for methyl acetate. The longitudinal position is relative to the total simulation cell center of mass.

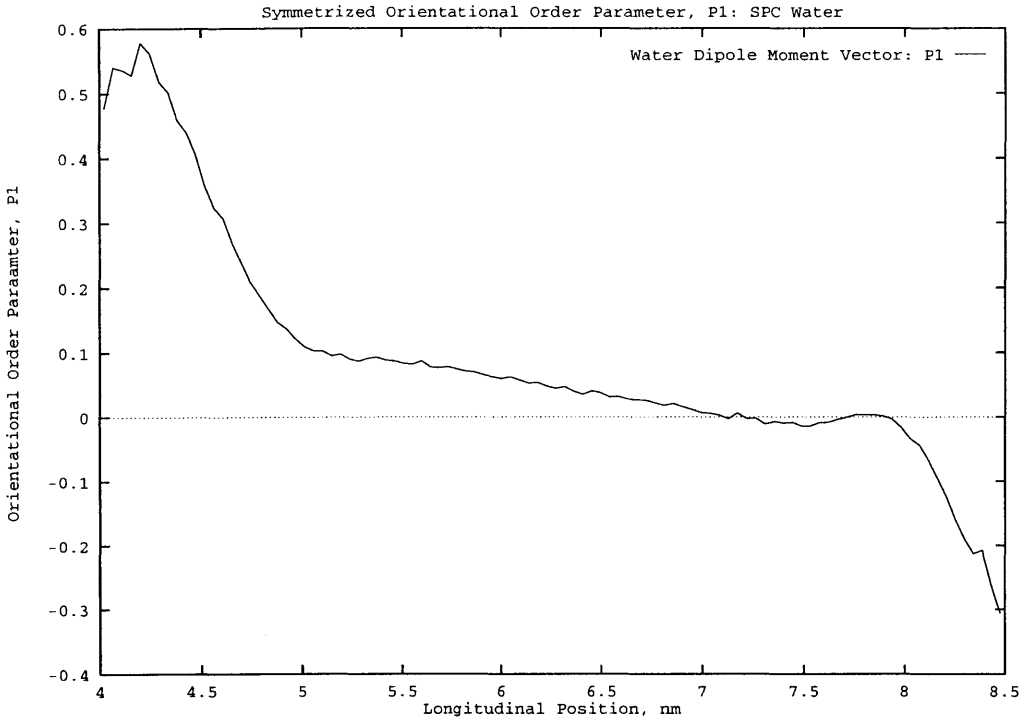


Figure 5-21: Symmetrized P1 profile for SPC water. The longitudinal position is relative to the total simulation cell center of mass.

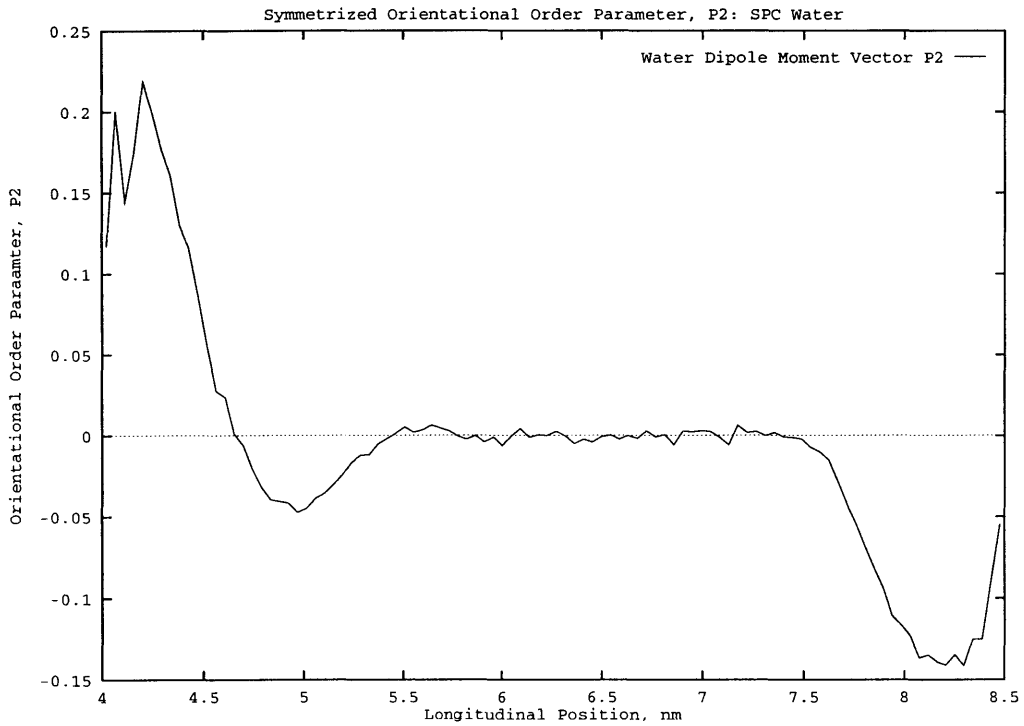


Figure 5-22: Symmetrized P2 profile for SPC water. The longitudinal position is relative to the total simulation cell center of mass.

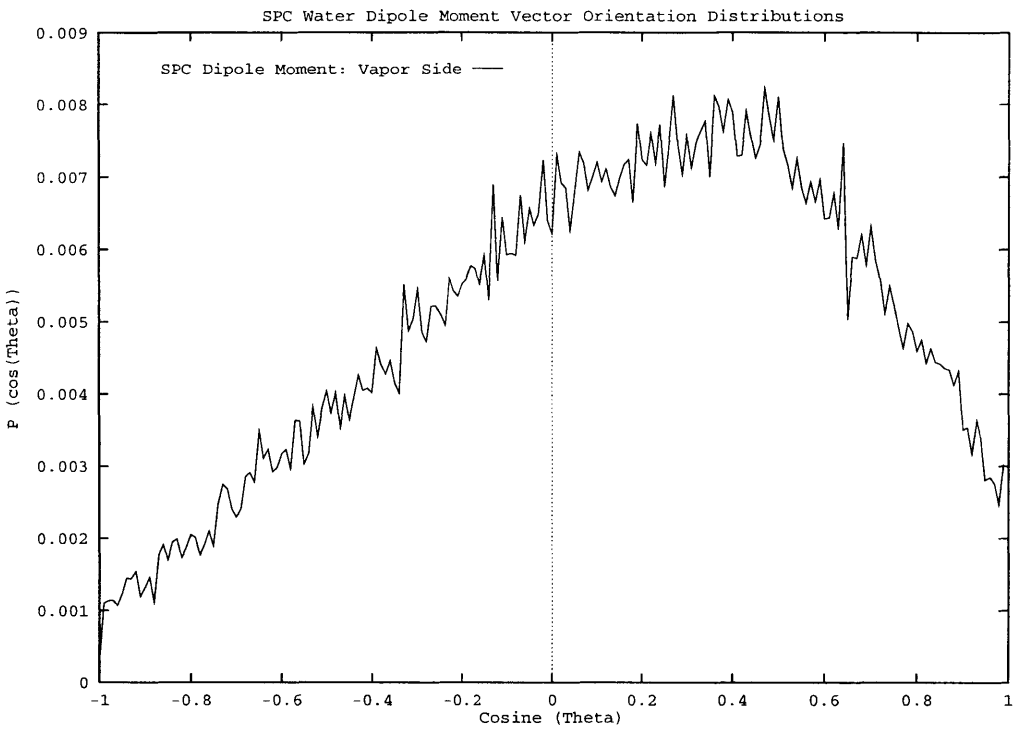


Figure 5-23: Distribution of the angle between the water dipole and interface normal: vapor side of the liquid-vapor interface

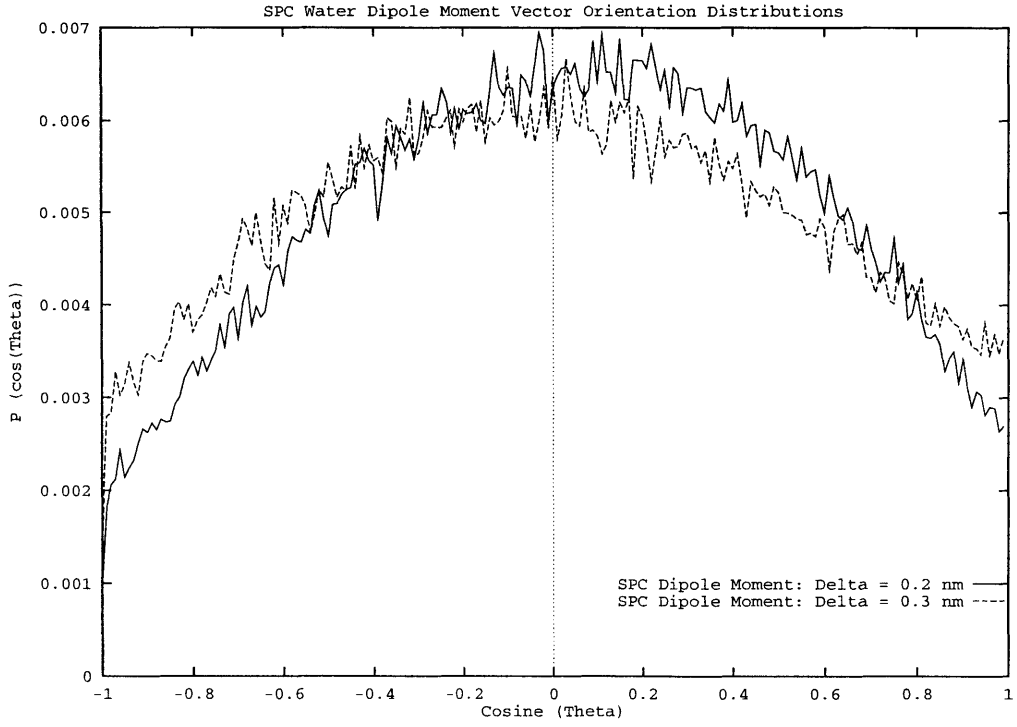


Figure 5-24: Distribution of the angle between the water dipole and interface normal: 2 and 3 angstroms from the outer interfacial region

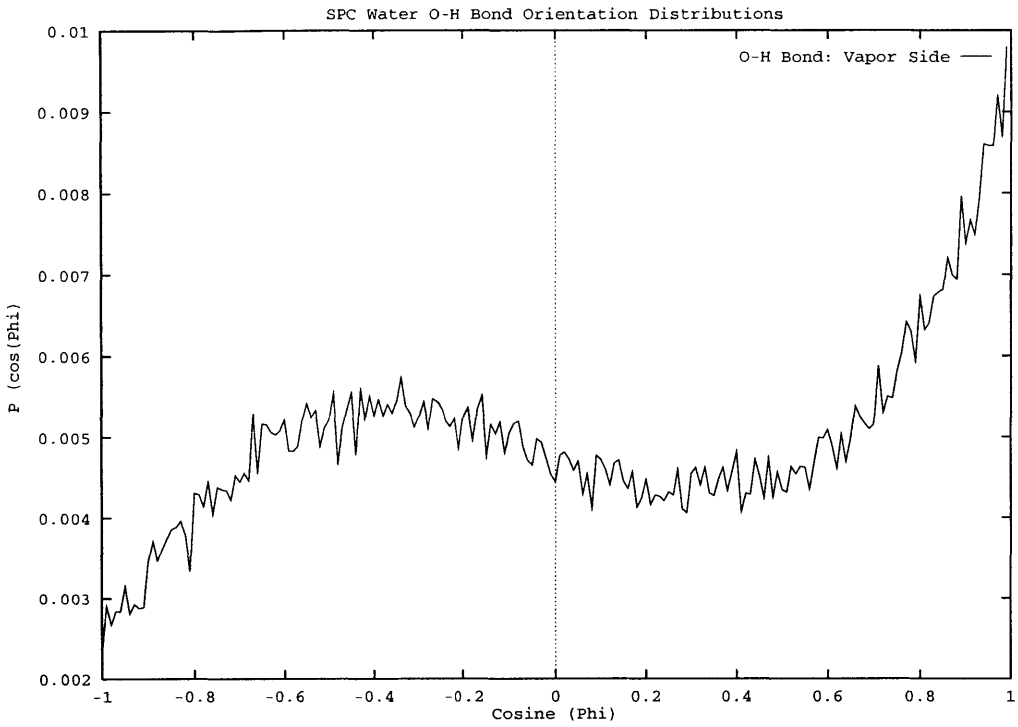


Figure 5-25: Distribution of the angle between the O-H bond and liquid-vapor interface normal: vapor side of the L-V interface

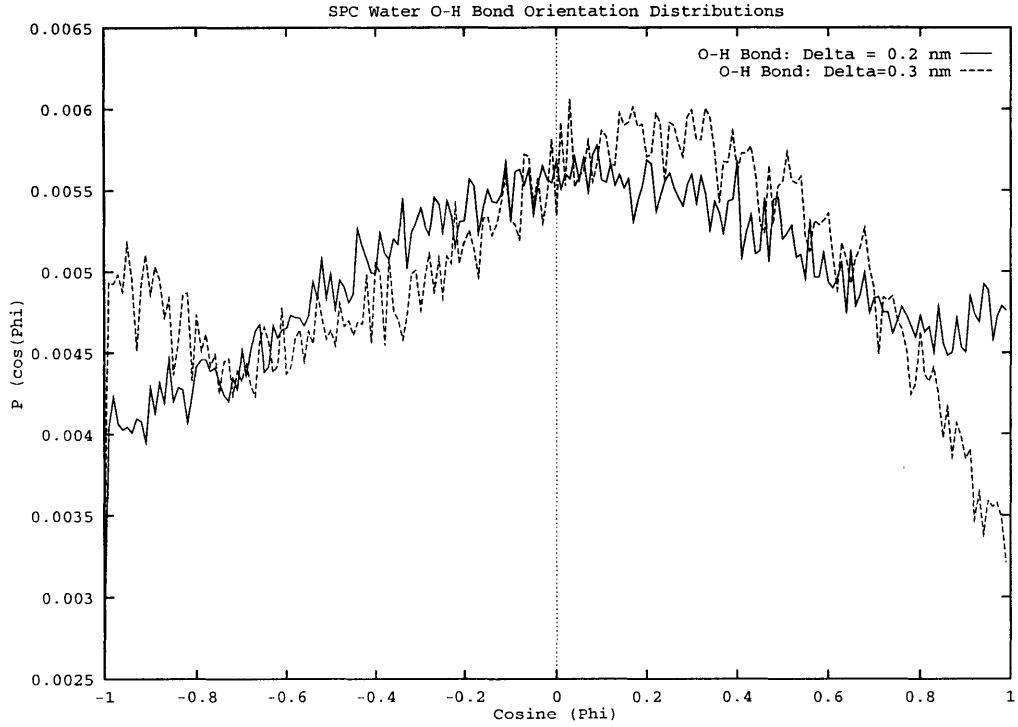


Figure 5-26: Distribution of the angle between the O-H bond and interface normal: 2 and 3 angstroms from the outer interfacial region

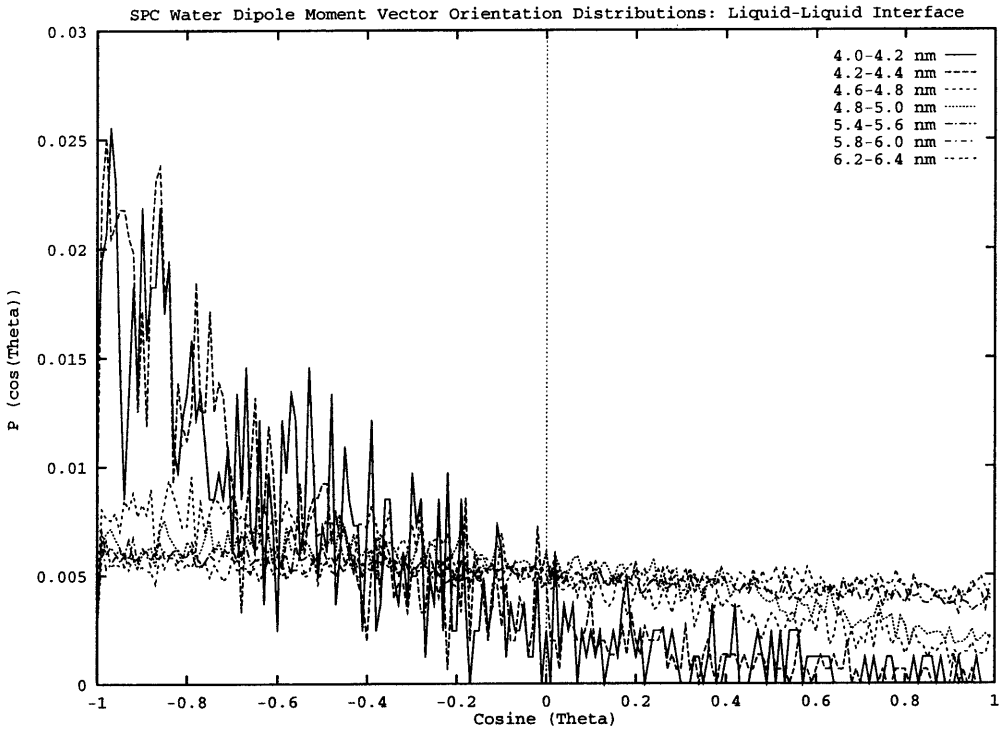


Figure 5-27: Distributions of the angle between the water dipole and interface normal showing the evolution of three distinct regions of characteristic structure.

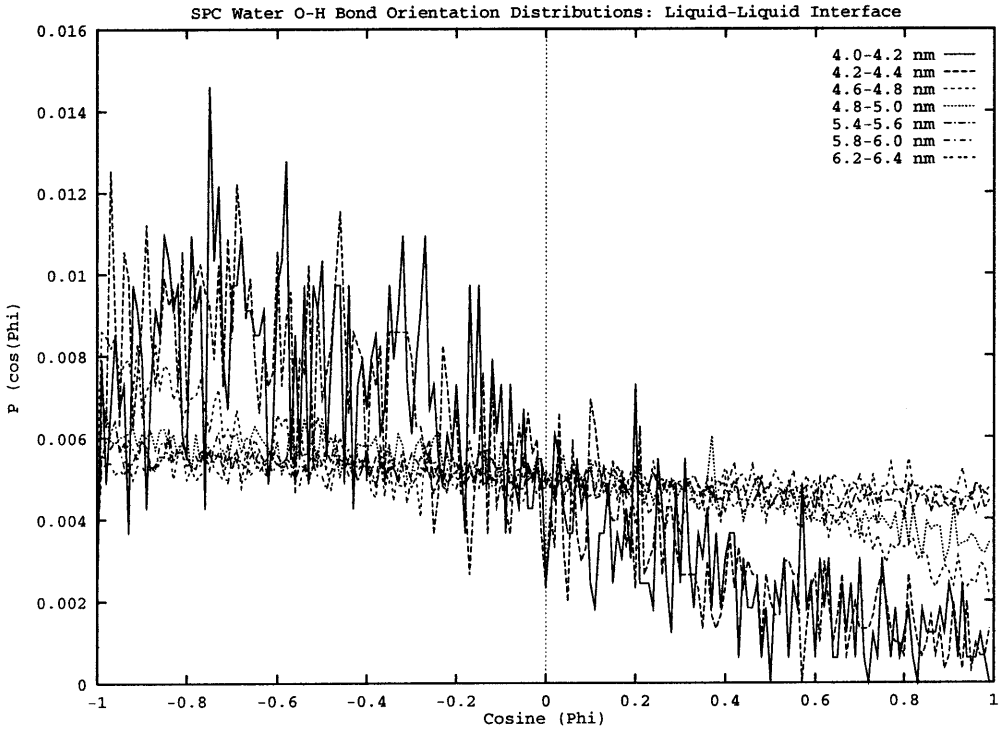


Figure 5-28: Distributions of the angle between the O-H bond and interface normal showing the evolution of three distinct regions of characteristic structure.

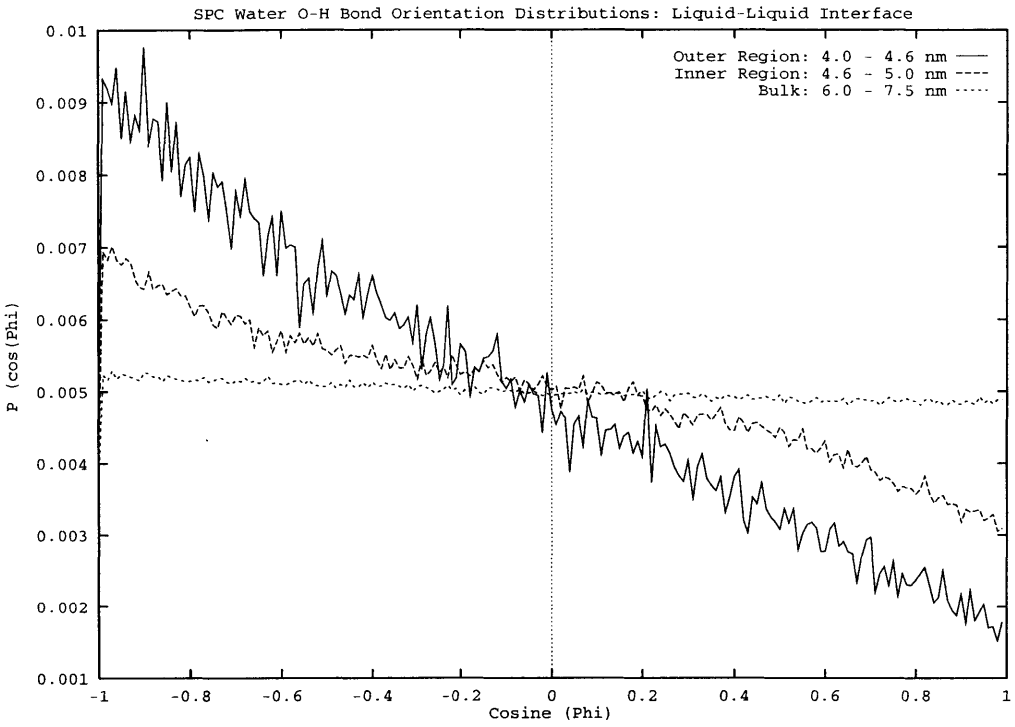


Figure 5-29: Distribution of the angle between the O-H bond and interface normal at the ester/water liquid-liquid interface.

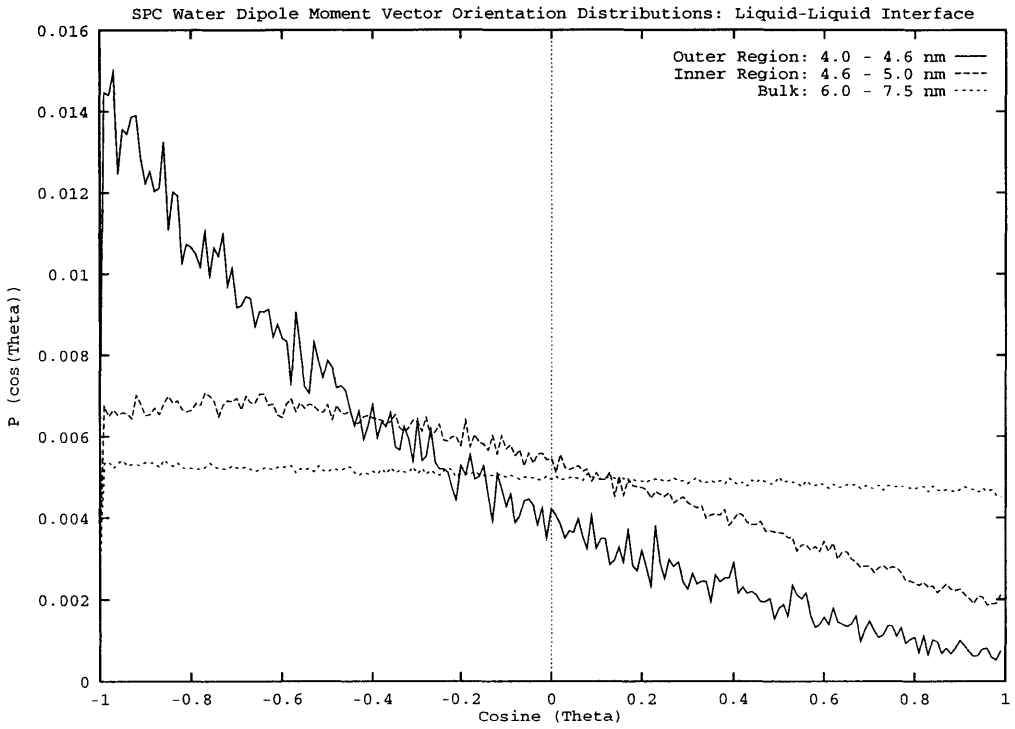


Figure 5-30: Distribution of the angle between the water dipole and interface normal at the ester/water liquid-liquid interface.

Comparison of Orientation at Liquid-Liquid and Liquid-Vapor Interfaces

Much like the effect of the type of interface on the water molecular orientation, Figures 5-31 though 5-36 demonstrate the change in orientation of the methyl acetate bond vectors. As discussed earlier, the reorientation of the surface groups is driven by the electrostatic interactions. The carbonyl oxygen - carbonyl carbon bond vector is dramatically flipped between the two interfaces, as the increased hydrogen-bonding is energetically favorable, overcompensating for any entropic loss due to the structuring.

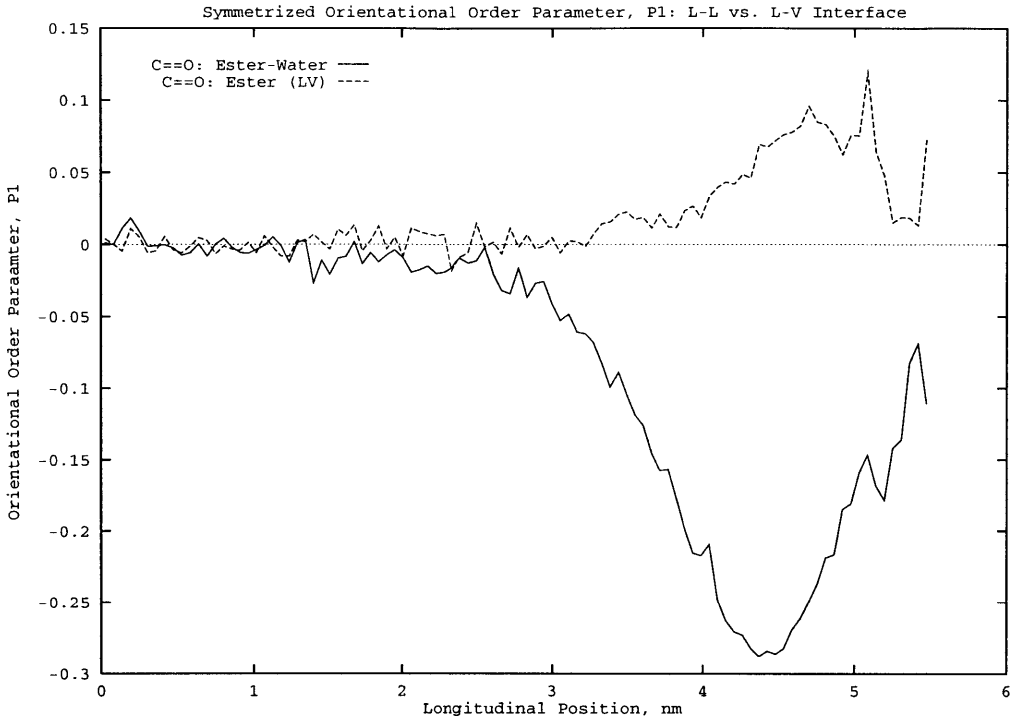


Figure 5-31: Symmetrized Methyl Acetate P1 profiles for the liquid-liquid and liquid-vapor interfaces: Carbonyl Bond Vector.

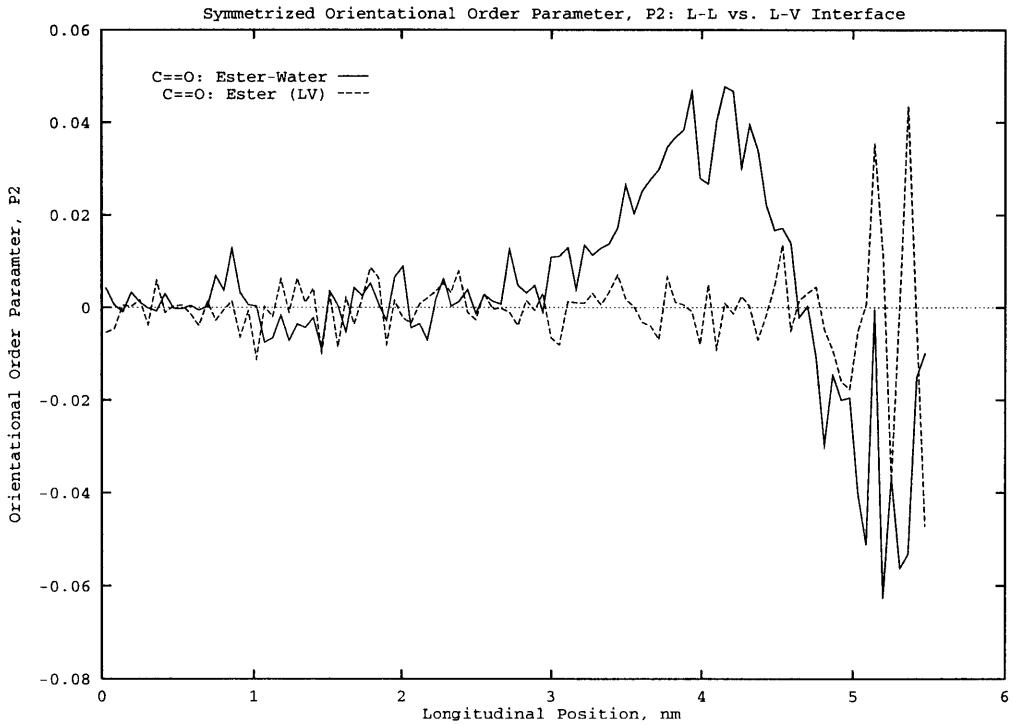


Figure 5-32: Symmetrized Methyl acetate P2 profiles for the liquid-liquid and liquid-vapor interfaces: Carbonyl Bond Vector.

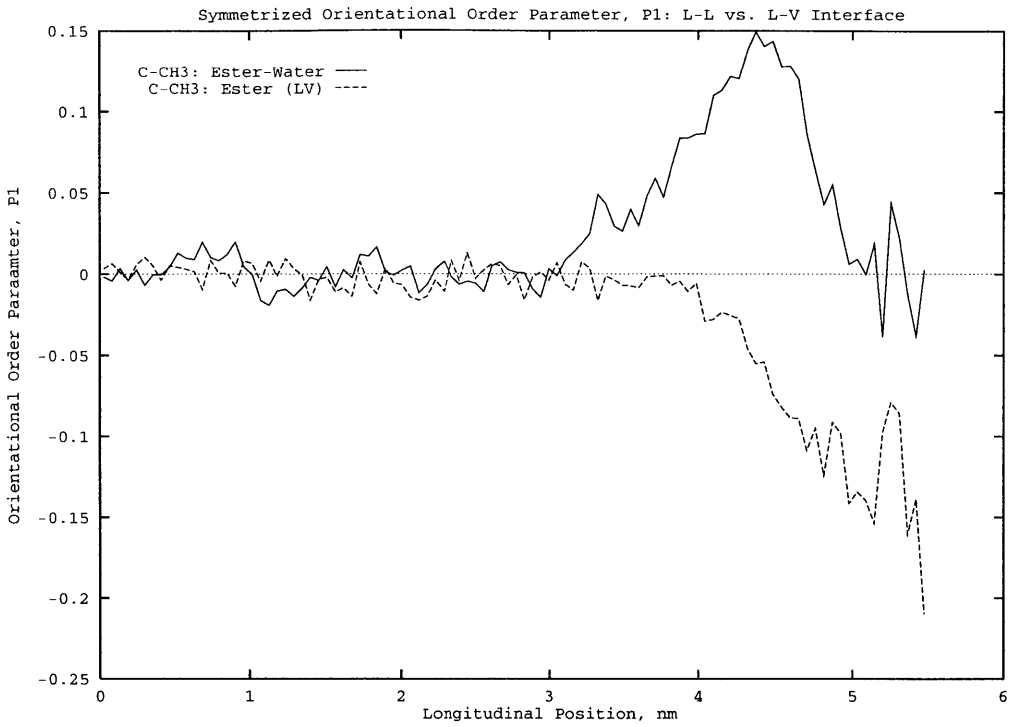


Figure 5-33: Symmetrized methyl acetate P1 profiles for the liquid-liquid and liquid-vapor interfaces: Carbonyl Methyl Group.

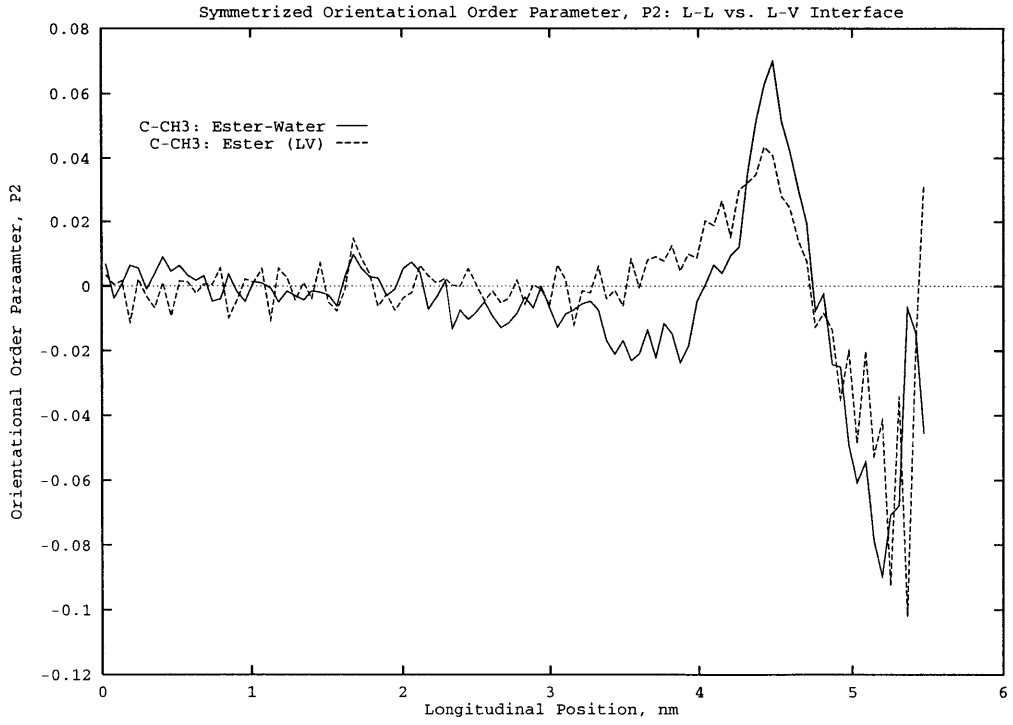


Figure 5-34: Symmetrized methyl acetate P2 profiles for the liquid-liquid and liquid-vapor interfaces: Carbonyl Methyl Group.

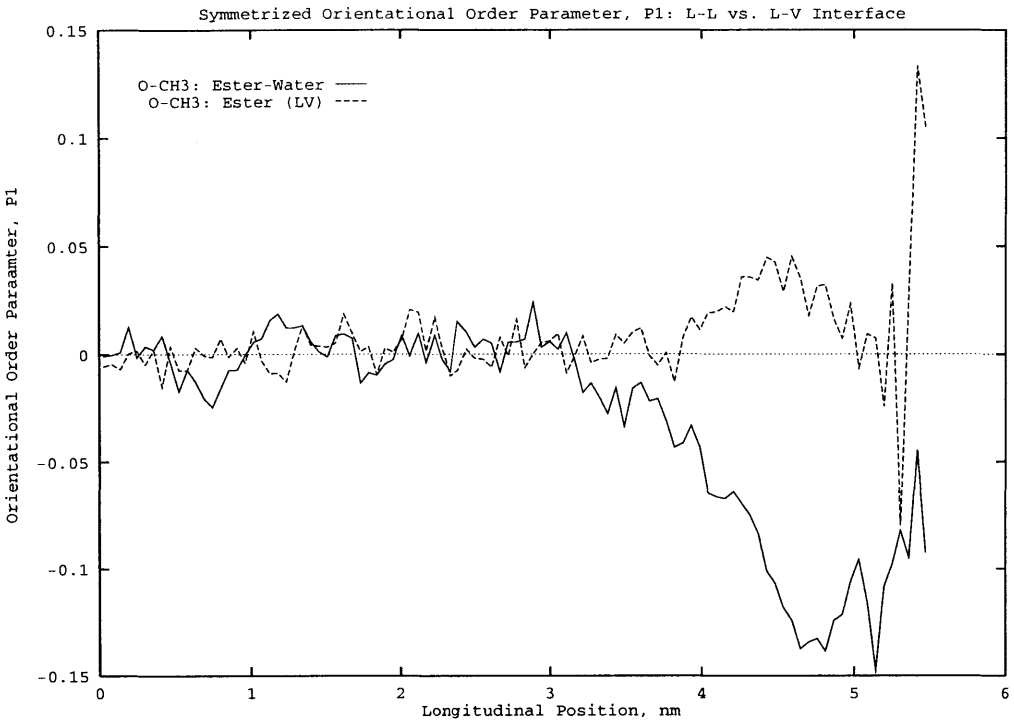


Figure 5-35: Symmetrized methyl acetate P1 profiles for the liquid-liquid and liquid-vapor interfaces: Methoxy Methyl Group.

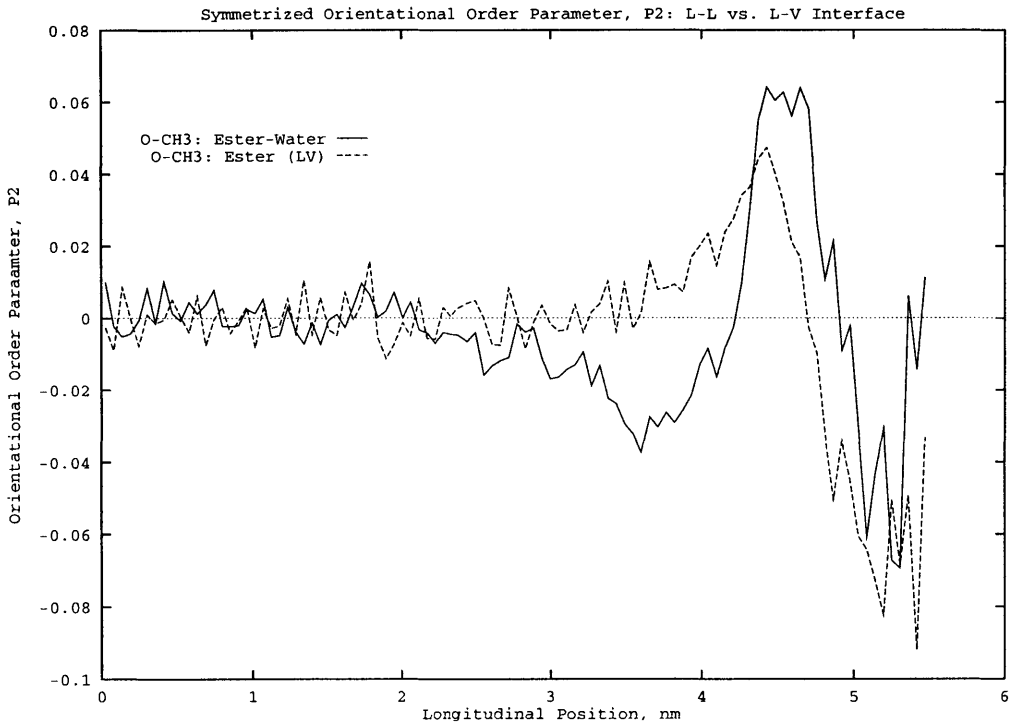


Figure 5-36: Symmetrized methyl acetate P2 profiles for the liquid-liquid and liquid-vapor interfaces: Methoxy Methyl Group.

5.3.3 Radial Distribution Functions

In this section the radial distribution functions computed for various slices along the normal direction are presented in order to investigate the local environment changes in moving from the bulk to the interface region. Furthermore, the emergence of ester carbonyl oxygen and water hydrogen pair correlations further elaborates the picture of the orientation of molecules in the interface via a hydrogen bonding mechanism. The pair correlation function for a slice is calculated from the histogram of the separations between all pairs of particle types within the slice. The normalized pair correlation function is given by [131]:

$$g(R) = \frac{2N(R, \Delta R)}{V_{int} N_T^2} V_s \tag{5.3}$$

where $N(R, \Delta R)$ is the average number of pairs of particles within the section with separation between R and $R + \Delta R$, V_{int} is the average volume of the intersection of the section with the space between spheres of radii R and $R + \Delta R$, V_s is the total slice volume, and N_T is the average number of particles in the slice. Note that for proper analysis, section widths should be equal when comparing the distributions in different regions of a system.

For methyl acetate, radial distribution functions (rdf's) are calculated for the bulk (the central cubic section of the system with dimensions 3.085 nm by 3.085 nm by 0.5 nm, an intermediate slice from 3.25 nm to 3.75 nm out from the center of mass, the inner interface region from 4.0 to 4.5 nm, and the outer interface region from 4.5 to 5.0 nm. Rdf's are computed for the following types of pairs: carbonyl oxygen - carbonyl oxygen (O-O), ether oxygen - ether oxygen (Oe-Oe), carbonyl oxygen - ether oxygen (Oe-O), carbonyl oxygen - ester methyl group (O-MeE), and ether oxygen - ester methyl group (Oe-MeE). In the interfacial regions, rdf's with water atom types are also computed. These are: carbonyl oxygen - water hydrogen (Hw-O) and ether oxygen - water hydrogen (Hw-Oe). Figures 5-38 through 5-42 show the methyl acetate rdf's for the defined slices. These are typical liquid state rdf's, quite featureless, and as Jorgensen reports, show a lack of any significant organized structure. The bulk rdf's

for all pair types show good agreement with those computed by Jorgensen from Monte Carlo simulations of bulk methyl acetate [67]. Thus,structurally, one is confident that the bulk region is equilibrated. Furthermore, as the interface is approached from the bulk, the rdf's reflect the fall in density with the increasing peak heights, but the liquid structure is not lost. Within the statistical error of the computation, the profiles for the bulk, intermediate, and inner interfacial regions are practically equivalent. This is consistent with the observation that the average densities in these regions are not too disparate from one another. Once the outer interfacial region is encountered, one sees a dramatic increase in the first peak (for the profiles for all pair distributions). Furthermore, it is the first peak which is enhanced the greatest; subsequent peaks are not affected significantly. The peak enhancment effect has been observed in prior simulations of water clusters, as will be discussed below with respect to the water radial distributions functions which also display such behavior.

More interesting to note are the ester-water rdf's shown in Figures 5-44 and 5-45 . Figure 5-45 clearly indicates hydrogen bonding interactions between the ester carbonyl oxygen and the water hydrogen with a characteristic strong first peak at 0.19 nm with a first minimum at 0.25 nm. This interpretation of enhanced hydrogen bonding within the interfacial region is further supported by a set of interesting theoretical calculations by Rablen and coworkers [140]. This group has performed density functional *ab initio* calculations with the Becke3LYP method [140, 141, 142] using the correlation functional of Lee,Yang, and Parr [142, 143, 140]. The authors report hydrogen-bonding geometries for two anti-periplanar geometries of methyl acetate, the Z1 and Z2 structures as shown in Figure 5-37,

The hydrogen-bond geometry for the Z1 structure is such that the carbonyl oxygen to water hydrogen distance is 0.1938 nm and the angle made by the carbonyl oxygen, water hydrogen, and water oxygen is 170.8; for the Z2 structure, the hydrogen bond distance is 0.1916 nm and the bond angle is 164.5. The water is the hydrogen donor to the ester carbonyl in both complexes. The distortion of the geometry from the idealized "linear/trigonal planar" structure is reported for the class of compounds containing the carbonyl functionalities including carboxylic acids, esters,

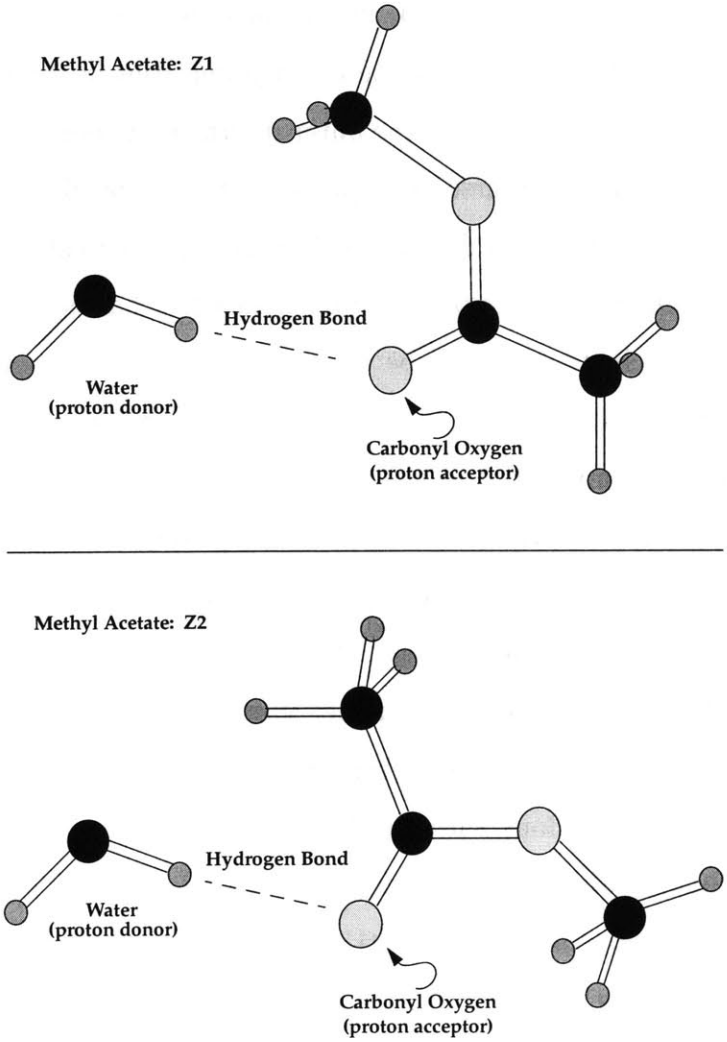


Figure 5-37: Becke3LYP/6-31+G(d(X+),p) optimized geometries of the Z1 and Z2 Hydrogen-Bond Complexes of Methyl Actate and Water

amides, aldehydes, and ketones. It is proffered that the interaction of the individual total molecular dipoles affects the resulting geometries. For hydrogen-bonded complexes of these functionalities with water, the water molecule distorts from the idealized geometry (with the $O \cdots H - - - O$ angle being 180 degrees and the $C = O \cdots H$ angle being 120 degrees in the position of the sp2-hybridized lone pair) to one in which the water molecule's dipole moment aligns more closely antiparallel to that of the carbonyl bond. The present MD simulations reveal the hydrogen bond distance to be 0.19 nm which is in excellent agreement with these higher level calculations; furtheremore, calculated distributions of the angle made by the carbonyl oxygen, the water hydrogen, and the water oxygen are peaked at an angle of 160.1 degrees. This result supports the picture of a hydrogen bonded structure based on interactions between the ester carbonyl oxygen and water in the interfacial region.

From the P_1 and P_2 profiles for the three methyl acetate bond vectors discussed earlier, one would anticipate that the typical conformation of the hydrogen bond complex is of the Z2 type as described by Rablen and coworkers. This is by no means an absolute structure in the sense that the Z2 geometry as given by the *ab initio* calculations should be the one resulting from the MD simulations; the theoretical geometries are strongly indicative of gas-phase hydrogen-bond complexes as the effects of surrounding media are neglected in these calculations (incorporation of these types of interactions may perhaps be affected by reaction field calculations on such systems). The fact that the present values of the first peak in the carbonyl oxygen-water hydrogen radial distribution function and the peak in the angle distribution profiles are in such agreement (at the very least in a qualitative sense) with the theoretical geometries is strong indication of a hydrogen-bonding effect contributing to the structure evident at the interface. Fundamentally, it is evident that the drive for the structuring, which leads to entropic losses, is the energetic gain due to maximization of interactions; the mechanism through which this occurs is the hydrogen-bonding.

This element of structure adds to the picture of the water and ester orientation at the interface resulting from the favorable hydrogen bonding type interaction between the species. Furthermore, there is virtually no hydrogen-bonding type interaction

with the ether oxygen as shown by Figure 5-44, an observation also reported by Jorgensen *et al* [67] for monte carlo simulations of acetic acid bulk liquid. Finally, one can note that there are fewer hydrogen bonds with the carbonyl oxygen as one moves to the outer interface where the ester density falls off sharply.

For water, radial distribution functions (rdf's) are calculated for the bulk (the central rectangular slice with dimensions of 3.085 nm by 3.085 nm by 0.5 nm) starting with a z-coordinate value of 6.25 and moving outward to a value of 6.75 nm from the center of mass, the inner and outer ester-water interface regions, and the water liquid-vapor interface which is here taken to extend from 8.0 to 8.5 nm from the center of mass. Figure 5-43, Figure 5-46, and Figure 5-47 show the water rdf's between the oxygen-oxygen and hydrogen and oxygen. Here there is nothing out of the ordinary in terms of the structure signified by the curves. They hydrogen-oxygen rdf's indicate hydrogen-bonding interaction via the strong first peak. Even the liquid-vapor interface retains the characteristic liquid structure, and in terms of local environment, is not much different from the inner ester-water interface region.

At this point, some comments are made about the behavior of the radial distribution functions as the interface is approached from the bulk region. In the case of the SPC water, Figure 5-46 and Figure 5-47 show the increase in the first peak heights as one moves to lower density regions. This same effect is seen for two other water models, the ST2 and LSR (Lemberg-Stillinger-Rahman) potentials, applied to the study of clusters of water molecules by Townsend and Rice [53]. The authors present oxygen-oxygen, oxygen-hydrogen, and hydrogen-hydrogen radial distribution functions for shells encompassing a bulk region and various positions within the inhomogeneous region (the interfacial region characterized by monotically decreasing density profiles). For both potentials studied by the authors, the oxygen-oxygen correlations became stronger as the density decreased, with a threshold for increasing amplitude of the first peak of the oxygen-oxygen distribution being about ninety percent of the bulk density. Furthermore, the second peak of the oxygen-oxygen distribution is independent of the position, showing a slight enhancement for densities less than 25 percent of the bulk. The oxygen-hydrogen and hydrogen-hydrogen

functions are shown to parallel the oxygen-oxygen functions. Based on these results, the authors claim that the behavior of the distribution functions suggest that with a breakup of the hydrogen-bonded structure of the bulk liquid in the lower density interfacial region, there is precipitated dimer formation; it is this dimerized structure which yields the stark enhancement in the first peaks of radial distribution functions.

In the present case, we observe very much the same effects. Figure 5-47 shows the oxygen-oxygen distributions for the bulk water region, the inner and outer liquid-liquid interfacial regions, and the water liquid-vapor interface. For the inner liquid-liquid region, the first and second peaks are dramatically enhanced relative to the bulk; this is a manifestation of the low average water density of this region (this region has the lowest average water density of the regions for which profiles are shown in this figure). As one moves into the outer liquid-liquid interface, where the water density is higher, the first peak is still enhanced, but the second peak shows a much smaller enhancement. Finally, the liquid-vapor region again shows enhancement of the first peak, much as the outer liquid-liquid region. The first peak is higher for the liquid-vapor region simply because of the lower density associated with this region. The general qualitative effect is the same. In the case of the oxygen-hydrogen distribution functions, Figure 5-46,we find that the density affects the first two peaks, with the third peak showing enhancement at very low densities as encountered in the inner liquid-liquid interface. For the present study, we have not proceeded to quantify the bounds on the density where the effects on the first, second, and higher peaks occur. The fact that this behavior seems to be an inherent characteristic of water potentials is an interesting result nevertheless.

Finally, turning to ester site distribution functions, Figures 5-38 to Figure 5-42, we observe a similar peak enhancing effect with decreasing density. Furthermore, these figures show a greater affect on all but the carbonyl oxygen - carbonyl oxygen function. This is understood if one notes the component density profile. This shows that the carbonyl oxygen is enriched in the outer interfacial region compared to the other substituents; thus, we are confident to see this smaller density difference of the carbonyl oxygen reflected in the associated radial distribution function. Now, as for

the underlying mechanism giving rise to this structure, we can say little with the information presented. It would be interesting to determine if, even in the case of the ester, there was a dimerization phenomenon taking place (or in a very general sense, some type of clustering event). This would tend to suggest that such a 'clustering' phenomenon may be a general characteristic of structural changes as a function of density in inhomogeneous systems.

At the end of this Chapter, we present the same radial distribution functions presented up to this point, but given with the associated statistical errors to demonstrate that the phenomena we see are actually significant. The reader is referred to that section, but its omission will not incur any loss of meaning for the rest of the Thesis.

One caveat to this analysis is the fact that there is an inherent limitation in the nature of the calculation of the *interfacial* radial distribution functions. One can note that the density over a small section within the monotonically decaying section of the density profile is dramatically changing. Now, the calculation of the rdf for this region requires normalizing the histogram of pair separations; this is done by employing a *uniform density* for this region. In so doing, the smaller pair separations are weighted more, and thus, one observes an enhancement in the first peak of the rdf in the interfacial region relative to the bulk. This effect becomes more pronounced as the degree of inhomogeneity within the region becomes more severe, since this causes the uniform density used for normalization to decrease substantially. This is a topic for further investigation and will not be considered further in this work; it would be interesting to reinvestigate the phenomenon of water clustering at the free surface (based on the previous studies which seem to show very similar results in terms of the rdf behavior) with the intent to consider if this effect is indeed 'real'.

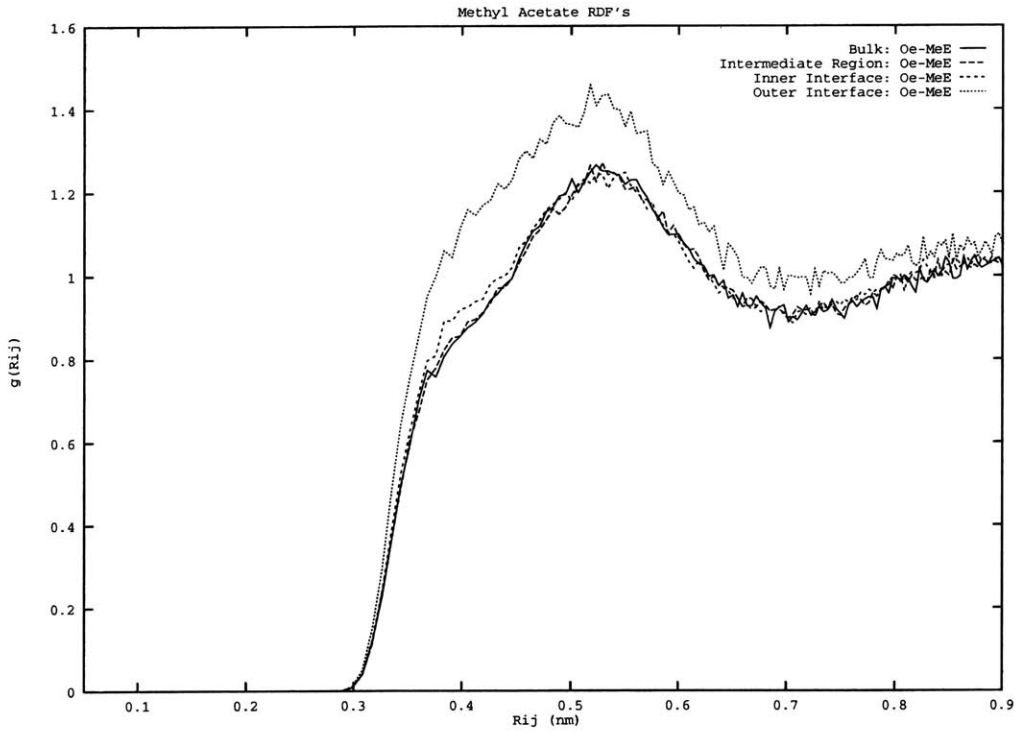


Figure 5-38: Methyl acetate radial distribution functions: ether oxygen - alkoxy methyl group

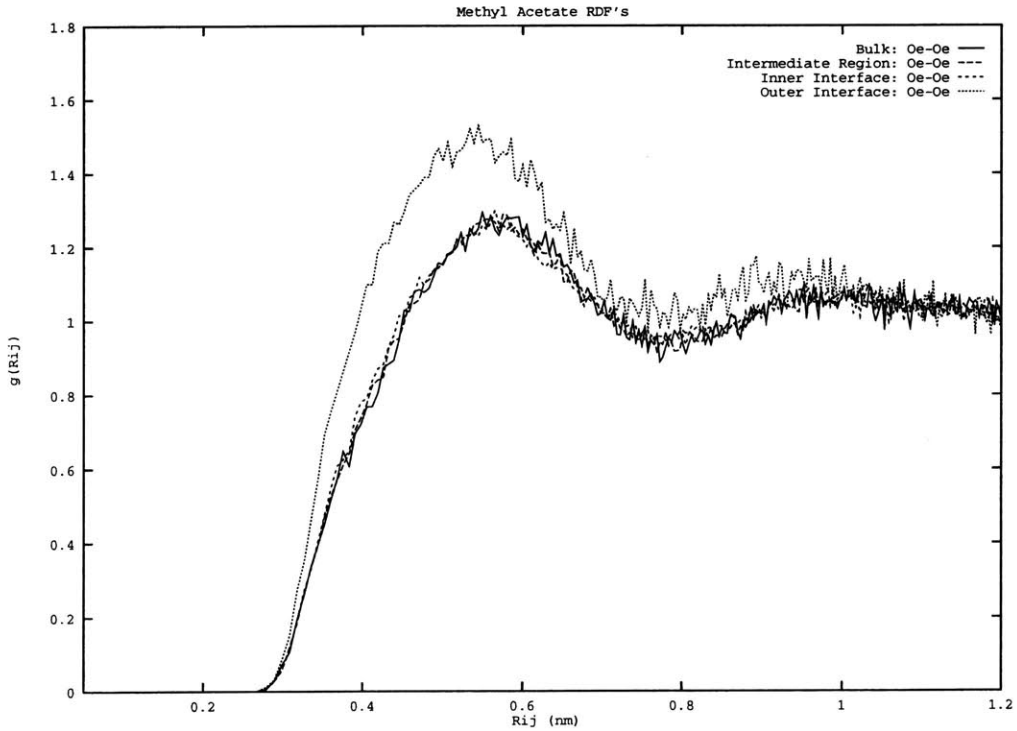


Figure 5-39: Methyl acetate radial distribution functions: ether oxygen - ether oxygen

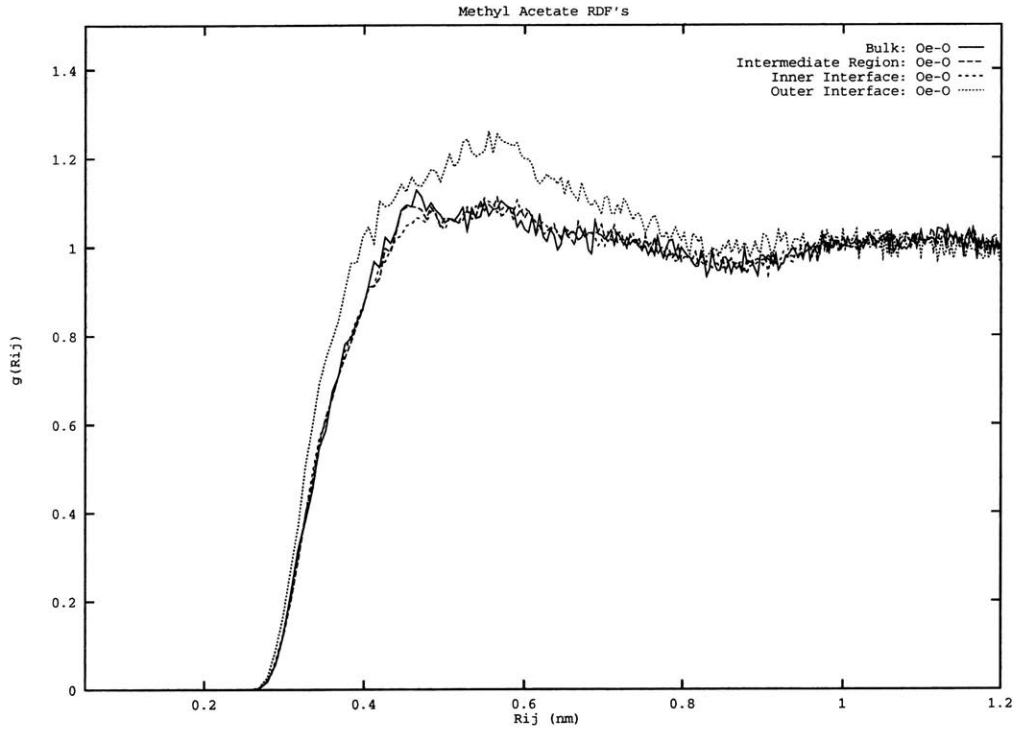


Figure 5-40: Methyl acetate radial distribution functions: ether oxygen - carbonyl oxygen

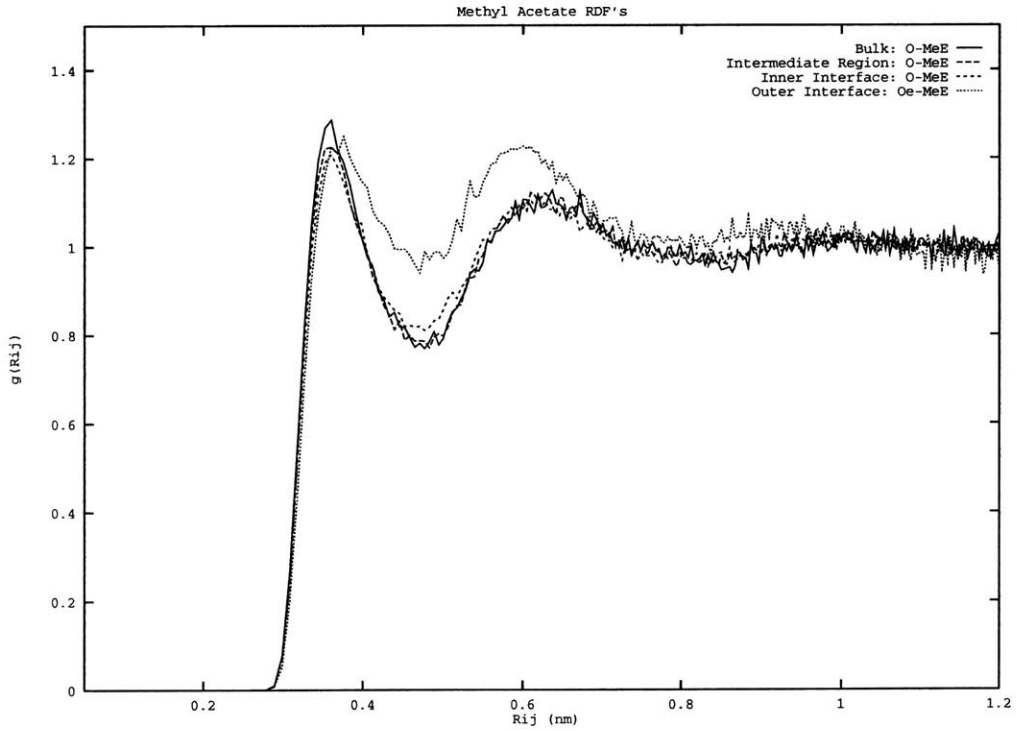


Figure 5-41: Methyl acetate radial distribution functions: carbonyl oxygen - alkoxy methyl group

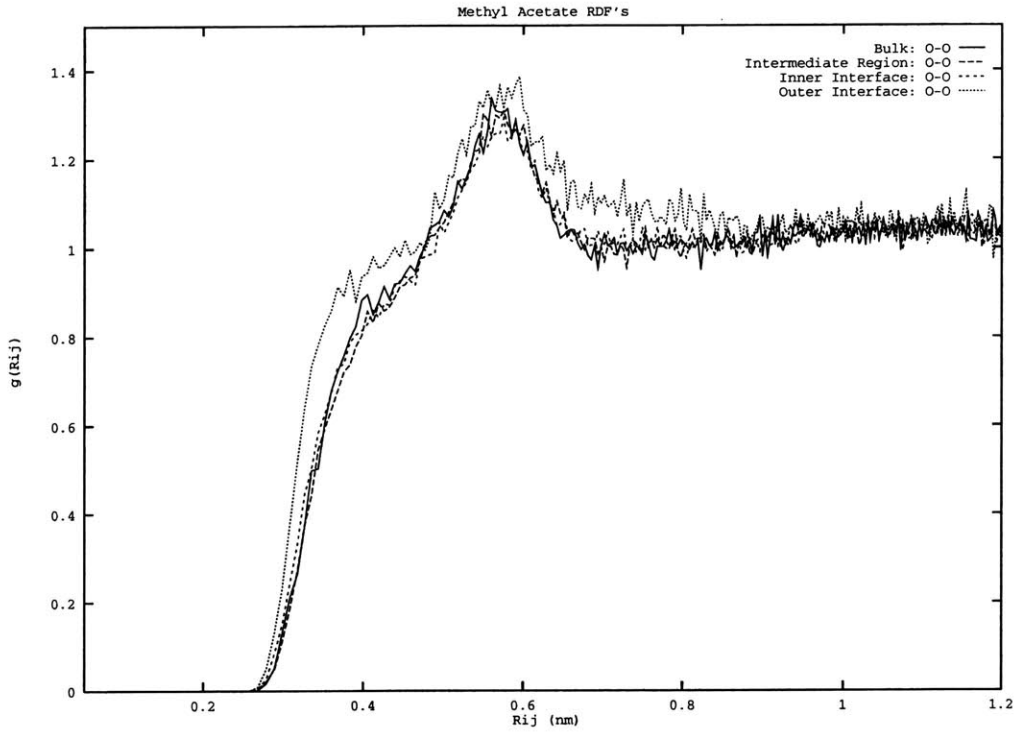


Figure 5-42: Methyl acetate radial distribution functions: carbonyl oxygen - carbonyl oxygen

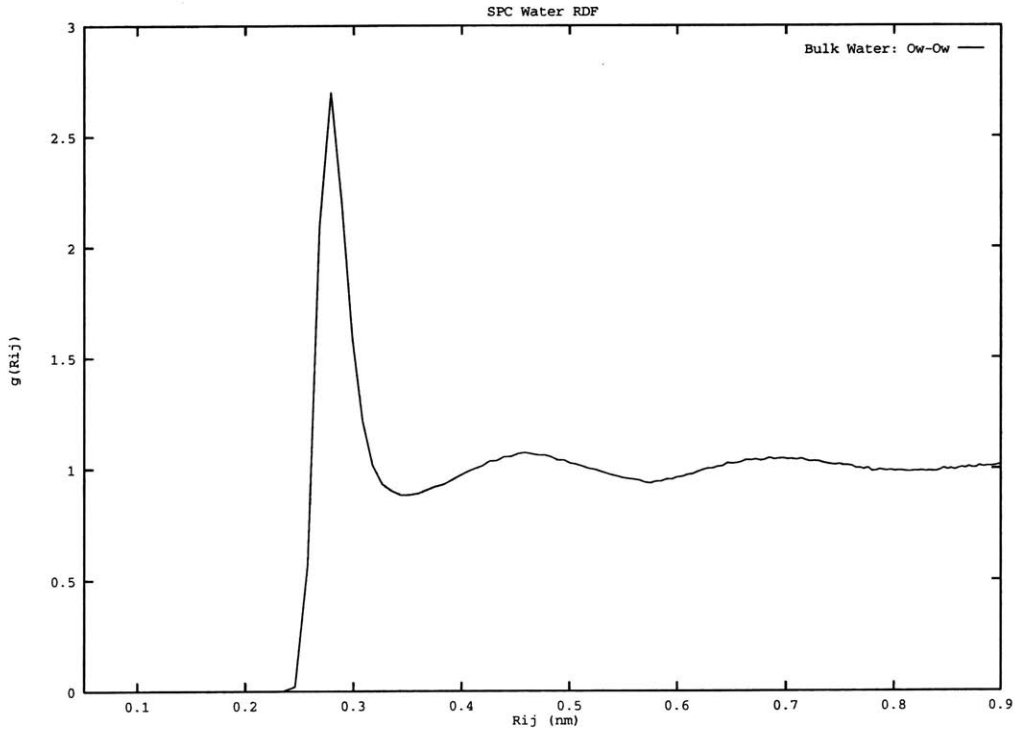


Figure 5-43: Bulk Water oxygen-oxygen radial distribution function

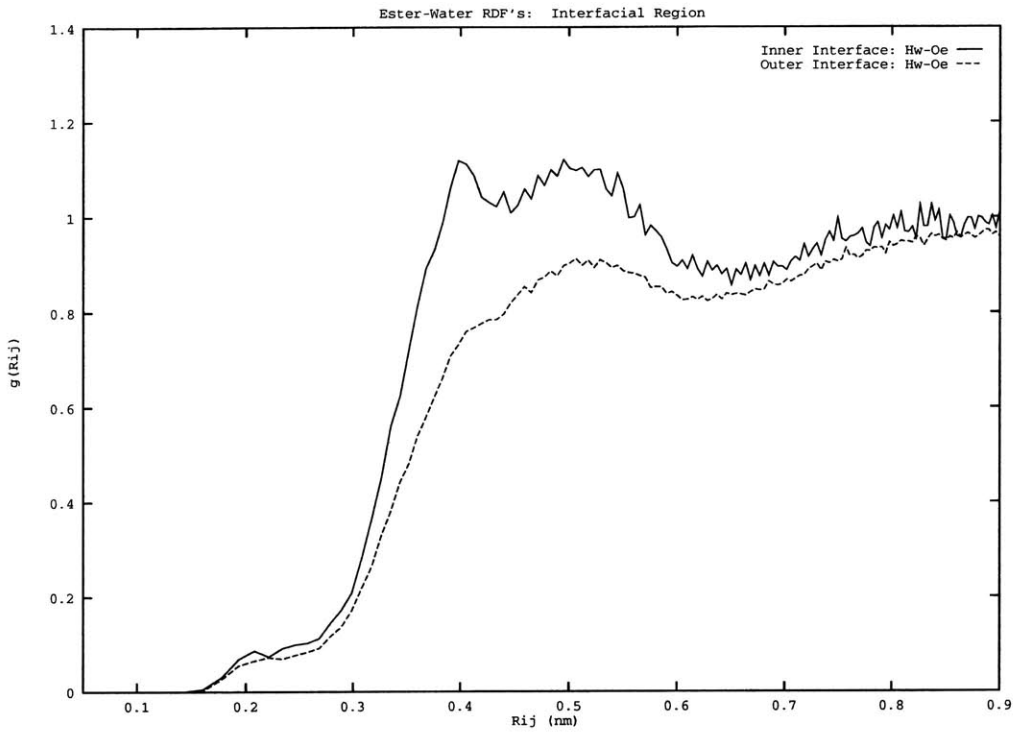


Figure 5-44: Water hydrogen - ester ether oxygen radial distribution function in the inner and outer liquid-liquid interfacial regions.

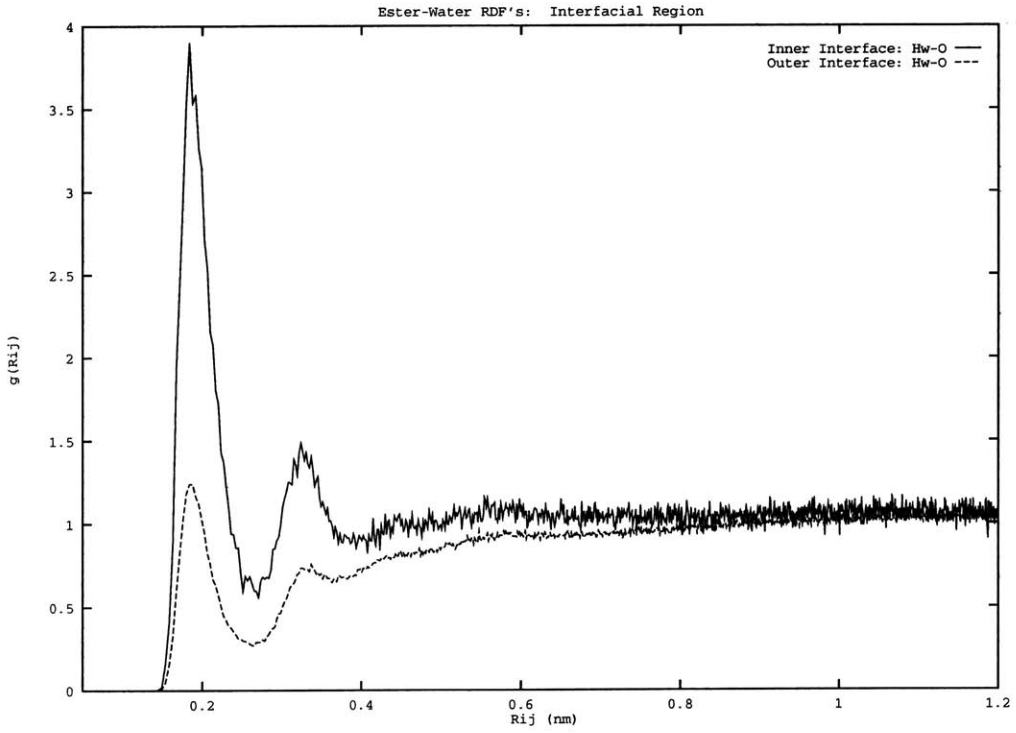


Figure 5-45: Hydrogen - carbonyl oxygen radial distribution function in the inner and outer liquid-liquid interfacial regions.

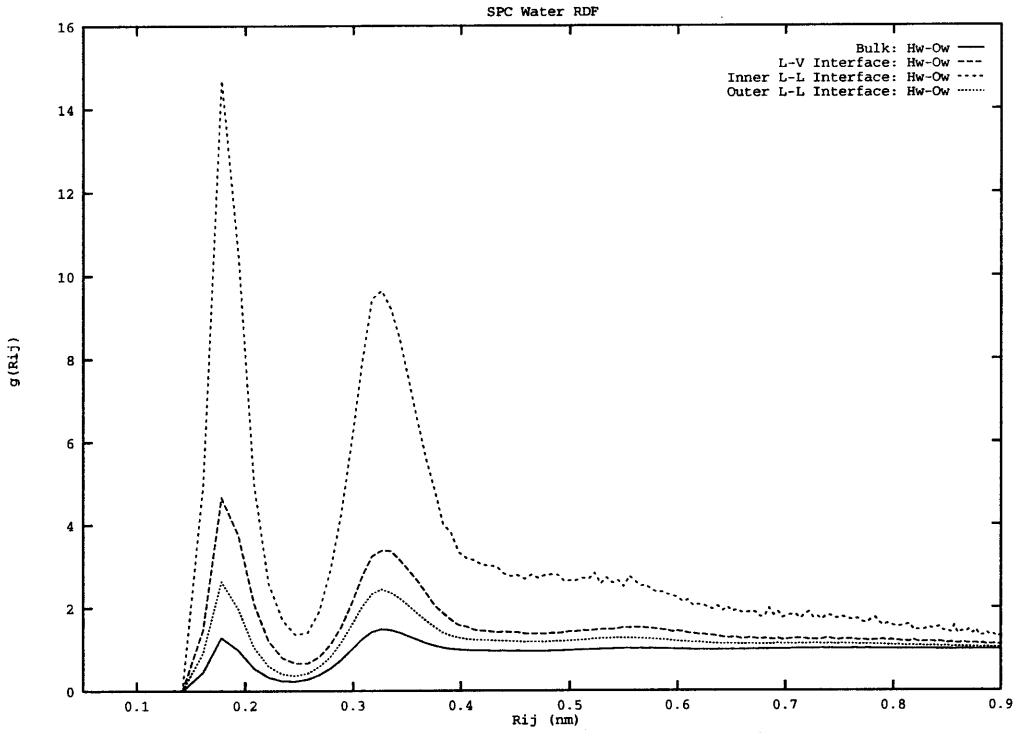


Figure 5-46: Hydrogen - water oxygen radial distribution functions.

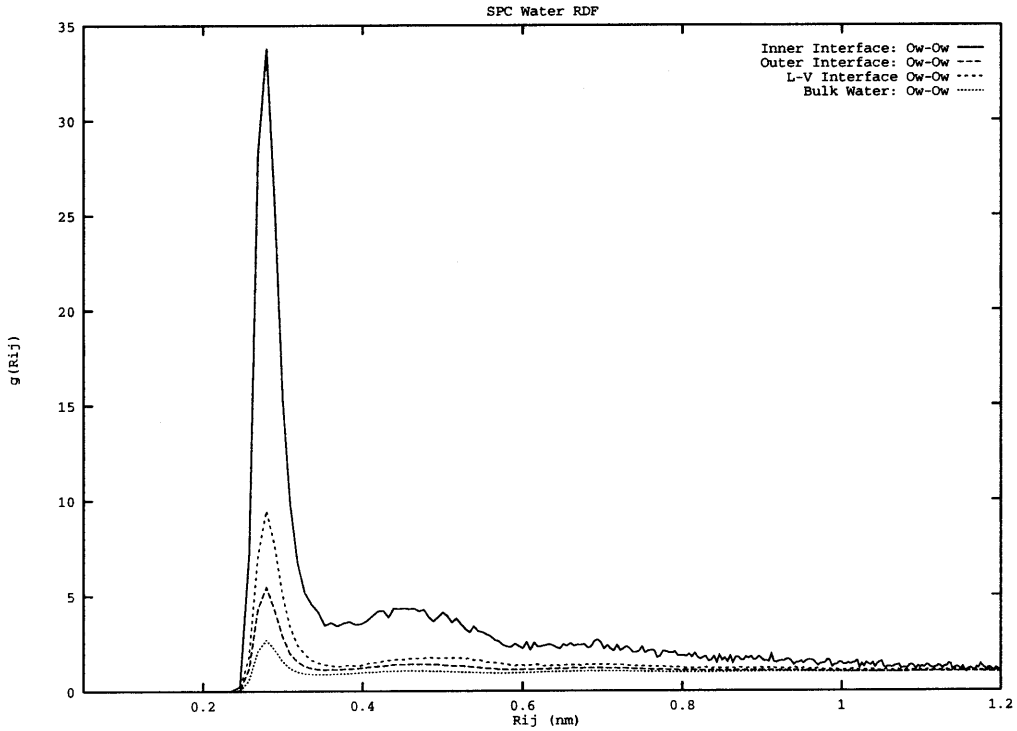


Figure 5-47: Water radial distribution functions: oxygen - oxygen

5.4 Excess Chemical Potential Via Widom Insertion Molecular Dynamics

As one measure of the equilibration of the system, the excess chemical potential of water is computed as a function of longitudinal position. This calculation gives an energetic description of the nature of the fluid along the system geometry. In the following sections, the theory, implementation, testing, and results of the Widom insertion calculations are presented.

5.4.1 Widom Insertion Formalism

Fundamental thermodynamics tells us that the chemical potential of a species is constant through a system (single or multiphase) at equilibrium [144]. Furthermore, it is the total chemical potential that is constant; its contributing elements, the ideal and excess chemical potentials can vary with system inhomogeneity (density fluctuation, compositional fluctuation, and external potential interaction) [145]:

$$\mu = \mu_{id} + \mu_{excess}(z) \tag{5.4}$$

By applying Widom's insertion theory (also called potential distribution theory), one can compute the excess chemical potential as a function of spatial position. Note that the effect of system heterogeneities on the chemical potential arises due to the interaction potential between constituents, and it is this dependence which is exploited in the computation of excess (or residual) chemical potential. To see this, formally, one begins with the configurational partition function for a fluid of N particles in a volume V at temperature T [146]:

$$Q_N = \int_V \cdots \int_V e^{-\frac{\Phi_N}{kT}} d\tau_1 \cdots d\tau_N \tag{5.5}$$

where k is Boltzmann's constant, $d\tau_1 \cdots d\tau_N$ is a volume element of the configuration space of the N particles, Φ_N is the total potential energy of interaction of the N particles (a function of particle positions). This expression can be rewritten as:

$$Q_N = \int_V \cdots \int_V e^{-\frac{\psi}{kt}} e^{-\frac{\Phi_{N-1}}{kt}} d\tau_1 \cdots d\tau_N \tag{5.6}$$

$$= Q_{N-1} V \left\langle e^{-\frac{\psi}{kT}} \right\rangle \tag{5.7}$$

where ψ is the interaction of a single particle with the remaining $N-1$ particles. The brackets denote a canonical ensemble average over the N-particle system [146]. An elucidative interpretation of the average is as follows. Consider a fluid in thermodynamic equilibrium in which the motion of all particles suddenly stops. Now, an $(N+1)'th$ particle is introduced and allowed to sample the entire volume of the N particle system. At each point of this 'wandering', the interaction of the particle with the rest of the system as well as the corresponding Boltzmann factor is computed. The average value of this function of position is thus given by the bracket notation [147].

Now, the connection to the excess chemical potential is made by noting that the chemical potential for the canonical ensemble is:

$$\mu_{excess} = \left(\frac{\partial A}{\partial N}\right)_{V,T} = \lim_{N\to\infty} (A(N,V,T) - A(N-1,V,T)) \tag{5.8}$$

$$\mu_{excess} = -kT \lim_{N\to\infty} ln \left(\frac{Q_N}{Q_{N-1}}\right) \tag{5.9}$$

The limit of taking large N is simply the thermodynamic limit.

Combining Equations 5.7 and 5.9, an expression for the excess chemical potential results:

$$\mu_{excess} = -kT\, ln \left(\left\langle e^{\frac{-\psi}{kT}} \right\rangle\right) \tag{5.10}$$

For an inhomogeneous system, the spatial dependence is given by:

$$\mu_{excess}(\mathbf{r}) = -kT\, ln \left(\left\langle e^{\frac{-\psi(\mathbf{r})}{kT}} \right\rangle\right) \tag{5.11}$$

and for a z-dependent heterogeneity, the **r** is replaced with a z-dependence [146].

5.4.2 Widom Insertion Molecular Dynamics: Implementation and Testing

The approach to calculating the excess chemical potential as a function of position within an inhomogeneous system involves insertions of the relevant particle/molecule within a selected slice of the slab; the insertion locations may be random or can be predetermined as in the case of a grid of points at which insertions are to be made (the grid obviously detailed enough to sufficiently represent the entire simulation volume). For the ester-water system studied in this chapter, we have selected to perform water insertions into both the water and ester phases. This is based on a practical consideration as water is the smaller molecule of the two components, thus not presenting any problems with respect to convergence of the insertions into the less dense ester; it is felt that attempting to insert methyl acetate into the fairly dense water phase would require a longer convergence time (as a result of higher overlap probablilites in the water phase). For water insertion, the oxygen atom is placed at a random point and the molecular orientation is chosen randomly. For the analysis of the methyl acetate/water system, insertions were made into configurations saved every 100 timesteps during the simulation. For testing purposes, calculations on a shift-force Lennard-Jones system (following the work of Powles *et al* [146]) were performed on configurations from short simulations of the LJ-SF3 fluid. Note that the LJ-SF3 fluid is simply a system of Lennard-Jones spheres, with the interparticle interactions shifted such that the interaction energy and resulting particle-particle force go to zero at the spherical cutoff distance. The cutoff is taken to be three times the Lennard-Jones diameter, σ, from which the numerical value of 3 derives in the name.

The shifted force Lennard-Jones potential is given as:

$$U_{LJ}^{SF} = U_{LJ}(r_{ij}) - U_{LJ}(r_{ij} = r_c) - \left(\frac{\partial U_{LJ}}{\partial r_{ij}}\right)_{r_{ij}=r_c} (r_{ij} - r_c) \qquad r_{ij} < r_c \qquad (5.12)$$

and is zero for separations outside of the cutoff. The cutoff is 3σ, where σ is the Lennard-Jones diameter. The potential and force (negative gradient of the potential) are both continuous and zero at the cutoff distance.

To make sure that the present coding for the LJ-SF3 fluid was correct and that the proper state points were being considered, a pressure-density isotherm was reproduced from the Powles *et al* work [146]. Bulk LJ-SF3 fluid MD simulations were performed for a system of 256 particles in a cubic simulation cell of dimensions to give the required density. Canonical ensemble (NVT) simulations at a reduced temperature of 1.28 (153.3 Kelvin) were performed using a Nose-Andersen thermostat with thermal inertia parameter of 2.5 and a timestep of 10 femtoseconds; simulations of 50,000 timesteps were sufficient to equilibrate the systems and generate acceptable statistics for equilibrium property calculation. Figure 5-48 shows the $P(\rho)$ isotherm (all in Lennard-Jones reduced units) at a reduced temperature of 1.28. Note that the reduced units for the Lennard-Jones system are defined as

$$\rho^* = \rho\sigma^3 \qquad (5.13)$$

$$T^* = \frac{k_B T}{\epsilon} \qquad (5.14)$$

$$P^* = \frac{P\sigma^3}{\epsilon} \qquad (5.15)$$

Also shown is the curve for the pure Lennard-Jones fluid with the standard analytic corrections for pressure [13, 146]. The points represent the work of Powles *et al.* Furthermore, the system energy per particle as a function of reduced density is shown in Figure 5-49 and is seen to be an almost linear function of density over a wide range of densities; this result is equivalent to those presented by Powles [146].

To evaluate the insertion algorithm, first the two distribution functions, f(u) and g(u), for the fictitious inserted particles and the real particles, are computed for a reduced density of 0.05 and reduced temperature of 1.02 as shown in Figure 5-50 .

These distribution functions are integral components within the formalism of potential distribution theory and a slight digression is presented here. The fictitious and real particle energy distribution functions, $f(u_f)$ and $g(u_r)$, respectively, are defined as

$$f(u_f) = \frac{\int \delta(u_f - u(N:N-1))e^{-\beta U_{N-1}} dq^N}{\int e^{-\beta U_{N-1}} dq^N} \tag{5.16}$$

and

$$g(u_r) = \frac{\int \delta(u_r - u(N;N-1))e^{-\beta U_N} dq^N}{Q_N} \tag{5.17}$$

The energies u_f and u_r represent the fictitious and real particle energies. For the fictitious particle, i.e. the particle which is allowed to explore the phase space of the (N-1)-particle system, the energy is the resultant of all the interactions of the fictitious particle with the rest of the system. For the real particle, the energy gives the interaction of the N'th particle with the N-1 particles of the N-particle system. For both cases, the energy is the negative of the energy required to extract the N'th particle leaving the (N-1) fixed; this energy is equal to twice the 'shared energy' arising from all pairwise interactions.

Now, the relevance of the distributions becomes evident if one follows a route to the chemical potential that is different than the formulation presented in the previous section. Consider the ratio of the partition functions of two systems with N particles each, one which is called system 0 and the other system 1,

$$\frac{Q_0}{Q_1} = \frac{\int e^{-\beta U_0} dq^N}{\int e^{-\beta U_1} dq^N} \tag{5.18}$$

Equation 5.18 can be recast in a form resulting in an average over each system as

$$\frac{Q_0}{Q_1} = \frac{\int e^{-\beta(U_0-U_1}e^{-\beta U_1}dq^N}{\int e^{-\beta U_1}dq^N} = \left\langle e^{-\beta(U_0-U_1)}\right\rangle_1 \tag{5.19}$$

or

$$\frac{Q_0}{Q_1} = \frac{\int e^{-\beta u_0}dq^N}{\int e^{-\beta(U_1-U_0)}e^{-\beta U_0}dq^N} = \frac{1}{\langle e^{+\beta(U_0-U_1)}\rangle_0} \tag{5.20}$$

Introducing a normalized distribution function for the coordinate-dependent quantity $(U_0 - U_1)$ for each system gives,

$$f_1(U_0 - U_1) = \frac{\int \delta[(U_0 - U_1) - (U_0(q^N) - U_1(q^N))]e^{-\beta U_1(q^N}dq^N}{Q_1} \tag{5.21}$$

$$f_0(U_0 - U_1) = \frac{\int \delta[(U_0 - U_1) - (U_0(q^N) - U_1(q^N))]e^{-\beta U_0(q^N}dq^N}{Q_0} \tag{5.22}$$

The δ 's represent the Dirac delta function. Now, writing $u = (U_0 - U_1)$, the ratio of the two distributions, f_1 and f_0 gives the important general relation,

$$\frac{f_0(u)}{f_1(u)} = e^{-\beta u}\frac{Q_1}{Q_0} \tag{5.23}$$

For the case where one is considering two systems, one of N particles and the other of N-1 particles, the definitions and distributions can be more specifically delineated. Consider,

$$U_0 = U_N(q^N) \tag{5.24}$$

$$U_1 = U_{N-1}(q^{N-1}) \tag{5.25}$$

for which case

$$U_0 - U_1 = U_N - U_{N-1} = u(N; N-1) \tag{5.26}$$

where $u(N; N-1)$ is the energy of the N'th particle in the N-particle system.

The ratio of the partition functions of the two systems then becomes,

$$\frac{Q_0}{Q_1} = \frac{Q_N}{Q_{N-1}} = \frac{\int e^{-\beta U_{N-1}}[V^{-1}e^{-\beta u}dq_N]dq^{N-1}}{\int e^{-\beta U_{N-1}}dq^{N-1}} \tag{5.27}$$

Note that this is simply the canonical average of the factor $e^{-\beta u}$ over the system of $N-1$ particles,

$$\frac{Q_0}{Q_1} = \frac{Q_N}{Q_{N-1}} = \left\langle e^{-\beta u} \right\rangle_{N-1} \tag{5.28}$$

Furthermore, in the thermodynamic limit ($N \to \infty$) , we have

$$\frac{Q_0}{Q_1} = \frac{Q_N}{Q_{N-1}} = e^{-\beta \mu^r} \tag{5.29}$$

The distribution functions for the N and N-1 particle systems are

$$f_N(u_r) = \frac{\int \delta[u_r - u(N; N-1)]e^{-\beta U_N}dq^N}{Q_N} \tag{5.30}$$

$$f_{N-1}(u_f) = \frac{\int \delta[u_f - u(N; N-1)]e^{-\beta u_{N-1}}dq^{N-1}}{Q_{N-1}} \tag{5.31}$$

At this point, we call the f_0 distribution the real particle energy distribution, g(u), and the f_1 distribution the fictitious particle energy distribution, f(u).

From Equation 5.23, we obtain the relation,

$$\frac{g(u)}{f(u)} = \frac{Q_{N-1}}{Q_N}e^{-\beta u} \tag{5.32}$$

Substituting Equation 5.29 into the last equation gives the result

$$e^{-\beta \mu^r} g(u) = e^{-\beta u} f(u) \tag{5.33}$$

Thus, we can see that the excess/residual chemcial potential, μ^r, may be determined from a knowledge of $f(u)$ and $g(u)$, and furthermore, for a value of u where the two distributions overlap significantly [146].

Rearranging Equation 5.33 gives,

$$e^{-\beta\mu^r} = e^{-\beta u}\frac{f(u)}{g(u)} \tag{5.34}$$

A convenient form to extract the residual chemical potential is

$$L(u) \equiv ln\left(\frac{f(u)}{g(u)}\right) = \beta u - \beta\mu^r \tag{5.35}$$

which shows that a plot of $L(u)$ versus u gives the slope as β and the intercept as $-\beta\mu^r$.

In the present study, the excess chemical potential is determined from the Widom result, namely, from a canonical average of the Boltzmann factor for a water insertion, but the above digression points out the significance and relevance of the distributions that will be computed in order to test the algorithm. Note, however, that it is reported in the literature that the distribution function approach suffers more from statistical uncertainty due to the inherently substantitial statistical uncertainties in $f(u)$ for u large and negative and in $g(u)$ for u large and positive. The former arises from the physical picture that a fictitious particle randomly inserted into a bulk fluid will have difficulty finding a hole surrounded by many particles at a sufficient separation to give rise to a large, negative (favorable) overall energetic interaction. The latter, quite analogously, is indicative of the fact that in a finite system, a real particle will not explore that portion of phase space in which there is a resultant large, positive interaction energy (as in the case of overlaps). Thus, from a practical standpoint, the Widom result is invariably applied in some modified form or other.

The two distribution functions, f(u) and g(u), for the fictitious inserted particles and the real particles, are computed for a reduced density of 0.05 and reduced temperature of 1.02 as shown in Figure 5-50.

Note that the computed energy distribution functions are in excellent agreement

with those of Powles *et al* given in Figure 5-50.

For the low particle density case, $\rho_{reduced} = 0.05$, $f(u)$ tends to peak at $u = 0$, indicating test particles inserted outside the range of interaction (beyond the specified cutoff). For the real particle energy distribution, $g(u_r)$, there is a similar peak at $u = 0$, but this is accompanied by a secondary peak at $u = -1$ corresponding to an energetically favorable situation of a pair of particles separated at the appropriate separation.

Finally, Figure 5-51 shows the excess chemical potential, in the form of μ_1. The modified chemical potential, μ_1 is:

$$\mu_1 = \mu_{excess} + TlnT - \frac{3}{2}TlnT \tag{5.36}$$

In the figures, the points represent the results of Powles et al. The agreement with the calculations of Powles is not exact, but this is attributed to the fact that the previous work uses a temperature correction for μ_1 due to the fact that there is a temperature fluctuation even for constant temperature molecular dynamics. In the present case, the correction is neglected, and it is evident that the error generated is not great. With these results, there is confidence in the current implementation of the insertion code. In the next section the results of the insertions in the ester-water system are shown.

5.4.3 Widom Insertion: Results

Longitudinal profiles of the residual chemical potential of water were computed over blocks of 100 configurations. For each configuration, the simulation cell was divided into 72 layers into which the water molecule was inserted; each layer was of thickness 2.5 angstroms with the transverse dimensions being equal to the box lengths in these directions. For each layer, 100 random water insertions were performed. The position of the water oxygen was selected randomly, and the locations of the hydrogens were determined so as to achieve random orientations of the water dipole relative to the simulation cell-centered frame of reference.

Figures 5-52 through 5-54 show the results of the insertion calculations. Figures 5-52 and 5-53 show the time evolution of the excess chemical potential profiles over the system length. Figure 5-52 shows the initial time intervals over which the system is equilibrating. Comparing the two figures, the system appears to have equilibrated by about 250 picoseconds, and certainly the profiles after 300 picoseconds show no changing trend. Finally, Figure 5-54 shows the average insertion energy of a single water molecule which is seen to be symmetric across the system profile. Furthermore, the average insertion energy into water is roughly one third that for insertion into the ester which translates to a low solubility of water in ester (the experimental solubility is 0.266 mole fraction water in methyl acetate corresponding to 8.1 $\pm$ 0.1 grams of water per 100 grams of solution at 298K and atmospheric pressure, Solubility Data Series, volume 48, pp 8.). For the present work, a quantitative analysis of the solubility is not made as this is used mainly as one check on equilibration. The proper trend in the excess chemical potential is noted as well as the invariance with time for the duration of the sampling period.

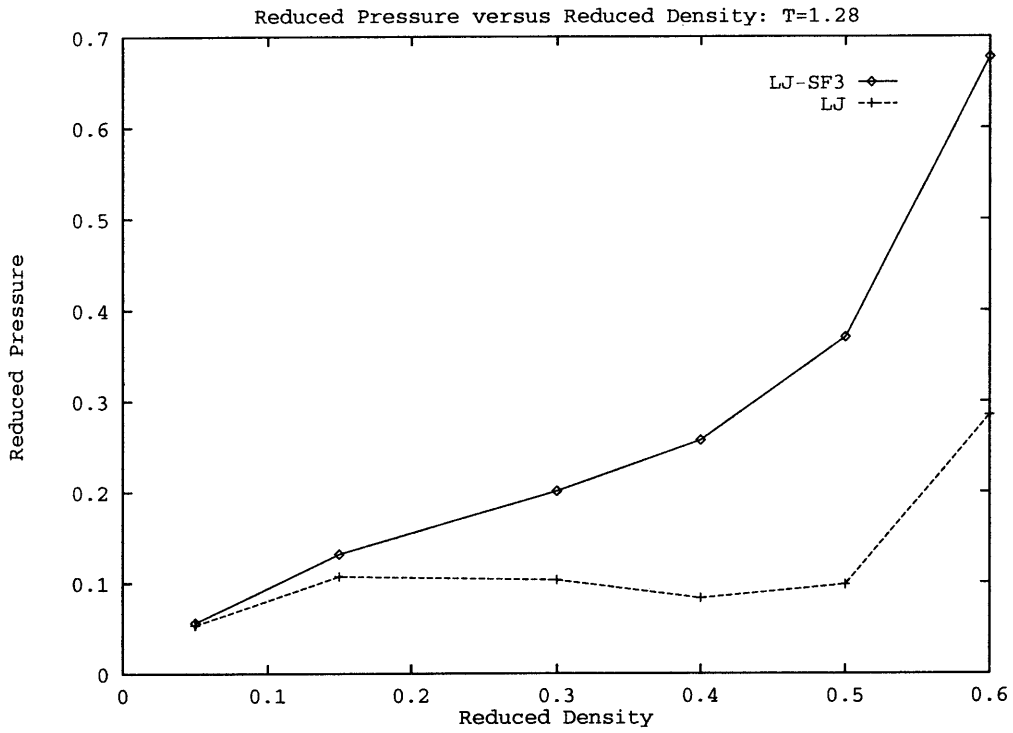


Figure 5-48: Pressure-Density isotherm for the Lennard-Jones SF3 fluid at a reduced temperature of 1.28

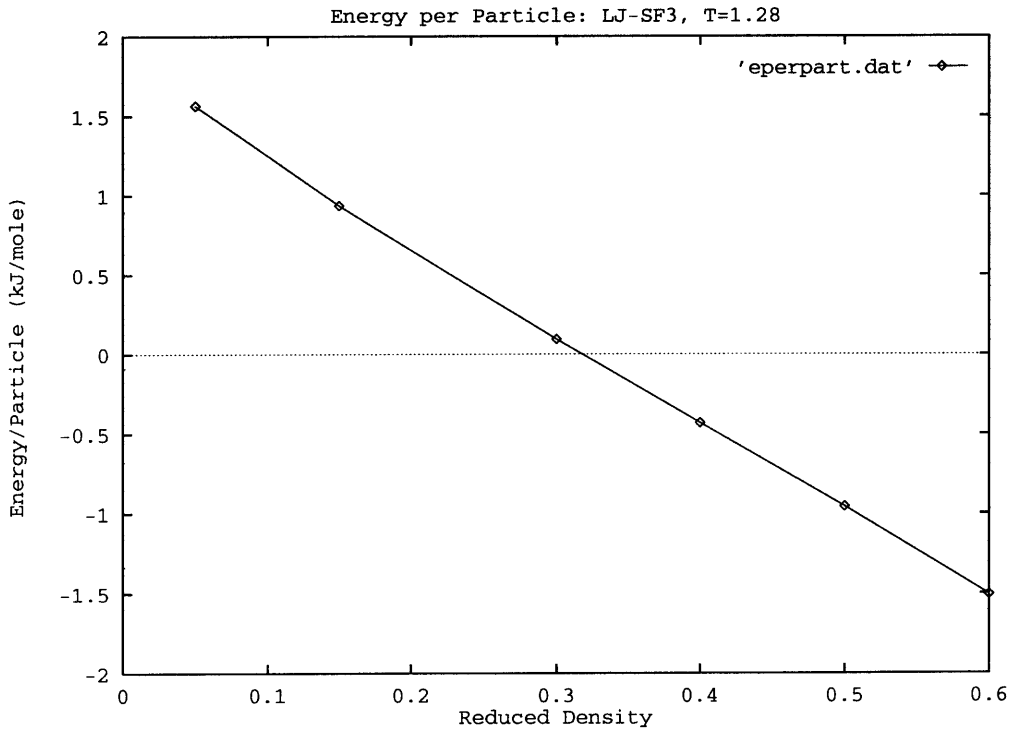


Figure 5-49: Energy per particle as a function of reduced density for the Lennard-Jones SF3 fluid

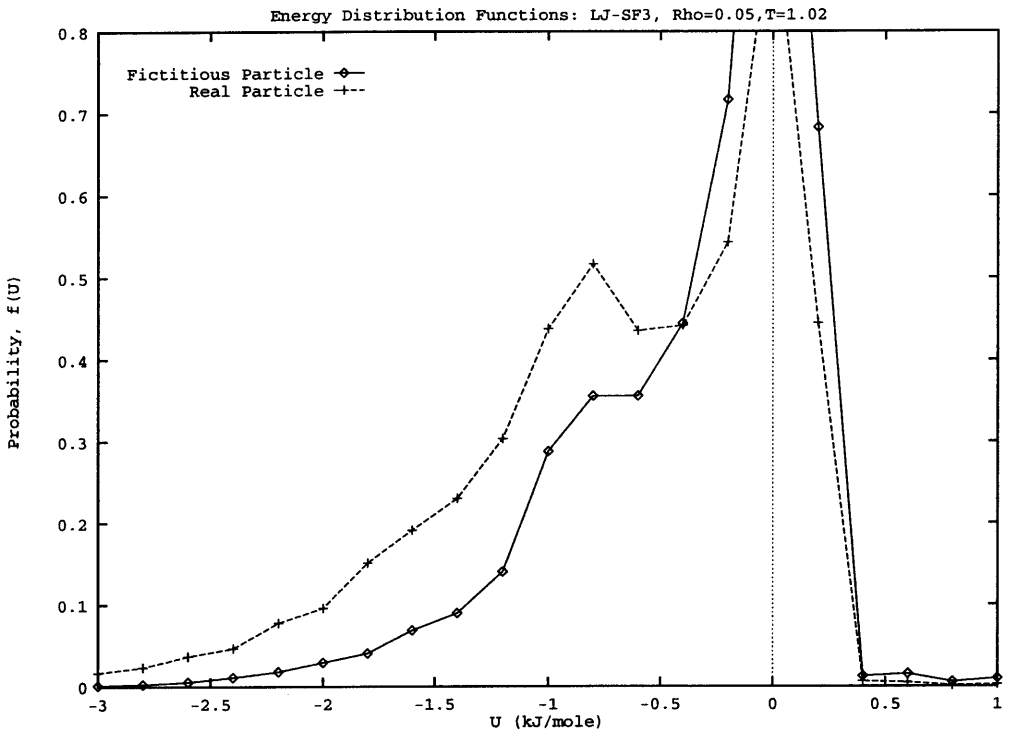


Figure 5-50: Real and fictitious particle energy distribution functions for the Lennard-Jones SF3 fluid at $\rho^* = 0.05$ and $T^* = 1.02$.

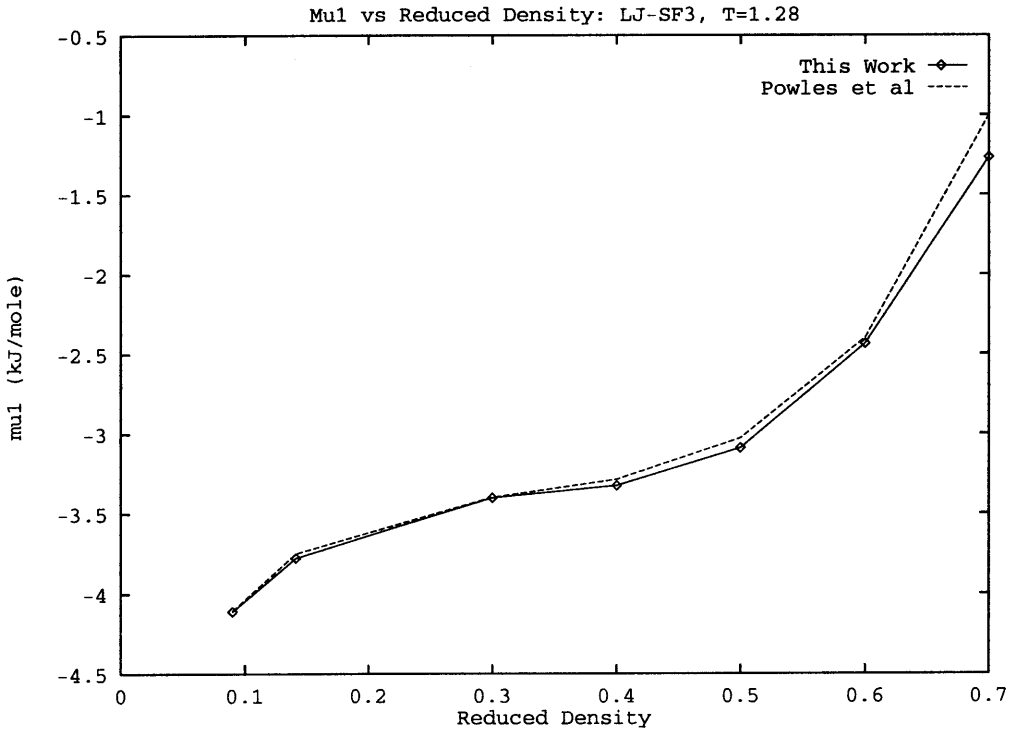


Figure 5-51: Modified excess chemical potential, μ_1 versus reduced density for the Lennard-Jones SF3 fluid

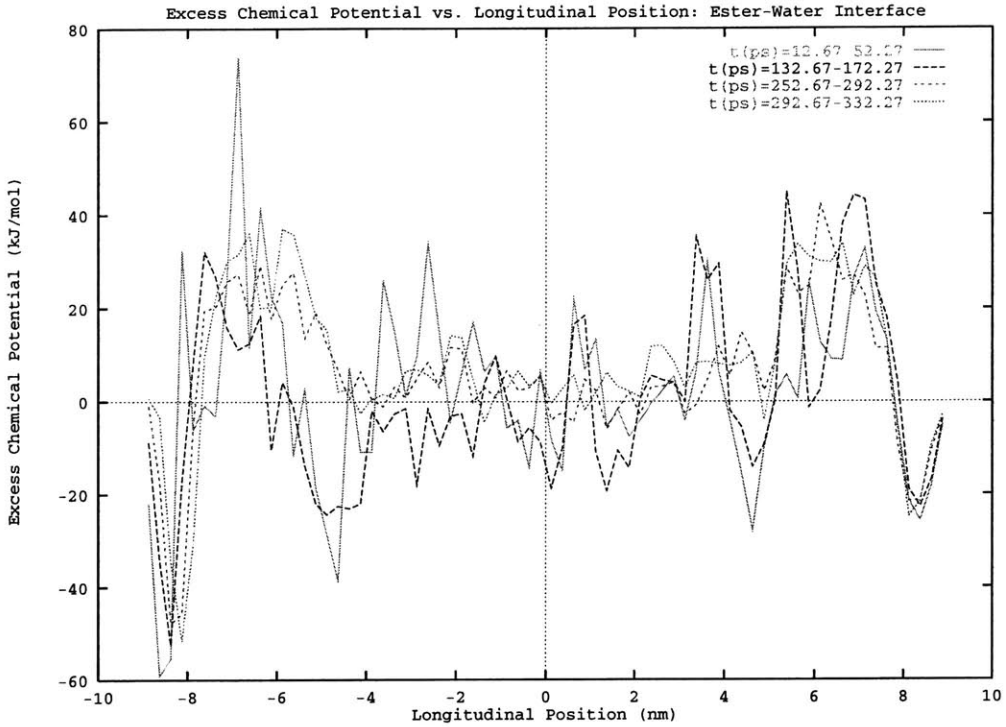


Figure 5-52: Excess chemical potential for the ester-water system: profiles are averages over blocks of 100 configurations spanning the shown time periods of the simulation.

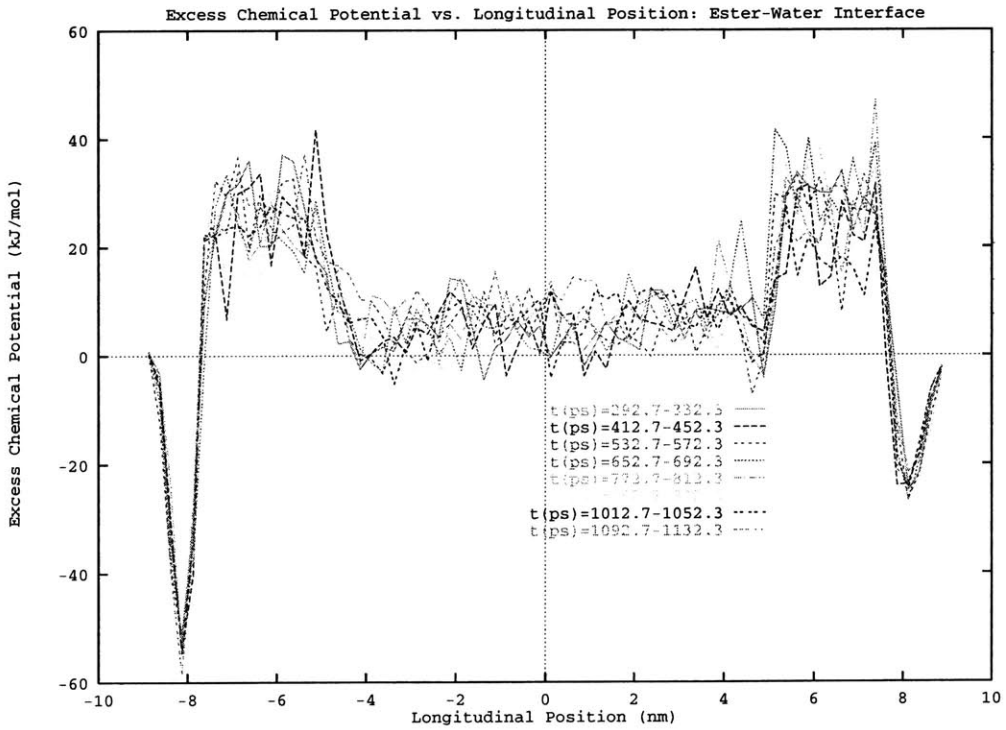


Figure 5-53: Same as previous figure but applies for later time periods

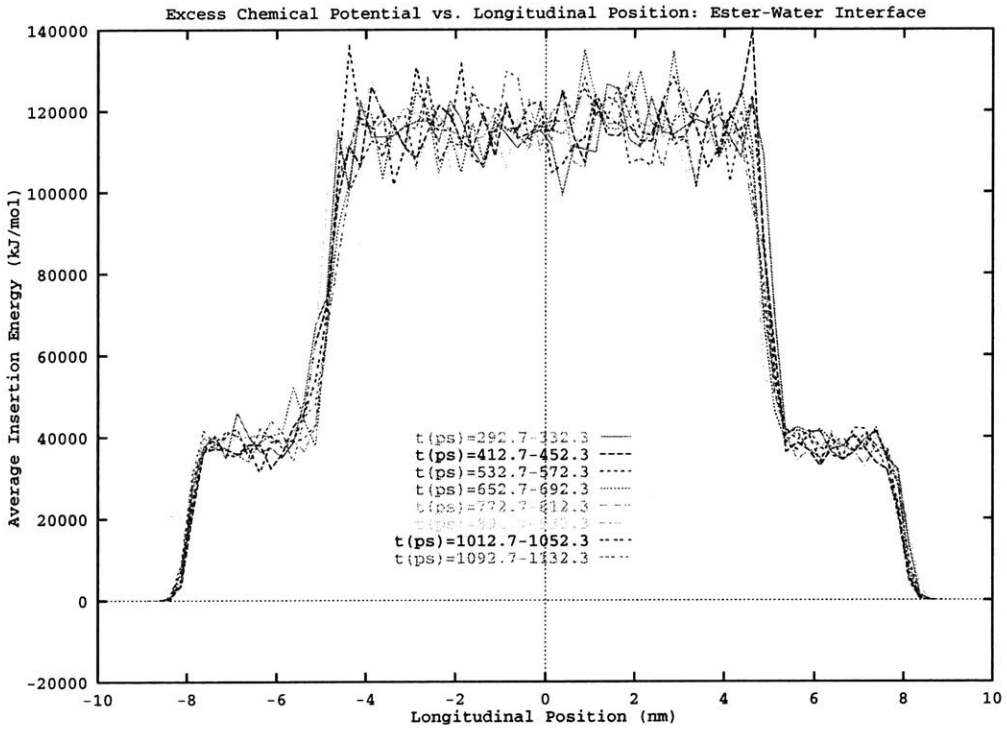


Figure 5-54: Profiles of the average insertion energy of a water molecule for various time periods of the simulation.

5.5 Pressure Tensor Component Profiles

In order to evaluate the consistency of the OPLS ester potential parameters for interfacial thermodynamics, the molecular virial is computed to obtain the molecular surface tension via [60, 126, 132]:

$$\gamma = \frac{1}{2A}(2V_{zz} - V_{xx} - V_{yy}) \tag{5.37}$$

where $V_{\alpha\alpha}$ is the diagonal molecular virial component and $A = 2L_xL_y$ is the total surface area of the slab. The molecular virial is:

$$V_{\alpha\alpha} = \sum_{a<b} F_{ab}^{\alpha} R_{ab}^{\alpha} = \sum_{a} F_{a}^{\alpha} R_{a}^{\alpha} \tag{5.38}$$

where the sum is over all molecules a and b and R_{ab}^{α} and F_{ab}^{α} are the $\alpha = x, y, or z$ component of the molecular center of mass separation and total force,

$$F_{ab}^{\alpha} = -\sum_{k,j=1}^{m} \frac{r_{abkl}^{\alpha}}{|r_{abkl}|} \phi^{'}(r_{abkl}) \tag{5.39}$$

The profile of the normal-tangential pressure differential is computed as:

$$\gamma(z) = \frac{1}{2}\int_{-\inf}^{z} dz^{'}(P_N(z) - P_T(z)) \tag{5.40}$$

where $P_N(z)$ and $P_T(z)$ are the normal and tangential components of the stress across a plane parallel to the interface at the longitudinal position z. γ(inf) is the surface tension.

The pressure tensor components following the Irving-Kirkwood definition of the pressure tensor are [132, 60]

$$P_T(z) = <\rho(z)> kT - \frac{1}{A} < \sum_{j>i}^{N} \sum_{a}^{m} \sum_{b}^{m} \frac{(x_{ij}x_{iajb} + y_{ij}y_{iajb})}{2r_{iajb}} \frac{dU(r_{iajb})}{dr_{iajb}} \frac{1}{|z_{ij}|} \theta(\frac{z - z_i}{z_{ij}}) \theta(\frac{z_j - z}{z_{ij}}) > \tag{5.41}$$

and

$$P_N(z) = < \rho(z) > kT - \frac{1}{A} < \sum_{j>i}^{N} \sum_{a}^{m} \sum_{b}^{m} \frac{z_{ij} z_{iajb}}{r_{iajb}} \frac{dU(r_{iajb})}{dr_{iajb}} \frac{1}{|z_{ij}|} \theta(\frac{z - z_i}{z_{ij}}) \theta(\frac{z_j - z}{z_{ij}}) > \tag{5.42}$$

In these equations, $\rho(z)$ is the total molecular density at z (its average over all configurations is used), k is Boltzmann's constant, and T is the absolute temperature. $\theta(q)$ is the Heavyside step function which is zero for arguments less than zero and equal to unity for arguments greater than or equal to zero. For clarification, the x_{ij} , y_{ij} , and z_{ij} are the molecular center of mass separation components, whi le the x_{iajb} , y_{iajb} , and z_{iajb} are the site-site separation components. The sums are over all N molecule pairs and over the m sites on each molecule (as shown here, these equations reflect a system where, all molecules are taken to be equivalent, but the extension to a multi-component system is straightforward).

Figures 5-56 and 5-55 show the results of the Irving-Kirkwood calculations. Figure 5-55 shows the longitudinal profile of the differential of the normal and tangential components of the pressure tensor. The pressure compoents are given in program units of $kJ/mol\,cm^3$. There are two clearly defined interfacial regions for the water liquid-vapor region, and two interfacial regions for the ester-water inhomogeneous zones. The former spans the range from about $|7.5nm|$ to $|8.0nm|$. For the ester-water interface, the relevant region is centered about $|5nm|$. The profile is computed over blocks of 100 configurations spanning the last picoseconds of the production run. The rather large fluctuations in this profile is indicative of two things. First, the nature of the computed pressure is that of a strongly fluctuating property; this is an inherent drawback of the MD method. Secondly, the fluctuation is representative of the small sample used for the computation. Again, one is interested in the difference of the tangential stress at the interfacial regions. This requires accounting of the force interactions between molecules in the interfacial region. For a simulation of this type, since such a small fraction of the total number of molecules are resident in this region, one is relegated to accumulating exorbitant amounts of data. In the present case, not all of the data from the production run is computed since we are not interested in a

quantitative analysis of the surface tension. We accept *a priori* that the combination of the OPLS and SPC potentials will most likely not quantitatively reproduce experimental surface tension data. Nevertheless, this computation indicates that with the models selected, we are able to develop a stable interfacial system which most importantly captures the requisite physics–that is, we don't observe mixing of the two phases.

Figure 5-56 represents the integrated normal-tangential pressure differential curve. Although the curve is not completely flat in the bulk regions, we still obtain a nice indication that the system has equilibrated for the purposes of this investigation. Based on the limited sample size, one must accept the observed fluctuations. Note that the contributions from the two pairs of interfaces is self-consistent; that is, the contributions from the two water liquid-vapor interfaces are equivalent to one another as are those from the two liquid-liquid interfaces.

5.6 Conclusions

This Chapter has presented results of direct molecular dynamics simulations of the methyl acetate/water liquid-liquid interface to probe the molecular level structure and orientation within the interfacial region. One observes the evolution of a stable interface using the OPLS intermolecular potentials for the methyl acetate and the Simple Point Charge model for water. In terms of molecular structure, one observes the ester molecules adopting an average orientation in which the carbonyl oxygen atom extends towards the water bulk phase. Complementary to the ester orientation, the water molecules within the interfacial region adopt a configuration which allows the hydrogen atoms to participate in hydrogen bonding interactions with the carbonyl oxygen of the ester. Moreover, the water orientation varies across three regions of the interface. Again, in the outer region closest to the ester, the water dipole is aligned with the interface normal and the hydrogens extending into the ester phase. In the inner region of the interface (more in proximity to the bulk water), the water dipole aligns further away from the interface normal with one hydrogen atom extending into

the outer interfacial region where there is a layer of water oxygens available for hydrogen bonding. Radial distribution functions indicate a hydrogen bonding interaction with the first peak in the carbonyl oxygen - water hydrogen rdf at 0.19 nanometers. This corresponds to density functional calculations of gas phase bimolecular complexes of water and methyl acetate. Note however, that the hydrogen bonding angle is smaller than that computed via theory; this may be due to the fact that the present calculations are for the condensed phase.

The orientation of the water in the interfacial region relative to that in the bulk is interesting in the context that it is observed both theoretically (molecular dynamics simulations) and experimentally [25, 26] that the properties of liquid water in the vicinity of solid interfaces differ from those in the bulk phase. The vicinal structure of water is held to extend 3-6 nanometers from the interface [26], and as such, this long-ranged ordering phenomenon is critical in biological systems, for example. In the context of the present work, we observe a rich vicinal structure for SPC water. One can consider the interfacial structure of the ester as a similar phenomenon, as we see not only a variation in molecular orientation from bulk to liquid-liquid interface, but also a difference between the liquid-liquid and liquid-vapor interface; this equally applies to the water/vapor surface as well.

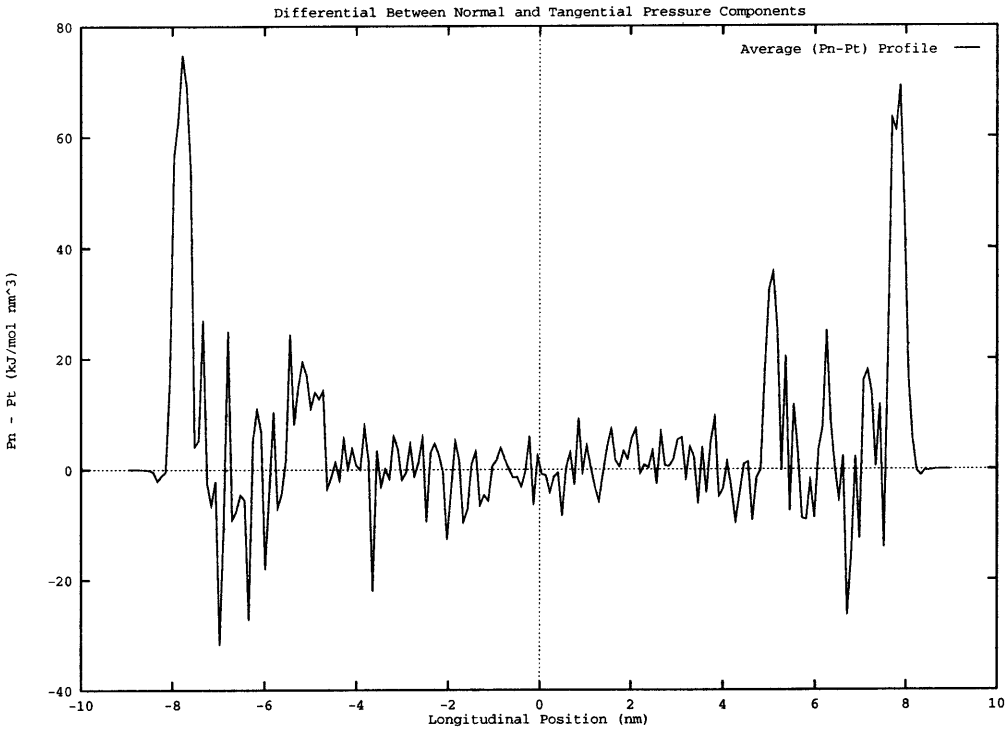


Figure 5-55: Profile of the differential between normal and tangential pressure tensor components

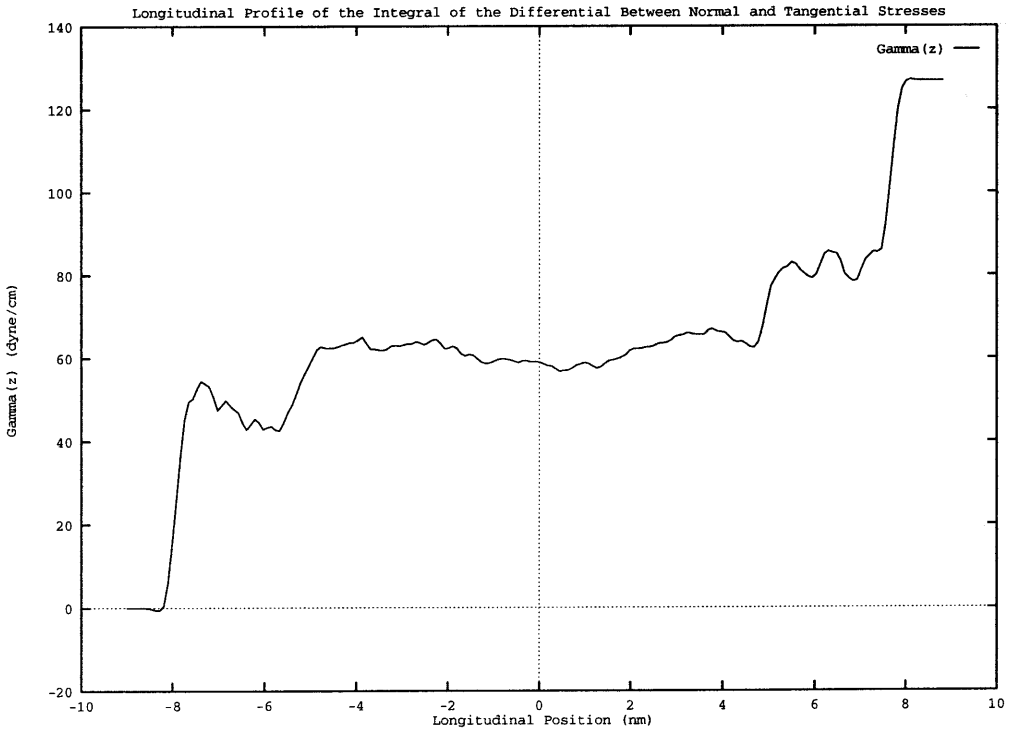


Figure 5-56: Profile of the integral of the differential between normal and tangential pressure tensor components

5.7 Statistical Significance of Observed Radial Distribution Function Profiles

Figures 5-57 through 5-66 show the various pair correlation functions discussed earlier along with the associated error bounds. The average profiles are calculated as block averages over blocks of 100 configurations (each configuration stored every 100 timesteps, with a timestep being 0.004 picoseconds). The average is at the 95 percent confidence level.

For methyl acetate, the bulk region is taken from -1.0 to 1.0 nanometers relative to the simulation cell center of mass. The 'inner interface' spans from 4.0 to 4.5 (and analogously from -4.0 to -4.5 for the z ¡ 0 regions) nanometers from the cell center of mass. The 'outer interface' extends from 4.5 to 5.0 nanometers (again the symmetry applies). Finally, the 'extreme outer' interface reaches from 4.8 to 5.3 nanometers. For the water layer, the bulk is taken to be from 6.0 to 7.0 nanometers, and the liquid-vapor interface is lumped into one region from 8.0 to 8.5 nanometers.

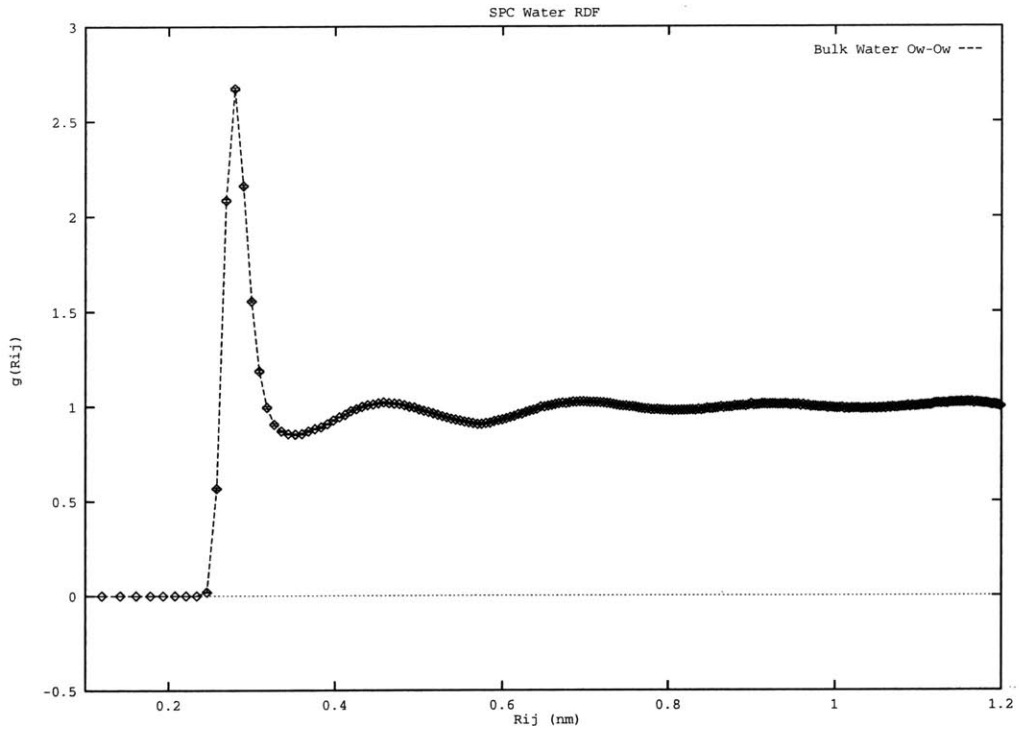


Figure 5-57: Bulk Water oxygen-oxygen radial distribution function

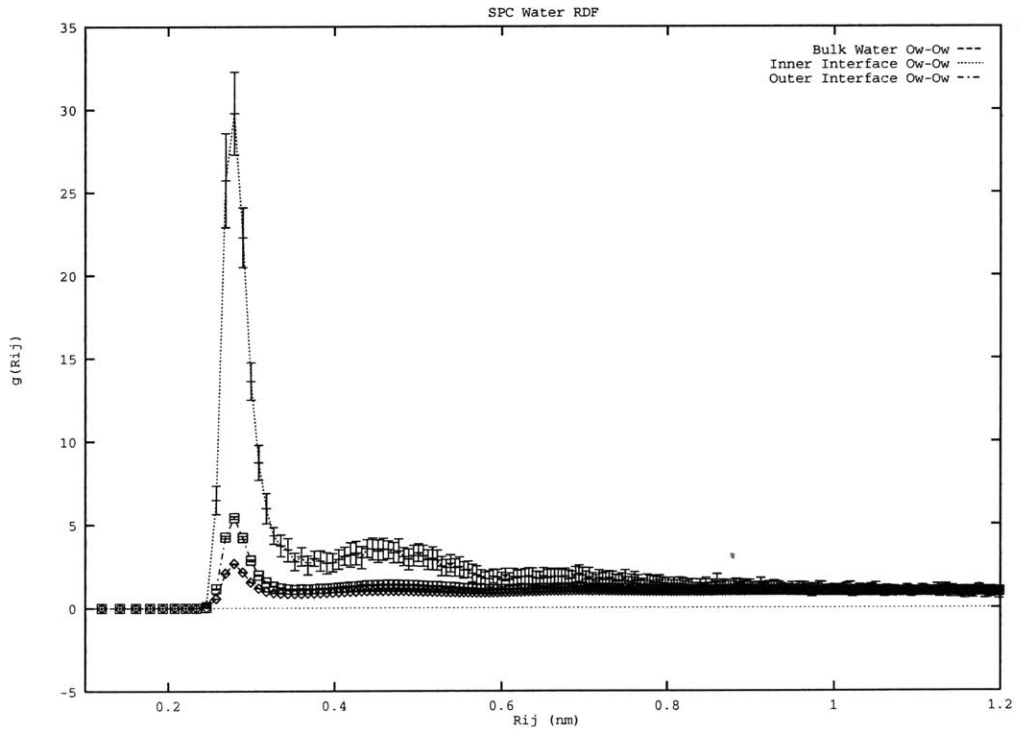


Figure 5-58: Water oxygen-oxygen radial distribution functions

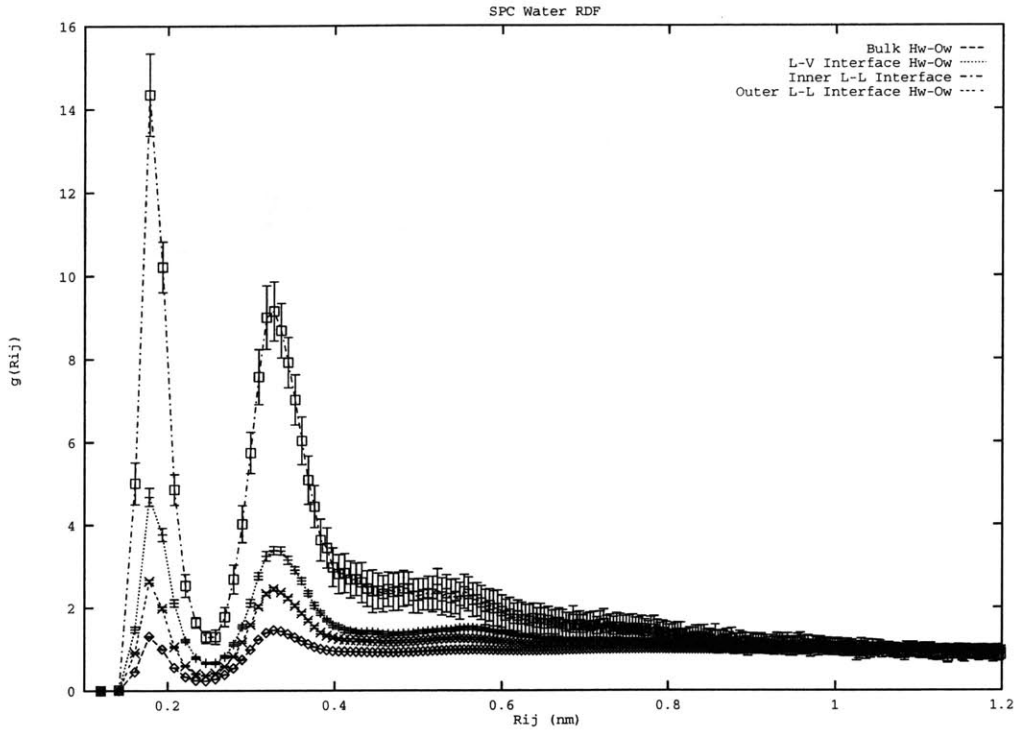


Figure 5-59: Water oxygen-hydrogen radial distribution functions

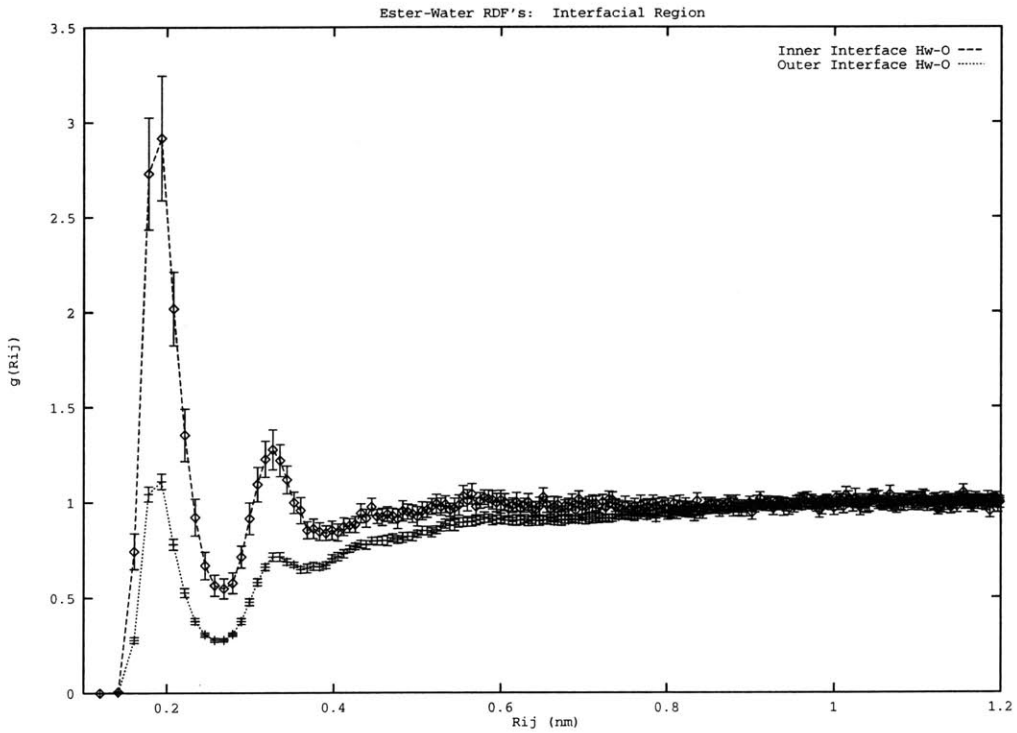


Figure 5-60: Water hydrogen - ester carbonyl oxygen radial distribution functions

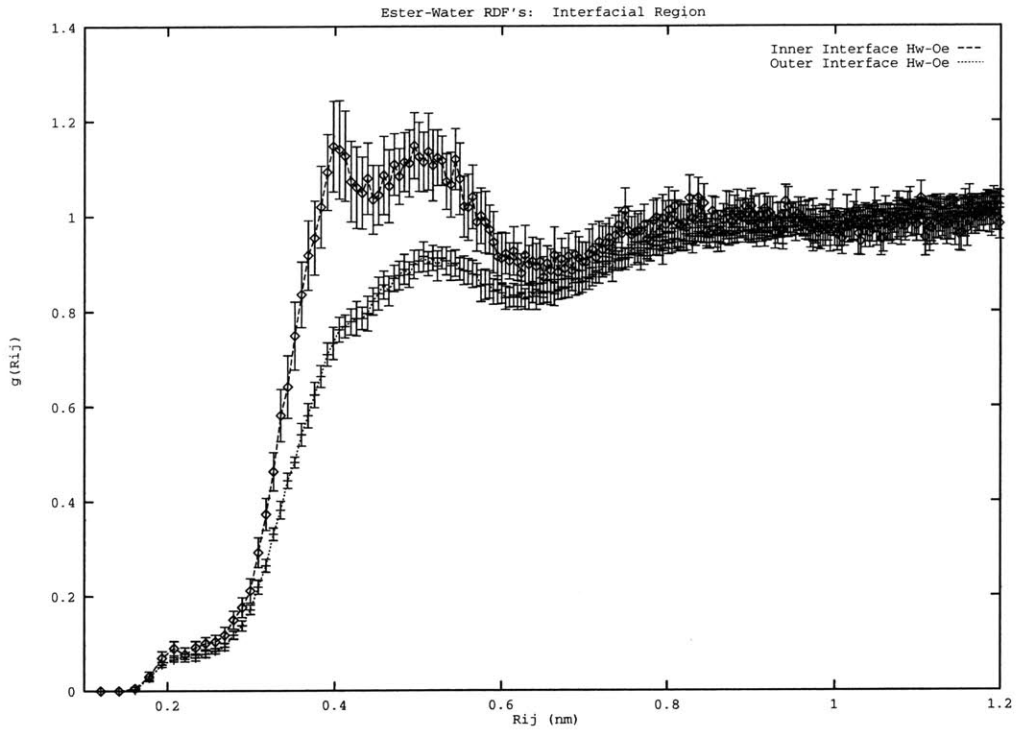


Figure 5-61: Water hydrogen - ester ether oxygen radial distribution functions

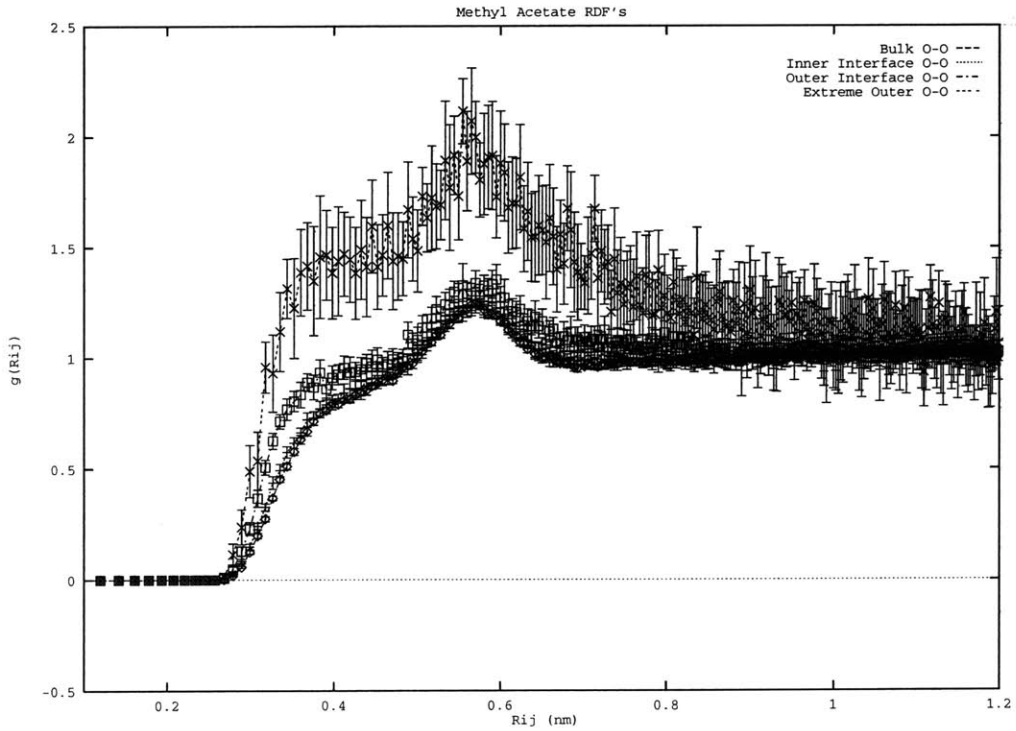


Figure 5-62: Methyl acetate radial distribution functions: carbonyl oxygen - carbonyl oxygen

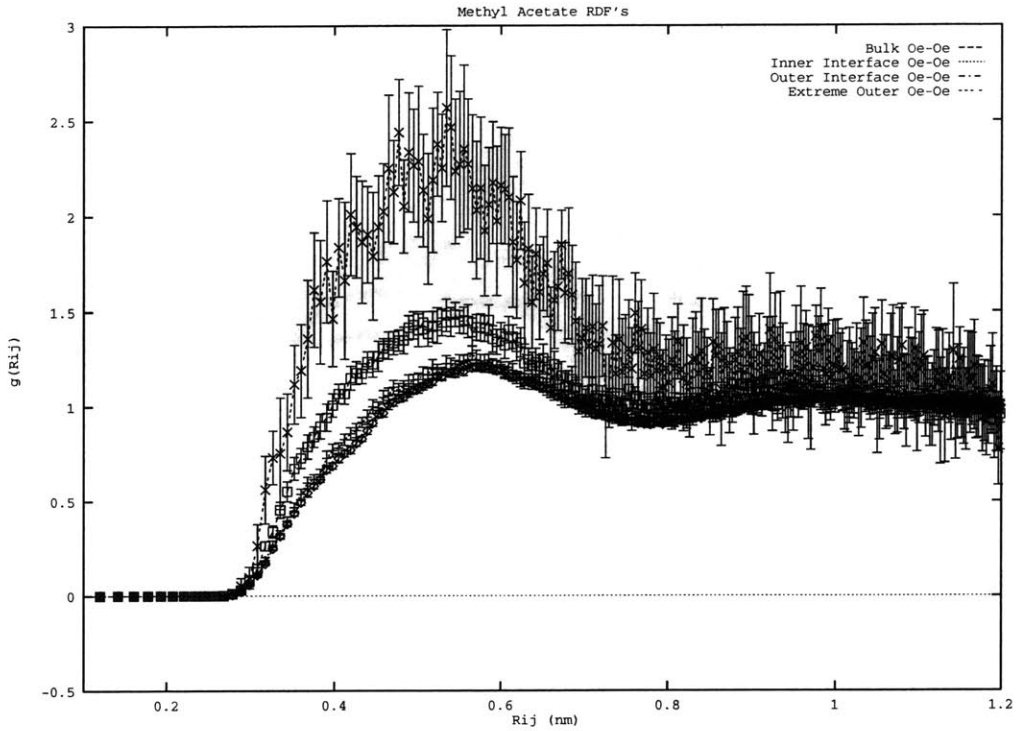


Figure 5-63: Methyl Acetate radial distribution functions: ether oxygen - ether oxygen

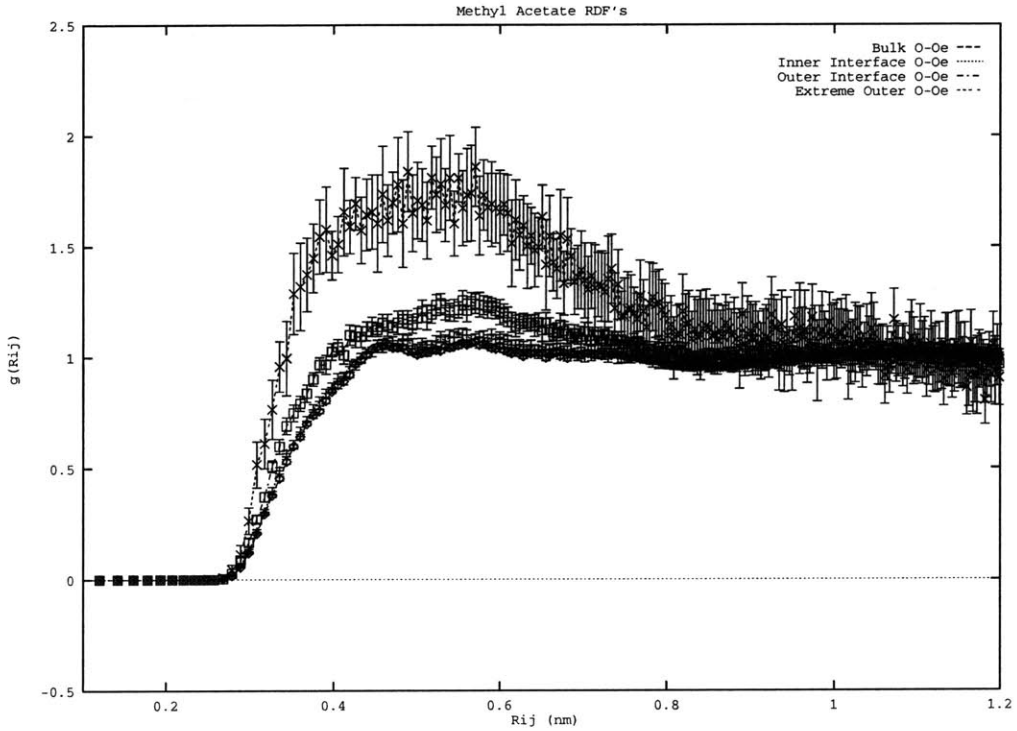


Figure 5-64: Methyl Acetate radial distribution functions: ether oxygen - carbonyl oxygen

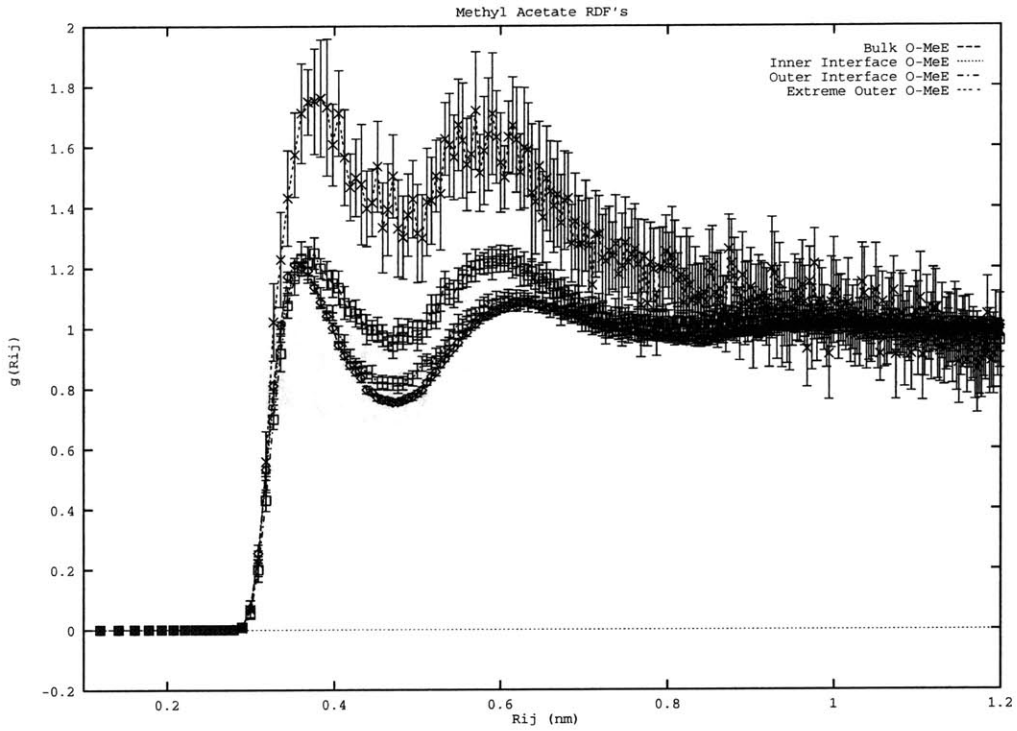


Figure 5-65: Methyl Acetate radial distribution functions: carbonyl oxygen - alkoxy methyl

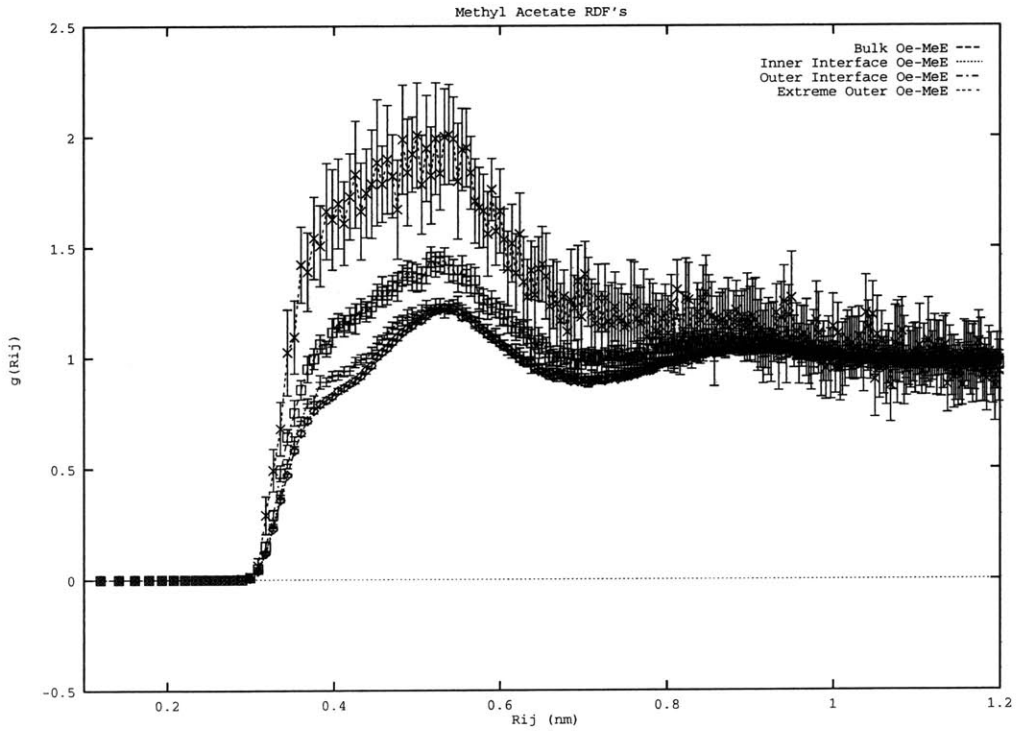


Figure 5-66: Methyl Acetate radial distribution functions: ether oxygen - alkoxy methyl

Chapter 6

Parameterization of United-Atom Oligomer Model

6.1 Background and Motivation

The previous three chapters have dealt with small molecule ester interfaces, both the liquid-vapor and liquid-liquid; based on these computations, the interfacial structure in terms of the orientation of functional groups and hydrogen bonding characteristics is seen to be rich and dependent on the nature of the interface. The fact that such effects are seen for small molecules is quite interesting, and more importantly, bodes well for the appearance of a more complex structure for longer chain molecules where the effects of functional group character and chain connectivity can interact. A rich interfacial structure is observed from simulations of alkane oligomers [60] and similar phenomena are predicted by Monte Carlo simulations of lattice polymer models [148]. Furthermore, theoretical studies of compressible diblock and triblock chain molecules in the vane of self-consistent models also show a very rich interfacial structure in terms of surface segregation caused by differences in the interaction energies between two components [149, 150]. With this in mind, the next segment of work attempts to study the interface of simple linear ester oligomers representative of PLA as introduced in Chapter 1.

As the interface is studied via molecular dynamics simulation, as discussed in

Chapter 2, a forcefield representation of the molecular system is required. For the PLA ester oligomer, this requires a set of non-bond parameters (to handle both inter- and intramolecular dispersion and electrostatic interactions) and a set of internals to represent two-, three-, and four-body interactions within the molecule (these are the standard bond, angle, torsion, and out-of-plane potentials discussed earlier). For the present study, the dispersion and electrostatic interactions are modeled using the OPLS parameters used in the small molecule simulations of the previous chapters. For the internals, it was initially considered that parameters for simple potentials from existing forcefields would be applied and merged with the OPLS non-bond parameters to arrive at a total ester oligomer forcefield. This approach of merging parameters is not new, having been applied in the literature and even in the development of the more common forcefields available.

Figure 6-1 shows the PLA ester oligomer model targeted for study. Based on a simple persistence length calculation using a repeat unit length of $l_u = 0.37$ nanometers calculated by Flory *et al* [151] and a characteristic ratio of $C_{infty} = 2.0$ determined from light scattering and intrinsic viscosity measurements of solutions of two different molecular weight polymers [151],

$$L = \frac{1}{2} l_u (C_\infty + 1) \tag{6.1}$$

$$L = 0.555nm = 1.5 l_u \tag{6.2}$$

(the persistence length is roughly 1.5 times the repeat unit length), the oligomer is taken to be five repeat units, the repeat unit shown in Figure 6-1. This is felt to be a good trade-off between the need to represent a true 'chain-like' structure (for which conformational properties averaged over the entire chain will have statistical meaning) and the need to maintain the total number of particles involved to a tractable value. This is not to say that this chain length will be representative of much larger chains, but it will at least afford a model from which are derived statistically significant conformational properties. The relevant internal coordinates (bond angles,

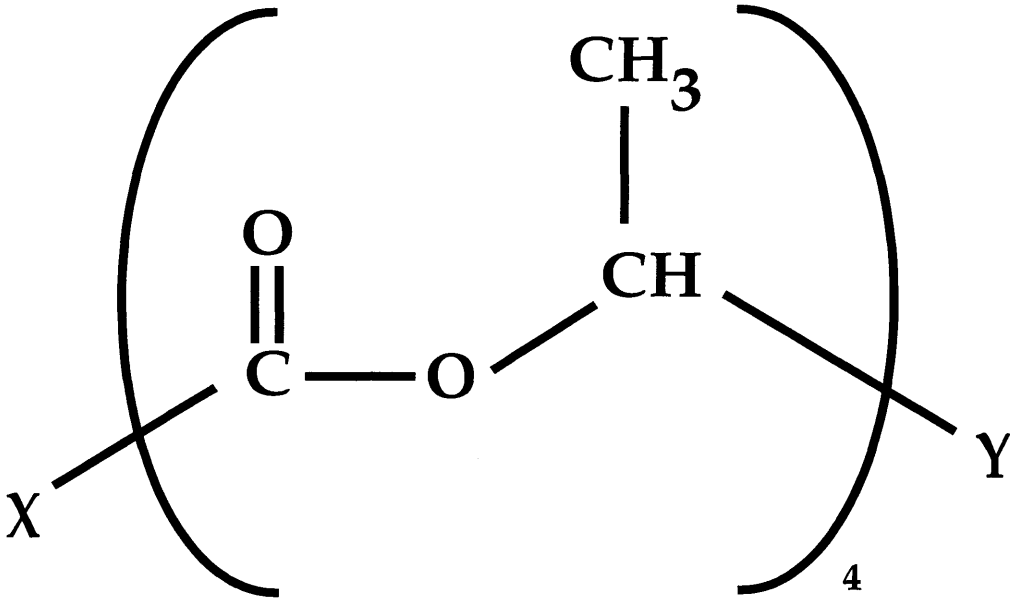
Table 6.1: United-Atom Internals for Ester Oligomer Model

Type	Atoms	Quantity
Bond Angles	$CH_3 - C^* - O'$	1
	$CH_3 - C^* - O$	1
	$O' - C^* - O$	5
	$C^* - O - CH$	4
	$O - CH - CH_3$	4
	$O - CH - C^*$	4
	$CH_3 - CH - C^*$	4
	$CH - C^* - O'$	4
	$CH - C^* - O$	4
	$C^* - O - CH_3$	1
Torsion Angles	$CH_3 - C^* - O - CH$	1
	$CH_3 - CH - C^* - O'$	4
	$CH_3 - CH - C^* - O$	4
	$CH - C^* - O - CH$	3
	$O' - C^* - O - CH$	4
	$C^* - O - CH - CH_3$	4
	$C^* - O - CH - C^*$	4
	$O - CH - C^* - O'$	4
	$O - CH - C^* - O$	4
	$CH - C^* - O - CH_3$	1
	$O' - C^* - O - CH_3$	1
Out-of-Plane Angles *	$CH - C^* - O - O'$	1
	$CH_3 - C^* - O - O'$	1

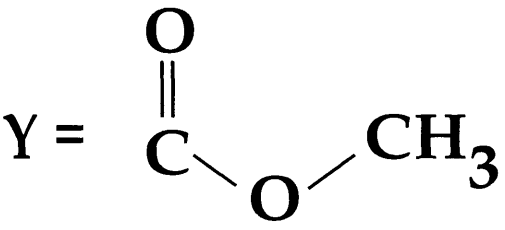
torsions, and out-of-plane angles) are shown in Table 6.1 along with a description of the nomenclature used to define the internals.

Unfortunately, a search of the available united-atom forcefields yielded an incomplete set of parameters for the internal potentials. This was surprising as the system under study is a fairly simple one relative to the types currently studied using molecular simulation approaches. Consider that aromatic polymers, cyclic aromatic and non-aromatic group containing polymers, nylons, polyamides, DNA, proteins, and the like are routinely investigated via MD. Yet for the simple linear esters studied here, one cannot find the relevant internals in the literature. With this being the situation, the methodology espoused was that introduced in Chapter 3 for the parameterization of the internals for methyl acetate. An all-atom forcefield would be used to generate distributions for the required internals, and united-atom potentials would be fit to

PLA Oligomer Model



$X = CH_3$



United-Atom Representation for -CH and -CH3 Groups

Figure 6-1: Oligomer Model

these distributions. The all-atom forcefield selected for the parameterization was the CFF91 potential; however, as it turns out, even this Class II next-generation forcefield was lacking in a particular torsion parameter set. At this point, several options were available. First, it was suggested [152] that an analogous torsion parameter set within the CFF91 forcefield be used to model the torsion. Second, one could attempt to find another all-atom parameter set to use for the fitting; this option was rejected as the CFF91 forcefield was felt to be adequate in terms of its ability to model the internals accurately. Thirdly, the missing torsion could be generated via *ab initio* computations on a model compound, this torsion then to be introduced into the CFF91 parameter set which would be used to generate internals distributions to which a united-atom parameter set would be fitted. This last is the approach taken for the current work. The remainder of this chapter discusses 1), the *ab initio* computations, 2), the fitting of a potential to the quantum mechanically derived torsional profile, 3), the integration of the *ab initio* torsion into the CFF91 forcefield and subsequent generation of internals distributions via single-molecule Hybrid Monte Carlo, and 4), the fitting of a united-atom potential to the distributions.

The flow diagram for the torsion parameterization is shown pictorially in Figure 6-2 . The following sections discuss aspects of this flow diagram.

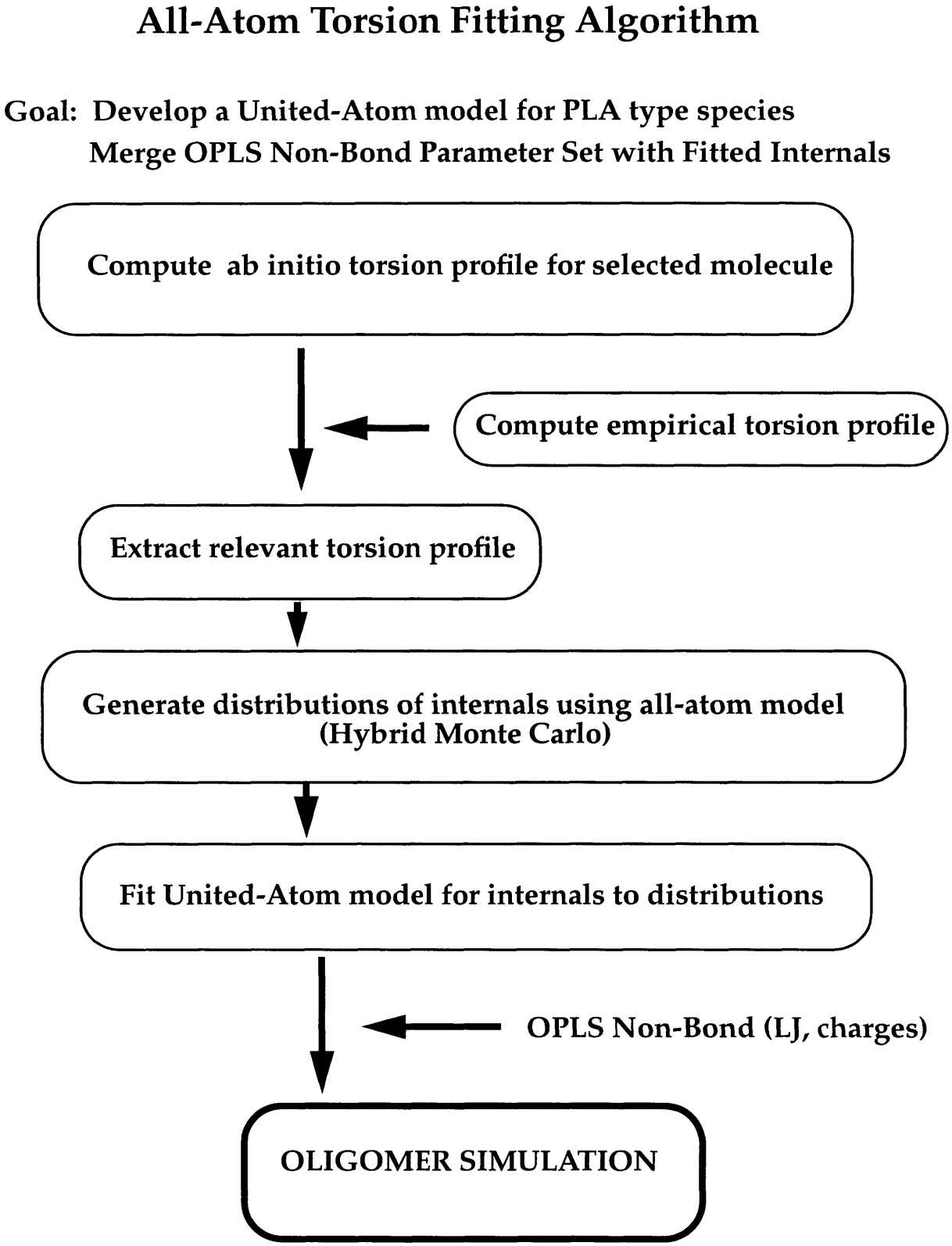


Figure 6-2: United Atom Model for Ester Oligomer Simulation: Fitting Algorithm

6.2 *Ab Initio* Computations: Quantum Mechanical Torsion Profile for Model Ester Compound

6.2.1 Applied Quantum Mechanics: Application of *ab initio* Methods to Chain Conformational Energetics

Over the last decade, the methods of *ab initio* quantum mechanics have been increasingly applied to the study of chain molecules and polymers. Moreover, information accessible via these techniques has been used to ascertain functional forms of the various contributions to the molecular potential energy surface, especially with respect to the development and validation of more flexible, highly transferrable force fields. Relevant to polymers, as well as medium to large molecules and oligomeric/macromolecular species, in the liquid and solid phases, is the chain conformation; it is this component which directly relates to the bulk structure which in turn dictates macroscopic, rather 'practical', properties such as optical, mechanical, and electrical properties. Thus, from a theoretical/computational perspective, one requires a sound, accurate description of the rotational behavior of the molecule; more specifically, the energetic differences between stable energy conformers and the relevant barrier heights between them. Experimentally, data on conformational energetics for small molecules comes from microwave spectra of the gas phase, infrared spectroscopy of both gas and liquid phases, and vibrational spectra and dielectric and/or mechanical relaxation phenomena. The latter approach entails tedious work in assigning observed loss peaks to specific molecular rotations. Vibrational spectra require assignment of all relevant vibrational frequencies as well as study of deuterated species. Furthermore, the experimental techniques will not give information about the 'exact' form of the potential energy curve describing the conformational energy surface. Thus, applied quantum mechanical methods allow one to extract a detailed potential energy profile which can then be compared to quantitative experimental data.

The literature is replete with *ab initio* studies of the conformational energetics of a wide range of molecules. *Ab initio* studies on perfluoroalkanes, alkyl ethers, 1,2-dimethoxyethane [153], poly(p-hydroxybenzoic acid) (PHBA) and poly(ethylene terephthalate) (PET) [154]. In general, these investigations do not generate the full potential energy profile, focusing more on relative energies between low-lying conformers which can then be used in rotational isomeric state (RIS) type models [155]. Even force field development work invariably neglects the full torsional profile, opting instead to determine the stable configurations and fitting parameters to a selected model to represent the energy profile (this is necessarily due to the large number of time-intensive calculations that would be needed).

6.2.2 Model System and Computational Details

In this section, the model ester compound used in the *ab initio* calculations is presented. Also discussed are the details of the computations in terms of basis sets, minimization algorithms, algorithm performance.

Figure 6-3 shows the ester compound selected to probe the o-c-c*-o torsion (indicated by the arrow in the same figure). This was selected as it is small enough to capture the local environment of the torsion while keeping the number of heavy atoms small enough to make the computation tractable. Furthermore, since the specific torsional component was to be extracted by subtraction of the CFF91 partial forcefield from the total electronic torsion profile, no new unparameterized torsions could be introduced in the model compound. All the necessary potential parameters for this compound were available in the CFF91 forcefield.

The *ab initio* calculations were performed using the GAUSSIAN92 software package. Calculations were performed at the Hartree-Fock self-consistent field (SCF) level with 6-31G** and 6-311G** basis sets. For polar systems, the conformational energetics are substantially influenced by the size of the basis set as well as inclusion of electron correlation effects. Naturally, the time required for a given calculation varies with the size of the basis set as well–computational time scales as the fourth power of the number of basis functions describing the system. These computations involved

Ab Initio Ester Compound Model

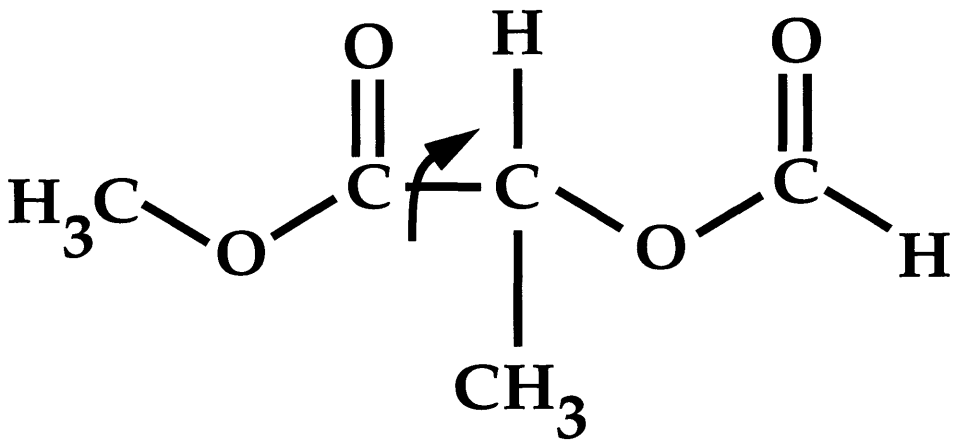


Figure 6-3: *Ab Initio* model ester compound for torsion fitting. The arrow indicates the O-C-C*-O torsion.

geometry optimizations at a constant value of the dihedral angle. Electron correlation corrections were computed via Moller-Plesset perturbation theory (many-body perturbation theory) up to second order using the 6-311G** and D95+ basis sets; the electron correlation calculations were performed only for the 6-311G** optimized geometries. Calculations of the harmonic normal mode vibrational frequencies for each of the energy minimizations were used to confirm that an energy minimum was reached; all normal mode eigenvalues are positive at the energy minimum (positive definite diagonalized Hessian matrix). The basis functions applied are split valence with polarization functions added for all atoms. The use of polarization functions has been cited in the literature as being important for an accurate description of dispersion interactions. Further concerning electron correlation effects, for conformational energetics studies, the primary contribution comes from dispersion energy which is sufficiently handled by second order perturbation theory; for instance, the inclusion of higher levels of perturbation does not seem to enhance the trans-gauche energy difference for 1,2-difluoroethane [153]. Thus, for the present work, only second or-

der contributions are considered. Furthermore, test calculations at the MP4 level of correction performed as part of the current work show that the fourth order corrections are two orders of magnitude smaller than second-order contributions, and thus substantiate the neglect of these terms.

At this point, some mention is made of the basis sets used and the number of such functions involved for the various basis sets employed in this work. The basis functions used, as mentioned above, are the split-valence sets of 6-31G**, 6-311G**, and Dunning's D95+ basis set.

The simplest basis set, or the minimal basis set, is the STO-3G set, which is a Slater-type spatial orbital represented as a linear combination of 3 primitive gaussian functions. The minimal basis set, although useful qualitatively, comes with several inherent inadequacies. One of the major drawbacks is that such basis sets cannot account rigorously for anisotropy of the molecular charge distribution (i.e., polarization of the electron cloud). In response to these deficiencies, split-valence basis sets with polarization functions are used to describe systems with polar character (as well as systems in which resonance plays a significant role in stabilization). These are composed of inner, core, electron shells similar to the equivalent to the minimal basis set, and valence shells separated into two or three spatial regions determined by the exponent of the radial component. For the present case, we consider the 6-31G** and 6-311G** basis sets. For the former, inner shells are represented by contractions, or linear combinations, of 6 primitive gaussian functions, or gaussian primitives (much in the vane of STO-KG basis sets, with K = 6 in this case). The valence shells are split into two regions, the inner region represented by a contracted gaussian of 3 primitives, and the outer region by a single primitive. The double asterisks indicate that hydrogen atoms are given 3 additional p-type gaussian primitives in the valence shell, and heavy atoms are given 6 3d-type gaussian primitives in the valence shell. For the latter basis set (6-311G**), the inner shells are again represented by contractions of 6 gaussian primitives while the valence shells are split into three regions, the inner region a contraction of 3 gaussian primitives, and the middle and outer regions single gaussians. Hydrogen atoms again receive an additional 3 p-type gaussian primitives

in the valence shell; however, heavy atoms are assigned only 5 d-type gaussians as opposed to the 6 for the 6-31G** set.

The basis functions for the 6-31G set are as follow:

$$\phi_{1s}(\mathbf{r}) = \sum_{i=1}^{6} d_{i,1s} g_{1s}(\alpha_{i,1s}, \mathbf{r}) \tag{6.3}$$

$$\phi'_{2s}(\mathbf{r}) = \sum_{i=1}^{3} d'_{i,2s} g_{1s}(\alpha'_{i,2sp}, \mathbf{r}) \tag{6.4}$$

$$\phi''_{2s}(\mathbf{r}) = g_{1s}(\alpha''_{i,2sp}, \mathbf{r}) \tag{6.5}$$

$$\phi'_{2px}(\mathbf{r}) = \sum_{i=1}^{3} d'_{i,2px} g_{2px}(\alpha'_{i,2sp}, \mathbf{r}) \tag{6.6}$$

$$\phi''_{2px}(\mathbf{r}) = g_{2px}(\alpha''_{2sp}, \mathbf{r}) \tag{6.7}$$

$$\phi'_{2py}(\mathbf{r}) = \sum_{i=1}^{3} d'_{i,2py} g_{2py}(\alpha'_{i,2sp}, \mathbf{r}) \tag{6.8}$$

$$\phi''_{2py}(\mathbf{r}) = g_{2py}(\alpha''_{2sp}, \mathbf{r}) \tag{6.9}$$

$$\phi'_{2pz}(\mathbf{r}) = \sum_{i=1}^{3} d'_{i,2pz} g_{2pz}(\alpha'_{i,2sp}, \mathbf{r}) \tag{6.10}$$

$$\phi''_{2pz}(\mathbf{r}) = g_{2pz}(\alpha''_{2sp}, \mathbf{r}) \tag{6.11}$$

There are 6 d-type polarization functions for heavy atoms and 3 p-type functions for hydrogens; this description gives the 6-31G** basis set. Thus, for hydrogen atoms, there are five basis functions with 7 gaussian primitives; for carbon atoms, there are 15 basis functions with 28 gaussian primitives; and for oxygen atoms, there are similarly 15 basis functions with 28 gaussian primitives. These figures include the

requisite polarization functions for the atom. For the oligomer model, there are thus a total of 175 contracted basis functions with a total of 308 primitive gaussians.

For the 6-311G basis set, the individual basis functions are similar to the 6-31G functions given above with the addition of an additional gaussian primitive for the valence shell (thus giving a triple valence shell).

$$\phi'''_{2s}(\mathbf{r}) = g_{1s}(\alpha'''_{2sp}, \mathbf{r}) \tag{6.12}$$

$$\phi'''_{2px}(\mathbf{r}) = g_{2py}(\alpha'''_{2sp}, \mathbf{r}) \tag{6.13}$$

$$\phi'''_{2py}(\mathbf{r}) = g_{2py}(\alpha'''_{2sp}, \mathbf{r}) \tag{6.14}$$

$$\phi'''_{2pz}(\mathbf{r}) = g_{2py}(\alpha'''_{2sp}, \mathbf{r}) \tag{6.15}$$

Thus, for the 6-311G** basis, hydrogens have 6 basis functions with 8 gaussian primitives; carbons and oxygens have 18 basis functions with 31 gaussian primitives. For the oligomer model, there are thus a total of 210 contracted basis functions with a total of 343 primitive gaussians.

The D95+ basis set, developed by Dunning and coworkers [156], is a split-valence, double-zeta basis set. Double zeta basis sets are the next level of sophistication above minimal STO-3G type sets; the double-zeta nomenclature refers to the fact that two functions are used for each function of the mininal basis. The orbital exponents of the two functions are taken to be slightly above and below the optimal minimal basis set exponent. This allows for expansion or contraction of the orbital spatial extent by variation of the linear coefficients rather than non-linearly through the exponents. The SCF algorithm will weight the contributions of the two functions so as to effect a contraction of expansion of the molecular orbital depending on the molecular environment. For the present oligomer model, the D95+ basis set entails 220 basis functions with 362 primitive gaussians. For further detailed discussion of

basis functions and *ab initio* theory, the reader is referred to Appendix A and the references therein.

Computationally, each of the constrained minimizations cost roughly 14 hours of CPU time on a DEC Alpha 600 at the 6-311G** SCF level. Each MP2 calculation using the same basis set cost approximately 7957.9 cpu seconds. Standard exponents and contractions for each basis set were employed; these are given in Appendix C [157, 158, 159].

6.2.3 *Ab Initio* Torsion Profile: Results

Figure 6-4 shows the o-c-c*-o torsion profile computed at the HF-SCF level using the 6-31G** and 6-311G** basis functions along with the electron correlation calculations on the 6-311G** data. The MP2 results are for 6-311G** and D95+* basis functions. Table 6.2 shows results of total geometry optimizations at the HF-SCF 6-31G**, 6-311G**, 6-31G+**, and 6-311G+** levels. This table shows that for the present system, addition of the diffuse functions (signified by (+)) does not give rise to as significant a change in the electronic energy as does the inclusion of polarization functions (signified by (*)). This is reassuring since diffuse functions are generally important in cases where electronic charge is significant far away from the atomic center, as in the case of anions. Based on calculations at the 3-21G, 3-21+G, 6-31G*, and 6-31+G* levels of one-heavy-atom hydrides, Hehre et al show that the energy-lowering effect is dramatically more significant for the negatively charged species (deprotonated) of the associated neutral parent hydride. Thus, based on this result, further consideration was given to the 6-311G** profile.

Figure 6-5 shows the conformations of the model compound at various values of the torsion for illustrative purposes. In each configuration shown, the carbonyl carbon to backbone aliphatic carbon bond is projecting out at an angle from the plane of the figure. The four atoms comprising the dihedral are labelled respectively. The conformations at -180 and +40 degrees correspond to the two minimum energy regions on the ensuing torsion energy profiles, and the conformations at -60 and +120 degrees represent the structures at the barrier heights. From this data it is evident that the

Table 6.2: Effect of Polarization and Diffuse Functions

Basis Set	Electronic Energy a.u.	Time (hours)
6-31G**	-493.473101048	5.5
6-31G+**	-493.484915842	7.5
6-311G**	-493.588702731	7.5
6-311G+**	-493.5980072	14

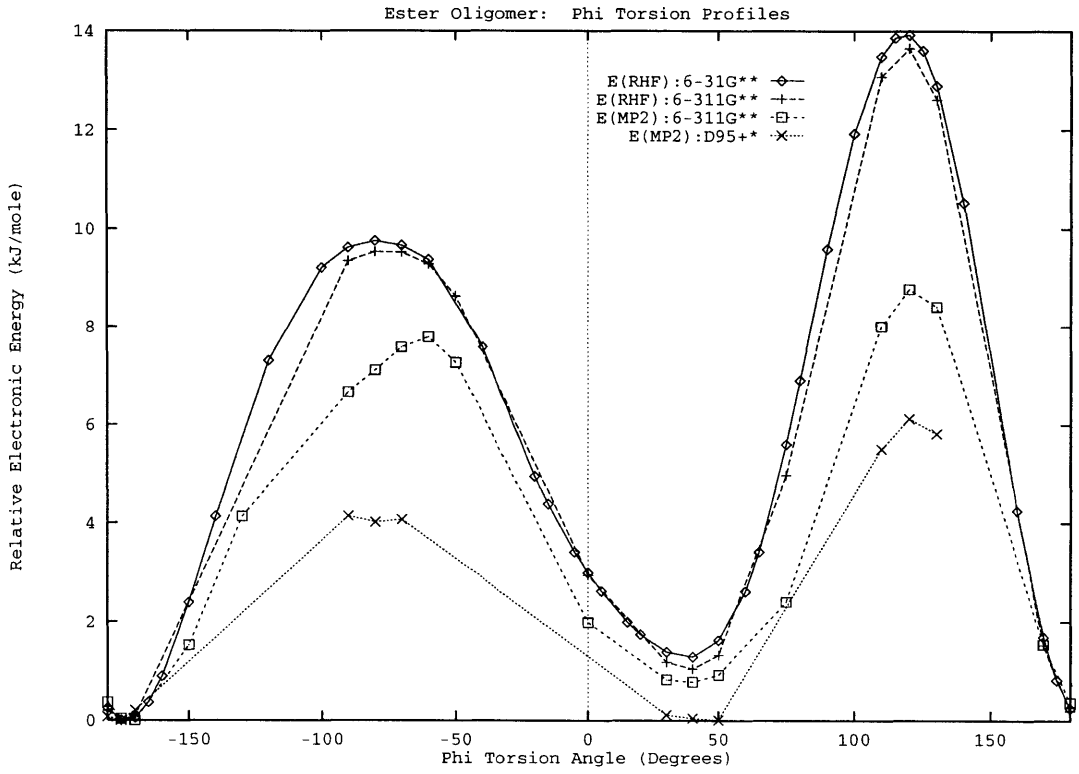


Figure 6-4: *Ab Initio* Torsion Profiles at Varying Levels of Theory

minimum energy conformer corresponds to the ether oxygens almost trans to each other. There is not much difference between the 6-311G** and 6-31G** profiles, but the addition of the electron correlation energy for the 6-311G** geometry significantly reduces the barrier heights while leaving the relative minimum energy difference less perturbed. The D95+* MP2 correlation correction dramatically reduces the barrier heights to about half of the HF-SCF values, and gives virtually two equally stable conformers, the gauche form at 50 degrees becoming equivalent to the trans form at -175 degrees. This reduction in the energy difference is documented in the literature for this basis set. There is a slight difference of roughly 5 degrees in the location of the higher energy minimum computed for the RHF/6-311G** and MP2 levels. However, the difference in the energies of the two conformers is a small fraction of a kJ/mole, and a difference of such small magnitude is not overly debilitating for the purposes of the current work. More importantly, both levels of theory agree very well on the energy difference between the minima. Based on the current calculations, the MP2/6-311G**//HF/6-311G** profile is used for the torsion development.

6.3 Treatment of Electron Correlation in Molecular Conformation Studies

As there has arisen some controversy in the literature as to the correct approach to inclusion of electron correlation in optimization calculations, some thoughts are presented here. The paradigm at the time of this work involved the methodology employed; full optimizations are performed at a lower level of theory followed by the calculation of an estimate of the correlation effects using configuration interaction or perturbation theoretical methods. Some researchers suggest that this approach leads to inconsistencies and *inaccuracies* since the geometric and energetic properties evaluated using full MP2 optimization are different than those computed by performing electron correlation calculations after Hartree-Fock computations [160]. However, full MP2 optimizations are practical only for small systems, on the order of 10 or fewer

heavy atoms [161]. Thus, one is forced to weigh computational resources against desired level of theory. Although it has been suggested in the literature that non-bond dispersion interactions giving rise to correlation effects are important in the study of conformational energetics of large molecules [161], thus implying that full optimizations are more rigorous, the current work has accepted this drawback in the methodology in order to perform the necessary computations in a reasonable amount of time. We note that it has also been cited in the literature that the effects of electron correlation on rotation barriers is small (on the order of 0.1 kcal/mole); we however, do not see this to be true from the data presented earlier, and thus have opted to account for correlation at some level. It would be interesting to perform the torsion profile calculations again using density functional methods which have in the last five years come to be a powerful tool in studying electronic and structural properties of metals and polymeric materials, chemical reactions and related processes, molecular geometries, and relative conformational energies [162]. These methods are faster relative to *ab initio* methods, even in the case where full optimizations including electron correlation are performed. Furthermore, in the case of n-alkanes, DFT conformational energies and internal rotation barrier heights agreed better with experimental data than did *ab initio* calculations [163]. Another study suggested that gradient-corrected density functional methods led to more accurate conformational energetics (with respect to experimental data) than corresponding *ab initio*-based methods [162, 164]; it is held that the non-local correction is required to accurately account for dispersion interactions.

6.4 Fitting of Torsion Potential to *ab initio* Calculations

In this section, the fitting of a functional form to the inherent torsion profile is discussed. Having computed a quantum mechanical torsion profile as outlined in the previous section, the same model compound was simulated using the DISCOVER

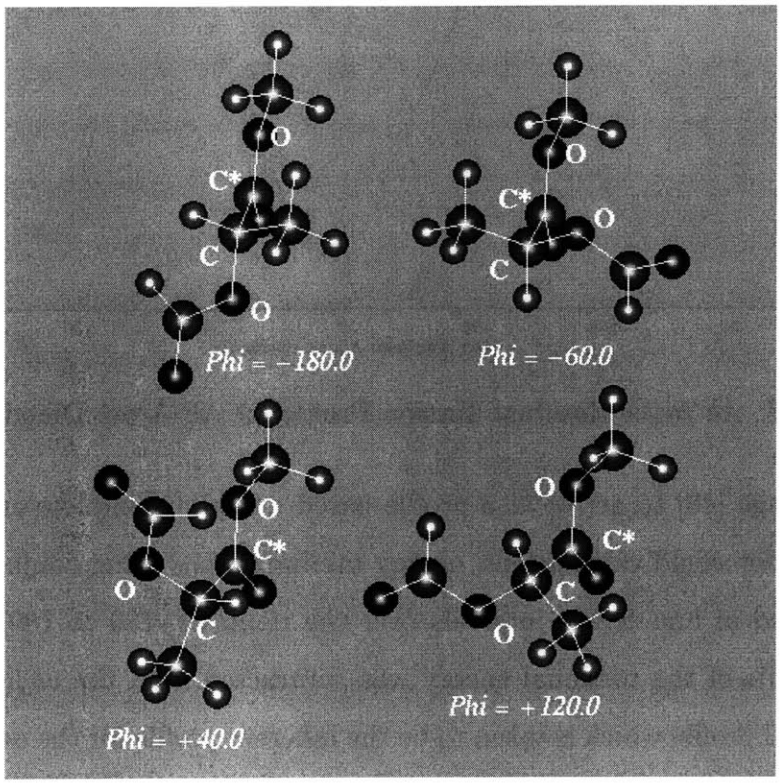


Figure 6-5: *Ab initio* Model Compound Shown with Varying Values of the O-C-C*-O Torsion Angle

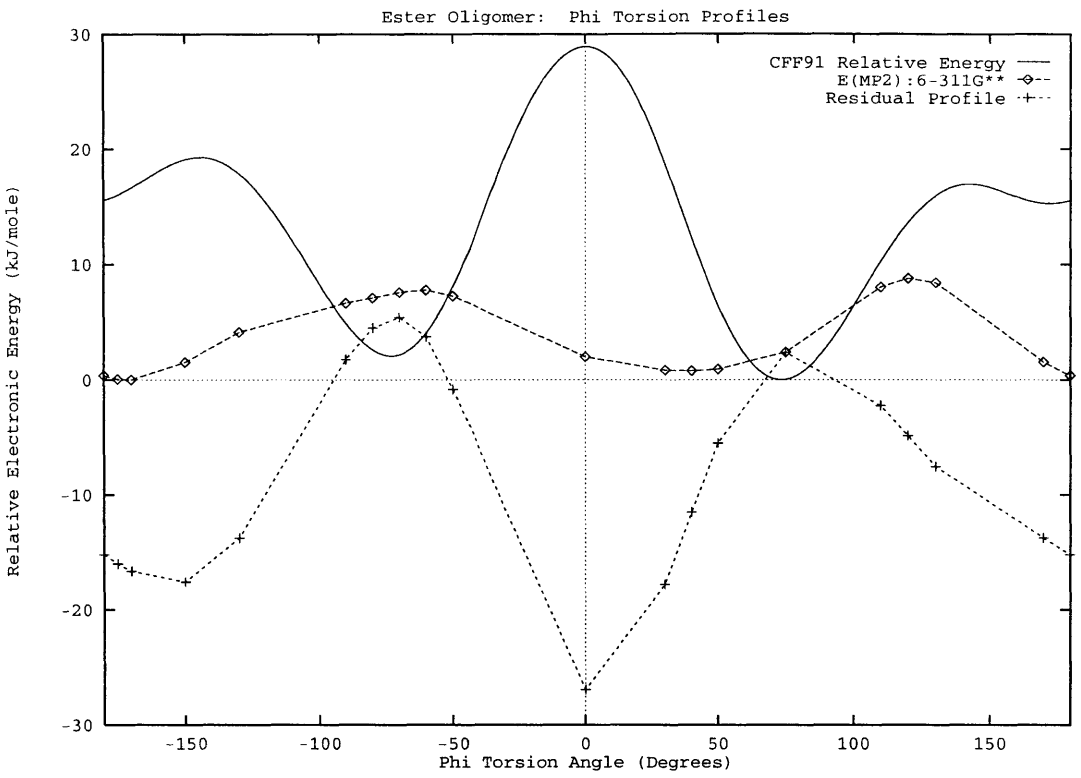


Figure 6-6: *Ab Initio* Residual Torsion Profile for All-Atom Dihedral Model

software package [48] to arrive at a profile based on the incomplete empirical forcefield. For the forcefield calculations, energy minimizations using conjugate gradients were performed in one-degree intervals over the range of -180 to 180 degrees. The empirical profile of the torsional energy was subtracted from the *ab initio* profile to yield a residual profile which is taken to be the inherent profile for the o-c-c*-o torsion. Figure 6-6 shows the *ab initio* , empirical, and residual profiles.

The functional form of the torsion potential is chosen to be :

$$U(\phi) = A(1 + cos(\phi + \phi_{1o})) + B(1 + cos(2\phi + \phi_{2o})) + C(1 + cos(3\phi + \phi_{3o})) \quad (6.16)$$

The functional form of the CFF91 forcefield is retained. The fitting of the parameters A , B , C , ϕ_{1o} , ϕ_{2o} , and ϕ_{3o} was performed via weighted linear least-squares using a singular value decomposition algorithm [165]. The approach involved systematically varying the angle variables for which the χ^2 error was computed. That is,

Table 6.3: All-Atom Torsion Parameters

Parameter	Value	
A	-0.3092 kJ/mole	-
B	14.187 kJ/mole	-
C	5.586 kJ/mole	-
ϕ_{1o}	0.6807 rad	39.35
ϕ_{2o}	2.862 rad	164.0
ϕ_{3o}	3.14159 rad	180.0

each of the phase angles was varied in the range of $-\pi$ to $+\pi$ in increments of one degree. At each set of phase angle values, the linear least squares fit for the linear parameters A, B, and C was performed; this calculation resulted in a value for χ^2 as a function of the parameters. The weighting scheme involved using a Boltzmann weight for each torsion angle; the energy used in the exponential factor was actually relative to the mininum (in the global sense over the domain of the torsion profile) ,

$$W_\phi = exp\left(\frac{-(U_\phi - U_{min})}{kT}\right) \tag{6.17}$$

The set of values of the three phase angles and three linear parameters giving the lowest χ^2 error was accepted as the best fit in the chi-squared sense. One may question the choice of interval in the angle search, in this case one degree. We feel that this is not too coarse a value since the error of a fraction of a degree incurred with this selection is negligible relative to the magnitude of the cumulative errors in the preceding calculations—the *ab initio* and MP2 correction calculations. For the MP2 corrected profile, the fitting gives the parameters shown in Table 6.3 .

Figure 6-7 shows a plot of the MP2 residual profile and the fitted function,

$$U(\phi) = A(1 + cos(\phi + \phi_{1o})) + B(1 + cos(2\phi + \phi_{2o})) + C(1 + cos(3\phi + \phi_{3o})) \tag{6.18}$$

The fit to the residual profile is quite good, particularly in the region of the barrier heights. This is a significant outcome in light of the fact that an exponential (Boltz-

mann) weight was assigned to each computed energy point; this scheme is necessary in order to ensure that the regions in the vicinity of the local minima are well-fit since the molecule will for the most part reside in these low-lying conformational states at the conditions we are considering. Note though, how well the singular value decomposition algorithm has allowed fitting of the curve over the energy barriers. We note here that the barrier heights are overestimated and this may be cause for concern with respect to reaching conformational equilibrium. However, we feel that this does not pose a terrible problem based on the following reasoning. Since two barriers are present, with one being lower than the other, equilibration will not suffer. There is a pathway between the two local minima. Furthermore, very few molecules will reside in conformational states along the barriers, and practically none at the barrier heights. Thus, the current fit is accepted.

Thus, at this point, a complete all-atom CFF91 parameter set for the ester oligomer is available and the next step is to perform a single molecule Hybrid Monte Carlo simulation to generate distributions of the internals to which united-atom distributions will be fit.

6.5 Single Molecule Hybrid Monte Carlo: All-Atom Internal Coordinate Distributions

To generate the distributions of bond angles, out-of-plane angles, and torsion angles, a single molecule hybrid Monte Carlo calculation of the sort described in Chapter 3 was performed. A slight addition to the algorithm included a crankshaft move to allow the various torsions to jump barriers to other minima along the respective energy surfaces. The crankshaft follows the standard move found in the literature; the current implementation utilizes a local Flory coordinate system [166] to perform a random move of the relevant site along the circumference of a cone. A brief discussion of the torsion angle selection algorithm is given here. Figure 6-8 shows the schematic of the Flory local coordinate system used as reference for the following discussion.

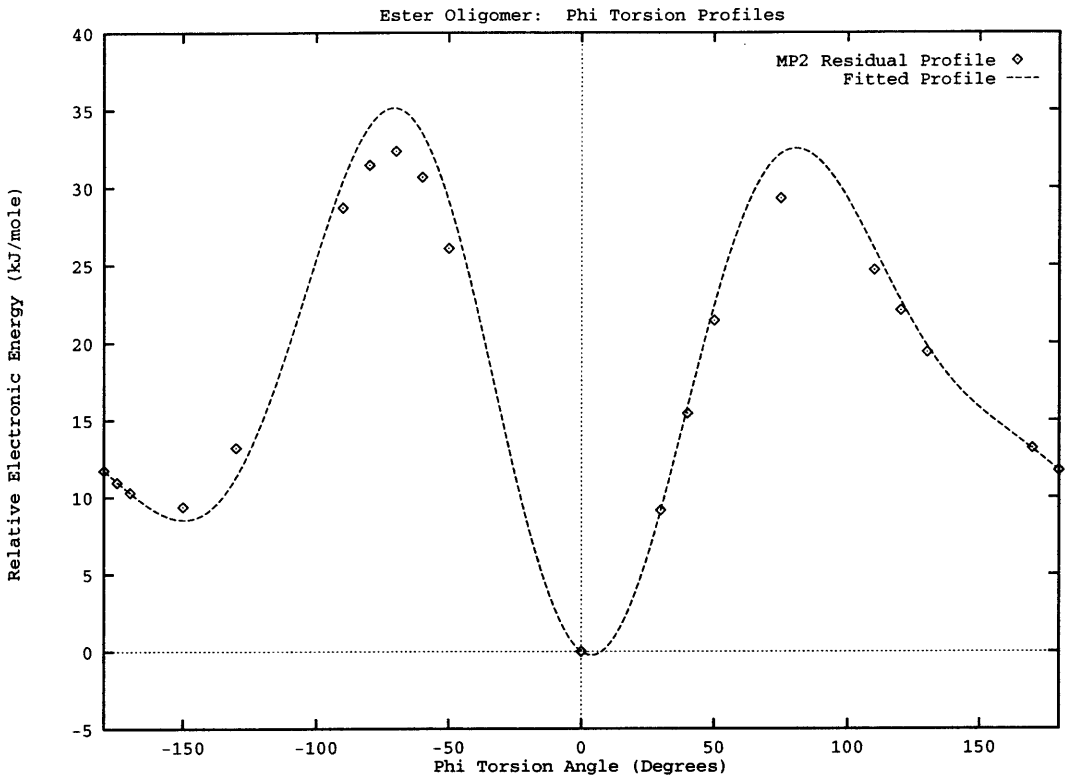


Figure 6-7: MP2 Residual Profile (points) and Fitted Function (solid line0

To perform a random change in a particular torsion, the algorithm first selects at random one of the torsions defined in the molecular topology of the molecule. This definition consists of the four atoms which make up the four-body interaction; for the present case, we are dealing with sites h, i, j, k. Next, the Flory local coordinate system for the bond vector directed from site i to site k is established, and the coordinates of site j are determined within this system. We are interested in moving the selected site j along the circumference of a cone whose axis is the vector connecting sites i and k; this is the k-i bond vector and contains the x-coordinate in the local coordinate system of site i.

Once the coordinates of the j particle are determined in the reference frame of the k-i bond vector, a matrix operation is used to transform from the local coordinates to those in the laboratory, fixed-frame coordinate system. The matrix which allows this transformation is [167],

$$\mathbf{T}_{ik}^{lab} = (\mathbf{x}_{ik}, \mathbf{y}_{ik}, \mathbf{z}_{ik}) \tag{6.19}$$

where the column matrices $\mathbf{x}_{ik}$, $\mathbf{y}_{ik}$, $\mathbf{z}_{ik}$ are defined as,

$$\mathbf{x}_{ik} = \mathbf{u}_{ik} \equiv \frac{\mathbf{r}_{ik} - \mathbf{r}_{hi}}{|\mathbf{r}_{ik}|} \tag{6.20}$$

$$\mathbf{y}_{ik} = \mathbf{z}_{ik} \times \mathbf{x}_{ik} \tag{6.21}$$

$$\mathbf{z}_{ik} = \frac{\mathbf{u}_{ik} \times \mathbf{u}_{hi}}{sin(\theta_{hi})} \tag{6.22}$$

The vectors subscripted with ik pertain to the bond vector i-k. The $\mathbf{y}_{ik}$ and $\mathbf{z}_{ik}$ vectors represent the local coordinate axes; the x-axis is directed along the $i\,k$ bond vector. The supplement of the bond angle at skeletal element h is given by θ_{hi}; it derives from the relation,

$$cos\theta_{hi} = \mathbf{u}_{hi} \cdot \mathbf{u}_{ik} \tag{6.23}$$

The coordinates of the j site after the move are given simply by,

$$\mathbf{r}^{laboratoryframe} = \mathbf{T}_{ik}^{lab} \mathbf{r}^{local} \tag{6.24}$$

Before moving on, we offer some words on the recipe by which a random position of the j'th particle is obtained.

From simple geometric relations, the projection of the k-i bond on the local y- and z-axes are determined. The projections allow the calculation of the angle between the y-z plane projection of the bond and the local y-axis (see Figure 6-9). This angle is then changed by a random amount (the value of the change ranges from $+\pi$ to $-\pi$). The new local projections of the bond are now known, and the transformation matrix is applied to obtain the final fixed, laboratory frame coordinates.

The fitting procedure for bond angles, out-of-plane angles, and torsion angles is

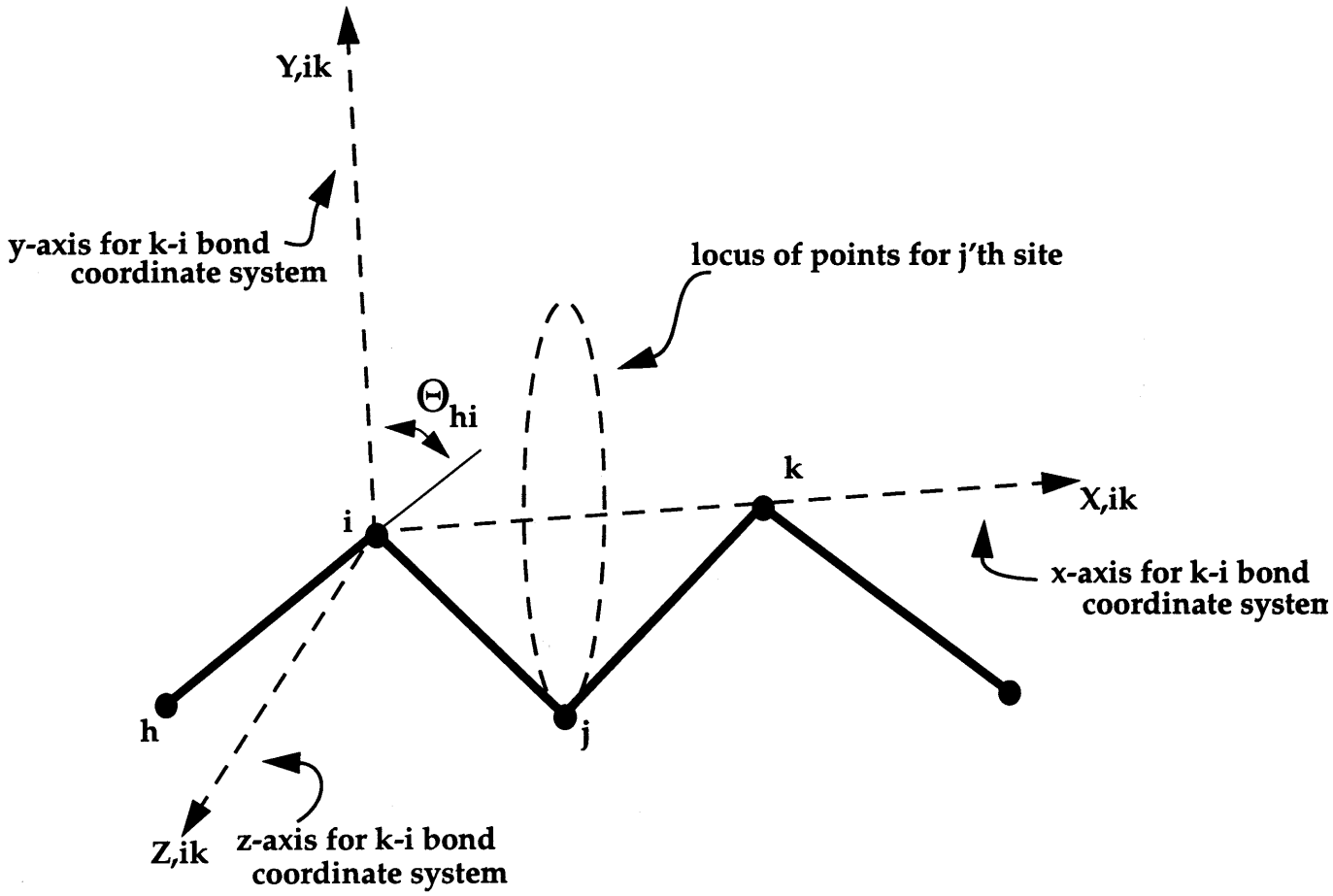


Figure 6-8: Local coordinate system for performing crankshaft move

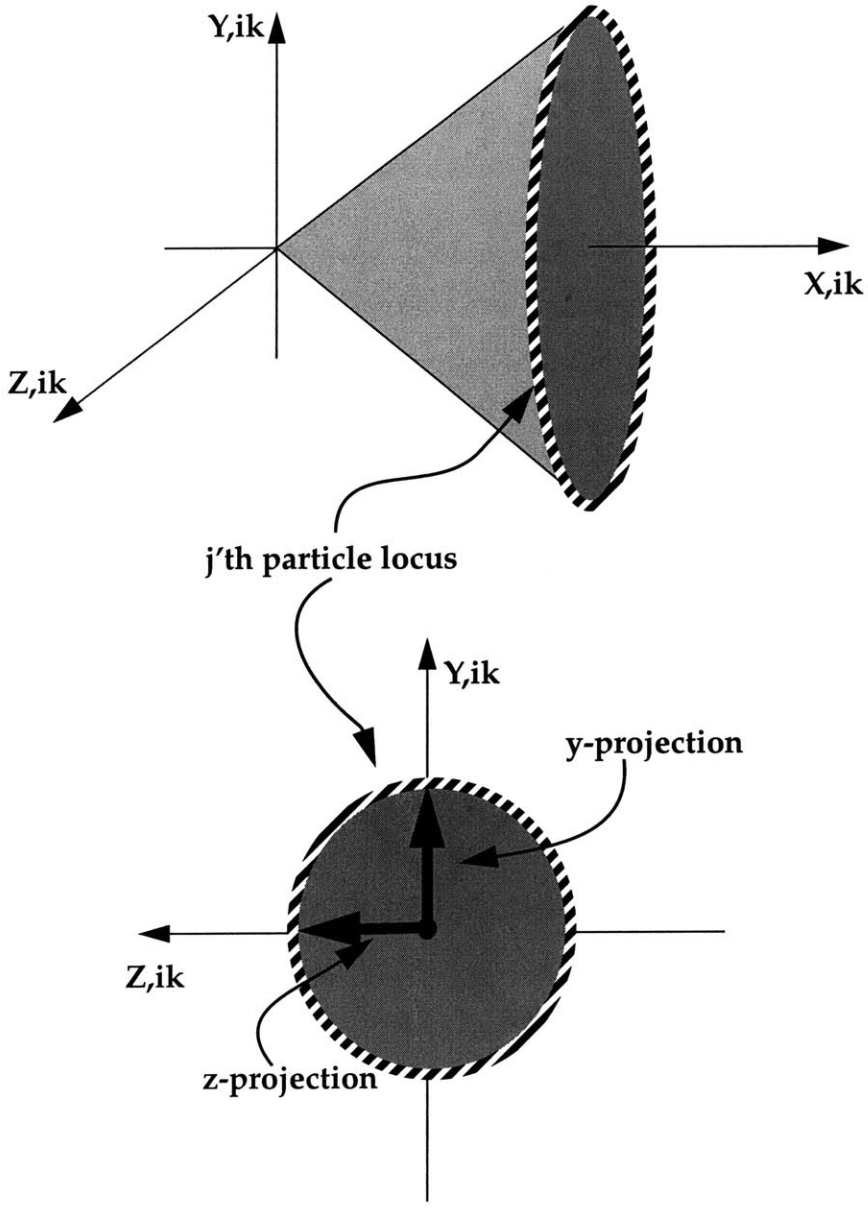


Figure 6-9: Y- and Z-axis projections of j'th particle. The locus of valid points resides on the circular edge of the cone

briefly described here, as much of it follows the algorithm prescribed in Chapter 3 (for the bond angles and out-of-plane angles).

The functional form for the angle bending is:

$$U_\theta = \frac{1}{2} K_\theta (\theta - \theta_o)^2 \tag{6.25}$$

and that for the Wilson out-of-plane motion:

$$U_\chi = K_\chi \chi^2 \tag{6.26}$$

The potentials were parameterized in the variables K_θ and K_χ by performing a least-squares fit of the Boltzmann distribution of angles:

$$P(\theta) = \frac{1}{N_\theta} e^{[(-\frac{1}{2}K_\theta(\theta-\theta_o)^2)/RT]} \tag{6.27}$$

$$P(\chi) = \frac{1}{N_\chi} e^{[(-\frac{1}{2}K_\chi(\chi^2)/RT]} \tag{6.28}$$

where the N_θ and N_χ are normalizing factors:

$$N_\theta = \sum_{i=1}^{M_\theta} (e^{[(-\frac{1}{2}K_\theta(\theta_i-\theta_o)^2)/RT]}) \tag{6.29}$$

$$N_\chi = \sum_{i=1}^{M_\chi} (e^{[(-\frac{1}{2}K_\chi(\chi^2)/RT]}) \tag{6.30}$$

The fitting involved:

1. Picking a value of K
2. Calculating P(θ)
3. Calculating sum of squares (of deviations)
4. Minimizing sum of squares with respect to K

Initial values of K were selected based on typical values of similar bond angles found in the literature.

The functional form for the torsions was taken as a Fourier cosine series:

$$U(\phi) = \sum_{n=1}^{3} A_n cos(n\phi + \delta_n) \tag{6.31}$$

Again, for values of the A_i, the sum of squares of the error between the all-atom distribution and the fitted distribution were computed, and the set yielding the lowest sum of squares is taken as the fit.

The distributions and normalization factors are computed straightforwardly as:

$$P(\phi) = \frac{1}{N_\phi} e^{[(-(\sum_{n=1}^{3} A_n cos(n\phi+\delta_n))/RT]} \tag{6.32}$$

$$N_\phi = \sum_{i=1}^{M_\phi} e^{[-(\sum_{n=1}^{3} A_n cos(n\phi+\delta_n))/RT]} \tag{6.33}$$

Figures 6-10 through 6-31 show the all-atom torsion, angle, and out-of-plane distributions along with the fitted distributions . Tables 7.3, 7.2, 7.4 give the parameters for the functional forms of the various internals.

Figures 6-32 through 6-53 show comparisons of the distributions of internals derived from single molecule hybrid monte carlo simulations using the CFF91 all-atom forcefield and the united-atom parameter set obtained from fitting to the former distributions.

Finally, Table 7.1 shows the Lennard-Jones and charge parameters for the oligomer model. Figure 6-1 shows the neutral repeat unit for the oligomer model. The ends are capped with a united-atom methyl group on each end. The repeating chemistry is neutral and allows a constant residual charge before end-capping (this is not dependent on chain-length, thus allowing one to probe multiple molecular weights). Local neutrality is robust and less vulnerable to erratic electrostatic effects during simulation [168]. Hence, the electrostatic model is built up from a repeat unit perspective. The bond lengths are constrained at the values taken for the OPLS methyl acetate model from Chapters 3, 4, and 5. The Lennard-Jones parameters are taken from OPLS parameters for alkanes and esters, and the charges are from the alkane, ether, and methyl acetate parameter set [67, 80, 82].

Table 6.4: United-Atom Non-Bond Parameters: Ester Oligomer Model

Group Type	ϵ $\frac{kJ}{mole}$	σ nm	charge e	mass (a.m.u.)
C^* , Carbonyl Carbon		0.375	0.55	12.01115
O', Carbonyl Oxygen		0.296	-0.45	15.9994
O, Ether		0.3	-0.4	15.9994
CH		0.38	0.30	13.01912
CH_3, pendant		0.391	0.0	15.03506
CH_3, carbonyl		0.391	0.05	15.03506
CH_3, ether		0.380	0.25	15.03506

Table 6.5: United-Atom Bond Angle Parameters: Ester Oligomer Model

Bond Angle	K_θ $\frac{kJ}{mol\,rad^2}$	θ_o degrees
$CH_3 - C^* - O'$	801.1	123.9308
$CH_3 - C^* - O$	704.7	110.3288
$O' - C^* - O$	1079.2	124.1247
$C^* - O - CH$	582.2	118.2380
$O - CH - CH_3$	577.8	111.2130
$O - CH - C^*$	770.9	108.8752
$CH_3 - CH - C^*$	341.1	113.0788
$CH - C^* - O'$	812.6	124.6914
$CH - C^* - O$	687.0	111.2485
$C^* - O - CH_3$	553.2	117.4483

Table 6.6: United-Atom Torsion Angle Parameters: Ester Oligomer Model

Torsion Angle	A_1 $\frac{kJ}{mol}$	A_2 $\frac{kJ}{mol}$	A_3 $\frac{kJ}{mol}$	δ_1 degrees	δ_2 degrees	δ_3 degrees
$CH_3 - C^* - O - CH$	25.0	-22.0	-5.0	0.0	0.0	0.0
$CH_3 - CH - C^* - O'$	5.0	1.0	-2.0	-30.0	-30.0	-30.0
$CH_3 - CH - C^* - O$	-6.0	1.0	2.0	-25.0	-25.0	-25.0
$CH - C^* - O - CH$	34.0	15.0	12.0	0.0	0.0	0.0
$O' - C^* - O - CH$	-46.0	25.0	-17.0	0.0	0.0	0.0
$C^* - O - CH - CH_3$	6.0	10.0	3.0	0.0	0.0	0.0
$C^* - O - CH - C^*$	11.0	-11.0	-10.0	0.0	0.0	0.0
$O - CH - C^* - O'$	-3.0	-3.0	-1.0	-23.0	-23.0	-23.0
$O - CH - C^* - O$	4.0	-4.0	0.0	0.0	0.0	0.0
$CH - C^* - O - CH_3$	50.0	-50.0	-21.0	0.0	0.0	0.0
$O' - C^* - O - CH_3$	-49.0	23.0	-15.0	0.0	0.0	0.0

Table 6.7: United-Atom Out-of-Plane Angle Parameters: Ester Oligomer Model

Out-of-Plane Angle	K_χ $\frac{kJ}{mol\,rad^2}$	χ_o degrees
$CH - C^* - O - O'$	444.6	0.0
$CH_3 - C^* - O - O'$	444.6	0.0

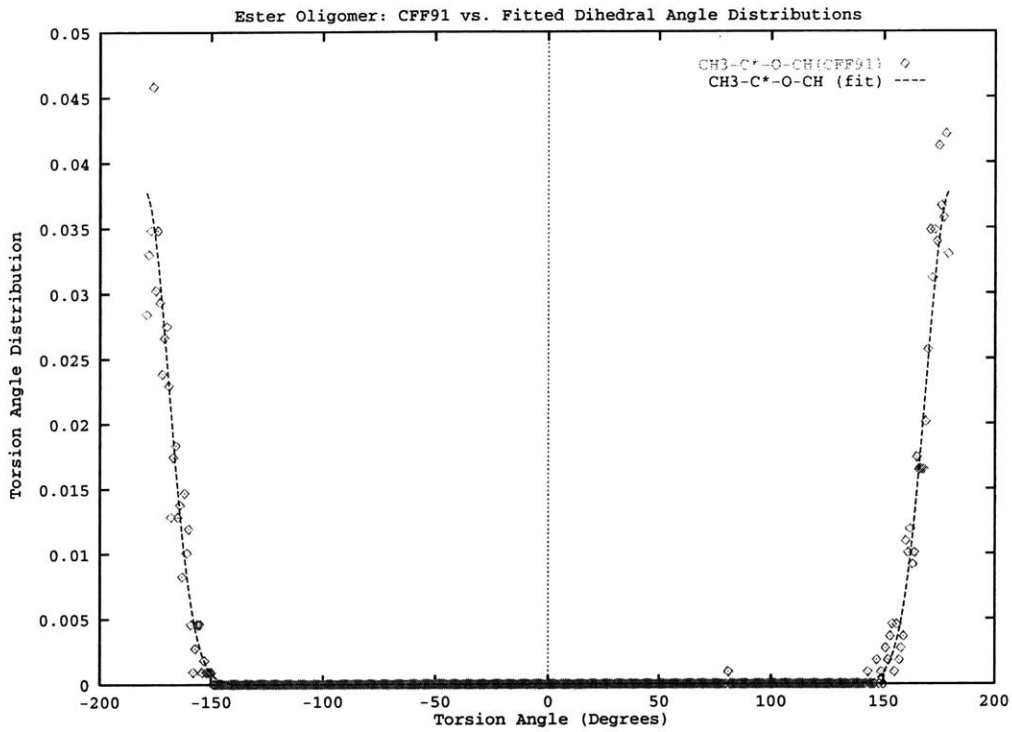


Figure 6-10: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

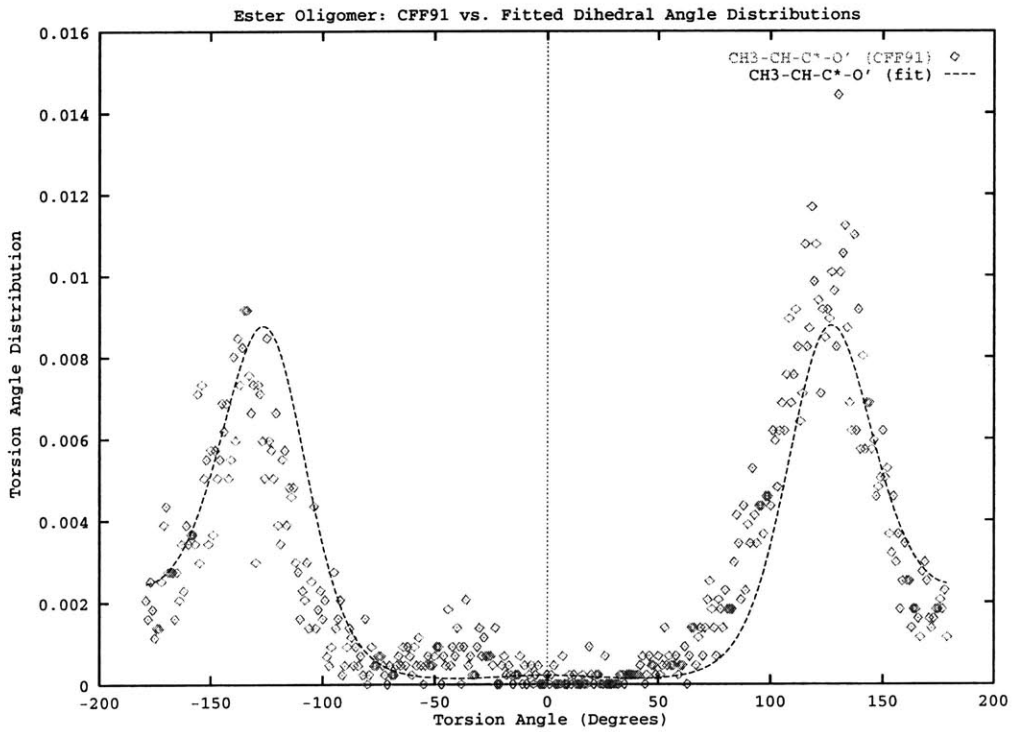


Figure 6-11: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

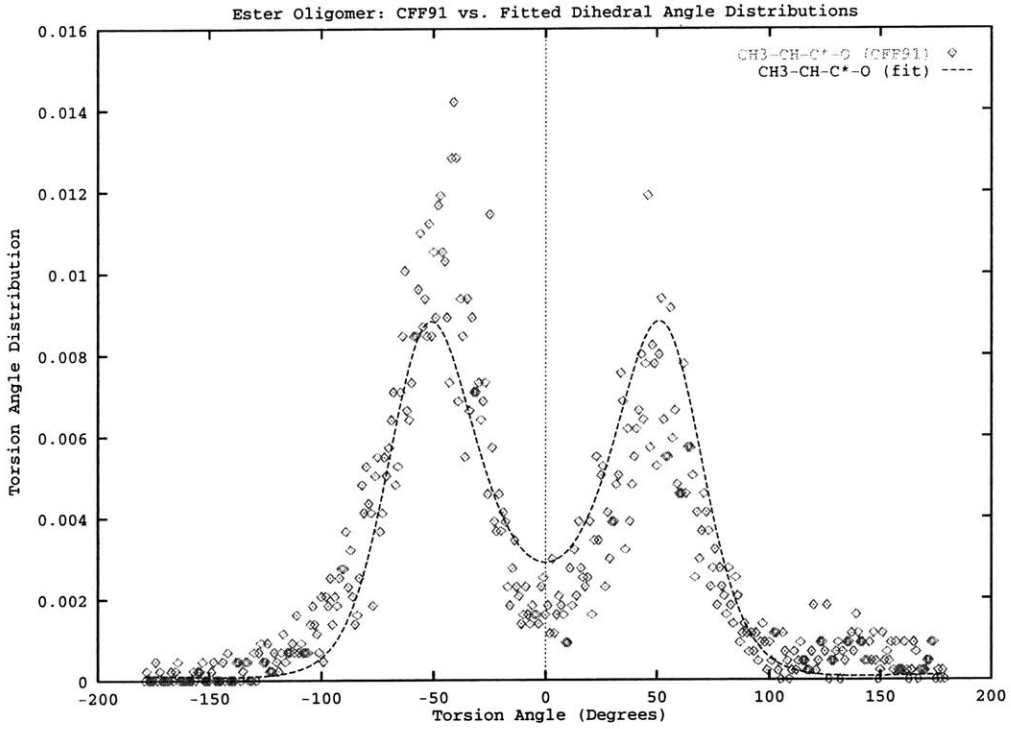


Figure 6-12: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

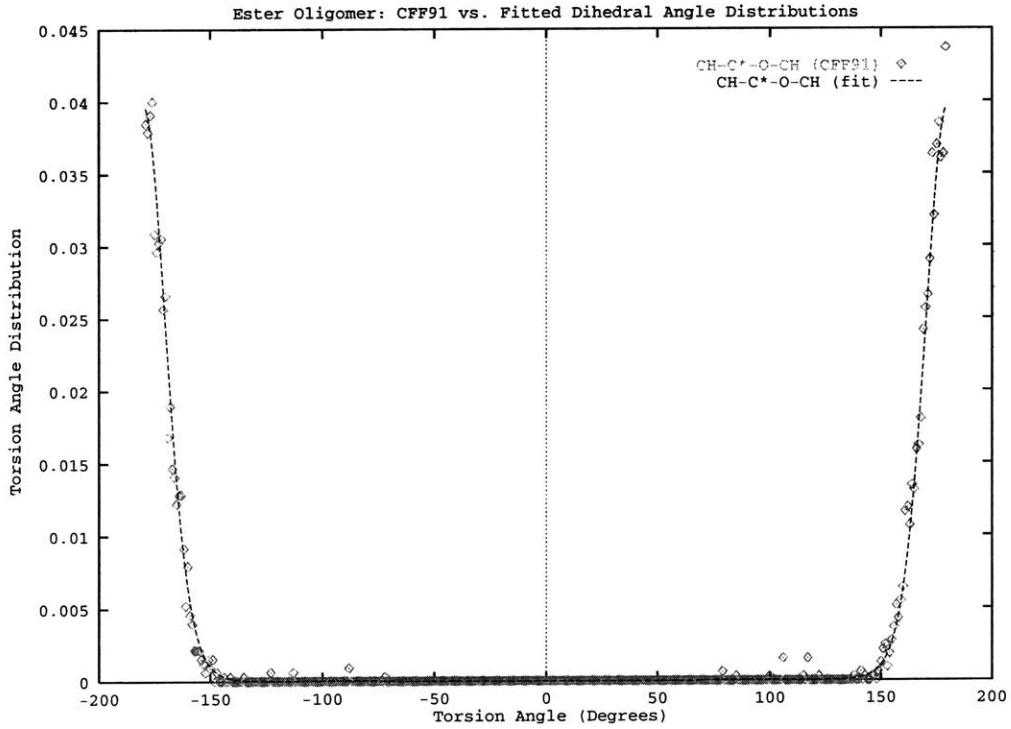


Figure 6-13: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

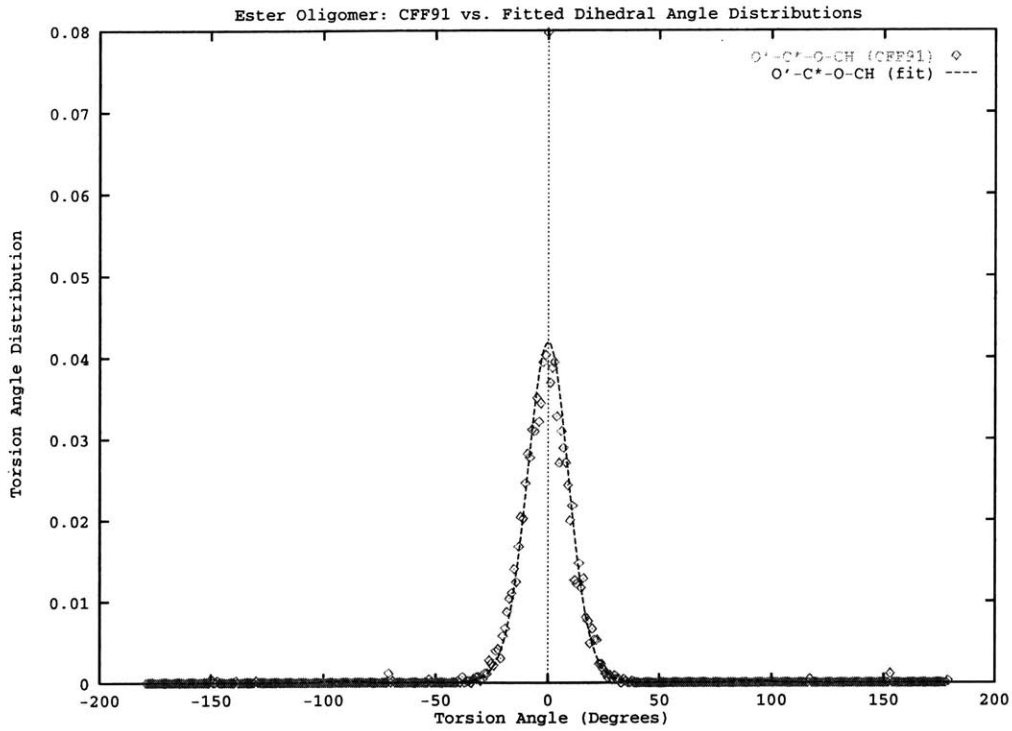


Figure 6-14: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

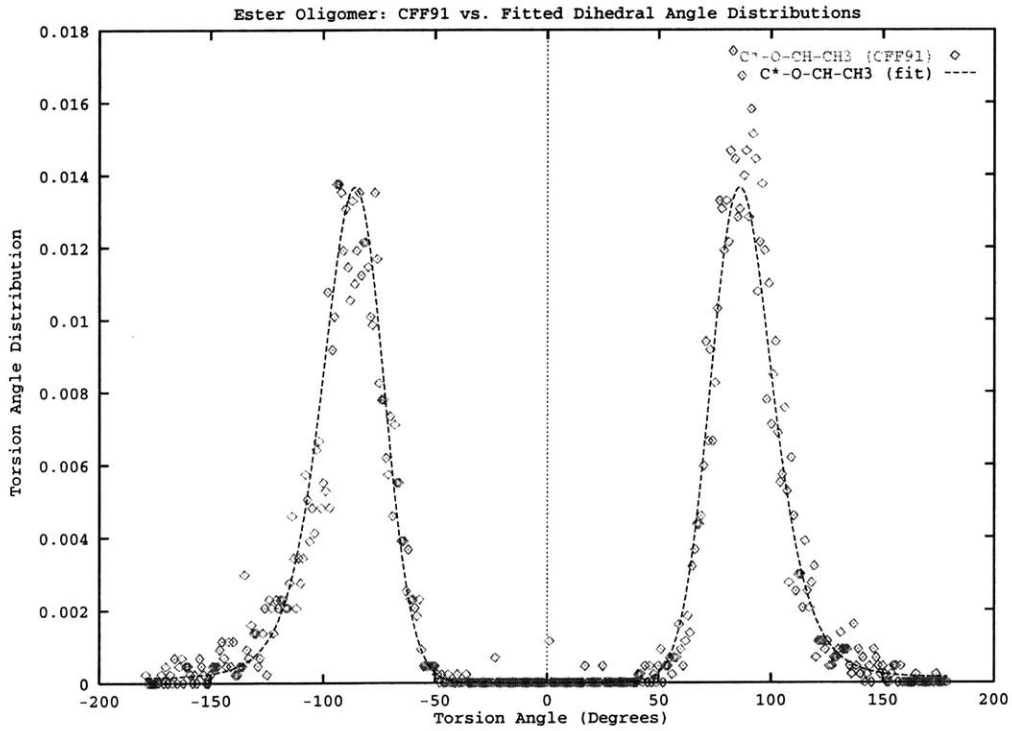


Figure 6-15: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

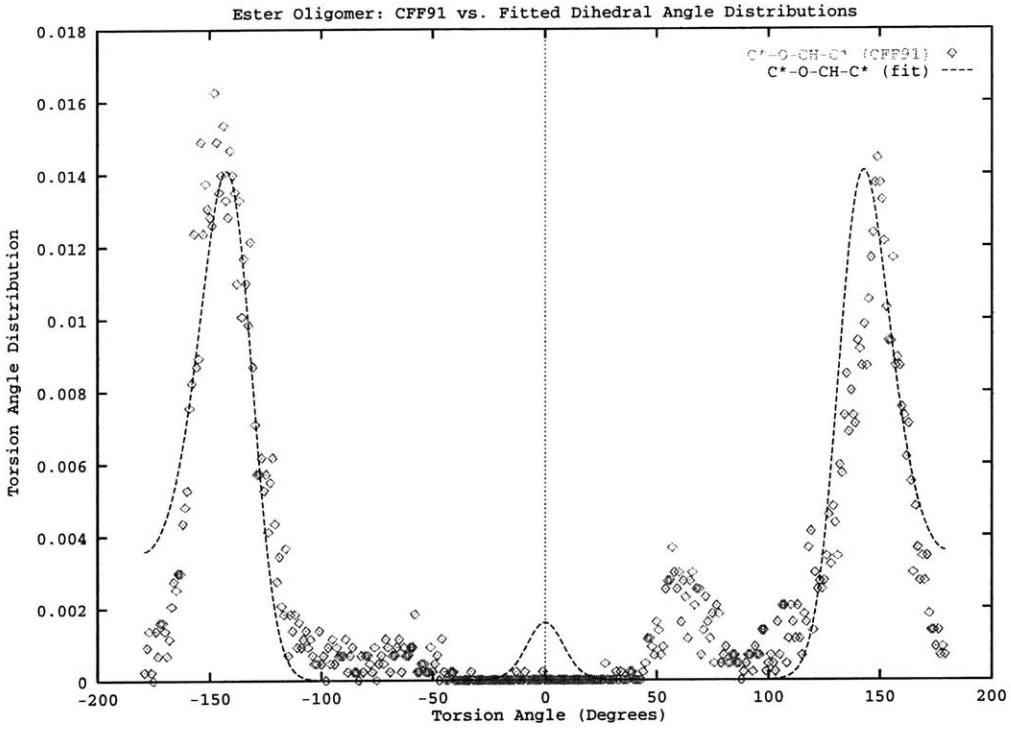


Figure 6-16: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

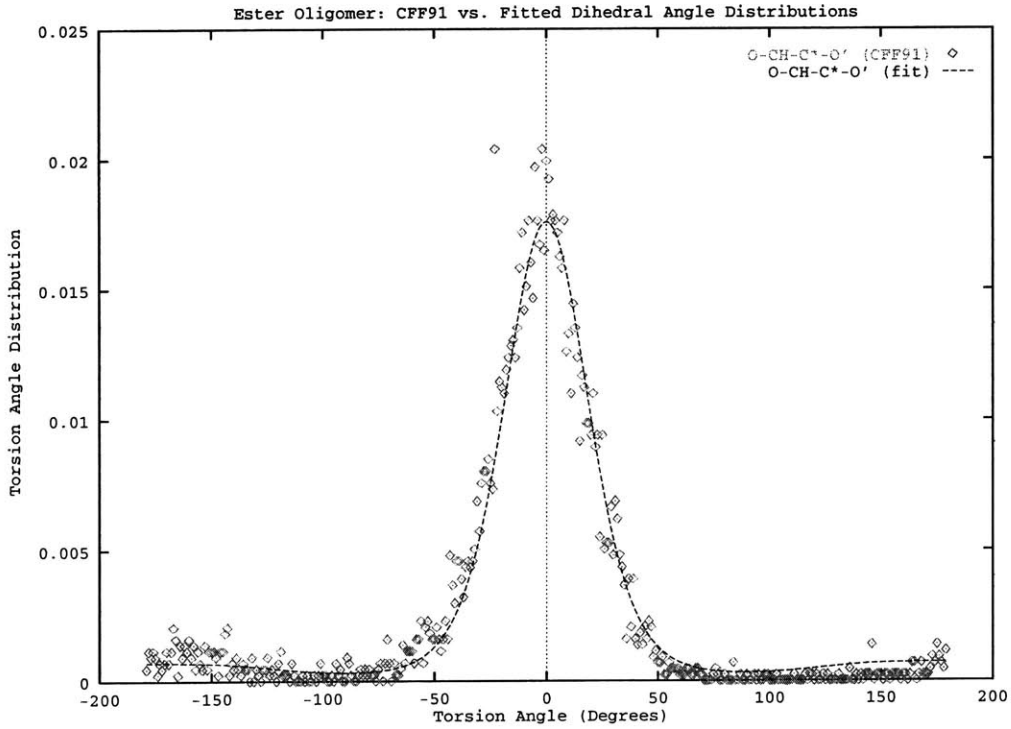


Figure 6-17: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

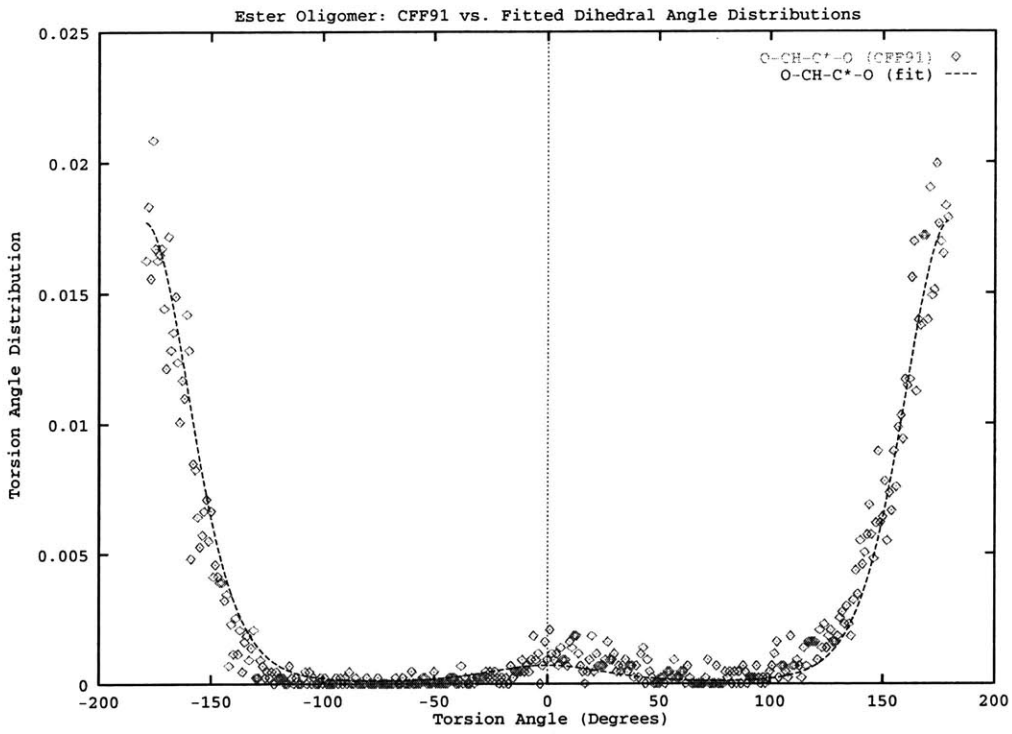


Figure 6-18: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

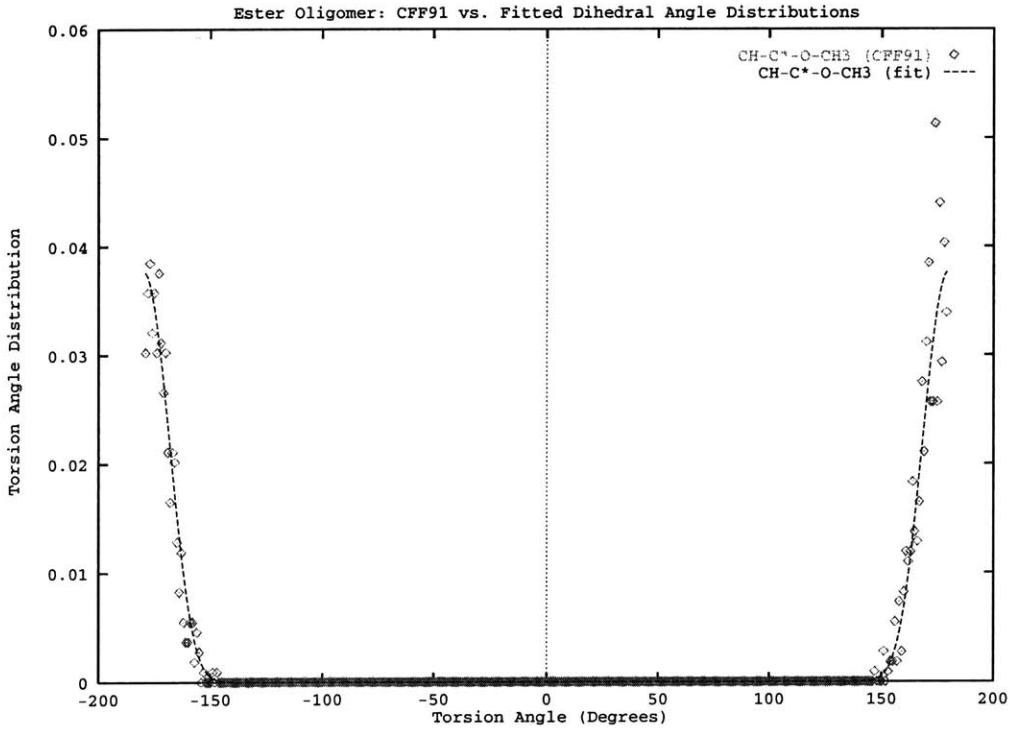


Figure 6-19: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

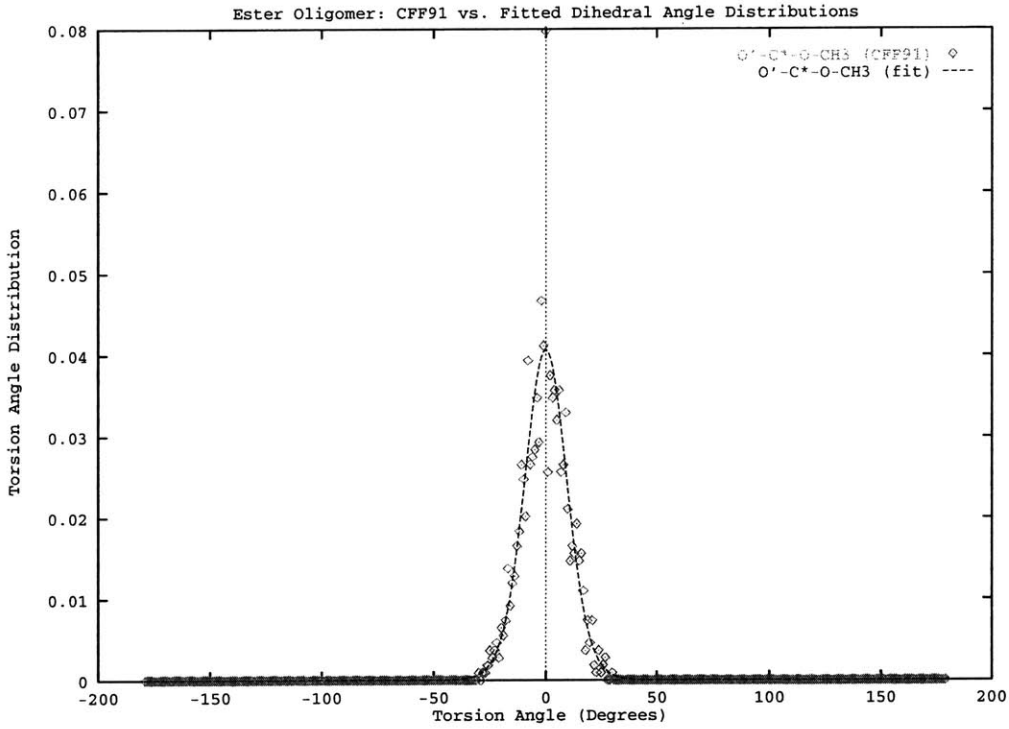


Figure 6-20: CFF91 all-atom torsion angle distribution (points) and fitted function (dashed line)

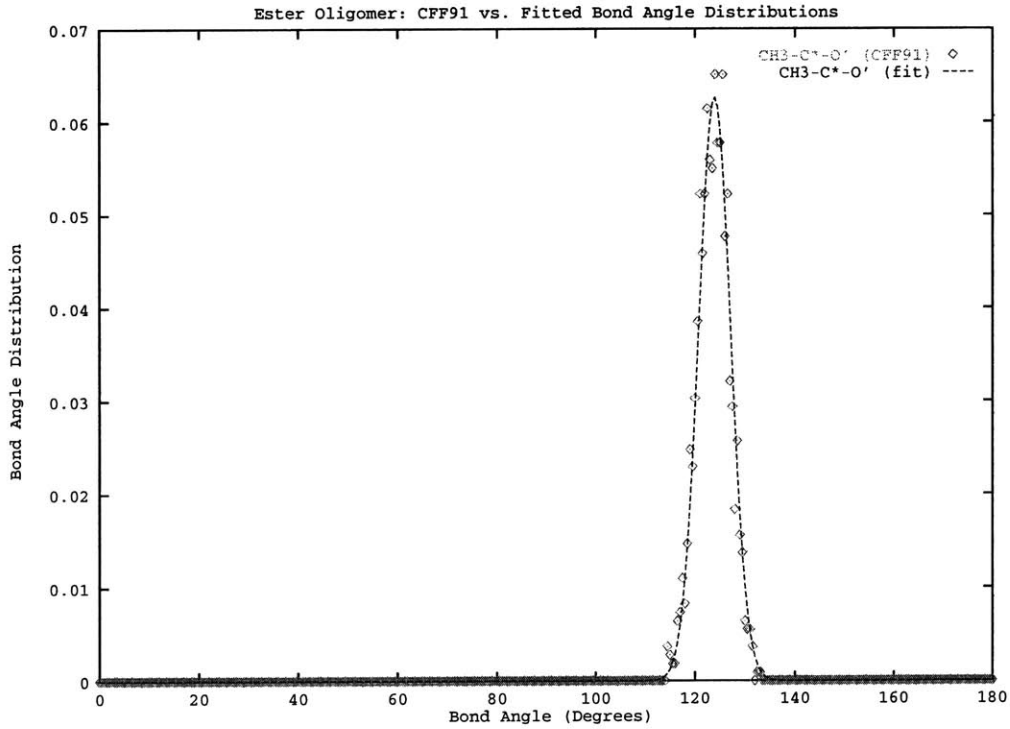


Figure 6-21: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

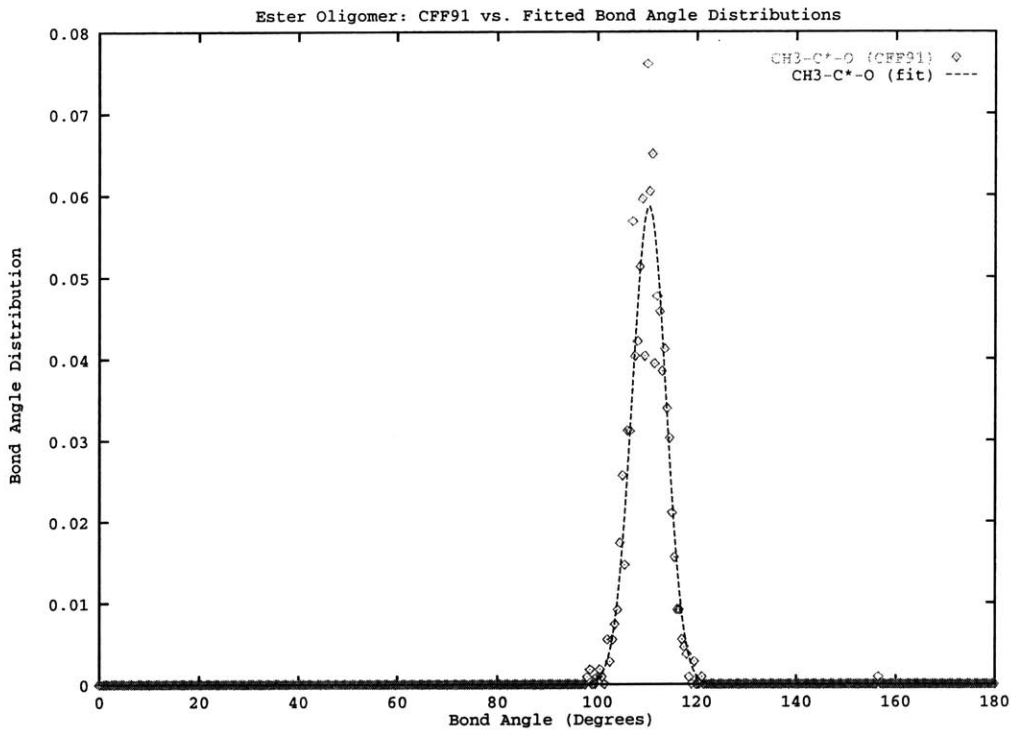


Figure 6-22: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

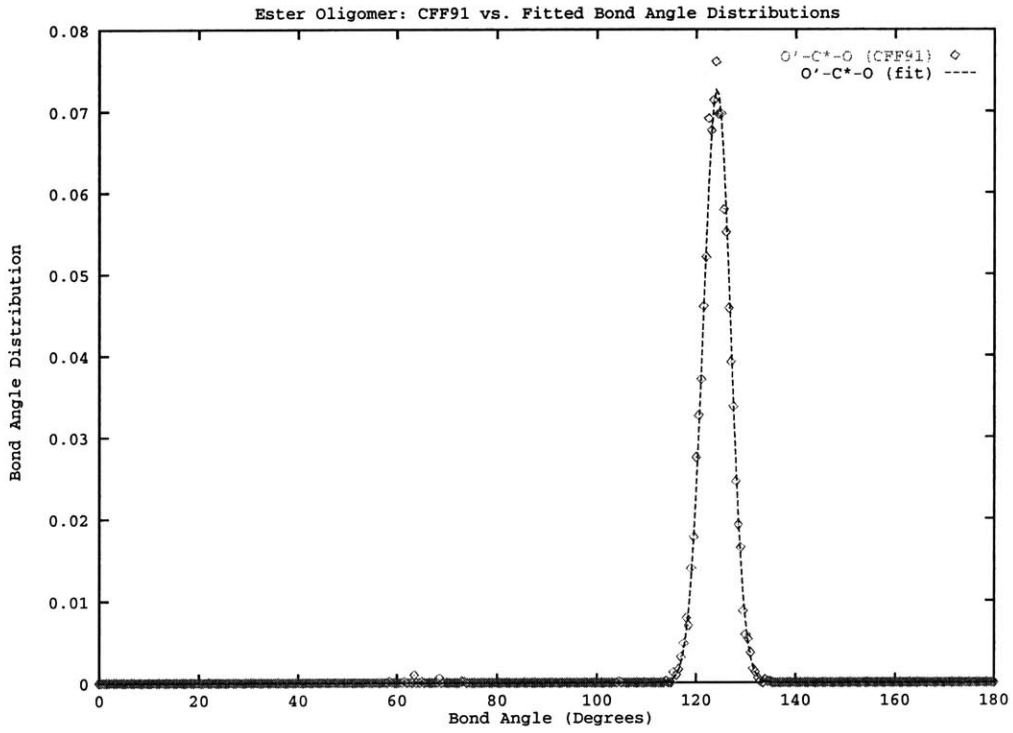


Figure 6-23: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

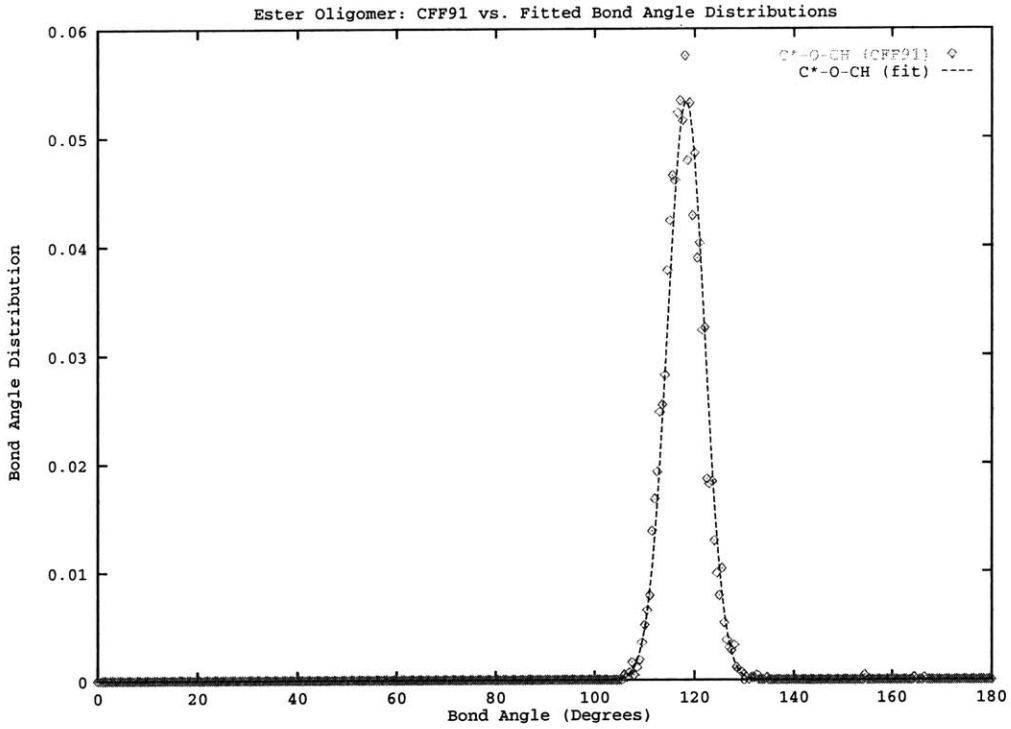


Figure 6-24: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

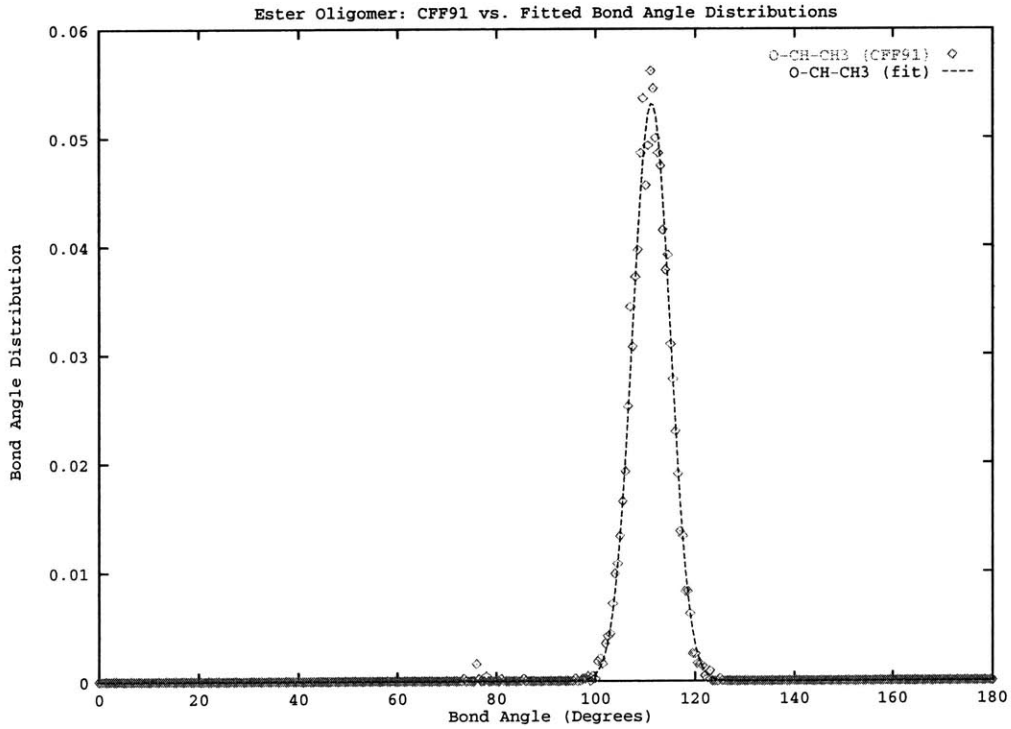


Figure 6-25: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

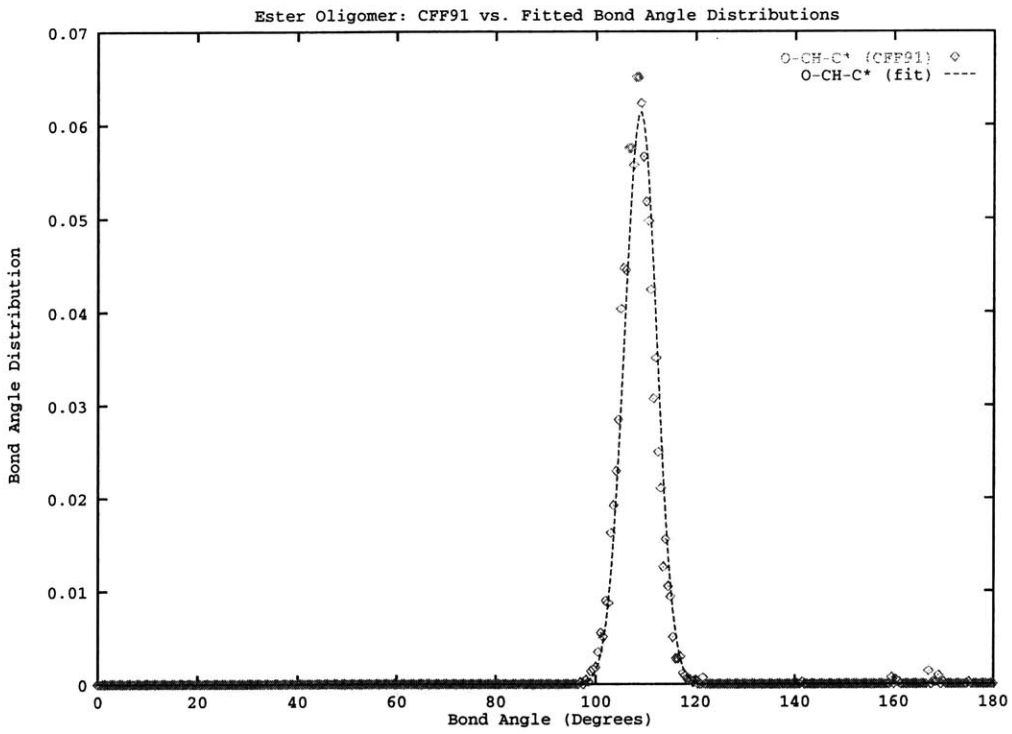


Figure 6-26: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

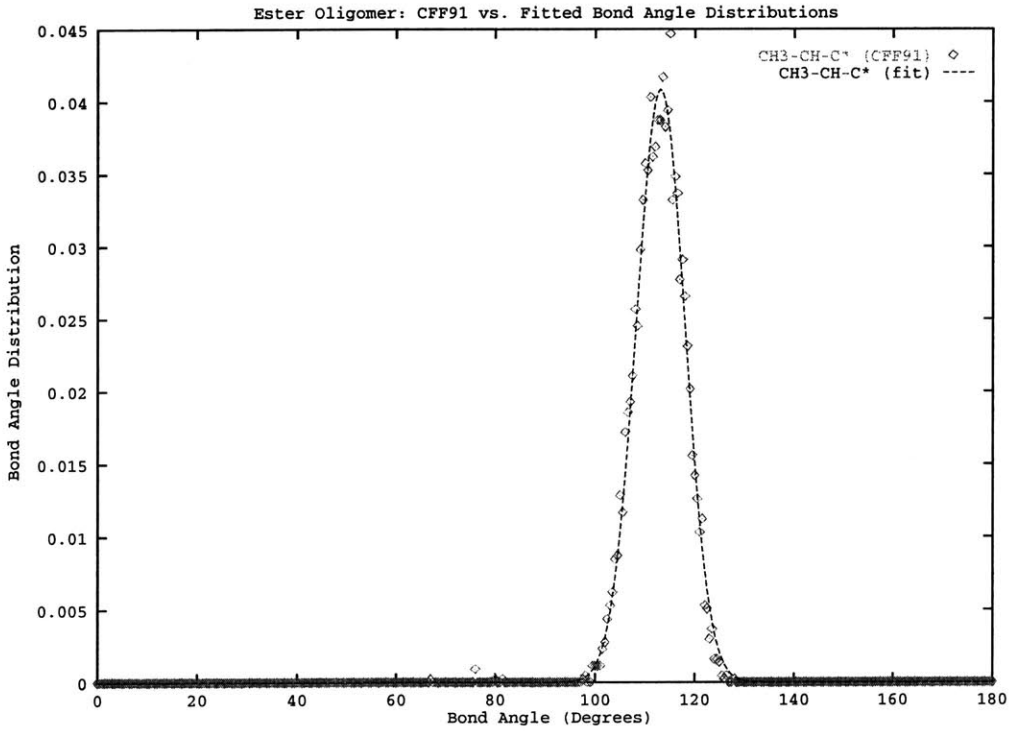


Figure 6-27: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

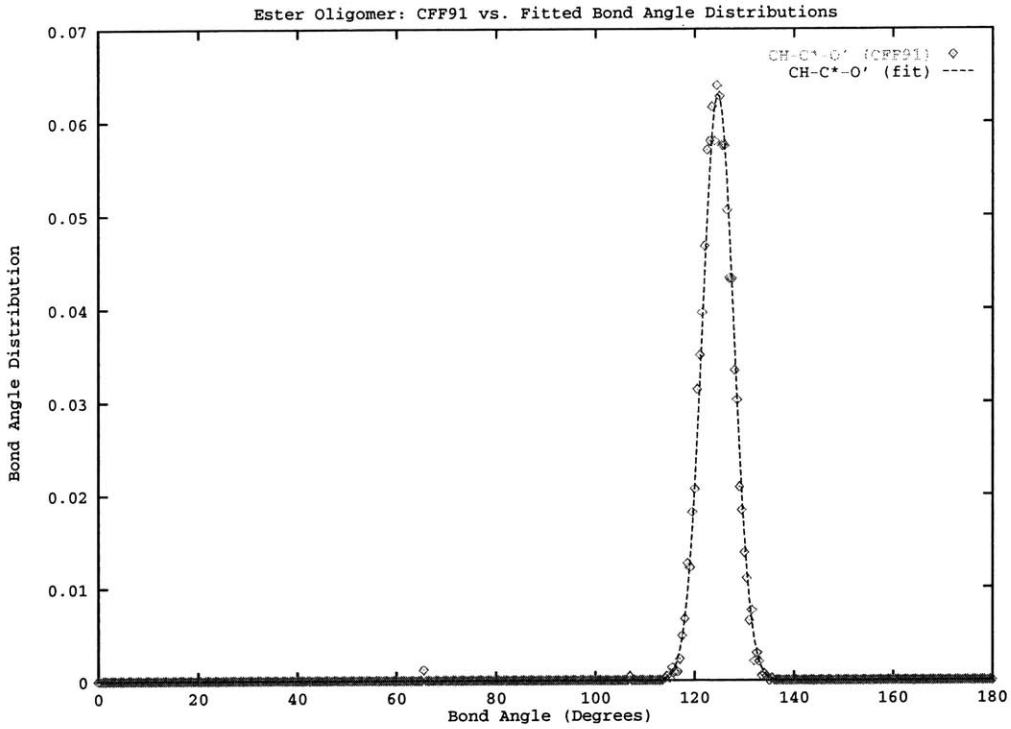


Figure 6-28: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

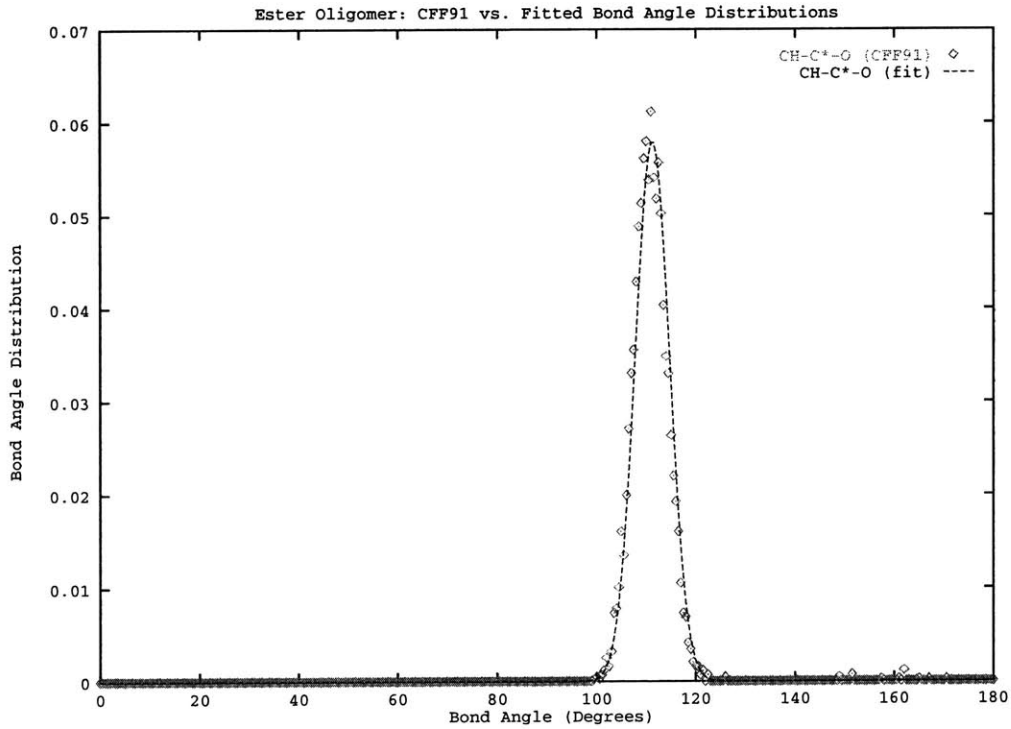


Figure 6-29: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

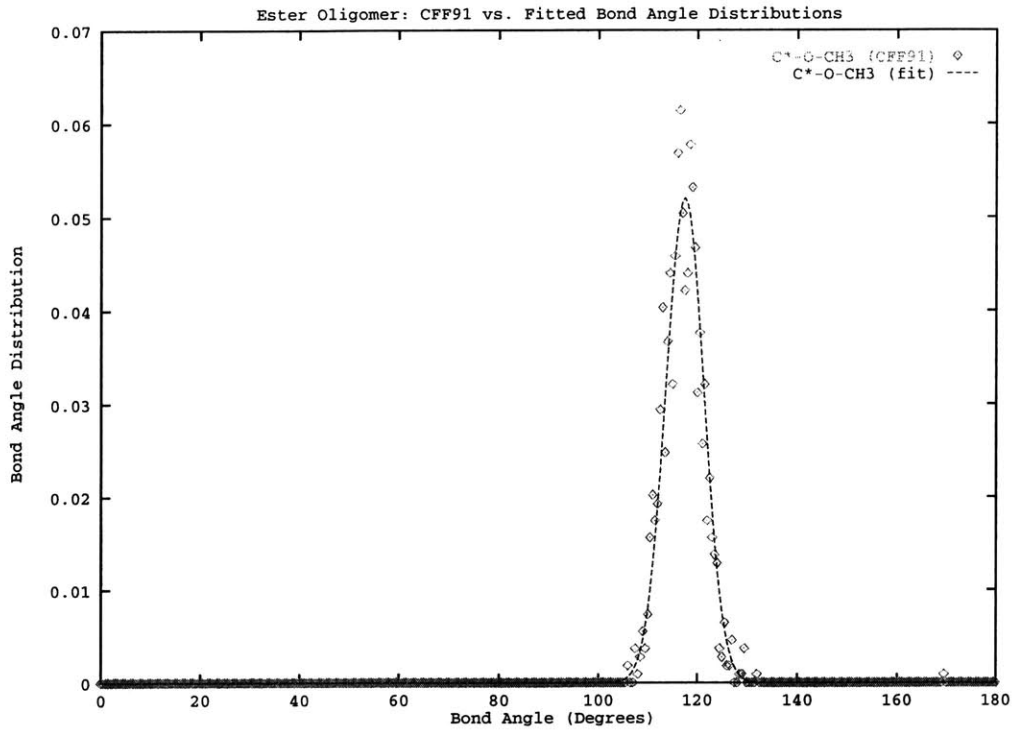


Figure 6-30: CFF91 all-atom bond angle distribution (points) and fitted function (dashed line)

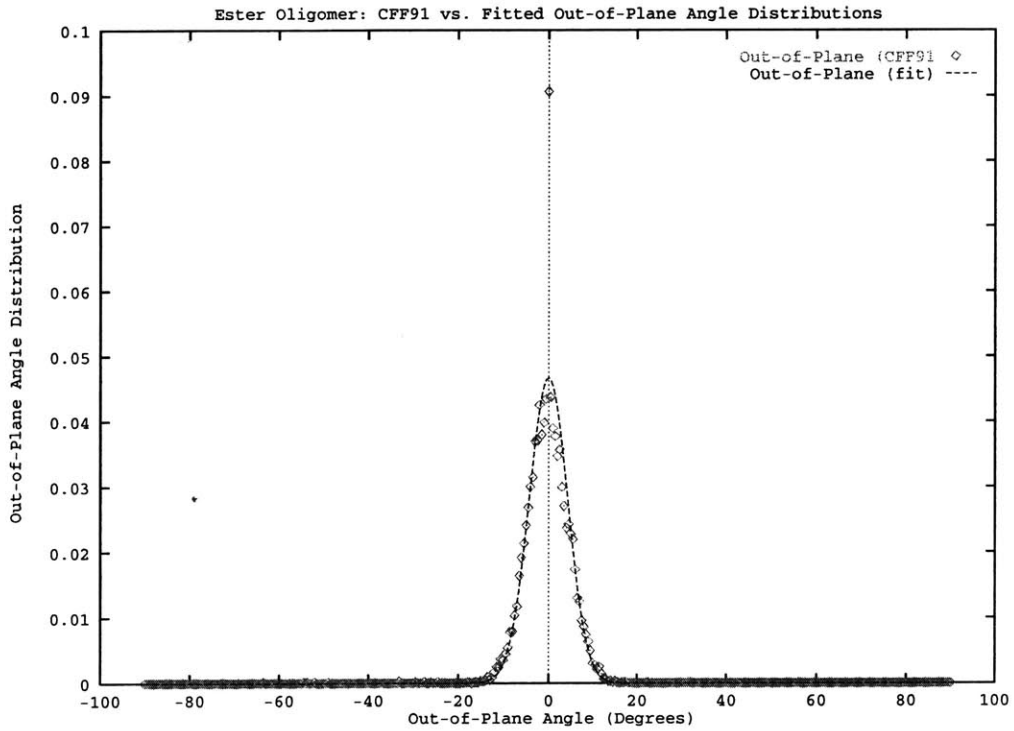


Figure 6-31: CFF91 all-atom out-of-plane angle distribution (points) and fitted function (dashed line)

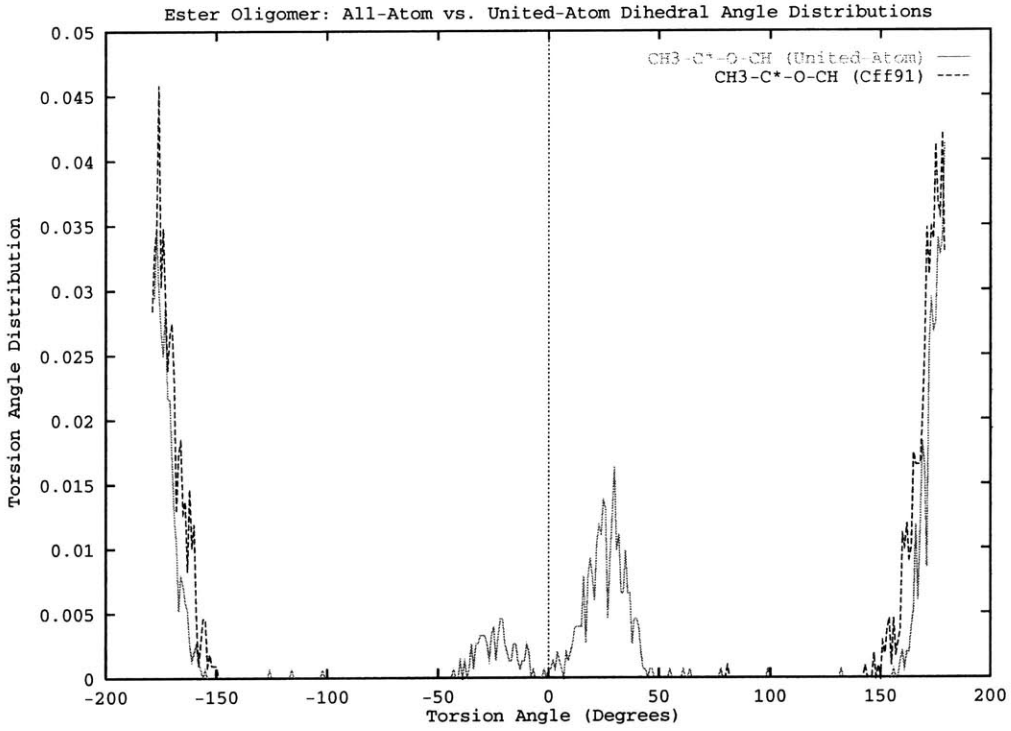


Figure 6-32: CFF91 all-atom versus fitted united-atom torsion angle distributions

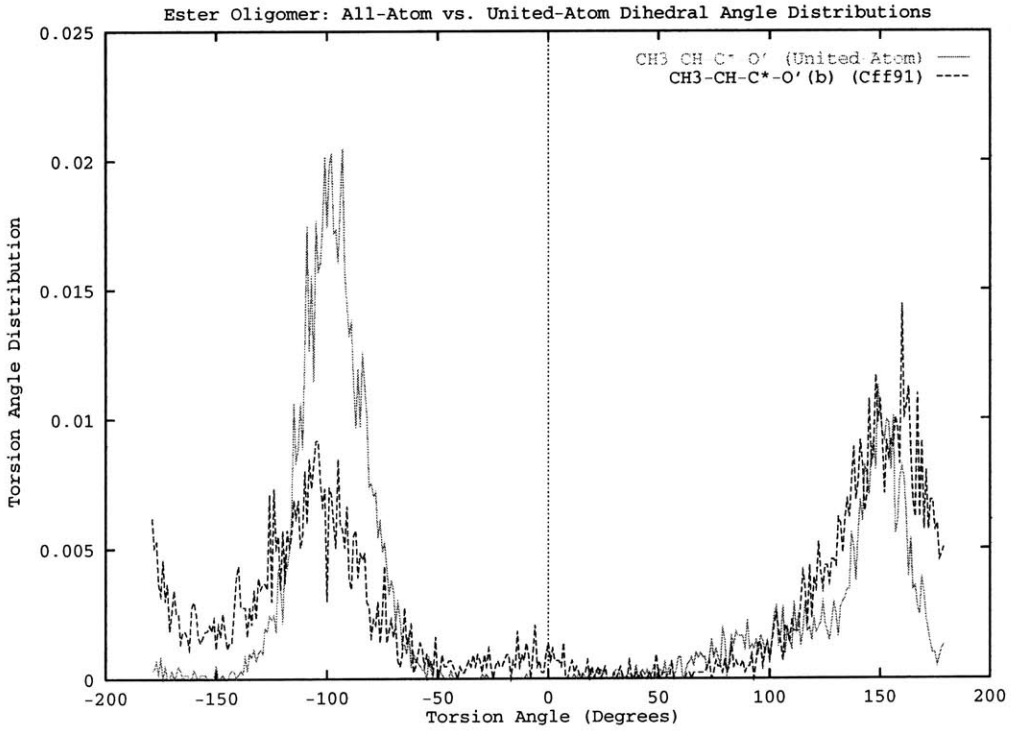


Figure 6-33: CFF91 all-atom versus fitted united-atom torsion angle distributions

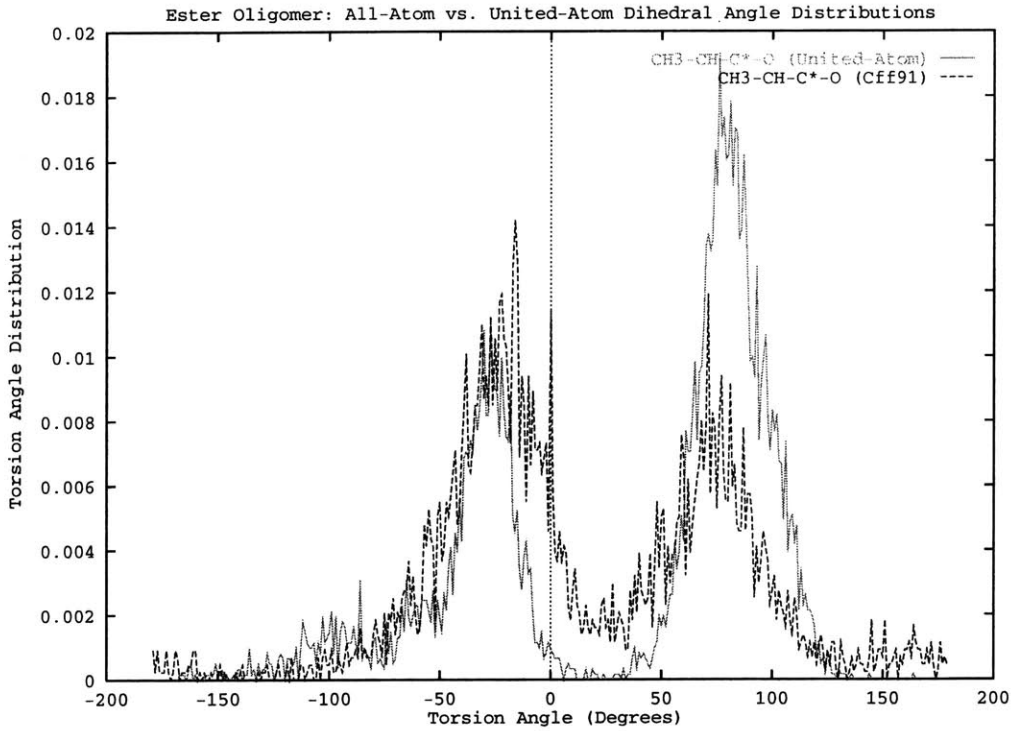


Figure 6-34: CFF91 all-atom versus fitted united-atom torsion angle distributions

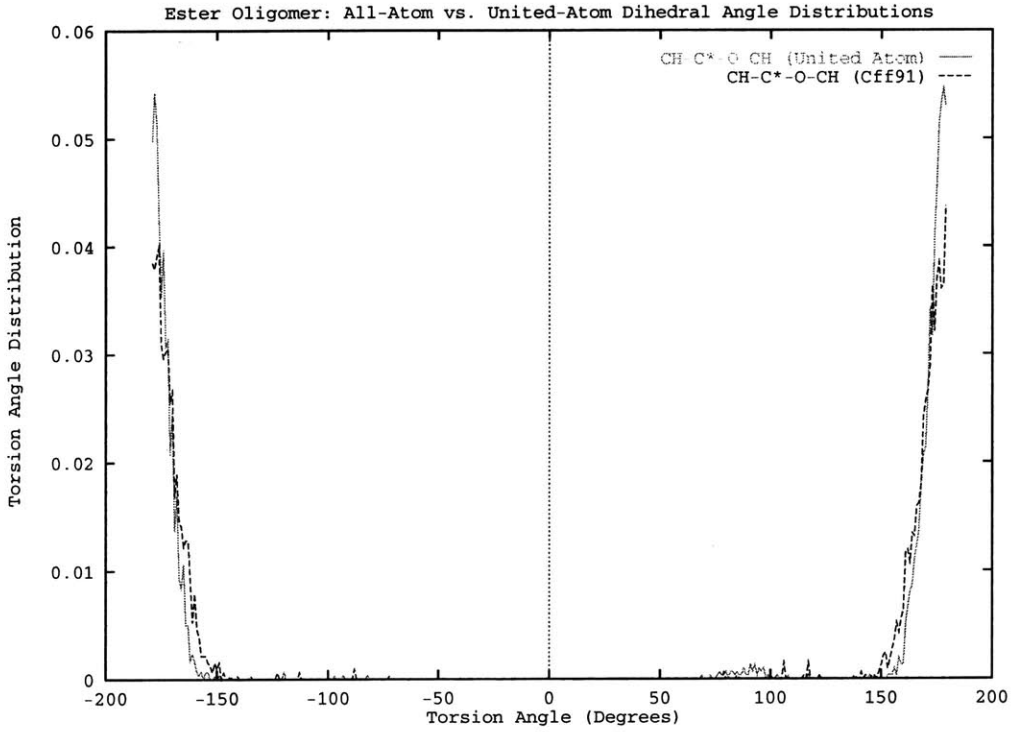


Figure 6-35: CFF91 all-atom versus fitted united-atom torsion angle distributions

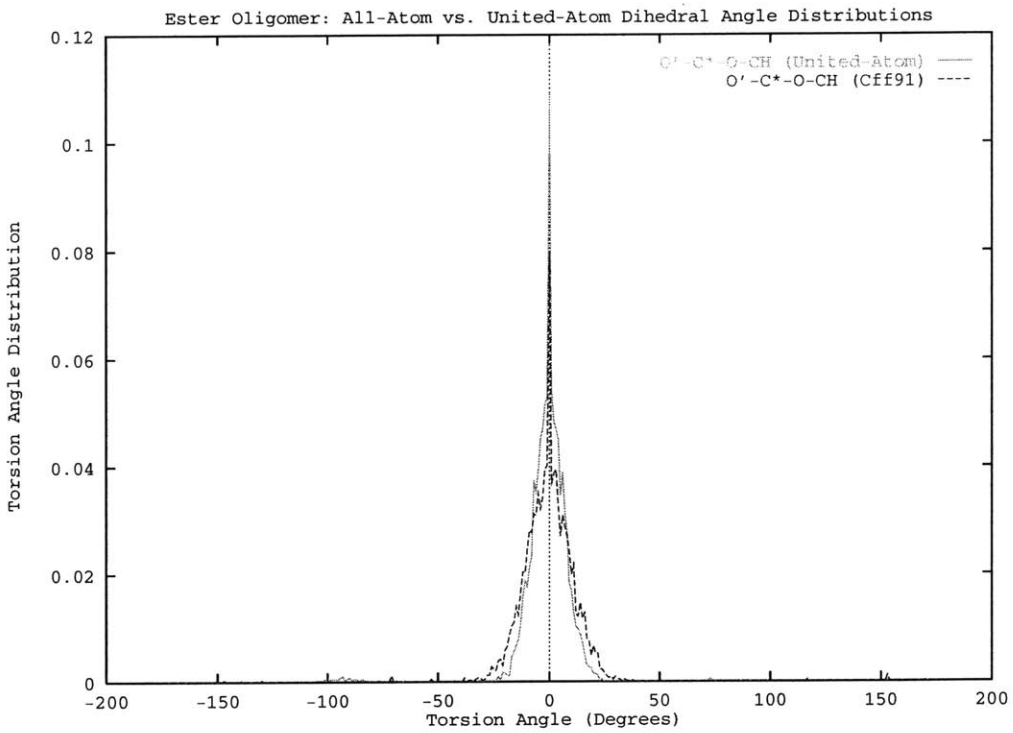


Figure 6-36: CFF91 all-atom versus fitted united-atom torsion angle distributions

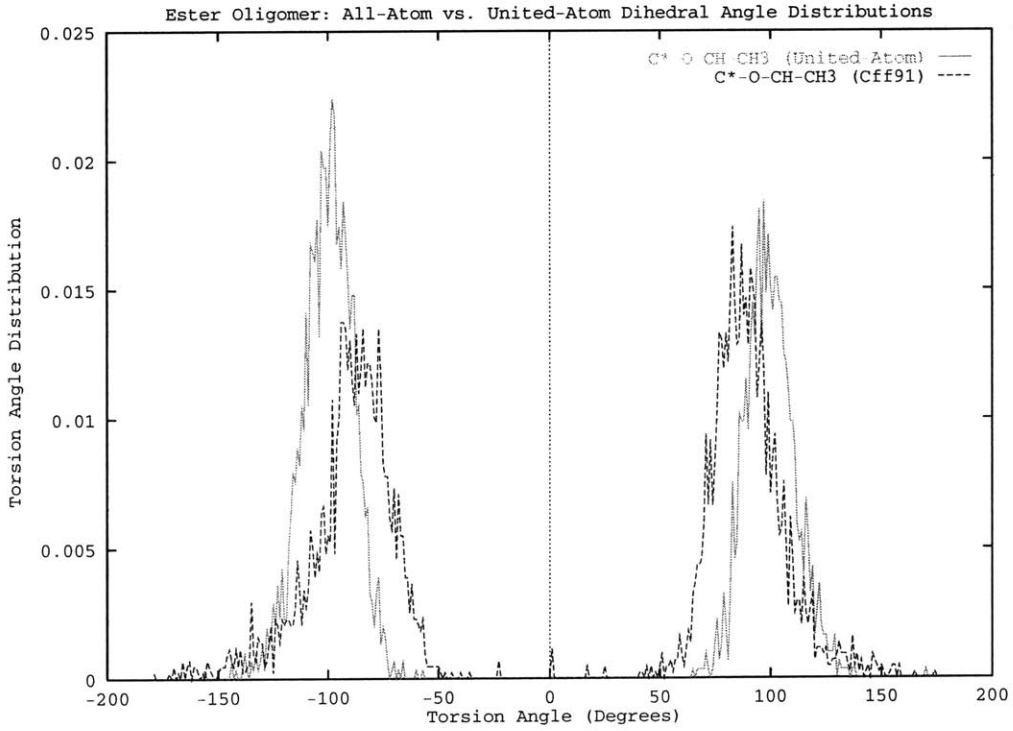


Figure 6-37: CFF91 all-atom versus fitted united-atom torsion angle distributions

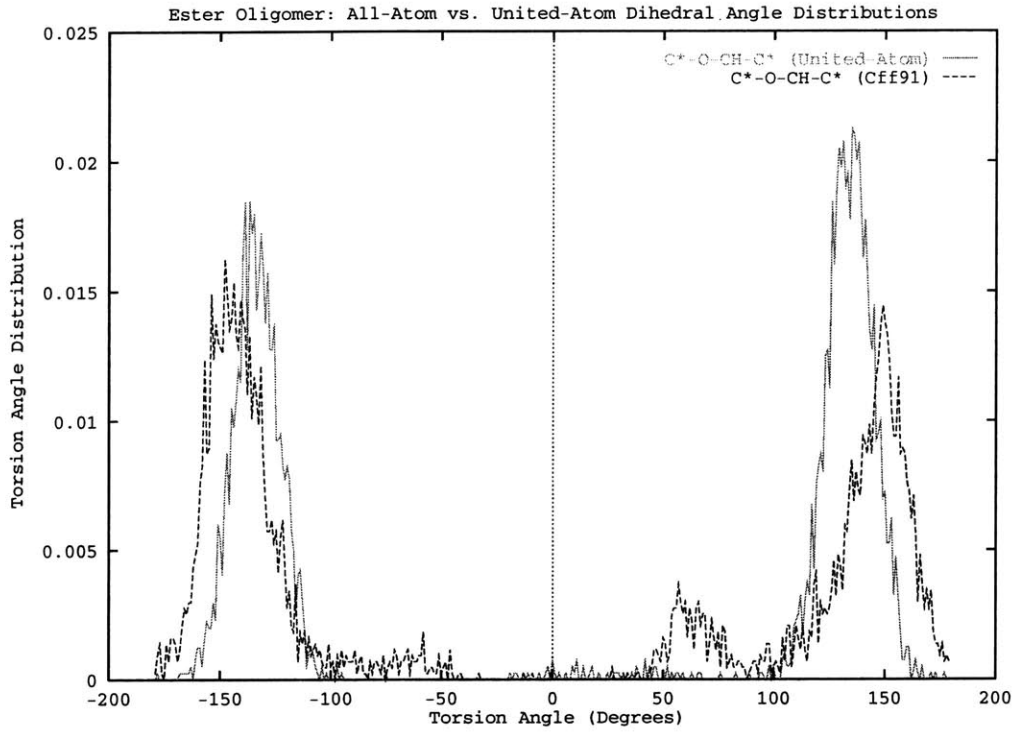


Figure 6-38: CFF91 all-atom versus fitted united-atom torsion angle distributions

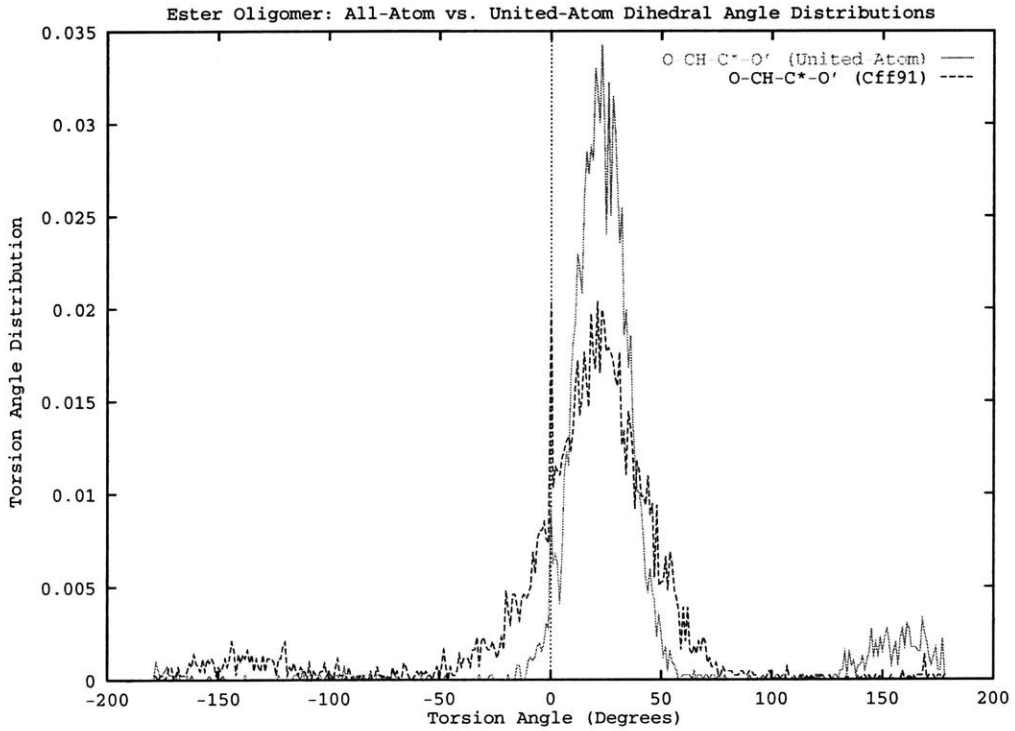


Figure 6-39: CFF91 all-atom versus fitted united-atom torsion angle distributions

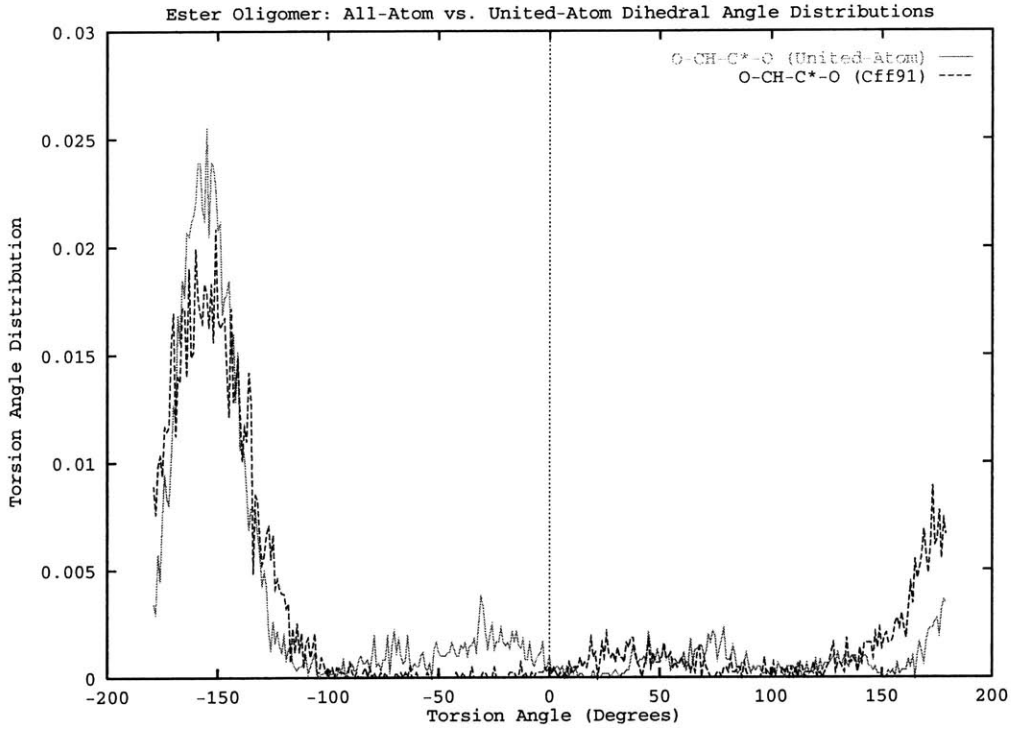


Figure 6-40: CFF91 all-atom versus fitted united-atom torsion angle distributions

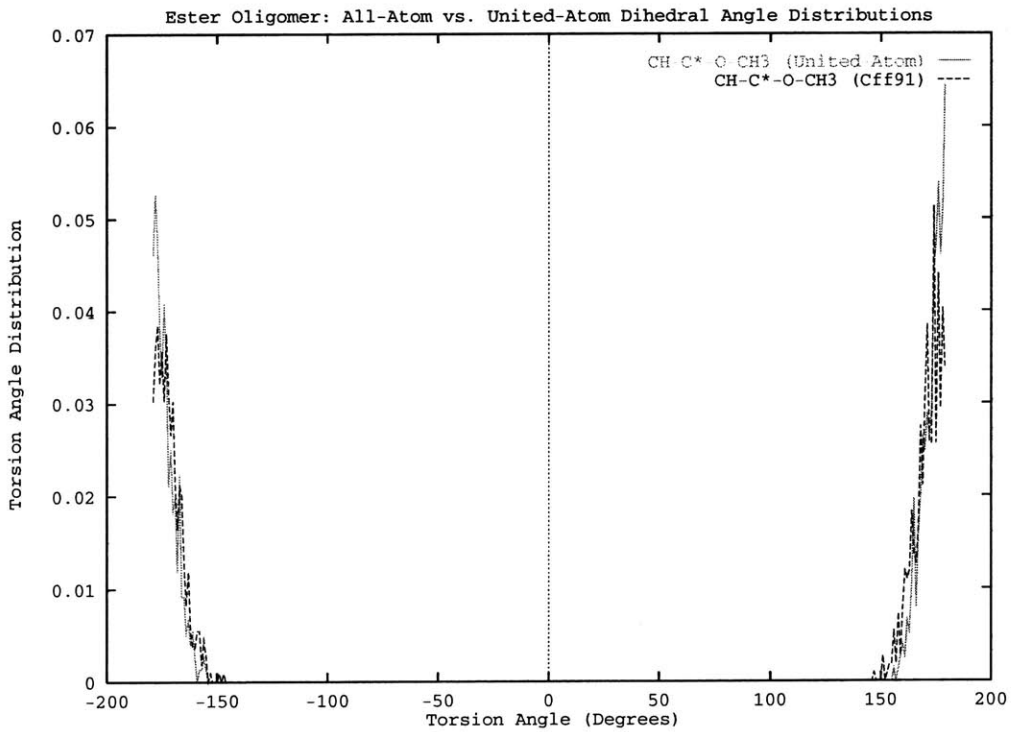


Figure 6-41: CFF91 all-atom versus fitted united-atom torsion angle distributions

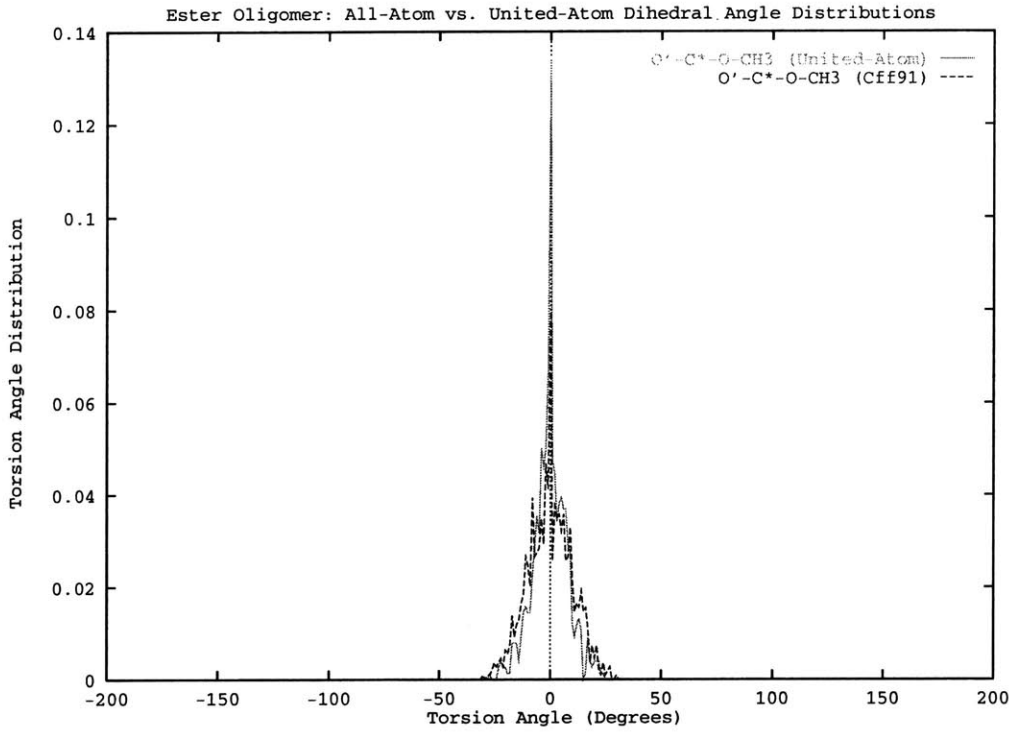


Figure 6-42: CFF91 all-atom versus fitted united-atom torsion angle distributions

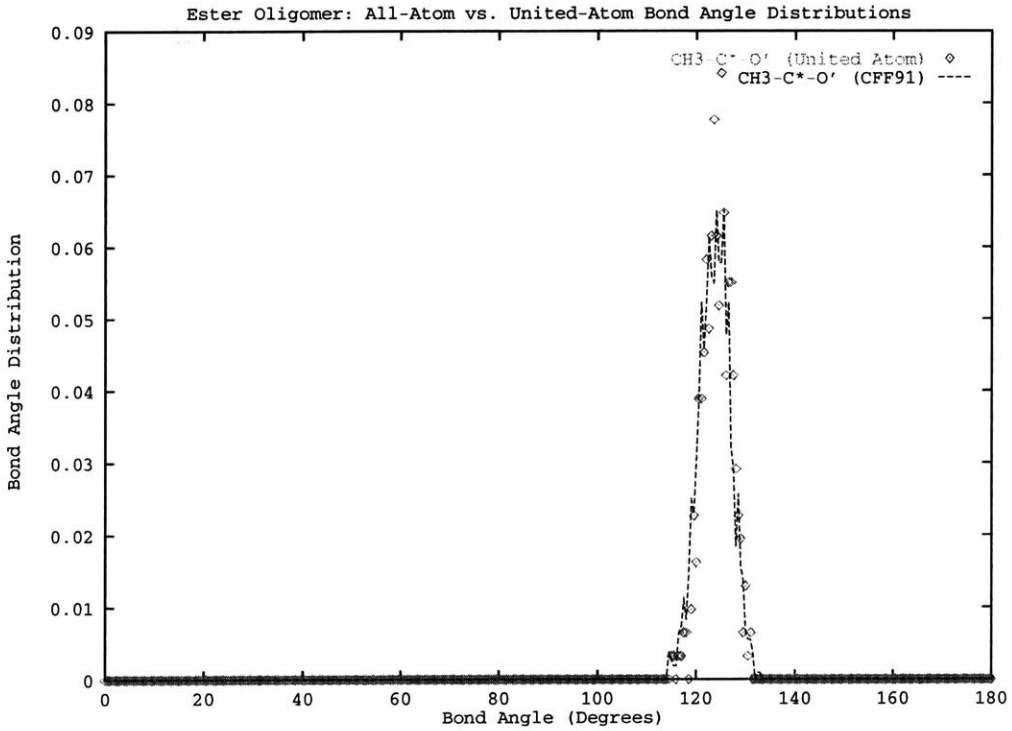


Figure 6-43: CFF91 all-atom versus fitted united-atom bond angle distributions

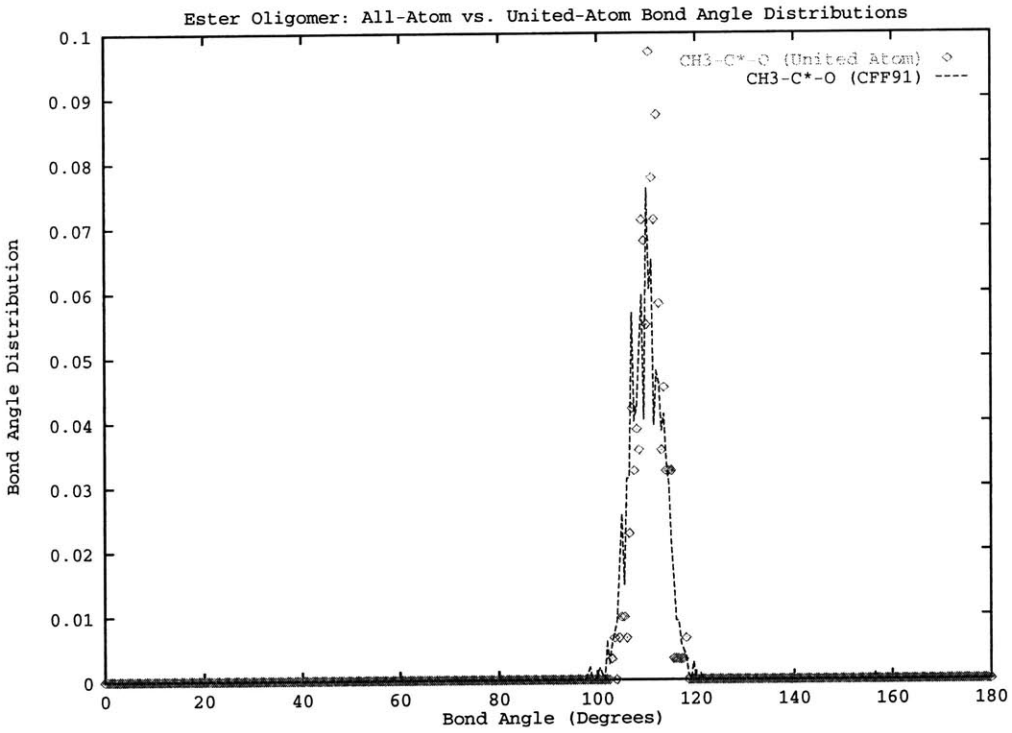


Figure 6-44: CFF91 all-atom versus fitted united-atom bond angle distributions

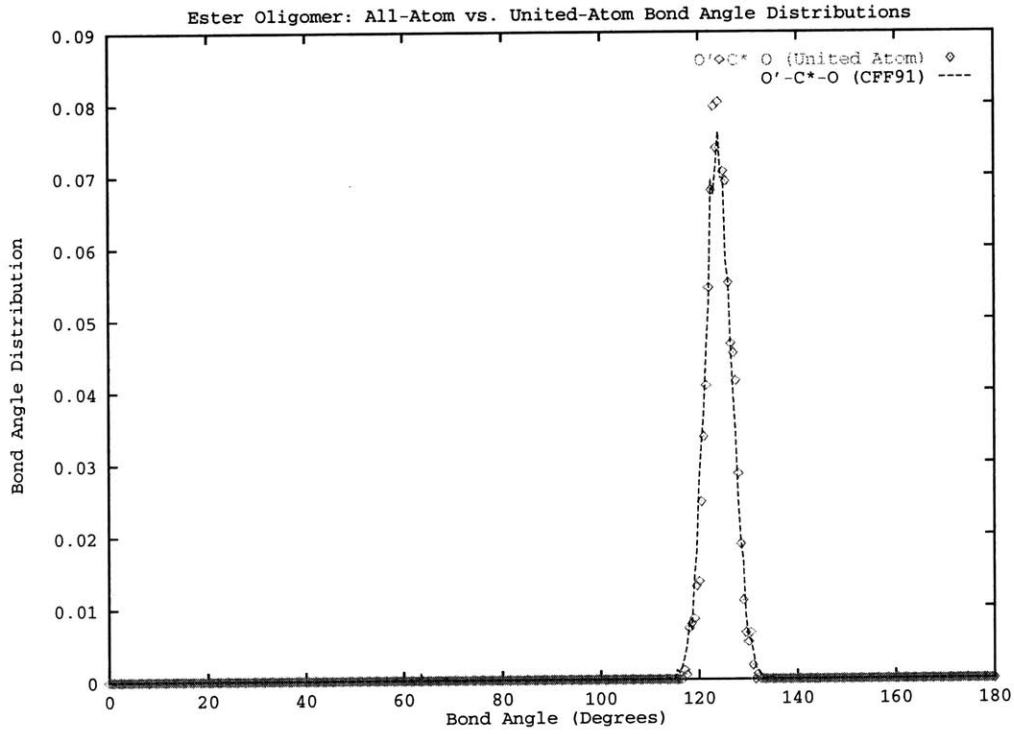


Figure 6-45: CFF91 all-atom versus fitted united-atom bond angle distributions

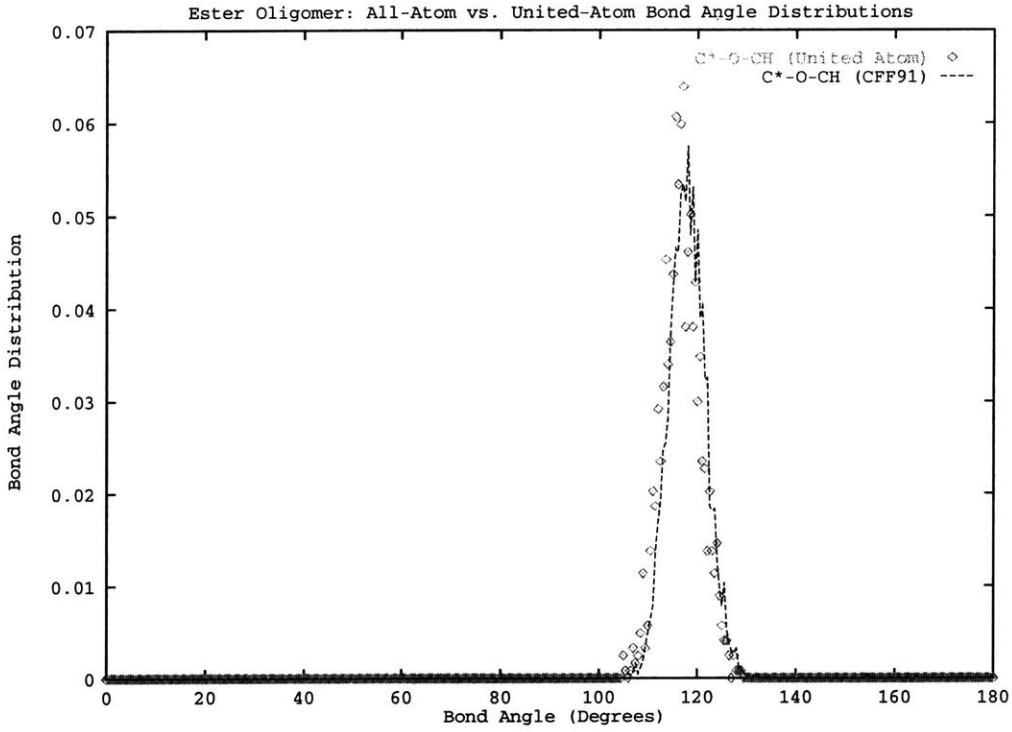


Figure 6-46: CFF91 all-atom versus fitted united-atom bond angle distributions

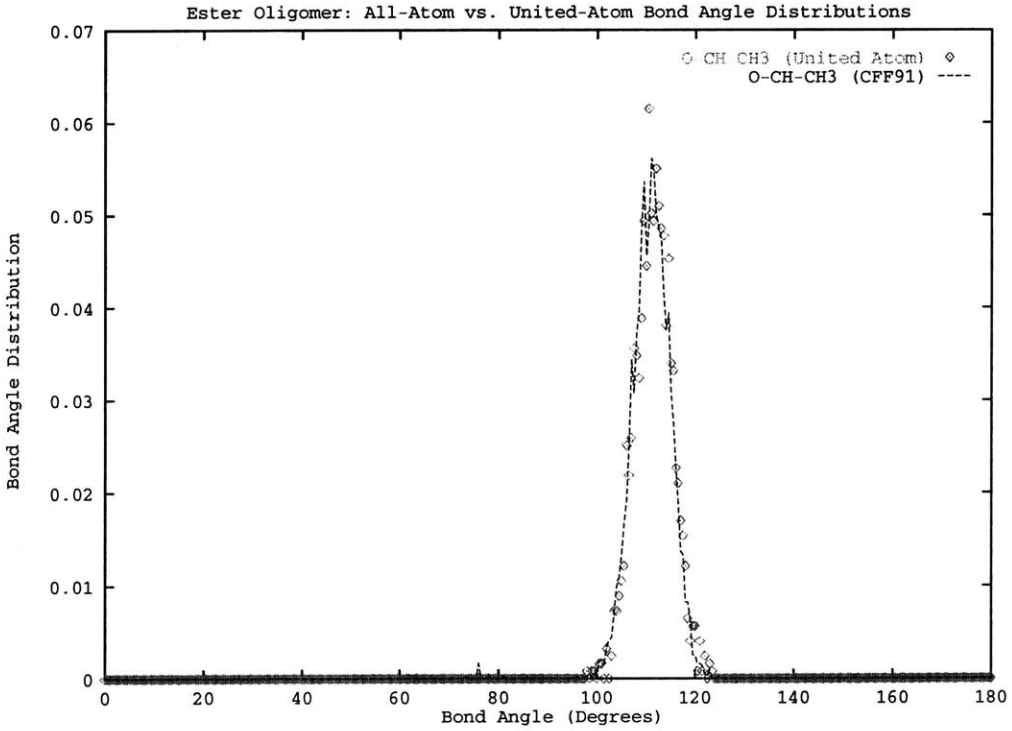


Figure 6-47: CFF91 all-atom versus fitted united-atom bond angle distributions

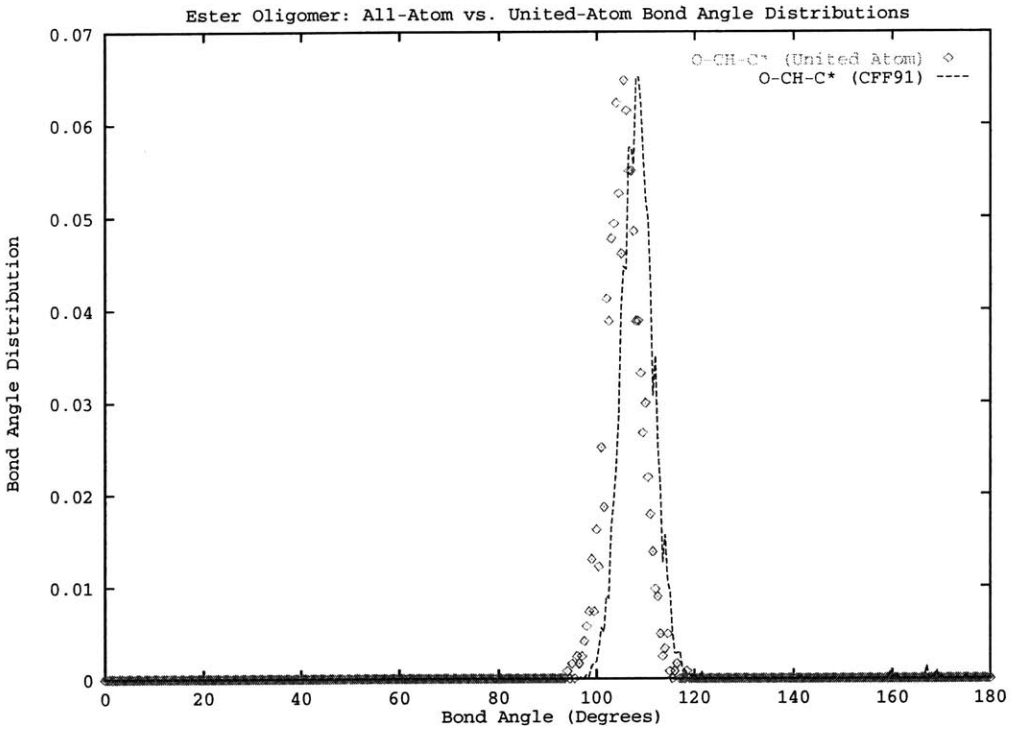


Figure 6-48: CFF91 all-atom versus fitted united-atom bond angle distributions

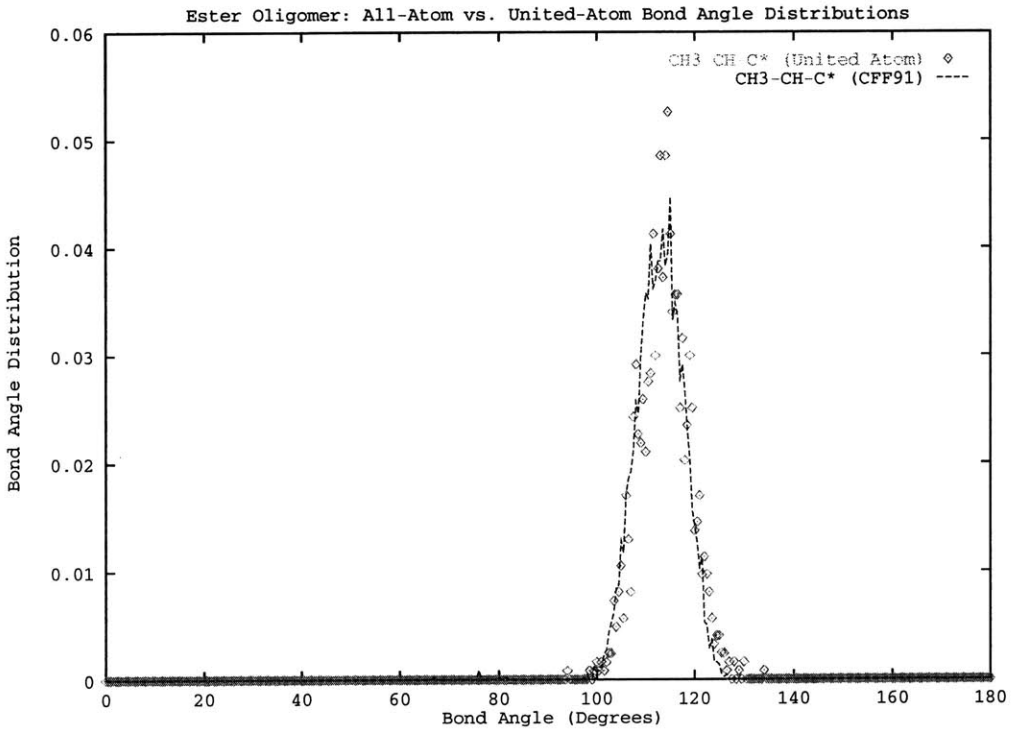


Figure 6-49: CFF91 all-atom versus fitted united-atom bond angle distributions

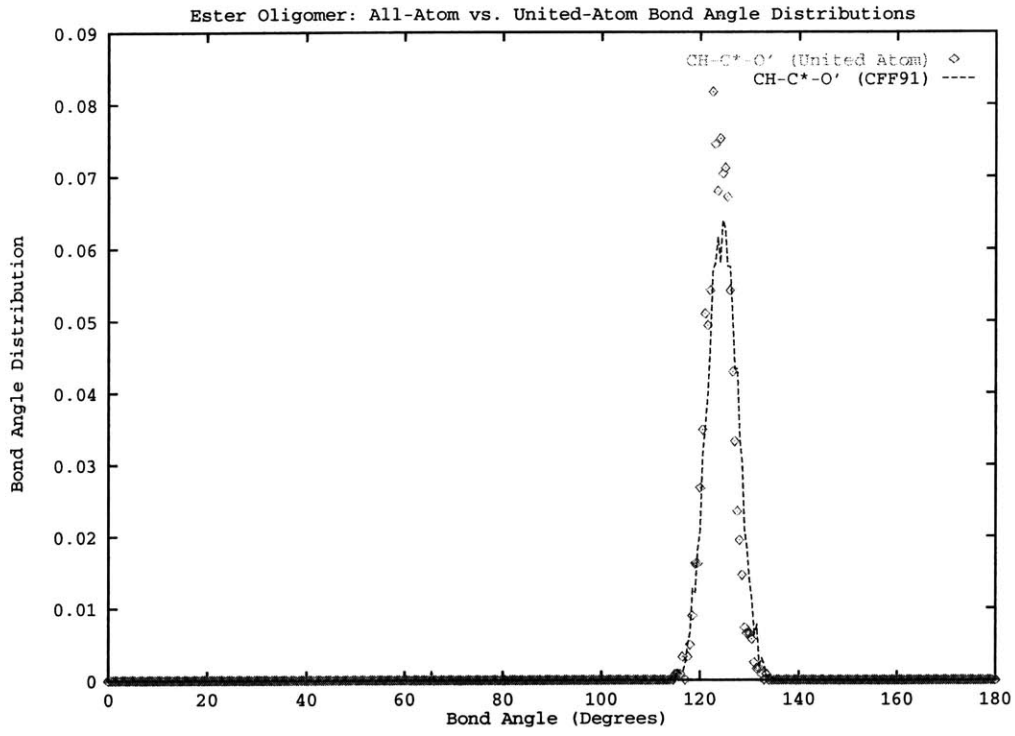


Figure 6-50: CFF91 all-atom versus fitted united-atom bond angle distributions

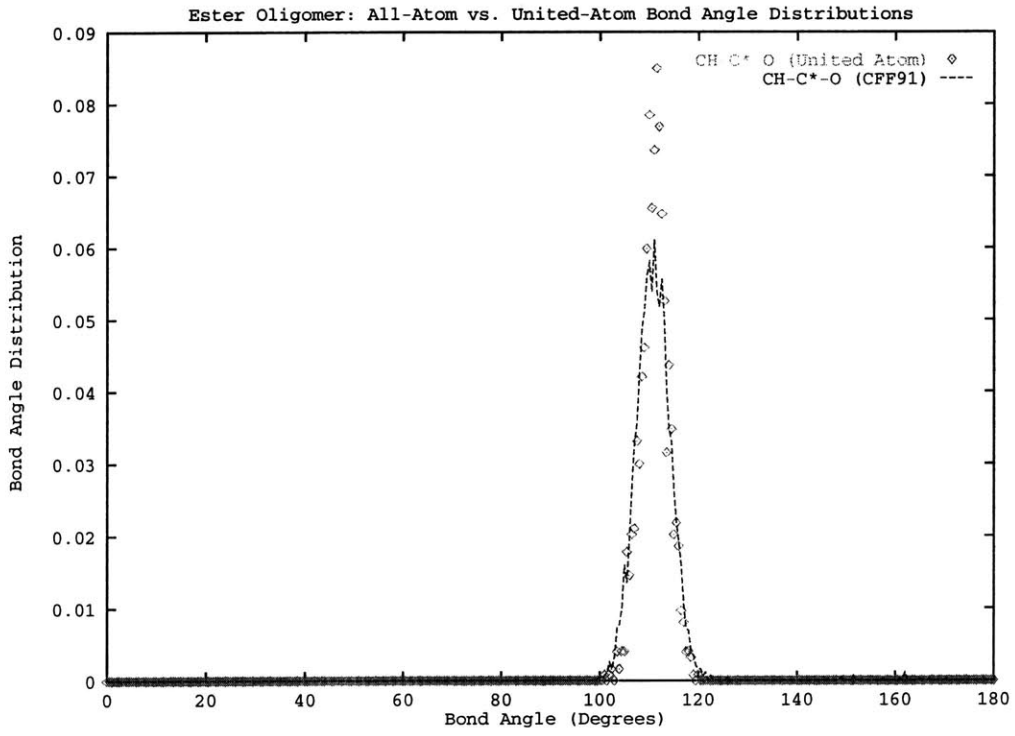


Figure 6-51: CFF91 all-atom versus fitted united-atom bond angle distributions

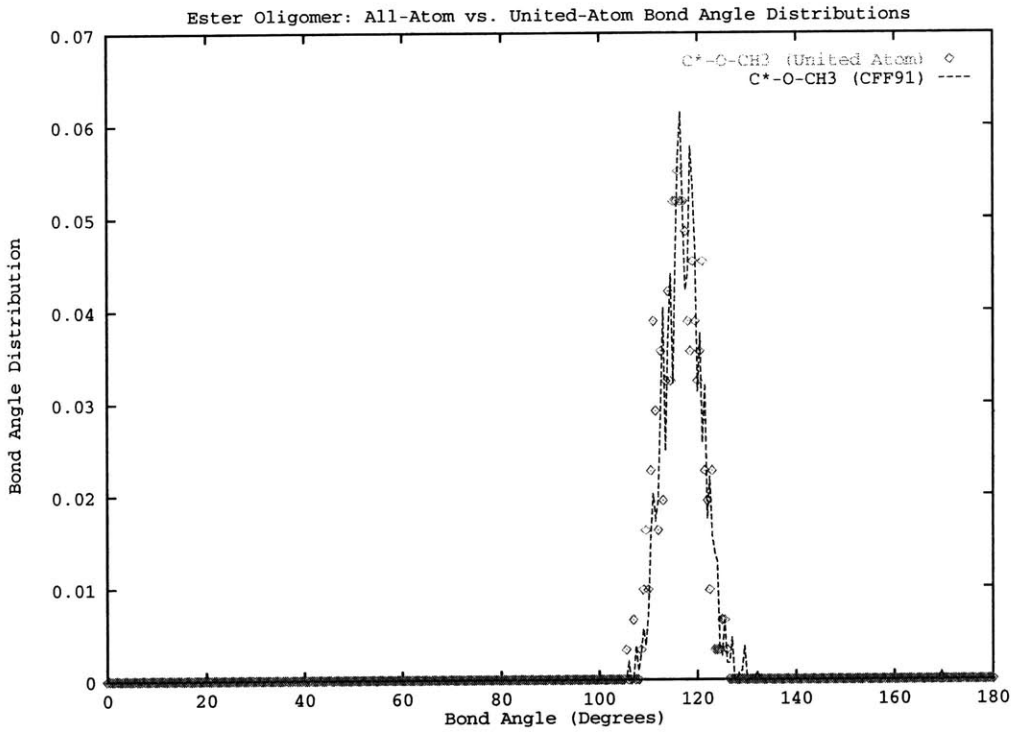


Figure 6-52: CFF91 all-atom versus fitted united-atom bond angle distributions

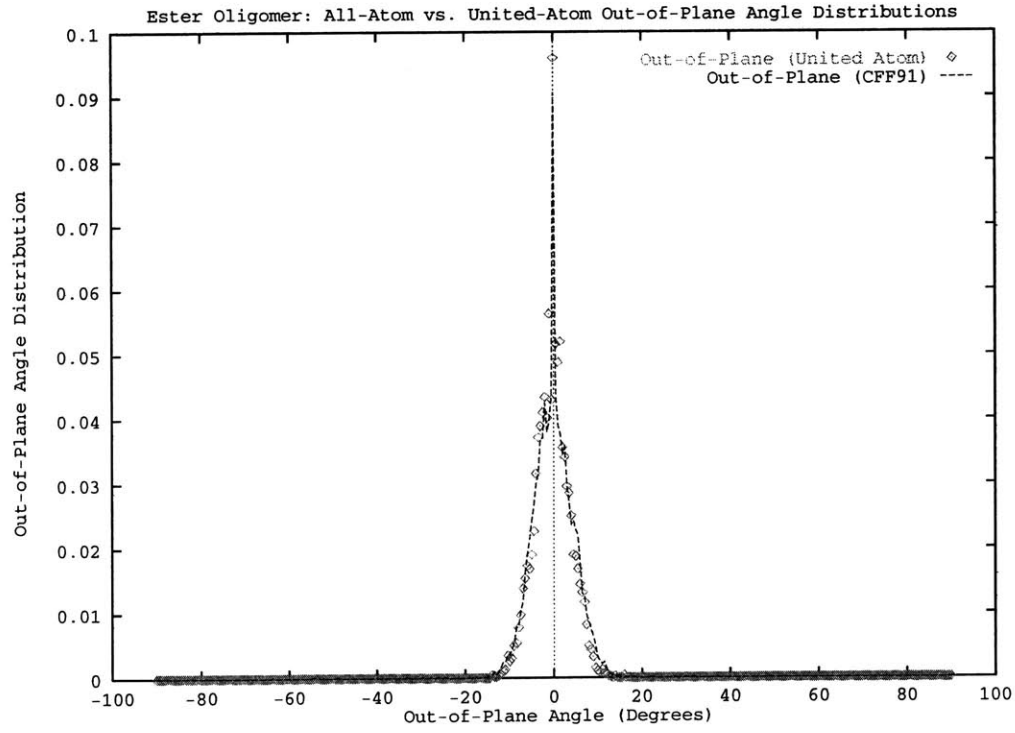


Figure 6-53: CFF91 all-atom versus fitted united-atom out-of-plane angle distributions

Chapter 7

Bulk and Interfacial Aspects of PLA Oligomers via Molecular Dynamics Simulation

7.1 Introduction and Objectives

Numerous oligomeric species have been studied via molecular dynamics simulations with respect to their bulk and interfacial properties. These systems allow one to probe certain phenomena within the time constraints imposed by the molecular dynamics method while still maintaining the essence of the polymeric species in terms of important functional groups. Thus, one can investigate the effect of connectivity and long-chain character on very local molecular properties such as structure, orientation, dynamics, as well as intramolecular phenomena such as group relaxations. Furthermore, in the particular case of interfacial properties, oligomeric models have been shown to capture polymeric physics such as surface segregation of chain ends from computational [60] and theoretical [149, 150, 60] approaches.

This chapter addresses the bulk and interfacial properties of a PLA ester oligomer via molecular dynamics simulations using the forcefield model described in Chapter 6. The aims of the current study are as follows. First, having established a set

of intermolecular and intramolecular potential parameters, we will determine the P-V-T behavior of the model; in particular, we are concerned with the density of the model at ambient conditions of 1 atmosphere pressure and temperature of 298K. This will entail constant-pressure, constant-temperature molecular dynamics simulations to determine the equilibrium density. From this data, further canonical ensemble simulations for the bulk and interfacial properties will be performed.

Second, for the bulk PLA oligomer liquid, several properties will be determined. Radial distribution functions will be computed to determine bulk liquid structure. Further quantities of interest will be bulk radius of gyration and the mean end-to-end distance of chains in the bulk. These can be compared to results of theory, and thus offer insight into the chain configurations in the bulk. An important quantity will be a collective order parameter describing the orientational distribution of the collection of molecules; this will be used to assess the time requirement for the decay of temporal orientational correlation. This will be a useful check for the equilibration of the interfacial simulations.

Finally, the liquid-vapor interface of the PLA oligomer will be analyzed from MD data. Again, as in previous studies of the liquid-vapor interface, we will be concerned with density profiles (molecular, component, chain segment), orientational order parameter profiles (P1 and P2), overall molecular orientations at the bulk compared to the interface, radial distributions in the bulk and interface, and in terms of a thermodynamic property, the surface tension of this model with no corrections will be computed.

7.2 Model and Methods

7.2.1 Molecular Model

The PLA oligomer model is a five-mer of the repeating unit shown in Figure 6-1 with the ends capped with methyl groups. The model is a united-atom ester oligomer model with the CH_3 and CH groups treated as single units (explicit hydrogens are

excluded). With the united-atom approximation, the number of "atoms" per oligomer molecule is drastically reduced from 47 to 25, almost a fifty percent reduction; this is attractive from the force/energy computation aspect of the MD method.

Intermolecular atom pairs and intramolecular pairs beyond torsion quartets interact via a Lennard-Jones potential:

$$\psi_{lj}(r) = 4\epsilon[(\frac{r}{\sigma})^{-12} - (\frac{r}{\sigma})^{-6}] \tag{7.1}$$

where σ is the Lennard-Jones core diameter, ϵ is the well depth, and r is the interatomic distance. Interaction parameters for atoms of different types are generated from the homogeneous parameters via a geometric mean. For computational efficiency, the interactions are truncated at a cutoff of $R_c = 1.2nm$.

Electrostatic interactions are computed via a reaction field treatment with a spherical cutoff of the same magnitude as that for the Lennard-Jones interactions; the intramolecular treatment is handled analogously:

$$U_{ij}^{coulomb}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_o}\left[\frac{1}{r_{ij}} + \frac{r_{ij}^2}{2R_c^3} - \frac{3}{2R_c}\right], \qquad r_{ij} \leq R_c \tag{7.2}$$

$$U_{ij}^{coulomb}(r_{ij}) = 0, \qquad r_{ij} > R_c \tag{7.3}$$

Bond lengths are constrained to the appropriate values. The intramolecular potential, as described in Chapter 6, includes bond angle bending, out-of-plane bending, and dihedral angle interactions. The bond angle potential is quadratic

$$\psi_\theta = \frac{1}{2}k_\theta(\theta - \theta_o)^2 \tag{7.4}$$

where the parameters k_θ and θ_o are those parameterized by fits to internals distributions of an all-atom potential, namely the CFF91 potential (see Chapter 6). The out-of-plane angle is defined in the Wilson sense (reference) and is also quadratic,

$$\psi_\chi = \frac{1}{2}k_\chi\chi^2 \tag{7.5}$$

where the force constant k_χ is again derived from fits to an all-atom potential. For the dihedral interactions, a Fourier series of the form

$$\psi_\phi = k_{phi,1}[1+cos(\phi+\phi_{1,O}] + k_{phi,2}[1+cos(2\phi+\phi_{2,O}] + +k_{phi,3}[1+cos(3\phi+\phi_{3,O}] \quad (7.6)$$

Note that with the current model, intramolecular interaction between the end sites of a dihedral angle are implicitly included in the torsion description, thus untreated with any Lennard-Jones type interaction.

Tables 7.1, 7.2, 7.3, 7.4 list the Lennard-Jones, charge, and intramolecular potential parameters.

Table 7.1: United-Atom Non-Bond Parameters: Ester Oligomer Model

Group Type	ϵ $\frac{kJ}{mole}$	σ nm	charge e	mass (a.m.u.)
C^* , Carbonyl Carbon		0.375	0.55	12.01115
O', Carbonyl Oxygen		0.296	-0.45	15.9994
O, Ether		0.3	-0.4	15.9994
CH		0.38	0.30	13.01912
CH_3, pendant		0.391	0.0	15.03506
CH_3, carbonyl		0.391	0.05	15.03506
CH_3, ether		0.380	0.25	15.03506

Table 7.2: United-Atom Bond Angle Parameters: Ester Oligomer Model

Bond Angle	K_θ $\frac{kJ}{mol\,rad^2}$	θ_o degrees
$CH_3 - C^* - O'$	801.1	123.9308
$CH_3 - C^* - O$	704.7	110.3288
$O' - C^* - O$	1079.2	124.1247
$C^* - O - CH$	582.2	118.2380
$O - CH - CH_3$	577.8	111.2130
$O - CH - C^*$	770.9	108.8752
$CH_3 - CH - C^*$	341.1	113.0788
$CH - C^* - O'$	812.6	124.6914
$CH - C^* - O$	687.0	111.2485
$C^* - O - CH_3$	553.2	117.4483

Table 7.3: United-Atom Torsion Angle Parameters: Ester Oligomer Model

Torsion Angle	A_1 $\frac{kJ}{mol}$	A_2 $\frac{kJ}{mol}$	A_3 $\frac{kJ}{mol}$	δ_1 degrees	δ_2 degrees	δ_3 degrees
$CH_3 - C^* - O - CH$	25.0	-22.0	-5.0	0.0	0.0	0.0
$CH_3 - CH - C^* - O'$	5.0	1.0	-2.0	-30.0	-30.0	-30.0
$CH_3 - CH - C^* - O$	-6.0	1.0	2.0	-25.0	-25.0	-25.0
$CH - C^* - O - CH$	34.0	15.0	12.0	0.0	0.0	0.0
$O' - C^* - O - CH$	-46.0	25.0	-17.0	0.0	0.0	0.0
$C^* - O - CH - CH_3$	6.0	10.0	3.0	0.0	0.0	0.0
$C^* - O - CH - C^*$	11.0	-11.0	-10.0	0.0	0.0	0.0
$O - CH - C^* - O'$	-3.0	-3.0	-1.0	-23.0	-23.0	-23.0
$O - CH - C^* - O$	4.0	-4.0	0.0	0.0	0.0	0.0
$CH - C^* - O - CH_3$	50.0	-50.0	-21.0	0.0	0.0	0.0
$O' - C^* - O - CH_3$	-49.0	23.0	-15.0	0.0	0.0	0.0

Table 7.4: United-Atom Out-of-Plane Angle Parameters: Ester Oligomer Model

Out-of-Plane Angle	K_χ $\frac{kJ}{mol\,rad^2}$	χ_o degrees
$CH - C^* - O - O'$	444.6	0.0
$CH_3 - C^* - O - O'$	444.6	0.0

7.2.2 Simulation Technique

Bulk Simulations: Constant Pressure and Temperature

In order to determine the density of the model at a temperature of 298 K and atmospheric pressure, constant-pressure, constant-temperature MD simulations were performed on a system of 64 oligomer molecules.

For the bulk PLA oligomer simulations, a system of 64 molecules (1600 "atoms") enclosed in a cubic box with periodic boundary conditions is used. A Nose-Andersen thermostat [44] with thermal inertia parameter, W, of $100.0\frac{kJps^2}{mol}$ and pressure inertia parameter, Q, of $10.0\frac{kJps^2}{mol\,nm^3}$ for temperature. The choice of the inertia parameters is made more-or-less by trial and error. One must make sure that the selected parameter set allows proper damping of the fluctuations in the pressure and temperature over a relevant time scale. Furthermore, in the case of NPT simulations, care must be taken to select parameters such that the time-scale for the system temperature to equilibrate following each volume change(as per the constant-pressure algorithm) is sufficient to attain overall system temperature equilibrium. There is offered a rationale for selecting these parameters based on the speed of sound in the fluid. It is suggested that in order to approximate a small volume element of a fluid, the time scale for box size fluctuations should be on the order of the speed of sound in the fluid [38]. This criterion, of course, requires *a priori* knowledge of fluid properties. In the final analysis, one often turns to a trial and error approach or selects parameters based on experience with similar systems. RATTLE, the velocity Verlet method for systems with bond constraints is used to integrate the equations of motion; a timestep of 0.003 ps is used. A spherical cutoff of 1.2 nanometers is applied for both the non-bond Lennard-Jones and reaction field electrostatic interactions. A Verlet neighbor list is used to track 0 within a spherical volume of 1.4 nanometers around each particle.

The initial configuration was a cubic lattice with individual molecules located at evenly-spaced separations in a cube of dimensions of $L_x = L_y = L_z = 5.0nm$. All the chains are started from the same conformational geometry, which is itself a random distribution of torsion and bond angles; there was no effort made to start all

chains from an all-trans or all-gauche type conformation as is often done. This will not affect the final outcome of the distribution of internal coordinates in the equilibrium bulk; this can be verified by computing the internals distributions in the equilibrium bulk generated from different starting configurations. The initial system is much larger than the equilibrium volume attained; this was done so as not to collapse the system initially via interparticle overlaps, as well as to allow the molecules to explore conformational states (essentially traversing a gas-phase trajectory) before settling into an average liquid-state conformation.

Results of Constant-Pressure, Constant-Temperature Bulk Simulations

This section gives the results of the bulk MD simulations of the 64-molecule oligomer system described in the previous section. The results of average pressure, average temperature, and average system volume are given in Table 7.6 along with the associated errors. The averages are computed by blocking the time data into segments of 100 data points, with every two data point being 0.03 picoseconds in separation. Each block thus represents 0.3 picoseconds of simulation time. The total simulation time over which the average is computed is roughly 200 picoseconds. The time-average system volume is 31.859 cubic nanometers which corresponds to a density of 1.209 $gr\,cm^{-3}$. For the purposes of applying the oligomer system as a model for the PLA polymer, this value of the density is in excellent agreement with the experimental range of densities, 1.1 to 1.3 $gr\,cm^{-3}$, reported for PLA (Table 7.5 shows a range of the values determined by some groups using methods such as air-comparison pycnometer).

One anticipates some fluctuation of this value upon using a different ensemble method for MD simulations of the model, but the value thus obtained is well within an acceptable range of values and we are hopeful that upon going to a canonical ensemble will not introduce much change in the density. Note that the canonical ensemble is chosen since the preferred geometry for performing the interface simulations (oligomer liquid-vapor) will be a constant volume, rectangular simulation box. This will be further described below within the appropriate section.

Table 7.5: Experimental Values of Bulk PLA, PGA, and PLGA Densities

Material	Density (gr/cm^3)
PLA (amorphous)[a]	1.25
PLA (crystalline)[a]	1.36
PLGA 50:50, PLGA 75:25[b]	1.29
PLA[b]	1.26
L-PLA[c]	1.24
DL-PLA[c]	1.25
PGA[c]	1.53
PLA[d]	1.05-1.1

The computed pressure, which averages to somewhat higher than the desired 1 bar value, is well within the error bounds associated with NPT molecular dynamics. The pressure is notoriously vulnerable to large fluctuations (since the quantities from which it derives fluctuate significantly and are quite sensitive to timestep—this last ultimately is an unavoidable consequence of the numerical integration) and is often very difficult to control within the extended Hamiltonian formulation of system dynamics. The system dynamics are sensitive to the choice of inertia parameters used (more so to the pressure parameter for constant pressure dynamics) and this often requires accepting a certain level of fluctuation in the pressure rather than performing a rather tedious, and invariably time-consuming optimization of the inertia parameter. Furthermore, the nature of the origin of the computed pressure–the sum of the forces between interacting particles–is itself inherently extremely sensitive to quite small fluctuations in density. Overwhelmingly large pressure fluctuations are a well-documented phenomenon associated with molecular dynamics [61, 60, 126]. Thus, in the final analysis, the results of the present simulations are quite acceptable.

Table 7.6: Temperature, Pressure, Volume for 64-Oligomer Bulk System: NPT MD

	Temperature (K)	Pressure (bar)	Volume (nm^3)
Average	298.03	1.65	31.859
Standard Deviation	0.374	39.9	0.074
Standard Deviation (Mean)	0.040	3.72	0.0079

Bulk Simulations: Constant Volume and Temperature

Having determined a system density based on NPT molecular dynamics, canonical ensemble (NVT) molecular dynamics was performed for a system of 64 oligomer molecules at the average density of 1.2091217 $gr\,cm^{-3}$. A cubic simulation cell of edge lengths of 3.17 nm was used to contain the molecules. The initial configuration was taken to be the final configuration from the NPT simulation described in the last section. A Nose-Andersen thermostat with thermal inertia parameter of $100.0\frac{kJps^2}{mol}$ was used to maintain the temperature at 298 K. The dynamics trajectory was generated using RATTLE with a timestep of 0.003 picoseconds. A production run of 96 picoseconds resulted in the average values for pressure and temperature listed in Table 7.7. The pressure obtained at this value of the bulk density averages to -37.93 bar, well out of the physical bounds of this thermodynamic variable. One could in theory accept this value if it were within the statistical error of the NPT simulations; however, the lower statistical bound for the NPT pressure is around -2.0 bar, while the upper statistical limit for the NVT simulations at a density of 1.209 $gr\,cm^{-3}$ is -22.6 bar. Clearly, even the large pressure fluctuations do not allow the two results to remain statistically equivalent for the purposes of determining a self-consistent set of thermodynamic variables defining the state point of interest, namely, standard conditions of temperature and pressure (T = 298K, and P = 1 bar) and the corresponding bulk density.

From a theoretical viewpoint, the ensemble averages of properties should be independent of the ensembles which are used to calculate them; this is rigorous in the thermodynamic limit. An equivalent statement is that the thermodynamic potentials used to represent an equilibrium system all contain the same information. Transforming between the different ensembles simply represents alternate (but equivalent) ways of describing a system in terms of conjugate extensive and intensive variables. In classical thermodynamics, the Legendre transformation presents the mathematical framework within which one transforms between the various ensembles. We note here that for a finite system, the equivalence is not obvious unless one considers the formal

statistical thermodynamic formalism to transformations between ensembles. Without delving into the details of the derivation [13], the error incurred for finite systems is evident by comparing the average values of a property computed in constant-F and constant-f ensembles [13],

$$\langle \mathcal{A} \rangle_F = \langle \mathcal{A} \rangle_f + \frac{1}{2} \frac{\partial}{\partial f} \left(\frac{\partial f}{\partial F} \right) \frac{\partial}{\partial f} \langle \mathcal{A} \rangle_f \tag{7.7}$$

With the stipulation that F is an extensive quantity and f intensive (the two being a conjugate pair), the magnitude of the correction term (second term) in Equation 7.7 amounts to $\mathcal{O}(N^{-1})$, where N is the system size. Thus, for a small, finite system, one should not be surprised to observe inconsistencies in average properties computed in various ensembles. Certainly, in the present case, one can argue that the system size of 64 molecules may not be sufficient to mitigate finite-size effects on averages between ensembles. However, as mentioned above, the average value of a property is computed within the fluctuation bounds of other variables; in this case, the bulk system pressure is computed within the bounds of the fluctuations in the system volume. Thus, before trying to determine in detail the contribution to the error in Equation 7.7, the current approach will be to evaluate the effect of the volume fluctuation on the system pressure.

Table 7.7: Temperature, Pressure Results for 64-Oligomer Bulk System: NVT MD $\rho = 1.2091217 \frac{gr}{cm^3}$

	Temperature (K)	Pressure (bar)
Average	298.05	-37.9
Standard Deviation	0.285	87.2
Standard Deviation (Mean)	0.050	15.4

Based on the fact that the pressure at the specified density was too low (and in fact, negative), a second simulation was performed at the lower bound of the system volume (or equivalently, a higher bulk system density) computed from the NPT simulations. Although negative pressures are physically feasible, for a bulk liquid system one must obtain a positive pressure under the conditions considered in

this work. As a first guess (and for lack of a better approach to estimate what would be an acceptable density), the value for the system volume was chosen to be 31.85077 $gr\,cm^{-}3$, which corresponds to the block average volume of the NPT simulation less the standard deviation of the mean (for exact values, refer to Table 7.6). Thus, once again, a cubic simulation cell was used to simulate 64 oligomer molecules. The initial configuration for the simulations was taken to be the last configuration from the NPT calculations. Since the volume had to be reduced, an isotropic scaling of the box dimensions was performed to obtain a cubic cell of length 3.1699 nm. A Nose-Andersen thermostat with thermal inertia parameter of $100.0\frac{kJps^2}{mol}$ was used to maintain a temperature of 298 K. The equations of motion were integrated using the RATTLE algorithm with a timestep of 0.003 picoseconds; the calculations were performed for 150 picoseconds of simulation time.

The simulation results for temperature and pressure are shown in Table 7.8. With the slightly higher density representing the upper statistical bound of the NPT density, we approach a pressure close to the desired 1 bar; at this point, we feel that this is sufficiently representative of the desired state point. Again, the absolute average pressure obtained is slightly lower than 1 bar, but desired value is well within the statistical limits shown in Table 7.8. Thus, for further work with the oligomer system, we shall use the results of this simulation. It is worth noting that the effected order of magnitude increase in the system pressure was the result of a change in total system volume of less than 0.05 percent. This quite nicely shows the sensitivity of the pressure, as calculated from MD methods, to volume (density); this has critical implications in controlling system pressure as is well known.

Table 7.8: Temperature, Pressure Results for 64-Oligomer Bulk System: NVT MD, $\rho = 1.209434$

	Temperature (K)	Pressure (bar)
Average	297.99	0.979
Standard Deviation	0.307	81.8
Standard Deviation (Mean)	0.042	11.3

7.3 PLA Oligomer Liquid-Vapor Interface: Constant Volume and Temperature I.

Canonical ensemble (NVT) simulations of the PLA Oligomer model were performed using a rectangular simulation cell of dimensions $L_x = L_y = 4.5nm$ and $L_z = 20.0nm$, the longitudinal (z-coordinate) dimension being roughly 4-5 times the size of the tangential dimensions; L_z is selected to be large enough so as to prevent interactions between molecules at opposite surfaces across the vapor phase. Figure 7-1 shows the cell geometry used for the current simulations. The equations of motion are integrated using RATTLE, the velocity Verlet algorithm adapted to systems with bond constraints; a timestep of 0.003 picoseconds is used. A Verlet neighbor list is used to monitor neighbors of each particle; an outer skin thickness of 0.2 nm is used to generate the list of neighbors. A Nose-Andersen thermostat with thermal inertia parameter of $100.0\frac{kJps^2}{mol}$ is used to maintain the temperature about a mean of 298.0 Kelvin. The geometry of the liquid-vapor interface simulation is generated by situating two equilibrated bulk slabs of the oligomer liquid (both slabs of dimensions 4.5 by 4.5 by 4.718 nanometers) adjacent to each other to generate a system from which will evolve two liquid-vapor interfaces and a definitive bulk region. The interfaces are established once again by imposing an asymmetry in the box dimension along the longitudinal axis (perpendicular to the interface).

7.3.1 PLA Oligomer Liquid-Vapor Interface Results I.

This section will give some results of the MD simulation of the liquid-vapor interface, and the discussion is prefaced by several comments. The ensuing results indicated that the system size was rather too large; that is, a smaller system size would have sufficed in the analysis of the oligomer interface. Also, the diffusion of the molecules was observed to be quite slow. In the time available, the combination of large system size and slow molecular diffusion would have been unfeasible. In the next sections, we briefly discuss these results, and then move on to describing a modified system

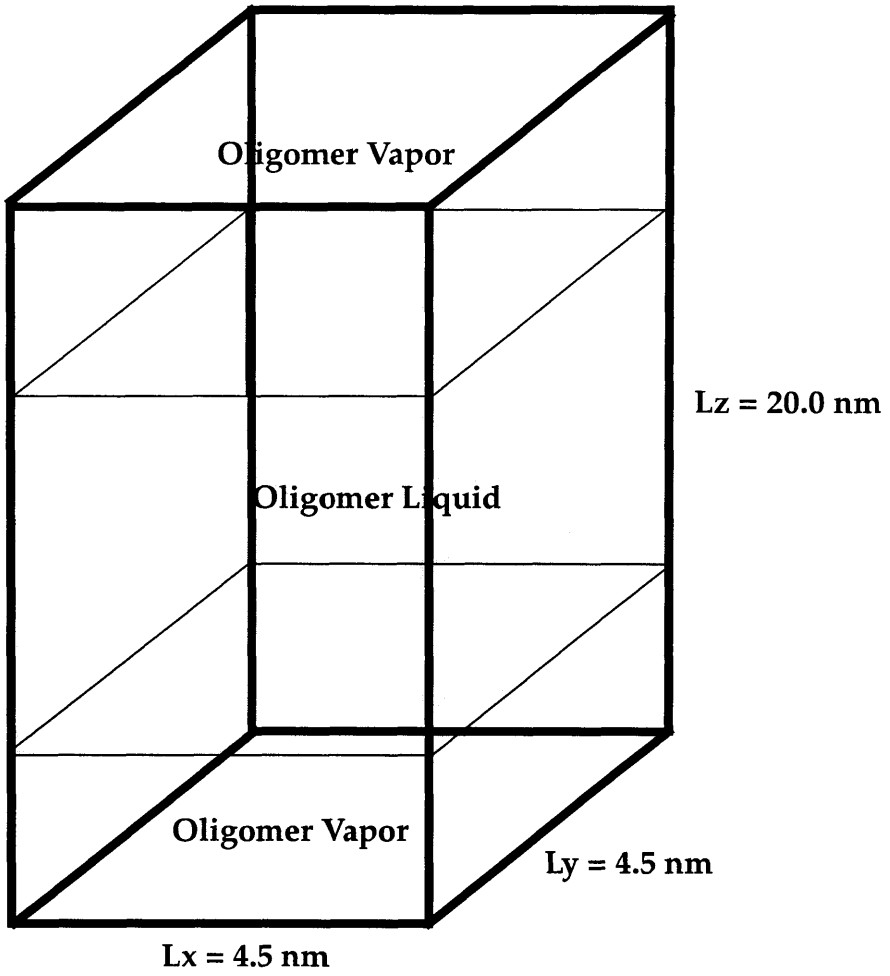


Figure 7-1: PLA Oligomer Liquid-Vapor Interface: Simulation Cell Geometry

and the MD results for that system.

Figure 7-2 shows total molecular number density profiles averaged over several distinct time intervals. Each profile is an average over 200 configurations (representing 60 picoseconds of simulation time). It is apparent that the profiles do not change significantly over the span of the simulation. Also, the bulk density is consistently lower than the required density of 1.209434 $gr\,cm^{-3}$. The average of the bulk density over the last 150 picoseconds is 1.188 $gr\,cm^{-3}$; the bulk region is taken to be the within 1.5 nanometers from the center of mass. Furthermore, Figure 7-3 shows the root mean squared center-of-mass displacement profile averaged over all molecules. This curve indicates an extremely small self-diffusion coefficient. Consequently, at this

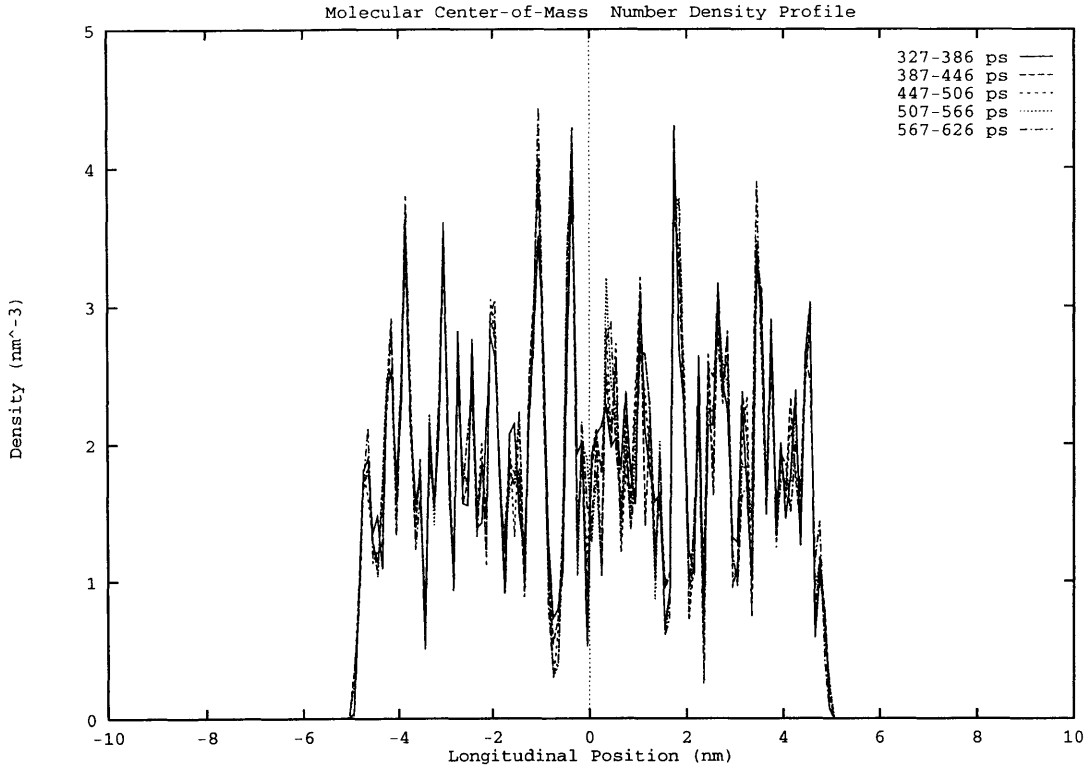


Figure 7-2: Molecular density profiles averaged over the time intervals indicated. The longitudinal position is relative to the system center of mass.

point, we felt that the best route forward was to select a smaller system size. This is reasonable since the density profiles show that the current size is more than enough to allow a bulk and interfacial region to evolve. Furthermore, to simulate 6 picoseconds requires 6.8 hours of real time. This translates to about 47 days (barring any system failures and assuming constant operation) to simulate one nanosecond of simulation time. This may not seem too extravagant, but considering the time scales associated with the diffusion of molecules, it seems that one would not see much motion from the initial state after such a long run time. This is an exorbitant computational expense for very little gains in information by using such a large system. Thus, the next section describes a smaller system selected for study and the preliminary results obtained from the simulations of that system.

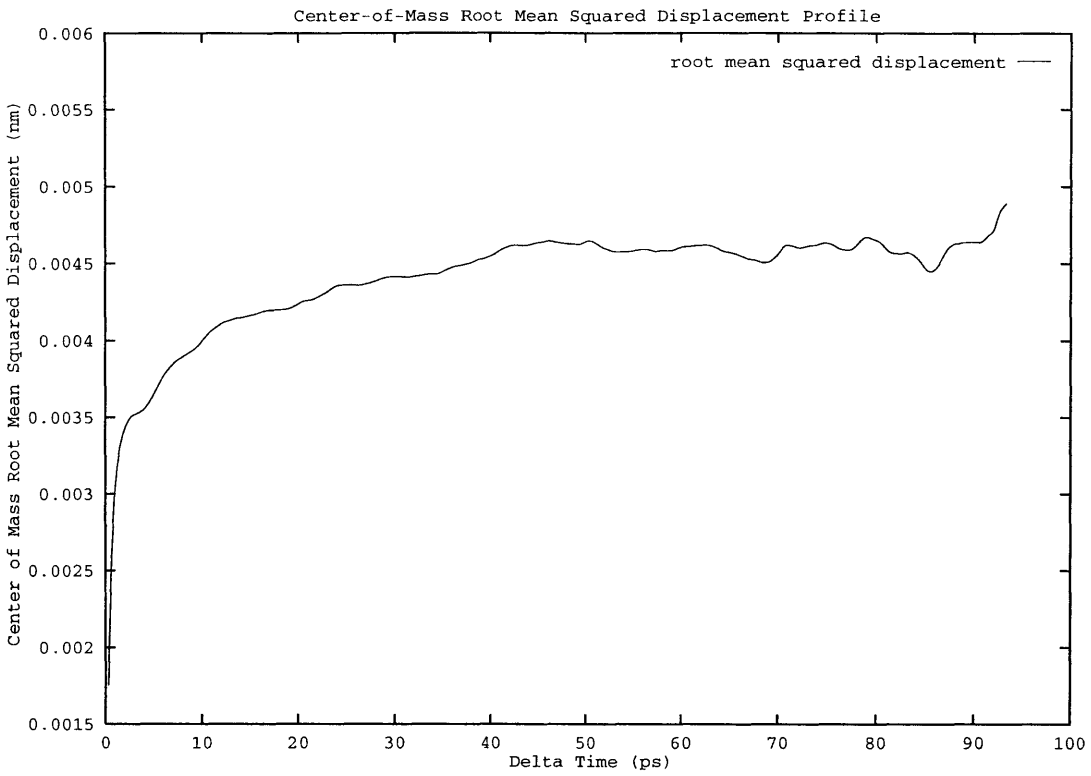


Figure 7-3: Center-of-Mass root-mean-squared displacement profile for oligomer.

7.4 PLA Oligomer Liquid-Vapor Interface: Constant Volume and Temperature II.

A smaller system consisting of 231 oligomer molecules was generated by randomly placing molecules in a total volume so as to give a density slightly smaller than that determined earlier for 298K and 1 bar pressure. The density chosen was 1.144 $gr\,cm^{-3}$; this presented no problems in terms of the final density as the system was equilibrating towards a higher density as expected (this will be shown below). Furthermore, with a lower initial density, it was felt that the system would be able to equilibrate more rapidly in terms of orientational and translational order with a higher degree of individual molecule motion possible due to the reduction in density.

As for the simulation cell, all was kept exactly as before except that the longitudinal dimension was reduced to 12 nanometers from the original 20. This again presented no inconsistencies in the approach. Thus, this modified system contains

5775 particles instead of the original 9600, and in the longitudinal dimension spans 6 nanometers as compared to the original 10. The simulation methodology is retained from the first simulation (larger system). The reader is referred to that section for details.

7.4.1 PLA Oligomer Liquid-Vapor Interface: Results II.

This section describes preliminary results of simulations of the liquid-vapor interface of PLA oligomers. Profiles of the total molecular center of mass, total site, and individual group number densities are presented. Profiles of the components of the radius of gyration are reported. Bulk density time profiles are shown to indicate the equilibration of the system; furthermore, we show the translational order parameter as a further check on whether a *liquid* state is achieved. We note here that these are preliminary results since they are based on short simulation times. Due to constraints in real time, this study is restricted to roughly 1000 picoseconds. To be rigorous, one must actually sample such a system for at least a nanosecond to obtain sound statistics. In the present case, however, some interesting behavior and trends can be discussed, noting the context within which they emerge.

Density Profiles and Bulk Density

Number density profiles for molecular center of mass and atom number density (all relative to the simulation cell center of mass) are computed as averages over 500 configurations (representing 30 picoseconds of simulation time) which are then block averaged; that is, profiles computed as averages over 500 configurations are averaged to obtain the final block-averaged profile. Furthermore, exploiting the system symmetry, the profiles are symmetrized; only the symmetrized profiles are shown for the density profile discussion.

Figure 7-4 shows the symmetrized molecular center of mass number density profile. The bulk center of mass density averages to $1.980 nm^{-3}(0.011)$; this is 1.5 percent less than the molecular density determined from the upper statistical limit of the density

from NPT simulations at a temperature of 298 Kelvin and 1 bar pressure. However, again, this decrease in density is typical and expected due in part to the relaxation allowed by extending the longitudinal system dimension and because of the spherical truncation applied in the calculation of the non-bond interactions; in fact, this result is quite acceptable in terms of reproducibility of the density. Although there appears a very slight shoulder within the interfacial region, this is most probably due to insufficient statistics to smooth out the profile; one observes a great deal of fluctuation in the profile throughout the whole range. Overall, the profile is monotonic over the region. Figure 7-5 shows the symmetrized total site and center of mass number density profiles; the center of mass profile is scaled to fit on the same axis as the site density profile. This figure serves to show the inherent noise in the two data sets; this is a result of the fact that for each center of mass, there are 25 individual sites, thus leading to better statistical averaging for the site density profile. Note in this figure how the center of mass density is more compact relative to the total site density. This, of course, is to be expected since the center of mass cannot be farther out than any individual component of the molecule. Finally, we note that the 10-90 width of the total density profile is 0.45 nanometers. This is a very compact region, particularly compared to the methyl acetate liquid-vapor interface (10-90 thickness of 0.6 nm). The width is computed as the longitudinal distance over which the *total site density* falls from 90 percent of its average bulk value to 10 percent of the same. The bulk average is computed, once again, as a block average over 18,000 configurations spanning roughly the last 1.0 nanosecond of the simulation. Figure 7-9 shows the time profile of the total site number density in the central 3.0 nanometers of the simulation cell (bulk). The average over the last nanosecond is 49.38 (0.073) nm^{-3}, an error of 1.7 percent relative to the NPT value of 50.23.

Figure 7-6 shows the symmetrized density profiles of the individual atoms of the oligomer. The definition of the atoms considered is derived from Figure 6-1 showing the PLA oligomer model. The carbonyl methyl group is that terminating the 'left' side of the chain represented in the Figure; it corresponds to the 'X' group. The end carbonyl oxygen atoms refer to the first and last carbonyl oxygens of the chain

(I being the first and II the last). The ether oxygen and -CH groups represent the middle of the chain. Finally, the end methoxy methyl group refers to the methyl group terminating the 'right' end of the chain in Figure 6-1. The ordering of the profiles in the interfacial region shows a marked spatial distribution of the various atoms (in an average sense) which is also distinct from the random nature of the bulk. Based on earlier MD simulations of alkane oligomers [60], one would interpret this ordering as indicating that the ends of the chain (in this case, the carbonyl oxygen atoms and the methyl groups) are enriched in the outer region of the interface, while the middle segments (in this case the -CH and ether oxygen atom) reside in the layer just beneath (note also that the carbonyl oxygen atoms are counter-intuitively 'segregated' in the outer region of the interface). For normal alkanes, structure is held to be more energetically favorable (both based on simulation and theoretical arguments). In the present case, we do observe that the middle segments, represented by atoms located in the more central regions of the chain, are enriched in the region between 2.5 and 3.0 nanometers from the center of mass (immediately underneath the outer interfacial region). In the outer interface, the end segments seem to be more concentrated. This spatial distribution is also predicted by mean-field approaches [169]. Figure 7-7 shows the number density profiles of the middle groups along with the total site density; Figure 7-8 shows the analogous for the end groups. These two figures further show that up to 3 nanometers from the cell center of mass, the total site density is dominated by the middle groups, Figure 7-7 (which interestingly follow rather close to the molecular center of mass profile), beyond which the total density is weighted towards the end groups,Figure 7-8. The overall conformation of the chain is apparently of a horseshoe nature, with the middle of the chain more-or-less parallel to the interfacial plane and the ends oriented normal to it. In order to investigate the validity of this interpretation, one can look at quantitative measures of the orientation of sections of the chain as well as that of the entire chain. For this purpose, we consider in the next section the orientational order parameter $P_2(z)$ which describes the effect of the interface on the chain orientation.

Finally, we show a profile of the translational order parameter profile which is

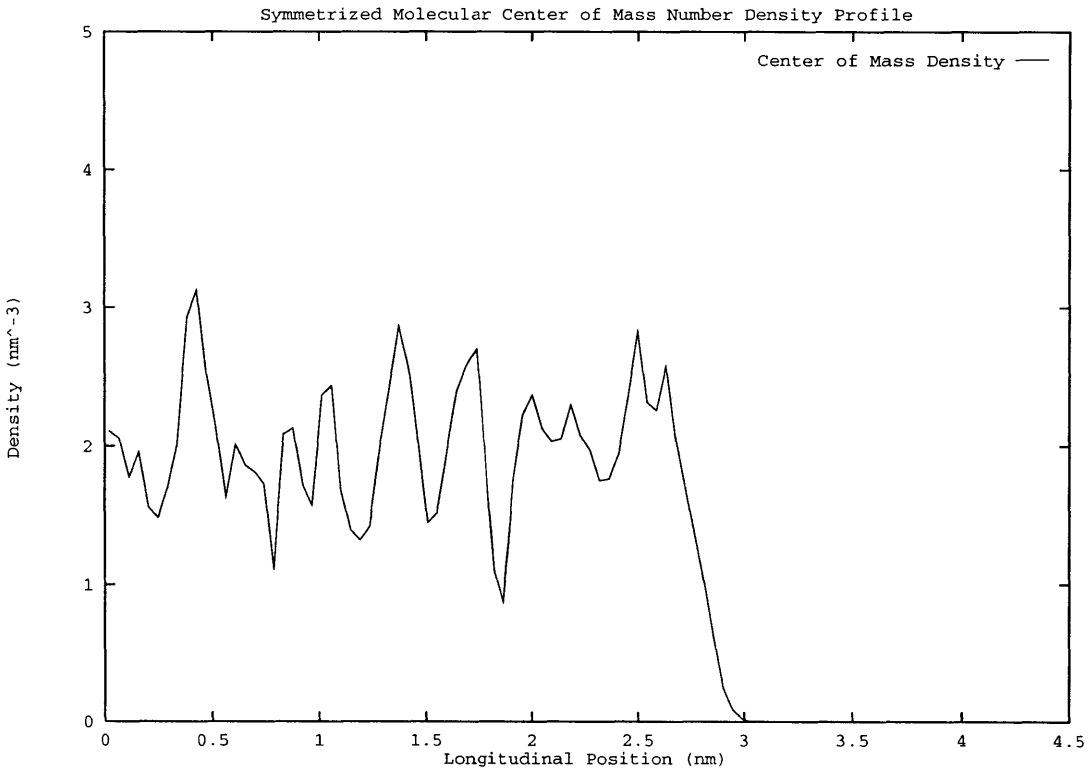


Figure 7-4: Total molecular number density profile. The longitudinal position is relative to the system center of mass.

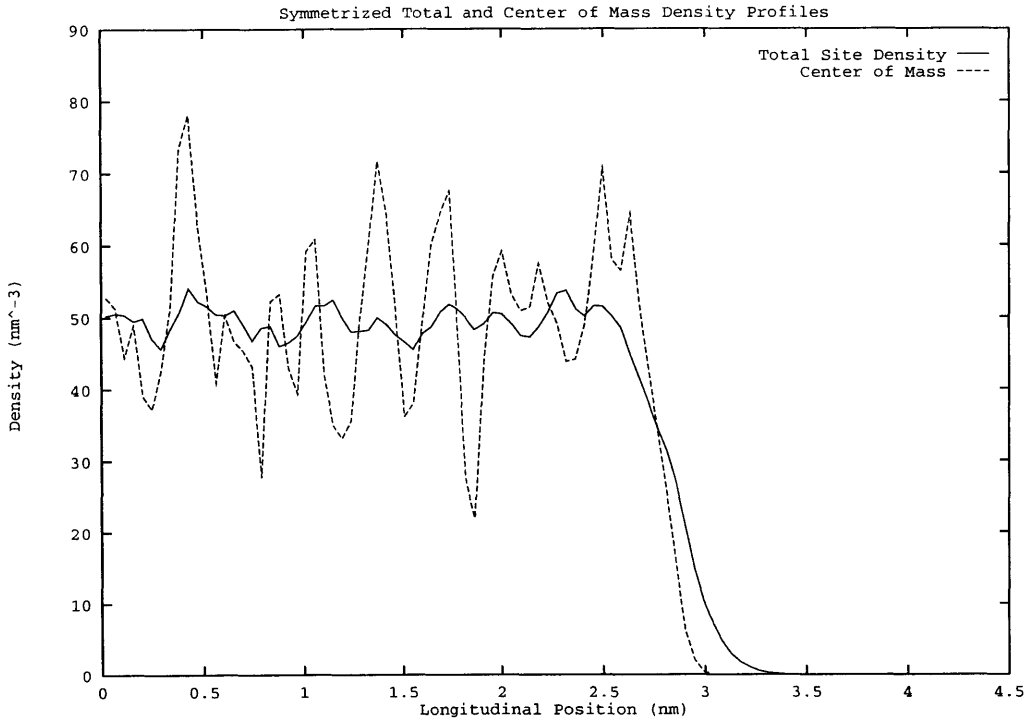


Figure 7-5: Total site and molecular center of mass number density profile. The longitudinal position is relative to the system center of mass.

useful in interpreting the nature of the fluid studied. For a liquid, the profile should fluctuate about zero; Figure 7-10 shows this to be the case. We are thus satisfied that we are not trapped in a 'solid' type structure which has long-range translational structure (such as a crystal).

Molecular Conformational Behavior: Orientational Order Parameter Profiles

In this section, we consider the effect of the interface on chain conformation. An informative order parameter for this is the orientational order parameter, $P_2(z)$ defined by the angle between the interface normal and a local vector. For the present work, the vector is taken to be that connecting atoms j and $j+2$ along the backbone of the chain. This definition excludes the carbonyl oxygens and the pendant methyl groups (which here are considered as side-groups). With θ being the angle between the interface normal and a given vector, the order parameter $P_2(z)$ is taken to be,

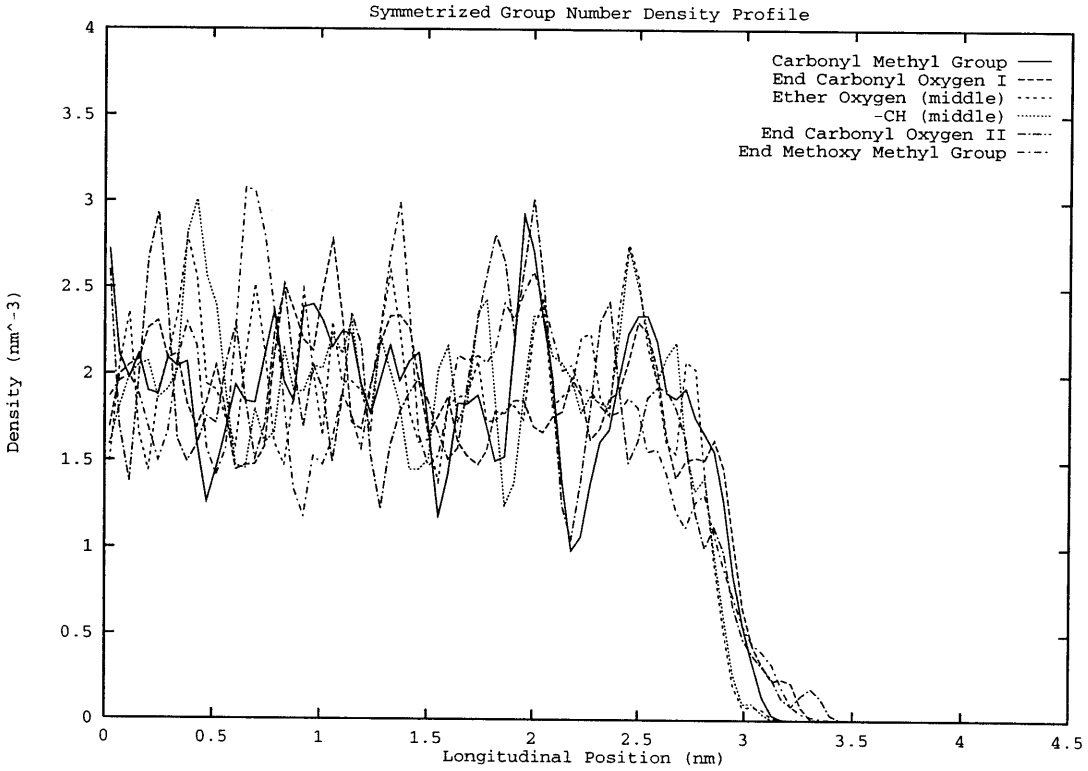


Figure 7-6: Atom number density profiles. The longitudinal position is relative to the system center of mass.

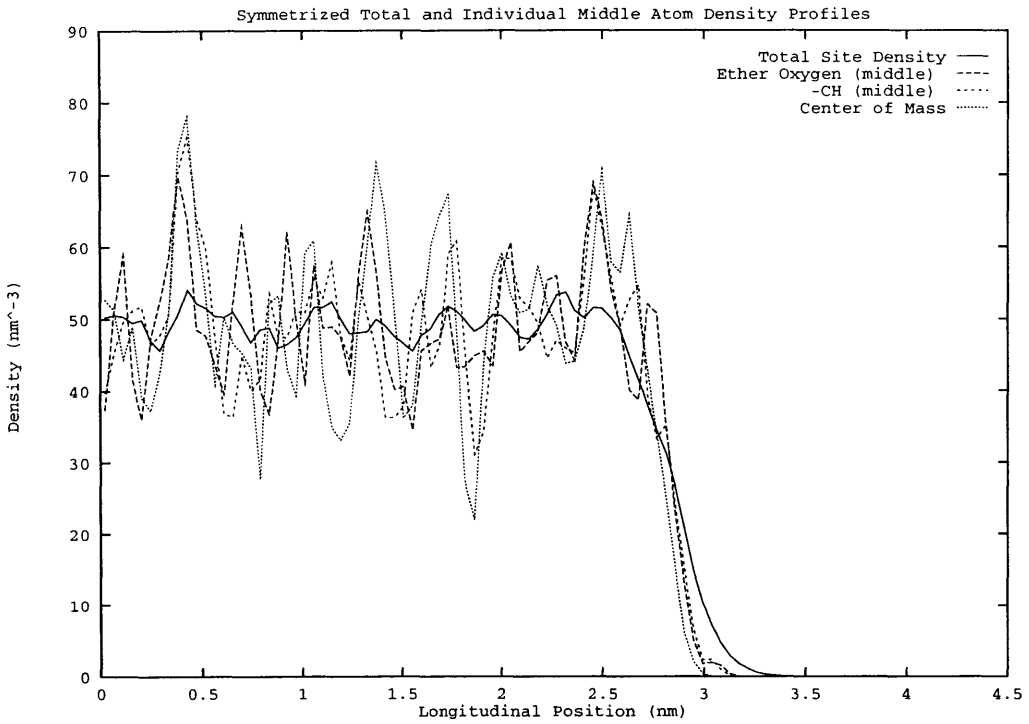


Figure 7-7: Total site and middle group number density profiles. The longitudinal position is relative to the system center of mass.

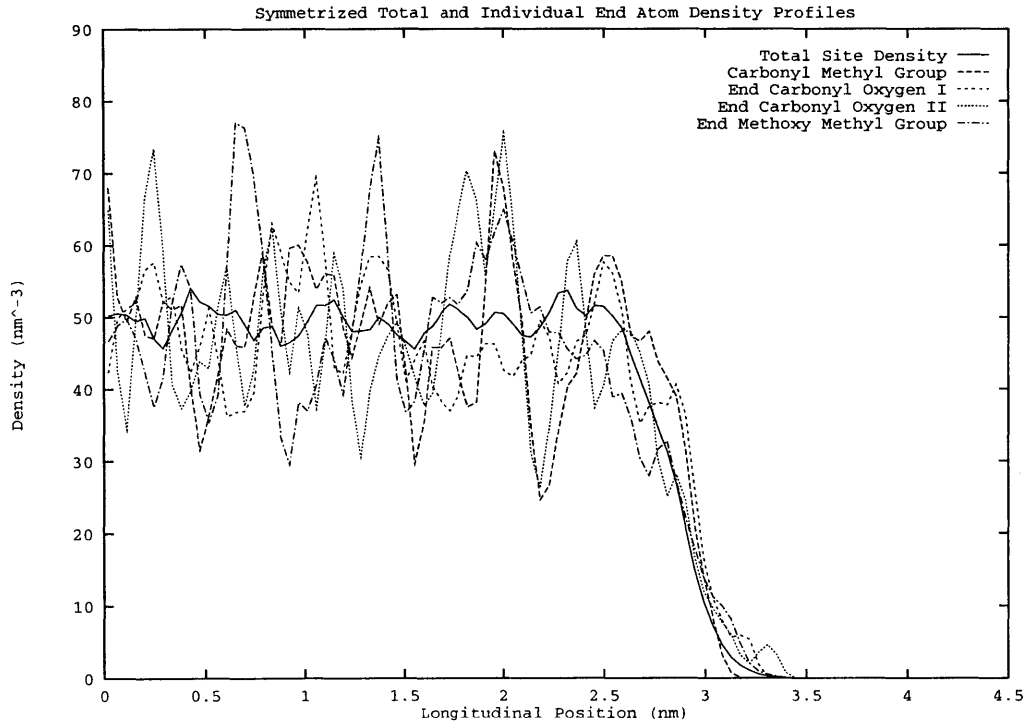


Figure 7-8: Total site and end group number density profiles. The longitudinal position is relative to the system center of mass.

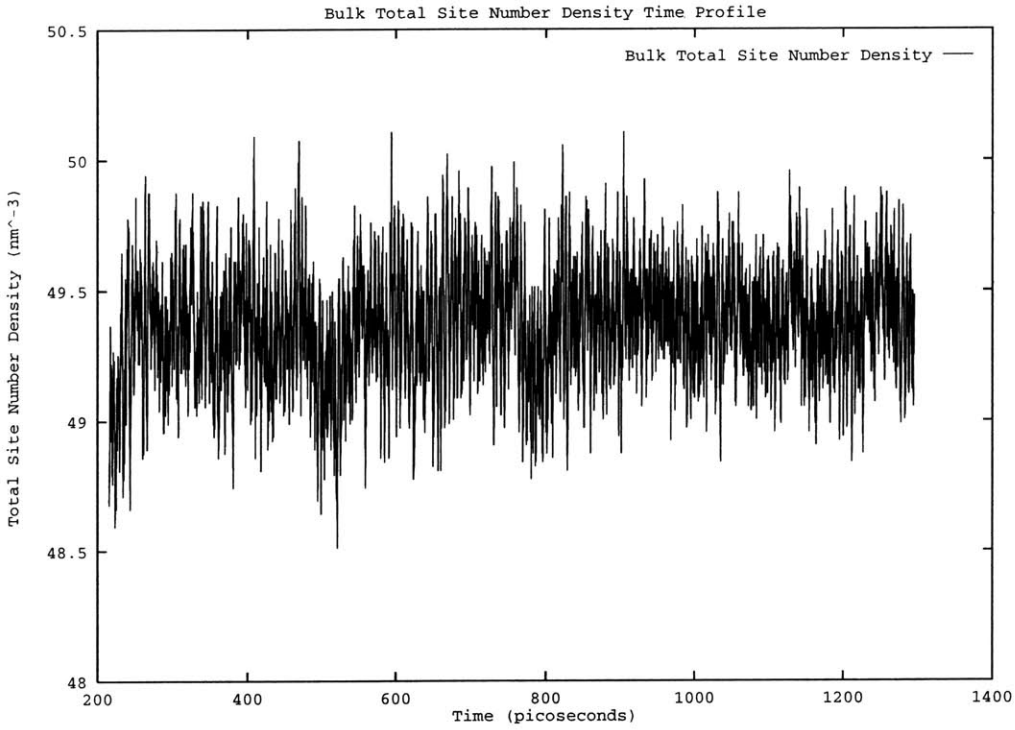


Figure 7-9: Total site number density time profile.

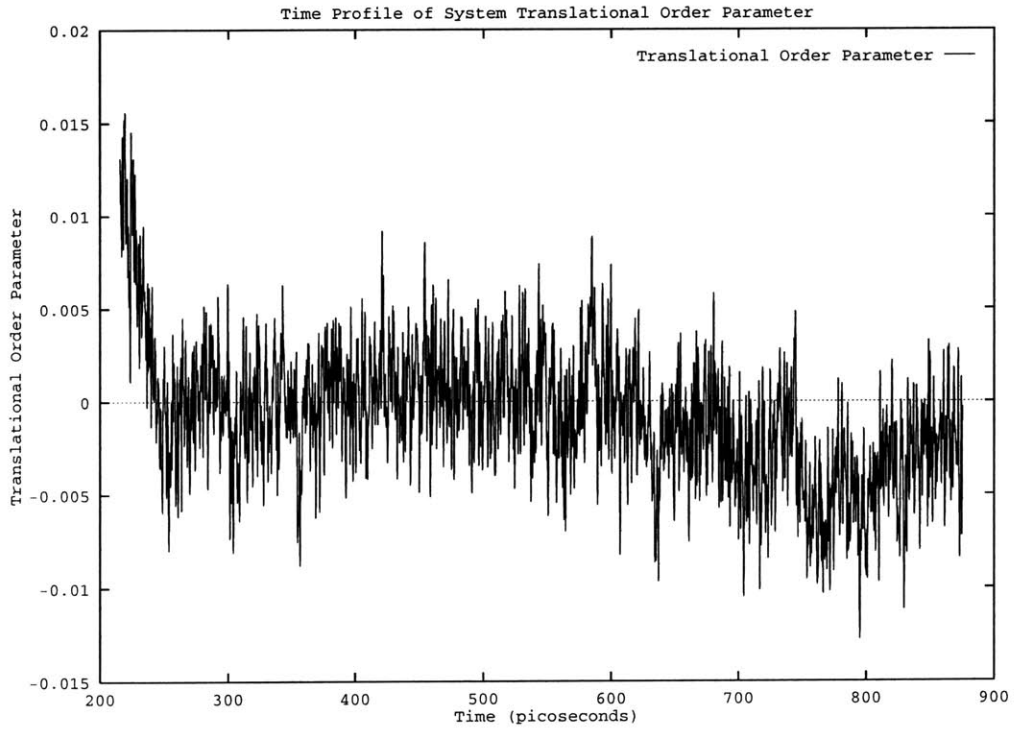


Figure 7-10: Time profile of system translational order parameter. A profile fluctuating about z ero indicates equilibration to a liquid state.

$$P_2(z) = \frac{1}{2}\left\langle 3(cos\theta)^2 - 1.0 \right\rangle \tag{7.8}$$

where the average is taken over all vectors and configurations. The region a vector is assigned to is determined by the location of the midpoint of the two atoms defining the vector. If the vectors in a region are on average randomly oriented, $P_2 = 0$. If they are normal to the interface, $P_2 = 1$, and if they lie in the interfacial plane, $P_2 = -0.5$.

We first consider vectors located at the ends of the chain; if the notion that the ends are jutting out towards the vapor is correct, these vectors should yield positive order parameters. The vectors considered are those between atoms 1 and 4 (end 1) and 25 and 22 (end 2). See Figure 7-13 for a schematic describing the vectors discussed. Due to the fact that the oligomer model does not possess molecular symmetry, we must consider each vector independently. We also target vectors in the middle region of the chain. In this case, we look at one vector between atoms 10 and 14 (middle 1) and 15 and 12 (middle 2) (refer to Figure 7-13 for description). Again, we look at these independently. Finally, we compute the order parameter profiles averaging over all vectors of the molecule. Figure 7-11 shows the order parameter profiles for the end vectors and the average over all vectors. These profiles unequivocally indicate a planar orientation for the two end vectors, and an average planar conformation of the entire molecule. It is interesting that the ends do not strongly segregate to the outer interface. Figure 7-12 shows the analogous profiles for the two middle vectors and the average profile over all vectors. Although both middle vector profiles show rather large fluctuations in the bulk, there is a discernible tendency of the vectors to adopt a planar orientation within the interfacial region. This means locally, the chain is flattened. Thus, based on this information, we are inclined to regard the interfacial chain conformation as being rather flattened; chain ends do not segregate to the outer surface with the middle sections of the chain (due to connectivity) enriched in the layer immediately beneath. This is in contrast to MD simulations [60] of the free interface of decane and eicosane, which indicate that in the outer interfacial layer,

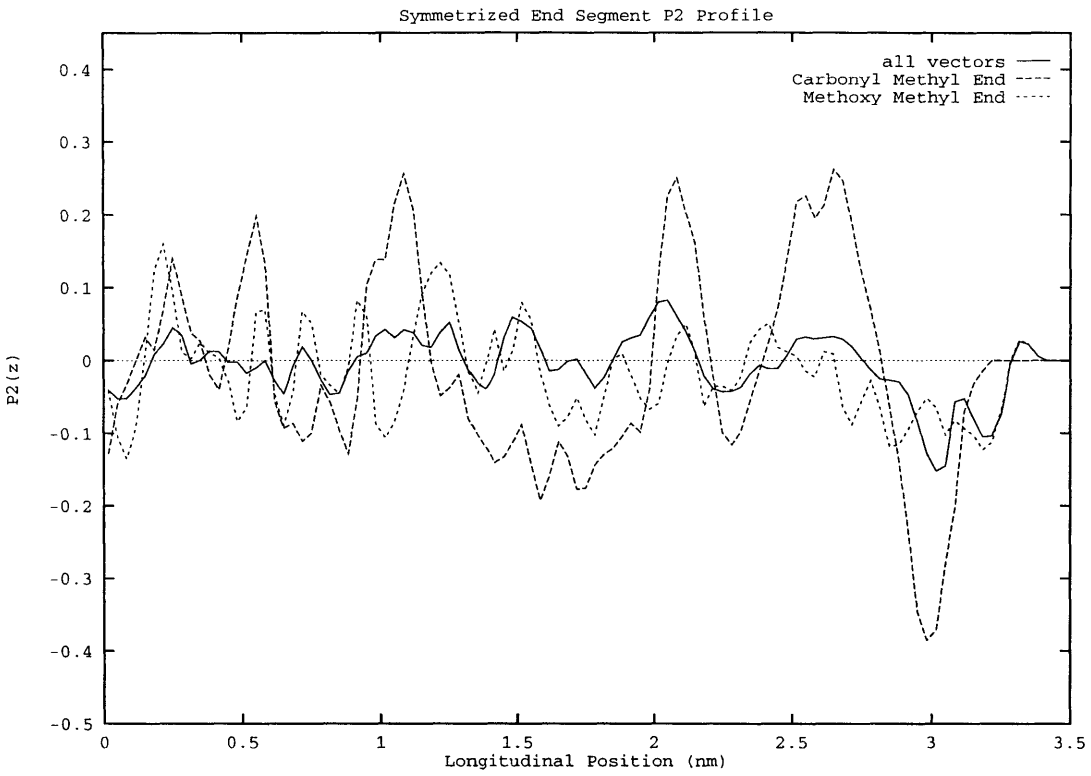


Figure 7-11: Order parameter profiles averaged over the end vectors and all vectors

there is an anisotropy in the behavior of chain ends and middle segments. Here, one sees that the ends are normal to the interface ('poking out'), and the middle segments are parallel to it. This is also observed in mean-field theories [170] and off-lattice Monte Carlo simulations [148].

With respect to the issue of chain flattening, we can compute components of the radius of gyration as a function of position from the system center of mass. Figure 7-14 shows the profiles of the x-, y-, and z-components of the radius of gyration in symmetrized form. Again, the final profile is a block average of profiles computed as averages over 500 configurations. In the interfacial region near 3 nanometers out from the center of mass, one observes a flattening effect of the molecule, with the longitudinal component of the radius of gyration smaller than either of the transverse components. Similar effects are seen in off-lattice MD simulations of alkane oligomers [60] as well as in equilibrium simulations of lattice models [148]. Interestingly, simulations of decane and eicosane performed by Harris [60] show what appears to be a

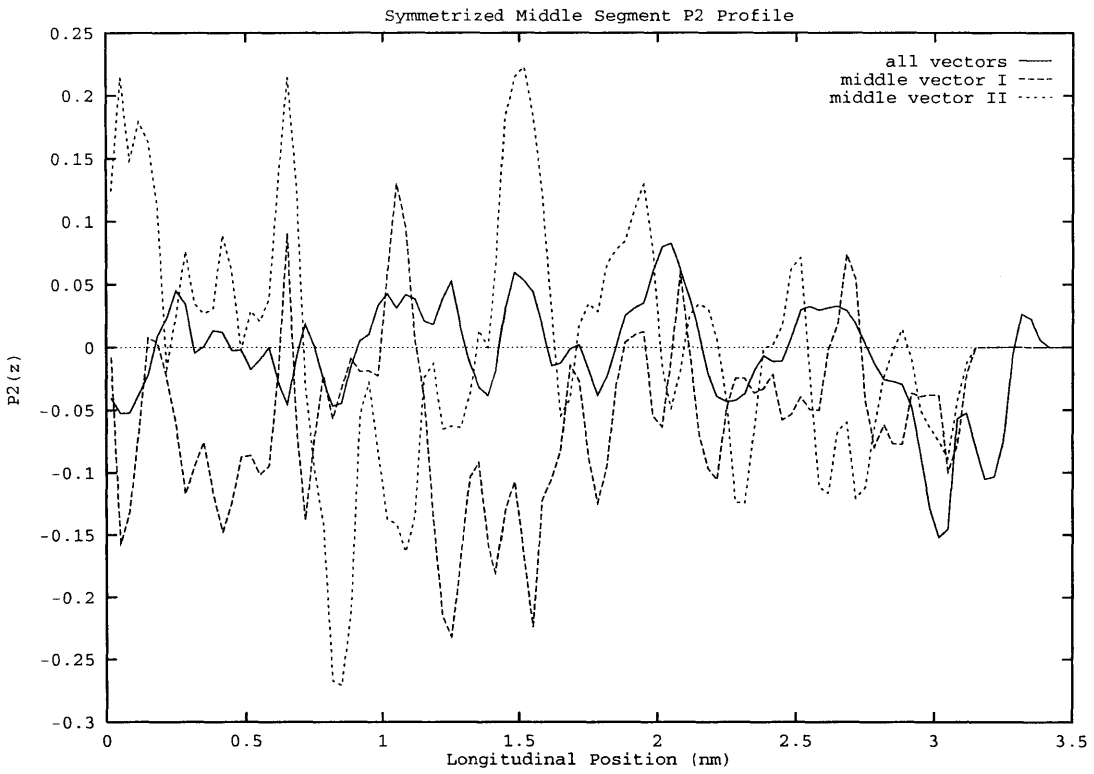


Figure 7-12: Order parameter profiles averaged over the middle vectors and all vectors

richer structuring in terms of chain conformational behavior. He shows that in the outer interfacial region of the alkane liquid-vapor interface, there is indeed a flattening of the chain as evidenced by similar radius of gyration profiles. However, in the region immediately beneath this outer region, he observes a layer in which the chains are more extended along the interface normal. In the present simulations, we fail to see this inner layer. It could be that the present system of 5-mers is still short enough to not allow such an effect; there may be some critical chain length where this phenomenon arises. Of course, there is the possibility that continuing the production phase of the simulation to much longer times may allow the evolution of this structuring.

One can further explore the question of chain flattening by examining the behavior of the mean squared direction cosines as one moves longitudinally from the bulk to the interface. Figure 7-15 shows the average of the squares of the direction cosines of the chain end-to-end vectors as a function of position from the bulk. The direction cosines

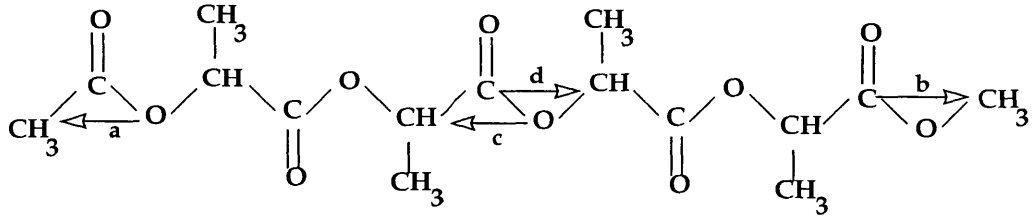


Figure 7-13: Schematic depicting the end and middle vectors.

are defined for a given chain end-to-end vector with Cartesian vector components, v_x, v_y, and v_z,

$$cos^2(\alpha) = \frac{v_x}{v_x^2 + v_y^2 + v_z^2} \tag{7.9}$$

$$cos^2(\beta) = \frac{v_y}{v_x^2 + v_y^2 + v_z^2} \tag{7.10}$$

$$cos^2(\gamma) = \frac{v_z}{v_x^2 + v_y^2 + v_z^2} \tag{7.11}$$

In a three-dimensional,isotropic environment, these three parameters average to $\frac{1}{3}$, while in a two-dimensional space, the existing dimensions would give parameters of value $\frac{1}{2}$ with the third parameter vanishing [148]. Figure 7-15 shows nicely the anisotropic behavior of the three parameters in the interfacial region between 2.5 and 3.0 nanometers from the bulk center. Despite the strong fluctuations in the bulk (which apparently do fluctuate about a value of one-third as expected), the difference in behaviors at the surface is evident; the transverse parameters (in the x and y-dimensions) tend to deviate towards a larger value, while the normal component attenuates. In the two-dimensional limit, the transverse parameters would take on values of $\frac{1}{2}$.

Now, with the picture of flattened chains at the interface compared to the bulk, one can ask whether the chains are 'extended' (elongated) in the interfacial plane,

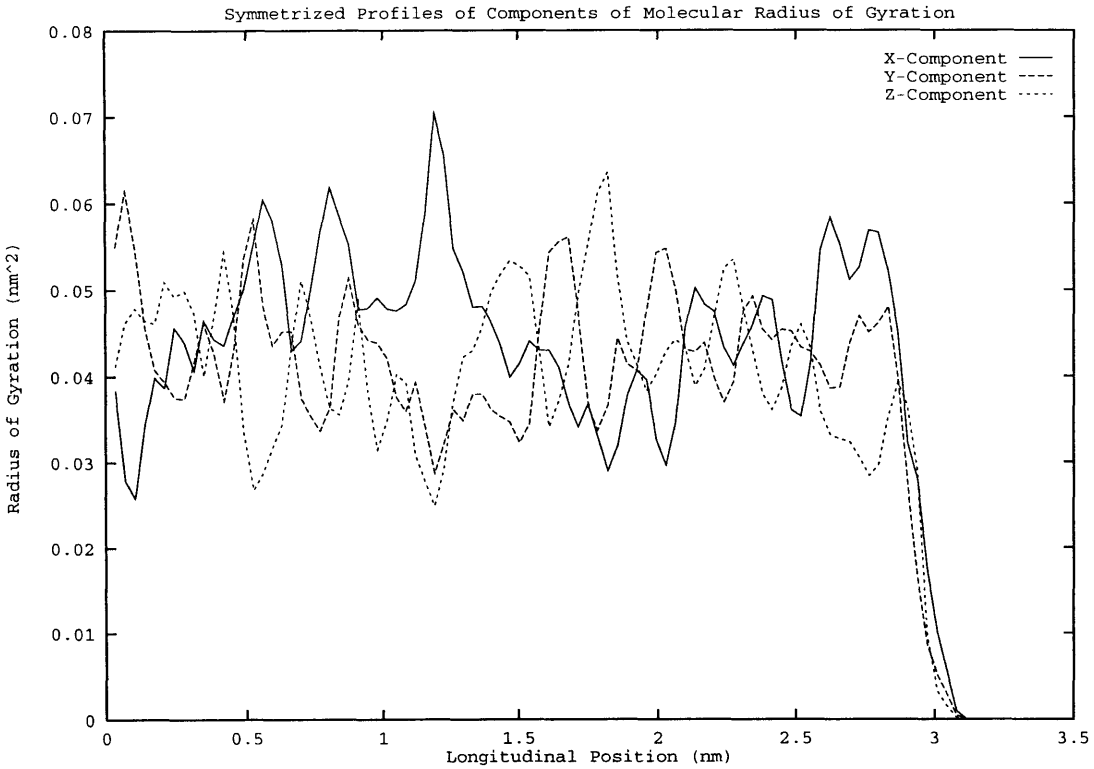


Figure 7-14: X-, y-, and z-components of the radius of gyration

or adopt a more compact planar structure such as a ring or a compressed spiral (much like a spring under compression). To answer this question, we look at the profile of the chain end-to-end vector, and more specifically, its transverse and longitudinal components. Figure 7-16 show symmetrized, block-averaged profiles of the three components of the chain end-to-end vector. We observe that in the interfacial region, the longitudinal component (z-component) decreases sharply. Concurrently, the transverse components, x- and y-components, rise sharply, or in the case of the y-component, hold the average bulk value well into the interfacial region; this indicates an average elongation of the chain relative to the bulk. This emphatically describes an extended chain at the interface. Again, this is in sharp contrast to previous MD simulations of alkane chains which show chains flattening but with their ends 'poking' out into the vapor region. Here we do not observe a strong chain end segregation.

Finally, returning to the interpretation of the density profiles of Figure 7-6, we now abandon the earlier notion of chain ends jutting into the vapor region and proffer

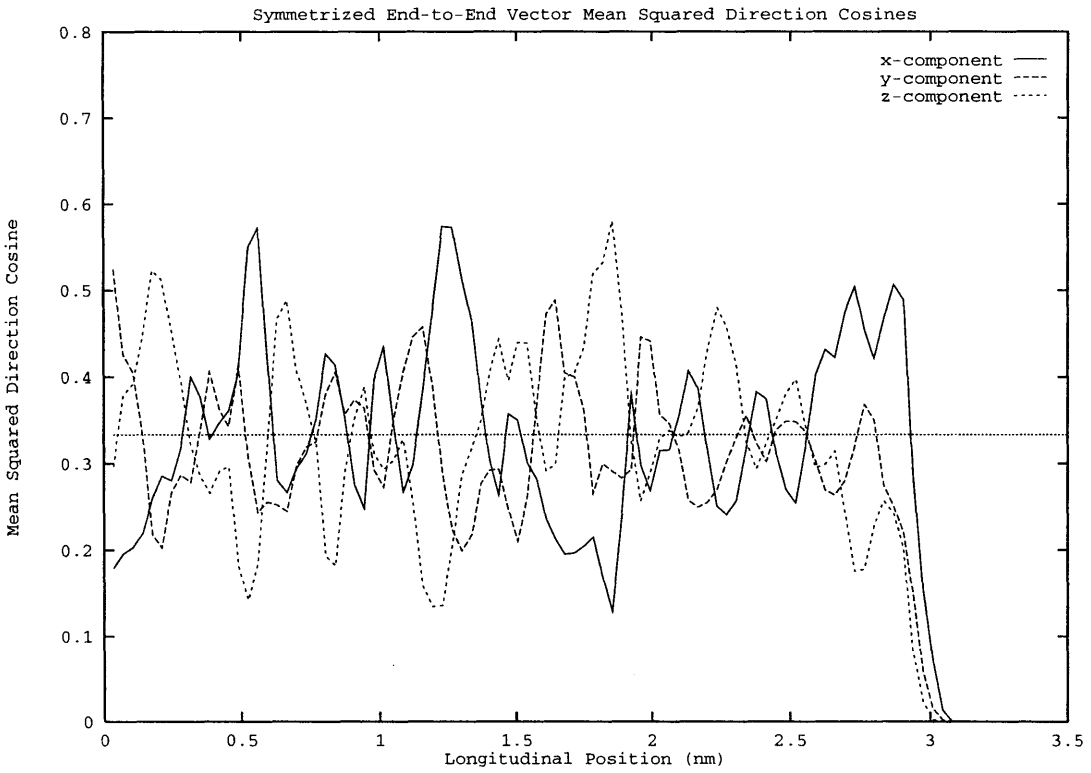


Figure 7-15: Mean squared direction cosines of the chain end-to-end vector as a function of longitudinal position (relative to the system center of mass.

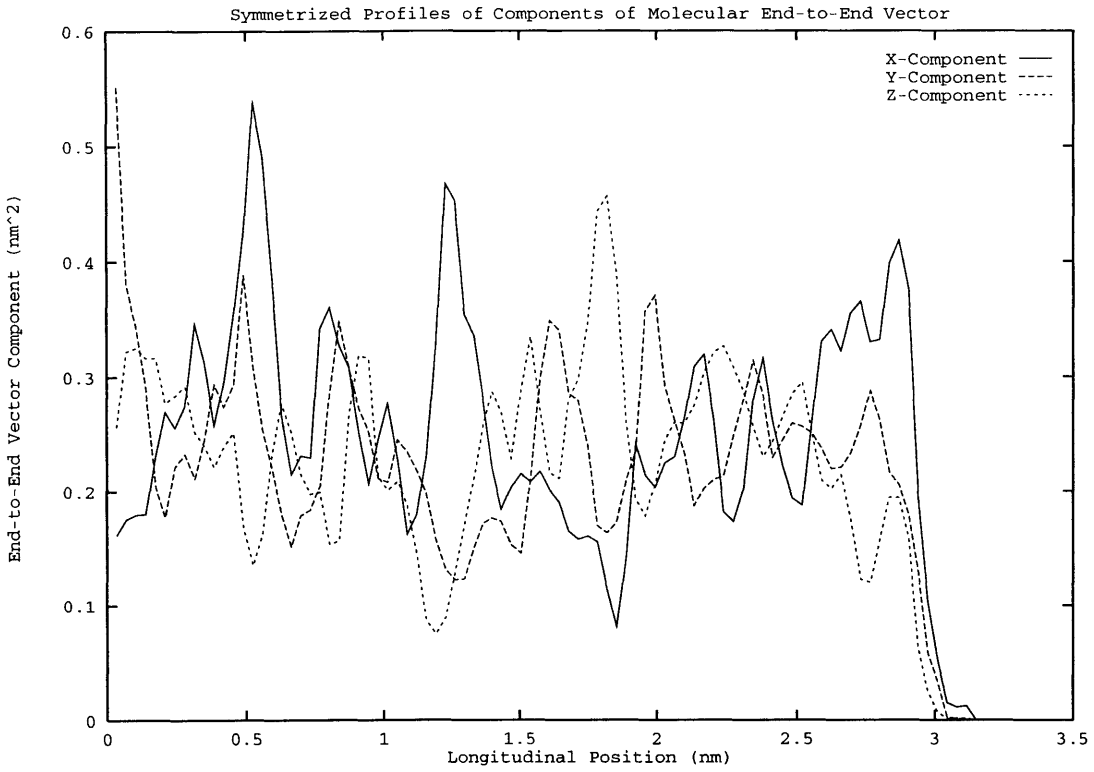


Figure 7-16: X-, y-, and z-components of the chain end-to-end vector

an alternative. The chains within the interface on average are elongated, but tilted at an angle to the interface. This allows one end of the chain to remain at a lower longitudinal position relative to the other end; furthermore, due to local conformations about individual dihedral angles, there is allowed the possibility for certain middle segments to lie lower longitudinally than either of the ends, albeit not too much lower. This interpretation appears consistent with the observed density profiles. The nature of the chain tilt can be seen in Figure 7-17 which shows the $P_2(z)$ profile for the chain end-to-end vector along with the $P_1(z)$ profile which in this case gives simply the cosine of the angle between the interface normal and the end-to-end vector. This figure shows that this vector is more or less planar, with slight deviations as shown by the P_1 curve. This deviation from total planarity results in the proposed average tilt of the chain. One can argue that despite the indication of the orientational order parameter profiles of the end and middle vectors being on average in a planar configuration, the chain may still be exhibiting end-segregation; it may be that the *degree* to which this

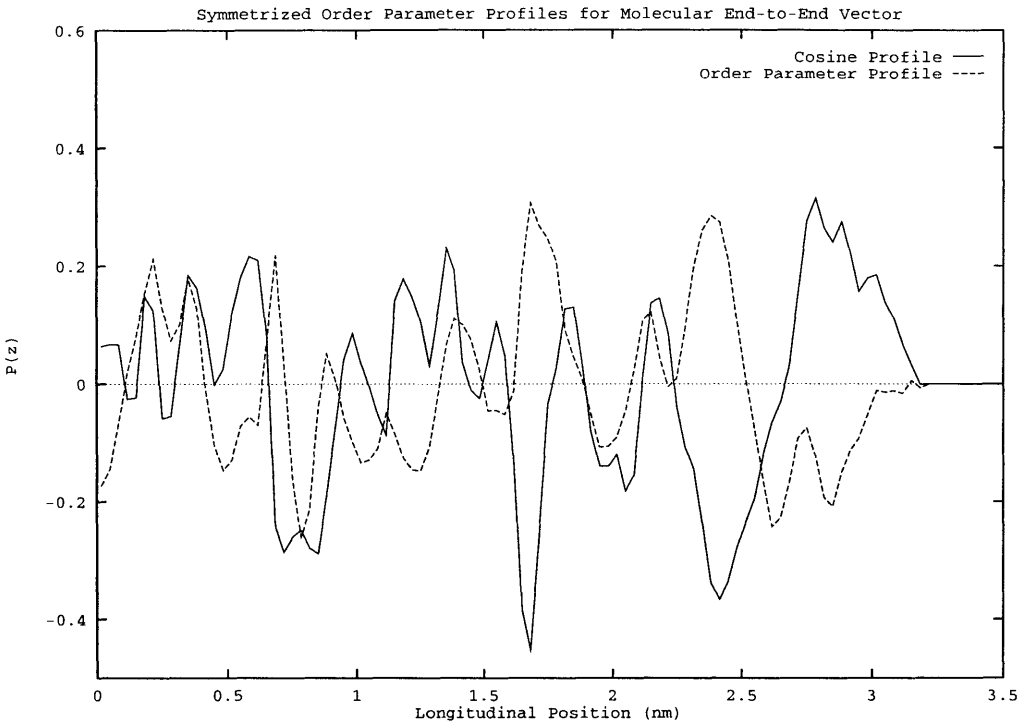


Figure 7-17: Order parameter profile of the chain end-to-end vector

segregation is occurring in this system is not strong enough to be dominated by the *local* orientational nature of the chain. It could well be that simulating a *longer* chain will give more unequivocal indication of chain-end segregation; at least one would hope that for much longer chains, the entropic effect associated with end segregation would be the more dominant.

For the present work, the observations discussed thus far are interesting, but from the point of view of biomaterials, one would like to inquire as to the nature of the interface with respect to what types of functional groups are present. With this in mind, in the next section, we investigate the 'appearance' of the surface (as let's say, an ambient peptide might see) via the behavior of orientations of carbonyl bonds in the bulk and at the interface.

7.5 The Nature of the Oligomer Liquid-Vapor Interface: Carbonyl Bond Vector Orientations

At this point, we can ask about the local orientations of bond vectors of the molecule. In particular, we are interested in the orientation of the carbonyl bonds as these are the focal point of hydrolytic processes in the actual polymer. Again, to study the bond orientation, we compute profiles of the average angle between the interface normal and the carbonyl bond vector, as well as the orientational order parameter, P_2 associated with this angle (and defined above). Note that the carbonyl bond vector is defined as $\mathbf{r}_{CO} = \mathbf{r}_C - \mathbf{r}_O$ with the vector directed toward the carbonyl carbon atom. Once again, a P_2 value of 1 indicates a sharp orientation along the interface normal, while a value tending toward -0.5 demonstrates orientation in the interfacial plane; an average value of zero implies a random orientation. Furthermore, concerning the angle between the interface normal and an individual bond vector, a positive value indicates the carbon atom to be in closer proximity to the vapor phase than the oxygen; a negative value indicates the opposite.

Figure 7-18 shows the orientational order parameter profile for the five individual carbonyl bond vectors along the chain. The end vectors represent the two outermost groups; the middle vectors are the internal groups. The figure shows a distinct grouping of the vectors with respect to orientational behavior. The two end vectors show a significant orientation along the interface normal, while the inner bonds are either in the interfacial plane (this is probably more a small magnitude fluctuation from the planar orientation) or protruding slightly out of it. Furthermore, based on the P_1 profiles shown in Figure 7-19, we see that the end carbonyl oxygen atoms are oriented towards the vapor phase, while for the internal atoms, the degree of orientation out of the interfacial plane is smaller (and relatively insignificant based on the P_2 profiles). The behavior of the end carbonyl groups seems to be counterintuitive, particularly based on the results of the methyl acetate simulations showing that at the free interface, the carbonyl groups tend to orient with the oxygens situated away from the vapor side. Furthermore, if one considers the notion of the system minimizing uncom-

pensated charge, the more favorable conformation would be to have all the carbonyl oxygen atoms orienting towards the bulk (or in the very least, away from the vapor phase), with the end and pendant methyl groups showing a greater tendency for orienting towards the vapor phase. At this point, it is difficult to say what the driving force for this orientational behavior is specifically, although a few factors may be involved. First, we note the possibility of the intramolecular potentials being deficient in terms of contributing the correct physics; that is, the the parameterized torsion and angle potentials could be totally inappropriate for this system as they are all based on parameterizations of small molecule systems. Second, there is the possibility of the intramolecular non-bond interactions contributing erroneously to the total physics. This indeed is not too implausible since, for example, the charges assigned on the associated groups are adjusted simply based on arguments of balancing total molecular charge to maintain neutrality. There is no *a priori* justification, *for this specific system*, why the charges *must* be those that are employed. Again, these were taken, along with the intramolecular Lennard-Jones interaction parameters, from fits to small molecule physical properties. Furthermore, we have made no allowance for the fact that possibly the intramolecular interactions may be slightly different than intermolecular interactions. Finally, there is the problem of a coupling between the two—the effect of the non-bond and bonded interactions which has been treated in a 'lumped' manner via the fits to internals distributions. At this point, we offer for the current model that at the free interface, the end carbonyl oxygen atoms tend to orient towards the vapor phase with the internal atoms showing a more expected behavior in orienting, on average, in the interfacial plane (or very slightly out of it due to thermal fluctuations). Interestingly, if one considers chain-length effects, one may find that for longer chains, the effect of the ends becomes negligible and so this *anomalous* behavior of the end oxygen atoms would be inconsequential to the total energetics of the system, particularly if the chains remain in an elongated conformation. In the same manner, however, with a longer chain, end segregation effects may become more pronounced; in this case, the surface would be dominated by chain ends. In this case, one might expect to observe the opposite behavior from end carbonyl

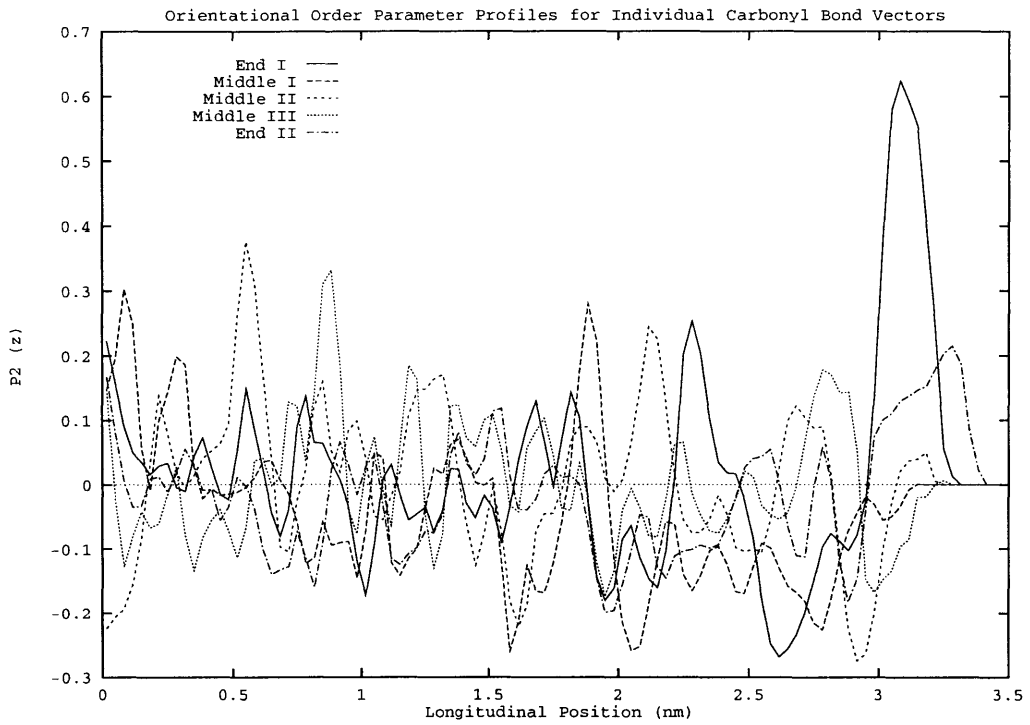


Figure 7-18: Symmetrized order parameter profile of the individual C=O vectors

groups. The above comments are quite speculative, and only by rigorous testing of the potential and further simulations (using longer chains, different contact media) can one be obtain a more precise view of the interfacial physics for this material.

7.6 Conclusions

This chapter has addressed an oligomer system selected as a model for polylactic acid (PLA). Based on roughly one nanosecond fo simulation time (production phase of simulations), preliminary data indicates that the surface chains orient, on average, in the interfacial plane; furthermore, there is a trend toward chains being more extended in this region relative to the bulk as evidenced by longitudinal profiles of the molecular radius of gyration and end-to-end vector. This observation is in keeping with previous simulations (MD and Monte Carlo) and mean-field theories for alkane chains interacting via Van der Waals or hard-sphere potentials. The interesting aspect here is that we see a similar effect for a system with a strong electrostatic interaction.

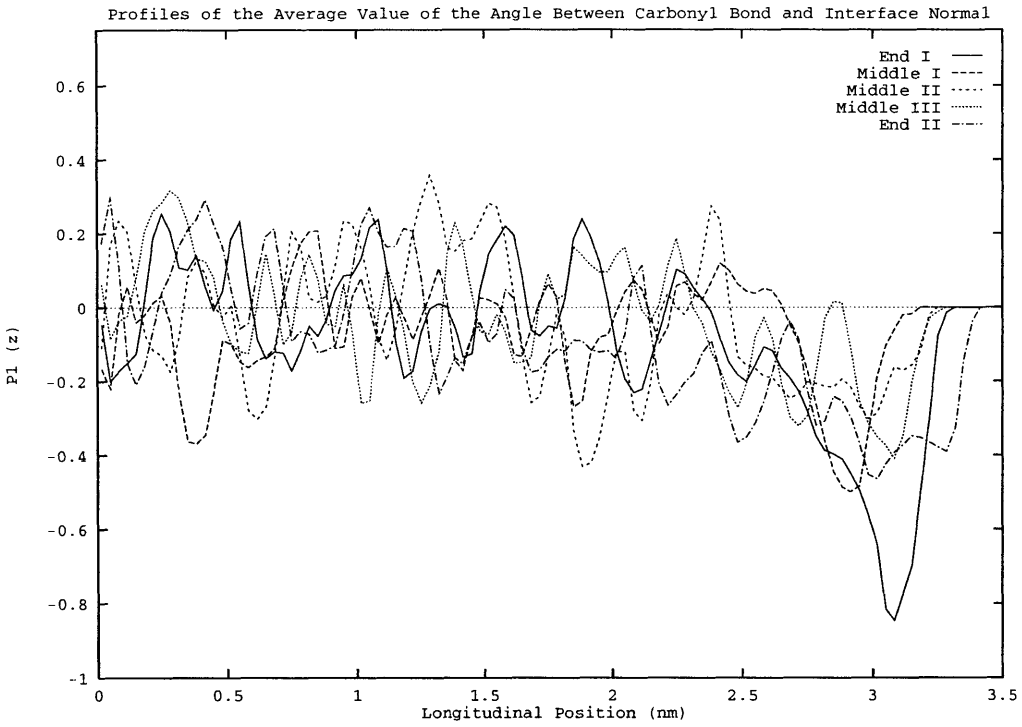


Figure 7-19: Symmetrized profiles of the average angle between interface normal and C=O vector

Figure 7-20: Proposed average molecular structure of interfacial oligomer chains

However, we fail to see a region of chains orienting normal to the interfacial plane as observed in the same previous MD and Monte Carlo simulations; this is also not observed in mean-field theories. The results of the present simulations also depart from earlier simulations and theoretical work with respect to the absence of chain-end segregation effects. The segregation of chain ends is accepted as the result of the interplay between entropic and enthalpic losses and gains; we note that both simulations and theory tend to predict this phenomenon for systems lacking strong dipolar character. The driving force for the observed surface enhancement in chain ends is the minimization of interfacial free energy through the minimization of enthalpic and entropic penalties for introducing a species into the interfacial region. Accommodating a chain end in the lower density interfacial region entails a smaller cohesive energy loss compared to a middle segment due to the *connectivity* of the latter with a *greater* number of neighbors [170]. Moreover, there is a less severe conformation restriction in having an end jutting into the interface instead of a larger middle portion of the chain; this is intimately related to the conformational entropy of the system [170]. We note here that in the case of the lattice model, the surface enhancement becomes more dramatic as one moves from an n-decane system to a 50-unit poly(dimethylsiloxane) system; for the n-decane system, 69 percent of the segments in the interface are chain ends (with ends making up 17 percent of the total chain mass), while in the PDMS case, 51.5 percent of the interfacial population is chain ends despite the fact that only 0.8 percent of the total polymer mass is represented by the ends [170]. This is an interesting observation, as it tends to suggest that the chain length selected for this study may be too short for such an effect to be strongly manifest. Furthermore, lattice Monte Carlo simulations of the melt-vacuum interface of a lattice polymer suggest that at a constant total bead density, the conformational structure of longer chains is more significantly affected than that of shorter chains [148]. In light of these observations, it would be interesting to determine whether there is associated a 'critical' chain length above which such structural phenomena are seen within the interfacial region. We do note that experimentally, there is evidence that contradicts this hypothesis. For example, in the case of polystyrene with perfluorinated

end groups, that there is an excess of perfluoro end groups at the interface that is inversely proportional to the chain length; these results are based on ion scattering spectroscopy (ISS), static secondary ion mass spectroscopy (static SIMS), and x-ray photoelectron spectroscopy (XPS) [171].

A further issue of contention in comparing the present results to previous work is the influence of the energetic characteristics of the constituents of the chain. For example, it is not hard to imagine that if a particular site on the chain is a high-energy site, the system would incur a large energetic penalty by forcing it to the low-density region, where essentially, it would not be sufficiently 'energetically compensated'; and in particular, if it is a high-energy end, one should not be too surprised to observe a suppression of this end segregating to the surface. In the present case, we employ a model with strong partial charge assignments; the energetic penalty of having uncompensated charge jutting toward the vapor phase may be sufficient to prevent chain-end segregation. Adding to this vane of thought are experimental observations. Consider that x-ray photoelectron spectroscopy and neutron reflectivitty measurements on fluorocarbon-capped polystyrenes show a surface excess of the low-energy fluorocarbon ends with a simultaneous lowering of surface tension [172]. On the other hand, for polystyrenes terminated with high-energy components such as *carboxylic acid* end groups, a surface *depletion* is observed [172].

Finally, with respect to the orientation of individual functional groups, and in particular the carbonyl groups, we assert that our model seems to suggest that there are two characteristic orientational behaviors associated with the *end* and *internal* carbonyl groups. We observe that the carbonyl vectors associated with each chain end are preferentially oriented towards the vapor phase–that is, the carbonyl oxygen atom is directed towards the low-density region. The vectors along the 'middle' section of the oligomer chain on average lie in the interfacial plane. This result seems counter-intuitive in light of the results of orientational behavior of the carbonyl vector in methyl acetate. At this time, we can only offer the possibility that our potential model suffers from severe deficiencies. Furthermore, we must actually determine experimentally the nature of the ester surface in order to assess the validity of the

model.

Chapter 8

Conclusions and Future Work

8.1 Recapitulation of Goals

This thesis has attempted to study the interfacial structure of small-molecule and oligomeric esters. The motivation for such an endeavor stems from the importance of the nature of the interface in biomedical applications, and in particular, where foreign materials are introduced into environments composed of aqueous media containing various and sundry physiological components (proteins, ions, cells, etc.). Furthermore, this work has addressed the structure and, in part, the thermodynamics of the liquid-liquid and liquid-vapor interfaces of a small molecule ester (methyl acetate) as well as the liquid-vapor interface of an ester oligomer representing PLA via molecular dynamics simulations. Toward this end, suitable models to represent the relevant physics of ester molecules have been selected, tested, and validated. The results of simulations provide information on the molecular-level structure of the ester interface in terms of orientations of particular functional groups, spatial distributions of molecular components, and the response of these to changes in contacting medium (in the present work, from a vapor to water medium); in parallel, the effect of the interface on the contacting medium is studied. Furthermore, the conformational distributions of bond, torsion, and out-of-plane angles have been studied in order to derive a clearer picture of molecular geometry in the interface relative to the bulk. Fluid structure has been investigated via analysis of pair correlation functions in the bulk and interfacial regions; these have been used to determine hydrogen-bonding interactions in systems including aqueous media. In terms of interfacial thermodynamics, the surface tension of methyl acetate is computed.

8.2 Major Findings and Conclusions

Along the road to ultimately investigating the interfaces of an oligomer model of polylactic acid ester (PLA), several topics have been broached, each providing some interesting and important finding allowing the next step along the journey. Chapter 3 discusses the selection and validation of a forcefield for molecular dynamics simula-

tions of methyl acetate in the condensed state. Based on the criteria that the chosen forcefield efficiently and accurately reproduces experimental conformational energy differences and barrier heights between rotational conformers, as well as reproduces experimental enthalpies of vaporization, the OPLS model was chosen as the potential for use in the subsequent small-molecule liquid-vapor and liquid-liquid simulations. An interesting finding of the work on bulk liquid energetics is the discrepancy between the reaction field and Ewald summation approaches to accounting for electrostatics in the prediction of bulk liquid enthalpy of vaporization for the case of a system with intramolecular non-bond interactions.

Chapter 4 discusses the MD simulations of the methyl acetate liquid-vapor interface using the OPLS potential. This is the first such study to be reported (manuscript submitted) in the literature. The work of this chapter shows that the OPLS model does indeed allow the evolution of a stable interfacial system defined by a molecularly sharp interface (typical of small-molecule interfaces resulting from forcefield calculations as reported in the literature) with a 10-90 thickness of 6 angstroms. A significant structuring is apparent at the interface, even for this small molecule. The carbonyl carbon - carbonyl oxygen bond vector align antiparallel to the interface normal with the oxygen adopting an orientation away from the vapor phase. In terms of interfacial thermodynamics, the atomic and molecular surface tensions of 22.7 and 21.06 dyne/cm are within 10 and 15 percent, respectively, of the experimental value of 24.73 dyne/cm. This inadequacy of the parameter set is not too surprising as interfacial thermodynamics, and in particular surface tensions, have traditionally proved to be intractable via molecular simulations, especially in the case of polymers and oligmers.

Chapter 5 details results of MD simulations of the methyl acetate / water liquid-liquid interface using the OPLS and SPC potential for the ester and water, respectively. The results of this work demonstrate that the applied potentials are sufficient to allow the evolution of a stable liquid-liquid interface with a 10-90 thickness of 7.45 angstroms (based on the water density profile). More importantly, we observe a change in interfacial structure of the ester and the water relative to the liquid-vapor

interface (for both species). For the ester, we observe the reorientation of the carbonyl oxygen more towards the aqueous layer, the opposite to the orientation adopted at the liquid-vapor interface as seen in Chapter 4. Furthermore, the SPC water model captures a rich surface structuring of the water molecules, again showing the reversal of molecular orientation from the liquid-vapor to the liquid-liquid interface. This mutual effect of the presence of each species on the other is driven by hydrogen bonding interactions evidenced by the first peak in the carbonyl oxygen - water hydrogen radial distribution functions; furthermore, the angle formed by the carbonyl oxygen, hydrogen, and water oxygen prefers a value of 160.1 degrees, which is close to the value of this angle computed for gas-phase complexes of water and methyl acetate via density functional theory.

Chapter 7 addresses the oligomer system selected as a model for polylactic acid (PLA). Based on roughly one nanosecond of simulation time (for the production phase of the simulation), preliminary data indicates that the surface chains orient, on average, in the interfacial plane; furthermore, there is a trend toward chains being more extended in this region relative to the bulk. This behavior is in contrast to prior MD simulations of alkane liquid-vapor interfaces which show that chains adopt a configuration in which the ends jut towards the vapor region, aligned more along the interface normal. The middle segments align more in the plane of the interface. In the present case, we observe a slight tilt which allows for the slight longitudinal separation of atomic species seen in the component density profiles. It is difficult to say at this point whether this is just an artifact of the short simulation time, or whether this phenomenon is the equilibrium behavior.

8.3 Future Work

Although this thesis has attempted to shed light on some aspects of the interfaces of ester compounds, many questions can still be posed and many aspects of the work could be extended or refined. In this section, we offer some words to this effect.

Concerning the methyl acetate liquid-vapor interface, connection of the results

of molecular structure to experimental observations would prove a rigorous test of the potential model. Although molecular-level detail of liquid-vapor interfaces has been traditionally difficult to achieve (at least to the level where one can obtain detail of molecular bond orientation), recent advances in non-linear optical techinques, particularly sum frequency generation and second harmonic generation would be able to provide the complementary data. These techniques generate infrared spectra of the interface molecules; spectra computed from dynamics trajectories could be used to directly compare to experimental data. This type of analysis has been used in the study of water and methanol liquid-vapor interfaces. A further study would be to study the effect of the method used for electrostatics accounting; Ewald summation in two dimensions (or a derivative such as particle mesh ewald [173], which we have shown for bulk methyl acetate gives equivalent results to the reaction field but is still slower by a factor of two-thirds). This study would be relevant to all the simulations performed in this work. The motivation for this work is the discrepancy between the reaction field and Ewald results for the bulk liquid energetics study, as well as the fact that the reaction field assumes an isotropic dielectric medium around each point charge; the latter is certainly not the case when considering an interface and as such, the effects of this assumption need to be further studied via comparison to these more rigorous (yet more computationally intensive) techinques.

Concerning the *ab initio* work, only one test molecule was targeted for generating the requisite torsion profile. It would be more rigorous to study several compounds and determine a more general torsion profile. Furthermore, noting the discussion on the effects of electron correlation and its inclusion in traditional Hartree-Fock approaches to *ab initio* electronic structure calculations, it would be interesting to perform the next set of calculations using the newer density functional methods. These have been documented to yield equivalent if not superior results relative to Hartree-Fock theory in a fraction of the time. Moreover, the fidelity of the torsion potential in reproducing experimental IR spectrometric data should be investigated. From powerspectra computed as Fourier transforms of selected velocity autocorrelation function, one effectively has a spectrum to which to compare to. By varying the

parameters of the potential, one can pinpoint the wavelength of the associated mode and proceed with the analysis.

With respect to the oligomer simulations, several recommendations for future work arise. First, it appears that the current chain size is probably too small to observe any of the behavior of truly long-chain molecules. Thus, perhaps trying simulations of chains of 10 or 20 repeat units would prove more useful. Second, with respect to the results presented in this work, it is imperative to allow further simulation time to accumulate statistics. The results shown at present are obtained from simulations of roughly one nanosecond, whereas the timescale of relevance is on the order of two or three nanoseconds. Third, moving on to one of the goals of this research, one could perform simulations of the oligomer-water interface. This would be interesting in the sense of allowing one to observe the effect of the water on the orientations of individual bonds and molecules as a whole. Furthermore, one could also attempt to look at the dynamics associated with this phenomenon. This dynamical aspect of surface reorientation has been neglected for the small-molecule systems in this thesis. It is felt that the time-scales for such processes would be quite short; however, with the presence of chain connectivity induced by the oligomers, a longer time-scale is more practical.

8.4 Final Thoughts

Any undertaking such as a thesis begins with a simple question and inevitably ends with orders of magnitude more; such is the nature of the beast. The work presented in the preceding pages is not immune from such plagues. In attempting to study the structure of the interface on a molecular level, of what I initially believed to be a relatively simple system, I realized the complexity of the task at hand. Invariably, this complexity begs of simplifications–in models, assumptions, and approaches. With respect to models, one of the biggest assumptions required is that of what interaction potential model to use; unfortunately, just about all that emerges from subsequent simulations hinges on the accuracy of this model. At present, I feel that

the state of the art is not at the point where one can take a generic model, or even a model designed for a specific molecular (or even atomic system in the case of simple atomic crystals), and apply it to the study of all properties of the system, with particular interest in the quantitative predictive abilities of the model. The situation is further complicated by the fact that in implementing the potential itself, one adopts methods needed to facilitate the computation. Hence, we are adding more assumptions/simplifications to the system. In the present case, one can look to the electrostatic nature of the system and consider the use of the reaction field approach a gross approximation with respect to a study of the interface. For rigor, it would be one small step to investigate the same system using an alternative technique, perhaps Ewald summation, or some more efficient algorithm such as particle mesh or cell-multipole; however, these themselves introduce certain approximations and assumptions, again adding to the convolution of the problem.

Returning to the issue of quantitative accuracy briefly. In part, this work has shown the behavior of 'water' molecules at the liquid-vapor and water-ester interface. Specifically, distributions of dipole moment vectors and O-H bond vectors have been computed. In interpreting this data, one must always be wary of the effect of the potential used. By this, I mean that we see a general, qualitative trend, for instance, of one O-H vector jutting into the vapor and the other in the plane of the liquid-vapor interface. However, the average *value* of the angle of orientation of this configuration really makes no sense, or is irrelevant in the sense that if one were to use another model for water, there is the possiblity of determining a different value, or maybe a set of values (multimodal distribution). In the final analysis, what we learn from a simulation such as those presented in this work, are general, fundamental behaviors of the model systems we *plug* into the computer. To transfer these behaviors to *real* systems is quite tenuous and requires a tremendous leap of faith in the assumptions and simplifications underlying all results (although from the tone of this thesis, it seems as though I claim to have 'modelled' PLA oligomers on the computer!).

The point I would like to argue is that I believe that for complex systems, simulations are capable of providing information about very general phenomena; to seek

high quantitative fidelity (with respect to experiment) with some 'universal' or 'transferrable' model is asking too much, particularly when this fidelity is with respect to a spectrum of properties not necessarily related in any way (and as the complexity of the system increases, one may as well forget about rigorous accuracy in all property estimations—one only has to consider the trials and tribulations of modelling the range of properties for water (bulk, interface, supercritical phase, solid phase) with a single model). At the heart of the matter is probably an understanding of how to correctly place in a model the fundamental underlying physics—and by this physics I mean the fundamental physical mechanisms which determine the properties of a *class* of materials; this can be further refined to a *multi-scale* perspective, where one now considers the relevant physics over appropriate length and time scales (this is currently an active and interesting area of research).

In the context of the chemical engineering discipline, I describe this work as a *miniscule* step in understanding the interfacial behavior of one class of materials, that of the simple polyesters typified by PLA. As a nucleation point for future study of this class of materials via simulations, I believe this thesis has contributed a model which *must* be more rigorously tested and refined before one can say that it has any pretense of representing the *true* polyester system. However, along the way to determining this model, some very interesting interfacial phenomena have been encountered, and these for systems which are important industrially (small-molecule esters are used in diverse applications from fragrances to flavorings to starting points for making high-tech fabrics such at those used by elite (and not-so-elite) athletes).

From an engineering perspective, this work has demonstrated, I think, that molecular modelling techinques can be a practical tool in gaining insight into materials properties not experimentally accessible (at least without much difficulty and debilitating cost). This is nothing new; one only has to survey a small fraction of the literature to come across similar investigations. The novelty comes from applying existing methods (of course, modified to fit certain needs) to a new, practical system.

I must conclude with an apology if the above statements sound too grandiose in describing the contributions of this work. Moreover, I refer to a recent article ad-

dressing the question of the predictive value of computer simulations, with specific focus on the properties of that ubiquitous and ever-elusive fluid, water [174]. The author paints a very pessimistic picture of the art of molecular simulations as a predictive tool. The description of "...microscopic effects by classical potentials, and, in particular, the description of hydrogen bonding, is *simplistic* and hopelessly naive...", according to this esteemed researcher [174]. Furthermore, in answer to the question of why, then, computer simulations have become so popular, the author adds, '...because of their conceptual simplicity. Such simulations allow one to solve almost every important problem in the theory of water simulations, however complex, using skills which are not much different from those necessary for computerized book-keeping. All theoretical and experimental information of condensed state physics and chemistry accumulated during the last two centuries is considered as practically unnecessary in many of such calculations," [174]. Well, whether this is too harsh a judgment on the state of the art and its application is open to intense debate. In the final analysis, my hope is that this work is not received as merely glorified book-keeping.

Appendix A

Applied Quantum Chemistry

A.1 Introduction

Quantum mechanics tells us that the energy and properties of a molecular stationary state are given by the solution of the non-relativistic, time-independent Schrodinger partial differential equation [45]

$$\mathcal{H}\Psi = E\Psi \tag{A.1}$$

where $\mathcal{H}$ is the Hamiltonian operator whose operation on the stationary wavefunction Ψ gives the total system Energy, E. The energy E itself is relative to that of the state where the system particles, both nuclei and electrons, are infinitely separated and at rest. Ψ is the stationary wavefunction dependent on the coordinates of all particles and also spin coordinates which take on finite values according to the projection of the spin angular momentum vector along a particular direction. Central to the probabilistic interpretation of quantum mechanics is the identification of Ψ^2 or $|\Psi|^2$ if Ψ is complex is a measure of the probability distribution of the particles within the molecule [45].

The Hamiltonian $\mathcal{H}$ is the sum of kinetic and potential contributions for a system of N electrons and M nuclei in atomic units is[47]

$$\mathcal{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \quad \text{(A.2)}$$

This equation requires some description in terms of the geometrical definitions used. The particle coordinates are given by position vectors $\mathbf{R}_A$ and $\mathbf{r}_i$ which are the nuclear and electronic position vectors, respectively. The distance between the i'th electron and A'th nucleus is $r_{iA} = |\mathbf{r}_{iA}| = |\mathbf{r}_i - \mathbf{R}_A|$. The distance between the i'th and j'th electrons is $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, and similarly the separation between nuclei A and B is $R_{AB} = |\mathbf{R}_A - \mathbf{R}_B|$ [47] .

In Equation A.2, M_A is the ratio of the mass of nucleus A to the mass of an electron, and Z_A is the atomic number of nucleus A. The Laplacian operators ∇_i^2 and ∇_A^2 signify differentiation with respect to the electronic and nuclear coordinates, respectively. In Cartesian coordinates, the operators are [45]

$$\nabla^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \quad \text{(A.3)}$$

The first and second terms in Equation A.2 represent the electronic and nuclear kinetic energies, respectively. The third term is the coulomb potential between electrons and nuclei, while the fourth term is the coulomb interaction between electrons and the fifth term the coulomb interaction between nuclei. The coulomb terms are generally referred to as the potential energy and the first two terms are the kinetic energy if a division of the Hamiltonian into these classes of contributions (as in classical dynamics) must be made [45].

A.2 Decoupling of Nuclear and Electronic Motion: Born-Oppenheimer and Consequences

Solving the Schrodinger equation presents a formidable challenge, and to achieve a solution for practical systems, simplifications must invariably be made. The starting

point for all attempts at solving this problem is the separation of the nuclear and electronic motions which ultimately leads to two sub-problems. This is the adiabatic or Born-Oppenheimer approximation [45, 46]. Since the nuclei are three orders of magnitude heavier than electrons, nuclear motion can be effectively neglected and at any instant, the electrons can be considered moving in the fixed field of the nuclei. In other words, the electrons in a molecule are able to respond instantaneously to any change in nuclear positions, thus making it reasonable to take the electron distribution to depend only on the instantaneous nuclear positions and not on their velocities. Thus, the general problem can be first posed as an electronic problem–solving for the electronic energy as a function of relative nuclear coordinates. This effective electronic energy is then used as a potential for the nuclear motion. For a diatomic molecule, there is only one relative nuclear coordinate, and the effective electronic energy is given by a curve. For polyatomic systems, the effective electronic energy dependence on relative nuclear coordinates is represented as the potential energy surface [45] .

With the Born-Oppenheimer approximation, the nuclear kinetic energy term in the Hamiltonian can be neglected, and the nuclear repulsion term can be considered a constant (for a given nuclear configuration). Note that since this is a matrix eigenvalue problem, the presence of a constant in the operator does not affect the eigenvectors (or electronic wavefunctions) and contributes a constant to the eigenvalues (effective electronic energies) [47]. An electronic hamiltonian is now defined as [47, 45]

$$\mathcal{H}_{elec} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{A.4}$$

The eigenvalue problem now becomes

$$\mathcal{H}_{elec}\Phi_{elec} = \mathcal{E}_{elec}\Phi_{elec} \tag{A.5}$$

where the electronic wavefunction,

$$\Phi_{elec} = \Phi_{elec}(((\mathbf{r}_i);(\mathbf{R}_A)) \tag{A.6}$$

describes the electronic dynamics, depending explicitly on the electronic positions and parametrically on the nuclear coordinates; likewise, the electronic energy is also a parametric function of the nuclear coordinates, by parametric meaning that for every nuclear configuration, the electronic energy is a *different* function of the electronic coordinates [47, 45].

The total energy is now just the electronic energy plus the constant nuclear repulsion term. To now solve the nuclear problem, since the electrons are moving much faster than the nuclei, the electronic coordinates are replaced by average values over the wavefunction, thereby generating a nuclear Hamiltonian for the motion of the nuclei in the *average* field of the electrons [47, 45],

$$\begin{aligned}
\mathcal{H}_{nuc} &= -\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2 + < -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} > + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \\
&= -\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2 + \mathcal{E}_{elec}(\mathbf{R}_A) + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \\
&= -\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2 + \mathcal{E}_{tot}(\mathbf{R}_A) \qquad \text{(A.7)}
\end{aligned}$$

The total energy $\mathcal{E}_{tot}$ is the potential energy surface for the nuclear dynamics. Solutions to the nuclear Schrodinger equation

$$\mathcal{H}_{nucl}\Phi_{nucl} = \mathcal{E}\Phi_{nucl} \qquad \text{(A.8)}$$

describe the vibration, rotation, and translation of the molecule, and the $\mathcal{E}$ gives the electronic, vibrational, rotational, and electronic energy [47].

Applied quantum chemical methods generally only consider the electronic, searching for the potential energy surface on which the nuclear dynamics evolve; this is certainly the case for the current work, as *ab initio* methods are applied to study the torsional profile of a model compound, the torsional profile representing the changing nuclear configurations for which the electronic problem is solved. Thus, for the remaining discussion of this Appendix, only the electronic problem is considered.

A.3 Molecular Orbital Theory

The solution of the electronic Schrodinger equation begins with some approximation to the total wavefunction. In general, this amounts to writing the wavefunction as an expansion in some primitive functions which incorporate the physics of the structure; the accuracy of the calculation then becomes a matter of the adequacy of the functions in describing fundamental physics and the truncation of the expansion (an infinite expansion is the limiting case of retrieving the true solution to the electronic problem). Molecular orbital theory approaches the electronic problem by using one-electron functions, or *orbitals*, to approximate the full wavefunction. A one-electron wavefunction, or single-particle molecular orbital is a function of the coordinates and spin of the particle. The spatial component is dependent only on the coordinates and its square modulus is interpreted as the probability density distribution of the particle in space (or within the molecular environment). If the spatial orbitals were complete, any function can be expanded exactly in these functions as [45]

$$f(\mathbf{r}) = \sum_{i=1}^{infty} a_i \psi_i(\mathbf{r}) \tag{A.9}$$

with the a_i constants representing the projection of the original function onto the spatial functions. As shown, the set of functions would have to be infinite to be complete; in practice, one works only with finite sets which span only a portion of the total space represented by the exact function, and the solutions obtained with any given set are exact within the subspace spanned by the finite set of spatial orbitals [47].

The spin component depends on the spin coordinate which takes a value of $\pm\frac{1}{2}$; this measures the z-component of the spin angular momentum. The spin functions for spin aligned along the positive and negative z-axes are

$$\alpha(\frac{1}{2}) = 1 \quad \alpha(-\frac{1}{2}) = 0 \tag{A.10}$$

$$\beta(\frac{1}{2}) = 0 \quad \beta(-\frac{1}{2}) = 1 \tag{A.11}$$

The complete single-particle wavefunction, the spin orbital, is thus

$$\chi(\mathbf{r}, \eta) = \psi(\mathbf{r})\alpha(\eta) \tag{A.12}$$

$$\chi(\mathbf{r}, \eta) = \psi(\mathbf{r})\beta(\eta) \tag{A.13}$$

Note that both the spatial and spin functions are orthogonal [45].

Now, the question of what an appropriate N-electron wavefunction arises. One option is the simple product of N single-particle spin orbitals, referred to as a Hartree product [47],

$$\Psi^{HP} = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)\cdots\chi_N(\mathbf{x}_N) \tag{A.14}$$

where the $\mathbf{x}_i$ represent the spatial and spin coordinates taken together as the particle coordinates. Note that the Hartree product is an independent electron, or uncorrelated wave-function because the square modulus of Ψ^{HP}, which is the simultaneous probability of finding electron one in volume element $d\mathbf{x}_1$ centered at $\mathbf{x}_1$, electron two is $d\mathbf{x}_2$, etc., is equivalent to the product of the individual probabilities [47]

$$|\Psi^{HP}|^2 = |\chi_i(\mathbf{x}_1)|^2 d\mathbf{x}_1 \, |\chi_j(\mathbf{x}_2)|^2 d\mathbf{x}_2 \cdots |\chi_k(\mathbf{x}_N)|^2 d\mathbf{x}_N \tag{A.15}$$

that electron one is in $\mathbf{x}_1$ times the probability that electron two is in $d\mathbf{x}_1$, etc.

Assuming independent electrons as well as a Hamiltonian that is the sum of individual electron hamiltonians, the Hartree product approximation to the wavefunction is deficient in the sense that it fails to obey the antisymmetry principle for fermionic particles (a generalization of the Pauli Exclusion principle) which requires that the wavefunction be antisymmetric with respect to interchange of spatial and spin coordinates of two particles. In other words, the Hartree product distinguishes between

individual particles, whereas the antisymmetry principle does not call for this [47].

In response to this, the traditional procedure is to construct a determinantal form for the wavefunction which obeys the antisymmetry principle. The Slater determinant representing the full many-electron molecular orbital wavefunction for the closed-shell ground state of a molecule is then [45]

$$\Psi_{determinant} = (N!)^{-1/2} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix} \tag{A.16}$$

This determinant contains N electrons occupying N spin orbitals without specifically placing any one electron in a particular orbital. The rows determine the electron, and the columns the spin orbital. The Slater determinant meets the requirement of antisymmetry since changing the labels of any two electrons means interchanging the rows of the determinant; this leads to a sign change of the determinant (and thus the wavefunction) as required. Furthermore, having two electrons in the same spin orbital corresponds to two equal columns of the determinant, and this lead to a vanishing determinant/wavefunction. This is in keeping with the Pauli exclusion principle which forbids any two electrons from occupying the same spin orbital.

Expanding the determinant gives a sum of products of spin orbitals [45]

$$\Psi_{determinant} = \sum_{P} (-1)^P \mathcal{P}[\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)\cdots\chi_N(\mathbf{x}_N)] \tag{A.17}$$

where $\mathcal{P}$ is a permutation operator which reassigns the coordinates $\mathbf{x}_i$ according to any of the $N!$ possible permutations among the N electrons, and $(-1)^P$ is ± 1 for even or odd permutations.

Finally, the Slater determinant is normalized with the factor $(N!)^{-1/2}$.

A.3.1 Slater Determinants and Exchange Correlation

Although the Slater determinantal form of the full molecular orbital addresses the issues of antisymmetry and Pauli exclusion, the single determinantal form given above introduces exchange correlation effects. That is, the motion of two electrons with parallel spins is correlated (as it should be); however, the motion of electrons of opposite spins is uncorrelated, and for this reason, the single Slater determinant is considered an uncorrelated wavefunction [47].

This is easily seen by comparing the probability distribution functions for the cases where two electrons with same spin occupy different spatial orbitals and where both have opposite spins in different spatial orbitals [47].

For the first case, expanding the full 2-electron determinant to get the modulus gives

$$|\Psi^2|d\mathbf{x}_1 d\mathbf{x}_2 = |\psi_1(\mathbf{r}_1)\alpha(\omega_1)\psi_2(\mathbf{r}_2)\beta(\omega_2) - \psi_1(\mathbf{r}_2)\alpha(\omega_2)\psi_2(\mathbf{r}_1)\beta(\omega_1)|^2 d\mathbf{x}_1 d\mathbf{x}_2 \quad \text{(A.18)}$$

which gives the simultaneous probability of electron one being in $d\mathbf{x}_1$ and electron two being in $d\mathbf{x}_2$. By integrating over the spin coordinates (the spin functions are orthogonal) gives the simultaneous probability of finding electron one in $d\mathbf{r}_1$ and electron two in $d\mathbf{r}_2$; more importantly, the orthogonality of the spin functions forces the contributions from the cross terms to vanish, leaving

$$\begin{aligned} P(\mathbf{r}_1, \mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 &= \int d\omega_1 d\omega_2 |\Psi|^2 d\mathbf{r}_1 d\mathbf{r}_2 \\ &= \frac{1}{2}[\psi_1(\mathbf{r}_1)|^2\psi_2(\mathbf{r}_2)|^2 + \psi_1(\mathbf{r}_2)|^2\psi_2(\mathbf{r}_1)|^2 d\mathbf{r}_1 d\mathbf{r}_2 \end{aligned} \quad \text{(A.19)}$$

The first term is the product of the probability of finding electron one in $d\mathbf{r}_1$ at $\mathbf{r}_1$ multiplying the probability of finding electron two in $d\mathbf{r}_2$ at $\mathbf{r}_2$ with electron one in ψ_1 and electron two in ψ_2 . The second term is the same, only now the electrons have switched the spatial orbitals. Due to the indistinguishability of the two electrons, the

final probability is the average of the two probabilities. The form of these probabilities is that of the independent electron picture, so we see that for opposite spin electrons, there is no spatial correlation. The deficiency is pronounced when one considers that $P(\mathbf{r}_1, \mathbf{r}_1) \neq 0$ which means that *there is a finite probability of finding electrons of opposite spin at the same point in space.*

For the case where the electrons are of same spin, the probability $P(\mathbf{r}_1, \mathbf{r}_2)$ has cross terms [47]

$$\begin{aligned} P(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{2}[\psi_1(\mathbf{r}_1)|^2\psi_2(\mathbf{r}_2)|^2 + \psi_1(\mathbf{r}_2)|^2\psi_2(\mathbf{r}_1)|^2 \\ &\quad - \psi_1^*(\mathbf{r}_1)\psi_2(\mathbf{r}_1)\psi_2^*(\mathbf{r}_2)\psi_1(\mathbf{r}_2) + \psi_1(\mathbf{r}_1)\psi_2^*(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\psi_1^*(\mathbf{r}_2)] \end{aligned} \quad \text{(A.20)}$$

which signify the correlation. The cross terms arise from the non-orthogonality of the parallel spin functions. A *Fermi hole* is created around an electron in which another electron of same spin cannot be found. Thus, within the Slater determinantal description of the full molecular orbital wavefunction, the motion of electrons of same spin is correlated and of those with opposite spins is uncorrelated [47]. The methods used to estimate the correlation correction will be addressed below.

A.4 Hartee-Fock Theory and Variational Solution of the Electronic Problem

Now that a form for the full molecular orbital has been set up, this section deals with the theory of the solution of the electronic Schrodinger equation. For the present discussion, we note in brief that each of the molecular orbitals used in the construction of the Slater determinant is expressed as a linear combination of a finite set of N prescribed one-electron functions called basis functions [47],

$$\psi_i = \sum_{i=1}^{N} c_{\mu i}\phi_\mu \quad \text{(A.21)}$$

where the $c_{\mu i}$ are the molecular orbital expansion coefficients. More will be said about the basis functions later. At present, we focus on the expansion coefficients and how they are determined via Hartree-Fock calculations; it is the domain of Hartree-Fock to determine these coefficients [47].

Hartree-Fock theory is founded on the variational method of quantum mechanics. For any antisymmetric, normalized function of electronic coordinates, Φ, the energy expectation value corresponding to this wavefunction is [45, 47]

$$E' = \int \Phi^* \mathcal{H} \Phi d\tau \tag{A.22}$$

where the integration is over all electronic coordinates, and the asterisk denotes complex conjugation. If Φ is the exact wavefunction, Ψ , for the ground state of the system, it will satisfy the Schrodinger equation, and because of its normalization in space, E' will identically be equal to E, the exact electronic ground state energy,

$$E' = E \int \Psi^* \Psi d\tau = E \tag{A.23}$$

If Φ is an antisymmetric, normalized wavefunction other than the exact, then it can be shown that the variational energy, E' is higher than the exact energy, E [45, 47],

$$E' = \int \Phi^* \mathcal{H} \Phi d\tau > E \tag{A.24}$$

It follows that if the normalized, antisymmetric function Φ is the Slater determinant form introduced in the last section, then the variational energy associated with it will be higher than the exact value.

Now, the variational principle can be applied to compute the optimum expansion coefficients for the single-determinant wavefunctions. The approach is to defined a basis set of which the orbitals are written as linear combinations, and to adjust the coefficients of the expansion to minimize the expectation energy, E'. The optimum value of E' will be as close to the exact value within the confines of the single-determinant wavefunction and the basis set applied. In effect, the variational method gives the

'best' wavefunction of the form postulated (ie, represented as the linear expansion) in terms of the energy, by a minimization of E' with respect to the expansion coefficients; thus, the variational equations are generated with [45, 47]

$$\frac{\partial E'}{\partial c_{\mu i}} = 0 \quad (all\, \mu, i) \tag{A.25}$$

A.4.1 The Hartree-Fock Equations: Functional Variation and Canonical Form

The two major types of calculations performed with *ab initio* methods are restricted closed shell Hartree-Fock and unrestricted Hartree-Fock (open-shell) calculations. The equations for each approach differ slightly, but start from the same initial machinery. In this section, the Hartree-Fock equations are presented, beginning with the variational minimization of the single-determinant energy and leading up to the canonical Hartree-Fock equations. Along the way the Fock operator is introduced. For more detailed derivations, the reader is referred to the literature [45, 47].

Variational Minimization of the Slater Energy

For the single Slater determinant wavefunction

$$\Psi_{determinant} = \sum_{P}(-1)^{P}\mathcal{P}[\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)\cdots\chi_N(\mathbf{x}_N)] \tag{A.26}$$

the energy E' is a functional of the molecular spin orbitals, $[\chi_i]$. The Hartree-Fock equations are derived by minimizing the $E'[\chi_i]$ with respect to the spin orbitals subject to the constraint that the spin orbitals are mutually orthogonal,

$$\int d\mathbf{x}_1 \chi_a^*(\mathbf{x}_1)\chi_b(\mathbf{x}_1) = [a|b] = \delta_{ab} \tag{A.27}$$

The constraints are framed in the form [45, 47]

$$[\chi_a|\chi_b] - \delta_{ab} = 0 \tag{A.28}$$

Thus, the functional of spin orbitals $\mathcal{L}[\chi_a]$ is considered for functional variation

$$\mathcal{L}[\chi_a] = E'[\chi_a] - \sum_{a=1}^{N}\sum_{b=1}^{N}\epsilon_{ba}([\chi_a|\chi_b] - \delta_{ab}) \tag{A.29}$$

where the ϵ_{ab} are Lagrange multipliers. Their physical significance will be pointed out below.

The energy E' is the expectation value for the single Slater determinant energy

$$E'[\chi_a] = \sum_{a=1}^{N}[\chi_a|h|\chi_a] + \frac{1}{2}\sum_{a=1}^{N}\sum_{b=1}^{N}[\chi_a\chi_a|\chi_b\chi_b] - [\chi_a\chi_b|\chi_b\chi_a] \tag{A.30}$$

In the last equation, the bracketed terms are short-hand notation for one- and two-electron integrals encountered and are introduced now for clarity.

The one-electron integral notation is [47]

$$[\chi_a|h|\chi_a] = \int d\mathbf{x}_1\chi_a^*(\mathbf{x}_1)h(\mathbf{r}_1)\chi_a(\mathbf{x}_1) \tag{A.31}$$

where $h(\mathbf{r}_1)$ is a one-electron core-hamiltonian

$$h(\mathbf{r}_1) = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}} \tag{A.32}$$

which gives the one-electron kinetic energy and electron-nuclear potential energy. The two-electron integrals are [47]

$$[\chi_a\chi_a|\chi_b\chi_b] = \int \mathbf{x}_1\mathbf{x}_2\chi_a^*(\mathbf{x}_1)\chi_b(\mathbf{x}_1)r_{12}^{-1}\chi_b^*(\mathbf{x}_2)\chi_b(\mathbf{x}_2) \tag{A.33}$$

$$[\chi_b\chi_a|\chi_a\chi_b] = \int \mathbf{x}_1\mathbf{x}_2\chi_a^*(\mathbf{x}_1)\chi_b(\mathbf{x}_1)r_{12}^{-1}\chi_b^*(\mathbf{x}_2)\chi_a(\mathbf{x}_2) \tag{A.34}$$

The first integral represents the action of the coulomb operator and the second is the result of the exchange operator.

Now, the minimization of E' with respect to the spin orbitals and subject to the orthonormality constraints is performed by varying the spin orbitals by an arbitrary infinitesimal amount,

$$\chi_a \rightarrow \chi_a + \delta\chi_a \tag{A.35}$$

and setting the first variation in $\mathcal{L}$ to zero

$$\delta\mathcal{L} = \delta E' - \sum_{a=1}^{N}\sum_{b=1}^{N}\epsilon_{ba}\delta[\chi_a|\chi_b] = 0 \tag{A.36}$$

After some straightforward manipulation [47], this gives the non-canonical form of the Hartree-Fock equations:

$$[h(\mathbf{r}_1) + \sum_{b=1}^{N}\mathcal{J}_b(\mathbf{x}_1) - \mathcal{K}_b(\mathbf{x}_1)]\chi_a(\mathbf{x}_1) = \sum_{b=1}^{N}\epsilon_{ba}\chi_b(\mathbf{x}_1) \quad a = 1, 2, \cdots N \tag{A.37}$$

The quantity in square brackets is the Fock operator defined in terms of the one-electron core hamiltonian and the coulomb and exchange operators [47],

$$f(\mathbf{x}_1) = h(\mathbf{r}_1) + \sum_{b=1}^{N}\mathcal{J}_b(\mathbf{x}_1) - \mathcal{K}_b(\mathbf{x}_1) \tag{A.38}$$

And so, in a more compact form, the non-canonical Hartree-Fock eigenvalue problem is

$$f(\mathbf{x}_1)\chi_a = \sum_{b=1}^{N}\epsilon_{ba}\chi_b \tag{A.39}$$

The reason why the final result up to this point does not fit the standard (canonical) eigenvalue form is that the single determinant wavefunction composed of the sum of products of spin orbitals retains a flexibility in how the various orbitals are mixed, each combination of mixing giving rise to the same variational energy. Thus, the standard procedure is to find a unitary transformation matrix which allows the transformation of the old set of spin orbitals to a new set and to make sure that the Fock operator is invariant to such a transformation matrix. Finally, the Hermitian matrix of Lagrange multipliers is shown to be diagonizable under the action of some unitary transformation, thus allowing the Hartree-Fock equation to be rewritten in

terms of the canonical spin orbitals [47],

$$f\chi_a^{'} = \epsilon_{ba}^{'}\chi_a^{'} \tag{A.40}$$

or dropping the primed notation and understanding that we are dealing with canonical spin orbitals,

$$f\chi_a = \epsilon_{ba}\chi_a \tag{A.41}$$

Some comments on the physical interpretation of the canonical representation of the Hartree-Fock equations.

For a single determinant wavefunction, an expectation value is invariant to an arbitrary unitary transformation of the spin orbitals. The importance of this is that the *spin orbitals which make the total energy stationary are not unique*, — any set of spin orbitals generated from a unitary transformation of another set are valid within the single determinant description. Thus, the concept of 'localized' orbitals is not rigorously correct, although none less 'physical' than the notion of delocalized spin orbitals [47].

A.4.2 Closed-Shell Hartree-Fock: Restricted Spin Orbitals

The previous section introduced the Hartree-Fock equations in terms of a set of general spin orbitals, $[\chi_a]$. In this section, more specific orbital functions are considered for the actual calculation of the Hartree-Fock (HF) wavefunctions. The types of calculations for which equations are presented in this section are those for systems with closed electronic shells. The molecular states considered have only an even number of N electrons, with all electrons paired such that there are $n = N/2$ spatial orbitals which are doubly occupied. These are restricted spin orbitals which have the same spatial orbital for spin up (α) and spin down (β) electrons. Unrestricted calculations with separate spatial functions for α and β electrons [45, 47].

Closed shell calculations are calculated using the closed-shell Fock operator which is simply the spin orbital Fock operator introduced in the last section from which the

spin coordinates are integrated over. Without going into the details of the integration, the closed-shell Fock operator is obtained as [45, 47]:

$$f(\mathbf{r}_1) = h(\mathbf{r}_1) + \sum_a^{N/2} \int d\mathbf{r}_2 \psi_a^*(\mathbf{r}_2)(2 - \mathcal{P}_{12}) r_{12}^{-1} \psi_a(\mathbf{r}_2) \tag{A.42}$$

or, more compactly,

$$f(\mathbf{r}_1) = h(\mathbf{r}_1) + \sum_a^{N/2} 2J_a(\mathbf{r}_1) - K_a(\mathbf{r}_1) \tag{A.43}$$

where the closed-shell coulomb and exchange operators are defined as

$$J_a(\mathbf{r}_1) = \int d(\mathbf{r}_2) \psi_a^*(\mathbf{r}_2) r_{12}^{-1} \psi_a(\mathbf{r}_2) \tag{A.44}$$

$$K_a(\mathbf{r}_1)\psi_i(\mathbf{r}_1) = \int d(\mathbf{r}_2) \psi_a^*(\mathbf{r}_2) r_{12}^{-1} \psi_i(\mathbf{r}_2)] \psi_a(\mathbf{r}_1) \tag{A.45}$$

This gives the closed-shell eigenvalue equation as [45, 47]

$$f(\mathbf{r}_1)\psi_i(\mathbf{r}_1) = \epsilon_i \psi_i(\mathbf{r}_1) \tag{A.46}$$

Although this equation can be solved via numerical techniques, the standard approach to its solution is that of Roothan [175, 176] which converts the differential equation into a set of algebraic equations which can be solved via standard matrix techniques. The scheme begins by writing the spatial orbitals as linear expansions in a set of given basis functions, the expansion coefficients being solved for as the results of the calculation.

For the expansion

$$\psi_i = \sum_{\nu=1}^{K} c_{\nu i} \phi_\nu \tag{A.47}$$

one obtains a matrix equation for the expansion coefficients by substituting the expansion in the closed-shell Hartree-Fock equations,

$$f(\mathbf{r}_1)\sum_{\nu} c_{\nu i}\phi_{\nu}(\mathbf{r}_1) \;=\; \epsilon_i \sum_{\nu} c_{\nu i}\phi_{\nu}(\mathbf{r}_1) \tag{A.48}$$

By multiplying by $\phi^*_{\mu}(\mathbf{r}_1)$ on the left and integrating over the spatial coordinates of a single electron, the matrix equation emerges as

$$\sum_{\nu} c_{\nu i}\int d\mathbf{r}_1)\phi^*_{\mu}(\mathbf{r}_1)f(\mathbf{r}_1)\phi_{\nu}(\mathbf{r}_1) \;=\; \epsilon_i \sum_{\nu} c_{\nu i}\int d\mathbf{r}_1)\phi^*_{\mu}(\mathbf{r}_1)\phi_{\nu}(\mathbf{r}_1) \tag{A.49}$$

The preceding equation generates two matrices, the overlap and Fock matrices.

The overlap matrix, **S**, is

$$S_{\mu\nu} \;=\; \int d\mathbf{r}_1\phi^*_{\mu}(\mathbf{r}_1)\phi_{\nu}(\mathbf{r}_1) \tag{A.50}$$

It is Hermitian and represents the overlap between the generally non-orthogonal basis functions. The diagonal elements are unity and the off-diagonal elements are numbers in the range $0 \leq |S_{\mu\nu}| \leq 1$. This Hermitian property of the overlap matrix will be exploited later to diagonalize the matrix via a unitary transformation.

The Fock matrix,**F** is

$$F_{\mu\nu} \;=\; \int d\mathbf{r}_1\phi^*_{\mu}(\mathbf{r}_1)\{(\mathbf{r}_1)\phi_{\nu}(\mathbf{r}_1) \tag{A.51}$$

The Fock matrix is the matrix representation of the Fock operator introduced above,

$$f(\mathbf{r}_1) \;=\; h(\mathbf{r}_1) \;+\; \sum_{a}^{N/2} 2J_a(\mathbf{r}_1) \;-\; K_a(\mathbf{r}_1) \tag{A.52}$$

in the basis ϕ_{μ}.

The first term will give the core-Hamiltonian matrix representing the one-electron kinetic and nuclear potential energies,

$$H^{core}_{\mu\nu} \;=\; \int d\mathbf{r}_1\phi^*_{\mu}(\mathbf{r}_1)\langle(\mathbf{r}_1)\phi_{\nu}(\mathbf{r}_1) \tag{A.53}$$

where the one-electron operator $h(\mathbf{r}_1)$ is

$$h(\mathbf{r}_1) = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|} \tag{A.54}$$

The coulomb and exchange terms of the Fock operator give rise to the two-electron part of the Fock matrix which is presented here in terms of the density matrix, $P_{\lambda\sigma}$,

$$\begin{aligned} F_{\mu\nu} &= H_{\mu\nu}^{core} + \sum_a^{N/2} \sum_{\lambda\sigma} c_{\lambda a} c_{\sigma a}^* [2(\mu\nu|\sigma\lambda) - (\mu\lambda|\sigma\nu)] && \text{(A.55)} \\ &= H_{\mu\nu}^{core} + \sum_{\lambda\sigma} P_{\lambda\sigma} [(\mu\nu|\sigma\lambda) - \frac{1}{2}(\mu\lambda|\sigma\nu)] && \text{(A.56)} \\ &= H_{\mu\nu}^{core} + G_{\mu\nu} && \text{(A.57)} \end{aligned}$$

The two-electron integrals are defined as,

$$(\mu\nu|\lambda\sigma) = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_\mu^*(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1) r_{12}^{-1} \phi_\lambda^*(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2) \tag{A.58}$$

and the one-electron density matrix, $P_{\lambda\sigma}$ is

$$P_{\lambda\sigma} = 2 \sum_a^{N/2} c_{\mu a} c_{\nu a}^* \tag{A.59}$$

For all HF calculations, it is the two-electron integrals which require the most cpu time and thus pose the most difficulty in terms of calculation and manipulation (storage) during the computation. Also note that the Fock matrix depends on the density matrix, or equivalently, on the expansion coefficients [45, 47].

The Roothan equations are non-linear since the Fock matrix is dependent on the expansion coefficients via the density matrix. The solution is therefore iterative, and since the resulting molecular orbitals are generated from their own effective potential, the solution scheme is common known as self-consistent field (SCF) theory.

The integrated HF equations are now,

$$\sum_\nu F_{\mu\nu} c_{\nu i} = \epsilon_i \sum_\nu S_{\mu\nu} c_{\nu i} \quad i = 1, 2, \cdots, K \tag{A.60}$$

or more compactly,

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{A.61}$$

In the next section, the equations for unrestricted, open-shell systems are presented, following which the solution procedure is outlined. Finally, the basis functions commonly used are discussed.

A.4.3 Open-Shell Hartree-Fock: Unrestricted Spin Orbitals

The Roothan equations are modified for systems in which all electrons are not paired, leading to doublet (one extra α electron) and triplet (two extra α electrons) states. One type of molecular orbital theory commonly applied to such systems is spin-unrestricted Hartree- Fock (UHF) theory. Within the UHF formalism, spatial orbitals associated with the α and β spin functions are independent–that is, different. Consequently, there emerge two separate set of molecular orbitals, ψ_i^α and ψ_i^β (where $i = 1, 2, \cdots, N$). Thus, the doubly occupied orbital ψ_1 in restricted Hartree-Fock is now replaced by two functions, ψ_1^α and ψ_1^β. Note that since the UHF wavefunctions are composed of more basis functions, the UHF energy in the variational sense will be lower than the RHF energy [45, 177].

The two sets of α and β spatial orbitals are,

$$\psi_i^\alpha = \sum_{\mu=1}^{N} c_{\mu i}^\alpha \phi_\mu \,; \quad \psi_i^\beta = \sum_{\mu=1}^{N} c_{\mu i}^\beta \phi_\mu \tag{A.62}$$

The coefficients are optimized via functional variation leading to the UHF generalizations of the Roothan equations [45, 177],

$$\sum_{\nu=1}^{N} = (F_{\mu\nu}^\alpha - \epsilon_i^\alpha S_{\mu\nu}) c_{\mu i}^\alpha = 0 \quad , \mu = 1, 2, \cdots, N \tag{A.63}$$

$$\sum_{\nu=1}^{N} = (F_{\mu\nu}^\beta - \epsilon_i^\beta S_{\mu\nu}) c_{\mu i}^\beta = 0 \quad , \mu = 1, 2, \cdots, N \tag{A.64}$$

The requisite Fock matrices are defined as,

$$F^{\alpha}_{\mu\nu} = H^{core}_{\mu\nu} + \sum_{\lambda=1}^{N}\sum_{\sigma=1}^{N}[(P^{\alpha}_{\lambda\sigma} + P^{\beta}_{\lambda\sigma})(\mu\nu|\lambda\sigma) - P^{\alpha}_{\lambda\sigma}(\mu\lambda|\nu\sigma)] \tag{A.65}$$

$$F^{\beta}_{\mu\nu} = H^{core}_{\mu\nu} + \sum_{\lambda=1}^{N}\sum_{\sigma=1}^{N}[(P^{\alpha}_{\lambda\sigma} + P^{\beta}_{\lambda\sigma})(\mu\nu|\lambda\sigma) - P^{\beta}_{\lambda\sigma}(\mu\lambda|\nu\sigma)] \tag{A.66}$$

The density matrices are,

$$P^{\alpha}_{\mu\nu} = \sum_{i=1}^{\alpha_{occ}} c^{\alpha *}_{\mu i} c^{\alpha}_{\nu i} \tag{A.67}$$

$$P^{\beta}_{\mu\nu} = \sum_{i=1}^{\beta_{occ}} c^{\beta *}_{\mu i} c^{\beta}_{\nu i} \tag{A.68}$$

The overlap, core-Hamiltonian, and $(\mu\nu|\lambda\sigma)$ matrices are the same as those defined for the closed-shell Roothan equations.

A.5 Computational Approach

In this section is presented a very brief discussion of the implementation of the Hartree-Fock equations within an algorithmic scheme. As one would expect, there are various and sundry details involved with the efficient implementation of the Roothan-Hall equations, and the reader is referred to the literature for further discussion. As we will be concerned with geometry optimization calculations, the scheme for such computations is discussed.

Figure A-1 shows a schematic of the optimization of molecular geometry for a given basis set. The calculation begins with the specification of a starting geometry, usually in the form of a Z-matrix, the calculation of symmetry properties of the molecule, and the specification of a basis set. Following, all one-electron overlap, kinetic energy, and potential energy integrals, as well as the two-electron repulsion integrals are computed. All one-electron integrals are stored as matrices in, while only the non-zero elements of the two-electron integrals are sequentially stored with adequate labelling information to identify a particular integral. Prior to the SCF calculations, an initial

guess for the wavefunction and density matrix is made. The general approach for this is to use a wavefunction from a semi-empirical molecular orbital procedure, ideally, from a procedure which has been parameterized to the results of a particular *ab initio* basis set. Less favorable from a computatio nal view is that one can use a wavefunction generated from the diagonalized one-electron core-Hamiltonian. Next, a loop for the geometry optimization is begun. The SCF equations are solved for the total energy and wavefunction, which in turn is used to compute an energy gradient–the first derivatives of the energy with respect to displacements in the nuclear coordinates. If the gradient is within a preset tolerance, the loop is terminated and the post-SCF calculations are performed. If the gradient is too large, the initial geometry is varied (according to the rules of some search algorithm), and a new calculation of integrals, SCF energy, and energy gradient follows. This loop continues until the desired tolerance is achieved. Empirically, it is observed that for a system with n degrees of freedom, it takes between n and $2n$ loops to achieve convergence in bond lengths and bond angles to within 0.001 angstroms and 0.1 degrees, respectively [45].

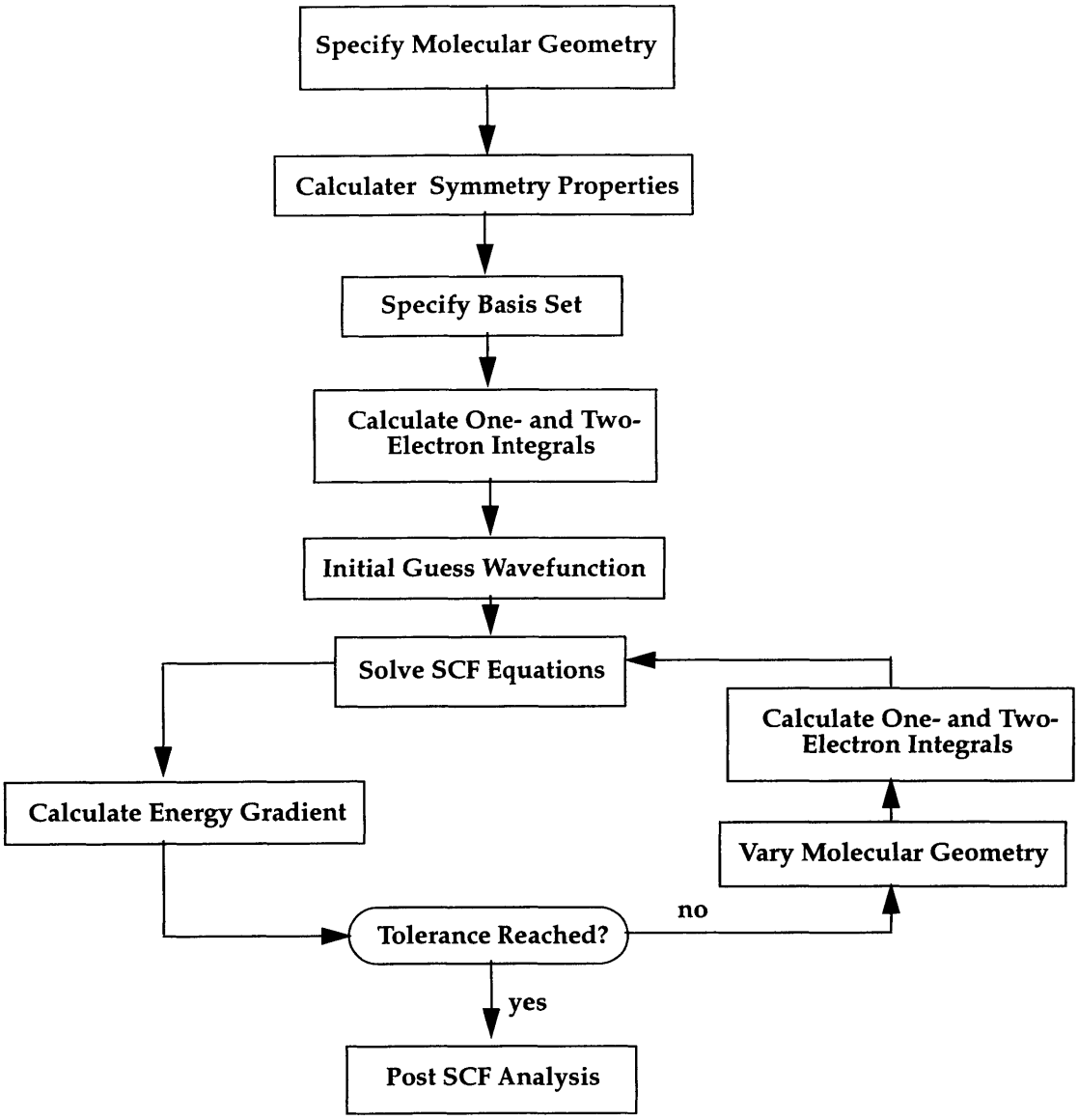


Figure A-1: Schematic of Hartree-Fock Self-Consistent Field Algorithm (adopted from Hehre *et al*)

A.6 Basis Functions

The Hartree-Fock single determinantal wavefunction is built up from a set of molecular orbitals, ψ_i, which in turn are represented as linear expansions in a set of *basis functions*, and written as,

$$\psi_i = \sum_{\mu=1}^{N} c_{\mu o} \phi_\mu \tag{A.69}$$

where the μ_i are the individual *nuclear-centered* basis functions, and the $c_{\mu i}$ are the molecular orbital expansion coefficients. The basis functions are associated with each nucleus and therefore depend only on the nuclear charge; the functions will also have symmetry and angular properties representative of the atomic orbitals they correspond to. Two fundamental types of basis functions historically applied in HF calculations are Slater-type atomic orbitals (STO's) and gaussian-type atomic functions [45, 47].

Slater-type orbitals have exponential radial components and are labelled like hydrogen atom orbitals, $1s, 2s, 2p_x, \cdots$ and are in normalized form,

$$\phi_{1s} = (\frac{\zeta_1^3}{\pi})^{1/2} e^{(-\zeta_1 r)} \tag{A.70}$$

$$\phi_{2s} = (\frac{\zeta_2^5}{96\pi})^{1/2} r e^{(\frac{-\zeta_2 r}{2})} \tag{A.71}$$

and so on for functions or higher angular momentum. The ζ values are constants which influence the radial width of the orbitals. Although STO's are fairly accurate representations of atomic orbitals, they are not computationally amenable and so are not used so widely [45, 47].

Gaussian functions are polynomials of x, y, z, multiplied by a factor of $e^{(-\alpha r^2)}$ with α a constant setting the radial extent of the function. A few exemplary gaussian functions of varying angular momentum are [45, 47],

$$g_s(\alpha, \mathbf{r}) = (\frac{2\alpha}{\pi})^{3/4} e^{(-\alpha r^2)} \tag{A.72}$$

$$g_x(\alpha, \mathbf{r}) = (\frac{128\alpha^5}{\pi^3})^{1/4} x\, e^{(-\alpha r^2)} \tag{A.73}$$

$$g_{xy}(\alpha, \mathbf{r}) = (\frac{2048\alpha^7}{\pi^3})^{1/4} x\, y\, e^{(-\alpha r^2)} \tag{A.74}$$

Gaussian type orbitals were introduced by Boys (ref in Hehre, Pople, et al, p 19). They are not as accurate a representation of atomic orbital functions as they do not have a singularity at the origin (the gradient diverges at the origin). However, their advantage lies in the fact that the gaussian function is analytically integrated [45, 47].

Finally, before moving on to discuss specific basis function sets, it is noted that a third approach to basis sets is that of using linear combinations of gaussians of certain angular momentum to build up an overall basis function. The overall function is the *contracted gaussian* and the functions used in the linear expansion are *primitive gaussians* [45, 47].

A.7 Minimal Basis Sets: STO-KG

The simplest level of *ab initio* theory uses minimal basis sets of nuclear-centered functions. This model includes only enough functions to accommodate all of the electrons of the atom and maintain spherical symmetry. Thus, with this framework, hydrogen and helium, for example, would have one s-type function; lithium and beryllium a pair of such functions; and second-row elements would have the $1s, 2s, 2p$ inner-shell functions along with the appropriate valence shell s- and p-type functions. Although this would be truly a *minimal* basis set, in practice, the low-lying p- and d-type functions are added even though they are unoccupied. Consequently, hydrogen and helium have one function, elements from lithium to neon have five functions, elements sodium to argon have nine functions, potassium and calcium have 13 functions, elements scandium to krypton have 18 functions, and so on. The minimal basis set is too small to be of practical value in a quantitative sense; however, for qualitative anal ysis, it can be useful as a starting point for further, higher level calculations [45, 47].

Because the expansion in a minimal basis set is quite short, the functions used for the approximation must be accurate. For this reason, single Gaussian functions are not used, but rather, contracted Gaussian functions which are expansions of gaussian functions of given symmetry and angular properties fitted to Slater functions. These contracted Gaussian functions are known as STO-LG (Slater-type orbitals contracted with 'L' Gaussian functions) [45, 47].

The Slater-type orbital in terms of 'K' primitive gaussian functions is:

$$\phi_{nl}^{SF}(\mathbf{r}) = \sum_{k=1}^{K} d_{nl,k} g_l(\alpha_{n,k}, \mathbf{r}) \tag{A.75}$$

where the subscripts n and l refer to the principal and angular quantum numbers, and the g_l are primitive gaussian functions having angular properties corresponding to the l angular quantum number. Consider, for instance,

$$\phi_{1s}^{SF}(\zeta = 1.0) = \sum_{i=1}^{K} d_{i,1s} g_{1s}(\alpha_{i,1s}) \tag{A.76}$$

$$\phi_{2s}^{SF}(\zeta = 1.0) = \sum_{i=1}^{K} d_{i,2s} g_{1s}(\alpha_{i,2sp}) \tag{A.77}$$

$$\phi_{2p}^{SF}(\zeta = 1.0) = \sum_{i=1}^{K} d_{i,2p} g_{2p}(\alpha_{i,2sp}) \tag{A.78}$$

which are the contracted gaussian expansions for the $1s, 2s, 2p$ Slater functions. The contraction coefficients, d's, and exponents, α's are obtained via a least squares fit to the Slater function which minimize the integral [45, 47],

$$\epsilon_{nl} = \int d\mathbf{r} [\phi_{nl}^{SF}(\mathbf{r}) - \phi_{nl}^{CGF}(\mathbf{r})]^2 \tag{A.79}$$

For the example of the three functions introduced above, the error integrals are,

$$\epsilon_{1s} = \int d\mathbf{r} [\phi_{1s}^{SF}(\mathbf{r}) - \phi_{1s}^{CGF}(\mathbf{r})]^2 \tag{A.80}$$

$$\epsilon_{2s} + \epsilon_{2p} = \int d\mathbf{r}[\phi_{2s}^{SF}(\mathbf{r}) - \phi_{2s}^{CGF}(\mathbf{r})]^2 + \int d\mathbf{r}[\phi_{2p}^{SF}(\mathbf{r}) - \phi_{2p}^{CGF}(\mathbf{r})]^2 \qquad \text{(A.81)}$$

The important idea to note about the STO-KG fitting is that the contraction exponents are common to the functions of a particular shell; this is done so that the functions within a shell of quantum number n have identical radial behavior and can be integrated as one function. The constant normalization factor is associated after the radial integration is performed. This grouping adds tremendously to the efficiency of the algorithm. Although longer contraction lengths would seem to yield better fits, it is found empirically that a contraction length of $K = 3$ is sufficient to reproduce faithfully all results of a Slater function calculation. The STO-3G is the standard for minimal basis set calculations [45, 47].

Note that to fit Slater functions of different ζ values, all one has to do is scale the α values fitting the $\zeta = 1$ Slater functions by a factor of ζ^2 [45, 47].

A.8 Extended Basis Sets: Double Zeta and Split-Valence Basis Functions

The minimal basis, although useful qualitatively, comes with several inherent inadequacies. Because fixed Gaussian exponents are used, the individual orbitals cannot expand or contract spatially in response to changing molecular environments. This is due to the single valence function used for each shell and the fact that there is no exponent optimization that will allow the orbitals to adjust sizes . A second deficiency of the minimal basis set is that it cannot account rigorously for anisotropy of the molecular charge distribution (ie, polarization of the electron cloud). Finally, since the number of atomic basis functions is not allocated according to number of electrons, elements near the right of the Periodic Table tend to be less accurately described relative to those further left [45, 47].

Remedies to these deficiencies include allowing more than one valence function for

each symmetry and angular type in the basis set. This would increase the number of basis functions of all elements beyond what is actually needed, and multiple valence functions would introduce greater flexibility for the radial size to be determined via different weighting of the individual functions in the variational scheme. For instance, if an s-type function were replace by a more diffuse and a more contracted function, the final description would be able to weight the two so as to attain a radial behavior in between the contracted and diffuse extremes determined by the functions. To address the issue of describing anisotropic molecular charge distributions, two approaches are feasible. One is to allow the x,y, and z p-components of the valence region to have different radial distributions—in effect, use anisotropic minimal basis sets. This approach, which attributes different exponents for individual Cartesian directions, is useful for molecules with high symmetry. The second, more attractive approach is to use multiple valence p- and d-type functions of an *isotropic* nature. This is equivalent to the remedy for the radial size problem. Again, the use of two or more sets of isotropic valence functions with varying radial extents allows independent optimization of the individual functions within the SCF procedure [45, 47].

Thus, although the isotropic minimal basis set constrains the radial behavior of valence functions to be identical, the extended models allow more flexibility in representing more realistically molecular environments due to the adjustability of individual components in the variational scheme.

Two ways to extend the molecular orbital expansion are to double all the functions of a minimal basis set or, to double only the valence the functions, leaving the inner, core, functions as those for the minimal set. The first approach is commonly referred to as a double zeta basis, and the latter a split valence basis. The split-valence approach is justified by reasoning that although the core electrons are important to the total energetics of the molecule, molecular bonding as well as dipole moments, valence ionization potentials, charge densities, dissociation energies, etc. are not affected [45, 47].

Commonly used split-valence representations are the 4-31G and 6-31G basis sets. The 4-31G set, defined for all first-row elements and the second-row elements phos-

phorous, sulfur, and chlorine, uses inner-shell expansions of four Gaussians and two valence functions, one being a single primitive Gaussian, and the inner function a contraction of three Gaussian primitives. Related to this is the 6-31G basis set, which differs in that the inner shells are contractions of six Gaussian primitives [45, 47].

As an example, for the atoms lithium to fluorine, the contractions for the 6-31G basis set are,

$$\phi_{1s}(\mathbf{r}) = \sum_{i=1}^{6} d_{i,1s} g_{1s}(\alpha_{i,1s}, \mathbf{r}) \tag{A.82}$$

$$\phi'_{2s}(\mathbf{r}) = \sum_{i=1}^{3} d'_{i,2s} g_{1s}(\alpha'_{i,2sp}, \mathbf{r}) \tag{A.83}$$

$$\phi''_{2s}(\mathbf{r}) = g_{1s}(\alpha''_{i,2sp}, \mathbf{r}) \tag{A.84}$$

$$\phi'_{2p}(\mathbf{r}) = \sum_{i=1}^{3} d'_{i,2p} g_{2p}(\alpha'_{i,2sp}, \mathbf{r}) \tag{A.85}$$

$$\phi''_{2p}(\mathbf{r}) = g_{2p}(\alpha''_{i,2sp}, \mathbf{r}) \tag{A.86}$$

A.9 Polarization Basis Sets

Following split-valence basis sets, the next level of sophistication comes with polarization functions used to allow descript ion of highly polar molecules and small strained ring systems by allowing nonuniform displacement of charge away from nuclear centers.

One approach would be to add basis functions not associated with any single center, although this would not be size consistent (the number of non-nuclear centered functions does not necessarily increase in direct proportion with molecular size). Also, it may be difficult to unambiguously define where to place these functions. The generally applied technique is to add functions of higher angular quantum number

(d-type functions on heavy-atoms and p-type on hydrogen). Consider the mixing of a valence s-function with a p_x function for the hydrogen atom. This would lead to displacement of the function center away from the hydrogen nucleus along the x-axis; analogously, mixing with p_y and p_z functions would lead to displacements in the corresponding directions. For shifting p-type functions away from their origins, d-type functions can be added, and for displacing d-type functions, one adds f-type functions [45, 47].

Thus, basis sets with functions of higher angular momentum than required by the atomic ground state are termed polarization basis sets and these effectively allow for charge polarization, the displacement of charge away from nuclear centers.

A.9.1 Polarized Basis Sets: 6-31G* and 6-31G**

The simplest polarization basis sets are the 6-31G* and 6-31G** representations. These functions are built up by adding to the 6-31G split valence functions a set of six second-order (d-type) Gaussian primitives on the heavy-atoms–the 6-31G* set, and a single set of three Gaussian p-type primitives on hydrogen and helium–the 6-31G** set. Note that the six d-type functions are uncontracted 3d primitive Gaussians. There are six primitives per atom, the six being linear combinations of the usual five 3d functions and a single 3s function [45, 47].

A.9.2 Polarized Basis Sets: 6-311G* and 6-311G**

Two larger but totally analogous polarization function sets are the 6-311G* and 6-311G** basis sets. For these functions, the inner-shell functions are contractions of 6 Gaussian primitives and a valence region split into three regions, one represented by a contraction of three Gaussians, and the other two by two different single primitives. The asterisks as before indicate the addition of valence d- and p-type functions for heavy-atoms and hydrogen. The triple split allows for more flexibility and improved description of the valence region. The valence d-functions added are a set of 5 primitives instead of the six used in the split-valence sets. This is not a problem as the last

d-type function represents an s-type symmetry and is effectively redundant [45, 47].

A.10 Electron Correlation Correction Methods: Moller-Plesset Perturbation Theory

As discussed earlier, the Hartree-Fock approach neglects electron correlation between particles of opposite spin–the HF treatment is thus uncorrelated. Configuration interaction (CI) methods attempt to incorporate these effects (ref), but they are not size consistent and truly rigorous when all possible excitations are included in the trial wavefunctions (full CI). A non-variational, size-consistent alternative is perturbation theory (PT) in which the total, correct Hamiltonian is decomposed into a zeroth-order part, $\mathcal{H}_o$, which has known eigenfunctions and eigenvalues, and a perturbation, $\mathcal{V}$. The exact energy and wavefunctions are then expressed as expansions of corrections to the zeroth-order components. Terms of similar order are grouped together and represent n-th order corrections to the energy and wavefunctions [45, 47].

The following is a brief discussion of Moller-Plesset theory.

The correct total Hamiltonian of a system may be written as

$$\mathcal{H}\Phi_i = (\mathcal{H}_o + \mathcal{V})\Phi_i = \mathcal{E}_i\Phi_i \tag{A.87}$$

where the zero-th order eigenfunctions and eigenvalues of $\mathcal{H}_o$ are,

$$\mathcal{H}_o\Psi_i^{(0)} = E_i^{(0)}\Psi_i^{(0)} \tag{A.88}$$

If the perturbation $\mathcal{V}$ is small, the actual eigenfunctions and energies, Φ_i and $\mathcal{E}_i$ will be close to the zeroth-order eigenfunctions and energies. The perturbation approach is a systematic way to obtain improvements to the zeroth-order components to approach the correct total Hamiltonian eigenfunctions and eigenvalues. Thus, one begins by writing the total Hamiltonian, $\mathcal{H}$ as,

$$\mathcal{H} = \mathcal{H}_o + \lambda\mathcal{V} \tag{A.89}$$

The eigenfunctions and eigenvalues are expanded in Taylor series in λ ,

$$\mathcal{E}_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots \tag{A.90}$$

$$\Phi_i = \Psi_i^{(0)} + \lambda \Psi_i^{(1)} + \lambda^2 \Psi_i^{(2)} + \cdots \tag{A.91}$$

The energy $E_i^{(n)}$ is the n-th order energy correction. The problem is now to express the various order energies and eigenfunctions in terms of the zero'th order components and matrix elements of the perturbation, $\mathcal{V}$ between the unperturbed wave functions.

The reader is referred to the literature for the mathematical details of the derivations, but the second-order energy is calculated as [45, 47]

$$E_i^{(2)} = \sum_n{}' \frac{| < i|\mathcal{V}|n > |^2}{E_i^{(0)} - E_n^{(0)}} \tag{A.92}$$

where the prime on the sum indicates that $n \neq i$ and the squared term is the matrix element of the perturbation between the zeroth-order wave functions $\Psi_i^{(0)}$ and $\Psi_n^{(0)}$.

Appendix B

Experimental Techniques Probing Liquid-Vapor and Liquid-Liquid Interfacial System Properties

There are a host of surface analytical techniques currently applied in the study of solid interfaces and surfaces. These include Auger electron spectroscopy (AES), electron spectroscopy for chemical analysis (ESCA) or x-ray photoelectron spectroscopy (XPS), low energy electron diffraction (LEED), scanning tunneling microscopy (STM), and scanning probe microscopy (SPM) such a atomic force (AFM) or lateral force (LFM) microscopy), as well as a host of others not mentioned here. Although each of these methods is capable of giving compositional and/or structural information of the interfacial region, they are difficult to apply to liquid-vapor or liquid-liquid interfaces. AES, ESCA, XPS, and STM must be carried out under ultra-high vacuum conditions; furthermore, in the case of STM, the surface must be electrically conducting so as to generate a tunneling current from the tip to sample. Also, surface probing techniques such as STM and SPM require bringing a very fine probe tip close to the sample, a feat that is virtually impossible with liquid surfaces or interfaces between immiscible liquids. Techniques which are more amenable to studying fluid interfaces, such as x-ray reflectometry, neutron and x-ray reflectivity, and ellipsometry give information on interfacial or thin film thicknesses, roughness,

density, and composition (and its profile as a function of depth in the interface or thin film). Fundamentally, none of these techniques can give details as to the *orientations* of molecules at an interface; furthermore, one is at a loss if the very specific *atomistic* description of the system is desired. All these techniques give average pictures of the underlying molecular/atomic structure. For this objective, one must turn to the more recently applied non-linear optical techniques of sum frequency generation (SFG) and second harmonic generation (SHG), the latter being a special case of the more general sum frequency methods.

In the following, the SFG and SHG methods will be discussed briefly as they are an important experimental tool complementing the results of theoretical/computational studies of fluid-fluid interfaces to give a clearer understanding of the physics of these inhomogeneous systems, particularly with respect to the orientation of molecules.

In non-linear media, the interaction of waves leads to wave mixing, thus generating waves at sum and difference frequencies. Sum frequency generation is the technique which uses the sum frequencies for analysis. Consider two laser pulses/beams at frequencies ω_1 and ω_2 interacting in a nonlinear medium to generate a polarization (nonlinear itself) $\mathbf{P}^{(2)}(\omega_3 = \omega_1 + \omega_2$. The polarization, itself being a collection of oscillating dipoles, radiates at ω_3. Although in the general case this radiation can occur in all directions, if the polarization is correlated in a certain direction, the radiation pattern can be peaked in that direction. The direction is then determined by phase-matching, which is essentially conservation of energy and momentum. The energy conservation leads to $\omega_3 = \omega_1 + \omega_2$, and the momentum conservation requires that the wave-vectors be related as $\mathbf{k}_3 = \mathbf{k}_1 + bfk_2$; this is merely the phase-matching direction in which the sum frequency radiation is most effectively generated. For second harmonic generation, which is simply a special case of sum frequency generation, the initial frequencies are equal, $\omega_1 = \omega_2 = \omega$ and the sum frequency becomes the second harmonic, $\omega_3 = 2\omega$.

The SFG and SHG methods are attractive spectroscopic techniques for studying fluid-fluid interfaces because they are surface-sensitive techniques. A major difficulty of conventional spectroscopic methods applied to liquid-liquid interfaces is the separa-

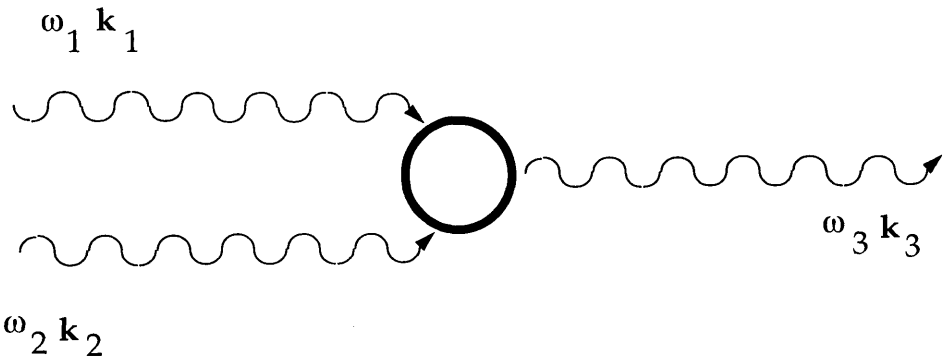


Figure B-1: Schematic of the Sum Frequency Generation Method

tion of the optical signal from the bulk and interfacial regions; that is, decoupling the optical response of the interface from the optical response of the bulk media is non-trivial. However, SFG and SHG are electric-dipole forbidden in *centrosymmetric* media; thus, liquids and gases, which are centrosymmetric (isotropic), as well as some centrosymmetric solids, do not generate appreciable signals from the bulk. At an interface, inversion symmetry is broken, and the secondary frequency becomes dipole-allowed. In effect, only the contributions from the interface are observed. Furthermore, in the electric-dipole approximation, for a molecular liquid, the interface contributes to the SFG spectrum only if the molecules are polar-oriented; the polarization dependence of the SFG spectrum then gives information on the molecular orientation of surface/interface molecules.

The SFG and SHG techniques have been applied to the study of numerous interfacial systems ranging from the pure water-water vapor interface to molecular adsorption at the interface between two immiscible electrolyte solutions (ref).

In terms of thermodynamic properties, the surface tension and surface potential are often studied; these of course give very little information in terms of detailed molecular structure. A widely-used non-optical method for obtaining surface tension is the measurement of contact angles of fluids on solid surfaces (Andrade, 1985). Although the interpretation and analysis of results rests heavily on a set of controversial assumptions, results afford at least a preliminary base for investigating polymer surfaces. Fundamentally, contact angle techniques measure contact angles of liquids on surfaces, and, by coupling these angles to equilibrium thermodynamics relations,

attempt to arrive at surface and interfacial free energies. Numerous techniques for measuring contact angles are available. These include the Wilhelmy plate, captive bubble, sessile drop, and capillary rise methods. The thermodynamic component is based on Young's equation which describes the mechanical and energetic equilibrium between an ideally non-deformable solid, a vapor, and liquid (Andrade, 1985),

$$\gamma_{sl} = \gamma_{sv} - \gamma_{lv} cos(\theta) \tag{B.1}$$

where γ_{sv} , γ_{sl}, and γ_{lv} are the solid-vapor, solid-liquid, and liquid-vapor interfacial free energies. Fowkes presented a theory of the interfacial tension as arising from various classes of intermolecular forces such as dispersion, dipole-dipole, induction, and hydrogen-bonding forces (Fowkes, 1964). The Fowkes expression for the interfacial tension is,

$$\gamma = \gamma^{dispersion} + \gamma^{hydrogen-bonding} + \gamma^{dipole-dipole} + \cdots \tag{B.2}$$

where γ is the total interfacial tension and the sum is over the various contributions. Note that although these individual components are not thermodynamically defined, Fowkes' formalism regards them as unique physical properties of the material.

Finally, a few comments on critical surface tension approaches. Zisman *etal* proposed a quantity, the critical surface tension, γ_c, obtained from a plot of contact angle versus surface tension of pure liquids places on a solid surface (Andrade, 1985). The authors found that the values of $cos(\theta)$ plotted as a function of the surface tensions of a homologous series of liquids (e.g., alkanes) fell along a narrow, linear band, which when extrapolated to $cos(\theta) = 1$ (that is, complete wetting of the solid) yields the critical surface tension of wetting of a solid. Finally, for surface tensions of polymers below the T_g, extrapolated values of γ may be obtained based on the observed linear relation between temperature and surface tension in the melt (Wu, 1982).

Appendix C

Contraction Coefficients and Gaussian Exponents

C.1 Specification of Basis Functions: 6-311G**

The 6-311G** basis set for heavy atoms is defined in terms of normalized Gaussian primitive functions as,

$$\phi_{1s} = \sum_{k=1}^{6} d_{1s,k} g_s(\alpha_{1k}) \ ; \tag{C.1}$$

$$\phi'_{2s} = \sum_{k=1}^{3} d'_{2s,k} g_s(\alpha'_{2k}) \ ; \tag{C.2}$$

$$\phi'_{2p} = \sum_{k=1}^{3} d'_{2p,k} g_p(\alpha'_{2k}) \ ; \tag{C.3}$$

$$\phi''_{2s} = g_s(\alpha''_2) \ ; \tag{C.4}$$

$$\phi''_{2p} = g_p(\alpha''_2) \ ; \tag{C.5}$$

$$\phi_{2s}''' = g_s(\alpha_2''') \ ; \tag{C.6}$$

$$\phi_{2p}''' = g_p(\alpha_2''') \ ; \tag{C.7}$$

$$\phi_d = g_d(\alpha_d) \tag{C.8}$$

In the above relations, the $g_s(\alpha)$, $g_p(\alpha)$, and $g_d(\alpha)$ are normalized s-, p-, and d-type primitive Gaussian functions; the exponent is given by α. For the hydrogen atom, only the split valence functions are needed,

$$\phi_s' = \sum_{k=1}^{3} d_k' g_s(\alpha_k') \ ; \tag{C.9}$$

$$\phi_s'' = g_s(\alpha'') \ ; \tag{C.10}$$

$$\phi_s''' = g_s(\alpha''') \ ; \tag{C.11}$$

$$\phi_p = g_p(\alpha_p) \tag{C.12}$$

The constants and exponents which completely describe the basis are given in Table C.1 and Table C.2. Note that for heavy atoms, a single set of 5 d-type gaussians is used; the exponent for this set is $\alpha_d = 0.626$ for carbon and $\alpha_d = 1.292$ for oxygen. For hydrogen, the p-type exponent is $\alpha_p = 0.75$. Furthermore, in practice, one needs to consider the px, py, pz orbitals, thus necessitating primed, double-primed, and triple-primed analogs for the px, py, and pz wave functions separately.

Table C.1: 6-311G Gaussian Basis Functions for Carbon, Oxygen

Atom	α_1	d_{1s}	$\alpha_2^{'}$	$d_{2s}^{'}$	$d_{2p}^{'}$	$\alpha_2^{''}$	$\alpha_2^{'''}$
Carbon	4563.24	0.00196665	20.9642	0.114660	0.0402487	0.483456	0.145585
	682.024	0.0152306	4.80331	0.919999	0.237594		
	154.973	0.0761269	1.45933	-0.00303068	0.815854		
	44.4553	0.260801					
	13.0290	0.616462					
	1.82773	0.221006					
Oxygen	8588.50	0.00189515	42.1175	0.113889	0.0365114	0.905661	0.255611
	1297.23	0.0143859	9.62837	0.920811	0.237153		
	299.296	0.0707320	2.85332	-0.00327447	0.819702		
	87.3771	0.240001					
	25.6789	0.594797					
	3.74004	0.280802					

Table C.2: 6-311G Gaussian Basis Functions for Hydrogen

Atom	$\alpha^{'}$	$d^{'}$	$\alpha^{''}$	$\alpha^{'''}$
Hydrogen	33.8650	0.0254938	0.325840	0.102741
	5.09479	0.190373		
	1.15879	0.852161		

C.2 Specification of Basis Functions: 6-31G**

For the 6-31G** basis set, there is only the minor change of one less valence split compared to the 6-311G** basis. Thus, there are no components corresponding to the triple primed exponents and coefficients. Table C.3 and Table C.4 give the 6-31G parameters for carbon, oxygen, and hydrogen atoms. The exponents for the polarization functions (d-type for heavy atoms and p-type for hydrogen) are the same as for the 6-311G** basis.

Table C.3: 6-31G Gaussian Basis Functions for Carbon, Oxygen

Atom	α_1	d_{1s}	$\alpha_2^{'}$	$d_{2s}^{'}$	$d_{2p}^{'}$	$\alpha_2^{''}$
Carbon	3047.525	0.001834737	7.868272	-0.1193324	0.06899907	0.155986
	457.3695	0.01403732	1.881289	-0.1608542	0.316424	
	103.9487	0.06884262	0.5442493	1.143456	0.7443083	
	29.21016	0.2321844				
	9.286663	0.4679413				
	3.163927	0.362312				
Oxygen	5484.672	0.00189515	15.85513	-0.1107775	0.07087427	0.2811389
	825.2349	0.01395017	3.673027	-0.1480263	0.3397528	
	188.047	0.06844508	1.034345	1.130767	0.7271586	
	52.9645	0.2327143				
	16.89757	0.4701929				
	5.799635	0.3585209				

Table C.4: 6-31G Gaussian Basis Functions for Hydrogen

Atom	$\alpha^{'}$	$d^{'}$	$\alpha^{''}$
Hydrogen	13.00773	0.0334946	0.1219492
	1.9620790	0.234727	
	0.444529	0.8137573	

Bibliography

[1] A. S. Hoffman. Applications of synthetic polymeric biomaterials in medicine and biotechnology. In E. Piskin and A. S. Hoffman, editors, *Polymeric Biomaterials.* Martinus Hijhoff, Dordrecht, 1986.

[2] I. Engelberg and J. Kohn. Physico-mechanical properties of degradable polymers used in medical applications: a comparative study. *Biomaterials*, 12:292, 1991.

[3] R. L. Dunn, J. P. English, J. D. Strobel, D. R. Cowsar, and T. R. Tice. Preparation and evaluation of lactide/glycolide copolymers for drug delivery. In C. Migliaresi, editor, *Polymers in Medicine III.* Elsevier, Amsterdam, 1988.

[4] D. K. Gilding and A. M. Reed. Biodegradable polymers for use in surgery–polyglycolic/poly(lactic) acid homo- and copolymers:1. *Polymer*, 20:1459, 1979.

[5] Su Ming Li, H. Garreau, and M. Vert. Structure-property relationships in the case of degradation of massive aliphatic poly-(alpha-hydroxy acids) in aqueous media. part 1. *Journal of Materials Science: Materials in Medicine I*, ?:123–130, 1990.

[6] T. Nakamura, S. Hitomi, S. Watanabe, Y. Shimizu, K. Jamshidi, S-H. Hyon, and Y Ikada. Bioabsorption of polylactides with different molecular properties. *J. Biomed. Mat. Res.*, 23:1115–1130, 1989.

[7] D. K. Gilding and A. M. Reed. Biodegradable polymers for use in surgery–polyglycolic/poly(lactic) acid homo- and copolymers:2 in vivo degradation. *Polymer*, 20:1459, 1979.

[8] Ajit Sadana. Protein adsorption and inactivation on surfaces. influence of heterogeneities. *Chem. Rev.*, 92:1799–1818, 1992.

[9] D. Rigby and R. J. Roe. Molecular dynamics simulation of polymer liquid and glass. i. glass transition. *J. Chem. Phys.*, 87(12):7285, 1987.

[10] D. Rigby and R. J. Roe. Molecular dynamics simulation of polymer liquid and glass. i. short range order and orientation correlation. *J. Chem. Phys.*, 89(8):5280, 1988.

[11] D. Rigby and R. J. Roe. Molecular dynamics simulation of polymer liquid and glass. 3. chain conformation. *Macromolecules*, 22:2259–2264, 1989.

[12] K. F. Mansfield and D. N. Theodorous. Molecular dynamics simulation of a glassy polymer surface. *Macromolecules*, 24:6283–6294, 1991.

[13] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids.* Oxford University Press, New York, 1987.

[14] D. J. Mooney, L. G. Cima, R. Langer, L. Johnson, L. K. Hansen, D. E. Ingber, and J. P. Vacanti. Principles of tissue engineering and reconstruction using polymer-cell constructs. In L. G. Cima and Eyal S. Ron, editors, *Tissue-Inducing Biomaterials: MRS Symposium Proceedings*, pages 345–352. Materials Research Society, Pittsburgh, PA, 1992.

[15] R. Singhvi, G. N. Stephanopoulos, and D. I. C. Wang. Effect of substratum morphology on animal cell adhesion and behavior. In L. G. Cima and Eyal S. Ron, editors, *Tissue-Inducing Biomaterials: MRS Symposium Proceedings*, pages 237–245. Materials Research Society, Pittsburgh, PA, 1992.

[16] Jean-Pierre Cazenave. Interaction of surfaces and blood cells: Platelet reaction. In Steen Dawids, editor, *Polymers: Their Properties and Blood Compatibility*, page ? Kluwer Academic Publishers, ?, ?

[17] J. D. Andrade and V. Hlady. Protein adsorption and materials biocompatibility: A tutorial review and suggested hypotheses. In *Advances in Polymer Science.* Springer-Verlag, Berlin, 1986.

[18] Anna Ludwicka, B. Jansen, T. Wadstrom, L. M. Switalski, G. Peters, and G. Pulverer. Attachment of staphylococci to various polymers. In A. S. Hoffman, S. W. Shalaby, B. D. Ratner, and T. A. Horbett, editors, *Polymers as Biomaterials.* Plenum Press, New York, 1984.

[19] R. A. Miller, J. M. Brady, and D. E. Cutright. Degradation rates of oral resorbable implants (polylactates and polyglycolates): Rate modification with changes in pla/pga copolymer ratios. *J. Biomed. Mater. Res.*, 11:711–719, 1977.

[20] J. D. Andrade, D. E. Gregonis, and L. M. Smith. Polymer surface dynamics. In J. D. Andrade, editor, *Surface and Interfacial Aspects of Biomedical Polymers.* Plenum Press, New York, 1985.

[21] H. Yasuda, E. J. Charlson, E. M. Charlson, T. Yasuda, M. Miyama, and T. Okuno. Dynamics of surface property change in response to changes in environmental conditions. *Langmuir*, 7:2394–2400, 1991.

[22] Frank J. Holly and Miguel F. Refojo. Wettability of hydrogels i. poly(2-hydroxyethyl methacrylate). *J. Biomed. Mater. Res.*, 9:315–326, 1975.

[23] Eli Ruckenstein and Sathyamurthy V. Gourisankar. Environmentally induced restructuring of polymer surfaces and its influence on their wetting characteristics in an aqueous environment. *Journal of Colloid and Interface Science*, 107(2):488–502, 1985.

[24] Yasuyuki Tezuka, Hideki Kazama, and Kiyokazu Imai. Environmentally induced macromolecular rearrangement on the surface of polyurethane-

polysiloxane block copolymers. *J. Chem. Soc. Faraday Trans.*, 87(1):147–152, 1991.

[25] F.M. Etzler and W. Drost-Hansen. A role for water in biological rate processes. In W. Drost-Hansen and James S. Clegg, editors, *Cell-Associated Water*. Academic Press, New York, 1979.

[26] F.M. Etzler. *Langmuir*, 4:878, 1988.

[27] J. M. Haile. *Molecular Dynamics Simulation, Elementary Methods*. John Wiley and Sons, Inc., New York, 1992.

[28] H. Goldstein. *Classical Mechanics*. Addison-Wesley, Reading, MA, 2 edition, 1980.

[29] G. D. Birkhoff. Proof of a recurrence theorem for strongly transitive systems. *Proc. Nat. Acad. Sci.*, 17:650, 1931.

[30] G. D. Birkhoff. Proof of the ergodic theorem. *Proc. Nat. Acad. Sci.*, 17:656, 1931.

[31] J. von Neumann. Proof of the quasi-ergodic hypothesis. *Proc. Nat. Acad. Sci.*, 18:70, 1932.

[32] J. von Neumann. Physical applications of the ergodic hypothesis. *Proc. Nat. Acad. Sci.*, 18:263, 1932.

[33] D. C. Rapaport. *The art of molecular dynamics simulation*. Cambridge University Press, Cambridge, 1995.

[34] Hans C. Andersen. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52:24–34, 1983.

[35] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. *Journal Chemical Physics*, 76:637, 1982.

[36] J-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. *J. Comput. Phys.*, 23:327, 1977.

[37] 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. New molecular dynamics methods for various ensembles. *Rapport d'activite scientifique du CECAM*, pages 82–115, 1984.

[38] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72:2384–2393, 1980.

[39] T. A. Andrea, W. C. Swope, and H. C. Andersen. The role of long ranged forces indetermining the structure and properties of liquid water. *J. Chem. Phys.*, 79:4576–4584, 1983.

[40] D. M. Heyes. Molecular dynamics at constant pressure and temperature. *Chem. Phys.*, 82:285–301, 1983.

[41] S. Nose. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, 52(2):255–268, 1984.

[42] T. Schneider and E. Stoll. Molecular dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev*, B17:1302–1322, 1978.

[43] D. L. Ermak and Y. Yeh. Equilibrium electrostatic effects on behaviour of polyions in solution: polyion-mobile ion interaction. *Chem. Phys. Lett.*, 24:243–248, 1974.

[44] S. Nose. A unified formulation of the constant termperature molecular dynamics methods. *J. Chem. Phys.*, 81(1):511, 1984.

[45] Warren J. Hehre, Leo Radom, Paul v.R. Schleyer, and John A. Pople. *AbInitio Molecular Orbital Theory.* John Wiley and Sons, Inc., New York, 1986.

[46] M. Born and J. R. Oppenheimer. *Ann. Physik*, 84:527, 1927.

[47] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry, Introduction to Advanced Electronic Structure Theory*, volume ? of *?* McGraw-Hill Publishing Company, New York, ? edition, 1989.

[48] *Discover User Guide, Part 1, version 2.9.5.* Biosym Technologies, San Diego, 1994.

[49] Tsun-Mei Chang and Liem X. Dang. Molecular dynamics simulations of ccl4-h2o liquid-liquid interface with polarizable potential models. *Journal of Chemical Physics*, 104(17):6772–6783, 1996.

[50] Per Linse. Monte carlo simulation of liquid-liquid benzene-water interface. *Journal of Chemical Physics*, 86(7):4177–4187, 1987.

[51] S. M. Thompson and K. E. Gubbins. Molecular orientation at a vapor-liquid interface: Theoretical and computer simulation results for a model of chlorine. *Journal of Chemical Physics*, 70(11):4947–4951, 1979.

[52] M. Meyer, Mareschal M., and M. Hayoun. *Journal of Chemical Physics*, 89:1067, 1988.

[53] R. M. Townsend and Stuart A. Rice. Molecular dynamics studies of the liquid-vapor interface of water. *Journal of Chemical Physics*, 94(3):2207–2218, 1991.

[54] Ramona S. Taylor, Liem X. Dand, and Bruce C. Garrett. Molecular dynamics simulations of the liquid/vapor interface of spc/e water. *Journal of Physical Chemistry*, 100(28):11720, 1996.

[55] M. Matsumoto and Y. Kataoka. Study of the liquid-vapor interface of water. i. simulational results of thermodynamic properties and orientational structure. *Journal of Chemical Physics*, 88(5):3233–3245, 1988.

[56] M. Matsumoto and Y. Kataoka. Molecular orientation near liquid-vapor interface of methanol: Simulational study. *Journal of Chemical Physics*, 90(4):2398–2407, 1989.

[56] M. Matsumoto and Y. Kataoka. Molecular orientation near liquid-vapor interface of methanol: Simulational study. *Journal of Chemical Physics*, 90(4):2398–2407, 1989.

[57] Yuhong Zhang, Scott E. Feller, Bernard R. Brooks, and Richard W. Pastor. Computer simulation of liquid/liquid interfaces. i. theory and application to octane/water. *Journal of Chemical Physics*, 103(23):10252–10266, 1995.

[58] Aldert R. van Buuren, Siewert-Jan Marrink, and Herman J. C. Berendsen. A molecular dynamics study of the decane/water interface. *Journal of Physical Chemistry*, 97(36):9206–9212, 1993.

[59] Ilene Locker Carpenter and Warren J. Hehre. A molecular dynamics study of the hexane/water interface. *Journal of Physical Chemistry*, 94(2):531–536, 1990.

[60] Jonathan G. Harris. Liquid-vapor interfaces of alkane oligomers. structure and thermodynamics from molecular dynamics simulations of chemically realistic models. *Journal of Physical Chemistry*, 96(12):5077–5086, 1992.

[61] Mounir Tarek, Douglas J. Tobias, and Michael L. Klein. Molecular dynamics simulation of tetradecyltrimethylammonium bromide monolayers at the air/water interface. *Journal of Physical Chemistry*, 99(5):1393–1402, 1995.

[62] T. K. Xia, J. Ouyang, W. M. Ribarsky, and U. Landman. *Interfacial Alkane Films, preprint.*

[63] M. Vacatello, D. Y. Yoon, and B. Laskowski. *preprint.*

[64] K. F. Mansfield and D. N. Theodorous. *Macromolecules*, 23:4430–4445, 1990.

[65] J. R. Maple, M.-J. Hwang, T. P. Stockfisch, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler. Derivation of class ii force fields. i. methodology and quantum force field for the alkyl functional group and alkane molecules. *Journal of Computational Chemistry*, 15(2):162–182, 1994.

[66] Terrell L. Hill. *An Introduction to Statistical Thermodynamics.* Dover Publications, Inc., New York, 1986.

[67] James M. Briggs, Toan B. Nguyen, and William L. Jorgensen. Monte Carlo Simulations of Liquid Acetic Acid and Methyl Acetate with the OPLS PotentialFunctions. *Journal of Physical Chemistry*, 95(8):3315–3322, 1991.

[68] Grant D. Smith and Richard H. Boyd. A force field for conformational energy calculations on ester grop containing polymers. *Macromolecules*, 23(5):1527–1530, 1990.

[69] H. Sun. Force field for computation of conformational energies, structures, and vibrational frequencies of aromatic polyesters. *Journal of Computational Chemistry*, 15(7):752–768, 1994.

[70] S. Lifson, A. T. Hagler, and Dauber P. Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 1. carboxylic acids, amides, and the c=o–h hydrogen bonds. *J. Am. Chem. Soc.*, 101(18):5111–5121, 1979.

[71] A. T. Hagler, E. Huler, and S. Lifson. Energy functions for peptides and proteins. i. derivation of a consistent force field including the hydrogen bond from amide crystals. *J. Am. Chem. Soc.*, 96(17):5319–5335, 1974.

[72] A. T. Hagler, Dauber P., and S. Lifson. Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 3. the c=o—h-o hydrogen bond and the analysis of the energetics and packing of carboxylic acids. *J. Am. Chem. Soc.*, 101(18):5131–5141, 1979.

[73] P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler. Structure and energetics of ligand binding to proteins: E. coli dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins: Structure, Function, and Genetics*, 4:32–47, 1988.

[74] Bernard R. Brooks, Robert E. Bruccoleri, Barry D. Olafson, David J. States, S. Swaminathan, and Martin Karplus. Charmm: A program for maromolecular

energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4(2):187–217, 1983.

[75] Scott J. Weiner, Peter A. Kollman, David A. Case, U. Chandra Singh, Caterina Ghio, Giuliano Alagona, Salvatore Jr. Profeta, and Paul Weiner. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, 106(3):765–784, 1984.

[76] Scott J. Weiner, Peter A. Kollman, Dzung T. Nguyen, and David A. Case. An all atom force field for simulations of proteins and nucleic acids. *Journal of Computational Chemistry*, 7(2):230–252, 1986.

[77] Norman L. Allinger, Zhi-qiang S. Zhu, and Kuohsiang Chen. Molecular mechanics (MM3) studies of carboxylic acids and esters. *J. Am. Chem. Soc.*, 114(15):6120–6133, 1992.

[78] Norman L. Allinger, Young H. Yuh, and Jenn-Huei Lii. Molecular mechanics. the mm3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.*, 111(23):8551–8566, 1989.

[79] Norman L. Allinger and S. H. M. Chang. Conformational analysis-cxxiii. carboxylic acids and esters in force field calculations. *Tetrahedron*, 33:1561–1567, 1977.

[80] William J. Jorgensen, Jeffry D. Madura, and Carol J. Swenson. Optimized intermolecular potential functions for liquid hydrocarbons. *J. Am. Chem. Soc.*, 106(22):6638–6646, 1984.

[81] William J. Jorgensen and Julian Tirado-Rives. The opls potential functions for proteins. energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110(6):1657–1666, 1988.

[82] M. Briggs, James, Tooru Matsui, and William L. Jorgensen. Monte carlo simulations of liquid alkyl ethers with the opls potential functions. *Journal of Computational Chemistry*, 11(8):958–971, 1990.

[83] William L. Jorgensen, James M. Briggs, and M. Leonor Contreras. Relative partition coefficients for organic solutes from fluid simulations. *J. Phys. Chem.*, 94(4):1683–1686, 1990.

[84] Do Y. Yoon, Grant D. Smith, and Tsunetoshi Matsuda. A comparison of a united atom and an explicit atom model in simulations of polymethylene. *J. Chem. Phys.*, 98(12):10037, 1993.

[85] S. Toxvaerd. Molecular dynamics calculation of the equation of state of alkanes. *J. Chem. Phys.*, 93(6):4290, 1990.

[86] P. V. Krishna Pant, Jie Han, Grant D. Smith, and Richard H. Boyd. A molecular dynamics simulation of polyethylene. *J. Chem. Phys.*, 99(1):597, 1993.

[87] Jie Han and Richard H. Boyd. Molecular packing and small-penetrant diffusion in polystyrene: a molecular dynamics simulations study. *Polymer*, 37(10):1797–1804, 1996.

[88] Yong Jin and Richard H. Boyd. Subglass chain dynamics and relaxation in polyethylene: A molecular dynamics simulation study. *J. Chem. Phys.*, 108(23):9912–9923, 1998.

[89] S. Lifson and A. Warshel. *J. Chem. Phys.*, 49:5116, 1968.

[90] Marvin Waldman and A. T. Hagler. New combining rules for rare gas van der waals parameters. *Journal of Computational Chemistry*, 15(9):1077–1084, 1993.

[91] C. E. Blom and Hs. H. Gunthard. Rotational isomerism in methyl formate and methyl acetate: A low-temperature matrix infrared study using thermal molecular beams. *Chemical Physics Letters*, 84(2):267–271, 1981.

[92] Tatsuo Miyazawa. Internal rotation and low frequency spectra of esters, mono-substituted amides and polyglycine. *J. Am. Chem. Soc.*, 34(5):691–696, 1961.

[93] W. Pyckhout, C. van Alsenoy, and H. J. Geise. Structure of gaseous methyl acetate as determined by joint analysis of electron diffraction, microwave and

infrared spectroscopy, supplemented by a valence force field and constraints from geometry relaxed ab initio calculations. *Journal of Molecular Structure*, 144(3):265–279, 1986.

[94] Rose-Marie Moravie and Jacques Corset. Conformational behaviour and vibrational spectra of methyl propionate. *Chemical Physics Letters*, 26(2):210–214, 1974.

[95] Kenneth B. Wiberg and Keith E. Laidig. Barriers to rotation adjacent to double bonds. 3. the c-o barrier in formic acid, methyl formate, acetic acid, and methyl acetate. the origin of ester and amide "resonance". *J. Am. Chem. Soc.*, 109(20):5935–5943, 1987.

[96] V. Hernandez, C. Castiglioni, and Zerbi G. Hyperconjugation from infrared intensities: the case of methyl acetate and of its selectively deuterated derivatives. *Journal of Molecular Structure*, 324:189–198, 1994.

[97] D. Y. Yoon, M. Vacatello, and G. D. Smith. Simulations studies of polymer melts at interfaces. Research report, IBM Research Division, Almaden Research Center, San Jose, CA, September 1994.

[98] M. Watanabe and M. Karplus. Dynamics of molecules with internal degrees of freedom by multiple time-step methods. *Journal of Chemical Physics*, 99(10):8063–8073, 1993.

[99] M. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *Journal of Chemical Physics*, 97(3):1990–2001, 1992.

[100] M. Tuckerman, B. J. Berne, and G. J. Martyna. Molecular dynamics algorithm for multiple time scales: Systems with long range forces.

[101] M. Tuckerman, B. J. Berne, and Angelo Rossi. Molecular dynamics algorithm for multiple time scales: Systems with disparate masses. *Journal of Chemical Physics*, 94(2):1465–1469, 1991.

[102] M. Tuckerman and B. J. Berne. Molecular dynamics in systems with multiple time scales: Systems with stiff and soft degrees of freedom and with short and long range forces. *Journal of Chemical Physics*, 95(11):8362–8364, 1991.

[103] M. Tuckerman, G. J. Martyna, and B. J. Berne. Molecular dynamics algorithm for condensed systems with multiple time scales. *Journal of Chemical Physics*, 93(2):1287–1291, 1990.

[104] M. Tuckerman, B. J. Berne, and G. J. Martyna. Molecular dynamics algorithm for multiple time scales: Systems with long range forces. *Journal of Chemical Physics*, 94(10):6811–6815, 1991.

[105] H. F. Trotter. On the product of semi-groups of operators. *Proc. Am. Math. Soc.*, 10:545, 1959.

[106] A. Brass, B. J. Pendleton, Y. Chen, and B Robson. Hybrid monte carlo simulations theory and initial comparison with molecular dynamics. *Biopolymers*, 33:1307–1315, 1993.

[107] A. Irback. Hybrid monte carlo simulation of polymer chains. *Journal of Chemical Physics*, 101(2):1661–1667, 1994.

[108] B. M. Forrest and U. W. Suter. Hybrid monte carlo simulation of dense polymer chains. *Journal of Chemical Physics*, 101(3):2616–2629, 1994.

[109] S. W. de Leeuw, J. W. Perram, and E. R. Smith. Simulation of electrostatic systems in periodic boundary conditions. i. lattice sums and dielectric constants. *Proc. R. Soc. Lond. A.*, 373:27–56, 1980.

[110] L. Onsager. Electric moments of molecules in liquids. *J. Am. Chem. Soc.*, 58:1486–1493, 1936.

[111] M. Neumann. The dielectric constant of water. computer simulations with the mcy potential. *J. Chem. Phys.*, 82(12):5663–5672, 1985.

[112] E. B. Wilson, J. C. Decius, and P. S. Cross. *Molecular Vibrations.* Dover, New York, 1980.

[113] J. G. Harris and S. T. Cui. The structure and phase equilibria of salt-water solution at supercritical conditions. *International Journal of Thermophysics*, 16(2):493–502, 1995.

[114] J. G. Harris and Kwong H. Yung. Carbon dioxide's liquid-vapor coexistance curve and critical properties as predicted by a simple molecular model. *Journal of Physical Chemistry*, 99(31):12021, 1995.

[115] J. K. Lee, J. A. Barker, and G. M. Pound. *J. Chem. Phys.*, 60:1976, 1974.

[116] D. C. Holcomb, P. Clancy, and J. A. Zollweg. *Mol. Phys.*, 78:437, 1993.

[117] M. Matsumoto, Y. Takaoka, and Y. Kataoka. *J. Chem. Phys.*, 98:1464, 1993.

[118] Susan Budavari, Maryadele J. O'Neil, Ann Smith, Patricia E. Heckelman, and J. F. Kinneary. *Merck Index.* Merck and Co., Inc., Whitehouse Station, NJ, 1996.

[119] C. D. Stanners, Q. Du, R. P. Chin, P. Cremer, G. A. Somorjai, and Y.-R. Shen. *Chemical Physics Letters*, 232:407–413, 1995.

[120] R. Superfine, J. Y. Huang, and Y. R. Shen. Nonlinear optical studies of the pure liquid/vapor interface: Vibrational spectra and polar ordering. *Physical Review Letters*, 66(8):1066–1069, 1991.

[121] Q. Du, R. Superfine, E. Freysz, and Y. R. Shen. Vibrational spectroscopy of water at the vapor/water interface. *Physical Review Letters*, 70(15):2313–2316, 1993.

[122] R. J. Good. Surface entropy and surface orientation of polar liquids. *Journal of Physical Chemistry*, 61:810, 1957.

[123] Arthur W. Adamson. *Physical Chemistry of Surfaces.* John Wiley and Sons, Inc., New York, 4 edition, 1982.

[124] R. C. Reid, J. M. Prausnitz, and B. E. Poling. *The Properties of Gases and Liquids.* McGraw-Hill, New York, 4 edition, 1987.

[125] V. Zakharov, E. N. Brodskaya, and A. Laaksonen. Molecular dynamics simulation of methanol clusters. *J. Chem. Phys.*, 109(21):9487–9493, 1998.

[126] Jose Alejandre, Dominic J. Tildesley, and Gustavo A. Chapela. Molecular dynamics simulation of the orthobaric densities and surface tension of water. *Journal of Chemical Physics*, 102(11):4574–4583, 1995.

[127] D. Beaglehole. Experimental studies of liquid interfaces. In Clive A. Croxton, editor, *Fluid Interfacial Phenomena.* John Wiley and Sons Ltd, New York, 1986.

[128] H. T. Davis. Capillary waves and the mean field theory of interfaces. *J. Chem. Phys.*, 67(8):3636, 1977.

[129] J. S. Rowlinson and B. Widom. *Molecular Theory of Capillarity.* Clarendon, Oxford, 1982.

[130] F. P. Buff, R. A. Lovett, and F. H. Stillinger. *Phys. Rev. Lett.*, 15:621, 1965.

[131] J. G. Harris, J. Gryko, and S. A. Rice. Self-consistent monte carlo simulations of the electron and ion distributions on inhomogeneous liquid alkali metals. i. longitudinal and transverse density distributions in the liquid-vapor interface of a one-component system. *Journal of Chemical Physics*, 87(5):3069–3081, 1987.

[132] J. P. R. B. Walton, D. J. Tildesley, and J. S. Rowlinson. The pressure tensor at the planar surface of a liquid. *Molecular Physics*, 48(6):1357–1368, 1983.

[133] Daniel A. Higgins and Robert M. Corn. Second harmonic generation studies of adsorption at a liquid-liquid electrochemical interface. *J. Phys. Chem.*, 97(2):489–493, 1993.

[134] J. Gao and W. L. Jorgensen. *J. Phys. Chem.*, 92:5813, 1988.

[135] I. Benjamin. *J. Phys. Chem.*, 97:1432, 1992.

[136] Siewert-Jan Marrink and Herman J. C. Berendsen. *Journal of Physical Chemistry*, 98:4155, 1994.

[137] M. Hayoun, M. Meyer, and P. Turq. *J. Phys. Chem.*, 98:6626, 1988.

[138] K. J. Schweighofer and I. Benjamin. *J. Phys. Chem.*, 99:9974, 1995.

[139] D. Levesque, J. J. Weis, and J. P. Hansen. Simulation of classical fluids. In K. Binder, editor, *Monte Carlo Methods in Statistical Physics.* Springer-Verlag, New York, 1986.

[140] Paul R. Rablen, Jeffrey W. Lockman, and William L. Jorgensen. *ab initio* study of hydrogen-bonded complexes of small organic molecules with water. *J. Phys. Chem.*, 102(21):3782–3797, 1999.

[141] A. D. Becke. *J. Chem. Phys.*, 98:5648–5652, 1993.

[142] C. Lee, W. Yang, and R. G. Parr. *Phys. Rev. B*, 37:785–789, 1988.

[143] B. Miehlich, A. Savin, H. Stoll, and H. Preuss. *Chem. Phys. Lett.*, 157:200–206, 1989.

[144] Jefferson W. Tester and Michael Modell. *Thermodynamics and Its Applications.* Prentice Hall PTR, Upper Saddle River, New Jersey, 3 edition, 1997.

[145] J. G. Powles, S. E. Baker, and W. A. B. Evans. The chemical potential in atomically inhomogeneous fluids in external force fields by computer simulation. *J. Chem. Phys.*, 101(5):4098–4102, 1994.

[146] J. G. Powles, W. A. B. Evans, and N. Quirke. Non-destructive molecular-dynamics simulation of the chemical potential of a fluid. *Mol. Phys.*, 46(6):1347–1370, 1982.

[147] B. Widom. Some topics in the theory of fluids. *J. Chem. Phys.*, 39(11):2808–2812, 1963.

[148] W. G. Madden. Monte carlo studies of the melt-vacuum interface of a lattice polymer. *J. Chem. Phys.*, 87(2):1405, 1987.

[149] Arvind Hariharan and Jonathan G. Harris. Theory of the free surface of copolymer melts: Application to symmetric diblock copolymers. *J. Chem. Phys.*, 101(4):3353–3366, 1994.

[150] Arvind Hariharan and Jonathan G. Harris. Effect of copolymer sequence distribution upon surface structure and interfacial tension as predicted by a self-consistent field lattice model. *J. Phys. Chem.*, 99(9):2788–2796, 1995.

[151] Alan Antonelli and P. J. Flory. The configuration statistics of random poly(lactic acid) chains. i. experimental results. *Macromolecules*, 2(3):225–227, 1969.

[152] Inc. Biosym Technologies. *personal communication.*

[153] Richard L. Jaffe, Grant D. Smith, and Do Y. Yoon. Conformations of 1,2-dimethoxyethane from ab initio electronic structure calculations. *J. Phys. Chem.*, 97(49):12745–12751, 1993.

[154] P. Lautenschlager, J. Brickman, Jippe van Ruiten, and Robert J. Meier. Conformations and rotational barriers of aromatic polyesters. *Macromolecules*, 24(6):1284–1292, 1991.

[155] G. C. Rutledge. Rotational isomeric state approach to the single-chain behavior of aromatic polyesters. *Macromolecules*, 25(15):3984–3995, 1992.

[156] T. H. Dunning and P. J. Hay. In H. F. Schaefer, editor, *Methods of Electronic Structure Theory.* Plenum Press, New York, 1977.

[157] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople. Self-consistent molecular orbital methods. xx. a basis set for correlated wave functions. *Journal of Chemical Physics*, 72:650–654, 1980.

[158] Raymond Poirier, R. Kari, and I. G. Csizmadia. *Handbook of Gaussian Basis Sets.* Elsevier, New York, 1985.

[159] S. Huzinaga, J. Andzelm, M. Klobukowski, E. Radzio-Andzelm, Y. Sakai, and H. Tatewaki. *Gaussian Basis Sets for Molecular Calculations.* Elsevier, New York, 1984.

[160] M. Ramek, F. A. Momany, D. M. Miller, and L. Schafer. On the importance of full geometry optimization in correlation-level *ab initio* molecular conformational analyses. *Journal of Molecular Structure*, 375:189–191, 1996.

[161] B. J. Teppen, Ming Cao, R. F. Frey, C. van Alsenoy, D. M. Miller, and L. Schafer. An investigation into intramolecular hydrogen bonding: impact of basis set and electron correlation on the *ab initio* conformational analysis of 1,2-ethandiol and 1,2,3-propanetriol. *Journal of Molecular Structure*, 314:169–190, 1994.

[162] In Michael Springborg, editor, *Density-Functional methods in chemistry and Materials Science.* John Wiley and Sons Ltd, New York, 1997.

[163] S. Tsuzuki, T. Uchimaru, and K. Tanabe. Conformational analysis of n-alkanes using density functional theory. comparison with *ab initio* calculations. *Chem. Phys. Lett.*, 246:9–12, 1995.

[164] G. I. Csonka and L. Sztraka. *Chem. Phys. Lett.*, 233:611, 1995.

[165] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in FORTRAN.* Cambridge University Press, Cambridge, 2 edition, 1992.

[166] P. J. Flory. *Statistical Mechanics of Chain Molecules.* Oxford University Press, New York, 1969.

[167] L. R. Dodd, T. D. Boone, and D. N. Theodorou. A concerted rotation algorithm for atomistic Monte Carlo simulation of polymer melts and glasses. *Molecular Physics*, 78(4):961–996, 1993.

[168] G. C. Rutledge. *personal communication*, M.I.T. Deptartment of Chemical Engineering.

[169] D. N. Theodorou. *Macromolecules*, 21:1391–1400, 1987.

[170] D. N. Theodorou. Variable-density model of polymer melt surfaces: Structure and surface tension. *Macromolecules*, 22:4578–4589, 1989.

[171] S. Affrossman, P. Bertrand, M. Hartshorne, T. Kiff, D. Leonard, R. A. Pethrick, and R. W. Richards. Surface segregation in blends of polystyrene and perfluorohexand double end capped polystyrene studied by static sims, iss, and xps. *Macromolecules*, 29:5432–5437, 1996.

[172] T. F. Schaub, G. J. Kellogg, A. M. Mayes, R. Kulasekere, J. F. Ankner, and H. Kaiser. Surface modification via chain end segregation in polymer blends. *Macromolecules*, 29:3982–3990, 1996.

[173] U. Essmann, L. Perera, M. Berkowitz, T. Darden, H. Lee, and L. Pedersen. Smooth particle mesh Ewald method. *J. Chem. Phys.*, 103(19):8577, 1995.

[174] Anatol Brodsky. Is there predictive value in water computer simulations? *Chemical Physics Letters*, 101(4):3353–3366, 1996.

[175] C. C. J. Roothaan. *Rev. Mod. Phys.*, 23:69, 1951.

[176] C. G. Hall. *Proc. Roy. Soc. (London)*, A205:541, 1951.

[177] J. A. Pople and R. K. Nesbet. *J. Chem. Phys.*, 22:571, 1954.

[178] G. Hummer, D. M. Soumpasis, and M. Neumann. Pair correlations in an nacl-spc water model. simulations versus extended rism computations. *Molecular Physics*, 77(4):769–785, 1992.

[179] M. Neumann. Dipole moment fluctuation formulas in computer simulations of polar systems. *Mol. Phys.*, 50(4):841–858, 1983.

[180] D. M. Heyes. Molecular dynamics of ionic solid and liquid surfaces. *Physical Review B*, 30(4):2182–2201, 1984.

[181] Geoffrey C. Maitland, Maurice Rigby, E. Brian SMith, and William A. Wakeham. *Intermolecular Forces, Their Origin and Determination.* Oxford University Press, New York, 1981.

[182] W. Pauli. *Z. Physik*, 31:765, 1925.

[183] J. C. Slater. *Phys. Rev.*, 34:1293, 1929.

[184] J. C. Slater. *Phys. Rev.*, 36:57, 1930.

[185] S. F. Boys. *Proc. Roy. Soc. (London)*, A200:542, 1950.

[186] C. C. J. Roothan. *Rev. Mod. Phys.*, 32:179, 1960.

[187] J. S. Binkley, J. A. Pople, and P. A. Dobosh. *Mol. Phys.*, 28:1423, 1974.

[188] H. Yasuda, Ashok K. Sharma, and Takeshi Yasuda. Effect of orientation and mobility of polymer molecules at surfaces on contact angle and its hysteresis. *Journal of Polymer Science: Polymer Physics Edition*, 19:1285–1291, 1981.

[189] T. Yasuda, M. Miyama, and H. Yasuda. Dynamics of the surface configuration change of polymers in response to changes in environmental conditions. 2. comparison of changes in air and in liquid water. *Langmuir*, 8(28):1425–1430, 1992.

[190] J. D. Andrade, V. Hlady, and A. P. Wei. Adsorption of complex proteins at interfaces. *Pure and Appl. Chem.*, 64(11):1777–1781, 1992.

[191] Michael Szycher. Thrombosis, hemostasis, and thrombosis at prosthetic interfaces. In M. Szycher, editor, *Biocompatible Polymers, Metals, and Composites*, chapter 1, pages 1–33. TECHNOMIC Publishing o. Inc., Lancaster, PA, 1983.

[192] Leo Vroman. Protein/surface interaction. In M. Szycher, editor, *Biocompatible Polymers, Metals, and Composites*, chapter 4, pages 81–88. TECHNOMIC Publishing o. Inc., Lancaster, PA, 1983.

[193] J. L. Brash. Mechanism of adsorption of proteins to solid surfaces and its relationship to blood compatibility. In M. Szycher, editor, *Biocompatible Polymers, Metals, and Composites*, chapter 2, pages 35–52. TECHNOMIC Publishing o. Inc., Lancaster, PA, 1983.

[194] Masakatsu Watanabe and Martin Karplus. Dynamics of molecules with internal degrees of freedom by multiple time-step methods. *J. Chem. Phys.*, 99(10):8063–8074, 1993.

[195] Wen-Bin Liau and Richard H. Boyd. Structure and packing in crystalline aliphatic polyesters. *Macromolecules*, 23(5):1531–1539, 1990.

[196] Marvin Waldman and A. T. Hagler. New combining rules for rare gas van der waals parameters. *Journal of Computational Chemistry*, 14(9):1077–1084, 1993.

[197] Jon R. Maple, Uri Dinur, and A. T. Hagler. Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proc. Natl. Acad. Sci. USA*, 85:5350–5354, 1988.

[198] M. Depner and B. L. Schurmann. Computer simulation of aromatic polyesters including molecular dynamics.

[199] Kenneth B. Wiberg and Ming Wah Wong. Solvent effects. 4. effect of solvent on the e/z energy difference for methyl formate and methyl acetate. *J. Am. Chem. Soc.*, 115(3):1078–1084, 1993.

[200] R. F. Curl. Microwave spectrum, barrier to internal rotation, and structure of methyl formate. *Journal of Chemical Physics*, 30(6):1529–1536, 1959.

[201] J. K. Wilmshurst. A vibrational assignment for methyl formate and methyl acetate. *Journal of Molecular Spectroscopy*, 1:201–215, 1957.

[202] Alan E. Tonelli and Paul J. Flory. The configuration statistics of random poly(latctic acid) chains. i. experimental results. *Macromolecules*, 2(3):225–227, 1969.

[203] Bernd Kuhn, Michael Ehrig, and Reinhart Ahlrichs. Atomistic modeling of amorphous polymer bulk based on an ab initio optimized force field. *Macromolecules*, 29(11):4051–4059, 1996.

[204] Grant D. Smith, Richard L. Jaffe, and Do Y. Yoon. Conformations of perfluoropoly(oxyethylene) from ab initio electronic structure calculations on model molecules. *Macromolecules*, 28(17):5804–5810, 1995.

[205] D. N. Theodorou. Variable-density model of polymer melt/solid surfaces: Structure, adhesion tension, and surface forces. *Macromolecules*, 22:4589–4597, 1989.

[206] D. N. Theodorou. Structure and thermodynamics of bulk homopolymer/solid interfaces: a site lattice model approach. *Macromolecules*, 21:1400–1410, 1988.

[207] T. Darden, D. York, and L. Pedersen. Particle mesh Ewald: an n log (n) method for Ewald sums in large systems. *J. Chem. Phys.*, 261:563–568, 1993.

[208] Kenneth B. Eisenthal. *Equilibrium and Dynamic Processes at Interfaces by Second Harmonic Generation and Sum Frequency Generation*, volume 43 of *Annual Review of Physical Chemistry*, pages 627–661. Annual Reviews, Inc., Palo Alto, California, 1992.

[209] Y. R. Shen. *The Principles of Nonlinear Optics.* John Wiley and Sons, INC., New York, 1984.