

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

**COMPUTATIONAL STUDIES OF POLYMER ELECTROLYTE
COMPLEXES BY MOLECULAR DYNAMICS AND MOLECULAR
VIBRATIONS BY PRINCIPAL COMPONENT ANALYSIS**

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

Haitao Dong

Norman, Oklahoma

2003

UMI Number: 3109053

Copyright 2003 by
Dong, Haitao

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3109053

Copyright 2004 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

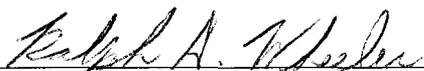
ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© Copyright by Haitao Dong 2003
All Rights Reserved.

**COMPUTATIONAL STUDIES OF POLYMER ELECTROLYTE
COMPLEXES BY MOLECULAR DYNAMICS AND MOLECULAR
VIBRATIONS BY PRINCIPAL COMPONENT ANALYSIS**

A Dissertation APPROVED FOR THE
DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

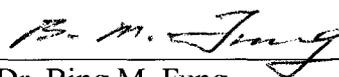
BY



Dr. Ralph A. Wheeler



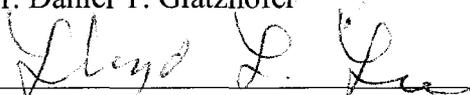
Dr. Roger Frech



Dr. Bing M. Fung



Dr. Daniel T. Glatzhofer



Dr. Lloyd L. Lee

Acknowledgements

There are many people who made the completion of this dissertation possible and others who have made my time as a graduate student less stressful. First of all, I would like to sincerely thank my major adviser, Dr. Ralph Wheeler for coordinating my research and helping me with my graduate study. His direction has been invaluable and it is really through his efforts that I was able to accomplish my goals. I would also thank the other members of my Graduate Committee, Professors Roger Frech, Bing Fung, Daniel Glatzhofer, Lloyd Lee, and Wai Tak Yip. I appreciate all the instructions and advice you have given me. My second thanks go to all professors in the department for your kind help and to the ladies in the department office especially Ms. Arlene Crawford. Thank you very much for all the help you have provided for me since I first entered this country and this city. Next, I would like to say thank you to all past and present members of my research group and collaborator group especially Dr. Jin-Kee Hyun, Dr. Christopher Rhodes, and Scott Boesch for their contributions to my research projects and helpful discussions and suggestions. Dr. Jin-Kee Hyun and I set up computer simulations of tetraglyme and tetraglyme:LiTf systems (see Chapter 3). He also analyzed one part of these two systems and provided the results, which are presented in Table 3.4 and Figure 3.1, 3.2, and 3.8. Dr. Christopher Rhodes kindly provided his spectroscopic study results, which are presented as the IR sections in Table 3.3 and 3.6 and Figure 3.3, 3.6, 3.11, and 3.13. Scott Boesch provided the force field parameters for the triflate anion and dihedral angle of the poly(ethylenimine) chain (see Chapter 2). I would also like to thank the Department of Chemistry and

Biochemistry, Graduate College, and the International Student Office. Finally, I would like to thank my parents, my sister, and all my friends. Without your support, I could not have the opportunity to go abroad and pursue my academic goals. Specific acknowledgements regarding financial supports and supercomputer time allocations can also be found in the end of each of the chapters.

Table of Contents

List of Figures.....	ix
List of Tables.....	xiii
Abstract.....	xv

Part I

Chapter 1. Introduction of computer-aided molecular dynamics simulations and structural analyses of poly(ethylene oxide), poly(ethyleneimine), and their polymer:salt complexes.	1-13
1.1. Solid polymer electrolytes.	1
1.2. Molecular dynamics simulation and model polymer electrolytes.	6
1.3. Methods for analyzing polymer structural properties.	10
1.4. References.....	12
Chapter 2. Molecular dynamics simulations and structural comparisons of amorphous poly(ethylene oxide) and poly(ethyleneimine) models.	14-47
2.1. Introduction.	14
2.2. Simulation methodology.	18
2.2.1. Models.	18
2.2.2. Simulation details.	18
2.2.3. Force field parameters.	20
2.3. Results and discussions.	23
2.3.1. Chain dimensions.	23
2.3.2. Dihedral angle distribution.	25
2.3.3. Conformational triads population analysis.	27
2.3.4. Radial distribution function analysis for PEI-4.....	29
2.3.5. Helical structure analysis of PEO-4 and PEI-4 chains.	40
2.4. Conclusions.	42
2.5. Acknowledgements.	45
2.6. References.	46
Chapter 3. Molecular dynamics simulations and vibrational spectroscopic studies of local structure in amorphous $\text{CH}_3\text{O}(\text{CH}_2\text{CH}_2\text{O})_4\text{CH}_3$: XCF_3SO_3 (X=Li, Na) systems.	

.....	48-108
3.1. Introduction.	48
3.2. Simulation methodology.	53
3.3. Results and discussion (I) – the tetraglyme:LiCF ₃ SO ₃ complex.	56
3.3.1. Chain dimensions.	56
3.3.2. Conformation of pure tetraglyme and its changes upon complexation.	57
3.3.3. Li ⁺ ion complexation.	64
3.3.4. Ionic association.	69
3.4. Conclusions on the tetraglyme:LiCF ₃ SO ₃ system.	76
3.5. Results and discussion (II) – the tetraglyme:NaCF ₃ SO ₃ complex.	79
3.5.1. Tetraglyme chain dimensions.	79
3.5.2. Conformations of tetraglyme chains.....	81
3.5.3. Na ⁺ -ether oxygen complexation.	89
3.6. Comparison of ionic association in two systems.	93
3.7. Conclusions on the tetraglyme:NaCF ₃ SO ₃ system.	100
3.8. Acknowledgements.	103
3.9. References.	104

Part II

Chapter 4. Calculating accurate vibrational frequencies: the method of principal component analysis and its application to molecular dynamics simulations.	109-154
4.1. Quasiharmonic analysis and background information of principal component analysis.	109
4.2. PCA comparison with conventional spectrum estimation methods.	112
4.2.1. PCA and normal mode analysis.	112
4.2.2. PCA and Fourier transform.	115
4.2.3. Numerical and water vibrational tests of PCA. ...	116
4.3. Methodology.	123
4.3.1. The theory of principal component analysis and its application to molecular vibrations.....	123
4.3.2. A modification in calculating vibrational frequencies.	127
4.3.3. Translation and rotation present in molecular trajectories.	130

4.3.4. Frequencies and modes of molecules with flexible structures.	134
4.4. Computational details.	139
4.5. Results and discussion.	140
4.6. Conclusions.....	149
4.7. Acknowledgement.	150
4.8. References.	151
Chapter 5. Vibrational energy relaxation in water studied by principal mode analysis of quantum mechanical/molecular mechanical (QM/MM) simulations.	155-200
5.1. Introduction.	155
5.2. Principal mode analysis: principal component analysis and its application to molecular vibrations.	161
5.3. Molecular dynamics or Qm/MM simulations of vibrational energy relaxation.	163
5.4. Computational method.	167
5.5. Results and discussion.....	169
5.5.1. Principal mode analysis of isolated water molecules and liquid water.	169
5.5.2. Vibrational energy transfer between vibrational modes/IVR in isolated water molecules.	172
5.5.3. Vibrational energy relaxation of liquid water....	179
5.5.4. VER from the solvent's perspective.	188
5.6. Conclusions.	196
5.7. References.	199
Chapter 6. Summary and prospective research in the application of principal component analysis to molecular vibrations.	203-207

List of Figures

- Figure 2.1.** (a, upper) Average population density distribution of dihedral angles along the C-X-C-C backbone atom sequence for PEO-4 and PEI-4 at 300 K. (b, lower) Average population density distribution of dihedral angles along the X-C-C-X backbone atom sequence for PEO-4 and PEI-4 at 300 K.26
- Figure 2.2.** Population analysis of conformational triads along the C-X-C-C-X-C backbone atom sequence for PEO-4 and PEI-4 at 300 K.28
- Figure 2.3.** The total radial distribution functions (RDFs) of N...H pairs and its components for PEI-4 at 300K. The adjacent (NH groups are separated by one ethyl groups), non-adjacent (NH groups are separated by at least two ethyl groups but are in the same molecule), and intermolecular (NH groups are from different molecules).31
- Figure 2.4.** The total radial distribution functions of H...H pairs and its components for PEI-4 at 300K. The adjacent, non-adjacent, and intermolecular RDF are represented as dashed, dash dotted, and dotted lines, respectively.....32
- Figure 2.5.** Diagrams of two distinguishable Hydrogen-bonding structures in C-(NH)-C-C-(NH)-C units in PEI-4.35
- Figure 2.6.** The contributions from the double H-bonding and the single H-bonding structures to the adjacent N...H RDF.37
- Figure 2.7.** The H-bond angle distribution for the double H-bonding and the single H-bonding structures.39
- Figure 3.1(a).** Average population density distributions of tetraglyme backbone dihedral angles around C-C bonds at 400K. The dashed line is for the torsions within the inside regions of tetraglyme:LiTf melts.59
- Figure 3.1(b).** Average population density distributions of tetraglyme backbone dihedral angles around C-O bonds at 400K. The dashed line is for the torsions within the inside regions of tetraglyme:LiTf melts.59
- Figure 3.2.** Average population density distributions of tetraglyme backbone dihedral angles around C-O bonds at 400K. The dashed line is for the closest torsion of the outside region of tetraglyme:LiTf melts.61

Figure 3.3. Raman spectra of pure tetraglyme and tetraglyme:LiTf at 295K and 400K in the region 780 cm^{-1} to 900 cm^{-1}	63
Figure 3.4. Population analysis of conformational triads (C-O-C-C-O-C atom sequences) for tetraglyme and tetraglyme:LiTf melts at 300K and 400K. Error bars at the top of each column are calculated standard deviations.	65
Figure 3.5. Radial distribution function between Li^+ and all oxygen atoms in a tetraglyme:LiTf melt at 300K. The first minimum occurs at $r = 2.8 \text{ \AA}$ and is used as the cutoff distance of the first coordination shell.	67
Figure 3.6. Infrared spectra of tetraglyme:LiTf (with an ether oxygen:LiTf ratio of 10:1) in the $\delta_s(\text{CF}_3)$ region at 295K and 400K. Curve-fitted peaks are shown below the recorded spectra.	70
Figure 3.7. Radial distribution functions between Li^+ and the triflate center of mass at 300K and 400K. The first peaks occur at $r = 3.4 \text{ \AA}$ and the first minima at $r = 4.5 \text{ \AA}$ at both temperatures.	72
Figure 3.8. The number of Li^+ ions as a function of the distance between Li^+ and the center of mass of triflate at 300K and 400K. The flat parts of the curves give the numbers of Li^+ ions associated with one triflate ion in the first ion association shell.	75
Figure 3.9. Time-average population density distributions of tetraglyme backbone torsions around C-C bonds in pure tetraglyme (solid line), tetraglyme:NaTf (dashed line), and tetraglyme:LiTf (dotted line) systems at 400K.	82
Figure 3.10. Populations of conformational triads (C-O-C-C-O-C backbone atom sequences) for tetraglyme and tetraglyme:NaTf complex at 300 and 400K. Error bars on each column are calculated standard deviations.	86
Figure 3.11. Infrared spectra in the 800-900 cm^{-1} region of tetraglyme and tetraglyme-NaTf 10:1 solutions.	88
Figure 3.12. Radial distribution functions (RDF) between Na^+ and all oxygens (solid line), Na^+ and ether oxygens (dotted line), and Na^+ and triflate oxygens (dashed line) in tetraglyme: NaTf complex at 400K.	90
Figure 3.13. Curve-fit IR spectrum in the 740-780 cm^{-1} region of a tetraglyme-NaTf 10:1 solution indicating the presence of different triflate species.	94

Figure 3.14. Radial distribution functions (RDF) between Na⁺ and the triflate center of mass in tetraglyme:NaTf complex at 300 (dashed line) and 400K (solid line).97

Figure 4.1. The schematic sketch of the potential energy surface (solid line) being approximated by PCA (dashed line) and NMA (normal mode analysis, dotted line).114

Figure 5.1. Total vibrational energy of the bending, symmetric stretching, and antisymmetric stretching modes for water in isolation with the symmetric stretching mode excited at the beginning of the simulation.174

Figure 5.2. The autocorrelation functions for the total vibrational energy presented in Figure 6.1 for water in isolation. The insets are the power spectra of the correlation functions calculated by fast Fourier transform (FFT).175

Figure 5.3. Total vibrational energy of the bending, symmetric stretching, and antisymmetric stretching modes for water in isolation with the antisymmetric stretching mode excited at the beginning of the simulation.177

Figure 5.4. The autocorrelation functions for the total vibrational energy presented in Figure 6.3 for water in isolation. The insets are the power spectra of the correlation functions calculated by fast Fourier transform (FFT).178

Figure 5.5. The normalized total vibrational energy of the bending (A), symmetric stretching (B), and antisymmetric stretching (C) modes for one sample QM/MM trajectory of liquid water with the symmetric stretching mode of the QM water excited at the beginning of the simulation.....181-182

Figure 5.6. The normalized total vibrational energy of the bending (A), symmetric stretching (B), and antisymmetric stretching (C) modes for one sample QM/MM trajectory of liquid water with the antisymmetric stretching mode of the QM water excited at the beginning of the simulation.....184-185

Figure 5.7. The time scales as functions of the transition frequency calculated from Fourier transformed force-force correlation function for liquid water with the symmetric stretching mode excited. The forces are from the solvent acting on

each of the solute vibrational coordinates. The inset shows the low frequency part of the functions.190

Figure 5.8. The time scales as functions of the transition frequency calculated from Fourier transformed force-force correlation function for liquid water with the antisymmetric stretching mode excited. The forces are from the solvent acting on each of the solute vibrational coordinates. The inset shows the low frequency part of the functions.194

List of Tables

Table 2.1. Force Field Parameters for PEO-4.	21
Table 2.2. Force Field Parameters for PEI-4.	22
Table 2.3. MD results of mean-square radii of gyration, $\langle S^2 \rangle$, and end-to-end distances, $\langle R^2 \rangle$, and characteristic ratio, C_n , for PEO-4 and PEI-4 at 300K.	24
Table 3.1. The force field parameters for Li^+ , Na^+ , and Tf^- ions.....	55
Table 3.2. The positions of the first peaks (1^{st} p) and the first minima (1^{st} m) in radial distribution functions for tetraglyme: LiCF_3SO_3 from MD simulations performed at 300K and 400K.	66
Table 3.3. Comparison of the percent contributions from different Li-Tf species to the total integrated intensity of the $\delta_s(\text{CF}_3)$ band observed in IR experiments with percentages of Li^+ cations bound up in different Li-Tf species in molecular dynamics (MD) simulations.	71
Table 3.4. The mean-square radius of gyration, $\langle S^2 \rangle$, mean-square backbone end-to-end distance, $\langle R^2 \rangle$, and characteristic ratio C_n for pure tetraglyme, tetraglyme: NaTf , and tetraglyme: LiTf samples at $T=300$ and 400 K.	80
Table 3.5. The coordination numbers of Na^+ and Li^+ in tetraglyme: NaTf and tetraglyme: LiTf samples at 300K and 400K.	91
Table 3.6. Comparison of calculated populations of Na^+ - Tf^- associated “anion aggregates” (see text) in tetraglyme: NaTf samples from MD results at 300K and 400K with IR experimental data at 398K in parentheses.	95
Table 3.7. Coordination geometries for Na coordinated to triflate in tetraglyme: NaTf calculated from MD trajectories at 300 K and 400 K.	99
Table 4.1. Comparison of vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) with results by standard quantum chemical calculation and experiment for isolated water, water dimer, and water molecules in liquid phase.	121

Table 4.2. Comparison of frequencies calculated by using the fast Fourier transform (FFT), maximum entropy method (MEM), and principal component analysis (PCA) with analytical solutions to harmonic, quartic, and Langevin oscillators or experimental results for a single water molecule in isolation and in liquid phase.	122
Table 4.3. Vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) of QM/MM simulation trajectories of isolated water molecules with various modifications.	146
Table 4.4. Vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) of QM/MM simulation trajectories of isolated methane molecules with various modifications as the same presented in Table 4.3.	146
Table 4.5. Vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) of QM/MM simulation (with B3LYP/6-311++G(2d,2p) method/basis) trajectories of isolated water dimers with and without dividing the whole flexible structure in parts.	147
Table 4.6. Vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) of QM/MM simulation trajectories of isolated ethane molecules with and without dividing the whole flexible structure in parts.	148
Table 5.1. Frequencies for the bending, symmetric stretching, and antisymmetric stretching modes calculated by PCA of QM/MM trajectories of isolated water and liquid water molecules.	171
Table 5.2. Time scales of vibrational energy relaxation of water in the gas and liquid phase.	187

Abstract

The first three chapters of this dissertation are concerned with the structural properties of solid polymer electrolyte systems. Model polymers of poly(ethylene oxide) (PEO) and poly(ethylenimine) (PEI) containing four repeat units were studied by using molecular dynamics simulations. The chain dimensions, dihedral angle distributions, triad conformations were analyzed for PEO, PEI, and PEO:salt systems. The special ability of PEI chains to form two kinds of intra-chain hydrogen bonding, the single and double hydrogen bonding, makes the conformations of PEI chains more compact than PEO chains. The structural properties of PEO:LiCF₃SO₃ (lithium triflate) and PEO:NaCF₃SO₃ (having an ether oxygen:salt ratio of 10:1) complexes including the effects of oxygen-salt coordination on the polymer conformations and the ionic aggregation were also analyzed. The results show that the PEO chains contribute fewer oxygens (1.6 to Li⁺ and 2.2 to Na⁺ at 300K) to the cations than the triflate ions (4.6 and 4.9) and high temperature weakens PEO chain-cation coordination further. The species and the populations of different ionic aggregates calculated from MD simulations agree well with experimental values. All the results will help the understanding of ionic conductivity mechanisms in polymers.

The second part (Chapter 4 and 5) focuses on computing vibrational frequencies and modes by using a new technique, principal component analysis (PCA) which is broadly used in signal processing, pattern recognition, and multivariate analysis. The advantages of PCA to incorporate anharmonicity in the calculated spectra over the conventional normal mode analysis and Fourier transform are presented in Chapter 4.

PCA-derived frequencies of harmonic, anharmonic (with a quartic term in the potential energy), and Langevin oscillators and water molecules are more accurate than the frequencies calculated by normal mode analysis and Fourier transforms, in comparison with analytical or experimental results. New procedures in order to improve the PCA method including the determination of the vibrational energy, removal of the translations and rotations from the trajectories, and the treatment of flexible molecules based on fragments are proposed. More test cases of water, methane, water dimer, and ethane are provided to show that the frequencies are greatly improved by applying these new procedures in addition to the original PCA method. The last chapter is a demonstration of one of PCA's abilities to study molecular vibrational problems. Vibrational energy relaxation (VER) in liquid water has been previously studied by spectroscopy and molecular dynamics because of its importance in understanding aqueous and biological processes. PCA is able to set up a vibrational mode frame with some anharmonicity included. In this frame, by examining the vibrational energy possessed by each of the vibrational modes as a function of time, two distinguishable VER time scales of 0.6-0.7 ps and 0.2-0.3 ps in liquid water were obtained that agree well in magnitude with experiment. However, PCA studies indicate that the longer time scale is attributed to energy transfer out of the initially excited vibrational modes and the shorter time scale refers to energy transfer out of the secondarily excited mode, which is different from the mechanism inferred spectroscopically.

Part I

Chapter 1

Introduction to molecular dynamics simulations and structural analyses of poly(ethylene oxide), poly(ethyleneimine), and their polymer:salt complexes.

1.1. Solid polymer electrolytes.

Our daily life has been greatly changed by the widespread applications of polymers. Solid polymer electrolytes (SPE), as one substitute for conventional electrolyte solution-based materials, have been used in the fabrication of high energy-density batteries. [1-4] Compared to conventional liquid-based electrolyte systems, SPE's have desirable mechanical properties to be used as the intermediate materials between electrodes, and therefore reduces problems of leakage and improves operation safety. However, several shortcomings have hindered the industrial production of high energy-density batteries using SPE. As an ideal polymer host for SPE, the polymer matrix must have relatively strong interactions with the salts to dissolve and separate the ions. On the other hand, these interactions cannot be strong enough to trap the charge-carrying ions in the local polymer matrix. In a good SPE candidate, the polymer-ion interactions must have reached a balance between these two aspects to produce both strong enough interactions to dissolve the salt and weak enough interactions to ensure high ionic conductivity.

Poly(ethylene oxide) (PEO, $(-\text{CH}_2\text{CH}_2\text{O}-)_n$) was first suggested as a candidate polymer host for SPE more than 20 years ago. [5] The strong interactions between the electronegative ether oxygens and the salt cations enable PEO to dissolve a variety of inorganic salts. Due to the linear chain structure of PEO, one would expect a high regularity in the entire structure of the polymer matrix. Indeed, it has been reported that the melting point of crystalline PEO is relatively high at 338 K and the crystallinity can be as high as 85%. [3] The role that the crystalline polymer plays in ionic conduction through the polymer matrix is still not fully understood. But it is generally considered true that the amorphous portions of polymer contribute significantly to the conduction. Unlike conventional materials in which all particles would be fixed in the “crystal grids” below the melting temperature, the crystalline and amorphous phases can coexist in the temperature below the melting temperature, which is a unique property of polymers. When a polymer material is in between the melting temperature and the glass transition temperature (for PEO, this temperature is 213 K [3]), a small-scale local motion of polymer chains in the amorphous phase is allowed. Because of the local mobility of these “glassy” portions, ionic conduction through the polymer matrix is made possible even though the polymer is not fully liquid-like. There is some NMR evidence to support this consideration. [6] However, the ionic conductivity of PEO-based SPE is still too small for any practical use. For PEO:LiX (X=ClO₄, CF₃SO₃), the ionic conductivity is in the range of 10⁻³-10⁻⁴ S cm⁻¹ at 373K but falls to 10⁻⁶-10⁻⁸ S cm⁻¹ at room temperature, which is four orders in magnitude smaller than the conventional liquid electrolytes. [7] Therefore,

improvement of the polymer's ionic conduction performance has been heavily pursued by altering the structures of polymer hosts and species of inorganic salts. Before a desirable polymer structure can be designed, an in-depth understanding of the mechanisms of ionic conduction needs to be obtained. Due to the simplicity in structures of PEO and PEO alkali salt systems, they serve as a practical beginning stage for making connections between ionic conduction and structural and dynamic properties of polymer hosts.

An alternative polymer to PEO, poly(ethylenimine) (PEI, $(-\text{CH}_2\text{CH}_2\text{NH}-)_n$) with the ether oxygen in PEO replaced by an NH group, is a step further toward structural complexity. It would be expected that PEI behaves similarly in ionic conduction to PEO since the electronegative nitrogen interacts with salt cations in a fashion similar to the oxygen in PEO. However, the hydrogen-bonding ability of the NH groups makes PEI a very different system in terms of both structural and dynamic features. [8] Moreover, the nitrogen allows a third covalent bond to be formed thus making side-chain structures possible in the polymer. All these added flexibilities in the PEI structure have made PEI interesting for comparison with PEO. The melting point for a semicrystalline PEI sample is 332 K and the glass transition temperature is 250 K, comparable with PEO. [9] The amorphous state of PEI, however, may not resemble the structure of PEO of the same phase. The NH groups in PEI can act as both H-bond donors and acceptors. This additional interaction therefore will change the intra-chain configurations and modify inter-chain forces. The lack of the H-bonding groups

in PEO may then cause substantial differences between PEO- and PEI-based polymer electrolytes.

The main focus of this part of my research is on the structural properties of the two polymers, PEO and PEI, and the inorganic salt-dissolved polymer electrolyte systems. In Chapter 2, the structures of amorphous PEO and PEI are studied using molecular dynamics simulations. Several structural properties including the average polymer chain dimensions, dihedral angle distributions along different chain atomic sequences, triad conformational analysis, and the intra- and inter-chain hydrogen bonding are reported and compared between the two polymers. Radial distribution function analysis reveals that H-bonds exist in a number of different ways between two adjacent NH groups in the same PEI chains. All intra-chain H-bonds are found to be nonlinear and longer than typical ones. These results are consistent with an *ab initio* study of a methyl capped monomer of PEI, dimethylethylenediamine. [10] Finally, the conformations of PEO and PEI at different temperatures are discussed.

In Chapter 3 the molecular dynamics simulation study begins with the calculations of chain dimensions and identification of conformations of PEO and then the examination of the chain conformational changes caused by dissolving LiCF_3SO_3 (lithium trifluoromethanesulphonate or lithium triflate) in the polymer amorphous phase. The concentration of this polymer:salt complex has a 10:1 ether oxygen: Li^+ ratio. Dihedral angle population density distributions for PEO chains show that the *trans* conformation is favored around C-O bonds while the more compact *gauche* conformation is favored around C-C bonds and is enforced by the coordination of

adjacent oxygens to Li^+ . Calculated populations of the conformational triads indicate that the most stable conformation of the O-C-C-O dihedral angle sequence around Li^+ -oxygens is tgt along PEO chains. The Li^+ coordination numbers are reported and the populations of Li^+ - CF_3SO_3^- associated species are compared with vibrational spectral data. Again, the temperature dependence of these properties is studied by comparing them at two different temperatures. Another system reported in Chapter 3 is PEO: NaCF_3SO_3 10:1 (ether oxygen: Na^+) complex. The conformational analyses are the same as in the lithium salt system and results from both polymer:salt complexes are compared. The two systems show very similar trends in the structural properties and their temperature dependence. One interesting finding is that unlike Li^+ in the tetraglyme: LiCF_3SO_3 system, in which the monodentate coordination of Li^+ increases at higher temperature, the bidentate coordination of Na^+ by CF_3SO_3^- becomes more favorable at higher temperature, with a decrease in monodentate and tridentate coordinations.

1.2. Molecular dynamics simulation and model polymer electrolytes.

Molecular dynamics (MD) is a technique to solve Newton's equations of motion for a system with N particles

$$\frac{d}{dt} \frac{\partial K}{\partial \dot{\mathbf{r}}_i} + \frac{\partial V}{\partial \mathbf{r}_i} = 0 \quad i = 1, 2, \dots, N, \quad (1.1)$$

where \mathbf{r} 's are position vectors in Cartesian coordinates for the particles and V and K are the system's potential and kinetic energy, which are considered to be functions of the coordinates \mathbf{r} 's and their time derivatives, respectively. With the usual definition of kinetic energy, eq. (1.1) becomes

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i = -\frac{\partial V}{\partial \mathbf{r}_i} \quad (1.2)$$

with m_i the mass and \mathbf{f}_i the force acting on the particle i . A standard method for solving eq. (1.2) is the finite difference approach. [11] With the initial positions and velocities of all particles known at t_0 , the new positions and velocities at a later time $t_0 + \delta t$ can be obtained through calculating the forces from the initial positions using eq. (1.2). The new positions are then used to calculate next step's forces and the equations of motion are solved on a step-by-step basis. The step size δt must be set small enough to allow a high accuracy in integrating eq. (1.2) for positions and velocities. After a large number of steps are calculated, any system's structural and thermodynamic properties of interest can be obtained by averaging over all time steps.

One of the most commonly used methods of integrating eq. (1.2) is the so-called "leapfrog" scheme of the Verlet algorithm. [12-14] With the forces evaluated at a

time t , the new positions and velocities are calculated from the current positions $\mathbf{r}(t)$, the mid-step velocities $\mathbf{v}(t-1/2\delta t)$, and the force-derived accelerations $\mathbf{a}(t)$ as

$$\begin{aligned}\mathbf{r}(t + \delta t) &= \mathbf{r}(t) + \delta t \mathbf{v}(t + 1/2\delta t) \\ \mathbf{v}(t + 1/2\delta t) &= \mathbf{v}(t - 1/2\delta t) + \delta t \mathbf{a}(t)\end{aligned}\tag{1.3}$$

In this algorithm the velocities at every whole number step are never calculated and the time differential between the velocities and positions is always a half time step.

This is where the algorithm gets the name “leapfrog”. The current time velocities can be calculated as

$$2\mathbf{v}(t) = \mathbf{v}(t + 1/2\delta t) + \mathbf{v}(t - 1/2\delta t),\tag{1.4}$$

if both positions and velocities are required for evaluating any properties of the system.

Using molecular dynamics to solve the equations of motions exactly is not only possible because of the simplicity and the repeatability of the algorithm introduced above, is but also necessary when the systems become large. Eq. (1.1) and (1.2) are sets of $3N$ equations (if the positions and velocities are in Cartesian coordinates) and exact solutions are impossible to obtain when N becomes large, which is often the case for molecular systems of interests. MD simulations are therefore very valuable to experimentalists because they make connections between microscopic and macroscopic properties. For many problems in statistical mechanics that are only solvable by approximate methods or cannot be described by existing theories, computer simulation has been proven to be a useful tool. The roles played by computer simulation in solving those problems are twofold: starting with a model

constructed based on theories or hypotheses for the molecular system of interest, computer simulation gives the positions and velocities of microscopic particles and calculates macroscopic properties. These results may be compared with experimental results to test the validation of the model and the theories behind it. On the other hand, if the model is a good one, insights to the system of interest can be offered to assist experimentalists in the interpretation of experimental data and predictions can also be made by computer simulations.

The construction of the model used in a computer MD simulation is very critical in making a successful connection between the theory and experiment as previously discussed. The goal of generating a good model is to mimic the real potentials at a relatively high level of accuracy for the systems with a set of simple equations to be easily solved by computer. For the polymer electrolyte systems, there are five specific categories of interactions present between atoms: the bond stretching, angle bending, torsion twisting, van der Waals interaction, and electrostatic attraction and repulsion between charged atoms. A typical equation set for evaluating these potentials used by AMBER, a commonly used MD simulation software package, [15,16] is

$$\begin{aligned}
 V = & \frac{1}{2} \sum_{\text{bonds}} K_r (r - r_{eq})^2 + \frac{1}{2} \sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2 \\
 & + \frac{1}{2} \sum_{\substack{\text{dihedral} \\ \text{angles}}} k_\phi [1 + \cos(n\phi - \gamma)] + \sum_{\substack{\text{nonbonded} \\ \text{atom pairs}}} \left[\frac{A}{r^{12}} - \frac{B}{r^6} + \frac{q_1 q_2}{r} \right]
 \end{aligned} \tag{1.5}$$

where r_{eq} is the equilibrium bond distance, θ_{eq} is the equilibrium bond angle, ϕ is the dihedral angle, r is the distance between two atoms, q_1 and q_2 are the atomic charges, K 's are force constants, and A and B are constants determined by the atomic species.

The values of the constants (K 's, r_{eq} , θ_{eq} , etc.) are called a “force field” meaning a whole description of the forces (which are the first derivatives of potential energy with respect to the atomic positions) and are provided in the AMBER package or in the literature.[17] With eq. (1.5) set up in a computer simulation, the initial positions of atoms in a simulation “box”, which includes a sufficiently large number of molecules to represent the amorphous polymers (for the PEO and PEI studies, they are 50 four repeat-unit, short-chain oligomers), are input in the computer. The initial velocities necessary to integrate eq. (1.2) using the leapfrog-Verlet algorithm, are assigned according to a Gaussian distribution at a desired temperature. Then the computer simulation is executed step by step using eq. (1.3) and (1.4). There are more details in the technical aspects of MD simulations including periodic boundary conditions, cutoffs, restraints, and particle mesh Ewald sums for electrostatic potentials. Thorough discussions on these topics can be found elsewhere. [11,18]

1.3. Methods for analyzing polymer structural properties.

Polymer materials, in contrast to other materials, do not have exact chemical formulas due to the fact that individual molecular chain lengths and structures may vary. Therefore special quantities have been introduced in polymer systems, such as weight-average molecular weight, polydispersity, mean square radius of gyration, and mean square end-to-end distance. The latter two, expressed as $\langle S^2 \rangle$ and $\langle R^2 \rangle$, are used to determine the overall spatial extent of the model polymers in MD simulations, for comparison with experiment. [19,20] The mean square radius of gyration, which is an average of individual atoms' mass weighted distances from the center of mass, accounts for the spatial distribution of atoms in a polymer chain. The mean square end-to-end distance, being exact as its name indicates, simply measures the length spanned by the polymer chain. These quantities increase as the polymer chains become longer but the characteristic ratio, defined as $C_n = \langle R^2 \rangle / n l^2$ with n the number of repeat units and l the average single bond length in the chain, allows a direct comparison between polymer chains with different sizes. This ratio is theoretically predicted as 1 for the random walk-like polymer chains and usually greater than 1 for real systems. [19] In the computer simulations of PEO and PEI, the characteristic ratios of the model oligomers are compared with results from neutron scattering experiment to test the validity of the models of four repeat units used to represent the amorphous polymers.

Because in a computer simulation the atomic positions and velocities are given at any time, direct analyses of a molecule's structures are possible to understand a

system's macroscopic properties at the molecular level. Radial distribution functions, [11] among the most commonly used mathematical tool for direct structural and dynamic analyses, are used to analyze the computer simulations of PEO and PEI to reveal polymer chain-alkali cation coordination structures, cation-anion aggregations, and hydrogen bonding structures in model PEI systems. A pair radial distribution function $g(\mathbf{r})$ is

$$g(\mathbf{r}) = \rho^{-2} \left\langle \sum_i \sum_{j \neq i} \delta(\mathbf{r}_i) \delta(\mathbf{r}_j - \mathbf{r}) \right\rangle = \frac{V}{N^2} \left\langle \sum_i \sum_{j \neq i} \delta(\mathbf{r} - \mathbf{r}_{ij}) \right\rangle \quad (1.6)$$

and gives the probability of finding a pair of atoms with a distance \mathbf{r} between them, compared to the probabilities if those atoms are evenly distributed in space. From the computer simulated atomic positions, the populations of atom pairs of interest with a distance \mathbf{r} apart are calculated and compared with the ideal case in which the interactions between the atoms are uniform so the atoms would fill the space in a completely even manner. In the PEO and PEI polymer and polymer:salt systems, the hydrogen bond between NH groups and the Coulombic interactions between charged ions and partially charged groups are readily studied by the pair radial distribution analysis.

1.4. References.

- [1] Gray, F. M. *Solid Polymer Electrolytes. Fundamentals and Technological Applications*; VCH: New York, 1991.
- [2] Gray, F. M. *Polymer Electrolytes*; The Royal Society of Chemistry: Cambridge, UK, 1997.
- [3] MacCallum, J. R.; Vincent, C. A. *Polymer Electrolyte Reviews*; Elsevier: New York, 1987; Vol. 1 & 2.
- [4] Ratner, M. A.; Shriver, D. F. *Chem. Rev.* **1988**, *88*, 109.
- [5] Armand, M. B.; Chabagno, J. M.; Duclot, M. *Second International Meeting on Solid Electrolytes. St Andrews, Scotland 1978*, Extended Abstract.
- [6] Berthier, C.; Gorecki, W.; Minier, M.; Armand, M. B.; Chabagno, J. M.; Rigaud, P. *Solid State Ionics* **1983**, *11*, 91.
- [7] Albinsson, I.; Mellander, B.-E.; Stevens, J. R. *Polymer* **1991**, *32*, 2712.
- [8] Chatani, Y.; Kobatake, T.; Tadokoro, H. *Macromolecules* **1982**, *15*, 170.
- [9] Tanaka, R.; Fujita, T.; Nishibayashi, H.; Saito, S. *Solid State Ionics* **1993**, *60*, 119.
- [10] Boesch, S. E.; York, S. S.; Frech, R.; Wheeler, R. A. *J Chem Soc, Phys Chem Comm* **2001**, *1*, (electronic publication).
- [11] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.
- [12] Verlet, L. *Phys. Rev.* **1967**, *159*, 98.
- [13] Hockney, R. W. *Methods Comput. Phys.* **1970**, *9*, 136.
- [14] Potter, D. *Computational Physics*; Wiley: New York, 1972.
- [15] Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 5*; University of California-San Francisco: San Francisco, 1997.

- [16] Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham III, T. E.; DeBolt, S.; Ferguson, D. M.; Seibel, G. L.; Kollman, P. A. *Comp. Phys. Comm.* **1995**, *91*, 1.
- [17] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J Am Chem Soc* **1995**, *117*, 5179.
- [18] Cramer, C. J. *Essentials of Computational Chemistry*; John Wiley & Sons, Inc.: New York, 2002.
- [19] Flory, P. J. *The Statistical Mechanics of Chain Molecules*; Hanser: New York, 1988.
- [20] Painter, P. C.; Coleman, M. M. *Fundamentals of Polymer Science*; Technomic Publishing Company, Inc.: Lancaster, 1997.

Chapter 2

Molecular dynamics simulations and structural comparisons of amorphous poly(ethylene oxide) and poly(ethylenimine) models. ^a

2.1. Introduction.

The prospect of high energy-density batteries has been greatly improved by the application of solid polymer electrolytes (SPE) because of the polymeric materials' mechanical properties. A candidate polymer matrix should have strong enough interactions with ions to dissolve and dissociate salts into the polymer matrix. [1] NMR evidence has shown that ion movement takes place mainly in an amorphous phase in polymers, [2] which is generally present at temperatures above the glass transition temperature. When a polymer system is above this temperature, local motion of polymer chains is less restricted and can become associated with ion movement. Therefore, SPE and SPE-salt complexes with glass transition temperatures at or below ambient temperatures are desirable.

It has been more than 20 years since poly(ethylene oxide) (PEO, $(-\text{CH}_2\text{CH}_2\text{O}-)_n$) was first suggested as one possible polymer matrix. [3] The melting point of crystalline PEO is 338K, and due to the regularity of the linear structure of PEO, its crystallinity can be as high as 85%. [4] X-ray diffraction studies have revealed that linear PEO adopts a $(7/2)$ helical structure, indicating seven repeat units and two turns

^a This chapter is reproduced with permission from : Dong, H.; Hyun, J-K.; Durham, C.; Wheeler, R. A. *Polymer* **2001**, *42*, 7809-7817. © 2001 Published by Elsevier Science Ltd.

per crystal period. [5] The glass transition temperature of amorphous PEO is well below room temperature, at approximately 213K. [4] The strong interactions between the electronegative ether oxygen and salt cations enable PEO to dissolve a variety of inorganic salts. PEO has therefore become a common host polymer matrix in high energy-density batteries. In fact, PEO and PEO-based SPEs have been subjected to intensive study in recent years.

Structural properties of polymer matrices are one important aspect of a full understanding of ionic conductivity mechanisms in the SPEs. Molecular dynamics (MD) simulations have shown their power to represent real systems and provide insight into molecular structures. [6-9] An MD simulation of a model for crystalline PEO including sixteen 14-repeat unit chains in 8 crystallographic cells (2×2×2) was carried out at 300K and the equilibrium structure was compared with the X-ray crystal structure. [6] The good agreement suggested that the force field parameters reproduce well the crystalline PEO/PEO interactions. In subsequent MD work, analogues of PEO, $C_2H_5-O(-CH_2CH_2O-)_n C_2H_5$ with $n=1-9, 12, 15, 32, 65, 165,$ and 332, were simulated at 400K to represent the pure melts. [7] The study showed that the dihedral angles along the C-O-C-C atom sequence favored the trans over the gauche conformation, while the gauche conformation dominated the dihedral angles along the O-C-C-O sequence. Lin *et al.* used a new procedure to generate PEO models by simulating the polymerization process of dimethyl ether in order to represent a multiply-dispersed sample of polymeric materials. [8] The authors reported structural and dynamical properties by presenting radial distribution

functions and comparisons of calculated vibrational spectra with experiment.

Poly(oxyethylene) (POE, $H(-CH_2CH_2O-)_nH$), has the same chain structure as PEO. A 12-repeat unit model of POE was simulated by Smith *et al.* at 300K, 333K, 373K, and 450K. [9] The mean square radii of gyration $\langle S^2 \rangle$ and end-to-end distances $\langle R^2 \rangle$ of this model agreed with small-angle neutron scattering (SANS) results for POE samples and followed the same temperature dependence. All the work above indicates that molecular dynamics simulations have become a fruitful approach to understand macromolecular structure.

Poly(ethylenimine) (PEI, $(-CH_2CH_2NH-)_n$) is an alternative to PEO because of their similar structures and ability to dissolve inorganic salts. In addition, the hydrogen atoms directly bonded to nitrogen add more complexity to PEI structures and give different ionic mobilities. The melting point for a semicrystalline PEI sample is 332K and the glass transition temperature is 250K. [1] The complexity of PEI structure is expected since the NH groups can behave as both proton donors and acceptors to form hydrogen bonds. X-ray diffraction studies of linear PEI showed a different crystal structure than PEO. Chatani *et al.* revealed that the linear PEI chains adopt a double-stranded helical structure. [10] Each strand takes a (5/1) helical form and two strands are connected by inter-chain hydrogen bonds. In each NH group, the N and H atoms, as hydrogen bond acceptor and the donor respectively, each form one H-bond with NH groups from another chain. The inter-atomic distance for $N \cdots N$ pairs is 3.16Å. For each single chain, dihedral angles along the C-N-C-C backbone atom

sequence are 160° and those along the N-C-C-N sequence are 13° . Intra-chain hydrogen bonds do not exist in the crystal structure.

Computer simulation is a good way to understand the structural characteristics of PEO, but prior to our work there were no published MD results for models of PEI or PEI:salt complexes. In this work, MD simulations on models of the amorphous phase of PEO and PEI at 300K are presented. Chain dimensions, dihedral angle distribution, and conformational triads population analysis are compared with each other. Radial distribution function analyses on the PEI model system is done to study the presence and the character of hydrogen bond structures.

2.2. Simulation methodology.

2.2.1. Models.

Low molecular weight oligomeric analogues of PEO have been widely used to model PEO and PEO:salt complexes as addressed in the introduction. Short oligomers can easily adopt a fully amorphous phase at accessible temperatures where ion transport takes place, so the structural analyses can be simplified by studying single-phase systems. Thus in this work, we use a four repeat-unit oligomer ($\text{CH}_3\text{X}(\text{CH}_2\text{CH}_2\text{X})_4\text{CH}_3$, $\text{X}=\text{O}$ in PEO-4 and NH in PEI-4) as the simulation model for the single amorphous phase of PEO and PEI. For convenience, we shall call the tetramer models PEO-4 (it is conventionally called tetraglyme) and PEI-4. We filled PEO-4 and PEI-4 sample boxes with 50 chains and then simulated at 300K, near room temperature. This is below the melting temperature but above the glass transition temperature for both PEO and PEI. This was done to obtain the amorphous phase in each.

2.2.2. Simulation details.

The AMBER 5 computer program [11,12] was used for MD simulations. The initial configurations of PEO-4 and PEI-4 model molecules were generated by taking Cartesian coordinates of a PEO chain with the same length as the model molecules in the known crystal structure. The coordinates of oxygens were also used for nitrogens in PEI-4. The hydrogen in the NH groups was started along one tetrahedral direction

with the correct N-H bond length. Periodic boundary conditions were applied and the SHAKE algorithm was used to constrain bond lengths, while bond angles and torsional angles remained flexible. The Berendsen coupling algorithm was used to maintain a constant temperature of $300\text{K}\pm 30\text{K}$ in the NPT and NVT ensembles [13] by scaling the kinetic energy every 10ps. The particle mesh Ewald (PME) method [14] was used to calculate long-range electrostatic interactions with direct-space cutoff values of 10\AA for PEO-4 and 12\AA for PEI-4. [14] Nonbonded interactions were truncated beyond these ranges. In all cases, the equations of motion were integrated by using the leap-frog version of the Verlet algorithm with a time step of 1 fs. The sample boxes of PEO-4 and PEI-4 were started at a very low density of $0.1\text{g}/\text{cm}^3$ and then simulated at 1 atmosphere for 1 ns in the NPT ensemble. Equilibrium densities for PEO-4 ($1.0\text{g}/\text{cm}^3$) and PEI-4 ($0.93\text{g}/\text{cm}^3$) were calculated by averaging densities over 0.2ns. Calculated densities for PEO-4 and PEI-4 are close to experimental densities of PEO-4 and $\text{H}_2\text{N}(\text{CH}_2\text{CH}_2\text{NH})_3\text{CH}_2\text{CH}_2\text{NH}_2$, which are $1.009\text{g}/\text{cm}^3$ and $0.998\text{g}/\text{cm}^3$, respectively. [15] Further simulations were performed at 300K in the NVT ensemble for 1.5 ns to reach equilibrium. To test for equilibration, total energies were monitored over the last 500ps of the 1.5ns run. The total energies fluctuated less than 1.6% and moving average energies calculated every 10ps increased less than 0.05% over the 500ps for both systems. The cosine torsional angle autocorrelation function [16] was also calculated. For the C-C and C-O bonds in PEO-4, the relaxation times were 25.8ps and 21.4ps. In PEI-4, the relaxation times were 16.5ps and 24.3ps for the C-C and C-N bonds, respectively. After thus determining that our

systems were at equilibrium, 500ps of atomic trajectory data were collected for PEO-4 and PEI-4.

2.2.3. Force field parameters

The AMBER force field model uses a potential energy function composed of five different terms: harmonic bond length stretching, harmonic bond angle bending, sinusoidal dihedral angle twisting, a 6-12 Lennard-Jones interaction, and an electrostatic interaction: [17]

$$E_{\text{total}} = \sum_{\text{bonds}} K_r (r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right] + \sum_{i < j} \frac{q_i q_j}{\epsilon R_{ij}} \quad (2.1)$$

A dielectric constant, ϵ , of unity is used in the electrostatic interaction term for both systems. The force field parameters used for PEO-4 are taken from the AMBER force field for ethers reported in the literature. [17] The force field parameters for the nitrogen atoms in PEI-4 were taken from a solvation free energy calculation on small aliphatic amines. [18] One missing parameter, the C-C-N bond angle bending force constant, is taken from a similar bond type in the database. The dihedral angle torsional parameter for the N-C-C-N atom sequence is obtained by fitting an energy profile from a density functional calculation for dimethylethylenediamine (DMEDA) into the dihedral angle torsion functional form. [19] All parameters used in this study are given in Table 2.1 and 2.2. Note that $A_{ij} = e^*(R_{ij})^{12}$ and $B_{ij} = 2*e^*(R_{ij})^6$ in the van der Waals parameter section.

Table 2.1. Force field parameters for PEO-4. [17] (Scott E. Boesch provided the partial atomic charges. The charges were obtained fitting atom-centered charges to reproduce the electrostatic potential from a B3LYP/6-31G(d) calculation.)

Bond Parameters				
bond	K_r (kcal/molÅ ²)	r_{eq} (Å)		
CT-CT	310.0	1.526		
CT-HC	340.0	1.090		
CT-OS	320.0	1.410		
Angle Parameters				
angle	K_θ (kcal/mol rad ²)	θ_{eq} (deg)		
CT-CT-HC	50.0	109.50		
CT-CT-OS	50.0	109.50		
CT-OS-CT	60.0	109.50		
HC-CT-HC	35.0	109.50		
Torsional Parameters				
torsion	no. of paths	$V_n/2$ (kcal/mol)	γ (deg)	n
X -CT-CT-X	9	1.40	0.0	3.0
X -CT-OS-X	3	1.15	0.0	3.0
CT-CT-OS-CT	1	0.383	0.0	-3.0
CT-CT-OS-CT	1	0.1	180.0	2.0
OS-CT-CT-OS	1	0.144	0.0	-3.0
OS-CT-CT-OS	1	1.00	0.0	2.0
van der Waals Parameters and Partial Charges				
atom type	R^* (Å)	c^* (kcal/mol)	q	
CT	1.9080	0.1094	0.103	
HC	1.4870	0.0157	0.0355	
OS	1.6837	0.1700	-0.348	

Table 2.2. Force field parameters for PEI-4. (Scott E. Boesch provided the partial atomic charges and force field parameters for torsions of the PEI-4 chains. The torsional parameters were obtained by fitting the quantum mechanically determined energy of PEI-4 chains with different dihedral angles to a Fourier series.)

Bond Parameters				
bond	K_r (kcal/molÅ ²)	r_{eq} (Å)		
CT–NT [18]	367.0	1.471		
NT–HN [17]	434.0	1.010		
Angle Parameters				
angle	K_θ (kcal/mol rad ²)	θ_{eq} (deg)		
CT–CT–NT [17]	80.0	111.20		
CT–NT–HN [18]	35.0	109.50		
CT–NT–CT [18]	60.0	109.50		
HC–CT–NT [18]	35.0	109.50		
Torsional Parameters				
torsion	no. of paths	$V_n/2$ (kcal/mol)	γ (deg)	n
X –CT–NT–X [18]	6	1.0	0.0	3.0
NT–CT–CT–NT[19]	1	0.6	0.0	-3.0
NT–CT–CT–NT	1	0.4	0.0	-2.0
NT–CT–CT–NT	1	0.6	180.0	1.0
van der Waals Parameters and Partial Charges				
atom type	R^* (Å)	e^* (kcal/mol)	q	
NT [18]	1.875	0.1700	-0.614	
HN [18]	1.689	0.0157	0.309	

2.3. Results and discussion.

The chain dimensions, dihedral angle distributions, and conformational triad populations were calculated and compared for PEO-4 and PEI-4 at 300K based on their atomic trajectories. Radial distribution functions were used to probe hydrogen bonding structures for the PEI-4 system.

2.3.1. Chain dimensions.

The overall polymer chain dimensions can be described by the mean square radius of gyration, $\langle S^2 \rangle$, and the mean square backbone end-to-end distance, $\langle R^2 \rangle$. The mean square radius of gyration, a weight averaged expression of polymer chain size, is a useful parameter in polymer studies because it can be measured directly by scattering methods. The characteristic ratio, defined as $C_n = \langle R^2 \rangle / nl^2$, where n is the number of backbone bonds and l^2 is the mean-square bond length, is theoretically predicted as 1 for a freely jointed polymer chain model of any length and is used to compare chain dimensions of polymer samples with different chain lengths. [20] A real polymer chain has a C_n value greater than 1 and it reaches a limit as $n \rightarrow \infty$. For poly(oxyethylene) (HO-(CH₂CH₂O)_n-H, POE, an analogue of PEO), the limiting value of C_n is approximately 5.2. [21] Calculation results are listed in Table 2.3 for PEO-4 and PEI-4 at 300K.

The characteristic ratio calculated in our work for PEO-4 is 5.7 ± 0.2 . A small-angle neutron scattering (SANS) study on a nearly monodisperse POE sample gave the characteristic ratio as 5.7-5.5 in the temperature ranging of 347-459K. [9] Another

SANS result for a POE sample with the polydispersity of 1.5 at 353K gave a value of 6.9. [22] A characteristic ratio value for PEO with the molecular weight M_w from 6×10^3 to 1.1×10^7 in aqueous solution was reported to be 5.2 ± 0.1 at 298K. [23] Our result for the monodisperse short chain sample is therefore in good agreement with the monodisperse POE C_n value, while it is larger than POE's limiting value and smaller than the value for polydispersed POE.

Table 2.3. MD results of mean-square radii of gyration, $\langle S^2 \rangle$, and end-to-end distances, $\langle R^2 \rangle$, and characteristic ratio, C_n , for PEO-4 and PEI-4 at 300K.

300K	$\langle S^2 \rangle$	$\langle R^2 \rangle$	C_n
PEO-4	21.3 ± 0.3	143 ± 4	5.7 ± 0.2
PEI-4	17.7 ± 0.2	95 ± 3	3.6 ± 0.1

Smaller values of $\langle S^2 \rangle$ and $\langle R^2 \rangle$ for PEI-4 than for PEO-4 indicate a more compact spatial shape for PEI-4. The more compact shape in PEI-4 implies that there are relatively stronger intra-molecular attractions or inter-molecular repulsions within the PEI-4 sample. Experimental data for the characteristic ratio are lacking, so our MD result for PEI-4 cannot be compared to any experiment. A theoretical calculation of linear poly(ethylenimine), in which each dihedral angle was allowed to sample a small number of discrete torsional states according to the rotational isomeric state (RIS) theory, gave a value of 5.10-6.56 for the characteristic ratio at 300K. [24] Our result of 3.6 ± 0.1 for PEI-4 is smaller than this theoretical prediction for the long chain PEI polymer.

2.3.2. Dihedral angle distribution.

For the backbone atoms of PEO-4 and PEI-4 chains, there are two different types of dihedral angles defined by C-X-C-C and X-C-C-X atom sequences, where X=O in PEO-4 and NH in PEI-4. In order to examine the conformations of PEO-4 and PEI-4 chains, the time-averaged population density distribution of these dihedral angles (θ) were calculated at 300K and the results are plotted in Figure 2.1.

Figure 2.1(a) compares the population density distribution of dihedral angles along the C-X-C-C sequence in PEO-4 and PEI-4. By our definition, the conformations are eclipsed when $\theta=0^\circ$ and are labeled trans (denoted t) if $120^\circ < \theta < 240^\circ$, gauche plus (g^+) if $0^\circ < \theta < 120^\circ$, and gauche minus (g^-) if $240^\circ < \theta < 360^\circ$. Not surprisingly, these two dihedral angle types give a similar distribution because of the similarity of oxygen and NH groups. Both cases show that a trans conformation is predominant, while gauche conformations (g^+ and g^-) give much smaller peaks near 80° and 280° , respectively. Neither g^+ nor g^- is favored. Moreover, PEO-4 shows a larger population of t (90%) while g^+ and g^- (10% total) are less populated compared to PEI-4 (t 84%, g 16%).

Figure 2.1(b) is the population density distribution of dihedral angles along the X-C-C-X atom sequence for PEO-4 and PEI-4. The g^+ and g^- conformations are much more favored than t. This distribution is attributed to “the gauche effect”. It is well accepted that when atoms at the two ends of a dihedral angle are very electronegative, the dihedral angle favors a gauche (g^+ or g^-) over the trans conformation despite the Coulombic repulsion between the two atoms. [25] The gauche effect has been found

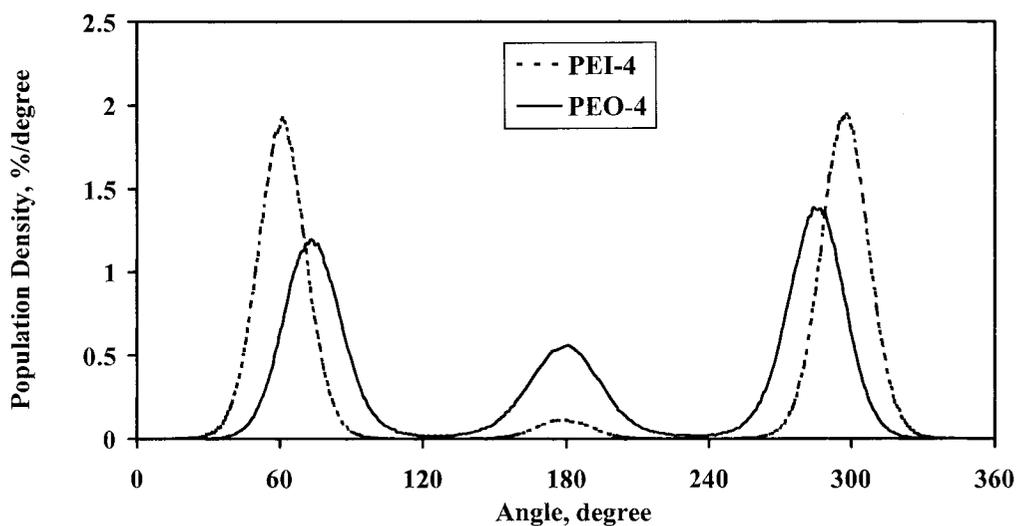
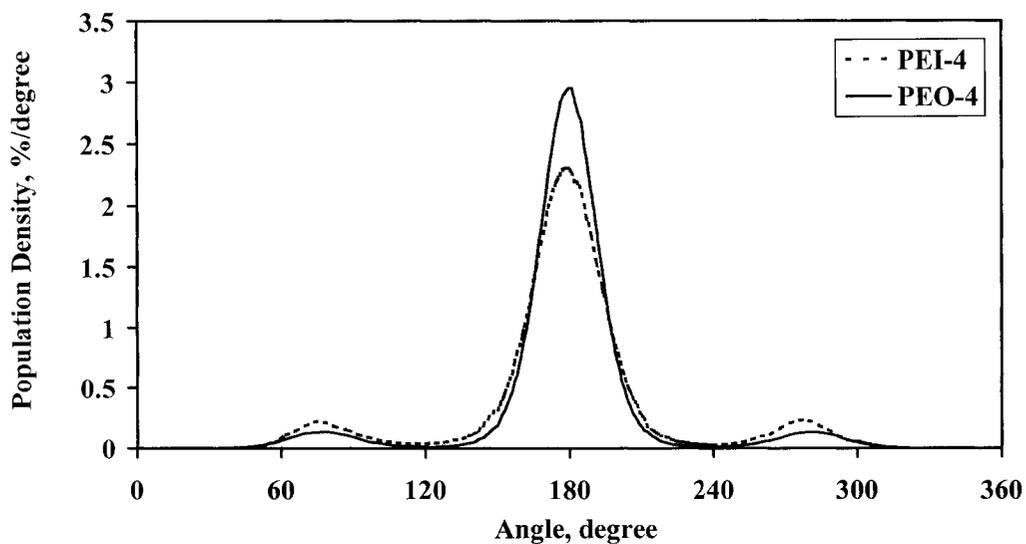


Figure 2.1. (a, upper) Average population density distribution of dihedral angles along the C-X-C-C backbone atom sequence for PEO-4 and PEI-4 at 300 K. (b, lower) Average population density distribution of dihedral angles along the X-C-C-X backbone atom sequence for PEO-4 and PEI-4 at 300 K. For both figures, X=O in PEO-4 and NH in PEI-4.

in high molecular-weight PEO and in lighter oligomeric crystals, melts, and solutions. By comparison, PEI-4 shows a more significant “gauche effect” than PEO-4. In PEI-4, the dihedral angles are almost exclusively in the g^+ and g^- conformations (97% of the total compared to 78% in PEO-4), whereas the t (3%) conformation has a much lower population in PEI-4 than in PEO-4 (22%). The most favored dihedral angles of PEI-4 shift to 60° and 300° , toward the eclipsed conformation, compared to 75° and 285° in PEO-4. These significant differences suggest that there may be a force between the two NH groups other than the usual “gauche effect” that pulls them closer. Hydrogen bonding is one likely possibility that will be described later.

2.3.3. Conformational triads population analysis.

Conformational triads are a combination of three consecutive dihedral angles along the C-X-C-C-X-C atom sequence. The population analysis of triads is one way to examine local structures of polymers. In expressing the geometry of triads, g' indicates a change in sign of the gauche conformation. For example, tgg' corresponds either to the tg^+g^- or tg^-g^+ conformation. The result of the conformational triads analysis at 300K is shown in Figure 2.2. In PEO-4, the tgt conformation has the largest population ($59\pm 3\%$), with smaller, but still significant percentages of ttt ($20\pm 2\%$) and tgg' ($10\pm 1\%$) triads. The result qualitatively agrees with an MD simulation on monoglyme and diglyme in aqueous solutions at 318K, [26] for monoglyme:water and diglyme:water mole fractions of 0.004-0.180 and 0.004 to 0.130, respectively. This study reported 60%-70% tgt populations. In PEI-4, the tgt

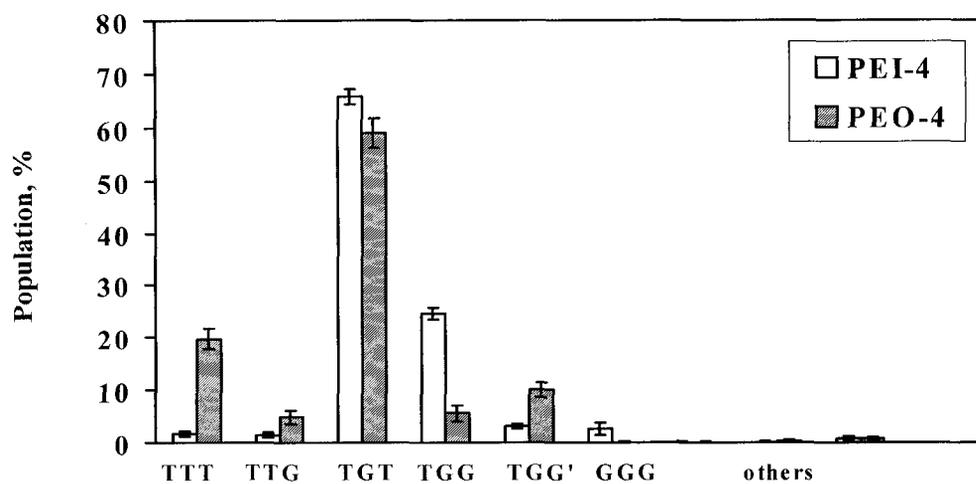


Figure 2.2. Population analysis of conformational triads along the C-X-C-C-X-C backbone atom sequence for PEO-4 and PEI-4 at 300 K. X=O in PEO-4 and NH in PEI-4. The standard deviations are shown as error bars.

conformation is even more favorable ($66\pm 2\%$) than in PEO-4. Among all conformations, the largest differences in populations between PEO-4 and PEI-4 are tgg ($24\pm 1\%$ in PEI-4 and $6\pm 2\%$ in PEO-4) and ttt ($1.6\pm 0.4\%$ in PEI-4 and $20\pm 2\%$ in PEO-4). The different tgt and ttt populations are understandable, since the dihedral angle distribution analysis shows that the central dihedral angle of triads along the N-C-C-N sequence favors the gauche over the trans conformation much more in PEI-4 compared to PEO-4.

2.3.4. Radial distribution function analysis for PEI-4.

Radial distribution function (RDF) calculations provide insights into the PEO-4 and PEI-4 structures by reporting the relative density of atom pairs separated by a distance r :

$$g(r) = \frac{V}{N^2} \left\langle \sum_i \sum_{j \neq i} \delta(r - r_{ij}) \right\rangle \quad (2.2)$$

where i and j refer to the i^{th} and j^{th} molecules and the angle brackets imply averaging over different configurations. [27] These functions give a way to understand interactions between atoms. Smith *et al.* reported the pair radial distribution functions for POE models. [9] Our calculation gives similar results, so they are not addressed here. The PEI system, with its NH group instead of the ether oxygen in PEO, has more interesting structural characteristics. As discussed in the introduction, the NH group is key to forming hydrogen bonding interactions.

We have chosen to focus on three possible ways to form a hydrogen bond between the two NH groups, depending on their relative positions. Adjacent NH groups are separated by one ethyl group, non-adjacent groups are separated by at least two ethyl groups but are in the same molecule, and intermolecular NH groups are from different molecules. All those component RDFs and the total RDF between nitrogen and hydrogen atoms (N...H) and between two hydrogen atoms (H...H) are plotted in Figure 2.3 and 2.4. The hydrogen atoms are only those within NH groups and, in the adjacent RDF calculation, the hydrogen and nitrogen pair that forms a covalent bond is omitted.

The intermolecular RDFs are plotted in Figure 2.3 and 2.4 to examine intermolecular hydrogen bonding. The non-adjacent N...H and H...H RDFs only show very small and indistinct structure in Figure 2.3 and 2.4. This result indicates that minimal non-adjacent hydrogen bonding occurs. In Figure 2.3, the first peak of the intermolecular N...H RDF occurs at 2.3 Å. The intermolecular N...N RDF (not shown) gives 3.3 Å as the most probable distance, as compared to the twice van der Waals distance of 3.75 Å for nitrogen atoms. Since the covalent bond length of N-H is 1.01 Å and the N...H distance is 2.3 Å, the intermolecular hydrogen bonding is linear. Thus, H-bonds with the classical 180° N-H...N angle, N...H distances near 2.3 Å, and N...N distances less than the sum of van der Waals radii are formed between different PEI-4 chains. The populations of a single PEI-4 chain connected with other chains by hydrogen bonding were calculated. The most common case (39.0%) is that of a single PEI-4 chain that hydrogen bonds to only one other chain. Single chains which

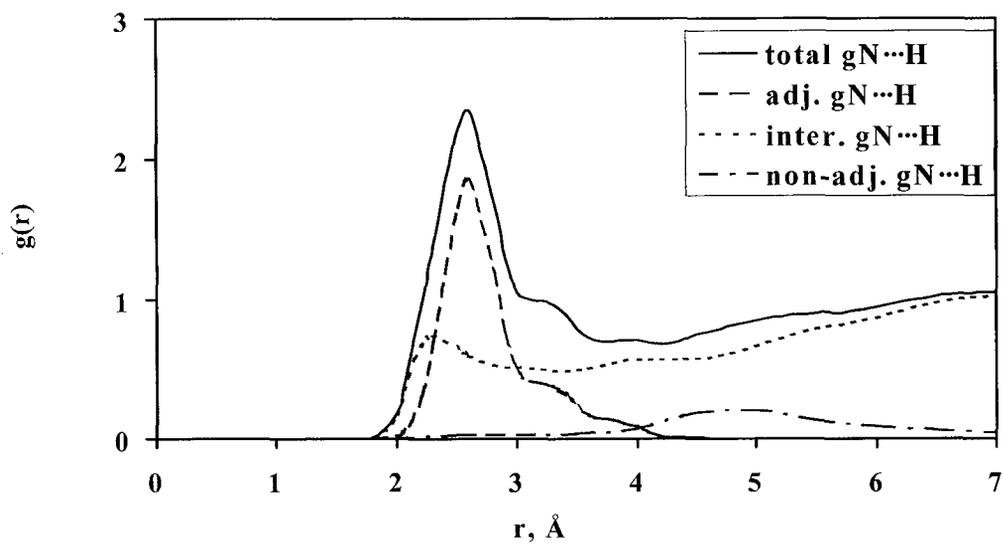


Figure 2.3. The total radial distribution functions (RDFs) of N \cdots H pairs and its components for PEI-4 at 300K. The adjacent (NH groups are separated by one ethyl groups), non-adjacent (NH groups are separated by at least two ethyl groups but are in the same molecule), and intermolecular (NH groups are from different molecules) RDF are represented as dashed, dash dotted, and dotted lines, respectively. Only hydrogen atoms in NH groups are considered.

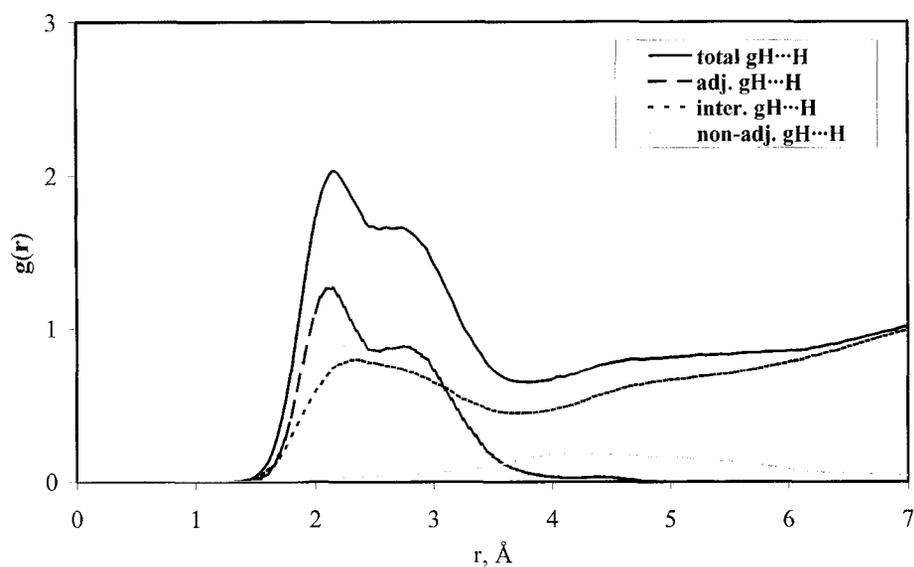


Figure 2.4. The total radial distribution functions of H...H pairs and its components for PEI-4 at 300K. The adjacent, non-adjacent, and intermolecular RDF are represented as dashed, dash dotted, and dotted lines, respectively. Only hydrogen atoms in NH groups are considered. The adjacent RDF clearly shows two peaks at 2.2Å and 2.8Å implying two distinguishable structures. These structures are shown in this figure labeled (A), double H-bonding for the first peak, and (B), single H-bonding structure for the second peak. The first minimum between the first and second peak, at 2.55Å, is set as the cutoff to distinguish the two structures.

hydrogen bond to 2 or 3 chains also have large populations (38.1% and 13.8%). The population for chains having no hydrogen bonds with others is only 7.0%. But this intermolecular N...H RDF does not account for the majority of the total RDF. It implies that intermolecular H-bonding is not the most favorable structure.

We now analyze the extent of H-bonding between NH groups separated by only one ethyl group. In Figure 2.3 the adjacent N...H RDF shows a strong peak at 2.6Å, with a distinct shoulder near 3.3Å. This peak implies an attraction between the nitrogen and hydrogen, but the distance is not short enough to represent a classical hydrogen bond, whose N...H distance is approximately 2.0Å. [28] The peak position in the adjacent N...H RDF is slightly longer than than 2.3Å, the most probable length of an intermolecular H-bond. Because the adjacent NH pairs have to overcome the restraint of bond angles, the N...H distances between adjacent NH pairs cannot get as close as those between different molecules. For convenience, we shall not distinguish this hydrogen-bond-like interaction between adjacent NH groups from classical hydrogen bonding. It can explain the dihedral angle distribution of PEI-4, as it stabilizes the gauche conformations for dihedral angles along N-C-C-N. It also causes the shift of these dihedral angles toward the eclipsed geometry, relative to PEO-4, because the two NH groups can get closer to each other. The presence of the shoulder (3.3Å) is an interesting feature and its origin is described later.

Figure 2.4 shows H...H RDFs for PEI-4. As in Figure 2.3, the total RDF has the same shape as the RDF for hydrogens in adjacent NH groups in the distance range below 4Å. There are two peaks that occur at 2.2Å and 2.8Å, separated by a shallow

minimum at 2.55Å. This means there are two distinguishable (NH)-C-C-(NH) structures. Since 97% of the dihedral angles along this atom sequence are in a gauche conformation, there are two distinguishable gauche conformations. (They are not g^+ and g^- because g^+ and g^- are statistically indistinguishable.) A number of local structures have been examined for adjacent H...H distances of 2.2Å and 2.8Å to identify typical structures. Two distinguishable gauche structures were found. One is a double hydrogen bonding structure in which both N-H bonds are oriented toward the adjacent nitrogens. (See Figure 2.5(A)) The structure is symmetrical and the distance between the two hydrogen atoms is near 2.2Å, so this corresponds to the first peak in the adjacent H...H RDF. Conformational constraints prevent the two N-H bonds from pointing directly at the adjacent nitrogen to form a 180° N-H...N angle. In the other structure, one of the hydrogen atoms points toward the adjacent nitrogen and one points away from the adjacent nitrogen (Figure 2.5(B)). The distance between the two hydrogen atoms on adjacent nitrogen atoms is 2.8Å, which gives the second peak in the adjacent H...H RDF. Once again, the N-H...N angle of this structure cannot be 180°. A third gauche structure without H-bonds is also possible with both H atoms pointing away from the adjacent N atoms. In this case, the H...H distance would be longer than that in either the singly or the doubly hydrogen bonded structures. The absence of any peaks in the adjacent H...H RDF at longer distances than 2.8Å suggests that the *tgt* conformation without H-bonding is unfavorable.

Referring back to the shoulder of the adjacent N...H RDF at 3.3Å in Figure 2.3, the double hydrogen bonding structure, structure (A), cannot cause this feature since

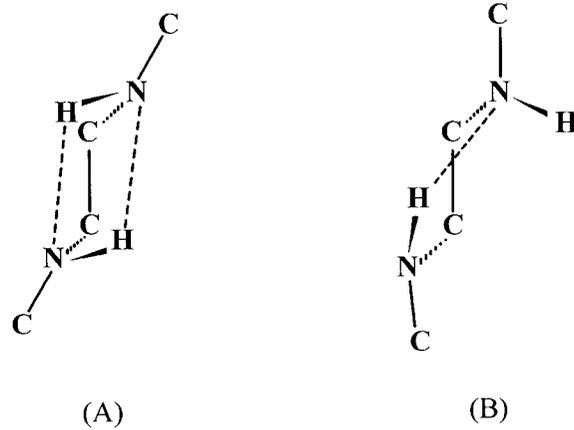


Figure 2.5. Diagrams of two distinguishable Hydrogen-bonding structures in C-(NH)-C-C-(NH)-C units in PEI-4. (A) Double H-bond and (B) Single H-bond. Carbon hydrogens are omitted.

its N...H distance is too short. The single hydrogen bonding structure, structure (B), could cause this shoulder because there is one non-H-bonded N...H pair in this structure and this explanation is verified by Figure 2.6. Figure 2.6 shows the contributions from the structure in Figure 2.5(A) and structure in Figure 2.5(B) to the adjacent N...H RDF. The cutoff value for distinguishing the two structures is set at 2.6 Å, the minimum between the first and second peak in the adjacent H...H RDF. If the H...H distance is less than 2.6 Å, the structure is considered to adopt structure (A), and the N...H RDF is calculated to give its contribution to the total adjacent N...H RDF. The same procedure is repeated for structure (B) where the H...H distance is greater than 2.55 Å. In Figure 2.6, the different shapes of the two structural contributions to the RDF are clearly shown. The single peak in the contribution from structure (A) implies that the two hydrogen bonds are formed approximately within the same distance, 2.6 Å. Thus, the structure is symmetrical. The split contribution from structure (B) shows its asymmetrical character in the relative positions of N and H. The first peak at 2.6 Å is the length of the H-bond and the second peak at 3.3 Å is the distance between the non-H-bonded N...H pair. Thus, based on analyses of the adjacent N...H and H...H RDFs, we conclude that there are two distinguishable H-bonding structures responsible for the predominance of the gauche conformation of the (NH)-C-C-(NH) angle, the double and the single hydrogen bonding structures. An integration of the adjacent H...H RDF gives relative populations of the two structures as 37% double H-bond and 63% single H-bond structures.

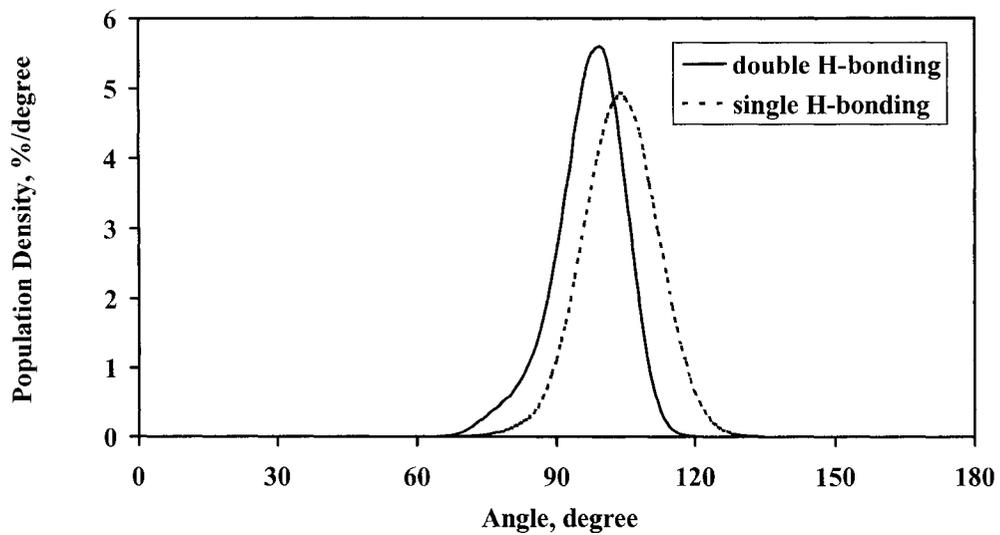


Figure 2.6. The H-bond angle distribution for the double H-bonding and the single H-bonding structures.

Hydrogen-bonding angle distributions for the structures (A) and (B) are calculated and plotted in Figure 2.7. Each structure simply gives one peak at 99° for structure (A) and 104° for (B). The values indicate that the adjacent H-bond is far from linear, probably because of conformational constraints of the chain.

The mean lifetime of the adjacent H-bonding was also calculated to ensure that our results are statistically meaningful. In Figure 2.6, the first minimum of the adjacent N \cdots H RDF occurs at 3.0\AA and any N \cdots H pairs within this distance will be considered to form H-bonds. The H-bonds will be considered to break when the N \cdots H distances become larger than 3.0\AA . The calculation gives a time much shorter than the 500 ps length of the trajectory, 11.4ps, as the mean H-bond lifetime.

The MD results demonstrating hydrogen bonding structures between adjacent NH groups is consistent with an *ab initio* study of dimethylethylenediamine ($\text{CH}_3\text{-NH-CH}_2\text{-CH}_2\text{-NH-CH}_3$, DMEDA). [19] This work implies that the tgt structure is the most stable and also gives the single H-bonding structure as the lowest energy structure and the double H-bonding structure as the second lowest energy from among 8 different structures calculated. This agrees qualitatively with our populations of the two structures (63% for single and 37% for double). The *ab initio* double H-bonding structure gives N \cdots H distances of 2.59\AA , H \cdots H distances of 2.20\AA , and the N-H \cdots N angles of 95.0° close to our MD results of 2.6\AA , 2.2\AA , and 99° , respectively. The *ab initio* single H-bonding structure gives N \cdots H distances of 2.45\AA and 3.31\AA , H \cdots H distances of 2.85\AA , and N-H \cdots N angles of 107.4° . Once again, our MD results

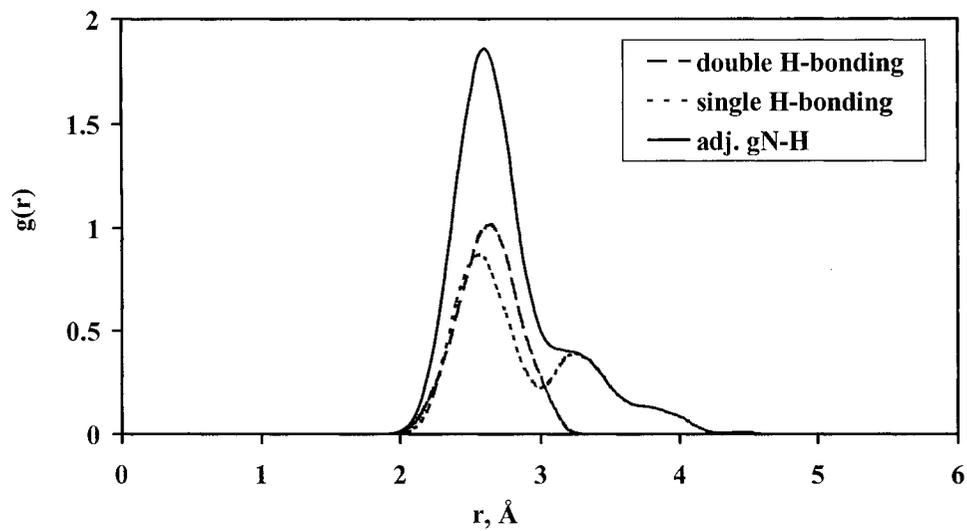


Figure 2.7. The contributions from the double H-bonding and the single H-bonding structures to the adjacent N...H RDF.

of 2.6Å, 3.3Å, 2.8Å, and 104° agree closely with values from *ab initio* quantum chemistry.

2.3.5. Helical structure analysis of PEO-4 and PEI-4 chains.

The dihedral angle distribution and the conformational triads population analysis of PEO-4 and PEI-4 provide information regarding their local structures. Recall that PEO and PEI crystal structure both show long-range conformational order: PEO chains adopt a (7/2) single helical, [5] while PEI chains adopt a (5/1) double-stranded helical structure. [10] Inter-chain H-bonding stabilizes the tightly formed double helix of crystalline PEI. In the amorphous state, it is not known how much of the helical structure may be retained. Since crystalline PEO executes 2 helical turns for every 7 repeat units, four monomeric units plus end-capping methyl and methoxy groups are sufficient for PEO-4 to form one helical turn. Likewise crystalline PEI executes one helical turn every 5 monomer units but adopts a double helical structure. Since our calculations imply that PEI-4 is more compact than PEO-4, four monomeric units plus the end-caps should be sufficient to form locally helical structures for PEI-4. Our results show, however, that only small parts of PEO-4 (5.6%) and PEI-4 (8.2%) adopt helical structures, with left- and right-handed helices being equally populated. For PEI-4, calculations indicate that the single H-bonding is favored over the double H-bonding structure within helices.

The intermolecular H-bonding analysis can indicate the extent of double-stranded helical structure in PEI-4. For our four repeat unit PEI model, a double helix

extending the entire length of one chain requires 8 H-bonds exclusively with one other chain. Our result implies that only 2.0% of the total PEI-4 chains adopt this structure. So very little double stranded helical structure exists in the amorphous state of PEI-4.

2.4. Conclusions.

Molecular dynamics simulations of four repeat unit models of poly(ethylene oxide) (PEO) and poly(ethylenimine) (PEI), denoted PEO-4 and PEI-4, were performed at 300K. The chain dimensions, dihedral angle distributions, and conformational triad populations were calculated and compared with available experiments to show the similarities and differences between equilibrium amorphous structures of PEO-4 and PEI-4.

The chain dimension calculations gave values of the mean-square radius of gyration, $\langle S^2 \rangle$, and mean-square end-to-end distance, $\langle R^2 \rangle$, both measures of spatial extent of a polymer, for the two systems. The $\langle S^2 \rangle$ value of 17.7 ± 0.2 and $\langle R^2 \rangle$ value of 95 ± 3 for PEI-4 are smaller than the corresponding values for PEO-4, 21.3 ± 0.3 and 143 ± 4 , and indicate a more compact form for amorphous PEI-4 at 300K. The characteristic ratio is 5.7 ± 0.2 for PEO-4 and 3.6 ± 0.1 for PEI-4 and are comparable to results from small-angle neutron scattering experiments, [9,22] other MD simulations, [9] and theoretical predictions [24] for similar systems related to PEO-4.

The dihedral angle distribution analysis reveals that PEO-4 and PEI-4 have similar distributions for the two types of dihedral angles, C-X-C-C and X-C-C-X (X=O in PEO-4 and X=NH in PEI-4). The C-X-C-C dihedral angles favor a trans conformation (t) while the gauche conformation (g^+ and g^-) is predominant for the X-C-C-X dihedral angles. For the C-X-C-C dihedral angles, PEO-4 has 90% T conformations and 10% g^+ and g^- compared to 84% t and 16% g^+ and g^- conformations for PEI-4. For the X-C-C-X dihedral angles, PEI-4 almost exclusively

favors g^+ or g^- conformations (97% population). By comparison, 78% of the X-C-C-X dihedral angles in PEO-4 adopt g^+ and g^- conformations. The most probable angles for g^+ or g^- conformations are also different. In PEO-4, they are 75° and 285° while in PEI-4 they shift to 60° and 300° . The large populations of g^+ or g^- and the shift of the most probable dihedral angles toward a more eclipsed conformation suggest a stronger attraction between the adjacent NH groups in PEI-4.

A triad population analysis implies that the *tgt* conformation is the most stable local structure for PEO-4 and PEI-4 because of its large population ($59\pm 3\%$ in PEO-4 and $66\pm 2\%$ in PEI-4). PEO-4 shows a secondary preference for *ttt* conformations ($20\pm 2\%$) and PEI-4 has secondary preference for *tgg* conformations ($24\pm 1\%$). The *ttt* conformer is negligibly populated ($1.6\pm 0.4\%$) in PEI-4. It is easy to understand the decrease of the *ttt* population in PEI-4 because only 3% of N-C-C-N dihedral angles adopt the *t* conformation.

The presence of NH groups in PEI-4 brings more complexities to its structure. One of the most important features is the possibility of hydrogen bond formation, since the NH groups can behave as both proton donors and proton acceptors. Total radial distribution functions (RDFs) and their component adjacent, non-adjacent, and intermolecular RDFs were calculated for N...H, H...H, and N...N pairs of PEI-4. Two distinguishable hydrogen bonding structures between adjacent NH groups have been found. One is the double H-bonding structure in which each member of an NH pair acts as an H-bond donor and acceptor to the other adjacent NH of the pair (Figure 2.5(A)). Another is the single H-bonding structure in which only one NH group forms

an H-bond with an adjacent N (Figure 2.5(B)). In both cases the H-bonds are long (N...H distances are 2.6Å for both), and non-linear (N-H...N angles are 99° for the double H-bond and 104° for the single). The single H-bonding structure is more prevalent (63%) than the double (37%) and is apparently the most stable local structure for the (NH)-C-C-(NH) unit.

Intermolecular hydrogen bonding was examined by analyzing the intermolecular N...H RDF for PEI-4. Its small contribution to the total N...H RDF suggests that intermolecular H-bonding is not as common as H-bonding between adjacent NH groups in the same chain, but the shorter N...H distance implies that this interaction is stronger than the adjacent H-bonding. The RDF results also show that the intermolecular H-bonding structure is nearly linear.

Only a small percentage of PEO-4 and PEI-4 adopt a helical structure (5.6% for PEO-4 and 8.2% for PEI-4). This means that PEO-4 and PEI-4 chains are much less ordered than in the polymer crystalline phases, where PEO is helical and PEI adopts a double helical structure. A population analysis of PEI-4 chains displaying intermolecular H-bonds was done to test the existence of double-stranded helices. The result indicates that this structure exists to an insignificant degree in the amorphous state of PEI-4 at 300K.

2.5. Acknowledgments.

We are grateful to the National Science Foundation/National Resource Allocations Committee for a supercomputer time award at the University of Illinois (award number MCA96-N019 to R.W.) and to the NSF for financial support for Curtis Durham (NSF-REU# CHE-9820544). We thank Prof. Daniel T. Glatzhofer for administering the REU/SURF program, Scott E. Boesch for providing force field parameters for torsions of the PEI-4 chain and for sharing results of his quantum chemical calculations, and an anonymous reviewer for clarifying several issues concerning characteristic ratios.

2.6. References.

- [1] Tanaka, R.; Fujita, T.; Nishibayashi, H.; Saito, S. *Solid State Ionics* **1993**, *60*, 119.
- [2] Berthier, C.; Gorecki, W.; Minier, M.; Armand, M. B.; Chabagno, J. M.; Rigaud, P. *Solid State Ionics* **1983**, *11*, 91.
- [3] Armand, M. B.; Chabagno, J. M.; Duclot, M. *Second International Meeting on Solid Electrolytes. St Andrews, Scotland* **1978**, Extended Abstract.
- [4] MacCallum, J. R.; Vincent, C. A. *Polymer Electrolyte Reviews*; Elsevier: New York, 1987; Vol. vol. 1.
- [5] Takahashi, Y.; Tadokoro, H. *Macromolecules* **1973**, 672.
- [6] Neyertz, S.; Brown, D.; Thomas, J. O. *J Chem Phys* **1994**, *101*, 10064.
- [7] Neyertz, S.; Brown, D. *J Chem Phys* **1995**, *102*, 9275.
- [8] Lin, B.; Boinske, P. T.; Halley, J. W. *J Chem Phys* **1996**, *105*, 1668.
- [9] Smith, G. D.; Yoon, D. Y.; Jaffe, R. L.; Colby, R. H.; Krishnamoorti, R.; Fetters, L. J. *Macromolecules* **1996**, *29*, 3462.
- [10] Chatani, Y.; Kobatake, T.; Tadokoro, H. *Macromolecules* **1982**, *15*, 170.
- [11] Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 5*; University of California-San Francisco: San Francisco, 1997.
- [12] Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham III, T. E.; DeBolt, S.; Ferguson, D. M.; Seibel, G. L.; Kollman, P. A. *Comp. Phys. Comm.* **1995**, *91*, 1.
- [13] Berendsen, H. T. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J Chem Phys* **1984**, *81*, 3684.
- [14] Darden, T.; York, D.; Pedersen, L. *J Chem Phys* **1993**, *98*, 10089.
- [15] *Aldrich Handbook*, 2000-2001.

- [16] McQuarrie, D. A. *Statistical Mechanics*; Harper & Row: New York, 1976.
- [17] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J Am Chem Soc* **1995**, *117*, 5179.
- [18] Morgantini, P.; Kollman, P. A. *J Am Chem Soc* **1995**, *117*, 6057.
- [19] Boesch, S. E.; York, S. S.; Frech, R.; Wheeler, R. A. *J Chem Soc, Phys Chem Comm* **2001**, *1*, (electronic publication).
- [20] Flory, P. J. *The Statistical Mechanics of Chain Molecules*; Hanser: New York, 1988.
- [21] Abe, A.; Furuya, H.; Mitra, M. K.; Hiejima, T. *Comput Theor Polym Sci* **1998**, *8*, 253.
- [22] Kugler, J.; Fischer, E. W.; Peuscher, M.; Eisenbach, C. D. *Makromol Chem* **1983**, *184*, 2325.
- [23] Kawaguchi, S.; Imai, G.; Suzuki, J.; Miyahara, A.; Kitano, T.; Ito, K. *Polymer* **1996**, *38*, 2885.
- [24] Wang, S.; DeBolt, L.; Mark, J. E. *Polym Prepr* **1993**, *34*, 478.
- [25] Abe, A.; Mark, J. E. *J Am Chem Soc* **1976**, *98*, 6468.
- [26] Bedrov, D.; Borodin, O.; Smith, G. D. *J Phys Chem B* **1998**, *102*, 5683.
- [27] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987.
- [28] Schuster, P.; Zundel, G.; Sandorfy, C. *The Hydrogen Bond: Recent Developments in Theory and Experiments*; North-Holland Publishing Company: Amsterdam, 1976; Vol. Vol. 2.

Chapter 3

Molecular dynamics simulations and vibrational spectroscopic studies of local structure in amorphous $\text{CH}_3\text{O}(\text{CH}_2\text{CH}_2\text{O})_4\text{CH}_3:\text{XCF}_3\text{SO}_3$ ($\text{X}=\text{Li}, \text{Na}$) systems. ^a

3.1. Introduction.

Solid polymer electrolytes (SPE), usually comprised of alkali salts dissolved in a polymer matrix, are subject to intense study because they are crucial components of modern electronic devices, including high energy-density batteries, fuel cells, and electrochromic displays. [1-5] For instance, SPE have been suggested to replace the liquid electrolyte in lithium-ion rechargeable batteries because SPE can provide variable geometries, large active areas, flexibility of operation, and improved safety. In lithium batteries, SPE serve as both the intermediate between the anode and cathode and as a medium for transporting lithium ions in the charging/discharging cycle of the cell.

Because of its ability to solvate a variety of inorganic salts leading to polymer electrolytes with significant values of ionic conductivity, poly(ethylene oxide) or “PEO” is one of the most widely studied host polymer matrix for potential use as an electrolyte in high energy-density batteries. [4-6] It is now well known that ion transport occurs primarily through amorphous regions of PEO:salt complexes. [7,8]

^a This chapter is reproduced with permission from: Hyun, J-K.; Dong, H.; Rhodes, C. P.; Frech, R.; Wheeler, R. A. *J. Phys. Chem. B* **2001**, *105*, 3329-3337 and Dong, H.; Hyun, J-K.; Rhodes, C. P.; Frech, R.; Wheeler, R. A. *J. Phys. Chem. B* **2002**, *106*, 4878-4885. © 2001 2002 American Chemical Society

Ionic conductivity in PEO-based SPE, however, is low at room temperature, because large amounts of crystalline phases are present. [9] For instance, for PEO:LiX (X=ClO₄, CF₃SO₃), ionic conductivity is in the range of 10⁻³ - 10⁻⁴ S cm⁻¹ at 100 °C, while it falls to the range of 10⁻⁶ - 10⁻⁸ S cm⁻¹ at room temperature. [10] This limits the practical operating temperatures to the 80-140 °C range. Thus, developing SPE with high conductivity along with acceptable mechanical properties at room temperature is highly desirable. Molecular-level understanding of ion transport in polymer:salt complexes is required to engineer systems with higher ionic conductivity at lower temperatures for improved device performance.

Although ionic motion through molecular liquids occurs through simple diffusion, ion-polymer interactions, polymer segmental motions, and ionic association complicate ion transport mechanisms in PEO-based SPE. Understanding each of these aspects and their interrelations in amorphous PEO is necessary to engineer PEO-based SPE with higher ionic conductivity. Numerous experimental and computational studies on these aspects have been performed. [11-17] Here I will focus the discussions on models for PEO-based SPE.

X-ray diffraction has been used to determine the structure of crystalline PEO and PEO:salt complexes. The crystal structure of pure PEO, [18] for example, shows that individual PEO molecules adopt a 7/2 helical structure, where 7 monomeric units make 2 turns. Structures of PEO complexed with a variety of salts such as NaClO₄, NaSCN, HgCl₂, LiCF₃SO₃, NH₄SCN, and KSCN have been determined by X-ray diffraction. [19-22] These studies show that the cations are coordinated to both the

anions and the oxygens of PEO and the overall coordination number is dependent on the cation and the anion. It is also concluded from these studies that the conformation of PEO chains may be altered upon complexation with metal salts.

Vibrational spectroscopy has been employed to study such conformational changes in crystalline and amorphous systems. Conformational changes of PEO upon complexation with metal cations have been observed in PEO:NaX (X=Br, I, SCN, BF₄, and CF₃SO₃) and PEO:MSCN (M=K, Rb, and Cs) complexes by vibrational spectroscopy. [23,24] Furthermore, salt concentration and chain length effects on the conformational changes for oligomer:salt complexes such as glyme:MCF₃SO₃ have been examined using vibrational spectroscopy. [11,25,26] (M=Li, Na, and K; glyme is defined as CH₃O(CH₂CH₂O)_nCH₃. For example, the n = 1 oligomer is described as monoglyme, n = 2 as diglyme, n = 3 as triglyme, and n = 4 as tetraglyme or tetraethylene glycol dimethyl ether.)

Ionic association in PEO:salt complexes and low molecular weight oligomer:salt systems has been investigated via vibrational spectroscopy [27-33] and X-ray absorption fine structure. [34-36] Ion association has been observed in PEO:MCF₃SO₃ and glyme:MCF₃SO₃ (M=Li and K) solutions using vibrational spectroscopy. Different types of ionic aggregates in these complexes—including free ions, ion pairs, and larger aggregates—have been identified by monitoring changes in the SO₃ symmetric stretching mode and the CF₃ symmetric deformation mode of the triflate ion (triflate is trifluoromethanesulfonate, CF₃SO₃⁻, sometimes abbreviated as Tf) as temperature and salt concentrations vary. [11,27,28] Although vibrational

spectroscopy can provide excellent ways to identify the species of ionic aggregates, their sizes and structures can only be inferred indirectly. Conductance measurements for a variety of PEO:salt and oligomer:salt complexes also show that ionic conductivity increases with increasing salt concentration up to a certain level (dependent upon the species of salt) and then decreases at high salt concentrations, [10,37] where ionic mobility is presumably reduced by the lower mobility of larger aggregate species.

Published molecular dynamics (MD) and Monte Carlo (MC) simulations show the power of each method for constructing structural models for crystalline and amorphous PEO and PEO:salt complexes. [38-42] Several MD simulations of models of PEO:MX (MX=NaI, LiI, LiBr, ZnBr₂) have been performed to examine ionic association and conductivity. [43-49] Oligomer chain lengths in these studies vary from 1 to 142 monomeric CH₂CH₂O units. Different types of ionic aggregates were identified and the dependence of ionic association on concentration was studied. Aggregates ranging in number of ions from 1 to 7 were detected and their sizes increase as MX concentration increases. [44] Diffusion coefficients were calculated for PEO:NaI complexes and a schematic picture of diffusion has been suggested by Neyertz and Brown. [40,48] Muller-Plathe *et al.* [44] have also calculated diffusion coefficients and conductivity for PEO:LiI and have produced the same behavior of conductivity versus salt concentration as that from experiment at 400 K. Also, some features of ion diffusion for metal salts in very short oligomers have been examined by MD simulations. Payne *et al.*, [45-47] for example, have calculated diffusion

coefficients and compared many features of ion diffusion for NaI in dimethyl ether and in monoglyme.

Low molecular weight oligomers analogous to PEO have also been widely used to model PEO:salt complexes in experimental studies. Oligomers more easily adopt a fully amorphous state at accessible temperatures, so single phase systems can be more easily prepared and analyzed. Among those, glymes have been extensively used in spectroscopic studies by two of the authors. [11,50] LiCF_3SO_3 is one of the most widely used salts, in view of its applicability in practical PEO-based electrochemical cells. An additional advantage to CF_3SO_3^- is the ability to identify distinct aggregate species present in a system by Raman and infrared spectroscopy. [51]

MD simulation studies of tetraglyme: MCF_3SO_3 for comparison with experimental data are currently lacking, so MD simulations of amorphous tetraglyme: LiCF_3SO_3 and tetraglyme: NaCF_3SO_3 are reported here. Advantages of studying these systems include (1) tetraglyme and tetraglyme:salt complexes are more easily equilibrated, experimentally and computationally, than longer chain oligomers and (2) the LiCF_3SO_3 salt is actually used in some commercial lithium batteries. Moreover, halide salts previously studied cannot capture the monodentate, bidentate, or tridentate coordinate of Li^+ and Na^+ offered by CF_3SO_3^- . Dynamical properties will also be the subject of future work. This contribution thus emphasizes the temperature dependence of conformational changes upon metal cation complexation, Li^+ and Na^+ coordination by CF_3SO_3^- , and their relation to the extent of ionic aggregation.

3.2. Simulation methodology.

The models of tetraglyme:LiTf and tetraglyme:NaTf contains 50 tetraglyme chains, 25 Li^+ or Na^+ , and 25 Tf^- ions to satisfy the ether oxygen:salt ratio of 10:1. The initial tetraglyme chain configuration was generated from the known crystal structure for PEO. [18] The Li^+ , Na^+ , and Tf^- ions were randomly distributed in the tetraglyme matrix. All parameters of the inter- and intra-molecular interactions of tetraglyme and Li^+ and Na^+ ion were taken from the AMBER all-atom force field [52] while their partial charges were obtained from the literature. [53] The Lennard-Jones parameters for triflate ion were also from the AMBER all-atom force field. Partial atomic charges for the triflate ion were derived by a standard procedure. First, hybrid Hartree-Fock/density functional quantum chemical calculations [54] were performed and then the program CHELPG [55] was used to perform a least-squares fit to find charges that best reproduce the electrostatic potential on a grid of points. The grid was chosen to include at least 13,000 points spaced 0.3 Å apart and located outside the van der Waals radius of each atom. This procedure has been used in all the presented research to determine atomic charges for p-benzoquinones and their semiquinone anions and found to give one-electron reduction potentials within approximately 0.1 eV of their experimental values. [56-58] The force field parameters for ions are enclosed in Table 3.1 and the parameters for tetraglyme can be found in Table 2.A1 in Chapter 2.

Before the model boxes of tetraglyme:salt complexes were made, the equilibrium state of the model box of pure tetraglyme was verified. To do this, first the box was

kept at 300 K and 1 atmosphere for approximately 1 ns in the NPT ensemble, where density close to the experimental density for tetraglyme (1 g/cm^3) was obtained. Then total energies and conformational changes were monitored over the last 500 ps of a sequential 2 ns NVT simulation. Moving average total energies calculated every 10 ps increased 1.0% at 300 K and 0.2% at 400 K over this 500 ps period. Conformational changes (e.g. gauche to trans) occurred every 14.1 ps, on average, around C-C bonds and 17.7 ps around C-O bonds at 300 K and at 400 K, conformations changes every 4.3 ps around C-C bonds and 3.9 ps around C-O bonds. The cosine dihedral angle autocorrelation functions around C-C and C-O bonds were calculated from the molecular trajectories and the relaxation times were 15.9 ps and 16.4 ps at 300 K and 2.8 ps and 2.5 ps at 400 K. These time scales were short compared to the simulation time and therefore the equilibrium state could be verified.

Molecular dynamics simulations of the model tetraglyme:LiTf and tetraglyme:NaTf were performed at 300 and 400K using the AMBER 5.0 MD program. [59] Temperature was held constant in the NPT and NVT ensembles by the Berendsen coupling algorithm. [60] Bond lengths were constrained by SHAKE algorithm, while valence angles and dihedral angles remained flexible. Periodic boundary conditions were applied and the particle mesh Ewald (PME) method was used to evaluate the long-range electrostatic interactions. [61,62] A direct-space cutoff of 12 \AA was used for Lennard-Jones and the electrostatic potentials for both temperatures. The equations of motion were integrated with a time step of 1fs. The models were kept at 300 and 400K and 1.0 atm for 1ns in the NPT ensemble

calculation. Stable densities were obtained, yielding 1.13 g/cm³ and 1.04 g/cm³ at 300 K and 400 K for the lithium system and 1.13 g/cm³ at 300K and 1.01 g/cm³ at 400 K for the sodium system, respectively. Further equilibration was continued in the NVT ensemble for another 2 ns and atomic trajectories were collected for the last 500 ps, for analysis.

Table 3.1. The force field parameters for Li⁺, Na⁺, and Tf⁻ ions used in AMBER. [52-54] (Scott E. Boesch provided the partial atomic charges and force field parameters for dihedral angle of the triflate ion. For the procedure of obtaining these parameters please see Tables 2.1. and 2.2. on pages 21 and 22.)

Bond parameters				
Bond	K_r (kcal/mol Å ²)	r_{eq} (Å)		
C-S	187.56	1.868		
C-F	423.48	1.352		
S-O	621.77	1.481		
Angle parameters				
Angle	K_θ (kcal/mol rad ²)	θ_{eq} (deg)		
F-C-S	50.39	111.92		
O-S-C	81.38	115.52		
F-C-F	50.03	106.91		
O-S-O	49.65	102.39		
Dihedral angle parameters				
Dihedral angle	No. of paths	$V_n/2$ (kcal/mol)	γ (deg)	n
F-C-S-O	9	14.41	0.00	3.00
van der Waals parameters and charges				
Atom type	R^* (Å)	ϵ^* (kcal/mol)	q	
O	1.6612	0.2100	-0.573	
C	1.9080	0.1094	0.282	
F	1.7500	0.0610	-0.152	
S	2.0000	0.2500	0.895	
Li ⁺	1.1370	0.0183	1.000	
Na ⁺	1.8680	0.00277	1.000	

3.3. Results and discussion (I) – the tetraglyme:LiCF₃SO₃ complex.

A series of MD simulations of amorphous tetraglyme and tetraglyme:LiCF₃SO₃ (at ether oxygen:Li=10:1 composition) were performed at 300 K and 400 K. The changes in the local structures and conformations of tetraglyme chains upon Li⁺ complexation with ether oxygen and oxygen of CF₃SO₃⁻ were examined and ionic aggregation was analyzed from MD results. Structural inferences from MD calculations were also compared with vibrational spectroscopic experiments.

3.3.1. Chain dimensions.

As discussed in the introduction, local structure and conformation of PEO may be altered by metal cation complexation, so the dimensions of polymer chains may consequently be changed. The configuration of polymer chains can be characterized by the mean-square radius of gyration $\langle S^2 \rangle$, a measure of the size of the spatial domain of polymer chains, the mean-square backbone end-to-end distance $\langle R^2 \rangle$, a measure of the spatial distribution of polymer chains, and the characteristic ratio C_n , which can be used to compare polymer chains with different sizes. The results for the pure tetraglyme and tetraglyme:salt complexes are presented in Section 3.5.1. where a thorough discussion can also be found on the chain dimensions and their changes by cation coordination and temperatures.

3.3.2. Conformation of pure tetraglyme and its changes upon complexation.

In polyether chains, torsions along C-C bonds (an O-C-C-O atom sequence) prefer gauche states over trans states. This is referred to as the oxygen gauche effect. [63,64] The gauche effect has also been found in PEO and low weight oligomer crystals, melts, and in solution. It is believed that a pendant oxygen atom acts to lower the energy of the gauche conformation despite the Coulombic repulsion between oxygen atoms in ethers and polyethers and therefore makes the gauche state the preferred conformation over the trans state. [63,64] Conformational characteristics of ethers and polyethers including short-chain analogues of PEO such as 1,2-dimethoxyethane (DME or monoglyme) have been a subject of several NMR vicinal coupling experiments and computational studies. [65-68]

To examine the conformations of tetraglyme chains, the time-averaged population density distributions of backbone dihedral angles (θ) around C-O bonds (a C-C-O-C atom sequence) and C-C bonds were calculated from MD simulation results at T=400 K. Comparisons of dihedral angles for tetraglyme and tetraglyme:LiTf are shown in Figures 3.1 and 3.2. To generate Figures 3.1 and 3.2, we divided the tetraglyme chain into two parts, an “inside” region and an “outside” region. The inside region of a chain includes all atoms within rings defined by adjacent oxygen atoms coordinating a Li^+ cation. The outside region is the rest of the chain. For example, if two adjacent oxygen atoms coordinate a Li^+ , the inside region includes only one O-C-C-O subunit, if three adjacent oxygens coordinate a Li^+ , the inside region includes an O-C-C-O-C-O unit. If all five oxygen atoms of a tetraglyme chain complex a single Li^+ , the

whole chain would define an inside region. On the other hand, if only one oxygen atom of a tetraglyme chain complexes with a Li^+ , there should be only an outside region. For pure tetraglyme, Figure 3.1(a) shows that a gauche conformation (g^+ and g^-) along C-C bonds is predominant with a lesser, but still substantial fraction of chains adopting the trans conformation (labeled t). Our convention is to define the conformation with heavy atoms eclipsed as $\theta=0^\circ$ and label conformations g^+ if the torsional angle θ has a value $0^\circ < \theta < 120^\circ$, g^- if $240^\circ < \theta < 360^\circ$, and t if $120^\circ < \theta < 240^\circ$. Among the two gauche conformations, neither g^+ nor g^- is preferred. Figure 3.1(b) shows that the trans conformation predominates along C-O bonds.

It has been observed from vibrational spectroscopy and X-ray diffraction studies that the conformation of pure PEO is altered upon complexation of metal salts, as discussed in the introduction section. So Li^+ complexation with ether oxygen atoms of tetraglyme chains is expected to alter the conformation of tetraglyme chains.

The time-averaged population density distributions of backbone dihedral angles (θ) around C-O bonds and C-C bonds for the inside region of tetraglyme are shown in Figure 3.1, for comparison with those of pure tetraglyme. Figure 3.1(a) indicates that for dihedral angles of tetraglyme: LiTf along the C-C bond, the g^+ conformation is centered around 60° , while the g^- conformation is centered around 300° , instead of the 75° and 285° angles calculated in pure tetraglyme. The population of both g^+ and g^- increases relative to those of pure tetraglyme while the t state vanishes. Compared to the distribution in pure tetraglyme, ether oxygen coordination to Li^+ apparently enforces a distribution of C-C conformations around only two discrete conformational

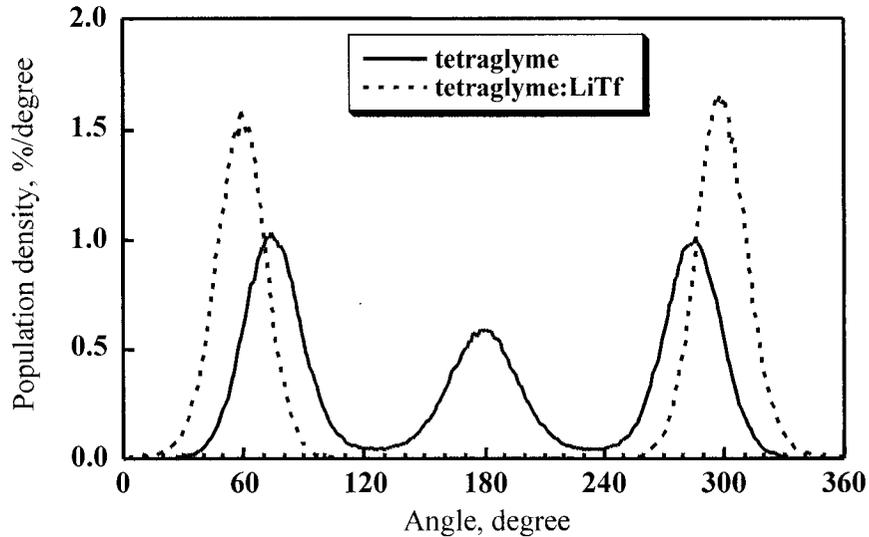


Figure 3.1(a). Average population density distributions of tetraglyme backbone dihedral angles around C-C bonds at 400K. The dashed line is for the torsions within the inside regions of tetraglyme:LiTf melts (defined in the text).

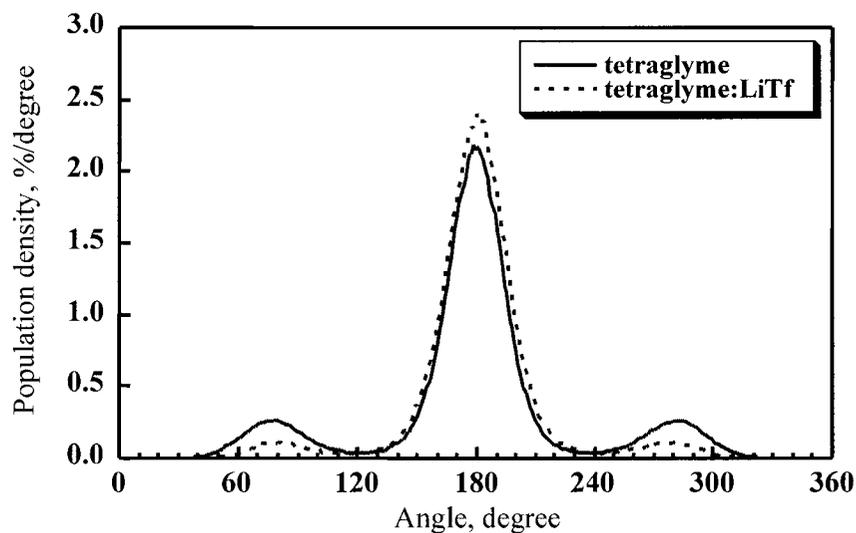


Figure 3.1(b). Average population density distributions of tetraglyme backbone dihedral angles around C-O bonds at 400K. The dashed line is for the torsions within the inside regions of tetraglyme:LiTf melts (defined in the text). (These figures were made by Dr. Jin-Kee Hyun.)

states, with angles near 60° and 300° . The distribution along the C-O bond of the inside region shown in Figure 3.1(b) shows a slightly higher population of trans conformations and less of the gauche conformations as compared to pure tetraglyme.

In Figure 3.2, only the dihedral angle population density distribution of the closest torsional angle of the outside region with respect to the inside region, or the complexing oxygen is plotted because the comparison of further parts of the outside region to pure tetraglyme shows that the torsions stay the same. For tetraglyme chains, the closest torsion should be along a C-O bond. Figure 3.2 shows that the gauche population increases upon Li^+ complexation, while the trans population decreases. Furthermore, non-complexed tetraglyme chains show the same conformational population as pure tetraglyme (not shown). It is thus concluded that the complexation of Li^+ metal cations with ether oxygens is a short range interaction and influences only the local structures of tetraglyme chains.

Vibrational spectroscopy has been used to determine conformational sequences present in ethylene oxide chains and the conformational changes occurring upon cation complexation. Bands between 800 and 950 cm^{-1} have been assigned to modes that are a mixture of CH_2 rocking and C-O-C stretching motions. [67] The frequencies and intensities of these bands are dependent on the local backbone conformation and conformational sequences. [25] The Raman spectra of pure tetraglyme and tetraglyme: LiCF_3SO_3 with a ratio of O:M=10:1 at 295 K and 400 K in the region 780 - 900 cm^{-1} are presented in Figure 3.3. The band at 850 cm^{-1} is attributed to a tgt sequence for the C-O-C-C-O-C bonds and the band at 807 cm^{-1} is

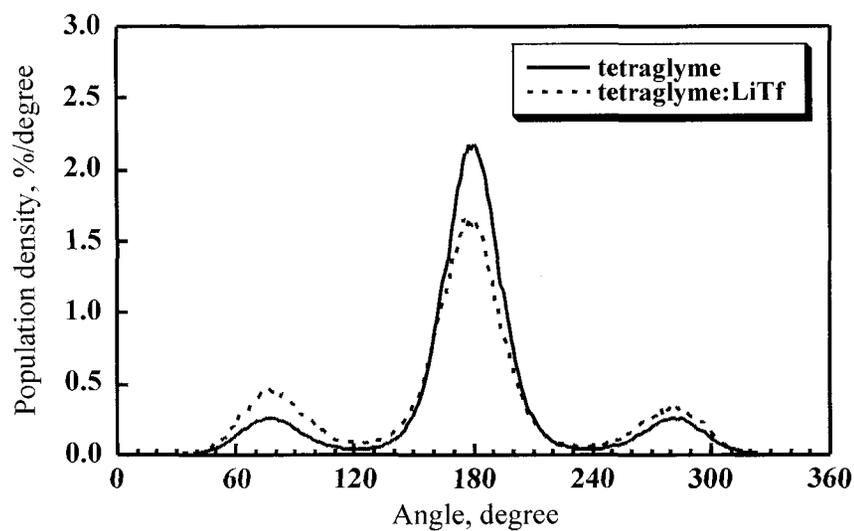


Figure 3.2. Average population density distributions of tetraglyme backbone dihedral angles around C-O bonds at 400K. The dashed line is for the closest torsion of the outside region of tetraglyme:LiTf melts (defined in the text). (This figure was made by Dr. Jin-Kee Hyun.)

attributed to a xtx conformational sequence, where x denotes a number of possible conformations. [69,70] The comparison of the Raman spectra of pure tetraglyme at 295 K and 400 K shows a band shift from 850 cm^{-1} to 845 cm^{-1} at higher temperature. This band shift suggests that elevated temperatures cause changes in the value and distribution of torsional angles.

Spectroscopic evidence for the decrease in the O-C-C-O torsional angles upon interaction of a cation with the tetraglyme oxygen atoms can be seen in the Raman spectra shown in Figure 3.3. The appearance of the high frequency band at 870 cm^{-1} signals a decrease in the value of the O-C-C-O torsional angle, in agreement with the MD results and consistent with behavior noted in other glyme:salt systems. No significant changes in band shape are observed for the tetraglyme: LiCF_3SO_3 sample at 295 K compared to 400 K, suggesting that major changes in the value and distribution of torsional angles are not occurring over this temperature range.

Populations of tetraglyme triads (C-O-C-C-O-C dihedral sequences) in pure tetraglyme and tetraglyme: LiCF_3SO_3 complexes were calculated from the MD simulation results at 300 K and 400 K. Results are shown in Figure 3.4. For tetraglyme: LiCF_3SO_3 complexes, only the tetraglyme chains that participated in complexation with Li^+ were considered. The nomenclature used to describe conformations is defined in the beginning of this section. In addition, it should be noted that g' implies a change in sign of the gauche conformation. For example, tgg' implies either tg^+g^- or tg^-g^+ conformations.

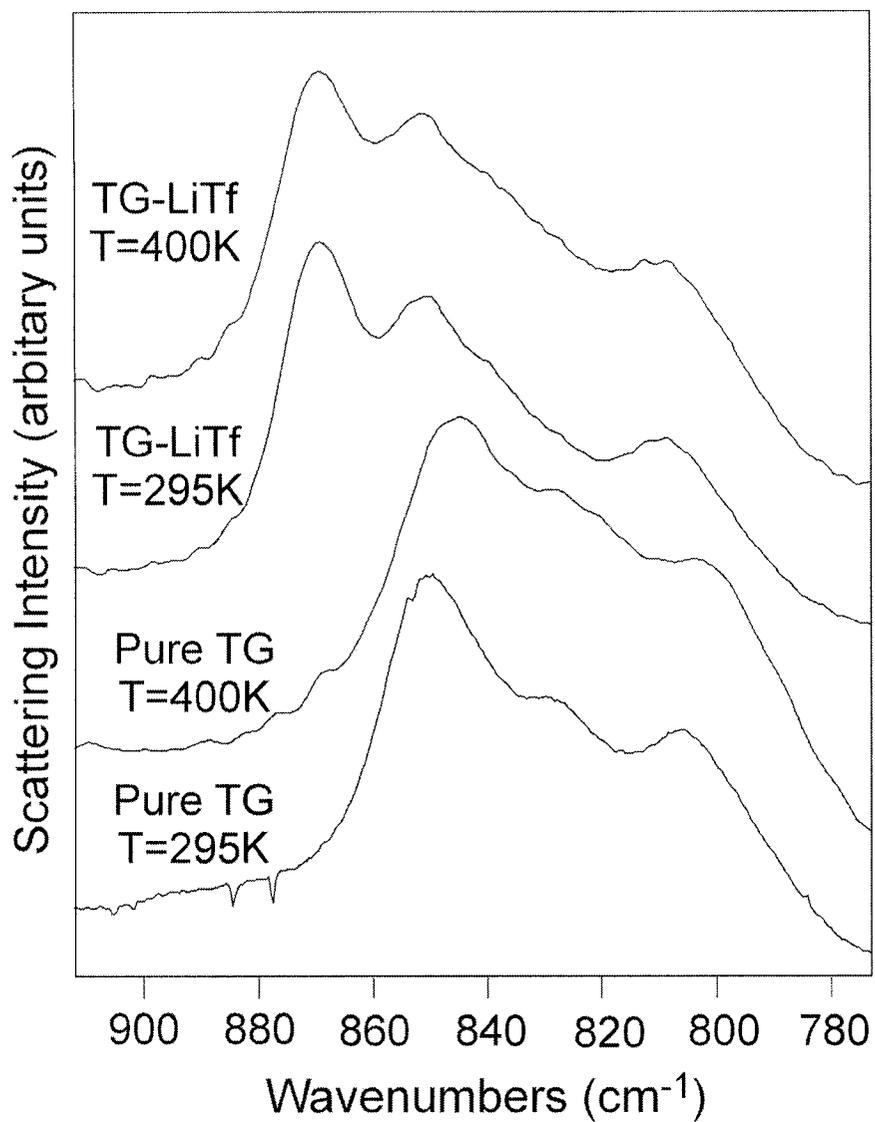


Figure 3.3. Raman spectra of pure tetraglyme and tetraglyme:LiTf (with an ether oxygen:LiTf ratio of 10:1) at 295K and 400K in the region 780 cm⁻¹ to 900 cm⁻¹. (These spectra were taken by Dr. Christopher Rhodes.)

In Figure 3.4, for pure tetraglyme, *tgt* conformers predominate and their population decreases with increasing temperature. Meanwhile the population of *ttg*, *tgg*, and *tgg'* conformers increases. Also interesting to note is the fact that the population of *ttt* conformer has no significant dependence on temperature, within the error limits of the calculation. All conformational triad populations for tetraglyme: LiCF_3SO_3 complexes show the same trend as pure tetraglyme upon increasing the temperature from 300 K to 400 K. We note that the presence of more *gauche* than *trans* conformers along the C-C bond at both 300 K and 400 K is consistent with shorter end-to-end chain lengths described in Section A, because the *gauche* conformation implies a more folded polymer chain. The decrease of the population of the most stable complexing structure, the *tgt* conformer, with increasing temperature indicates that the Li^+ -tetraglyme oxygen complexation somehow weakens at higher temperature.

3.3.3. Li^+ ion complexation.

It is found from X-ray diffraction studies on structures of PEO:salt complexes that metal cations are coordinated to both the salt anions and oxygens of PEO. For example, the crystal structure in crystalline $\text{P}(\text{EO})_3:\text{LiCF}_3\text{SO}_3$ shows that Li^+ cations are located within one helix and coordinated in an approximately trigonal bipyramidal arrangement by three adjacent PEO oxygen atoms and one oxygen atom from each of two different CF_3SO_3^- ions. [20] No cationic or anionic bridges are observed between PEO chains in $\text{P}(\text{EO})_3:\text{LiCF}_3\text{SO}_3$.

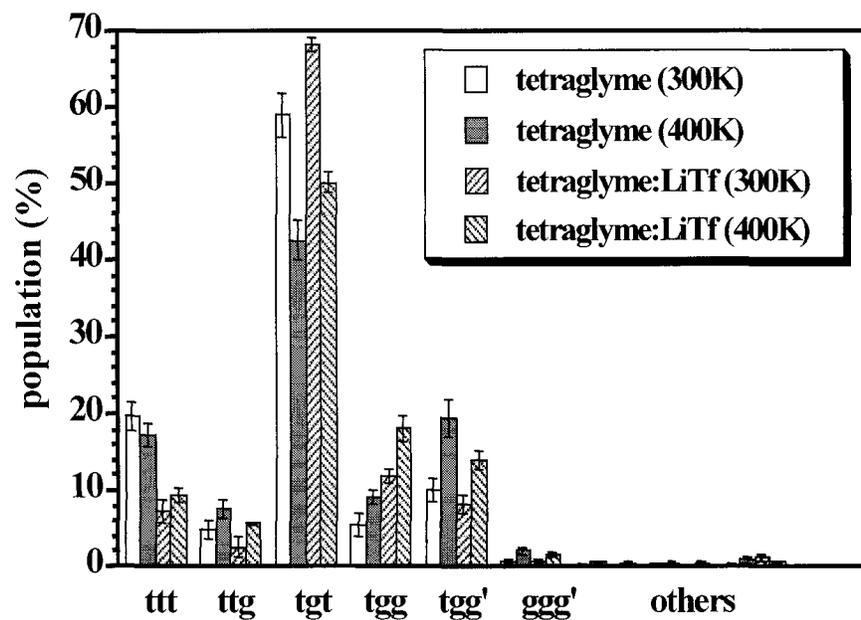


Figure 3.4. Population analysis of conformational triads (C-O-C-C-O-C atom sequences) for tetraglyme and tetraglyme:LiTf melts at 300K and 400K. Error bars at the top of each column are calculated standard deviations.

To examine how Li^+ ions are coordinated to the oxygen atoms from both tetraglyme chains and CF_3SO_3^- in amorphous tetraglyme: LiCF_3SO_3 complexes, radial distribution functions (RDF) between Li^+ and all oxygen atoms, $g_{\text{Li-O}}(r)$, were calculated from MD simulation results at $T=300$ K. Results are displayed in Figure 3.5. Figure 3.5 shows that the first strong peak occurs at about 2.0 \AA followed by a minimum at about 2.8 \AA . The RDF at 400K (not shown) has a similar shape and gives the first peak at 2.0 \AA and the first minimum at 2.9 \AA as reported in Table 3.2.

Table 3.2. The positions of the first peaks (1st p) and the first minima (1st m) in radial distribution functions for tetraglyme: LiCF_3SO_3 from MD simulations performed at 300K and 400K. The coordination numbers shown in parentheses were calculated by integrating the appropriate radial distribution functions from zero to the position of the first minimum.

	Li-O(all)		Li-O(tetraglyme)		Li-O(triflate)	
	1 st p, \AA	1 st m, \AA	1 st p, \AA	1 st m, \AA	1 st p, \AA	1 st m, \AA
300K	2.0	2.8 (6.2)	2.1	3.8 (2.1*)	2.0	2.8 (4.6)
400K	2.0	2.9 (6.2)	2.1	3.5 (1.4*)	2.0	2.8 (4.8)

*Since the first minima of $g_{\text{Li-O(tetraglyme)}}(r)$ are greater than those of $g_{\text{Li-O(all)}}(r)$, the coordination numbers of tetraglyme oxygen are larger than their contributions to the total coordination numbers, which are the numbers of oxygen atoms within the first minimums of $g_{\text{Li-O(all)}}(r)$. In order to get the correct contributions from tetraglyme oxygen, the $g_{\text{Li-O(tetraglyme)}}(r)$ are integrated from 0 to 2.8 \AA giving 1.6 and 1.3 at 300K and 400K, respectively.

The total oxygen atom coordination number for the first coordination shell was calculated by integrating $g_{\text{Li-O}}(r)$, from zero to the radius of the first coordination shell (2.8 \AA). All positions of peaks and minima in the RDF, as well as Li^+ coordination numbers are listed in Table 3.2. The large oxygen coordination numbers for Li^+ (e.g. 6.2 at 300 K) initially surprised us because four- and five-coordinate lithium ions are most common, but six-coordinate Li^+ ions are also common and coordination

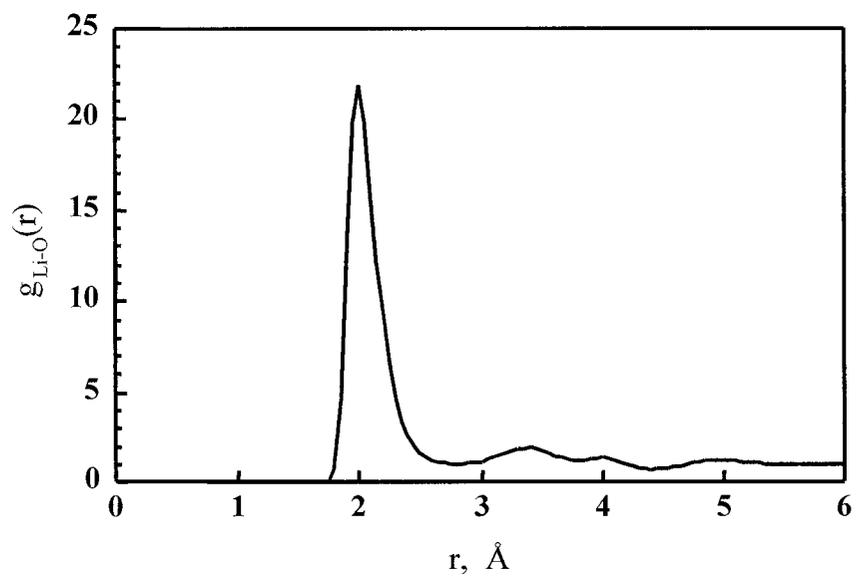


Figure 3.5. Radial distribution function between Li^+ and all oxygen atoms in a tetraglyme:LiTf melt at 300K. The first minimum occurs at $r = 2.8 \text{\AA}$ and is used as the cutoff distance of the first coordination shell.

numbers as high as eight are known for Li^+ . [71] The higher coordination numbers are in fact favored by charged, bidentate or tridentate ligands with relatively inflexible “bite angles”, similar to CF_3SO_3^- . Radial distribution functions between ether oxygen and Li^+ and between triflate oxygen and Li^+ were also calculated separately (not shown). They were integrated to distinguish the tetraglyme’s and triflate’s contribution to Li^+ complexation. This calculation shows that an average of 4.6 CF_3SO_3^- oxygen atoms coordinate Li^+ while an average of 1.6 ether oxygen atoms coordinate each Li^+ . This result indicates that there is a competition in complexation of Li^+ between ether oxygens and CF_3SO_3^- oxygens, and that the degree of coordination of Li^+ to CF_3SO_3^- oxygens is stronger than that of ether oxygens. This observation is in contrast to the crystalline $\text{P}(\text{EO})_3:\text{LiCF}_3\text{SO}_3$ compound, where a greater number of ether oxygens than triflate oxygens coordinate to Li^+ . At 400 K, the coordination number remains the same at 6.2. The average contribution from CF_3SO_3^- oxygens is 4.8, while tetraglyme oxygens contribute 1.3. This shows that the triflate oxygen is more competitive than tetraglyme oxygen in complexing with lithium ions at higher temperature. Fewer tetraglyme oxygens involved in Li^+ coordination qualitatively agrees with the higher mean-square end-to-end distance at higher temperature, since the tetraglyme chains should expand as they coordinate fewer Li^+ ions and become free to adopt trans rather than gauche conformations more frequently.

3.3.4. Ionic association.

The population of Li^+ -triflate associated species was calculated in order to compare with IR experimental data. Figure 3.6 shows IR spectra of the CF_3 symmetric deformation mode, $\delta_s(\text{CF}_3)$, for the tetraglyme: LiCF_3SO_3 sample with an ether oxygen: Li^+ ratio of 10:1 at 295 K and 400 K. The frequency of this vibrational mode has been shown to be sensitive to the nature of cation interactions with the SO_3 portion of the triflate anion. [51,72] The spectra at each temperature can be curve-fit with a linear baseline and four mixed Lorentzian-Gaussian bands at 765, 761, 757, and 753 cm^{-1} . These frequencies are assigned to the aggregate species $[\text{Li}_3\text{Tf}]^{2+}$, the triple ion $[\text{Li}_2\text{Tf}]^+$, the Li-Tf contact ion pair, and a spectroscopically “free” Tf^- ion, respectively. These assignments are based on numerous spectroscopic studies of the triflate ion vibrations in a variety of ethylene-oxide based solutions. The relative contributions of each component band to the total integrated band intensity in the $\delta_s(\text{CF}_3)$ region are summarized in Table 3.3. The integrated band intensities show that the LiTf ion pair predominates at 300 K, but significant populations of Tf^- , $[\text{Li}_2\text{Tf}]^+$, and some $[\text{Li}_3\text{Tf}]^{2+}$ also exist. At 400 K, ionic association increases, as the contributions from the larger $[\text{Li}_2\text{CF}_3\text{SO}_3]^+$ and $[\text{Li}_3\text{CF}_3\text{SO}_3]^{2+}$ aggregates increase and that from the LiTf ion pair decreases. The percentage of Li^+ ions involved in each type of aggregate observed in the MD simulations is also reported in Table 3.3 and agrees remarkably well with the contributions of each component band to the total integrated IR band intensity. It should be noted that the percentages reported in Table 3.3 were derived from MD simulations by integrating from zero to the first

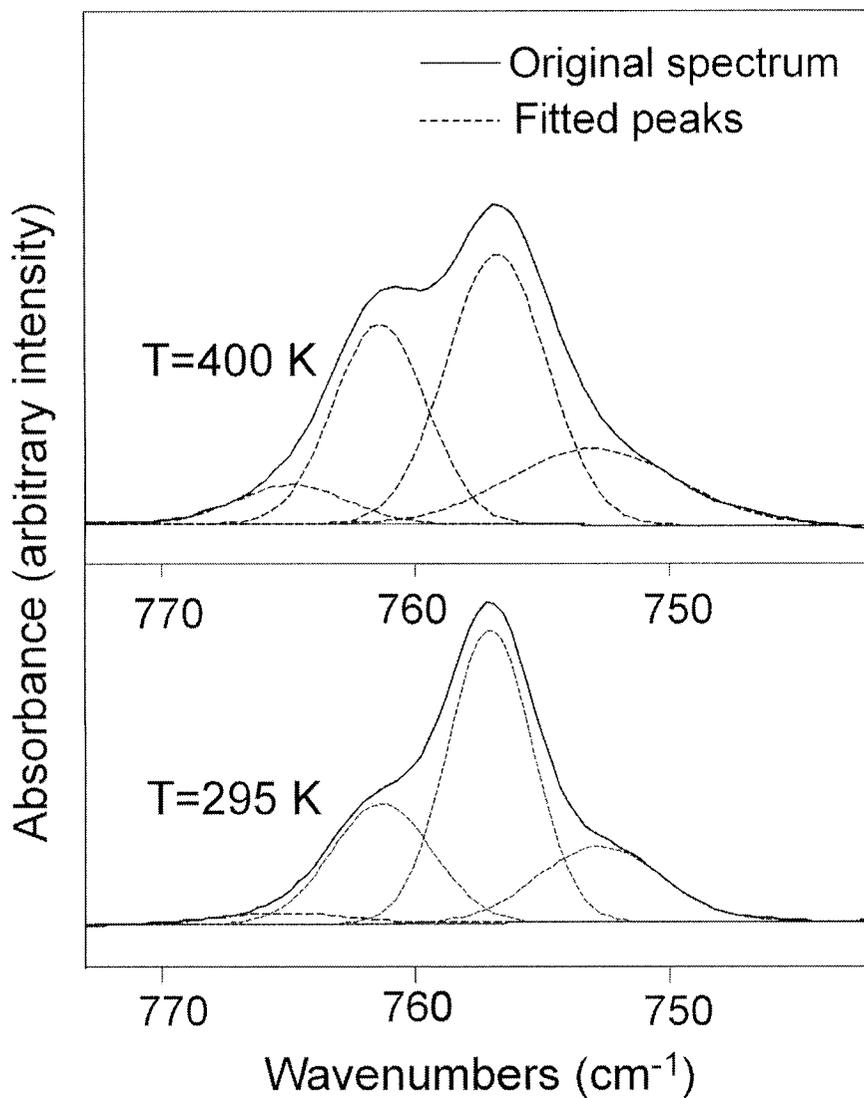


Figure 3.6. Infrared spectra of tetraglyme:LiTf (with an ether oxygen:LiTf ratio of 10:1) in the $\delta_s(\text{CF}_3)$ region at 295K and 400K. Curve-fitted peaks are shown below the recorded spectra. (These spectra and fitting were done by Dr. Christopher Rhodes.)

maximum in the lithium-triflate oxygen RDF (2.0 Å). The excellent agreement of the MD calculation with the experimental data indicates that the IR absorption of the $\delta_s(\text{CF}_3)$ mode may only be sensitive to the most probable Li-O distance. In summary, the MD simulations are consistent with IR spectral data and directly demonstrate an increase in population of larger aggregates at the higher temperature, a result that is also consistent with our MD analysis of the spatial extent of tetraglyme chains, their conformations, and Li^+ coordination by oxygen.

Table 3.3. Comparison of the percent contributions from different Li-Tf species to the total integrated intensity of the $\delta_s(\text{CF}_3)$ band observed in IR experiments with percentages of Li^+ cations bound up in different Li-Tf species in molecular dynamics (MD) simulations. (The IR results were provided by Dr. Christopher Rhodes.)

		Tf^-	LiTf	$[\text{Li}_2\text{Tf}]^+$	$[\text{Li}_3\text{Tf}]^{2+}$
300K	IR*	20	51	27	2
	MD	20	46	28	6
400K	IR	21	44	28	7
	MD	18	44	30	8

* This IR experiment was carried out at 295K.

Figure 3.7 shows the radial distribution functions between Li^+ and the triflate center of mass, $g_{\text{Li}^+-\text{Tf}}(r)$, at 300 K and 400 K to illustrate the temperature dependence of the cation-anion association. The strong first peak occurs at about 3.4 Å, followed by the minimum at about 4.5 Å. The small peaks occurring between the first, largest peaks and the first minima at both temperatures are unusual. The slight difference in the positions of the larger and smaller peaks (3.4Å vs. 3.8Å), suggests that there may be two separate geometries for lithium-triflate oxygen complexation. To investigate

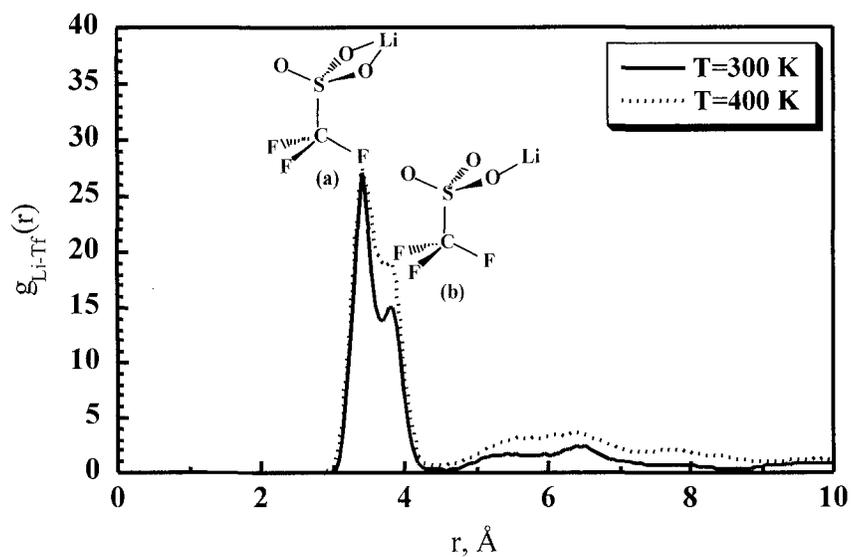


Figure 3.7. Radial distribution functions between Li^+ and the triflate center of mass at 300K and 400K. The first peaks occur at $r = 3.4\text{ Å}$ and the first minima at $r = 4.5\text{ Å}$ at both temperatures. The larger peaks are due to bidentate coordination of Li^+ by one CF_3SO_3^- ion (indicated in (a)) and the smaller peaks at slightly longer distances are due to monodentate coordination (indicated in (b)).

this possibility, we calculated radial distribution functions between Li^+ and triflate oxygen if the distances between lithium and the triflate center of mass are 3.4 Å or 3.8 Å at both temperatures. The integration of these functions must give 3 because each triflate molecule has three oxygen atoms. At 400 K, the radial distribution function for the shorter Li^+ -triflate center of mass distance (3.4 Å) gives two peaks, one at 2.0 Å and one at 3.4 Å with areas of 2.0 and 1.0, respectively. This implies that two of three triflate oxygens are bridged by one lithium ion within relatively the same distance and the third triflate oxygen is farther away from the lithium. This “bidentate” coordination of Li^+ by triflate is shown in Figure 3.7 as picture (a). The Li^+ -O radial distribution function for the longer Li^+ -triflate center of mass distance (3.8 Å) also gives two different peaks, at 1.95 Å and 4.0 Å, with areas of 1.0 and 2.0, respectively. This result clearly suggests a different geometry: a lithium ion complexes with only one triflate oxygen and the other two oxygens remain uncomplexed to that particular Li^+ ion. The picture (b) in Figure 3.7 represents this “monodentate” coordination geometry. At 300 K, the reason for the double peak in the Li^+ -triflate center of mass RDF is the same. Therefore at both temperatures, the first larger peak in $g_{\text{Li}^+-\text{TF}}(r)$ is caused by the bidentate geometry described in Figure 3.7(a) and the second, smaller peak is due to the monodentate geometry in 3.7(b). The higher first peak means that the bidentate geometry is more probable than the monodentate geometry. Upon heating from 300 K to 400 K, the height of the first peak remains approximately the same but the height of the smaller peak increases, as the monodentate geometry becomes more favorable. We note that a comparison of

the frequency splitting of the antisymmetric SO_3 stretching mode in the IR spectra of tetraglyme:LiTf (not shown) with ab initio quantum chemical calculations [72] suggests that the predominant ion pair structure in the tetraglyme:LiTf solution is monodentate rather than bidentate. However, the analysis is complicated by the presence of aggregates of different sizes. Some minor changes in the antisymmetric SO_3 stretching frequency are observed with temperature, but it is unclear if the ion pair coordination geometry is changing with temperature in solution.

The functions, $g_{\text{Li}^+-\text{Tf}}(r)$, were next integrated to yield the average number of Li^+ cations around a CF_3SO_3^- ion, $n_{\text{Li}}(r)$, as a function of distance between Li^+ and the CF_3SO_3^- center of mass. This is plotted in Figure 3.8 for both temperatures. Figure 3.8 shows that the average number of Li^+ cations around each triflate slightly increases at higher temperature and the coordination number of Li^+ in the first association shell is about 2.6 at 300 K and 3.0 at 400 K. This implies an increasing size of ionic aggregates with increasing temperature. Therefore we conclude that ionic aggregation increases at the higher temperature as the extent of monodentate coordination of Li^+ by triflate (Figure 3.7(b)) becomes more probable. The increase in monodentate coordination of Li^+ by triflate apparently allows Li^+ to coordinate to more different triflate ions.

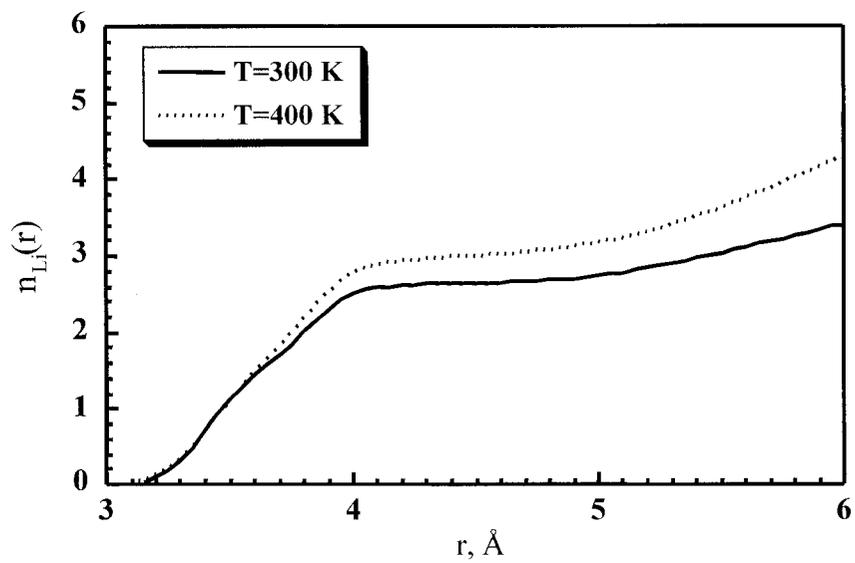


Figure 3.8. The number of Li^+ ions as a function of the distance between Li^+ and the center of mass of triflate at 300K and 400K. The flat parts of the curves give the numbers of Li^+ ions associated with one triflate ion in the first ion association shell. (This figure was made by Dr. Jin-Kee Hyun.)

3.4. Conclusions on the tetraglyme:LiCF₃SO₃ system.

Molecular dynamics simulations of pure tetraglyme and tetraglyme:LiCF₃SO₃ complexes were performed at 300 K and 400 K for comparison with structures inferred from vibrational spectroscopic measurements. Structural analyses were done to investigate tetraglyme conformations, triflate coordination by Li⁺, and ionic aggregation, and the structures at the two different temperatures were compared. The calculated mean-square radius of gyration and end-to-end distance, measures of overall chain dimensions, show that pure tetraglyme chains become less extended at the higher temperature. Upon lithium complexation by ether oxygens the chains also become more compact, but the chain dimensions of tetraglyme: LiCF₃SO₃ complexes increase at 400 K, showing that higher temperatures weaken Li⁺-ether oxygen complexation. The characteristic ratios of tetraglyme chains have also been calculated and shown to be qualitatively consistent with values obtained by small-angle neutron scattering measurements and published MD simulations of related systems.

The influence of Li⁺-O complexation on the distribution of tetraglyme chain torsional angles and the populations of conformational triads was also examined. For both temperatures, the trans conformation (t) of dihedral angles along the C-O bond of tetraglyme chains is predominant while the gauche conformation (g⁺ and g⁻) of dihedral angles along the C-C bond is favorable. Two parts of tetraglyme chains complexed by lithium cations can be described: regions “outside” the oxygens coordinated to a lithium ion and regions “inside”. For the inside region, the gauche conformation along C-C bonds is still predominant but the most probable angles are

shifted from 75° (g^+) and 285° (g^-) for pure tetraglyme to 60° and 300° after Li^+ -O complexation. For the outside region, only the torsional angles closest to the inside region are influenced significantly by Li^+ -O complexation. Raman spectra of the CH_2 rocking region provide supporting evidence for the calculated change in the O-C-C-O torsional angle resulting from the interaction of the Li^+ ion with the tetraglyme oxygens. Analysis of conformational triads (analysis of C-O-C-C-O-C dihedral angle sequences) shows that the *tgt* conformation is the most stable structure for Li^+ -O complexation and its population decreases at 400 K. This trend implies that oxygens of the tetraglyme chains compete less effectively than triflate oxygens for Li^+ at the higher temperature.

Radial distribution functions between lithium and oxygen have been calculated to investigate coordination of Li^+ by oxygen atoms of triflate ions and tetraglyme chains. Integrating the Li^+ -O radial distribution function over the first coordination shell gives an average 6.2 oxygen atoms coordinate to each lithium cation at both temperatures. The contribution from tetraglyme oxygens to the total coordination number decreases from 1.6 at 300 K to 1.3 at 400 K, while the contribution from triflate oxygens increases from 4.6 to 4.8. This result confirms that triflate oxygens are more competitive than tetraglyme oxygens in complexing with Li^+ at the higher temperature. This conclusion is consistent with the analyses of conformational triads discussed previously.

Ionic association has also been examined. The population of lithium-triflate associated species was calculated from MD results for direct comparison with IR

data. Computational and experimental results agree that the population of LiTf ion pairs decreases as temperature increases while populations of $[\text{Li}_2\text{CF}_3\text{SO}_3]^+$ and $[\text{Li}_3\text{CF}_3\text{SO}_3]^{2+}$ aggregates increase. Two geometries of cation-anion association were detected. In one geometry a lithium cation bridges two of three triflate oxygens (bidentate coordination of Li^+ by triflate); in the other geometry a lithium cation complexes with only one triflate oxygen. The bidentate coordination geometry predominates at both temperatures, but the monodentate coordination increases upon going from 300 K to 400 K.

In summary, all of our analysis is consistent with the view that ionic aggregates become larger and aggregation becomes more extensive at 400 K than at 300 K. Triflate coordination of Li^+ includes more monodentate coordination at the higher temperature and thus accommodates more bridges between two triflate ions. At the higher temperature, tetraglyme oxygens compete less effectively than triflate oxygen atoms for Li^+ ions and the gauche conformation about C-C bonds (enforced by Li^+ coordination by adjacent oxygen atoms of tetraglyme) becomes less predominant. As a result, the tetraglyme chains adopt more extended conformations in tetraglyme: LiCF_3SO_3 at 400K than at 300K, in contrast to the temperature dependence of chain conformations evident for pure tetraglyme.

3.5. Results and discussion (II) - the tetraglyme:NaCF₃SO₃ complex.

3.5.1. Tetraglyme chain dimensions.

The mean-square radius of gyration $\langle S^2 \rangle$ and mean-square backbone end-to-end distance $\langle R^2 \rangle$ describe the spatial extent of polymer chains. In this simulation, these quantities are used to investigate conformational changes of tetraglyme chains which are affected by cation-ether oxygen complexation. $\langle S^2 \rangle$ and $\langle R^2 \rangle$ can be easily calculated from the trajectories in MD simulations. $\langle R^2 \rangle$ can also be directly observed experimentally using techniques such as light scattering, and therefore provide a way to evaluate the accuracy of the applied force field. For convenience, a quantity called the characteristic ratio, defined as $C_n = \langle R^2 \rangle / nl^2$, where n is the number of backbone bonds and l^2 is the mean-square bond lengths, is calculated for comparing polymer chains with different lengths. [63] For a freely jointed polymer chain model, this ratio is theoretically predicted as 1. For a real polymer chain, this ratio may vary because of the presence of intra- and inter-molecular forces and will reach a limit as $n \rightarrow \infty$. For example, the C_∞ of PEO in the melt is approximately 5.2. [73] Smith et al. reported C_∞ values of 5.7-5.5 for nearly monodisperse PEO melts at 347-459K by small-angle neutron scattering. [42] Our pure tetraglyme result gives C_n values of 4.9 at 300K and 4.4 at 400K, which are in qualitative agreement with the experimental data and show similar temperature dependence. The $\langle S^2 \rangle$, $\langle R^2 \rangle$, and C_n of tetraglyme chains in pure tetraglyme (taken from a separate publication [74]) are listed in Table 3.4 at 300K and 400K. In order to examine the effects of cation-ether

oxygen complexation on the chain dimensions, tetraglyme:NaTf and tetraglyme:LiTf are also listed in Table 3.4.

Table 3.4. The mean-square radius of gyration, $\langle S^2 \rangle$, mean-square backbone end-to-end distance, $\langle R^2 \rangle$, and characteristic ratio C_n for pure tetraglyme, tetraglyme:NaTf, and tetraglyme:LiTf samples at T=300 and 400 K. (Dr. Jin-Kee Hyun analyzed the tetraglyme and tetraglyme:LiTf samples and provided the data.)

	T (K)	$\langle S^2 \rangle$ (\AA^2)	$\langle R^2 \rangle$ (\AA^2)	C_n
tetraglyme	300	21.3 ± 0.3	142.7 ± 4.3	4.9
	400	20.0 ± 0.4	129.5 ± 6.3	4.4
tetraglyme:NaTf	300	17.6 ± 0.4	88.7 ± 4.8	3.0
	400	17.6 ± 0.6	95.1 ± 8.7	3.3
tetraglyme:LiTf	300	18.2 ± 0.5	101.6 ± 6.4	3.5
	400	18.6 ± 0.8	110.9 ± 10.3	3.8

Compared to pure tetraglyme, all values of $\langle S^2 \rangle$ and $\langle R^2 \rangle$ in tetraglyme:MTf (M=Na, Li) are smaller at both temperatures. This indicates that either the intrachain attractions or interchain repulsions are stronger when compared to pure tetraglyme. The complexation between the metal cations and the ether oxygens is one obvious attractive force. The smaller sizes of tetraglyme chains imply that the complexation causes the chains to form a more compact shape. In the two tetraglyme:salts systems, the complexation of sodium ions with ether oxygens results in smaller values of $\langle S^2 \rangle$ and $\langle R^2 \rangle$ than complexation of lithium ions with ether oxygens at both temperatures. The calculated differences are close to the statistical error however and may not be significant. Of further interest are the characteristic ratios of tetraglyme:MTf, which show an opposite temperature dependence to what is found in pure tetraglyme for both systems, C_n increases with temperature. This effect implies that the cation

complexation with ether oxygens from the same chains becomes weaker at high temperature. The weakened cation-ether oxygen complexation may be because the cation-anion association (discussed further below) becomes more favorable in the competition of ether oxygens and triflate oxygens for sodium ions.

3.5.2. Conformations of tetraglyme chains.

Conformational analysis provides more detailed information about the tetraglyme chains and complements the preceding chain dimension calculations. In this analysis, dihedral angles of the tetraglyme backbone bonds are calculated to show the chain conformation resulting from the cation-ether oxygen complexation. Each tetraglyme chain has two kinds of dihedral angles. One is a dihedral angle around C-C bonds (an O-C-C-O atom sequence) and the other is around C-O bonds (a C-O-C-C atom sequence). The time-averaged population density distribution of the dihedral angles around the C-C bonds for pure tetraglyme at 400K is shown in Figure 3.9 (solid line, θ denotes the dihedral angles). This distribution gives three peaks as the two bigger peaks are at the two ends centered around 75° and 285° and between them the smaller peak is centered around 180° . By our definition, the dihedral angles around C-C bonds favor gauche plus (g^+ , $0^\circ < \theta < 120^\circ$) and gauche minus (g^- , $240^\circ < \theta < 360^\circ$) over trans (t, $120^\circ < \theta < 240^\circ$) conformations. This is referred to as the oxygen gauche effect. In polyether chains, a pendant oxygen atom is known to lower the energy of the gauche conformation despite the coulombic repulsion between ether oxygens, the so-called gauche effect. [63,64] For the dihedral angles around C-O bonds (not shown),

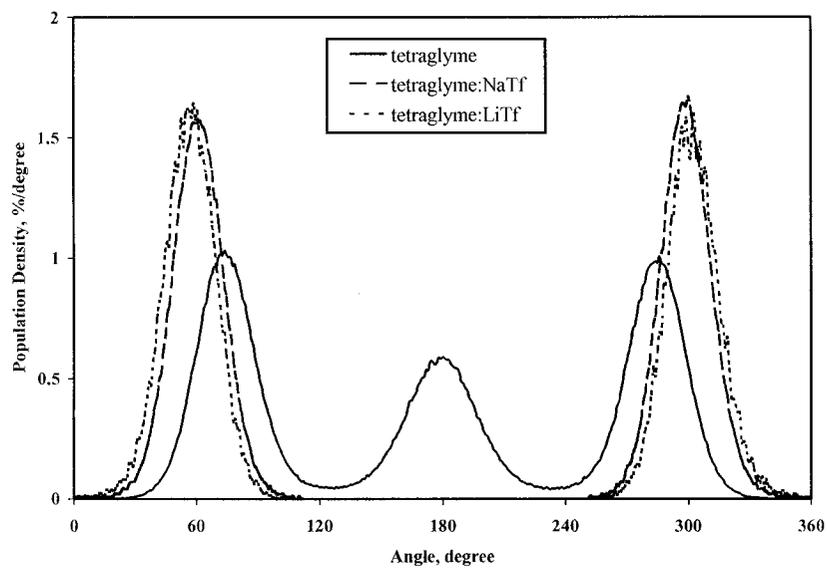


Figure 3.9. Time-average population density distributions of tetraglyme backbone torsions around C-C bonds in pure tetraglyme (solid line), tetraglyme:NaTf (dashed line), and tetraglyme:LiTf (dotted line) systems at 400K.

the distributions indicate that the t conformation is predominant. All distributions are symmetric about 180° and therefore neither g^+ nor g^- is favored over the other.

Because the two adjacent oxygens in one chain can stay in the relatively low-energy gauche conformation, they are in a geometry which allows two oxygens to coordinate a metal cation. The coordination to a metal cation affects the conformation of the O-C-C-O atom sequence as well. In order to examine the conformational changes, the time-averaged population density distributions of the dihedral angles around the C-C bonds of tetraglyme:NaTf systems at 400K are also shown in Figure 3.9 (dashed line). To preferentially investigate conformations of portions of tetraglyme chains which are complexed to sodium ions, we divided the whole chain into two parts, an “inside” region and an “outside” region. The inside region of a chain includes at least two adjacent oxygen atoms that are coordinated to a Na^+ , so this portion of chain forms a local structure of Na^+ -O complexation. The outside region is the rest of the chain. If only one oxygen atom of a tetraglyme chain complexes with a Na^+ , there should be only an outside region. If all five oxygen atoms of a tetraglyme chain complex with a Na^+ , the whole chain would define only an inside region. It should be noted that only dihedral angles of those O-C-C-O atom sequences in which both oxygens are complexed with the same metal cations (the “inside” regions) are calculated for the population density distributions. Figure 3.9 shows that g^+ and g^- conformations around C-C bonds are predominant for tetraglyme:NaTf samples. The populations of g^+ and g^- become larger than their populations in pure tetraglyme while the t conformation vanishes for the portions of

tetraglyme coordinated to sodium ions. This is because the Na^+ -ether oxygen complexation forces the two oxygens to stay in a g^+ or g^- conformation. In the tetraglyme: NaTf , the two distribution peaks of g^+ and g^- are centered around 60° and 300° , instead of 75° and 285° in pure tetraglyme. The shifts of the most probable dihedral angles of the g^+ and g^- conformations clearly show that the Na^+ -O complexation forces the two oxygens to be closer to each other to fit in two coordinate positions. To compare different cation species' influence on tetraglyme chain conformations, the distribution of the inside regions in tetraglyme: LiTf is shown in Figure 3.9 (dotted line). The two distribution peaks of the inside regions in the complexes show very small differences from each other, but the distribution of C-C angles in tetraglyme: LiTf is shifted to a smaller angle (toward the eclipsed state). This may be because the lithium ion's smaller size allows the two oxygens to be even closer to the lithium ion. However, the size does not affect the conformation very much and the complexation of sodium and lithium with ether oxygen shows a very similar influence on tetraglyme chain conformation around the C-C bonds.

Conformational triads are three consecutive dihedral angles along a C-O-C-C-O-C tetraglyme chain sequence. The three dihedral angles are around O-C, C-C, C-O bonds and their dihedral angles, as described in the previous discussion, favor t, g^+ (or g^-), and t, respectively. Therefore, the most probable triads for pure tetraglyme should be tg^+t and tg^-t (in this case, g will be used to represent either g^+ or g^- and g' to indicate a change in sign of the gauche conformation. For example, tgg' implies either tg^+g^- or tg^-g^+). Populations of tetraglyme triads in pure tetraglyme and in

tetraglyme:NaTf at 300K and 400K are also shown in Figure 3.10. In order to clarify conformational changes upon the Na^+ -O complexation, only the tetraglyme chains that participated in the complexation are considered.

In both pure tetraglyme and tetraglyme:NaTf, the tgt conformation is predominant. For pure tetraglyme, the total population of tgt is 59.0% at 300 K and 42.6% at 400K, and in tetraglyme:NaTf the population of tgt (in the “inside” regions) is 67.0% at 300K and 54.6% at 400K. At a given temperature, the population of tgt is higher in tetraglyme:NaTf compared to pure tetraglyme. The preference for the tgt conformation in tetraglyme:NaTf indicates that the Na^+ -O complexation favors this local chain conformation. It can be easily understood that a gauche angle around the C-C bond enables the two oxygens to coordinate with Na^+ , and the t state implies a further distance between the big end groups and the Na^+ which thus lower the repulsions. For tetraglyme:NaTf, the tgg and tgg' have the second (15.7% at 300K and 17.9% at 400K) and third largest (6.4% at 300K and 11.1% at 400K) populations.

Compared to pure tetraglyme (as observed in Figure 3.10), the tetraglyme chains that are involved in complexation with Na^+ have more tgg and fewer tgg' conformations. In our work, the overall ether oxygen:salt ratio of 10:1 is used, but since only complexing tetraglyme chains are considered, the ether oxygen:salt ratio will be more concentrated in these local areas. Andreev et al. reported the crystal structure for a PEO:NaTf complex with ether oxygen:salt ratio of 1:1. The structure shows that the PEO chain adopts a zigzag conformation in which the triads (C-O-C-C-O-C sequence) are tgg. [75] Similarly, the crystalline phase of tetraglyme:NaTf

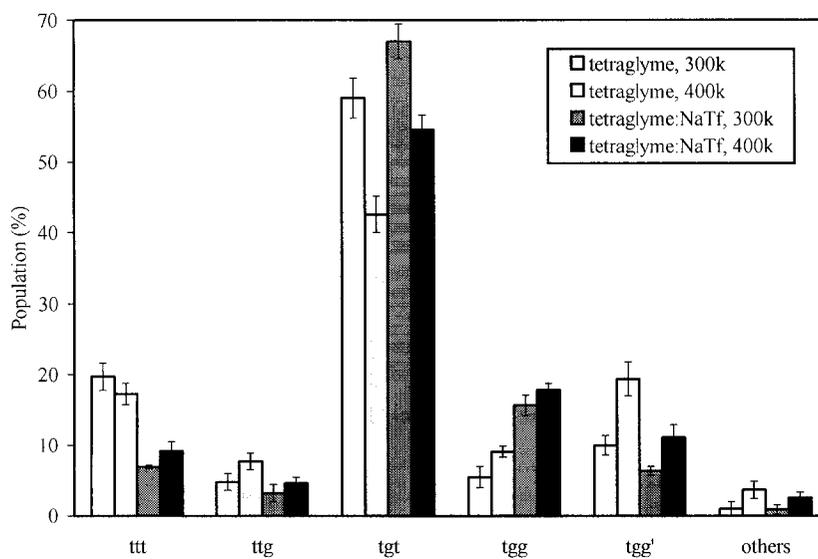


Figure 3.10. Populations of conformational triads (C-O-C-C-O-C backbone atom sequences) for tetraglyme and tetraglyme:NaTf complex at 300 and 400K. Error bars on each column are calculated standard deviations.

with an ether oxygen:salt ratio of 5:1 shows tgt and tgg conformational sequences, which is consistent with our computational results. [76] Another difference in triads population between pure tetraglyme and tetraglyme:NaTf occurs in the ttt conformation. The Na⁺-O complexation decreases its population from 19.7% and 17.3% in pure tetraglyme to 6.9% and 9.2% at 300K and 400K, respectively. The ttt conformation is the most extended chain structure and the decrease in its population implies a shorter end-to-end distance for the tetraglyme chains. This is an agreement with the chain dimension results.

Experimental evidence for the conformational changes of tetraglyme due to coordination to cations can be observed in the vibrational spectra. The bands in the 800-1000 cm⁻¹ region are particularly sensitive to conformation and have been used in PEO-salt systems to investigate conformational changes induced by the salt. [25,69] IR spectra of pure tetraglyme at 400 K and a tetraglyme:NaTf solution with ether oxygen:salt = 10:1 at 400 K are presented in Figure 3.11. The spectrum of pure tetraglyme shows a band centered at 851 cm⁻¹ and a high frequency shoulder. The band at 851 cm⁻¹ is due to ethylene oxide units or sequences which contain gauche O-C-C-O dihedral angles. The spectrum of the tetraglyme:NaTf solution shows a peak centered at 848 cm⁻¹ and a defined shoulder at 864 cm⁻¹. The 864 cm⁻¹ band is due to predominantly CH₂ wagging and CH₂ twisting motions of C-O-C-C-O-C units or sequences in tgt conformations which interact with sodium atoms, and the band has a fairly weak IR intensity. [11,77] Previous studies have shown that preferred interaction of ethylene oxide units with cations in glyme-salt solutions is for C-O-C-

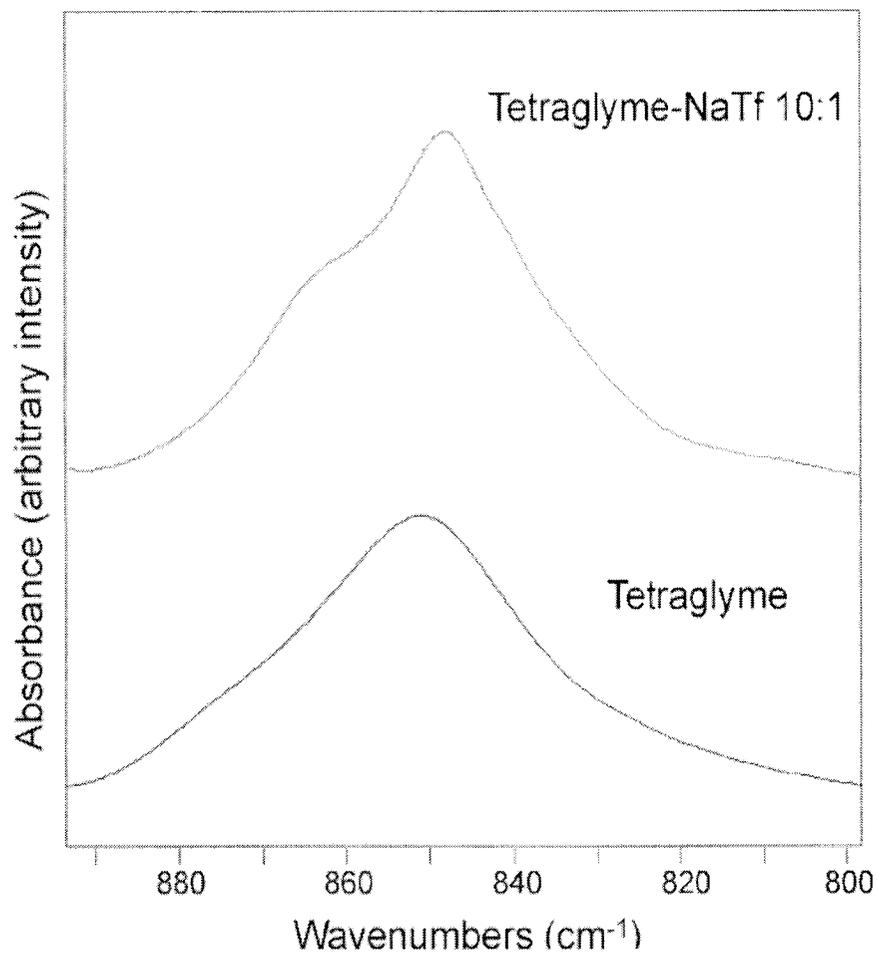


Figure 3.11. Infrared spectra in the 800-900 cm⁻¹ region of tetraglyme and tetraglyme-NaTf 10:1 solutions. (These spectra were taken by Dr. Christopher Rhodes.)

C-O-C bond sequences in trans-gauche-trans conformations. [25] The interaction of the ethylene oxide unit with a sodium ion results in a different O-C-C-O dihedral angle than is present without the sodium atom. The frequency shift of the center of the band from 851 cm^{-1} in pure tetraglyme to 848 cm^{-1} in the tetraglyme:NaTf solution is attributed to the change in the relative populations of different conformers.

3.5.3. Na^+ -ether oxygen complexation.

As seen in crystal structures of PEO:MTf (M=Na, Li) complexes, the cation can complex with oxygen from both the triflate ions and the polyether chains. [20,75] In PEO:NaTf (ether oxygen:salt = 1:1), the total coordination number for Na^+ is six, consisting of four oxygens from four different triflate ions and two oxygen from a polyether chain. In the crystal structure of tetraglyme:NaTf (ether oxygen:salt = 5:1), the total coordination number for Na^+ is seven, consisting of two oxygens from two triflate ions and five ether oxygens from a single tetraglyme chain. [76]

Radial distribution functions (RDF) are used to examine the ether oxygen and triflate oxygen coordination to Na^+ in amorphous tetraglyme:NaTf. The RDFs of Na-O(total), Na-O(triflate), and Na-O(tetraglyme) at 400K were calculated from the trajectories and are displayed in Figure 3.12. The Na-O(total) RDF shows a strong first peak at 2.4 \AA followed by a minimum at 3.4 \AA . The Na-O(triflate) RDF has a similar shape, with Na-O(total) as the first peak and minimum are located at 2.4 \AA and 3.2 \AA , respectively. The Na-O(tetraglyme) RDF has a much smaller peak at 2.5 \AA and no obvious minimum. The first peak positions, the most probable distances between

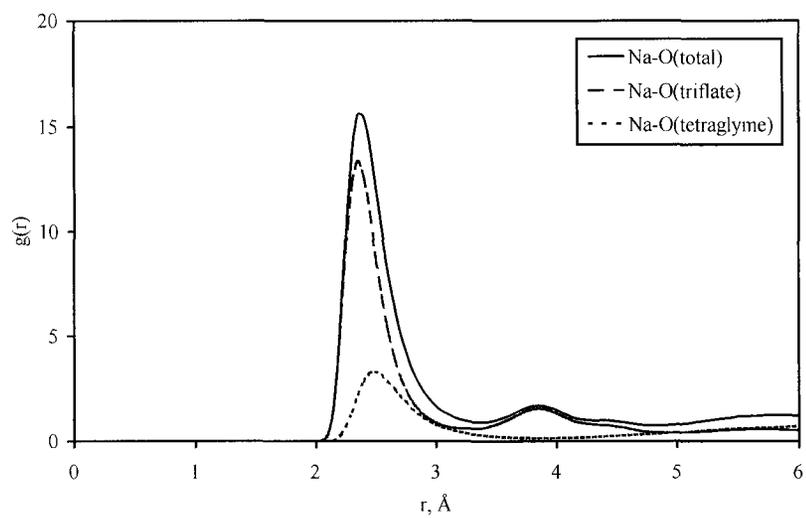


Figure 3.12. Radial distribution functions (RDF) between Na^+ and all oxygens (solid line), Na^+ and ether oxygens (dotted line), and Na^+ and triflate oxygens (dashed line) in tetraglyme: NaTf complex at 400K.

Na⁺ and oxygens, are in agreement with the mean Na-O distance of 2.46Å observed experimentally for crystalline tetraglyme:NaTf . [76]

The larger peak of the Na-O(triflate) RDF suggests that Na⁺ is coordinated by more triflate oxygens than tetraglyme oxygens. An integration of the RDFs from zero to the first minimum (the radius of the first coordination shell) provides coordination numbers. All calculated results are list in Table 3.5. The total coordination number for Na⁺ is 7.3 at 400K. This coordination number is slightly larger than the six or the seven-coordinated Na⁺ in the crystal structures. The results also give an average of 2.0 ether oxygens and 5.3 triflate oxygens coordinating each Na⁺. The greater number

Table 3.5. The coordination numbers of Na⁺ and Li⁺ in tetraglyme:NaTf and tetraglyme:LiTf samples at 300K and 400K. The total coordination numbers, the numbers of ether oxygens, and the numbers of triflate oxygens are calculated by integrating the appropriate radial distribution functions from zero to the position of the first minimum.

	O(total)		O(tetraglyme)		O(triflate)	
	Na ⁺	Li ⁺	Na ⁺	Li ⁺	Na ⁺	Li ⁺
300K	7.1	6.2	2.2	1.6	4.9	4.6
400K	7.3	6.2	2.0	1.3	5.3	4.8

of triflate oxygens indicates the Na-triflate oxygen complexation is more favorable than Na-ether oxygen complexation. This result qualitatively agrees with the PEO:NaTf crystal structure but is in contrast to the crystalline tetraglyme:NaTf structure with ether oxygen:salt = 5:1 where ether oxygens are more favored than triflate oxygens.

The RDFs of Na-O(total), Na-O(tetraglyme), and Na-O(triflate) at 300K have similar shapes (only the first minimum of the Na-O(total) RDF occurs at 3.3Å, slightly different from 3.4Å at 400K) and therefore are not shown. The coordination numbers obtained by integrating these RDFs are listed in Table 3.5. As the temperature is increased from 300K to 400K, the total coordination number increases from 7.1 to 7.3, the contribution of triflate oxygens increases from 4.9 to 5.3, and the contribution of ether oxygens decreases from 2.2 to 2.0. This result suggests that at higher temperatures the triflate oxygens compete more successfully with ether oxygens for Na⁺.

This is consistent with the mean-square end-to-end distance results of tetraglyme:NaTf complex presented above. The larger extent of tetraglyme chains at 400K implies a weaker Na-ether oxygen complexation. This result agrees with the temperature dependence shown by tetraglyme oxygen's coordination numbers.

For comparison with tetraglyme:LiTf samples, which is reported in a previous paper, [78] the coordination numbers of Li⁺ are also listed in Table 3.5. The lithium's result also shows a stronger cation-triflate oxygen complexation (4.6 at 300K and 4.8 at 400K) than cation-ether oxygen complexation (1.6 at 300K and 1.3 at 400K). The temperature dependence of coordination numbers in the two complexes also agrees each other although the coordination numbers of Li⁺ are smaller than the corresponding numbers of Na⁺. This can be attributed to the lithium ion's smaller size, which does not allow as many coordination sites as the larger sodium ion.

3.6. Comparison of ionic association in PEO-4:NaTf and PEO-4:LiTf.

Crystal structure studies of PEO:MTf complexes have shown that the metal ions strongly associate with the SO₃ portion of the triflate ions. For example, in the PEO:NaTf (ether oxygen:salt=1:1) crystal structure each Na⁺ associates with four oxygens from four triflate ions while each triflate ion also associates with four Na⁺ ions. In crystalline PEO:NaTf, two of the three oxygens in the SO₃ end of the triflate each coordinate with a Na⁺ ion, and the third oxygen bridges two Na⁺ ions. [75] In amorphous oligomer PEO:MTf system, estimates of the populations of free ion and ion associated species such as ion pairs and aggregates can be obtained from spectroscopic analysis of the symmetric deformation mode of the CF₃ portion in triflate, $\delta_s(\text{CF}_3)$, through the dependence of the frequency of the $\delta_s(\text{CF}_3)$ mode on the nature of cation-anion association. [72] A partial IR spectrum of tetraglyme:NaTf with an ether oxygen:Na ratio of 10:1 at 400K are shown in Figure 3.13. The spectrum can be curve-fit with a linear baseline and four mixed Lorentzian-Gaussian bands at 765, 761, 757, and 753 cm⁻¹, corresponding to the aggregate species [Na₃Tf]²⁺, the triple ion [Na₂Tf]⁺, the ion pair [NaTf], and the spectroscopically “free” ion Tf⁻. The frequencies of these species are consistent with studies of triflate anions coordinated to Li⁺ and Na⁺ ions in a variety of solvents. [11,25,50,72] The areas of the four bands are integrated to yield estimates of the populations of these ion associated species and are presented in Table 3.6.

The populations of ionic species are also calculated from the trajectories collected in MD simulations of the tetraglyme:NaTf sample. To investigate Na⁺-Tf⁻

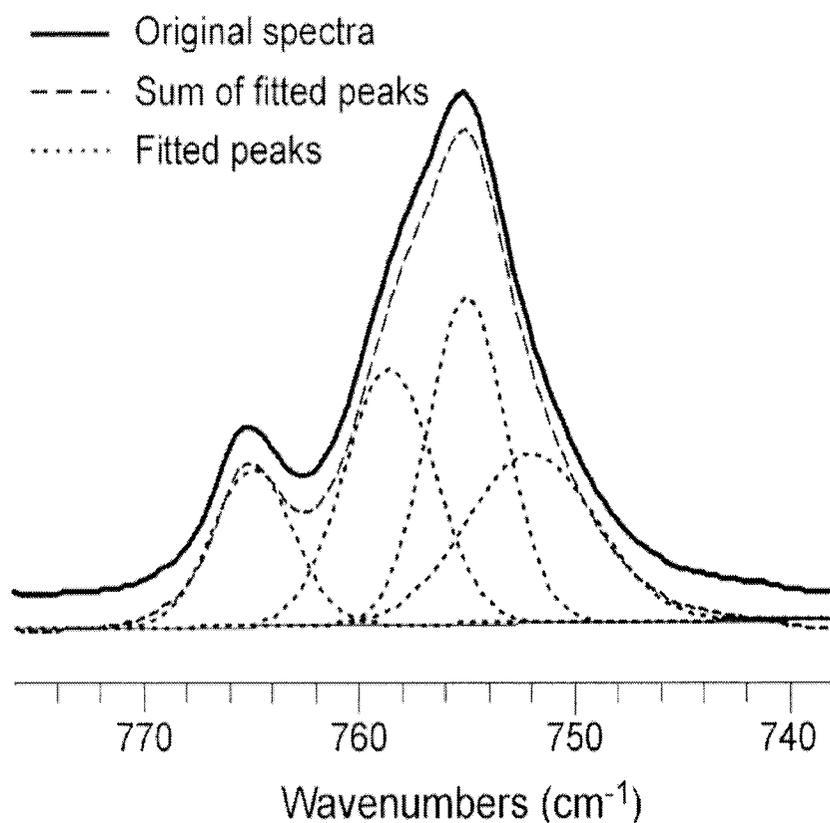


Figure 3.13. Curve-fit IR spectrum in the 740-780 cm⁻¹ region of a tetraglyme-NaTf 10:1 solution indicating the presence of different triflate species "free" triflate ions (Tf⁻) at 752 cm⁻¹, ion-pairs (NaTf) at 755 cm⁻¹, aggregate species ([Na₂Tf]⁺) at 758 cm⁻¹ and aggregate ([Na₃Tf]²⁺) at 765 cm⁻¹. (This spectrum and fitting was done by Dr. Christopher Rhodes.)

association, we examined structure from the perspective of the cation and the anion separately. We count the number of sodium ions around a triflate anion by using the

Table 3.6. Comparison of calculated populations of Na⁺-Tf⁻ associated “anion aggregates” (see text) in tetraglyme:NaTf samples from MD results at 300K and 400K with IR experimental data at 398K in parentheses. MD results for tetraglyme:LiTf samples are also listed for comparison. The ‘M’ in the species names denotes metal ions (either Li⁺ or Na⁺). (The IR results were provided by Dr. Christopher Rhodes.)

		Tf ⁻	MTf	[M ₂ Tf] ⁺	[M ₃ Tf] ⁺
300K	Na	42	40	15	3
	Li	20	46	28	6
400K	Na	9(24)	39(23)	38(35)	13(13)
	Li	18	44	30	8

first peak positions of the Na-O(triflate) RDFs, 2.5Å at both temperatures. All IR (in parentheses) and MD results are summarized in Table 3.6. At 300K, the “free” triflate ion and the ion pair are the predominant species in tetraglyme:NaTf samples as they account for 42% and 40% of the total ion associated species. The triple ion has 15% of the total populations and only a very small amount (3%) of [Na₃Tf]²⁺ aggregates is found. However, the populations of these species change dramatically as temperature increases. Only 9% of total species are found to be “free” triflate ions at 400K. The ion pair still has a large population, 39% (40% at 300K) while the triple ion and the [Na₃Tf]²⁺ aggregates demonstrate a large growth in their populations: from 15% to 38%, and 3% to 13%, respectively. The changes indicate that the cation-anion association becomes more extensive at higher temperature. This agrees with the

coordination number calculations and the chain dimension results. The results from two methods, MD simulations and IR spectroscopy, show some inconsistency for the “free” ion and ion pair, but agree well for the more highly associated species.

The populations of triflate anion species in tetraglyme:LiTf samples calculated from MD trajectories are also listed in Table 3.6. These populations agree very well with IR results. [78] Among all ionic species, the Li^+ - Tf^- ion pair is the most populous species (46% at 300K, 44% at 400K). The next most populous species are the triple ion (28%), the “free” triflate ion (20%), and the aggregate (6%) at 300K. Their populations change slightly at 400K, as they are 30%, 18%, and 8%, respectively. In tetraglyme:NaTf there are more “free” ions (42%) than triple ions (15%) at 300K, whereas in tetraglyme:LiTf there are more triple ions (38%) than “free” ions (9%). However in tetraglyme:NaTf aggregation appears highly temperature dependent, but in tetraglyme:LiTf the populations of Li^+ - Tf^- species show only a slight temperature dependence.

Radial distribution functions between Na^+ and the triflate ion’s center of mass are calculated and shown in Figure 3.14. These functions are expected to give more information of the Na^+ - Tf^- association. In our previous study of tetraglyme:LiTf system, the Li^+ - Tf^- center of mass RDF showed an interesting feature, as there was a strong first peak followed by a smaller peak at both 300K and 400K. After further study, the two peaks turned out to be two distinguishable Li^+ - Tf^- coordination geometries, the bidentate and the monodentate structure. The bidentate geometry was more populous in tetraglyme:LiTf. In tetraglyme:NaTf, each of the Na^+ -center of

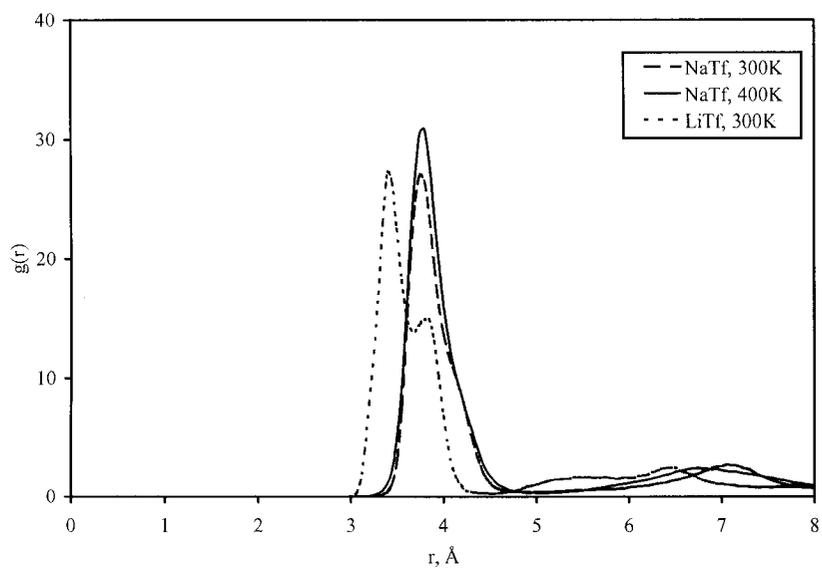


Figure 3.14. Radial distribution functions (RDF) between Na^+ and the triflate center of mass in tetraglyme:NaTf complex at 300 (dashed line) and 400K (solid line). A RDF between Li^+ and the triflate center of mass in tetraglyme:LiTf complex at 300K is shown for comparison.

mass of triflate RDFs at 300K (solid line) and 400K (dashed line) gives a big peak at 3.8Å with an indistinct shoulder.

A population analysis for cation environment has been done to further address the question of ionic association geometries. In contrast to the analysis from the triflate ion's point of view presented above, the populations shown in Table 3.7 focus on how many triflate ions are around a Na^+ and what kind of geometry is formed between the Na^+ and the triflate's three oxygens. The cutoff value used in this calculation is 3.4Å, the first minimum of Na^+ -O(triflate) RDFs. It should be noticed that this cation environment analysis only considers triflate anions which are in the first coordination shell of Na^+ . Some of the triflate anions around the Na^+ are coordinated to other Na^+ ions, and therefore possibly form more complicated ionic aggregates with several cations and anions. The result shows that the aggregate, $[\text{NaTf}_3]^{2-}$, is the dominant aggregate with 52.0% of the population at 300K and 58.9% at 400K while the triple ion, $[\text{NaTf}_2]^-$, is second most populous with 36.4% at 300K and 33.1% at 400K. The NaTf ion pair accounts for the rest of the population. The highly associated species $[\text{NaTf}_3]^{2-}$ becomes more highly populated at 400K indicating that the cation-anion association is favored with increasing temperature.

For each of the aggregates, details of the Na^+ -triflate oxygen aggregate geometries are summarized in Table 3.7. Mono, bi, and tri in the parentheses means that the Na^+ associates with oxygens through a mono, bi, and tridentate geometry, respectively. In the NaTf pair, the monodentate geometry disappears at 400K while the bidentate and

tridentate remain unchanged (Geometries with no significant temperature dependence are not listed in the table). In $[\text{NaTf}_2]^-$, the population of (mono,tri) coordination

Table 3.7. Coordination geometries for Na coordinated to triflate in tetraglyme:NaTf samples calculated from MD trajectories at 300 K and 400 K. Three species, the NaTf ion pair, $[\text{NaTf}_2]^-$, and $[\text{NaTf}_3]^{2-}$, are found for the system. The association geometries between Na^+ and triflate oxygens are listed in parentheses. Mono, bi, and tri denote monodentate, bidentate, and tridentate coordination geometries, respectively. Only geometries with significant temperature dependence are listed.

	NaTf		$[\text{NaTf}_2]^-$	
	(mono)	(mono,tri)	(bi,bi)	
300K	3.4±0.3	11.1±0.5	13.5±0.7	
400K	0	0.6±0.4	20.7±2.1	
$[\text{NaTf}_3]^{2-}$				
	(mono,mono,bi)	(mono,bi,bi)	(mono,bi,tri)	(bi,bi,bi)
300K	15.1±0.5	9.5±0.4	11.3±1.0	6.6±0.4
400K	11.6±2.6	17.8±2.4	5.2±0.9	15.5±3.5

dramatically decreases from 11.1% to 0.6% at 400K. Meanwhile, the (bi,bi) coordination becomes more highly populated (20.7%, comparing to 13.5%) with increasing temperature. In $[\text{NaTf}_3]^{2-}$, populations of (mono,mono,bi) and (mono,bi,tri) decrease at high temperature. But (mono,bi,bi) and (bi,bi,bi) show an opposite temperature dependence as their populations increase from 9.5% to 17.8% and from 6.6% to 15.5%, respectively. In general, this population analysis shows that the higher temperature favors the bidentate Na^+ - Tf^- association geometry.

3.7. Conclusions on the tetraglyme:NaCF₃SO₃ system.

Molecular dynamics simulations and vibrational spectroscopy are described to investigate structural properties of tetraglyme:NaTf system with an ether oxygen:salt ratio of 10:1 at 300 and 400K. Mean-square radii of gyration and end-to-end distances of tetraglyme chains are calculated from the MD simulations. Smaller values of these quantities for tetraglyme chains coordinating Na⁺ at both temperatures suggest that the tetraglyme chains form more compact shapes upon Na⁺-ether oxygen complexation. The population density distributions of dihedral angles around C-C bonds in tetraglyme from the MD simulations are found to favor the gauche (g⁺ or g⁻) over the trans (t) conformation while the dihedral angles around C-O bonds favor the t more than the g⁺ and g⁻ conformations. Upon Na⁺-ether oxygen complexation, dihedral angles around C-C bonds in which both oxygens coordinate to the same Na⁺ (defined as the inside regions) show predominant g⁺ and g⁻ conformations while the t conformation vanishes. Moreover, the most probable values of those angles are shifted from non-coordinating tetraglyme's 75° (g⁺) and 285° (g⁻) to 60° and 300°. The population analysis of the conformational triads (C-O-C-C-O-C sequences) in tetraglyme chains shows that tgt is the most populous conformation (59.0%). The tgt conformation is stabilized by the Na⁺-ether oxygen complexation as its population increases to 67.0% in tetraglyme:NaTf at 300K. At higher temperature, the tgt conformation in tetraglyme:NaTf decreases to 54.6%, showing that the Na⁺-ether oxygen complexation is less favored at the high temperature.

The coordination numbers of Na^+ are calculated by integrating Na^+ -O(triflate) and Na^+ -O(tetraglyme) radial distribution functions (RDFs). Overall, the sodium ions are coordinated by approximately seven oxygens (7.1 oxygen atoms at 300K and 7.3 at 400K). The triflate oxygens are clearly the major contributors to the first coordination shell (4.9 at 300K) and become even more favored (5.3) at 400K. Meanwhile, the ether oxygen's contribution decreases from 2.2 to 2.0 with increasing temperature. A comparison of the Na^+ coordination calculated in this work with the two crystal structures relevant to this work is a bit surprising. In the MD results the sodium ion is coordinated to 2.2 ether oxygen atoms and 4.9 triflate oxygen atoms at 300K. As noted earlier, this is consistent with the crystal structure of P(EO):NaTf wherein the sodium ion is coordinated to 2 ether oxygen atoms and 4 triflate oxygen atoms. However this stands in sharp contrast to the single crystal structure of tetraglyme:NaTf, where the sodium ion is coordinated to 5 ether oxygen atoms and two triflate oxygen atoms. We had expected that if a significant difference between MD and crystal structure results would occur, the discrepancy would be with the high molecular weight PEO system because of the relatively greater difficulty in packing a continuous polyether chain into the crystal. However a closer examination of the tetraglyme:NaTf crystal structure shows chains of sodium ions completely "wrapped" by all oxygen atoms of a tetraglyme molecule with triflate ions bridging adjacent "wrapped" cations. Therefore the crystal structure of the tetraglyme system is most unusual and quite unrelated to the structure of the high molecular weight PEO system that our MD simulation more closely mimics.

A number of IR bands have been assigned for $\text{Na}^+\text{-Tf}^-$ association species in tetraglyme: NaTf systems and their populations are obtained by integrating the areas. The results from MD simulations agree well with IR experiments for $[\text{Na}_2\text{Tf}]^+$ and $[\text{Na}_3\text{Tf}]^{2+}$ at 400K. The MD calculations also show that the population of free ions, Tf^- , dramatically decreases from 42% to 9% with increasing temperature while $[\text{Na}_2\text{Tf}]^+$ and $[\text{Na}_3\text{Tf}]^{2+}$ increase from 15% to 38% and from 3% to 13%. The ion pair, $[\text{NaTf}]$, remains the most populous species at both temperatures (40% at 300K and 39% at 400K). Compared to tetraglyme: LiTf system (presented previously), [78] tetraglyme: NaTf has more “free” Tf^- at 300K and more $[\text{Na}_2\text{Tf}]^+$ and $[\text{Na}_3\text{Tf}]^{2+}$ at 400K.

Monodentate, bidentate and tridentate coordination geometries of Na^+ by Tf^- were identified. For Na^+ which is coordinated to two triflate ions, $[\text{NaTf}_2]^-$, a distribution of coordination geometries exists and a double-bidentate geometry in which the Na^+ bridges four oxygens, two from each triflate ions, is favored at the higher temperature. In $[\text{NaTf}_3]^{2-}$, $\text{Na}^+\text{-Tf}^-$ association through mono-bi-bidentate and bi-bi-bidentate geometries becomes more prominent with increasing temperature. It is important to emphasize that the notation of $[\text{NaTf}_2]^-$ and $[\text{NaTf}_3]^{2-}$ includes only triflate anions which are in the first coordination shell of Na^+ . Some of the triflate anions in these species are coordinated to other Na^+ ions, and therefore exist as part of ionic aggregates rather than as isolated species. The percentages of these association geometries as a function of temperature show that the bidentate structure is favored at the higher temperature.

3.8. Acknowledgments.

We are grateful to the National Science Foundation for supporting this work through grant number EPS-9720651 and to the National Science Foundation/National Resource Allocations Committee for a supercomputer time award at the University of Illinois (award number MCA96-N019 to R.W.). We thank Scott E. Boesch for providing force field parameters for the triflate anion, Dr. Jin-Kee Hyun for providing the figures indicated and for contributing to their analysis and Dr. Christopher Rhodes for providing experimental spectra.

3.9. References.

- [1] Armand, M. B.; Chabagno, J. M.; Duclot, M. J. Polyethers as solid electrolytes. In *Fast Ion Transport in Solids*; Vashishta, P., Mundy, J. N., Shenoy, G. K., Eds.; North-Holland: New York, 1979; pp 131.
- [2] Gray, F. M. *Solid Polymer Electrolytes. Fundamentals and Technological Applications*; VCH: New York, 1991.
- [3] Gray, F. M. *Polymer Electrolytes*; The Royal Society of Chemistry: Cambridge, UK, 1997.
- [4] MacCallum, J. R.; Vincent, C. A. *Polymer Electrolyte Reviews*; Elsevier: New York, 1987; Vol. 1.
- [5] MacCallum, J. R.; Vincent, C. A. *Polymer Electrolyte Reviews*; Elsevier: New York, 1989; Vol. 2.
- [6] Ratner, M. A.; Shriver, D. F. *Chem. Rev.* **1988**, 88, 109.
- [7] Minier, M.; Berthier, C.; Gorecki, W. *J. Phys.* **1984**, 45, 739.
- [8] Berthier, C.; Gorecki, W.; Minier, M.; Armand, M. B.; Chabagno, J. M.; Rigaud, P. *Solid State Ionics* **1983**, 11, 91.
- [9] Bruce, P. G.; Vincent, C. A. *J. Chem. Soc. Faraday Trans.* **1993**, 89, 3187.
- [10] Albinsson, I.; Mellander, B.-E.; Stevens, J. R. *Polymer* **1991**, 32, 2712.
- [11] Frech, R.; Huang, W. *Solid State Ionics* **1994**, 72, 103.
- [12] Torell, L. M.; Jacobson, P.; Sidebottom, D.; Petersen, G. *Solid State Ionics* **1992**, 53-56, 1037.
- [13] Xie, L.; Farrington, G. C. *Solid State Ionics* **1992**, 53-56, 1054.
- [14] Catlow, C. R. A.; Mills, G. E. *Electrochimica Acta.* **1995**, 40, 2057.
- [15] Borodin, O.; Smith, G. D. *Macromolecules* **1998**, 31, 8396.
- [16] Borodin, O.; Smith, G. *Macromolecules* **2000**, 33, 2273.
- [17] Ennari, J.; Neelov, I.; Sundholm, F. *Polymer* **2000**, 41, 4057.

- [18] Takahashi, Y.; Tadokoro, H. *Macromolecules* **1973**, 672.
- [19] Hibma, T. *Solid State Ionics* **1983**, 9-10, 1101.
- [20] Lightfoot, P.; Mehta, M. A.; Bruce, P. G. *Science* **1993**, 262, 883.
- [21] Lightfoot, P.; Mehta, M. A.; Bruce, P. G. *J. Mater. Chem.* **1992**, 2, 379.
- [22] Lightfoot, P.; Nowinski, J. L.; Bruce, P. G. *J. Am. Chem. Soc.* **1994**, 116, 7469.
- [23] Papke, B. L.; Ratner, M. A.; Shriver, D. F. *Chem. Solids* **1981**, 42, 493.
- [24] Dupon, R.; Papke, B. L.; Ratner, M. A.; Whitmore, D. H.; Shriver, D. F. *J. Am. Chem. Soc.* **1982**, 104, 6247.
- [25] Frech, R.; Huang, W. *Macromolecules* **1995**, 28, 1246.
- [26] Bernson, A.; Lindgren, J.; Huang, W.; Frech, R. *Polymer* **1995**, 36, 4471.
- [27] Peterson, G.; Jacobsson, P.; Torell, L. M. *Electrochim. Acta* **1992**, 37, 1495.
- [28] Huang, W.; Frech, R. *Polymer* **1994**, 35, 235.
- [29] Chintapalli, S.; Frech, R. *Electrochimica Acta.* **1995**, 40, 2093.
- [30] Huang, W.; Frech, R.; Johansson, P.; Lindgren, J. *Electrochimica Acta.* **1995**, 40, 2147.
- [31] Bishop, A. G.; MacFarlane, D. R.; McNaughton, D.; Forsyth, M. *J. Phys. Chem.* **1996**, 100, 2237.
- [32] Muhuri, P. K.; Das, B.; Hazra, D. K. *J. Phys. Chem. B* **1997**, 101, 3329.
- [33] Chintapalli, S.; Quinton, C.; Frech, R.; Vincent, C. A. *Macromolecules* **1997**, 30, 7472.
- [34] Bandara, H. M. N.; Sachlindwein, W. S.; Latham, R. J.; Linford, R. G. *J. Chem. Soc. Faraday Trans.* **1994**, 90, 3549.
- [35] Yang, X. Q.; Lee, H. S.; McBreen, J.; Xu, Z. S.; Skotheim, T. A.; Okamoto, Y.; Lu, F. *J. Chem. Phys.* **1994**, 101, 3230.
- [36] Chintapalli, S.; Frech, R.; Grady, B. *Polymer* **1997**, 38, 6189.

- [37] Davies, G. R.; Ward, I. M. *Polymer* **1999**, *40*, 3945.
- [38] Neyertz, S.; Brown, D.; Thomas, J. O. *J. Chem. Phys.* **1994**, *101*, 10064.
- [39] Neyertz, S.; Brown, D.; Thomas, J. O. *Electrochimica Acta.* **1995**, *40*, 2063.
- [40] Neyertz, S.; Brown, D. *J. Chem. Phys.* **1995**, *102*, 9275.
- [41] Lin, B.; Boinske, P. T.; Halley, J. W. *J. Chem. Phys.* **1996**, *105*, 1668.
- [42] Smith, G. D.; Yoon, D. Y.; Jaffe, R. L.; Colby, R. H.; Krishnamoorti, R.; Fetters, L. J. *Macromolecules* **1996**, *29*, 3462.
- [43] Laasonen, K.; Klein, M. L. *J. Chem. Soc., Faraday Trans.* **1995**, *91*, 2633.
- [44] Muller-Plathe, F.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *103*, 4745.
- [45] Payne, V. A.; Xu, J.-H.; Forsyth, M.; Ratner, M. A.; Shriver, D. F.; de Leeuw, S. W. *Electrochimica Acta* **1995**, *40*, 2087.
- [46] Payne, V. A.; Xu, J.-h.; Forsyth, M.; M.A., R.; Shriver, D. F.; de Leeuw, S. W. *J. Chem. Phys.* **1995**, *103*, 8734.
- [47] Payne, J. A.; Xu, J.-H.; Forsyth, M.; Ratner, M. A.; Shriver, D. F.; de Leeuw, S. W. *J. Chem. Phys.* **1995**, *103*, 8746.
- [48] Neyertz, S.; Brown, D. *J. Chem. Phys.* **1996**, *104*, 3797.
- [49] Londono, J. D.; Annis, B. K.; Habenschuss, A.; Borodin, O.; Smith, G. D.; Turner, J. Z.; Soper, A. K. *Macromolecules* **1997**, *30*, 7151.
- [50] Rhodes, C. P.; Frech, R. *Solid State Ionics* **1999**, *121*, 91.
- [51] Huang, W.; Wheeler, R. A.; Frech, R. *Spectrochim. Acta A* **1994**, *50*, 985.
- [52] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- [53] Muller-Plathe, F. *Acta Polymer.* **1994**, *45*, 259.
- [54] Boesch, S. E.; Wheeler, R. A., Unpublished work.

- [55] Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.
- [56] Wheeler, R. A. *J. Am. Chem. Soc.* **1994**, *116*, 11048.
- [57] Raymond, K. S.; Grafton, A. K.; Wheeler, R. A. *J. Phys. Chem. B* **1997**, *101*, 623.
- [58] Wise, K. E.; Grafton, A. K.; Wheeler, R. A. *J. Phys. Chem. A* **1997**, *101*, 1160.
- [59] Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 5*; University of California-San Francisco: San Francisco, 1997.
- [60] Berendsen, H. T. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
- [61] Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
- [62] Essmann, V.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577.
- [63] Flory, P. J. *The Statistical Mechanics of Chain Molecules*; Hanser: New York, 1988.
- [64] Abe, A.; Mark, J. E. *J. Amer. Chem. Soc.* **1976**, *98*, 6468.
- [65] Jaffe, R. L.; Smith, G. D.; Yoon, D. Y. *J. Phys. Chem.* **1993**, *97*, 12745.
- [66] Inomata, K.; Abe, A. *J. Phys. Chem.* **1992**, *96*, 7934.
- [67] Abe, A.; Tasaki, K. *J. Mol. Struct.* **1986**, *145*, 309.
- [68] Abe, A.; Inomata, K. *J. Mol. Struct.* **1991**, *245*, 399.
- [69] Matsuura, H.; Fukuhara, K. *J. Poly. Sci.: Part B: Polymer Physics* **1986**, *24*, 1383.
- [70] Yoshida, H.; Matsuura, H. *J. Phys. Chem. A* **1998**, *102*, 2691.
- [71] Olsher, U.; Izatt, R. M.; Bradshaw, J. S.; Dalley, N. K. *Chem. Rev.* **1991**, *91*, 137.

- [72] Huang, W.; Frech, R.; Wheeler, R. A. *J. Phys. Chem.* **1994**, *98*, 100.
- [73] Kawaguchi, S.; Imai, G.; Suzuki, J.; Miyahara, A.; Kitano, T.; Ito, K. *Polymer* **1997**, *38*, 2885.
- [74] Dong, H.; Hyun, J.-K.; Durham, C.; Wheeler, R. A. *Polymer* **2001**, *42*, 7809.
- [75] Andreev, Y. G.; MacGlashan, G. S.; Bruce, P. G. *Phys. Rev. B* **1997**, *55*, 12011.
- [76] Rhodes, C. P.; Khan, M.; Frech, R. *J. Phys. Chem. B* **2002**, *106*, 10330.
- [77] Boesch, S. E.; Rhodes, C. P.; Frech, R.; Wheeler, R. A. *Phys. Chem. Comm.* **2002**, submitted.
- [78] Hyun, J.-K.; Dong, H.; Rhodes, C., P.; Frech, R.; Wheeler, R. A. *J. Phys. Chem. B* **2001**, *105*, 3329.

Part II

Chapter 4

Calculating accurate vibrational frequencies: the method of principal component analysis and its application to molecular dynamics simulations.

4.1. Quasiharmonic analysis and background information of principal component analysis.

Quasiharmonic analysis, as an alternative to the conventional normal mode analysis, has been used to obtain vibrational frequencies, modes, and configurational entropies of large systems from MD simulations using empirical force fields. [1-7] This method first computes the variances and covariances (mean square displacements) of atomic positions in a molecular trajectory to construct an effective force constant matrix. Then a diagonalization of this matrix is followed to give eigenvalues and eigenvectors, from which the vibrational frequencies, modes and other quantities can be directly calculated.

The theory of quasiharmonic analysis is adopted from a statistical technique, commonly known as principal component analysis (PCA) which has been extensively used in multichannel mapping in geology, weather forecasting in meteorology, and signal processing and pattern recognition in engineering. [8-10] PCA primarily serves as a dimensionality reduction technique in dealing with a large group of experimental or computational data, which usually forms a multivariate time series (the original data

matrix). The eigenvectors of the covariance matrix constructed from this data matrix define an orthogonal coordinate transformation, which transforms the original data matrix from the observation coordinate system to a new coordinate system. In the new coordinate space, the column vectors are aligned in a pattern such that the first vector specifying a particular “direction” in which accounts for most of the total variances, and the next vector specifies a direction perpendicular to the previous and accounts for most of the remaining variances, and so on. For a large set of original data, the major part of the total variances in the observables can be summed or “compressed” in a few vectors and the remaining variance can be truncated without introducing a significant amount of error. This special set of orthonormal vectors given by PCA is the set of eigenvectors of a covariance matrix calculated from the original data. PCA is often given other names, e.g., factor analysis, Karhunen-Loeve transform, and singular value decomposition (SVD). In the area of multivariate time series analysis, these techniques are equivalently employed and their relationships and further discussions can be found in Ref. [11-13].

The application of PCA to molecular vibrations is based on the special feature that in the transformed data the variances are completely separate. In terms of atomic movements (variances) in a piece of molecular trajectory, this ability of PCA makes the separation of vibrational modes possible. Since a Cartesian coordinate system is the natural choice for running molecular dynamics simulations but very inconvenient for expressing molecular vibrations, a coordinate transformation is desired in order to study the vibrational modes. PCA provides this transformation in matrix form in which the column vectors are eigenvectors of the covariance matrix. This coordinate-

transforming treatment in principle is the same as used in normal mode analysis and the similarities and differences between these methods will be described in the following section.

4.2. PCA comparison with conventional spectrum estimation methods. ^a

4.2.1. PCA and normal mode analysis.

Theoretical calculation of vibrational frequencies and modes has greatly assisted mode assignment in experimental spectroscopic analysis. The conventional normal mode calculation involves solving an eigenvalue problem of the force constant matrix constructed by evaluating the second derivative of potential energy with respect to molecular configurations. [14] This method is limited by the pure harmonic approximation to the real potential energy, e.g., using a “parabola” with constant curvature to approximate the shape of the potential energy surface, and therefore is only valid when the system stays at the energy minimum with an infinitesimal amount of vibrations. Another difficulty in calculating accurate frequencies emerges from the condensed phase calculation. The non-covalent effects on molecular vibrations are anharmonic. Quenched normal mode analysis, [15,16] instantaneous normal mode analysis (INM), [17-19] or combined quantum mechanical/molecular mechanical (QM/MM) [20,21] simulations have been performed to replicate solute-solvent interactions. However, these methods are still limited by the harmonic approximation.

The name “quasi-harmonic”, e.g., applying PCA to molecular vibrations, implies that this method can incorporate anharmonicity in the harmonic approximation frame. This method distinguishes itself from normal mode analysis by calculating the variances and covariances (the second order moments if considering the molecular

^a This chapter is partially reproduced with permission from: Wheeler, R. A.; Dong, H. *ChemPhysChem* **2003**, *4*, 382-384 and Wheeler, R. A.; Dong, H. *ChemPhysChem* **2003**, *4*, 1227-1230. © 2003 WILEY-VCH Verlag GmbH&Co. KGaA, Weinheim.

trajectory as a spatial population distribution of atomic positions) instead of the second derivatives. To be more exact, the covariance matrix is inversely proportional to the force constant matrix (consisting of the second derivatives of energy with respect to coordinates). Moreover, a single low energy-optimized geometry is used to calculate the second derivatives in normal mode analysis and quenched normal mode analysis. (In INM, each of the instantaneous structures, still a single point on the potential energy surface, is used.) The calculated frequencies are only valid for small vibrations around this point and do not reflect any anharmonicity. PCA, on the other hand, calculates the second order moments based on the spatial distribution of atomic positions, which is not a single point but the actual part of the energy surface sampled during the simulation. This can best be seen in a schematic sketch for a one-dimensional case in Figure 4.1. If the potential energy is harmonic, the distribution is a Gaussian and the second order moments are sufficient to describe it. The fitting given by PCA should be exactly the same as from normal mode analysis. When there is some anharmonicity present in the potential energy surface, PCA will give a fit of the best Gaussian distribution to the sampled portion that may contain non-Gaussian characters, while normal mode analysis still remains harmonic, based on the curvature of energy as a function of coordinates, evaluated the energy minimum point.

Quasiharmonic analysis/PCA has found direct application in solving vibrations of large systems and condensed-phase mode analysis. The second order moments are evaluated much faster than the second derivatives when the system size gets even moderately large. In condensed phases, all the normal mode analysis-based methods

like INM require a matrix diagonalization at every time interval during the entire trajectory in order to study the vibrations about the energy minimum point. Even for modern computers this can be extremely time-consuming when macromolecular systems (e.g. protein folding or enzyme mechanisms) are studied in condensed phases. In contrast, quasiharmonic analysis calculates a time average of the effective force constant matrix (inverse covariance matrix) and then performs only one diagonalization of the final matrix. This procedure therefore considerably reduces computational loads. All these advantages have made PCA a favorable choice over the conventional normal mode analysis for large and condensed-phase systems. Indeed, PCA has been used to identify several large-scale, low frequency motions of proteins. [22-24] However, not all modes, particularly high frequency modes were calculated in their work.

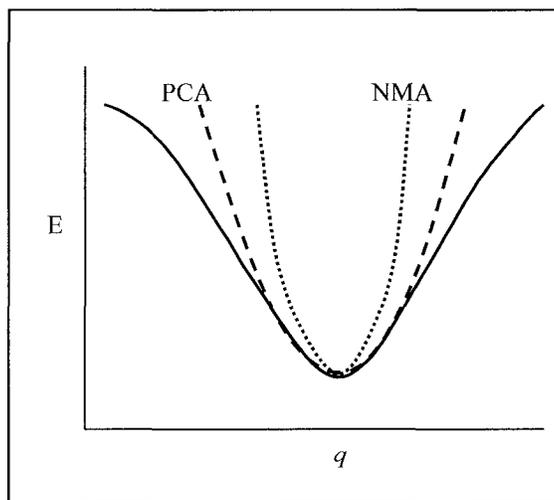


Figure 4.1. A schematic sketch of the potential energy surface (solid line) being approximated by PCA (dashed line) and NMA (normal mode analysis, dotted line). PCA with the ability to incorporate anharmonicity fits the real surface better than NMA which fits the surface with a pure parabola.

4.2.2. PCA and Fourier transform.

Fourier transform, one of the most widely used numerical analysis methods, relates two different representations of the same function, one in the time domain and the other in the frequency domain. In molecular vibrations, spectrum estimation is traditionally done by computing Fourier transforms of time correlation functions. [25-27] In the time domain, the correlation functions measure the degree to which two dynamic properties of the molecular system are correlated over a period of time. The Fourier transforms of these functions give amplitudes at a series of frequencies to reveal the dynamics of the system in the frequency representation. Important time correlation functions for molecular systems include the atomic velocity correlation function whose Fourier transform is the “power” spectrum in theoretical calculations of molecular vibrations. Other examples directly related to experiment are the dipole moment, polarizability, and magnetization correlation functions, whose Fourier transforms are infrared, Raman, and NMR line shapes, respectively. [25] However, the applications of Fourier transforms to spectrum estimation are restricted by assumptions of periodic or stationary time series and a linear system so that data may be projected onto complex sinusoidal basis functions. [28,29] PCA, on the other hand, although sharing many limitations of Fourier transforms, is more flexible than Fourier transforms because it gives data-adaptive basis functions in addition to the amplitudes of frequency components. A theoretical treatment of the relationship between PCA and Fourier transform can be found in Ref. [30].

4.2.3. Numerical and water vibrational tests of PCA.

To test PCA's ability to calculate vibrational frequencies accurately, a simple system, water in gas and liquid phases is used in this section. An isolated water molecule has three fundamental vibrations, the H-O-H bending mode at 1593 cm^{-1} , the symmetric O-H stretching mode at 3657 cm^{-1} , and the antisymmetric O-H stretching mode at 3756 cm^{-1} . [31] The spectrum of the water dimer in gas phase has been studied by matrix isolation technique and rotation-vibration-tunneling spectroscopy. [32-41] The vibrations of the water dimer are shown to roughly have two ranges, the high-frequency (above 1600 cm^{-1}) intramolecular motions and the low-frequency (well below 1000 cm^{-1}) intermolecular motions which are combinations of each individual water molecule's translational and rotational motions. Liquid water's spectrum shows that the vibrations of water undergo substantial changes as the bending mode blue-shifts to 1635 cm^{-1} and the two stretching modes red-shift to 3240 cm^{-1} and 3446 cm^{-1} . These frequencies are listed in Table 4.1 for comparison with results of PCA.

The computer simulations of the water systems were set up using the procedures below. For the isolated water molecule, 12 vibrationally excited starting geometries were generated by selecting random water molecules from a box of 216 water molecules, equilibrated at constant temperature using periodic boundary conditions and the TIP3P force field. [42] For each starting geometry, 1 ps constant energy quantum MD trajectories were generated by using the B3LYP/6-31G(d) method, [43] a 1-fs time step, and the QMD option of the computer program NWChem. [44] Next, PCA was used to calculate vibrational frequencies and modes for each of the 12

trajectories, average the frequencies for each mode, and calculate standard deviations to estimate errors (the averaging procedure was judged necessary because anharmonicity of the energy surface renders calculated frequencies amplitude dependent). Time steps as low as 0.1 fs were tested to find a balance between accuracy and computer time required. Mode assignments were made by using the XMol program [45] and projecting modes obtained by PCA onto conventional normal modes. [46] Simulations for water dimer were performed in exactly the same way, but using the 6-311++G(2d,2p) basis set because it reportedly gives a qualitatively correct geometry for (H₂O)₂. [47] For liquid water, a single water molecule was defined as the QM region and the other 215 molecules as the MM region. Then a 1 ps trajectory was generated using the CHARMM/GAMESS combination, [20] the HF/6-31G(d) method, a TIP3P water model, [42] and a 1 fs time step, in the NVE ensemble. “Standard quantum chemical” results in Table 4.1 were obtained by using the same quantum chemical methods as for generating trajectories, but the vibrational problem was solved by using the normal mode analysis method at the optimized geometries of H₂O and (H₂O)₂. Although vibrational frequencies calculated from standard quantum chemical methods are sometimes multiplied by empirical scaling factors to correct for anharmonicities and errors inherent in the methods, [48] the table presents unscaled frequencies.

In Table 4.1, for two of water’s three modes, PCA-derived frequencies are within one standard deviation of the standard quantum chemical results and the two O-H stretching frequencies are closer to experiment than the quantum chemical results.

Vibrational frequencies of $(\text{H}_2\text{O})_2$ in Table 4.1 show that with one exception intramolecular vibrations are closer to experiment than those derived from standard quantum chemistry. The one exception is important, as the relative ordering of the acceptor and donor bending modes is reversed and the large standard deviation implies significant error in the calculated bending frequencies. Calculated frequencies for intermolecular vibrations are much closer to experiment than harmonic frequencies derived from standard quantum chemistry (but not quite as close as those from MP2 calculations corrected explicitly for anharmonicity [49]). It is noted that modern spectroscopic experiments to probe intermolecular dynamics at frequencies less than approximately 150 cm^{-1} are interpreted using rotation-vibration-tunneling Hamiltonians [50] rather than harmonic oscillator equations. Calculated frequencies for liquid water in Table 4.1 show good agreement with experiment. The large standard deviation in each frequency reflects the inhomogeneous spectral broadening observed experimentally. Overall, results for water, water dimer, and liquid water imply that PCA of quantum MD or QM/MM trajectories (or analogous Monte Carlo simulations) shows promise for going beyond the harmonic approximation to model intra- and inter-molecular vibrations, as well as condensed-phase molecular vibrations. It also allows the attachment of realistic error estimates to calculated frequencies.

To compare PCA with Fourier transform-based methods for extracting frequencies from time series, both numerical and water vibrational tests are presented. The three numerical models are harmonic, anharmonic with a quartic term in the potential, [51] and Langevin oscillators. The Langevin oscillator is a harmonic oscillator subject to

frictional damping and a randomly fluctuating force and is frequently used as a model for condensed-phase processes. [52] The equations of motion of these oscillators are analytically solvable and therefore serve as ideal test cases. The initial conditions of the oscillators used in computer simulations (with the “leapfrog” scheme of the Verlet algorithm, [53] to integrate the equations of motion. Also see Section 1.2) can be found as the footnotes in Table 4.2. The frequencies from their analytical solutions are listed in Table 4.2 as well. For water test cases, one representative trajectory for the isolated and liquid water molecules are borrowed from the 12 trajectories averaged to present results in Table 4.1, to show the differences in the frequencies calculated by PCA, Fast-Fourier transform (FFT), and maximum entropy method (MEM, a popular Fourier transform-based technique). [54] The trajectories (initially 1000 steps for PCA) were extended to 30 times longer for FFT analysis to give a resolution in wavenumbers comparable to PCA and MEM. The experimental results for water are discussed previously and listed in Table 4.2 for comparison.

Since the equations of motion of the harmonic oscillator are pure harmonic and can be solved exactly by Fourier transform techniques, accurate frequencies are expected from FFT, MEM, and PCA. Indeed PCA gives frequencies slightly closer to the exact solutions for both the harmonic and anharmonic oscillators. Moreover, results in Table 4.2 indicate that the Langevin oscillator is damped too quickly for FFT to give a well defined frequency, but PCA gives a result slightly closer to the analytical frequency than the MEM technique. For the water molecule and liquid water, Table 4.2 shows that PCA of the quantum MD and QM/MM trajectories gives

frequencies significantly closer to experimental values than either the FFT or MEM results. The superior accuracy of PCA is particularly striking for the higher frequency O-H stretching modes, whose frequencies calculated by using PCA are ~15% closer to experiment (approximately 490 cm^{-1} closer) than the next best method (MEM).

In conclusion, principal component analysis is demonstrated to give overall more accurate frequencies than the conventionally used spectrum estimation methods such as normal mode analysis and Fourier transform-based techniques. By constructing a much simpler covariance matrix than the force constant matrix, PCA calculates the frequencies much faster and more accurately due to its ability to incorporate anharmonicity, compared to normal mode analysis. Although the FFT algorithm makes Fourier transforms faster than PCA (but the FFT method needs much longer trajectories to have reasonable resolution), numerical tests verify the superior accuracy of PCA compared to FFT and MEM. In addition, PCA gives the eigenvectors needed for vibrational mode assignments. Moreover, PCA is a very general technique for a variety of multivariate data series and therefore may be applied to MD, quantum MD, QM/MM trajectories, or even theoretically, Monte Carlo simulations to estimate optimal spectra.

Table 4.1. Comparison of vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) with results by standard quantum chemical calculation and experiment for isolated water, water dimer, and water molecules in liquid phase. PCA frequencies are averages over 12 trajectories.

Mode description	PCA ave [a]	PCA std dev	Standard Quant Chem	Experiment
H₂O				
Antisym stretch	3838	6	3850	3756 [31]
Sym stretch	3719	9	3728	3657 [31]
Bend	1719	12	1713	1593 [31]
(H₂O)₂				
Acceptor antisym stretch	3896	9	3912	3715-3745 [32-35]
Donor free O-H stretch	3881	9	3894	3714-3730 [32-35]
Acceptor sym stretch	3804	11	3814	3600-3638 [32-35]
Bridge O-H stretch	3683	54	3704	3530-3601 [32-35]
Acceptor H-O-H bend	1568	37	1641	1601 [37]
Donor H-O-H bend	1519	90	1660	1619,1611 [38-41]
Bridging H out-of-plane bend	402	134	645	520 [38-41]
Donor in-plane rock	296	102	375	243,320,290 [38-41]
Acceptor twist	205	93	154	??
Acceptor wag	181	64	156	??
O...O stretch	178	25	191	155,147,150 [38-41]
H-O...O-H torsion	153	53	130	??
H₂O (liq)				
Antisym stretch	3542	110		3446 [55]
Sym stretch	3438	107		3240 [55]
Bend	1693	81		1635 [55]

[a] Averages include data deemed statistically significant at the 99% confidence level by applying Dixon's Q test. [56]

Table 4.2. Comparison of frequencies calculated by using the fast Fourier transform (FFT), maximum entropy method (MEM), and principal component analysis (PCA) with analytical solutions to harmonic, quartic, and Langevin oscillators or experimental results for a single water molecule in isolation and in liquid phase.

	FFT	MEM	PCA	Analytical or experiment
Harmonic [a]	5.008	5.001	5.001	5.000
Quartic [b]	5.454	5.435	5.447	5.446 [51]
Langevin [c]	-- [d]	4.995	4.926	4.899 [52]
Water (cm ⁻¹)	1718	1724	1714	1593 [31]
	3803	3804	3716	3657 [31]
	3929	3978	3824	3756 [31]
Liquid water [e] (cm ⁻¹)	1970	1958	1793	1635 [55]
	3925	3798	3305	3240 [55]
	4120	4046	3337	3446 [55]

[a] The equation of motion of a one-dimensional harmonic oscillator is $mx'' + kx = 0$. The parameters and initial conditions used are $m = 0.5$, $k = 12.5$, $x(0) = -1.0$, and $x'(0) = 0.0$.

[b] The model of a quartic oscillator is defined by $mx'' + kx(1 + \beta x^2) = 0$ with $m = 0.5$, $k = 12.5$, $\beta = 0.25$, $x(0) = -1.0$, and $x'(0) = 0.0$.

[c] A one-dimensional Langevin equation $mx'' + cx' + kx = R(t)$ with $m = 0.5$, $k = 12.5$, $c = 1.0$, $x(0) = -1.0$, and $x'(0) = 1.0$ is used as the numerical model. $R(t)$ is the random number term with a Gaussian distribution.

[d] The Langevin oscillator is damped quickly and the random force becomes dominant before a sufficiently long trajectory can be generated for the FFT method to give a well-defined peak.

[e] Average frequencies calculated by PCA of 12 trajectories for isolated water and liquid water can be found in Table 4.1. With one exception, FFT- and MEM-derived frequencies for the tests differ from experiment more than PCA frequencies differ from experiment by at least three times the standard deviations. Large standard deviations for liquid water in Table 4.1 reflect inhomogeneous spectral broadening.

4.3. Methodology.

4.3.1. Theory of principal component analysis and its application to molecular vibrations.

Principal component analysis, a commonly used dimensionality reduction technique, extracts the “principal components” from an immense amount of data, discards the insignificant components with minimized truncation error in order to achieve the desirable goal of data compression. [10,57-60]

Consider that a set of D measurements made for a system at t discrete time intervals forms the original data matrix \mathbf{X} by setting the D measurements at every time as the column vectors $\mathbf{x}(t)$:

$$\mathbf{X} = [\mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(t) \dots], \text{ in which} \quad (4.1)$$
$$\mathbf{x}^T(t) = [x_1(t) - \langle x_1 \rangle, x_2(t) - \langle x_2 \rangle, \dots, x_D(t) - \langle x_D \rangle]$$

It should be noted that in this expression the vector elements are actually the displacements of each individual measurement away from the time averages.

Expanding the vector \mathbf{x} as a linear combination of any set of orthonormal basis vectors, for instance, $\Phi = [\phi_1 \ \phi_2 \ \dots \ \phi_i \ \dots \ \phi_D]$ with $\phi_i^T = [\phi_{i1} \ \phi_{i2} \ \dots \ \phi_{iD}]$, the following is true:

$$\mathbf{x} = y_1 \phi_1 + y_2 \phi_2 + \dots + y_D \phi_D = \sum_{i=1}^D y_i \phi_i, \quad (4.2)$$

the y 's are the coefficients which also form a matrix \mathbf{Y} for all time t , whose column vector \mathbf{y} is related to \mathbf{x} expressed in matrix representation:

$$\mathbf{x} = \Phi \mathbf{y} \quad \mathbf{X} = \Phi \mathbf{Y} \quad (4.3a)$$

$$\mathbf{y} = \mathbf{\Phi}^T \mathbf{x} \quad \mathbf{Y} = \mathbf{\Phi}^T \mathbf{X}. \quad (4.3b)$$

Eq. (4.3a, 4.3b) can also be considered to represent an orthogonal coordinate transformation from one set of coordinates to another where each row vector $\boldsymbol{\phi}_i^T$ is the new coordinate axes' projections on the old coordinate axes. This coordinate transformation scheme is a very powerful technique to transform interesting subjects to a coordinate system that is easier to understand.

Often the size of the original data matrix \mathbf{X} is extremely large so that a reduction in size after the transformation is desirable. One method of doing this is to truncate the linear combination in eq. (4.2) after d terms (usually $d \ll D$), to get a new vector $\hat{\mathbf{x}}$. The mean-square error ε between the original matrix and the truncated data matrix, averaged over all time t , is

$$\begin{aligned} \varepsilon &= \left\langle (\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T \right\rangle \\ &= \left\langle \left(\sum_{i=1}^D y_i \boldsymbol{\phi}_i - \sum_{i=1}^d y_i \boldsymbol{\phi}_i \right) \left(\sum_{i=1}^D y_i \boldsymbol{\phi}_i - \sum_{i=1}^d y_i \boldsymbol{\phi}_i \right)^T \right\rangle \\ &= \left\langle \sum_{i=d+1}^D y_i^2 \right\rangle = \frac{1}{t} \mathbf{Y} \mathbf{Y}^T \end{aligned} \quad (4.4)$$

Using eq. (4.3b), this error is expressed as

$$\varepsilon = \frac{1}{t} \mathbf{\Phi}^T \mathbf{X} \mathbf{X}^T \mathbf{\Phi} = \mathbf{\Phi}^T \mathbf{R}_x \mathbf{\Phi}, \quad (4.5)$$

where $\mathbf{R}_x = (1/t) \mathbf{X} \mathbf{X}^T$ is the time-averaged covariance matrix (for simplicity the term time-averaged will be omitted hereafter) of the original time series in which the elements of \mathbf{R}_x are:

$$R_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle. \quad (4.6)$$

Now the problem becomes to search a special set of $\boldsymbol{\phi}$'s to minimize the error ε . Referring to the proof found elsewhere [30], the $\boldsymbol{\phi}$'s are the eigenvectors of \mathbf{R}_x

$$\mathbf{R}_x \boldsymbol{\phi}_i = \lambda_i \boldsymbol{\phi}_i, \quad (4.7)$$

where λ 's are the associated eigenvalues. If the $\boldsymbol{\phi}$'s are lined up in a sequence so that their associated eigenvalues are in a descending order, then the first eigenvector defines a direction in which the total variances in \mathbf{X} are maximally accounted for since the error ε is minimized if the truncation is made after the first term in eq. (4.2). In a similar fashion the second eigenvector accounts for the maximum remaining variance in a direction orthogonal to the first eigenvector (because all eigenvectors are orthogonal) and so on.

Another interesting feature of this coordinate transformation can be seen when the covariance matrix for the transformed matrix \mathbf{Y} is constructed and converted as

$$\mathbf{R}_y = \mathbf{Y}\mathbf{Y}^T = \boldsymbol{\Phi}^T \mathbf{X}\mathbf{X}^T \boldsymbol{\Phi} = \boldsymbol{\Phi}^T \mathbf{R}_x \boldsymbol{\Phi}. \quad (4.8)$$

Since $\boldsymbol{\Phi}$ is the eigenvector matrix that diagonalizes \mathbf{R}_x , eq. (4.8) implies a diagonalized covariance matrix for \mathbf{Y} . This result can be further interpreted that the row vectors of \mathbf{Y} are totally uncorrelated since all $Y_{ij} = 0$ when $i \neq j$.

Molecular trajectories given by computer simulations are no different than a group of $3N$ (N is the total number of atoms in the system) atomic positions at a series of time steps and therefore can be treated by PCA. Usually a trajectory is generated in Cartesian coordinates. PCA will find an orthogonal transformation matrix, $\boldsymbol{\Phi}$, that

transforms the trajectory from Cartesian to a new set of coordinates. Eq. (4.8) states that the variances in the newly-transformed trajectory are completely uncorrelated. In other words, the motions expressed in the new coordinates are orthogonal to each other. Thus the column vectors of this transformation matrix can be thought as a set of vibrational modes and eq. (4.3b) gives the way to do the transformation. As explained in the chapter introduction, this new set of vibrational modes may or may not be the same as given by normal mode analysis depending on the specific shape of the potential energy surface sampled by the trajectory.

In order to relate the covariance matrix \mathbf{R}_x in principal component analysis to the effective force constant matrix in quasiharmonic analysis, an equilibrium probability distribution of the atomic positions, $P(\mathbf{X})$, of the molecular trajectory can be expressed as a normalized multivariate Gaussian distribution:

$$P(\mathbf{X}) = \frac{1}{(2\pi)^{3N/2}} |\mathbf{R}_x|^{1/2} \exp\left(-\frac{1}{2} \mathbf{X}^T \mathbf{R}_x^{-1} \mathbf{X}\right), \quad (4.9)$$

where $|\mathbf{R}_x|$ is the determinant of the covariance matrix of the mass-weighted atomic positions \mathbf{X} . The potential energy, $V(\mathbf{X})$, sampled by this particular trajectory determines the configurational partition function from the probability distribution:

$$P(\mathbf{X}) = \frac{1}{(2\pi)^{3N/2}} |\mathbf{R}_x|^{1/2} \exp(-V(\mathbf{X})/k_B T), \quad (4.10)$$

where k_B and T are the Boltzmann constant and temperature, respectively. Next, truncating the potential energy at the second order,

$$V(\mathbf{X}) = \frac{1}{2} \mathbf{X}^T \mathbf{F}_{\text{eff}} \mathbf{X}, \quad (4.11)$$

and combining eq. (4.11) with eq. (4.10) and comparing with eq. (4.9) gives the effective force constant matrix constructed in quasiharmonic analysis is as

$$\mathbf{F}_{\text{eff}} = k_B T \mathbf{R}_x^{-1}. \quad (4.12)$$

In a similar fashion as solving the eigenproblem of the force constant matrix in normal mode analysis, the frequencies and modes can be calculated from the eigenvalues and eigenvectors of this effective force constant matrix.

4.3.2. A modification in calculating vibrational frequencies.

Because of the heavy computational load required for calculating the inverse matrix using eq. (4.12) when N , the number of atoms in the system, becomes large, this step is skipped and the covariance matrix itself is directly diagonalized. The reason for doing so is that the set of eigenvectors of \mathbf{F}_{eff} is orthonormal and each eigenvector forms a column of Φ that diagonalizes \mathbf{F}_{eff} ,

$$\Phi^T \mathbf{F}_{\text{eff}} \Phi = \Lambda \quad (4.13)$$

where Λ is a diagonal matrix in which the eigenvalues λ lay along the diagonal.

Inverting eq. (4.13) follows as

$$\Lambda^{-1} = (\Phi^T \mathbf{F}_{\text{eff}} \Phi)^{-1} = \Phi^{-1} \mathbf{F}_{\text{eff}}^{-1} (\Phi^T)^{-1} = \frac{1}{k_B T} \Phi^T \mathbf{R}_x \Phi = \frac{1}{k_B T} \Lambda'. \quad (4.14)$$

The property that the inverse of an orthonormal matrix is its transpose is used here. Since Λ^{-1} is a diagonal matrix (whose diagonal terms are just λ^{-1}), then the resultant matrix, Λ' , should also be diagonal, with the diagonal elements different than those in Λ^{-1} only by a factor of $1/k_B T$. For the same reason, the expression second from the

right in eq. (4.14) should also be diagonal, in other words, Φ diagonalizes \mathbf{R}_x as well.

Eq. (4.14) indicates that a matrix and its inverse have a common set of eigenvectors.

Therefore the covariance matrix \mathbf{R}_x can be directly diagonalized without calculating its inverse and the eigenvalues of \mathbf{R}_x , λ'_i , are used to calculate the eigenvalues of the effective force constant matrix by

$$\lambda_i = \frac{k_B T}{\lambda'_i}. \quad (4.15)$$

Eq. (4.6) and (4.12) compose the basis of the methodology of PCA or quasiharmonic analysis for molecular vibrations; however, in practical use of these equations a number of unforeseen challenges have been overcome to allow this new method to show its ability to calculate accurate vibrational frequencies. These challenges will be addressed and corresponding modifications will be described in the methodology section following.

The first issue to resolve is that for a molecular trajectory with sampling time often limited in picoseconds or shorter, in calculating frequencies by eq. (4.12) or (4.15) the conversion factor $k_B T$ is not commonly suitable for all of the vibrational modes. One important assumption of using eq. (4.9) to derive eq. (4.12) is that the entire potential energy surface for the system is sampled completely and evenly. With the time scale at the macroscopic level for real systems, this assumption is usually satisfied and the factor $k_B T$ (as the total vibrational energy) is good for every vibrational mode. However, this macroscopic time scale (even though it can be as short as sub- μ s for simpler systems) is still well above the limitation of computer simulation techniques.

In a typical computer simulated molecular trajectory on the order of $ps \sim ns$, the amount of energy is probably distributed unevenly among the vibrational modes. Therefore $k_B T$ is not necessarily the exact amount of energy associated with every mode for this short simulation time. The previously introduced eq. (4.15) needs to be modified to reflect the specific amount of vibrational energy possessed by each individual mode.

Consider a one-dimensional harmonic oscillator, whose total vibrational energy is given as the sum of kinetic and potential energy terms:

$$E_{vib} = \frac{1}{2} \mu \dot{\mathbf{q}}^2 + \frac{1}{2} \omega^2 \mathbf{q}^2 \quad (4.16)$$

where \mathbf{q} is the vibrational coordinate, μ is the reduced mass, and ω is the frequency of the oscillator. During each cycle of the oscillation the kinetic and potential parts keep exchanging energy with each other but their time averages are the same and each accounts for exactly one half of the total vibrational energy. Therefore, either the time average of $\mu \dot{\mathbf{q}}^2$ or $\omega^2 \mathbf{q}^2$ can be used to specify E_{vib} .

PCA or quasiharmonic analysis actually decomposes the molecular motions into a set of separate one-dimensional quasiharmonic oscillators. In the i^{th} oscillators, the total vibrational energy can be calculated as $\langle \mu_i \dot{\mathbf{q}}_i^2 \rangle$ instead of $k_B T$ in eq. (4.15). In practice, this term is calculated by projecting the mass-weighted Cartesian velocities ($\sqrt{m_j} \dot{x}_j, j = 1, 2, \dots, 3N$) onto the i^{th} PCA mode and then summing the squares over all time steps:

$$E_{vib} = \langle \mu_i \dot{\mathbf{q}}_i^2 \rangle = \left\langle \left(\sum_j^{3N} \sqrt{m_j} \dot{x}_j \phi_{ji} \right)^2 \right\rangle \quad (4.17)$$

with ϕ_{ji} is the j^{th} element in the i^{th} eigenvector ϕ_i in eq. (4.7).

4.3.3. Translation and rotation present in molecular trajectories.

It has been found in testing principal component analysis of molecular vibrations that the translational and rotational motions of the system can affect the calculated modes and frequencies. In the normal mode analysis the force constant matrix is set up based on a single low energy-optimized structure without actual movements.

However, PCA treats a series of molecular structures generated by a simulation in which the molecule is subject to translation and rotation. The six external motions, 3 translations and 3 rotations, are very likely present in atomic displacements. When the covariance matrix of the original trajectory is constructed, the equilibrium or time-averaged atomic positions can be shifted and altered if the system as a whole undergoes a movement. In Cartesian coordinates the average position for each atom is usually determined by taking a time average of each of the coordinate variables. The shifting in the average positions caused by translation actually plays no role in affecting the covariances because the translational motions uniformly shift the positions in Cartesian space. However, if the same Cartesian space is used, the rotational motions do not change all atomic positions uniformly and these changes are structure dependent. If the rotational motions present in a trajectory are significant, the time-averaged positions in Cartesian space will be altered unevenly and therefore the

covariances (which depend on the displacements from the averages) will also be changed unevenly. One way of dealing with this problem used in this chapter is to remove both translational and rotational motions before calculating the averages and covariances.

It is well known that the whole motions of a system can be approximately separated into three groups, translation, rotation, and internal vibration when a carefully chosen coordinate system is used. [14] This coordinate system includes the three space-fixed Cartesian coordinate axes on the center of mass of the system, the three body-fixed rotational axes which are coincident with the principal axes of inertia of the undistorted structure, and $3N-6$ vibrational coordinate axes in the body-fixed system specifying the vibrational modes. By using this consideration, the task of removing translation and rotation from vibration involves searching for these special coordinate systems from the original trajectory, overlaying them for all time steps, and then only the pure “vibrational” motions will be kept.

The translational motions in the space-fixed Cartesian coordinates can first be separated from the other motions. Consider tracing an arbitrary point in the molecule in the space-fixed coordinates at each time step, then set three orthogonal axes on this point. There will be a series of this kind of coordinates in the original Cartesian space and they are moving along with the trajectory. But if the point is viewed in each of these moving coordinates, then it does not have any net translational motions. In other words, the point is translation free. A conventional choice of this arbitrary point is the center of mass of the molecule because it is convenient for setting the rotational and

vibrational coordinates later. The condition of setting the center of mass as the origin of the moving coordinate system is given as

$$\sum_{n=1}^N m_n \mathbf{r}'_n = 0 \quad (4.18)$$

where m is the atomic mass and \mathbf{r}'_n is the atomic coordinate in this new “center-of-mass” coordinate system. Each \mathbf{r}'_n can be obtained by subtracting the center-of-mass coordinates, \mathbf{R}_{com} , from atomic coordinates \mathbf{r} , all expressed in the original space-fixed coordinates:

$$\mathbf{r}'_n = \mathbf{r}_n - \mathbf{R}_{com} = \mathbf{r}_n - \frac{\sum_{n=1}^N m_n \mathbf{r}_n}{\sum_{n=1}^N m_n} \quad (4.19)$$

Putting this equation back in eq. (4.18) one can see the condition is satisfied. In a PCA calculation, the procedure to remove translation is therefore simply to use eq. (4.19) to calculate \mathbf{r}'_n and replace the original \mathbf{r}_n by \mathbf{r}'_n at each time step.

The removal of rotation can be achieved in a similar fashion by setting up a series of “body-fixed” coordinate systems (they are moving relative to the original space-fixed coordinate system) at each time interval in a way so that the orientation of the system is unchanged in time. Then overlaying these coordinate axes will eliminate the rotational motions. With the center of mass located at the origin of this set of body-fixed coordinate axes, a set of principal axes of inertia serves as a good choice to specify the directions of rotation. The moment of inertia of an atom with mass m located at \mathbf{r} (r_i , $i=1,2,3$ for x, y, and z axis in Cartesian space) from the center of mass with respect to i and j axes is defined by

$$I_{ij} = m(\delta_{ij}r_k r_k - r_i r_j) \quad i, j, k = 1, 2, 3. \quad (4.20)$$

The total moment of inertia of the system can form a matrix \mathbf{I} where its entries are $\sum I_{ij}$ (sum over N atoms). [61] The eigenvector matrix (the column vectors are eigenvectors of \mathbf{I}) will diagonalize matrix \mathbf{I} to make the off-diagonal entries, the products of inertia, become all zeroes. These eigenvectors are called principal axes which completely separate the moment of inertia along with these three axes. This method works perfectly for rigid bodies (e.g. the internal distances between particles remain constant). However, for the molecules in which the internal motions (vibration) are the greatest interest, the principal axes of inertia deviate from the true rotation axes depending on the magnitude of the vibrational motions. When the molecule undergoes a small vibration a set of conditions similar to eq. (4.18) has been given by Eckart for specifying the true rotation axes, [62]

$$\sum_{n=1}^N m_n \mathbf{a}_n \times \mathbf{r}_n = 0 \quad (4.21)$$

where \mathbf{a} is the equilibrium or average positions of the atoms. In the same paper Eckart presented a procedure to find the rotation axes starting with constructing a matrix \mathbf{B} :

$$B_{ij} = \sum_{n=1}^N m_n a_{ni} r_{nj} \quad i, j = 1, 2, 3, \quad (4.22)$$

then searching for an orthogonal transformation \mathbf{T} to transform \mathbf{B} to a symmetric matrix \mathbf{S}

$$\mathbf{BT} = \mathbf{S}. \quad (4.23)$$

If \mathbf{S} is symmetric then the entries across the diagonal in \mathbf{S} are the same:

$$\sum_{n=1}^N m_n \alpha_{ni} r'_{nj} = \sum_{n=1}^N m_n \alpha_{nj} r'_{ni} \quad (4.24)$$

and this will satisfy eq. (4.21). Therefore the new coordinate axes described by \mathbf{T} are the true rotation axes. In the same reference the details of finding \mathbf{S} were suggested and \mathbf{T} could be obtained by multiplying the inverse matrix \mathbf{B}^{-1} on the both sides of eq. (4.23).

One problem with this procedure of finding the rotation axes that has been pointed out is that the inverse matrix of \mathbf{B} does not always exist. This singularity problem can be solved by the well-suited technique, singular value decomposition (SVD), and a modification to Eckart's procedure given by Pickett and Strauss is used in this chapter. [63]

4.3.4. *Frequencies and modes of molecules with flexible structures.*

Once the translational and rotational motions are removed from a trajectory, the covariance matrix that is a measure of the pure vibrational motions is constructed for solving frequencies and modes. When the structure of a molecule shows some flexibility, for instance, the dihedral angle motions of the main chain atoms or the distorted motions of the side groups are present, some atoms particularly far away from the center of mass of the molecule can travel in a large area of space. The assumption of small vibrations is probably not valid for this situation. In practice it has been found that if these motions (often related to low frequencies) are present in the trajectory, PCA likely gives a strong mixing of these motions with the usual high-

frequency bending and stretching motions. The mixing also results in inaccurate frequencies (bending and stretching modes that are lower than they should be). This problem can be explained very similarly to the translation's and rotation's effect on PCA discussed in the previous section. In the molecules with flexible structures, the vibrations associated with the large-scale low-frequency motions can be considered as combinations of translations and rotations of smaller fragments of the molecules. For example, the C-C bond stretch in the ethane molecule is a perfect internal vibration but it can also be seen as a combination of translations of the two methyl groups in opposite directions. The twisting motion around the same C-C bond is a dihedral angle-changing motion of H-C-C-H groups. It can also be seen as the relative rotation of the two methyl groups along their common C_3 axes. This consideration implies that when these "internal" translations and rotations are present in the trajectory, the determination of average structures and covariances would be affected in the same manner as they are affected by the true external translations and rotations. One way to eliminate these effects is to treat the fragments affected by these "internal" translational and rotational motions as separated smaller "molecules".

Prior to the determination of average structures and covariances, the entire molecule needs to be divided into smaller fragments. For each fragment, the proper set of coordinates is determined individually in the same way as described in the previous section. However, a further calculation is necessary in this stage: the specific atomic displacements caused by these "internal" translations and rotations of the fragments are needed since they will later be a part of the internal vibrations for the whole

molecule. For the translational motions of an individual fragment, again the center of mass of the fragment can be the origin of the body-fixed coordinate system. Since in a Cartesian coordinate system translation shifts the atomic positions evenly, the displacements of each the atoms caused by translation are equal to the displacements of the center of mass:

$$\Delta \mathbf{r}_{trans} = \Delta \mathbf{R}_{com}. \quad (4.25)$$

The next step is to find the atomic displacements exclusively associated with the rotational motions. For a finite rotation it has been proven that the change in positions cannot be represented by a single vector (e.g. the displacement between a position vector before and after the rotation). [64] It is obvious that this displacement is not perpendicular to either the initial or the end position vectors. However, when the time interval is set small enough in the trajectory that the rotation in this amount of time is infinitesimal, the displacements are close enough to represent the rotation. The rotation matrix \mathbf{T} (see eq. (4.23)) for such an infinitesimal rotation is nearly diagonalized with very small (antisymmetric) off-diagonal entries:

$$\mathbf{T} = \begin{pmatrix} 1 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 1 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 1 \end{pmatrix}. \quad (4.26)$$

Then the displacements caused by this infinitesimal rotation are the differences between the original \mathbf{r} and transformed vector \mathbf{r}' ,

$$\Delta \mathbf{r}_{rot} = \mathbf{r}' - \mathbf{r} = \mathbf{T} \mathbf{r} - \mathbf{r}. \quad (4.27)$$

The actual matrix multiplication can be carried out to show that if a vector $d\Omega$ is set to have components $d\Omega_1$, $d\Omega_2$, and $d\Omega_3$, $\Delta\mathbf{r}_{rot}$ is

$$\Delta\mathbf{r}_{rot} = \mathbf{r} \times d\Omega . \quad (4.28)$$

After the displacements associated with the “internal” translations and rotations of the fragment are calculated and stored, they can be subtracted from the atomic positions and the remaining are pure vibrations within the fragment. Once the average structure is determined for this fragment using the translation and rotation-free trajectory, a covariance matrix is constructed and diagonalized for eigenvectors and eigenvalues. Eq. (4.3b) is then used here again to transform the Cartesian displacements to a vibrational coordinate system based on the body-fixed axes on this fragment. In total $3N_i$ (N_i is the number of atoms in fragment i) dimensions of this transformed displacement matrix, six of them should be zeroes because they are associated with three translations and three rotations which have already been removed. These matrix elements need to be replaced by dot products of the stored $\Delta\mathbf{r}$'s with the six vectors specifying translations and rotations. The whole procedure is merely the same as introduced earlier with the only difference being that it is for smaller fragments and has accounted for the “internal” translations and rotations.

If this procedure is repeated for each of the fragments, the transformed matrices can be combined to form a $3N \times t$ block diagonalized matrix in the following way,

$$\Psi = \begin{pmatrix} \Phi_1 & \cdots & \Phi_{3N_1} & 0 & & 0 \\ & 0 & & \Phi'_1 & \cdots & \Phi'_{3N_2} & 0 & \cdots \\ & 0 & & & & 0 & \Phi''_1 & \cdots & \Phi''_{3N_3} \\ & & & \cdots & & & & & \cdots \end{pmatrix}, \quad (4.29)$$

and then do the transformation for the whole molecule. In this equation $\boldsymbol{\phi}$'s are eigenvectors for the individual fragment covariance matrices. They take the $3N_1 \times 3N_1$, $3N_2 \times 3N_2$, ... blocks along the diagonal where N_1, N_2, \dots are the number of atoms in the 1st, 2nd, ... fragments, respectively. The rest of the matrix is padded with zeroes. By examining the column vectors of this $3N \times 3N$ matrix, $\boldsymbol{\Psi}$, one finds that it is orthonormal and therefore defines a unitary transformation for the initial whole molecule's covariance matrix \mathbf{R}_x :

$$\mathbf{P}_x = \boldsymbol{\Psi}^T \mathbf{R}_x \boldsymbol{\Psi}. \quad (4.30)$$

The resultant matrix \mathbf{P}_x , now transformed in a fragment-based coordinate system, should be nearly block diagonalized since the covariances between different fragments should be small. This treatment shown in eq. (4.30) is in principle very similar to the symmetry-aided treatment in the conventional normal mode analysis in which the force constant matrix is also block diagonalized first by a unitary transformation. [14] The column vectors of this unitary transformation matrix are linear combinations of internal coordinates for the whole molecule according to symmetry operations.

The final step follows the same step as the regular PCA procedure. \mathbf{P}_x , the modified covariance matrix will be diagonalized to give eigenvalues (frequencies) and eigenvectors (modes) for the whole molecule. It needs to be noted that the modes at this time are linear combinations of the column vectors of $\boldsymbol{\Psi}$, which are the translational, rotational, and vibrational modes of the fragments. The transformation from these fragment-based modes to the modes expressed in Cartesian coordinate system is straightforward.

4.4. Computational details.

Test cases include isolated water, methane, water dimer, and ethane molecules. Their initial structures for QM/MM simulations were randomly selected from previously generated trajectories of these molecules equilibrated at constant temperature 300K by molecular dynamics. A constant energy QM/MM trajectory was generated for each test molecule and structure using the CHARMM/GAMESS combination, [20] for 1 ps at a 0.5 fs time interval. The velocity assignment was only performed once at the beginning of the simulations at 300K. The entire molecules were defined as the QM region using the B3LYP hybrid Hartree-Fock/density functional method with a 6-31G(d) basis set, except for the water dimer test case, in which the 6-311++G(2d,2p) basis set was used because it reportedly gives a qualitatively correct geometry. [47] After the trajectories were generated, the plain PCA method and PCA with modifications to calculate vibrational energy in specific modes and to remove translation and rotation were applied to calculate vibrational frequencies and modes for all test molecules. We shall call the procedures PCA, PCA-I (energy corrected), and PCA-II (energy corrected, with translation and rotation removed). The new PCA procedure of treating the molecules with flexible structures in individual fragments, PCA-III was also used for the water dimer and ethane tests. All results from PCA methods were then compared with standard quantum chemical calculation results using the same method/basis set, in which the frequencies and modes were solved by diagonalizing the force constant matrix based on the optimized structures. Finally, experimental results were provided for further comparison.

4.5. Results and discussion.

Table 4.3 shows the calculated frequencies for isolated water molecules, calculated by using various methods, and experimental values. The first columns of the table, water1 and water2 are the frequencies of two sample water trajectories by the plain PCA method without any modifications as described in eq. (4.9-15). The frequencies are far from the experiment and are often more than one order of magnitude lower. This suggests that the original PCA formalism fails to predict the frequencies even for the simplest molecules. Two water tests with different initial structures are presented here to show that the amount of energy distributed in each vibration is different. For instance, the frequency for O-H antisymmetric stretch of water 1 is much greater than the frequency of the same mode of water 2 even though they were both started at 300K. This means that the universal energy factor $k_B T$ is not appropriate for calculating vibrational frequencies from simulations of limited length. The second columns of the two water tests are the frequencies calculated using eq. (4.15) and (4.17) to correct the energy term (the PCA-I method) but without removing the translation and rotation from the same trajectories. The frequencies are improved from the original PCA. The antisymmetric stretch of water 1 and all three frequencies of water 2 are reasonably close to experiment but the H-O-H bend and O-H symmetric stretch of water 1 are still very low. These two modes are found to be highly mixed with rotational motions and the other four are not. The frequencies calculated using eq. (4.15) and (4.17) are thus in better agreement with experiment if there is no rotation-vibration mixing. The third columns of frequencies are given by PCA with energy

term corrected and translation and rotation removed (PCA-II method). The results are considerably improved for each of the vibrational modes. The frequencies for the two stretches are consistently close to experiment and for the bend are 7% higher than experiment but still better than the frequencies given by the conventional quantum chemical calculations. Therefore the suggested modifications (Section 4.3.2 and 4.3.3) greatly improve PCA's application to molecular vibrations.

Table 4.4 presents the same tests on a different system, methane (two sample methane trajectories, methane 1 and 2), which is more complicated than water. Again as shown in the water tests, the plain PCA fails to calculate a consistent set of accurate frequencies for methane. When eq. (4.17) is used to correct the specific amounts of energy associated with individual modes, the removal of translation and rotation can make a difference for the frequencies of methane. Table 4.4 shows that out of total 9 vibrations there is only 1 frequency (asymmetric deformation, 1534 cm^{-1}) for methane 1 and 1 frequency (C-H symmetric stretch, 2973 cm^{-1}) for methane 2 by PCA-I significantly closer to experiment than PCA-II (1627 cm^{-1} and 3011 cm^{-1}). There are 2 frequencies (symmetric deformation, 1335 cm^{-1} and 1364 cm^{-1} and asymmetric deformation 1491 cm^{-1} and 1585 cm^{-1}) for methane 1 and 3 frequencies (symmetric deformation, 1370 cm^{-1} and 1371 cm^{-1} , asymmetric deformation, 1587 cm^{-1} and 1555 cm^{-1} , and C-H asymmetric stretch, 3097 cm^{-1} and 3109 cm^{-1}) for methane 2 that are consistent regardless of the presence of translation and rotation. However, the translation and rotation in trajectories strongly affects the rest of the frequencies, either giving high (for bend-type modes) or low (for stretch-type modes) wavenumbers.

Again, the modified PCA method calculates frequencies comparable and often closer to experiment than the standard quantum chemical method. The observation here confirms that the suggested modifications to the original PCA formalism improve the accuracy of the predicted frequencies.

One interesting feature of the PCA derived frequencies can be seen in methane tests. Methane is a highly symmetric molecule and some vibrational modes are triply degenerate. This degeneracy is reflected in the conventional quantum chemical calculations. But it is not displayed by the PCA calculation. Instead the frequencies are approximately grouped as the same pattern shown in quantum chemical calculation and experiment, with fine differences between them. This can be explained as when the molecule undergoes vibrational motions, the symmetry of the instantaneous structures is lowered, thus breaking the degeneracy. The quantum chemical calculation is based on the energy-optimized structure of methane which retains the high level symmetry resulting in degenerate modes and frequencies. In experiment there is a large number of molecules being observed, the experimental frequencies are actually the averages over all possible configurations of methane molecules and therefore show the expected degeneracy in a statistical sense. It is expected that the degeneracy in the PCA-calculated frequencies will be improved as the trajectories become longer. In one group member's work, this is actually verified from the results of long simulations (up to 50 ps) for ammonia and methane. [65]

In a similar fashion to the effect on PCA derived frequencies of the external whole-molecule motions, the "internal" translations and rotations of fragments of a

flexible molecule can also spoil the construction of the covariance matrix. In order to test the modified procedure proposed in Section II. 4, two test systems are selected that would in effect “double” the structure of the previously tested water and methane molecules. Table 4.5 and 4.6 show the results of the water dimer system and the ethane molecule, respectively.

The total vibrational motions of the water dimer are in two groups, the high-frequency motions of each water’s bend and stretches, and the low-frequency related combinations of the “internal” translations and rotations of the two water parts. Therefore it is natural to set each water unit as a fragment and calculate the translations, rotations, and vibrations separately at each time step. The final PCA then can be performed to regroup the two sets of motions together to find frequencies and modes. Table 4.5 shows the calculated results using PCA with only the global translation and rotation removed (PCA-II) and the PCA with the fragment setup (PCA-III) methods as described in Section 4.3.4. In the first column of the table, PCA-II, there are only 5 modes that can be clearly assigned and two of them (the bridge O-H stretch and donor free O-H stretch) are very low compared to experiment. The frequencies of other modes are simply listed in an ascending order in the rest of the entries, and it should be noted that these strongly mixed modes do not correspond to the listed modes. The frequencies in the second column, PCA-III, however, are much closer to experiment than those in the first column. Indeed, the high frequency bending and stretching modes given by PCA-III are all better than the conventional quantum chemical calculation. At the low frequency end, PCA-III gives two closer frequencies

to experiment for the O···O stretch and donor in-plane rock, one poorer for the donor out-of-plane bend, and the rest are not comparable due to the lack of experimental data. In general, the improvement in predicting frequencies for the water dimer by treating the system as fragments is obvious in Table 4.5.

Table 4.6 presents the same treatment of the ethane molecule which can be considered as two methyl groups linked by the C-C bond. Each methyl group has its own set of translational, rotational, and vibrational motions. Then the two sets of motions are regrouped by PCA to find the cooperative motions between the two methyl group fragments and related frequencies. In the table, the first column is the frequencies by the PCA-II method. For the high frequency C-H stretch modes, this method tends to give very low frequencies except one asymmetric stretch (3005 cm^{-1}). For the same stretches the results with the extra dividing step (PCA-III method) are in very good agreement with experiment and all are better than conventional quantum chemical calculations. For the motions based on H-C-H bends, e.g. the deformations, PCA-III again predicts the frequency better than PCA-II and conventional quantum chemical methods. Next, for the motions considered as combinations of the “internal” translations and rotations, the frequencies of the methyl group rocking motions by PCA-III are closer to experiment than PCA-II. The frequencies of the C-C twist and C-C stretch by both PCA methods are approximately same. With three exceptions, the C-C twist observed at 289 cm^{-1} (PCA 260 cm^{-1} , quantum chemical 315 cm^{-1}), a methyl rock at 822 cm^{-1} (PCA 1075 cm^{-1} , quantum chemical 832 cm^{-1}) and a methyl antisymmetric deformation at 1469 cm^{-1} (PCA 1589 cm^{-1} , quantum chemical

1538 cm^{-1}), results of PCA methods are closer to experiment than quantum chemical results. From the comparison presented in Table 4.6 it is shown that the PCA-III method modified from the original PCA, in which the fragments of flexible molecules are treated separately prior to the construction of covariance matrices, has the ability to predict vibrational frequencies accurately.

Table 4.3. Vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) of two QM/MM simulation trajectories of isolated water molecules (Water 1 and Water 2) with various modifications. The frequencies are compared with standard quantum chemistry and experiment results. Simulation details are described in the text.

	PCA of Water 1			PCA of Water 2			Std. Quant.	Exp. [31]
	PCA ^a	PCA-I ^b	PCA-II ^c	PCA ^a	PCA-I ^b	PCA-II ^c		
bend	120	482 ^d	1704	55	1701	1701	1713	1593
sym. str.	115	146 ^d	3663	146	3633	3655	3728	3657
antisym. str.	1855	3335	3759	328	3659	3763	3850	3756

^a PCA without energy correction (see the text). ^b PCA without removing translation and rotation from the trajectory. ^c PCA with energy correction and translation and rotation removed. ^d These modes are highly mixed with rotation modes.

Table 4.4. Vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) of two QM/MM simulation trajectories of isolated methane molecules with various modifications as the same presented in Table 4.3.

	PCA of Methane 1			PCA of Methane 2			Std. Quant.	Exp. [66]
	PCA ^a	PCA-I ^b	PCA-II ^c	PCA ^a	PCA-I ^b	PCA-II ^c		
sym def.	149	1335	1364	46	1370	1371	1374	1306
sym. def.	29	1844 ^d	1372	1116	1698	1402	1374	1306
dym. def.	447	2009	1421	741	1891	1422	1374	1306
asym. def.	44	1491	1585	601	1587	1555	1594	1534
asym. def.	255	1534	1627	63	1700	1587	1594	1534
C-H sym. str.	179	2694	2989	82	2973	3011	3051	2917
C-H asym. str.	762	1763	3006	2440	1799	2806	3161	3019
C-H asym. str.	203	1887	3110	853	2024	3058	3161	3019
C-H asym. str.	73	1934	3113	127	3097	3109	3161	3019

^{a-c} Same as in Table 4.3. ^d The deformation modes are highly mixed with C-H stretch modes without removing translational and rotational motions.

Table 4.5. Vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) of a single QM/MM simulation (with B3LYP/6-311++G(2d,2p) method/basis) trajectory of isolated water dimers with and without dividing the whole flexible structure in parts (described in the text). The frequencies are compared with the treatment of standard quantum chemistry and experiment.

	PCA of water dimer		Standard Quant Chem	Exp. [32- 35,37-41]
	PCA-II ^a	PCA-III ^b		
O...O torsion	170	165	130	-
O...O str.	140	137	191	147-155
accept. wag	127	320	156	-
accept. twist	205 ^c	70	154	-
don. i. p. rock	252 ^c	303	375	243-290
don. o. p. bend	425 ^c	253	645	520
accept. HOH bend	506 ^c	1632	1641	1601
don. HOH bend	1021 ^c	1642	1660	1611-1619
bridge OH str.	2181	3673	3704	3530-3601
accept. sym. str.	1176 ^c	3738	3814	3600-3638
don. free OH str.	1392	3818	3894	3714-3730
accept. antis. str.	1589 ^c	3825	3912	3715-3745

^a Same as in Table 4.3 and 4.4. ^b Modified PCA with fragments setup. ^c These modes are strongly mixed with others and therefore cannot be assigned. The frequencies of these modes given by PCA-II are listed in an ascending order.

Table 4.6. Vibrational frequencies (cm^{-1}) calculated by principal component analysis (PCA) of a single QM/MM simulation trajectory of isolated ethane molecules with and without dividing the whole flexible structure in fragments (described in the text). The frequencies are compared with the treatment of standard quantum chemistry and experiment.

	PCA of ethane		Standard Quant Chem	Exp. [66]
	PCA-II ^a	PCA-III ^b		
CH ₃ twist	262	260	315	289
CH ₃ rock	862	860	832	822
CH ₃ rock	1217	1075	832	822
C-C str.	983	987	1009	995
CH ₃ rock	1210	1183	1235	1190
CH ₃ rock	1226	1207	1235	1190
CH ₃ sym. def.	1383	1416	1432	1379
CH ₃ sym. def.	1588	1426	1454	1388
CH ₃ antisym. def.	1546	1445	1531	1468
CH ₃ antisym. def.	1591	1507	1531	1468
CH ₃ antisym. def.	1591	1511	1538	1469
CH ₃ antisym. def.	1851	1589	1538	1469
C-H sym. str.	1918	2947	3046	2896
C-H sym. str.	1958	2991	3047	2954
C-H asym. str.	2390	3015	3098	2969
C-H asym. str.	2597	3040	3098	2969
C-H asym. str.	2615	3047	3122	2985
C-H asym. str.	3005	3059	3122	2985

^a Same as in Table 4.3 and 4.4. ^b Modified PCA with fragment setup.

4.6. Conclusions.

The formalism of principal component analysis (PCA) is introduced and its application to molecular vibrations is discussed in detail for determining average positions, constructing the covariance matrix, and solving the eigenproblem to give vibrational modes and frequencies. A proof of that the inverse covariance matrix from a trajectory effectively represents the force constant matrix is carried out. Compared with the conventional normal mode analysis, this method shows advantages of incorporating anharmonicity and saving computer time. The problems in accurately predicting frequencies in the actual PCA calculations are addressed and explained from the aspect of the energy associated with individual motions and the mixing between the external and internal motions. To each problem a modification in the application of PCA to analyze molecular trajectories is provided to improve PCA's performance. For the energy term to convert the eigenvalues of the covariance matrix to frequencies, a new quantity related to the specific amount of energy in each mode is suggested to replace the universal $k_B T$ term (see eq. (4.15) and (4.17)). A pre-treatment of the trajectory is also suggested to remove the apparent mixing between vibrational modes and the six translational and rotational motions. In this pre-treatment, the original trajectory in Cartesian space is transformed to a new set of coordinates including three Cartesian axes with the origin at the center of mass of the system, three rotation axes to satisfy the Eckart condition in eq. (4.21), and $3N-6$ axes for the pure internal vibrations. By identifying and overlapping these new coordinates at each time interval the translational and rotational motions are eliminated from the trajectory.

Finally, for molecules with flexible structures, a modified procedure of treating the fragments first and then regrouping the whole molecule is proposed in this chapter. Test trajectories generated by QM/MM simulations for isolated water, methane, the water dimer, and ethane systems are presented to show the improvement made by these modifications. All frequencies calculated by the original PCA and modified PCA methods are compared to illustrate the overall improved accuracy. It is also shown that the frequencies are closer to experiment than the standard quantum chemical calculation results.

4.7. Acknowledgements.

This work was supported by the U.S. Department of Energy Contract No. DE-FG03-01ER15164, NSF/NRAC supercomputer time allocation number MCA96-N019, and supercomputer time from the Oklahoma Supercomputing Center for Education and Research (OSCER).

4.8. References.

- [1] Amadei, A.; Linssen, A. B.; Berendsen, H. J. *Proteins* 1993, 17, 412.
- [2] Case, D. A. *Curr. Opin. Struc. Biol.* 1994, 4, 285.
- [3] Karplus, M.; Kushick, J. N. *Macromolecules* 1981, 14, 325.
- [4] Levy, R. M.; Karplus, M.; Kushick, J.; Perahia, D. *Macromolecules* 1984, 17, 1370.
- [5] Hayward, S.; Go, N. *Annu. Rev. Phys. Chem.* 1995, 46, 233.
- [6] Brooks, B. R.; Janezic, D.; Karplus, M. *J. Comput. Chem.* 1995, 16, 1522.
- [7] Balsera, M. A.; Wriggers, W.; Oono, Y.; Schulten, K. *J. Phys. Chem.* 1996, 100, 2567.
- [8] Pratt, W. K. *Digital Image Processing*; John Wiley: New York, 1978.
- [9] Fukunaga, K. *Introduction to Statistical Pattern Recognition*; Academic: New York, 1972.
- [10] Morrison, D. F. *Multivariate Statistical Methods*; McGraw-Hill: New York, 1990.
- [11] Lanczos, C. *Linear Differential Operators*; D. Van Nostrand Company Ltd.: London, 1961; Vol. Sec. 3.9.
- [12] Golub, G. H.; Van Loan, C. F. *Matrix Computations*; The Johns Hopkins University Press: Baltimore, 1989.
- [13] Gerbrands, J. J. *Pattern Recognition* 1981, 14, 375.
- [14] Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; Dover Publications, Inc.: New York, 1955.
- [15] Stillinger, F. H.; Weber, T. A. *Science* 1984, 225, 983.
- [16] Ohmine, I.; Tanaka, H.; Wolynes, P. G. *J. Chem. Phys.* 1988, 89, 5852.
- [17] Stratt, R. M. *Acc. Chem. Res.* 1995, 28, 201.

- [18] Cho, M.; Fleming, G. R.; Saito, S.; Ohmine, I.; Stratt, R. M. *J. Chem. Phys.* 1994, *100*, 6672.
- [19] Keyes, T. *J. Phys. Chem. A* 1997, *101*, 2921.
- [20] Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* 1990, *11*, 700.
- [21] Cui, Q.; Karplus, M. *J. Chem. Phys.* 2000, *112*, 1133.
- [22] Hayward, S.; Kitao, A.; Hirata, F.; Go, N. *J. Mol. Biol.* 1993, *234*, 1207.
- [23] Hayward, S.; Kitao, A.; Go, N. *Protein Science* 1994, *3*, 936.
- [24] Tama, F.; Miyashita, O.; Kitao, A.; Go, N. *Eur. Biophys. J.* 2000, *29*, 472.
- [25] McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA, 2000.
- [26] Berne, B. J.; Harp, G. D. *Adv. Chem. Phys.* 1970, *17*, 63.
- [27] Gordon, R. G. *Adv. Mag. Res.* 1968, *3*, 1.
- [28] Priestley, M. B. *Spectral Analysis and Time Series, Vol. 1: Univariate Series, Chapt. 7*; Academic Press: London, 1981.
- [29] Percival, D. B.; Walden, A. T. *Spectral Analysis for Physical Applications*; Cambridge University Press: Cambridge, 1993.
- [30] Wheeler, R. A.; Dong, H. *ChemPhysChem* 2003, *4*, 1227.
- [31] Smith, D. F.; Overend, J. *Spectrochim. Acta A, Part A* 1972, *28*, 471.
- [32] Page, R. H.; Frey, J. G.; Shen, Y. R.; Lee, Y. T. *Chem. Phys. Lett.* 1984, *106*, 373.
- [33] Coker, D. F.; Miller, R. E.; Watts, R. O. *J. Chem. Phys.* 1985, *82*, 3554.
- [34] Huang, Z. S.; Miller, R. E. *J. Chem. Phys.* 1989, *91*, 6613.
- [35] Huisken, F.; Kaloudis, M.; Kulcke, A. *J. Chem. Phys.* 1996, *104*, 17.
- [36] Walrafen, G. E.; Yang, W. H.; Chu, Y. C. *J. Phys. Chem. B* 1999, *103*, 1332.

- [37] Paul, J. B.; Provencal, R. A.; Chapo, C.; Roth, K.; Casaes, R.; Saykally, R. J. *J. Phys. Chem. A* 1999, *103*, 2972.
- [38] Fredin, L.; Nelander, B.; Ribbegard, G. *J. Chem. Phys.* 1977, *66*, 4065.
- [39] Nelander, B. *J. Chem. Phys.* 1978, *69*, 3670.
- [40] Bentwood, R. M.; Barnes, A. J.; Orville-Thomas, W. J. *J. Molec. Spectrosc.* 1980, *84*, 391.
- [41] Dyke, T. R.; Mack, K. M.; Muenter, J. S. *J. Chem. Phys.* 1977, *66*, 498.
- [42] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* 1983, *79*, 926.
- [43] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* 1994, *98*, 11623.
- [44] Harrison, R. J.; Straatsma, T. P.; Dupuis, M.; Bylaska, E. J.; Fann, G. I.; Windus, T. L.; Apra, E.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; de Jong, B.; Deegan, M.; Dylla, K.; Elwood, D.; Fruchtl, H.; Glendenning, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kendall, R.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nieplocha, J.; Niu, S.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wolinski, K.; Wong, A.; Zhang, Z. *NWChem, A Computational Chemistry Package for Parallel Computers, Version 4.0*; Pacific Northwest National Laboratory: Richland, WA 99352-0999, USA, 2000.
- [45] Wasikowski, C.; Klemm, S. *XMol*; Research Equipment, Inc. d.b.a. Minnesota Supercomputer Center, Inc., 1993.
- [46] Grafton, A. K.; Wheeler, R. A. *J. Comput. Chem.* 1998, *19*, 1663.
- [47] Jursic, B. S. *J. Mol. Struct.* 1998, *434*, 29.
- [48] Scott, A. P.; Radom, L. *J. Phys. Chem.* 1996, *100*, 16502.
- [49] Hobza, P.; Bludsky, O.; Suhai, S. *Phys. Chem. Chem. Phys.* 1999, *1*, 3073.
- [50] Keutsch, F. N.; Saykally, R. J. *Proc. Nat. Acad. Sci. USA* 2001, *98*, 10533.
- [51] Dinca, F.; Teodosiu, C. *Nonlinear and Random Vibrations*; Academic Press: New York, 1973.

- [52] Reichl, L. E. *A Modern Course in Statistical Physics*; John Wiley & Sons: New York, 1998.
- [53] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.
- [54] Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*; Cambridge University Press: Cambridge, 1996.
- [55] *The Aldrich Library of FT-IR Spectra*; 1st ed.; Pouchert, C. J., Ed.; Aldrich Chemical Company: Milwaukee, 1985; Vol. 1.
- [56] Rorabacher, D. B. *Anal. Chem.* 1991, 63, 139.
- [57] Harman, H. H. *Modern Factor Analysis*; The University of Chicago Press: Chicago, 1972.
- [58] Malinowski, E. R. *Factor Analysis in Chemistry*; John Wiley & Sons, Inc.: New York, 1991.
- [59] Gnanadesikan, R. *Methods for Statistical Data Analysis of Multivariate Observations*; John Wiley & Sons, Inc.: New York, 1977.
- [60] Galka, A. *Topics in Nonlinear Time Series Analysis with Implications for EEG Analysis*; World Scientific: Singapore, 2000.
- [61] Young, E. C. *Vector and Tensor Analysis*; Marcel Dekker: New York, 1993.
- [62] Eckart, C. *Phys. Rev.* 1935, 47, 552.
- [63] Pickett, H. M.; Strauss, H. L. *J. Am. Chem. Soc.* 1970, 92, 7281.
- [64] Goldstein, H. *Classical Mechanics*; Addison-Wesley Publishing Company, Inc.: Reading, MA, 1959.
- [65] Boesch, S. E. *unpublished work* 2003.
- [66] Shimanouchi, T. *Table of Molecular Vibrational Frequencies*; Natl. Stand. Ref. Data Ser., National Bureau of Standards: Washington, D.C., 1972; Vol. I.

Chapter 5

Vibrational energy relaxation in water studied by principal mode analysis of quantum mechanical/molecular mechanical (QM/MM) simulations.

5.1. Introduction.

The unique structural and dynamic properties of liquid water are the key to understanding the solvation process and aqueous reaction chemistry. Water is the universal solvent in life and forms a complicated hydrogen bonding network in interacting with biological systems. By allowing energy transfer through this network, water helps to stabilize reactive intermediates and lower the energy barrier in biological processes. [1-3] The details of how water receives energy and then distributes it among the hydrogen bonded continuum are fascinating aspects of liquid water dynamics. Spectroscopic studies of vibrational energy relaxation (VER) in aqueous systems can provide direct insights into the energy transfer processes between the solute and the bath and therefore has been intensively investigated for years. [4-6] Among these studies, the energy relaxation time scales, the mechanisms of energy transfer, and the rearrangement of surrounding water molecules due to the extra energy introduced into the system are all of great interest. Furthermore, for polyatomic molecules, intramolecular vibrational energy redistribution (IVR) provides an alternative path for energy transfer to VER. Because of the solute-bath interactions through hydrogen bonding, IVR among water's three vibrational modes

is considered to be a competitive mechanism to VER and therefore makes the whole picture of energy transfer in liquid water even more interesting. [7]

Experimentally, with the recent advances in laser technology and time-resolved spectroscopy, a particular vibration of an interesting molecule can be excited and its relaxation can be observed directly even at the femtosecond level. Saykally *et al.* used far-infrared vibration-rotation-tunneling spectroscopy to measure the relaxation times following excited intermolecular vibrations in a series of water systems. [8-10] They found that the lifetime for the hydrogen bond was strongly affected by the excited librational motions and the hydrogen bond breaking was closely related to the reorientation relaxation and dielectric relaxation processes in the liquid phase. The high frequency water O-H stretching vibrations in liquid water and HDO/D₂O were also studied by femtosecond mid-infrared pump-probe spectroscopy. [11-16] By exciting the O-H stretch with pulses around 3400 cm⁻¹ and monitoring the intensity changes at a number of different frequencies, Lock and Bakker reported a lifetime of 0.26 ps for the O-H stretch excitation and a thermal equilibrium constant of ~0.55 ps for surrounding water to reposition. [16] They suggested that the H-O-H bending overtone was the energy-accepting intermediate and that the energy was thermally distributed among all degrees of freedom of surrounding water molecules. The temperature dependence of the time scales were also studied to show an unusual increase in the 0.26 ps time scale with increasing temperature. The authors suggested that the weaker hydrogen bond at high temperature lowered the coupling between the O-H stretch and H-O-H bend states and therefore elongated the energy transfer time

scale. They also indicated that a local temperature jump caused the “uphill” spectral diffusion by weakening the local hydrogen bonding.

Dlott *et al.* applied an ultrafast anti-Stokes Raman technique to the same water O-H stretching motions excited from 3000 to 3600 cm^{-1} and observed two comparable time scales of 0.4ps and 0.7-1.1ps by analyzing the responses of the absorption band at the same frequency range. [17-21] However, their interpretations of these time scales were different from Lock and Bakker’s. Dlott *et al.* attributed the short-time process to spectral diffusion which is due to “local” liquid structure and the longer-time process to energy transfer from O-H stretching to H-O-H bending vibrations. [17] For the mechanisms of energy relaxation in water, both groups agreed that hydrogen bonding played a central role in which the weaker hydrogen bond reduced the coupling between the stretch and the accepting intermediates and bath. Dlott *et al.* also further suggested that the specific relaxation intermediates might be responsible for breaking hydrogen bonds. [20] Besides these pump and probe vibrational spectroscopic studies, other VER experimental work includes vibrational saturation studies ($0.3 < \tau < 0.6\text{ps}$), [22] polarization resolved infrared double resonance spectroscopy, [23] and a laser induced fluorescence/heated flow tube method. [24] Among these studies, there are few direct measurements of the time scales of IVR in liquid water and its mechanisms are much less well understood.

VER in condensed phase has also been studied theoretically. [25-35] Kenkre *et al.* presented a complete quantum mechanical treatment of both the internal degrees of freedom of the excited solutes and the bath degrees of freedom based on the

assumption that the coupling between the two sets of degrees of freedom is weak. With the use of Fermi's golden rule a general expression for the relaxation rate was given for polyatomic molecules in liquid and its temperature dependence was also predicted. [25] Oxtoby developed a semiclassical theory by replacing the quantum correlation functions with their classical counterparts. Skinner *et al.* further developed this theory by introducing quantum-correction factors (QCF) for multiple types of VER processes. [26,27,32-35] In a different approach, Stratt *et al.* suggested that the VER dynamics could be described simply by instantaneous-pair theory which only involves the relaxing solute and a single mutual-nearest-neighbor solvent molecule. By using instantaneous normal mode (INM) analysis they verified this assumption, but only for those high-frequency vibrations relative to the solute-solvent intermolecular instantaneous normal modes. As the excited vibrational frequency decreases, effects from multiple solvent molecules needed to be accounted for in the relaxation dynamics. [29-31]

Classical molecular dynamics (MD) or mixed quantum-classical MD simulations have been applied to VER in the liquid phase, particularly in water solution systems. [7,36-44] Abseher *et al.* and Bizzarri and Cannistraro investigated the structural and dynamical properties of water on a protein surface. Through their MD simulations, relaxation behaviors of water in the first solvation shell including the diffusive mobility, rotational reorientation, and residence times were analyzed. [36,37] Rey and Hynes studied the excited O-H stretch relaxation in HOD/D₂O using MD. They reported a long relaxation time of ~8 ps and suggested an O-H stretch relaxation

pathway through the H-O-H bend overtone as the intermediate. [36,40] Other theoretical calculations of the rates of VER in water systems include path integral influence functional methods [7,42] and path integral centroid molecular dynamics. [43,44]

In the present chapter, we provide a different approach to the conventional normal mode analysis to analyze the vibrations of water. Combined quantum mechanical/molecular mechanical (QM/MM) methods are used to simulate liquid water. [45] These methods provide a much more accurate description of the potential energy surface for liquid water than the empirical force fields used to determine nuclear wavefunctions in the previous theoretical studies. The VER in water is then studied by using principal mode analysis (PMA) of the QM/MM trajectories. [46,47] PMA is based on the statistical analysis method of principal component analysis (PCA). PCA finds wide use for compressing and representing data in signal processing, [48] pattern recognition, [49] and multivariate time series statistical analysis. [50] In our previous work, [46,47] we have shown that PMA, unlike conventional normal mode analysis, is able to incorporate anharmonicity and solvent effects for studying molecular vibrations in liquid phase. For instance, the frequencies of isolated water molecules, water dimers, and liquid water calculated by PMA are much closer to experiment than frequencies calculate by using conventional quantum chemical calculations. PMA also gives vibrational modes for liquid water that are different from the normal modes of water in the gas phase, because it incorporates some anharmonicity. In this study, the principal modes of water will be used for separating

the total vibrational energy into individual modes. By monitoring the time dependence of the vibrational energy in every mode of a single “solute” water molecule, it will ideally give more accurate time scales and mechanisms for VER and IVR in liquid water since the anharmonic QM/MM potentials can be optimally analyzed in terms of principal vibrational modes.

5.2. Principal mode analysis: principal component analysis and its application to molecular vibrations.

Principal component analysis originated from Pearson's fundamental work of fitting lines and planes by a means of orthogonal least squares [51] and was later developed by Hotelling. [52] In its century-long development history, the method has been given various names such as factor analysis, [53] the Karhunen-Loève transform, [48,49] or singular value decomposition. [54] Besides subtle differences, these techniques all involve searching for a set of orthonormal vectors determined in a sequence so that each vector accounts for a maximum remaining variance in a multivariate data set. The detailed technical aspect of PCA and the related methods are not the subject of this chapter and thorough discussions can be found elsewhere. [50,55]

For a piece of a molecular trajectory given by a computer simulation, which forms a multivariate time series, PCA involves first calculating the covariance matrix \mathbf{R}_x from the instantaneous atomic positions $x(t)$:

$$R_{ij} = \langle (x_i(t) - \langle x_i \rangle) (x_j(t) - \langle x_j \rangle) \rangle, i, j = 1, 2, \dots, 3N \quad (5.1)$$

where i and j are atomic indices and N is the total number of atoms in the system.

This matrix has been proven to represent an effective force constant matrix (constructed in normal mode analysis) in the following manner [56,57]

$$\mathbf{F}_{\text{eff}} = k_B T \mathbf{R}_x^{-1}, \quad (5.2)$$

where k_B and T are the Boltzmann constant and temperature, respectively. Then following the standard solution of the eigenvalues equation constructed from the force constant matrix, the frequencies and modes can be obtained from the diagonalization of \mathbf{R}_x^{-1} . The details of applying PCA to the molecular vibrations and necessary modifications to the standard procedures can be found in Ref. [58].

This method, principal mode analysis (PMA), provides an alternative to the conventional normal mode analysis with favorable advantages. Since the second order moments (covariances) are utilized instead of the second derivatives in normal mode analysis for the approximation of the potential energy surface, PMA is able to incorporate some anharmonicity in constructing the matrix. Moreover, PMA only diagonalizes the covariance matrix once because it calculates a time average of the matrix prior to solving the eigenproblem. Indeed, PCA has been used to investigate several lowest-frequency, large-scale motions of proteins. [59-61] In the present chapter, however, our focus is on providing evidence that PCA is able to accurately extract the effective motions from QM/MM molecular trajectories showing some anharmonic character. With the principal modes obtained, the specific amount of energy in each of these separate motions can be traced as a function of time. Then useful information about time scales and mechanisms of vibrational energy relaxation will be readily revealed.

5.3. Molecular dynamics or QM/MM simulations of vibrational energy relaxation.

The processes of VER have been theoretically studied using molecular dynamics (MD) simulations. At the molecular level, MD or QM/MM (MD with quantum chemically derived potentials) simulations help to reveal the mechanisms of VER for a number of solute-bath systems and give time scales in good agreement with experiment. [5,7,38-42,62-64] Considering the canonical ensemble, [65] the probability of finding a vibrational mode in the i^{th} vibrational state is

$$P_i = \exp(-E_i / kT) / \sum_i \exp(-E_i / kT) \quad (5.3)$$

where E_i is the energy of the i^{th} state and the total vibrational energy possessed by this mode is then the sum over all the possible states

$$E_{vib} = \sum_i E_i P_i . \quad (5.4)$$

If some excess vibrational energy is input into this vibrational mode at the beginning of an MD simulation, the nonequilibrium energy will relax back to equilibrium with time or equivalently, the probability P_i will change with time as well. The time derivative of $P_i(t)$ can be given as the sum of all energy transitions with higher energy states minus the sum of all transitions with lower energy states. [66] If one can safely assume that the populations of the mode at higher excited states than the 1st are negligible and ignores the upward transitions from the 1st excited state, following the standard treatment, [67] an expression for VER can be obtained:

$$\frac{dE_{vib}(t)}{dt} = \sum_i E_i \frac{dP_i(t)}{dt} = -(k_{1 \rightarrow 0} + k_{0 \rightarrow 1})(E_{vib}(t) - \langle E \rangle_{ens}), \quad (5.5)$$

with the k 's being the transition rate constants and $\langle E \rangle_{ens}$ the equilibrium vibrational energy. Eq.(5.5) implies that the total vibrational energy decays exponentially with the characteristic time $\tau = 1/(k_{1 \rightarrow 0} + k_{0 \rightarrow 1})$. Using the detailed balance condition, $k_{0 \rightarrow 1} = \exp(-\hbar\omega_{10} / kT)k_{1 \rightarrow 0}$, where ω_{10} is the frequency difference between the two states, $k_{0 \rightarrow 1}$ would be much less than $k_{1 \rightarrow 0}$ (it is usually the case when ω_{10} is larger than 300 cm^{-1} at $T=300 \text{ K}$), the vibrational energy relaxation in eq.(5.5) will be a monoexponential decay with only the transition rate from the first excited state to the ground state dominating the process. When plotting the vibrational energy versus time, a normalized term, $(\bar{E}(t) - \langle E \rangle_{ens}) / (\bar{E}(0) - \langle E \rangle_{ens})$, where $\bar{E}(t)$ is an average energy over a short time period (0.1 ps in analyzing the energy profiles in this work), is commonly used for getting the characteristic time τ .

In the introduction it is suggested that PMA is able to calculate actual vibrational modes of liquid water in which some anharmonic effects are included. These PMA-derived modes are not necessarily the same as the ones derived from normal mode analysis and thus might show more accurate and more interesting relaxation dynamics. In the PMA vibrational mode frame, the total vibrational energy possessed by the mode can be expressed in the quasiharmonic approximation as

$$E_{vib}(t) = \sum_i \frac{1}{2} \dot{q}_i(t)^2 + \sum_i \frac{1}{2} \omega_i^2 q_i(t)^2 \quad (5.6)$$

where $q_i(t)$ and $\dot{q}_i(t)$ are the PMA mode coordinates and velocities obtained by projecting the mass-weighted Cartesian coordinates and velocities onto the i^{th} principal mode.

In liquid water, VER and IVR provide competitive paths for energy transfer. In order to further understand these processes, water in the gas phase can be studied by PMA to provide useful information about IVR. For an isolated water molecule, pure IVR should be possible because of the anharmonic couplings between vibrational modes. These IVR processes can be examined in the vibrational energy versus time plots to give time scales of the energy transfer path within water molecules. The IVR time scales can also be compared with the time scales of VER for liquid water to find out how competitive the IVR processes are with VER.

In VER processes, solvent water molecules play an important role since their low frequency intermolecular motions form a continuum of states that are the final destination of the solute's excess vibrational energy. A time-dependent perturbation approach (Fermi's golden rule) has been introduced to calculate the rate constants for such VER processes from the solvent's perspective. This approach involves a solute-bath force correlation function calculated from an MD or QM/MM simulation and its Fourier transform at the transition frequency ω_{ij} is: [32,33,67]

$$k_{i \rightarrow j} = \frac{1}{2\mu\hbar\omega_{ij}} \int_{-\infty}^{\infty} \exp(i\omega_{ij}t) \langle F(t)F(0) \rangle_{qni} dt \quad (5.7)$$

in which $k_{i \rightarrow j}$ is the rate constant for a transition from the state i to j , the μ is reduced mass, and $\hbar\omega_{ij}$ the energy difference between state i and j . $F(t)$ is the fluctuating

force acting on the vibrational mode coordinate by the solvent molecules and $\langle \dots \rangle_{qm}$ implies a quantum mechanical time correlation function. Usually this quantum mechanical correlation function can be replaced by its classical analogue, $\langle \dots \rangle_{cl}$, multiplied by a quantum-correction factor (QCF). For a vibration-translation and/or vibration-rotation energy transfer scheme between the solute and bath, a QCF has been suggested: [33]

$$QCF(\omega) = \exp(\beta\hbar\omega/4) (\beta\hbar\omega / (1 - \exp(-\beta\hbar\omega)))^{1/2}. \quad (5.8)$$

This equation combined with eq. (5.7) has therefore provided a way to study VER through classical molecular dynamics simulations. With the PMA-derived vibrational modes for liquid water, the solute-bath forces can be transformed from a Cartesian coordinate system to the principal mode coordinate system and the calculation of the rate constants is straightforward. Thorough discussions on various aspects of this formalism can be found in Ref. [25,26,32].

5.4. Computational method.

A box of 216 water molecules was first equilibrated using classical molecular dynamics (MD) at constant pressure and temperature with periodic boundary conditions and the TIP3P force field and the CHARMM MD program. [68] Then the combined QM/MM simulation technique was used to mimic the liquid phase with a relatively low computational load. For each of 12 QM/MM simulations, a single water molecule was randomly selected from the box as the QM region and the remaining 215 molecules were chosen as the MM region in which the bond lengths and angles were held rigid using SHAKE. The CHARMM/GAMESS combination [69] was used with the B3LYP/6-311++G (2d,2p) method/basis set which reportedly gives a correct hydrogen bond structure in the water dimer. [70] The 12 water boxes were then equilibrated for 1 ps of QM/MM simulations with the same conditions as used in MD followed by another 1 ps for trajectory collection for calculating PMA frequencies.

An isolated water molecule has three vibrational modes: the antisymmetric stretch ν_1 with an experimentally measured frequency of 3756 cm^{-1} , the symmetric stretch ν_3 of 3657 cm^{-1} , and the bend ν_2 of 1593 cm^{-1} . [71] However, the absorption spectrum of liquid water differs substantially from the gas phase spectrum due to hydrogen bonding. The two stretching modes overlap and red-shift to form a broad band centered at $\sim 3400\text{ cm}^{-1}$ with a width of about 400 cm^{-1} and the bending mode blue-shifts to 1635 cm^{-1} . [72] In Dlott's ultrafast IR-Raman spectroscopic studies of VER, the O-H stretching mode(s) of water were excited in the $2850\text{-}3670\text{ cm}^{-1}$ range.

[17,20] Similarly, Bakker *et al* pumped the water O-H stretch(es) at 3250-3400 cm^{-1} in their femtosecond mid-IR pump-probe spectroscopic experiments. [13,16] Due to the large overlapping between the two O-H stretching modes shown in the spectrum, both the antisymmetric and the symmetric stretch were probably excited in those two sets of experiments. We have calculated the frequencies of liquid water with the B3LYP/6-31++G(2d,2p) method and then both the antisymmetric and symmetric O-H stretching modes are excited based on the newly calculated principal modes.

The starting geometries of the excited water molecules were prepared first using the QM water structures averaged over the 1 ps trajectories. Then by manually changing the atomic positions according to the principal modes, the antisymmetric and symmetric stretching modes were excited by an amount corresponding to one quantum of vibrational energy. Finally, these new structures were used to replace the structures of the QM waters at the last step of the previous QM/MM simulations. After the preparation, 8 ps trajectories of the QM water and the remaining MM water molecules were generated with a step size of 0.5 fs for each water box. Meanwhile, in order to study VER in a water molecule itself without solvent effects, the same QM waters were simulated under the same conditions, but excluding the 215 classical water molecules.

5.5. Results and discussion.

5.5.1. Principal mode analysis of isolated water molecules and liquid water.

The pathway of vibrational energy relaxation in water involves two possible distinguishable processes: the intramolecular vibrational redistribution (IVR) in which the vibrational energy is transferred directly between the same water molecules's three modes, and vibrational energy relaxation (VER) in which the excess energy may be distributed among adjacent water molecules. Before these processes can be studied, the vibrational frequencies and modes need to be determined to set up a vibrational mode frame in which the total vibrational energy can be separated. The average frequencies calculated by PMA of the 12 NVE trajectories are presented in Table 5.1 as $1627\pm 4\text{ cm}^{-1}$ for the H-O-H bend, $3734\pm 21\text{ cm}^{-1}$ for the O-H symmetric stretch, and $3824\pm 11\text{ cm}^{-1}$ for the O-H antisymmetric stretch, respectively. This excellent agreement with experiment is a demonstration of PMA's ability to incorporate anharmonicity and therefore calculate more accurate frequencies than standard quantum chemical methods. The PMA frequencies from different trajectories are consistently closer to experiment by approximately 100 cm^{-1} than the quantum chemical result for the stretching modes. For the bending frequency, the PMA method gives a better value than the quantum chemistry method as well. This result also shows a very accurate description of the potential energy surface for gas phase water by the B3LYP/6-311++G(2d,2p) method.

In putting vibrational energy into the molecules, these principal modes are used to indicate the directions in which the atomic positions are changed. The orthogonality

of the principal modes ensures only one particular mode is excited at the beginning of each simulation. This clear separation in mode excitation will make it easier to understand the mechanisms of IVR/VER in terms of the actual molecular motions.

The single-mode excited water molecules were simulated with the same conditions as in the equilibration stage. PMA was again applied to the trajectories to calculate frequencies and modes. The results are summarized in Table 5.1 under the isolation/excited column. The PMA frequencies are 3695 ± 6 and 3764 ± 6 cm^{-1} for the two stretches. The bending mode frequency stays the same as the unexcited water. The excellent agreement with experimental frequencies again indicates that PMA can provide an accurate description of molecular vibrations for water. The extra excitation energy in the stretching modes means more areas of the energy surface being sampled during the simulation and thus leads to a more precise approximation of the energy surface than the unexcited water.

The same starting geometries of water molecules were used in the liquid phase QM/MM simulations as well. The average PMA-derived vibrational frequencies with standard deviations are also listed in Table 5.1 for comparison. When the molecules were not excited, the frequencies of the stretching modes represent the experimentally observed “red-shift” in wavenumbers relative to frequencies measured for gas-phase water very well. The symmetric stretch goes from 3734 to 3455 cm^{-1} while the antisymmetric stretching frequency changes from 3824 to 3507 cm^{-1} . However, the bending frequency calculated by PMA decreases from 1627 to 1502 cm^{-1} , which is a trend opposite to the one shown in experiment for liquid water. The standard

Table 5.1. Frequencies for the bending, symmetric stretching, and antisymmetric stretching modes calculated by PMA of isolated water and liquid water molecules. A water box was first equilibrated by using molecular dynamics. Then a water molecule was randomly selected and defined as the QM region from the box for each of total 12 QM/MM simulations with the B3LYP/6-311++G (2d,2p) method. At the beginning of each simulation, the geometry of the water molecule was either not excited (no exc), or symmetrically or antisymmetrically excited (exc, the two case were averaged together). The PMA frequencies are compared with the standard quantum chemical method (quantum) and experiment.

	isolation				liquid		
	no exc	exc	quantum	exp	No exc	exc	exp
bend	1627±4	1626±12	1640	1593	1520±121	1582±30	1635
sym str	3734±21	3695±6	3823	3657	3455±82	3316±82	3240
antis str	3824±11	3764±6	3924	3756	3507±70	3407±56	3446

deviations of those frequencies become much larger than in the gas phase, a reflection of inhomogeneous spectral broadening typical in the liquid phase. When the OH stretching modes were excited at the beginning of the simulations, the frequencies of all modes are also in excellent agreement with experimental values. The bending mode is calculated at $1582\pm 30\text{ cm}^{-1}$ and that is also much closer to the experiment value than the frequency given from the unexcited liquid water. Although compared to the PMA frequency ($1626\pm 12\text{ cm}^{-1}$) in isolation this bending frequency does not show the expected “blue-shift”, both frequencies are almost statistically identical. The frequencies of OH stretches are given at $3316\pm 82\text{ cm}^{-1}$ and $3407\pm 56\text{ cm}^{-1}$. They also agree with the experimental values within one standard deviation.

5.5.2. Vibrational energy transfer between vibrational modes / IVR in isolated water molecules.

Before going into the condensed phase vibrational energy relaxation, IVR can be first examined in the gas phase where it is clearly a pure IVR process. Having set the molecular vibrations of water in the frame of principal vibrational modes given by PMA, the total vibrational energy can be separated by projecting the atomic positions and velocities in Cartesian coordinates onto the principal modes. Then the vibrational energy in each mode is obtained using eq. (5.6). By plotting these vibrational energy in each mode as functions of time, the paths of the energy transfer between modes and their rates flowing in and out of each mode can be found.

Figure 5.1 shows the total vibrational energy of the three modes for an example water molecule in isolation with the symmetric OH stretch excited. The symmetric stretch has the highest energy and three different oscillations can be seen in this energy profile. There are one with the largest amplitude change and the smallest time period, one with a small amplitude change and relatively fast oscillation, and finally one with a large amplitude change and the more frequent oscillation. The bend and antisymmetric stretch have approximately the same amount of starting energy which is much lower than the symmetric stretch since only the latter mode is excited. In the bend an oscillation appears having the same time period as the second oscillation present in the symmetric stretch energy profile. Meanwhile, the large scale slow envelope oscillation in the antisymmetric mode corresponds to the first oscillation in the symmetric stretch. This plot clearly indicates that the vibrational energy is transferred between the two stretches and between the symmetric stretch and the bend, respectively. An autocorrelation function is calculated for each of the energy profiles and fast Fourier transform (FFT) is used to generate the power spectrum. Figure 5.2 shows the autocorrelation functions and the insets are the power spectra given by FFT. Three characteristic time periods are obtained at 3746 cm^{-1} for the pure symmetric stretch, 130 cm^{-1} for the energy transfer between the two OH stretches, and 521 cm^{-1} for the energy transfer between the symmetric stretch and the bend. The stretching frequency is in agreement with experiment but is not as accurate as the PMA result of 3687 cm^{-1} for this particular trajectory. The reason the total vibrational energy oscillates at the frequency of the symmetric stretch is probably that along the

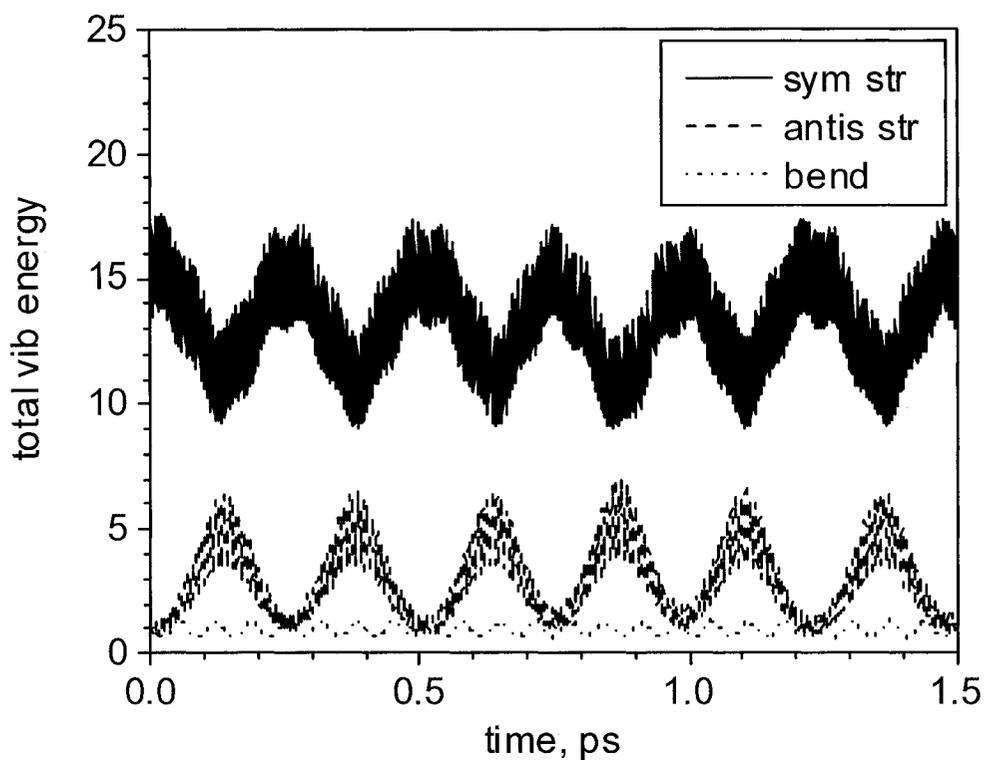


Figure 5.1. Total vibrational energy of the symmetric stretching (top, solid line), antisymmetric stretching (middle, dashed line), and bending (bottom, dotted line) modes for water in isolation with the symmetric stretching mode excited at the beginning of the simulation.

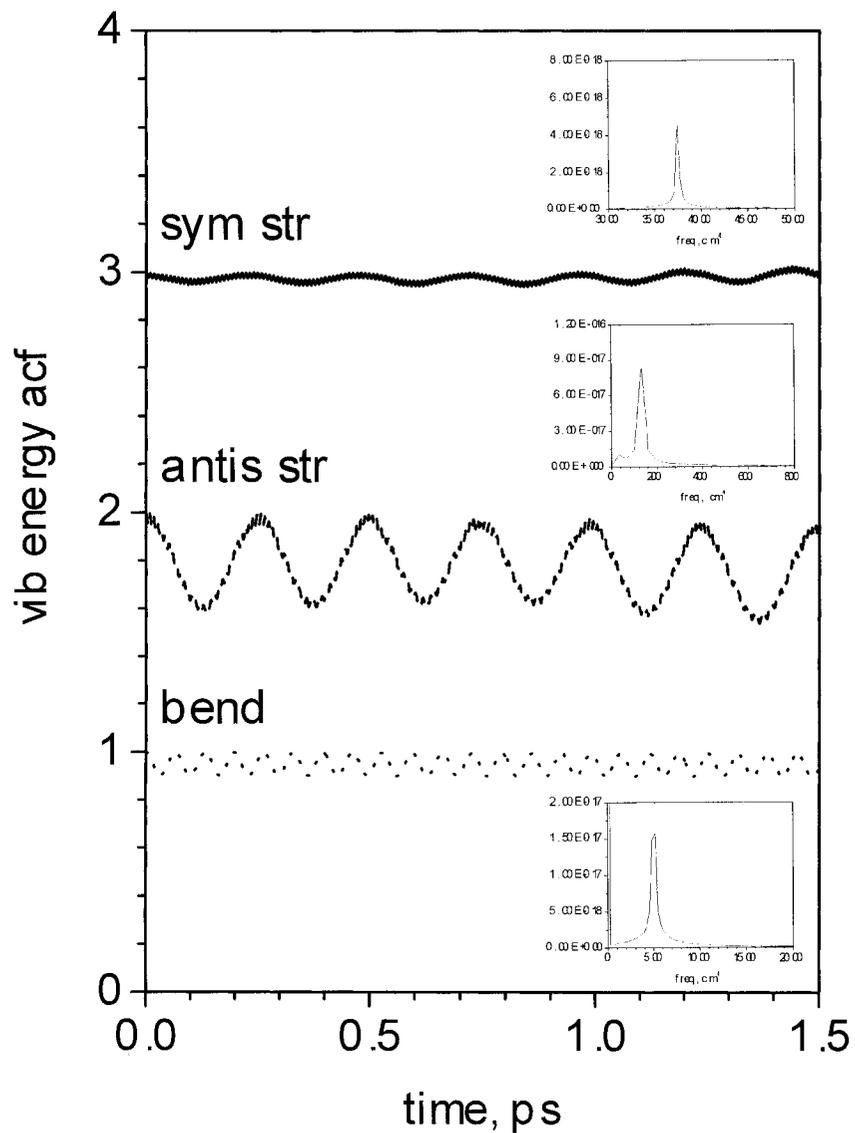


Figure 5.2. The autocorrelation functions for the total vibrational energy presented in Figure 5.1 for water in isolation. The insets are the power spectra of the correlation functions calculated by fast Fourier transform (FFT).

vibration the configuration of atoms changes, the electronic energy also changes at the same rate. In order to keep the total energy constant for the isolated water molecule, there has to be a certain amount of energy exchanged between the vibrational energy and the electronic energy.

The total vibrational energy vs. time plots for all 12 trajectories with the symmetric stretch excited show similar features and their averages are presented in Table 5.2. The two frequencies of 87 cm^{-1} and 516 cm^{-1} correspond to 0.38 ps and 0.065 ps and indicate that these IVR processes are fast for water in isolation and could compete with VER processes in liquid water. To our knowledge, IVR time scales for a single gas-phase water molecule have not been measured, because water forms clusters in the gas phase. [73] The major energy transfer is between the two OH stretches as the amount of energy being transferred accounts for approximately 25% of the total amount of energy possessed by the symmetric stretch (the difference between the maximum and minimum amplitudes of the symmetric stretch is $\sim 25\%$ of the maximum amplitude of the same oscillation in Figure 5.1). This significant energy transfer process also suggests a possible mechanism through IVR for the “uphill” spectral diffusion observed in liquid water, i.e. the direct transfer of energy from a lower frequency to a higher frequency mode. [17]

When the vibrational energy is put in the antisymmetric stretching mode, which has a slightly higher frequencies than the symmetric stretch, the transfer processes are studied in the same manner and all results are presented in Figure 5.3 and 5.4 and

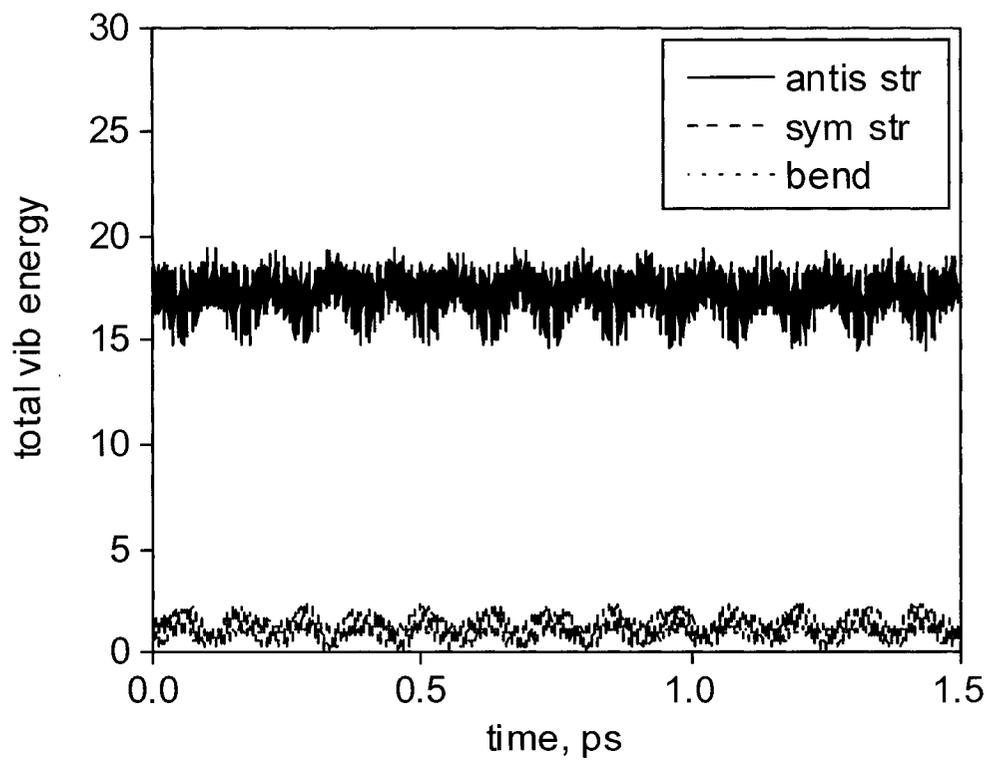


Figure 5.3. Total vibrational energy of the antisymmetric stretching (top, solid line), symmetric stretching (bottom, dashed line), and bending (bottom, dotted line overlapped with symmetric stretching) modes for water in isolation with the antisymmetric stretching mode excited at the beginning of the simulation.

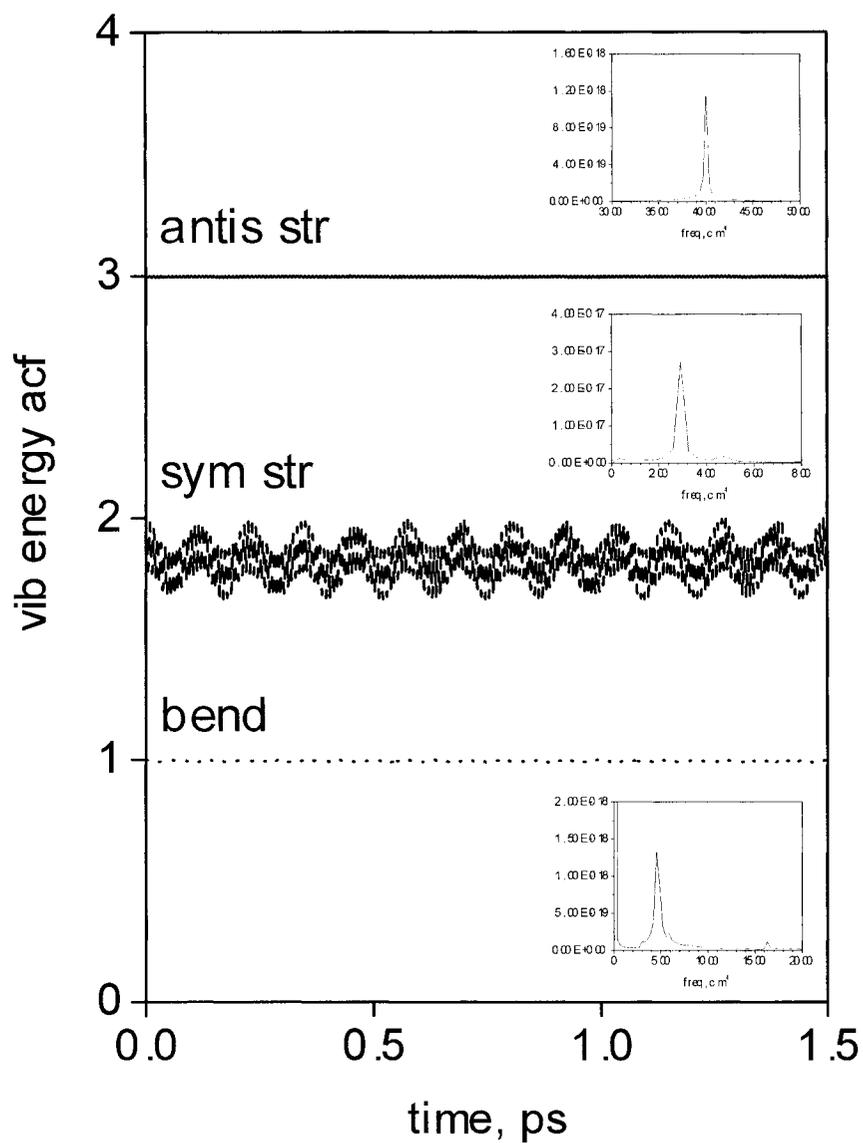


Figure 5.4. The autocorrelation functions for the total vibrational energy presented in Figure 5.3 for water in isolation. The insets are the power spectra of the correlation functions calculated by fast Fourier transform (FFT).

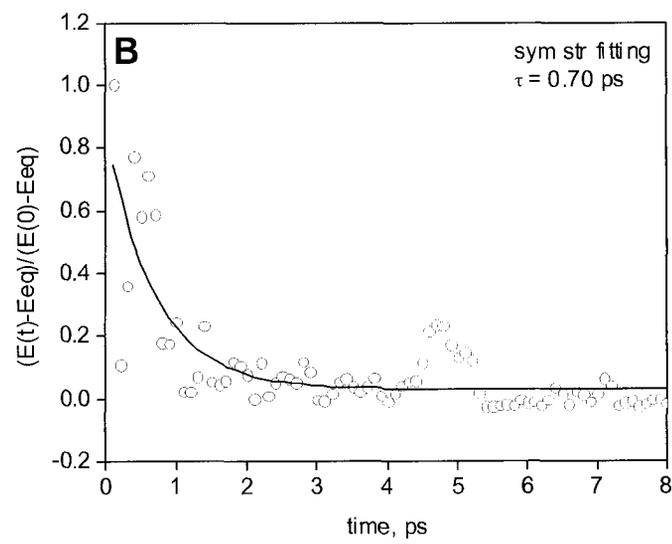
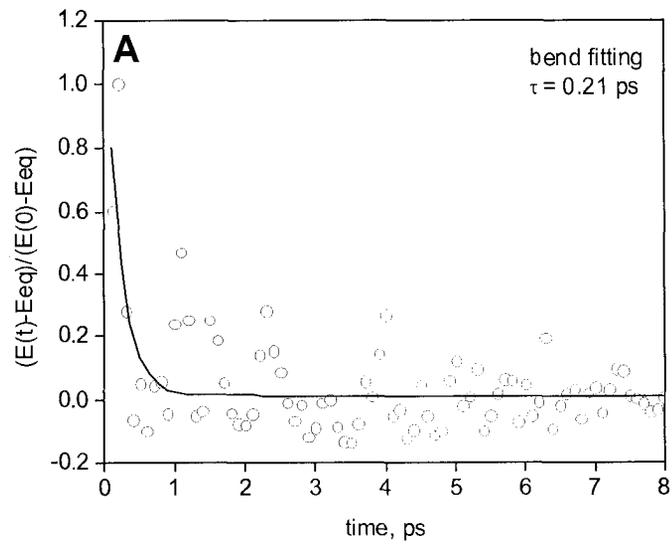
Table 5.2. The energy plot appears quite different from the one previously discussed. The vibrational energy exchange between the two stretches occurs at a more frequent rate but with a smaller variation of energy ($\sim 12\%$ of the total) than the same processes shown in Figure 5.1. The time scale for this energy transfer is calculated as 290 cm^{-1} or 0.12 ps from both the symmetric and antisymmetric stretch energy profiles. This time scale is shorter than the same process when the symmetric stretch is initially excited. From Figure 5.3 it is shown again that the vibrational energy transfer primarily occurs between the stretches. The insignificant amount of energy (less than 5%) being exchanged between either one of the stretches and the bend (even though the power spectrum of the bend shows a 473 cm^{-1} peak which is close to 516 cm^{-1} in the symmetric stretch excited case) indicates weak couplings between the stretching and bending modes for the isolated water molecule.

5.5.3. *Vibrational energy relaxation of liquid water.*

The vibrational energy relaxation of water in the liquid phase is expected to be more complicated than in gas phase. The formation of a hydrogen bonding network represents the strong and complex interactions between the solute water molecules and bath. The vibrations of the water molecules are then strongly perturbed by the vibration-bath coupling. The IVR processes could be altered to show different characters from that in the gas phase. Moreover, the bath water molecules provide a continuum of energy-accepting states (including translational and rotational states) for the solute's excess vibrational energy, so that VER processes are also possible.

The overall effect on the vibrational energy decay by the two competing processes are studied using PMA of QM/MM trajectories of liquid water and then the results are compared with experiment.

Figure 5.5 shows a representative plot of total vibrational energy (more precisely, a normalized excess vibrational energy, see Section 5.3) as a function of time for the three modes of the QM water molecule when the symmetric stretch is first excited. Very different from the same type of the plot for isolated water, the amounts of energy in the modes do not remain constant but rather show an overall decrease over the length of the entire simulation. This energy loss has to be due to the energy transfer from the solute to the bath. The energy profiles also oscillate at times but there is no clear oscillatory pattern of energy exchange between vibrational modes similar to the isolated water case. At the beginning of the trajectory, the symmetric stretch has the highest amplitude and the other two have much lower energy, which is no surprise since energy is input in the symmetric stretch. However, after a short time the two unexcited modes show a large increase while the symmetric stretch shows a decrease in amplitude. This short time can be seen as about 0.2 ps from Figure 5.5 since each point in this figure is an average over 0.1 ps (see Section 5.3). This energy increase cannot be fitted to an exponential growth because of the limited number of points. The amount of energy the bend gains accounts for a significant portion, almost 40% of the initial amount of energy in the symmetric stretch (calculated from the unnormalized values of vibrational energy). Compared to the very small amount of



(Continued on the next page.)

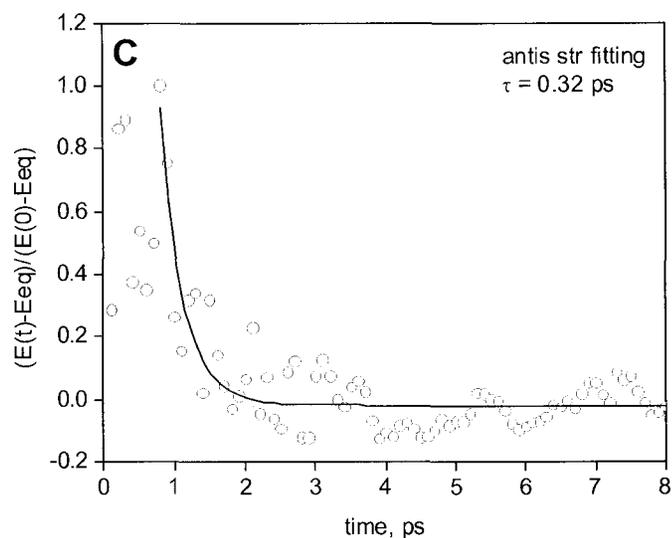
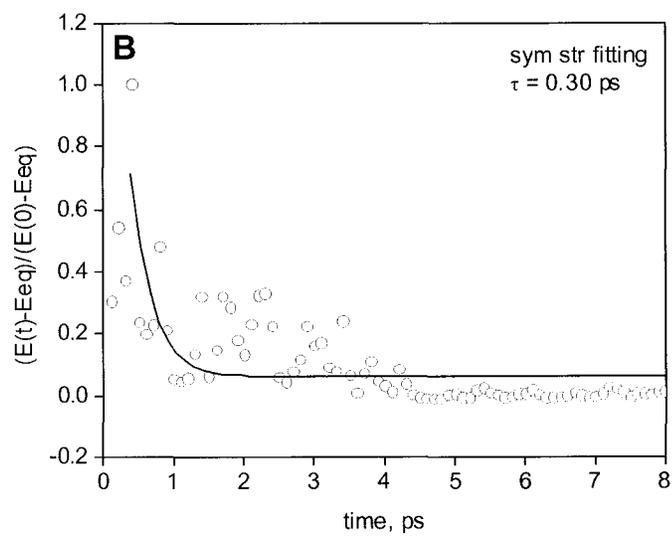
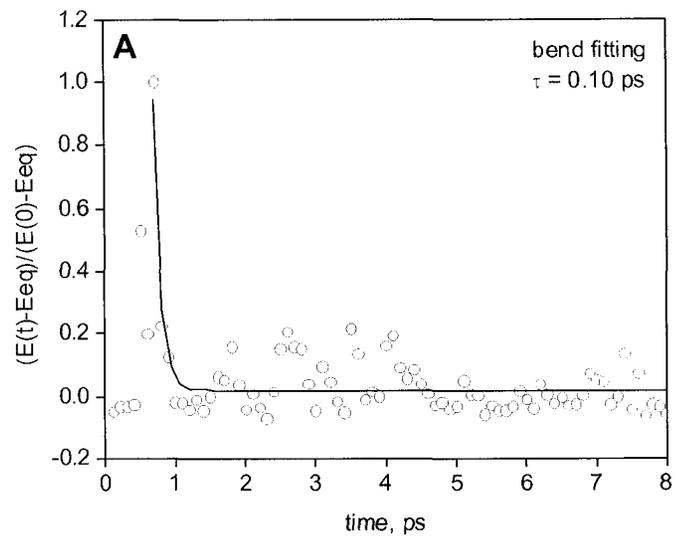


Figure 5.5. The normalized total vibrational energy of the bending (A), symmetric stretching (B), and antisymmetric stretching (C) modes for one sample QM/MM trajectory of liquid water with the symmetric stretching mode of the QM water excited at the beginning of the simulation. The solid lines are fittings with the characteristic times, τ , shown on the figures. Only the fittings with R^2 values larger than 0.6 are used to calculate the averages and standard deviations listed in Table 5.2.

energy transferred between the stretching and the bending modes for the isolated water molecule, this result suggests that this energy transfer process must be different from pure IVR, because of the existence of the solvent water molecules. As the simulation time evolves, the total amounts of energy in the modes gradually drain out. This overall trend is fitted as an exponential decay with a single characteristic time scale and the actual fittings for all the mode are shown in Figure 5.5. The time scales for all the three vibrational modes when the symmetric stretch is excited, averaged over 12 trajectories, are included in Table 5.2 as 0.62 ± 0.23 ps for the symmetric stretch, 0.28 ± 0.12 ps for the antisymmetric stretch, and 0.23 ± 0.16 ps for the bend. It can be seen that there are roughly two time scales for the vibrational energy relaxation, the longer (0.6 ps) one for the primarily excited mode and the shorter one (~ 0.2 ps) for the secondarily excited modes. It should be noted that for the two secondarily excited modes, the exponential decay fitting is done on the energy profiles after ~ 0.3 ps, i.e. after energy transfer into the modes.

Figure 5.6 represents the same vibrational energy versus time plot as in Figure 5.5 but the antisymmetric stretch is excited for the QM water molecule. An example fitting of the vibrational energy in the antisymmetric stretch is shown in the figure. Averaging over 12 trajectories for this excitation, the time scale of overall VER is 0.70 ± 0.31 ps for the antisymmetric stretch. For the secondarily excited modes, after a energy growth range of ~ 0.4 ps, the decay times are 0.25 ± 0.10 ps for the symmetric stretch and 0.16 ± 0.12 ps for the bend. There are again clearly two time scales that are statistically distinguishable. The two-time-scale result calculated based on the



(Continued on the next page.)

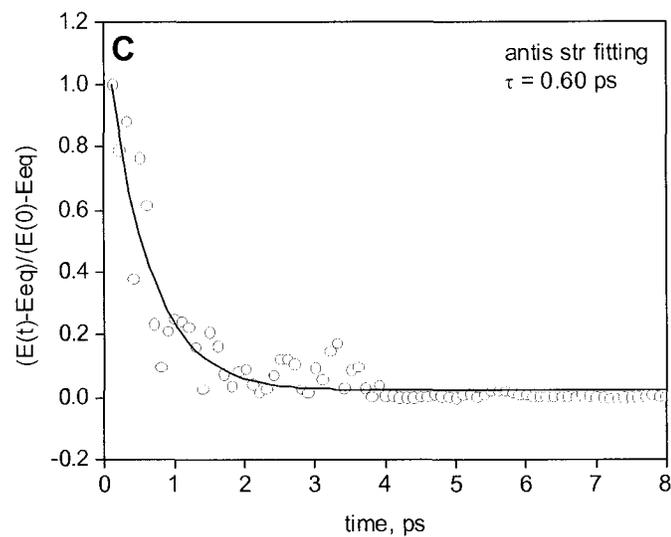


Figure 5.6. The normalized total vibrational energy of the bending (A), symmetric stretching (B), and antisymmetric stretching (C) modes for one sample QM/MM trajectory of liquid water with the antisymmetric stretching mode of the QM water excited at the beginning of the simulation. The solid lines are fittings with the characteristic times, τ , shown on the figures. Only the fittings with R^2 values larger than 0.6 are used to calculate the averages and standard deviations listed in Table 5.2.

principal modes agrees well with what has been found in the pulse-probe spectroscopic experiments. Bakker *et al.* observed two time scales of 0.26 ps and 0.55 ps from their experiments by examining the intensities at several frequencies in the O-H stretching band as functions of delay time following an excitation. The shorter time scale T_1 was attributed to the energy transition from the OH stretch to an intermediate state, which is a specific combination of excited intra- and intermolecular vibrations, while the longer one was for a “thermalization” process which is a vibrational energy redistribution among the bath’s degrees of freedom. [13,16] Dlott’s group also obtained two time scales, 0.4 ps and 0.7-1.1 ps, using a different spectroscopic technique but suggested that the shorter time would result from a “spectral diffusion” as the energy migrates from one end to the other of the OH stretching band in liquid water. For the longer time scale they suggested the transition from the stretching to an intermediate state, which is believed to be the bending overtone. [17,20] Our calculations are in better agreement with Bakker’s in magnitude but with a different explanation. By exciting the stretching modes individually, we find that the vibrational energy relaxation process in the initially excited mode accounts for the longer time scale. However, by studying the overall energy decay it is not possible to determine which of two experimentally measured time scales, “thermalization” or that labeled T_1 , corresponds to this calculated time scale. This question will be answered after the force-force correlation functions are analyzed in the next section. The short energy decay time found in the secondarily excited modes, on the other hand, can be clearly attributed to the spectral diffusion

Table 5.2. Time scales of IVR and VER of water in isolation and liquid phase. For each of the water molecules, either the symmetric stretching (abbreviated as sstr in the table) or the antisymmetric stretching (astr) mode was initially excited. For water molecules in isolation, the time scales for the total vibrational energy transfer rates between modes were calculated by Fourier transforming the energy autocorrelation functions and expressed as the frequency inverse. In liquid, the total vibrational energy in each of the three modes as functions of time were plotted and fitted to an exponential decay with a characteristic time τ . The averages and standard deviations are listed in the table for all 12 trajectories. The approximate time scales of energy growth for the secondarily excited modes (sec. exc. growth) are also listed for comparison with experiment.

isolation				
	sstr excited		astr excited	
	bend/sstr	sstr/astr	bend/astr	sstr/astr
freq ⁻¹ , ps	0.065±0.002	0.38±0.15	0.071±0.003	0.12±0.01
liquid, sstr excited				
τ , ps	sec. exc. growth	bend decay	sstr decay	astr decay
	~0.3	0.23±0.16	0.62±0.23	0.28±0.12
liquid, astr excited				
τ , ps	~0.4	0.16±0.12	0.25±0.10	0.70±0.31
experiment				
τ , ps	spectral diffusion	T ₁	thermalization	
	[13,16]	0.4	0.26	0.55
[17,20]		0.7-1.1		

process observed in Dlott's experiment. Our simulations show that there is a vibrational energy growth process for the secondarily excited O-H stretch in the time range of approximately 0.3-0.4 ps followed by an energy decay of 0.2-0.3 ps. These two time scales are very close suggesting that there is an energy exchange process between the secondarily excited stretch and, most likely the initially excited stretch. This energy redistribution between the two stretches can be seen as the changes in spectroscopic intensity and therefore our interpretation matches with Dlott's for the shorter time scale.

5.5.4. VER from the solvent's perspective.

Vibrational energy transfer processes involve vibrational states of both solute and bath water molecules through the interactions between the two. In contrast to the discrete vibrational states of the solute, the vibrational states of the bath are practically continuous in the low frequency range below 800 cm^{-1} . These motions are mainly the combinations of solvent water molecules' translations and rotations. In one of our previous papers, we calculated frequencies and modes by using PMA for the water dimer and found that the low frequency modes ($100\text{-}300\text{ cm}^{-1}$) are the combinations of each individual water molecule's translations and rotations relative to one other. [46] It is expected that as the size of water clusters increases, there will be more low frequency modes until a near continuum of modes is formed in liquid water. The importance of this continuum is that for every possible transition occurring in the solute, there is a group of states in the bath that have the

corresponding energy change and can be involved as the energy accepting states during VER processes. Therefore, by investigating the solvent dynamics (through the correlation function of forces on the solute by the bath), indirect information about the transitions within the solute can be studied and the transition rate constants can be calculated by using eq. (5.7) and eq. (5.8).

Figure 5.7 shows the time scales ($\tau = k^{-1}$) calculated from the Fourier transformed classical force-force correlation functions (eq. 5.7) for each vibrational mode of a symmetric stretch-excited water molecule in the liquid phase, plotted as a function of frequency. This function of the VER transition frequency is an average over all 12 QM/MM trajectories and corrected with the QCF in eq. (5.8). This QCF is used because it has been claimed that it is appropriate for VER processes between the solute's vibrations and bath's translations and rotations. [33] In our QM/MM simulations, SHAKE was applied to solvent molecules. Therefore the energy accepting states in the bath are restricted to translations, rotations, and their combinations. In the figure, large bands are clearly seen around 2800-4000 cm^{-1} for the two O-H stretching modes while there are smaller bands below 1000 cm^{-1} for all three modes. The high frequencies are in qualitative agreement with the PMA-calculated frequencies for liquid water's O-H stretching modes and the shortest time scale of 0.7 ps occurs at 3500 cm^{-1} for the symmetric stretching mode. Since in our QM/MM simulations the solvent water molecules have a rigid structure, there are no internal vibrational modes that can serve as the energy accepting states in the bath. Therefore, these must be multiple translations and rotations in the bath that

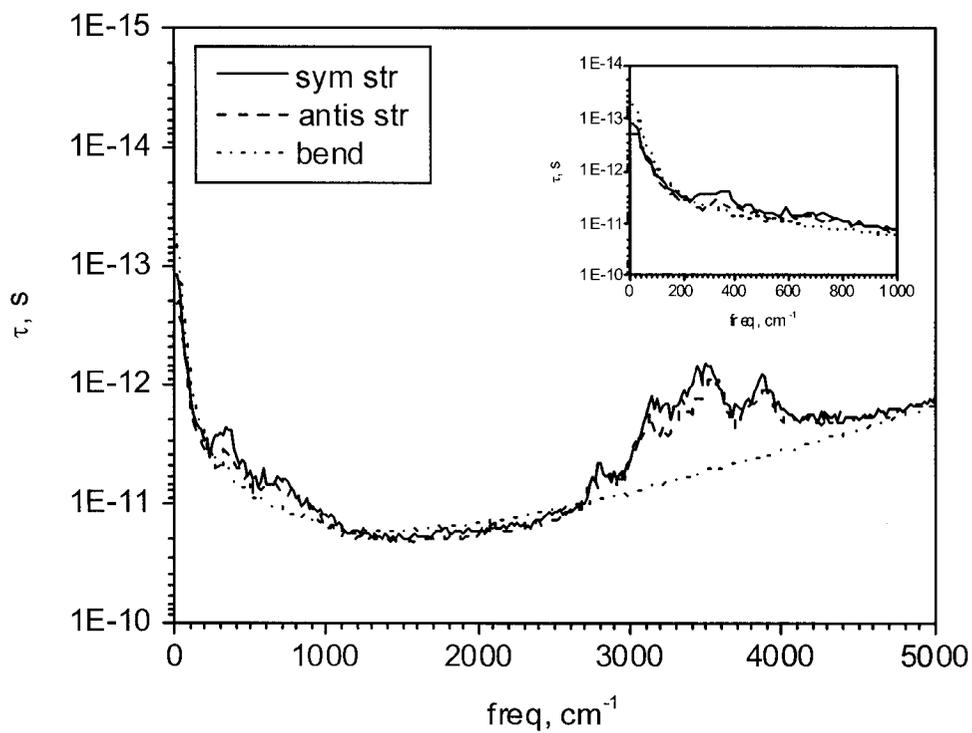


Figure 5.7. The time scales as functions of the transition frequency calculated from Fourier transformed force-force correlation function for liquid water with the symmetric stretching mode excited. The forces are from the solvent acted on each of the solute vibrational coordinates. The inset shows the low frequency part of the functions.

simultaneously participate in these VER processes to account for the high frequency difference. The time scale of 0.7 ps for the symmetric stretch is in good agreement with the result obtained from the previous section (0.62 ± 0.23 ps). This implies that the major VER pathway is the transition from the first excited to the ground symmetric stretch state with a group of translations and rotations of the bath as the energy accepting states. For the antisymmetric stretch, the shortest time scale in this frequency range is found to be 0.9 ps, which is contrary to 0.28 ps obtained by analyzing the total vibrational energy decay. Therefore this discrepancy implies that for the antisymmetric stretch, the transition from the first excited to the ground state may not be a significant contributor to VER. Instead, the transition between the antisymmetric and the symmetric stretch, whose frequency difference is below 100cm^{-1} , according to a fast Fourier transform (FFT) of the velocity autocorrelation function for the solute, is possibly the major path. In Figure 5.7, the time scales for both stretches at this low frequency range are in the 0.2-0.4 ps range indicating fast energy exchange processes involving the two stretching modes. Therefore vibrational energy in the antisymmetric stretch could be preferentially transferred to the symmetric stretch with an excitation of one or more translational or rotational states in the bath at a shorter time scale (0.2-0.4 ps) than the 0.9 ps time scale at 3500cm^{-1} . This path is suggested from analyzing the overall vibrational energy decay in the previous section and is consistent with ultrafast IR-Raman spectroscopy experiments. [17] A similar time scale of 0.4 ps has been reported by this experiment interpreted as spectral diffusion, which is indeed a vibrational energy exchange process between the

two O-H stretching modes. Compared to the time scale for IVR process, 0.38 ± 0.15 ps, obtained by analyzing the energy decay for isolated water molecules, it can be concluded both IVR and VER can well compete with each other at this time scale.

An interesting feature in Figure 5.7 is that the high frequency bands in the 2800-4000 cm^{-1} range are roughly composed of four peaks. The number of peaks matches with the number of solvent water molecules, two hydrogen bond acceptors and two hydrogen bond donors, with a direct contact to the central excited solute water molecule. The slight different local structures where these four solvent molecules are located may be the cause of the four distinguishable peaks.

The low frequency part (below 1000 cm^{-1}) in Figure 5.7 is enlarged in the inset to show more details. For the symmetric stretching mode, a shoulder from 240 to 380 cm^{-1} can be seen in the inset with a time scale of 2.6 ps. Assuming this VER process only happens between the solute and a close solvent water molecule, our previously calculated water dimer results imply that this frequency falls in the range for the hydrogen bonding distortions. [46] However, it should be kept in mind that the liquid water structure may be substantially different from the water dimer in the gas phase and therefore these modes may not be the energy accepting modes. In the solute molecule, on the other hand, this frequency range matches the energy gap between the stretch and the bend overtone considering anharmonicity present in the latter state, which is approximately 200 cm^{-1} by FFT. Therefore, this time scale accounts for the VER process between the symmetric stretch and the bend overtone in the solute. The energy accepting state in the bath possibly includes the hydrogen bonding distortions

but it needs to be further studied. Compared to the 0.7 ps time scale for the transition from the first excited to the ground symmetric stretch, this transition path is much slower and thus does not make a major contribution to VER on the sub-picosecond time scale observed experimentally. However, from analyzing the total vibrational energy in the previous section, a time scale of ~ 0.3 ps, (which is much shorter than 2.6 ps), for energy growth is obtained for the bending mode. This implies that besides the slower VER process, there must be a faster IVR process, which is enhanced by the existence of the solvent probably by increase the coupling between the stretch and bend.

The same VER time scale as a function of transition frequency is also generated for the water molecules with the antisymmetric stretch excited (see Figure 5.8). This plot shows very similar features as seen in Figure 5.7 for the three vibrational modes and therefore is not discussed in detail. The values of the time scales, though, are slightly shorter than the ones obtained from Figure 5.7. At 3500 cm^{-1} , the plot gives 0.4 ps for the antisymmetric stretch, which is shorter than 0.70 ± 0.31 ps given by analyzing the total vibrational energy decay but the two time constants are still within one standard deviation of each other. For the symmetric stretch, the secondarily excited mode in this situation, the time scale is 0.6 ps at 3500 cm^{-1} and 0.1-0.4 ps when the frequency is lower than 100 cm^{-1} . Then the major VER process when the antisymmetric stretch is excited occurs from the first excited state to the ground state for the primarily excited mode, with the energy transferred to multiple external states

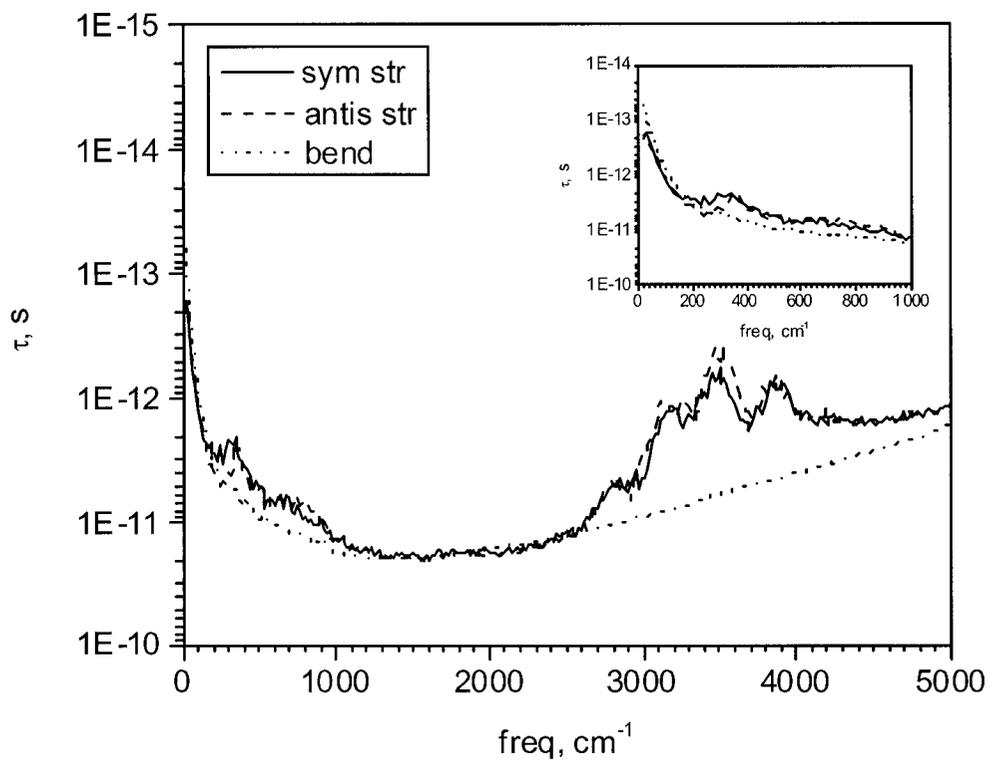


Figure 5.8. The time scales as functions of the transition frequency calculated from Fourier transformed force-force correlation function for liquid water with the antisymmetric stretching mode excited. The forces are from the solvent acting on each of the solute vibrational coordinates. The inset shows the low frequency part of the functions.

of solvent water molecules. For the secondarily excited mode, the energy is likely transferred through a faster path between the two stretches (the spectral diffusion process). The time scale for the stretch-to-bend overtone path is again longer (2 ps) than the 0.7 ps or 0.4 ps time scales indicating it does not significantly contribute to the sub-picosecond VER processes.

In both Figure 5.7 and 5.8, there is no peak at the 1500-1700 cm^{-1} frequency range, where the bending mode is expected to appear. Therefore, direct vibrational energy transfer processes involving the bending mode, e.g., from the first excited state to the ground state of the bend, are insignificant in the simulations reported here, with rigid solvent water molecules.

5.6. Conclusions.

The vibrational energy transfer processes including intramolecular vibrational redistribution (IVR) and vibrational energy relaxation (VER) in water are studied by using principal mode analysis (PMA) to calculate frequencies and vibrational modes from QM/MM trajectories of water molecules. The superior ability to predict frequencies for water both in the gas and liquid phase that PMA has shown illustrates PMA's advantages over the conventional normal mode analysis. By analyzing QM/MM trajectories of water molecules in the gas and liquid phase with specific OH stretching modes excited, PMA sets up an accurate frame of vibrational coordinates for the liquid. Within this frame, the total vibrational energy can be separated along each of the vibrational modes as functions of time. Then the mechanisms and rates of energy transfer are readily studied.

The isolated water molecules and liquid water show very different behaviors in terms of the transfer pathways and characteristic times depending on the excitation conditions and solvent effects. In isolation, the primary IVR process is between the symmetric and the antisymmetric stretching modes at rates of 0.38 ± 0.15 ps and 0.12 ± 0.01 ps, depending on whether the symmetric stretch or the antisymmetric stretch is excited at the beginning of the simulations, respectively. The magnitude of the 0.38 ps energy transfer process is larger than the 0.12 ps process and both of them are more significant than the transfer processes between the stretches and the bend, which has the fastest time scale of 0.07 ± 0.03 ps.

When the same processes are studied in the liquid phase, the perturbation due to the solvent water molecules (primarily hydrogen bonding) strongly alters both the pathways and the time scales. The energy exchange processes between the modes are significantly different from the processes in isolated water while the excess vibrational energy of the excited water molecules constantly flows into the bath's degrees of freedom. Two distinguishable time scales are found for the overall VER processes from the calculation of the vibrational energy as functions of time fitted with monoexponential decay. They are 0.6-0.7 ps for the initially excited stretching mode and 0.2-0.3 ps for the secondarily excited mode, including both the other stretch and the bend. There is also an energy growth time of approximately 0.3-0.4 ps for the two secondarily excited modes that is too short to fit accurately. These time scales agree very well with the values of ultrafast pulse-probe spectroscopic experiments but our interpretation is different than those given in the experimental works.

An approach to VER through the force-force correlation function derived from applying time-dependent perturbation theory to the solute-bath interactions is presented to provide detailed information about the transitions between vibrational states. A Fourier transform is used to calculate the VER time scales as a function of the transition frequency corrected by the quantum-correction factor (QCF) for the vibration-translation and vibration-rotation scheme. A major VER process is identified as the transition from the first excited state to the ground state in the solute for the primarily excited mode. A time scale is calculated as 0.4-0.7 ps for this process and is in good agreement with the total vibrational energy decay results as

well as the experimental value for T_1 . Possible energy accepting states are multiple translations, rotations, or their combinations for the solvent water molecules. For the secondarily excited modes, the energy is transferred back to the primarily excited modes because its time scale, 0.1-0.4 ps, is shorter than a direct transition to the ground state. This time scale is near the same magnitude as the time scale of an IVR process between the two O-H stretches. This suggests that IVR and VER processes are competitive with each other. This time scale is also in the same range as a time scale of 0.4 ps observed in ultrafast IR-Raman spectroscopic studies for spectral diffusion. Our study suggests an agreement with experiment in the interpretation of the process occurring on this time scale. Another possible VER process is the stretch-to-bend overtone path, which is proposed by several spectroscopic studies, but the long time scales (2-3 ps) from our calculation make this process unlikely on a sub-picosecond time scale.

5.7. References.

- [1] Franks, F. *Biophysics of water*; Wiley-Interscience: New York, 1982.
- [2] Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water*; Oxford University: Oxford, 1969.
- [3] Walrafen, G. E. *The Physics and Physical Chemistry of Water*; Plenum: New York, 1972.
- [4] Yardley, J. T. *Introduction to Molecular Energy Transfer*; Academic Press: New York, 1980.
- [5] *Ultrafast Infrared and Raman Spectroscopy*; Fayer, M. D., Ed.; Marcel Dekker: New York, 2001.
- [6] Bakker, H. J. Femtosecond Mid-Infrared Spectroscopy of Water. In *Ultrafast Hydrogen Bonding Dynamics and Proton Transfer Processes in the Condensed Phase*; Elsaesser, T., Bakker, H. J., Eds.; Kluwer Academic Publishers: Dordrecht, 2002.
- [7] Vikhrenko, V. S.; Heidelberg, C.; Schwarzer, D.; Nemtsov, V. B.; Schroeder, J. *J. Chem. Phys.* **1999**, *110*, 5273.
- [8] Keutsch, F. N.; Fellers, R. S.; Brown, M. G.; Viant, M. R.; Petersen, P. B.; Saykally, R. J. *J. Am. Chem. Soc.* **2001**, *123*, 5938.
- [9] Keutsch, F. N.; Saykally, R. J. *Proc. Nat. Acad. Sci. USA* **2001**, *98*, 10533.
- [10] Keutsch, F. N.; Saykally, R. J.; Wales, D. J. *J. Chem. Phys.* **2002**, *117*, 8823.
- [11] Nienhuys, H. K.; Woutersen, S.; van Santen, R. A.; Bakker, H. J. *J. Chem. Phys.* **1999**, *111*, 1494.
- [12] Woutersen, S.; Bakker, H. J. *Nature* **1999**, *402*, 507.
- [13] Lock, A. J.; Woutersen, S.; Bakker, H. J. *J. Phys. Chem. A* **2001**, *105*, 1238.
- [14] Kropman, M. F.; Nienhuys, H. K.; Woutersen, S.; Bakker, H. J. *J. Phys. Chem. A* **2001**, *105*, 4622.
- [15] Bakker, H. J.; Nienhuys, H. K.; Gallot, G.; Lascoux, N.; Gale, G. M.; Leicknam, J. C.; Bratos, S. *J. Chem. Phys.* **2002**, *116*, 2592.

- [16] Lock, A. J.; Bakker, H. J. *J. Chem. Phys.* **2002**, *117*, 1708.
- [17] Deak, J. C.; Rhea, S. T.; Iwaki, L. K.; Dlott, D. D. *J. Phys. Chem. A* **2000**, *104*, 4866.
- [18] Deak, J. C.; Iwaki, L. K.; Rhea, S. T.; Dlott, D. D. *J. Raman Spectrosc.* **2000**, *31*, 263.
- [19] Dlott, D. D. *Chem. Phys.* **2001**, *266*, 149.
- [20] Pakoulev, A.; Wang, Z.; Dlott, D. D. *Chem. Phys. Lett.* **2003**, *371*, 594.
- [21] Wang, Z.; Pakoulev, A.; Pang, Y.; Dlott, D. D. *Chem. Phys. Lett.* **2003**, *378*, 281.
- [22] Vodopyanov, K. L. *J. Chem. Phys.* **1991**, *94*, 5389.
- [23] Graener, H.; Seifert, G. *J. Chem. Phys.* **1993**, *98*, 36.
- [24] Zittel, P. F.; Masturzo, D. E. *J. Chem. Phys.* **1989**, *90*, 977.
- [25] Kenkre, V. M.; Tokmakoff, A.; Fayer, M. D. *J. Chem. Phys.* **1994**, *101*, 10618.
- [26] Oxtoby, D. W. *Adv. Chem. Phys.* **1981**, *47*, 487.
- [27] Adelman, S. A.; Muralidhar, R.; Stote, R. H. *J. Chem. Phys.* **1991**, *95*, 2738.
- [28] Skrebkov, O. V.; Smirnov, A. L. *Chem. Phys.* **1995**, *198*, 297.
- [29] Ladanyi, B. M.; Stratt, R. M. *J. Phys. Chem. A* **1998**, *102*, 1068.
- [30] Deng, Y.; Stratt, R. M. *J. Chem. Phys.* **2002**, *117*, 1735.
- [31] Deng, Y.; Ladanyi, B. M.; Stratt, R. M. *J. Chem. Phys.* **2002**, *117*, 10752.
- [32] Egorov, S. A.; Skinner, J. L. *J. Chem. Phys.* **1996**, *105*, 7047.
- [33] Skinner, J. L.; Park, K. *J. Phys. Chem. B* **2001**, *105*, 6716.
- [34] Lawrence, C. P.; Skinner, J. L. *J. Chem. Phys.* **2002**, *117*, 5827.
- [35] Lawrence, C. P.; Skinner, J. L. *Chem. Phys. Lett.* **2003**, *369*, 472.

- [36] Abseher, R.; Schreiber, H.; Steinhauser, O. *Proteins* **1996**, *25*, 366.
- [37] Bizzarri, A. R.; Cannistraro, S. *J. Phys. Chem. B* **2002**, *106*, 6617.
- [38] Rey, R.; Hynes, J. T. *J. Chem. Phys.* **1996**, *104*, 2356.
- [39] Rey, R.; Hynes, J. T. *J. Chem. Phys.* **1998**, *108*, 142.
- [40] Rey, R.; Moller, K. B.; Hynes, J. T. *J. Phys. Chem. A* **2002**, *106*, 11993.
- [41] Terashima, T.; Shiga, M.; Okazaki, S. *J. Chem. Phys.* **2001**, *114*, 5663.
- [42] Okazaki, S. *Adv. Chem. Phys.* **2001**, *118*, 191.
- [43] Jang, S.; Pak, Y.; Voth, G. A. *J. Phys. Chem. A* **1999**, *103*, 10289.
- [44] Voth, G. A. *Adv. Chem. Phys.* **1996**, *93*, 135.
- [45] Cui, Q.; Karplus, M. *J. Chem. Phys.* **2000**, *112*, 1133.
- [46] Wheeler, R. A.; Dong, H.; Boesch, S. E. *ChemPhysChem* **2003**, *4*, 382.
- [47] Wheeler, R. A.; Dong, H. *ChemPhysChem* **2003**, *4*, 1227.
- [48] Pratt, W. K. *Digital Image Processing*; John Wiley: New York, 1978.
- [49] Fukunaga, K. *Introduction to Statistical Pattern Recognition*; Academic: New York, 1972.
- [50] Morrison, D. F. *Multivariate Statistical Methods*; McGraw-Hill: New York, 1990.
- [51] Pearson, K. *Phil. Mag.* **1901**, *series 2*, 559.
- [52] Hotelling, H. *JEP* **1933**, *24*, 417.
- [53] Harman, H. H. *Modern Factor Analysis*; The University of Chicago Press: Chicago, 1972.
- [54] Golub, G. H.; Van Loan, C. F. *Matrix Computations*; The Johns Hopkins University Press: Baltimore, 1989.
- [55] Gerbrands, J. J. *Pattern Recognition* **1981**, *14*, 375.

- [56] Levy, R. M.; Karplus, M.; Kushick, J.; Perahia, D. *Macromolecules* **1984**, *17*, 1370.
- [57] Brooks, B. R.; Janezic, D.; Karplus, M. *J. Comput. Chem.* **1995**, *16*, 1522.
- [58] Dong, H.; Boesch, S. E.; Wheeler, R. A. *In Preparation* **2003**.
- [59] Amadei, A.; Linssen, A. B.; Berendsen, H. J. *Proteins* **1993**, *17*, 412.
- [60] Case, D. A. *Curr. Opin. Struc. Biol.* **1994**, *4*, 285.
- [61] Hayward, S.; Go, N. *Annu. Rev. Phys. Chem.* **1995**, *46*, 233.
- [62] Heidelberg, C.; Vikhrenko, V. S.; Schwarzer, D.; Schroeder, J. *J. Chem. Phys.* **1999**, *110*, 5286.
- [63] Frankland, S. J.; Maroncelli, M. *J. Chem. Phys.* **1999**, *110*, 1687.
- [64] Pan, R.; Raff, L. M. *J. Phys. Chem.* **1996**, *100*, 8085.
- [65] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science Publications: Oxford, 1987.
- [66] Oppenheim, I.; Shuler, K. E.; Weiss, G. H. *Stochastic Processes in Chemical Physics: The Master Equation*; MIT Press: Cambridge, MA, 1977.
- [67] *Vibrational energy relaxation in liquids and supercritical fluids*; Skinner, J. L.; Egorov, S. A.; Everitt, K. F., Eds.; Marcel, Dekker, Inc.: New York, 2001.
- [68] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- [69] Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
- [70] Jursic, B. S. *J. Mol. Struct.* **1998**, *434*, 29.
- [71] Smith, D. F.; Overend, J. *Spectrochim. Acta A, Part A* **1972**, *28*, 471.
- [72] *The Aldrich Library of FT-IR Spectra*; 1 ed.; Pouchert, C. J., Ed.; Aldrich Chemical Company: Milwaukee, 1985; Vol. 1.
- [73] Liu, K.; Cruzan, J. D.; Saykally, R. J. *Science* **1996**, *271*, 929.

Chapter 6

Summary and prospective research in the application of principal component analysis to molecular vibrations.

Principal component analysis provides an alternative to the conventional normal mode analysis and Fourier transform to calculate vibrational spectra. [1,2] It uses second order moments from a distribution of molecular structures in coordinate space instead of the second derivative of the potential energy with respect to a single structure. The second order moments or covariances form a covariance matrix effectively equivalent to the force constant matrix calculated in normal mode analysis. Following a diagonalization, vibrational frequencies and modes can be obtained from eigenvalues and eigenvectors of the covariance matrix. PCA was previously used to study the non-Gaussian character of a few large scale, low frequency motions of proteins because of its ability to incorporate some anharmonicity. [3,4] However, in these studies the frequencies of other motions were not calculated and the proof of PCA's ability to calculate accurate vibrational frequencies and modes was lacking. By using PCA in the field of molecular vibrations, I propose new procedures in addition to the original PCA methodology. First, the specific amount of vibrational energy in each of the modes instead of the original $k_B T$ (Boltzmann constant times temperature) term is used to obtain frequencies. Next, the translational and rotational motions are removed prior to the construction of the covariance matrix to eliminate their possible mixing with

vibrations. For flexible molecules, PCA is first applied to each fragment of the whole structure. Then the individual fragments' translations, rotations, and vibrations form a basis, in which the whole structure's vibrational modes can be expressed as linear combinations of these fragments' motions. This procedure helps to untangle the mixing between the motions from different fragments of the whole molecule. In order to show the improvement in PCA's performance by using these procedures, tests on water, water dimer, liquid water, methane, and ethane are presented in Chapter 4. The results show that PCA with the new procedures consistently calculates frequencies more accurately than the original PCA method and the conventional normal mode analysis.

PCA accurately calculates vibrational frequencies and modes due to its ability to incorporate anharmonicity. Within this more accurate PCA vibrational frame, processes relating to molecular vibrations such as vibrational energy relaxation VER can be examined. Liquid water VER processes have been an active research area both spectroscopically and computationally because of their central role in solvation, aqueous reaction, and protein folding mechanisms. [5-8] In Chapter 5 I present a VER study by applying PCA to QM/MM trajectories of liquid water. QM/MM simulations can provide a more precise description of the system's potential energy surface than molecular dynamics simulations using empirical force fields. Moreover, some anharmonic character in the energy surface can be accounted for in the PCA frame. The major result presented in Chapter 5 is that two distinguishable time scales of 0.6-0.7 ps and 0.2-0.3 ps in liquid water were obtained for the VER processes

involving energy transfer out of O-H stretching modes. These time scales agree well in magnitude with experiment but an interpretation different from the one given by experiment comes from the analysis of the vibrational energy as a function of time. PCA studies indicate that the longer time scale is attributed to energy transfer out of the initially excited vibrational modes and the shorter time scale is due to energy transfer out of the secondarily excited mode. The main energy transition path is from the first excited state to the ground state of the excited O-H stretching mode with multiple translational and rotational modes of the surrounding water molecules as the energy accepting modes.

Despite the good agreement with experiment and more information about VER in liquid water provided in Chapter 5, this PCA study is only at its preliminary stage. In the QM/MM simulation setup, only one water molecule is treated quantum mechanically for moderate computational cost, so the description of the bath water molecules may not be accurate. Moreover, the bath water molecules are rigid during the simulations. This accelerates the computation but also prevents the bath water molecules' vibrational modes from being the energy accepting modes. Therefore, the next step of this liquid water VER research would include increasing the quantum mechanical region to include more than just one water molecule and eliminating the constraint on the internal coordinates of the bath water in QM/MM simulations. Then these trajectories can be studied using the same methodologies as presented in Chapter 5.

Another possible direction of research in Chapters 4 and 5 will lead to methods better than PCA. Independent component analysis (ICA) is a candidate to replace PCA in the field of molecular vibrations. [9-11] ICA calculates moments higher than second order and separates them as completely as possible. This high-order statistics-based method is particularly suitable for calculating the molecular vibrations including cubic and quartic terms in the system's potential energy and therefore goes beyond the "quasiharmonic" approximation. More accurate vibrational frequencies and modes can be expected from ICA than PCA.

References.

- [1] Levy, R. M.; Karplus, M.; Kushick, J.; Perahia, D. *Macromolecules* **1984**, *17*, 1370.
- [2] Brooks, B. R.; Janezic, D.; Karplus, M. *J. Comput. Chem.* **1995**, *16*, 1522.
- [3] Case, D. A. *Curr. Opin. Struc. Biol.* **1994**, *4*, 285.
- [4] Hayward, S.; Go, N. *Annu. Rev. Phys. Chem.* **1995**, *46*, 233.
- [5] Deak, J. C.; Rhea, S. T.; Iwaki, L. K.; Dlott, D. D. *J. Phys. Chem. A* **2000**, *104*, 4866.
- [6] Lock, A. J.; Bakker, H. J. *J. Chem. Phys.* **2002**, *117*, 1708.
- [7] Rey, R.; Hynes, J. T. *J. Chem. Phys.* **1996**, *104*, 2356.
- [8] Lawrence, C. P.; Skinner, J. L. *J. Chem. Phys.* **2002**, *117*, 5827.
- [9] Comon, P. *Signal Processing* **1994**, *36*, 287.
- [10] Lathauwer, L. D.; Moor, B. D.; Vandewalle, J. *J. Chemometrics* **2000**, *14*, 123.

- [11] *Independent Component Analysis: Principles and Practice*; Roberts, S.; Everson, R., Eds.; Cambridge University Press: Cambridge, UK, 2001.