

**MODELLING AND PREDICTING THE FREE ENERGY AND SOLUBILITY
BEHAVIOR OF LOVASTATIN AND SIMVASTATIN USING MOLECULAR
SIMULATIONS**

by

FABIAN F CASTEBLANCO

A Thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Masters of Science

Graduate Program in Chemical and Biochemical Engineering

written under the direction of

Yee C. Chiew Ph.D.

and approved by

New Brunswick, New Jersey

October, 2013

ABSTRACT OF THE THESIS

MODEL TO PREDICT ACTIVITY COEFFICIENT RATIO AT INFINITE DILUTION

OF LOVASTATIN AND SIMVASTATIN DRUG SOLUTIONS

By

FABIAN F CASTEBLANCO

Thesis Director:
Professor Yee C. Chiew Ph.D

In this master thesis, exploration is done in employing molecular modelling methods to predict drug solubility in liquid solvents, compare the simulation results with experimental data to validate the molecular force field model, and explain certain solubility behaviour of the statin compounds. Experimental data shows that lovastatin solubility increases in a family of alkanols and reaches a peak as the non-polarity nature of the solvent increases from water (polar) to 1-butanol (nonpolar), and the trend reverses from 1-pentanol to 1-octanol. This study investigates this interesting behaviour and provides insight on why this can be occurring. In this study, the CHARMM General Force Field is utilized to model lovastatin and simvastatin in different liquid solvents. The free energy and thermodynamic properties of these systems are calculated using molecular dynamics techniques. Periodic boundary conditions are used with electrostatics treated with Particle-mesh Ewald (PME), using a short-range cutoff of 1.2 nm, while having van der Waals interactions switched off between 1.0 to 1.2 nm. The Bennet Acceptance Ratio (BAR)

method is implemented as a means of estimating the Gibbs free energy of decoupling the drug molecule in the system. We report results obtained from two studies. In Study 1, the free energy of de-coupling of a drug molecule in different liquid solvents is computed. These results yield the ratio of the infinitely dilute activity coefficient and solubility of the drug compounds in two different solvents (including water and a family of alcohols up to 1-octanol) The simulation results are observed to predict a peak in the lovastatin solubility in alkanols, with a peak for 1-butanol. The simulated data is analyzed further by calculating the energies between polar and nonpolar groups between the drug and solvent. In Study 2, we employ a specific thermodynamic model that involves mutating the unique methyl group in simvastatin to hydrogen, essentially converting simvastatin to lovastatin in liquid solvents and vacuum. The free energy of mutation can be shown to be related to the activity coefficients of lovastatin and simvastatin in a solvent. Results from Study 1 showed that the simulated lovastatin solubility ratios are in agreement with experimental data in that a solubility peak in 1-butanol is estimated (Fig1). Further analysis of the energy of interactions of the polar and nonpolar groups between the drug and solvent molecules shows that the nonpolar interactions become stronger with increasing alcohol carbon chain length. However, the interactions between the polar functional groups of the drug and solvent appear to reach a peak in strength at 1-butanol and seem to reverse trend (Fig2). The polar and polar interactions contribution to the overall enthalpy thus changes direction and provides a weakening effect after the organic alcohol chain length surpasses that of 1-butanol. Results obtained from Study 2 show that the simulation provides a reasonably adequate prediction of the activity coefficient ratios between the two drugs within the same solvent. In this study, it was shown that CHARMM General Force Field is a reasonably good model for lovastatin, simvastatin and common liquid solvents. Our simulation study was found to give

good estimates and capture the behavior of the solubility of lovastatin in liquid solvents. Once the drug solvent mixtures were further analyzed and broken into groups based on polarity, analysis on the energy of interactions provides scientific insights and explanation on why solubility of lovastatin reaches a peak with increasing organic alcohol chain length.

ACKNOWLEDGEMENTS

This master thesis is written during the time-period of January 2011 through August 2012 under the teaching supervision of Dr. Yee Chiew from Rutgers University in Piscataway, NJ. The intent of this thesis was to explore molecular dynamics and design a model that would allow to predict certain thermodynamic properties that are crucial to drug solubility. The purpose was to demonstrate that molecular dynamics, a field that continues to advance with computational power advancements, can be used to explain certain phenomena that is physically seen at the lab scale.

I would like to thank Dr. Yee Chiew for his excellent guidance and motivation towards this thesis and his expertise that he openly shared with faculty and staff. I would also like to thank Dr. Alexei Kotelnikov for his help with accessibility to the Mphase computer cluster and his help with Linux-related issues. A great thanks is also noticed for the amazing undergraduate students that I was able to work with for my teaching assistant assignments for thermodynamics as it greatly helped me to understand the importance of teaching and communication. These are skills that greatly contributed to the writing of this thesis.

Table of Contents

Abstract.....	ii
Acknowledgement.....	v
Tables.....	viii
Figures.....	ix
<u>CHAPTER 1</u>	
INTRODUCTION	1
1.1 Applications.....	2
1.2 Methods to Determine Solubility.....	2
1.3 Drugs In Focus.....	4
1.4 Studies Performed.....	8
<u>CHAPTER 2</u>	
THEORY	9
2.1 Applications.....	10
2.2 Molecular Dynamics.....	10
2.3 Force Fields.....	12
2.4 Simulation Run Parameters [7].....	14
2.5 Free Energy Calculations [8].....	15
<u>CHAPTER 3</u>	
METHODS	19
3.1 Study 1 – The Thermodynamic Cycle.....	20
<i>Solid/Liquid Equilibrium</i>	24
3.2 Study 2 – The Thermodynamic Cycle.....	27
<u>CHAPTER 4</u>	
EXPERIMENTAL DATA	29
4.1 Experimental Solubility Charts (298 K).....	30
<u>CHAPTER 5</u>	
RESULTS	33
5.1 Study 1.....	34
5.2 Study 2.....	43
<u>CHAPTER 6</u>	
CONCLUSION AND FUTURE DIRECTION	46
6.1 Study 1.....	47
6.2 Study 2.....	48
6.3 Future Direction.....	48
<u>CHAPTER 7</u>	

LIST OF NOTATIONS	50
<u>CHAPTER 8</u>	
MOLECULE BUILDING	52
<u>CHAPTER 9</u>	
LIST OF CITATIONS	69
<u>CHAPTER 10</u>	
CIRRICULUM VITAE	72

List of Tables

Table 1. Drugs used for the study are known as the ‘statins’. They are powerful lipid-lowering drug compounds used to treat patients with cholesterol health problems.....	5
Table 2. Solvents used for this study ranged from polar solvents such as water or ethanol to largely non-polar solvents such as 1-octanol.....	6
Table 3. Data from NTI-GYABAAH (* = Sun et al).....	31
Table 4. Data from NTI-GYABAAH.....	32
Table 5. Comparison of simulated data to experimental data.....	35
Table 6. Comparison of simulated data to experimental data.....	37
Table 7. Experimental vs Simulated dG values.....	44

Table of Figures

Figure 1. The difference between a macroscopic and microscopic scale.	11
Figure 2. An intermediate state (labeled 1) can improve the degree of overlap in phase space and lead to improved sampling [9].....	16
Figure 3. Example of Intermediate States [8].....	17
Figure 4. Free energy for Coulombic decoupling of Lovastatin in Ethanol test case.....	17
Figure 5. Thermodynamic cycle used in Study 1 to estimate activity coefficient ratio at infinite dilution.....	21
Figure 6. Thermodynamic cycle used for Study 2 to calculate activity coefficient ratios between drugs in a solvent.	27
Figure 7. Solubility ratio (solubility of solvent divided by solubility of 1-butanol) versus the number of carbons on the alcohol starting with water (zero carbons) and ending with 1-octanol (eight carbons). Experimental values are shown in diamonds (orange) while simulated results are in circles (purple).....	36
Figure 8. Solubility ratio (solubility of solvent divided by solubility of 1-butanol) versus various solvents such as ethanol (ETH), acetone (ACE), ethyl acetate (ETA), and water (H ₂ O). Experimental values are shown in diamonds (orange) while simulated results are in squares (purple).....	38
Figure 9. G(r) analysis can help us identify where parts of the drug have strong interactions with surrounding atoms. Shown here is an oxygen on the lovastatin interacting with an hydrogen on an ethanol molecule.....	39
Figure 10. Lovastatin polar groups.....	40
Figure 11. Group Interaction Energies. Three sets have consistent trends. Polar drug and pol solvent has a trend that reverses at 1-butanol.....	42
Figure 12. Sum of all interaction energies produces a concave plot having its minimum at 1-butanol.....	43
Figure 13. Free energy difference of mutation. Y-axis corresponds to the right of Equation I. (X-axis to the left of equation)	45

CHAPTER 1

INTRODUCTION

1.1 Applications

An important challenge in applied thermodynamics that can be used to make crucial decisions from the earliest stages of drug discovery is the prediction of solubility behaviour without the use of experimental data. Solubility is one of the most fundamental physiochemical properties that is particularly useful to a wide variety of applications important to biological, chemical and pharmaceutical industries[1]. In the earlier stages of development, solubility characterization is used to determine whether these compounds are soluble enough for structure-activity relationship screens so predicting solubility earlier in the timeline is beneficial[2]. It is therefore quite useful to be able to model the estimated solubility relevant to other potential solvents. This can also allow us to explain certain solubility trends such as those experienced in the drug compounds known as the ‘statins’.

1.2 Methods to Determine Solubility

Measuring experimental solubility of a solid in a solvent is no easy task as it requires much time to stir the solution to ensure equilibrium was reached. Adding to the patience required is also the fact that high purity crystalline material is needed which might not always be readily available especially for drug screening. The amount of compound required to measure the thermodynamic solubility can be overwhelming during the discovery phase. Kinetic solubility is another method used but it is often misleading as it measure precipitation rate rather than a direct measurement of solubility. These difficulties set up the stage to encourage the use of molecular dynamics to take short cuts in determining relative solubility.

There are various simulation methods out there that have been attempted and some give reasonable results. Quantitative structure property relations (QSPR) are commonly used to predict aqueous solubility, the method based on statistical correlations between properties. Training sets of molecules are used to come up with a statistical correlation that aids in determining estimates for other compounds. This method, although simple and quick, often fails when the compound of interest has a different chemical structure type than those used to create the correlation. There have been many attempts on various methods in molecular dynamics to calculate the free energy of solvation of molecules, some with varying levels of success. Three common techniques used are the Thermodynamic Integration Method, Exponential Reweighting Method, and the Bennett Acceptance Ratio. Thermodynamic Integration is a common technique that compares the difference in free energy between two given states using sampling from molecular dynamics or Monte Carlo simulations. This technique requires integration of the enthalpy state but can give large errors since you are integrating forward or backwards in phase space. The Exponential Method suffers from the same drawbacks as the Thermodynamic Integration Method. The Bennett Acceptance Ratio has a major advantage as it samples various energy states. The advantages of this method are described in Section 2.5

1.3 Drugs In Focus

Lovastatin and simvastatin belong to a class of the most powerful lipid lowering drug compounds that are crucial to rate-limiting cholesterol biosynthesis in the body. Their mode of action is through inhibition of the (3S)-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase[3]. Lovastatin is a natural product, derived from fermentation of *Aspergillus terreus* and is a key raw material for synthesis of simvastatin[3]. In this paper, we employ molecular modelling methods to these specific drugs to predict activity coefficient ratios in liquid solvents. The following step is to compare the simulation results with experimental data to validate the molecular force field model to see if it can help explain certain solubility behaviour of the statin compounds. Experimental data by other several authors conclusively shows that lovastatin solubility increases in a family of alkanols and reaches a peak as the non-polarity nature of the solvent increases from water (polar) to 1-butanol (nonpolar), and the trend reverses from 1-pentanol to 1-octanol. If the simulations can show this same trend, perhaps it can lead to insight on why this solubility behaviour occurs. This investigation dives into molecular dynamics simulations to attempt to observe in simulations the same interesting behaviour occurring through experiment and analyse the data to provides insight on why this solubility behaviour is observed.

Table 1. Drugs used for the study are known as the ‘statins’. They are powerful lipid-lowering drug compounds used to treat patients with cholesterol health problems.

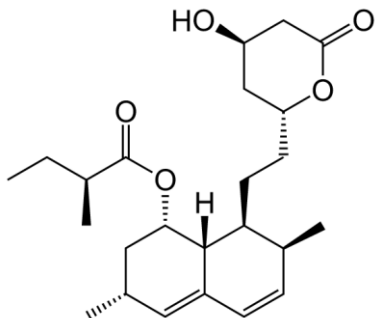
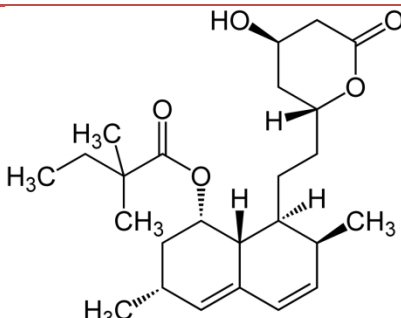
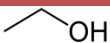


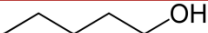
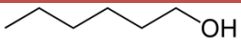
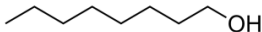
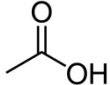

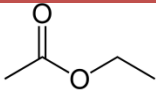
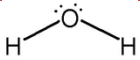
DRUG MOLECULES			
Lovastatin		Formula	C ₂₄ H ₃₆ O ₅
		Mol. mass	404.55 g/mol
Simvastatin		Formula	C ₂₅ H ₃₈ O ₅
		Mol. mass	418.57 g/mol

Table 2. Solvents used for this study ranged from polar solvents such as water or ethanol to largely non-polar solvents such as 1-octanol.

SOLVENT MOLECULES			
Ethanol			Molecular formula $\text{C}_2\text{H}_6\text{O}$
			Molar mass 46.07 g mol^{-1}
			Density 0.789 g/cm^3
1-Propanol			Molecular formula $\text{C}_3\text{H}_8\text{O}$
			Molar mass 60.1 g mol^{-1}
			Density 803.4 mg cm^{-3}
1-Butanol			Molecular formula $\text{C}_4\text{H}_{10}\text{O}$
			Molar mass 74.12 g mol^{-1}
			Density 0.81 g cm^{-3}
1-Pentanol			Molecular formula $\text{C}_5\text{H}_{12}\text{O}$
			Molar mass 88.15 g mol^{-1}
			Density 814.4 mg cm^{-3}
1-Hexanol			Molecular formula $\text{C}_6\text{H}_{14}\text{O}$
			Molar mass $102.17 \text{ g mol}^{-1}$
			Density 813.6 mg cm^{-3}
1-Octanol			Molecular formula $\text{C}_8\text{H}_{18}\text{O}$
			Molar mass $130.23 \text{ g mol}^{-1}$
			Density 0.824 g/cm^3
Acetic Acid			Molecular formula $\text{C}_2\text{H}_4\text{O}_2$

		Molar mass	60.05 g mol ⁻¹
		Density	1.049 g cm ⁻³
Acetone		Molecular formula	C ₃ H ₆ O
		Molar mass	58.08 g mol ⁻¹
		Density	0.791 g cm ⁻³
Ethyl Acetate		Molecular formula	C ₄ H ₈ O ₂
		Molar mass	88.105 g/mol
		Density	0.897 g/cm ³ , liquid
Water		Molecular formula	H ₂ O
		Molar mass	18.01528 g/mol
		Density	1000 kg/m ³ , liquid (4 °C)
			917 kg/m ³ , solid

1.4 Studies Performed

Molecular Dynamics is a multidisciplinary field that employs computer science algorithms and theories from mathematics, physics, and chemistry that allow atoms and molecules to interact virtually for long periods of time under the existing laws of physics [4]. Force fields can be chosen according to the system of study to parameterize inter and intra molecular interactions. In this study, the CHARMM General Force Field is utilized to model lovastatin and simvastatin in different liquid solvents. The free energy and thermodynamic properties of these systems are calculated using molecular dynamics techniques. In Study 1, the free energy of de-coupling of a drug molecule in different liquid solvents is computed. These results yield the ratio of the infinitely dilute activity coefficient and solubility of the drug compounds in two different solvents (including water and a family of alcohols up to 1-octanol). In Study 2, we employ a specific thermodynamic model that involves mutating the unique methyl group in simvastatin to hydrogen, essentially converting simvastatin to lovastatin in liquid solvents and vacuum. The free energy of mutation can be shown to be related to the activity coefficients of lovastatin and simvastatin in a solvent.

CHAPTER 2

THEORY

2.1 Applications

There are several different types of applicable theories and methods that must be understood prior to discussing the thermodynamic methods used to obtaining results. The first discusses the role of molecular dynamics in the experiments. Understanding the importance of the force fields chosen is also discussed as well as the theory behind free energy calculations.

2.2 Molecular Dynamics

Through the use of statistical and classical mechanics in a molecular dynamics (MD) simulation one can achieve an accurate estimate of the Gibbs decoupling free energy which is required to estimate the activity coefficient ratio. Molecular dynamics simulations generate information at the microscopic level, including atomic positions and velocities. The conversion of this microscopic information to macroscopic observables such as pressure, energy, heat capacities, etc., requires statistical mechanics. With molecular dynamics simulations, one can study both thermodynamic properties and/or time dependent (kinetic) phenomenon. [4]. The classical mechanics are what describes the physical motion of each particle using Newton's laws of motion as the governing equations to determine past and future particle positions.

Statistical Mechanics - Definitions [4]

The thermodynamic, or macroscopic, state of a system is usually defined by a small set of parameters, for example, the temperature, T , the pressure, P , and the number of particles, N . Other thermodynamic properties may be derived from the equations of state and other fundamental thermodynamic equations.

The mechanical, or microscopic, state of a system is defined by the atomic positions, q , and momenta, p ; these can also be considered as coordinates in a multidimensional space called phase space. For a system of N particles, this space has $6N$ dimensions. A single point in phase space, denoted by G , describes the state of the system. An ensemble is a collection of points in phase space satisfying the conditions of a particular thermodynamic state. A molecular dynamics simulation generates a sequence of points in phase space as a function of time; these points belong to the same ensemble, and they correspond to the different conformations of the system and their respective momenta.

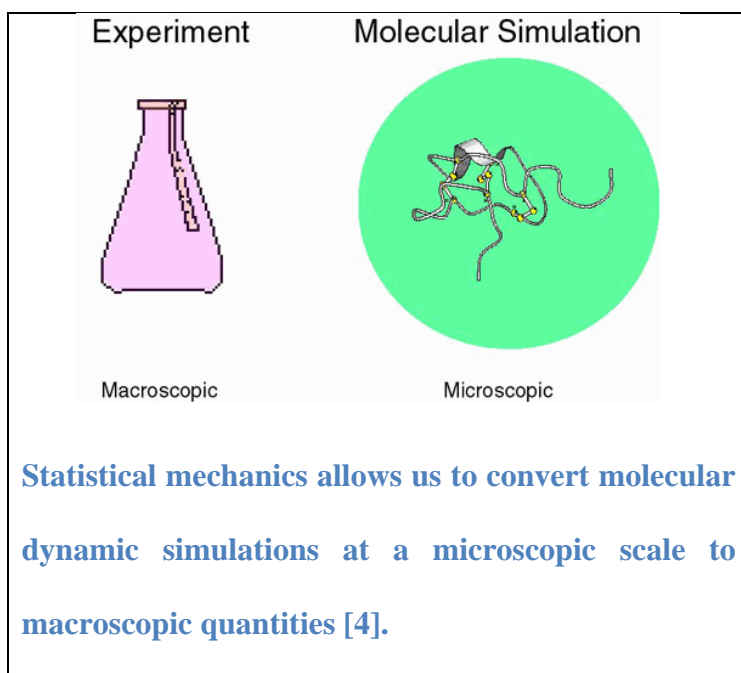


Figure 1. The difference between a macroscopic and microscopic scale.

An ensemble is a collection of all possible systems which have different microscopic states but have an identical macroscopic or thermodynamic state. Specifically for this paper, we are looking at typical atmospheric conditions so an Isobaric-Isothermal Ensemble (NPT) is selected which is characterized by a fixed number of atoms, N , a fixed pressure, P , and a fixed

temperature, T . A simulation must be long enough to sample the possible representative conformations of the system.

Classical Mechanics - Definitions [4]

The molecular dynamics simulation method is based on Newton's second law or the equation of motion, $F=ma$, where F is the force exerted on the particle, 'm' is its mass and 'a' is its acceleration. From knowledge of the force on each atom, it is possible to determine the acceleration of each atom in the system. Integration of the equations of motion then yields a trajectory that describes the positions, velocities and accelerations of the particles as they vary with time.

MD simulations solve Newton's equations of motion for a system of N interacting atoms:

$$m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = \mathbf{F}_i, \quad i = 1 \dots N. \quad [5]$$

2.3 Force Fields

Force fields (FFs) provide approximate parameters for force calculations and are not really part of the simulation method but will affect the results. They can be user-modified but mainly are chosen based on the desired system of study. GROMACS is a user-friendly engine capable of performing molecular dynamics simulations using its very own GROMOS-96 FF. GROMACS also has the capability of running several other popular force fields such as AMBER, OPLSAA, CHARMM, etc. which allows users the flexibility of GROMACS to run simulations by choosing from among the various different force fields available.

The CHARMM General Force Field (CGenFF) was chosen for this study as it provides an organic force field explicitly aimed at simulating drug-like molecules in a biological environment represented by the CHARMM additive biomolecular force fields [6]. Essentially, CGenFF uses the same potential energy functions as the other CHARMM FFs and the same recommendations apply but it is optimized to represent tested drug interactions. For this test case, the system will be tested on a combination of CGenFF(drug and organic solvent). The drug molecule lovastatin and simvastatin were represented by the several CGenFF atom types available but several parameters were left unrepresented and had to be estimated given other similar parameters or omitted if none could be found. This does present certain possibilities for error; however, the assumption is that the error will cancel out by the use of solubility ratios and also by realizing that the drug molecule will be the same for all solvents.

The form of the CHARMM potential energy function used to calculate $V(r)$, where 'r' represents the Cartesian coordinates of the system:

<p>Intramolecular (internal, bonded terms)</p> $\sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\phi(1 + \cos(n\phi - \delta)) + \sum_{\substack{\text{improper} \\ \text{dihedrals}}} K_\varphi(\varphi - \varphi_0)^2 + \sum_{\text{Urey-Bradley}} K_{UB}(r_{1,3} - r_{1,3,0})^2$
<p>Intermolecular (external, nonbonded terms)</p> $\sum_{\text{nonbonded}} \frac{q_i q_j}{4\pi D r_{ij}} + \varepsilon_{ij} \left[\left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^6 \right] \quad [6]$

The intramolecular portion of the potential energy function includes terms for the bonds, valence angles, torsion or dihedral angles, improper dihedral angles and a Urey-Bradley 1,3-term, where b_0 , θ_0 , φ_0 , and $r_{1,3,0}$ are the bond, angle, improper, and Urey-Bradley equilibrium

terms, respectively, 'n' and 'd' are the dihedral multiplicity and phase and the K's are the respective force constants. The intermolecular terms include electrostatic and van der Waals (vdW) interactions, where q_i and q_j is the partial atomic charge of atom i and j, respectively, e_{ij} is the well depth, $R_{\min,ij}$ is the radius in the Lennard-Jones (LJ) 6–12 term used to treat the vdW interactions, and r_{ij} is the distance between i and j [6].

2.4 Simulation Run Parameters [7]

Choosing the run parameters is crucial to using the force field in most optimized way which can be achieved by utilizing the similar conditions that were used when developing CHARMM. Steepest descent was used for the initial molecular energy minimization. Electrostatics are treated with Particle-mesh Ewald (PME), using a short-range cutoff of 1.2 nm, while having van der Waals interactions switched off between 1.0 to 1.2 nm. V-rescale is used for the temperature coupling as well as periodic boundary conditions and isotropic pressure-coupling with the Parrinello-Rahman barostat. A suggested temperature and pressure coupling constant is given as 1 ps. Langevin dynamics, or stochastic dynamics (SD) should be used as they are required for proper sampling of the (nearly) decoupled state. Simulations were run using a 2 fs time step with the neighbor list being updated at least every 20 fs [7].

2.5 Free Energy Calculations [8]

Free Energy Calculations are unfortunately a difficult quantity to obtain for systems such as liquids because the associated quantities such as entropy and chemical potential are very difficult to calculate given that it requires adequate sampling from higher-energy regions whereas molecular dynamics sampling seeks out the lower-energy regions of phase space [9]. These calculations can be done using a variety of methods but the latest GROMACS v4.5 utilizes the latest tool function, *g_bar*, which relies on the Bennett Acceptance Ratio (BAR) method for calculating free energy differences. This method was tested against the older Thermodynamic Integration (TI) approach using the methane in water experiment by *Shirts et al.* The BAR method estimation strategy for free energy differences between two canonical ensembles depends on the extent of overlap between the two ensembles and on the smoothness of the density-of-states as a function of the difference potential [10]. Using ensembles that have sufficient overlap between their phase space will lead to improved sampling.

There are two types of non-bonded interactions that need to be fully decoupled in order to get the Gibbs free energy of the drug being decoupled from its interactions with its surrounding solvent molecules. The Coulombic (long-range) forces need to first be decoupled from a state where the drug molecule is completely charged (COU: on, LJ: on) to a state where they are completely uncharged (COU: off, LJ :on). The same needs to be done with the van der Waal (short-range) interactions where an uncharged state (COU: off, LJ: on) completely loses any van der Waal interaction until it is essentially a dummy drug molecule (COU: off, LJ: off).

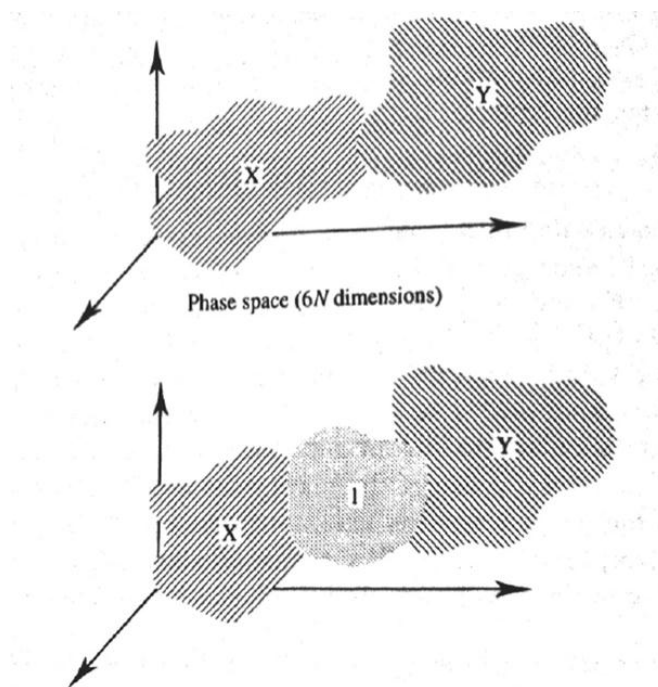


Figure 2. An intermediate state (labeled 1) can improve the degree of overlap in phase space and lead to improved sampling [9].

To allow for better degree of overlap between phase spaces, one introduces intermediate states that allow for more accurate free energy values between smaller differences in states. As in the figure above, the top portion shows two separate states X and Y that have no intermediate state which results in poor overlap between phase spaces which gives inaccurate values for the free energy. By introducing the intermediate in the bottom portion of the figure above, you allow better overlap resulting in accurate estimates of free energy. By utilizing this technique, the Coulombic and LJ decoupling simulations are broken down into 20 intermediate states of $\Lambda=0, 0.05, 0.1, \dots, 0.95, \text{ and } 1.0$ to allow sufficient phase space overlap.

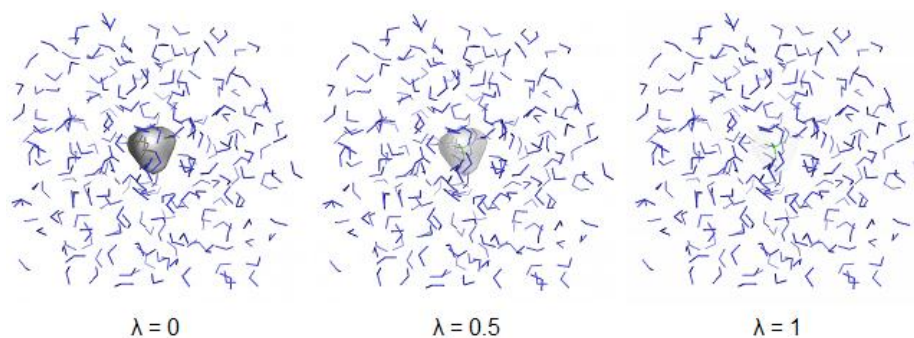


Figure 3. Example of Intermediate States [8].

For each intermediate state, the output will result in a portion of the overall free energy. Using the GROMACS function *g_bar*, one can obtain several plots that show the overall free energy for the entire decoupling simulation as long as all simulation ran to completion, from Lambda 0 to Lambda 1. An example of this result is shown in the figure below.

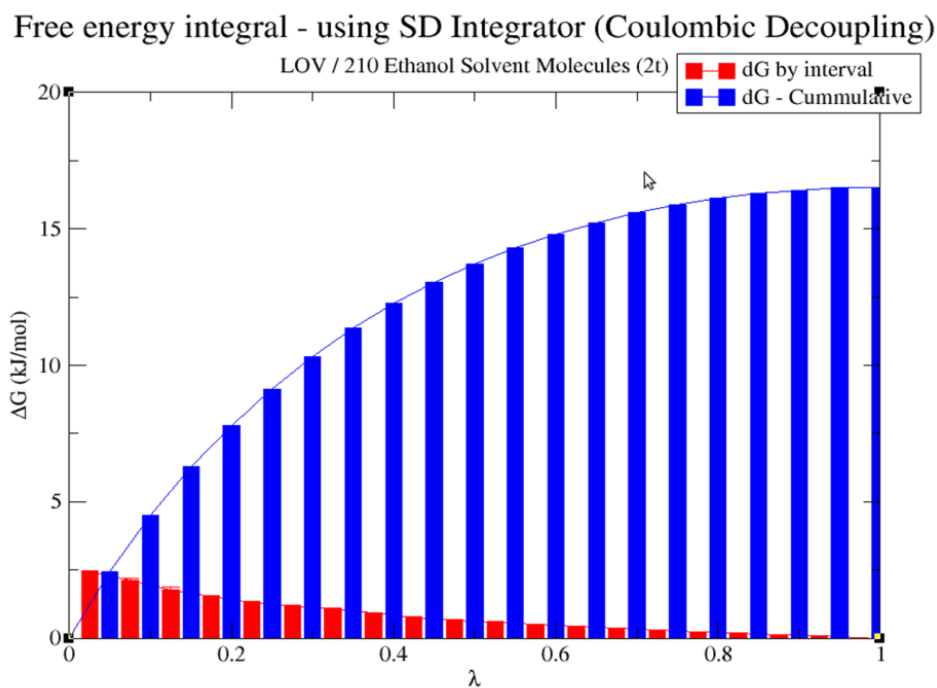


Figure 4. Free energy for Coulombic decoupling of Lovastatin in Ethanol test case.

The last remaining output of *g_bar* is the histogram which will show the extent of overlap between intermediate states. This can be used to validate that sufficient overlap occurred on all simulations. Once both Coulombic and van der Waal decoupling free energies are found, the sum will give you the overall decoupling Gibbs free energy for the drug molecule in a specific solvent, essentially providing the values of ΔG_{13} and ΔG_{24} in the thermodynamic cycle in Figure 5.

CHAPTER 3

METHODS

3.1 Study 1 – The Thermodynamic Cycle

In order to link the molecular dynamics modeled data to real world experimental values, a thermodynamic cycle must be created to allow reasonable validation of our model. This thermodynamic cycle will allow direct validation by means of using experimental values obtained from various experiments from other authors, to the model developed using molecular dynamics. Both studies have similar yet independent thermodynamic cycles and by calculating free energies, their values can be used to find the thermodynamic values in question.

Gibbs free energy is a state function, a property whose value does not depend on the path taken to reach the specific value. If a thermodynamic cycle can be achieved incorporating feasible paths that can be performed on an MD simulation, one can estimate the activity ratio coefficient. The following is a suggested path given under the assumption that we are dealing with an infinitely diluted solution and given an experimental solubility value from the literature. The specified case below is for a drug in two different types of solvents, with ΔG_{13} and ΔG_{34} both representing the Gibbs free energy of the drug being completely decoupled from its surrounding solvent molecules.

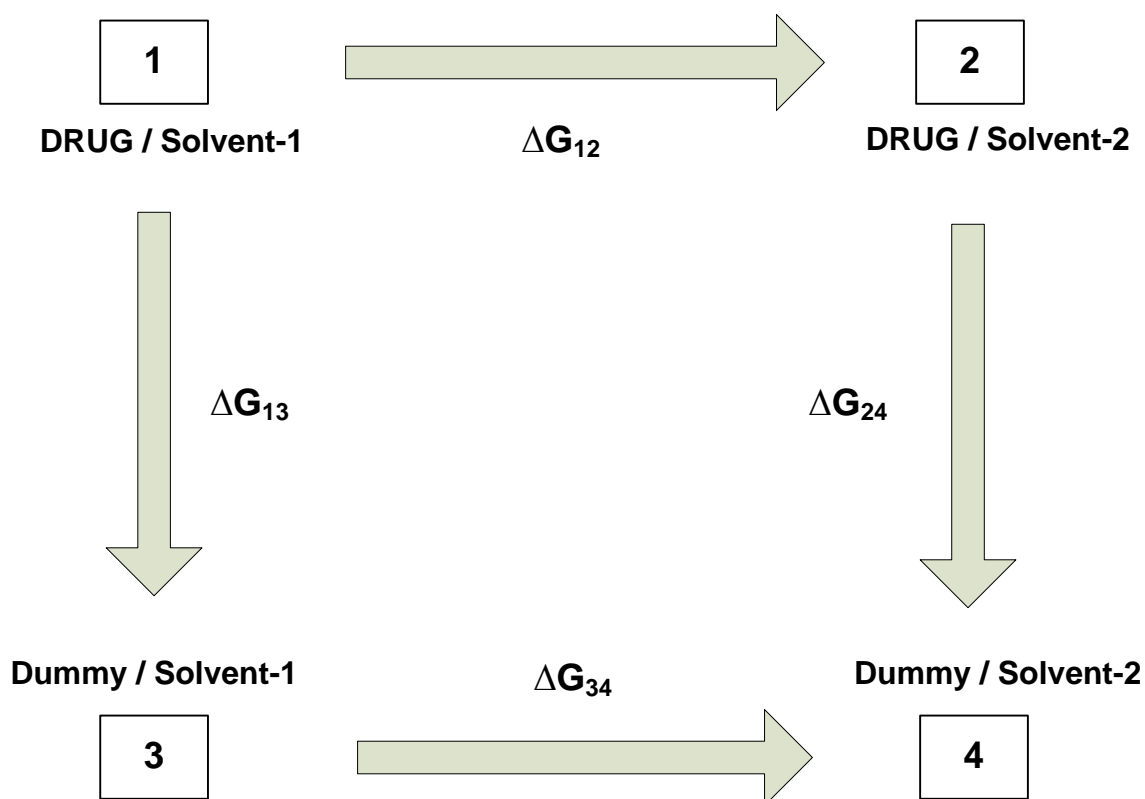


Figure 5. Thermodynamic cycle used in Study 1 to estimate activity coefficient ratio at infinite dilution.

where the thermodynamic cycle states that...

$EQ(A)$	$\Delta G_{12} + \Delta G_{24} = \Delta G_{13} + \Delta G_{34}$
---------	---

Sample Case –

- **DRUG:** Lovastatin
- **Solvent 1:** Ethanol
- **Solvent 2:** Propanol

The change in Gibbs free energy between the two states can be expressed as an equation with the sum of the partial molar Gibbs for each mixture.

$EQ(B)$	$\Delta G_{12} =$ $[N_{LOV} * \overline{G}_{LOV/pro} + N_{pro} * \overline{G}_{pro/LOV}] - [N_{LOV} * \overline{G}_{LOV/eth} + N_{eth} * \overline{G}_{eth/LOV}]$
$EQ(C)$	$\Delta G_{34} =$ $[N_{Dum} * \overline{G}_{Dum/pro} + N_{pro} * \overline{G}_{pro/Dum}] - [N_{Dum} * \overline{G}_{Dum/eth} + N_{eth} * \overline{G}_{eth/Dum}]$

Under the following limiting cases, the following is true:

Limit as $X_{LOV} \rightarrow 0$, $X_{Dum} \rightarrow 0$, & $X_{eth} \rightarrow 1$, $X_{pro} \rightarrow 1$ which is ...Infinite dilution[∞]

(1) $\lim \Delta G_{12} =$

$$[N_{LOV} * \overline{G}^{\infty}_{LOV/pro} + N_{pro} * g_{pro} - N_{LOV} * \overline{G}^{\infty}_{LOV/eth} - N_{eth} * g_{eth}]$$

(2) $\lim \Delta G_{34} =$

$$[N_{Dum} * \overline{G}^{\infty}_{Dum/pro} + N_{pro} * g_{pro} - N_{Dum} * \overline{G}^{\infty}_{Dum/eth} - N_{eth} * g_{eth}]$$

Solving for the difference between the two limits gives us the following:

$\lim [\Delta G_{12} - \Delta G_{34}] =$	$[N_{LOV} * \overline{G}^{\infty}_{LOV/pro} + N_{pro} * g_{pro} - N_{LOV} * \overline{G}^{\infty}_{LOV/eth} - N_{eth} * g_{eth}]$ $- [N_{Dum} * \overline{G}^{\infty}_{Dum/pro} + N_{pro} * g_{pro} - N_{Dum} * \overline{G}^{\infty}_{Dum/eth} - N_{eth} * g_{eth}]$
--	--

Simplifying further by subtracting like terms leads to:

$EQ(C)$	$\lim [\Delta G_{12} - \Delta G_{34}] = [\overline{G}^{\infty}_{LOV/pro} - \overline{G}^{\infty}_{LOV/eth}] - [\overline{G}^{\infty}_{Dum/pro} - \overline{G}^{\infty}_{Dum/eth}]$
---------	--

‘Dummy’ molecules do not interact with the solvent molecules ‘eth/pro’ so it’s free energy is independent on the solvent present.

EQ(D)	$\overline{G}^{\infty}_{\text{Dum/pro}} = \overline{G}^{\infty}_{\text{Dum/eth}}$
-------	---

Using Equation D, one can simplify to the following:

EQ(E)	$\lim [\Delta G_{12} - \Delta G_{34}] = [\overline{G}^{\infty}_{\text{LOV/pro}} - \overline{G}^{\infty}_{\text{LOV/eth}}]$
-------	--

The partial molar Gibbs thermodynamic property equations as a function of infinitely diluted activity coefficients are as follows:

- $\overline{G}_{\text{LOV/pro}}(T, P, X_{\text{LOV}}, X_{\text{pro}}) = g_{\text{LOV}}(T, P) + kT * \ln(\gamma^{\infty}_{\text{LOV/pro}} * X_{\text{LOV}})$
- $\overline{G}_{\text{LOV/eth}}(T, P, X_{\text{LOV}}, X_{\text{eth}}) = g_{\text{LOV}}(T, P) + kT * \ln(\gamma^{\infty}_{\text{LOV/eth}} * X_{\text{LOV}})$

Therefore,

$$\bullet \quad \overline{G}_{\text{LOV/pro}} - \overline{G}_{\text{LOV/eth}} = kT * \ln\left(\frac{\gamma^{\infty}_{\text{LOV/pro}} * X_{\text{LOV}}}{\gamma^{\infty}_{\text{LOV/eth}} * X_{\text{LOV}}}\right) = kT * \ln\left(\frac{\gamma^{\infty}_{\text{LOV/pro}}}{\gamma^{\infty}_{\text{LOV/eth}}}\right)$$

Since Equation A can be rewritten as, $\Delta G_{12} - \Delta G_{34} = \Delta G_{13} - \Delta G_{24}$, the following can be simplified as follows:

$EQ(F)$	$\Delta\Delta G = [\Delta G_{12} - \Delta G_{34}] = [\Delta G_{13} - \Delta G_{24}] = [\overline{G}^{\infty}_{LOV/pro} - \overline{G}^{\infty}_{LOV/eth}]$ $= kT * \ln\left(\frac{\gamma^{\infty}_{LOV/pro}}{\gamma^{\infty}_{LOV/eth}} \right)$
---------	---

The values of ΔG_{13} , ΔG_{24} can be found directly through a molecular dynamics simulation. Each ΔG can be found by converting a lovastatin molecule into a ‘dummy’ molecule, completely devoid of electrostatic and Lennard Jones interactions, by slowly decoupling the solute until it has no remaining interactions with its surroundings.

Rearranging Equation F to calculate the ratio of the infinitely diluted activity coefficient ratio,

$EQ(G)$	$e^{\frac{\Delta\Delta G}{kT}} = \frac{\gamma_{LOV/pro}}{\gamma_{LOV/eth}}$
---------	---

Solid/Liquid Equilibrium

Treating the mixture between the drug solute and the organic alcohol solvents as a solid/liquid mixture in equilibrium, thermodynamics states that the fugacities must be equal in each phase.

$$\hat{f}_i^{\alpha} = \hat{f}_i^{\beta}$$

The symbols ‘ α ’ and ‘ β ’ represent the solid and liquid phases respectively. Since ‘ α ’ is a pure species solid lovastatin and ‘ β ’ is a mixture of lovastatin infinitely diluted into an organic alcohol as a liquid, the following is true for the fugacities of lovastatin at equilibrium:

$$f_{LOV}^S = \hat{f}_{LOV}^L$$

The fugacity of the lovastatin liquid mixture can be described as a function of mole fraction and the activity coefficient.

$$f_A^S(T, P) = \hat{f}_A^L = X_A^L * \gamma_{A/S1}^L(T, X_{A/S1}) * f_A^L(T, P)$$

For each solvent mixture in the case mentioned in the sample case, the equations are as follows:

$$f_{LOV}^S = X_{LOV}^L * \gamma_{LOV/eth}^L * f_{LOV}^L$$

$$f_{LOV}^S = X_{LOV}^L * \gamma_{LOV/pro}^L * f_{LOV}^L$$

Setting the equations equal to each other results in the following:

$$\frac{f_{LOV}^S}{f_{LOV}^L} = X_{LOV}^L * \gamma_{LOV/eth}^L = X_{LOV}^L * \gamma_{LOV/pro}^L$$

Relating it to Equation G above, we can relate the activity coefficient ratio to the solubility of lovastatin in each solvent. At infinite dilution, the following equation relates the thermodynamic cycle to the solubility in each solution:

$EQ(H)$	$\frac{X_{LOV/eth}}{X_{LOV/pro}} = \frac{\gamma_{LOV/pro}^{\infty}}{\gamma_{LOV/eth}^{\infty}}$
---------	---

As mentioned in the thermodynamic cycle section, the activity coefficient at infinite dilution can be found through the simulated MD runs by using Equation G. A model must be set to most accurately represent the solute-solvent interactions in order to best estimate the values of ΔG_{13} , ΔG_{24} which will best approximate the activity coefficient ratios. Experimentally this model can be validated by utilizing Equation H. The mole fraction solubility ratio can be based off experimental data to validate that the activity coefficient is within reason.

3.2 Study 2 – The Thermodynamic Cycle

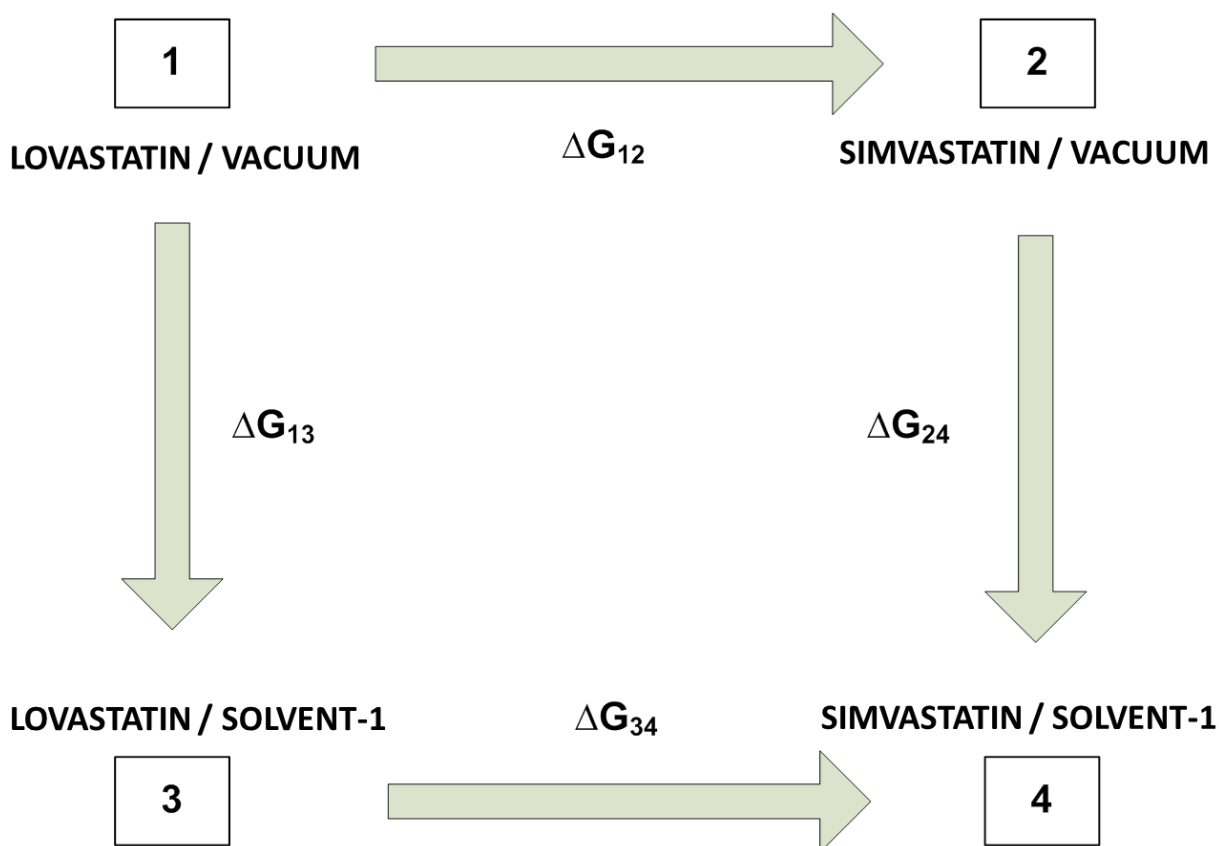


Figure 6. Thermodynamic cycle used for Study 2 to calculate activity coefficient ratios between drugs in a solvent.

Study 2 employs a similar thermodynamic cycle except ΔG_{12} describes the mutation of lovastatin to simvastatin in a vacuum, while ΔG_{34} describes the mutation of lovastatin to simvastatin in one specific solvent. The first simulation of ΔG_{12} is done by dividing the simulation into three steps. The first step removes the charge on the hydrogen group on the

lovastatin; the second step mutates the molecule from an hydrogen group to a methyl group; the third step recharges the methyl group essentially completing the transformation from lovastatin to simvastatin. The same process is repeated for ΔG_{34} in solvent. The same theory is applied resulting in Equation I:

$EQ(I)$	$\Delta\Delta G = [\Delta G_{12} - \Delta G_{34}] = kT * \ln(\frac{\gamma_{LOV}^{\infty}}{\gamma_{SIM}^{\infty}})$
---------	--

Instead of comparison between the activity coefficients of lovastatin in different solvents, this equation yields a comparison between the coefficients of lovastatin and simvastatin within the same solvent. Again, this value is compared to existing experimental data to validate that the simulation gives reasonable results.

CHAPTER 4

EXPERIMENTAL DATA

4.1 Experimental Solubility Charts (298 K)

The following is experimental solubility measurements for the drug and solvent mixtures. All temperatures are at 298 K and have been interpolated using the following empirical equation:

$EQ(H)$	$\ln x = A + \frac{B}{(T/K)} + C \ln(T/K)$
---------	--

The equation was fitted on MATLAB using all the solubility data points provided from the literature. In some cases, the literature already provided the A, B, and C parameters or the solubility was given at 298 K. The existing results were as follows.

LOVASTATIN

Table 3. Data from NTI-GYABAAH (* = Sun et al)

SOLVENT	$X_{LOV} \times 10^3$
Ethanol*	3.390224
Ethanol	3.449803
Propanol	5.900127
Butanol	6.424681
Pentanol	4.453902
Hexanol	4.397839
Octanol	4.029594
Acetone*	12.22976
Ethyl_Acetate*	5.544865

SIMVASTATIN

Table 4. Data from NTI-GYABAAH

SOLVENT	$X_{SIM} \times 10^3$
Ethanol	19.10651
Propanol	23.77167
Butanol	24.8447
Pentanol	28.70269
Hexanol	33.00223
Octanol	28.24027

The experimental solubility data provided above will provide an accurate validation of the model as the solubility approaches zero since at low solubility the probability of infinite dilution is much higher. If the solubility were to be high, then the model cannot be accurately used as they provide estimates of infinitely diluted activity coefficients. Looking at the data above, the model would be of most accuracy for lovastatin solutions since they are all well below 0.006. The solubility data provided will determine the values on the left-hand side of Equation H.

CHAPTER 5

RESULTS

5.1 Study 1

Results from Study 1 showed that the simulated lovastatin solubility ratios are in agreement with experimental data in that a solubility peak in 1-butanol is observed (Table 5. Comparison of simulated data to experimental data.). Experimental and simulated solubility ratios are found in Table 5. The table provides the experimental mole fraction solubility of lovastatin in various solvents over a reference solvent of 1-butanol (since experimentally it should be the solvent where lovastatin is most soluble). The Gibbs free energy differences are also listed as what value should be found through simulation based on experimental data as well as the observed simulated value (taken from the same reference of 1-butanol). This can be found using Equation F to solve for the $\Delta\Delta G_{\text{exp}}$ value that is required for an exact validation of the model given no errors, making it easier for comparison against simulated results. As can be seen in the last column, the simulated solubility ratios do not give values near those given by previous experiment; however, they still seem to predict the same overall trend as seen in Figure 7. Solubility ratio (solubility of solvent divided by solubility of 1-butanol) versus the number of carbons on the alcohol starting with water (zero carbons) and ending with 1-octanol (eight carbons). Experimental values are shown in diamonds (orange) while simulated results are in circles (purple).

Table 5. Comparison of simulated data to experimental data.

LOVASTATIN							
<u>SOLVENT</u>	<u>#C</u>	<u>X_{LOV} × 10³</u>	<u>x_i/x₄</u>	<u>x_i/x₄</u>	<u>ΔΔG_{exp}</u>	<u>ΔΔG_{sim}</u>	<u>% Error</u>
			Experimental	Simulation	(J/mol)	(J/mol)	(based on solubility ratio)
Ethanol*	2	3.3902	0.5277	0.8824	-1584	-310	67.2
Ethanol	2	3.4498	0.5370	0.8824	-1541	-310	64.3
Propanol	3	5.9001	0.9184	0.9880	-211	-30	7.6
Butanol	4	6.4247	1.0000	1.0000	0	0	0.0
Pentanol	5	4.4539	0.6932	0.1770	-908	-4290	-74.5
Hexanol	6	4.3978	0.6845	0.1633	-939	-4490	-76.1
Octanol	8	4.0296	0.6272	0.1904	-1156	-4110	-69.6

The figure below shows that in using an appropriate force field and carefully modeling the drug to match its real world physical structure, a trend can be seen to mimic the experimental data of solubility ratios. Much of the actual data will rely on careful choosing of the parameters that will model the drug and solvents. The free energy calculations done by the *g_{bar}* method are very sensitive to the values of the parameters that are chosen.

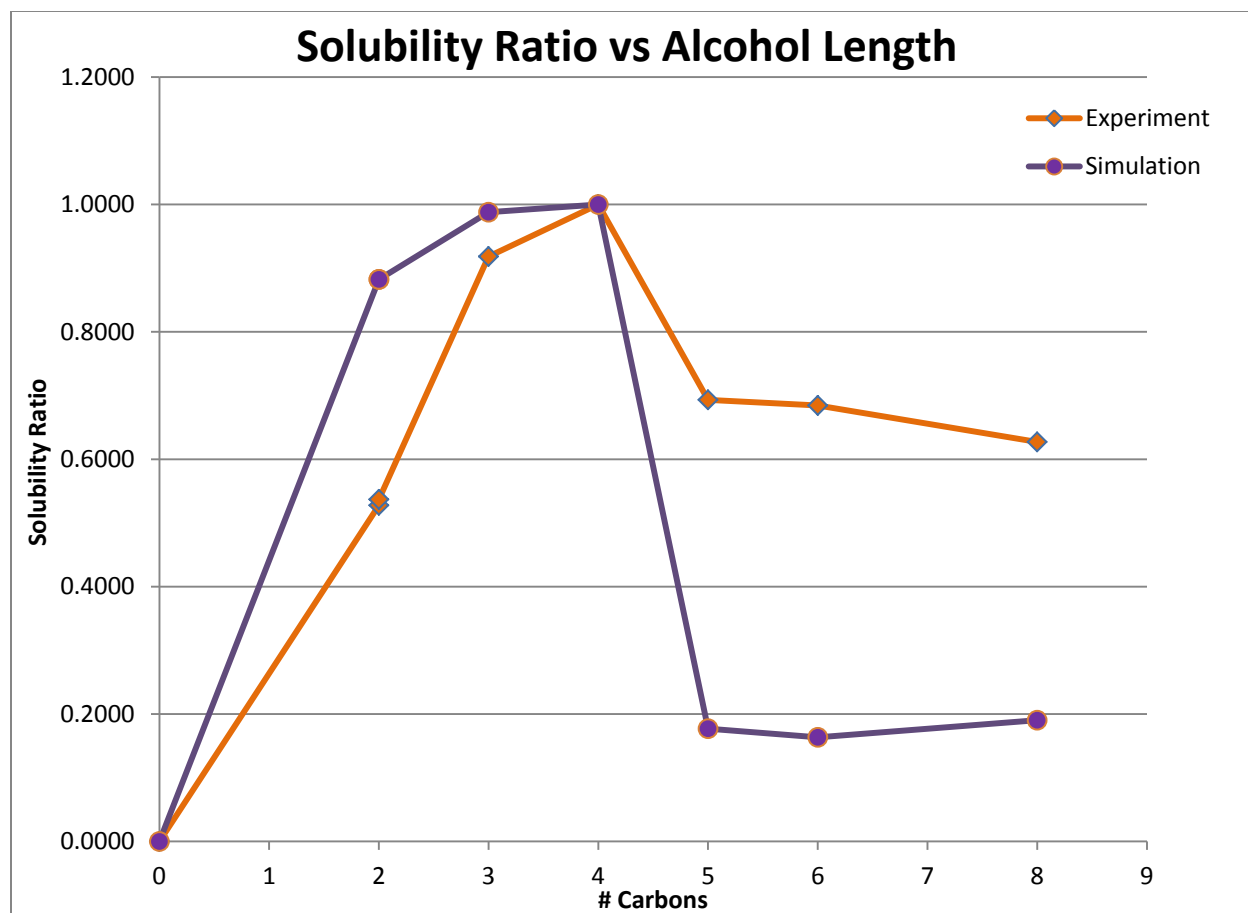


Figure 7. Solubility ratio (solubility of solvent divided by solubility of 1-butanol) versus the number of carbons on the alcohol starting with water (zero carbons) and ending with 1-octanol (eight carbons). Experimental values are shown in diamonds (orange) while simulated results are in circles (purple).

Similarly, results for other solvents were found as well to follow the similar experimental trends as acetone is observed to be more soluble relevant to various other solvents. Lovastatin is also observed to be nearly insoluble in water as is observed experimentally. It is also noted that as experimental solubility is larger, the less likely the model will predict the solubility ratio since it is valid for infinitely diluted solutions which is why various other solvents are omitted. Some values of the free energy did tend to oscillate due to the large nature of the drug and it's

decoupling from its surroundings and the possible parameter estimates while creating the topology of the drug mentioned previously. This is the main reason the solubility ratios are not in the best of agreement with experimental data but overall the consistent trends were of most importance.

Table 6. Comparison of simulated data to experimental data.

LOVASTATIN							
<u>SOLVENT</u>	<u>#C</u>	$X_{LOV} \times 10^3$	x_i/x_4	x_i/x_4	$\Delta\Delta G_{exp}$	$\Delta\Delta G_{sim}$	% Error
			Experimental	Simulation	(J/mol)	(J/mol)	(based on solubility ratio)
ETHANOL*	2	3.3902	0.2772	0.4099	-3179	-2210	-47.8
ACETONE	2	12.2298	1.0000	1.0000	0	0	0.0
ETA	3	5.5449	0.4534	0.0610	-1960	-4455	-86.5
WATER	4	~0	-	0.0000	-	-39360	-

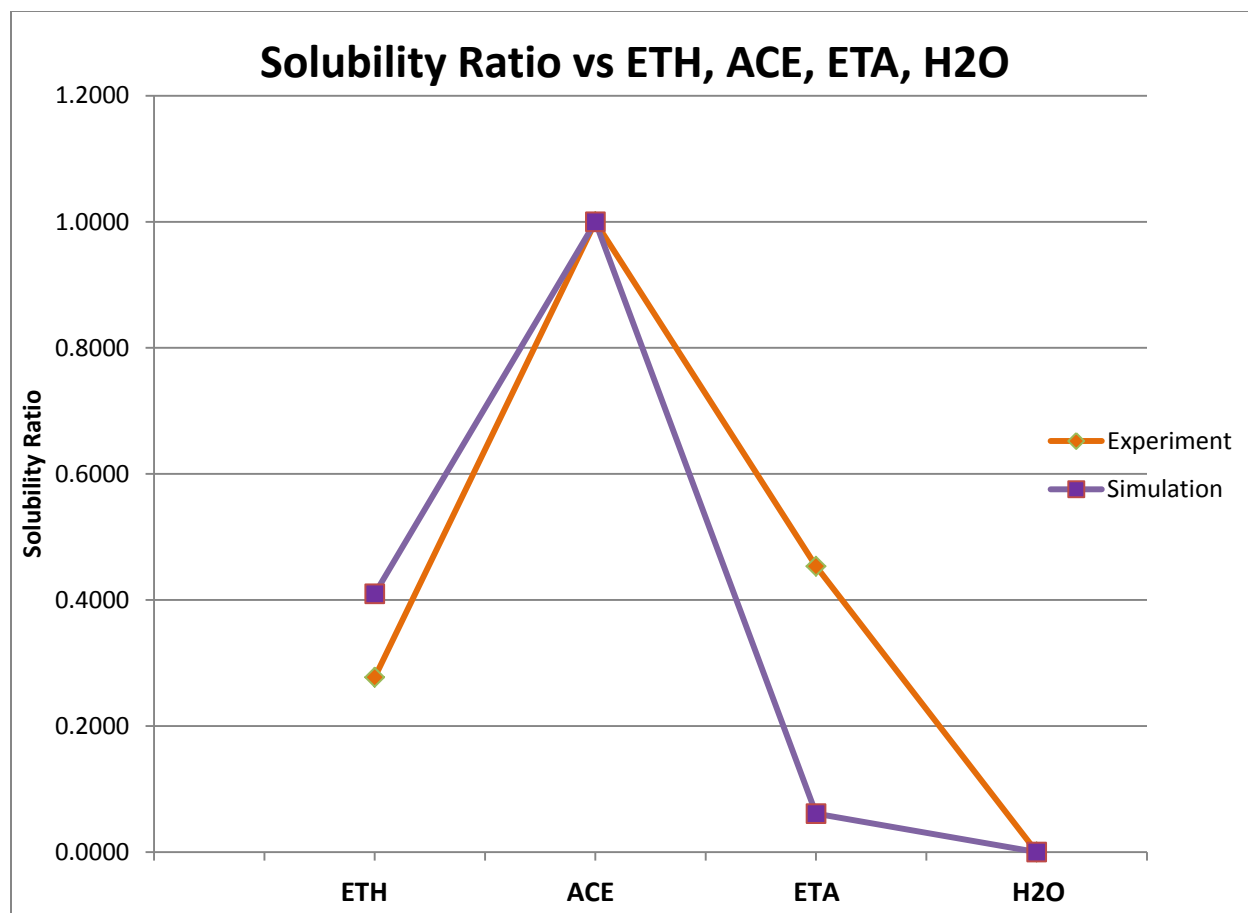


Figure 8. Solubility ratio (solubility of solvent divided by solubility of 1-butanol) versus various solvents such as ethanol (ETH), acetone (ACE), ethyl acetate (ETA), and water (H2O). Experimental values are shown in diamonds (orange) while simulated results are in squares (purple).

Once the same experimental solubility trends are observed through simulation, the radius distribution functions, $g(r)$, can help us further break down the molecule to begin analysis on why the solubility reaches a peak at 1-butanol. Radius distribution functions describe how the density varies a distance r from a reference atom which can help us determine which interactions are strongest in order to identify the polar and non-polar parts to the drug and solvent. The drug was observed to have five polar groups, three of them fairly strong and two fairly weak. As

expected, the drug is for the most part fairly non-polar so the non-polar interactions with the alcohol molecules should increase as the non-polarity increases by adding carbons to the chain.

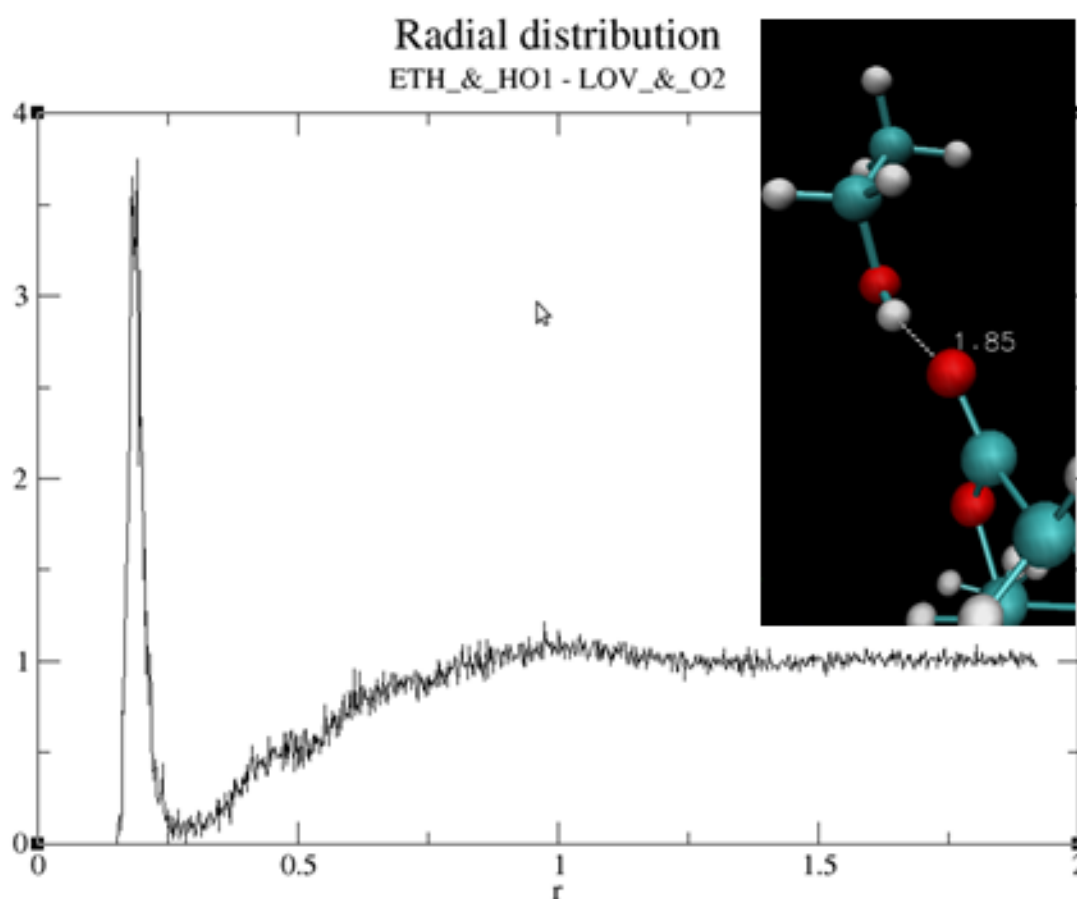


Figure 9. G(r) analysis can help us identify where parts of the drug have strong interactions with surrounding atoms. Shown here is an oxygen on the lovastatin interacting with an hydrogen on an ethanol molecule.

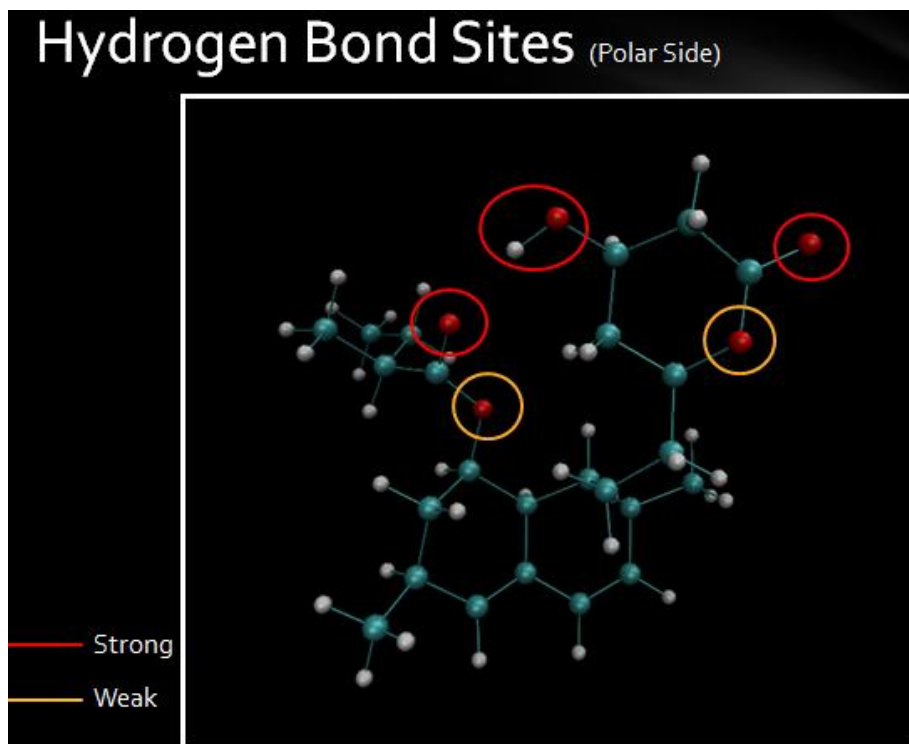


Figure 10. Lovastatin polar groups.

Interaction energies are now calculated between polar and non-polar groups to determine whether they provide strengthening or weakening trends. As can be seen in Figure 11. Group Interaction Energies. Three sets have consistent trends. Polar drug and polar solvent has a trend that reverses at 1-butanol., the various energy groups and their interactions are plotted according to solvent. As was predicted, the interactions between the nonpolar groups of the drug and nonpolar group of the solvent strengthened as the non-polarity of the solvent increased from ethanol to 1-octanol. Further analysis also demonstrates weakening of interactions between the non-polar drug and polar solvent and polar drug non-polar solvent. These three sets of interactions appear to have a consistent trend on the plot as the solvent molecules gets longer in length. The last set of interactions; however, appear to have a non-consistent trend. As the solvent molecule increases to the size of 1-butanol, the polar drug group and the polar solvent

group show an increase in strength. Further increase in solvent size towards 1-octanol shows that the polar interactions between both groups have reached a peak and slowly decrease in strength. This point is validated further as the average number of hydrogen bonds per timeframe appears to be near a steady value of three, however, reach a number of two past 1-hexanol suggesting that somehow certain hydrogen bonds are interacting less. Adding up all the interaction energies gives us a concave plot showing that at 1-butanol the interaction energies reach a peak and slowly reverse in strength possibly giving us a reason on its experimentally determined solubility behavior. One must keep in mind that the interaction energies do not take into account the entropic effect on the Gibbs free energy. The study overall shows that the polar and polar interactions contribution to the overall enthalpy thus changes direction and provides a weakening effect after the organic alcohol chain length surpasses that of 1-butanol.

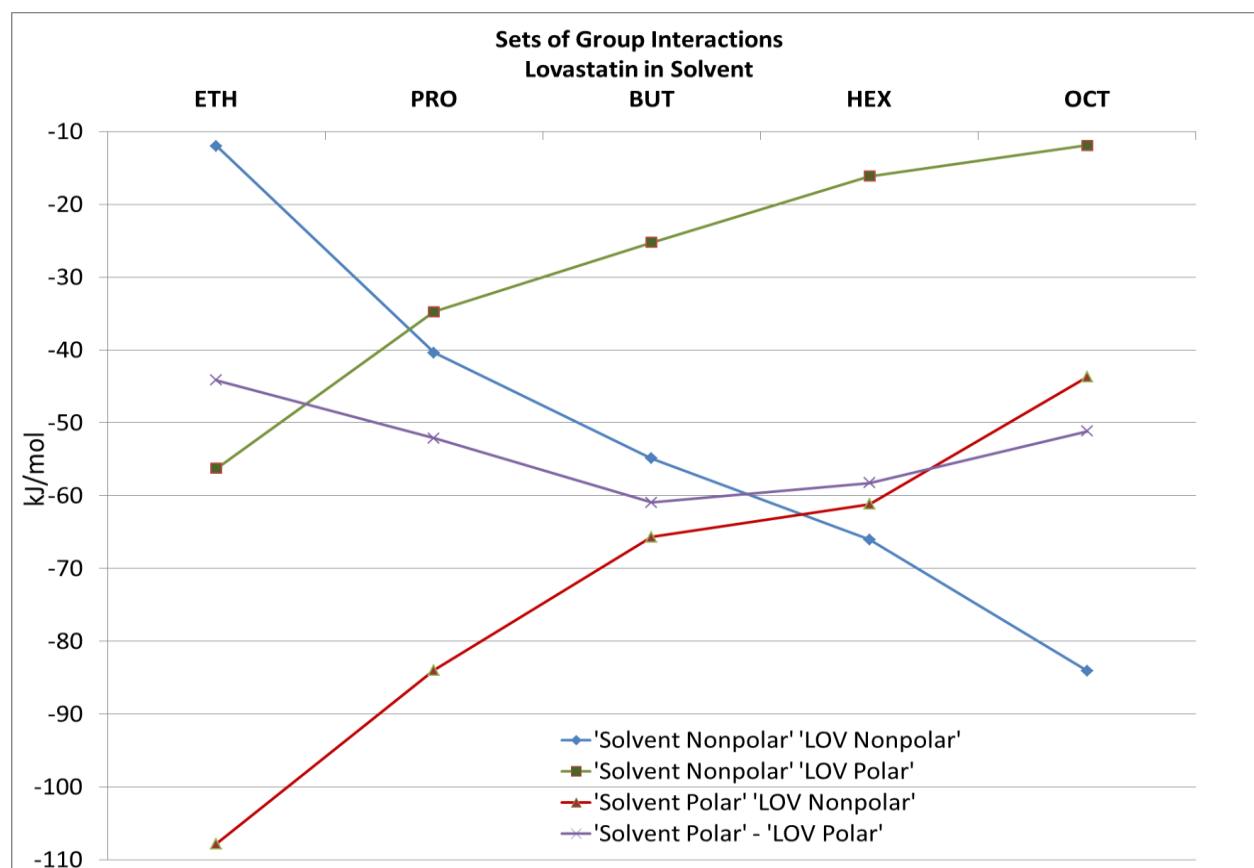


Figure 11. Group Interaction Energies. Three sets have consistent trends. Polar drug and polar solvent has a trend that reverses at 1-butanol.

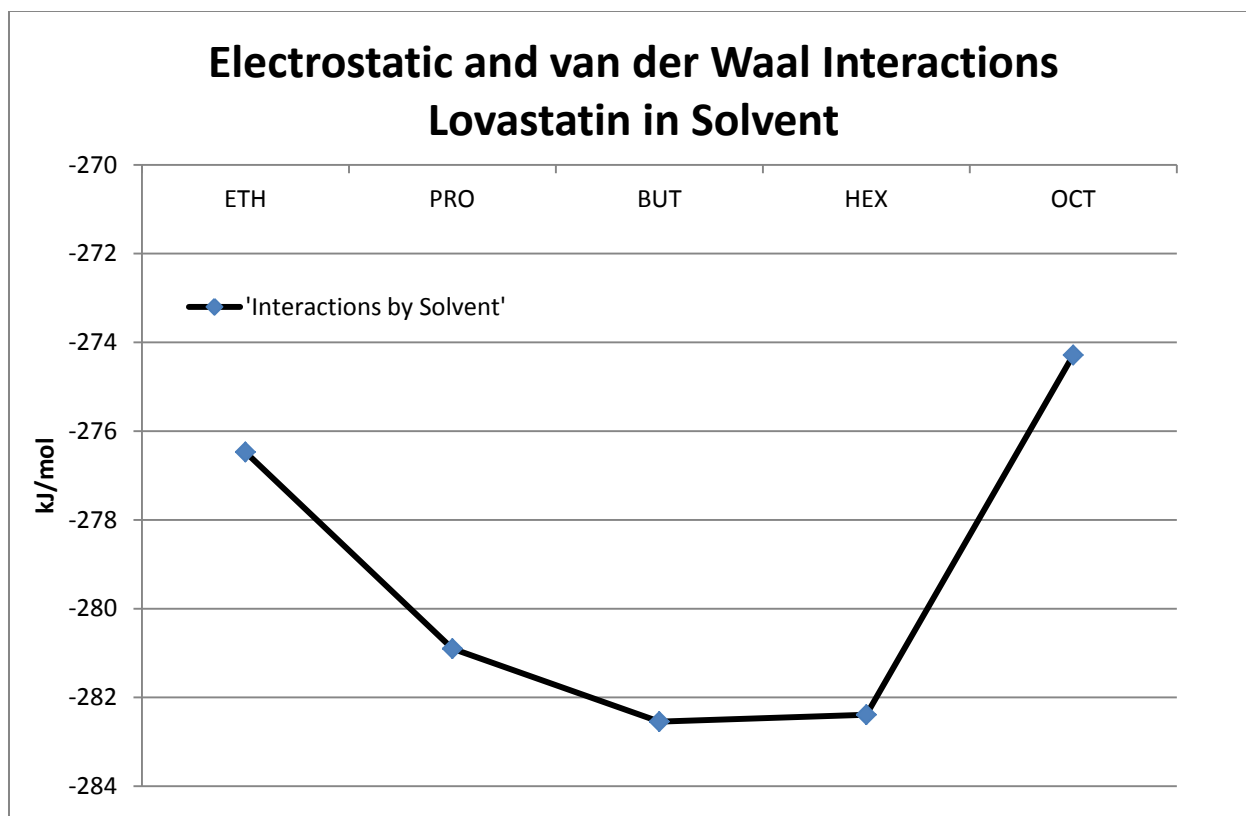


Figure 12. Sum of all interaction energies produces a concave plot having its minimum at 1-butanol.

5.2 Study 2

Results obtained from Study 2 show that the simulation provides a reasonably adequate prediction of the activity coefficient ratios between the two drugs within the same solvent. As can be seen in Figure 13. Free energy difference of mutation. Y-axis corresponds to the right of Equation I. (X-axis to the left of equation), the data obtained from experiment (Nti-Gyabaah [3]) on the y-axis of the plot compares reasonably well to the simulated data on the x-axis. Simulations were each repeated three times to ensure a precise value of the Gibbs free energy of mutation. Values were much less oscillatory compared to Study 1 possibly due to the much

smaller molecule group that goes through decoupling and mutations, ultimately ensuring the averages are much more precise. Thus it is clear that the results can be used to obtain the values of the infinitely diluted activity coefficients given a reference compound.

Table 7. Experimental vs Simulated dG values

dG12-dG34	Experimental	Simulated
Vacuum	-	-
Ethanol	0.77	0.73
Acetone	0.82	0.84
Butanol	0.88	0.85
ETA	0.95	0.99
Hexanol	1.17	1.25

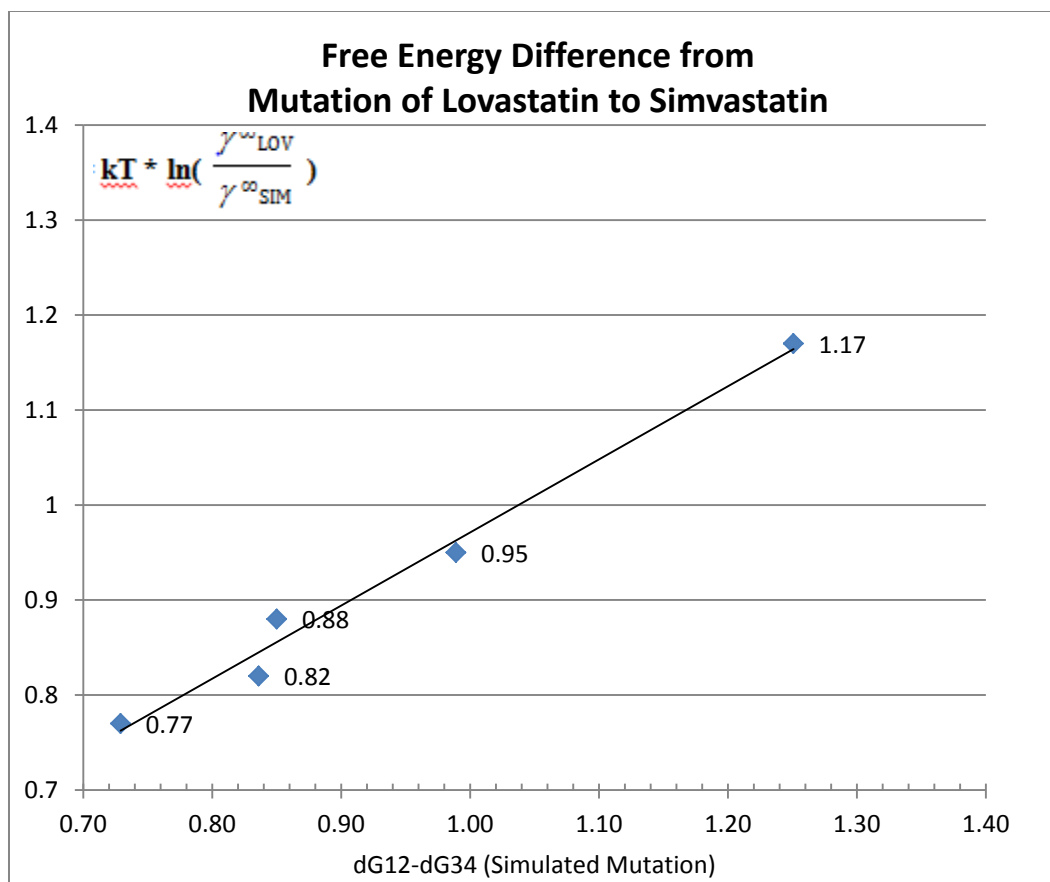


Figure 13. Free energy difference of mutation. Y-axis corresponds to the right of Equation

I. (X-axis to the left of equation)

It can be seen that this technique could prove of use to predict solubility of drug compounds in liquid solvents which could reduce solubility measurement experiments as long as an appropriate force field is chosen.

CHAPTER 6

CONCLUSION AND FUTURE DIRECTION

6.1 Study 1

In this study, it was shown that CHARMM General Force Field is a reasonably good model for lovastatin, simvastatin and common liquid solvents. Our first simulation study was found to give good estimates and capture the behavior of the solubility of lovastatin in liquid solvents. Once the drug solvent mixtures were further analyzed and broken into groups based on polarity, analysis on the energy of interactions provides scientific insights and explanation on why solubility of lovastatin reaches a peak with increasing organic alcohol chain length. Although the nonpolar interactions dominate and strengthen for increasing solvent chain length, the polar and polar interactions of the drug with solvent reach a peak at 1-butanol and gradually weaken. Simvastatin also appears experimentally to experience the same effect as lovastatin so the time consumption for additional simulations on simvastatin was not needed. The main difference is that simvastatin reaches its peak at 1-pentanol.

6.2 Study 2

The second study allowed us to predict the activity coefficient ratios for the different drugs in the same solvent. Results shown give conclusive evidence that this technique can be applied possibly to other drugs that have similar structure differences. In conclusion, molecular dynamics provided excellent insight to understanding solubility behavior of the statins in organic alcohols and aided in the estimation of the activity coefficient ratio that can help in predicting solubility behavior.

6.3 Future Direction

The purpose of this paper was to explore the idea of using molecular simulations using the GROMACS package on predicting certain drug behavior within a solvent. Although various theories on free energy perturbation are available on extracting the necessary data, the focus here was on using the g_{bar} method to approximate the Gibbs free energy between various states within the model. The method saw lovastatin and simvastatin simulated using the CHARMM and CGenFF general force fields within various solvents and used these free energy differences to approximate the activity coefficient ratios. Although the actual ratios differed quite a bit from their experimental values, the overall trend was observed to be similar to experiment. This similarity allowed us to use the molecular dynamics tool to further study the interactions to lead us to a possible explanation on the solubility phenomena seen for lovastatin within various alcohols at the experimental level. The future holds much promise as computational methods continue to improve and better force field models are made. Improvements in these areas can

lead to possibly simulating closer activity coefficient ratios to those seen by experiment.

Modeling the drug also needs to improve as building drug molecules tends to be more difficult as the drugs get more complex. Building these complex drugs often finds the users with a lack of the certain atom types needed to accurately model the drug using one specified force field.

Either way, the promise of utilizing molecular dynamics to explain more thermodynamic properties of drugs such as solubility is there, leading to faster, efficient and cost-effective ways of reducing the number of solubility measurement experiments needed for thermodynamic predictions on drug compounds.

CHAPTER 7
LIST OF NOTATIONS

ΔG	- Change in Gibbs Free Energy
N	-Number of Moles
\overline{G}	-Partial Gibbs Free Energy
X	-Mole Fraction
\overline{G}^{∞}	-Partial Gibbs Free Energy at Infinite Dilution
g	-Pure Gibbs Free Energy
k	-Boltzmanns constant
γ	-Activity Coefficient
f	-Fugacity

CHAPTER 8

MOLECULE BUILDING

CHARMM vs CGenFF Solvent Comparison of Thermodynamic Properties:

The table below is an example of some comparison data between solvent molecules in both CHARMM and CGenFF force fields.

CHARMM							CGENFF					
n-Alcohol	Experimental Density [kg/m3]	Simulated Density [kg/m3]	Density - % Error	Experimental dH,vap [kJ/mol]	Simulated dH,vap [kJ/mol]2	dH,vap [kJ/mol]3 - % Error	Volume [nm3]	CGENFF Simulated Density [kg/m3]	Density - % Error	Simulated dH,vap [kJ/mol]2	dH,vap [kJ/mol]3 - % Error	Volume [nm3]
Methanol	791	784	0.884956	35.2	34.09	0.03	67.80	788	0.41719343	32.89	0.07	67.55
Ethanol	789	784	0.633714	38.60	39.46	-0.02	97.59	808	2.44613435	41.02	-0.06	94.64
Propanol	803	788	1.930262	47.5			126.70	810	0.84682441	47.86	-0.01	123.20
Butanol	810	798	1.495185	49.4			154.30	820	1.2345679	51.41	-0.04	150.00
Pentanol	814	804	1.240786				182.10	826	1.42506143			177.28
Hexanol	814	809	0.651106				209.80	831	2.03783784			204.20
Heptanol	819	813	0.728938				237.34	835	1.91697192			231.00
Octanol	824	818	0.788835				264.50	838	1.66990291			258.10

CHARMM Atom Types:

Table 28. Atom types--CHARMm (Page 1 of 4)

general class	atom type	description
<i>hydrogen types</i>		
	H	hydrogen bonding hydrogen (neutral group)
	HA	aliphatic or aromatic hydrogen
	HC	hydrogen bonding hydrogen (charged group)
	HMU	mu-bonded hydrogen for metals and boron-hydride
	HO	hydrogen on an alcohol oxygen
	HT	TIPS3P water-model hydrogen
<i>carbon types</i>		

	C	carbonyl or guanidinium carbon
	C3	carbonyl carbon in 3-membered aliphatic ring
	C4	carbonyl carbon in 4-membered aliphatic ring
	C5R	aromatic carbon in 5-membered ring
	C5RP	for aryl-aryl bond between C5R rings
	C5RQ	for second aryl-aryl bond between C5RP rings (ortho)
	C6R	aromatic carbon in a 6-membered ring
	C6RP	for aryl-aryl bond between C6R rings
	C6RQ	carbon of C6RP type ortho to C6RP pair
	CF1	carbon with one fluorine
	CF2	carbon with two fluorines
	CF3	carbons with three fluorines
	CM	carbon in carbon monoxide or other triply bonded carbon
	CP3	carbon on nitrogen in proline ring
	CPH1	CG and CD2 carbons in histidine ring
	CPH2	CE1 carbon in histidine ring
	CQ66	third adjacent pair of CR66 types in fused rings
	CT	aliphatic carbon (tetrahedral)

	CT3	carbon in 3-membered aliphatic ring, usually tetrahedral
	CT4	carbon in 4-membered aliphatic ring, usually tetrahedral
	CUA1	carbon in double bond, first pair
	CUA2	carbon in double bond, second conjugated pair
	CUA3	carbon in double bond, third conjugated pair
	CUY1	carbon in triple bond, first pair
	CUY2	carbon in triple bond, second conjugated pair
<i>extended-atom carbon types</i>		
	C5RE	extended aromatic carbon in 5-membered ring
	C6RE	extended aromatic carbon in 6-membered ring
	CH1E	extended-atom carbon with one hydrogen
	CH2E	extended-atom carbon with two hydrogens
	CH3E	extended-atom carbon with three hydrogens
	CR55	aromatic carbon-merged 5-membered rings
	CR56	aromatic carbon-merged 5- or 6-membered rings
	CR66	aromatic carbon-merged 6-membered rings
	CS66	second adjacent pair of CR66 types in fused rings
<i>nitrogen types</i>		

	N	nitrogen: planar-valence of 3, i.e., nitrile, etc.
	N3	nitrogen in a 3-membered ring
	N5R	nitrogen in a 5-membered aromatic ring
	N5RP	for aryl-aryl bond between 5-membered rings
	N6R	nitrogen in a 6-membered aromatic ring
	N6RP	for aryl-aryl bond between 6-membered rings
	NC	charged guanidinium-type nitrogen
	NC2	for neutral guanidinium group - Arg sidechain
	NO2	nitrogen in nitro or related group
	NP	nitrogen in peptide, amide, or related group
	NR1	protonated nitrogen in neutral histidine ring
	NR2	unprotonated nitrogen in neutral histidine ring
	NR3	nitrogens in charged histidine ring
	NR55	N at fused bond between two 5-membered aromatics
	NR56	N at fused bond between 5- and 6-membered aryls
	NR66	N at fused bond between two 6-membered aromatics
	NT	nitrogen (tetrahedral), i.e., amine, etc.
	NX	proline nitrogen or similar

<i>oxygen types</i>		
	O	carbonyl oxygen for amide or related structures
	O2M	oxygen in Si-O-Al or Al-O-Al bond
	O5R	oxygen in 5-membered aromatic ring-radicals, etc.
	O6R	oxygen in 6-membered aromatic ring-radicals, etc.
	OA	carbonyl oxygen for aldehydes or related
	OAC	carbonyl oxygen for acids or related
	OC	charged oxygen
	OE	ether oxygen / acetal oxygen
	OH2	ST2 water-model oxygen
	OK	carbonyl oxygen for ketones or related
	OM	oxygen in carbon monoxide or other triply bonded oxygen
	OS	ester oxygen
	OSH	massless O for zeolites or related cage compounds
	OSI	oxygen in Si-O-Si bond
	OT	hydroxyl oxygen (tetrahedral) or ionizable acid
	OW	TIP3P water-model oxygen
<i>sulfur types</i>		

	S5R	sulfur in a 5-membered aromatic ring
	S6R	sulfur in a 6-membered aromatic ring
	SE	thioether sulfur
	SH1E	extended-atom sulfur with one hydrogen
	SK	thioketone sulfur
	SO1	sulfur bonded to one oxygen
	SO2	sulfur bonded to two oxygens
	SO3	sulfur bonded to three oxygens
	SO4	sulfur bonded to four oxygens
	ST	sulfur, general: usually tetrahedral
<i>phosphorus types</i>		
	P6R	phosphorous in aromatic 6-membered ring
	PO3	phosphorous bonded to three oxygens
	PO4	phosphorous bonded to four oxygens
	PT	phosphorous, general: usually tetrahedral
	PUA1	double-bonded phosphorous
	PUY1	triple-bonded phosphorus

Example of Ethanol Molecule in GROMACS format:

This is an example of a topology file built for GROMACs to use the CHARMM force field.

When building molecule, be sure to verify that the overall net charge sums up to 0. Some premade topologies can be found online at several websites but for more complex molecules one needs to make their own. The topology below comes built-in with the CHARMM files on GROMACs.

```
;
;RESI ETOH          0.00 ! Ethanol, adm jr.
;
;  H21  H11 H12
;      \   \ /
; H22--C2--C1
;      /   \
;  H23      O1--HO1
;
; Ethanol, Jorgensen et al. JACS 118 pp. 11225 (1996)
;
[ moleculetype ]
; name  nrexcl
ETH    3

[ atoms ]
;  nr      type   resnr  residu    atom    cgnr      charge
mass
    1         CT2         1      ETH      C1        1  0.05
    2         OH1         1      ETH      O1        2 -0.66
    3          H          1      ETH      HO1       3  0.43
    4         HA          1      ETH      H11       4  0.09
    5         HA          1      ETH      H12       5  0.09
    6         CT3         1      ETH      C2        6 -0.27
    7         HA          1      ETH      H21       7  0.09
    8         HA          1      ETH      H22       8  0.09
    9         HA          1      ETH      H23       9  0.09

[ bonds ]
;  ai      aj funct          b0          kb
    1      6      1
    1      2      1
    1      4      1
    1      5      1
    2      3      1
    6      7      1
    6      8      1      1
    6      9      1

[ pairs ]
```

```

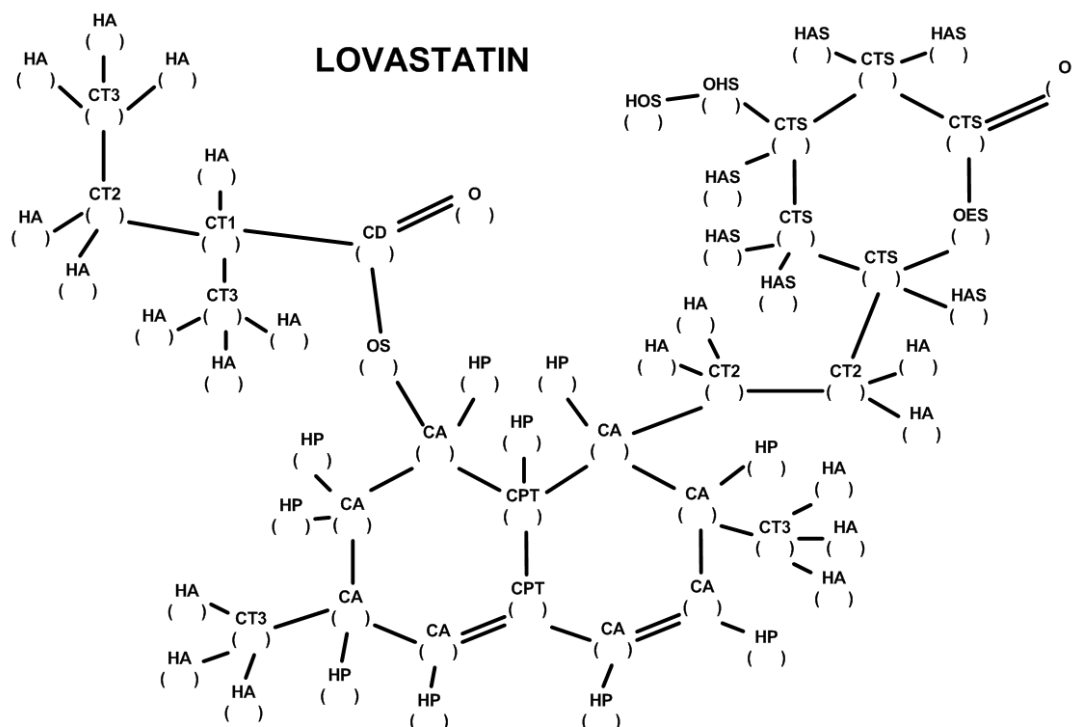
; i   j   func
7   4   1
7   5   1
7   2   1
8   4   1
8   5   1
8   2   1
9   4   1
9   5   1
9   2   1
6   3   1
4   3   1
5   3   1

[ angles ]
; ai   aj   ak funct      theta      ktheta
; atom 1 with CH3 side of molecule H(x)-C-C
7   6   1   5
8   6   1   5
9   6   1   5
; atom 8 with other hydrogen and carbon H(8)-C-H(x)
8   6   7   5
; atom 9 with other hydrogens and carbon H(9)-C-H(x)
9   6   8   5
9   6   7   5
; atom 6 with CH2OH side of molecule C(6)-C-H(x)
6   1   4   5
6   1   5   5
6   1   2   5
; atoms 1,2,3 C(1)-O-H
1   2   3   5
; atom 4 with other hydrogens and carbon H(4)-C-H(x)
4   1   5   5
4   1   2   5
; atom 7 with other hydrogen and carbon H(5)-C-H(2)
5   1   2   5

[ dihedrals ]
; ai   aj   ak   al   funct      phi      kphi      n
7   6   1   4   9
8   6   1   4   9
9   6   1   4   9
7   6   1   5   9
8   6   1   5   9
9   6   1   5   9
7   6   1   2   9
8   6   1   2   9
9   6   1   2   9
6   1   2   3   9
4   1   2   3   9
5   1   2   3   9

```


Lovastatin in CHARMM format:



The lovastatin molecule was made using CHARMM atom types below. Try to break up the molecule into smaller parts to start assigning atom types. This makes it easy to assign charges as well. When combining the molecule together, be sure to carry over charges to have overall compound have a zero net charge. The combination of the smaller parts is the difficult part when creating this larger molecule. Some atom types don't necessarily fit in with the force field which create problems when GROMACS uses the default pair, angle, and dihedral parameters. For example, for the two ring connections below, there are no angle types for CA –CPT – CPT. For more detailed simulations that do not contain ratios in the theory, this will cause more concern for error in calculations.

Two things can be done help optimize the molecule. One is to use ab-initio calculations to try to find the parameters using the methods described in the appropriate force field paper for

CHARMM. The other method is to make an educated estimate by comparing other molecules in literature and see how those parameters work with your molecule.

CGENFF Atom Types:

```

MASS  256 HGA1      1.00800  ! alphatic proton, CH
MASS  257 HGA2      1.00800  ! alphatic proton, CH2
MASS  258 HGA3      1.00800  ! alphatic proton, CH3
MASS  259 HGA4      1.00800  ! alkene proton; RHC=
MASS  260 HGA5      1.00800  ! alkene proton; H2C=CR
MASS  261 HGA6      1.00800  ! aliphatic H on fluorinated C, monofluoro
MASS  262 HGA7      1.00800  ! aliphatic H on fluorinated C, difluoro
MASS  263 HGAAM0    1.00800  ! aliphatic H, NEUTRAL trimethylamine (#)
MASS  264 HGAAM1    1.00800  ! aliphatic H, NEUTRAL dimethylamine (#)
MASS  265 HGAAM2    1.00800  ! aliphatic H, NEUTRAL methylamine (#)
!(#) EXTREME care is required when doing atom typing on compounds that
look like this. Use ONLY
!on NEUTRAL METHYLAMINE groups, NOT Schiff Bases, but DO use on 2 out of 3
guanidine nitrogens
MASS  266 HGP1      1.00800  ! polar H
MASS  267 HGP2      1.00800  ! polar H, +ve charge
MASS  268 HGP3      1.00800  ! polar H, thiol
MASS  269 HGP4      1.00800  ! polar H, neutral conjugated -NH2 group (NA
bases)
MASS  270 HGP5      1.00800  ! polar H on quarternary ammonium salt
(choline)
MASS  271 HGPAM1    1.00800  ! polar H, NEUTRAL dimethylamine (#)
MASS  272 HGPAM2    1.00800  ! polar H, NEUTRAL methylamine (#)
MASS  273 HGPAM3    1.00800  ! polar H, NEUTRAL ammonia (#)
!(#) EXTREME care is required when doing atom typing on compounds that
look like this. Use ONLY
!on NEUTRAL METHYLAMINE groups, NOT Schiff Bases, but DO use on 2 out of 3
guanidine nitrogens
MASS  274 HGR51     1.00800  ! nonpolar H, neutral 5-mem planar ring C, LJ
based on benzene
MASS  275 HGR52     1.00800  ! Aldehyde H, formamide H (RCOH); nonpolar H,
neutral 5-mem planar ring C adjacent to heteroatom or + charge
MASS  276 HGR53     1.00800  ! nonpolar H, +ve charge HIS hel(+1)
MASS  277 HGR61     1.00800  ! aromatic H
MASS  278 HGR62     1.00800  ! nonpolar H, neutral 6-mem planar ring C
adjacent to heteroatom
MASS  279 HGR63     1.00800  ! nonpolar H, NAD+ nicotineamide all ring CH
hydrogens
MASS  280 HGR71     1.00800  ! nonpolar H, neutral 7-mem arom ring, AZUL,
azulene, kevo
!carbons
MASS  281 CG1T1     12.01100  ! alkyn C

```


MASS 282 CG1N1 12.01100 ! C for cyano group
 MASS 283 CG2D1 12.01100 ! alkene; RHC= ; imine C
 MASS 284 CG2D2 12.01100 ! alkene; H2C=
 MASS 285 CG2D1O 12.01100 ! double bond carbon adjacent to heteroatom.
 In conjugated systems, the atom to which it is double bonded must be
 CG2DC1.
 MASS 286 CG2D2O 12.01100 ! double bond carbon adjacent to heteroatom.
 In conjugated systems, the atom to which it is double bonded must be
 CG2DC2.
 MASS 287 CG2DC1 12.01100 ! conjugated alkenes, R2C=CR2
 MASS 288 CG2DC2 12.01100 ! conjugated alkenes, R2C=CR2
 MASS 289 CG2DC3 12.01100 ! conjugated alkenes, H2C=
 MASS 290 CG2N1 12.01100 ! conjugated C in guanidine/guanidinium
 MASS 291 CG2N2 12.01100 ! conjugated C in amidinium cation
 MASS 292 CG2O1 12.01100 ! carbonyl C: amides
 MASS 293 CG2O2 12.01100 ! carbonyl C: esters, [neutral] carboxylic
 acids
 MASS 294 CG2O3 12.01100 ! carbonyl C: [negative] carboxylates
 MASS 295 CG2O4 12.01100 ! carbonyl C: aldehydes
 MASS 296 CG2O5 12.01100 ! carbonyl C: ketones
 MASS 297 CG2O6 12.01100 ! carbonyl C: urea, carbonate
 MASS 298 CG2O7 12.01100 ! CO2 carbon
 MASS 299 CG2R51 12.01100 ! 5-mem ring, his CG, CD2(0), trp
 MASS 300 CG2R52 12.01100 ! 5-mem ring, double bound to N, PYRZ,
 pyrazole
 MASS 301 CG2R53 12.01100 ! 5-mem ring, double bound to N and adjacent
 to another heteroatom, purine C8, his CE1 (0,+1), 2PDO, kevo
 MASS 302 CG2R61 12.01100 ! 6-mem aromatic C
 MASS 303 CG2R62 12.01100 ! 6-mem aromatic C for protonated pyridine
 (NIC) and rings containing carbonyls (see CG2R63) (NA)
 MASS 304 CG2R63 12.01100 ! 6-mem aromatic amide carbon (NA) (and other
 6-mem aromatic carbonyls?)
 MASS 305 CG2R64 12.01100 ! 6-mem aromatic amidine and guanidine carbon
 (between 2 or 3 Ns and double-bound to one of them), NA, PYRM
 MASS 306 CG2R66 12.01100 ! 6-mem aromatic carbon bound to F
 MASS 307 CG2R67 12.01100 ! 6-mem aromatic carbon of biphenyl
 MASS 308 CG2RC0 12.01100 ! 6/5-mem ring bridging C, guanine C4,C5, trp
 MASS 309 CG2R71 12.01100 ! 7-mem ring arom C, AZUL, azulene, kevo
 MASS 310 CG2RC7 12.01100 ! sp2 ring connection with single bond(!),
 AZUL, azulene, kevo
 MASS 311 CG301 12.01100 ! aliphatic C, no hydrogens, neopentane
 MASS 312 CG302 12.01100 ! aliphatic C, no hydrogens, trifluoromethyl
 MASS 313 CG311 12.01100 ! aliphatic C with 1 H, CH
 MASS 314 CG312 12.01100 ! aliphatic C with 1 H, difluoromethyl
 MASS 315 CG314 12.01100 ! aliphatic C with 1 H, adjacent to positive
 N (PROT NTER) (+)
 MASS 316 CG321 12.01100 ! aliphatic C for CH2
 MASS 317 CG322 12.01100 ! aliphatic C for CH2, monofluoromethyl
 MASS 318 CG323 12.01100 ! aliphatic C for CH2, thiolate carbon
 MASS 319 CG324 12.01100 ! aliphatic C for CH2, adjacent to positive N
 (piperidine) (+)
 MASS 320 CG331 12.01100 ! aliphatic C for methyl group (-CH3)
 MASS 321 CG334 12.01100 ! aliphatic C for methyl group (-CH3),
 adjacent to positive N (PROT NTER) (+)

MASS 322 CG3AM0 12.01100 ! aliphatic C for CH₃, NEUTRAL trimethylamine
 methyl carbon (#)
 MASS 323 CG3AM1 12.01100 ! aliphatic C for CH₃, NEUTRAL dimethylamine
 methyl carbon (#)
 MASS 324 CG3AM2 12.01100 ! aliphatic C for CH₃, NEUTRAL methylamine
 methyl carbon (#)
 !(#) EXTREME care is required when doing atom typing on compounds that
 look like this. Use ONLY
 !on NEUTRAL METHYLAMINE groups, NOT ETHYL, NOT Schiff Bases, but DO use on
 2 out of 3 guanidine nitrogens
 MASS 325 CG3C31 12.01100 ! cyclopropyl carbon
 !MASS 326 CG3C41 12.01100 ! cyclobutyl carbon RESERVED!
 MASS 327 CG3C50 12.01100 ! 5-mem ring aliphatic quaternary C
 (cholesterol, bile acids)
 MASS 328 CG3C51 12.01100 ! 5-mem ring aliphatic CH (proline CA,
 furanoses)
 MASS 329 CG3C52 12.01100 ! 5-mem ring aliphatic CH₂ (proline CB/CG/CD,
 THF, deoxyribose)
 MASS 330 CG3C53 12.01100 ! 5-mem ring aliphatic CH adjacent to
 positive N (proline.H⁺ CA) (+)
 MASS 331 CG3C54 12.01100 ! 5-mem ring aliphatic CH₂ adjacent to
 positive N (proline.H⁺ CD) (+)
 MASS 332 CG3RC1 12.01100 ! bridgehead in bicyclic systems containing
 at least one 5-membered or smaller ring
 !(+) Includes protonated Schiff base (NG3D5, NG2R52 in 2HPP) but NOT
 amidinium (NG2R52 in IMIM), guanidinium
 !nitrogens
 MASS 333 NG1T1 14.00700 ! N for cyano group
 MASS 334 NG2D1 14.00700 ! N for neutral imine/Schiff's base (C=N-R,
 acyclic amidine, guanidine)
 MASS 335 NG2S0 14.00700 ! N,N-disubstituted amide, proline N
 (CO=NRR')
 MASS 336 NG2S1 14.00700 ! peptide nitrogen (CO=NHR)
 MASS 337 NG2S2 14.00700 ! terminal amide nitrogen (CO=NH₂)
 MASS 338 NG2S3 14.00700 ! external amine ring nitrogen
 (planar/aniline), phosphoramidate
 MASS 339 NG2O1 14.00700 ! NITB, nitrobenzene
 MASS 340 NG2P1 14.00700 ! N for protonated imine/Schiff's base
 (C=N(+)H-R, acyclic amidinium, guanidinium)
 MASS 341 NG2R50 14.00700 ! double bound neutral 5-mem planar ring,
 purine N7
 MASS 342 NG2R51 14.00700 ! single bound neutral 5-mem planar (all atom
 types sp²) ring, his, trp pyrrole (fused)
 MASS 343 NG2R52 14.00700 ! protonated schiff base, amidinium,
 guanidinium in 5-membered ring, HIS, 2HPP, kevo
 MASS 344 NG2R53 14.00700 ! amide in 5-memebered NON-SP² ring (slightly
 pyramidized), 2PDO, kevo
 MASS 345 NG2R60 14.00700 ! double bound neutral 6-mem planar ring,
 pyr1, pyzn
 MASS 346 NG2R61 14.00700 ! single bound neutral 6-mem planar ring
 imino nitrogen; glycosyl linkage
 MASS 347 NG2R62 14.00700 ! double bound 6-mem planar ring with
 heteroatoms in o or m, pyrd, pyrm

MASS 348 NG2RC0 14.00700 ! 6/5-mem ring bridging N, indolizine, INDZ,
 kevo
 MASS 349 NG301 14.00700 ! neutral trimethylamine nitrogen
 MASS 350 NG311 14.00700 ! neutral dimethylamine nitrogen
 MASS 351 NG321 14.00700 ! neutral methylamine nitrogen
 MASS 352 NG331 14.00700 ! neutral ammonia nitrogen
 MASS 353 NG3C51 14.00700 ! secondary sp3 amine in 5-membered ring
 MASS 354 NG3N1 14.00700 ! N in hydrazine, HDZN
 MASS 355 NG3P0 14.00700 ! quarternary N+, choline
 MASS 356 NG3P1 14.00700 ! tertiary NH+ (PIP)
 MASS 357 NG3P2 14.00700 ! secondary NH2+ (proline)
 MASS 358 NG3P3 14.00700 ! primary NH3+, phosphatidylethanolamine
 !oxygens
 MASS 359 OG2D1 15.99940 ! carbonyl O: amides, esters, [neutral]
 carboxylic acids, aldehydes, uera
 MASS 360 OG2D2 15.99940 ! carbonyl O: negative groups: carboxylates,
 carbonate
 MASS 361 OG2D3 15.99940 ! carbonyl O: ketones
 MASS 362 OG2D4 15.99940 ! 6-mem aromatic carbonyl oxygen (nucleic
 bases)
 MASS 363 OG2D5 15.99940 ! CO2 oxygen
 MASS 364 OG2N1 15.99940 ! NITB, nitrobenzene
 MASS 365 OG2P1 15.99940 ! =O in phosphate or sulfate
 MASS 366 OG2R50 15.99940 ! FURA, furan
 MASS 367 OG3R60 15.99940 ! O in 6-mem cyclic enol ether (PY01, PY02)
 or ester
 MASS 368 OG301 15.99940 ! ether -O- !SHOULD WE HAVE A SEPARATE ENOL
 ETHER??? IF YES, SHOULD WE MERGE IT WITH OG3R60???
 MASS 369 OG302 15.99940 ! ester -O-
 MASS 370 OG303 15.99940 ! phosphate/sulfate ester oxygen
 MASS 371 OG304 15.99940 ! linkage oxygen in
 pyrophosphate/pyrosulphate
 MASS 372 OG311 15.99940 ! hydroxyl oxygen
 MASS 373 OG312 15.99940 ! ionized alcohol oxygen
 MASS 374 OG3C51 15.99940 ! 5-mem furanose ring oxygen (ether)
 MASS 375 OG3C61 15.99940 ! DIOX, dioxane, ether in 6-membered ring
 !SHOULD WE MERGE THIS WITH OG3R60???
 !sulphurs
 MASS 376 SG2D1 32.06000 ! thiocarbonyl S
 MASS 377 SG2R50 32.06000 ! THIP, thiophene
 MASS 378 SG311 32.06000 ! sulphur, SH, -S-
 MASS 379 SG301 32.06000 ! sulfur C-S-S-C type
 MASS 380 SG302 32.06000 ! thiolate sulfur (-1)
 MASS 381 SG301 32.06000 ! sulfate -1 sulfur
 MASS 382 SG302 32.06000 ! neutral sulfone/sulfonamide sulfur
 MASS 383 SG303 32.06000 ! neutral sulfoxide sulfur
 !halogens
 MASS 384 CLGA1 35.45300 ! CLET, DCLE, chloroethane, 1,1-
 dichloroethane
 MASS 385 CLGA3 35.45300 ! TCLE, 1,1,1-trichloroethane
 MASS 386 CLGR1 35.45300 ! CHLB, chlorobenzene
 MASS 387 BRGA1 79.90400 ! BRET, bromoethane
 MASS 388 BRGA2 79.90400 ! DBRE, 1,1-dibromoethane
 MASS 389 BRGA3 79.90400 ! TBRE, 1,1,1-dibromoethane

```

MASS    390 BRGR1    79.90400  ! BROB, bromobenzene
MASS    391 IGR1    126.90447  ! IODB, iodobenzene
MASS    392 FGA1    18.99800  ! aliphatic fluorine, monofluoro
MASS    393 FGA2    18.99800  ! aliphatic fluorine, difluoro
MASS    394 FGA3    18.99800  ! aliphatic fluorine, trifluoro
MASS    395 FGP1    18.99800  ! anionic F, for ALF4 AlF4-
MASS    396 FGR1    18.99800  ! aromatic flourine
!miscellaneous
MASS    397 PG0     30.97380  ! neutral phosphate
MASS    398 PG1     30.97380  ! phosphate -1
MASS    399 PG2     30.97380  ! phosphate -2
MASS    400 ALG1    26.98154  ! Aluminum, for ALF4, AlF4-

MASS    402 CG25C1  12.01100  ! same as CG2DC1 but in 5-membered ring with
exocyclic double bond
MASS    403 CG25C2  12.01100  ! same as CG2DC2 but in 5-membered ring with
exocyclic double bond
MASS    404 CG25I0  12.01100  ! same as CG2D10 but in 5-membered ring with
exocyclic double bond
MASS    405 CG2520  12.01100  ! same as CG2D20 but in 5-membered ring with
exocyclic double bond

!MASS    410 HGTIP3   1.00800  ! polar H, TIPS3P WATER HYDROGEN
!MASS    411 OGTIP3  15.99940  ! TIPS3P WATER OXYGEN
!MASS    412 DUM      0.00000  ! dummy atom
!MASS    413 HE       4.00260  ! helium
!MASS    414 NE      20.17970  ! neon

```

Lovastatin in CGenFF format:

While creating the molecule, be sure to use the default molecules created in the topology file.

This is very helpful when putting together the larger drug compound. When in doubt, the force

CHAPTER 9
LIST OF CITATIONS

1. J. Nti-Gyabaah, R.C., V. Chan, Y.C. Chiew, *Solubility of lovastatin in a family of six alcohols: Ethanol, 1-propanol, 1-butanol, 1-pentanol, 1-hexanol, and 1-octanol*. International Journal of Pharmaceutics, 2008(359): p. 111-117.
2. Shobha N. Bhattachar, L.A.D.a.J.A.W., *Solubility: It's not just for physical chemists*. Drug Discovery Today, 2006. **11**: p. 21/22.
3. Nti-Gyabaah, J., *Solubility and Activity Coefficient of Pharmaceutical Compounds in Liquid Organic Solvents*, in *Graduate Program in Chemical and Biochemical Engineering* 2009, Rutgers, The State University of New Jersey.
4. Stote, D., and Kuznetsov. *Molecular Dynamics Simulations CHARMM*. October 26, 1999; Available from: http://www.ch.embnet.org/MD_tutorial/.
5. Hess, B., Kutzner, C., van der Spoel, D. and Lindahl, E., *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*. J. Chem. Theory Computation., 2008. **4**: p. 435-447.
6. K. Vanommeslaeghe, E.H., C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell Jr., *CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields*, in *Department of Pharmaceutical Sciences, School of Pharmacy* 2009, University of Maryland, Baltimore, Maryland 21201: Wiley Periodicals, Inc.
7. P. Bjelkmar, P.L., M. Cuendet, B. Hess, and E. Lindahl, *Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models*. Journal of Chemical Theory and Computation, 2009. **2010**(6): p. 459 - 466.

8. Lemkul, J. *Free Energy Calculations: Methane in Water*. GROMACS Tutorials 2008-2011; Available from: http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/free_energy/01_theory.html.
9. Leach, A.R., *Molecular Modelling: Principles and Applications*. Second Edition ed2001: Pearson Education Limited 1996.
10. Bennet, C.H., *Efficient Estimation of Free Energy Differences from Monte Carlo Data*. *Journal of Computational Physics*, 1976. **22**: p. 245-268.

CHAPTER 10
CIRRICULUM VITAE

CIRRICULUM VITAE

FABIAN F CASTEBLANCO

EDUCATION:

- **New Jersey Institute of Technology, 2005-2010, Chemical Engineering, B.S**
- **Rutgers University, 2010-2013, Chemical Engineering, M.S.**

INDUSTRIAL EXPERIENCE:

- **Dow Chemical, Coatings & Latex Business, Kankakee, IL**
 - **September 2012 – Present, Run Plant Engineer**
- **General Electric, GE Power & Water, Schenectady, NY**
 - **May 2010 – August 2010, Process Systems Engineering Coop**
- **Infineum USA, Bayway Chemical Plant, Linden, NJ**
 - **September 2009 – December 2009, Process Engineering Coop**
- **General Electric, GE Energy, Greenville, SC**
 - **May 2009 – August 2009, Process Controls Engineering Coop**
- **General Electric, GE Energy, Somersworth, NH**
 - **May 2008 – August 2008, Manufacturing Engineering Coop**
- **General Electric, GE Energy, Houston TX**
 - **January 2007 – May 2007, Process Engineering Coop**