SOLUBILITY AND ACTIVITY COEFFICIENT OF PHARMACEUTICAL COMPOUNDS IN LIQUID ORGANIC SOLVENTS

by

JOSEPH NTI-GYABAAH

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Chemical and Biochemical Engineering

Written under the direction of

Professor Yee C. Chiew Ph.D

and approved by:

_____

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2009

ABSTRACT OF THE DISSERTATION

Molecular Thermodynamics Modeling of Solubility and Activity Coefficient of Pharmaceutical Compounds in Liquid Organic Solvents

by

JOSEPH NTI-GYABAAH

Dissertation Director:

Professor Yee C. Chiew Ph.D

The production of pharmaceuticals and oligo-sized biochemicals involves liquid solvent selection as a function of solubility, for purification, chemical reaction and formulation. Selecting the optimum solvent for a particular application is of critical importance to developing efficient process; choosing solvents for pharmaceutical processes has been based on experience and empirical description of experimental results. Therefore, rapid and reliable prediction of drug solubility is needed for design and optimization of cost-effective manufacturing processes.

A semi-automated, two-millimeter-scale method, was designed and validated for rapid measurement of equilibrium solubility of crystalline solids in liquid solvents. Solubilities and stability of model compounds such as lovastatin, simvastatin, and artemisinin, were measured in several solvents and exhibited a maximum deviation of 5% compared to

literature data. Extrapolation of the solubility in the form of activity coefficient was determined to agree with the experimental data.

Calculation of the ideal solubility of crystalline solid in liquid solvent requires knowledge of the difference in molar heat capacity of the solid and super-cooled liquid solute at the solution temperature. This is a hypothetical parameter and therefore, can not be measured directly, and hence, three assumptions are commonly used in the literature for its estimation. Evaluation of the assumptions revealed some thermodynamic inconsistencies. A new strategy was explored to estimate the difference in molar heat capacity, allowing the experimental solubility data to be fitted to the Two-Liquid-Non-Random (NRTL) activity coefficient equation to obtain the model binary energetic interaction parameters. The binary interaction parameters were successfully used to estimate solubility of the model compounds in mixed solvents.

In free energy perturbation (FEP) methods, mutation of lovastatin to simvastatin in five different solvents were carried out to obtain the free energy differences, and hence the ratio of the activity coefficients at infinite dilution. The FEP calculation reproduced the experimental solubility results quite well, and provides a basis for development of molecular modeling for estimation of pharmaceutical compounds in liquid solvents.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 1

INTRODUCTION

Research Motivation and Objectives

The importance of solubility of drugs in liquid solvents can be recognized at all steps of the drug discovery and development processes. Experimental determination of drug solubility occurs multiple times along the drug discovery, and manufacturing process.

For the drug discovery phase compounds are synthesized and tested for solubility for bioavailability; whereas for drug manufacturing process development various solvents are screened for solubility, for extraction, crystallization, and chromatography. Compounds that lack solubility have a higher potential of failure during discovery as insufficient solubility may compromise other property assays, mask additional undesirable properties, influence both pharmacokinetic and pharmacodynamic properties of the compound, and finally may impact the potential to further develop the compound. Thus, in ideal case, solubility issues must be known before the compound can be evaluated for compound development. Therefore in drug discovery, during the period from lead identification to clinical candidate selection, hundreds to thousands of different compounds are tested for solubility in order to select and optimize for the most suitable compound.

As compounds leave the lead identification phase and enter the compound optimization space, solubility screens are also performed in organic solvents. These are required to

support initial salt and polymorph screening activities with a goal to aim to pure, stable, crystalline material in the most desirable physical form of the compound as early as possible for development. Additionally, organic solvent solubility data are very useful for formulators, process chemists, and analytical chemists for formulation, scale-up, and analytical method development, respectively.

In much later phases, the solubility profile in organic solvents is necessary to develop efficient cleaning validation protocols for the cleaning of pilot plant and equipment, in compliance with regulatory requirements. For drug manufacturing processes the choice of solvent, based on solubility, for purification (e.g., extraction / crystallization) have huge impact on product recovery, as well as purity. For drug manufacturing processes, experiments are performed to screen solvents and identify suitable operating ranges. The required solubility data multiplies rapidly as a function of chosen solvent, solvent mixture/ composition, and temperature. In addition, as drug discovery techniques continue to improve, the number of potential drug candidates and associated intermediates continues to increase. The task of generating solubility data becomes overwhelming even with powerful automation tools. Although automated screening is appealing, it can become resource and compound intensive as experimental conditions are, in most instances, chosen arbitrarily or simply based on experience. An alternate approach is to apply thermodynamic modeling to predict or estimate solubility which can, in turn, result in a significant reduction of compound consumption, particularly during the early stages of discovery when material is very scarce. For example, prediction of solubility can be very important in guiding synthetic routes, where a range of reaction schemes and building blocks can be narrowed down considerably on the basis of

theoretical predictions of solubility prior to incurring the costs of synthesis and testing, and modifications to improve the solubility could be incorporated into the next design cycle.

Computational methods, for prediction of solubility for drug-like compounds presents a challenge, since these compounds are usually complex: (i.e., they contain a number of different features of rings, and chiral centers). The presently used solution models, depend on precise determination of various kinds of phase equilibria, which provide the meams to obtain the required model parameters through rigorous thermodynamic data reduction. Particularly, the interaction parameters for the popularly used activity coefficient models, such as, UNiversal QUAsiChemical (UNIQUAC) by Abrams and Prausnitz (1975), the Wilson's equation (Wilson, 1968), and the Non-random Two-liquid activity coefficient (NRTL) model (Renon, 1968)) were all derived from vapor-liquid-equilibria (VLE) data. Efforts to apply these models for solid-liquid-equilibria (SLE) for pharmaceutical application have not been successful. For instance, it has been shown that a different set of parameters is needed for SLE (Magnussen et al., 1981).

Additionally, when the temperature range departs substantially from that of the original data which was employed to generate the parameters, predictions given by these activity coefficient models do not always compare favorably with the experimental data, as is the case with many multi-component SLE of several multi-component organic mixtures (Gmeling et al., 1978, Carta and Dernini, 1983). Thus, insights into why these models fail in the low temperature range, where drug compounds are processed (because they are labile) is needed, the precise and rapid determination of the SLE conditions must be

practically taken into consideration. With these in mind, this thesis has three objectives, and these are:

1) Develop an automated material-conserving solubility measuring technique for rapid measurement of equilibrium solubility of pharmaceutical compounds, using lovastatin, simvastatin, and artemisinin as model compounds.

2) Evaluate various techniques used for calculation of ideal solubility of drug compounds, examining all the assumptions for thermodynamic consistency.

3) Model the solubility data to the Non-random Two-Liquid (NRTL) activity coefficient equation to the solute-solvent solubility data to obtain the model binary interaction parameters. Use the binary interaction parameters to predict solubility of the model compounds in ternary mixtures (i.e., solute-binary solvent mixtures); and as needed, propose correction to the interaction parameters.

4) Finally, explore free energy perturbation theory to develop a mathematical tool for *in sillico* predicting the ratio of activity coefficients, using a reference compound. Presentation in the thesis follows the following order:

Review of various solubility measurement techniques, and the design and implementation of the solubility measurement technique in single solvents and solvent mixtures is described in Chapter 2.

In Chapter 3, various methods used for calculation of ideal solubility is reviewed. Various assumptions used for the calculations are examined for thermodynamic consistency.

Modeling of the experimentally measured solubility to the NRTL activity coefficient equation is presented in Chapter 4. The use of the model parameters to predict solubility in mixed solvents is also discussed.

In Chapter 5, the use of free energy perturbation calculations (FEP) in a mutation of lovastatin to simvastatin to obtain the ration of the activity coefficients at infinite dilution is presented.

In Chapter 6, the overall summary of the research and recommendations for future directions are overviewed.

CHAPTER 2

MATERIAL-CONSERVING SOLUBILITY MEASURING METHOD

2.1 Introduction

Accurate solubility data are needed for process and product design including production and purification of pharmaceutical compounds, formulation, controlled drug delivery systems, bio-separations, precipitation/crystallization processes, chemical reaction systems, pollution prevention and remediation, and food processing. There is a vast amount of literature reporting the results of drug solubility measurements in organic solvents (Jouyban et al., 1998; 2002, Kolar et al., 2002, Frank et al., 1999; Tiziani et al., 2007, Mirmehrabia et al., 2004, Huang et al., 2005, Ruckenstein and Shulgin, 2002, 2003, 2003, 2004, Rubino and Obeng, 1991, Adjei et al., 1980). But, many more combinations of solvent and solute remain to be investigated. At the interface between drug discovery and preclinical candidate selection, solubility measurement and correlation have to satisfy the needs of both drug discovery and product development. Especially for low solubility compounds, extended solubility data, in both aqueous and organic pharmaceutically relevant solvents are needed to select suitable pre-formulation strategy for subsequent pharmacokinetics and pharmacodynamics (PK/PD) and toxicology studies (Tiziani et al., 2007). Crystallization as a unit operation in the pharmaceutical industry serves the dual purpose of isolation and purification of the active pharmaceutical ingredient (API). And, because the compounds of interest are often labile, solution crystallization is the primary method of crystallization in comparison to other crystallization techniques such as heat melt. With increasing complexity of API molecules, enhanced analytical techniques to

detect impurities and ever-decreasing timelines for drug development, the challenges of selecting appropriate solvents with optimum selectivity for rejection in drug purification have grown significantly over the years.

In this chapter, various solubility measuring techniques are reviewed, and the method employed in for the current study is discussed. The method was employed to measure the solubility of lovastatin, simvastatin, artemisnin in different pure solvents and solvent mixtures. Background of these model compounds are provided in the next two sections.

2.2 Background of lovastatin and simvastatin

Lovastatin, and simvastatin (structures shown in Figure 2.1 and Figure 2.2, respectively) belong to a class of the most powerful lipid lowering drug compounds, called the statins (Istvan and Deisenhofer, 2001). Their mode of action is through inhibition of the (3S)-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase. Using competition, statins specifically inhibit HMG-CoA reductase, the enzyme that catalyzes the conversion of HMG-CoA to mevolanate, which is an early rate-limiting step in cholesterol biosynthesis in the body. The resulting decrease in intracellular cholesterol results in compensatory increase in cholesterol uptake by means of low density lipoprotein (LDL) receptors and concomitant decrease in plasma cholesterol. Statins are the treatment of choice for management of hypercholestrolaemia because of their proven efficacy and safety profile (Palinski, 2000, Javernik et al., 2003, Greenberg et al., 2004, Kim et al., 1999, Vincenzi et al., 2003, Strode et al., 1999, Shacter, 2004, Elder, 1988, Sutherland et al., 2001, Manzoni et al., 1998). Seven statins are currently approved for clinical use (Shacter, 2004): Lovastatin, Simvastatin, Fluvastatin, Atorvastatin, Rosuvastatin,

Prevastatin, and Pitavastatin. Lovastatin, (a.k.a. butanioc acid 2-methyl-1,2,3,7,8,8a-hexahydro-3,7-dimethyl-8-[2-(tetrahydro-4-hydroxy-6-oxo-2H-pyran-2-yl) ethyl]-1-naphalenyl ester) is a natural product, a secondary metabolite, derived from fermentation of *Aspergillus terreus* (Elder, 1988, Sutherland et al., 2001, Manzoni et al., 1998). It is also a key raw material for synthesis of simvastatin. Following isolation of the lovastatin crude product, the compound is purified by various crystallization sequences prior to milling to finished product, and subsequent formulation as drug (Elder, 1988).

Very little has been reported in the literature about the solubility of statins in organic solvents. Recently, Sun *et al.* (2005) employed the synthetic method, using a laser monitoring observation technique to measure the solubility of lovastatin in acetone, ethyl acetate, butyl acetate, ethanol, and methanol at different temperatures.

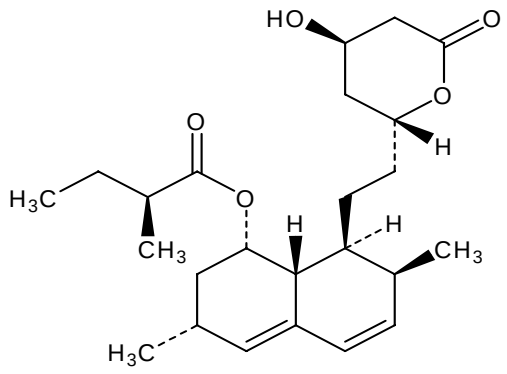


Figure 2.1: Structure of lovastatin.

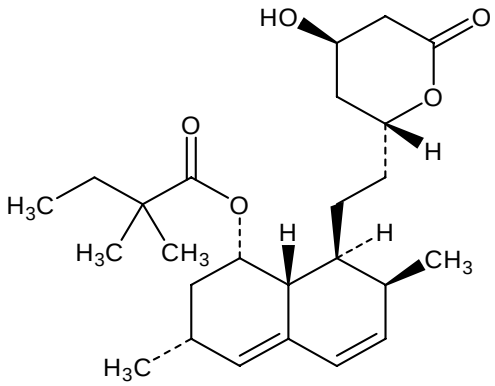


Figure 2.2: Structure of simvastatin.

2.3 Background of artemisinin

Artemisinin, known in Chinese as Qinghaosu, is an important bioactive component in *Artemisia annua* leaves and flowers, which has been used as a traditional Chinese medicine in the treatment of fever for a long time (Kumar, 2005). It has been recently found to have activity against malaria, and as matter, is the most efficacious anti-malaria drug at the moment. Artemisinin itself has physical properties such as poor bioavailability that limit its effectiveness. Therefore, semi-synthetic derivatives of artemisinin, including artemether and artesunate, have been developed (Noedl et al, 2008). However, their activity is not long lasting, with significant decreases in effectiveness after one to two hours. To counter this drawback, artemisinin is typically given with lumefantrine (also known as benflumetol) to treat uncomplicated *falciparum* malaria. Lumefantrine has a half-life of about 3 to 6 days and prevents the disease from returning. The treatments are called "ACT" (artemisinin-based combination therapy); other examples are artemether-lumefantrine, artesunate-mefloquine, artesunate-amodiaquine,

and artesunate-sulfadoxine/pyrimethamine. Recent trials have shown that these therapies are more than 90% effective, with a recovery from symptoms after three days, especially for the chloroquine-resistant *Plasmodium falciparum* (Mutabingwa et al., 2005).

Malaria is a vector-borne infectious disease caused by protozoan parasites. It is widespread in tropical and subtropical regions, including parts of the Americas, Asia, and Africa. According to the World Health Organization (WHO) each year, there are approximately 350–500 million cases of malaria, killing between one and three million people, the majority of whom are young children in Sub-Saharan Africa. Ninety percent of malaria-related deaths occur in Sub-Saharan Africa. Malaria is commonly associated with poverty, but is also a cause of poverty and a major hindrance to economic development (Hamm, 2009).

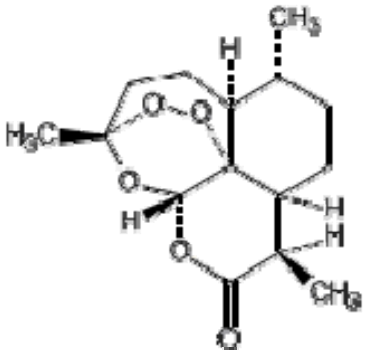


Figure 2.3: Structure of artemisinin (MW = 282.33).

Most of the *Artemisia annua* grown worldwide is currently processed via solvent extraction, using warm hexane and/or petroleum ether (Kumar, 2005). The only other considered alternative has been the use of supercritical carbon dioxide (Brunner et al., 2005). In as much as both hexane and petroleum ether are cheap solvents, they present a considerable safety and environmental hazards. The lack of accurate solubility data to aid selection of optimum solvent for the extraction process, makes it difficult for new and existing artemisia producers to assess the efficiency, financial viability, safety and environmental impacts.

Recent advances in large-scale natural product isolation process development using conventional extraction and purification protocols have enabled optimization to high recovery yields, in the range of 80 – 90 % for a variety natural products of interest in the pharmaceutical industry. These development approaches are based on rationale selection and optimization of solvents, co-solvents, chromatographic resins, operating conditions within the context of physico-chemical properties of the natural products, interfering substances and impurities as well as application of thermodynamic principles of phase equilibra such as solubilities, and partition coefficients. Rapid solvent screening methods coupled with predictive models will enable optimization of conventional extraction and purification protocols into scaleable processes for product delivery within realistic development timeframes.

2.4 Various solubility measurement techniques

2.4.1 Batch method

In this commonly used method, the solvent is agitated or stirred with solid solute in a vessel, afterwhich an aliquot of the saturated solution is removed, then filtered or centrifuged to remove the excess solute before the solution is analyzed for equilibrium concentration (Jouyban et al., 2001). Here, it is important to ascertain whether the solid phase of the solute changes during the equilibration to produce a different polymorph or solvate (Yalkowsky, 1981). Additionally, precaution must be taken when removing the excess solids from the solution. For example, if filter is employed, the filter media must be compatible with the solvent and its pore size must be small enough to retain the smallest solid particles. Furthermore, care must be taken to avoid losses of the dissolved solute by adsorption to the filter media, and/or onto vessels, pipettes, and syringes. If centrifugation is employed any disturbance and carryover of the un-dissolved solid solute must be monitored and avoided.

2.4.2 Flow column method

This analytical method was developed by Wasik (1981). Here, a suitable column is packed with the solid solute or with suitable supporting material, such as glass beads, onto which the solute has previously been adsorbed by evaporation of a suitable solution. The solvent is pumped through the column. The increased surface area of contact facilitates rapid attainment of equilibrium so that the effluent from the column is saturated; and it analyzed just as in the previous method. The flow column has useful

advantages. Manipulation of the system prior to analyses is minimized, so there are less problems such as adsorption or evaporation that may result from separation of the excess solids from saturated solution are reduced. This method is rapid, precise and particularly valuable for sparingly soluble systems, such as hydrophobic solutes in water. However, there are some disadvantages. For example, there is a possibility of the solute undergoing phase transition during the evaporation process to coat the solute onto the solid support. Furthermore, a strong interaction between the adsorbed solute and the support material may hinder the thermodynamic activity of the adsorbed solute below that of the crystalline solids, to the extent that, the measured solubility may be reduced below that of the batch agitation method.

2.4.3 Synthetic method

For this method, either a weighed amount of the solute (or a finite amount of the solvent) is placed in a suitable vessel (Wang et al., 2005). While stirring the solution at a fixed temperature, known amounts of solvent (or the solute) are gradually added until the solution becomes turbid. Appropriate care must be taken to ensure that the system has reached or is very close to reaching equilibrium at the solution temperature.

2.5 Method employed in this thesis

The experimental procedure used here is somewhat similar to approach used by Shiu and co-workers (Shiu et al., 1988), except that, here, an analytical high performance liquid Chromatographic (HPLC) method, coupled with the batch solubility measurement

technique, was designed to measure the solubility of three model compounds (i.e. lovastatin, simvastatin, and artermisinin) in several solvents over a wide range of temperature. Key advantages of this method include:

1. Small amount of material is required (i.e., over fifty-fold reduction in sample size compared to the conventional synthetic method);
2. Automated in-situ sampling of solution, and hence significant reduction with issues associated with sample manipulation; and
3. The presence or absence of the degradation products could be determined.

2.6 Experimental methods

2.6.1 Material

Crystalline solids of the three model compounds: lovastatin ($C_{24}H_{36}O_5$; MW 404.54); simvastatin ($C_{25}H_{38}O_5$; MW 418.58); and artemisnin ($C_{15}H_{22}O_5$; MW 282.33) were obtained from Alexis Biochemicals (San Diego, CA) with mass purity determined by HPLC, of 99.8 % wt were used as model compounds for this dissertation study. HPLC analytical grade reagent solvents (each > 99.5 % purity): methanol, ethanol, 1-propanol, 1-butanol, 1-hexanol, 1-pentanol, 1-hexanol, and 1-octanol, methyl acetate, ethyl acetate, propyl acetate, isopropyl acetate, butyl acetate, isobutyl acetate, sec butyl acetate, tert butyl acetate, acetone, and 2-butanoe, methylene chloride, methyl tert-butyl ether (MTBE) were obtained from Fisher Scientific. The solvents were dried with molecular

sieves before use, and the purities confirmed by gas chromatography to be > 99.5 %. Water content was determined by Karl Fisher titration to be < 0.005 %wt.

2.6.2 Equipment

Wrist Action, Burrel, Model 75 mechanical shaker; Mettler AE 160 digital analytical balances, sensitivity 0.01 mg; 2910 Modulated DSC, TA Instruments differential scanning calorimeter; DSC822e, Mettler-Toledo differential scanning calorimeter. Analytical scale solubility experiments were performed using an Agilent HP-1100 HPLC system composed of a quaternary pump, column and auto sampler thermostat and variable wavelength detector. A set of five standard stock solutions of the pertinent solute were prepared by appropriate dilution of a stock solution. These were then used to generate a calibration curve (with regression coefficient better than 0.999). The calibration curve was used to determine the equilibrium concentrations of the pertinent solute upon sampling and analysis.

2.6.3 Solubility measurements

Solubility of the three model compounds, namely lovastatin, simvastatin, and artemisinin, were measured in the following sixteen organic solvents: methanol, ethanol, propanol, butanol, pentanol, hexanol, octanol, methyl acetate, ethyl acetate, isopropyl acetate, butyl acetate, 2-butanone, methyl tert-butyl ether, acetone, and toluene. The solvents were chosen to examine solubility of the model compounds over a wide range of solvent polarity.

In each pure solvent, or solvent mixture, about 150 mg of the pertinent model compound (an excess of substance) was added to several 2 mL HPLC vials, containing 1.5 mL of the pertinent solvent. The mixtures were stirred in a mechanical shaker, maintained at 40 ± 0.1 ºC, for 24 hours. Visual inspection was carefully made to ensure there were excess crystalline solids, indicating saturation had been reached. The vials were then loaded into the theromostat-temperature-controlled autosampler of the HPLC (setup shown in Figure 4) and the temperature was lowered to the desired temperature (at a cooling rate of 0.25 ºC/hr). Upon reaching the desired temperature, the mixture was allowed to equilibrate for 24 to 48 hours (although our experimental results indicated that 12 hours was sufficient for complete equilibration and settling of un-dissolved solute). Thereafter, the solution was sampled then analyzed via the reversed phase method to determine the equilibrium concentration, as well as to ensure the compound was stable in the pertinent solvent.

To avoid any potential differential temperature driven precipitation upon sampling, the HPLC sampling needle was stored in the thermost-temperature-controlled HPLC auto-sampler compartment with the samples. This ensured that its temperature was same as that of the sample. Additionally, the needle was positioned to allow careful sampling of 2 μL solution from the top middle portion of the vial; this ensured that the settled solids were not disturbed. Furthermore, each vial was sampled and analyzed in triplicates to ensure that the system had reached equilibrium at the point of sampling. The method was validated by comparing our results with literature values for equilibrium solubility of lovastatin in ethanol (Sun et al., 2005). The solvent mixtures were prepared by weight to within $5x10^{-5}$ g.

Figure 2.4: HPLC setup for solubility measurements.

2.6.4 Reversed-phase analytical HPLC methods

All samples were analyzed by reversed phase analytical HPLC with UV detection. The column used for the reversed phase analysis (Symmetry®, 4.6mm I.D. x 50 mm, packed with silica-C-8, 3.5 µm particle diameter) was obtained from Waters Corporation, and maintained at 60°C. All elutions were carried out at 4.5 mL/min; mobile phase conditions were: started isocratically with 70% 0.01M H_3PO_4 (in water) and 30% acetonitrile for 1 minute, followed by a linear gradient to 70 % acetonitrile in 3 minutes, afterwhich the column was flushed with 100% acetonitrile for 1 minutes, then re-equilibrated with the 70% 0.01M H_3PO_4 (in water) and 30% acetonitrile for 2 minutes prior to the next injection (i.e. total run time was 7 minutes). For each run, the mobile was directed through the sampling needle sample loop into to the column, to ensure complete loading of the sample to the column. Concentration of the solute was

calculated based on a calibration curve, and the value was used to calculate the equilibrium solubility mole fraction, x_1, in the binary mixtures as:

$$x_1 = \frac{M_1}{M_1 + M_2} \tag{2.1}$$

where M_1 and M_2 represent the moles of the solute and solvent, respectively.

And, in the solute-binary solvent mixtures:

$$x_1 = \frac{M_1}{M_1 + M_2 + M_3} \tag{2.2}$$

where M_1, M_2, and M_3 represent the moles of the solute, solvent # 1, and solvent #2 , respectively.

2.7 Results and discussion

2.7.1 Solubility of lovastatin in alcohols.

The solubility measurement technique in this study was validated by comparing the temperature-dependent equilibrium mole fraction of lovastatin in ethanol, with literature results (Table 2.1). At elevated temperature (> 300 K), a second peak was observed for the lovastatin-methanol system. The spectra of the second peak varied slightly from that obtained with the pure lovastatin, suggesting modification to the compound has occurred

at these temperatures. An isolate of the second peak was analyzed via NMR and mass spectroscopy, and found to be a derivative of the original compound, indicating that lovastatin is not stable in methanol, at elevated temperatures, for extended period. Solubility of lovastatin in methanol exists in the literature (Sun et al., 2005); it is likely that the degradation was not observed because they employed the synthetic method, which did not require analysis on HPLC column. Lovastatin was relatively stable in all the other alcohols employed in this study. For all cases, solubility of lovastatin increases with increasing temperature (Figure 2.5). It is interesting to note that, solubility of lovastatin increased with increasing alkyl chain length of the alcohol (from ethanol to 1-butanol) then decreased as the carbon chain length increased.

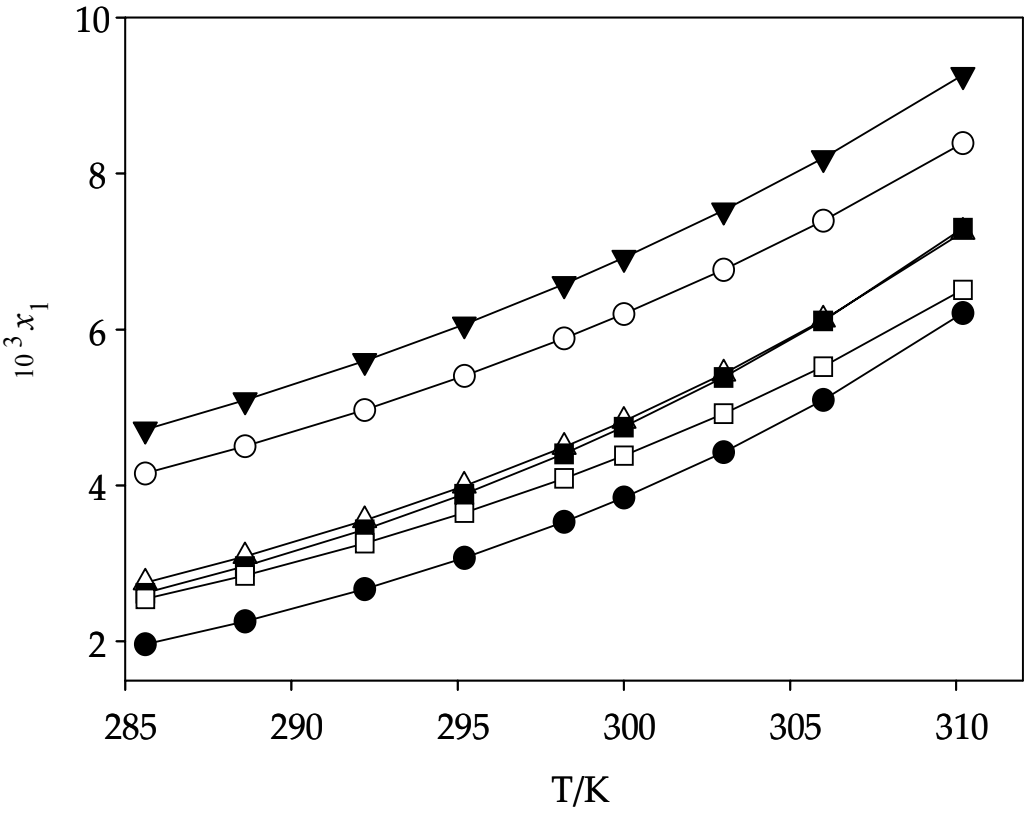


Figure 2.5: Equilibrium solubility of lovastatin in alcohols at different temperatures.

Table 2.1

Mole fraction solubility data of lovastatin, $10^3\ x_1$, in various alcohols.

T/K	Ethanol♣	T/K	**Ethanol** ♢	T/K	Propanol	T/K	Butanol	T/K	Pentanol	T/K	Hexanol	T/K	Octanol
286.2	2.0574	278.3	1.4020	286.2	4.2104	285.7	4.6046	289.2	3.1213	293.9	3.6267	285.7	2.5518
289.2	2.3045	283.3	1.7970	289.2	4.5126	288.7	4.9877	291.1	3.3974	295.9	3.9958	288.7	2.7800
291.1	2.5077	288.3	2.2560	291.1	4.8848	290.7	5.2554	295.5	4.0441	298.9	4.5493	290.6	3.1255
296.7	3.2270	307.8	5.5360	301.7	6.4980	296.2	6.1126	299.2	4.6819	300.7	4.9080	296.2	3.9135
301.7	4.0839	313.1	7.0940	304.6	7.1873	301.2	7.0227	304.6	5.7108	303.9	5.6535	301.2	4.6741
304.6	4.6648	318.7	9.3260	305.7	7.3603	304.1	7.6196	305.6	6.1291	308.9	7.0345	304.1	4.9997
307.2	5.4305	307.8	5.5360	307.2	7.5945	305.2	7.8606	307.2	6.4004	312.0	7.6794	305.2	5.1493
309.2	5.8507	313.1	7.0940	309.2	8.0885	306.7	8.2028	309.2	6.9139			306.7	5.6685
310.6	6.2596	318.7	9.3260	310.5	8.4415	310.1	9.0410	310.6	7.3787			310.1	6.5927

♣ Thesis data

♢ Literature data (Sun et al, 2005).

2.7.2 Solubility of lovastatin in acetates and ketones

The experimentally measured mole fraction temperature-dependent equilibrium solubility of lovastatin in methyl acetate, ethyl acetate, propyl acetate, isopropyl acetate, n-butyl acetate, sec butyl acetate, isobutyl acetate, tert butyl acetate, acetone, and 2-butanone, between 279 K and 313 K, are presented in Table 2.2, and graphically displayed in Figure 2.6 and Figure 2.7. For each solvent studied, the equilibrium solubility mole fraction of lovastatin increased with temperature (Nti-Gyabaah and Chiew, 2008). Generally, the order of decreasing solubility of lovastatin, in the solvents studied was, acetone > 2-butanone > ethyl acetate > n-butyl acetate > n-propyl acetate > sec butyl acetate > iso butyl acetate > methyl acetate > tert butyl acetate. We observed a trend of decreasing solubility for the straight chain, and branched acetates.

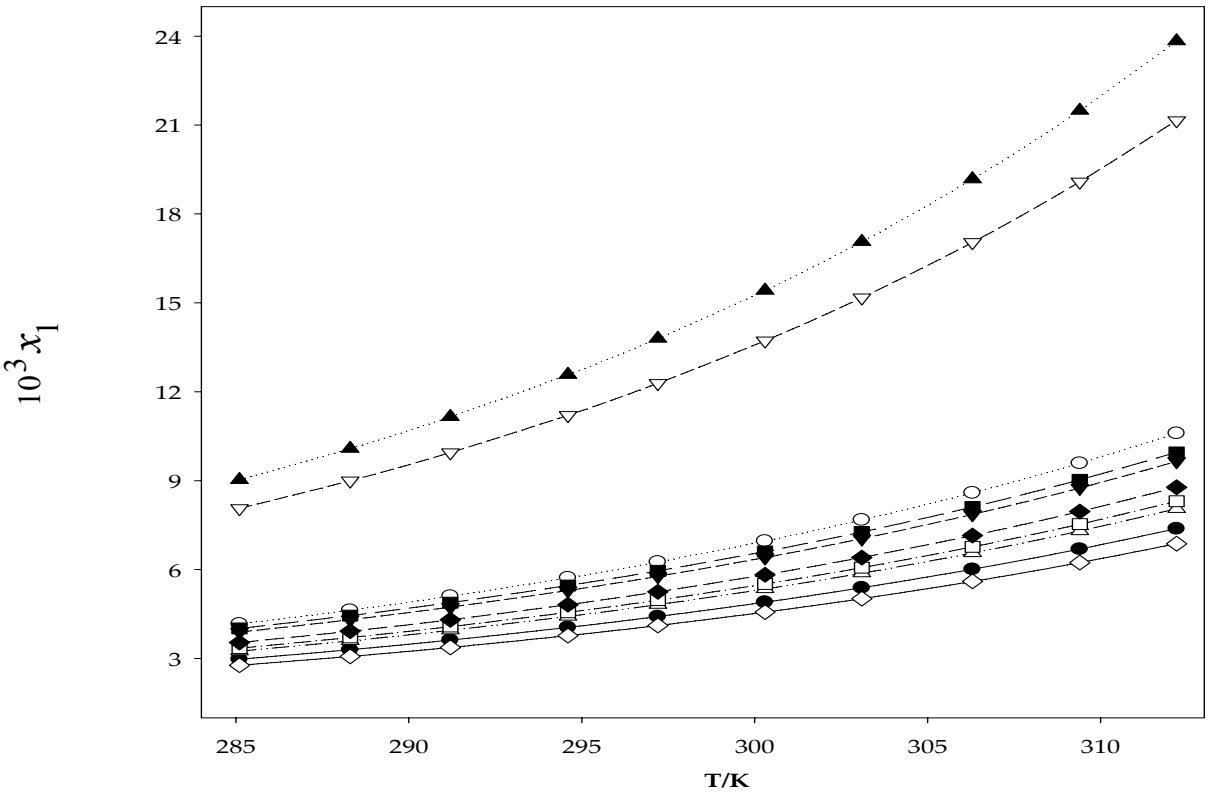


Figure 2.6: Equilibrium solubility of lovastatin in different solvents at different temperatures.

Table 2.2

Mole fraction solubility data of lovastatin, $10^3 x_1$, in different organic solvent mixtures and at different temperatures.

T/K	Methyl acetate	Ethyl acetate	Propyl acetate	Isopropyl acetate	Butyl Acetate	*Iso*butyl acetate	*sec* Butyl acetate	*tert bu*tyl acetate	Acetone	2-Butanone
285.1	3.09	4.17	3.95	3.35	4.07	3.45	3.65	2.75	8.31	7.52
288.3	3.40	4.70	4.38	3.72	4.52	3.83	4.01	3.14	9.36	8.78
291.2	3.76	5.17	4.81	4.12	4.94	4.22	4.46	3.46	10.43	9.18
294.6	4.17	5.86	5.44	4.59	5.57	4.70	5.00	3.86	11.84	11.05
297.2	4.47	6.56	5.87	4.97	6.11	5.04	5.45	4.20	13.04	11.51
300.3	5.15	7.06	6.58	5.53	6.79	5.70	6.06	4.79	14.64	13.24
303.1	5.43	7.78	7.24	6.06	7.48	6.27	6.66	5.10	16.26	14.70
306.3	6.19	8.70	8.03	6.77	8.36	7.02	7.45	5.69	18.32	17.51
309.4	6.74	9.67	9.01	7.56	9.31	7.77	8.30	6.38	20.56	19.72
312.2	7.46	10.78	9.97	8.32	10.32	8.64	9.15	6.93	22.83	21.86

2.7.3 Solubility of simvastatin in different organic solvents

Experimentally measured temperature-dependent equilibrium mole fractions of simvastatin in the acetates (i.e., methyl acetate, ethyl acetate, n-propyl acetate, iso-propyl acetate, n-butyl acetate, iso-butyl acetate, sec-butyl acetate, tert-butyl acetate); and in the alcohols (i.e., ethanol, 1-propanol, 1-butanol, 2-butanol, 1-pentanol, 1-hexanol, and 1-octanol) are shown in Table 2.3 and Table 2.4, respectively. For each solvent studied herein, the equilibrium solubility mole fraction of simvastatin increased with temperature (Nti-Gyabaah et al., 2009). Throughout the entire range of temperature studied (279 K to 315 K), the order of increasing mole fraction solubility of simvastatin in the acetates was: n-butyl acetate > ethyl acetate > n-propyl acetate > sec-butyl acetate > *iso*-butyl acetate > methyl acetate > iso-propyl acetate > tert-butyl acetate (Figure 2.7). Whereas the order of increasing mole fraction solubility of simvastatin in the alcohols was: 1-hexanol > 1-octanol > 2-butanol > 1-pentanol > 1-butanol > 1-propanol > ethanol, at temperatures less than 296 K, beyond which, there are some few crossovers: 1-hexanol > 1-pentanol > 1-octanol > 1-butanol > 1-propanol > 2-butanol > ethanol (Figure 2.8).

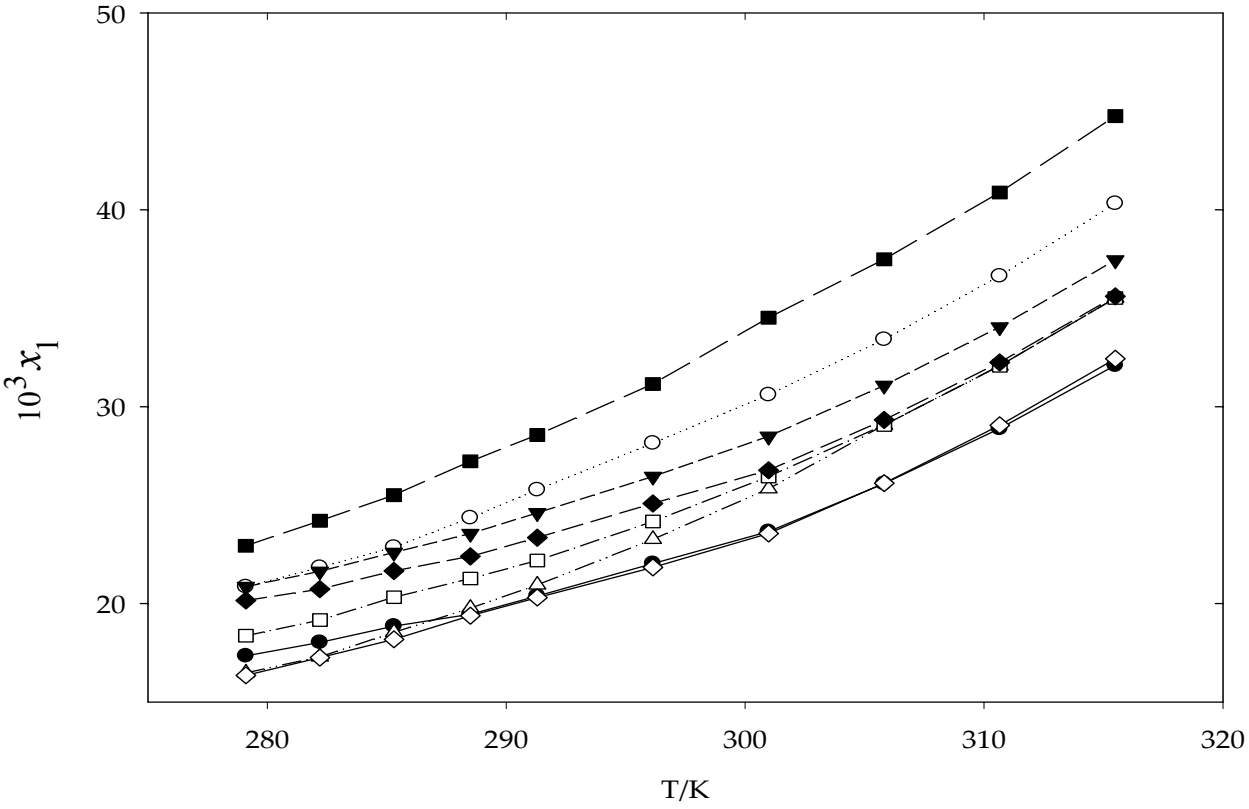


Figure 2.7: Equilibrium solubility of simvastatin in acetates at temperatures.

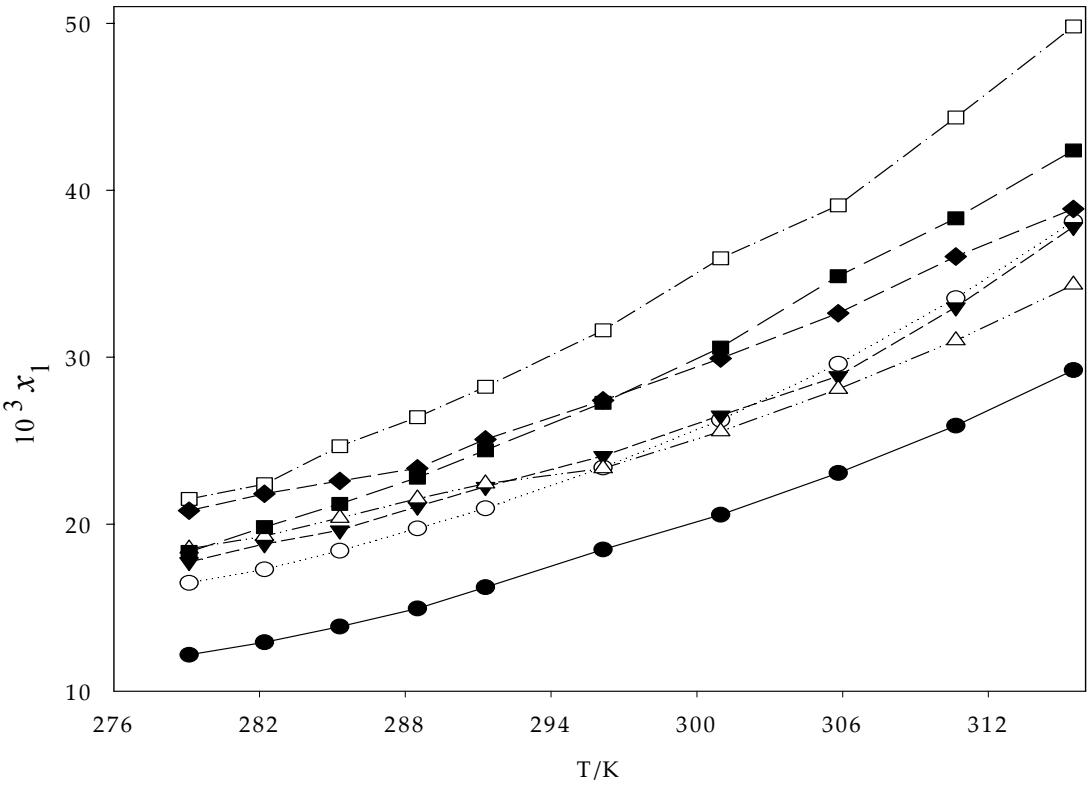


Figure 2.8: Equilibrium solubility of simvastatin in alcohols at different temperatures.

Table 2.3

Mole fraction solubility data of simvastatin, $10^3 x_1$, in different organic solvents, and at different temperatures.

T/K	Methyl Acetate	Ethyl acetate	*n*-Propyl acetate	*iso*propyl acetate	*n*-Butyl acetate	*iso*butyl acetate	*sec*butyl acetate	*tert*butyl acetate
279.1	17.34	20.86	20.86	16.48	22.94	18.37	20.16	16.36
282.2	18.03	21.84	21.64	17.29	24.21	19.16	20.73	17.25
285.3	18.86	22.86	22.59	18.53	25.51	20.32	21.66	18.19
288.5	19.46	24.36	23.55	19.77	27.23	21.27	22.40	19.38
291.3	20.38	25.79	24.60	20.95	28.57	22.18	23.35	20.31
296.1	22.04	28.15	26.46	23.27	31.16	24.17	25.09	21.84
301.0	23.66	30.61	28.49	25.83	34.52	26.45	26.77	23.56
305.8	26.10	33.42	31.08	29.07	37.49	29.07	29.33	26.12
310.7	28.89	36.64	34.05	32.08	40.88	32.08	32.25	29.06
315.5	32.10	40.33	37.45	35.52	44.76	35.52	35.60	32.44

Table 2.4

Mole fraction solubility data of simvastatin, $10^3 x_1$, in alcohols at different temperatures.

T/K	Ethanol	1-Propanol	1-Butanol	*2*-Butanol	1-Pentanol	1-Hexanol	1-Octanol
279.1	12.18	16.48	17.74	18.52	18.34	21.51	20.81
282.2	12.93	17.29	18.81	19.29	19.82	22.40	21.82
285.3	13.87	18.41	19.64	20.37	21.21	24.66	22.59
288.5	14.95	19.74	21.06	21.53	22.79	26.41	23.35
291.3	16.23	20.95	22.25	22.43	24.43	28.22	25.08
296.1	18.48	23.39	24.09	23.34	27.28	31.61	27.41
301.0	20.56	26.25	26.51	25.55	30.58	35.92	29.92
305.8	23.06	29.60	28.87	28.09	34.85	39.09	32.64
310.7	25.90	33.53	32.99	31.00	38.33	44.36	36.04
315.5	29.23	38.13	37.81	34.33	42.39	49.81	38.89

2.7.4 Solubility of artemisinin in different solvents

Experimentally measured temperature-dependent equilibrium mole fractions of artemisinin in methanol, acetone, ethyl acetate, acetonitrile, hexane, heptane, methylene chloride, 2-butanone (MEK), methyl tert-butyl ether (MTBE), ethanol, butanol, isopropyl acetate, and toluene are presented in Table 2.5 and Table 2.6, and graphically displayed in Figure 2.9. For each solvent studied herein, the equilibrium solubility mole fraction of artemisinin increased with temperature. Generally, solubility decreased according to the following order, methylene chloride > MEK > toluene > propyl acetate > ethyl acetate > acetone > ethanol > MTBE > butanol > acetonitrile > methanol > hexane > heptane.

As previously reviewed, some of the current manufacturers use hexane to extract the compound (Kumar, 2005). This is surprising considering that solubility of artemisinin is very low in this solvent. This may explain the need to use warm hexane for the extraction. Additionally, the waxes in the plant would have very high solubility in hydrocarbons (e.g., hexane and heptane), and perhaps, is partially the reason why waxes co-extract with the product during the bulk initial extraction, and hence the penalty on yield during the re-crystallization. Based on the present data, one proposed approach to minimize the problem, and optimize the yield, might be to selectively extract the waxes with hexane or heptane (at low temperature), extract the product with small volumes of one of the solvents in which artemisinin has been identified to have high solubility (e.g., methylene chloride, MEK, or toluene), then use hexane or heptane as anti-solvent to crystallize out the product. We envision that the proposed approach will lead to a more

efficient manufacturing process for artemisinin and help address some of the cost and supply issues.

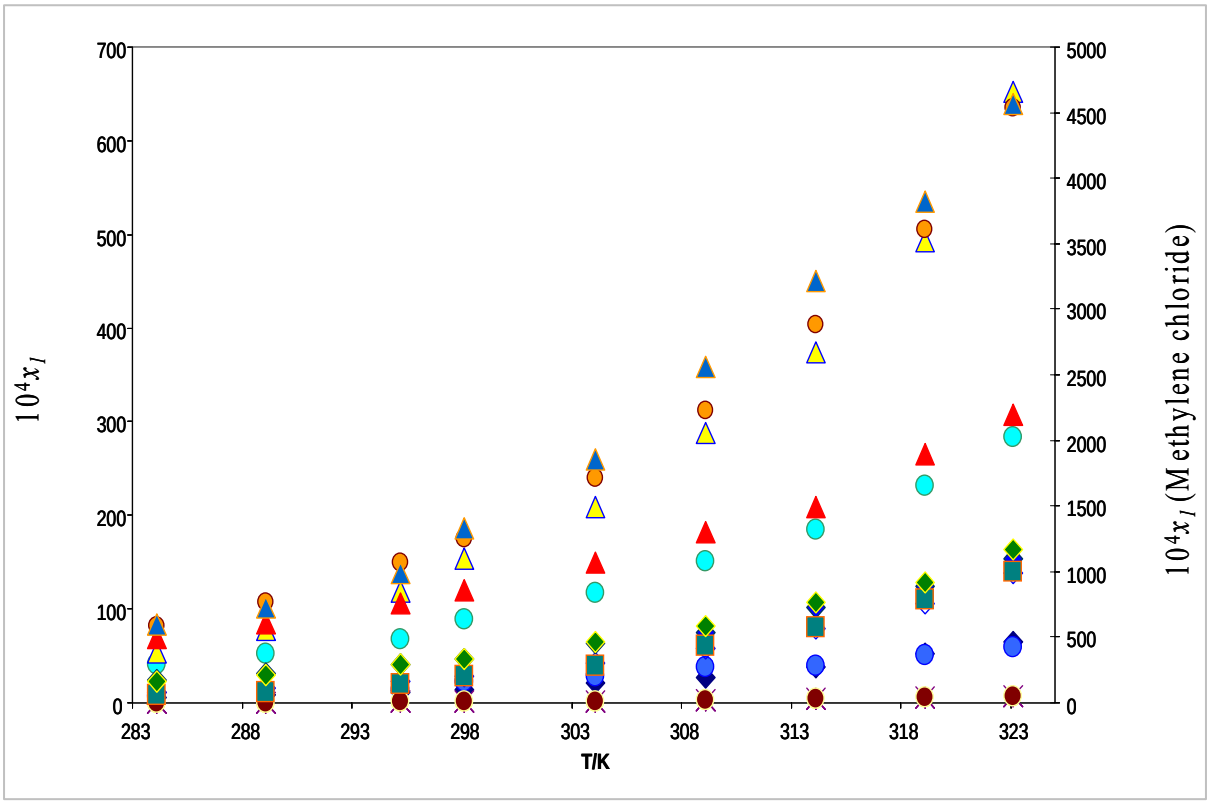


Figure 2.9: Equilibrium solubility of artemisinin in different solvents at different temperatures.

Table 2.5

Mole fraction solubility data of artemisnin, 10^4x, various organic solvents at different temperatures.

T/K	Methanol	Acetone	Ethyl acetate	Acetonitrile	Hexane	Heptane	Methylene Chloride	MTBE	MEK
284.1	5.7	41.3	53.2	7.2	0.5	0.5	592.1	11.1	162.6
289.1	7.8	52.1	77.4	11.2	0.7	0.7	721.6	15.3	211.6
295.21	11.6	68.3	118.3	14.6	1.0	1.1	823.3	23.1	293.2
298.1	14.2	88.3	154.3	22.2	1.3	1.3	1526.2	28.4	336.8
304.1	20.6	117.2	209.2	28.5	1.9	1.8	1858.2	42.1	461.2
309.1	27.2	151.2	288.5	38.4	2.8	2.8	2561.2	57.8	586.4
314.1	38.2	185.3	374.1	39.4	3.8	3.8	3217.3	79.2	758.4
319.1	51.6	231.45	492.3	51.2	5.3	5.4	3823.5	105.8	920.9
323.1	64.5	284.15	651.75	59.4	6.9	6.7	4567.9	137.7	1174.5

Table 2.6

Mole fraction solubility data of artemisnin, 10^4x, various organic solvents at different temperatures.

T/K	Ethanol	Butanol	Isopropyl acetate	Toluene
284.1	23.9	8.6	69.1	81.9
289.1	30.9	11.4	84.3	107.8
295.2	41.1	19.9	105.7	149.9
298.1	46.8	28.1	119.4	175.1
304.1	63.6	39.7	150.2	240.0
309.1	75.1	60.4	181.9	311.4
314.1	101.6	80.0	209.1	403.5
319.1	124.2	109.4	265.4	504.7
323.1	153.5	139.5	308.1	635.7

2.7.5 Solubility of lovastatin, simvastatin, or artemisinin in mixed solvents

Solubility of lovastatin in binary solvent mixtures of ethyl acetate and ethanol, ethyl acetate and acetone, and acetone and ethanol, are graphically displayed in Figures 2.10, 2.11, and 2.12, respectively.

Solubility of simvastatin in binary solvent mixtures of acetone and ethanol, ethyl acetate and ethanol, and ethyl acetate and acetone, are graphically displayed in Figures 2.13, 2.14, and 2.15, respectively.

Solubility of artemisinin in binary solvent mixtures of acetone and ethyl acetate, ethyl acetate and ethanol, and acetone and ethanol, are graphically displayed in Figures 2.16, 2.17, and 2.18, respectively.

As expected, solubility of each solute increased with increasing temperature. In the acetone-containing mixtures, solubility of each solute increased when the amount of acetone was increased, whereas in the ethyl acetate ethanol mixtures, solubility increased with increasing amounts of ethyl acetate.

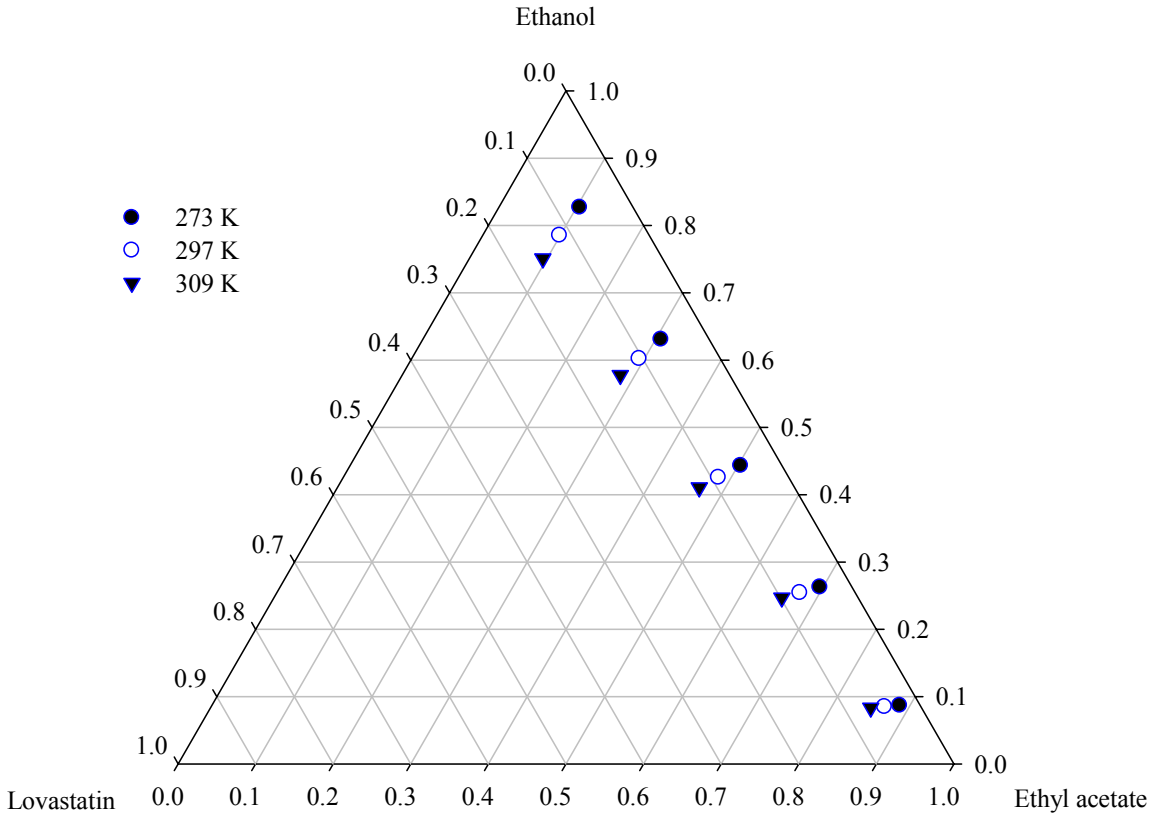


Figure 2.10: Solubility of lovastatin in ethyl acetate and ethanol binary solvent mixture at different temperatures.

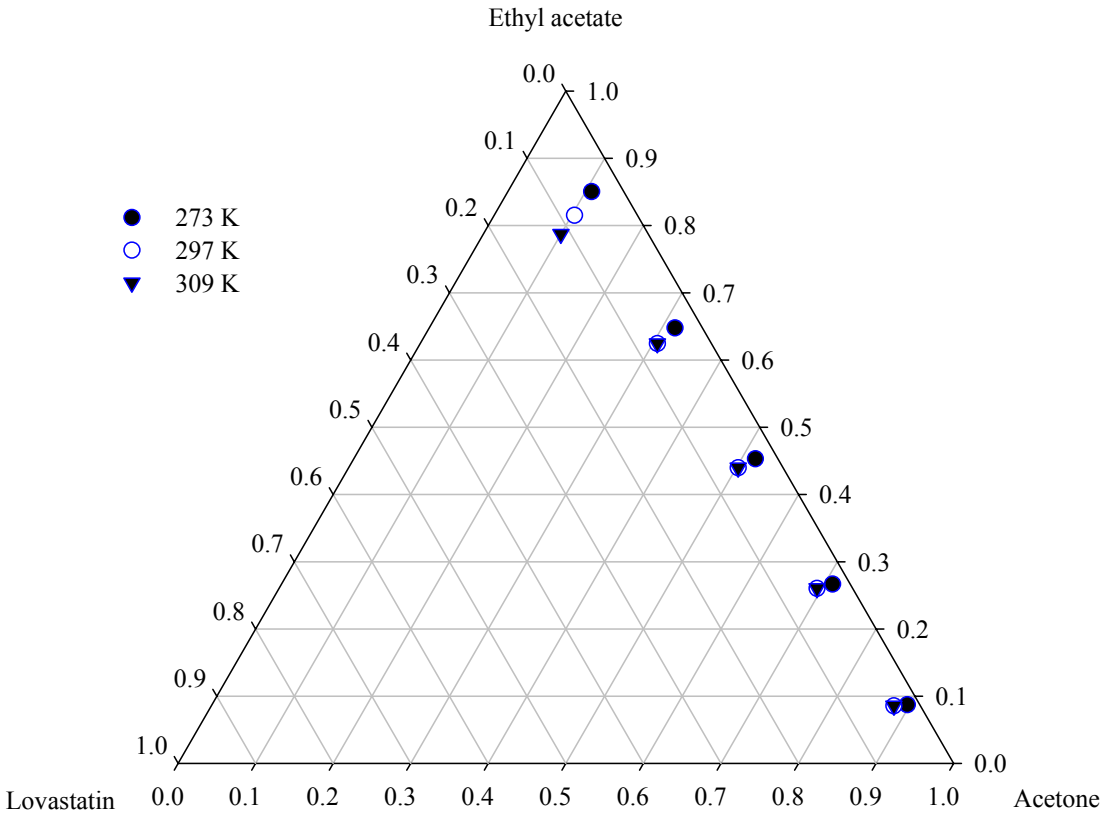


Figure 2.11: Solubility of lovastatin in ethyl acetate and acetone binary solvent mixture at different temperatures.

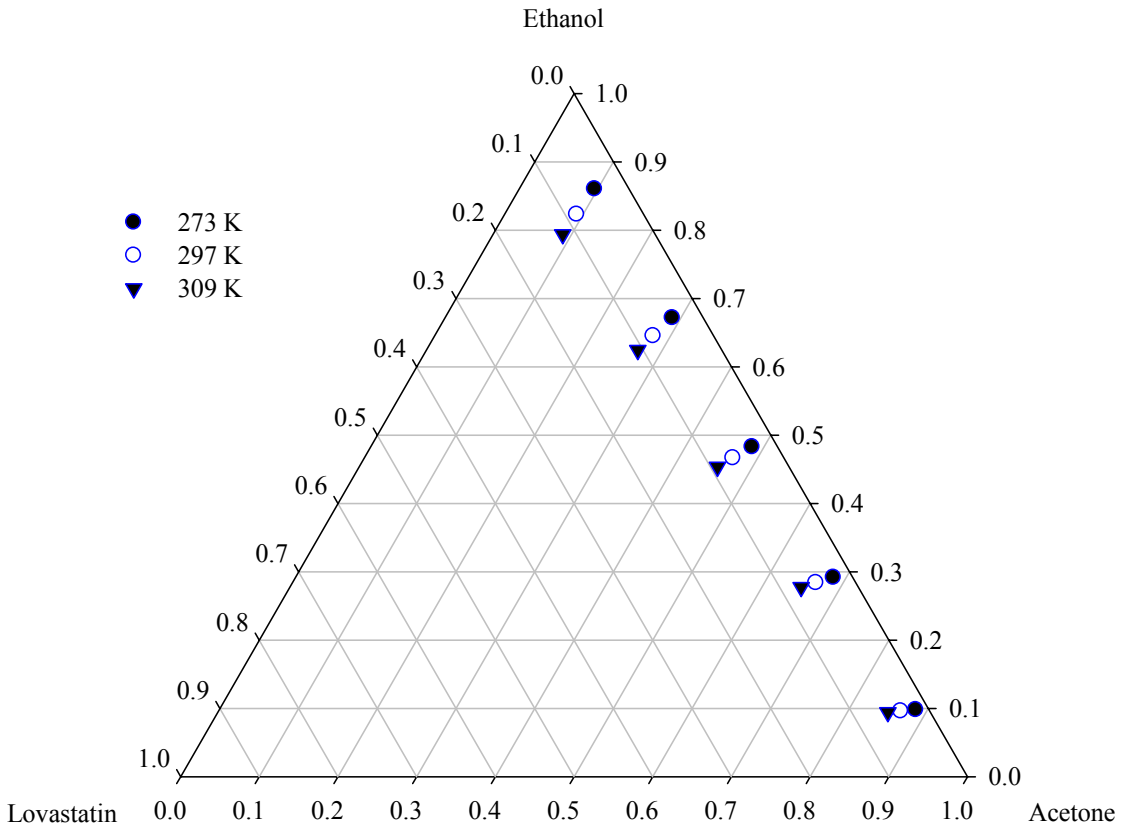


Figure 2.12: Solubility of lovastatin in ethanol and acetone binary solvent mixture at different temperatures.

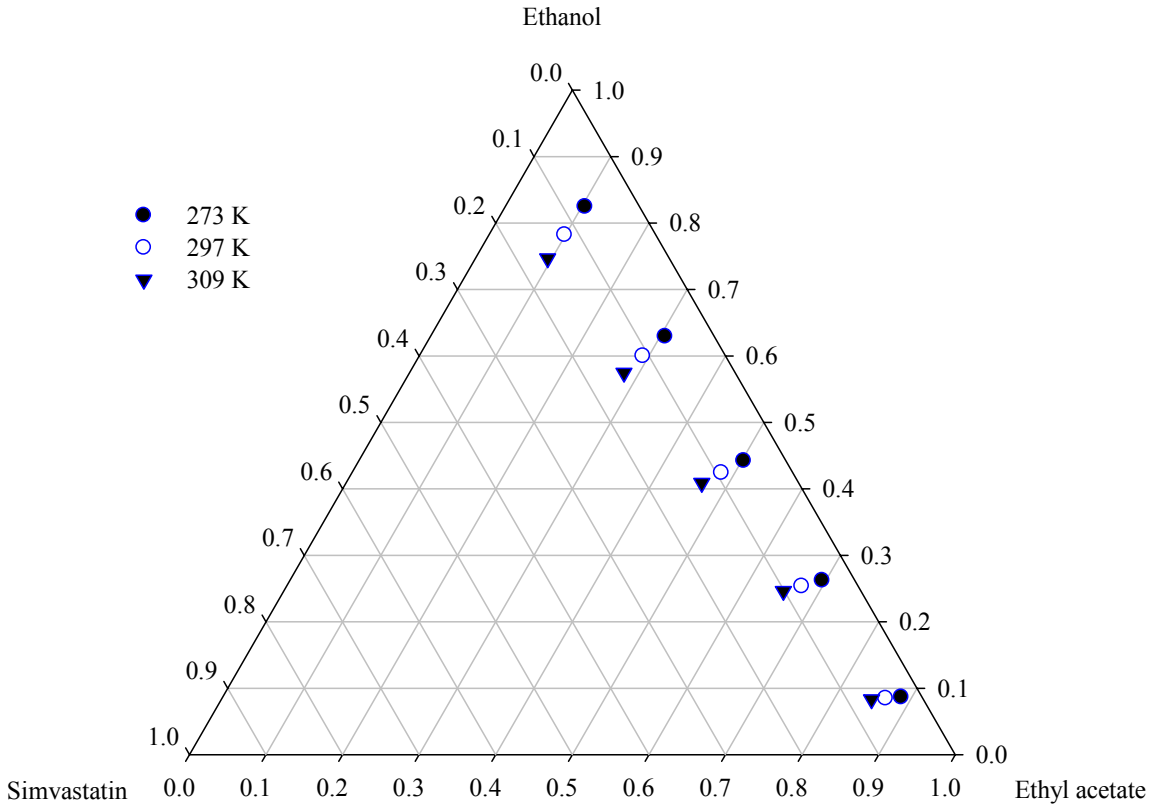


Figure 2.13: Solubility of simvastatin in ethanol and ethyl acetate binary solvent mixture at different temperatures

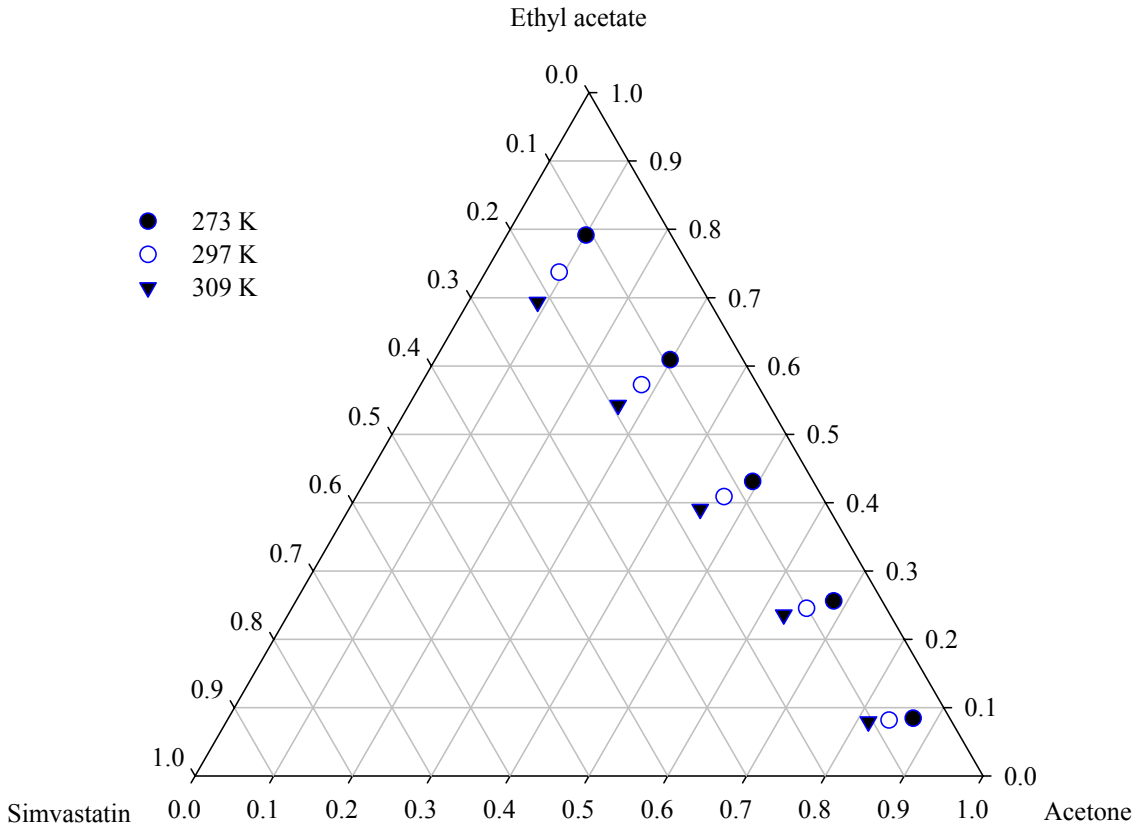


Figure 2.14: Solubility of simvastatin in ethyl acetate and acetone binary solvent mixture at different temperatures.

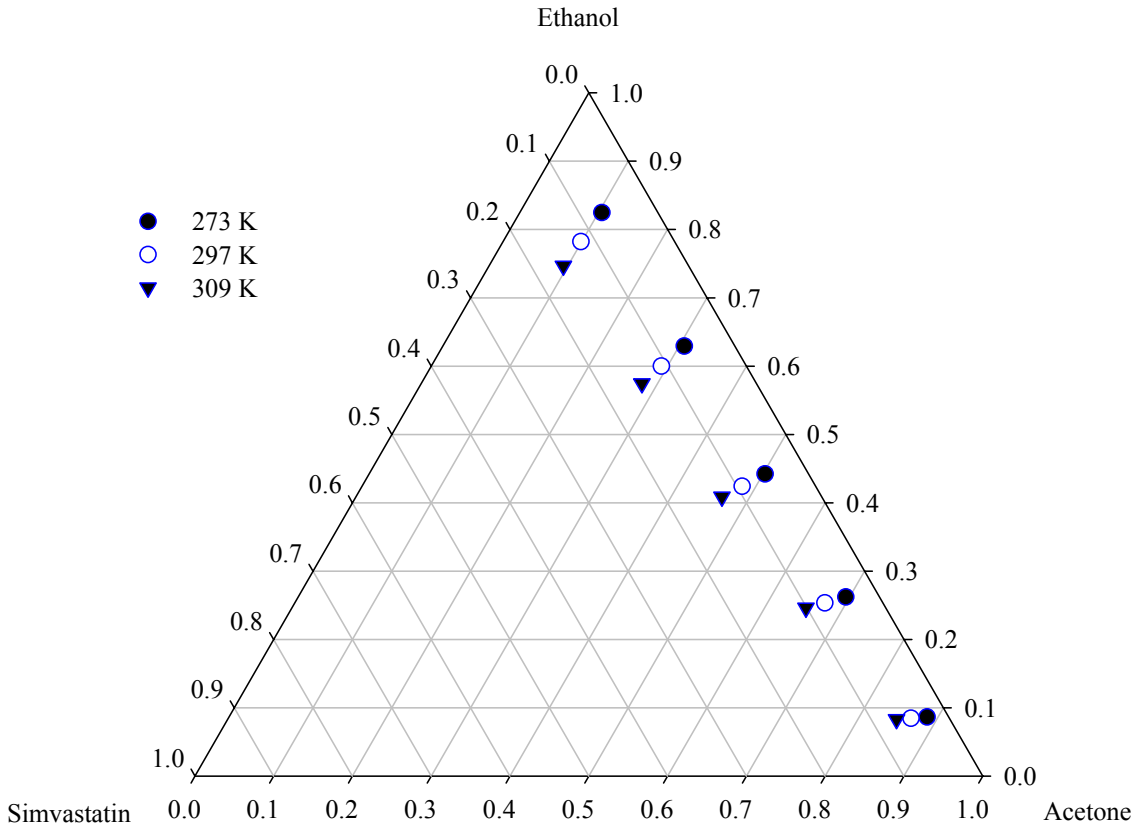


Figure 2.15: Solubility of simvastatin in ethyl acetate and ethanol binary solvent mixture at different temperatures.

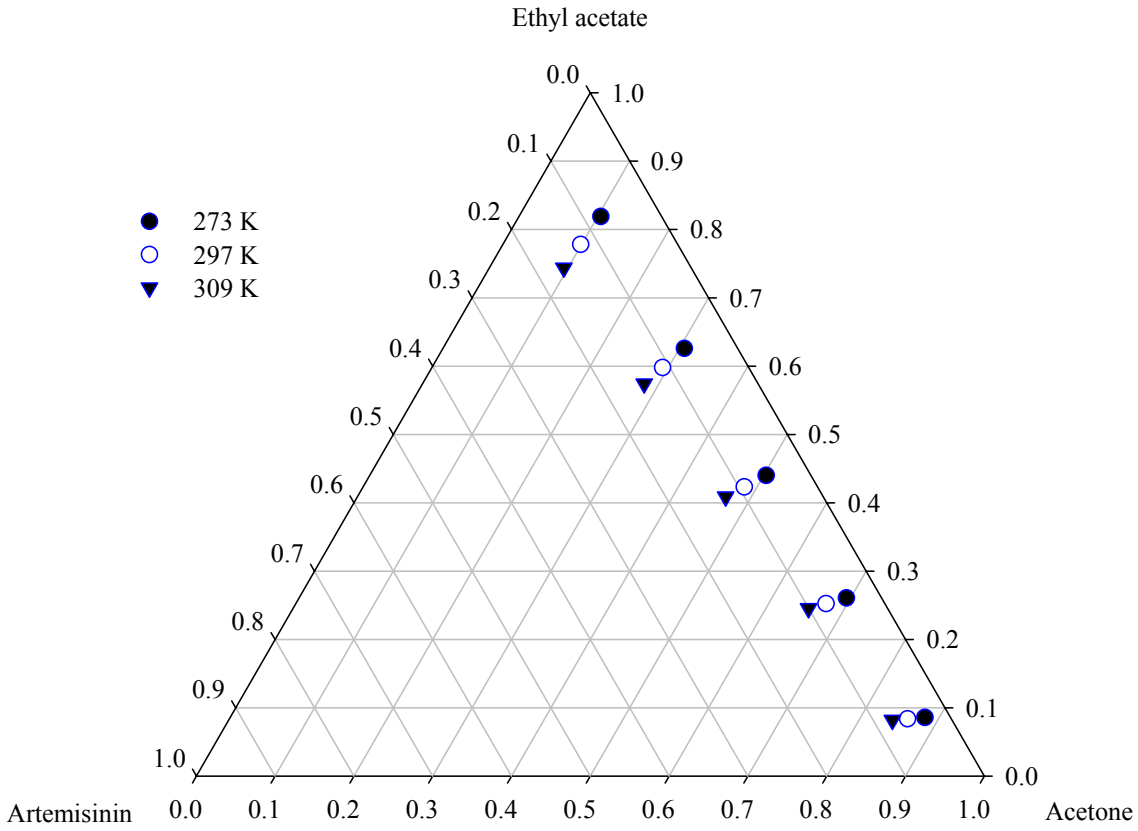


Figure 2.16: Solubility of artemisinin in ethyl acetate and acetone binary solvent mixture at different temperatures.

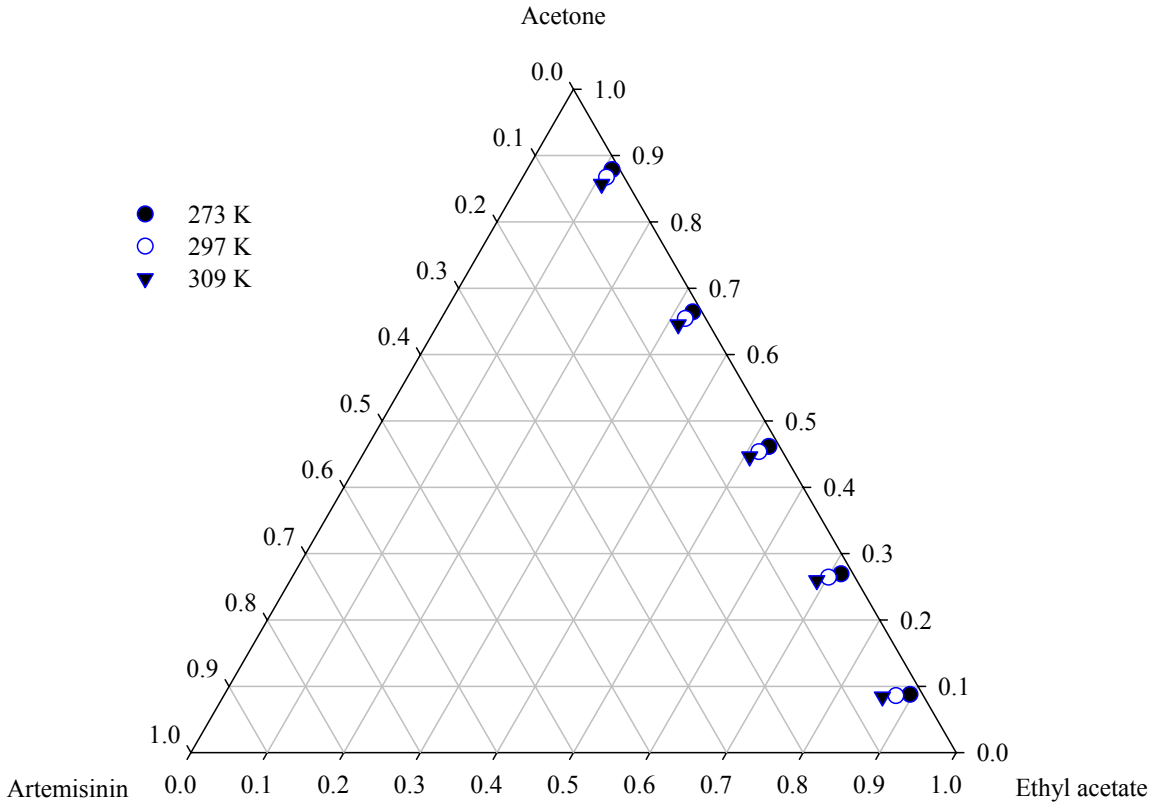


Figure 2.17: Solubility of artemisinin in ethyl acetate and ethanol binary solvent mixture at different temperatures.

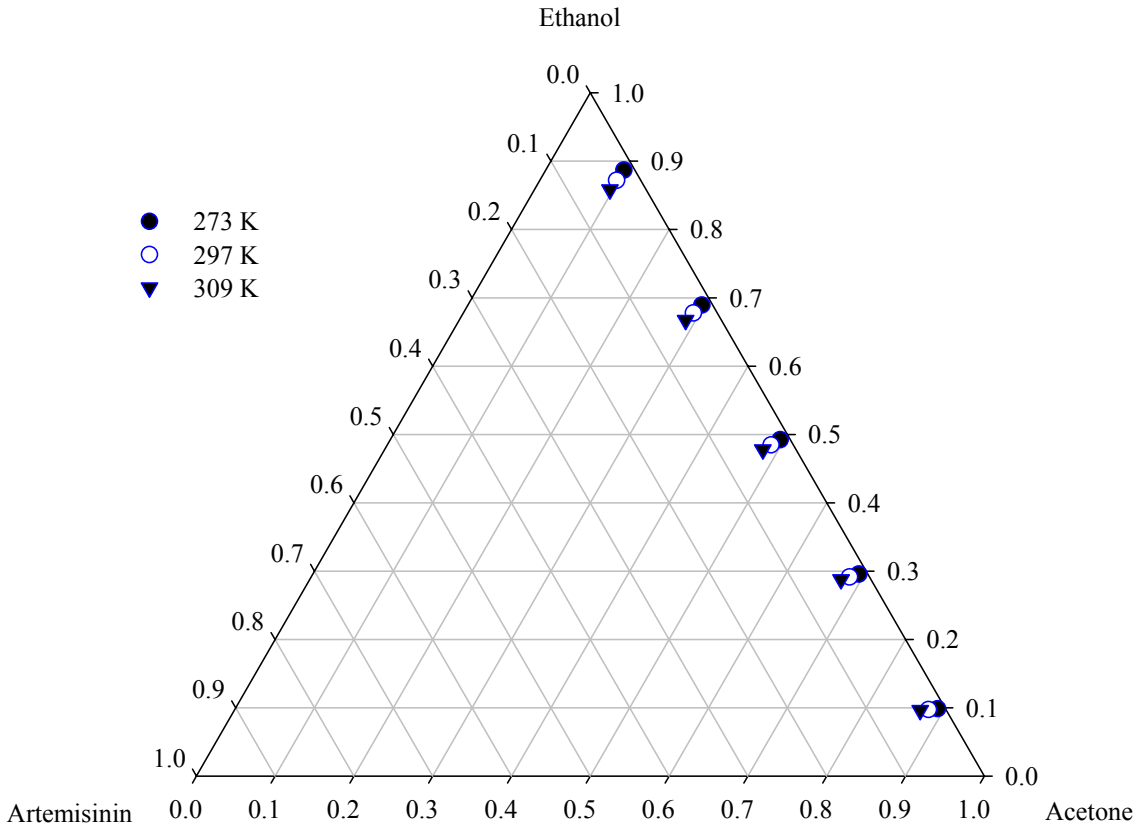


Figure 2.18: Solubility of artemisinin in acetone and ethanol binary solvent mixture at different temperatures.

2.8 Conclusion

The new scale down method that uses over fifty-fold less material compared to the conventional method has been designed to successfully measure the solubility of lovastatin, simvastatin, and artemisnin in several organic solvents. Results from the new protocol uncovered degradation of lovastatin in methanol.

Based on results of the solubility measurements, a we propose the use of hexane or heptane to extract the waxes from the *artemisia annua* plant, after which, artemisia can be extracted with toluene or methylene chloride. For re-crystallization of artemisinin, we propose the use of hexane or heptane as antisolvent (in conjunction with toluene or methylene chloride as co-solvent) to improve the cyrstallization yield. We envisioned that the use of new solvents will improve recovery for purification of artemisnin.

The solubility data was used to test and validate various assumptions involving ideal solubility calculation, as well as for activity coefficient and molecular modeling, which will be discussed in the next two chapters.

CHAPTER 3

IDEAL SOLUBILITY CALCULATION

Accurate computation of the ideal solubility of a crystalline solute in a liquid solvent requires knowledge of the difference in the molar heat capacity at constant pressure of the solid and the supercooled liquid solute, ΔCp. Since this is a hypothetical parameter, three assumptions are commonly used in the literature to estimate the parameter. In this chapter, the validity of the assumptions in calculating the ideal solubility of pharmaceutical compounds is examined for thermodynamic consistency. A simple approach of using solubility data to estimate the differential molar heat capacity is proposed and discussed.

3.1.1 Theoretical framework

The phase equilibrium equation for a pure solid solute, which dissolves in a liquid solvent, at temperature, *T,* and pressure *P*, given in terms of fugacities is as follows (Prausnitz et al., 1986):

$$f_i^S(T,P) = x_i \gamma_i(T,P,x_i) f_i^L(T,P) \tag{3.1}$$

The subscript S and L represent the solid and liquid, respectively.

Where $f_i^S(T,P)$, and $f_i^S(T,P)$ refer to the fugacity of the pure species as solids and as liquid respectively, at the temperature and pressure of the saturated solution, and x_i is the saturated mole fraction of the solid solute in the solvent.

By introducing the activity coefficient, γ_i, of the solute in the liquid phase, the equilibria equation becomes:

$$x_i \gamma_i(T, P, x_i) = \frac{f_i^S(T, P)}{f_i^L(T, P)} \tag{3.2}$$

If the temperature of the mixture is equal to the normal melting temperature of the pure solids, T_m, then:

$$f_i^S(T_m) = f_i^L(T_m) \tag{3.3}$$

and

$$\frac{f_i^S(T, P)}{f_i^L(T, P)} = x_i \gamma_i(T, P, x_i) = 1 \tag{3.4}$$

Thus, the equilibrium solubility mole fraction of a solute in a liquid at the solute's melting point temperature is simply equal to the reciprocal of its activity coefficient. Otherwise, as usually the case, the solid is below its melting point (Prausnitz et al., 1986).

$$f_i^S \leq f_i^L \tag{3.5}$$

Therefore Eq. (6) is not valid for predicting solute solubility in solvent. To predict the solubility in this case, Eq. (5) must be used with some estimates for the fugacity ratio:

$$\frac{f_i^S(T,P)}{f_i^L(T,P)} \tag{3.6}$$

Relating the sublimation (vapor) pressure to the vapor pressure:

$$f_i^S(T,P) \approx P^{sat}(T) \tag{3.7}$$

Note: since the sublimation (vapor) pressure of solid is generally small, the fugacity coefficient (Prausnitz et al., 1986):

$$\left(\frac{f_i^S}{P_i^{sat}}\right)_{sat,T} = 1 \tag{3.8}$$

If heat capacity data is available for both the solid and the solvent, and heat (enthalpy) of fusion is available, one can directly compute the Gibbs free energy change which is related to the fugacity ratio as follows:

$$\frac{\Delta G^{SL}(T,P)}{RT} = \frac{\Delta G_i^L(T,P) - \Delta G_i^S(T,P)}{RT} = \ln\frac{f_i^L(T,P)}{f_i^S(T,P)} \tag{3.9}$$

Thus,

$$\ln[x_i\gamma_i(T,P,x_i)] = \frac{-\Delta G^{SL}(T,P)}{RT} \tag{3.10}$$

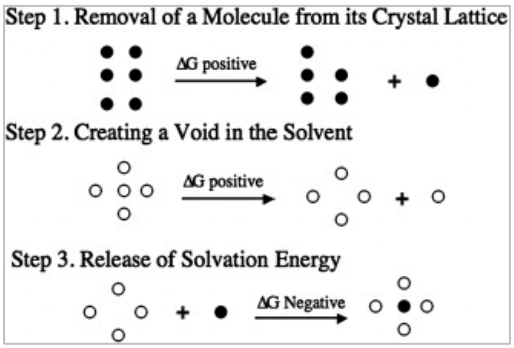
$\Delta G^{LS}(T)$ *is* calculated by separately calculating $\Delta H^{LS}(T)$ *and* $\Delta S(T)$, and then employing the relation $\Delta G^{LS}(T) = \Delta H^{LS}(T) - T\Delta S^{LS}(T)$ to compute the enthalpy and entropy changes of fusion.

Now, supposing that the change in state from solid (S) to liquid (L) of a solid occurs at temperature and pressure below the melting point to form a liquid is carried in the following three-step isobaric process Hildebrand and Scott, (1952); Bettini, (2004):

(1) The solute is heated at a fixed pressure from the temperature T to its normal melting temperature T_m.

(2) The solute is then melted at the melting temperature to form a liquid.

(3) The molten solute liquid is rapidly cooled from T_m back to T without solidification, then mixed with solvent as, as illustrated in the scheme below:



Scheme 1: An illustration of the three steps needed for drug solubility.

The enthalpy and entropy changes for this process can be mathematically expressed as:

$$\Delta H^{SL}(T) = \int_{T}^{T_m} C_P^S dT + \Delta H^{SL}(T_m) + \int_{T_m}^{T} C_P^L dT = \Delta H^{SL}(T_m) + \int_{T_m}^{T} \Delta C_P dT \tag{3.11}$$

$$\Delta S^{SL}(T) = \int_{T}^{T_m} \frac{C_P^S}{T} dT + \Delta S^{SL}(T_m) + \int_{T_m}^{T} \frac{C_P^L}{T} dT = \Delta H^{SL}(T_m) + \int_{T_m}^{T} \frac{\Delta C_P^L}{T} dT \tag{3.12}$$

And,

$$\Delta S^{fus}(T) = \int_{T}^{T_m} \frac{C_P^S}{T} dT + \Delta S^{fus}(T_m) + \int_{T_m}^{T} \frac{C_P^L}{T} dT = \Delta S^{fus}(T_m) + \int_{T_m}^{T} \frac{\Delta C_P}{T} dT \tag{3.13}$$

Note that,

$$\Delta C_p(T) = C_P^L(T) - C_P^S(T) \tag{3.14}$$

It is important to note that the last two equations relate the enthalpy and entropy changes of fusion at any temperature, *T,* to those changes at the melting at same pressure.

Now, since $G = H - TS$, and $\Delta G^{SL}(T = T_m) = 0$, Eq. (14) can be rewritten as:

$$\Delta S^{SL}(T) = \frac{\Delta H^{SL}(T_m)}{T_m} + \int_{T_m}^{T} \frac{\Delta C_P}{T} dT \tag{3.15}$$

and therefore,

$$\Delta G^{SL}(T) = \Delta H^{SL}(T) - T\Delta S^{SL}(T) \tag{3.16}$$

$$= \quad \Delta H^{fus}(T_m)\left[1 - \frac{T}{T_m}\right] + \int_{T_m}^{T} \Delta C_P \, dT \; - T \int_{T_m}^{T} \frac{\Delta C_P}{T} dT \tag{3.17}$$

$$= RT \ln \frac{f_i^{L}(T, P)}{f_i^{S}(T, P)}$$

Substitution of Eq. (3.9) into Eq. (3.17) yields,

$$\ln x_i \gamma_i = \frac{-\Delta H^{SL}(T_m)}{RT}\left[1 - \frac{T}{T_m}\right] - \frac{1}{RT}\int_{T_m}^{T} \Delta C_P dT + \frac{1}{R}\int_{T_m}^{T} \frac{\Delta C_P}{T} dT \tag{3.18}$$

where x_1, γ_1, T_m, $\Delta H^{fus}(T_m)$, $\Delta C_P(T)$, R, and T represent the solubility mole fraction of the solute (denoted as component 1) in solution, activity coefficient of the solute in solution, melting temperature of the solute, enthalpy of fusion of the pure solute at melting temperature, differential molar heat capacity of the pure solute (that is, the difference between the molar heat capacity of the solid and the liquid at temperature, T, their melting temperature), gas constant, and temperature, respectively.

The right hand side of Eq. (3.18) is identified as the ideal solubility, $x^{id}(T)$, and hence may be re-arranged to give:

$$\ln x_1(T) = \ln x_1^{id}(T) - \ln \gamma_1(T) \tag{3.19}$$

It is seen that the actual solubility, x_1, is directly impacted by the both ideal solubility and activity coefficient. Thus, the mole fraction solubility of solids in various solvents can be expressed as the sum of two terms: the ideal solubility, and the activity coefficient of the solute in the solution. As solutions become more ideal, the activity coefficient approaches unity. However, only rarely does the experimentally measured actual solubility compare favorably with the calculated ideal solubility. Hence, the activity coefficient which depends on the type of both the solute and the solvent, as well as the solution temperature, must be accounted for, in order to ensure good agreement between estimated and experimentally measured solubility. The ideal solubility of a solid solute is dependent on its fusion properties, like melting point, entropy of fusion, and heat-capacity change on melting. The melting point is one of the first properties measured during the early phase of drug compound development. Hence, the melting point of a compound is almost always available. The entropy of fusion is determined by measuring the heat of fusion at the melting point. The parameter that is most difficult to obtain from the literature or to experimentally measure is the heat capacity change on melting (Finke et al., 1977, Cassellato et al., 1973).

In what follows, we will discuss methods to calculate $x_1^{id}(T)$ accurately, and modeling of activity coefficient is discussed in Chapter 4.

3.1.2 Calculation of ideal solubility calculation

As reflected in Eq. (3.19), to calculate $x_1^{id}(T)$, it is necessary to know T_m, ΔH^{SL}, and $\Delta Cp(T)$ of the solute. The solid-liquid differential molar heat capacity, ΔCp, is required to accurately estimate the ideal solubility of crystalline solids in liquid solvents. And, accurate ideal solubility is paramount for successful modeling of solids solubility in liquid solvents. However, ΔCp is a hypothetical value, and therefore very difficult to measure experimentally. Therefore, three assumptions are commonly used to estimate the value:

Assumption I: Several researchers, including Mishra and Yalkowsky, Prausnitz *et al.,* and Abildskov and O'Connell (2003), have assumed that, the quantity ΔC_P is negligible, and therefore can be ignored. This simplifies Eq. (3.18) to:

$$\ln x_1^{id} = \frac{\Delta_{fus} H}{RT_m}\left(1 - \frac{T_m}{T}\right) \tag{3.20}$$

Assumption II: Neau and co-workers (1990) have assumed that ΔC_P can not be ignored, but be approximated by assuming it to be independent of temperature and, equals to the value at the melting temperature of the solute:

$$\Delta C_P(T) = \Delta C_P(T_m) = C_P^L(T_m) - C_P^S(T_m)$$

which simplifies Eq. (3.18) to:

$$\ln x_1^{id} = \frac{\Delta_{fus}H}{RT_m}\left[1-\frac{T_m}{T}\right]+\frac{\Delta C_P}{R}\left[\frac{T_m}{T}-1+\ln\left(\frac{T}{T_m}\right)\right] \tag{3.21}$$

Assumption III: Finally, other researchers such as Hildebrand *et al.,* (1970), Mauger *et al.,* (1972) and Grant *et al.* (1984) have agreed with Neau and co-workers (1990) that ΔC_p can not be ignored, but can alternatively approximated to the entropy of fusion at the solute melting point, ΔS_m. This simplifies Eq. (3.18) to:

$$\ln x_1^{id} = \frac{-\Delta_{fus}H}{RT_m}\ln\left(\frac{T_m}{T}\right) \tag{3.22}$$

Here, the validity of the assumptions in predicting the ideal solubility of the three compounds discussed in the previous chapters, was evaluated.

To calculate the ideal solubility for each compound, T_m, and $\Delta_{fus}H(T_m)$ were measured by differential scanning calorimetry (DSC). Solid and liquid heat capacities of each compound near its melting point were also determined by DSC, and linear equations describing the heat capacities were extrapolated to the melting point to obtain $\Delta C_p(T_m)$.

3.2 Experimental

3.2.1 Equipment

Mettler AE 160 digital analytical balances with sensitivity of 0.01 mg; 2910 Modulated DSC, TA Instruments differential scanning calorimeter; DSC822e, Mettler-Toledo differential scanning calorimeter.

3.2.2 Differential scanning calorimetric (DSC) measurements

Determination of melting points, enthalpies of fusion, and purity analysis were performed by differential scanning calorimetry (DSC). All measurements were carried out at a heating rate of 10 °C min^{-1} in a dynamic nitrogen atmosphere (50 mL min^{-1}). About 4 mg of the model compound was used for each DSC experiment. The equipment was calibrated using indium as a standard. All thermal analyses were carried out in tripplicates. The heat capacity measurement procedure involved placing an empty crucible and lid in the sample compartment to serve as reference. The reference or sample was initially equilibrated isothermally, then heated at 0.5°C/min for 10°C. Thereafter it was equilibrated isothermally at the higher temperature for 60 minutes. Since heat capacity measurements closer to the melting temperature may present some anomalies due to phase change effect, sample heat capacities were measured at about 20 degrees away from the melting point temperature. For each ten minute heating period, the thermogram showed a trapezoidal peak, which represents the differential flow of heat (dH/dt) to the sample that was necessary to keep the sample and the reference side at essentially the same temperature at every point during the run. Integration of this peak area with respect to time, as obtained by the software, gave the total heat provided to the

empty sample pan and its lid (ΔH_{P+l}). Once the temperature range of interest was covered, about 5 mg of the pertinent powder was carefully weighed into the sample pan, and the lid was then crimped on. This unit was then placed into the sample compartment of the calorimeter, and runs similar to the ones described for the empty sample and lid were carried out. Thus, the heat introduced into the unit consisting of the sample pan, lid and the sample (ΔH_{P+l+s}) was determined and the corresponding (ΔH_{P+l}) for a specific temperature range equals the enthalpy need of the sample (ΔH_s) This is related to the specific molar heat capacity by the expression:

$$\Delta H_S = \frac{m \overline{C_P} \Delta T}{M} \qquad (3.23)$$

where *m* is the mass of the solute powder, *M* is the molecular weight of the solute and ΔT is the temperature range of each run (10 degrees for this work). $\overline{C_P}$ is the specific molar heat capacity of the sample at constant pressure, and the value of $\overline{C_P}$ obtained for the 10 degree range was assigned to the mean temperature (for example, the heat capacity determined for run from 380K to 390K was assigned to 385K). To test the validity of the calorimeter to accurately measure heat capacities of crystalline organic compounds, anthracene heat capacities were determined in the temperature range where lovastatin heat capacity was measured. Biphenyl was used as corresponding for liquid heat capacity measurement. The experimental heat capacities of these test samples were compared to literature values (Elder, 1986). Heat capacities for lovastatin and test samples were measured in triplicates. Statistical analysis of the data was executed, and if necessary,

individual heat capacities were excluded if they were more than two standard deviations away from the mean heat capacity for that temperature range.

3.3 Results and discussion

3.3.1 Calorimetric data for lovastatin

The melting temperature and heat of fusion of the lovastatin crystalline powders were measured by differential scanning calorimetry (DSC) to be 445.6 ± 0.5K, and 43,169 ± 500 J/mol, respectively. The values are in good agreement with literature values (Elder, 1986). The differential molar heat capacity, ΔC_P , was also measured by DSC, employing a similar protocol used by Neau *et al.* (1990), and linearly extrapolated to the melting point to be 177 ± 5 J/mol/K.

3.3.2 Calorimetric data for simvastatin

The melting temperature and heat of fusion of the simvastatin crystalline powders were measured by DSC to be to be 412.6 ± 0.5K, and 32,170 ± 400 J/mol, respectively, respectively, and the ΔC_P *(Tm)* was linearly extrapolated to the melting temperature to be 149 ± 5 J/mol/K.

3.3.3 Calorimetric data for artemisinin

The melting temperature and heat of fusion of the simvastatin crystalline powders were measured by DSC to be to be 429.6 ± 0.5 K, and 24300 ± 400 J/mol, respectively, respectively. Extrapolation to the melting point to estimate ΔC_P *(Tm)* could not be carried as the compound appeared to severely degrade at less than 10 K after melting.

3.4 Calculated ideal solubilities of lovastatin and simvastatin

Eq. (3.20), (3.21), and (3.22), were used to calculate the ideal solubilities of lovastatin and simvastatin, as a function of temperature for assumptions I, II, and III, respectively. The results are presented in Figures 3.1 and Figure 3.2, for lovastatin and simvastatin, respectively. Consistent with observation made by Neau and co-workers (1990), ΔC_P has significant impact on calculated ideal solubility, especially at temperatures far below the melting temperature. Both assumption I ($\Delta C_P(T) = 0$) and assumption III ($\Delta C_P(T) = \Delta S(T_{m)}$)) give lower ideal solubility when compared to assumption III ($\Delta C_P(T) = \Delta C_P(T_m)$).

On the other hand, as illustrated in 3.2, at temperatures close to the melting temperature of simvastatin, all three assumptions converge to a common ideal solubility. As most pharmaceutically-important compounds are often temperature-labile, the processes for purifying them (e.g., extraction, crystallization, etc) are often carried out at temperatures far below the solute melting point. Therefore, the importance of using the appropriate

assumption when modeling solubility data for pharmaceutical compounds, can not be over-emphasized.

In the next section, van't Hoff-like analyses were carried out to assess which assumption gives thermodynamically consistent ideal solubility.

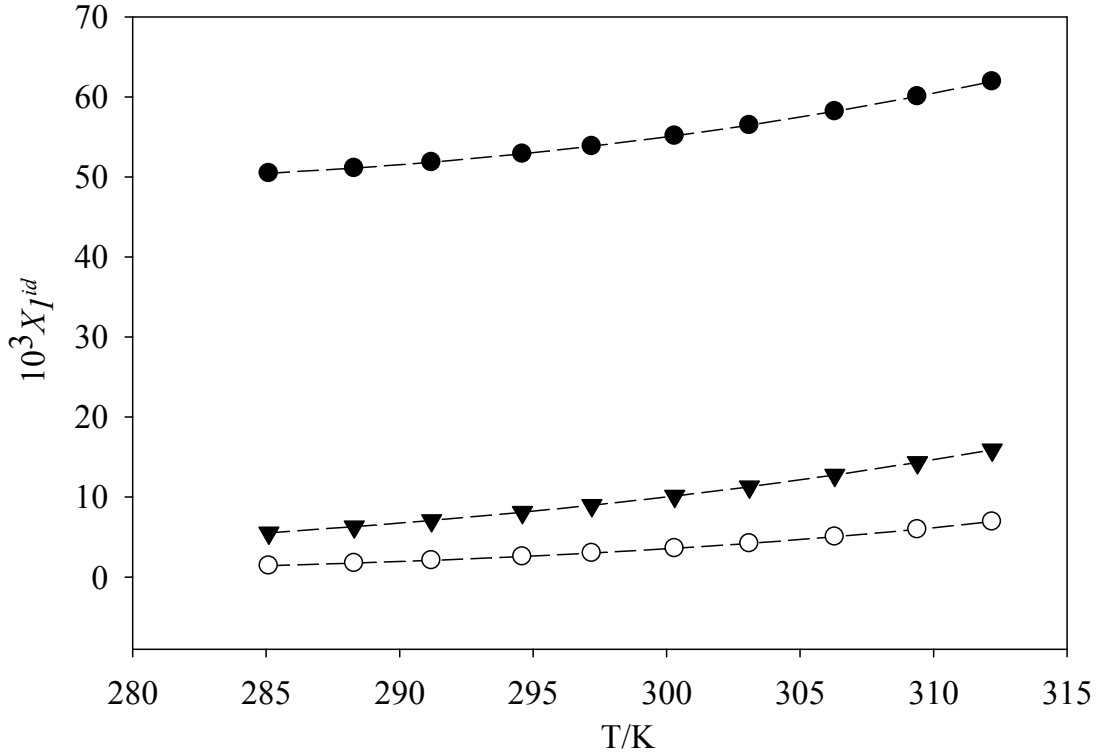


Figure 3.1: Calculated ideal solubility of lovastatin as function of temperature using various assumptions involving ΔCp.

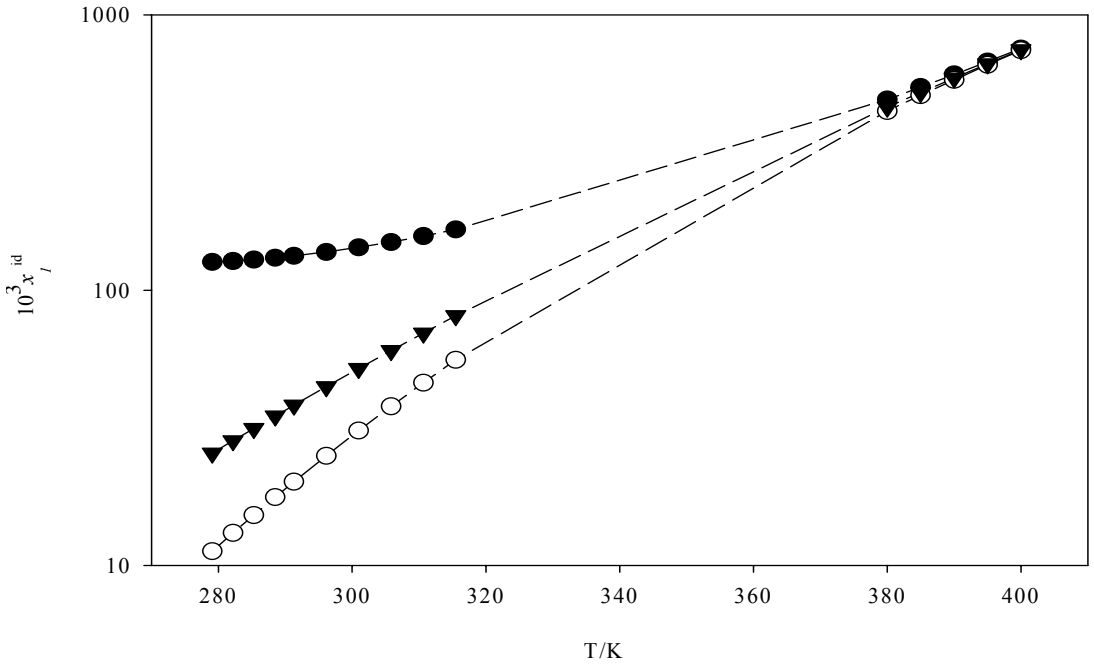
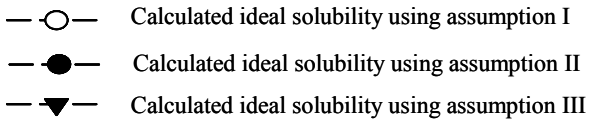


Figure 3.2: Calculated ideal solubility of simvastatin as function of temperature using various assumptions involving *ΔCp*.



3.5 A van't Hoff-like analysis of the activity coefficient data

The activity coefficient of crystalline solute, in liquids, as a function of temperature, was calculated from the experimentally measured solubility data and calculated ideal solubility from each assumption:

$$\ln \gamma_1 = \ln x_1^{id} - \ln x_1^{Exp} \tag{3.25}$$

And, for a narrow temperature range, $\overline{H}_1^E$ can be assumed to be constant, and $\ln \gamma_1$ can be expressed as:

$$\ln \gamma_1 = \frac{\overline{H}_1^E}{RT} - \frac{\overline{S}_1^E}{R} \tag{3.26}$$

where $\overline{H}_1^E$ and $\overline{S}_1^E$ represent the partial molar excess enthalpy, and partial molar excess entropy, respectively, and are assumed to be temperature independent.

The calculated values were used to obtain van't Hoff plots of $\ln \gamma_1$ versus $1/T$, which are displayed in Figures 3.3, 3.4, and 3.5, and each slope of the plot represent estimates of $\left(\overline{H}_1^E / R \right)$. And, for dissolution of lovastatin and simvastatin in the solvents studied, since the enthalpy of mixing, ΔH_{mixing}, were measured to be positive, it would be reasonable to expect $\overline{H}_1^E$ to be positive.

However, when the ideal solubility calculated from Eq. (3.20) (Assumption I, ΔC_P=0) was used to determine $\ln \gamma_1$, the resulting activity coefficients of lovastatin in the various alcohols are displayed in Figure 3.3. It is worthy of note that, the van't Hoff plot of $\ln \gamma_1$ versus $1/T$ shows linear correlation, however, the slope is negative. This implies that assumption I yields $\overline{H}_1^E$ that is inconsistent with the experimental value of $\overline{H}_1^E > 0$.

On the other hand, for assumption II, a good fit of the plot was obtained, ant the slope for alcohol were positive, indicating that $\overline{H}_1^E$ is greater than zero. This is consistent with the fact that measured ΔH_{mixing} was positive (Figure 3.4).

Finally, in the case of Assumption III, ($\Delta C_P = \Delta S$), the van't Hoff plot of $\ln \gamma_1$ is scattered with some of the solvents exhibiting negative slope, whereas others had positive slope (Figure 3.5).

From the data generated, we found that assumption II gives the most thermodynamically consistent calculated ideal solubility. However, a major limitation with this assumption is that it would be difficult to extrapolate experimental $\Delta C_P(T)$ data from temperatures above the melting point for compounds that degrade when heated above the melting point, as that was the case artemisnin. To this end, a new approach is proposed to obtain the differential heat capacity. The technique is discussed in the next section.

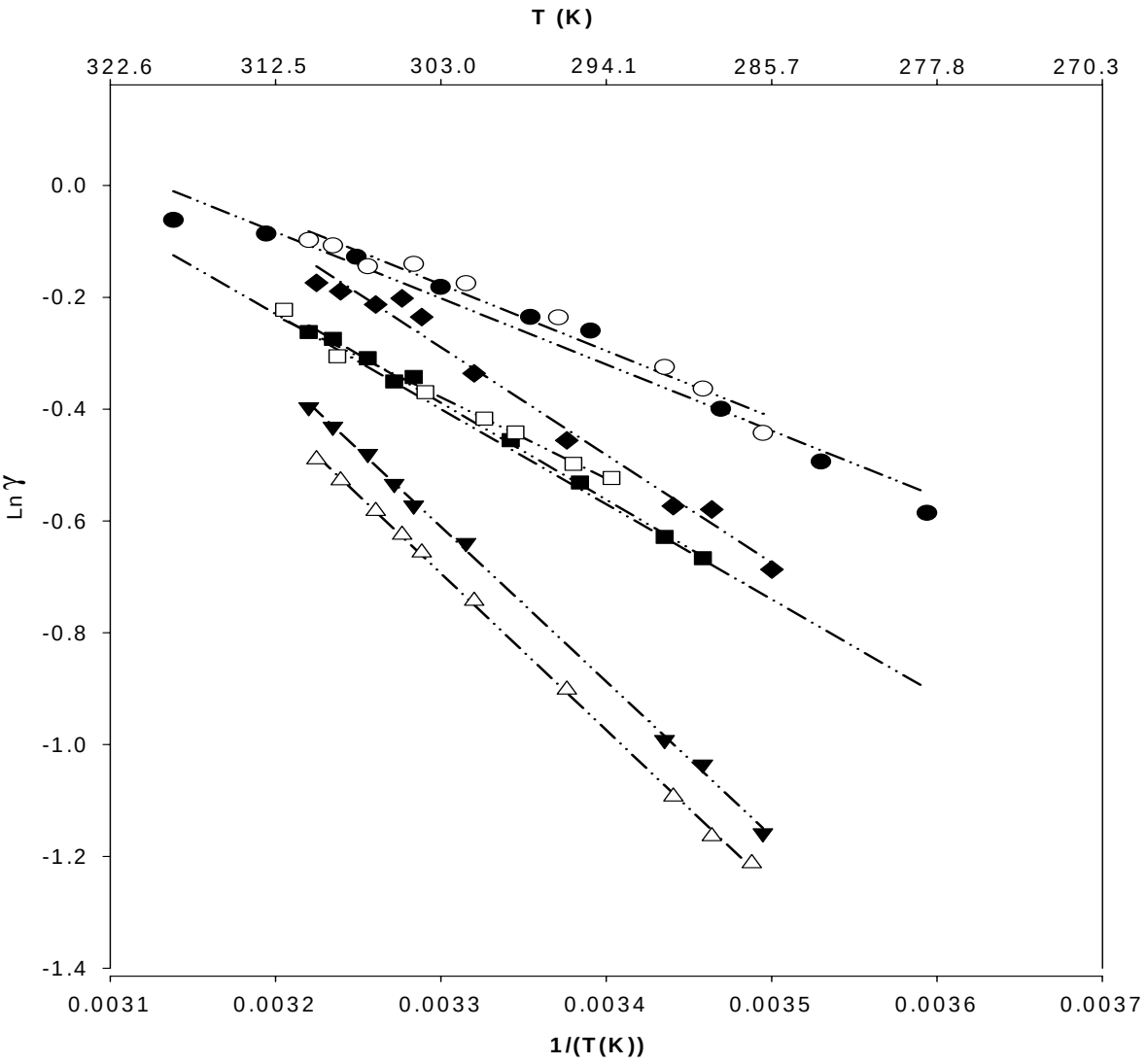


Figure 3.3: van't Hoff plot temperature-dependent activity coefficient at infinite dilution data for lovastatin in different alcohols using the assumption I ($\Delta C_p = 0$; Eq. (3.20).

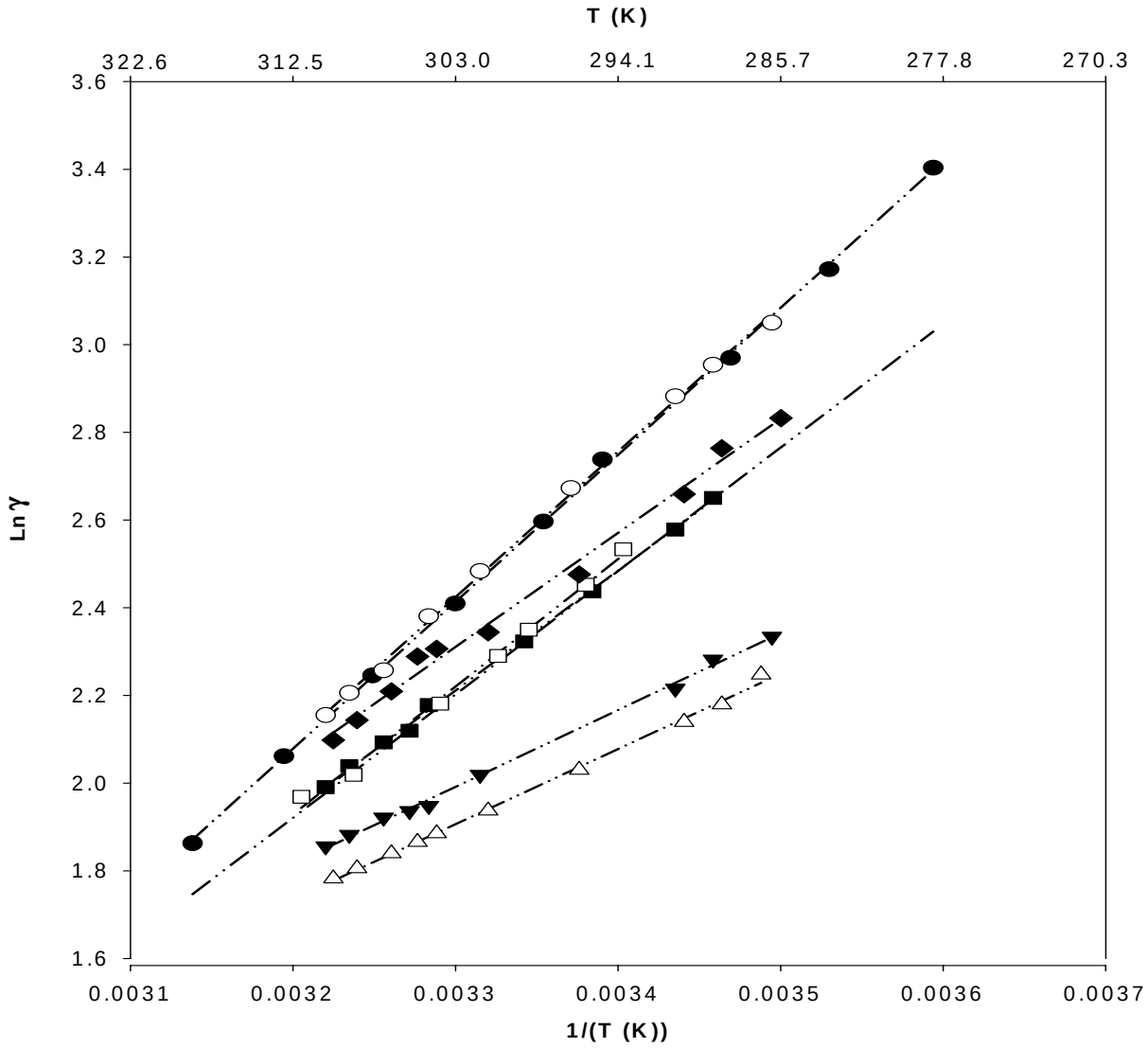


Figure 3.4: Van't Hoff plot of temperature-dependent activity coefficient at infinite dilution data for lovastatin in different alcohols using the assumption II ($\Delta C_p = \Delta C_p(Tm)$; Eq. (3.21)).

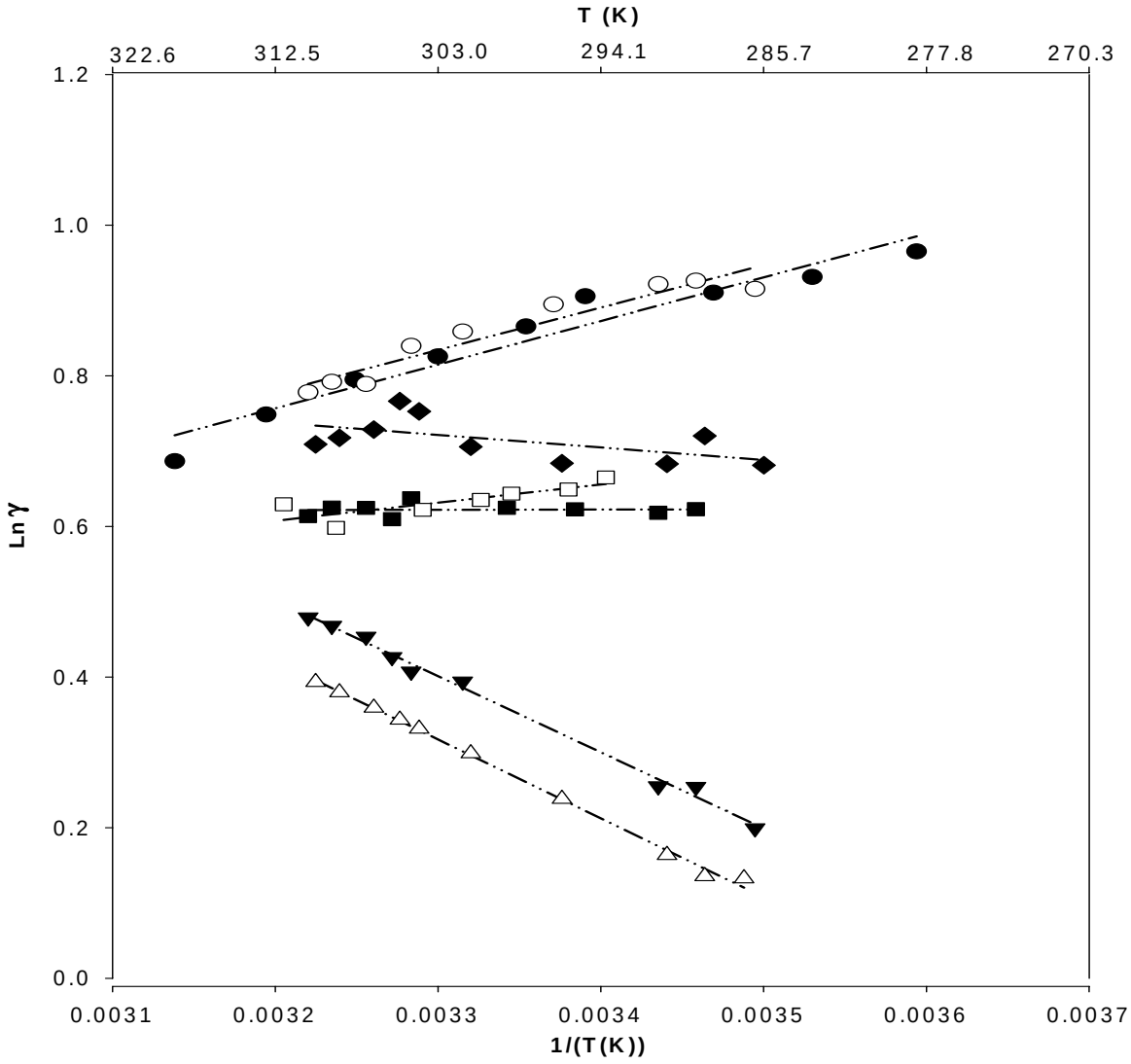


Figure 3.5: van't Hoff plot of temperature-dependent activity coefficient at infinite dilution data for lovastatin in different alcohols using assumption II ($\Delta C_p = \Delta S(Tm)$; Eq. (3.22)).

3.6 Revised technique for obtaining the differential molar heat capacity for artemisinin.

In this technique, since the solubility of artemisinin is very low in both heptane or hexane (order 10^{-4}), an assumption of infinite dilution was invoked to estimate $\Delta C_P(Tm)$.

Thus expressing Eq. (3.23) in terms empirical constants, *A* and *B* yields:

$$\ln \gamma_1^{\infty} = \frac{\overline{H}_1^{E,\infty}}{RT} - \frac{\overline{S}_1^{E,\infty}}{R} = \frac{A}{T} + B \tag{3.27}$$

Combining Eq. (3.18) with Eq. (3.27) and rearranging, derives:

$$\ln x_1^{\infty} = \left[\frac{\Delta H^{fus}(T_m)}{RT_m} - \frac{\Delta C_P(T_m)}{R}\left[1 + \ln T_m\right] - A\right] + \left[\left(\frac{\Delta C_P(T_m).T_m}{R} - \frac{\Delta H^{fus}(T_m)}{R}\right) - B\right]\frac{1}{T} + \frac{\Delta C_P(T_m)}{R}\ln T \tag{3.28}$$

Or,

$$\ln x_1 = a + \frac{b}{T} + c \ln T \tag{3.29}$$

where

$$a = \frac{\Delta H^{fus}(T_m)}{RT_m} - \frac{\Delta C_P}{R}\left[1 + \ln T_m\right] - A\,, \tag{3.30}$$

$$b = \left(\frac{\Delta C_P \cdot T_m}{R} - \frac{\Delta H^{fus}(T_m)}{R}\right) - B\,, \tag{3.31}$$

$$c = \frac{\Delta C_P}{R} \tag{3.32}$$

Here, it is proposed that the parameters *a*, *b* and *c* can be obtained by regressing Eq. (3.31) against experimental solubility data in which solubility of the solute is infinitely dilute. For the case of artemisinin, solubility in hexane and heptane were used in the regression, using Sigma Plot.

3.6.1 Results and discussion

The regression gave excellent fit, as illustrated in Figures 3.6, and 3.7, yield an average $\Delta Cp(T_m)$ value of 77 J/mol/K; the values of $\Delta Cp(T_m)$ are displayed in Table 3.1. The regression report are displayed in the Appendix.

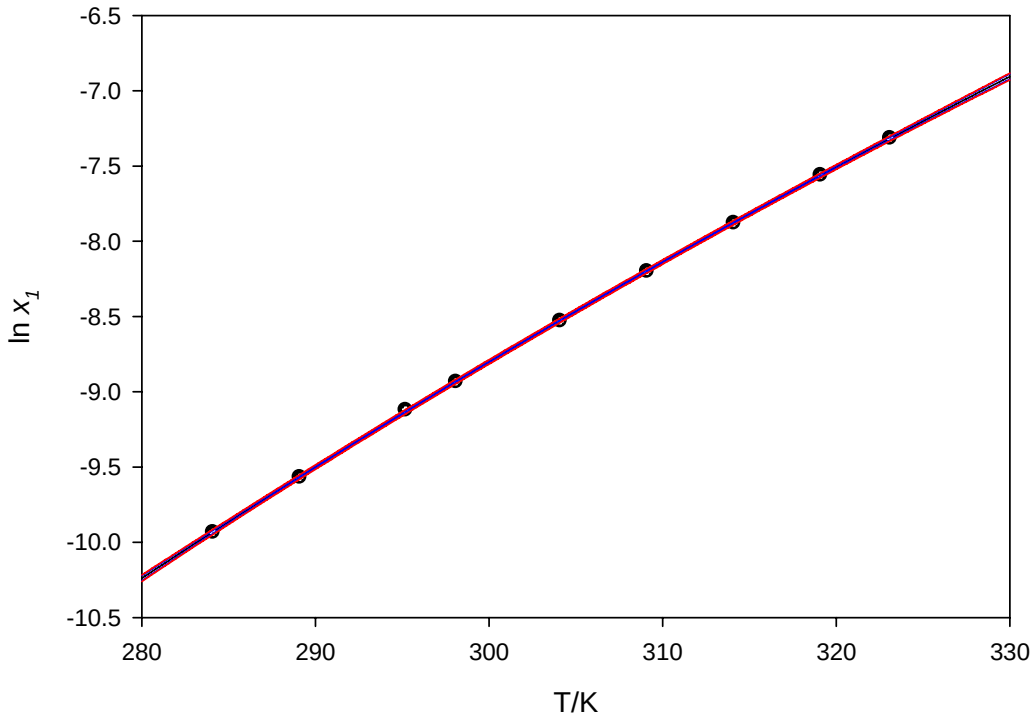


Figure 3.6: Non-linear regression of solubility mole fraction artemisinin in hexane.

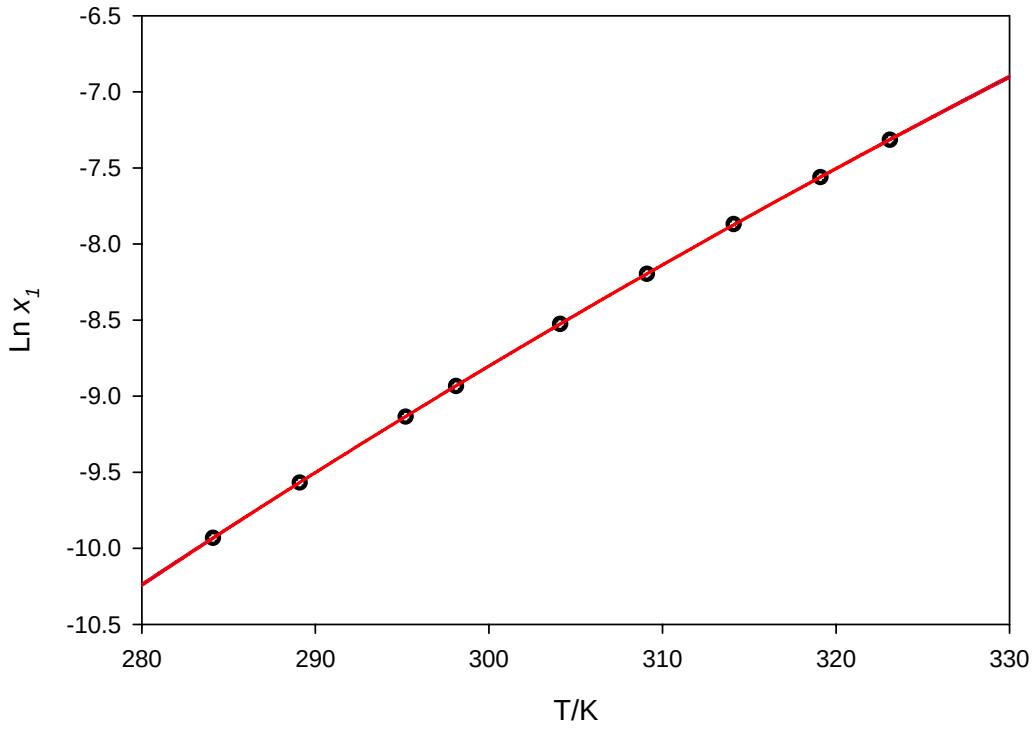


Figure 3.7: Non-linear regression of solubility mole fraction artemisinin in hexane.

3.7 Conclusion

Pharmaceutical compounds are labile and therefore are purified at temperatures, far lower than their melting points. And, because each of the equations involving assumptions regarding differential molar heat capacity yielded relatively different values of the calculated ideal solubility, it is recommended that the differential heat capacity term should be validated, and accounted when modeling solubility of pharmaceutical compound in liquid solvents. A new thermodynamically rational approach has been used to successfully obtain the term for artemisinin. The method uses regression of infinitely mole fraction solubility data to the absolute temperature.

The thermodynamically consistent calculated ideal solubility (Eq. (3.28)) was used for activity coefficient modeling of the solubility; this is discussed in the next chapter.

CHAPTER 4

ESTIMATING SOLIBILITY OF DRUGS IN SOLVENT MIXTURES

4.1 Introduction

The knowledge of solubility is important in the pharmaceutical area, because it allows process development scientists and engineers to choose optimum solvents for drug manufacturing processes. The prediction of ternary liquid-liquid equilibria by means of the NRTL-equation and using the parameters obtained by the correlation of the limiting binary systems is examined.

In principle, equations-of state-state methods can be applied for liquid phase equilibria as well as gas phase, much has been published about development of *PVT* equations of state for both phases (Frank, 1999, Kolar 2002). However, a widely used alternative for the liquid phase is application of excess properties (Tung et al., 2007). The Gibbs energy, G^E, is of paramount importance, because its canonical variables temperature, T, pressure, P, and composition, x_1, are always needed for process design and calculations.

Modern theoretical development of equations that use the excess properties are often based on the concept of *local composition* theory, which is presumed to account for the short-range order and non-random molecular orientations that result from differences in molecular size and intermolecular forces. Introduced with publication of model of G^E behavior known as the Wilson equation (1964), it prompted the development of alternative local composition models, most notably the Non-random Two-liquid (NRTL)

equation of Renon and Prausnitz (1968) and the Universal Quasi-chemical (UNIQUAC) equation of Abrams and Prausnitz of Abrams and Prausnitz (1975). There are two general assumptions in the application of the local composition models. The first is that binary phase equilibrium data is sufficient to obtain the model parameters. The second is that the model parameters are independent of temperature since the models are purported to have built in temperature dependence. These two principles are also applicable to equations of state.

The advantages and disadvantages of these equation for predicting solubility of pharmaceutical compounds in liquid solvents is discussed in the next three sections.

4.1.1 The Wilson's equation

Using on molecular considerations, Wilson (1964) presented the following expression for the excess Gibbs energy of binary solutions:

$$\frac{g^E}{RT} = x_1 \ln(x_1 + \Lambda_{12} x_2) - x_2 \ln(x_2 + \Lambda_{21} x_1) \tag{4.1}$$

and the activity coefficient derived from this equation becomes

$$\ln\gamma_1 = -\ln(x_1 + \Lambda_{12} x_2) + x_2 \left(\frac{\Lambda_{12}}{x_1 + \Lambda_{12} x_2} - \frac{\Lambda_{21}}{x_2 + \Lambda_{21} x_1} \right) \tag{4.2}$$

and

$$\ln\gamma_2 = -\ln(x_2 + \Lambda_{21} x_1) - x_1 \left(\frac{\Lambda_{12}}{x_1 + \Lambda_{12} x_2} - \frac{\Lambda_{21}}{x_2 + \Lambda_{21} x_1} \right) \tag{4.3}$$

As can be seen, the Wilson equation has two adjustable parameters, Λ_{12} and Λ_{21}, which are related to the pure component molar volumes and to characteristic energy differences by:

$$\Lambda_{21} = \frac{\upsilon_2}{\upsilon_1} \exp\left(\frac{\lambda_{12} - \lambda_{11}}{RT} \right) \tag{4.4}$$

$$\Lambda_{21} = \frac{\upsilon_2}{\upsilon_1} \exp\left(\frac{\lambda_{12} - \lambda_{11}}{RT} \right) \tag{4.5}$$

Where υ_i is the molar liquid volume of pure component *i* and the $\lambda's$ represent energies of interactions. To good approximation, the differences in the characteristic energies are constant over a modest temperature range. Therefore, an advantage of the Wilson equation is that, not only does it give an expression for the activity coefficient, but also and estimate of the variation of the activity coefficient with temperature. This may provide some practical advantage in isobaric calculations where temperature varies with composition.

Wilson's equation is reported to provide a good representation of excess Gibb's energies for a variety of miscible mixtures, and is particularly useful for solutions of polar or associating components such as alcohols in non-polar solvents. However, Wilson's equation has two main disadvantages. First, it is not useful for systems where the logarithmic plot of the activity coefficient versus the mole faction solubility exhibit a maxima or minima. The second and most serious disadvantage of Wilson's equation is inherent in its inability to predict limited immiscibility. Therefore, Wilson's equation can

only be used for solutions that are completely miscible or else for those limited regions of partially miscible systems where only single liquid phase exists. The difficulty in using the Wilson's equation is that the molar volume of drug compounds at the solution temperature is often not available.

4.1.2 The UNIQUAC equation

Abrams and Prausnitz (1975) developed the UNIQUAC model from statistical mechanics, and showed that it gives good representation of both vapor-liquid and liquid-liquid equilibria for binary and multi-component mixtures containing a variety of nonelectrolytes such as hydrocarbons, ketones, esters, water, amines, alcohols, nitriles, etc. A basic assumption of the model is that it is possible to define local mole fractions representative of concentrations on a microscopic scale. The UNIQUAC equation treats the $\frac{g^E}{RT}$ as comprising of two additive parts, namely the *combinatorial* term, g^c, accounting for the dominant entropic contributions, such as composition, molecular size and shape differences; it requires only pure component data. The *residual* term, g^c, accounts for the molecular interactions that are responsible for mixing enthalpy:

$$\frac{g^E}{RT} = \frac{g^C}{RT} + \frac{g^R}{RT} \tag{4.6}$$

The g^C term contains pure-species parameters only, whereas the g^R term contains two *binary* parameters for each pair of molecules. The UNIQUAC equation is applicable to a wide variety of non-electrolyte liquid mixtures containing polar or non-polar fluids such as hydrocarbons , alcohols, nitriles, ketones, aldehydes, organic acids, etc and water, as well as partially miscible mixtures. The first main advantage of UNIQUAC is that it is relatively simple because it requires only two adjustable parameters, and a second advantage is that it has a wide range of applications. However, the requirement of knowledge of the *combinatorial* term for complex drug compounds, which is often not available, makes it not favorable for pharmaceutical application.

4.1.3 Non-random Two-liquid (NRTL) activity coefficient model

As previously reviewed, the basic idea of Wilson's derivation of the of Eq. (4.1) follows from the local composition concept. This concept prompted (Renon, 1968) to derive the NRTL equation; however, unlike Wilson's equation, the NRTL equation is applicable to partially miscible systems as well as completely miscible mixtures.

For a binary mixture the following equations are used:

$$\ln \gamma_1 = x_2^2 \left[\frac{\tau_{12} G_{12}}{\left(x_2 + x_1 G_{12}\right)^2} + \tau_{21} \left(\frac{G_{21}}{x_1 + x_2 G_{21}} \right)^2 \right] \tag{4-7}$$

Here,

$$G_{12} = \exp(-\alpha\tau_{12}), \text{ and } G_{21} = \exp(-\alpha\tau_{21}) \quad (4\text{-}8)$$

and

$$\tau_{12} = \frac{b_{12}}{RT} \text{ and } \tau_{21} = \frac{b_{21}}{RT} \quad (4.9)$$

where b_{12} and b_{21}, are interaction parameters specific to a particular pair of species, independent of temperature and composition, and the parameter α is a measure of the non-randomness of the mixture; when α is zero, the mixture is said to be completely random. The NRTL equation provides good representation of the binary vapor-liquid equilibrium and it readily generalized for multicomponent mixtures, with only the binary parameters. This equation is superior to the Wilson's equation in the sense that it can represent the liquid-liquid equilibrium (LLE). In addition, it is simpler in form than the UNIQUAC equation. The main disadvantage with the NRTL equation is that has three adjustable parameters for each pair of components. From both practical and theoretical standpoints, it is desirable to minimize the number of parameters needed to describe as wide variety of systems as possible. To reduce the number of adjustable parameters, from three to two in the NRTL equation and to overcome the fore cited disadvantage, Renon and Prausnitz (1968) recommended an α of 0.20 for partially miscible systems. For completely miscible systems, Renon and Prausnitz (1968), from reduction of experimental data, determined that α varies from 0.20 to 0.47.

Given the aforementioned limitation in using the Wislon's equation or the UNIQUAC for pharmaceutical application, this study focused on the use of NRTL equation for

estimation of solubility of drug compounds in solvent mixtures. The research focuses on the use of the interaction parameters from pure solute-solvent binary system, with the appropriate calculated ideal solubility, to predict solubility in solute-binary solvent mixtures.

4.2 Solubility of pharmaceutical compounds in binary solvent mixtures

Mixtures of non-aqueous solvents are important from a pharmaceutical viewpoint, since the mixtures could be used as synthesis medium or re-crystallization solvents for purification of drugs. Accurate mathematical models are needed for rational selection of solvent mixtures as function of solubility for the design of efficient drug manufacturing processes. At present, the log-linear model of Yalkowsky (1987) is the most commonly model for predicting solubility of drugs in water and organic solvent mixtures. The model uses aqueous solubility of the drug for the prediction. On the other hand, the Jouyban-Acree model can be used to estimate solubility in solvent mixtures. However, the model requires several experimental data points, which can be time and resource consuming.

Recently, Tung *et al.,* (2007) attempted to use the NRTL model to predict the solubility of lovastatin and simvastatin in solvent mixtures; agreement between their predicted and experimental solubility data was poor. We believe this to be likely due to the fact that they neglected the differential molar heat capacity when they calculated the ideal solubility. Here, the solubility data was modeled to the NRTL activity coefficient model, this time, properly accounting for the differential heat capacity. The interaction parameter obtained was then used to predict solubility of the model compounds in solvent mixtures. The results agreed favorably with the experimental data.

4.3 Methods

4.3.1 Predicting solubility in solvent mixtures using the multi-component NRTL equation with binary interaction parameters.

The solubility data obtained from Chapter 2 were non-linearly regressed with Eq. (4.7); a value of α of 0.40 was used throughout the regression procedure. For lovastatin and simvastatin, the calculated ideal solubility from assumption II (Chapter 3; Section 3.2) was used, whereas for artemisinin the ideal solubility was calculated using the differential molar heat capacity for the present methodology, previously explained in Chapter 3 (Section 3.6).

4.4 Results and discussion.

The best fit values of the interaction parameters are displayed in Tables 4.1, 4.2, 4.3, and 4.4, and also, are graphically displayed in Figures 4.1, 4.2, 4.3, 4.4, and 4.5. The values agreed well with the experimental values, as reflected by the average relative deviation (ARD) defined as:

$$ARD = \frac{1}{N}\sum\left[\left|\frac{x_1 - x_{1\,(\mathrm{Calc})}}{x_1}\right|\right] \qquad (4\text{-}5)$$

where N is the number of data points obtained in each set which equal the number of temperatures used, and $x_{1\,(\mathrm{Calc})}$ is the calculated value from the NRTL equation.

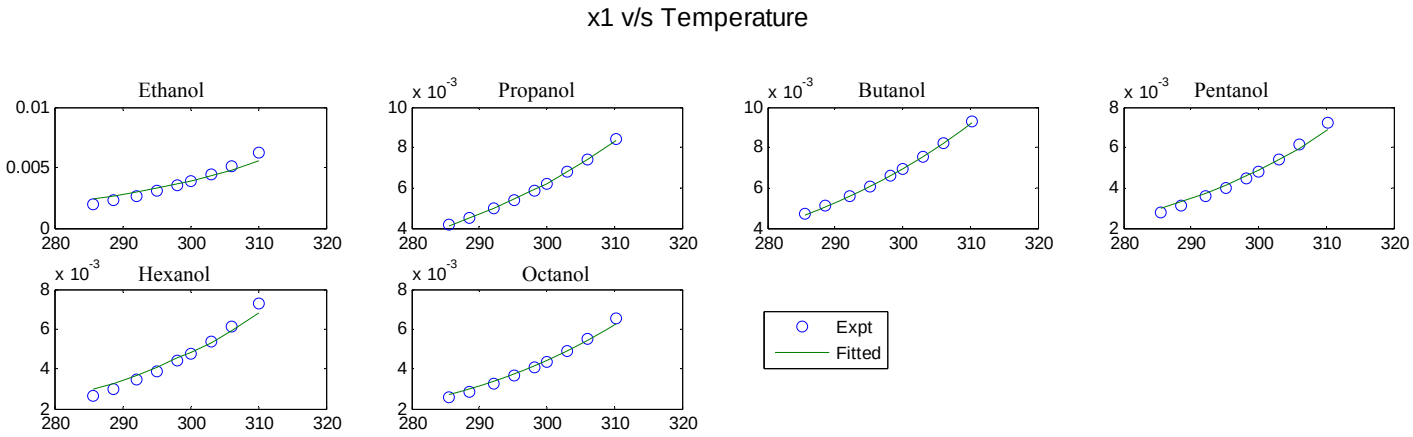


Figure 4.1: NRTL correlation of solubility of lovastatin in alcohols

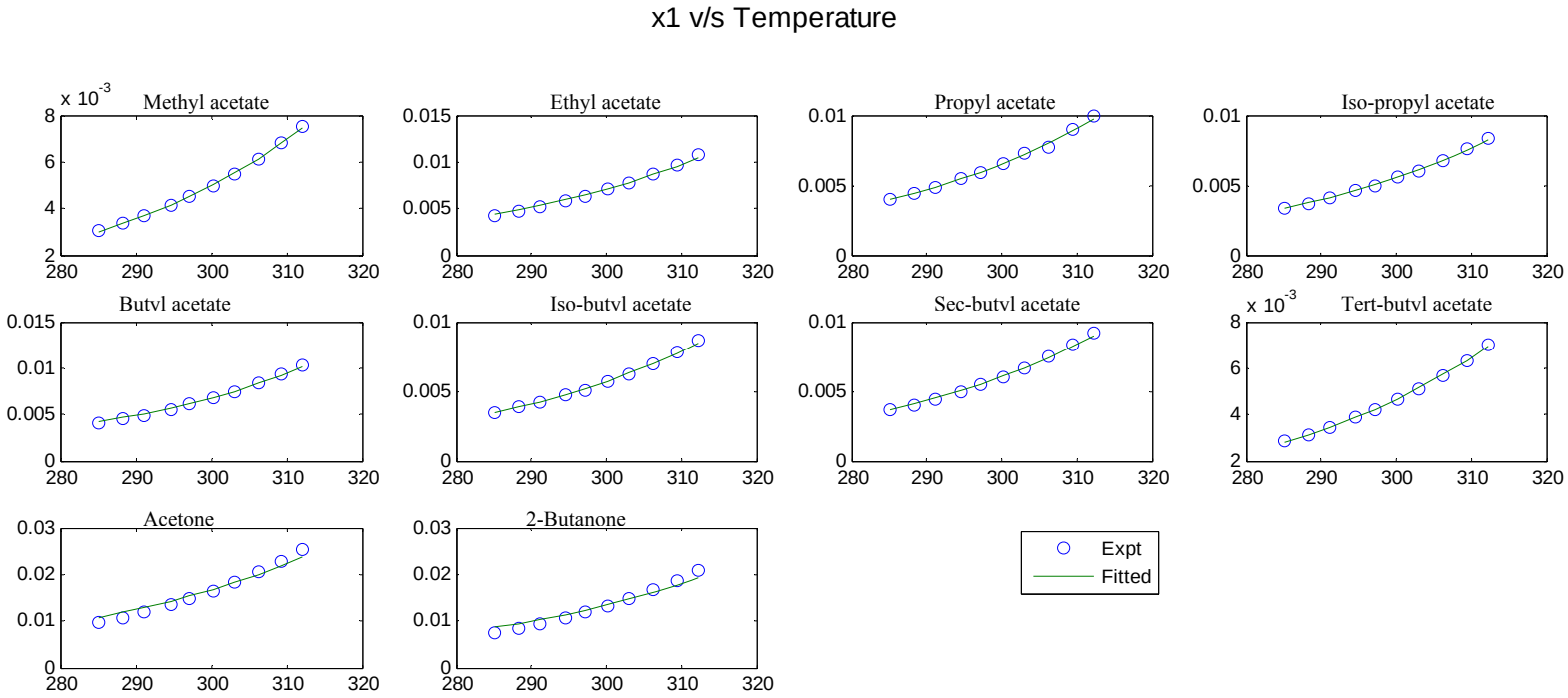


Figure 4.2: NRTL correlation of solubility of lovastatin in acetates and ketones

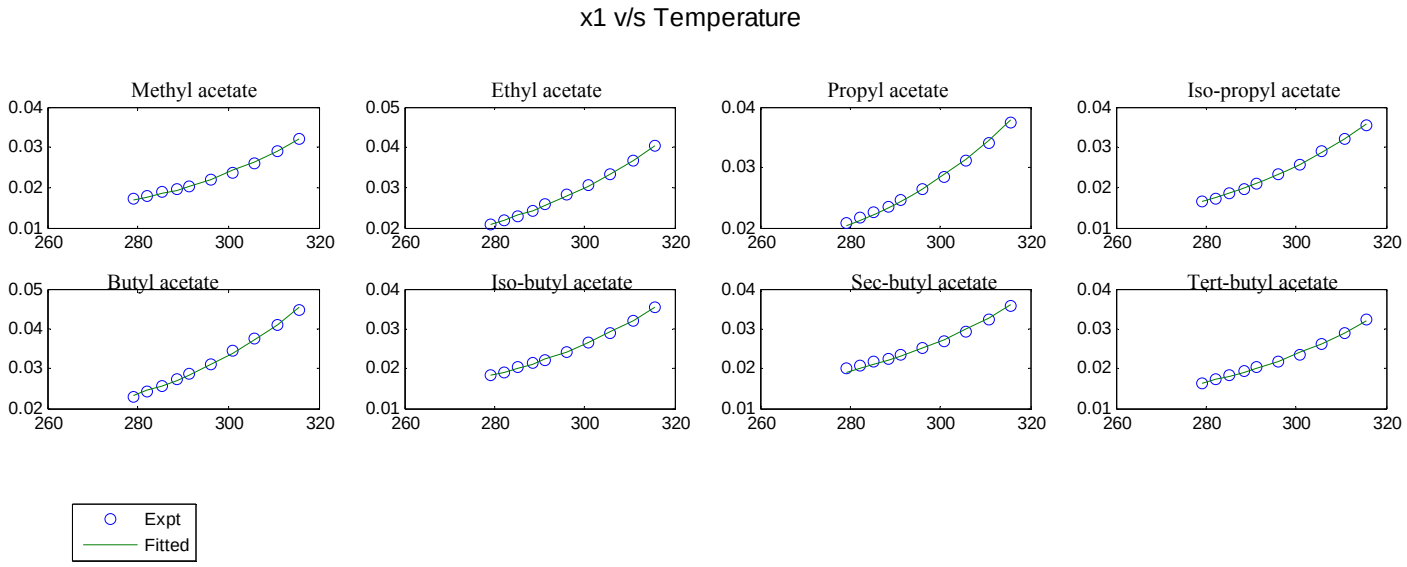


Figure 4.3: NRTL correlation of solubility of simvastatin in acetates and ketones.

x1 v/s Temperature

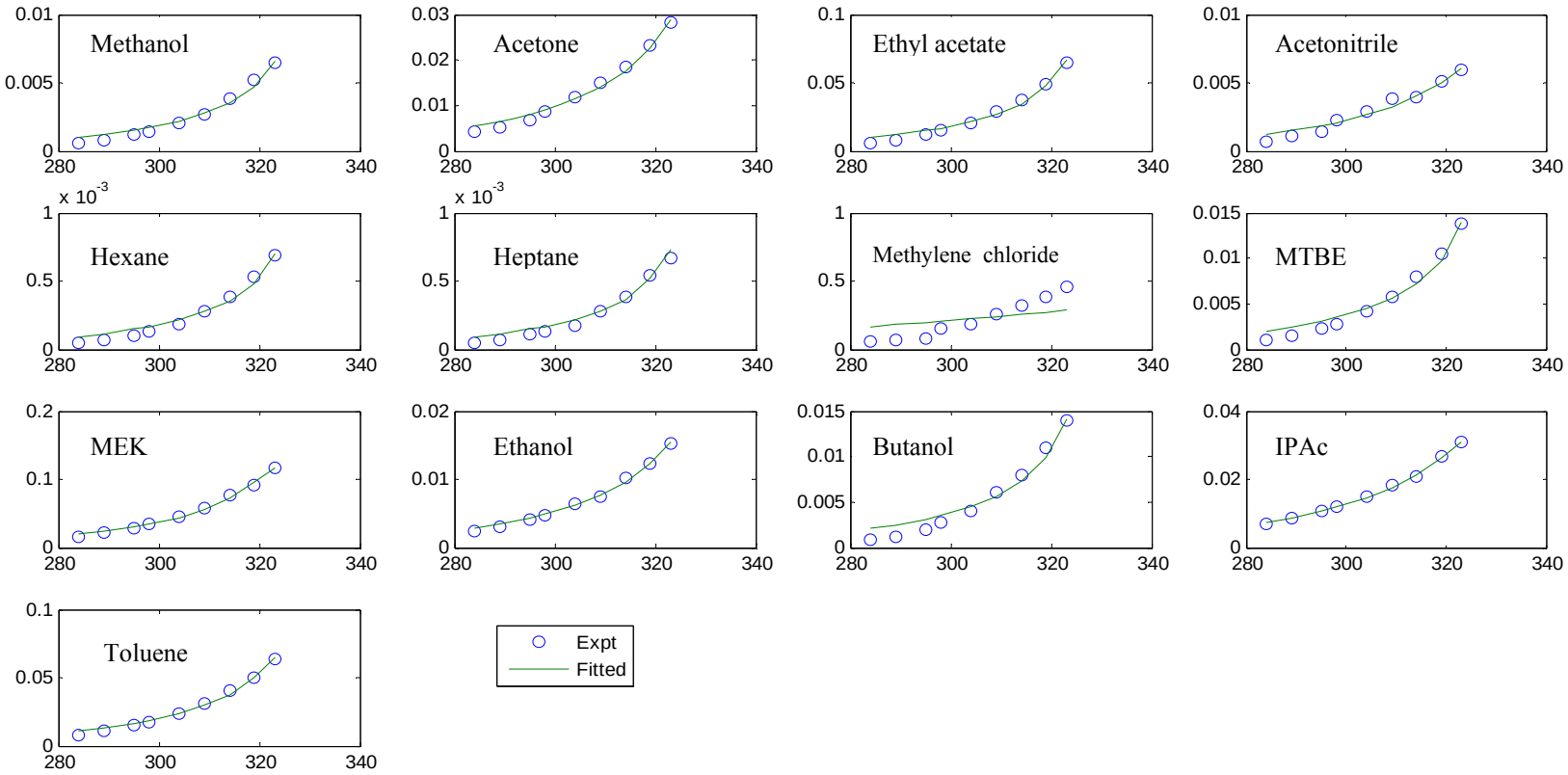


Table 4.1

Values of the binary interaction parameters for the NRTL model and average relative deviation (ARD) from the measured equilibrium mole fraction of lovastatin in alcohols.

	b_{12}	b_{21}	*%ARE*
Ethanol	-597	4424	1.2
Propanol	-622	3140	0.91
Butanol	-709	3456	1.98
Pentanol	-862	4093	2.51
Hexanol	-806	3938	1.45
Octanol	-731	3679	0.48

Table 4.2

Values of the binary interaction parameters for the NRTL model and average relative deviation (ARD) from the measured equilibrium mole fraction of lovastatin in acetates and ketones.

	b_{12}	b_{21}	*%ARE*
Methyl acetate	315	2905	0.55
Ethyl acetate	-509	2897	0.76
Propyl acetate	-583	3190	0.98
Isopropyl acetate	-715	3790	0.50
Butyl acetate	-521	3013	0.45
Isobutyl acetate	-635	3607	0.88
sec butyl acetate	-563	3358	0.67
tert Butyl acetate	-111	2476	0.68
Acetone	-37	171	1.93
2-Butanone	-177	866	2.51

Table 4.3

Values of the binary interaction parameters for the NRTL model and average relative deviation (ARD) from the measured equilibrium mole fraction of simvstatin.

	b_{12}	b_{21}	*100 ARD*
Methyl acetate	9554	2245	0.51
Ethyl acetate	5256	1243	0.79
Propyl acetate	9652	1804	0.88
iso-Propyl acetate	2064	2278	0.67
Butyl acetate	4359	1001	0.82
iso-Butyl acetate	6170	1621	0.91
sec-Butyl acetate	9626	1934	0.83
tert-Butyl acetate	6395	1896	0.75
Ethanol	358	4062	0.45
1-Propanol	-215	4031	0.98
1-Butanol	2725	1917	0.39
2-Butanol	960	2059	0.78
1-Pentanol	-47	3469	0.76
1-Hexanol	-851	4207	0.86
1-Octanol	6364	1340	0.93

Table 4.4

Values of the binary interaction parameters for the NRTL model and average relative deviation (ARD) from the measured equilibrium mole fraction of artemisnin.

	b_{12}	b_{21}	*100 ARD*
Methanol	-2585	15287	0.92
Acetone	-1950	9812	2.16
Ethyl acetate	-2114	8973	0.87
Acetonitrile	-2084	13387	4.28
Hexane	-3813	24785	5.21
Heptane	-3886	25142	2.56
MTBE	-2302	12876	0.68
Ethanol	-2046	11438	1.85
Butanol	-2304	12876	0.45
Isopropyl acetate	-1607	8309	1.02
Toluene	-2074	8566	0.87

4.4.1 Estimation of the molar excess enthalpies in the binary systems

Waals (1985) has comprehensively reviewed the NRTL activity coefficient model. Based on this review, the excess enthalpy, h^E, which indicates the temperature and composition dependence of the Gibbs energy, and hence also the activity coefficients can be determined using the Gibbs-Helmholtz equation.

$$-\frac{h^E}{RT^2}=\left(\frac{\partial\left(G^E/RT\right)}{\partial T}\right)_{P,x} \qquad 4.10$$

Based on the binary parameters and the temperature independent non-randomness parameter, α, the Gibbs-Helmholtz relation gives the following expression for the excess enthalpy.

$$h^E = RTx_1x_2\left[\frac{\tau_{21}G_{21}}{x_1+x_2G_{21}}+\frac{\tau_{12}G_{12}}{x_2+x_1G_{12}}-\alpha\left(\frac{x_1\tau^2{}_{21}G_{21}}{(x_1+x_2G_{21})^2}+\frac{x_2\tau^2{}_{12}G_{12}}{(x_2+x_1G_{12})^2}\right)\right] \qquad 4.11$$

Eq. (4.11) was used to obtain temperature-dependent values of h^E as function of the smole fraction lovastatin, simvastatin, and artermisinin at 298 K in each of the solvents studied herein.

In Figure 4.5, and 4.6, the calculated values of h^E are plotted versus mole ratio of lovastatin in alcohols, and (acetates and ketones), respectively

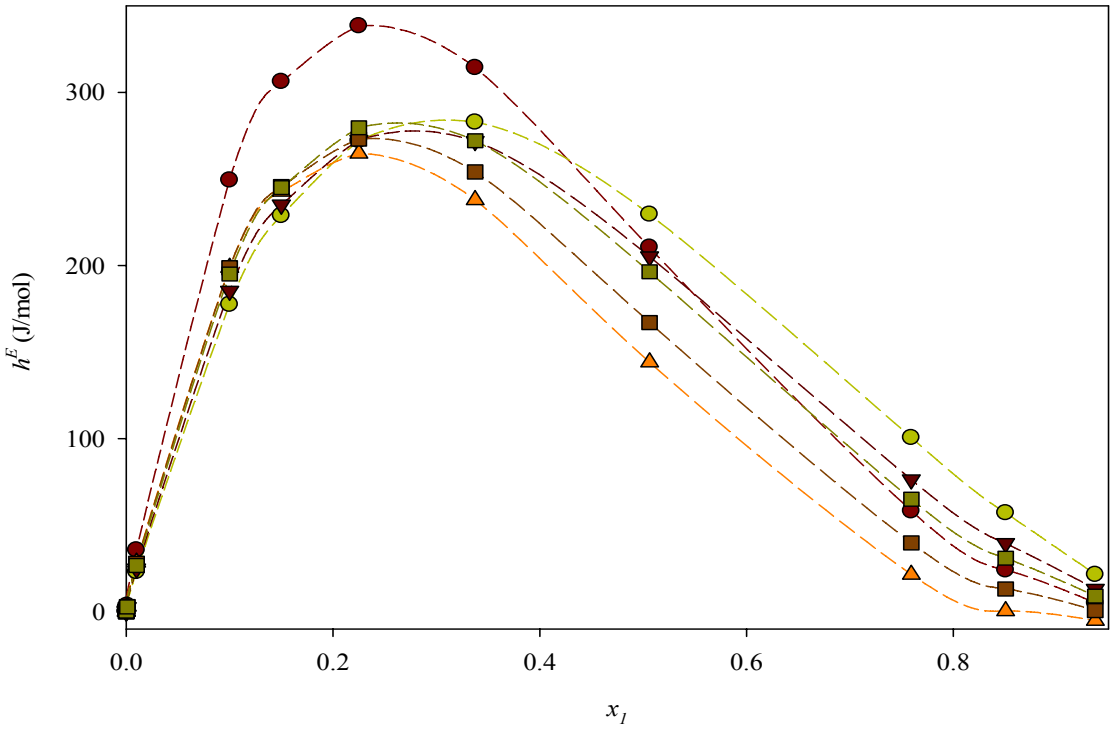


Figure 4.5. Excess molar enthalpies of the binary system of lovastatin (1) and:

Ethanol Butanol Hexanol
Propanol Pentanol Octanol

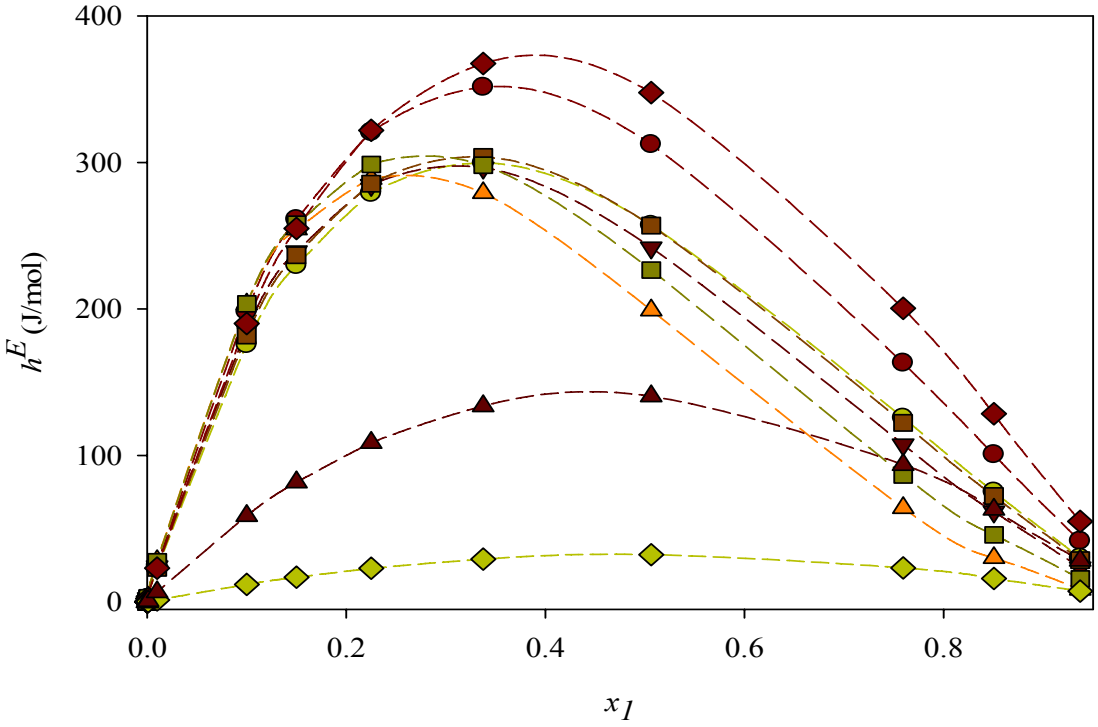


Figure 4.6. Excess molar enthalpies of the binary system of lovastatin (1) and:

Methyl acetate
Ethyl acetate
Propyl acetate
Iso propyl acetate
Butyl acetate
Iso butyl acetate
Sec butyl acetate
Tert butyl acetate
Acetone
Acetone

In Figure 4.7, and 4.8, the calculated values of h^E are plotted versus mole ratio of simvastatin in alcohols, and acetates, respectively.

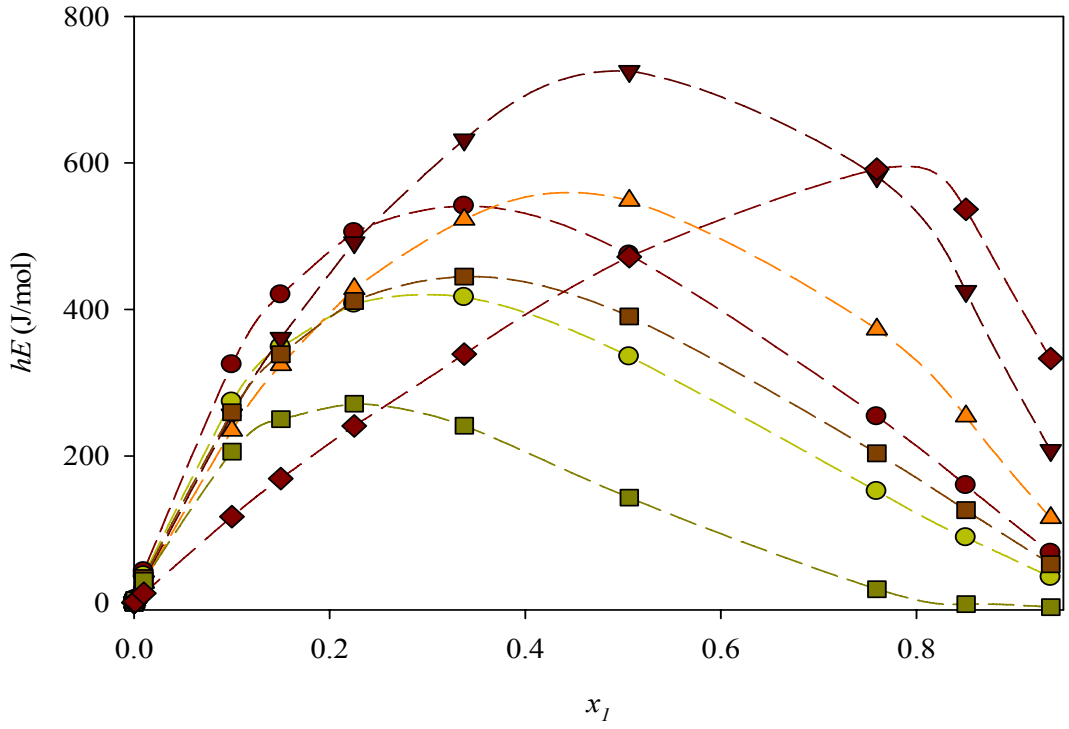


Figure 4.7. Excess molar enthalpies of the binary system of simvastatin (1) and:

Ethanol
1-Propanol
1-Butanol
2-Butanol
1-Pentanol
1-Hexanol
1-Octanol

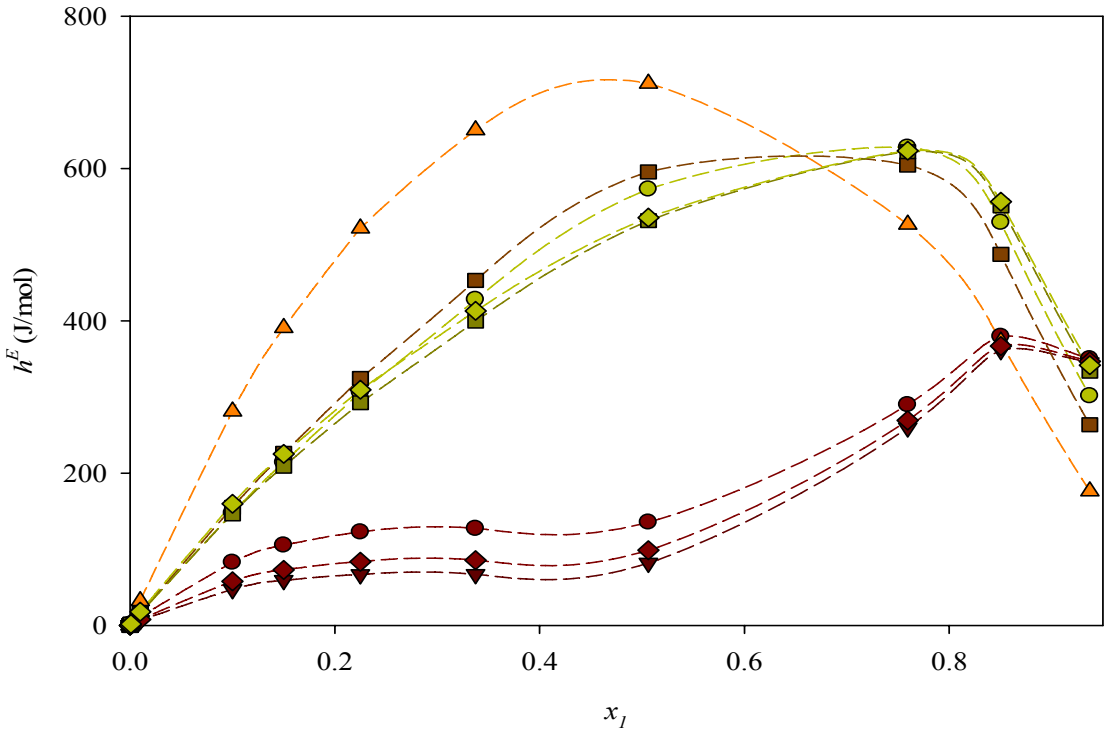


Figure 4.8. Excess molar enthalpies of the binary system of simvastatin (1) and:

In Figure 4.9 calculated values of h^E are plotted versus mole ratio of artemisinin in alcohols, and other solvents.

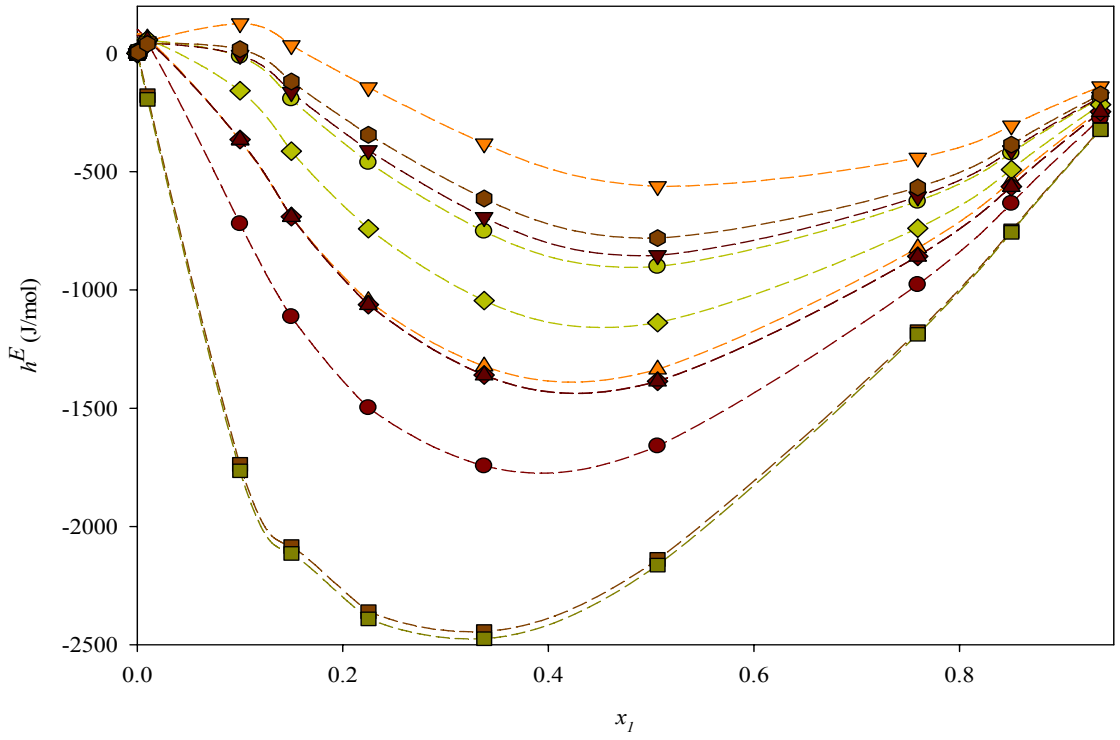


Figure 4.9. Excess molar enthalpies of the binary system of artemisinin (1) and:

Methanol
Acetone
Ethyl acetate
Acetonitrile
Hexane
Heptane
MTBE
Ethanol
Butanol
Iso propyl acetate
Toluene

4.4.2 Estimation of the molar excess entropies in the binary systems

As previously reviewed Waals (1985) has comprehensively reviewed the NRTL activity coefficient model. Based on this review, the excess entropy, s^E, can be determined by:

$$-s^E = \left(\frac{\partial \left(g^E \right)}{\partial T} \right)_{P,x} \qquad 4.12$$

Based on the binary parameters and the temperature independent non-randomness parameter, α, the Gibbs-Helmholtz relation gives the following expression for the excess entropy, s^E:

$$s^E = \frac{-Rx_1 x_2}{\alpha} \left[\frac{x_1 G_{21} \tau^2{}_{21}}{\left(x_1 + x_2 G_{21} \right)^2} + \frac{x_2 G_{12} \tau^2{}_{12}}{\left(x_2 + x_1 G_{12} \right)^2} \right] \qquad 4.13$$

Eq. (4.13) was used to obtain temperature-dependent values of s^E as function of the mole fraction lovastatin, simvastatin, and artermisinin at 298 K in each of the solvents studied herein.

In Figure 4.10, and 4.11, the calculated values of s^E are plotted versus mole ratio of lovastatin in alcohols, and (acetates and ketones), respectively

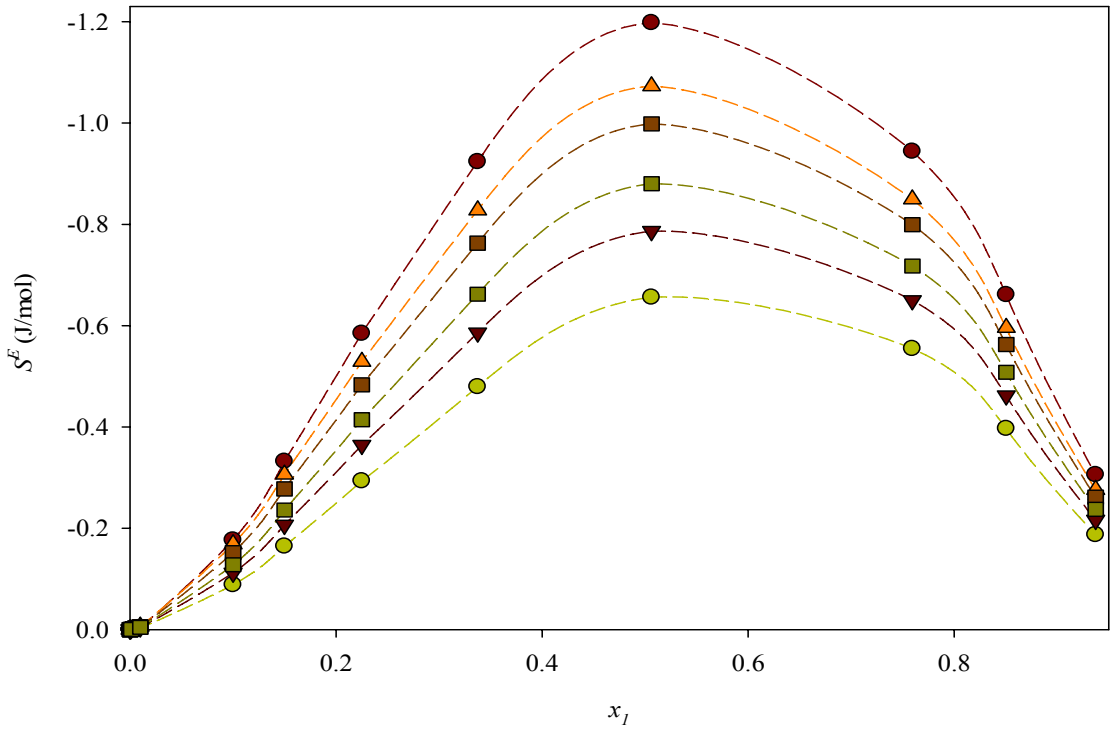


Figure 4.10. Excess molar entropies of the binary system of lovastatin (1) and

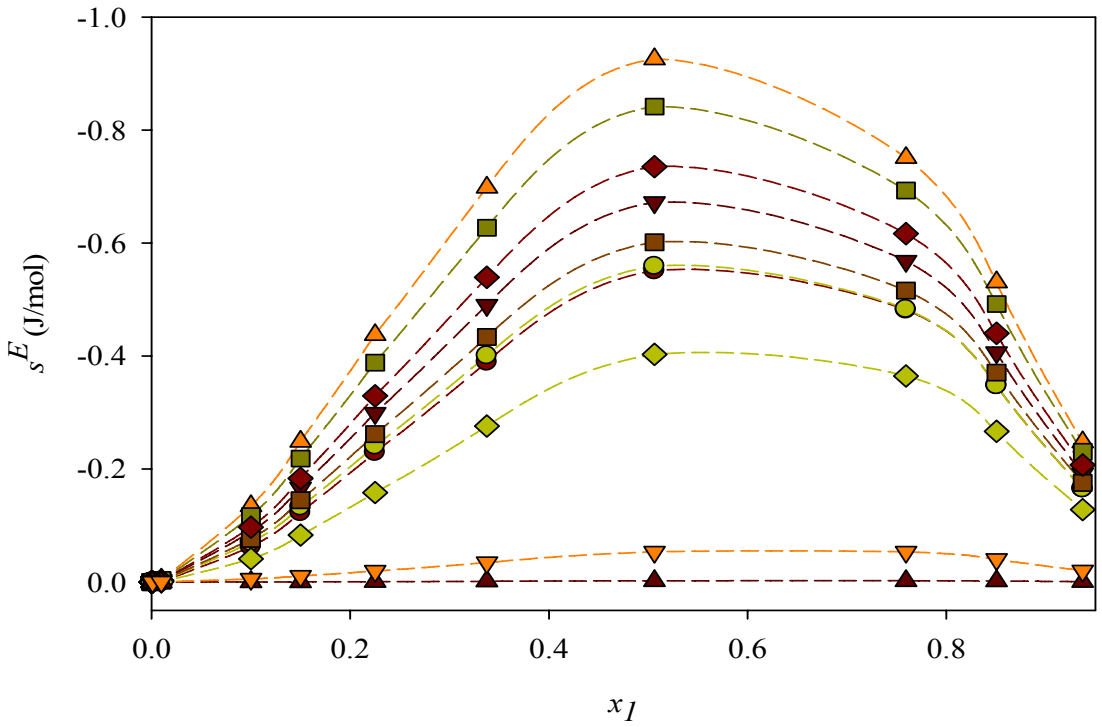


Figure 4.11. Excess molar entropies of the binary system of lovastatin (1) and:

Methyl acetate
Ethyl acetate
Propyl acetate
Iso propyl acetate
Butyl acetate
Iso butyl acetate
Sec butyl acetate
Tert butyl acetate
Acetone
Acetone

In Figure 4.12, and 4.13, the calculated values of s^E are plotted versus mole ratio of simvastatin in alcohols, and acetates, respectively

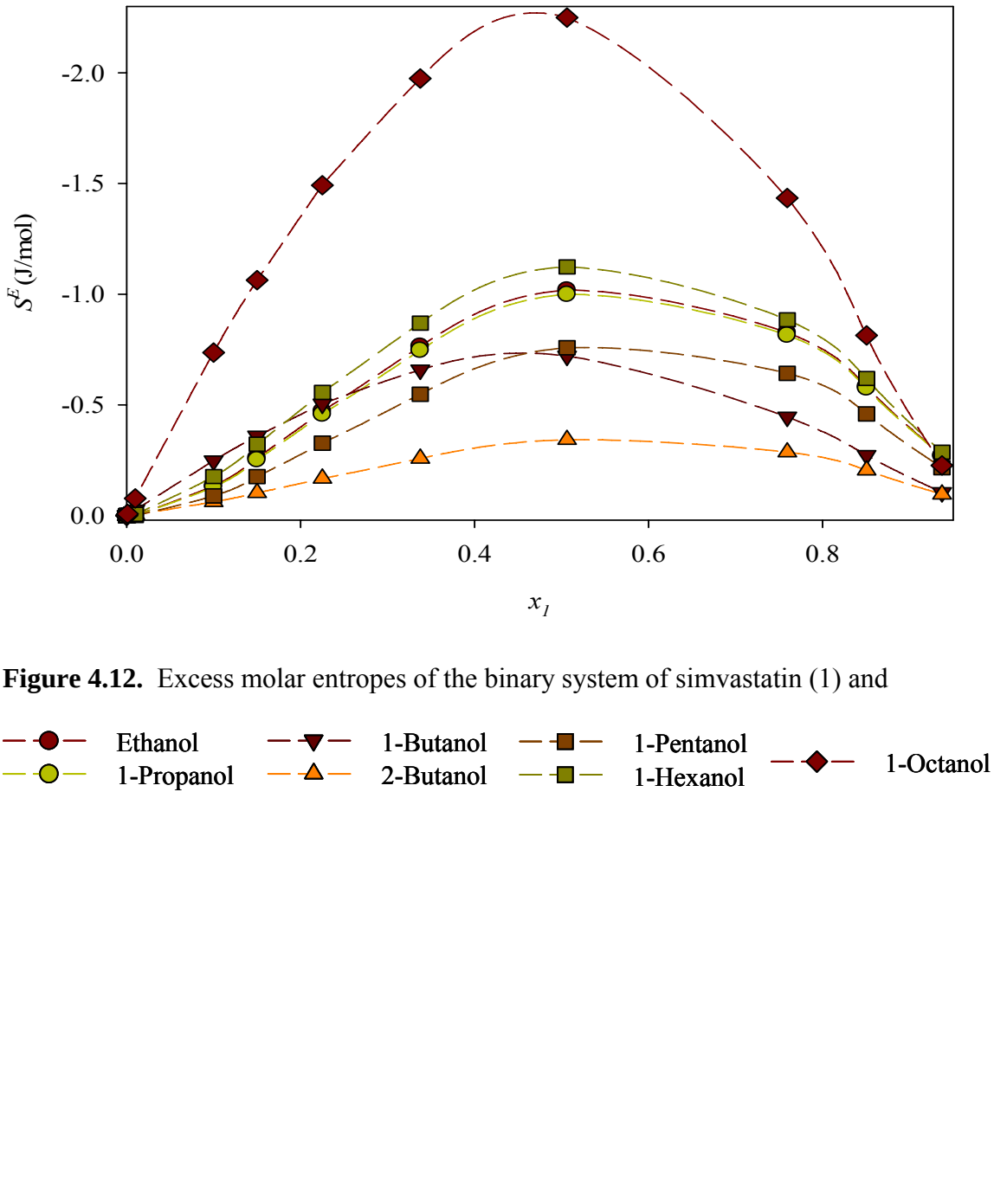


Figure 4.12. Excess molar entropes of the binary system of simvastatin (1) and

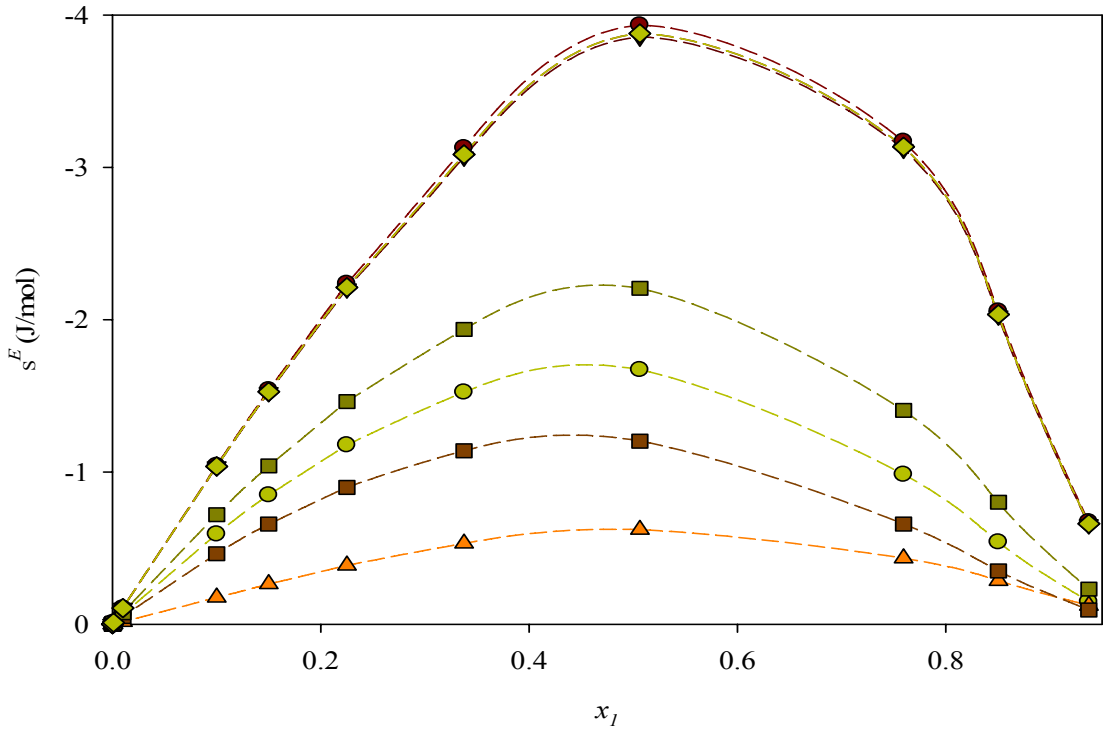


Figure 4.13. Excess molar entropies of the binary system of simvastatin (1) and:

Methyl acetate
Ethyl acetate
Iso propyl acetate
Iso propyl acetate
Butyl acetate
Iso butyl acetate
Sec butyl acetate
Tert butyl acetate

In Figure 4.14 calculated values of s^E are plotted versus mole ratio of artemisinin in alcohols, and other solvents.

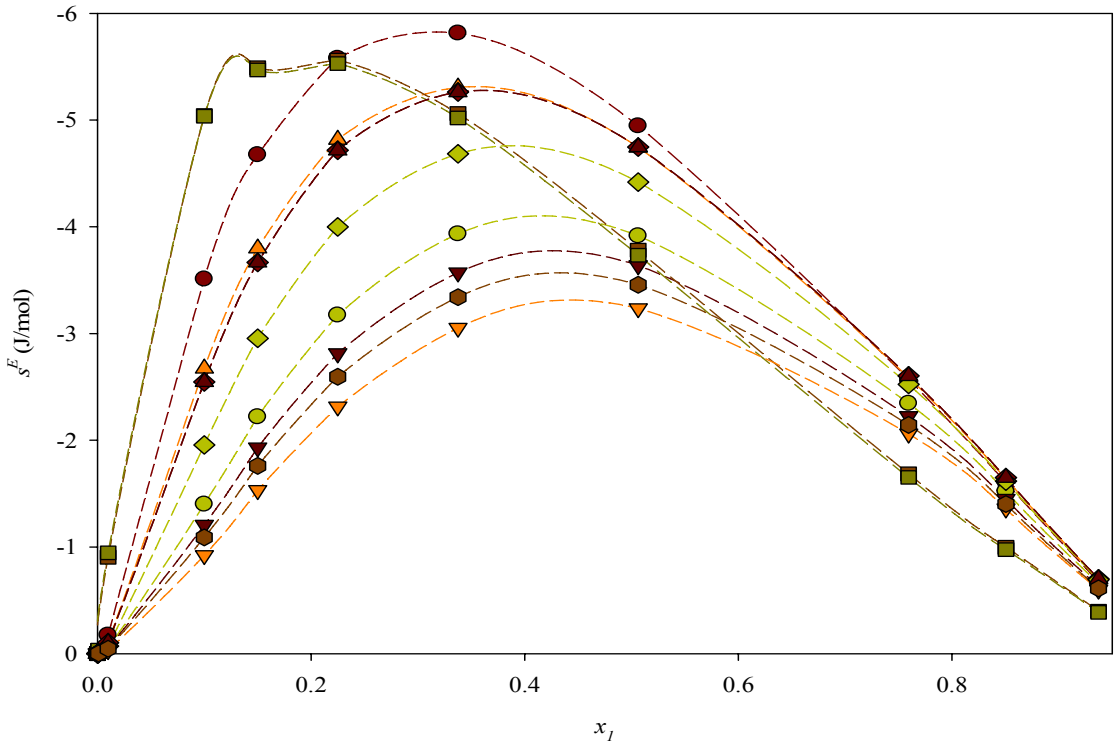


Figure 4.14. Excess molar enthalpies of the binary system of artemisinin (1) and:

Methanol
Acetone
Ethyl acetate
Acetonitrile
Hexane
Heptane
MTBE
Ethanol
Butanol
Iso propyl acetate
Toluene

4.4.3 Discussion of the excess enthalpy and entropy results.

Putting the results summarized in Figures 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11, 4.12, 4.13, and 4.14 into the context of the engineering "modeling" coordinates (Shukla et al., 1988) would suggest at 298 K:

1. The dissolution of lovastatin and simvastatin in all the solvents studied is enthalpy dominated, because $h^E > 0$, $s^E < 0$, and $g^E > 0$.

2. The dissolution of artemisinin in all the solvents studied is also enthalpy dominated because $h^E < 0$, $s^E < 0$, and $g^E < 0$.

4.5 Non-random Two-liquid activity coefficient for multi-component mixtures

The activity coefficient for any given component *i*, the NRTL equation is given by:

$$\frac{g^E}{RT} = \sum_{i=1}^{m} x_i \frac{\sum_{j=1}^{m} \tau_{ji} G_{ji} x_j}{\sum_{l=1}^{m} G_{li} x_l} \tag{4.14}$$

where

$$\tau_{ji} = \frac{g_{ji} - g_{ii}}{RT} \tag{4.15}$$

and,

$$G_{ji} = \exp(-\alpha_{ji}\tau_{ji}) \qquad (\alpha_{ji} = \alpha_{ij}) \qquad (\alpha_{jj} = \alpha_{ii}) = 0 \tag{4.16}$$

Thus, the activity coefficient for any component *i* is given by:

$$\ln \gamma_i = \frac{\sum_{j=1}^{m} \tau_{ji} G_{ji} x_j}{\sum_{l=1}^{m} G_{li} x_l} + \sum_{j=1}^{m} \frac{x_j G_{ij}}{\sum_{l=1}^{m} G_{lj} x_l} \left[\tau_{ij} - \frac{\sum_{r=1}^{m} x_r \tau_{rj} G_{rj}}{\sum_{l=1}^{m} G_{lj} x_l} \right] \tag{4-17}$$

As is seen, all the parameters in Eq. (4.17) can be obtained from binary system. Renon (1968) used only binary data in Eq. (4.17) to successfully predict ternary vapor-liquid data. However, it generally fails when used to estimate the solubility of pharmaceutical compounds solids in ternary solvent mixtures. Hypothesizing that one of the main reason why it fails to predict solubility in the solvent mixtures is because the correct ideal

solubility is generally not used, we set out to use the thermodynamically consistent ideal solubility to non-linearly regress our solubility data, and then use the interaction parameters to predict solubility of lovastatin, simvastatin, and artemisinin in ethyl acetate, acetone, ethanol binary solvent mixtures.

4.6 Methods

4.6.1 NRTL modeling of the binary experimental data

The interaction parameters obtained from the previous section was used in equation Eq. (4.14) to calculate the activity coefficient of lovastatin, simvastatin, and artemisinin in binary solvent mixtures of: (ethyl acetate / acetone), or (acetone/ethanol), or (ethyl acetate/ethanol).

For the solute-solvent interactions, an α value of 0.40 was used, and the binary solvent-solvent interaction parameters, τ_{ij}, were taken from the literature (Demirel, 1990), with α set to 0.30 for all the solvents, 0.40 for each of the solutes. The calculated activity coefficient was combined with calculated ideal solubility to estimate solubility of the model compound. The data was compared to the experimentally measured solubility data previously reported in Chapter 2. The MARLAB code used for the calculation is provided in the Appendix.

4.7 Results and discussions.

The calculated solubility excellently agreed with the experimentally measured solubility, as reflected in Figures 4.5, 4.6, and 4.7. This clearly demonstrates that NRTL interaction parameters from binary system can be used to predict solubility in ternary solute-binary solvent system provided the correct ideal solubility is used.

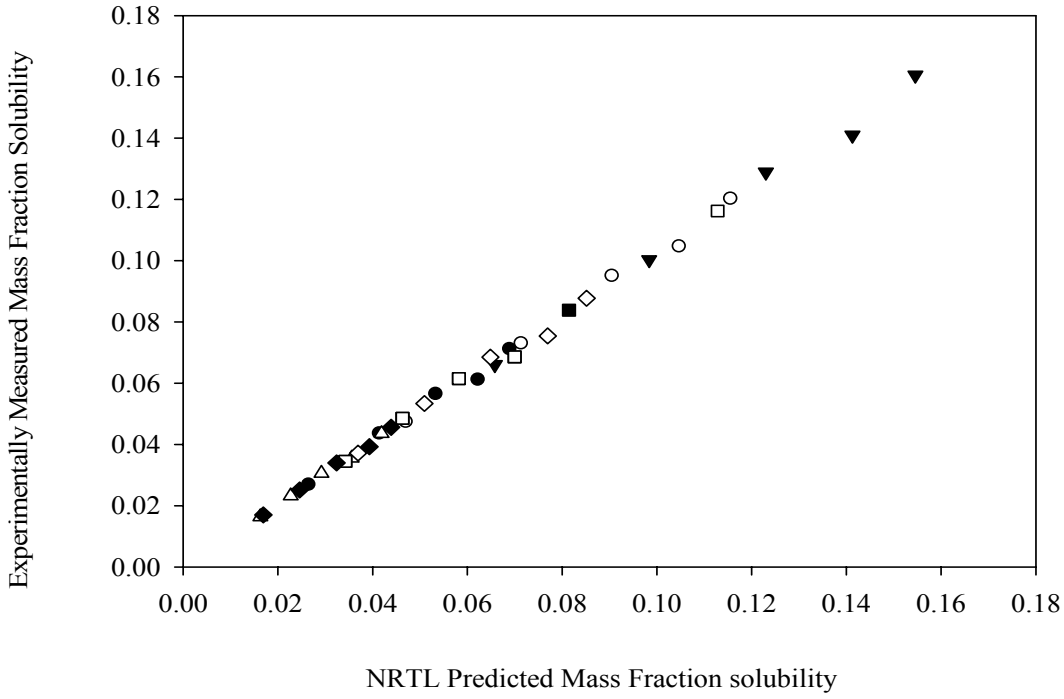


Figure 4.5: Plot of predicted solubility of lovastatin in acetone, ethyl acetate, and acetone solvent mixtures at different temperatures.

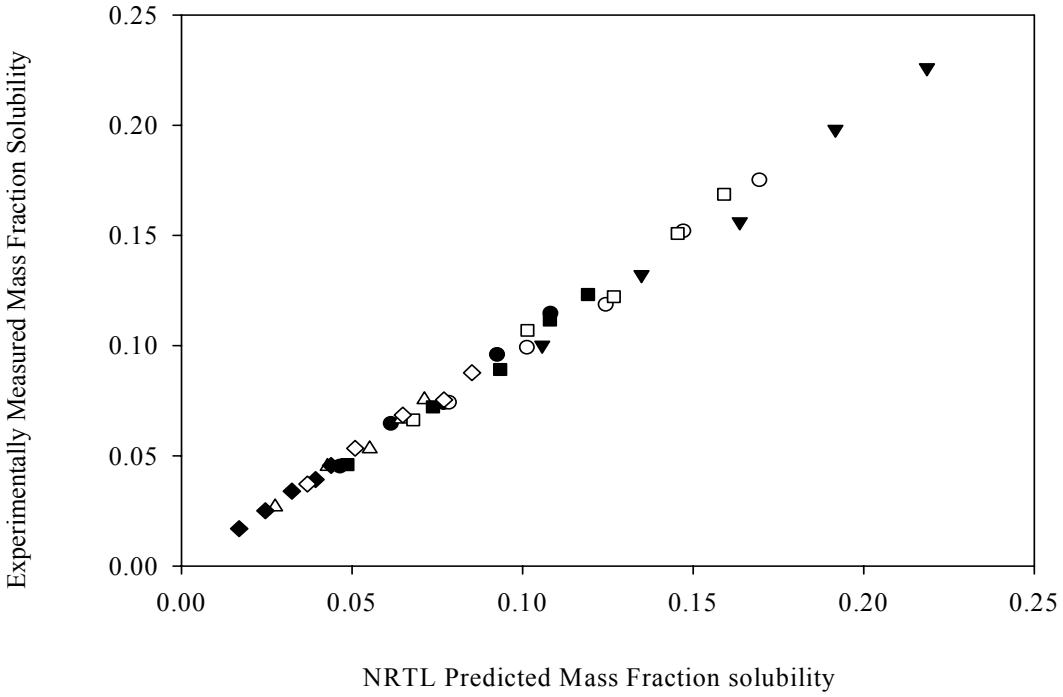


Figure 4.6: Plot of predicted solubility of simvastatin in acetone, ethyl acetate, and acetone solvent mixtures at different temperatures.

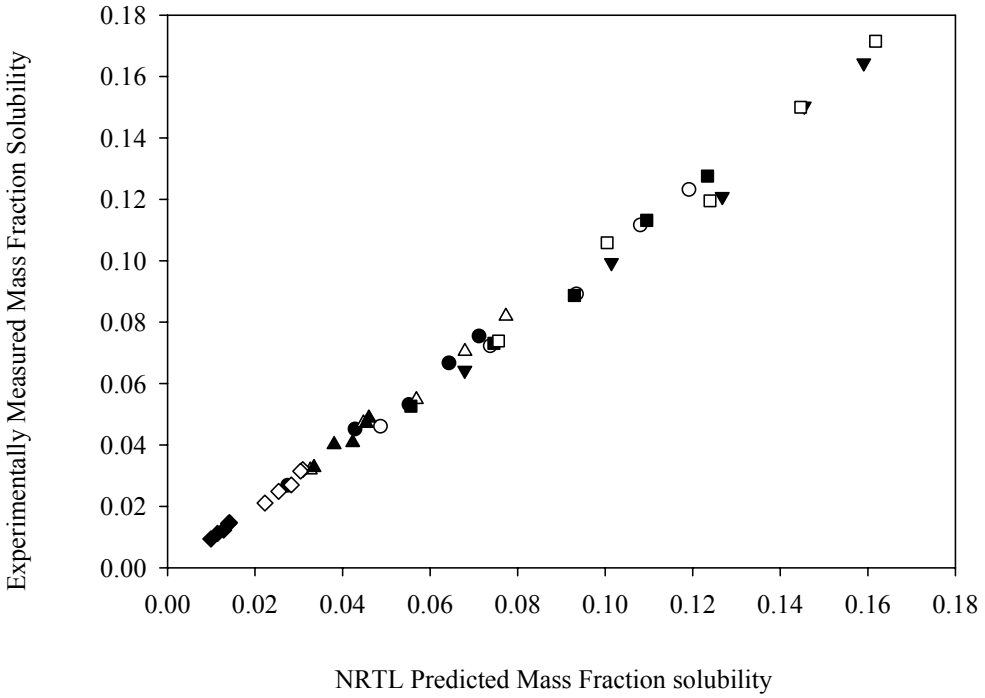


Figure 4.7: Plot of predicted solubility of artemisinin in acetone, ethyl acetate, and acetone solvent mixtures at different temperatures.

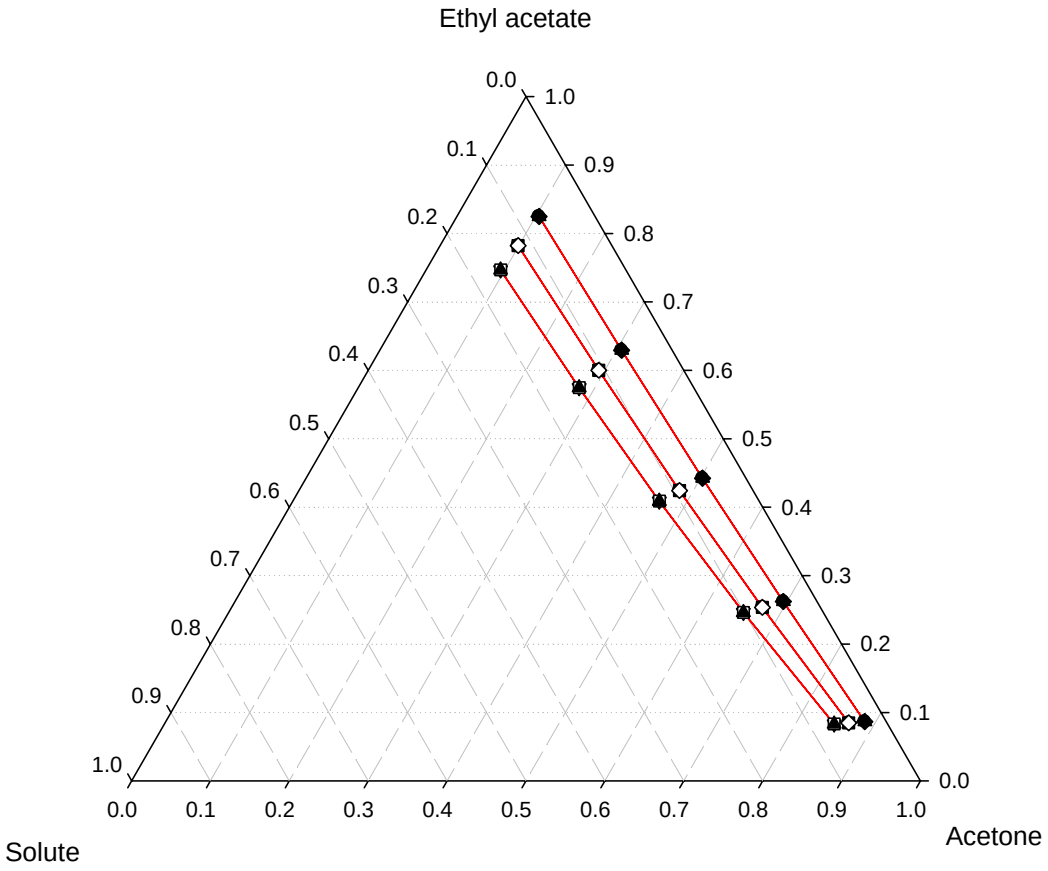


Figure 4.8: Predicted solubility of artemisinin, lovastatin, or simvastatin in ethyl acetate and acetone solvent mixture.

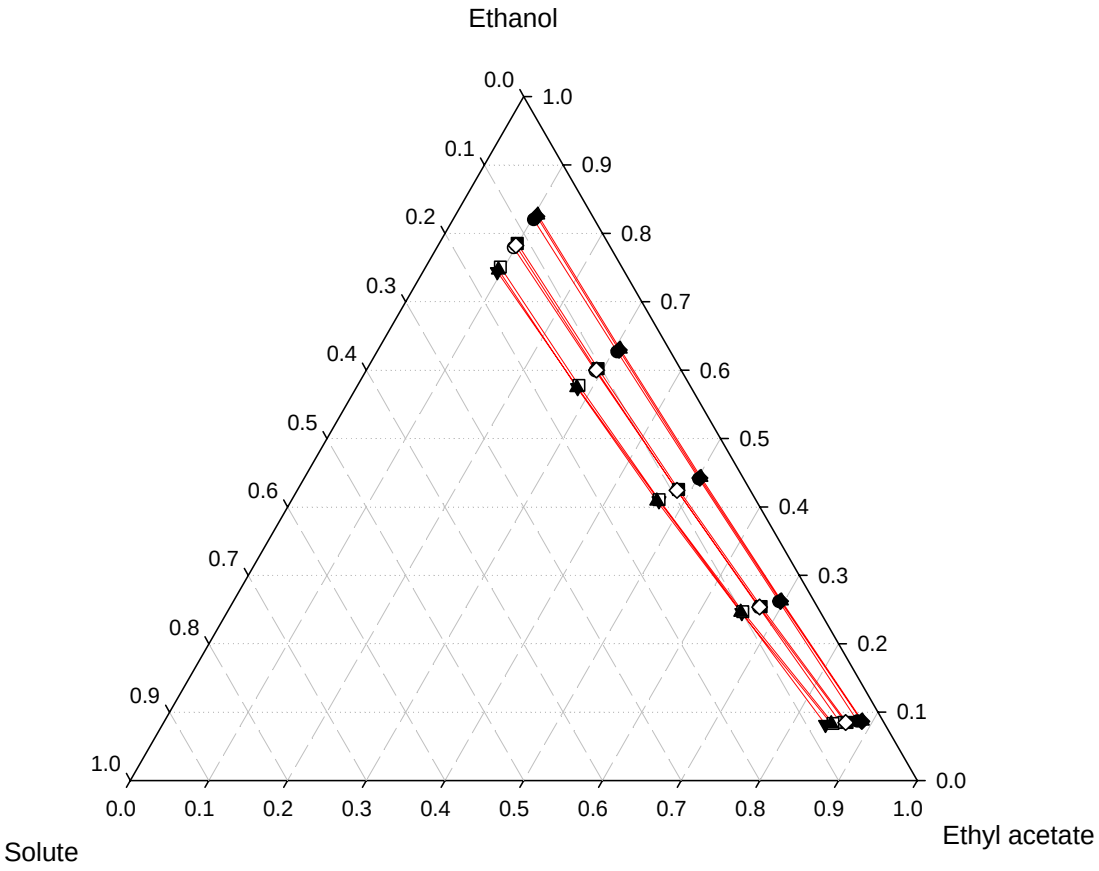


Figure 4.9: Predicted solubility of artemisinin, lovastatin, or simvastatin in ethyl acetate and acetone solvent mixture.

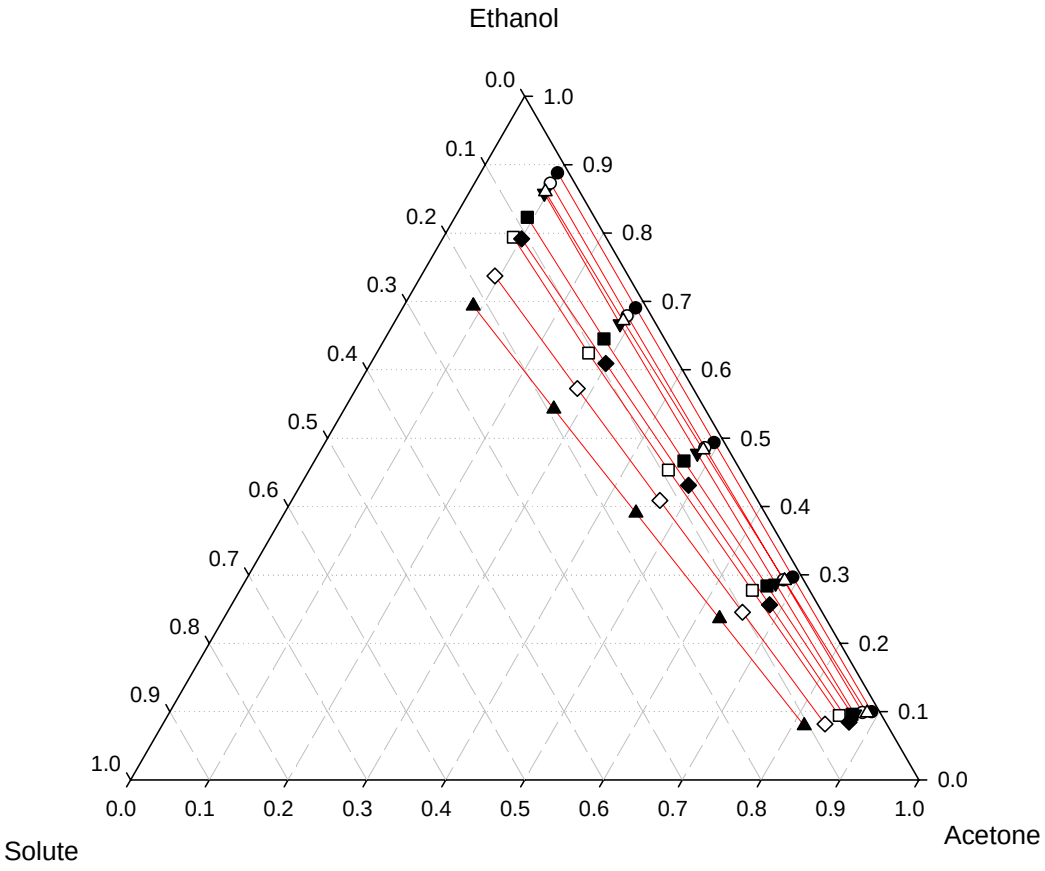


Figure 4.10: Predicted mass fraction solubility of artemisinin, lovastatin, or simvastatin in ethanol and acetone mixture.

4.8 Conclusion

NRTL activity coefficient interaction parameters have been obtained for binary systems of: lovastatin, simvastatin, or atermisinin, and: 1-butanol, 1-hexanol, 1-pentanol, 1-hexanol, and 1-octanol, methyl acetate, ethyl acetate, propyl acetate, isopropyl acetate, butyl acetate, isobutyl acetate, sec butyl acetate, tert butyl acetate, acetone, and 2-butanoe, methylene chloride, or methyl tert-butyl ether (MTBE).

Analysis of the binary interaction parameters in the context of engineering "modelling: coordinates suggest the follow:

1. The dissolution of lovastatin and simvastatin in all the solvents studied is enthalpy dominated, because $h^E > 0$, $s^E > 0$, and $g^E > 0$.

2. The dissolution of artemisinin in all the solvents studied is also enthalpy dominated because $h^E < 0$, $s^E < 0$, and $g^E < 0$.

The NRTL activity coefficient interaction parameters from the binary system were used to predict the equilibrium mole fraction solubility of lovastatin, simvastatin, or atermisinin in binary solvent mixtures of ethyl acetate/ethanol, or ethyl acetate/acetone, or ethanol acetone. The predicted results compared very favorably with the experimentally measured solubility values.

CHAPTER 5

MOLECULAR SIMULATION

5.1 Introduction

Molecular modeling combine the physical laws of motion with models of atomistic attraction and repulsion to calculate the positions and trajectories of atoms within a molecular structure (Allen and Tildesely, 1987) and molecules within a compound. From the simulation, measured energies, velocities, and positions are used to compute physical properties. Molecular modeling have the potential for being the most broadly used technique for estimating physical properties of compounds. Currently, research in this area continues to expand the number of properties and classes of chemicals that can be estimated by this technique (Frenkel, 2002).

The application of molecular modeling tools has proved to be of great benefit in the pharmaceutical industry in several ways. Within the drug discovery process, modeling techniques have allowed the medicinal chemist to probe complex receptor sites, which has facilitated the design of molecules with specific therapeutic advantages (Damewood, 1996). The use of molecular modeling is valuable complement to existing experimental methods and consequently can streamline and accelerate the drug development process.

The first articles applying the numerical technique of molecular dynamics to a protein were in the mid 1970s, beginning with articles such as those by Levitt and Warshel (1975) and by Karplus and McCammon (1977). In this technique, the classical equations of motion $F=ma$ are integrated forward in time, with the force F being determined from the gradient of a phenomenologically determined potential. Much effort has been devoted

to determine potentials suitable for studying proteins, with AMBER (Kollman, 1995) and CHARMM (Brooks et al., 1983) being two of the most widely used. In the last ~30 years, molecular dynamics has become the standard technique for studying the motion of proteins, with over 10,000 articles published containing the words "molecular dynamics simulations" and "proteins". In this study, molecular simulation was employed to develop a mathematical tool for estimating the ratio of activity coefficients.

5.2 Theoretical framework for molecular dynamics simulation

Molecular dynamic simulation is based on Newton's second law, or the equation of motion. From knowledge of the force on each atom, it is possible to determine the acceleration of each molecule in the system. Integrating the equations of motion then yields a trajectory that describes the positions, velocities and accelerations of the particles as they vary with time. From this trajectory, the average thermodynamic properties of the compounds can be calculated.

5.3 Free energy perturbation theory

Theoretically, a difference in free energy is determined either from the relative probability of finding the system in a given state, or from the reversible work required to transform the system from one initial state into another. Computationally, both approaches can be used. In practice, a perturbation is applied to force the transition from one state to another. Then, statistical procedures are used to calculate the work done on the system by the perturbation. The free energy is a state function, which means that the

free energy difference is only depending on the initial and end state, no matter what path is taken to go from one to the other. As a consequence, one can choose any nonphysical path to perform ones calculations as long as one is able to relate them through thermodynamic cycles to the physical process one is interested in.

5.4 FEP Method

With a given set of potential functions one can simulate experimentally, and observed macroscopic properties using microscopic models. According to the theory of statistical mechanics, one should consider all the quantum mechanical energy levels of the system in order to evaluate different average properties (McQuarrie, 2000). However, in the classical limit it is possible to approximate the average of a given property, A (which is independent of the momentum of the system), by (McQuarrie, 2000):

$$\langle A\rangle = \frac{1}{z(U)}\int A(r)\exp\{-U(r)\beta\}dr \qquad (5\text{-}1)$$

Where dr designates the $z(U)=\int\{-U(r)\beta\}dr$ volume element of the complete space spanned $3n$ vector r associated with n atoms of the system. The evaluation of Eq. (5.1) requires one to explore all points in the entire configuration space of the given system. Such a study of solvated macromolecule is clearly impossible with any of the available computers. However, one can evaluate to see if the average over a limited number of configurations will give similar results to those obtained from an average over the entire space. With this working hypothesis we experiment to look for an efficient way of

spanning phase space. Evaluation of free energies by statistical mechanical approaches are extremely time consuming due to sampling problems. Fortunately, it is possible in some cases to obtain meaningful results using perturbation approaches. Such approaches exploit the fact that many important properties depend on local changes in the macromolecules so that the effect of the overall macromolecular potential cancels out. Such calculations are usually done by the so-called free-energy perturbation (FEP) method (Zwanzig, 1954; Valleau and Torrie, 1977) and the related umbrella sampling method (Valleau and Torrie, 1977). This method evaluates the free energy associated with the change of the potential surface from by gradually changing the potential surface using the relationship

$$U_m(\lambda_m) = U_1(1-\lambda_m) + U_2\lambda_m \tag{5-2}$$

The free-energy increment $\delta G(\lambda_m \Rightarrow \lambda_{m'})$ associated with the change $U_m \Rightarrow U_{m'}$ can described as:

$$\exp\{-\delta G(\lambda_m \Rightarrow \lambda_{m'})\beta\} = \langle \exp\{-(U_{m'} - U_m)\beta\}\rangle_m \tag{5-3}$$

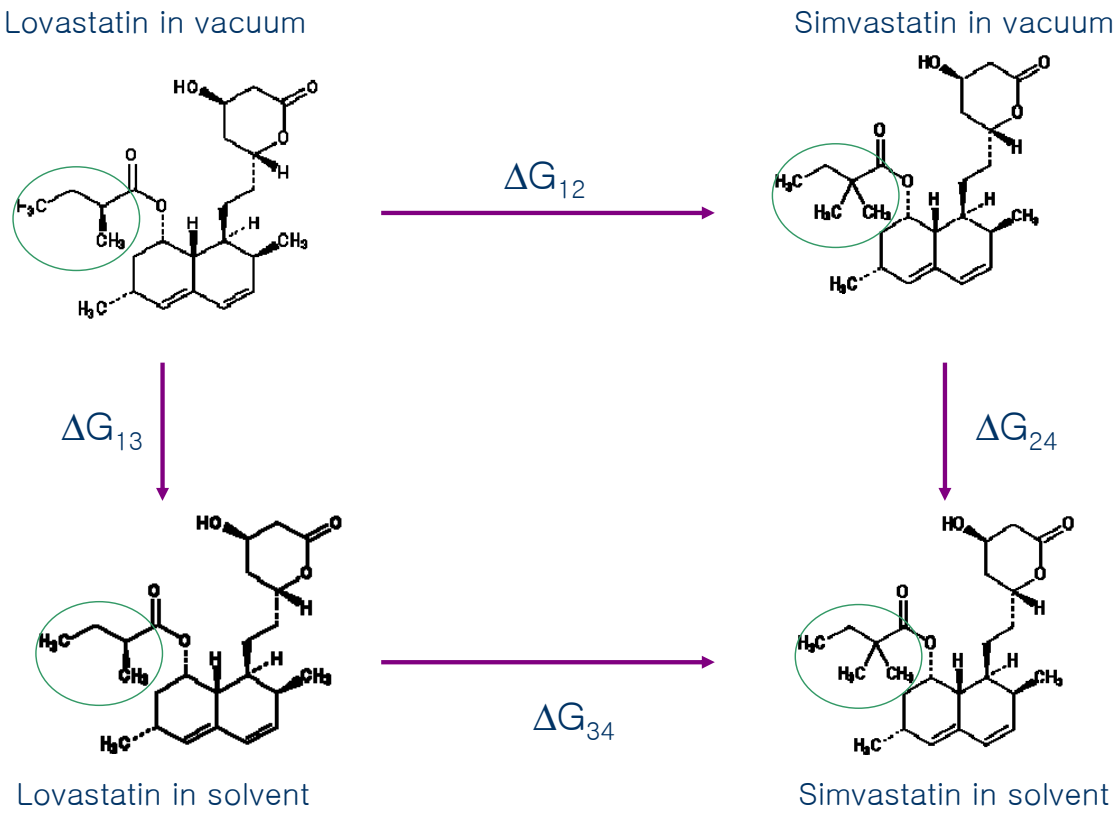
where $\langle\rangle_m$ represents that the given average is evaluated by propagating trajectories over U_m. The overall free energy change can be obtained by changing the λ_m in *n* equal increments and evaluating the sum of the corresponding δG:

$$\Delta G(U_m \Rightarrow U_{m'}) = \sum_{m=0}^{n-1} \delta G(\lambda_m \Rightarrow \lambda_{m+1}) \qquad (5\text{-}4)$$

5.5 Simulation methods

For each simulation, the NAMD Software using the realistic empirical CHARMM force field (Brooks et al., 1983) within the molecular systems was subjected to a constant-moles, constant-pressure-constant-temperature (NPT) ensemble (in which the dynamics are modified to allow the system to exchange heat with the environment at a controlled temperature, maintained by the Nosé-Hoover thermostat (Brooks et al., 1983). An atom based summation method was employed with the non-bonded interactions cutoff set to a distance of 11.0 Å, accompanied by a spline width of 1.0 Å and a buffer width of 0.5 Å. The conformation of the initial models with 3D periodic boundary conditions was first minimized using the steepest descent convergence method, followed by the conjugate gradient method until the convergence reaches 0.1 kcal/mol.Å. The NPT system was then subjected to equilibration (for 100 ps with a time step of 1 fs). The molecular dynamic simulation was allowed to progress with the same operational conditions for up to 400 ps, and recording the atomic coordinates at 1 ps intervals. Coordinate file from the MD simulation was used as starting point for the EP run. Each run was carried out in triplicate at 273K.

From mutation of lovastatin to simvastatin, the following thermodynamic cycle and FEP calculations scheme was employed to obtain the free energy difference:



Scheme 1: Thermodynamic cycle to obtain relative free energy.

Expanding on Scheme 1 translates to,

$$\Delta G_{12} + \Delta G_{24} = \Delta G_{13} + \Delta G_{34} \qquad (5.5)$$

$$\Delta G_{13} - \Delta G_{24} = \Delta G_{12} - \Delta G_{34} \qquad (5.6)$$

$$G_1 = N_A g_A + N_S g_S \qquad (5.7)$$

$$G_2 = N_A g_B + N_S g_S \qquad (5.8)$$

g_A = Gibbs energy of pure A

g_B = Gibbs energy of pure B

$$G_3 = N_A \overline{G_A} + N_S \overline{G_S} \qquad (5.9)$$

$$G_4 = N_A \overline{G_B} + N_S \overline{G_S} \qquad (5.10)$$

$$\Delta G_{12} = G_2 - G_1 = N_A g_B + N_S g_S - N_A g_A - N_S g_S \qquad (5.11)$$

$$\Delta G_{34} = N_A \overline{G_B}^{\infty} + N_S \overline{G_S} - N_A \overline{G_A}^{\infty} - N_S \overline{G_S} \qquad (5.12)$$

Lim of infinite dilution; $N_A = 1$

$$\Delta G_{12} - \Delta G_{34} = N_A g_B + N_S g_S - N_A g_A - N_S g_S - \left(N_A \overline{G_B}^{\infty} + N_S \overline{G_S} - N_A \overline{G_A}^{\infty} - N_S \overline{G_S}\right)$$

$$= g_B - g_A - \overline{G_B}^{\infty} + \overline{G_A}^{\infty}$$

$$(\Delta G_{12} - \Delta G_{34}) = \overline{G_A} - g_A - (\overline{G_B}^{\infty} - g_B)$$

$$\text{Lim } N_A = 1$$

Since $\overline{G_A}(T, P, x) = g_A(T, P) + RT \ln(\gamma_A x_A)$ (5.14)

$$\overline{G_A}(T, P, x) - g_A(T, P) = RT \ln(\gamma_A x_A) \qquad (5.15)$$

Hence: $$\overline{G_A}(T, P, x) - g_A(T, P) = RT \ln(\gamma_A x_A) \qquad (5.16)$$

$\text{Lim } N_A = 1$ $\qquad$ $\text{Lim } x_A = 0$

$$(\Delta G_{12} - \Delta G_{34}) = RT \ln(\gamma_A x_A) - RT \ln(\gamma_B x_B)$$
$$\text{Lim } N_A = 1 \qquad \text{Lim } x_A = x_B = 0 \tag{5.17}$$

$$(\Delta G_{12} - \Delta G_{34}) = RT \ln\left(\frac{\gamma_A^{\infty}}{\gamma_B^{\infty}}\right)_{\text{in the same solvent}}$$
$$\text{Lim } N_A = 1 \qquad \text{Lim } x_A = x_B = 0 \tag{5.18}$$

$$(\Delta G_{12} - \Delta G_{34}) = (1.9858775 \times 10^{-3}\ \text{kCal mol}^{1}\text{K}^{-1} * 273\ \text{K}) * \ln\left(\frac{\gamma^{\infty}_{Lovastatin}}{\gamma^{\infty}_{Simvastatin}}\right) \tag{5.1}$$

Three dimensional structure of lovastatin, simvastatin, and the hybrid molecule are displayed in Figures, 5.1, 5.2, and 5.3, respectively.

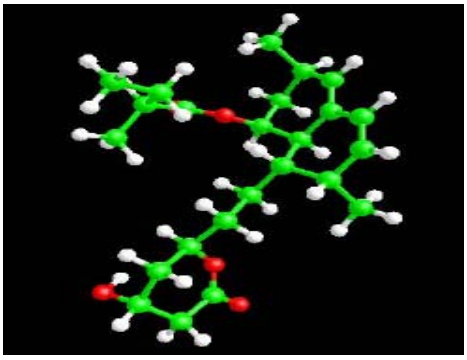


Figure 5.1: Three dimensional structure of lovastatin.

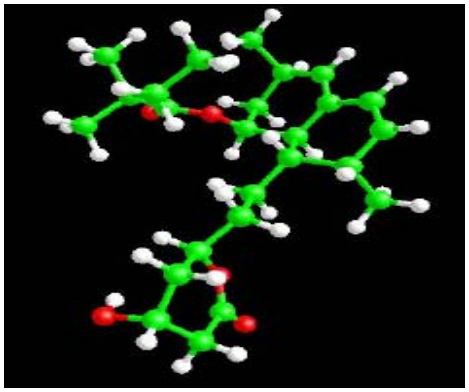


Figure 5.2: Three dimensional structure of simvastatin.

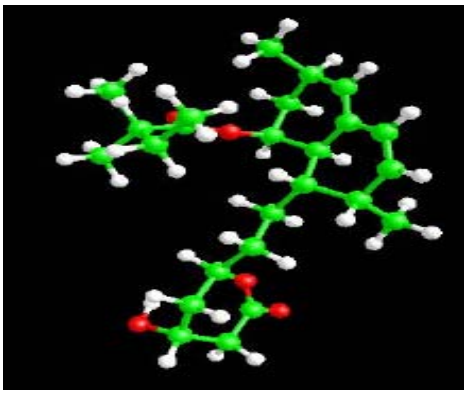


Figure 5.3: Three dimensional structure of lovastatin/simvastatin hybrid molecule.

The CHARMM atom types and charges for the hybrid molecule is presented in Table 5.1; with the atom numbers displayed in Figure 5.4.

From the FEP simulation, and using the thermodynamic cycle, the Gibbs free energy change ΔG was calculated for the mutation of lovastatin to simvastatin in the pertinent solvent, from which the ratio of the activity coefficients were obtained. The results are summarized in Table x.

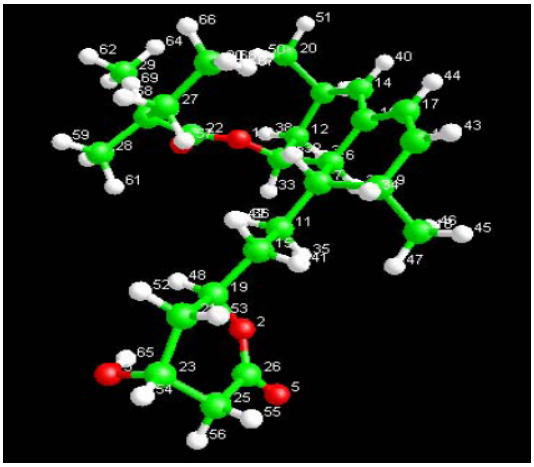


Figure 5.4: Three dimensional structure of the hybrid molecule illustrating the atom numbers.

Table 5.1: CHARMM atom types and charges of the hybrid molecule used in the simulation.

Atom number	Element	Atom Type	Charge	Atom number	Element	Atom Type	Charge
1	O	O	-0.34000	36	H	HA	0.09000
2	O	O	-0.34000	37	H	HA	0.05750
3	O	OH1	-0.66000	38	H	HA	0.05750
4	O	O	-0.26467	39	H	HP	0.11500
5	O	OE	-0.26487	40	H	HP	0.11500
6	C	CPT	-0.09500	41	H	HA	0.09000
7	C	CA	-0.11500	42	H	HA	0.09000
8	C	CA	-0.11500	43	H	HP	0.11500
9	C	CA	-0.11500	44	H	HP	0.11500
10	C	CPT	-0.02000	45	H	HA	0.09000
11	C	CT2	-0.18000	46	H	HA	0.09000
12	C	CT2	-0.11500	47	H	HA	0.09000
13	C	CA	-0.11500	48	H	HP	0.11500
14	C	CA	-0.11500	49	H	HA	0.09000
15	C	CT2	-0.18000	50	H	HA	0.09000
16	C	CA	-0.11500	51	H	HA	0.09000
17	C	CA	-0.11500	52	H	HA	0.05750
18	C	CT3	-0.27000	53	H	HA	0.05750
19	C	CT	0.22500	54	H	HA	0.34500
20	C	CT3	-0.27000	55	H	HA	0.05750
21	C	CT2	-0.11500	56	H	HA	0.05750
22	C	CC	0.26527	57	H	HA	0.09000
23	C	CT1	12.01070	58	H	HA	0.09000
24	C	CT	-0.09125	59	H	HA	0.09000
25	C	CT2	-0.11500	60	H	HA	0.09000
26	C	CC	0.26491	61	H	HA	0.09000
27	C	CT2	-0.18000	62	H	HA	0.09000
28	C	CT3	-0.27000	63	H	HA	0.09000
29	C	CT3	-0.27000	64	H	HA	0.09000
30	H	CT3	-0.27000	65	H	H	0.43000
31	H	HP	0.11500	66	H	HA	0.09000
32	H	HP	0.11500	67	H	HA	0.09000
33	H	HP	0.11500	68	H	HA	0.09000
34	H	HP	0.11500	69	H	HA	0.09000
35	H	HA	0.09000				

5.6 Results and discussion

The Gibbs energy difference, as function of λ, is displayed in Figure 5.5. The calculated overall Gibbs energy difference (left hand side of Eq. (5.19)) for each of the triplicate runs and the averaged value are displayed in Table 5.2. The values compares favorably with the experimentally determined ratio of the activity coefficients at infinite dilution (the right hand side of Eq. (5.19)) – as reflected in Figure 5.6 (slope > 0.99).

It is clear from the simulation results that free energy mutation can be used to obtain the ratio of activity coefficients at infinite dilution. Thus, if the activity coefficient at infinite dilution for a reference compound is known, then mutating the reference compound then from mutating the reference compound to a compound of interest, the activity coefficient at infinity dilution (of the compound of interest) can be obtained via Eq. (5.19).

We were interested checking to see if we could use the activity coefficient at infinite dilution to predict the solubility of simvastatin in solvents used in the simulation. As can be seen from Figure 5.7, a reasonable agreement between the simulated results and experimentally measured values. It is clear from this results that free energy perturbation methods can be used to further reduce the number of solubility measurement experiments needed to predict solubility of drug compounds in liquid solvents.

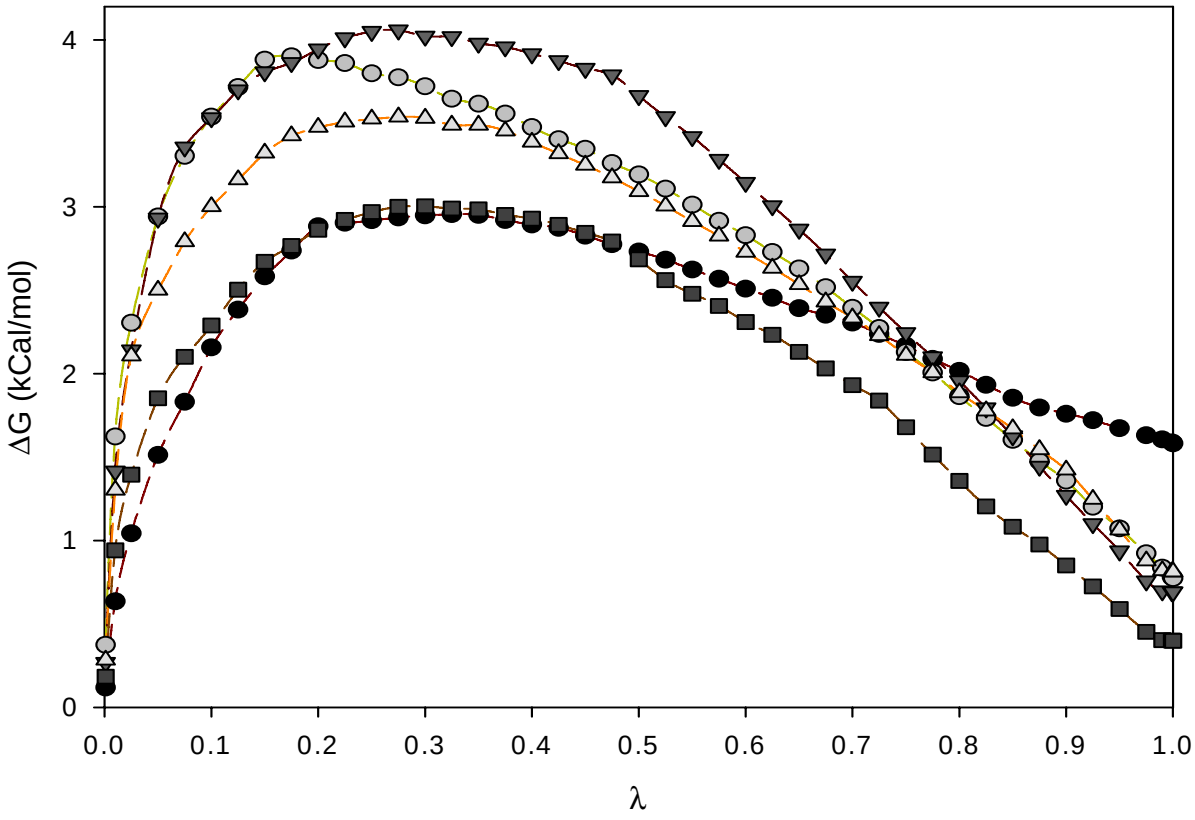


Figure 5.5: Free energy difference from mutation of lovastatin to simvastatin in vacuum and in different solvents at 273 K.

Table 5.2:

Free energy differences from mutation of lovastatin to simvastatin at 273 K in-vacuum, and in different solvents.

Solvent	Run #1 kCal/mol	Run #2 kCal/mol	Run #3 kCal/mol	Average kCal/mol
In vacuum	**1.49**	**1.62**	**1.63**	**1.58**
Ethanol	**0.76**	**0.81**	**0.83**	**0.80**
Ethyl acetate	**0.98**	**0.95**	**1.07**	**1.00**
Butanol	**0.86**	**0.91**	**0.89**	**0.87**
Hexanol	**1.12**	**1.09**	**1.21**	**1.14**
Acetone	**0.78**	**0.83**	**0.85**	**0.82**

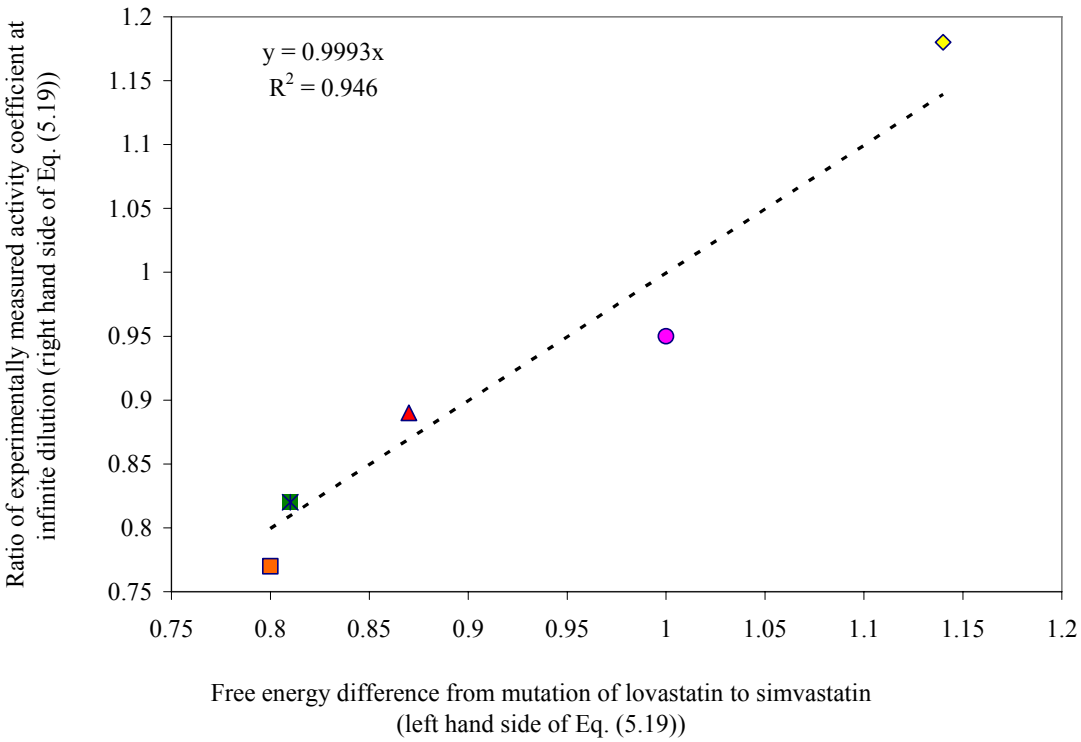


Figure 5.6: Ratio of experimentally measured activity coefficient of lovastatin to simvastatin versus the Gibb's free energy difference from mutation of lovastatin to simvastatin in:

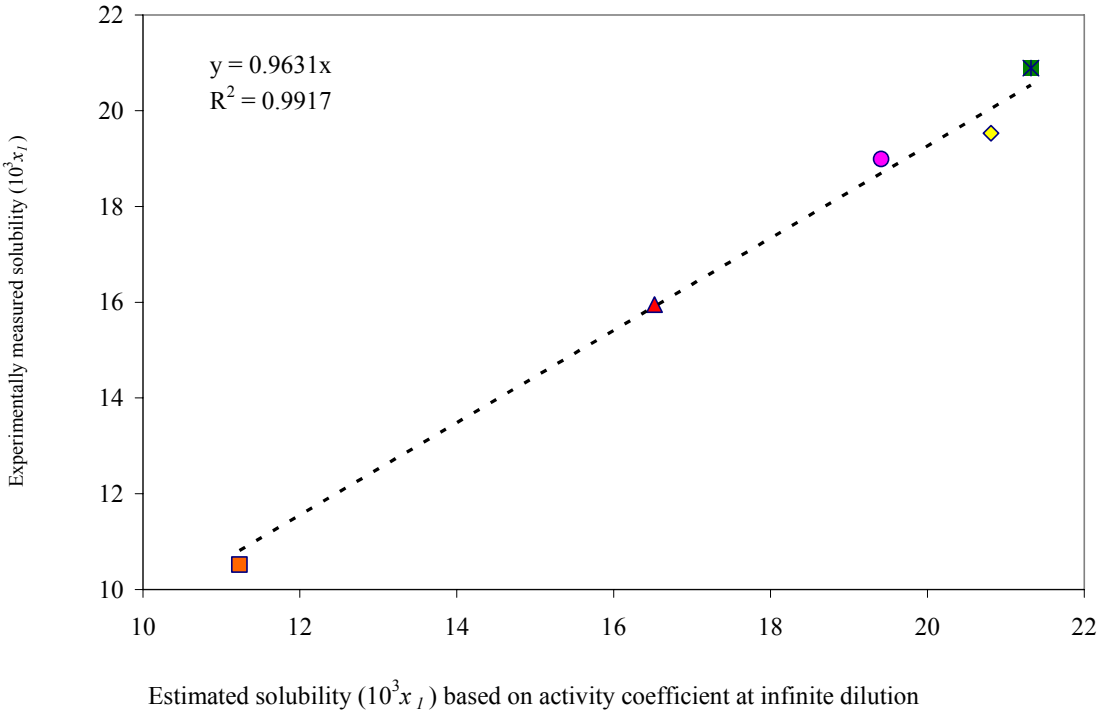


Figure 5.7: Predicted versus experimentally measured solubility of simvastatin in:

5.7 Conclusion

Application of molecular simulation methods for modeling solubility of in liquid solvents is explored. Free Energy Perturbation (FEP) theory has been successfully employed to estimate the ratio of activity coefficients at infinite dilution of lovastatin to simvastatin in five different organic solvents (i.e., ethanol, acetone, butanol, hexanol, and ethyl acetate) at 273 K. The FEP calculation reproduced the experimental solubility results quite well. Additionally, the estimated activity coefficient at infinite dilution for simvastatin was

used to estimate its solubility in each of the solvents studied at 273 K. Pleasantly, the results excellently agreed with the experimentally measured solubility.

CHAPTER 6

SUMMARY AND FUTURE DIRECTIONS

Summary

The primary objective of this thesis was to investigate whether the Non-random Two-Liquid (NRTL) local composition-based activity coefficient model could be used to predict solubility of drug compounds in solvent mixtures using binary interaction parameters. A secondary objective was to explore free energy perturbation calculation for prediction of drug solubility in liquid solvents. For this purpose, both experimental and computational studies were performed. A new small-scale analytical method for experimental solubility determination of crystalline compounds was devised. This method was used to experimentally determine solubility values used for the computational model assessment.

The results show that high quality solubility data of crystalline compounds can be obtained by the new method (Nti-Gyabaah et al., 2008, Nti-Gyabaah and Chiew, 2008, Nti-Gyabaah et al., 2009). One key advantage with the present method is that, significantly less material is required for solubility measurement less material. For example, less than 10 mg of material was needed to measure the equilibrium solubility of lovastatin in acetone, compared to a requirement of more than 1000 mg by Sun and co-workers (2005) when they measured solubility of the same compound in the same

solvent). The other advantage is that, it allows determination of solvent-mediated compound degradation, which is a very important component of solvent selection.

Furthermore, calculation of the ideal solubility of crystalline solid in liquid solvent requires knowledge of the difference in molar heat capacity of the solid and super-cooled liquid solute at the solution temperature. This is a hypothetical parameter and therefore, can not be measured directly. Hence three assumptions are commonly used in the literature for its estimation. Evaluation of the assumptions revealed thermodynamic inconsistencies, especially at temperatures far below the solute melting point. A new strategy was explored to estimate the parameter, allowing the experimental solubility data to be fitted to the Two-Liquid-Non-Random (NRTL) activity coefficient model to obtain model energetic interaction parameters. The interaction parameters were successfully used to estimate solubility of the model compounds in mixed solvents over a temperature range.

Application of molecular dynamics (MD) simulation for modeling solubility of crystalline solids in liquid solvents is explored. Molecular dynamics simulation has been coupled with Free Energy Perturbation (FEP) theory to estimate activity of coefficient at infinite dilution of pharmaceutical compounds. The technique provides basis for developing molecular modeling tool to estimate solubility of pharmaceutically-important compounds in liquid solvents. FEP and MD simulations on mutation of lovastatin to simvastatin in five different solvents, were carried out to obtain the free energy

differences, and the free energy of mixing and activity coefficient at infinite dilution. The FEP calculation reproduced the experimental solubility results quite well.

Future Directions

This thesis focused on modeling of non-electrolyte drug compounds. The established arguments are thermodynamically sound, and should be applicable to electrolyte drug solids. It is expected that the methods established should have wide application for not only drug compounds, but also, for non-drug compounds which are processed at moderately low temperatures (e.g., fine chemicals).

Although this thesis has highlighted the need to use accurate differential molar heat capacity for modeling solubility, additional research is needed to refine the technique used to obtain the parameter. One option is to explore molecular simulation techniques for estimating the parameter. Whichever technique is used to obtain this parameter, we strongly recommend that one should examine it for thermodynamic consistency before attempt is made to use it to model solubility data. One way to do this is reconciling it with heat of mixing data, using a van't Hoff-like plot.

For the FEP method we only used mutation of the reference solute to another solute in the same solvent. This provides a basis for exploring mutation of one solvent to the other in the presence of the solute. Such approach will allow one to estimate solubility of the same compound in different solvents. It is expected that the high quality experimental

data that have been generated through this thesis should be helpful for validation when this option is explored.

LISTS OF NOTATIONS

x_1	Mole fraction solubility of solute
x_2	Mole fraction solubility of solvent
x_1^{id}	Ideal mole fraction solubility of solute
T_m	Melting point temperature of solute
T	Absolute temperature of solution
ΔC_P	Difference in the molar heat capacity (at constant pressure).
R	Gas constant
C_P^L	Molar heat capacities of the liquid solute
C_P^S	Molar heat capacity solid solute
$\Delta_{fus}H$	Change enthalpy of melting
$\overline{H_1^{E,\infty}}$	Limiting partial molar excess enthalpy
$\overline{S}_1^{E,\infty}$	Limiting partial molar excess enthalpy
$\Delta_{fus}S$	Entropy of fusion at solute melting point
G_{12}	Parameter in the NRTL model
G_{21}	Parameter in the NRTL model

LITERATURE CITED

Abrams, D.S., and Prausnitz, J.M. Statistical Thermodynamics of Liquid Mixtures: A New Expression for the Excess Gibbs Energy of Partial or Completely Miscible Systems. *AIChE Journal*, 1975, 21, (1) 116-128.

Abildskov, J.; O'Connell, J.P. Predicting the Solubility of Complex Chemicals I. Solutes in Different Solvents. *Ind. Chem. Eng. Res.* 2003, 42, 5622.

Acree W.E. Jr, McCargar J.W., Zvaigzne AI, Teng IL. Mathematical representation of thermodynamicproperties. Carbazole solubilities in binary alkane + dibutyl ether and alkane + tetrahydropyran solvent mixtures: *Phys Chem Liq*, 1991, 23: 27-35.

Acree W.E. Jr. Mathematical representation of thermodynamic properties. Part 2. Derivation of thecombined nearly ideal binary solvent (NIBS) /Redlich-Kister mathematical representation from a two-body and three-body interactional mixing model: *Thermochimica Acta*, 1992, 198: 71-79.

Álvarez, V. H, and Valderrama, J. O. Método Modificado Lydersen-Joback-Reid para Estimar Propiedades Críticas de Biomoléculas. *Alimentaria,* 2004, 354, 55-66.

Adjei A., Newburger J., Martin A. (1980) Extended hildebrand approach: Solubility of caffeine in dioxane-water mixtures: *J. Pharm. Sci.* 1980, 69, 659-661.

Amidon N.A. Excess free energy approach to the estimation of solubility in mixed solvent systems: *J. Pharm. Sci.* 1990, 73, 9-13.

Apelblat, A., Manzurola, E. Solubilities of o-Acetylsalicylic, 4-Aminosalicylic, 3,5-Dinitrosalicylic, and p-Toluic Acid, and Magnesium-DL-Aspartate in Water from (278 to 348) K. *J. Chem. Thermodyn.* 1999, 31, 85–91.

Bevan C. D., Lloyd R. S. A high-throughput screening method for the determination of aqueous drug solubility using laser nephelometry in microtiter plates. *Anal. Chem*. 2000, 72:1781–1787.

Barzegar-Jalali M, Hanaee J. Model for solubility estimation in mixed solvent systems. *Int. J. Pharm*. 1994, 109: 291-295.

Barzegar-Jalali M, Jouyban-Gh A. Models for calculating solubility in binary solvent systems. *Int J. Pharm* 1996, 140: 237-246.

Bettini R., Bertolini G., Frigo E., Rossi A., Casini I., Pasquali I. and F. Giordano, *J. Therm. Anal. Calorim.* 2004, 77: 625–638.

Constantinou, L., and Gani, R. New Group Contribution Method for Estimating Properties of Pure Compounds. *AIChE Journal*. 1994, 40, (10) 1697-1710.

Demirel Y. Calculation of infinite dilution activity coefficients by the NRTL and UNIQUAC models, *Canadian J. of Chem. Eng*. 1990, 68: 697-701.

Ettmayer P., Amidon G.L., Clement B., Testa B. Lessons learned from marketed and investigational prodrugs, *J. Med. Chem*. 2004, 47: 2393–2404.

Frank, T. C., Downey, J. R., Gupta, S. K. Quickly screen solvents for organic solids. *Chem. Eng. Prog.* 1999, 95: 41-61.

Frank T.C., Downey J.R., Gupta S.K. Quickly screening solvents for organic solvents. *Chem Eng. Prog.* 1999, 41: 175-182.

Fredenslund, A., Jones, R.L., and Prausnitz, J.M. Group Contribution Estimation of Activity Coefficients in Nonideal Liquid Mixtures, *AIChE Journal*, 1975, 21: 1086-1099.

Fleisher D., Bong R., and Stewart B.H. Improved oral drug delivery: solubility limitations overcome by the use of prodrugs, *Adv. Drug Deliv. Rev.* 1996, 19: 115–130.

Gabas, N.; Laguérie, C. Predictions with UNIFAC of liquid-solid phase diagrams: applications water-sucrose-glucose, water-sucrose-fructose and water-xylose-mannose. *J. Cryst. Growth.* 1993, 128: 1245-1249.

Gracin, S., Brinck T., Rasmuson, A. C. Prediction of solubility of solid organic compounds in solvents by UNIFAC. *Ind. Eng. Chem. Res*. 2002, 41: 5114-5124.

Grant D.J.W., Mehdizadeh M, Chow AHL, Fair-Brother JE. Non-linear van't Hoff solubility temperature plots and their pharmaceutical interpretation. *Int J. Pharm.* 1984, 18: 25-38.

Gupta, R. B.; Heidesmann, R. A. Solubility models for amino acids and antibiotics. *AIChE J.* 1990, 36 (3): 333-341.

Gmehling, J. G.; Andersson, T. F.; Prausnitz, J. M. Solid-liquid equilibria using UNIFAC. *Ind. Eng. Fundam*. 1978, 17: 269-273.

Steve Hamm. Creative Capitalism' versus Malaria. *Business Week:* 2009, 083

Huang L.F., and Tong W.Q. Impact of solid state properties on developability assessment of drug candidates, *Adv. Drug Deliv. Rev*. 2004, 56: 321–334.

Herrador, M.A. and Gonzalez, A.G. Solubility prediction of caffeine in aqueous *N*,*N*-dimethylformamide mixtures using the extended Hildebrand solubility approach. *Int. J. Pharm.* 1997, 156: 239–244.

Hildebrand J.H. and Scott R.L., Entropy of solution of nonelectrolytes, *J. Chem. Phys.* 1952, 20: 1520–1521.

Jiang, Q.; Gao, G.-H.; Yu, Y.-X.; Qin, Y. Solubility of Sodium Dimethyl Isophthalate-5-sulfonate in Water and in Water + Methanol Containing Sodium Sulfate. *J. Chem. Eng. Data.* 2000, 45: 292-294.

Jouyban, A., Chew, N.Y.K., Chan, H.K., Sabour, M and Acree, W.E. Jr. A unified cosolvency model for calculating solute solubility in mixed solvents, *Chemical & Pharmaceutical Bulletin*. 2005, 53: 634-637.

Jouyban-Gh A. The modified Wilson model and predicting drug solubility in water cosolvent mixtures. *Chem Pharm Bul.* 1998, 46: 1058-1061.

Jouyban-Gh A, York P, Hanna M, Clark BJ. Solubility prediction of salmeterol xinafoate in waterdioxane. *Int J Pharm* 2001, 216: 33-41.

Jouyban, A., Khoubnasabjafari, M., Chan, H.K., Clark, B.J. and Acree Jr., W.E. Solubility prediction of anthracene in mixed solvents using a minimum number of experimental data. *Chem Pharm. Bull*. 2002, 50: 21-25.

Jouyban-Gh A, Acree WE Jr. Comparison of models for describing multiple peaks in solubility profiles. *Int J Pharm.* 1998, 167: 177-182.

Jouyban-Gharamaleki, A., Barzegar-Jalali, M. and Acree, W.E. Jr. Solubility correlation of structurally related drugs in binary solvent mixtures, *International Journal of Pharmaceutics*. 1998, 166: 205-209.

Jouyban A., Romero S., Chan H., Clark B., Bustamante A. A co-solvency Model to predict solubility of drugs at several temperatures from a limited number of solubility measurements. *Chem Pharm. Bull.* 50 (5) 594 – 599 (2002).

Kolar P., Shen J.W., Tsuboi A, Ishikawa T., Solvent selection for pharmaceuticals. *Fluid Phase Equilibria*. 2002, 194: 42-48.

Kibbey C.E., Poole S.E., Robinson B., Jackson J.D., and Durham D. An integrated process for measuring the physicochemical properties of drug candidates in a preclinical discovery environment. *J. Pharm. Sci.* 2001, 90:1164–1175.

Kuramochi, H.; Noritomi, H.; Hoshino, D.; Nagahama, K. Measurements of solubility of two amino acids in water and prediction by the UNIFAC model. *Biotechnol. Prog*. 1996, 12, 371-379.

Kuramochi, H.; Noritomi, H.; Hoshino, D.; Nagahama, K. Representation of activity coefficient of fundamental biochemicals in water by the UNIFAC model. *Fluid Phase Equilib*. 1997, 130, 117-132.

Kretchmer C.B., and Wiebe R., Thermodynamics of Alcohol-hydrocarbon mixtures., *J. Chem. Phys*. 1954, 22: 1697-1699.

Larsen, B. L., Rasmussen, P., and Fredenslund, A. (1987), A Modified UNIFAC Group-Contribution Model for Prediction of Phase Equilibria and Heats of Mixing. *Ind. Eng. Chem. Res*. 1987. 26: 2274-2286.

Lipinski C.A. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.* 2001, 46: 3–26.

Liu. B.S., Gong J.B., Wang J.K. and Jia C.Y. Solubility of potassium clavulate in ethanol, 1-propanol, and 2-methyl-1-propanol between 273 K and 305 K. *J. Chem. Eng. Data.* 2005, 50, 1684 – 1686.

Martin A, Newburger J, Adjei A. Extended Hildebrand solubility approach: Solubility of theophylline in polar binary solvents. *J Pharm Sci*. 1980, 69: 487-491.

May, W. E.; Wasik, S. P.; Freeman, D. H. Determination of the Aqueous Solubility of Polynuclear Aromatic Hydrocarbons by a Coupled Column Liquid Chromatographic Technique. *Anal. Chem.* 1978, 25: 17-23.

Mohan R., Lorenz H., and Myerson A.S. Solubility measurement using differential scanning calorimetry. *Ind. Eng. Chem. Res.* 2002, 41:4854–4862.

Mirmehrabia M., Rohani S., Murthy S.K., and Radatusb B. Solubility, dissolution rate and phase transition studies of ranitidine hydrochloride tautomeric forms. *Int. J. of Pharm.* 2004, 282, 1-2, (2004): 73-85.

Mishra D.S. and S.H. Yalkowski. Ideal solubility of a solid solute: Effect of Heat Capacity Assumptions. *Pharm. Res.,* 1992, 9: 82-88.

Mutabingwa TK, Anthony D, Heller A. Amodiaquine alone, amodiaquine+sulfadoxine-pyrimethamine, amodiaquine+artesunate, and artemether-lumefantrine for outpatient treatment of malaria in Tanzanian children: a four-arm randomised effectiveness trial. *Lancet*. 2005, 365: 1474–8

Nass, K. K. Rational solvent selection for cooling crystallization. *Ind. Eng. Chem. Res*. 1994, 33: 1580-1584.

Neau, S. H.; Flynn, G. N. Solid and liquid heat capacities of n-alkyl para-aminobenzoates near the melting point. *Pharm. Res*. 1990, 7 (11): 1157-1163.

Neau, S. H.; Flynn, G. N.; Yalkowsky, S. H. The influence of heat capacity assumptions on the estimation of solubility parameters from solubility data. *Int. J. Pharm*. 1989, 49: 223-229.

Neau, S. H.; Bhandarkur, S. V.; Helmuth, E. W. Differential molar heat capacities to test ideal solubility estimations. *Pharm. Res.* 1997, 14 (5): 601-605.

Nordstrom, F. L.; Rasmuson, A. C. Solubility and melting properties of salicylic acid. *J. Chem. Eng. Data.* 2006, 51: 1668-1671.

Niea Q, and Wang J.K. Solubility of 16,17-Epoxyprogesterone in Six Different Solvents. *J. Chem. Eng. Data.* 2005, 50: 1750 -1752.

Niea Q., Wang J.K., and Yina Q. Separation and purification of two isomorphic steroids by a novel extractive drowning out crystallization. *Separation and Purification Technology*, 2006, 50: 342-346.

Niea Q., Wang J.K., Wang Y.-L., Wang S. Solubility of 11α-hydroxy-16α,17α-epoxyprogesterone in different solvents between 283 K and 323 K. *J. Chem. Eng. Data*. 2005: 989-992.

Nti-Gyabaah J., Antia F.D., Dahlgren M.E. and Göklen K.E.. Normal phase high-performance liquid chromatography of pneumocandins; In situ modification of silica with L-proline to separate structural analogs. *J. Biotech Prog.,* 2006, Vol. 22: 538-546.

Nti-Gyabaah J., Antia F.D., Dahlgren M.E. and Göklen K.E.. High-performance liquid chromatography of pneumocandins; modification of silica with L-proline to separate structural analogs. *J. Bioprocess R&D*, 2005, 6: 35-45.

Nti-Gyabaah, J., Chan, V., Chiew Y.C. Solubility of and Limiting Activity Coefficient of Simvastatin in Different Organic Solvents. *Fluid Phase Equilib.* 2009, 280: 35-41.

Nti-Gyabaah, J. and Chiew, Y.C. Solubility of Lovastatin in Ethyl Acetate, Butyl Acetate, Acetone, Methanol, Ethanol, and 2-Butanone. *J. Chem and Eng. Data,* 2008, 53: 2060-2065.

Nti-Gyabaah, J, Chmmielowski, J., Chan, V., and Chiew Y.C. Solubility of Lovastatin in a Family of Six Alcohols: Ethanol, 1-Propanol, 1-Butanol, 1-Pentanol, 1-Hexanol, and 1-Octanol. *Int. J Pharm*. 2008, 339: 111-117.

Noedl H, Se Y, Schaecher K, Smith BL, Socheat D, Fukuda MM. "Evidence of artemisinin-resistant malaria in western Cambodia". *N. Engl. J. Med.* 2008, 359 (24): 2619–20.

Ochsner, A. B., Sokoloski, T. D. Prediction of Solubility in Nonideal Multicomponent Systems using the UNIFAC Group Contribution Model. *J. Pharm. Sci.* 1985, 74, 634-637.

Pinho, S. P.; Silva, C. M.; Macedo, E. A. Solubility of amino acids: A Group-Contribution Model involving phase and chemical equlibria. *Ind. Eng. Chem. Res.* 1994, 33, 1341-1347.

Prausnitz, J.M., Lichtenthaler, R.N., Gomes de Azevedo, E., 1986. Molecular Thermodynamics of Fluid-Phase Equilibria, second ed. Prentice-Hall, Englewood Cliffs, NJ.

Paul E.L., H.-H. Tung, M. Midler. Organic crystallization processes. *Powder Technology*. 2005, 150: 133- 143.

Pan L., Ho Q., Tsutsui L., and Takahashi L. Comparison of chromatographic and spectroscopic methods used to rank compounds for aqueous solubility. *J. Pharm. Sci.* 2001, 90:521–529.

Prankerd RJ. Solid-state properties of drugs. I. Estimation of heat capacities for fusion and thermodynamic functions for solution from aqueous solubility-temperature dependence measurements. *Int. J. Pharm* 1992; 84: 233-244.

Reid R.C., Prausnitz,and Sherwood T.K. (1977). The properties of Gases and Liquids. 3rd Edition, McGraw Hill Book Company, New York.

Ren G.B and Wang J.K. Solubility of DL-p-hydroxyphenylglycine sulfate in binary acetone+water solvent mixtures*, J. Chem. Eng. Data.* 2004, 49: 1376–1378.

Ren, G. B.; Wang, J. K.; Yin, Q. X.; Zhang, M. J. Solubilities of proxetine hydrochloride hemihydrate between 286 K and 363 K. *J. Chem. Eng. Data*. 2004, 49, 1671-1674.

Renon H., Prausnitz J.M., Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE Journal*. 1968, 14: 153 – 144.

Regosz A., Pelplinska T., Kowalski P., Thiel Z. Prediction of solubility of sulfonamides in water and organic solvents based on the extended regular solution theory. *Int J Pharm.* 1992, 88: 437-442.

Shukla K.P., Chiavlo., Haile J.M. Thermodynamics Excess Properties in Binary Fluid Mixtures. *Ind. Eng. Chem. Res.* 1988, 27: 664-683.

Smith, J.M., van Ness, H.C., and Abbott, M.M. (1996), Introduction to Chemical Engineering Thermodynamics, McGraw-Hill, New York.

Shiu, W. Y., Doucette, W., Gobas, A. P. C., Andren, A., Mackay, D. Physical-Chemical Properties of Chlorinated Dibenzo-p-dioxins. Environ. *Sci. Technol.* 1988, 22, 651-658.

Sun H., Gong J., and Wang J.K. Solubility of Lovastatin in Acetone, Methanol, Ethanol, Ethyl Acetate, and Butyl Acetate between 283 K and 323 K. *J. Chem. Eng. Data.* 2005, 50, 1389.

Tetko, V.Y. Tanchuk, T.N. Kasheva and A.E.P. Villa, *J. Chem. Inf. Comput. Sci.* 2001, 41: 1488–1493.

Techene S., Autret J.M., Biscans B. Determination of solubility of eflucimibe polymorphs: Experimental and modeling. *J. Pharm Sci*. 2006, 4: 871-882.

Tung H., Tabora J., Variankaval N., Bakken D., and Chen C-C. Prediction of Pharmaceutical Solubility via NRTL-SAC and COSMO-SAC. *J. Pharm. Sci.,* 2007: 1 – 6.

Tiziana, F.; Montanes, F.; Olano, A.; Ibanez, E. Modeling ssolubilities of sugars in alcohols based on original experimental data. *J AIChE.* 2007, 53, 9, 2411-2418.

Wang, S.; Wang J.K.; Yin, Q.X. Measurement and correlation of solubility of 7-aminocephalosporanic acid in aqueous acetone mixtures. *Ind. Eng. Chem. Res.* 2005, 44, 3783-3787.

Wasik, S.P. Octanol/Water Partition Coefficient and Aqueous Solubilities of Organic Compounds. NBS Report NBSIR 81-2406. Washington, DC: National Bureau of Standards, U.S. Department of Commerce (1981).

Welch C.J., DaSilva J.O., Nti-Gyabaah J., Antia F.D, Göklen K.E., and Boyd R.F., Preparation and evaluation of novel stationary phases for improved chromatographic purification of pneumocandin. B_0. *J. Chromatogr.* 2006, 1101: 204–221.

Williams NA, Amidon GL. Excess free energy approach to the estimation of solubility in mixed solvent systems: I. Theory. *J Pharm Sci* 1984, 1997; 152: 247-250.

Wu, H. S.; Sandler, S. I. Use of *ab Initio* Quantum Mechanics Calculation in Group Contribution Methods. Theory and Basis for Group Identifications. *Ind. Eng. Chem. Res*. 1991, 30, 881-889.

Yalkowsky S.H. Aqueous Solubility. Methods of Estimation for Organic Compounds, Marcel Dekker, New York (1992).

Yalkowsky SH, Roseman T.J. Solubilization of drugs by cosolvents. In: Yalkowsky SH, ed., Techniques of solubilization of drugs, New York; Marcel Dekker, 1981: 91-134.

Yu X, Zipp GL, Davidson III GWR. The effect of temperature and pH on the solubility of quinolone compounds: Estimation of heat of fusion. *Pharm Res.* 1994; 11: 522-527.

APPENDIX

MATLAB code used for solubility prediction

```
clc
close all
clear all

% Constants
   R     = ;
   Tm    = ;
   delH  = ;
   delCP = ;

%*********************************************************************
*********
% Data that have to be changed for different mixtures

% Data that have to be changed for different mixtures

% Interaction parameters Aij for NRTL equation (Gmehling et al. 1977-
1990)
A11=0;                   % Pure component 1 (Joule/mol)
A22=0;                   % Pure component 2 (Joule/mol)
A33=0;                   % Pure component 3 (Joule/mol)
A12=-596.8;               % Binary mixture 1-2 (Joule/mol
A21=4424;                % Binary mixture 1-2 (Joule/mol)
A13=-509;               % Binary mixture 1-3 (Joule/mol)
A31=2897;               % Binary mixture 1-3 (Joule/mol)
A23=368.91;                 % Binary mixture 2-3 (Joule/mol)
A32=265.80;                 % Binary mixture 2-3 (Joule/mol)

% Alpha interaction parameter
alpha11=0;
alpha22=0;
alpha33=0;
alpha12=0.48;                     % Binary mixture 1-2 (Joule/mol)
alpha21=alpha12;                  % Binary mixture 2-1 (Joule/mol)
alpha13=0.48;                     % Binary mixture 1-3 (Joule/mol)
alpha31=alpha13;                  % Binary mixture 3-1 (Joule/mol)
alpha23=0.30;                     % Binary mixture 2-3 (Joule/mol)
alpha32=alpha23;                  % Binary mixture 3-2 (Joule/mol)

T = input (' Temperarure(K) = ');
% nA=1;

nB=9.1629305941774;
nC=1.7126112437595;
```

```
% End of data that have to be changed for different mixtures
%*********************************************************************
*********

% Binary interaction parameters 'tafji' for NRTL equation
taf11=A11./(R*T');            % Pure component 1
taf22=A22./(R*T');            % Pure component 2
taf33=A33./(R*T');            % Pure component 3
taf12=A12./(R*T');            % Binary mixture 1-2
taf21=A21./(R*T');            % Binary mixture 1-2
taf13=A13./(R*T');            % Binary mixture 1-3
taf31=A31./(R*T');            % Binary mixture 1-3
taf23=A23./(R*T');            % Binary mixture 2-3
taf32=A32./(R*T');            % Binary mixture 2-3

% Binary interaction parameters 'Gji' for NRTL equation
G11=exp(-alpha11*taf11);            % Pure component 1
G22=exp(-alpha22*taf22);            % Pure component 2
G33=exp(-alpha33*taf33);            % Pure component 3
G12=exp(-alpha12*taf12);            % Binary mixture 1-2
G21=exp(-alpha21*taf21);            % Binary mixture 1-2
G13=exp(-alpha13*taf13);            % Binary mixture 1-3
G31=exp(-alpha31*taf31);            % Binary mixture 1-3
G23=exp(-alpha23*taf23);            % Binary mixture 2-3
G32=exp(-alpha32*taf32);            % Binary mixture 2-3

% Logarithm of activity coefficients 'LNGAMMA'
nA=linspace(0.00001,0.5,4500);

LNGAMMA1=(taf11*G11*nA./(nA+nB+nC)+taf21*G21*nB./(nA+nB+nC)+taf31*G31*(
nC./(nA+nB+nC))./(G11*nA./(nA+nB+nC)+G21*nB./(nA+nB+nC)+G31*(nC./(nA+n
B+nC)))...
    +
nA./(nA+nB+nC)*G11./(G11*nA./(nA+nB+nC)+G21*nB./(nA+nB+nC)+G31*(nC./(nA
+nB+nC))).*(taf11 -
(nA./(nA+nB+nC)*taf11*G11+nB./(nA+nB+nC)*taf21*G21+(nC./(nA+nB+nC))*taf
31*G31)/(G11*nA./(nA+nB+nC)+G21*nB./(nA+nB+nC)+G31*(nC./(nA+nB+nC))))...
    +
nB./(nA+nB+nC)*G12./(G12*nA./(nA+nB+nC)+G22*nB./(nA+nB+nC)+G32*(nC./(nA
+nB+nC))).*(taf12 -
(nA./(nA+nB+nC)*taf12*G12+nB./(nA+nB+nC)*taf22*G22+(nC./(nA+nB+nC))*taf
32*G32)/(G12*nA./(nA+nB+nC)+G22*nB./(nA+nB+nC)+G32*(nC./(nA+nB+nC))))...
    +
nC./(nA+nB+nC)*G13./(G13*nA./(nA+nB+nC)+G23*nB./(nA+nB+nC)+G33*(nC./(nA
+nB+nC))).*(taf13 -
(nA./(nA+nB+nC)*taf13*G13+nB./(nA+nB+nC)*taf23*G23+(nC./(nA+nB+nC))*taf
33*G33)/(G13*nA./(nA+nB+nC)+G23*nB./(nA+nB+nC)+G33*(nC./(nA+nB+nC))));

% Activity coefficients 'GAMMA'

yA=LNGAMMA1;                          % Component 1

lnxideal=-( delH*(1- T/Tm)/(R*T) + delCP*( 1 - Tm/T + log(Tm/T) )/R );
```

```
yI=lnxideal;

lnx=log(nA./(nA+nB+nC));

yX=lnx;

final = yX - yI + yA;

[vmin,II]=min(abs(final));
figure(1)
plot(nA,final,'--r'),hold on
plot(nA(II),vmin,'sk')
xlabel('nA')
ylabel('Final')
text(nA(II),vmin,['\leftarrow
nA=',num2str(nA(II))],'HorizontalAlignment','left')
```

MATLAB CODE USED FOR NRTL BINARY INTERACTION PARAMETERS IN PURE SOLVENTS

```
function SolubilityRegressionforLVinEtOH()

clc

% Constants
   R     = ;
   Tm    = ;
   delH  = ;
   delCP = ;
   alpha = ;

   bo    = [100 100]   ;    % initial condition

   opt1 = optimset('Display', 'off');

% Experimental data - Column 1 : T ; Column 2-end : x1

%% data set 1 - comment out data-set 1 or data-set 2 depending on which
set
%% is being used for the fitting
%
%% data set-2 -
%
% data = [279.1 0.01734 0.02086 0.02086 0.01648 0.02294 0.01837 0.02016
0.01636
%         282.2 0.01803 0.02184 0.02164 0.01729 0.02421 0.01916 0.02073
0.01725
%         285.3 0.01886 0.02286 0.02259 0.01853 0.02551 0.02032 0.02166
0.01819
%         288.5 0.01946 0.02436 0.02355 0.01977 0.02723 0.02127 0.0224
0.01938
%         291.3 0.02038 0.02579 0.0246  0.02095 0.02857 0.02218 0.02335
0.02031
%         296.1 0.02204 0.02815 0.02646 0.02327 0.03116 0.02417 0.02509
0.02184
%         301.0 0.02366 0.03061 0.02849 0.02583 0.03452 0.02645 0.02677
0.02356
%         305.8 0.0261  0.03342 0.03108 0.02907 0.03749 0.02907 0.02933
0.02612
%         310.7 0.02889 0.03664 0.03405 0.03208 0.04088 0.03208 0.03225
0.02906
%         315.5 0.0321  0.04033 0.03745 0.03552 0.04476 0.03552 0.0356
0.03244];
%
```

```
%%  Non-linear fitting

figure;

for j = 2:size(data,2)

    k = nlinfit(data(:,1),data(:,j), @myfun, bo, opt1)
    subplot(4,4,j-1), plot(data(:,1),data(:,j), 'o',
data(:,1),myfun(k,data(:,1)));

end

suptitle(' x1 v/s Temperature');
legend('Expt', 'Fitted');

%% Function to be fitted
% Since nlinfit accepts a function of the form y = f(x,parameters), and
in
% this case we do not have a x1 as a explicit function of T, we have to
% use a numerical solution to obtain x1(T) at a particular guess value
of
% [b12 b21].

      function f = myfun(b,T)

          C1 = exp(-alpha*b(1)/R./T) ;  % temporary variables
          C2 = exp(-alpha*b(2)/R./T) ;  % temporary variables

          f = zeros(size(T));             % initializes numerical
solution for x1

          for i = 1: length(T)            % Numerically solve for x at
each value of T

              f1 = @(x1) log(x1)+( delH*(1- T(i)/Tm)/(R*T(i)) +
delCP*( 1 - Tm/T(i) + log(Tm/T(i)) )/R ) + ((1-x1)^2)*( b(1)*C1(i)/(( 1
- x1 + x1*C1(i) )^2)/R/T(i) + b(2)/R/T(i)*( C2(i)/( x1 + (1-x1)*
C2(i) ) )^2);
              f(i) = lsqnonlin(f1, data(i,j), 0, 1, opt1);      %
numerical solution for x1
          end
      end

end
```

CARICULUM VITAE

JOSEPH NTI-GYABAAH

EDUCATION:

Rutgers University	2006 - 2009	Chem. Eng.	Ph.D
Rutgers University	2006 - 2009	Chem. Eng.	M.S
Rutgers University	1991 - 1996	Chem. Eng.	B.S

- **INDUSTRIAL EXPERIENCE:**

Merck Research Laboratories August, 1997 - Present

- **PEER-REVIEWED PUBLICATIONS:**

J. Nti-Gyabaah, R. Chmmielowski, V. Chan, and Y.C. Chiew. Solubility of Lovastatin in a Family of Six Alcohols: Ethanol, 1-Propanol, 1-Butanol, 1-Pentanol, 1-Hexanol, and 1-Octanol. *Int. J Pharm*. (2008) 339: 111-117.

J. Nti-Gyabaah, and Y.C. Chiew. Solubility of Lovastatin in Ethyl Acetate, Butyl Acetate, Acetone, Methanol, Ethanol, and 2-Butanone. *J. Chem and Eng. Data*, 53: 2060-2065.

J. Nti-Gyabaah, V. Chan, and Y.C. Chiew. Solubility of and Limiting Activity Coefficient of Simvastatin in Different Organic Solvents. *Fluid Phase Equilib*. (2009) 280: 35-41.

B. Kosjek, J. Nti-Gyabaah, K. Telari, L. Dunne, and J.C. Moore. Preparative Asymmetric Synthesis of 4,4-dimethoxytetrahydro-2h-pyran-3-ol with a Ketone reductase and *In-situ* Cofactor Recycling Using Glucose Dehydrogenase. *J. Org. Proc. Res. & Dev*., 2008, 12 (4): 584–588.

J. Nti-Gyabaah, F.D. Antia, M.E. Dahlgren and K.E. Göklen. Normal phase high-performance liquid chromatography of pneumocandins; In situ modification of silica with L-proline to separate structural analogs. *J. Biotech Prog.*, (2006) 22, 2: 538-546.

C.J. Welch, J.O. DaSilva, J. Nti-Gyabaah, F.D. Antia, K.E. Göklen, and R.F. Boyd. Preparation and evaluation of novel stationary phases for improved chromatographic purification of pneumocandin B0. J. Chromatogr., 2006, 1101: 204–221.

J. Nti-Gyabaah, F.D. Antia, M.E. Dahlgren and K.E. Göklen. Normal phase high-performance liquid chromatography of pneumocandins; In situ modification of silica with L-proline to separate structural analogs. J. Bioprocess R&D, 2005, Vol. 6, Issue 1: 35-45.

M. Iammarino, J. Nti-Gyabaah, M. Chandler, D. Roush, and K. E. Göklen. Impact of Cell Density and Viability on Primary Clarification of Mammalian Cell Broth. An Analysis Using Disc-Stack Centrifugation and Charged Depth Filtration. BioProc. Int., 2007 (5) 10: 38-50.

- **PATENTS:**

US Patent No: 6,764,842 B2: Enantioselective bioreduction for the preparation of integrin receptor antagonist intermediates.

US Patent No: 7,241,866 B2*:* Using amines or amino acids as mobile phase modifiers in chromatography.

WO2006098959*:* Process for the preparation of (R)-4,4-Dialkoxy-pyran-3-ols such as (R)-4,4-Dimethoxy-pyran-3-ol.

WO2005026323. Stationary phases and a purification process using the stationary phases.

Merck Patent Application # 22461PV: High Performance Liquid Chromatographic (HPLC) Purification of Proteins Using Semi-compressible Resins.