Thermodynamic Modeling for Protein Extraction in Water - PEG 8000 – Dextran 75 System using MATLAB

by

Jazlin Ernida binti Zulkurnain

Dissertation submitted in partial fulfillment of
the requirement for the
Bachelor of Engineering (Hons)
(Chemical Engineering)

JANUARY 2009

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

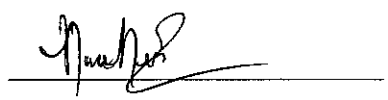
Thermodynamic Modeling for Protein Extraction in Water - PEG 8000 – Dextran 75 System using MATLAB

by

Jazlin Ernida binti Zulkurnain

A project dissertation submitted to the
Chemical Engineering Programme
Universiti Teknologi PETRONAS
in partial fulfillment of the requirement for the
Bachelor of Engineering (Hons)
(CHEMICAL ENGINEERING)

Approved by,

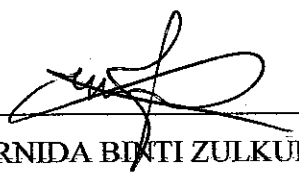


(Dr. Murni Melati Ahmad)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
January 2009

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.



JAZLIN ERNIDA BINTI ZULKURNAIN

ABSTRACT

This report basically discusses the research done and basic understanding of the chosen topic, which is Thermodynamic Modeling for Protein Extraction in Aqueous Polymer System. The objective of the project is to study on extending the application of polymer thermodynamic model, namely Flory-Huggins Theory [1,2], to capture and predict the partitioning behaviour of a sample of different types of proteins in an Aqueous Two-Phase Extraction System (ATPES). ATPES is a liquid-liquid extraction technique that can be applied as a protein separation method. At equilibrium, these systems form water-rich two liquid phases which a target protein selectively partitions into one of the phases according to its ATPES' characteristics. The protein partitioning behavior can be captured using the above mentioned thermodynamic model. Examples of two-phase systems are either two polymers (e.g., PEG/dextran) or one polymer in high-salt concentration (e.g., PEG/salt). The scope for this project will be the protein extraction in aqueous polymer system. To observe the efficiency of ATPES to extract certain types of proteins, a modeling approach by Ahmad [3] based on MATLAB will be used. Using MATLAB software as a helping hand, equilibrium compositions and the partition coefficient of a target protein in the system can be computed. The model will be used to analyze the variations in the partitioning trends and patterns for different types of proteins. The results will be the partitioning of the respective target proteins, either to the top or to the bottom phase in the ATPES, producing partitioning coefficient of greater or less than the value of 1 respectively. In this report, the equilibrium compositions and the partition coefficient of human hemoglobin and lysozyme from chicken egg white in water – PEG 8000 – Dextran 75 system are calculated using the modeling approach by Ahmad [3]. The result obtained is then compared to empirical data [30].

TABLE OF CONTENT

ABSTRACT v

CHAPTER 1 INTRODUCTION

1.1 Background of Study 1

1.2 Problem Statement 9

1.3 Objectives & Scope of Study 10

CHAPTER 2 LITERATURE REVIEW

2.1 ATPES as a Protein Extraction Technique 11

2.2 Modeling of LLE Behavior in ATPES 12

2.3 Protein of Choice for ATPES Study 18

CHAPTER 3 METHODOLOGY

3.1 Modeling using MATLAB 22

3.2 Initial Guess for Modeling Parameter's Values 26

3.2 .1 Human Hemoglobin 26

3.2 .2 Lysozyme from Chicken Egg White 27

CHAPTER 4 RESULTS AND DISCUSSION

4.1 Human Hemoglobin 28

4.1.1 Initial Guess of Interaction Parameters 28

4.1.2 Minimization of Gibbs Energy – Model Objective Function 30

4.1.3 Protein Partitioning Behavior 31

4.1.4 Phase Diagram Formation Patterns 32

4.1.5 Adjustment of Scaling Factors 33

4.2 Lysozyme from Chicken Egg White 36

4.2.1 Initial Guess of Interaction Parameters 36

4.2.2 Minimization of Gibbs Energy – Model Objective Function 37
4.2.3 Protein Partitioning Behavior 39
4.2.4 Phase Diagram Formation Patterns 39

4.3 Errors Deviations in Results 41
4.3.1 Human Hemoglobin 42
4.3.2 Lysozyme from Chicken Egg White 43

CONCLUSION AND RECOMMENDATION 45

REFERENCES 47

APPENDICES 50
FYP 1 Gantt Chart A1
FYP 2 Gantt Chart A2
Calculation of initial guess FH parameters of human hemoglobin used for the modeling and simulation for aqueous polymer mixture system containing Water + PEG-8000 + Dextran75 + Human Hemoglobin [30], interpolated using the data for Ovalbumin and Phosphofructokinase [32] B
Example of calculation of volume fraction from the information of weight fraction C
Example of data file D
Example of main script run file E

LIST OF FIGURES

Figure 1: Aqueous Two Phase Polymer System 3
Figure 2: Phase Diagram of an ATPES 5
Figure 3: A structure of human hemoglobin 20
Figure 4: A structure of lysozyme from chicken egg white 21
Figure 5: Steps in the extended approach of Ahmad [3] on predicting the LLE behavior based on an aqueous polymer system 22
Figure 6: Manual partial enumeration strategy in determining the model parameters used in the ATPES 23
Figure 7: Algorithm of the MATLAB solver 25
Figure 8: Phase diagrams of Dextran75 –PEG-8000 at 23°C 29
Figure 9: Gibss Energy Difference Between One - Phase and Two - Phase For An ATPES of PEG8000 + Dextran 75 + Water + Human Hemoglobin at 23°C 30
Figure 10: Partition coefficients in 4-component ATPES system of water + PEG 8000 + Dextran 75 + human hemoglobin before adjustment of scaling factors 31
Figure 11: Phase diagrams of an ATPES of Water + PEG8000 + Dextran 75 + Human Hemoglobin at 23°C before adjustment of scaling factors 32
Figure 12: Gibbs energy difference between one-phase and two-phase for an ATPES of Water + PEG8000 + Dextran 75 + Human Hemoglobin at 23˚C after adjustment of scaling factors 34
Figure 13: Partition coefficient of all components in 4-component ATPES of Water + PEG8000 + Dextran 75 + Human Hemoglobin at 23˚C after adjustment of scaling factors 34
Figure 14: Phase Diagram for an ATPES of Water + PEG8000 + Dextran 75 + Human Hemoglobin at 23 ˚C after adjustment of scaling factors 35
Figure 15: Phase diagram of ATPES of interest consisting of water + PEG 8000 + Dextran 75 + lysozyme from chicken egg white at 23 37
Figure 16: Gibbs Energy Difference Between one-phase and two-phase for an ATPES of Water + PEG8000 + Dextran 75 + lysozyme from chicken egg white (at 23˚C) 38

Figure 17: Partition coefficients in 4-component ATPES system of water + PEG8000 + Dextran 75 + lysozyme from chicken egg white at 23°C 38
Figure 18: Phase diagrams of an ATPES of water + PEG8000 + Dextran 75 + lysozyme from chicken egg white at 23°C 40
Figure 19: Comparison of phase diagram of water + PEG 8000 + Dx75 + human hemoglobin between empirical tie lines [30] and tie line simulated by modeling approach by Ahmad [3] on common plotting plane 42
Figure 20: Comparison of phase diagram of water + PEG 8000 + Dx75 + lysozyme from chicken egg white between empirical tie lines [30] and tie line simulated by modeling approach by Ahmad [3] on common plotting plane 43

LIST OF TABLES

Table 1: Major Components in Water-PEG6000-DexT500 ATPES 4
Table 2: Final Year Project Time Frame 10
Table 3: ATPES used in the experiment to examine the partition coefficients for a variety of proteins 18
Table 4: Major Properties of Human Hemoglobin [30] 19
Table 5: Major Properties of Lysozyme from Chicken Egg White [30] 20
Table 6: Partition coefficients of the proteins in ATPES of dextran75 – PEG-8000 21
Table 7: Initial guess of FH parameters of chitinase used for the modeling and simulation for aqueous polymer mixture system containing PEG8000+Dextran75+Water+Human Hemoglobin [30], interpolated using the data for Ovalbumin and Phosphofructokinase [32] 27
Table 8: The initial guess for the components' scaling factors and fixed FH parameters used for the modeling for aqueous polymer mixture system containing Human Hemoglobin. 1 = water, 2 = PEG8000, 3 = Dx75, p = Human Hemoglobin 28
Table 9: The scaling factors of each component calculated for LLE calculations using the Gibbs energy of mixing minimization approach for the water + PEG 8000 + Dextran 75 + Human Hemoglobin system 33
Table 10: The initial guess for the components' scaling factors and fixed FH parameters used for the modeling for aqueous polymer mixture system containing lysozyme from chicken egg white. 1 = water, 2 = PEG8000, 3 = Dx75, p = lysozyme from chicken egg white 36
Table 11: Summary of degree of polymerization and Flory Huggins interaction parameters used for human hemoglobin and lysozyme from chicken egg white partitioned in water – PEG 8000 – Dx 75 system at 23°C 46
Table 12: Summary of scaling factors of components in water – PEG 8000 – Dx75 system at 23°C 46

CHAPTER 1
INTRODUCTION

1.1 BACKGROUND OF STUDY

Like other biological macromolecules such as nucleic acids and polysaccharides, proteins are essential parts of organisms and take part in every process within cells. Many proteins are enzymes, vital to metabolism and catalyze biochemical reactions. Proteins have structural or mechanical functions, such as myosin and actin in muscle and the proteins in the cytoskeleton, which form a system of scaffolding that maintains cell shape. Other proteins are important in cell signaling, immune responses, cell adhesion, and the cell cycle. Proteins are also necessary in animals' diets, since animals cannot synthesize all the amino acids they need and must obtain essential amino acids from food. Through the process of digestion, animals break down ingested protein into free amino acids that are then used in metabolism [4].

The production of proteins using a selected host with the necessary posttranslational modifications is one of the key successes in modern biotechnology. This methodology allows the industrial production of proteins that otherwise produced in small quantities. However, the major production costs (50-90%) for a typical biological product resides in the purification strategy [4]. There is a need for efficient, effective and economic large scale bioseparation techniques to achieve high purity and high recovery of the target protein while maintaining the biological activity of the molecule [5].

Chromatography is one of the common methods which can be used in protein separation[11]. It involves passing a mixture dissolved in a "mobile phase" through a stationary phase, which separates the analyte to be measured from other molecules in the mixture and allows it to be isolated. Example of chromatography is affinity chromatography. It is based on selective non-covalent interaction between an analyte and specific molecules. It is very specific, but not very robust. It is often used in biochemistry in the purification of proteins. There are many types chromatography but all these

methods require very high maintenance for the column and rapid turn-over of packaging material.

Besides that, ultrafiltration technique is also used to purify proteins. Ultrafiltration (UF) is a variety of membrane filtration in which hydrostatic pressure forces a liquid against a semi permeable membrane. Suspended solids and solutes of high molecular weight are retained, while water and low molecular weight solutes pass through the membrane. This separation process is used in industry and research for purifying and concentrating macromolecular (10^3 - 10^6 Da) solutions, especially protein solutions.

This research relates to liquid two-phase extraction systems, particularly, to systems for separating proteins using a two-phase protein extraction system. Several low cost two-phase systems are known which can handle protein separations on a large scale. These systems use polyethylene glycol (PEG) as the upper phase-forming polymer and dextran [6], a concentrated salt solution [7] as the lower phase-forming polymer.

The use of aqueous two-phase systems for the purification of proteins is a very powerful technique for the primary downstream processing steps [5]. Aqueous two-phase systems have some advantages in comparison with other commonly used separation and purification techniques.

The main advantages have been summarized by Albertsson (1986) and are given below:

1. Both phases of the system are of aqueous nature, which has high water content (70-85%, w/w), hence provide a suitable environment for the preservation of biological activity.
2. High biocompatibility and low interfacial tension of about 400-fold less than that between water and an immiscible organic solvent allowing small droplet size, large interfacial areas, efficient mixing under very gentle stirring and rapid partition, which facilitates the migration of biomolecules through the interface [8].
3. Rapid mass transfer and mixing until equilibrium requires little energy input.
4. Low probability of denaturation and degradation.

5. Technique facilitates the processing of solid containing streams.
6. Polymers stabilize the proteins.
7. Good resolution and high separation yield. Separation may be achieved in a few minutes, minimizing the harmful action of endogenous proteases.
8. Separation can be made selective.
9. Scale-up from small laboratory experiments is easy and reliable [9].
10. Relatively high capacity.
11. Continuous operation is possible.
12. Technique is cost effective due to low material cost.
13. Possibility of polymer and salt recycling [15].

For these reasons, aqueous two-phase systems have been widely studied on a laboratory scale for the partitioning of many biomolecules.

Aqueous Two Phase System (ATPES)

Many water soluble polymers are incompatible with each other, or with the salt solutions that are high in ionic strength. Thus, if one of such polymer is mixed with another polymer or with a salt solution, aqueous two phase system is formed. Such partitioning systems are employable to separate proteins, from cell debris or to other impurities.

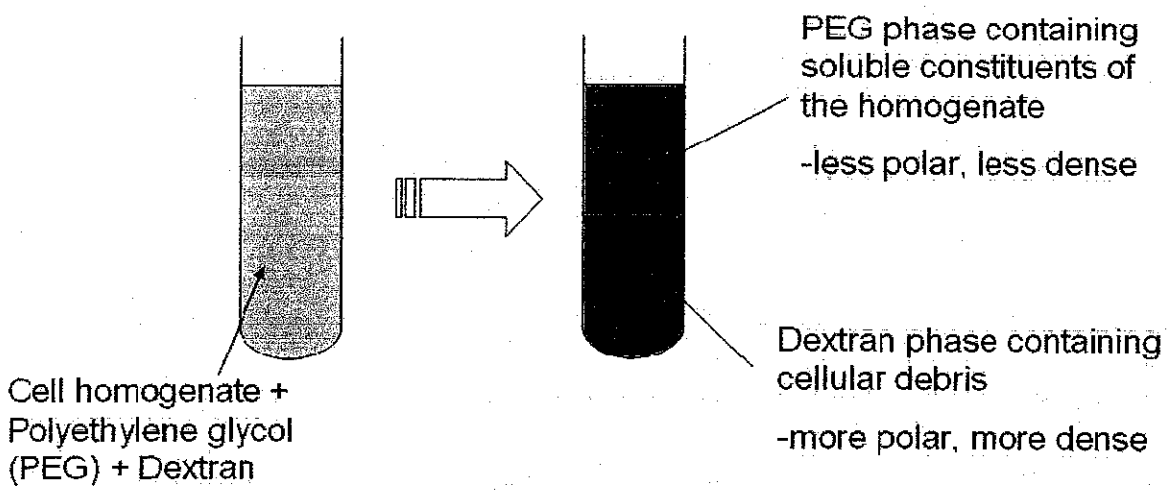


Figure 1: Aqueous two phase polymer system

From Figure 1, the utilization of Polyethylene Glycol (PEG) and Dextran (Dx) results in the cell debris partitioning to the lower, more polar and more dense phase, while the soluble desired proteins tends to partition to the top, less polar, and less dense phase.

This is similar for a polymer-salt system, where the soluble desired proteins with partition into one of the phases. Nonetheless, the focus of this whole study would be on ATPES in comparative analysis for different types of proteins.

In general, aqueous solutions, being polar, are immiscible with non-polar organic solvents (chloroform, toluene, hexane, etc) and form a two-phase system. However, in an ATPES, both immiscible polymers are water-based.

ATPES differs from the traditional liquid-liquid extraction system as it is not performed in the pre-existing immiscible organic and aqueous phases, but two-water soluble and at the same time mutually incompatible materials. Example of a typical ATPES which comprises of two-polymers in aqueous solution is a water – Polyethylene Glycol (PEG) – Dextran (Dx) system as shown in Table 1 below.

Table 1: Major Components in Water – PEG 6000 – Dx 500 ATPES [3]

System	Water-PEG-Dex
Temperature	20°C
Overall Composition (volume fraction)	86 vol% water, 6 vol% PEG, 8 vol% Dx

The formation of the distinct phases is affected by the pH, temperature and ionic strength of the two polymers, and separation occurs when the amount of a polymer present exceeds a certain limiting concentration.

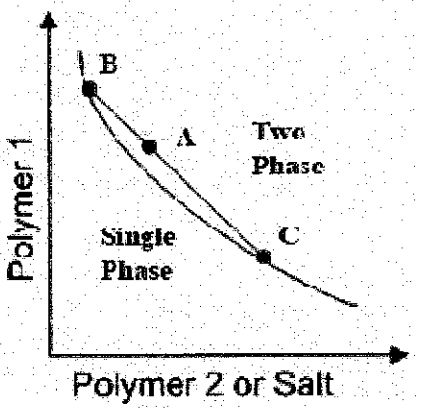


Figure 2: Phase Diagram of an ATPES

Figure 2 shows a phase diagram of a polymer-polymer system. At low concentrations, the mixtures form a single phase, but separation into two incompatible phases will take place at higher concentrations. For instance, in the case of a mixture at composition A, the separation of the two-phase region will result in two different phases which compositions are given by points B and C of the tie line passing through A. The ratio of the line segments *AB* and *AC* can be utilized in calculating the relative volumes of these two phases.

As the protein partition between the two aqueous phases, this characteristic can be represented by partition coefficient, K, which is the ratio of concentrations of the target molecule (protein) in the top PEG-rich phase and bottom dextran-rich phases, C_T/C_B.

$$K = \frac{C_T}{C_B} \qquad (1)$$

where C_t and C_b represent the concentrations in the top and bottom phases respectively. The K-value for each solute was determined as the slope of the concentration in the upper phase plotted as a function of the concentration in the bottom phase. If value of K is greater than 1, the target protein is concentrated in the top phase, and on the other hand, if the value of K is less than 1, the target protein partitioned into the bottom phase. The yield and efficiency of the separation is determined by the relative amounts of material in

the two phases and therefore depends on the volume ratio (V_t/V_b). The partition coefficient is exponentially related to the surface area, molecular weight and surface charge of the particles in addition to the difference in the electrical potential and hydrophobicity of the phases. Proteins and larger particles are normally partitioned into one phase whereas smaller molecules are distributed more evenly between phases. Typical partition coefficients for proteins are 0.01-100 whereas the partition coefficients for cells and cell debris are effectively zero [10].

Both the phase separation and protein partitioning behaviour are based on the equilibrium behavior of the whole system. This liquid-liquid equilibrium behaviour can be captured and predicted using the thermodynamic models. Different thermodynamic model has its own advantages and disadvantages.

There are several models for liquid-liquid equilibrium behaviour prediction. Some, such as Non-Random Two-Liquid (NRTL) Model, use the local composition concept; and others, such as UNIQUAC, have a more theoretical basis; and finally, UNIFAC use the group contribution method.

Non-Random Two-Liquid (NRTL) Model

The NRTL model is based on the local composition, and it is applicable to partially miscible systems. The concept is based on the hypothesis that the local concentration around a molecule will be different from the bulk concentration when there is a difference between the interaction energy of the central molecule with the molecules of its own kind and that with the molecules of the other kind. This difference introduces a non-randomness at the molecular level [27].

The NRTL model is an activity coefficient model that correlates the activity coefficients with the composition of a mixture of chemical compounds, expressed by mole fractions. Mole fractions have been traditionally used in this model, but they are not suitable for polymeric systems because the mole fraction of a polymer, due its large molecular mass, is an extremely small quantity.

Instead, mass fraction can be used for calculation of the activity coefficient of a solvent in polymeric solutions with the UNIQUAC and the UNIFAC methods.

Universal Quasi-Chemical (UNIQUAC) Model

UNIQUAC model is an activity coefficient model in which the activity coefficients of the components in a chemical mixture can be related through their molar fraction. The model is a lattice model and this one has been be derived from a first order approximation of interacting molecule surfaces in statistical thermodynamics.

Although UNIQUAC is mathematically more complex than NRTL it has following advantages [23, 24]:

1. It contains only two adjustable parameters instead of three parameters in NRTL.
2. The parameters are less dependent on temperature.
3. UNIQUAC is applicable to solutions containing small as well as large molecules.

However, UNIQUAC requires two basic underlying parameters that make the calculation of activity coefficient tedious and slow, which are:

1. Relative surface and volume fractions are chemical constants, which must be known for all chemicals.
2. An empirical parameter between components that describes the intermolecular behaviour. This parameter must be known for all binary pairs in the mixture, and the number rapidly increases with additional chemical components. The empirical parameters are derived from experimental activity coefficients, or from phase diagrams, from which the activity coefficients themselves can be calculated [2,24].

An alternative is to obtain activity coefficients with a method such as UNIFAC, and the UNIFAC parameters can then be simplified by fitting to obtain the UNIQUAC parameters [23, 24]. This method allows for the more rapid calculation of activity coefficients, rather than direct usage of the more complex method.

UNIversal Functional Activity Coefficient (UNIFAC) Model

UNIFAC Model is a semi-empirical system for the prediction of non-electrolyte activity estimation in non-ideal mixtures [26]. It has several advantages over other NRTL and UNIQUAC model, which are:

1. Size and binary interaction parameters are available for wide range of types of functional groups.
2. An open system, and more functional groups and more parameters can be used.
3. Shortening the measurement time of liquid-liquid equilibrium behaviour in multiphase system.

UNIFAC uses the functional groups present on the molecules that make up the liquid mixture to calculate activity coefficients. By utilising interactions for each of the functional groups present on the molecules, as well as some binary interaction coefficients, the activity of each of the solutions can be calculated. Besides that, by using UNIFAC, it is gives more flexibility to vary liquid phases compositions [26].

Flory-Huggins Theory

Flory-Huggins theory is a mathematical model of the thermodynamics of polymer solutions which takes account of the great dissimilarity in molecular sizes in adapting the usual expression for the entropy of mixing. The result is an equation for the Gibbs free energy change ΔG_m for mixing a polymer with a solvent. Although it makes simplifying assumptions, it generates useful results for interpreting experiments.

Aqueous two phase systems are of considerable value to biotechnology. ATPES would therefore be the ideal answer to protein purification as the aqueous nature in ATPES offers the ideal solubility solution for proteins. Overall, ATPES are particularly useful as the two phases can be constituted to span for the relatively small, yet sensitive differences in desired protein.

1.2 PROBLEM STATEMENT

Aqueous Two-phase Extraction System (ATPES) is basically a liquid-liquid extraction technique that can be applied as a protein separation method. When the system achieves equilibrium, water-rich two liquid phases will be formed and the target protein will selectively partitions into either phases based on the behavior of its own and that of the ATPES [3]. All these results could be achieved via full-scale laboratory work in the laboratory. However, if a particular protein separation were to be performed manually in a laboratory scale, it is exhaustive to explore for many possible combination of polymers to find the suitable ATPES. Different proteins partition differently in different ATPES. Moreover, the materials to extract a target protein which are pharmaceutical grade polymers/salt and proteins required are expensive, and the whole process itself being extremely time consuming.

Being both a cost-effective and faster alternative, modeling is a good way to represent and predict system behaviour and performance. Systematic computational approach is needed to model the LLE behavior and to predict partitioning behavior. The protein partitioning behavior can be captured using a thermodynamic model for polymer systems, namely Flory-Huggins Theory [1,2].

The advantages of Flory-Huggins Theory are [13]:

1. It is simple with only a limited number of parameters required.
2. It provides good mathematical insights into the phase formation phenomena.
3. It enables qualitative prediction or correlation for measured phase compositions, and partitioning behaviors.

This study focuses on extending the application [3] of a polymer thermodynamic approach by Ahmad [3], which is based on Flory-Huggins Theory [1,2], to capture and predict the LLE behaviour of the ATPES and protein partitioning in ATPES. At the same time, the applicability of the approach by Ahmad [3] is tested using the empirical data reported by Pedro et al [30].

1.3 OBJECTIVE AND SCOPE OF STUDY

This study focuses on extending the modeling approach proposed by Ahmad [3] based on a polymer thermodynamic model, namely Flory-Huggins Theory [1,2], to capture and predict the partitioning behavior of a sample of different types of proteins in an ATPES.

The scope of this project is:

1. Familiarize with the concepts and theories behind the Aqueous Two-phase Extraction System (ATPES) as a liquid-liquid extraction technique.
2. Master MATLAB applications.
3. Study all the thermodynamics concepts incorporated in the modeling approach on ATPES by Ahmad [3] using MATLAB.
4. To perform literature research on suitable polymer system and application of study.
5. To calculate LLE data using modeling approach of different ATPES for different polymer combinations.
6. To present complete analysis on the trend or effect of different polymers on the LLE behavior.

The total time period allocated for this Final Year Project (FYP) would be as shown in Table 2 below.

Table 2: Final Year Project Time Frame

Semester	Duration	Total Credit Hour
July 2008	14 weeks	2
January 2009	14 weeks	4

CHAPTER 2
LITERATURE REVIEW

2.1 ATPES as a Protein Extraction Technique

Liquid-liquid extraction utilizing aqueous two-phase systems are formed spontaneously upon mixing two aqueous solutions of structurally different polymers, above a certain critical concentration [8]. This phenomenon was first observed by Beijernick, in 1896, where in this case two phases were formed when agar was mixed with soluble starch or gelatin. However it was not until 1956 that the potential use of these systems as an important separation technique in biotechnology was realized [11]. When two water-soluble polymers or a polymer and a strong electrolyte were dissolved in water, a two-phase system is formed. The work of Albertsson in the mid 1950s established that biological solutes added to theses systems tend to partition unevenly between the phases. Since then, many two phase aqueous systems have been found; the most thoroughly investigated being the aqueous dextran-polyethylene glycol system (e.g. 10% polyethylene glycol 4000/2% dextran T500), where dextran forms the more hydrophilic, denser, lower phase and polyethylene glycol the more hydrophobic, less dense, upper phase.

In recent years, the separation of proteins and other biomolecules were studies by many researchers [12]. Since then, aqueous two-phase extraction has been successfully used for the purification of biomolecules, such as proteins, nucleic acids and peptides, as well as for the fractionation of biological membranes, cell organelles and cells [8]. For instance, Yang et al. investigated Cephalosporin C partitioning in ATPS [13]. Marcos et al. worked on the purification of Penicillin acylase by these systems [14]. According to the results of these investigations, extraction by aqueous two phase system is a powerful technique and allows the integration of clarification, separation, concentration, and purification of biomolecules and pharmaceutical products [15, 16].

Aqueous Two-Phase Extraction System (ATPES) is a liquid-liquid extraction technique that can be applied as a protein separation method. At equilibrium, these systems form water-rich two liquid phases which a target protein selectively partitions into one of the phases according to its and the ATPES' characteristics. Since both bulk phases consist mainly of water (usually more than 85% w/w in water), have similar densities and low interfacial tensions [8], and most polymers have a stabilising effect on the protein tertiary structure [17], these systems form a gentle and mild environment for proteins. Furthermore, this process is cost-effective and its scale-up is very easy and reliable [18]. ATPES represents wholly aqueous systems that are safe, nontoxic, and nonflammable, and thus, they represent relatively environmentally benign extraction media. Such systems could be employed as alternatives to traditional aqueous-organic systems for the separation of, inter alia, small organic molecules. Liquid-liquid extraction of ATPS is considered to be environmentally friendly due to no use of traditional volatile organic solvents (VOCs) in the whole process and therefore has already been used to separate and purify various biological products from the complex mixtures in which they are produced.

2.2 Modeling of LLE Behaviour in ATPES

When a system reaches equilibrium, the energy content is at the lowest possible state. This applies to an ATPES, where the equilibrium state attained would enable it to partition into two distinguishable phases, namely the top and the bottom phase, with respective components in each phase achieving the following:

1. Minimize in the difference of Gibbs energy of mixing (min $\Delta G^{two-phase}$) [3, 32]
2. Equal chemical potentials $\left(\mu_{component\ i}^{top-phase} = \mu_{component\ i}^{bottom-phase}, where\ i = 1,2,3,\dots,n\right)$
3. Equal fugacity $\left(f_{component\ i}^{top-phase} = f_{component\ i}^{bottom-phase}, where\ i = 1,2,3,\dots,n\right)$

A suitable modeling approach has to be adopted to calculate the composition of each component in each phase at equilibrium. Flory-Huggins theory is able to capture and calculate for the minimization of the Gibbs energy of mixing level of the components at equilibrium in the ATPES. This modeling approach, based on Gibbs free energy is chosen because:

1. It is simple with only a few parameters required.
2. It provides good mathematical insights into the phase formation phenomena.
3. It enables qualitative prediction or correlation for measured phase compositions, and partitioning behaviors.

A non-linear programming (NLP) formulation in minimizing the Gibbs energy of mixing difference ($\Delta G^{two\text{-}phase}$ - $\Delta G^{one\text{-}phase}$) is developed by Ahmad [3].

The total Gibbs energy of mixing for system k, ΔG^{k}_{mix} or ΔG is given by the product $N^k.\overline{\Delta G}^k$. The total Gibbs energy of mixing for a two-phase system is given by:

$$\Delta G^{two-phase} = N^{top-phase}.\overline{\Delta G}^{top-phase} + N^{bottom-phase}.\overline{\Delta G}^{bottom-phase} \qquad (2)$$

N^k represents the total moles of lattice in phase k and is calculated using Equation (3).

$$N^k = \sum_i^m n_i^k . r_i \qquad (3)$$

where : n_i^k - moles of component i in phase k

The volume fraction of component i in phase k is calculated by using Equation (4).

$$\emptyset_i^k = \frac{n_i^k . r_i}{N^k} \qquad (4)$$

At phase equilibrium, the chemical potential of each component in the phase is equal and similarly the changes in chemical potential on mixing are equal,

$$\Delta\mu_i^{top-phase} = \Delta\mu_i^{bottom-phase} \qquad (5)$$

Flory [1] and Huggins [2] suggested lattice theory where a liquid mixture is modeled to be a solid-like as quasi-crystalline lattice [27].

One of the theoretical treatments of protein partitioning in aqueous two phase systems is based on the Flory-Huggins theory.[28] Flory-Huggins theory is a mathematical model of the thermodynamics of polymer solutions which takes account of the great dissimilarity in molecular sizes in adapting the usual expression for the entropy of mixing. The result is an equation for the Gibbs free energy change ΔG_m for mixing a polymer with a solvent. Although it makes simplifying assumptions, it generates useful results for interpreting experiments. The thermodynamic equation for the Gibbs free energy change accompanying mixing at constant temperature and (external) pressure is

$$\Delta G_m = \Delta H_m - T\Delta S_m \qquad (6)$$

The objective is to find explicit formulas for ΔH_m and ΔS_m, the enthalpy and entropy increments associated with the mixing process.

Based on lattice theory, on a molar basis, for an aqueous system containing n types of polymer, Flory-Huggins expressed the Gibbs energy of mixing as

$$\overline{\Delta G}^k = \sum_i \sum_{j \neq i} \omega_{ij} \emptyset_i^k \emptyset_j^k + RT \sum_i \frac{\emptyset_i^k}{r_i} ln \emptyset_i^k \qquad (7)$$

where:

k – top phase or bottom phase

$\emptyset_i^k$ - volume fraction of component i in phase k

r_i – degree of polymerization of component i

ω_{ij} – effective pair-wise interchange energy between component i and j

R – gas constant

T – absolute temperature

Therefore, based on Equation (7), the chemical potential change on mixing for each component in the phase is given by:

$$\Delta\mu_i^k = N^k\left(-r_i\sum_{j=1(j\neq i)}^{m-1}\sum_{l=j+1(l\neq i)}^{m}\omega_{j,l}\emptyset_j^k\,\emptyset_l^k + r_i\sum_{j=1(j\neq i)}^{m}\omega_{i,j}\emptyset_j^k\,(1-\emptyset_i^k)\right) + N^k RT\left(ln\emptyset_i^k + 1 - r_i\sum_{j=1}\frac{\emptyset_j^k}{r_i}\right) \quad (8)$$

The LLE compositions are the $\emptyset_i^{top-phase}$ and $\emptyset_i^{bottom-phase}$ that give the minimum $\Delta G^{two-phase}$ or the maximum $(\Delta G^{one-phase} - \Delta G^{two-phase})$, or equivalent chemical potential changes on mixing between the phases $(\Delta\mu^{top-phase} = \Delta\mu^{bottom-phase})$.

The m components is accounted simultaneously by using a vector of design variables with the size of 1 x m. Mole fractions of the components in the bottom phase, f_i, is the design variable and the total moles in the system is pre-specified as N. The relationship between f_i and N is defined by Equation (9) and (10).

$$N = N^{top-phase} + N^{bottom-phase} \quad (9)$$

$$N^{top-phase} = (1 - f_i).N \quad (10)$$

where,

$N^{top-phase}$ = total moles in top phase

$N^{bottom-phase}$ = total moles in bottom phase

The total moles in respective phases are defined as $N^{top-phase}$ and $N^{bottom-phase}$. The mole of component i in each phase is calculated using Equation (3) and using Equation (4) to calculate the volume fraction of component i. The NLP is formulated [3] as:

$$min(\Delta G^{two-phase} - \Delta G^{one-phase}) + \sum_i w_i\left(\frac{\Delta\mu_i^{top-phase} - \Delta\mu_i^{bottom-phase}}{\Delta\mu_i^{top-phase}}\right)^2 \quad (11)$$

$$\left(\frac{\Delta\mu_i^{top-phase}\Delta\mu_i^{bottom-phase}}{\Delta\mu_i^{top-phase}}\right)^2 \leq tol_i \quad \forall i = 1, \ldots, n \quad (12)$$

$$\Delta G^{two-phase} < \Delta G^{one-phase} \quad (13)$$

$$\emptyset_2^{top-phase} - \emptyset_2^{bottom-phase} > tol_2 \quad (14)$$

$$\emptyset_3^{bottom-phase} - \emptyset_3^{top-phase} > tol_3 \quad (15)$$

$$0 \leq \emptyset_i^k \leq 0.3 \quad \forall i = 1, \ldots, n \quad k = top - phase, bottom\ phase \quad (16)$$

where :

μ_i^k - change in the chemical potential on mixing for component *i* in phase k

w_i - static weighing factor for the chemical potential difference between the phases for component *i*.

As shown in Equation 11 above, the chemical potential difference terms are added into the right hand side as "penalty term" to the original objective function on the left. This is due to the Gibbs energy of mixing difference surface alone being too flat, hampering the MATLAB written model in looking for the global minimum [3].

The most important factor for phase separation is the chemical nature of both polymers. The partitioning behavior of a biomolecule depends on the composition of the two phases in equilibrium. When the composition of the two phases in equilibrium changes, the partition coefficient changes accordingly – for instance, usually the partitioning becomes more one sided when the length of the tie-line increases [19].

In ATPES, the phase separation is due to small repulsive interactions between the two types of monomers in the solution. The total interaction between the two polymers is large (high energy) because each one is composed of several monomers [5]. Hence, for an ATPES with the composition lying in the two phase possible region, the immiscible polymers prefer to separate from each other, arriving to an energetically more favorable lower energy state, thus achieving equilibrium.

The thermodynamics of polymeric solutions has been studied extensively. Among the earlier pioneered modeling approaches, each has its very own list of advantages and disadvantages. Edmond and Ogston [20] modeled nonidealities with a truncated osmotic

virial expansion based on McMillan – Mayer theory [20]. Local composition models UNIversal-QUAsi-Chemical (UNIQUAC), UNIFAC, NRTL and NRF have also been used to describe the thermodynamics of polymer solutions. Kang and Sandler [22,23] and Hartounian et al. [24] used the UNIQUAC solution model to deal with the phase behavior of polymer–polymer aqueous two-phase systems. The UNIFAC model [25] was extended to polymer solutions by Oishi and Prausnitz [26].

Besides that, a model representing a synergistic combination of the Flory-Huggins description for the configurational entropy of mixing molecules of different sizes and the NRTL or NRF theory for the local composition contribution from mixing solvents and segments of polymers is developed by Zafarani-Moattar and Sadeghi [27]. This model was applied for the correlation of the phase behavior of some aqueous two phase polymer systems and the results show that the model can accurately predict the liquid-liquid equilibrium behaviour.

Brooks [29] developed the first lattice model for solute partitioning applying the Flory theory of polymer-solvent mixing to aqueous two-phase systems, treating the solute being partitioned as a third polymer component [28]. The expression for the solute partition coefficient was derived from the equality of the solute chemical potentials in the two phases using only first order term in the polymer concentration differences between the phases [29]. According to this expression, the solute partition behavior is governed by the molecular volumes of the components of the system and the Flory interaction parameters χ describing the solute interactions with the solvent and each phase polymer. It shows that the treatment used predicts that the protein partitioning should be more one-sided with increasing protein molecular weight or increasing total polymer concentrations in the system and the protein partitioning should increase to the phase rich in the polymer with decreased molecular weight.

2.3 Protein of Choice for ATPES Study

The types of proteins to be used in this study would be selected based on the following criteria:

i. Accessibility to the information on the protein properties/data.
ii. Suitability of the protein type.
iii. Availability of the protein in the market.
iv. Industrial significance of the particular protein.

Pedro et al. [30] reported the partitioning coefficients for a variety of proteins which were measured in ATPES. The ATPES used in this experiment is as shown in Table 3 below.

Table 3: ATPES used in the experiment to examine the partition coefficients for a variety of proteins [30]

System	Water-PEG-Dex
Major components	Water-PEG8000-Dex75
Temperature	23°C
Overall Composition (volume fraction)	86 vol% water, 6 vol% PEG, 8 vol% Dex

In ATPES described above the distribution coefficients for the proteins were correlated according to the Collander linear equation, which is a solvent regression equation [30], given in Equation 17.

$$\ln K_i = a_{ij} ln K_j + b_{ij} \qquad (17)$$

where K_i and K_j are partition coefficients for any protein in the i-th and j-th two-phase systems. Coefficients a_{ij} and b_{ij} are constants, the values of which depend upon the particular compositions of the two-phase systems under comparison.

Pedro et al. [30] concluded that it is still early at this stage to draw any general conclusion in regard to the existence of a linear correlation between the logarithms of the partition coefficients of different proteins in ATPES since there are very few studies of the same proteins partitioned in ATPES formed by polymers of different chemical nature, and even then the purity and other characteristics of the proteins preparations used by different authors may differ significantly.

ATPES is highly suited to macromolecular polyelectrolytes such as proteins and enzymes as they can be optimized for partition on the basis of charge, hydrophobicity, and molecular weight. It has also been discovered by previous studies that proteins with molecular weight of 20 – 100 kDa, would partition more effectively in ATPES [31]. Hence, the first two proteins to be chosen in carrying-out this study would be preferably the ones with the molecular weight that are within the above mentioned range.

The identified target proteins to be used in this study would come from two particular protein-type groups, namely the hemoglobin and enzymes.

The main reason for this selection is due to the availability of proven empirical performance data on the protein partitioning in an ATPES in literature [30]. Also, as the empirical nature of the ATPES in this paper matches to that of our model, it therefore facilitates the verification our modeling outcome by comparing it to the experimental results.

The first target protein selected is human hemoglobin. Some of the useful properties of hemoglobin for this modeling approach are as shown below:

Table 4: Major Properties of Human Hemoglobin [30]

Properties	Values
Molecular Weight	68000 Da
Isoelectric	6.8 pI

Hemoglobin is the iron-containing oxygen-transport metalloprotein in the red blood cells of vertebrates. In mammals, the protein makes up about 97% of the red cell's dry content, and around 35% of the total content. Hemoglobin is responsible for red blood cells' red color. Its normal concentration in erythrocytes is 34%. Hemoglobin is the most important respiratory protein of vertebrates by virtue of its ability to transport oxygen from the lungs to body tissues, and to facilitate the return transport of carbon dioxide. Anomalous globins in which various amino acids have been substituted with others, or in which certain amino acids are missing entirely from the normal sequence, comprise 153 abnormal hemoglobin species. Some of these are responsible for diseases, the most common of which is sickle cell anemia.

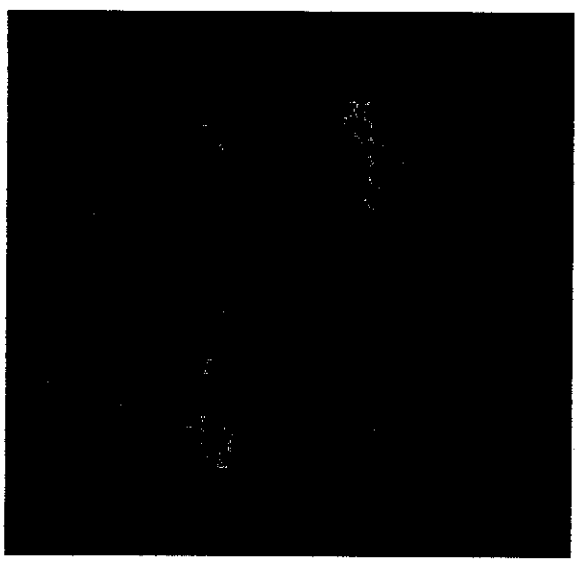


Figure 3: A structure of human hemoglobin

The second target protein selected is one from the enzyme class, i.e. chicken egg lysozyme. Some of the useful properties of lysozyme for this modeling approach are as shown below:

Table 5: Major Properties of Lysozyme from Chicken Egg White [30]

Properties	Values
Molecular Weight	14307 Da
Isoelectric	11.35 pI

In viruses (or bacteriophages), lysozyme is used as an agent to break into the host bacterial cell. Lysozyme from the tail of the virus (or bacteriophage) destroys the peptidoglycan bacterial cell wall and then virus can injects its DNA. After multiplication in bacteria, many lysozyme molecules are created to lyse the bacterial cell wall and release new viruses.

Figure 4: A structure of lysozyme from chicken egg white

Table 6 listed the partition coefficients for the target proteins examined in the present work in ATPES by Pedro et al [30].

Table 6 : Partition coefficients of the proteins in the aqueous two-phase systems Dextran 75 – PEG 8000 [30]

Protein	Partition coefficient, K
Human hemoglobin	0.131
Lysozyme	2.36

From the value of partition coefficients, it shows that human hemoglobin is partitioned into Dextran 75 and lysozyme from chicken egg white is partitioned into PEG 8000.

CHAPTER 3
METHODOLOGY

3.1 Modeling Using MATLAB

MATLAB is used as a modeling simulation language to calculate the prediction of the LLE behavior in an ATPES system. MATLAB is a numerical computing environment and programming language. Steps to be taken for this project are given in Figure 6:

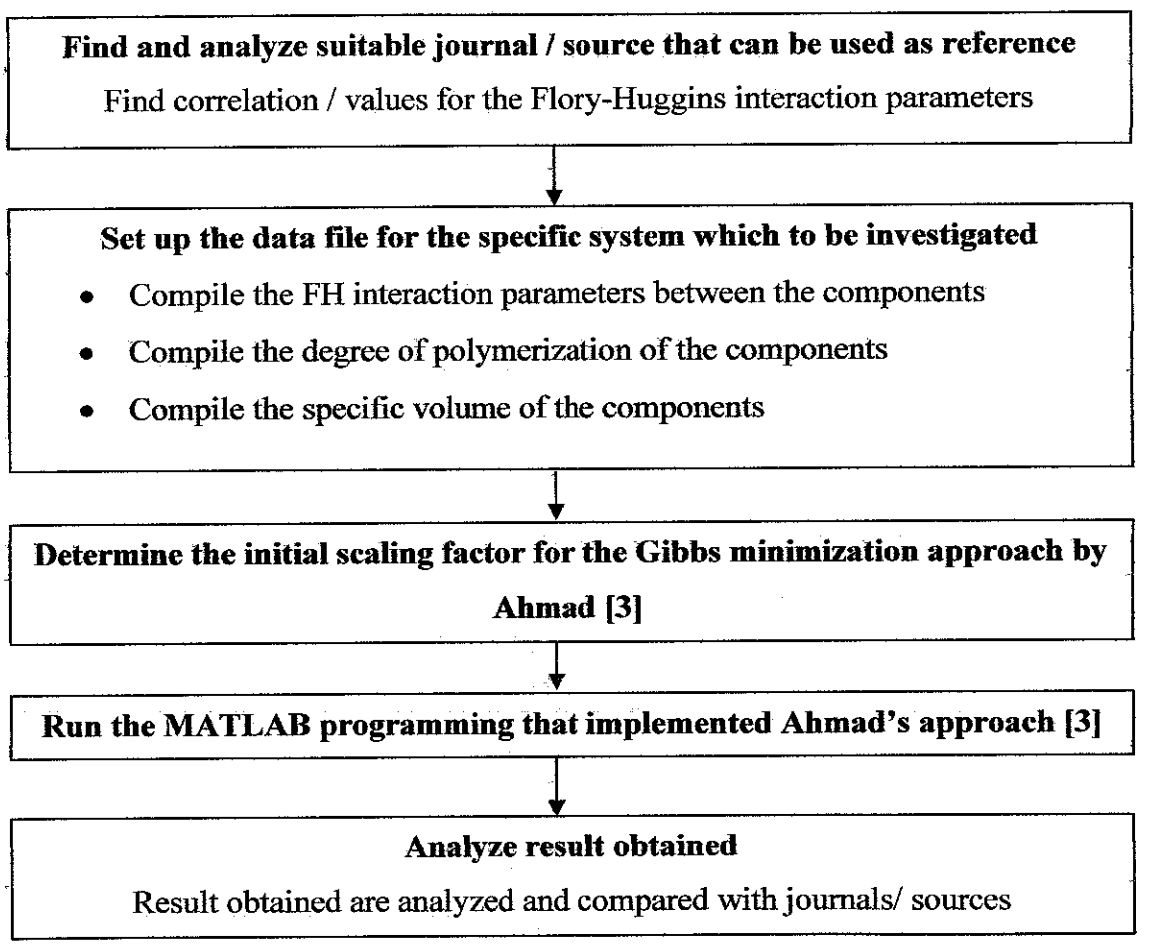


Figure 5: Steps in extending the approach by Ahmad [3] on predicting the LLE behavior based on an aqueous polymer system.

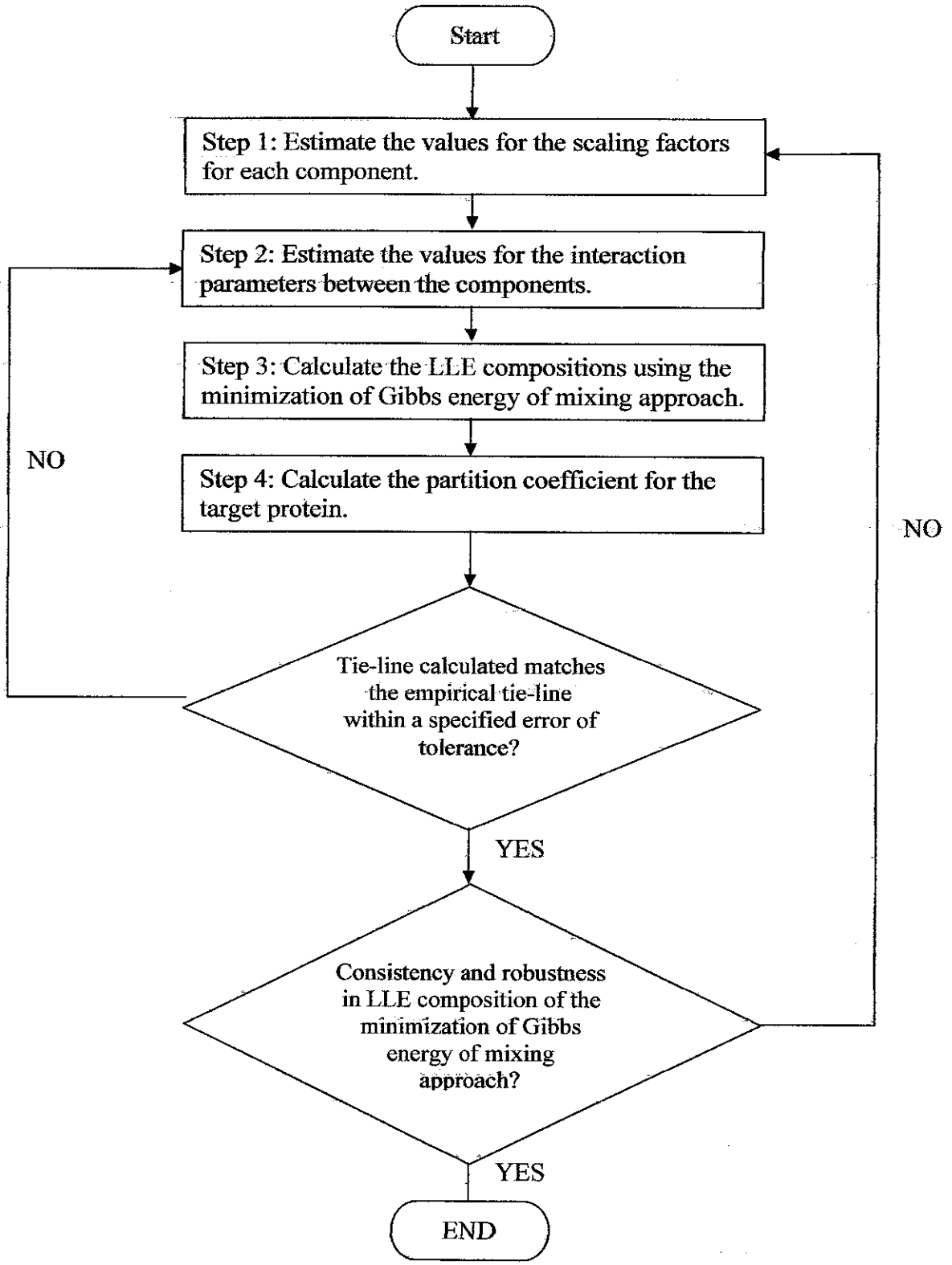


Figure 6: Manual partial enumeration strategy in determining the model parameters used in the ATPES modeling.

Figure 5 shows the steps in extending the approach of Ahmad [3] on predicting the LLE behaviour based on aqueous two phase polymer system. The first step is to find and analyze suitable journals or sources that can be used as reference. The correlation or value for the Flory-Huggins (FH) interaction parameters is determined. Then, the data file for the system is set up. The FH interaction parameters between the components, the degree of polymerization of the components and the specific volume of the components is declared. Next, the main script file for the system is set up where the system name, type, number of component, type of component and temperature is specified. The initial points for the solver and composition of the mixture are also specified.

Figure 6 shows the manual partial enumeration strategy in determining the model parameters used in the ATPES modeling. It explains the calculation steps that are used in extending the approach by Ahmad [3]. First, the scaling factors values are estimated for each component in the particular ATPES system in study. This is used as a starting point to determine the correct scaling factor and interaction parameters for each component for an ATPES system. Next, the interaction parameters values between the components are estimated. By using the value that been estimated, the LLE compositions and the partition coefficient for the target polymer system are calculated. The tie-line produced by modeling approach is compared with the empirical tie-line whether or not it is within a specified error tolerance.

Hence, the applicability of minimization of Gibbs energy of mixing approach in calculation of the LLE behavior of the ATPES and protein partitioning behavior by Ahmad [3] is tested using the empirical data from Pedro et al [30]. The model parameters used for the four components in the aqueous polymer mixture of PEG8000 ($M_W \approx 8000$) + Dextran 75 ($M_W \approx 75000$) + water + protein (at 23°C), including the scaling factor for the objective function of the model are to be predetermined via a manual partial enumeration strategy as shown in Figure 6.

Being one of the main model parameters, the FH interaction parameter between components i and j, ω_{ij} (see Equation 7) plays an important role in modeling the overall

performance of the protein partitioning behaviour. With the values for the interaction parameter of proteins with each of the respective polymer component acting as the manipulated variables, the following variables are set to be constant:

i. Degree of polymerization of protein
ii. Specific volume of protein
iii. Temperature of system
v. Molecular weight of protein
vi. Initial guess of overall composition for ATPES system of PEG8000 + Dextran 75 + water + protein
vii. Interaction parameter of water-protein

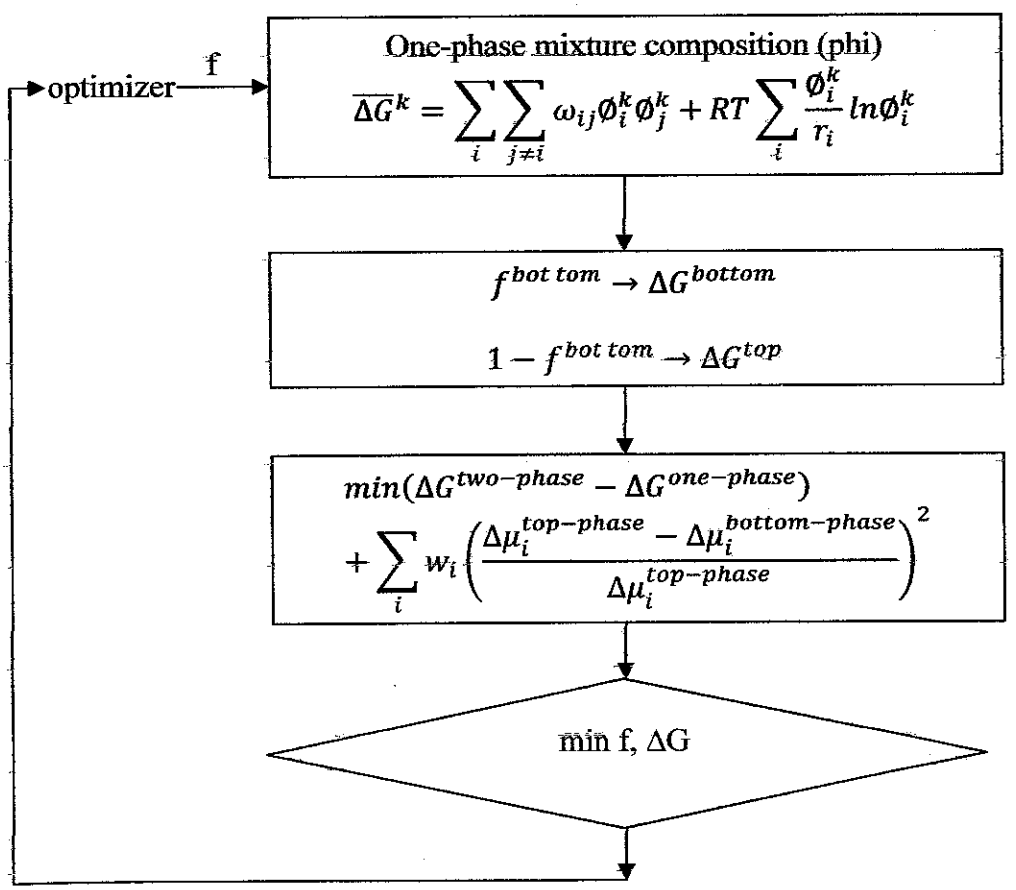


Figure 7: Algorithm of the MATLAB solver [3].

Figure 7 shows how the MATLAB coding is written in order to solve for the minimum ΔG. By providing the initial value of f, which is the mole fraction of each component in bottom phase compared to top phase, the solver will solve by running iteratively until the value of f is optimized.

3.2 Initial Guess for Modeling Parameters' Values

3.2.1 Human Hemoglobin

In varying the manipulated variable of interaction parameters of human hemoglobin with each of the respective polymer components, initial guesses would be important in determining the later suitable values to arrive at the desired outcome of the protein partitioning. The first set of the initial guess values directly interpolated from two readily available reference proteins, i.e. ovalbumin and phosphofructokinase as shown in Table 3 [37] is utilized before we embark on the manual partial enumeration strategy. Calculation is shown in Appendix B.

Table 7 : Initial guess FH parameters of human hemoglobin used for the modeling and simulation for aqueous polymer mixture system containing PEG 8000 + Dextran 75 + Water + Human Hemoglobin [30], interpolated using the data for Ovalbumin and Phosphofructokinase [32].

Protein	Ovalbumin	Human hemoglobin	Phosphofructokinase
MW (Da)	43000	68000	85000
r.prot	100	2720	4500
omega.water_prot	0	0	0
omega.PEG_prot	0	89.3	150
omega.Dx_prot	100	40.5	0
omega.prot_prot	0	0	0

3.2.2 Lysozyme from chicken egg white

Table 8 : Initial guess FH parameters of human hemoglobin used for the modeling and simulation for aqueous polymer mixture system containing PEG 8000 + Dextran 75 + Water + lysozyme from chicken egg white [30], interpolated using the data for ovalbumin and phosphofructokinase [32].

Protein	Lysozyme	Ovalbumin	Phosphofructokinase
MW (Da)	14307	43000	85000
r.prot	10	100	4500
omega.water_prot	0	0	0
omega.PEG_prot	0	0	150
omega.Dx_prot	150	100	0
omega.prot_prot	0	0	0

Once the results are obtained, scaling factors of the original objective function is to be fine-tuned to make sure the tie-lines produced are parallel and constant with each other and comparable with the experimental results [30]. The consistency of the tie lines indicates the robustness of our model in modeling ATPES within the desired composition range.

Chapter 4
RESULTS AND DISCUSSION

4.1 Human Hemoglobin

Analysis is carried out on the partitioning of human hemoglobin in an ATPES consisting of Water – PEG 8000 – Dextran 75 – human hemoglobin. According to the manual partial enumeration strategy (See Figure 6) proposed earlier, the initial steps involve the scaling factor and Flory-Huggins (FH) interaction parameters determination for the respective components.

4.1.1 Initial Guess of Interaction Parameters

The initial guesses are calculated based on the approach proposed in section 3.1, on the simulation parameters used in the modeling for ATPES of Water – PEG 8000 – Dextran 75 - human hemoglobin are as shown in Table 8.

Table 8: The initial guesses for the components' scaling factors and fixed FH parameters used for the modeling for aqueous polymer mixture system containing human hemoglobin. Subscripts: 1 = water, 2 = PEG 8000, 3 = Dextran 75, p = human hemoglobin

System	Water – PEG8000-Dx75
Temperature	23°C
Overall composition (volume fraction)	86 vol% water, 5 vol% PEG, 8 vol% Dx, 1 vol% human hemoglobin
Scaling factors	Water: 1, PEG8000: 1e-1, Dx75: 1e-2, Human Hemoglobin: 1e-1
Degree of polymerization	$r_1 = 1$, $r_2 = 100$, $r_3 = 4500$, $r_{human\ hemoglobin} = 2720$
Constant Flory Huggins interaction parameters (J/mol)	$\omega_{12} = 100$, $\omega_{13} = 0$, $\omega_{23} = 345$, $\omega_{1p} = 0$

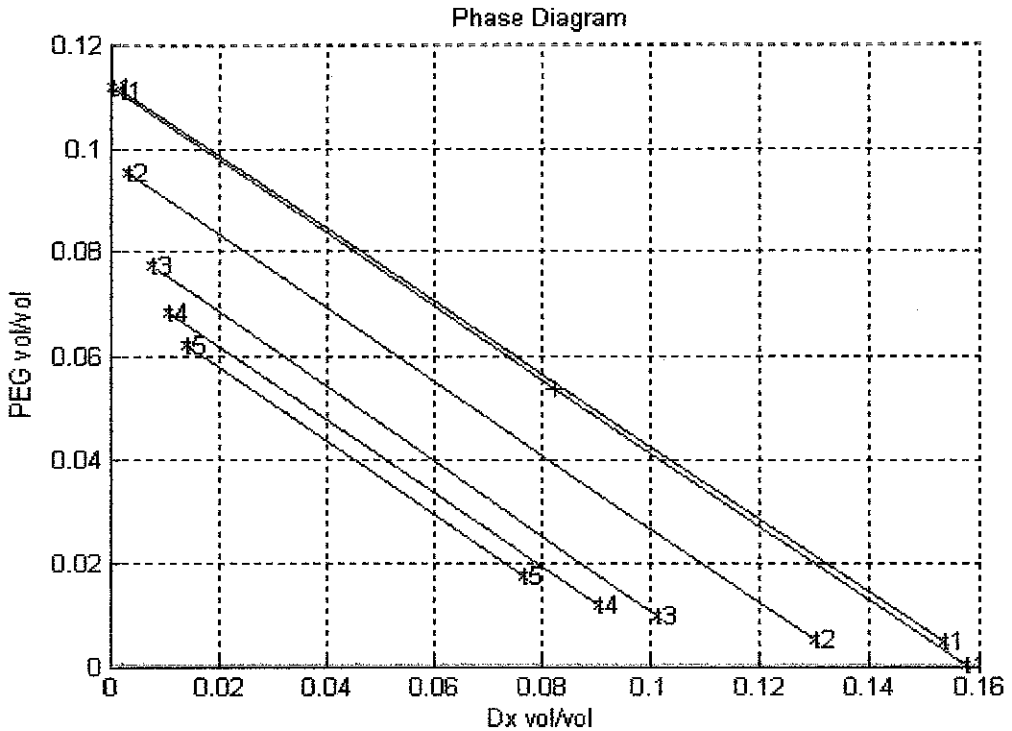


Figure 8: Phase diagrams of PEG 8000 - Dextran 75 at 23°C [30].

Figure 8 shows the phase diagrams of ATPES of interest consisting of Dextran 75 – PEG 8000 at 23°C [30] in volume fraction. Weight fraction is converted into volume fraction as in Appendix C.

Prediction of the constant FH interaction parameter values between water and PEG (ω_{12}) and PEG and Dx (ω_{23}) are made by manipulating both values until the red line almost intercept one of the tie-lines. By trial and error, the values of ω_{12}, ω_{23} and ω_{1p} are found.

$$\omega_{12} = -1200$$

$$\omega_{23} = 2450$$

$$\omega_{1p} = 19000$$

4.1.2 Minimization of Gibbs Energy – Model's Objective Function

The LLE composition is calculated using the minimization of Gibbs Energy approach. LLE is achieved when the system reaches equilibrium and forms the two-phase system, and the target protein partitions to either the top or bottom phase. At the same time, the value of the system's Gibbs energy of difference level between one-phase and two-phase is also at the lowest.

Hence, it is therefore relevant for us to look into the outcome of our simulation in terms of its Gibbs energy of mixing minimization. In the results from obtained the runs carried out, the minimization of Gibbs energy difference between one-phase and two-phase was met throughout all the attempts. Nonetheless, the respective trend of the Gibbs energy of difference between one-phase and two-phase among the simulation attempts have some inconsistencies.

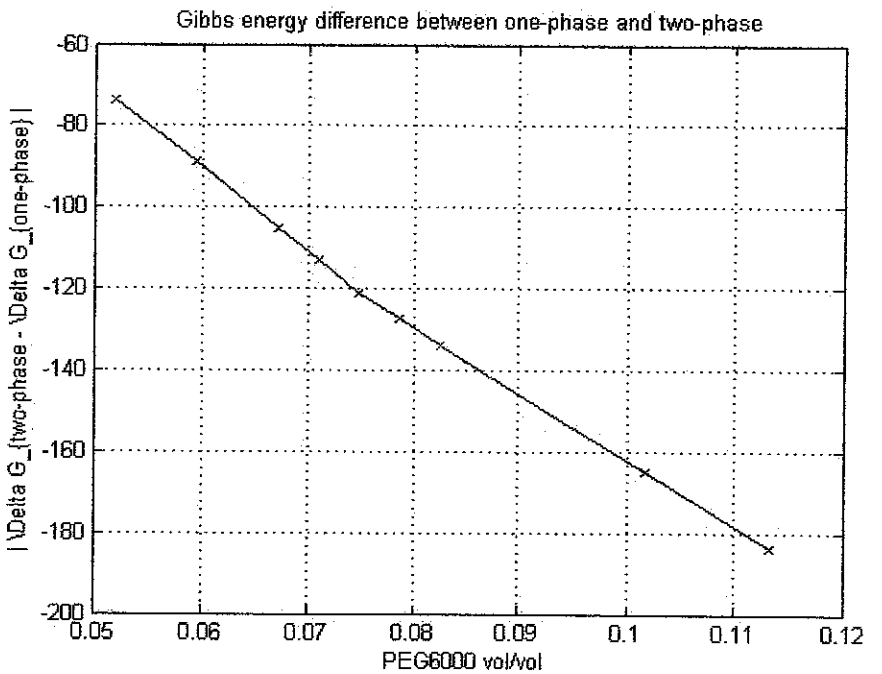


Figure 9: Gibbs Energy Difference Between one-phase and two-phase for an ATPES of Water + PEG 8000 + Dextran 75 + Human Hemoglobin at 23°C produced before adjustment of scaling factors.

Figure 9 shows the plot of the minimization of Gibbs energy difference between one phase and two-phase. However, the LLE minimization trend achieved has missing points and is not successfully converged. This plot is not consistent and therefore, the scaling factors need to be changed until the plot has 25 points all converged.

4.1.3 Protein Partitioning Behavior

Being at the bottom of the two-phase system, Dextran is of denser and more polar than the PEG phase which is at the top phase [3]. Human hemoglobin is reported to partitioned into bottom phase in an ATPES system which consists of PEG 8000 + Dextran 75 + Water at 23°C, achieving a partition coefficient, K of around 0.131 [30].

Figure 10 shows the partition coefficients in 4-component ATPES system of water, PEG 8000, Dextran 75 and human hemoglobin. PEG partitions to the top, and Dextran separates almost entirely to the bottom phase. Water partitioned between the phases. However, from the graph generated, hemoglobin is partitioned into PEG phase and this contradicted with experimental result, where hemoglobin is supposed to partition into Dextran phase. Therefore, adjustment of scaling factors values is needed.

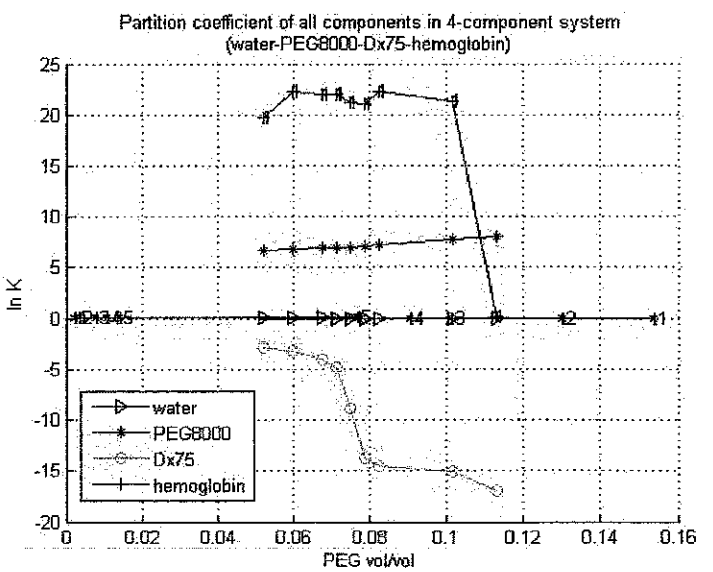


Figure 10: Partition coefficients in 4-component ATPES system of water + PEG8000 + Dx75 + human hemoglobin before adjustment of scaling factors.

4.1.4 Phase Diagram (Binodal Curve) Formation Patterns

The following stage involves matching the tie-lines from the simulation to those reported in [30]. In the event that inconsistencies do occur between the simulated results with the reported empirical results in and in order for the simulation results to be acceptable or relevant, the requirement is such that the differences (when comparing the modeling outcome to empirical results) are to be within a specified error of tolerance.

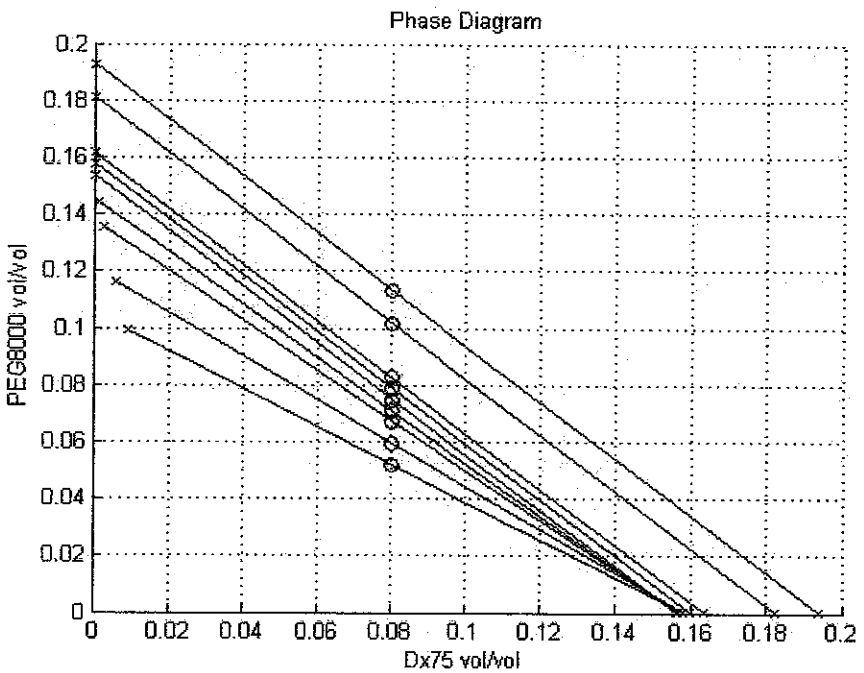


Figure 11: Phase Diagram for an ATPES of Water + PEG8000 + Dextran 75 + Human Hemoglobin at 23 °C before adjustment of scaling factors.

However, from Figure 11, we can see that the simulated tie lines are not consistent. Failing to comply with the required minimum range of error tolerance will result in the rejection of the modeling outcome, and re-diversion of the decision making process flow back to Step 2 of the Manual Partial Enumeration Strategy.

4.1.5 Adjustment of Scaling Factors

To make the result more consistent, scaling factors of the components are changed accordingly depending on the size of the molecules. The scaling factors values act as penalty terms to counter the non-uniform sensitivity of the difference in overall Gibbs energy of mixing to each component. The simplest way of dealing with this scaling problem was to select scaling factors that reduce or increase the order of magnitude of each term so that they have similar orders of magnitude [3].

The scaling factors values are found from trial and error, until the results and plots are consistent. Table 9 shows the values of scaling factors of each component.

Table 9 : The scaling factors of each component calculated for LLE calculations using the Gibbs energy of mixing minimization approach for the water + PEG 8000 + Dextran 75 + Human Hemoglobin system.

Components	ATPES + Human Hemoglobin
Water	1
PEG-8000	1e-1
Dx75	1e-7
Human hemoglobin	1e-8

Figure 12 shows Gibbs energy difference between one-phase and two-phase for an ATPES of Water + PEG 8000 + Dextran 75 + Human Hemoglobin at 23˚C after scaling factors are changed. The Gibbs energy of mixing difference also changes monotonically and smoothly as overall PEG concentration increases. The presence of target protein does not significantly change the Gibbs energy of mixing difference [3].

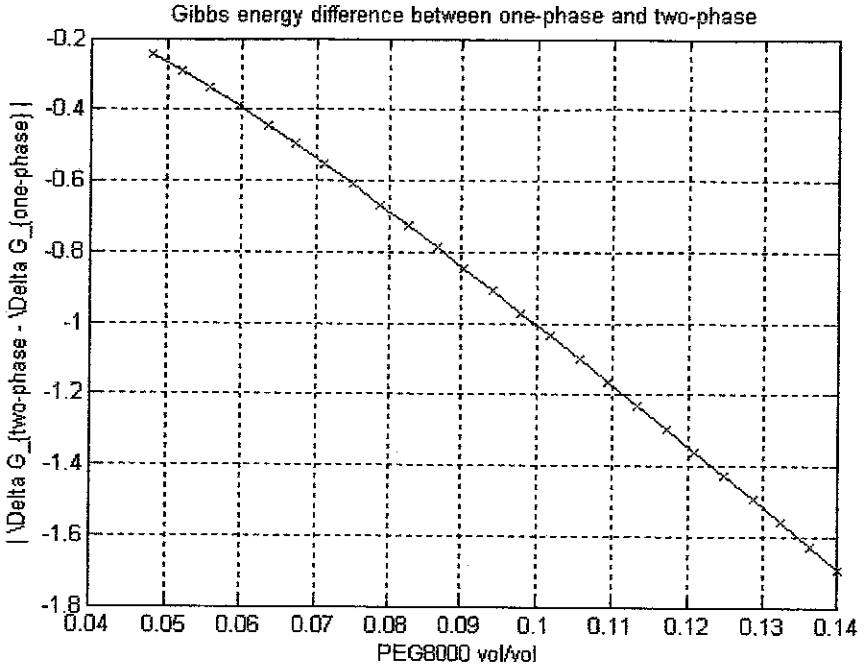


Figure 12 : Gibbs energy difference between one-phase and two-phase for an ATPES of Water + PEG8000 + Dextran 75 + Human Hemoglobin at 23˚C after adjustment of scaling factors.

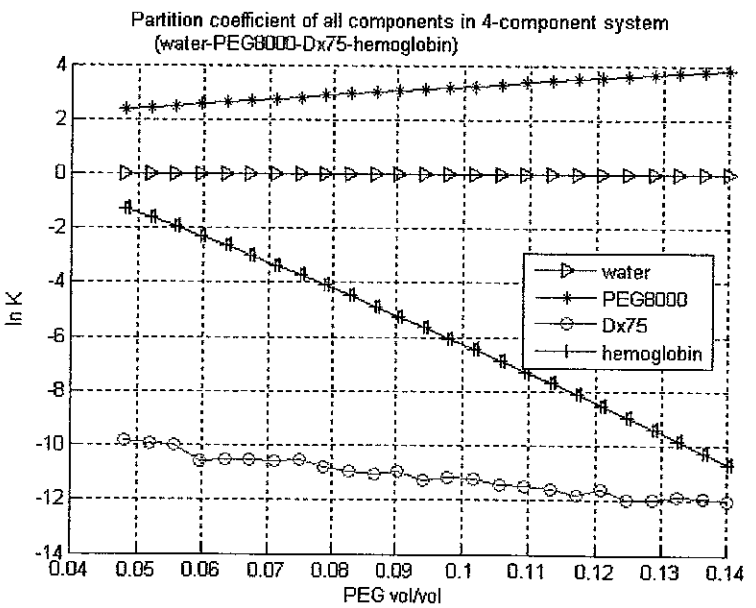


Figure 13 : Partition coefficient of all components in 4-component ATPES of Water + PEG8000 + Dx75 + Human Hemoglobin at 23˚C after adjustment of scaling factors.

Figure 13 shows the partition coefficient calculated for each component. Water partitions equally between the phases, PEG8000 partitions to the top phase and Dextran 75 occupies the bottom phase. In the model, human hemoglobin partitions into the bottom phase, due to the unfavorable interaction with PEG8000. By changing scaling factors accordingly, the plot is more consistent and all 25 points are plotted.

Figure 14 shows the phase diagram of the ATPES system after changing the values of scaling factors. It shows the tie lines and LLE compositions calculated for the mixtures that are indicated with the red points. We see that in a confined range, the tie lines are roughly parallel and the length increases with increasing overall PEG composition in the system. As predicted by the theory and observed by experiment, the LLE behavior is not significantly changed by the presence of a small amount of target proteins [3].

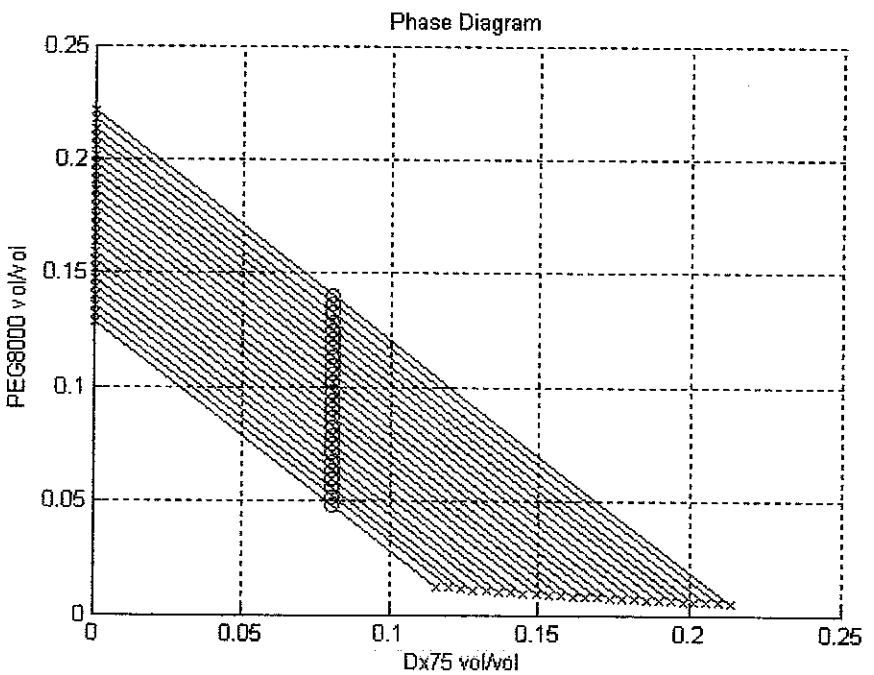


Figure 14: Phase Diagram for an ATPES of Water + PEG8000 + Dextran 75 + Human Hemoglobin at 23 °C after adjustment of scaling factors.

4.2 Lysozyme from chicken egg white

Analysis is carried out on the partitioning of human hemoglobin in an ATPES consisting of Water-PEG8000-Dx75–lysozyme. Initial guesses for the components' scaling factors and FH parameters are made. Since lysozyme from chicken egg white has a relatively small molecular weight as compared to other components, the degree of polymerization of lysozyme from chicken egg white is assumed to be small as well.

4.2.1 Initial Guess for Interaction Parameters

Table 10: The initial guesses for the components' scaling factors and fixed FH parameters used for the modeling for aqueous polymer mixture system containing lysozyme from chicken egg white. Subscripts: 1 = water, 2 = PEG8000, 3 = Dx75, p = lysozyme from chicken egg white

System	Water – PEG8000-Dx75
Temperature	23°C
Overall composition (volume fraction)	86 vol% water, 5 vol% PEG, 8 vol% Dx, 1 vol% lysozyme from chicken egg white
Scaling factors	Water: 1, PEG8000: 1e-1, Dx75: 1e-7, Lysozyme from chicken egg white: 1e-6
Degree of polymerization	$r_1 = 1$, $r_2 = 100$, $r_3 = 4500$, $r_{lysozyme} = 100$
Constant Flory Huggins interaction parameters (J/mol)	$\omega_{12} = -1200$, $\omega_{13} = 0$, $\omega_{23} = 2450$, $\omega_{1p} = 100$

Figure 15 shows the phase diagram of ATPES of interest consisting of Dextran75 – PEG-8000 –lysozyme from chicken egg white at 23°C [30] in volume fraction. The red line has similar slope as compared to the blue line, therefore the initial guess of interaction parameters of the components are acceptable. From the model, the partition coefficient

for lysozyme from chicken egg white is 3.255076 (ln Kp = 1.180216). Experimentally, the partition coefficient is 2.36 [30].

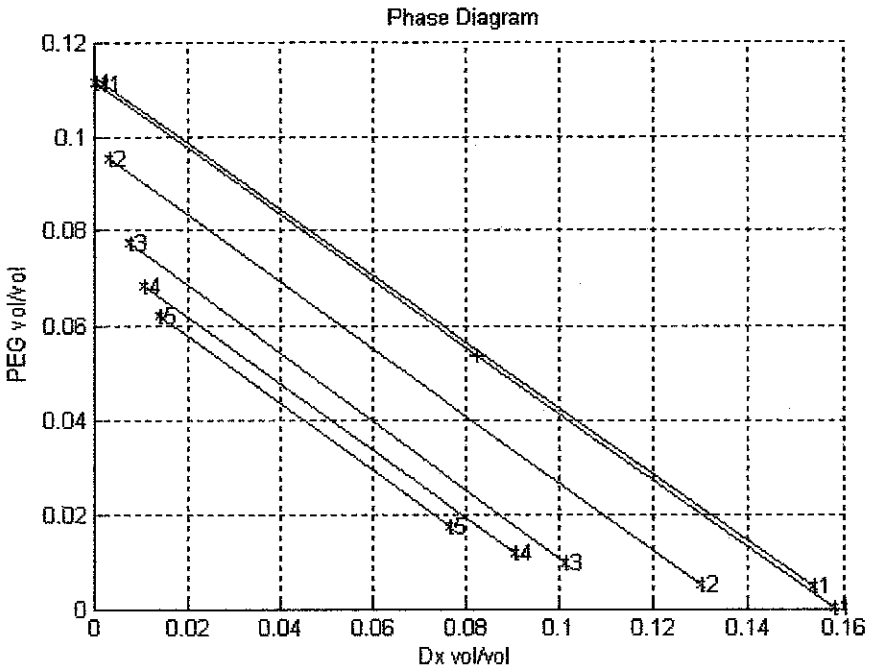


Figure 15 : Phase diagram of ATPES of interest consisting of water + PEG 8000 + Dextran 75 + lysozyme from chicken egg white at 23°C [30]

4.2.2 Minimization of Gibbs Energy – Model's Objective Function

Figure 16 shows the plot of the minimization of Gibbs energy difference between one phase and two-phase. Although there are some missing plots, it still clear that the Gibbs energy of mixing difference changes monotonically and smoothly as overall PEG concentration increases. The presence of target protein does not significantly change the Gibbs energy of mixing difference [3].

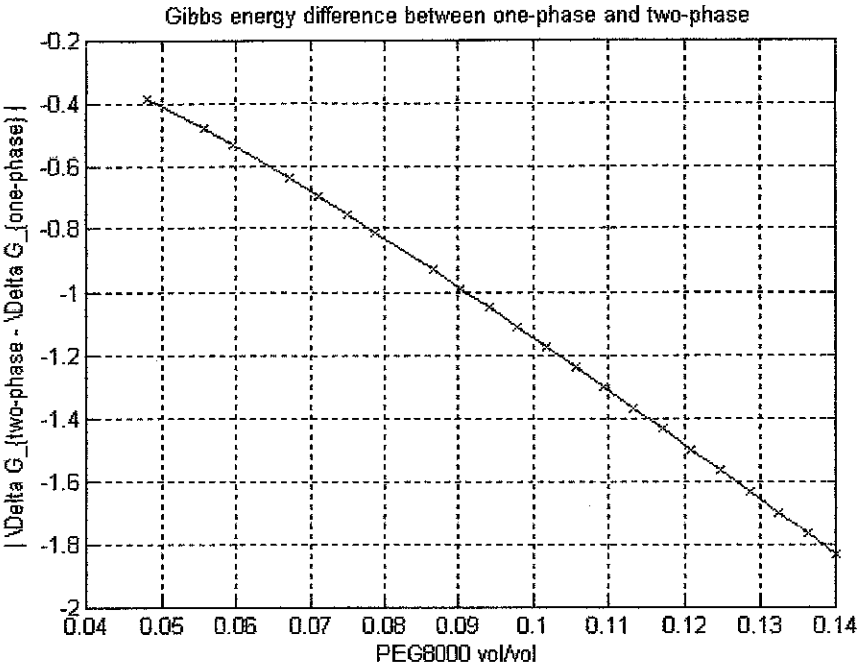


Figure 16: Gibbs Energy Difference Between one-phase and two-phase for an ATPES of Water + PEG8000 + Dextran 75 + lysozyme from chicken egg white (at 23˚C)

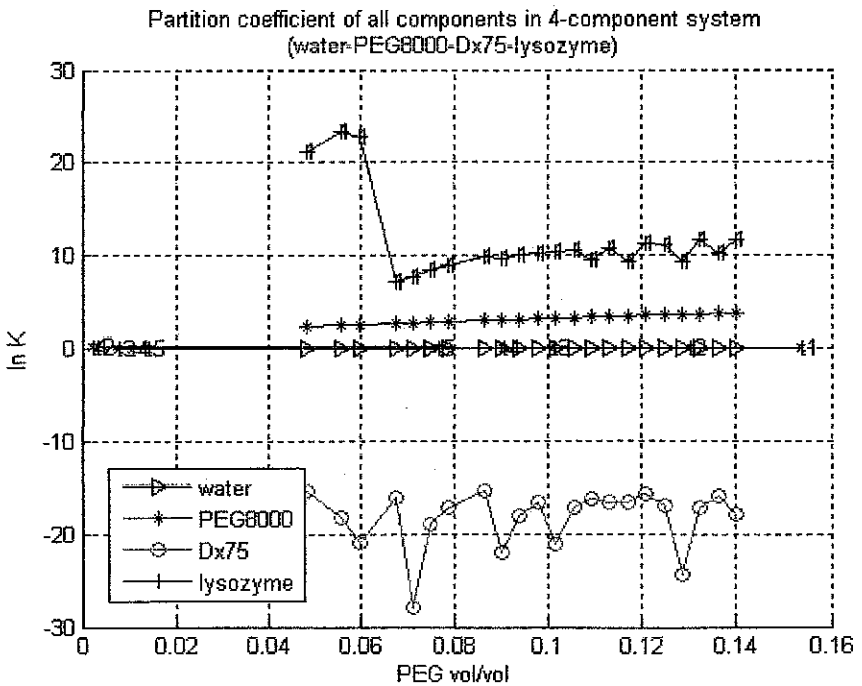


Figure 17: Partition coefficients in 4-component ATPES system of water + PEG8000 + Dextran 75 + lysozyme from chicken egg white at 23°C.

4.2.3 Protein Partitioning Behavior

Figure 17 shows the partition coefficients in 4-component ATPES system of water, PEG8000, Dextran 75 and lysozyme from chicken egg white. PEG8000 partitions to the top, and Dextran 75 separates entirely to the bottom phase. Water partitioned between the phases. From the model, lysozyme from chicken egg white partitioned into PEG phase and this correspond with experimental result [30].

However, the plots of Dx75 and lysozyme from chicken egg white are not smooth. This might be the reason because lysozyme from chicken egg white has small molecular weight as compared to other molecules. Besides that, it is also due to the unfavorable interaction with Dextran-75 as well as the small degree of polymerization of lysozyme from chicken egg white relative to the other components [3]. Lysozyme from chicken egg white partitions into the top phase because the self-energy in the top phase of this ATPES is much higher than and therefore overcoming the repulsive interaction between PEG and lysozyme from chicken egg white [3].

4.2.4 Phase Diagram (Binodal Curve) Formation Patterns

From Figure 18, we can see that the simulated tie lines are not consistent, and there are some missing tie lines. However, the plot is still consistent with almost similar slope. The missing tie lines might be because of the lysozyme from chicken egg white is a relatively small protein and has small molecular weight compared to other molecules and therefore, some points cannot converge and form tie lines.

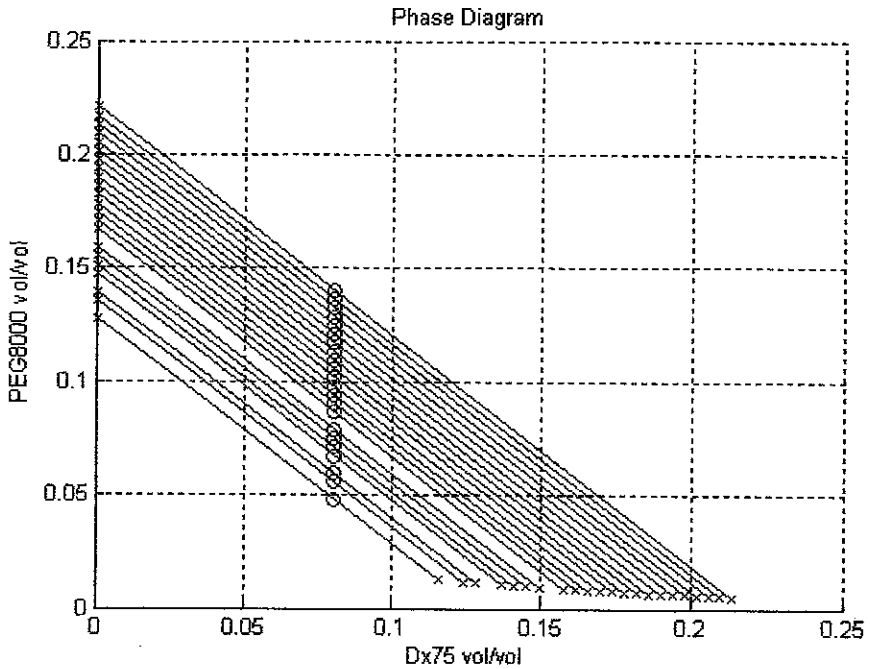


Figure 18: Phase Diagram for an ATPES of water + PEG8000 + Dextran 75 + lysozyme from chicken egg white at 23 ˚C

4. 3 Error Deviations in Results

An acceptable degree deviation of the simulation's outcome as compared to the experimentally reported results [30] would reflect the reliability of this modeling approach in substitution of laboratory studies. This is measured via the following proposed formula:

Error of composition at the top phase

$$error = \sqrt{\left(x_{top,PEG} - x_{top,PEG,lit}\right)^2 + \left(x_{top,Dx} - x_{top,Dx\,lit}\right)^2}$$

where:

$x_{top,PEG}$ = composition of PEG 8000 at the top phase from model,

$x_{top,PEG,lit}$ = composition of PEG 8000 at the top phase obtained experimentally,

$x_{top,Dx}$ = composition of Dx 75 at the top phase from model, and

$x_{top,Dx,lit}$ = composition of Dx 75 at the top phase obtained experimentally.

Error of composition at the bottom phase

$$error = \sqrt{\left(x_{bottom,PEG} - x_{bottom,PEG,lit}\right)^2 + \left(x_{bottom,Dx} - x_{bottom,Dx\,lit}\right)^2}$$

where:

$x_{bottom,PEG}$ = composition of PEG 8000 at the bottom phase from model,

$x_{bottom,PEG,lit}$ = composition of PEG 8000 at the bottom phase obtained experimentally,

$x_{bottom,Dx}$ = composition of Dx 75 at the bottom phase from model, and

$x_{bottom,Dx,lit}$ = composition of Dx 75 at the bottom phase obtained experimentally.

4.3.1 Human Hemoglobin

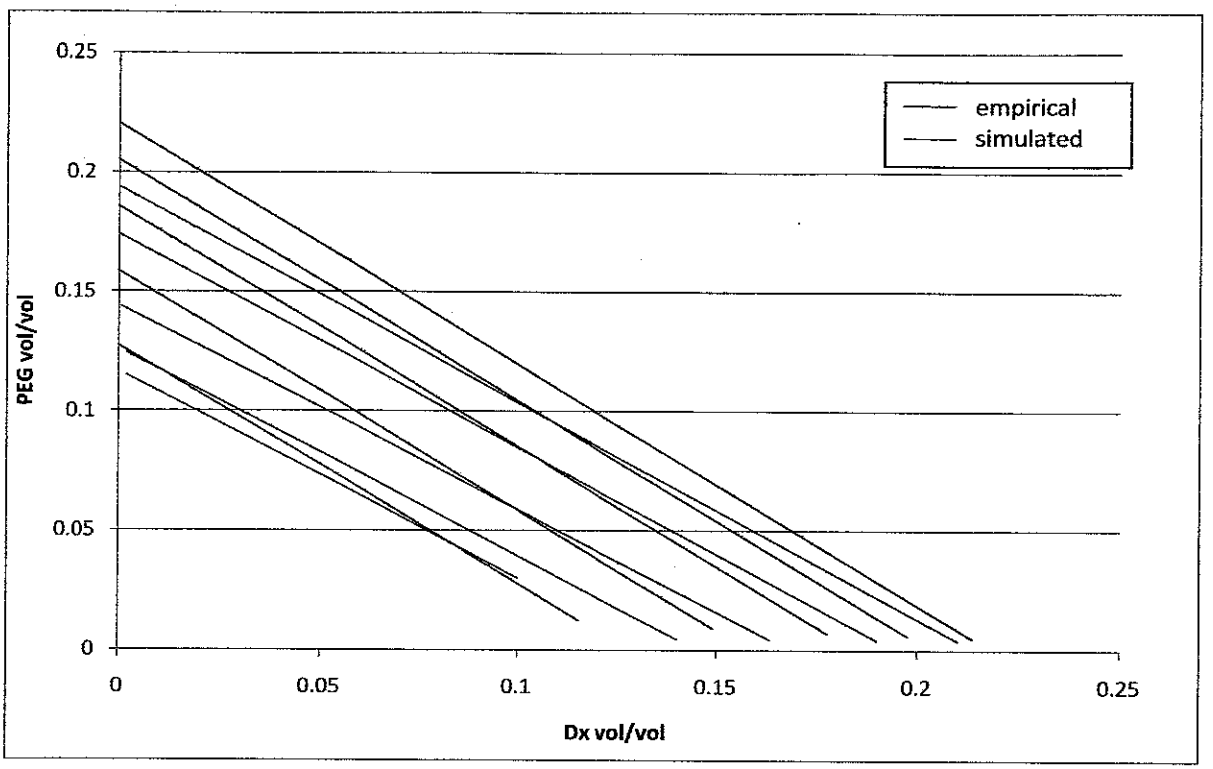


Figure 19 : Comparison of phase diagram of water + PEG 8000 + Dx75 + human hemoglobin between empirical tie lines [30] and tie line simulated by modeling approach by Ahmad [3] on common plotting plane.

Error of composition at the top phase

$$error = \sqrt{\left(x_{top,PEG} - x_{top,PEG,lit}\right)^2 + \left(x_{top,Dx} - x_{top,Dx\ lit}\right)^2}$$

$$error = \sqrt{(0.1553 - 0.1203)^2 + (3.871 \times 10^{-6} - 0.001976)^2}$$

$$error = 0.035 = 3.5\%$$

Error of composition at the bottom phase

$$error = \sqrt{\left(x_{bottom,PEG} - x_{bottom,PEG,lit}\right)^2 + \left(x_{bottom,Dx} - x_{bottom,Dx\,lit}\right)^2}$$

$$error = \sqrt{(0.009312 - 0.004820)^2 + (0.1451 - 0.1555)^2}$$

$$error = 0.0113 = 1.13\%$$

4.3.2 Lysozyme from Chicken Egg White

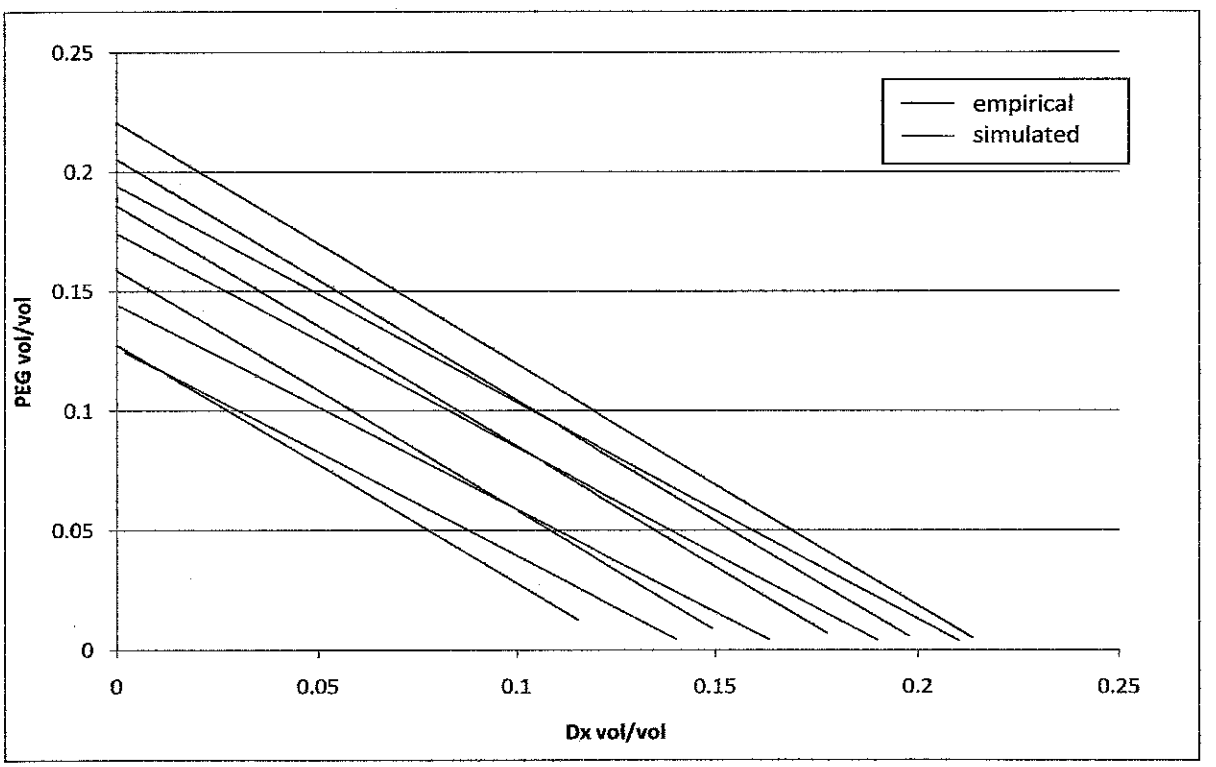


Figure 20 : Comparison of phase diagram of water + PEG 8000 + Dx75 + lysozyme from chicken egg white between empirical tie lines [30] and tie line simulated by modeling approach by Ahmad [3] on common plotting plane.

Error of composition at the top phase

$$error = \sqrt{(0.1589 - 0.1109)^2 + (5.331 \times 10^{-9} - 0.001986)^2}$$

$$error = 0.048 = 4.8\%$$

Error of composition at the bottom phase

$$error = \sqrt{(0.009163 - 0.004818)^2 + (0.1493 - 0.1535)^2}$$

$$error = 6.04 \times 10^{-3} = 0.604\%$$

Results obtained thus far have set the groundwork for further study of protein partitioning behavior in ATPES for other types of system and proteins. Besides that, the modeling parameters' fine-tuning efforts in making sure the parallelism of our modeling outcome to those reported from a basis of empirical nature has also been meticulously established.

In addition, the applicability of this modeling approach that is based on Gibbs energy of minimization to determine the behavior of the protein partitioning which are human hemoglobin and lysozyme from chicken egg white, in addition to calculating equilibrium behavior of an aqueous two-phase system, *i.e.* water – PEG 8000 – Dextran 75 has been verified.

CONCLUSION AND RECOMMENDATION

ATPES comprises of a liquid-liquid extraction technique that takes advantage of phase separation phenomenon in order to recover protein. In this water-rich system, proteins will selectively partitions into one of the phases. ATPES allows continuous steady-state operation with high capacity, easy engineering scale-up, and also offers high yield for protein recovery. Despite the fact that ATPES has the potential to fulfill the demand by industry as an efficient, large-scale; primary downstream process to extract protein, there are still a limited number of applications at industrial scale. This is due to the economic concerns regarding the amount of polymer required to extract a small amount of protein as some of the polymers are quite expensive.

Due to the advantages of Flory-Huggins theory in capturing the protein partitioning behaviour, this project is carried out to focus on extending the application [3] of a polymer thermodynamic model since there are limited designs methods available. MATLAB is used as a modeling simulation language to calculate the prediction of the LLE behaviour in an ATPES system.

In this project, using the modeling approach by Ahmad [3], the LLE equilibrium behaviour of ATPES, namely a polymer-polymer system of water-PEG 8000-Dextran 75 are successfully predicted. The human hemoglobin and lysozyme from chicken egg white partitioning behavior in ATPES mentioned are also successfully predicted. Table 11 shows the summary of degree of polymerization and Flory Huggins interaction parameters used for human hemoglobin and lysozyme from chicken egg white for this model. Table 12 shows summary of scaling factors of each component in ATPES. The results obtained from this model tally with the results found experimentally. Therefore, we can conclude that this model can be used to predict protein partitioning behavior in ATPES.

Table 11: Summary of degree of polymerization and Flory Huggins interaction parameters used for human hemoglobin and lysozyme from chicken egg white partitioned in water – PEG 8000 – Dx 75 system at 23°C.

Parameter	ATPES + Human Hemoglobin	ATPES + Lysozyme from chicken egg white
Degree of polymerization	$r_1 = 1, r_2 = 100, r_3 = 4500,$ $r_{human\ hemoglobin} = 2720$	$r_1 = 1, r_2 = 100, r_3 = 4500,$ $r_{lysozyme} = 100$
Constant Flory Huggins interaction parameters (J/mol)	$\omega_{12} = -1200, \omega_{13} = 0,$ $\omega_{23} = 2450, \omega_{1p} = 0,$ $\omega_{2p} = 19000, \omega_{3p} = 40.5$	$\omega_{12} = -1200, \omega_{13} = 0,$ $\omega_{23} = 2450, \omega_{1p} = 0,$ $\omega_{2p} = 100, \omega_{3p} = 20$

Table 12: Summary of scaling factors of components in water – PEG 8000 – Dx75 system at 23°C.

Components	ATPES + Human Hemoglobin	ATPES + Lysozyme from chicken egg white
Water	1	1
PEG-8000	1e-1	1e-1
Dx75	1e-7	1e-6
Target protein	1e-8	1e-2

For future work, the study can maybe be expanded to other polymeric biomolecules such as nucleic acids and polysaccharides. Besides that, model can also be improved by also including program that can predict the time taken for the ATPES from mixture to partition into two phases.

REFERENCES

[1] **Flory, P.J.** Thermodynamics of high-polymer solutions. J.Chem.Phys., pages 51-61, 1942.

[2] **Huggins, M.L.** Some properties of solutions of long-chain compounds. J. Phys.Chem., 46:151-158, 1942.

[3] **Ahmad, Murni M.** Design of Aqueous Two-Phase Protein Extraction Process. Ph. D. Thesis, Carnegie Mellon University, Pittsburgh, USA, 2006.

[4] **M.H. Hariri, J.F. Ely, G.A. Mansoori,** Bio. Technol. 7 (1989) 686.

[5] **M.T. Cunha, M.R Aires-Barros, and J.M.S. Cabral**, Isolation and Purification of Proteins, Extraction for Rapid Protein Isolation, 9:322-372, 2003

[6] **Kroner et al.**, Biotechnology Bioengineering, 24, 1015-1045, 1982

[7] **Kula et al.,** Adv. Biochem. Bioeng. 24, 73-118, 1982

[8] **P.A. Albertsson,** Partition of Cells, Particles and Macromolecules, 3rd ed., Wiley Interscience, New York, 1986

[9] **M.R. Kula, K.H. Kroner, H. Husted, H. Schütte,** Ann. N.Y. Acad. Sci. 369 (1981) 341

[10] **M. Chaplin,** 20 Dec 2004 <http://www.lsbu.ac.uk/biology>

[11] **P.A. Albertsson,** Chromatography partition of cells and cell fragments, Nature 177 (1956) 771–774.

[12] **H. Walter, G. Johansson,** Methods in Enzymology, 228, Academic Press, New York, 1994.

[13] **W.Y. Yang, C.D. Lim, I.M. Chu, C. Lee,** Separation of Desacetyl Cephalosporin C by Aqueous Two Phase Partition, J. Biotechnol, Bioeng. 35 (1994) 439-445.

[14] **J.C. Marcos, L.P. Fonseca, M.T. Ramah, J.M.S. Cabral,** Partial purification of penicillin acylase from *Escherichia coli* in poly(ethylene glycol)-sodium citrate aqueous two-phase systems, J. Chromatogr. B 734 (1999) 15-22.

[15] **B. Mokhtarani, R. Karimzadeh, M.H. Amini, S.D. Manesh,** Partitioning of Ciprofloxacin in aqueous two phase system of poly(ethylene glycol) and sodium phosphate, Biochem. Eng. J. 38 (2008) 241-247

[16] **K. Schugerl, J. Hubbuch,** Integrated bioprocesses, Curr. Opin. Microbiol. 8 (2005) 294–300.

[17] **A.M. Azevedo, L.P. Fonseca, D.M.F. Prazeres,** Stability and stabilisation of penicillin acylase, J. Chem. Technol. Biotechnol. 74 (1999) 1110–1116.

[18] **P.A.J. Rosa, A.M. Azevedo, M.R. Aires-Barros,** Application of central composite design to the optimisation of aqueous two-phase extraction of human antibodies, J. Chromatogr. A 1141 (2007) 50–60.

[19] **P.A. Pessôa Filho, R.S. Mohamed,** Thermodynamic Modeling of the partitioning of biomolecules in aqueous two phase systems using a modified Flory-Huggins equation, P. Biochem 39 (2004) 2075-2083.

[20] **E. Edmond, A.G. Ogston,** Biochem. J. 109 (1968) 569-576.

[21] **W. G. McMillan, J. E. Mayer,** J. Chem. Phys. 13 (1945) 276-284.

[22] **C.H. Kang, S.I. Sandler,** Fluid Phase Equilib. 38 (1987) 245–272.

[23] **C.H. Kang, S.I. Sandler,** Ind. Eng. Chem. Res. 28 (1989) 1537–1542.

[24] **H. Hartounian, E. Floeter, E.W. Kaler, S.I. Sandler,** AIChE J. 39 (1993) 1976–1984.

[25] **A. Fredenslund, R.L. Jones, J.M. Prausnitz,** AIChE J. 21 (1977) 1086–1099.

[26] **T. Oishi, J.M. Prausnitz,** Ind. Eng. Chem. Proc. Des. Dev. 17 (1978) 333–339.

[27] **M.T. Zafarani-Moattar, R. Sadeghi,** Fluid Phase Equilib. 202 (2002) 413-422

[28] **B.Y. Zaslavsky**, Aqueous Two Phase Partitioning : Physical Chemistry and Bioanalytical Applications, 1995

[29] **D.E. Brooks, K.A. Sharp, D.Fisher**, Partitioning in Aqueous Two Phase Systems : Theory, Methods, Uses and Applications to Biotechnology (H. Walter, D.E.Brooks, D. Fisher, eds.), Academic Press, Orlando, Fla, 1985, pp.11-84.

[30] **P. P. Madeira, J.A. Teixeira, E.A. Macedo, L.M. Mikheeva, B.Y. Zaslavsky,** Correlations between distribution coefficients of various biomolecules in different polymer/ polymer aqueous two phase systems, Fluid Phase Equilib. 267 (2008) 150-157.

[31] **Singh and Rizvi (1995).** Bioseparation Processes in Foods, (1), New York, Marcel Dekker, Inc.

[32] **Johansson, H. O., Karlstrom G., Tjerneld F., & Haynes, C. A. (1998).** Driving force for phase separation and partitioning in aqueous two-phase systems. Journal of Chromatography 711: 3-17.

[33] **R. Collander**, Acta Physiol. Scand. 13 (1947) 363-381

APPENDICES

Appendix A.1 : FYP 1 Gantt Chart

No.	Detail/Week	1	2	3	4	5	6	7	8	9	10		11	12	13	14
1	Selection of project topic	x	x													
2	Submission of proposal		x													
3	Preliminary research work		x	x	x											
4	Submission of preliminary report				x											
5	Literature review on ATPES as protein separation technique				x	x	x	x	x	x	x	x	x	x	x	
6	Literature review on modeling and experimental approach				x	x	x	x	x	x	x	x	x	x	x	
7	Literature research on suitable polymer-polymer ATPES model and application for study				x	x	x	x	x	x	x	x	x	x	x	
8	Familiarize with application of modeling and simulation by Ahmad to predict the LLE behaviour of polymer-salt ATPES				x	x	x	x	x	x	x	x	x	x	x	
9	Familiarize with MATLAB tools				x	x	x	x	x	x	x	x	x	x	x	
10	Submission of progress report									x						
11	Seminar										x					
12	Submission of interim report													x		
13	Final oral presentation															x

Appendix A.2 : FYP 2 Gantt Chart

No.	Detail/Week	1	2	3	4	5	6	7	8	9	10		11	12	13	14
1	Project Work Continue	x	x	x	x	x	x	x	x	x	x		x	x		
2	Calculation of LLE data using modeling approach of ATPES system (Water-PEG 8000-Dx75 system)	x	x	x	x	x	x	x	x	x	x		x	x		
3	Prediction of protein partitioning behavior	x	x	x	x	x	x	x	x	x	x		x	x		
2	Submission of Progress Report 1				x											
3	Submission of Progress Report 2												x			
4	Pre-EDX/ Poster Exhibition													x		
5	Submission of Dissertation (soft bound)													x		
6	Oral Presentation														x	
7	Submission of Project Dissertation (Hard Bound)															x

Appendix B : Calculation of initial guess FH parameters of human hemoglobin used for the modeling and simulation for aqueous polymer mixture system containing PEG-8000 + Dextran75 + Water + Human Hemoglobin [30], interpolated using the data for Ovalbumin and Phosphofructokinase [32].

Table A1 : Interpolation of initial guess FH parameters of human hemoglobin

Protein	Ovalbumin	Human hemoglobin	Phosphofructokinase
MW (Da)	43000	68000	85000
r.prot	100	x	4500
omega. PEG_prot	0	y	150
omega.Dx_prot	100	z	0

Calculation of degree of polymerization of human hemoglobin

Let x be the degree of polymerization, r.prot of human hemoglobin;

$$\frac{x-100}{4500-100}=\frac{68000-43000}{85000-43000}$$

$$x=2720$$

$\therefore$ Degree of polymerization, r.prot of human hemoglobin = 2720

Appendix C : Example of calculation of volume fraction from the information of weight fraction

Table A2 : Overall composition of ATPES in weight fraction

System	Water – PEG8000-Dx75
Overall composition (weight fraction)	82% (w/w) water, 6% (w/w) PEG, 12% (w/w) Dx

Density of the mixture, ρ is given by:

$$\rho = \frac{1}{\sum v.w}$$

where v is specific volume (cm^3/g) and w is the weight fraction of the component.

Given;

Specific volume of water: 1 cm^3/g

Specific volume of PEG: 0.832 cm^3/g

Specific volume of Dx: 0.626 cm^3/g

Therefore,

$$\rho = \frac{1}{(1 \times 0.82) + (0.832 \times 0.06) + (0.626 \times 0.12)}$$

$$\rho = 1.060\ g/cm^3$$

Volume fraction, ϕ of component i is calculated by:

$$\phi_i = \rho v_i w_i$$

Example :

$\phi_{water} = 1.060 \times 1 \times 0.82 = 0.864$ vol% water

Appendix D: Example of data file

```
% Data file for WaterPEG8000DxT75hemoglobin
% reference: Pedro et al.

% FH interaction parameter between water-water
omega.water_water = 0;

% FH interaction parameter between water-PEG
omega.water_PEG = -1200;

% FH interaction parameter between water-Dx
omega.water_Dx = 0;

% FH interaction parameter between PEG-PEG
omega.PEG_PEG = 0;

% FH interaction parameter between PEG-Dx
omega.PEG_Dx = 2450;

% FH interaction parameter between Dx-Dx
omega.Dx_Dx = 0;

% specifying the FH interaction parameter between water and protein
omega.water_prot = 0;

% specifying the FH interaction parameter between PEG and protein
omega.PEG_prot = 0;

% specifying the FH interaction parameter between Dx and protein
omega.Dx_prot = 19000;

% specifying the FH interaction parameter between protein and protein
omega.prot_prot = 0;

% specifying number of component
no_comp = 4;

% specifying the component
comp.c1 = 'water';
comp.c2 = 'PEG';
comp.c3 = 'Dx';
comp.c4 = 'hemoglobin';

% degree of polymerisation
r.water = 1;
```

```
r.PEG = 100;
r.Dx = 4500;
r.hemoglobin = 2720;

% specific volume [cm3/g]
% Kang and Sandler 1987
v.water = 1;
v.PEG = 0.832;
v.Dx = 0.626 ;
v.hemoglobin = 0.73;

% source:
% /results_LLEgrid_3comp_type2_tocompare_mylog.mat

w_lit.mix = [ 1-0.0606-0.1241  0.0606  0.1241
    1-0.0522-0.1075   0.0522   0.1075
    1-0.0434-0.0899   0.0434   0.0899
    1-0.0408-0.0842   0.0408   0.0842
    1-0.0382-0.0785   0.0382   0.0785
    1-0.0357-0.0729   0.0357   0.0729];


w_lit.top = [1-0.1302-0.0031  0.1302  0.0031
    1-0.1121-0.0048   0.1121  0.0048
    1-0.0912-0.0121  0.0912  0.0121
    1-0.0805-0.0164  0.0805  0.0164
    1-0.0728-0.0215  0.0728  0.0215];


w_lit.bot = [1-0.0053-0.2244  0.0053   0.2244
1-0.0057-0.1923   0.0057   0.1923
1-0.0111-0.1521   0.0111   0.1521
1-0.0132-0.1367   0.0132   0.1367
1-0.02-0.1166     0.0200   0.1166];

[m,n] = size(w_lit.mix);
para.v = [v.water v.PEG v.Dx];

% loop to convert each row of w into phi
for i = 1:m

    % mixture composition
    phi_lit.mix(i,:) = Calc_w2phi(w_lit.mix(i,:),para);

end

[m,n] = size(w_lit.top);
```

```
for i = 1:m

    % top composition
    phi_lit.top(i,:) = Calc_w2phi(w_lit.top(i,:),para);

    % bottom composition
    phi_lit.bot(i,:) = Calc_w2phi(w_lit.bot(i,:),para);

end

% assigning the system type
type = 1;

% plotting the results
h = plot_LLE(1,phi_lit,'bx-',type,comp,1,1);

% saving the data in a mat file
save WaterPEG8000Dx75hemoglobin;
```

Appendix D: Example of main script run file

```
% script file to run file to_split
% in order to calculate split compositions
% for a system containing
% water, polymer1, polymer2 and target protein

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% clearing memory and closing all figures
clear all;
close all;

t=127;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% setting the save data file (as mat file)
SaveFile = sprintf('results_LLE_4comp_varyPEG_polym_tgt_thesis_%d',t);

% checking if data file exist
existflag = check_file(SaveFile);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% declaring global variables
global R;
global TolF;
global TolX;
global Toldmu;
global MaxIter;

% specifying the universal gas constant
R = 8.314;

% specifying the tolerance for function evaluation in solver
TolF = 1e-8;

% specifying the tolerance for variable evaluation in solver
TolX = 1e-6;

% specifying the tolerance for chemical potential difference
Toldmu = [ 1 1 1 1 ];

% specifying the maximum iteration for SolvOpt
MaxIter = 10000000;
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% specifying the number of components
no_comp = 4;

% specifying the name of components
comp_name.c1 = 'water';
comp_name.c2 = 'PEG8000';
comp_name.c3 = 'Dx75';
comp_name.c4 = 'hemoglobin';

% specifying the ATPS and proteins
atps_sys = sprintf('%s%s%s',comp_name.c1,comp_name.c2, comp_name.c3);
tgt_protein = comp_name.c4;
type = 1;

% specifying the system temperature
T = 273.15 + 23;

% loading the parameters for ATP-protein-contaminant system
parameter = GetParameters_4comp(atps_sys,type,tgt_protein,T);
parameter.omega(2,4) = 150;

% specifying initial total amount of the mixture (one-phase system)
N0 = 1;

% specifying the composition for the mixture (one-phase system)
vol.PEG = linspace(0.048,0.14,25);
vol.Dx = 0.08;
vol.hemoglobin = 0.0001;

% generating the composition matrix
for i = 1:length(vol.PEG)
  phi(i,:) = [ (1-vol.PEG(i)-vol.Dx-vol.hemoglobin) ...
               vol.PEG(i)  vol.Dx  vol.hemoglobin];
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% specifying the initial points for solver
s0 = [ 0.5 0.1 0.9 0.9 ];

% solving the fractions of mole in the bottom phase at equilibrium
for i = 1:length(vol.PEG)
  [f_split,phi_split,mole_split,Ntot_split,dG_split,dU_split, ...
   K_split,options] =
Calc_split_solvopt(phi(i,:),N0,parameter,s0,type);
```

```
  f_store(i,:) = f_split;
  phi_split_store.top(i,:) = phi_split.top;
  phi_split_store.bot(i,:) = phi_split.bot;
  dG_split_store(i) = dG_split;
  K_split_store(i,:) = K_split;

  % assigning flag for convergence
  if ( options(9) > 0 )
    convflag(i) = 1;
  else
    convflag(i) = 0;
  end
end

% extracting the results
cnt = 1;
for i = 1:length(vol.PEG)
  if ( convflag(i) )
    phi_one_plot(cnt,:) = phi(i,:);
    phi_split_plot.top(cnt,:) = phi_split_store.top(i,:);
    phi_split_plot.bot(cnt,:) = phi_split_store.bot(i,:);
    K_plot(cnt,:) = K_split_store(i,:);
    dG_plot(cnt) = dG_split_store(i);
    flag_plot(cnt) = convflag(i);
    cnt = cnt+1;
  end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% plotting partition coefficient
g(1) = figure(1);
%subplot(3,1,3),
hold on;
h(1) = plot(phi_one_plot(:,2),log(K_plot(:,1)),'k>-');
h(2) = plot(phi_one_plot(:,2),log(K_plot(:,2)),'*-','Color',[1  0  0]);
for i = 1:length(K_plot)
  text(phi_one_plot(i,2),log(K_plot(i,4)),sprintf('%d',flag_plot(i)));
end
h(3) = plot(phi_one_plot(:,2),log(K_plot(:,3)),'o-','Color',[0  1  0]);
h(4) = plot(phi_one_plot(:,2),log(K_plot(:,4)),'+-','Color',[0  0  1]);
grid on;
ylabel('ln K');
xlabel('PEG vol/vol');
title(sprintf('Partition coefficient of all components in %d-component
system\n(%s-%s-%s-
%s)',no_comp,comp_name.c1,comp_name.c2,comp_name.c3,comp_name.c4));
```

```
eval(sprintf('legend(h,''%s'',''%s'',''%s'',''%s'',''Location'',''Best'
')',comp_name.c1,comp_name.c2,comp_name.c3,comp_name.c4));

% plotting the Gibbs energy difference
g(2) = figure(2);
%subplot(3,1,1),
plot(phi_one_plot(:,2),dG_plot,'bx-');
%for i = 1:length(dG_plot)
%  text(phi_one_plot(i,2),dG_plot(i),sprintf('%d',flag_plot(i)));
%end
grid on;
ylabel('| \Delta G_{two-phase - \Delta G_{one-phase} |');
xlabel('PEG8000 vol/vol');
title('Gibbs energy difference between one-phase and two-phase');

% plotting phase diagram
g(3) = figure(3);
%subplot(3,1,2),
hold on;
grid on;
plot(phi_one_plot(:,3),phi_one_plot(:,2),'ro');
%for i = 1:length(flag_plot)
%
text(phi_one_plot(i,3),phi_one_plot(i,2),sprintf('%d',flag_plot(i)));
%end
plot_LLE(3,phi_split_plot,'bx-',1,comp_name,1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% saving the results
save(SaveFile);
saveas(g(3),sprintf('LLE_4comp_varyPEG_polym_prot_thesis_%d',t),
'fig');
saveas(g(1),sprintf('lnK_4comp_varyPEG_polym_prot_thesis_%d',t),
'fig');
saveas(g(2),'dG_4comp_varyPEG_polym_prot_thesis_%d', 'fig');
```