INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

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

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
RICE UNIVERSITY

Effect of Polar Functional Groups on the Phase Behavior of Amino Acids, Small Peptides, Solvents, and Polymers

by

Sharon Gail Sauer

A THESIS SUBMITTED IN
PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Walter G. Chapman, Professor, Chair
Chemical Engineering

Ka-Yiu San, Professor
Bioengineering

Clarence A. Miller, Professor
Chemical Engineering

Gustavo E. Scuseria, Professor
Chemistry

HOUSTON, TEXAS

DECEMBER, 2001
DECEMBER, 2001

ABSTRACT

Effect of Polar Functional Groups on the Phase Behavior of Amino Acids, Small Peptides, Solvents, and Polymers

by

Sharon Gail Sauer

Dipolar interactions significantly influence the phase behavior of many systems of interest to the biochemical, chemical, petroleum and polymer industries. For example, the solution behavior of amino acids, small peptides, polar solvents and co-polymers have potential applications for biochemicals, water-soluble polymers for paints and coatings, and surfactants. By considering molecular-level interactions, the phase behavior of a large range of systems can be predicted. Using structural analysis and thermodynamics, the essential role of polar functional groups on solubility of small biochemicals is established. An accurate model for fluid mixtures with multiple polar functional groups is developed.

Results from a systematic experimental study on the aqueous solubility of amino acids and dipeptides as a function of temperature, salt type, and salt concentration are analyzed. Changes in temperature and residue sequence have the most substantial effect on solubility. Structural analysis shows that intra and intermolecular association
largely influence the behavior. For small molecules, end effects dominate the behavior but should be less important for many biochemicals. Values for enthalpic changes from the solid to the infinitely dilute liquid state for the dipeptides of asp and gly are reported.

An accurate model for mixtures of polar fluids, in which any number of groups and/or any component of the mixture may be polar, is developed by applying the u-expansion to a reference fluid mixture of polar and non-polar spheres. An in-depth parameter study of this model, named Polar SAFT, for a homologous series of ketones indicates that the model parameters have physically reasonable values. A methodology is proposed for developing a group-contribution approach for the model.

The ability of Polar SAFT to accurately predict the effect of multiple dipolar groups and molecular shape on the phase behavior of binary mixtures of polar and non-polar components is exemplified by application to a series of ketone/alkane mixtures. Using only pure component parameters, Polar SAFT accurately represents these systems, indicating the predictive capability of the model and the importance of explicitly accounting for polar interactions. For alkane/copolymer (poly(ethylene-co-methyl acrylate)) solutions, Polar SAFT accurately predicts the polar co-monomer content and solvent effects on cloud point behavior.
To my children

*Samuel, John, and Jemima*

…I remind you to stir into flame the gift of God that you have…

For God did not give us a spirit of cowardice

but rather of power and love and

self-control.

So do not be ashamed of your testimony to our Lord, nor of me…

but bear your share of hardship for the gospel

with the strength that comes from God.

II Timothy 1:6-8
ACKNOWLEDGEMENTS

My eternal gratitude is given to my Lord and Savior Jesus Christ for blessing me with the opportunities to pursue life-long dreams and putting people in my life to assist and encourage me along the way. I am most indebted to my husband, Bob, who has been so encouraging, urging me on when I wasn’t sure I wanted to continue, and who has so willingly loved and supported me. I am grateful to my children, Sam, John, and Jemima for all their love and support. I pray that these years are an example to you of the importance of working hard for the goals the Lord puts in your life, and that you will have the courage to pursue your God given dreams, even when it takes many years to accomplish. I am very thankful to my parents, Johnie and Lillian Gross, for their love and for the work ethic that they instilled in me as a child. Many thanks are due my dearest friends, Tina Forster and Emma Meehan, without whose prayers and companionship this journey would have been much darker.

My most sincere gratitude is given to Professor Walter G. Chapman, my research advisor, for being a very positive influence in my life, guiding me in both research and in teaching. Thank you for being so very patient with my endless questions, for being helpful, encouraging and understanding through all the ups and downs of research and of life throughout these five years. Many thanks to Professor Ka-Yiu San for serving as my co-advisor and to Professor Clarence A. Miller and Professor Gustavo E. Scuseria for serving on my thesis committee.
Professor Pedro E. Arce, my undergraduate research advisor, deserves much credit for his belief in me and my potential during my undergraduate work. His encouragement to pursue a graduate education and his continuing confidence in me as a researcher and educator are greatly appreciated. And thanks to Pedro for continuing to be both my mentor and dear friend.

Thanks to all in my research group for your help and laughter, most especially Dave Ting for his kindness and friendship. The following people contributed to the FORTRAN code used in this work: Professor Chapman, Victor Perez, Prasanna Jog, David Ting, and Auleen Ghosh. Yea-Tyng Yang and Amit Mistry performed the solubility experiments that are analyzed in this work. Many thanks to Dr. Yang for answering all my questions regarding the solubility experiments. I also want to acknowledge each of the course instructors that I had the privilege of sitting under during my graduate education: chemical engineering (Dr. Armeniades, Dr. Chapman, Dr. Hellums, Dr. Hightower, Dr. McIntire, and Dr. Zygourakis), chemistry (Dr. Scuseria and Dr. Weisman), computational and applied math (Dr. Symes and Dr. Hribar), and physics (Dr. Hulet). I must also acknowledge the friendship and fun of Professor Jacqueline Goveas. Thanks to Diana, Lindy, Seryu, and Virginia for your willingness to always answer my trivial business related questions and assisting me.

I acknowledge the National Science Foundation for a graduate fellowship, the Robert A. Welch Foundation for financial support, and the Office of the Dean of Engineering at Rice University for two Teaching Assistantships.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER 1: INTRODUCTION</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Why Study Aqueous Solutions of Amino Acids?</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Separation Techniques and Aqueous Two-Phase Systems</td>
<td>6</td>
</tr>
<tr>
<td>1.3 Models of ATPS and Aqueous Systems for Peptides and Amino Acids</td>
<td>19</td>
</tr>
<tr>
<td>1.4 Overview of our Approach</td>
<td>21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 2: BACKGROUND INFORMATION</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Thermodynamics and Statistical Mechanics</td>
<td>24</td>
</tr>
<tr>
<td>2.1.1 Determining molecular properties of condensed fluids with statistical mechanics.</td>
<td>25</td>
</tr>
<tr>
<td>2.1.2 Brief Historical Overview of Thermodynamic Perturbation Theories for Atomic and Molecular Fluids.</td>
<td>28</td>
</tr>
<tr>
<td>2.1.3 Statistical Mechanics-Based Perturbation Theory.</td>
<td>34</td>
</tr>
<tr>
<td>2.2 Chain Equations of State for Polar Fluids</td>
<td>42</td>
</tr>
<tr>
<td>2.3 Quantum Theory</td>
<td>43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 3: EFFECTS OF TEMPERATURE AND SALT CONCENTRATION ON THE SOLUBILITY OF AMINO ACIDS AND DIPEPTIDES</th>
<th>46</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>46</td>
</tr>
<tr>
<td>3.2 Experimental Methods</td>
<td>50</td>
</tr>
<tr>
<td>3.3 Results and Discussion</td>
<td>53</td>
</tr>
</tbody>
</table>
3.3.1 Temperature Study: Amino Acids. 53
3.3.2 Temperature Study: Peptides. 61
3.3.3 Salt Type and Concentration Study: Amino Acids. 67
3.3.4 Salt Type and Concentration Study: Reverse Dipeptides. 71

3.4 Determination of Change in Enthalpy 77
3.5 Conclusions 81

CHAPTER 4: A DIPOLAR CHAIN EQUATION OF STATE 83
4.1 Introduction 83
4.2 The Statistical Associating Fluid Theory 86
4.2.1 The Reference Fluid and Dispersion Contributions 87
4.2.2 The Contribution from Associating Fluids 90
4.2.2.1 The fraction of monomers for fluids with multiple association sites 93
4.2.2.2 Extension to Mixtures of Associating Fluids 97
4.2.3 The Chain Term 98
4.2.4 Impact of SAFT 102
4.3 Comparison of the Molecular Sphere and Segment Approaches 102
4.4 Development of the Polar Term Applicable to Mixtures 107
4.5 Conclusions 111

CHAPTER 5: A PARAMETRIC ANALYSIS OF DIPOLAR SAFT 112
5.1 Introduction 112
5.2 Method

5.3 Parametric Analysis

5.4 Parameter Predictions

5.5 Effect of explicit inclusion of long-range polar interactions.

5.6 Functional Group Parameters

5.7 Conclusions

CHAPTER 6: DIPOLAR SAFT: APPLICATIONS

6.1 Introduction

6.2 Phase Equilibria in Acetone-Alkane Binary Systems

6.3 Phase Equilibria in Other Ketone/Alkane Systems

6.4 Phase Equilibria in Polar Copolymer-Solvent Systems

6.5 Conclusions

CHAPTER 7: CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

7.2 Future Work

REFERENCES

APPENDIX A: AMINO ACIDS

APPENDIX B: SOLUBILITY OF A SOLID AT INFINITE DILUTION
LIST OF TABLES

CHAPTER 3

Table 3.1. Solubility Temperature Dependence of Amino Acids and Dipeptides 55

Table 3.2. Salt Type and Concentration Dependence of Amino Acid and Dipeptide Solubility at Constant Temperature (25°C) 69

Table 3.3. Enthalpic Changes and Melting Temperatures 79

CHAPTER 5

Table 5.1 Parameters for the Polar CK-SAFT model for ketones. 114

Table 5.2 Parameters for the Polar PC-SAFT model for ketones. 114

Table 5.3 Predicted Polar CK-SAFT Parameters 127

CHAPTER 6

Table 6.1. CK-SAFT parameters for alkane solvents and ketones. 140

Table 6.2. Optimized binary interaction parameters. 146

Table 6.3. Binary interaction parameters for ketone/alkane mixtures. 149

Table 6.4. Molecular characterization of the polymers. 162

Table 6.5. Polar CK-SAFT Parameters for Polymers. 165

Table 6.6. Binary Interaction Parameters for Polymers in Solvents. 168
LIST OF FIGURES

CHAPTER 2

Figure 2.1. Dispersion potential energy as a function of distance with 0 being at the particle center. 37

CHAPTER 3

Figure 3.1. Structures of the amino acids a) glycine, b) alanine, and c) (α acid is charged) and d) (β acid is charged) aspartic acid. These structures are built with Molecule 3D using the bond lengths, bond angles, and torsional angles reported for optimized structures in the corresponding references. 49

Figure 3.2. a) Solubility (log(g solute/kg solvent)) of glycine (●), alanine (●), and aspartic acid (◆) as a function of temperature (°C) compared with data reported by Fasman (—). b) Solubility (g solute/kg solvent) of the dipeptides gly-asp (◆), asp-gly (■), β-asp-gly (▲) and aspartic acid (●) as a function of temperature (°C). 54

Figure 3.3. Structures of aspartic acid at pH 7.0: (the α carbon is in italics and the β carbon in bold. a) Predominant Form has a net charge of -1, b) the α acid is in a charged state, and c) the β acid is in a charged state. 57
Figure 3.4. Structures of the dipeptides a) gly-asp, b) asp-gly, and c) \(\beta\)-asp-gly. Built and optimized using Molecule 3D.

Figure 3.5. a) Solubility of glycine in NaCl (●) and KCl (■): Results of Khoshkbarchi and Vera at 25°C, Pfeiffer and Wugler at 21°C, and this work at 25°C. b) Solubility (g solute/kg solvent) of glycine at 25°C in aqueous solutions of NaCl (●), KCl (■), and LiCl (▲).

Figure 3.6. Solubility (g solute/kg solvent) of alanine at 25°C in aqueous solutions of NaCl (●), KCl (■). Comparison of DL-Alanine from Khoshkbarchi and Vera (open symbols) and L-Alanine of this work (filled symbols).

Figure 3.7. Solubility (g solute/kg solvent) of the dipeptides gly-asp (◇) and asp-gly (■) at 25°C in aqueous solutions of NaCl.

Figure 3.8. Solubility (g solute/kg solvent) of the dipeptides a) gly-ala and b) ala-gly at 25°C in aqueous solutions of NaCl (◇), KCl (■), and LiCl (▲).

Figure 3.9. Structures of the dipeptides a) gly-ala and b) ala-gly. Built and optimized using Molecule 3D.

Figure 3.10. Henry's Law (Log of mole fraction, \(x\), as a function of inverse temperature (1/T)) applied to the dipeptides gly-asp (▲), asp-gly (■), and \(\beta\)-asp-gly (◇).
CHAPTER 4

Figure 4.1 A cartoon representation of the free energy expansion. Associating molecules are essentially modeled by removing the association sites and covalent bonds to form a reference fluid of segments. The theory builds the associating molecules beginning with this reference fluid of segments, which is typically modeled as a hard sphere fluid with a dispersion contribution as a perturbation. The change in free energy due to chain formation and association are then added as perturbations to the segment free energy.

Figure 4.2 a) Segment volume based on a temperature independent diameter $\sigma$. b) $u$ is the segment-segment interaction energy, and c) $m$ represents the number of segments within a chain molecule.

Figure 4.3 Cartoon representation of possible association. While each site can bond only once, a given molecule can bond as many times as it has sites, thus forming clusters of molecules.

Figure 4.4 Large circles represent molecules. The smaller dark circles denote association sites. a) The square-well potential model with a hard sphere core and associating interaction energy of $-\epsilon$. b) If a site is bonded, it cannot bond with an additional site. c) Two molecules with association sites approaching each other in the correct orientation and within the minimum
distance will associate. (Adapted from Chapman.)

Figure 4.5  Cartoon representation of the covalent bonding scheme. By allowing each species to bond with only one or two other spheres and restricting how the spheres may bond, a chain is formed.

Figure 4.6.  Representative alkanone: a) segment approach, b) effective molecular sphere approach.

Figure 4.7  A cartoon representation of the free energy expansion for a polar-associating fluid. As before, the associating molecules (now with polar sites) are modeled by removing the association sites and covalent bonds to form a reference fluid of a mixture of polar and non-polar segments. The theory builds the polar-associating molecules beginning with this reference fluid of polar segments, which is now modeled as a hard sphere fluid with the polar and dispersion contributions as a perturbation.

CHAPTER 5

Figure 5.1.  The natural log of the vapor pressure of the ketones as a function of inverse temperature. The lines are least squares fit to the experimental data.

Figure 5.2.  a) The natural log of the vapor pressure of the 2-ketones as a function of inverse temperature. b) Same as a) with the 2-Octanone results replaced by the predicted results from
Polar CK-SAFT. The data is from Smith and Srivastava. 116

Figure 5.3. a) The natural log of the vapor pressure of the mid-ketones as a function of inverse temperature. b) Same as a) with the 4-Heptanone results replaced by the predicted results from Polar CK-SAFT. The data is from Smith and Srivastava. 117

Figure 5.4 Molecular volume as a function of molecular weight. a) CK-SAFT results for alkanes, alkanols, and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes and Polar PC-SAFT results for alkanols and ketones. 120

Figure 5.5 Chain length as a function of molecular weight. a) CK-SAFT results for alkanes, alkanols, and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes and Polar PC-SAFT results for alkanols and ketones. 122

Figure 5.6 Segment dispersion energy as a function of molecular weight. a) CK-SAFT results for alkanes, alkanols, and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes and Polar PC-SAFT results for alkanols and ketones. 123

Figure 5.7 Number of polar segments per chain as a function of molecular weight for the ketones. a) Polar CK-SAFT. b) Polar PC-SAFT. 125

Figure 5.8. Vapor pressure curve for 8-Pentadecanone. Comparison of experimental data and Polar CK-SAFT prediction. 128

Figure 5.9. Two possible conformations of 2,4-Pentadione. Built and
optimized using Molecule 3D.

Figure 5.10. Vapor pressure curve for 2,4-Pentadione. Comparison of experimental data and Polar CK-SAFT prediction.

Figure 5.11 a) Molecular volume and b) Chain length as a function of molecular weight. Comparison of CK-SAFT and Polar CK-SAFT results for ketones.

Figure 5.12 a) Segment dispersion energy and b) Molecular dispersion energy as a function of molecular weight. Comparison of CK-SAFT and Polar CK-SAFT results for ketones.

Figure 5.13 The change in chain length between two members in a homologous series. a) CK-SAFT results for alkanes, alkanols, and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes and Polar PC-SAFT results for ketones.

Figure 5.14 The change in molecular volume between two members in a homologous series. a) CK-SAFT results for alkanes, alkanols, and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes and Polar PC-SAFT results for ketones.

CHAPTER 6

Figure 6.1. Vapor-liquid equilibrium of acetone-pentane system. Polar SAFT (—) and SAFT (---) at $k_{12}=0$; Experimental data (●).

Figure 6.2. Vapor-liquid equilibrium of acetone-hexane system. Polar
SAFT (—) and SAFT (---) at \( k_{12} = 0 \); Experimental data (•).

Figure 6.3. Vapor-liquid equilibrium of acetone-decane system. Polar
SAFT (—) and SAFT (---) at \( k_{12} = 0 \); Experimental data (•).

Figure 6.4. Vapor-liquid equilibrium of acetone-dodecane system. Polar
SAFT (—) and SAFT (---) at \( k_{12} = 0 \); Experimental data (•).

Figure 6.5. Vapor-liquid equilibrium of acetone-hexane system. Polar
SAFT (—) and SAFT (---) at optimized \( k_{12} \); Experimental data (•). \( k_{12} \) is correlated to phase composition data.

Figure 6.6. Vapor-liquid equilibrium of acetone-dodecane system. Polar
SAFT (—) and SAFT (---) at optimized \( k_{12} \); Experimental data (•). \( k_{12} \) is correlated to phase composition data.

Figure 6.7. Vapor-liquid equilibrium of acetone-hexane system. Polar
PC-SAFT (—) at \( k_{12} \) and at optimized \( k_{12} \); Experimental data (•).

Figure 6.8. 2-Butanone in Heptane at 45 C. Comparison of experimental data and model predictions. a) CK-SAFT, b) Polar CK-SAFT, c) Polar PC-SAFT.

Figure 6.9. 2-Butanone in Decane at 70 °C. Comparison of experimental data and model predictions. a) CK-SAFT, b) Polar CK-SAFT, c) Polar PC-SAFT.

Figure 6.10. 3-Pentanone in Heptane. Comparison of experimental data and Polar CK-SAFT model predictions at 40.05 °C and 80.0 °C.
Figure 6.11. 3-Pentanone in Heptane. Comparison of experimental data and model predictions at 40.05 °C. a) CK-SAFT, b) Polar CK-SAFT, c) Polar PC-SAFT.

Figure 6.12. 3-Pentanone in Heptane. Comparison of experimental data and model predictions at 80.0 °C. a) CK-SAFT, b) Polar CK-SAFT, c) Polar PC-SAFT.

Figure 6.13. 4-Heptanone in Hexane. Comparison of experimental data and model predictions at 65.0 °C.

Figure 6.14. 5-Nonanone in Hexane. Comparison of experimental data and model predictions at 60.0 °C.

Figure 6.15. 2,4-Pentadione in Cyclohexane. Comparison of experimental data and model predictions at 25.0 °C.

Figure 6.16. Cloud point curves of 5 wt % polyethylene (PE), and EMA copolymers with methyl acrylate mol % of 31 (EMA₃₁) and 41 (EMA₄₁) in butane. Comparison of experimental data (•) and model predictions (—).

Figure 6.17. Cloud point curves of 5 wt % polyethylene (PE), and EMA copolymers with methyl acrylate mol % of 31 (EMA₃₁) and 10 (EMA₁₀) in a) propane and b) ethane. Comparison of experimental data (•) and model predictions (—).
CHAPTER 1: INTRODUCTION

The objectives of this work are to understand from a molecular perspective the behavior of small molecules in aqueous based separation processes. Structural analysis and thermodynamics are used to study the interactions of small biological molecules, such as amino acids and dipeptides, in aqueous solution. Further goals are, with the insight gained from this investigation, to progress toward a comprehensive theoretical model to be used in both the analysis of experimental data and the design of separation processes.

1.1 Why Study Aqueous Solutions of Amino Acids?

Water is the primary solvent for biologically active compounds. As such, an understanding of the physical and chemical characteristics of aqueous solutions of biochemicals is imperative for the pharmaceutical, food, agricultural, and other industries. The relationship of the thermodynamic properties of these systems, such as the dependence of the solubility of the biochemical (e.g. amino acid, peptide) on temperature and electrolyte type and concentration, are essential to understanding the complex interactions of the materials and in devising optimum separation processes for production of the desired biochemical.

When attempting to model solutions of large molecules, such as polymers or proteins, many approximations regarding the nature of the solvent and/or solute are often necessary. Numerous interactions may be occurring
simultaneously due to the presence of various forces. Some of these include Coulombic forces between ions and/or between permanent dipoles, quadrupoles, and higher multipoles; inductive forces of permanent dipoles or quadrupoles with induced dipoles; attractive and repulsive forces; and chemical association or complex formation (e.g. hydrogen bonding, ionic complexes). In addition, hydrophobic, hydrophilic, and surface interactions can be significant. Because of the large number of possible simultaneous interactions, many modelers seek to determine which of these forces contribute the largest to the thermodynamic properties of interest and which forces (or interactions) can be regarded as negligible. These of course vary with each type of system studied and, therefore, result in a need for a different set of assumptions, simplifications and often even the model for each type of system.

Many of the current models for biochemical systems are largely empirical with little or no theoretical basis. Most of the models are simple correlations of experimental data rather than being derived from first principles. However, with the advances in statistical mechanics and molecular simulation, newer models are being produced with sound theoretical basis, which incorporate contributions from several of these molecular level interactions.

One well-known separation process used for the partitioning of biological macromolecules is the Aqueous Two-Phase System (ATPS) which typically consists of two immiscible aqueous polymer solutions. In these systems, there are typically four or five substances and their interactions for which to account (two polymers, water, two or more types of solute to be separated, and often a
salt which is used to either control the system pH or influence the degree of separation (salting out/in)). Various approaches to modeling the system have surfaced. These include a wide variety from Flory-Huggins type approaches to activity coefficient models. While each model may reasonably capture some aspect of the separation process, it is difficult, at best, to capture the affects of the myriad two-body, three-body, and higher interactions.

Although the ATPS systems are widely used for some biological separation processes, a more fundamental knowledge of the interactions of the individual components is needed to gain a better understanding of the physics involved. This knowledge can then lead to the enhancement of existing processes and/or the development of new, more robust processes. For example, by beginning with a study of hydrocarbons or amino acids in water, the effects of the side chains, their sequence and interactions, and various other phenomena can be studied in a systematic way. An extension of the study to small peptides (i.e., dipeptides, tripeptides), provides a basis for comparison of the single amino acid (mono-peptide) with that of this same amino acid as a residue in a peptide chain. The information gained from this type of study can give insight into the behavior and characteristics of the larger macromolecular entity and bring clarity to the information already gathered regarding larger protein systems. While these small systems cannot form the secondary, tertiary, and higher structures often seen in the native conformations of proteins, they can provide information regarding the very basics from which these more structured systems are built.
To gain an in-depth understanding of any process requires looking at not only the "big picture" but also at the very minute details – at the molecular level. For example, how does hydrogen bonding of one type of side group with water affect the next member of the polymer or peptide chain? Does intramolecular hydrogen bonding occur and if so, how does it affect solution properties? What basic changes in the molecular interactions occur when a simple alkane is replaced with an alcohol in water? How does this affect solution thermodynamic properties? What occurs when the chain is longer, contains a polar group in the middle of the chain (e.g. ketone or ester), two or more polar groups are present (e.g. di-ketone, amide), polar groups of opposite charge are at the end termini with an additional intra-molecular polar group (e.g. aspartic acid)? Similarly, what occurs when a ketone is added to an amine, compared to the addition of an aldehyde, compared to the addition of an acid? By looking at these differences, the effects of individual functional groups can be isolated, as well as how these interactions change in the presence of other functional groups.

One of the largest time consumption areas in determining optimal operating conditions for separation processes such as aqueous two phase partitioning is in the use of the current “trial-and-error” approach.23,24 In order for industry to be able to design and operate these separation processes more cost effectively, a theoretical model is needed which will provide, as a first step, the information needed to guide the trial and error process. For example, if a researcher could sit at his/her computer and determine that potassium chloride in the amount of 0.1M would cause the desired peptide to salt out while 0.01M
would cause salting-in, he or she would then have an order of magnitude at which to begin the pilot experiments. Furthermore, since the separation processes for biochemicals sometimes represent up to 90% of their total manufacturing costs,\textsuperscript{25} any improvement to the separation process can result in substantial improvements in the economics for the production of these compounds.

The study of small peptide systems can provide much needed information for the biomedical and pharmaceutical industries. For example, RGD peptides are used for blocking platelet aggregation. Some as small as eight member cyclic peptides have been identified \((\text{CRGDxxxC, CxxxRGDC}).\textsuperscript{13}\) (Appendix A lists the twenty naturally occurring amino acids along with their symbolic representations.) Other studies are investigating the use of small peptides to target and/or modify specific metabolic activities. Analysis of the \textit{Conus} venoms revealed myriad diverse pharmacologically active small peptides, with lengths of the most active in the range of ten to thirty residues.\textsuperscript{26} The authors assert that these substances could prove "to have a pharmaceutical potential comparable to plant alkaloids" and suggest that these venoms be screened when searching for a ligand targeted for a particular receptor or ion channel. In order for these small peptides to be available commercially, effective separation processes must be developed or the peptides chemically synthesized, which is a very costly procedure.\textsuperscript{27}

Moreover, any insight into the effects of functional groups on solution properties or separation processes is considered invaluable to the chemical,
polymer, and petroleum industries. Each of these industries rely heavily on the ability to separate compounds based on various thermodynamic properties (e.g. distillation based on boiling temperature, two-phase separations based on solubility, etc.). In particular, the polymer industry seeks to add various functional groups to existing polymers in order to produce specific properties, for instance, in the making of carpet materials and plastics.

1.2 Separation Techniques and Aqueous Two-Phase Systems

Various separation techniques for biochemical molecules are available. Counter-current chromatography and distribution, electrophoresis, aqueous two-phase systems, liquid-liquid extraction, affinity chromatography, precipitation, reverse micelles, and membrane-based ligand separations are some of the techniques currently used for biochemical separations.\textsuperscript{27,28} Other systems of current interest include countercurrent chromatography based on the coil planet centrifugation principle,\textsuperscript{29} the use of ferric particles in one phase to induce a magnetic effect,\textsuperscript{30} application of an electric field,\textsuperscript{31} continuous countercurrent extraction (e.g. spray columns),\textsuperscript{32} and column partition chromatography.\textsuperscript{33} Each of these methods has advantages and disadvantages, which often vary with the specific biochemical system requiring separation. Other more traditional methods often utilize operating conditions (e.g., high temperature, shear rates, non-aqueous solvents, etc.) that result in significant loss in the biological activity of the desired molecule. Methods for separating large-scale volumes of amino acids or small peptides are very costly in terms of time and the number of
separations that must be completed in order to retrieve a pure sample. Typically, it is the end stage in the purification process that requires the highest degree of selectivity where target molecules must be separated from molecules with very similar structures and properties. Properties such as the chirality or amino acid sequence can be the objective of the end stage process.\textsuperscript{28}

Therefore, a method that is both biocompatible and highly selective is needed, especially for the end stage of the bio-separation process. Since the aqueous two-phase systems (ATPS) are one of the most promising separation techniques, an in-depth look at the work accomplished for this method is worthwhile.

When two mutually incompatible components such as two polymers (e.g. dextran (Dex) and poly(ethylene glycol) (PEG) or Dex and Ficoll (Fic)) or a polymer and a salt (e.g. PEG and sodium phosphate) are dissolved together in water, an aqueous two-phase system (ATPS) results. The phase separation occurs when the concentrations of the two components are above the phase boundary.\textsuperscript{34} Since the resulting two-phase system is normally eighty percent or more water, the term “aqueous two-phase system” has been coined. The resulting system is usually suitable for the separation of biological materials, from enzymes and nucleic acids to cells and viruses, without loss of biological activity.\textsuperscript{35} The measure of separation is the partition coefficient, which is the ratio of the concentration of the solute in the upper phase to its concentration in the lower phase.\textsuperscript{34} For a given mixture of biomolecules, the partition coefficient of each type of biomolecule is unique.
A brief overview of the development of aqueous two-phase systems as a means of separation follows and is based on the more detailed historical perspective of Albertsson.36 Aqueous two-phase systems (ATPS) have appeared in the literature since the late nineteenth century beginning with Beijerinck's (1896, 1910) work. Beijerinck, a Dutch microbiologist, described the formation of an aqueous two-phase system of gelatin and agar with most of the agar in the bottom phase and most of the gelatin in the top phase. Dobry and Boyer-Kawenoki reported phase separations of polymer mixtures in both aqueous and organic solvents in 1947. They described a systematic study on the miscibility of seventy-eight polymer-polymer-solvent systems and concluded that phase separation in these systems was the norm and complete miscibility the exception.

Albertsson (1956) was the first to utilize an aqueous two-phase system to partition biomolecules. It was actually a “mistake” that began Albertsson’s pioneering work in this area. He decided to try adding a detergent to his experiments in which he was attempting to separate chloroplasts. In a chromatographic experiment, the chloroplast particles adsorbed to hydroxylapatite but could not be de-sorbed. Recalling the name poly(ethylene glycol) (PEG) from a book on detergents, and since PEG was available in his lab, he tried it. He combined an aqueous solution of PEG and the hydroxylapatite sediment from the chromatographic experiment in 1M potassium phosphate buffer. The result: a clean separation of the chloroplasts into the PEG rich phase. Of course, he later found that PEG isn't a detergent but is used in the
manufacture of detergents. Continued experimentation with the ATPS showed that the partitioning is selective and particles collect either in one of the two bulk phases or at the interface (Albertsson, 1956). It was the earlier work of Bronsted and De Courdes that influenced Albertsson in the interpretation of his experimental results. Bronsted observed that there is a direct relationship between the molecular weight of the solutes and the level of partitioning achieved in a two liquid phase system (1931). The partition coefficients for cell particles, which are of very high molecular weight, are either zero or infinity. Hence, if all the cell particles are identical, then it is expected that all of the particles would separate to a single phase. However, in practice adsorption at the interface does occur in some cases. This is related to the particle stability, which depends on the interfacial tensions between the particle and the two phases and on the interfacial tension between the two phases (De Courdes, 1898).

In an attempt to find a more "physiological" system, Albertsson tested the use of a second polymer, for example, dextran, in place of the potassium phosphate. Over the course of the next several years, Albertsson studied various effects on partitioning of cell particles and proteins in polymer ATPS. These include 1) the influence of the molecular weight of the polymers (Albertsson, 1958a,b) and 2) the effect of ionic composition (Albertsson and Nyns, 1959, 1961). 3) He determined the linear relation between the surface area of proteins and viruses and that of the logarithm of the partition coefficient (Albertsson, 1958b, 1962; Albertsson and Frick, 1960). 4) He also considered the linear relation between the sedimentation constant of nucleic acids, as well as 5)
the logarithm of the partition coefficient (Lif et al., 1961). In 1960, Albertsson published a monograph based on his Ph.D. work, which contains the phase diagrams of many ATPS.

Walter and colleagues studied the effects of electrophoretic mobility (Walter et al., 1965, 1970a) and polymer charge and type of salt on partitioning (Walter and Selby, 1967; Walter et al., 1968). Johansson (1970) showed that salts did not partition equally between the phases as had been assumed in the earlier work but that significant differences, although small, occurred between similar salts like lithium, sodium, and potassium chloride. Albertsson (1971) concluded that this unequal partitioning of the salts created an electrical potential between the phases and substantially affects partitioning of many biomolecules, which often contain a significant number of charges. This electrostatic potential difference results in the preferential partitioning of negatively charged solute into the top phase and positively charged solute into the bottom phase for PEG-dextran systems (Johansson, 1970; Reitherman et al., 1973; King et al., 1988).\(^{37}\) Johansson (1971) showed that there exists a direct relationship between the logarithm of the partition coefficient and the net charge for proteins (determined by titration), while Reitherman and Johansson used electrodes to estimate the potential (Reitherman et al., 1973; Johansson, 1974). Therefore, cross partitioning could be used to determine the isoelectric point of proteins (Albertsson, 1970; Albertsson et al., 1970) and for the estimation of net charge of proteins (Blomquist, 1976). Johansson (1970b) showed that a much higher selectivity in separation of proteins could be attained by increasing the interfacial
potential through covalent binding of charged groups to PEG. The exponential relationship between the partition coefficient and both the size of molecules and net charge of molecules is a strong indication of the high degree of selectivity the method affords. Various additional experimental work using partitioning includes the following: a) separations: virus mutants (Bengtsson et al., 1962; Walter et al., 1970a), chloroplasts (Albertsson and Baltscheffsky, 1963; Karlstam and Albertsson, 1969; Larsson et al., 1971), erythrocytes of different age in the cells' life span (Walter et al., 1964, 1965; Walter and Selby, 1966), Chlorella cells in different stages of the growth cycle (Walter et al., 1966; Brunette et al., 1968); b) membrane modification of cells (Walter and Coyle, 1968); and c) comparison with electrophoretic mobility (Brooks et al., 1971). These studies have clearly shown the surface dependence of the partitioning method and affirmed its viability as a separation technique. The high water content and low interfacial tension (Ryden and Albertsson, 1971) as well as the general protective effect of the polymers (Albertsson, 1985) are characteristics of the polymer ATPS which make the system amenable to fragile organelles and cells.

The advantages of utilizing an aqueous two-phase system for biochemical separation include the following. 1) Aqueous systems provide a highly biocompatible environment. 2) The systems have low interfacial tension (several orders of magnitude lower than organic-organic systems or aqueous-organic systems). 3) They can be made isotonic if needed. 4) The resolution and yields are good. 5) They have high capacity and 6) linear scale-up. 7) Yields and resolution can be increased by the addition of an affinity ligand when needed. 8)
They are inexpensive, time and labor effective;\textsuperscript{35} and 9) provide unique information as well as being highly sensitive.\textsuperscript{35} Indeed, slight surface modifications of cells that occur during \textit{in vivo} processes or \textit{in vitro} treatments can be detected by doing multiple extractions (e.g., counter-current distribution).\textsuperscript{5}

The disadvantages include a lack of understanding of the mechanisms involved in the partitioning processes in ATPS and the generality of the information gained from the technique. While the method will separate, for example, two different samples of glycoproteins it does not provide specific information on what the differences are in the glycoproteins such as backbone structures.\textsuperscript{35}

Various applications of ATPS to separation processes have been studied. These include biomolecule purification, affinity partitioning, extractive bioconversion, and liquid-liquid partition chromatography.\textsuperscript{16} For example, Kula and co-workers (Kula, \textit{et al.}, 1982; Kroner, \textit{et al.}, 1978; Kula, 1979) have shown that it is beneficial and economical to replace centrifugation (high shear) with PEG-salt-water two-phase systems for the removal of cell debris in both intra- and extra-cellular biological purification processes. Additional analytical applications for the ATPS technique include the following: 1) The study of hydrophobicity of biopharmaceuticals, which would allow the exploration of their corresponding quantitative structure-activity relationships and targeted modifications of the molecules; 2) Clinical biochemistry for diagnosis and treatment monitoring, for example in the diagnosis of breast tumor malignancy; 3) In monitoring the manufacture of recombinant protein; and 4) In the study of
chemical modifications of proteins and other biopolymers. The most frequent biotechnological use of the aqueous polymer two-phase systems is in the downstream processing of enzymes, in the purification of fermentation products, and large-scale recovery.

Experiments. While there has been significant interest in aqueous two-phase systems among experimentalists and theoreticians, there have been very few systematic experimental studies to provide the needed data with which to develop models for the separation processes. The most noted studies for small peptides (di, tri) include the work of Grossmann and co-workers, Diamond and co-workers, and Khoshkbarchi and Vera (aqueous solutions with electrolytes). The work by Grossmann and Tintinger is one of the more complete studies. They have determined partition coefficients in an ATPS of PEG and Dextran for four amino acids (gly, glu, phe, lys), their mono-dipeptides and tripeptides (e.g., glu-glu, gly-gly-gly), nine combination dipeptides (e.g., gly-glu, gly-phe, glu-gly) and seven combination tripeptides (e.g., gly-gly-phe, gly-phe-gly).

The partition coefficient is affected by many physical parameters including solution properties, polymer characteristics, the presence of salts and their characteristics, the elective presence of affinity ligands, as well as the characteristics of the solute itself. Solution properties that affect the partitioning include temperature, pH, osmotic pressure, interfacial tension and hydrophobicity. The molecular weight, charge, and concentration of each of the polymers used in the ATPS influence the partitioning. Similarly, the type and
concentration of salt in the system affect the separation achieved. For the solute, the charge, molecular weight, type and concentration are important properties in designing an appropriate aqueous two-phase separation system. It has also been proposed that an electrostatic potential often exists between the two-phases due to the presence of buffering ions and/or the unequal partitioning of salts (Johansson, 1970), which will also influence the partitioning.

Several studies have been done to determine the effects of various conditions on the partition coefficients of numerous proteins, peptides, or other biological molecules. A choice of polymers and/or salts with which to create the ATPS requires some investigation. For example, the partitioning of small peptides can be separated in PEG/dextran systems or in a PEG/MgSO₄ system. However, the latter is preferred for the small peptides when its desirable to magnify the effect of the side groups on the amino acid residues. The PEG/MgSO₄ system provides a high change in the free energy of methylene group transfer between the phases since the components are quite dissimilar. The salt also acts somewhat as a buffer so that the system pH is nearly independent of different concentrations and thus prevents the need for additional salts. In PEG/salt systems, most proteins favor the lower salt-rich phase.

Various methods are used to determine the polymer phase concentration. For example, freeze drying each phase followed by extraction of PEG from the residue using warm acetone for a PEG/MgSO₄ system, high performance liquid chromatography (HPLC) (e.g. for the determination of PEG concentration of a PEG/dextran system), and size exclusion HPLC (SEC) are used. Dextran
concentration can also be determined using a polarimeter\textsuperscript{38} at an optical rotation of 1.99 (\%w/v dm).\textsuperscript{37}

To determine solute concentration, the sample extracted from each phase is diluted and analyzed by UV absorbance using a spectrophotometer.\textsuperscript{37,38} Reverse-phase chromatography and a derivatization method\textsuperscript{49} are also sometimes used.\textsuperscript{6} According to Kula (1979),\textsuperscript{38} the protein partition coefficient is independent of its own concentration up to approximately thirty weight percent. Salt partition coefficients can be determined via liquid scintillation.\textsuperscript{38} King et al. (1988)\textsuperscript{38} used Low-Angle Laser Light-Scattering (LALLS) to determine interaction parameters (osmotic virial coefficients) and voltammetry to determine electric-potential differences between the phases.

Eiteman studied the pH difference in a PEG/citrate ATPS and the influence of sodium chloride.\textsuperscript{50} He found that the pH difference between the two phases increased with increasing PEG concentration regardless of whether or not the salt was added. However, the degree of increase was substantially less when the two percent sodium chloride was added. The pH difference between the phases was positive for all variations of the system (the pH was greater in the PEG phase than the citrate phase). The same trend was seen in the PEG/phosphate system studied by Eiteman and Gainer.\textsuperscript{7} Partition coefficients for oppositely charged analogous compounds, indole 3-acetic acid and tryptamine, increased with increasing pH with or without the addition of salt. However, the negatively charged compound had partition coefficients up to four times that of the positively charged one in the no salt system. The magnitude of
the difference in the partition coefficients is directly related to the pH, or tie-line length. In the presence of the two percent sodium chloride system, the partition coefficients of the negatively charged component were only slightly greater than the positively charged one. This was due primarily to a significant increase in the partition coefficient of the positively charged amine in the presence of the salt.

Temperature can be used to affect both the formation of two-phase systems and the economics of the system. For example, some systems of PEG and salts, such as sodium chloride, do not exclude PEG at room temperature but do at elevated temperatures. Some copolymers of ethylene oxide and propylene oxide have an upper cloud point in water that allows them to be recovered after use in a two-phase system.\(^5\)

**Hydrophobicity.** Eiteman and Gainer determined the hydrophobicity of several amino acids and short peptides ranging from dipeptides to a pentapeptide.\(^6\) They intentionally omitted from their study amino acids that contain a charged side chain. They also developed a correlation between the partition coefficients and the hydrophobicity of the solute in a PEG-MgSO\(_4\) ATPS.

Polymer fractionation, the differing of the average molecular weight of a polymer in the two phases, is thought to influence partitioning.\(^3\) The higher the molecular weight of PEG, the greater is its partitioning coefficient. Since polymer molecular weight has been shown to affect the system partitioning capabilities, it is inferred that polymer fractionation affects not only its own partitioning coefficient but also that of the system.
The partition coefficients of neutral solutes are directly proportional to the tie-line length. Therefore, addition of a salt such as sodium chloride is expected to increase the partition coefficient of a neutral solute since it increases the tie-line length (given other conditions remain constant and the added salt replaces an equal amount of water). The effect of the salt on partitioning depends in part on the characteristics of the solute itself. Depending on the type of solute, the charge of the solute may vary with pH. For example, proteins or amino acids, along with numerous other biological molecules, will be positive, negative, or neutral depending on the solution pH relative to the isoelectric point of the protein or amino acid. For some solutes, the effect of the salt is related to electrostatic interactions (longer-range Coulombic forces) and for others it is related to the hydrophobicity (double layer interactions). For example, the hydrophobic characteristic of partitioning of vancomycin, a glycopeptide antibiotic, appears to be affected by the presence of a high concentration of water structure-making salts in the ATPS. "Water making salts" are defined as those salts which dissolved in water result in a greater organization of the water molecules relative to the normal water structure. Furthermore, these salts are known to increase hydrophobic interactions in polymer-polymer ATPS.

Selective partitioning through the use of affinity ligands, referred to as affinity partitioning, is of significant current interest. An affinity ligand for a specific protein is introduced into one of the two phases, usually through binding to a polymer. The structure of the ligand-carrying polymer and the nature of the phase-forming polymers influence the efficacy of a ligand in binding and
extracting a protein. The use of a polymer-polymer versus a polymer-salt system depends on the type of interaction between the ligand and the protein of interest. For electrostatic interactions in the ligand-protein binding, the polymer-polymer systems are generally preferred since the affinity partitioning is very sensitive to salt concentration. If the binding is due to a hydrophobic effect, then high salt concentrations enhance ligand-protein interaction and the salt-polymer systems are used. Metal chelates as affinity ligands have been studied and shown to have significant potential. The chelate is added by binding iminodiacetic acid (IDA) to PEG and then loaded with divalent ions of the transition elements. For example, proteins with clustered histidine residues can be effectively extracted with copper chelate bound to one of the phase polymers.

Other considerations, especially for large-scale applications, include the cost of the polymer and disposal requirements. For example, dextran is a more expensive polymer and is sometimes replaced by less expensive polysaccharides such as pullulan, dextrin, maltodextrin, and ethylhydroxyethylcellulose (EHEC). The PEG/sodium phosphate system has the undesirable attribute of waste phosphate. The disposal expenditure is greater for phosphate, due to environmental concerns, than for some other salts such as sodium sulfate, magnesium sulfate, potassium carbonate and citrate which have been used in ATPS.

Apparatuses for large-scale applications of two-phase extraction originally designed for other systems, for example, aqueous-organic solvents, can often be used for the aqueous polymer-polymer two-phase systems depending on the
viscosity of the phases and settling times.\textsuperscript{5} Commercially, centrifugal separators and mixer-settler systems have been used in ATPS successfully.\textsuperscript{24} On a laboratory scale, the primary apparatus for multi-step partitioning has been the thin-layer counter-current distribution (CCD) system designed by Albertsson.\textsuperscript{34}

1.3 Models of ATPS and Aqueous Systems for Peptides and Amino Acids

A better understanding of the aqueous two-phase systems and partitioning phenomena is expected to lead to more efficient separations and more economical processes.\textsuperscript{5} Most of the theoretical developments for the polymer-polymer ATPS\textsuperscript{2-4} have their basis in the Flory-Huggins lattice theory,\textsuperscript{52} which traditionally has been applied to polymer solutions in unstructured solvents for the prediction of general properties. Other approaches include osmotic virial expansions,\textsuperscript{38} UNIQUAC,\textsuperscript{17} and a statistical mechanics based model.\textsuperscript{19} For polymer-salt ATPS, approaches have included fluctuation theory\textsuperscript{18} and exclusion mechanisms.\textsuperscript{38} Aqueous solutions of amino acids and small peptides have been modeled using primitive perturbation models\textsuperscript{20,21} and chain equations of state accounting for polar contributions using an effective sphere approach.\textsuperscript{22}

Eiteman and Gainer\textsuperscript{6} developed a model to relate hydrophobicity of the solute with its partitioning coefficient for a PEG/MgSO\textsubscript{4} ATPS. The model predicted most trends well for the experimental data with which it was compared. They limited their study to amino acids that do not have a charged side group in an attempt to isolate the hydrophobic effect. The beginning point for their model was the correlation that Zaslavsky \textit{et al.}\textsuperscript{10,11} established for relating partition
coefficients of compounds of a homologous series. Zalavsky's work was based on the difference in free energy in each compound in the homologous series. Although Eiteman and Gainer began a systematic study of amino acids using ATPS, the complexity of the system prevented the development of a theoretically based model.

Eiteman and Gainer \(^7\) also studied the effect of pH difference between the phases on partitioning in an ATPS of PEG and phosphate. Since the charge of proteins as well as other biomolecules depends on the system pH, a pH difference between the phases is expected to affect partitioning. To model this affect, they developed an equation that relates the partition ratio to the pK\(_a\) of the acid, the pH difference between the phases, and the activity coefficient ratios in both phases. They tested the model by comparing the effect of phase pH on the partitioning of acids and corresponding alcohols. At a given pH, the differences in partitioning coefficients, according to Eiteman and Gainer, should only be related to the presence of the charge on the acid. The model predicts that the acids and alcohols will have the same partition ratio at low pH and that the partition ratio for acids will increase with increasing pH at a faster rate than for the alcohols. These trends were observed experimentally. They concluded that the pH difference between the phases in the PEG/phosphate system could account for differences in partition coefficients between charged and uncharged solutes. Eiteman and Gainer also developed more general models\(^8,9\) for predicting the partition coefficients in ATPS.
1.4 Overview of our Approach

The present study of the separation of biological molecules is from a very fundamental aspect. While it is desirable to accurately model ATPS, it is necessary to first understand the basis on which the success of the method lies. In other words, in order to understand what makes a clock "tick", one must consider both the "whole" mechanism and the intricate workings of the individual "parts". While the "whole" has previously been considered in some detail, this work focuses on the individual "parts". Considering only an aqueous solution, with or without electrolytes, eliminates numerous variables and interactions from the system. Since most of the biochemical separation processes are aqueous-based, and in particular ATPS, the properties of these molecules in, and their interactions with, water need to be fully understood as a basis for extending the work to more complicated systems. The contribution of the individual amino acids and/or functional groups to the system properties can be determined through a systematic investigation of aqueous solutions of amino acids and more complex systems, such as using longer chains with and without additional side groups. Based on this type of study, insight into the specific molecular interactions as well as how, and to what degree, each of these interactions contribute to the various solution properties can be gained.

In cooperation with other research groups, a systematic experimental study of aqueous solutions of amino acids and dipeptides has been initiated. Through this study the details of the physical system give insight to the molecular interactions that we seek to model. A thermodynamic and structural analysis of
this new data for amino acids and dipeptides in solution is presented. Recent
quantum studies for amino acids and small peptides also assist in the
interpretation of the experimental results.

One of the strongest contributors to the thermodynamic properties of
these systems is the polar interactions (i.e. dipole – dipole). A new approach to
include polar interactions in a statistical mechanics based model for fluid
properties is presented. The theoretical study begins with molecules containing a
single dipolar group internal to the molecule and is then extended to multiple
internal polar groups. As a more difficult test of this stage of the theoretical
development, the model is applied to solutions of polar co-polymers in non-polar
solvents. This is the first step towards developing a theoretically sound model for
the more complicated ATPS systems.

In the following chapters, the results and analysis of the experimental
studies are detailed (chapter 3); the theoretical development of a segment-based
approach to the explicit inclusion of dipolar effects is given along with a
comparison to previous approaches in chapter 4. In chapter 5, a parametric
analysis is presented for the polar model using two forms of the well-known
Statistical Associating Fluid Theory (SAFT). Results with CK-SAFT\textsuperscript{53-56} are
compared with PC-SAFT\textsuperscript{57-59} The two forms differ in their approach to the
dispersion interactions. Applications of the segment-based approach to binary
mixtures of alkanes with ketones or di-ketones as well as a polar co-polymer in
an alkane solvent are given in chapter 6. In chapter 7, the conclusions of this
work and directions for future related research are outlined. In order to aid the
reader in understanding the concepts put forth in this work, a brief review of statistical mechanics and quantum theory is needed. This is where we begin in chapter 2.
CHAPTER 2: BACKGROUND INFORMATION

2.1 Thermodynamics and Statistical Mechanics

The development of classical thermodynamic laws has been independent of any reliance on the existence of matter in the form of particles, atoms, etc. Rather, classical thermodynamics, based on fundamental postulates, provides relationships between various bulk system properties (e.g., temperature, pressure, and volume) without a direct means of calculating these properties. Statistical mechanics strives to ascertain the origin of these laws based on fundamental atomic theory and provide theories and models for the direct calculation of thermodynamic properties. For instance, statistical mechanics demonstrates how the laws of thermodynamics are a result of the quantum theory postulates in conjunction with a statistical postulate. Although not available in pure classical thermodynamics, means of computing properties such as heat capacities, free energies, etc. from spectroscopic data, are found in statistical mechanics.62

When considering a closed system in classical thermodynamics, the equilibrium state is one in which the state (macroscopic) properties are constant in time. However, to view this same system from the statistical point of view, it must be realized that this equilibrium state is the resulting average from a very dynamic molecular environment. For example, the macroscopic properties such as temperature, pressure and entropy are the average value of multiple thousands of molecular movements (i.e., the system temperature is a result of
the average kinetic energy of the numerous molecules that comprise the substance of interest). These movements give rise to what is referred to as fluctuations in the thermodynamic properties at the molecular level. While for a bulk system these fluctuations from the mean are unappreciable, for a very small system such as a colloidal particle, the fluctuations are observable (e.g. Brownian motion).  

2.1.1 *Determining molecular properties of condensed fluids with statistical mechanics.*

We know that matter is comprised of molecules in constant motion and that it is the combined interactions of the particles attraction, repulsion, and motion that determine the properties of fluids. The relative degree of these forces allows a given compound to change from gas to liquid to solid, and vice versa. For example, vapor-liquid phase transitions are absent when only repulsive forces are present. Thus, the attractive forces are entirely responsible for these phase transitions. Attractive forces have little effect on the pair correlation function, \( g(r) \), at high densities. However, when fluids are near the vapor-liquid critical point, and hence less dense, the attractive forces have a significant effect on the fluid structure, \( g(r) \). Nevertheless, they are responsible for large contributions to thermodynamic functions and dynamical properties at high densities. Therefore, in order to accurately model a liquid, both attractive and repulsive forces must be taken into account. Of course, there are numerous state conditions that also affect the phase changes (e.g., temperature, pressure).
Here the focus is on the liquid state and thus the concerns are with properly accounting for the intermolecular forces that make a liquid remain in the liquid state rather than solid or gaseous state.

Early attempts to model condensed phases applied such methods as the virial expansion around a “dense gas” and lattice theories. However, these methods are limited in that they cannot capture the true physics of the liquid state. A virial expansion around a dense gas fails to account sufficiently for the attractive forces needed to form the liquid state. Lattice theories limit the movement of a particle to a small cell or lattice point by the repulsive forces of its neighbors, essentially overcompensating for the repulsion. The lattice-based theories may be applicable to very viscous liquids but are not a realistic model for most liquids.\textsuperscript{64}

There are several statistical mechanics approaches that have been developed during the last century for modeling liquids, including integral equation methods, perturbation theories, scaled-particle theory, as well as molecular simulations.\textsuperscript{64} For example, Kirkwood and Yvon (1930's) introduced the integral equation methods. This more fundamental approach begins with an exact equation for the molecular distribution function of interest (e.g. the pair function). Approximations are then introduced, typically motivated by mathematical simplicity, to effect a solution. The validity of the approximations relies on a posteriori agreement with either computer simulation or experimental data. The Yvon, Born, and Green (YBG), the Percus and Yevick (PY), and the hypernetted chain (HNC) approximation are some of the integral equation methods that have
been developed. By providing the distribution functions directly, these methods are applicable to a wide variety of properties.\textsuperscript{64}

In thermodynamic perturbation theory, the properties of the fluid of interest, with an intermolecular potential energy $U$, are related to the properties of a reference fluid, with a potential $U_0$, through an appropriate expansion. A reference fluid is chosen that will most closely model the real fluid and one for which the properties are well known, either from molecular simulation or integral equation theory. These theories give excellent quantitative results for liquids and dense gases with the exception of the region near the critical point.\textsuperscript{64} The most commonly used potential models include the hard sphere model, with a hard sphere diameter $d$, given by

$$U = \begin{cases} \infty & r \leq d \\ 0 & r > d \end{cases}$$  \hspace{1cm} (2.1)

the square well model which is expressed as

$$U = \begin{cases} \infty & r \leq d \\ \varepsilon & d \leq r \leq Rd \\ 0 & r > Rd \end{cases} ,$$  \hspace{1cm} (2.2)

where $\varepsilon$ is the potential well depth and $d(R-1)$ is the potential well width; and the Lennard Jones (LJ) potential model with a characteristic diameter of $\sigma$ and potential well depth of $\varepsilon$:

$$U = 4\varepsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right] .$$  \hspace{1cm} (2.3)

Scaled-particle theory (Reiss, et al., 1959, 1965) is another statistical mechanics based approach developed to study liquids.\textsuperscript{64} Since it does not
produce the molecular distribution functions, scaled-particle theory is not considered a complete theory. However for hard spherical or convex molecules, it gives good results for the thermodynamic properties.

Molecular simulations, either Monte Carlo or molecular dynamics, are invaluable to the development of new statistical mechanics based theories. These simulations are computer experiments with precisely defined intermolecular potentials used to determine fluid structure, compressibility factors, or other properties of interest. Since the same intermolecular potential model can be used in both theory and simulation, the simulation results are used to test the validity of the theory and its assumptions. Furthermore, to test the accuracy and/or validity of the intermolecular potential model, the simulation results can be compared with experimental data of real fluids. Comparison of the theory with experimental data alone is an insufficient test of the theory since the theory includes both the intermolecular potential model and the theoretical assumptions. Hence, simulation provides a way to rigorously test the theory.

2.1.2 Brief Historical Overview of Thermodynamic Perturbation Theories for Atomic and Molecular Fluids

Perturbation theories of liquids have developed using two primary approaches: expanding in terms of the free energy and expansions for the distribution functions. These approaches have developed along similar lines with applications to atomic fluids and to molecular fluids. The following is a brief overview of these developments.
In 1951, Longuet-Higgins related the free energy of a non-ideal solution to an ideal one, with the free energy expanded in powers of the potential difference between the real and ideal solution. This theory was applied to liquids and gave good results only for weakly non-ideal solutions. In 1954, Zwanzig related the free energy of a dense atomic fluid to that of a hard sphere fluid through the potential. Considering the fluid as one interacting with an isotropic pair potential \( u(r) \), he related this to the potential of a hard sphere fluid \( u_0 \) by defining the potential as \( u = u_0 + u_1 \), and then expanding in powers of \( u_1/kT \). When applied to calculating the equation of state for a Lennard-Jones (12,6) fluid, he obtained good results at high temperatures. However, Zwanzig's approach does not properly account for the softness of the repulsive part of the intermolecular potential. Subsequently, Rowlinson used an expansion in powers of a softness parameter, \( n^{-1} \), about \( n^{-1} = 0 \) for a potential of the form \( u(r) = f(r^{-n}) \). While this method did give some improvement over Zwanzig's method in accounting for the repulsion, it failed to give good results for the attractive part of the potential. Barker and Henderson\(^{67,68}\) (1967) used a second-order perturbation of the potential, which gave reasonable quantitative results for the thermodynamic properties of a Lennard-Jones (LJ) liquid. Weeks, Chandler, and Anderson\(^{69}\) (1976) used the repulsive part of the full potential as their reference system and performed a first order perturbation about this reference fluid. They then related the properties of the reference fluid to a hard sphere fluid using a first order perturbation theory. Their method gives good results for dense liquids with the first-order theory and converges more rapidly than Barker-Henderson's.
McQuarrie and Katz (1966) introduced the use of the $r^{-12}$ term from the LJ potential as a reference system.\(^{64}\)

For the pair correlation function expansion, we begin with the work of Kirkwood et al. (1952), who applied perturbation theory to the Born-Green integral equation for the pair correlation function $g(r)$. Buff and Schindler (1958) derived an exact expression for the first-order perturbation term in the distribution function expansion of order $h$. Then, along with the suggestion that repulsive forces determine fluid structure at high densities, Weeks et al.\(^{65}\) proposed a theory of $g(r)$ for fluids with purely repulsive forces. Smith et al. (1971) and Gubbins et al. (1971) applied perturbation expansions of $g(r)$ to fluids with both attractive and repulsive potentials.\(^{64}\)

Now we turn to molecular fluids. In 1951, Barker expanded the partition function for a polar fluid about that for an isotropic molecular fluid. When neglecting induction interactions, the first-order perturbation term vanishes. Barker derived the second-order free energy term for the dipole-dipole interaction and demonstrated how to express it in terms of the reference fluid thermodynamic properties. In Pople's treatment of polar liquids, he evaluated the perturbation terms within the framework of cell theory. Additionally, he derived the expressions for the second-order perturbation terms for multipole interactions up to quadrupole. The extension of this work to include polarizable dipoles was completed by Barker. In 1953, Pople gave the expression for the second-order contribution to the Helmholtz free energy ($A_2$) in terms of the reference fluid distribution functions for multipole-like potentials. At this same time, Cook and
Rowlinson (1953) noted that the free energy for some types of orientation dependent potentials (e.g., dipole-dipole) can be equated to that of a fluid of isotropic molecules, where these molecules interact with an effective temperature dependent potential. In 1954, Pople presented the rigorous expression for the second order term in the Helmholtz perturbation expansion \( A_2 \) for an intermolecular potential of general form. This expression included both two- and three-body integrals over the reference system distribution functions. For multipole-like potentials the three-body term vanishes.\(^{64}\) Gubbins and Gray\(^{70}\) (1972) later introduced perturbation theory for the angular pair correlation function, using an approach similar to Pople's. A quite successful approach is the Padé approximant to the Barker-Pople free energy series proposed by Stell \textit{et al.}\(^{71,72}\) in 1973.

There are a variety of approaches to modeling molecular fluids. While most approaches utilize a spherical reference fluid as discussed above, an expansion about a non-spherical reference potential can be made. While this type of expansion generally converges more rapidly, it is typically more difficult to evaluate numerically.\(^{64}\) In the mid-1980's, Wertheim developed a thermo-dynamic perturbation theory (TPT) for associating fluids.\(^{73-76}\) Wertheim approached the problem of associating fluids by considering monomers, dimers, and higher-mers as distinct chemical entities and gave precise statistical-mechanical definitions of what constitutes a monomer, dimer, etc.\(^{77}\) In the last half of the 1980's, Chapman developed an equation of state for associating fluids based on Wertheim's TPT.\(^{53-56,77,78}\) This equation of state, known as the
Statistical Associating Fluid Theory (SAFT), was the first to account for both molecular shape (non-sphericity) and association effects. SAFT is an extension of the van der Waals type equation of state to associating fluids and is accurate even in the limit of spheres bonding to form a disphere fluid. A detailed review of SAFT applications was recently published by Müller and Gubbins.

Some of the more recent applications of SAFT are of specific interest to the study here. For example, some studies have been done with carboxylic acids where the electrostatic interactions are accounted through association. They developed a procedure to relate the SAFT association parameters ($u^0/k$, $\varepsilon^{AB}/k$) to the entropic and enthalpic changes of dimerization determined from quantum mechanical methods and using only low-pressure data. They were able to relate the molar volumes determined from molecular orbital calculations to the segment size ($v_{oo}$) and chain length ($m$) SAFT parameters. By doing so, they reduced the number of parameters needed for associating systems from five to three. However, the relationship between the quantum mechanical results and the SAFT parameters are dependent on the specific association scheme and thus must be derived independently for each system studied.

Other studies have considered aqueous solutions of alkanolamines and also modeled the electrostatic interactions as associations. Differing from prior approaches, they chose to use four interaction sites for water and, for simplicity, used only one association site for the two sets of lone pairs on the alkanolamine oxygen. Since the alkanolamines have one or more alcohol functional groups and an amine functional group, they are capable of forming
various hydrogen bonds, both with the nitrogen and oxygen atoms. However, this distinction is not made in order “to limit the number of adjustable parameters”. They applied SAFT to cross-associating mixtures and found reasonable agreement with experimental data for the systems studied. In addition, they determined improved parameter values for carbon dioxide and water.

Khoshkbarchi and Vera\textsuperscript{21} developed a very simple model for aqueous solutions of amino acids using dipole moments determined from a quantum mechanical approach available in software (HyperChem). They used a primitive perturbation model accounting for short-range repulsion through a hard sphere reference fluid, long-range attraction via an LJ dispersion term, and dipole-dipole interactions from the Keesom equation. Inclusion of the dipolar perturbation term did not increase the number of adjustable parameters needed since the simple polar model they used needed only the dipole moment. Rather than making this an “adjustable” parameter, they determined the dipole moments of the compounds using the afore mentioned software.

One of the most recent contributions to the study of polar fluids has been that of Jog and Chapman\textsuperscript{12} in their work, they apply the Barker-Pople (or u-expansion) to spherical segments of a chain-like molecule and assume the dipole is oriented perpendicular to the molecular axis. They apply the method to pure fluids using a hard sphere reference fluid and the SAFT chain equation of state.
2.1.3 Statistical Mechanics-Based Perturbation Theory.

As noted above, there are a long series of applications of perturbation theory to atomic and molecular liquids, with expansions in either the intermolecular potential or free energy. Here, we begin with a perturbation expansion of a general property \( \mathcal{B} \). We then narrow the focus to the so-called \( u \)-expansion or “Barker-Pople expansion”.

Following Gray and Gubbins,\(^{64}\) in doing an expansion for a general property, no assumptions are made regarding either the reference system or the parameterization method. The grand canonical ensemble is applied to a uniform fluid with a potential energy of \( U_\lambda \), where \( \lambda \) is a perturbation parameter, \( U_0 \) is the potential in the reference system, \( U \) is the potential of the real system and the following relations hold:

\[
\begin{align*}
U_{\lambda=0} &= U_0 \\
U_{\lambda=1} &= U.
\end{align*}
\]

Expanding the property \( \langle \mathcal{B} \rangle \) at fixed temperature, \( T \), density, \( \rho \), and volume, \( V \), in a Taylor series about \( \lambda = 0 \), yields

\[
\langle \mathcal{B} \rangle_\lambda = \langle \mathcal{B} \rangle_0 + \frac{\partial \langle \mathcal{B} \rangle}{\partial \lambda} \bigg|_{\lambda=0} \lambda + \frac{1}{2} \left( \frac{\partial^2 \langle \mathcal{B} \rangle}{\partial \lambda^2} \right) \bigg|_{\lambda=0} \lambda^2 + \ldots
\]

(2.5)

For the general property of a real system, by the above definitions, we have

\[
\langle \mathcal{B} \rangle_\lambda = \langle \mathcal{B} \rangle_0 + \langle \mathcal{B} \rangle_1 + \langle \mathcal{B} \rangle_2 + \ldots
\]

(2.6)

where \( \langle \mathcal{B} \rangle_0 \) is the value of the property \( \langle \mathcal{B} \rangle \) of the reference system at the same temperature, density, and volume as the real system; the first order perturbation term is defined as
\[
\langle B \rangle_1 = \left( \frac{\partial \langle B \rangle}{\partial \lambda} \right)_{\lambda=0}
\]  

(2.7)

and the second order term is given by

\[
\langle B \rangle_2 = \left( \frac{\partial^2 \langle B \rangle}{\partial \lambda^2} \right)_{\lambda=0}
\]  

(2.8)

The mathematical details for obtaining the final form of the first and second order perturbation terms can be found in Gray and Gubbins.64

Significant simplifications of these terms can be made depending on the choice of reference potential and parameterization. Using the Barker-Pople reference system, where the expansion is in terms of the anisotropic potential, \( U_a \), the total intermolecular potential, \( U_\lambda \), is given by

\[
U_\lambda(r^N \omega^N) = U_0(r^N) + \lambda U_a(r^N \omega^N).
\]  

(2.9)

\( U_0 \) is the unweighted average of \( U \equiv U_\lambda \) over the orientations:

\[
U_0(r^N) = \langle U(r^N \omega^N) \rangle_{\omega^N}
\]  

(2.10)

such that

\[
\langle U_a(r^N \omega^N) \rangle_{\omega^N} = 0.
\]  

(2.11)

The unweighted average is denoted by \( \langle \cdots \rangle_{\omega^N} \), and defined as

\[
\langle \cdots \rangle_{\omega^N} = \langle \cdots \rangle_{\omega_1 \cdots \omega_n} = \frac{1}{\Omega^n} \int d\omega_1 \cdots d\omega_n (\cdots) = \frac{1}{\Omega^n} \int d\omega^n (\cdots)
\]  

(2.12)

with \( \Omega \) given by

\[
\Omega = \int d\omega = 8\pi^2 \text{ (non-linear)}
\]

\[
= 4\pi \text{ (linear)}.
\]  

(2.13)
Furthermore, pair-wise additivity of the total potential energy is assumed. As a matter of convenience, the total potential energy of a given configuration of \( N \) molecules is separated into a sum of terms, where all two-body, three-body, and higher interactions are included. When the assumption of pair-wise additivity is made, only the sum of isolated pair interactions is considered as non-zero. The degree to which this assumption affects the accuracy of the results is uncertain for two reasons. One, the non-additive terms are usually small, and negligibly so for very dilute systems. Second, the pair potential is not known exactly. Therefore, any discrepancies with experimental data can not be attributed entirely to the pair-wise additive assumption.\(^{66}\) Now, the intermolecular pair potential, \( u_{\lambda} \), can be written as

\[
u_{\lambda}(r\omega_{1}\omega_{2}) = u_{0}(r) + \lambda u_{a}(r\omega_{1}\omega_{2}).\tag{2.14}\]

The pair-potential energy is the sum of contributions arising from repulsive forces (\( + \)) and attractive forces (\( - \)). The mathematical relationship between the potential energy and the forces is given by

\[
f = -\nabla u,\tag{2.15}\]

where \( f \) is the intermolecular force and \( u \) is the intermolecular potential between two molecules. Consider the potential energy curve in figure 2.1. The potential minimum occurs at a distance \( r_0 \) from the particle center and has a value of \(-\varepsilon\), which is referred to as the maximum depth of the potential or potential well depth. The distance at which the potential energy curve crosses zero, representing a complete balance of attractive and repulsive forces, is called \( \sigma \). Using this notation, we can define three levels of forces that are the primary contributors to
the potential energy. For short-range \((r \leq \sigma)\), the repulsive forces arising from valence energies of chemically saturated molecules make a positive contribution to the potential energy. In the short and intermediate-range \((0 < r \leq r_0)\), the residual valence interactions are important. Hydrogen bonding molecules and electron donor-acceptor complexes exhibit residual valence interactions and are referred to as *associated* molecules. The third division of forces is the long-range interactions \((r \geq r_0)\) and are subdivided into three primary categories: fluctuation (or dispersion), polarization (or induction), and direct electrostatic forces.\(^{66}\)

![Diagram of potential energy}

**Figure 2.1.** Dispersion potential energy as a function of distance with 0 being at the particle center.
When two molecules are in close proximity of each other, such that their electron clouds overlay, a very strong Coulombic repulsion between the negatively charged electron clouds prevent them from collapsing onto each other.\textsuperscript{63} This results in a very strong positive contribution to the potential energy, increasing to infinity at the minimum separation distance. This minimum distance is defined differently depending on the potential model. For example, for a hard sphere potential, the minimum separation distance is the radius of the hard sphere. For softer potential models, some degree of overlap of the particles is allowed such that the minimum separation distance is less than that of the hard sphere radius.

The residual valence interactions are more difficult to model. Typically, exact or near exact models can be developed for the short and long range forces, but the intermediate range is much more difficult.\textsuperscript{66}

The fluctuation or dispersion energies, one of the long-range contributors, resulting from van der Waal forces, arise from quantum effects that are also associated with the dispersion of light by matter. Physically, these effects can be described as arising from fluctuating induction effects as the spatial location of a molecule's electrons slightly polarizes a neighboring molecule. These changes occur very rapidly and are only maintained instantaneously.\textsuperscript{63} Polarization or induction energies arise when the charge distribution of one molecule is disturbed by the multipole moments of a second molecule. The magnitude of this effect is related to the polarizability of the molecule (i.e. the molecule’s propensity for a change in the charge distribution due to the presence of a neighboring polar
molecule). Classic coulombic interactions give rise to the direct electrostatic energies. Dipole moments, quadrupole moments and the higher electric moments of a molecule interact with those of other molecules giving rise to electrostatic forces. These forces along with the polarization and induction, and their corresponding energies, can be formulated from classical electrostatic theory. The valence, residual valence, and dispersion energies must be derived from quantum theory. Quantum mechanics can be used to formulate the total interaction.  

In order to distinguish between types of molecules in the following discussion, three categories are defined based on the above description of intermolecular interactions. The total interaction energies of non-polar molecules (e.g., argon (Ar), diatomic chlorine (Cl₂), and methane (CH₄)) arise from short-range valence energies, dispersion energies, and direct electrostatic energies from multipole moments higher than the dipole moment. The negative and positive charge centers of non-polar molecules are coincident. Associated molecules (e.g., water (H₂O), ammonia (NH₃), and methanol (CH₃OH)) have all the types of energies contributing to its total interaction energy and are distinguished from the other categories on the basis of the inclusion of the residual valence energies responsible for hydrogen-bonding and formation of electron donor-acceptor complexes. In polar molecules (e.g., acetone (CH₃COCH₃), methylchlorine (CH₃Cl), and acetic acid (CH₃COOH)), the positive and negative charge centers are permanently separated, albeit a partial separation, by some finite distance and thus constitute a dipole and a dipole
moment. The total interaction energies of polar molecules include all those of the non-polar plus the polarization energies. As a distinction, all associating molecules are polar to some extent but not all polar molecules associate. In this work, the concern is primarily with mixtures of a polar component in a non-polar solvent, a polar co-polymer in a non-polar solvent, and the solution of a polar and associating solute (NH$_2$–R–COOH) in a polar/associating (H$_2$O) solvent.

In general, equations of state derived from statistical-mechanical perturbation theory are expressed as expansions in either the compressibility or Helmholtz free energy. For example,

$$A = A_{av} + A_{	ext{pert}}. \quad (2.16)$$

The perturbation term can account for any number of effects that can be considered as a “perturbation” from the reference fluid. For example, the van der Waals attraction forces are typically accounted for in a “dispersion” term that is treated as a perturbation of the reference fluid. The more accurately the reference fluid represents the actual fluid of interest the better representation of the real fluid by any of the perturbation methods. Therefore, the choice of reference fluid is most critical to the success of any equation of state arising from statistical-mechanical perturbation theory.

As stated above, any number of perturbations to the reference fluid can be included. In this approach, each perturbation term is assumed to be independent of all other terms. For example, we can write

$$A = A_{av} + A_{\text{chain}} + A_{\text{disp}} + A_{\text{assoc}} + A_{\text{polar}} + A_{\text{ind}} + A_{\text{mut}} + \ldots \quad (2.17)$$
where $hs$ is the contribution of a hard sphere reference fluid; $chain$ represents the contribution due to the formation of covalent bonds giving rise to a chain-like molecule from individual spheres, or segments; $disp$ represents dispersion forces often modeled by a Lennard-Jones tail or square well model; the tendency of molecules to associate, such as in hydrogen-bonding or formation of complexes, is given by $assoc$; the long-range Coulombic forces due to the interaction of molecules with dipoles is included in the $polar$ term; $ind$ represents any contributions from the induction of dipoles by permanent dipolar molecules; and $multi$ includes the contributions from interactions of dipoles with higher moments as well as the interactions of the higher moments among themselves. Any additional effects can equally be added as perturbations.

The type of perturbations included in a given model, or equation of state, depends on the particulars of the system of interest. For example, in the liquid phase the long and intermediate-range forces must be accounted for in order to produce a reasonable model. When dealing with liquid mixtures, information regarding the structure of the liquid mixture and the molecular forces driving the like-like and like-unlike interactions is required. Since the information needed is incomplete, varying assumptions are made in the theoretical developments aimed at modeling these systems. The assumptions made are dependent upon the specific interactions relevant to the system of interest. So, while a general theory for all liquid mixtures has not been developed, various theories that model some subset are readily available. While many of the early theories account for only the short-range repulsion and long-range attraction (i.e. van der
Waals type equations of state), several newer methods have been developed to account for some of the intermediate forces. For example, SAFT, which additionally considers the formation of a chain and molecular association as perturbations, is often applied to polymeric systems while the Scatchard-Hildebrand regular-solution theory is typically applied to mixtures of non-polar liquids. The development of equations of state for polar fluids continues to be an active area of research.

### 2.2 Chain Equations of State for Polar Fluids

In chain equations of state, the reference fluid is one that accounts for molecular repulsion (i.e. hard-sphere type interactions) and the formation of chemical bonds (i.e., covalent bonds) resulting in a chain of spheres or segments:

\[
A_{\text{ref}} = A_{\text{hs}} + A_{\text{chain}}. \tag{2.18}
\]

There have been several efforts to extend the basic chain equation of state to account for the various interactions that are often treated as perturbations to the reference fluid. For example, SAFT is a chain equation of state, applicable to pure fluids or mixtures, that also accounts for association and molecular attraction (dispersion).\textsuperscript{53-55,78} Liu, et al.,\textsuperscript{22} using the SAFT chain term, proposed a model for mixtures that also accounts for dipolar interactions with a LJ type dispersion term.\textsuperscript{84,85} Following the same approach as Liu et al., in addition to the polar term, Xu et al.\textsuperscript{86} included terms for dipole-induced dipole interactions and a multipole term accounting for dipole-quadrupole and quadrupole-quadrupole...
interactions applicable to pure chain and polar fluids. Cong _et al._,\textsuperscript{87} in their development of a model for mixtures, include the terms of Xu _et al._ as well as a contribution from what he refers to as “perturbation internal energy”. However, this term is omitted from the calculations since it is cancelled in the differentiation of the residual Helmholtz free energy to obtain activities (the goal of Cong's work). Another significant difference in Cong's approach is that the parameters are fit on a segment basis rather than a molecular basis. This allows for multi-sized segments rather than an averaged-size applied to all segments. While the results are good for the systems he studied, it has not been tested against other systems to show whether or not his “segment-based” parameters are applicable outside the limited systems he studied. Jog and Chapman\textsuperscript{12} developed a chain equation of state applicable to pure polar fluids, where the reference fluid is a chain of polar hard spheres such that

\[
A_{ref} = A_{polar_{ns}} + A_{chain}. \tag{2.19}
\]

Each of these equations use the chain term developed by Chapman _et al._ in the SAFT equation of state.\textsuperscript{53-55,78}

### 2.3 Quantum Theory

As noted above, quantum theory has been successfully used in the determination of parameters for the SAFT model and for dipole moments of amino acids. Furthermore, recent work utilizes quantum theory to determine the optimum structures of some amino acids and di-peptides.
Quantum mechanics is a set of laws describing the behavior of very small particles, such as electrons and nuclei, which do not obey the laws of classical mechanics. These laws can be used to determine or gain insight into a wide variety of interactions of these small particles, such as the properties of reaction intermediates, mechanisms of chemical reactions, and intermolecular forces. By combining quantum mechanics with statistical mechanics, thermodynamic properties (e.g., entropy, heat capacity) can be determined. Quantum mechanics can also be used for the theoretical calculation of molecular properties (e.g., bond lengths, bond angles, dipole moments, multipole moments) and stable, or equilibrium, geometry.

The properties in which we are interested include the equilibrium geometry, hydrogen bonding strength and range, and dipole and higher moments in a solvent model. The nuclear configuration that minimizes the molecular electronic energy and nuclear repulsion is the “equilibrium geometry” of a molecule’s set of electrons and nuclei. The hydrogen bonding strength is given by a change in molecular energy between two molecules existing independently and when the two molecules “dimerize” by forming a “hydrogen bond”. The range, or length, of the bond is determined from an optimized structure of the dimer. The dipole and multipole moments can be thought of as the response of the wave function to the presence of an external electric field in the limit of the field strength approaching zero.

The theoretical levels of interest here include the standard Hartree-Fock (H-F), perturbation methods such as MP2, density functional theory (DFT)
methods, and hybrid approaches, for example B3LYP. Since the Fock operator is dependent on all occupied molecular orbitals (MOs), the Hartree-Fock equations are a set of pseudo-eigenvalue equations. This necessitates an iterative procedure. Self-Consistent Field (SCF) orbitals are a set of functions that are a solution to the Hartree-Fock equations. Electron correlation is not accounted for in the H-F SCF method. The second order correction of Møller-Plesset, MP2, is a perturbation method which accounts for electron correlation by a perturbation of the H-F wave function.\textsuperscript{88,89}

Density functional theory is based on the idea of the existence of a one-to-one correspondence between the electron density of a system and the energy (as proved by Hohenberg and Kohn, 1964). Since the electron density is given by the integration of the wave function squared over N-1 electron coordinates, it only depends on three coordinates, independent of the number of electrons. Thus, increasing the system size, and hence the complexity of the wave function, does not change the number of electron density variables. B3LYP, a hybrid method, is one of the proposed density functionals that includes exact exchange. However, current functionals poorly represent the weak dispersion interactions,\textsuperscript{89} in which we are interested.
CHAPTER 3: EFFECTS OF TEMPERATURE AND SALT CONCENTRATION ON THE SOLUBILITY OF AMINO ACIDS AND DIPEPTIDES

3.1 Introduction

Developing new separation processes and improving existing purification processes for proteins and smaller molecules, such as small chain peptides, is a constant goal of the pharmaceutical industry. In order to accomplish this, numerous lab scale studies must be done to determine optimum operating conditions, including the effects of temperature and various salt types and concentration on the biochemical of interest. Furthermore, the study of small peptide systems can provide valuable information for the biomedical and pharmaceutical industries in their efforts to develop highly target-specific products. For example, RGD peptides are used for blocking platelet aggregation. Some as small as eight member cyclic peptides have been identified (CRGDxxxC, CxxxRGDC). Other studies are investigating the use of small peptides to target and/or modify specific metabolic activities. Analysis of the Conus venom revealed myriad diverse pharmacologically active small peptides, with the most active in the range of ten to thirty residues. The authors assert that these substances could prove “to have a pharmaceutical potential comparable to plant alkaloids” and suggest that these venom be screened when searching for a ligand targeted for a particular receptor. In order for these small peptides to be available commercially, effective separation processes must be
developed or the peptides chemically synthesized, which is a very costly procedure.\textsuperscript{27}

Furthermore, a literature survey for solubility data for the amino acids, and in particular, aqueous electrolyte solutions, revealed some inconsistencies in the reported data. For example, recent work by Khoshkbarchi and Vera\textsuperscript{48} show that the effect of potassium chloride (KCl) on the solubility of glycine is greater than that of sodium chloride (NaCl), while Cohn and Edsall,\textsuperscript{91} reporting the results of early studies by Pfeiffer and Würgler,\textsuperscript{92} show the opposite effect of the salts. Moreover, it is rare to find cases in the literature for aqueous two-phase systems (ATPS) in which a series of peptides or amino acids are studied under a consistent set of conditions, e.g. polymer type, salt type, etc. A systematic study of these systems is critical to the development of accurate and more general models.\textsuperscript{23,35}

A solubility study of amino acids and dipeptides can lead to significant insight into the solubility patterns of these small peptides as well as more complex molecules. Additionally, by studying pairs of reverse dipeptides, the effects of amino acid sequence and end groups on the thermodynamics of separation processes can be detected. Varying the amino acids within a small peptide allows the explicit study of the effect of functional groups on solubility. Also, a complete experimental solubility study of the amino acids and small peptides would provide an invaluable database to industry and academia. Based on the experimental results, a theoretical model could be developed to assist with
the design and improvement of separation processes for larger, more complex systems.

We are applying a two-fold approach: first, an experimental study of the solubility of amino acids and dipeptides as a function of solution temperature, salt type, salt concentration, and sequence; and secondly, the interpretation of the experimental results in terms of theoretical studies, in particular *ab initio* quantum mechanics. The future goal is to develop a statistical mechanics based solution model.

As the interest here is in understanding the contributions of each residue to the behavior of peptides in solution, three amino acids where the side chains can be classified as neutral (glycine), hydrophobic (alanine), and hydrophilic (aspartic acid) were selected. The side chain of glycine is a single hydrogen; alanine has a hydrophobic methyl group; and the carboxylic acid side chain of aspartic acid is hydrophilic. The configurations are shown in figure 3.1. By considering these amino acids both as individual entities and as residues within a small peptide, we can begin to understand how the behavior of these molecules differs in relation to their existence as single molecules versus as a residue in a peptide. In this study, the effects of temperature and salt concentration on the solubility of the amino acids glycine (*gly*), alanine (*ala*), and aspartic acid (*asp*), which is a precursor to the sweetener aspartame, are considered. Also investigated are the reverse dipeptides *asp-gly*, *β-asp-gly*, *gly-asp*, *gly-ala*, and *ala-gly*. Two experimental methods are used in the analysis of the solubility studies: 1) A dry weight method and 2) reverse phase HPLC using an internal
Figure 3.1. Structures of the amino acids a) glycine,\textsuperscript{93} b) alanine,\textsuperscript{94} and c) (\(\alpha\) acid is charged) and d) (\(\beta\) acid is charged) aspartic acid.\textsuperscript{95} These structures are built with Molecule 3D\textsuperscript{96} using the bond lengths, bond angles, and torsional angles reported for optimized structures in the corresponding references.
standard. Additionally, we report values for the change in enthalpy in going from the solid state to infinite dilution in the liquid state for the dipeptides of asp and gly. Approximate values for the melting temperatures of these compounds are also reported.

In the next section, we describe the experimental methods employed in this study. The results and discussion section is divided as follows: the temperature studies of the amino acids and dipeptides respectively; the salt studies of the amino acids and dipeptides, respectively. Following the results and discussion, a theoretical analysis of the temperature study in terms of Henry's Law is given and the main conclusions of the work are stated.

3.2 Experimental Methods

All chemicals (sodium chloride, potassium chloride, lithium chloride, glycine, alanine, aspartic acid, and the dipeptides: gly-asp, asp-gly, gly-ala, ala-gly) were obtained from Sigma-Aldrich with a purity of >99.99% for the salts, and >99.0% purity for the amino acids and dipeptides with the exception of >98% for gly-ala. The salts and amino acids were stored at room temperature, and the dipeptides were kept in a freezer at -20°C. Prior to making the solutions, the salts and amino acids were dried for three days (72 hrs) in a dessicator (VWR1500E) to remove any moisture. Dipeptides were not dried due to degradation at high temperature. Salt solutions (0.0N, 0.5N, 1.0N and 1.5N) were prepared with milli-Q water and filtered using a nylon 0.22 μm syringe filter. The amino acids or peptides were dissolved in the solutions and over saturated by at
least ten percent. Initial solutions of 25 ml for the amino acids and 3 ml for the
dipeptides were prepared. The solutions were placed in an orbital shaker (Lab-
line Orbit Environ-Shaker) for 3 hrs at 30°C and 250 rpm followed by 48 hrs at
25°C and 250 rpm. The temperature was maintained constant by the internal
thermostat of the shaker and verified with a standard thermometer. The
solutions were then allowed to stand for 7 hrs at 25°C in an oven or water bath
for later experiments at lower or higher temperature.

Analysis. Dry-weight method. The solution was filtered (nylon 0.22 μm
syringe filter) into a previously weighed aluminum dish. The dish containing the
solution was then weighed and dried in the oven (VWR) at 55°C for 24 hrs and at
70°C for 72 hrs. The dried sample was then weighed. From this data and the
known normality of the salt solutions, the amount of amino acid in the solution
was determined. Khoshkbarci and Vera,48 using a very similar dry weight
method, have shown that no appreciable amount of the salts is adsorbed by the
amino acid in the solid phase. Similar measurements were performed in this work
for glycine and salt solutions showing also that no appreciable amount of the
salts were absorbed. Thus, the amount of amino acid can be accurately
determined from this method.

HPLC Method. For analysis of the dipeptide systems, an accurate method
for smaller sample size and lower concentrations was developed. After filtration
and equilibration, a 200 μl sample of the solution was pipetted into new vials and
diluted with 400 μl of a 0.1-0.4% tryptophan (trp) solution. The vials were then
capped and stored under refrigeration at 4°C until the HPLC was available. The
storage time was varied up to two weeks. Controls (standards) did not show any sign of denaturing after prolonged storage under refrigeration. The HPLC (from Millipore Corporation/Waters Chromatography Division, Marlborough, MA with an 18C-5μm Luna column (Phenomenex, Torrance, CA) and PDA detector scanning for wavelengths of 200 to 254 nm) was used to determine the concentration of the samples. Trp was included in the sample as a marker of known wavelength and concentration. For the systems studied here, there was no notable interaction of the trp with the amino acids and dipeptides, as determined by variation of the trp concentration in gly standards and vice versa. These measurements were performed for each of the amino acid and dipeptide solutions. The column solvent was a 68:32 mixture of 10 mM sodium acetate to acetonitrile with a flow rate of 1ml/min. Both the acetonitrile and the sodium acetate solutions were degassed for 30 minutes, the injector purged and the column conditioned with 1ml/min acetonitrile for 1 hr prior to use. The column was then allowed to equilibrate for 1 hour at the solvent flowrate (68:32 NaAc:Acn) of 1 ml/min before processing the samples. After the completion of the processing of the samples, the column was reconditioned for an hour with the acetonitrile at a rate of 1ml/min. The amount of dipeptide or amino acid in solution was determined from the area under the curve at the specified wavelengths (204 nm for asp and β-asp-gly; 212 nm for gly, asp-gly, ala, and gly-ala; 220 nm for gly-asp, ala-gly, gly-ala; and 230 nm for trp). All experiments were replicated at least three times. Based on the verification of irp as an internal standard, the uncertainty is 2-3%.
3.3 Results and Discussion

As is stated in the introduction, we are interested in understanding the contributions of each residue to the behavior of peptides in solution. Recall that the three amino acids of this study represent neutral (glycine), hydrophobic (alanine), and hydrophilic (aspartic acid) side chains. The configurations are shown in figure 3.1 and are in the zwitterionic form, the form in which amino acids exist in aqueous solution. We begin with the results of our temperature study presented in figure 3.2, considering first the amino acids followed by the peptide systems.

3.3.1 Temperature Study: Amino Acids.

Figure 3.2a shows the effects of temperature on the solubility of glycine (gly), alanine (ala), and aspartic acid (asp). The results published by Fasman are included for comparison. The comparison serves as a verification of our experimental method as well as providing us the confidence in the accuracy of the published data. The only notable difference in our results and the previously published data is for asp, which has a very low solubility and is thus difficult to measure. Whether our data or the earlier data are the more reliable is unclear at this point. The experimental values from this work are listed in table 3.1. Figure 3.2a indicates that increasing the solution temperature significantly enhances the absolute magnitude of gly solubility, with a relative magnitude increase of approximately a factor of two. As is clearly observable on the log scale, the
Figure 3.2. a) Solubility (log(g solute/kg solvent)) of gly (■), ala (●), and asp (◇) as a function of temperature (°C) compared with data reported by Fasman (—).97 b) Solubility (g solute/kg solvent) of the dipeptides gly-asp (◆), asp-gly (●), β-asp-gly (▲) and asp (●) as a function of temperature (°C).
relative solubility of \textit{asp} increases by a factor of three over this same temperature range.

Table 3.1. Temperature Dependence of the Aqueous Solubility of Amino Acids and Dipeptides

<table>
<thead>
<tr>
<th>Temp (\degree C)</th>
<th>Solubility g/kg water</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Asp</td>
</tr>
<tr>
<td>5</td>
<td>2.11</td>
</tr>
<tr>
<td>10</td>
<td>2.26</td>
</tr>
<tr>
<td>15</td>
<td>3.35</td>
</tr>
<tr>
<td>20</td>
<td>4.00</td>
</tr>
<tr>
<td>25</td>
<td>4.72</td>
</tr>
<tr>
<td>30</td>
<td>6.38</td>
</tr>
</tbody>
</table>

As the temperature increases, the density of water decreases, as well as the degree of hydrogen bonding, while the coordination number increases.\textsuperscript{98} These changes within the hydrogen-bonding network of the water increase the probability for the amino acid molecules to be solvated. In addition, any intramolecular hydrogen bonding among the amino acids themselves is also weakened as the temperature rises, thus allowing an increase in the solubility.

By taking a cross-section from the amino acid temperature studies at 25\degree C in figure 3.2a, we see that the order of solubility for these amino acids is \textit{asp}<\textit{ala}<\textit{gly}. This trend is inconsistent with “hydropathy” charts for the amino acids,\textsuperscript{27} which indicates that \textit{asp} is the most hydrophilic and \textit{ala} the most hydrophobic of these three amino acids, as we would expect. However, the hydropathy values are indicative of how the amino acids affect the native, or minimum energy, conformation of a large peptide or protein. They are used as a guide to which
sections of a large chain will be interior (hydrophobic, due to the inability to
hydrogen bond with the water solvent) and which will be exterior (hydrophilic, due
to their ability to hydrogen bond with water) in a folded structure. It should be
noted that, when a chain forms, the connectivity affects the conformation of the
amino acids. The neighboring residues may cause steric hindrance and/or an
increase (or decrease) in attractive or repulsive forces thus resulting in a
minimum energy conformation of the residue which differs from its conformation
as a single molecule.

Consider again the structures of these molecules (figure 3.1) and recall
that glycine has a neutral side group and is shown to be the most soluble of the
three. In figure 3.2a, we see that alanine is slightly less soluble than glycine.
This phenomenon is due to the added hydrophobic group which may act in two
respects: a) the change in free energy since the methyl group is not "liked" by
water; and b) steric hindrance whereby the group interferes with hydrogen bonds
forming between the ionic end groups and water. Although as we stated earlier,
the side chain of asp is hydrophilic, it has a solubility of an order of magnitude
less than that of glycine and alanine. For this phenomenon, we propose that the
behavior is due to intramolecular hydrogen bonding of the side group with the
terminal amine, which would prevent solvation by water. Since asp is a di-acid
and has a pI of 2.98,\textsuperscript{99} at neutral pH, it is predominantly in a negatively charged
state. The three possible states are shown in figure 3.3.\textsuperscript{91} In the following
discussion, we will show that our hypothesis is supported by molecular structures
determined using Raman and FTIR spectroscopy and quantum mechanics.
Figure 3.3. Structures of aspartic acid at pH 7.0:\textsuperscript{91} (the $\alpha$ carbon is in \textit{italics} and the $\beta$ carbon in bold. a) Predominant Form has a net charge of -1, b) the $\alpha$ acid is in a charged state, and c) the $\beta$ acid is in a charged state.

There has been significant study of a few amino acids using quantum mechanical approaches,\textsuperscript{93-95,100-115} with glycine being the most frequent object of the study.\textsuperscript{93} However, the most numerous studies have been in the gaseous phase in which the zwitterionic form does not exist as a minimum energy conformation.\textsuperscript{104} Also, in the gaseous phase the amino acids exhibit some level of intramolecular hydrogen bonding, including glycine and $L$-alanine which are known not to do so in aqueous solution.\textsuperscript{93} Császár\textsuperscript{110,111} performed thorough potential energy surface (PES) searches for the conformations of gaseous glycine and $\alpha$-alanine and found eight minimum energy conformations for glycine and thirteen corresponding minimum energy conformations for alanine. His work indicates that the addition of the methyl group of alanine has little effect on the preferred geometries. Moreover, in an experimental and \textit{ab initio} study of gaseous alanine and small peptides, Cassady \textit{et al.}\textsuperscript{115} demonstrated that the position of the methyl group near the N-terminus served to stabilize the protonated amine, but had little effect on the geometry.
Including the solvent effects is challenging and has been performed for a few systems.\textsuperscript{94,95,105,107,109,112-114} Nagaoka \textit{et al.}\textsuperscript{114} and Tuñón \textit{et al.}\textsuperscript{113} both considered the problem of intramolecular proton transfer of glycine in water. Kikuchi \textit{et al.}\textsuperscript{107} investigated the solvent effects of water on the conformation of zwitterionic forms of glycine, alanine and serine, using a continuum model consisting of the generalized Born formula for the solvation energy and the Lowdin population analysis for evaluating the atomic fractional charges. They find two conformers of glycine to be equally probable, differing only in a 60° rotation around the carbon-nitrogen bond. Ding and Krogh-Jespersen\textsuperscript{105} studied a 1:1 glycine zwitterion-water complex and found only one true minimum on the PES. In this complex, the water molecule, bridged between the NH$_3^+$ and COO$^-$ groups of glycine, serves as both a hydrogen bond acceptor and donor. Chakraborty and Manogaran\textsuperscript{93} are able to reproduce experimental results for the vibrational frequencies of glycine by using the Onsager reaction field model to account for the aqueous environment, and their optimized structure is reproduced here in figure 3.1a.

Sambrano \textit{et al.}\textsuperscript{109} studied the intramolecular proton transfer of α-alanine in an aqueous medium using \textit{ab inito} methods at various levels from HF/6-31G* to MP2/6-31++G** and included the solvent effects using self-consistent reaction field theory (SCRF). They selected conformers from the work of Császár\textsuperscript{111} which would be in a conformational form conducive to intramolecular proton transfer. Their results indicate that the formation of the zwitterionic form from an intramolecular proton transfer is much more energetically favorable than from the
configurational changes necessary for the proton transfer to take place with water molecules. Frimand et al.\textsuperscript{94} performed a variety of studies on aqueous solutions of L-alanine. They determined, in order to gain agreement with experimental vibrational absorption (VA) and vibrational circular dichroism (VCD) intensities, that a combination of solvating the solute with explicit water molecules and a dielectric medium using the Onsager model was required. They found that including the solvent effects differentiated between the backbone conformation of possible minimum energy conformations of the solute. The final conformation of the zwitterion in water of Frimand et al.\textsuperscript{94} is shown in figure 3.1b.

Studies of \textit{asp} in solution are more limited for several reasons. Since \textit{asp} is a larger molecule, it is more difficult to model and more computationally expensive. Furthermore, because \textit{asp} has two acid groups, two zwitterionic forms, or tautomers, exist at its pI (isoelectric point). Nagy and Noszály\textsuperscript{95} have performed a comprehensive study of the possible conformations of each of the tautomers using various levels of theory from HF/6-31G* to QCISD(T)/6-31G# and MP2/6-311++G**. The particular conformations they considered were each dihydrated species (i.e., two water molecules were hydrogen bonded to a single \textit{asp} molecule). They found two forms to be energetically favorable, with the \textit{\alpha}-gauche 1 being slightly more favorable than the \textit{\beta}-gauche 1. The \textit{\beta}-gauche form contains a very strong hydrogen bond (1.90Å in solution) between the \textit{\beta}-COO\textsuperscript{-} group and the NH\textsubscript{3}\textsuperscript{+}. This configuration repels water molecules from this space of
the molecule and reduces the coordination number to 7.5 as compared to 8.8 for the α-gauche form. These two conformations are shown in figure 3.1c,d.

A Raman and FTIR difference spectroscopy study of the water structural changes in aqueous solutions of glycine and alanine was performed by Fischer and Eysell\textsuperscript{16} at their pI (6.1). By comparing the results of 2 M and 1M solutions of glycine, they show that the degree of hydrogen bonding for glycine with the aqueous medium is proportional to the solute concentration. They found intense broad bands in the IR spectra (around 3100 cm\textsuperscript{-1}) indicative of hydrogen bonding in their solutions of amino acids. Also, they found evidence of electrostatic and hydrophobic interactions as well as hydrogen bonding. Furthermore, they conclude that point charges modify the symmetry of the water molecules while no symmetry changes were observed for what they refer to as hydrophobic interactions. Alanine showed very similar qualitative trends to those of glycine in terms of these effects. The absorbance indicative of hydrogen bonding was very slightly increased (~0.002 a.u.) for alanine at the same concentration of solute. One important note is that in their study, which included carboxylic acids, they found no alteration of the symmetry of the water molecules in proximity to the carboxylic acid as compared to the bulk water. They propose that the density of the negative charge of the functional group is equivalent to one of the oxygen atoms of water thus not interrupting the water symmetry. Additionally, they show that hydrophobicity becomes important with increasing solute concentration and with increasing the aliphatic chain length, as would be expected since the aliphatic chains are known to be hydrophobic in nature.
Fischer and Eysel's FTIR and Raman difference spectra studies of aqueous solutions of glycine and alanine indicate that hydrogen bonding with water occurs for both glycine and alanine. The results of Cassady et al. indicate that the methyl side group of alanine serves to stabilize the protonated form. The studies of Nagy and Noszá indicate that in aqueous solution aspartic acid forms strong hydrogen bonds with itself (intramolecular), with one of the two dominant forms having an intramolecular bond strength energetically favorable to that which would be formed with water. Therefore, this bond does not break when put in solution. At pH of 7.0, the di-acid asp exists in a charged state thus increasing the probability of intramolecular hydrogen bonding. This configuration excludes water from that region of the molecule thus reducing the potential bonding sites with the aqueous medium. Hence, these studies support our hypothesis regarding the relationship between amino acid solubility and intramolecular hydrogen bonding as well as intermolecular hydrogen bonding with the aqueous medium. Or, equivalently, our experimental results support the findings of the quantum studies.

3.3.2 Temperature Study: Peptides.

In figure 3.2b, the effect of temperature on dipeptides of gly and asp is shown. The mono-peptide, asp, is included simply as a point of reference; gly is not included due to the large difference in scale needed. The solubility for the dipeptides is significantly greater than asp throughout the temperature range studied with the largest increase occurring at the higher temperatures. However,
when the $\beta$ form of asp is used in the dipeptide, the solubility as compared to the amino acid asp is only slightly affected by temperature increases. Also, the effect of reversing the dipeptide sequence on solubility is demonstrated from this study. When gly is the leading residue (N-terminus), the solubility is much greater than when asp is.

Therefore, when asp is in a chain, the presence of its neighboring residues affect the degree to which asp behaves as more hydrophobic or hydrophilic or, more correctly, the degree to which it hydrogen bonds with the water. This is evident in the change in solubility of asp as compared to the asp-gly and gly-asp dipeptides. This change in solubility can be explained in part by observing the structural changes. Note that in figure 3.2b, the dipeptide $\beta$-asp-gly is only slightly more soluble than the “mono-peptide” asp. Comparison of the structural differences of asp-gly and $\beta$-asp-gly (see figure 3.4), reveals that the acid side chain moves from the $\alpha$ carbon to a $\beta$ carbon. This change places the $\beta$-carboxyl group significantly closer to the terminal amine, thus increasing the probability of intramolecular hydrogen bonding between the amine and carboxyl group. When a molecule hydrogen bonds with itself, those bonding sites are no longer available to bond with water or other solute molecules. Going from $\beta$-asp-gly to asp-gly again we see an increase in solubility throughout the temperature range. The shift of the side chain from a $\beta$ to an $\alpha$ position places the terminal amine and the $\beta$-carboxyl group farther apart than with the $\beta$ form of asp-gly. Again, while intramolecular hydrogen bonding is still possible, its probability is less due to steric hindrance of the methyl group, thus increasing the probability
Figure 3.4. Structures of the dipeptides a) gly-asp, b) asp-gly, and c) β-asp-gly. Built and optimized using Molecule 3D.96
for hydrogen bonding of the peptide with water. By reversing the sequence in the dipeptide, the solubility increased more dramatically. Again, comparing the structural differences in *asp-gly* and *gly-asp* reveals that the probability of intramolecular hydrogen bonding of the terminal amine with the β-carboxyl (side) group of *gly-asp* is very low since the β-carboxyl group is now in proximity to the terminal carboxyl group.

While only one temperature point (25°C) for the dipeptides formed from alanine and glycine is reported (see table 3.2), there are some observations that need to be stated. The solubility of *gly-ala* is greater than that of its reversed sequence by a factor of 1.94 indicating the significant dependence of solubility on the sequence. We propose that the much lower solubility of *ala-gly* results from the steric hindrance of the methyl group, which hinders solvation of the N-terminus.

The solubility of the dipeptides in terms of glycine as the C-terminal residue can also be compared. In relation to the solubility of the mono-peptide (*ala* or *asp*), the dipeptide solubility increases for both *ala-gly* and *asp-gly*. The degree to which this occurs, however, is significantly different for these residues. By forming a dipeptide through the addition of *gly* as a C-terminal residue to *asp*, the solubility increases from ~5 g/kg H₂O for *asp* to ~24 g/kg H₂O for the *asp-gly* dipeptide, an increase of a factor of 4.8. By forming the analogous dipeptide with *ala*, the solubility increases from ~164 g/kg H₂O for *ala* to ~193 g/kg H₂O for the *ala-gly* dipeptide, an increase of 1.18 times. First, we see that the quantitative increase in solubility for the *ala-gly* peptide is slightly greater than for the *asp-gly*
peptide, as compared to the solubility of their respective mono-peptides. However, the magnitude of the solubility of the asp dipeptide has increased 480% compared to 118% for the ala dipeptide. Thus, we see that forming a dipeptide by adding the most neutral of the amino acids as the C-terminal residue, the solubility of the N-terminal residue (or mono-peptide) is increased substantially and more dramatically for the residue with a polar side group versus the residue with a non-polar side group.

In terms of using quantum mechanics to assist in the understanding of the physically observed solubility trends for the dipeptides, there are relatively few studies available. Dipeptides of alanine and glycine have been considered by Loeffler et al. and Cassady et al. One interesting result of the gaseous phase study by Cassady et al. is that although the methyl side group of alanine had "almost no effect on the structures", it does play a role in stability of the molecules. They found that in dipeptides and tripeptides the location of this methyl side group near the N-terminus had the greatest impact on basicity indicating the ability of the substitute substituent near the protonation site to stabilize the ionic form. In other words, a methyl side group near the N-terminus tended to stabilize the ionic (or protonated) form of the N-terminus. Loeffler et al. showed that there is a significant dependence of the free energy of hydration on sequence for the ala-gly and gly-ala reverse dipeptides. They found that in aqueous solution and in zwitterionic form gly-ala is 2.7 kcal/mol more stable than ala-gly. The only quantum study available of dipeptides of aspartic acid and glycine is that of Alemán, who performed a conformational study of a
dipeptide of α-L-aspartate and glycine in the gaseous phase. And, as he states, the results are very different than what would be expected for the corresponding aqueous system and are therefore not considered here.

The results of these quantum studies support our findings that gly-ala is more soluble than ala-gly in that the gly-ala sequence is appreciably energetically favorable as compared to ala-gly in aqueous solution. Furthermore, the effect of the side chains being in proximity to the N-terminus having a greater effect on the basicity of the molecules supports the effects of sequence that we have noted in that the dipeptides are more soluble when glycine is the N-terminus residue. The protonated amine group is less stable without the side groups and therefore has a higher propensity to hydrogen bond with the water molecules. With glycine at the N-terminus, steric hindrance preventing solvation of the N-terminus is no longer a consideration. And, for these small peptide chains, the probability of the terminal amine intramolecularly hydrogen bonding with a side group is greatly diminished with gly at the N-terminus. Moreover, as we noted in the discussion of the amino acids, asp has a propensity to form intramolecular hydrogen bonds and this does not appear to be strongly affected when asp is the N-terminal residue of a dipeptide. In asp-gly, the asp can readily intramolecularly hydrogen bond. In the gly-asp sequence, however, the side group of asp acts more like a carboxylic acid. And, as indicated by the study of Fischer and Eysel, the charge on the carboxylic acid functional group is similar to that of a water oxygen and, consequently more favorable to the hydrogen-bonding network of water.
It would be interesting to note the change when a tri-peptide, such as gly-asp-gly, is formed and compared with the dipeptides studied here. Since gly is the most “neutral” of the amino acids, the changes in solubility as compared to the gly-asp and asp-gly peptides are expected to be primarily a result of transforming the terminal group of asp into an internal residue with a carboxyl side chain. Thus, asp would be in the same form as in a long peptide or protein and its solubility trends within a chain could be studied without the inclusion of end effects (position at the N or C terminus) and independent of any interactions with side chains on neighboring amino acids. In addition, it would be of great benefit to compare these results with those of other tripeptides such as gly-gly-gly and gly-ala-gly, as well as longer chains (e.g. gly-gly-asp-gly-gly).

3.3.3 Salt Type and Concentration Study: Amino Acids.

As mentioned in the introduction, published results are inconsistent regarding the effects of NaCl and KCl on the solubility of glycine. Therefore, this is where we began our investigation. Using the method indicated by recent work, solubility experiments for gly in solutions of various normality of NaCl and KCl were performed. Our results are compared with those of the previous studies in figure 3.5a. In order to clarify the trend of cation size, we extended the study to include lithium chloride (LiCl). By maintaining the anion constant, the effect of the cation size on solubility is shown for K⁺, Na⁺, and Li⁺. As clearly depicted in figure 3.5b, in the presence of these salts at constant temperature (25°C), the solubility of gly increases with increasing salt concentration and with
Figure 3.5. a) Solubility of glycine in NaCl (○) and KCl (■): Results of Khoshkbarci and Vera at 25°C,48 Pfeiffer and Wügler at 21°C,92 and this work at 25°C. b) Solubility (g solute/kg solvent) of glycine at 25°C in aqueous solutions of NaCl (○), KCl (■), and LiCl (△).
decreasing size of the cation, with the anion (Cl') held constant. This trend is in agreement with studies of the structural properties of aqueous electrolyte solutions, which show that the ions cause a significant perturbation in the hydrogen-bonded network of water, and the degree of this disturbance increases with the concentration of the ionic species.\textsuperscript{121} Our experimental values are given in table 3.2.

**Table 3.2. Salt Type and Concentration Dependence of Amino Acid And Dipeptide Solubility at Constant Temperature (25°C)**

<table>
<thead>
<tr>
<th>Salt Type</th>
<th>Conc (N)</th>
<th>Solubility g/kg solvent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Gly</td>
</tr>
<tr>
<td>None</td>
<td>0.0</td>
<td>240.01</td>
</tr>
<tr>
<td>LiCl</td>
<td>0.5</td>
<td>246.49</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>258.40</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>270.71</td>
</tr>
<tr>
<td>NaCl</td>
<td>0.5</td>
<td>226.30</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>223.95</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>234.13</td>
</tr>
<tr>
<td>KCl</td>
<td>0.5</td>
<td>198.95</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>242.57</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>245.23</td>
</tr>
</tbody>
</table>

While there is a significant difference in solubility when changing the salt from KCl to NaCl, the most significant change occurs when Li\textsuperscript+ is the cation. We know from molecular simulation that in electrolyte solutions the water molecules are strongly oriented by the anions in the first hydration shells.\textsuperscript{121} Furthermore, we know that the hydration number of the cation increases with increasing cation size while the hydration number of Cl' decreases as the size of the counterion
increases.\textsuperscript{121} Thus, the lithium ion has the lowest hydration number (~4-5) of the cations in this study and the chloride ion has its highest hydration number (~8) with Li\textsuperscript+ as the counterion. The increase in solubility of glycine in the LiCl solution could be due to the asymmetrical changes in the structure of the water hydrogen-bonding network by the ions, which would make more of the hydrogen bonding sites of the water available to the glycine.

The Hofmeister series, which lists ions in the order of their effectiveness in stabilizing proteins, parallels the capability of the ions to salt out proteins and is largely independent of the specific protein.\textsuperscript{37} This series indicates that the effect of the three cations considered in this study are K\textsuperscript+, Na\textsuperscript+ > Li\textsuperscript+, which means that the K\textsuperscript+ and Na\textsuperscript+ are more effective in stabilizing proteins and hence less likely to salt in (or more likely to salt out) a protein as compared to Li\textsuperscript+. Li\textsuperscript+ is one of the chaotropic ions (ions that tend to denature, or salt in, proteins) in the Hofmeister series. Chaotropic ions increase the solubility of nonpolar substances in water through their ability to disrupt hydrophobic effects.\textsuperscript{37} As a point of clarification, we use the term "hydrophobic effects" to indicate the inability of certain molecular groups to form hydrogen bonds with water. The term "hydrophobic interactions" is somewhat of a misnomer in that the effect it is describing is a minimization of free energy rather than a true "molecular interaction" and can be more accurately described as a lack of hydrogen bonding.

With single amino acids and small peptides, the role of the functional groups in determining solubility, and other solution properties, is magnified. Whereas in a large protein the various functional groups behave as part of a
larger entity, the effect of each individual functional group cannot readily be
determined. Since the native conformation of a protein is typically one in which
the surface groups are largely neutral or hydrophilic with sparse occurrences of
hydrophobic groups, the ions are exposed to approximately the same electrolytic
forces regardless of the specific protein studied.27 Therefore, we expect that,
depending upon the specific side groups, the interactions of single amino acids,
or peptides consisting of a few residues, with electrolytic solutions would be quite
different from those of large proteins.

Our study indicates that potassium and sodium chloride have a relatively
small affect on Ala solubility at 25°C, as shown in figure 3.6. The effect of KCl is
essentially negligible at the concentrations we considered. The results of this
study are compared with those of Khoshkbarchi and Vera48 who considered the
DL form of alanine whereas we looked at the L form. It is possible that the
differences in magnitude could be purely attributed to the fact that these two
molecules are stereoisomers. However, we do expect the trends of the salt type,
in terms of cation size, on solubility to be similar. Based on our extended study
with glycine as well as our results for alanine, we believe the trends of our results
are the more physically reasonable.

3.3.4 Salt Type and Concentration Study: Reverse Dipeptides.

The results for the dipeptides are shown in figures 3.7-8. NaCl shows little
effect on the solubility of asp-gly at the salt concentrations considered. On the
other hand, gly-asp salted in with the solubility increasing almost linearly with
Figure 3.6. Solubility (g solute/kg solvent) of alanine at 25°C in aqueous solutions of NaCl (●), KCl (▲). Comparison of D,L-Alanine from Khoshkbarzhi and Vera* (open symbols) and L-Alanine of this work (filled symbols).

NaCl concentration. Gly-ala and ala-gly have significantly different solubilities regardless of the presence/absence of salt, salt type, or salt concentration. Ala-gly salts in slightly with the three salts tested. Gly-ala salts out slightly with both NaCl and KCl and salts in with LiCl.

With the reverse dipeptides, the solubility trends vary with the type and concentration of salt and with the sequence of the residues. There are several comparisons to be made with this data: a) effect of reversing sequence gly-asp to asp-gly, b) effect of reversing sequence gly-ala to ala-gly, c) changes in the side
Figure 3.7. Solubility (g solute/kg solvent) of the dipeptides gly-asp (●) and asp-gly (●) at 25°C in aqueous solutions of NaCl.

group of the C-terminus residue from polar to non-polar (gly-asp to gly-ala), and
d) changes in the side group of the N-terminus residue from polar to non-polar
(asp-gly to ala-gly). Note that references made to specific amino acids as being
polar or non-polar correspond to the classification given in appendix A.

a) By reversing the sequence of the gly-asp peptide to asp-gly (figure 3.7),
the solubility of the molecule decreases significantly as also noted in the
temperature study. By adding NaCl to the system, the solubility of the gly-asp
peptide increases significantly with increases in salt concentration throughout the
range studied. In contrast, there is little change in solubility for the *asp-gly* peptide at any of the salt concentrations studied. Although the addition of salt increases the disturbance of the hydrogen-bonding network for the water, its effect on the structure of *asp-gly* is minimal, possibly due to strong intramolecular hydrogen bonding.

b) As in the case of the glycine and aspartic acid dipeptides, reversing the sequence of *gly-ala* to *ala-gly* leads to a large decrease in solubility for any of the electrolytes studied (figure 3.8a,b). As noted in the discussion of the temperature effects, when *gly* is the N-terminal residue there are no steric or associative interactions to prevent the N-terminus from being available to bond. And, when *ala* is the C-terminal group there is a steric hindrance for interactions of the carboxyl group with water due to the presence of the methyl side group. Moreover, *ala-gly* salts in with each of the salts studied due to the ability of the salts to disturb the hydrogen bonding network of water. In contrast, *gly-ala* salts out with KCl and NaCl but salts in with LiCl. We attribute the strong increase in solubility of *gly-ala* in the presence of LiCl to the chaotropic nature of the lithium ion. The salting out of *gly-ala* in the KCl and NaCl solutions is most likely a result of a stabilizing effect of these salts on the dipeptide, an effect which does not occur with *ala-gly*.

c) When *gly* is the N-terminus residue, the solubility of the peptides, as compared to their respective reverse dipeptide, is much greater at any concentration or type of salt, which we also saw in the temperature studies. However, in the NaCl solutions, *gly-asp* salts in (figure 3.7) while *gly-ala* salts out
(figure 3.8b). This implies that the salting in/out behavioral differences are due to the molecular differences of the solutes and their interactions with the electrolytes rather than, or in addition to, the effect of the electrolyte on the structure of the water. In figures 3.4 and 3.9, optimized structures of these dipeptides are shown.

d) The effect of NaCl on the solubility of both xxx-gly dipeptides is minimal, with the more significant change in solubility occurring with the ala-gly form.

Figure 3.9. Structures of the dipeptides a) gly-ala and b) ala-gly. Built and optimized using Molecule 3D.96
Figure 3.8. Solubility (g solute/kg solvent) of the dipeptides a) gly-ala and b) ala-gly at 25°C in aqueous solutions of NaCl (●), KCl (▲), and LiCl (△).
Again, for the *asp-gly* dipeptide, the potential for the formation of strong intramolecular hydrogen bonds exists. For the *ala-gly* form, we expect that the presence of the salt in some way reduces the hydrophobic effects associated with the methyl side group of alanine, thus showing a small increase in solubility.

### 3.4 Determination of Change in Enthalpy

Based on thermodynamics, we know that the solubility of any solute in the liquid phase depends not only on its behavior in the liquid state, but also the stability of the solute in the solid phase. This stability of the solid phase is related to the change in enthalpy between the solid state and that of an infinitely dilute solution. The phase behavior of binary systems in the region where one component is infinitely dilute is described by Henry’s Law. The Henry’s law constant, *K*, is obtained from the limiting slope of the fugacity, *f̂*, versus mole fraction, *x*, curve (*f̂* = *xK* as *x* → 0). The temperature dependence of the *Henry’s Law constant*, for the case where the solid (the amino acid or dipeptide) goes to infinite dilution in the liquid (water) is given by:

\[
\left( \frac{\partial \ln (K / f^s)}{\partial T} \right)_{P,x} = -\frac{\Delta H^{s→∞}}{RT^2}
\]  

where *f*^s^ is the fugacity of the pure solid, *T* is the temperature in Kelvin, *R* is the gas law constant, *ΔH* is the change in enthalpy, with the superscript *s*→ *∞* signifying that the enthalpy change is for the solid state going to a state of infinite dilution, and the subscripts *P*,*x* indicating that the relation is valid for constant
pressure and composition. A more detailed mathematical description can be found in Appendix B.

We are interested in how the solubility changes as a function of the temperature in this limiting case. Therefore we integrate equation 3.1, assuming that the change in enthalpy, $\Delta H^{\text{g-v}}$, is independent of temperature over the temperature range of interest to obtain:

$$\ln(x) = -\frac{\Delta H^{\text{g-v}}}{R}\left(\frac{1}{T} - C\right)$$  \hspace{1cm} (3.2)

where $T$ is the solution temperature, $x$ is the mole fraction of the solute, and $C$ is a constant of integration.

Since the solubility of the amino acid and dipeptide systems considered in this work is very low, Henry's law should be applicable. In figure 3.10 the natural log of the mole fraction as a function of reciprocal absolute temperature is plotted and indeed we do obtain a straight line. From the slope, the change in enthalpy for the amino acids and dipeptides on going from a solid to infinite dilution in water can be determined. While the accuracy of these results is signified by the $R^2$ values of these lines being very close to one, the applicability is limited to the range of temperatures studied and should be applied cautiously outside this range. To extend the range of applicability, additional solubility studies over a larger temperature range are needed. In table 3.3, the enthalpic changes ($\Delta H^{\text{g-v}}$) for the amino acids determined from our experimental data are reported and compared with the values reported in the CRC Handbook of Physics and Chemistry\textsuperscript{122} and Lange's Handbook of Chemistry.\textsuperscript{123} For the dipeptides, data
Figure 3.10. Henry’s Law (Log of mole fraction, $x$, as a function of inverse temperature (1/T)) applied to the dipeptides gly-asp (▲), asp-gly (■), and β-asp-gly (♦).

Table 3.3. Enthalpic Changes and Melting Temperatures

<table>
<thead>
<tr>
<th>Enthalpic Change (kcal/mol)</th>
<th>Gly</th>
<th>Asp</th>
<th>Gly-Asp</th>
<th>Asp-Gly</th>
<th>β-Asp-Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>This work</td>
<td>4.18</td>
<td>7.56</td>
<td>10.75</td>
<td>12.29</td>
<td>7.48</td>
</tr>
<tr>
<td>Calorimetric$^{97}$</td>
<td>3.8</td>
<td>7.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solubility$^{97}$</td>
<td>3.37</td>
<td>7.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melting Temperature (K)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This work</td>
<td>496.0</td>
<td>697.7</td>
<td>414.2</td>
<td>423.3</td>
<td>678.0</td>
</tr>
<tr>
<td>CRC$^{122}$</td>
<td>563</td>
<td>543</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lange$^{123}$</td>
<td>506</td>
<td>543</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
were not available for comparison and we are reporting here, for the first time (to our knowledge), the enthalpy changes for these compounds.

If we assume that the solute forms a nearly ideal solution with water, the constant of integration in equation 3.2 is the inverse of the solute melting temperature, which is given by:

$$T_m = \left( \frac{1}{T} - \ln(x) - \frac{\rho}{\Delta H^\infty} \right)^{-1}$$  \hspace{1cm} (3.3)

This assumption of ideal solution behavior serves as a test for our analysis in that the values of the integration constant for the real system would be expected to be in the range of that if it were an ideal solution. In table 3.3 the melting temperature for the amino acids and dipeptides are reported along with data, where available, from the literature. Because of the approximations made, we consider the new values reported here for the melting temperature of the dipeptides to be only rough approximations.

Based on equation 3.2 and the results of the temperature study that is shown in figure 3.2, we can assess the validity of our results. Our results experimentally show that gly has a significantly higher solubility than asp throughout the temperature range. Since asp has a significantly higher melting point as determined from this temperature range, the solubility of asp is expected to be lower. Also, recall that the slope of solubility as a function of temperature increased for asp greater than that for gly (-3:2). The ratio of the enthalpic changes, which determines the degree to which solubility is affected by changes in temperature to a first approximation,\(^1\) of asp to gly is approximately 3.8:2.1
from our results. Recall that there is very little difference in the solubility of β-asp-gly as compared to the mono-peptide asp. Note also that the change in enthalpy and the melting temperatures of these two compounds are very close and hence we expect that they should have similar solubilities. For the gly-asp and asp-gly reverse dipeptide pair, the melting temperatures are similar. The somewhat lower melting temperature for gly-asp indicates a higher solubility than its reverse dipeptide at any temperature. There is a notable difference in the enthalpic change with a ratio of asp-gly to gly-asp of 1.15:1.0 and the effect of temperature on solubility of these dipeptides as seen from figure 3.2b and table 3.1 is 1.4:1.0. Hence the trends that are observed in the theoretical analysis do correspond to what is expected physically, based not only on our data but also of the previously reported data for amino acids.

### 3.5 Conclusions

We have initiated a systematic study of solution behavior of amino acids and small peptides, where the effects of temperature, salt type, and salt concentration are considered. By choosing glycine, alanine, and aspartic acid, entities with neutral, hydrophobic, and hydrophilic side chains, respectively, are investigated. The changes in temperature have a significantly greater effect on solubility than do salt type and concentration. Nonetheless, salt type and concentration can be very important variables depending on the applications. We have resolved the contradictory results reported in the literature for studies of glycine solubility in NaCl and KCl aqueous solutions. Based on our extended
study with glycine as well as our results for alanine, we believe the trends of our results are the more physically reasonable. The effect of sequence on solution behavior is clearly identified. \textit{Ab initio} quantum calculations as well as experimental FTIR and Raman spectroscopy results support our hypothesis regarding the role of intramolecular hydrogen bonding in the solution thermodynamics of these systems. Nonetheless, additional quantum mechanical studies, in particular those of aqueous solutions of aspartic acid, would be helpful to further understand the behavior at neutral pH, the pH at which biological processes normally occur. Values for enthalpic changes for the dipeptides of \textit{asp} and \textit{gly} are reported, and it is demonstrated that the solid state is also an important consideration in evaluating the solubility trends of these substances. Rough approximates for the melting temperatures of the \textit{asp} and \textit{gly} dipeptides are also given.

Our goal is to be able to capture these effects in a theoretical model that is much less computationally intensive than the present quantum methods available. Statistical mechanics based approaches have been developed to include the effect of molecular shape (chain-like and/or branched chains) and hydrogen bonding on solution properties. By developing a suitable database of experimental results, a theoretical approach can be designed and verified.
CHAPTER 4: A DIPOLAR CHAIN EQUATION OF STATE

4.1 Introduction

Dipolar interactions are known to have a significant effect on the phase behavior of numerous systems from those containing esters and ketones to polymeric systems of polar polymers and co-polymers as well as a host of biochemical systems. This is evident from the non-ideal behavior seen in mixtures in which one of the components is polar and another non-polar. For example, 2-butane/n-heptane and 3-pentanone/n-heptane both form azeotropic mixtures while a n-heptane/n-hexane mixture or a 2-butane/n-3-pentanone mixture is nearly ideal. Many equations of state, including the original SAFT equation of state due to Chapman et al.\textsuperscript{53-55} and many of its modified forms, do not include a separate long-range polar term. If such an equation of state is applied to polar fluids, the polar interaction is effectively included in the van der Waals attraction resulting in an artificially large pure fluid attraction energy, $U_i$. As we will show in chapters 5 and 6, this causes poor prediction of properties for mixtures of polar and non-polar fluids.

For the interaction between two components designated as 1 and 2, the cross van der Waals attraction is typically estimated from the geometric mean of the two pure fluid attraction energies, $U_{12} = \sqrt{U_{11} U_{22}} (1 - k_{12})$, where $k_{12}$ is a binary interaction parameter. The binary interaction parameter is used to correct for any deviation of the mixture properties from the geometric mean of the pure fluid attraction energies. Let component 1 be a polar fluid and component 2 a non-
polar fluid. Then, if \( U_{11} \) is artificially large due to the neglect of polar interactions, the binary interaction parameter, \( k_{12} \), must then be large and possibly composition dependent to compensate for the fact that there is no multipolar interaction between components 1 and 2. This reduces the predictive ability of the equation of state since experimental data would be necessary to determine the value of \( k_{12} \) for each system of interest. The necessity for this large, state dependent \( k_{12} \) can be removed by correctly including the effect of the multipolar interaction.

Among the statistical mechanics based perturbation theories developed for polar fluids, the \( u \)-expansion is the most widely applied.\(^{64,124}\) However, the \( u \)-expansion is valid only for mixtures of polar spheres;\(^{64}\) attempts to extend the theory to non-spherical molecules have failed. In previous efforts to develop a theory for non-spherical polar fluids, researchers primarily studied diatomic molecules with the multipole moment oriented along the molecular axis of the molecule. This was considered to be the simplest realistic model to study since it is linear and representative of many small polar molecules. For example, Wojcik \textit{et al.}\(^{125}\) investigated the thermodynamics and structure of mixtures of quadrupolar hard diatomics and non-polar molecules, calculating the full angular distribution functions by molecular simulation. Although the convergence of various perturbation theories was tested with molecular simulation, an accurate theory was not found.

While no verifiably rigorous theory has previously been developed, the necessity of modeling non-spherical, polar molecules for engineering applications has resulted in the development and use of approximate models. The model that
has primarily been used approximates the dipolar non-spherical molecule as a large sphere with the dipole positioned at the center and with the sphere diameter chosen such that the molecular volume is preserved.\textsuperscript{22,60,70,84-86,126-128} Although this model maintains the idea that the contribution to the fluid properties of a molecule with a single dipolar functional group becomes weaker as the size of the molecule increases, it is known to be in poor agreement with molecular simulation results.\textsuperscript{64,125} Another crucial limitation of this approach is that it allows only a single polar group per molecule thus excluding molecules with multiple polar sites such as di- and tri-ketones, di- and tri-ethers, polar co-polymers and many biochemicals.

Other approaches have used site-site perturbation theory accounting for molecular shape through an interaction volume parameter or diagrammatic expansion techniques. The PACT approach derived by Walsh \textit{et al.}\textsuperscript{129} uses an interaction site perturbation theory with polar interactions. This model accounts for the shape of the molecule through an interaction volume parameter that depends on the molecular shape. Sear\textsuperscript{130} applied a diagrammatic expansion technique similar to that of Wertheim to study the effect of chain formation by dipolar hard spheres at low density, which is energetically favorable. This approach is applicable only at low temperature (or high dipole moment) and low density because it neglects the long-range multi-body interactions. In a model for water, Müller and Gubbins applied the u-expansion to a fluid of spheres and allowed these spheres to associate using Wertheim's associating fluid theory.\textsuperscript{131}
The theory showed reasonable agreement with molecular simulation results and in comparison with experimental data for water.

In the present work a method is proposed to account for dipolar interactions in chain equations of state for mixtures and the differences between this new approach and previous models is demonstrated. Our approach is to use physical insight concerning the molecular structure of fluids to suggest reasonable approximations to theories of molecular fluids. A statistical mechanical theory for chains with dipolar functional groups based on Wertheim's thermodynamic perturbation theory of first order\textsuperscript{73}-\textsuperscript{76} is presented. Since Wertheim's theory is the basis for the statistical associating fluid theory (SAFT), the new approach is similar to that used in developing the SAFT equation of state. Jog and Chapman showed that for pure dipolar hard sphere chains with the dipole oriented perpendicular to the molecular axis the theoretical results for thermodynamics are in excellent agreement with molecular simulation results.\textsuperscript{12} This model is realistic for many compounds such as ketones, ethers, and some polar polymers. In this work, the approach of Jog and Chapman is extended to mixtures of dipolar chains. First a brief overview of SAFT is given, followed by a comparison of the molecular sphere and segment approaches to the dipolar contribution, and the development of the segment approach for polar mixtures.

4.2 The Statistical Associating Fluid Theory

Recall from chapter 2 that the statistical mechanics based perturbation theories of liquids are typically expansions in terms of either the Helmholtz free
energy or the distribution function. The Statistical Associating Fluid Theory (SAFT) is an expansion in terms of the former. Here, we present a low-density derivation of the SAFT equation of state.

4.2.1 The Reference Fluid and Dispersion Contributions

The total residual Helmholtz free energy is given in terms of a perturbation expansion:

\[
a^{\text{res}} = a^{\text{hs}} + a^{\text{dispersion}} + a^{\text{chain}} + a^{\text{association}}
\]  

(4.1)

where the Helmholtz free energy is residual to an ideal gas at the same temperature and density as the fluid of interest. As depicted in figure 4.1, associating molecules are essentially modeled by removing the association sites and covalent bonds to form a reference fluid of segments. The theory builds the associating molecules beginning with this reference fluid of segments, which is typically modeled as a hard sphere fluid with a dispersion contribution as a perturbation. The change in free energy due to chain formation and association are then added as perturbations to the segment free energy. The equation of state can be used with a variety of reference fluids, with the hard-sphere fluid most commonly applied. Here, we use the hard sphere contribution \(a^{\text{hs}}\) for mixtures as given by Mansoori et al.\textsuperscript{132}

\[
a^{\text{hs}}_{/RT} = \frac{6}{\pi \rho} \left[ \frac{\zeta^{3}}{\zeta(1-\zeta)^{2}} - \left( \frac{\zeta^{3}}{\zeta_{0}^{3}} \right) \ln \left( \frac{\zeta}{\zeta_{0}} \right) \right],
\]  

(4.2)

where \(\zeta_{\omega03}\) is a function of the molar density and given by\textsuperscript{55}

\[
\zeta_{x} = \frac{\pi}{N_{A}} \rho \sum x_{i} m_{i} d_{ii}^{x}.
\]  

(4.3)
$N_A$ is Avagadro's number, $\rho$ is the number density, $x_i$ is the mole fraction of component $i$, $m_i$ is the number of segments per molecule of component $i$ and $d_i$ is the diameter of segments of type $i$.

Various approaches to the dispersion term ($a_{\text{dispersion}}$) can be used. For example, CK-SAFT\textsuperscript{58,59} uses the Chen & Krewlewski\textsuperscript{57} (CK) dispersion term, LJ-SAFT\textsuperscript{133,134} uses a Lennard-Jones potential for this contribution, and PC-SAFT\textsuperscript{61,135} uses a perturbed chain (PC) term. In this work, the CK and PC forms of SAFT are used. In the CK form, the equation of state for spheres, which was originally fit to a power series by Alder \textit{et al.}\textsuperscript{136} and the constants refitted by Chen & Krewlewski,\textsuperscript{57} is given by

$$\frac{a_{\text{dispersion}}}{RT} = \overline{m} \sum_i \sum_j D_{ij} \left( \frac{u}{kT} \right)^{\left( \frac{\eta}{\tau} \right)}.$$ \hfill (4.4)

![Diagram of free energy expansion](image)

**Figure 4.1** A cartoon representation of the free energy expansion. Associating molecules are modeled by removing the association sites and covalent bonds to form a reference fluid of segments. The theory builds the associating molecules beginning with this reference fluid, which is typically modeled as a hard sphere fluid with a dispersion energy as a perturbation. The change in free energy due to chain formation and association are then added as perturbations to the segment free energy.
The constants \( D_i \) were fitted to experimental data for argon and \( \eta \) is a reduced density given by \( \eta = \tau \rho m v^0 \), where a close-packed arrangement of segments has been assumed \( (\tau = \pi \sqrt{2}/6) \). \( v^0 \) is the segment molar volume, based on a temperature dependent diameter, and \( v^{(x)} \) is the corresponding temperature independent segment volume at \( T=0 \). The interaction energy, \( u \), is determined from a set of mixing rules and \( \bar{m} \) is the average number of segments per molecule in the mixture. The parameters \( v^{(x)} \), \( u \), and \( m \) are depicted in figure 4.2.

\[ v^{(x)} = \frac{\pi N_A \sigma^3}{6\tau} \]

**Figure 4.2** a) Segment volume based on a temperature independent diameter \( \sigma \). b) \( u \) is the segment-segment interaction energy. c) \( m \) represents the number of segments within a chain molecule.

We also use the PC version of SAFT in which the dispersion term is given by\(^{61}\)

\[ \frac{A_{\text{disp}}}{RT} = \frac{A_1}{RT} + \frac{A_2}{RT}, \quad (4.5) \]

where \( A_1 \) and \( A_2 \) are given by

\[ \frac{A_1}{RT} = -2\pi \rho \ell (\eta, \bar{m}) \sum_i \sum_j x_i x_j \bar{m}_i (\varepsilon_{ij} / kT) \sigma^3 \quad (4.6) \]
and
\[
\frac{A_v}{RT} = -\pi \rho \bar{m} \left( 1 + Z^{hc} + \rho \frac{\partial Z^{hc}}{\partial \rho} \right)^{-1} l_z(\eta, \bar{m}) \sum \sum x_i x_j m_i m_j \left( \frac{\xi_{ij}}{kT} \right) \sigma_{ij}^2
\] (4.7)

with
\[
\left( 1 + Z^{hc} + \rho \frac{\partial Z^{hc}}{\partial \rho} \right)^{-1} = 1 + \frac{8\eta - 2\eta^3}{(1 - \eta)^4} + \frac{20\eta - 27\eta^2 + 12\eta^5 - 2\eta^6}{[(1 - \eta)(2 - \eta)]^2}. \ 	ag{4.8}
\]

\( l_1 \) and \( l_2 \) are power series in density (\( \eta \)) and the coefficients of the series are functions of chain length. The coefficients in the \( l_1 \) and \( l_2 \) expressions were fit to a large set of experimental pure-component data for the n-alkanes. The parameters in this case are \( \sigma, \varepsilon, \) and \( m. \) \( \sigma, \) a temperature independent diameter, and \( m, \) the number of segments in a chain-like molecule, are the same as those depicted in figure 4.2a and 4.2c respectively. \( \varepsilon \) is the potential well depth of a modified square-well potential model and the dispersion energy, based on equations 4.6-4.8, depends on the average chain length.

4.2.2 The Contribution from Associating Fluids

SAFT is an equation of state that, unlike many others, explicitly and accurately accounts for association of molecules (i.e., any aggregation of molecules due to intermolecular interactions such as hydrogen bonding). Any aqueous system will involve some degree of association as will numerous gas phase systems such as low-density carboxylic acids. The association term in SAFT is based on Wertheim's thermodynamic perturbation theory of first order. The following is a brief derivation of the expression for the Helmholtz free energy for mixtures of fluids with multiple association sites.
Consider an associating fluid such as carboxylic acids at low density. Here, "low density" refers to conditions at which a non-associating fluid would be considered an ideal gas. Since the fluid is at low density, the non-ideality in the fluid is primarily due to association. In order to determine the thermodynamic properties of such a fluid, consider a system with a fixed number of molecules (N), volume (V), and temperature (T), which are the natural variables of the thermodynamic potential given by the Helmholtz free energy (A). For a low-density mixture at fixed N, V, and T, the Helmholtz free energy is expressed as

$$A = N\mu - PV$$ \hspace{1cm} (4.9)

where $\mu_i$ is the chemical potential of species $i$ in the mixture and $P$ is the system pressure. The pressure of a low-density fluid is given by $P = \bar{N}kT/V$, where $\bar{N}$ is the number of independent or "free" molecules. For a non-associating fluid, the number of molecules equals the number of "free" molecules. For an associating fluid, however, any molecules bonded to each other form a single entity. Thus, association reduces the number of "free" molecules in the system.

We will use the notation that the association sites on a molecule are labeled as A, B, C, etc. and $X_A$ is defined as the fraction of molecules not bonded at site A. We also define $\Gamma$ as the set of association sites ($\Gamma=\{A,B,C,\ldots\}$). In figure 4.3, large circles represent molecules and association sites are denoted by the small circles, with each different shaded small circle indicating the various site types in $\Gamma$. By associating, these molecules form clusters of various shapes and sizes. For example, molecules with only one association site will form dimers if they associate with other molecules with a single association site. Or,
they can be located at the ends of chain-like structures. Molecules with two association sites can form chain-like structures, while molecules with more than two association sites serve as branch points in larger tree-like clusters. The types of structure formations are limited to linear and branched tree-like structures (i.e., no rings) and only single bonds between molecules occur. Thus, the number of "free" molecules is reduced by one for each bond that is formed. The quantity of \((1-X_A)\) is then the fraction of molecules bonded at sites of type A and, since two molecules form a single bond, the fraction of bonds formed is given by \((1/2)\sum_A (1-X_A)\). The number of "free" molecules for an associating mixture can be expressed as:

\[
\bar{N} = N \left[ 1 - \frac{1}{2} \sum_A (1-X_A) \right]
\]  

(4.10)

where the summation is over the set of association sites.

Figure 4.3 Cartoon representation of possible association. While each site can bond only once, a given molecule can bond as many times as it has sites, thus forming clusters of molecules.

Since this mixture of associated clusters is being treated as an ideal gas mixture, the chemical potential of a monomer can be written from statistical mechanics. From chemical equilibria we know that the chemical potential of a
component as a monomer is the same as the chemical potential of that component in any associated cluster. Therefore, for the low-density approximation,

\[ \mu_i = \mu_i^{\text{monomer}} = \mu_i^{\text{inner}} = \mu_i^{\text{inner}} = \ldots = kT \ln(\rho_o \Lambda^3), \]  

(4.11)

where \( \rho_o \) is the number density of monomers and \( \Lambda \) is the de Broglie wavelength.

From eq 4.9, we can write the Helmholtz free energy for our associating fluid and for a non-associating reference fluid. The change in Helmholtz free energy due to association is then given by

\[ \frac{A - A^{\text{Ref}}}{NkT} = \ln(X_o) - \frac{1}{2} \sum_{\alpha \in \Gamma} (X_A - 1) \]  

(4.12)

where \( X_o = \rho_o / \rho = \text{the fraction of monomers}. \)

Assuming the bonding at one site is independent of the bonding state at other sites on a given molecule, the fraction of monomers can be written as

\[ X_o = \prod_A X_A. \]  

(4.13)

Note that by making this assumption, steric hindrances are not taken into account. Let M to be the number of association sites on a molecule; then the Helmholtz free energy becomes\(^{53}\)

\[ \frac{A - A^{\text{Ref}}}{NkT} = \sum_{\alpha \in \Gamma} \left( \ln(X_A) - \frac{X_A}{2} \right) + \frac{M}{2}. \]  

(4.14)

### 4.2.2.1 The fraction of monomers for fluids with multiple association sites

Consider a model dimerizing fluid. A simple pair potential of a hard sphere core with a square well association interaction that is dependent on
orientation can be used. Figure 4.4 is a cartoon depicting the association, where
the large circles represent molecules and the small filled circles denote
association sites. As shown in figure 4.4a, the range of the square well
interaction is chosen such that a molecule can associate with only one
other molecule at a time. The energy of this association is defined as \(-\varepsilon\). When two
molecules are associated, additional molecules are prevented from associating
with the dimer as depicted in figure 4.4b. When two of the monomers are in the
correct orientation and distance from each other, they associate to form a dimer
(figure 4.4c). In this case, there are two methods that can be used to calculate
the change in internal energy on association. For this model fluid, the change in
internal energy due to association is equal to the number of molecules in the
bonding orientation multiplied by \((-\varepsilon/2)\). Also, thermodynamic differentiation of
the Helmholtz free energy gives an expression for the change in internal energy.
By applying both methods and setting the equations equal, a differential equation
for \(X_o\) is obtained.

The number of molecules in the hydrogen bonding orientation equals the
number of molecules that are associated plus the number of monomers that are
in the hydrogen bonding orientation. The monomers in the hydrogen bonding
orientation must be included since, even under conditions of no association (e.g.,
infinite temperature or zero energy of association), molecules (or monomers) are
in the hydrogen bonding orientation. Wertheim's precise definition of a monomer
allows monomers to be in the hydrogen bonding orientation and remain as
monomers. This is a necessary requirement in order for the fraction of monomers
Figure 4.4 Large circles represent molecules. The smaller dark circles denote association sites.  

a) The square-well potential model with a hard sphere core and associating interaction energy of $-\varepsilon$.  
b) If a site is bonded, it cannot bond with an additional site.  
c) Two molecules with association sites approaching each other in the correct orientation and within the minimum distance will associate. (Adapted from Chapman.\textsuperscript{54})

to become exactly one when the association energy is zero. The number of molecules in the bonding orientation is given by

$$N' = N(1 - X_o) + N X_o \rho_o \int_{\text{Bond Volume}} g_{\infty}(12) \, d(12)$$  \hspace{1cm} (4.15)

where the first term on the right hand side is the number of associated molecules and the second term is the number of monomers that are in the bonding orientation. In the second term, $g_{\infty}(12)$ is the monomer-monomer pair correlation function and the integral represents an unweighted angle average over all possible bonding orientations and an integration over all possible
bonding separations between a pair of monomers. The integral times $\rho_o$ represents the number of monomers in the bonding orientation of a monomer and $N X_o$ is the number of monomers in the fluid.

In the perturbation form of Wertheim's theory, the monomer-monomer pair correlation function is approximated by the pair correlation function of the non-associating reference fluid. In the present case, the reference fluid is a hard sphere fluid. With this approximation, the change in internal energy due to association can be written as

$$\Delta U^{\text{assoc}} = -\frac{\varepsilon}{2} \left[ N (1 - X_o) + N N \rho X_o^2 \int g_{\text{Ref}}(12) \, d(12) \right]$$ \hspace{1cm} (4.16)

By differentiating the Helmholtz free energy, the change in internal energy on association can also be obtained:

$$\Delta U^{\text{assoc}} = \left( \left. \frac{\partial (A - A^{\text{Ref}})}{\partial (1/T)} \right|_{N,V} \right)_{N,V} \quad \text{.}$$ \hspace{1cm} (4.17)

Thus, by differentiating the expression for the Helmholtz free energy (eq 4.14), and equating the two expressions for the change in internal energy, the differential equation for the fraction of monomers results. By integrating this expression with the boundary condition that the fraction of monomers goes to one as the temperature goes to infinity ($X_o \rightarrow 1$ as $T \rightarrow \infty$), a quadratic equation in $X_o$ results:

$$X_o + \rho X_o^2 \left[ \exp(\varepsilon/kT) - 1 \right] \int g_{\text{Ref}}(12) \, d(12) = 1 \quad \text{.}$$ \hspace{1cm} (4.18)
This equation is a statement of the fact that the fraction of monomers plus the
fraction of associated molecules equals one.

Wertheim's theory may be generalized for fluids of multiple association
sites. In the notation of Chapman et al.,\textsuperscript{53} this gives

\[
X_A + \rho \sum_{B \in \Gamma} X_B \left[ \exp(\epsilon_{AB}/kT) - 1 \right] \int_{\text{Bond Volume}} g_{\text{Ref}}(12) \, d(12) = 1. \tag{4.19}
\]

The fraction of molecules not bonded at site A plus the fraction of molecules
bonded at site A to the various sites B on another molecule must equal one.

Solving for the fraction of molecules not bonded at site A yields

\[
X_A = \frac{1}{1 + \rho \sum_{B \in \Gamma} X_B \left[ \exp(\epsilon_{AB}/kT) - 1 \right] \int_{\text{Bond Volume}} g_{\text{Ref}}(12) \, d(12)}. \tag{4.20}
\]

4.2.2.2 Extension to Mixtures of Associating Fluids

Chapman et al.\textsuperscript{53} extended the theory to mixtures of associating fluids.

The change in Helmholtz free energy due to association is then

\[
\frac{A - A_{\text{Ref}}}{NkT} = \sum_i x_i \left[ \sum_{A \in \Gamma^i} \left( \ln(X_A^{(i)} - \frac{X_A^{(i)}}{2}) + \frac{M_i}{2} \right) \right], \tag{4.21}
\]

where \( x_i \) is the mole fraction of molecules of component \( i \), \( X_A^{(i)} \) is the fraction of
molecules of component \( i \) that are not bonded at site A, and \( M_i \) is the number of
association sites on a molecule of component \( i \). The first sum is over all
components and the second sum is over all sites on a molecule of component \( i \).

For mixtures, the fraction of molecules not bonded at site A of species \( i \) is given by
\[ X_A^{(i)} + X_A^{(i)} \sum_i \rho \cdot x_i \sum_{\text{Bond}} X_B^{(j)} \left[ \exp \left( \frac{\epsilon_{AB}^{(i,j)}}{kT} \right) - 1 \right] \int_{\text{Volume}} g_{\text{Ref}}^{(i,j)}(12) \, d(12) = 1 \quad (4.22) \]

where \( \epsilon_{AB}^{(i,j)} \) is the energy of interaction between site A on species \( i \) and site B on species \( j \) and \( g_{\text{Ref}}^{(i,j)} \) is the reference fluid pair correlation function between molecules of species \( i \) and \( j \). This equation can be interpreted as the fraction of molecules of type \( i \) not bonded at site A plus the number of molecules of type \( i \) associated at site A to a site B on any of the various components \( j \) must equal one. This interpretation is used to calculate the extent of like pair and unlike pair association in mixtures. Comparisons with molecular simulation results show excellent agreement.\(^{54}\) Solving for the fraction of component \( i \) not bonded at site A yields

\[
X_A^{(i)} = \frac{1}{1 + \sum_i \rho \cdot x_i \sum_{\text{Bond}} X_B^{(j)} \left[ \exp \left( \frac{\epsilon_{AB}^{(i,j)}}{kT} \right) - 1 \right] \int_{\text{Volume}} g_{\text{Ref}}^{(i,j)}(12) \, d(12)} \quad (4.23)
\]

4.2.3 The Chain Term\(^{53,54}\)

The derivation of the chain term is based on the associating fluid theory where some of the association sites are replaced with covalent-bonding sites. The covalent bonding sites differ from association sites in that the former has an infinitely strong bonding potential which forces complete bonding. Consider a fluid of \( M \) components (i.e. \( M \) types of segments) with each type of segment having two bonding sites, with the exception of the 1\(^{st}\) and \( M \)^{th} species which have only one bonding site. By restricting the type of segment that each one can bond with, a chain is formed. For example, segments of type 1 can only bond with
segments of type 2; segments of type 2 can only bond with segments of type 1 and 3... and segments of type $m$ can only bond with segments of type $m$. (See figure 4.5.) Considering a stoichiometric ratio of types of segments and imposing a restriction that the fraction of monomers of these chain-bonding sites is zero results in chain formation. Note that we are distinguishing between covalent bonding and associating sites. There have been no assumptions made to prevent the presence of additional sites that are associating. Once the chains are formed, the chains can associate with other chains through their association sites. For the formation of a fluid of chains of length $m$, eq. 4.21 becomes

$$\frac{A - A_{\text{Ref}}}{N_{\text{segments}} kT} = \sum_{\beta=1}^{m} \frac{1}{m} \left[ \sum_{\Gamma \in \Gamma_{(\beta)}} \left( \ln(X_{A}^{(\beta)}) - \frac{X_{A}^{(\beta)}}{2} \right) + \frac{M_{\beta}}{2} \right]$$

(4.24)

since $x_1 = x_2 = \cdots = x_m = 1/m$. The first summation is over all components $\beta$ and the second is over all sites of type $A$ where $\Gamma$ is the set of site types. Note that $m$ is independent of the summation and that $m/N_{\text{segments}} = N_{\text{chain}}$. For a mixture of chains, the change in the Helmholtz free energy due to chain formation, or covalent bonding, can be expressed:

$$\frac{\Delta A_{\text{chain}}}{N_{\text{chain}} kT} = \sum_{i} \rho_{c}^{(i)} \left[ \sum_{\beta=1}^{m} \left[ \sum_{\Gamma \in \Gamma_{(\beta)}} \left( \ln(X_{A}^{(\beta)}) - \frac{X_{A}^{(\beta)}}{2} \right) + \frac{M_{\beta}}{2} \right] \right].$$

(4.25)

![Figure 4.5 Cartoon representation of the covalent bonding scheme. By allowing each species to bond with only one or two other spheres and restricting how the spheres may bond, a chain is formed.](image)

Figure 4.5 Cartoon representation of the covalent bonding scheme. By allowing each species to bond with only one or two other spheres and restricting how the spheres may bond, a chain is formed.
\[
\frac{\Delta A_{\text{chain}}}{N_{\text{chain}} kT} = \sum_{i=1}^c \rho_c^{(i)} \left\{ \sum_{\beta=1}^m \left[ \sum_{\mathbf{A}, \mathbf{r}, \mathbf{p}} \left( \ln X_A^\beta - \frac{1}{2} X_A^\beta + \frac{M_\beta}{2} \right) \right] \right\}. \tag{4.25}
\]

Since we are interested in the case when all the bonding sites are bonded, i.e. \(X_A^\beta = 0\), it is more mathematically convenient to work in terms of the compressibility factor. Recall that the relationship between the compressibility factor and the Helmholtz free energy is given by

\[
\Delta \mathcal{P} \quad \frac{\rho}{kT} = \rho \left( \frac{\partial A_{\text{chain}}}{\partial \rho} \right)_{\text{NkT}}. \tag{4.26}
\]

Thus, taking the derivative of eq. 4.25, gives

\[
\frac{\Delta \mathcal{P}_{\text{chain}}}{\rho kT} = \sum_{i=1}^c \rho_c^{(i)} \left\{ \sum_{\beta=1}^m \left[ \sum_{\mathbf{A}, \mathbf{r}, \mathbf{p}} \left( \frac{\partial X_A^{(\beta)}}{\partial \rho} \right) \left( \frac{1}{X_A^{(\beta)}} - \frac{1}{2} \right) \right] \right\}. \tag{4.27}
\]

All segments on a given chain are assumed to be of equal size and with equal interaction energies. The imposition of a constraint that each site can bond with only one type of site indicates that all \(X_i\)'s behave in the same way, independent of site type. Therefore, the \(X_i\)'s satisfy the same defining equation and they will no longer be distinguished as to what site is being considered. Thus, eq. 4.23 can be written as

\[
X^{(i)} = \frac{1}{1 + \rho_c^{(i)} X^{(i)} \Delta^{(i)}} \tag{4.28}
\]

where

\[
\Delta^{(i)} = \int_{\text{Bond}} \int_{\text{Volume}} \mathcal{g}_{\text{Ref}} (ij) \left( \exp \left( \frac{\mathcal{E}_{\text{assoc}}^{(ij)}}{kT} \right) - 1 \right) d(ij). \tag{4.29}
\]
$g_{ref}$ is the pair correlation function of the reference fluid and the integral is over all orientations, $\omega$, and separations, $r$, of segments $i$ and $j$, where $i$ and $j$ are segments of the same chain. Taking the density derivative of eq. 4.28 and substituting into eq. 4.27 yields

$$\frac{\Delta P_{\text{chain}}^{\text{chain}}}{\rho kT} = \frac{1}{2} \sum_{i=1}^{r} \rho_c^{(i)} \sum_{j=1}^{m} \sum_{\alpha \in \Gamma^{(i)}} \left( (X^{(i)})^2 \left[ \frac{\Delta^{(j)} \rho_c^{(i)}}{\rho} + \rho_c^{(i)} \left( \frac{\partial \Delta^{(j)}}{\partial \rho} \right) \right] \right). \tag{4.30}$$

Since each term in the summand is independent of site type and type of segment component, the two inner summations can be directly evaluated. Two spheres have a single site, which yields a factor of 2, and $(m-2)$ segments have two sites each, which yields $2(m-2)$. Summing these results gives a factor of $2(m-1)$. Thus, in the limit of total bonding with $X^{(i)} = 0$ and after some algebra, eq. 4.30 becomes

$$\frac{\Delta P_{\text{chain}}^{\text{chain}}}{\rho kT} = \sum_{i=1}^{r} \rho_c^{(i)} (m_i - 1) \left[ 1 + \rho \left( \frac{\partial \ln \Delta^{(i)}}{\partial \rho} \right) \right]. \tag{4.31}$$

The only term in eq. 4.29 for $\Delta^{(i)}$ that depends on the density is the reference fluid pair correlation function. This term is general so that any intramolecular bonding function can be used in $\Delta^{(i)}$, which allows the study of the thermodynamics of non-rigidly bonded molecules. Moreover, the effects of bond-bending and bond-vibrations can also be studied independently provided that the appropriate intramolecular potential function is used. If we use a hard sphere reference fluid and restrict the bonding potential to be nonzero only at hard sphere contact, then eq. 4.31 simplifies to
\[
\frac{\Delta P_{\text{chain}}}{\rho \ kT} = -\sum_{i} \rho_{i}^{(t)} (m_{i} - 1) \left[ 1 + \rho \left( \frac{\partial \ln g_{i}^{(t)}(d_{\text{contact}})}{\partial \rho} \right) \right],
\]

where \( g_{i}^{(t)}(d_{\text{contact}}) \) is the hard sphere pair correlation function for the interaction of two spheres \( i \) and \( j \) of equal size in a mixture of spheres and evaluated at the hard sphere contact. The corresponding contribution to the Helmholtz free energy due to bonding \( (a_{\text{chain}}) \) is expressed as

\[
a_{\text{chain}} \frac{RT}{\rho} = \sum_{i} \chi_{i} (1 - m) \ln(g_{i}^{\text{HS}}(d_{i})).
\]

4.2.4 Impact of SAFT

SAFT has proved to be an invaluable equation of state in that it is one of few that can accurately be applied to a wide range of systems, including both associating and non-associating, from alkane/alkane mixtures\(^{133,138,139}\) to polymers\(^{140-145}\) to high-pressure equilibria\(^{146,147}\) and a variety of other highly nonideal systems.\(^{79}\) One of the most significant benefits arising from the development of SAFT by Chapman\(^{54}\) is the establishment of a methodology allowing application to a plethora of systems including water\(^{126,127,131}\) and electrolytes\(^{148-150}\). More detailed information regarding the variations and applications of the SAFT approach can be found in a recent review by Müller & Gubbins.\(^{79}\)

4.3 Comparison of the Molecular Sphere and Segment Approaches

In this work, chain-like molecules that contain one or more polar functional groups are considered. Most of the prior approaches to account for dipolar
interactions in chain equations of state treat the dipolar molecule as a sphere of volume equal to the molecular volume with an ideal dipole at the molecular center. Thus, as the molecule becomes larger, the effect of the dipole weakens because of the larger effective separation of the dipoles. The effect of non-sphericity of the molecule on the polar term is not considered. In addition, when a molecule has multiple polar groups, the molecular dipole moment is used and as such there is no way to account for differing polarity of functional groups within a molecule. We call this the "molecular sphere approach". In contrast, the approach developed in this work, which we refer to as the "segment approach", explicitly accounts for multiple dipolar functional groups and the non-spherical shape of the molecule in the polar term. The dipolar contribution to the Helmholtz free energy is obtained by dissolving all the bonds in a chain and then applying the $u$-expansion to the resulting mixture of polar and non-polar spherical segments.

We illustrate the difference between the two approaches quantitatively with an example. Consider a pure fluid system, for example a representative alkanone that can be modeled as a chain of five segments as depicted in figure 4.6a. Now, compare this to a molecular sphere with the same volume as shown in figure 4.6b. It is obvious that in both models, the non-polar segments dilute the dipolar interaction. This is consistent with the idea that polar interactions are less significant in long chain than in short chain alkanones. Nonetheless, this effect is artificially exaggerated in the molecular sphere model since the distance of closest approach of dipoles is greater than in the segment approach.
We can demonstrate this by considering the contributions to the Helmholtz free energy of the two methods for pure dipolar fluids. In the $u$-expansion, the dipolar contribution is an infinite series of terms of order second, third and higher. The second and third order terms are calculated explicitly\textsuperscript{151} and the higher order terms are estimated by the Padé approximant of Rushbrooke \textit{et al.}\textsuperscript{71} Using the Padé approximate, the change in free energy due to polar interactions can be written as

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.6.png}
\caption{Representative alkanone: a) segment approach, b) effective molecular sphere approach.}
\end{figure}
\[ A_{\text{polar}} = \frac{A_2}{1 - A_1/A_2} \]  

where \( A_2 \) is the second order term in the perturbation expansion and \( A_3 \) is the third order term.

For the dipolar molecular sphere approach, \( A_2 \) is given by

\[ \frac{A_{2\text{molecule}}}{NkT} = -\frac{2}{9} \pi \left( \frac{1}{kT} \right)^2 \rho \frac{\mu_{\text{molecule}}^3}{d_{\text{molecule}}^3} I_2(\rho^*) \]  

where \( \mu_{\text{molecule}} \) is the molecular dipole moment, \( N \) is the number of molecules, \( k \) is the Boltzmann constant, \( T \) is the absolute temperature, and \( d_{\text{molecule}} \) is the diameter of a sphere with the same volume as the molecule. This diameter is given by

\[ d_{\text{molecule}} = \left( m d_{\text{segment}}^3 \right)^{1/3} \]  

where \( d_{\text{molecule}} \) is the diameter of the effective spherical molecule, \( m \) is the number of segments in the actual molecule and \( d_{\text{segment}} \) is the diameter of the segments. \( I_2 \) is the integral over the pair correlation function of the reference fluid (here a hard sphere fluid):

\[ I_2(\rho^*) = \frac{3\sigma^3}{4\pi} \int g_{\text{HS}}(r, \rho^*) \frac{1}{r^6} dr. \]  

\( \rho \) is the number density of molecules and \( \rho^* = \rho d_{\text{molecule}}^3 = \rho m d_{\text{segment}}^3 \) is the reduced density.

Using the segment approach, we define \( x_\rho \) as the fraction of dipolar segments on a molecule. In the case of a single dipolar segment in a molecule, assuming that all segments contribute equally to the molecular size and
interaction energy, then the fraction of polar segments is inversely proportional to the number of segments in a chain molecule (\( x_\rho = 1/m \)). For the segment approach, the second order contribution, \( A_2 \), in the perturbation expansion for dipolar molecules is given as

\[
\frac{A_2^{\text{segment}}}{NkT} = -\frac{2}{9} \pi \left( \frac{1}{kT} \right)^2 \rho \frac{\mu_{\text{segment}}^4}{d_{\text{segment}}^2} I_2(\rho^*) ,
\]

where \( \rho^* = \rho m d_{\text{segment}}^3 \) is the reduced segment density, in keeping with the segment basis of the dipolar free energy contribution. For molecules with a single dipolar site, the molecular dipole moment is the same as the segment dipole moment.\(^{123}\) For this example of an alkanone, the segment approach and the molecular sphere approach give different results. From equations 4.35-4.38 it can be seen that

\[
\frac{A_2^{\text{molecule}}}{A_2^{\text{segment}}} = \frac{1}{m} .
\]

By also analyzing the third order contribution to the free energy and the Padé, a simple relationship is obtained between the dipolar contribution to the free energy from the two approaches for pure fluids:

\[
\frac{A_2^{\text{molecule}}}{A_2^{\text{segment}}} = \frac{1}{m} .
\]

Therefore, the contribution to the dipolar free energy based on the molecular sphere approach is smaller than that from the segment approach. For molecules with a single polar segment, the contribution of the polar term to the total Helmholtz free energy differs by a factor of the number of segments and hence
the difference between the two approaches is more pronounced as the chain molecule becomes longer. For mixtures, the relationship between the two approaches is more complex. Given that our segment approach has been shown to be in excellent agreement with molecular simulation results, we conclude that the commonly applied molecular sphere approach underestimates the effect of dipolar interactions for chain like molecules. In the following section the segment approach is extended to mixtures and the predictive capabilities of the resulting equation of state are demonstrated in chapters 5 and 6.

4.4 Development of the Polar Term Applicable to Mixtures

As has been shown previously for pure polar fluids, the dipolar term can be added to the perturbation expansion of the chain equation of state used by SAFT. The total residual Helmholtz free energy is then given by

\[ a^{res} = a^{hs} + a^{dispersion} + a^{polar} + a^{chain} + a^{association} \]  \hspace{1cm} (4.41)

In section 4.2, each of these contributions is described, with the exception of the polar term. The hard sphere \((a^{hs})\) and the chain \((a^{chain})\) terms in the original SAFT equation of state are unaffected by the presence of a dipole term as shown by Jog and Chapman. The species in this work are non-associating so that the association \((a^{association})\) contribution to the Helmholtz free energy is zero. For non-associating, non-polar fluids, SAFT has three pure component parameters: namely segment volume \((v^0)\), segment dispersion energy \((\ell^0/K)\) and chain length \((m)\). In order to account for multiple dipolar segments within a molecule,
the functional group dipole moment is required. We also define an additional parameter, $\chi_p$, as the fraction of dipolar segments in a chain. For the polar contribution ($A^{\text{polar}}$), the $u$-expansion is applied to the mixture of polar and non-polar spheres resulting from the dissolving of all the bonds in a chain. This is depicted in the cartoon representation of the free energy expansion for a polar-associating fluid in figure 4.7. The associating molecules with polar sites are modeled by removing the association sites and covalent bonds to form a reference fluid of a mixture of polar and non-polar segments. The theory builds the polar-associating molecules beginning with this reference fluid of polar segments, which is modeled as a hard sphere fluid with the polar and dispersion contributions as a perturbation. The chain and associating terms are then applied to this reference fluid mixture of polar and non-polar spheres.

$$A = A^{\text{segment \_ polar}} + \Delta A^{\text{chain}} + \Delta A^{\text{association}}$$

Figure 4.7 A cartoon representation of the free energy expansion for a polar-associating fluid. As before, the associating molecules (now with polar sites) are modeled by removing the association sites and covalent bonds to form a reference fluid of a mixture of polar and non-polar segments. The theory builds the polar-associating molecules beginning with this reference fluid of polar segments, which is now modeled as a hard sphere fluid with the polar and dispersion contributions as a perturbation.
As shown previously,\(^1\)\(^2\) the dipolar contribution to the Helmholtz free energy is accurately obtained by dissolving all the bonds in a chain and then applying the u-expansion to the resulting mixture of polar and non-polar spherical segments. The polar contribution is written in the Padé approximant form of Rushbrooke et al.\(^7\)\(^1\)

\[
a_{\text{polar}} = \frac{a}{1 - a/a_n}. \tag{4.42}
\]

For mixtures, \(A_s\) and \(A_{st}\), allowing for multiple dipolar segments, gives

\[
\frac{A_s}{N k T} = -\frac{2\pi}{9} \frac{\rho}{(kT)^2} \sum_i \sum_i x_i m_i x_j m_j \frac{\mu_i^2 \mu_j^2}{\sigma_i^2} l_{ij} \tag{4.43}
\]

and

\[
\frac{A_{st}}{N k T} = \frac{5}{162} \frac{\rho^2}{(kT)^3} \sum_i \sum_j x_i x_j m_i m_j x_k x_k \frac{\mu_i^2 \mu_j^2 \mu_k^2}{\sigma_i \sigma_j \sigma_k} l_{ijk}, \tag{4.44}
\]

where \(l_{ij}\) and \(l_{ijk}\) are the angular pair and triplet correlation functions, respectively, and given by

\[
l_{ij} = \frac{3d_i^3}{4\pi} \int g_{HS_i} (r, \rho^{-}) r^{-6} \, dr = \frac{3}{4} \int g_{HS_i} (r^{-}, \rho^{-}) r^{-6} \, dr \tag{4.45}
\]

and

\[
l_{ijk} = \frac{192}{5} \left( \frac{14\pi}{5} \right)^{1/2} \int_0^r \int_0^r \int_0^r r_{ij}^{-6} r_{ij}^{-6} r_{ij}^{-6} g_{ijk} (r_{ij}, r_{ik}, r_{jk}) \psi_{ijk} (\alpha_1, \alpha_2, \alpha_3) \tag{4.46}
\]

The function \(\psi_{ijk} (\alpha_1, \alpha_2, \alpha_3)\) is a well-known function of angles \(\alpha_1, \alpha_2\) and \(\alpha_3\) of the triangle formed by the centers of the three molecules.\(^1\)\(^5\)\(^1\) These integral equations, \(l_{ij}\) and \(l_{ijk}\), are cast in terms of a reduced radius and in this reduced
form are independent of component at a given reduced density. Therefore, the integrals $l_{2,i}$ and $l_{3,ik}$ in equations 4.43 and 4.44 can be related to the corresponding pure fluid integrals, $l_2$ and $l_3$, by

$$l_{2,i} = l_2(\rho d_x^i)$$

(4.47)

and

$$l_{3,ik} = l_3(\rho d_x^i),$$

(4.48)

where $d_x^i = \sum_i m_i x_i d_x^i$. Another approach that should give similar results would be to use a van der Waals one fluid theory approximation for $d_x$.

Other thermodynamic quantities such as pressure and chemical potential can be obtained by differentiation of the Helmholtz free energy equation. The four pure component parameters, the segment size ($v^0$), the segment dispersion energy ($u^0/k$), number of segments per chain ($m$) and the fraction of polar segments in a chain ($x_p$), are fitted to the saturated liquid density and vapor pressure data of the pure components. For molecules with a single polar site, the experimental value for the dipole moment of either the molecule or the functional group can be used. If experimental values are not available, the dipole moment of the functional group calculated from quantum mechanics can also be used. For components with a single polar site and in a homologous series, we use the average value for the dipole moment of the series. For systems with multiple dipolar groups, we use the value of the functional group.
4.5 Conclusions

We have presented a motivation for the study of polar fluids and presented a low-density derivation of the SAFT equation of state on which this work is built. Further we have shown that for pure fluids, the molecular approach significantly underestimates the contribution of the dipolar effects and that the error increases as the chain length increases. We have also extended the pure fluid dipolar chain equation of state to mixtures in which any number of the components may contain polar functional groups as well as allowing multiple polar groups within any of the components.
CHAPTER 5: A PARAMETRIC ANALYSIS OF DIPOLAR SAFT

5.1 Introduction

Parameters are fit to pure component data with the idea that the model itself will be able to sufficiently capture the physics of a mixture using knowledge only of the pure components. This is one of the motivations for inclusion of the polar contribution in SAFT. Although SAFT has been shown to work well for a large variety of systems, it has not been particularly successful with mixtures of polar and non-polar components. Therefore, by explicitly including the long-range polar effects, as outlined in chapter 4, the model is expected to work well for these polar/non-polar mixtures.

The goal of the present chapter is to give a detailed parametric analysis of Polar SAFT, based primarily on a homologous series of ketones, from acetone to tridecanone. The motivation for this analysis is to lay the groundwork for the development of a group contribution approach for SAFT. By considering the differences in the parameters in various homologous series, knowledge regarding the parameter values of functional groups can be gained. In this work, two approaches to the dispersion term (CK and PC) have been used and are included in the present analysis.

5.2 Method

The parameters are fit to pure component vapor pressure and liquid density data by minimizing the deviation between the experimental values and
the model predicted values. The vapor pressures tend to be the more crucial of
the two. Average absolute deviations of less than five percent in the pure
component vapor pressure are needed to give reasonable results for most binary
systems. The values of the four model parameters along with the average
absolute deviations are listed in table 5.1 for Polar CK-SAFT and in table 5.2 for
Polar PC-SAFT. Note that the average absolute deviations in vapor pressures
for 2-octanone and 4-heptanone are significantly greater than the other
compounds. This occurs with both dispersion terms.

Concern regarding these differences led to an investigation of the
accuracy of the reported experimental data. In figure 5.1, the vapor pressure
data for the compounds from 2-butanone through 6-undecanone are plotted as
the natural log versus the inverse of the temperature. A comparison of the vapor
pressure data, along with linear fits, of each compound in the series is presented.
It can be readily seen that the compounds differing only in the position of the
carbonyl group lie very close to each other, for example, 2- and 3-pentanone, 2-
and 3-hexanone, and 2- and 6-undecanone. The vapor pressure curves for the
heptanone and octanone isomers do not lie as neatly. In figure 5.2a, only the 2-
alkanones are shown. Note that the slopes of the majority of the curves
decrease with increasing molecular weight of the compounds. However, the
slope for the 2-octanone data is greater than that of 2-heptanone and 2-
hexanone. In figure 5.2b, the Polar CK-SAFT predicted results of 2-octanone are
shown to have a similar trend as the other 2-alkanones. Thus, the 2-octanone
parameters presented in table 5.1 do agree well with data for the other
### Table 5.1 Parameters for the Polar CK-SAFT model for ketones.

<table>
<thead>
<tr>
<th>Ketones</th>
<th>MW (g/mol)</th>
<th>Temp (K)</th>
<th>v00 (ml/mol)</th>
<th>u0/k (K)</th>
<th>m</th>
<th>Mu (D)</th>
<th>xp</th>
<th>Avg Abs Dev (%) Vap Pres</th>
<th>Liq den</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetone</td>
<td>58.08</td>
<td>275-478</td>
<td>13.24</td>
<td>232.97</td>
<td>2.94</td>
<td>2.7</td>
<td>0.1695</td>
<td>2.2605</td>
<td>1.4404</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>72.107</td>
<td>270-374</td>
<td>14.31</td>
<td>238.58</td>
<td>3.37</td>
<td>2.7</td>
<td>0.1465</td>
<td>2.4038</td>
<td>2.1448</td>
</tr>
<tr>
<td>2-Pentanone</td>
<td>86.13</td>
<td>282-353</td>
<td>14.66</td>
<td>239.40</td>
<td>3.81</td>
<td>2.7</td>
<td>0.1315</td>
<td>1.5459</td>
<td>0.6191</td>
</tr>
<tr>
<td>3-Pentanone</td>
<td>86.13</td>
<td>285-347</td>
<td>14.46</td>
<td>239.45</td>
<td>3.81</td>
<td>2.7</td>
<td>0.13125</td>
<td>1.0599</td>
<td>0.8918</td>
</tr>
<tr>
<td>2-Hexanone</td>
<td>100.16</td>
<td>299-332</td>
<td>14.53</td>
<td>239.58</td>
<td>4.35</td>
<td>2.7</td>
<td>0.11621</td>
<td>0.5687</td>
<td>2.7126</td>
</tr>
<tr>
<td>3-Hexanone</td>
<td>100.16</td>
<td>299-324</td>
<td>14.53</td>
<td>237.85</td>
<td>4.33</td>
<td>2.7</td>
<td>0.1147</td>
<td>0.2663</td>
<td>2.4228</td>
</tr>
<tr>
<td>2-Heptanone</td>
<td>114.19</td>
<td>306-416</td>
<td>14.51</td>
<td>237.99</td>
<td>4.945</td>
<td>2.7</td>
<td>0.1015</td>
<td>0.942</td>
<td>1.1800</td>
</tr>
<tr>
<td>4-Heptanone</td>
<td>114.19</td>
<td>275-384</td>
<td>14.49</td>
<td>232.91</td>
<td>4.91</td>
<td>2.7</td>
<td>0.1001</td>
<td>9.377</td>
<td>1.8816</td>
</tr>
<tr>
<td>2-Octanone</td>
<td>128.22</td>
<td>319-420</td>
<td>14.45</td>
<td>231.50</td>
<td>5.59</td>
<td>2.7</td>
<td>0.0891</td>
<td>10.3032</td>
<td>1.2538</td>
</tr>
<tr>
<td>2-Nonane</td>
<td>142.24</td>
<td>340-422</td>
<td>14.41</td>
<td>233.94</td>
<td>6.20</td>
<td>2.7</td>
<td>0.08065</td>
<td>1.5693</td>
<td>1.1559</td>
</tr>
<tr>
<td>5-Nonane</td>
<td>142.24</td>
<td>301-356</td>
<td>14.59</td>
<td>232.38</td>
<td>6.12</td>
<td>2.7</td>
<td>0.08305</td>
<td>2.132</td>
<td>0.8209</td>
</tr>
<tr>
<td>2-Undecanone</td>
<td>170.3</td>
<td>309-427</td>
<td>14.36</td>
<td>232.68</td>
<td>7.4</td>
<td>2.7</td>
<td>0.06736</td>
<td>3.4147</td>
<td>1.2358</td>
</tr>
<tr>
<td>6-Undecanone</td>
<td>170.3</td>
<td>303-352</td>
<td>14.32</td>
<td>228.68</td>
<td>7.4</td>
<td>2.7</td>
<td>0.0692</td>
<td>2.4177</td>
<td>0.6930</td>
</tr>
<tr>
<td>2-Undecanone</td>
<td>198.35</td>
<td>341-427</td>
<td>14.20</td>
<td>231.03</td>
<td>8.58</td>
<td>2.7</td>
<td>0.0614</td>
<td>2.8127</td>
<td>0.9413</td>
</tr>
</tbody>
</table>

### Table 5.2 Parameters for the Polar PC-SAFT model for ketones.

<table>
<thead>
<tr>
<th>Ketones</th>
<th>MW (g/mol)</th>
<th>Temp (K)</th>
<th>v00 (ml/mol)</th>
<th>u0/k (K)</th>
<th>m</th>
<th>Mu (D)</th>
<th>xp</th>
<th>Avg Abs Dev (%) Vap Pres</th>
<th>Liq den</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetone</td>
<td>58.078</td>
<td>253-463</td>
<td>20.00</td>
<td>259.99</td>
<td>2.22</td>
<td>2.7</td>
<td>0.2258</td>
<td>1.3381</td>
<td>4.7255</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>72.1</td>
<td>270-374</td>
<td>21.38</td>
<td>270.00</td>
<td>2.41</td>
<td>2.7</td>
<td>0.2074</td>
<td>0.4975</td>
<td>0.6468</td>
</tr>
<tr>
<td>2-Pentanone</td>
<td>86.13</td>
<td>282-353</td>
<td>21.75</td>
<td>264.97</td>
<td>2.826</td>
<td>2.7</td>
<td>0.1770</td>
<td>0.0325</td>
<td>0.2352</td>
</tr>
<tr>
<td>3-Pentanone</td>
<td>86.13</td>
<td>285-347</td>
<td>21.67</td>
<td>265.83</td>
<td>2.812</td>
<td>2.7</td>
<td>0.1774</td>
<td>0.1972</td>
<td>0.1365</td>
</tr>
<tr>
<td>2-Hexanone</td>
<td>100.16</td>
<td>299-332</td>
<td>22.11</td>
<td>263.80</td>
<td>3.232</td>
<td>2.7</td>
<td>0.1543</td>
<td>0.0125</td>
<td>0.0246</td>
</tr>
<tr>
<td>3-Hexanone</td>
<td>100.16</td>
<td>299-324</td>
<td>22.15</td>
<td>262.10</td>
<td>3.202</td>
<td>2.7</td>
<td>0.1560</td>
<td>0.0098</td>
<td>0.0356</td>
</tr>
<tr>
<td>2-Heptanone</td>
<td>114.19</td>
<td>306-416</td>
<td>22.12</td>
<td>259.17</td>
<td>3.704</td>
<td>2.7</td>
<td>0.1360</td>
<td>0.6384</td>
<td>0.5560</td>
</tr>
<tr>
<td>4-Heptanone</td>
<td>114.19</td>
<td>275-384</td>
<td>22.37</td>
<td>256.72</td>
<td>3.651</td>
<td>2.7</td>
<td>0.1350</td>
<td>9.7031</td>
<td>0.8636</td>
</tr>
<tr>
<td>2-Octanone</td>
<td>128.22</td>
<td>319-420</td>
<td>22.50</td>
<td>257.07</td>
<td>4.002</td>
<td>2.7</td>
<td>0.1254</td>
<td>12.33</td>
<td>1.5231</td>
</tr>
<tr>
<td>3-Octanone</td>
<td>128.22</td>
<td>316-360</td>
<td>21.93</td>
<td>254.12</td>
<td>4.134</td>
<td>2.7</td>
<td>0.1200</td>
<td>3.6711</td>
<td>0.3274</td>
</tr>
<tr>
<td>2-Nonane</td>
<td>142.24</td>
<td>340-422</td>
<td>22.34</td>
<td>256.81</td>
<td>4.550</td>
<td>2.7</td>
<td>0.1101</td>
<td>0.1546</td>
<td>0.0988</td>
</tr>
<tr>
<td>5-Nonane</td>
<td>142.24</td>
<td>301-356</td>
<td>22.74</td>
<td>256.25</td>
<td>4.448</td>
<td>2.7</td>
<td>0.1125</td>
<td>0.0125</td>
<td>0.0526</td>
</tr>
<tr>
<td>2-Undecanone</td>
<td>170.3</td>
<td>309-427</td>
<td>22.33</td>
<td>253.80</td>
<td>5.441</td>
<td>2.7</td>
<td>0.0919</td>
<td>0.1436</td>
<td>0.0471</td>
</tr>
<tr>
<td>6-Undecanone</td>
<td>170.3</td>
<td>303-352</td>
<td>22.38</td>
<td>251.65</td>
<td>5.399</td>
<td>2.7</td>
<td>0.0924</td>
<td>1.3053</td>
<td>0.1830</td>
</tr>
<tr>
<td>2-Undecanone</td>
<td>198.35</td>
<td>341-427</td>
<td>22.68</td>
<td>253.34</td>
<td>6.250</td>
<td>2.7</td>
<td>0.0799</td>
<td>0.1765</td>
<td>0.0223</td>
</tr>
</tbody>
</table>
alkanones, which gives reason to question the accuracy of the experimental vapor pressure data for 2-octanone.

Figure 5.3a represents the alkanones with the carbonyl group in a position other than on the second carbon. Most of these compounds are what will be referred to as “mid-ketones” meaning that there are an odd number of carbons and the oxygen is on the middle carbon. Note that the slopes of the majority of
Figure 5.2. a) The natural log of the vapor pressure of the 2-ketones as a function of inverse temperature. b) Same as a) with the 2-Octanone results replaced by the predicted results from Polar CK-SAFT. The data is from Smith and Srivastava.\textsuperscript{153}
Figure 5.3. a) The natural log of the vapor pressure of the mid-ketones as a function of inverse temperature. b) Same as a) with the 4-Heptanone results replaced by the predicted results from Polar CK-SAFT. The data is from Smith and Srivastava.153
these curves also decrease with increasing molecular weight of the compounds. Nevertheless, the slopes of 3-octanone and 4-heptanone are slightly out of sync with the other compounds. The slope of the vapor pressure curve for 3-octanone is smaller than that of 3-hexanone and greater than that of 5-nonanone. Although the slope appears to be somewhat out of proportion with the other compounds, it does fit the trend of decreasing magnitude with increasing molecular weight. However, the slope for 4-heptanone is greater than that of 3-hexanone. Figure 5.3b shows that the Polar CK-SAFT prediction for 4-heptanone corresponds with the trend of the other compounds. Numerical comparisons of the slopes of the vapor pressure curves also indicate that there is some discrepancy in the data for heptanone and octanone. For example, the 2- and 3-pentanone slopes differ by 1.8 %, the 2- and 3-hexanone slopes differ by 0.7 %, the 2- and 5-nonanone slopes differ by 3.8%. However, the slopes of 2- and 4-heptanone differ by 8.1% and the slopes of 2- and 3-octanone differ by 12.9%. Thus, there is reason to question the accuracy of the 4-heptanone and 2-octanone experimental data.

5.3 Parametric Analysis

We have fitted the four Polar SAFT parameters to the saturated liquid density and vapor pressure data for the homologous series of ketones, from acetone to tridecanone. The values are cited in tables 5.1 and 5.2 for Polar CK-SAFT and Polar PC-SAFT, respectively. Here, the parameters are examined in regards to their physical representation of molecular characteristics.
Since the change in molecular size, from one member to the next in a homologous series is that of a methylene group, it is expected that the molecular volume is a linear function of molecular weight. In figure 5.4a, the molecular volume, which is given by the product of the number of segments per molecule and the segment volume \(m^*V^0\), is shown as a function of molecular weight for the CK parameter set. The results for the ketones with Polar SAFT are compared with results of the alkanes and alkanols for non-polar SAFT using the parameters fit by Huang and Radosz. Note that the molecular volume of the alkanols and alkanones are very similar as a function of molecular weight. This is expected since the molecular formulae differ only by two hydrogen atoms. For example, 1-pentanol and 2-pentanone have molecular formulae of \(C_5H_{12}O\) and \(C_5H_{10}O\) and molecular weights of 88.15 g/mol and 86.13 g/mol, respectively. Hence we would expect them to have approximately the same molecular volume. Also, as the chain length increases, the effect of the single oxygen atom on the molecular volume is expected to become less and less important and, at sufficiently high molecular weight, the ketone molecular volume should approach that of the alkanes. We find this limiting behavior to occur around a molecular weight of 175 g/mol. In figure 5.4b, the molecular volumes for the ketones based on Polar PC-SAFT are compared with results of the alkanol parameters of Ghosh and Chapman and the alkanes for non-polar PC-SAFT using the parameters fit by Gross and Sadowski. Again the molecular volumes of the alkanols and alkanones agree quite well, although the limiting behavior of approaching the molecular volume of the alkanes is not observed.
Figure 5.4 Molecular volume as a function of molecular weight. a) CK-SAFT results for alkanes, alkanols, and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes and Polar PC-SAFT results for alkanols and ketones.
Since the change in molecular size, both volume and length, from one member to the next in a homologous series is that of a methylene group, it is expected that both the molecular volume and chain length are linear functions of molecular weight. In figures 5.5a,b, the chain length for the alkanes, alkanols, and ketones are shown as a function of molecular weight, for CK- and PC-SAFT, respectively. Indeed, for each method and series considered, the chain length is a linear function of molecular weight. For CK-SAFT (figure 5.5a), the alkanol results approach the alkane results while the ketone results appear to diverge from the alkanes. For PC-SAFT (figure 5.5b), the linear function representing the alkanols has a significantly higher slope than that of either the alkanones or alkanes, while the alkanones and alkanes cross at a molecular weight of approximately 150 g/mol and have similar slopes. To more accurately determine the limiting behavior of the ketones, additional data for longer chains are needed.

The segment dispersion energy, which accounts for the van der Waals type interaction between segments, is shown as a function of molecular weight in figures 5.6a,b for CK- and PC-SAFT, respectively, for the alkanes, alkanols and ketones. Note that in both figures and for each series considered the segment dispersion energies appear to be approaching a limiting value. For CK-SAFT (figure 5.6a), parameters for longer chain alkanes and ketones are needed before a conclusion can be drawn regarding whether or not, and at what molecular weight, the ketones approach the limiting value for the segment dispersion energy of the alkanes. For the PC-SAFT model (figure 5.6b), the ketones appear to reach the limiting value of the alkanes at a molecular weight of
Figure 5.5 Chain length as a function of molecular weight. a) CK-SAFT results for alkanes,\textsuperscript{58} alkanols,\textsuperscript{58} and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes\textsuperscript{61} and Polar PC-SAFT results for alkanols\textsuperscript{54} and ketones.
Figure 5.6 Segment dispersion energy as a function of molecular weight. a) CK-SAFT results for alkanes,\textsuperscript{58} alkanols,\textsuperscript{58} and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes\textsuperscript{61} and Polar PC-SAFT results for alkanols\textsuperscript{154} and ketones.
approximately 200 g/mol.

The polar contribution to the Helmholtz free energy is related to the product of the fraction of polar segments \((x_p)\) and the square of the dipole moment \((\mu^2)\). In order to account for the polar effects, the functional group dipole moment, or in the case of a homologous series with a single polar group, the average molecular dipole moment, is needed. These values can be obtained either from experimental data available in the literature or determined using quantum mechanics. In this study, the average molecular dipole moment of 2.7 Debye is used for the homologous series of the ketones. Further, to allow multiple polar groups in a given molecule requires a parameter representing the fraction of segments in a chain that are polar. It was found that for several of the ketones the number of polar segments per chain \((x_p \cdot m)\) is one-half. Therefore, this constant value was used for this homologous series and held for both the CK and PC dispersion models, as shown in figures 5.7a,b. Others have shown that for the primary alkanols, which have both short-range electrostatic interactions via hydrogen bonding and long-range polar interactions, one polar segment per chain is appropriate.\(^\text{154}\)

Recall that we are considering the dipole moment of the carbonyl functional group rather than the molecule. The actual reported gas phase experimental values for the molecular dipole moment for the series of ketones that are considered here varies from about 2.6 to 2.8 Debye, and an average value of 2.7 Debye is used in this work. For a given ketone, for example, heptanone, we would expect \(x_p\) to be the same regardless of the position of the
Figure 5.7 Number of polar segments per chain as a function of molecular weight for the ketones. a) Polar CK-SAFT. b) Polar PC-SAFT.
carbonyl group. The number of ketone groups is the same, so "logically" $\chi_0$ should be the same. However, the degree to which the polarity affects the solution is affected by the position of the ketone, as is evident from reported experimental dipole moments: 2-heptanone is 2.61D, 3-heptanone is 2.81D, and 4-heptanone is 2.74D.\textsuperscript{123} The mere fact that these differences exist indicates that we should expect some variation in the contributions of the polar component to the overall results.

Both the segment dispersion energy and the polar interactions are attractive forces so that the parameters for these interactions are where the variations are expected. While these differences are small in magnitude, they are evident in the $\chi_0$ parameter as shown in figure 5.7 and in the segment dispersion energy, $U^p/k$, as shown in figure 5.6. This same type of analysis for ethers, esters and other polar, non-associating molecules is needed to gain a clearer understanding of how polar interactions affect solution properties. In order to explicitly include alcohols in this type of analysis, other series of compounds that are polar and self-associating are required.

During the fitting process for the four pure component parameters in this model, it was noted that a broad minimum exists such that a wide range of parameters will accurately represent the experimental pure component data. However, the ability of the pure component parameters to predict the phase behavior of mixtures provides a more rigorous test of the validity of the parameter values, in view of the fact that the mixture results are sensitive to the values of the parameters.
5.4 Parameter Predictions

By developing correlations of the model parameters with molecular weight, parameters for other members of the ketone homologous series can be predicted. Using the correlations for the molecular volume, chain length, and setting the number of polar segments to one-half, the parameters for 8-pentadecanone are predicted. The value of the segment dispersion energy ($u^2/k$) is estimated based on the approaching of a limiting value exhibited by the higher molecular weight ketones. The results for the vapor pressure curve, predicted and experimental, are shown in figure 5.8. The difference in predicted and experimental values are 1.6% for the liquid densities and 7.6% for the vapor pressures. The parameter values are listed in table 5.3.

Furthermore, the segment volume and interaction energy parameters for 2,4-pentadione can be determined based on molecular weight using values for that of 3-hexanone, the ketone of similar molecular weight. The fraction of polar segments and chain length were adjusted to correlate with experimental data. Since we now have two carbonyl groups, we expect the chain length to be larger than that of the five-carbon molecule 3-pentanone but not as long as the six-carbon molecule 3-hexanone. A chain length of the average of these two

<table>
<thead>
<tr>
<th>Table 5.3 Predicted Polar CK-SAFT Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compound</strong></td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>8-Pentadecanone</td>
</tr>
<tr>
<td>2,4-Pentadione</td>
</tr>
</tbody>
</table>
compounds was determined to be appropriate. It was found that a value of 0.69 for the number of polar segments fit the data well. This is a 37% increase over the number of polar segments for 3-pentanone. While we might expect the number of polar segments to double for a diketone, the two carbonyl groups in 2,4-pentadione tend to orient such that the effect of the dipoles are partially canceled by each other as can be seen from the molecular structures shown in figure 5.9. Hence, for this case, the effect is increased but not doubled. The vapor pressure curve for 2,4-pentadione is shown in figure 5.10 and the parameters are listed in table 5.3.
Figure 5.9. Two possible conformations of 2,4-Pentadione. Built and optimized using Molecule 3D.\textsuperscript{36}

Figure 5.10. Vapor pressure curve for 2,4-Pentadione. Comparison of experimental data\textsuperscript{153} and Polar CK-SAFT prediction.
5.5 Effect of explicit inclusion of long-range polar interactions.

In one of the early works of developing parameters for the SAFT model, parameters were fit for acetone, 2-butanone, and 2- and 3-pentanone.\textsuperscript{58} A comparison of these parameters developed for non-polar SAFT with those of Polar SAFT can yield some insight into the benefits of explicitly accounting for the polar interactions. Figures 5.11a,b compare the molecular volume and chain length as a function of molecular weight, respectively. Whereas in most homologous series fit by Huang and Radosz\textsuperscript{58} (e.g. alkanes, alkanols) the molecular volume is a linear function of the molecular weight (shown above in figure 5.4a), that was not the case for the ketones as depicted in figure 5.11a. By explicitly including the polar interaction, the contribution of the polar group to the molecular volume ($m^* \nu^p$) becomes a constant addition rather than the highly varied one. Similarly, the chain length ($m$) (figure 5.11b) becomes a well-behaved function of molecular weight as a result of the explicit accounting for the polar effects.

Moreover, as we noted in chapter 4, in order to account for polar effects using a non-polar model the dispersion interaction energy must be artificially large. Comparison of the segment-segment interaction energies, shown in figure 5.12a, reveals that the polar segment actually has higher segment dispersion interaction energy than the corresponding non-polar segment. However, the difference in the number of segments results in a significant difference in the molecular dispersion interaction energies of the two models for a given polar compound. The molecular dispersion interaction energy, shown in figure 5.12b,
Figure 5.11 a) Molecular volume and b) Chain length as a function of molecular weight. Comparison of CK-SAFT$^5$ and Polar CK-SAFT results for ketones.
Figure 5.12 a) Segment dispersion energy and b) Molecular dispersion energy as a function of molecular weight. Comparison of CK-SAF\textsuperscript{T} and Polar CK-SAF results for ketones.
decreases significantly (45-70%) by explicitly accounting for the polar interactions, thus removing the artificial inflation of the segment energy in the non-polar model.

5.6 Functional Group Parameters

One of the goals of this in-depth study of the SAFT parameters is to establish a basis for the development of a group contribution parameter system for the SAFT equation of state. We know that in reality segments do not actually represent functional groups, for if they did the chain length parameter, \( m \), would be a whole number. And, not all functional groups are "created equal". Each one has a different contribution to the various fluid properties, and consequently the segment volume (\( \nu^0 \)) or segment dispersion energy (\( u^f/k \)) parameters are really reflections of the "averages" of what each segment contributes. Since the change in chain length and molecular volume from one compound to the next in a homologous series represents the addition of a methylene group to the chain, we expect this value to be very nearly constant. One exception is perhaps for the compounds with carbon numbers of three or less where the fluid behavior is known to differ significantly from the higher molecular weight components in a series.

In figure 5.13a,b, the incremental change in chain length (\( m_{c_n} - m_{c_{n-1}} \)) is presented as a function of carbon number. For Polar CK-SAFT, the majority of the ketones show an incremental change around 0.6 with the average value of 0.557±0.077; for the alkanol parameters\(^{58} \) of non-polar CK-SAFT, the value is
Figure 5.13 The change in chain length between two members in a homologous series. a) CK-SAFT results for alkanes,\textsuperscript{58} alkanols,\textsuperscript{58} and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes\textsuperscript{61} and Polar PC-SAFT results for ketones.
0.702±0.051; for the alkane parameters\textsuperscript{58} of non-polar CK-SAFT, the average incremental change is 0.697±0.111. For Polar PC-SAFT, the alkanes yield an average value of 0.354±0.083 and the ketones 0.433±0.079. Since these parameters represent the \textit{average} segment contribution to the molecule, it is expected that the incremental change of the ketones, alcohols, and alkanes would differ from each other.

In figure 5.14a,b, the change in molecular volume (\(v^{00} \cdot m_{c_n} - v^{00} \cdot m_{c_{n-1}}\)) between sequential members of a homologous series is presented as a function of carbon number. For CK-SAFT, the average change in molecular volume for the ketones is 8.30±0.72, 7.68±1.63 for the alkanes, and for the alkanols, 8.35±0.59, where we've consider n=4 and higher. For PC-SAFT, the average values for ketones are 9.95±1.09 and 10.30±0.85 for the alkanes.

For CK-SAFT, we can deduce that the segment volume of a methylene group is this average change in molecular volume in the homologous series of alkanes (7.68±1.63). The difference in average molecular volume for the ketones is larger (by 0.62 ml/mol) than that of the alkanes. However, this does not provide the information needed for the size of a carbonyl segment. Nonetheless, by subtracting the molecular volume of an alkane of carbon number \(n-1\) from that of an alkanone of carbon number \(n\), the result is an estimate of the \textit{segment} volume of a carbonyl group (C=O). Using values from 2-butanone to 2-nonanone, the average segment volume of a carbonyl group is 13.16±1.47 ml/mol. Similarly, using values from 1-butanol to 1-decanol, an estimate for the
Figure 5.14 The change in molecular volume between two members in a homologous series. a) CK-SAFT results for alkanes,\textsuperscript{58} alkanols,\textsuperscript{58} and Polar CK-SAFT results for ketones. b) PC-SAFT results for alkanes\textsuperscript{61} and Polar PC-SAFT results for ketones.
segment volume for an alcohol group \((\text{CH}_2\text{OH})\) is \(12.69\pm0.93\) ml/mol. The range in the series considered in estimating these group values omit the lower molecular weight compounds, since as stated earlier, they tend to deviate from the behavior of the heavier compounds in a series. Furthermore, the upper end of the range considered is truncated at the last sequential compound for which we have parameters. For PC-SAFT, similar calculations yield a value of \(10.30\pm0.85\) for the segment volume of a methylene segment and \(9.76\pm/1.15\) for a carbonyl group. It is interesting that the two forms of SAFT yield different trends for the segment volume of a carbonyl group as compared to a methylene group. CK-SAFT predicts that the segment volume of a carbonyl group is larger than that of a methylene while PC-SAFT predicts the reverse trends. While these results are preliminary, the need for a very thorough and systematic parametric study in order to develop a group contribution based parameter set for SAFT is demonstrated.

Once we can determine accurately the contributions to these parameters on a functional group basis, a group contribution type approach can be developed which would allow us to study a much broader range of molecules and strongly increase the predictive capability of the Polar SAFT model. In order to accomplish this, a thorough analysis of parameter sets for a wide range of linear and branched compounds (e.g. primary alkanols, alkynes, alkenes, carboxylic acids, amines, esters, ethers, amides, etc.) is needed. For the analysis to result in accurate functional group parameters, the series parameters must be determined in a very careful and systematic way.
5.7 Conclusions

In this chapter, results of an in-depth parameter study of Polar SAFT, using both the CK and PC dispersion terms, are presented. Based on the analysis, the reliability of the experimental data for the vapor pressures of 4-heptanone and 2-octanone is determined to be questionable. The Polar SAFT parameters are shown to be well-behaved functions of molecular weight. Moreover, it is showed that the explicit inclusion of polar effects provides a systematic set of parameters for a homologous series of ketones. Furthermore, a methodology for developing a group-contribution approach for the SAFT equation of state is suggested. Preliminary estimates for segment volume for methylene and carbonyl functional groups are proposed for Polar CK-SAFT and Polar PC-SAFT, as well as for the alcohol group for non-polar CK-SAFT.
CHAPTER 6: DIPOLAR SAFT: APPLICATIONS

6.1 Introduction

We now apply the Polar SAFT model to molecules containing single or multiple dipolar segments. Firstly, we consider the phase behavior of acetone/alkane mixtures. This tests the ability of the theory to model polar/non-polar mixtures and to model differences in molecular size of the alkane. Secondly, ketone/alkane mixtures in which the size of either the ketone or alkane can vary are modeled. Thirdly, we calculate the cloud point of poly (ethylene-co-methyl acrylate) in butane, propane, and ethane solvents as a function of the polar co-monomer concentration. This tests the ability of the model to account for multiple dipolar functional groups in a complex system. Thus, modeling these systems provides a strong test of the predictive capabilities of Polar SAFT.

6.2 Phase Equilibria in Acetone-Alkane Binary Systems

To demonstrate the effect of the dipolar term in Polar CK-SAFT, we have performed calculations for the acetone/alkane systems both with and without including the polar contribution. Acetone contains a carbonyl group, which has a dipole moment perpendicular to the molecular axis. The polar model equation of state parameters for acetone were fitted to the saturated liquid density and vapor pressure data for acetone.\textsuperscript{155} Huang and Radosz parameters\textsuperscript{58} are used for the alkane solvents with zero values for the dipole moment and the fraction of polar segments ($x_p$). The solvents we chose are pentane, hexane, decane and
dodecane. The solvent parameters are listed in table 6.1. Both non-polar SAFT and Polar SAFT have an adjustable binary interaction parameter, $k_{ij}$. To test the predictive capability of Polar SAFT, this binary interaction parameter is set to zero.

Table 6.1. CK-SAFT parameters for alkane solvents and ketones.\textsuperscript{58}

<table>
<thead>
<tr>
<th>Component</th>
<th>$\nu^\infty$ (mL/mol)</th>
<th>$\xi/k$ (K)</th>
<th>$m$</th>
<th>$\mu$ (D)</th>
<th>$\chi_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alkanes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethane</td>
<td>14.460</td>
<td>191.44</td>
<td>1.941</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Propane</td>
<td>13.457</td>
<td>193.03</td>
<td>2.696</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Butane</td>
<td>12.599</td>
<td>195.11</td>
<td>3.458</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pentane</td>
<td>12.533</td>
<td>200.02</td>
<td>4.091</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hexane</td>
<td>12.475</td>
<td>202.72</td>
<td>4.724</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Heptane</td>
<td>12.282</td>
<td>204.61</td>
<td>5.391</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Decane</td>
<td>11.723</td>
<td>205.46</td>
<td>7.527</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dodecane</td>
<td>11.864</td>
<td>205.93</td>
<td>8.921</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cyclohexane</td>
<td>13.502</td>
<td>236.41</td>
<td>3.970</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Ketones</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetone</td>
<td>7.765</td>
<td>210.92</td>
<td>4.504</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>11.871</td>
<td>229.99</td>
<td>4.193</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3-Pentanone</td>
<td>10.510</td>
<td>235.24</td>
<td>4.569</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figures 6.1 through 6.4 show the predictions of phase behavior for acetone/alkane mixtures in the form of P-x-y diagrams for both polar and non-polar CK-SAFT. The Polar CK-SAFT results show remarkable agreement with the experimental data for each of these systems, where the phase behavior is dominated by the polar effects. However, CK-SAFT does not give acceptable results at zero $k_{ij}$. In fact, CK-SAFT predicts slight negative deviation from ideal solution behavior for each of these systems, while the experimental data shows strong positive deviation from ideality. This strong positive deviation from ideality
Figure 6.1. Vapor-liquid equilibrium of acetone-pentane system. Polar SAFT (---) and SAFT (-----) at $k_{ij}=0$; Experimental data$^{156}$ (*).

Figure 6.2. Vapor-liquid equilibrium of acetone-hexane system. Polar SAFT (---) and SAFT (-----) at $k_{ij}=0$; Experimental data$^{156}$ (*).
Figure 6.3. Vapor-liquid equilibrium of acetone-decane system. Polar SAFT (—) and SAFT (—) at $k_{ij} = 0$; Experimental data$^{156}$ (○).

Figure 6.4. Vapor-liquid equilibrium of acetone-dodecane system. Polar SAFT (—) and SAFT (—) at $k_{ij} = 0$; Experimental data$^{156}$ (○).
is due to the polar interactions, or more correctly, the lack of polar interactions between the two components. In each plot, two temperatures of the mixture are considered with temperature differences of 40, 10, 20, and 20 Kelvin respectively. No significant model dependence on temperature is observed. Note that in the acetone-pentane system, Polar SAFT shows a region of liquid-liquid equilibrium at the lower temperature considered (258 K) whereas the experimental data at this temperature does not. However, at sufficiently low temperatures, we would expect a region of liquid-liquid immiscibility for this system. Thus, we conclude that Polar SAFT predicts the onset of liquid-liquid equilibrium at a higher temperature for this system than would be observed experimentally.

In the second set of calculations, the binary interaction parameter $k_{12}$ that best fits the P-x-y data at different temperatures is selected. The $k_{12}$ parameter is obtained by minimizing the function

$$\min f(k_{12}) = \sum_{i} \frac{|x_i - x_{i,\text{exp}}|}{\min(x_{i,\text{exp}},1-x_{i,\text{exp}})} + \sum_{i} \frac{|y_i - y_{i,\text{exp}}|}{\min(y_{i,\text{exp}},1-y_{i,\text{exp}})}$$

(6.1)

where $N$ is the number of experimental data points, $x$ is the liquid phase molar composition and $y$ is the vapor phase molar composition of a consistently chosen component at the experimental temperature and pressure for a particular data point. The suffix ‘exp’ indicates the experimental composition. The optimization is minimizing the difference in predicted and experimental values so that any deficiency in the model parameters affect the value determined to be the optimum $k_{12}$. Thus, the $k_{12}$ value compensates, not only for deficiencies in the
theory, but also in the parameters. Since the pure component vapor pressures for the pure alkanes are not reproduced accurately by the Huang and Radosz parameters for the CK-SAFT model, this is of particular concern and will be discussed further in the next section. In certain cases where the data had very close liquid and vapor phase compositions, as in the case of azeotropic mixtures, the best \( k_{12} \) is determined manually by repeating calculations at different values of \( k_{12} \).

The \( k_{12} \) parameter was fitted both for non-polar and Polar CK-SAFT. Table 6.2 lists the optimized \( k_{12} \) values for the systems considered here. Polar CK-SAFT needs consistently smaller values of \( k_{12} \) than SAFT. In the case of acetone-pentane, attempts to adjust the binary interaction parameter in the CK-SAFT model result in phase splitting at the lower temperature (258.15K) similar to that shown in figure 6.1 for Polar CK-SAFT with a \( k_{12} \) of zero. Figures 6.5 and 6.6 show the results for the acetone/hexane and acetone/dodecane binary systems at optimized \( k_{12} \), respectively. Polar CK-SAFT consistently gives better agreement with the experimental results than the non-polar model. For the acetone/hexane system, Polar CK-SAFT gives accurate azeotropic composition and pressure. By adjusting \( k_{12} \) for the acetone/dodecane mixture a more accurate fit is obtained for both models. However, CK-SAFT is unable to reproduce the correct shape of the liquid composition curves. Acetone parameters were also fit for the Polar PC-SAFT model. Results for the acetone/hexane system are shown in figure 6.7, where a \( k_{12} \) of 0.022 is needed to give reasonable agreement with the experimental data. Overall, Polar SAFT predicts the shape of the
Figure 6.5. Vapor-liquid equilibrium of acetone-hexane system. Polar SAFT (—) and SAFT (---) at optimized $k_{ij}$; Experimental data$^{156}$ (*). $k_{ij}$ is correlated to phase composition data.

Figure 6.6. Vapor-liquid equilibrium of acetone-dodecane system. Polar SAFT (—) and SAFT (---) at optimized $k_{ij}$; Experimental data$^{156}$ (*). $k_{ij}$ is correlated to phase composition data.
Figure 6.7. Vapor-liquid equilibrium of acetone-hexane system. Polar PC-SAFT (—) at $k_{ij}$ and at optimized $k_{ij}$; Experimental data\textsuperscript{156} (*).

pressure-composition curves quite well and shows a significant improvement compared to non-polar CK-SAFT for acetone-alkane mixtures.

Table 6.2. Optimized binary interaction parameters.

<table>
<thead>
<tr>
<th>System</th>
<th>CK-SAFT\textsuperscript{152}</th>
<th>Polar CK-SAFT\textsuperscript{152}</th>
<th>Polar PC-SAFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetone/Pentane</td>
<td>0.090</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Acetone/Hexane</td>
<td>0.092</td>
<td>0.004</td>
<td>0.022</td>
</tr>
<tr>
<td>Acetone/Decane</td>
<td>0.075</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Acetone/Dodecane</td>
<td>0.078</td>
<td>0.005</td>
<td></td>
</tr>
</tbody>
</table>

6.3 Phase Equilibria in Other Ketone/Alkane Systems

In the previous section, the effect of the alkane solvent chain length with mixtures of acetone is demonstrated. Now, the effect of molecular size of the ketones is considered using mixtures of ketones in n-heptane and n-hexane. The
temperature effects are shown using the 3-pentanone/heptane systems. Moreover, the effect of having two carbonyl groups on the same molecule is considered in the system of 2,4-pentadione in cyclohexane. To further establish the importance of explicitly including the dipolar term in SAFT, calculations for ketone/alkane systems both with and without the explicit inclusion of the polar contribution are performed. In addition, both the CK and PC approaches to the dispersion term are considered. As in the previous section, Huang and Radosz parameters are used for the alkane solvents, and for the ketones for comparison, with zero values for the dipole moment and the fraction of polar segments \( \langle x_p \rangle \). The solvents used include hexane, heptane, decane and cyclohexane and are listed in table 6.1. Parameters for Polar CK- and PC-SAFT are cited in chapter 5, tables 5.1 and 5.2, respectively. Since each SAFT model has an adjustable binary interaction parameter, \( k_{ij} \), this parameter is initially set to zero in order to test the predictive capability of Polar SAFT. This binary interaction parameter is then adjusted as needed.

In figures 6.8 through 6.14, the predictions of CK-SAFT and Polar CK- and PC-SAFT for the P-x-y diagrams for several ketone/alkane mixtures are compared with experimental data. Overall, the Polar SAFT model is able to capture the physics of these systems, in which the phase behavior is dominated by the polar effects. However, non-polar SAFT does not give acceptable results at zero \( k_{ij} \) and, in most cases, predicts negative deviation from ideal solution behavior. For the ketone-rich phase this is largely due to the fact the CK-SAFT parameters were fit to a model that does not explicitly include the polar
interactions, as we discussed in chapter 5. As in the systems of the previous section, these mixtures show strong positive deviation from ideality due to the lack of polar interactions between the two components. Furthermore, using the parameters of Huang and Radosz, the CK-SAFT model does not accurately reproduce the pure component vapor pressures for several of the alkanes or ketones at the temperatures of this study.

The $k_2$ parameter was determined manually for each SAFT model. A manual approach to fitting the binary interaction parameter was used due to an inherent problem with the ability of the Huang and Radosz parameters for CK-SAFT to accurately reproduce pure component vapor pressures. Optimization of the binary interaction parameter based on equation 6.1 results in less acceptable predictions for both phases while a manual adjustment demonstrates the ability of Polar SAFT to capture the physics of the ketone rich phase, and likewise highlights the integral problems with the Huang and Radosz parameters. Table 6.3 lists the $k_2$ values for each of the models and systems considered here. Polar SAFT needs consistently smaller values of $k_2$ than the non-polar model. By adjusting $k_2$, a more accurate fit is obtained for each model for most systems. Overall, Polar SAFT predicts the shape of the pressure-composition curves quite well, showing qualitative agreement with zero $k_2$ and quantitative agreement with binary interaction parameters typically four to five times less than that required for non-polar SAFT for these systems.

Figure 6.8a,b,c compares the results of the three models for 2-butanone in heptane at 45°C. The non-polar version of CK-SAFT predicts nearly ideal
Table 6.3. Binary interaction parameters for ketone/alkane mixtures.

<table>
<thead>
<tr>
<th>System</th>
<th>Temp (°C)</th>
<th>Polar CK-SAFT</th>
<th>Polar CK-SAFT</th>
<th>Polar PC-SAFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Butanone/Heptane</td>
<td>45.0</td>
<td>0.050</td>
<td>0.009</td>
<td>0.012</td>
</tr>
<tr>
<td>2-Butanone/Decane</td>
<td>70.0</td>
<td>0.05</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td>3-Pentanone/Heptane</td>
<td>40.05</td>
<td>0.035</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>3-Pentanone/Heptane</td>
<td>80.0</td>
<td>0.035</td>
<td>0.006</td>
<td>0.012</td>
</tr>
<tr>
<td>4-Heptanone/Hexane</td>
<td>65.0</td>
<td></td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>5-Nonanone/Hexane</td>
<td>60.0</td>
<td></td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2,4-Pentadione/Cyclohexane</td>
<td>25.0</td>
<td></td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

solution behavior. Note also that the pure component vapor pressure of the mixture data for 2-butane and the CK-SAFT predictions using the Huang and Radosz parameters do not agree. A $k_i$ of five percent gives only somewhat reasonable results for the 2-butane rich phases. In contrast, with a zero binary interaction parameter ($k_{ij}$), both Polar CK- and PC-SAFT predict the strongly non-ideal behavior of the formation of an azeotrope. Nonetheless, without adjusting $k_{ij}$ the agreement is only qualitative. Small $k_{ij}$'s are needed for each polar method ($\sim 1\%$) to obtain very accurate results.

Results for 2-butane in decane at 70°C are shown in figures 6.9a,b,c. CK-SAFT (figure 6.9a) predicts nearly ideal solution behavior and requires a binary interaction parameter of 5% to accurately model the system. Polar SAFT (figures 6.9b,c) predicts the positive deviations from ideality without adjusting $k_i$, and for quantitative agreement with experimental data requires a $k_i$ of 0.012 for both Polar CK- and PC-SAFT.

The difficulty of accurately modeling polar mixtures is well demonstrated by the systems in figure 6.10. Here we have 3-pentanone in heptane at two
Figure 6.8. 2-Butanone in Heptane at 45 C. Comparison of experimental data\textsuperscript{157} and model predictions. a) CK-SAFT, b) Polar CK-SAFT
Figure 6.8. 2-Butanone in Heptane at 45 °C. Comparison of experimental data and model predictions. c) Polar PC-SAFT.

Figure 6.9. 2-Butanone in Decane at 70 °C. Comparison of experimental data and model predictions. a) CK-SAFT.
Figure 6.9. 2-Butanone in Decane at 70 °C. Comparison of experimental data\textsuperscript{156} and model predictions. b) Polar CK-SAFT, c) Polar PC-SAFT.
temperatures showing strong non-ideal behavior. Since the pure component vapor pressures are so similar, they form azeotropic mixtures over a very small pressure range. To show the details of the ability of the Polar SAFT model to capture these significant changes in phase behavior, a closer look at these systems is warranted.

In figures 6.11a,b, and c the results of 3-pentanone in heptane at 40.05°C are shown for $k_{12}$ of zero. Non-Polar CK-SAFT predicts a very slight positive deviation from ideality and does not give the correct pure component vapor pressure for 3-pentanone or heptane (figure 6.11a). Thus, while adjusting $k_2$ will give the correct form of non-ideal behavior, it does not accurately predict the

Figure 6.10. 3-Pentanone in Heptane. Comparison of experimental data$^{158,159}$ and Polar CK-SAFT model predictions at 40.05°C and 80.0°C.
Figure 6.11. 3-Pentanone in Heptane. Comparison of experimental data\textsuperscript{158} and model predictions at 40.05 °C. a) CK-SAFT, b) Polar CK-SAFT.
Figure 6.11. 3-Pentanone in Heptane. Comparison of experimental data and model predictions at 40.05 °C. c) Polar PC-SAFT.

compositions of the 3-pentanone rich phases. Both forms of Polar SAFT predict positive deviations from ideality and the formation of an azeotrope. For Polar CK-SAFT (figure 6.11b) a very small adjustment to $k_2$ (0.003) is needed to obtain excellent agreement with the data. Polar PC-SAFT (figure 6.11c) requires no adjustment of the binary interaction parameter and gives an excellent representation of the heptane rich phase and slightly over-predicts the compositions in the 3-pentanone rich phases.

The P-x-y representation of the 3-pentanone/heptane mixture at 80°C is presented in figures 6.12a,b,c. As in the 40°C case, CK-SAFT predicts very
nearly ideal solution behavior with a value of zero for the binary interaction parameter and does not accurately reproduce the pure component vapor pressures for either component. Using the same value of \( k_2 \) as for the 40°C case, the non-polar model does again predict the correct form of the non-ideal behavior but does not accurately predict the compositions of either phase. Both forms of Polar SAFT again predict positive deviations from ideality and the formation of an azeotrope. For Polar CK-SAFT (figure 6.12b) a small adjustment to \( k_2 \) (0.006) gives the correct azeotropic composition but over-predicts the vapor pressures for the 3-pentanone rich phases. Due at least in part to the poor prediction of the Huang and Radosz parameters for the pure heptane vapor pressures, a value of \( k_2 \) that gives excellent agreement with the experimental data for the heptane rich phases was not determined. Polar PC-SAFT (figure 6.12c), which accurately represented the pure component vapor pressures, needs a one percent correction to the interaction energy to give accurate representation of the 3-pentanone rich phases. However, again, a value of \( k_2 \) that gives excellent agreement with the experimental data for the heptane rich phase was not determined.

A thermodynamic consistency test of the experimental data for the 3-pentanone in heptane at 80°C revealed that there is significant scatter in the reliability of the data, especially in the 3-pentanone concentration ranges of less than 10% and greater than 90%. Of the twenty-three mixture data points, only five are within 5% and twelve within 10%. Typically, deviations of three per cent or less indicate a high degree of consistency and less than ten percent are
Figure 6.12. 3-Pentanone in Heptane. Comparison of experimental data\textsuperscript{159} and model predictions at 80.0 °C. a) CK-SAFT, b) Polar CK-SAFT.
considered acceptable for most systems. Thus, the difficulty of accurately reproducing the experimental data with the polar model is most likely a reflection of the data reliability rather than the predictive capability of the model.

For the mixtures in which the ketone has a carbon number greater than five, there are no comparisons with non-polar SAFT, since the corresponding parameters have not previously been fitted. We did not deem worthwhile fitting additional parameters for polar fluids for the non-polar model.

The Polar CK- and PC-SAFT model predictions and experimental data for 4-heptanone at 65°C and 5-nonanone at 60°C in hexane are depicted in figures 6.13 and 6.14, respectively. Both systems exhibit nearly ideal solution behavior
for the liquid phases, with only slight positive deviations from ideality. While the pure component vapor pressure for hexane is over predicted in both cases by the Polar CK-SAFT model, due to the inherent problem with the Huang and Radosz parameters as was discussed in chapter 5, the model does capture the shape of the curves. It agrees quantitatively with essentially no deviation from the experimental values for the ketone rich phase compositions, and that with no adjustments of the binary interaction parameter. The Polar PC-SAFT model gives excellent agreement with the experimental data for both systems at all compositions and also with no adjustments of the binary interaction parameter. This nearly ideal solution behavior for the 4-heptanone/hexane and 5-nonanone/hexane systems is expected since there is a significant difference in the pure component vapor pressures of the two species. It would be very interesting to investigate these systems at lower temperatures where the pure component vapor pressures are more similar. This would provide information on the degree to which the polar functional group is affecting the phase behavior and whether or not it is approaching some limiting behavior such as dilution of the carbonyl group so that the polar effects are no longer a significant contribution to the phase behavior.

To test the idea of using a functional group dipole moment, application to mixtures in which one component contains more than one carbonyl group is needed. While data for these systems are limited, a comparison of Polar CK-SAFT predictions and experimental data for 2,4-pentadione in cyclohexane at
Figure 6.13. 4-Heptanone in Hexane. Comparison of experimental data\textsuperscript{160} and model predictions at 65.0 °C.

Figure 6.14. 5-Nonanone in Hexane. Comparison of experimental data\textsuperscript{161} and model predictions at 60.0 °C.
25°C is presented in figure 6.15. The polar model captures the physics of the system giving quantitative agreement with the experimental data with no adjustments to the binary interaction parameter. Sufficient data for di- and higher ketones are not available at present for further testing.

![Graph showing 2,4-Pentadione in Cyclohexane](image)

**Figure 6.15.** 2,4-Pentadione in Cyclohexane. Comparison of experimental data\textsuperscript{162} and model predictions at 25.0 °C.

### 6.4 Phase Equilibria in Polar Copolymer-Solvent Systems

Polar copolymer–solvent systems provide a difficult test of our approach by demonstrating the effect of multiple dipolar functional groups on the properties of polymeric fluids. Hasch et al.\textsuperscript{144} showed that SAFT can not predict the phase behavior for polar copolymer solutions and can only be used to correlate the
experimental cloud point data. In order to illustrate the predictive capability of Polar SAFT, we apply the model to some of the systems studied by Hasch et al.\textsuperscript{145}

The polymer considered is poly(ethylene-co-methyl acrylate) (EMA). The goal is to predict the phase behavior of this copolymer in different solvents and for different concentrations of the polar comonomer (methyl acrylate) in the polymer. Hasch et al.\textsuperscript{145} modeled this system by using the same segment volume and chain length parameters as those for polyethylene. They obtained the dispersion energy parameter by fitting the cloud point data of EMA with a suitable solvent. Hence, the dispersion energy parameter became a function of methyl acrylate content of the polymer; and attempts to model the effect of polarity were through adjusting this dispersion energy parameter.

The polymers considered here are EMA\textsubscript{10}, EMA\textsubscript{31}, and EMA\textsubscript{41}, where the subscript indicates the molar percentage of methyl acrylate in the polymer. The molecular characterization of these polymers is given in table 6.4.

\begin{table}
\centering
\caption{Molecular characterization of the polymers.\textsuperscript{145}}
\begin{tabular}{lcc}
\hline
Polymer & Mw & Mole \% MA \\
PE & 20,100 & 0 \\
EMA\textsubscript{10} & 17,000 & 10 \\
EMA\textsubscript{31} & 33,000 & 31 \\
EMA\textsubscript{41} & 29,000 & 41 \\
\hline
\end{tabular}
\end{table}

Our goal is to demonstrate the predictive ability of our theory by minimizing the number of available adjustable parameters. Therefore, the segment volume ($\nu^o$), segment energy ($u/k$), and chain length ($m$) parameters
are set to those of polyethylene. The chain length parameters are determined based on the relationship for PE proposed by Huang and Radosz:\textsuperscript{58} \[ m = 0.05096 \times M_n, \] where \( M_n \) is the number average molecular weight of the polymer. The binary interaction parameter, \( k_{ij} \), is set to the corresponding value for the polyethylene/solvent system, which is based on fitting the experimental data. In addition to these parameters, Polar SAFT needs the dipole moment of the functional group (or monomer) and the fraction of polar segments \( (x_p) \). The experimental value for the dipole moment of methyl acrylate is used as the segment dipole moment.

The parameter \( x_p \) is a function of the segment fraction of methyl acrylate in the polymer. \( x_p \) is adjusted to fit the cloud point data for the EMA\textsubscript{3}--butane system shown in figure 6.16. To predict the phase behavior for other comonomer concentrations, we recognize that \( x_p \) is proportional to the segment fraction of methyl acrylate and we assume that the contribution of a functional group to the segment number is proportional to its molecular weight. Thus, \( x_p \) is given by

\[ x_p = \frac{y_{MA}}{y_{MA} + (M_{\text{ethyl}} / M_{MA})(1 - y_{MA})} \]  \hspace{1cm} (6.2)

where \( x_p \) is the fraction of polar segments for pure methyl acrylate (determined using this equation and \( x_p \) for the EMA\textsubscript{3}--butane system), \( y_{MA} \) is the mole fraction of methyl acrylate in the polymer, \( M_{\text{ethyl}} \) is the molecular weight of the ethyl group and \( M_{MA} \) is the molecular weight of the methyl acrylate group.

With \( x_p \) fitted to the EMA\textsubscript{3}--butane system, the phase behavior of EMA\textsubscript{4}, in butane can be predicted. Figure 6.16 shows the cloud point pressures calculated
by Polar CK-SAFT for EMA$_{41}$ and EMA$_{31}$ in butane as a function of temperature for a solution of 5% by weight of the polymer. The binary interaction parameter is set to that of the polyethylene-butane system ($k_{12} = 0.008$). The Polar SAFT prediction of the cloud point curve for EMA$_{41}$ agrees well with experimental data. The values of the parameter $x_\phi$, calculated using equation 6.2, and the other Polar CK-SAFT parameters for the polymers are listed in table 6.5 and the parameters for the solvents are included in table 6.1. The solvent parameters and the parameters for polyethylene are taken from Huang and Radosz.$^{58}$

![Figure 6.16. Cloud point curves of 5 wt % polyethylene (PE), and EMA copolymers with methyl acrylate mol % of 31 (EMA$_{31}$) and 41 (EMA$_{41}$) in butane. Comparison of experimental data$^{145}$ (•) and model predictions (—).]
Table 6.5. Polar CK-SAFT Parameters for Polymers.\textsuperscript{152}

<table>
<thead>
<tr>
<th>Polymer</th>
<th>( v^\infty ) (mL/mol)</th>
<th>( u/k ) (K)</th>
<th>( m )</th>
<th>( \mu ) (D)</th>
<th>( \chi_p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td>12.0</td>
<td>216.15</td>
<td>1,024.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EMA\textsubscript{10}</td>
<td>12.0</td>
<td>216.15</td>
<td>866.3</td>
<td>1.77</td>
<td>0.067</td>
</tr>
<tr>
<td>EMA\textsubscript{31}</td>
<td>12.0</td>
<td>216.15</td>
<td>1,681.7</td>
<td>1.77</td>
<td>0.15</td>
</tr>
<tr>
<td>EMA\textsubscript{41}</td>
<td>12.0</td>
<td>216.15</td>
<td>1,523.7</td>
<td>1.77</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Using the fraction of polar segments, \( \chi_p \), determined from eq 6.2, the cloud point curves for EMA in propane and ethane can also be predicted (figures 6.17a,b). As in the case of butane, the same segment volume, interaction energy, and chain length parameters for the PE polymer are used. Figure 6.17a shows the cloud point curves of PE, EMA\textsubscript{31} and EMA\textsubscript{10} in propane. For quantitative agreement the \( k_{12} \) value for the EMA\textsubscript{10} in propane system needs to be adjusted to 0.0125 and to 0.01 for the EMA\textsubscript{31} in propane. For the PE/propane system, we determined a binary interaction parameter of 0.0165 was necessary to give the correct trend and quantitative agreement with the experimental data whereas Hasch \textit{et al.}\textsuperscript{145} used a \( k_{12} \) of 0.015 for this system. The results for PE, EMA\textsubscript{31} and EMA\textsubscript{10} in ethane are shown in figure 6.17b. For quantitative agreement the \( k_{12} \) value for the EMA\textsubscript{10} in ethane system needs to be adjusted to 0.024 and to 0.02 for the EMA\textsubscript{31} in ethane. For the PE/ethane system, we determined a binary interaction parameter of 0.027 was necessary to give quantitative agreement with the experimental data whereas Hasch \textit{et al.}\textsuperscript{145} used a \( k_{12} \) of 0.025 for this system. The binary interaction parameter for each system considered is listed in table 6.6. For the non-polar model the binary interaction
Figure 6.17. Cloud point curves of 5 wt % polyethylene (PE), and EMA copolymers with methyl acrylate mol % of 31 (EMA₃₁) and 10 (EMA₁₀) in a) propane and b) ethane. Comparison of experimental data (•) and model predictions (—).
parameter increases relative to the poly(ethylene) results as the mole per cent of
the polar co-monomer, methyl acrylate, is increased. However, for the polar
model which explicitly accounts for the polar interactions in the solution behavior,
the binary interaction parameter values actually decrease as the mole per cent of
methyl acrylate is increased.

In comparison with the SAFT calculations of Hasch et al.,\textsuperscript{145} where they
adjusted the dispersion energy as a function of methyl acrylate co-monomer
content and fitted $k_z$ for different solvents, the Polar SAFT calculations are much
more predictive. In our approach, we took the simplistic view that the segment
volume, dispersion energy, and chain length of EMA are the same as those of
polyethylene. Although in reality these parameters, especially the dispersion
energy, are different for the methyl acrylate group than that of the ethyl group,
our rather simplistic approach gives good results as compared to experimental
data. In the work of Hasch et al.,\textsuperscript{145} increases in both the dispersion energy and
the binary interaction parameters were required to give quantitative agreement
with the experimental data. Here, the additional attractive forces of the polar co-
monomer are accounted for by the explicit inclusion of the polar interactions in
the model. Since the binary interaction parameters in our study are less for the
polar co-polymer systems than the non-polar polymer systems, we may be over
predicting the polar contribution and need to re-evaluate the manner in which we
define $x_p$. Nonetheless, the method does give qualitative agreement with the
experimental data, which is quite impressive for the approximations being made.
The study also emphasizes the importance of dipolar effects in the phase behavior of EMA-alkane solutions.

Table 6.6. Binary Interaction Parameters for Polymers in Solvents.

<table>
<thead>
<tr>
<th>Solvent</th>
<th>PE</th>
<th>EMA_{10}</th>
<th>EMA_{11}</th>
<th>EMA_{n}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Butane</td>
<td>.008</td>
<td>.008</td>
<td>.008</td>
<td></td>
</tr>
<tr>
<td>Propane</td>
<td>.0165</td>
<td>.0125</td>
<td>.0100</td>
<td></td>
</tr>
<tr>
<td>Ethane</td>
<td>.027</td>
<td>.024</td>
<td>.020</td>
<td></td>
</tr>
</tbody>
</table>

6.5 Conclusions

Dipolar interactions have a significant effect on the phase behavior of polar fluids. Predicting the phase behavior of non-spherical molecules with even a single polar group, much less multiple polar groups, has long been an unsolved problem in statistical mechanics based perturbation theory. We have solved this problem, in part, by incorporating the polar term in the SAFT chain equation of state utilizing a segment approach in which the dipoles are assumed to be located on the spherical segments of a chain. This approach is applicable to molecules with multiple dipolar groups; the theory has been validated versus molecular simulation results.\(^{12}\) We have shown that the Polar SAFT equation of state for mixtures accurately predicts the effect of multiple dipolar groups and molecular shape on the phase behavior of mixtures of polar and non-polar components.

We applied Polar SAFT to three different types of systems. For the acetone/alkane mixtures, the Polar SAFT results are in much better agreement
with the experimental data than non-polar CK-SAFT. Even at a $k_2$ of zero, Polar SAFT gives good results for these systems, indicating the predictive capability of the model and the importance of explicitly accounting for the polar interactions directly in the theory rather than attempting to account for it by superficially large interaction energies and binary interaction parameters. In the ketone/alkane systems in which the chain length of the ketone was allowed to vary, Polar CK- and Polar PC-SAFT were able to capture the minute details of the phase behavior. For the alkane solutions of the copolymer EMA (poly (ethylene-co-\methyl acrylate)), Polar SAFT accurately predicts the effect of the polar methyl acrylate co-monomer content and of the solvent on the cloud point behavior. This is possible only because our approach accounts for multiple polar segments. Finally, the accurate small chain and copolymer results demonstrate the predictive capability of Polar SAFT for modeling phase equilibria in a wide range of fluid mixtures from polar monomers to polymers.
 CHAPTER 7: CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

In this work, we considered the solution behavior of polar molecules from amino acids and small peptides to polar solvents and polar co-polymers. By considering the molecular level interactions, we are able to accurately predict phase behavior of a significant range of systems. For the analysis of the interactions of small biological molecules, such as amino acids and dipeptides, in aqueous solution, we utilized structural analysis and thermodynamics. Based on these investigations, we determined the role of the polar functional groups to be of significant importance in the study of biochemicals and proceeded to focus our theoretical work towards the development of an accurate model for polar fluids.

We initiated a systematic experimental study of the solution behavior of amino acids and small peptides, where we consider the effects of temperature, salt type, and salt concentration. By choosing glycine, alanine, and aspartic acid, we considered entities with neutral, hydrophobic, and hydrophilic side chains, respectively. The changes in temperature are shown to have a significantly greater effect on solubility than salt type and salt concentration. Nonetheless, salt type and concentration can be very important variables depending on the specific application. We resolved the contradictory results reported in the literature for studies of glycine solubility in NaCl and KCl aqueous solutions. Based on our extended study with glycine to include LiCl as well as our results for alanine, we believe the trends of our results are more physically reasonable than those
reported by Khoshbarchi and Vera.\textsuperscript{48} The effect of sequence on solution behavior is clearly identified for the systems studied in this work, noting that the placement of a residue with a side group (i.e., any amino acid other than glycine) at the N-terminus significantly reduces the dipeptide solubility. \textit{Ab initio} quantum calculations as well as experimental FTIR and Raman spectroscopy results support our hypothesis that intramolecular hydrogen bonding plays a significant role in the solution thermodynamics of amino acids and dipeptides. We reported values for enthalpic changes from the solid state to the infinitely dilute liquid state for the dipeptides of \textit{asp} and \textit{gly} and demonstrated that the solid state is also an important consideration in evaluating the solubility trends of these substances.

Since biochemical systems contain a significant number of polar groups, forms intramolecular hydrogen bonds, and actively hydrogen bonds with water, we developed a method to very accurately model mixtures of polar and non-polar fluids. Predicting the phase behavior of non-spherical molecules with even a single polar group, much less multiple polar groups, has long been an unsolved problem in statistical mechanics based perturbation theory. We have solved this problem, in part, by developing a polar model utilizing a segment approach in which the dipoles are assumed to be located directly on the spherical segments of a chain. By applying the $u$-expansion to a reference fluid mixture of polar and non-polar spherical segments, this approach is applicable to molecules with multiple dipolar groups. We included the resulting polar model in the SAFT equation of state, which is well-known to accurately model chain and associating systems.
We presented a motivation for the study of polar fluids and a low-density derivation of the SAFT equation of state on which this work is built. Furthermore, we compared our approach to modeling polar fluids to the molecular equivalent sphere approach that has traditionally been applied to account for dipolar interactions. We show that for pure fluids, the molecular approach significantly under-estimates the contribution of the dipolar effects and that the error increases as the chain length increases. For molecules with a single polar segment, the contribution of the polar term to the total Helmholtz free energy differs by a factor of the number of segments and hence the difference between the two approaches is more pronounced as the chain molecule becomes longer. Given that our segment approach has been shown to be in excellent agreement with molecular simulation results,\(^1\)\(^2\) we conclude that the commonly applied molecular sphere approach under-estimates the effect of dipolar interactions for chain like molecules. This molecular sphere approach is limited to the case of a single polar group on a molecule. We extended the pure fluid segment approach, i.e., the dipolar chain equation of state, to mixtures in which any number of the components may contain polar functional groups and allowing for the first time multiple polar groups within any of the components.

We performed an in-depth parameter study of Polar SAFT, using both the CK and PC dispersion terms, and presented the results. Based on the parameter analysis, the reliability of the experimental data for the vapor pressures of 4-heptanone and 2-octanone are determined to be questionable. The Polar SAFT parameters are shown to be well-behaved functions of
molecular weight. Moreover, it is shown that the explicit inclusion of polar effects provides a systematic set of parameters for a homologous series of ketones, from acetone to 2-tridecanone. These correlations based on molecular weight can be used to predict parameters for longer chains within the series and for some diketones. Furthermore, a methodology for developing a group-contribution approach for the SAFT equation of state is suggested. And, preliminary estimates for the segment volume of the methylene \((\text{CH}_2)\) and carbonyl \((\text{C}=\text{O})\) functional groups are proposed for Polar CK-SAFT and Polar PC-SAFT, as well as for the alcohol group \((\text{CH}_3\text{OH})\) for the non-polar CK-SAFT model.

We have shown that the Polar SAFT equation of state for mixtures accurately predicts the effect of multiple dipolar groups and molecular shape on the phase behavior of binary mixtures of polar and non-polar components. We applied Polar SAFT to three different types of systems. For the acetone/alkane mixtures in which the chain length of the alkane is allowed to vary, the Polar SAFT results are in much better agreement with the experimental data than non-polar SAFT. Even with a binary interaction parameter of zero, Polar SAFT gives good results for these systems. This indicates both the predictive capability of the model and the importance of explicitly accounting for the polar interactions directly in the theory, rather than attempting to account for it by superficially large interaction energies and binary interaction parameters. In the ketone/alkane systems in which the chain length of the ketone was allowed to vary, Polar CK- and Polar PC-SAFT were able to capture the minute details of the phase
behavior of several binary systems. For the alkane solutions of the copolymer EMA (poly (ethylene-co-methyl acrylate)), Polar CK-SAFT accurately predicts the effect of the polar methyl acrylate co-monomer content and of the solvent on the cloud point behavior. This is possible only because our approach accounts for multiple polar segments. Finally, the accurate small chain and copolymer results demonstrate the predictive capability of Polar SAFT for modeling phase equilibria in a wide range of fluid mixtures from polar monomers to polymers.

7.2 Future Work

While this work is significant progress towards an accurate model for the solution behavior of components that have both polar functional groups and intra and intermolecular hydrogen bonding, such as biochemicals, much work remains to be done, both experimentally and theoretically. In order to fully understand the effect of the amino acid side chains, experimental and theoretical analysis of tri-peptides and penta-peptides, such as gly-asp-gly, gly-ala-gly, and gly-gly-asp-gly-gly, are needed. As was shown in this study, end effects are a significant determinant in the solution behavior of the dipeptides. By building the chain of glycine, the most neutral of the amino acids, and including the amino acid residue of interest as the center residue of the chain, the end effects of the residue of interest are eliminated. Furthermore, these small chains will allow the effect of the side chain functional groups to be isolated and exhibit little or no folding. Studies of tri-gly and penta-gly would be required as standards on which to base any comparisons. For example, using the tri-peptide gly-asp-gly or the
penta-peptide *gly-gly-asp-gly-gly*, the aspartic acid residue would be in the same form as it would be in a long peptide or protein. This would allow a study of its solubility trends without the inclusion of end effects (position at the N or C terminus) and independent of any interactions with side chains on neighboring residues. Based on the current study, valine and glutamic acid are highly recommended as the next immediate steps to include additional amino acids. Glutamic acid is expected to intramolecularly hydrogen bond but at a slower rate than aspartic acid, since *glu* has an extra methyl group separating the carboxylic acid group from the backbone chain. As such, it is believed that *glu* would behave as a more hydrophilic compound than *asp* under the same conditions. Moreover, similar studies of each of the amino acids are needed to do a thorough and complete analysis of the solubility trends as a function of side group. To extend the experimental study to multiple non-*gly* residues within a chain will require a very systematic variation in the neighboring residues within a small peptide. One objective in considering a single non-*gly* residue within a chain is to totally isolate the effects of that particular functional group or groups. Then, to utilize the knowledge gained from this type of study to be able to predict the thermodynamic behavior of chains, or peptides, in which *multiple* functional groups are present. By developing a suitable database of experimental results, a theoretical approach to modeling biochemical separation processes can be designed and verified.

Quantum mechanical studies provide a great deal of insight into the molecular structure and solution behavior of the small amino acids in an aqueous
solvent. Additional quantum studies of these systems, and in particular those of aqueous solutions of aspartic acid, would be very useful for gaining further understanding of the behavior at neutral pH, the pH at which biological processes normally occur. Moreover, using quantum mechanics to study changes in both structural and molecular properties, such as dipole moments, enthalpy, and entropy, between a single amino acid, or monomer, and that of a dimer or trimer would certainly bring insight into the solution thermodynamics of these systems.

Because the pure component vapor pressures of the 4-heptanone/hexane and 5-nonanone/hexane systems are more similar at lower temperatures, it would be very interesting to investigate the phase behavior and predictive ability of the model at temperatures lower than considered in the present study. The nearly ideal solution behavior seen in these systems is expected since there is a significant difference in the pure component vapor pressures of the species at the temperatures considered. However, by considering these same systems at conditions more favorable to the existence of non-ideal solution behavior, information regarding the degree to which the polar functional groups of these longer-chain, or higher molecular weight, ketones affects the phase behavior can be gained. Furthermore, whether or not the mixture is approaching some limiting behavior, such as what chain length is necessary such that the polar functional group is sufficiently diluted to no longer be a significant contribution to the phase behavior, could also be evaluated. This would have implications also in verification of the parameter analysis by noting at what molecular weight the ketones do indeed behave as alkanes.
For the polar co-polymer systems, we noted that the binary interaction parameters in our study are less for the polar co-polymer systems than the non-polar polymer systems. Since in non-polar models the binary interaction parameter is increased with increasing concentration of the polar monomer to obtain reasonable agreement, we may be over predicting the polar contribution and need to re-evaluate the manner in which we define \( x_p \). We presently believe that using the PEMA/propane system to determine \( x_p^* \) will result in a more accurate representation of all of the systems and reduce, or possibly eliminate, any need to adjust the binary interaction parameter for these systems. Moreover, an additional test of the predictive ability of the model would be to consider systems of non-polar polymers in polar solvents as well as polar co-polymers in polar solvents.

It is highly desirable to develop a group contribution approach to the Polar SAFT equation of state because it would allow us to study a much broader range of molecules and strongly increase the predictive capability of the model. Once we can determine accurately the contributions to the model parameters on a functional group basis, a group contribution type approach can be developed. In order to accomplish this, a thorough analysis of parameter sets for a wide range of linear and branched homologous series of compounds (e.g. primary alkanols, alkynes, alkenes, carboxylic acids, amines, esters, ethers, amides, etc.) is needed. For the analysis to result in consistent and accurate functional group parameters, the series parameters must be determined in a very careful and systematic way.
In order to extend the current polar SAFT model to aqueous solutions of amino acids or peptides, the model needs to be thoroughly tested with less complex systems that involve both polarity and hydrogen bonding, such as alcohols and carboxylic acids. While such a study for alcohols has begun,\textsuperscript{154} the parameters are not as well-behaved functions of molecular weight as other homologous series. Further, no comparison has yet been made with non-polar SAFT, which accounts only for the association interactions. These comparisons are needed to aid in the understanding of the relative importance of the two effects. A similar test of this relative importance would be to compare the contributions of the polar interactions and association interactions to the thermodynamic properties. However, since the magnitude of these contributions is strongly influenced by the model parameters, it would not necessarily be a conclusive test. Moreover, water is a difficult compound to model and a better representation of the thermodynamic properties of water by the model, i.e., a more accurate parameter set, is needed prior to delving into an in-depth study of aqueous solutions. Once the water parameters are well-established, studies of aqueous solutions of amines, carboxylic acids, and amides would be advisable prior to modeling aqueous solution of amino acids and small peptides.
REFERENCES


APPENDIX A: AMINO ACIDS

Amino Acids with Non-Polar Side Chains

Glycine

\[
\begin{align*}
\text{Gly} & \quad \text{G} \\
\text{H} & \quad \text{C} \quad \text{H} \\
& \quad \text{NH}_3^+ \\
\end{align*}
\]

Alanine

\[
\begin{align*}
\text{Ala} & \quad \text{A} \\
\text{H} & \quad \text{C} \quad \text{CH}_3 \\
& \quad \text{NH}_3^+ \\
\end{align*}
\]

Valine

\[
\begin{align*}
\text{Val} & \quad \text{V} \\
\text{H} & \quad \text{C} \quad \text{CH} \\
& \quad \text{NH}_3^+ \quad \text{CH}_3 \\
\end{align*}
\]

Leucine

\[
\begin{align*}
\text{Leu} & \quad \text{L} \\
\text{H} & \quad \text{C} \quad \text{CH}_2 \quad \text{CH} \\
& \quad \text{NH}_3^+ \quad \text{CH}_3 \\
\end{align*}
\]

Isoleucine

\[
\begin{align*}
\text{Ile} & \quad \text{I} \\
\text{H} & \quad \text{C} \quad \text{C}^* \quad \text{CH}_2 \quad \text{CH}_3 \\
& \quad \text{NH}_3^+ \quad \text{H} \\
\end{align*}
\]
Methionine  Met  M  
\[
\text{H} \quad \text{C} \quad \text{CH}_2\text{CH}_2\text{S}\text{CH}_3
\]
\[
\text{NH}_3^+
\]

Proline  Pro  P  
\[
\text{H}_2\quad \text{C}\quad \text{CH}_2\quad \text{COO}^-
\]
\[
\text{N}\quad \text{H}_2\quad \text{CH}_2\quad \text{H}
\]

Phenylalanine  Phe  F  
\[
\text{H} \quad \text{C} \quad \text{CH}_2\quad \text{NH}_3^+
\]

Tryptophan  Trp  W  
\[
\text{H} \quad \text{C} \quad \text{CH}_2\quad \text{NH}_3^+
\]

**Amino Acids with uncharged polar side chains**

Serine  Ser  S  
\[
\text{H} \quad \text{C} \quad \text{CH}_2\text{OH}
\]
\[
\text{NH}_3^+
\]
Amino Acids with charged polar side chains

Lysine

\[
\text{Lys (K)} \quad \text{COO}^- \quad \text{NH}_3^+ \quad \text{NH}_2^+ \\
H \quad \text{C} \quad \text{CH}_2 \quad \text{CH}_2 \quad \text{CH}_2 \quad \text{CH}_2 \quad \text{NH}_3^+ \\
\text{NH}_3^+ 
\]

Arginine

\[
\text{Arg (R)} \quad \text{COO}^- \quad \text{NH}_2^+ \\
H \quad \text{C} \quad \text{CH}_2 \quad \text{CH}_2 \quad \text{CH}_2 \quad \text{CH}_2 \quad \text{NH}_2^+ \\
\text{NH}_3^+ 
\]

Histidine

\[
\text{His (H)} \quad \text{COO}^- \quad \text{NH}_3^+ \\
H \quad \text{C} \quad \text{CH}_2 \quad \text{CH}_2 \quad \text{C} \quad \text{NH}_3^+ 
\]

Aspartic Acid

\[
\text{Asp (D)} \quad \text{COO}^- \quad \text{O} \\
H \quad \text{C} \quad \text{CH}_2 \quad \text{C} \quad \text{NH}_3^+ \quad \text{O} 
\]

Glutamic Acid

\[
\text{Glu (E)} \quad \text{COO}^- \quad \text{O} \\
H \quad \text{C} \quad \text{CH}_2 \quad \text{CH}_2 \quad \text{C} \quad \text{NH}_3^+ \quad \text{O} 
\]
APPENDIX B: SOLUBILITY OF A SOLID AT INFINITE DILUTION

The phase behavior of binary systems in the region where one component is infinitely dilute is described by Henry’s Law. The Henry’s law constant, $K_i$, is defined as

$$\lim_{x_i \to 0} \left( \frac{f_i}{x_i} \right) = K_i.$$  \hfill (A1)

We also know that the fugacity of component $i$ in terms of the Gibbs energy is given by

$$\ln(f_i) = \frac{G_i}{RT} + \ln(P).$$ \hfill (A2)

so that at constant pressure,

$$\left( \frac{\partial \ln(f_i)}{\partial T} \right)_p = \left( \frac{\partial \left( \frac{G_i}{RT} \right)}{\partial T} \right)_p = -\frac{H^R}{RT^2}.$$ \hfill (A3)

For the pure solid phase and the infinitely dilute liquid phase to be in equilibrium, the fugacity of the amino acid in each phase must be equal: $f_i^s = f_i^l$. Applying equation A3 to both phases, noting that both phases are residual to the same reference fluid, and subtracting the results for the solid phase from the liquid, we have

$$\left( \frac{\partial \ln(K/f_i^s)}{\partial T} \right)_{p,x} = -\frac{\Delta H^{s \to \infty}}{RT^2}.$$ \hfill (A4)

where $f_i^s$ is the fugacity of the pure solid, $T$ is the solution temperature in Kelvin, $R$ is the gas law constant, $\Delta H$ is the change in enthalpy, with the superscript $s \to \infty$.
signifying that the enthalpy change is for the solid state going to a state of infinite
dilution, and the subscripts $P,x$ indicating that the relation is valid for constant
pressure and composition.

Integration of equation A4 from the melting temperature of the solute to
the solution temperature, assuming that the change in enthalpy, $\Delta H^{s\rightarrow s}$, is
independent of temperature over the temperature range of interest, we obtain:

$$\ln(x,)=\frac{\Delta H^{s\rightarrow s}}{R}\left(\frac{1}{T}-\frac{1}{T_{m,}}\right)-\ln\left[\frac{f^s(T_{m,},P)}{K(T_{m,},P)}\right]$$  \hspace{1cm} (A5)

where $T$ is the solution temperature, $x$ is the mole fraction of the solute, $T_{m,}$ is the
melting temperature of the solute, and the last term is the natural log of the ratio
of the fugacity of the solute as a solid at its melting temperature and system
pressure to the Henry’s law constant of the solute at the same state conditions.

For an ideal solution, the fugacity and Henry’s constant at the solute
melting temperature and same pressure are equal and the last term of equation
A5 is zero. For this limiting behavior, the solute melting temperature is given by:

$$T_{m,}=\left(\frac{1}{T}-\ln(x)^{P'/-\cdot\Delta H^{s\rightarrow s}}\right)^{-1}$$  \hspace{1cm} (A6)