

Modelling the Solubility of the 20 Proteinogenic Amino Acids with Experimentally Derived Saturation Data

Nathan A. Bowden

Thesis committee

Promotor

Prof. Dr. J.P.M. Sanders

Emeritus chair of Biobased Chemistry

Wageningen University & Research

Co-promotor

Dr. ir. M.E. Bruins

Senior Researcher at Wageningen Food & Biobased Research

Wageningen University & Research

Other members

Dr.ir. M.L.F. Giuseppin, Avebe

Prof. Dr. Ir. J van der Gucht, Wageningen University & Research

Dr. ir. A.E.M. Janssen, Wageningen University & Research

Dr. A. Plijter-Schuddemat, Cereal & Food Expertise

This research was conducted under the auspices of the Graduate School Food Technology
Agrobiotechnology Nutrition and Health Science

Modelling the Solubility of the 20 Proteinogenic Amino Acids with Experimentally Derived Saturation Data

Nathan A. Bowden

Thesis

submitted in fulfilment of the requirements for the degree of doctor at Wageningen University by the authority of the Rector Magnificus, Prof. Dr. A.P.J. Mol, In the presence of the Thesis Committee appointed by the Academic Board to be defended in public on Wednesday, 16 May, 2018 at 4 P.M. in the Aula

Nathan A. Bowden

Modelling the Solubility of the 20 Proteinogenic Amino Acids with Experimentally Derived Saturation Data, 261 pages.

Ph.D. thesis, Wageningen University, Wageningen, the Netherlands (2018)

With references, with summary in English

ISBN 978-94-6343-861-2

DOI 10.18174/446739

Contents

Chapter 1 Introduction	1
Chapter 2 Solid-liquid phase equilibria and thermodynamic properties of mixtures of 20 α -amino acids in a model system of industrial residues	19
Chapter 3 Solubility of the proteinogenic α -amino acids in water, ethanol and ethanol-water mixtures ...	49
Chapter 4 The effects of ethanol on the solubility of the proteinogenic α -amino acids in mixtures	79
Chapter 5 Modelling the effects of ethanol on the solubility of the proteinogenic amino acids with the NRTL, Gude and Jouyban-Acree models	129
Chapter 6 Molecular-thermodynamic Correlation of Solubility Data for 20 Amino Acids in Water, in Ethanol and in Water-Ethanol Mixtures	165
Chapter 7 Fractional precipitation of amino acids from agro-industrial residues using ethanol	197
Chapter 8 General Discussion.....	223
Summary	253

Chapter 1

Introduction: Amino Acids in the Bio-based Economy

1.1 Introduction to the Bio-based Economy

Until recently, the human species depended solely on the use of biological and renewal materials for food, fuel and along with mineral resources all the products that were used in our daily lives. The shift to non-renewable resources for energy and products was made about 250 years ago with coal during the industrial revolution [1]. Today, coal, along with other non-renewable resources serve as the basis of our energy and many of the products that we use (e.g. plastic).

While the use of petroleum has many benefits, there are a few side effects that make their wide-use unattractive. First, the use of non-renewable fuels contributes to the enhancement of the natural greenhouse effect that we have on Earth. This leads to raising global temperatures and negative impacts on human lives, global ecosystems and the planet [2]. Second, as the non-renewable reservoirs that are most easily accessible are used up, we must go to extreme locations to get to new reservoirs. Extracting petroleum from these extreme locations has led in the last few years to several global-scale disasters, the incidence of which will only increase if we continue. Third, the extraction of non-renewable fuels is itself a contaminating process for ecosystems and techniques such as hydraulic fracturing (also known as “fracking”) has been shown to contaminate both surface and groundwater reservoirs [3]. Fourth, due to the geographic imbalance in the location of the petroleum reserves, there have been political and economic developments that many believe have contributed to tensions between human settlements [4]. Fifth, fluctuating oil prices will slow down investments [5].

An alternative to the petroleum-based economy could be to initiate a paradigm shift back to a bio-economy. In the bio-economy, many of the products and energy that we rely on would come from biological material.

1.1.1 Bio-based value pyramid

The use of biological materials can be illustrated using the value pyramid [6, 7]. Shown in Figure 1.1, the bio-based value pyramid shows that the value of the biological material that is used is determined by its application. At the bottom of the pyramid, the application has its lowest value but also demands the largest amount of biological material to produce it. Also at this level, the most fossil fuel is replaced, having the greatest impact on greenhouse emission reduction. As

the applications move up the pyramid, the higher the value of the application and also the less amount of biological material is used.

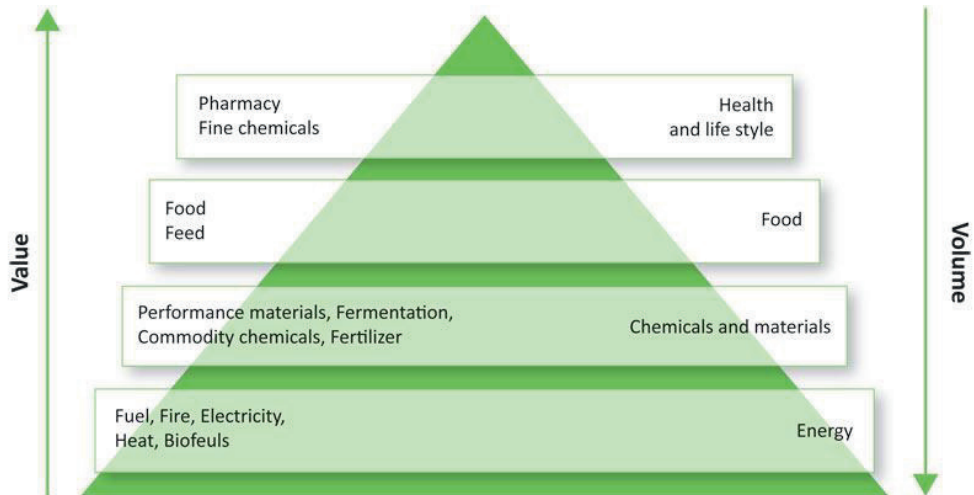


Figure 1.1: The bio-based value pyramid. At the top, the products are the most valuable and less biomass is used. At the bottom the products are less valuable and more biomass is used (8)

At the bottom of the pyramid are the applications of fuel, fire, electricity, heat and biofuels. Here, any biological material can be used. Processes to transform the biological material into energy resources include pyrolysis, gasification among others.

The step second from the bottom of the bio-based value pyramid are performance materials, fermentation, commodity chemicals and fertilizer. Less biological material is needed for this and the value of the products created are worth more. However, the impact in replacing non-renewable resources is also lower. Examples of biological materials that can be used for in these processes include cellulosic, hemi-cellulosic, other carbohydrates and protein-rich sources. Processes that are useful here include but are not limited to steam-explosion and fermentation technologies.

Applications on the step second from the top of the bio-based value pyramid include products such as food for humans and feed for animals. Similar material and processes can be used here, as in the step below it in the bio-based value pyramid. Present food chains are far from efficient and this helps to close the waste and residues associated with this.

The top position in the bio-based value pyramid is for pharmacy and fine chemicals. Here the values of the products are the highest and the amount of biological material that is used is the least. However, only certain biological molecules can be used in these applications. These include, peptides, proteins, amino acids, their derivatives and other secondary metabolites.

1.1.2 The potential for amino acids in bio-based products

Amino acids are a potential feedstock for the production of nitrogen containing pharmaceuticals and fine chemicals. An example of this is the chemical 1,2-ethanediamine that is used as a corrosion inhibitor, an anti-scaling agent and in lubricants [8]. 1,2-Ethanediamine is made by adding ammonia to ethanolamine. Currently, ethanolamine is produced by the oxidative conversion of ethylene to ethylene oxide. Then, the ethylene oxide is then converted to ethanolamine with the addition of ammonia. However, another route would be to convert an amino acid, such as L-serine, to ethanolamine directly [9]. This route would require one less step in the production of 1,2-ethanediamine and is energetically more efficient, as shown in Figure 1.2.

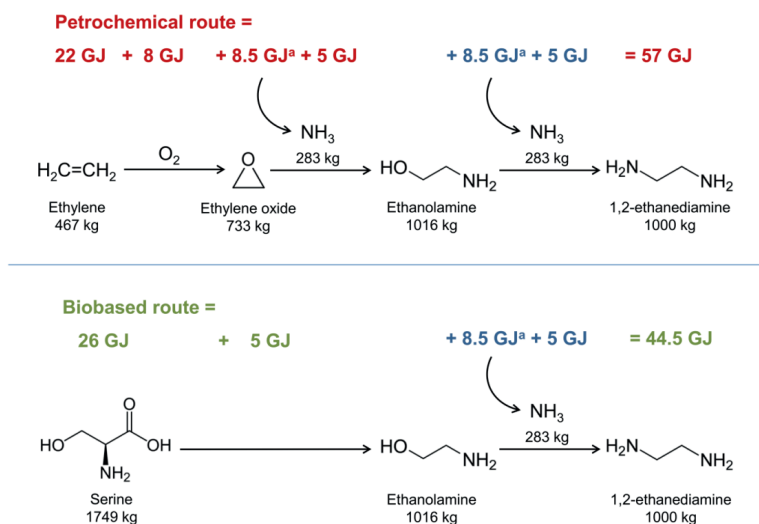


Figure 1.2: Production of 1,2-ethanediamine [9-11]. On the top is the petrochemical route. On the bottom is the bio-based route. The petrochemical route takes more energy to produce (57 GJ) than the bio-based route (44.5 GJ).

However, the efficiency of this example is illustrated in the caloric value of the feedstocks ethylene and L-serine as well as the production processes. Yet, we must take into account the energy needed to produce the feedstock (ethylene or L-serine). Ethylene is produced by steam cracking of naphtha and requires 20.4 GJ/ton [12]. L-Serine is currently produced through fermentation using *Methylobacterium* sp [13]. The production of similar amino acids, L-glutamic acid and L-lysine, through fermentation show that a fermentative production of L-serine would consume energy around 24 GJ/tonne [14]. Since 1.7 tons of L-serine are needed for the production of 1.0 tonne of 1,2-ethanediamine but only 0.5 tons of ethylene, the production of 1,2-ethylamine is energetically more efficient using naphthalene. The use of L-serine as a feedstock would only be useful if it took less than 15GJ/ton to produce.

Therefore, it is important to find an energetically more efficient method to produce an amino acids in order to use to produce bio-based products. One solution could be to use agricultural residues that contain high amounts of amino acids and proteins as part of a circular economy.

1.1.3 Sourcing amino acids in a circular economy and the food vs. fuel debate

One method of sourcing proteins and amino acids is to produce them through agricultural. Examples of current agricultural protein production includes casein from milk and soy protein from soy plants. If we were to switch production of products to biological materials, such as proteins and amino acids, then there is some concern that we would either have to use more land for agriculture or that land that is currently being used for food production will have to be switched to agricultural practices that produce high amounts of proteins and amino acids. This is, in essence, the food vs. fuel debate.

While it has been shown that the food vs. fuel debate could be solved by more efficient use of agricultural land, a reduction in food waste and a more efficient food logistics network [15], this would not be a concern if the source of proteins and amino acids were to come from agricultural waste. This approach is part of the circular economy. A circular economy is among other things an economy where waste is reused as a feedstock for another product.

In this approach, land can be used for the production of food (e.g. corn), but the waste (stalks, leaves, roots, stems, *etc.*) are used for the production of energy and other products [16].

Waste, or residue as it is called in this research, does not exist in a circular economy, it is simply another resource that is used for another process.

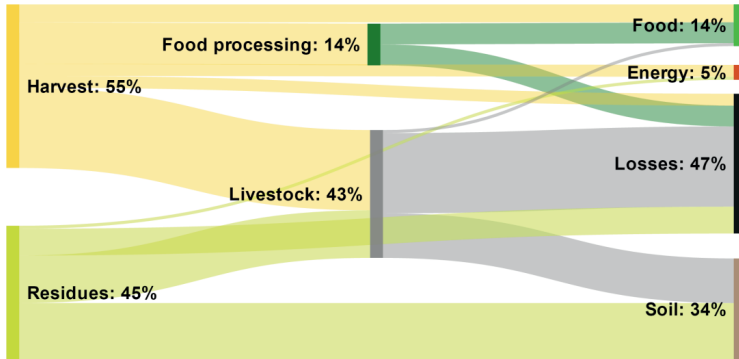


Figure 1.3 Distribution of harvest and residues [11,17] based on carbon-weight. On the left is the original agricultural products as carbon-source, on the right is the eventual product at the end of the product life-cycle. The majority of the carbon ends either as losses or left in the soil.

As seen in Figure 1.3, of all land-based agricultural production, only 55% of it, based on carbon-weight, is harvested and 45% is residue. Of the harvest and residue, approximately half of each goes to feed livestock. From the livestock, the majority of the mass, or 47% of the original carbon source, is lost. These losses are also often high-protein sources such as hair, bone, and feathers. Another large portion of agricultural production, 34%, is left on the soil. The recovery of just a portion of the losses from food processing, the harvest given to feedstock and/or the residue that is lost or left on the soil, would amount to a feedstock capable of supporting a small-scale bio-refinery [11].

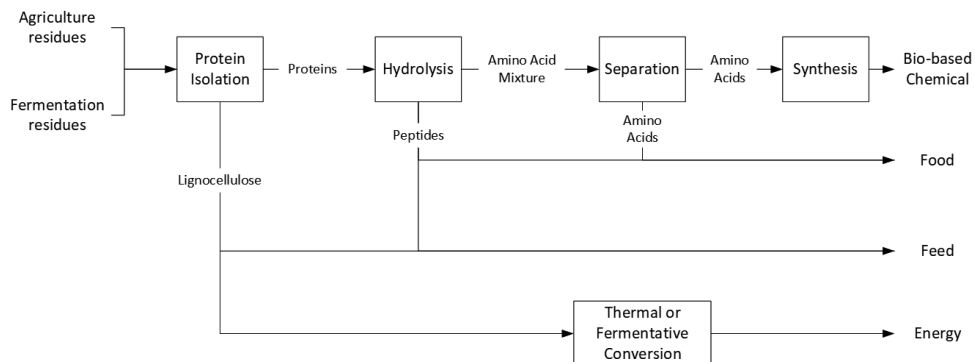


Figure 1.4 Conceptual model of a bio-refinery to produce bio-based energy and products. On the left, is the source of the biomass. On the right is the eventual product after the individual steps of the bio-refinery. In the bio-refinery, many products can be made in parallel adding more value than a single product (e.g. energy).

Agricultural residue is also a low-cost feedstock for a refinery process. It would make an ideal candidate for recovering amino acids, such as L-serine in the example of the previous section. In this way, we can imagine a process to recover amino acids, peptides and proteins from agricultural residue. As shown in Figure 1.4, the agricultural residues and fermentation residues would first go through a form of protein isolation. Then the proteins could either be given as feed to livestock or the proteins could be hydrolyzed. During hydrolysis, the proteins are divided into amino acids and peptides. The peptides could be used either for food for humans or feed for livestock. Conversely, the amino acids could be separated into individual amino acids, which could be used as food/supplements or converted into bio-based chemicals, such as L-serine 1,2-ethanediamine.

1.1.4 An introduction to amino acids

20 amino acids are coded by DNA to form proteins. All α -amino acids contain an amine group and a carboxyl group attached to a single carbon. To this carbon, called the α -carbon, a side chain is attached in between the amine and carboxyl groups. The exceptions to this are glycine, which has no side chain, and L-proline, which has a side chain that forms a 4-carbon ring (including the α -carbon) with the nitrogen atom of the amine group [18,19].

The side chains of these 20 amino acids all have different physical properties [19]. The physical properties include the form (e.g. cyclic or non-cyclic) and the atomic composition of the side chain. The form of the amino acids can be cyclic or non-cyclic, branched, non-branched or form rings. Some amino acids have a combination of these properties. The atomic composition also varies between the amino acids. Some amino acid side chains contain only carbon and hydrogen (defined here as aliphatic), while other amino acids contain oxygen, Sulphur, nitrogen or a combination of these. An inventory of the physical and atomic composition of the 20 α -amino acids is shown in Table 1.1.

Table 1.1: Physical properties of the side chains of the 20 proteinogenic α -amino acids [18]. In the first column, the amino acids are shown. In the subsequent columns, an 'x' denotes the presence of a functional group of that quality. In the far right column, the length of the side chain in carbons is given. These physical properties are used to group the amino acids.

Amino Acid	Ringed	Phenylic	Branched	Aliphatic	Hydroxylic	Oxygenic	Sulphurous	Nitrogenous	Number of Carbons
ARG								x	4
LYS								x	4
GLN						x		x	3
ASN						x		x	2
GLU						x			3
ILE			x	x					4
ASP						x			2
LEU			x	x					4
MET							x		3
THR					x	x			2
ALA				x					1
CYS							x		1
HIS	x							x	4
PHE	x	x		x					7
SER					x	x			1
TRP	x	x						x	9
TYR	x	x			x	x			7
VAL			x	x					3
GLY									0
PRO	x								3

The cyclic amino acids that do not contain a phenyl ring are L-histidine and L-proline. The cyclic amino acids that are cyclic and also contain a phenyl ring are L-phenylalanine, L-tryptophan and L-tyrosine. The aliphatic amino acids that are also branched include L-isoleucine, L-leucine and l-valine. The unbranched aliphatic amino acids are L-alanine, L-glycine and L-

proline. The amino acids containing a hydroxyl group are L-threonine, L-serine and L-tyrosine. The oxygen containing amino acids are L-glutamine, L-asparagine, L-glutamic acid and L-aspartic acid. The amino acids that contain sulphur are L-methionine and L-cysteine. The amino acids that contain amides or amines are L-arginine, L-lysine, L-asparagine and L-glutamine. A visual representation of the α -amino acids characterized by the physical properties of their side chains in a Venn Diagram is shown in Figure 1.5.

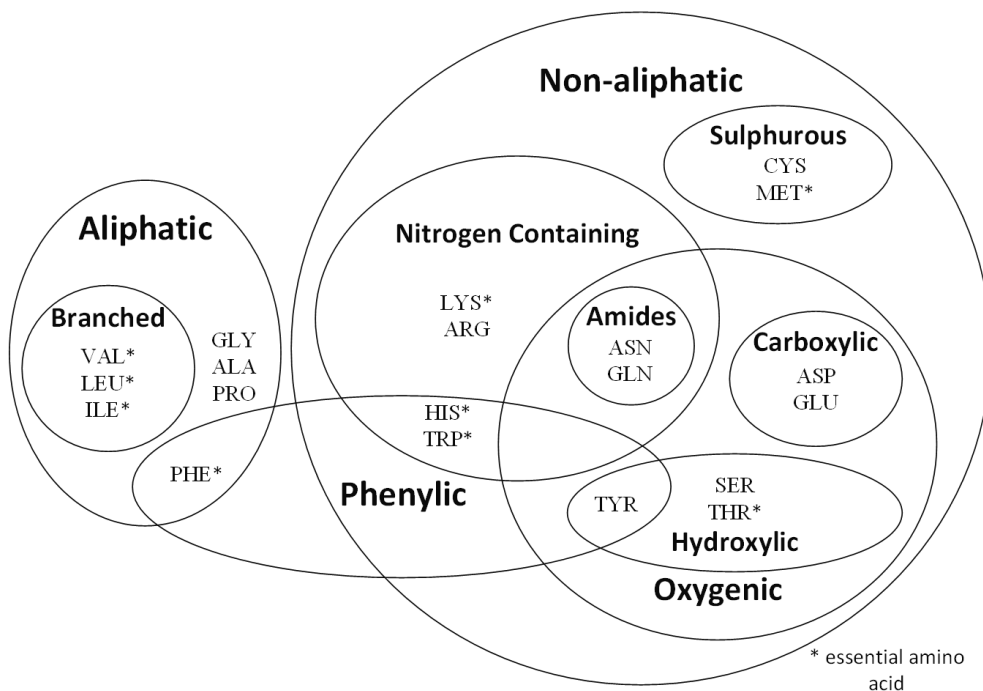


Figure 1.5: Proteinogenic amino acids grouped by their physical properties. In this Venn diagram, the qualities of the functional groups from Table 1.1 are grouped. This Venn diagram is used to summarize the conclusions from the experimental work in other chapters in this thesis.

1.1.5 Amino acid solubility

Amino acids can be extracted from agricultural residues [20,21]. Examples of such residues include green leafy biomass (e.g., grass, leaves, tea leaves), algae, cyanobacteria, press cake from plant-based oil and residue from juice production (e.g., seeds, grapes, peels). These residues often contain proteins and amino acids that can be recovered. When amino acids are

extracted from residue streams using mechanical, chemical or thermochemical processes, they are in solution. In order to design a process to separate the amino acids from a liquid, their solubility must be understood.

Solubility of amino acids is dependent on the temperature of the solution [22, 23]. Co-solvents may also be added to a solution to change the solubility of the amino acids [24]. In Figure 1.6 and Figure 1.7 general models of the effects of temperature and a co-solvent are shown. As shown in Figure 1.6, the solubility of the amino acid increases as temperature increases [25]. The amino acid can be crystallized by lowering the temperature to where the concentration is above the saturation point. As shown in Figure 1.7, in general, the solubility of an amino acid in water is lowered as the concentration of an organic co-solvent increases. The amino acid will crystallize when the concentration of the co-solvent in the solution is increased to the point where the solubility of the amino acid is below the concentration of the amino acid in the solution [25].

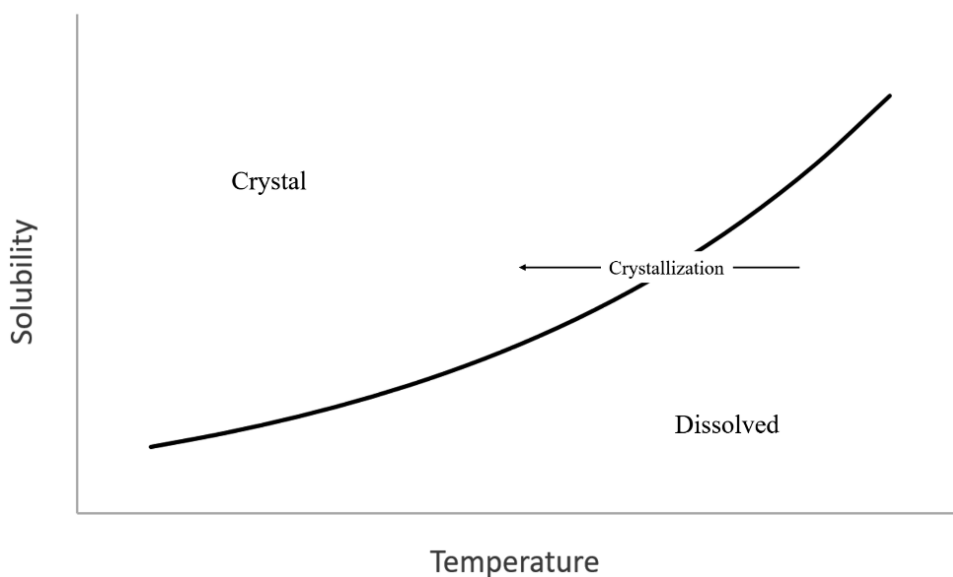


Figure 1.6: A general model of the effect of temperature on the solubility of an amino acid. As the temperature increases, the solubility of the amino acid increases. Crystallization occurs when the temperature of the system is decreased past the point where the concentration of the solute in the solvent is higher than the maximum solubility of that solute in the solvent at that temperature.

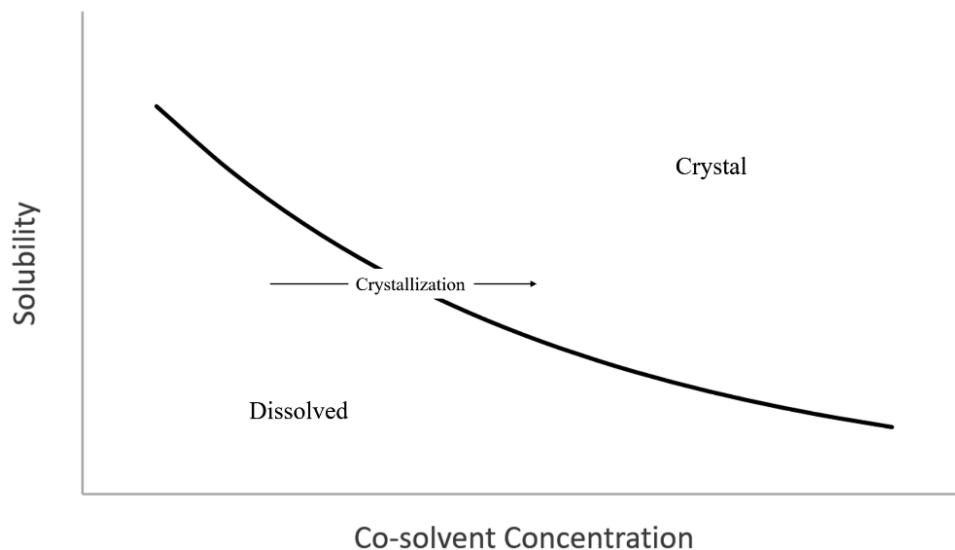


Figure 1.7: A general model of the effect of an organic co-solvent on the solubility of an amino acid in water. As the concentration of the organic co-solvent increases, the solubility decreases. Crystallization occurs when the concentration of the organic co-solvent increases to the point where the maximum solubility of the solute in the mixed two-solvent system is lower than the solubility of the solute in the system.

1.2 Aim

The aim of this research is to be able to understand some of the physical and thermodynamic properties of α -amino acids and their intermolecular interactions in order to separate α -amino acids that have been extracted from agricultural residue. Once the α -amino acids have been extracted, they can be processed for products (pharmaceuticals, fine chemicals, food/feed, *etc.*). In order to do this, three approaches have been used. These are:

1. Measure the impact of temperature and co-solvents on the solubility of α -amino acids in a model mixture of all 20 α -amino acids and individual α -amino acids.
2. Apply existing solubility models and propose a new solubility model to the solubility data on α -amino acid solubility in both the literature and measured for this thesis from approach 1.
3. Develop a process for extracting and isolating proteins and amino acids from a few agricultural residues

1.3 Approaches

1.3.1 Single amino acids versus mixtures of amino acids

Amino acids in agricultural residues and hydrolyzed proteins, are not in single solution [20]. They are found in mixtures of all 20 amino acids. In order to understand how mixtures of amino acids in agricultural residues interact, we work with a model mixture of 20 amino acids. In this model mixture, all 20 amino acids are present, 11 of which at solubility. The other 9 amino acids are represented in the mixture at 3.0 g / 100 g of solvent. In this way, every amino acid is present, but no gel is formed. By doing this, and measuring the solubility of these amino acids in a model mixture, we can study how temperature and ethanol, our chosen anti-solvent, effect the solubility of the amino acids. We thereafter can study the influence that the interaction between amino acids have on their solubility in a mixture similar to agricultural residues by comparing them to the solubilities of the 20 amino acids in single mixture.

1.3.2 Ethanol as anti-solvent

The intermolecular forces between the α -amino acids and to organic solvents are not fully understood. Furthermore, the effects of organic solvents on the solubility of α -amino acids is incomplete [26]. In order to apply models on the effects of organic solvents to single amino acids in aqueous mixture and mixtures of amino acids in aqueous mixture, we must first have a complete data set. In this research, the solubility of the amino acids both in single solution and in mixtures is measured by bringing the solutions to saturation at varying temperatures and pH levels. The solubility is then measured using the gravimetric method and/or the UPLC method [27]. These data are then compared and when applicable added to the literature data available on amino acid solubility.

1.3.3 Models

Solubility models can be split into three categories. The categories are group contribution models, activity coefficient models and empirical models [18]. These models can then either be predictive or regressed. In this research we use three models. The models that we use are the Non-Random Two Liquid (NRTL) model, the Jouyban-Acree model, the Gude model and our own model.

The NRTL and Gude models are regressed activity coefficient models [28, 29]. The NRTL regresses two parameters and the Gude model regresses 1 parameter. The difference between the two is that the NRTL model not only regresses a parameter for the amino acid solute but also the interaction between the molecules.

The Jouyban-Acree model is a predictive empirical model [30]. The model is based on the estimation of Hansen parameters and previously regressed parameters that together predict the solubility of a solute in solution.

The model that we have developed in this research is a predictive model that uses both activity coefficients as well group contribution. The group contribution is using the Marrero and Gani technique [31] to estimate certain thermodynamic parameters such as melting temperature among others. Then, the activity coefficients are calculated using the van Laar equation in the binary [32]. Thereafter, the van Laar equation in the ternary is used to predict the solubility of the solute.

1.3.4 Protein extraction and hydrolysis

Extraction of proteins and amino acids from residues and their subsequent hydrolysis can take place under different methods. These methods can be done apart or in combination. They are physical methods (e.g. milling), chemical/physiochemical methods and biological methods. These are methods that can be used in the first step of a bio-refinery as conceptually proposed in Figure 1.4.

Physical methods can be grouped between grinding, irradiation and other [33]. Under grinding, there are several types of milling technologies such as ball milling, hammer milling, colloid milling and vibro-energy milling among others. Irradiating can be done with gamma rays, electron-beam and microwave. Other physical techniques include hydrothermal, high pressure steam, expansion and pyrolysis among others.

There are several chemical/physiochemical methods that can be employed [34]. Alkaline extraction techniques often use sodium hydroxide and ammonia. Commonly used acids include sulfuric acid and hydrochloric acids. Oxidators like hydrogen peroxide and oxygen have been used. Extraction of lignin with solvents include ethanol-water, benzene-water, ethylene glycol, butanol-water among many others. Organo-ionic solvents have also been used in the past.

Biological methods include fungi and enzymes [35]. Enzymes that are used range from generic cellulases to more specific enzymes such as lignin peroxidase, manganese peroxidase and laccase. Each of these is used to separate one type of biological material such as cellulose using cellulase.

The extraction techniques that are employed in this research include various methods. Of the physical methods, we have used grinding, hydrothermal and vibro-energy milling. Of the chemical/physiochemical methods, both alkaline and acid extraction has been used as well as ethanol-water extraction and freeze-drying. No biological methods were used in this research.

1.4 Chapter Outline

In chapter 2, the solubility and thermodynamic properties of α -amino acids in a model system of industrial residues are discussed. Then, in chapter 3, the solubilities of the single proteinogenic α -amino acids in water, ethanol-water mixtures and ethanol are measured and compared. In chapter 4, the effects of ethanol on the solubility and thermodynamic properties of α -amino acids in a model system of industrial residues is shown. In chapter 5, the effects of ethanol on the solubility of the proteinogenic amino acids is modelled using the NRTL, Gude and Jouyban-Acree models. In chapter 6, we develop a predictive model based on the Marrero and Gani group contribution technique and the van Laar solubility equation. In Chapter 7, the extraction of amino acids from several agriculture residue streams is discussed as well as using ethanol as an anti-solvent for amino acid separation. Lastly, in chapter 8, a general summary and conclusions of all of the previous chapters is presented as well as new research lines such as binary interactions of amino acids, a proposal to apply our model to model mixtures and applications of the protein and amino acid extraction protocol to another biomass.

References

- [1] Smil, V. World history and energy. In *Encyclopedia of Energy*; Cleveland, C. J., Ed.; Elsevier: New York, **2004**; Vol. 6, pp 549–561.
- [2] International Energy Agency. *World Energy Outlook 2012*; **2012**.
- [3] Jackson, R. B.; Vengosh, A.; Carey, J. W.; Davies, R. J.; Darrah, T. H.; O’Sullivan, F.; Pétron, G. The environmental costs and benefits of fracking. *Annu. Rev. Environ. Resour.* **2014**, 39 (1), 327–362.
- [4] Thienpont, J. R.; Kokelj, S. V.; Korosi, J. B.; Cheng, E. S.; Desjardins, C.; Kimpe, L. E.; Blais, J. M.; Pisaric, M. F.; Smol, J. P. Exploratory hydrocarbon drilling impacts to Arctic lake ecosystems. *PLoS ONE* **2013**, 8 (11), e78875.
- [5] Cheaper oil: many winners, a few bad losers *The Economist* October 25th 2014
- [6] Werkgroep Businessplan Bioeconomy (WBBE). Een Punt Op de Horizon: Aan Zet voor Een Intersectoraal Businessplan Bioeconomy. *Rijksoverheid Nederland* **2011**
- [7] Langeveld, J. W. A.; Sanders, J. P. M. General introduction. In *The Biobased Economy: Biofuels, Materials and Chemicals in the Post-Oil Era*; Langeveld, J. W. A., Meeusen, M., Sanders, J. P. M., Eds.; Earthscan: London, UK, **2010**; pp 3–17.
- [8] National Center for Biotechnology Information. PubChem Compound Database; CID=3301, Ethylenediamine <http://pubchem.ncbi.nlm.nih.gov/compound/3301> (accessed Jan 2018).
- [9] Scott, E.; Peter, F.; Sanders, J. P. M. Biomass in the manufacture of industrial products—the use of proteins and amino acids. *Appl. Microbiol. Biotechnol.* **2007**, 75 (4), 751–762
- [10] Rabou, L.; Deurwaarder, E. P.; Elbersen, H. W.; Scott, E. L. *Biomass in the Dutch Energy Infrastructure in 2030*; Platform Groene Grondstoffen, **2006**.
- [11] Widayarani; Biorefinery of Proteins from Rubber Plantation Residues Wageningen, **2015**
- [12] Patel, M. Cumulative energy demand (CED) and cumulative CO₂ emissions for products of the organic chemical industry. *Energy* **2003**, 28 (7), 721–740.
- [13] Ivanov, K.; Stoimenova, A.; Obreshkova, D.; Saso, L. Biotechnology in the production of pharmaceutical industry ingredients: Amino acids. *Biotechnol. Biotechnol. Equip.* **2013**, 27 (2), 3620–3626. Alexandratos, N.; Bruinsma, J. *World Agriculture Towards 2030/2050: The 2012 Revision*; ESA Working Paper 12-03; FAO: Rome, **2012**.

- [14] Ajinomoto. Ajinomoto Group Sustainability Report 2014-Additional Documents and Data http://www.ajinomoto.com/en/activity/csr/pdf/2014/additonal_document_en.pdf (accessed Jan 2018).
- [15] Tilman, D.; Balzer, C.; Hill, J.; Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci.* **2011**, *108* (50), 20260–20264.
- [16] Bajželj, B.; Richards, K. S.; Allwood, J. M.; Smith, P.; Dennis, J. S.; Curmi, E.; Gilligan, C. A. Importance of food-demand management for climate mitigation. *Nat. Clim. Change* **2014**, *4* (10), 924–929.
- [17] Bajželj, B.; Richards, K. S.; Allwood, J. M.; Smith, P.; Dennis, J. S.; Curmi, E.; Gilligan, C. A. Importance of food-demand management for climate mitigation. *Nat. Clim. Change* **2014**, *4* (10), 924–929.
- [18] Bowden, N.A.; Se villano, D.M.; Sanders, J.P.M.; Bruins, M.E., Modelling the effects of ethanol on the solubility of the proteinogenic amino acids with the NRTL, Gide and Jouyban-Acree models. Accepted for publication *Fluid Phase Equilib.* **2018**
- [19] Creighton, T. H.; "Chapter 1". *Proteins: structures and molecular properties*. San Francisco: W. H. Freeman. **1993**
- [20] T.M. Lammens, M.C.R. Franssen, E.L. Scott, J.P.M. Sanders (2012) "Availability of protein-derived amino acids as feedstock for the production of bio-based chemicals" *Biomass and bioenergy*, **44**, 168-181
- [21] Widyarani; Bowden, N. A.; Sanders, J. P. M.; Kolf schoten, R. C.; Bruins, M. E., Fractional precipitation of amino acids from agro-industrial residues using ethanol. *Ind. Eng. Chem. Res.* **2016**, *55*, 7462-7472.
- [22] Dalton, J.B.; Schmidt, C.L.A. The solubilities of certain amino acids in water, the densities of their solutions at twenty-five degrees, and the calculated heats of solution and partial molal volumes *J. Biol. Chem.* **1933**, *103*, 549–578
- [23] Dunn, M.S.; Ross, F.J.; Read, L.S. The solubility of the amino acids in water *J. Biol. Chem.* **1933**, *103*, 579–595
- [24] Cohn, E.J.; McMeekin, T.J.; Edsall, J.T.; Weare, J.H. Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. II. The Solubility of α -Amino Acids in Water and in Alcohol—Water Mixtures *J. Am. Chem. Soc.* **1934**, *56*, 2270–2282

- [25] Eggeling L, Pfefferle W, Sahm H **2006** Chapter 14 Amino acids. 3rd edition, *Basic Biotechnology*, ed. C. Ratledge and B. Kristiansen, Cambridge University Press, Page 337-338
- [26] Bowden, NA; Sanders, JPM and Bruins, ME Solubility of the Proteinogenic α -Amino Acids in Water, Ethanol, and Ethanol–Water Mixtures *J. Chem. Eng. Data* **2018** DOI: 10.1021/acs.jced.7b00486
- [27] Meussen B, van Zeeland AT, Bruins M, Sanders J.P.M. A Fast and Accurate UPLC Method for Analysis of Proteinogenic Amino Acids. *Food Anal. Methods* **2014**, 7:1047-55
- [28] Renon H., Prausnitz J. M., Local Compositions in Thermodynamic Excess Functions for Liquid Mixtures, *AIChE J.*, **1968** 14 (1), S.135–144
- [29] Gude, M. T.; van der Wielen, L. A. M.; Luyben, K. C. A. M., Phase behavior of α -amino acids in multicomponent aqueous alkanol solutions. *Fluid Phase Equilibria* **1996**, 116 (1–2), 110-117.
- [30] Jouyban-Gharamaleki, A.; Valaee, L.; Barzegar-Jalali, M.; Clark, B. J.; Acree Jr, W. E., Comparison of various cosolvency models for calculating solute solubility in water–cosolvent mixtures. *International Journal of Pharmaceutics* **1999**, 177 (1), 93-101.
- [31] Marrero, J.; Gani, R.; Group-contribution based estimation of pure component properties *Fluid Phase Equilibria* **2001**, 183-208
- [32] Prausnitz, J. M., Lichtenthaler, R. N.; Gomes de Azevedo, E., *Molecular Thermodynamics of Fluid-Phase Equilibria* 3rd edition, **1999**
- [33] Demirbas, A., *Biorefineries for Biomass Upgrading Facilities*, Springer, 2010
- [34] Zhang, C., Sanders, J.P.M., Bruins, M.E., Critical parameters in cost-effective alkaline extraction for high protein yield from leaves. *Biomass Bioenergy* **2014** 67, 466-472
- [35] Widyaningrum; Ratnaningsih, E.; Sanders, J. P. M.; Bruins, M. E. Biorefinery methods for separation of protein and oil fractions from rubber seed kernel. *Ind. Crops Prod.* **2014**, 62, 323–332

Chapter 2

Solid-liquid phase equilibria and thermodynamic properties of mixtures of 20 α -amino acids in a model system of industrial residues

In preparation as: Bowden, N.A., Sanders, J.P.M., Bruins, M.E., Solid-liquid phase equilibria and thermodynamic properties of mixtures of 20 α -amino acids in a model system of industrial residues

Abstract:

Crystallizing individual α -amino acids from a mixture of α -amino acids would improve their separation in a bio-refinery, but the solubility of the 20 α -amino acids in a mixture of 20 α -amino acids is unknown. In order to understand the solubility of α -amino acids in mixtures of 20 α -amino acids, we first modeled the solubility of the 20 α -amino acids in single solution using the van 't Hoff and Sober equations and data available in the literature. The results of these equations were compared and their coefficients are reported here. Then we measured the solubility of the 20 α -amino acids in a model system of an industrial residue containing all 20 α -amino acids and found that only L-tyrosine, L-tryptophan, L-aspartic acid, L-glutamic acid and L-asparagine dissolved in higher concentrations in the model system of 20 α -amino acids than in aqueous solution by themselves. The solubility of all of the aliphatic α -amino acids was lower in the mixture. The longer the side chain length of the aliphatic α -amino acids, the greater the decrease in solubility in the mixture as compared to being in solution by itself. The maximum solubility of L-lysine and L-glutamine showed the most decrease when in the mixture in comparison to in aqueous solution alone. Our results demonstrate that α -amino acids with similar physical structures have similar changes in solubility when in a mixture of α -amino acids. We assert that α -amino acids can be grouped by their physical structure.

2.1 Introduction

Bio-based products made from proteins and α -amino acids, such as bioplastics, pharmaceuticals, and fine chemicals, could become increasingly important if an effective way of separating them from solution is found [1-3]. One of the challenges is to find a way to separate proteins and α -amino acids that are dissolved in liquid industrial residues or produced through fermentation [4]. Understanding the solubility of α -amino acids in mixtures is integral to designing technologies for solid-liquid separation of individual α -amino acids from mixtures of α -amino acids in these residue streams.

Research on the solubility of amino acids has proceeded over several decades. The initial work on the solubility of individual amino acids focused on their solubility in water [5-7]. This work was thereafter expanded to include ethanol [8-10] and other solvents [11-16] and additives [17-22]. In the subsequent decades, more amino acids were discovered and their solubilities were compared and modelled by their activity coefficients [23].

However, the experiments and models in the literature are focused on single or pairs of α -amino acids in solution. The challenge is to be able to measure, model and predict the solubilities of amino acids in mixtures of all amino acids, like those found in industrial residues and through fermentation. The results of our experiments can be used in designing protocols and developing technologies for the separation of amino acids in mixtures of amino acids.

2.2 Creating a model industrial residue

In order to have broad scientific impact and yet be able to apply this research to industry, a model industrial residue was created. The model residue is not based on any one industrial residue (e.g. sugarbeet vinasse), because in industrial residues, only one amino acid can be brought to maximum solubility. If a real industrial residue had been chosen, then it would not be applicable to other residues [2]. Furthermore, most amino acids found in industrial residues are α -amino acids. For this reason, a model residue of 20 α -amino acids was designed that had as many of the 20 α -amino acids at maximum solubility as possible.

Initial research showed that attempting to bring all 20 α -amino acids to maximum solubility produced a gel. Further research showed that adding 3.0 grams of each of the 20 α -

amino acids to 100 g of water, no gel was formed. Furthermore, there were 9 α -amino acids at maximum solubility at this concentration.

The 9 α -amino acids that were brought to maximum solubility in the model residue are applicable to food, feed and pharmaceutical industries. These 9 α -amino acids are L-tyrosine, L-tryptophan, L-phenylalanine, L-methionine, L-valine, L-isoleucine, L-leucine, L-lysine and L-glutamine. The other 11 α -amino acids (L-arginine, L-threonine, L-aspartic acid, L-glutamic acid, L-asparagine, L-histidine, L-cysteine, L-proline, L-serine, L-alanine and glycine) were in the model industrial residue at 3.0 g/100 g H₂O, which is below their maximum solubility.

2.3 Solubility of single α -amino acids

The solubility data of α -amino acids have not all been measured at the same temperatures [24]. The information needed to calculate the solubilities of the α -amino acids at any temperature are incomplete [25]. Therefore, in this article, we calculate the solubility coefficients using the van 't Hoff equation and Sober equation in order to be able to calculate the solubilities of the α -amino acids at any temperature. The literature data that we used can be found in Table A.3

The van 't Hoff equation [26] has been used widely in the literature [27-28] and has been applied to other models of solubility [29]. The van 't Hoff equation may be expressed at constant pressure in the exact form [28]

$$d \ln \alpha_i^s / dT = \Delta H / (R * \{T/K\}^2), \quad (2.1)$$

where α_i^s is the unitless activity at saturation of the solute, i , at equilibrium. ΔH represents the partial molar enthalpy of the solute in solution. Furthermore,

$$\alpha_i^s = (\gamma_i^s / \text{Kg H}_2\text{O} \cdot \text{g amino acid}^{-1}) * (x_i^s / \text{g amino acid} \cdot \text{Kg H}_2\text{O}^{-1}) \quad (2.2)$$

where x_i^s is the solubility of the solute in solution. The activity coefficient of solute i is given by γ_i^s . This is usually applied to the van 't Hoff equation in its approximate form, becoming [28]

$$d \ln x_i^s / dT = \Delta H' / (R * \{T/K\}^2), \quad (2.3)$$

where $\Delta H'$ is the apparent partial molar enthalpy of solution. In this case, $\Delta H'$ does not equal ΔH , since $\Delta H' = \Delta H(\partial \ln x_i^s / \partial \ln \gamma_i^s)_T$. We now assume that $\Delta H'$ is a linear function of temperature, as

$$\Delta H' = a + bT/K \quad (2.4)$$

in which a and b are constants. Therefore, a is considered the hypothetical value of $\Delta H'$ at absolute zero and b is the heat capacity of the solute at constant pressure. Coefficient b in the application to this work is, as shown in Table A.1, negligible. As shown in earlier work [28], leaving out b , combining Equations (2.3) and (2.4) and integrating becomes

$$\ln x_i^s = - \frac{(a'/J \cdot \text{g amino acid} \cdot \text{mol}^{-1} \cdot \text{Kg H}_2\text{O}^{-1})}{(R \cdot T/K)} + c/\text{g amino acid} \cdot \text{Kg H}_2\text{O}^{-1} \quad (2.5)$$

Where coefficients a' and c are constants.

Another equation that is used in standard reference texts for amino acids is the Sober equation [24, 30]. This solubility equation is

$$\ln x_i^s = \delta/\text{g amino acid} \cdot \text{Kg H}_2\text{O}^{-1} + (\theta/\text{g amino acid} \cdot \text{Kg H}_2\text{O}^{-1} \cdot (T/K)^{-1}) * T/K, \quad (2.6)$$

where the coefficients δ and θ are constant for each amino acid.

The solubility constants a'/R , c , δ and θ were estimated by minimizing the normalized root mean square variance (NRMSV) of the solubility value in Equation (5) and Equation (6). The minimized NRMSV for all calculations are reported in Table A.1. The NRMSV was calculated using Equation (2.7), where \hat{y}_T is the predicted solubility, y_T is the solubility reported in the literature and y_{max} and y_{min} are the maximum and minimum solubility reported in the literature in the temperature range from 273.15 K to 373.15 K.

$$NRMSV = \sqrt{\frac{\sum_{i=T}^n (\hat{y}_T/K/\text{g amino acid}\cdot\text{Kg } H_2O^{-1} - y_T/K/\text{g amino acid}\cdot\text{Kg } H_2O^{-1})^2}{\sum_{i=T}^n (\hat{y}_T/\text{g amino acid}\cdot\text{Kg } H_2O^{-1} - y_T/\text{g amino acid}\cdot\text{Kg } H_2O^{-1})^2}} \quad (2.7)$$

The results of the estimation of coefficients a'/R , c , δ and θ are shown in Table 2.2.

2.4 Solubility of multiple α -amino acids

Comparing the solubilities of the α -amino acids is a challenge due to the magnitudes in difference in solubility between α -amino acids. The solubilities of α -amino acids at 278.15 K range from 0.011 g of L-cysteine \cdot 100 g H_2O^{-1} , to 162.3 g of L-proline \cdot 100 g H_2O^{-1} .

To aid comparison and since solubility is measured at equilibrium, the standard free energy of transfer, $\Delta G^{\circ}_{tr}/J\cdot\text{mol}^{-1}$ [14] can be calculated. This is the relative solubility between α -amino acids in mixtures compared to their solubility in single solution. This has been calculated using Equation (2.8). In Equation (2.8), the solubility x_i^s was converted to m_i , the molar concentration of the amino acid at maximum solubility in single solution, which was calculated using Equation (2.5) and plotted in Fig. 1.

$$\Delta G^{\circ}_{tr} = R/J\cdot\text{mol}^{-1}\cdot K^{-1} * T/K \ln \frac{m_m/\text{moles amino acid}\cdot\text{Kg } H_2O^{-1}}{m_i/\text{moles amino acid}\cdot\text{Kg } H_2O^{-1}} \quad (2.8)$$

In Equation (2.8), m_m is the maximum solubility in molar concentration of each individual amino acid in a mixture of 20 α -amino acids. The solubility of the amino acids was measured through experimentation and shown in Table 2.1.

2.5 Method

The α -amino acids that were used in this article were purchased from Sigma-Aldrich. These α -amino acids were at least 97% pure (mole fraction). Table 2.1 lists the supplier and purity of the chemicals used in this work.

3.0 Grams of each amino acid was put into 100 ml of demineralized water in a 250 ml glass Schott bottle. This was done again for a duplicate. The sealed Schott bottles were mixed and added to a jacketed water bath on a stir plate and set to 277.15 K and 0.1 MPa and left to mix continuously at 80 rpm. Both the samples and the water in the water bath were continuously

monitored with temperature probes. The pH of all solutions was measured to be 5.3. This stable pH is consistent with the buffering nature of amino acids.

The amino acids were allowed to mix until the maximum solubility of the amino acids had reached equilibrium. Equilibrium was determined as follows. Duplicate samples were taken daily, filtered and the total dry weight was determined. The amino acids were said to have reached equilibrium when successive dry weights, 24-hours apart, yielded a concentration within the variation of the balance. All measurements were performed in duplicate.

Table 2.1: Description of chemicals and solvents used. All chemicals were ordered from Sigma-Aldrich with purities of at least 0.97 mole fraction.

Chemical Name	Source	Mole Fraction Purity	Purification Method
glycine	Sigma-Aldrich	1.00	None
L-alanine	Sigma-Aldrich	0.98	None
L-arginine	Sigma-Aldrich	0.98	None
L-asparagine	Sigma-Aldrich	1.00	None
L-aspartic acid	Sigma-Aldrich	0.99	None
L-cysteine	Sigma-Aldrich	0.97	None
L-glutamic acid	Sigma-Aldrich	0.99	None
L-glutamine	Sigma-Aldrich	0.99	None
L-histidine	Sigma-Aldrich	0.99	None
L-leucine	Sigma-Aldrich	1.00	None
L-lysine	Sigma-Aldrich	0.97	None
L-methionine	Sigma-Aldrich	0.98	None
L-phenylalanine	Sigma-Aldrich	0.98	None
L-proline	Sigma-Aldrich	0.99	None
L-tryptophan	Sigma-Aldrich	0.98	None
L-tyrosine	Sigma-Aldrich	0.98	None

In parallel, liquid samples were filtered and diluted for amino acid profile analysis. Concentrations of each amino acid after dilution ranged from 0.0100 to 2.00 mM and with a total amino acid concentration of not more than 10.00 mM per sample. The samples were analysed in

duplicate by the UPLC method [43]. The UPLC method is based on automated pre-column derivatisation in the injection needle of the amino acids in an autosampler using o-phthalaldehyde (OPA) reagent in combination with 9-fluorenylmethyl chloroformate (FMOC) that enables the amino acids to fluoresce. Separation was achieved with a Dionex RSLC system using an Acquity UPLC BEH C18 reversed-phase column. Sample analysis was performed with an UltiMate 3000 Rapid Separation pump and autosampler. Derivatised amino acids were detected at 263 nm (FMOC derivative of L-proline) and 338 nm (OPA derivatives of the other amino acids).

The experiments and analyses were performed again using the methods stated above at temperatures 293.15, 313.15 and 323.15 K. Both experiments and analyses were done in duplicate.

2.6 Results

The average maximum solubility for each of the 9 α -amino acids in the model industrial mixture, their standard deviations and their solubility coefficients for Equation (2.5) are reported in Table 2.2 and 2.3. The column labelled +/- represents one standard deviation from the calculated average maximum solubility.

All of the solubilities increased as the temperature increased. The average standard deviation of the concentrations was 4%. The only standard deviations above 10% were l-glutamine at 293.15 K and l-tyrosine at 335.15 K, due to their relatively low concentrations in solution.

In Figures 2.1 to 2.5, the solubilities of the amino acids are plotted as a function of T/K. Three lines were fitted through the data in Table 2.2 and 2.3. The solid line represents the van 't Hoff Equation (2.5) fitted through the literature data (closed circles). The dotted line represents the Sober Equation (2.6) fitted through the literature data (closed circles). The dashed line represents the van 't Hoff Equation (2.5) fitted through α -amino acids measured in the mixtures. The crosses represent the solubility measured for the amino acids in the model system of industrial residues. The open circles represent the concentrations interpolated from the van 't Hoff Equation (2.5). The fits of the other 11 α -amino acids are shown in the appendix of this chapter in Fig. A2.1 to A2.6.

Table 2.2: Measured solubility of the α -amino acids in the model industrial residue mixture with the standard deviation noted by the column \pm . The solubility is given in g amino acid/100 g H₂O. These solubilities were converted to molar concentration (m_i) and are used to calculate the standard free energies using equation (2.8).

Amino Acid	277.15 K		293.15 K		315.15 K		335.15 K	
	Solubility g amino acid · 100 g H ₂ O ⁻¹	\pm	Solubility g amino acid · 100 g H ₂ O ⁻¹	\pm	Solubility g amino acid · 100 g H ₂ O ⁻¹	\pm	Solubility g amino acid · 100 g H ₂ O ⁻¹	\pm
L-Lysine	2.419	0.178	2.261	0.135	2.428	0.044	2.575	0.063
L-Glutamine	0.007	0	0.010	0.002	0.034	0.003	0.02	0.002
L-Tyrosine	0.232	0.013	0.280	0.007	0.249	0.001	0.425	0.055
L-Tryptophan	2.474	0.072	2.414	0.038	2.866	0.006	3.014	0.132
L-Methionine	1.297	0.021	1.404	0.049	1.958	0.026	2.343	0.137
L-Phenylalanine	1.584	0.064	1.819	0.013	2.477	0.027	2.725	0.133
L-Leucine	0.762	0.031	0.799	0.012	0.904	0.018	1.095	0.06
L-Valine	2.795	0.11	2.743	0.018	2.906	0.077	3.035	0.137
L-Isoleucine	1.320	0.5	1.349	0.009	1.441	0.028	1.795	0.073

Table 2.3: The solubility coefficients a^*/R and c derived from Equation (2.5) and the coefficients δ and θ derived from Equation (6) for calculating the solubilities of each of the 20 proteinogenic α -amino acids in single solution. These coefficients are used to calculate the solubilities of the individual amino acids at any temperature using equations (2.5) and (2.6).

Amino Acid	Coefficient a^*/R g α -amino acid \cdot K \cdot Kg H ₂ O ⁻¹	Coefficient c g α -amino acid \cdot Kg H ₂ O ⁻¹	Coefficient δ g α -amino acid \cdot Kg H ₂ O ⁻¹	Coefficient θ g α -amino acid \cdot K ⁻¹ \cdot Kg H ₂ O ⁻¹
L-Alanine	1180.439	9.058	1.873	0.01085
L-Cysteine	2758.141	7.033	-11.549	0.03123
L-Aspartic Acid	4221.993	15.534	-8.816	0.03497
L-Glutamic Acid	4475.994	16.929	-8.950	0.03723
L-Phenylalanine	1788.809	9.355	-1.439	0.01617
Glycine	1666.603	11.108	0.560	0.01661
L-Histidine	1836.889	9.911	-0.945	0.01590
L-Isoleucine	846.452	6.392	0.978	0.00863
L-Lysine	1759.514	11.374	0.970	0.01525
L-Leucine	1056.396	6.698	0.074	0.01032
L-Methionine	1401.158	8.727	-0.828	0.01626
L-Asparagine	2930.182	13.043	-4.168	0.02508
L-Proline	1036.993	10.870	4.506	0.00968
L-Glutamine	2906.456	13.469	-3.610	0.02490
L-Arginine	3316.840	16.354	-3.095	0.02830
L-Serine	3136.077	16.543	-5.031	0.03706
L-Threonine	1542.152	9.722	0.588	0.01341
L-Valine	728.103	6.514	1.733	0.00783
L-Tryptophan	2231.585	9.812	-3.543	0.01986
L-Tyrosine	4042.381	12.550	-10.802	0.03359

Only L-tyrosine and L-tryptophan had increased solubilities in the model industrial residue compared to single solution. The solubility of L-phenylalanine was slightly lower in the model industrial residue compared to single solution. The solubility of the aliphatic α -amino acids L-valine, L-isoleucine, L-leucine and L-methionine were also lower in the model residue compared to single solution. The solubility of L-lysine and L-glutamine decreased the most of all α -amino acids in this study in the model residue compared to single solution.

The ΔG°_{tr} as calculated from Equation (8) at $T/K = (277.15, 293.15, 313.15 \text{ and } 323.15)$ are shown in Figure 2.. The α -amino acids are grouped by the structure of their side chains: phenylic, aliphatic and nitrogen containing.

The only α -amino acids with a positive ΔG°_{tr} are L-tyrosine and L-tryptophan. These α -amino acids also have phenyl rings in their side chains. L-Tyrosine has a hydroxide group attached to the phenyl group and L-tryptophan has additionally an indole group attached to the

phenyl group. L-Phenylalanine is the only α -amino acid with a phenylic ring that has a negative ΔG°_{tr} , however only slightly.

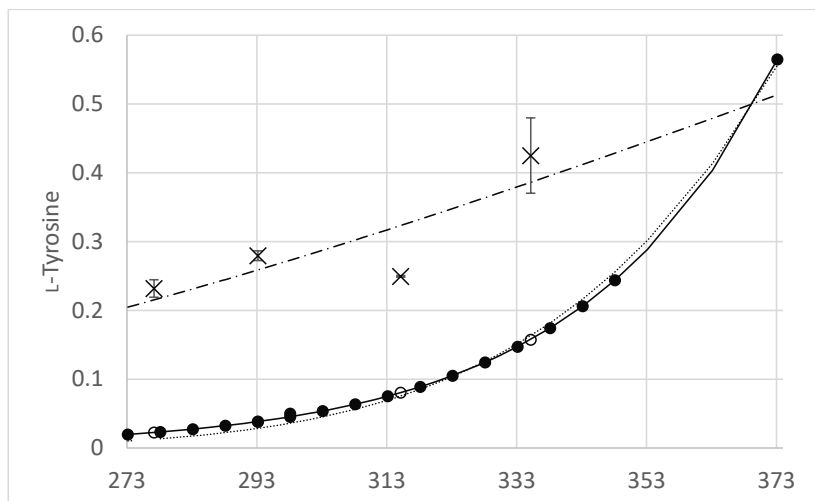


Figure 2.1: Solubility of L-tyrosine at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. Empty circles represent the solubility data of the single amino acids as calculated from Equation 2.5. Crosses represent data measured in the model mixture of 20 amino acid. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data. The dot-dash line is the van 't Hoff model applied to the measured mixture data.

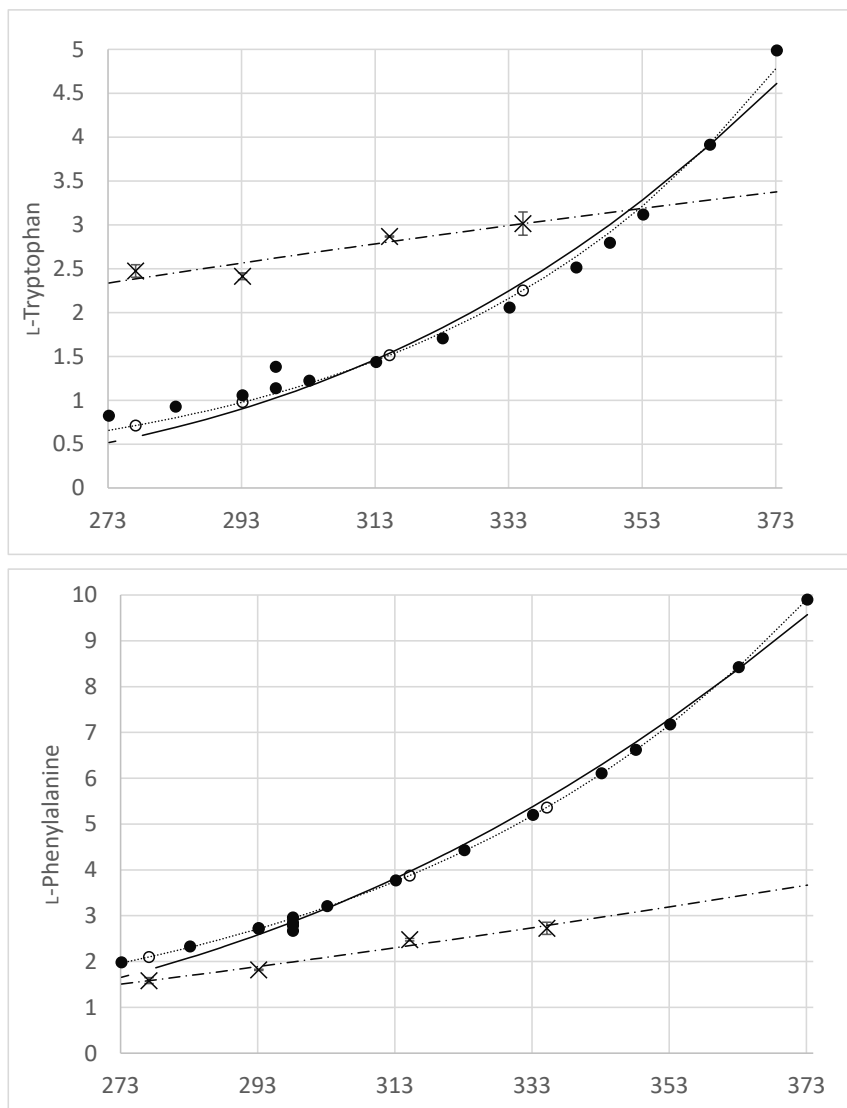


Figure 2.2: Solubility of L-tyrosine and L-phenylalanine at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. Empty circles represent the solubility data of the single amino acids as calculated from Equation 2.5. Crosses represent data measured in the model mixture of 20 amino acid. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data. The dot-dash line is the van 't Hoff model applied to the measured mixture data.

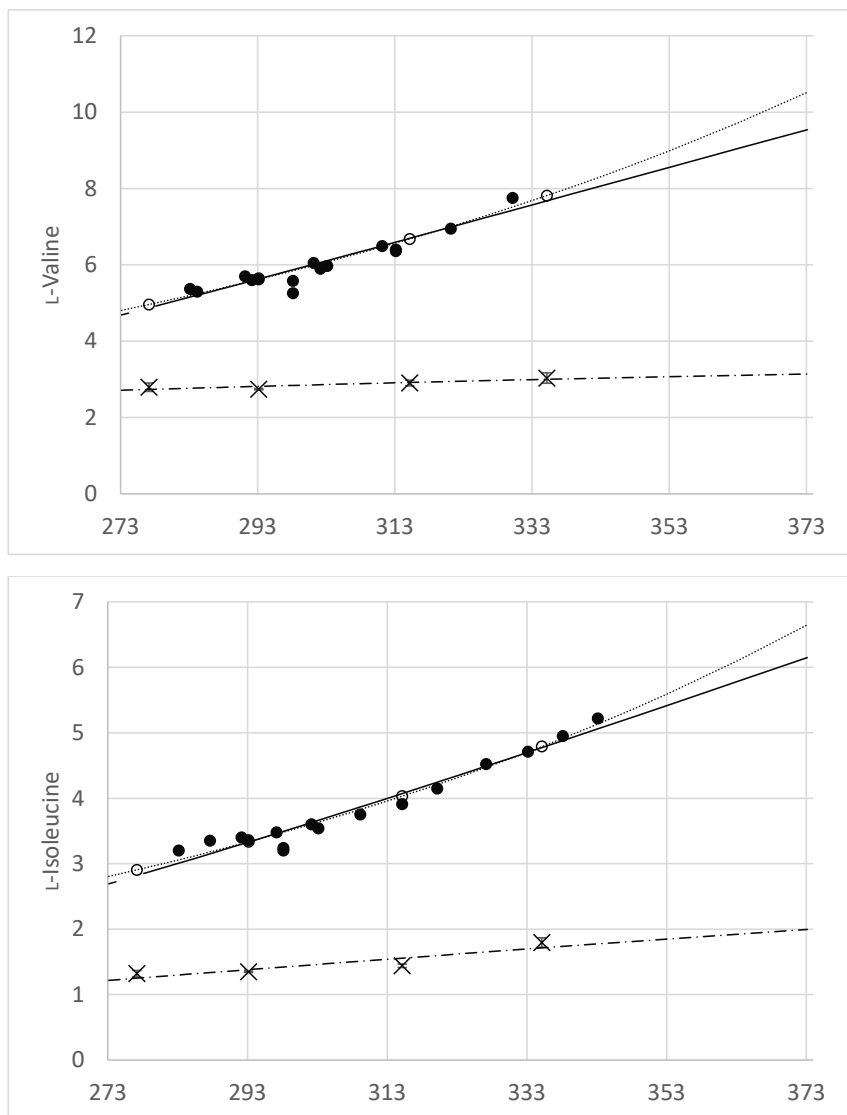


Figure 2.3: Solubility of L-valine and L-isoleucine at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. Empty circles represent the solubility data of the single amino acids as calculated from Equation 2.5. Crosses represent data measured in the model mixture of 20 amino acid. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data. The dot-dash line is the van 't Hoff model applied to the measured mixture data.

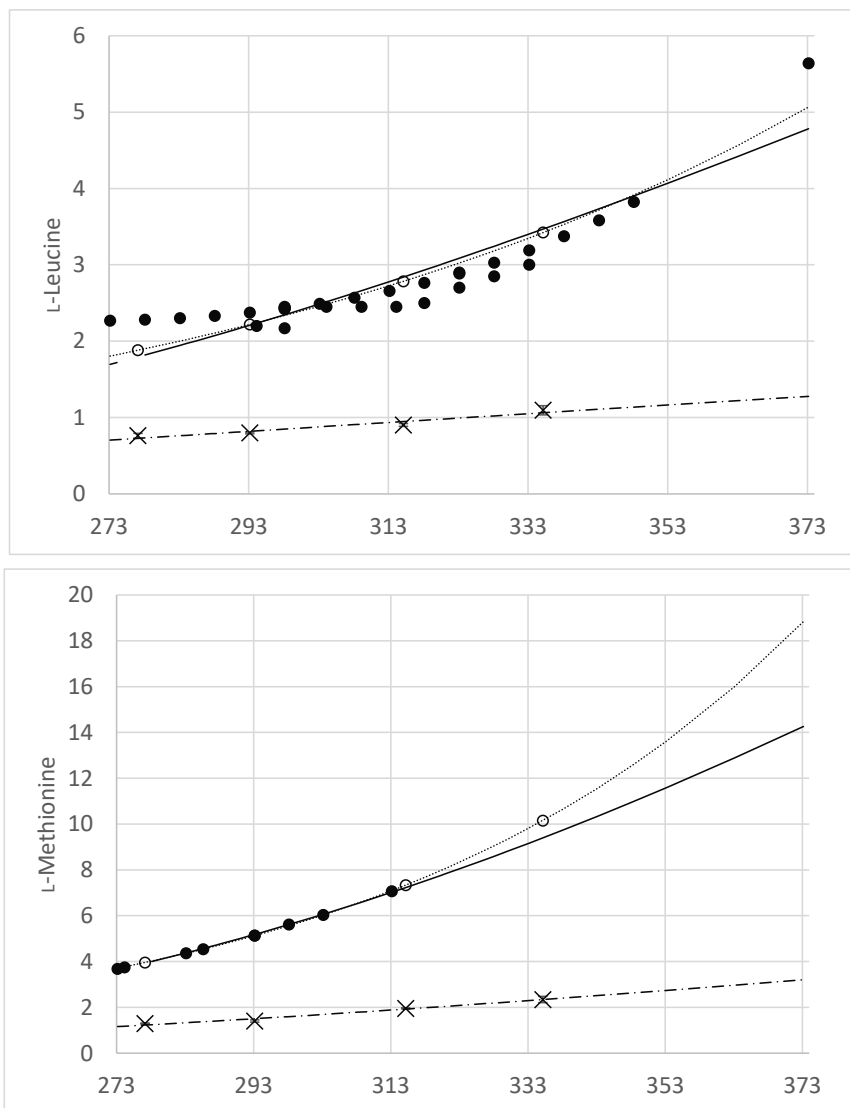


Figure 2.4: Solubility of L-leucine and L-methionine at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. Empty circles represent the solubility data of the single amino acids as calculated from Equation 2.5. Crosses represent data measured in the model mixture of 20 amino acid. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data. The dot-dash line is the van 't Hoff model applied to the measured mixture data.

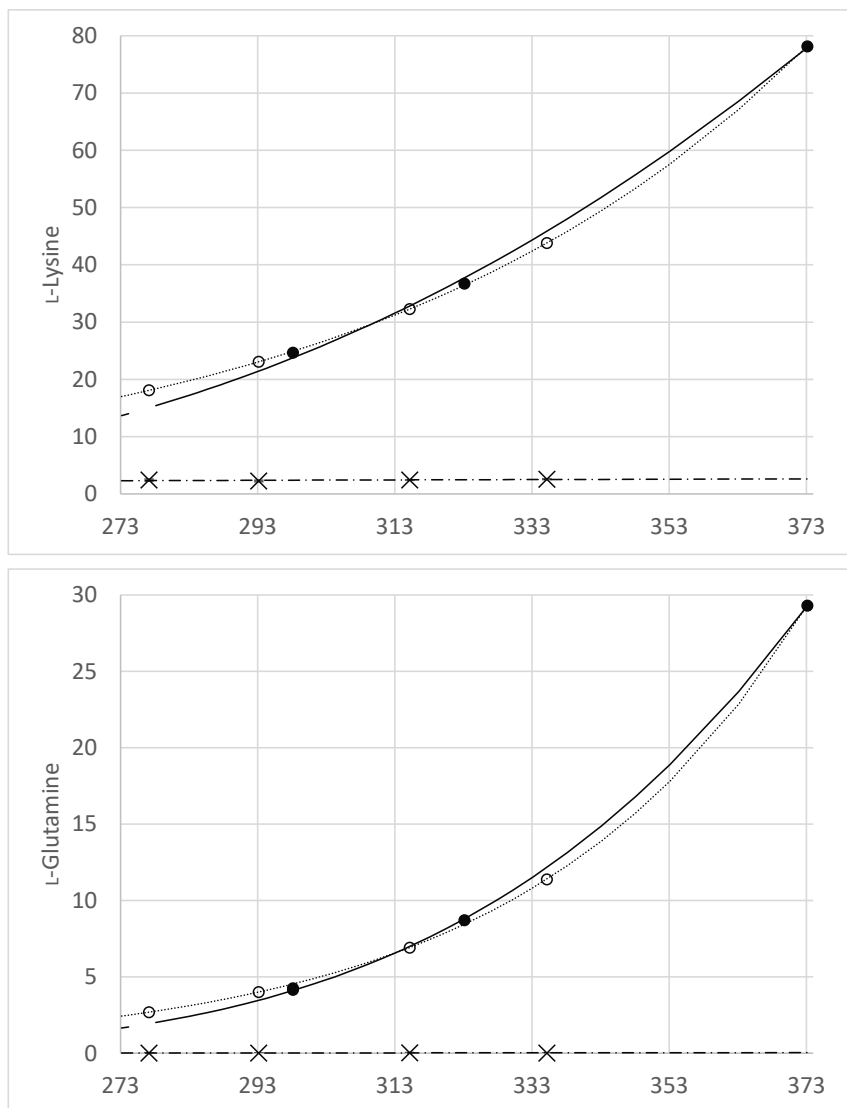


Figure 2.5: Solubility of L-lysine and L-glutamine at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. Empty circles represent the solubility data of the single amino acids as calculated from Equation 2.5. Crosses represent data measured in the model mixture of 20 amino acid. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data. The dot-dash line is the van 't Hoff model applied to the measured mixture data.

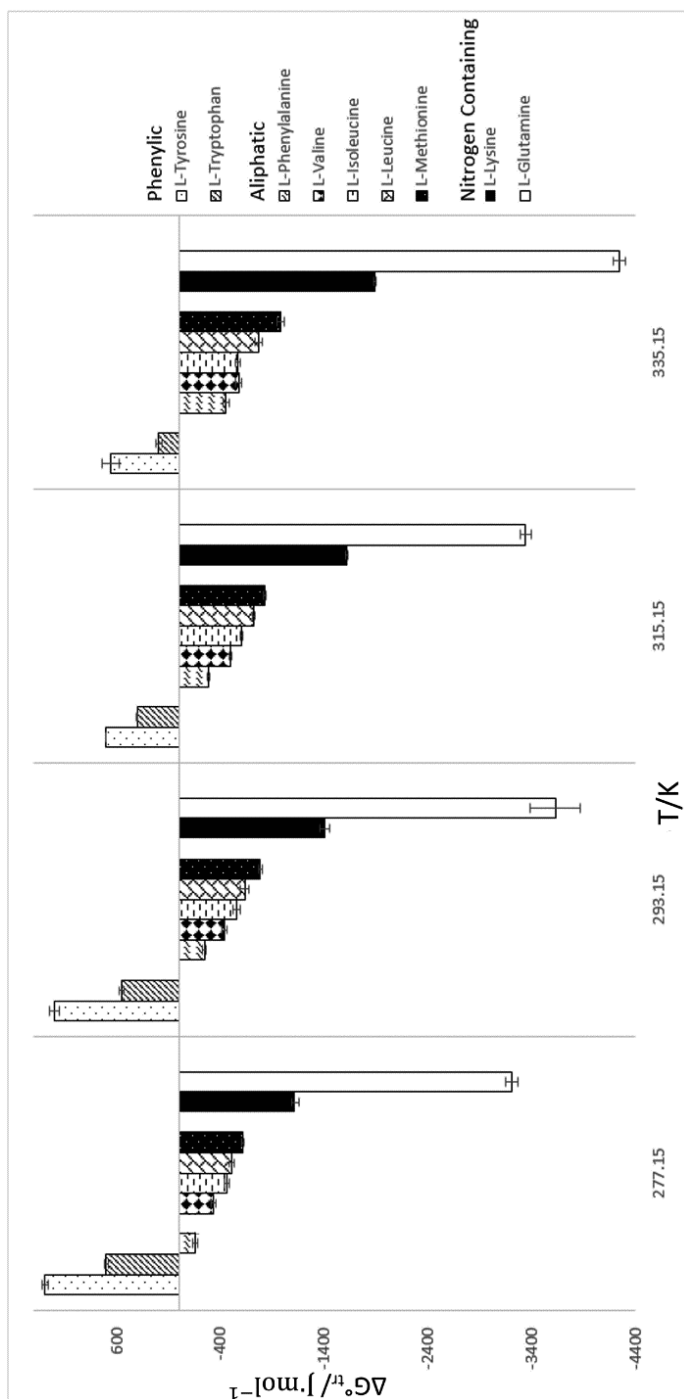


Figure 2.10: The free energy of transfers, ΔG°_{tr} , for the α -amino acids grouped by their side chain physical characteristics at 277.15, 293.15, 315.15, and 335.15 K. The amino acids are grouped by the characteristics of the functional groups on their side chains as phenyllic (L-tyrosine and L-tryptophan), aliphatic (L-phenylalanine, L-valine, L-isoleucine, L-leucine and L-methionine) and nitrogen containing (L-lysine and L-glutamine). The phenyllic amino acids have a positive ΔG°_{tr} , while all other amino acids have negative ΔG°_{tr} at all temperatures. Amino acids that have nitrogen in their side chains had the most negative ΔG°_{tr} at all temperatures.

All of the aliphatic α -amino acids have negative ΔG°_{tr} . The aliphatic α -amino acid with the shortest side chain has the least negative ΔG°_{tr} . The aliphatic α -amino acid with the longest side chain has the most negative ΔG°_{tr} . The longer the aliphatic side chain the more negative the ΔG°_{tr} .

Both of the α -amino acids that contain an additional nitrogen have the most negative ΔG°_{tr} values. L-Lysine has a side chain with a lysyl group ($[\text{CH}_2]_4\text{NH}_2$) on the end. Furthermore, because L-lysine is the only α -amino acid under consideration in this article with an isoelectric point above the pH of the mixture, the solubility of the individual α -amino acid would be even higher than calculated at the pH of the model industrial residue [44]. Therefore, taking this into consideration, the ΔG°_{tr} value of L-lysine would be even lower if the solubility was calculated at pH 5.3 as in the model residue. L-Glutamine has both an oxygen and an amide at the end of a three-carbon side chain. It has the lowest ΔG°_{tr} value of all the α -amino acids in this study.

The amino acids L-aspartic acid, L-glutamic acid and L-asparagine, were measured at concentrations of 3.0 g·100 g of H₂O⁻¹ in the mixture. This is above the solubility values in single solution reported in the literature, but not quantifiable as no more than 3 grams was added. Further research using more specific model industrial residues tailored for these α -amino acids need to be developed in order to understand the solubility of these α -amino acids in mixtures.

2.7 Concluding discussion

The solubilities of all α -amino acids in mixtures can be described by the same thermodynamic relations as those for single α -amino acids in model industrial residues, but with different coefficients. Furthermore, the ΔG°_{tr} values of the α -amino acids plotted against temperature can be described linearly and with a negative slope (Fig. A.2). This means that the solubilities of all the α -amino acids in the model system increase less as a function of temperature than the same α -amino acid in single solution.

The changes in the solubility of α -amino acids in this study when in mixtures as compared to being in single solution are similar for α -amino acids with similar molecular structures. By grouping the amino acids by their side chain structure, general conclusions can be drawn on the solubility of amino acids.

The solubility of aliphatic α -amino acids (L-valine, L-leucine and L-isoleucine) decreases in mixed solution in comparison to in single solution. L- Methionine, often regarded as an aliphatic α -amino acid, also behaves as an aliphatic α -amino acid in this respect. The solubility of all of these α -amino acids was lowered and the results show that the longer the side chain, the more the solubility was lowered in solutions of mixed α -amino acids.

Our results also show that an α -amino acid with a side chain that has a phenyl ring has an increased solubility when in a mixture of α -amino acids. Three α -amino acids in this study have side chains with aromatic, homocyclic, phenyl rings. Both L-tryptophan and L-tyrosine have increased solubilities in mixed solution. L-Phenylalanine is phenylic but also only has carbon and hydrogen in the side chain. L-Phenylalanine shows a solubility change when in a mixture that is between the solubility increase of the phenylic α -amino acids and the slight decrease of the aliphatic α -amino acids.

Solubility increases for the α -amino acids with oxygen in their side chains. Results show this for the 5 α -amino acids that have side chains containing oxygen. These amino acids have side chains that contain either a hydroxylic group, such as L-tyrosine; or a carboxyl group, like L-aspartic acid and L-glutamic acid; or an oxygen atom, such as L-asparagine and L-glutamine. All of these α -amino acids showed elevated solubility in a mixture with the exception of L-glutamine.

Having additional nitrogen in the side chain is a poor predictor of how the solubility of an α -amino acid will change in a mixture. Four α -amino acids in this study have side chains that contain additional nitrogen. L-Lysine and L-glutamine have nitrogen in their side chains and both of them had decreased solubility in the model system. L-Asparagine and L-tryptophan showed increases in solubility.

The simplification of the van 't Hoff equation and Sober equation both describe the data well.

References

- [1] E. Scott, F. Peter, J.P.M. Sanders, *Appl. Microbiol. Biotechnol.* 75 (2007) 751-762.
- [2] T.M. Lammens, M.C.R. Franssen, E. Scott, J.P.M. Sanders, *Biomass Bioenergy* 44 (2012) 168-181.
- [3] T.M. Lammens, *Bio-based Industrial Chemicals from Glutamic Acid*, (2011) PhD Dissertation.
- [4] W.B. Widayarani, N. Bowden, J.P.M. Sanders, R. Kolfschoten, M.E. Bruins, *Ind. Eng. Chem. Res.* (2016).
- [5] J. B. Dalton, C. L. A. Schmidt, *J. Biol. Chem.* 103 (1933) 549-578.
- [6] J. B. Dalton, C. L. A. Schmidt, *J. Biol. Chem.* 109 (1935) 8.
- [7] M. S. Dunn, F. J. Ross, L. S. Read, *J. Biol. Chem.* 103 (1933), 579-595.
- [8] E.J. Cohn, T.L. McMeekin, J.T. Edsall, J.H. Weare, *JACS* 56 (1934) 2270-2282.
- [9] T.L. McMeekin, E.J. Cohn, E. J. J.H. Weare, *JACS* 57 (1935) 626-633.
- [10] E.J. Cohn, T.L. McMeekin, J.D. Ferry, M.H. Blanchard, *J. Phys. Chem* (1939) 43.
- [11] V.E. Bower, R.A. Robinson, *J. Res. Nat. B. St. Sect. a-Phys. Chem.* 69 (1965) 131-137.
- [12] M.G. Brown, R.W. Rousseau, *Biotechnol. Progr.* 10 (1994) 253-257.
- [13] M.K. Khoshkbarchi, J.H. Vera, *Ind. Eng. Chem. Res.* 36 (1997) 2445-2451.
- [14] M.N. Islam, R. K.Wadi, *Physics and Chemistry of Liquids* 39 (2001) 77-84.
- [15] R. Badarayani, A. Kumar, *Fluid Phase Equilib.* 201 (2002) 321-333.
- [16] L. A. Ferreira, E. A. Macedo, S. P. Pinho, *J. Chem. Thermodyn.* 41 (2009), 193-196.
- [17] Q. Yuan, Z.-F. Li, B.-H. Wang, *J. Chem. Thermodyn.* 38 (2006) 20-33.
- [18] C. Held, T. Reschke, R. Muller, W. Kunz, G. Sadowski, *J. Chem. Thermodyn.* 68 (2014) 1-12.
- [19] T.S.K. Banipal, P.K. Banipal, G. Singh, *J. Chem. Thermodyn.* 39 (2007) 371-384.
- [20] R.Sadeghi, A. Gholamireza, *J. Chem. Thermodyn.* 43 (2011) 200-215.
- [21] T.S.K. Banipal, P.K. Banipal, G. Singh , *J. Chem. Thermodyn.* 43 (2011) 200-215.
- [22] R. Sadeghi, A. Gholamireza, *J. Chem. Thermodyn.* 43 (2011) 200-215.
- [23] C. Held, L.F. Cameretti,; G. Sadowski, *Ind. Eng. Chem. Res.* 50 (2011) 131-141.
- [24] R.d.F. Ferreira, *Separation and Purification of Amino Acids*. 2008. PhD Dissertation.
- [25] R.L. Lundblad, F.M. MacDonald, *Handbook of Biochemistry and Molecular Biology*. 2010, 65.

- [26] J.H.v. 't Hoff, *Etudes de Dynamique Chimique*. (1884), 242.
- [27] S.H. Yalkowsky; S.C. Valvani; G.L. Amidon, *J. Pharm. Sci.* 65 (1976) 6.
- [28] D.J.W. Grant, M. Mehdizadeh, A.H.L. Chow, J.E. Fairbrother, *Int. J. Pharm.* 18 (1984) 25-38.
- [29] A. Apelblat, E. Manzurola, *The J. Chem. Thermodyn.* 31 (1999) 85-91.
- [30] Sober, H., *Handbook of Biochemistry; Selected Data for Molecular Biology* (1968).
- [31] E.P.K. Hade, Ph.D. Dissertation. University of Chicago Press (1962).
- [32] J.P. Amend, H. C. Helgeson, *Pure Appl. Chem.* 69 (1997) 935-942.
- [33] T.S. Tomiyama, C.L.A. Schmidt, *J. Gen. Physiol.* 19 (1935) 139.
- [34] J.B. Dalton, C.L.A Schmidt *J. Gen. Physiol.* 19 (1936) 767-771.
- [35] R.C. Zumstein, R.W. Rousseau, *Ind. Eng. Chem. Res.* 1989, 28 (3), 334-340.
- [36] M. Gatewood Brown, R.W. Rousseau, *Biotechnol. Progr.* 10 (1994) 253-257.
- [37] G. Fasman, *The Development of the Prediction of Protein Structure* (1973).
- [38] I. Kurosawa, A.S. Teja, R.W. Rousseau, *Fluid Phase Equilib.* 224 (2004) 245-249.
- [39] J.C. Givand, A.S. Teja, R.W. Rousseau, *AIChE J.* 47 (2001) 2705-2712.
- [40] Y. Nozaki, C. Tanford, *J. Biol. Chem.* 246 (1971) 2211-2217.
- [41] R. Carta, G. Tola, *J. Chem. Eng. Data* 41 (1996) 414-417.
- [42] C.J. Orella, D.J. Kirwan, *Biotechnol. Progr.* 5 (1989) 89-91.
- [43] B. Meussen, A.T. van Zeeland, M.E. Bruins, J.P.M. Sanders, *Food Anal. Methods* 7 (2014) 1047-1055.
- [44] H.-C. Tseng, C.-Y. Lee, W.-L. Weng, I.M. Shiah, *Fluid Phase Equilib.* 285 (2009) 90-95.

Appendix A

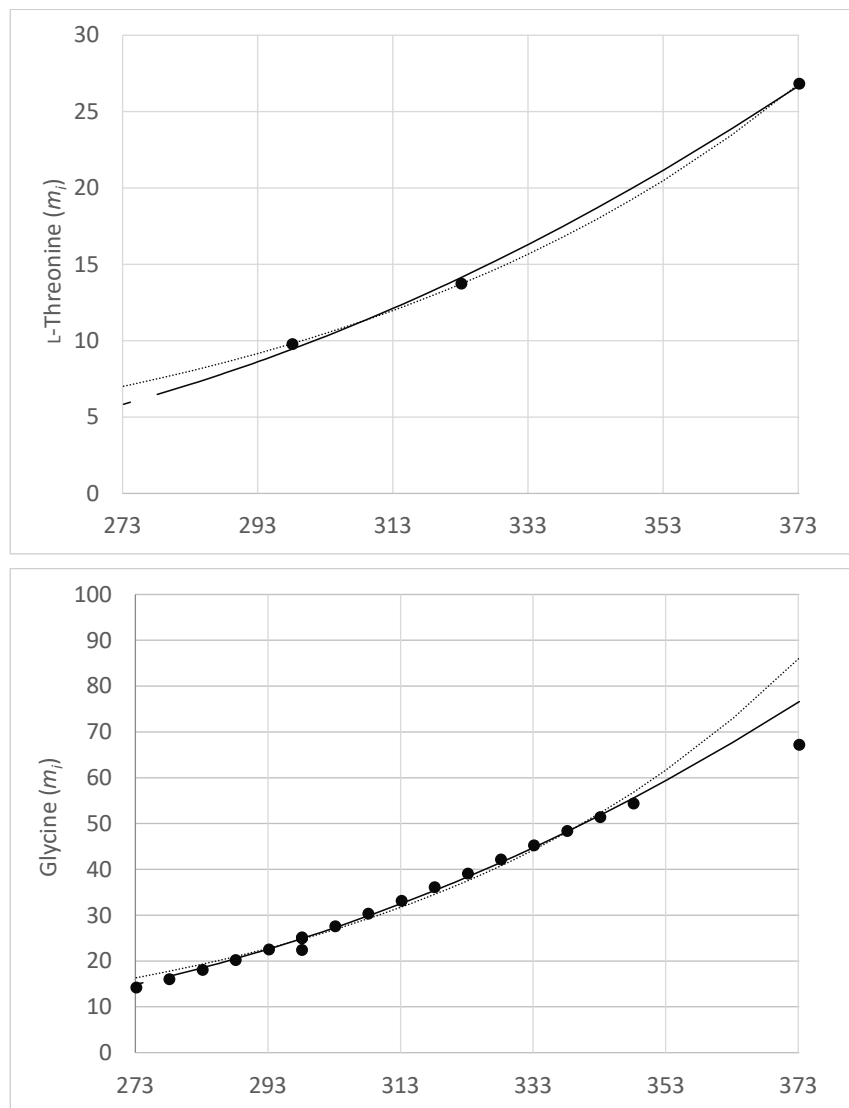


Fig. A2.1: Solubility of L-threonine and glycine at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data.

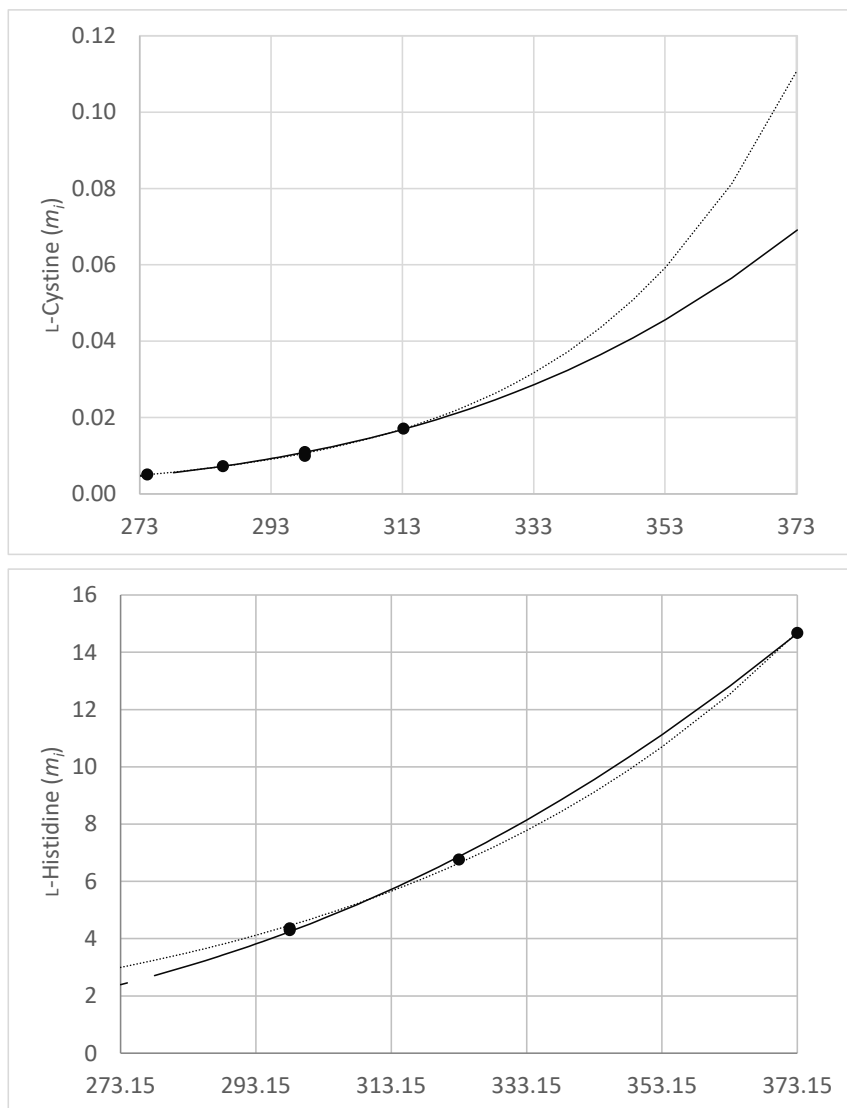


Fig. A2.2: Solubility of L-cystine and L-histidine at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data.

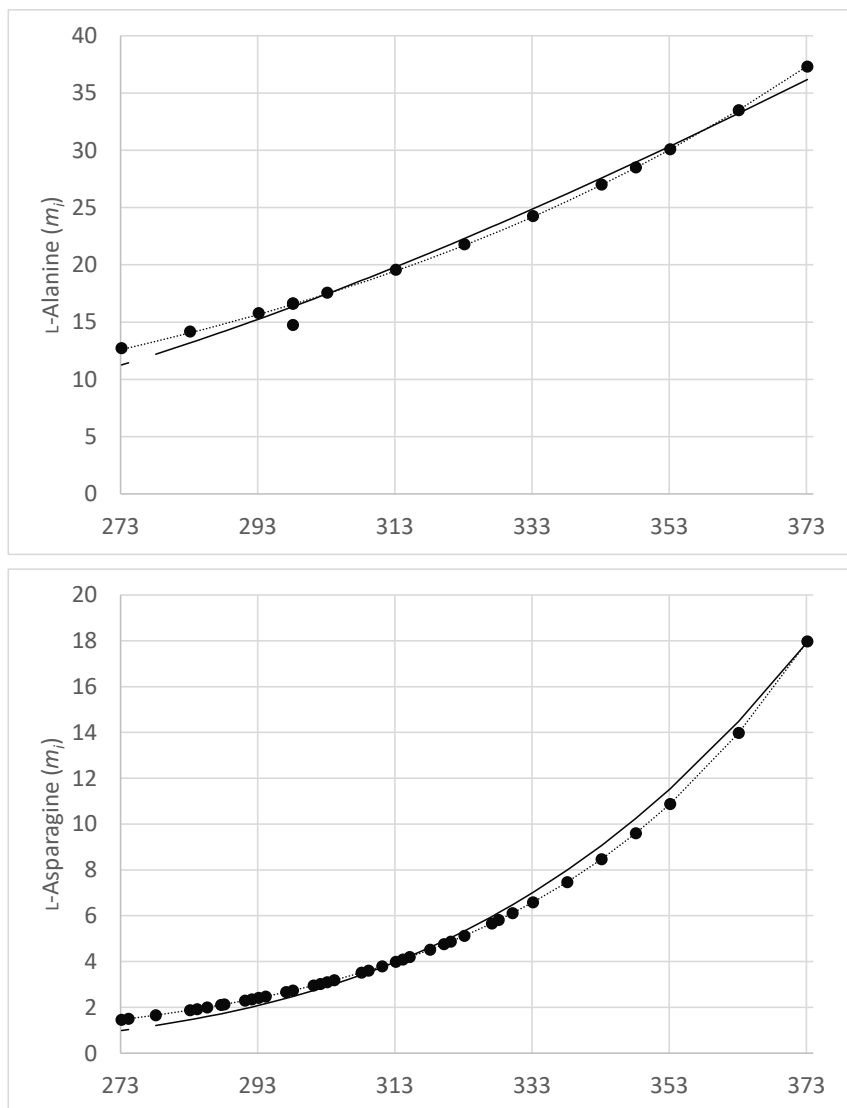


Fig. A2.3: Solubility of L-alanine and L-asparagine at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data.

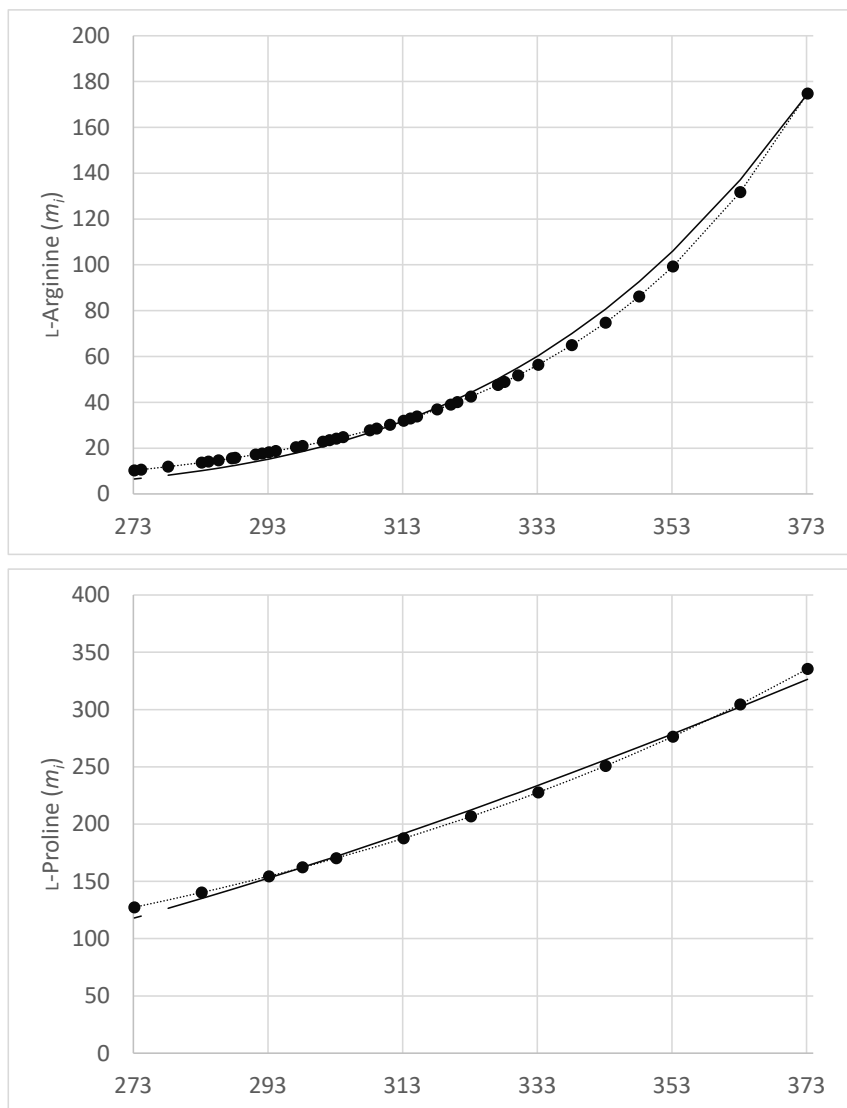


Fig. A2.4: Solubility of L-arginine and L-proline at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data.

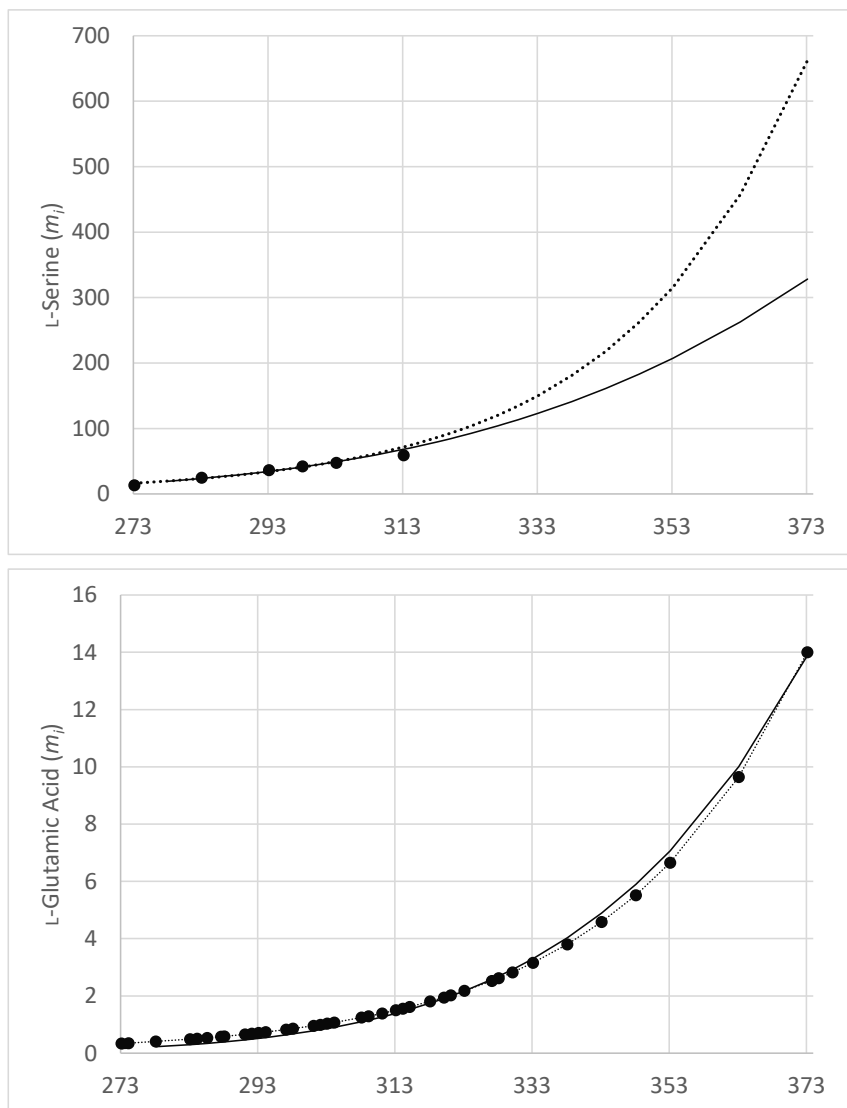


Fig. A2.5: Solubility of L-serine and L-glutamic acid at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data.

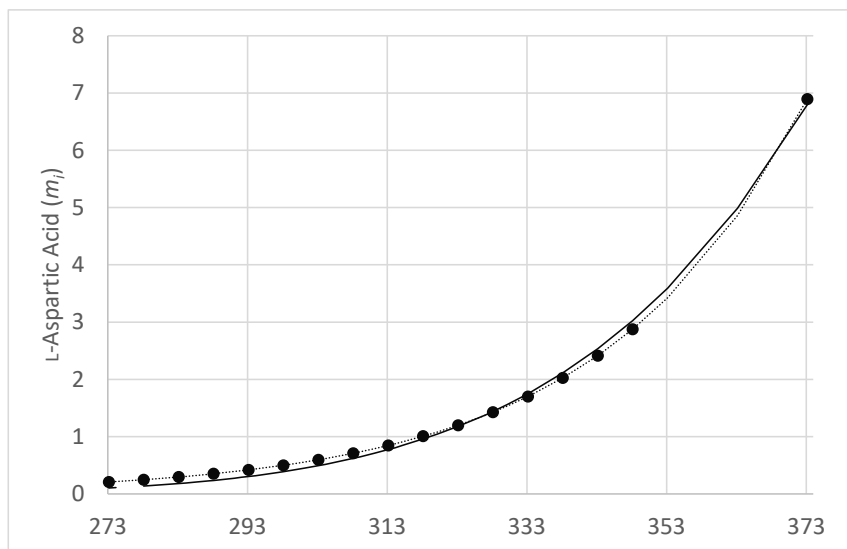


Fig. A2.6: Solubility of L-aspartic acid at various temperatures (K). Solid circles represent the solubility data of the single amino acids taken from the literature. The solid line represents the van 't Hoff model (Equation 2.5) applied to the literature data. The dotted line represents the Sober model (Equation 2.6) of the literature data.

The three models for solubility; Equations (4), (5) and (6), were fit to the literature data. The fit was performed by regressing the coefficients to minimize the NRMSV. For the amino acids L-alanine, L-cystine, L-aspartic acid, L-glutamic acid, L-phenylalanine, L-isoleucine, L-lysine, L-leucine, L-proline, L-threonine, L-valine, L-tryptophan and L-tyrosine, Equation (6) published in the reference text by Sober had a better fit. For glycine, L-histidine, L-methionine, L-asparagine, L-glutamine, L-arginine and L-serine, either the simplified van 't Hoff Equation (5) or both the simplified and full van 't Hoff equation (4) provided a better fit. This is shown in Table A.1.

Table A2.1: Normalized Root Mean Square of Variation (NRMSV) for all 20 α -amino acids calculated with the Sober equation (Equation 6), the approximated van 't Hoff equation (Equation 2.5) and the van 't Hoff equation (Equation 2.4). The Sober model has a better fit than the van 't Hoff model for all amino acids with the exception of glycine, L-histidine, L-asparagine, L-glutamine and L-serine. The simplified van 't Hoff equation and the full van 't Hoff equation show similar accuracies.

	Sober	van 't Hoff	
	$\ln S = \delta + \theta T$	$\ln x_i^s = -\frac{a}{RT} + c$	$\ln x_i^s = -\frac{a}{RT} + \frac{b}{R} \ln T + c$
L-Alanine	0.007	0.024	0.027
L-Cysteine	0.008	0.027	0.021
L-Aspartic Acid	0.000	0.051	0.053
L-Glutamic Acid	0.007	0.097	0.103
L-Phenylalanine	0.012	0.056	0.056
Glycine	0.039	0.017	0.017
L-Histidine	0.043	0.027	0.031
L-Isoleucine	0.061	0.075	0.075
L-Lysine	0.017	0.059	0.065
L-Leucine	0.208	0.270	0.270
L-Methionine	0.008	0.008	0.009
L-Asparagine	0.061	0.005	0.009
L-Proline	0.000	0.052	0.052
L-Glutamine	0.054	0.013	0.017
L-Arginine	0.040	0.028	0.033
L-Serine	0.042	0.035	0.035
L-Threonine	0.005	0.073	0.079
L-Valine	0.058	0.068	0.069
L-Tryptophan	0.095	0.166	0.166
L-Tyrosine	0.003	0.052	0.054

Showing the ΔG°_{tr} along a temperature gradient as a line graph shows more clearly that increasing temperature has a greater effect on increasing solubility in single solution than in the model industrial mixture of α -amino acids. In the Fig. A.2, you see that all of the slopes are similarly negative as temperature increases.

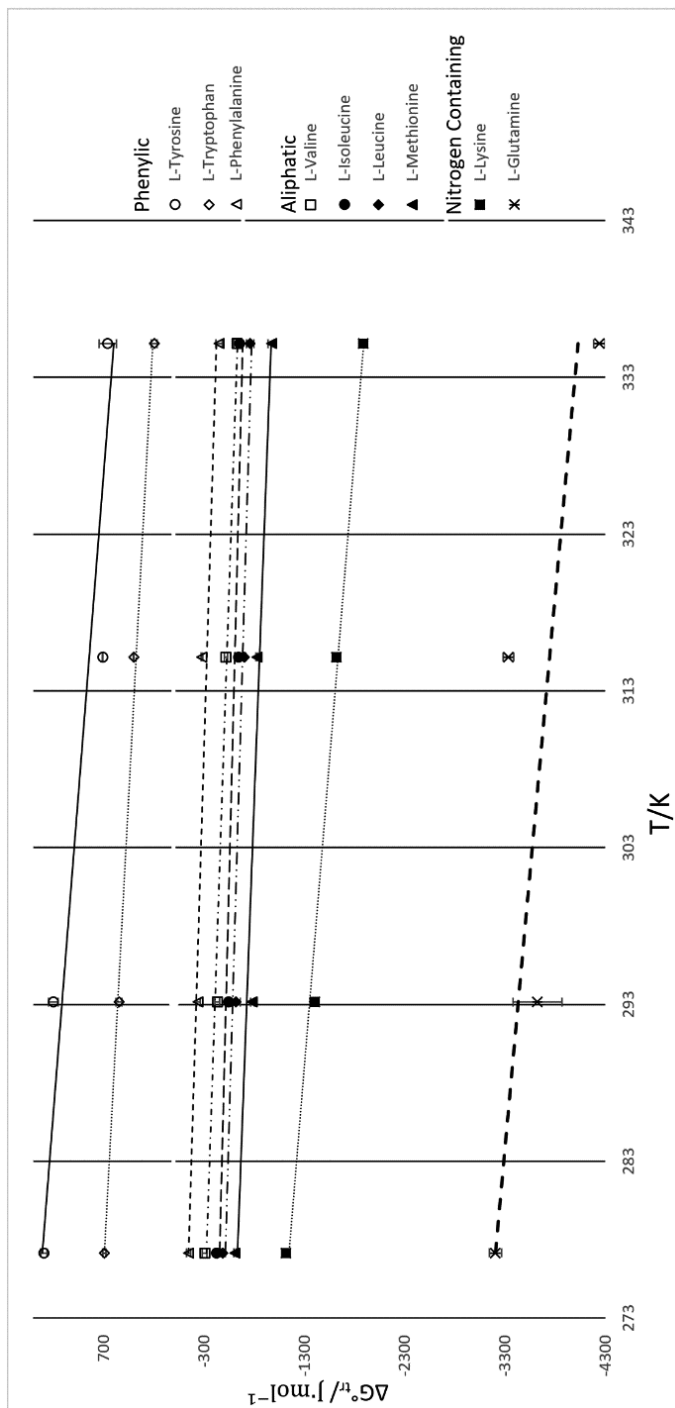


Fig. A.27: The standard free energy of transfers, ΔG°_{tr} , for the α -amino acids grouped by their side chain physical characteristics. The slopes are similar showing an equal effect of temperature on the solubility of all amino acids

Table A2.2 Solubility reported in the literature for each amino acid. Sources: a Dalton and Schmidt [5], b Hade [31], c Amend and Hegelson [32], d Dalton and Schmidt [6], e Tomiyama and Schmidt [33], f Dalton and Schmidt [34]

Temp (K)	ALA	CYS	ASP	GLU	PHE	GLY	HIS	ILE	LYS	LEU
273.15	12.73	-	0.21	0.34	1.98	14.18	-	-	-	2.27
274.15	-	0.01	-	-	-	-	-	-	-	-
278.15	-	-	0.25	-	-	16.03	-	-	-	2.28
283.15	14.17	-	0.30	0.50	2.33	18.04	-	3.20	-	2.30
285.65	-	0.01	-	-	-	-	-	-	-	-
288.15	-	-	0.35	-	-	20.20	-	-	-	2.33
293.15	15.78	-	0.42	0.72	2.74	22.52	-	3.36	-	2.37
298.15	16.65	0.01	0.50	0.86	2.97	24.99	4.36	-	24.66	2.43
303.15	17.57	-	0.60	1.02	3.21	27.59	-	3.54	-	2.49
308.15	-	-	0.71	-	-	30.32	-	-	-	2.57
313.15	19.57	0.02	0.85	1.47	3.77	33.16	-	3.75	-	2.66
318.15	-	-	1.01	-	-	36.10	-	-	-	2.76
323.15	21.79	-	1.20	2.19	4.43	39.10	6.76	-	36.72	2.89
328.15	-	-	1.43	-	-	42.18	-	-	-	3.03
333.15	24.26	-	1.70	-	5.20	45.26	-	-	-	3.19
338.15	-	-	2.03	-	-	48.35	-	-	-	3.37
343.15	27.02	-	2.41	-	6.11	51.39	-	-	-	3.58
348.15	28.51	-	2.88	5.53	6.62	54.39	-	-	-	3.82
353.15	30.08	-	-	-	7.18	-	-	-	-	-
363.15	33.50	-	-	-	8.43	-	-	-	-	-
373.15	37.30	-	6.89	14.00	9.90	67.17	14.68	-	78.15	5.64
Source	a, b	b	a	a, b	d	a	c	b	c	a

Table A2.3 Solubility reported in the literature for each amino acid. Sources: a Dalton and Schmidt [5], b Hade [31], c Amend and Hegelson [32], d Dalton and Schmidt [6], e Tomiyama and Schmidt [33], f Dalton and Schmidt [34]

Temp (K)	MET	ASN	PRO	GLN	ARG	SER	THR	VAL	TRP	TYR
273.15	3.68	-	127.40	-	-	13.30	-	-	0.82	0.02
274.15	3.75	-	-	-	-	-	-	-	-	-
278.15	-	-	-	-	-	-	-	-	-	0.02
283.15	4.37	-	140.30	-	-	24.70	-	5.37	0.93	0.03
285.65	4.54	-	-	-	-	-	-	-	-	-
288.15	-	-	-	-	-	-	-	-	-	0.03
293.15	5.14	-	154.50	-	-	36.20	-	5.65	1.06	0.04
298.15	5.62	2.51	162.30	4.25	19.60	42.20	9.79	-	1.14	0.05
303.15	6.04	-	170.30	-	-	47.60	-	5.97	1.22	0.05
308.15	-	-	-	-	-	-	-	-	-	0.06
313.15	7.07	-	187.60	-	-	59.20	-	6.36	1.44	0.08
318.15	-	-	-	-	-	-	-	-	-	0.09
323.15	-	5.31	206.70	8.71	43.45	-	13.75	-	1.71	0.11
328.15	-	-	-	-	-	-	-	-	-	0.12
333.15	-	-	227.70	-	-	-	-	-	2.06	0.15
338.15	-	-	-	-	-	-	-	-	-	0.17
343.15	-	-	250.80	-	-	-	-	-	2.51	0.21
348.15	-	-	-	-	-	-	-	-	2.80	0.24
353.15	-	-	276.40	-	-	-	-	-	3.12	-
363.15	-	-	304.50	-	-	-	-	-	3.91	-
373.15	-	17.95	335.50	29.30	174.74	-	26.83	-	4.99	0.57
Source	b	c	e	c	c	b	c	b, f	d	a

Chapter 3

Solubility of the proteinogenic α -amino acids in water, ethanol and ethanol-water mixtures

Published as: Bowden, NA; Sanders, JPM and Bruins, ME (2018) Solubility of the Proteinogenic α -Amino Acids in Water, Ethanol, and Ethanol–Water Mixtures J. Chem. Eng. Data DOI: 10.1021/acs.jced.7b00486

Abstract:

The addition of organic solvents to α -amino acids in aqueous solution could be an effective method in crystallization. We reviewed the available data on the solubility of α -amino acids in water, water-ethanol mixtures and ethanol at 298.15 K and 0.1MPa. The solubility of L-alanine, L-proline, L-arginine, L-cysteine and L-lysine in water and ethanol mixtures and the solubility of L-alanine, L-proline, L-arginine, L-cysteine, L-lysine, L-asparagine, L-glutamine, L-histidine and L-leucine in pure ethanol systems were measured and are published here for the first time. The impact on the solubility of amino acids that can convert in solution, L-glutamic acid and L-cysteine, was studied. At lower concentrations, only the ninhydrin method and the UPLC method yield reliable results. In the case of α -amino acids that convert in solution, only the UPLC method was able to discern between the different α -amino acids and yields reliable results. Our results demonstrate that α -amino acids with similar physical structures have similar changes in solubility in mixed water/ethanol mixtures. The solubility of L-tryptophan increased at moderate ethanol concentrations.

3.1 Introduction

Bio-based products made from proteins and α -amino acids (e.g. bioplastics, pharmaceuticals, fine chemicals) could become increasingly important [1-4]. One of the challenges is to find a way to separate α -amino acids from industrial residues. The literature data on the use of anti-solvents, such as ethanol, to lower the solubility of the α -amino acids to promote crystallization is incomplete. An understanding on the impact of anti-solvents on α -amino acids in solution is integral in designing technologies for separating α -amino acids from solution.

Many articles report the solubility measurements of α -amino acids in water [5-31] and in mixtures of alcohol and water [32-40]. Subsequent research focused on calculating the activity coefficients of these α -amino acids in water, water-ethanol and ethanol [41-43]. Recently, the effect of the addition of ethanol on the solubility of amino acids has also been applied to the crystallization of amino acids after protein hydrolysis [44]. However, the solubility of α -amino acids in a two-solvent system cannot be described by a first-degree exponential function with a discrete partition coefficient as espoused in the earliest solubility studies. This is due to the ternary interactions of the solvents to each other and with the solute.

In response to this, models have been developed to explain the solubility of a few α -amino acids in water, water-ethanol, ethanol and other two solvent systems [7, 45-52]. However, for several α -amino acids, no data have been published on their solubility in water-ethanol or ethanol systems. For many other α -amino acids, the data is incomplete or unreliable. For this reason, applying the models to all α -amino acids is not possible.

The goal of this article is to understand the solubility of the 20 proteinogenic α -amino acids in water, water-ethanol mixtures and ethanol. In order to achieve this goal, three research objectives are pursued. First, the methodologies of bringing the α -amino acids to maximum solubility and the analytical technique of measuring these concentrations in solutions of water, water-ethanol mixtures and ethanol are evaluated. Second, since recent evidence shows that some α -amino acids can convert to other α -amino acids, the solubility data in solutions of water, water-ethanol mixtures and ethanol of these α -amino acids at 298.15 K and 0.1MPa are re-evaluated. Third, the solubility of the α -amino acids L-alanine, L-arginine, L-lysine, L-proline and L-cysteine in solutions of water, water-ethanol mixtures at 298.15 K and 0.1MPa is measured. Furthermore, the solubility of L-alanine, L-arginine, L-lysine, L-proline, L-cysteine, L-asparagine,

L-aspartic acid, L-glutamine, L-histidine and L-leucine at 298.15 K and 0.1MPa are measured in ethanol.

3.1.1 Review of amino acid solubility data and methodologies

Data on the solubility of glycine, L-valine, L-serine, L-isoleucine, L-tryptophan, L-tyrosine, L-phenylalanine and L-threonine in water, water-ethanol mixtures and ethanol were found in peer reviewed journals. These data were obtained using disparate methodologies in both the dissolution of the α -amino acids as well as in their measurement.

Gravimetric measurement of the dry weight of a solute is a technique that is often used in measuring solubility [7]. However, the solubility of some α -amino acids (e.g. L-tyrosine) are extremely low. Furthermore, the solubility of all α -amino acids in ethanol are low. Measuring amino acids with low solubility gravimetrically would consume excessive amounts of ethanol to produce a few milligrams of the solute. Therefore, a spectrophotometric analytical technique using ninhydrin was developed to measure the concentrations of α -amino acids [45]. This article will evaluate these two measurement techniques and use a third technique, the UPLC method [54]. The UPLC method is able to detect concentrations of 2.3 μ M.

3.1.2 Impact of amino acid conversions on their solubilities

L-Cysteine can form a sulphur bond with itself upon oxidation to form the dimer cysteine [55]. There is only one piece of solubility data in the literature for the monomer L-cysteine [8]. However, the authors do not mention in their work that they took the oxidation reaction with L-cysteine into account when measuring the solubility. This article reports data measured on the solubility of L-cysteine in water, water-ethanol mixtures and ethanol under sealed oxygen-poor conditions. Furthermore, after measuring the solubility, the samples of L-cysteine in this work were analysed through mass-spectrometry to show that the formation of the dimer cystine was negligible.

Data on the solubility of L-glutamic acid in ethanol-water system were found by McMeekin *et al.* and expanded by other authors [12, 23]. However, L-glutamic acid has been shown to convert to L-pyroglutamic acid. The conversion to L-pyroglutamic acid increases as the temperature of the solution increases [56,57]. This was not considered in the initial solubility data. To account for this possibility, in this study, the solubility of L-glutamic acid was

determined by measuring the concentration of both L-glutamic acid and L-pyroglutamic acid in the same sample by using the UPLC method [53].

3.1.3 Incomplete solubility data of amino acids

Data on the solubility of L-asparagine, L-aspartic acid, L-glutamine, L-histidine and L-leucine in water, water and ethanol mixtures were published in peer reviewed journals, but did not include data in ethanol solutions [32, 35]. The solubilities of these α -amino acids were measured for this work ethanol using the UPLC method.

Data on the solubility of glycine in water, water-ethanol mixtures and ethanol are conflicting. Reports show the solubility of glycine in water to be 4.25 and 2.733 g per 100 ml. For this reason, the solubility of glycine in water, water-ethanol mixtures and ethanol was re-measured.

No published data could be found on the solubility of L-alanine, L-arginine, L-lysine, L-proline and L-cysteine in water-ethanol mixtures and ethanol. Their solubility in these systems was measured and reported here.

3.2 Experimental

The α -amino acids that were used in this article were purchased from Sigma-Aldrich. These α -amino acids were at least 99% pure. Table 3.1 lists the supplier and purity of the chemicals used in this work.

For L-arginine, L-lysine, L-proline, L-methionine, L-cysteine, L-glutamic acid and L-phenylalanine in water-ethanol mixtures, excess amounts of these α -amino acids were added to 15.0 ml Greiner tubes in duplicate. Then 0.0%, 25.0%, 50.0%, 75.0% and 100.0% ethanol (g/g) solutions in water were added to the Greiner tubes and sealed. The tubes were mixed and added to a jacketed shaking water bath set to 298.15 K and 0.1 MPa and left to mix at 80 rpm until they had reached equilibrium. Both the samples and the water in the water bath were continuously monitored. The amino acids were said to have reached equilibrium when successive measurements, 24-hours apart, yielded a concentration within the variation of the balance. All measurements were performed in duplicate. The solubility of glycine, L-asparagine, L-aspartic acid, L-glutamine, L-histidine and L-leucine was measured in ethanol using the same procedure.

The amino acids with low solubilities produced results below the detection limits of the balance. For this reason, these samples were analyzed by UPLC. For L-glutamic acid, all systems were measured with the UPLC as it enabled simultaneous measurement of L-pyroglutamic acid.

The UPLC method is based on automated pre-column derivatisation in the injection needle of the amino acid in an autosampler using o-phthalaldehyde (OPA) reagent in combination with 9-fluorenylmethyl chloroformate (Fmoc) that enables the amino acids to fluoresce. Separation was achieved with a Dionex RSLC system using an Acquity UPLC BEH C18 reversed-phase column. Sample analysis was performed with an UltiMate 3000 Rapid Separation pump and autosampler. Derivatised amino acids were detected at 263 nm (Fmoc derivative of L-proline) and 338 nm (OPA derivatives of the other amino acids).

Table 3.1: Description of Chemicals and Solvents Used

Chemical Name	Source	Mole Fraction Purity	Purification Method
glycine	Sigma-Aldrich	1.00	None
L-alanine	Sigma-Aldrich	0.98	None
L-arginine	Sigma-Aldrich	0.98	None
L-asparagine	Sigma-Aldrich	1.00	None
L-aspartic acid	Sigma-Aldrich	0.99	None
L-cysteine	Sigma-Aldrich	0.97	None
L-glutamic acid	Sigma-Aldrich	0.99	None
L-glutamine	Sigma-Aldrich	0.99	None
L-histidine	Sigma-Aldrich	0.99	None
L-leucine	Sigma-Aldrich	1.00	None
L-lysine	Sigma-Aldrich	0.97	None
L-methionine	Sigma-Aldrich	0.98	None
L-phenylalanine	Sigma-Aldrich	0.98	None
L-proline	Sigma-Aldrich	0.99	None
L-tryptophan	Sigma-Aldrich	0.98	None
L-tyrosine	Sigma-Aldrich	0.98	None
ethanol	Sigma-Aldrich	> 0.99	None

For the samples that were measured using the gravimetric analytical technique, approximately 3 grams of each solution was filtered through a sterile 0.45µm Minisart filter. Then, the filtered sample was added to a pre-dried and pre-weighed drying tin and weighed again using a AB204 Analytical Balance from Mettler Toledo to +/- 0.0001 grams. All samples were filtered and weighed in duplicate. The samples were put in a drying oven set at 315.15 K and 0.1 MPa for 5 days and weighed again. 24 hours later, the dry samples were weighed once more. This procedure was repeated until the weights were within the error range of the analytical balance, and the sample was assumed to be at equilibrium.

Additionally, samples of L-cysteine were measured on a LCQ Fleet Ion Trap mass spectrometer from Thermo ScientificTM. This was done to ensure there was undetectable L-cystine formation.

The saturated mole fraction solubility of all amino acids was calculated by Eq. (1), while the mole fraction composition of the solvent mixture was calculated by Eq. (2):

$$x_1 = \frac{m_1/M_1}{m_1/M_1 + m_2/M_2 + m_3/M_3} \quad (1)$$

$$x_2 = \frac{m_2/M_2}{m_2/M_2 + m_3/M_3} \quad (2)$$

Where m_1 , m_2 and m_3 are the mass of the amino acid, ethanol and water and M_1 , M_2 and M_3 are the molecular mass of the amino acid, ethanol and water.

3.3 Results

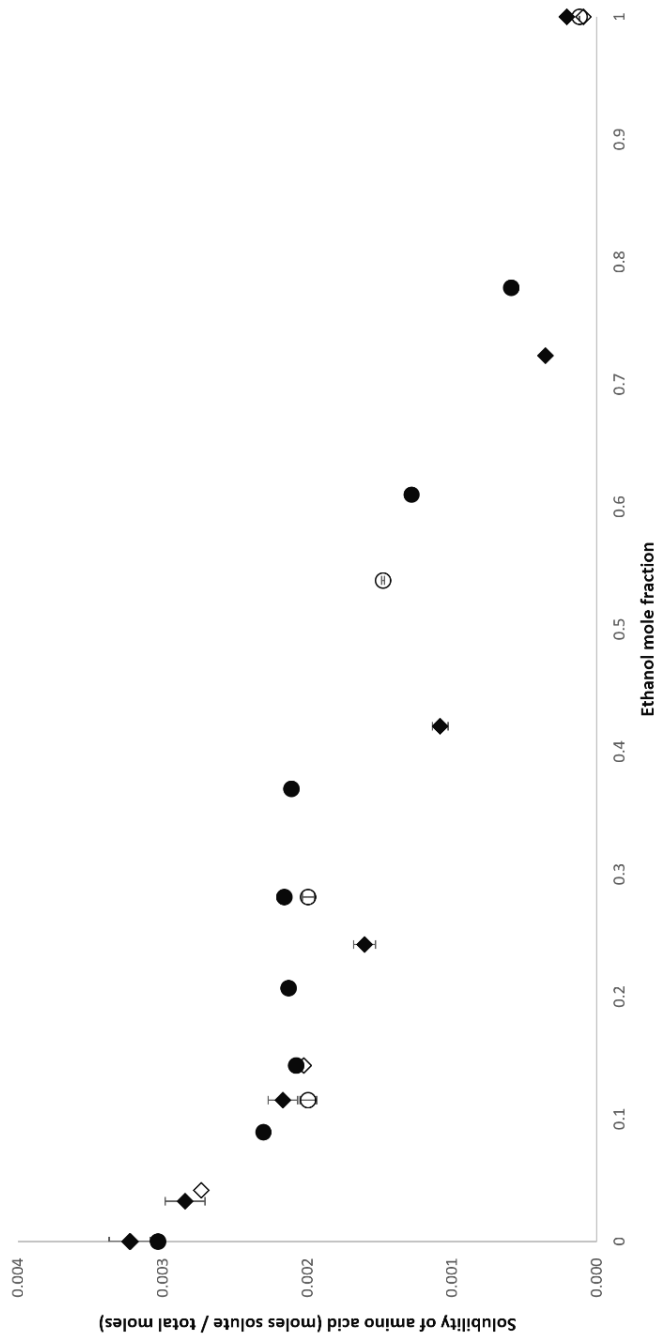
3.3.1 Review of amino acid solubility data and methodologies

All results from the analyses in this work are shown in Table 3.2. The method by which the samples were measure is also shown in Table 3.2.

Table 3.2: Experimental Mole Fraction Solubilities, x_1 , of Amino Acids, the Ethanol Solvent Mole Fraction, x_2 , the Method of Detection and the Relative Uncertainty at 298.15 K and $p = 0.1$ MPa

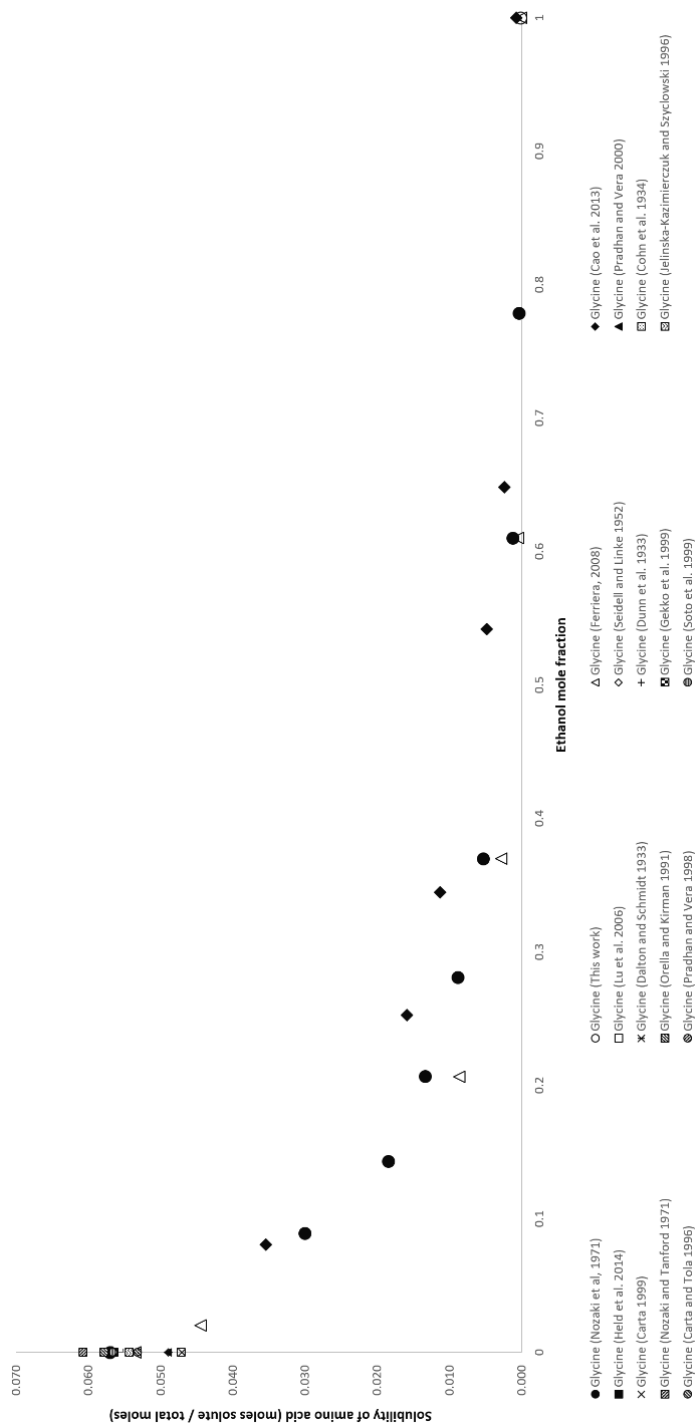
Amino Acid	Method	Amino Acid	Method	Amino Acid	Method									
L-alanine	UPLC	2.4976	0	0.05	L-glutamine	UPLC	0.0001	1	0.01	L-proline	Gravimetric	21.6583	0.1153	0.01
L-alanine	UPLC	1.3635	0.0891	0.07	L-histidine	UPLC	0.0004	1	0.04	L-proline	Gravimetric	15.5698	0.2811	0.03
L-alanine	UPLC	0.7857	0.2068	0.01	L-leucine	UPLC	0.0095	1	0	L-proline	Gravimetric	13.9836	0.5398	0.03
L-alanine	UPLC	0.4204	0.3697	0.01	L-arginine	Gravimetric	1.8614	0	0.02	L-proline	Gravimetric	0.5176	1	0.84
L-alanine	UPLC	0.1013	0.61	0.03	L-arginine	Gravimetric	1.3238	0.1153	0.04	L-tryptophan	Gravimetric	0.1034	0	0.18
L-alanine	UPLC	0.0282	1	0.01	L-arginine	Gravimetric	0.6303	0.2811	0.06	L-tryptophan	Gravimetric	0.1047	0.1153	0.02
L-glutamic acid	UPLC	0.1	0	0	L-arginine	Gravimetric	0.1292	0.5398	0.09	L-tryptophan	Gravimetric	0.1691	0.2811	0.03
L-glutamic acid	UPLC	0.0434	0.0891	0.03	L-arginine	Gravimetric	0.0143	1	0.11	L-tryptophan	Gravimetric	0.125	0.5398	0.02
L-glutamic acid	UPLC	0.0273	0.2068	0.08	L-lysine	Gravimetric	14.6733	0.1153	0	L-tryptophan	Gravimetric	0.0243	1	0.19
L-glutamic acid	UPLC	0.0225	0.3697	0.07	L-lysine	Gravimetric	5.0789	0.2811	0.05	L-tyrosine	Gravimetric	0.0035	0	0.19
L-glutamic acid	UPLC	0.0183	0.61	0.09	L-lysine	Gravimetric	2.4035	0.5398	0.01	L-tyrosine	Gravimetric	0.0031	0.1153	0.06
L-glutamic acid	UPLC	0.0009	1	0.02	L-lysine	Gravimetric	0.618	1	0.02	L-tyrosine	Gravimetric	0.0015	0.2811	0.45
L-pyroglutamic acid	UPLC	0.0033	0	0	L-methionine	Gravimetric	0.6768	0	0.03	L-tyrosine	Gravimetric	0.002	0.5398	0.26
L-pyroglutamic acid	UPLC	0.0019	0.0891	0.09	L-methionine	Gravimetric	0.3135	0.1153	0	L-tyrosine	Gravimetric	0.0031	1	1.41
L-pyroglutamic acid	UPLC	0.0024	0.2068	0.05	L-methionine	Gravimetric	0.2014	0.2811	0.01	L-cysteine	Gravimetric	1.2992	0	0
L-pyroglutamic acid	UPLC	0.0025	0.3697	0.06	L-methionine	Gravimetric	0.0859	0.5398	0.05	L-cysteine	Gravimetric	0.4652	0.1153	0.08
L-pyroglutamic acid	UPLC	0.0037	0.61	0.04	L-methionine	Gravimetric	0.0057	1	0.09	L-cysteine	Gravimetric	0.2546	0.2811	0
L-pyroglutamic acid	UPLC	0.0029	1	0.06	L-phenylalanine	Gravimetric	0.1994	0.1153	0.03	L-cysteine	Gravimetric	0.1318	0.5398	0.08
L-pyroglutamic acid	UPLC	0.0046	1	0	L-phenylalanine	Gravimetric	0.1994	0.2811	0.02	L-cysteine	Gravimetric	0.0113	1	0.21
Glycine	UPLC	0.0002	1	0.05	L-phenylalanine	Gravimetric	0.1478	0.5398	0.01					
L-asparagine	UPLC	0.0002	1	0.03	L-phenylalanine	Gravimetric	0.0118	1	0.02					
L-aspartic acid	UPLC	0.0002	1	0.03										

Solubility data of L-phenylalanine comes from four literature sources. The initial solubility measurement in water was conducted by Dalton and Schmidt (1935), then in water and water-ethanol mixtures by Needham (1970), Nozaki *et al.* (1971) and Lu *et al.* (2012). In this article, an additional set of data was collected in water-ethanol mixtures. All data was collected by the gravimetric method except for Dalton and Schmidt, who used the dissolution method. The data of Needham, Dalton and Schmidt, Nozaki *et al.* (2012) and the experimental data collected for this article are similar, as can be seen in Figure 3.1. After an initial decrease in solubility, as ethanol mole fraction increases from 0 to 0.100 mole fraction, the solubility of L-phenylalanine is greater between an ethanol mole fraction of 0.100 and 0.400 than below 0.100. The solubility of L-phenylalanine decreases again at an ethanol mole fraction above 0.400. The exception to this is Lu *et al.* (2012), who did not measure an increase in solubility between 0.100 and 0.400.



○ L-phenylalanine (This work) ● L-phenylalanine (Nozaki, 1971) ◇ L-phenylalanine (Dalton and Schmidt, 1935) ◇ L-phenylalanine (Needham, 1970) ◆ L-phenylalanine (Lu et al. 2012)

Figure 3.1 Solubility of L-phenylalanine (moles solute/mole solvent) in water, ethanol and ethanol water mixtures



In Figure 3.2, the concentration of glycine is plotted against the mole fraction of ethanol from 0.000 to 1.000. Several authors have measured the solubility of glycine in water. Cao *et al.* (2013), Ferreira (2008), Nozaki *et al.* (1971) and this work have measured the solubility of glycine in various binary solutions of ethanol and water and in ethanol. Solubilities reported by Nozaki *et al.* (1971) are higher than those of Ferreira (2008), while the solubilities measured by Cao *et al.* (2013) are the highest reported. At a solvent mole fraction of 1.000 ethanol, the solubility reported by Ferreira (2008), using the ninhydrin method were within the standard deviation measured by this work, using the UPLC method. These were $4.59 \cdot 10^{-5}$ and $5.52 \cdot 10^{-5}$ respectively. Cao *et al.* (2013), using the gravimetric method, reported a solubility mole fraction of 0.0007.

The solubility measured by Nozaki *et al.* (1971) and our own measurements of L-tryptophan are shown in Figure 3.3. Except for in water, at all mole fractions of ethanol, the solubility measured by Nozaki *et al.* (1971) was higher than the new data reported in this article. The solubility of L-tryptophan peaks between ethanol mole fractions of 0.281 and 0.540. The highest solubility of L-tryptophan was measured by Nozaki *et al.* at an ethanol mole fraction of 0.371.

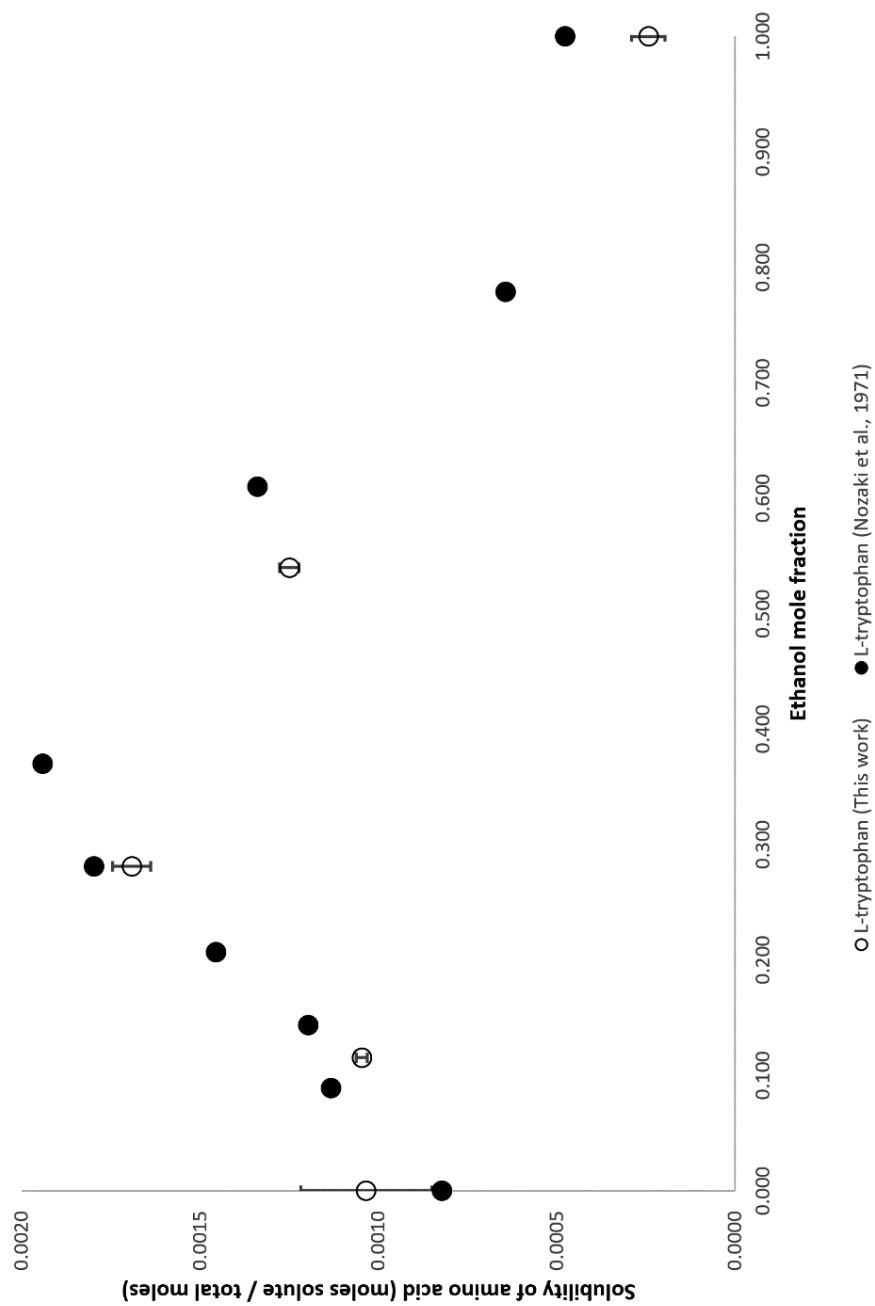


Figure 3.3 Solubility of L-tryptophan (moles solute/mole solvent) in water, ethanol and ethanol water mixtures

3.3.2 Impact of amino acid conversions on their solubilities

It has been shown that L-glutamic acid can form L-pyroglutamic acid in solution⁵³. Previous solubility studies did not take this into consideration. In Figure 3.4, the solubility of L-glutamic acid from Dunn and Ross, 1938, who used the gravimetric method of dissolution, is compared to the simultaneous measurement of L-glutamic acid and L-pyroglutamic acid in this work. Furthermore, the work of Mo *et al.* shows that the form of the crystal influences the solubility of the α -amino acid. The β -crystal forms follow the solubility data from Dunn and Ross, 1938, and this work closely. At pure water, the solubility reported by Dunn and Ross is approximately equal to the combined L-glutamic acid and L-pyroglutamic acid solubility collected experimentally in this work. At higher ethanol mole fraction, the difference between the data presented by Dunn and Ross and this work increases. The solubilities reported by Dunn and Ross fall below the standard deviation of those in this report at ethanol mole fractions above 0.370. Models on the solubility of amino acids in water and ethanol mixtures⁵³ show that the data generated in this work, shown in Figure 3.4, fit better than the data by Dunn and Ross, 1938. Note should be taken that the measurements of L-pyroglutamic acid in Figure 3.4 are not at maximum solubility.

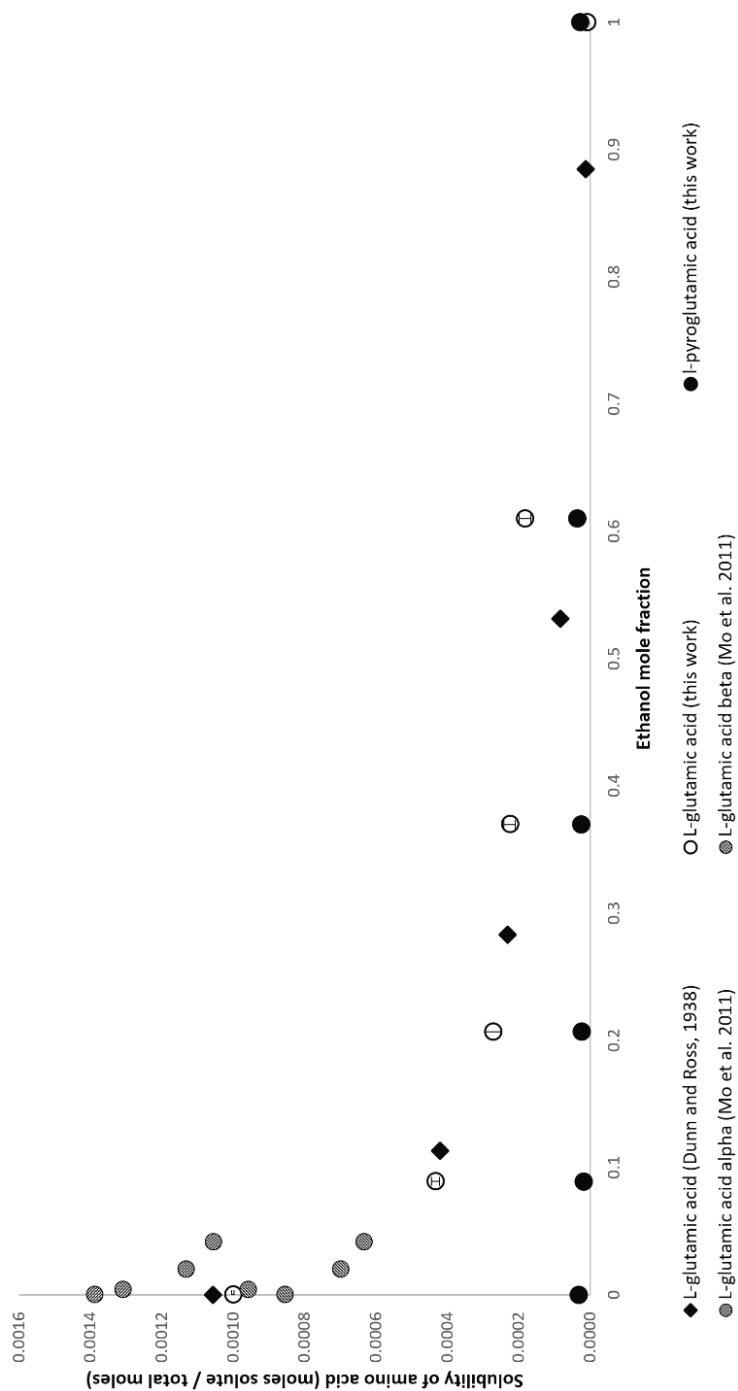


Figure 3.4 Solubility of L-glutamic acid and L-pyroglutamic acid (moles solute/mole solvent) in water, ethanol and ethanol water mixtures

While the L-cysteine trials were kept in a low oxygen environment, there is still the possibility that oxidation to L-cystine took place, which in turn could affect the solubility of L-cysteine. Therefore, the L-cysteine (molar mass = 121.16 g/mol) solubility trials were checked by mass spectrometer in that negligible amounts of L-cystine (molar mass = 240.3 g/mol) were formed. In order to accommodate for any build up on the detector of the mass spectrometer, the detector was cleaned before each measurement. In both measurements, only trace amounts of L-cystine were found, Figure S1, leading to the conclusion that the trace amounts of L-cystine do not affect the solubility data presented of L-cysteine in this work. The measured solubility data is presented in Figure 3.7 and discussed in the subsequent section.

3.3.3 Incomplete solubility data of amino acids

Previous work has published the solubility of DL-alanine but not L-alanine in various ethanol mole fractions. Furthermore, the reported solubilities of L-alanine in water vary widely. In Figure 3.5, the solubilities of the published DL-alanine, L-alanine and the newly measured solubilities of L-alanine are compared. In water, more DL-alanine dissolved than L-alanine for all reported data. It is unclear from the literature what the individual fractions of D-alanine and L-alanine are in the DL-alanine mixture. Measured as a mixture, the DL-alanine measurements are only slightly more soluble than L-alanine alone at 0.00 and 0.100 ethanol mole fraction. This gives evidence that the chiral forms have a negative impact on the other's solubility. At 0.200 ethanol mole fraction and higher, L-alanine has a higher solubility than the DL-alanine mixture. The solubility data of L-alanine in water, water-ethanol mixtures and ethanol measured for this work were measured using the UPLC.

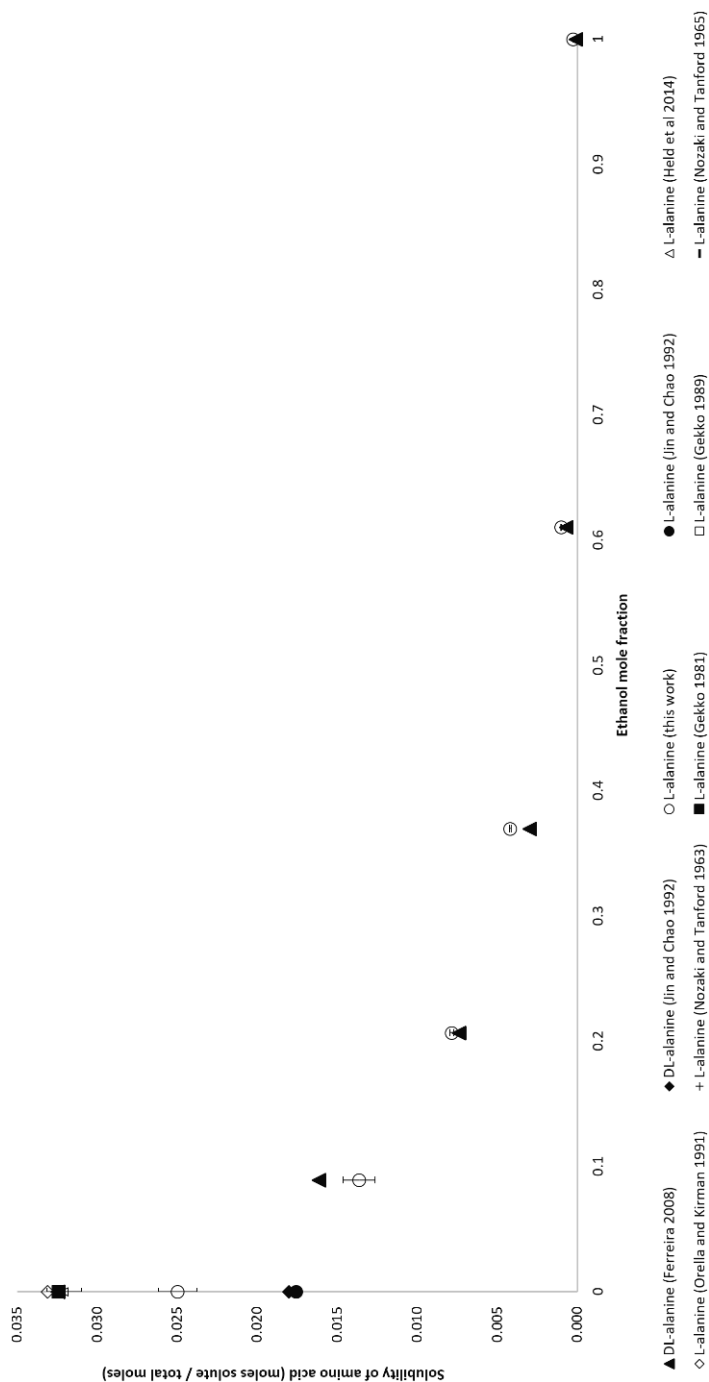


Figure 3.5 Solubility of L-alanine (moles solute/mole solvent) in water, ethanol and ethanol water mixtures

The results for L-proline and L-lysine are shown in Figure 3.6. Both α -amino acids have similar solubility ranges. Their solubilities are lower at higher ethanol mole fractions. The L-proline solubilities in water reported by Held *et al.* are similar to those reported by Tomiyama and Schmidt. The solubility of L-proline in water reported by El-Dossoki are much lower than the solubility of L-proline of reported by Amend and Hegelson. Furthermore, the solubility of L-proline in water reported by El-Dossoki is lower than the solubility of L-proline measured for this work at 0.100 ethanol mole fraction.

The solubilities of L-cysteine and L-arginine are shown in Figure 3.7. Also, both L-cysteine and L-arginine have similar solubility ranges, but the solubility of L-arginine decreases faster than L-cysteine as the ethanol mole fraction increases. The solubility of L-cysteine in water that was reported by El-Dossoki and El-Damarany is higher than the solubility measured for this work. El-Dossoki and El-Damarany do not report that their measurements were taken in a sealed, oxygen-poor environment. This could account for elevated experimentally measured solubility due to the formation of the dimer cystine.

The solubility of L-methionine is shown in Figure 3.8 and is the α -amino acid with the lowest solubility of the α -amino acids for which new data is being presented. The measurements of Zhang *et al.*, Sawamura and Kunimasa and this work are similar.

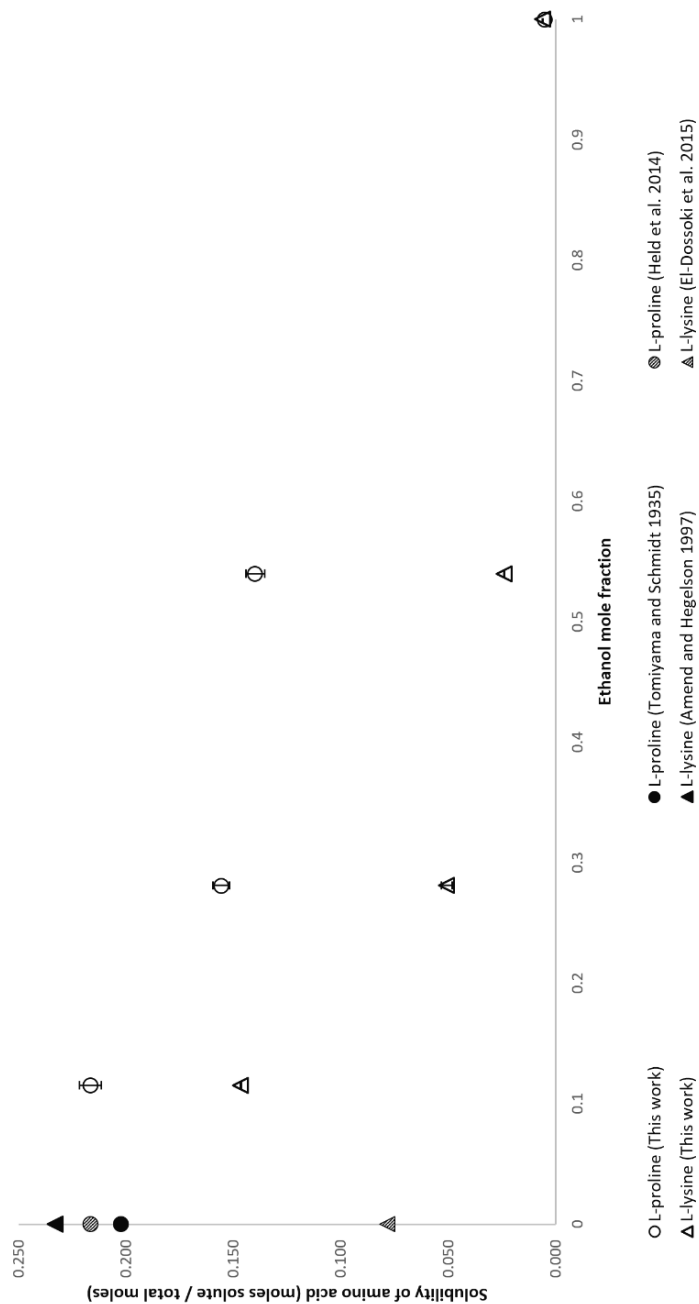


Figure 3.6 Solubility of L-proline and L-lysine (moles solute/mole solvent) in water, ethanol and ethanol water mixtures

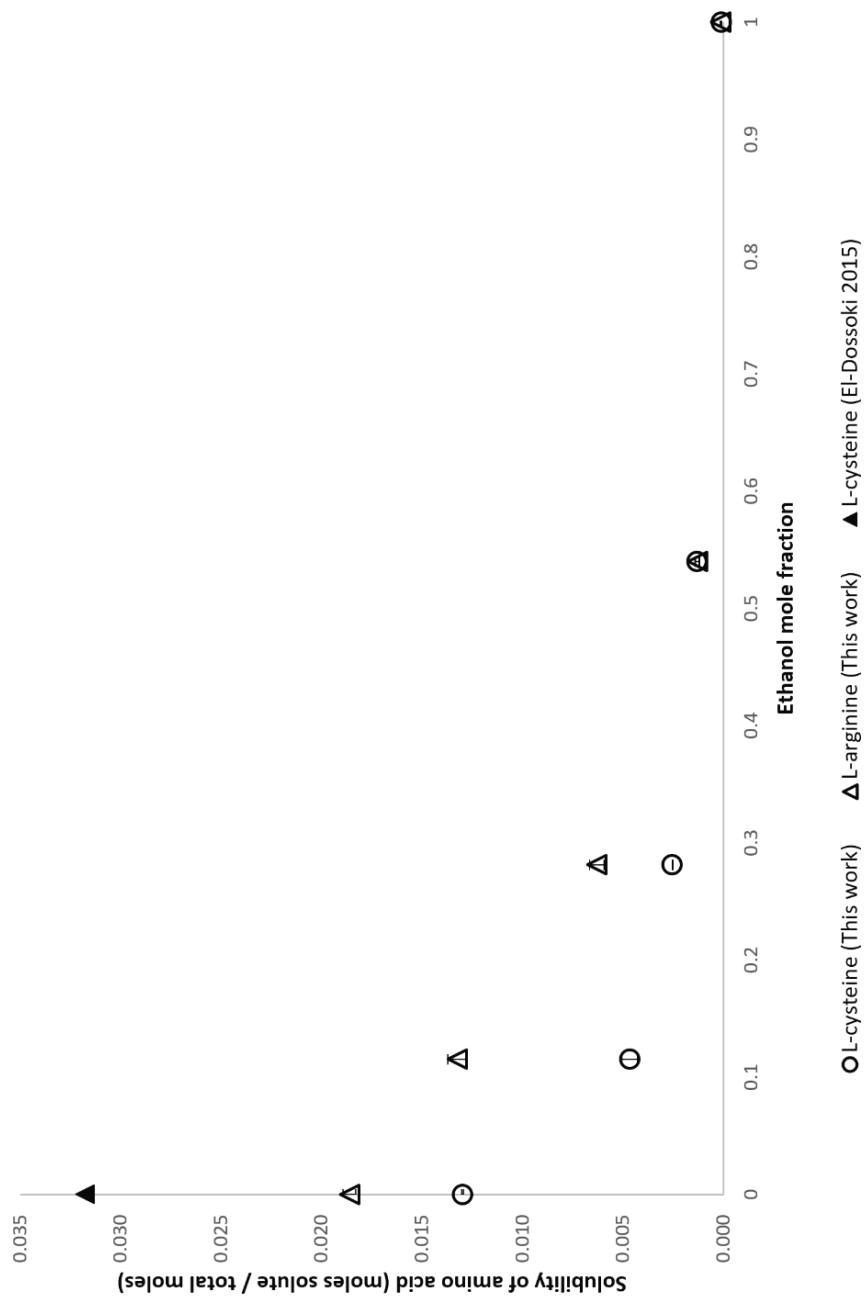
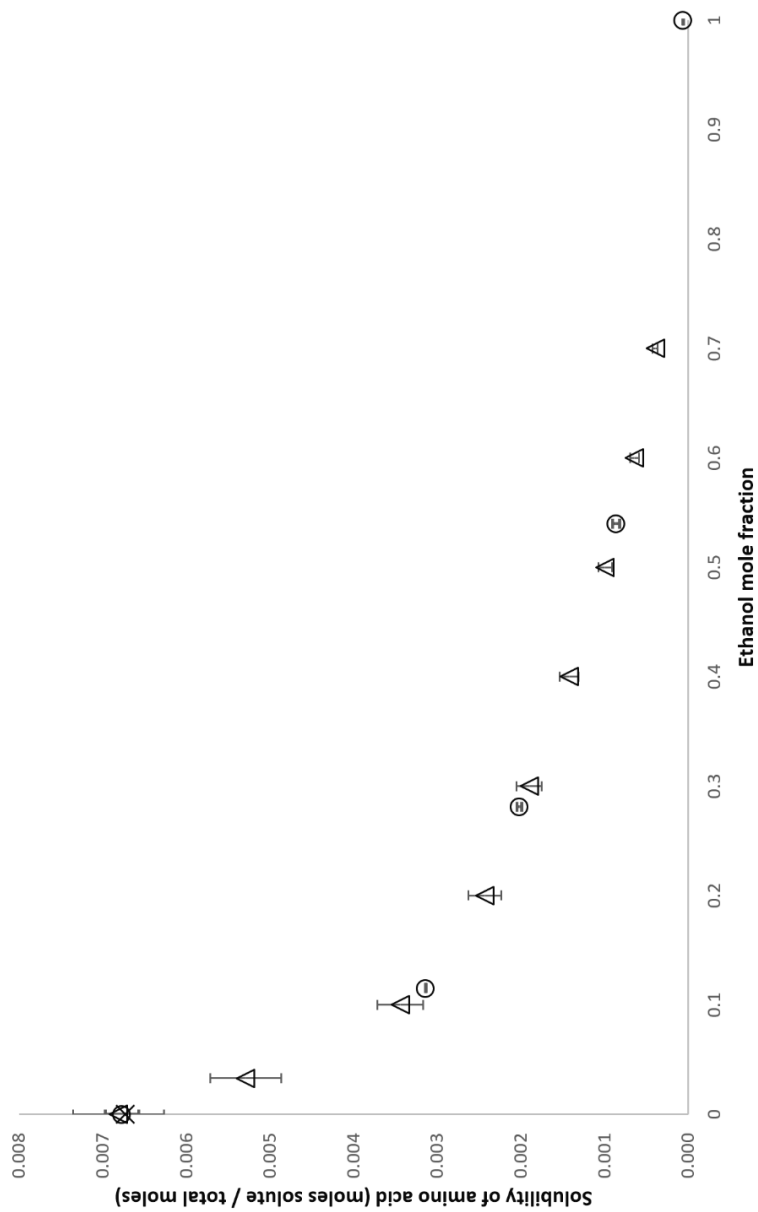


Figure3.7 Solubility of L-cysteine and L-arginine (moles solute/mole solvent) in water, ethanol and ethanol water mixtures



○ L-Methionine (This work) △ L-methionine (Zhang et al. 2016) X L-methionine (Sawamura and Kunimasa, 2014)

Figure 3.8 Solubility of L-methionine (moles solute/mole solvent) in water, ethanol and ethanol water mixtures

3.3 Discussion

The influence of ethanol on the solubility of amino acids is not the same for all amino acids. Most amino acids have a lower solubility when their solvent is at a higher ethanol mole fraction. All amino acids have a loss in solubility above a mole fraction of 0.5. The change in solubility is not the same for all amino acids in the range of 0-0.5 mole fraction ethanol. This difference between the amino acids is most pronounced at ethanol mole fractions around 0.2. The effect of ethanol on the solubility of amino acids can be characterized by the groups found in their side chains.

Five amino acids have a ring in the side chain. These amino acids are L-tryptophan, L-tyrosine, L-proline, L-phenylalanine and L-histidine. These rings include either phenyl, pyrrolidine or imidazole. The amino acids with rings in the side chains had the least decrease in solubility as ethanol is added. The average decrease in solubility of these amino acids at an ethanol fraction of 0.2 was only -3.8%. In the case of L-tryptophan, the solubility was even increased by 105.6%. We hypothesize that the rings of these amino acids are ethanolphilic, while the amino and carboxylic groups on these amino acids are ethanolphobic. Moderate ethanol fractions between 0.2 and 0.4 increase the solubility of these amino acids. The water and ethanol molecules arrange themselves at the respective groups of the molecule, creating a lattice around the amino acids. Higher ethanol mole fractions lower the solubility of these amino acids, because the ethanol molecules surround the amino acid molecule and disrupt the water molecules surrounding the amino and carboxylic groups on the amino acid molecule.

The aliphatic amino acids, L-phenylalanine, L-isoleucine, L-leucine, L-alanine, L-methionine and L-valine, show initially a low to medium decrease in solubility at an ethanol mole fraction of approximately 0.2. The decreases range from -31.0% to -71.6%. The aliphatic amino acids showed an average decrease of -54.8% solubility. L-Phenylalanine is both aliphatic and phenylic and shows a lower decrease in solubility, -33.9%, than the average for the aliphatic group in this range. This decrease could be possibly mitigated by the phenyl ring.

The hydroxyl containing amino acids, L-tyrosine, L-serine and L-threonine, show a medium decrease in solubility. Together, L-serine and L-threonine have an average decrease of -68.5% at ethanol mole fraction levels around 0.2. L-Tyrosine, which is both hydroxylic and

phenylic, has only a low decrease in solubility -1.3%. Here, as in the case of L-phenylalanine, the decrease in solubility is mitigated by the phenyl ring.

The amide containing amino acids, L-glutamine and L-asparagine, show a high decrease in solubility of -75.1 and -77.4% respectively at ethanol mole fraction around 0.2. The average for amide containing amino acids increases slightly to -72.9% when L-arginine is added to this group. L-Arginine contains both an amide group and is positively charged, is added to this group.

A high decrease in solubility is seen in the charged amino acids L-glutamic acid, L-aspartic acid and L-lysine. The average decrease at an ethanol mole fraction of 0.2 was -78.1%. The lower decrease in solubility of the charged amino acids L-histidine (-66.1%) and L-arginine (-66.1%) seemed to be mitigated by either their ring containing side chain, imidazole, or amide respectively.

L-Cysteine is the only amino acid that contains sulphur. It has the highest decrease in solubility at ethanol mole fraction of 0.2. The decrease was -80.4%. L-Methionine also contains a sulphur molecule but is normally considered aliphatic. L-Methionine solubility decrease is -70.2%.

Glycine, containing no side chain, had the largest decrease in solubility. Glycine solubility decrease by -83.9% at ethanol mole fraction around 0.2.

3.4 Conclusion

For most α -amino acids, the gravimetric, ninhydrin and UPLC methods produced similar solubility data. Exceptions to this are α -amino acids that convert to other forms (L-glutamic acid and L-cysteine) and α -amino acids at very low solute concentrations (e.g. in pure ethanol).

The two α -amino acids that are the exceptions in the previous paragraph are L-glutamic acid, which has been shown to convert to L-pyroglutamic acid, and L-cysteine, which can convert to the dimer cystine. Therefore, all amino acids that have a possibility to convert to other amino acids in solution should be analysed by a technique that takes this into account. The UPLC technique used in this work shows reliable results.

At low concentrations (e.g. α -amino acids in pure ethanol), using gravimetric analytical techniques to measure amino acid solubility is not always reliable. However, the UPLC method used in this work was reliable at low concentrations. The data produced by the UPLC were also within the variation of the data published by Ferreira, 2008. Ferreira used the ninhydrin method

of analysis. While the ninhydrin method produced data consistent to data in this work for several α -amino acids, it is not able to detect and differentiate multiple amino acids in solution, like the UPLC method.

Most data points of several amino acids by Nozaki *et al.* (1971) and the data for Cao *et al.* for glycine were higher when compared to this work and the work of Ferreira (2008), Dalton and Schmidt (1933) and Needham (1970). A possible explanation for these results includes, but is not limited to, the samples being measured when the solutions were over-saturated or when dissolved from a crystal of another shape (e.g. α -crystal versus β -crystal) as shown by Mo *et al.*

This work gives a new more complete look at the solubility of all 20 proteinogenic α -amino acids. The new data published in this work doubles the peer-reviewed data on α -amino acid solubility in water, water/ethanol and ethanol systems. This data is the first published data in ethanol and ethanol/water systems for L-alanine, L-proline, L-arginine, L-cysteine and L-lysine solubility. Furthermore, this work gives the first data for the solubility of L-asparagine, L-glutamine, L-histidine and L-leucine in pure ethanol.

Lastly, the side chain of an amino acid has an effect on the solubility of that amino acid when ethanol is added. This is shown at ethanol mole fractions around 0.2. Side chains containing rings show the least decrease in solubility when water is replaced by a water-ethanol mixture due to the ethanolphilic properties of these rings. This is followed in descending order by the aliphatic amino acids, hydroxyl containing amino acids, amide containing amino acids, charged amino acids, sulphur containing amino acids and the amino acid with no side chain. Amino acids with side chains of two characteristics, such as L-tyrosine, which is both phenylic and containing a hydroxyl group, show a decrease in solubility in between both of their groups.

References

- [1] Scott, E.; Peter, F.; Sanders, J., Biomass in the Manufacture of Industrial Products – The use of Proteins and Amino Acids. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 751-762.
- [2] Lammens, T. M.; Franssen, M. C. R.; Scott, E. L.; Sanders, J. P. M., Availability of protein-derived amino acids as feedstock for the production of bio-based chemicals. *Biomass Bioenergy* **2012**, *44*, 168-181.
- [3] Lammens, T. M., Bio-based Industrial Chemicals from Glutamic Acid. **2011** PhD Thesis, Wageningen University Press, Wageningen, the Netherlands
- [4] Lammens, T. M.; Potting, J.; Sanders, J. P. M.; De Boer, I. J., Environmental comparison of biobased chemicals from glutamic acid with their petrochemical equivalents. *Environ. Sci. Technol.* **2011**, *45*, 8521-8.
- [5] Jin, X. Z.; Chao, K.-C. Solubility of four amino acids in water and of four pairs of amino acids in their water solutions *J. Chem. Eng. Data*, **1992**, *37*, 199-203
- [6] Matsuo, H.; Suzuki, Y.; Sawamura, S. Solubility of -amino acids in water under high pressure: glycine, -alanine, -valine, -leucine, and - isoleucine *Fluid Phase Equilib.*, **2002**, *200*, 227-237
- [7] Held, C.; Reschke, T.; Muller, R.; Kunz, W.; Sadowski, G. Measuring and modeling aqueous electrolyte/amino-acid solutions with ePC-SAFT *J. Chem. Thermodyn.*, **2014**, *68*, 1-12
- [8] El-Dossoki, F.; El-Damarany, M. Solvation of Basic and Neutral Amino Acids in Aqueous Electrolytic Solutions: Measurements and Modeling *J. Chem. Eng. Data*, **2015**, *60*, 2989-2999
- [9] Sawamura, S.; Kunimasa, N. High-Pressure Solubility of L-Methionine in Water *J. Solution Chem.*, **2014**, *43*, 1810-1815
- [10] Lu, J.; Lin, Q.; Rohani, S. Solubility of L-Phenylalanine Anhydrous and Monohydrate Forms: Experimental Measurements and Predictions *J. Chem. Eng. Data*, **2012**, *57*, 1492-1498
- [11] Lu, J.; Wang, X.-J.; Yang, X.; Ching, C.-B. Solubilities of Glycine and Its Oligopeptides in Aqueous Solutions *J. Chem. Eng. Data*, **2006**, *51*, 1593-1596
- [12] Mo, Y.; Dang, L.; Wei, H. Solubility of alpha-form and beta-form of L-glutamic acid in different aqueous solvent mixtures *Fluid Phase Equilib.*, **2011**, *300*, 105-109

- [13] Zhou, X.; Fan, J.; Li, N.; Du, Z.; Ying, H.; Wu, J.; Xiong, J.; Bai, J. Solubility of L-phenylalanine in water and different binary mixtures from 288.15 to 318.15 K *Fluid Phase Equilib.*, **2012**, 316, 26-33
- [14] Seidell, A.; Linke, W. F. Solubility of Inorganic and Organic Compounds, 3rd ed.; Van Nostrand Co.: New York, **1952**
- [15] Pradhan, A. A.; Vera, J. H. Effect of acids and bases on the solubility of amino acids. *Fluid Phase Equilib.* **1998**, 152, 121-132
- [16] Carta, R.; Tola, G. Solubilities of L-cystine, L-tyrosine, L-leucine, and glycine in aqueous solutions at various pHs and NaCl concentrations. *J. Chem. Eng. Data* **1996**, 41, 414-417
- [17] Orella, C.J.; Kirman, D.J. Correlation of amino acid solubilities in aqueous aliphatic alcohol solutions *Ind. Eng. Chem. Res.* **1991**, 30, 1040–1045
- [18] Gekko, K. Mechanism of polyol-induced protein stabilization: solubility of amino acids and diglycine in aqueous polyol solutions *J. Biochem.* **1981**, 90, 1633–1641
- [19] Gekko, K.; Idota, Y. Acid Solubility and Protein Stability in Aqueous malitol Solutions *Agric. Biol. Chem.* **1989**, 53, 89–90
- [20] Nozaki, Y.; Tanford, C.T. The Solubility of Amino Acids and Related Compounds in Aqueous Urea Solutions *J. Biol. Chem.* **1963**, 238, 4074–4081
- [21] Nozaki, Y.; Tanford, C.T. The solubility of amino acids and related compounds in aqueous ethylene glycol solutions *J. Biol. Chem.* **1965**, 240, 3568–3573
- [22] Dalton, J.B.; Schmidt, C.L.A. The solubilities of certain amino acids in water, the densities of their solutions at twenty-five degrees, and the calculated heats of solution and partial molal volumes *J. Biol. Chem.* **1933**, 103, 549–578
- [23] Dunn, M.S.; Ross, F.J.; Read, L.S. The solubility of the amino acids in water *J. Biol. Chem.* **1933**, 103, 579–595
- [24] Cohn, E.J.; McMeekin, T.J.; Edsall, J.T.; Weare, J.H. Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. II. The Solubility of α -Amino Acids in Water and in Alcohol—Water Mixtures *J. Am. Chem. Soc.* **1934**, 56, 2270–2282
- [25] Gekko, K.; Ohmae, E.; Kameyama, K.; Takagi, T. Acetonitrile-protein interactions: amino acid solubility and preferential solvation *Biochim. Biophys. Acta* **1998**, 1387, 195–205

- [26] Jelińska-Kazimierczuk, M.; Szyclowski, J. Solubility of α -amino acids in water under high pressure: glycine, l-alanine, l-valine, l-leucine, and l-isoleucine *J. Sol. Chem.* **1996**, *25*, 1175–1184.
- [27] Carta, R.; Tola, G. Solubilities of l-Cystine, l-Tyrosine, l-Leucine, and Glycine in Aqueous Solutions at Various pHs and NaCl Concentrations *J. Chem. Eng. Data* **1996**, *41*, 414–417
- [28] Soto, A.; Arce, A.; Khoshkbarchi, M.K.; Vera, J.H. Measurements and modelling of the solubility of a mixture of two amino acids in aqueous solutions *Fluid Phase Equilib.* **1999**, *158–160*, 893–901
- [29] Hlynka, I. Thesis, California Institute of Technology, Pasadena, California, **1939**
- [30] Stoddard, M.P.; Dunn, M.S. Quantitative investigations of amino acids and peptides *J. Biol. Chem.* **1942**, *142*, 329–343
- [31] Carta, R. Solubilities of l-Cystine, l-Tyrosine, l-Leucine, and Glycine in Their Water Solutions *J. Chem. Thermodynamics* **30** (1998) 379–387
- [32] McMeekin, T. L.; Cohn, E. J.; Weare, J. H., Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. III. The Solubility of Derivatives of the Amino Acids in Alcohol—Water Mixtures. *J. Am. Chem. Soc.* **1935**, *57*, 626–633.
- [33] Dunn, M. S.; Ross, F. J., Quantitative Investigations of Amino Acids and Peptides: IV. The Solubilities of the Amino Acids in Water-Ethanol Mixtures. *J. Biol. Chem.* **1938**, *125*, 309–332.
- [34] Needham, T. E.; Paruta, A. N.; Gerraugh, R.J., Solubility of Amino Acids in Pure Solvent Systems. *J. Pharm. Sci.* **1971**, *60*, 565.
- [35] Nozaki, Y.; Tanford, C., The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions: Establishment of a hydrophobicity scale. *J. Biol. Chem.* **1971**, *246*, 2211–2217.
- [36] Conio, G.; Curletto, L.; Patrone, E., Temperature coefficient of solubility of some cyclic peptides in water-ethanol mixtures. *J. Biol. Chem.* **1973**, *248*, 5448–5450.
- [37] Cao, Z.; Hu, Y.; Li, J.; Kai, Y.; Yang, W. Solubility of glycine in binary system of ethanol + water solvent mixtures: Experimental data and thermodynamic modeling *Fluid Phase Equilib.*, **2013**, *360*, 156–160

- [38] Zhang, T.; Li, Z.; Wang, Y.; Li, C.; Yu, B.; Zheng, X.; Jiang, L.; Gong, J. Determination and correlation of solubility and thermodynamic properties of L-methionine in binary solvents of water + (methanol, ethanol, acetone) *J. Chem. Thermodyn.*, **2016**, 96, 82-92
- [39] Yu, Q.; Ma, X.; Xu, L. Solubility, dissolution enthalpy and entropy of L-Glutamine in mixed solvents of ethanol+water and acetone+water *Thermochim. Acta*, **2013**, 558, 6-9
- [40] Liu, Y.; Wang, Y.; Liu, Y.; Xu, S.; Chen, M.; Du, S.; Gong, J. Solubility of L-histidine in different aqueous binary solvent mixtures from 283.15 K to 318.15 K with experimental measurement and thermodynamic modelling *J. Chem. Thermodyn.*, 2017, 105, 1-14
- [41] Gude, M. T.; Meuwissen, H. H. J.; van der Wielen, L. A. M.; Luyben, K. C. A. M., Partition Coefficients and Solubilities of α -Amino Acids in Aqueous 1-Butanol Solutions. *Ind. Eng. Chem. Res.* **1996**, 35, 4700-4712.
- [42] vanBerlo, M.; Gude, M. T.; vanderWielen, L. A. M.; Luyben, K., Partition coefficients and solubilities of glycine in the ternary solvent system 1-butanol plus ethanol plus water. *Ind. Eng. Chem. Res.* **1997**, 36, 2474-2482.
- [43] Gude, M. T.; van der Wielen, L. A. M.; Luyben, K. C. A. M., Phase behavior of α -amino acids in multicomponent aqueous alkanol solutions. *Fluid Phase Equilib.* **1996**, 116, 110-117.
- [44] Widyarani; Bowden, N. A.; Sanders, J. P. M.; Kolfschoten, R. C.; Bruins, M. E., Fractional precipitation of amino acids from agro-industrial residues using ethanol. *Ind. Eng. Chem. Res.* **2016**, 55, 7462-7472.
- [45] Ferreira, L. A.; Macedo, E. A.; Pinho, S. P., Solubility of amino acids and diglycine in aqueous-alkanol solutions. *Chem. Eng. Sci.* **2004**, 59, 3117-3124.
- [46] Ferreira, L. A.; Pinho, S. P.; Macedo, E. A., Solubility of L-serine, L-threonine and L-isoleucine in aqueous aliphatic alcohol solutions. *Fluid Phase Equilib.* **2008**, 270, 1-9.
- [47] Fuchs, D.; Fischer, J.; Tumakaka, F.; Sadowski, G., Solubility of amino acids: influence of the pH value and the addition of alcoholic cosolvents on aqueous solubility. *Ind. Eng. Chem. Res.*, **2006**, 45, 6578-6584.
- [48] Grosse Daldrup, J. B.; Held, C.; Ruether, F.; Schembecker, G.; Sadowski, G., Measurement and Modeling Solubility of Aqueous Multisolute Amino-Acid Solutions. *Ind. Eng. Chem. Res.*, **2010**, 49, 1395-1401.

- [49] Held, C.; Sadowski, G.; Carneiro, A.; Rodriguez, O.; Macedo, E. A., Modeling thermodynamic properties of aqueous single-solute and multi-solute sugar solutions with PC-SAFT. *AICHE J.*, **2013**, 59, 4794-4805.
- [50] Ji, P.; Zou, J.; Feng, W., Effect of alcohol on the solubility of amino acid in water. *J. Mol. Catal. B: Enzym.* **2009**, 56, 185-188.
- [51] Zhou, X.; Fan, J.; Li, N.; Du, Z.; Ying, H.; Wu, J.; Xiong, J.; Bai, J., Solubility of L-phenylalanine in water and different binary mixtures from 288.15 to 318.15 K. *Fluid Phase Equilib.* **2012**, 316, 26-33.
- [52] Bowden, N. A.; Méndez Seviliano, D.; Sanders, J. P. M.; Bruins, M. E., Effects of ethanol on the solubility of the proteinogenic amino acids validated with the NRTL, Gude and Jouyban-Acree models. *Fluid Phase Equilib.*, submitted for publication.
- [53] Meussen, B.; van Zeeland, A. T.; Bruins, M. E.; Sanders, J. P. M., A Fast and Accurate UPLC Method for Analysis of Proteinogenic Amino Acids. *Food Anal. Methods* **2014**, 7, 1047-1055.
- [54] Edward C. Kendall, F. F. N., Reversible Oxidation-Reduction Systems of Cysteine-Cystine and Reduced and Oxidized Glutathione. *J. Biol. Chem.* **1926**, 69, 295-337.
- [55] Kumar, A.; Bachhawat, A. K., Pyroglutamic acid: throwing light on a lightly studied metabolite. *Curr. Sci.* **2012**, 102, 288-297.
- [57] Haitinger, L., Vorlauge Mittheilung uber Glutaminsaure und Pyrrol. *Monatsh. Chem.* **1882**, 3, 228-229.

Chapter 4

The effects of ethanol on the solubility of the proteinogenic α -amino acids in mixtures

In preparation as: Bowden, N.A., Sanders, P.M. and Bruins, M.E. The effects of ethanol on the solubility of the proteinogenic α -amino acids

Abstract:

Research on the effects of organic solvents on the solubility of α -amino acids is incomplete. In this study, the solubilities of each of the 20 proteinogenic α -amino acids in a mixture of all the 20 proteinogenic α -amino acids in 0%, 15%, 30%, 50%, 70% and 80% (g/g) ethanol-water solutions and at 277.15, 297.15, 315.15 and 335.15 K are reported and compared to literature values of single α -amino acids. The solubility of the individual α -amino acids in water, ethanol and water-ethanol mixtures are discussed in chapter 3. The results of this research show that amino acids can be grouped according to the structure of their side chains. When branched aliphatic, hydroxylic, phenylic and carboxylic groups are on the side chain of an amino acid, then the solubility of that amino acid will increase when in a mixture of 20 amino acids between 30-70% (g/g) ethanol-water solution. If Sulphur containing and amine/amide containing groups are on the side chain of an amino acid, then the solubility of that amino acid will decrease in a mixture of 20 amino acids between 30-70% (g/g) ethanol-water solution compared to the solubility as a single amino acid.

4.1 Introduction

Bio-based products can potentially be made from proteins and α -amino acids (e.g. bioplastics, pharmaceuticals, fine chemicals) [1-3]. One of the challenges in this line of research, is to find a way to separate α -amino acids from industrial residues. However, there is a dearth of data on the molecular interactions between the α -amino acids in solution and even less on the use of anti-solvents. This understanding of the interactive forces and anti-solvents is integral in designing technologies for separating α -amino acids from residue streams.

In this work, the solubility of each of the 20 α -amino acids in solutions of mixtures of all α -amino acids in an ethanol/water system and at various temperatures and at a constant pH are derived experimentally. The solubility for each of these α -amino acid is compared to the literature values for each of these individual α -amino acids in single solution.

Research on the solubility of amino acids has proceeded over several decades. The initial work on the solubility of α -amino acids in water at various temperatures showed the differences in hydropathy between the amino acids [4-10]. Based on this research, hydropathy scales were created to distinguish the relative hydropathy and hydrophobicity of each individual α -amino acid [11-15].

The pH of a solvent is a parameter that influences the solubility of an amino acid. All α -amino acids are composed of an amine group and a carboxyl group attached to a carbon atom, to which a side chain may be attached. Therefore, they may exist as either positively charged, when the pH of the solvent is lower than their isoelectric point, or negatively charged when the pH of the solvent is above their isoelectric point. When at their isoelectric point, they exist as a zwitterion and have their lowest solubility. Charged amino acids have an increased solubility [16, 17].

The pH of the isoelectric points for the α -amino acids vary between 3.0 and 11. For 15 of the 20 α -amino acids in this study, the average pH of the isoelectric point is 6.0 ± 0.6 . Two amino acids have an isoelectric point at pH = 3.0, one amino acid has an isoelectric point at pH = 10 and two amino acids have an isoelectric point at pH = 11 [18]. While the pH of the system will influence the solubility of the amino acids, in this research, we do not vary the pH of the solvent.

There has been some research on the solubility of α -amino acids in mixtures of alcohol and water. The first basic solubility measurements reported these solubilities [7; 19-20].

Subsequent research focused on calculating the partition coefficients of the solubility of these α -amino acids and their phase behavior [21-23].

While α -amino acids exist as mixtures in residue streams, research on the impact of the interactions of α -amino acids in mixtures in water/ethanol systems is lacking. In binary systems of α -amino acids, the interactions between two α -amino acids have been studied. In these studies, large variations in solubility have been seen in L-cystine and L-tyrosine [24, 25] as well as decreasing solubility with L-isoleucine in the presence of L-alanine [26]. Furthermore, L-leucine was only slightly affected by the addition of L-valine, while the solubility of L-valine was increased significantly in the presence of L-leucine [27]. Meanwhile, some work has also been done on the thermochemical properties of mixtures of α -amino acids [28].

Studies on the interactions between α -amino acids in an ethanol/water system are lacking in peer reviewed journals. This work seeks to begin to fill the gap.

4.2 Material and Methods

The α -amino acids that were used in this article were purchased from Sigma-Aldrich. These α -amino acids were at least 97% pure (mole fraction). Table 4.1 lists the supplier and purity of the chemicals used in this work.

3.0 Grams of each amino acid was put into 100 ml of demineralized water in a 250 ml glass Schott bottle. This was done again for a duplicate. The solutions were stirred and the pH of both mixtures was measured to be 5.3.

Then, 3.0 grams of each amino acid was put into 100 ml of 15%, 30%, 50%, 70% and 80% (g/g) ethanol-water in separate 250 ml glass Schott bottle. These were all done again in duplicate. The solutions were brought to $\text{pH } 5.5 \pm 0.2$ by 5M HCl addition. The pH was measured with a Sigma-Aldrich pH strip with a resolution of 0.5 pH unit.

The sealed Schott bottles were mixed and added to a jacketed water bath on stir plates and set to 277.15 K and 0.1 MPa and left to mix continuously at 80 rpm. Both the samples and the water in the water bath were continuously monitored with temperature probes.

The amino acids were allowed to mix until the maximum solubility of the amino acids had reached equilibrium. Equilibrium was determined as follows. Duplicate samples were taken daily, filtered and the total dry weight was determined. The amino acids were said to have

reached equilibrium when successive dry weights, 24-hours apart, yielded a concentration within the variation of the balance. All measurements were performed in duplicate.

Table 4.1: Description of chemicals and solvents used. All chemicals were ordered from Sigma-Aldrich with purities of at least 0.97 mole fraction.

Chemical Name	Source	Mole Fraction Purity	Purification Method
glycine	Sigma-Aldrich	1.00	None
L-alanine	Sigma-Aldrich	0.98	None
L-arginine	Sigma-Aldrich	0.98	None
L-asparagine	Sigma-Aldrich	1.00	None
L-aspartic acid	Sigma-Aldrich	0.99	None
L-cysteine	Sigma-Aldrich	0.97	None
L-glutamic acid	Sigma-Aldrich	0.99	None
L-glutamine	Sigma-Aldrich	0.99	None
L-histidine	Sigma-Aldrich	0.99	None
L-leucine	Sigma-Aldrich	1.00	None
L-lysine	Sigma-Aldrich	0.97	None
L-methionine	Sigma-Aldrich	0.98	None
L-phenylalanine	Sigma-Aldrich	0.98	None
L-proline	Sigma-Aldrich	0.99	None
L-tryptophan	Sigma-Aldrich	0.98	None
L-tyrosine	Sigma-Aldrich	0.98	None
ethanol	Sigma-Aldrich	> 0.99	None

In parallel, liquid samples were filtered and diluted for amino acid profile analysis. Concentrations of each amino acid after dilution ranged from 0.0100 to 2.00 mM and with a total amino acid concentration of not more than 10.00 mM per sample. The samples were analysed in duplicate by the UPLC method [43]. The UPLC method is based on automated pre-column derivatisation in the injection needle of the amino acids in an autosampler using o-phthalaldehyde (OPA) reagent in combination with 9-fluorenylmethyl chloroformate (FMOC) that enables the amino acids to fluoresce. Separation was achieved with a Dionex RSLC system using an Acquity UPLC BEH C18 reversed-phase column. Sample analysis was performed with

an UltiMate 3000 Rapid Separation pump and autosampler. Derivatised amino acids were detected at 263 nm (FMOC derivative of L-proline) and 338 nm (OPA derivatives of the other amino acids).

The experiment was performed again using the methods stated above at temperatures 293.15, 313.15 and 323.15 K. Both the experiments and the analyses were done in duplicate.

4.3 Results and Discussion

The results of the experiments have been divided according to the structure of the side chains of the amino acids. These amino acid groups are explained in chapter 1.

4.3.1 Unbranched aliphatic: Glycine, L-Alanine

Glycine is the only non-enantiomeric amino acid because it does not have a side chain. As shown in Figure 4.1, the solubility of glycine is the same in a mixture of 20 amino acids as in single solution in an ethanol-water system. The reason for this could be that glycine does not specifically interact with other amino acids since it does not have a side chain.

L-Alanine is similar to glycine but has a side chain of one methyl group. In Figure 4.2, the solubility of L-alanine is higher between at 40%, 50% and 70% (g/g) ethanol-water solutions when mixed with other amino acids relative to being in solution without any other amino acids present. This shows that L-alanine may interact with other amino acids with its methyl side chain, which increases its solubility slightly.

However, the comparison in Figure 4. shows that glycine is still relatively more soluble than L-alanine. The hypothesis is that while the methyl group side chain of L-alanine can form hydrogen bonds with other amino acids, it less readily forms hydrogen bonds with water. At temperatures where more than three data points could be recorded, the effect of an incremental increase in ethanol concentration decreases for L-alanine, showing a trend line that is concave upwards, in a mixture of amino acids. This is the same for glycine. This trend line is also concave upwards for both amino acids in single solution.

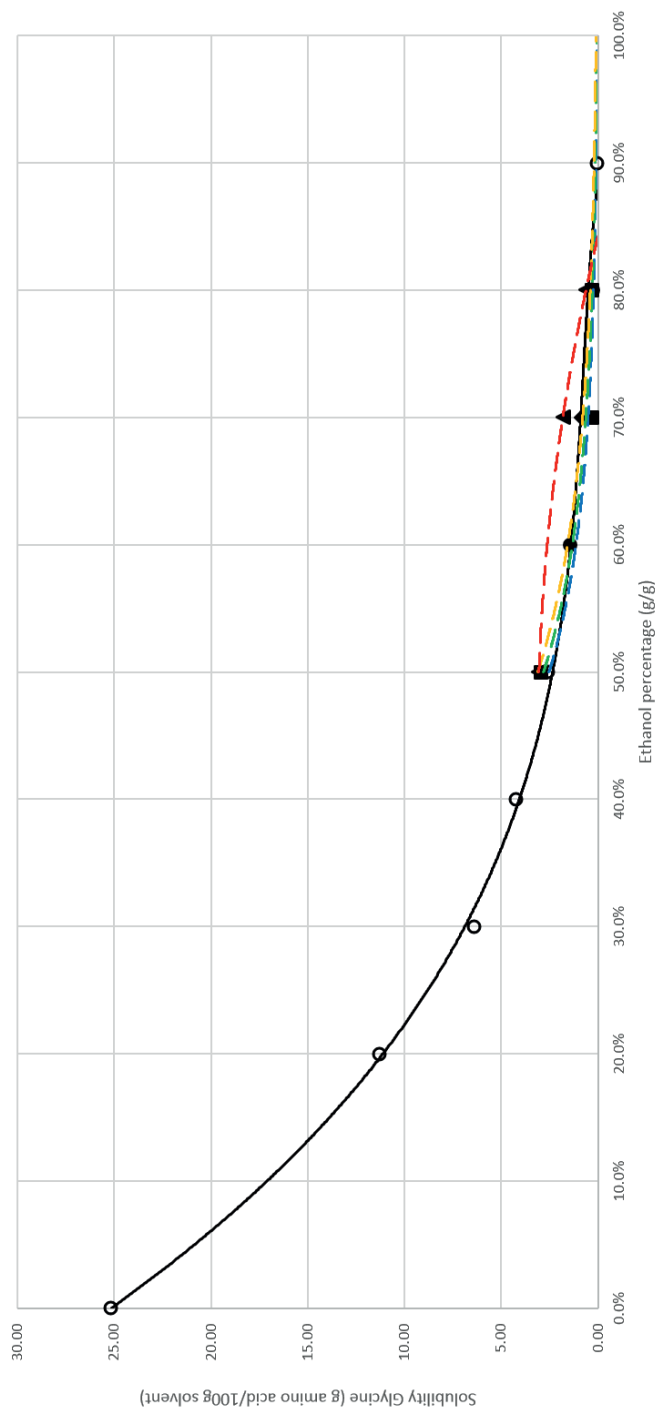


Figure 4.1: Solubility of glycine as a function of ethanol solvent concentration (in g of ethanol /g of solvent) in single solution at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

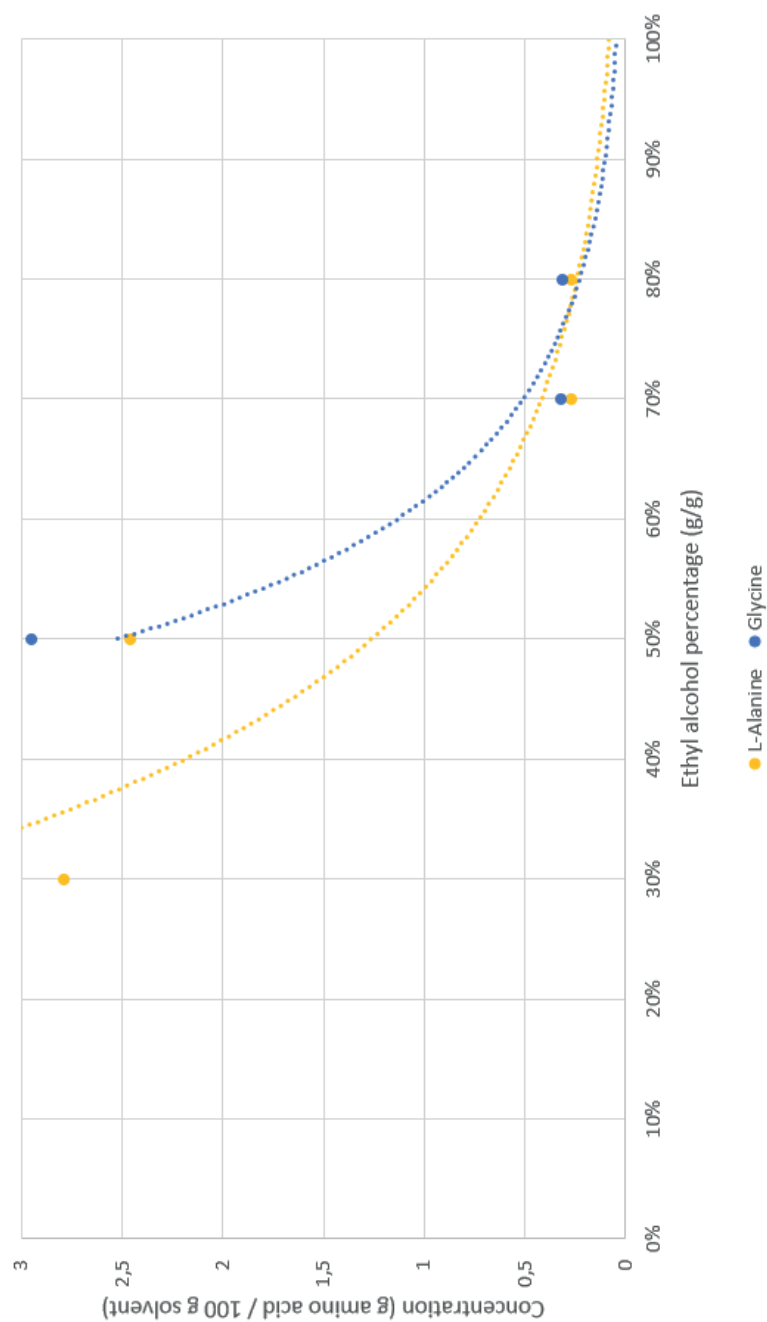


Figure 4.2: Solubility of glycine and L-alanine as a function of ethanol solvent concentration (in g of ethanol /g of solvent) in a mixture of 20 α -amino acids at 277.15 K in an ethanol/water system.

4.3.2 Branched aliphatic series: L-Valine, L-Leucine, L-Isoleucine

Three α -amino acids have aliphatic side chains that branch: L-valine, L-leucine and L-isoleucine. L-Valine is similar to L-alanine but with two branched methyl groups off of the methyl group of L-alanine's side chain. L-Isoleucine is one carbon atom longer on one of the branched methyl groups. L-Leucine has two branched methyl groups off of the second carbon of the side chain.

As shown in Figure 4.3, L-valine becomes more soluble in a mixture of 20 α -amino acids between 30% and 70% (g/g) ethanol-water solutions than if it were alone in the same ethanol/water solution. However, as seen in Figure 4.4 and Figure 4.5, L-leucine is less soluble in a mixture of α -amino acids than by itself and this is even more pronounced in L-isoleucine. As in the hypothesis made with glycine/ L-alanine, one additional methyl group makes the amino acid slightly more interactive with other amino acids, increasing its solubility. However, an increase of a carbon in the side chain will negatively impact this ability to interact with other α -amino acids. The longer the carbon chain, the less interactive the α -amino acid will be to other α -amino acids.

Furthermore, as shown in Figure 4., L-valine remains the most soluble in relation to the other 2 branched. As the concentration of ethanol increases, L-leucine, which is the least soluble in this series, "catches up" to the solubility of the other 2 and surpasses L-isoleucine at 50% and above. L-Isoleucine decreases solubility relative to the other two amino acids in this series. L-Alanine is at all points relatively more soluble than all of the branched series. With the exception of L-isoleucine at 277.15 K and 293.15 K, at all temperatures, as the ethanol concentration increases, the rate at which the solubility of the amino acid decreases increases for all branched aliphatic amino acids in mixtures. This is shown by a trendline that is concave downwards. In single solution, this is the reverse. In single solution, as the ethanol concentration increases, the rate at which the solubility of the amino acid decreases decreases, producing a trendline that is concave upwards.

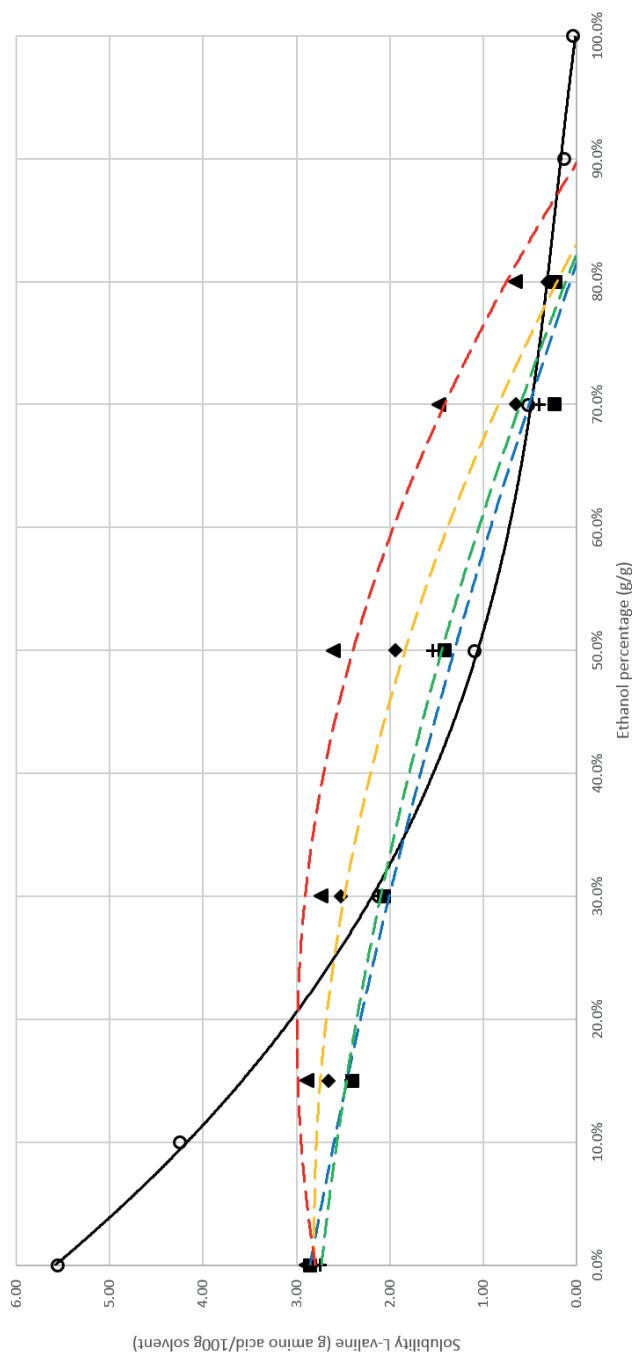


Figure 4.3: Solubility of L-valine as a function of ethanol solvent concentration (in g of ethanol /g of solvent) in single solution at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

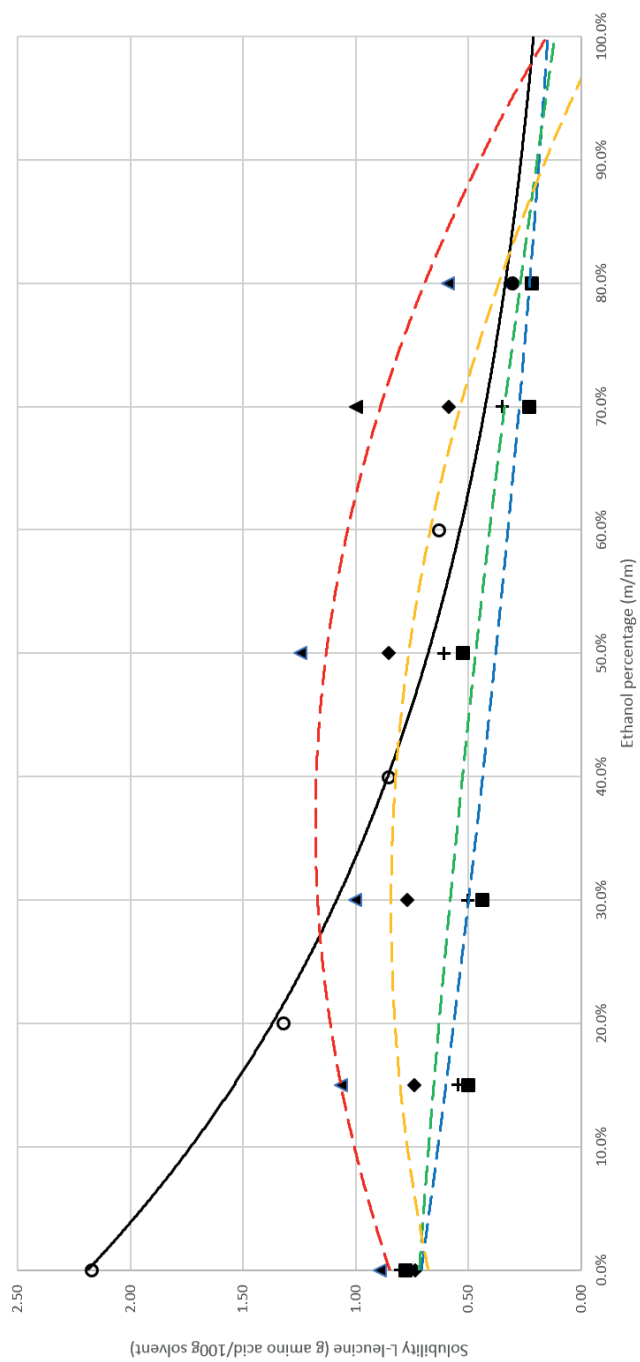


Figure 4.4: Solubility of L-leucine as a function of ethanol solvent concentration (in g of ethanol / g of solvent) in single solution at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

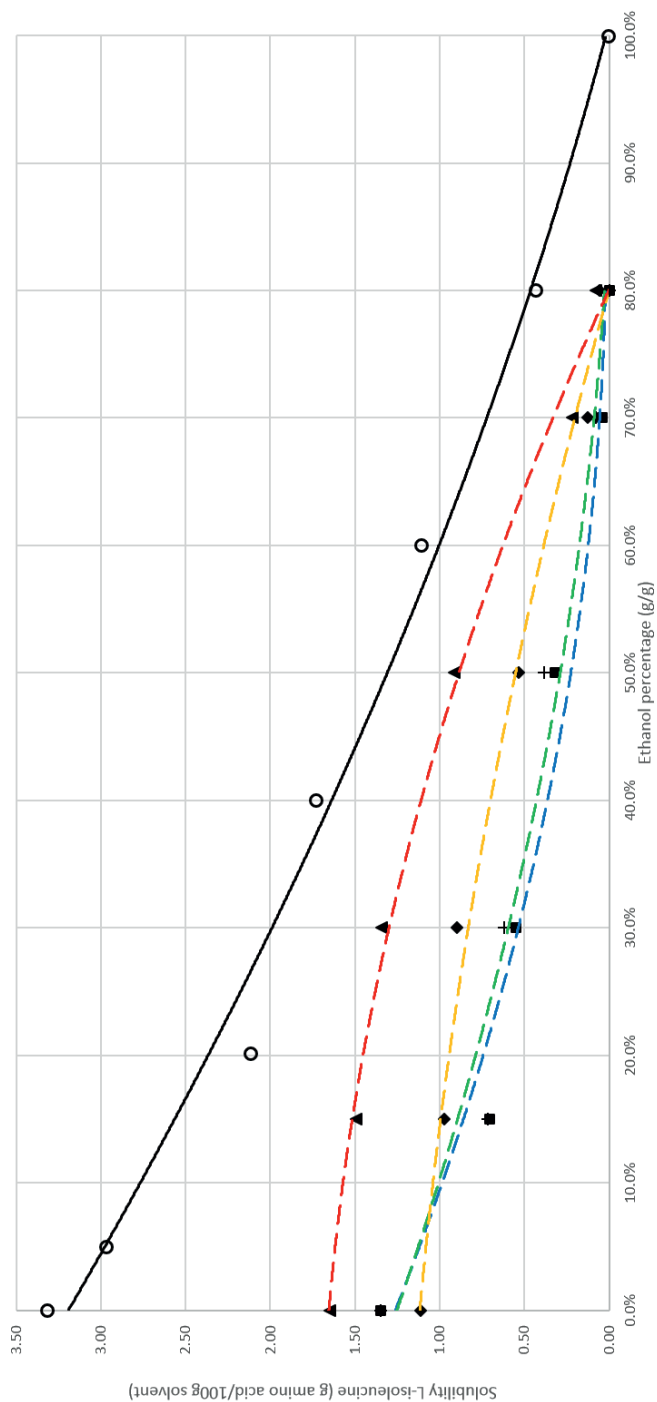


Figure 4.5: Solubility of L-isoleucine as a function of ethanol solvent concentration (in g of ethanol /g of solvent) in single solution at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

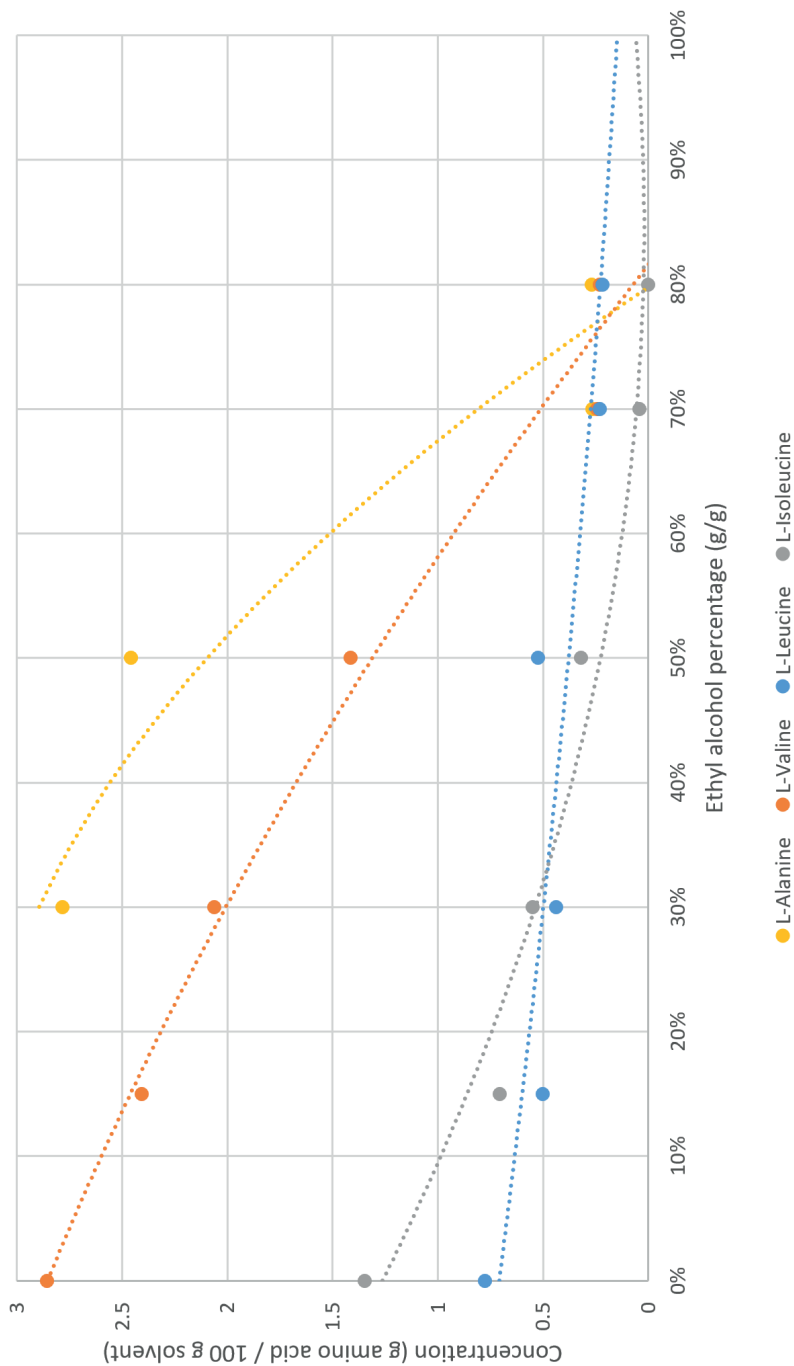


Figure 4.6: Solubility of L-alanine, L-valine, L-leucine and L-isoleucine in a mixture of 20 α -amino acids as a function of ethanol solvent concentration (in g of ethanol / g of solvent) at 27.15 K in an ethanol/water system.

4.3.3 Hydroxylic group: L-Serine, L-Threonine, L-Tyrosine

There are three α -amino acids with a hydroxyl group on its side chain. L-Serine is similar to L-alanine but with a hydroxyl group attached to the methyl group. L-Threonine has a side chain of two carbons with the methyl group attached to the first carbon of the side chain. L-Tyrosine is, like L-serine, similar to L-alanine, but with a phenyl group attached to the single carbon in the side chain and thereupon the hydroxyl group is attached.

As shown in Figure 4.7, the presence of the hydroxyl group has little effect on the solubility of L-serine in a solution of 20 α -amino acids. Little difference can be shown between L-serine in an ethanol/water solution by itself or in the presence of other α -amino acids. As shown in Figure 4.7, L-threonine has an elevated solubility in a solution with 20 α -amino acids in comparison to being in an ethanol-water solution by itself. However, this could be due to the presence of an additional methyl group, as compared to L-serine. This additional methyl group has been shown in previous series to increase solubility in groups of amino acids. Figure 4.9 demonstrates a remarked increase in solubility for L-tyrosine in conjugation with other α -amino acids compared to single solution. However, the presence of a phenyl group could be the reason for the increase in solubility and not the hydroxyl group.

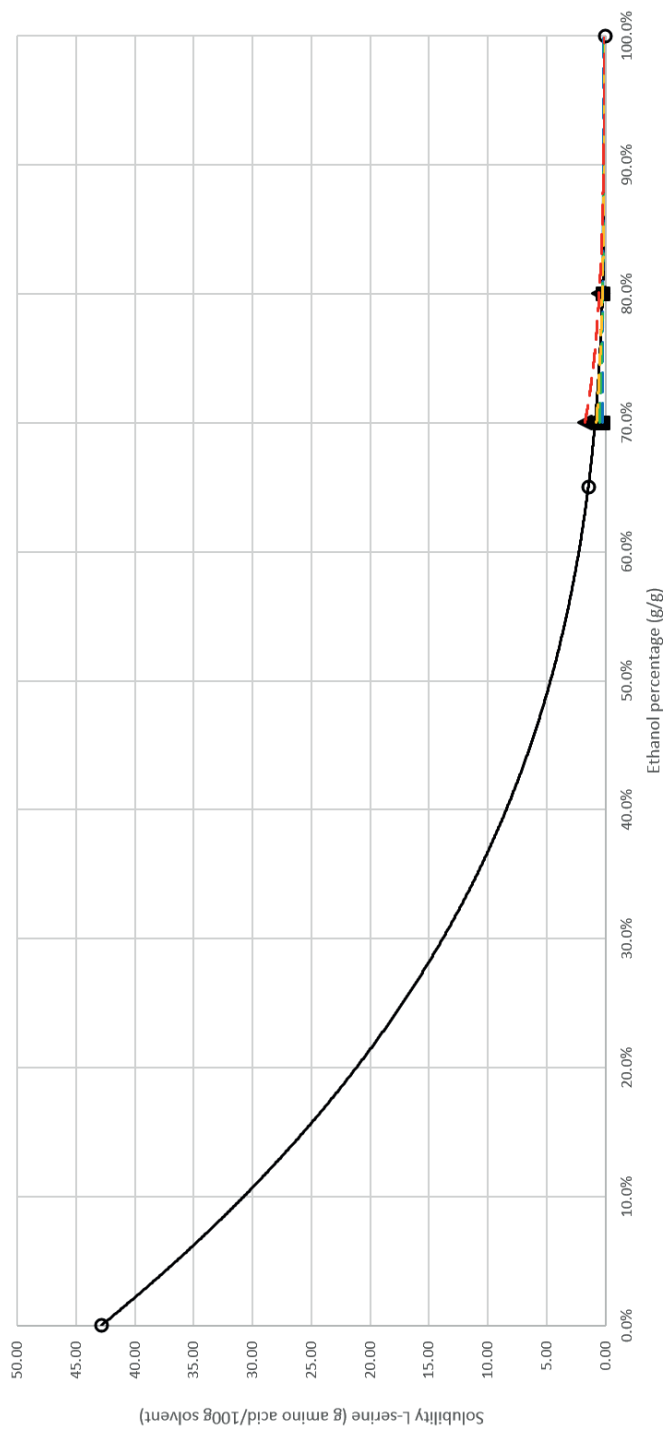


Figure 4.7: Solubility of L-serine in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

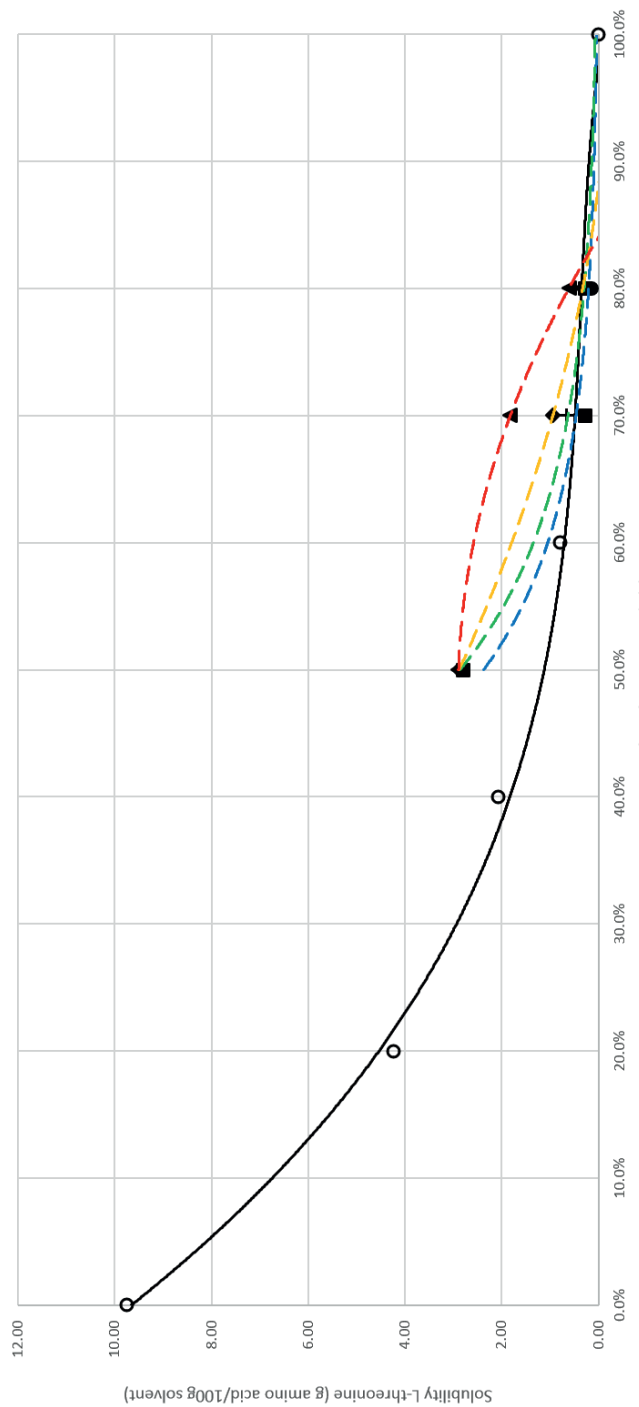


Figure 4.8: Solubility of L-threonine in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

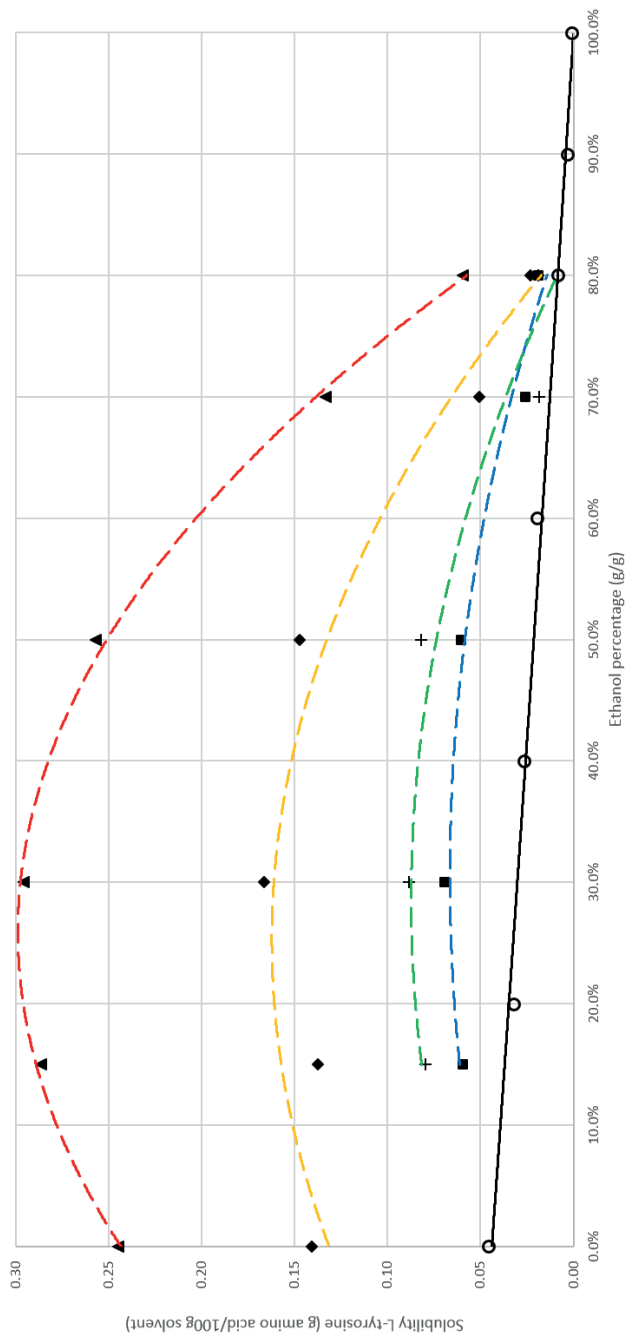


Figure 4.9: Solubility of L-tyrosine in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

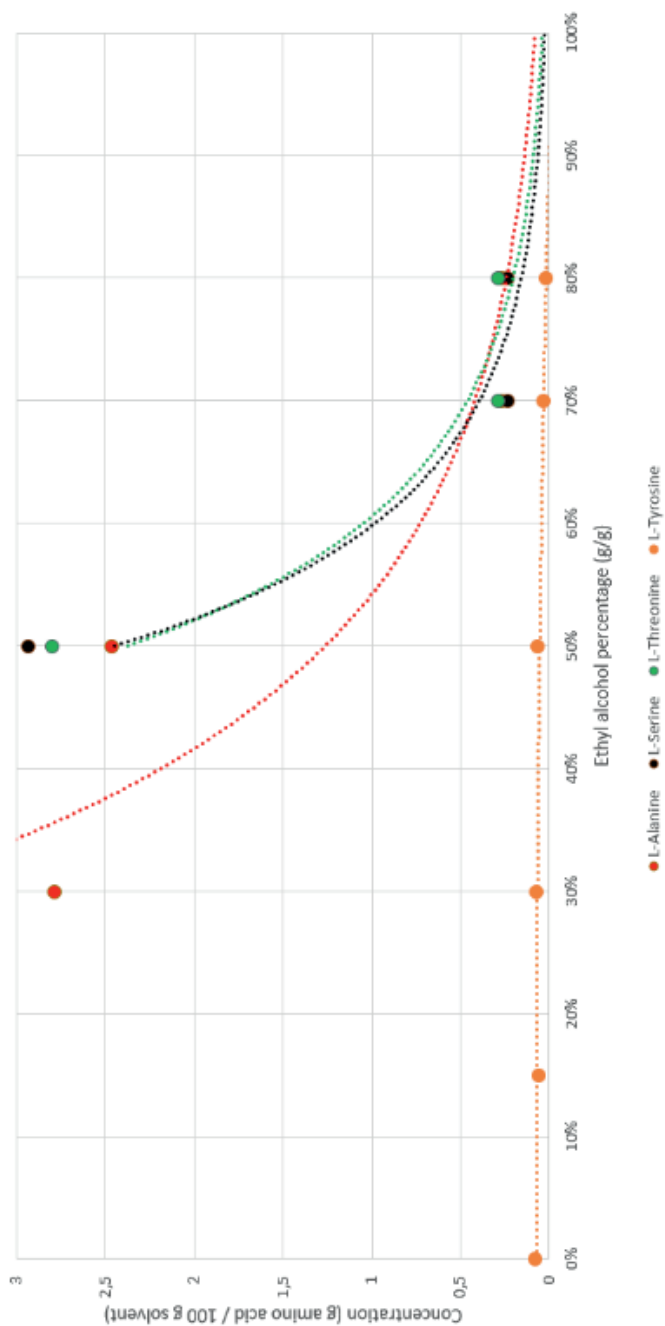


Figure 4.10: Solubility of L-alanine, L-serine, L-threonine and L-tyrosine as a function of ethanol solvent concentration (in g of ethanol / g of solvent) in a mixture of 20 α -amino acids at 277.15 K in an ethanol/water system.

With no other amino acid in solution, L-alanine has a relatively higher solubility than the other three in this series. However, in a solution with other amino acids, both L-serine and L-threonine have solubilities higher than L-alanine at 30% and 50% ethanol (g/g). L-Threonine has furthermore, a higher solubility at all points. L-Tyrosine, still has a relatively lower solubility than the other amino acids in this series despite the increase in solubility relative to itself in a solution with no other amino acids present. With the exception of L-tyrosine at all temperatures and L-threonine at 335.15 K, as the ethanol concentration increases, the rate at which the solubility of the amino acid decreases decreases for amino acids with a hydroxylic group. This produces a trendline that is concave upwards.

4.3.4 Phenyl series: L-Phenylalanine, L-Tyrosine, L-Tryptophan

Three α -amino acids have a phenyl group in their side chain. L-Phenylalanine is similar to L-alanine but with a phenyl group attached to the single carbon in the side chain. As shown in 4.11, L-phenylalanine does not show an increase in solubility in a group of α -amino acids.

L-Tyrosine is similar to L-phenylalanine but with a hydroxyl group attached to the 4-carbon. As shown in Figure 4.9, L-tyrosine has a marked increase in solubility when in a solution with α -amino acids present.

L-Tryptophan, as seen in Figure 4.12, sees an increase in solubility similar to L-tyrosine. L-Tryptophan is similar in structure to L-phenylalanine, but with a pyrrol group attached between the carbon atom in the side chain and the phenyl group.

Since L-phenylalanine showed no increase in solubility, while L-alanine did show a slight increase in solubility, the conclusion cannot be drawn that the addition of a phenol group has an impact on interactions between amino acids. The addition of an OH group or a pyrrol group could explain the large increase to the interactions with other amino acids in L-tyrosine and L-tryptophan respectively.

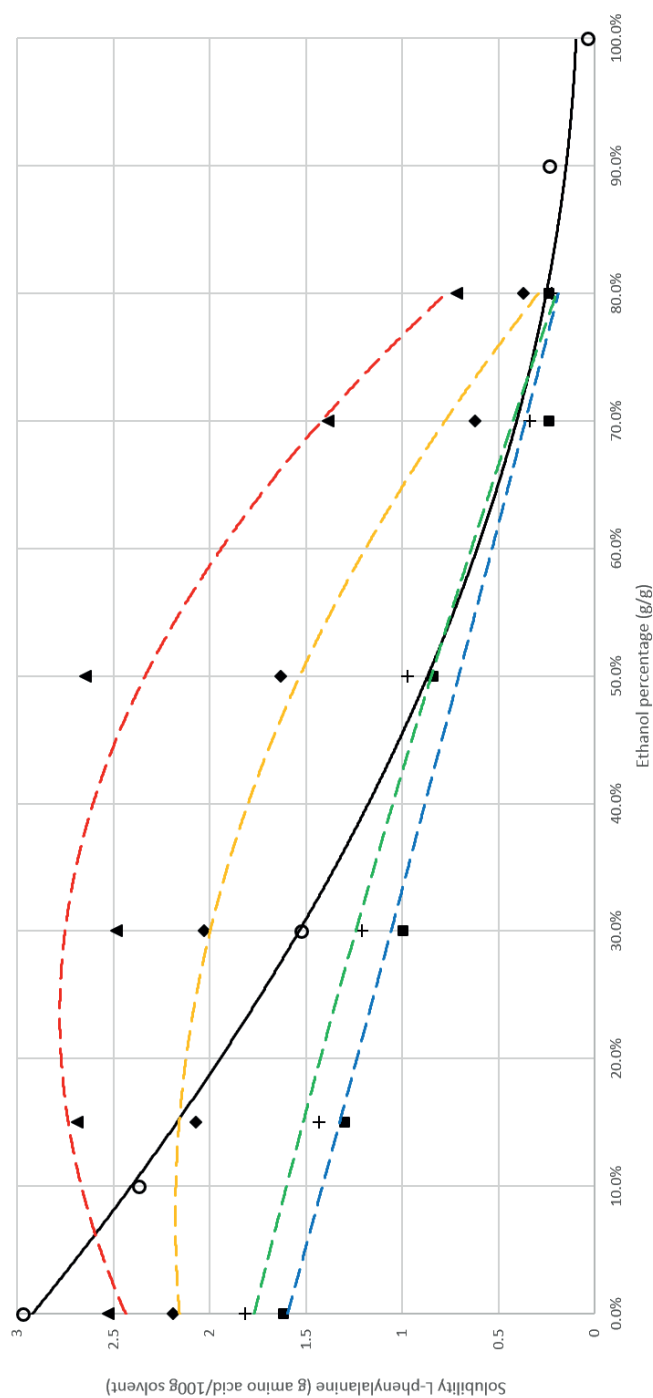


Figure 4.11 : Solubility of L-phenylalanine in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

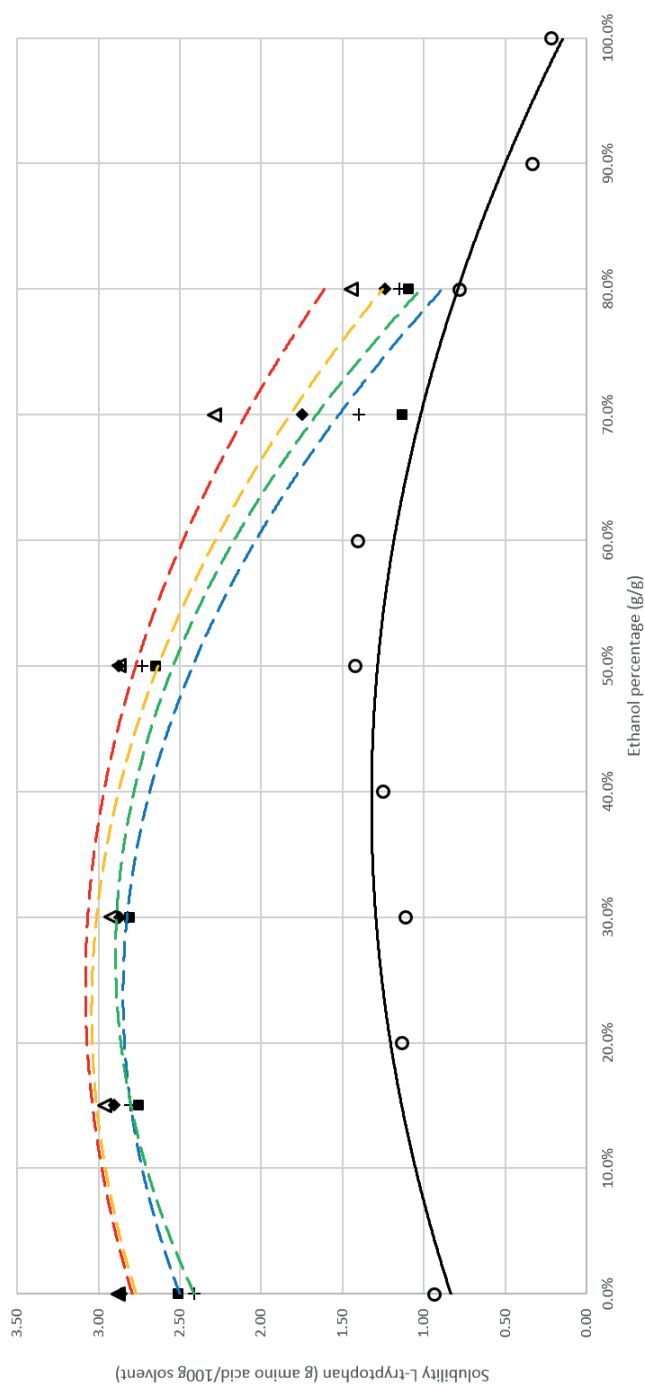


Figure 4.12: Solubility of L-tryptophan in single solution as a function of ethanol solvent concentration (in g of ethanol / g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

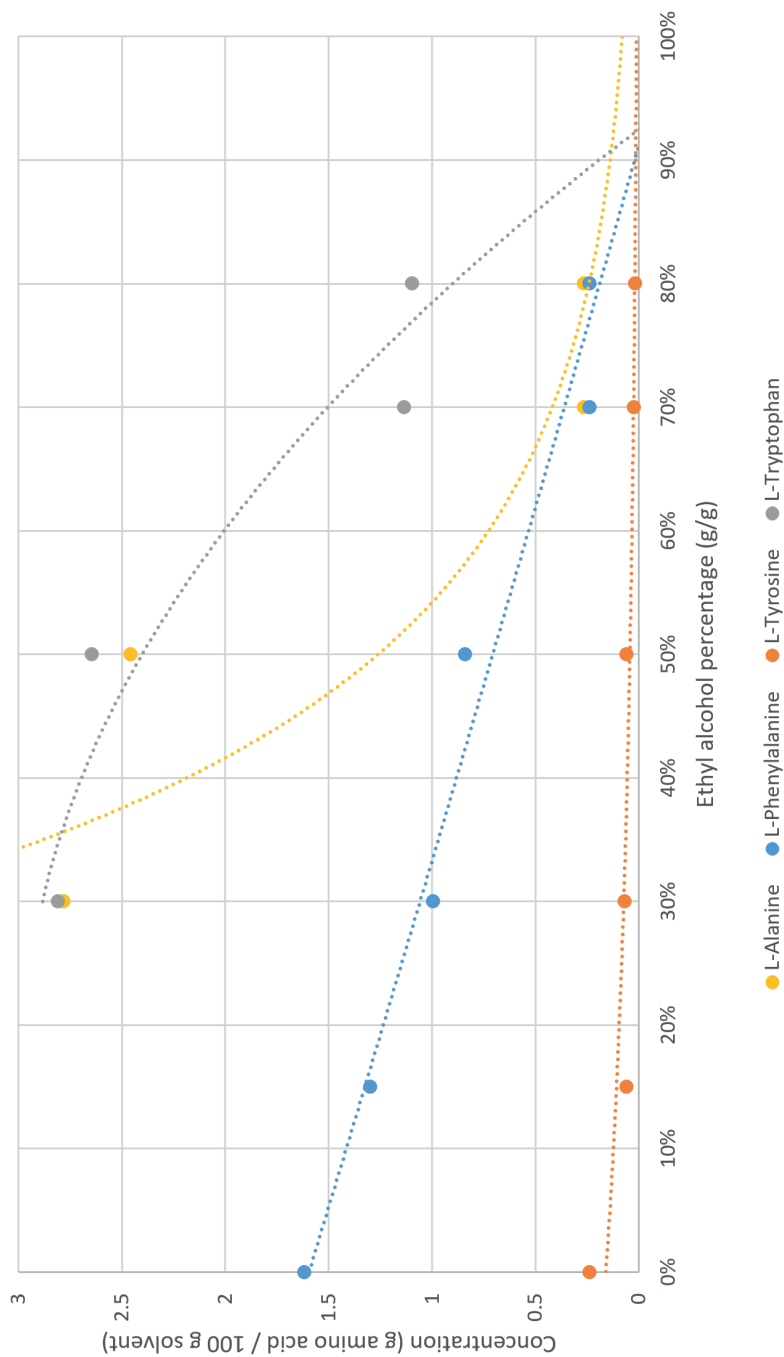


Figure 4.13: Solubility of L-alanine, L-tyrosine and L-tryptophan in a mixture as a function of ethanol solvent concentration (in g of ethanol / g of solvent) of 20 α -amino acids at 277.15 K in an ethanol/water system.

However, L-tyrosine is still the least soluble in this group relative to the other three as shown in Figure 4.. L-Phenylalanine is always more soluble than L-tyrosine but always less soluble than L-alanine and L-tryptophan. L-Tryptophan is the most soluble with only a few exceptions. At all temperatures for L-tryptophan and for L-phenylalanine at 293.15 K and 315.15 K, as the ethanol concentration increase, the rate of decrease in solubility of the amino acid increases. This causes a trendline that is concave downwards. The same trendline is produced by these amino acids in single solution with the addition of ethanol as seen in chapter 3.

4.3.5 Sulphur series: L-Cysteine and L-Methionine

The two α -amino acids that contain Sulphur are L-cysteine and L-methionine. L-Cysteine is similar to L-alanine but with a Sulphur at the end of the side chain. L-Methionine is also similar to L-alanine but has an additional CH_2SCH_3 group.

As seen in Figure 4.14, L-cysteine has a lower solubility when other α -amino acids are present than alone. Furthermore, a mass balance of the precipitate and solutes in a similar study, see section 8.3, showed a 98% recovery of all amino acids, making consumption of L-cysteine in a secondary reaction unlikely. In Figure 4.15, L-methionine also shows a sharp decrease in solubility, although less so than L-cysteine. The hypothesis could be made that sulphur decreases the ability of α -amino acids to form bonds with other amino acids. Furthermore, additional methyl groups help mitigate this insolubility effect.

Furthermore, as shown in Figure 4., the relative solubilities of the two α -amino acids in this series are lower than L-alanine. This indicates that the presence of sulphur inhibits their solution. Furthermore, L-cysteine when not in a mixture is relatively more soluble than L-methionine, however, when other amino acids are present the solubility of L-cysteine is lower than L-methionine. For L-methionine at all temperatures, as the ethanol increases the rate of decrease in solubility of L-methionine increases. This is shown by a trendline that is concave downward. In single solution, the trendline is concave upward, see chapter 3.

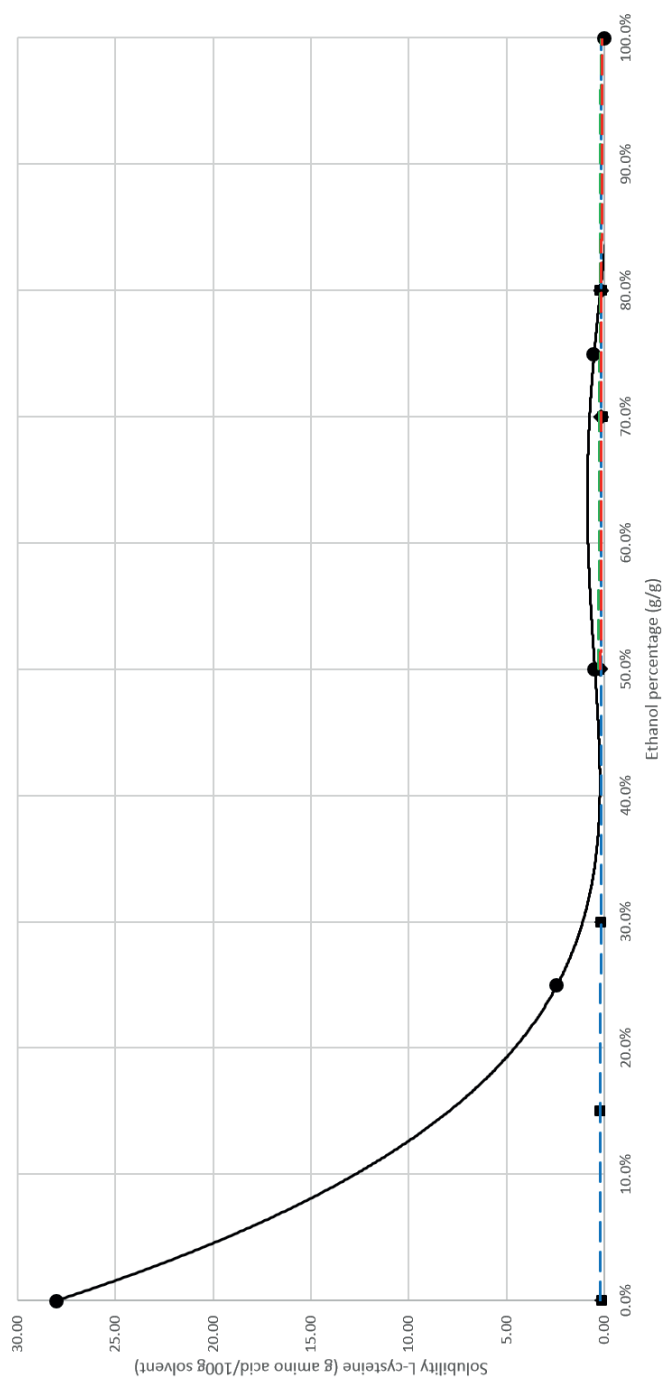


Figure 4.14: Solubility of L-cysteine in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

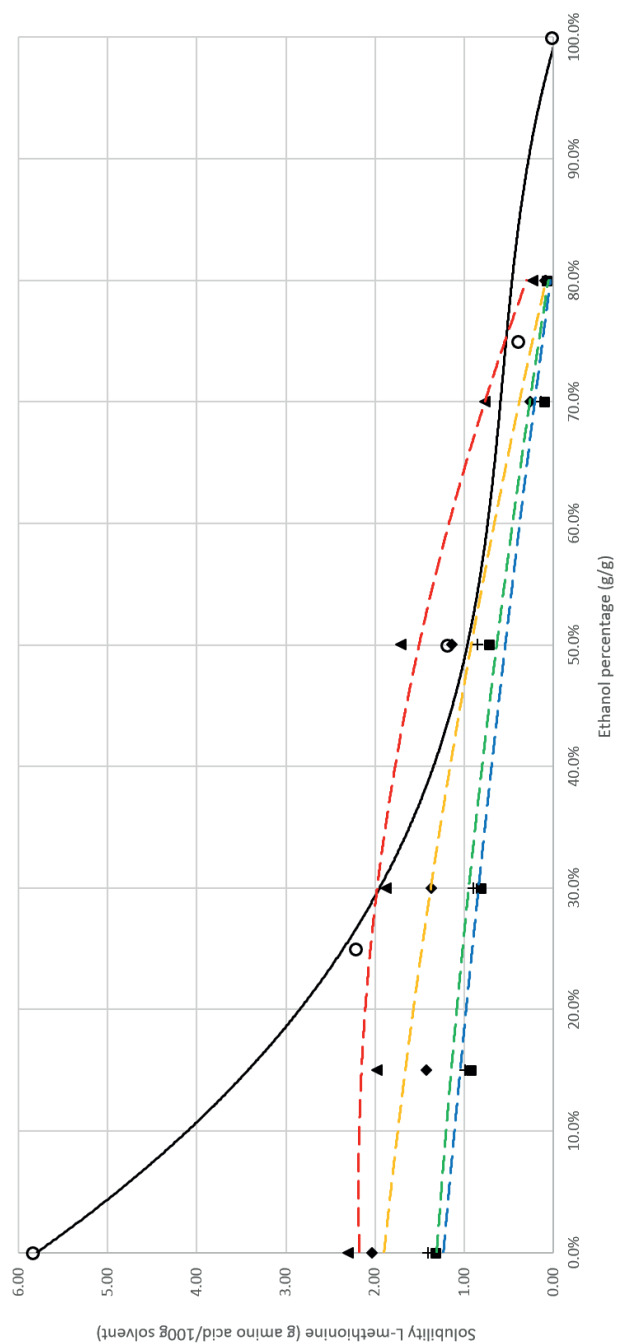


Figure 4.15: Solubility of L-methionine in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

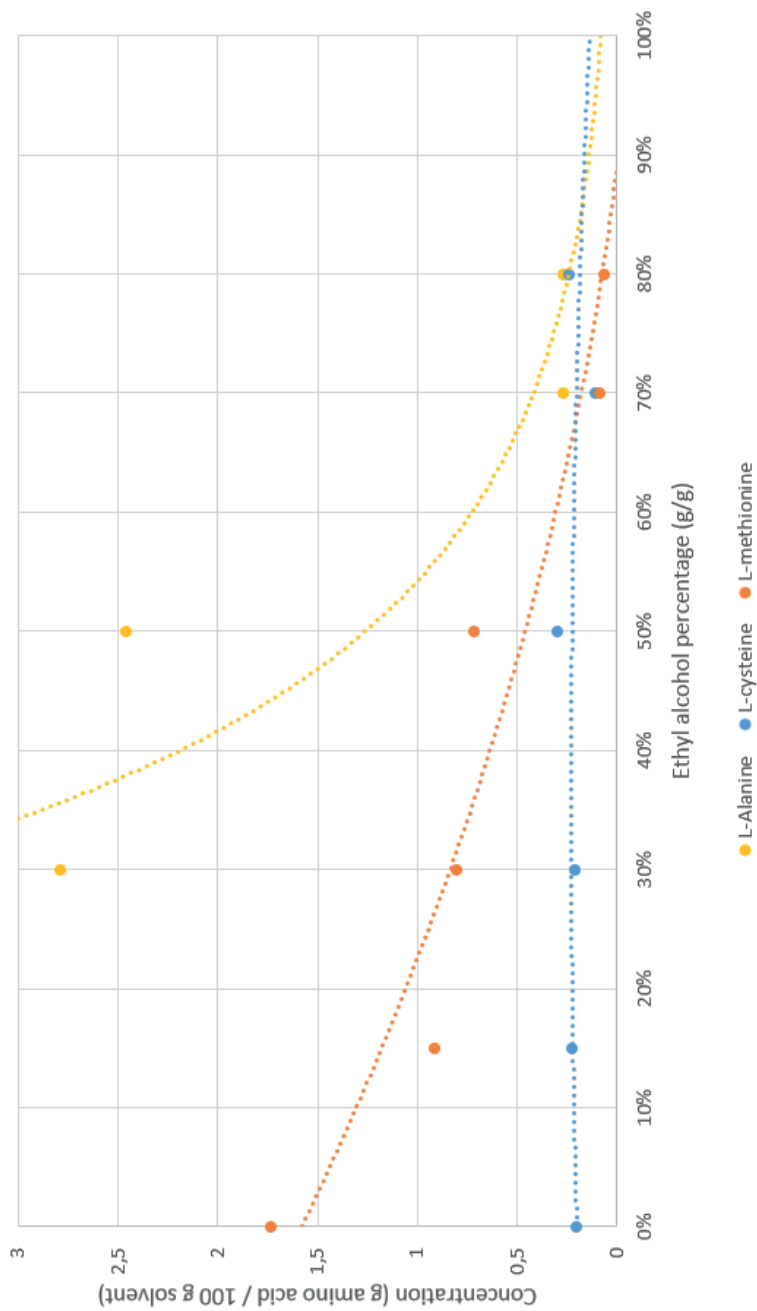


Figure 4.16: Solubility of L-alanine, L-cysteine and L-methionine in a mixture of 20 α -amino acids as a function of ethanol solvent concentration (in g of ethanol / g of solvent) at 277.15 K in an ethanol/water system.

4.3.6 Amine/Amide series: L-Lysine, L-Arginine, L-Asparagine, L-Glutamine

Four α -amino acids contain either an amine or an amide in their side chains. L-Lysine is similar to L-alanine but with three additional carbons in the side chain and ending in a positively charged amine group. L-Arginine has a side chain three carbons in length then a nitrogen atom followed by a carbon that is bonded with two amines. L-Asparagine and L-glutamine both have side chains consisting of a number of carbons, L-asparagine has one and L-glutamine has 2, ending in an amide group.

As seen in Figure 4.17 and Figure 4.18, the solubility of L-lysine and L-arginine is reduced, resulting in a hypothesis can be made that the presence of an amine group inhibits the ability of α -amino acids to interact and form bonds with other α -amino acids. At all temperatures, for L-lysine and L-arginine, as the ethanol concentration increases, the rate of decrease in solubility increases. This causes a trendline that is concave downward.

The effect of the presence of amides can be seen in Figure 4.18 and Figure 4.. For L-asparagine, the solubility was increased at all temperatures when in a model mixture. However, for L-glutamine, the solubility was decreased. L-Glutamine is one methyl group longer than L-asparagine. The hypothesis could be made that a methyl group decreases the reactivity of the amino acid in a mixture. At all temperatures, for L-asparagine and L-glutamine, as the ethanol concentration increase, the rate at which the solubility of the amino acids decrease decreases. This is shown by trendlines that are concave upward. For all of these amino acids in single solution, the trendlines are also concave upward, see chapter 3.

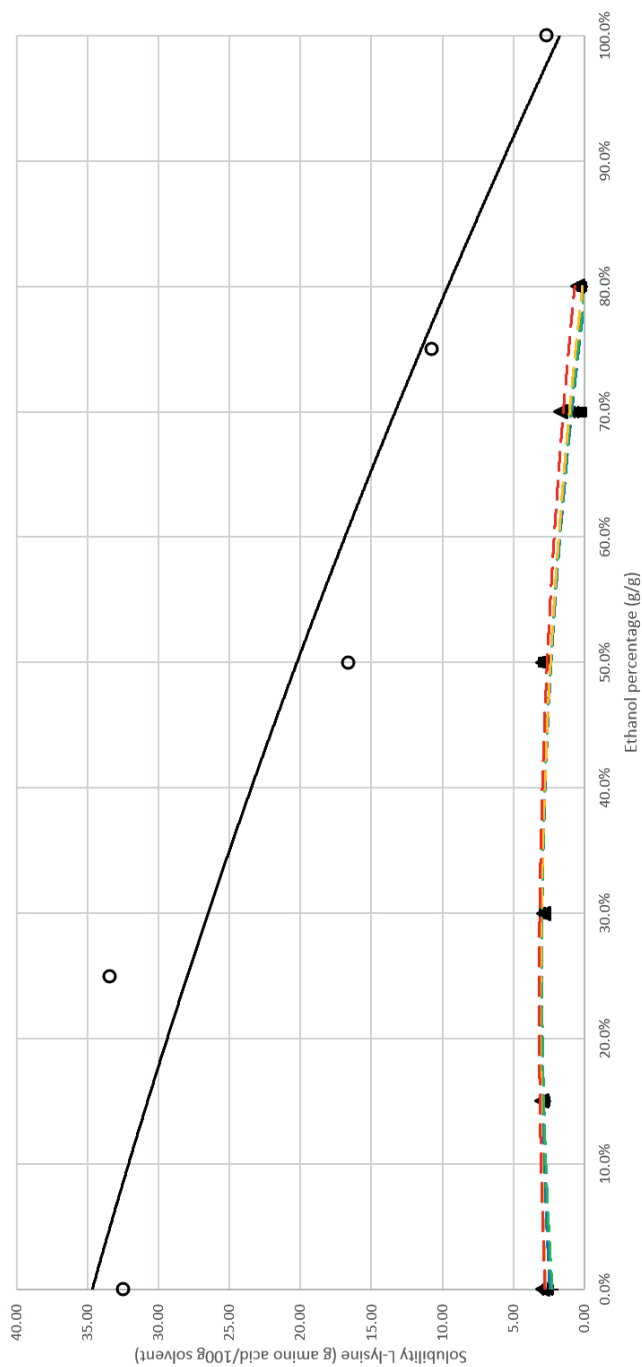


Figure 4.17: Solubility of L-lysine in single solution at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

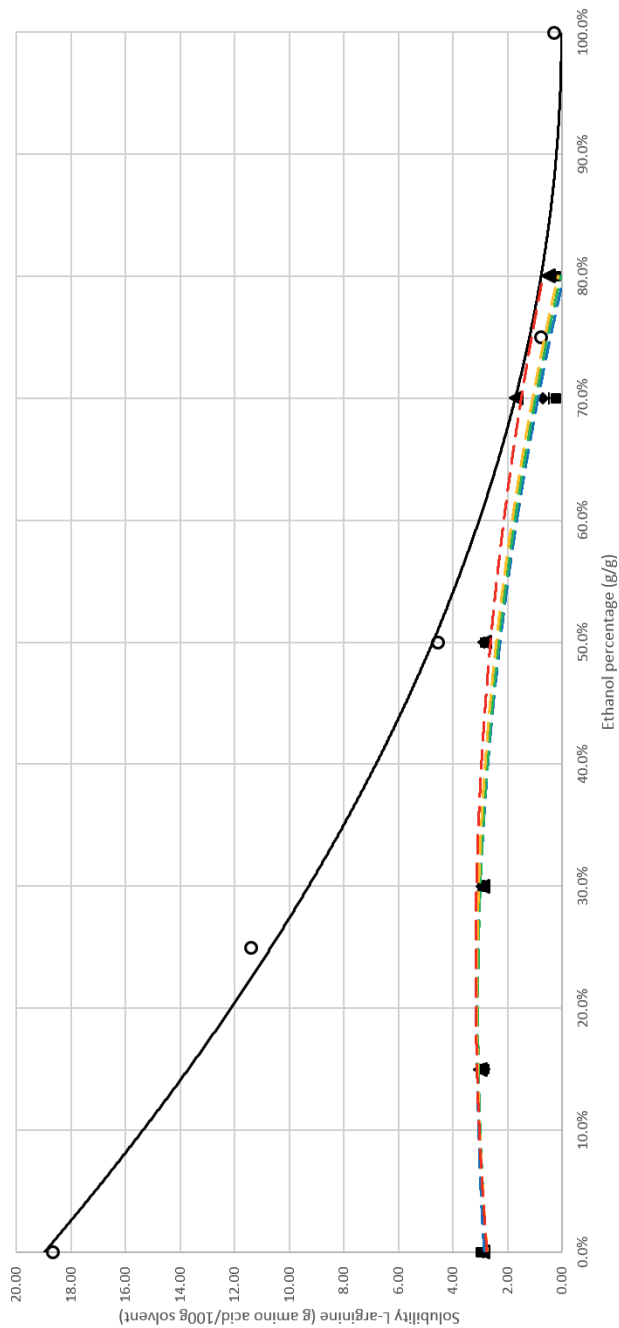


Figure 4.18: Solubility of L-arginine in single solution at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

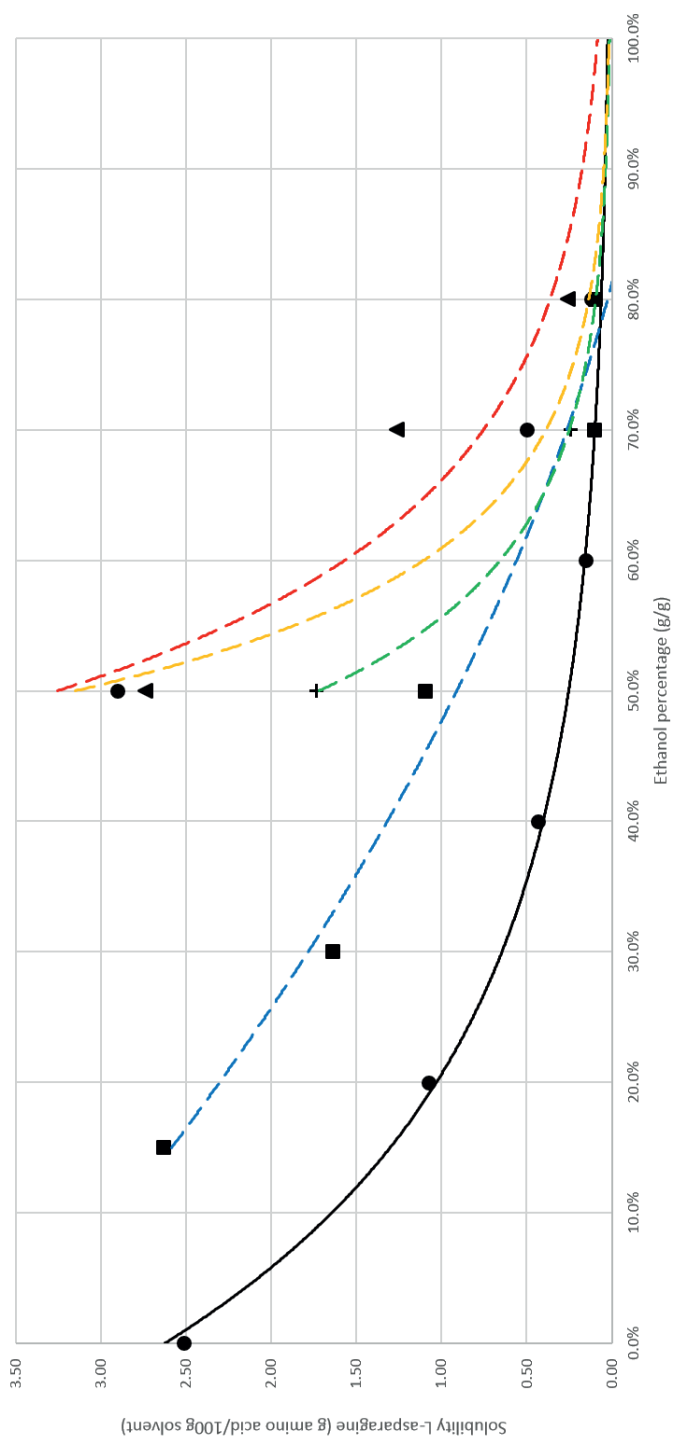


Figure 4.19: Solubility of L-asparagine in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

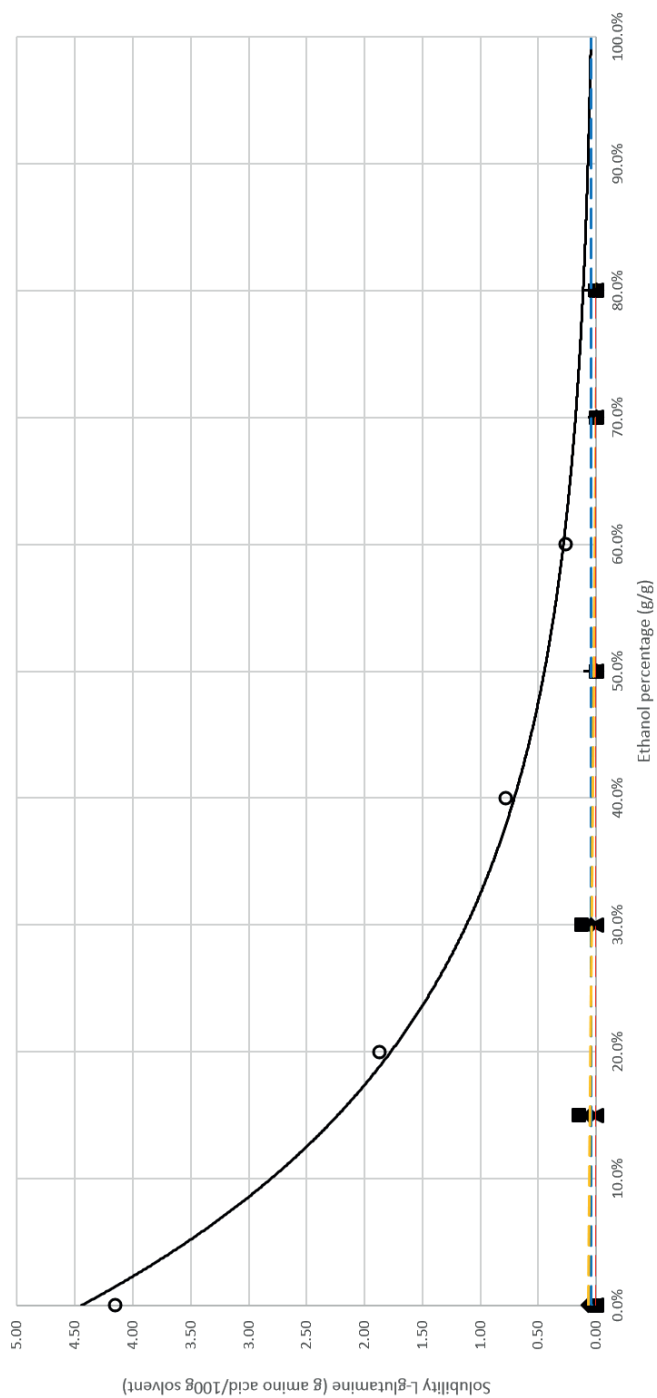


Figure 4.20: Solubility of L-glutamine in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

4.3.7 Carboxylic Acid Containing Series: L-Aspartic Acid, L-Glutamic Acid

In Figure 4.22 and Figure 4. the solubilities of L-aspartic acid and L-glutamic acid are shown respectively. Both amino acids are more soluble at all temperatures in a model mixture of all 20 amino acids than in solution by themselves. This is true for all ethanol concentrations. At all temperatures, for L-aspartic acid, as the ethanol concentration increases, the rate at which the solubility of the amino acid decreases increases. This is shown in trendlines that are concave downward. This is the opposite effect for the solubility in single solution.

As you can see in Figure 4., the solubility of L-aspartic acid and L-glutamic acid are higher than the solubility of L-alanine until 50% ethanol. At 70% ethanol, L-alanine has a higher solubility than both L-aspartic acid and L-glutamic acid. This leads to the hypothesis that both L-aspartic acid and L-glutamic acid are more ethanolphobic than L-alanine. This means that side chains with carboxylic acid groups are more ethanolphobic when in mixtures of amino acids than side chains that do not have carboxylic acid groups.

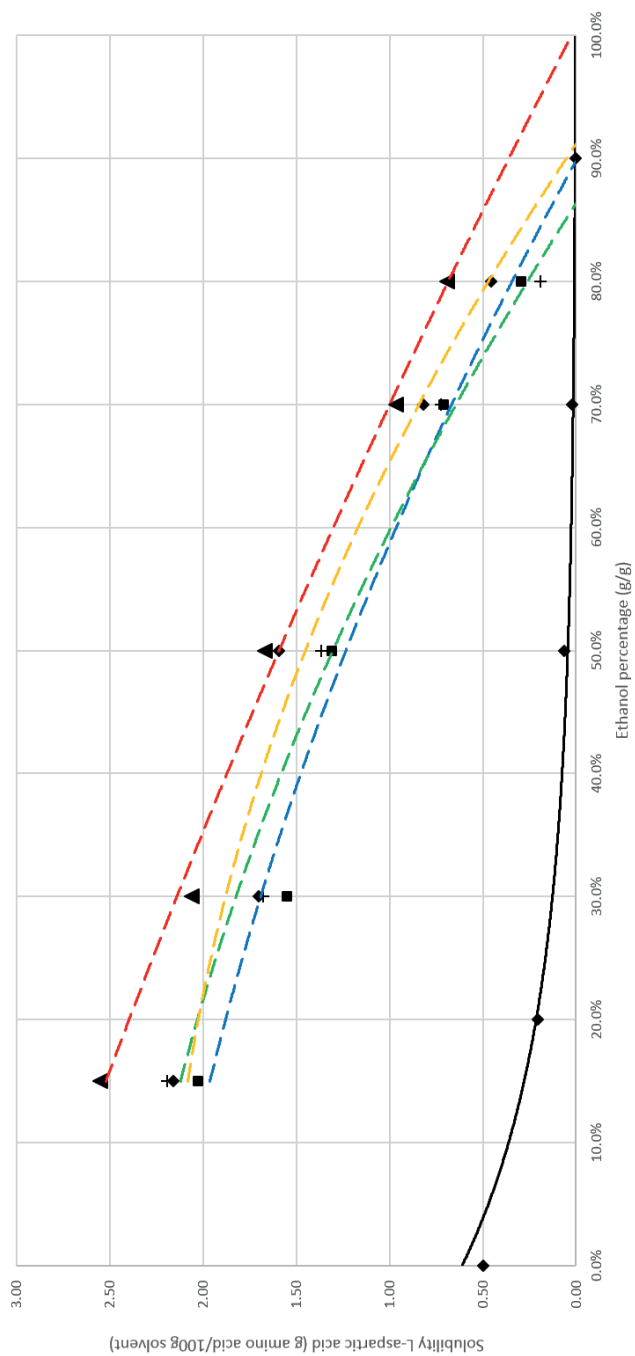


Figure 4.21 : Solubility of L-aspartic acid in single solvent as a function of ethanol solvent concentration (in g of ethanol / g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

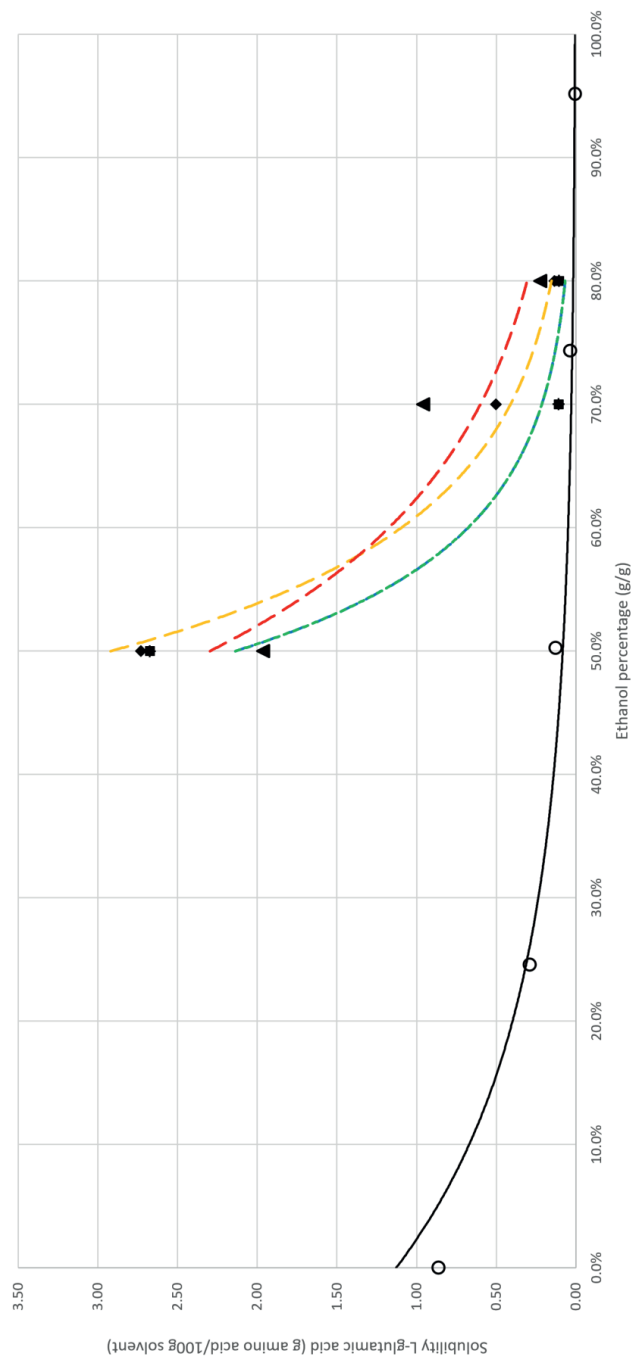


Figure 4.22: Solubility of L-glutamic acid in single solution as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

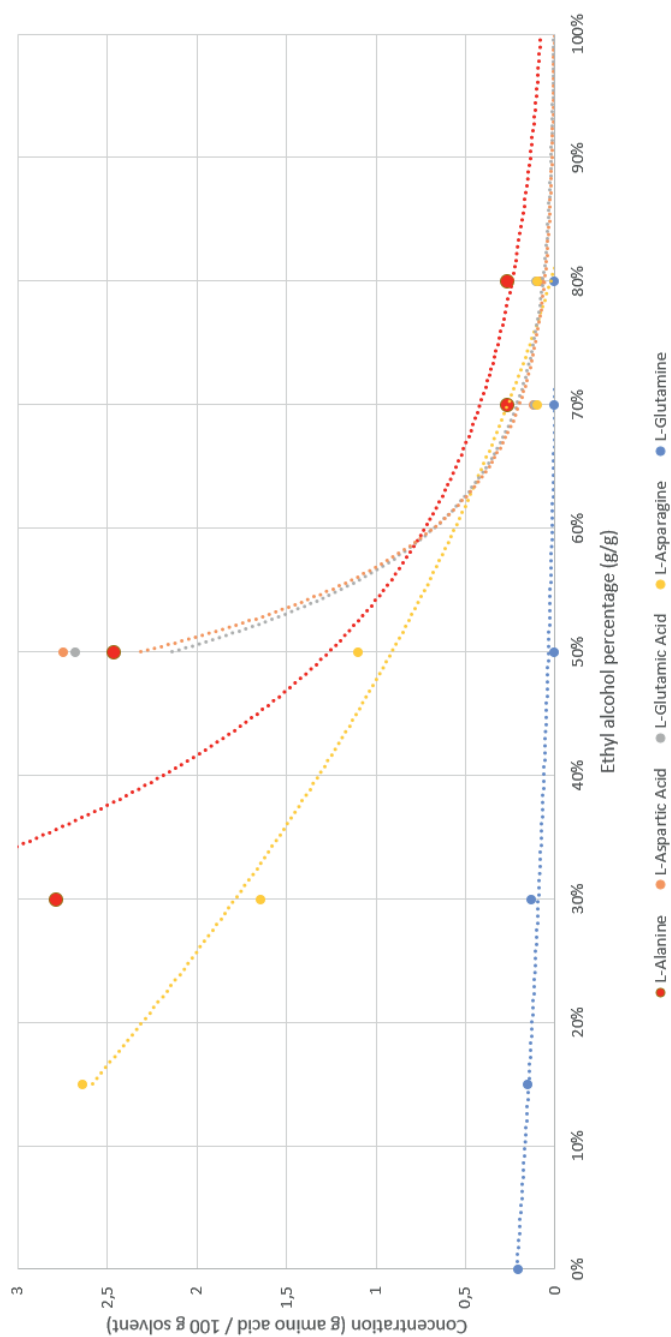


Figure 4.23: Solubility of amino acids with carboxylic acid groups in their side chain in a mixture of 20 amino acids compared to L-alanine in an ethanol/water system as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 277.13 K.

4.3.8 Heterocyclic Series: L-Histidine, L-Proline

L-Histidine and L-proline are amino acids with heterocyclic groups. The solubility of these amino acids can be seen in Figure 4.24 and Figure 4. respectively. We see that L-histidine has a higher solubility in a mixture of all 20 amino acids and at all temperatures until 90% ethanol. L-Proline has a lower solubility when in a group of amino acids at all temperatures and at all concentrations of ethanol.

As shown in Figure 4., L-proline has a higher solubility than L-alanine at all concentrations of ethanol when compared to L-alanine. L-Histidine has a higher concentration than L-alanine until 70% ethanol concentration. At 80 and 90% ethanol, both L-histidine and L-alanine have approximately the same solubility. This means that not all heterocyclic side chains have the same effect in a mixture of amino acids when ethanol is added.

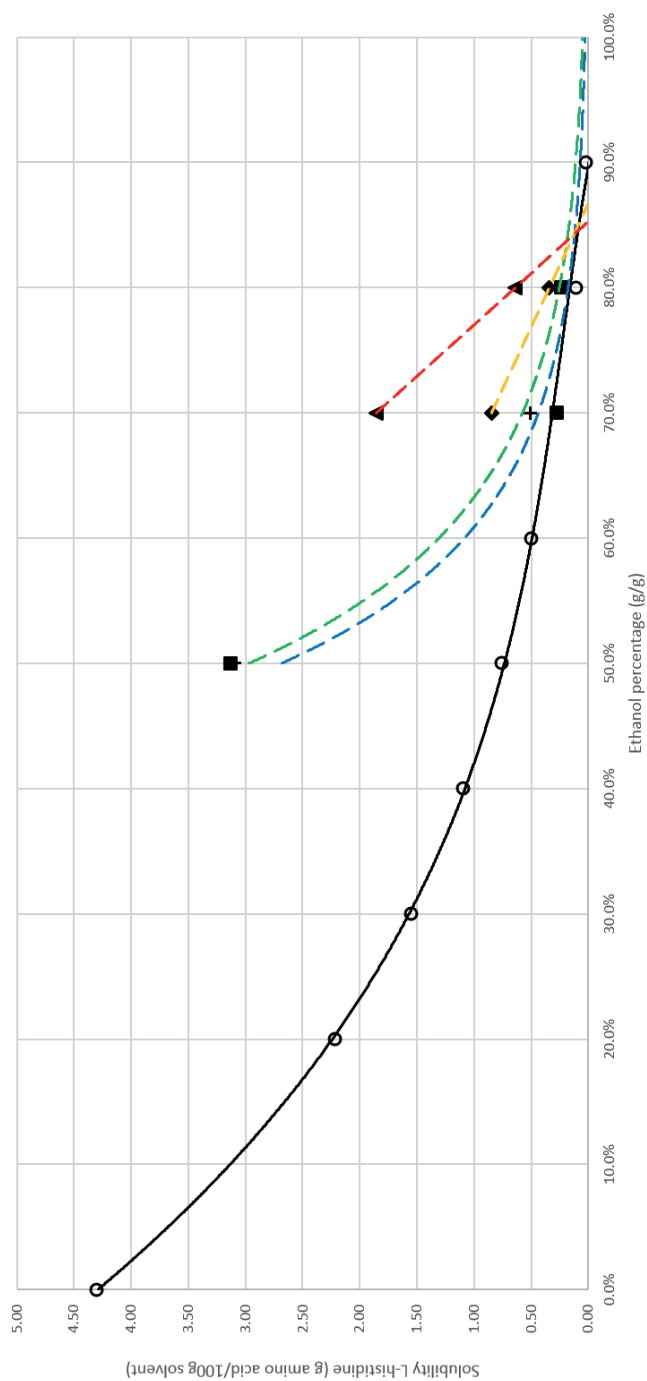


Figure 4.24: Solubility of L-histidine in single solution at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system

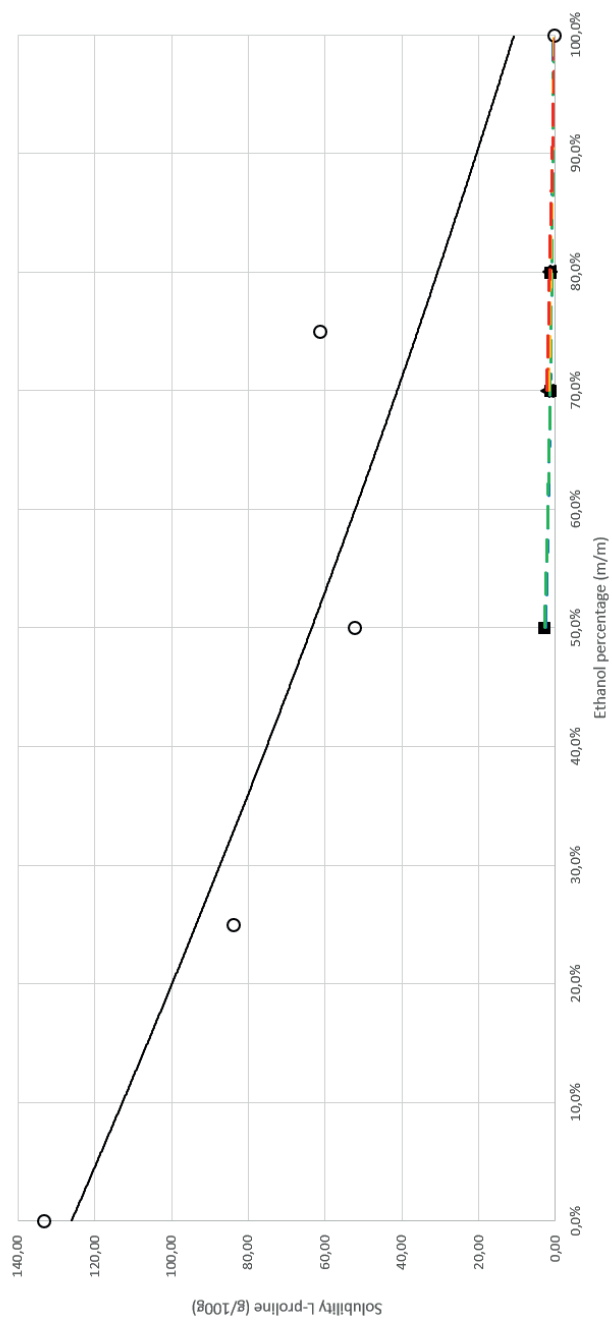


Figure 4.25: Solubility of L-proline in single solution at 298.15 K (solid line) and in a mixture of 20 amino acids (striped line) as a function of ethanol solvent concentration (in g of ethanol /g of solvent) at 277.15 K (blue), 293.15 K (green), 315.15 K (yellow) and 335.15 K (red) in an ethanol/water system.

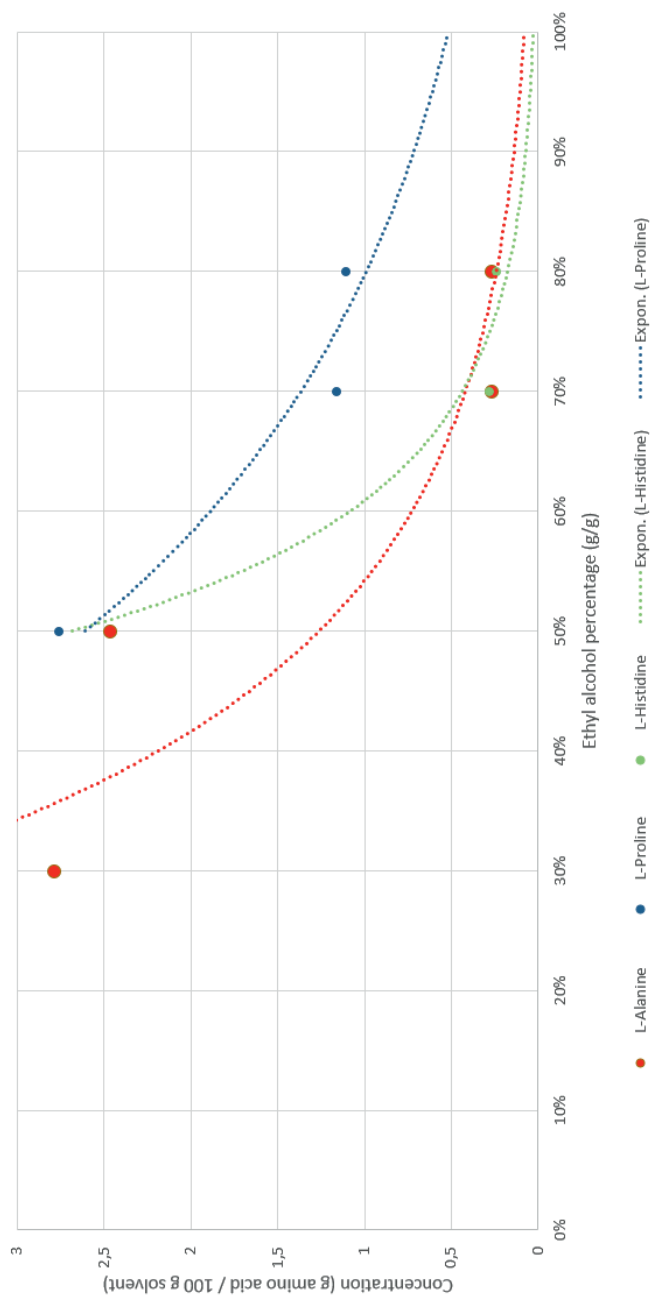


Figure 4.26: Solubility of the amino acids with heterocyclic side chains in a mixture of 20 amino in an ethanol/water system compared to L-alanine as a function of ethanol solvent concentration (in g of ethanol / g of solvent) at 277.13 K.

4.3.9 Standard free energy of transfer

To aid comparison and since solubility is measured at equilibrium, the standard free energy of transfer, ΔG_{tr}° in $J \cdot mol^{-1}$ can be calculated. This is the relative solubility between an α -amino acid in a mixture at a lower ethanol concentration compared to its solubility in a mixture at a higher ethanol concentration. This has been calculated using:

$$\Delta G_{tr}^\circ = R/J \cdot mol^{-1} \cdot K^{-1} * T/K \ln \frac{m_l / \text{moles amino acid} \cdot Kg \ H_2O^{-1}}{m_h / \text{moles amino acid} \cdot Kg \ H_2O^{-1}} \quad (1)$$

and shown in Figure 4.27, Figure 4.28, Figure 4.29 and Figure 4.30. Where m_l is the molar concentration of the amino acid at maximum solubility in a mixture at the lower ethanol concentration, m_h is the molar concentration of the amino acid at maximum solubility in a mixture at the higher ethanol concentration, R is the gas constant in $J \cdot mol^{-1} \cdot K^{-1}$ and T is temperature in K.

At 277.15 K, ΔG_{tr}° is low for all amino acids except for between 50% and 70% ethanol. At 277.15 K and between 50% and 70%, the decrease in solubility of all the amino acids is lower in the lower ethanol concentration than the decrease in solubility in the higher ethanol concentration. The ΔG_{tr}° between the other concentrations of ethanol at 277.15 K are near zero, showing that the decrease between both ethanol concentrations are the same.

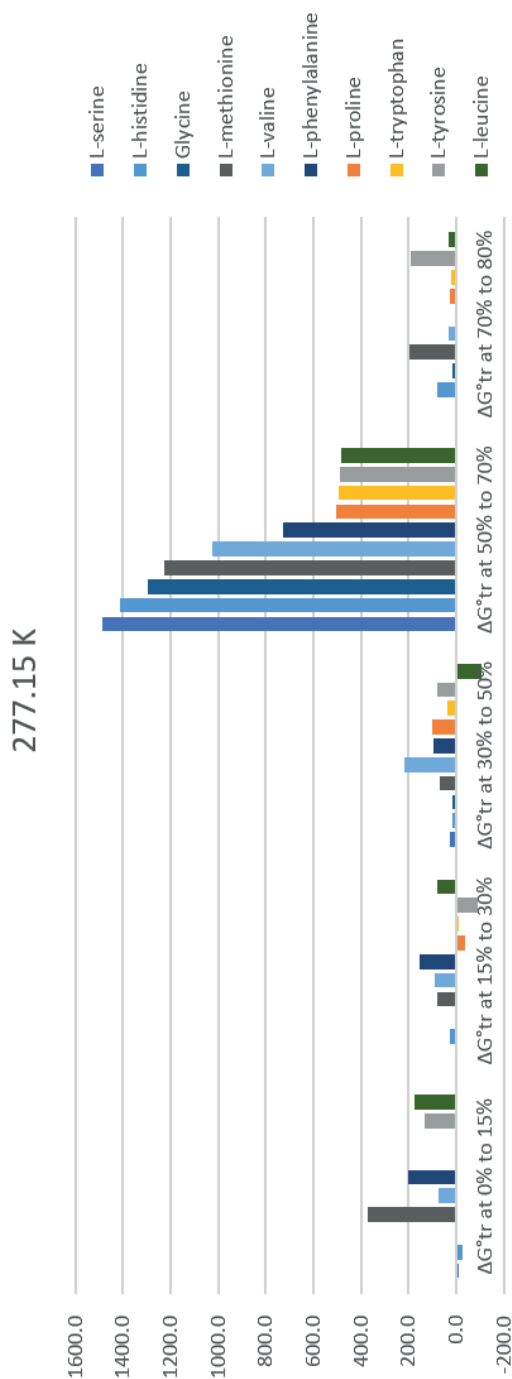


Figure 4.27: ΔG_{tr}° for each amino acid in a mixture at 277.15 K transferring from a lower ethanol solvent concentration (in g of ethanol / g of solvent) to a higher ethanol concentration as calculated in equation (1). At 277.15 K, elevated ΔG_{tr}° is seen between ethanol concentrations of 50 and 70% showing that the decrease in amino acid solubility in a mixture of amino acids is higher than in single solution. Aliphatic and non-phenyl amino acids (L-serine, L-histidine, glycine, L-methionine, L-valine, L-proline and L-leucine) show a higher decrease in solubility than phenyl amino acids (L-phenylalanine, L-tryptophan and L-tyrosine).

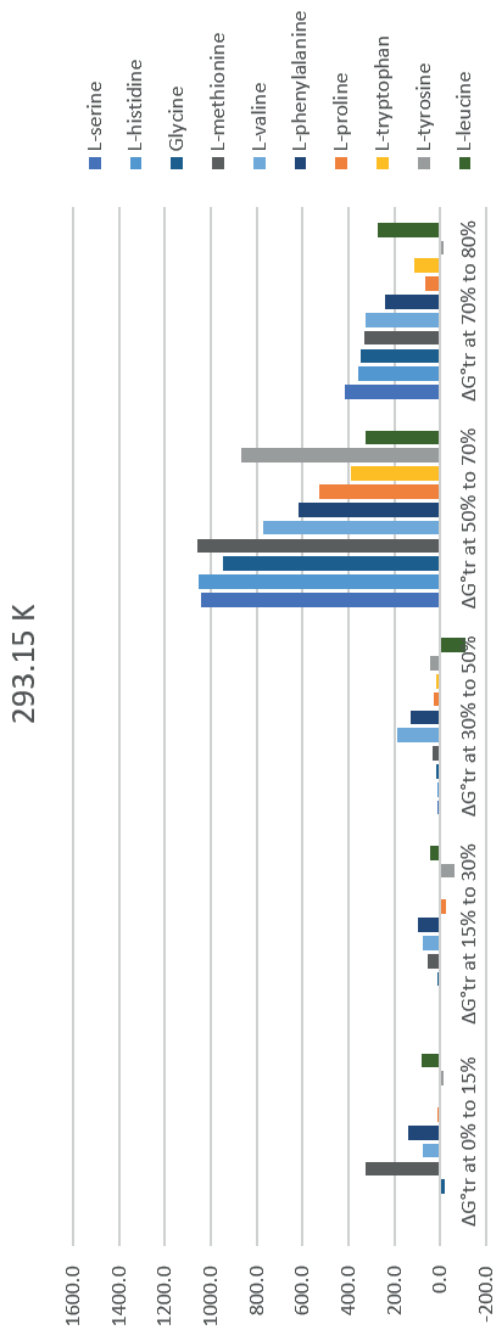


Figure 4.28 ΔG_{tr}° for each amino acid in a mixture at 293.15 K transferring from a lower ethanol solvent concentration (in g of ethanol / g of solvent) to a higher ethanol concentration as calculated in equation (1). At 293.15 K, elevated ΔG_{tr}° is seen between ethanol concentrations of 50 and 70% and again between ethanol concentrations 70 and 80% showing that the decrease in amino acid solubility in a mixture of amino acids is higher than in single solution. Aliphatic and non-phenyllic amino acids (L-serine, L-histidine, glycine, L-methionine, L-valine, L-proline and L-leucine) show a higher decrease in solubility than phenyllic amino acids (L-phenylalanine, L-tryptophan and L-tyrosine).

315.15 K

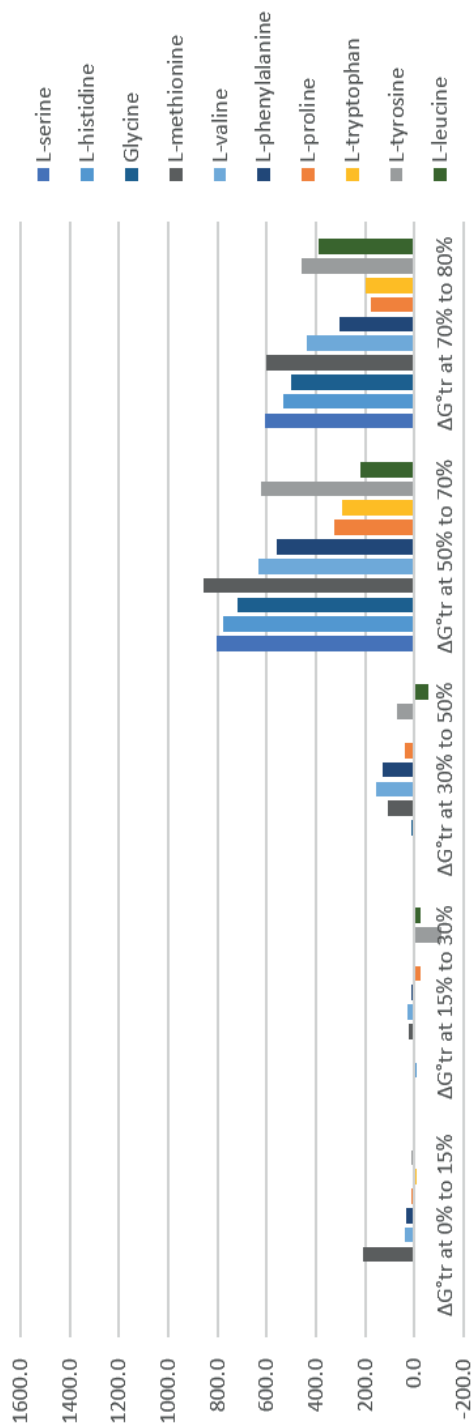


Figure 4.29: ΔG_{tr}° for each amino acid in a mixture at 315.15 K transferring from a lower ethanol solvent concentration (in g of ethanol / g of solvent) to a higher ethanol concentration as calculated in equation (1). At 315.15 K, elevated ΔG_{tr}° is seen between ethanol concentrations of 50 and 70% and again between ethanol concentrations 70 and 80% showing that the decrease in amino acid solubility in a mixture of amino acids is higher than in single solution. Aliphatic and non-phenylic amino acids (L-serine, L-histidine, glycine, L-methionine, L-valine, L-proline and L-leucine) show a higher decrease in solubility than phenylic amino acids (L-phenylalanine and L-tryptophan) with the exception of L-tyrosine (phenylic) and L-leucine (aliphatic).

335.15 K

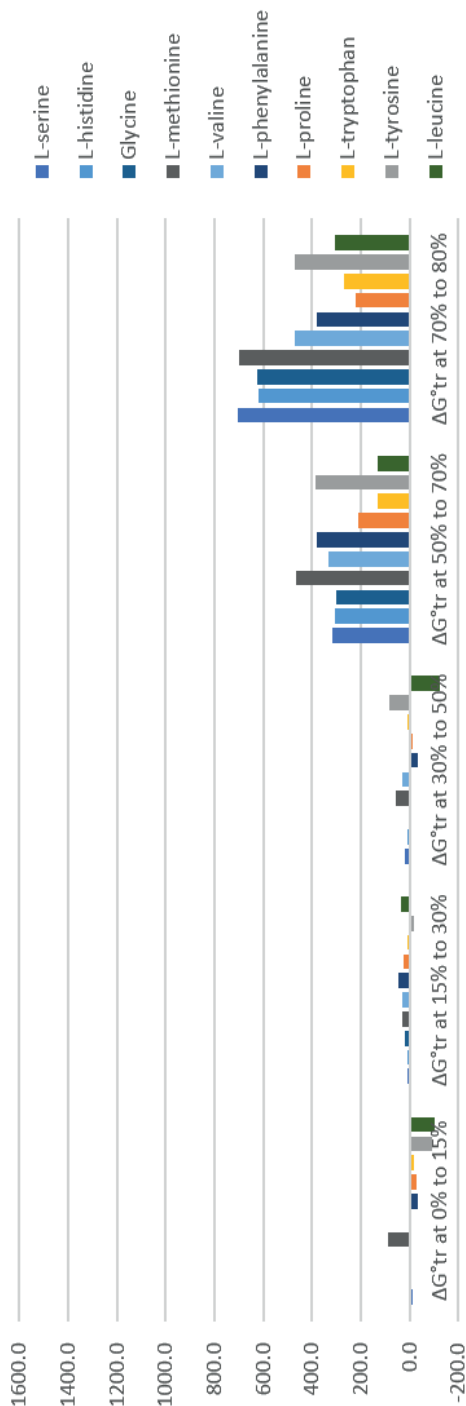


Figure 4.30: ΔG_{tr}° for each amino acid in a mixture at 335.15 K transferring from a lower ethanol solvent concentration (in g of ethanol / g of solvent) to a higher ethanol concentration as calculated in equation (1) . At 335.15 K, elevated ΔG_{tr}° is seen between ethanol concentrations of 50 and 70% and again between ethanol concentrations 70 and 80% showing that the decrease in amino acid solubility in a mixture of amino acids is higher than in single solution. Aliphatic and non-phenylic amino acids (L-serine, L-histidine, glycine, L-methionine, L-valine, L-proline and L-leucine) show a higher decrease in solubility than phenylic amino acids (L-phenylalanine and L-tryptophan) with the exception of L-tyrosine (phenylic) and L-leucine (aliphatic).

At 293.15 K, ΔG_{tr}° is low for all amino acids between 0% to 15% ethanol and 15% to 30% ethanol. Showing that the decrease in solubility is similar at all of these ethanol concentrations. Between 50% and 70% ethanol, the decrease in solubility of all the amino acids is lower in the lower ethanol concentration than the decrease in solubility in the higher ethanol concentration. Between 70% and 80%, the ΔG_{tr}° is lower for all amino acids than between 50% and 70%, but higher than the lowest ethanol concentrations. At 315.15 K, ΔG_{tr}° is still low for the ethanol concentrations from 0% to 50%. However, at 315.15 K, the ΔG_{tr}° is approximately equal for the ethanol concentrations 50% to 70% ethanol and between 70% to 80%. At 335.15 K, ΔG_{tr}° is still low for the ethanol concentrations from 0% to 50%. However, at 335.15 K, the ΔG_{tr}° is lower for the 50% to 70% ethanol concentrations and higher for the 70% to 80%. This shift in ΔG_{tr}° shows a trend that as the temperature increases, the difference in solubility between the increasing ethanol concentration also increases.

4.4 Conclusions

We can group the amino acids according to their side chains and draw conclusions. These groups are branched aliphatic, hydroxylic, phenylic, sulphurous, amide/amine containing, carboxylic acid containing and heterocyclic.

An analysis of the solubility data brings a few general conclusions. Amino acids with branched aliphatic side chains have a lower solubility when in a mixture of amino acids than in solution alone. The exception for this is when the side chain includes 3 methylene groups or less. Having a carboxylic acid, hydroxylic or phenylic group in the side chain of the amino acid leads to an increase in the solubility of the amino acid when in a mixture of amino acids and ethanol. Having a side chain that contains Sulphur or an amide and/or an amine group in their side chain leads to a lower solubility when in a group of 20 amino acids at all ethanol concentrations than in solution alone unless the side chain is short (e.g. L-asparagine).

For all amino acids in solution by themselves, the decrease in solubility for each incremental increase in ethanol (e.g. from 10% ethanol to 20% ethanol versus 20% ethanol to 30% ethanol) decreases and this is shown by a trend line that is concave upwards. However, when L-aspartic acid, L-glutamic acid, L-lysine, L-arginine, L-methionine, L-phenylalanine, L-

tryptophan, L-tyrosine, L-isoleucine, L-valine, L-leucine and L-alanine are in a mixture of 20 amino acids, the decrease in solubility for each incremental increase in ethanol concentration increased. This is shown in downward concave trend lines.

The ΔG_{tr}° shows that the greatest decrease in amino acid solubility is at 50-70% (g/g) ethanol/water solutions at 277.15 K. As the temperature of the solution increases, the ΔG_{tr}° shows a shift in the point at which the solubility decreases to higher ethanol concentrations. This indicates that as more energy is added to the system, the amino acid solubility also increases.

References

- [1] Scott E.L., Peter F., Sanders J.P.M. Biomass in the Manufacture of Industrial Products – The use of Proteins and Amino Acids. *Applied Microbiology and Biotechnology* **2007** 75:751-62
- [2] Lammens T.M., Franssen M.C.R., Scott E.L., Sanders J.P.M. Availability of protein-derived amino acids as feedstock for the production of bio-based chemicals. *Biomass and Bioenergy* **2012** 44:168-81
- [3] Lammens T.M. Ph.D. Thesis 2011. Bio-based Industrial Chemicals from Glutamic Acid.
- [4] Cohn E.J., Edsall J.T. **1943** Proteins Amino Acids and Peptides as Ions and Dipolar Ions.
- [5] Cohn E.J., McMeekin T.L., Edsall J.T., Weare J.H. Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. II. The Solubility of α -Amino Acids in Water and in Alcohol—Water Mixtures. *Journal of the American Chemical Society* **1934** 56:2270-82
- [6] Cohn E.J., McMeekin T.L., Ferry J.D., Blanchard M.H. Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. XII. Interactions between Dipolar Ions in Aqueous Solution. *Journal of Physical Chemistry* **1939** 43
- [7] McMeekin T.L., Cohn E.J., Weare J.H. Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. III. The Solubility of Derivatives of the Amino Acids in Alcohol—Water Mixtures. *Journal of the American Chemical Society* **1935** 57:626-33
- [8] Dalton J.B., Schmidt C.L.A. The Solubilities of Certain Amino Acids in Water, the Densities of their Solutions at 25°, and the Calculated Heat of Solution and Partial Molal Volumes. *Journal of Biological Chemistry* **1933** 103:549-78
- [9] Dalton J.B., Schmidt C.L.A. The Solubilities of Certain Amino Acids in Water, the Densities of their Solutions at 25°, and the Calculated Heat of Solution and Partial Molal Volumes. Volume II. *The Journal of Biological Chemistry* **1935** 109:8
- [10] Dunn M.S., Ross F.J., Read L.S. The Solubility of the Amino Acids in Water. *Journal of Biological Chemistry* **1933** 103:579-95
- [11] Kyte J., Doolittle R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982** 157:105-32

- [12] Wolfenden R., Andersson L., Cullis P.M., Southgate C.C.B. Affinities of amino-acid side-chains for solvent water. *Biochemistry* **1981** 20:849-55
- [13] Rose G., Geselowitz A., Lesser G., Lee R., Zehfus M. Hydrophobicity of amino acid residues in globular proteins. *Science* **1985** 229:834-8
- [14] Janin J. Surface and inside volumes in globular proteins. *Nature* **1979** 277:491-2
- [15] Nozaki Y., Tanford C. The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions: Establishment of a hydrophobicity scale. *Journal of Biological Chemistry* **1971** 246:2211-7
- [16] Pradhan A.A., Vera J.H. Effect of acids and bases on the solubility of amino acids. *Fluid Phase Equilibria* **1998** 152:121-32
- [17] Tseng H.-C., Lee C.-Y., Weng W.-L., Shiah I.M. Solubilities of amino acids in water at various pH values under 298.15 K. *Fluid Phase Equilibria* **2009** 285:90-5
- [18] Daldrup J.-B.G., Held C., Sadowski G., Schembecker G. Modeling pH and Solubilities in Aqueous Multisolute Amino Acid Solutions. *Industrial & Engineering Chemistry Research* **2011** 50:3503-9
- [19] Dunn M.S., Ross F.J. Quantitative Investigations of Amino Acids and Peptides: IV. The Solubilities of the Amino Acids in Water-Ethyl Alcohol Mixtures. *Journal of Biological Chemistry* **1938** 125:309-32
- [20] Needham T.E., Paruta A.N., Gerraugh Rj. Solubility of Amino Acids in Pure Solvent Systems. *Journal of Pharmaceutical Sciences* **1971** 60:565-
- [21] Conio G., Curletto L., Patrone E. Temperature coefficient of solubility of glycyl peptides in water-ethanol mixtures. *Journal of Biological Chemistry* **1973** 248:5448-50
- [22] Gude M.T., Meuwissen H.H.J., van der Wielen L.A.M., Luyben K.C.A.M. Partition Coefficients and Solubilities of α -Amino Acids in Aqueous 1-Butanol Solutions. *Industrial & Engineering Chemistry Research* **1996** 35:4700-12
- [23] van Berlo M., Gude M.T., van der Wielen L.A.M., Luyben K. Partition coefficients and solubilities of glycine in the ternary solvent system 1-butanol plus ethanol plus water. *Industrial & Engineering Chemistry Research* **1997** 36:2474-82
- [24] Gude M.T., van der Wielen L.A.M., Luyben K.C.A.M. Phase behavior of α -amino acids in multicomponent aqueous alkanol solutions. *Fluid Phase Equilibria* **1996** 116:110-7

- [25] Carta R. Solubilities of l-Cystine, l-Tyrosine, l-Leucine, and Glycine in Their Water Solutions. *Journal of Chemical & Engineering Data* **1999** 44:563-7
- [26] Kamei T., Hasegawa K., Kashiwagi T., Suzuki E., Yokota M. Solid-Liquid Equilibria in an L-Isoleucine plus L-Alanine plus Water System. *Journal of Chemical and Engineering Data* **2008** 53:2801-6
- [27] Kurosawa I., Teja A.S., Rousseau R.W. Solid-liquid equilibria in L-leucine plus L-valine plus water. *Fluid Phase Equilibria* **2004** 224:245-9
- [28] Dhondge S.S., Paliwal R.L., Bhawe N.S., Pandhurnekar C.P. Study of thermodynamic properties of aqueous binary mixtures of glycine, L-alanine and beta-alanine at low temperatures (T=275.15, 279.15, and 283.15) K. *Journal of Chemical Thermodynamics* **2012** 45:114-21
- [29] Meussen B., van Zeeland A.T., Bruins M.E., Sanders J.P.M. A Fast and Accurate UPLC Method for Analysis of Proteinogenic Amino Acids. *Food Anal. Methods* **2014** 7:1047-55

Chapter 5

Modelling the effects of ethanol on the solubility of the proteinogenic amino acids with the NRTL, Gude and Jouyban-Acree models

Published as: Bowden, N.A., Mendez Sevilano, D., Sanders, J.P.M. and Bruins, M.E. Modelling the effects of ethanol on the solubility of the proteinogenic amino acids with the NRTL, Gude and Jouyban-Acree models *Fluid Phase Equilibria* **2017** DOI: 10.1016/j.fluid.2017.11.036

Abstract:

The addition of organic solvents, such as ethanol, to molecules in solution is an effective process for crystallization and is used in industrial settings (i.e. pharmaceutical production, downstream processing, etc.). In this study, we use solubility data of all proteinogenic α -amino acids in binary ethanol/water systems to model their excess solubility. We use the empirical and regressive models of Gude and NRTL and the predictive Jouyban-Acree model. Based on the results, we hypothesize that amino acids that are spherical and lack a reactive side chain show little or no excess solubility. Being rod-like and/or having a reactive side chain leads to a positive excess solubility in a mixed solvent of ethanol and water. The empirical and regressed models, NRTL and Gude, fit the data well and the predictive Jouyban-Acree model, not originally intended to be used for small molecules, is less accurate but offers insights into the thermodynamic properties of the amino acids.

5.1 Introduction

In the future, products that are currently being produced using non-renewable resources (e.g. plastics, pharmaceuticals and fine chemicals) could be made from bio-based sources, such as proteins and α -amino acids [1-3]. One of the challenges in this line of research, is to find a way to separate α -amino acids from industrial residues so that the production of bio-based products can begin. This research is applicable to the industrial challenges of separating amino acids from solution.

Industrial residues can be used as a feedstock for the extraction of amino acids and other biomolecules. When amino acids are extracted, they need to be separated from aqueous solution. Currently, the most common method of separating many amino acids from solution is by using industrial chromatography. An alternative to chromatography could be to crystallize the amino acids using an anti-solvent, such as ethanol.

The structure of every amino acid contains a carboxyl group attached to an α -carbon. This α -carbon is also attached to an amino group. The amino acids studied in this article are α -amino acids, which all have side chains also attached to the α -carbon. The exception is glycine which does not have a side chain. The side chains of α -amino acids include aliphatic groups, aromatic and non-aromatic rings, hydroxyl groups, sulphur and charged groups (e.g. a second carboxyl group, lysyl group, guanidinium group). The amino and carboxyl groups attached to the α -carbon will be charged at a pH that is not the isoelectric point. At the isoelectric point, the amino acid has a neutral charge and is called a zwitterion. All measurements in this manuscript were taken at the isoelectric point.

There has been some research on the solubility of α -amino acids in mixtures of alcohol and water [4-7]. Basic solubility measurements were reported and subsequent research focused on calculating the partition coefficients of the solubility of these α -amino acids and their phase behavior [8]. Recently, complete and reliable data has been published on the solubility of α -amino acids in ethanol/water systems [9] and mixtures of α -amino acids [10].

Many models have been proposed to model the solubility of amino acids in aqueous solution. These models include calculating partition coefficients [11], using regressed coefficients [12], examining non-ideality [13], measuring and modelling activity coefficients [14-17], activities [18] and applying a modification of the Wilson model [19]. Other models have been applied to model the solubility of amino acids in salt solutions [20-27]. Only a few models

have been proposed to describe the solubility of α -amino acids in ethanol/water systems, but these manuscripts focus on a single model and only a few α -amino acids [28-30]. This article will model all proteinogenic α -amino acids using solubility data that is available in the literature.

We use three models that represent two different modelling approaches. Of these three, two of the models use regressed parameters. The models that we use that have regressed parameters are the Gude model and the Non-Random Two Liquid (NRTL) model. While models that use regressed parameters have in general given excellent results, they do not explain what thermodynamic properties of the molecules lead to their results. The third model that we use is the Jouyban-Acree model, which is a predictive model. Predictive solubility models are based on thermodynamic properties of the molecules that they are modelling. While the thermodynamic properties of the molecules explain the results of the predictive models, predictive models have been less accurate than regressed models.

Using the different approaches allows conclusions to be made on whether the predictive model (Jouban-Acree) provides sufficient accuracy to model amino acid solubility or if a regressed model (Gude or NRTL) should be used. Other solubility models [31-36] were considered for this article, but due to their complexity were left out in favour of models with fewer variables.

The Gude [12] and NRTL [54] models were chosen in this research for their accuracy in the literature and the minimum number of parameters they use. Both the NRTL and Gude models furthermore acknowledge the lattice and therefore entropic nature of liquids, first investigated by Flory [37] and Huggins [38]. The Gude model has one parameter that is regressed to fit the data and the NRTL has two parameters that are regressed to fit the data. For this reason, it is expected that the NRTL model will have a lower error. However, it is preferential to use a regressive model with the least number of regressed parameters. In the case where both models have similar errors, the Gude model could be used.

While the Gude and NRTL models will be accurate, in comparison, the Jouyban-Acree model is predictive and based on the bonds and forces of the molecules being modelled. The version of the Jouyban-Acree model that is used in this research has nine regressed constants. These constants are used in conjunction with Hansen solubility parameters, which are based on physical chemistry group contribution data. While the Jouyban-Acree model uses more parameters than the Gude and NRTL models, the parameters are predictive, not regressed. The

Jouyban-Acree model has been shown to perform well with relatively large pharmaceutical solutes in ternary systems [39]. A version of this model with regressed parameters has been applied to only a few amino acids in ternary solution, but no α -amino acids in water and ethanol mixtures, with the exception of glycine [40]. We use the Jouyban-Acree model without regressed parameters in this research in order to evaluate the use of group contribution data to amino acid solubility models. In the future, data from this work could contribute to refining the non-regressed Jouyban-Acree parameters for amino acids.

5.2 Theory

5.2.1 Thermodynamic modelling of excess solubility

The addition of organic solvents, e.g. ethanol, to aqueous solutions of amino acids lowers the solubility of the amino acid solutes. This allows for precipitation and crystallization. The solubility of the amino acids is often lowered by organic solvents by more than 1000 times its solubility in water alone [9]. Industrial applications using organic solvents can only be designed when this effect on the solubility is understood. This presents a challenge for chemical engineers in modelling their solubility.

Data is taken from the literature [4-7,9] and modelled with two empirical and regressive models and with one predictive model. The two empirical and regressive models are the Gude [12] and NRTL [41-45] models and the semi-empirical and predictive model is the Jouyban-Acree model [46-50].

In order to effectively compare the performance of the models, excess solubility has been chosen as the output of the model. This decision aligns with literature [51-52] in the specific case of binary solvent mixtures. Excess solubility, represented by the mole fraction x_{aa}^E , can be calculated using equation (5.1).

$$\ln x_{aa}^E \equiv \ln x_{aa,mix} - \sum_{i=1}^N x'_i \ln x_{aa,i} \quad (5.1)$$

in which case $x_{aa,mix}$ and $x_{aa,i}$ are the mole fractions of the amino acid solute (aa) in a mixed solvent and pure solvent, i , respectively. The mole fraction of the solvent i without the solute is denoted by x'_i .

When assuming a pure solvent phase as a standard state, such as in this research, at standard system pressure and temperature, the chemical potential of the solute is not dependent on the solvent composition. Therefore, the excess solubility can be rewritten as:

$$\ln x_{aa}^E \equiv -\ln \gamma_{aa,mix} + \sum_{i=1}^N x'_i \ln \gamma_{aa,i} \quad (5.2)$$

where the dimensionless activity coefficients of the solute in saturated solutions of the mixed solvent and pure solvent are represented by $\gamma_{aa,mix}$ and $\gamma_{aa,i}$.

Cohn and Edsall [53] noted that the solubility of the solute in these systems is low. Therefore, it can be assumed that the solute is infinitely dilute and approximated as:

$$\ln x_{aa}^E \equiv -\ln \gamma_{aa,mix}^{\infty} + \sum_{i=1}^N x'_i \ln \gamma_{aa,i}^{\infty} \quad (5.3)$$

5.2.2 Gude Model

Gude [6] developed a simplified equation to model the behaviour of amino acids in mixed solvents. This model uses 2 constants. The constant for the interaction between the solvents, $A_{j,i}$, was set to 1.55 for ethanol/water in the work of Gude and is applied in this work. The constant for the interaction between the amino acid and the solvent mixture, $C_{j,i,aa}$, is specific to each amino acid. This interaction parameter, $C_{j,i,aa}$ (mol·L⁻¹), is constant for the system and found by fitting the model to the data. Equation (4) describes the model:

$$\ln x_{aa}^E \equiv \ln r' - \sum_{j=1}^N x'_j \ln r_j + r_{aa} \left(\frac{1}{r'} - \sum_j \frac{x'_j}{r_j} \right) + \sum_j \sum_i [A_{j,i} x'_j x'_i (1 + C_{j,i,aa})] \quad (5.4)$$

where subscripts j and i relate to solvents and subscript aa relates to the solute. The values of the UNIFAC variable r were set at 0.92 for water and 2.11 for ethanol and calculated individually for the amino acids [12]. Values for r' are the solute free value of r . The $C_{j,i,aa}$ are fitted for each amino acid from Equation (5.4) and are shown in Table 5.2.

5.2.3 NRTL Model

Based on the hypothesis of Wilson, that the local concentration of solvent molecules in a two-solvent system around a molecule of the solute are not the same as the concentration in the solution in general, Renon and Prausnitz [54] developed the NRTL model to calculate the interaction parameters between these molecules. In the case of this research, the mixed solvent is comprised of only two solvents, so the activity coefficient $\gamma_{aa,i}$ equation (5.5):

$$\ln \gamma_{aa,i} = \frac{\sum_{i=1}^n x'_i \tau_{i,aa} G_{i,aa}}{\sum_{i=1}^n x'_i G_{i,aa}} + \sum_{i=1}^n \frac{x'_i G_{aa,i}}{\sum_{j=1}^n x'_j G_{j,i}} \left(\tau_{aa,i} - \frac{\sum_{j=1}^n x'_j \tau_{j,i} G_{j,i}}{\sum_{j=1}^n x'_j G_{j,i}} \right) \quad (5.5)$$

can be substituted with the NRTL equation, which yields equation (6):

$$\ln x_{aa}^E = \sum_{i=1}^N (\tau_{i,aa} + \tau_{i,aa} G_{i,aa}) x'_i - \frac{\sum_{i=1}^n x'_i \tau_{i,aa} G_{i,aa}}{\sum_{i=1}^n x'_i G_{i,aa}} - \sum_{i=1}^n \frac{x'_i G_{aa,i}}{\sum_{j=1}^n x'_j G_{j,i}} \left(\tau_{aa,i} - \frac{\sum_{j=1}^n x'_j \tau_{j,i} G_{j,i}}{\sum_{j=1}^n x'_j G_{j,i}} \right) \quad (5.6)$$

where $G_{mn} = \exp(-\alpha_{mn} \tau_{mn})$ and the dimensionless interaction parameters τ_{mn} , τ_{nm} and the non-randomness parameter α_{nm} are represented for each system of two solvents.

The interaction parameters, τ , and the non-randomness parameters, α , for the solvents have previously been published⁴². These are $\tau_{ethanol,water} = -406.47$ and $\tau_{water,ethanol} = 1413$ at 298.15K, $\alpha_{water,ethanol} = 0.1830$ and $\alpha_{aa,water} = 0.05$ and $\alpha_{aa,ethanol} = 0.02$. Furthermore, in this research we have assumed that the unitless interaction parameters for the system amino acid-solvent, $\tau_{aa,i}$, and solvent-amino acid, $\tau_{i,aa}$, are the same. The $\tau_{aa,i}$ for each amino acid is calculated by regressing Equation (6) and are shown in Table 5.2.

5.2.4 Jouyban-Acree Model

Jouyban and colleagues developed a model for the excess solubility [39] based on the log-linear model developed by the group of professor Sadowski [32]. This model uses as input the Hansen solubility parameters which can be calculated from group contribution models [55].

There are several versions of the Jouyban-Acree model. The version that we use here [49], shown in equation (7), uses nine previously regressed constants that can be found in Table 5.1 to calculate the solubility in the mixture of solvents. Once that is calculated, equation (5.1) can be used to calculate the excess solubility and compare the performance with the aforementioned models.

$$\begin{aligned} \log x_{aa,mix} = & f_c \log x_{aa,c} + f_w \log x_{aa,w} + \left(\frac{f_c f_w}{T}\right) \left[A_0 \delta_{d,aa} (\delta_{d,c} - \delta_{d,w})^2 + \right. \\ & A_1 \delta_{p,aa} (\delta_{p,c} - \delta_{p,w})^2 + A_2 \delta_{hb,aa} (\delta_{hb,c} - \delta_{hb,w})^2 \left. \right] + \left(\frac{f_c f_w (f_c - f_w)}{T}\right) \left[A_3 \delta_{d,aa} (\delta_{d,c} - \delta_{d,w})^2 + \right. \\ & A_4 \delta_{p,aa} (\delta_{p,c} - \delta_{p,w})^2 + A_5 \delta_{hb,aa} (\delta_{hb,c} - \delta_{hb,w})^2 \left. \right] + \left(\frac{f_c f_w (f_c - f_w)^2}{T}\right) \left[A_6 \delta_{d,aa} (\delta_{d,c} - \delta_{d,w})^2 + \right. \\ & A_7 \delta_{p,aa} (\delta_{p,c} - \delta_{p,w})^2 + A_8 \delta_{hb,aa} (\delta_{hb,c} - \delta_{hb,w})^2 \left. \right] \end{aligned} \quad (5.7)$$

Where subscripts w, c, p, d and hb stand for water, co-solvent, polar, dispersion and hydrogen bonding respectively. Furthermore, δ and f stand for the Hansen solubility parameter, in MPa^{0.5}, and volume fraction respectively. The Hansen solubility parameters were calculated as discussed previously and are shown in Table 5.2. The solubility parameters are constant and could be included in the A values. The A parameters show the effect of the forces in the solvent system on the amino acid. In this case, the solvent system is water and ethanol. The solubility parameters, A₀-A₈, are shown in Table 5.1.

Table 5.1: Jouyban-Acree constants

Constant	Value
A ₀	0.0000
A ₁	0.6060
A ₂	0.0130
A ₃	-8.6960
A ₄	0.3760
A ₅	0.0130
A ₆	9.2770
A ₇	-0.4610
A ₈	0.0170

5.3 Materials and Methods

Matlab version 9.0.0341360 was used for the regression and calculations. All graphical objects in Figure 5.22 were obtained from Wikimedia and have been released to the public domain worldwide.

The data from the literature that is used in all of the models is shown in the supplementary data. In this table, the solubility of each of the 20 proteinogenic amino acids in mole fraction is given, along with the ethanol mole fraction in the solvent without the solute, the standard deviation (labelled “+/-”) and the source of the data. The standard deviation was calculated by the root of the sum of the square of the difference between each of the measurements and the average of the measurements, divided by the number of measurements minus one. All data were measured at the isoelectric point. This means that the amino acids are present as neutral zwitterions and therefore carry no net charge.

The interaction parameters of the NRTL and Gude models are regressed by minimizing the normalized root-mean-square error (NRMSE). The NRMSE was calculated for all three models by equation (8), where x'_i is the mole fraction of ethanol in the solute free solvent, $\hat{y}_{x'_i}$ is the predicted excess solubility, $y_{x'_i}$ is the measured excess solubility and y_{max} and y_{min} are the maximum and minimum excess solubility. Normalizing the root-mean-square-error by dividing by the range facilitates the comparison between amino acids that are on different scales.

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^n (\hat{y}_{x'_i} - y_{x'_i})^2}{n}}}{y_{max} - y_{min}} \quad (5.8)$$

5.4 Results and discussion

The regression coefficients, $\tau_{aa,i}$ and $\tau_{i,aa}$, of the NRTL model for the interaction between the amino acid and ethanol and the amino acid and water are shown in Table 5.2. The regression coefficients of the Gude model for each amino acid, $C_{j,i,aa}$, are also shown in Table 5.2. These coefficients were calculated by minimizing the NRMSE of the excess solubility values that were modelled to the excess solubility measured. The Jouyban-Acree parameters that were calculated are shown in Table 5.2.

The modelled fits of the Gude and NRTL models and the application of the Jouyban-Acree model are shown along with the data points in Figure 5.20-5.20 for all 20 proteinogenic amino acids. If the standard deviation of the data was available, this was included in the figures. If multiple data were available for ethanol mole fractions of 0.000 and 1.000, then preference was given to the data that has been shown to be more accurate [5]. A fit where the excess solubility was equal to 0 was added to each of the amino acids in Figure 5.20-5.20 to guide the eye.

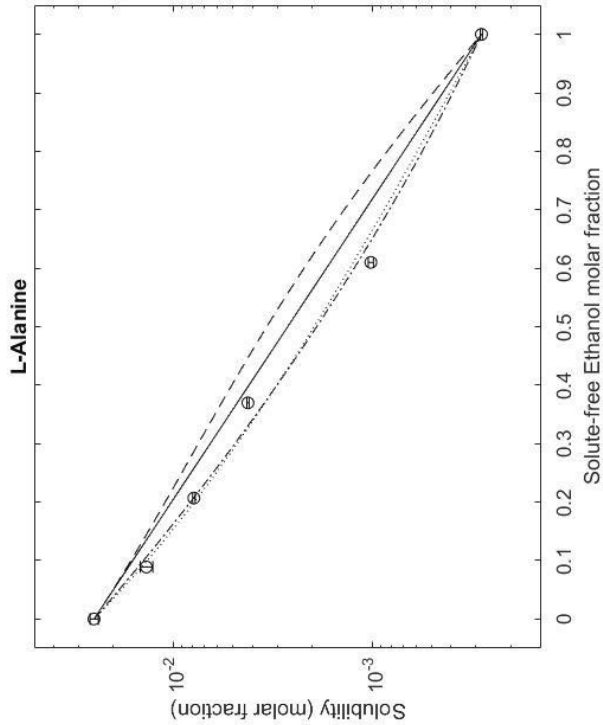


Figure 5.1: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

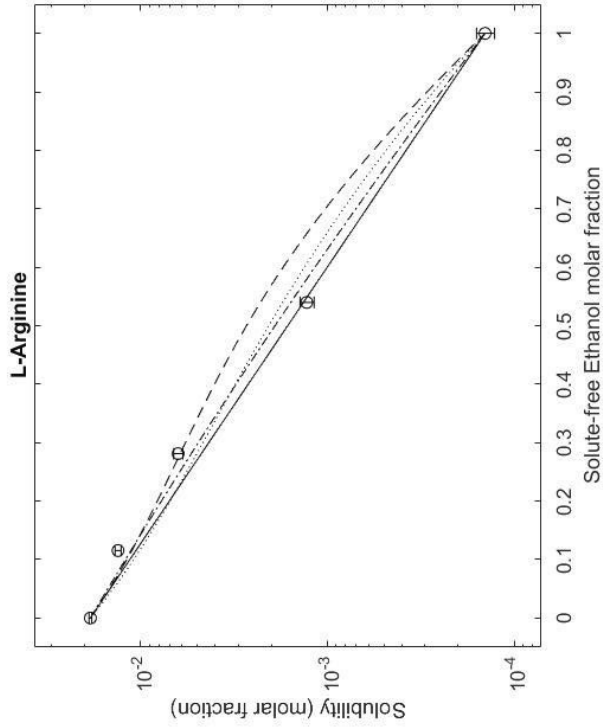


Figure 5.2: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

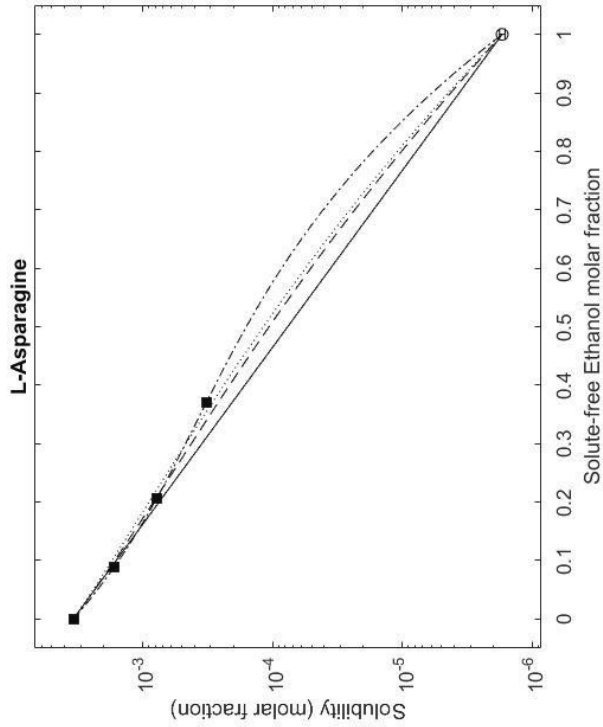


Figure 5.3: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

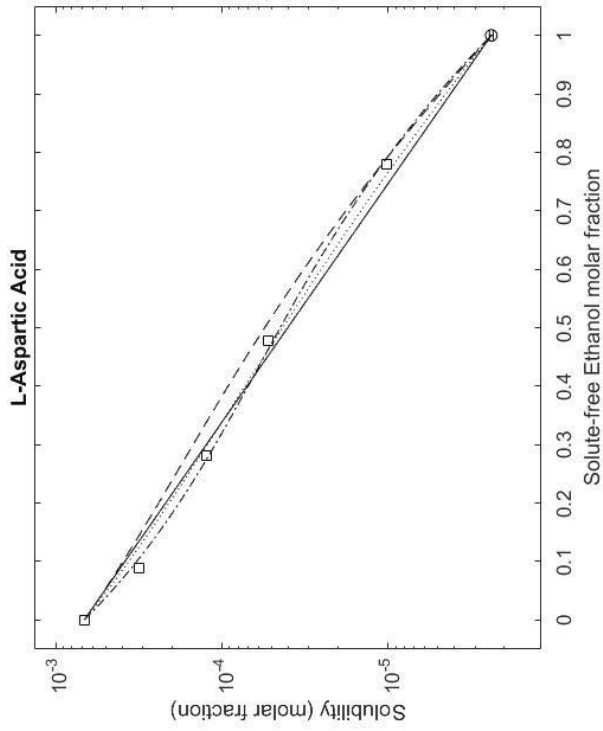


Figure 5.4: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

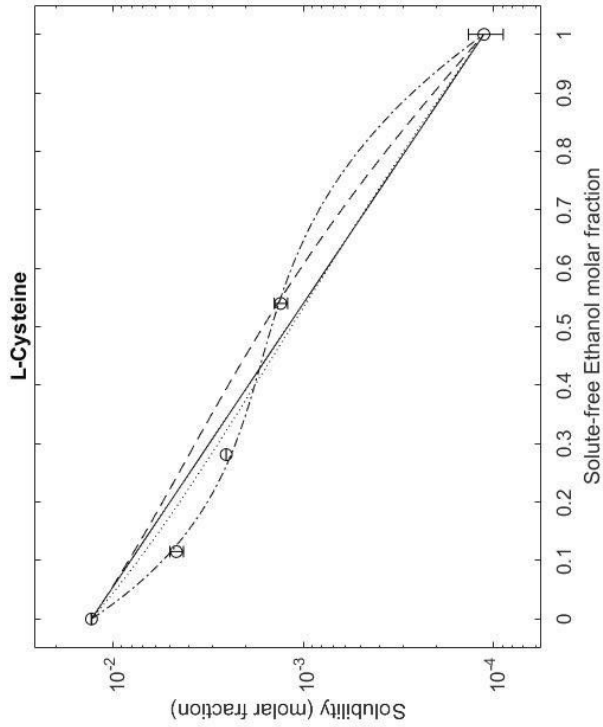


Figure 5.5: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

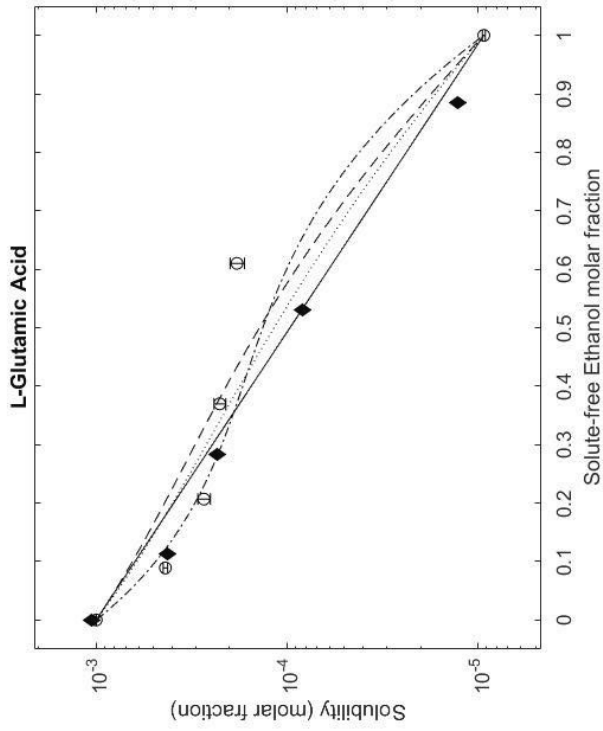


Figure 5.6: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

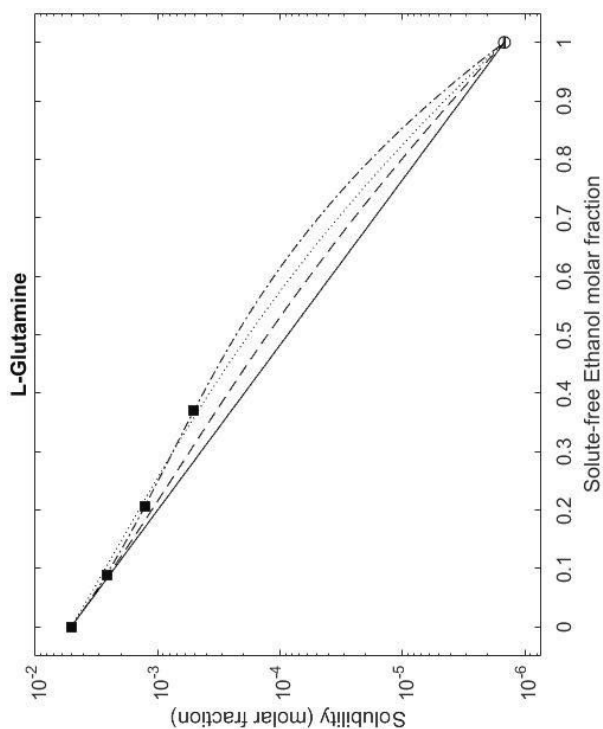


Figure 5.7: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

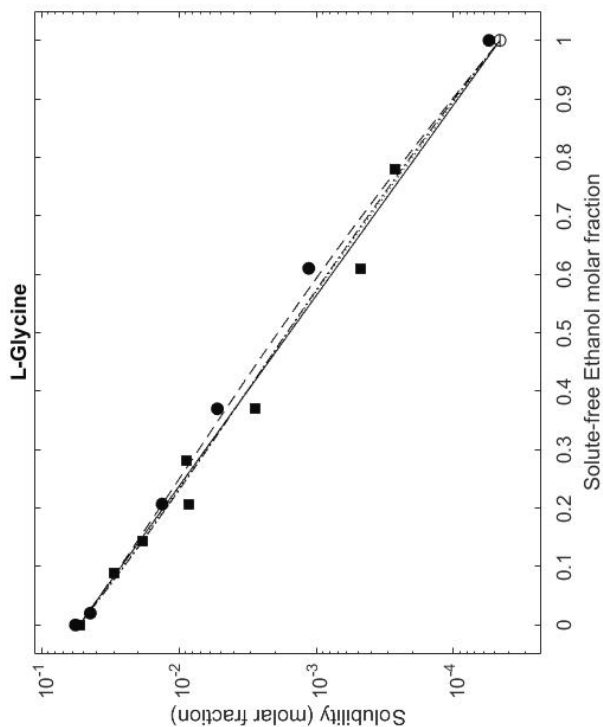


Figure 5.8: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

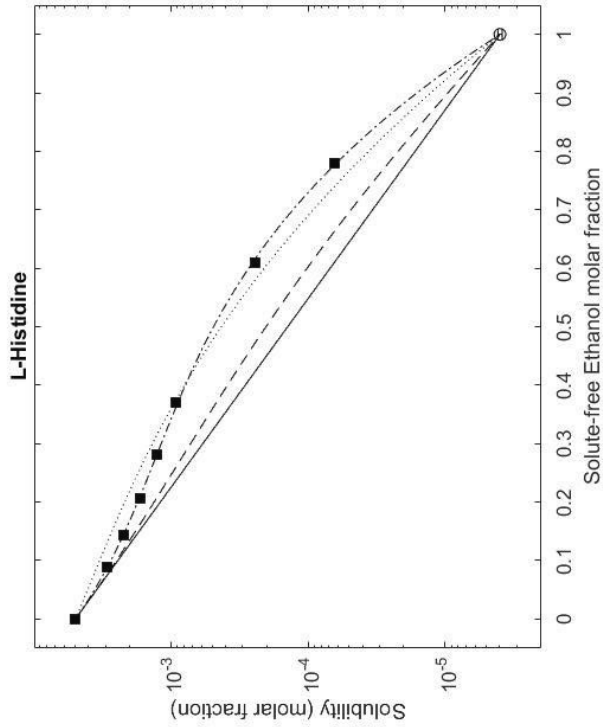


Figure 5.9: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

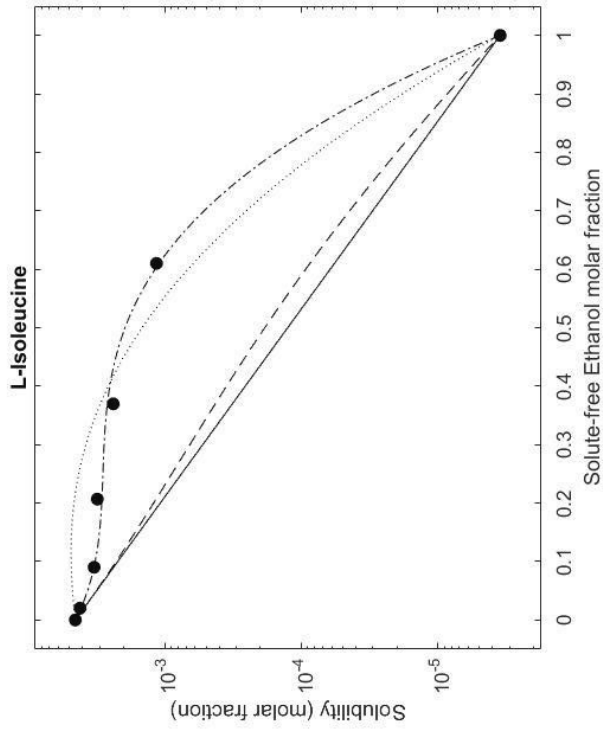


Figure 5.10: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

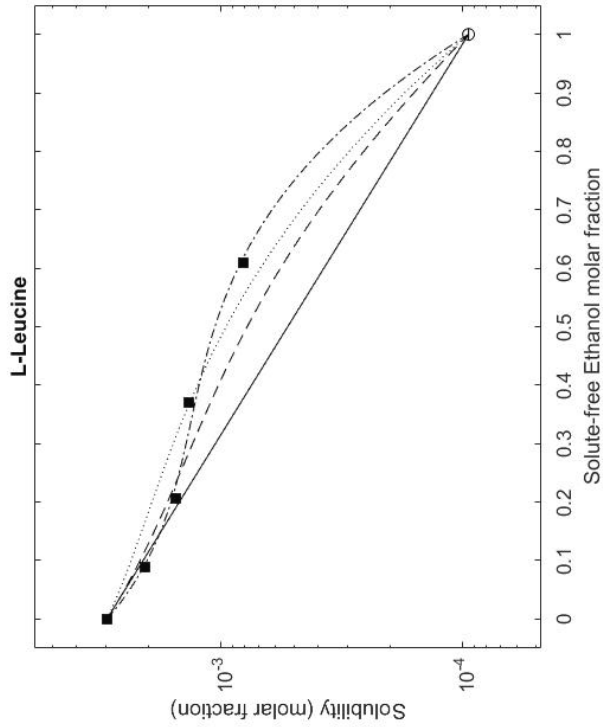


Figure 5.11: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

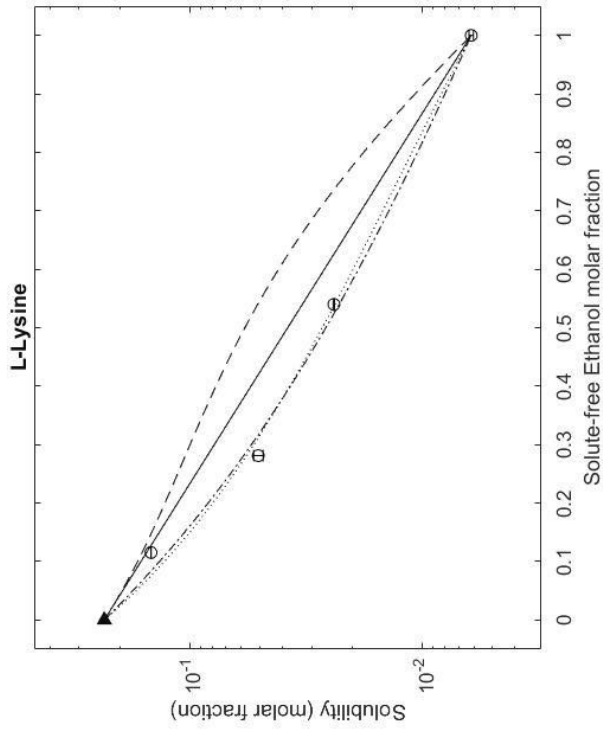


Figure 5.12: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

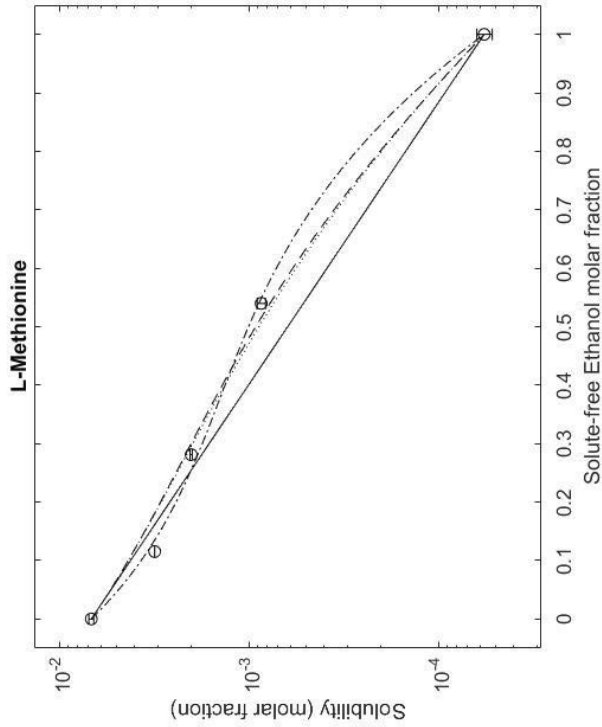


Figure 5.13: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

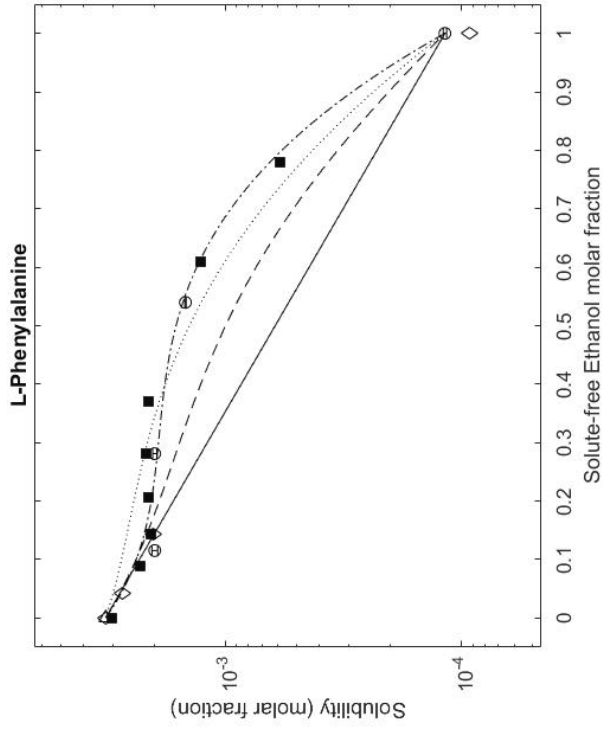


Figure 5.14: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

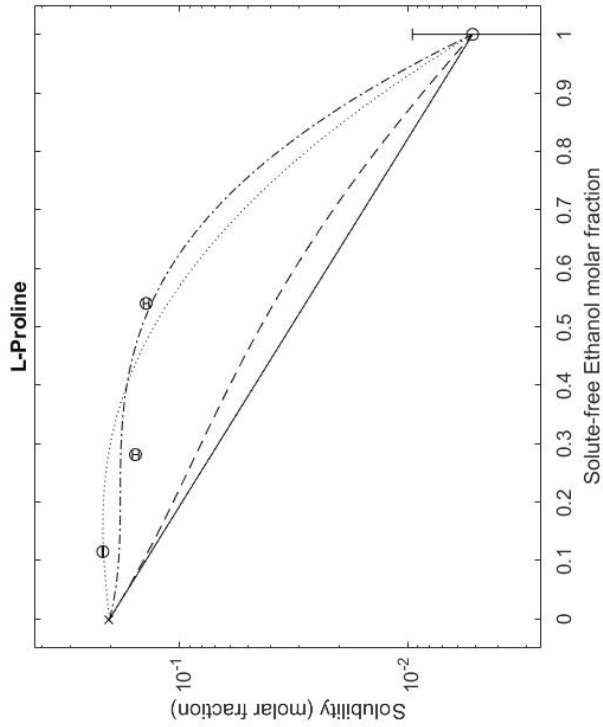


Figure 5.15: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

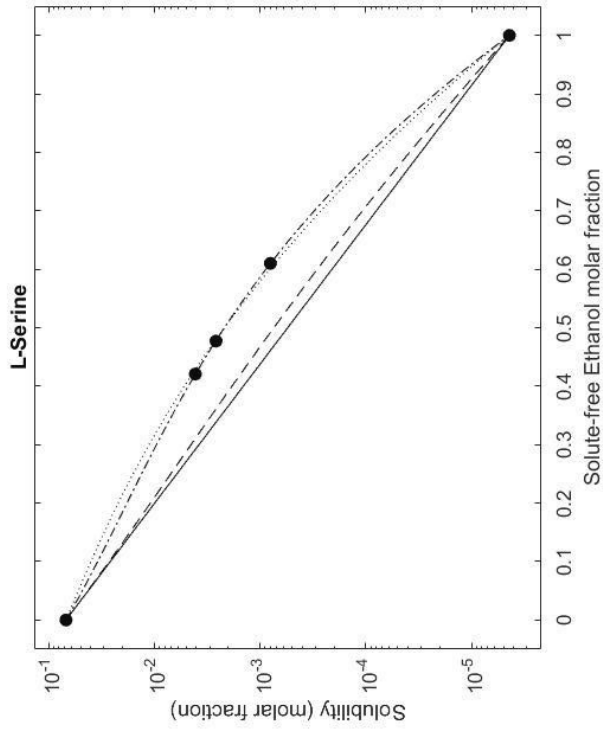


Figure 5.16: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

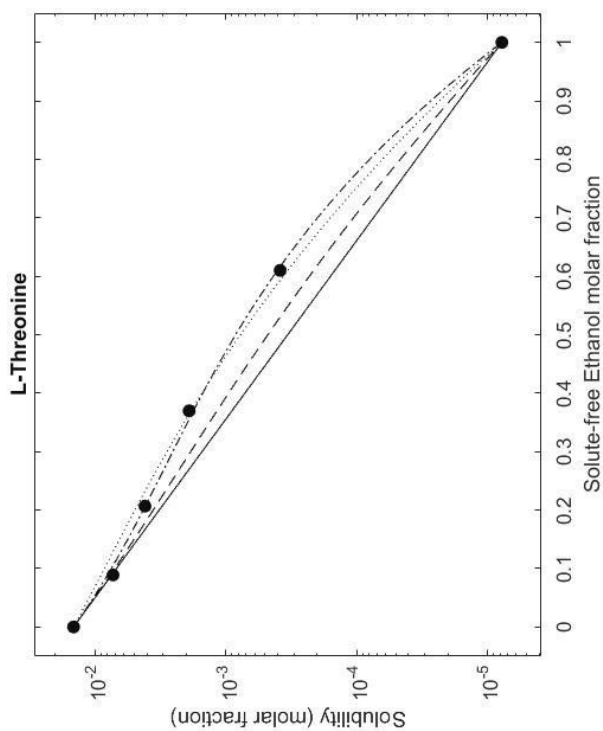


Figure 5.17: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

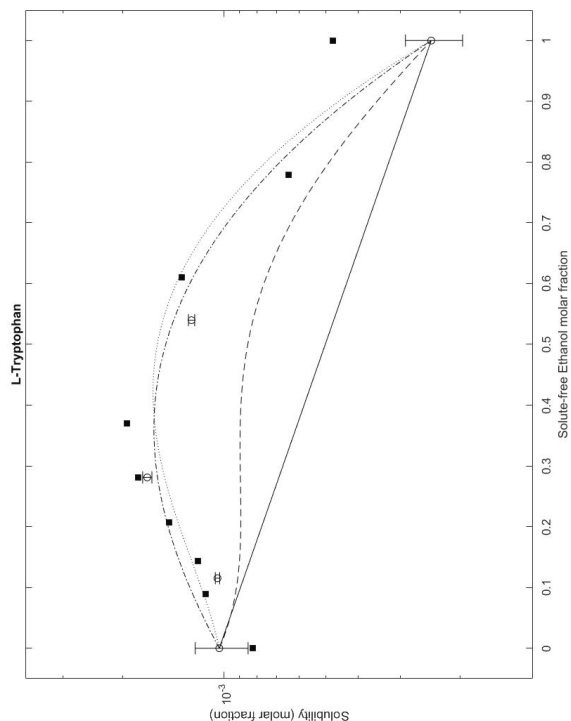


Figure 5.18: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

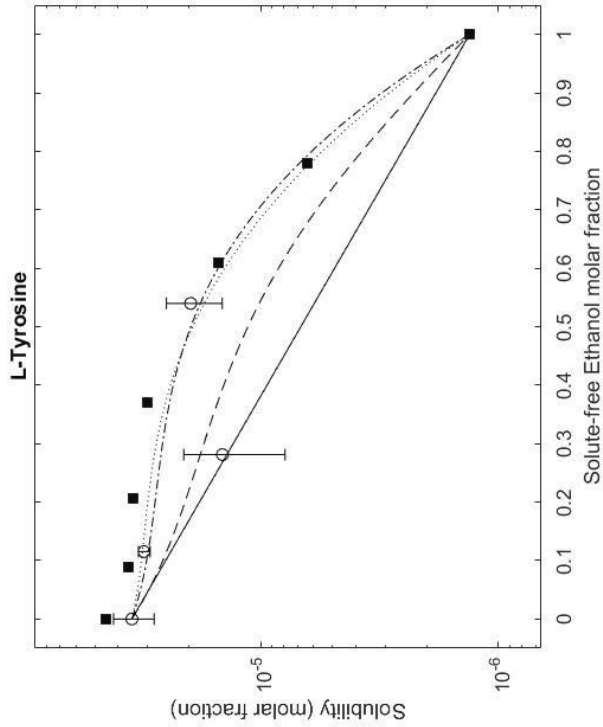


Figure 5.19: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

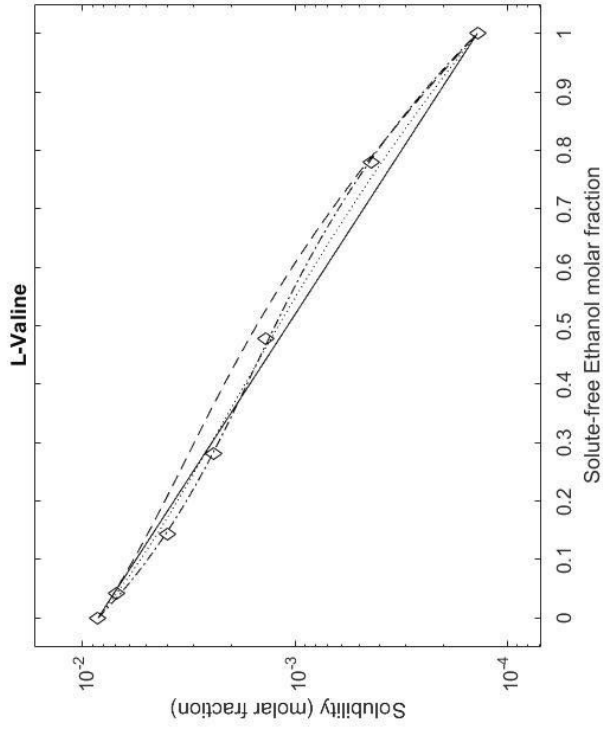


Figure 5.20: Plots of excess solubility = 0 (solid line), Gude model (dotted line), NRTL model (dot-dash line) and Jouyban-Acree model (dashed line) of the proteinogenic amino acids. Data are from the authors (open circles) or from the literature (closed squares).

Table 5.2: Calculated parameters for the Jouyban-Acree model and regressed parameters for the Gude and NRTL models for each amino acid

Model	Jouyban-Acree			Gude	NRTL	
Parameter	δ_d MPa ^{0.5}	δ_p MPa ^{0.5}	δ_{hb} MPa ^{0.5}	$C_{j,i,aa}$ mol·L ⁻¹	τ (water, aa) * 10 ⁶	τ (ethanol, aa) * 10 ⁶
L-Arginine	18.2312	8.0426	18.7229	1.5926	1.7003	4.2508
L-Cysteine	18.2152	6.2829	16.6663	-0.0542	0.9855	2.4638
Glycine	16.3684	10.0170	14.8238	-0.3007	1.2510	3.1276
L-Alanine	16.0719	5.1966	12.4649	-0.9696	1.6393	4.0982
L-Asparagine	16.8666	13.1746	17.4297	1.3097	1.0379	2.5947
L-Aspartic Acid	16.7254	7.2224	17.7194	0.3348	1.0962	2.7404
L-Glutamic Acid	16.6985	6.9179	17.3075	0.8557	1.0147	2.5369
L-Glutamine	16.8397	12.8701	17.0178	2.3001	1.0566	2.6416
L-Histidine	19.2245	4.8443	14.8368	3.2647	1.0297	2.5743
L-Isoleucine	15.7186	3.8964	11.0699	6.7822	0.9472	2.3681
L-Leucine	15.7646	3.8983	11.3848	1.9626	1.0476	2.6190
L-Serine	16.7016	8.5020	19.1997	3.6126	1.0840	2.7100
L-Threonine	16.4021	7.8108	18.6285	2.4094	1.0718	2.6796
L-Valine	15.7915	4.2028	11.7967	0.4935	1.1135	2.7837
L-Lysine	16.3246	7.5725	18.0542	-0.2720	1.2858	3.2146
L-Methionine	17.0776	5.3406	11.4124	1.3551	1.0421	2.6053
L-Phenylalanine	17.7072	4.5880	10.6483	3.0520	1.0343	2.5857
L-Proline	19.1658	6.1022	13.9127	3.6895	1.0573	2.6430
L-Tryptophan	20.3128	5.1780	8.4406	4.1462	1.2889	3.2223
L-Tyrosine	17.2033	3.2604	18.1645	3.8473	1.0968	2.7420
Water	15.6	16	42.3	N/A	N/A	N/A
Ethanol	15.8	8.8	19.4	N/A	N/A	N/A

5.4.1 Comparing regressed to predictive models of excess solubility

The NRMSE values and the number of measurements, n , for all of the models for each amino acid are shown in Table 5.3. The model with the lowest NRMSE value is the most accurate. For some amino acids, the number of data points were low, with only 5 or 6 data

points. Some of these amino acids with only 5 or 6 data points show the highest NRMSE values and therefore the most error. However, other amino acids with 5 data points (e.g. L-serine, L-methionine) had low error values. It is possible to compare the accuracy of the models for each amino acid since all models used the same data points. However, since the number of data points for some amino acids is limited, we cannot draw conclusions on the amino acids by comparing the NRMSE values.

For all amino acids, the NRTL model had the lowest error and is therefore the most accurate. The second most accurate for all amino acids, except for L-methionine, was the Gude model. The predictive Jouyban-Acree model was more accurate than the Gude model for L-methionine. Both the NRTL and Gude models had lower error values for all (in the case of NRTL) or most (in the case of Gude) amino acids. The predictive Jouyban-Acree model had a higher error value for all amino acids when compared to the NRTL model. The Jouyban-Acree model had a higher error value for all amino acids except L-methionine when compared to the Gude model.

The NRTL model described the empirical data well for all of the amino acids. All error values for the NRTL model were below 0.500, except for L-arginine, which had only 5 data points.

While the Gude model fits had higher NRMSE values than the NRTL model, the values of the error of the Gude model were under 0.500 for 19 of the 20 proteinogenic amino acids. The exception is L-arginine (NRMSE = 0.816). Since the errors are low, the Gude model could be used for drawing conclusions as we do in the next section. However, when more accurate calculations are needed, e.g. when designing an industrial process, we advise using the NRTL model.

Of the 20 amino acids, 14 of the amino acids modelled by the Jouyban-Acree model were under 0.500 except for L-arginine, L-cysteine, L-alanine, L-serine, L-lysine and L-proline. These 6 amino acids had only 5 or 6 data points each and were some of the most soluble amino acids. Furthermore, 5 of these 6 amino acids with NRMSE values above 0.500 in the Jouyban-Acree model had low NRMSE values using one or both of the other models. Even without using regressed parameters, the Jouyban-Acree model predicts the amino acid solubility for most of the amino acids well, but not as well as the Gude and NRTL models. The Jouyban-Acree model could be used when there are no or few solubility data available.

Table 5.3: NRMSE values for each amino acid for the Gude, NRTL and Jouyban-Acree models

Amino Acid	n	Gude	NRTL	Jouyban-Acree
L-Arginine	5	0.816	0.531	1.060
L-Cysteine	5	0.401	0.070	0.522
L-Glycine	15	0.286	0.285	0.310
L-Alanine	6	0.423	0.379	1.270
L-Asparagine	5	0.210	0.009	0.255
L-Aspartic Acid	6	0.284	0.161	0.476
L-Glutamic Acid	11	0.257	0.217	0.264
L-Glutamine	5	0.125	0.003	0.413
L-Histidine	9	0.182	0.016	0.483
L-Isoleucine	7	0.131	0.020	0.499
L-Leucine	6	0.191	0.042	0.260
L-Serine	5	0.360	0.021	5.470
L-Threonine	6	0.147	0.067	0.402
L-Valine	7	0.217	0.069	0.436
L-Lysine	5	0.304	0.280	1.320
L-Methionine	5	0.237	0.098	0.227
L-Phenylalanine	17	0.134	0.073	0.214
L-Proline	5	0.181	0.118	0.773
L-Tryptophan	14	0.174	0.170	0.354
L-Tyrosine	11	0.222	0.215	0.407

5.4.2 Effect of molecular shape on excess solubility of amino acids

As discussed earlier, the work of Flory-Huggins shows that liquids, similar to solids, have an entropic and lattice structure. Due to this entropy, Prausnitz *et al* [56] showed that the shape of a solute has an effect on the solubility of the solute. In their work, they used the relative van der Waals variables Q , surface area, and r , radius of the molecule, to describe the shape of the molecule and therefore how it influences this entropic and lattice structure. The shape of spherical solutes ($Q/r = 1.00$) showed no effects on the excess solubility of a solute. Straight-chain solutes

($Q/r = 0.788$) showed strong effects on the excess solubility of the solute, while rod-like solutes ($Q/r = 0.394$) showed an even greater effect on the excess solubility of the solute.

In Figure 5.21, the UNIFAC variables Q/r for each α -amino acid are plotted against the regressed constant in the Gude model, $C_{j,i,aa}$. A Q/r ratio close to unity means that the molecule is spherical and a lower ratio means that the molecule is rod-like. The $C_{j,i,aa}$ denotes the degree of excess solubility. A $C_{j,i,aa}$ close to 0 means that there is no excess solubility. A positive $C_{j,i,aa}$ means there is positive excess solubility and negative means there is negative excess solubility.

Spherical α -amino acids, like glycine, L-alanine and L-aspartic acid, with Q/R ratios from 0.89 to 0.92, react with less molecules of solvent. The spherical amino acids are surrounded by less water molecules than the rod-like amino acids, as their local concentration of ethanol is close to the concentration of the whole solution. As an organic anti-solvent is added, the lattice structure of these amino acids in solution is disrupted. This leads to little or no excess solubility.

Some rod-like α -amino acids show slightly positive excess solubility. The α -amino acids L-arginine, glycine, L-leucine, L-methionine and L-asparagine have Q/r ratios ranging from 0.81 to 0.85 and positive excess solubilities. The evidence supports the conclusion that they have a lower concentration of ethanol molecules around them locally than in the solution in general because of their shape. This would lead to their higher solubility than expected.

Even more pronounced rod-like amino acids, L-tyrosine, L-tryptophan, L-histidine, L-phenylalanine and L-proline, with Q/r ratios between 0.49 and 0.81, could react with even more molecules of solvent, due to their shape.

However, the shape of the amino acid molecules and therefore their effect on the entropic and lattice structure is only a part of the effect that the side chain of the amino acid has on its excess solubility. In Figure 5.21 there are exceptions to the general trend of the Q/r ratio of the amino acid and its excess solubility. These exceptions are the amino acids with reactive side chains. Therefore, in the next two sections we will examine the effect of the reactivity of the side chain to the excess solubility.

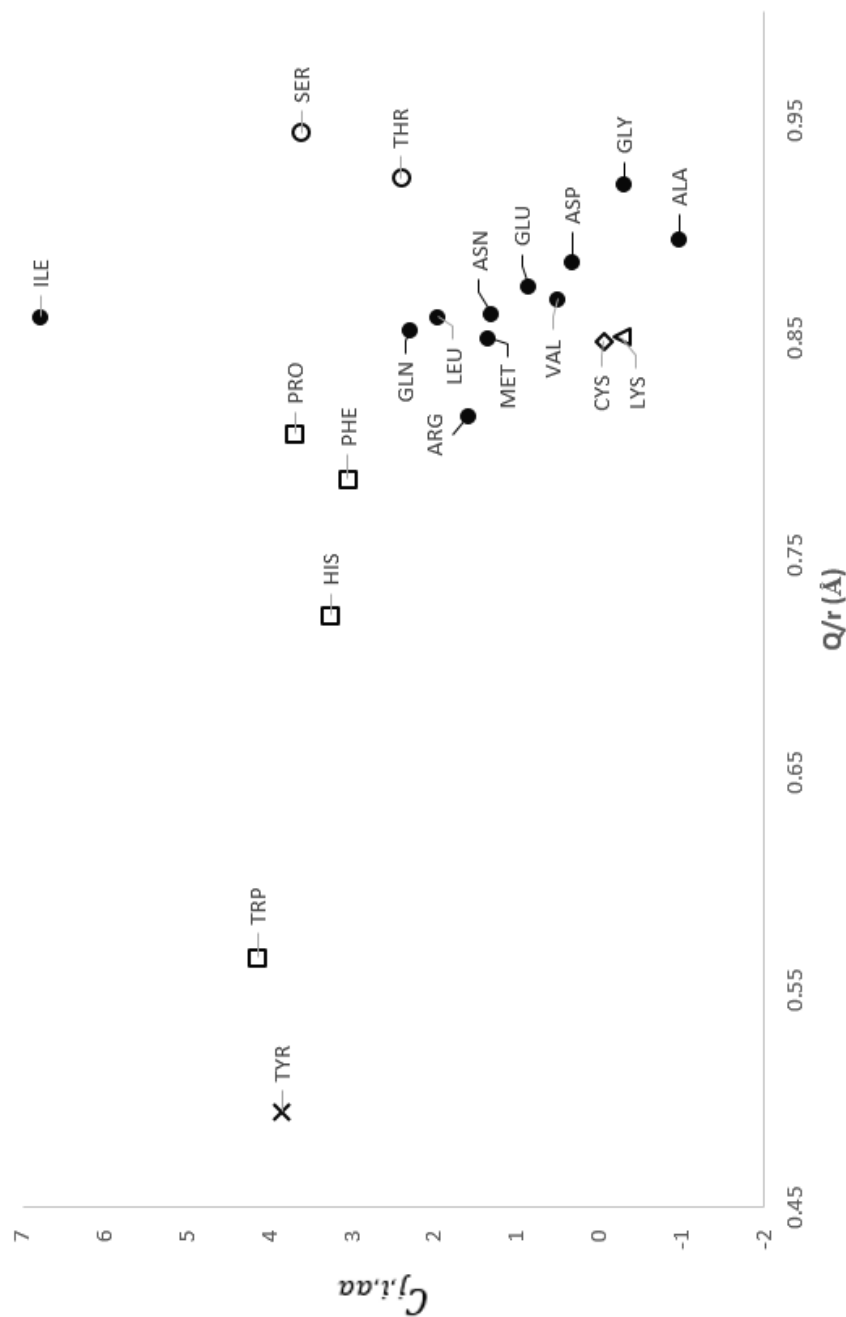


Figure 5.21: Regressed Gude model solubility parameter, $G_{j,aa}$, in relation to UNIFAC surface and radius parameters, Q/r showing non-reactive polar and aliphatic side chains (solid circles), hydroxyl side chains (open circles), lysyl side chain (open triangle), ringed side chains (open square), sulphur (open diamond) and hydroxyl ringed side chains (cross)

5.4.3 Amino Acids with non-reactive side chains

Eleven amino acids were identified as having non-reactive side chains. Non-reactive side chains are defined here as side chains that are either aliphatic or as measured at their isoelectric point, such as the data in this article, do not have a charge. These are shown in Figure 5.22 as black circles.

Glycine shows no excess solubility. Glycine has no side chain and has only an amino group and a carboxyl group. This supports the conclusion that lacking a reactive side chain, glycine follows the solubility predicted by the mole fraction of the solubility of both solvents. All other amino acids can be classified as glycine and a side chain. Glycine is therefore the null amino acid from which the change in excess solubility, not explained by its shape, due to the side chain can be discussed.

L-Glutamine, L-asparagine and L-arginine show little excess solubility. The first two amino acids have an amide in the side chain, while the last one has a guanidinium group in its side chain. At maximum solubility, the solution is at the isoelectric point, meaning that the side chains would not have a charge. Building on the evidence of glycine, the addition of an amide group or an amine group also has little effect on the excess solubility. Their slight increase in excess solubility could be explained by their shape alone as shown by the Q/r ratio.

L-Aspartic acid and L-glutamic acid are negatively charged amino acids. However, as discussed previously with L-arginine, since by definition, maximum solubility is measured at the isoelectric point, L-aspartic acid and L-glutamic acid would not be charged. This could mean that having no charge and being mostly spherical with a non-reactive side chain has no effect on the excess solubility in a two-solvent system. Similar to the previous amino acids, any small increase in excess solubility could possibly be explained by their slightly rod-like shape.

L-Alanine, L-valine, L-methionine, L-leucine and L-isoleucine are aliphatic amino acids. L-alanine has only one methylene group, L-valine and L-methionine have three and L-leucine and L-isoleucine have four. L-Methionine is slightly longer than L-valine because of a sulphur atom in between the second and third methylene. These amino acids show increasing excess solubility in order of their decreasing Q/r ratios. This means that as they become more rod-like, their

excess solubility has been shown to increase. However, this does not explain why L-isoleucine has an even higher increased solubility than L-leucine. Further research should be focused on the effect of the position of the branching on the side-chain to understand its effects on excess solubility.

5.4.4 Amino acids with reactive side chains

Nine amino acids have reactive side chains. These amino acids therefore would not follow the trend of higher Q/r ratios leading to lower excess solubility.

The only amino acid to show a large negative excess solubility is L-lysine. L-Lysine has a lysyl group in its side chain. This negative excess solubility is most pronounced around equal mole fractions of ethanol and water. The lysyl group is less attractive to the solvents as the water and ethanol are to each other, leading to lower solubility than expected.

All five amino acids with rings on their side chain have high positive excess solubilities. These amino acids include all three phenylic amino acids: L-phenylalanine, L-tryptophan and L-tyrosine. L-Histidine, which has imidazole on its side chain, shows positive excess solubility as well as L-proline, which has pyrrolidine as a side chain. It is possible that the two solvents act as affinity molecules, bringing these amino acids further into solution. However, it is also possible that their rod-like shape is causing this effect.

The three amino acids with a hydroxylic side chain show positive excess solubility. These include L-tyrosine, which is also has a phenyl group, L-serine and L-threonine. A side chain with a hydroxyl group leads to a preferential reaction to the solvents ethanol and water than ethanol to water. This cannot be explained by the shape of the amino acids, since both L-serine and L-threonine are spherical. Therefore, it may be concluded that an addition of a hydroxyl group leads to a marked increase in excess solubility.

5.5 Conclusion

The results support a hypothesis that both the shape of an amino acid and the activity of the side chain of an amino acid influence the solubility of the amino acid in mixed solvent solutions. Results support the conclusion that if the amino acid is spherical and does not have a reactive side chain, then there will be no change in the excess solubility as expected from the solvent mole fraction of ethanol and water. Spherical amino acids with reactive side chains, like

L-serine and L-threonine, will have positive excess solubilities. Rod-like amino acids with either a long side chain or a reactive side chain, such as the presence of a phenyl group and/or hydroxyl group, react preferentially to water and ethanol than water and ethanol do to each other and will have the greatest positive excess solubilities.

This hypothesis is artistically rendered in Figure 5.22 for four amino acids. In all four amino acids, the mole fraction of ethanol is 0.2. In the top left, L-alanine, a spherical amino acid ($Q/r = 0.90$; $C_{j,i,aa} = -0.97$) with a non-reactive side chain, is shown. Here the ethanol disrupts the water molecule lattice and there is a slight decrease in excess solubility. In the top right, L-serine, a spherical amino acid ($Q/r = 0.94$; $C_{j,i,aa} = 3.61$) with a reactive hydroxyl group on its side chain, is shown. The ethanol does not disrupt the lattice, rather it joins the lattice, being attracted to the hydroxyl group. Given small to medium molar concentrations of ethanol, there is marked positive excess solubility. In the bottom left, L-arginine, a rod-like amino acid ($Q/r = 0.81$; $C_{j,i,aa} = 1.59$) with a non-reactive side chain, is shown. Here, the lattice of water molecules is not disrupted, because it has contact with many water molecules. Given small molar concentrations of ethanol, there is a small amount of excess solubility. In the bottom right, L-tyrosine, a rod-like amino acid ($Q/r = 0.49$; $C_{j,i,aa} = 3.85$) with a reactive ring and hydroxyl groups on its side chain, is shown. Here, the ethanol and the water form a tight lattice around the molecule. In this case, even at medium concentrations of ethanol, there will be great excess solubility. At low concentrations of ethanol, the relative solubility has even been shown to increase.

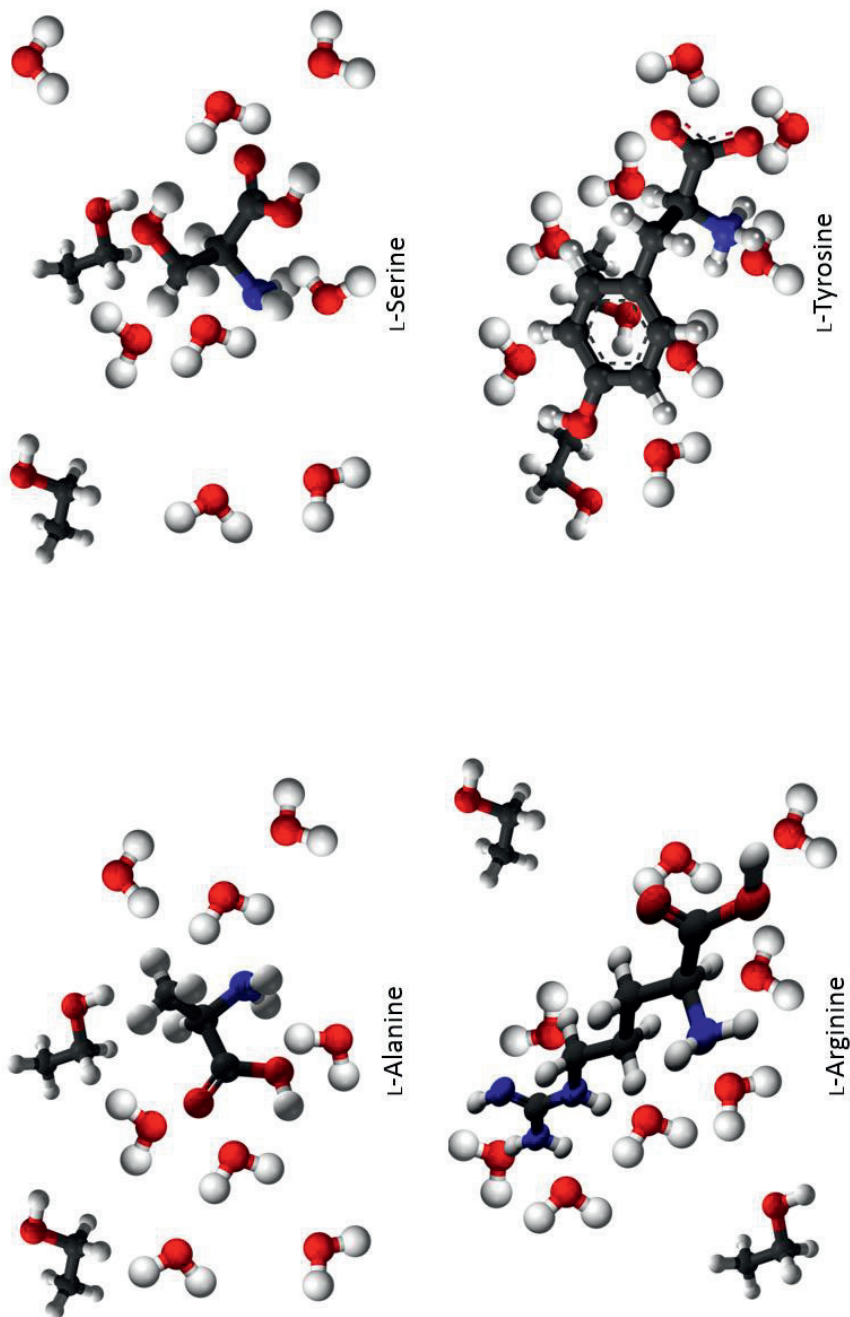


Figure 5.22: A depiction of the effects of amino acid shape and side chain composition in solution. Top left, L-alanine, spherical and non-reactive. Top right, L-serine, spherical and non-reactive. Bottom left, L-arginine, rod-like and reactive. Bottom right, L-tyrosine, rod-like and reactive.

Regressed models describe the solubility of the amino acids well. The NRTL model is better than the Gude model in this regard. However, since the Gude model has only one regressed parameter, it may be preferential to use it. The predictive Jouyban-Acree model performs well for some amino acids but not as well as both the Gude and NRTL models. Future research on group contribution in amino acid side chains is encouraged, in order to improve the accuracy of predictive models. The model that the end-user should use depends on the accuracy that is required. If the highest accuracy is required and solubility data is abundant, then a regressed model could be used. If the highest accuracy is not required, and there is no or few data, then a predictive model could be used.

The effect of the charge of an amino acid on the solubility of the amino acid has not been studied in this research. All the solubility data were taken at the isoelectric point, meaning that the amino acid was not charged. Further work on the effect of ethanol on a charged amino acid is encouraged.

References

- [1] Scott, E.; Peter, F.; Sanders, J., Biomass in the Manufacture of Industrial Products – The use of Proteins and Amino Acids. *Applied Microbiology and Biotechnology* **2007**, *75*, 751-762
- [2] Lammens, T. M.; Franssen, M. C. R.; Scott, E. L.; Sanders, J. P. M., Availability of protein-derived amino acids as feedstock for the production of bio-based chemicals. *Biomass and Bioenergy* **2012**, *44* (0), 168-181;
- [3] Lammens, T. M., Bio-based Industrial Chemicals from Glutamic Acid. **2011**. PhD Thesis, Wageningen University Press, Wageningen, the Netherlands
- [4] Dunn, M.S.; Ross, F.J.; Quantitative Investigations of Amino Acids and Peptides: IV. The Solubilities of the Amino Acids in Water-Ethyl Alcohol Mixtures *Journal of Biological Chemistry* **1938**, *125*, 309-322
- [5] McMeekin, T.L.; Cohn, E.J.; Weare, J.H.; Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. III. The Solubility of Derivatives of the Amino Acids in Alcohol—Water Mixtures *Journal of the American Chemical Society* **1935** *57* (4), 626-633
- [6] Ferreira, L. A.; Pinho, S. P.; Macedo, E. A., Solubility of L-serine, L-threonine and L-isoleucine in aqueous aliphatic alcohol solutions. *Fluid Phase Equilibria* **2008**, *270* (1-2), 1-9
- [7] Nozaki, Y; Tanford, C; The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions *Journal of Biological Chemistry* **1971** (246) 7, 2211-2217
- [8] Gude, M. T.; Meuwissen, H. H. J.; van der Wielen, L. A. M.; Luyben, K. C. A. M., Partition Coefficients and Solubilities of α -Amino Acids in Aqueous 1-Butanol Solutions. *Industrial & Engineering Chemistry Research* **1996**, *35* (12), 4700-4712;
- [9] Bowden, N.; Sanders, J. P. M.; Bruins, Marieke, Solubility of the proteinogenic α -amino acids in water, ethanol-water mixtures and ethanol. *Journal of Chemical Engineering Data* **2017**
- [10] Bowden, N.; Sanders, J. P. M.; Bruins, Marieke, Solid-liquid phase equilibria and thermodynamic properties of mixtures of 20 α -amino acids in a model system of industrial residues. *Journal of Chemical Thermodynamics* **2017** .

- [11] van Berlo, M.; Gude, M. T.; van der Wielen, L. A. M.; Luyben, K., Partition coefficients and solubilities of glycine in the ternary solvent system 1-butanol plus ethanol plus water. *Industrial & Engineering Chemistry Research* **1997**, 36 (6), 2474-2482;
- [12] Gude, M. T.; van der Wielen, L. A. M.; Luyben, K. C. A. M., Phase behavior of α -amino acids in multicomponent aqueous alkanol solutions. *Fluid Phase Equilibria* **1996**, 116 (1–2), 110-117.
- [13] Figueiredo, B. R.; Da Silva, F. A.; Silva, C. M., Non-ideality and Solubility Modeling of Amino Acids and Peptides in Aqueous Solutions: New Physical and Chemical Approach. *Industrial & Engineering Chemistry Research* **2013**, 52 (45), 16044-16056;
- [14] Held, C.; Cameretti, L. F.; Sadowski, G., Measuring and Modeling Activity Coefficients in Aqueous Amino-Acid Solutions. *Industrial & Engineering Chemistry Research* **2011**, 50 (1), 131-141;
- [15] Hempel, S.; Sadowski, G., Water activity coefficients in aqueous amino acid solutions by molecular dynamics simulation: 1. Force field development. *Mol. Simul.* **2012**, 38 (2), 132-138;
- [16] Khoshkbarchi, M. K.; Vera, J. H., A Simplified Perturbed Hard-Sphere Model for the Activity Coefficients of Amino Acids and Peptides in Aqueous Solutions. *Industrial & Engineering Chemistry Research* **1996**, 35 (11), 4319-4327;
- [17] Lee, B.-S.; Kim, K.-C., Study on the activity coefficients and solubilities of amino acids in aqueous solutions with perturbed-chain statistical associating fluid theory. *Korean Journal of Chemical Engineering* **2010**, 27 (1), 267-277;
- [18] Rudakov, A. M.; Sergievskii, V. V.; Khomchenko, E. O., Modeling component activities in binary aqueous solutions of some amino acids. *Russian Journal of Physical Chemistry* **2004**, 78, S77-S81;
- [19] Sadeghi, R., Thermodynamic representation of phase equilibrium behavior of aqueous solutions of amino acids by the modified Wilson model. *Fluid Phase Equilibria* **2007**, 260 (2), 266-274;
- [20] Soto, A.; Arce, A.; Khoshkbarchi, M. K., Experimental data and modelling of apparent molar volumes, isentropic compressibilities and refractive indices in aqueous solutions of glycine+NaCl. *Biophysical Chemistry* **1998**, 74 (3), 165-173.

- [21] Gao, W. C.; Li, Z. B., Determination and Chemical Modeling of Phase Equilibria for the Glycine-KCl-NaCl-H₂O System and Its Application To Produce Crystals with Anticaking Characteristics. *Industrial & Engineering Chemistry Research* **2012**, *51* (24), 8315-8325;
- [22] Held, C.; Prinz, A.; Wallmeyer, V.; Sadowski, G., Measuring and modeling alcohol/salt systems. *Chemical Engineering Science* **2012**, *68* (1), 328-339;
- [23] Khoshkbarchi, M. K.; Vera, J. H., Measurement and modeling of activities of amino acids in aqueous salt systems. *Aiche Journal* **1996**, *42* (8), 2354-2364;
- [24] Seyfi, S.; Pazuki, G.; Aghamiri, S. F.; Beheshti, M., On the prediction of equilibrium phase behavior of amino acids in aqueous and aqueous-electrolyte solutions using SAFT equation of state. *Fluid Phase Equilibria* **2009**, *287* (1), 15-22;
- [25] Soto, A.; Arce, A.; Khoshkbarchi, M. K., Experimental data and modelling of apparent molar volumes, isentropic compressibilities and refractive indices in aqueous solutions of glycine plus NaCl. *Biophysical Chemistry* **1998**, *74* (3), 165-173;
- [26] Tome, L. I. N.; Jorge, M.; Gomes, J. R. B.; Coutinho, J. A. P., Toward an Understanding of the Aqueous Solubility of Amino Acids in the Presence of Salts: A Molecular Dynamics Simulation Study. *Journal of Physical Chemistry B* **2010**, *114* (49), 16450-16459;
- [27] Tome, L. I. N.; Pinho, S. P.; Jorge, M.; Gomes, J. R. B.; Coutinho, J. A. P., Salting-in with a Salting-out Agent: Explaining the Cation Specific Effects on the Aqueous Solubility of Amino Acids. *Journal of Physical Chemistry B* **2013**, *117* (20), 6116-6128.
- [28] Ferreira, L. A.; Breil, M. P.; Pinho, S. P.; Macedo, E. A.; Mollerup, J. M., Thermodynamic Modeling of Several Aqueous Alkanol Solutions Containing Amino Acids with the Perturbed-Chain Statistical Associated Fluid Theory Equation of State. *Industrial & Engineering Chemistry Research* **2009**, *48* (11), 5498-5505;
- [29] Fuchs, D.; Fischer, J.; Tumakaka, F.; Sadowski, G., Solubility of Amino Acids: Influence of the pH value and the Addition of Alcoholic Cosolvents on Aqueous Solubility. *Industrial & Engineering Chemistry Research* **2006**, *45* (19), 6578-6584;
- [30] Ji, P.; Zou, J.; Feng, W., Effect of alcohol on the solubility of amino acid in water. *Journal of Molecular Catalysis B: Enzymatic* **2009**, *56* (2-3), 185-188.

- [31] Cameretti, L. F.; Sadowski, G., Modeling of aqueous amino acid and polypeptide solutions with PC-SAPT. *Chem. Eng. Process.* **2008**, *47* (6), 1018-1025;
- [32] Gross, J.; Sadowski, G., Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules. *Industrial & Engineering Chemistry Research* **2001**, *40* (4), 1244-1260;
- [33] GrosseDaldrup, J.-B.; Held, C.; Ruether, F.; Schembecker, G.; Sadowski, G., Measurement and Modeling Solubility of Aqueous Multisolute Amino-Acid Solutions. *Industrial & Engineering Chemistry Research* **2010**, *49* (3), 1395-1401;
- [34] Held, C.; Reschke, T.; Muller, R.; Kunz, W.; Sadowski, G., Measuring and modeling aqueous electrolyte/amino-acid solutions with ePC-SAFT. *Journal of Chemical Thermodynamics* **2014**, *68*, 1-12;
- [35] Ji, P. J.; Feng, W., Solubility of amino acids in water and aqueous solutions by the statistical associating fluid theory. *Industrial & Engineering Chemistry Research* **2008**, *47* (16), 6275-6279;
- [36] Klamt, A.; Krooshof, G. J. P.; Taylor, R., COSMOSPACE: Alternative to conventional activity-coefficient models. *AIChE Journal* **2002**, *48* (10), 2332-2349.
- [37] Flory, P. J., Thermodynamics of High Polymer Solutions. *the Journal of Physical Chemistry* **1942**, *10* (51).
- [38] Huggins, M. L., Solutions of Long Chain Compounds. *Journal of Physical Chemistry* **1941**, *9* (440).
- [39] Jouyban-Gharamaleki, A.; Valaee, L.; Barzegar-Jalali, M.; Clark, B. J.; Acree Jr, W. E., Comparison of various cosolvency models for calculating solute solubility in water–cosolvent mixtures. *International Journal of Pharmaceutics* **1999**, *177* (1), 93-101.
- [40] Jouyban, A.; Khoubnasabjafari, M.; Chan, H.K.; Acree, W.E. Jr., Mathematical representation of amino acids in binary aqueous-organic solvent mixtures at various temperatures using the Jouyban-Acree model. *Pharmazie* **2005** *61*: 789-792
- [41] Soto, A.; Arce, A.; K. Khoshkbarchi, M.; Vera, J. H., Measurements and modelling of the solubility of a mixture of two amino acids in aqueous solutions. *Fluid Phase Equilibria* **1999**, *158–160* (0), 893-901;
- [42] Ferreira, L. A.; Macedo, E. A.; Pinho, S. P., Solubility of amino acids and diglycine in aqueous–alkanol solutions. *Chemical Engineering Science* **2004**, *59* (15), 3117-3124;

- [43] Sadeghi, R., Modification of the NRTL and Wilson models for the representation of phase equilibrium behavior of aqueous amino acid - electrolyte solutions. *Can. J. Chem.-Rev. Can. Chim.* **2008**, *86* (12), 1126-1137;
- [44] Rivera, A.; Espinosa, A., Analysis of the NRTL parameters for amino acids in ethanol-water mixtures, using the approach of relative solubility. *Afinidad* **2016**, *73* (573), 37-42.
- [45] Chen, C. C.; Crafts, P. A., Correlation and Prediction of Drug Molecule Solubility in Mixed Solvent Systems with the Nonrandom Two-Liquid Segment Activity Coefficient (NRTL-SAC) Model. *Ind. Eng. Chem. Res.* **2006**, *45* (13), 4816.
- [46] Jouyban, A.; Khoubnasabjafari, M.; Chan, H. K.; Acree, W. E., Jr., Mathematical representation of solubility of amino acids in binary aqueous-organic solvent mixtures at various temperatures using the Jouyban-Acree model. *Pharmazie* **2006**, *61* (9), 789-792.
- [47] Jouyban, A.; Shayanfar, A.; Panahi-Azar, V.; Soleymani, J.; Yousefi, B. H.; Acree, W. E.; York, P., Solubility prediction of drugs in mixed solvents using partial solubility parameters. *Journal of Pharmaceutical Sciences* **2011**, *100* (10), 4368-4382.
- [48] Jouyban, A., Review of the cosolvency models for predicting solubility of drugs in water-cosolvent mixtures. *J. Pharm. Pharm. Sci.* **2008**, *11* (1), 32;
- [49] Jouyban, A.; Shayanfar, A.; Panahi-Azar, V.; Soleymani, J.; Yousefi, B. H.; Acree, W. E.; York, P., Solubility prediction of drugs in mixed solvents using partial solubility parameters. *J. Pharm. Sci.* **2011**, *100* (10), 4368;
- [50] Shayanfar, A.; Fakhree, M. A. A.; Acree, W. E.; Jouyban, A., Solubility of lamotrigine, diazepam, and clonazepam in ethanol + water mixtures at 298.15 K. *J. Chem. Eng. Data* **2009**, *54* (3), 1107.
- [51] Ellegaard, M. D.; Abildskov, J.; O'Connell, J. P., Molecular Thermodynamic Modeling of Mixed Solvent Solubility. *Industrial & Engineering Chemistry Research* **2010**, *49* (22), 11620-11632;
- [52] Se villano, D. M.; van der Wielen, L. A. M.; Trifunovic, O.; Ottens, M., Model Comparison for the Prediction of the Solubility of Green Tea Catechins in Ethanol/Water Mixtures. *Industrial & Engineering Chemistry Research* **2013**, *52* (17), 6039-6048.
- [53] Cohn, E. J.; Edsall, J. T., Proteins Amino Acids and Peptides as Ions and Dipolar Ions. *J. Chem. Educ.* **1943**, *20* (8), 415.

- [54] Renon H., Prausnitz J. M., Local Compositions in Thermodynamic Excess Functions for Liquid Mixtures, *AIChE J.*, **1968** 14 (1), S.135–144
- [55] Stefanis, E.; Panayiotou, C., Prediction of Hansen Solubility Parameters with a New Group-Contribution Method. *Int. J. Thermophys.* **2008**, 29 (2), 568.
- [56] Prausnitz, J. M.; Lichtenthaler, R. N. ; Gomes de Azevedo, E. *Molecular Thermodynamics of Fluid-Phase Equilibria* **1999** Prentice Hall PTR, Prentice-Hall Inc., Upper Saddle River, New Jersey, USA 07458

Chapter 6

Molecular-thermodynamic Correlation of Solubility Data for 20 Amino Acids in Water, in Ethanol and in Water-Ethanol Mixtures

In preparation as: Bowden, N.A., Bruins, M.E., Sanders, J.P.M. and Prausnitz, J. Molecular-thermodynamic Correlation of Solubility Data for 20 Amino Acids in Water, in Ethanol and in Water-Ethanol Mixtures

Abstract:

A predictive molecular-thermodynamic model for ternary liquid systems is useful for designing a process to separate α -amino acids from aqueous solution by using an anti-solvent, e.g. ethanol. Recently, reliable data have been published on the solubilities of many α -amino acids in mixtures of ethanol and water. To predict α -amino-acid solubility in a solution of water and ethanol, we use the van Laar equation for the molar excess Gibbs energy. To obtain meaningful activity coefficients from the solubility data using the van Laar equation, we need the ratio of the fugacity of the solid α -amino acid to that of the subcooled liquid amino acid. That ratio is obtained from estimated melting temperatures and enthalpies of fusion. The ternary van Laar equation provides a predictive model for obtaining the solubility of an α -amino acid in an ethanol-water solvent. The normalized root mean square variances (NRMSV) for 16 of the 20 solubility predictions are below 0.100, indicating very good agreement with the solubility data. The NRMSV of the other four predictions are below 0.220, indicating good agreement with the α -amino-acid solubility data. Six of the 20 amino acids could be calculated using previously published group contribution data. Additional group contribution data is reported here for seven amino acids.

6.1 Introduction

In the future, products that are currently produced using non-renewable resources (e.g. plastics, pharmaceuticals and fine chemicals) may be made from bio-based sources, such as proteins and α -amino acids [1-3]. One of the challenges in specifying a bio-based process is to design a procedure for separating α -amino acids from a pertinent industrial residue. Examples of such residues include green leafy biomass (e.g., grass, leaves, tea leaves), algae, cyanobacteria, press cake from plant-based oil and residue from juice production (e.g., seeds, grapes, peels). Waste from such processes often contain proteins and amino acids that can be recovered.

After amino acids are extracted from a residue, they need to be separated from each other. Currently, the most common method for separating amino acids from aqueous solution is industrial chromatography. A possibly less expensive procedure may be to crystallize each amino acid using an anti-solvent such as ethanol. To study this possible process, pertinent initial solubility data were reported earlier [4]. These data comprise the solubilities of 20 amino acids in water, in ethanol and in water-ethanol mixtures at 25 °C. In this work, we present a molecular-thermodynamic correlation of the solubility data. This correlation predicts the solubility of each amino acid in a water-ethanol mixture.

This research concerns all 20 proteinogenic α -amino acids. Every molecule of α -amino acid contains a carboxyl group attached to an α -carbon. This α -carbon is also attached to an amino group. The amino acids studied here are α -amino acids with side chains attached to the α -carbon between the carboxyl and amino groups. The exception is glycine that does not have a side chain. The side chains include aliphatic groups, aromatic or non-aromatic rings, hydroxyl groups, sulphur and charged groups [e.g. a second carboxyl group, or a lysyl group ($[\text{CH}_2]_4\text{NH}_2$), or a guanidinium group ($\text{HNC}[\text{NH}_2]_2$)]. A side chain on an amino acid strongly influences its solubility.

Several publications have reported studies on the solubility of α -amino acids in mixtures of alcohol and water [5-9]. Solubility measurements were reported and subsequent research focused on calculating their phase behavior [10]. Applications of this research have been used to separate proteins and α -amino acids that are dissolved in liquid industrial residues [11].

Models have been proposed for amino acid solubility in aqueous solution [12-19] and for the solubility of amino acids in salt solutions [20-27]. A few empirical models with regressed

parameters have been proposed to describe the solubility of α -amino acids in ethanol/water systems [28-33]. In this work, we apply the van Laar equation for the molar excess Gibbs energy to all 20 proteinogenic amino acids in ethanol, in water and in ethanol-water mixtures.

6.2 Thermodynamics of solid-liquid equilibria

For a solid solute, I , in equilibrium with a liquid solvent,

$$f_1^s = f_1^L, \quad (6.1)$$

where f is fugacity, superscript s stands for solid and superscript L stands for liquid. If the solid is pure, f_1^s is the fugacity of the pure solid.

For the liquid phase,

$$f_1^L = \gamma_1^s x_1^s f_{1,pure}^L, \quad (6.2)$$

where x_1^s is the mole fraction of amino acid (1) at saturation (superscript s), in the liquid phase and γ_1^s is the activity coefficient of amino acid at x_1^s . The standard state for the activity coefficient is pure subcooled liquid amino acid at system temperature.

The ratio, $(f^s/f^L)_{pure}$, can be obtained by a thermodynamic cycle as shown in many textbooks (e.g. Prausnitz *et al.* [34]). The molar Gibbs energy of fusion, Δg , is related to enthalpy of fusion, Δh , and entropy of fusion, Δs by

$$\Delta g = \Delta h - T\Delta s. \quad (6.3)$$

The enthalpy of fusion is related to temperature by

$$\Delta h = \Delta h_{T_m} + \int_{T_m}^T \Delta c_p \, dT, \quad (6.4)$$

where T_m is the melting temperature and Δc_p is the change in heat capacity upon melting. The entropy of fusion is related to temperature by

$$\Delta s = \Delta s_{T_m} + \int_{T_m}^T \frac{\Delta c_p}{T} dT. \quad (6.5)$$

At T_m ,

$$\Delta s_{T_m} = \frac{\Delta h_{T_m}}{T_m}. \quad (6.6)$$

Δg of fusion is related to the fugacity ratio by

$$\Delta g = RT \ln(f^L/f^S)_{pure}. \quad (6.7)$$

Substituting Equations (6.4-6.7) into Equation (6.3) yields

$$\ln(f^L/f^S)_{pure} = \frac{\Delta h_{T_m}}{RT_m} \left(\frac{T_m}{T} - 1 \right) - \frac{\Delta c_p}{R} \left(\frac{T_m}{T} - 1 \right) + \frac{\Delta c_p}{R} \ln \frac{T_m}{T} = -\ln(f^S/f^L)_{pure}. \quad (6.8)$$

The sum of the last two terms in Equation (6.8) is much smaller than the first term. A simplified, approximate form of Equation (6.8) is

$$-\ln(f^S/f^L)_{pure} = \frac{\Delta h_{T_m}}{RT_m} \left(\frac{T_m}{T} - 1 \right). \quad (6.9)$$

The fugacity ratio $(f^S/f^L)_{pure}$ can be calculated for each amino acid using Equation (6.9). However, because all amino acids disintegrate before they melt, there are no experimental data for Δh_{T_m} or T_m . Fortunately, these properties can be estimated.

6.3 Methods

6.3.1 Δh_{T_m} and T_m for amino acids

For each amino acid, T_m and Δh_{T_m} are estimated using the group contribution method proposed by Marrero and Gani [35]. The estimations are calculated using

$$e^{T_m/T_{m0}} = \sum_i N_i T_{m1i} + \sum_j M_j T_{m2j} + \sum_k O_k T_{m3k} \quad (6.10)$$

and

$$\Delta h_{T_m} - \Delta h_{T_0} = \sum_i N_i \Delta h_{T_{m1i}} + \sum_j M_j \Delta h_{T_{m2j}} + \sum_k O_k \Delta h_{T_{m3k}}, \quad (6.11)$$

where $T_{m0} = 147.450 \text{ K}$, $\Delta h_{T_0} = -2.806 \text{ kJ}\cdot\text{mol}^{-1}$, T_{m1i} and Δh_{m1i} are the contributions of the first-order groups, i, for the corresponding properties N_i times, T_{m2j} and Δh_{m2j} are the contributions of the second-order groups, j, for the corresponding properties M_j times and T_{m3k} and Δh_{m3k} are the contributions of the third-order groups, k, for the corresponding properties O_k times. Contributions for these groups are reported in the supplementary data of Marrero and Gani [35]. T_m and Δh_{T_m} are given in Table 2 as well as the ratio $(f^s/f^L)_{pure}$ at 25°C calculated using Equation (6.9)..

For seven α -amino acids, the second-order group contribution data in the literature only covers 66% or less, by mass percent, of the amino acid. This makes accurate estimations difficult. For these amino acids a regression analysis was performed to obtain the contribution of the groups that are not covered in the literature, then a modified Δh_{T_m} and T_m are calculated and reported in Table 6.2. Table 6.1 shows the group contributions, Δh_{T_m} and T_m , and the groups for these amino acids.

Table 6.1: Additional T_m and Δh_{T_m} second-order group contribution data not found in Marrero and Gani [35] for α -amino acids as calculated from solubility data

Amino Acid	Group	$T_{m2} \text{ (K)}$	$\Delta h_{T_{m2}} \text{ (kJ}\cdot\text{mol}^{-1})$
L-Arginine	$(\text{CH}_2)_2(\text{NH})_2\text{CNH}_2$	77.60	-17.4
L-Glutamic Acid	$(\text{CH}_2)_2\text{COOH}$	7.014	0.73
L-Glutamine	$(\text{CH}_2)_2\text{CONH}_2$	6.071	1.99
L-Histidine	$\text{CH}_2(1\text{H-imidazol-4})$	111.6	4.61
L-Isoleucine	$(\text{CH}_3)_2\text{CCH}_2$	323.8	7.75
L-Leucine	$\text{CH}(\text{CH}_3)_2$	36.04	3.27
L-Lysine	$(\text{CH}_2)_3\text{NH}_2$	-56.71	-26.7

6.3.2 Determining the interaction parameters of an amino acid in water or ethanol and for amino-acid-free water and ethanol for the van Laar model

Before we can predict the solubility of an amino acid in mixtures of water and ethanol, we must first determine the interaction parameters, A_{ij} , of an amino acid in water and ethanol and of water and ethanol in a solute free solution, where i and j are all possible combinations of solute (1), water (2) and ethanol (3). We used the van Laar equation obtained from Wohl's expansion for the molar excess Gibbs energy of a binary solution, Equation (6.12) for these determinations. The van Laar equation is useful for solutions of non-electrolyte components whose effective molecular surface area, q (\AA^2), are somewhat but not excessively different. For van Laar, the molar excess Gibbs energy is

$$\frac{g^E}{RT} = \frac{2A_{12}x_1x_2q_1q_2}{x_1q_1+x_2q_2}. \quad (6.12)$$

As discussed in standard thermodynamic books (e.g. Prausnitz et al. [34]), upon differentiating Equation (6.12), the activity coefficient of amino acid (1) in water (2) is

$$\ln \gamma_{12}^s = \frac{2q_1A_{12}}{\left(1 + \frac{q_1x_1}{q_2x_2}\right)^2}. \quad (6.13)$$

Similarly, the activity coefficient of amino acid (1) in ethanol (3) is

$$\ln \gamma_{13}^s = \frac{2q_1A_{13}}{\left(1 + \frac{q_1x_1}{q_3x_3}\right)^2}. \quad (6.14)$$

For the water-ethanol system, the activity coefficient of water and ethanol are

$$\ln \gamma_{23}^s = \frac{2q_2A_{23}}{\left(1 + \frac{q_2x_2}{q_3x_3}\right)^2} \quad \text{and} \quad \ln \gamma_{32}^s = \frac{2q_3A_{32}}{\left(1 + \frac{q_3x_3}{q_2x_2}\right)^2}. \quad (6.15-6.16)$$

At fixed temperature, all A_{ij} ($i = 1, 2$ or 3 ; $j = 1, 2$ or 3) and all q 's are constants. Pure-component parameters q_1 , q_2 , and q_3 roughly reflect the molecular areas of solute-solvent

interaction. However, in practice, they are pure-component empirical parameters obtained from binary and ternary experimental data.

In the van Laar model,

$$A_{21} = \frac{q_2}{q_1} A_{12}, \quad (6.17)$$

$$A_{31} = \frac{q_3}{q_1} A_{13}, \quad (6.18)$$

and

$$A_{32} = \frac{q_3}{q_2} A_{23}. \quad (6.19)$$

For each amino acid, at 25 °C, we calculated $(f^s/f^L)_{pure}$ and we have an experimental result for solubility, x_1^s . From Equations (6.2) and (6.9), we obtain γ_1^s for each amino acid in water and again for each amino acid in ethanol. We assumed that in solution each amino acid is at its isoelectric point. We use γ_1^s and x_1^s in Equations (6.13) and (6.14) to obtain parameters A_{12} and A_{13} . For that calculation, we need assigned values for q_1 , q_2 , and q_3 . For q_2 we use 0.92 and for q_3 we use 1.00 as suggested by Anderson and Prausnitz [37]. The parameter q_1 was calculated for each amino acid using UNIFAC group contribution data [38-40]. Table 6.2 gives all required parameters for calculating γ_{12} and γ_{13} in water or ethanol.

Data for activity coefficients for the binary system water-ethanol determine constants A_{23} and A_{32} . Assuming that the vapor phase is ideal, activity coefficients for water and for ethanol are calculated from vapor-liquid-equilibrium data [41] by

$$\frac{y_2 P}{x_2 P_2^s} = \gamma_{23}^s, \text{ and } \frac{y_3 P}{x_3 P_3^s} = \gamma_{32}^s, \quad (6.20-6.21)$$

where y is the vapor-phase mole fraction, P^s is the vapor pressure of pure liquid, x is the liquid-phase mole fraction and P is the total pressure. We used $A_{23} = 0.778$ and, $A_{32} = 0.846$.

Calculated solubilities of amino acids in water-ethanol mixtures are not sensitive to A_{23} and A_{32} .

6.4. Predicting the solubilities of amino acids in water-ethanol mixtures

The solubility of an amino acid in an aqueous solution of ethanol can be predicted by applying the van Laar model to a ternary solution. For a ternary system of amino acid, water and ethanol, the van Laar molar excess Gibbs energy is

$$\frac{\bar{g}^E_{\text{ternary}}}{RT} = \frac{2q_2A_{12}x_1x_2 + 2q_3A_{13}x_1x_3 + \frac{2q_2q_3}{q_1}A_{23}x_2x_3}{x_1 + \frac{q_2}{q_3}x_2 + \frac{q_3}{q_1}x_3}. \quad (6.22)$$

Differentiation of Equation (6.22) gives the activity coefficient of an amino acid in a mixture of water and ethanol.

$$\ln \gamma_1^s = \frac{x_2^2 A_{12} \left(\frac{A_{21}}{A_{12}}\right)^2 + x_3^2 A_{13} \left(\frac{A_{31}}{A_{13}}\right)^2 + x_2 x_3 \frac{A_{21} A_{31}}{A_{12} A_{13}} (A_{12} + A_{13} - A_{32}) \left(\frac{A_{13}}{A_{31}}\right)}{\left(x_1 + x_2 \frac{A_{21}}{A_{12}} + x_3 \frac{A_{31}}{A_{13}}\right)^2} \quad (6.23)$$

The ratio $x_2/x_3 = x'_2/x'_3$, where the prime indicates amino-acid-free. We fix this ratio that reflects the amino-acid-free solvent composition. We then obtain the saturated mole fraction of the amino-acid solute, x_1^s by simultaneously solving Equation (6.24) with the mass balance $x_1 + x_2 + x_3 = 1.00$. The ratio $(f^s/f^L)_{\text{pure}}$ at 25 °C is obtained from Equation (6.9).

$$x_1^s = (f^s/f^L)_{\text{pure}} \left(\exp \left[\frac{x_2^2 A_{12} \left(\frac{A_{21}}{A_{12}}\right)^2 + x_3^2 A_{13} \left(\frac{A_{31}}{A_{13}}\right)^2 + x_2 x_3 \frac{A_{21} A_{31}}{A_{12} A_{13}} (A_{12} + A_{13} - A_{32}) \left(\frac{A_{13}}{A_{31}}\right)}{\left(x_1 + x_2 \frac{A_{21}}{A_{12}} + x_3 \frac{A_{31}}{A_{13}}\right)^2} \right] \right)^{-1} \quad (6.24)$$

This calculation is not explicit in x_1^s . Solution of Equation (6.24) is achieved by a simple iterative program.

6.5 Results and discussion

6.5.1 Comparison of predicted and measured solubilities

The normalized root-mean-square variance (NRMSV) between measured and predicted solubilities was calculated for all 20 proteinogenic amino acids. Here, \hat{c}_{x_i} is the predicted solubility of the amino acid and c_{x_i} is the measured solubility of the amino acid at a given solute-

free mole fraction of ethanol in the mixed solvent. The maximum measured solubility of the amino acid in water, in ethanol or in the water-ethanol mixed solvent is c_{max} and the minimum measured solubility in the water, ethanol or water-ethanol mixed solvent is c_{min} .

$$NRMSV = \frac{\sqrt{\frac{\sum_{i=1}^n (c_{x_i} - c_{x_i})^2}{n}}}{c_{max} - c_{min}} \quad (6.25)$$

6.5.2 Model parameters and results

Table 6.2 gives experimental binary x_1^s for amino acids in pure water and pure ethanol. Table 6.2 also gives Δh_{T_m} , T_m , and $(f^s/f^L)_{pure}$. R is the gas constant and is equal to $8.314 \text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$.

Table 6.2: Estimated model parameters (q_1 , Δh_{T_m} and T_m) using Marrero and Gani group contribution method equations (6.10 and 6.11) [35], modified estimated model parameters (Mod. Δh_{T_m} and Mod. T_m) from additional regressed group contributions reported in Table 6.1, intermediate results ($(f^s/f^L)_{pure}$, A_{12} and A_{13}), solubility data (x_1^s) in pure water and pure ethanol and the normalized root-mean-square variance (NRMSV) for each of the 20 proteinogenic amino acids.

Amino Acid	$\Delta h_{T_m}/R$ (K)	Mod. $\Delta h_{T_m}/R$ (K)	T_m (K)	Mod. T_m (K)	$(f^s/f^L)_{pure}$ (-)	x_1^s in water $\cdot 10^3$ (-)	x_1^s in ethanol $\cdot 10^3$ (-)	A_{12} (-)	A_{13} (-)	q_1 (\AA^2)	NRMSV (-)
L-Alanine	2905	-	587	-	8.251	2.496	0.282	-1.283	3.384	2.996	0.050
L-Arginine	3219	-	642	-	3.088	1.861	0.143	-2.103	3.077	4.416	0.022
L-Asparagine	5048	-	619	-	0.155	0.341	0.002	-3.177	4.505	3.936	0.024
L-Aspartic Acid	4308	4843	605	680	0.109	0.068	0.002	-1.837	3.817	3.912	0.033
L-Cysteine	3325	-	577	-	4.551	1.299	0.113	-1.167	3.700	3.528	0.054
L-Glutamic Acid	4626	-	606	-	0.378	0.106	0.009	-1.038	3.353	4.452	0.141
L-Glutamine	4958	-	628	-	0.161	0.509	0.001	-3.614	4.698	4.476	0.025
L-Glycine	2550	-	579	-	15.84	5.677	0.046	-1.655	5.845	2.460	0.046
L-Histidine	4379	-	703	-	0.213	0.497	0.004	-3.271	3.981	3.752	0.036
L-Isoleucine	4198	-	855	-	0.104	0.453	0.004	-3.925	3.389	4.384	0.124
L-Leucine	3832	-	664	-	0.841	0.297	0.095	-1.298	2.183	4.612	0.083
L-Lysine	1200	-	491	-	206.1	22.32	6.180	-0.408	3.718	5.004	0.041
L-Methionine	4087	-	588	-	1.160	0.677	0.057	-1.876	3.008	4.596	0.074
L-Phenylalanine	4509	3957	583	633	0.890	0.322	0.118	-1.332	2.025	5.348	0.126
L-Proline	2403	2203	362	451	82.11	19.59	5.176	-2.958	2.873	3.468	0.118
L-Serine	3388	-	588	-	3.686	6.815	0.004	-4.812	6.736	3.888	0.015
L-Threonine	3286	3786	588	669	0.877	1.452	0.008	-3.185	4.731	4.424	0.048
L-Tryptophan	4818	4194	580	656	0.464	0.103	0.043	-0.809	0.646	4.159	0.103
L-Tyrosine	4948	5769	609	785	0.006	0.003	0.001	-1.734	1.540	3.975	0.025
L-Valine	3121	3331	588	645	2.454	0.847	0.138	-1.326	2.882	4.072	0.028

Estimates for $\Delta h_{T_m}/R$ range from 1200 K for L-lysine to 5048 K for L-asparagine. In general, the amino acids with the least number of methylene groups (e.g. L-proline, glycine, L-alanine) have the lowest Δh_{T_m} . Amino acids with the same number of methylene groups but with an OH, ring, O, lysyl or amine group in the side chain have a higher Δh_{T_m} .

Estimated melting temperatures, T_m , range from 451 K for L-proline to 855 K for L-isoleucine. Our estimates for T_m are within the range of T_m previously reported for similar amino acids [36]. The more methylene groups in the amino acids, the higher the T_m .

Using the estimates of Δh_{T_m} and T_m resulted in a NRSMV under 0.150 for 13 of the amino acids. For the other seven amino acids, a modified Δh_{T_m} and T_m was calculated. The optimized Δh_{T_m} for these amino acids ranged from -12% of the original estimate for L-phenylalanine to 17% for L-tyrosine. The optimized T_m for these amino acids ranged from 9% above the original estimate for L-phenylalanine to 29% above the original estimate for L-tyrosine.

At 25 °C, fugacity ratio $(f^s/f^L)_{pure}$ ranges from $0.60 \cdot 10^{-5}$ for L-tyrosine to 0.206 for L-proline. In general, if the Δh_{T_m} is relatively high, then it follows that the $(f^s/f^L)_{pure}$ is relative low. However, this is not always the case (e.g. L-cysteine).

The solubility of an amino acid in a single solvent is determined by two separate quantities, the ratio $(f^s/f^L)_{pure}$ and the interaction parameter, A_{12} (or A_{13}). The ratio $(f^s/f^L)_{pure}$ depends only on temperature and on the properties of the pure amino acid, but A_{12} (or A_{13}) depends on solute-solvent interactions.

Interaction parameters for amino acids in water, A_{12} , range from -9.80 for L-lysine to 2.71 for L-tyrosine. The lower the A_{12} , the stronger the attractive solute-water interaction and therefore, the more hydrophilic the amino acid. However, this does not necessarily mean that a more hydrophilic amino acid will have a higher solubility in water because the solubility of an amino acid in water also depends on the ratio $(f^s/f^L)_{pure}$. For example, L-glutamine A_{12} is -4.91, lower than that for L-valine, -2.79; yet, L-glutamine has a solubility in water lower than that for L-valine because $(f^s/f^L)_{pure}$ for L-valine is higher than that for L-glutamine. The ratio $(f^s/f^L)_{pure}$ is 0.525 for the former and 0.039 for the latter.

Interaction parameters for amino acids in ethanol, A_{13} , range from -4.34 for L-lysine to 9.17 for L-histidine. A negative A_{13} indicates that the amino acid is ethanolphilic. L-Lysine, is the only proteinogenic amino acid with a negative A_{13} . In most cases, the lower the A_{13} , the higher the solubility in ethanol. This is not always the case, because solubility in ethanol is also affected by the $(f^s/f^L)_{pure}$ ratio of the amino acid, as indicated in the previous paragraph for water.

For the 20 amino acids, the surface-area parameters of interaction, q , range from 0.111 for L-valine to 3.47 for L-proline. Previous studies have shown that the effective surface area of a molecule indicates its solubility [33]. Parameter q is lower than the geometric surface area of the entire molecule for all amino acids. The smaller q indicates the area of the amino acid that interacts with the solvent. At the isoelectric pH, for amino acids like glycine, L-proline, L-cysteine and L-alanine, nearly the entire molecule interacts with the solvent because the q is similar to the area of the entire amino-acid molecule. For other amino acids, q is less than the area of the entire amino-acid molecule. For amino acids like L-valine, L-tryptophan, L-lysine, L-leucine and L-tyrosine, the interactive surface area is limited to a small portion of the molecule. The physical significance of q is approximate. Essentially, q is an empirical fitting parameter.

Figures 6.1-6.20 compare predicted and measured χ_1^s . The NRMSV varies from 0.015 for L-serine to 0.217 for L-tryptophan. The NRMSV for 16 of the 20 proteinogenic amino acids were lower than 0.100, indicating very good agreement with the solubility data. This group of amino acids includes L-tyrosine, whose maximum solubility occurs when the mixed solvent contains 91.1 mole% water and 8.90 mole% ethanol, higher than that at 100 mole% water. For L-isoleucine, L-tryptophan, L-phenylalanine and L-proline, NRMVs are lower than 0.220, indicating good agreement with the solubility data.

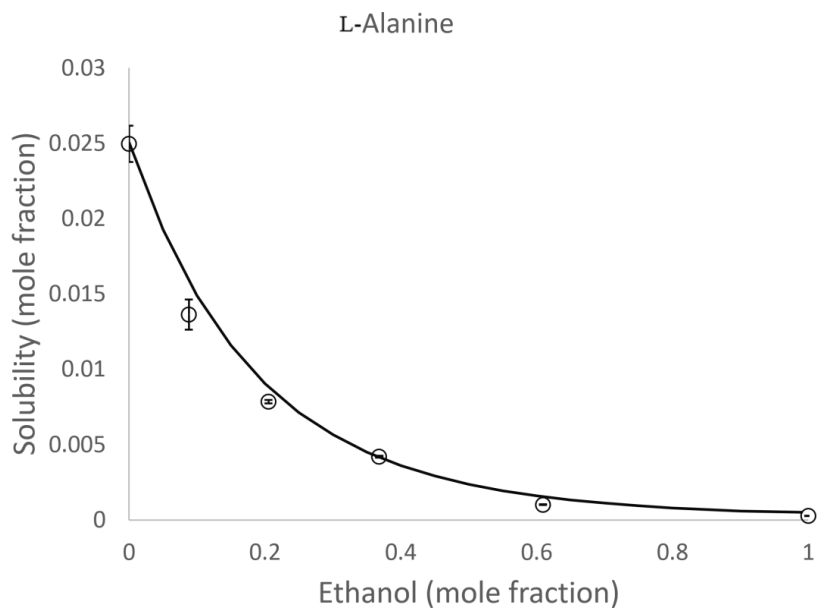


Figure 6.1: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-alanine as a function of ethanol mole fraction at 25 °C and 0.1 MPa.

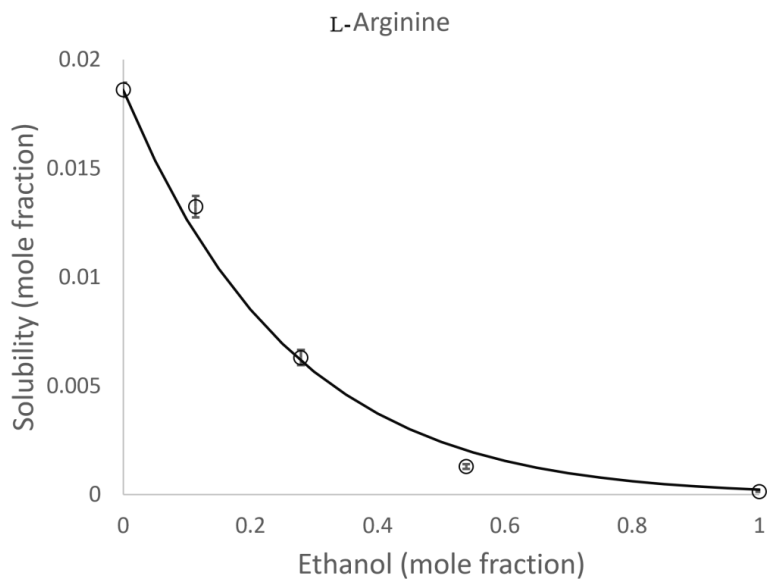


Figure 6.2: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-arginine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

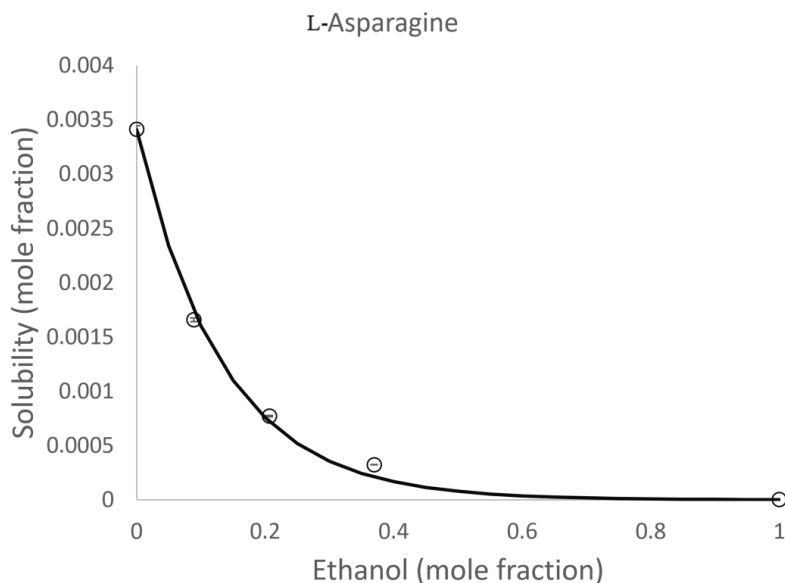


Figure 6.3: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-asparagine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

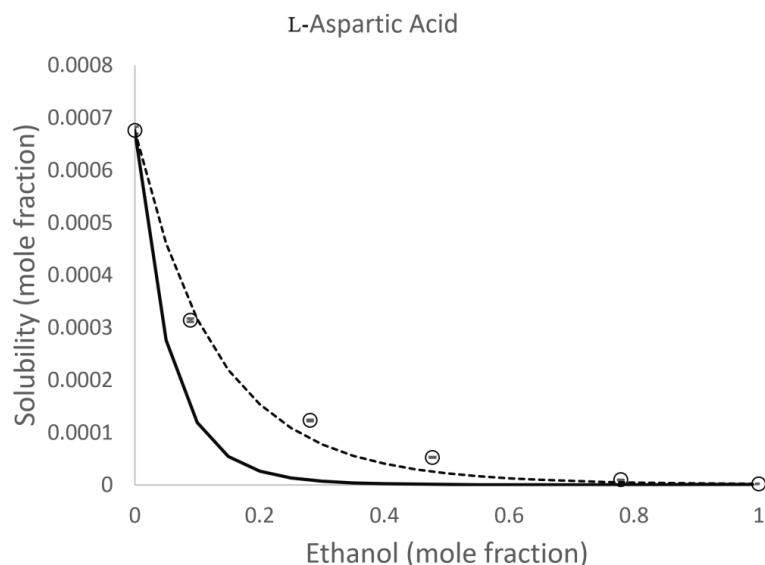


Figure 6.4: Solubility as predicted by the model without the additional group contribution data in Table 6.1 (solid line), solubility as predicted by the model including the modified Δh_{T_m} and T_m in Table 6.1 (dashed line) and experimental solubility data (circles) in mole fraction of L-aspartic acid as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

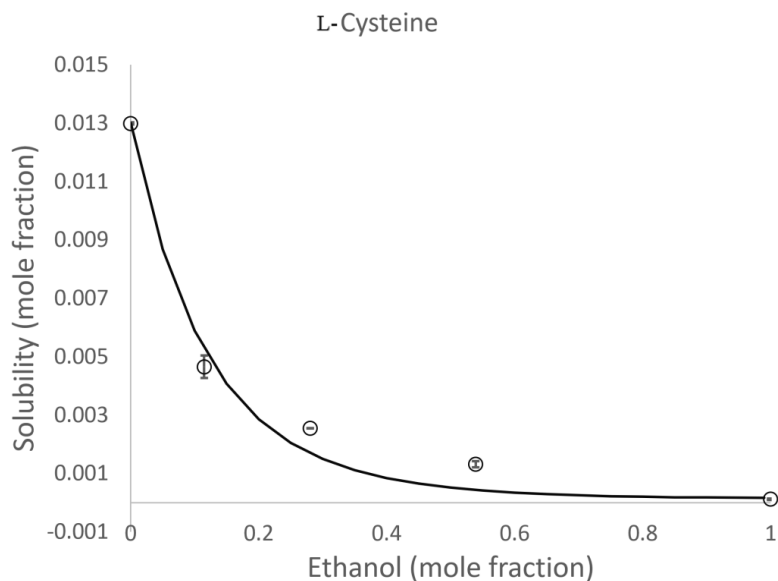


Figure 6.5: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-cysteine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

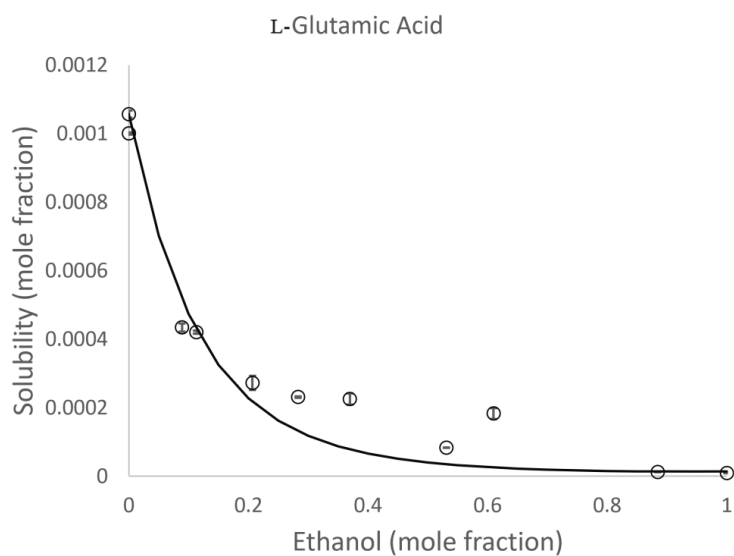


Figure 6.6: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-glutamic acid as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

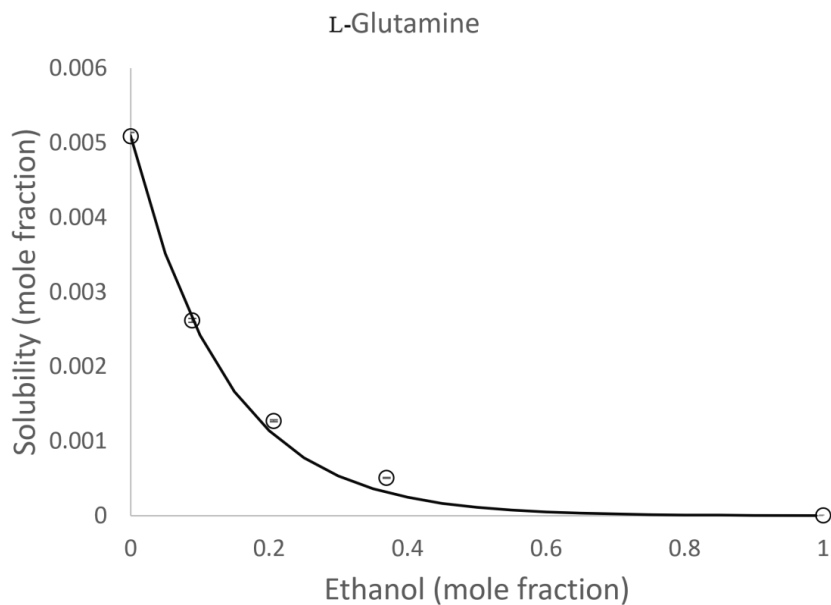


Figure 6.7: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-glutamine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

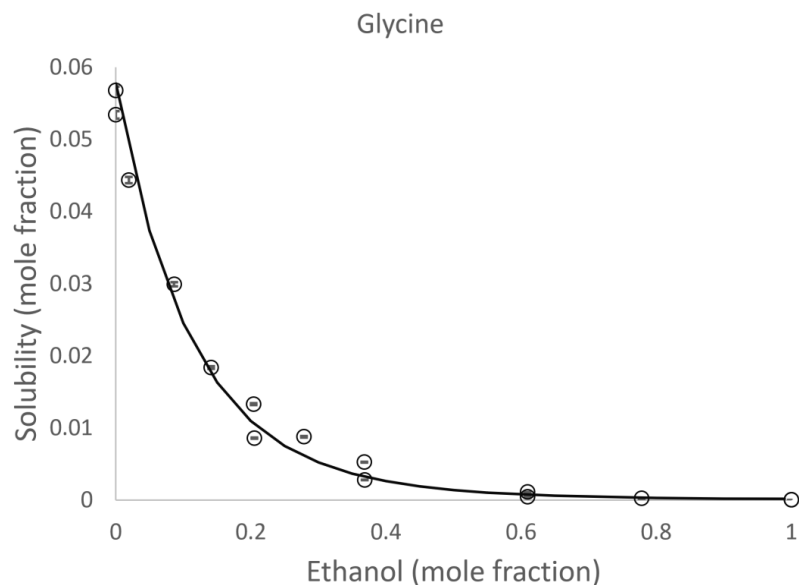


Figure 6.8: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of glycine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

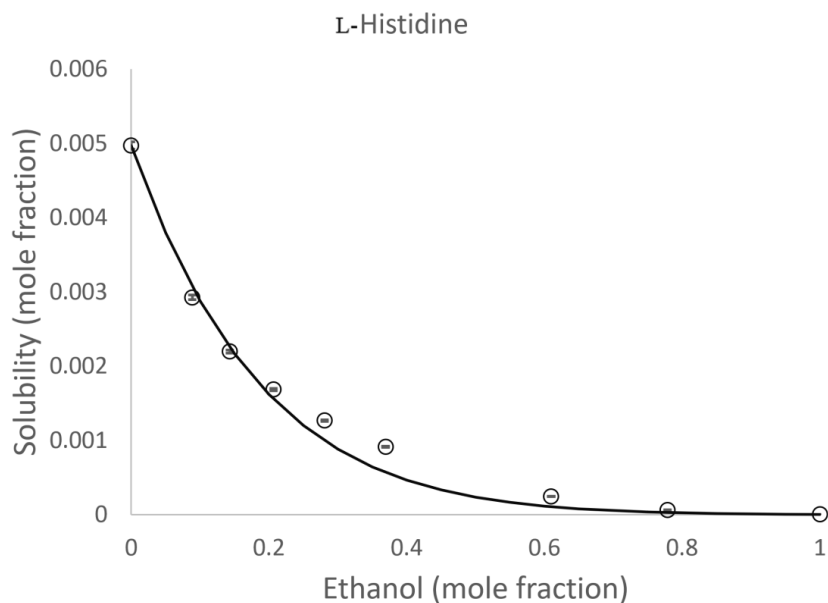


Figure 6.9: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-histidine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

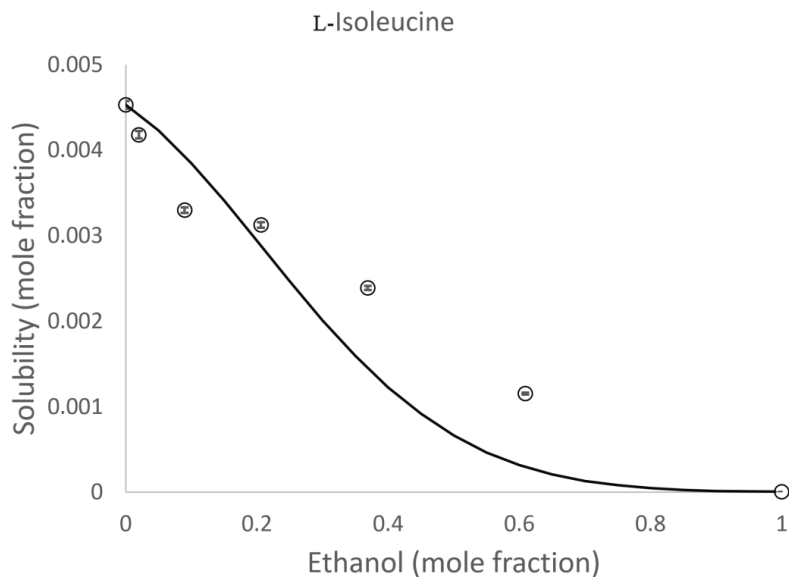


Figure 6.10: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-isoleucine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

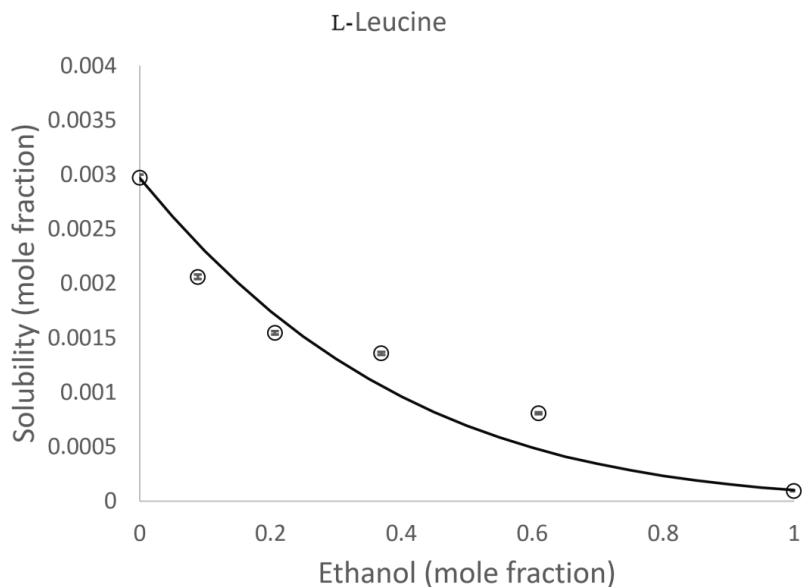


Figure 6.11: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-leucine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

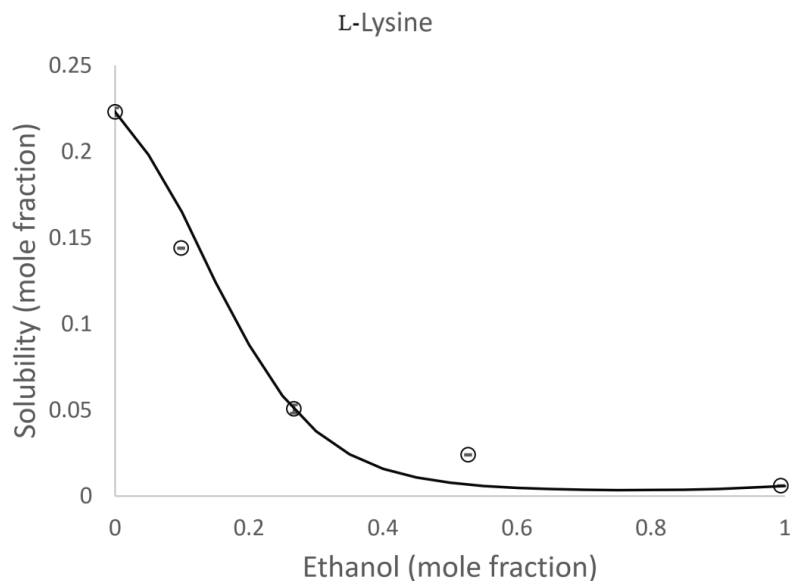


Figure 6.12: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-lysine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

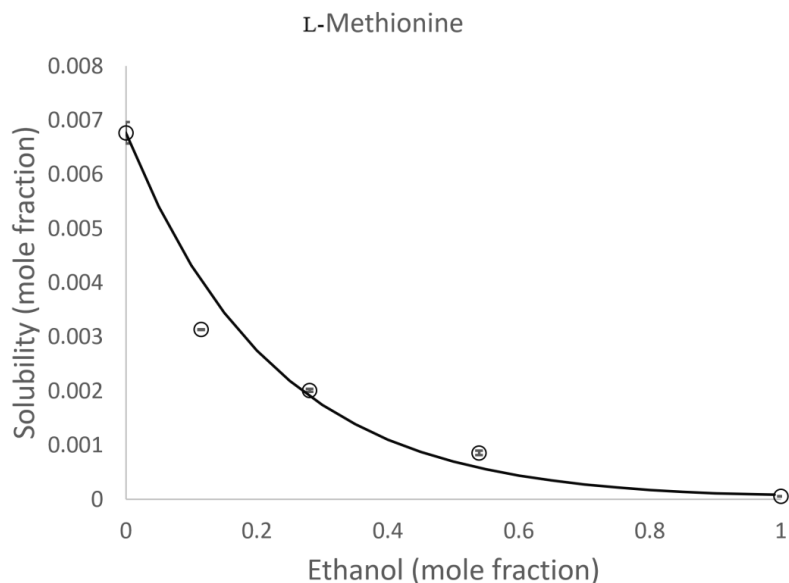


Figure 6.13: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-methionine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

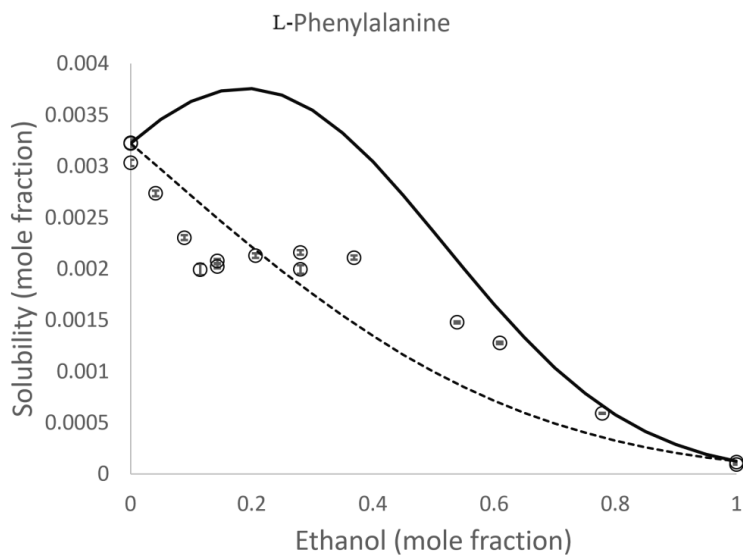


Figure 6.14: Solubility as predicted by the model without the additional group contribution data in Table 6.1 (solid line), solubility as predicted by the model including the modified Δh_{T_m} and T_m in Table 6.1 (dashed line) and experimental solubility data (circles) in mole fraction of L-phenylalanine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

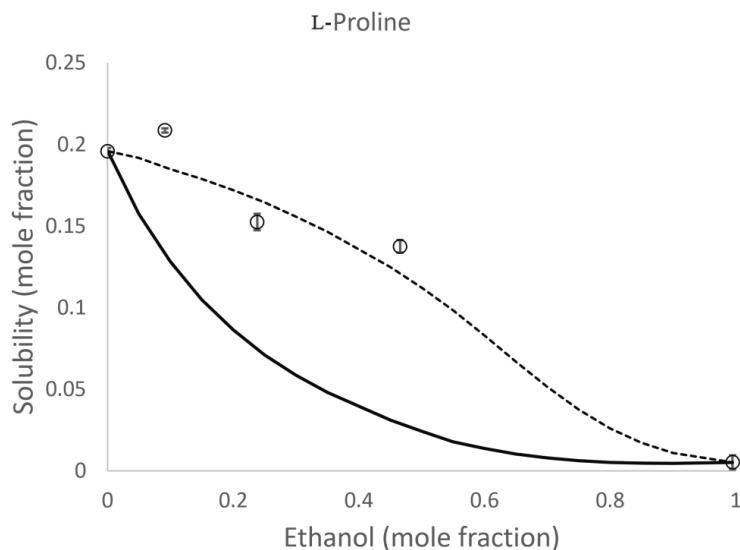


Figure 6.15: Solubility as predicted by the model without the additional group contribution data in Table 6.1 (solid line), solubility as predicted by the model including the modified Δh_{T_m} and T_m in Table 6.1 (dashed line) and experimental solubility data (circles) in mole fraction of L-proline as a function of ethanol mole fraction in the solvent at 25 °C and 0.1MPa.

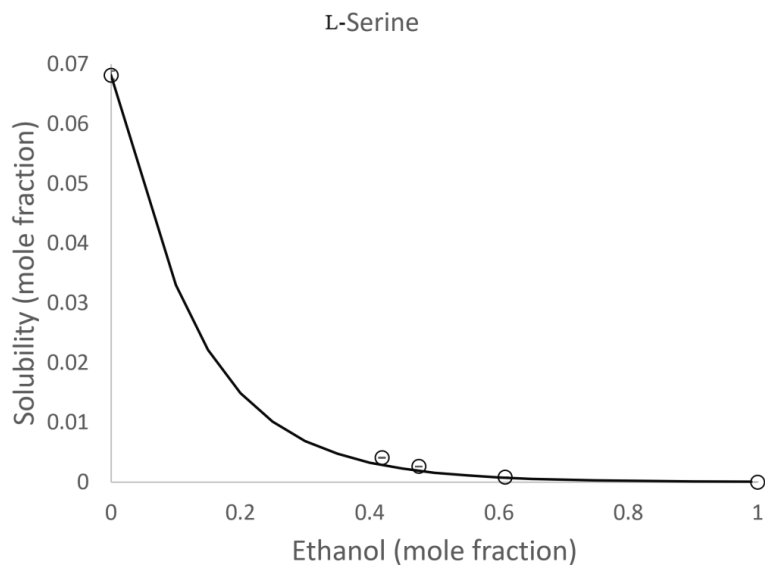


Figure 6.16: Solubility as predicted by the model (solid line) and experimental solubility data (circles) in mole fraction of L-serine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

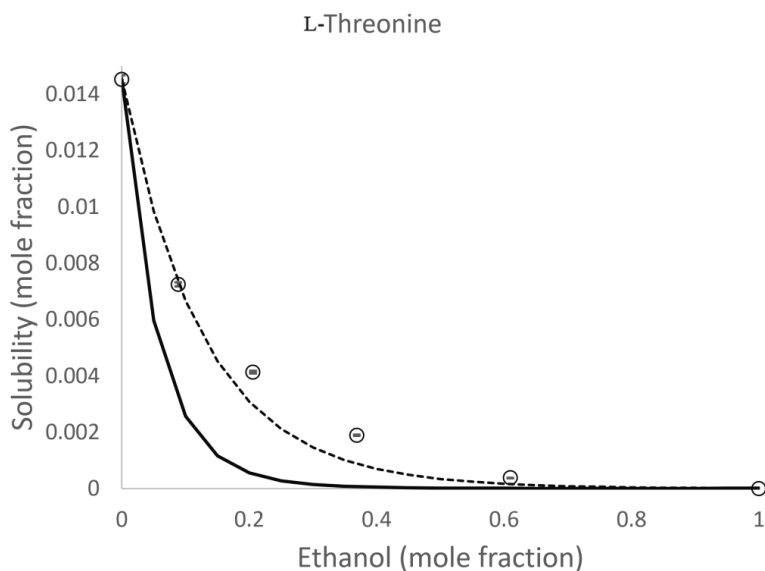


Figure 6.17: Solubility as predicted by the model without the additional group contribution data in Table 6.1 (solid line), solubility as predicted by the model including the modified Δh_{T_m} and T_m in Table 6.1 (dashed line) and experimental solubility data (circles) in mole fraction of L-threonine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

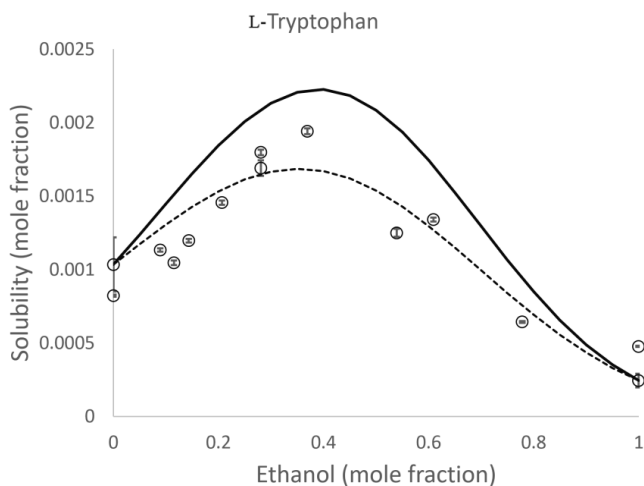


Figure 6.18: Solubility as predicted by the model without the additional group contribution data in Table 6.1 (solid line), solubility as predicted by the model including the modified Δh_{T_m} and T_m in Table 6.1 (dashed line) and experimental solubility data (circles) in mole fraction of L-tryptophan as a function of ethanol mole fraction in the solvent at 25 °C and 0.1 MPa.

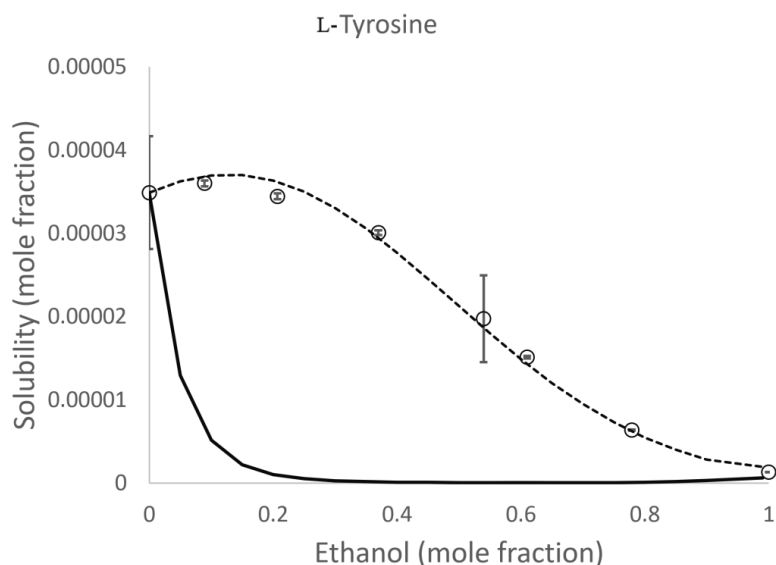


Figure 6.19: Solubility as predicted by the model without the additional group contribution data in Table 6.1 (solid line), solubility as predicted by the model including the modified Δh_{T_m} and T_m in Table 6.1 (dashed line) and experimental solubility data (circles) in mole fraction of L-tyrosine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1MPa.

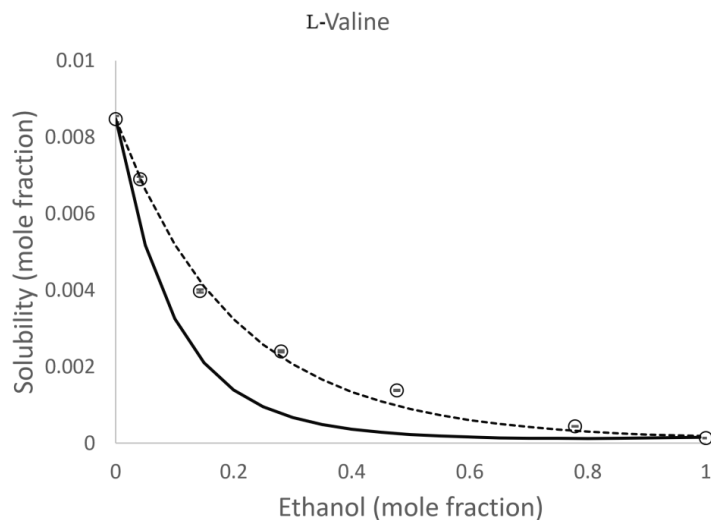


Figure 6.20: Solubility as predicted by the model without the additional group contribution data in Table 6.1 (solid line), solubility as predicted by the model including the modified Δh_{T_m} and T_m in Table 6.1 (dashed line) and experimental solubility data (circles) in mole fraction of L-valine as a function of ethanol mole fraction in the solvent at 25 °C and 0.1MPa.

6.6 Conclusions

The ternary form of the van Laar equation for the molar excess Gibbs energy is useful for predicting the solubilities of amino acids in water-ethanol mixtures at 25 °C and 0.1MPa. The calculated solubilities of some amino acids show maxima in a water-ethanol mixed solution instead of in pure water or ethanol. These calculated maxima, predicted by the van Laar model, are in agreement with experimental data.

Amino acids with the lowest interaction parameters, A_{12} or A_{13} , are not necessarily the most soluble. The solubility of an amino acid depends on both its $(f^s/f^L)_{pure}$ ratio and its interaction with the solvent or solvent mixture.

To illustrate, consider the solubility of glycine. The glycine $(f^s/f^L)_{pure}$ ratio is 0.016, indicating low activity. For glycine, $A_{12} = -1.66$, indicating that it is relatively hydrophilic, and A_{13} of 5.85, indicating that it is relatively ethanolphobic. Glycine has a relatively large $q = 2.46$. Compared to other α -amino acids, glycine is only slightly soluble in water because it has a low $(f^s/f^L)_{pure}$ ratio, indicating low solubility, but a relatively low A_{12} , suggesting that glycine is relatively soluble. Glycine has a low solubility in ethanol because glycine has a low $(f^s/f^L)_{pure}$ ratio, indicating low solubility and a relatively high A_{13} , again suggesting a low solubility. The q for glycine is larger than those for both water and ethanol. Because A_{12} is negative, the maximum solubility of glycine is at 100 mole% water.

Consider L-tryptophan. L-Tryptophan has a low $(f^s/f^L)_{pure}$ ratio. L-tryptophan is relatively hydrophobic (the second highest A_{12} of all the α -amino acids in this work) and ethanolphobic (positive A_{13}); therefore, they are relatively insoluble in water and in ethanol. L-Tryptophan has q much larger than the q for water and that for ethanol. L-Tryptophan has their maximum solubilities above that in pure water and pure ethanol.

Finally, consider L-lysine, that has the highest $(f^s/f^L)_{pure}$ ratio of any α -amino acid, more than twice that of the second highest $(f^s/f^L)_{pure}$ ratio, indicating high activity. L-Lysine has one of the highest A_{12} of any α -amino acid, indicating that it is one of the most hydrophobic of the α -amino acids. However, since it has a high activity, it is still the most soluble α -amino acid in this work. Conversely, L-proline has an activity that is less than the half of L-lysine, however, it has a lower A_{12} (hydrophilic) and therefore a solubility in water similar to L-lysine.

The advantage of the technique that is used in this work is that an accurate solubility of a solute in a binary solvent can be calculated based on only two experimental measurements. One measurement in each pure solvent is enough to run the model. The other parameters can be calculated and estimated using current thermodynamic data, such as UNIFAC data or group contributions of Marrero and Gani. However, this brings with it the disadvantage that the group contribution data is not complete and has itself a variance. In this work, we add group contribution data to the literature and find that the variance was acceptable for most solute solubilities.

References

- [1] Scott, E.; Peter, F.; Sanders, J. P. M., Biomass in the Manufacture of Industrial Products – The use of Proteins and Amino Acids. *Applied Microbiology and Biotechnology* **2007**, 75, 751-762
- [2] Lammens, T. M.; Franssen, M. C. R.; Scott, E. L.; Sanders, J. P. M., Availability of protein-derived amino acids as feedstock for the production of bio-based chemicals. *Biomass and Bioenergy* **2012**, 44 [0], 168-181
- [3] Lammens, T. M., *Bio-based Industrial Chemicals from Glutamic Acid*. **2011**. PhD Dissertation. Wageningen University Press. Wageningen, the Netherlands
- [4] Bowden, N.; Sanders, J. P. M.; Bruins, M. E., Solubility of the proteinogenic α -amino acids in water, ethanol-water mixtures and ethanol. *Journal of Chemical & Engineering Data* **2017**
- [5] Gude, M. T.; Meuwissen, H. H. J.; van der Wielen, L. A. M.; Luyben, K. C. A. M., Partition Coefficients and Solubilities of α -Amino Acids in Aqueous 1-Butanol Solutions. *Industrial & Engineering Chemistry Research* **1996**, 35 [12], 4700-4712
- [6] Dunn, M.S.; Ross, F.J.; Quantitative Investigations of Amino Acids and Peptides: IV. The Solubilities of the Amino Acids in Water-Ethyl Alcohol Mixtures *Journal of Biological Chemistry* **1938**, 125, 309-322
- [7] McMeekin, T.L.; Cohn, E.J.; Weare, J.H.; Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. III. The Solubility of Derivatives of the Amino Acids in Alcohol—Water Mixtures *Journal of the American Chemical Society* **1935** 57 [4], 626-633
- [8] Ferreira, L. A.; Pinho, S. P.; Macedo, E. A., Solubility of L-serine, L-threonine and L-isoleucine in aqueous aliphatic alcohol solutions. *Fluid Phase Equilibria* **2008**, 270 [1-2], 1-9;
- [9] Nozaki, Y; Tanford, C; The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions *Journal of Biological Chemistry* **1971** [246] 7, 2211-2217
- [10] Gude, M. T.; Meuwissen, H. H. J.; van der Wielen, L. A. M.; Luyben, K. C. A. M., Partition Coefficients and Solubilities of α -Amino Acids in Aqueous 1-Butanol Solutions. *Industrial & Engineering Chemistry Research* **1996**, 35 [12], 4700-4712
- [11] Widarani; Bowden, N.; Sanders, J. P. M.; Bruins, M. E., Fractional Precipitation of Amino Acids from Agro-industrial Residues Using Ethanol. *Industrial & Engineering Chemistry Research* **2016** 55[27], 7462-7472
- [12] van Berlo, M.; Gude, M. T.; van der Wielen, L. A. M.; Luyben, K., Partition coefficients and solubilities of glycine in the ternary solvent system 1-butanol plus ethanol plus water. *Industrial & Engineering Chemistry Research* **1997**, 36 [6], 2474-2482

- [13] Figueiredo, B. R.; Da Silva, F. A.; Silva, C. M., Non-ideality and Solubility Modeling of Amino Acids and Peptides in Aqueous Solutions: New Physical and Chemical Approach. *Industrial & Engineering Chemistry Research* **2013**, 52 [45], 16044-16056
- [14] Held, C.; Cameretti, L. F.; Sadowski, G., Measuring and Modeling Activity Coefficients in Aqueous Amino-Acid Solutions. *Industrial & Engineering Chemistry Research* **2011**, 50 [1], 131-141
- [15] Hempel, S.; Sadowski, G., Water activity coefficients in aqueous amino acid solutions by molecular dynamics simulation: 1. Force field development. *Mol. Simul.* **2012**, 38 [2], 132-138
- [16] Khoshkbarchi, M. K.; Vera, J. H., A Simplified Perturbed Hard-Sphere Model for the Activity Coefficients of Amino Acids and Peptides in Aqueous Solutions. *Industrial & Engineering Chemistry Research* **1996**, 35 [11], 4319-4327
- [17] Lee, B.-S.; Kim, K.-C., Study on the activity coefficients and solubilities of amino acids in aqueous solutions with perturbed-chain statistical associating fluid theory. *Korean Journal of Chemical Engineering* **2010**, 27 [1], 267-277
- [18] Rudakov, A. M.; Sergievskii, V. V.; Khomchenko, E. O., Modeling component activities in binary aqueous solutions of some amino acids. *Russian Journal of Physical Chemistry* **2004**, 78, S77-S81
- [19] Sadeghi, R., Thermodynamic representation of phase equilibrium behavior of aqueous solutions of amino acids by the modified Wilson model. *Fluid Phase Equilibria* **2007**, 260 [2], 266-274
- [20] Soto, A.; Arce, A.; Khoshkbarchi, M. K., Experimental data and modelling of apparent molar volumes, isentropic compressibilities and refractive indices in aqueous solutions of glycine+NaCl. *Biophysical Chemistry* **1998**, 74 [3], 165-173
- [21] Gao, W. C.; Li, Z. B., Determination and Chemical Modeling of Phase Equilibria for the Glycine-KCl-NaCl-H₂O System and Its Application To Produce Crystals with Anticaking Characteristics. *Industrial & Engineering Chemistry Research* **2012**, 51 [24], 8315-8325
- [22] Held, C.; Prinz, A.; Wallmeyer, V.; Sadowski, G., Measuring and modeling alcohol/salt systems. *Chemical Engineering Science* **2012**, 68 [1], 328-339
- [23] Khoshkbarchi, M. K.; Vera, J. H., Measurement and modeling of activities of amino acids in aqueous salt systems. *Aiche Journal* **1996**, 42 [8], 2354-2364
- [24] Seyfi, S.; Pazuki, G.; Aghamiri, S. F.; Beheshti, M., On the prediction of equilibrium phase behavior of amino acids in aqueous and aqueous-electrolyte solutions using SAFT equation of state. *Fluid Phase Equilibria* **2009**, 287 [1], 15-22

- [25] Soto, A.; Arce, A.; Khoshkbarchi, M. K., Experimental data and modelling of apparent molar volumes, isentropic compressibilities and refractive indices in aqueous solutions of glycine plus NaCl. *Biophysical Chemistry* **1998**, *74* [3], 165-173
- [26] Tome, L. I. N.; Jorge, M.; Gomes, J. R. B.; Coutinho, J. A. P., Toward an Understanding of the Aqueous Solubility of Amino Acids in the Presence of Salts: A Molecular Dynamics Simulation Study. *Journal of Physical Chemistry B* **2010**, *114* [49], 16450-16459
- [27] Tome, L. I. N.; Pinho, S. P.; Jorge, M.; Gomes, J. R. B.; Coutinho, J. A. P., Salting-in with a Salting-out Agent: Explaining the Cation Specific Effects on the Aqueous Solubility of Amino Acids. *Journal of Physical Chemistry B* **2013**, *117* [20], 6116-6128
- [28] Gude, M. T.; van der Wielen, L. A. M.; Luyben, K. C. A. M., Phase behavior of α -amino acids in multicomponent aqueous alkanol solutions. *Fluid Phase Equilibria* **1996**, *116* [1–2], 110-117
- [29] Ferreira, L. A.; Breil, M. P.; Pinho, S. P.; Macedo, E. A.; Mollerup, J. M., Thermodynamic Modeling of Several Aqueous Alkanol Solutions Containing Amino Acids with the Perturbed-Chain Statistical Associated Fluid Theory Equation of State. *Industrial & Engineering Chemistry Research* **2009**, *48* [11], 5498-5505
- [30] Fuchs, D.; Fischer, J.; Tumakaka, F.; Sadowski, G., Solubility of Amino Acids: Influence of the pH value and the Addition of Alcoholic Cosolvents on Aqueous Solubility. *Industrial & Engineering Chemistry Research* **2006**, *45* [19], 6578-6584
- [31] Ji, P.; Zou, J.; Feng, W., Effect of alcohol on the solubility of amino acid in water. *Journal of Molecular Catalysis B: Enzymatic* **2009**, *56* [2–3], 185-188
- [32] Jouyban, A.; Khoubnasabjafari, M.; Chan, H.K.; Acree, W.E. Jr., Mathematical representation of amino acids in binary aqueous-organic solvent mixtures at various temperatures using the Jouyban-Acree model. *Pharmazie* 2005 *61*: 789-792
- [33] Bowden, N. A.; Mendez Sevilano, D.; Sanders, J. M. P.; Bruins, M. E. Modelling the effects of ethanol on the solubility of the proteinogenic amino acids with the NRTL, Gude and Jouyban-Acree models. *Fluid Phase Equilibria* **2017**
- [34] Prausnitz, J. M.; Lichtenthaler, R. N.; Gomes de Azevedo, E., *Molecular Thermodynamics of Fluid-Phase Equilibria* 3rd edition, **1999**
- [35] Marrero, J.; Gani, R.; Group-contribution based estimation of pure component properties *Fluid Phase Equilibria* **2001**, 183-208
- [36] Daldrup-Grosse, J. B.; Held, C.; Feelly, R.; Schembecker, G.; Sadowski, G., Measurement and Modeling Solubility of Aqueous Multisolute Amino-Acid Solutions. *Industrial & Engineering Chemistry Research* **2010**, *49* [3], 1395-140

- [37] Anderson, T.F.; Prausnitz, J. M., Application of the UNIQUAC Equation to Calculation of Multicomponent Phase Equilibria. 2. Liquid-Liquid Equilibria, *Industrial & Engineering Chemistry Process Design and Development* **1978**, 17 [4], 561-567
- [38] Hansen, H. K.; Rasmussen, P.; Fredenslund, A.; Schiller, M. and J Gmehling Vapor-liquid equilibria by UNIFAC group contribution. 5. Revision and extension *Ind. Eng. Chem. Res.* **1991** [30], 2352-2355
- [39] Wittig, R.; Lohmann, J. and Gmehling, J. Vapor-Liquid Equilibria by UNIFAC Group Contribution. 6. Revision and Extension *Ind. Eng. Chem. Res.* **2003** [42], 183-188
- [40] Balslev, K. and Abildskov, UNIFAC Parameters for Four New Groups *J. Ind. Eng. Chem. Res.* **2002** [41], 2047-205
- [41] Gmehling, J.; Onken, U. *Vapor-Liquid Equilibrium Data Collection, Aqueous-Organic Systems* Vol. 1, Part 1 **1977**

Supplemental Data

Example: Glycine

In pure liquid water, the solubility of glycine, x_1^s , has been measured as $5.67 \cdot 10^{-2}$ mole glycine·mole H_2O^{-1} . In pure liquid ethanol, the solubility of glycine, x_1^s , has been measured as $4.59 \cdot 10^{-5}$ mole glycine·mole ethanol⁻¹. In a water-ethanol mixed solvent with an ethanol solvent mole fraction, x_2 , of 0.629, the solubility of glycine, x_1^s , has been measured as $2.82 \cdot 10^{-3}$ mole glycine·mole solvent⁻¹. This will be used to validate the model.

The estimations of $T_m = 579$ K and $\Delta h_{T_m} = 2550$ kJ·mol⁻¹ were calculated using the Marrero and Gani method³⁵. Using Equation (9), $(f^s/f^L)_{pure} = 1.58 \cdot 10^{-2}$ at 298.15 K.

Using Equation (6.2) and the experimental data and the just calculated $(f^s/f^L)_{pure}$, $\gamma_{12}^s = 2.79 \cdot 10^{-1}$ and $\gamma_{13}^s = 3.45 \cdot 10^{-2}$. Using Equations (6.13) and (6.14) and $q = 2.46$, $A_{12} = -1.66$ and $A_{13} = 5.85$. Using experimental vapor-liquid equilibrium data for amino-acid-free water-ethanol mixtures³⁶ and Equations (6.15) and (6.21), $A_{23} = 0.778$. Using equations (6.17), (6.18) and (6.19), $A_{21} = -0.673$, $A_{31} = 2.16$ and $A_{32} = 0.846$.

Using Equation (6.23) the activity coefficient of glycine in a water-ethanol mixed solvent, γ_1 , can be calculated at a given x_2 . At $x_2 = 0.629$, then $\gamma_1^s = 3.53 \cdot 10^{-2}$. Using Equation (6.2), the calculated solubility of glycine, $x_1^s = 2.79 \cdot 10^{-3}$ glycine·mole solvent⁻¹, only $0.03 \cdot 10^{-3}$ mole glycine·mole solvent⁻¹ lower than the experimental data.

All calculation steps are shown schematically in Figure S1.

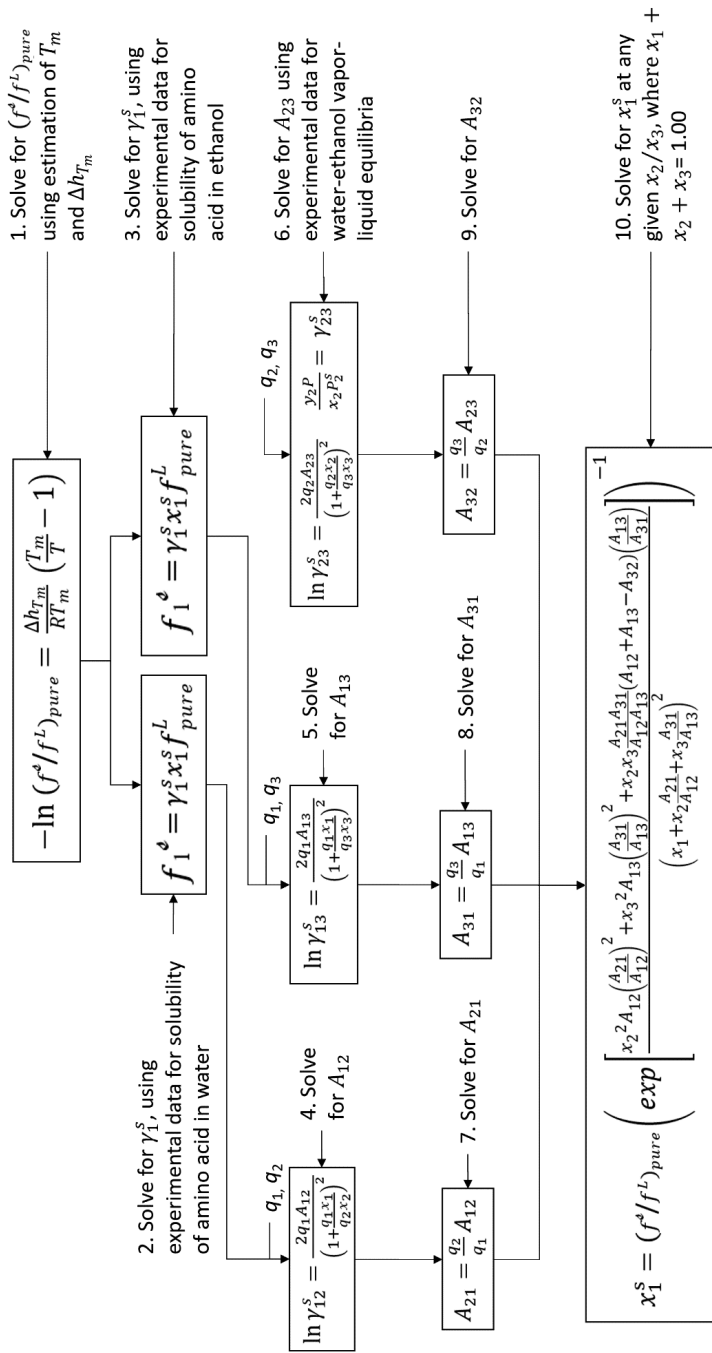


Figure S1: Schematic representation of the calculations used in predicting solubility, x_1^S , of an amino acid in a mixed solvent of water-ethanol as presented in this work. Explanations of each variable and the constants q_1 , q_2 , and q_3 are given in the main text of this work.

Chapter 7

Fractional precipitation of amino acids from agro-industrial residues using ethanol

Published as: Widyarani, Bowden, N.A., Kolfschoten, R.C., Sanders, J.P.M. and Bruins, M.E. Fractional precipitation of amino acids from agro-industrial residues using ethanol. *Industrial & Engineering Chemistry Research* **2016** 55, 7462-7472

Abstract:

Amino acids are important in human and animal diet, as well as being potential feedstocks for chemical production. Amino acids can be obtained from protein after hydrolysis. In addition, several agro-industrial residues already contain a mixture of free amino acids. The objective of this study was to develop a method for amino acids separation, starting from mixtures containing amino acids, and using anti-solvent precipitation with ethanol. Protamylasse™, rubber seed protein hydrolysates, and grass juice were used in the experiments, representing existing and potential agro-industrial residues. Our results show that in a water-ethanol system, some amino acids had lower solubility in mixtures than as a single component, thereby facilitating precipitation. A sufficiently high total amino acid concentration in the mixture is needed to achieve precipitation, therefore a concentration step is sometimes required. Ethanol precipitation can be applied as a pre-treatment to separate mixtures into groups of amino acids or a polishing step to increase purity.

7.1 Introduction

Value-added products can be obtained from proteinaceous fractions in a biorefinery framework. Native proteins, functional peptides, and essential amino acids can be used for food or feed applications. Their presence is often limiting and therefore determines the quality and price. Amino acids also have functionalities that already resemble traditional chemical products, therefore they are interesting as intermediate building blocks for nitrogen-containing chemicals [1].

Amino acids can be produced by hydrolysis of proteins from biofuel residues, e.g. oil seed press cake or dried distiller grains with solubles [2,3]. This assures feedstock availability as biofuel production is also increasing [4]. In addition, several agro-industrial residues already contain a mixture of free amino acids. Examples of these agro-industrial residues include vinasse from sugar cane and sugar beet refinery [5], and Protamylasse™ that is the concentrated potato juice from potato-starch production [6]. However, since the amino acids from hydrolysis or in agro-industrial residues are present as a mixture containing multiple amino acids, a separation process is required.

Separation of single amino acids from a mixture can be performed by crystallization [7,8], reactive extraction [9], chromatography [10,11], and electrodialysis [12]. Combination of two or more of these methods is often needed to get pure compounds. These methods are mostly applied on fermentation broth that has a high concentration of one or two amino acids [13,14]. Amino acid crystallization is one of the most applied processes in industrial amino acid production through fermentation. One of the notable amino acid fermentation products is monosodium glutamate [8]. Chromatography has been applied for separation of phenylalanine from impurities, most notably tyrosine [10,11]. Enzymatic reaction or thermo-chemical treatment may alter specific amino acid properties to aid its separation from a mixture [15].

Crystallization can be performed by water removal, whether or not combined with pH shifting or the addition of an anti-solvent. The solubility of individual amino acids is the most important parameter in crystallization [16]. When using hydrolysates as amino acid source, one of the challenges is the aqueous system the amino acids are present in. These aqueous systems are often at a much lower concentration than the maximum solubility of the amino acids.

The solubility of amino acid also changes in the presence of other amino acids (Table S1 in the Supporting Information), and the mechanisms are still not fully understood [17,18]. In only water, the presence of a polar amino acid seems to increase the solubility of other polar amino acids, alanine, glycine, and cysteine [19-21]. On the other hand, the influence of

hydrophobic amino acid is not as clear. The presence of a hydrophobic amino acid does not seem to influence the solubility of other hydrophobic amino acids, except for the increase of the solubility of valine in the presence of leucine [18]. The presence of tyrosine and leucine also does not influence the solubility of glycine [17], while the presence of phenylalanine increases glycine's solubility [21]. The solubility of alanine decreases in the presence of leucine or valine and vice versa [22,23]. In these studies, the influence of other components that are naturally present in bio-based resources or are added to influence solubility is not discussed.

In water-ethanol systems, the solubility of individual amino acids decreases at different selectivity [24-26]. These differences can be used to separate groups of amino acids, however, limited studies have been done on the simultaneous influence of ethanol addition and interactions between amino acids. A study shows that when soybean or fish protein hydrolysates were mixed with aqueous ethanol followed by centrifugation, the relative amount of hydrophobic amino acids in the solution increased with increasing ethanol concentration [27]. However, there was no distinction between free and peptide-bound amino acids.

The objective of this study was to develop an energy-efficient method for amino acid separation from aqueous system containing amino acid mixtures. The effectiveness of ethanol as an anti-solvent was investigated. Experiments were performed using Protamylasse™, hydrolysate of rubber seed protein, and grass juice. This is the first article describing the interactions between amino acids in aqueous-ethanolic system, to be applied on complex materials containing multiple amino acids. Protamylasse™ contains 150-180 g of proteins, peptides, and free amino acids per liter; asparagine is notably abundant [6,28]. Interest is growing for using rubber seeds for oil and protein production [29,30]. Rubber seed proteins (RSP) contain high amounts of aspartic acid, glutamic acid, arginine, valine, and leucine. Hydrolysis of these proteins with proteases, however, results in mixtures with different free amino acid composition. Grass juice is the liquid product after grass pressing and protein precipitation. The liquid still contains a mixture of amino acids that are interesting to use for further valorization.

7.2 Materials and Methods

Protamylasse™ is a residue from potato starch production. Potato juice is obtained after separation of the starch and fiber from the potatoes. Part of the proteins from the potato juice is separated via steam coagulation. The remaining liquid fraction is concentrated via

water evaporation to obtain Protamylasse™. Protamylasse™ (51% dry matter) for the experiments was obtained from AVEBE (the Netherlands) that produces 120,000 m³ Protamylasse™ annually [6].

Rubber seed protein (RSP) hydrolysates were prepared via hydrolysis of RSP concentrate using protease and dilute acid³. Protein concentrate was immersed in water at 5% w-protein/w-water and 0.01 M Britton-Robinson buffer was used to keep the pH at 7. Pronase and Peptidase R (1% w-protease/w-protein each) were added at t = 0.5 h and t = 2 h, respectively, and total incubation time was 24.5 h. To stop protease activity, the mixture was incubated at 90°C for 10 min. Part of the mixture was centrifuged at 3000 x g, 4°C, 20 min, the supernatant was removed and filtered through Schleicher and Schuell filter No. 604 to get RSP enzymatic hydrolysate. Another part of the mixture that had not been centrifuged was mixed with 6 M HCl at a ratio of 5:1 (w-mixture/w-acid) and incubated at 95°C for 48 h to further hydrolyze the mixture. After the incubation, this mixture was let cool until room temperature and the pH was neutralized using 6 M and 0.1 M NaOH; the final pH was 7.6. The mixture was filtered through Schleicher and Schuell filter No. 604 to get RSP combined hydrolysate. The latter shows a broader and extended amino acid pattern compared to the RSP enzymatic hydrolysate, and the glutamine and asparagine have been converted to glutamic acid and aspartic acid due to the high acid concentrations.

Grass juice is the residue from protein production from grass. Grass is ground, pulped, and extruded to separate the protein-rich liquid fraction from the fiber-rich solid fraction. Protein in the liquid fraction is separated via steam coagulation or isoelectric precipitation. The amino acids are not separated during this process and remain in the liquid (grass juice). Grass juice for the experiments was obtained from the pilot plant of Grassa BV (the Netherlands) that produces 0.6 m³ grass juice for every tonne grass processed³¹. The Netherlands is projected to process 5 million tonnes fresh grass/year (3 million m³ grass juice) in 2021^{31,32}.

Chemicals used in the experiments and analyses were of analytical grade. Pronase was obtained from Roche Diagnostics (Germany). Peptidase R was obtained from Amano (Japan).

Precipitation with fixed starting material concentration and varying ethanol concentration

Prior to the experiment, RSP enzymatic hydrolysate was concentrated in a rotary evaporator to a concentration of 410 µmol/g free amino acids. Protamylasse™ or concentrated hydrolysate was added into empty tubes at a fixed amount equivalent to 5% w/w in the starting mixture before precipitation. Milli-Q water was added at decreasing amounts to each

tube. Subsequently, ethanol was added at increasing amounts to obtain mixtures with 0-95% w/w ethanol concentration. The tubes were mixed by vortex for 10 s, then immediately centrifuged at 7000 x g, 20°C for 5 min. The supernatant was transferred into empty containers using graduated pipettes, and stored at -18°C until analyzed.

Precipitation with fixed ethanol concentration and varying starting material concentration

Protamylasse™ was added into empty tubes at amounts equivalent to 5-40% w/w in the starting mixture before precipitation. Milli-Q water was added at decreasing amounts to make the total water fraction of 40% w/w. Subsequently, ethanol was added to obtain mixtures with 60% w/w ethanol concentration. The tubes were mixed by vortex for 10 s, then immediately centrifuged at 7000 x g, 20°C for 5 min. The supernatant was transferred into empty containers using graduated pipettes, and stored at -18°C until analyzed.

Precipitation with both ethanol and starting material concentrations varied

Prior to experiment, Protamylasse™ was diluted with Milli-Q water to a concentration of 213 µmol/g free amino acids. Grass juice was concentrated with rotary evaporator to a concentration of 225 µmol/g free amino acid. RSP hydrolysates were used without any pre-treatment.

Diluted Protamylasse™, RSP hydrolysate, or concentrated grass juice was added into empty tubes at decreasing amounts. Subsequently, ethanol was added at increasing amounts to each tube to obtain starting mixtures with 10-95% w/w ethanol concentration. The tubes were centrifuged at 7000 x g, 20°C for 5 min. The supernatant was transferred into empty containers using graduated pipettes, and stored at -18°C until analyzed.

Amino acid analysis

Prior to measurement, frozen samples were thawed at room temperature and mixed. To measure free amino acid contents, the starting materials and supernatants from the experiments were dissolved in methanol and filtered through 0.2 µm Minisart filter. Norleucine (0.04 mM) was used as the internal standard. The filtered solutions were loaded onto Ultra-HPLC Dionex RSLC (Dionex Corporation, USA), and detections were performed at 263 nm and 338 nm³³.

An amino acid was considered to be precipitated when the concentration in the supernatant was lower than in the starting mixture. The concentrations were compared using Student's t-test; $p < 0.05$ was regarded as significant.

7.3 Results and Discussion

7.3.1 Starting material properties

ProtamylasseTM contained 48 g/kg free amino acids (Table 7.1), equivalent to 9% of the dry matter and 36% of the total protein. It also contained 129 g of sugars [28], 205 g of ash [28], and 174 g of organic acids (own measurement) per kg. Citric acid (94 g/kg) was the most abundant organic acid. Malic acid, lactic acid, and acetic acid concentrations were 24 g/kg, 29 g/kg, and 27 g/kg, respectively. ProtamylasseTM also contained 8 g of γ -aminobutyric acid (GABA) per kg.

RSP enzymatic hydrolysate contained 13 g/kg free amino acids (Table 7.1), equivalent to 22% of the dry matter and 56% of the total protein. RSP combined hydrolysate contained 23 g/kg free amino acids (Table 7.1), equivalent to 23% of the dry matter and 44% of the total protein. GABA concentrations in both hydrolysates were less than 0.1 g/kg. Sugars and lipid contents were not determined in this study, but it was assumed that they were present in (partially) hydrolyzed forms.

Grass juice contained 4 g/kg free amino acids (Table 7.1), equivalent to 6% of the dry matter. It also contained 3 g of sugars [34], 12 g of ash [34], 0.9 g of GABA, and 45 g of organic acids (own measurement) per kg. Malic acid (19 g/kg) and lactic acid (25 g/kg) were the most abundant organic acids.

Table 7.1. Free amino acid composition of the starting materials

Amino acid	Amino Acid	pI ^a	Concentration in starting material (μmol/g)			
			Protamy-lasse™ (pH = 5.8)	RSP enzymatic hydrolysate (pH = 5.7)	RSP combined hydrolysate (pH = 7.6)	Grass juice (pH = 6)
Polar uncharged:						
Asparagine	Asn	5.4	135 ± 3	3 ± 0	0 ± 0	0.9 ± 0.1
Glutamine	Gln	5.7	0 ± 0	6 ± 0	0 ± 0	0.4 ± 0.0
Serine	Ser	5.7	13 ± 0	4 ± 0	11 ± 0	2.4 ± 0.0
Threonine	Thr	5.6	7 ± 3	3 ± 0	6 ± 0	1.6 ± 0.1
Negative:						
Aspartic acid	Asp	2.8	69 ± 2	3 ± 0	29 ± 0	5.0 ± 0.0
Glutamic acid	Glu	3.2	31 ± 1	6 ± 0	28 ± 0	2.5 ± 0.0
Positive:						
Arginine	Arg	10.8	21 ± 0	1 ± 0	5 ± 0	1.1 ± 0.0
Histidine	His	7.6	3 ± 0	2 ± 0	3 ± 0	0.2 ± 0.0
Lysine	Lys	9.7	10 ± 1	3 ± 0	4 ± 0	1.3 ± 0.1
Hydrophobic:						
Isoleucine	Ile	6.0	6 ± 0	7 ± 0	6 ± 0	1.3 ± 0.0
Leucine	Leu	6.0	5 ± 0	17 ± 0	15 ± 0	2.4 ± 0.0
Valine	Val	6.0	17 ± 0	21 ± 0	18 ± 0	2.3 ± 0.1
Phenylalanine	Phe	5.5	6 ± 0	7 ± 0	7 ± 0	1.2 ± 0.1
Tryptophan	Trp	5.9	0 ± 0	2 ± 0	0 ± 0	0.2 ± 0.0
Tyrosine	Tyr	5.7	5 ± 0	0 ± 0	1 ± 0	0.7 ± 0.0
Methionine	Met	5.7	2 ± 0	2 ± 0	3 ± 0	0.5 ± 0.0
Proline	Pro	6.3	5 ± 1	5 ± 0	9 ± 0	1.1 ± 0.0
Special:						
Alanine	Ala	6.0	30 ± 1	9 ± 0	16 ± 0	5.4 ± 0.1
Glycine	Gly	6.0	4 ± 0	3 ± 0	18 ± 0	1.5 ± 0.1
Cysteine	Cys	5.1	0 ± 0	0 ± 0	2 ± 0	0.0 ± 0.0
Total (μmol/g)			368 ± 8	104 ± 0	181 ± 0	32 ± 1
Total (g/kg)			48 ± 1	13 ± 0	23 ± 0	4 ± 0

^a Isoelectric point at 25°C³⁵

7.3.2 Amino acid precipitation at fixed starting material concentration

For experiments at fixed starting material concentration, the starting mixture contained 5% w/w starting material and water-ethanol at various concentrations. Figure 7.1 shows that overall, RSP enzymatic hydrolysate and Protamylasse™ were completely soluble up to 50% and 60% ethanol, respectively. However, different amino acids precipitated at different ethanol concentrations. The amino acid fractionation between the supernatant and the precipitate were also different. The concentration of an amino acid in each supernatant was related to both solubility and initial concentration in the mixture. In complex mixtures such as Protamylasse™ and RSP hydrolysate, the maximum solubility of individual amino acids was different from their solubility in mixtures that only contain one amino acid.

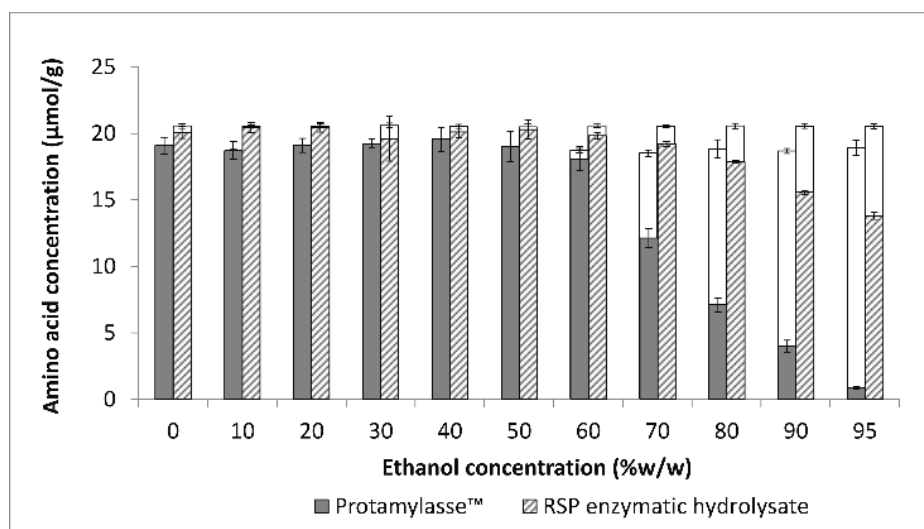


Figure 7.1. Total amino acid concentrations in the starting mixtures (unfilled bars) and supernatants from Protamylasse™ and RSP enzymatic hydrolysate at different ethanol concentrations, from experiments with fixed starting material concentration and varying ethanol concentration.

Figure 7.2a, 7.2b, and 7.2c show the amino acid concentrations in the starting mixture and supernatant at 50%, 70%, and 95% ethanol, respectively. Amino acid concentrations at other ethanol concentrations are presented in Table S2-S5 in the Supporting Information. The concentrations were compared to the maximum solubilities of individual amino acids from literature, as shown in Figure 7.2. As bio-based sources, amino acids in Protamylasse™ and RSP enzymatic hydrolysate were most likely present as l-isomers. Consistent data on maximum solubility in aqueous ethanol is scarce because most studies only focused on a few

amino acids. Therefore, the data has been compiled from several sources. Data for methionine, arginine, lysine, and cysteine are not available.

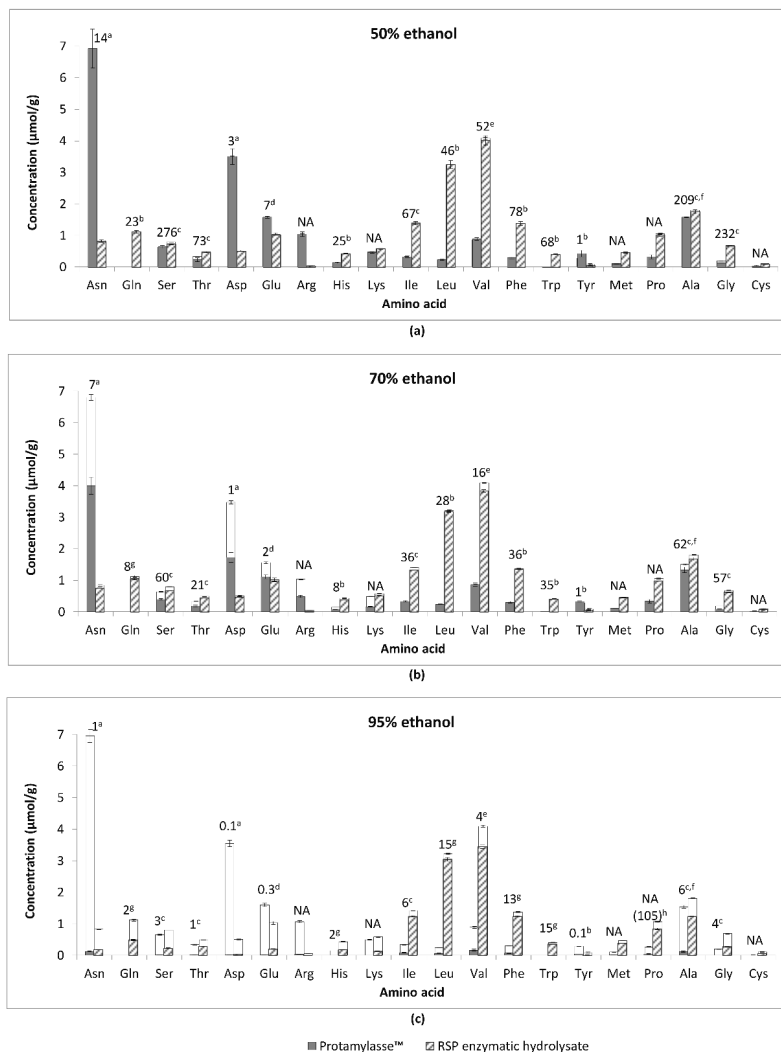


Figure 7.2. Amino acid concentrations in the starting mixtures (unfilled bars) and supernatants of Protamylasse™ and RSP enzymatic hydrolysate at 50% (a), 70% (b), and 95% (c) ethanol, from experiments with fixed starting material concentration and varying ethanol concentration. Numbers above bars indicate maximum solubility of l-isomer at 25°C, unless otherwise stated. NA = data not available. ^a Interpolated from McMeekin *et al.*[36]; ^b Interpolated from Nozaki and Tanford [26]; ^c Interpolated from Ferreira *et al.*[24,25]; ^d Interpolated from Dunn and Ross [37]; ^e Interpolated from Zhang *et al.*[38,39], maximum solubility at 20°C; ^f Data for DL-alanine; ^g Extrapolated from Nozaki and Tanford [26]; ^h Data not available. Number between brackets shows maximum solubility in 100% ethanol at 19 °C [40].

Almost all amino acids in our experiments precipitated at concentrations below their maximum solubility as single amino acids (Figure 7.2). The exception was aspartic acid from Protamylasse™ that at 50% ethanol already exceeded its maximum solubility based on data for the single amino acid (Figure 7.2a), but only precipitated at 70% ethanol and higher (Figure 7.2b). The solubility of some amino acids increases in each other's presence, particularly for polar amino acids [19-21]. The decrease we observed, therefore, might be due to different interaction patterns between amino acids in the presence of ethanol or due to the presence of non-amino acid components.

At 70% ethanol, all polar amino acids from Protamylasse™ started to precipitate while the hydrophobic ones stayed in the solution (Figure 7.2b). The combined aspartic acid/asparagine fraction increased from 56% (mol/mol) in the starting material to 69% in the precipitate. The combined aspartic acid/asparagine fraction in the precipitate decreased at higher ethanol concentration as more amino acids precipitated, and was as low as 57% again at 95% ethanol where practically all amino acids precipitated (Figure 7.2c).

At 70% ethanol, all amino acids from RSP enzymatic hydrolysate precipitated except aspartic acid, glutamic acid, tyrosine, tryptophan, phenylalanine, and leucine (Figure 7.2b). The precipitated amino acids were 6-18% of the amount in the starting material except for arginine (58%) and cysteine (32%); both were present at low concentrations and therefore uncertainty in measurements was high. At 95% ethanol, however, 79-95% hydrophobic amino acids were still present in the supernatant (Figure 7.2c), except tyrosine that has lower solubility than the other amino acids.

The relative abundance of hydrophobic amino acids in the supernatant of RSP enzymatic hydrolysate suggests that at high ethanol concentrations, interactions between hydrophobic amino acids resulted in the increase of overall solubility of hydrophobic amino acids, which is similar to the influence of polar amino acids interactions in water-only solution [17-23]. However, this was not observed in Protamylasse™ at 95%. RSP enzymatic hydrolysate had 59% (mol/mol) hydrophobic amino acids in the starting material, much higher than Protamylasse™ that only had 13%. This suggests that the increase of hydrophobic amino acid solubility in high ethanol concentration may only occur in starting materials with a sufficiently high fraction of hydrophobic amino acids.

The experiments were performed at 20°C. Incubation at low temperature may result in lower solubility and consequently, more amino acids should be precipitated. Experiments with Protamylasse™, however, did not show any differences between mixtures directly separated and mixtures incubated at 4°C for two weeks.

7.3.3 Influence of starting material concentration

When maximum solubility is the sole factor that determines the amount of amino acids that remain in the supernatant, it is expected that the amino acids concentration in the supernatant is constant regardless of the concentration in the starting mixture. Therefore, we performed an experiment with different starting material concentrations in a fixed ethanol concentration at which changes would be most measurable. At 50% ethanol, all amino acids from Protamylasse™ were still present in the supernatant (Figure 7.2a) while at 70% ethanol, some amino acids already precipitated but some were still present in the supernatant (Figure 7.2b). This shows that the 60% ethanol concentration was the most sensitive to changes.

When precipitation occurred at 60% ethanol, the total amino acid concentrations from Protamylasse™ in supernatants were always around $27 \pm 2 \mu\text{mol/g}$ as expected. The change was relatively small compared to the increase in Protamylasse™ concentration (Figure 7.3a). For individual amino acids, this was not always the case. At the start not all amino acids were at maximum solubility yet. Also after precipitation, the relative composition changed and this influenced the solubility of individual amino acids in the water-ethanol mixture.

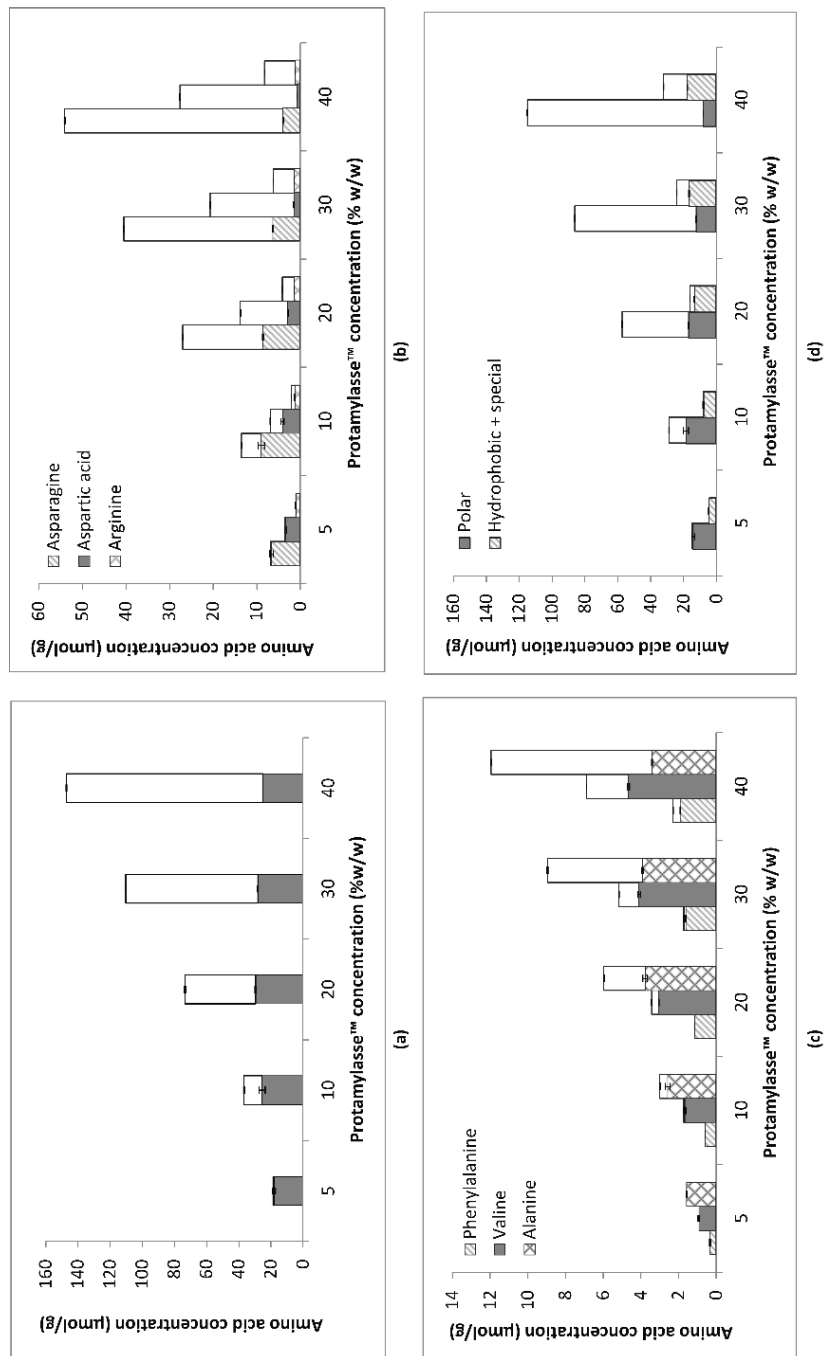


Figure 7.3. Total amino acid (a), selected polar amino acid (b), selected hydrophobic and special amino acid (c), and combined polar amino acids and hydrophobic and special amino acids (d) concentrations in the starting mixture (unfilled bars) and supernatant at 60% ethanol and different Protamylasse™ concentration. Data for other amino acids are presented in Table S6 and S7 in the Supporting Information.

Polar amino acids

At 60% ethanol, the maximum solubility of asparagine that present as a single amino acid is 10 $\mu\text{mol/g}$ [36]. At 5% Protamylasse™ concentration, asparagine was completely soluble in 60% ethanol (Figure 7.3b). Precipitation was observed at 10% Protamylasse™; the asparagine concentration in the supernatant was $9 \pm 1 \mu\text{mol/g}$, similar to the maximum solubility. At higher Protamylasse™ concentrations, however, the amino acid concentrations in the supernatant decreased despite more amino acids were present in the starting mixture.

Aspartic acid concentration in the supernatant at 60% ethanol and 10% Protamylasse™ was 4 $\mu\text{mol/g}$, higher than the maximum solubility as single amino acid (2 $\mu\text{mol/g}$ [36]). At higher Protamylasse™ concentrations, the amino acid concentrations in the supernatant also decreased and were even lower than its maximum solubility at 30% and 40% Protamylasse™ (Figure 7.3b).

Other polar amino acids: arginine, lysine, histidine, glutamic acid, and serine also started to precipitate at 10% Protamylasse™. Precipitation occurred despite the maximum solubility was higher than the starting mixture concentration for some amino acids, e.g. glutamic acid and serine. Glycine, which is considered to be a non-polar amino acid, also showed similar pattern.

Hydrophobic amino acids

The maximum solubilities of phenylalanine and valine as single amino acid in 60% ethanol are 65 $\mu\text{mol/g}$ and 84 $\mu\text{mol/g}$, respectively [26,39]. Figure 7.3c shows that phenylalanine and valine concentrations were lower than the maximum solubility. Precipitation occurred at 20% Protamylasse™ for valine and 30% Protamylasse™ for phenylalanine, higher than the required Protamylasse™ concentration for precipitation of polar amino acids. Even then, the supernatant concentration still increased at higher Protamylasse™ concentrations. Similar patterns were observed for other hydrophobic amino acids: proline, tyrosine, methionine, tryptophan, isoleucine, and leucine.

Alanine

At 60% ethanol, the maximum solubility of dl-alanine as a single amino acid is 105 $\mu\text{mol/g}$ [24]; the solubility of L-alanine may be lower but the difference should not be more than one order of magnitude. Even though alanine concentrations in the starting materials were lower than their maximum solubility, alanine started to precipitate at 10%

Protamylasse™ concentration (Figure 7.3c). With the increase of Protamylasse™ concentration, the concentration in the supernatant first increased and then decreased again when Protamylasse™ concentration increased further. The pattern was between that observed for the polar and hydrophobic amino acids, suggesting what would happen to hydrophobic amino acids if higher Protamylasse™ concentration can be applied.

Changes in composition as a result of higher starting material concentrations

With increasing Protamylasse™ concentration, the concentration of polar amino acids in the supernatant decreased, while the concentration of hydrophobic amino acids increased (Figure 7.3b-d). Consequently, amino acid composition in the supernatant changed. At 5% Protamylasse™, the supernatant consisted of 76% polar amino acids and 24% hydrophobic and special amino acids while at 40% Protamylasse™, the supernatant consisted of 30% polar amino acids and 70% hydrophobic and special amino acids. This again suggests the increase of overall solubility of hydrophobic amino acids in water-ethanol solution due to interactions between hydrophobic amino acids. However, as the influence of single amino acids was not measured independently, the solubility might also be influenced by the other components that also increased when more Protamylasse™ was present. This effect will be discussed in subsection 3.5.

7.3.4 Precipitation by variation of ethanol and starting material concentrations

A more practical approach to anti-solvent separation is the simple addition of ethanol to the starting material. This will enable precipitation at lower ethanol concentration than the previous experiments since no water is added to the mixture. Figure 7.4a-d show that precipitation occurred for all starting materials, but precipitation started at different ethanol concentrations. The amino acid fractionation between the supernatant and the precipitate were also different.

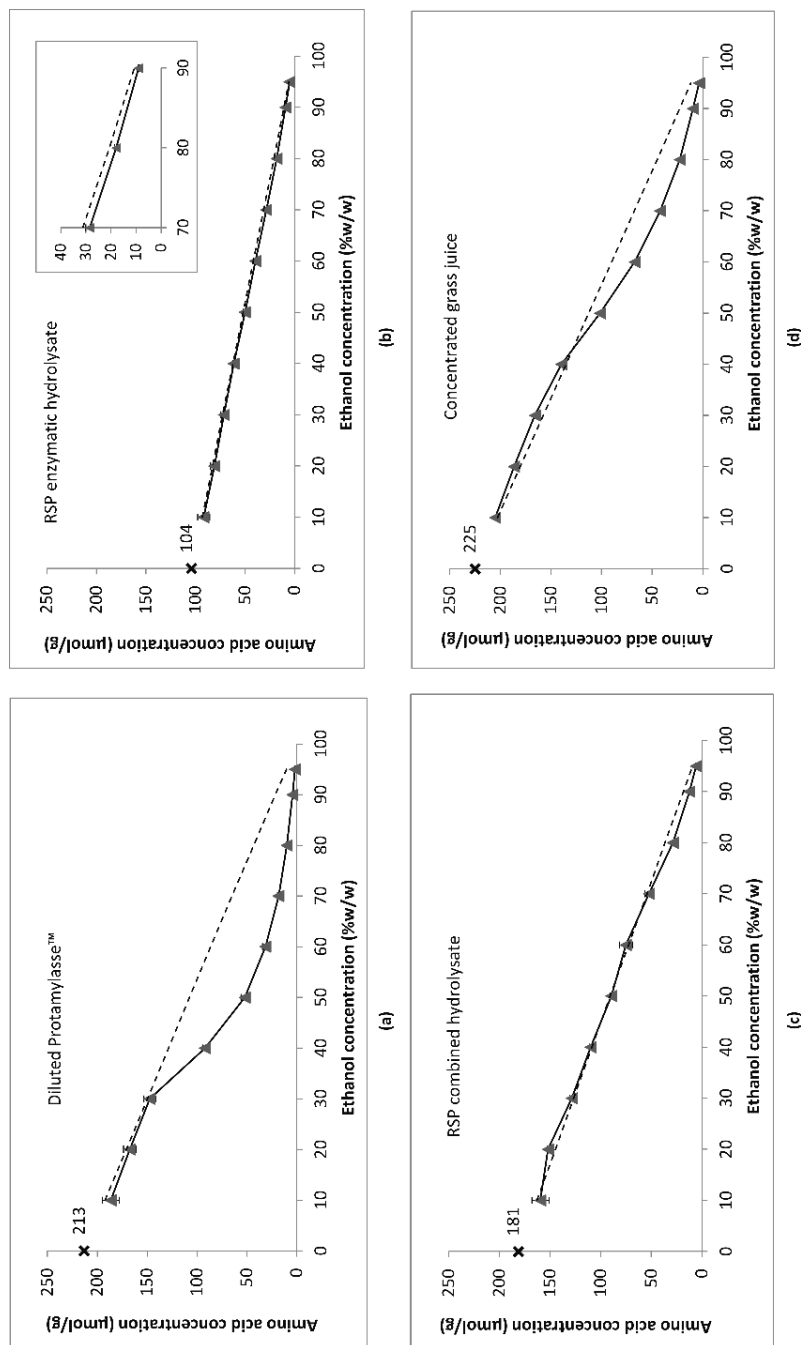


Figure 7.4. Amino acid concentration in the supernatant from diluted Protamylasse™ (a), RSP enzymatic hydrolysate (b), RSP combined hydrolysate (c), and concentrated grass juice (d). Dashed lines indicate amino acid concentrations in the starting mixtures. x = Total amino acid concentration in the starting material before ethanol addition. Axis titles on the inset of Figure 4b are similar to the main graph.

Protamylasse™ started to precipitate at 40% ethanol, and at this concentration already 28% of total amino acids in the starting mixture precipitated (Figure 7.4a). RSP enzymatic hydrolysate, on the other hand, only started to precipitate at 80% ethanol and only 19% of total amino acids in the starting mixture precipitated (Figure 7.4b). This is probably due to the low amino acid concentrations in the mixture. Protamylasse™ and grass juice had the highest and comparable total amino acid concentration; they had more precipitation compared to RSP hydrolysates (Figure 7.4a-d). This may suggest that there was a minimum total amino acid concentration that was required to achieve precipitation and below this value, precipitation did not occur or occurred only at high ethanol concentration. This was the case with Protamylasse™, which is industrially concentrated potato juice that is more than ten times thicker than the original material. This was also illustrated in RSP enzymatic hydrolysate precipitation, which could precipitate at lower ethanol concentration (60%) when concentrated starting material was used (Figure 7.1). The minimum total amino acid requirement, however, may be different for different starting materials.

The total amino acid concentration of RSP combined hydrolysate was roughly 20% lower than Protamylasse™ and grass juice (Figure 7.4a, 4c, 4d). Unlike these two materials, RSP combined hydrolysate started to precipitate at much higher ethanol concentration (80%), similar to RSP enzymatic hydrolysate (Figure 7.4b, 4c). The large difference between precipitations of RSP combined hydrolysate and Protamylasse™ or grass juice suggests that there were influences of other factors, particularly starting material composition.

With increasing ethanol concentrations, the composition of the amino acid groups in the supernatant shifted from polar to hydrophobic amino acids. The shift was most apparent in Protamylasse™ (Figure 7.5a), which had the least hydrophobic amino acids in the starting material. Also, the supernatant from grass juice showed an apparent shift and had even higher hydrophobic amino acids content at 90% and 95% ethanol compared to Protamylasse™ (Figure 7.5d). RSP enzymatic hydrolysate had relatively the most hydrophobic amino acids in the starting material, but the low mixture concentration resulted in only small changes in supernatant composition at higher ethanol concentration (Figure 7.5b). The low mixture concentration also influenced RSP combined hydrolysate (Figure 7.5c), where precipitation mostly occurred on aspartic and glutamic acid.

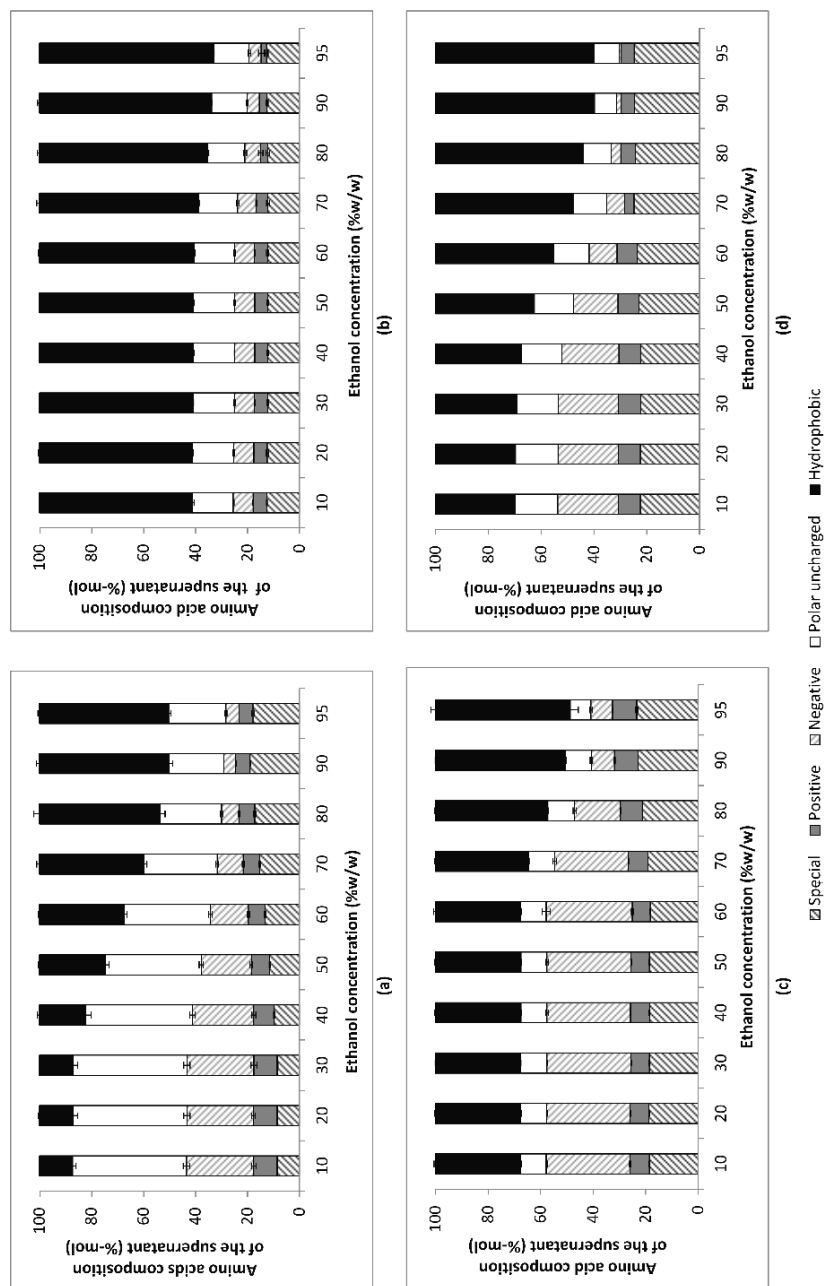


Figure 7.5. Distribution of special, positive, negative, polar uncharged, and hydrophobic amino acid groups in the supernatant after precipitation of diluted Protamylase™ (a), RSP enzymatic hydrolysate (b), RSP combined hydrolysate (c), and concentrated grass juice (d) at different ethanol concentrations.

On a mole-fraction basis, grass juice contained more hydrophobic amino acids than Protamylasse™ (Table 7.1). This might be the reason the precipitation started at higher ethanol concentration, even though the starting mixture concentration was higher (Figure 7.4a, 4d). On the other hand, the absolute concentrations of hydrophobic amino acids were almost similar in the two RSP hydrolysates (Table 7.1), despite the different relative compositions due to the higher concentrations of aspartic and glutamic acid in the RSP combined hydrolysate (Figure 7.5b, 5c). The similarity of hydrophobic amino acids content might therefore have resulted in similar precipitation behavior. Next to amino acids, the presence of non-amino acids component might also influence the precipitation from these materials.

7.3.5 Influence of non-amino acid components in the mixture

Next to amino acids, the starting materials used in our experiments also contained other components shows that all amino acids except polar charged amino acids have isoelectric points between 5.0 and 6.3 [35], which were close to the pH of all starting materials except RSP combined hydrolysate that had a slightly higher pH (7.6). At pH 5, metal salts can form insoluble complex with amino acids, but the bonds are likely broken at pH 7 [46]. Our experiments showed that polar uncharged and hydrophobic amino acids showed different precipitation patterns. Furthermore, both positively-charged amino acids (pI 7.5-10.8) and negatively-charged amino acids (pI 2.8-3.2) precipitated with similar pattern as polar uncharged amino acids. This suggests that for dilute mixtures at pH close to 7, hydrophobicity of the amino acid side chains, ethanol concentration, and starting material concentration had more influence than pH.

7.3.6 Applications

The results show that ethanol can be best applied in amino acid separation as either a pre-treatment to separate amino acid groups or a polishing step to increase purity. The parameters that need to be considered are the minimum amino acid concentration in the mixture, the ethanol concentration, and the mixture composition, especially the ratio between hydrophobic and polar amino acids. Furthermore, the presence of non-amino acid components also may influence the separation.

In our cases, a mixture with total amino acid concentration of 200 $\mu\text{mol/g}$ or higher is needed to achieve precipitation at 50% ethanol. For mixtures with lower concentrations, a pre-treatment step to concentrate is necessary. This might however result in a viscous mixture. For instance, while the original potato juice only has 4% dry matter, Protamylasse™ in our experiment had 368 $\mu\text{mol/g}$ amino acids and 51% dry matter. Based on visual observation, the material was very viscous. On the other hand, concentrated RSP enzymatic hydrolysate had 410 $\mu\text{mol/g}$ amino acids and was still sufficiently fluid with only 23% dry matter. The high viscosity of Protamylasse™, therefore, was likely due to the abundant presence of other components like sugars. Viscosity reduction can be achieved by ethanol addition, which also increases diffusion rates. In practice, the mixture of the starting material and ethanol should have 15% dry matter or less to enable easy processing.

Amino acid composition determines the required ethanol concentrations for precipitation, and the window of operation may vary between starting materials. Two possible processing scenarios are proposed:

For mixtures that are rich in polar amino acids, 50-70% ethanol can be applied to obtain a precipitate that is rich in polar amino acids. For the case of Protamylasse™, the combined fraction of aspartic acid and asparagine increased from 56% (mol/mol) in the starting material to 69% in the precipitate. Ethanol concentrations of 90% or higher should be avoided, as at these concentrations, hydrophobic amino acids may also precipitate.

For mixture rich in hydrophobic amino acids, 90% ethanol or higher is required to increase the fraction of hydrophobic amino acid in the liquid. In these concentrations, most polar amino acids will precipitate. For the case of RSP enzymatic hydrolysate with prior concentrating step, hydrophobic amino acid fraction increased from 59% (mol/mol) in the starting material to 76% in the supernatant.

Alternatively, a two-step precipitation can be applied: Step 1 to precipitate most (>90%) polar amino acids, while a small number of hydrophobic ones also precipitate. Step 2 can be used to further purify the precipitated polar amino acids, by re-solubilizing the hydrophobic amino acids.

7.4 Conclusions

Ethanol can be applied in the fractionation of amino acids from protein hydrolysate and agro-industrial residues. In such complex mixtures, interactions between amino acids influence the solubility. Our results show that in a water-ethanol system, some amino acids have lower solubility in a mixture than as a single component, facilitating precipitation.

Precipitation of polar amino acids mostly occurs at lower ethanol concentration compared to the hydrophobic ones. Meanwhile, interactions between hydrophobic amino acids in ethanol presence may lead to the increase of overall solubility of hydrophobic amino acids.

Ethanol precipitation can be applied as a pre-treatment to separate mixtures into groups of amino acids or as a polishing step to increase purity. For dilute mixture at a pH close to 7, precipitation is determined by hydrophobicity of the amino acid side chains, ethanol concentration, and starting material composition. In the case of very dilute stream, a concentration step is required to get a sufficiently high mixture concentration to achieve precipitation.

References

- [1] Scott, E.; Peter, F.; Sanders, J. P. M. Biomass in the Manufacture of Industrial Products—the Use of Proteins and Amino Acids. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 751–762.
- [2] Lammens, T. M.; Franssen, M. C. R.; Scott, E. L.; Sanders, J. P. M. Availability of Protein-Derived Amino Acids as Feedstock for the Production of Bio-Based Chemicals. *Biomass Bioenergy* **2012**, *44*, 168–181.
- [3] Sari, Y. W.; Alting, A. C.; Floris, R.; Sanders, J. P. M.; Bruins, M. E. Glutamic Acid Production from Wheat by-Products Using Enzymatic and Acid Hydrolysis. *Biomass Bioenergy* **2014**, *67*, 451–459.
- [4] International Energy Agency. *World Energy Outlook 2014 Factsheet*; 2014.
- [5] Hidalgo, K. Vinasse in Feed: Good for Animal and Environment. *Feed Tech* **2009**, *13* [5], 18–20.
- [6] Mooibroek, H.; Oosterhuis, N.; Giuseppin, M.; Toonen, M.; Franssen, H.; Scott, E.; Sanders, J. P. M.; Steinbüchel, A. Assessment of Technological Options and Economical Feasibility for Cyanophycin Biopolymer and High-Value Amino Acid Production. *Appl. Microbiol. Biotechnol.* **2007**, *77*, 257–267.
- [7] Black, S. N.; Davey, R. J. Crystallisation of Amino Acids. *J. Cryst. Growth* **1988**, *90* [, 136–144.
- [8] Sano, C. History of Glutamate Production. *Am. J. Clin. Nutr.* **2009**, *90*, 728S–732S.
- [9] Cascaval, D.; Oniscu, C.; Galaction, A.-I. Selective Separation of Amino Acids by Reactive Extraction. *Biochem. Eng. J.* **2001**, *7* [3], 171–176.
- [10] de Hollanda e Vasconcellos, A. M.; Neto, A. L. C. S.; Grassiano, D. M.; de Oliveira, C. P. H. Adsorption Chromatography of Phenylalanine. *Biotechnol. Bioeng.* **1989**, *33* [10], 1324–1329.
- [11] Otani, M.; Sano, C.; Kusumoto, I. Refining Phenylalanine. US4584400 A, April 22, 1986.
- [12] Sandeaux, J.; Sandeaux, R.; Gavach, C.; Grib, H.; Sadat, T.; Belhocine, D.; Mameri, N. Extraction of Amino Acids from Protein Hydrolysates by Electrodialysis. *J. Chem. Technol. Biotechnol.* **1998**, *71* [3], 267–273.
- [13] Ivanov, K.; Stoimenova, A.; Obreshkova, D.; Saso, L. Biotechnology in the Production of Pharmaceutical Industry Ingredients: Amino Acids. *Biotechnol. Biotechnol. Equip.* **2013**, *27* [2], 3620–3626.

- [14] Shiratsuchi, M.; Kuronuma, H.; Kawahara, Y.; Yoshihara, Y.; Miwa, H.; Nakamori, S. Simultaneous and High Fermentative Production of L-Lysine and L-Glutamic Acid Using a Strain of *Brevibacterium Lactofermentum*. *Biosci. Biotechnol. Biochem.* **1995**, *59* [1], 83–86.
- [15] Teng, Y.; Scott, E. L.; Sanders, J. P. M. The Selective Conversion of Glutamic Acid in Amino Acid Mixtures Using Glutamate decarboxylase—A Means of Separating Amino Acids for Synthesizing Biobased Chemicals. *Biotechnol. Prog.* **2014**, *30* [3], 681–688.
- [16] Charmolue, H.; Rousseau, R. W. L-Serine Obtained by Methanol Addition in Batch Crystallization. *AIChE J.* **1991**, *37* [8], 1121–1128.
- [17] Carta, R. Solubilities of L-Cystine, L-Tyrosine, L-Leucine, and Glycine in Their Water Solutions. *J. Chem. Eng. Data* **1999**, *44* [3], 563–567.
- [18] Kurosawa, I.; Teja, A. S.; Rousseau, R. W. Solid–liquid Equilibria in L-Leucine + L-Valine + Water. *Fluid Phase Equilibria* **2005**, *228–229*, 83–87.
- [19] Cohn, E. J.; McMeekin, T. L.; Ferry, J. D.; Blanchard, M. H. Studies in the Physical Chemistry of Amino Acids, Peptides, and Related Sub Stances. XII. Interactions between Dipolar Ions in Aqueous Solution. *J. Phys. Chem.* **1939**, *43* [1], 169–188.
- [20] Jin, X. Z.; Chao, K. C. Solubility of Four Amino Acids in Water and of Four Pairs of Amino Acids in Their Water Solutions. *J. Chem. Eng. Data* **1992**, *37* [2], 199–203.
- [21] Soto, A.; Arce, A.; K. Khoshkbarchi, M.; Vera, J. H. Measurements and Modelling of the Solubility of a Mixture of Two Amino Acids in Aqueous Solutions. *Fluid Phase Equilibria* **1999**, *158–160*, 893–901.
- [22] Grosse Daldrup, J.-B.; Held, C.; Ruether, F.; Schembecker, G.; Sadowski, G. Measurement and Modeling Solubility of Aqueous Multisolute Amino-Acid Solutions. *Ind. Eng. Chem. Res.* **2010**, *49* [3], 1395–1401.
- [23] Kuramochi, H.; Noritomi, H.; Hoshino, D.; Nagahama, K. Measurements of Solubilities of Two Amino Acids in Water and Prediction by the UNIFAC Model. *Biotechnol. Prog.* **1996**, *12* [3], 371–379.
- [24] Ferreira, L. A.; Macedo, E. A.; Pinho, S. P. Solubility of Amino Acids and Diglycine in Aqueous–alkanol Solutions. *Chem. Eng. Sci.* **2004**, *59* [15], 3117–3124.
- [25] Ferreira, L. A.; Pinho, S. P.; Macedo, E. A. Solubility of L-Serine, L-Threonine and L-Isoleucine in Aqueous Aliphatic Alcohol Solutions. *Fluid Phase Equilibria* **2008**, *270* [1–2], 1–9.
- [26] Nozaki, Y.; Tanford, C. The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions. *J. Biol. Chem.* **1971**, *246* [7], 2211–2217.

- [27] Lalasidis, G. Four New Methods of Debittering Protein Hydrolysates and a Fraction of Hydrolysates with High Content of Essential Amino Acids. *Ann. Nutr. Aliment.* **1978**, 32 [2–3], 709–723.
- [28] Elbahloul, Y.; Frey, K.; Sanders, J. P. M.; Steinbuchel, A. Protamylasse, a Residual Compound of Industrial Starch Production, Provides a Suitable Medium for Large-Scale Cyanophycin Production. *Appl. Environ. Microbiol.* **2005**, 71 [12], 7759–7767.
- [29] Widyarani; Ratnaningsih, E.; Sanders, J. P. M.; Bruins, M. E. Biorefinery Methods for Separation of Protein and Oil Fractions from Rubber Seed Kernel. *Ind. Crops Prod.* **2014**, 62, 323–332.
- [30] Zhu, Y.; Xu, J.; Li, Q.; Mortimer, P. E. Investigation of Rubber Seed Yield in Xishuangbanna and Estimation of Rubber Seed Oil Based Biodiesel Potential in Southeast Asia. *Energy* **2014**, 69, 837–842.
- [31] Sanders, J. P. M. Grassa BV, Friesland. Personal Communications, 2016.
- [32] Cornelissen, J. M. R.; De Haan, M. H. A.; Hin, C. J. A.; Zijerveld, E. J. M.; Philipsen, A. P. Exploitable yield potential of grasslands in the Netherlands <http://www.nvww.nl/en/document/exploitable-yield-potential-grasslands-netherlands> [accessed Jun 5, 2016].
- [33] Meussen, B. J.; van Zeeland, A. N. T.; Bruins, M. E.; Sanders, J. P. M. A Fast and Accurate UPLC Method for Analysis of Proteinogenic Amino Acids. *Food Anal. Methods* **2014**, 7 [5], 1047–1055.
- [34] Schaffenberger, M.; Ecker, J.; Koschuh, W.; Essl, R.; Mandl, M. G.; Boechzelt, H. G.; Steinmueller, H.; Schnitzer, H. Green Biorefinery Production of Amino Acids from Grass Silage Juice Using an Ion Exchanger Device at Pilot Scale. *Chem. Eng. Trans.* **2012**, 29, 505–510.
- [35] Haynes, W. M. *CRC Handbook of Chemistry and Physics, 92nd Edition*; CRC Press, 2011.
- [36] McMeekin, T. L.; Cohn, E. J.; Weare, J. H. Studies in the Physical Chemistry of Amino Acids, Peptides and Related Substances. III. The Solubility of Derivatives of the Amino Acids in Alcohol-Water Mixtures. *J. Am. Chem. Soc.* **1935**, 57 [4], 626–633.
- [37] Dunn, M. S.; Ross, F. J. Quantitative Investigations of Amino Acids and Peptides IV. The Solubilities of the Amino Acids in Water-Ethyl Alcohol Mixtures. *J. Biol. Chem.* **1938**, 125 [1], 309–332.

- [38] Zhang, C.; Liu, B.; Wang, X.; Wang, H. Measurement and Correlation of the Solubilities of L-Valine in Water, Ethanol, *N,N*-Dimethylformamide, Acetone, and Isopropyl Alcohol between [293.15 and 343.15] K. *J. Chem. Eng. Data* **2014**, *59* [9], 2704–2708.
- [39] Zhang, C.; Liu, B.; Wang, X.; Wang, H.; Zhang, H. Measurement and Correlation of Solubility of L-Valine in Water + [Ethanol, *N,N*-Dimethylformamide, Acetone, Isopropyl Alcohol] from 293.15 K to 343.15 K. *J. Chem. Eng. Data* **2014**, *59* [9], 2732–2740.
- [40] Belitz, H.-D.; Grosch, W.; Schieberle, P. Amino Acids, Peptides, Proteins. In *Food Chemistry*; Springer Berlin Heidelberg, 2009; pp 8–92.
- [41] Bull, H. B.; Breese, K. Interaction of Alcohols with Proteins. *Biopolymers* **1978**, *17* [9], 2121–2131.
- [42] Moore, B. D.; Deere, J.; Edrada-Ebel, R.; Ingram, A.; van der Walle, C. F. Isolation of Recombinant Proteins from Culture Broth by Co-Precipitation with an Amino Acid Carrier to Form Stable Dry Powders. *Biotechnol. Bioeng.* **2010**, *106* [5], 764–773.
- [43] Daneshfar, A.; Baghlani, M.; Sarabi, R. S.; Sahraei, R.; Abassi, S.; Kaviyan, H.; Khezeli, T. Solubility of Citric, Malonic, and Malic Acids in Different Solvents from 303.2 to 333.2 K. *Fluid Phase Equilibria* **2012**, *313*, 11–15.
- [44] Oliveira, M. L. N.; Malagoni, R. A.; Franco Jr., M. R. Solubility of Citric Acid in Water, Ethanol, *N*-Propanol and in Mixtures of Ethanol + Water. *Fluid Phase Equilibria* **2013**, *352*, 110–113.
- [45] Yuan, Y.; Leng, Y.; Shao, H.; Huang, C.; Shan, K. Solubility of DL-Malic Acid in Water, Ethanol and in Mixtures of Ethanol + Water. *Fluid Phase Equilibria* **2014**, *377*, 27–32.
- [46] van Koningsveld, G. A.; Gruppen, H.; de Jongh, H. H. J.; Wijngaards, G.; van Boekel, M. A. J. S.; Walstra, P.; Voragen, A. G. J. The Solubility of Potato Proteins from Industrial Potato Fruit Juice as Influenced by pH and Various Additives. *J. Sci. Food Agric.* **2002**, *82* [1], 134–142.
- [47] Khoshkbarchi, M. K.; Vera, J. H. Effect of NaCl and KCl on the Solubility of Amino Acids in Aqueous Solutions at 298.2 K: Measurements and Modeling. *Ind. Eng. Chem. Res.* **1997**, *36* [6], 2445–2451.
- [48] El-Dossoki, F. I. Effect of the Charge and the Nature of Both Cations and Anions on the Solubility of Zwitterionic Amino Acids, Measurements and Modeling. *J. Solut. Chem.* **2010**, *39* [9], 1311–1326.

- [49] El-Dossoki, F. I.; El-Damarany, M. M. Solvation of Basic and Neutral Amino Acids in Aqueous Electrolytic Solutions: Measurements and Modeling. *J. Chem. Eng. Data* **2015**, *60* [10], 2989–2999.
- [50] Yokoyama, C.; Terui, M.; Takahashi, S. Salt Effect on Partition Coefficient of Glycine, L-Valine, and L-Phenylalanine in N-Octanol-Water System. *Fluid Phase Equilibria* **1993**, *82*, 283–290.

Chapter 8

General Discussion

8.1 Introduction

The objective of the research presented in this thesis was to understand the physical and thermodynamic properties of α -amino acids and their intermolecular interactions to each other and to solvents and co-solvents. This would aid us in separating α -amino acids that have been extracted from agricultural residue from solution. The α -amino can subsequently be processed for products (pharmaceuticals, fine chemicals, food/feed, *etc.*). In order to reach these objectives, three approaches have been used. These were:

1. Measure the impact of temperature and co-solvents on the solubility of α -amino acids in a model mixture of all 20 α -amino acids and individual α -amino acids.
2. Apply existing solubility models and propose a new solubility model to the solubility data on α -amino acid solubility in both the literature and measured for this thesis from approach 1.
3. Develop a process for extracting and isolating proteins and amino acids from a few agricultural residues

The first research approach was applied in chapters 2, 3 and 4. The second research approach was applied in chapters 5 and 6. The models developed and applied in chapters 5 and 6 used either single amino acids or mixtures of amino acids, but not actual biomass residues. The third research approach was applied in chapter 7 and in this chapter, where a methodology was developed to study the extraction and isolation of amino acids and proteins from biomass.

8.2 Summary of previous chapters

It is crucially important to understand the solubility of a solute (e.g. amino acids) in order to separate them out of solution by crystallization. In the past, solubility data was incomplete for many of the amino acids, making designing processes for amino acid crystallization difficult. In this thesis, data on amino acid solubility and models for understanding and predicting amino acid solubility for crystallization has been filled in.

As shown in previous chapters in this thesis, amino acids can be grouped by the properties of their side chains. The physical structure of the amino acid determines their solubility and the physical interaction with other amino acids. In section 8.2, an overview of the solubility of the amino acids in mixtures, the effects of ethanol on single amino acids and

the effects of ethanol on amino acids in mixtures is given. These amino acids will be grouped by their side chains structured as described in the Venn diagram showed in Chapter 1 and in Figure 8.1.

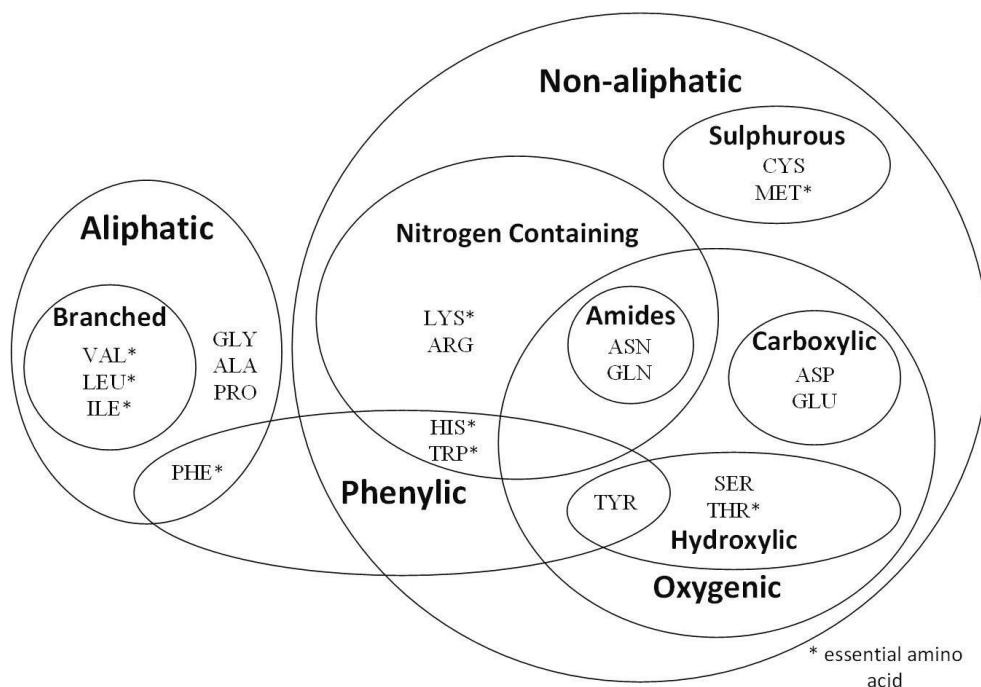


Figure 8.1: Venn diagram of the proteinogenic amino acids grouped by the characteristics of their functional groups

8.2.1 Venn Diagram Mixtures

In Chapter 4, the effects on the solubility of an individual amino acid in water when in a mixture of 20 amino acids were analyzed. The results are summarized in a Venn diagram in Figure 8.2. In this figure, blue denotes that the amino acids in this group increased in solubility compared to being in single solution. Green denotes that the amino acids in this group were only slightly less soluble than in single solution. Yellow denotes that the amino acids in this group were less soluble in a mixture than in single solution. Large decreases in solubility (would have been red in Venn diagram) were not observed under these conditions.

The amino acids that showed the most decrease in a mixture of amino acids were amino acids with phenylic side chains. Of these amino acids with side chains that have phenyl groups, L-phenylalanine shows the least decrease in solubility. L-Phenylalanine is also an

aliphatic amino acid, containing only carbon and hydrogen in the side chain. Being both phenylic and aliphatic could signal a mitigation in the decrease in solubility for L-phenylalanine.

All other amino acids, showed an increase in solubility when in a mixture of amino acids. These groups include all nitrogen containing amino acids with the exception of L-tryptophan, both carboxylic amino acids and sulphur containing amino acids.

These results suggest that non-aliphatic non-phenylic amino acids and to a lesser extent aliphatic non-phenylic amino acids could be separated from solution through crystallization with the addition of another amino acid or amino acids to the solution. This would be the case when the phenylic amino acid is in single solution or in solution with few other amino acids.

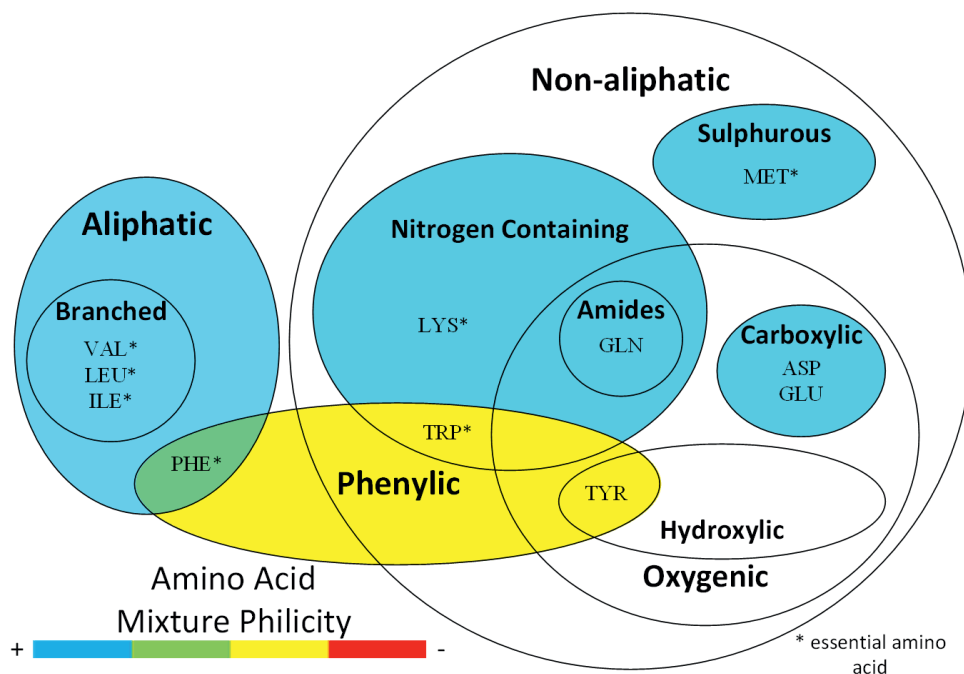


Figure 8.2: Venn diagram of the effects on the solubility of amino acids when in an aqueous mixture of amino acids compared to being in solution by themselves. Blue indicates increased solubility when in a group. Green indicates slightly decreased solubility when in a group. Yellow indicates a moderately decreased solubility when in a group. Red indicates decreased solubility when in a group.

8.2.2 Venn Diagram Ethanol

As discussed in Chapter 3, the effects of ethanol on the solubility of amino acids in water can be classified by the properties of the side chains of the amino acids. In Figure 8.3,

the amino acids have been grouped by their physical properties. The colors in this diagram are assigned in the same manner as in Figure 8.3 with blue showing an increase and green and yellow being a slight decrease and a moderate decrease in solubility respectively, but red denotes a greater decrease in solubility.

Only the phenylic amino acids show an increase in solubility when ethanol is introduced. This group is shown in blue and green. However, above 60% (g/g), even the solubility of these amino acids decreased. L-Phenylalanine, also being aliphatic, was the only phenylic amino acids to not increase in solubility at any concentration of ethanol.

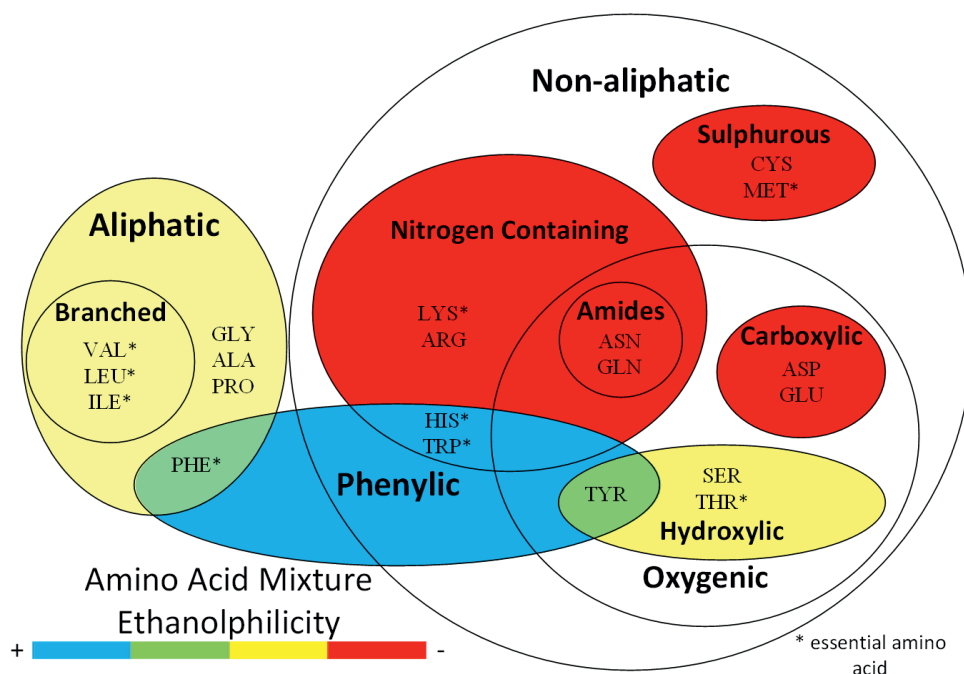


Figure 8.3: Venn diagram of the effects on the solubility of amino acids when in aqueous solution compared to being in a solution with ethanol. Blue indicates ethanolphilicity. Green indicates slightly decreased solubility when in a group. Yellow indicates a moderately decreased solubility when in a group. Red indicates ethanolphobicity.

The aliphatic and hydroxylic amino acids, shown in yellow, show a medium decrease in solubility when ethanol is added. L-Tyrosine showed the least decrease when ethanol is added among the hydroxylic group.

The amino acids that have side chains that include amides/nitrogen, sulphur and an additional carboxylic acid group show large decreases in solubility when ethanol is added. All of these amino acids showed increases when in a mixture.

These results suggest that the phenylic amino acids and to a lesser extent, the aliphatic and hydroxyllic amino acids are less sensitive to the addition of ethanol. Therefore, another solvent should be investigated as an anti-solvent when they are in single solution.

8.2.3 Venn Diagram Mixtures and Ethanol

The effects of ethanol on the solubility of amino acids in mixtures was discussed in Chapter 4. In Figure 8.4, these effects are shown. The colors in this diagram are assigned in the same manner as in Figure 8.3 with blue showing an increase and green and yellow being a slight decrease and a decrease in solubility respectively and red denoting a greater decrease in solubility.

The solubility of ethanol on mixtures of amino acids is similar to the effect of ethanol on single amino acids. Phenylic amino acids have a marked decrease in solubility.

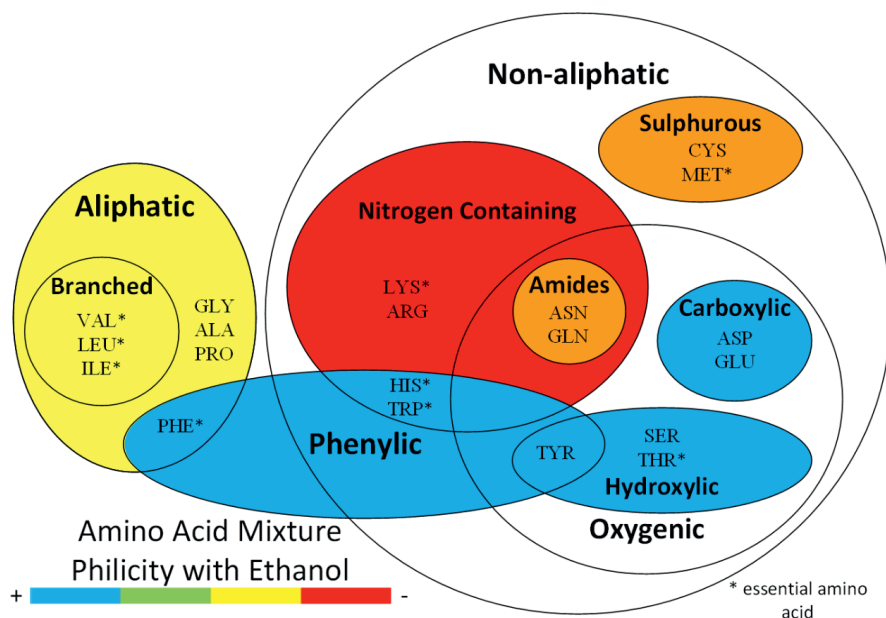


Figure 8.4: Venn diagram of the effects of ethanol on the solubility of amino acids in aqueous solution when ethanol is added to mixtures of amino acids. Blue indicates ethanolphilicity when in a group. Green indicates slight ethanolphilicity when in a group. Yellow indicates slight ethanolphobicity when in a group. Red indicates ethanolphobicity when in a group.

However, there are a few differences. Unlike when in single solution, the amino acids that contain carboxylic acid or a hydroxylic acid in their side chain show an increase in

solubility when ethanol was added to a mixture instead of a decrease when present as a single amino acid.

The amino acids with sulphur and amides in the side chain showed less decrease than when in single solution alone. The aliphatic amino acids only showed a slight decrease in solubility when ethanol was added to a mixture of amino acids. This is similar to the decrease in single solution when ethanol was added.

These results suggest that nitrogen containing and sulphurous amino acids could be separated from solution when they are in a mixture with the addition of ethanol. The other amino acids would proportionally become less insoluble and would therefore precipitate proportionally less when ethanol is added.

8.3 Discussion

The results presented in this thesis are specific to the solubility behavior of amino acids in single solution and in mixtures. Furthermore, the effect of temperature has also been explored and discussed. However, there are other specific circumstances that have not been addressed, that have impact on amino acid solubility. Important influences include the pH of the system, the interactions between individual amino acids and the use of two solvents that do not mix completely.

8.3.1 Influence of pH

The pH of the system will affect the solubility of an amino acid. All amino acids have at least two charged groups: the amine group ($pK_a = 2$) and carboxylic acid group ($pK_a = 9$) attached to the alpha carbon. The point at which there is no charge, and the zwitterion is present, is called the isoelectric point. This means that if the pH of the system is under the isoelectric point of that amino acid, that amino acid will be protonated and positively charged. At a pH higher than the isoelectric point of that amino acid, that amino acid will be deprotonated and therefore negatively charged. For most amino acids this isoelectric point is approximately pH 5. However, some amino acids have a charged side chain. The amino acids with acidic side chains are L-aspartic acid ($pK_a = 4$) and L-glutamic acid ($pK_a = 4$). The amino acids with positively charged side chains are L-arginine ($pK_a = 12$), L-histidine ($pK_a = 6$), L-lysine ($pK_a = 11$), L-cysteine ($pK_a = 8$) and L-tyrosine ($pK_a = 10$) [1].

The further away that the system is from the isoelectric point of the amino acid, the higher the solubility that amino acid will be. This is an exponential equation and the solubility of the amino acid in relation to pH can be calculated by:

$$x_A = \frac{K_x}{\gamma_{A^\pm}^*} \left[1 + \frac{[H^+]}{K_{1c}} + \frac{K_{2c}}{[H^+]} + \frac{K_{2c}K_{Rc}}{[H^+]^2} \right] \quad (8.1)$$

where x_A is the amino acid mole fraction, $\gamma_{A^\pm}^*$ is the activity coefficient of the amino acid zwitterions. K_x is the solubility equilibrium constant, K_{1c} , K_{2c} and K_{Rc} are the equilibrium constants of the dissociation reactions of the amine group (1), carboxyl group (2) and the side chain (R) [2, 3].

In single solution, when the amino acid is at saturation, the pH of the system is at the isoelectric point. If an acid, base or charged molecule (e.g. another amino acid with an isoelectric point above or below the isoelectric point of the saturated amino acid) were added to this system, then the pH of the system would change and the solubility of the saturated amino acid would increase due to this pH change and would therefore no longer be saturated. Conversely, if an amino acid were saturated in a system where the pH was above or below the saturation point and the pH was brought closer to the isoelectric point of that amino acid, then the solubility of the amino acid would be lowered and the amino acid would be oversaturated and begin to precipitate and crystallize [4].

In mixed solutions of amino acids, the pH of the system is highly buffered around 5 by the amino acids themselves. Theoretically, the addition or loss of small amounts of amino acids, or all of a single amino acid, from this system would probably not cause a shift in pH where the solubility of the other amino acids would be greatly affected. Furthermore, the research in this thesis has shown that amino acids in mixture raise the solubility of the amino acids themselves, increasing the buffering of the system. However, this is theoretical conjecture and should be taken into account in future research and industrial processes. This is especially the case when one amino acid is dominant in the system, or the system includes only a few amino acids, like in the next section.

8.3.2 Binary interactions

In this research, amino acids have been studied in single solution and in mixtures of all amino acids. However, in many systems, one amino acid may be in much higher relation to the other amino acids (e.g. L-asparagine in protamylasse, chapter 7), or only a few amino acids are present. In order to understand the effects that individual amino acids have on each other, an examination of two amino acids in water was carried out according to protocols in the literature [5]. Careful consideration has been given in these binary amino acid interaction

experiments to the pH of the system, since the amino acids chosen for this represent groups with disparate isoelectric points. The amino acids represent the following groups: positively charged, negatively charged, phenylic and nitrogen containing.

The binary solutions of amino acids are made by bringing L-aspartic acid to saturation by adding an excess of crystal L-aspartic acid to water. After the L-aspartic acid comes to equilibrium, we took a dilution series (0%, 25%, 50%, 75% and 100%) of the saturated L-aspartic acid solution to a Greiner tube. An excess of solid second amino acid (L-tyrosine, L-tryptophan, L-glutamic acid, L-histidine and L-glutamine) was added to the Greiner tubes of the dilution series of L-aspartic acid. The solutions Greiner tubes were sealed and put in a thermomixer set at 25 °C for a 48-hour period. The pH of these solutions was measured.

Aliquots of the solutions were filtered, weighed and dried to find the amount of total dissolved amino acids in solution. Another aliquot of the solution was filtered and diluted and analyzed by HPLC to find the concentration of both amino acids that were dissolved. The crystals that were formed were dried and then dissolved in water and then the molar concentrations were measured by HPLC.

The solubility of amino acids (L-tyrosine, L-tryptophan, L-glutamic acid, L-aspartic acid, L-histidine, L-glutamine *etc.*) increased in solution when other amino acids were present. However, looking at mixtures of 20 amino acids does not answer the question of which intermolecular interactions between which amino acids (binary interaction) is causing this increase in solubility. In order to begin investigating this, we measured the interactions that these 6 amino acids have on each other in binary solution in reference to their solubility. We used L-aspartic acid as the second amino acid. We brought L-aspartic acid to maximum solubility, then took aliquots of this at various concentrations. To all concentrations, including 100% solubility, we added the other 5 amino acids individually (L-tyrosine, L-tryptophan, L-glutamic acid, L-histidine and L-glutamine) in duplicate in excess. We then measured the pH and the concentration of both amino acids in all samples.

The results of the experiments are shown in Figure 8.5-Figure 8.9. The solubilities of the amino acids are shown as a function of pH. The point at the highest pH value is the 0% L-aspartic acid solution. As the L-aspartic concentration of the system increases, the pH of the solution decreases. Therefore, in Figure 8.5-Figure 8.8, as the concentration of L-aspartic acid increases, the concentration of the other amino acid also increases. The calculated solubility of the amino acid without L-aspartic acid at any given pH has been calculated with equation (8.1) and is shown as a solid line in the figures. Furthermore, the pH scale of the figures has been kept to approximately 2 in order to aid comparison between the amino acids. For L-

tyrosine and L-histidine, the 0% L-aspartic acid is not shown, because the scale of the axis does not allow for a clear view of the other four points. The exception is Figure 8.9, since both amino acids, L-aspartic acid and L-glutamic acid, have similar isoelectric points and the variation in pH in all of the results is low.

In Figure 8.5, L-tyrosine has a similar measured solubility when compared to the modeled solubility. The measured solubility at the highest concentration of L-aspartic acid has a variance that includes the modeled solubility.

In Figure 8.6, L-glutamine shows a slightly higher solubility at all concentrations of L-aspartic acid. The solubility increase was measured to be between 5% and 8% higher than the modeled solubility.

In Figure 8.7, L-histidine shows an increase in solubility as the concentration of L-aspartic acid increases. The increase in solubility increased as the concentration of the L-aspartic acid increased. At the highest concentration of L-aspartic acid, the solubility increase was measured to be approximately 10%.

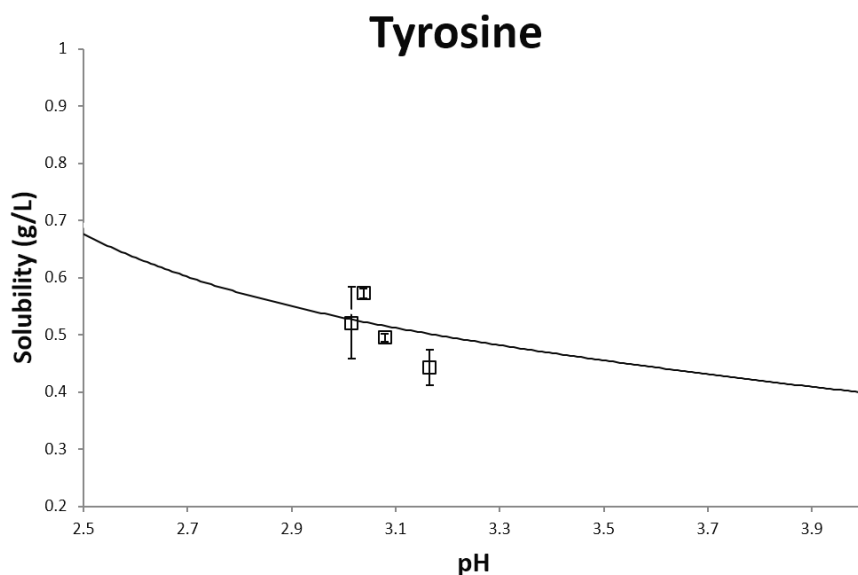


Figure 8.5: Solubility of L-tyrosine in g/L at various concentrations of L-aspartic acid (squares) and the predicted solubility of L-tyrosine in single solution (solid line) using equation 8.1 in relation to the pH.

Glutamine

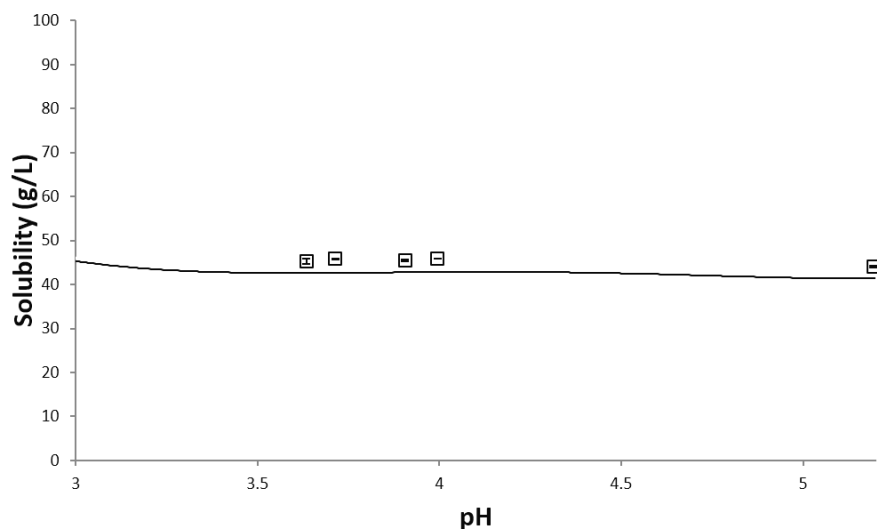


Figure 8.6: Solubility of L-glutamine in g/L at various concentrations of L-aspartic acid (squares) and the predicted solubility of L-glutamine in single solution (solid line) using equation 8.1 as a function of pH.

Histidine

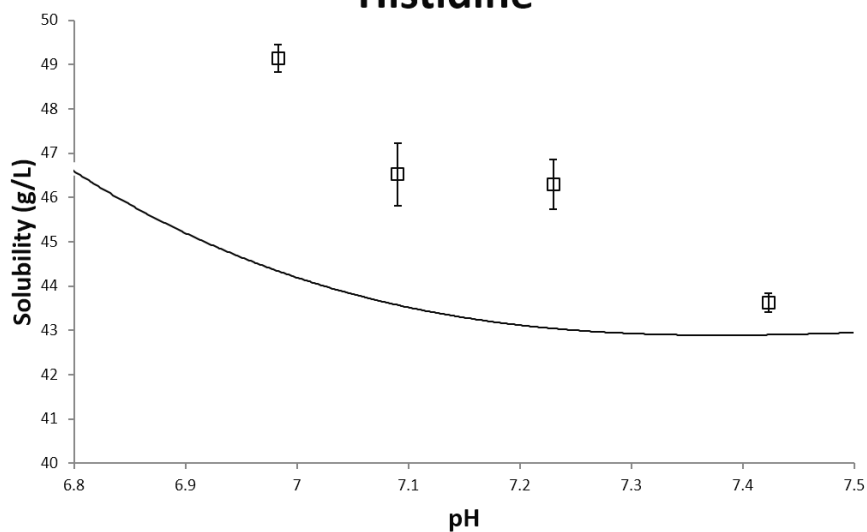


Figure 8.7: Solubility of L-histidine in g/L at various concentrations of L-aspartic acid (squares) and the predicted solubility of L-histidine in single solution (solid line) using equation 8.1 as a function of pH.

Tryptophan

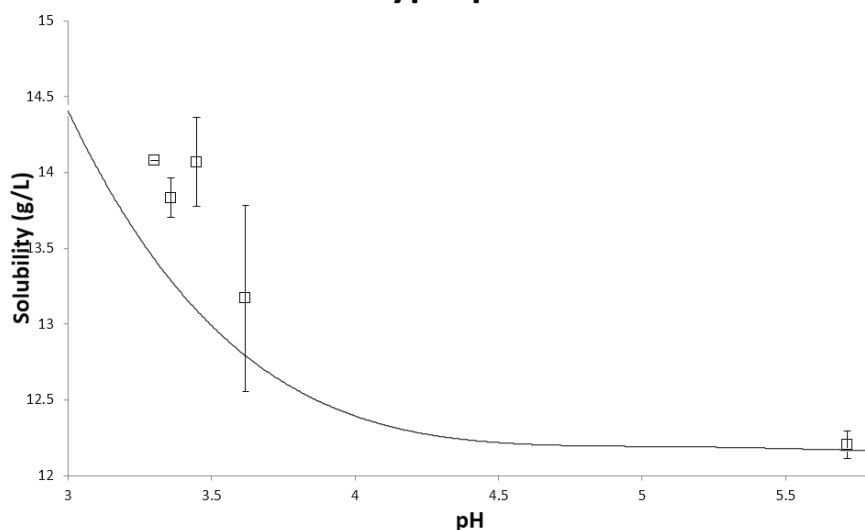


Figure 8.8: Solubility of L-tryptophan in g/L at various concentrations of L-aspartic acid (squares) and the predicted solubility of L-tryptophan in single solution (solid line) using equation 8.1 as a function of pH.

Glutamic acid

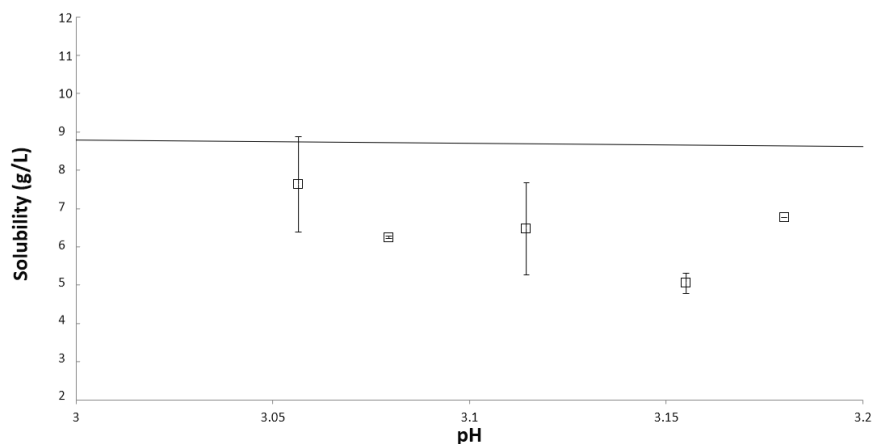


Figure 8.9: Solubility of L-glutamic acid in g/L at various concentrations of L-aspartic acid (squares) and the predicted solubility of L-glutamic acid in single solution (solid line) using equation 8.1 as a function of pH.

In Figure 8.8, L-tryptophan shows an increase in solubility as the concentration of L-aspartic acid increases. The increase in solubility remained approximately constant as the

concentration of the L-aspartic acid increased. The solubility increase was measured to be between 2% and 8% higher than the modeled solubility.

In Figure 8.9, L-glutamic acid shows a decrease in solubility at all concentrations of L-aspartic acid. The decrease in solubility was measured to be between -15% and -40%.

The results suggest that for L-tryptophan, L-glutamine and L-histidine, a decrease in pH is not the only factor that increases the solubility of the amino acid. Furthermore, L-tryptophan showed an increase in solubility in the model mixture of 20 amino acids (chapter 2). The conclusion could be that the interaction between the L-aspartic acid and the second amino acid (e.g. L-tryptophan) causes the increase in solubility of the second amino acid. This suggests that as the pH decreases and the L-aspartic acid becomes uncharged and zwitterionic that it is able to interact with the increasingly charged L-tryptophan by, for example, decreasing the distance between the amino acids. However, L-glutamine showed a decrease in solubility in a model mixture of 20 amino acids. The conclusion could be that the interaction between the L-aspartic acid and the L-glutamine causes increase in solubility of the second amino acid, however, another amino acid in the mixture is causing a decrease in the solubility of L-glutamine.

The results for L-tyrosine show that L-aspartic acid does not show a measured increase or decrease in the solubility of L-tyrosine. In the model mixture of 20 amino acids, however, L-tyrosine showed an increase in solubility. The conclusion could be that an amino acid other than aspartic acid causes the increase in solubility for L-tyrosine in the model mixture of 20 amino acids.

The results for L-glutamic acid suggest that L-aspartic acid decreases the solubility of L-glutamic acid. In the model mixture of 20 amino acids, the solubility of L-glutamic acid is increased. The conclusion could be that an amino acid or acids are the cause for the increase of solubility in L-glutamic acid in the model mixture of 20 amino acids.

8.3.3 Two-liquid system

Another technique that can be used to separate amino acids, or any solute, is using a two-liquid system [6]. In this system, two liquids that do not mix are used to divide the amino acids into different fractions. An example of this liquid-liquid extraction is if two amino acids were dissolved in water and then a non-miscible liquid were added to the solution and the solution mixed, then when the system were allowed to settle, a portion of either or both of the amino acids would migrate to the non-miscible liquid. In a system where two amino acids with dissimilar fractionation coefficients to the non-miscible liquid, then one amino acid

would be found in high concentration in the one solvent and the other in high concentration in the other solvent.

It has been shown that ethanol and water can be separated if charged ions are present [7]. This is known as “salting out” ethanol from water. This can also be achieved with amino acids, if the pH of the system is increased above the isoelectric point of the amino acids and at temperatures around 80° C. In this way the solubility of the amino acids is highly increased and the amino acids are also charged.

12.0 grams of each of the 20 proteinogenic amino acids were added to 5 1.0 L Schott bottles. 200.0 ml of a solution of 50% ethanol and 50% water (g/g) were added to each Schott bottle and each were brought to pH levels 4.0, 5.5, 5.5 duplicate, 9.0 and 10.0. The pH was measured with a Sigma-Aldrich pH strip with a resolution of 0.5 pH unit. The Schott bottles were placed in a shaking incubator set to 333.15 K. The pressures in the Schott bottles were equilibrated by quickly opening and resealing the bottles to maintain a pressure of 0.1 MPa 3 times while the solution warmed up. Once the bottles were at 333.15 K, they were left in the shaking incubator for an additional 2 hours. The bottles were taken out of the incubator, and before cooling, two 10.0 ml samples were taken via syringe from each bottle. The sealed syringes were allowed to cool off and then put into a centrifuge at 2500 RPM for 15 minutes. The results of the experiment are shown in the syringe are showed in Figure 8.10.

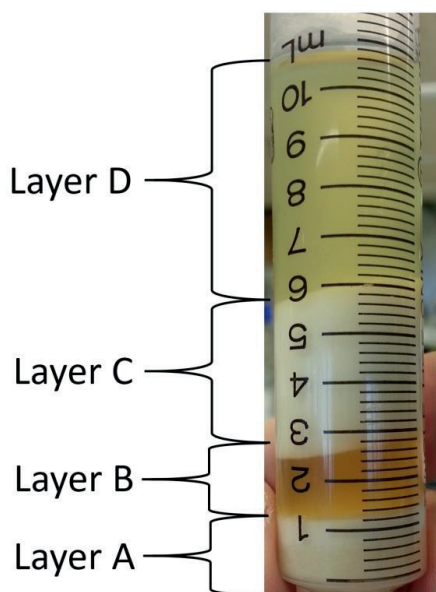


Figure 8.10: Photograph of the layers formed in the two-liquid system. Layers D and C are mostly liquid and layers C and A are mixtures of crystal amino acids.

The liquid layers B and D, shown in Figure 8.10, were analyzed using the HPLC technique (described in chapters 2, 3 and 4) to determine the concentration of amino acids. The liquids were also analyzed using a GC to measure the concentration of ethanol by using the head space technique. The crystal layers are A and C in Figure 8.10, The concentration of ethanol in the liquid layers is shown in Figure 8.11. The concentration of the 20 amino acids in the liquid layers is shown in Figure 8.12 and Figure 8.13.

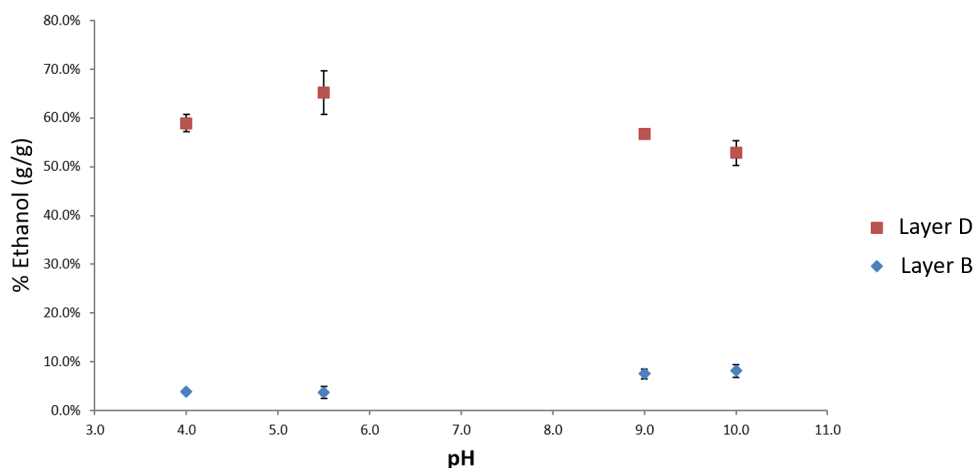


Figure 8.11: Ethanol concentration in % g ethanol /g solvent in the two liquid layers as a function of pH. The top liquid layer, layer D, is shown in red squares, the bottom liquid layer, layer B, is shown in blue diamonds.

Layer D had an average volume of 4.5 ml in the 10 ml samples. Layer C had an average volume of 1.5 ml in the 10 ml samples. Layer B was viscous and had a light brown color. The average ethanol concentration at pH 4.0 was 58% (g/g) in layer D and 4% (g/g) in Layer B. The average ethanol concentration at pH 5.5 was 65% (g/g) in layer D and (g/g) 4% in layer B. The average ethanol concentration at pH 9.0 was 56% (g/g) in layer D and 6% (g/g) in layer B. The average ethanol concentration at pH 10.0 was 53% (g/g) in layer D and 7% (g/g) in layer B.

In Figure 8.12, we see the concentration of the amino acids in layer D at the different pH levels. We see that all amino acids are at maximum solubility, because 60.0 g/L were added initially to the Schott bottles and no sample reached that concentration. Therefore, all amino acids were at maximum solubility. The amino acids all had their lowest concentrations

at pH 5.5, with the exception of the acid amino acids, that had their lowest concentration at pH 4. The highest concentration was at pH 10 for all amino acids.

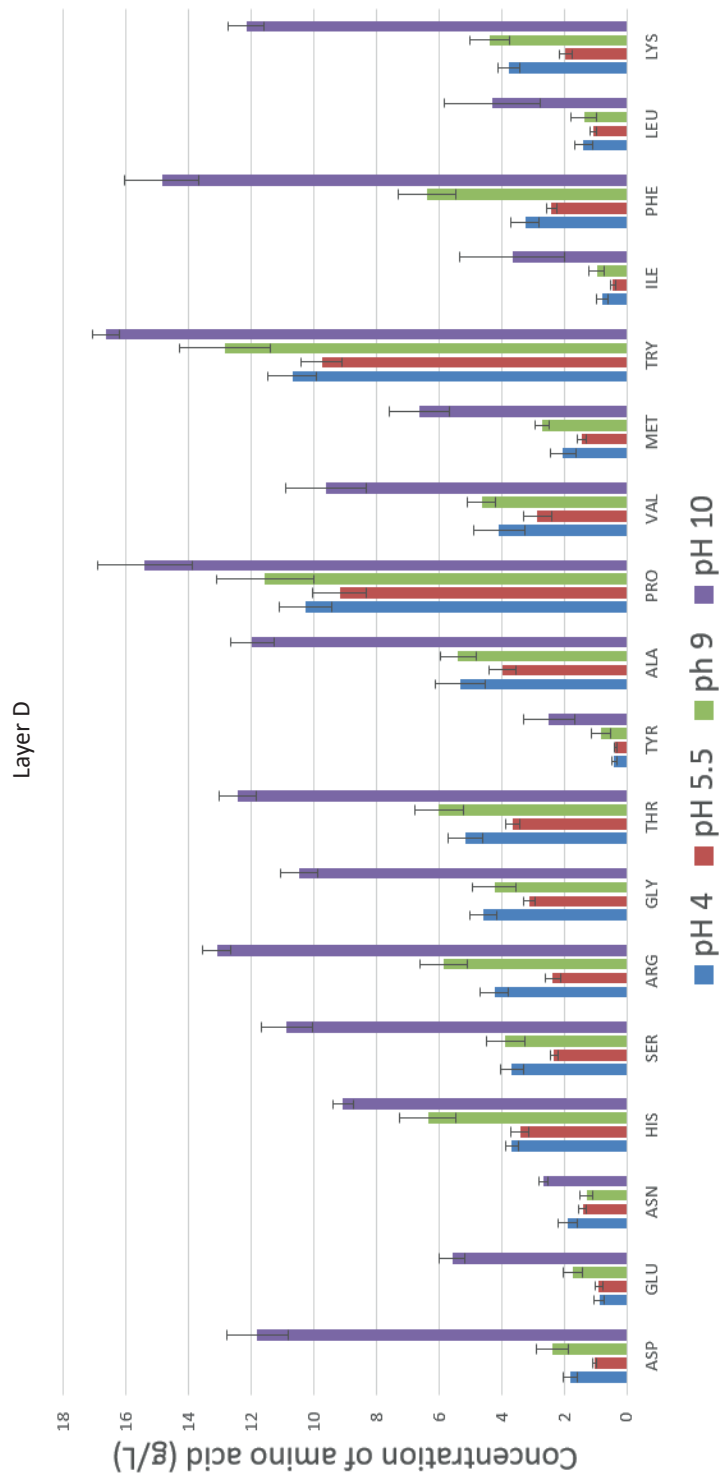


Figure 8.12: The concentration of amino acids in mixed solution in layer D at pH 4.0 (blue), 5.5 (red), 9.0 (green) and 10 (purple)

In Figure 8.13, the concentration of the amino acids in layer B is shown. Again, all of the amino acids are at maximum solubility, since all of them have concentrations under 60.0 g/L. Contrary to layer D, the highest concentrations for most of the amino acids were at pH 4.0 or 5.5, with the exception of L-glutamic acid. This is remarkable, because amino acids show their lowest solubilities at their isoelectric points. Since most amino acids have an isoelectric point around 5.0, we would expect the lowest concentrations to be between pH 4.0 and 5.5.

In all samples, the concentration of ethanol in layer D was higher than in the original solution. The samples at pH 5.5 had the highest concentration of ethanol and pH 10 had the lowest ethanol concentration. This is in contrast to the amino acid concentrations, which had their lowest concentrations at pH 5.5.

In all samples, the concentration of ethanol in layer B was lower than in the original solution. All samples had ethanol concentrations under 8% ethanol (g/g). The low concentrations of ethanol coincide with high concentrations of amino acid.

The concentration of every amino acid was higher in layer B, the predominantly water layer, than in layer D, the predominantly ethanol layer. Furthermore, the concentrations were the highest around the isoelectric points of the amino acids. This means that the zwitterion of these amino acids would be dominant. It is possible that this zwitterion is acting as a salt to separate the liquid into two phases.

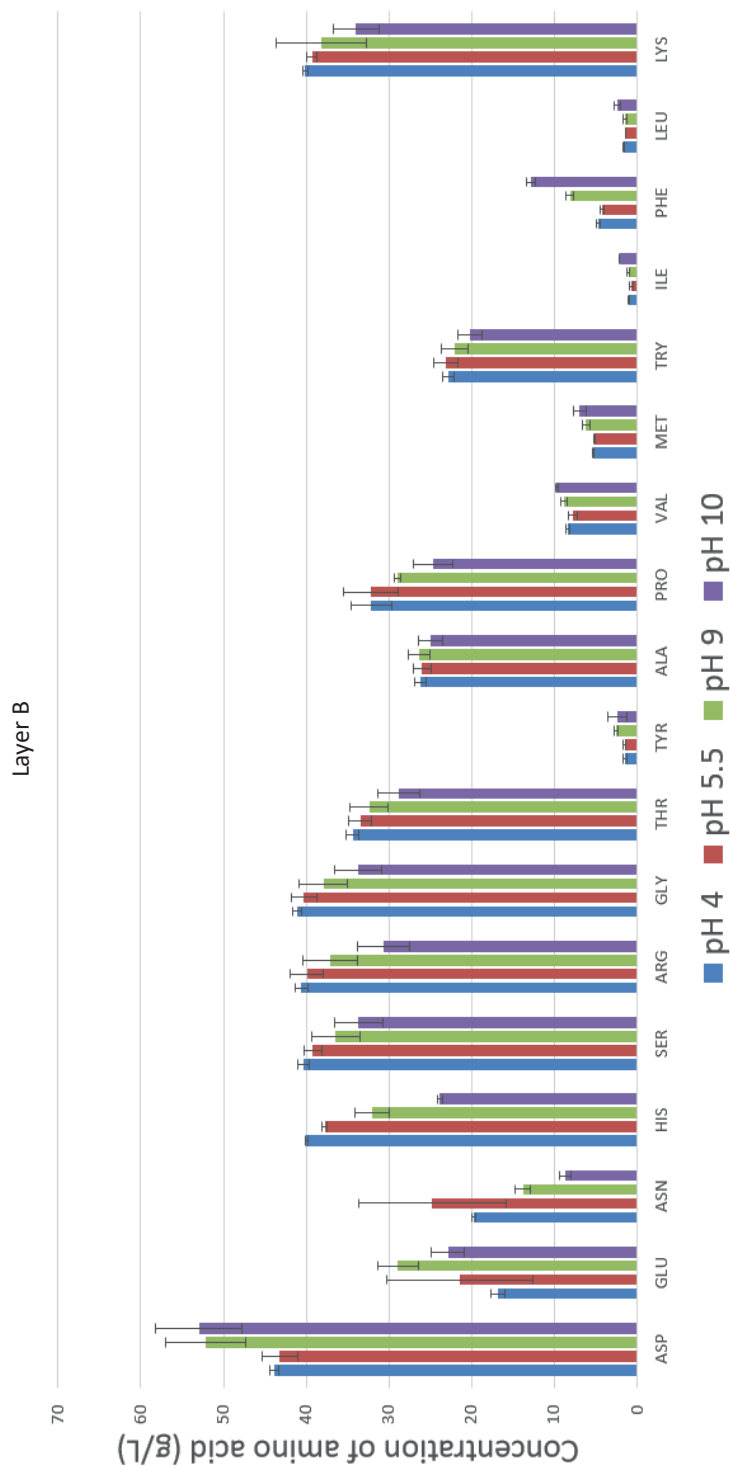


Figure 8.13: The concentration of amino acids in mixed solution in layer B at pH 4.0 (blue), 5.5 (red), 9.0 (green) and 10 (purple)

As discussed in chapter 2, the interactions between the amino acids is what is elevating their solubility in both the predominantly water layer and the predominantly ethanol layer. Furthermore, the effect of the ethanol, as discussed in chapter 4, is lower than if the amino acids were in individual solutions of ethanol-water and not in a mixture of 20 amino acids.

In conclusion, the amino acids that were separated into the ethanol dominant liquid layer, layer D, were the phenylic and branched aliphatic amino acids. These amino acids are specifically, L-tyrosine, L-valine, L-methionine, L-isoleucine, L-phenylalanine and L-leucine. This was especially the case when the pH was furthest away from the pI. The amino acids that were separated into the water dominant layer, layer A, were all other amino acids.

8.4 Integration towards increased selectivity

The knowledge of amino acid selectivity individually, their reaction to anti-solvents individually and in mixtures and how to predict this is an important step in a bio-refinery for amino acids. However, there are other steps before the amino acid separation that increases the integration of amino acid selectivity in a bio-refinery.

One way to increase the selectivity in a bio-refinery, would be to choose the correct biomass for the amino acid that you want. Not all biomasses contain the same amino acid profile [8]. Choosing a biomass with free amino acids and proteins with a predominant amino acid that you are looking to separate, aids the bio-refinery.

A second way of preselecting an amino acid before it goes into solution would be to selectively hydrolyze for a specific amino acid, or a group of amino acids from a protein. In this first fraction, the selected amino acids would be in high concentration facilitating their separation. Furthermore, the successive fractions would have less of the amino acid or acids selectively hydrolyzed and therefore would also facilitate their further separation.

A third way of increasing the selectivity of amino acid recovery would be to convert an amino acid or groups of amino acids, that is a precursor for the production of a chemical, to a form that no longer has the same properties [9]. In this way, they would either be less soluble and precipitate out of solution or make the dissolution of other amino acids easier either directly (e.g. bringing down their solubility due to their absence) or indirectly (e.g. altering the pH of the solution).

8.4.1 Biomass choice

Another way to recover amino acids from proteins is to hydrolyze a protein rich agriculture residue stream. These protein-rich agricultural residues are a source of the amino acids that will be separated.

Once the protein has been extracted, the proteins would be hydrolyzed by cleaving the bonds between the amino acids in the proteins. In this way, the proteins and peptides are broken up into the individual amino acids.

The hydrolysate of several biomasses have been analyzed for their amino acid concentration [8]. The biomasses represent residues from agricultural industry such as dried distillers grains and solids (DDGS), Vinasse (a by-product of the fermentation industry) and press-cake meals left over after the pressing of oils (e.g. palm) and juices (e.g. soy milk). This analysis is given in Figure 8.14.

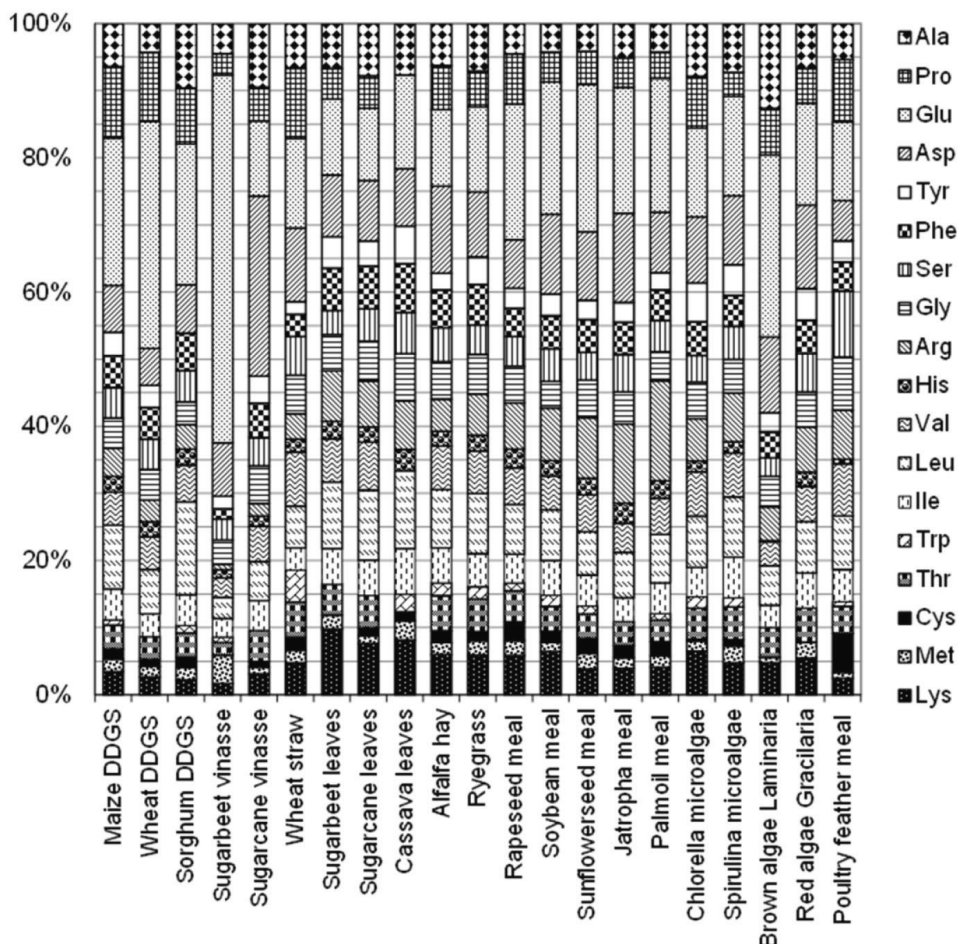


Figure 8.14 The relative presence of the different amino acids after hydrolysis of the proteins in a variety of agricultural residues. References are shown between brackets. Because of applied hydrolysis conditions, L-glutamic acids includes L-glutamine and L-aspartic acid includes L-asparagine [8].

An illustrative example from the biomasses shown in Figure 8.14 is L-glutamic acid in sugarbeet Vinasse. Once the proteins of sugarbeet vinasse have been hydrolyzed, approximately 60% of the amino acids are glutamic acid. It could thereafter be possible to lower the solubility of L-glutamic acid (e.g. lower the pH of the system to the isoelectric point of L-glutamic acid) and make it precipitate out.

8.4.2 Hydrolysis

Protein hydrolysis presented in the previous section was based on a full hydrolysis of proteins to their constituent amino acids. Another way to hydrolyze would be to do it step-wise, dividing the hydrolysate into fractions. The theory is that the bonds between some of the amino acids are stronger than the bonds between other amino acids. If a mild acid at a low temperature was used, the weakest bonds would be cleaved first. In this manner, certain amino acids could be selectively hydrolyzed. This idea was tested on poultry feather meal. Poultry feather meal is made up of 97% of the protein β -keratin (g/g) [8].

Two grams of the poultry feather meal used in this research was dried at 70°C in a vacuum oven until no weight loss could be observed. This appeared to be after 48 hours drying. The loss in weight was then used to calculate the moisture content. The moisture content of the used PFM was 3.8%.

Four hydrochloric acid solutions were made with pH values of 0.999, 1.475, 2.050 and 2.989 and with a biomass concentration of 50g/l PFM (see table 2). These solutions were hydrolyzed for 48 hours at 60°C using the Starfish reaction system (see Figure A8.1).



Figure 8.15: Hydrolysis setup of Poultry Feather Meal

The results from the fractional techniques are seen in Figure 8.. After 48 hours, the highest concentration of L-aspartic acid was found to be at 60 °C and at pH 2.050. 2.43 g/l of L-aspartic acid was measured. This represents 55% of the total L-aspartic acid in the Poultry Feather Meal calculated from Lammen [8] and shown in Figure 8.14. All concentrations of the amino acids are shown in Figure 8.16.

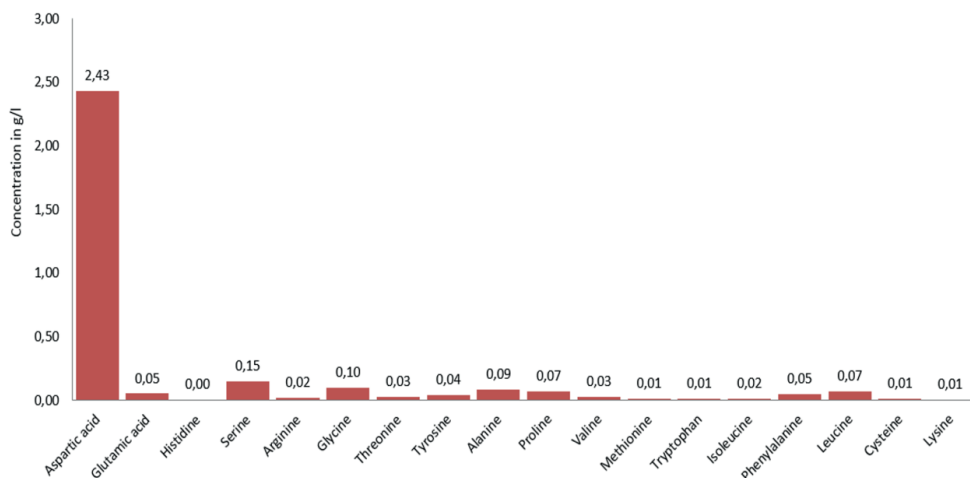


Figure 8.16: Concentration in grams of each amino acid per liter of distillate of the hydrolysis of poultry feather meal using a mild acid at 60 °C and initial pH 2.050. 2.43 g of L-aspartic acid was measured, compared to 0.15 g/l or lower for all other amino acids.

Using a mild acid at low temperatures, the amino acid L-aspartic acid was the first to be cleaved from the protein. In this fraction 55% of the total L-aspartic acid in the poultry feather meal has been hydrolyzed. If this fraction were taken away, then the L-aspartic acid could be separated. Furthermore, if the rest of the proteins were hydrolyzed, then L-glutamic acid would be the only negatively charged amino acid left in the second fraction. Bringing the system to the isoelectric point for L-glutamic acid would facilitate its precipitation. This would especially be the case if the second fraction had large concentrations of L-glutamic acid, such as sugarbeet Vinasse, shown in Figure 8.14 [9].

8.4.3 Conversion

Once the amino acids are in solution, it would be possible to convert an amino acid or a group of amino acids. When the amino acids are converted, their effects on the system are changed. This aids in their removal or in the removal of other amino acids [10]. Two examples are given below.

Of the two negatively charged amino acids, L-glutamic acid can be converted to L-pyroglutamic acid. The carboxylic acid on the side chain of L-glutamic acid bonds with the amine group on the alpha carbon forming a ring. This is done at temperatures around 160 °C. L-Pyroglutamic acid is not charged. The only other negatively charged amino acid is L-aspartic acid. If L-glutamic acid is converted, then if the pH of the system is brought to the isoelectric point of L-aspartic acid, then L-aspartic acid will precipitate. Furthermore, L-pyroglutamic acid can be reconverted back to L-glutamic acid and itself be precipitated out of solution [10].

Simultaneous decarboxylation of L-serine and deamination of L-phenylalanine can be done by using the enzymes L-serine decarboxylase and L-phenylalanine ammonia lyase. These enzymatic reactions convert the amino acids to ethanolamine and trans-cinnamic acid respectively. L-Serine and L-phenylalanine are not charged amino acids. However, ethanolamine and trans-cinnamic acid are respectively positively charged and negatively charged [11]. Now these two can be selected out by their charge, or other non-charged amino acids can be selected by their lack of charge (e.g. electrodialysis) and that the same conversion can aid separation on the basis of solubility [12].

8.5 Future Research

In this thesis, several topics were discussed. These topics included the solubility of amino acids with and the effects of co-solvents and being in mixtures. Thereafter, the amino acid solubilities were modeled in order to both compare the existing equilibrium models of solubility but also to present a new model. However, future research could focus on filling in the knowledge gaps of the solubility for the applications in a bio-refinery and in applying the models to the bio-refinery concept for amino acid recovery.

8.5.1 Solubility Data

While Chapters 2, 3 and 4 aimed at filling the gap in the data on amino acid solubility, there are still several areas where experimental data is required. These data are on both the extraction of amino acid from the biomass as well as the solubility of the data in solution.

Experiments in protein extraction and amino acid hydrolysis

Some proteins are hydrophobic [13]. Hydrophobic proteins may not allow the acid to hydrolyze the protein in a predictable manner. Surfactants like sodium dodecyl sulfate (SDS) are used to denature proteins [14]. This could help to hydrolyze hydrophobic proteins.

The solubility of other miscible and non-miscible organic solvents in amino acid separation could be studied. In the case of miscible solvents, the methodology in chapters 2,3 and 4 could be used. For non-miscible solvents, the liquid-liquid extraction technique could be used to calculate the partition coefficients.

This thesis has focused almost exclusively on L-amino acids. The reason for this was the L-amino acids are the amino acids found in proteins. There is some evidence that isomers of the same amino acid have an effect on their solubility (as discussed in Chapter 3). The effects of the isomers on solubility and crystallization could be looked at and thereafter modelled.

With the exception of the data in chapter 7 and 8, there is no data reported in the literature on the effects of pH on amino acid solubility in agro-industrial residues. It would be a good step to take these measurements in a model mixture, add them to the model and then test it on agro-industrial residues.

In this chapter, we look at the effects on solubility of an amino acid with another amino acid, but only focussed on a few amino acids. There is few other data on binary amino acid interactions and their effect on solubility and crystallization [15]. In order to further model on the interactions between amino acids, this should have a high priority.

One of the reasons why models on solubility are sometimes not accurate when applied to industrial solutions is because of impurities. Samples of impurities that are common in agro-industrial mixtures include sugars, fats, nutrients, salts, fatty acids among many others. The effects of these impurities on amino acids solubility should be studied.

Lastly, some of the parameters (e.g. Δh_{T_m} and T_m) that are used in solubility models cannot be measured. These parameters can be estimated. In the model in chapter 6, we used these estimations. However, a reliance on these estimations is a weakness in this model. The model needs these parameters to calculate the ratio of the fugacity of the liquid and the solid. However, another way to obtain this value is to measure the vapor pressure if the solute in the solvent [17]. These measurements could be carried out to no longer rely on the estimations in the model.

8.5.2 Models

The models developed in this thesis are not exhaustive. There are several lines of research that could be taken to improve the knowledge on solubility of amino acids and other molecules.

Expand model to two amino acids and water

The model in chapter 6 has been applied to a single solute and two solvents. However, this model can be used in a system of two amino acids and one solvent. In this way the interaction parameter between the amino acids can be calculated. In order to do this, more data on the effects of a secondary amino acid to the solubility of a primary amino acid would have to be produced.

Kinetic models

All of the models used in this thesis were equilibrium models. However, the use of the kinetic models could be significant in the application and study of the solubility and crystallization of amino acids in several ways.

A kinetic model could be developed for looking at the crystallization of amino acids. Parameters that could be included are temperature and pH and/or the addition of co-solvents. The rate of crystallization of amino acids are not the same. It could be possible that time could be a method to separate amino acids. For example, cool down the solution, one of the amino acids would crystallize and would be removed. Then, the temperature of the solution is increased again before the second amino acid crystallizes.

Add pH to Model

The model and experimentation in this thesis did not look at the effects of pH. This choice was made because pH has already been modeled for single amino acids extensively. This is a well-studied and relatively simple model as shown in Equation (8.1). However, especially for amino acids expanding the model by adding a variable for pH would expand the applicability of the model.

Model Expansion to non-amino acids

There is still a limited amount of data on amino acid solubility and the effects of co-solvents. Furthermore, some properties that the model in chapter 6 uses (e.g. Δh_{T_m} and T_m) are not measurable. However, for other molecules, there is already solubility data that is readily

available and that have measurable thermodynamic properties. It would be possible to test this model with the data of another solute and to confirm and improve it.

8.6 Conclusions and Recommendations

Finally, the proposed future work will give both insights into theory but also to the applications. The theory would be advanced by both to models, but also to how amino acids interact with each other and to solvents. The applications in the future could be an operational bio-refinery that produces individual amino acids.

In conclusion, this thesis provided novel insights in protein extraction, protein hydrolysis, amino acid separation and the use of models for solid-liquid interactions. In this way, future researchers benefit not only from the data and protocols developed from this research but also the knowledge gained from its application.

References

- [1] Sigma-Aldrich Amino acid reference chart <https://www.sigmaaldrich.com/life-science/metabolomics/learning-center/amino-acid-reference-chart.html> (accessed Jan. 2018)
- [2] Held, C.; Reschke, T.; Muller, R.; Kunz, W.; Sadowski, G. Measuring and modeling aqueous electrolyte/amino-acid solutions with ePC-SAFT *J. Chem. Thermodyn.*, 2014, 68, 1-12
- [3] Pradhan, A. A.; Vera, J. H. Effect of acids and bases on the solubility of amino acids. *Fluid Phase Equilib.* 1998, 152, 121-132
- [4] Carta, R.; Tola, G. Solubilities of L-cystine, L-tyrosine, L-leucine, and glycine in aqueous solutions at various pHs and NaCl concentrations. *J. Chem. Eng. Data* 1996, 41, 414-417
- [5] Jin, X. Z.; Chao, K.-C. Solubility of four amino acids in water and of four pairs of amino acids in their water solutions *J. Chem. Eng. Data*, 1992, 37, 199-203
- [6] Rezaee, M.; Assadi, Y.; Milani Hosseini, M.-R.; Aghaee, E.; Ahmadi, F.; Berijani, S. (2006). Determination of organic compounds in water using dispersive liquid-liquid microextraction. *Journal of Chromatography A*. 1116 (1-2): 1-9.
- [7] Shaoqu Xie, Wenli Song, Conghua Yi, Xueqing Qiu, Salting-out extraction systems of ethanol and water induced by high-solubility inorganic electrolytes, *Journal of Industrial and Engineering Chemistry*, 56, 2017, 145-150,
- [8] Lammens, T.; Franssen MCR; Scott, EL; Sanders, JPM; Availability of protein-derived amino acids as feedstock for the production of bio-based chemicals *Biomass and Bioenergy* 44 (2012) 168-181
- [9] Widyarani, E Ratnaningsih, JPM Sanders, ME Bruins; Biorefinery methods for separation of protein and oil fractions from rubber seed kernel *Industrial Crops and Products* 62, 323-332
- [10] Teng, Y; Scott, EL; van Zeeland, ANT and Sanders, JPM; The use of L-lysine decarboxylase as a means to separate amino acids by electrodialysis, 2011, *Green Chem.* 3, 624-630
- [11] Elisseeva, TV; Shaposhnik, VA; Luschik, IG; Demineralization and separation of amino acids by electrodialysis with ion-exchange membranes, *Desalination* 149, 2002, 405-409,

- [12] Teng, Y; Scott, EL and Sanders, JPM; The Selective Conversion of Glutamic Acid in Amino Acid Mixtures Using Glutamate Decarboxylase—A Means of Separating Amino Acids for Synthesizing Biobased Chemicals *Biotech Prog.* 30, 2014
- [13] Chandler, D; Interfaces and the driving force of hydrophobic assembly 2005 *Nature* 437 640–7.
- [14] Bhuyan, AK; On the mechanism of SDS-induced protein denaturation *Biopolymers* 2010 93 186-99
- [15] Carta, R.; Solubilities of L-Cystine, L-Tyrosine, L-Leucine, and Glycine in Their Water Solutions *J. Chem. Eng. Data* 1999, 44, 563-567
- [16] Barton, AFM CRC Handbook of Solubility Parameters and Other Cohesion Parameters, Second Edition, CRC Press 1991

Summary

In this thesis, experimental work and thermodynamic modelling has been presented. The application of this is the extraction of proteins from agro-industrial residue streams, the hydrolysis of the proteins to their constituent amino acids and the separation and crystallization of amino acids. The development of this process will benefit the reduction of waste and the worlds reliance on non-renewable resources.

Crystallizing individual α -amino acids from a mixture of α -amino acids would improve their separation in a bio-refinery, but the solubility of the 20 α -amino acids in a mixture of 20 α -amino acids is unknown. In order to understand the solubility of α -amino acids in mixtures of 20 α -amino acids, in chapter 2 the solubility of the 20 α -amino acids in single solution was studied using the van 't Hoff and Sober equations and data available in the literature. The results of these equations were compared and their coefficients are reported. Then we measured the solubility of the 20 α -amino acids in a model system of an industrial residue containing all 20 α -amino acids and found that only L-tyrosine, L-tryptophan, L-aspartic acid, L-glutamic acid and L-asparagine dissolved in higher concentrations in the model system of 20 α -amino acids than in aqueous solution by themselves. The solubility of all of the aliphatic α -amino acids was lower in the mixture. The longer the side chain length of the aliphatic α -amino acids, the greater the decrease in solubility in the mixture as compared to being in solution by itself. The maximum solubility of L-lysine and L-glutamine showed the most decrease when in the mixture in comparison to in aqueous solution alone. Our results demonstrate that α -amino acids with similar physical structures have similar changes in solubility when in a mixture of α -amino acids. The assertion was that α -amino acids can be grouped by their physical structure.

In chapter 3, the addition of organic solvents to α -amino acids in aqueous solution was shown to be an effective method in crystallization. The available data on the solubility of α -amino acids in water, water-ethanol mixtures and ethanol at 298.15 K and 0.1MPa were reviewed. The solubility of L-alanine, L-proline, L-arginine, L-cysteine and L-lysine in water and ethanol mixtures and the solubility of L-alanine, L-proline, L-arginine, L-cysteine, L-lysine, L-asparagine, L-glutamine, L-histidine and L-leucine in pure ethanol systems were measured and are published here for the first time. The impact on the solubility of amino acids that can convert in solution, L-glutamic acid and L-cysteine, was studied. At lower concentrations, only the ninhydrin method and the UPLC method yielded reliable results. In

the case of α -amino acids that convert in solution, only the UPLC method was able to discern between the different α -amino acids and yields reliable results. The results demonstrated that α -amino acids with similar physical structures have similar changes in solubility in mixed water/ethanol mixtures. The solubility of L-tryptophan increased at moderate ethanol concentrations.

Since the knowledge on the effects of organic solvents on the solubility of α -amino acids is incomplete, the intermolecular forces between the α -amino acids are not fully understood. In chapter 4, the solubilities of each of the 20 proteinogenic α -amino acids in a mixture of all the 20 proteinogenic α -amino acids in 0%, 15%, 30%, 50%, 70% and 80% (g/g) ethanol-water solutions and at 277.15, 297.15, 315.15 and 335.15 K are reported and compared to literature values of single α -amino acids. The solubility of the individual α -amino acids in water, ethanol and water-ethanol mixtures were discussed in chapter 3 and are compared with this data. The results showed that amino acids can be grouped according to the structure of their side chains. If branched aliphatic, hydroxylic, phenylic and carboxylic groups were on the side chain of an amino acid, then the solubility of that amino acid increased when in a mixture of 20 amino acids between 30-70% (g/g) ethanol-water solution. If sulphur containing and amine/amide containing groups were on the side chain of an amino acid, then the solubility of that amino acid would decrease in a mixture of 20 amino acids between 30-70% (g/g) ethanol-water solution compared to the solubility as a single amino acid.

In chapter 5, the solubility data of all proteinogenic α -amino acids in binary ethanol/water systems was used to model their excess solubility. The empirical and regressive models of Gude and NRTL and the predictive Jouyban-Acree model were applied. Based on the results, it is hypothesized that amino acids that are spherical and lack a reactive side chain show little or no excess solubility. Being rod-like and/or having a reactive side chain lead to a positive excess solubility in a mixed solvent of ethanol and water. The empirical and regressed models, NRTL and Gude, fit the data well and the predictive Jouyban-Acree model, not originally intended to be used for small molecules, was less accurate but offers insights into the thermodynamic properties of the amino acids.

In chapter 6, in order to predict α -amino-acid solubility in a solution of water and ethanol, the van Laar equation for the molar excess Gibbs energy was used. To obtain meaningful activity coefficients from the solubility data using the van Laar equation, the ratio of the fugacity of the solid α -amino acid to that of the subcooled liquid amino acid was needed. That ratio is obtained from estimated melting temperatures and enthalpies of fusion.

The ternary van Laar equation provided a predictive model for obtaining the solubility of an α -amino acid in an ethanol-water solvent. The normalized root mean square variances (NRMSV) for 16 of the 20 solubility predictions were below 0.100, indicating very good agreement with the solubility data. The NRMSV of the other four predictions were below 0.220, indicating good agreement with the α -amino-acid solubility data. Additional group contribution data was reported here for seven amino acids.

Amino acids can be obtained from protein after hydrolysis. In addition, several agro-industrial residues already contain a mixture of free amino acids. The objective of chapter 7 was to develop a method for amino acids separation, starting from mixtures containing amino acids, and using anti-solvent precipitation with ethanol. Protamylasse™, rubber seed protein hydrolysates and grass juice were used in the experiments, representing existing and potential agro-industrial residues. The results show that in a water-ethanol system, some amino acids had lower solubility in mixtures than as a single component, thereby facilitating precipitation. A sufficiently high total amino acid concentration in the mixture was needed to achieve precipitation, therefore a concentration step is sometimes required. Ethanol precipitation can be applied as a pre-treatment to separate mixtures into groups of amino acids or as a polishing step to increase purity.

In conclusion, this thesis provided novel insights in protein extraction, protein hydrolysis, amino acid separation and the use of models for solid-liquid interactions. In this way, future researchers benefit not only from the data and protocols developed from this research but also the knowledge gained from its application.

Acknowledgements

This Ph.D. thesis covers my work for the past four and a half years. I would not have been successful without the support of many people and organizations. I would like to express my sincere gratitude to all of them.

The people that I would first like to thank are my supervisors Johan Sanders and Marieke Bruins. It was Johan who accepted me into his research chair and I am truly grateful. You opened my world to the possibilities of bio-refineries. I was also truly lucky to have Marieke as my closest collaborator during my time as a Ph.D. student at Wageningen University & Research. Even though you switched groups, you remained an excellent supervisor, reading the many versions of my articles and thesis, correcting the almost endless problems in my research and giving me clear guidance. You were truly invaluable in many ways.

I am thankful to all of the student assistants that have helped me over the years. Their work was always valuable to my research. I thank you Krisztina, Nigel, Brendan, Rania and Thalita.

There are three people who helped me on three different articles that are printed in this thesis. They are Prof. Dr. John Prausnitz, Dr. Widyarani and Dr. David Mendez Sevilano. John was an inspiration to work with and opened my eyes and my thinking to the molecular thermodynamics of liquids. Widyarani showed me the way on how to complete a high quality Ph.D. thesis. David, on the other hand, was fearless in guiding me in the ways of Matlab. I am afraid I was not always the best student, but we got there eventually (mostly due to your brilliance).

I would like to thank the Netherlands Organisation for Scientific Research who granted me funding for my research.

I would like to thank the Management Board of Avans University of Applied Sciences and the management team of the School of Life Sciences that granted me the time to work on my research that lead to this thesis. I would also like to thank my colleagues at Avans University who were always there in different capacities. I especially would like to thank the Friday Night Crew (Ellen, Paul, Laurens and Wilfred) who listened to my complaints in a festive manner. I would also like to thank my mother for her wisdom.

My fellow Ph.D. students and the staff at Biobased Chemistry and Technology at Wageningen University & Research were and remain a constant support to me. It has been at times a perilous voyage where both the path and the destination were unknown, confusing and

unexpected ...and in the case of the Ph.D. trip to China, I mean this quite literally. To Susan, without whom all the analytical work would have been much more difficult, to Gerda who aided me administratively in too many ways to count, to my office mates Piet and Sanne who didn't seem to mind me too much. To Elvira, David, Farnoosh, Daniel, Roxani, Yu, Andrada, Gwen, Frits, hurry up already, you can't let this old man show you up. To the newer Ph.D. students, I wish you the best of luck, but can offer few words of wisdom other than to follow your own inner guiding voice, especially if that voice says to take a coffee break or a walk outside.

Lastly, I would like to remember Marja Teekens. You were my role model and we miss you.

Curriculum Vitae

Nathan Bowden was born on July 1, 1974 in pastoral southern Indiana, USA. He attained his a Master of Science degree from Indiana University, Bloomington, Indiana USA in analytical chemistry in Environmental Science, researching the effects of agricultural practices on riverine water quality in 2000.

From 2000 to 2009, Nathan was a researcher at several research organizations including the Smithsonian Institute, the Technical University of Delft and TNO. In 2009, Nathan joined the Avans University of Applied Science as a lecturer in Environmental Science where he is still working today.

Nathan started his Ph.D. research with funding from Avans University of Applied Sciences and the Netherlands Organisation for Scientific Research under the supervision of Prof. Dr. Johan P. M. Sanders and Dr. ir. Marieke E. Bruins in 2003.

Nathan can be contacted at n.a.bowden@gmail.com.



List of Publications

Bowden, N.A.; Sanders, J.P.M. and Bruins, M.E. (2018) Solubility of the Proteinogenic α -Amino Acids in Water, Ethanol, and Ethanol–Water Mixtures *J. Chem. Eng. Data* DOI: 10.1021/acs.jced.7b00486

Bowden, N.A.; Sanders, J.P.M. and Bruins, M.E. (2018) Solid-liquid phase equilibria and thermodynamic properties of mixtures of 20 α -amino acids in a model system of industrial residues Submitted

Bowden, N.A., Bruins, M.E., Sanders, J.P.M. and Prausnitz, J. (2018) Molecular-thermodynamic Correlation of Solubility Data for 20 Amino Acids in Water, in Ethanol and in Water-Ethanol Mixtures Submitted

Bowden, NA; Mendez Sevilano, D; Sanders, JPM and Bruins, ME (2017) Modeling the effects of organic solvents on the solubility of proteinogenic amino acids in aqueous solution *Fluid Phase Equilibria* DOI: 10.1016/j.fluid.2017.11.036

Widyarani; Bowden, NA; Kolfshoten, RC; Sanders, JPM and Bruins, ME (2016) Fractional precipitation of amino acids from agro-industrial residues using ethanol. *Ind. Eng. Chem. Res.* 55, pp 7462–7472 DOI: 10.1021/acs.iecr.6b00054

Overview of Completed Training Activities

Discipline specific activities

- Biorefinery and Biomolecules, WUR/VLAG, Wageningen, the Netherlands, 2015
- Symposium bio-refinery for food, fuel and materials, WUR, Wageningen, the Netherlands, 2014
- Sustainability analysis in food and bio-based production, WUR/VLAG, Wageningen, the Netherlands, 2015
- DNA Techniques, Avans University of Applied Sciences, Breda, the Netherlands, 2011
- AIChE Conference 2016, American Institute for Chemical Engineering, San Francisco, USA, 2016
- Chemometrics, WUR/VLAG, Wageningen, the Netherlands, 2016
- Downstream Process Engineering, TU Delft, Delft, the Netherlands, 2016
- Symposium Biobased Economy, Center of Expertise Biobased Economy, the Netherlands, 2015-2018

General Courses

- IPMA D (project management), International certificate, 2008
- Didactic Training, Avans University of Applied Science, Breda, the Netherlands, 2012
- Creating an international program, EAIE, 2013

Optional activities

- Preparation of research proposal, Wageningen, the Netherlands, 2013
- Ph.D. trip to China, WUR, 2015
- Theme meetings BCH, WUR, Wageningen, the Netherlands, 2013-2018
- Fellowship UC Berkeley, University of California, Berkeley, CA, USA 2017
- Lectorate Biobased Energy, Avans University of Applied Science, Breda, the Netherlands, 2014-2017

The research described in this thesis was financed by Avans University of Applied Science and the Doctoral Grant for Teachers from the Netherlands Organisation for Scientific Research

