THERMODYNAMIC AND KINETIC INVESTIGATION OF DEPSIPEPTIDE OLIGOMERIZATION AND DEGRADATION UNDER PREBIOTIC EARTH CONDITIONS

A Dissertation
Presented to
The Academic Faculty

By

Kelvin Howard Smith

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Engineering
Department of Chemical and Biomolecular Engineering

Georgia Institute of Technology

December 2022

THERMODYNAMIC AND KINETIC INVESTIGATION OF DEPSIPEPTIDE OLIGOMERIZATION AND DEGRADATION UNDER PREBIOTIC EARTH CONDITIONS

Thesis committee:

Dr. Martha Grover
Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Anant Paravastu
Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Charles Liotta
Chemistry
Georgia Institute of Technology

Dr. Fani Boukouvala
Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Michael Filler
Chemical and Biomolecular Engineering
Georgia Institute of Technology

Date approved: June 17, 2022

I believe everyone on the planet has their thing and, especially in my experience, autistic people all have a tremendous gift. It's a matter of finding that gift and nurturing it.

Edie Brannigan, Mother to runner, Mikey Brannigan

For my mother LaRita Smith

For my grandfather Howard Smith

ACKNOWLEDGMENTS

I owe my entire Ph.D. journey to primarily four people: my mother LaRita Smith, my advisor Dr. Martha Grover, my co-advisor Dr. Charles Liotta, and my colleague Dr. Moran Frenkel-Pinter.

When I started my Ph.D. program, I had quite a difficult time finding an advisor. Compounding the matter is the fact that I am on the autism spectrum which limits my social skills greatly. I began to lose hope until one day Dr. Martha Grover sent an e-mail to me saying that even though she didn't have an open spot, that doesn't mean that one couldn't be created for me. That did it. Dr. Grover saw something in me. She believed that I could be a great asset to the Grover Group. I transitioned into a computational role which I was strongest in. She wanted me to succeed and I give great thanks to her. I've always been successful being a consumer of knowledge. Dr. Grover and Dr. Liotta showed me that I can be a creator of knowledge and still be myself. I am thankful for the opportunity to contribute significantly to the scientific field with my thermodynamic and kinetic models.

I would like to thank all of my thesis committee members for their support throughout the years. I also want to thank Dr. Carson Meredith for his support and guidance, Dr. Anant Paravastu for a wonderful TA experience, Dr. Julie Champion for my Tech to Teaching Capstone experience, and all of the faculty and staff in the Tech to Teaching program including Dr. Kate Williams, Dr. Tammy McCoy, and Dr. David Lawrence. I would also like to thank Ms. Stephanie Reikes for the volunteer lecturer experience she gave me in the Peer-Led Undergraduate Study (PLUS) program.

I would like to thank two undergraduate researchers which include Kaitlyn Jacobson and Sara Olivieros. Their lab experiments assisted me greatly in my research.

I also thank the Center for Chemical Evolution (CCE) headed by Dr. Christine Conwell and Shantel Floyd. I met Dr. Moran Frenkel-Pinter there. She has been a dear friend and colleague to me and has been strongly supportive during my journey. I was very happy

to work together with her during my modeling research. She was a great person to talk to when I needed help. There are countless others in the CCE who have supported me including Dr. David Fialho, Dr. Jay Forsythe, and Dr. Martin Solano.

Next, I want to thank Ms. Jacqueline Mohalley Snedeker for taking time to mentor me on improving my technical writing skills.

Finally, my biggest cheerleader of all is my mother. She has seen it all. The happiness, the sadness, the triumphs, the challenges. I took her on a roller coaster ride of a lifetime for me to pursue my Ph.D. studies. She believed I could do it and could have easily given up and urged me to quit, but she didn't. I dedicate this Ph.D. to her as well as my deceased grandfather. My mother and grandfather have taught me to be the man I am today. Without them, none of this possible: no college, no grad school, no dissertation, no Ph.D., nothing.

The Grover group has been extremely accepting of me as well. I thank Chiamaka Obianyor, Aaron Liu, Stefani Kocevska among others. It has been an honor presenting, talking, and sharing quality time with each of you. I always felt included and accepted, and that means the world to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The formation of polypeptides on the early Earth has been a long-standing problem in the field of prebiotic chemistry. Although it is generally accepted that amino acids were present on the prebiotic Earth based on the Miller-Urey experiment, the plausible mechanism to form long chain polypeptides is still unclear. This mechanism of polypeptide formation is known to be slow unless high temperatures or metal catalysts are used because of the high activation energies and the formation of side-products such as diketopiperazine (DKP).

This thesis describes a simple system containing hydroxy acids and amino acids that combine to form depsipeptides, which are hypothesized to be precursors to polypeptides. Hydroxy acids form metastable oligoesters in an oscillating (hot dry/cool wet) environment and transform into mixed copolymers via the ester-amide exchange reaction. This pathway enables the amide bond formation in lower temperature and leads to a library of oligomers with random sequences containing catalytic functions crucial for sustaining life.

The main theme of this thesis is thermodynamic and kinetic modeling of depsipeptide degradation, esterification, and exchange of prebiotic molecules such as lactic acid (a), glycolic acid (g), and glycine (G). The goal is to be able to explain the chemical driving forces of depsipeptide oligomerization and degradation under mild conditions using classical thermodynamics and kinetics. The specific aims are as follows: 1) determine whether models can be used to visualize how Gibbs free energy of lactic acid esterification and ester-amide exchange changes as a function of pH and temperature; 2) assess what type of behavior is anticipated from oligoester degradation, and finally 3) use the models to extract important parameters from experimental data such as the pH range that favors depsipeptide oligomerization. The answers to all these questions can be found in the five chapters of this thesis.

Chapter 1 introduces the thesis. It discusses the background information of esterifica-

tion and the ester-amide exchange reaction and provides a thorough literature review of the reaction mechanisms and existing theories including the gaps. After reading the introduction, the reader will understand the motivation and objectives behind my research.

Chapter 2 identifies and investigates the relevant kinetic mechanisms of depsipeptide degradation. Population balance kinetic models are created to form concentration profiles that yield valuable kinetic information such as half-lives and Arrhenius behavior. Depsipeptides are copolymers of hydroxy acids and amino acids. The study focuses on ester bonds since the amide bonds are more stable under plausible prebiotic conditions. This type of research is important for two reasons. The first reason is that once these mechanisms of depsipeptides are identified, we can predict behavior of similar peptides used to create biopolymers for pharmaceutical purposes such as drug delivery and tissue engineering. The second reason is that depsipeptides are much easier to form than traditional polypeptides. To investigate the kinetic mechanisms of ester reactions in depsipeptides, high performance liquid chromatography (HPLC) data of two test molecules are collected and quantified at different pH values and temperatures: multiple oligomers of an HO-dipeptide called glycolic acid alanine (gA) and a methylated molecule of the gA dimer called propionic acid alanine glycolic acid alanine ((PA)AgA).

Chapter 3 provides a cohesive framework that explains the thermodynamic driving force of oligomerizations involving ester bond linkages under plausibly prebiotic conditions. The framework predicts thermodynamics of oligomerizations of dimers and trimers as well as cyclic esters as function of pH and temperature for any hypothetical solution. Density functional theory (DFT) is used to compute thermochemical data and a benchmark compound is used to fix a critical parameter, pK_a of the homodimer of lactic acid.

Chapter 4 demonstrates a first-principles thermodynamic model that inputs key parameters such as pK_a and allows visualization of a free energy map that shows the spontaneity of the exchange reaction gg + G $\leftrightarrow$ gG + g as a function of temperature and pH. This model also shows the thermodynamic driving forces of the exchange reaction. A key result is that

the reaction seems insensitive, and thus robust, over a range of dimer pK_a's. A grid search of the unknown pK_a values of gG show what the possible pK_a values could be for each dimer at 25°C, to favor the reaction.

Chapter 5 concludes the thesis, summarizes its key points, and discusses potential future work.

It is hoped that the reader gains a great appreciation for fundamental chemical modeling of a longstanding puzzle of the Earth: the origin of life. A secondary motivation of this dissertation is the pharmaceutical application of tissue engineering, hydrogels, and drug delivery.

The models provide a computational and chemical engineering view of the problem. As the famous statistician George Box has stated, "All models are wrong, some are useful." I believe that after seeing the models presented in this dissertation, they will be useful in the origins of life field.

CHAPTER 1
INTRODUCTION AND BACKGROUND

Direct oligomerization of amino acids to form peptides is difficult under prebiotic conditions [1–6]. In the 1950s, Stanley Miller and Harold Urey performed experiments that showed how amino acids may have formed in the early Earth [7, 8]. The only materials allowed during the experiments were water, methane, ammonia, and nitrogen, as they are believed to be representative of the early Earth's atmosphere. Figure 1.1 shows their experiment. The compounds were all sealed and circulated inside a sterile array of glass tubes and flasks with one flask half-full of liquid water and another flask containing a pair of electrodes. The liquid water was heated to add water vapor to the chemical mixture, and the resulting gases were circulated around the apparatus, simulating the Earth's atmosphere. Sparks were fired between the electrodes to simulate lightning storms (believed to be common on the early Earth) through the water vapor, and then the vapor was cooled again so that the water could condense (simulating the oceans) and trickle back into the first water flask in a continuous cycle. At the end of the experiment, Miller and Urey observed that the cooled water contained organic molecules including amino acids. They successfully demonstrated that the simplest α-amino acids and α-hydroxy acids formed in plausibly prebiotic conditions as they understood the conditions at that time [3, 7, 9–11].

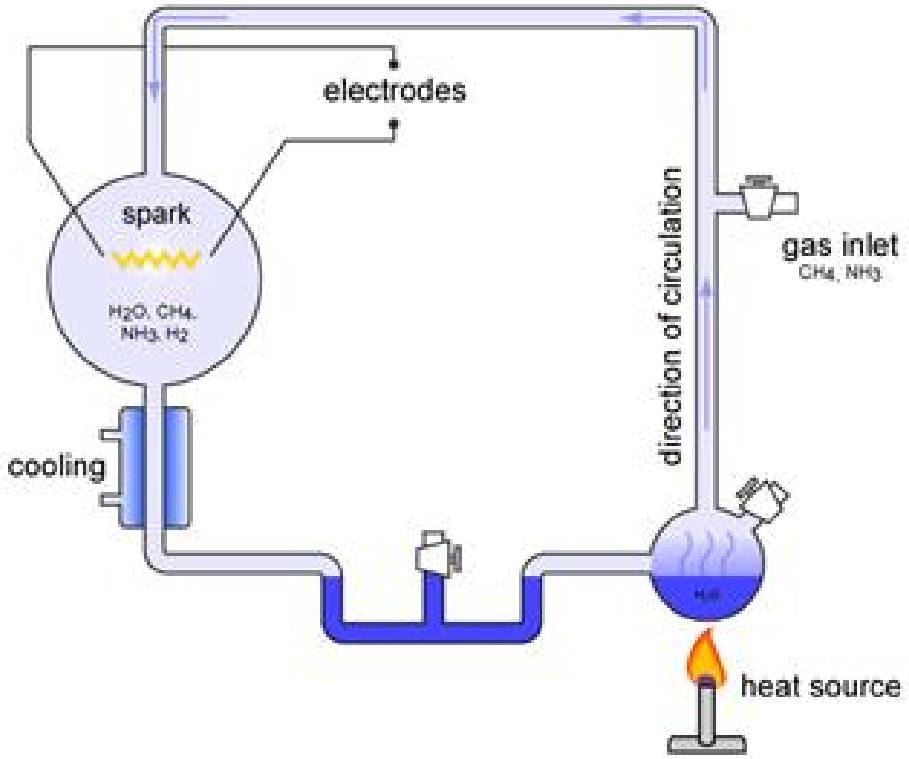

Figure 1.1: Miller-Urey Experiment. "From Ref. [7]. Reprinted with permission from AAAS."

Some compounds from the Miller-Urey experiment are seen in Figure 1.2 and Figure 1.3. As groundbreaking as this experiment is, there is still an important fundamental question that the experiment fails to address: What mechanisms allow these amino acids to form peptides in prebiotic conditions?

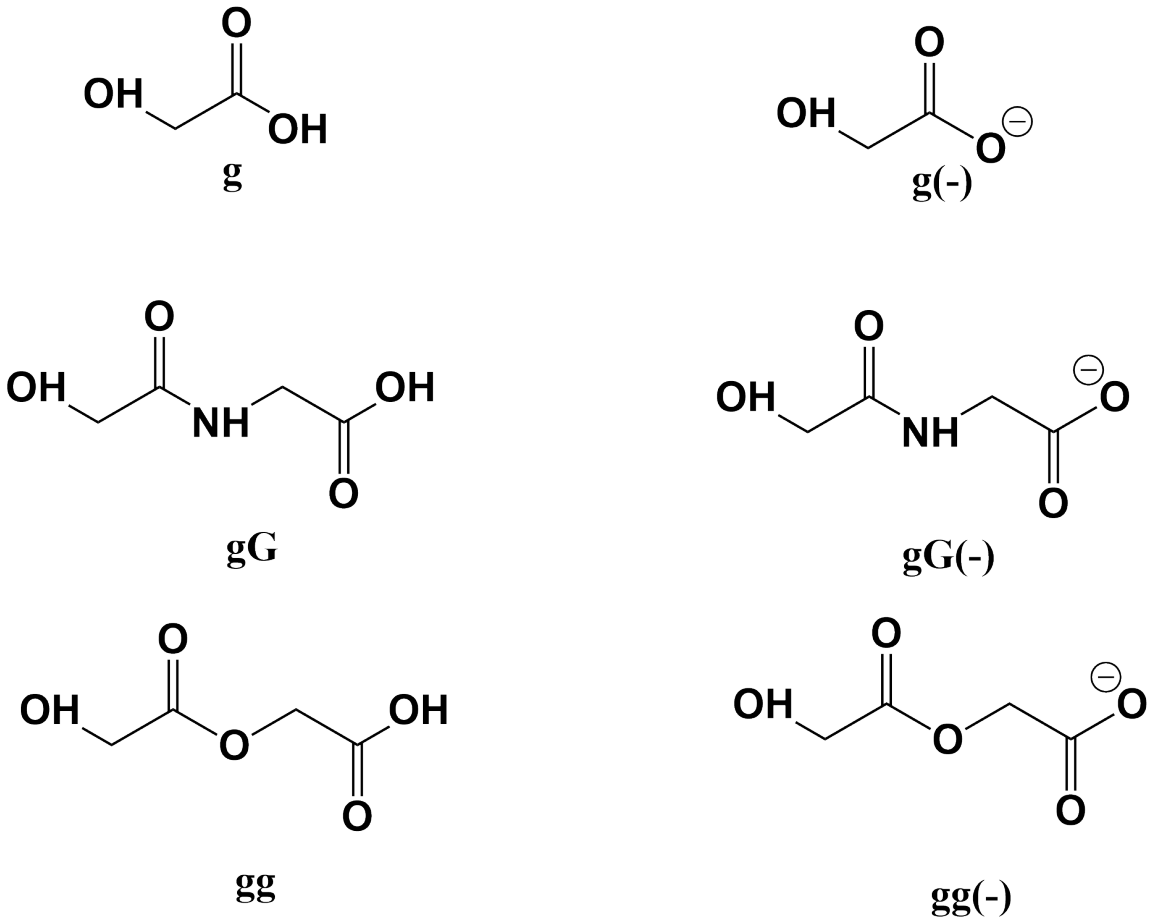

Figure 1.2: Hydroxy-acids and derivatives discussed in dissertation. From these compounds, only glycolic acid was found in the Miller-Urey experiment.

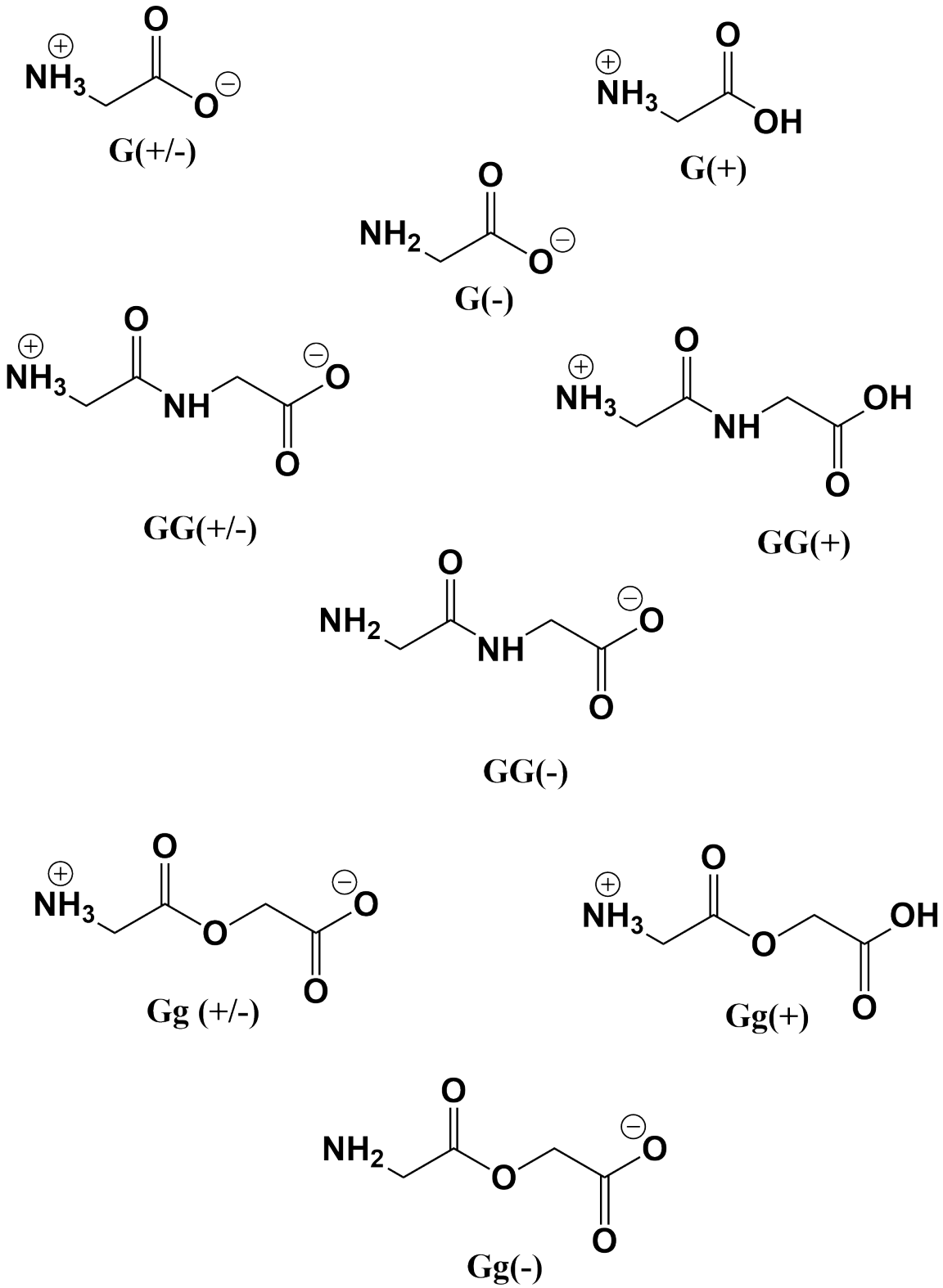

Figure 1.3: Amino acids and derivatives discussed in dissertation. From these compounds, only glycine was found in the Miller-Urey experiment.

For two amino acids to form a peptide, an amide bond needs to be formed, but the formation of this bond is thermodynamically unfavorable at neutral pH since its Gibbs free energy is positive [9, 12]. However, the literature has reported computational calculations of peptide bond formation to be negative [13]. This disagreement provides motivation to delve deeper into the thermodynamic driving forces of amide bond formation.

Even though today's peptides can be formed through enzymes, catalysts, and other compounds, enzymes did not exist in the prebiotic Earth [2, 14, 15]. So how did early peptides form on the prebiotic Earth?

To answer this question, we must first ask why the amide bond formation is difficult. One obstacle of direct oligomerization of amino acids is the formation of diketopiperazine (DKP), a cyclic amide that acts as a thermodynamic sink thus hindering the reaction [1, 3, 4, 16]. Once DKP forms, the reaction largely ceases because it is extremely difficult to perform a ring opening on DKP due to its stability.

Direct oligomerization of α-hydroxy acids is relatively easy under plausibly prebiotic conditions, which helps provide alternate pathways to polypeptide formation [1–6]. In the recent literature, α-hydroxy acids such as glycolic acid and lactic acid have been found to provide an alternate possible pathway to peptide bond formation on the prebiotic Earth [6]. The reason is that α-hydroxy acids polymerize easily into polyesters because when they polymerize, they form a cyclic ester, not a cyclic amide. This distinction is important because unlike cyclic amides, cyclic esters readily hydrolyze, and they undergo ring-opening polymerization, so the reaction does not cease [17]. From this observation, it was hypothesized that polyesters would have existed before polypeptides in the prebiotic Earth [6]. However, the goal is to form polypeptides not polyesters, meaning that amide bonds must be introduced.

Through a process called ester-amide exchange, indirect oligomerization of amino acids can be achieved by first synthesizing polyesters through direct oligomerization of α-hydroxy acids followed by exchanging the ester bond with an amide bond [3, 6, 18]. Now we have

found a prebiotic method of introducing peptide bonds during polymerization, but degradation, the reverse of polymerization, is another reaction that must be investigated.

Why should we care about the degradation reaction and its underlying mechanisms? When there are multiple amino acids, peptides can form many different sequences. This collection of polymers is known as the sequence space [7]. Only a small subset of these sequences would have catalytic functions suitable for biomedical and pharmaceutical applications such as drug delivery or tissue engineering [19, 20], or for catalysis. Polymerization by itself is necessary but not sufficient for building more complex early proteins capable of evolving and performing important biological functions such as sustaining life. Therefore, there is a need to break down and reform new sequences to discover more functional sequences.

A balance between these two mechanisms of building up and breaking down is required, so quantifying the polymerization and degradation rates is critical to the understanding of how these early peptides can self-assemble to form early proteins capable of sustaining life.

Knowing the mechanism and energetic pathways of these reactions under various environmental conditions is key to understanding what factors affect the thermodynamics and kinetics of reactions [3, 21]. The environmental conditions in question are pH, temperature, salt, and hydration levels. Another pathway to oligopeptides is wet/dry cycling. These cycles intend to simulate what plausibly occurred during a daily cycle on the prebiotic Earth and are hypothesized to be responsible for creating oligopeptides [3, 6, 14, 21, 22]. The wet and dry cycles each lasts 12 hours. During the wet phase, degradation of oligopeptides occurs while oligopeptide formation occurs during the dry phase. This dissertation primarily focuses on the thermodynamic and kinetic driving forces of depsipeptide chemistry during the wet phase. However, the dissertation offers a brief analysis of reaction thermodynamics when water is driven out of solution.

The general problem is that thorough thermodynamic studies of the oligomerization of simple α-amino acids [9, 12, 13] and oligomerization of α-hydroxy acids [13, 23, 24]

are scant at best. The dearth of literature studies of thermodynamic driving forces extends to depsipeptides. Depsipeptides are mixtures, or co-polymers, of α-amino acids and α-hydroxy acids. Little work has been done to elucidate the mechanism and kinetics of degradation of depsipeptides. Yu [14], however, provided a kinetic framework of the polymerization and degradation of a depsipeptide composed of the α-amino acid valine and α-hydroxy acid lactic acid. He found that both the polymerization and the ester degradation rates followed an Arrhenius expression, but he did not consider the pH dependence of either polymerization or degradation. He also created a free energy landscape containing activation parameters for both esterification and exchange and found that the former is slightly favorable while the latter was slightly unfavorable.

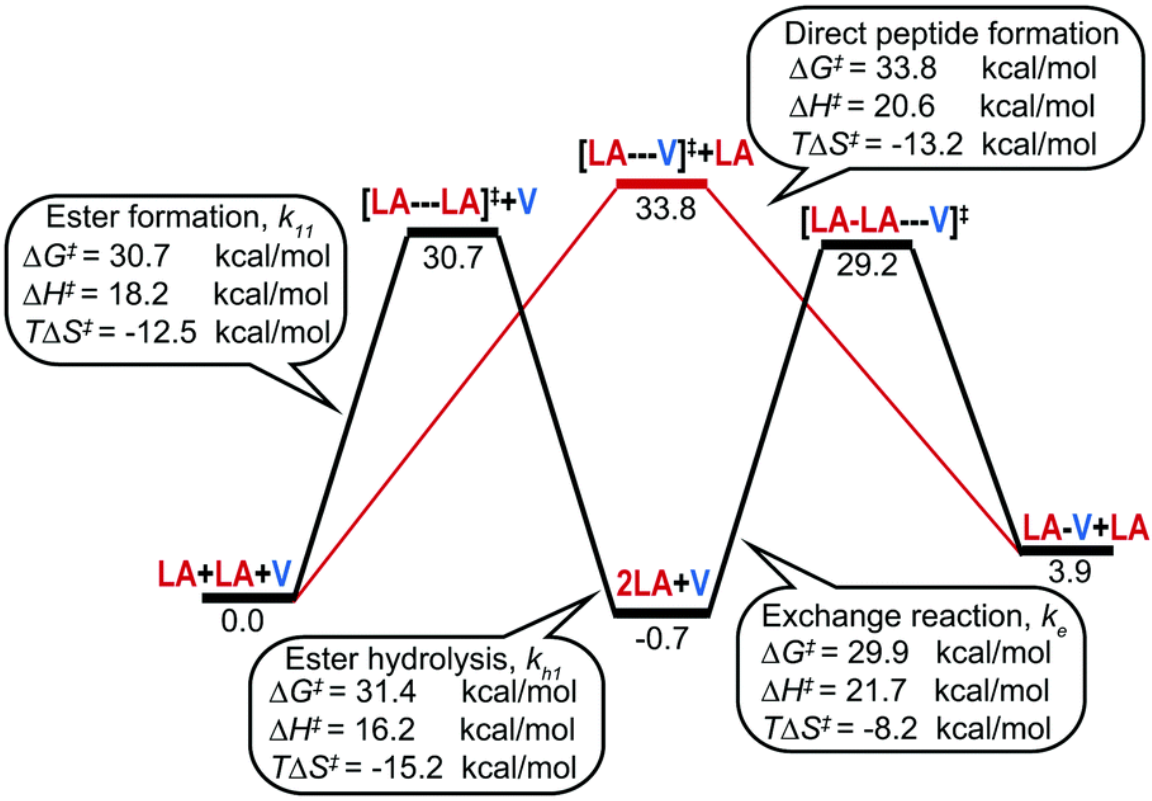

Figure 1.4: Free energy landscape for the polymerization kinetics and thermodynamics of lactic acid and valine. The numbers below the species' transition states and ground states represent the Gibbs free energy of reaction. Reproduced from Ref. [14] with permission.

Lactic acid is the most extensively studied α-hydroxy acid in the literature [19, 23–28], but literature studies primarily deal with the kinetics of lactic acid oligomerization. There is thermodynamic data of lactic acid and other relevant prebiotic molecules, but they mainly exist at standard temperature and pressure with pH unadjusted. This is a major limitation of the existing thermodynamic data.

Polylactic acid (PLA) is a commonly used biomaterial [29–33] and is a biorenewable feedstock [23, 24, 34]. Its monomer lactic acid is a biological metabolite and has been proposed as a key participant in the origin of polypeptides on the early Earth about four billion years ago [6, 7, 22, 35–37]. Due to its importance as a biomaterial, the hydrolytic kinetics have been studied as a function of temperature and pH [19, 25–28]. The polycondensation reaction for lactoyllactic acid, the linear lactic acid dimer, is seen in Figure 3.1 and is the primary reaction studied in this chapter.

To determine the thermodynamic driving force of amide bond formation in depsipeptides, the first step is to form an ester linkage using α-hydroxy acids. These reactions are reported to be thermoneutral, meaning that the enthalpy change is near zero [24, 38]. There is a disagreement in the ΔG sign of the ester linkage formation: Some papers report a slightly positive ΔG [13, 23, 24, 26] while others report a slightly negative ΔG [6, 28, 38]. Even though the literature is not clear on the sign of ΔG, the value of zero is within the confidence interval of ΔG and ΔH. No error analysis has been done on either parameter, so this is a possibility. Another possibility is that the experiments performed in the literature occur at only one pH. It is necessary to view Gibbs free energy and enthalpy from a range of pH values which the thesis discusses.

After the ester linkage is formed, it can be exchanged for an amide linkage in a process called ester-amide exchange. This process has been recently studied in the origins of life field at Georgia Tech because it is well-suited to form a peptide bond under prebiotic conditions, and it bypasses the formation of a diketopiperazine (DKP), a cyclic amide that acts as a thermodynamic sink halting the direct amidation reaction [6, 13, 14, 18, 22]. The

thermodynamic driving force of the exchange reaction is not fully understood yet. For example, Kua and Sweet [13] performed a DFT analysis on reactions involving glycolic acid and glycine including exchange reactions at standard conditions (pH = 7, T = 25°C, and P = 1 atm) by computing the Gibbs energies of formation of these species. They did not take dissociation states into account although they used a zwitterionic correction to those energies.

Very few thermodynamic models have been implemented in determining the thermodynamics of reactions involving α-amino acids and α-hydroxy acids [9, 12–14]. What researchers know is a plausible mechanism on how peptides are formed in the early Earth and the experimental thermodynamic data of oligomerization of simple α-amino acids and α-hydroxy acids [6, 18]. What we do not know is if there is a model that can accurately describe the thermodynamics of oligomerizations under plausible prebiotic conditions [39]. Researchers are interested in how temperature and pH affect the thermodynamic driving forces of the oligomerization reactions [4, 19, 25, 40].

The purpose of this thesis is to investigate the thermodynamics and kinetics of relevant prebiotically plausible reactions of simple α-amino acids and α-hydroxy acids in aqueous solutions seen in Figure 1.5 using models. This purpose is subdivided into two main goals: 1) provide a computational and statistical model approach towards calculating degradation rate constants of depsipeptides and 2) to describe the thermodynamic driving force of creating an ester bond linkage and an amide bond linkage under prebiotic conditions using a cohesive theoretical model involving thermodynamics and density functional theory (DFT) to compute Gibbs free energies and overall conversion as a function of pH and temperature.

Tackling these objectives will help elucidate clearly what degradation mechanism is most dominant and how well these modified models simulate depsipeptide chemistry. Polymerization of depsipeptides was not modeled due to the difficulty of taking into account volume changes and the dry state in the modeling. Another factor is the availability of quantitative data for polymerization.

As a result, this thesis provides a piece to the large puzzle that is the origin of life, as well as a better understanding behind the reaction mechanisms that play a major role in many scientific fields such as origins of life, astrobiology, pharmaceutics, and catalysis.

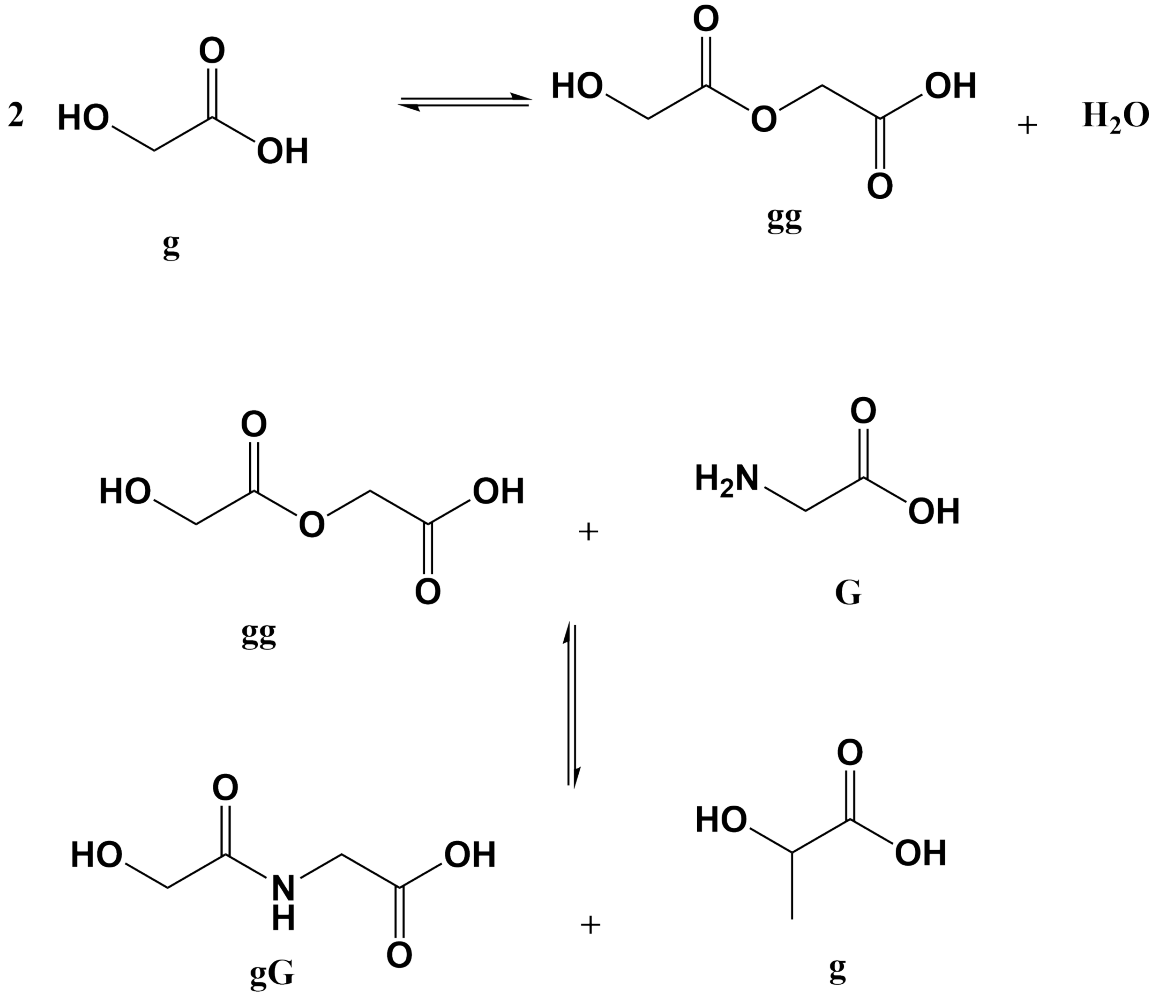

Figure 1.5: Top: Esterification reaction (g + g $\leftrightarrow$ gg + H_2O). Bottom: Exchange reaction (gg + G $\leftrightarrow$ gG + g).)

CHAPTER 2
MECHANISTIC INVESTIGATION OF DEPSIPEPTIDES IN THE EARLY EARTH THROUGH COMPUTATIONAL MODELING

2.1 Abstract

This chapter focuses on the identification and investigation of the relevant kinetic mechanisms of ester bond degradation in depsipeptides. Depsipeptides are co-polymers of hydroxy acids and amino acids. The kinetic driving forces were identified through rate constant analysis of two major mechanisms: backbiting and scission. Considering knowledge gaps from the literature, the immediate research goal is to propose a kinetic model that not only predicts degradation rate constants but also the most likely mechanism of degradation as a function of pH, buffer concentration, and temperature.

Depending on temperature and pH, the degradation of prebiotic molecules can shift towards backbiting, scission, or both. Both backbiting and scission follow a clear Arrhenius relationship across a pH range of 2–8. Each process has a constant activation energy but different prefactors suggesting the kinetics of these mechanisms are driven primarily by entropic effects.

2.2 Introduction

To elucidate the mechanism and kinetics of degradation of depsipeptides, one must first look at the degradation of polyesters composed of multiple units of α-hydroxy acids. In the literature, two of the common α-hydroxy acids were lactic acid [19, 25–27] and malic acid [17]. Degradation is subdivided into two processes: scission and backbiting. Scission occurs at a random location along a homopolymer when water cleaves any ester moiety in a polyester at equal probability. Backbiting occurs when the hydroxyl terminus of a

polyester performs an intermolecular attack on the carbonyl ester, forming a cyclic diester and a molecule that is two degrees of polymerization less than the original starting material. For example, a trimer of polylactic acid can backbite to form a cyclic diester and a lactic acid monomer. Both mechanisms are outlined in Section 2.3. The literature hypothesizes that at acidic pH values, the mechanism of degradation is scission while at neutral and basic pH values, the mechanism of degradation is backbiting [19, 25].

De Jong and van Nostrum made these observations using mass spectrometry and high performance liquid chromatography (HPLC) [19, 25]. They discovered that backbiting always reduced a longer polyester by two degrees of polymerization, while scission reduces the degree of polymerization by a random amount since the ester cleavage can happen anywhere. Another observation is that backbiting requires a hydroxyl terminus in the compound. If there is a protecting group in the compound, scission is the only process that can occur initially. However, scission exposes a hydroxyl terminus and backbiting is the dominant mechanism again [19]. At high pH conditions for polyesters, particularly polylactic acid, backbiting is seen to be at least four orders of magnitude faster than scission [19]. It has also been observed that degradation follows an Arrhenius relationship [25]. The question now becomes "Do depsipeptides degrade in a similar fashion?"

Little work has been done to elucidate the mechanism and kinetics of degradation of depsipeptides. Yu has provided a kinetic framework of the polymerization and degradation of a depsipeptide composed of the α-amino acid valine and α-hydroxy acid lactic acid [14]. He found that both the polymerization and the ester degradation rates followed an Arrhenius expression, but he did not consider the pH dependence on either polymerization or degradation. My research continues from where Yu left off but with a different "monomer", the HO-dipeptide of glycolic acid alanine (gA). Under the experimental conditions studied here, the amide bond between g and A is stable, such that gA is the unit of polymerization via ester bond formation.

My major research goal is to provide a computational and statistical approach towards

calculating degradation rate constants of depsipeptides and identifying pH dependency in these rate constants. To begin answering the question, a theoretical framework is needed to model the chemistry of the degradation of any depsipeptide with repeating monomeric units. The literature provides this framework in the form of a chain mass balance to model the polymerization and degradation processes [17, 28]. The mass balance models the polycondensation and degradation of malic acid and lactic acid oligomers. Polycondensation reactions combine two monomers and create a polymer, releasing water. This model is not useful for the research goal for two reasons. One reason is that the model does not consider backbiting and ring opening polymerization, which are critical processes in depsipeptide degradation and sequence space expansion. The second reason is that the reported scission rate constants do not show the pH dependence.

The literature [25, 41, 42] attempts to address the deficiencies of the chain mass balance model by proposing kinetic equations that depends on the catalytic activity of protons, hydroxyls, solvents, degrading species, and buffer species in acidic, neutral, and basic media. Although de Jong [25] addresses pH dependence of polymerization and hydrolysis of hydroxy acids to form polyesters, this model may or may not describe the desired behavior of the polymerization and hydrolysis of depsipeptides. Additionally, de Jong's equation does not account for possible buffer species contributions and multiple deprotonation equilibria of degrading species. De Jong [25] provides the starting point for the studies in Chapter 2 on ester bond degradation in depsipeptides.

2.2.1 Catalysis Types

After the most likely degradation mechanisms were determined and their rate constants were calculated, the rate constants were deconstructed to identify the correct catalysis type.

There are two main categories of catalysis: acid and base. These categories subdivide into two subcategories: general and specific. In specific acid (base) catalysis, rates of reaction depend only on the conjugate acids (bases) from the solvent. Since water is the

solvent in this case, hydronium ions (acid) or hydroxide ions (base) are considered. In general acid (base) catalysis, rates of reaction depend on all acids (or bases) in solution. The solution includes both the solvent and the buffer. According to Jelinska [41], the degradation rate constants have a linear relationship (Equation 2.1) with the total buffer concentration. The y-intercept is the rate constant associated with the solvent catalytic activity (k_{pH}) and the slope is the rate constant associated with the buffer catalytic activity (k_B). By performing linear regression on experimental data when pH and temperature remain constant while varying buffer concentrations, these two parameters can be obtained and thus be used to determine the type of catalysis. In this experiment, phosphate buffers of varying concentrations (75 mM, 100 mM, 150 mM) were used with or without salt. The salt affects ionic strength of the buffer.

The salt concentration is such that the sum of the buffer concentration and salt concentration is 150 mM. For example, if the buffer concentration is 100 mM, the salt concentration is 50 mM. One can notice that if the buffer concentration is 150 mM, the salt concentration is 0 mM, meaning that the observed rate constant would be the same with or without salt added.

$$k_{obs} = k_{pH} + k_B[B] \tag{2.1}$$

where k_{obs} is the overall degradation rate constant at T = 65°C and $[B]$ is the buffer concentration.

Figure 2.1 shows these two parameters along with linear fits of real experimental data. The observed rate constants were directly from HPLC trace data from gA oligomer degradation to form gA monomer. Dr. Moran Frenkel-Pinter performed these experiments in triplicate.

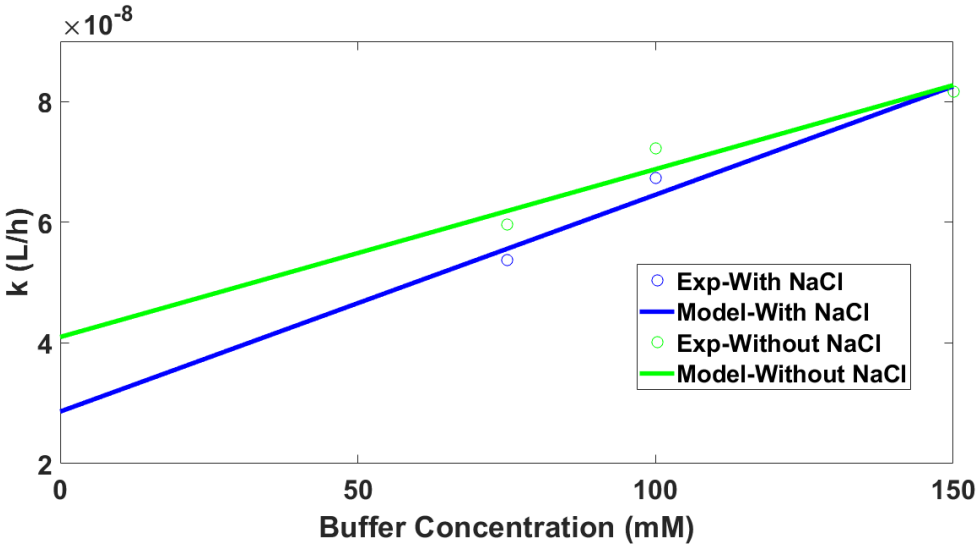

Figure 2.1: Linear fits of overall gA oligomer degradation rate constants at $65°$C with and without NaCl. The phosphate buffer concentrations are 75 mM, 100 mM, and 150 mM. The experimental data point with NaCl at a 150 mM buffer concentration is overlapping the one without NaCl at the same buffer concentration.

Table 2.1 displays the rate constants k_{pH} and k_B gleaned from linear regression of the experimental data and their 90% confidence intervals. According to Table 2.1, k_{pH} and k_B have the same order of magnitude whether salt is in the buffer or not. This observation strongly suggests ionic strength plays a minor role in degradation. Since k_{pH} accounts for approximately half of the k_{obs}, the dominant type of catalysis is ambiguous.

Table 2.1: Upper and Lower Bounds of k_{pH} and k_B in addition to the actual value. Units of k_{pH} and k_B are L/h and $\frac{L}{h \times mM}$, respectively. The bounds are 90% confidence intervals.

Parameter(Condition)	Lower Bound ($\times 10^{-10}$)	Upper Bound ($\times 10^{-10}$)	Actual Parameter ($\times 10^{-10}$)
k_{pH} (With NaCl)	273	300	286
k_B (With NaCl)	3.47	3.72	3.59
k_{pH} (Without NaCl)	396	424	410
k_B (Without NaCl)	2.66	2.91	2.78

Knowing the catalysis type has significant mechanistic implications [43]. For example, in general catalysis, proton transfer from a reactant to its conjugate acid is part of the rate-determining step. However, the equilibrium is dictated by the pH of solution only in specific catalysis. Proton transfers are not part of the rate-determining step.

2.2.2 Computational Goals and Polymerization Obstacles

The aim is to model equations that simulate backbiting and scission of gA oligomer degradation. An error analysis on the simulation results is also performed. A kinetic model is developed and implemented to simulate depsipeptide degradation. Tackling this aim will help elucidate clearly what degradation mechanism is most dominant and how well these modified models simulate depsipeptide chemistry. Polymerization was not modeled due to the difficulty of taking into account volume changes and the dry state in the modeling. Another factor is the availability of quantitated data for polymerization.

Kinetic modeling of polymerization is a good topic for future research because when combined with thermodynamics, the driving forces behind other similar phenomenon in pharmaceutics become significantly clearer. One also gains a better understanding of the stability of the desired product and what conditions are necessary to obtain that product. As discussed previously, both polymerization and degradation are necessary to expand the sequence space of peptides in the prebiotic Earth, so deep knowledge of these two processes is critical.

Another application of this research is green chemistry because.performing chemistry in water is green. Green chemistry refers to the sustainable and environmental friendly methods in producing products useful to society. According to the literature, synthetic polymers such as polyethylene and polystyrene have attractive properties such as low density and good tensile strength. However, their inability to biodegrade has become increasingly unattractive in the field of green chemistry [28]. PLA, on the other hand, is biodegradable and can be industrially produced through direct polycondensation and ring opening polymerization [23, 28, 30, 44]. Another reason for green chemistry is the conversion from organic solvents to cleaner solvents like water.

2.3 Chemistry of gA oligomers

2.3.1 Mechanism Descriptions Using gAgA as a Guide

Ester-Amide Exchange

To form a depsipeptide, a nucleophilic attack of the amino group on the carbonyl on diglycolic acid is followed by an ester-amide exchange. Then, a monomeric unit of glycolic acid leaves and finally reacts with the carboxylic acid moiety to form an ester bond, elongating the molecule. It could also reversibly give back the original glycolic acid alanine depsipeptide through hydrolysis of the ester bond. Ester-amide exchange is shown in Figure 2.2.

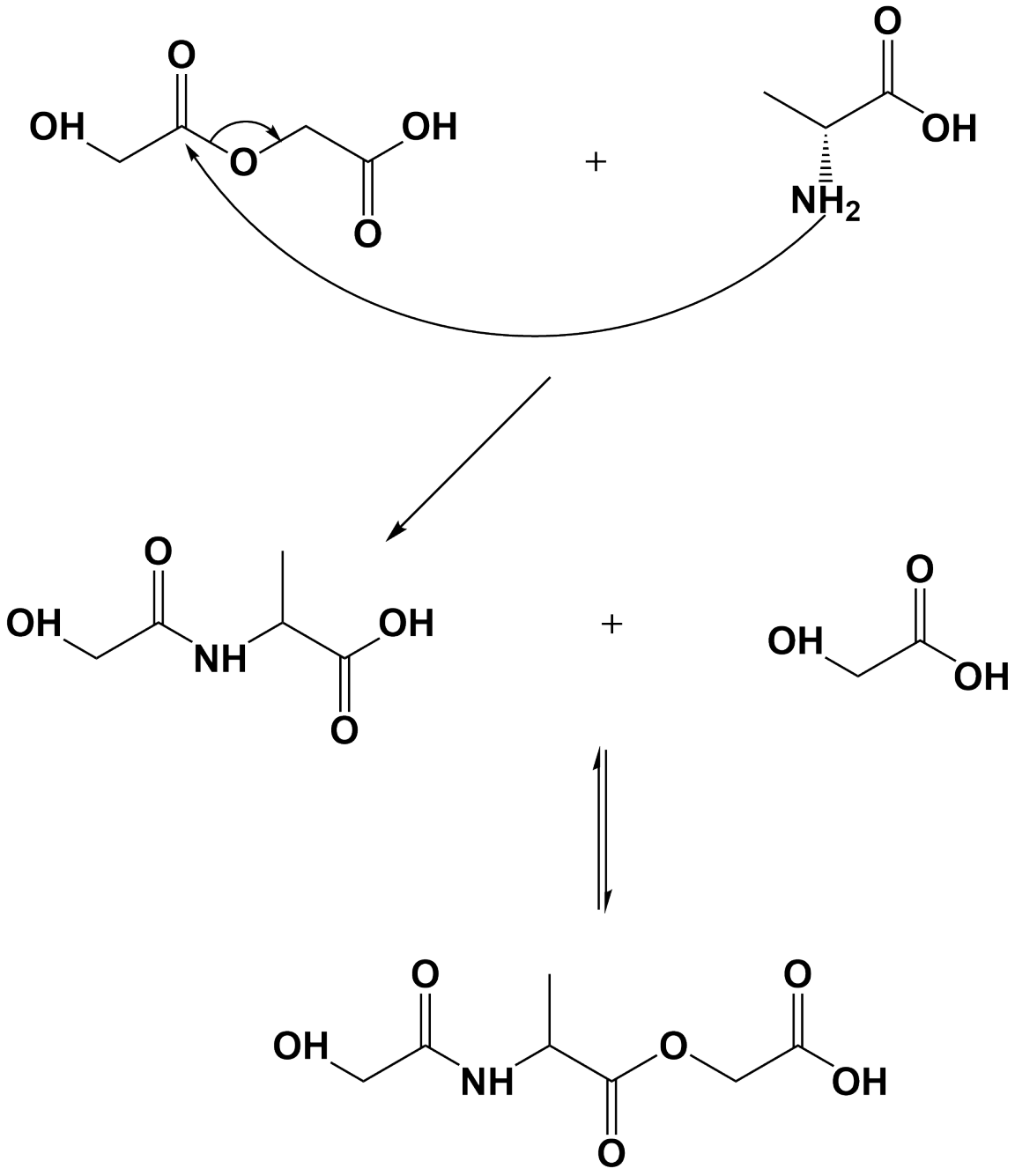

Figure 2.2: The most prevailing mechanism of the formation of a glycolic acid alanine dimer: ester-amide exchange.

Acid-Catalyzed Ester Scission

This type of hydrolysis, shown in Figure 2.3 and Figure 2.4, undergoes a similar mechanism as the base-catalyzed version shown in Figure 2.5. Instead of hydroxyl ions, water

molecules attack the electrophilic carbon in the ester. However, water is not a strong enough nucleophile to attack the ester. Before the nucleophilic attack occurs, the ester must first be activated by protonating the oxygen atom on the carbonyl. Water attacks the ester forming a tetrahedral intermediate just like the base-catalyzed version.

Next, water attacks one of the hydrogen atoms in the water moiety in the intermediate thereby breaking an OH bond and pushing the electrons onto the oxygen atom on the water. This event stabilizes charge. The alkoxy group needs to leave, but it is not a good leaving group in this circumstance. The alkoxy group needs to be protonated first to form a depsipeptide oligomer of length n which is a good leaving group. A lone pair of electrons on the oxygen atom are pushed to form a double bond with the carbon reforming the carbonyl. This event forces the leaving group to detach from the intermediate. Finally, water attacks a hydrogen atom adjacent to the carbonyl oxygen atom breaking the OH bond and pushing the electrons back to the carbonyl oxygen atom stabilizing the charge and forming another depsipeptide oligomer of length k-n. The variable k represents the original degree of polymerization of a depsipeptide before hydrolyzing.

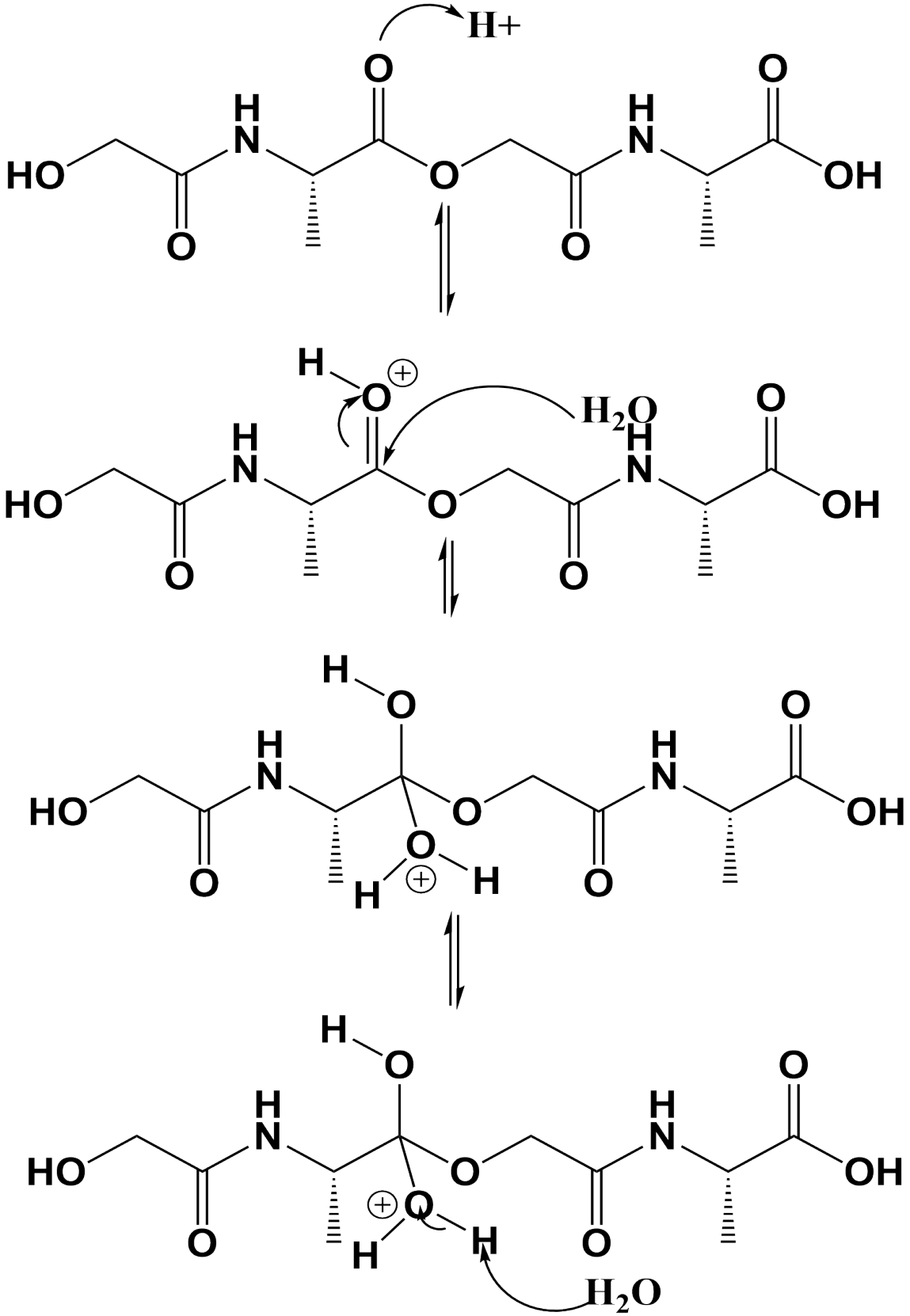

Figure 2.3: Part 1 of acid-catalyzed ester scission of a gAgA compound.

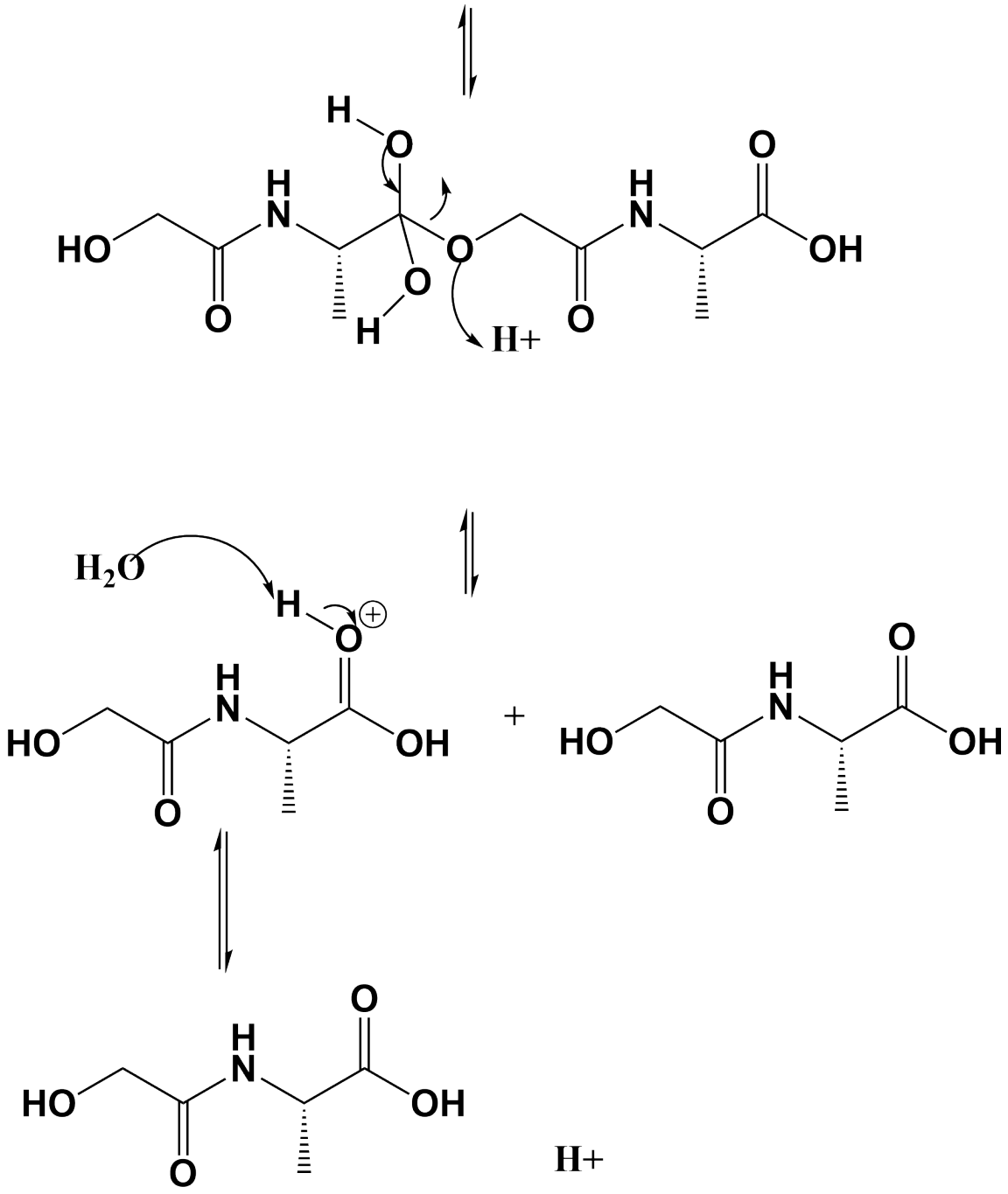

Figure 2.4: Part 2 of acid-catalyzed ester scission of a gAgA compound.

Base-Catalyzed Ester Scission

Hydroxyl ions are strong nucleophiles that attack an electrophilic carbon on the carbonyl group. This event breaks one of the double bonds in the carbonyl and pushes the electrons

on the oxygen atom as shown in Figure 2.5. A tetrahedral intermediate is formed. The extra lone pair of electrons on the oxygen atom goes back to forming a carbonyl making an alkoxy group leave. Then, the nucleophilic alkoxy group attacks a hydrogen atom on the carboxyl group breaking the OH bond and the electrons from the bond are pushed onto the oxygen atom in the carboxyl group. The two final products are a depsipeptide of degree of polymerization n and a conjugate base of a depsipeptide of length $k - n$.

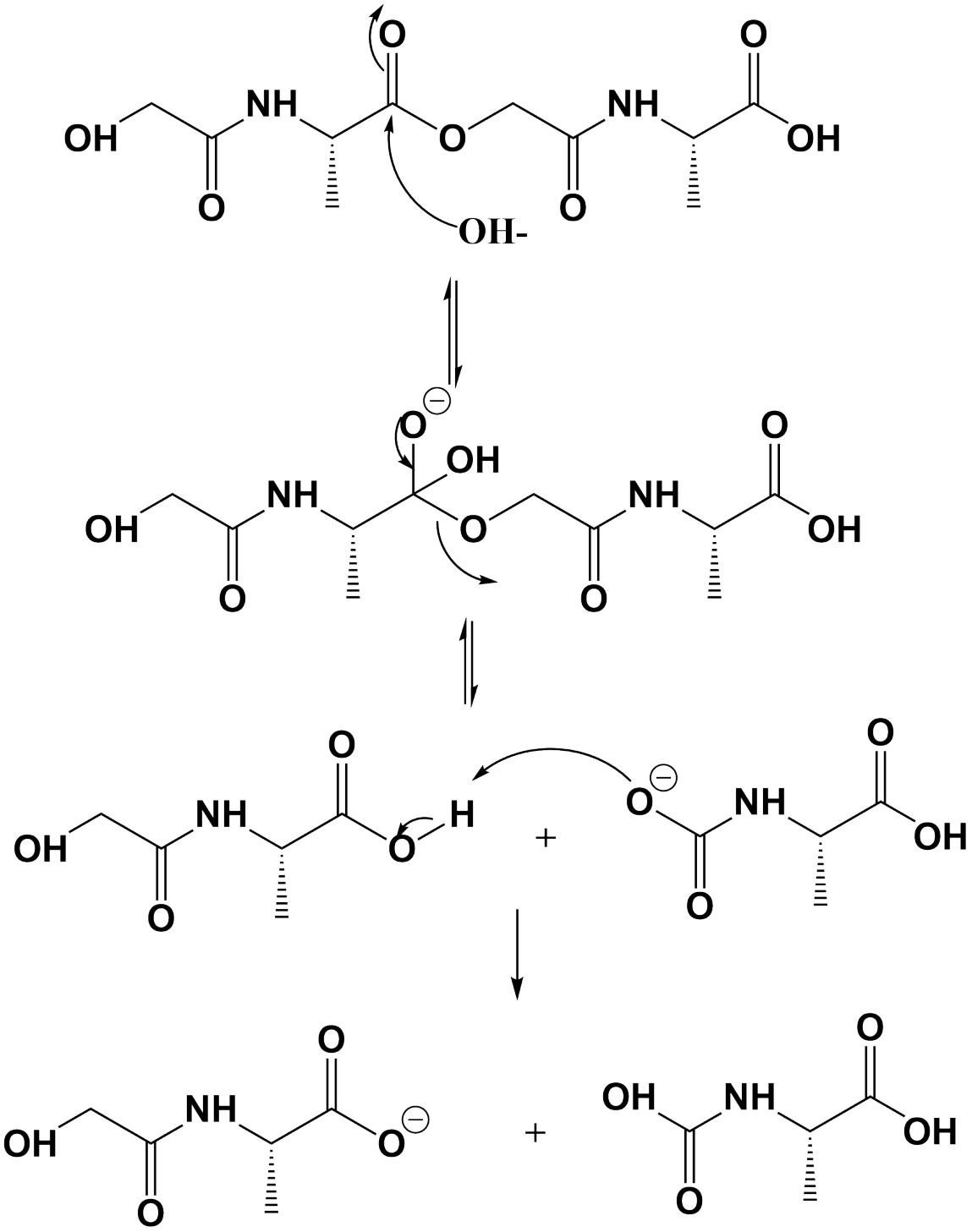

Figure 2.5: Base-catalyzed ester scission of a gAgA compound.

Backbiting

Backbiting (shown in Figure 2.6) is a base-catalyzed process. Hydroxyl ions attack a hydrogen atom on a hydroxyl terminus of a depsipeptide forming a water-bonded molecule. Next, the bond between the water and oxygen atom breaks and performs an intermolecular attack with an electrophilic carbon on the carbonyl forming a tetrahedral intermediate containing a six-member ring. This event also breaks the double bond and pushes the electrons to the oxygen atom. The extra lone pair of electrons are pushed back reforming the double bond thereby pushing off an alkoxy leaving group. The remaining products are a conjugate base of a depsipeptide of length $k-2$ and a cyclic ester called a morpholine-2,5-dione. Because cyclic esters easily ring open, the morpholine-2,5-dione can open to form a depsipeptide of length two.

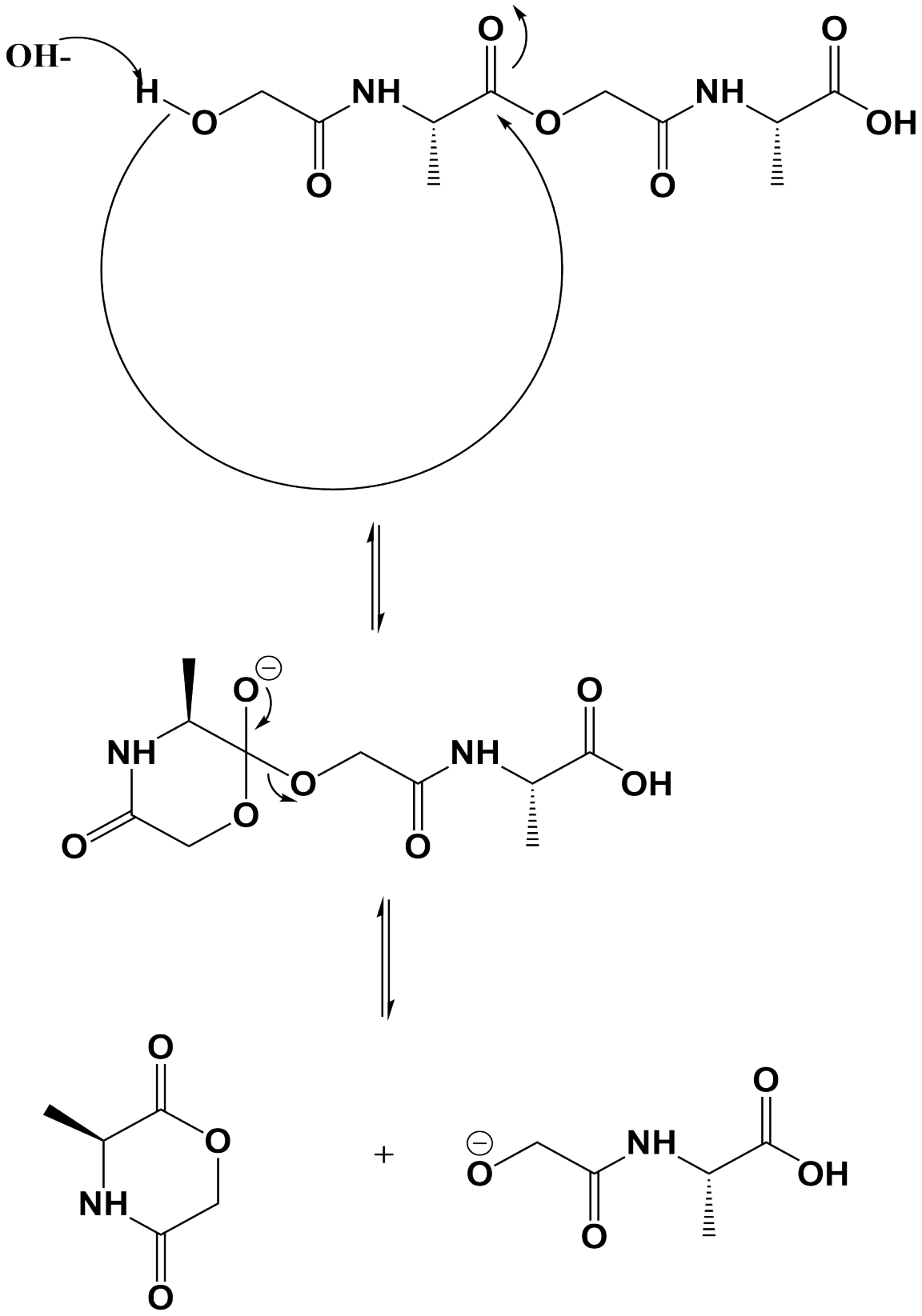

Figure 2.6: Backbiting of a gAgA oligomer

2.4 Kinetic Modeling

2.4.1 Mass Balance

The kinetics of the degradation of gA is based on the model in Ref. [17, 28] outlined below:

$$\frac{dz_i}{dt} = \frac{1}{V}\left(2k_{sc}\sum_{j=i+1}^{n} z_j \; - \; k_{sc}\,(i-1)\,z_i + k_{bb}(z_{i+1} - z_i\right) \; (i = 1, ...n) \tag{2.2}$$

$$\frac{dW}{dt} = \; -k_{sc}\sum_{i=1}^{n}(i-1)\,z_i - \; K_p x_w \; - k_{bb}\sum_{j=2}^{n} z_j \tag{2.3}$$

$$\frac{dz_n}{dt} = \; \frac{1}{V}\left(-k_{sc}\,(n-1)\,z_n - \; k_{bb} z_n\right) \tag{2.4}$$

$$\frac{dz_1}{dt} = \; \frac{1}{V}\left(2k_{sc}\sum_{j=2}^{n} z_j \; + \; k_{bb}\left(\sum_{j=2}^{n} z_j + z_2\right)\right) \tag{2.5}$$

where z_i is the number of moles of i^{th}-mer of gA; t is time; V is the total volume of the system; k_{sc} is the scission rate constant; W is the number of moles of water; K_p is the mass transfer coefficient which includes pressure terms; x_w is the mole fraction of water; k_{bb} is the backbiting rate constant; and n is the highest oligomer length that is modeled. The model is not expressed in concentrations since it is assumed that volume is constant. To enable future modeling of dry-down reactions with changing V, one could use heat and mass transfer correlations from [45].

These differential equations form a mass balance of various oligomer chain lengths. For example, a dimer can form from polymerization of two monomers. It can also form from hydrolysis of a trimer into a dimer and monomer. There are many other ways to form $P-2$ oligomers from higher oligomer chain lengths. In Equation 2.2, variable i can range from two to $n-1$. The experiments were designed so that this approximation is valid, such

that n is much greater than the longest observed oligomer.

The scission model assumes that all scission events happen with equal probability. For example, a pentamer can perform scission to a tetramer and a monomer just as easily as to a trimer and dimer.

The kinetics of the degradation of a homopolymer of chain length two (dimer) simplify to the following:

$$\frac{dz_1}{dt} = \frac{1}{V}\left(2k_{sc}z_2 + 2k_{bb}z_2\right) \tag{2.6}$$

$$\frac{dz_2}{dt} = \frac{1}{V}\left(-k_{sc}z_2 - k_{bb}z_2\right) \tag{2.7}$$

$$\frac{dW}{dt} = -\left(k_{sc} + k_{bb}\right) z_2 \tag{2.8}$$

If $n = 3$ (trimer), the equations simplify to:

$$\frac{dz_1}{dt} = \frac{1}{V}(2k_{rs}(z_2 + z_3) + k_{bb}(2z_2 + z_3)) \tag{2.9}$$

$$\frac{dz_2}{dt} = \frac{1}{V}\left(2k_{sc}z_3 - k_{sc}z_2 + k_{bb}(z_3 - z_2)\right) \tag{2.10}$$

$$\frac{dz_3}{dt} = \frac{1}{V}\left(-2k_{sc}z_3 - k_{bb}z_3\right) \tag{2.11}$$

$$\frac{dW}{dt} = -k_{rs}\left(z_2 + 2z_3\right) - k_{bb}(z_2 + z_3) \tag{2.12}$$

The pattern repeats for higher oligomers.

Key issues affecting the use of this model are as follows: First, I will have a system of $n + 1$ differential equations to solve which become more computationally expensive as n

increases. Additionally, the monomer ($n = 1$) includes not only the linear homopolymer but also the cyclic gA known as morpholine-2,5-dione (seen in Figure 2.7).

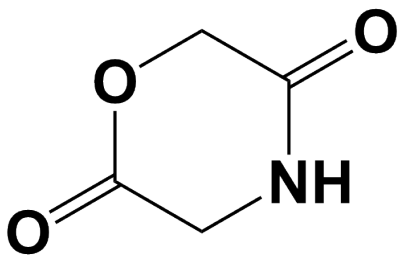

Figure 2.7: Morpholine-2,5-dione

The reason why morpholine-2,5-dione is included in the monomer differential equation is the assumption that the ring opening of morpholine-2,5-dione to a linear monomer gA molecule is sufficiently rapid in water that it is not the rate limiting step of the degradation. Finally, the model cannot differentiate between backbiting and scission for the dimer specifically. The rate constant it predicts will be the sum of the two rate constants. This issue is addressed in the "(PA)AgA and gAgA" subsection.

MATLAB's patternsearch function was used to estimate the rate constants that minimize the sum squared error (SSE). Table 2.4 shows the optimized rate constants. The patternsearch function is a global optimization package that selects several initial values for each differential equation and finds the one that minimizes the error. This technique increases the likelihood that the global optimum was found and not just a local optimum.

2.4.2 How Experimental Data is Collected for the Model

Now that the kinetic model for gA degradation is presented, experimental data needs to be acquired to fit it. The high performance liquid chromatography (HPLC) data for a variety of pH values, temperatures, and time points is recorded in .csv files. Then, I use a data analysis program called Igor to read these .csv files, and it outputs gA oligomer peaks at unique retention times. Using the Peak Integration tool within Igor, I determine the areas under these peaks using multiple Gaussian fits. These areas give rise to concentrations

according to Beer-Lambert's Law:

$$A = \epsilon bc \tag{2.13}$$

where A (unitless) is absorbance (which in this case is the integrated peak area); ϵ is the molar extinction coefficient in $\frac{1}{M \times cm}$; b is the path length in cm; and c is the concentration in M.

Then, the kinetic model (Equations 2.2-2.5) was solved in MATLAB using ode15s to create a degradation profile of gA oligomers and the evolution of the monomer, and the concentrations gleaned from Igor were fit using MATLAB.

2.4.3 Experimental Conditions tested (gA Oligomer Mixture/Distribution)

- pH: 2–8. For the phosphate buffer, the phosphate concentration is 100mM. Here is the recipe for the phosphate buffer at a pH range between 6–8. From pH 2–5, citric acid buffers were used, but only pH 3.0 and pH 5.0 are known.
  - pH 3.0: the final concentration of Na2HPO4 was 20.55mM and citric acid was 39.725mM
  - pH 5.0: the final concentration of Na2HPO4 was 51.5mM and citric acid was 24.425mM
  - pH 6.0, the final concentration of Na_2HPO_4 is 12mM and for NaH_2PO_4 is 88mM
  - pH 7.0, the final concentration of Na_2HPO_4 is 57.7mM and for NaH_2PO_4 is 42.3mM
  - pH 8.0, the final concentration of Na_2HPO_4 is 93.2mM and for NaH_2PO_4 is 6.8mM
- NaCl is added, and all sodium ions come from the phosphate salts. The Na^+ concen-

tration is different depending on the pH, and should be calculated as follows: At a pH = 6, the phosphate concentration was 100mM and the sodium concentration was 112mM (88mM $\times$ 1 (only one sodium atom in monobasic phosphate salt) + 12mM $\times$ 2 (two sodium atoms in dibasic phosphate salt).

- Temperature (°C): 25, 37, 50, 65
- Time Scale: varies depending on the rate of gA degradation
- Each pH-temperature combination was tested at least 3 times. Vials containing aqueous solution of gA compounds were capped. The volume is 40 μL and the total concentration is 10 mM gA monomer. HPLC data taken after each hydrolysis is completed. The experiments were performed by Dr. Moran Frenkel-Pinter, reported in Ref. [46].

After the HPLC data was taken, the Peak Integration tool in Igor and Beer Lambert's Law (Equation 2.13) were used to calculate concentrations of oligomers at each time point. Each of the 3 replicates contains 7 timepoints and 7 types of oligomer (monomer-heptamer) making a 7x7 matrix of data. Then, the 3 replicates were averaged and the standard deviation of the replicates was calculated resulting in two matrices: a matrix with averaged values and a matrix containing standard deviations.

(PA)AgA and gAgA

Scission and backbiting mechanisms of gA oligomer degradation were discussed. However, there is one special type of gA oligomer that must be addressed: the gAgA dimer. If one looks at Equation 2.6 – Equation 2.8, both k_{bb} and k_{sc} cannot be found simultaneously with the model because both mechanisms are not independently separated and only the sum of the mechanisms would be found.

To remedy this problem, one of these mechanisms must be suppressed. Since backbiting can only occur with an OH terminus, backbiting can be suppressed by blocking this

functional group. The molecule (PA)AgA, or propionic acid alanine glycolic acid alanine, is synthesized (by Dr. Luke Leman) for this purpose. (PA)AgA is a gA dimer derivative where the only difference is that a methyl group replaces the OH terminus. The same modeling exercise is performed with (PA)AgA and the model predicts the scission rate constant while setting the backbiting rate constant to zero because backbiting theoretically should not happen in (PA)AgA.

Next, the model predicts the backbiting rate constants by simulating gAgA's kinetic degradation. The scission rate constant for gAgA is the same as that for (PA)AgA. The results are found in Figure 2.8. It shows the natural logarithm of the predicted rate constants for (PA)AgA vs. pH at a variety of temperatures. Note the U-shaped behavior suggesting minimum scission at around a pH of 3–4, and the same type of behavior is found in the literature for lactic acid [7].

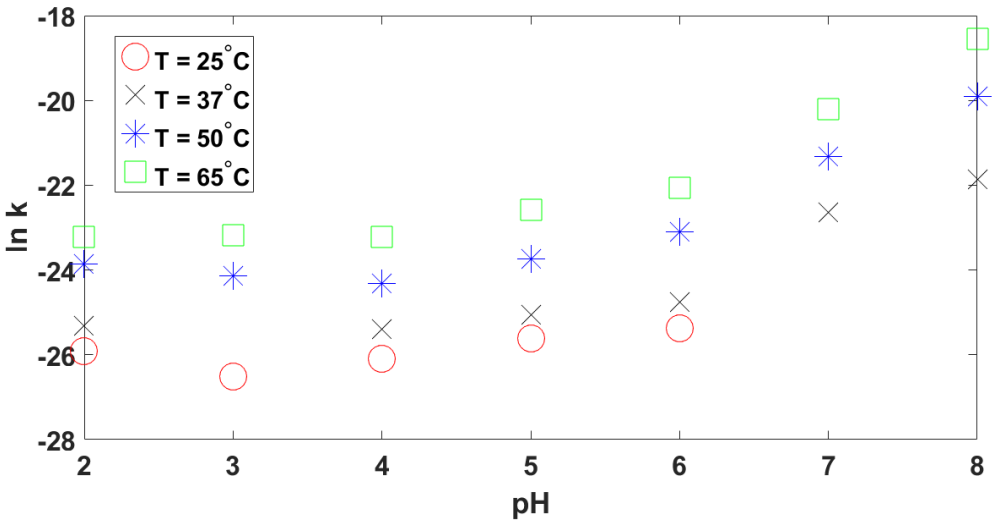

Figure 2.8: k-pH Curve of (PA)AgA scission rate constants. Some dots are missing because no experiments have been done at those conditions. Circle: 25°C; X: 37°C ; Stars: 50°C; Squares: 65°C

Figure 2.9 shows the retention times and absorbances of the gA monomer, gAgA dimer, (PA)AgA, and the cyclic gA. From Equation 2.13, the higher the absorbance, the higher the concentration of species in solution. This figure demonstrates the kinetics of the gAgA

degradation into monomer as time progresses.

Figure 2.10 shows backbiting and hydrolysis (scission) of gAgA and (PA)AgA. Backbiting cannot occur with a methylated protecting group on the (PA)AgA, but hydrolysis can expose a hydroxyl group that can undergo further backbiting.

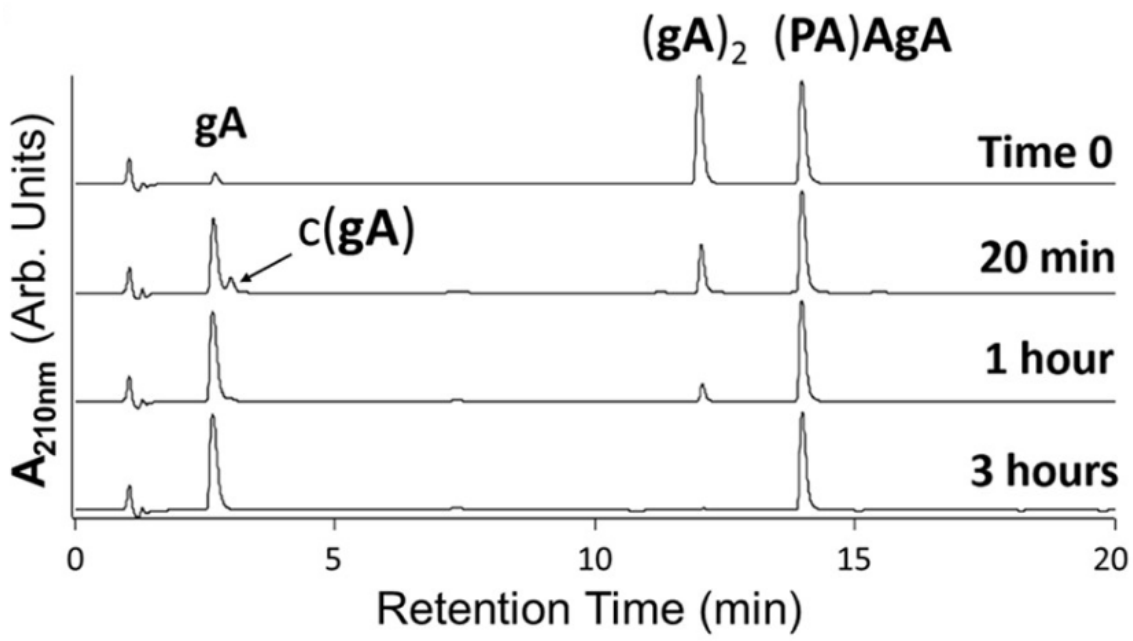

Figure 2.9: Standards $(gA)_2$ and (PA)AgA subject to pH 6, 65°C aqueous incubation. Image taken from Ref. [46] with permission.

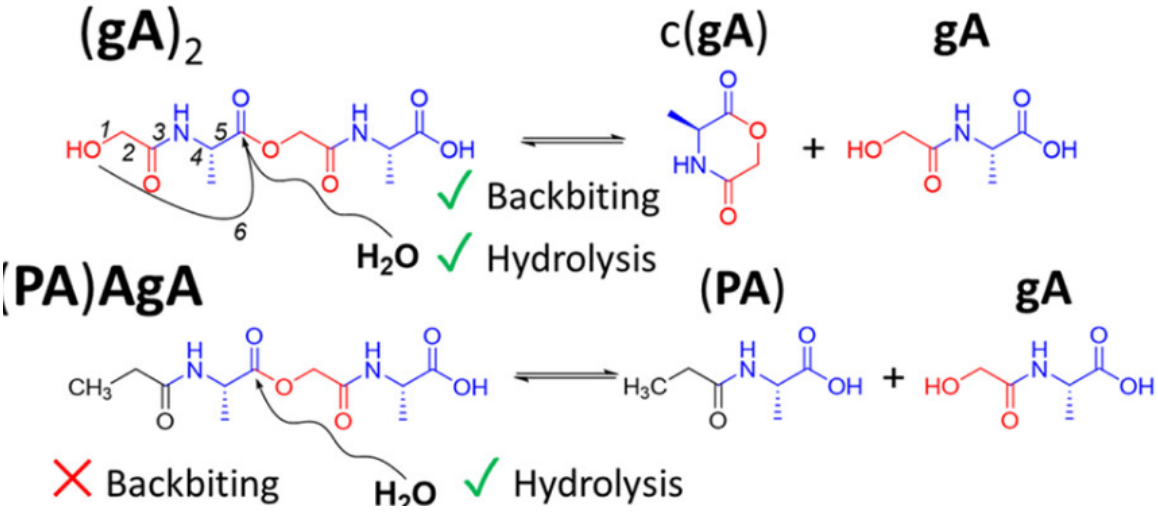

Figure 2.10: Standards $(gA)_2$ and (PA)AgA proceeding through plausible degradation routes. Image taken from Ref. [46] with permission.

2.4.4 Sample Model Fits for gAgA and (PA)AgA

Figure 2.11 shows the optimized fits of experimental data for gAgA and (PA)AgA. This figure is the best possible fit when optimizing the scission rate constant because scission is the only relevant degradation mechanism in (PA)AgA. Figure 2.11 also demonstrates the model fit of the most optimal combination of both scission and backbiting in gAgA. The scission rate constant from the (PA)AgA is the same as the scission rate constant as gAgA. The experimental data does not satisfy the mass balance, but the model does.

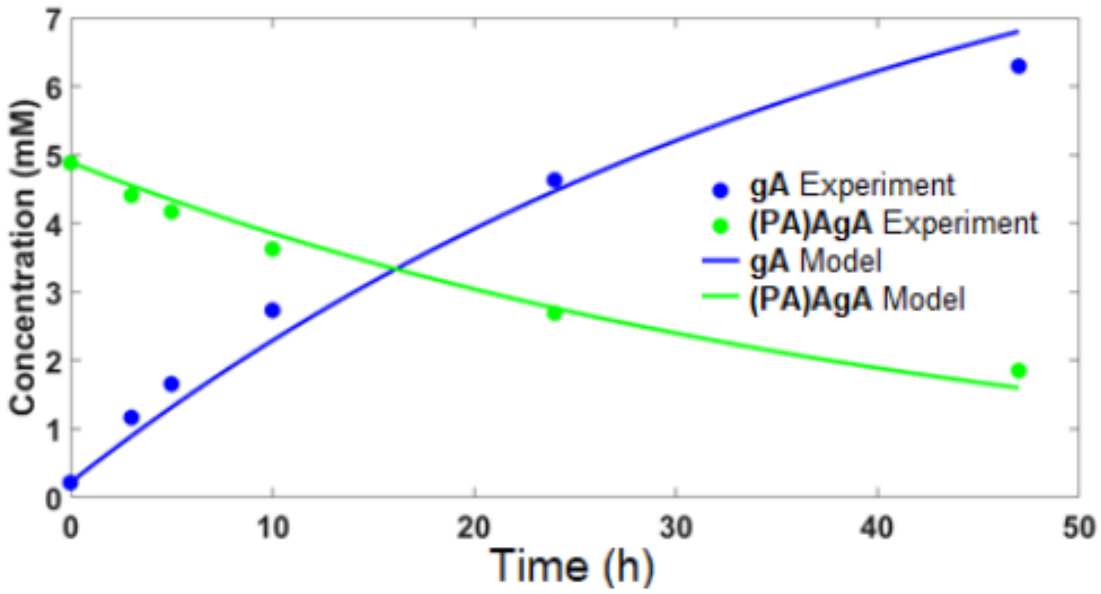

Figure 2.11: Model Fit of gAgA and (PA)AgA Degradation at pH = 6, T = 65°C

Figure 2.12 shows strong Arrhenius behavior of gAgA and (PA)AgA degradation for several pH values. The similar slopes suggest the same activation energy (difference in transition state enthalpy and ground state enthalpy) for all pH values but the different y-intercepts suggest highly different prefactors. Arrhenius prefactors are directly related to the frequency of molecular collisions. Higher numbers of collisions increase the molecular motion of a system which expands the number of possible microstates a system can take. A microstate is simply one way that molecules can rearrange themselves. From Boltzmann's

equation ($S = k_B ln(\Omega)$), which relates entropy (S) and the number of microstates in the system (Ω), the more microstates a system can take, the higher the entropy. Boltzmann's constant is represented by k_B. Figure 2.12 also clearly shows an Arrhenius relationship of the optimum (PA)AgA constants. To clearly see how the prefactors directly suggest strong entropic effects, one can expand the Arrhenius equation to the Eyring equation:

$$k = \frac{k_B T}{h} e^{\frac{\Delta S}{R}} e^{-\frac{\Delta H}{RT}} \tag{2.14}$$

where k is the rate constant, T is the absolute temperature, h is Planck's constant, ΔS is the entropy of activation, ΔH is the enthalpy of activation (also known as activation energy), and R is the universal gas constant.

One notices that the prefactor known as A in the Arrhenius equation is directly related to the entropy of activation in the Eyring equation justifying the conclusion that different prefactors are related to entropic effects.

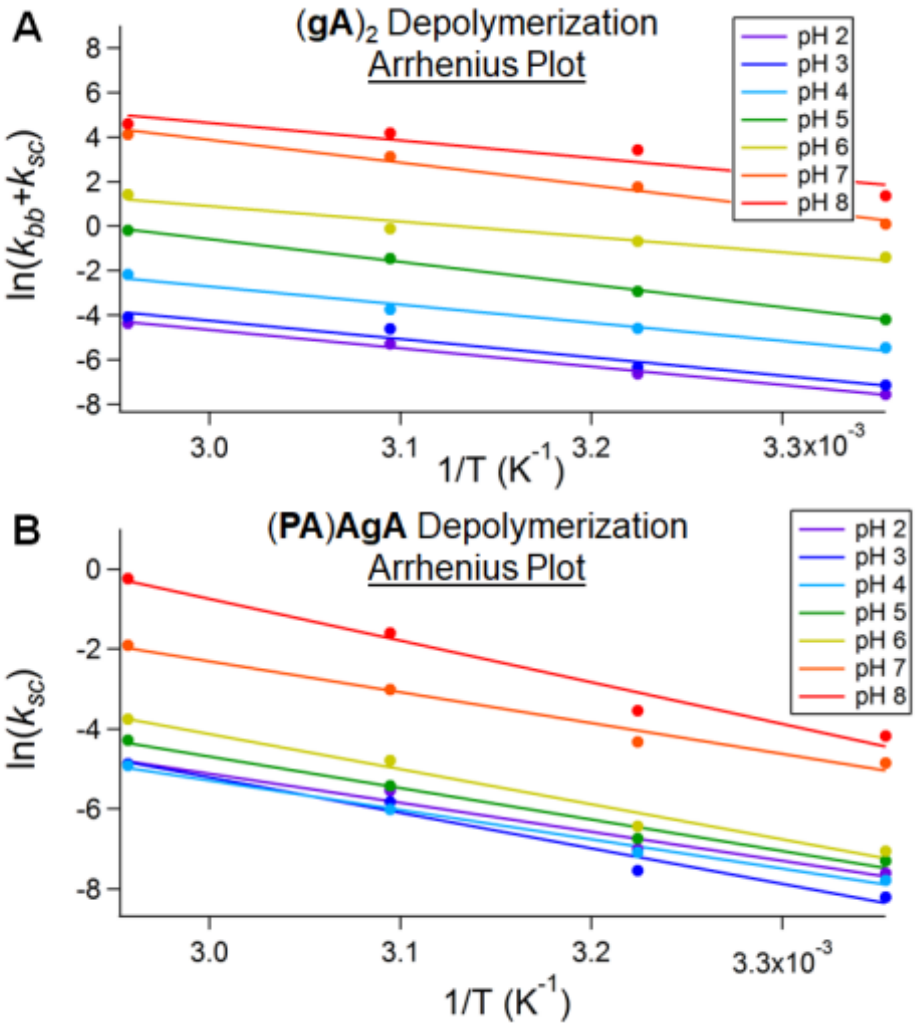

Figure 2.12: Arrhenius Plot for (PA)AgA Degradation Rate Constants at 4 different pH values; Lines are the Arrhenius Model. The terms depolymerization and degradation are interchangeable.

2.4.5 Sample Model Fits for gA Oligomers

Figure 2.13, Figure 2.14 and Figure 2.15 show the experimental data and model fits for scission only, backbiting only, and both, respectively at an experimental condition of pH = 7 and T = 65°C. Dr. Moran Frenkel-Pinter determined from HPLC that the initial total concentration of gA oligomers was 10 mM. As time elapses, the longer oligomers start to degrade until mostly monomer remains. However, some material was lost during the experiment, which explains why the data do not follow a mass balance but the models do.

Two interesting features exist with the gA oligomer fits. In these three cases, the dimer model predicts slightly increasing concentrations at the beginning then decreasing slightly afterwards during scission while the dimer model predicts monotonically decreasing concentrations during backbiting. The second interesting feature is how well the backbiting mechanism qualitatively describes the data better than scission. Both features suggest that backbiting is dominant at a pH of 6–8.

This conclusion is quantitatively justified in Section 2.5.

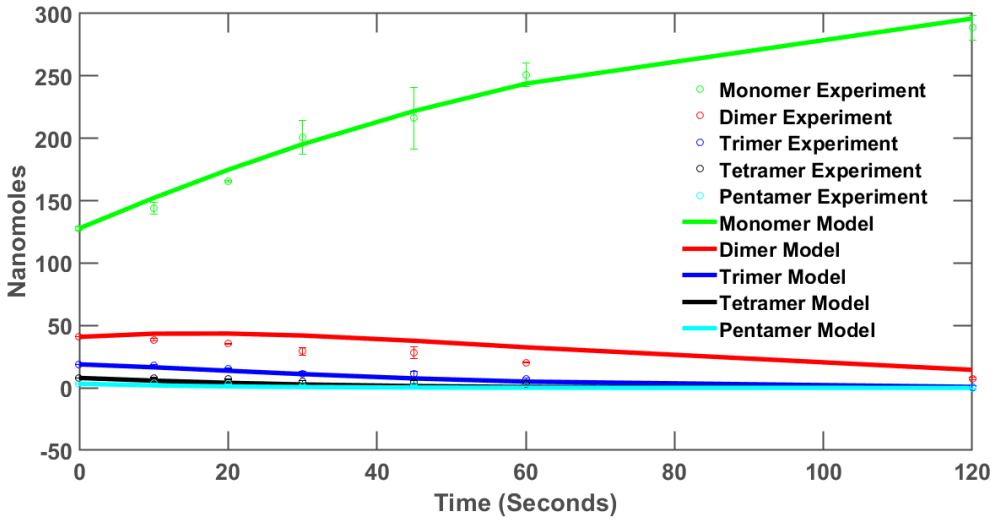

Figure 2.13: Scission Fits of gA Oligomer Distribution at pH = 7, T = 65°C. The dots are the experimental data while the lines are the model predictions. The bars represent the sample standard deviation of the amount of oligomer since the experiments were done in triplicate.

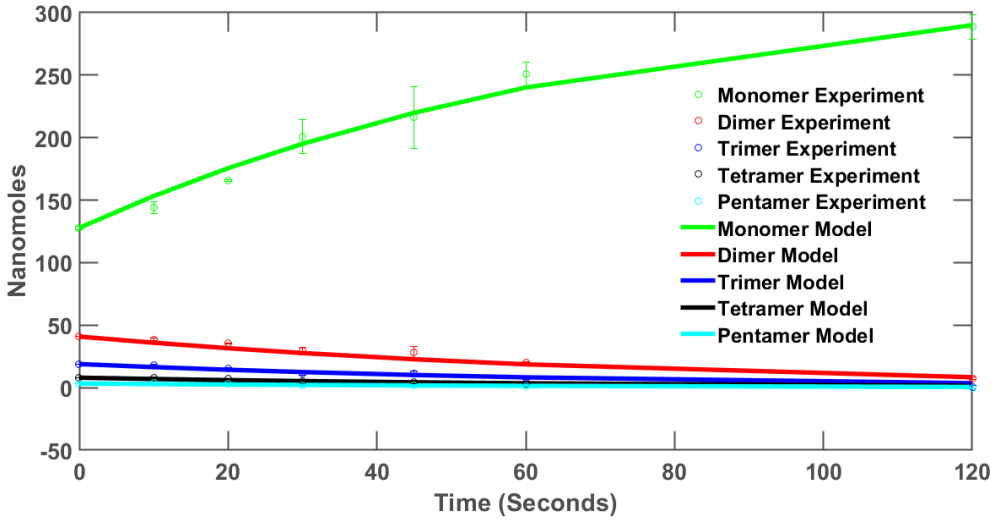

Figure 2.14: Backbiting Fits Rate Constants of gA Oligomer Distribution at pH = 7, T = 65°C. The dots are the experimental data while the lines are the model predictions. The bars represent the sample standard deviation of the amount of oligomer since the experiments were done in triplicate.

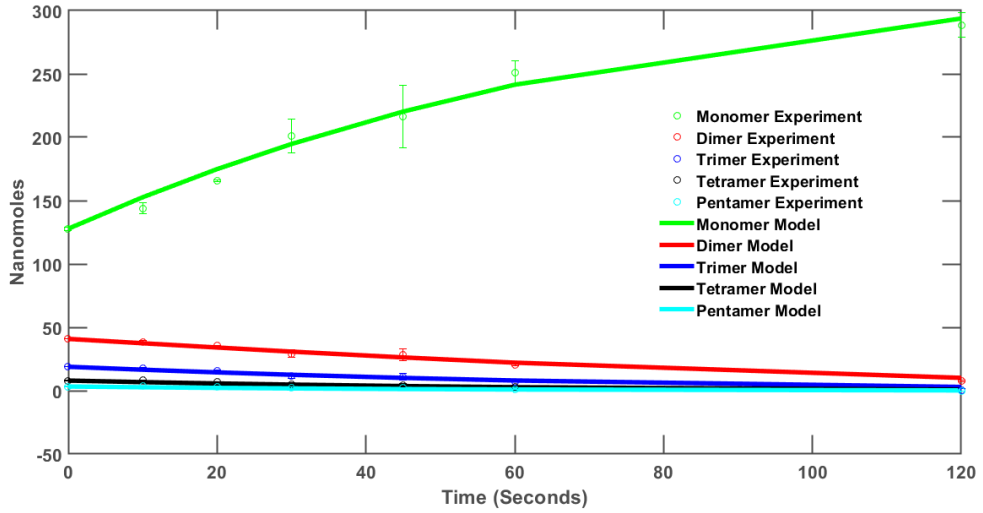

Figure 2.15: Both Mechanisms Best Fit Rate Constants of gA Oligomer Distribution at pH = 7, T = 65°C. The dots are the experimental data while the lines are the model predictions. The bars represent the sample standard deviation of the amount of oligomer since the experiments were done in triplicate.

2.4.6 Rate Constants for gAgA and (PA)AgA Oligomers

Table 2.2 shows the degradation rate constants and half-lives for gAgA and (PA)AgA. As pH increases, the half-lives for both compounds decrease while the rate constants increase suggesting higher instability. These compounds are highly susceptible to degradation at neutral to alkaline pH values.

Table 2.2: Kinetic Degradation Data for $(gA)_2$ and (PA)AgA for a range of conditions. Adapted from [46] with permission

$(gA)_2$ Half-Lives (h)							
$T(°C)$	pH: 2	3	4	5	6	7	8
25	674	647	146	43.2	2.78	0.63	0.18
37	308	298	62.6	12.6	1.34	0.12	0.02
50	76.8	53.5	26.1	2.91	0.76	0.03	0.01
65	34.4	27.9	5.63	0.82	0.17	0.01	0.01
(PA)AgA Half-Lives (h)							
$T(°C)$	pH: 2	3	4	5	6	7	8
25	1394	2528	1655	1027	806	88.6	45.0
37	758	1300	826	583	432	51.9	23.9
50	175	232	281	157	83.2	14.1	3.43
65	93.3	90.6	94.0	49.8	29.4	4.63	0.88
$(gA)_2$ Depolymerization, ($k_{sc} + k_{bb}$)$[h^{-1}]$							
$T(°C)$	pH: 2	3	4	5	6	7	8
25	0.001	0.001	0.005	0.016	0.249	1.11	3.96
37	0.002	0.002	0.011	0.055	0.516	5.86	31.1
50	0.009	0.013	0.027	0.238	0.914	22.7	66.4
65	0.020	0.025	0.123	0.843	4.21	62.3	100
(PA)AgA Depolymerization, (k_{sc})$[h^{-1}x10^{-3}]$							
$T(°C)$	pH: 2	3	4	5	6	7	8
25	0.497	0.27	0.419	0.675	0.86	7.83	15.4
37	0.914	0.533	0.839	1.19	1.60	13.4	29.0
50	3.95	2.98	2.47	4.42	8.33	49.2	202
65	7.43	7.64	7.38	13.9	23.5	150	792

2.5 Statistical Analysis of Model Fits

Statistical analysis on these fits is crucial since they answer the most important scientific questions in research in the field of modeling and prediction:

- What is the error associated with each data point in Figure 2.13, Figure 2.14 and Figure 2.15?
- What does that error really mean?
- What is the confidence level of the rate constants obtained from the model fits?
- How can one tell that backbiting is the most appropriate mechanism of gA degradation at a particular pH value and temperature in the first place?

Each of these questions will be addressed one at a time.

2.5.1 Akaike Information Criterion (AIC)

There are three candidates for the most likely mechanism of gA oligomer degradation: scission, backbiting, and both. Each of these candidates is associated with a model fit, and it is desirable to know which candidate is the most appropriate. One way to determine which candidate to use is through a criterion called the Akaike Information Criterion (AIC). The AIC gives a systematic way of comparing the models with each other to see which one is the best among the three and is a function of the sum squared error, the number of observations, and the number of parameters. The AIC, by itself, is not conducive towards small sample sizes (usually < 30) because it will pick the model with the most parameters which is not always desirable, so a correction needs to be made to account for them. This correction is known as the AIC_c, and the formula is given in Equation 2.15.

$$\text{AIC}_c = 4K + nln\left(\frac{\text{SSE}}{n}\right) + \left(\frac{2K\left(K+1\right)}{n-K-1}\right) \tag{2.15}$$

where K is the number of parameters, n is the sample size, and SSE is the sum squared error.

Consider the gA oligomer degradation system at T = 65°C and pH = 7. That system has three figures associated with it (Figures 2.12–2.14), each one representing scission only,

backbiting only, or both. The scission model is denoted Model 1; the backbiting model is denoted Model 2; the model with both mechanisms is denoted Model 3. Table 2.3 shows that the AIC_c values for Model 1, 2, and 3 are 149, 112, and 109, respectively. Model 3 exhibited the lowest AIC_c and the lowest SSE. However, it is not always the case that one model has both characteristics due to the number of parameters in the model. Even if the SSE is the lowest, the model may not always have the lowest AIC_c if it is overfitting the data with too many parameters.

After the AIC_c values were computed, each candidate model must be compared with each other rather than on an individual basis. For this reason, another metric is calculated: the relative Akaike weight of each model denoted w_i.

The relative Akaike weights tell how much more likely one model describes a set of data compared to another. To determine the Akaike weights, use the following equation:

$$w_i = \frac{e^{-0.5(AIC_{c,i} - min(AIC_{c,i}))}}{\sum_{i=1}^{N} e^{-}0.5(AIC_{c,i} - min(AIC_{c,i}))} \tag{2.16}$$

The lowest AIC_c value was subtracted from the other two values to obtain Δi. In this example case Δ 1 = 40, Δ 2 = 3, and Δ 3 = 0. Each Δi value is multiplied by -0.5 and the exponential of this value is calculated. The resulting values are then added. The ratio of each result is taken and divided by the sum of the resulting values to evaluate the relative Akaike weights [47]. In Table 2.3, the relative Akaike weights for Models 1, 2, and 3 are 0, 0.18 and 0.82, respectively. To interpret these values, Model 3 is 0.82/0.18, or about 4.5 times more likely to describe this set of data than Model 2.

Table 2.3: Akaike results for three models for gA Degradation at T = 65°C and pH = 6. BB = Backbiting. SC = Scission. Both = Backbiting and Scission

Type of Model	*SSE*	*n*	*K*	AIC_c	Δi	exp(-0.5*Δi)	Sum	w_i
SC	942	49	1	149	40	0	1.23	0
BB	440	49	1	112	3	0.23		0.18
Both	380	49	2	109	0	1		0.82

The noteworthy column is the w_i column on the far right. The model with both mechanisms has a relative Akaike weight of 0.82. To interpret this value, one can say that using both mechanisms explains 82% of the experimental data in Figures 2.12–2.14. In Table 2.3, one can reject both random scission and both mechanism models under this criterion, leaving just backbiting.

2.5.2 Percentage Degradation

The goal of this section is to quantify the percentage of backbiting and scission of gA oligomer distributions at different temperatures and pH values. Quantifying the distributions gives an insight to how dominant one mechanism is over another for a particular temperature and pH.

In order to quantify the dominance of either backbiting or scission for a given experimental condition, it is necessary to calculate the rates and the number of bonds broken during backbiting or scission. Afterwards, the percentage of bonds broken during each of these processes can be calculated. This percentage represents the dominance of each process. Equation 2.17 shows the number of nanomoles of bonds broken during backbiting while Equation 2.18 shows the number of nanomoles of bonds broken during scission. Scission can take place at any length along the oligomer chain except the first length hence the $i-1$ correction. Based on the assumptions of the model, both variables k and V are constant.

$$\frac{k_{bb}}{V}\sum_{i=2}^{\infty}\int_{t=0}^{t=\,t_f} z_i dt \tag{2.17}$$

$$\frac{k_{rs}}{V}\sum_{i=2}^{\infty}\int_{t=0}^{t=\,t_f} z_i\,(i-1)\,dt \tag{2.18}$$

where k is the rate constant (either backbiting or scission); V is the volume; i is the length of the oligomer; t is time; t_f is the final time; and z is the number of nanomoles.

To solve both equations, the integrals must be evaluated. The method chosen to evaluate the integral is the trapezoid rule. This method approximates the integral by breaking the area underneath the function into many trapezoids whose areas are easily computed. The results are seen in Figure 2.16 and indicate that as pH increases, the mechanism for gA oligomers changes from scission, to both, to backbiting. The significance of this trend is that it corroborates the experimental findings in the literature suggesting that backbiting is most dominant at basic pH while scission is most dominant at acidic pH. One exception occurs at pH = 6, but it is an open question to determine why and if this is a robust finding.

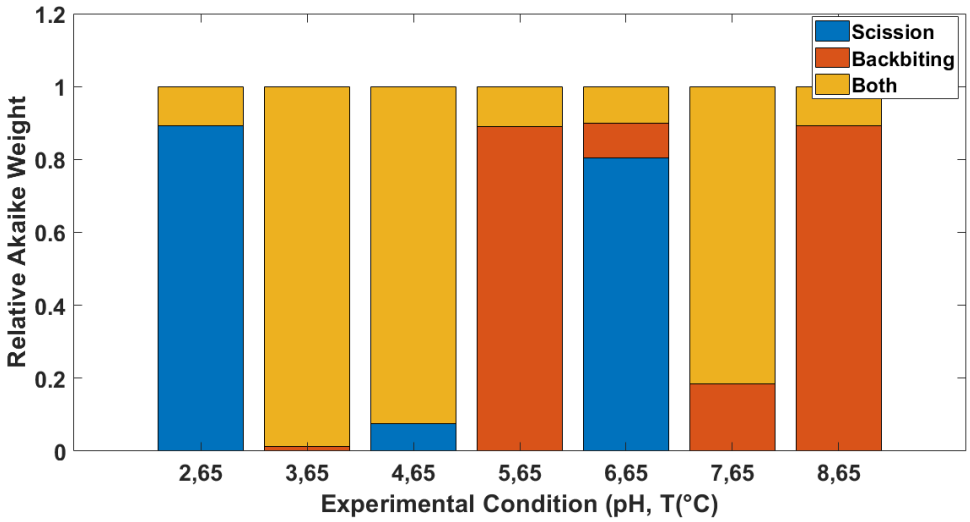

Figure 2.16: Relative Akaike Weights and Model Selection for pH = 2–8 for T = 65°C. The larger the section, the more probable the model structure.

Figure 2.17 shows the contribution of each degradation mechanism as a function of pH and temperature for gAgA. The redder the color, the more dominant the mechanism. The bluer the color, the less dominant the mechanism. According to the data, backbiting constitutes at least half of the degradation at all conditions. This observation is consistent with the literature [19, 25]. However, comparing the results at pH = 6 and T = 65°C in Figure 2.16 and Figure 2.18, there is an anomaly in the dominant mechanism type. This experimental condition has shown to be troublesome, and it will need to be redone to confirm the mechanism.

Percent Contribution of Backbiting

T (°C)	pH: 2	3	4	5	6	7	8
25	51.7	74.4	91.2	95.8	99.7	99.3	99.6
37	59.4	77.0	92.4	97.8	99.7	99.8	99.9
50	56.2	77.0	90.7	98.1	99.1	99.8	99.7
65	63.2	69.2	94.0	98.4	99.4	99.8	99.2
Scale (%)	50.0	58.0	66.0	75.0	84.0	92.0	100.0

Percent Contribution of Scission (hydrolysis)

T (°C)	pH: 2	3	4	5	6	7	8
25	48.3	25.6	8.83	4.21	0.35	0.71	0.39
37	40.6	23.0	7.59	2.15	0.31	0.23	0.09
50	43.8	23.0	9.30	1.86	0.91	0.22	0.30
65	36.8	30.8	6.00	1.65	0.56	0.24	0.79
Scale (%)	50.0	42.0	34.00	25.00	16.00	8.00	0.00

Figure 2.17: $(gA)_2$ (Dimer Only) Degradation Contributions. Contributions from backbiting (top) and scission (bottom). Values calculated from totals in Table 2.4. Cell color scale depicted below each data set. Image taken from Ref. [46] with permission.

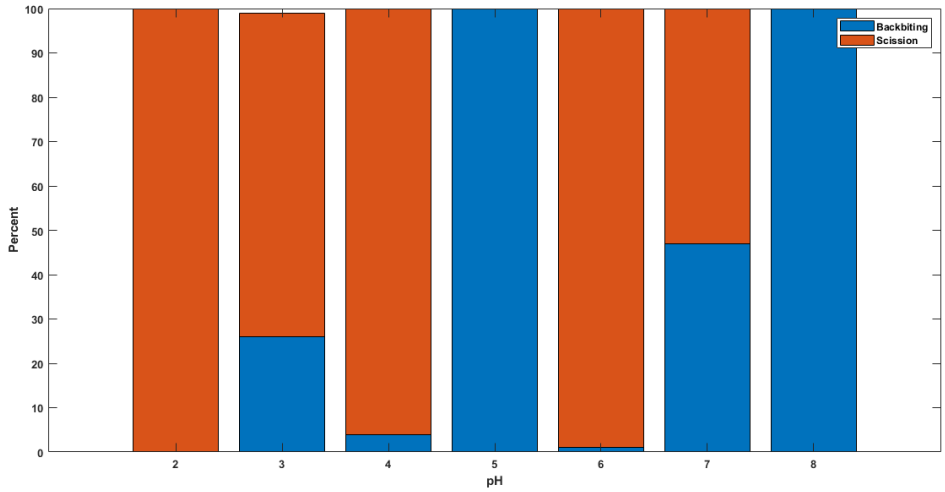

Figure 2.18: Comparisons of Percentage Degradation for gA Oligomers at T = 65°C. Orange bars represent scission while blue bars represent backbiting.

2.5.3 Error on the Data Points and Parameters

The sample standard error of estimate (SEE) (Equation 2.19) and sample standard deviation (Equation 2.20) are the metrics used to measure the error on each experimental data point for gA oligomers. The SEE is used to determine how accurate the predictions of the model are, and it is calculated by Equation 2.19, where SSE is the sum squared error; N is the number of observations; and K is the number of parameters.

The SEE is also known as a tolerance interval which is the range of values that the model predicts would be a data point for a certain timepoint with 95% confidence. The error bars associated with the data points are the sample standard deviations while the error bars associated with the rate constants are the 95% confidence intervals whose calculation is fundamentally different. The SEE could be used for experiments in which one replicate is performed such as the (PA)AgA experiments.

$$SEE = \sqrt{\frac{SSE}{N-K}} \tag{2.19}$$

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2} \tag{2.20}$$

2.5.4 Confidence Intervals

To determine the optimum backbiting rate constants, the MATLAB command patternsearch was used, but it, does not yield any information required to do confidence interval analysis. MATLAB offers an alternative which is the lsqcurvefit command which allows calculation of the best-fit rate constant and a Jacobian matrix. This matrix yields the information needed to put a confidence interval on the rate constants.

Table 2.4 shows the 95% confidence intervals for the backbiting and scission rate constants found in Table 2.2. In the scientific community, the most common confidence interval to use is the 95% variety. Confidence intervals contain the true value of a population pa-

rameter and it is a fixed value. Any interpretation of a confidence interval that mentions sample parameters or random population parameters is incorrect. There are two correct interpretations of the 95% confidence intervals shown in Table 2.4. Assume the interval of (7.06E-9, 1.09E-8) is obtained at pH = 6 and T = 37°C for the upcoming interpretations.

Table 2.4: Backbiting and Scission Constants at Several pH and Temperatures. Units are in L/h. All values in table including the intervals are divided by 10^{-10}.

pH, T°C	gAgA RS	gAgA BB	gA Oligomer RS	gA Oligomer BB
2,50	0.439 ± 0.0251	0.564 ± 0.113		
2,65	0.825 ± 0.597	1.42 ± 0.474	3.08 ± 2.14	0.0565 ± 2.63
3,50	0.331 ± 0.0290	0.853 ± 0.172		
3,65	0.850 ± 0.178	1.91 ± 1.15	1.25 ± 0.93	3.08 ± 1.34
4,50	0.274 ± 0.0164	3.04 ± 0.684		
4,65	0.819 ± 0.784	12.9 ± 10.4	15.3 ± 3.51	6.86 ± 4.41
5,50	0.492 ± 7.3E-3	25.5 ± 7.90		
5,65	1.55 ± 0.715	92.1 ± 27.5	0.0188 ± 10.7	125 ± 17.3
6,50	0.925 ± 0.111	155 ± 71		
6,65	2.62 ± 13.2	620 ± 325	921 ± 571	34 ± 693
7,50	5.47 ± 0.282	2840 ± 408		
7,65	16.6 ± 9.89	6890 ± 942	1240 ± 1010	7520 ± 1570
8,50	22.5 ± 2.39	8680 ± 2500		
8,65	88 ± 36.6	16200 ± 4630	10.2 ± 4520	20500 ± 8800

One correct interpretation of the 95% confidence interval is that if the gA degradation experiment was done 100 times, 95 of the 100 confidence intervals would capture the true backbiting rate constant. The second correct interpretation of the same confidence interval is that it captured the true backbiting rate constant at pH = 6, T = 37°C 95% of the time. Narrow confidence intervals are ideal because they show that the parameter estimation procedure is reasonably accurate. They are a result of good and sound experimental procedure as well as correct modeling procedure. Most of the confidence intervals in Table 2.4 are quite narrow though only 3 replicates were performed. Narrow confidence intervals always occur with infinite sampling, but this is impractical. For this reason, high quality experimentation is important for modeling because if experiments are done incorrectly, parameter estimation and confidence interval interpretation would be useless.

2.6 Chapter 3 Prelude

In Chapter 2, several equations and models were developed to predict and optimize degradation rate constants of gA. However, polymerization and degradation are key to increasing the sequence space of polymers in the prebiotic Earth. Chapter 3 introduces model equations for the esterification of polylactic acid as it is the first step in creating peptide bonds, so knowing the thermodynamics of this process is paramount.

CHAPTER 3
PH DEPENDENT THERMODYNAMICS OF LACTIC ACID OLIGOMERIZATION

3.1 Abstract

Recently, in experiments attempting to replicate conditions on early Earth, aqueous mixtures of hydroxy acids and amino acids were subjected to dry-down conditions to induce ester and amide bond formation of depsipeptides. These experiments were conducted at an unadjusted pH of 3–4 [6]. It is not clear what the thermodynamics of polylactic acid (PLA) formation are at this pH range. The pK_a of lactic acid is 3.86, so lactic acid is expected to be presented in both the neutral and anionic forms in this pH range. Kua and Sweet sought to calculate the thermodynamic driving force for ester and amide bond formation in depsipeptides [13]. However, they only calculated the neutral form for the hydroxy acid, despite reporting the Gibbs energies of reaction at "neutral pH" (pH = 7) where the anionic form of the hydroxy acid is clearly dominant. We instead explicitly calculate the formation energies and enthalpies of the neutral and anionic forms of lactic acid and its oligomers, including chirality. By explicitly calculating the energies of the ionic states, one can calculate how the thermodynamic driving force depends on pH.

This chapter quantifies the thermodynamic driving force for esterification of lactic acid by polycondensation over a range of pH values and hydration levels. The thermodynamic driving force for esterification of lactic acid has been studied at highly acidic pH values [23], but knowledge of esterification at higher pH is important to understand the role of hydroxy acids in the chemical origins of peptides. This topic has not been thoroughly explored in the literature.

Density functional calculations provide formation energies of various ionic states seen

at acidic, neutral, and basic pH. The Gibbs energy, equilibrium constant and species distribution are calculated over a range of pH and hydration levels. The model predictions agree with previous reports at unadjusted acidic pH. Model predictions at higher pH indicate that significant fractional conversions can be obtained at acidic, neutral, and basic pH, as long as the water content in the system is low.

The broader significance of this chapter relates to the first step in the prebiotic pathway of forming protopeptides. These protopeptides could be the precursor to the peptides ubiquitous for life on Earth. The foundation for deeper understanding of the thermodynamics of similar reactions in pharmaceutical applications such as drug delivery and tissue engineering can also be explored [19, 41, 48].

3.2 Introduction

Lactic acid has been extensively studied in the literature [19, 23–28, 49] including the kinetics of polylactic acid (PLA) degradation at different ranges of pH. Vu [24], Witzke [23], and Holten [49] provide foundational thermodynamic data that allows formulation of Gibbs free energy and conversion maps.

Holten [49] wrote a book about lactic acid including its physical chemistry. He presents thermodynamic data on pK_a of lactic acid and pH readings of solutions involving lactic acid and lactoyllactic acid in equilibrium. He determined pH using a pH-meter and a glass electrode system. These pH readings are the foundation of determining pH of other solutions involving lactic acid.

Vu [24] collected PLA oligomer distribution data using gas chromatography, HPLC, and mass spectrometry. He also presented a chemical model that consisted of an infinite series of esterification reactions between successive oligomers of lactic acid and discovered a thermoneutral equilibrium constant $K = 0.2023$. This equilibrium constant is vital to determining the pH of Vu's solution and comparing the thermodynamic model results.

Witzke [23] determined equilibrium relationships between water, lactic acid monomer,

and lactic acid oligomer chains. He used data tabulated in Holten [49] to find average degrees of polymerization and weight percentages of water, lactic acid, and polylactic acid. He developed general relationships about the equilibrium constants of homo-esterification of lactic acid such as Flory-Schulz. Like Vu, Witzke assumed all of the equilibrium constants are equal and independent of chain length. He then reported summary data including conversion, initial lactic acid monomer, average degree of polymerization, and weight fraction of water for several experiments. With this data combined with Holten, the pH of a lactic acid solution can be determined.

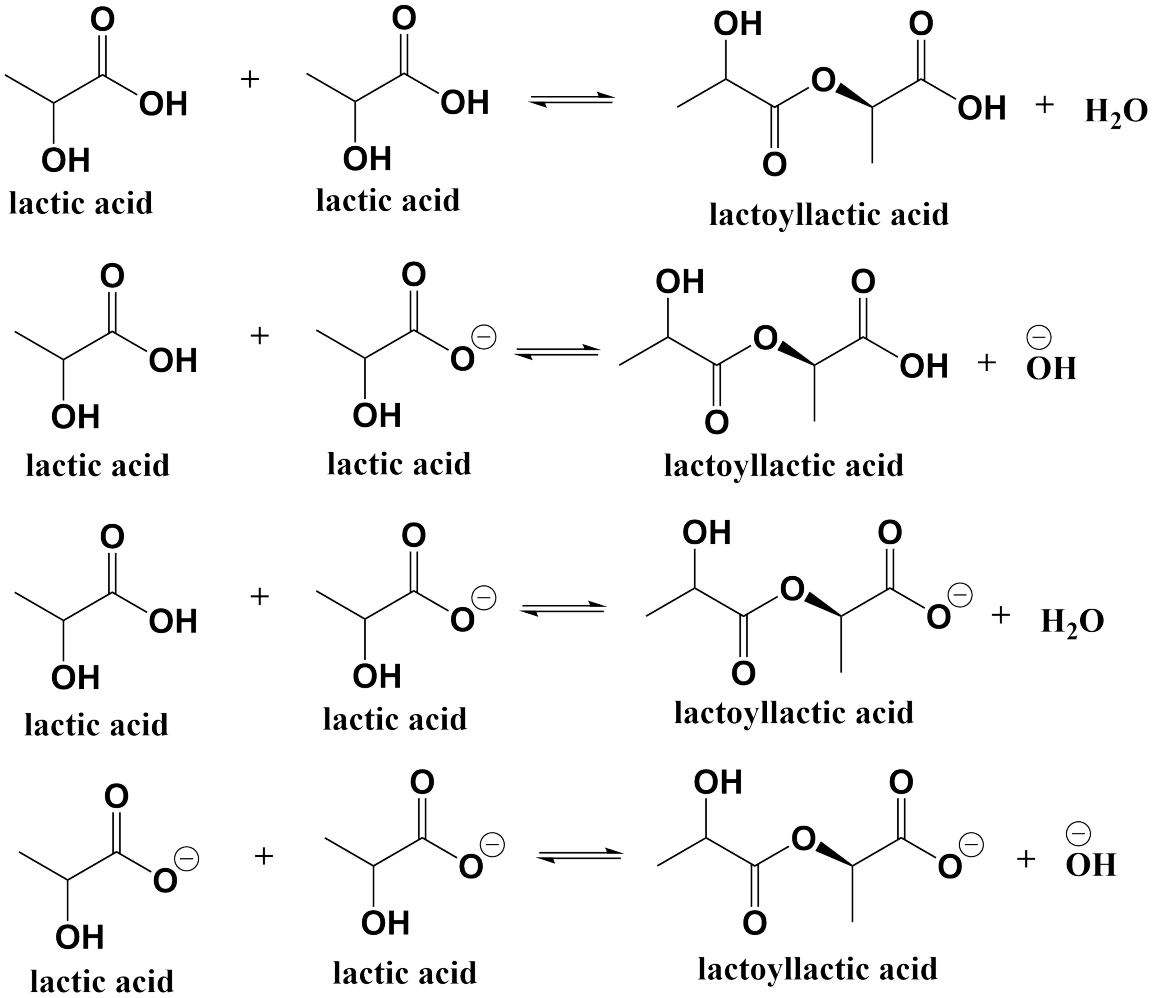

Figure 3.1: List of common polycondensation reactions of lactic acid (dimerization of lactoyllactic acid)

Polylactic acid is typically formed through ring-opening polymerization of the cyclic

lactide dimer in an organic solvent to form the polyester bond [44, 50]. Direct polycondensation of the lactic acid monomer provides an alternative method to ring-opening polymerization, generating PLA from the lactic acid monomer as exhibited in Figure 3.1 [23, 26, 28, 44, 51]. The thermodynamics of the polycondensation of lactic acid in water has been extensively studied under unadjusted pH conditions (less than 1) and as a function of temperature [23, 24, 49]. However, the thermodynamics of polylactic acid are not documented at pH conditions above 1. At dilute conditions, the pH would be higher, but conversion would be very low. There is a disagreement in the sign of ΔG for lactic acid polymerization. Some papers report a slightly positive ΔG [23, 24, 26, 50] while others report a slightly negative ΔG favoring esterification. [6, 14, 18, 28, 33, 51]. Since the statistical analysis was not done on these calculations, ΔG can still be near zero.

The primary goal of this chapter is to determine thermodynamic driving forces of esterification of lactic acid that the literature has not explored yet. A thermodynamic model is developed to calculate thermochemical properties as a function of pH, temperature, and hydration level. It is desired to compare these results to those in the existing literature.

3.3 Methods

3.3.1 Density Functional Theory Calculations

Density functional theory (DFT) is a quantum mechanical method of obtaining an approximate solution of the Schrödinger equation for a many-body system. An approximation is necessary since it is impossible to solve the Schrödinger equation for a many-body system. A many-body system is simply any system comprised of many interacting electrons, atoms and molecules. The information garnered from DFT includes structural and physical property data for a system.

The most widely used basis set to calculate thermodynamic properties, such as free energy, enthalpy and entropy, is called the 6-31G* while the most common hybrid functional is B3LYP [52].

A main alternative to DFT is called the Hartree-Fock (HF) method. The HF method assumes that electrons in a system move independently from each other, and there is no "electron correlation" which causes predictions of energies to become too high. As a result, calculations become too complicated and error-prone [53]. To remedy these complications, DFT was invented in the early 1960s and introduced a concept known as electron density which is theorized to contain all the necessary ground-state property data of a many-body system. It simplifies the HF method by significantly reducing the degrees of freedom of N electrons each with $3N$ spatial coordinates to only three spatial coordinates. This theory is known as the first Hohenberg-Kohn (HK) theorem [54].

For each species in a chemical reaction, the geometry is optimized such that energies are at a global minimum and that the field between solute and solvent is a self-consistent field. Each species is subject to a Cramer/Truhlar [55, 56] treatment such that the total energy is a summation of solvation energy and electrostatic energy. Next, the correction free energy due to internal energy is computed and added to the total energy. From statistical mechanics, this adjusted free energy is used to calculate the relevant thermochemical properties such as Gibbs free energies and enthalpies.

3.3.2 Thermodynamic Model

Table 3.1 displays the DFT free energies (and enthalpies) of formation of lactic acid species and water. These values are needed to compute the Gibbs free energies of reaction found in Table 3.2. The model presented in this chapter simulates a reaction network containing reactions in Table 3.2. From Table 3.2, the esterification reactions of both lactic acid and glycolic acid are spontaneous provided that neutral water is the product and at most one of the lactic acid species (either monomer or dimer) is anionic.

Table 3.1: DFT Free Energies and Enthalpies of Formation of Pure Species in atomic units at T = 25°C and P = 1 atm. Reminder: 2625.45 kJ/mol = 1 au. caa: cyclic dimer of lactic acid. Data taken from [57]

Pure Species	ΔG_f° (au)	ΔH_f° (au)
H_2O	-76.469296	-76.447867
$H_3O^{(+)}$	-76.855396	-76.833186
$OH^{(-)}$	-75.974337	-75.954784
a(R)	-343.696965	-343.658555
a(R)$^{(-)}$	-343.248054	-343.210257
aa(RS)	-610.925365	-610.874081
aa(RS)$^{(-)}$	-610.483643	-610.432735
aa(SS)	-610.925737	-610.873340
aa(SS)$^{(-)}$	-610.484749	-610.430539
aaa(RSS)	-878.155282	-878.086593
aaa(RRR)	-878.155201	-878.086497
aaa(RSS)$^{(-)}$	-877.703096	-877.642245
g	-304.393762	-304.358545
g$^{(-)}$	-303.944443	-303.909796
gg	-532.318961	-532.272090
gg$^{(-)}$	-531.87457	-531.830309
ggg	-760.229233	-760.179605
ggg$^{(-)}$	-759.785922	-759.736641
cgg	-455.854035	-455.816053
caa (RR)	-534.46209	-534.419209
caa(RS)	-534.463902	-534.420893

Table 3.2: Several key reactions and their Gibbs free energy of reaction. The R and S represent lactic acid's chirality. All reactions have charge and atoms balanced. Note: a = lactic acid residue; g = glycolic acid residue. Data taken from [57]

Reaction	ΔG° (kJ/mol)	ΔH° (kJ/mol)
Homodimer		
$a_R + a_R \Leftrightarrow a_R a_R + H_2O$	-2.896	-10.76
$a_R + a_R \Leftrightarrow a_R a_R^{(-)} + H_3O^{(+)}$	145.2	140.16
$a_R + a_R^{(-)} \Leftrightarrow a_R a_R^{(-)} + H_2O$	-19.71	-25.19
$a_R + a_R^{(-)} \Leftrightarrow a_R a_R + OH^{(-)}$	118.0	106.82
$a_R^{(-)} + a_R^{(-)} \Leftrightarrow a_R a_R^{(-)} + OH^{(-)}$	101.2	92.39
$g + g \Leftrightarrow gg + H_2O$	-1.982	-7.53
$g + g^{(-)} \Leftrightarrow gg^{(-)} + H_2O$	-14.86	-25.82
$g + g^{(-)} \Leftrightarrow gg + OH^{(-)}$	117.9	108.87
$g^{(-)} + g^{(-)} \Leftrightarrow gg^{(-)} + OH^{(-)}$	104.96	90.58
Heterodimer		
$a_R + a_S \Leftrightarrow a_R a_S + H_2O$	-1.919	-12.70
$a_R + a_S^{(-)} \Leftrightarrow a_R a_S^{(-)} + H_2O$	-20.79	-30.95
$a_R^{(-)} + a_S \Leftrightarrow a_R a_S + OH^{(-)}$	119.0	104.88
$a_R^{(-)} + a_S^{(-)} \Leftrightarrow a_R a_S^{(-)} + OH^{(-)}$	101.2	86.63
Trimer		
$a_R + a_S a_S \Leftrightarrow a_R a_S a_S + H_2O$	27.93	-6.73
$a_R + a_S a_S^{(-)} \Leftrightarrow a_R a_S a_S^{(-)} + H_2O$	32.96	-2.67
$a_R^{(-)} + a_S a_S \Leftrightarrow a_R a_S a_S^{(-)} + H_2O$	16.14	-17.1
$a_R + a_S a_S^{(-)} \Leftrightarrow a_R a_S a_S + OH^{(-)}$	165.6	125.28
$a_R^{(-)} + a_S a_S \Leftrightarrow a_R a_S a_S + OH^{(-)}$	148.8	110.85
$a_R^{(-)} + a_S a_S^{(-)} \Leftrightarrow a_R a_S a_S^{(-)} + OH^{(-)}$	153.85	114.91
$g + gg \Leftrightarrow ggg + H_2O$	37.27	8.30
Cyclic Dimer		
$a_R a_R \Leftrightarrow c a_R a_R + H_2O$	-14.83	16.45
$a_R a_R^{(-)} \Leftrightarrow c a_R a_R + OH^{(-)}$	122.88	148.46

Before creating the model, assumptions of the model must be stated. They are as follows:

- It is a closed solution, meaning no loss of water or aqueous species to air.
- The lactic acid monomer and oligomers are made up of only the neutral and anionic species.

- The anionic species of the monomer and oligomers exist only through acid dissociation.
- The only reactions are the acid dissociation reactions and the reversible polycondensation reactions.
- The volume of solution equals the volume of solvent plus the volume of the solute, i.e. volume additivity.
- The trimer pK_a is equal to the dimer pK_a.

Once these assumptions are made, several equations can be written that describe the esterification system in Table 3.2 starting with Equation 3.1, which is the acid dissociation of an oligomer chain of length i. This equation yields oligomer ratios between the neutral and anionic dissociation states. These ratios are then used to calculate the thermodynamic mole fractions:

$$pH = pK_{a(L_i)} + \log([{L_i}^{(-)}]/[L_i]) \tag{3.1}$$

where pH is the pH of solution; $[L_i]$ is the concentration of the neutral aqueous lactic acid oligomer of chain length i, $[L_i^{(-)}]$ is the concentration of the anionic aqueous lactic acid oligomer of chain length i, and $pK_{a(L_i)}$ is the pK_a of lactic acid oligomer length i.

3.3.3 Atom Balances

A critical aspect of modeling the esterification reaction network is the atom balance. It is necessary to balance the carbon, hydrogen, and oxygen atoms when using formation energies to calculate ΔG and ΔH. Each reaction class in Table 3.2 has its own atom balance, which can equivalently be considered as a charge balance. The reaction scheme is as follows:

$$x_1 L_1 + (1 - x_1) L_1^{(-)} + x_2 L_2 + (1 - x_2) L_2^{(-)} \leftrightarrow x_3 L_3 + (1 - x_3) L_3^{(-)} + \alpha H_2O + \beta H_3O^{(+)} + \gamma OH^{(-)} \tag{3.2}$$

where x_i is the mole fraction of species i calculated from Equation 3.1; and α, β, γ represents the stoichiometric coefficients of water, hydronium, and hydroxide, respectively.

A carbon balance is performed on Equation 3.2, and there are nine carbon atoms balanced on each side, three carbons for each lactic acid monomer.

A hydrogen balance on Equation 3.2 yields the following expression:

$$x_1 + x_2 - x_3 + 1 = 2\alpha + 3\beta + \gamma \tag{3.3}$$

The oxygen balance yields:

$$\alpha + \beta + \gamma = 1 \tag{3.4}$$

There are no unique solutions to this set of atom balances. However, one can obtain unique solutions by making an additional assumption: (β = 0) or (γ = 0). It is assumed that only one ionic species of water is produced. This assumption is necessary because β γ = K_w (acid dissociation constant of water = 10^{-14}) sometimes yields infeasible solutions since pH is not actually constant due to acid consumption. Each case thus yields a unique solution for α, β, and γ.

For the case of β = 0, $\alpha = x_1 + x_2 - x_3$, β = 0, and γ = 1 - α as long as $0 \leq \alpha \leq 1$. For the case of γ = 0, $\alpha = 1 - \beta$, $\beta = x_1 + x_2 - x_3 - 1$, and γ = 0 as long as $0 \leq \beta \leq 1$. To determine which case to use, each of the mole fractions x_1–x_3 is calculated using pH and pK_a data. The parameters α and β are obtained. Whichever parameter is between zero and one inclusive, that is the case to use. For example, α and β could be 0.88 and 1.12 for a particular pH. Because α is between zero and one, it is used in future calculations. There

is a point where α = β = 1, but these give the same solution for both cases.

3.3.4 Calculate Mole Fractions

The pH and $pK_{a(L_2)}$ parameters are specified in the model. The pH was varied from 0–8 and $pK_{a(L_2)}$ was varied from 2–4 in MATLAB. The benchmark value of the dimer pK_a is 2.74 [58], which is the predicted pK_a of a similar molecule seen in Figure 3.2.

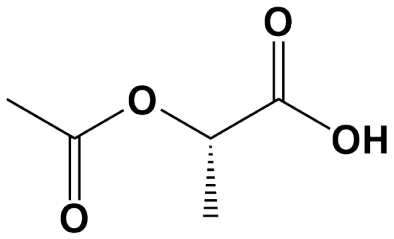

Figure 3.2: O-Acetyl Lactic Acid

For every pH and dimer pK_a, the general form of the mole fraction of any dissociation state of a species is in Equation 3.5. The equation assumes two dissociation states (neutral and anionic). One can expand the summation and derive an expression containing a ratio between the concentrations of anionic and neutral species. After the ratio is calculated from Equation 3.1, the mole fraction can be determined:

$$x_i = \frac{[M_i]}{\sum_{i=1}^{N}[M_i]} \tag{3.5}$$

where x_i is the mole fraction of the i^{th} dissociation state of species M and M_i is the concentration of the i^{th} dissociation state of species M. These mole fractions are inputs to the Gibbs free energy equations (and enthalpies) as shown in Equation 3.6 and Equation 3.7.

3.3.5 Computing Gibbs Energies of Reactions

The Gibbs free energy for the dimer reaction can be determined by the following equation:

$$\Delta G_r^\circ = \Delta G_{f,H_2O}^\circ + \Delta G_{f,aa}^\circ - 2\Delta G_{f,a}^\circ \tag{3.6}$$

where ΔG_r° is the standard Gibbs energy of reaction and each of the ΔG_f° terms are the standard Gibbs energy of formation of each species in reaction. At different pH values, the distribution of the lactic acid species' dissociation states (and thus ΔG_r°) will change. These changes are reflected in the mole fractions, x_i, and can be seen in the following modified equation for the dimer reaction:

$$\begin{aligned}\Delta G_r^\circ = (x_2 \Delta G_{f,L_2}^\circ + (1-x_2)\Delta G_{f,L_2^{(-)}}^\circ) + \Delta G_{f,H_2O}^\circ - \\ 2(x_1 \Delta G_{f,L_1)}^\circ + (1-x_1)\Delta G_{f,L_1^{(-)}}^\circ)\end{aligned} \tag{3.7}$$

$\Delta G_{f,H_2O}^\circ$ is defined as:

$$\Delta G_{f,H_2O}^\circ = \alpha \Delta G_{f,H_2O} + \beta \Delta G_{f,H_3O^{(+)}} + \gamma \Delta G_{f,OH^{(-)}} \tag{3.8}$$

3.3.6 Temperature Dependence

The third facet of the theoretical framework is to account for temperature changes. This step is important because the temperatures cited in literature or in lab experiments may not be the same as DFT, so adjustments are necessary. The van't Hoff equation (Equation 3.9) computes the equilibrium constants and is stated as follows:

$$\ln \frac{K_2}{K_1} = -\frac{\Delta H^\circ}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \tag{3.9}$$

where T_1 and T_2 are absolute temperatures, K_1 and K_2 are the equilibrium constants at T_1 and T_2, respectively, R is the universal gas constant, and ΔH° is the enthalpy change of reaction. The ΔH° value for the esterification reaction is calculated using the values in Table 3.1 and Equation 3.6 – Equation 3.7 since enthalpy and Gibbs free energy are state functions.

3.3.7 Literature Comparisons

Model predictions should be validated using experimental data. The literature values will come from three sources: Vu [24], Holten [49] and Witzke [23]. Holten [49] provides experimental data where lactic acid and lactoyllactic acid are in equilibrium. According to Holten, the pH of the solution is dependent on the weight percent of lactic acid. Equation 3.10 allows one to compute the weight fraction of monomer in solution at equilibrium.

$$w_1 = \frac{c_{LA}\frac{mol}{L}\rho_{LA}}{\left(c_{LA}\frac{mol}{L}\rho_{LA} + c_{H_2O}\ \frac{mol}{L}\rho_{H_2O}\right)} \tag{3.10}$$

where w_1 is the weight fraction of lactic acid monomer, c_{LA} is the lactic acid monomer concentration in solution at equilibrium, ρ_{LA} is the density of lactic acid, ρ_{H_2O} is the density of water, and c_{H_2O} is the concentration of water in solution at equilibrium. As solutions become more concentrated, c_{H_2O} becomes smaller. The density of water is taken to be 1 kg / L while the density of lactic acid is 1.206 kg / L.

Extracting pH from Experimental Data

The model accounts for pH and temperature, and one needs to find literature values for both of these parameters for any experimental condition in the model to be able to make comparisons. Temperature is not a problem; however, Holten [49] is the only source to provide pH values. The pH readings are a function of lactic acid weight percent. If one can calculate weight percent from experimental data, then Table 3.3 is cross checked to find pH. Both Witzke [23] and Vu [24] calculated weight fractions for lactic acid for a variety of experiments, so pH can be extracted from them.

Table 3.3: pH Readings of Aqueous Solutions of Lactic Acid at 25°C. The lactic acid is in equilibrium with lactoyllactic acid in solution. Reproduced from Ref. [49] with permission

Lactic Acid % by Weight	pH
1.0	2.28
6.29	1.87
10.0	1.75
12.19	1.72
25.02	1.45
37.30	1.23
54.94	0.85
77.7	0.39
84.0	0.20
88.6	0.00

Witzke [23] analyzed lactic acid equilibrium with only monomer and dimer at $[COOH_0]$ = 7.88 M and p = 0.06, where p and $[COOH_0]$ are defined as:

$$p = 1 - ([COOH]/[COOH]_0) \tag{3.11}$$

where p is the fractional conversion, $[COOH]$ is the total concentration of all acid endgroups at equilibrium, and $[COOH_0]$ is the initial concentration of acid endgroups. Only monomer exists in solution initially.

Witzke also studied highly concentrated lactic acid solutions containing longer oligomer chains. The second test solution from Witzke has the initial lactic acid monomer concentration, $[COOH_0]$ = 9.97 M and fractional conversion, p = 0.11. Under these parameters, the maximum oligomer length observed is three.

Weight fraction of lactic acid is important to make a proper comparison of the model to the literature. When dealing with concentrated oligomers, one needs an equation that can accommodate for more concentrated solutions. Witzke provides an equation on computing weight fraction of lactic acid in Eq. (1-13) [23]:

$$w_i = \frac{[COOH](72i + 18)(1 - p)p^{i-1}}{1000\rho_{solution}} \tag{3.12}$$

where w_i is the weight fraction of oligomer length i and $\rho_{solution}$ is the density of the lactic acid solution and will change depending on the amount of lactic acid in solution. This equation comes from the definition of weight fraction. It is defined as the concentration of all lactic acid monomer residues in solution divided by the overall solution concentration. The term $[COOH](1-p)(p^{i-1})$ represents the concentration of lactic acid oligomer length i from the Flory-Schulz distribution. The Flory-Schulz distribution equation (Equation 3.13) computes these concentrations, and it assumes that all endgroups in solution have equal reactivity.

To obtain the total amount of lactic acid residues, this value is multiplied by the molecular mass of the oligomer length i, but a water molecule is released each time an ester linkage is formed. This multiplicative factor is the $72i + 18$ term. As lactic acid elongates by one, the longer oligomer has 72 more atomic mass units than the shorter oligomer. The constant of 18 represents the molecular weight of the water molecule ejected from the oligomerization.

The individual mole fractions are summed. Then, the total is subtracted from 100 to obtain the weight percent of water. The sum of the mole fractions is the total weight percent to cross-check in Table 3.3. In the literature, both weight percents are available.

3.4 Results and Discussion

3.4.1 Calculating ΔG_r°

The inputs required to calculate ΔG_r° are pH, pK_a, concentrations of dissociation states (e.g. $a^{(-)}$), and mole fractions of the dissociation states. Table 3.1 is consulted to acquire the Gibbs energies of formation of relevant lactic acid compounds. Finally, the Gibbs energies are calculated using Equation 3.7 for dimer formation or similar equations for other reactions.

3.4.2 Comparison to experimental data

Unlike Vu [24], Witzke has more complete data regarding initial concentrations of lactic acid and fractional conversion.

In Witzke's [23] test solution where $[COOH]_0$ = 7.88 M and water concentration was 24.2 M, p = 0.06 so only 6% of the $[COOH]_0$ converted into other forms. In this solution, this "other form" is primarily the dimer. There are longer oligomers in this solution, but they represent only 0.34% of the total $[COOH]$ in solution justifying the assumption that only monomer and dimer exist. To obtain this value, one needs to compute $[COOH]$ using Equation 3.11 given p and $[COOH]_0$. The equilibrium concentrations of both monomer and dimer need to be obtained. One can arrive to the same conclusion using Equation 3.12 assuming a $\rho_{solution} \approx$ 1.1 kg/L and can calculate the sum of the weight fractions of the monomer and dimer. This sum is added to the known weight fraction of water which is 38% according to Witzke

$$[COOH](1-p)p^{i-1} = C_i \tag{3.13}$$

where C_i is the equilibrium concentration of lactic acid oligomer of length i.

Once C_1 and C_2 are calculated, this value is divided by $[COOH]$. Using the above values for p and $[COOH]_0$, one can see that 99.66% of the total lactic acid in equilibrium is in the monomer and dimer form leaving 0.34% for the other forms.

Using Witzke's dimer data, the weight percent of lactic acid (n = 1) is 61.64%. Cross-referencing the value to Table 3.3 and using linear interpolation, the pH of the solution would be 0.71.

In another test solution from Witzke where the maximum reported oligomer length is 3 (trimer), the water concentration was 16.25 M. The weight fraction of lactic acid is 0.75. Cross-referencing the value to Table 3.3, the pH of solution is approximately 0.4. The pH decreased as more acid is added. When Holten reports lactic acid, it is interpreted as the

total amount of lactic acid that is added, not the current monomer concentration.

From Table 3.4, the model predicts spontaneity of lactic acid oligomerization while the literature reports unspontaneity of lactic acid oligomerization. Because ΔG_r° of lactic acid oligomerization is near zero, both the model prediction and the literature reports may not be contradictory. An error analysis on the experimental values would be prudent.

Table 3.4: Table of esterification reactions from a variety of literature sources. The reaction is a + a $\leftrightarrow$ aa + water. * (The ΔG_r° will be the same regardless of what the temperature is.)

Source	pH	Temperature ($^\circ C$)	Racemic?	Reported ΔG_r°	Model Prediction
[24]	0.5	temperature independent*	Unknown	3.96 kJ/mol	-2.2 kJ/mol
[23]	0.71	25	Yes	3.75 kJ/mol	-1.76 kJ/mol
[23]	0.4	25	Yes	3.34 kJ/mol	-2.34 kJ/mol

3.4.3 Impact of Chirality

The similarity of the energies between SS and RS diastereomers in Table 3.1 suggests that chirality does not play a strong role in the driving forces of dimerization and subsequent oligomerizations. RR and SS are enantiomers as well as RS and SR. Enantiomers have identical thermodynamics; however, SS and RS are diasteromers which can be different thermodynamically. The abundances between the two diastereomers are approximately equal since the ΔG_r° of SS $\leftrightarrow$ RS is -1 kJ/mol. From this value and T = 25°C, the equilibrium constant equals 1.5. At an elevated temperature such as T = 75°C, the equilibrium constant equals 0.71.

The analysis on the lactic acid trimers can be performed similarly. The ΔG_r° of RRR $\leftrightarrow$ RSS is -0.21 kJ/mol. Assuming a temperature of 25°C, the equilibrium constant equals 1.0 suggesting a 50/50 racemic mixture of the two diastereomers. The calculations do not suggest a strong preference for homochirality in oligomerization of lactic acid.

Figure 3.3 is an energy map for the esterification reaction a + a $\Leftrightarrow$ aa + H_2O for different dimer chiralities. The two curves are practically superimposed on each other which matches the observation that chirality plays very little role in the oligomerization.

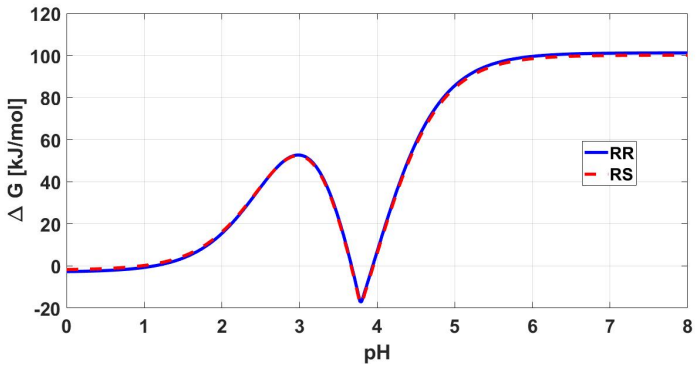

Figure 3.3: pH vs ΔG_r° of a + a $\Leftrightarrow$ aa + H_2O for different dimer chiralities. Dimer pK_a set to 3 and T = 25$^\circ C$.

3.4.4 ΔG_r° as a function of pH and pK_a

Figure 3.3, Figure 3.4 and Figure 3.6 show how the ΔG_r° of the a + a $\leftrightarrow$ aa + H_2O reaction changes with pH according to dimer pK_a, chirality, and temperature, respectively. The pK_a of the linear dimer of lactic acid is not reported in the literature, so a benchmark for the compound is established. The benchmark must be a compound chemically similar to the lactic acid dimer whose pK_a is detemined. That compound is O-Acetyl Lactic Acid (pK_a = 2.74) in Figure 3.2. It is difficult to measure the pK_a because the linear dimer is not stable in water as it degrades to monomer or cyclizes.

Dimer Energy Plot

There are some important insights to learn from Figure 3.4. At highly acidic pH, the reaction is very slightly spontaneous and monotonically increases in Gibbs free energy. The system is closed, beginning with lactic acid monomer and some water. pK_a is a critical parameter in determining how spontaneity changes as a function of pH. Between pH = 3 and pH = 3.8, the reaction becomes more spontaneous at a dimer pK_a between 2.74 and

3. It also reaches a minimum Gibbs free energy near the pK_a of the monomer lactic acid (3.86).

If one compares Table 3.2 to Figure 3.4, the pH can identify the most dominant reactions. At low pH, the Gibbs free energy of reaction is only slightly spontaneous as indicated by the completely neutral reaction of lactic acid esterification. As pH increases and species begin to ionize, the Gibbs free energy generally increases. This is when water begins to ionize to hydroxide, which significantly hinders the reaction. At certain pK_a values and pH ranges, the reaction's free energy dips considerably representing the reaction where only the dimer and one of the monomers are ionized. This reaction has the greatest spontaneity from Table 3.2.

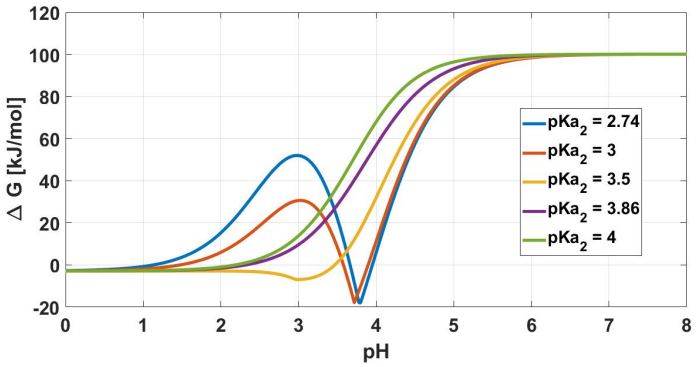

Figure 3.4: pH vs ΔG_r° of a + a $\Leftrightarrow$ aa + H_2O for several dimer pK_a values. $T = 25°C$. The dimer is homochiral.

Trimer Energy Plot

Figure 3.5 shows that the reaction is unfavorable regardless of pH and temperature, unless water is removed. However, the reaction is more favorable at highly acidic pH.

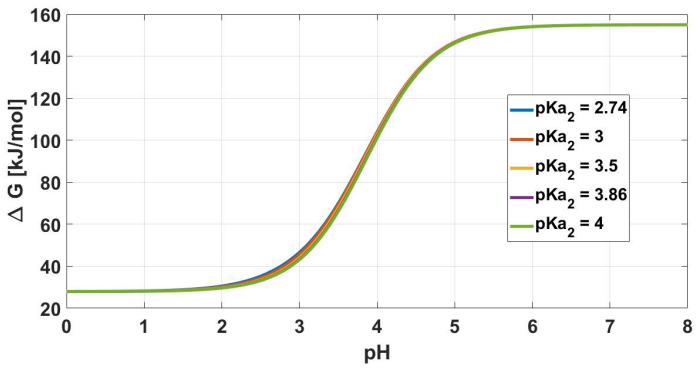

Figure 3.5: pH vs ΔG_r° of a + aa $\Leftrightarrow$ aaa + H_2O for several dimer pK_a values. The trimer pK_a is unknown and is assumed to be the same as dimer. The trimer's chirality is RSS.

van't Hoff

From Figure 3.6, temperature also plays a minor role in the thermodynamic driving force which is expected for a thermoneutral reaction.

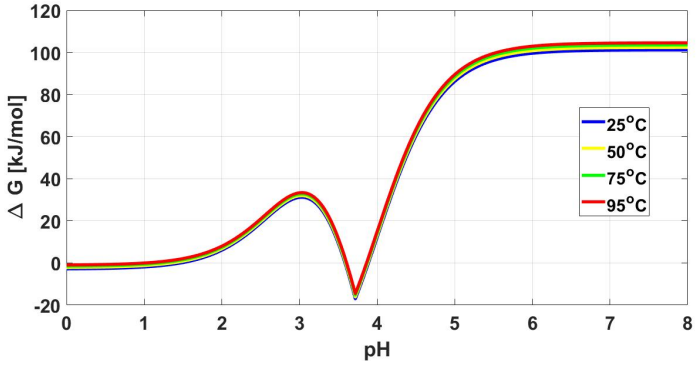

Figure 3.6: pH vs ΔG_r° of a + a $\Leftrightarrow$ aa + H_2O for several temperatures. Dimer pK_a set to 2.74 and the dimer is homochiral.

Cyclic Dimer DFT Calculation

Figure 3.7 shows the free energy map as a function of pH for the cyclization of lactic acid dimer. If there is any cyclic dimer to begin with, Figure 3.7 shows that cyclic dimer is unfavored at pH $>$ 1.8 assuming a linear dimer pK_a of 2.74. Note that this calculation is performed for the reaction aa ¡-¿ caa + H_2O with no water initially present. If water is initially present, the reaction would be more unfavorable, according to Le Chatelier's

principle. DFT calculations are in agreement with the cyclization reaction thermodynamics in the literature suggesting that the dimer should be in the linear form [23, 24, 26, 30, 33, 59, 60].

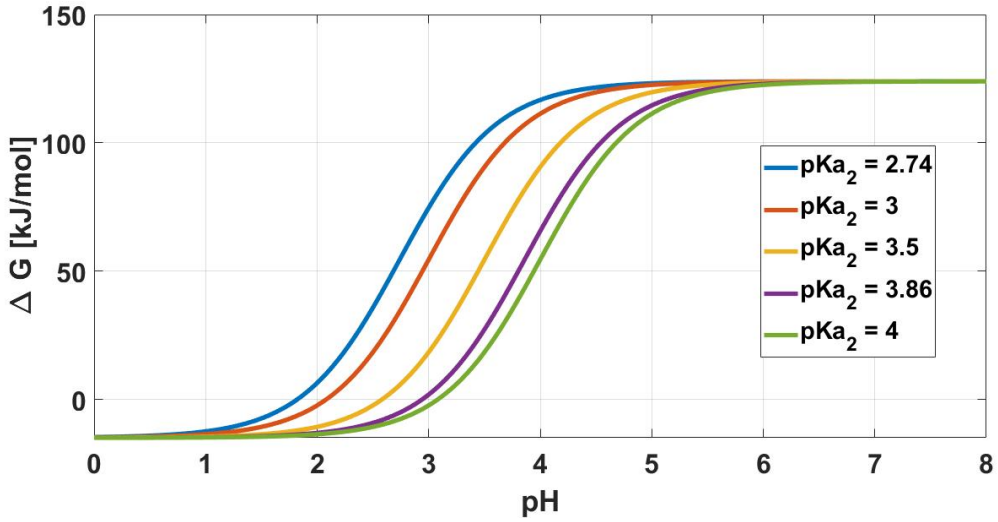

Figure 3.7: pH vs ΔG_r of aa $\leftrightarrow$ caa + H_2O for several dimer pK_a values. The temperature is $25°C$.

Considering a solution initially containing 1 M of lactoyllactic acid, one can assume that cyclization is the only relevant reaction, the pK_a of the lactoyllactic acid is 3, the equilibrium concentration of water is 55 M, the pH is 3, and T = 25°C. With these parameters, the equilibrium concentration of lactoyllactic acid is 1 M while no cyclic product is formed. In pure water, cyclic dimer will always be the dominant product even at extremely low pH ranges. In order for this reaction to produce linear dimer product, the water content must be reduced.

Although cyclic dimer is the thermodynamic product, the linear dimer may still be the most dominant product in aqueous solution due to kinetics making it the kinetic product. De Jong [25] and van Nostrum [19] show that the lactide (cyclic ester) forms but then quickly hydrolyzes into lactic acid in alkaline solutions. In acidic solutions, their mechanisms do not include lactide. Instead, the OH terminus of a long lactic acid oligomer (such as i = 7) is protonated, temporarily forming a 5-member ring intermediate. Finally, water attacks

the carbonyl group adjacent to the OH terminus forming a lactic acid oligomer one degree of polymerization (DP) lower, a lactic acid monomer, and hydronium. Lactide plays no role in acidic media. This mechanistic argument also explains why Witzke never reported seeing lactide in solution even at highly acidic solutions.

Combining the observations of Chapter 2 and Chapter 3, one can see that under highly acidic conditions (pH < 1.8), the kinetic driving force is more significant at our experimental timescales than the thermodynamic driving force. The linear dimer (kinetic product) is observed while the cyclic dimer (thermodynamic product) is not observed. Under alkaline conditions, the linear dimer is both the kinetic and thermodynamic product.

Energy Plot Behavior

In the previous sections, several energy plots were presented. Each of them displayed interesting behavior such as dips at certain pH ranges. This qualitatively suggests thermodynamic competitive forces at these ranges. Coincindentally, they happen at around the pK_a of the monomer. These forces arise from the dissociation state distribute dictated by the pH of solution. For the dimer energy plot, the working reaction becomes more spontaneous if either the products have a more negative Gibbs energy of formation or if the reactants have a more positive Gibbs energy of formation.

From Table 3.1, as pH rises, the water begins to ionize into hydroxide while the lactic acid dimer begins to ionize into its anionic state. Hydroxide has a more positive Gibbs energy of formation than water which contributes to the overall Gibbs free energy of reaction rising. The same is true for the lactic acid dimer. Meanwhile, the lactic acid monomer as a reactant actually lowers the overall Gibbs free energy of reaction as pH increases.

It is also clear how strong a role pK_a plays in determining whether the thermodynamic forces are competing or not because in some plots, the curves are either monotonically changing or there are dips. The pK_a and pH dictate these forces.

3.4.5 Impact of Water Concentration

Water is both a product and the solvent in this system. Figure 3.8 shows the conversion of lactic acid as a function of pH and initial water content ($[W_0]$). The contours represent the carboxylic acid conversion. As pH increases while holding water content constant, conversion decreases. As water content increases while holding pH constant, conversion also decreases according to Le Chatelier's Principle. These contour plots suggest that oligomerization is favorable when solutions are highly acidic and dry, which supports the notion that drydowns give higher lactic acid conversions. The initial lactic acid monomer concentration in Figure 3.8 was 7.88 M, from Witzke.

Another feature of Figure 3.8 is the red line that represents Holten's experimental data given the initial lactic acid concentration in solution and the measured pH. The lactic acid concentration could be used to back-calculate the initial water concentration given their densities and molecular weights. Points off the red line could be achieved by buffering the system.

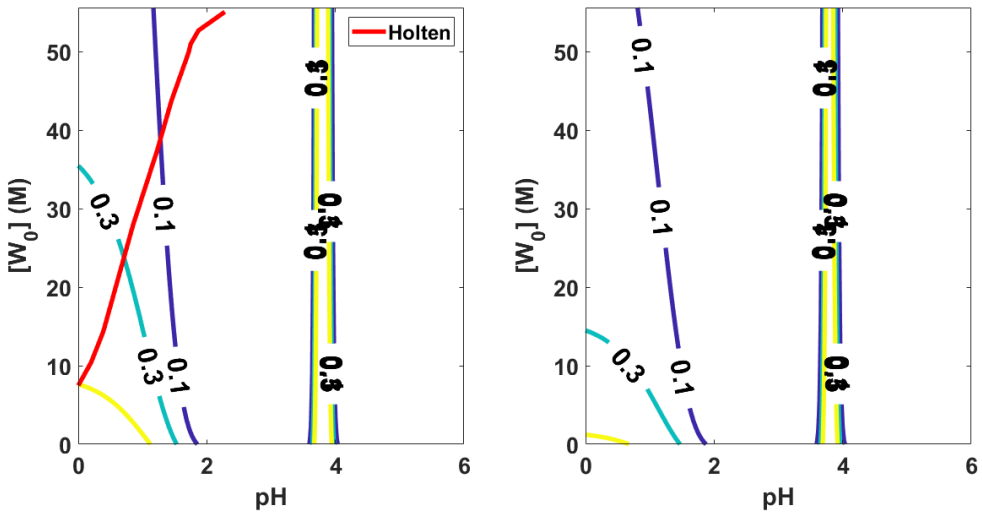

Figure 3.8: pH vs initial water content for dimerization of lactic acid. The colored curves represent the conversion of lactic acid computed from Harshe [28]. Initial concentration of lactic acid monomer is 7.88 M. The dimer pK_a is 2.74. Left: The temperature is $25°C$. Right: The temperature is $95°C$.

All of the results seen in this section are for closed systems, but by removing water produced by the reaction, arbitrary conversions could be achieved.

Effect of Water Addition and Removal

The system analyzed above has lactic acid as the reactant, with dimer and water existing in equal amounts as the product of the reaction. However, this is not the typical experimental scenario. Experiments such as those by Witzke and by Holten operate on a closed system, beginning with a concentrated aqueous lactic acid solution. Water is not only the product, but is also playing the role of a solvent. If the water W concentration is reduced, then the conversion x of monomer M into dimer D is observed to increase, according to Le Chatelier's principle:

$$K_{eq} = e^{(-\Delta G_r/RT)} = \frac{[D][W]}{[M]^2} = \frac{x([W]_0 + x[M]_0)}{([M]_0(1-x)^2)} \tag{3.14}$$

This equation is derived from reaction engineering principles. The concentration of each species is written on the basis of the limiting reactant which is the monomer M. These concentrations are as follows:

$$\begin{aligned} [D] &= [D_0] + [M_0]x \\ \text{[W]} &= \text{[W}_0\text{]} + [M_0]x \\ \text{[M]} &= \text{[M}_0\text{]}(1-x) \end{aligned} \tag{3.15}$$

In this system, no dimer exists initially so $[D_0]$ = 0. When Equation 3.15 is substituted into the third expression in Equation 3.14, one obtains the last expression. Because the pH information is already contained in the equilibrium constant, the conversion will be a function of pH.

An interesting note is that the last expression in Equation 3.14 is in implicit form. The explicit form of this equation is found in Harshe et al. [28] and is written in Equation 3.16. Both forms give identical answers.

$$X_{eq} = \frac{1}{2(K_{eq}-1)}\left(2K_{eq} + \frac{W_0}{\lambda_{1,0}} + \sqrt{4K_{eq}\left(\frac{W_0}{\lambda_{1,0}}+1\right) + \left(\frac{W_0}{\lambda_{1,0}}\right)^2}\right) \tag{3.16}$$

where $\lambda_{1,0}$ is the total concentration of monomer M charged to the solution initially; and X_{eq} is the equilibrium conversion and is identical to x for our purposes. By solving Equation 3.15 for x, one obtains Equation 3.16.

Once the overall Gibbs energy of reaction is known, then the equilibrium constant K_{eq} is obtained for the particular temperature of interest. Further specifying the initial monomer concentration $[M]_0$ and the initial water concentration $[W]_0$, the equilibrium dimer concentration [D] can thus be computed, as well as the monomer conversion.

Figure 3.8 shows the conversion of lactic acid as a function of pH and initial water content ($[W_0]$). The contours represent the carboxylic acid conversion. As pH increases while holding water content constant, conversion decreases. As water content increases while holding pH constant, conversion also decreases according to Le Chatelier's principle. These contour plots suggest that oligomerization is favorable when solutions are highly acidic and dry, which supports the notion that drydowns give higher lactic acid conversions. The initial lactic acid monomer concentration in Figure 3.8 was 7.88 M, from Witzke.

Another feature of Figure 3.8 is the red line that represents Holten's experimental data given the initial lactic acid concentration in solution and the measured pH. The lactic acid concentration could be used to back-calculate the initial water concentration given their densities and molecular weights. Points off the red line could be achieved by buffering the system.

The dimerization reaction can be further driven toward products in an open system, by removing the initial and the product water. Drying reactions have been used to polymerize lactic acid [28] as well as malic acid [17, 61].

Figure 3.9 illustrates the dependence of monomer conversion on the amount of water

removed, at various pH values and T = 25°C. A simple system contains 1 mol of dry lactic acid monomer initially under standard conditions (T = 25°C, P = 1 atm).

As water is removed, the new equilibrium distribution of species will change. This distribution will change the mole fractions of each of the species. In fact, more lactic acid dimer will be produced. Since lactic acid dimer has a lower pK_a than the monomer, the pH of solution should be lower. In Figure 3.9, one can see that a maximum conversion can be achieved for a certain pH and a certain amount of water removed because there is only so much water that one can remove from the equilibrium system due to stoichiometry and having the lactic acid be the limiting reagent.

This example can be extended to the contour plots in Figure 3.8. By considering one pH, one water concentration, and the equilibrium constant, one can use Le Chatelier's principle and ICE tables to see that conversions will increase as water is removed from the system.

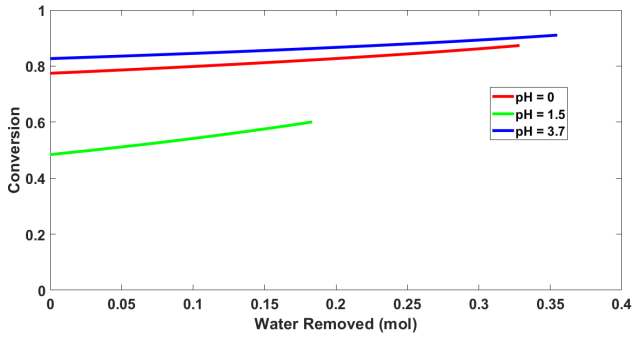

Figure 3.9: Conversion as a function of water removed (mol) from the equilibrium system that initially contained 1 mol of dry lactic acid. The temperature is $25°C$ and the pressure is 1 atm. Three representative pH values are chosen.

Derivation

To derive the above plot, one must start with a working reaction to start analysis.The working reaction of this system is 2M $\leftrightarrow$ D + W where M is the lactic acid monomer, lactic acid dimer, and water. It is desired to adjust the working reaction so that the limiting reactant has a stoichiometric coefficient of 1. Because M is the limiting reagent, the working re-

action is divided by 2 to yield the working reaction M $\leftrightarrow$ $\frac{1}{2}$ D + $\frac{1}{2}$ W. Since the working reaction has changed, so has its thermodynamics. The original working reaction has an equilibrium constant of 0.22 at pH = 1.5. However, the equilibrium constant changes with the new working reaction. To calculate this new value, one can set up two ΔG equations for each working reaction, the old and new one:

$$\begin{aligned} \Delta \mathrm{G}_{old} &= -RTln(K_{old}) \\ \Delta \mathrm{G}_{new} &= -RTln(K_{new}) \end{aligned} \tag{3.17}$$

Because ΔG_{new} is half of ΔG_{old} due to Hess's Law, the equilibrium constant for the new working reaction will also change. By substitution and solving for K_{new}, K_{new} is equal to the square root of K_{old}. In this case, K_{new} = 0.469. This is the equilibrium constant to use for future analysis.

The next step is writing the equilibrium expression for the new working reaction. One can construct an ICE (Initial, Change, Equilibrium) table for determining the equilibrium amount (in moles) of each species. From stoichiometry, the equilibrium amount of monomer is simply $M_0 - x$ while the equilibrium amounts of dimer and water are $\frac{1}{2}$x each assuming no dimer and water exist initially in solution. The variable x represents the amount of lactic acid consumed after equilibrium has occurred and M_0 is the initial moles of lactic acid monomer in solution. If M_0 = 1 and pH = 1.5, the equilibrium amounts of lactic acid monomer, lactic acid dimer, and water are 0.174 mol, 0.413 mol, and 0.413 mol, respectively by solving Equation 3.18. This calculation yields an equilibrium conversion of 48.4%.

$$K_{eq} = e^{(-\Delta G_r/RT)} = \frac{[D]^{0.5}[W]^{0.5}}{[M]} = \frac{(0.5x)^{0.5}(0.5x)^{0.5}}{M_0 - x} \tag{3.18}$$

The analysis continues from the equilibrium system described above. Water is forcibly removed from the system perturbing it. The reaction will respond according to Le Chatelier's principle. The working reaction is the same as before (M $\leftrightarrow$ $\frac{1}{2}$ D + $\frac{1}{2}$ W). This time,

some variables need to be denoted to dynamically handle various water removal rates. The variables are y and c which represent the amount of lactic acid consumed after perturbation and the amount of water removed, respectively.

One must revisit Equation 3.18 but this time, there are two major differences from before. The first difference is that the initial distribution of species is the equilibrium distribution calculated in the first part (D_0 = W_0 = 0.413 mol and M_0 = 0.174 mol). The second difference is that a changeable parameter must be introduced to dynamically solve Equation 3.19.

$$K_{eq} = e^{(-\Delta G_r/RT)} = \frac{[D]^{0.5}[W]^{0.5}}{[M]} = \frac{(D_0 + 0.5y)^{0.5}(W_0 - c + 0.5y)^{0.5}}{M_0 - y} \tag{3.19}$$

Finally, the overall conversion after initial equilibration and future perturbation is defined as:

$$X = 1 - \frac{M_{final}}{M_0} \tag{3.20}$$

where X is the overall conversion of lactic acid, M_{final} is the final concentration of lactic acid after perturbation of the equilibrium system, and M_0 is the initial lactic acid monomer concentration at the beginning of the analysis which is 1 mol.

This conversion is plotted for several water removal rates at three selected pH values in Figure 3.9.

It is important to note that Figure 3.9 is assuming a dry solution once all of the water is removed.

Volume Changes

In the previous section, volume changes were neglected, but when water is removed, the volume of solution changes. Therefore, one needs to determine the volume of solution for

each amount of water that is removed.

Consider a solution similar to Witzke. This solution initially contains 7.88 M of lactic acid. For simplicity, the volume of this solution is 1 L, so there are 7.88 moles of lactic acid. Next, consider the number of moles of water initially. By knowing the molar density of pure water (55 M) and the molar density of pure lactic acid (13.38 M), one can compute the moles of pure water by using the volume additivity assumption that the volume of lactic acid plus the volume of water equals the volume of solution (i.e. an ideal solution). From this information, one can calculate the initial number of moles of water to be 22.5 moles.

Now, some number of moles of water is removed from solution. The volume of pure water in solution will change, but the volume of lactic acid will not. The total volume of solution will decrease and the concentrations of each species will start to increase. For example, if 20 moles of water were removed, the initial number of moles for lactic acid monomer, dimer, and water are 7.88 moles, 0 moles, and 2.5 moles, respectively. However, the volume of solution decreased to 0.633 L. Each molar amount is divided by this volume yielding the initial concentrations before equilibrium. The procedure occurs exactly as before and the conversion was 36%. For several water removals, the conversion ranged from 11.8% to 48.4% at K_{new} = 0.469. It is interesting that the maximum conversion achieved is the same as the dry solution which is sensible since all water was removed from the system to begin with.

But water can be further removed because as water is driven off, water as a product is formed through Le Chatelier's principle. There is a maximum conversion of approximately 90% when all of the initial water and stoichiometric product water is driven off.

There is a natural extension from this section to Figure 3.8. At a K_{new} = 0.469 when pH = 1.5, a $[W_0]$ of 22.5 M and a $[M_0]$ of 7.88 M assuming a 1 L solution, the conversion was 11.8%. The contour plot seems to corroborate the value since the experimental point (pH = 1.5, $[W_0]$ = 22.5 M) lies closer to the 0.1 contour line.

3.5 Conclusion

One key takeaway from this chapter is that polycondensation of PLA can proceed in aqueous, concentrated systems. The thermodynamics of the polycondensation is driven primarily by the removal of water from the system. Significant fractional conversions can be seen in both acidic and basic conditions provided that water is removed from the system. The reaction network is most favorable under partly anionic conditions. Achieving high conversions of long oligomers is a key objective when forming the protopeptides that are the precursor to modern biology.

The results presented in this chapter demonstrate the importance of accounting for dissociation states, temperatures, pK_a's, pH, and free energies in one cohesive, theoretical framework. Dissociation states are critical in the understanding of the underlying driving force of the ester bond linkage. The thermodynamic driving force stems from the strong pK_a shift from the monomer to the dimer. This insight is key to explaining how ester bonds can form at mild temperatures and acidic pH values. The framework can be expanded to include other prebiotic compounds and may have implications in various other fields including origins of life, sustainability, and medicine.

CHAPTER 4

PH DEPENDENT THERMODYNAMICS OF ESTER-AMIDE EXCHANGE INVOLVING GLYCOLIC ACID AND GLYCINE

4.1 Abstract

A robust pathway has been shown to produce proto-peptides under model prebiotic reaction using drying mixtures of hydroxy and amino acids [6, 18]. A promising mechanism of forming amide linkages is the ester-amide exchange reaction. The formation of lactoyllactic acid (aa) occurs spontaneously through wet-dry cycling at temperatures below 100°C at a pH near three [6, 18]. This observation suggests two conclusions: Either the Gibbs free energy of the exchange reaction under these conditions must be negative or Le Chatelier's principle allows better yields. However, very little thermodynamic data exists in the literature that validates either conclusion.

The purpose of this chapter is to demonstrate a first-principles thermodynamic model that computes key parameters, such as Gibbs free energy and equilibrium constant, and allows visualization of Gibbs free energy that shows the spontaneity of the exchange reaction as a function of temperature, hydration level, and pH. This model also shows the thermodynamic driving forces of the exchange reaction.

A key result from this chapter is that pK_a is a major determining factor in dictating the spontaneity and driving forces of the exchange step. The exchange reaction is strongly favorable at pH $\geq$ 2.5 while it is strongly unfavorable at pH $<$ 2. This result provides a prebiotic "recipe" in producing amide bonds under prebiotic conditions.

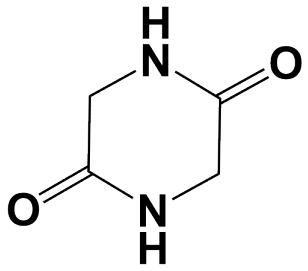

Figure 4.1: 2,5 Diketopiperazine (DKP)

4.2 Introduction

After an ester linkage is formed, it can be exchanged for an amide linkage in a process called ester-amide exchange. This process has been recently studied in the origins of life field because of its ability to form a peptide bond under plausibly prebiotic conditions and bypassing the formation of a 2,5 diketopiperazine (DKP), a cyclic amide that acts as a thermodynamic sink halting the amidation reaction [6, 13, 14, 22].

Sakata et al. [12] provides experimental data for glycine oligomerization kinetics at a variety of pH values and temperatures at atmospheric pressure. They also account for the dissociation states although they do not provide a framework for the Gibbs free energy of reaction of these oligomerizations. The gap with Sakata's paper is that the temperatures studied are extremely high. As a result, one of the main focuses of this chapter is on the thermodynamics of plausibly prebiotic amide bond formation in water through ester-amide exchange, at mild temperatures.

Kitadai et al. [9] performed a thorough analysis on the oligomerization of glycine monomers, dimers, and trimers. Kitadai used the Helgeson, Kirkham, Flowers method (HKF) [62–72], which is an empirical thermodynamic model based on electrostatic and structural forces among ions in aqueous solution. Thermodynamic data was compiled for specific dissociation states of glycine oligomers, and Gibbs free energy of reaction of oligomerization was calculated. Like Sakata, Kitadai did not report any exchange reaction thermodynamics involving hydroxy acids.

Kua and Sweet [13] performed DFT analyses on heterodimer formation from glycolic acid and glycine, but anionic species were not included. They only performed the reaction at one temperature and the pH is not well-defined. This chapter describes a larger spectrum of pH, temperature, and ionic states.

Yu [14] created an energy map of the thermodynamics of esterification and exchange between lactic acid and valine using experimental data. This formulation was performed only at one temperature and no explicit pH was considered. Yu used data from Shock [62] to acquire the thermodynamic parameters for his analysis. This chapter intends to expand on Yu's research with the same reasoning as Kua and Sweet: broaden the experimental conditions (pH, temperature, and ionic states) and determine if there are plausible prebiotic thermodynamic pathways for exchange to occur.

This literature review shows a need for a more accurate and detailed understanding of the thermodynamics of overall amidation of prebiotic compounds. To address this need, one must determine the thermodynamic driving force of amide bond linkages by calculating the Gibbs free energy of the ester-amide exchange reaction as a function of pH, hydration level, ionic state, and temperature.

The theoretical framework of the species' oligomerization reactions can be extended across different fields including drug delivery [41, 48] and origins of life chemistry [6, 17, 18, 20, 28, 73]. In drug delivery, peptide and depsipeptide oligomerization and self-assembly are crucial in producing nanomaterials for therapeutic medicines. In origins of life chemistry, the underpinnings of oligomerization thermodynamics of amino acids and peptides are one of the keys to solving the origin of life puzzle. The mechanisms of the important reactions in these fields depend on the thermodynamics of the oligomerization reactions outlined in the chapter.

4.3 Methods

4.3.1 DFT Model

See Chapter 3 for DFT explanation.

Unlike esterification from Chapter 3, the exchange reaction is an isodesmic reaction. Isodesmic reactions are chemical reactions in which the bonds broken in reactants are the same type as the bonds formed in products. This criterion is key since isodesmic reactions provide more accurate thermochemical data for DFT models [74–76].

For each species in a chemical reaction, the geometry is optimized such that energies are at a global minimum and that the field between solute and solvent is a self-consistent field. Each species is subject to a Cramer/Truhlar [55, 56] treatment such that the total energy is a summation of solvation energy and electrostatic energy. The correction free energy due to internal energy is computed and added to the total energy. From statistical mechanics, this adjusted free energy is used to calculate the relevant thermochemical properties such as Gibbs free energies and enthalpies.

4.3.2 Exchange Reaction Fundamentals

The fundamental mechanism of the exchange reaction is to trade an ester bond for an amide bond. It also serves as an alternate pathway to direct peptide bond formation [6, 18]. Generally, direct amidation is thermodynamically unfavorable because DKP acts as a "dead-end" and inhibits further polymerization [13, 15, 22].

The primary reason that DKP is a "dead-end" is that it is a cyclic amide consisting of two amide groups. Cyclic amides do not hydrolyze very easily and thus have great difficulty in ring opening. Exchange reactions do not form a cyclic amide but a cyclic "hybrid". This molecule is a cyclic dione containing an ester and an amide group which was the subject of study in Chapter 2. The ester group can easily hydrolyze and quickly ring open to form a heterodimer. Ester-amide exchange provides a critical pathway in forming amide bonds

under non-neutral conditions.

It is remarkable that Kua and Sweet [13] reported that direct amidation is favorable under neutral conditions through DFT. The key point is that amino acids are not dried in a closed system, but the ester-amide exchange reaction proceeds in differing hydration levels, considering Le Chatelier's principle. The reaction of interest is seen in Equation 4.1 and Figure 4.2. The homodimer gg and glycine react to form the heterodimer gG and glycolic acid monomer g.

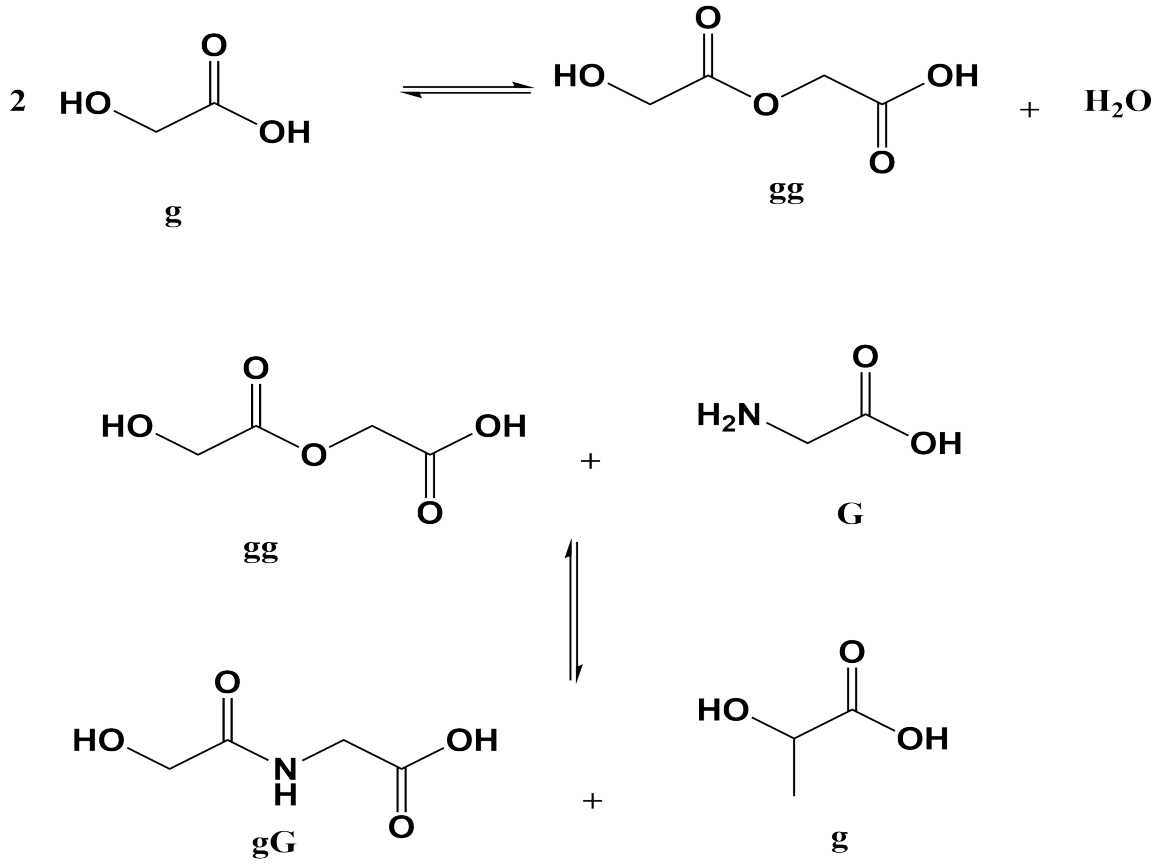

Figure 4.2: Scheme of exchange and esterification reactions of neutral species. Top: Esterification, Bottom: Exchange

$$x_1 gg + (1-x_1) gg^{(-)} + x_2 G^{(+)} + x_3 G^{(+/-)} + (1-x_2-x_3) G^{(-)} \leftrightarrow \\ x_4 gG + (1-x_4) gG^{(-)} + x_5 g + (1-x_5) g^{(-)} \tag{4.1}$$

where x_n is a mole fraction; g is neutral glycolic acid, gg is neutral glycolic acid dimer, gG is the neutral amide-linked heterodimer of glycolic acid and glycine; and G is glycine; $g^{(-)}$, $gg^{(-)}$, and $G^{(-)}$, represents the glycolic acid monomer, its dimer, and glycine in their anionic states, respectively; $G^{(+/-)}$ is the zwitterionic form of glycine; and $G^{(+)}$ is the cationic form of glycine.

Equation 4.1 would be valid at certain pH values higher than three, but at lower pH values, this equation would no longer be valid because the hydrogen atom balance is not satisfied. At pH < 3, the mole fraction of the cationic form of glycine will be highest, but gG and g cannot take a positive charge causing the unsatisfied hydrogen atom balance. Therefore, Equation 4.1 must be generalized to accommodate the missing positive charge. One way is to add water to one side and hydronium to the other as seen in Equation 4.2. This equation holds true if there are more hydrogen atoms in the reactants than in the products. The extra hydrogen atom on hydronium serves to counteract the excess hydrogen atoms on the reactant side.

$$\begin{aligned} \alpha H_2O + x_1 gg + (1-x_1)gg^{(-)} + x_2 G^{(+)} + x_3 G^{(+/-)} + (1-x_2-x_3)G^{(-)} \leftrightarrow \\ x_4 gG + (1-x_4)gG^{(-)} + x_5 g + (1-x_5)g^{(-)} + \alpha H_3O^{(+)} \end{aligned} \tag{4.2}$$

where α represents the stoichiometric coefficients of water and hydronium. The variable α cannot be greater than 1 because it represents the mole fraction of both water and hydronium in solution. The exact value for α is determined in the "Atom Balances" section and will be important in thermochemical calculations and contour plot formulations.

Another form of Equation 4.2 can occur if there are more hydrogen atoms on the products than in the reactants. Hydronium would be a reactant and water would be a product so that the additional hydrogen atoms counteract the excess hydrogen on the product side.

At pH > 3, Equation 4.2 will always balance and simplify to Equation 4.1.

4.3.3 Atom Balances

A critical aspect of modeling the exchange reaction is the atom balance. It is necessary to balance the carbon, hydrogen, nitrogen, and oxygen atoms on Equation 4.1 first. If the hydrogen balance is not fulfilled, then the correct form of Equation 4.2 is used and α is determined.

The carbon balance shows six carbon atoms balanced on each side. The nitrogen balance shows one nitrogen atom balanced on each side, and the oxygen balance shows seven oxygen atoms balanced on each side. Each balance is independent of x_n and the exchange reaction does not produce water. However, the hydrogen balance (Equation 4.3) tells a different story.

The hydrogen balance yields the following expression:

$$x_1 + 2x_2 + x_3 - x_4 - x_5 = 0 \tag{4.3}$$

If the left hand side of Equation 4.3 is not zero, the absolute value of the nonzero number is the α one uses in Equation 4.2. If the nonzero number is greater than zero, the reactants will have more hydrogen atoms than the products. Hydronium will be on the product side, and Equation 4.2 is used. If the nonzero number is less than zero, this means that the products will have more hydrogen atoms than the reactants, and Equation 4.2 is modified so that hydronium is on the reactant side. If Equation 4.3 is satisfied, Equation 4.1 is used.

4.3.4 Thermodynamic Model

The first step is to create a model that simulates a reaction network containing reactions in Table 4.2. All of the exchange reactions in Table 4.2 are favorable as these reactions occur at mildly acidic to basic pH. At highly acidic pH values, the exchange reaction becomes strongly unfavorable. Table 4.1 shows the formation energy data needed to calculate the energies.

Table 4.1: DFT Free Energies and Enthalpies of Formation of Pure Species in atomic units at T = 25° C and P = 1 atm. There are 2625.45 kJ/mol in 1 au.

Pure Species	**DFT ΔG_f° (au)**	**DFT ΔH_f° (au)**
H_2O	-76.469296	-76.447867
$H_3O^{(+)}$	-76.855396	-76.833186
$OH^{(-)}$	-75.974337	-75.954784
$G^{(+)}$	-284.950330	-284.914599
$G^{(+/-)}$	-284.514780	-284.475148
$G^{(-)}$	-284.058854	-284.023745
$GG^{(+)}$	-493.000055	-492.952826
$GG^{(+/-)}$	-492.560549	-492.513584
$GG^{(-)}$	-492.113992	-492.067418
gg	-532.318961	-532.272090
g	-304.393762	-304.358545
$g^{(-)}$	-303.944443	-303.909796
$gg^{(-)}$	-531.87457	-531.830309
gG	-512.446600	-512.399823
$gG^{(-)}$	-512.002367	-511.955637

Table 4.2: Several key reactions (direct amidation and exchange) and their Gibbs free energy of reaction.

Reaction	ΔG° (kJ/mol)
$G + G \Leftrightarrow GG + H_2O$	-0.75
$G^{(-)} + G^{(-)} \Leftrightarrow GG^{(-)} + OH^{(-)}$	128.47
$G^{(+)} + G \Leftrightarrow GG^{(+)} + H_2O$	45.13
$G^{(+)} + G \Leftrightarrow GG + H_3O^{(+)}$	187.39
$G^{(-)} + G \Leftrightarrow GG^{(-)} + H_2O$	30.91
$G^{(-)} + G \Leftrightarrow GG + OH^{(-)}$	153.07
$G^{(+)} + G^{(+)} \Leftrightarrow GG^{(+)} + H_3O^{(+)}$	177.01
$gg + G \Leftrightarrow gG + H_2O$	-17.38
$gg^{(-)} + G^{(-)} \Leftrightarrow gG^{(-)} + g^{(-)}$	-35.14
$gg + G^{(-)} \Leftrightarrow gG^{(-)} + g$	-48.08
$gg^{(-)} + G \Leftrightarrow gG^{(-)} + g$	-17.80

Before creating the model for calculating ΔG, assumptions of the model must be stated. They are as follows:

- It is a closed solution, meaning no loss of water or aqueous species to air.
- gg, gG, and g at equilibrium are made up of only the neutral and anionic species.

- G and GG at equilibrium are made up of cationic, zwitterionic, and anionic states.
- All dissociation states are a result of acid dissociation reactions.
- The only reactions in solution are the acid dissociation reactions and the reversible exchange reaction in Equation 4.1 and Equation 4.2.
- The volume of solution equals the volume of solvent plus the volume of the solute, i.e. volume additivity.

Based on these assumptions, several equations can be written that can describe the exchange system in Table 4.2 starting with Equation 4.4 and Equation 4.5. These equations represent the acid dissociation of exchange species i and yield species' ratios between the neutral and anionic dissociation states (or cationic dissociation states). The ratios are used to calculate the thermodynamic mole fractions.

$$pH = pK_{a,1(L_i)} + \log([{L_i}^{(-)}]/[L_i]) \tag{4.4}$$

$$pH = pK_{a,2(L_i)} + \log([L_i]/[{L_i}^{(+)}]) \tag{4.5}$$

where pH is the pH of solution; $[L_i]$ is the concentration of the neutral aqueous species i, $[L_i^{(-)}]$ is the concentration of the anionic aqueous species i, $[L_i^{(+)}]$ is the concentration of the cationic aqueous species i and $pK_{a,i}$ is the i^{th} pK_a of a species. α- hydroxy acids have one pK_a while amino acids have two pK_a's.

4.3.5 Calculate Mole Fractions

For every pH and dimer pK_a, the general form of the mole fraction of any dissociation state of a species is seen in Equation 4.6. The equation assumes three dissociation states (neutral, cationic, and anionic). One can expand the summation and derive an expression containing

a ratio between the concentrations of any pair of dissociation states using algebra. The mole fraction can then be calculated.

$$x_i = \frac{[M_i]}{\sum_{i=1}^{N}[M_i]} \tag{4.6}$$

where x_i is the mole fraction of the i^{th} dissociation state of species M and M_i is the concentration of the i^{th} dissociation state of species M. These mole fractions are inputs to the Gibbs free energy equations (and enthalpies) as shown in Equation 4.7 and Equation 4.9.

The pH and $pK_{a,gG}$ (heterodimer gG pK_a) parameters can be varied simultaneously in the model. The pH was varied from 0 to 10 and the $pK_{a,gG}$ was varied from 1 to 3 in MATLAB. The benchmark value of the homodimer gg pK_a is 2.74, which is the predicted pK_a of a similar molecule seen in Figure 4.3. The benchmark for the heterodimer gG is glycine (pK_a = 2.34). The evidence for this is the preliminary titration data showing that the pK_a is approximately equal to that of glycine.

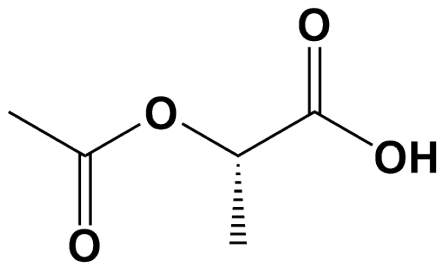

Figure 4.3: O-Acetyl Lactic Acid

4.3.6 Computing Gibbs Energies of Reactions

The Gibbs free energy is determined by the following equation:

$$\Delta G_r^\circ = \Delta G_{f,gG}^\circ + \Delta G_{f,g}^\circ + \Delta G_{water}^\circ - (\Delta G_{f,gg}^\circ + \Delta G_{f,G}^\circ) \tag{4.7}$$

where ΔG_r° is the standard Gibbs energy of reaction and each of the ΔG_f° terms are the standard Gibbs energy of formation of each species in reaction. The ΔG_f° values are found

in Table 4.1.

ΔG°_{water} is defined as follows:

$$\Delta G^{\circ}_{water} = \alpha(\Delta G^{\circ}_{f,H_2O} - \Delta G^{\circ}_{f,H_3O^{(+)}}) \tag{4.8}$$

Equation 4.8 is only valid if the hydrogen atom balance is greater than zero. Otherwise, the terms in the parentheses are reversed. ΔG°_{water} is needed to accommodate for any imbalances of hydrogen atoms. If the hydrogen balance equals zero, ΔG°_{water} equals zero because α equals zero.

At different pH values, the distribution of the exchange reaction species' dissociation states (and thus ΔG°_{r}) will change. These changes are reflected in the mole fractions (denoted x_i) and can be seen in the following modified form of Equation 4.7.

$$\begin{aligned} \Delta G^{\circ}_{r} = (x_4 \Delta G^{\circ}_{f,gG} + (1 - x_4)\Delta G^{\circ}_{f,gG^{(-)}} \\ + x_5 \Delta G^{\circ}_{f,g} + (1 - x_5)\Delta G^{\circ}_{f,g^{(-)}}) \\ + \Delta G^{\circ}_{water} - (x_1 \Delta G^{\circ}_{f,gg} + (1 - x_1)\Delta G^{\circ}_{f,gg^{(-)}} \\ + x_3 \Delta G^{\circ}_{f,G^{(+/-)}} + x_2 \Delta G^{\circ}_{f,G^{(-)}} + (1 - x_2 - x_3)\Delta G^{\circ}_{f,G^{(+)}}) \end{aligned} \tag{4.9}$$

The inputs required to calculate ΔG°_{r} of the exchange reaction are pH, pK_a, concentrations of dissociation states (e.g. $g^{(-)}$), and mole fractions of the dissociation states. Table 4.1 is consulted to acquire the Gibbs energies of formation of relevant exchange species. Finally, the Gibbs energies are calculated using Equation 4.7 and Equation 4.9.

4.3.7 Temperature Dependence

The third facet of the theoretical framework is to account for temperature since ΔG°_{r} are strongly dependent on temperature. This step is important because the temperatures cited in literature or in lab experiments may not be the same as DFT, so adjustments are necessary.

The van't Hoff equation (Equation 4.10) computes the equilibrium constants and is stated as follows:

$$\ln \frac{K_2}{K_1} = -\frac{\Delta H^\circ}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \tag{4.10}$$

where T_1 and T_2 are absolute temperatures, K_1 and K_2 are the equilibrium constants at T_1 and T_2, respectively, R is the universal gas constant, and ΔH° is the enthalpy change of reaction. The ΔH° value for the exchange reaction is calculated using the values in Table 4.1, Equation 4.7, and Equation 4.9 since enthalpy and Gibbs free energy are state functions. Equation 4.10 assumes that ΔH° is constant across the desired temperature range. The DFT calculated value for enthalpy at 25° is used here.

4.3.8 Estimating Unknown pK_a's

Banerjee [77] studied computational chemistry of α- hydroxycarboxylic acids (AHAs) and attempted to compute pK_a's of the OH groups for these compounds using complex thermodynamic cycles. The authors admit that "determination of pK_a values, especially for OH groups of α-hydroxycarboxylic acids, is challenging and rarely reported." This section focuses on a simpler method in predicting pK_a's inspired by [74]. This method is necessary because no published experimental data exists for gG, and pK_a is a critical parameter in computing the mole fractions of the species' dissociation states. However, there is some preliminary data performed at Georgia Tech that shows that the pK_a of gG is near that of glycine.

At first thought, one might directly compute the pK_a of the sample acid dissociation reaction in Equation 4.11 using the values in Table 4.1:

$$gG + H_2O \leftrightarrow gG^{(-)} + H_3O^{(+)} \tag{4.11}$$

The problem with this thinking stems from the fact that low acidity of hydroxyl groups makes experimental determination difficult and there are a lack of suitable test molecules

with known pK_a values that can probe a compound like gG [77]. The solution to this issue is to take advantage of one of DFT's greatest strengths: predicting thermochemical properties of isodesmic reactions.

It would be better to predict pK_a of gG using an isodesmic reaction. The method requires a reference material whose pK_a is well-known which will be the glycine (G) monomer. The procedure is as follows:

Step 1) The acid dissociation reaction for the reference material (Equation 4.12) is written.

$$G^{(+)} + H_2O \leftrightarrow G^{(+/-)} + H_3O^{(+)} \tag{4.12}$$

ΔG is computed from the known pK_a using Equation 4.13 and Equation 4.14. This value is denoted as ΔG_{ref}:

$$pK_a = -log(K_a) \tag{4.13}$$

$$\Delta G_{ref} = -RTln(K_a) \tag{4.14}$$

Step 2) The acid dissociation reaction for the target material (T) is written.

$$T + H_2O \leftrightarrow T^{(-)} + H_3O^{(+)} \tag{4.15}$$

If this compound has a known pK_a to compare, this value is denoted as the literature pK_a value ($pK_{a,ref}$).

Step 3) A reaction is written so that it yields Equation 4.15 when it is added to Equation 4.12. Hess's Law is used during this step:

$$G^{(+/-)} + T \leftrightarrow G^{(+)} + T^{(-)} \tag{4.16}$$

Step 4) The ΔG of Equation 4.16 is computed using values in Table 4.1 and is denoted

ΔG_{calc}. Equation 4.16 is an example of an isodesmic reaction, and DFT is most accurate when computing thermochemical properties from isodesmic reactions.

Step 5) ΔG_{calc} is added to ΔG_{ref}, and the sum is denoted ΔG_{target}.

Step 6) The pK_a is computed from ΔG_{target} using Equation 4.13 and Equation 4.14. This value is denoted $pK_{a,target}$.

Step 7) Finally, ΔpK$_a$ is computed by using Equation 4.17. Ideally, this value should be as close to zero as possible since it measures how close the target pK_a matches with the literature value. This step is only performed if the literature value of the pK_a for a particular species is known.

$$\Delta pK_a = |pK_{a,target} – pK_{a,ref}| \tag{4.17}$$

Unknown pK_a's

Some species of interest are found in Table 4.3. The parameter ΔpK_a decreases as the target species transitions from lactic acid to glycine and its derivatives. Lactic acid and glycine are chemically different suggesting that the method works best when both the reference and target species are similar. gG is functionally a hydroxy acid while glycine is an amino acid making glycine an inappropriate reference material. However, this method provides a starting point to the prediction of pK_a of compounds that have not been measured. Interestingly, if glycolic acid was the reference material, the target pK_a for gG would be 6.16.

Table 4.3: Comparison between target pK_a's and actual pK_a's of species. The temperature is 25°C. The reference species is the first ionization state of glycine where the pK$_a = 2.34$.

Target Species (T)	Target pK_a	Literature pK_a	ΔpK_a
Glycolic acid	1.49	3.83	2.34
Lactic acid	8.49	3.86	4.63
Glycine (pK_{a2})	11.71	9.6	2.11
Glycylglycine (pK_{a1})	4.16	3.140 [78]	1.02
Glycylglycine (pK_{a2})	7.40	8.265 [78]	0.86
Glycolic acid - Glycine (gG)	6.33	N/A	N/A

4.3.9 Conversions of Exchange Reactions

In Chapter 3, contour plots were constructed to show the relationship between conversion, pH, and initial water concentration. A similar procedure will occur in this chapter as well; however, the working reaction for exchange presents unique challenges that the working reaction for esterification did not have. The source of these challenges stems from the atom balances.

The equilibrium constant for the exchange reaction is as follows:

$$K_{eq} = \frac{[g][gG]}{[gg][G]} \tag{4.18}$$

where K_{eq} is the equilibrium constant for the exchange reaction and $[X]$ is the total equilibrium concentration of species X in molarity (M).

In Equation 4.2, the water and hydronium species were included to accommodate for any hydrogen atom imbalances, but these species are already lumped with K_{eq} since pH and α are known. Equation 4.18 contains no water terms because it is not produced in the exchange reaction, but exchange typically occurs in aqueous media. Without aqueous media or some type of proton donor (or acceptor), there are some pH ranges where exchange can never happen. From Equation 4.18, one can derive an equation that solves for the amount of reactant converted to product.

$$K_{eq} = \frac{([g_0] + x)([gG_0] + x)}{([G_0] - x)([gg_0] - x)} \tag{4.19}$$

where $[X_0]$ is the initial concentration of species X and x is the amount of reactant consumed.

Because Equation 4.19 is quadratic, it has two roots. The meaningful roots are extracted using MATLAB. They are real, positive, and must be less than or equal to the initial concentration of the limiting reactant. The variable x is not the conversion used in the contour plot because it is just the amount of reactant converted. To calculate the conversion, the

meaningful root is divided by the concentration of the limiting reactant.

4.4 Results/Discussion

4.4.1 ΔG_r° as a function of pH and dimer pK_a (Exchange)

Figure 4.4 shows how different gG pK_a's affect the thermodynamics of the exchange reaction. At pH < 3, hydronium must be added to accommodate for the positive charge on the glycine. The pK_a of the gG significantly dictates the thermodynamics of the exchange reaction, strongly motivating the experiments to be performed to determine it (See Appendix). At higher candidate pK_a values, the reaction is strongly spontaneous at a pH between 2–4. At lower candidate pK_a values, the reaction will not occur spontaneously at any pH range. This conclusion is independent of the exact value of the pK_a of gG.

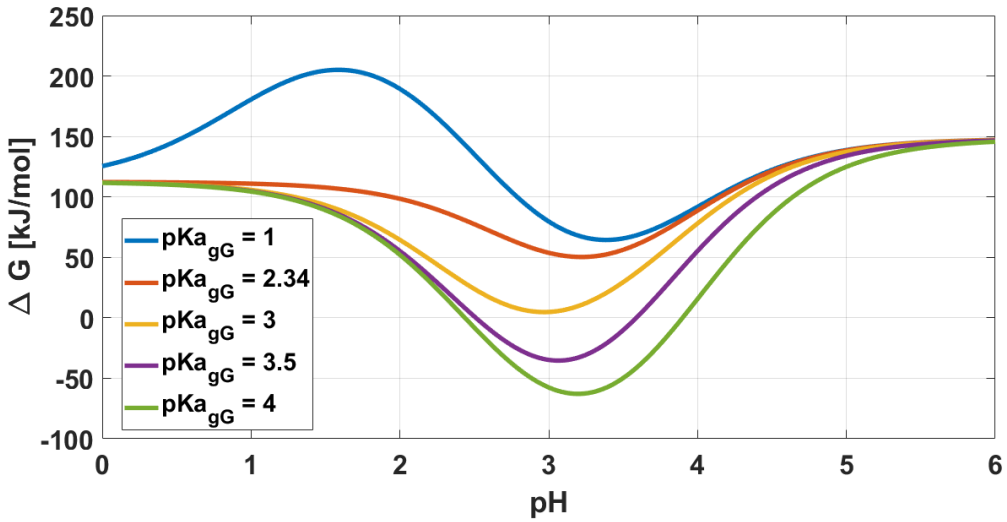

Figure 4.4: pH vs. ΔG_r° of gg + G $\Leftrightarrow$ gG + g for several gG pK_a values. $T = 25°C$. The pK_a of gg is 2.74.

Figure 4.5 shows how temperature affects the exchange reaction, and it demonstrates that temperature plays a very minor role in the thermodynamics of the ester-amide exchange reaction suggesting the reaction is thermoneutral.

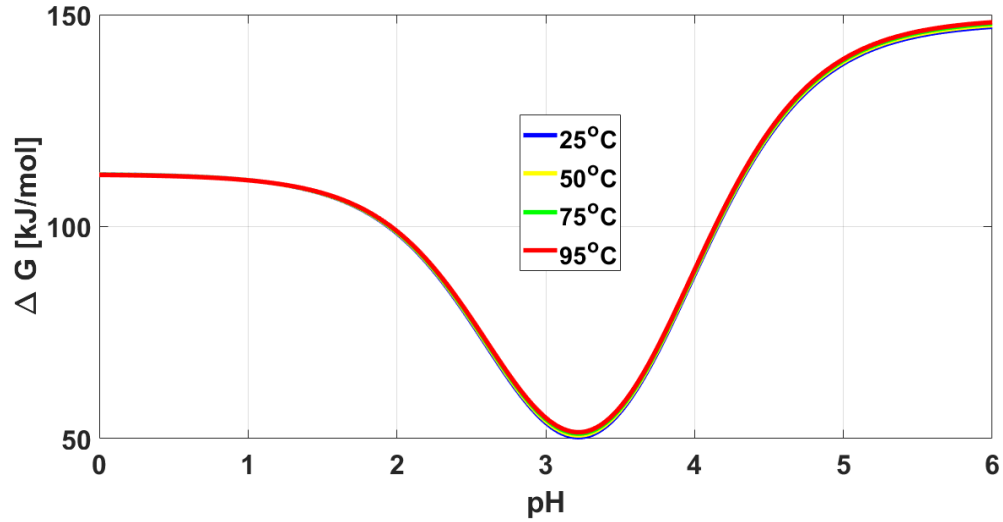

Figure 4.5: pH vs. ΔG_r° of gg + G $\Leftrightarrow$ gG + g for several temperatures. The pK_a of gg and gG are 2.74 and 2.34, respectively.

4.4.2 Conversion vs. pH

Figure 4.6 shows the conversion of the limiting reactant as a function of pH. Because the exchange reaction does not produce water, conversion is independent of initial water content. One can see a sharp increase in conversion at a pH value of around 3. Below this pH value, glycine takes a positive charge and hinders the reaction greatly, explaining the zero conversion. Once cationic glycine begins to convert to zwitterionic glycine, conversion increases substantially. However, the conversion is still miniscule due to a high positive Gibbs free energy.

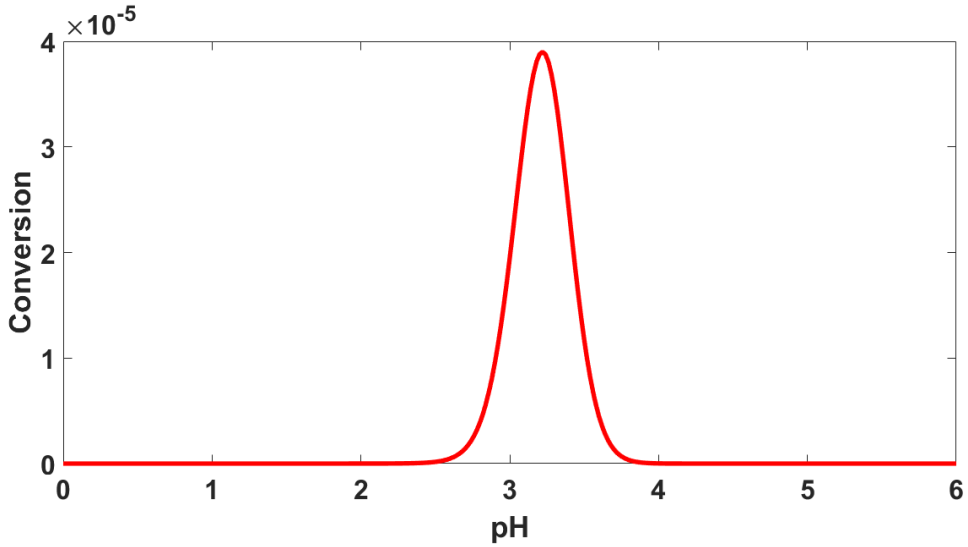

Figure 4.6: pH vs conversion for the exchange reaction. This model considers a 1:1 mixture of glycolic acid dimer and glycine and no product initially. $[gg_0]$ = $[G_0]$ = 1 M; pKa_{gg} = 2.74; pKa_{gG} = 2.34. The temperature is $25°C$.

4.5 Combined Esterification and Exchange

If the esterification reaction g + g ↔ gg + H_2O and the ester-amide exchange reaction gg + G ↔ gG + g are added, the direct amidation reaction g + G ↔ gG + H_2O is obtained. This reaction can be analyzed very similarly to the esterification reaction.

4.5.1 ΔG_r° as a function of pH and dimer pKa (Combined)

Figure 4.7 shows how the direct amidation reaction thermodynamics changes with pH by changing the gG pK_a. The Gibbs free energy of reaction is strongly dependent on pH and the heterodimer pK_a. At highly acidic pH values, the reaction is indeed unspontaneous corroborating the general consensus of the literature and points to the major challenge in prebiotic chemistry: forming peptide bonds under plausibly prebiotic conditions. The strong unfavorability of the combined reaction is primarily due to the cationic glycine rather than the hydronium. To make this determination, a set of hypothetical reactions is proposed:

Table 4.4: Several hypothetical combined reactions and their Gibbs free energy of reaction.

Reaction	ΔG° (kJ/mol)
$g + G^{(+/-)} \Leftrightarrow gG + H_2O$	-19.31
$g^{(-)} + G^{(+/-)} \Leftrightarrow gG^{(-)} + H_2O$	-32.66
$g^{(-)} + G^{(+)} \Leftrightarrow gG + H_2O$	1110.85
$g + G^{(+)} \Leftrightarrow gG + H_3O^{(+)}$	1276.8
$g + G^{(+/-)} \Leftrightarrow gG^{(-)} + H_3O^{(+)}$	133

According to Table 4.4, whenever cationic glycine is present and hydronium is not, the Gibbs free energy of reaction is greater than 1000 kJ/mol. If hydronium is present and cationic glycine is not, the Gibbs free energy is 133 kJ/mol. This observation demonstrates that cationic glycine has a much stronger correlation with the very high unfavorability of the combined reaction than hydronium.

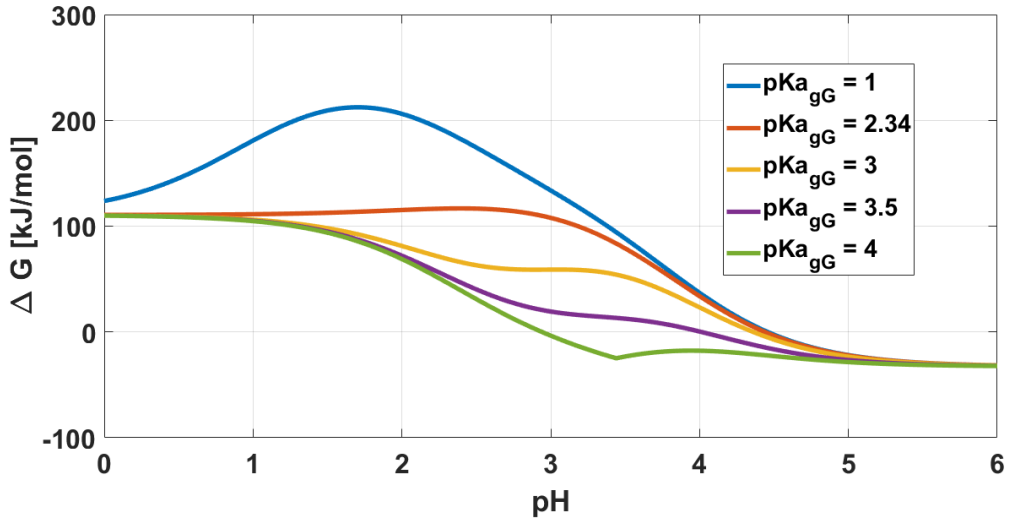

Figure 4.7: pH vs ΔG_r° of g + G $\Leftrightarrow$ gG + H_2O for several dimer pK_a values. $T = 25°C$.

According to Figure 4.7, amidation is strongly favored at pH between 2.2–3. This calculation casts doubt on the notion that direct amidation is always unfavorable regardless of early Earth conditions. The combined reaction is for a closed system initially containing no water. By removing water product, the reaction can be further promoted.

Figure 4.8 shows how temperature affects the exchange reaction. According to this figure, temperature does not noticeably affect the thermodynamics of the combined reaction

suggesting the combined reaction is thermoneutral.

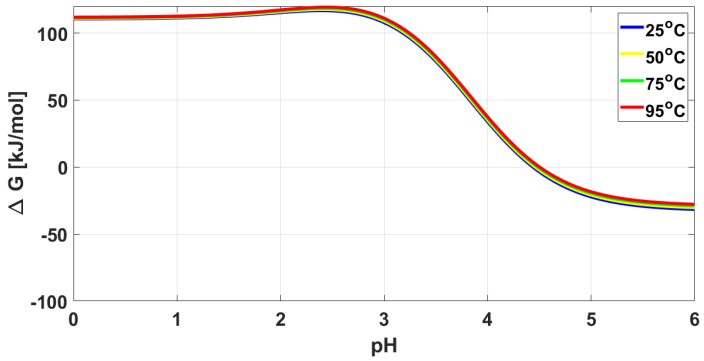

Figure 4.8: pH vs ΔG_r° of g + G $\Leftrightarrow$ gG + H_2O. The pK_a of gG is 2.34.

4.5.2 Impact of Water Concentration

The vertical contours at low pH suggest independence of water content, but the lines are slightly spread out more as pH slightly increases meaning that water content immediately becomes important when determining conversions. At dry conditions (low water content), conversion is highest.

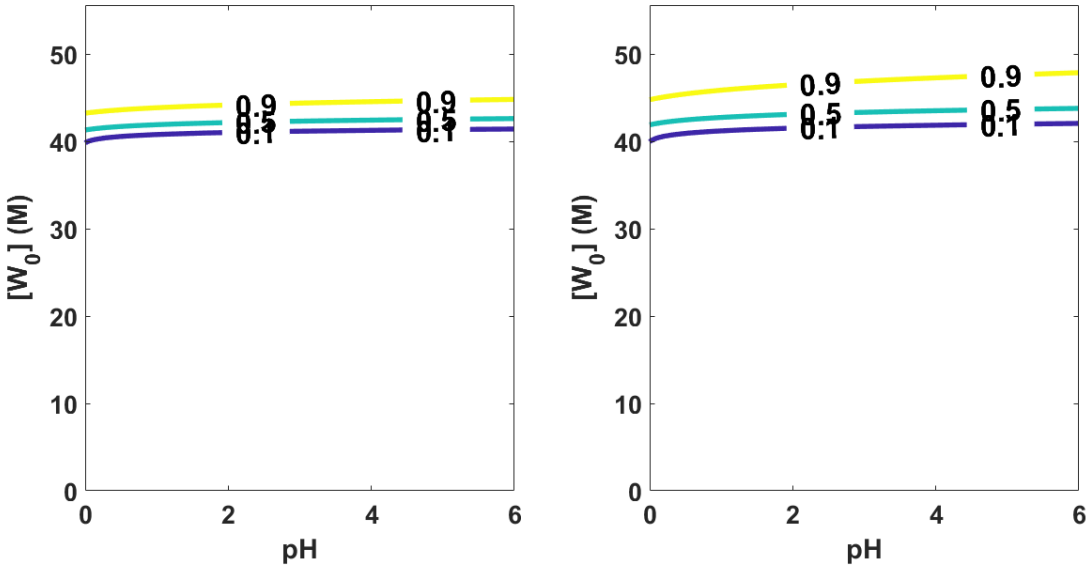

Figure 4.9: pH vs initial water content for the general reaction g + G $\leftrightarrow$ gG + H_2O. The colored curves represent the conversion. $[g_0]$ = $[G_0]$ = 1 M; pKa_{gG} = 3. Left: The temperature is $25°C$. Right: The temperature is $95°C$.

4.6 Conclusions

The results in this chapter show the importance of accounting for dissociation states, temperatures, pK_a's, pH, and free energies in one cohesive, theoretical framework. Dissociation states are critical in the understanding of the underlying driving force of the exchange reaction. The thermodynamic driving force is the pK_a shift from the homodimer to the heterodimer. This insight is key to explaining how amide bonds can form at mild temperatures and acidic pH values. The framework can be expanded to include other prebiotic compounds and will have implications in various other fields including origins of life and medicine.

CHAPTER 5
CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter will conclude my thesis by summarizing the key findings and insights from Chapters 2, 3, and 4 in relation to the research aims and research questions. It will also describe the value and contribution of my thesis and propose recommendations for future research.

5.2 Overall Findings

This thesis investigates the thermodynamics and kinetics of oligomerization and degradation of depsipeptides under prebiotic Earth conditions. The results indicate that the mechanisms of backbiting and scission are driven primarily by pH. Further results show that water content, dissociation states, and pK_a of homodimers and heterodimers play a critical role in the formation of depsipeptides on the prebiotic Earth.

There were some challenges during the investigation. One challenge is the limited amount of literature that exists for the thermodynamics and kinetics of hydroxy acids and amino acids as a function of pH. This thesis used the existing literature combined with DFT and modified chain mass balance models to determine dominant kinetic mechanisms of gA degradation, create concentration profiles for dipeptides, and calculate thermochemical properties of depsipeptides. Another challenge is the limited amount of experimental thermodynamic data critically needed for analysis of prebiotic reactions of depsipeptides. Witzke, Vu, and Holten provided formulations of weight percents and Gibbs free energies that made it possible to present model comparisons to the literature. The lack of a robust thermodynamic framework that is used to create free energy visualizations as a function of

pH for relevant prebiotic reactions presented another challenge. The thesis presents these visualizations that are used to solve questions that pre-existing literature could not answer before such as: What temperatures under 100°C induce favorable amide bond formation? What values of pH and initial water content does oligomerization become favorable? The answers to these questions point to a "recipe" for prebiotic peptide bond formation.

Chapter 2 describes the concentration profile of gA degradation and extracts its rate constants. Real-time spectral data from HPLC is collected at a pH from 2-8 and temperatures ranging from 25°C - 65°C. The spectral data can be converted to concentrations using Beer-Lambert's Law. A best-fit curve was fitted to the concentration data. This curve is a modified population balance model that describes the ideal degradation of dipeptides. Backbiting and scission rate constants were extracted from this curve, and backbiting was shown to be more dominant than scission at basic pH. The degradation rate constants followed a very clear Arrhenius relationship, and they suggest very strong entropic effects.

Chapter 3 describes esterification thermodynamics of lactic acid oligomers. Esterification proceeds in aqueous, concentrated systems, and the thermodynamics are driven primarily by the removal of water from the system as well as the pK_a shifts from the reactants to products. Atom balances provide a fundamental starting point for esterification analysis. Significant fractional conversions can be seen in both acidic and basic conditions provided the system is somewhat dry. These insights are key to explaining how ester bonds can form at mild temperatures and acidic pH values, which are standard prebiotic Earth conditions. The thermodynamic framework can be expanded to include other prebiotic compounds and will have implications in various other fields including origins of life and medicine.

Chapter 4 discusses ester-amide exchange reaction thermodynamics. Similar to Chapter 3, the thermodynamic driving force was shown to be dictated by the pK_a of the heterodimer gG. Atom balances were paramount when analyzing the exchange reaction. The introduction of glycine changed the thermodynamic formulation of the exchange reaction model because depending on the pH, the working reaction may or may not require water.

A new framework was developed to accommodate for hydrogen atom imbalances. There is a sigmoidal relationship between conversion and pH. Below pH = 2, the cationic glycine strongly hindered exchange while zwitterionic and anionic glycine facilitated exchange at pH > 2. Identification of pK_a of the linear dimers and heterodimers are critical in the understanding of the thermodynamic pathway of ester-amide exchange.

5.3 Value and Contribution of the Study

Although this study contributed feasible models for thermodynamic and kinetic analysis of plausibly prebiotic compounds, the big picture of this thesis goes beyond the models. This work touches many aspects of the chemical evolution field, including chemistry, astronomy, astrochemistry, and pharmaceutics, among others. The models presented in this thesis provide a fresh view on chemical evolution from a chemical engineering perspective.

This thesis fills in the literature gaps regarding the comprehensiveness of the thermochemical data. It challenges preconceived notions that direct peptide bond formation is always thermodynamically unfavorable. Chapter 4 demonstrates there are plausibly prebiotic experimental conditions where direct peptide bond formation can occur spontaneously.

The research findings presented in this thesis are useful in the real world. For instance, in pharmaceutics and drug delivery, polylactic acid can be used as a capsule for the drug to reach its intended target safely. Knowledge of polylactic acid stability and thermodynamics helps ensure that the capsule works as intended.

5.4 Future Work

I propose a few recommendations for future study.

The first recommendation is to experimentally determine what the unknown pK_a's are. Chapters 3 and 4 make it clear that these values are critical to know in understanding the thermodynamics of both esterification and ester-amide exchange. The esterification reaction networks primarily involved monomer, dimer, and trimer. What would the results

look like for tetramers and longer oligomers? They reveal motivation behind determining the pK_a of heterodimers and longer oligomers.

The second recommendation is to collect data on the thermochemical properties of the reactions other than DFT, if possible. DFT is a valid alternative when no real data is possible, but the values that DFT outputs give a guide on predicting the thermochemistry of reactions. Witzke [23] shows that it can be done with his studies on lactic acid oligomerization. One should follow his example and conduct similar studies for other prebiotic molecules such as glycolic acid and glycine. Kua and Sweet [13] have only performed DFT studies on these molecules, not experimental. The difficult task is experimentally determining the thermochemical data of oligomers.

It is hopeful that this data will be obtained, and the insights gained through chemical modeling will lead to a better understanding of the reaction mechanisms behind novel applications in the origins of life, green chemistry, pharmaceutics, and catalysis fields. According to van Nostrum [19], the knowledge of degradation rates of systems similar to those described in this thesis allows one to more accurately predict and control their hydrolysis rates for better drug capsule engineering. Today, industries are transitioning to a greener and more environmental way of manufacturing. Polylactic acid and their derivatives are popular for precisely this purpose, so knowledge of their properties are critical for chemical companies.

The third recommendation is addressing the kinetic polymerization of OH dipeptides such as gA. In Chapter 2, a population balance model was presented to create concentration profiles of degradation of gA oligomers, but polymerization was neglected. A good starting point is a paper by Moritz [45]. Moritz studied the influence on viscosity in heterogeneous and homogeneous polymerizations and presented empirical correlations. By incorporating these correlations, one may gain a good understanding in how heat and mass transfer could be used in polymerization.

5.5 Closing Summary

My thesis provides a comprehensive thermodynamic and kinetic framework that explains the hydration level dependence, pH, and temperature dependence on esterification and ester-amide exchange. The value of this work extends from astrobiology to pharmaceutics and its contribution stems from the thermodynamic and kinetic models addressing the missing gaps in the literature. Although there are a few limitations to the study, the results provide exciting opportunities for future work through determination of pK_a of unknown species and exploration of depsipeptide polymerization kinetics. I am hopeful that this thesis provides a strong foundation for more exciting research on depsipeptide oligomer thermodynamics and kinetics in the origins of life, pharmaceutics, and catalysis fields.

Appendices

.1 pH vs Lactic Acid Concentration

Figure 1 shows how pH changes with lactic acid monomer concentration in water. To construct this plot, one needs to create a system of equations. This system contains two acid dissociation equations (one for lactic acid and one for water), a charge balance, and a mass balance. Once solved, the hydronium concentration is obtained which can easily be used to find pH.

It is interesting to note that the more concentrated lactic acid, the lower the pH until it stagnates at around pH = 3.1.

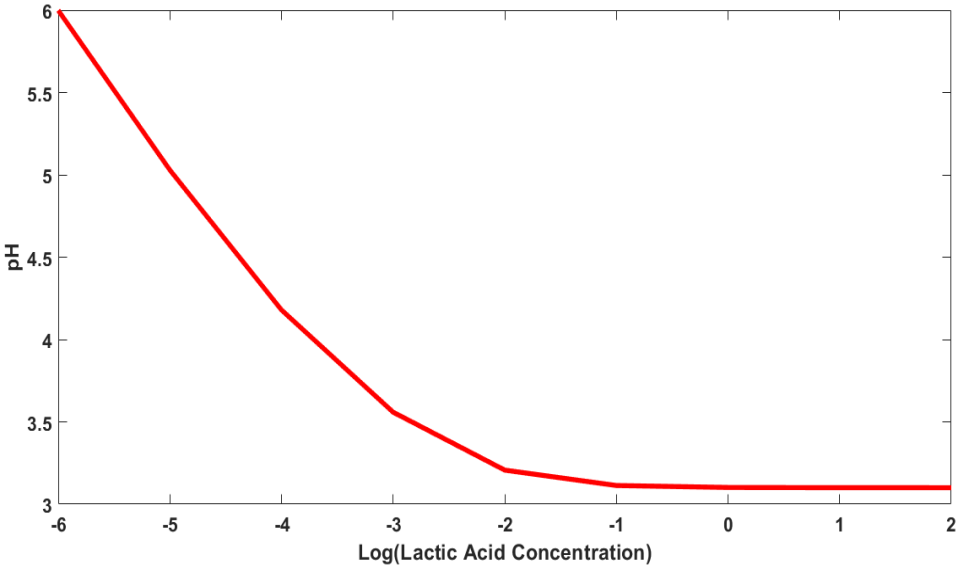

Figure 1: Semi-log plot of lactic acid concentration and pH.

.2 pH vs Glycolic Acid and Glycine

Figure 2 shows how pH changes with two hypothetical solutions: pure glycolic acid and a 1:1 mixture of glycolic acid and glycine. To construct this plot, one needs to create a system of equations. This system contains four acid dissociation equations (one for glycolic acid, two for glycine, and one for water), a charge balance, and two mass balances. Once solved, the hydronium concentration is obtained which can easily be used to find pH.

Like Figure 1, it is interesting to note that the more concentrated lactic acid, the lower the pH until it stagnates at around pH = 3.1. Adding glycine to the mixture leads to an overall lower pH than pure glycolic acid.

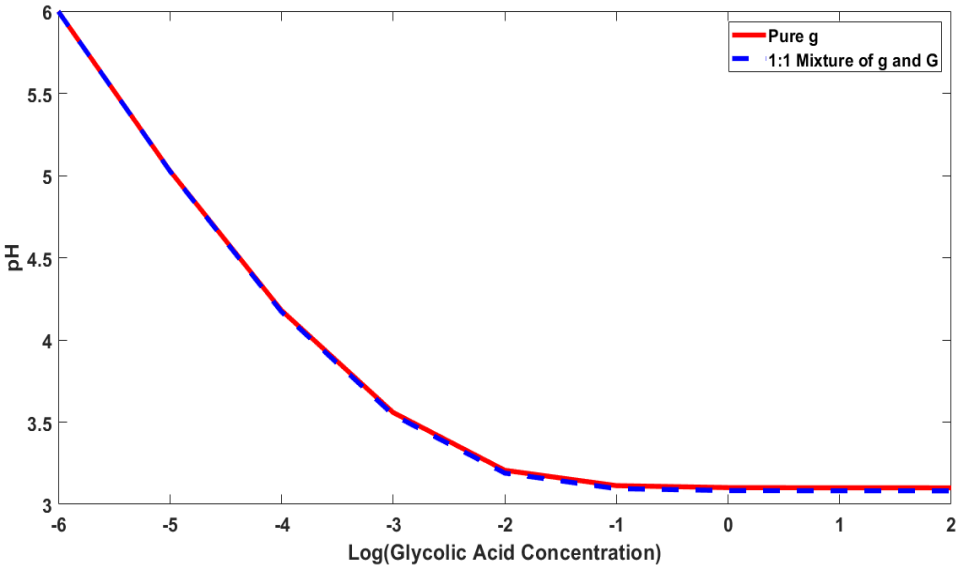

Figure 2: Semi-log plot of glycolic acid concentration and pH.

.3 DFT Coordinates for Selected Molecules

The free energies and enthalpies of formation for a series of glycolic acid, lactic acid, glycine and alanine monomer, dimers, cyclic dimers and several trimers.

Kohn-Sham density functional theory (DFT- B3LYP) has been used to calculate the free energies, the enthalpies of formation and the geometries for the following series of chiral hydroxy acids- lactic acid, the linear and cyclic diastereomeric dimers of lactic acid and the diastereomeric trimers of lactic acid- employing Spartan software (Spartan '16, Wavefunction, Inc., Irvine, CA). The basis set used in these calculations was 6-311+G(2df,2p). The most stable conformation associated with each of the species was determined by sequential rotation about each of the carbon-carbon single bonds in search of the most stable conformational structure. Cartesian coordinates are used to describe (in Angstroms) the ground state geometries for each of the molecules.

.3.1 Esterification Molecules

This subsection lists the DFT coordinates for molecules used in Chapter 3.

Figure 3: SS lactic acid dimer neutral. The labels represent specific atoms and correspond to the atoms located in Table 1

Table 1: Cartesian Coordinates (in angstroms) for the SS lactic dimer neutral

Atom	X	Y	Z
1 C C1	-1.6463358	0.0488635	-1.5469885
2 C C2	-0.6001617	-0.6057762	-0.6558536
3 O O2	-0.6786790	-1.7312504	-0.2238834
4 O O5	0.4133927	0.2303348	-0.3869796
5 C C3	1.4660396	-0.2517280	0.4659881
6 H H7	1.7935648	-1.2242618	0.0994364
7 C C4	0.9471632	-0.4067465	1.8959693
8 O O3	-0.0113349	0.1770869	2.3314653
9 O O1	1.6620715	-1.2144119	2.6911206
10 H H3	2.3855608	-1.6564343	2.2231012
11 C C5	-2.5112777	1.0140781	-0.7337808
12 H H1	-1.9081811	1.8112057	-0.3017493
13 H H10	-3.0210849	0.4844281	0.0721352
14 H H11	-3.2589395	1.4563985	-1.3898659
15 C C6	2.5987781	0.7643889	0.4022794
16 H H6	3.4238068	0.4501023	1.0394768
17 H H12	2.2525401	1.7443370	0.7275442
18 H H13	2.9609408	0.8347356	-0.6216025
19 H H14	-1.1242845	0.5975741	-2.3302758
20 O O4	-2.4160159	-0.9412195	-2.2049198
21 H H2	-2.7275633	-1.5817050	-1.5526175

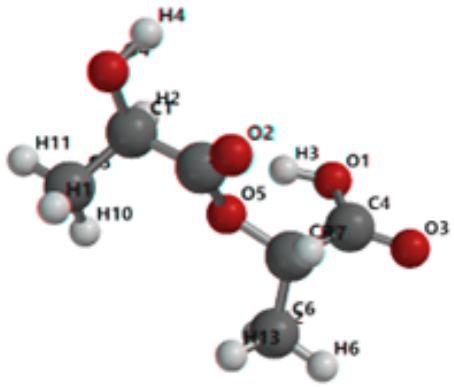

Figure 4: SR lactic acid dimer neutral. The labels represent specific atoms and correspond to the atoms located in Table 2

Table 2: Cartesian Coordinates (in angstroms) for the SR lactic dimer neutral

Atom	X	Y	Z
1 C C1	0.4754327	1.3679857	-1.5438013
2 H H2	0.0136275	2.1717244	-0.9623894
3 C C2	-0.1675596	0.0626074	-1.0799053
4 O O2	-0.6197645	-0.7917642	-1.7963670
5 O O5	-0.1625247	-0.0036572	0.2728751
6 C C3	-0.6701109	-1.1918906	0.9204803
7 H H7	-0.8525596	-1.9569300	0.1705409
8 C C4	-2.0088184	-0.8863994	1.5892991
9 O O3	-2.6903904	-1.7606086	2.0655952
10 O O1	-2.3918188	0.3908481	1.6514706
11 H H3	-1.7323151	0.9666945	1.2320003
12 C C5	1.9792690	1.3545752	-1.2892426
13 H H1	2.4522608	0.5490210	-1.8511867
14 H H10	2.1930778	1.2212449	-0.2313435
15 H H11	2.4022503	2.3036545	-1.6139933
16 C C6	0.3576374	-1.6617217	1.9417896
17 H H6	-0.0145733	-2.5479791	2.4513714
18 H H12	0.5535302	-0.8840684	2.6787062
19 H H13	1.2857641	-1.9139002	1.4326816
20 O O4	0.2708354	1.5695511	-2.9288957
21 H H4	-0.6732501	1.6410127	-3.1096855

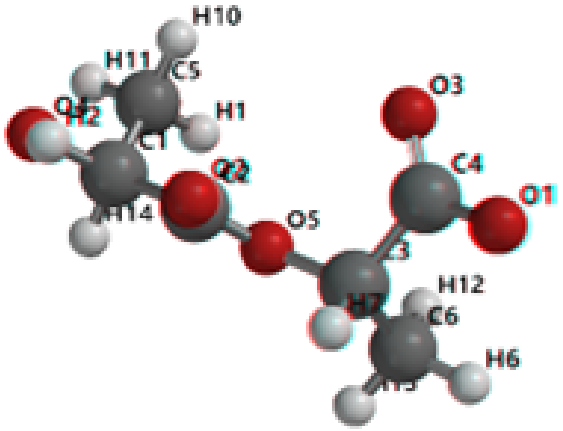

Figure 5: SS lactic acid dimer anion. The labels represent specific atoms and correspond to the atoms located in Table 3

Table 3: Cartesian Coordinates (in angstroms) for the SS lactic dimer anionic

Atom	X	Y	Z
1 C C1	-1.5943756	-0.0518969	-1.4068447
2 C C2	-0.5029852	-0.7071536	-0.5681332
3 O O2	-0.5105890	-1.8959568	-0.3138897
4 O O5	0.4349181	0.1467367	-0.1977064
5 C C3	1.5344584	-0.3463814	0.6223939
6 H H7	1.8689152	-1.2896950	0.1958921
7 C C4	1.0994641	-0.6015840	2.0900991
8 O O3	0.0134013	-0.1224713	2.4828732
9 O O1	1.9314396	-1.2631322	2.7587458
10 C C5	-2.3126648	1.0706441	-0.6635943
11 H H1	-1.6214074	1.8738365	-0.4147274
12 H H10	-2.7665081	0.6963402	0.2541674
13 H H11	-3.0972562	1.4737966	-1.3023149
14 C C6	2.6277306	0.7048959	0.5303978
15 H H6	3.4983375	0.3732483	1.0922341
16 H H12	2.2873008	1.6547603	0.9438194
17 H H13	2.9245671	0.8611642	-0.5063448
18 H H14	-1.1019312	0.3633808	-2.2925024
19 O O4	-2.5227384	-1.0459003	-1.8165858
20 H H2	-2.1900766	-1.8946318	-1.4879792

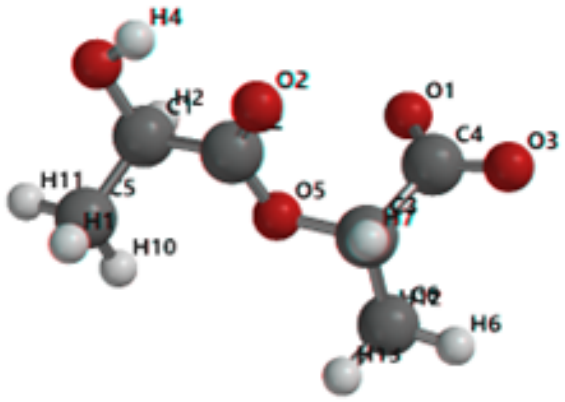

Figure 6: SR lactic acid dimer anion. The labels represent specific atoms and correspond to the atoms located in Table 4

Table 4: Cartesian Coordinates (in angstroms) for the SR lactic dimer anionic

Atom	X	Y	Z
1 C C1	0.3961362	1.5281462	-1.3803793
2 H H2	0.1165386	2.3430909	-0.7043229
3 C C2	-0.3286299	0.2843820	-0.8799859
4 O O2	-1.1403869	-0.3089817	-1.5628149
5 O O5	0.0263115	-0.0501989	0.3484183
6 C C3	-0.6364539	-1.1899293	0.9707029
7 H H7	-0.6746052	-1.9909061	0.2360182
8 C C4	-2.0869711	-0.8507909	1.4043755
9 O O3	-2.7935756	-1.8579450	1.6560981
10 O O1	-2.4136949	0.3529496	1.4982323
11 C C5	1.9135320	1.3664041	-1.3705968
12 H H1	2.2129749	0.5184951	-1.9872183
13 H H10	2.2801590	1.2149534	-0.3571881
14 H H11	2.3710309	2.2691878	-1.7730050
15 C C6	0.2286162	-1.5855230	2.1554008
16 H H6	-0.2044399	-2.4549305	2.6459450
17 H H12	0.2913045	-0.7727931	2.8799118
18 H H13	1.2364047	-1.8396650	1.8278635
19 O O4	-0.0571544	1.8376336	-2.6899122
20 H H4	-0.7370969	1.1864208	-2.9175429

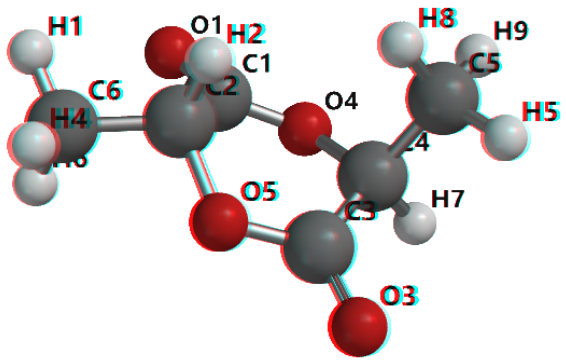

Figure 7: Cyclic Lactic Acid (RS). The labels represent specific atoms and correspond to the atoms located in Table 5

Table 5: Cartesian Coordinates (in angstroms) for Cyclic Lactic Acid (RS)

Atom	X	Y	Z
1 C C1	-0.3594081	-0.4207935	-1.3124469
2 O O1	-0.5554739	-0.5329854	-2.4961370
3 O O4	-0.0441491	-1.4964562	-0.5826422
4 C C2	-0.4353668	0.9080238	-0.5763942
5 H H2	0.5792418	1.3120038	-0.5330647
6 O O5	-0.8970370	0.7282883	0.7906272
7 C C3	-0.4440267	-0.2915387	1.5246164
8 O O3	-0.6814112	-0.3446072	2.7057181
9 C C4	0.3506478	-1.3672386	0.8085184
10 H H7	0.0637934	-2.3155210	1.2542998
11 C C5	1.8560920	-1.1661430	0.9336550
12 H H5	2.1362183	-1.1870426	1.9857107
13 H H8	2.1770419	-0.2176822	0.5051708
14 H H9	2.3665389	-1.9791491	0.4208360
15 C C6	-1.3619127	1.9057231	-1.2353556
16 H H1	-1.0120946	2.1200122	-2.2421150
17 H H4	-1.3620171	2.8304473	-0.6619073
18 H H6	-2.3766768	1.5146591	-1.2890895

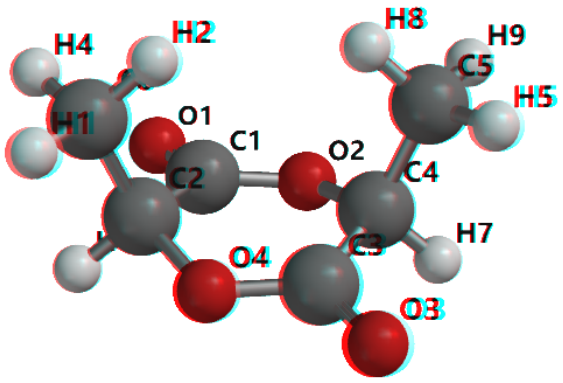

Figure 8: Cyclic Lactic Acid (SS). The labels represent specific atoms and correspond to the atoms located in Table 6

Table 6: Cartesian Coordinates (in angstroms) for Cyclic Lactic Acid (SS)

Atom	X	Y	Z
1 C C1	-0.6923874	-0.3013943	-1.4429546
2 O O1	-0.7743819	-0.3808265	-2.6446651
3 O O2	-0.4470693	-1.3927255	-0.7195258
4 C C2	-0.8807301	1.0245300	-0.7342765
5 O O4	-1.1886777	0.8995266	0.6769302
6 C C3	-0.7977068	-0.1454521	1.4050402
7 O O3	-0.9476535	-0.1356854	2.6027322
8 C C4	-0.1887690	-1.3447622	0.7061733
9 H H7	-0.7081764	-2.2155331	1.1009731
10 C C5	1.3064100	-1.4739280	0.9689964
11 H H5	1.4824821	-1.5411925	2.0409107
12 H H8	1.8545977	-0.6204783	0.5727712
13 H H9	1.6733857	-2.3831523	0.4969392
14 H H10	-1.7746104	1.4714301	-1.1639854
15 C C6	0.3077730	1.9584060	-0.9280116
16 H H2	1.2150586	1.5437815	-0.4911547
17 H H1	0.0900923	2.9142564	-0.4554150
18 H H4	0.4703630	2.1231996	-1.9914778

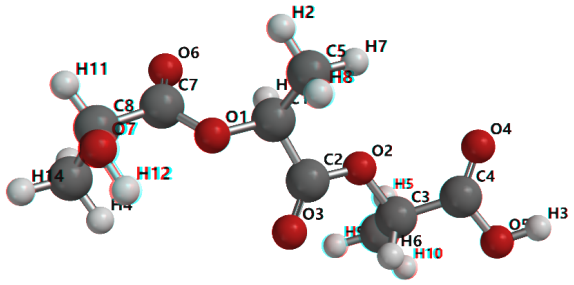

Figure 9: RRS neutral lactic acid trimer. The labels represent specific atoms and correspond to the atoms located in Table 7

Table 7: Cartesian Coordinates (in angstroms) for the RRS lactic acid trimer (neutral)

Atom	X	Y	Z
1 H H1	0.7564410	1.1914943	0.7946656
2 C C1	-0.1602043	0.9940157	0.2407511
3 O O1	0.0431175	1.3386441	-1.1411993
4 C C2	-0.4185223	-0.5067567	0.2915419
5 O O2	-0.5447551	-0.8926774	1.5680776
6 O O3	-0.4932878	-1.2407644	-0.6609538
7 C C3	-0.7821392	-2.2868106	1.8357134
8 H H6	-1.2934935	-2.7309949	0.9831062
9 C C4	-1.7199348	-2.3500106	3.0324332
10 O O4	-2.0737238	-1.4134147	3.7033927
11 O O5	-2.1041490	-3.6132235	3.2511548
12 H H3	-2.6814758	-3.6419173	4.0313417
13 C C5	-1.3111917	1.8166984	0.8070773
14 H H2	-1.0646278	2.8746436	0.7367016
15 H H7	-1.4722933	1.5663598	1.8529793
16 H H8	-2.2286335	1.6281564	0.2506663
17 C C6	0.5384946	-3.0039338	2.0967846
18 H H5	1.0488275	-2.5667526	2.9540748
19 H H9	1.1758659	-2.9153826	1.2194110
20 H H10	0.3579054	-4.0588853	2.2908555
21 C C7	1.0501207	2.1646011	-1.4535547
22 O O6	1.8302002	2.6279392	-0.6541363
23 C C8	1.1375308	2.4291954	-2.9558629
24 H H11	1.4603432	3.4645354	-3.0554120
25 O O7	-0.1173659	2.3514878	-3.6105412
26 H H12	-0.4246735	1.4373576	-3.6034578
27 C C9	2.1817513	1.5034617	-3.5818517
28 H H4	1.8615256	0.4629646	-3.5095559
29 H H13	3.1437910	1.6083487	-3.0816747
30 H H14	2.3045561	1.7616205	-4.6325284

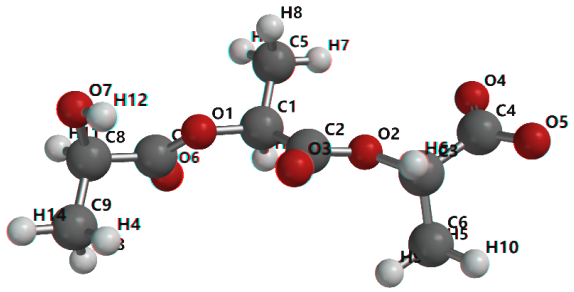

Figure 10: RRS anionic lactic acid trimer. The labels represent specific atoms and correspond to the atoms located in Table 8

Table 8: Cartesian Coordinates (in angstroms) for the RRS lactic acid trimer (anionic)

Atom	X	Y	Z
1 H H1	0.6488229	1.0832134	0.9367981
2 C C1	-0.2641691	0.8778866	0.3808150
3 O O1	-0.0595963	1.2328019	-1.0006679
4 C C2	-0.5160006	-0.6276089	0.4302792
5 O O2	-0.6220004	-1.0172734	1.6951631
6 O O3	-0.6047148	-1.3453356	-0.5393855
7 C C3	-0.8828887	-2.4238098	1.9834389
8 H H6	-1.3922493	-2.8475524	1.1217168
9 C C4	-1.8264662	-2.5398168	3.2064633
10 O O4	-2.0749317	-1.5211752	3.8884136
11 O O5	-2.2491643	-3.7076945	3.4018888
12 C C5	-1.4227645	1.6919788	0.9422268
13 H H2	-1.1835487	2.7523999	0.8824768
14 H H7	-1.5914070	1.4311561	1.9843580
15 H H8	-2.3349012	1.5022655	0.3771076
16 C C6	0.4484250	-3.1268888	2.2128446
17 H H5	0.9719463	-2.6879949	3.0628628
18 H H9	1.0829467	-3.0452604	1.3308846
19 H H10	0.2738296	-4.1810152	2.4180609
20 C C7	0.9606124	2.0381919	-1.3148806
21 O O6	1.7542453	2.4873850	-0.5196440
22 C C8	1.0500924	2.2989045	-2.8179732
23 H H11	1.3660259	3.3360043	-2.9217668
24 O O7	-0.2015110	2.2095258	-3.4779290
25 H H12	-0.5118254	1.2968517	-3.4472999
26 C C9	2.1026227	1.3770865	-3.4363087
27 H H4	1.7922651	0.3343819	-3.3531485
28 H H13	3.0636680	1.4957751	-2.9372612
29 H H14	2.2226365	1.6256171	-4.4895339

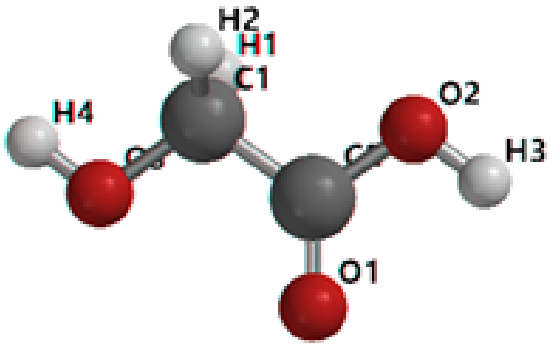

Figure 11: Neutral glycolic acid. The labels represent specific atoms and correspond to the atoms located in Table 9

Table 9: Cartesian Coordinates (in angstroms) for Neutral Glycolic Acid

Atom	X	Y	Z
1 H H1	0.1175681	0.4709346	-1.3673834
2 C C1	0.2393769	0.5649420	-0.2838138
3 H H2	1.3125857	0.6162765	-0.0765018
4 C C2	-0.2913339	-0.6983960	0.3495889
5 O O1	-1.1675346	-0.7622506	1.1765444
6 O O2	0.3503886	-1.7717790	-0.1360816
7 H H3	-0.0054572	-2.5717854	0.2822396
8 O O3	-0.4551019	1.6791719	0.2347256
9 H H4	-0.1004918	2.4728860	-0.1793178

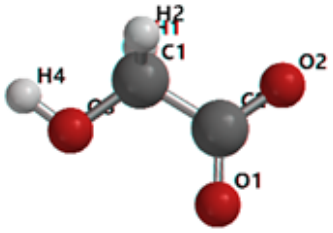

Figure 12: Anionic glycolic acid. The labels represent specific atoms and correspond to the atoms located in Table 10

Table 10: Cartesian Coordinates (in angstroms) for Anionic Glycolic Acid

Atom	X	Y	Z
1 H H1	0.1154550	0.1690463	-1.3325062
2 C C1	0.2336475	0.2371867	-0.2472211
3 H H2	1.3039850	0.3183510	-0.0364843
4 C C2	-0.2598542	-1.0863235	0.3612116
5 O O1	-1.1702256	-1.0646883	1.2206480
6 O O2	0.3314243	-2.1054187	-0.0854510
7 O O3	-0.4657200	1.3726034	0.2635838
8 H H4	-0.0887122	2.1592430	-0.1437809

Figure 13: Anionic glycolic acid dimer. The labels represent specific atoms and correspond to the atoms located in Table 11

Table 11: Cartesian Coordinates (in angstroms) for Anionic glycolic acid dimer

Atom	X	Y	Z
1 H H1	1.5798158	1.6801263	-0.5438569
2 C C1	0.7038851	1.5027456	-1.1751379
3 H H2	-0.0729861	2.2090930	-0.8686836
4 O O1	1.0104996	1.6237379	-2.5497902
5 H H4	1.3404644	2.5147189	-2.7062265
6 C C2	0.1992287	0.1043944	-0.8883128
7 O O2	-0.0793910	-0.0211298	0.4086630
8 O O3	0.0646187	-0.7711771	-1.7122897
9 C C3	-0.5815568	-1.2983453	0.8631428
10 H H5	0.1537149	-2.0686327	0.6363309
11 H H6	-1.4994554	-1.5320246	0.3253957
12 C C4	-0.8643828	-1.2928069	2.3755187
13 O O4	-0.6675184	-0.2510251	3.0372769
14 O O5	-1.2869369	-2.3996748	2.7979695

Figure 14: Anionic glycolic acid trimer. The labels represent specific atoms and correspond to the atoms located in Table 12

Table 12: Cartesian Coordinates (in angstroms) for Anionic glycolic acid trimer

Atom	X	Y	Z
1 H H1	0.6177201	-1.0016863	0.8338413
2 C C1	0.3094754	-0.3469070	0.0190062
3 H H2	0.9082907	0.5633775	0.0446845
4 O O1	0.4364494	-1.0088926	-1.2430874
5 C C2	-1.1473561	0.0312796	0.2097445
6 O O2	-1.2949101	0.5860996	1.4075484
7 O O3	-2.0255753	-0.1511848	-0.5970381
8 C C3	-2.5948855	1.1155718	1.7516433
9 H H5	-2.5909154	1.1773212	2.8365376
10 H H6	-3.3679329	0.4197160	1.4351758
11 C C4	-2.8760932	2.5137104	1.1603345
12 O O4	-1.9920666	3.0593717	0.4662918
13 O O5	-4.0062675	2.9662098	1.4679396
14 C C5	1.6270397	-1.4079159	-1.6994392
15 O O6	1.6901694	-1.9741465	-2.7676057
16 C C6	2.8663172	-1.1470296	-0.8699973
17 H H7	2.7467589	-1.6256971	0.1068233
18 H H8	2.9624944	-0.0710721	-0.6979349
19 O O7	4.0142776	-1.6403044	-1.5211288
20 H H4	3.7170097	-2.0578213	-2.3433394

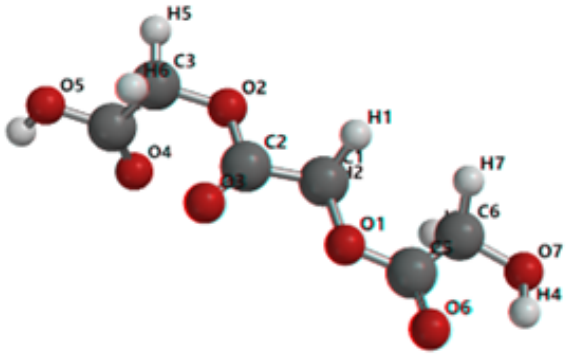

Figure 15: Glycolic acid trimer. The labels represent specific atoms and correspond to the atoms located in Table 13

Table 13: Cartesian Coordinates (in angstroms) for Glycolic acid trimer

Atom	X	Y	Z
1 H H1	0.8192510	-1.1607860	0.7992809
2 C C1	0.5275373	-0.5088547	-0.0242345
3 H H2	1.1346499	0.3962254	0.0036955
4 O O1	0.6522651	-1.1788851	-1.2795462
5 C C2	-0.9243123	-0.1179824	0.1487588
6 O O2	-1.0754554	0.4696319	1.3452282
7 O O3	-1.8079677	-0.3012414	-0.6449874
8 C C3	-2.3804626	0.9312282	1.6872691
9 H H5	-2.3943698	1.0326288	2.7700122
10 H H6	-3.1428432	0.2142365	1.3880023
11 C C4	-2.6713147	2.2784771	1.0556328
12 O O4	-1.9116686	2.8940945	0.3517019
13 C C5	1.8427824	-1.5910374	-1.7308529
14 O O6	1.9005375	-2.1702428	-2.7911283
15 C C6	3.0819709	-1.3231563	-0.9042745
16 H H7	2.9586656	-1.7831723	0.0808929
17 H H8	3.1828693	-0.2444068	-0.7521709
18 O O7	4.2288157	-1.8331162	-1.5434853
19 H H4	3.9343937	-2.2636882	-2.3597161
20 O O5	-3.8926913	2.6958661	1.4000393
21 H H3	-4.0626529	3.5641813	0.9998821

.3.2 Exchange Molecules

This subsection lists the DFT coordinates for molecules used in Chapter 4.

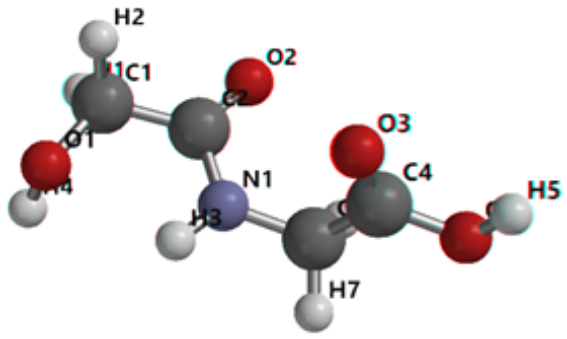

Figure 16: Glycolic acid-Glycine. The labels represent specific atoms and correspond to the atoms located in Table 14

Table 14: Cartesian Coordinates (in angstroms) for Glycolic acid-Glycine

Atom	X	Y	Z
1 H H1	-0.5769064	1.5110120	-2.7518040
2 C C1	-0.2448664	1.6720546	-1.7249519
3 H H2	0.6874394	2.2362289	-1.7619329
4 O O1	-1.1801724	2.4392514	-0.9823548
5 H H4	-2.0726086	2.2036381	-1.2574719
6 C C2	0.0841501	0.3088664	-1.1265349
7 O O2	0.8532716	-0.4531378	-1.7124485
8 N N1	-0.5074898	0.0227807	0.0451460
9 H H3	-1.0785036	0.7407936	0.4641243
10 C C3	-0.2790891	-1.2049223	0.7580178
11 H H6	-0.1446477	-2.0266257	0.0552881
12 H H7	-1.1486529	-1.4402975	1.3707694
13 C C4	0.9396571	-1.1627762	1.6652260
14 O O3	1.6619607	-0.2109867	1.8316284
15 O O4	1.1133413	-2.3431779	2.2756885
16 H H5	1.8931168	-2.2927015	2.8516104

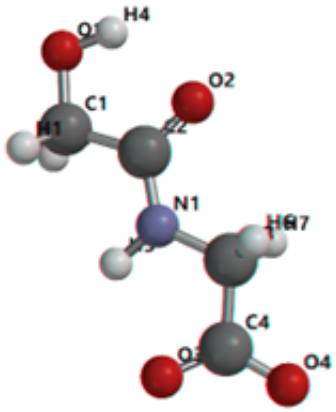

Figure 17: Glycolic acid-Glycine Anion. The labels represent specific atoms and correspond to the atoms located in Table 15

Table 15: Cartesian Coordinates (in angstroms) for Glycolic acid-Glycine (Anion)

Atom	X	Y	Z
1 H H1	0.3920360	2.3044794	-1.0318217
2 C C1	-0.4101086	1.6464483	-1.3772378
3 H H2	-1.3374024	1.9667628	-0.8959352
4 O O1	-0.5366092	1.7218003	-2.7835540
5 H H4	-0.3577410	0.8200631	-3.1017958
6 C C2	-0.0929393	0.2105265	-0.9622852
7 O O2	0.0429835	-0.6533717	-1.8393574
8 N N1	0.0223032	-0.0246210	0.3422350
9 H H3	-0.1036241	0.7175058	1.0195899
10 C C3	0.3258209	-1.3188869	0.9239192
11 H H6	1.2840186	-1.6915957	0.5576583
12 H H7	-0.4320676	-2.0535287	0.6459921
13 C C4	0.3856844	-1.2303339	2.4636076
14 O O3	0.1643992	-0.1104995	2.9872128
15 O O4	0.6532460	-2.3047488	3.0517722

Figure 18: Neutral glycine. The labels represent specific atoms and correspond to the atoms located in Table 16

Table 16: Cartesian Coordinates (in angstroms) for Neutral glycine

Atom	X	Y	Z
1 H H1	0.2562786	0.3854138	-1.4799129
2 C C1	0.3685993	0.4661066	-0.3968731
3 H H2	1.4420884	0.5345760	-0.2097919
4 N N1	-0.2960070	1.6609739	0.0829762
5 H H4	-1.2915086	1.6142068	-0.0992060
6 H H5	-0.1866654	1.7536378	1.0859089
7 C C2	-0.1079970	-0.8462169	0.2022139
8 O O1	-0.9786973	-0.9657811	1.0325650
9 O O2	0.5634239	-1.8918883	-0.3089973
10 H H3	0.2304852	-2.7110287	0.0911174

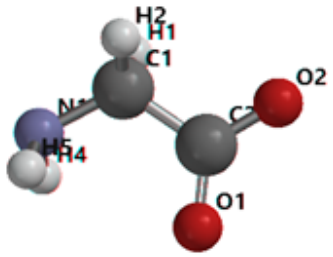

Figure 19: Anionic glycine. The labels represent specific atoms and correspond to the atoms located in Table 17

Table 17: Cartesian Coordinates (in angstroms) for Anionic glycine

Atom	X	Y	Z
1 H H1	0.2810866	0.1088619	-1.4774125
2 C C1	0.3990762	0.1635793	-0.3931343
3 H H2	1.4719690	0.2559299	-0.2115139
4 N N1	-0.2675177	1.3709460	0.0940622
5 H H4	-1.2636979	1.2996977	-0.0808598
6 H H5	-0.1770700	1.4213687	1.1027228
7 C C2	-0.0472318	-1.2001555	0.1867612
8 O O1	-0.9633548	-1.2161955	1.0459224
9 O O2	0.5667403	-2.2040326	-0.2665480

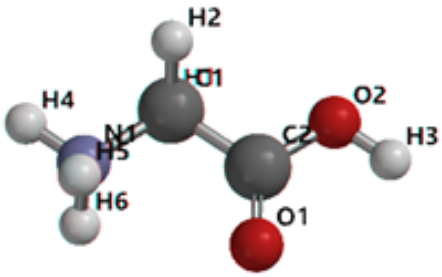

Figure 20: Cationic glycine. The labels represent specific atoms and correspond to the atoms located in Table 18

Table 18: Cartesian Coordinates (in angstroms) for Cationic glycine

Atom	X	Y	Z
1 H H1	0.2845957	0.1831131	-1.4858922
2 C C1	0.4179095	0.2135914	-0.4072712
3 H H2	1.4764715	0.3324843	-0.1884846
4 C C2	-0.1071094	-1.0554874	0.2237013
5 O O1	-0.9818875	-1.0830549	1.0535144
6 O O2	0.5269757	-2.1182132	-0.2613564
7 H H3	0.1805670	-2.9261613	0.1525261
8 N N1	-0.3135196	1.3905621	0.1291839
9 H H4	0.0411179	2.2533178	-0.2853572
10 H H5	-0.2119119	1.4705016	1.1421608
11 H H6	-1.3132089	1.3393463	-0.0727251

Figure 21: Zwitterionic glycine. The labels represent specific atoms and correspond to the atoms located in Table 19

Table 19: Cartesian Coordinates (in angstroms) for Zwitterionic glycine

Atom	X	Y	Z
1 H H1	0.0209123	-0.2239818	-1.5235611
2 C C1	0.3747816	-0.0896089	-0.5062510
3 H H2	1.4439994	0.0964462	-0.5235070
4 C C2	0.0614893	-1.3434601	0.3432963
5 O O1	-0.6817897	-1.1563892	1.3397523
6 O O2	0.5796521	-2.4057530	-0.0499193
7 N N1	-0.3177405	1.0904225	0.1028801
8 H H3	0.3257416	1.8196738	0.4018620
9 H H4	-0.8014936	0.7006430	0.9325424
10 H H5	-1.0055523	1.5120076	-0.5170949

.3.3 Other Molecules

This subsection lists the DFT coordinates and thermochemical data for molecules not used in either chapter.

Figure 22: Cyclic Lactic Acid-Glycine. The labels represent specific atoms and correspond to the atoms located in Table 20

Table 20: Cartesian Coordinates (in angstroms) for cyclic Lactic Acid-Glycine

Atom	X	Y	Z
1 C C1	-0.4616588	0.0403010	-1.6924857
2 O O1	-0.5215148	0.0109483	-2.8981164
3 O O2	-0.2007760	-1.0794609	-1.0124224
4 C C2	-0.6751961	1.3113536	-0.9043064
5 H H2	0.2501396	1.8931172	-0.9642835
6 H H4	-1.4523655	1.8814383	-1.4062921
7 N N1	-1.0687226	1.0659534	0.4709308
8 H H3	-1.5589434	1.8000806	0.9624071
9 C C3	-0.6911133	-0.0063804	1.1787461
10 O O3	-0.9345136	-0.1644902	2.3718981
11 C C4	0.0924002	-1.0635672	0.4148004
12 H H7	-0.2592663	-2.0307633	0.7635693
13 C C5	1.5930939	-0.9337777	0.6348907
14 H H5	1.8140128	-1.0182402	1.6978370
15 H H8	1.9639346	0.0264013	0.2759598
16 H H9	2.1104892	-1.7329139	0.1068671

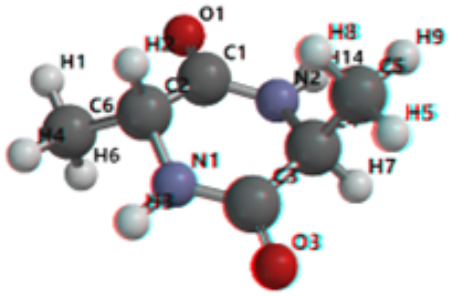

Figure 23: Cyclic Alanine Dimer (RS). The labels represent specific atoms and correspond to the atoms located in Table 21

Table 21: Cartesian Coordinates (in angstroms) for Cyclic Alanine Dimer (RS)

Atom	X	Y	Z
1 C C1	-0.1819977	-0.2417330	-1.3934008
2 O O1	-0.2679748	-0.2938365	-2.6201997
3 N N2	0.1082490	-1.3158453	-0.6424341
4 C C2	-0.3850974	1.0782866	-0.6562222
5 H H2	0.5522917	1.6318768	-0.7859863
6 N N1	-0.6301323	0.8777126	0.7712553
7 H H3	-1.0191554	1.6611750	1.2789984
8 C C3	-0.2523204	-0.1707070	1.5170552
9 O O3	-0.4047397	-0.2130418	2.7386109
10 C C4	0.3994358	-1.3332313	0.7882132
11 H H7	-0.0521971	-2.2349264	1.2010744
12 C C5	1.9093561	-1.3623723	1.0661686
13 H H5	2.0931107	-1.4435566	2.1361059
14 H H8	2.3897111	-0.4583506	0.6909037
15 H H9	2.3557422	-2.2243950	0.5715235
16 C C6	-1.5186246	1.8978313	-1.2661246
17 H H1	-1.3090797	2.1041175	-2.3120059
18 H H4	-1.6107451	2.8474637	-0.7389710
19 H H6	-2.4661046	1.3640896	-1.1930859
20 H H14	0.2902725	-2.1705574	-1.1514786

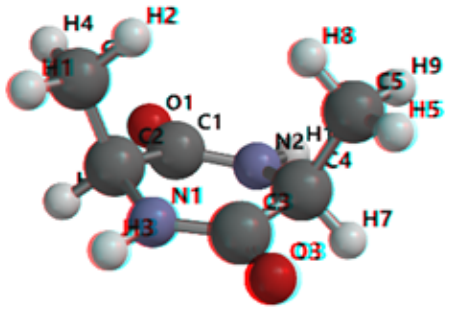

Figure 24: Cyclic Alanine Dimer (SS). The labels represent specific atoms and correspond to the atoms located in Table 22

Table 22: Cartesian Coordinates (in angstroms) for Cyclic Alanine Dimer (SS)

Atom	X	Y	Z
1 C C1	-0.6059818	-0.2366126	-1.4702765
2 O O1	-0.6906329	-0.2777386	-2.6987806
3 N N2	-0.3436207	-1.3174508	-0.7216679
4 C C2	-0.8045961	1.0829340	-0.7407940
5 N N1	-1.0580351	0.8931001	0.6851251
6 H H3	-1.4236855	1.6901618	1.1889942
7 C C3	-0.6910118	-0.1538590	1.4379368
8 O O3	-0.8259140	-0.1783948	2.6622568
9 C C4	-0.0954811	-1.3529706	0.7171746
10 H H7	-0.6208756	-2.2215395	1.1150632
11 C C5	1.3987596	-1.4935602	1.0388990
12 H H5	1.5436185	-1.5835177	2.1139034
13 H H8	1.9562241	-0.6307117	0.6749566
14 H H9	1.7932513	-2.3895407	0.5607129
15 H H14	-0.1948824	-2.1795848	-1.2290849
16 H H10	-1.7029604	1.5234664	-1.1736780
17 C C6	0.3775515	2.0286582	-0.9943896
18 H H2	1.3028044	1.6098310	-0.5987972
19 H H1	0.1909772	2.9840885	-0.5051921
20 H H4	0.4944910	2.2032409	-2.0623620

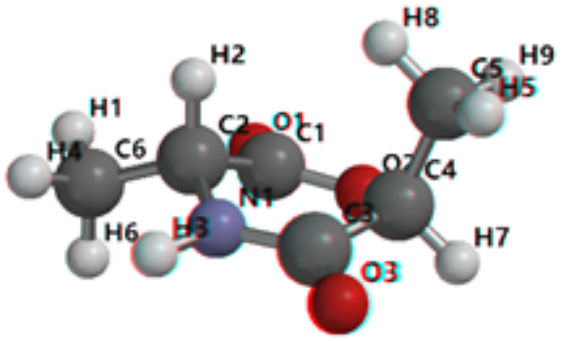

Figure 25: Cyclic Alanine Lactic Acid Dimer (trans). The labels represent specific atoms and correspond to the atoms located in Table 23

Table 23: Cartesian Coordinates (in angstroms) for Cyclic Alanine Lactic Acid Dimer (trans)

Atom	X	Y	Z
1 C C1	-0.1842090	-0.4149323	-1.3995409
2 O O1	-0.2952594	-0.5017657	-2.5991886
3 O O2	0.1423153	-1.5025418	-0.6929611
4 C C2	-0.3843197	0.8909628	-0.6419874
5 H H2	0.5944473	1.3868733	-0.6517803
6 N N1	-0.7906426	0.6306413	0.7370332
7 H H3	-1.2927958	1.3630079	1.2210049
8 C C3	-0.3669456	-0.4071002	1.4712075
9 O O3	-0.6018873	-0.5460353	2.6696615
10 C C4	0.4535287	-1.4482961	0.7273467
11 H H7	0.1445344	-2.4217085	1.0994773
12 C C5	1.9503922	-1.2570333	0.9293340
13 H H5	2.1833841	-1.3079373	1.9919592
14 H H8	2.2839681	-0.2936862	0.5437893
15 H H9	2.4913153	-2.0503833	0.4162815
16 C C6	-1.3974226	1.7981656	-1.3303195
17 H H1	-1.0592076	2.0443228	-2.3334098
18 H H4	-1.4994058	2.7239073	-0.7645794
19 H H6	-2.3717901	1.3135389	-1.3933281

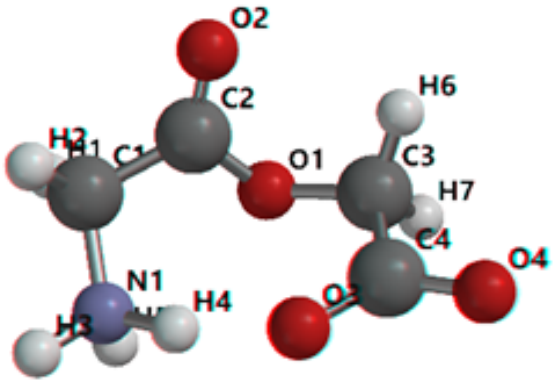

Figure 26: Glycine-glycolic acid (Zwitterionic). The labels represent specific atoms and correspond to the atoms located in Table 24

Table 24: Cartesian Coordinates (in angstroms) for Glycine-glycolic acid (Zwitterionic)

Atom	X	Y	Z
1 H H1	-1.3522345	1.0584657	-2.0019357
2 C C1	-0.6923193	1.2289065	-1.1541575
3 H H2	-0.2758688	2.2278746	-1.2176112
4 C C2	0.4404412	0.2202591	-1.1349009
5 O O2	1.5554557	0.4537013	-1.5256652
6 O O1	0.0307421	-0.9356923	-0.6005293
7 C C3	0.9650414	-1.6545578	0.2305154
8 H H6	1.9500220	-1.6764053	-0.2297499
9 H H7	0.5803728	-2.6668725	0.3085595
10 C C4	1.0442060	-1.0074210	1.6302595
11 O O3	0.3804575	0.0501490	1.8331976
12 O O4	1.7766658	-1.5994160	2.4425183
13 N N1	-1.4702579	1.1067562	0.1181750
14 H H3	-1.8564218	2.0057791	0.4001433
15 H H4	-0.8300680	0.7341556	0.8772481
16 H H5	-2.2462343	0.4543180	0.0239330

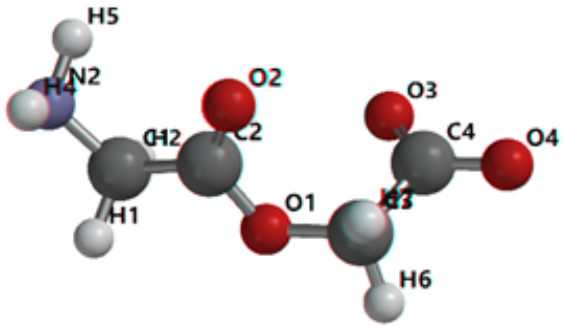

Figure 27: Glycine-glycolic acid (Anionic). The labels represent specific atoms and correspond to the atoms located in Table 25

Table 25: Cartesian Coordinates (in angstroms) for Glycine-glycolic acid (Anionic)

Atom	X	Y	Z
1 H H1	-0.6622941	2.2064227	0.8105602
2 C C1	-1.2202263	1.3212694	0.4996460
3 H H2	-1.5149129	0.8092718	1.4177485
4 C C2	-0.2514679	0.4026411	-0.2333092
5 O O2	-0.4256216	-0.0409659	-1.3469997
6 O O1	0.8273512	0.1565466	0.5151137
7 C C3	1.8147489	-0.7583241	-0.0008448
8 H H6	2.7233454	-0.5321519	0.5520781
9 H H7	1.9831674	-0.5606508	-1.0567149
10 C C4	1.4655320	-2.2486338	0.1943145
11 O O3	0.4089771	-2.5478226	0.7910918
12 O O4	2.3286980	-3.0299064	-0.2788532
13 N N2	-2.4020149	1.7184146	-0.2431066
14 H H4	-2.1406398	2.2002906	-1.0954571
15 H H5	-2.9346423	0.9035988	-0.5252672

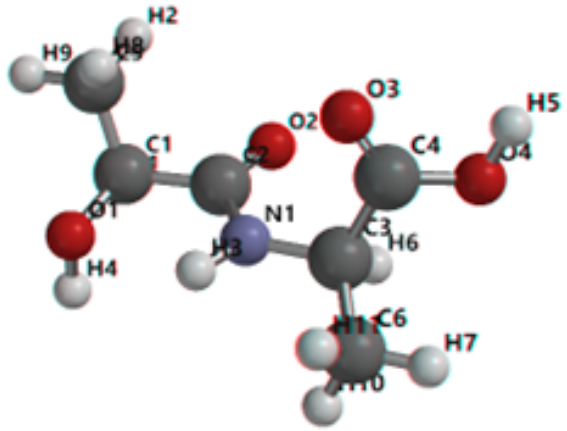

Figure 28: Lactic Acid Alanine. The labels represent specific atoms and correspond to the atoms located in Table 26

Table 26: Cartesian Coordinates (in angstroms) for Lactic Acid Alanine

Atom	X	Y	Z
1 H H1	-0.2915512	1.2435780	-2.8298593
2 C C1	-0.0983586	1.4695193	-1.7785188
3 O O1	-1.2194990	2.1445417	-1.2094929
4 H H4	-2.0306600	1.7652221	-1.5652579
5 C C2	0.1900235	0.1236087	-1.1033962
6 O O2	1.0081028	-0.6582094	-1.5904543
7 N N1	-0.4893415	-0.1291182	0.0278317
8 H H3	-1.1089056	0.5888071	0.3724597
9 C C3	-0.3233537	-1.3507031	0.7836112
10 H H6	-0.2818152	-2.1887546	0.0880872
11 C C4	0.9950616	-1.3521732	1.5565445
12 O O3	1.6231708	-0.3720250	1.8759808
13 O O4	1.3661454	-2.6010077	1.8717259
14 H H5	2.1809727	-2.5678127	2.3983936
15 C C5	1.1125569	2.3870716	-1.6678412
16 H H2	1.9988425	1.8937322	-2.0614779
17 H H8	1.2935598	2.6530711	-0.6257784
18 H H9	0.9371381	3.2987883	-2.2369814
19 C C6	-1.4970131	-1.5494224	1.7484422
20 H H7	-1.3783103	-2.4769417	2.3034815
21 H H10	-2.4269468	-1.5980312	1.1844566
22 H H11	-1.5598190	-0.7237407	2.4580432

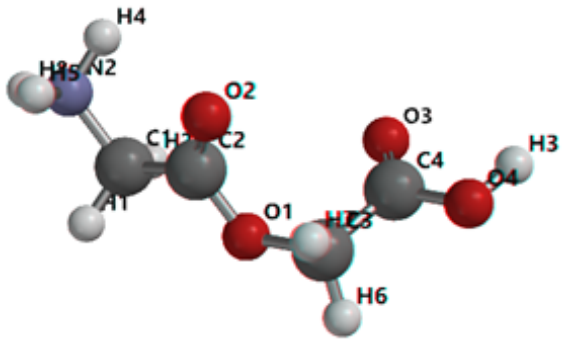

Figure 29: Cationic Glycine-glycolic acid. The labels represent specific atoms and correspond to the atoms located in Table 27

Table 27: Cartesian Coordinates (in angstroms) for Cationic Glycine-glycolic acid.

Atom	X	Y	Z
1 H H1	-0.7804129	1.8734672	1.5956427
2 C C1	-1.2083494	1.1995227	0.8578272
3 H H2	-1.7818699	0.4322532	1.3731637
4 C C2	-0.1210461	0.5623839	0.0242532
5 O O2	-0.0247822	0.6814692	-1.1704814
6 O O1	0.7026357	-0.1372396	0.8046908
7 C C3	1.7835644	-0.8247595	0.1710587
8 H H6	2.5050552	-1.0391514	0.9556439
9 H H7	2.2544105	-0.2003454	-0.5856070
10 C C4	1.3178246	-2.1264448	-0.4514351
11 O O3	0.1881822	-2.5437501	-0.4228806
12 O O4	2.3460914	-2.7525496	-1.0286034
13 H H3	2.0421431	-3.5917168	-1.4115495
14 N N2	-2.1308509	1.9716244	-0.0162508
15 H H4	-2.5352624	1.3855895	-0.7484905
16 H H5	-1.6571836	2.7497444	-0.4777716
17 H H8	-2.9001497	2.3599027	0.5307896

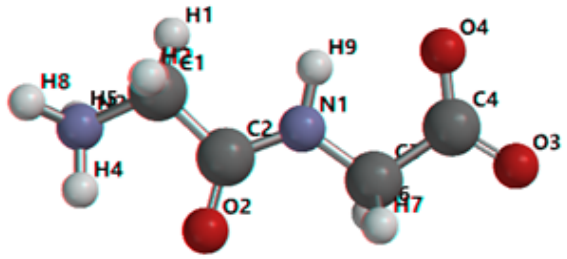

Figure 30: Zwitterionic Glycine Dimer. The labels represent specific atoms and correspond to the atoms located in Table 28

Table 28: Cartesian Coordinates (in angstroms) for Zwitterionic Glycine Dimer.

Atom	X	Y	Z
1 H H1	0.6442690	-1.0224608	-1.7620812
2 C C1	1.2414622	-0.3412669	-1.1650872
3 H H2	2.0386601	-0.8930988	-0.6718775
4 C C2	0.4287154	0.4081215	-0.1096476
5 O O2	0.7320437	1.5729416	0.1684093
6 N N1	-0.5386461	-0.2876598	0.4738648
7 C C3	-1.3568657	0.2034218	1.5685291
8 H H6	-1.9250205	1.0817179	1.2584640
9 H H7	-0.7305853	0.5023270	2.4101843
10 C C4	-2.3426211	-0.8881760	2.0391629
11 O O3	-3.0919329	-0.5600049	2.9877552
12 O O4	-2.3076310	-1.9906728	1.4387206
13 N N2	1.8773798	0.6690482	-2.0548924
14 H H4	2.0734418	1.5072104	-1.4951006
15 H H5	1.2612430	0.9487303	-2.8180203
16 H H8	2.7462859	0.3288772	-2.4641235
17 H H9	-0.7501982	-1.2390560	0.1957400

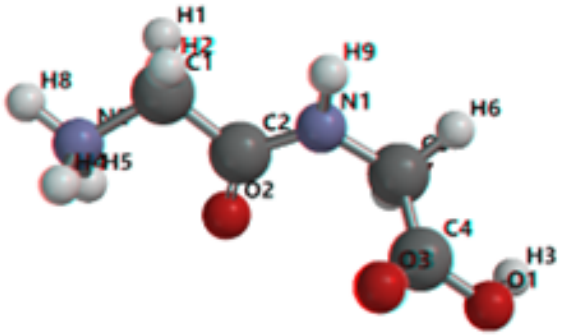

Figure 31: Cationic Glycine Dimer. The labels represent specific atoms and correspond to the atoms located in Table 29

Table 29: Cartesian Coordinates (in angstroms) for Cationic Glycine Dimer.

Atom	X	Y	Z
1 H H1	-1.0607842	0.2156819	-2.4250007
2 C C1	-1.2337334	0.6280751	-1.4332521
3 H H2	-0.9788950	1.6833780	-1.4351604
4 C C2	-0.4315383	-0.1712823	-0.4119296
5 O O2	-0.9724547	-1.0611302	0.2380356
6 N N1	0.8689484	0.1403934	-0.3313418
7 C C3	1.7744186	-0.5643304	0.5401097
8 H H6	2.7932730	-0.4274060	0.1788574
9 H H7	1.5587918	-1.6329420	0.5217181
10 C C4	1.7161112	-0.0822119	1.9859516
11 O O3	1.0501044	0.8421073	2.3754794
12 N N2	-2.6793803	0.4737361	-1.1168060
13 H H4	-2.9867848	1.1344295	-0.4028170
14 H H5	-2.8423585	-0.4686740	-0.7474991
15 H H8	-3.2640517	0.6111120	-1.9400712
16 H H9	1.2238439	0.9237703	-0.8575220
17 O O1	2.4905527	-0.7615134	2.8449246
18 H H3	2.9739369	-1.4831933	2.4163233

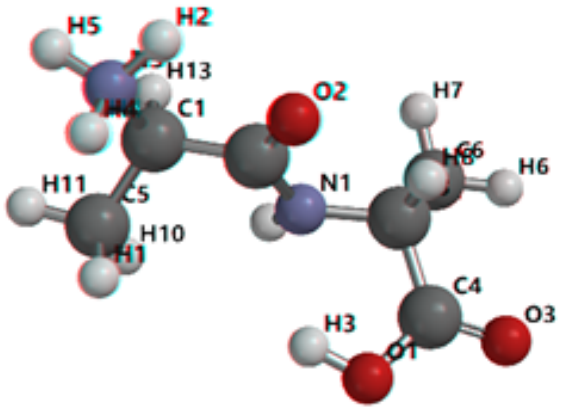

Figure 32: Cationic Alanine Dimer (RR). The labels represent specific atoms and correspond to the atoms located in Table 30

Table 30: Cartesian Coordinates (in angstroms) for Cationic Alanine Dimer (RR).

Atom	X	Y	Z
1 C C1	-1.4711942	0.3327463	-1.1422840
2 C C2	-0.3243505	-0.5162334	-0.5636116
3 O O2	-0.0179850	-1.5843283	-1.0901377
4 N N1	0.2863375	-0.0057489	0.5391059
5 C C3	1.2292533	-0.7954486	1.3350623
6 C C4	2.6290123	-0.1580949	1.3567306
7 O O3	3.5003713	-0.5308575	2.1140099
8 H H9	-0.1579397	0.7756730	1.0117000
9 O O1	2.8686611	0.8179786	0.4683342
10 H H3	2.0543688	0.9934123	-0.0455365
11 C C5	-1.1537764	1.8092777	-1.3733516
12 H H1	-0.2768049	1.9232922	-2.0190129
13 H H10	-0.9547302	2.3092447	-0.4222594
14 H H11	-2.0096232	2.3053856	-1.8406888
15 H H8	1.3438137	-1.7501203	0.8113746
16 C C6	0.7080039	-1.0422610	2.7538358
17 H H6	1.4409862	-1.6223181	3.3191823
18 H H7	-0.2322117	-1.5984532	2.7121488
19 H H12	0.5360936	-0.0941550	3.2755777
20 H H13	-2.3382351	0.2238391	-0.4831196
21 N N3	-1.8536621	-0.3229044	-2.4408281
22 H H2	-1.7493649	-1.3435580	-2.3585842
23 H H4	-1.2396225	-0.0260760	-3.2082822
24 H H5	-2.8174012	-0.1002919	-2.7093656

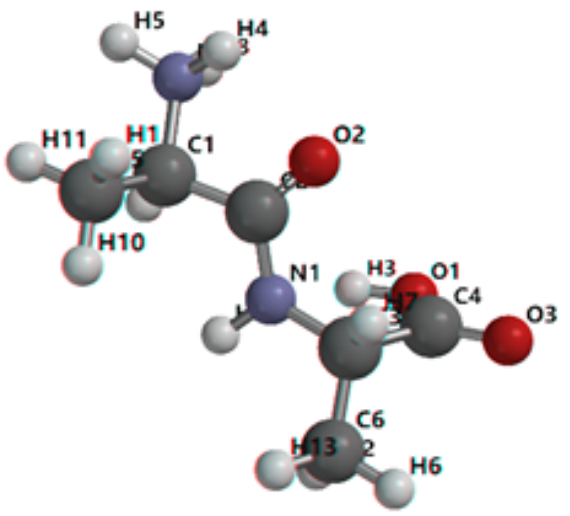

Figure 33: Cationic Alanine Dimer (SR). The labels represent specific atoms and correspond to the atoms located in Table 31

Table 31: Cartesian Coordinates (in angstroms) for Cationic Alanine Dimer (SR).

Atom	X	Y	Z
1 C C1	-1.3086544	0.6072255	-1.2110646
2 H H2	-2.1537800	0.5618274	-0.5215714
3 C C2	-0.1009957	-0.1713610	-0.6690899
4 O O2	0.6090083	-0.8291159	-1.4305632
5 N N1	0.1335935	-0.0528190	0.6583789
6 C C3	1.3836040	-0.5298334	1.2484958
7 H H7	2.1954749	-0.3358965	0.5419290
8 C C4	1.3845804	-2.0514157	1.4688697
9 O O3	2.4147623	-2.6751747	1.6118802
10 N N2	-1.7212387	-0.1096801	-2.4675274
11 H H4	-0.8723358	-0.4246935	-2.9608042
12 H H5	-2.2707944	0.4890471	-3.0930197
13 H H8	-2.2819663	-0.9454297	-2.2674856
14 H H9	-0.4643180	0.5490234	1.2152008
15 O O1	0.1936859	-2.6668213	1.5453558
16 H H3	-0.5241308	-2.0183775	1.4020366
17 C C5	-0.9371088	2.0570922	-1.5315124
18 H H1	-0.1104982	2.0974368	-2.2485005
19 H H10	-0.6292261	2.5644669	-0.6134708
20 H H11	-1.8014086	2.5852714	-1.9453362
21 C C6	1.6607109	0.1915403	2.5745200
22 H H6	2.6008194	-0.1699213	2.9982648
23 H H12	0.8566272	0.0093714	3.2959220
24 H H13	1.7435889	1.2682373	2.3990923

REFERENCES

[1] Carlton Levon Guidry. "Thermal synthesis of polypeptides in aqueous systems". Thesis. 1962.

[2] Ousama H Abbas. "Optimum conditions for prebiotic evolution in extraterrestrial environments". Thesis. 2004.

[3] Thomas D Campbell. "Potassium and the Prebiotic Formation of Polypeptides". Thesis. 2018.

[4] Rihe Liu. *Synthesis of biopolymers in aqueous environments*. University of California, San Diego, 1996.

[5] Robert Tattersall. "The energetics of peptide bond formation at elevated temperatures". Thesis. 1976.

[6] Jay G Forsythe et al. "Ester-mediated amide bond formation driven by wet–dry cycles: A possible path to polypeptides on the prebiotic Earth". In: *Angewandte Chemie International Edition* 54.34 (2015), pp. 9871–9875.

[7] Stanley L Miller and Harold C Urey. "Organic compound synthesis on the primitive earth". In: *Science* 130.3370 (1959), pp. 245–251.

[8] Eric. T Parker et al. "Conducting Miller-Urey Experiments". In: *The Journal of Visualized Experiments* 83 (2014), pp. 1–13.

[9] Norio Kitadai. "Thermodynamic prediction of glycine polymerization as a function of temperature and pH consistent with experimentally obtained results". In: *Journal of molecular evolution* 78.3-4 (2014), pp. 171–187.

[10] Henry Borsook. "Peptide bond formation". In: *Advances in protein chemistry*. Vol. 8. Elsevier, 1953, pp. 127–174. ISBN: 0065-3233.

[11] Stephen Cusumano. *Primordial and Modern Self Assembling Peptides: Synthesis and Characterization*. Southern Connecticut State University, 2017. ISBN: 0355188414.

[12] Kasumi Sakata, Norio Kitadai, and Tadashi Yokoyama. "Effects of pH and temperature on dimerization rate of glycine: evaluation of favorable environmental conditions for chemical evolution of life". In: *Geochimica et Cosmochimica Acta* 74.23 (2010), pp. 6841–6851.

[13] Jeremy Kua and Lauren M Sweet. “Preliminary oligomerization in a glycolic acid–glycine mixture: A free energy map”. In: *The Journal of Physical Chemistry A* 120.38 (2016), pp. 7577–7588.

[14] Sheng-Sheng Yu et al. “Kinetics of prebiotic depsipeptide formation from the ester–amide exchange reaction”. In: *Physical Chemistry Chemical Physics* 18.41 (2016), pp. 28441–28450.

[15] Rio Febrian. “PROSPECTIVE ROLES OF SALTS IN AQUEOUS PREBIOTIC ENVIRONMENTS”. Thesis. 2019.

[16] HJI Cleaves. “Studies in prebiotic synthesis and the origins of the metabolic pathways”. In: (2002).

[17] Irena Mamajanov et al. “Ester formation and hydrolysis during wet–dry cycles: generation of far-from-equilibrium polymers in a model prebiotic reaction”. In: *Macromolecules* 47.4 (2014), pp. 1334–1343.

[18] Jay G Forsythe et al. “Surveying the sequence diversity of model prebiotic peptides by mass spectrometry”. In: *Proceedings of the National Academy of Sciences* 114.37 (2017), E7652–E7659.

[19] Cornelus F van Nostrum et al. “Hydrolytic degradation of oligo (lactic acid): a kinetic and mechanistic study”. In: *Polymer* 45.20 (2004), pp. 6779–6787.

[20] Sara Imari Walker, Martha A Grover, and Nicholas V Hud. “Universal sequence replication, reversible polymerization and early functional biopolymers: a model for the initiation of prebiotic sequence evolution”. In: *PloS one* 7.4 (2012), e34166.

[21] Daniela Kroiss. “Minimalistic Peptide-Based Supramolecular Systems Relevant to the Chemical Origin of Life”. In: (2019).

[22] Sheng-Sheng Yu. “Ester-mediated amide bond formation: A possible path to protopeptides on the early earth”. Thesis. 2017.

[23] David R Witzke. “Introduction to properties, engineering, and prospects of polylactide polymers”. In: (1999).

[24] Dung T Vu et al. “Oligomer distribution in concentrated lactic acid solutions”. In: *Fluid phase equilibria* 236.1-2 (2005), pp. 125–135.

[25] SJ De Jong et al. “New insights into the hydrolytic degradation of poly (lactic acid): participation of the alcohol terminus”. In: *Polymer* 42.7 (2001), pp. 2795–2802.

[26] F Codari et al. “Kinetics of the hydrolytic degradation of poly (lactic acid)”. In: *Polymer degradation and stability* 97.11 (2012), pp. 2460–2466.

[27] S Lazzari et al. “Modeling the pH-dependent PLA oligomer degradation kinetics”. In: *Polymer degradation and stability* 110 (2014), pp. 80–90.

[28] Yogesh M Harshe et al. “Polycondensation kinetics of lactic acid”. In: *Macromolecular Reaction Engineering* 1.6 (2007), pp. 611–621.

[29] Suming Li. “Hydrolytic degradation characteristics of aliphatic polyesters derived from lactic and glycolic acids”. In: *Journal of Biomedical Materials Research* 48.3 (1999), pp. 342–353.

[30] Rajeev Mehta et al. “Synthesis of poly (lactic acid): a review”. In: *Journal of Macromolecular Science, Part C: Polymer Reviews* 45.4 (2005), pp. 325–349.

[31] Tatsuro Ouchi et al. “Synthesis and enzymatic hydrolysis of lactic acid–depsipeptide copolymers with functionalized pendant groups”. In: *Journal of Polymer Science Part A: Polymer Chemistry* 35.2 (1997), pp. 377–383.

[32] Hiroyuki Shirahama, Kouichi Umemoto, and Hajime Yasuda. “Synthesis and enzymatic degradation of optically active depsipeptide copolymers”. In: *Journal of Biomaterials Science, Polymer Edition* 10.6 (1999), pp. 621–639.

[33] Donald Garlotta. “A Literature Review of Poly(Lactic Acid)”. In: *Journal of Polymers and the Environment* 9.2 (2001), pp. 63–84.

[34] Navinchandra S. Asthana et al. “A Kinetic Model for the Esterification of Lactic Acid and Its Oligomers”. In: *Industrial Engineering Chemistry Research* 45.15 (2006), pp. 5251–5257.

[35] J Podlech. “Origin of organic molecules and biomolecular homochirality”. In: *Cellular and Molecular Life Sciences CMLS* 58.1 (2001), pp. 44–60.

[36] Jeremy Bailey. “Chirality and the origin of life”. In: *Acta Astronautica* 46.10-12 (2000), pp. 627–631.

[37] Leslie E. Orgel. “The Origin of Life on the Earth”. In: *Scientific American* 271.4 (1994), pp. 76–83.

[38] Kehinde Bankole and Gary Aurand. “Kinetic and Thermodynamic Parameters for Uncatalyzed Esterification of Carboxylic Acid”. In: *Research Journal of Applied Sciences, Engineering and Technology* 7 (2014), pp. 4671–4684.

[39] RA LUDICKY. “PROTEIN ANCESTORS: PEPTIDES FROM HYDROGEN-CYANIDE AND WATER (PREBIOTIC, NMR)”. In: (1987).

[40] D Scholefield. “Thermodynamic studies of peptide bond formation on clay mineral surfaces”. Thesis. 1974.

[41] Anna Jelińska, Leszek Dobrowolski, and Irena Oszczapowicz. “The influence of pH, temperature and buffers on the degradation kinetics of cefetamet pivoxil hydrochloride in aqueous solutions”. In: *Journal of pharmaceutical and biomedical analysis* 35.5 (2004), pp. 1273–1277.

[42] OAGJ Van der Houwen et al. “Systematic interpretation of pH-degradation profiles. A critical review”. In: *International journal of pharmaceutics* 155.2 (1997), pp. 137–152.

[43] Addison Ault. “General Acid and General Base Catalysis”. In: *Journal of Chemical Education* 84.1 (2007), pp. 38–39.

[44] Ashok Khanna et al. “Molecular modeling studies of poly lactic acid initiation mechanisms”. In: *Journal of Molecular Modeling* 14.5 (2008), pp. 367–374.

[45] Hans Ulrich Moritz. “Increase in viscosity and its influence on polymerization processes”. In: *Chemical Engineering Technology* 12.1 (1989), pp. 71–87.

[46] C. Martin et al. “Water-Based Dynamic Depsipeptide Chemistry: Building Block Recycling and Oligomer Distribution Control Using Hydration–Dehydration Cycles”. In: *JACS Au* 2.6 (2022), pp. 1395–1404.

[47] K.P. Burnham and D.R. Anderson. *Model selection and multimodel inference: A practical information-theoretic approach.* 2002.

[48] Michele Melchionna, Katie E Styan, and Silvia Marchesan. “The unexpected advantages of using D-amino acids for peptide self-assembly into nanostructured hydrogels for medicine”. In: *Current topics in medicinal chemistry* 16.18 (2016), pp. 2009–2018.

[49] Carl Henrik Holten. *Lactic acid. Properties and chemistry of lactic acid and derivatives*. Weinheim/Bergstr., W. Germany, Verlag Chemie GmbH, 1971. ISBN: 3527253440.

[50] Maria Ryner et al. “Mechanism of Ring-Opening Polymerization of 1,5-Dioxepan-2-one and l-Lactide with Stannous 2-Ethylhexanoate. A Theoretical Study”. In: *Macromolecules* 34.12 (2001), pp. 3877–3881.

[51] Yoshiharu Kimura. “Molecular, structural, and material design of bio-based polymers”. In: *Polymer journal* 41.10 (2009), pp. 797–807.

[52] Holger Kruse, Lars Goerigk, and Stefan Grimme. “Why the Standard B3LYP/6-31G* Model Chemistry Should Not Be Used in DFT Calculations of Molecular Thermochemistry: Understanding and Correcting the Problem”. In: *The Journal of Organic Chemistry* 77.23 (2012), pp. 10824–10834.

[53] M. Ya. Amusia, A.Z. Msezane, and V.R. Shaginyan. “Density Functional Theory versus the Hartree-Fock Method: Comparative Assessment”. In: *Physica Scripta* T68 (2003), pp. C133–C140.

[54] P Hohenberg and WJPR Kohn. “Density functional theory (DFT)”. In: *Phys. Rev* 136 (1964), B864.

[55] Christopher J Cramer and Donald G Truhlar. “Structure and Reactivity in Aqueous Solution: An Overview”. In: ACS Publications, 1994. ISBN: 1947-5918.

[56] Warren J Hehre. *A guide to molecular mechanics and quantum chemical calculations*. Vol. 2. Wavefunction Irvine, CA, 2003.

[57] Kelvin H Smith. “Thermodynamic and Kinetic Investigation of Depsipeptide Oligomerization and Degradation under Prebiotic Earth Conditions”. Thesis. 2022.

[58] Chemical Book. *(S)-(-)-2-ACETOXYPROPIONIC ACID*. 2022.

[59] Andrzej Duda and Stanislaw Penczek. “Thermodynamics of L-lactide polymerization. Equilibrium monomer concentration”. In: *Macromolecules* 23.6 (1990), pp. 1636–1639.

[60] T. G. Kulagina et al. “Thermodynamics of dl-lactide, polylactide and polymerization of dl-lactide in the range of 0–430K”. In: *Polymer Science U.S.S.R.* 24.7 (1982), pp. 1702–1708.

[61] David S. Ross and David Deamer. “Dry/Wet Cycling and the Thermodynamics and Kinetics of Prebiotic Polymer Synthesis”. In: *Life* 6 (3 2016), pp. 1–12.

[62] Everett L Shock. “Stability of peptides in high-temperature aqueous solutions”. In: *Geochimica et Cosmochimica Acta* 56.9 (1992), pp. 3481–3491.

[63] Everett L Shock and Harold C Helgeson. “Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: Correlation algorithms for ionic species and equation of state predictions to 5 kb and 1000 C”. In: *Geochimica et Cosmochimica Acta* 52.8 (1988), pp. 2009–2036.

[64] Everett L Shock and Harold C Helgeson. “Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: Standard

partial molal properties of organic species". In: *Geochimica et Cosmochimica Acta* 54.4 (1990), pp. 915–945.

[65] Everett L Shock, Harold C Helgeson, and Dimitri A Sverjensky. "Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: Standard partial molal properties of inorganic neutral species". In: *Geochimica et Cosmochimica Acta* 53.9 (1989), pp. 2157–2183.

[66] Jan P Amend and Harold C Helgeson. "Calculation of the standard molal thermodynamic properties of aqueous biomolecules at elevated temperatures and pressures Part 1L--Amino acids". In: *Journal of the Chemical Society, Faraday Transactions* 93.10 (1997), pp. 1927–1941.

[67] Jan P Amend and Harold C Helgeson. "Group additivity equations of state for calculating the standard molal thermodynamic properties of aqueous organic species at elevated temperatures and pressures". In: *Geochimica et Cosmochimica Acta* 61.1 (1997), pp. 11–46.

[68] Jan P Amend and Harold C Helgeson. "Calculation of the standard molal thermodynamic properties of aqueous biomolecules at elevated temperatures and pressures II. Unfolded proteins". In: *Biophysical chemistry* 84.2 (2000), pp. 105–136.

[69] Jeffrey M Dick, Douglas E LaRowe, and Harold C Helgeson. "Temperature, pressure, and electrochemical constraints on protein speciation: Group additivity calculation of the standard molal thermodynamic properties of ionized unfolded proteins". In: (2006).

[70] Harold C Helgeson and David H Kirkham. "Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures; I, Summary of the thermodynamic/electrostatic properties of the solvent". In: *American Journal of Science* 274.10 (1974), pp. 1089–1198.

[71] Vitalii A. Pokrovskii and Harold C. Helgeson. "Thermodynamic properties of aqueous species and the solubilities of minerals at high pressures and temperatures: the system Al2O3H2OKOH". In: *Chemical Geology* 137.3 (1997), pp. 221–242.

[72] George D Miron, Allan MM Leal, and Alina Yapparova. "Thermodynamic Properties of Aqueous Species Calculated Using the HKF Model: How Do Different Thermodynamic and Electrostatic Models for Solvent Water Affect Calculated Aqueous Properties?" In: *Geofluids* 2019 (2019).

[73] AB Meggy. "293. Glycine peptides. Part II. The heat and entropy of formation of the peptide bond in polyglycine". In: *Journal of the Chemical Society (Resumed)* (1956), pp. 1444–1454.

[74] S. Sastre, R. Casasnovas, and J. F. Munoz Frauac. In: *PCCP* 18 (2016), pp. 11202–11212.

[75] Bun Chan, Eric Collins, and Krishnan Raghavachari. In: *WIREs Comput. Mol Sci* 11 (2021), pp. 1–18.

[76] Oscar Ventura and Marc Segovia. In: *Chemical Physics Letters* 403 (2005), pp. 378–384.

[77] Snehasis Banerjee, Sunil K Bhanja, and Pijush Kanti Chattopadhyay. "Quantum chemical predictions of aqueous pKa values for OH groups of some -hydroxycarboxylic acids based on ab initio and DFT calculations". In: *Computational and Theoretical Chemistry* 1125 (2018), pp. 29–38.

[78] Robert N. Goldberg, Nand Kishore, and Rebecca M. Lennen. "Thermodynamic Quantities for the Ionization Reactions of Buffers". In: *Journal of Physical and Chemical Reference Data* 31.2 (2002), pp. 231–370.