

# **THE PH DEPENDENT MECHANISMS OF PEPTIDE BOND CLEAVAGE**

A Dissertation  
Presented to  
The Academic Faculty

by

Yi Sun

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Science in the  
School of Chemical & Biomolecular Engineering

Georgia Institute of Technology  
December 2019

**COPYRIGHT © 2019 BY YI SUN**

# **THE PH DEPENDENT MECHANISMS OF PEPTIDE BOND CLEAVAGE**

Approved by:

Dr. Martha Grover, Advisor  
School of Chemical & Biomolecular Engineering  
*Georgia Institute of Technology*

Dr. Charles Liotta  
School of Chemical & Biomolecular Engineering  
*Georgia Institute of Technology*

Dr. Anant Paravastu  
School of Chemical & Biomolecular Engineering  
*Georgia Institute of Technology*

Date Approved: [Nov 20, 2019]

## ACKNOWLEDGEMENTS

My special thanks goes to my thesis advisor, Dr. Martha Grover and Dr. Charles Liotta, who offered the great opportunity and provided tremendous support, mentorship, and encouragement throughout my entire graduate work at Georgia Tech. I feel very fortunate to be a student in their lab and grateful for their efforts on educating me to become a professional researcher. I have broadened my horizon and gained a wide range of scientific knowledge through this challenging but interesting project. The completion of this work would not have been possible without their support and I truly indebted to them throughout my lifetime.

I would also like to acknowledge the other advisors in the Center for Chemical Evolution (CCE): Dr. Nicholas Hud, Dr. Jay Forsythe at College of Charleston and Dr. Paul Bracher at Saint Louis University for their insightful discussions, valuable feedbacks and timely suggestions throughout my research project work. I appreciate the funding from this project from the Center for Chemical Evolution and thank all the members from the CCE proto-polypeptide team. Special thanks for Dr. Jay Forsythe on conducting related MS/MS analysis using Xevo G2 mass spectrometer and analyzing the depsipeptides pool sequence information.

Additionally, I would like to say a special thanks to my mentors, Dr. Shengsheng Yu, Dr. Zhao Li and Dr. Moran Pinter, for their tremendous assistance with technical trainings on the lab works, carrying out NMR, HPLC and LCMS studies, and for their helpful discussions on the related research project.

I am grateful for the supports and help from all current and former members in Dr. Grover's research group: Dr. Daniel Griffin, Dr. Xun Tang, Dr. Christine He, Dr. Ming-Chien Hsieh, Dr. Nils Persson, Dr. Michael McBride, Tristan Kernick, Matthew McDonald, Chiamaka Obianyo, Kelvin Smith, Youngjo Kim, Stefani Kocevaska, Patrick Harris, Colton Lagerman, and Aaron Liu. The friendly interaction and discussion with them are vital to my progress in this project and I will miss their accompany.

Above all, I would like to appreciate the people who mean a lot to me, my parents, for always supporting my dreams, respect all my decisions and giving me freedom to choose what I desired. I salute you all for the selfless love, care and sacrifice you did to shape my life, which I would never be able to pay back.

# TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS</b>	<b>iii</b>
<b>LIST OF TABLES</b>	<b>vi</b>
<b>SUMMARY</b>	<b>viii</b>
<b>CHAPTER 1. Introduction</b>	<b>1</b>
<b>1.1 Brief introduction on prebiotic chemistry</b>	<b>1</b>
<b>1.2 Peptide polymerization</b>	<b>3</b>
<b>1.3 Peptides degradation</b>	<b>6</b>
1.3.1 Preliminary experiments	6
<b>CHAPTER 2. The pH dependent mechanisms of Non-enzymatic peptide bond cleavage reaction</b>	<b>10</b>
<b>2.1 Background</b>	<b>10</b>
<b>2.2 Objective</b>	<b>13</b>
<b>2.3 Methods</b>	<b>15</b>
<b>2.4 Results</b>	<b>17</b>
2.4.1 Glycine dimer reaction kinetics	17
2.4.2 Glycine trimer (GGG) reaction kinetics	22
2.4.3 Rate constants based on proposed kinetics model	26
2.4.4 Alanine trimer (AAA) reaction kinetics	27
<b>2.5 Discussion</b>	<b>29</b>
<b>2.6 Conclusions</b>	<b>31</b>
<b>APPENDIX A. Supplements</b>	<b>33</b>
<b>A.1 Depsipeptides recycling preliminary experiment</b>	<b>33</b>
A.1.1 Method	33
<b>A.2 Peptide bond cleavage project</b>	<b>33</b>
A.2.1 Mathematical model	33
A.2.2 Confidence interval calculation	34
<b>A.3 Supporting materials</b>	<b>36</b>
<b>REFERENCES</b>	<b>71</b>

## LIST OF TABLES

Figure 1. 2. 1	- Strecker reaction pathways for prebiotic amino acids synthesis. <sup>19</sup>	3
Figure 1. 2. 2	- Ester amide exchange reaction mechanism. <sup>4</sup>	5
Figure 1. 3. 1	- Mass spectra for glycine dimer mixed with lactic acid, subjected to six wet (95°C)-dry (85°C) cycles, over a period of 6 days. MS/MS spectra for m/z at 260 and 374, performed on a Waters Xevo G2 mass spectrometer.	8
Figure 1. 3. 2	- Recycling of the glycine dimer and lactic acid experiment under different wet temperature conditions, characterized by MS.	9
Figure 2. 1. 1	- Transition state for the acid-catalyzed backbiting pathway as described in Ref. 39.	11
Figure 2. 2. 1	- Reaction pathways for the amide cleavage of the glycine trimer (GGG), the glycine dimer (GG), and the cyclic glycine dimer (cGG) and the ring closing reaction of the linear dimer (GG) along with the associated rate constants. $k_{sc}$ is the rate constant corresponding to the hydrolysis of the glycine oligomers through a scission pathway; $k_{bb}$ is the rate constant for the backbiting pathway of GGG, and $k_{rc}$ and $k_{ro}$ are the rate constants for cGG ring closing and opening, respectively.	14
Figure 2. 4. 1	- Stacked 1H NMR spectra for GG reactions at 95°C at pH 7. Each spectrum corresponds to a different sampling time, ordered in time from bottom to top.	18
Figure 2. 4. 2	- Glycine dimer (GG) (a-d) and cyclic glycine dimer (cGG) (e-h) reaction kinetics at pH 3, 5, 7 and 10 at 95°C over a period of five days. The squares represent experimental data obtained from 1H NMR measurements, while the solid lines represent the model prediction. (GG)=blue; (G)=black; (cGG)=red.	19
Figure 2. 4. 3	- Mechanistic pathway for GG hydrolysis at pH 3.	19
Figure 2. 4. 4	- Mechanistic pathway for GG hydrolysis at pH 10.	20
Figure 2. 4. 5	- Mechanistic pathway for the acid-catalyzed ring-opening of cGG.	21

Figure 2. 4. 6	- Acid-catalyzed backbiting mechanism for the formation of cGG and glycine.	23
Figure 2. 4. 7	- Glycine trimer (GGG) reaction kinetics at pH 3, 5, 7 and 10 at 95°C for a period of five days (a-d). The symbols denote <sup>1</sup> H NMR measured abundance and solid lines are the model predictions. (GGG)=green; (GG)=blue; (G)=black; (cGG)=red	24
Figure 2. 4. 8	- The estimated four rate constants for GG, cGG and GGG kinetics at pH 3, 5, 7 and 10 at 95°C. The confidence intervals are calculated following the chi-squared method (discussed in Supplemental Information) at the 95% confidence level.	26
Figure 2. 4. 9	- Alanine trimer (AAA) degradation reaction kinetics under pH 3, 5, 7 and 10, at 95°C, over 5 days (a-d). The symbols denote <sup>1</sup> H NMR measured quantifications and solid lines are the model predictions. (AAA)=green; (A)=black; (AA)=blue; (cAA)=red	27

## SUMMARY

The origin of life under prebiotic conditions has been an unsolved mystery for decades. Amino acids were available under prebiotic conditions, and different approaches of amino acids condensation into proto-polypeptides have been well designed, giving rise to a prebiotic soup with various peptide sequences.

The emergence of functional biopolymers involves not only polymerization into longer species, but also the selective process with some species being protected and enriched over time. In this project, we treated peptide bond cleavage as the driving force for the selection process, by reshuffling peptide sequences and thus increasing the rate of search through sequence space. As a result, understanding the reaction mechanisms and quantifying the degradation kinetics of various peptide species is necessary to design a prebiotically plausible system that can demonstrate chemical evolution.

In this project, we conducted fundamental research studies to understand the impact of pH on the peptide degradation reaction kinetics and mechanisms. The degradation rate of the amide bonds in oligopeptides in aqueous solution is pH-dependent and is suggested to involve two distinct mechanism: direct hydrolysis (herein termed “scission”) and backbiting. While amide degradation was studied previously using various peptides, no systematic study has been reported addressing the separate rates of amide bond degradation over a wide pH range via these two mechanisms. In this study, the degradation kinetics of several short oligopeptides, specifically the glycine dimer, trimer, and cyclic dimer, as well as the alanine trimer, were measured at 95°C over a range of pH conditions using  $^1\text{H}$  NMR. The rate constants were obtained by solving the differential equations based on mechanistic

models and elucidate the favored reaction pathway under acidic, neutral, and basic pH conditions.

The degradation rate of the glycine trimer is much faster than the dimer under the acidic and neutral pH conditions. The glycine dimer degradation rate is highest under acidic and basic conditions, while the glycine trimer degradation rate is highest under neutral pH conditions. The results suggest that while the glycine dimer undergoes ring opening purely through a scission reaction mechanism, the glycine trimer is degraded through both backbiting and scission reaction mechanisms. At an acidic pH of 3, both mechanisms are active, while at neutral pH backbiting is dominant. In contrast, at a basic pH of 10, scission dominates.

# CHAPTER 1. INTRODUCTION

## 1.1 Brief introduction on prebiotic chemistry

The emergence of life is a fascinating, unsolved mystery in the field of prebiotic chemistry. Over time, scientists have proposed many different theories and models to explain how life appears, but we are still far from deciphering the emergence of life from an inert chemical system into a proto-living state and, eventually, into a living organism. Many ideas have emerged and govern our thinking about this question. The successful identification of amino acids in the Miller-Urey experiment,<sup>1-4</sup> and the hypothesis of the RNA World<sup>5</sup> in which an RNA molecule catalyzes its own duplication<sup>6</sup> have led scientists to confront the origin of life problem from a chemistry perspective.

One large paradoxical question is the controversy between protein first and nucleic acid first scenarios. In current living systems, proteins are required for DNA replication while nucleic acids are needed for the biosynthesis of proteins. The self-catalysis and information storage roles of RNA have been proposed to validate the RNA world hypothesis.<sup>5,7</sup> However, the instability of RNA and the challenges to synthesize RNA with high yield under plausible prebiotic conditions prior to the emergence of enzymes suggest that the RNA world hypothesis is insufficient to explain the emergence of life. Instead, life might have originated from the interplay between small molecules<sup>8</sup>, peptides<sup>9</sup> and lipids.<sup>10</sup> To this day, the direct condensations of amino acids into longer proto-polypeptides are still important research topics in the field of prebiotic chemistry searching for the emergence of life. The primordial soup with mixtures of different amino acids and peptides is plausible to be generated under prebiotic conditions.<sup>9, 11-12</sup> These amino acids could then

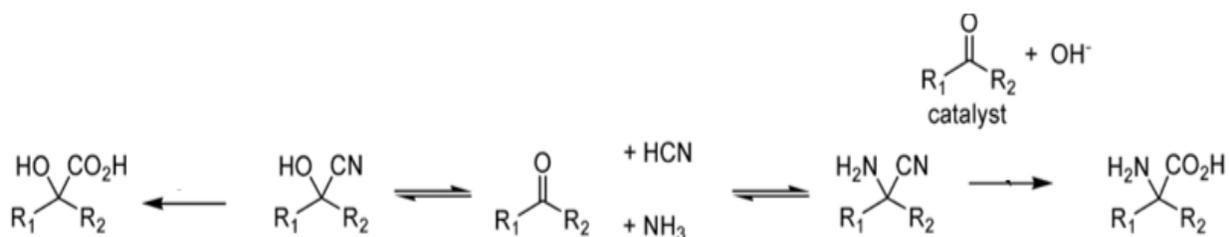
condense into peptide sequences, undergo further selection and evolution processes, and ultimately create molecular complexity.<sup>13</sup>

There are two factors worth approaching to demonstrate the emergence of an evolving prebiotic peptide system, including the condensation reactions of amino acids into longer proto-polypeptides and the degradation reactions of polymers back to shorter peptides or amino acids to regenerate limited starting materials under prebiotic conditions. In recent studies, peptide condensation studies compatible with different early Earth environmental conditions confirmed various reactions to synthesize proto-peptides abiotically, either through heating, wet-dry cycling<sup>14</sup> or with the aid of catalysts.<sup>15</sup> In recent studies, the length of the synthesized proto-polypeptides has been extended from dimers and trimers to much longer polymers, forming stable structures or assemblies and aggregates.<sup>13</sup> On the other hand, the recycling process of longer sequences back to short peptides or amino acids to regenerate the finite monomer resources and select for functional biopolymers is also vital for the evolving system as well.<sup>16</sup> In this project, a hydrated phase (hydrolysis) is added to the dehydrated condensation process, during which the polymers with reversible backbones are subjected to degradation reactions. We assumed that the peptide hydrolysis reactions are the driving forces for a reversible peptide polymerization system. The fundamental reaction mechanism study related to peptide degradations under different pH conditions are discussed in detail in the next few chapters.

The aim of this project is to systematically study the peptide hydrolysis reactions under different pH conditions. The amide bond degradation rate is pH dependent, related to two distinct reaction pathways, direct hydrolysis and backbiting. In this study, we want to learn the impact of pH on the degradation kinetics and the selection of reaction pathways in short oligopeptides.

## 1.2 Peptide polymerization

In the 1950s, Miller and Urey<sup>1</sup> have demonstrated the successful abiotic synthesis of amino acids by applying an electric charge to a chemical solution of methane, ammonia, hydrogen and water mixture, following the Strecker reaction pathway as shown in Figure 1.2.1. In a recent study<sup>2, 4</sup> that reanalyzed the archived samples of Miller's original experiment utilizing state-of-the-art analytical tools, over 40 amino acids and amines were identified, including abundant amount of glycine, alanine, valine, and serine.



**Figure 1. 2. 1 - Strecker reaction pathways for prebiotic amino acids synthesis.**<sup>19</sup>

In present living organisms, peptides are generated by association of the ribosome with mRNA. In organic chemistry, peptides are synthesized by solid-phase peptide synthesis (SSPS) on a rink amide resin using fmoc- $\alpha$ -amine-protected amino acid.<sup>17</sup> In both methods, those complex processes of synthesizing long peptides accurately require enzymes, which are unlikely to be present under prebiotic conditions. A challenge in prebiotic chemistry is to seek other simple methods that can form peptides in a reasonable prebiotic condition.

Various synthetic approaches have been postulated to promote peptide polymerization, either activated by a catalyst or with the addition of a condensing agent. One proposed idea brings about the formation of peptides under mild conditions by exposing  $\alpha$ -amino acids to one simple volcanic gas, carbonyl sulfide (COS),<sup>12</sup> through the formation of N-carboxyanhydride

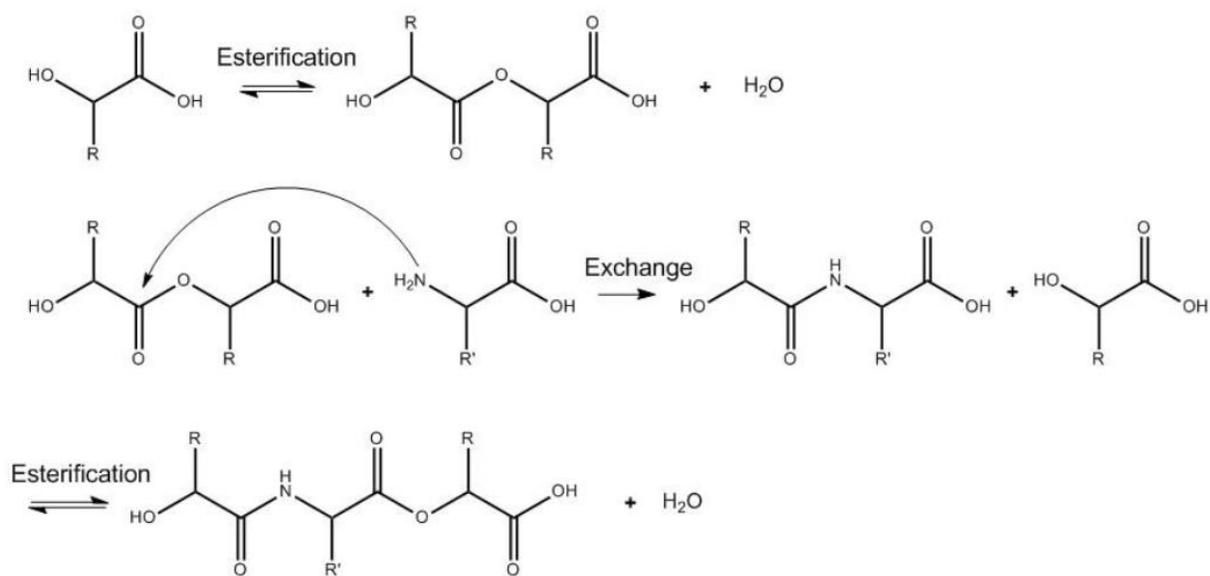
intermediates. Condensing agents are usually consumed as reactants and as a result, continuous feeding of a condensing agent, such as volcanic gas, is required to validate the proposed idea under prebiotic conditions. In addition, amino acids can be condensed into peptides by mineral catalysts, such as clays<sup>15</sup> (e.g., montmorillonite), silica and alumina,<sup>18</sup> and salts such as NaCl and Cu(II).<sup>19</sup> Catalysts may be more plausible explanations compared to condensing agents, due to the limited availability of continuous supply of feeding source under prebiotic conditions.

Another amino acid polymerization method is through direct polycondensation reaction, with the elimination of one water molecule under elevated temperature conditions. One of the challenges for direct poly-condensation<sup>20</sup> is that this reaction is thermodynamically unfavorable in aqueous solution. Another challenge is the accumulation of cyclic dipeptides (diketopiperazine, DKP) as a major side product, which hinders the reaction from further polymerization.

Previously, our group has proposed the polycondensation of polypeptides through ester amide exchange reactions.<sup>14</sup> The energy barrier for the polycondensation of hydroxyl acids to polyesters is much lower compared to the direct polycondensation of amino acids<sup>11</sup> and as a result, it is plausible to synthesize polyesters first through polycondensation reaction under mild temperature condition. Then the amino acids would replace hydroxyl acids on polyesters from the C-terminus and generate a polymer called depsipeptides, with alternating hydroxyl acid and amino acid backbones. The reaction mechanism is shown in Figure 1.2.2. The addition of hydroxyl acid lowers the overall activation energy barrier for amide bond formation.<sup>11</sup> Hydroxyl acids also form cyclic dimers as well, but unlike peptide cyclic dimers, ester linked cyclic dimers of hydroxyl acids are less stable towards hydrolysis. So, the ester bond linked ring structures could be reopened and recycled for further polymerization by gentle degradation. The design of this ester-amide exchange reaction system to synthesis proto-polypeptides overcomes the DKP energy sink challenge and

lowers the energy barrier with the addition of hydroxyl acids, and depsipeptides are treated as one plausible proto-polypeptides in the current prebiotic research area.

Recently, it is plausible to incorporate various amino acids, including nonpolar, polar and positive charged amino acids <sup>21</sup> into longer polymer with different sequences with hydroxyl acid as the catalyst, which increases the molecular complexity of the system. The elongated oligomerization could exhibit through both  $\alpha$  amines and side chain groups. These findings lead to another important topic in designing a selective system considering both degradation and recycling process. The hydrolysis step provides a chance for recycling, the reshuffling of limited monomer resources enabling a selection process.



**Figure 1. 2. 2 - Ester amide exchange reaction mechanism.<sup>4</sup>**

### 1.3 Peptides degradation

The emergence of a prebiotic peptide system involves not only polymerization of longer species, but also the selection process with some species being protected, while some other species degraded over time. In this case, the selection process is assumed to occur primarily through amide bond degradation from the prebiotic soup,<sup>22</sup> by recycling some of the peptide species, reshuffling peptide sequences, and thus increasing the rate of search for more stable peptides through sequence space.<sup>14, 23</sup>

#### 1.3.1 Preliminary experiments

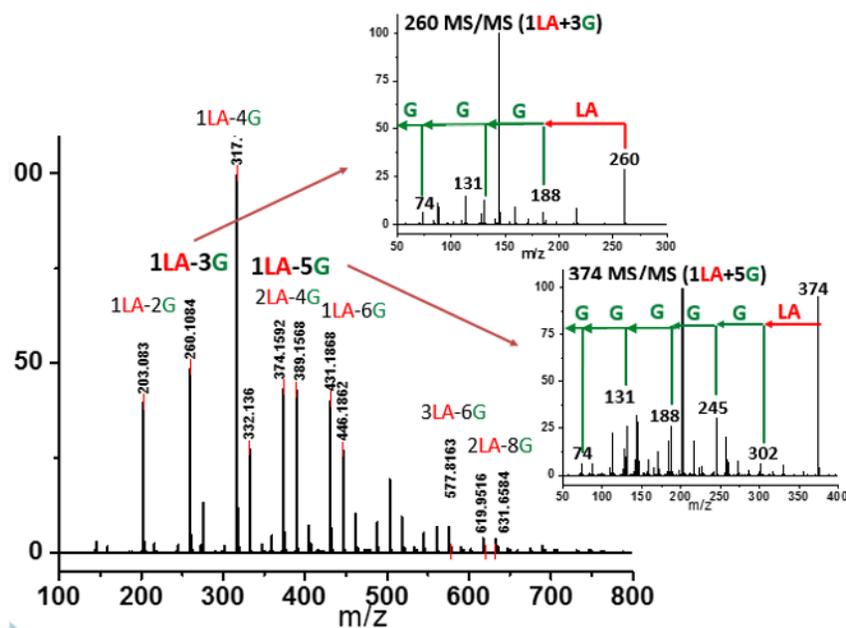
Based on the ester-amide exchange reaction system mentioned in Chapter 1.2, it is plausible to create a prebiotic soup with mixtures of different depsipeptides with various length by mixing hydroxyl acids and amino acids, subjected to wet-dry cyclings. The purpose of this project is to illustrate the recycling process and demonstrate that the amino acids and shorter peptides regenerated from the degradation reaction of longer peptides species during the wet phase, can be further polymerized in the subsequent cycles.

The traditional successive wet-dry cycles usually involve an alternation dry-hot phase mimicking the early Earth day time scenarios, followed by wet-cool periods mimicking the prebiotic night and rainy scenarios.<sup>14, 24</sup> Fluctuations on temperature and humidity conditions could be regular events on prebiotic Earth, which could be due to the self-rotation of the Earth (day-night cycling), the tidal impact of the Moon and the climate changes resulting (wet-dry cycling).<sup>25</sup> Cycling between dry and wet phases provides the environmental conditions required to selectively synthesize the depsipeptides through the condensation-dehydration process. During the hot-dry phase, the polymerization reactions are favored, while the wet-cool phase could promote the

cleavage and selection processes. In order to illustrate the recycling process, the wet phase temperature condition in this project is elevated to demonstrate the selective degradation of longer polymers into shorter ones, and the incorporation of those regenerated materials to form new polymers.

In this project, the glycine dimer and hydroxyl acid are mixed and subjected to six wet-dry cycles. The repeated condensation and dehydration reactions provide the chance of successive polymerization reactions with reversibility. If the original dimer is degraded during the wet phase and further polymerized into depsipeptides, it is plausible to observe odd number of glycine units in the final depsipeptide sequence library.

Figure 1.3.1 shows the mass spectra for the final species distribution of depsipeptides starting from glycine dimer and lactic acid and the MS/MS spectra of the selected depsipeptides, providing the peptide sequence information for the specific selected peptide mass. The appearance of 1LA-3G and 1LA-5G in the final depsipeptides library suggests that the original glycine dimers have been degraded and the regenerated glycine monomer could be further incorporated into depsipeptides with the addition of lactic acids in the subsequent cycles. The 1LA-3G and 1LA-5G MS/MS spectra confirm the synthesized depsipeptides backbones with the lactic acid on the N terminus and a continuous glycine sequence linked by amide bonds, based on the fragmentation information.

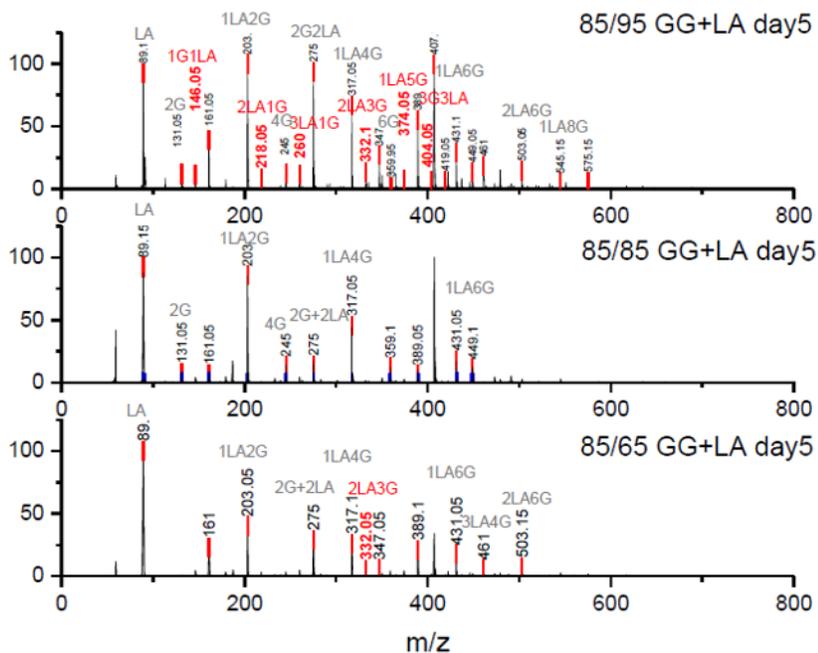


**Figure 1. 3. 1 - Mass spectra for glycine dimer mixed with lactic acid, subjected to six wet (95°C)-dry (85°C) cycles, over a period of 6 days. MS/MS spectra for m/z at 260 and 374, performed on a Waters Xevo G2 mass spectrometer.**

Additional control experiments were performed with lower wet temperatures at 65°C and 85°C, keeping dry phase at the same temperature condition at 85°C. Figure 1.3.2 has listed three mass spectra, with depsipeptide mass distribution under different wet phase temperature conditions at 65°C, 85°C and 95°C over five days.

If the temperature increases during the wet phase, the recycling rate of the glycine dimer would increase, regenerating higher amount of glycine monomer. As a result, the further recycling of glycine monomer into new depsipeptides with odd glycine units are more strongly observed. Depsipeptides with odd units (red labeled) of glycine in the depsipeptides backbone are detected mostly given the wet phase under 95°C and almost no such depsipeptides are observed under lower temperature conditions during the wet phase.

In conclusion, we have demonstrated that the amide bond could be degraded during the wet phase at 95°C and reused in the subsequent regeneration of new polymers through further polymerization reaction. In this case, a depsipeptide system including both regeneration and recycling processes could be generated, with some more stable peptides species being selected over time while some peptides with higher energy degraded. This reaction system is consistent with the limited amount of starting materials under the prebiotic conditions due to a regeneration of the monomer resources without a continuous supply of starting materials.



**Figure 1. 3. 2 - Recycling of the glycine dimer and lactic acid experiment under different wet temperature conditions, characterized by MS.**

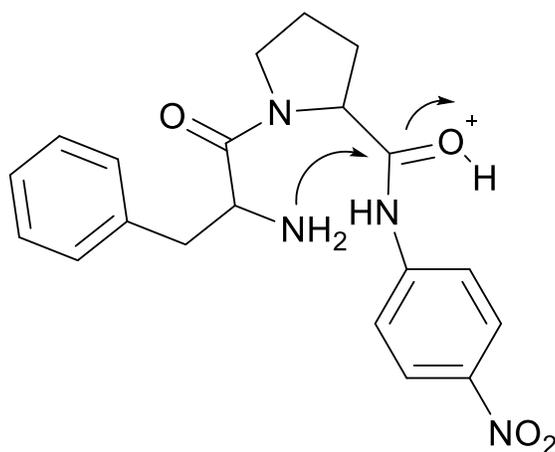
## CHAPTER 2. THE PH DEPENDENT MECHANISMS OF NON-ENZYMATIC PEPTIDE BOND CLEAVAGE REACTION

### 2.1 Background

Fundamental studies of the reaction kinetics and mechanisms associated with non-enzymatic peptide cleavage in aqueous solution are critical in a wide variety of scientific areas, including enzymatic catalysis,<sup>26-27</sup> peptide synthesis,<sup>17</sup> geochemistry<sup>28</sup> and prebiotic chemistry.<sup>9, 28</sup> On a practical note, these studies provide important comparisons for enzymatic peptide hydrolase reactivity studies,<sup>26-27, 29</sup> and information related to pharmaceutical storage procedures and conditions for peptide-based drugs.<sup>30-32</sup> From a prebiotic chemistry point of view, peptides have been shown to form under various conditions simulating environments on the early Earth, such as hydrothermal conditions mimicking deep-sea environments and shallow pools on land.<sup>11-14, 20, 33-35</sup> Investigation of the kinetics and mechanisms associated with the stepwise cleavage of the polypeptides back to the amino acid building blocks and smaller polypeptides are critical to understanding the survival and selection of functional polypeptides related to the origin of life on early Earth.<sup>23</sup>

The literature contains several reports addressing dipeptide hydrolysis within a range of pH and temperature conditions.<sup>36-37</sup> Wolfenden and co-workers<sup>27</sup> have reported the amide bond cleavage rates of diglycine under neutral pH conditions at temperatures ranging from 120°C to 200°C. The half-life of diglycine was determined to be approximately 350 years when the data were extrapolated to 25°C. Yokoyama and co-workers<sup>38</sup> showed that both pH and temperature affect the cleavage of diglycine and modeled the reaction kinetics taking into account the different ionization states of the dipeptide as the pH of the aqueous medium changed. Amide cleavage investigations of longer polypeptides have also been reported under various pH conditions. Bada and co-workers<sup>39-40</sup> investigated the decomposition of a tripeptide and a hexapeptide at elevated

temperatures (130°C) under neutral pH conditions and showed the formation of diketopiperzines. They proposed that the diketopiperzine was derived from an internal aminolysis (backbiting) mechanism where the N-terminal amine attacked a terminal carbonyl group via a 6-membered ring transition state cleaving the oligomer and forming the diketopiperazine. Goolcharran and Borchardt employed a simple model peptide, phenylalanine-proline-p-nitroaniline, to investigate the backbiting reaction pathway as a function of pH.<sup>40</sup> The overall amide cleavage rates increased with increasing pH. The backbiting pathway was found to dominate within the pH range 3-8 while the direct scission reaction pathway was dominant below pH 3 and above pH 8. Figure 2.1.1 shows the transition state for the acid-catalyzed backbiting process. It should be emphasized that a free amine group is necessary for this process to occur and that the amount of free amine decreases as the pH decreases. At the same time, the equilibrium protonation of the carbonyl oxygen, which increases the electrophilicity of the carbonyl carbon, increases with decreasing pH. In addition, to complete the formation of the diketopiperazine product, protonation of the amine leaving group is necessary. This last step in the overall process is facilitated as the pH decreases. Thus, a delicate pH balance must be achieved in order for substantial backbiting to take place.



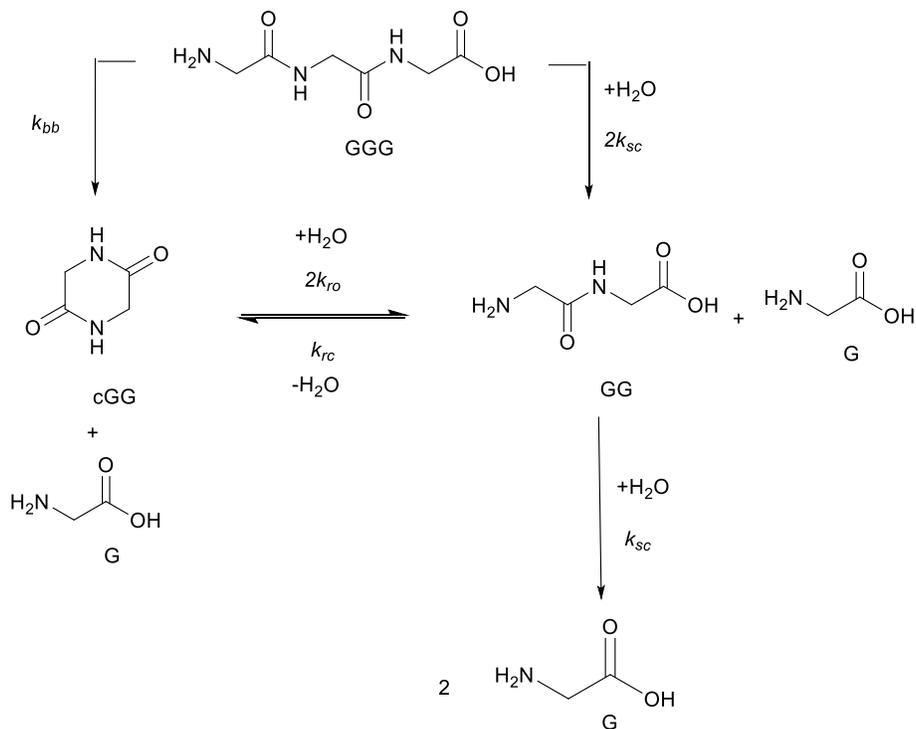
**Figure 2. 1. 1 - Transition state for the acid-catalyzed backbiting pathway as described in Ref. 39.**

Recently, Savage et al.<sup>41</sup> reported studies related to amide cleavage reactions of tetra-alanine as a function of temperature (170-230°C) and pH in high-pressure/high-temperature water to mimic hydrothermal vent conditions; they identified a kinetic model to describe their experimental results that contained both backbiting and scission pathways. It is well-known that the auto-ionization constant of water increases, and the dielectric constant of water decreases as the temperature of water increases. These two factors could potentially influence the relative reaction pathways (backbiting vs. scission) associated with amide bond cleavage when compared to similar reactions at substantially lower temperatures.<sup>42-43</sup> Moreover, acid catalysis was not included in their model, and could potentially play a crucial role affecting the interplay between the two mechanisms of amide bond cleavage. Indeed, as will be hereby shown in this paper, scission is significant at acidic pH.

## 2.2 Objective

The objectives of this paper are (a) to examine the effect of pH on diglycine (GG), cyclic glycine dimer (cGG) and triglycine (GGG) bond cleavage, and (b) to quantify the two competing reaction pathways (scission and backbiting) across acidic, neutral and basic pH conditions at 95°C. Herein we report the results of non-enzymatic peptide cleavage, addressing two mechanistic pathways. No previous studies on peptide cleavage have considered atmospheric pressure and temperatures below 100 °C. For comparison with GGG, we also investigated the effect of pH on alanine trimer (AAA) amide bond cleavage. Using quantitative <sup>1</sup>H NMR analysis, the amounts of each species were measured as a function of time and the rate constants for each reaction pathway were then estimated based on the proposed kinetic model (see Methods section). We characterized the cleavage kinetics of amide bonds both within the starting oligomer and also in all the accompanying products. Specifically, the hydrolysis reactions of GG, cGG, and GGG were conducted at 95°C at pH values ranging from 3 to 10. It is important to note that neither the decomposition of the glycine monomer nor polymerization of the glycine monomer nor any of the oligomers were observed under the reported experimental conditions. Only amide cleavage reactions and cyclization reactions to form diketopiperazine were observed. Figure 2.2.1 shows the postulated reaction pathways for GGG and GG amide bond cleavage, the opening of the cyclic dimer (cGG) and the closing of the linear dimer (GG) along with the accompanying rate constants. These are the rate processes which form the basis of the kinetic model reported herein. Two reaction pathways are possible for the hydrolysis of GGG, either forming G and GG through the scission pathway or forming G and cGG through the backbiting pathway. GG can react further to produce two Gs through amide scission or to produce cGG via a reversible cyclization. Reversibility for the ring opening of cGG or the ring closure of GG takes place only under acidic or neutral conditions. Under basic condition the ring opening process is irreversible. When fitting the four rate constants, multiple data sets including all GG, cGG and GGG hydrolysis data at a specific pH are fitted with a shared set of rate constants. It is assumed that the scission rate constant

is the same in both the trimer and the dimer. The model does not include the rate constants associated with every ionization state of G, GG, cGG, and GGG. Instead the more compact model is applied as outlined in Figure 2.2.1, having distinct rate constants at each of the four pH levels.



**Figure 2. 2. 1 - Reaction pathways for the amide cleavage of the glycine trimer (GGG), the glycine dimer (GG), and the cyclic glycine dimer (cGG) and the ring closing reaction of the linear dimer (GG) along with the associated rate constants.  $k_{sc}$  is the rate constant corresponding to the hydrolysis of the glycine oligomers through a scission pathway;  $k_{bb}$  is the rate constant for the backbiting pathway of GGG, and  $k_{rc}$  and  $k_{ro}$  are the rate constants for cGG ring closing and opening, respectively.**

Procedure and results in this chapter are reprinted with permission from Reference <sup>44</sup>.

## 2.3 Methods

Glycine monomer (Sigma G7126), glycine dimer (Sigma G1002), glycine trimer (Sigma G1377), cyclic glycine dimer (Sigma G7251), L-alanine trimer (Sigma A9627), hydrochloric acid, sodium hydroxide, potassium hydrogen phthalate (Sigma P1088), deuterium oxide (99.9 mol%) and HPLC-grade water were all obtained from Sigma-Aldrich.

Glycine dimer and glycine trimer were dissolved in water at a concentration of 200 mM. The initial concentration of cyclic glycine dimer and alanine trimer were 100 mM due to its low solubility. The initial pH of the solutions was measured using FiveEasy Benchtop F20 pH/mV Meter with an InLab Micro pH electrode probe from Mettler-Toledo.

In each experiment, peptide solutions were prepared and the initial pH was adjusted using HCl or NaOH to 3, 5, 7 or 10 at room temperature. Reactions were held in 2 mL glass vials (Supelco 29381-U) with a starting volume of 200  $\mu$ L. The vials were sealed and heated at 95°C for up to 120 hours in an oven. At various time points (0 h, 12 h, 24 h, 36 h, 48 h, 72 h, 96 h and 120 h), three replicates were removed from the oven. We noted that less than 10% of water evaporated during the heating process. All the results are expressed in moles instead of concentrations due to the volume changes during the heating process. Changes in pH of 0.2-1.0 pH units were observed during the heating process.

Before analysis, all the solutions were transferred into new tubes and water was removed using a Speedvac for 5 hours at ambient temperature to suppress the water signal during NMR analysis. The dried samples were then rehydrated with 600  $\mu$ L of D<sub>2</sub>O, and potassium hydrogen phthalate (25 mM final concentration) was added as an internal standard before <sup>1</sup>H NMR analysis was undertaken. The samples were then analyzed using a Bruker Avance IIIHD 700 spectrometer and the concentrations for G, GG, cGG, GGG, and AAA were determined. The data were collected using a 30-degree pulse program with a 15 second relaxation delay to ensure quantitative integration of the resonances. The signals for each species were separated on the <sup>1</sup>H NMR spectra

and the quantification of each species was based on the integrated signal intensities relative to the internal standard intensity. All the  $^1\text{H}$  NMR spectra were plotted and analyzed using MestReNova 9.1.

The kinetic model describing the degradation rates of the glycine oligomers is listed as below.

$$\frac{dn_G}{dt} = (2 k_{sc} n_{GG} + 2 k_{sc} n_{GGG} + k_{bb}n_{GGG}) \quad (1)$$

$$\frac{dn_{GG}}{dt} = (-k_{sc} n_{GG} - k_{rc}n_{GG} + 2 k_{ro} n_{cGG} + 2 k_{sc} n_{GGG}) \quad (2)$$

$$\frac{dn_{cGG}}{dt} = (k_{rc} n_{GG} - 2 k_{ro} n_{cGG} + k_{bb}n_{GGG}) \quad (3)$$

$$\frac{dn_{GGG}}{dt} = (-2 k_{sc} n_{GGG} - k_{bb}n_{GGG}) \quad (4)$$

where  $n_G$  is the amount of glycine,  $n_{GG}$  is the amount of linear diglycine,  $n_{GGG}$  is the amount of linear triglycine, and  $n_{cGG}$  is the amount of cyclic diglycine; the units are in  $\mu\text{mol}$ . The four rate constants are  $k_{sc}$ , the rate constant corresponding to the hydrolysis of the glycine oligomers through a scission pathway;  $k_{bb}$ , the rate constant for the backbiting pathway of GGG; and  $k_{rc}$  and  $k_{ro}$ , the rate constants for cGG ring closing and opening, respectively.

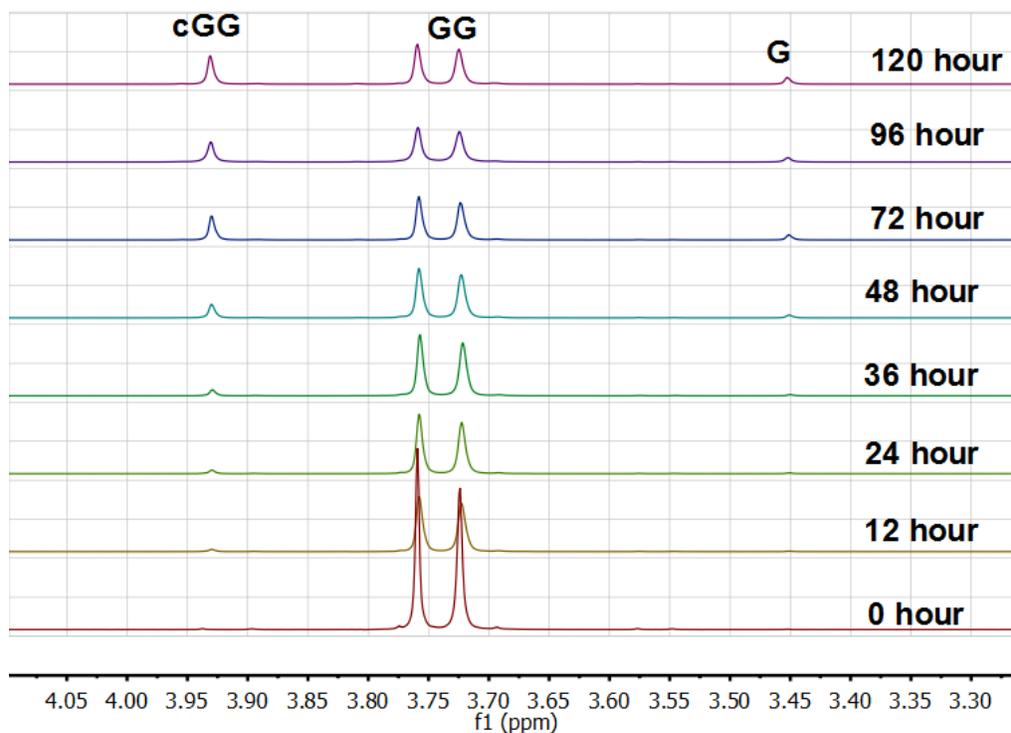
All the reactions are assumed to be pseudo-first order. As shown in Fig. 2, there are seven parameters in the fitting process, including the four rate constants, as well as the initial amounts of the reactants: GG, cGG and GGG. In this investigation, the initial amounts are estimated so as not to give inordinate weight to the first measurement compared to the subsequent measurements. The parameter estimates are obtained using MATLAB, using the ode45 function to solve the differential equations, and the patternsearch function to find the parameter values that minimize the overall sum-squared error of the model fit.

## 2.4 Results

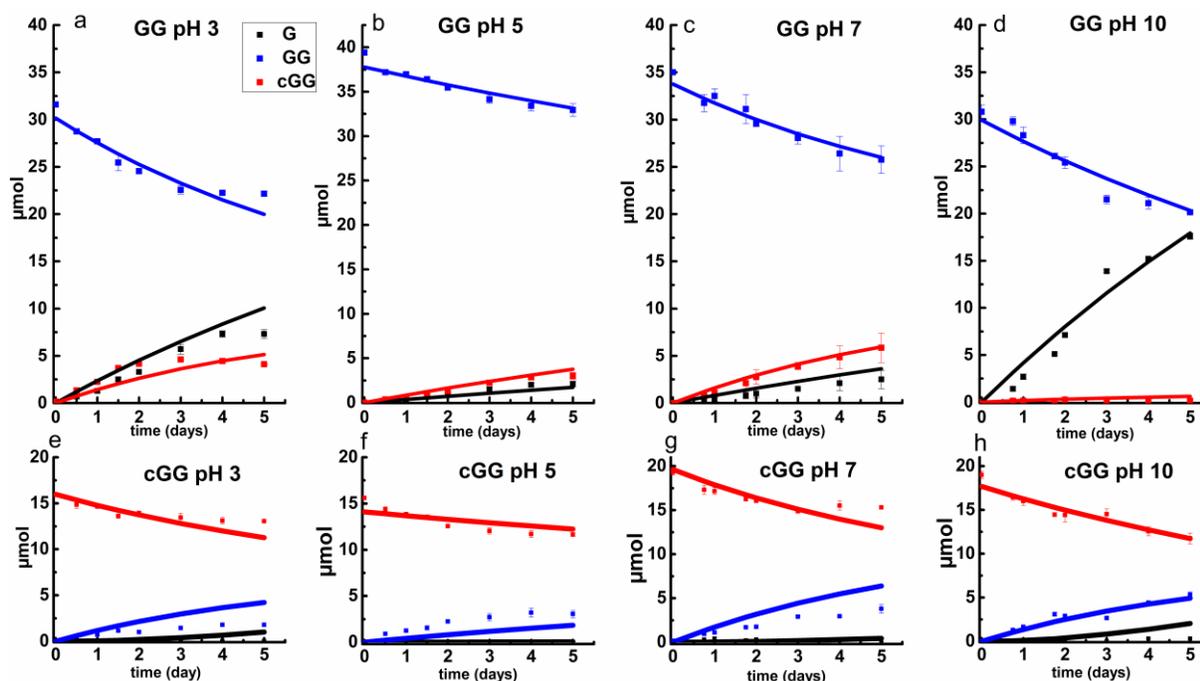
### 2.4.1 Glycine dimer reaction kinetics

The linear dimer GG can undergo two reactions (Figure 2.2.1): (a) The amide bond in GG can be cleaved via an acid-catalyzed attack by a water molecule at the carbonyl carbon or by a direct attack of hydroxide ion yielding two glycine monomers G, and (b) the terminal amino group of GG can react with the proximate carboxyl to form the six-membered ring cGG by an acid-catalyzed cyclization process.  $^1\text{H}$  NMR was used to monitor the reactions of GG at pH 3, 5, 7, and 10. As an example, Figure 2.4.1 shows the stacked  $^1\text{H}$  NMR spectra for GG degradation at pH 7 at 95°C over a period of five days. The signals for G, GG, and cGG are well resolved. The  $^1\text{H}$  NMR spectra for reactions of linear GG degradation at all four pH values are shown in Supplemental Information: Figures S1-S4. For linear GG degradation, the amount of the reactant (GG) and products (G and cGG) were determined from the integrated NMR signals, shown in Supplemental Information: Tables S1-S4. For cGG degradation, the related stacked  $^1\text{H}$  NMR spectra for the reactions of GG at pH 3, 5, 7, and 10 are shown in Supplemental Information: Figure S5-S8, and the amount of the reactant (cGG) and products (G and GG) were shown in Supplemental Information: Tables S5-S8. Figure 2.4.2 graphically illustrates the experimental (squares) and model-based (solid lines) profiles for GG, cGG, and G at pH 3, 5, 7, and 10 starting with GG and cGG. Overall, the experimental rate of reaction for GG is faster at pH 3 and 10 compared to pH 5 and 7. The mechanistic pathways for the reaction of GG under acidic and basic pH are outlined in Figure 2.4.3 and 2.4.4, respectively. In an acidic medium, the electrophilicity of the carbonyl carbon of the amide linkage is enhanced by the coordination of a proton with the carbonyl oxygen. The resulting increased electrophilicity of the carbonyl carbon facilitates reaction with weakly nucleophilic water in the formation of the tetrahedral intermediate which

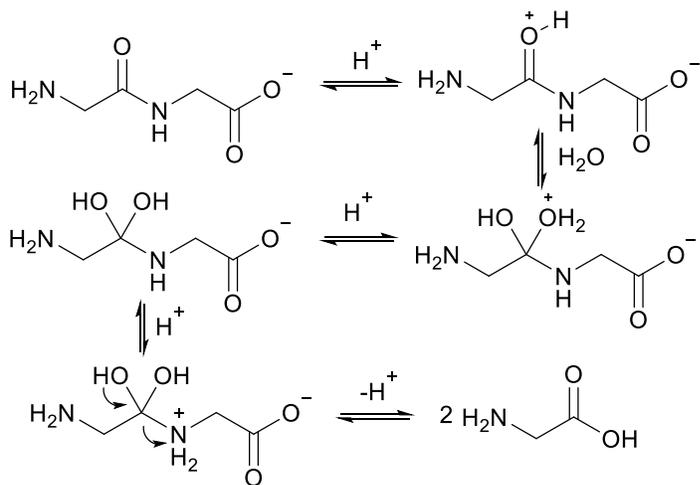
subsequently decomposes into the two glycine monomers. In contrast, under basic conditions, the hydroxide ion is a strong enough nucleophile to directly attack the carbonyl carbon to form an analogous tetrahedral intermediate. This intermediate then decomposes to form the glycine products.



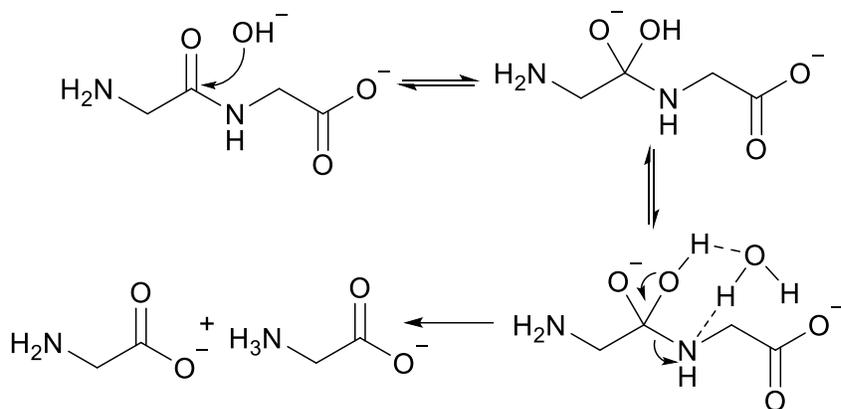
**Figure 2. 4. 1 - Stacked <sup>1</sup>H NMR spectra for GG reactions at 95°C at pH 7. Each spectrum corresponds to a different sampling time, ordered in time from bottom to top.**



**Figure 2. 4. 2 - Glycine dimer (GG) (a-d) and cyclic glycine dimer (cGG) (e-h) reaction kinetics at pH 3, 5, 7 and 10 at 95°C over a period of five days. The squares represent experimental data obtained from 1H NMR measurements, while the solid lines represent the model prediction. (GG)=blue; (G)=black; (cGG)=red.**

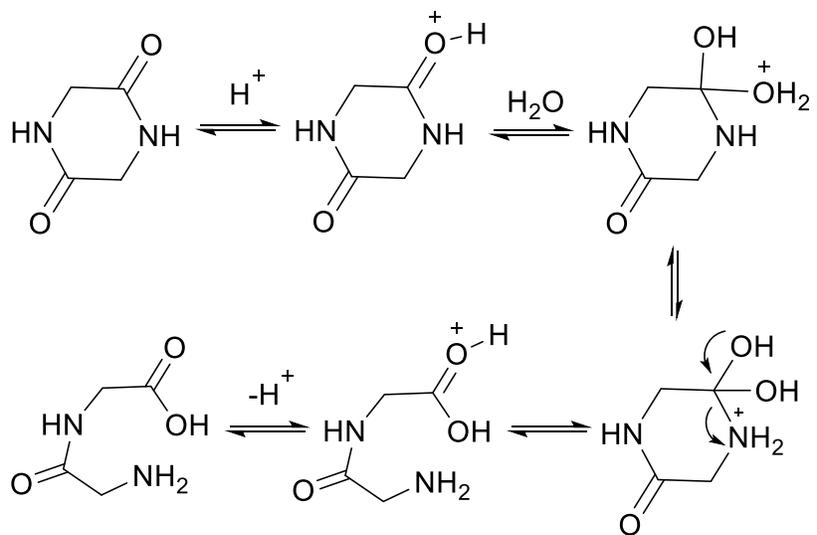


**Figure 2. 4. 3 - Mechanistic pathway for GG hydrolysis at pH 3.**



**Figure 2. 4. 4 - Mechanistic pathway for GG hydrolysis at pH 10.**

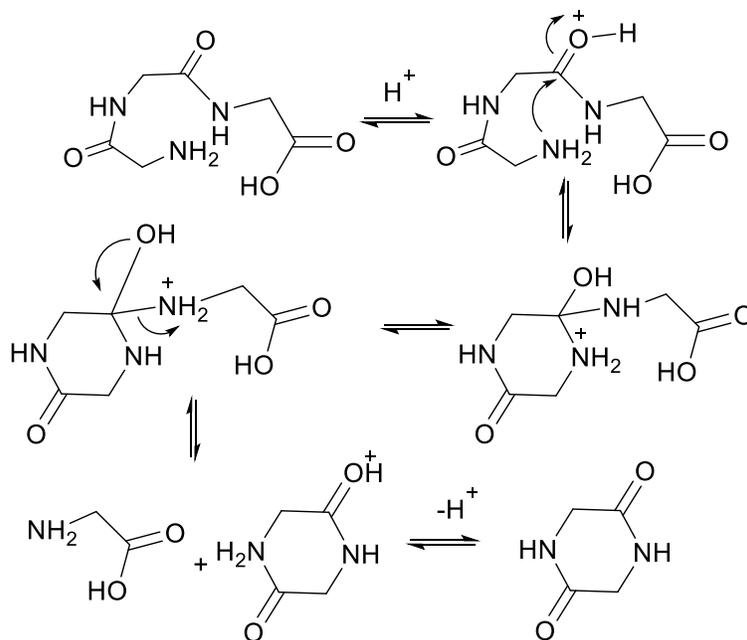
Figure 2.4.4 also demonstrates that the rates of ring opening are quite slow over the five-day time period, for the entire pH range studied. Earlier investigations have shown that the linear glycine dimer could be in equilibrium with the cyclic dimer.<sup>27,28,29</sup> This can only be the case under acidic or perhaps neutral conditions. Figure 2.4.5 shows the acid-catalyzed mechanistic pathway for the acid-catalyzed ring-opening of cGG; every step is reversible. The principle of microscopic reversibility dictates that the corresponding acid-catalyzed ring-closing mechanism is just the reverse of the pathway shown in Figure 2.4.5. In contrast, under basic conditions (pH 10), the product of reaction (the linear glycine dimer) has a terminal carboxylate anion which is not susceptible to nucleophilic attack by the proximate amino group to form cGG; the final step in the ring opening at pH 10 is irreversible. Indeed, at pH 10, the experimental formation of cGG is essentially zero (Figure 4d).



**Figure 2. 4. 5 - Mechanistic pathway for the acid-catalyzed ring-opening of cGG.**

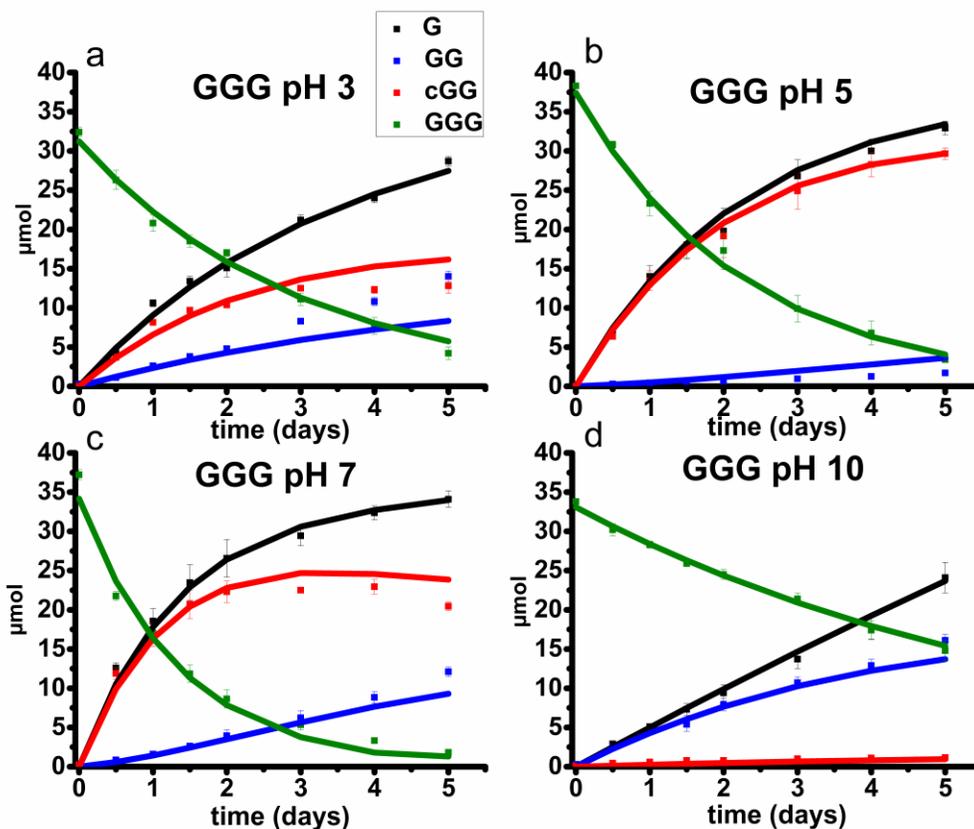
### 2.4.2 Glycine trimer (GGG) reaction kinetics

According to the reaction pathways displayed in Figure 2.2.1, there are two reaction pathways for the GGG reactions: (a) the formation of the cyclic dimer (cGG) and glycine (G) by means of a backbiting process and (b) the formation of the linear dimer and glycine via a direct scission of one of the amide linkages. The detailed description of the mechanistic pathway for each of these processes depends on the pH of the aqueous reaction medium. The acid-catalyzed backbiting process is described in Figure 2.4.6 while the mechanism for the competing acid-catalyzed scission to the linear glycine dimer and glycine is essentially the same as that shown in Figure 2.4.3. The corresponding backbiting and amide cleavage mechanisms in basic media are similar to Figure 2.4.4. The subsequent reactions of the glycine dimer have already been discussed (Section 1). It should be emphasized that the backbiting pathway is an intramolecular process which always begins at the N-terminal amino acid unit and, in order to proceed, the amine group must not be protonated. In contrast, the intermolecular amide hydrolysis pathway by water or hydroxide can, in principle, take place anywhere along the polypeptide chain.



**Figure 2. 4. 6 - Acid-catalyzed backbiting mechanism for the formation of cGG and glycine.**

The stacked  $^1H$  NMR spectra and the tabulated peak integration values for GGG are summarized in Supplemental Information: Figures S9-S12 and Tables S9-S12. Figure 2.4.7 graphically illustrates the experimental (squares) and model-based (solid lines) profiles for GGG, GG, and cGG with respect to time at 95°C at pH 3, 5, 7 and 10 starting with GGG. Interestingly, the degradation rate of GGG is fastest at pH 7 and appears to decrease as the medium becomes more acidic or basic. At pH 10 the rate is substantially slower compared to the other pH conditions studied. These results are in stark contrast to the GG kinetic profiles discussed in Section 1 where the fastest rates occurred at pH 3 and 10. These observations suggest that the dominant reaction pathway for GGG degradation may be different from that of GG.



**Figure 2. 4. 7 - Glycine trimer (GGG) reaction kinetics at pH 3, 5, 7 and 10 at 95°C for a period of five days (a-d). The symbols denote 1H NMR measured abundance and solid lines are the model predictions. (GGG)=green; (GG)=blue; (G)=black; (cGG)=red**

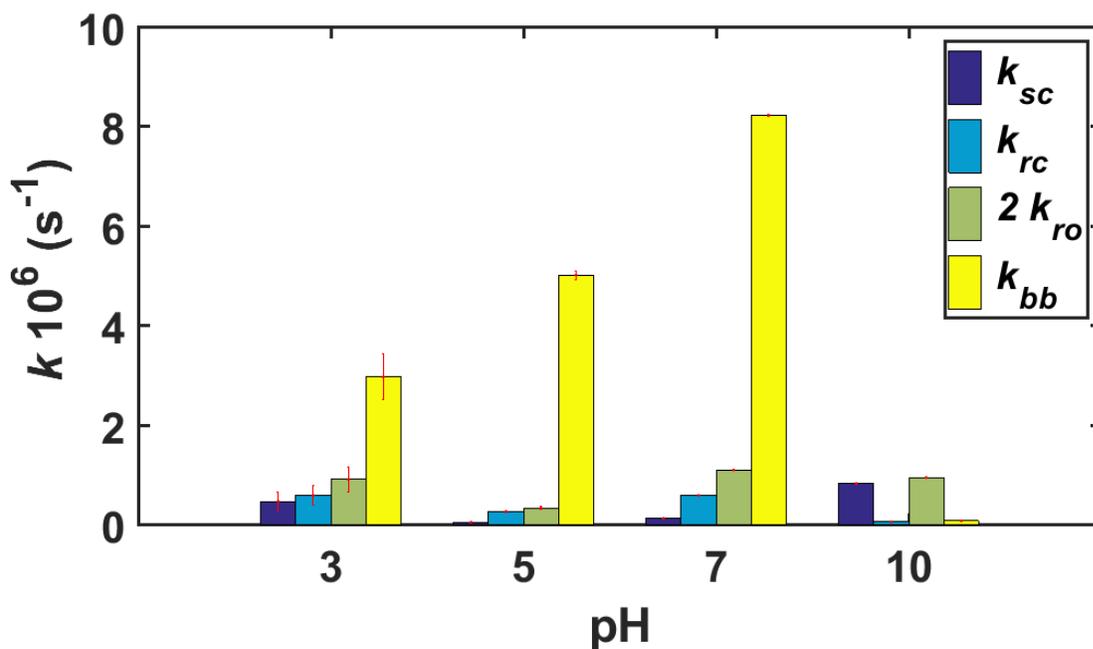
Figure 2.4.7 shows that the initial rates for the reaction of GGG via the backbiting process steadily increase from pH 3 to pH 7. At pH 5 and 7, cGG and G are initially produced in a 1:1 molar ratio which is consistent with the operation of the GGG backbiting pathway. As the reaction progresses, however, a deviation from the 1:1 ratio is observed due to the accompanying ring opening reaction of the cyclic dimer (cGG) producing the linear dimer (GG) which can subsequently form the

monomeric unit G. Overall these results clearly demonstrate that the backbiting reaction mechanism is favored under neutral pH conditions.

Backbiting at pH 10 appears to be negligible. At pH 10 the initial production of GG and G occurs in approximately a 1:1 molar ratio suggesting that the scission pathway is operating. The amount of cGG formed at pH 10 is negligible indicating that the scission mechanism is the favored pathway. At pH 3 both the scission and the backbiting mechanisms contribute to the reaction process. It is concluded that the competitive pathways (backbiting and random scission) for the reaction of GGG is strongly dependent on the pH of the aqueous medium and that backbiting is an important pathway for the GGG reaction in both acidic and neutral media.

### 2.4.3 Rate constants based on proposed kinetics model

Since all three sets of kinetic experiments (GG, cGG and GGG) share similar reaction pathways, the three sets of experimental data were fit with a shared set of rate constants at pH 3, 5, 7, and 10. Figure 2.4.8 graphically compares the pH-dependent rate constants shown in Figure 2.2.1 based on the fit to the proposed kinetic model. The rate constants and the initial amounts for the starting materials derived from the kinetic model for each of these reactions are also tabulated in Supplemental Information: Tables S13-S14.

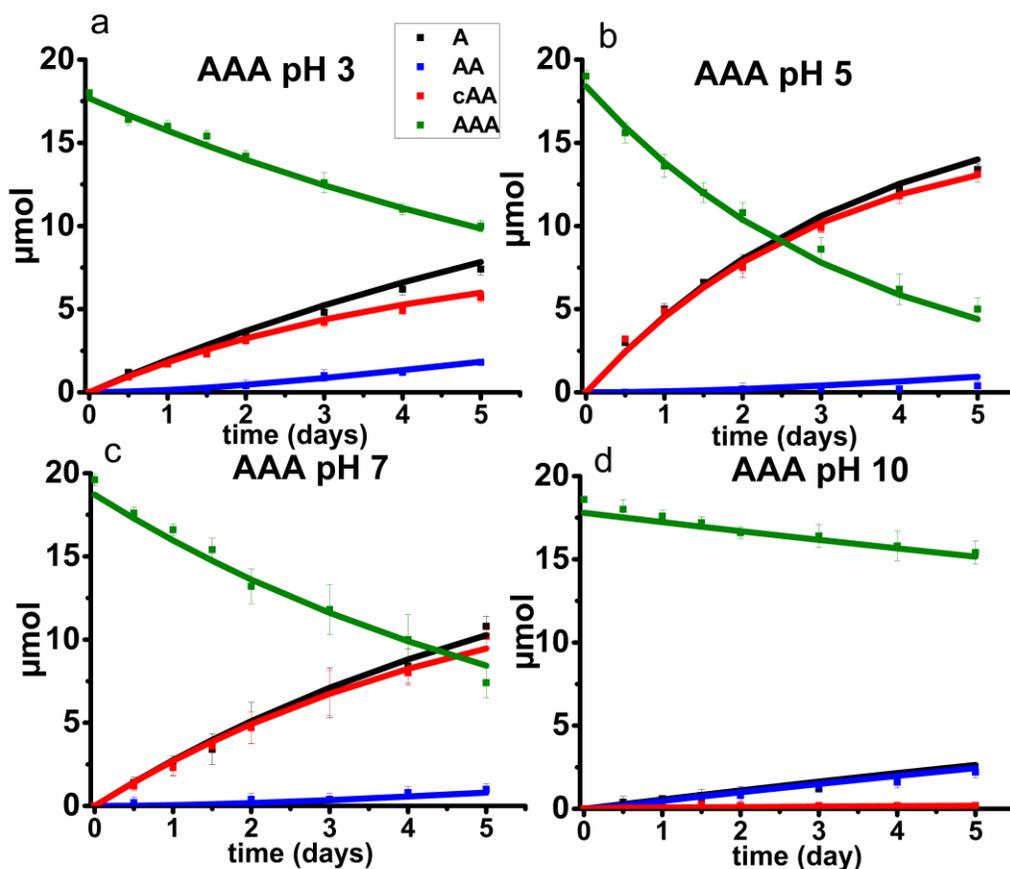


**Figure 2. 4. 8 - The estimated four rate constants for GG, cGG and GGG kinetics at pH 3, 5, 7 and 10 at 95°C. The confidence intervals are calculated following the chi-squared method (discussed in Supplemental Information) at the 95% confidence level.**

Overall, the highest rate constants are associated with backbiting, over the range of pH from 3-7, although the rate constant for backbiting is extremely low at pH = 10. The highest values of the scission rate constant are observed at acidic and basic pH, with lower rates near neutral pH.

#### 2.4.4 Alanine trimer (AAA) reaction kinetics

To support the generality of the glycine studies, the peptide tri-alanine (AAA) was also investigated. The results are shown in Figure 2.4.9. The stacked  $^1\text{H}$  NMR spectra and the tabulated peak integration values for AAA are summarized in Supplemental Information: Figure S13-S16 and Tables S15-S18. Alanine racemization is not significant at  $95^\circ\text{C}$  and was not included in the model.<sup>45</sup> The optimized reaction rates for each reaction pathway and the initial starting AAA amounts at each pH condition are listed in Supplemental Information: Tables S19-S20.



**Figure 2. 4. 9 - Alanine trimer (AAA) degradation reaction kinetics under pH 3, 5, 7 and 10, at  $95^\circ\text{C}$ , over 5 days (a-d). The symbols denote  $^1\text{H}$  NMR measured quantifications and solid lines are the model predictions. (AAA)=green; (A)=black; (AA)=blue; (cAA)=red**

Overall, the same trends in pH are observed for trialanine as for triglycine. Specifically, backbiting is the major reaction pathway at pH 3, 5, and 7 where a 1:1 molar ratio of cyclic alanine dimer (cAA) to monomer (A) are produced. This observation is similar to that of the GGG reaction within the same pH range. Only a limited amount of linear dimer (AA) is detected, suggesting slower rate of ring opening of the cAA compared with cGG (Figures 2.4.2 and Figure 2.4.9). In contrast, only the scission pathway is observed at pH 10. The overall reaction rates are slower for AAA compared to GGG, likely due to steric hindrance from the methyl group on the alpha carbon of alanine.

## 2.5 Discussion

Previous kinetic modeling studies of peptide cleavage have not used atmospheric pressure and temperatures below 100°C. Radzicka and Wolfenden<sup>27</sup> studied peptide cleavage at neutral pH with higher temperatures. Extrapolation of their degradation rate constant to 95°C yields  $k_{sc} = 1.5 \times 10^{-7} \text{ s}^{-1}$ , similar to our estimate of  $k_{sc} = 1.4 \times 10^{-7} \text{ s}^{-1}$ . Similarly, extrapolation of the model from Sakata *et al.*<sup>38</sup> at pH = 9.8 yields  $k_{sc} = 6.6 \times 10^{-7} \text{ s}^{-1}$ , compared to our estimate of  $k_{sc} = 8.4 \times 10^{-7} \text{ s}^{-1}$  at pH 10. Thus, the results presented here are consistent with past reports, while providing a comprehensive quantitation of the cleavage reaction network, from acidic to basic pH, and measured at atmospheric pressure.

While scission and backbiting are both significant in this study, the ring opening reaction of the diketopiperazines is very slow at all pH values considered. The ease of formation and the stability of the cyclic dimers presents one of the greatest obstacles in our understanding of the prebiotic origin of polypeptides.<sup>15, 46</sup> Once formed, the cyclic dimer is extremely stable and presents a dead-end for further polymerization under plausible prebiotic conditions. However, as shown here, basic conditions can be used to retard the ring-closure reaction.

In the model presented here, the rate of scission for each peptide bond is equal, independent of peptide length. Thus, the trimer will degrade by scission at twice the rate of the linear dimer, since the trimer has two peptide bonds. More generally, the degradation rate of any homopolymeric peptide with the length of  $n$  units could be described as the summation of two terms, the backbiting reaction rate and the random scission rate:

$$k_{obs} = k_{bb} + (n - 1)k_{rs} \quad (5)$$

The assumption of equal scission rates for all  $n$  would be valid when no macromolecular structures are formed. Thus, even though backbiting may appear to be quite dominant at pH 3-7 in Figure 10, scission might also be important for longer peptides, especially at acidic pH. A model of the reaction network, such as the one presented here, enables the quantification of these multiple competing mechanisms.

After obtaining the kinetics data for the two competing reaction pathways for longer trimer degradation, it is possible to look back on the preliminary experiment and relate the peptides degradation kinetics with the depsipeptides degradation. For the preliminary project described in Section 1.3.1, the depsipeptide pool is subjected to a total of 90 hours heating at 85°C for the dry phase and a total of 30 hours recycling at 95°C for the wet phase. As shown in Figure 1.3.2, most of the species formed in the final depsipeptide pool are depsipeptides with alternating LA and G. For 85 °C /95 °C dry/wet cycling condition shown in Figure 1.3.1, peptides (2G, 4G, 6G) with very limited amount appear in the MS spectrum as well. The degradation of those peptides should follow the same kinetics model as discussed in the Chapter 2 considering both backbiting and random scission reaction pathways. Based on the GG peptide degradation kinetics results, around 5.18% of GG could be degraded in 30 hours at 95°C at pH 3. The degradation of the depsipeptides should be limited to only random scission reaction pathways since the LA on the N-terminus cannot attack the carbonyl carbon, blocking the backbiting reaction mechanism.

## 2.6 Conclusions

The main purpose of this thesis is to study the degradation reactions related to the hydrolysis of longer polymers into shorter ones in prebiotically conditions. Selective of stable peptides involves not only the peptide polymerization reactions, another important pathway is the recycling of longer proto-peptides back into starting materials and reshuffling of peptide sequences, through the degradation of amide bond in the peptide backbone.

In this thesis, the depsipeptides recycling preliminary results shown in Chapter 1.3.1 have demonstrated that the glycine dimer could be degraded into two monomers under the wet phase, and then the regenerated monomer could further be polymerized into new depsipeptides species under the dry phase, coupled in the ester amide exchange reaction system. Therefore, it is plausible to show that the designed ester amide exchange reaction system involves polymerization reactions, selective degradations and regenerations of new depsipeptides driven by the alternating wet and dry environments.

As previously mentioned, the driving force for peptide recycling and reshuffling processes is assumed to be the amide bond degradation and therefore, I set up the second project for this thesis on the study of the selected reaction mechanisms for peptide bond degradations under different pH conditions. There are two reaction pathways for peptide degradation, either through direct scission or backbiting reaction mechanisms. I have traced GG, cGG, GGG and AAA degradation kinetics under 95°C for five days and optimized the reaction rates for each reaction pathways.

Dipeptide cleavage in water at atmospheric pressure occurs by direct scission, and is much faster at acidic and basic pH, compared to neutral pH. In contrast, tripeptide cleavage under the same conditions occurs by both scission and backbiting. The overall observed cleavage in tripeptides is

fastest at neutral pH, due to backbiting. At acidic pH, both backbiting and scission are active pathways.

Moving forward, it is necessary to search for some functional biopolymers selected from the prebiotic soup. One main plausible catalytic reactivity worth exploration is the aldol reaction, with the creation of a new carbon-carbon bond catalysed by the peptides free N terminus. Generally, the aldol reaction<sup>47</sup> is designed with the cyclohexanone and p-nitrobenzaldehyde as substrate. It is plausible to use NMR to calculate the yields and screen the catalytic reactivities corresponding to different short peptide catalysts. The literature<sup>47</sup> supports that glycine and alanine-related short peptides could catalyze this aldol reaction and also, proline related peptides<sup>48</sup> are able to catalyze the aldol reaction with stereospecific tetroses. The aldol reaction is one simple catalytic reaction that could demonstrate the emergence of functional peptides through the prebiotic soup and screen their reactivities.

## APPENDIX A. SUPPLEMENTS

### A.1 Depsipeptides recycling preliminary experiment

#### A.1.1 Method

The reaction typically started with mixing glycine dimer (100mM, 200uL) and lactic acid (LA) (100mM, 200uL), and then subjected to six wet (95°C)-dry (85°C) cycles. The solution was added into the 2 mL glass vials (Supelco 29381-U) and heated in the oven over six days. The temperature condition is adjusted by a temperature controller connected with the oven. The experiment started from the wet phase with the cap closed for 6 hours at 85°C. Then the dry phase started by adjusting the temperature controller to 95°C and keep the cap open for 18 hours. The second cycle started with the rehydration of 400uL water, and the whole process is repeated for six times. The initial and final pH conditions before and after six days of experiment are always around 3, without significant fluctuation. After six days of wet-dry cycles, the samples were diluted by a factor of 100 and then analyzed by the mass spectrometer.

The MS and MS/MS spectra in Figure 1.3.1 were analyzed with Waters Xevo G2 mass spectrometer with help from Dr. Jay Forsythe. The MS spectra in Figure 1.3.2 were analyzed using an Agilent 6130 single quadrupole mass spectrometer. For MS analysis, all data were obtained in negative mode electrospray ionization with a capillary voltage of 2.0 kV.

### A.2 Peptide bond cleavage project

#### A.2.1 Mathematical model

The kinetic model describing the degradation rate of glycine oligomer hydrolysis is listed as below as Equations (1)–(4). All the reactions are assumed to be pseudo-first order, since the water concentration is approximately constant at 55M.

$$\frac{dn_G}{dt} = (2 k_{sc} n_{GG} + 2 k_{sc} n_{GGG} + k_{bb}n_{GGG}) \quad (1)$$

$$\frac{dn_{GG}}{dt} = (-k_{sc} n_{GG} - k_{rc}n_{GG} + 2 k_{ro} n_{cGG} + 2 k_{sc} n_{GGG}) \quad (2)$$

$$\frac{dn_{cGG}}{dt} = (k_{rc} n_{GG} - 2 k_{ro} n_{cGG} + k_{bb}n_{GGG}) \quad (3)$$

$$\frac{dn_{GGG}}{dt} = (-2 k_{sc} n_{GGG} - k_{bb}n_{GGG}) \quad (4)$$

where  $n_G$  is the amount of glycine,  $n_{GG}$  is the amount of linear diglycine,  $n_{GGG}$  is the amount of linear triglycine, and  $n_{cGG}$  is the amount of cyclic diglycine; the units are in  $\mu\text{mol}$ . The four rate constants are  $k_{sc}$ , the rate constant corresponding to the hydrolysis of the glycine oligomers through a scission pathway;  $k_{bb}$ , the rate constant for the backbiting pathway of GGG; and  $k_{rc}$  and  $k_{ro}$ , the rate constants for cGG ring closing and opening, respectively. There are seven parameters in the optimization process, including four rate constants, as well as the actual initial amounts for the reactants. The parameters are obtained using MATLAB, with the ode45 function to solve the differential equations with a fourth-order Runge-Kutta method, and the patternsearch function to find the parameters that minimize the overall sum-squared error of the model fit.

### A.2.2 Confidence interval calculation

The uncertainty in the parameter estimates are calculated based on statistical measures.<sup>49</sup> The quantity  $N_m$  is the number of measured variables, which is equal to 4 here, since we have four measured variables: G, GG, cGG and GGG;  $N_d$  is the number of samples of each measured variable, which is equal to 7 in our case, since we have seven time points (not counting the initial

point);  $N_p$  is the number of parameters, which is equal to 7 in our case, for  $k_1-k_4$  and the three initial conditions for GG, cGG, and GGG.

The error covariance matrix ( $V_{ii}$ ) is calculated is calculated following Equations (5)–(6).

$$V_{ii} = \frac{1}{N_d} \sum_{j=1}^{N_d} e_{ij}^2(k) \quad (5)$$

$$e_{ij} = y_{ij} - \widetilde{y}_{ij} \quad (6)$$

where  $y_{ij}$  and  $\widetilde{y}_{ij}$  represent the actual and model prediction of the  $i^{\text{th}}$  measured variable and  $j^{\text{th}}$  sampling time. Assuming the model can be represented by linear functions, the Jacobian matrix is obtained as

$$B_{ij} = \left. \frac{\partial \widetilde{y}_{ij}}{\partial k} \right|_{k=k^*} \quad (7)$$

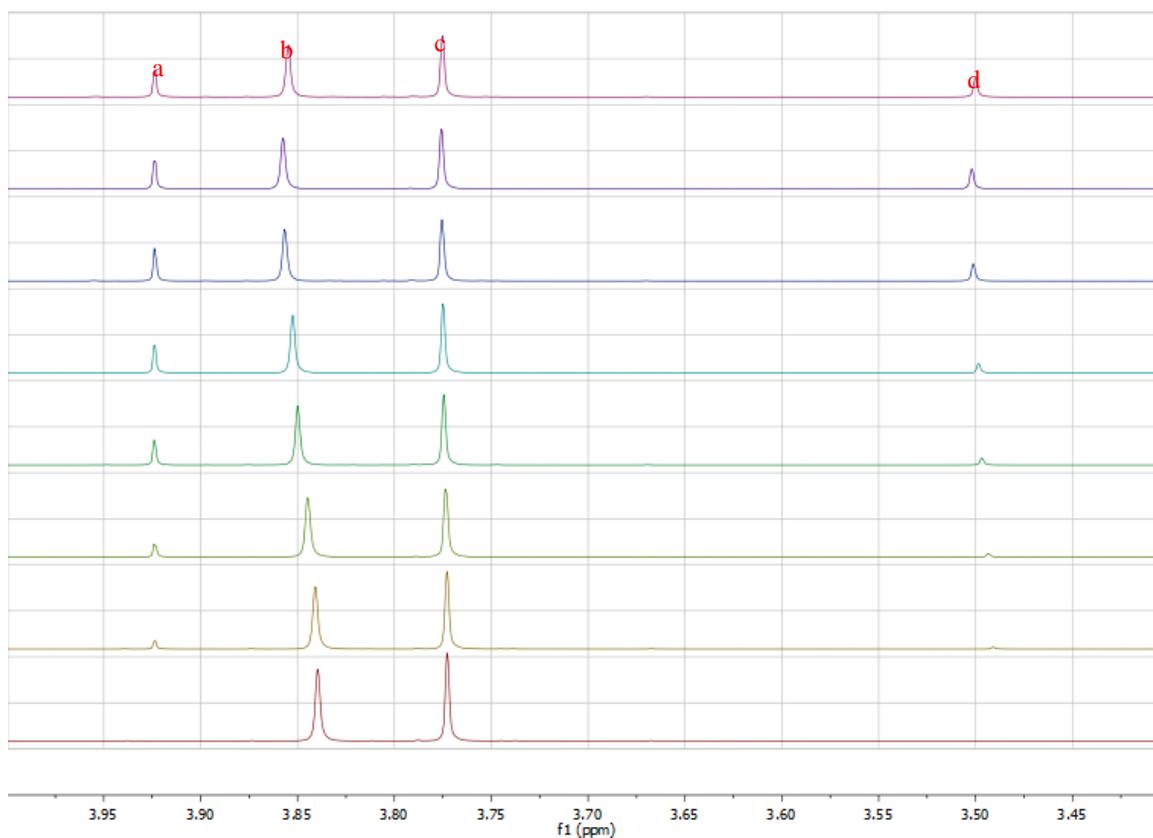
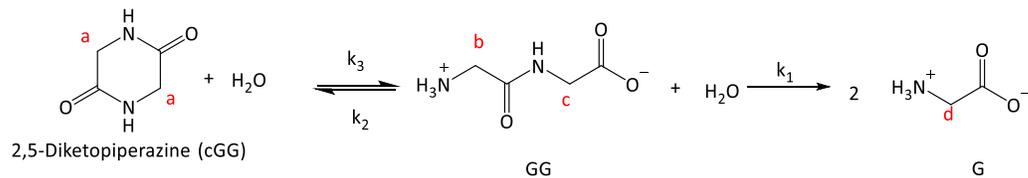
The 95% confidence interval for each parameter is obtained with  $\widehat{k} - k^*$  as the boundary following Equations (8)–(10).

$$V_{\theta}^{-1} = \sum_{i=1}^{N_m} \sum_{j=1}^{N_d} B_{ij}^T V^{-1} B_{ij} \quad (8)$$

$$(\widehat{k} - k^*)^T V_{\theta}^{-1} (\widehat{k} - k^*) = \chi_{N_p}^2(0.95) \quad (9)$$

$$(\widehat{k} - k^*) = \sqrt{\chi_{N_p}^2(0.95)/V_{\theta}^{-1}} \quad (10)$$

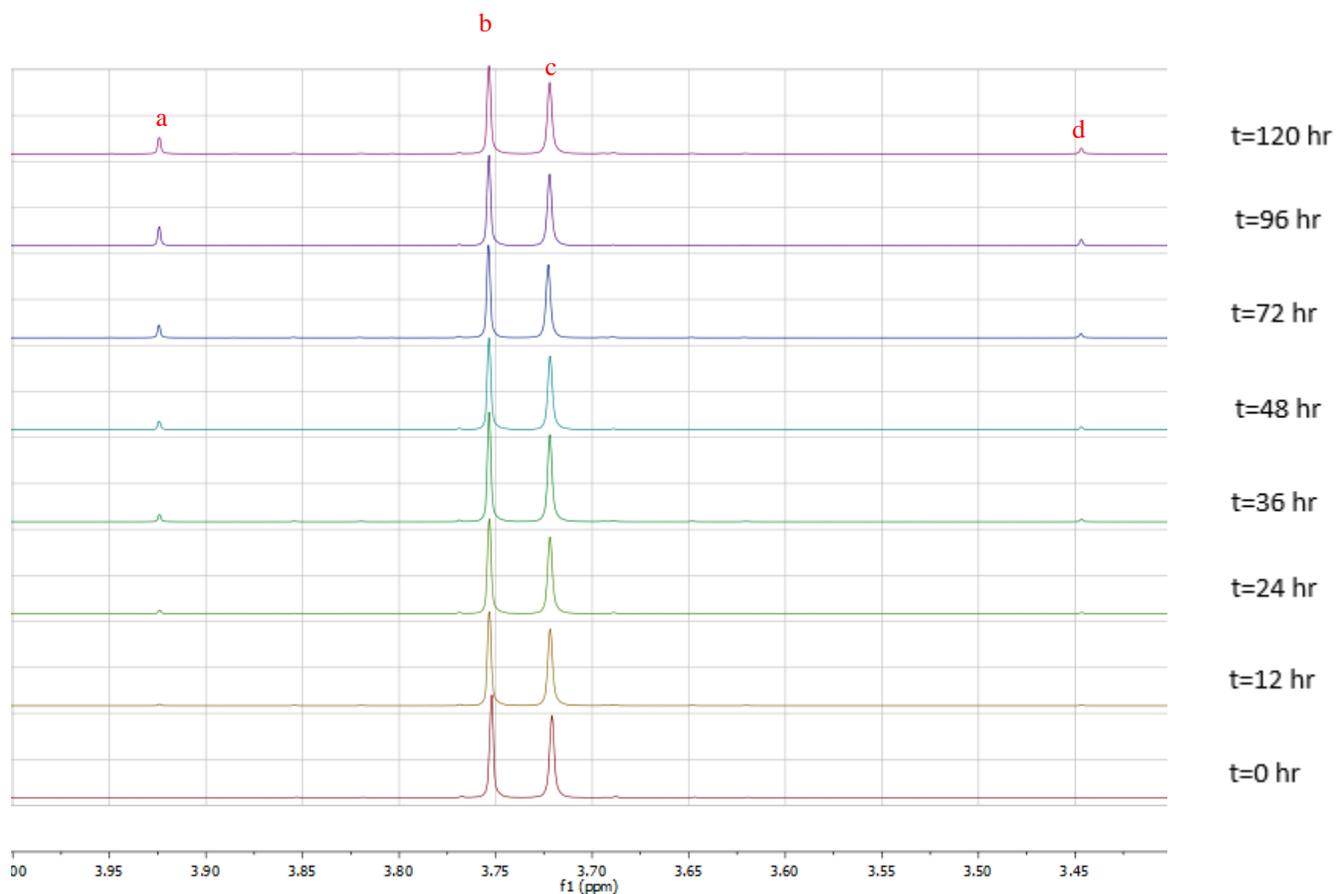
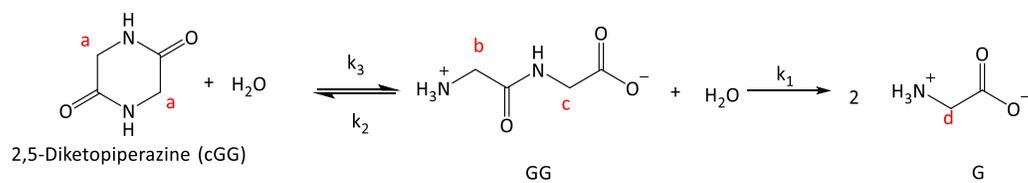
### A.3 Supporting materials



**Figure S1.** Stacked <sup>1</sup>H NMR spectra for GG decomposition at 95°C under pH 3. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.4ppm to 4.0ppm.

**Table S1.** GG degradation reactants and products amounts at varying reaction time under pH 3 at 95°C.

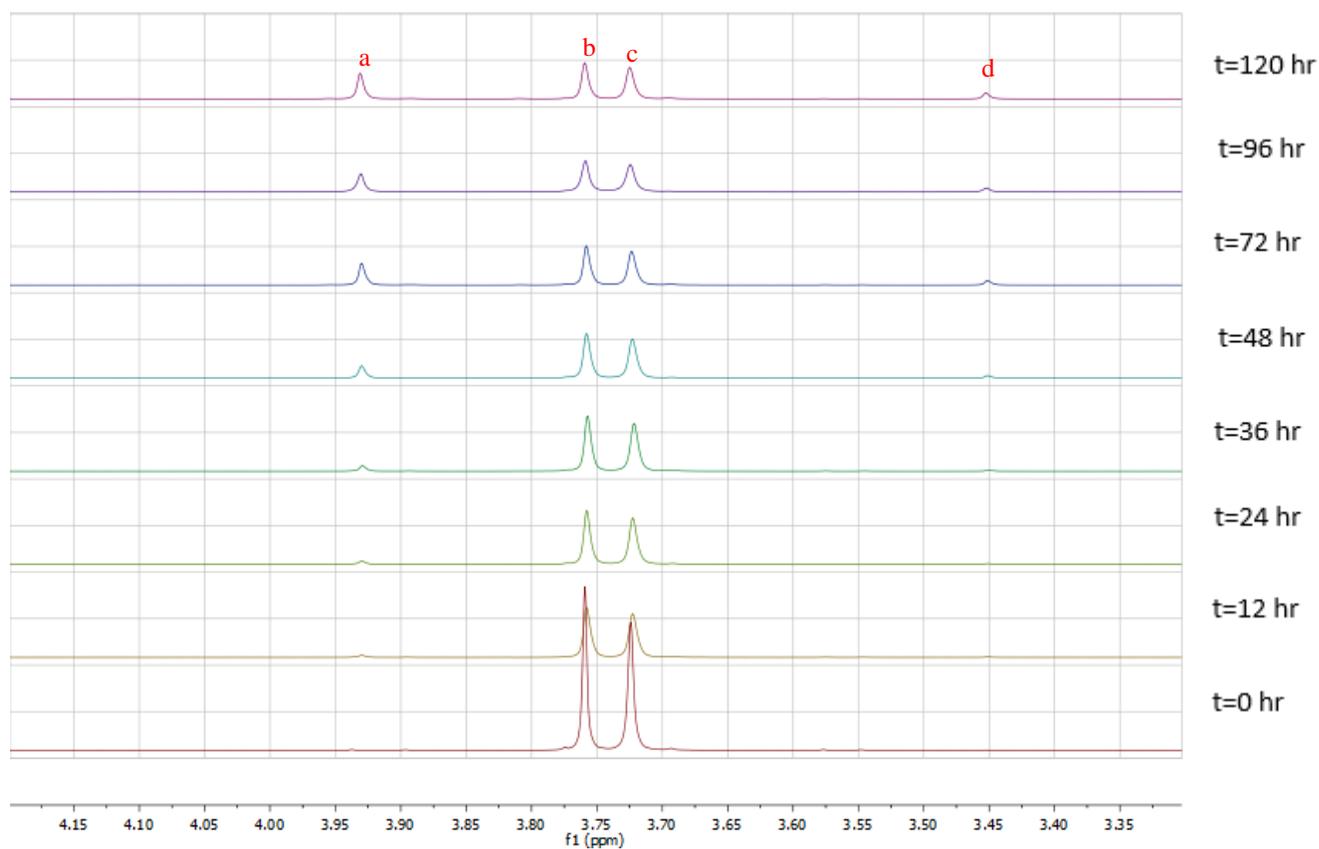
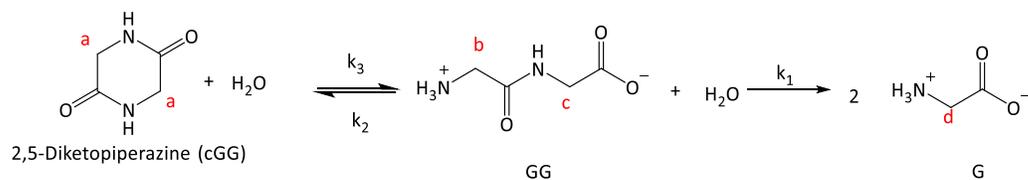
pH=3 GG analysis	NMR integrations				actual amount ( $\mu\text{mol}$ )		
time(h)	KHP	G	GG	cGG	G( $\mu\text{mol}$ )	GG( $\mu\text{mol}$ )	cGG ( $\mu\text{mol}$ )
0	1	0	2.1	0.01	0	31.5	0.15
0	1	0	2.1	0.01	0	31.5	0.15
0	1	0	2.12	0.01	0	31.8	0.15
12	1	0.02	1.9	0.08	0.6	28.5	1.2
12	1	0.03	1.93	0.1	0.9	28.95	1.5
12	1	0.02	1.92	0.09	0.6	28.8	1.35
24	1	0.04	1.85	0.14	1.2	27.75	2.1
24	1	0.05	1.84	0.16	1.5	27.6	2.4
24	1	0.04	1.85	0.15	1.2	27.75	2.25
36	1	0.08	1.68	0.25	2.4	25.2	3.75
36	1	0.09	1.65	0.25	2.7	24.75	3.75
36	1	0.08	1.76	0.24	2.4	26.4	3.6
48	1	0.11	1.63	0.27	3.3	24.45	4.05
48	1	0.11	1.63	0.28	3.3	24.45	4.2
48	1	0.11	1.65	0.28	3.3	24.75	4.2
72	1	0.2	1.47	0.31	6	22.05	4.65
72	1	0.17	1.52	0.3	5.1	22.8	4.5
72	1	0.2	1.52	0.31	6	22.8	4.65
96	1	0.25	1.49	0.3	7.5	22.35	4.5
96	1	0.25	1.47	0.3	7.5	22.05	4.5
96	1	0.23	1.49	0.29	6.9	22.35	4.35
120	1	0.23	1.49	0.27	6.9	22.35	4.05
120	1	0.26	1.46	0.27	7.8	21.9	4.05
120	1	0.24	1.48	0.28	7.2	22.2	4.2



**Figure S2.** Stacked <sup>1</sup>H NMR spectra for GG decomposition at 95°C under pH 5. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.4ppm to 4.05ppm.

**Table S2.** GG degradation reactants and products amounts at varying reaction time under pH 5 at 95°C.

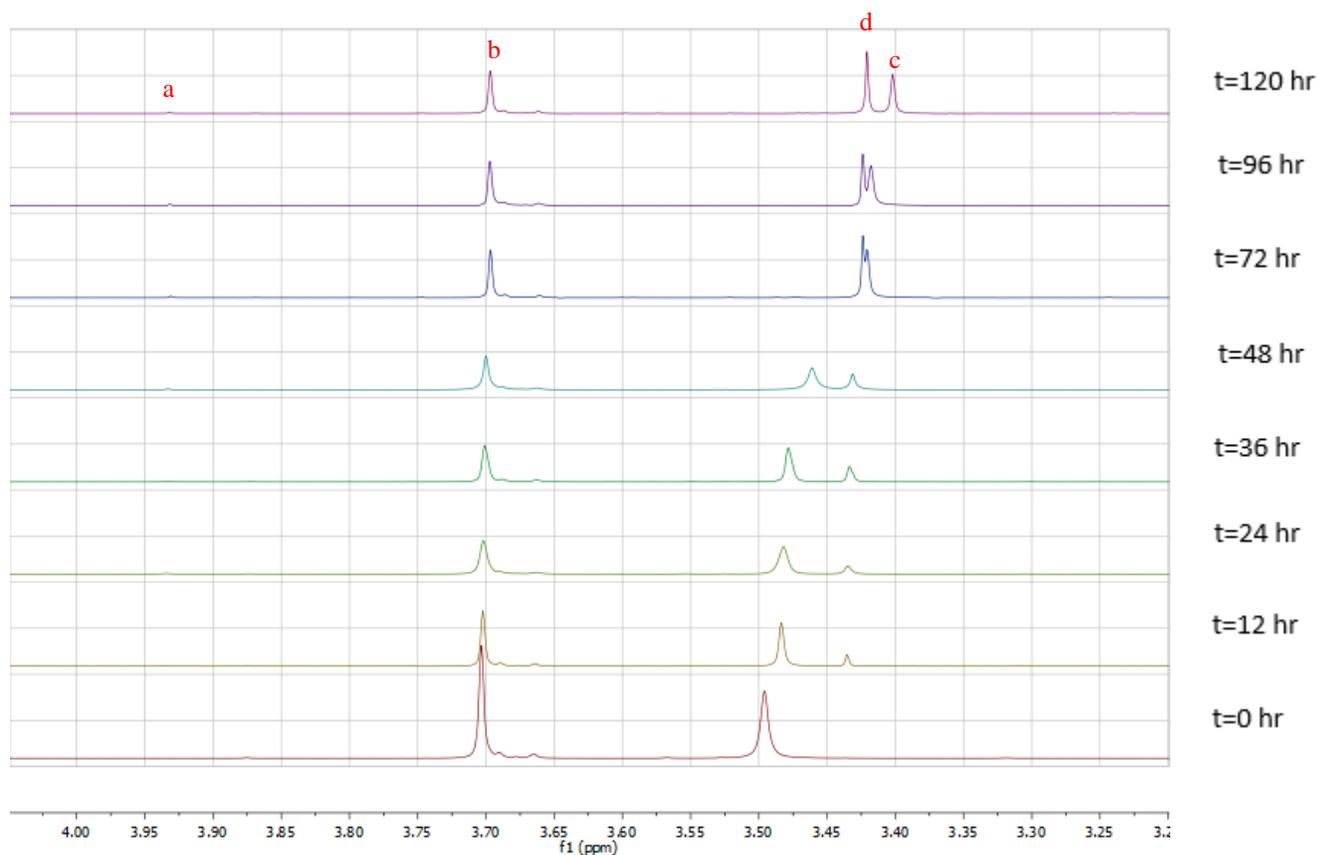
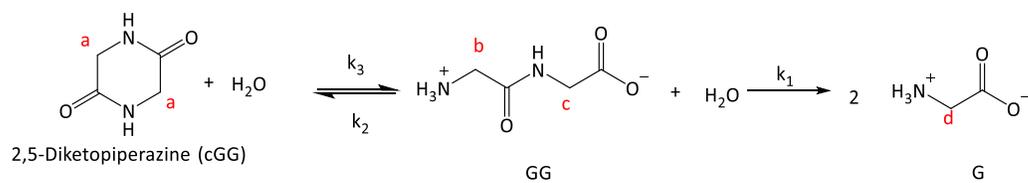
pH=5 GG analysis	NMR integrations				actual amount ( $\mu\text{mol}$ )		
time(h)	KHP	G	GG	cGG	G ( $\mu\text{mol}$ )	GG ( $\mu\text{mol}$ )	cGG ( $\mu\text{mol}$ )
0	1	0	2.62	0	0	39.3	0
0	1	0	2.62	0	0	39.3	0
0	1	0	2.64	0	0	39.6	0
12	1	0.01	2.5	0.02	0.3	37.5	0.3
12	1	0.01	2.47	0.03	0.3	37.05	0.45
12	1	0.01	2.47	0.02	0.3	37.05	0.3
24	1	0.01	2.46	0.04	0.3	36.9	0.6
24	1	0.01	2.45	0.03	0.3	36.75	0.45
24	1	0.01	2.48	0.03	0.3	37.2	0.45
36	1	0.02	2.43	0.07	0.6	36.45	1.05
36	1	0.02	2.42	0.06	0.6	36.3	0.9
36	1	0.03	2.43	0.08	0.9	36.45	1.2
48	1	0.03	2.36	0.09	0.9	35.4	1.35
48	1	0.03	2.36	0.07	0.9	35.4	1.05
48	1	0.04	2.37	0.1	1.2	35.55	1.5
72	1	0.04	2.25	0.13	1.2	33.75	1.95
72	1	0.06	2.27	0.17	1.8	34.05	2.55
72	1	0.05	2.31	0.15	1.5	34.65	2.25
96	1	0.07	2.2	0.19	2.1	33	2.85
96	1	0.07	2.21	0.21	2.1	33.15	3.15
96	1	0.06	2.27	0.17	1.8	34.05	2.55
120	1	0.06	2.25	0.18	1.8	33.75	2.7
120	1	0.07	2.18	0.19	2.1	32.7	2.85
120	1	0.08	2.16	0.23	2.4	32.4	3.45



**Figure S3.** Stacked <sup>1</sup>H NMR spectra for GG decomposition at 95°C under pH 7. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.3ppm to 4.2ppm.

**Table S3.** GG degradation reactants and products amounts at varying reaction time under pH 7 at 95°C.

pH=7 GG analysis	NMR integrations				actual amount		
time(h)	KHP	G	GG	cGG	G(μmol)	GG(μmol)	cGG (μmol)
0	1	0	2.34	0	0	35.1	0
0	1	0	2.32	0	0	34.8	0
0	1	0	2.34	0	0	35.1	0
16	1	0.01	2.13	0.04	0.3	31.95	0.6
16	1	0.01	2.05	0.04	0.3	30.75	0.6
16	1	0.01	2.17	0.07	0.3	32.55	1.05
24	1	0.01	2.18	0.07	0.3	32.7	1.05
24	1	0.02	2.11	0.09	0.6	31.65	1.35
24	1	0.01	2.21	0.06	0.3	33.15	0.9
40	1	0.02	2.19	0.11	0.6	32.85	1.65
40	1	0.02	2.03	0.12	0.6	30.45	1.8
40	1	0.04	2	0.19	1.2	30	2.85
48	1	0.05	1.95	0.24	1.5	29.25	3.6
48	1	0.03	1.97	0.17	0.9	29.55	2.55
48	1	0.02	1.99	0.14	0.6	29.85	2.1
72	1	0.05	1.89	0.24	1.5	28.35	3.6
72	1	0.05	1.9	0.29	1.5	28.5	4.35
72	1	0.09	1.65	0.42	1.5	27.3	3.75
96	1	0.1	1.62	0.41	3	24.3	6.15
96	1	0.06	1.84	0.31	1.8	27.6	4.65
96	1	0.05	1.82	0.25	1.5	27.3	3.75
120	1	0.12	1.61	0.51	3.6	24.15	7.65
120	1	0.06	1.8	0.32	1.8	27	4.8
120	1	0.07	1.74	0.34	2.1	26.1	5.1

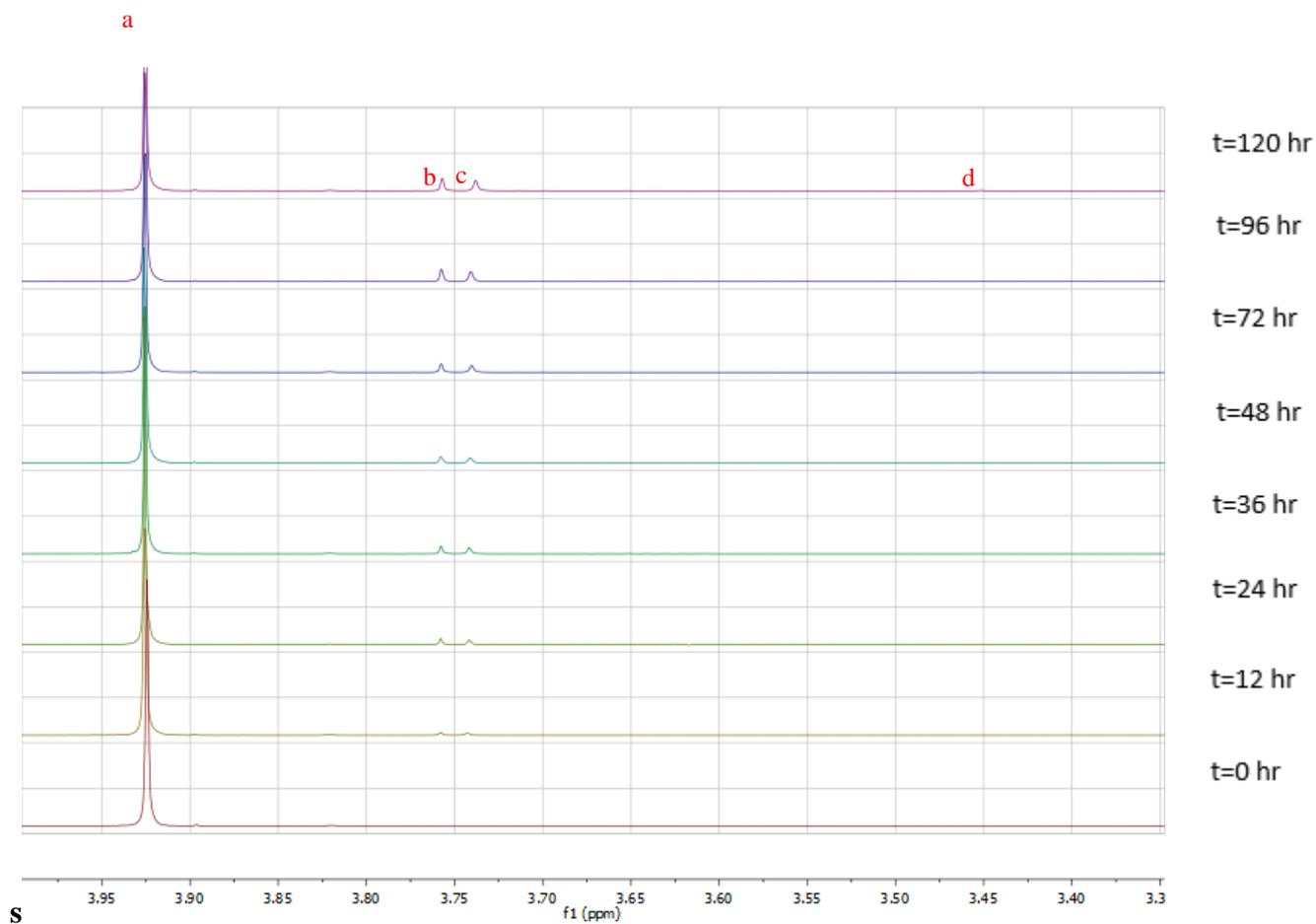
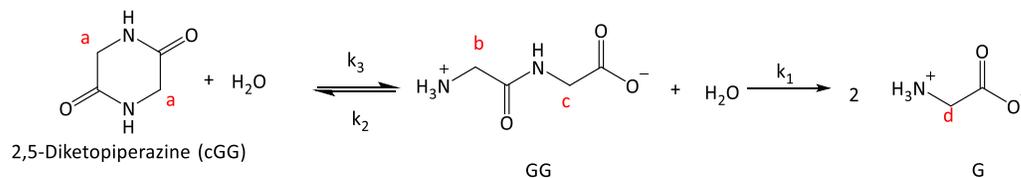


**Figure S4.** Stacked <sup>1</sup>H NMR spectra for GG decomposition at 95°C under pH 10. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.2ppm to 4.1ppm.

The chemical shift over time for peak assignment is due to pH changes.

**Table S4.** GG degradation reactants and products amounts at varying reaction time under pH 10 at 95°C

pH=10 GG analysis	NMR integrations				actual amount		
time(h)	KHP	G	GG	cGG	G(μmol)	GG(μmol)	cGG (μmol)
0	1	0	2.06	0	0	30.9	0
0	1	0	2	0	0	30	0
0	1	0	2.1	0	0	31.5	0
16	1	0.04	2.02	0.01	1.2	30.3	0.15
16	1	0.05	1.96	0.01	1.5	29.4	0.15
16	1	0.05	1.98	0.01	1.5	29.7	0.15
24	1	0.1	1.92	0.01	3	28.8	0.15
24	1	0.09	1.82	0.01	2.7	27.3	0.15
24	1	0.08	1.92	0.01	2.4	28.8	0.15
40	1	0.17	1.74	0.01	5.1	26.1	0.15
40	1	0.17	1.76	0.01	5.1	26.4	0.15
40	1	0.17	1.72	0.01	5.1	25.8	0.15
48	1	0.23	1.74	0.01	6.9	26.1	0.15
48	1	0.24	1.68	0.02	7.2	25.2	0.3
48	1	0.24	1.66	0.02	7.2	24.9	0.3
72	1	0.47	1.46	0.01	14.1	21.9	0.15
72	1	0.46	1.4	0.01	13.8	21	0.15
72	1	0.46	1.46	0.01	13.8	21.6	0.15
96	1	0.51	1.36	0.01	15.3	20.4	0.15
96	1	0.5	1.42	0.01	15	21.3	0.15
96	1	0.51	1.44	0.01	15.3	21.6	0.15
120	1	0.58	1.35	0.01	17.4	20.25	0.15
120	1	0.58	1.34	0.01	17.4	20.1	0.15
120	1	0.6	1.34	0.01	18	20.1	0.15

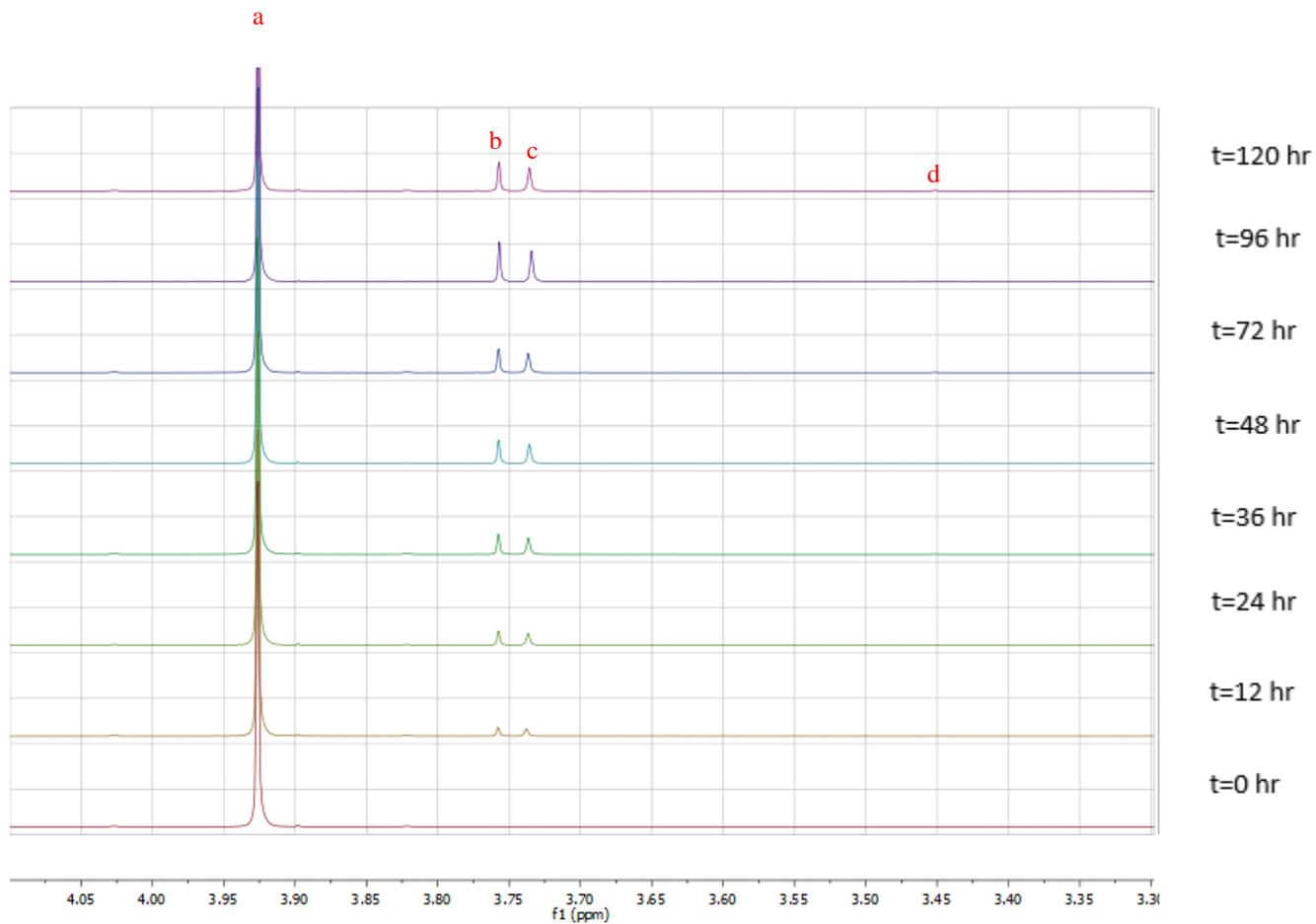
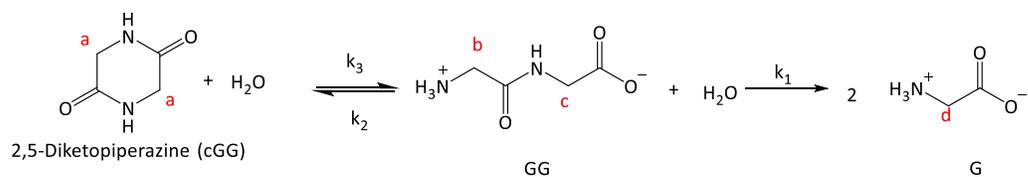


**Figure S5.** Stacked <sup>1</sup>H NMR spectra for cGG decomposition at 95°C under pH 3. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.35ppm to 4ppm.



**Table S5.** cGG degradation reactants and products amounts at varying reaction time under pH 3 at 95°C.

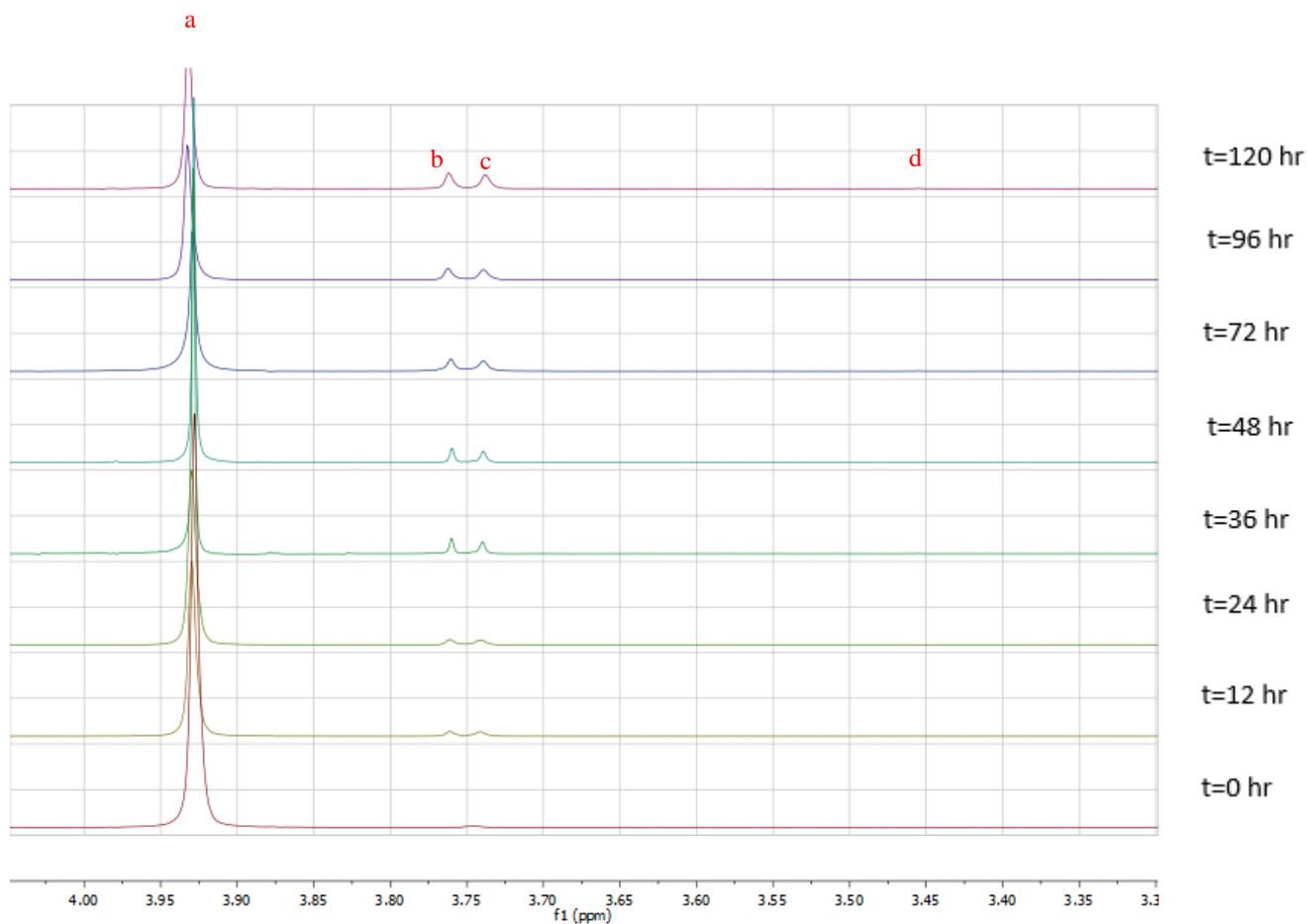
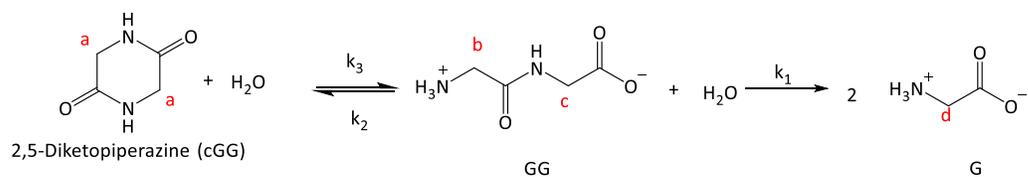
pH=3 cGG analysis	NMR integrations				actual amount			
	time(h)	KHP	G	GG	cGG	G( $\mu$ mol)	GG( $\mu$ mol)	cGG ( $\mu$ mol)
	0	1	0	0	1.07	0	0	16.05
	0	1	0	0	1.06	0	0	15.9
	0	1	0	0	1.05	0	0	15.75
	12	1	0	0.03	1.01	0	0.45	15.15
	12	1	0	0.03	1	0	0.45	15
	12	1	0	0.04	0.96	0	0.6	14.4
	24	1	0	0.05	0.96	0	0.75	14.4
	24	1	0	0.05	0.99	0	0.75	14.85
	24	1	0	0.04	0.98	0	0.6	14.7
	36	1	0	0.07	0.89	0	1.05	13.35
	36	1	0	0.08	0.91	0	1.2	13.65
	36	1	0	0.08	0.92	0	1.2	13.8
	48	1	0	0.07	0.93	0	1.05	13.95
	48	1	0	0.06	0.94	0	0.9	14.1
	48	1	0	0.07	0.92	0	1.05	13.8
	72	1	0	0.08	0.93	0	1.2	13.95
	72	1	0	0.1	0.89	0	1.5	13.35
	72	1	0	0.11	0.87	0	1.65	13.05
	96	1	0	0.12	0.9	0	1.8	13.5
	96	1	0	0.12	0.86	0	1.8	12.9
	96	1	0	0.09	0.9	0	1.35	13.5
	120	1	0	0.12	0.87	0	1.8	13.05
	120	1	0	0.11	0.89	0	1.65	13.35
	120	1	0	0.12	0.88	0	1.8	13.2



**Figure S6.** Stacked <sup>1</sup>H NMR spectra for cGG decomposition at 95°C under pH 5. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.3ppm to 4.1ppm.

**Table S6.** cGG degradation reactants and products amounts at varying reaction time under pH 5 at 95°C.

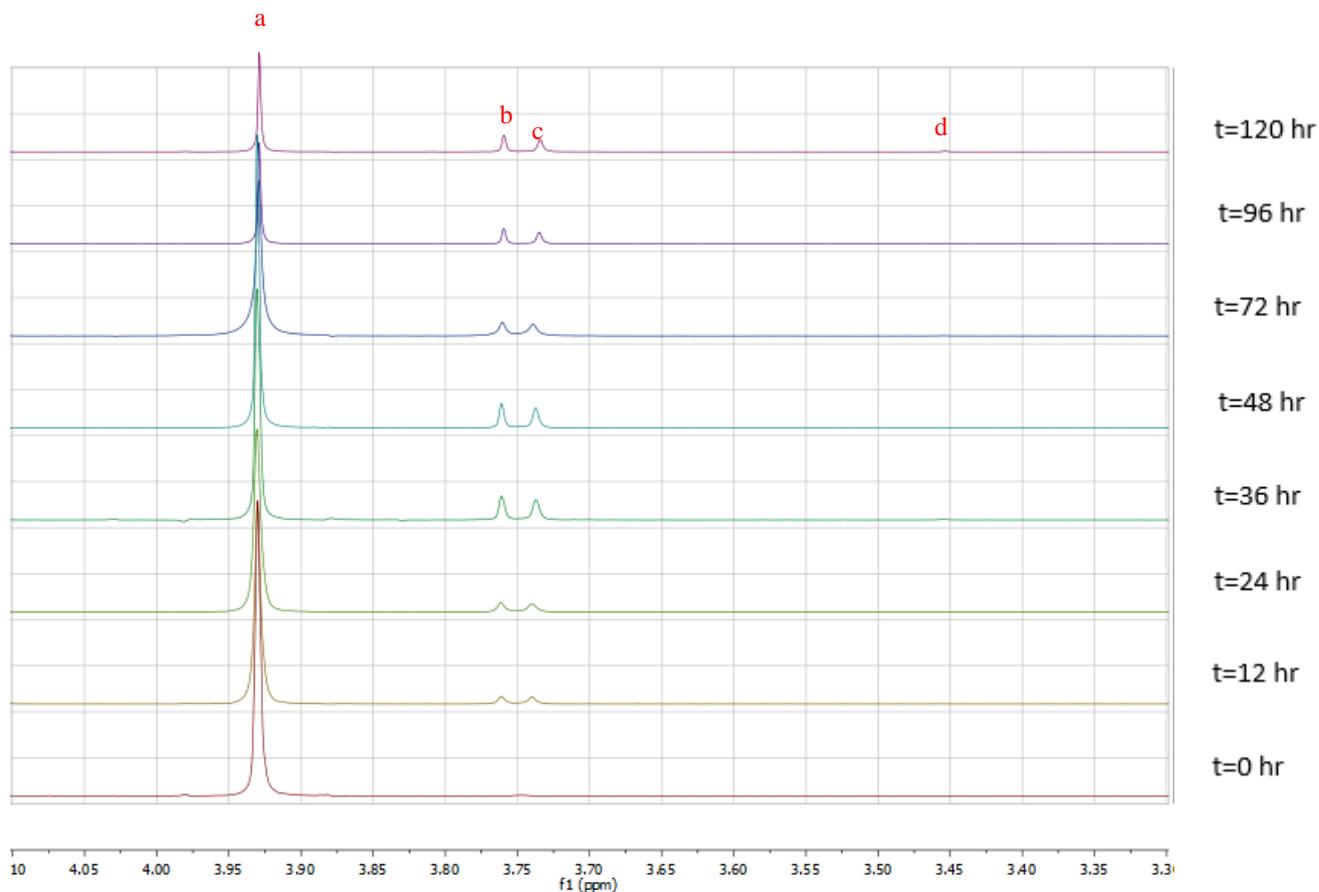
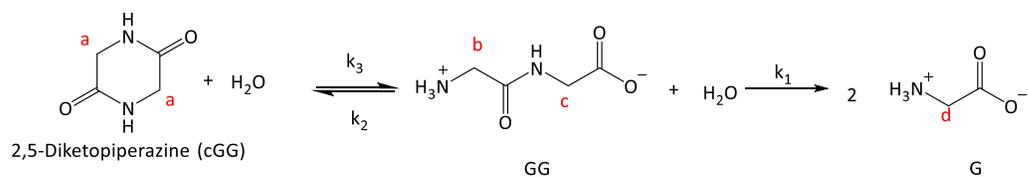
pH=5 cGG analysis time(h)	NMR integrations				actual amount		
	KHP	G	GG	cGG	G( $\mu$ mol)	GG( $\mu$ mol)	cGG( $\mu$ mol)
0	1	0	0	1.04	0	0	15.6
0	1	0	0	1.04	0	0	15.6
0	1	0	0	1.04	0	0	15.6
12	1	0	0.06	0.94	0	0.9	14.1
12	1	0	0.06	0.98	0	0.9	14.7
12	1	0	0.06	0.95	0	0.9	14.25
24	1	0	0.09	0.91	0	1.35	13.65
24	1	0	0.07	0.94	0	1.05	14.1
24	1	0	0.09	0.91	0	1.35	13.65
36	1	0	0.12	0.89	0	1.8	13.35
36	1	0	0.1	0.91	0	1.5	13.65
36	1	0	0.09	0.91	0	1.35	13.65
48	1	0	0.14	0.83	0	2.1	12.45
48	1	0	0.16	0.84	0	2.4	12.6
48	1	0	0.15	0.84	0	2.25	12.6
72	1	0	0.15	0.83	0	2.25	12.45
72	1	0	0.19	0.79	0	2.85	11.85
72	1	0	0.2	0.79	0	3	11.85
96	1	0	0.22	0.76	0	3.3	11.4
96	1	0	0.18	0.81	0	2.7	12.15
96	1	0	0.24	0.77	0	3.6	11.55
120	1	0	0.18	0.78	0	2.7	11.7
120	1	0	0.2	0.79	0	3	11.85
120	1	0	0.23	0.76	0	3.45	11.4



**Figure S7.** Stacked <sup>1</sup>H NMR spectra for cGG decomposition at 95°C under pH 7. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.3ppm to 4.05ppm.

**Table S7.** cGG degradation reactants and products amounts at varying reaction time under pH 7 at 95°C.

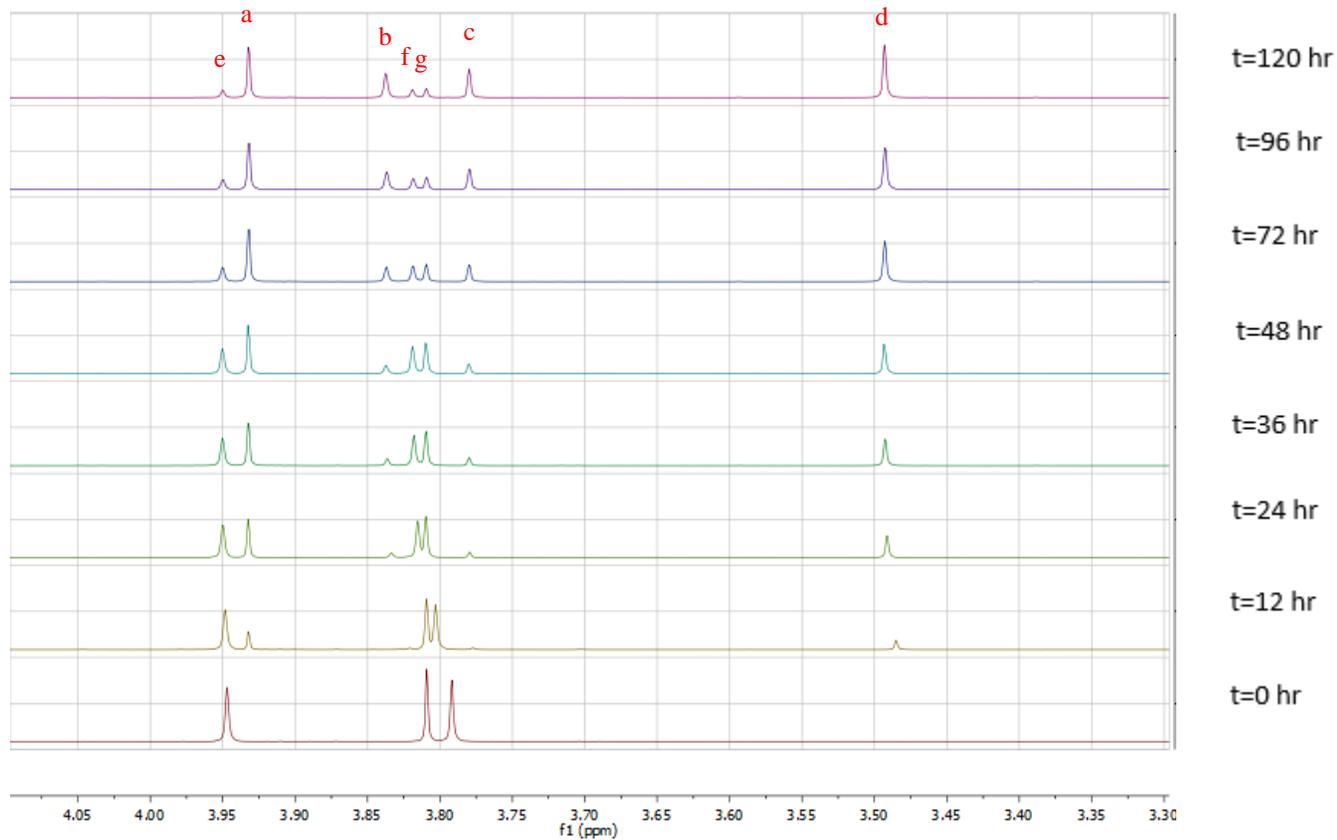
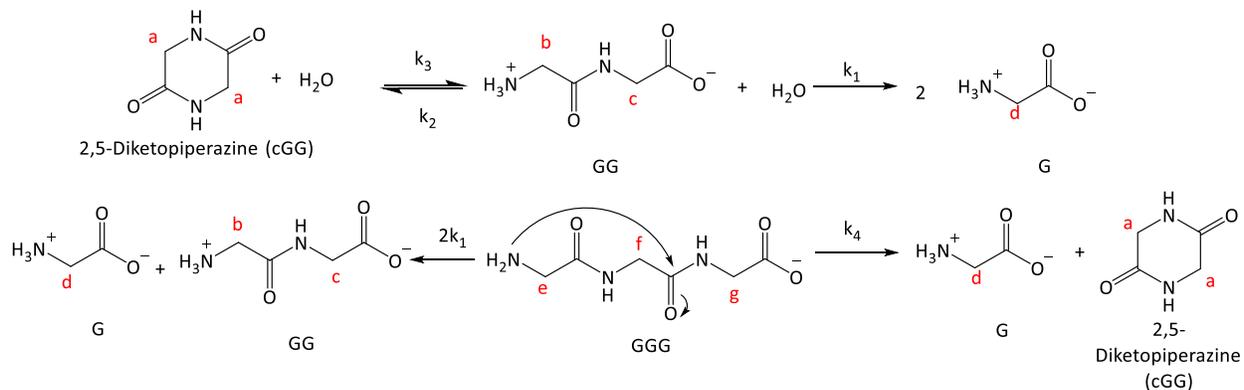
pH=7 cGG analysis	NMR integrations				actual amount			
	time(h)	KHP	G	GG	cGG	G( $\mu$ mol)	GG( $\mu$ mol)	cGG ( $\mu$ mol)
	0	1	0.01	0.01	1.3	0.3	0.15	19.5
	0	1	0.01	0.01	1.3	0.3	0.15	19.5
	0	1	0.01	0.01	1.26	0.3	0.15	18.9
	16	1	0.01	0.07	1.19	0.3	1.05	17.85
	16	1	0.01	0.07	1.12	0.3	1.05	16.8
	16	1	0.01	0.05	1.15	0.3	0.75	17.25
	24	1	0.01	0.08	1.14	0.3	1.2	17.1
	24	1	0.02	0.07	1.12	0.6	1.05	16.8
	24	1	0.01	0.07	1.17	0.3	1.05	17.55
	40	1	0	0.12	1.08	0	1.8	16.2
	40	1	0.01	0.11	1.1	0.3	1.65	16.5
	40	1	0.01	0.11	1.07	0.3	1.65	16.05
	48	1	0.01	0.11	1.07	0.3	1.65	16.05
	48	1	0.01	0.11	1.07	0.3	1.65	16.05
	48	1	0.01	0.13	1.07	0.3	1.95	16.05
	72	1	0	0.19	0.98	0	2.85	14.7
	72	1	0.01	0.2	1.25	0.3	3	18.75
	72	1	0	0.17	1	0	2.85	15.15
	96	1	0.01	0.2	1.06	0.3	3	15.9
	96	1	0.01	0.2	1.21	0.3	3	18.15
	96	1	0.01	0.19	1.01	0.3	2.85	15.15
	120	1	0	0.22	1.02	0	3.3	15.3
	120	1	0.01	0.28	1.16	0.3	4.2	17.4
	120	1	0.01	0.26	1.15	0.3	3.9	17.25



**Figure S8.** Stacked <sup>1</sup>H NMR spectra for cGG decomposition at 95°C under pH 10. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.3ppm to 4.1ppm.

**Table S8.** cGG degradation reactants and products amounts at varying reaction time under pH 10 at 95°C.

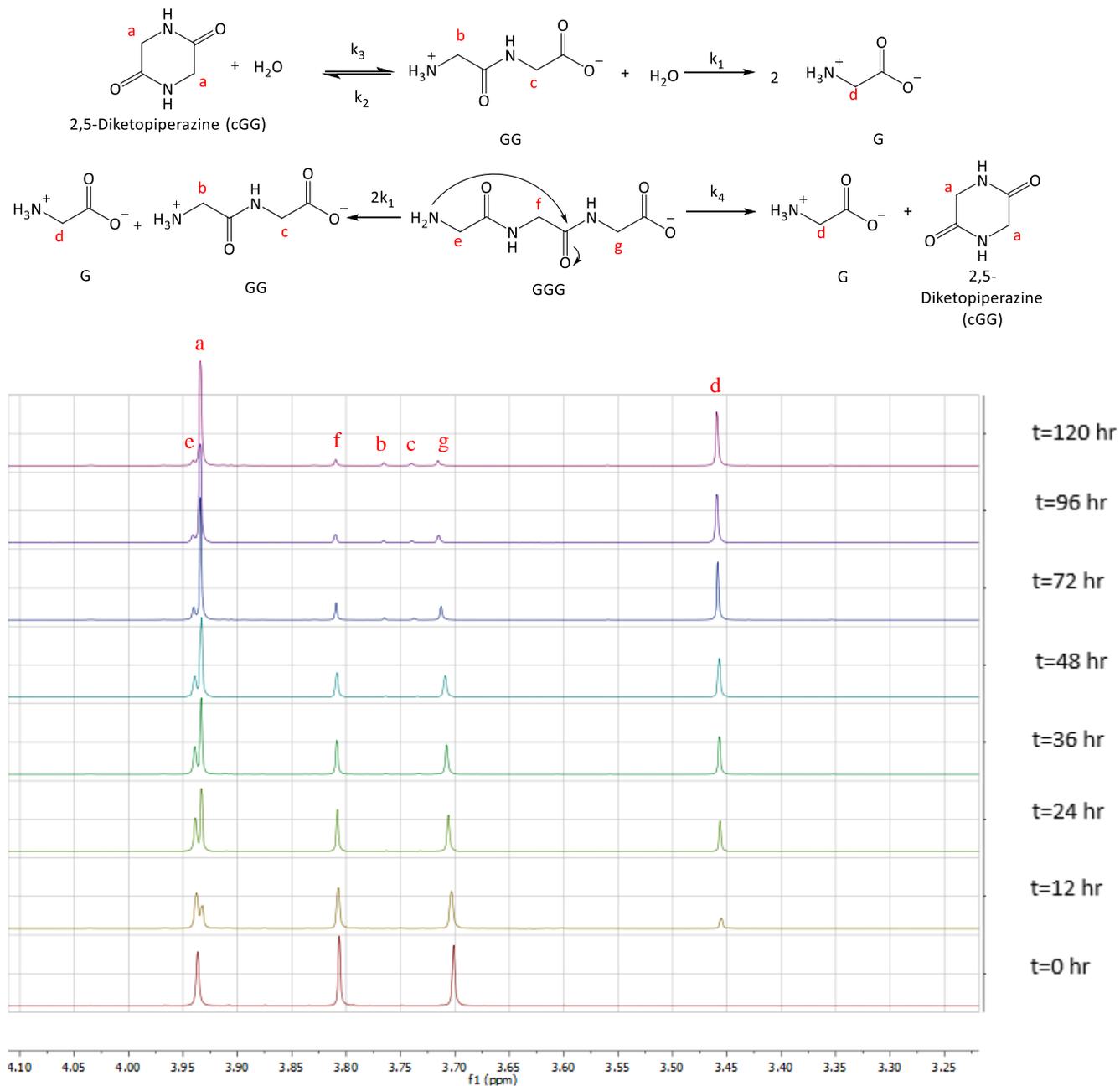
pH=10 cGG analysis	NMR integrations				actual amount			
	time(h)	KHP	G	GG	cGG	G( $\mu$ mol)	GG( $\mu$ mol)	cGG ( $\mu$ mol)
	0	1	0	0	1.3	0	0	19.5
	0	1	0	0	1.25	0	0	18.75
	0	1	0	0	1.25	0	0	18.75
	16	1	0	0.09	1.1	0	1.35	16.5
	16	1	0	0.09	1.07	0	1.35	16.05
	16	1	0	0.08	1.12	0	1.2	16.8
	24	1	0	0.11	1.03	0	1.65	15.45
	24	1	0	0.11	1.09	0	1.65	16.35
	24	1	0	0.11	1.08	0	1.65	16.2
	40	1	0	0.22	0.98	0	3.3	14.7
	40	1	0	0.21	0.96	0	3.15	14.4
	40	1	0	0.19	0.95	0	2.85	14.25
	48	1	0	0.2	1.02	0	3	15.3
	48	1	0	0.19	0.92	0	2.85	13.8
	48	1	0	0.19	0.94	0	2.85	14.1
	72	1	0	0.19	0.92	0	2.85	13.8
	72	1	0	0.17	0.99	0	2.55	14.85
	72	1	0	0.17	0.99	0	2.55	14.85
	96	1	0.01	0.3	0.8	0.3	4.5	12
	96	1	0.01	0.29	0.84	0.3	4.35	12.6
	96	1	0.01	0.29	0.87	0.3	4.35	13.05
	120	1	0.01	0.33	0.76	0.3	4.95	11.4
	120	1	0.01	0.38	0.83	0.3	5.7	12.45
	120	1	0.01	0.35	0.76	0.3	5.25	11.4



**Figure S9.** Stacked <sup>1</sup>H NMR spectra for GGG decomposition at 95°C under pH 3. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.3ppm to 4.1ppm.

**Table S9.** GGG degradation reactants and products amounts at varying reaction time under pH 3 at 95°C.

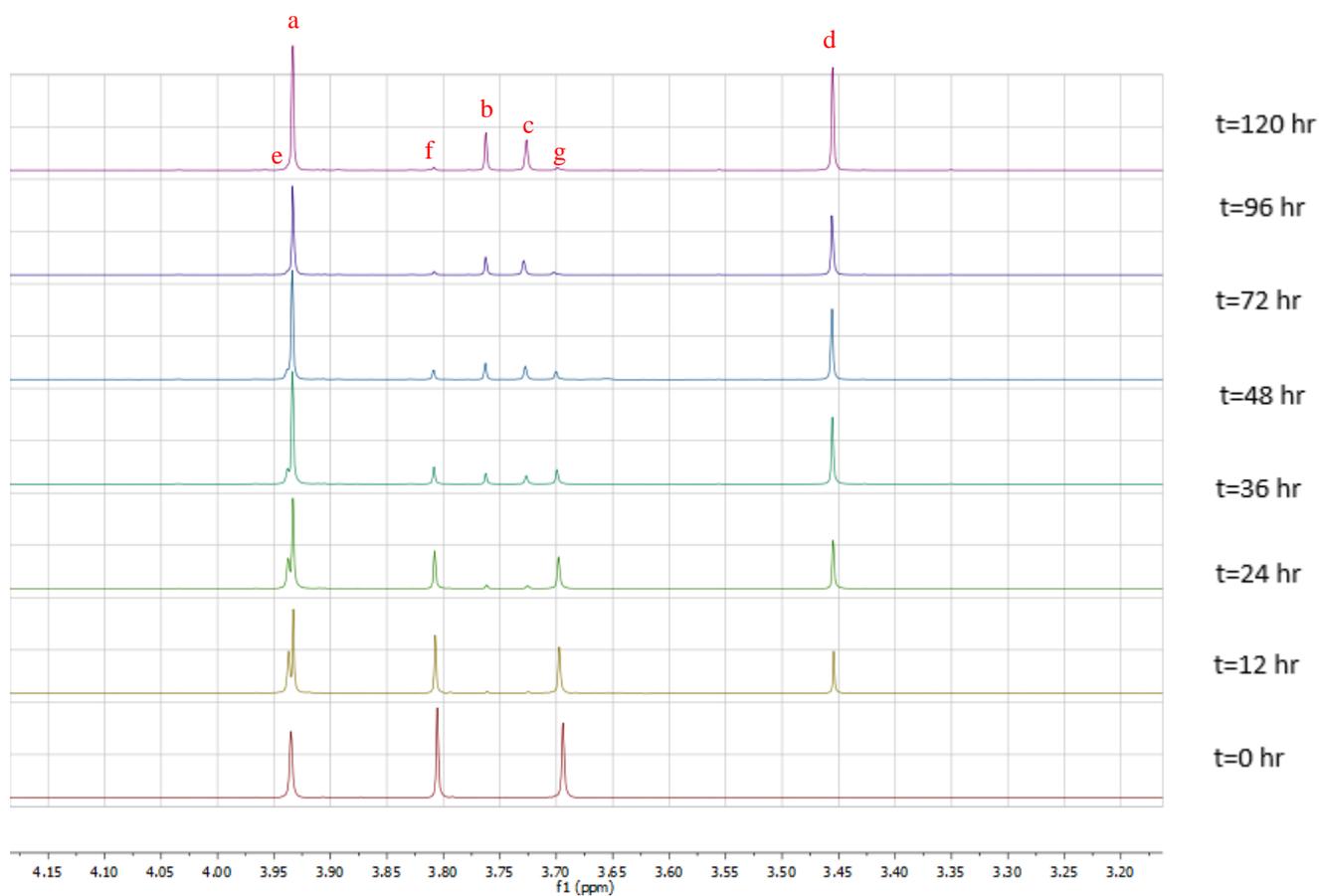
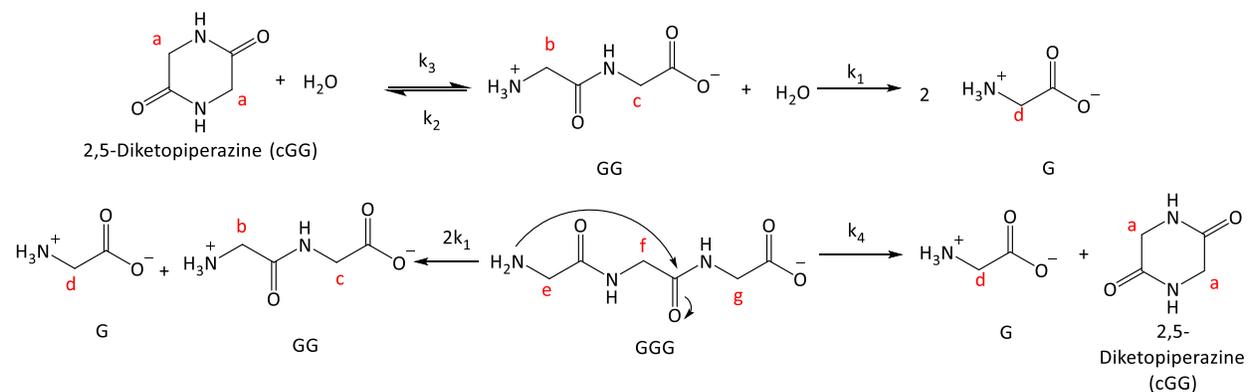
pH=3 GGG analysis	NMR peak integrations					Actual amount ( $\mu\text{mol}$ )				
	time(h)	KHP	G	GG	cGG	GGG	G ( $\mu\text{mol}$ )	GG ( $\mu\text{mol}$ )	cGG ( $\mu\text{mol}$ )	GGG ( $\mu\text{mol}$ )
	0	1	0	0.02	0	3.18	0	0.3	0	31.8
	0	1	0	0.02	0	3.24	0	0.3	0	32.4
	0	1	0	0.02	0	3.3	0	0.3	0	33
	12	1	0.15	0.08	0.28	2.52	4.5	1.2	4.2	25.2
	12	1	0.15	0.06	0.23	2.61	4.5	0.9	3.45	26.1
	12	1	0.14	0.08	0.23	2.76	4.2	1.2	3.45	27.6
	24	1	0.37	0.18	0.57	1.98	11.1	2.7	8.55	19.8
	24	1	0.35	0.16	0.53	2.07	10.5	2.4	7.95	20.7
	24	1	0.34	0.18	0.53	2.19	10.2	2.7	7.95	21.9
	36	1	0.47	0.28	0.67	1.77	14.1	4.2	10.05	17.7
	36	1	0.44	0.24	0.64	1.86	13.2	3.6	9.6	18.6
	36	1	0.43	0.24	0.63	1.92	12.9	3.6	9.45	19.2
	48	1	0.53	0.34	0.71	1.65	15.9	5.1	10.65	16.5
	48	1	0.52	0.32	0.71	1.71	15.6	4.8	10.65	17.1
	48	1	0.46	0.3	0.66	1.74	13.8	4.5	9.9	17.4
	72	1	0.73	0.58	0.84	1.02	21.9	8.7	12.6	10.2
	72	1	0.69	0.54	0.82	1.17	20.7	8.1	12.3	11.7
	72	1	0.7	0.54	0.84	1.14	21	8.1	12.6	11.4
	96	1	0.8	0.74	0.79	0.66	24	11.1	11.85	6.6
	96	1	0.82	0.74	0.85	0.78	24.6	11.1	12.75	7.8
	96	1	0.78	0.68	0.82	0.87	23.4	10.2	12.3	8.7
	120	1	0.94	0.96	0.78	0.45	28.2	14.4	11.7	4.5
	120	1	0.98	0.96	0.88	0.33	29.4	14.4	13.2	3.3
	120	1	0.95	0.88	0.9	0.48	28.5	13.2	13.5	4.8



**Figure S10.** Stacked <sup>1</sup>H NMR spectra for GGG decomposition at 95°C under pH 5. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.2ppm to 4.1ppm.

**Table S10.** GGG degradation reactants and products amounts at varying reaction time under pH 5 at 95°C.

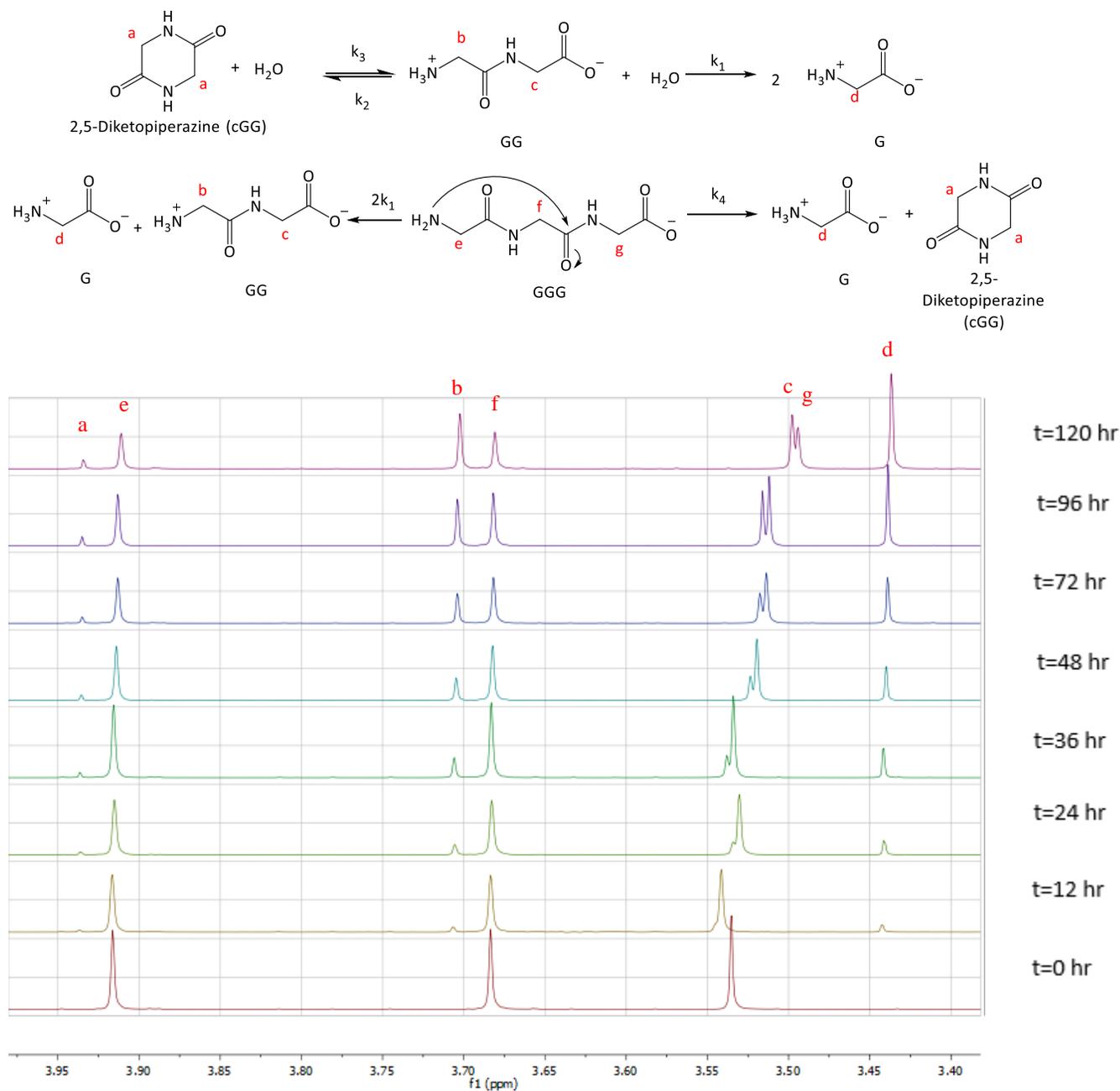
pH=5 GGG analysis	NMR peak integrations					Actual amount ( $\mu\text{mol}$ )				
	time(h)	KHP	G	GG	cGG	GGG	G ( $\mu\text{mol}$ )	GG ( $\mu\text{mol}$ )	cGG ( $\mu\text{mol}$ )	GGG ( $\mu\text{mol}$ )
	0	1	0	0	0	3.81	0	0	0	38.1
	0	1	0	0	0	3.84	0	0	0	38.4
	0	1	0	0	0	3.84	0	0	0	38.4
	12	1	0.24	0.02	0.44	3.03	7.2	0.3	6.6	30.3
	12	1	0.22	0.02	0.43	3.09	6.6	0.3	6.45	30.9
	12	1	0.21	0.02	0.4	3.12	6.3	0.3	6	31.2
	24	1	0.51	0.03	0.98	2.16	15.3	0.45	14.7	21.6
	24	1	0.47	0.03	0.9	2.37	14.1	0.45	13.5	23.7
	24	1	0.42	0.03	0.81	2.46	12.6	0.45	12.15	24.6
	36	1	0.63	0.04	1.23	1.77	18.9	0.6	18.45	17.7
	36	1	0.6	0.04	1.18	1.92	18	0.6	17.7	19.2
	36	1	0.54	0.04	1.2	1.62	16.2	0.6	18	16.2
	48	1	0.76	0.05	1.47	1.47	22.8	0.75	22.05	14.7
	48	1	0.65	0.04	1.25	1.77	19.5	0.6	18.75	17.7
	48	1	0.57	0.04	1.11	1.95	17.1	0.6	16.65	19.5
	72	1	0.96	0.07	1.81	0.84	28.8	1.05	27.15	8.4
	72	1	0.9	0.06	1.67	0.96	27	0.9	25.05	9.6
	72	1	0.82	0.06	1.5	1.17	24.6	0.9	22.5	11.7
	96	1	1.05	0.09	1.99	0.54	31.5	1.35	29.85	5.4
	96	1	1	0.08	1.87	0.66	30	1.2	28.05	6.6
	96	1	0.95	0.08	1.79	0.84	28.5	1.2	26.85	8.4
	120	1	1.08	0.11	1.95	0.36	32.4	1.65	29.25	3.6
	120	1	1.13	0.12	2.03	0.3	33.9	1.8	30.45	3
	120	1	1.07	0.13	1.94	0.39	32.1	1.95	29.1	3.9



**Figure S11.** Stacked <sup>1</sup>H NMR spectra for GGG decomposition at 95°C under pH 7. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.2ppm to 4.20ppm.

**Table S11.** GGG degradation reactants and products amounts at varying reaction time under pH 7 at 95°C.

pH=7 GGG analysis	NMR peak integrations					Actual amount ( $\mu\text{mol}$ )				
	time(h)	KHP	G	GG	cGG	GGG	G ( $\mu\text{mol}$ )	GG ( $\mu\text{mol}$ )	cGG ( $\mu\text{mol}$ )	GGG ( $\mu\text{mol}$ )
	0	1	0	0.02	0.01	3.51	0	0.32	0.16	37.44
	0	1	0	0.02	0.01	3.54	0	0.32	0.16	37.76
	0	1	0	0.01	0.02	3.42	0	0.16	0.32	36.48
	12	1	0.41	0.06	0.78	1.98	13.12	0.96	12.48	21.12
	12	1	0.4	0.05	0.76	2.07	12.8	0.8	12.16	22.08
	12	1	0.37	0.05	0.69	2.07	11.84	0.8	11.04	22.08
	24	1	0.63	0.11	1.15	1.59	20.16	1.76	18.4	16.96
	24	1	0.53	0.09	0.96	1.59	16.96	1.44	15.36	16.96
	24	1	0.58	0.1	1.05	1.62	18.56	1.6	16.8	17.28
	36	1	0.81	0.19	1.42	1.05	25.92	3.04	22.72	11.2
	36	1	0.72	0.16	1.26	1.05	23.04	2.56	20.16	11.2
	36	1	0.67	0.14	1.2	1.23	21.44	2.24	19.2	13.12
	48	1	0.9	0.29	1.47	0.72	28.8	4.64	23.52	7.68
	48	1	0.84	0.25	1.41	0.78	26.88	4	22.56	8.32
	48	1	0.75	0.2	1.3	0.93	24	3.2	20.8	9.92
	72	1	0.96	0.44	1.43	0.45	30.72	7.04	22.88	4.8
	72	1	0.92	0.4	1.4	0.48	29.44	6.4	22.4	5.12
	72	1	0.88	0.33	1.39	0.57	28.16	5.28	22.24	6.08
	96	1	1.05	0.54	1.6	0.36	31.5	8.1	24	3.6
	96	1	1.04	0.6	1.41	0.27	33.28	9.6	22.56	2.88
	96	1	1.01	0.55	1.39	0.33	32.32	8.8	22.24	3.52
	120	1	1.14	0.83	1.33	0.15	34.2	12.45	19.95	1.6
	120	1	1.17	0.83	1.4	0.18	35.1	12.45	21	1.92
	120	1	1.1	0.76	1.36	0.18	33	11.4	20.4	1.92



**Figure S12.** Stacked <sup>1</sup>H NMR spectra for GGG decomposition at 95°C under pH 10. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.35ppm to 4.0ppm.

Peak shifts for GGG and GG are observed comparing day0 and day5 spectra due to pH variation over reaction time.

**Table S12.** GGG degradation reactants and products amounts at varying reaction time under pH 10 at 95°C.

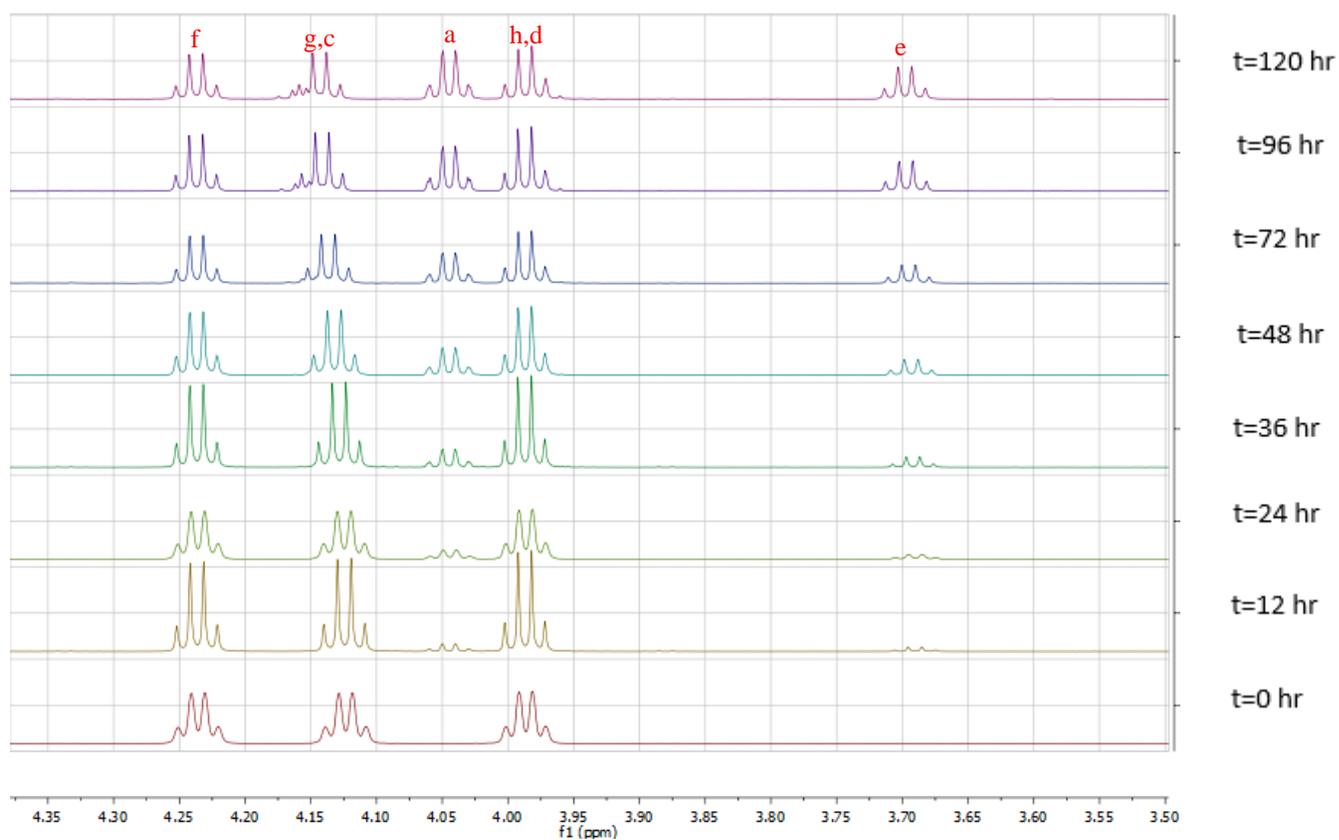
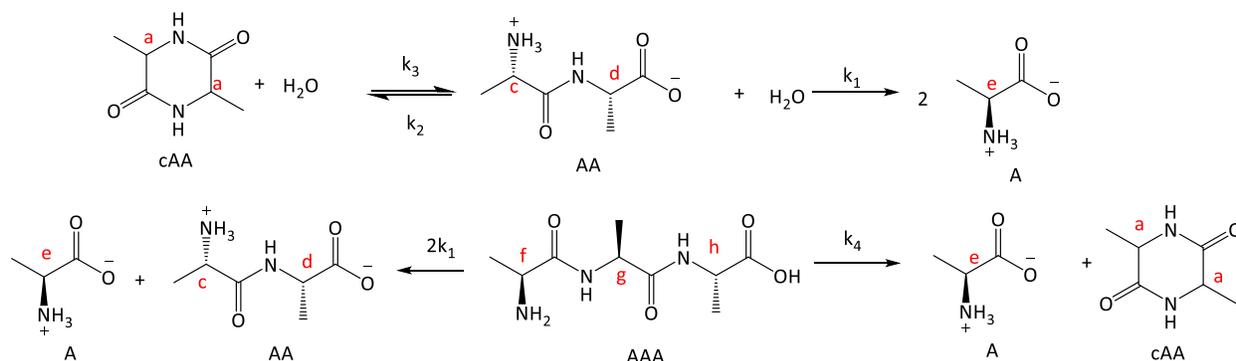
pH=10 GGG analysis	NMR peak integrations					Actual amount ( $\mu\text{mol}$ )				
	time(h)	KHP	G	GG	cGG	GGG	G ( $\mu\text{mol}$ )	GG ( $\mu\text{mol}$ )	cGG ( $\mu\text{mol}$ )	GGG ( $\mu\text{mol}$ )
	0	1	0.01	0.02	0	3.39	0.3	0.3	0	33.9
	0	1	0.01	0.02	0.01	3.39	0.3	0.3	0.15	33.9
	0	1	0.01	0.02	0.01	3.36	0.3	0.3	0.15	33.6
	12	1	0.1	0.16	0.03	2.94	3	2.4	0.45	29.4
	12	1	0.1	0.18	0.03	3.09	3	2.7	0.45	30.9
	12	1	0.09	0.16	0.03	3.03	2.7	2.4	0.45	30.3
	24	1	0.18	0.32	0.04	2.79	5.4	4.8	0.6	27.9
	24	1	0.17	0.3	0.04	2.82	5.1	4.5	0.6	28.2
	24	1	0.16	0.28	0.04	2.88	4.8	4.2	0.6	28.8
	36	1	0.24	0.42	0.05	2.58	7.2	6.3	0.75	25.8
	36	1	0.27	0.3	0.06	2.61	8.1	4.5	0.9	26.1
	36	1	0.22	0.36	0.05	2.58	6.6	5.4	0.75	25.8
	48	1	0.35	0.58	0.06	2.4	10.5	8.7	0.9	24
	48	1	0.31	0.52	0.05	2.43	9.3	7.8	0.75	24.3
	48	1	0.28	0.48	0.05	2.52	8.4	7.2	0.75	25.2
	72	1	0.5	0.76	0.07	2.07	15	11.4	1.05	20.7
	72	1	0.45	0.72	0.07	2.13	13.5	10.8	1.05	21.3
	72	1	0.42	0.66	0.06	2.22	12.6	9.9	0.9	22.2
	96	1	0.64	0.9	0.08	1.62	19.2	13.5	1.2	16.2
	96	1	0.6	0.88	0.07	1.74	18	13.2	1.05	17.4
	96	1	0.54	0.8	0.07	1.86	16.2	12	1.05	18.6
	120	1	0.87	1.12	0.08	1.35	26.1	16.8	1.2	13.5
	120	1	0.8	1.08	0.08	1.47	24	16.2	1.2	14.7
	120	1	0.74	1.02	0.07	1.62	22.2	15.3	1.05	16.2

**Table S13.** The estimated rate constants  $k_1$ - $k_4$  based on joint fitting results for GG, cGG and GGG kinetics, under pH 3, 5, 7 and 10, at 95°C. The confidence intervals follow the chi-squared method at the 95% confidence level.

	$k_{sc} \times 10^6$ (s <sup>-1</sup> )	$k_{rc} \times 10^6$ (s <sup>-1</sup> )	$2k_{ro} \times 10^6$ (s <sup>-1</sup> )	$k_{bb} \times 10^6$ (s <sup>-1</sup> )
pH=3	0.48(±0.183)	0.60(±0.205)	0.92(±0.255)	2.98(±0.458)
pH=5	0.06(±0.009)	0.27(±0.020)	0.35(±0.023)	5.03(±0.087)
pH=7	0.14(±0.004)	0.60(±0.008)	1.11(±0.012)	8.24(±0.031)
pH=10	0.84(±0.011)	0.07(±0.003)	0.96(±0.012)	0.09(±0.004)

**Table S14.** The estimated initial amount of GG, cGG and GGG based on joint fitting results, under pH 3, 5, 7 and 10, at 95°C. The confidence intervals are calculated following the chi-squared method at the 95% confidence level.

	GG <sub>0</sub> (μmol)	cGG <sub>0</sub> (μmol)	GGG <sub>0</sub> (μmol)
pH=3	30.15(±1.458)	15.98(±1.061)	31.23(±1.484)
pH=5	37.78(±0.240)	14.09(±0.146)	37.41(±0.239)
pH=7	33.76(±0.064)	19.63(±0.049)	34.17(±0.064)
pH=10	29.92(±0.068)	17.67(±0.053)	33.08(±0.072)

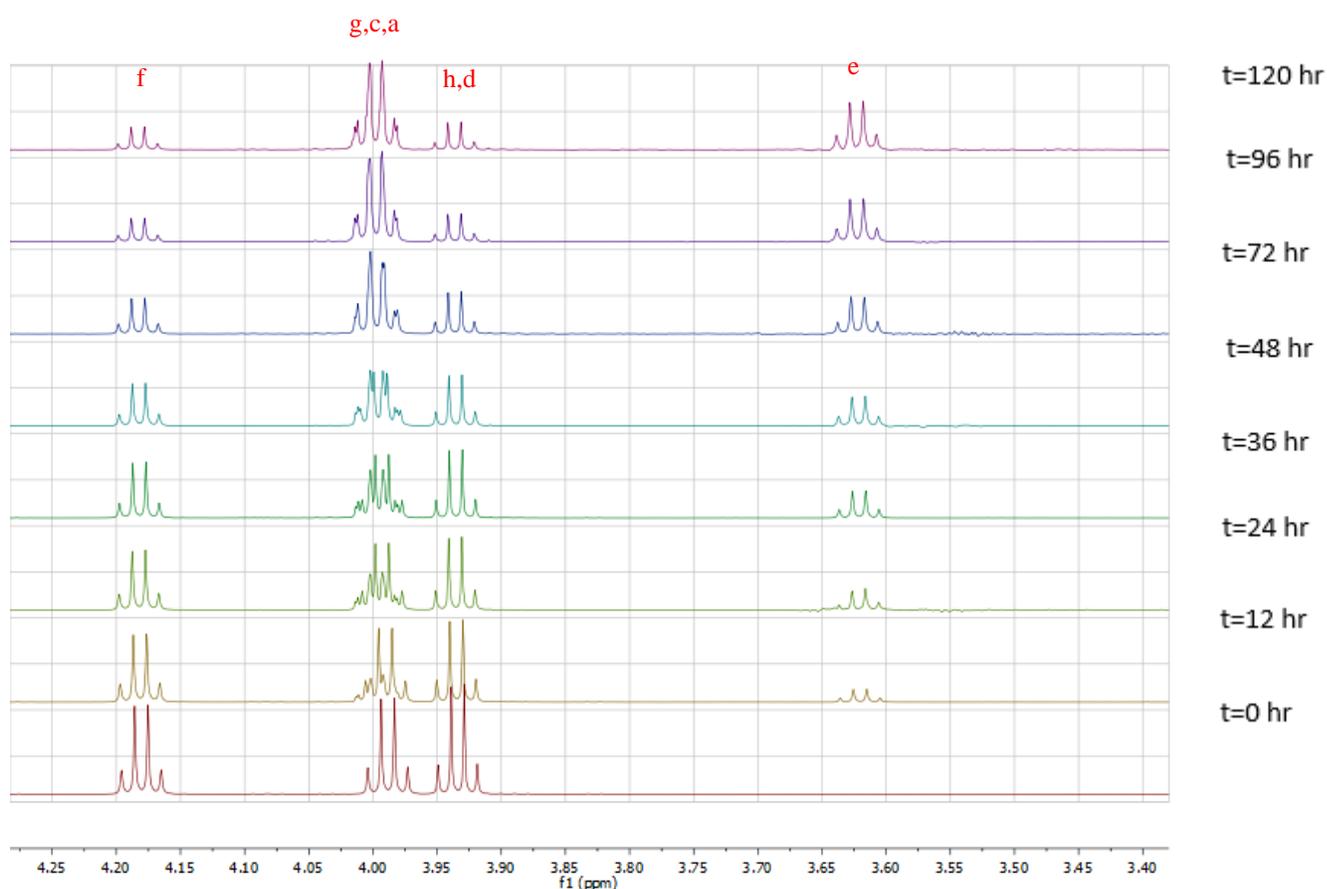
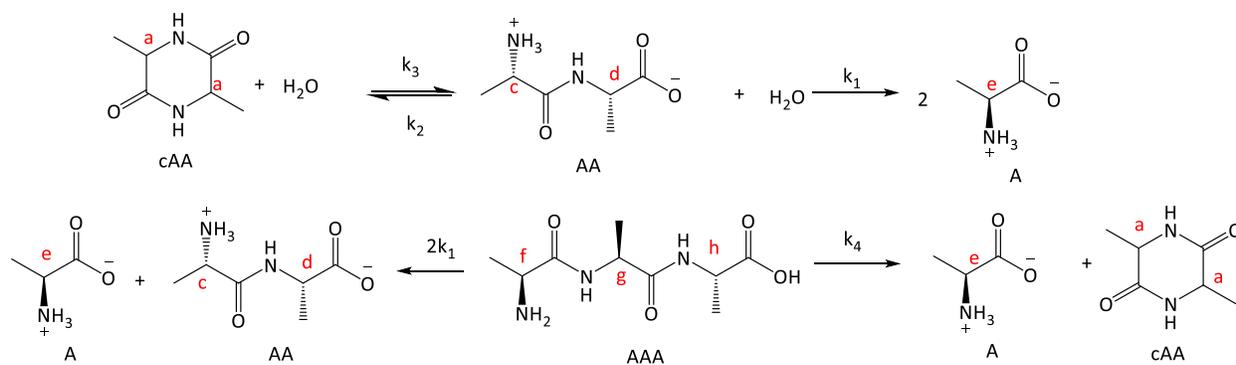


**Figure S13.** Stacked <sup>1</sup>H NMR spectra for AAA decomposition at 95°C under pH 3. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.5ppm to 4.4ppm.

For all the quantitative analysis of AAA degradation studies under different pH conditions, the calculations are based on the assignment of methylene protons on AAA.

**Table S15.** AAA degradation reactants and products amounts at varying reaction time under pH 3 at 95°C.

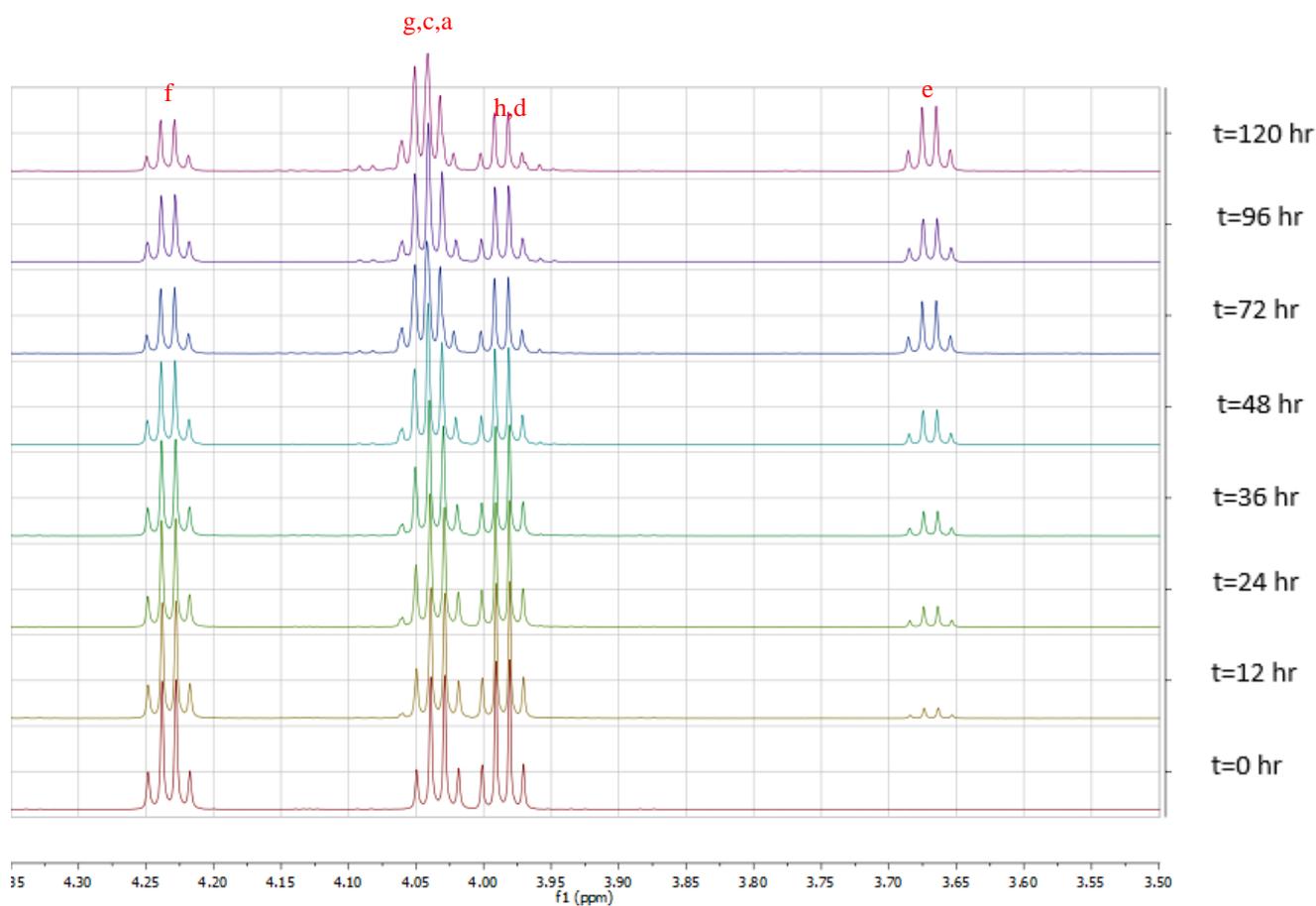
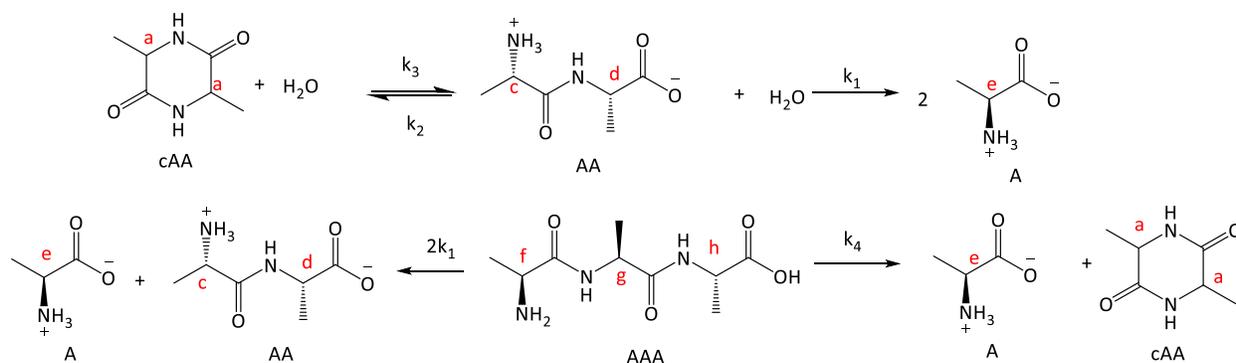
pH=3 AAA analysis	NMR peak integrations					Actual amount ( $\mu\text{mol}$ )				
	time(h)	KHP	A	AA	cAA	AAA	A ( $\mu\text{mol}$ )	AA ( $\mu\text{mol}$ )	cAA ( $\mu\text{mol}$ )	AAA ( $\mu\text{mol}$ )
	0	1	0	0	0	0.9	0	0	0	18
	0	1	0	0	0	0.9	0	0	0	18
	0	1	0	0	0	0.9	0	0	0	18
	12	1	0.02	0	0.03	0.81	1.2	0	0.9	16.2
	12	1	0.02	0	0.03	0.81	1.2	0	0.9	16.2
	12	1	0.02	0	0.03	0.84	1.2	0	0.9	16.8
	24	1	0.03	0	0.05	0.81	1.8	0	1.5	16.2
	24	1	0.03	0	0.06	0.78	1.8	0	1.8	15.6
	24	1	0.03	0	0.06	0.81	1.8	0	1.8	16.2
	36	1	0.04	0	0.07	0.78	2.4	0	2.1	15.6
	36	1	0.04	0	0.08	0.75	2.4	0	2.4	15
	36	1	0.04	0	0.08	0.78	2.4	0	2.4	15.6
	48	1	0.06	0	0.11	0.69	3.6	0	3.3	13.8
	48	1	0.06	0.02	0.1	0.72	3.6	0.6	3	14.4
	48	1	0.06	0.02	0.1	0.72	3.6	0.6	3	14.4
	72	1	0.09	0.04	0.15	0.6	5.4	1.2	4.5	12
	72	1	0.08	0.04	0.14	0.63	4.8	1.2	4.2	12.6
	72	1	0.07	0.02	0.13	0.66	4.2	0.6	3.9	13.2
	96	1	0.11	0.04	0.17	0.54	6.6	1.2	5.1	10.8
	96	1	0.1	0.04	0.16	0.54	6	1.2	4.8	10.8
	96	1	0.1	0.04	0.16	0.57	6	1.2	4.8	11.4
	120	1	0.13	0.06	0.2	0.48	7.8	1.8	6	9.6
	120	1	0.12	0.06	0.19	0.51	7.2	1.8	5.7	10.2
	120	1	0.12	0.06	0.18	0.51	7.2	1.8	5.4	10.2



**Figure S14.** Stacked <sup>1</sup>H NMR spectra for AAA decomposition at 95°C under pH 5. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.4ppm to 4.3ppm.

**Table S16.** AAA degradation reactants and products amounts at varying reaction time under pH 5 at 95°C.

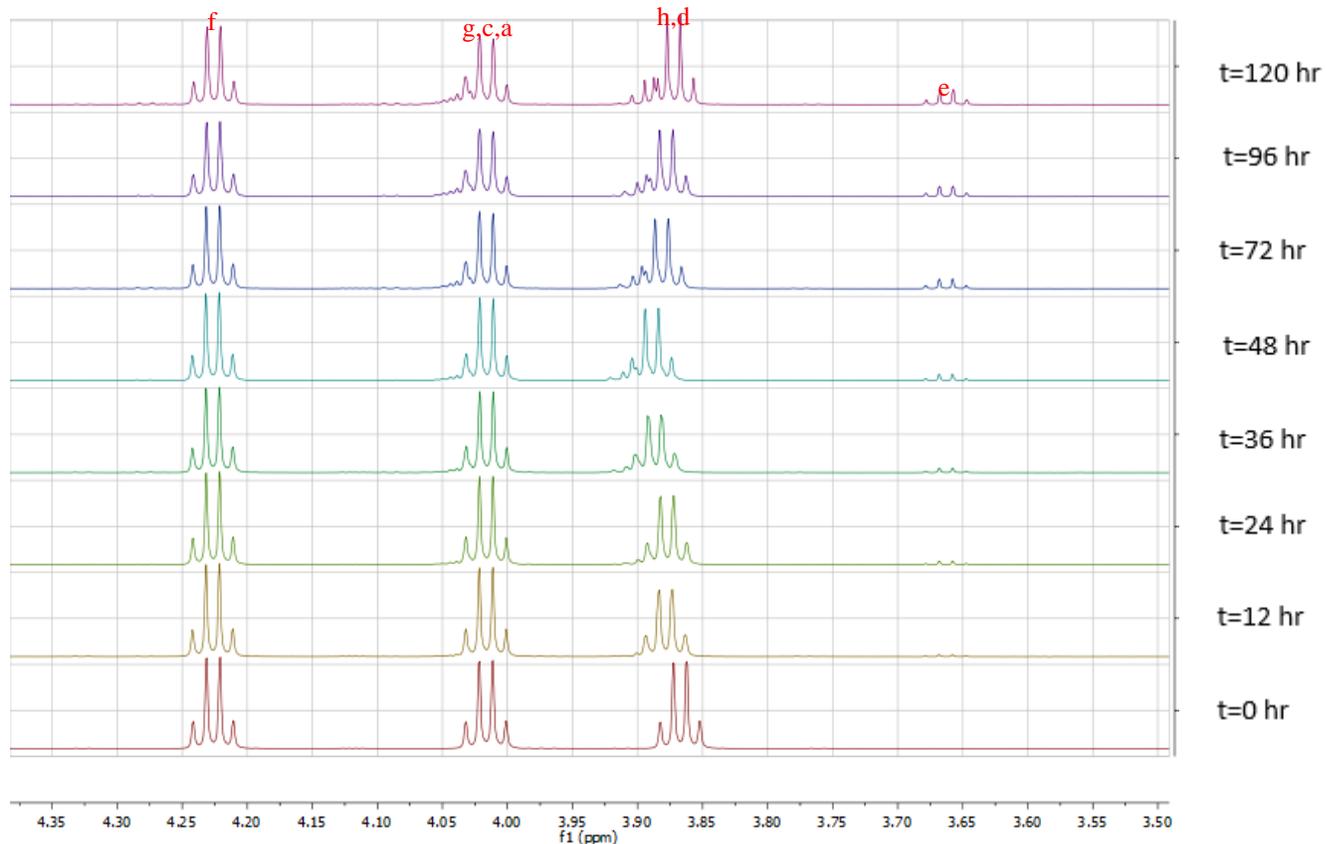
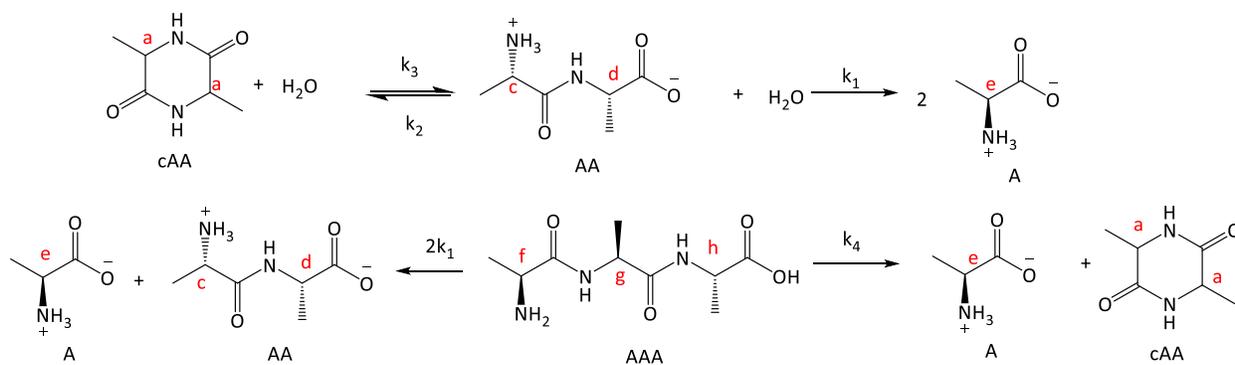
pH=5 AAA analysis	NMR peak integrations					Actual amount ( $\mu\text{mol}$ )				
	time(h)	KHP	A	AA	cAA	AAA	A ( $\mu\text{mol}$ )	AA ( $\mu\text{mol}$ )	cAA ( $\mu\text{mol}$ )	AAA ( $\mu\text{mol}$ )
	0	1	0	0	0	0.96	0	0	0	19.2
	0	1	0	0	0	0.96	0	0	0	19.2
	0	1	0	0	0	0.93	0	0	0	18.6
	12	1	0.05	0	0.11	0.78	3	0	3.3	15.6
	12	1	0.05	0	0.11	0.75	3	0	3.3	15
	12	1	0.05	0	0.1	0.81	3	0	3	16.2
	24	1	0.09	0	0.17	0.66	5.4	0	5.1	13.2
	24	1	0.08	0	0.16	0.66	4.8	0	4.8	13.2
	24	1	0.08	0	0.16	0.72	4.8	0	4.8	14.4
	36	1	0.11	0	0.22	0.57	6.6	0	6.6	11.4
	36	1	0.11	0	0.21	0.6	6.6	0	6.3	12
	36	1	0.11	0	0.21	0.63	6.6	0	6.3	12.6
	48	1	0.14	0.02	0.26	0.51	8.4	0.6	7.8	10.2
	48	1	0.12	0	0.24	0.54	7.2	0	7.2	10.8
	48	1	0.12	0	0.25	0.57	7.2	0	7.5	11.4
	72	1	0.17	0.02	0.33	0.45	10.2	0.6	9.9	9
	72	1	0.17	0	0.34	0.39	10.2	0	10.2	7.8
	72	1	0.16	0	0.32	0.45	9.6	0	9.6	9
	96	1	0.21	0.02	0.41	0.27	12.6	0.6	12.3	5.4
	96	1	0.2	0	0.39	0.3	12	0	11.7	6
	96	1	0.2	0	0.38	0.36	12	0	11.4	7.2
	120	1	0.22	0.02	0.42	0.27	13.2	0.6	12.6	5.4
	120	1	0.23	0.02	0.45	0.21	13.8	0.6	13.5	4.2
	120	1	0.22	0	0.44	0.27	13.2	0	13.2	5.4



**Figure S15.** Stacked  $^1\text{H}$  NMR spectra for AAA decomposition at  $95^\circ\text{C}$  under pH 7. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.5ppm to 4.4ppm.

**Table S17.** AAA degradation reactants and products amounts at varying reaction time under pH 7 at 95°C.

pH=7 AAA analysis time(h)	NMR peak integrations					Actual amount ( $\mu\text{mol}$ )			
	KHP	A	AA	cAA	AAA	A ( $\mu\text{mol}$ )	AA ( $\mu\text{mol}$ )	cAA ( $\mu\text{mol}$ )	AAA ( $\mu\text{mol}$ )
0	1	0	0	0	0.96	0	0	0	19.2
0	1	0	0	0	0.99	0	0	0	19.8
0	1	0	0	0	0.99	0	0	0	19.8
12	1	0.02	0	0.05	0.87	1.2	0	1.5	17.4
12	1	0.03	0.02	0.04	0.87	1.8	0.6	1.2	17.4
12	1	0.02	0	0.03	0.9	1.2	0	0.9	18
24	1	0.05	0	0.09	0.81	3	0	2.7	16.2
24	1	0.04	0	0.08	0.84	2.4	0	2.4	16.8
24	1	0.03	0	0.06	0.84	1.8	0	1.8	16.8
36	1	0.06	0	0.12	0.75	3.6	0	3.6	15
36	1	0.07	0	0.13	0.75	4.2	0	3.9	15
36	1	0.04	0	0.11	0.81	2.4	0	3.3	16.2
48	1	0.09	0.02	0.17	0.63	5.4	0.6	5.1	12.6
48	1	0.1	0.02	0.18	0.63	6	0.6	5.4	12.6
48	1	0.06	0	0.12	0.72	3.6	0	3.6	14.4
72	1	0.14	0.02	0.27	0.51	8.4	0.6	8.1	10.2
72	1	0.09	0.02	0.18	0.66	5.4	0.6	5.4	13.2
72	1	0.11	0	0.23	0.6	6.6	0	6.9	12
96	1	0.12	0.02	0.24	0.57	7.2	0.6	7.2	11.4
96	1	0.15	0.02	0.27	0.51	9	0.6	8.1	10.2
96	1	0.15	0.04	0.29	0.42	9	1.2	8.7	8.4
120	1	0.17	0.02	0.32	0.42	10.2	0.6	9.6	8.4
120	1	0.18	0.04	0.34	0.36	10.8	1.2	10.2	7.2
120	1	0.19	0.04	0.36	0.33	11.4	1.2	10.8	6.6



**Figure S16.** Stacked <sup>1</sup>H NMR spectra for AAA decomposition at 95°C under pH 10. Each spectrum corresponds to a different sampling time. From bottom to top spectrum, indicate increasing sampling times of 0, 12, 24, 36, 48, 72, 96 and 120 h. All the NMR spectra are zoomed in from 3.5ppm to 4.4ppm.

**Table S18.** AAA degradation reactants and products amounts at varying reaction time under pH 10 at 95°C.

pH=10 AAA analysis	NMR peak integrations					Actual amount ( $\mu\text{mol}$ )				
	time(h)	KHP	A	AA	cAA	AAA	A ( $\mu\text{mol}$ )	AA ( $\mu\text{mol}$ )	cAA ( $\mu\text{mol}$ )	AAA ( $\mu\text{mol}$ )
	0	1	0	0	0	0.93	0	0	0	18.6
	0	1	0	0	0	0.93	0	0	0	18.6
	0	1	0	0	0	0.93	0	0	0	18.6
	12	1	0.01	0	0	0.9	0.6	0	0	18
	12	1	0.01	0	0	0.93	0.6	0	0	18.6
	12	1	0	0.01	0	0.87	0	0.3	0	17.4
	24	1	0.01	0.02	0	0.87	0.6	0.6	0	17.4
	24	1	0.01	0	0.01	0.9	0.6	0	0.3	18
	24	1	0.01	0	0.01	0.87	0.6	0	0.3	17.4
	36	1	0.01	0.02	0.01	0.84	0.6	0.6	0.3	16.8
	36	1	0.02	0.02	0.01	0.87	1.2	0.6	0.3	17.4
	36	1	0.01	0.02	0.01	0.87	0.6	0.6	0.3	17.4
	48	1	0.02	0.04	0	0.81	1.2	1.2	0	16.2
	48	1	0.02	0.02	0.01	0.84	1.2	0.6	0.3	16.8
	48	1	0.01	0.02	0.01	0.84	0.6	0.6	0.3	16.8
	72	1	0.02	0.06	0.01	0.78	1.2	1.8	0.3	15.6
	72	1	0.02	0.04	0.01	0.84	1.2	1.2	0.3	16.8
	72	1	0.02	0.04	0	0.84	1.2	1.2	0	16.8
	96	1	0.03	0.06	0.01	0.78	1.8	1.8	0.3	15.6
	96	1	0.03	0.06	0.01	0.75	1.8	1.8	0.3	15
	96	1	0.02	0.04	0	0.84	1.2	1.2	0	16.8
	120	1	0.04	0.08	0	0.75	2.4	2.4	0	15
	120	1	0.04	0.08	0.01	0.75	2.4	2.4	0.3	15
	120	1	0.03	0.06	0.01	0.81	1.8	1.8	0.3	16.2

**Table S19.** The estimated rate constants  $k_1$ - $k_4$  based on joint fitting results for AAA kinetics, under pH 3, 5, 7 and 10, at 95°C.

	$k_{sc} \times 10^6 \text{ (s}^{-1}\text{)}$	$k_{rc} \times 10^6 \text{ (s}^{-1}\text{)}$	$2k_{ro} \times 10^6 \text{ (s}^{-1}\text{)}$	$k_{bb} \times 10^6 \text{ (s}^{-1}\text{)}$
pH=3	0.046	$2.89 \times 10^{-6}$	0.82	1.24
pH=5	$2.45 \times 10^{-6}$	2.01	0.32	3.18
pH=7	0.03	$1.52 \times 10^{-6}$	0.17	1.74
pH=10	$1.66 \times 10^{-6}$	2.24	63.40	0.37

The initial conditions for all AAA degradation rate constants are from GGG fitting results.

**Table S20.** The estimated initial amount of AAA fitting results, under pH 3, 5, 7 and 10, at 95°C.

	AAA <sub>0</sub> (μmol)
pH=3	17.75
pH=5	18.57
pH=7	18.813
pH=10	17.83

## REFERENCES

1. Miller, S. In *Which organic compounds could have occurred on the prebiotic earth?*, Cold Spring Harbor Symposia on Quantitative Biology, Cold Spring Harbor Laboratory Press: 1987; pp 17-27.
2. Bada, J. L.; Lazcano, A., Prebiotic soup--revisiting the miller experiment. *Science* **2003**, *300* (5620), 745-746.
3. Parker, E. T.; Cleaves, J. H.; Burton, A. S.; Glavin, D. P.; Dworkin, J. P.; Zhou, M.; Bada, J. L.; Fernández, F. M., Conducting miller-urey experiments. *JoVE (Journal of Visualized Experiments)* **2014**, (83), e51039.
4. Bada, J. L., New insights into prebiotic chemistry from Stanley Miller's spark discharge experiments. *Chemical Society Reviews* **2013**, *42* (5), 2186-2196.
5. Orgel, L. E., Evolution of the genetic apparatus. *Journal of molecular biology* **1968**, *38* (3), 381-393.
6. Orgel, L. E., Prebiotic chemistry and the origin of the RNA world. *Critical Reviews in Biochemistry and Molecular Biology*. **2004**, *39*, 99-123.
7. Gilbert, W., Origin of life: The RNA world. *nature* **1986**, *319* (6055), 618.
8. Higgs, P. G.; Lehman, N., The RNA World: molecular cooperation at the origins of life. *Nature Reviews Genetics* **2015**, *16* (1), 7.
9. Danger, G.; Plasson, R.; Pascal, R., Pathways for the formation and evolution of peptides in prebiotic environments. *Chemical Society Reviews* **2012**, *41* (16), 5416-5429.
10. Segré, D.; Ben-Eli, D.; Deamer, D. W.; Lancet, D., The lipid world. *Origins of Life and Evolution of the Biosphere* **2001**, *31* (1-2), 119-145.
11. Yu, S.-S.; Krishnamurthy, R.; Fernández, F. M.; Hud, N. V.; Schork, F. J.; Grover, M. A., Kinetics of prebiotic depsipeptide formation from the ester–amide exchange reaction. *Physical Chemistry Chemical Physics* **2016**, *18* (41), 28441-28450.
12. Leman, L.; Orgel, L.; Ghadiri, M. R., Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **2004**, *306* (5694), 283-286.
13. Martra, G.; Deiana, C.; Sakhno, Y.; Barberis, I.; Fabbiani, M.; Pazzi, M.; Vincenti, M., The formation and self-assembly of long prebiotic oligomers produced by the condensation of unactivated amino acids on oxide surfaces. *Angewandte Chemie International Edition* **2014**, *53* (18), 4671-4674.

14. Forsythe, J. G.; Yu, S. S.; Mamajanov, I.; Grover, M. A.; Krishnamurthy, R.; Fernández, F. M.; Hud, N. V., Ester-mediated amide bond formation driven by wet–dry cycles: A possible path to polypeptides on the prebiotic Earth. *Angewandte Chemie International Edition* **2015**, *54* (34), 9871-9875.
15. Lahav, N.; White, D.; Chang, S., Peptide formation in the prebiotic era: thermal condensation of glycine in fluctuating clay environments. *Science* **1978**, *201* (4350), 67-69.
16. Brack, A., From interstellar amino acids to prebiotic catalytic peptides: a review. *Chemistry & biodiversity* **2007**, *4* (4), 665-679.
17. Gisin, B. F.; Merrifield, R., Carboxyl-catalyzed intramolecular aminolysis. Side reaction in solid-phase peptide synthesis. *Journal of the American Chemical Society* **1972**, *94* (9), 3102-3106.
18. Bujdák, J.; Rode, B. M., Silica, alumina, and clay-catalyzed alanine peptide bond formation. *Journal of molecular evolution* **1997**, *45* (5), 457-466.
19. Schwendinger, M. G.; Rode, B. M., Possible role of copper and sodium chloride in prebiotic evolution of peptides. *Analytical Sciences* **1989**, *5* (4), 411-414.
20. Rodriguez-Garcia, M.; Surman, A. J.; Cooper, G. J.; Suárez-Marina, I.; Hosni, Z.; Lee, M. P.; Cronin, L., Formation of oligopeptides in high yield under simple programmable conditions. *Nature communications* **2015**, *6*, 8385.
21. Frenkel-Pinter, M.; Haynes, J. W.; Martin, C.; Petrov, A. S.; Burcar, B. T.; Krishnamurthy, R.; Hud, N. V.; Leman, L. J.; Williams, L. D., Selective incorporation of proteinaceous over nonproteinaceous cationic amino acids in model prebiotic oligomerization reactions. *Proceedings of the National Academy of Sciences* **2019**, *116* (33), 16338-16346.
22. Grover, M.; He, C.; Hsieh, M.-C.; Yu, S.-S., A chemical engineering perspective on the origins of life. *Processes* **2015**, *3* (2), 309-338.
23. Walker, S. I.; Grover, M. A.; Hud, N. V., Universal sequence replication, reversible polymerization and early functional biopolymers: a model for the initiation of prebiotic sequence evolution. *PloS one* **2012**, *7* (4), e34166.
24. Mamajanov, I.; MacDonald, P. J.; Ying, J.; Duncanson, D. M.; Dowdy, G. R.; Walker, C. A.; Engelhart, A. E.; Fernández, F. M.; Grover, M. A.; Hud, N. V., Ester formation and hydrolysis during wet–dry cycles: generation of far-from-equilibrium polymers in a model prebiotic reaction. *Macromolecules* **2014**, *47* (4), 1334-1343.
25. Maruyama, S.; Kurokawa, K.; Ebisuzaki, T.; Sawaki, Y.; Suda, K.; Santosh, M., Nine requirements for the origin of Earth's life: Not at the hydrothermal vent, but in a nuclear geyser system. *Geoscience Frontiers* **2019**, *10* (4), 1337-1357.

26. Lawrence, L.; Moore, W. J., Kinetics of the hydrolysis of simple glycine peptides. *Journal of the American Chemical Society* **1951**, *73* (8), 3973-3977.
27. Radzicka, A.; Wolfenden, R., Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases. *Journal of the American Chemical Society* **1996**, *118* (26), 6105-6109.
28. van Kleef, F. S.; de Jong, W. W.; Hoenders, H. J., Stepwise degradations and deamidation of the eye lens protein  $\alpha$ -crystallin in ageing. *Nature* **1975**, *258* (5532), 264.
29. Marshall-Bowman, K.; Ohara, S.; Sverjensky, D. A.; Hazen, R. M.; Cleaves, H. J., Catalytic peptide hydrolysis by mineral surface: Implications for prebiotic chemistry. *Geochimica et Cosmochimica Acta* **2010**, *74* (20), 5852-5861.
30. Battersby, J.; Hancock, W.; Canova-Davis, E.; Oeswein, J.; O'ONNOR, B., Diketopiperazine formation and N-terminal degradation in recombinant human growth hormone. *International journal of peptide and protein research* **1994**, *44* (3), 215-222.
31. Gu, L.; Strickley, R. G., Diketopiperazine formation, hydrolysis, and epimerization of the new dipeptide angiotensin-converting enzyme inhibitor RS-10085. *Pharmaceutical research* **1987**, *4* (5), 392-397.
32. Oyler, A. R.; Naldi, R. E.; Lloyd, J. R.; Graden, D. A.; Shaw, C. J.; Cotter, M. L., Characterization of the solution degradation products of histrelin, a gonadotropin releasing hormone (LH/RH) agonist. *Journal of pharmaceutical sciences* **1991**, *80* (3), 271-275.
33. Imai, E.-i.; Honda, H.; Hatori, K.; Brack, A.; Matsuno, K., Elongation of oligopeptides in a simulated submarine hydrothermal system. *Science* **1999**, *283* (5403), 831-833.
34. Greenwald, J.; Friedmann, M. P.; Riek, R., Amyloid aggregates arise from amino acid condensations under prebiotic conditions. *Angewandte Chemie* **2016**, *128* (38), 11781-11785.
35. Biron, J.-P.; Pascal, R., Amino acid N-carboxyanhydrides: activated peptide monomers behaving as phosphate-activating agents in aqueous solution. *Journal of the American Chemical Society* **2004**, *126* (30), 9198-9199.
36. Campbell, T. D.; Hart, C. A.; Febrian, R.; Cheneler, M. L.; Bracher, P. J., The opposite effect of  $K^+$  and  $Na^+$  on the hydrolysis of linear and cyclic dipeptides. *Tetrahedron Letters* **2018**, *59* (23), 2264-2267.
37. Smith, R. M.; Hansen, D. E., The pH-rate profile for the hydrolysis of a peptide bond. *Journal of the American Chemical Society* **1998**, *120* (35), 8910-8913.
38. Sakata, K.; Kitadai, N.; Yokoyama, T., Effects of pH and temperature on dimerization rate of glycine: evaluation of favorable environmental conditions for chemical evolution of life. *Geochimica et Cosmochimica Acta* **2010**, *74* (23), 6841-6851.

39. Steinberg, S. M.; Bada, J. L., Peptide decomposition in the neutral pH region via the formation of diketopiperazines. *The Journal of Organic Chemistry* **1983**, *48* (13), 2295-2298.
40. Goolcharran, C.; Borchardt, R. T., Kinetics of diketopiperazine formation using model peptides. *Journal of pharmaceutical sciences* **1998**, *87* (3), 283-288.
41. Sheehan, J. D.; Abraham, A.; Savage, P. E., Reaction pathways and kinetics for tetra-alanine in hot, compressed liquid water. *Reaction Chemistry & Engineering* **2019**.
42. Akiya, N.; Savage, P. E., Roles of water for chemical reactions in high-temperature water. *Chemical reviews* **2002**, *102* (8), 2725-2750.
43. Medina-Ramos, W.; Mojica, M. A.; Cope, E. D.; Hart, R. J.; Pollet, P.; Eckert, C. A.; Liotta, C. L., Water at elevated temperatures (WET): reactant, catalyst, and solvent in the selective hydrolysis of protecting groups. *Green Chemistry* **2014**, *16* (4), 2147-2155.
44. Sun, Y.; Frenkel-Pinter, M.; Liotta, C. L.; Grover, M. A., The pH Dependent Mechanisms of Non-enzymatic Peptide Bond Cleavage Reactions. *Physical Chemistry Chemical Physics* **2019**.
45. Yamada, S.; Hongo, C.; Yoshioka, R.; Chibata, I., Method for the racemization of optically active amino acids. *The Journal of Organic Chemistry* **1983**, *48* (6), 843-846.
46. Orgel, L. E., The origin of polynucleotide-directed protein synthesis. *Journal of molecular evolution* **1989**, *29* (6), 465-474.
47. Zou, W.; Ibrahim, I.; Dziedzic, P.; Sundén, H.; Córdova, A., Small peptides as modular catalysts for the direct asymmetric aldol reaction: ancient peptides with aldolase enzyme activity. *Chemical Communications* **2005**, (39), 4946-4948.
48. Weber, A. L.; Pizzarello, S., The peptide-catalyzed stereospecific synthesis of tetroses: a possible model for prebiotic molecular evolution. *Proceedings of the National Academy of Sciences* **2006**, *103* (34), 12713-12717.
49. Miller, S. M.; Rawlings, J. B., Model identification and control strategies for batch cooling crystallizers. *AIChE Journal* **1994**, *40* (8), 1312-1327.