



HAL
open science

A plausible prebiotic synthesis of thiol-rich peptides : the reaction of aminothiols with aminonitriles

Ibrahim Shalayel

► **To cite this version:**

Ibrahim Shalayel. A plausible prebiotic synthesis of thiol-rich peptides: the reaction of aminothiols with aminonitriles. Organic chemistry. Université Grenoble Alpes, 2018. English. NNT: 2018GREAV055 . tel-02131562

HAL Id: tel-02131562

<https://theses.hal.science/tel-02131562v1>

Submitted on 16 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : Chimie organique

Arrêté ministériel : 25 mai 2016

Présentée par

Ibrahim SHALAYEL

Thèse dirigée par **Yannick VALLEE**, UGA

préparée au sein du **Laboratoire Département de Chimie
Moléculaire**
dans l'**École Doctorale Chimie et Sciences du Vivant**

Synthèse prébiotique plausible de peptides riches en thiol : la réaction des aminothiols avec les aminonitriles

A plausible prebiotic synthesis of thiol-rich peptides: the reaction of aminothiols with aminonitriles

Thèse soutenue publiquement le **17 décembre 2018**,
devant le jury composé de :

Monsieur YANNICK VALLEE

PROFESSEUR, UNIVERSITE GRENOBLE ALPES, Directeur de thèse

Madame MARTINE DEMEUNYNCK

DIRECTRICE DE RECHERCHE, CNRS DELEGATION ALPES, Président

Monsieur JEAN-CLAUDE GUILLEMIN

DIRECTEUR DE RECHERCHE, CNRS DELEGATION BRETAGNE PAYS
DE LOIRE, Rapporteur

Monsieur STEPHANE PERRIO

MAITRE DE CONFERENCES, UNIVERSITE DE CAEN NORMANDIE,
Rapporteur



Dedicated to my beloved angel, Azadeh

ACKNOWLEDGMENTS

In the first place, I would like to thank my supervisor, Prof. Yannick VALLÉE for his enormous support and for his very kind personality. This work would not have been possible without his advice, guidance and brilliant ideas.

I am very grateful to my thesis committee: Prof. Martine DEMEUNYNCK, Prof. Jean-Claude GUILLEMIN and Prof. Stephane PERRIO for their time and their valuable comments and remarks on the manuscript. I would also like to thank Dr. Maxime BRIDOUX who was unfortunately not able to attend the defence.

I would like to acknowledge the “Consulat Général de France à Jérusalem”, Campus France and the “mairie de Gières” for the financial support of my thesis. Special thanks to George MORIN, Anne-Marie, Gilles and “L’association Gieres-Palestine” for their continuous help.

I would like to thank in a very special way Frederic MINASSIAN who helped me in every possible way to handle the difficulties I faced at the beginning of my master. It was not possible for me to keep going without his help. I would like to extend my thanks to Damien JOUVENOT and Rajaa FARRAN with whom I had my first internship in France.

I am so grateful to Prof. Yves GIMBERT and Prof. Jean-Luc DECOUT for their valuable comments throughout the review of my CSI reports. Many thanks go to Anne MILET and Maxime BRIDOUX for the nice corporation in our research. I should also thank Rodolphe GUÉRET, Laure FORT and Amelie DURAND for the time taken to analyse my samples.

I cannot let this opportunity pass without expressing my deep gratitude to my friends and colleagues in SeRCO team: Prof. Jean-Francoise POISSON (the director of our team), Sebastien, Bernard, Pierre-Yves (who was always ready to give a helping hand), Sandrine, Alice, Veronique, Florian, Pascale, Martine, Jerome, David, Pierre-Marc, Nathalie, Mathieu and Benjamin.

Many thanks to Jean Michel, Carlos, Pierre, Beatrice, Christian, Nicolas, Caroline, and all nice people in the DCM. Great thanks go to my friends in the lab with whom I shared nice moments and had hundreds of coffee breaks: Aladdin (le prince), Wathiq, Vianney, Evelyn, Laura and Ali (who designed my citation style on Mendeley). Many thanks to Salia (Le Roi), Anais (La fleure) and Thi (le soleil) for all wonderful moments and for making my stay

enjoyable. I wish the best luck for Naoual (la princesses), Sparta, and Seydou in their PhD journey with Yannick.

Thanks to everyone who participated in the project and shared the fume hoods with me: Ragha, Kieu-Dung, Audrey, Vnghia, Sarah, Alicia, Rudy, Filip, Bilel, Nicolas and Marielle. I can never forget Edgar; it is true that we miss him but he will be always in our thoughts.

I cannot forget to express my deep sense of gratitude to Erin, Jane, Alison, Roy, Ahmed and all my amazing colleagues in the English department at DLST.

Last but not least, I'm sincerely grateful to my beloved family in Palestine and to all my friends, especially Ali (le president) who was also my neighbour for 5 years in Gieres.

Table of Contents

Introduction	5
A. A brief history of evolution	5
B. Conditions on Earth 4 billion years ago	10
C. What molecules could come from space?	13
D. Synthesis of the “molecular bricks” of life on Earth	17
E. Prebiotic synthesis of peptides	23
F. Today’s role of thiols in biology	30
Chapter I: Prebiotic formation of thiol containing dipeptides from aminonitriles and aminothiols	39
I.1 Introduction	39
I.2 Synthesis of aminonitriles	40
I.2.1 Synthesis of racemic α-aminonitriles	40
I.2.2 Synthesis of L-aminonitriles	40
I.2.2 Synthesis of (S)-2-amino-1,4-butanedinitrile 32	45
I.2.3 Synthesis of 2-amino-4-cyanobutanoic acid 35	46
I.2.4 Synthesis of (S)-2-aminopentanedinitrile 38	47
I.2.6 Synthesis of Gly-Gly-CN 42	48
I.2.7 Synthesis of Gly-Gly-Gly-CN 45	48
I.3 Reactions between aminonitriles and aminothiols	49
I.3.1 Reaction of aminoacetonitrile with cysteine	49
I.3.2 Reaction of aminoacetonitrile with homocysteine	53
I.3.3 Reaction of alpha substituted aminonitriles with cysteine and homocysteine	55
I.3.4 N-Acetyl aminoacetonitrile with cysteine	57
I.3.5 Reaction of β-aminopropionitrile with cysteine and homocysteine	58
I.3.6 Competition between α and β-aminonitrile	60
I.3.6.a Competition between α and β-aminonitrile with cysteine	60
I.3.6.b Competition between α and β-aminonitrile with homocysteine	61
I.3.7 Reaction with 2-amino-1,4-butanedinitrile	61
I.4 Formation of Tri and Tetra-peptides	62
I.4.1 Reaction of cysteine with GlyGly-CN and GlyGlyGly-CN	62
I.4.2 Reaction of aminoacetonitrile with Cys-Gly	63
I.4.3 Competition between cysteine and Cys-Gly	64

Table of contents

I.5	Reactions with cysteine amide	65
I.5.1	Reaction of cysteine amide with aminoacetonitrile	66
I.5.2	Reaction of cysteine amide with aminoacetonitrile	67
I.6	Reaction of aminoacetonitrile with cysteamine	69
I.7	Reaction with penicillamine	70
I.8	Theoretical calculations	71
I.9	Observation of a catalytic behaviour of aa-Cys and aa-Hcy dipeptides	72
Chapter II: Prebiotic formation and a possible primitive role of homocysteine thiolactone		79
II.1	Introduction	79
II.2	Synthesis of homocysteine nitrile	79
II.3	Homocysteine nitrile behaviour in water	80
II.4	Prebiotic behaviour of Hcy-thiolactone	82
II.5	Aminoacyl transfer via a thioester	85
II.6	Formation of bis-thiol tetrapeptides	87
Chapter III: Prebiotic formation of polycysteine		91
III.1	Introduction	91
III.2	Synthesis of polycysteine precursors	92
III.2.1	Synthesis of cysteine nitrile 144	92
III.2.2	Synthesis of cysteine thioacid 145	94
III.2.3	Synthesis of cysteine ethyl thioester 146	95
III.2.4	Synthesis of cysteine propyl thioester 147, and cysteine benzyl thioester 148	96
III.3	Formation of polycysteine	96
III.3.1	Starting from cysteine nitrile	96
III.3.2	Starting from thioacid	101
III.3.3	Starting from cysteine thioesters	103
III.3.3.a	Cysteine ethyl thioester	104
III.3.3.b	Cysteine propyl and cysteine benzyl thioesters	107
Chapter IV: Prebiotic formation of thiol-rich polypeptides starting from aminothiols-containing amidonitriles		109
IV.1	Introduction	109
IV.2	Synthesis of aminothiols-containing amidonitriles	110
IV.2.1	Synthesis of Cys-Gly-CN	110
IV.2.2	Synthesis of Cys-Ala-CN, Cys-Val-CN and Cys-Met-CN	110
IV.2.3	Synthesis of Cys-Gly-Gly-CN	112

Table of contents

IV.3 Prebiotic Formation of thiol-rich polypeptides.....	113
IV.3.1 Formation of poly(Cys-Gly).....	113
IV.3.2 Formation of poly(Cys-Ala).....	114
IV.3.3 Formation of poly(Cys-Met).....	117
IV.3.4 Formation of poly(Cys-Val).....	122
IV.3.5 Formation of poly(Cys-Gly-Gly).....	126
IV.4 First attempts to observe a catalytic behaviour of the obtained polymers.....	127
Conclusion	129
Perspectives	132
Experimental Part.....	133
References.....	181

List of abbreviations and symbols

°C	degrees Celsius
aa	amino acid
Ala	alanine
APCI	Atmospheric Pressure Chemical Ionization
aq	aqueous
Asn	asparagine
Bn	benzyl
BOC	<i>tert</i> -butyloxycarbonyl
<i>ca.</i>	around
Cbz	carboxybenzyle
conc.	concentration
COSY	correlation spectroscopy
Cys	cysteine
DCC	<i>N,N'</i> -Dicyclohexylcarbodiimide
DCM	dichloromethane
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EA	ethyl acetate
Equiv.	equivalent
ESI	electrospray ionization
Et	ethyl
Fmoc	fluorenylmethyloxycarbonyl
Ga	giga-annum (a billion year)
Gln	glutamine
Gly	glycine
h	hour
Hcy	homocysteine
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High Performance Liquid Chromatography

List of abbreviations and symbols

HRMS	High-Resolution Mass Spectrometry
HSQC	Heteronuclear Single Quantum Coherence
Hz	hertz
IBCF	isobutyl chloroformate
<i>i</i> Pr	isopropyl
<i>J</i>	coupling constant in Hz
Leu	leucine
Liq	liquid
m.p.	melting point
MALDI	Matrix Assisted Laser Desorption Ionization
Me	methyl
MeOH	methanol
Met	methionine
min	minute
μ	micro
NMM	n-methylmorpholine
NMR	Nuclear Magnetic Resonance
Nu	nucleophile
Ph	phenyl
Phe	phenylalanine
ppm	part per million
Pr	propyl
Pro	proline
Py	piperidine
R _f	retention factor
RNA	ribonucleic acid
rt	room temperature
SM	starting material
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPS	triisopropylsilane
TLC	Thin Layer Chromatography
TMS	trimethylsilyl

List of abbreviations and symbols

TOF	time of flight
Trt	triphenylmethyl
UV	ultraviolet
Val	valine

List of abbreviations and symbols

Introduction

A. A brief history of evolution

“It is often said that all the conditions for the first production of a living organism are now present which could ever have been present. But If (and oh what a big if) we could conceive in some warm little pond with all sorts of ammonia and phosphoric salts, light, heat, electricity etc. present, that a protein compound was chemically formed, ready to undergo still more complex changes”. A letter from Charles Darwin (1809–1882) to his friend Joseph D. Hooker in 1871.¹

Since the early days of Darwin’s vision of his famous “warm little pond”, scientists (alongside with the public) have been inspired by this fascinating idea. Charles Darwin was an exceptional scientist who was experienced in biology and geology. However, experimental evidences of the possibility of emergence of life from such a “warm little pond” have become mainly the duty of chemists. Since some decades, many researchers have taken the responsibility to proof that it would be possible to form the components of life starting from simple molecules which might have been present on the primitive Earth (*Figure 1*).² In order to achieve this goal, researchers exploring the origin of life problem combine geological evidences, theoretical models, and laboratory experimentations that must be carried out under plausible prebiotic conditions.

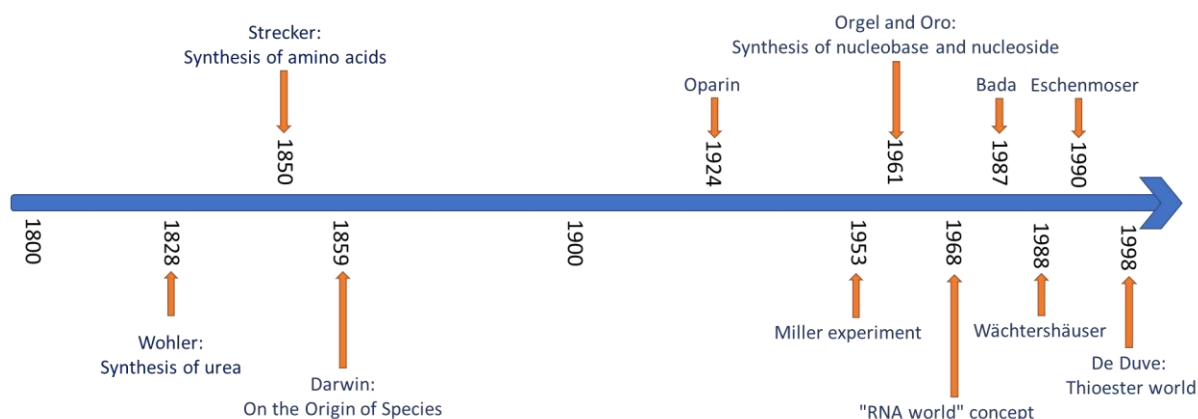


Figure 1: Some main achievements and theories related to the study of origin of life

Pre-Darwin evolutionary ideas

Humans have been always wondering about their surrounding environments and have been trying to answer questions related to phenomena they see. Little by little, they were able to find out many reasonable answers to their inquiries. However, two concerns remained mysterious: life and death. Religion, philosophy and science tried to give answers to these two issues throughout history. Creationists believe that humans were created separately, in other words, they did not undergo evolution. In view of this, it is assumed in the modern scientific communities that evolutionary theories did not exist before Darwin. This belief might arise from the contradiction between the idea of evolution and religious thoughts that dominated the old eras. However, several scholars between the eighth and fourteenth century proposed their ideas about evolution that have some similarities to Darwin's theory.

Al-Jahiz (776 – 868) was an Iraqi zoologist who might be the first to propose an evolutionary biological theory. He observed the behaviour of animals and classified them according to their characteristics and similarities. He explained that adaptation of organisms is directed by the environmental conditions. In spite of similarities between Al-Jahiz's idea and Darwin's theory, the former believed that God guides the evolution of species and that organisms are developing for survival as a part of God's plans to preserve nature in order.³

Another example is the Persian architect, philosopher, physician and scientist, Nasir al-Din Tusi (1201 – 1274). Tusi proposed his elementary theory of evolution about 600 years before Darwin. He believed that elements have evolved into plants, then animals. He also supposed that humans have evolved from advanced animals. He believed in the ability of organisms to change and vary according to environmental conditions, which agrees with the Darwin's concept of natural selection.⁴

An alternative example is Ibn Khaldun (1332 – 1406), the writer of 'Al-Muqaddimah' (Prolegomena) where he presented similar ideas to modern theory of evolution and highlighted human evolution. Ibn Khaldun thought that humans evolved from an ape/monkey ancestor. He rejected the religious belief that dark skin is a sign of God's curse upon sinful human beings. Similarly to Al-Jahiz, he attributed the dark pigmentation of southern people to the hot southern climates.³

For centuries, it was believed that organic matter could only be derived from biological sources or living organisms. The first synthesis of an organic compound (Urea) from inorganic starting materials was accomplished by F. Wohler in 1828. Along the same line, many efforts have been carried out with the purpose of obtaining organic compounds. Notably, A. Strecker found a way to synthesize amino acids starting from aldehydes, ammonia and cyanide.⁵

Oparin's idea went along with previous scenarios. The Russian biochemist first presented his ideas in 1924 in the first version of his booklet entitled "The Origin of Life" where he proposed that life must have arisen from a pool of organic compounds. Oparin claimed that life was the result of a smooth, gradual evolutionary developmental process. That is to say, it had come through chemical evolution.⁶

Chemical evolution

The term "chemical evolution" stands for all chemical processes that took place in the primitive Earth and lead to the emergence of life. Thus, chemical evolution must have preceded the Darwinian biological evolution through which modifications of living things took place over time. Even though the mysterious line between chemistry and life is still not well understood, it is widely believed that the first living cell came into existence through long and accumulated processes of chemical reactions that occurred about 4.5 – 3.5 billion years ago where complex molecules of life had been formed starting from simpler organic and inorganic molecules in the primitive oceans.⁶ No one knows if such process occurred also elsewhere in the universe.

Life definition

Generally speaking, the origin of life could be considered as either the very beginning of biology, or the very last chemical steps which had led to the first living system. In fact, the definition of life is still disputable even with the huge knowledge in biology. One famous definition considers a living system as a combination of cells that can undergo metabolism, grow, adapt to their environment, respond to external effects, and reproduce.⁷ This attempt to define life is indeed nothing but a way to show the characteristics and the observed properties of life.⁸ The absence of an exact definition of life indicates a difficulty to characterize it according to precise factors. Furthermore, some features of life can be observed in non-living

systems, such as polymerization, self-assembly, catalysis and exchange of matter and energy with surroundings. Thus, life cannot be identified by any individual characteristic. However, as explained by Oparin⁹, life is a specific and extremely complicated combination of a large number of properties that can be found individually in non-living materials. Considering the unidentified line between chemistry and life, nowadays definitions of life may be updated or reconsidered.

A common ancestor?

The story of life goes back to about 3.8 billion years ago, maybe more.¹⁰ Despite that the real scenarios for the creation of the first living thing(s) is still indefinite, it is widely accepted that it had been formed via series of chemical reactions that took place in parallel. These reactions might have existed in one place or might have started in different areas before they meet later to form the first living cell.¹¹ All past and present living things derive from one Last Universal Common Ancestor (LUCA). Being the first living thing in the tree of life does not mean that LUCA was a simple organism with narrow capacities; however, it may have been able to show an effective energetic metabolism and may have been also ready to face a wide range of environmental conditions.¹²

It is probable that some kind of “chemical and biological” evolution had taken place before the appearance of LUCA. Being the last common ancestor of all forms of life that we know nowadays, LUCA had to contain membranes, proteins, RNA and DNA. It was able to make replication, transcription, translation and already possessed a complex metabolism. These kinds of perfect and complicated combinations might have been the result of a huge amount of evolutionary processes that might have taken place between the “true” (but lost) origin of life and the appearance of LUCA.¹³ It is possible that many attempts were made before the final form of life, as we know it, appeared.

Biological (Darwinian) evolution

In his book, on the origin of species, Darwin has presented his theory of Evolution. He supposed that all forms of life are derived from one common ancestor and proposed that evolution occurs by the means of natural selection, which he described as “preservation of favoured races in the struggle for life”.¹⁴ Biological evolution stands for the “genetic”

changes over time. These changes can be passed to next generations and this process is necessary for the appearance of new species on the long run.

The tree of life

As aforementioned, it is thought that all living organisms were derived from one unique common ancestor. Every living thing is then connected to every living thing on Earth. This connection is expressed by the tree of life. The tree of life shows the evolutionary relationships between all known types of life. It was originally designed to classify living organisms according to their physical and metabolic properties. Since Darwin, relevant drawings of the tree of life have been updated and modified based on the new discoveries. Nowadays, the tree of life can be drawn in several ways that shows developing of each individual species as an independent branch. Nevertheless, a recent view of the tree of life has been presented. It shows total diversity, includes biological interactions between species and takes into account the linkage between various forms of life, such as the evolution of bacteria in relation to animals (*Figure 2*).¹⁵

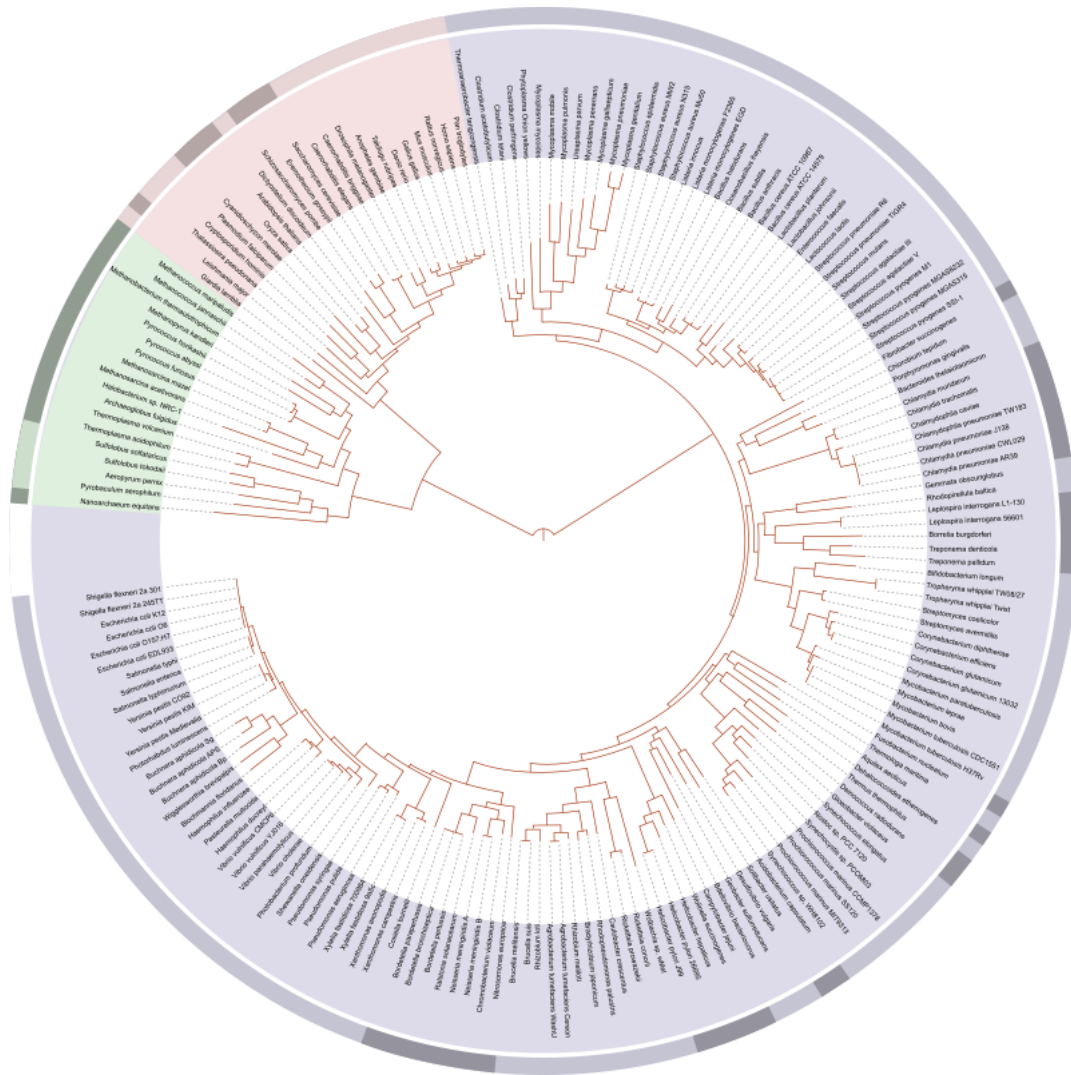


Figure 2: An example of the tree of life: Ivica Letunic, Mariana Ruiz Villarreal. The image was generated using iTOL: Interactive Tree Of Life, an online phylogenetic tree viewer and Tree Of Life resource

B. Conditions on Earth 4 billion years ago

Sir Arthur Conan Doyle once wrote: “When you have eliminated the impossible, whatever remains, however improbable, must be the truth”. Sherlock Holmes in “The Sign of the Four”, 1890

If the emergence of life on Earth was the case to be studied by Holmes, he would realize that the truth cannot be simply found after the elimination of the impossibilities. Nonetheless, several impossibilities can still be eliminated in this special case. For example, unreasonable starting materials or reactions that need conditions that would not be possible on Earth, must

be ignored. Considering liquid water as the solvent of life, any organic solvent must be excluded. As a consequence, proposed reactions must be performed within a limited range of temperatures, from around zero (or a little less considering the presence of salts in the ocean) and not more than 100 °C.

Answers to the questions “when, where and how” did life emerge should also be answered. The debate whether life exists elsewhere in the universe or if “seeds” of life were brought to our Earth from space will not be discussed in this thesis. However, it is our belief that chemical evolution started on Earth.

It was assumed that microbial life might have appeared in the Hadean, the geological era starting from the formation of Earth and ending, more or less, at the end of the heavy meteoritic bombardment. However, it is widely accepted that the story of life started about 3.8 billion years ago, just after the late heavy bombardment¹⁰ that lasted for few hundred million years during which huge number of meteorites struck the Earth. A new era, the Archean that witnessed the arising of life on Earth began near the end of this bombardment. In order to achieve a reasonable answer to how life started on the early Earth, we must first have a better understanding of the early Earth conditions.

Unfortunately, most of the geological record during the Hadean eon had been erased due to the heavy tectonic activity of the young Earth.¹⁶ However, some information about that period has been captured thanks to zircons grains ($ZrSiO_4$) found in Western Australia.¹⁷ These geological evidences demonstrate that the surface of the young primitive Earth was largely covered by liquid water that must have played a vital role in the synthesis of the first organic molecules and the appearance of life in the early Archean. As the moon was much closer to Earth than nowadays¹⁸, oceans were under vigorous motions and powerful tides that helped the surface of Earth to be mostly covered by water, or at least wet.

Despite the fact that sun was younger and weaker, Earth’s temperature might be higher than nowadays due to the high concentrations of carbon dioxide that enhanced the greenhouse effect.¹⁹ However, the temperature of Earth started to cool down before the birth of life, that is before 4 billion years ago, to be similar or just a little warmer than the present time. This could be explained by the declination of carbon dioxide in the atmosphere which helped in reducing the greenhouse effect.¹⁶

Today's ocean is slightly basic (pH 8.2). However, primitive ocean might have been slightly acidic due to the equilibrium with the large amount of carbon dioxide in the atmosphere.¹⁶

The composition of early Earth atmosphere has a degree of uncertainty; however, it can be estimated using thermodynamic modeling.²⁰ Active volcanoes, that were major participants in the chemistry of the primitive atmosphere, liberated large amounts of gases, including hydrogen sulfide, H₂S.²⁰ Beside H₂S and CO₂, early Earth atmosphere contained CH₄, H₂, H₂O, N₂ and NH₃.²¹ Each of these primeval atmospheric components was essential for the scenario of the synthesis of organic compounds in situ on Earth. Quick photochemical destruction of ammonia and methane in the atmosphere might appear as a serious problem against this scenario. However, this problem can be handled by presenting two issues. First, various works have shown that ammonia can be produced when introducing a spark into a mixture of CO₂, N₂ and water. Another argument is that the high altitude organic fog may have led to a shielding effect that protected both methane and ammonia from UV rays.²¹ Furthermore, significant amounts of ammonia could have dissolved in water as the pH of the early oceans was lower than the pK_a of NH₃ (about 9.2 at 25 °C).²²

It is widely accepted that oxygen did not exist in diatomic form, O₂. Conditions were globally reductive, thus iron was in its ferrous state, Fe⁺² and was soluble. Free oxygen started to appear only in the mid of Archean, probably 2.7 billion years ago, and became significant in the Proterozoic era, which followed the Archean.¹⁹ The arise of O₂ during the Archean, and the following eras, was due to photosynthesis. That is to say, the appearance of oxygen was a consequence of life. A relative timeline of different eras that accompanied the appearance of life on Earth is presented in *figure 3*.

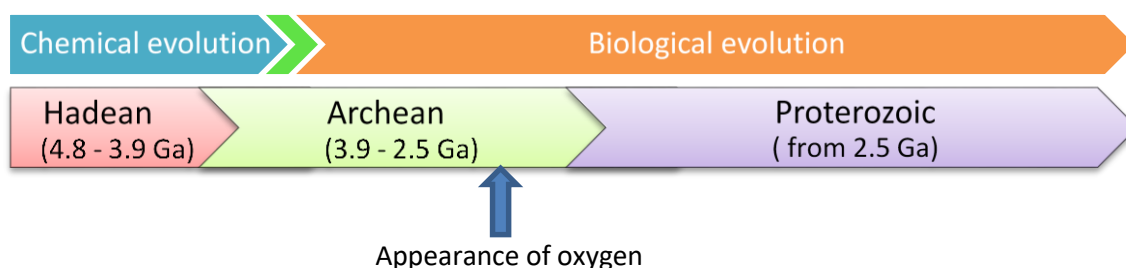


Figure 3: A relative timeline showing the different eras that accompanied the appearance of life on Earth

Beside iron, elements that were present in the primordial continents must have been present also in the primitive ocean including sodium, calcium, magnesium, manganese, phosphorous (as phosphates), potassium and chlorine.²³

C. What molecules could come from space?

One major plausible source of organic molecules on the primitive Earth is their delivery by extraterrestrial objects.²⁴ The organic matter in the present biosphere is estimated to be about 6×10^{14} kg, while the delivery of organic carbon from interplanetary dust particles 4 billion years ago was estimated to be 10^8 kg per year. Such a rate indicates that the role of extraterrestrial organic matter could be important.²⁵

Extraterrestrial objects might have brought inorganic nutrients such as ammonia and phosphates, beside a wide range of organic molecules, including nitriles, aldehydes, may be amino acids, nucleobases, sugar-like compounds and sulfur-containing molecules.²⁵

On one hand, the extraterrestrial compounds can be identified by looking at nowadays meteorites' components. These meteorites are the remains of asteroids that have fallen to Earth. Agencies such as the National Aeronautics and Space Administration (NASA) send expeditions to recover meteorites in Antarctica. These meteorites are less polluted than meteorites found in other environments.²⁶ They contain no terrestrial contaminants, and their composition can be precisely studied.²⁷

On the other hand, extraterrestrial molecules can be identified directly by observing space. The detection of many organic molecules in solar nebula analogues indicates that solar nebula itself (during the formation of Earth) was rich of these compounds. It is broadly believed that these organics were brought to the surface of our young Earth from space by comets and asteroids bombardment.²⁸

Furthermore, it has been suggested that Earth's water might have been captured from asteroids and comets that collided with the planet. In fact, water was one of the first molecules to be observed in the interstellar medium. Microwave radiation of water was observed from the direction of the Orion Nebula, Sagittarius B2 (Sgr B2), and W49.²⁹ In addition to H₂O, singly deuterated water (HDO) was also reported toward the Kleinmann-

Low nebula (KL) in Orion A.³⁰ Moreover, emissions of waters containing heavier oxygen isotopes (H_2^{17}O and H_2^{18}O) were detected as well.^{31,32}

Hydrogen sulfide is a common molecule in space; for example, it was detected in seven star-forming regions.³³ The corresponding isotopes HDS and D_2S were also detected.³⁴

Microwave emission of ammonia was first observed in the interstellar medium toward Sgr B2 in 1968.³⁵ This discovery was followed by observations of other ammonia isotopes. Microwaves of interstellar singly deuterated ammonia (NH_2D) were detected in 1978.³⁶ Later on, doubly deuterated ammonia (NHD_2) was detected in 2000.³⁷ Two year later, the triply deuterated ammonia ND_3 was also reported.³⁸ Moreover, $^{15}\text{NH}_3$ was observed toward the Orion Molecular Cloud.³⁹

Hydrogen cyanide is a simple volatile molecule that was observed in 1P/Halley comet in 1986.⁴⁰ Before this observation, radio emission spectra of the two molecular isotopes of interstellar hydrogen cyanide (H^{12}CN and H^{13}CN) were detected.⁴¹

It has been found (using quantum chemical and kinetic techniques) that HCN molecules cannot dimerize under interstellar conditions, and thus it is impossible to form more complex molecules, for instance adenine, via successive neutral-neutral reactions between HCN units.⁴² Despite the stability of HCN toward the formation of complex organic molecules, HCN assemblies “reported as polymers” were detected in many bodies of our solar system including asteroids, moons, planets and comets.⁴³

Besides HCN, more “complex” cyanides were detected as well. Spectral emission lines associated with the rotational states of acetonitrile (CH_3CN) and cyanoacetylene (HC_2CN) were observed in the protoplanetary disk surrounding the young star MWC 480 (which has roughly twice the mass of our sun and is brighter).²⁸ Longer HC_nCN (n being an even number up to 10) have also been identified.⁴⁴

In 2008, aminoacetonitrile ($\text{H}_2\text{NCH}_2\text{CN}$) was detected in the famous massive star-forming region Sagittarius B2(N).⁴⁵ Aminoacetonitrile is a possible precursor of the simplest amino acid, glycine. For this reason, its detection in the Interstellar Medium is of great interest. This interest is even greater for our team as aminonitriles are engaged directly in our research.

Another nitrile-containing molecule that has been recently detected with radio telescopes in the interstellar medium (in a cold molecular cloud of the Taurus region) is benzonitrile (C_6H_5CN). It may not have a big interest in the study of prebiotic chemistry; however, this nitrile is thought to be a precursor to more complex polycyclic aromatic hydrocarbons and polycyclic aromatic nitrogen heterocycles that are believed to have an extensive existence throughout the universe.⁴⁶

Many aldehydes were also observed in the interstellar medium. Formaldehyde (H_2CO) was first detected in 1969.⁴⁷ One of the first aldehydes to be detected in space was acetaldehyde (CH_3CHO) in 1973.⁴⁸ Since 1980's, more aldehydes have been reported in the literature: Propynal ($HCCCHO$)⁴⁹, propenal (acrolein, $H_2CCHCHO$), propanal (CH_3CH_2CHO)⁵⁰ and glycolaldehyde ($HOCH_2CHO$).⁵¹

Indeed, the detection of glycolaldehyde was of a big interest since it is structurally the simplest member of the monosaccharide sugars (monosaccharide sugars are carbohydrates with the formula $C_nH_{2n}O_n$).⁵¹

Amino acids, which are the monomers of proteins, could have been also delivered to the primitive Earth through comets and meteorites. Several amino acids were first detected in the Murchison meteorite, which fell to Earth near Murchison village in southeastern Australia in September 1969. The Murchison meteorite contained more than 100 different amino acids (ranging from two to nine carbon atoms). Among these amino acids, 80 have been clearly identified.⁵² Not surprisingly, the most abundant amino acid was glycine. Alanine, glutamic acid, valine and proline are other examples of proteinogenic amino acids that were detected. However, this meteorite also contained a number of non-proteinogenic amino acids, for instance α -aminoisobutyric acid.⁵³

Many amino acids (despite that most of them are not engaged in today's proteins formation) were also observed in other meteorites. These observations are a solid argument for the extraterrestrial origin of at least a part of amino acids. What is more interesting about few of meteoritic amino acids is the slight enantiomeric excess of the L- over the R-amino acids.²⁷ These meteoritic amino acids might have contributed to the origin of homochirality in life on Earth.⁵⁴

In carbonaceous chondrites (a rare class of meteorites) more than 80 amino acids were detected, 12 of them were protein-amino acids: alanine, aspartic acid, glutamic acid, glycine, isoleucine, leucine, phenylalanine, proline, serine, threonine, tyrosine, and valine.²⁵

Up to now, cysteine has not yet been observed in space, however, it has been proposed that cysteine precursor CysCN ($\text{HSCH}_2\text{CH}(\text{NH}_2)\text{CN}$) may be synthesized in interstellar dust clouds by radical combination reactions.⁵⁵

All organisms depend on nucleic acids to encode genetic information. Nucleobases are substituted one- or two-ring nitrogen heterocyclic compounds that are essential for this process. For this reason, it is strongly believed that nucleobases might have played an essential role in the early evolution of life. In fact, researches concerning the detection of nucleobases in meteorites are much less developed in comparison to those that were carried out for amino acids. However, the detection of some nucleobases in meteorites were reported in the literature.⁵⁶ Purines (Adenine, guanine, hypoxanthine, and xanthine) and one pyrimidine (uracil) have been identified in the Murchison, Murray and/or Orgueil meteorites.^{57,58} Other unusual nucleobase analogues that are not found in Earth's biology (purine, 2,6-diaminopurine, and 6,8-diaminopurine) were also identified in these meteorites.⁵⁶ The detection of various nucleobases in meteorites, support their extraterrestrial origin (Figure 4).²⁷

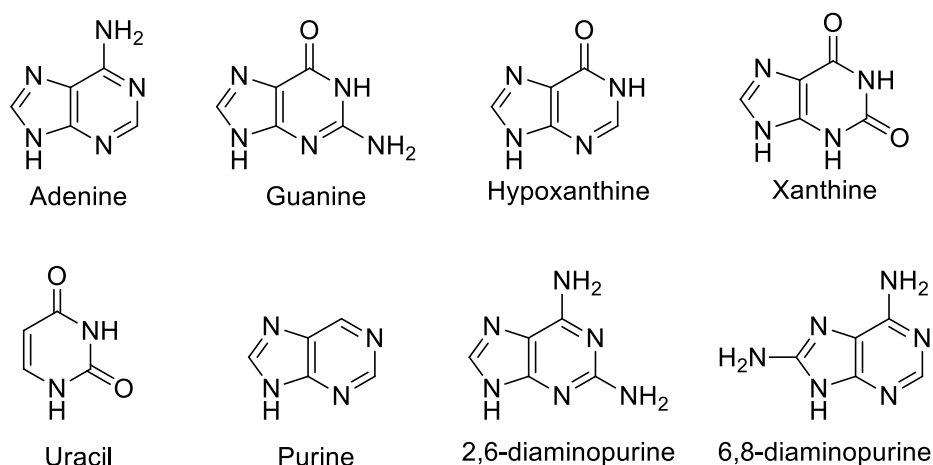


Figure 4: Some nucleobases that were detected in meteorites

D. Synthesis of the “molecular bricks” of life on Earth

Among all contributions in the 20th century, the Miller experiment (or Urey-Miller experiment) remains the most famous scientific breakthrough in the domain of prebiotic chemistry in which organic compounds including several amino acids have been detected. Miller ran his experiment in simulated primitive conditions by applying an electric discharge to a mixture of gases in a reducing atmosphere (CH₄, H₂, H₂O and NH₃).⁵⁹

Miller’s work has been continued by his student, J. Bada. Bada analysed/re-analysed archived samples of Miller’s original experiments found in his laboratory using more sensitive techniques. While a relatively small number of amino acids had been detected by Miller, Bada found that more than 40 amino acids and amines were synthesized.⁶⁰ Furthermore, seven organosulfur compounds in unreported 1958 Miller’s samples, in which H₂S gas was considered in the primitive atmosphere, were identified using HPLC and TOF mass spectroscopy.⁶¹ All of the amino acids that were identified in the Murchison meteorite have been also found in Miller experiment’s products.⁶²

In fact, Bada was not the first to detect sulphur containing amino acids. Miller himself have reported that methionine was a product of the action of an electric discharge on a simulated primitive reduced atmosphere containing H₂S or CH₃SH.^{61,63,64} Beside racemic methionine, other sulphur-containing organic compounds were detected again in unpublished 1958 Miller’s experiment samples.⁶¹

Miller experiments are not the only prebiotic process to yield amino acids. Plausible prebiotic synthesis of amino acids from HCN reactions in water has been reported.⁶⁵

It is strongly believed that amino acids are not the direct products of the electric discharge in Miller experiment. They probably arise from a Strecker synthesis (*Figure 5*), followed by the hydrolysis of transiently formed amino nitriles in aqueous phase.^{22,66}

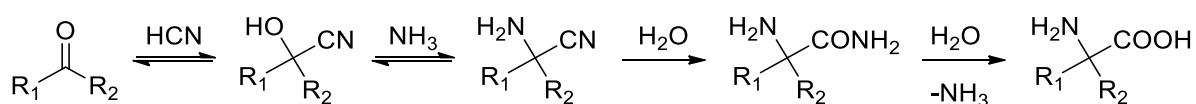


Figure 5: The Strecker mechanisms for the formation of amino acids from NH₃, HCN, and aldehydes (or ketones)

In 1990, a Strecker reaction was proposed to be the last step in a plausible prebiotic synthesis of histidine starting from erythrose and formamidine. In this case, the first product was Imidazole-4-acetaldehyde, which was then converted to histidine in the presence of hydrogen cyanide and ammonia.⁶⁷

Nucleobase

Miller's experiment have led to the modern era of prebiotic chemistry. From that time on, prebiotic chemists have been trying to synthesize biomolecules under plausible prebiotic conditions. J. Oro was able to report the first prebiotic synthesis of the nucleobase adenine, a basic constituent of nucleic acid, starting from hydrogen cyanide and ammonia in 1961.^{68,69}

Adenine can be regarded as a pentamer of HCN. The initial step in a possible pathway to adenine is the dimerization of HCN followed by further reactions to give HCN trimer (aminomalononitrile) and tetramer (diaminomaleonitrile). Then the tetramer undergoes photochemical isomerization followed by a ring closure reaction, and finally the ring reacts with another HCN molecule to provide adenine (*Figure 6*).⁷⁰

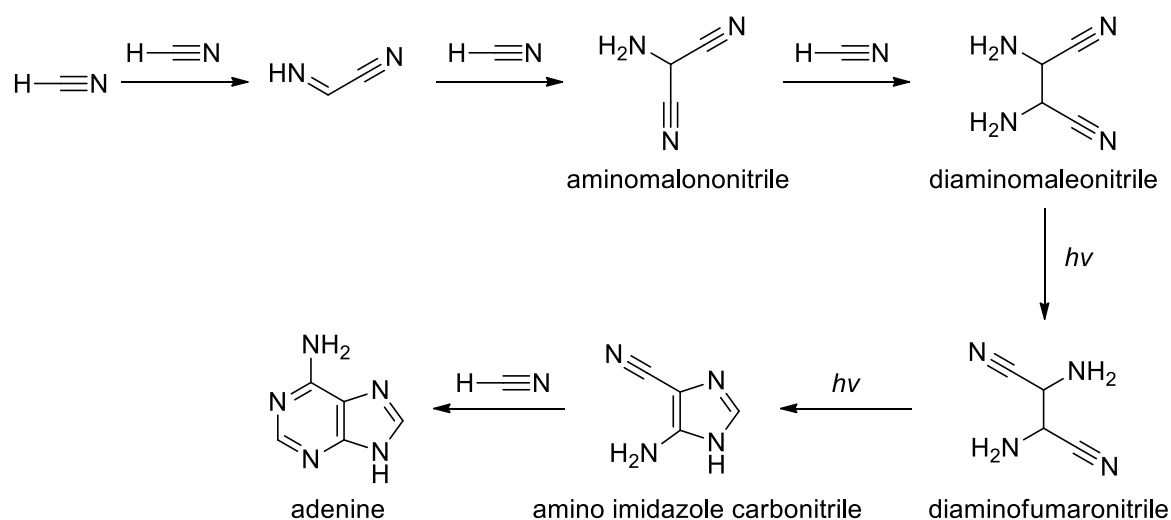


Figure 6: Proposed pathway for the formation of adenine in aqueous NH_4CN solution

Furthermore, the photochemical rearrangement of diaminomaleonitrile in sunlight to give aminoimidazole carbonitrile was demonstrated by Ferris and Orgel in 1966.⁷¹ The Imidazole ring can also give other purines in this mixture.²²

Beside purines, the prebiotic synthesis of Pyrimidines was also investigated. Uracil was produced from a mixture of malic acid and urea.⁷² Few years later, the synthesis of cytosine from cyanoacetylene (HCCCN) and cyanate (NCO⁻) was reported (*Figure 7*).^{22,73}

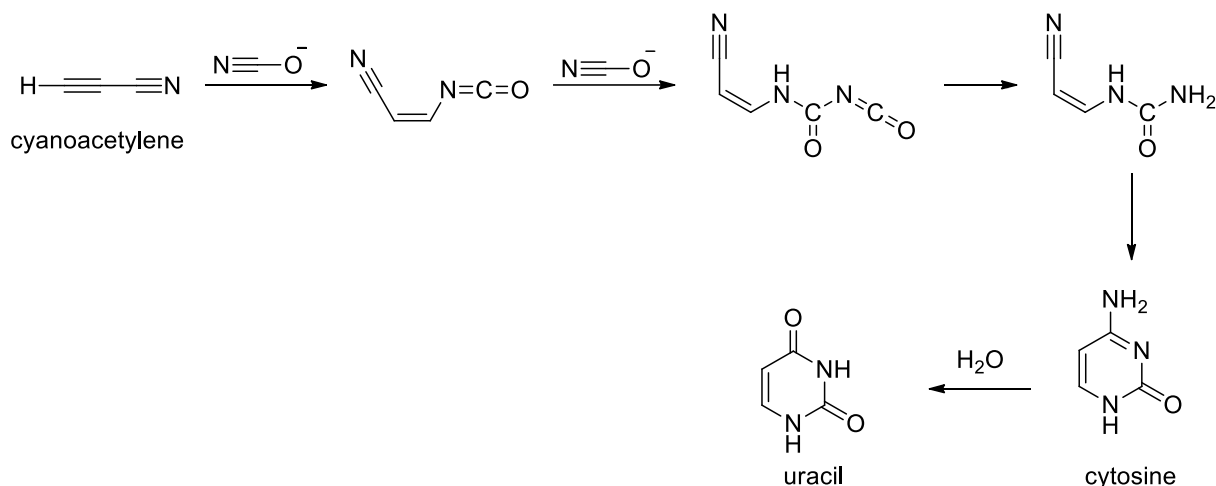


Figure 7: Proposed mechanisms for the prebiotic synthesis of pyrimidines

Thymine was obtained in a good yield from the reaction of uracil with formaldehyde and formate.⁷⁴

Sugars

The formose reaction (also known as Butlerov synthesis) has been considered as a prebiotic source of sugars. This reaction was discovered by Butlerov's in 1861.⁷⁵ It has a great importance concerning the study of the origin of life, as sugars (ribose and deoxyribose) are key molecules to construct RNA and DNA skeletons.

In the formose reaction, formaldehyde oligomerizes in the presence of alkaline earth hydroxides (*eg.* Ca(OH)₂ or CaCO₃). It is a simple reaction in which sugar elongation is achieved via repetitive additions of formaldehyde molecules. The importance of alkaline earth hydroxides as a catalyst for this reaction is not a problem from a prebiotic point of view as these alkalines can be found in (or replaced by) clays.²²

The presence of Kaolin, a white clay composed mainly of kaolinite (Al₂Si₂O₅(OH)₄), in the reaction mixture can catalyse the formation of sugars, including ribose.^{22,76} Pentoses were

formed from mixtures of formaldehyde, glyceraldehyde, and borate minerals such as colemanite ($\text{Ca}_2\text{B}_6\text{O}_{11}5\text{H}_2\text{O}$) or kernite ($\text{Na}_2\text{B}_4\text{O}_7$).⁷⁷

The formose reaction is favoured by the presence of glycolaldehyde, which is the first product of the reaction.²² The observation of glycolaldehyde in space endorses the importance that the formose reaction might have had as a major pathway in the synthesis of sugars on the early Earth.

Figure 8 shows a simple pathway to some of the common sugars, however, more complex mixtures of sugars would be obtained from the formose reaction. For example, Decker *et al.* reported more than 40 sugars in one reaction mixture.⁷⁸

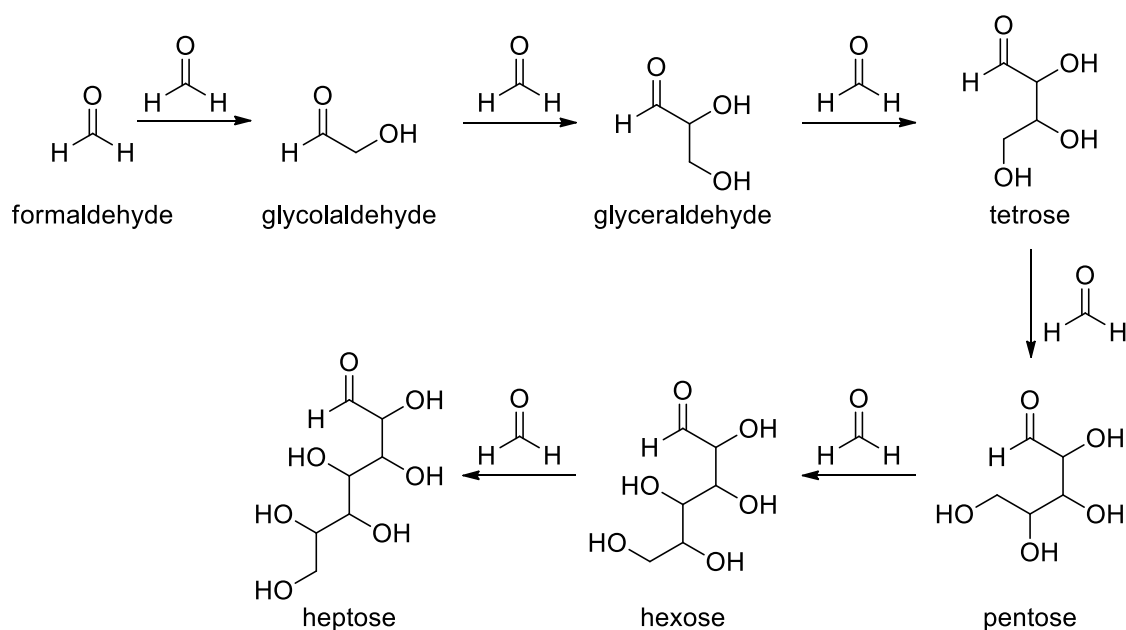


Figure 8: A simple pathway for sugar synthesis via the formose reaction

On the other hand, the formose reaction is not universally accepted for the prebiotic synthesis of sugars because of some weaknesses in this proposal. The difficulty to control the reaction (*i.e.* nothing would stop the elongation process under alkaline media) is one reason for this rejection. Another problem is that the conditions of the reaction may lead to the degradation of sugars.⁷⁹ For example, at pH seven, ribose and other sugars are found to be labile and are able to decompose.⁸⁰ Furthermore, sugars are able to undergo several reactions in a rate that would be faster than their accumulation in the environment.²²

As an alternative prebiotic synthesis, sugar synthesis was achieved by photoredox systems. It has been reported that simple sugars can be formed from hydrogen cyanide by ultraviolet irradiation in the presence of cyanometallates.⁸¹

Nucleosides and nucleotides

As prebiotic synthesis of both sugars and nucleotides are possible, it was logical to go further in the construction of more complex biomolecules. In this regard, we will discuss the attempts to obtain nucleosides and nucleotides as crucial examples of “more complex” biomolecules. Nucleosides are essential components in DNA and RNA, and understanding their development on Earth is fundamental in the study of the origin of life.

Of course, in today’s biology enzymes facilitate the synthesis of nucleotides, but it is not realistic to imagine that such evolved bio-catalysts were involved in the early stages of chemical evolution.⁸² Therefore, “nonenzymatic” prebiotic synthesis should be proposed.⁸³

Here, a lot of questions arise, for instance: why was ribose preferred over other sugars? Why not another sugar? It is also still unclear why the present RNA bases (but not other possible alternatives that were probably present on early Earth) would have been selected by prebiotic reactions.⁸⁴

The earliest attempts to obtain nucleosides in a prebiotic way was simply to heat a dry mixture of ribose and nucleobases (purines or pyrimidines). About 8 % of β -D-inosine was obtained when hypoxanthine was used in the presence of a mixture of some seawater salts. Mixtures of α - and β -isomers were obtained in low yields when adenine and guanine were used.⁸⁵ However, heating a mixture of ribose and uracil did not give the expected nucleosides, uridine. Similarly, cytidine was not produced when cytosine was used (*Figure 9*).²² The synthesis of pyrimidine nucleoside was achieved using a similar methodology.⁸⁶

Introduction

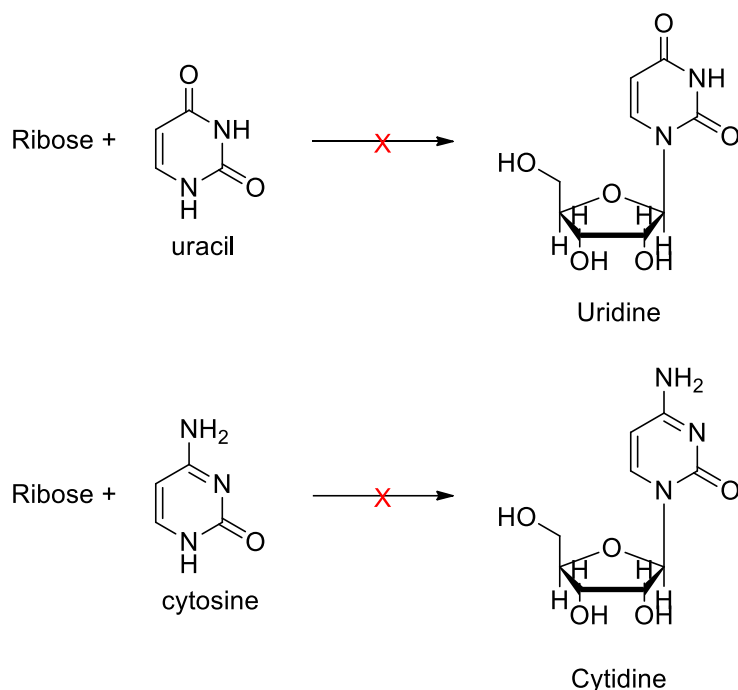


Figure 9: Heating ribose with uracil (and cytosine) did not give the corresponding nucleotides

Phosphates are believed to be an essential compound in primitive Earth. Small amounts of condensed phosphates are emitted in volcanic fumaroles.⁸⁷ Prebiotic formation of nucleotides were attempted through nucleosides phosphorylation. It was found that nucleosides can be phosphorylated upon dry heating with acidic phosphates.⁸⁸ Trimetaphosphate (TMP) has been found to be an active phosphorylating agent for nucleosides.²² Diiminosuccinonitrile (DISN) was considered as a potential prebiotic phosphorylating promoter.⁸⁹

Very interestingly, significant yields of ribonucleotides have been obtained in stepwise reactions, in which products were isolated after each reaction and submitted to next step.⁹⁰ This methodology was able to produce pyrimidine and purine ribonucleotides in good yields.⁹¹ Moreover, a prebiotic transformation of a pyrimidine nucleoside into a purine nucleoside has been reported.⁹²

A recent study has shown how a possible RNA ancestor (2,4,6-triaminopyrimidine (TAP)) can be glycosylated in water without the presence of ribose. This study supposed that ribose might not have been essential for prebiotic nucleoside formation at the very beginning of life development, as ancient RNA (reported as proto-RNA) might have used TAP (or other similar models).⁸⁴ It has also been reported that 2-aminoimidazole, that was found to be a

plausible prebiotic molecule, can act as a nucleotide activating group for nonenzymatic RNA copying.⁸³

E. Prebiotic synthesis of peptides

The big role proteins play in the present life suggests that amino acids and peptides might have had a huge contribution in the emergence and development of life.⁹³

It is very likely that amino acids were present in the primitive ocean. As we have previously reported it, amino acids are the most interesting products from Miller's experiment.⁵⁹ Strecker synthesis is a reasonable prebiotic reaction that leads to the formation of these compounds.⁵ Moreover, many amino acids have been detected among the organic contents of meteorites.²²

Peptide bond formation in water starting from α -amino acids is thermodynamically unfavorable, but not impossible. Indeed, spontaneous peptidization of several amino acids in aqueous medium, *eg.* Cysteine⁹⁴ and Serine⁹⁵, has been demonstrated. Considering a primitive soup full of different amino acid units, every single addition to the peptidic sequence would give the newly formed peptide a different reactivity toward the hydrolysis of some parts of it or promoting the synthesis of more complex peptidic sequence. Several attempts to obtain peptides under prebiotic conditions (*eg.* drying/wetting cycle experiments) ended up with diketopiperazines as major undesired byproducts.⁹³ However, it is assumed that these cyclic dipeptides can still be considered as precursors of dipeptides upon hydrolysis, and they can be opened by aminolysis, and thus lead to peptide elongation. This scenario would be a game of "trial and failure" processes toward self-organization and polymerization (*Figure 10*).⁹³

Introduction

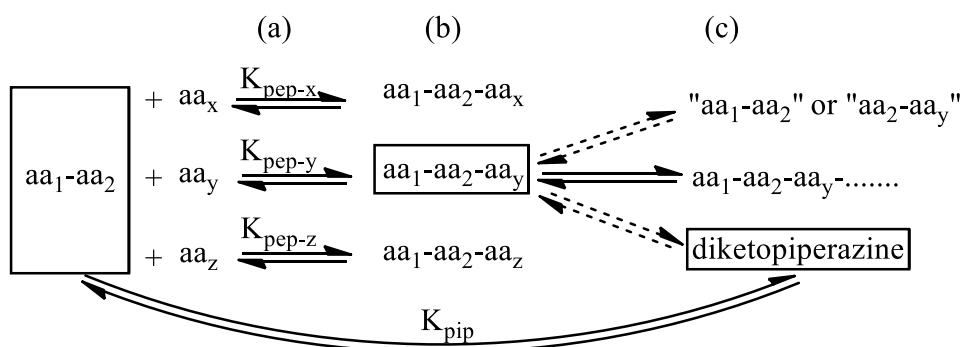


Figure 10: A plausible simple example; a dipeptide surrounded by different amino acids (x , y and z). Preferred routes depend on equilibrium constants for the formation of peptide bonds (a). Every resulting peptide (b) has a specific reactivity that control the following steps (c) (i.e., hydrolysis, elongation, or cyclization).

However, resulting peptides from spontaneous polymerization of amino acids would be obtained only with very low yields. Moreover, considerable amount of hydrolysis of peptide bonds under the postulated conditions in the primitive ocean must be taken into account. Warming up the reaction media will increase the equilibrium constants for the formation of peptides; however, very high temperatures are unreasonable from a prebiotic point of view. Therefore, it would be hard to form significant amounts of peptides without the existence of catalysts in order to fill the energy gap for peptide formation.⁹⁶ Adenosine triphosphate (ATP) is used as a catalyst and provides the required energy for peptide bond formation in nowadays living systems.

In a primitive model, activating agents simpler than ATP and/or activated forms of amino acids are required to promote peptide elongation. Among models that have been proposed the formation of N-carboxyanhydrides of amino acids seems reasonable. These compounds could have been formed in the presence of bicarbonate in high concentration.⁹⁶ It has been reported that this process can be achieved in the presence of cyanate by converting first amino acids into N-carbamoylamino acids which then can form N-carboxyanhydrides in the presence of carbonate (in equilibrium with saturated CO_2 in the aqueous medium) (Figure 11-A).⁹⁶ Under similar conditions, peptide bonds were also obtained from dicarboxylic acids, including aspartic acid, derivatives in which the extra carboxylic acid in these types of molecules would undergo an intramolecular cyclization in order to form succinic anhydrides that can be reacted with nucleophiles in following steps (Figure 11-B).⁹⁶

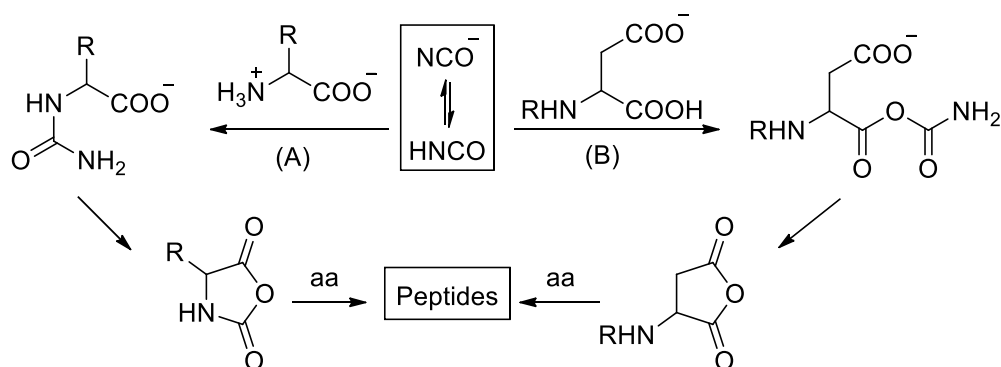


Figure 11: Elongation of peptides from either c-terminal (A) or N-terminal (B) activation of amino acids in the presence of cyanate

It has been proposed that amino acids can be activated and converted into peptides under hot, anaerobic, aqueous conditions. Indeed, in the presence of NiS, FeS, CO and H₂S (or CH₃SH) peptide bonds were formed under “geothermal conditions”.⁹⁷ These results endorse the scenario that first reactions that lead to the formation of life had been catalysed on the surfaces of metal sulfide minerals.

Indeed, some scientists who study the origin of life went farther with minerals. They proposed that they had the crucial role in the appearance of life.

Iron-sulfur world

Iron-sulfur world was one of the pioneering hypothesis in this domain. Iron-sulfur clusters have a vital role in metabolism and they are involved in a broad range of essential functions in all living cells such as DNA maintenance, gene expression regulation, protein translation, energy production, and antiviral response. Mitochondrion, endoplasmic reticulum, cytosol, and the nucleus contain Fe-S Proteins. Fe-S enzymes participate in the metabolism of amino acids in fungi.^{98,99} As a result of this essential role of Fe-S clusters as cofactors of proteins involved in biological cells, humans with a deficiency in the production of any these related protein suffer from serious diseases.¹⁰⁰

The fundamental role of Fe-S clusters (see p.34, Figure 19) of nowadays life can be considered as a biological heritage that is still kept for the earliest days of life.¹⁰¹

Wächtershäuser proposed his idea of an “iron-sulfur world” in 1988. He suggested that early life might have formed on the surface of iron sulfide minerals. It has been later demonstrated

that iron-sulfur clusters may have formed on early in the Earth ocean, with the help of UV light that drives their synthesis. These clusters may have coordinated and had been stabilized by a wide range of thiol-containing peptide.⁹⁸

Wächtershäuser and Huber have found that an activated thioester can be formed starting from CO and CH₃SH in the presence of FeS.¹⁰² Furthermore, thioacetic acid, which can be regarded as a simple acetyl-CoA analogue, was also formed in the presence of H₂S. The authors suggested that this reaction can be considered as the leading reaction for a chemoautotrophic origin of life in which, for example, amino acids would have been converted into simple peptides on FeS surfaces.

Unlike photosynthesis, the term “chemoautotrophic” was presented by Wächtershäuser to describe the system where simple forms of life obtain their nutrition via the oxidation of non-organic compounds, in this case FeS. He postulated a pioneer organism with a merged composition between an organic superstructure and an inorganic substructure. In this scenario, the growth of the organic superstructure would have been achieved thanks to the Fe-S surface that could afford a reproductive metabolism by an autocatalytic process. Hence, the theory of a chemoautotrophic origin of life proposed that a common origin of Bacteria, Archaea and Eukarya had arisen through a volcanic iron-sulfur world.¹⁰³

Iron and sulphide may be arranged in cage structures ranging from [Fe₂-S₂], [Fe₃-S₄], [Fe₄-S₄], up to [Fe₈-S₈] clusters.¹⁰¹ A familiar example of iron-sulfur structures is pyrite (FeS₂). Wächtershäuser proposed that life emerged on pyrite surfaces in volcanic or hydrothermal vents. On pyrite surfaces, peptides, nucleic acids and lipids would have been formed.¹⁰⁴

Rusell and Hall have gone farther and proposed that life had emerged from iron monosulphide bubbles that could have worked as primitive membranes in the Hadean ocean.¹⁰⁵

On the other hand, the iron-sulfur world theory was rejected by some scientists who believed that sequenced cellular genomes of nowadays living systems does not support Wächtershäuser proposal.¹⁰⁶

Zinc world

Alternatively, the zinc world hypothesis proposes that the emergence of life started on zinc sulfide (ZnS, sphalerite) surfaces of hydrothermal origin and was derived by UV-rich solar radiation. ZnS could precipitate at the surface of continents under the exposure of solar light.¹⁰⁷ These mineral structures submitted to UV radiations would have facilitated the formation of the first biomolecules. These mineral sulfides might have also absorbed the excess radiation and thus protect biomolecules from photo-dissociation. Zn²⁺ ions are expected byproducts in the photosynthesis of ZnS structures. As a result, life might have been evolved under high levels of Zn²⁺ ions. Indeed, this suggestion agrees with the fact that Zn²⁺ ions have important and specific roles in modern organisms.¹⁰⁸

Zinc world and iron-sulfur world scenarios promote the “metabolism-first” concept. This concept proposes that replication processes were preceded by many other chemical reactions that had been catalysed by minerals and had yielded a mixture of organic molecules. These compounds were able to react, polymerize, and undergo self-assembly in order to form proteins, RNA and DNA later.

Thioester world

Another hypothesis that is related to metabolism first concept is de Duve's thioester world.¹⁰⁹ Christian de Duve suggested that a thioester world might have preceded the RNA world (that we will discuss later) in the prebiotic evolution. He suggested that peptides might have been synthesized with the help of thioester bond energy, establishing a growing protometabolism.^{110,111}

Thioester bond has a high-energy character and thus is highly reactive toward amino acids. The thioester world hypothesis has gained a strong support since it has been first proposed.¹¹¹

The most famous example of biological thioesters is CoA (see p.33, Figure 16). This well-known coenzyme contributes in many biological processes in all domains of life. Indeed, CoA can be regarded as a molecular fossil from an early metabolic state and that has been kept in life's memory.¹¹² An ancient CoA-thioester model might have appeared in the very early form of life. It might have been even participated in the metabolism of LUCA itself!^{111,113}

RNA world

Alexander Rich first proposed the concept of the RNA world in 1962.¹¹⁴ He assumed that RNA might have played the key role in the early history of life after discovering that ribozymes, a type of RNA molecules, catalyse the formation of proteins beside being able to replicate itself. The “replication-first” concept thus suggests that the appearance of RNA, which is able to carry and transfer information, preceded the formation of other metabolisms. In order to form simple RNA strands, nucleotides had to form and polymerize, which might be possible in hydrothermal vents in the ocean or in “warm little ponds” in the primitive Earth.¹¹⁵

Even though, it is the nowadays most consensual theory, RNA world hypothesis has faced several objections due to some accompanying difficulties such as the very limited catalytic properties, the instability, and the high complexity of the RNA molecules, complexity that makes their prebiotic synthesis difficult to imagine.¹¹⁶

RNA world / protein world: A partnership?

In nowadays biochemistry, proteins catalyse chemical reactions and DNA stores the genetic information (with RNA transmitting the information from DNA, and allowing the synthesis of proteins). The problem whether proteins or nucleotides came first presents a famous chicken-or-egg dilemma.

Still believing in the RNA world, some scientists do not believe in an independent RNA world and they argue that the RNA world alone could have never existed without amino acids and peptides that “shaped” it.¹¹⁷

On one hand, RNA world concept relies on the fact that RNA is able to carry and transfer information. On the other hand, proteins-first concept depends on the catalytic properties of proteins for a wide range of chemical reactions. However, it has been proposed that both peptides and oligonucleotides might have contributed in the emergence of life in a complementary partnership.¹¹⁸ With this hypothesis, it is reasonable to assume that both metabolism and genetics shared a common origin.

It has also been assumed that RNA, protein and even lipid precursors might have arisen from common origins starting from hydrogen cyanide and hydrogen sulfide in the presence of

UV light.¹¹ This scenario suggested that reactions had been developed separately in different streams and that the resulting products had been mixed together later in pools, in a way very similar to what is now done in laboratories with flow chemistry.

Another scenario has nominated formamide as the possible simple starting material for the synthesis of a wide range of primordial biomolecules that are necessary for both genetic and metabolic processes such as nucleic bases, nucleotides, sugars, and amino acids (*Figure 12*).¹¹⁹

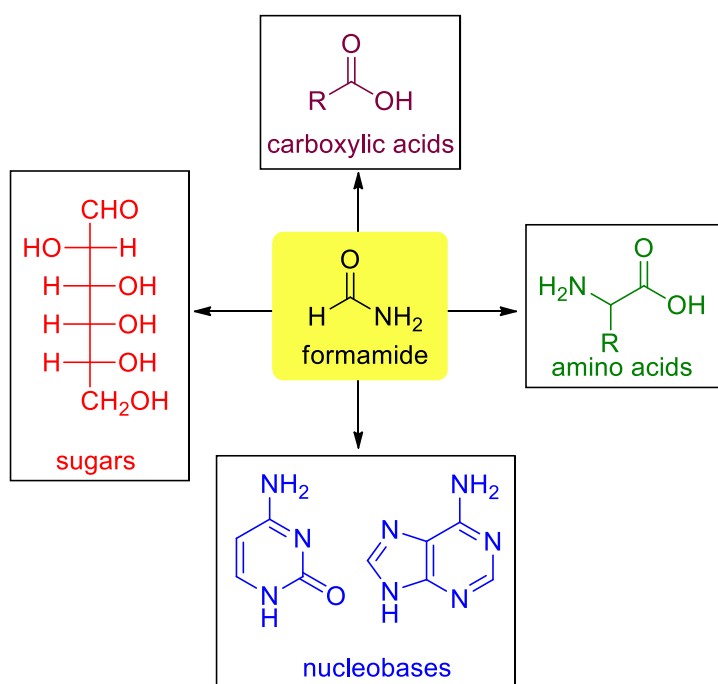


Figure 12: Syntheses of some biomolecules from formamide

Hence, both RNA and proteins might have been present and coevolved since the very first steps of life developed. Our work focus on the synthesis of peptides and thus highlights the protein-first concept.

F. Today's role of thiols in biology

Sulfur and oxygen share many chemical properties as they belong to the same group of elements in the periodic table. However, sulfur atom has a higher radius, which leads to a less electronegativity. Also, sulfur containing biomolecules are more hydrophobic than their oxygen counterparts.¹²⁰

The fact that SH is more nucleophilic than OH makes the thiol function one of the most reactive functional groups in biology. Moreover, thiols often occur as thiolates under physiological pH conditions, which make them even more reactive. Thiolates can easily chelate with different metals such as iron⁹⁸, zinc¹²¹, nickel¹²², copper¹²³, and cobalt¹²⁴. This chelating behaviour of thiols, especially in cysteine residues, might have participated in framing of the metalloproteins models that we know nowadays.

The presence of H₂S on primitive Earth support the fact that many sulfur-containing molecules were probably present in the Hadean/Archean ocean. Indeed, sulfur biomolecules have vital importance in all living organisms. They are involved, *inter alia*, in free radicals scavenging, functionality of enzymes, as well as in DNA methylation and repair.^{125,126} The diversity and the significance of active organosulfur molecules strongly suggests that sulfur-containing compounds, including cysteine and methionine residues, had an important role in the early evolutionary history of life.¹²⁷

Methionine is of crucial importance for mRNA translation. It is the initiating amino acid in the synthesis of eukaryotic and archaeal proteins, while *N*-formyl-Methionine plays this role in prokaryotes.¹²⁰

Cysteine-rich proteins are found in all living organisms and often play fundamental roles as hormones, growth factors, ion channel modulators and enzyme inhibitors in various biological pathways.¹²⁸ The similar functions that sulfur containing molecules play in plant and animal cells strongly support the hypothesis of a shared ancestor of these two kingdoms.¹²⁹

It has been proposed that primitive living organisms used elemental sulfur itself, or simple sulfur molecules, for energy metabolism (*i.e.* breathing) under oxygen-free atmosphere. This

belief is supported by the highly conserved sulfur metabolism in nowadays bacteria that can be regarded as “survivals” of ancient living organisms.¹³⁰

Sulfur metabolism in living cells

In humans, sulfur inorganic species and sulfur containing biomolecules are obtained from a normal equilibrated diet. The main sulfur-containing biomolecules active in human metabolism and their metabolism are presented in *Figure 13*.¹²⁰

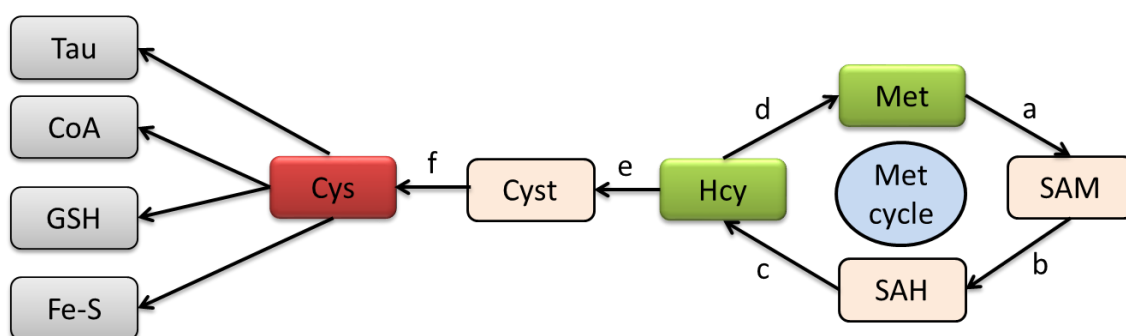


Figure 13: A simplified pathways of the main active sulfur-containing molecules in human metabolism. Met: methionine, SAM: S-adenosylmethionine, SAH: S-adenosylhomocysteine, Hcy: homocysteine, Cyst: cystathionine, Cys: cysteine, Tau: taurine, CoA: coenzyme A, GSH: glutathione, Fe-S: iron-sulfur clusters. a) methionine adenosyl transferase, b) x-methyl transferases, c) S-adenosylhomocysteine hydrolase, d) betaine:Hcy methyltransferase, e) cystathionine β -synthase, f) cystathionine γ -lyase

This Figure shows two connected pathways, the first part (right) is called the methionine cycle (or SAM cycle) in which homocysteine quantities are regulated in living systems. The other part of the figure (left) showed a simple presentation of the transsulfuration pathway in which cysteine is obtained in animal bodies.¹³¹

The methionine cycle

In the methionine cycle, methionine is converted to SAM (S-Adenosyl Methionine). SAM is a methyl donor, and it is converted to SAH (S-Adenosyl Homocysteine) after losing its methyl group. S-adenosylhomocysteine hydrolase converts SAH into homocysteine and adenosine (*Figure 14*).^{120,131}

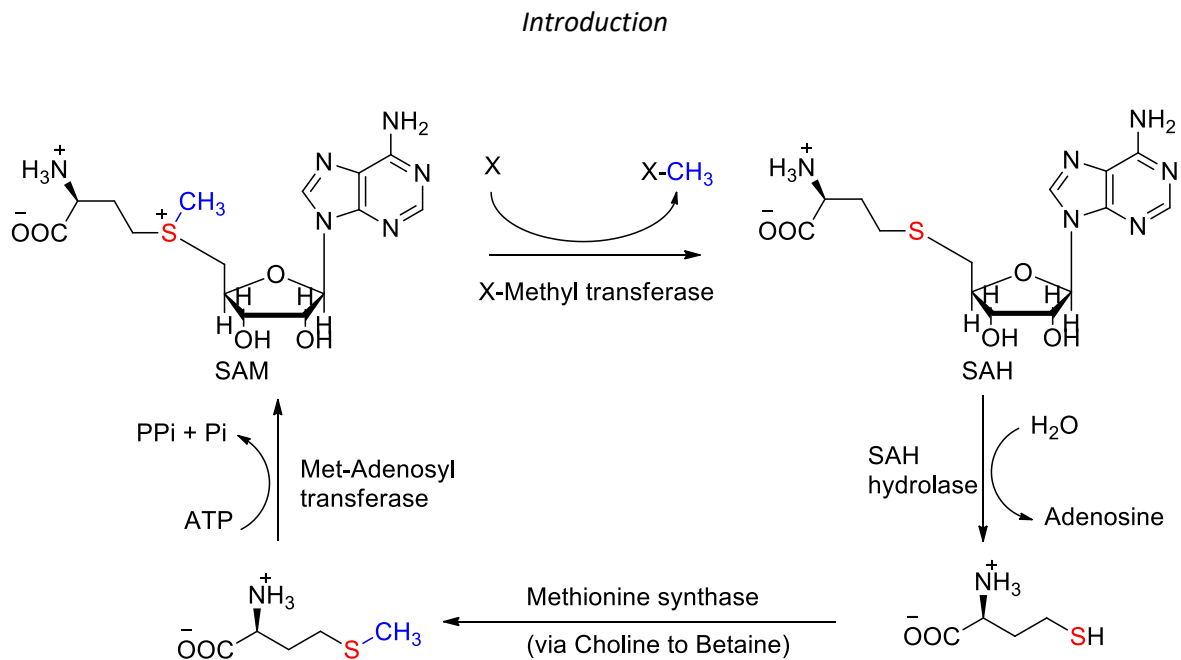


Figure 14: Summary of the methionine cycle

Homocysteine can either be recycled back to methionine through this cycle, or it can be used to form other sulfur-containing amino acids (*Figure 13*).

Biological metabolism of cysteine and homocysteine

In animals, the biosynthesis of cysteine starts with serine and homocysteine. Cystathionine β -synthase joins these two amino acids to form cystathionine. An enzyme, cystathionase, breaks down cystathionine into cysteine by releasing α -ketobutyrate and NH_3 . This process is called transsulfuration (*Figure 15*).^{120,132}

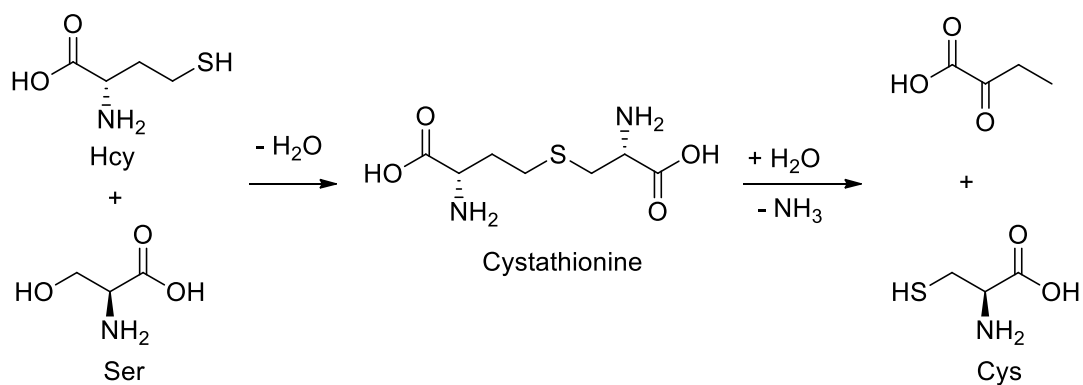


Figure 15: Transsulfuration pathway in mammals to obtain cysteine

The obtained cysteine can be used to produce other sulfur containing molecules such as CoA, taurine (Tau), glutathione (GSH, an intra-cell free radical scavenger), and iron-sulfur clusters (*Figure 13*).¹³²

The structure of **Coenzyme A** (CoA) was first revealed in 1953. CoA biosynthesis requires cysteine, vitamin B5, and adenosine triphosphate (ATP) (*Figure 16*).¹³³

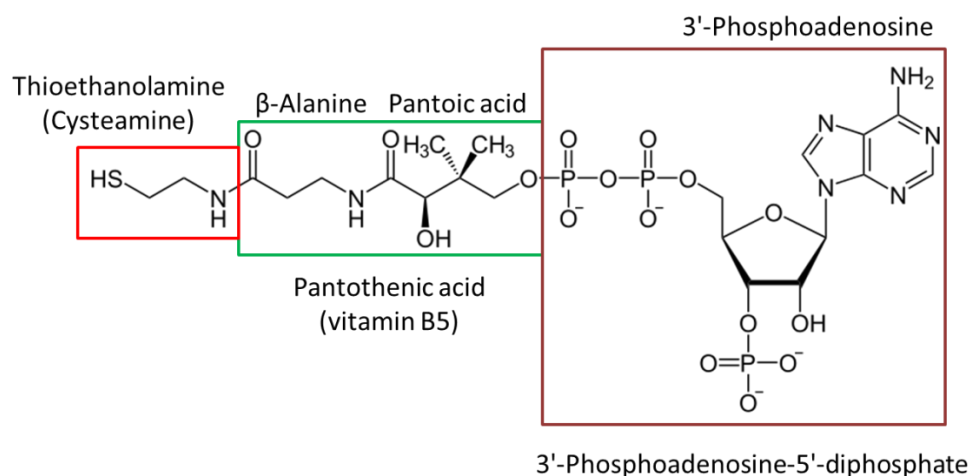


Figure 16: CoA structure

CoA plays a central role in the oxidation of pyruvate in the citric acid cycle (also known as Krebs cycle) (*Figure 17*). Aerobic organisms use this cycle to release stored energy in the form of ATP (adenosine triphosphate) via series of chemical reactions.¹³⁴

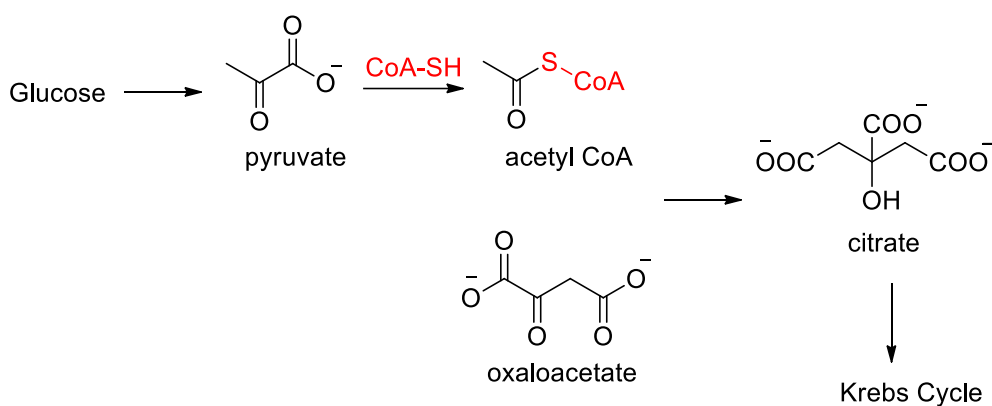


Figure 17: Simplified pathway showing the importance of CoA in Krebs cycle

Taurine is an amino sulfonic acid highly concentrated and widely distributed in animal and human tissues such as brain, retina, muscle, liver and kidney. Taurine has antioxidant activities in biological systems, besides regulation of Ca^{2+} levels (*Figure 18*).¹³⁵

Glutathione is the most abundant thiol compound in cells. It has important antioxidant activities in plants, animals, fungi, and bacteria. It binds to several toxic compounds for detoxification purposes. It also helps to maintain the suitable cell redox status (*Figure 18*).¹³⁶

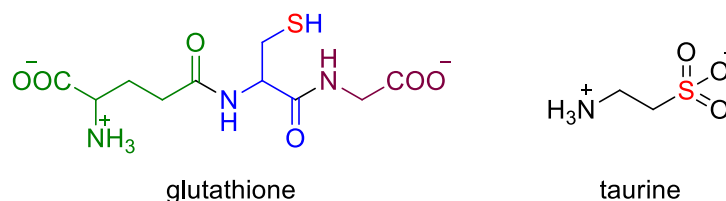


Figure 18: Taurine and glutathione structures

Iron-sulfur clusters are constructed of inorganic iron and sulfides arranged in a distinct caged structural makeup ranging from $[\text{Fe}_2\text{-S}_2]$, $[\text{Fe}_3\text{-S}_4]$, $[\text{Fe}_4\text{-S}_4]$ (up to $[\text{Fe}_8\text{-S}_8]$ clusters) (*Figure 19*). Their importance in the living cells arises from the fact that they are able to deliver electrons via the Fe^{2+} to Fe^{3+} oxidation reaction.^{101,137,138}

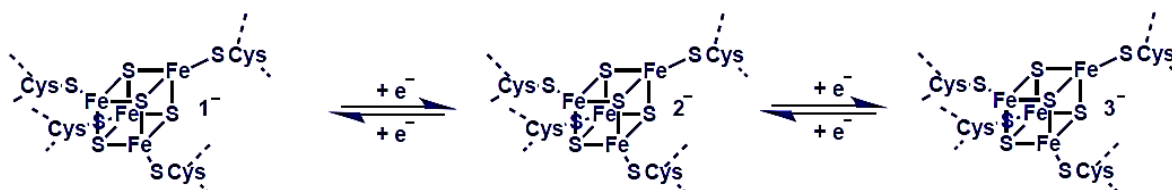


Figure 19: Iron clusters redox behaviour

The SMM cycle of flowering plants

Sulfur metabolism in plants presents some specificities. In this regard, we believe that the S-methylmethionine (SMM) cycle deserves to be considered. It was proposed that the SMM cycle contributes in the synthesis of methionine. SMM acts as a storage of methionine in plant tissues. SMM is first produced from methionine in non-seed tissues, and then transported through the phloem from leaves into seeds, where it is reconverted to methionine thanks to the SMM cycle (*Figure 20*).¹³⁹ In fact, methionine is a key metabolite in plant cells.

Introduction

It plays an essential role in the initiation of mRNA translation and the synthesis of proteins.^{139,140}

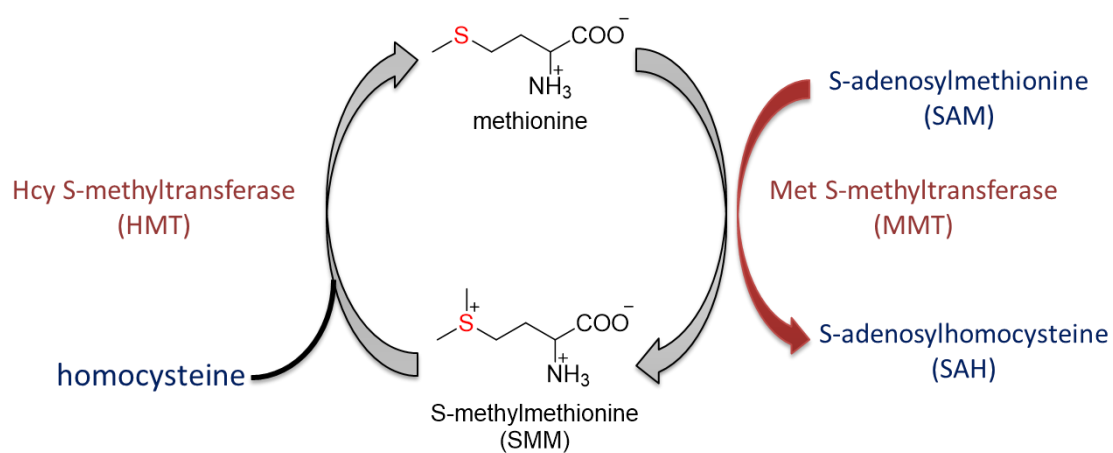


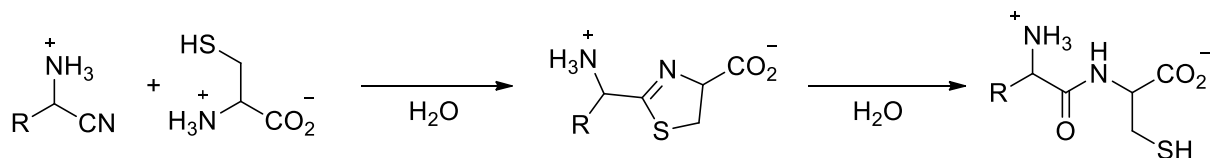
Figure 20: The SMM cycle of flowering plants

By this way, the SMM cycle helps plants to control Met, Hcy and AdoMet levels. SMM cycle is similar in both monocot and dicot plant families and it must have originated before the separation of these two families.¹⁴¹

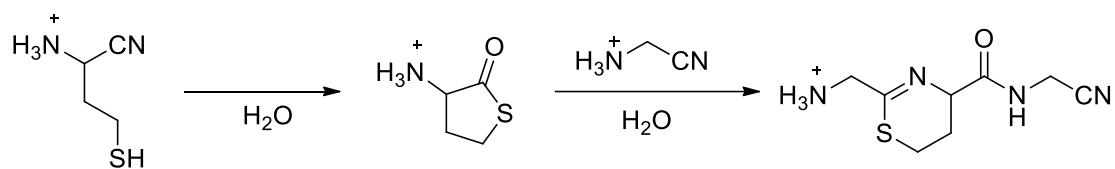
Introduction

Graphical abstract

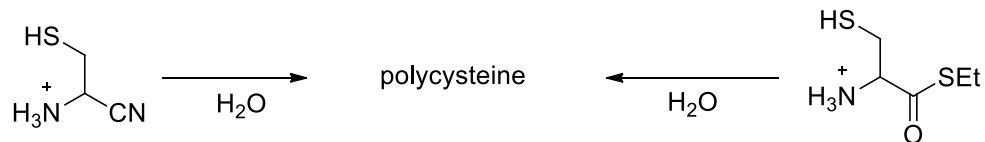
Chapter 1



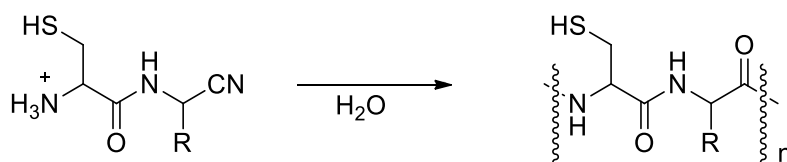
Chapter 2



Chapter 3



Chapter 4



Graphical abstract

Chapter I: Prebiotic formation of thiol containing dipeptides from aminonitriles and aminothiols

I.1 Introduction

As mentioned before, many simple aldehydes were probably present in the primitive ocean¹⁴² and they are plausible precursors for α -amino acids. Reacting with ammonia and hydrogen cyanide, they would have first give α -aminonitriles which upon hydrolysis would have delivered amino acids (*Figure 21*).¹⁴³ However, it has been reported that the reaction of nitriles with water is a slow process¹⁴⁴, so slow that, once formed in the ocean, aminonitriles would have had sufficient time to react with species more nucleophilic than water.

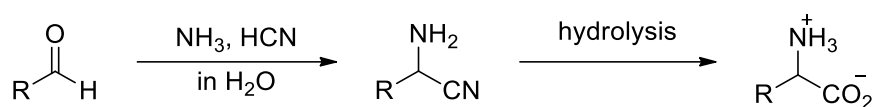


Figure 21: Strecker reaction

Nitriles are known to react with aminothiols to give thiazolines, which in turn can be hydrolysed to mercaptoamides.^{145,146} Starting from α -aminonitriles and cysteine, the expected products of this two-step process are dipeptides (*Figure 22*).

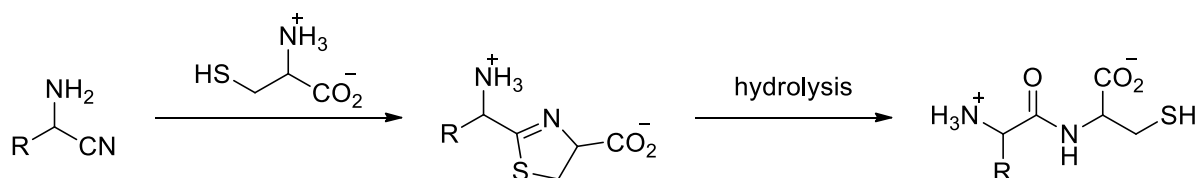


Figure 22: The reaction of aminonitriles with cysteine to obtain dipeptides

Compared to any activation process starting from acids, this nitrile scenario has the advantage of not necessitating any strong energy source. The acid does not need to be activated as it is delivered directly in an activated form by the Strecker reaction.

This chapter focuses on the synthesis and reactivity of different types of aminonitriles toward aminothiols, under plausible prebiotic conditions. In the first part of this chapter, we discuss the synthesis of aminonitriles that we used in reaction with aminothiols (cysteine, homocysteine and cysteine amide). The second part of the chapter studies the reactivity of these aminonitriles toward thiazoline formation. It includes some competitive reactions

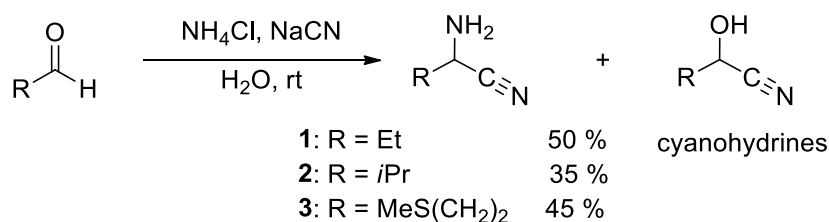
between different starting materials. The formation of active thioesters, nitrogen heteroaromatics, and other by-products of our reactions are also discussed. Some of these results are supported by theoretical calculations.

I.2 Synthesis of aminonitriles

Some nitriles that we used are commercially available, some are not. Their synthesis is described here.

I.2.1 Synthesis of racemic α -aminonitriles

We used Strecker synthesis to obtain racemic aminonitriles. In 1988, Denis and Guillemin used this reaction to obtain several aminonitriles starting from the related aldehydes.¹⁴⁷ We followed their procedure in order to obtain racemic 2-aminobutanenitrile **1**, Val-CN **2** and Met-CN **3** (*Scheme 1*).

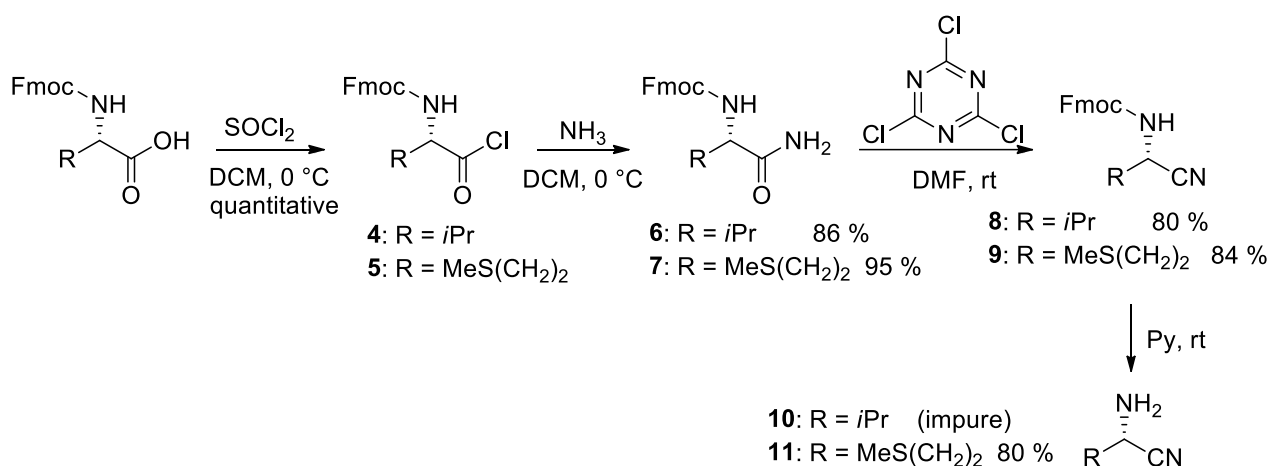


Scheme 1

The common by-products of this method were cyanohydrins. Denis and Guillemin have separated their aminonitriles by distillation. We used the same method. However, for reactions with smaller quantities we used column chromatography as another efficient method to recover the products.

I.2.2 Synthesis of L-aminonitriles

L-aminonitriles were obtained starting from the corresponding Fmoc protected L-amino acids. We activated the acids as acid chlorides using thionyl chloride, SOCl_2 . Then, we introduced ammonia gas to provide the corresponding amides. By using cyanuric chloride, we were able to dehydrate amides into nitriles. Fmoc group was finally removed using piperidine to obtain the required L-aminonitriles (*Scheme 2*).

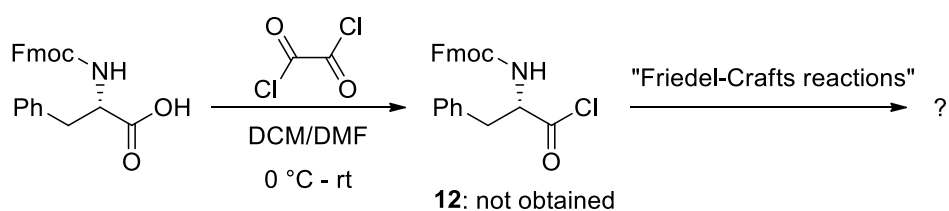


Scheme 2

In the last step, L-Methionine nitrile **11** was obtained upon deprotection. However, when we tried to deprotect Fmoc-L-Valine **8**, the resulting L-Valine nitrile **10** was found to be unstable and mixture of products has been obtained. As a consequence, we used only racemic Val-CN **2** in our reactions.

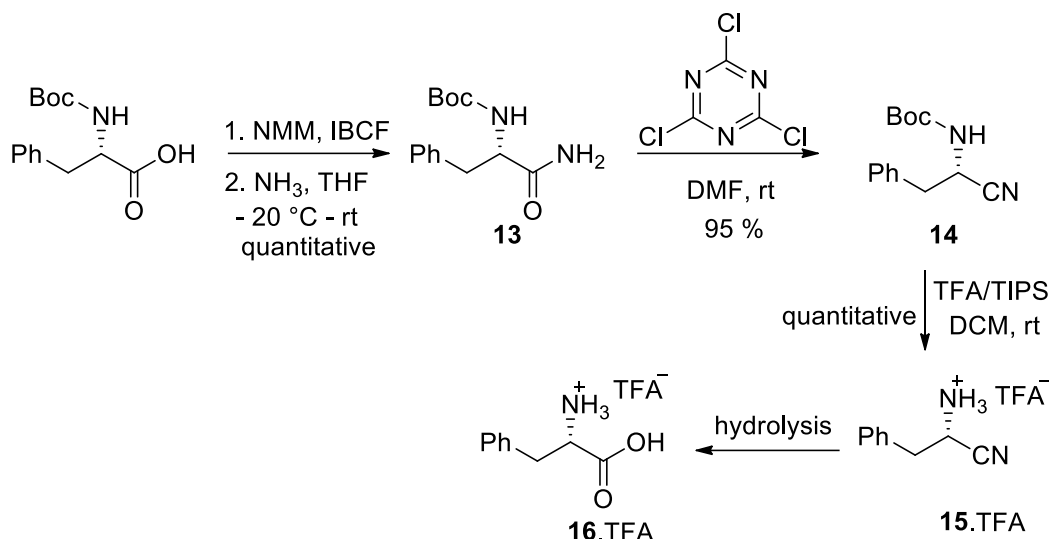
I.2.1.a Synthesis of L-Phe-CN **15**

On the other hand, we did not succeed to obtain the acid chloride of phenylalanine. When we started from Fmoc-Phe-OH, we found that the corresponding acid chloride **12** was not stable. We think that Friedel-Crafts by-products were formed. Similar results were obtained when we used oxalyl chloride, (COCl)₂, as an alternative to thionyl chloride (Scheme 3).



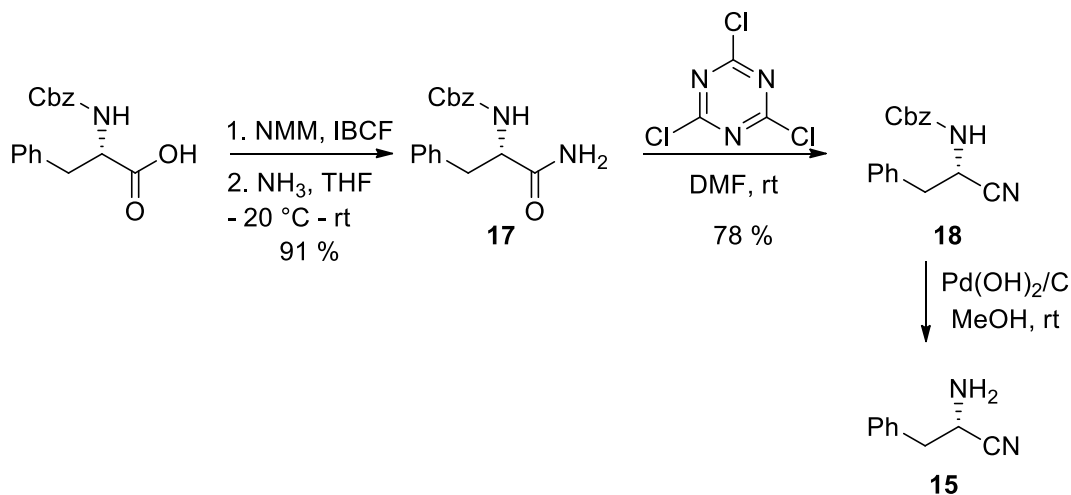
Scheme 3

Strangely, when Fmoc protecting group was replaced by Boc, we obtained the good product. In this case, we succeeded to obtain Boc-L-Phe-NH₂ **13** and Boc-L-Phe-CN **14** successively. However, this nitrile **15** was not stable under acidic conditions and it hydrolysed in the last deprotection step to form the corresponding acid, L-Phe-OH **16** (Scheme 4).



Scheme 4

Still trying to obtain the L-Phe-CN **15**, Cbz protected L-Phe was attempted. The acid was activated as a mixed anhydride, ammonia was introduced to form the amide **17**, and cyanuric chloride was used to provide the nitrile **18**. However, the last step was not successful and a mixture of products was produced in this case (Scheme 5).

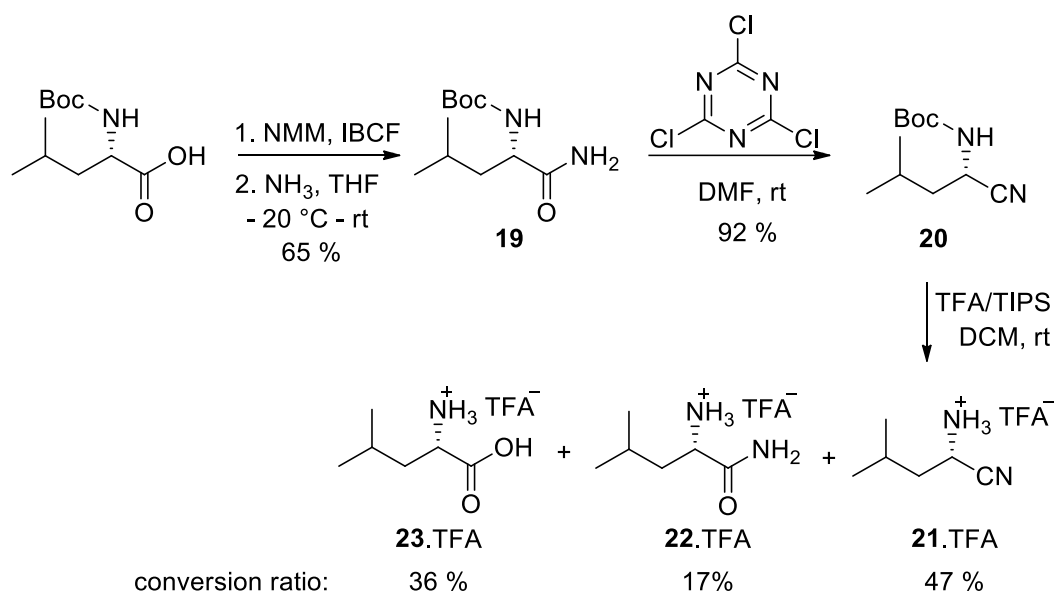


Scheme 5

I.2.1.b Synthesis of L-Leu-CN **21**

The observed stability of the phenylalanine nitrile under acidic conditions, in the previous example, makes the stability of similar simple aminonitriles questionable. In order to study the stability of these nitriles in acidic media, we decided to synthesize some of them in a

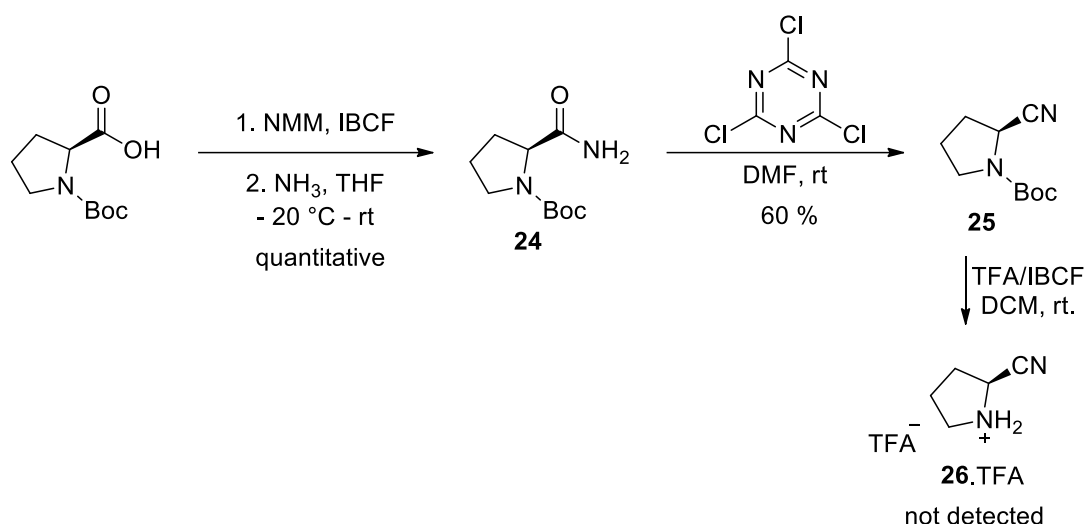
similar way starting from Boc protected L-amino acids. As an example, L-Leu-CN **21** was obtained. We found that it is only poorly stable under similar conditions. The last step was followed by extraction in D₂O. The ¹H-NMR showed a mixture of the nitrile **21**, the acid **22** and the amide (*Scheme 6*).



Scheme 6

I.2.1.c Synthesis of L-Pro-CN **26**

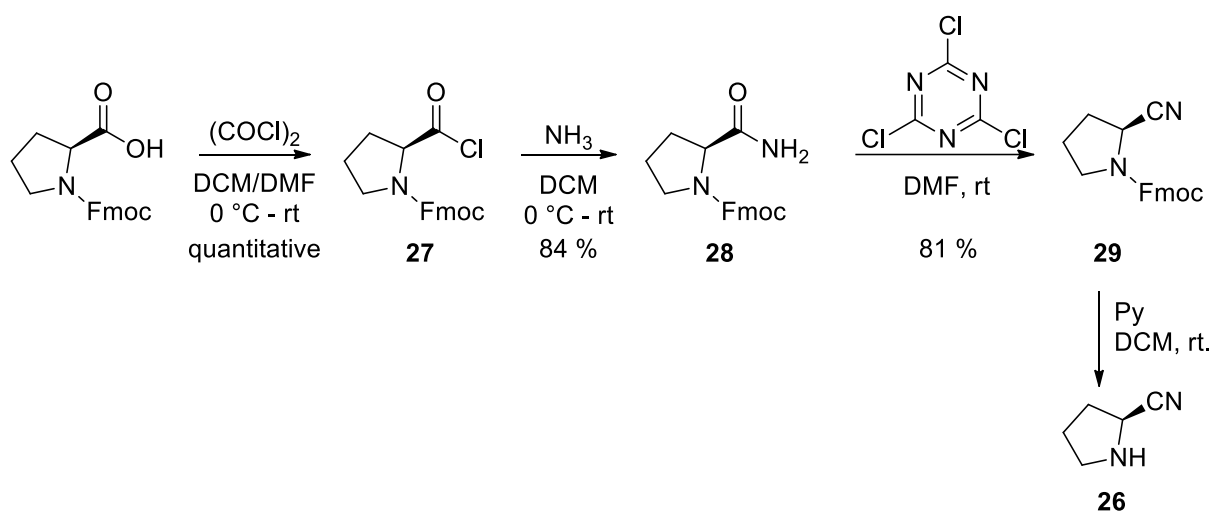
We also tried to synthesize the aminonitrile of proline, an example to study the stability of an amino nitrile with a secondary amine. We started this synthesis from Boc-L-Proline, we activated the acid as mixed anhydride, introduced ammonia to provide the amide **24**, converted the amide into nitrile **25** using cyanuric chloride, and finally deprotected the amine with TFA (*Scheme 7*). However, this nitrile **26** could not be isolated and a mixture of unidentified products was produced in this case.



Scheme 7

The instability of leucine and phenylalanine nitriles demonstrates that at least some simple aminonitriles (aa-CN) are poorly stable in strong acidic conditions. As we will see later, longer peptidic chains ending with nitriles (aa-(aa)_n-CN) are more stable under similar acidic conditions.

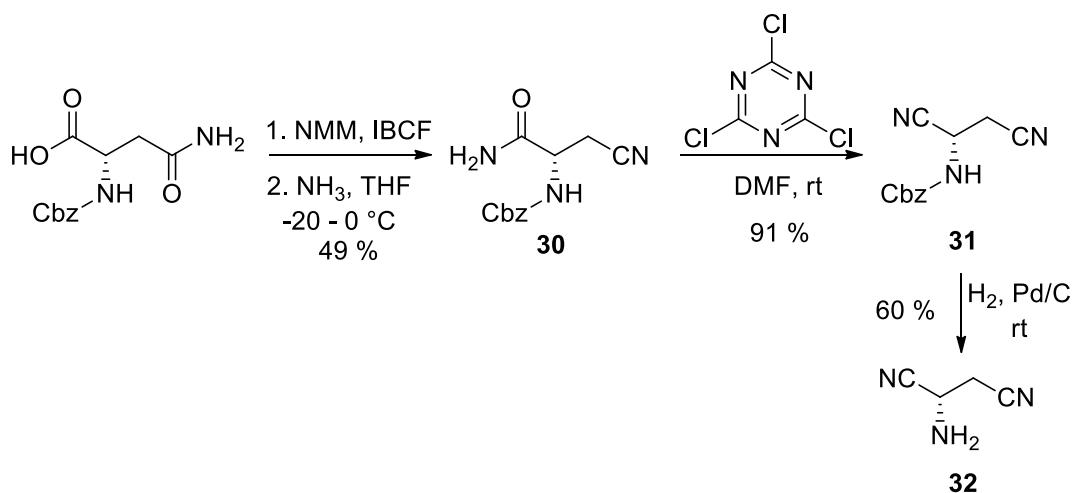
Despite that we could not obtain the nitrile starting from the Boc protected proline, we thought to attempt Fmoc as a protecting group of proline's amine as this protecting group can be removed under basic conditions to give the required nitrile **26**. The acid was activated as an acid chloride **27**, followed by amide formation **28**. The corresponding nitrile **29** was obtained using the cyanuric chloride. However, we were not able to isolate the product **26** in pure form when the protected nitrile was treated with piperidine (*Scheme 8*).



Scheme 8

The deprotection step was repeated using Et_3N , and by introducing ammonia gas in DCM. However, none of these methods was successful. Finally, the last step worked when we used methanol as a solvent and introduced ammonia to the solution to obtain Pro-CN **26**. However, this nitrile was not stable when passing through a silica column and thus it was difficult to purify it. Another idea for purification was to obtain the ammonium chloride salt by introducing HCl gas, but we found that the nitrile **26** did not survive such acidic condition and hydrolysed into the acid.

I.2.2 Synthesis of (S)-2-amino-1,4-butanedinitrile **32**



Scheme 9

In order to study the relative reactivity of α and β -aminonitriles, we decided to synthesize an α,β -aminonitrile, a compound having both an α -aminonitrile and a β -aminonitrile. The bis-nitrile **32** has been obtained starting from a protected asparagine (AsnCbz). We first tried the transformation of the carboxylic acid into amide. However, in this case, the obtained product was the amido-nitrile **30**. This unexpected nitrile formation in the first step can be explained by the following mechanism (Figure 23).¹⁴⁸ Indeed, this was lucky, as it shortens our synthesis.

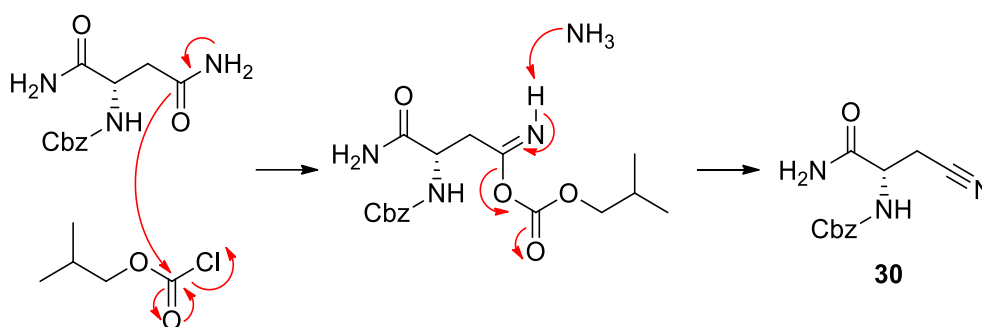
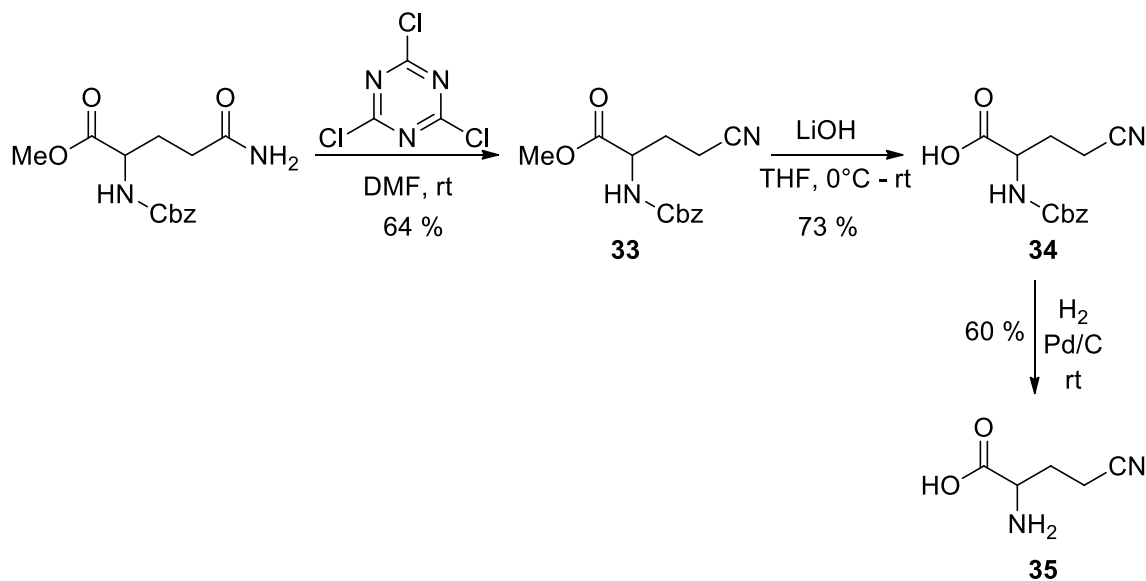


Figure 23: A reported mechanism for this unexpected nitrile formation

Cyanuric chloride was used to form the second nitrile group **31**. Finally, Cbz group was removed by hydrogenation (Scheme 9).

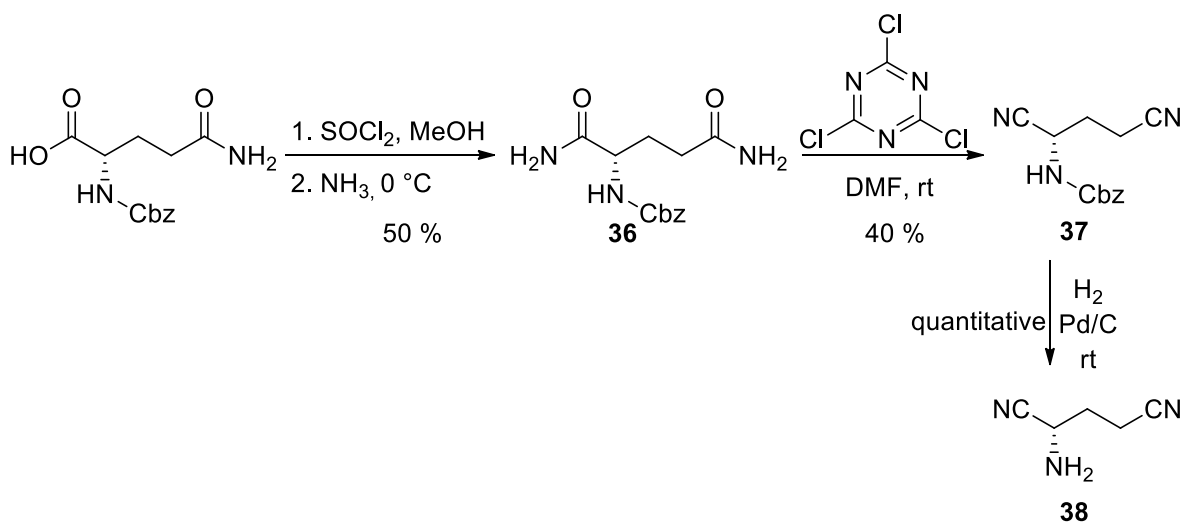
I.2.3 Synthesis of 2-amino-4-cyanobutanoic acid **35**



Scheme 10

γ -aminonitrile **35** was synthesized starting from protected glutamine, MeOGlnCbz. The amide was first transformed into the corresponding nitrile **33** using cyanuric chloride. Then, LiOH was used for the saponification of the ester into the carboxylic acid **34**, and finally, Cbz group was deprotected by hydrogenation to obtain the required nitrile **35** (*Scheme 10*).

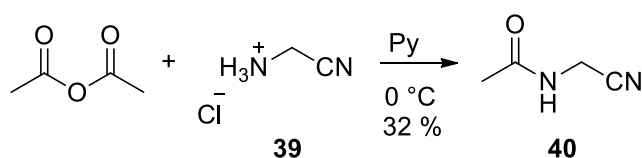
I.2.4 Synthesis of (S)-2-aminopentanedinitrile **38**



Scheme 11

An α,γ -amino bisnitrile, a bisnitrile which have both an α -aminonitrile and a γ -aminonitrile, was synthesized in order to study the competition between α and γ -amino nitriles. We have followed a previously described method to obtain the required bisnitrile **38**.¹⁴⁹ It was synthesized in 3 steps. Starting from GlnCbz, the acid was activated by SOCl₂ and then ammonia was introduced to form the amide **36**. Then, cyanuric chloride was used to form the bisnitrile **37** and the product **38** was obtained in the last step upon the deprotection of the Cbz group (*Scheme 11*).

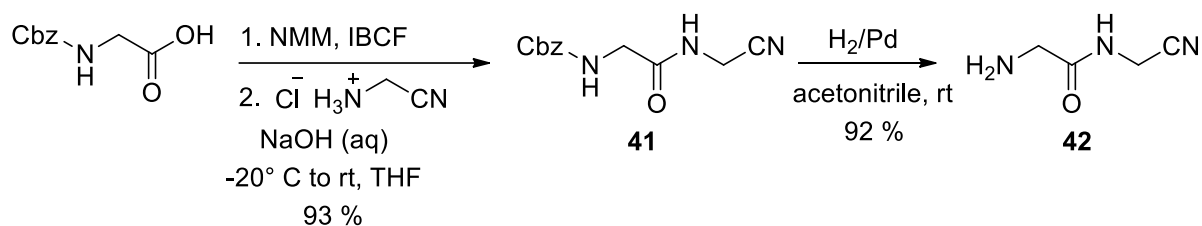
I.2.5 Synthesis of N-acetyl- α -aminonitrile **40**



Scheme 12

N-acetyl- α -aminonitrile is a good candidate as a simple model which will allow us to check the reactivity of a peptide chain ending with a nitrile. The required amino nitrile **40** was simply obtained starting from acetic anhydride and aminoacetonitrile **39** in pyridine at 0 °C for one day (*Scheme 12*).¹⁵⁰

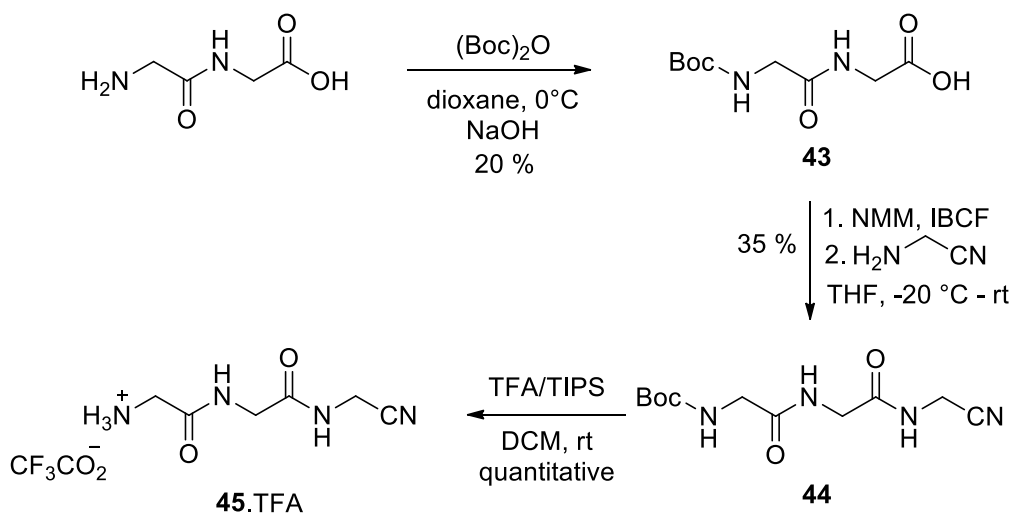
I.2.6 Synthesis of Gly-Gly-CN **42**



Scheme 13

We followed a previously described method to obtain Cbz-Gly-CN **41** starting from Cbz protected glycine in a good yield.¹⁵¹ Then, the Cbz group was removed by hydrogenation to give product **42** (*Scheme 13*).

I.2.7 Synthesis of Gly-Gly-Gly-CN **45**



Scheme 14

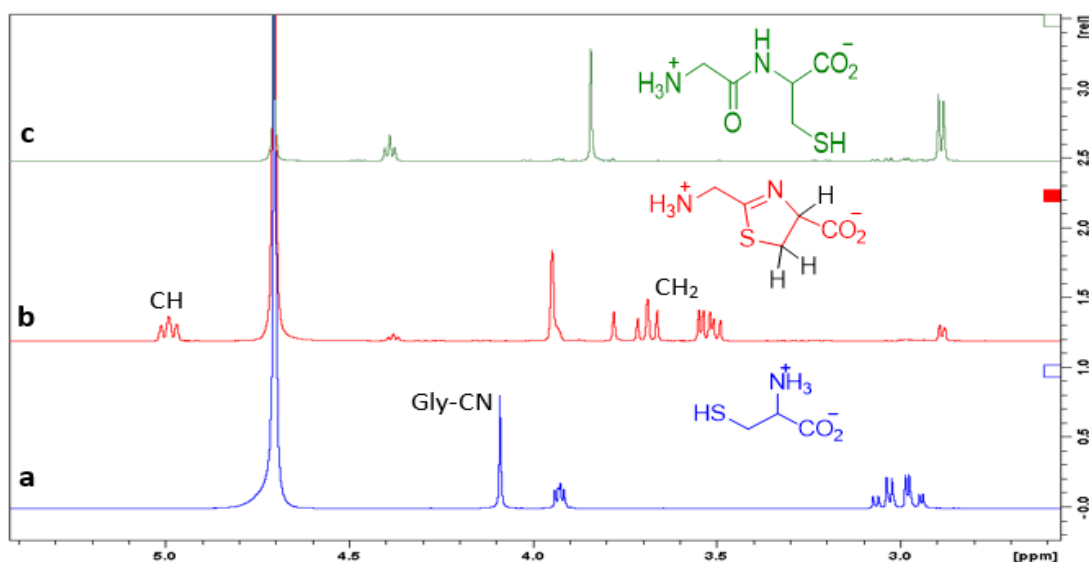
We started the synthesis from the commercially available diglycine. We protected this dipeptide with Boc group to give **43**.¹⁵² Then we activated the acid like in previous synthesis

and added aminoacetonitrile **39** to give the nitrile **44**.¹⁵¹ Finally, the Boc group was removed in acidic condition to give Gly-Gly-Gly-CN **45**, as its trifluoroacetate salt (*Scheme 14*).

I.3 Reactions between aminonitriles and aminothiols

I.3.1 Reaction of aminoacetonitrile with cysteine

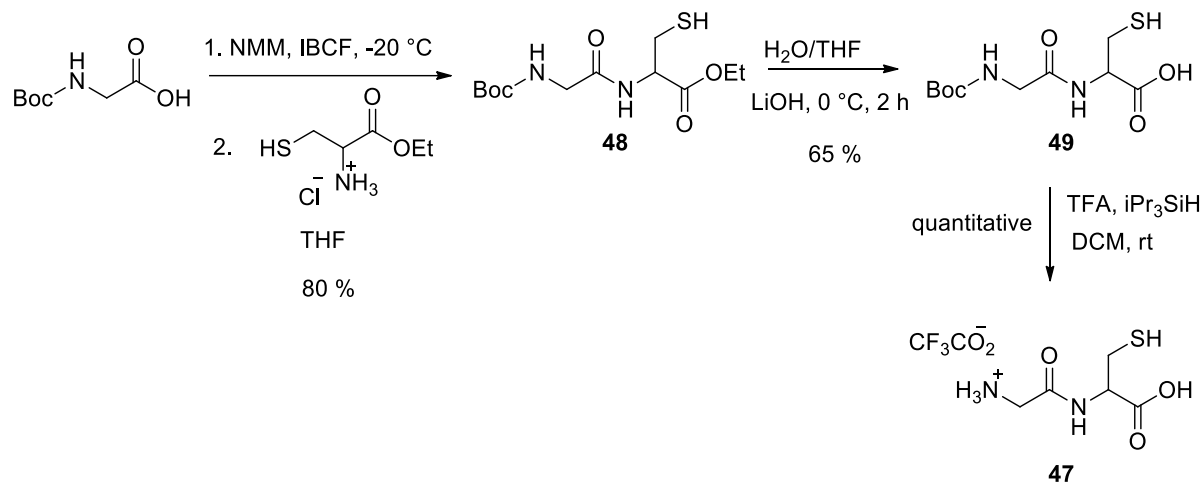
We first studied the reaction of aminoacetonitrile (Gly-CN **39**) with cysteine.¹⁵³ This reaction was conducted in D₂O solution and followed by NMR spectroscopy. Formation of the expected thiazoline ring **46** was evidenced by the apparition of signals at *ca.* 5 ppm (a triplet-like dd, H α to CO₂), and from 3.4 to 3.7 ppm (2 dd, SCH₂). After some time, new signals grew, including a triplet at 4.4 ppm (H α to C(O)N) and a thin doublet like signal at *ca.* 2.9 ppm (SCH₂), both characteristic of Gly-Cys **47** (*Figure 24*).



*Figure 24: Reaction of cysteine with aminoacetonitrile, Gly-CN **39** at rt. a) mixture of starting materials, b) mostly thiazoline ring **46**, c) Gly-Cys dipeptide **47**. Conditions: room temperature, pH 6.5, concentration 10^{-2} mol/L*

In order to confirm the structure of the obtained dipeptide **47**, an authentic sample was prepared and added to the reaction mixture. This synthesis started with activation of Boc protected glycine as a mixed anhydride, followed by adding cysteine ethyl ester in order to give Boc-Gly-Cys-OEt **48**. The ester group in **48** was converted into the corresponding acid **49** via saponification reaction. Finally, Trifluoroacetic acid (TFA) was added to **49** in order to remove the Boc group, giving the desired molecule Gly-Cys **47** (as a TFA salt) (*Scheme 15*).

Chapter I



Scheme 15: Synthesis of an authentic sample of Gly-Cys dipeptide 47

The evolution of this reaction was followed by ¹H-NMR for 15 hours in which we observe the declination of the starting materials and the evolution of the products (*Figure 25*).¹³⁸

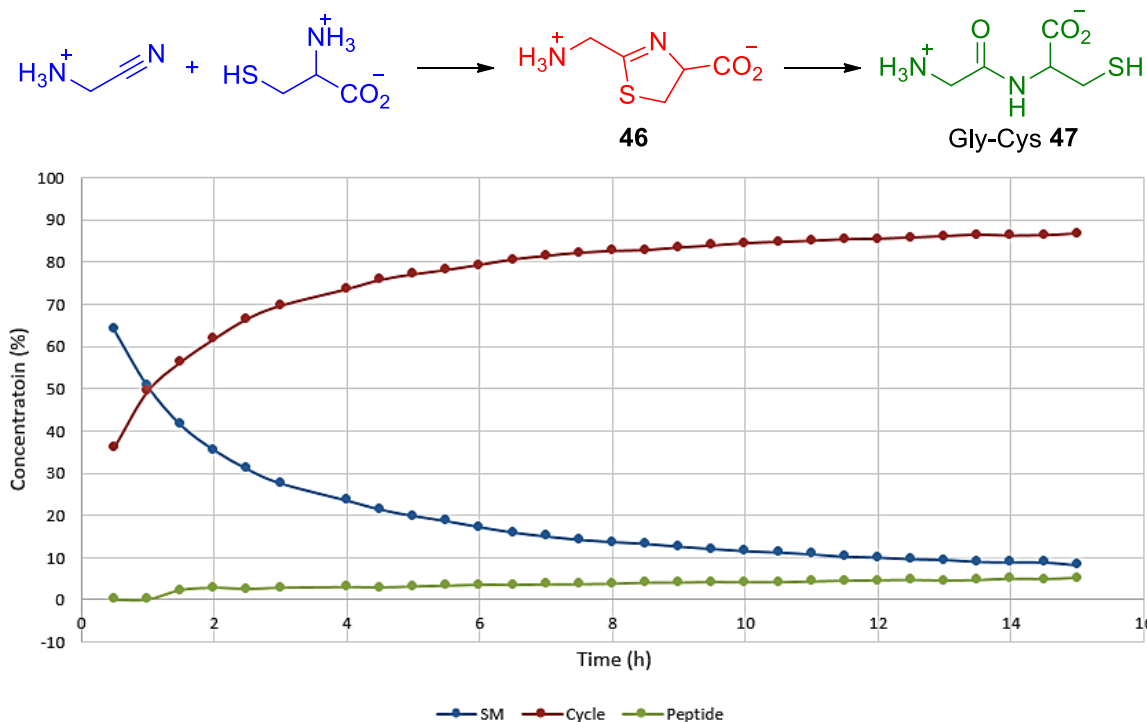


Figure 25: "Prebiotic" formation of Gly-Cys 47 at 24 °C (pH 6) followed by ¹H-NMR

The probable slight acidity of primitive oceans supports our work because acidity can favour Pinner reaction (*Figure 26*).¹⁵⁴ In this reaction, nitrile reacts with an alcohol (or a thiol) to give an imino-ether (or an imino-thioether).

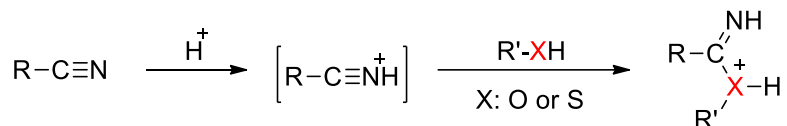


Figure 26: Pinner reaction

In the case of an aminonitrile reacting with cysteine, the mechanism starts with thiol attack on the nitrile (i.e. Pinner reaction). Then, the amine adds onto the obtained iminium double bond in **50** to form a non-observed thiazolidine **51**, which upon NH_3 loss gives the more stable thiazoline cycle **46**. This thiazoline is thus slowly hydrolysed to give the dipeptide **47** (*Figure 27*).

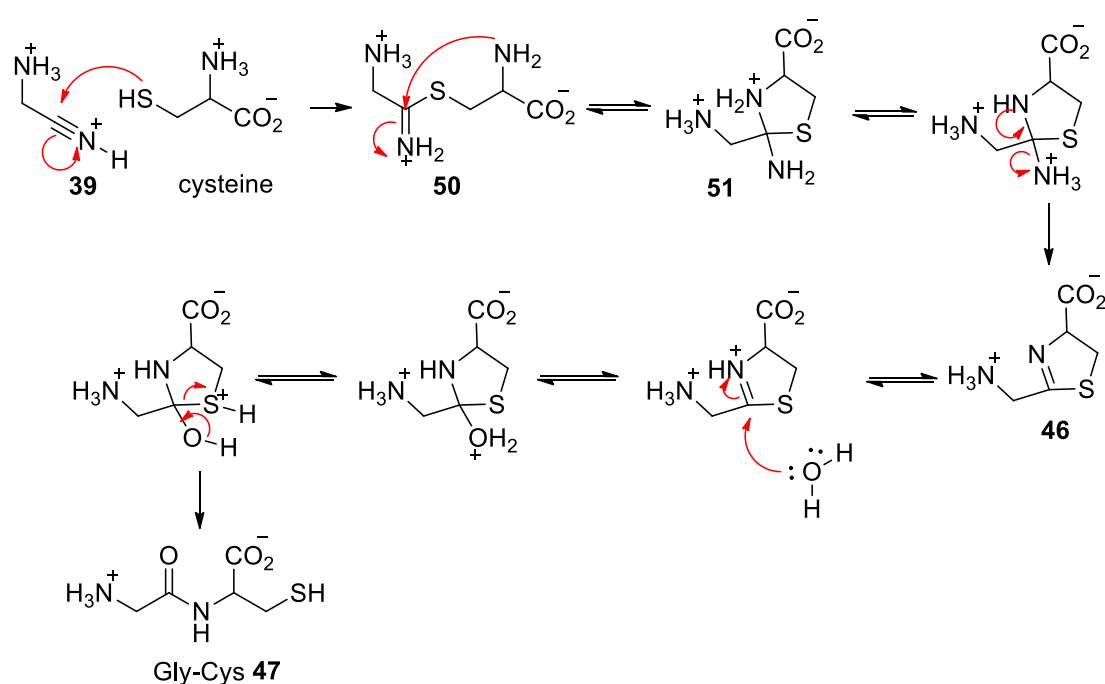


Figure 27: Proposed mechanism of the formation of Gly-Cys dipeptide 47 in our mixture

In addition, we performed this reaction in H_2O and followed it by mass spectrometry. This technique confirmed the formation of both the thiazoline ring **46** and Gly-Cys **47** (*Figure 28*).

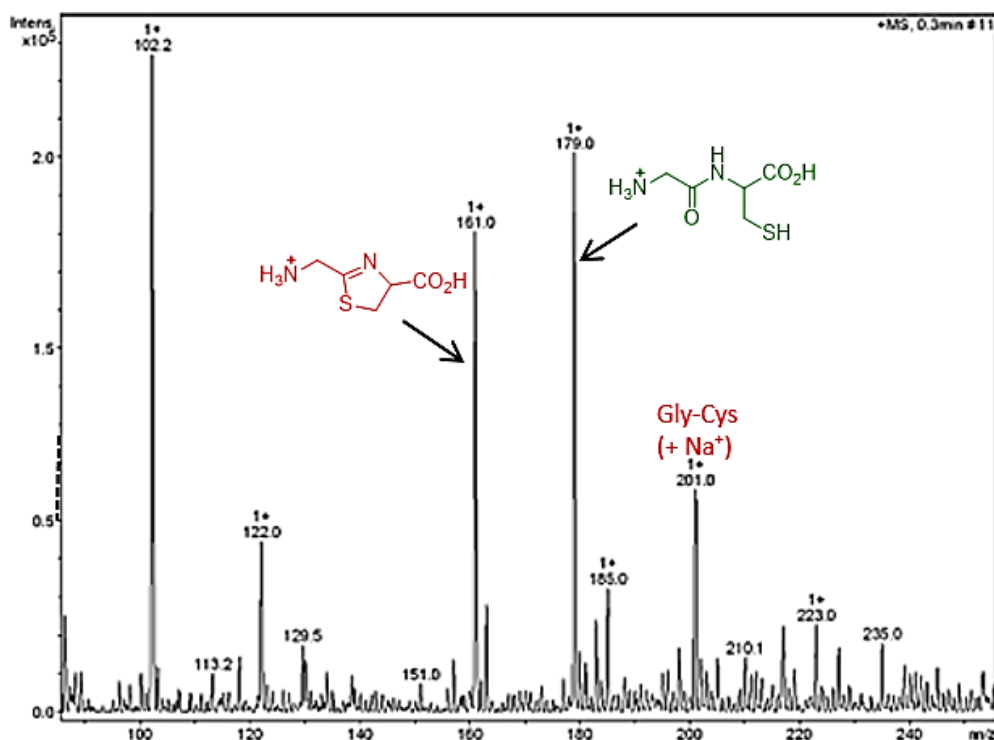


Figure 28: Reaction of cysteine with Gly-CN in H₂O; ESI spectrometry of the reaction mixture confirms the thiazoline and dipeptide formation

We observed the formation of the thiazoline **46** and Gly-Cys **47** at concentrations as low as 5×10^{-4} mol. L⁻¹. We have repeated the experiment in different temperatures and pH's and we have found that the process is quicker at higher temperatures but goes well also at room temperature. The ring formation is quicker under basic conditions. We believe that under such conditions the thiol function is deprotonated giving the more nucleophilic thiolate species. Under acidic condition the nucleophilic species is probably the thiol itself. We also noticed that the hydrolysis step is quicker under acidic conditions. This probably implies that the thiazoline ring is activated through protonation step before H₂O addition.

Although, these reactions were applicable under a wide range of pH's (4 – 9) and temperatures (20 – 70 °C), we (in most cases) choose 45 °C and a pH at *ca.* 5.5 – 6.5 as standard conditions for the following reactions, taking into account the primitive conditions that have been discussed in the introduction. The evolution of Gly-Cys formation at 45 °C and different pH's were followed by ¹H-NMR and is presented in (*Figure 29*). The figure shows that the formation of the cycle **46** is faster under basic conditions. These conditions gives more stability to the cycle **46**. The hydrolysis of this cycle is faster under more acidic conditions.

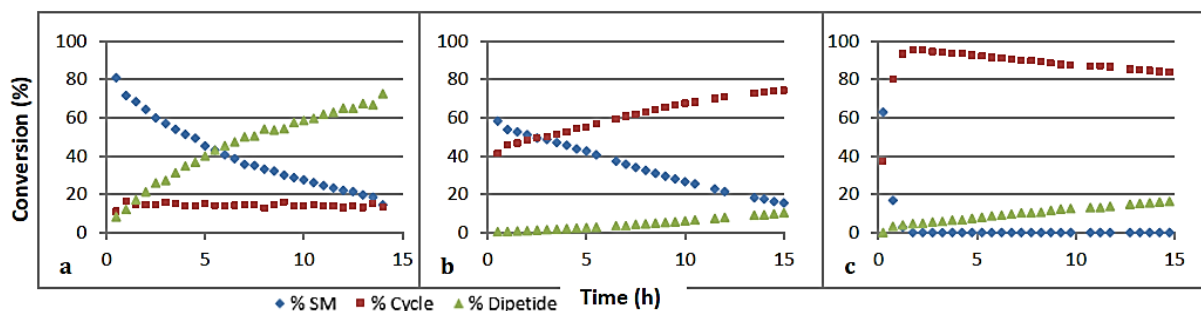


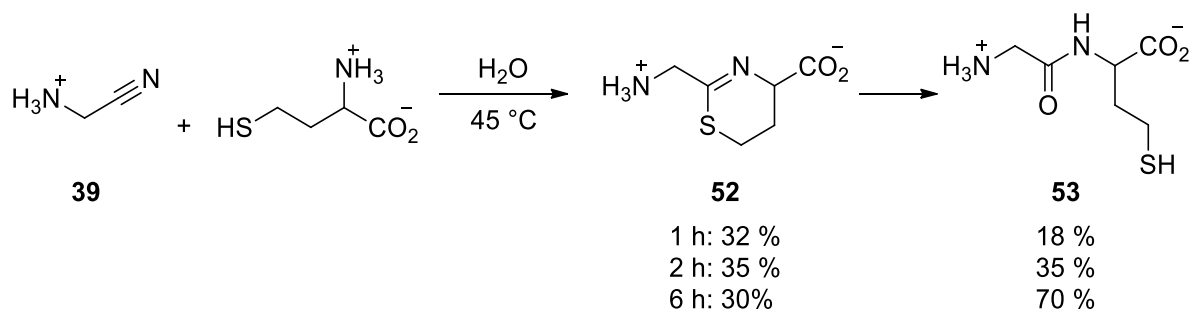
Figure 29: Evolution of a mixture of GlyCN **39** and cysteine in D₂O at 45 °C followed by ¹H-NMR, at various pH's: a) pH 4, b) pH 6, c) pH 8. SM: starting materials. Concentration 4×10^{-2} mol/L

We did not observe reactions between aminoacetonitrile and any other proteinogenic amino acid that we tested, including serine. In this case, it appears that the alcohol function of serine is not nucleophilic enough to attack the CN triple bond. Therefore, the presence of a thiol function is mandatory in our reactions.

Finally, in order to study the effect that minerals might have on these reactions, we perform this reaction in the presence of Zn (II), which is mandatory for the formation of zinc fingers. We found that addition of ZnCl₂ had no effect on the formation of the thiazoline ring **46** or the Gly-Cys dipeptide **47**.

I.3.2 Reaction of aminoacetonitrile with homocysteine

As it contains both thiol and amine functions, we expected homocysteine to react similarly to cysteine. Indeed, aminoacetonitrile **39** reacted with homocysteine in a similar way to that observed with cysteine. In this case, the intermediate is a six-membered ring **52**. The hydrolysis of this dihydrothiazine leads to the final product, Gly-Hcy dipeptide **53** (*scheme 16*).



Scheme 16

We performed the reaction in D₂O and followed it with ¹H-NMR (Figure 30). We found that homocysteine reacts slightly slower than cysteine. We believe that this is due to the fact that the formation of a five-membered ring, in the case of cysteine, is kinetically more favoured than the formation of a six-membered ring. Despite this remark, the difference was not much and it can be considered that both cysteine and homocysteine reacts in similar rates.

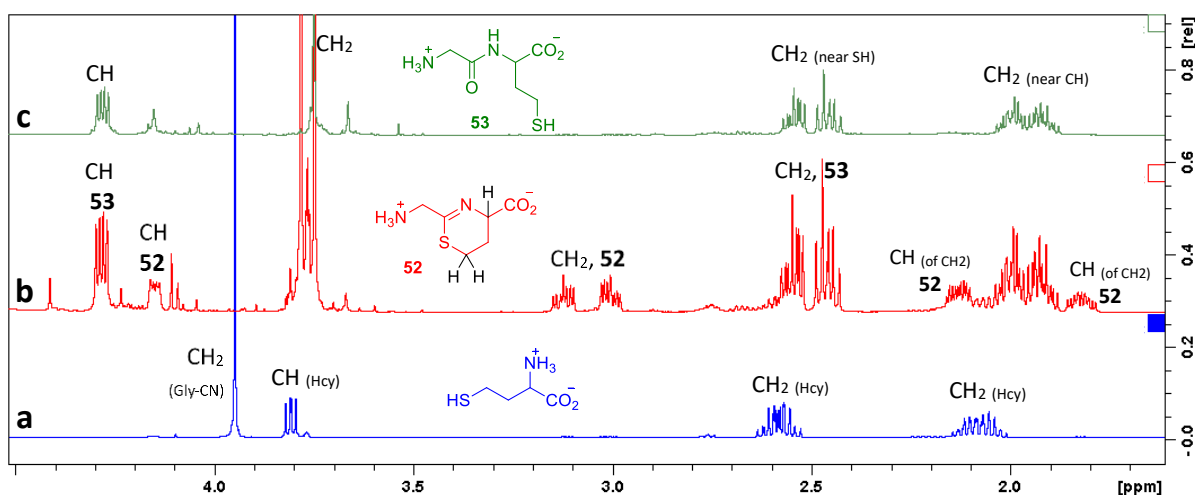
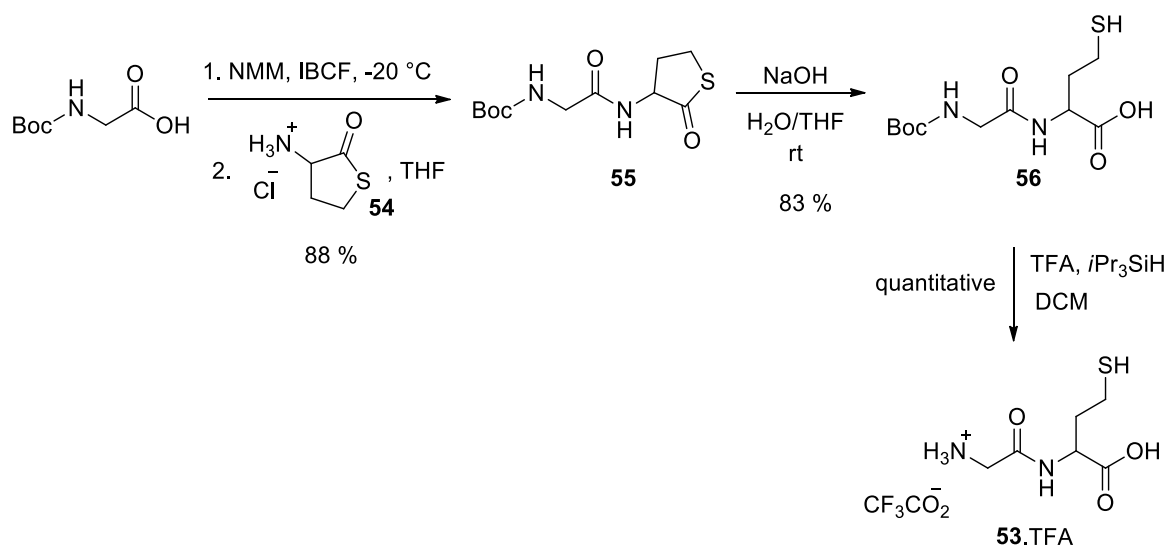


Figure 30: Reaction of homocysteine with Gly-CN **39** at 45 °C, pH = 6.5, Conc. = 10⁻² mol/L, followed by ¹H-NMR; a) starting mixture, b) reaction after 6 hours contains the cycle **52** and the dipeptide **53**, and c) Gly-Hcy **53** at the end of the reaction

Similarly to Gly-Cys **47**, we prepared an authentic sample of Gly-Hcy dipeptide **53**. We activated Boc protected glycine as a mixed anhydride and Hcy-thiolactone **54** was added in the first step to give the thiolactone **55**. In the second step the thiolactone **55** was opened by using NaOH to give Boc-Gly-Hcy **56**. Finally, Boc group was removed under acidic condition in the last step to give **53** (as a TFA salt) (Scheme 17).

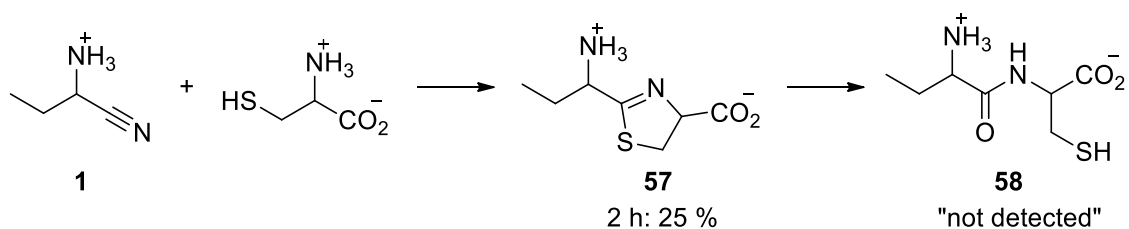


Scheme 17

Adding this sample to our “prebiotic” mixture confirmed the formation of Gly-Hcy **53**.

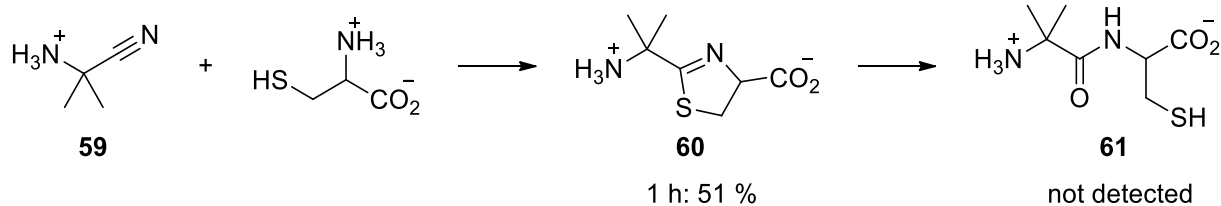
I.3.3 Reaction of alpha substituted aminonitriles with cysteine and homocysteine

We chose to study the substitution effect of aminoacetonitrile (in the alpha position) on our reactions. We first tried a reaction between α -aminobutanenitrile **1** and cysteine and we observed the formation of the thiazoline ring **57** on the $^1\text{H-NMR}$ spectrum (Scheme 18).



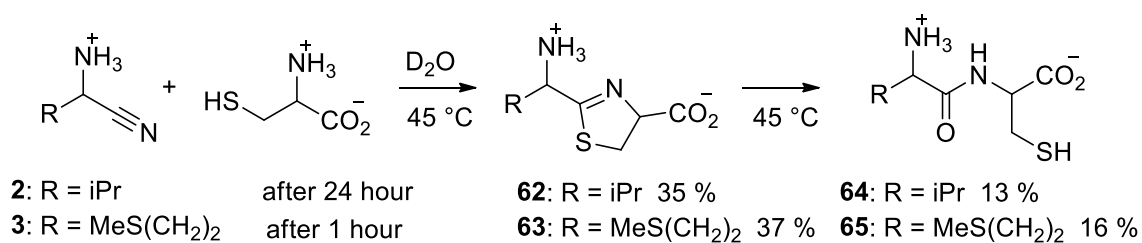
Scheme 18

We also tried the α -aminonitrile **59** in which two methyl groups are located on the aminoacetonitrile’s methylene, and observed the formation of the corresponding thiazoline **60** (Scheme 19).



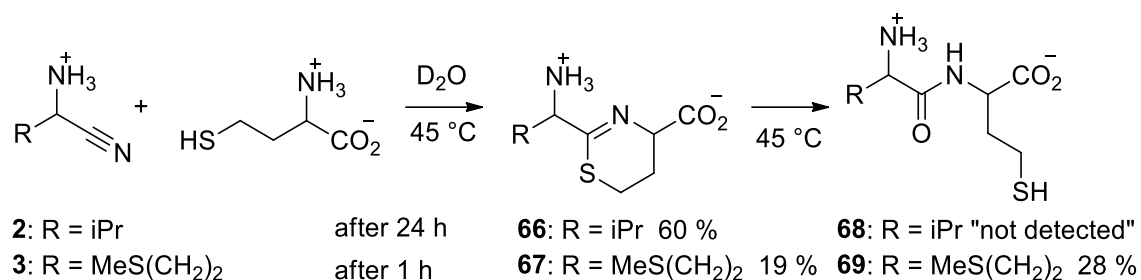
Scheme 19

However, the obtained dipeptides in the last two examples **58** and **61** are not composed of completely “proteinogenic” amino acids, and we thought that trying to obtain peptides consisting of proteinogenic amino acids would be more interesting to us. In order to obtain a wider library of “proteinogenic amino acid”-containing dipeptides, we introduced relevant aminonitriles to react with cysteine, and by this way, several aa-Cys examples were obtained. We followed these reactions by $^1\text{H-NMR}$ without purification. Furthermore, the expected products were detected in mass spectrometry of the mixtures when H_2O was used as a solvent. We chose to use the nitriles derived from methionine and valine. Both of these nitriles reacted with cysteine, and the expected thiazolines and the dipeptides were identified (Scheme 20).



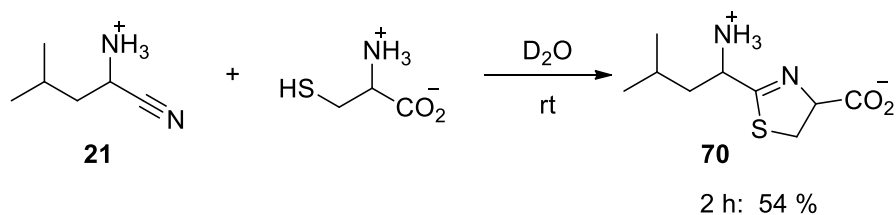
Scheme 20

Valine and methionine nitriles reacted in a similar way with homocysteine. In this case, the expected six-membered rings were obtained. Similarly to thiazolines, these six-membered rings are more stable in basic conditions (pH 7 – 8) and they hydrolyse faster in lower pH's, and/or upon heating, to give the corresponding dipeptides (Scheme 21).



Scheme 21

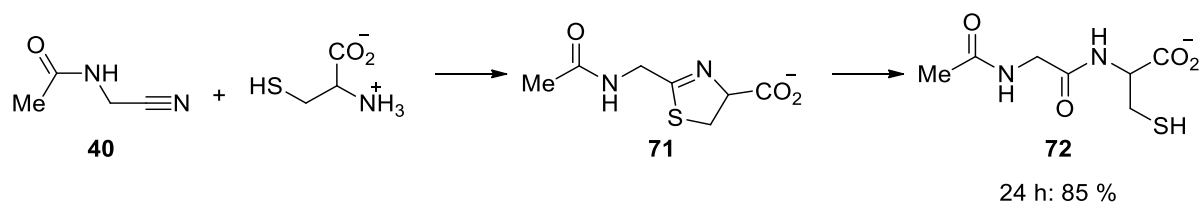
As we have previously discussed in the first part of this chapter, some of our nitriles were not stable under acidic conditions. One of these nitriles is L-Leu-CN **21** that was obtained in a mixture with the acid and amide. However, we tried to add cysteine to this mixture, despite it was not pure, as another example of our reactions. In this case, the expected thiazoline was observed (Scheme 22). Similarly to other thiazolines, this one was identified by the appearance of a nearly triplet peak at 4.97 ppm for CH on the cycle and an ABX system from 3.5 to 3.7 ppm corresponding to the CH₂.



Scheme 22

I.3.4 N-Acetyl aminoacetonitrile with cysteine

N-Acetyl aminoacetonitrile, which could be considered as a model for any other N-acyl acetonitrile, including peptidic nitriles, also reacted with cysteine in good rate (Scheme 23).



Scheme 23

The formation of the expected thiazoline ring **71** was confirmed by the appearance of a small triplet signal (4.9 ppm) corresponding to CH, and the ABX system (from 3.4 to 3.6 ppm) corresponding to CH₂ on the cycle. (*Figure 31*). In this example however, the hydrolysis step was quick, and the major signals observed on the ¹H-NMR were that of the dipeptide **72**.

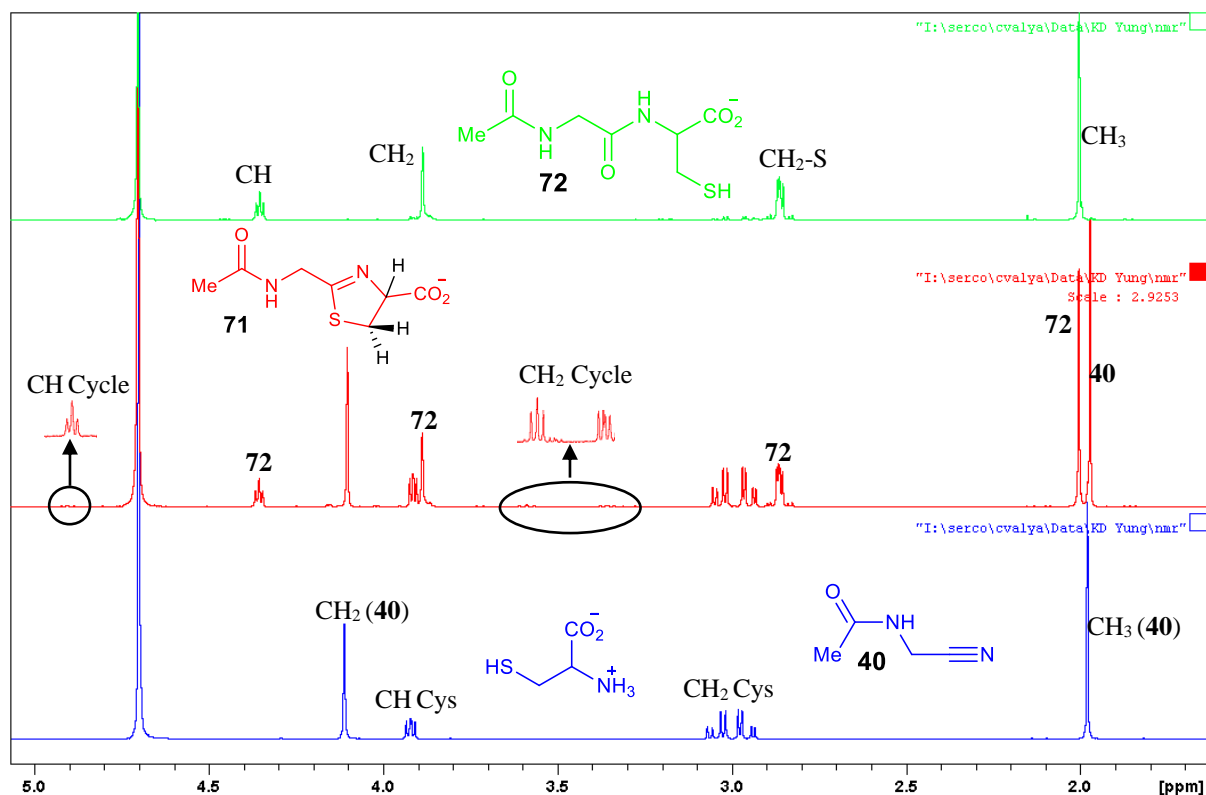
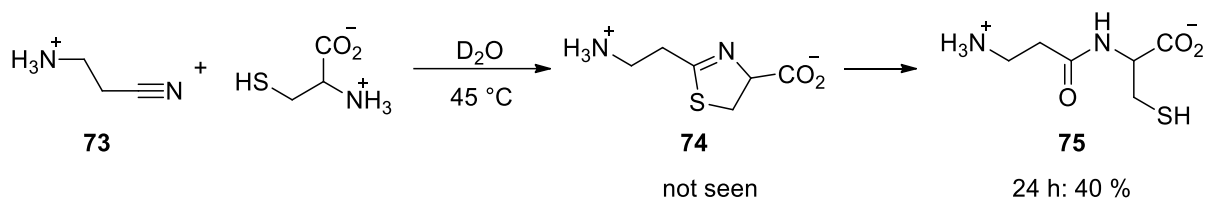


Figure 31: Reaction of cysteine with N-acetyl aminoacetonitrile

I.3.5 Reaction of β -aminopropionitrile with cysteine and homocysteine

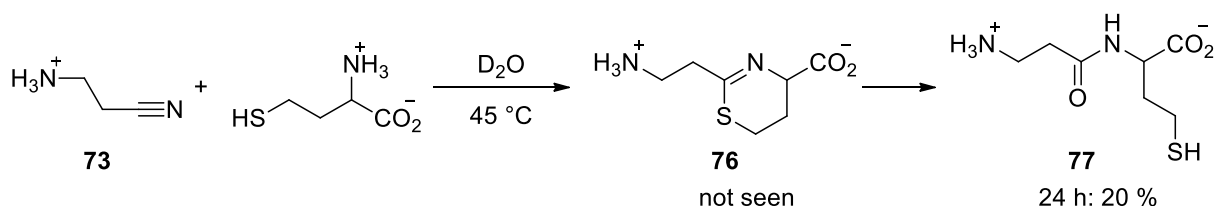
On the contrary, the reaction of cysteine was much slower when α -aminoacetonitrile **39** was replaced by β -aminopropionitrile **73**. In this case, the reaction took few days to complete but eventually yield the expected amide **75** (*Scheme 24*). This result agrees with our theoretical calculations that will be discussed at the end of this chapter. It is noteworthy that the corresponding thiazoline **74** in this example hydrolysed quickly and was not observed on ¹H-NMR spectra.

Chapter I



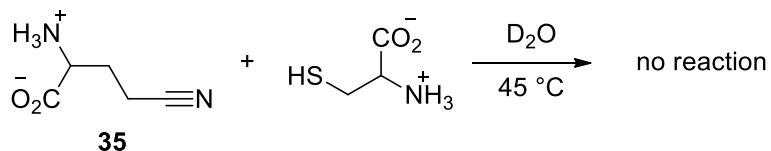
Scheme 24

As in the case of cysteine, β -aminopropionitrile **73** reacted with homocysteine much slower than aminoacetonitrile **39** (*Scheme 25*).



Scheme 25

Trying to test the effect of moving the nitrile farther from the amine, we used the γ -nitrile of glutamic acid **35**¹⁵⁵ and added it to a cysteine solution in D₂O. No reaction was observed in this case (*Scheme 26*).



Scheme 26

Thus, according to our findings, if a mixture of various α , β and γ -aminonitriles was present in the primitive ocean, α -aminonitriles would have had the best chance to react with aminothiols and would have led to the formation of “regular” peptides, consisting only of α -amino acid residues. Fortunately, these results agree with the living systems that we know nowadays, in which only α -amino acids take part in the formation of proteins.

I.3.6 Competition between α and β -aminonitrile

We used to preserve an acceptable range of pH's (5.5 – 6.5) in our reactions, and to maintain the temperatures at the selected values (45 °C or room temperature). However, our comparisons of the reactivity of two nitriles in separate containers are not very accurate as the pH (or even the temperature) may vary a little bit, and thus the conditions will not be exactly the same. In order to achieve error-free results, we performed competition reactions where both nitriles in the same container and react freely with a chosen aminothiols.

I.3.6.a Competition between α and β -aminonitrile with cysteine

We used an equimolar amounts of cysteine, aminoacetonitrile **39** and β -aminopropionitrile **73** in D₂O and followed the reactions by ¹H-NMR (*Figure 32*). Clearly, cysteine reacted with the α -aminonitrile much faster than with its β -analogue. The reaction with α -aminonitrile took few hours to complete. On the other hand, it took several days in the case of β -aminonitrile.

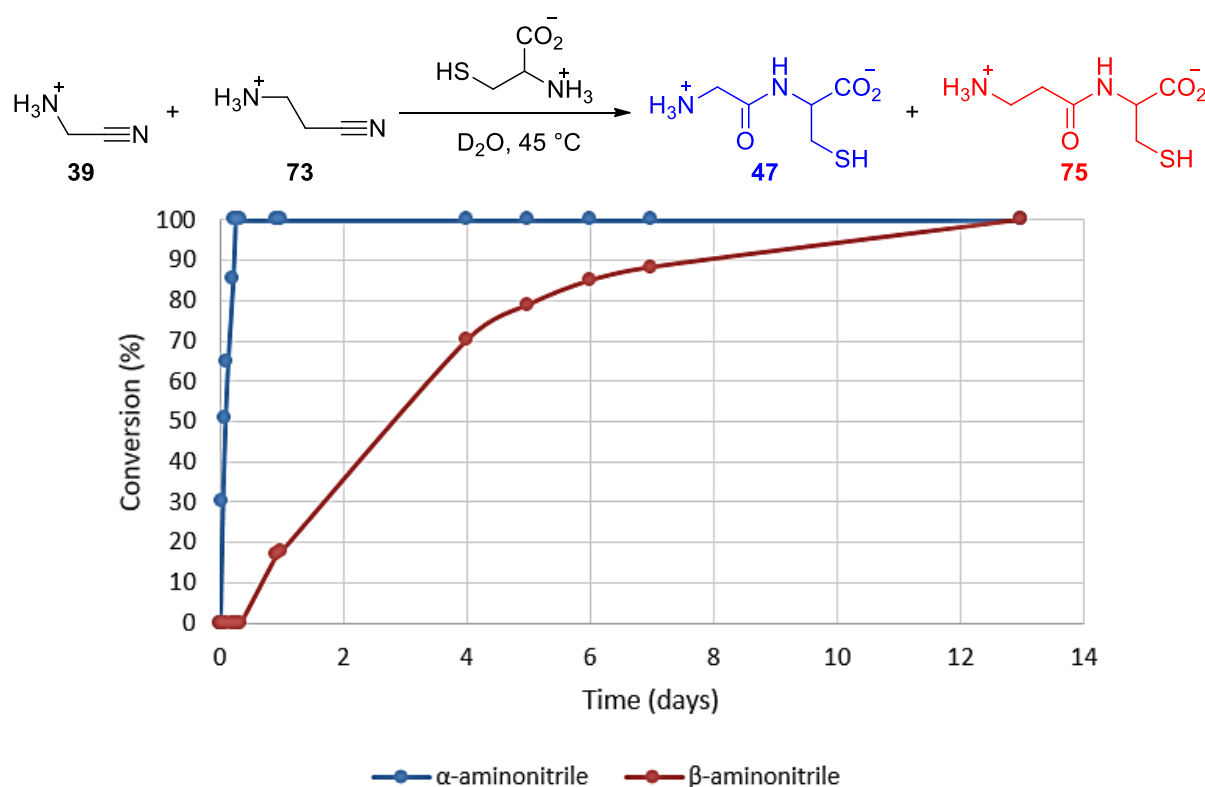


Figure 32: Competition between α and β -aminonitrile with cysteine. (Conversion = consumption of starting materials, forming thiazoline and the amide)

I.3.6.b Competition between α and β -aminonitrile with homocysteine

Similarly to cysteine, homocysteine reacted with aminoacetonitrile **39** much faster than with β -aminopropionitrile **73**. Here also, while the reaction of α -aminonitrile took few hours to complete, it took several days in the case of β -aminonitrile (Figure 33).

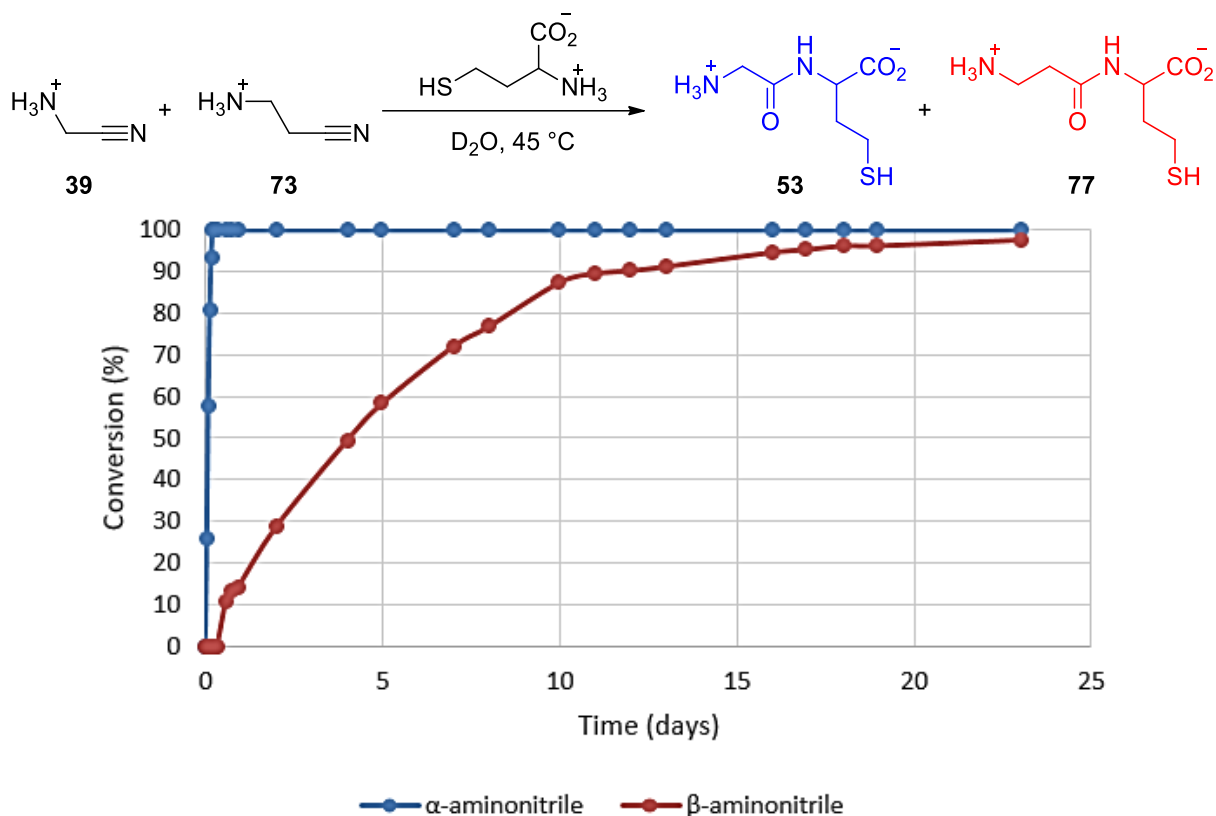


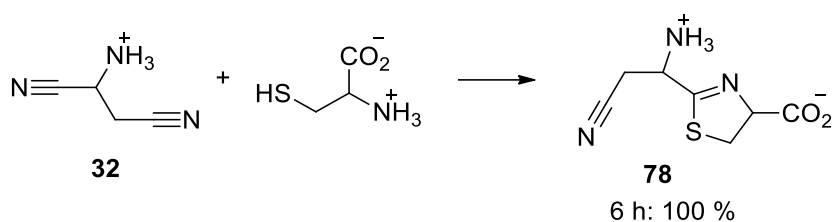
Figure 33: Competition between α and β -aminonitrile with homocysteine. (Conversion = consumption of starting materials, forming thiazoline and the amide)

I.3.7 Reaction with 2-amino-1,4-butanedinitrile

In the aforementioned examples, we have tested the chemoselectivity of aminothiols toward α - and β -aminonitriles (as separate molecules) in competition reactions and we found that reactions with α -aminonitriles are always much faster than with β -aminonitriles. As a complement, we investigated what happens when both nitriles were located on the same molecule, that is to say, when chemoselectivity can be also regarded as regioselectivity.

Kieu DUNG LY (during her M2R internship in our group) synthesized the bis-nitrile derived from aspartic acid **32**¹⁵⁶ and examined its reaction toward cysteine. In this case, only the α -

aminonitrile reacted with cysteine giving the corresponding thiazoline **78** which was stable under our conditions (*Scheme 27*).



Scheme 27

1D (^1H and ^{13}C -NMR) spectra demonstrated that only one nitrile reacted with cysteine to give a thiazoline cycle. The use of 2D NMR spectroscopy (HMBC) showed that the reaction occurred on the α -aminonitrile side, while the β -aminonitrile remained unreacted (an equimolar cysteine/bisnitrile mixture was used), once again demonstrating the selectivity toward α -aminonitriles (*Figure 34*).

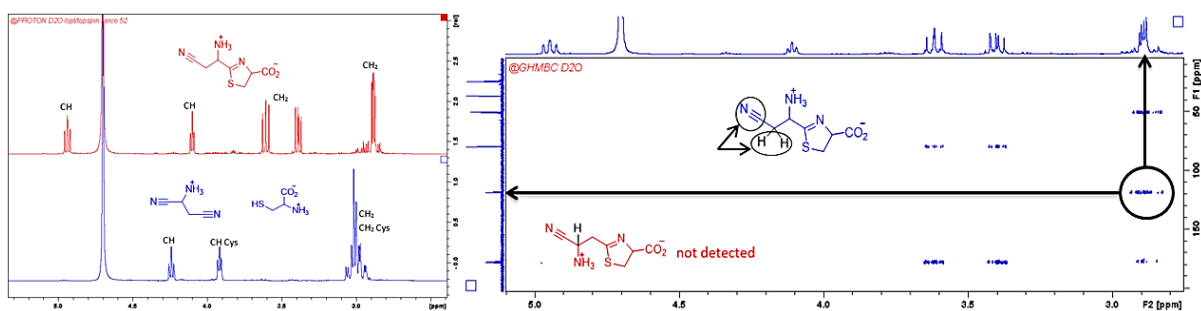


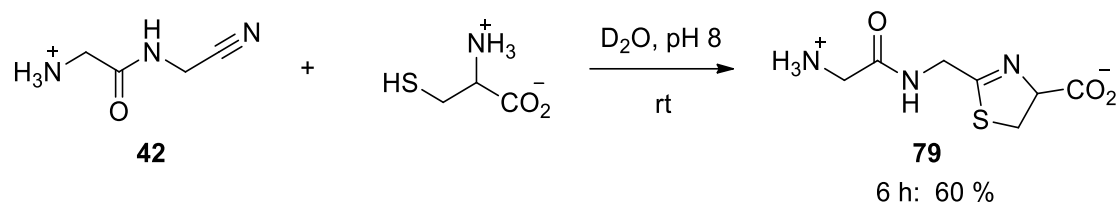
Figure 34: Reaction of aspartic acid bis-nitrile with cysteine, 1D NMR (left), 2D (HMBC) experiment (right) demonstrating the selectivity toward the α -nitrile

I.4 Formation of Tri and Tetra-peptides

I.4.1 Reaction of cysteine with GlyGly-CN and GlyGlyGly-CN

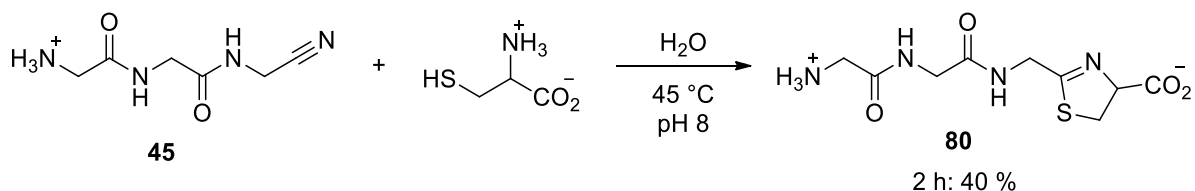
In order to obtain a simple example of a tripeptide, we mixed Gly-Gly-CN* with cysteine. The corresponding thiazoline **79** for this tripeptide was observed in this case and it was characterized in ^1H -NMR by its triplet (4.91 ppm) and its ABX system (3.38 to 3.6 ppm) relative to the CH_2 in the cycle (*Scheme 28*).

* We will demonstrate later that this nitrile may be regarded as a plausible prebiotic molecule.



Scheme 28

Similarly, when we mixed Gly-Gly-Gly-CN **45** with cysteine, we observed the formation of the thiazoline ring **80**. This formation was demonstrated by the appearance of a triplet (4.9 ppm) and the ABX system (3.3 to 3.6 ppm) in the ¹H-NMR. This reaction was relatively rapid at pH 8 and 45 °C. Under these conditions, 15 % of the thiazoline was obtained after 40 min and this conversion increased to 40 % after 2 hours (Scheme 29).

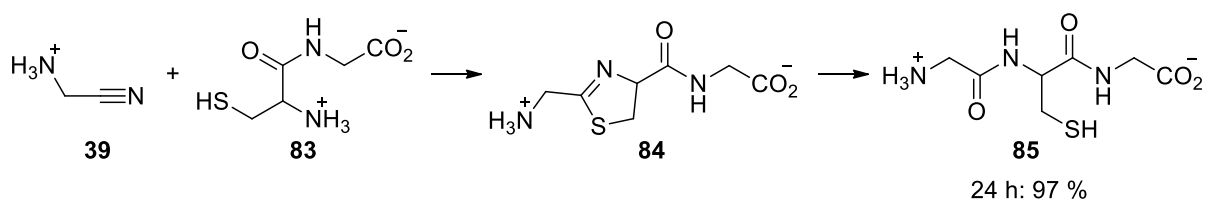


Scheme 29

In these cases, we used slightly basic conditions to be sure to observe thiazoline rings. Under acidic conditions, hydrolysis of thiazolines **79** and **80** would be quicker leading to the corresponding peptides Gly-Gly-Cys **81** and Gly-Gly-Gly-Cys **82**.

I.4.2 Reaction of aminoacetonitrile with Cys-Gly

When Cys-Gly **83** was used, the tripeptide Gly-Cys-Gly **85** was obtained with a very good conversion (Scheme 30).



Scheme 30

I.4.3 Competition between cysteine and Cys-Gly

We have also studied the competition between Cys-Gly dipeptide and cysteine itself. The competition between these two aminothiols was followed by $^1\text{H-NMR}$ (Figure 35). It appeared that the difference between reaction rates toward dipeptide formation was small. Indeed, the figure shows that the reaction with Cys-Gly was slightly faster than that one with cysteine.

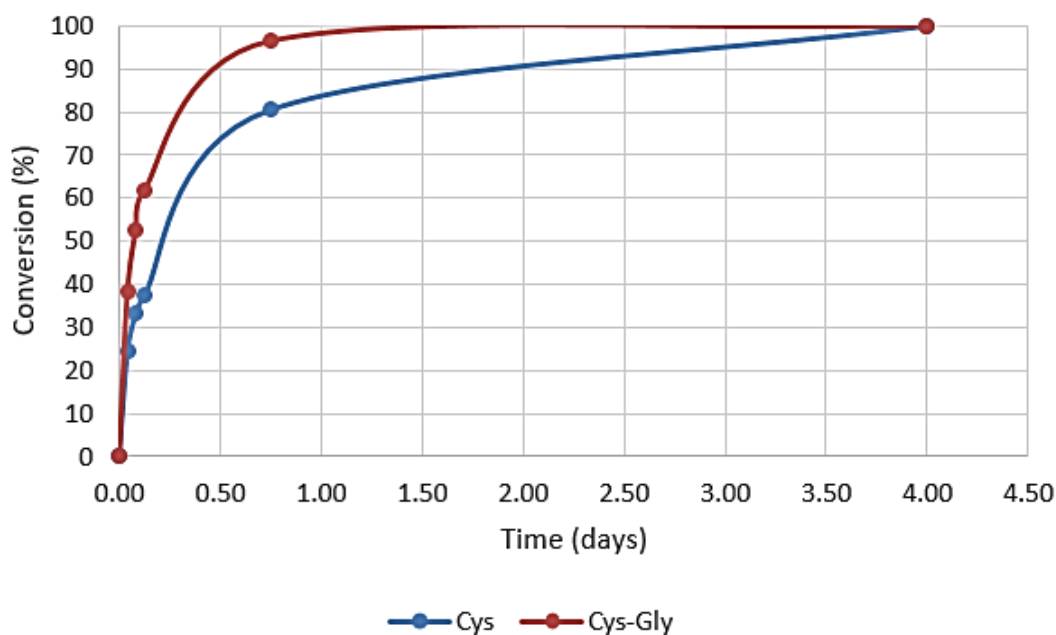
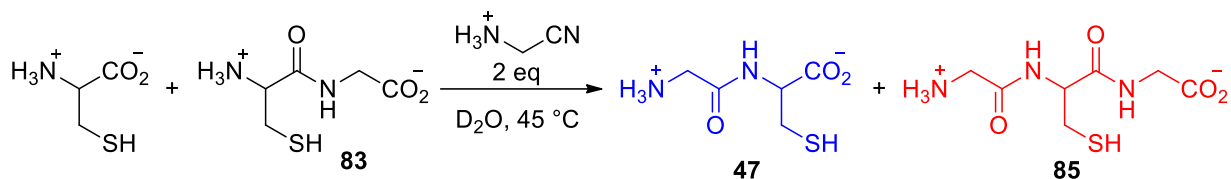


Figure 35: (Conversion = consumption of starting materials, forming thiazoline and the amide)

This shows that the acid function of cysteine can be replaced by a secondary amide. In other words, any longer peptide with an available aminothiols function at one of its NH_2 -ends would react with Gly-CN (or any other α -aminonitrile).

These results confirm our suggestion that longer peptides ending with an aminonitrile would react with cysteine (or any peptidic chain starting with aminothiols). Furthermore, this reaction

would link two peptidic chains together if they meet the needed requirements (*i.e.* NH₂-end of one chain is an aminothiols, the CO₂H-end of the other chain is replaced by CN) (*Figure 36*).

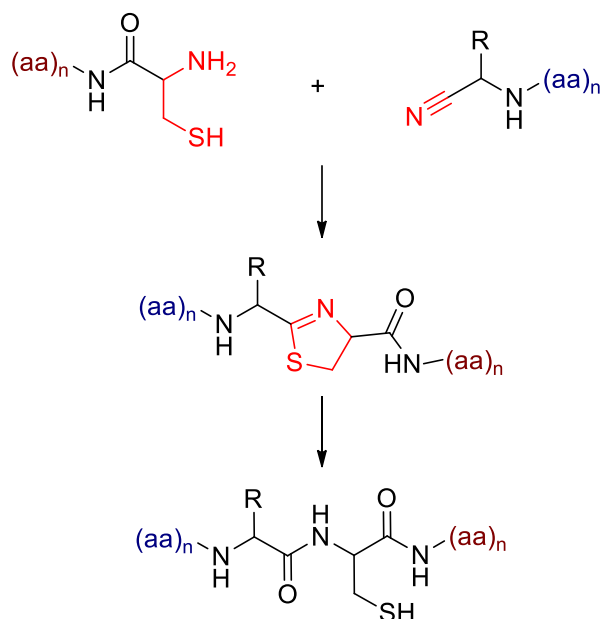
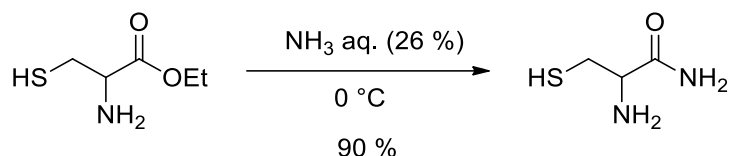


Figure 36: Connecting two peptidic chains together via thiazoline ring formation

I.5 Reactions with cysteine amide

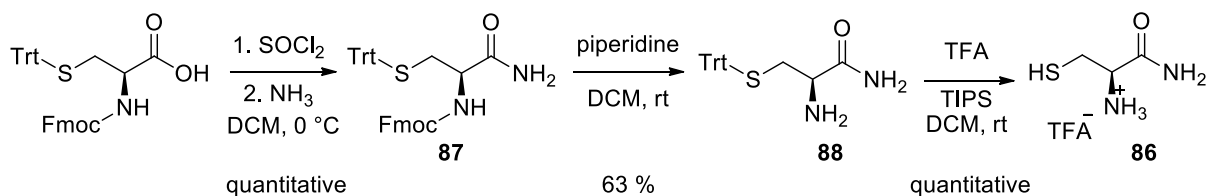
In common with other amino acids, cysteine is a polar water-soluble molecule. Consequently, reactions between cysteine and aminonitriles lead to highly polar products, which are also soluble in water. This makes it difficult to purify our products by the common chromatography column method (in normal phase). For this reason, we decided to test reactions with cysteine amide. In this case, we expected the isolation of products to be an easy process.

We used two methods to synthesize cysteine amide **86**. The first (and simplest) method was a one-step method starting from cysteine ethyl ester. It was enough to mix the ester with a 26 % ammonia solution for one day to obtain the amide with a good “yield” (*Scheme 31*).



Scheme 31

In spite of the apparent good “yield”, the product was not pure enough and contained some remaining starting material that hydrolysed to give cysteine, and cystine. From this mixture we were not able to obtain a “good-quality” cysteine amide. For this reason, we used an alternative three-step method to obtain it. Starting from Fmoc-Cys(Trt)-OH, we obtained the amide **87**. Then, we cleaved the Fmoc group using piperidine to provide **88** and finally, the trityl group was removed by TFA to provide the desired amide as a TFA salt **86** (Scheme 32).



Scheme 32

The second “multistep” method provided the amide in high purity, and thus it was used to obtain cysteine amide that we used in our reactions.

I.5.1 Reaction of cysteine amide with aminoacetonitrile

As an aminothiols, cysteine amide reacted with aminoacetonitrile in the same manner as cysteine.

First, with the aim of studying the selectivity between these two aminothiols, we added aminoacetonitrile, as an example of α -aminonitrile, to an equimolar mixture of cysteine and cysteine amide **86**.

This competition reaction was performed in D₂O at 45 °C and followed by ¹H-NMR. The aminonitrile was totally consumed after 26 hours. We found that both candidates reacted in similar rates, indicating that the presence of the amide group, in place of the acid, has no significant effect on the reaction (Figure 37).

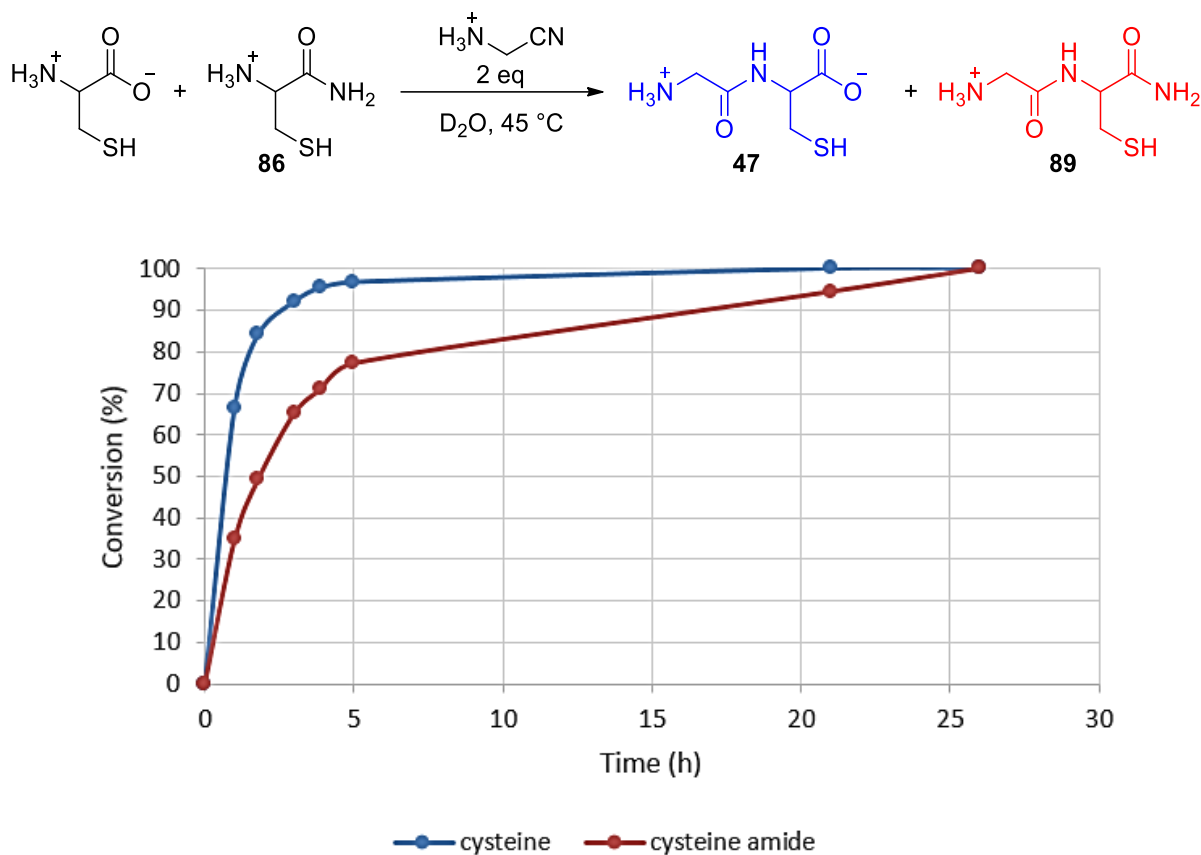
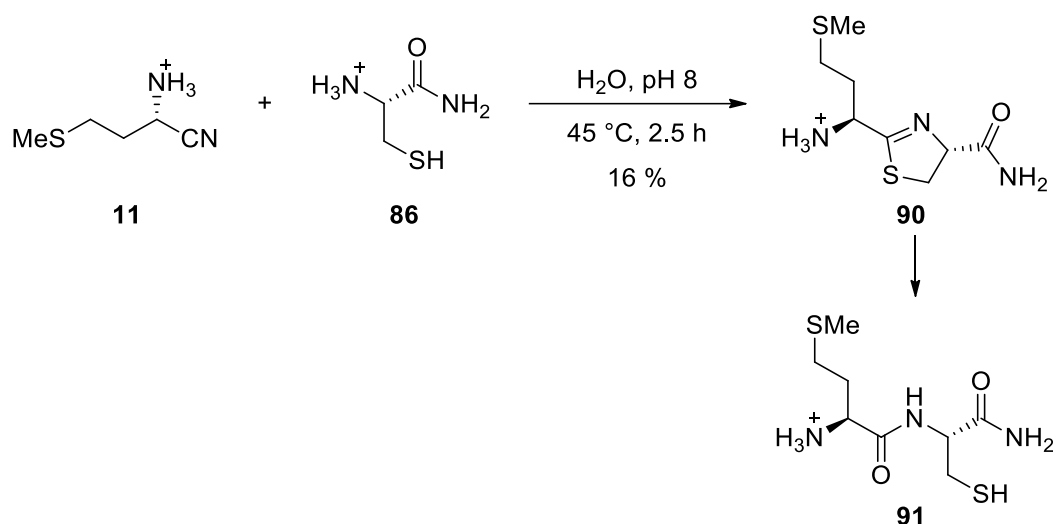


Figure 37: Competition between cysteine and cysteine amide **86** with aminoacetonitrile.
(Conversion = consumption of starting materials)

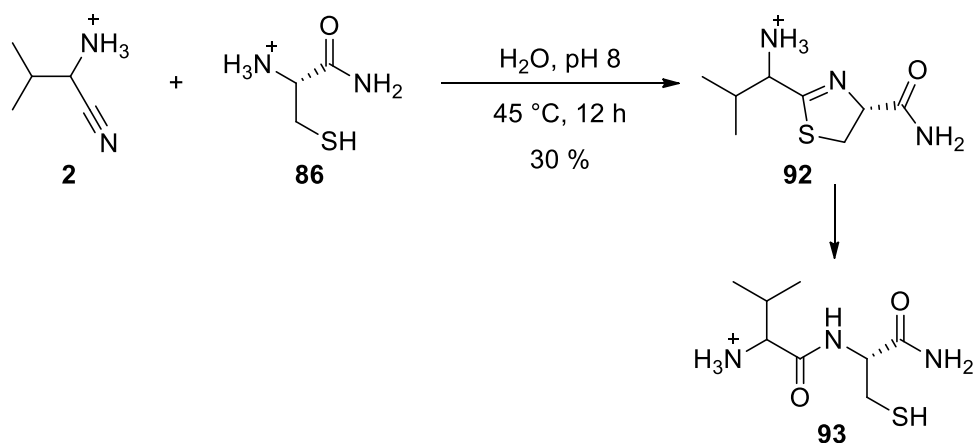
I.5.2 Reaction of cysteine amide with aminoacetonitrile

As reported before, both methionine and valine nitriles reacted with cysteine and the expected thiazolines and dipeptides were identified. However, none of these products was isolated. From this perspective, we repeated these reactions with cysteine amide. In this case, we perform the reactions in slightly basic water solutions (pH 8) in order to guarantee the formation and stability of heterocyclic rings. Using these conditions, starting with L-methionine nitrile used in an enantiopure form, we were able to isolate the corresponding thiazoline **90**, but only in 16 % yield (*Scheme 33*).



Scheme 33

We also used Val-CN **2** (racemic mixture). In this case, a one to one mixture of diastereoisomers of the thiazoline **92** was isolated (30 % yield) (Scheme 34).



Scheme 34

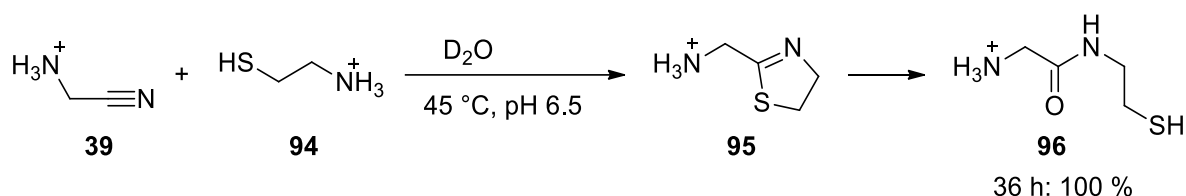
Both isolated yields are deceptive*: we believed that this is due to some degradation of products on silica.

It should be noticed that the hydrolysis of the valine derived thiazoline **92** to obtain the dipeptide **93** was found to be very slow process, indeed much slower than the hydrolysis of the methionine derivative **90** (and of the simplest glycine derivative **46**). This is probably due to the presence of the bulky isopropyl group in **92**.

* **90** and **92** were isolated as free amines.

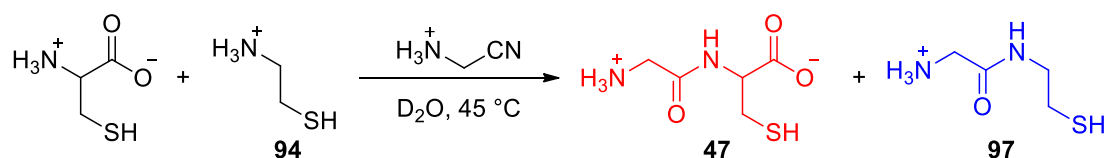
I.6 Reaction of aminoacetonitrile with cysteamine

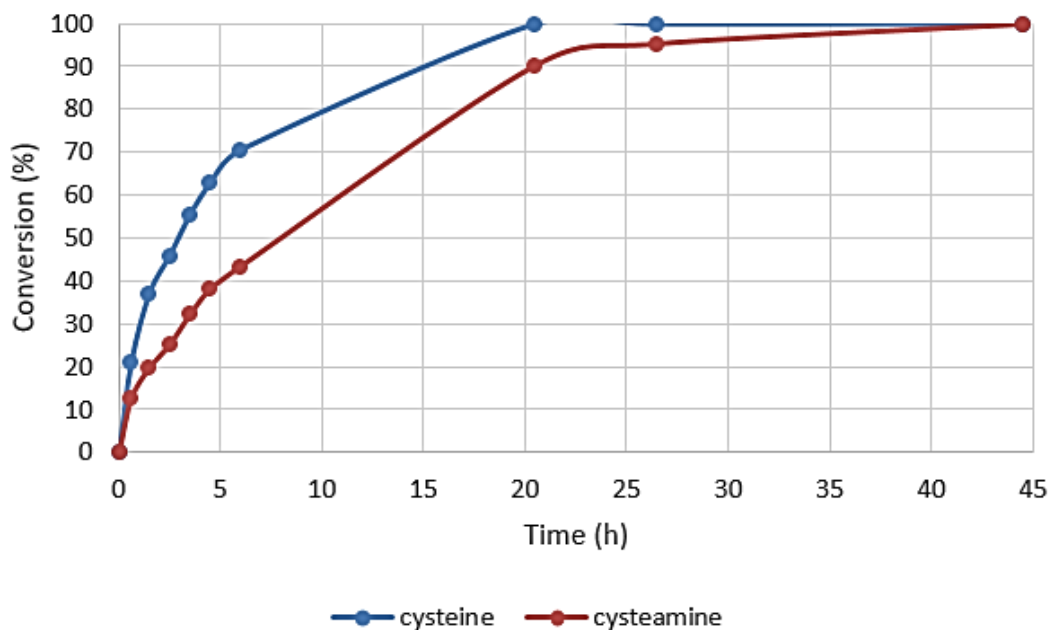
We also tested the reaction of α -aminoacetonitrile **39** with cysteamine **94**, the simplest stable aminothiols. In this case, the corresponding thiazoline ring **95** should be formed and be hydrolysed to give the amide **96** (*scheme 35*).



Scheme 35

As we did with cysteine amide, we studied the competition between cysteamine **94** and cysteine. We introduced both aminothiols in a D₂O solution containing aminoacetonitrile **39**, and followed the reaction by ¹H-NMR (*Figure 38*). The results did not show a big difference between both competitors toward thiazoline and amide formation. Even though cysteine reacted slightly faster than cysteamine, the result indicates once again (as with cysteine amide) that the presence of the carboxyl group in aminothiols is not compulsory for the formation of thiol-containing dipeptides.

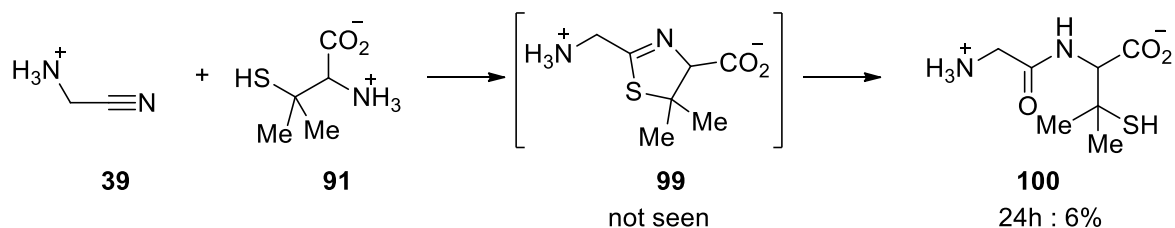




*Figure 38: Competition between Cysteine and cysteamine with aminoacetonitrile.
(Conversion = consumption of starting materials)*

I.7 Reaction with penicillamine

Finally, we tested the reactivity of penicillamine **91**, a sterically hindered aminothiol (*Scheme 36*). In this case, the reaction was very slow, probably because of the bulkiness of the two methyl substituents.



Scheme 36

Furthermore, the thiazoline was not seen on the $^1\text{H-NMR}$ and the only detected product was the final dipeptide (*Figure 39*). This might be due to the electron donating property of the methyl groups, making the nitrogen atom of the intermediate thiazoline ring more basic. Protonation of this nitrogen atom would thus be easier, hence the hydrolysis step quicker.

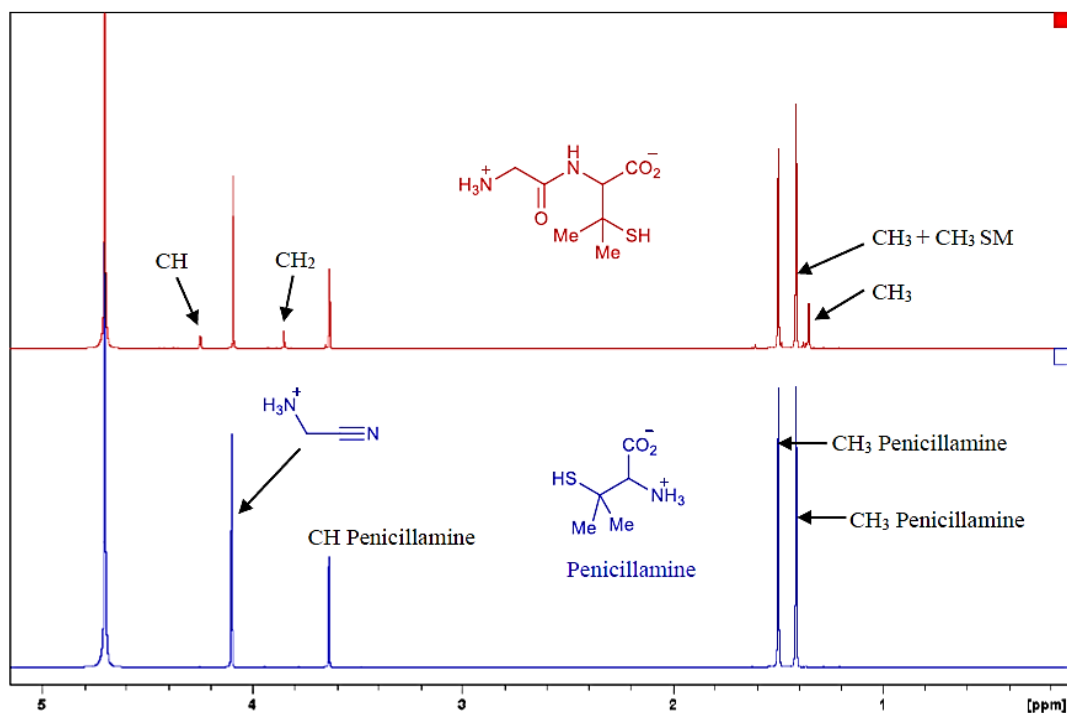


Figure 39: Reaction of aminoacetonitrile with penicillamine

I.8 Theoretical calculations

In order to explain the observed selectivity, Anne-Milet (from theoretical chemistry team in our lab) calculated the level of the π^* orbital of a series of nitriles (*Table 1*).

As shown in the table, the lowest calculated orbital was that of the protonated form of aminoacetonitrile **39**. Such level would explain its greater reactivity compared to other nitriles. The non-protonated form of aminoacetonitrile **39** is predicted to be much less reactive and so is probably not involved in the reaction mechanism.

The π^* orbital of β -aminopropionitrile **73** is higher (it is less reactive) and the simplest γ -aminonitrile **102** is predicted to be even less reactive (no reaction from glutamic nitrile **35**). In contrast, α -substitution of aminoacetonitrile, as in α -aminopropionitrile **101**, should not alter significantly its reactivity (*Table 1*).

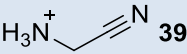
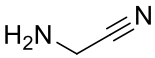
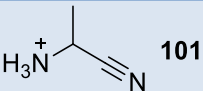
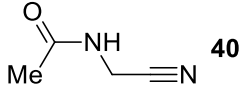
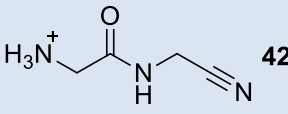
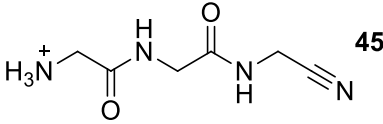
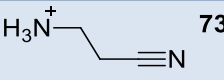
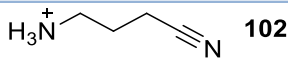
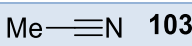
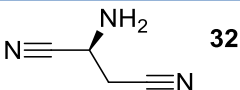
Nitrile	π^* value	reaction rate	
 39	-0.03632	quick	
	0.01698	no reaction (?)	
 101	-0.03010	quick	
 40	0.00504	slower	
 42	0.00353	quick	
 45	-0.00547	quick	
 73	0.00616	slower	
 102	0.01556	no reaction	
 103	0.03491	no reaction	
 32	α -CN	-0.03566	reacts at α -CN
	β -CN	-0.00500	

Table 1

I.9 Observation of a catalytic behaviour of aa-Cys and aa-Hcy dipeptides

On the basis of our experiments, we propose that aa-Cys and aa-Hcy dipeptides were over-represented in the primitive ocean (compared to non-thiol containing dipeptides). These dipeptides are thiols and as such could be major players in a “thioester world”.¹⁰⁹

Indeed, when we mixed Gly-Cys **47** with an excess of aminoacetonitrile **39** in D₂O solution at 45°C, a peak was observed at 194.16 ppm in the ¹³C NMR of the reaction mixture. Such chemical shift is characteristic of the thioester function (*Figure 41*). Thioester **104** was observed also in ¹H-NMR (*Figure 40*). We believe that it belongs to compound **104**. We also noticed the formation of glycine amide **105**. These products would derive from the first formed C=N double-bonded addition product **106**. The thioester was partly hydrolysed to

give glycine **107**. We were also able to characterize among the reaction products the amidonitrile Gly-GlyCN **42** (^{13}C NMR: 27.58, 40.34, 116.74, 167.58 ppm), meaning that the thioester reacted with the non-protonated amino group of GlyCN **39** (which is possible because of the low pKa of GlyCN: 5.55).¹⁵⁷

For instance, in an experiment in which we used globally 4 eq. of GlyCN **39** (relative to cysteine), after 2 days at 45°C, the observed GlyOH **107**/GlyNH₂ **105**/GlyGlyCN **42** ratio was found to be 21/37/42. This demonstrates that Gly-Cys **47** is able to promote the formation of a peptide bond from a nitrile (*Figure 40*). Similar results were obtained from Gly-Hcy **53**.

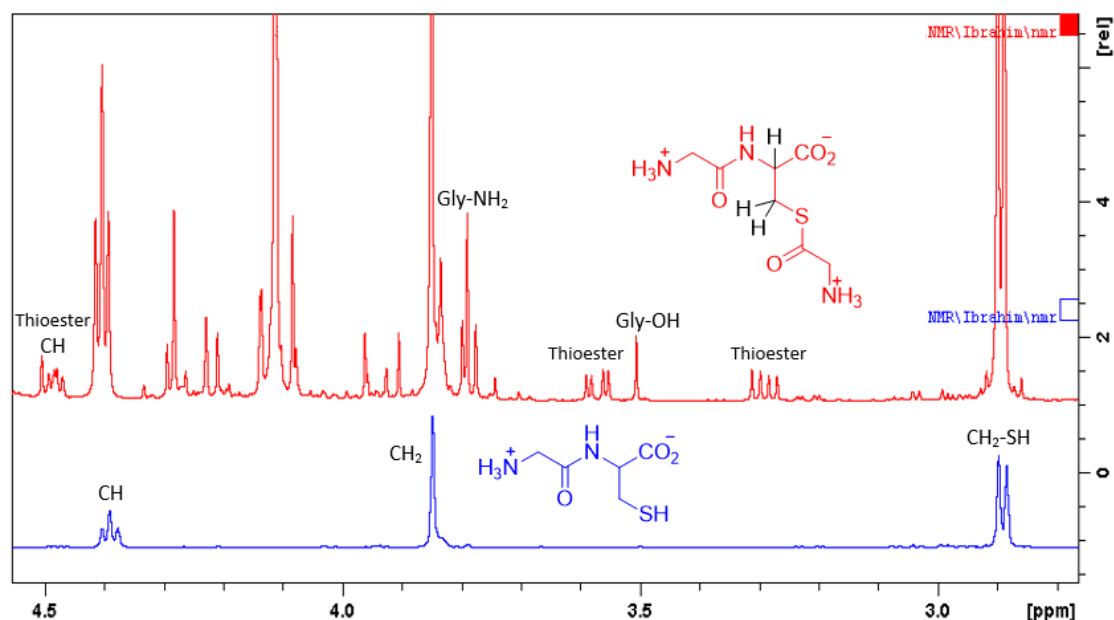


Figure 40: ^1H -NMR showing the formation of thioester xx upon addition of excess GlyCN to Gly-Cys

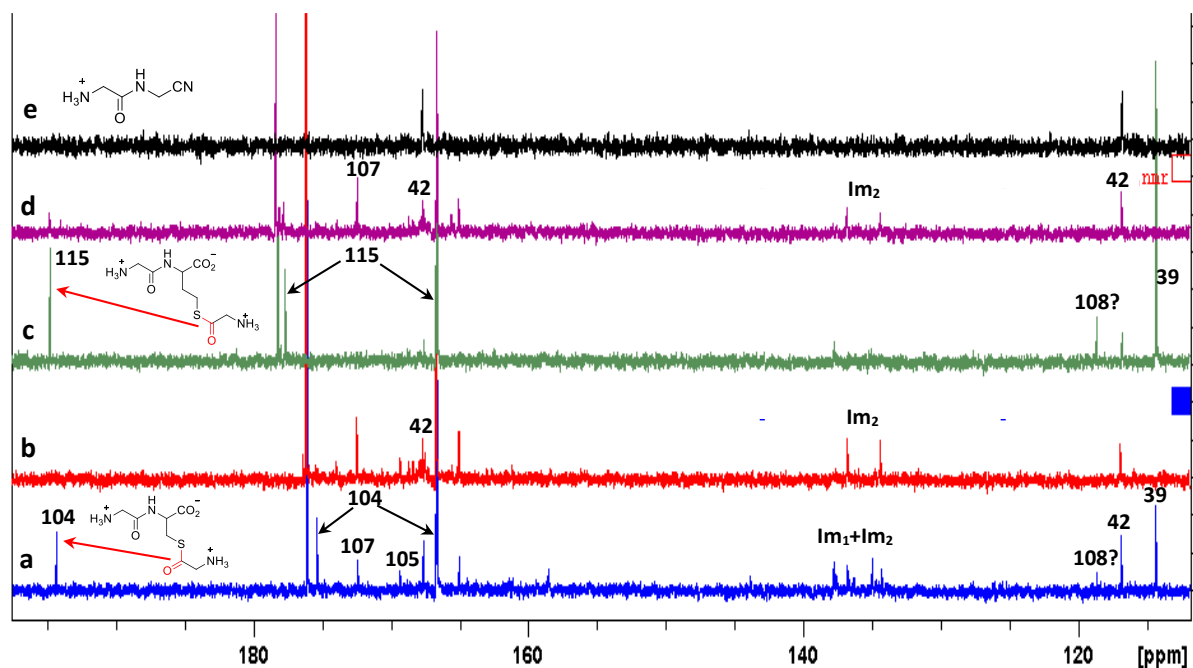


Figure 41: ^{13}C NMR spectra recorded during the reaction of an excess GlyCN **39** with GlyCys **47** (from **39.HCl** + cysteine) or GlyHcy **53** (from **39.HCl** + homocysteine) at 45°C , pH 6.5. *a:* with **47** 20h after mixing **39.HCl** and cysteine, *b:* after 70h, *c:* with **53** 20h after mixing **39.HCl** and homocysteine, *d:* after 70h, *e:* reference spectrum of 1h. glycine **107**. Peaks corresponding to at least two products were detected near 135-140 ppm. They might correspond to two different imidazoles (named Im_1 and Im_2). **115**: the homocysteine thioester analogue of **104**. Big peaks at 166.65 (166.66) and 176.07 (178.33) belong to **47** (and **53**). One peak of both **104** and **115** sticks to the foot of the 166.6 ppm peak of **47** and **53**

The formation of other products was also evidenced in the reaction of cysteine with excess GlyCN **39**. In the ^{13}C NMR spectra peaks were observed at 130-140 ppm (*Figure 41*).

The mass spectrum of a reaction mixture in water also showed the formation of various products (*Figure 42*). This mass spectrum first confirmed the presence of Gly-GlyCN **42** (protonated, found 114.0656, calcd 114.0667). Another mass was detected at 113.0817. It could be attributed to the amidine **108** (calcd 113.0817). However, in accordance with the observed ^{13}C NMR of the mixture, we propose that this amidine **108** cyclized and that this mass peak should be attributed to the imidazole **110**. Indeed, the cyclization of an amidinonitrile similar to **108** into an aminoimidazole (5-amino-2-methyl-1H-imidazole) has already been reported.¹⁵⁸ At least some of the other products observed in the mass spectrum would be evolution products of **110**. For instance, this imidazole could lose ammonia to give

the stabilized cation **111** that would in turn react with **110** (in its free amine form) to yield the bis-imidazole **112** ($M+H$, found 208.1300, calcd 208.1310). **110** itself could react with thioester **104** to give the amide **113** (or **114**) (Figure 43). Similar results were obtained in case of addition of **39** to **53**. In this case, the thioester **115** was identified.

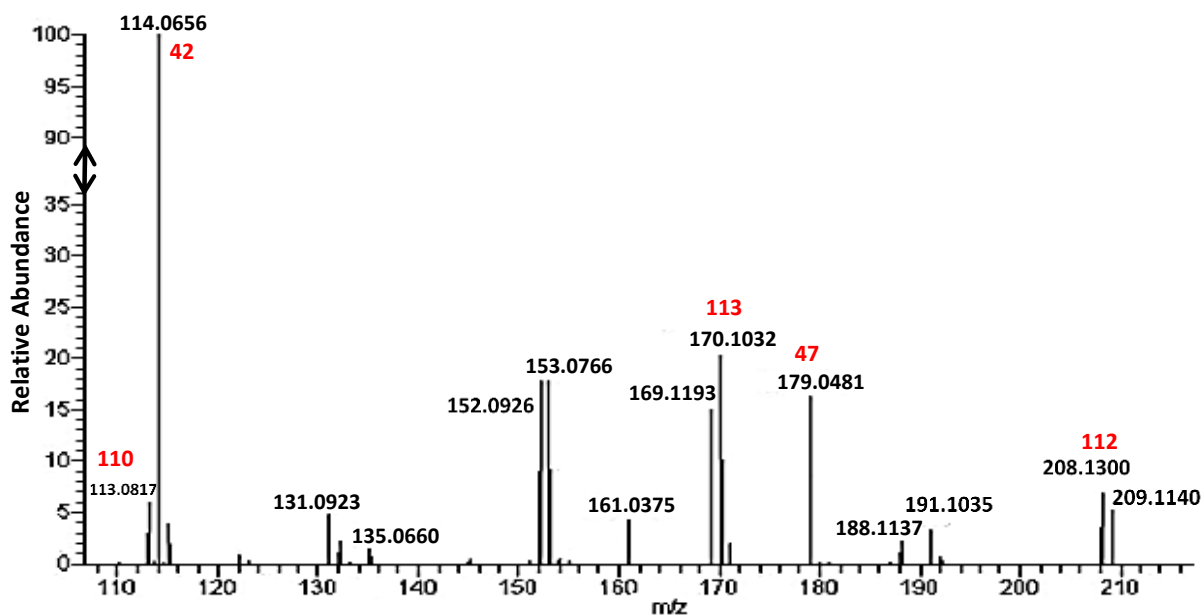


Figure 42 : Mass spectrum of a reaction of an excess GlyCN **39** with GlyCys **47**

Chapter I

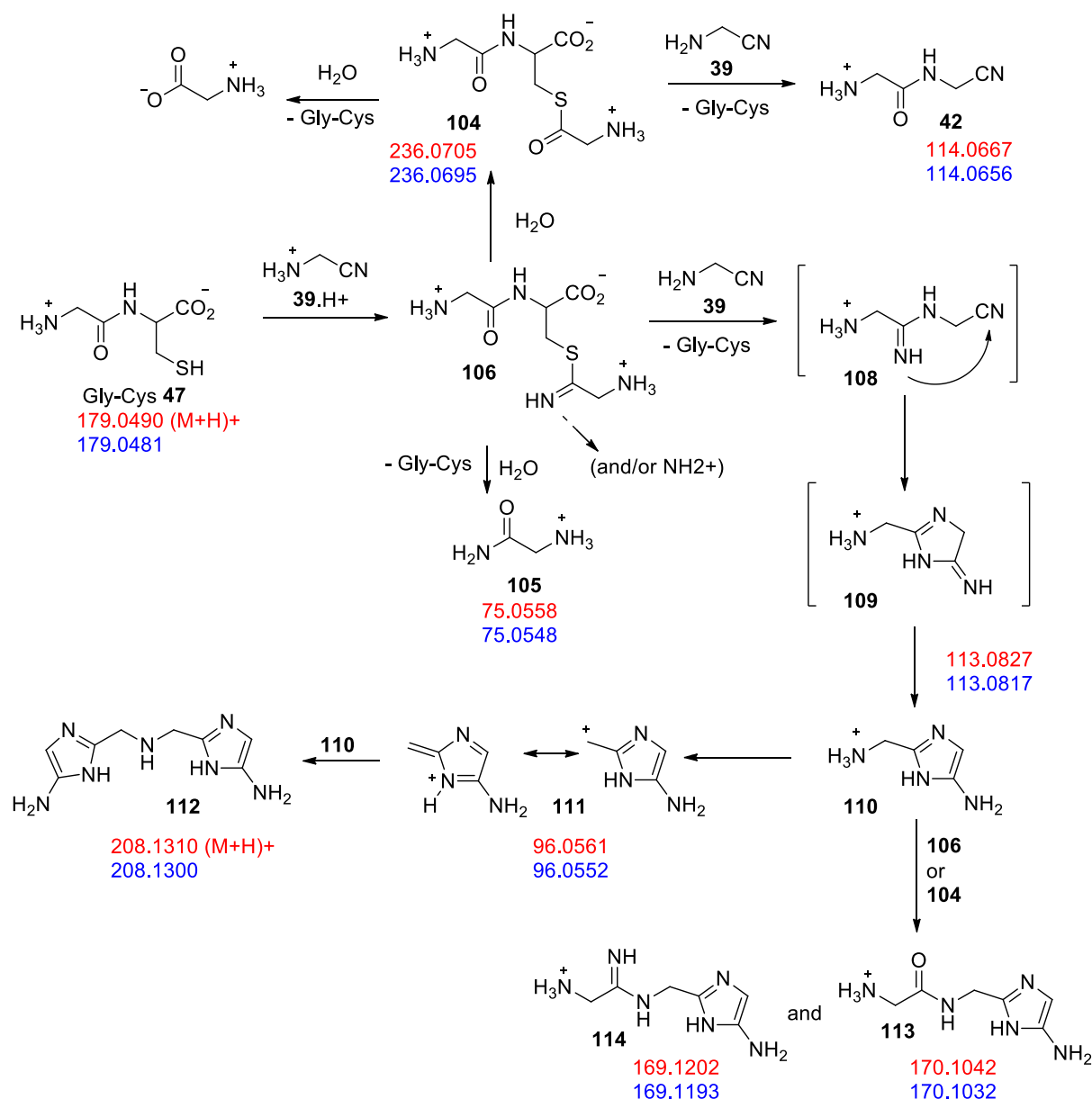
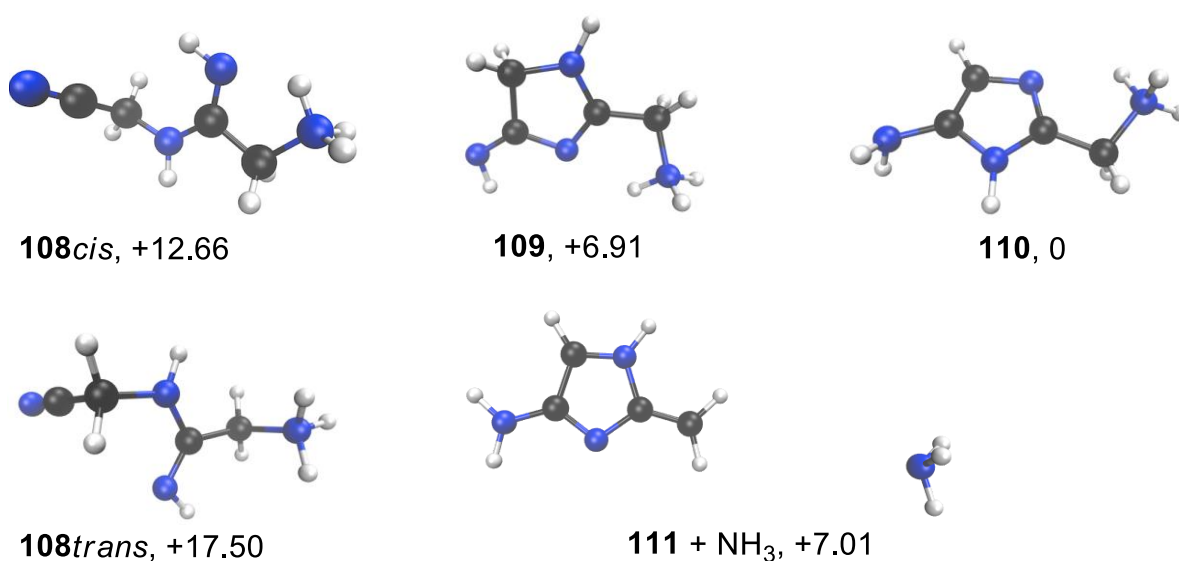


Figure 43: Proposed pathways in the reaction of excess GlyCN **39** with GlyCys **47**. Mass attribution (red, calculated; blue, found)

In order to further ascertain the cyclisation of the imino-compound **108** Anne-Milet, once again, calculated its stability in comparison to cyclized forms. We used the strategy described previously for table 1 results, with the 6-31+G(d,p) basis set (Fig. 11). Not surprisingly **110** was calculated to have free enthalpy 12.7 kcal/mol lower than **108** and 6.9 kcal/mol more stable than **109**.

Interestingly, it was found that the dissociation of **110** into **111** + NH₃ cost only around 7 kcal/mol. The ΔH of dissociation is nearly 19 kcal/mol (18.67 kcal/mol) but due to the dissociative character of the process, the ΔG value drops to 7.01 kcal/mol. This process does not show a well-defined TS. Thus we think that the formation of cation **111** proposed in (Figure 44) is a plausible event.

We studied more precisely the cyclisation step from **108** to **109** (Figure 45). H₃O⁺ was used to promote the reaction and to give a proton to the nitrogen atom of the nitrile group, which becomes part of the exocyclic imine of **109**. Two explicit molecules of water were introduced in addition to the water continuum. It appeared that the cyclisation step should be exocyclic and quick, with a low level TS (activation energy of 6.6 kcal/mol). This is another confirmation that the compound of mass 113.0817 we observed was not **108**, but indeed the imidazole **110** (resulting from a simple proton migration from **109**). It is noticeable that similar calculations for the potential cyclization of the amide **42** into an oxazole indicated that this reaction should not happen, *i.e.* in sharp contrast with **108**, **42** (GlyGlyCN) is stable, an obvious result, amply confirmed the experimental observation!



*Figure 44: Minimized conformation for compounds 108 to 111. B3LYP/6-31+G** with a continuum to mimic the solvent effect of water. ΔG 's relative to 110 in kcal/mol*

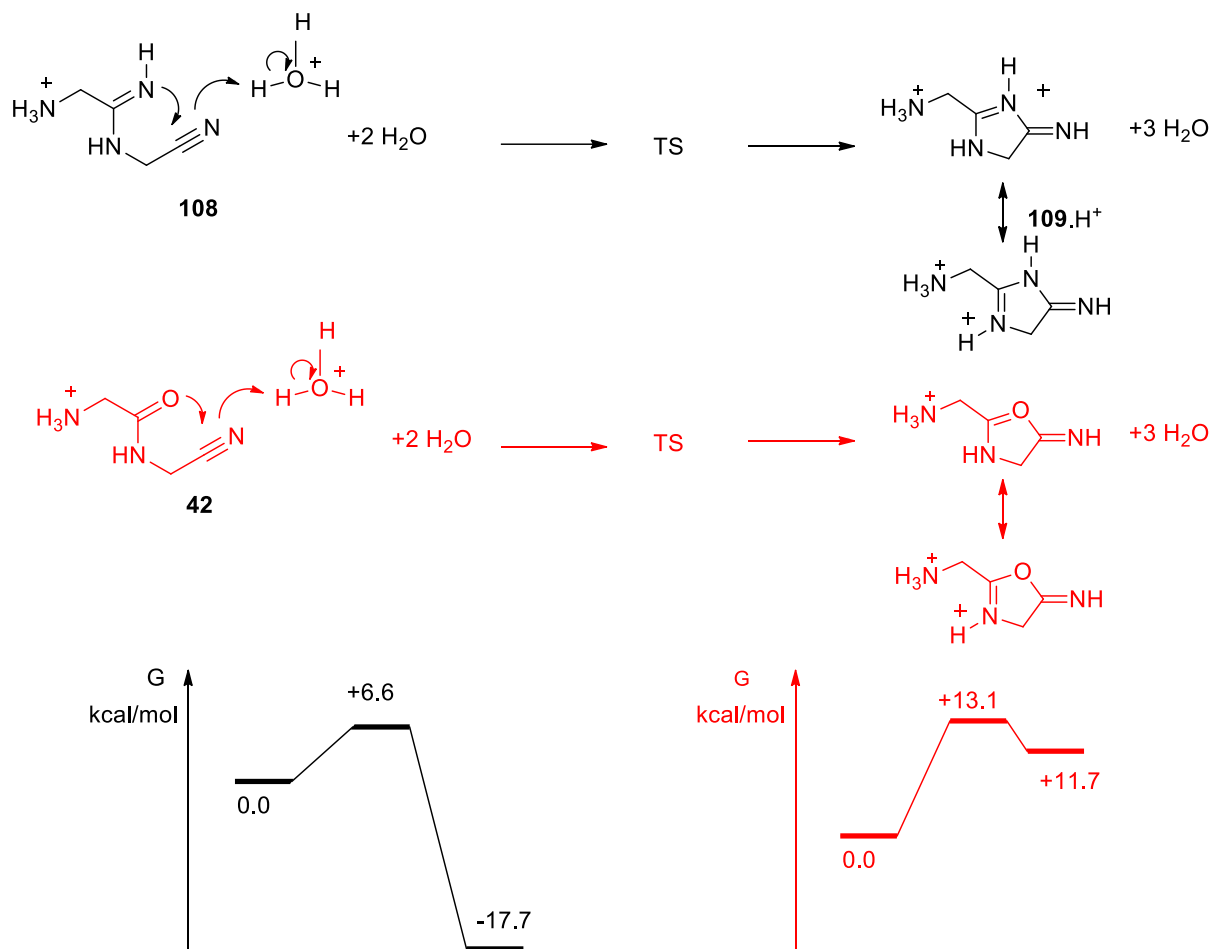
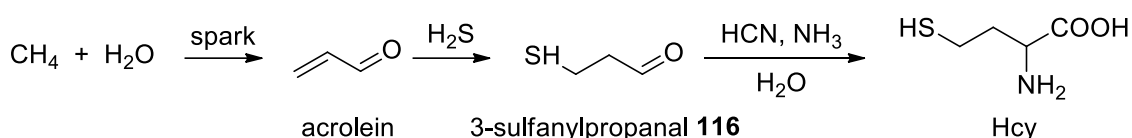


Figure 45: Reaction pathways for the cyclisation step from **108** (black), and **42** (red). Level of calculation B3LYP/6-31+G**/SCRF (water)

Chapter II: Prebiotic formation and a possible primitive role of homocysteine thiolactone

II.1 Introduction

According to our hypothesis, aa-Hcy dipeptides could have been easily formed in the primitive ocean. The presence of homocysteine was highly probable due to the large amounts of H₂S blown up from volcanoes and therefore the needed mercapto-aldehyde precursor **116** could be formed from acrolein and would lead to homocysteine in the presence of ammonia and hydrogen cyanide via Strecker reaction. Furthermore, acrolein is a product of an electric spark on a mixture of methane and water (*Scheme 37*).⁶¹

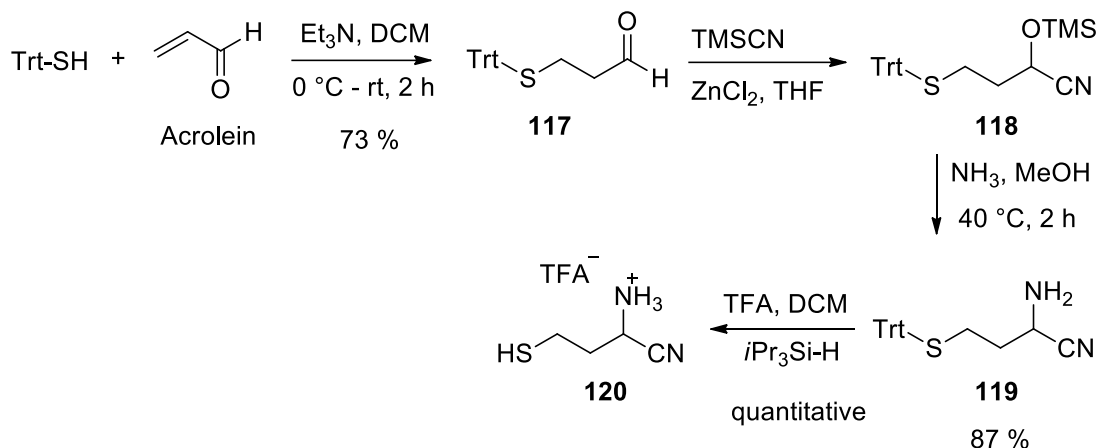


Scheme 37

However, 3-mercaptopropanal (3-sulfanylpropanal) is known to be very difficult to isolate because of autocondensation reactions. Thus, we synthesized homocysteine nitrile (Hcy-CN), in a non-prebiotic way in order to study the behavior that it might have had under primitive conditions.

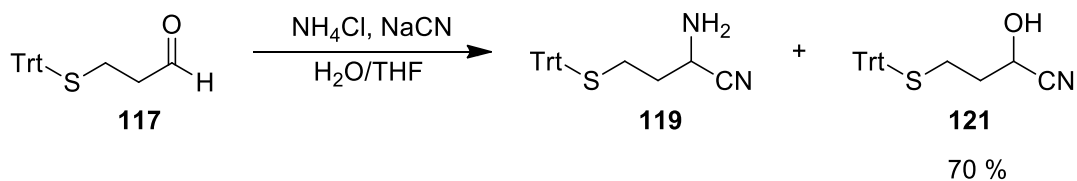
II.2 Synthesis of homocysteine nitrile

We started the synthesis from acrolein and triphenylmethanethiol to obtain 3-tritylsulfanylpropanal **117**.¹⁵⁹ The aldehyde **117** was isolated and submitted to a modified Strecker condensation.¹⁶⁰ The last step was deprotection of the Trt group under acidic conditions to obtain Hcy-CN **120**, which was finally recovered in D₂O solution as a salt with trifluoroacetic acid (*Scheme 38*).



Scheme 38

Traditional Strecker reaction has been also tried to obtain the product **119**. However, a considerable amount of cyanohydrin **121** was obtained in this case (Scheme 39).



Scheme 39

II.3 Homocysteine nitrile behaviour in water

Hcy-CN **120** was stable under strong acidic conditions. No hydrolysis was observed. It could be expected that it would be able to polymerize in a similar way to Cys-CN under neutral pH (see chapter 3). However, Hcy-CN **120** showed a completely different behaviour. Instead of polymerization, it cyclized to form homocysteine thiolactone **54** (Figure 46).

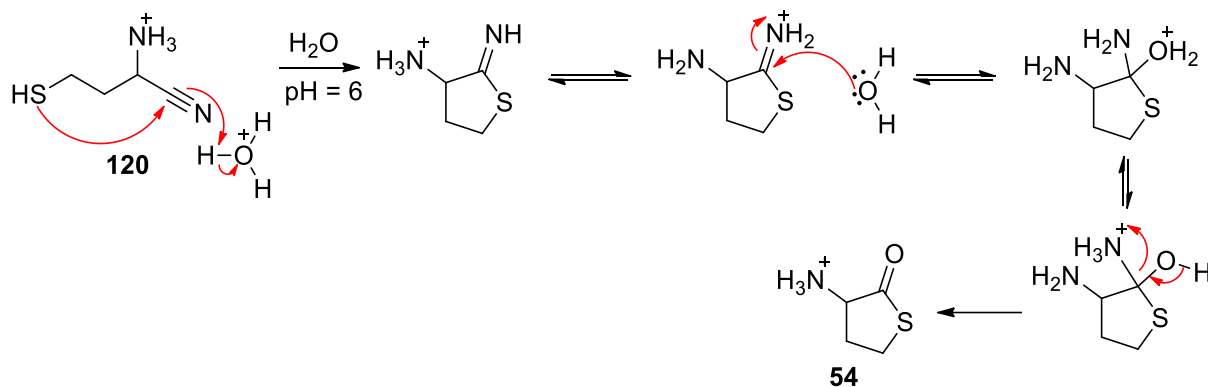


Figure 46: Homocysteine nitrile cyclization

This cyclization was followed by ¹H-NMR spectroscopy. Thiolactone formation was complete after about 30 hours at room temperature (*Figure 47*), while it was much faster at 45 °C, a temperature at which the cyclization was complete after 3 hours.

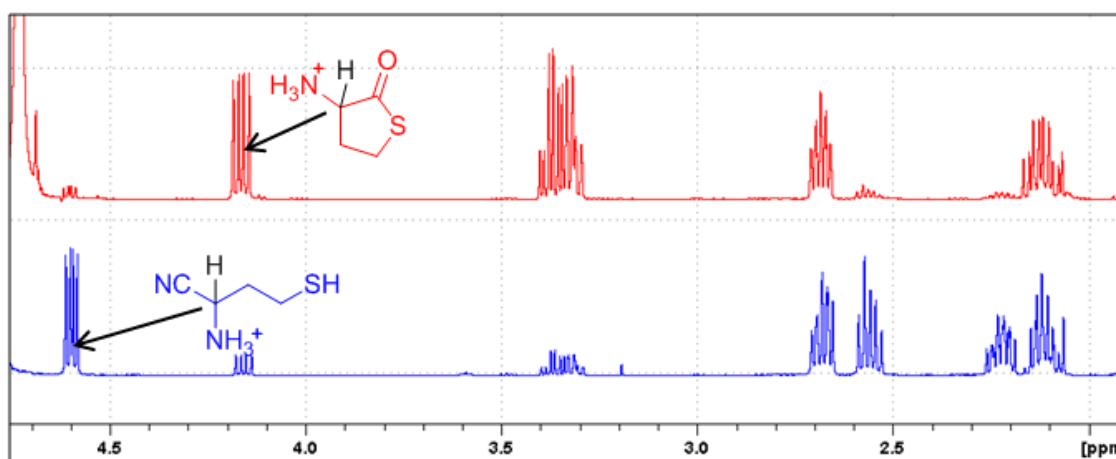


Figure 47: ¹H-NMR monitoring of Hcy-CN cyclization at rt. a) just after deprotection, b) after 3 h

Of course, the observed cyclization is easy and quick due to the formation of a five-membered ring **54**. In contrast, in the case of Cys-CN a similar mechanism would lead to a four-membered thiolactone, and thus is a highly disfavored pathway. The formation of Hcy-thiolactone **54** under our conditions argues for the existence of this thiolactone in the primitive ocean, and it can then be engaged in our scenario.

II.4 Prebiotic behaviour of Hcy-thiolactone

It has been reported in the literature that, under slightly basic conditions, Hcy-thiolactone undergoes a dimerization reaction to form Hcy-Hcy cyclic dipeptide **122**, the diketopiperazine of homocysteine, or disulfide polymers **123** depending on the used conditions (*Figure 48*).¹⁶¹

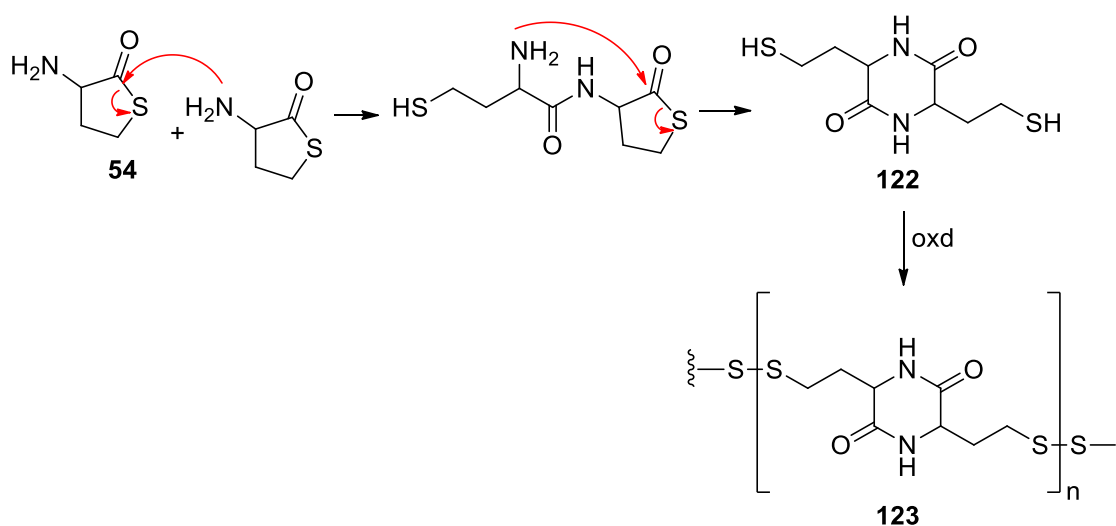
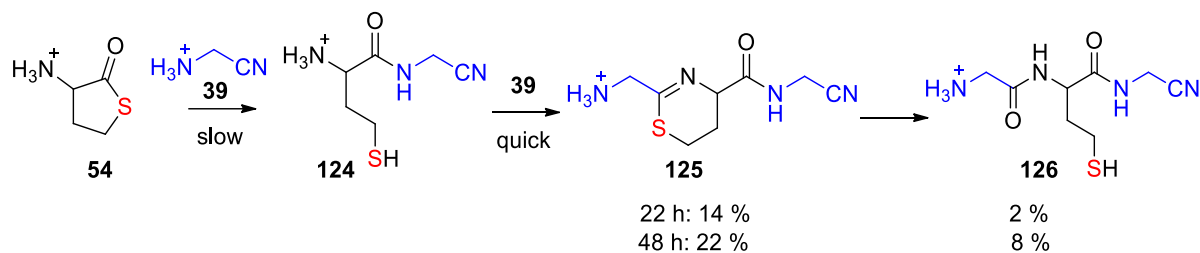


Figure 48: Opening of Hcy-thiolactone ring by another molecule

In spite of its possible dimerization, we were able to condense this thiolactone with aminonitriles. For instance, when we mixed Hcy-thiolactone **54** with aminoacetonitrile (Gly-CN **39**), a product was formed, whose formation can be explained this way: first, the thiolactone ring **54** was opened and Hcy-Gly-CN **124** was formed. Then the obtained product reacted quickly with another equivalent of Gly-CN **39** to form product **125**, a dihydrothiazine cycle which then hydrolysed to form Gly-Hcy-Gly-CN **126**, a precursor of Gly-Hcy-Gly tripeptide (*Scheme 40*). The obtained products were not isolated but characterized in D₂O solution by ¹H-NMR (*Figure 49*)*. This means that Hcy-thiolactone **54** showed a double reactivity, thiolactone opening by amines followed by aminothiols condensation reaction with nitriles. Having a great interest in the formation of Zn complexes (Zn fingers), we repeated this reaction in the presence of Zn²⁺ and we observed a similar reactivity. This indicates that zinc ions have no obvious effect in this particular case.

* As the reaction was accompanied by a decrease in the pH, NaHCO₃ was added stepwise to keep the pH neutral.

Chapter II



Scheme 40

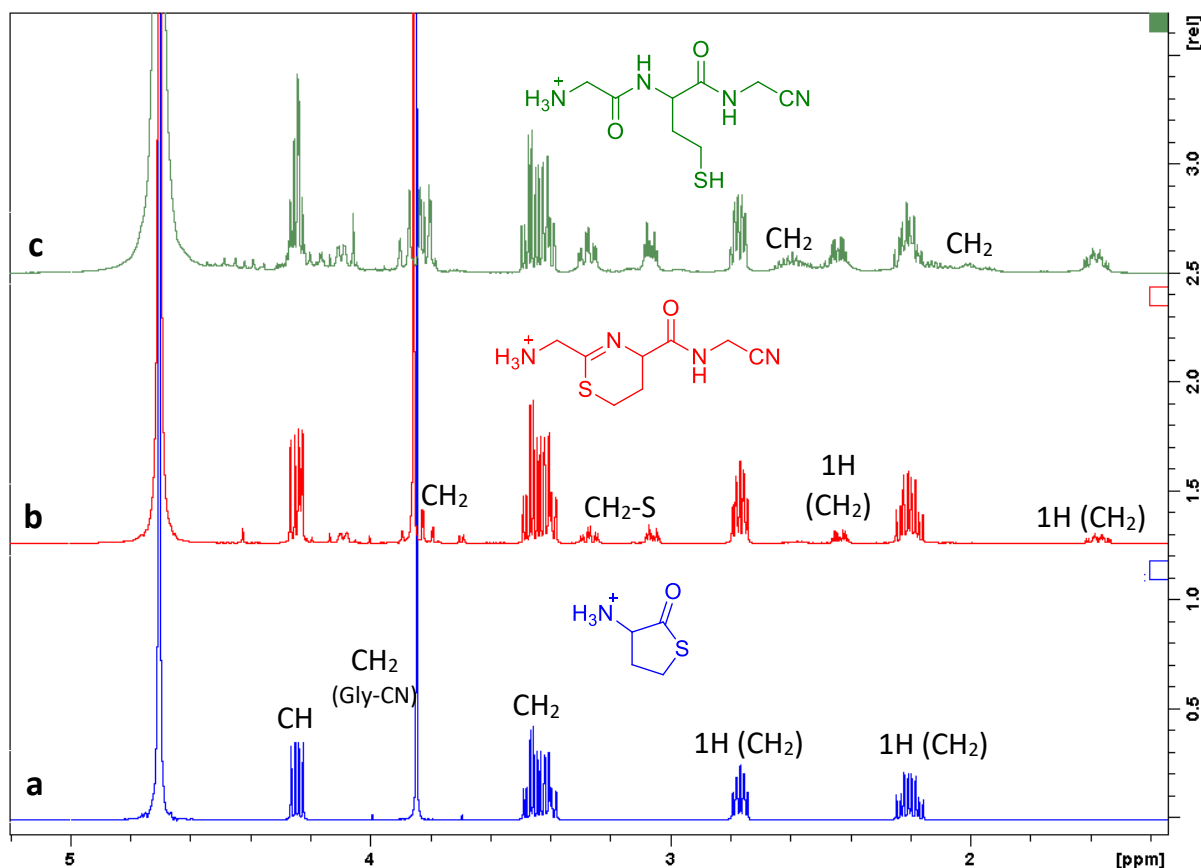
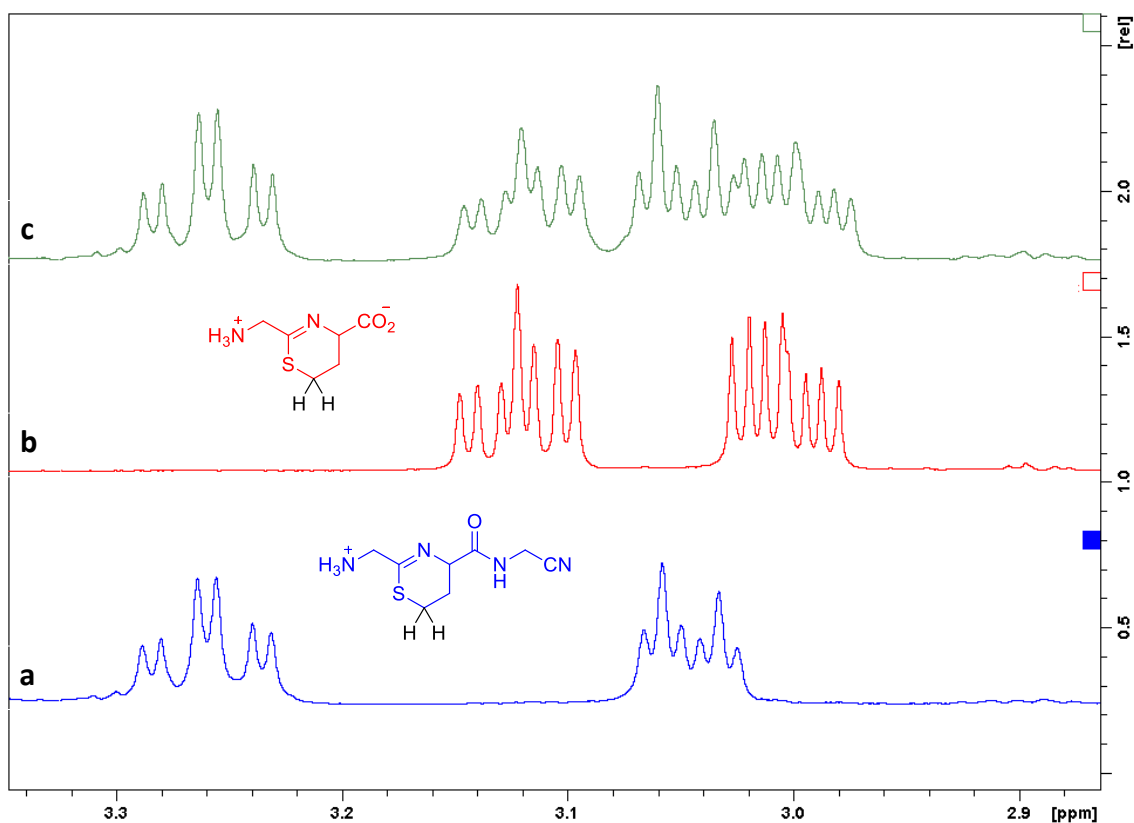


Figure 49: Products were characterized in D₂O solution by ¹H-NMR. a) mixture of starting materials, b) after 22 h, c) after 48 h

We found that Hcy-thiolactone **54** is stable in slightly acidic water at 45 °C for a long time. However, we wanted to make sure that no hydrolysis happened during this reaction. If this was the case, homocysteine, which is the product of the hydrolysis of thiolactone **54**, is expected to react with aminoacetonitrile to give dihydrothiazine cycle **52** (see chapter 1). We wanted to confirm that the observed cycle in the previous NMR is not the cycle **52**.

To do this, we prepared two solutions in which the first contained the cycle **52** (by mixing homocysteine and aminoacetonitrile **39** in D₂O at 45 °C for 2 hours) and the cycle **125** (by mixing Hcy-thiolactone **54** with aminoacetonitrile **39** in D₂O at 45 °C for 3 hours). Mixing the two solutions together confirmed that the observed SCH₂ signals did not belong to the same molecule (*Figure 50*).



*Figure 50: ¹H-NMR of (CH₂-S) protons: a) dihydrothiazine **125**, b) dihydrothiazine **52**, c) a mixture of **125** and **52***

In a similar way, Gly-Gly-Hcy-Gly-Gly-CN **128** was detected when Hcy-thiolactone **54** was reacted with Gly-Gly-CN **42** (*Figure 51*). In the primitive ocean, this nitrile **128** would then either been hydrolysed to obtain the corresponding pentapeptide **129** or it could react with cysteine (for instance) to yield the hexapeptide **130**. As we have demonstrated (see previous chapter, p. 73) that Gly-Gly-CN **42** can be produced from Gly-CN **39** in the primitive ocean, these new results show that tetrapeptides (and longer peptides) could have been produced under “our” prebiotic conditions.

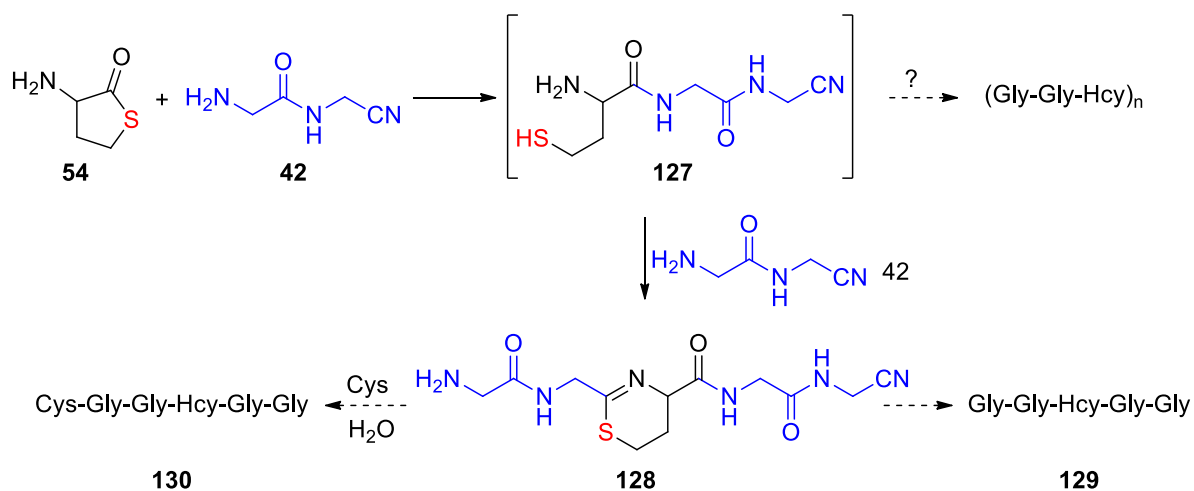


Figure 51: Reaction between Hcy-thiolactone 54 and Gly-Gly-CN 42

According to these results, it can be postulated that Hcy-thiolactone **54** can be considered as a molecular “staple” which would permit the connection of two peptide chains (one containing a nitrile and the other containing an amine) together (*Figure 52*).

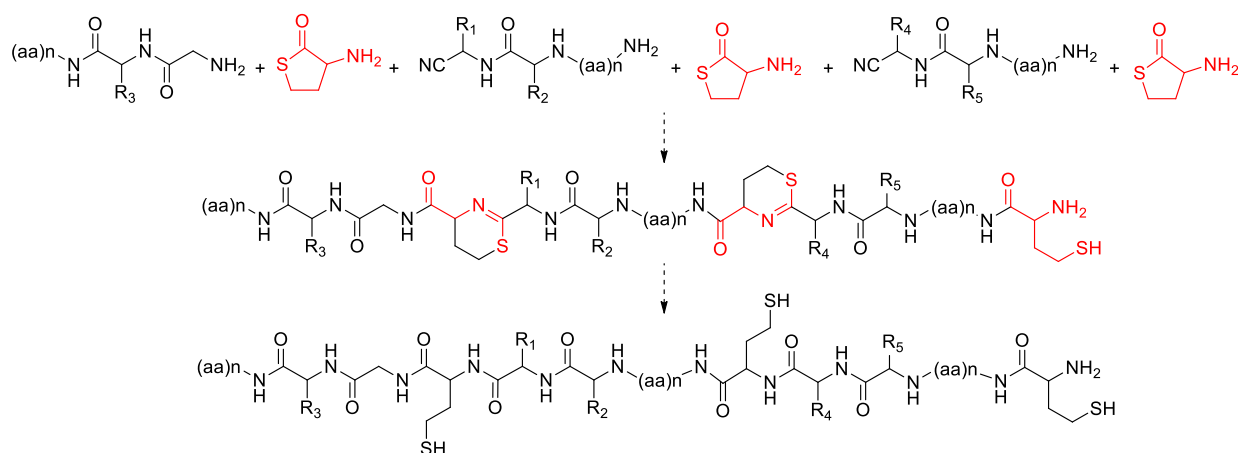


Figure 52: Suggested example for homocysteine thiolactone as a molecular "staple"

II.5 Aminoacyl transfer via a thioester

The reactivity of the obtained tripeptide **126** has been studied. Similarly to Gly-Cys **47** and Gly-Hcy **53**, Gly-Hcy-Gly-CN **126** is a thiol-containing small peptide and can be expected to give thioesters when reacted with nitriles. From the obtained ¹³C-NMR of the mixture, we were able to detect the formation of a thioester, most probably compound **132**.

Similarly to the previous examples, we studied the reactivity of this thioester **132**. In this case, thioester **132** reacted with the amino group of the starting thiolactone **54** to give Gly-Hcy-thiolactone **133** which was the dominant product resulting from the thioester (*Figure 53*). This is our second example of an “aminoacyl” transfer via a thioester (after Gly-Gly-CN **42** formation). Indeed, Gly-CN **39** and Hcy-thiolactone **54** have similar pKa’s (near 6), so that a significant amount of free amines are present in both cases under our conditions. However, the expected Gly-GlyCN **42** was not observed in this reaction mixture but only the thiolactone **133**.

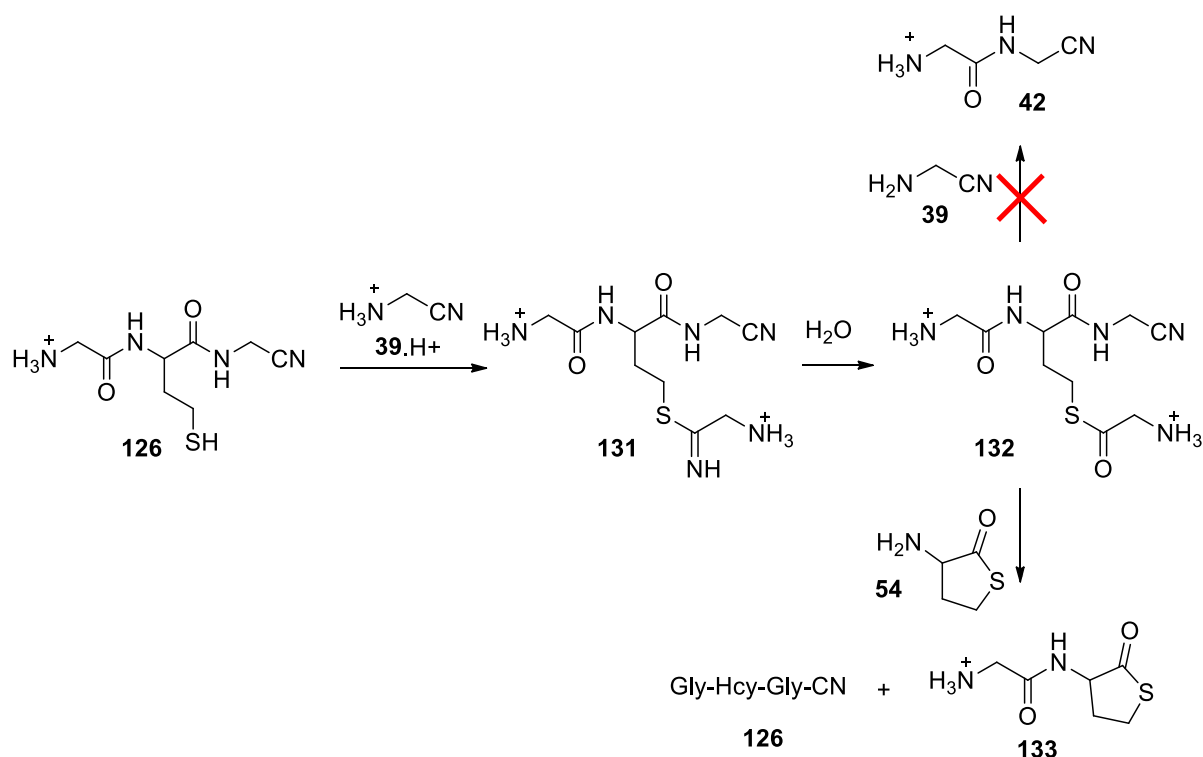
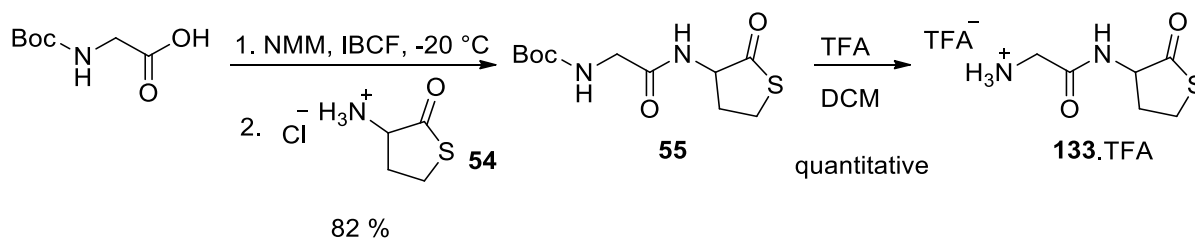


Figure 53: “Aminoacyl” transfer via a thioester

This result gives extra evidence that our thiol-containing peptides are able to promote the formation of peptide bonds via thioester bond formation and thus might have played a role in a “thioester world”.

We synthesized an authentic sample of Gly-Hcy-thiolactone **133** in order to compare its NMR spectra with the spectra of our prebiotic mixtures (*Scheme 41*).

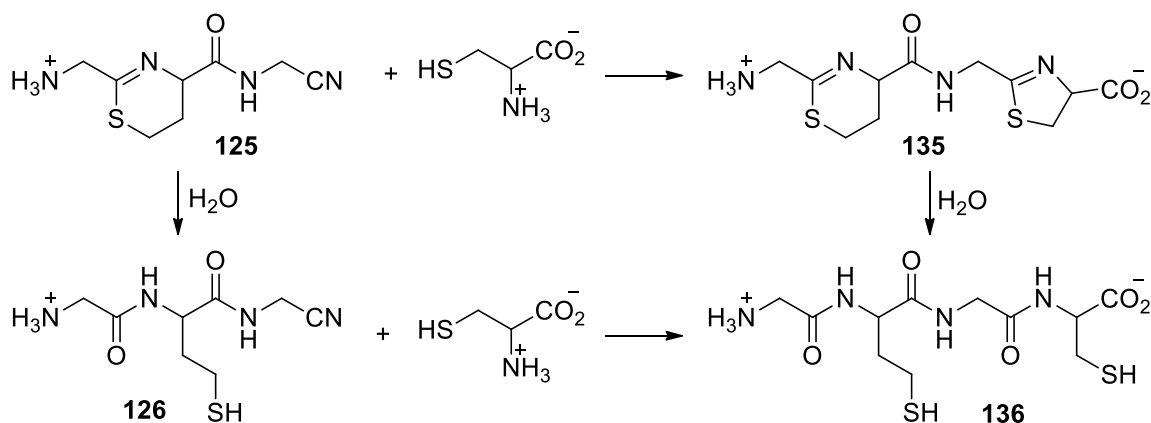


Scheme 41

The reactivity of the obtained Gly-Hcy-thiolactone **133** has also been studied. As the amine of Hcy thiolactone is poorly basic, we postulated that this amine could act as leaving group (which is of course unusual for amines). If so, **133** would be an activated form of Glycine and therefore, for instance, Gly-Cys would be formed from it in the presence of cysteine. In order to test this idea, cysteine was added to a pure sample of this thiolactone in D₂O. However, no Gly-Cys was formed.

II.6 Formation of bis-thiol tetrapeptides

The reactivity of the obtained nitriles **125** and **126** towards aminothiols group has been studied. It has been found that upon the addition of cysteine to a mixture of these nitriles, the bis-thiol tetrapeptide Gly-Hcy-Gly-Cys **136** was formed. It has been characterized by ¹H-NMR (Figure 54). When both nitriles were present, the thiazine-containing tripeptide **125** was found to be more reactive toward the formation of tetrapeptide **136** than the hydrolysed one **126**.

Figure 54: Prebiotic formation of bis-thiol tetrapeptide Gly-Hcy-Gly-Cys **136**

Similarly, Gly-Hcy-Gly-Hcy **138** was formed when Hcy was added (*Figure 55*).

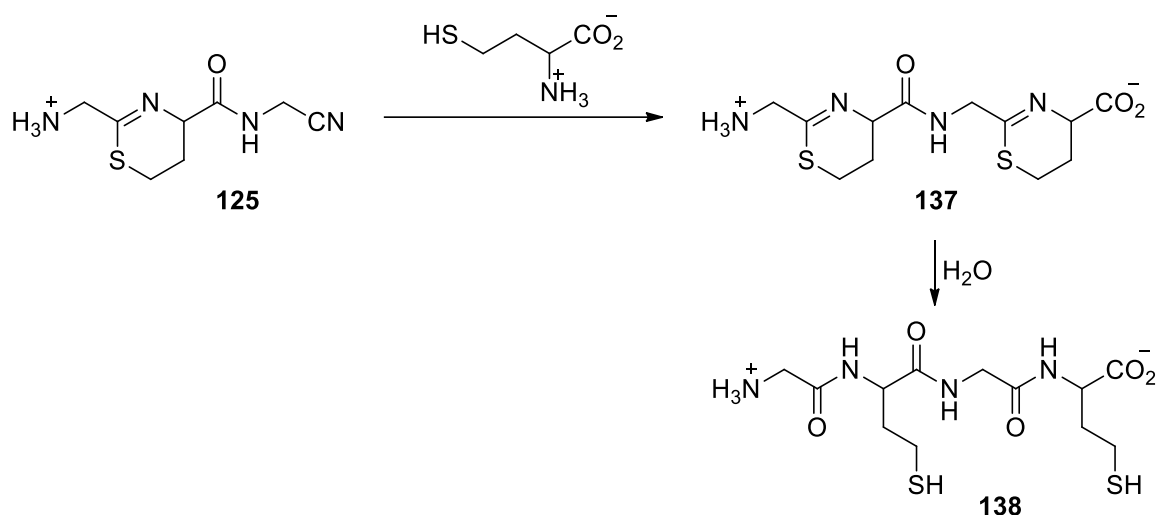


Figure 55: Prebiotic formation of bis-thiol tetrapeptide Gly-Hcy-Gly-Hcy 138

After some days, these bis-thiols were oxidized to form intramolecular sulfur bridges (*Figure 56*). However, when we analysed our disulfide samples by high resolution mass spectrometry (HMRS), we found that the detected exact masses were not those of the expected acids but those of the corresponding amides (*Figure 56*).

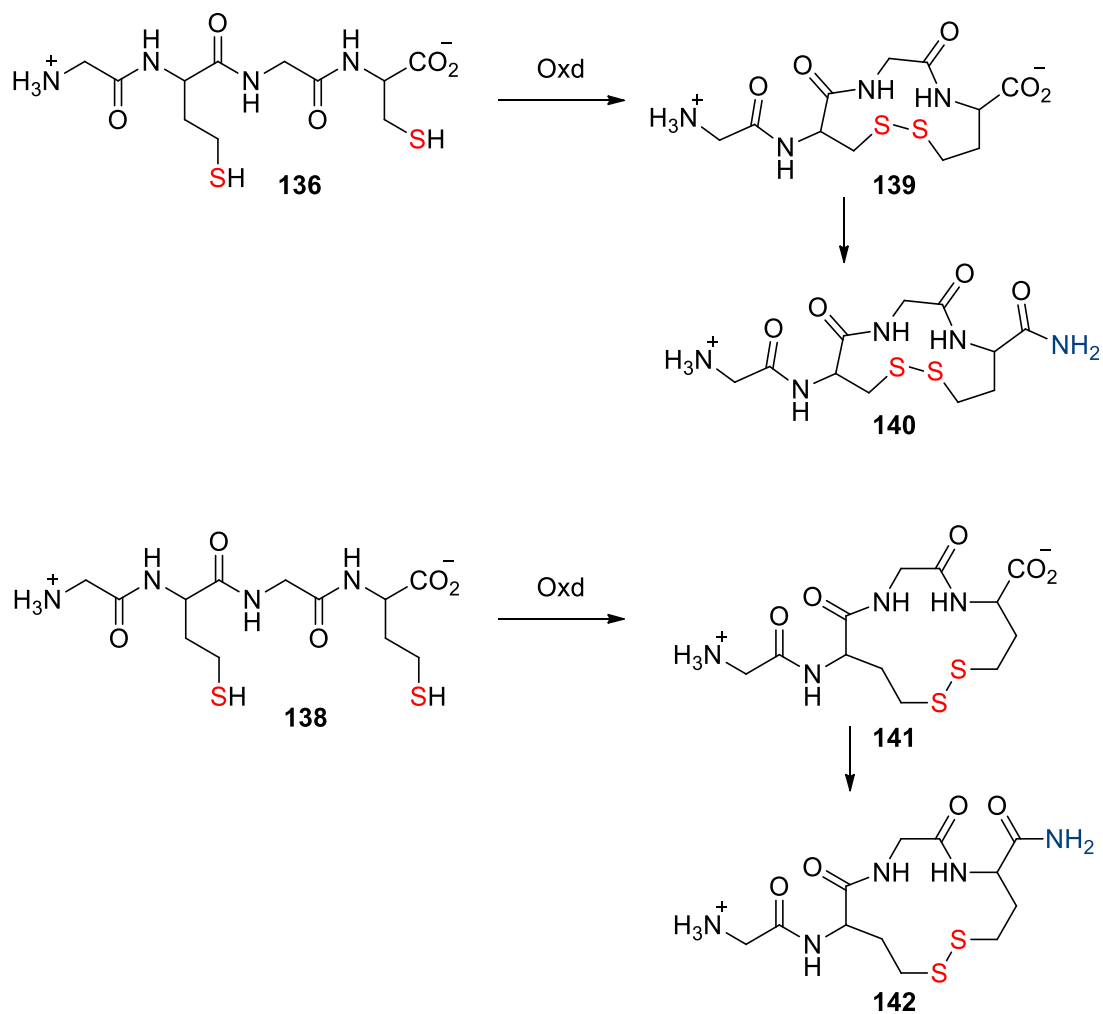


Figure 56: Intramolecular (S-S) bond formation

We have no definitive explanation for this unexpected result. Indeed, the acid to amide transformation necessitate an activation of the acid. A very tentative mechanism is presented in the following scheme (given from Gly-Hcy-Gly-Hcy as an example) (*Figure 57*).

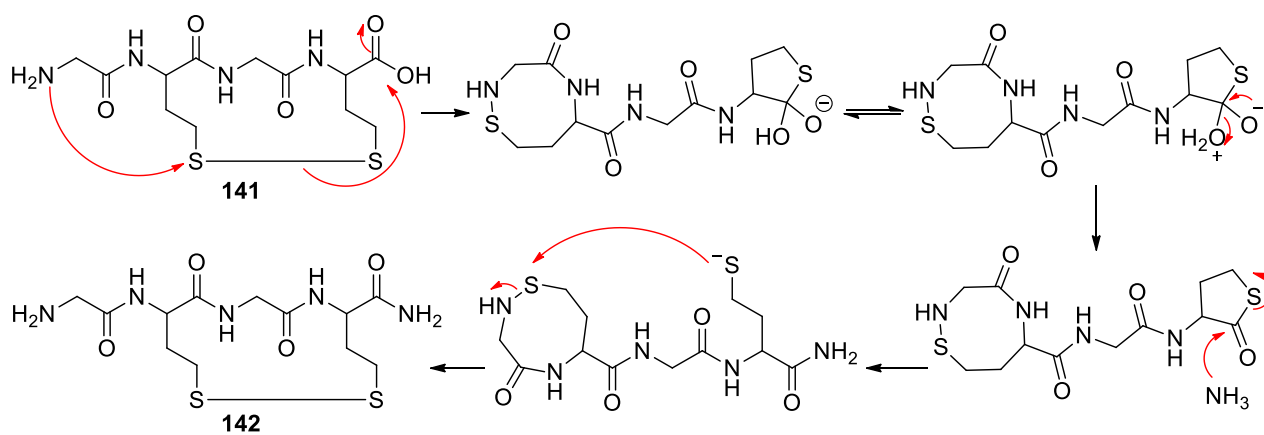


Figure 57: Tentative explanation of amide formation

Chapter III: Prebiotic formation of polycysteine

III.1 Introduction

Cysteine residues play a major role in proteins. This is due in part to the fact that thiol group in cysteine is deprotonated in physiological pH (pK_a of cysteine's thiol range from 7.4 to 9.1)¹⁶². Thus, the thiol group presents as thiolate form, which is considered as one of the most reactive function in peptides. Furthermore, the presence of thiols allows cysteine residues to participate in protein folding and redox reactions via the formation of disulfide bonds. They are involved in the formation of metalloproteins, being part of iron clusters, for instance in ferredoxins (four-cysteine residues involved)¹⁰¹ and of zinc fingers (in which up to six-cysteine units are involved) which are important in many DNA, RNA, protein interactions.¹⁶³ Even more cysteine residues are found in proteins. For instance, metallothioneins are small proteins containing about 60 residues of which up to 20 are cysteine units units.¹⁶⁴ They can complex zinc cations, and other metals such as copper and cadmium.^{165,166} The presence of cysteine in antioxidant glutathione and its use in the synthesis of coenzyme A are other examples of the widespread scope of this sulfur amino acid.¹³²

Prebiotic synthesis of cysteine is not so obvious. Still, two plausible pathways starting from glycine might have participated in this synthesis (*Figure 58*).⁶¹

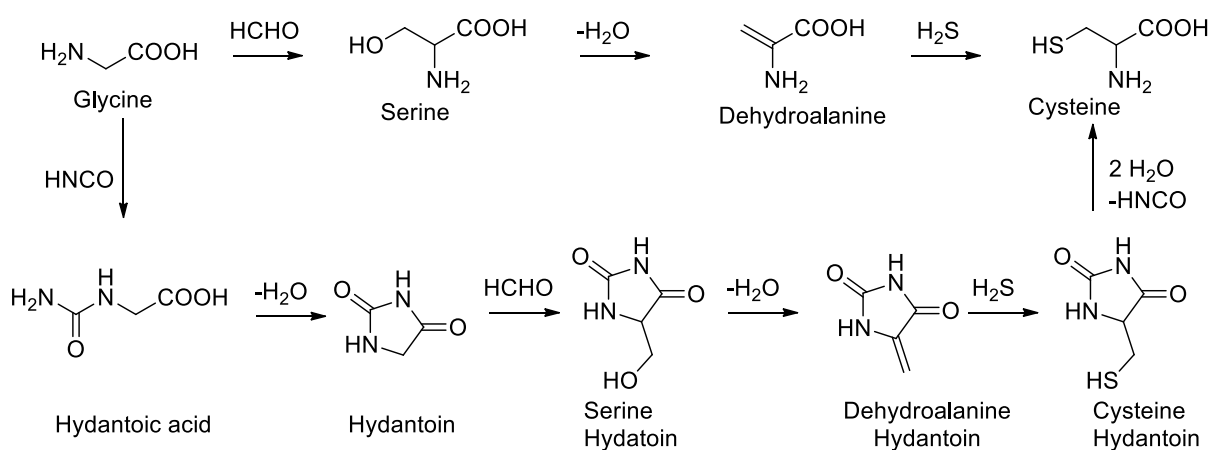


Figure 58: Two proposed mechanisms for prebiotic synthesis of cysteine starting from Glycine

In addition to the previous possibilities, cysteine might have also been formed via Strecker synthesis similarly to homocysteine (*Figure 59*).

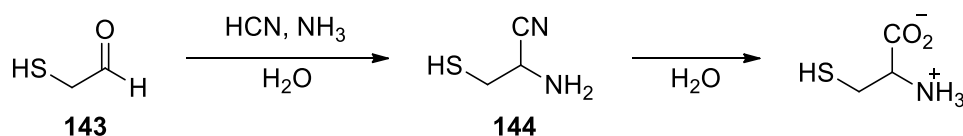


Figure 59: Cysteine Strecker synthesis

We have previously argued that cysteine residues in peptides may have participated in important catalytic sites when life appeared on Earth. The peptide richest in cysteine is polycysteine, Cys_n.¹⁶⁷ In this chapter, we present our results about synthetic ways to form this polymer. Of course, it is hard to believe that long chain polycysteines were formed in the prebiotic ocean. However, the presence of many cysteine residues in more diversified proteins are plausible, and the regular cysteine polymer can be regarded as a simplified (even though longer) model for such proteins.

To form polycysteine under plausible prebiotic conditions, in water, we needed activated cysteine derivatives. We choose to synthesize nitrile **144**, thioacid **145** and three thioesters **146**, **147** and **148** as plausible models that might be presented in the ocean 4 billion years ago (*Figure 60*).

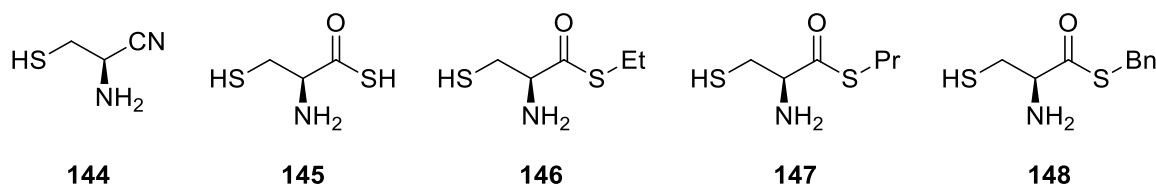


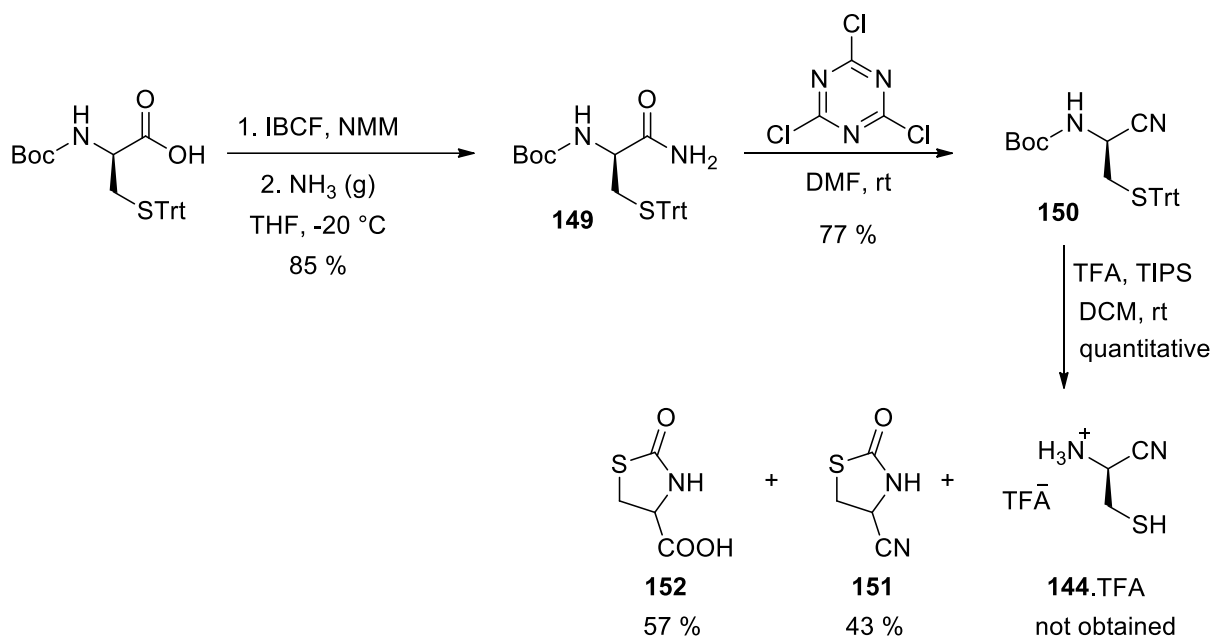
Figure 60: Activated cysteine derivatives

III.2 Synthesis of polycysteine precursors

III.2.1 Synthesis of cysteine nitrile 144

We first proposed a pathway toward the synthesis of cysteine nitrile starting from Boc-Cys(Trt)-OH as described in (*Scheme 42*). The advantage of starting from this protected cysteine is that the removal of both Boc group on the amine and the Trt group on the thiol would be performed in a single step. We successfully obtained the amide **149**, then the nitrile

150, with good yields. TFA was added to **150** and this deprotection process was followed by extraction in D₂O. However, the last step did not provide cysteine nitrile **144** as we expected. Instead, a mixture was obtained. The ¹³C-NMR spectra of this mixture allowed us to identify the cyclic nitrile **151** beside its hydrolysis product **152** (*Scheme 42*). Upon time, the nitrile **151** disappeared and the acid **152** became dominant in the mixture.



Scheme 42

This result was attributed to the removal of the trityl group before Boc leading to molecule **153**. The free thiol in **153** was able to react on the carbamate group and cyclized rapidly to give thiocarbamate **151**, which hydrolysed fast to obtain the corresponding acid **152** (*Figure 61*).

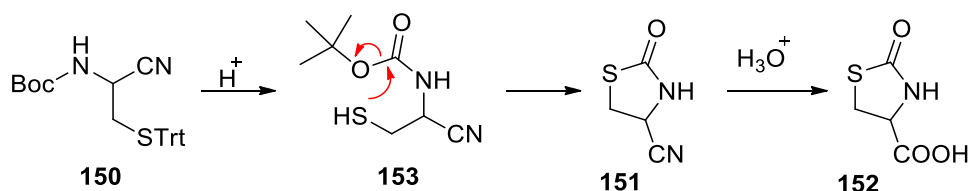
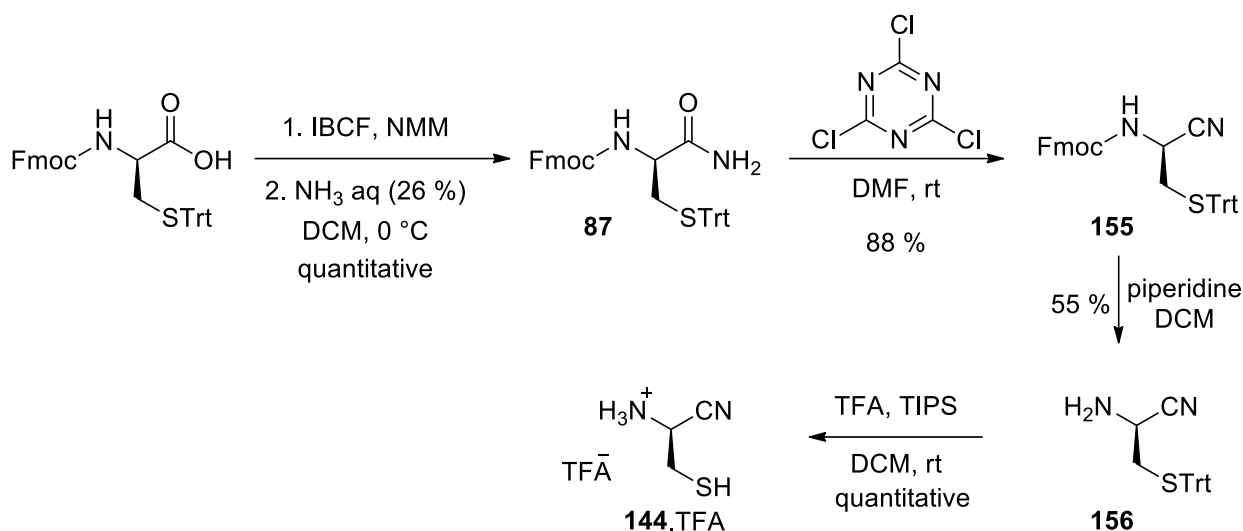


Figure 61: Explanation of the unexpected cyclisation

In order to avoid the formation of these by-products, we replaced the Boc group by a Fmoc protecting group. Starting from Fmoc-Cys(Trt)-OH, we were able to form the corresponding

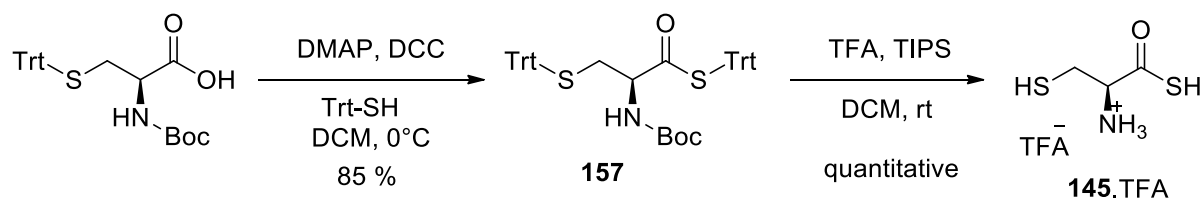
amide **87** and nitrile **155** with good yields. After removal of Fmoc group with 20 % piperidine solution in DCM to give Cys(Trt)-CN **156**, the cleavage of the Boc group was performed upon addition of TFA to obtain the free thiol **144**. Cys-CN **144** was extracted in water (or D₂O) for testing its behaviour under our primitive conditions (*Scheme 43*).



Scheme 43

III.2.2 Synthesis of cysteine thioacid **145**

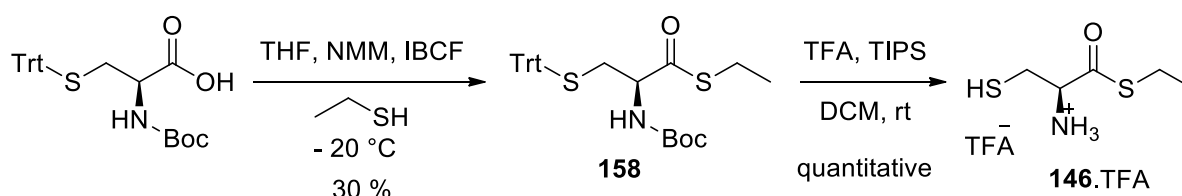
We were also interested in the study of the behaviour of the thioacid of cysteine **145**. To synthesize it, we started from Boc,Trt-protected cysteine. The acid was first activated with DCC, and then triphenylmethanethiol was introduced to provide the thioester **157**. Finally, the product **145** containing the free amine, thiol, and thioacid functions was obtained in a single deprotection step (*Scheme 44*). In this case, no hydrolysis to thiocarbamate cycle was noticed.



Scheme 44

III.2.3 Synthesis of cysteine ethyl thioester **146**

In order to obtain cysteine ethyl thioester, Boc-Cys(Trt)-OH was activated as a mixed anhydride and ethanethiol was then introduced to provide Boc-Cys(Trt)-SEt **158**. Then, the product **146** was formed in a single deprotection step in which protecting groups on cysteine were successfully removed without damaging the thioester group (*Scheme 45*).



Scheme 45

However, in this case a by-product was formed: the ester **159**, whose formation is explained in *figure 62*. The R_f , rates of flow on TLC paper, for product **158** was very close to that of this ester. For this reason, we were not able to obtain the product **158** in high purity. Indeed, having such an impurity in our “prebiotic soup” should not be a problem, as the ester **159** will not polymerize. However, we tried to overcome this difficulty by repeating the reaction with the use of excess amount of ethanethiol for a short duration (15 minutes); however, similar results were obtained, that is to say, the alcohol always reacted with the thioester.

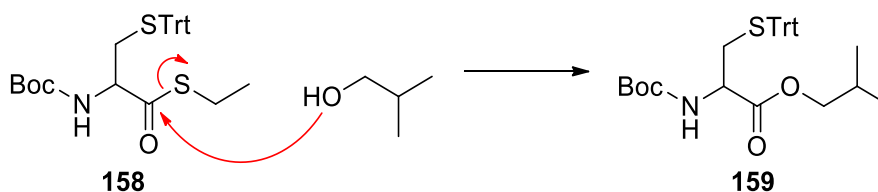
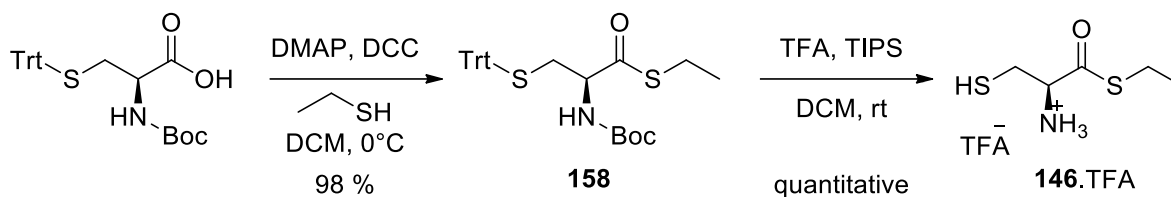


Figure 62: Formation of the undesired ester 159

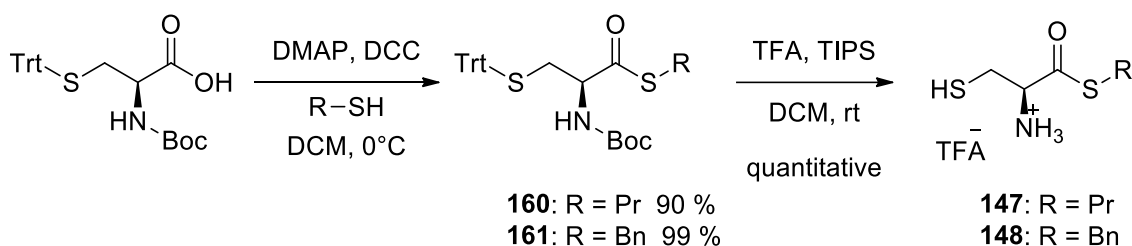
In order to obtain the thioester **158** in pure form, we intended to avoid the formation of isobutyl alcohol. To do so, we activated the acid using DCC with catalytic amount of DMAP. This way, the thioester **158** was obtained in a good yield and a good purity (*Scheme 46*).



Scheme 46

III.2.4 Synthesis of cysteine propyl thioester **147**, and cysteine benzyl thioester **148**

The same method was used to synthesize propyl thioester **147** and benzyl thioester **148** in a high yield. Despite the fact that cysteine benzyl thioester **148** can hardly be regarded as a primitive component, we considered it as an additional example that might lead to polycysteine formation in water. Both protecting groups (Boc and Trt) were removed in one step to provide the final cysteine thioesters **147** and **148** (Scheme 47).



Scheme 47

III.3 Formation of polycysteine

III.3.1 Starting from cysteine nitrile

In the first chapter, we have discussed the possibility to form a dipeptide between two amino acid precursors in which one holds a nitrile group and the other has an aminothioli moiety. Cysteine nitrile contains both nitrile and aminothioli functions, it should be able to polymerize. The nitrile precursor of cysteine is also interesting as it is a plausible interstellar molecule.⁵⁵

Indeed, if Cys-CN **144** was found to be stable when it was kept in acidic water (pH 1), the formation of its dimer **162** was observed as soon as pH was fixed at 5 – 6. Another product was observed to which we tentatively assigned a trimeric structure **163**, and after few hours, a non-soluble white polymer precipitated out of the solution (Scheme 48).

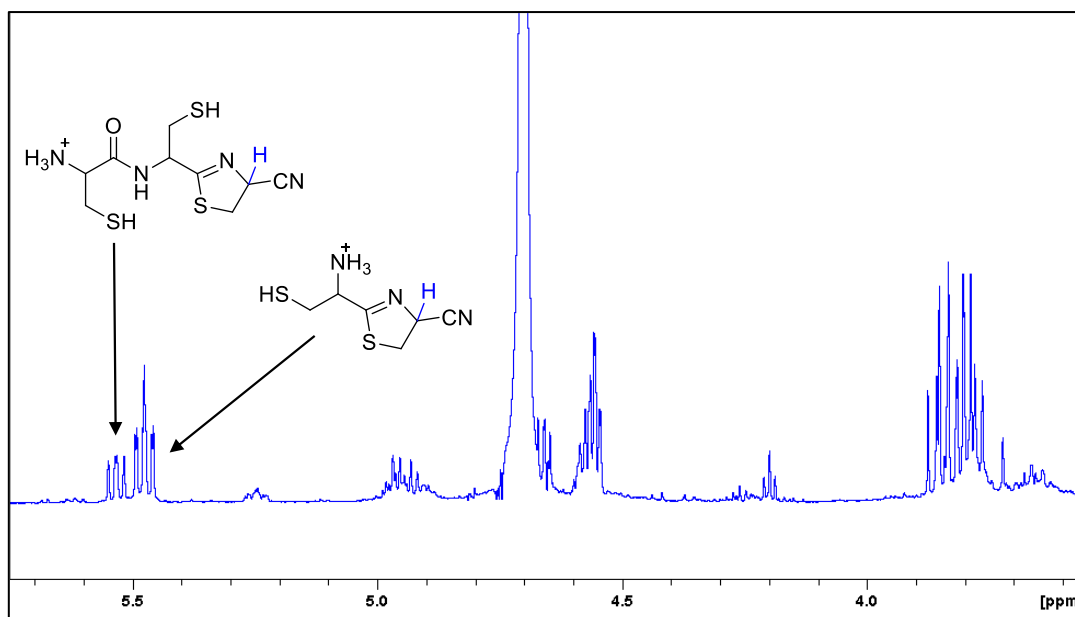


Figure 64: $^1\text{H-NMR}$ spectrum showing the dimer and the trimer of cysteine residues

In this experiment, we were not able to characterize the dimer **165** with two adjacent thiazolines. We believe that it was not formed. The results indicated that the thiazoline of the dimer **162** was first hydrolysed to give the dimer **164** before it reacted with another Cys-CN unit (at least in this experiment) (Figure 65).

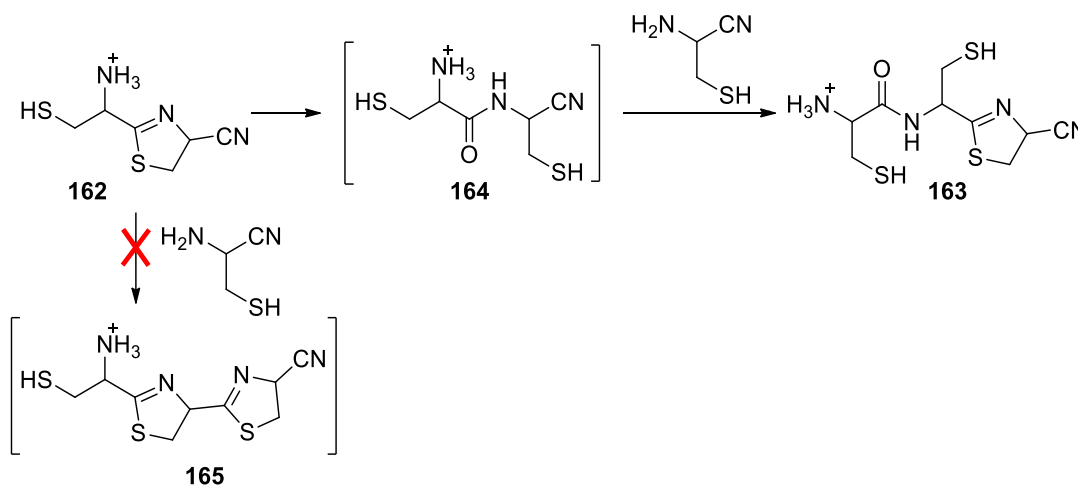


Figure 65: Formation of the trimer **163** via the dimer **164**

The reaction did not stop at this level of oligomerization. The monomer, dimer, and trimer, all contain nitrile and aminothiols moieties and thus reacted further to give longer chains. These longer chains were not soluble in D_2O and thus started to precipitate.

We wanted to study this polymer using MALDI-TOF mass spectrometry. So, we repeated the reaction in degassed (to avoid as much as possible any disulfide bridge formation) H₂O (to avoid any deuteration). In order to prepare the sample for mass analysis, the precipitate was washed several times with water to remove traces of TFA, before drying. The obtained solid was then dissolved in DMSO and a MALDI analysis was undertaken.

The expected polycysteine masses should give values that are separated by 103, the mass of a cysteine residue. However, the observed results did not agree with our expectation, the difference between the observed masses being 192, so much higher than 103 (without being 206, which would correspond to 2 Cys residues)

Believing that the polymer consisted of polycysteine assemblies, we tried to interpret our results by proposing some plausible structures, the question being: what is this the mass 192 related to? This value could be explained starting with two thiazole rings, hence two cysteine units, with mass equal to 170. These “dimers” could lose a proton and each formed anion would be neutralized by a sodium cation: $170 - 1 + 23 = 192$.

Therefore, we propose that the observed masses may correspond to the chelation of polythiazoline rings with sodium cations in which each ion chelates with two cycles (*Figure 66*). The presence of conjugated thiazoline rings would help the formation of this repetitive structure. If we are true, the polymer was thus mainly composed of these rings.

Two examples of these structures **166** (with 10 thiazoline rings) and **167** (with 12 thiazoline rings) are shown in *figure 66*. (Indeed our proposed structure for trimer **163** implies the hydrolysis of the ring).

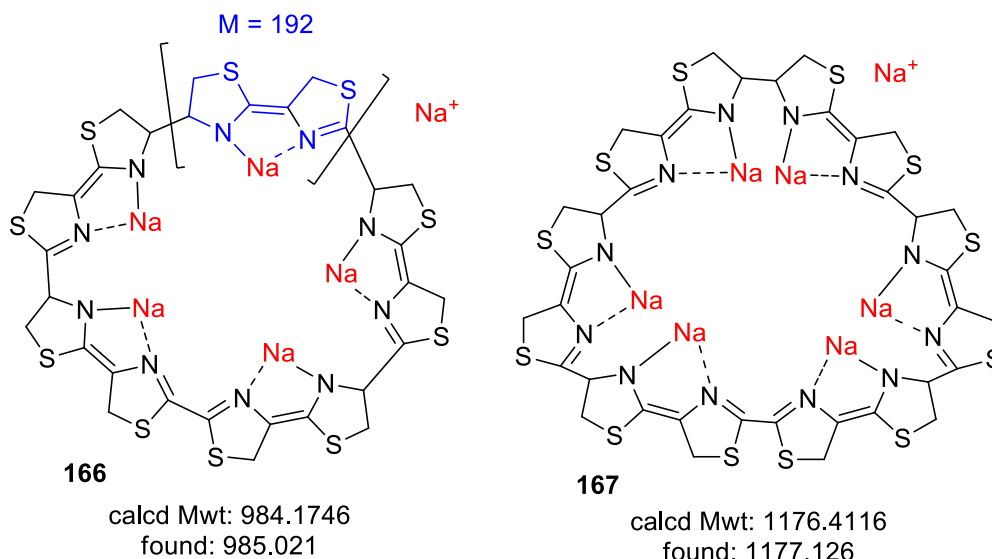


Figure 66: Proposed structures of the observed masses (166 for $M = 985.021$ and 167 for $M = 1177.126$)

Indeed, some ring hydrolysis might have occurred, then some thiols could have been oxidized to disulfides, facts that would explain some other masses that we observed.

Polythiazoline structures have already been described in the literature. For example, linear polythiazolines were obtained from polythioureas.¹⁶⁸ Cyclic oligothiazolines were formed by the head-to-tail cyclooligomerization of doubly deprotected linear fragments.¹⁶⁹ Moreover, poly(thiazole) amino acids, which are not far from our rings, are found in nature. The biosynthesis of thiazole itself can be achieved through cyclo-dehydration of cysteine moiety of cysteine-containing peptides followed by an oxidation step, hence via thiazoline structures (*Figure 67*).¹⁷⁰

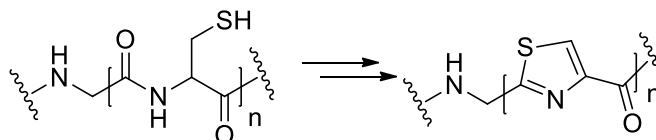


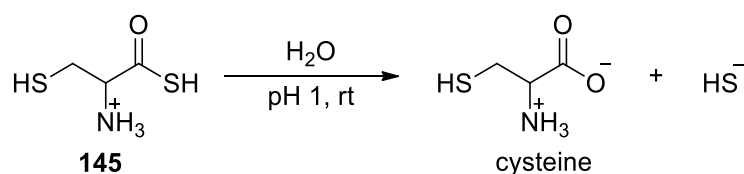
Figure 67 : Biosynthesis of thiazole peptides

However, we still consider our poly(bis-thiazole-Na) structure as tentative. Further experiments should confirm this proposal.

III.3.2 Starting from thioacid

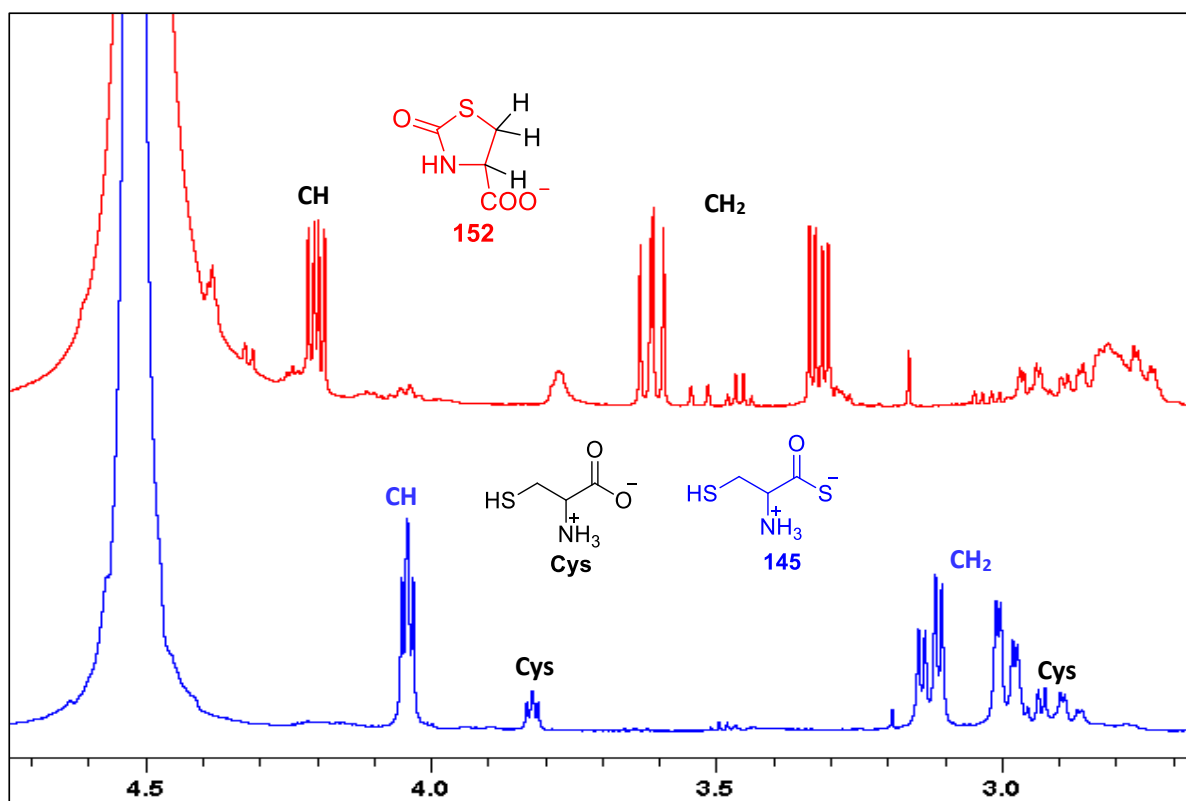
Like any other proteinogenic amino acid, cysteine alone does not significantly oligomerize in water. This can be explained by the fact that the hydroxyl group is a bad leaving group. In contrast, the sulfhydryl group is a good leaving group. Thus, we expected that replacing cysteine by its thioacid analogue would allow an easy polycysteine formation.

However, in contrast with what was observed for cysteine nitrile, we found that cysteine thioacid **145** is not stable in strong acidic conditions (pH 1). Under these conditions the thioacid was completely hydrolysed into cysteine within few hours (*Scheme 49*).



Scheme 49

To avoid this hydrolysis we extracted a newly formed thioacid **145** in D₂O and quickly added sodium carbonate to reach pH 7. Under these neutral conditions, the hydrolysis process was slower, but another product was formed to which we assigned the cyclic structure **152** (*Figure 68*).



*Figure 68: ¹H-NMR demonstrated the formation of the cyclic molecule **152** resulted from cysteine thioacid **145** at pH 7*

Indeed, we did not expect to obtain such structure which is not formed when carbonate and cysteine are mixed. Thus, we believe that the presence of hydrosulfide ions, which were released in the solution upon thioacid hydrolysis, promoted the observed process (*Figure 69-a*). A plausible mechanism is presented in (*Figure 69-b*). We assume that some thiocarbonate and/or dithiocarbonate ions were formed via the reaction of carbonate ions with HS⁻. These thio- and dithio-carbonic acid ions (or some O=C=S resulting from their decomposition) then reacted with the thiol and amine functions of cysteine to give the cyclic thiocarbamate **152**.

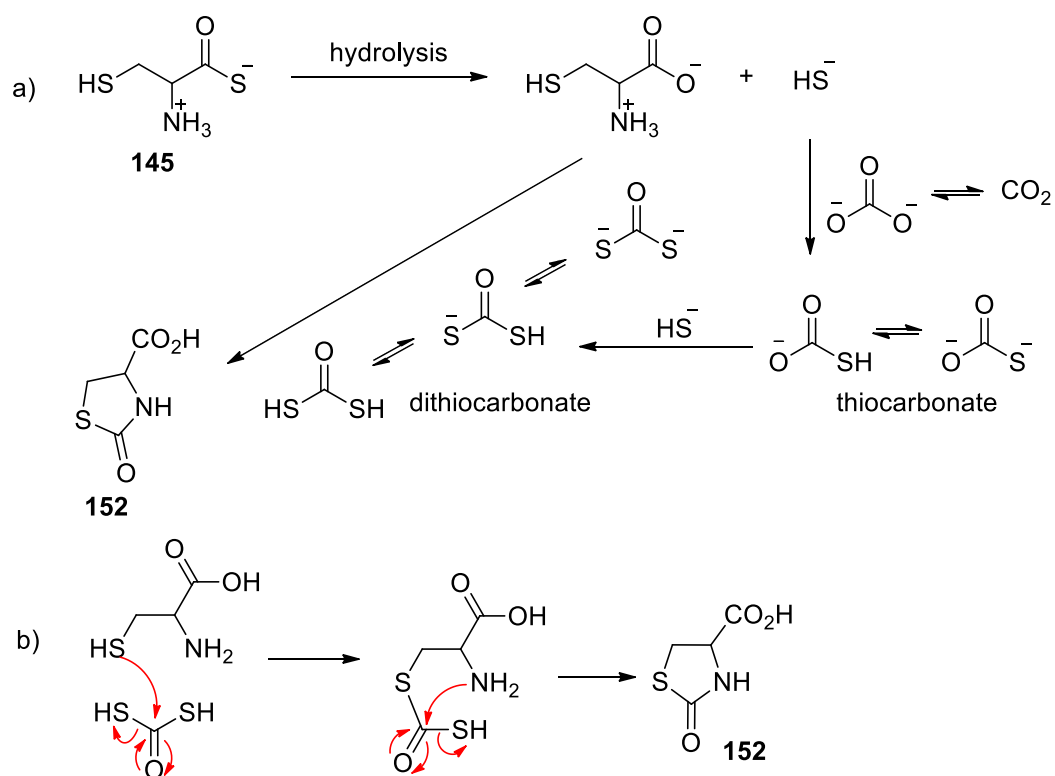


Figure 69: Formation of the cyclic thiocarbamate 152 starting from cysteine thioacid 145

III.3.3 Starting from cysteine thioesters

Thioesters are well-known activated forms of amino acids. A thioester function would react with an aminothiols nucleophile to give a peptide bond. Most probably, the thiol group would start the nucleophilic substitution process on the thioester, and then the amine reacts to form the peptide bond (*Figure 70*). This process is used in the native ligation procedure to synthesize polypeptides.¹⁷¹ Starting from any thioester of cysteine, polycysteine should be formed. For this reason, we synthesized ethyl, propyl and benzyl thioesters. Unlike cysteine nitriles, no thiazoline rings are expected in this case. The predicted masses for the polymer should be separated by one cysteine unit (103).

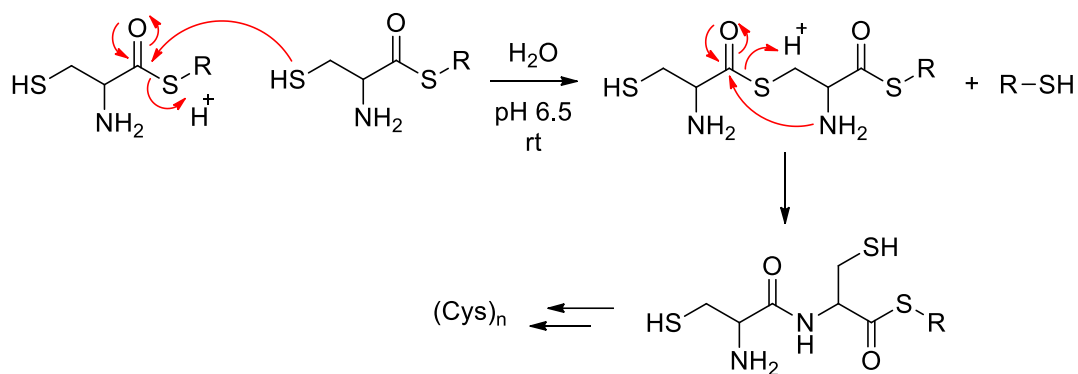
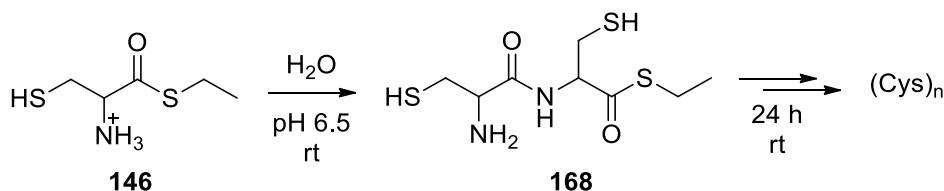


Figure 70: Cysteine thioester is a way to obtain polycysteine

III.3.3.a Cysteine ethyl thioester

Similarly to cysteine nitrile, cysteine ethyl thioester was stable in acidic D₂O. When we changed the pH to slightly acidic (around 6), we noticed on ¹³C-NMR the disappearance of the starting thioester **146**, and the appearance of new C=O peaks, both amides and thioesters. After few hours, the solution in the NMR tube became turbid, indicating the formation of an insoluble polymer, most probably polycysteine. We repeated this process in water in which the precipitated solid was recovered after 24 hours for mass spectrometry analysis (*Scheme 50*).



Scheme 50

In accordance with our expectations, MALDI analysis showed nice series of masses corresponding to cysteine polymer, clearly separated by 103 (*Figure 71*). This time, only single peaks were observed due to the absence of thiazoline rings.

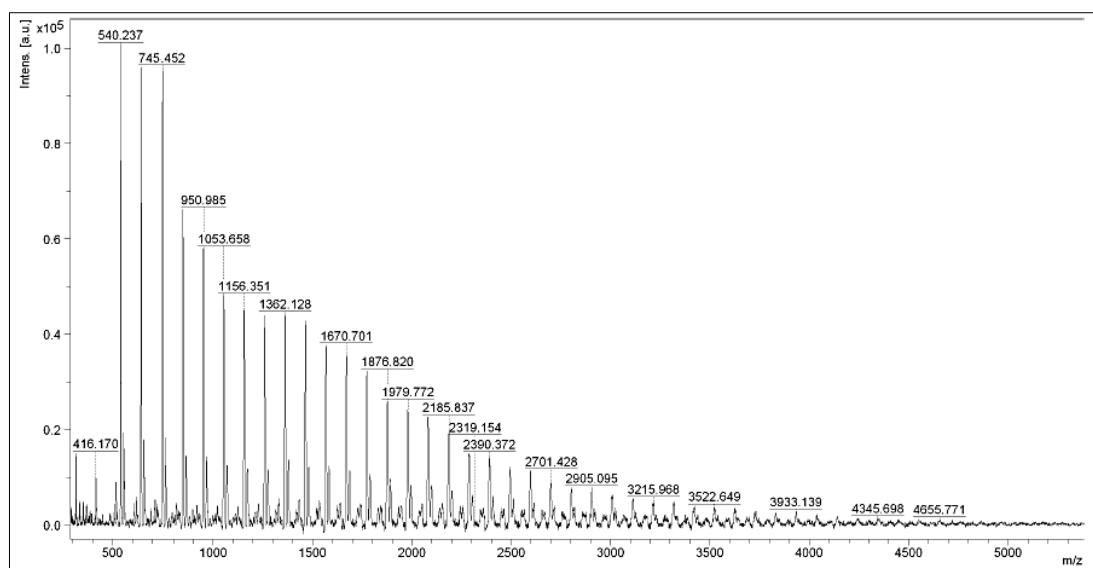


Figure 71: MALDI MS-spectrometry of polycysteine

The obtained masses may correspond to cyclic polycysteine structures. For example, the observed peak 2493.134 matches with **169**, (Cys)₂₄ (calculated exact mass = 2495.210). The observed difference between these values can be attributed to the presence of one or two disulfide (S-S) bonds (*Figure 72*).

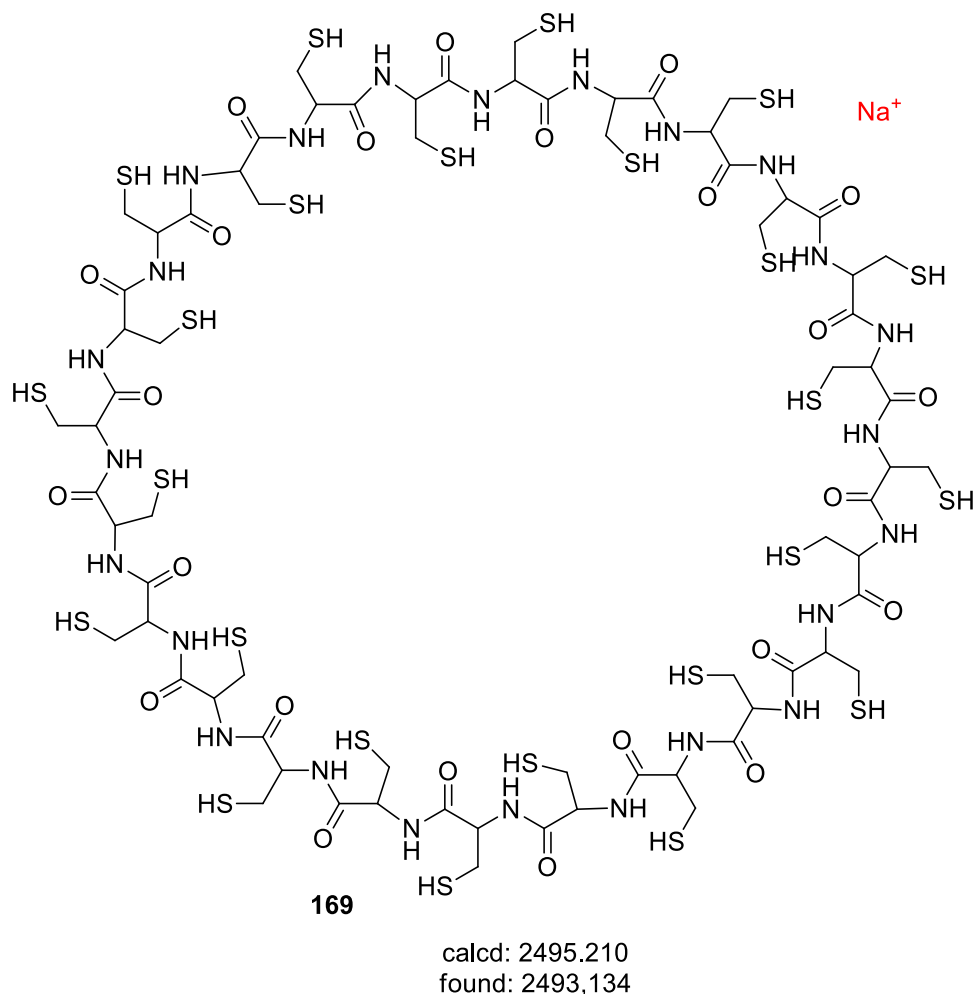


Figure 72: A proposed polycysteine structure for $M = 2493.134$

There is no doubt that the formation of small cyclic structures is possible. However, we believe that the presence of “big” macrocyclic polycysteine structures is uncertain. Therefore, we presented an alternative possible structure for the observed “bigger” masses. We proposed that small cyclic structures, considering that they have free thiols, might form thioester bonds. For instance, the previous value (2493.134) may correspond to **170**, a cyclo(Cys)₅ in which each thiol holds various cysteine chains via a thioester bond (*Figure 73*).

Of course, each mass may correspond to several possible (but similar) structures. That it so say, one cyclic polycysteine, or a cycle that link with other cysteine residues.

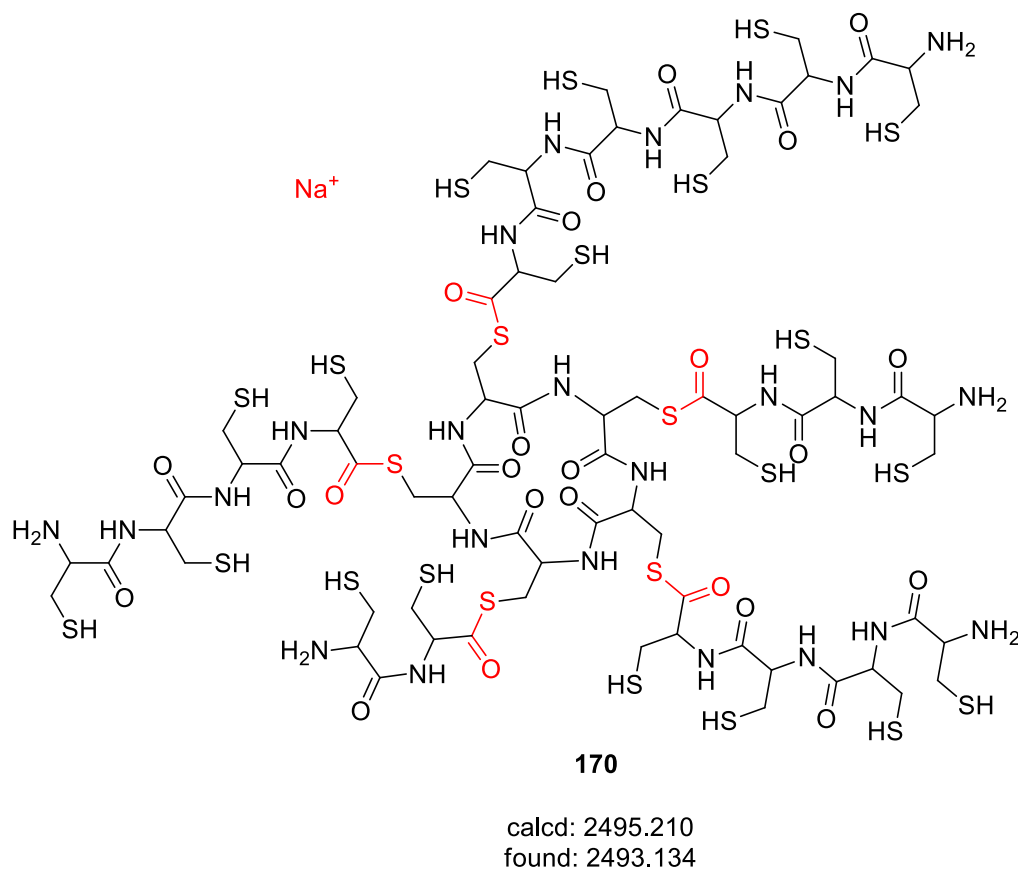
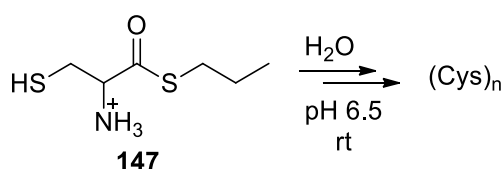


Figure 73: A proposed polycysteine structure for $M = 2493.134$

III.3.3.b Cysteine propyl and cysteine benzyl thioesters

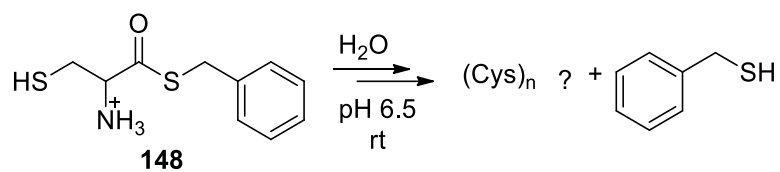
Cysteine propyl thioester **147** was also stable in acidic water. Polycysteine started to precipitate out of the solution around neutral pH at room temperature (*Scheme 51*). However, MALDI analysis did not give sufficient results in this case. This might be because of very low concentrations for example.



Scheme 51

When benzyl thioester **148** was synthesized and kept in water under our prebiotic conditions, some turbidity was observed in the mixture indicating the formation of insoluble products (*Scheme 52*). However, the precipitation process was slow in this case. This might be due to

the low solubility of benzyl containing thioester in water. Finally, we observed a milky emulsion-like product which was hard to recover. As a result, MALDI analysis were not sufficient in this case. However, benzyl thiol “by-product” was detected when the mixture was extracted with DCM. This confirms that some polymers were certainly formed.



Scheme 52

Chapter IV: Prebiotic formation of thiol-rich polypeptides starting from aminothioli-containing amidonitriles

IV.1 Introduction

In the previous chapter, we have reported that cysteine nitrile was stable under strong acidic conditions, and that it started to polymerize in slightly acidic water (pH 5 – 6). After complete hydrolysis of thiazoline rings, the final polymer in this case would consist exclusively of cysteine residues. As we discussed before, the formation of this polymer, (Cys)_n cannot be regarded as probable in the primitive ocean, as it would have required a pool with very high cysteine nitrile concentrations. However, it would be more reasonable to imagine that thiol rich (but not the richest) peptides might have formed in the primitive ocean and, if we are true, might have engaged in the process of appearance of life. Our idea was to synthesize aminonitrile models that would polymerize and lead to the formation of polycysteine units separated by one (or more) amino acids, that is to say, poly(Cys-aa) or poly(Cys-aa₁-aa₂). If such polymerizations happen to work, then it will be reasonable to propose that thiazoline formation from longer Cys-terminated amidonitriles permitted the synthesis of long thiol-containing peptides, with a “reasonable” number of cysteine residues.

For this project, we needed to synthesize Cys-aa-CN (and Cys-aa₁-aa₂-CN) samples. We expected these nitriles to have a behaviour similar to that found with cysteine nitrile, so that we predicted that linear (and/or cyclic) peptides would form from them in slightly acidic or neutral water (*Figure 74*).

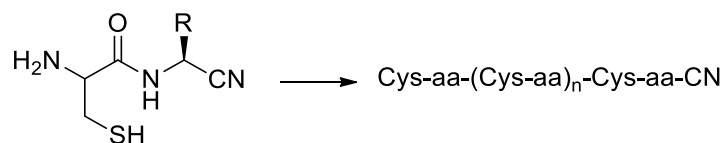


Figure 74: Aminothioli-containing amidonitriles are plausible precursors for thiol-rich peptides

We chose Cys-Gly-CN, Cys-Ala-CN, Cys-Val-CN and Cys-Met-CN as models to obtain peptidic chains. Furthermore, as we aimed to study the possibility to obtain polycysteine units that would be separated by two amino acids, we also decided to synthesize the simplest candidate, Cys-Gly-Gly-CN.

In this work, we will use amidonitriles produced from classical chemical synthesis. However, we believe that it is reasonable to propose that they may have been present in the primitive ocean, a prerequisite. As an example, a possible prebiotic synthesis of Cys-Gly-CN is presented in *figure 75*. First, cysteine could be activated as a thioester, which would condense with GlyCN to provide Cys-Gly-CN **171**.

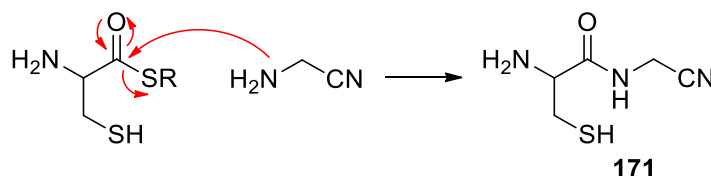
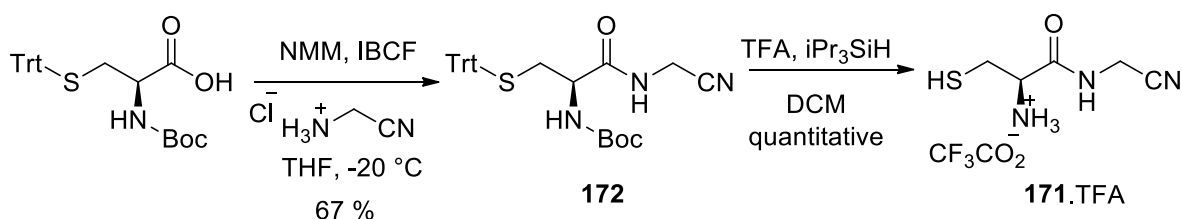


Figure 75: A possible prebiotic synthesis of Cys-Gly-CN

IV.2 Synthesis of aminothiols-containing amidonitriles

IV.2.1 Synthesis of Cys-Gly-CN

The easiest way to obtain Cys-Gly-CN **171** was to couple a protected cysteine residue directly with aminoacetonitrile. Thus, we treated Boc-Cys(Trt)-OH with isobutyl chloroformate (IBCF) in the presence of NMM as a base. Aminoacetonitrile was added after this activation process to give the expected nitrile **172**. Then, the removal of both Boc group on the amine and the Trt group on the thiol was performed in a single step (*Scheme 53*). The salt was recovered in water or D₂O for mass and NMR experiments.

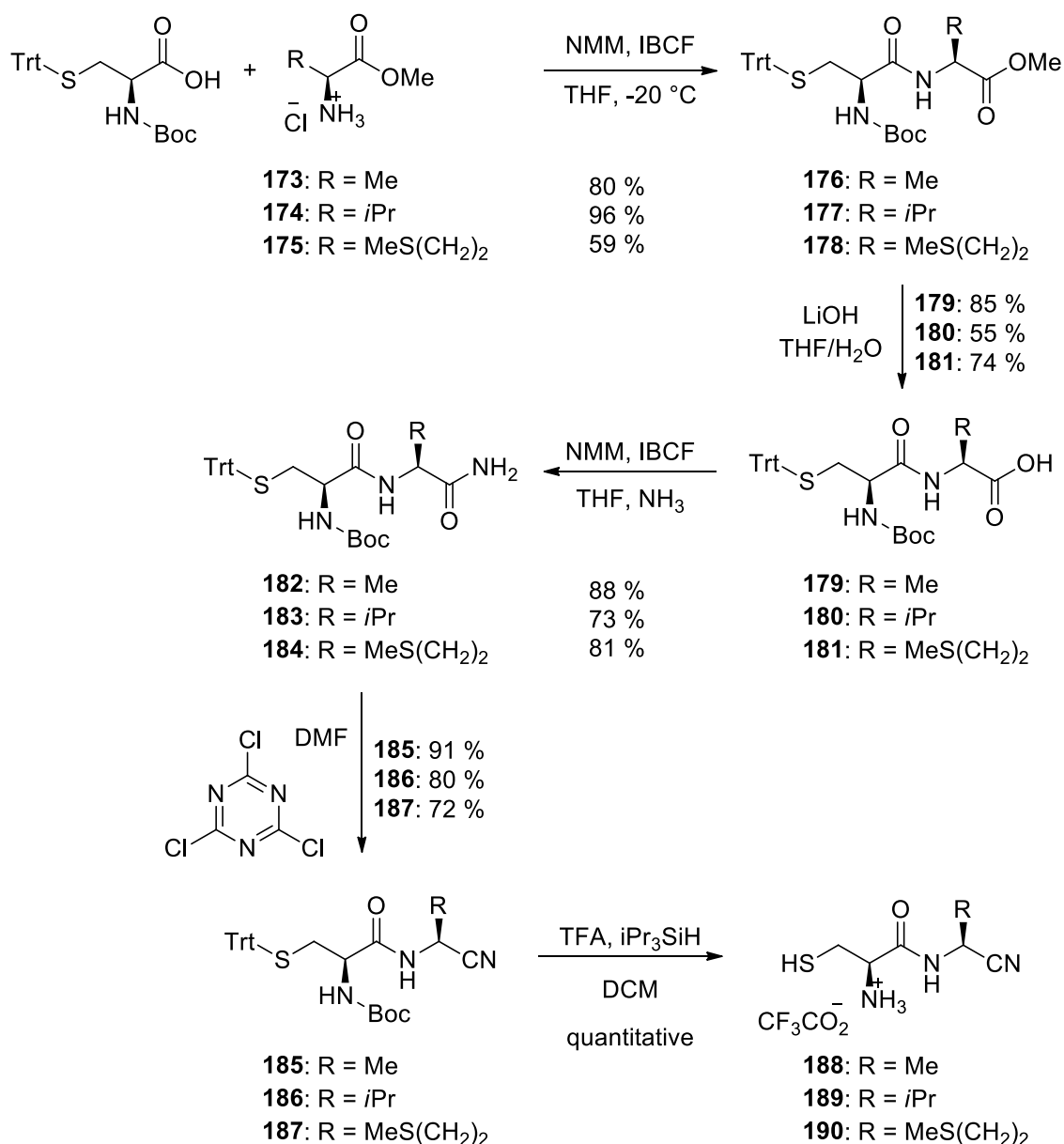


Scheme 53

IV.2.2 Synthesis of Cys-Ala-CN, Cys-Val-CN and Cys-Met-CN

In these cases, we treated Boc-Cys(Trt)-OH with IBCF in the presence of NMM as a base. This activation process was followed by the introduction of alanine, valine, or methionine methyl ester to give Boc-Cys(Trt)-Ala-OMe **176**, Boc-Cys(Trt)-Val-OMe **177** and Boc-

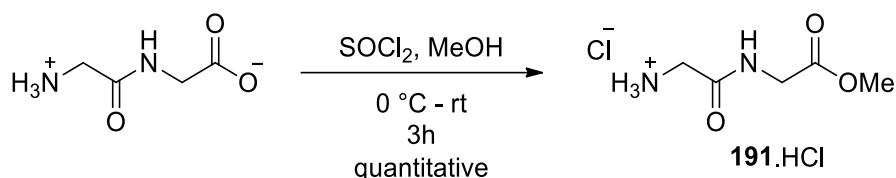
Cys(Trt)-Met-OMe **178**. These esters were transformed into the corresponding acids via saponification reactions. The obtained acids were activated again with IBCF and NMM, and then ammonia gas was introduced to afford the corresponding amides **182**, **183** and **184**. These amides were then subjected to a final dehydration step, using cyanuric chloride in DMF, and gave the corresponding nitriles **185**, **186** and **187**. Deprotection of Boc and Trt groups yielded the targeted nitriles **188**, **189** and **190** (*Scheme 54*). Once again, the final salts were recovered in water or D₂O.



Scheme 54

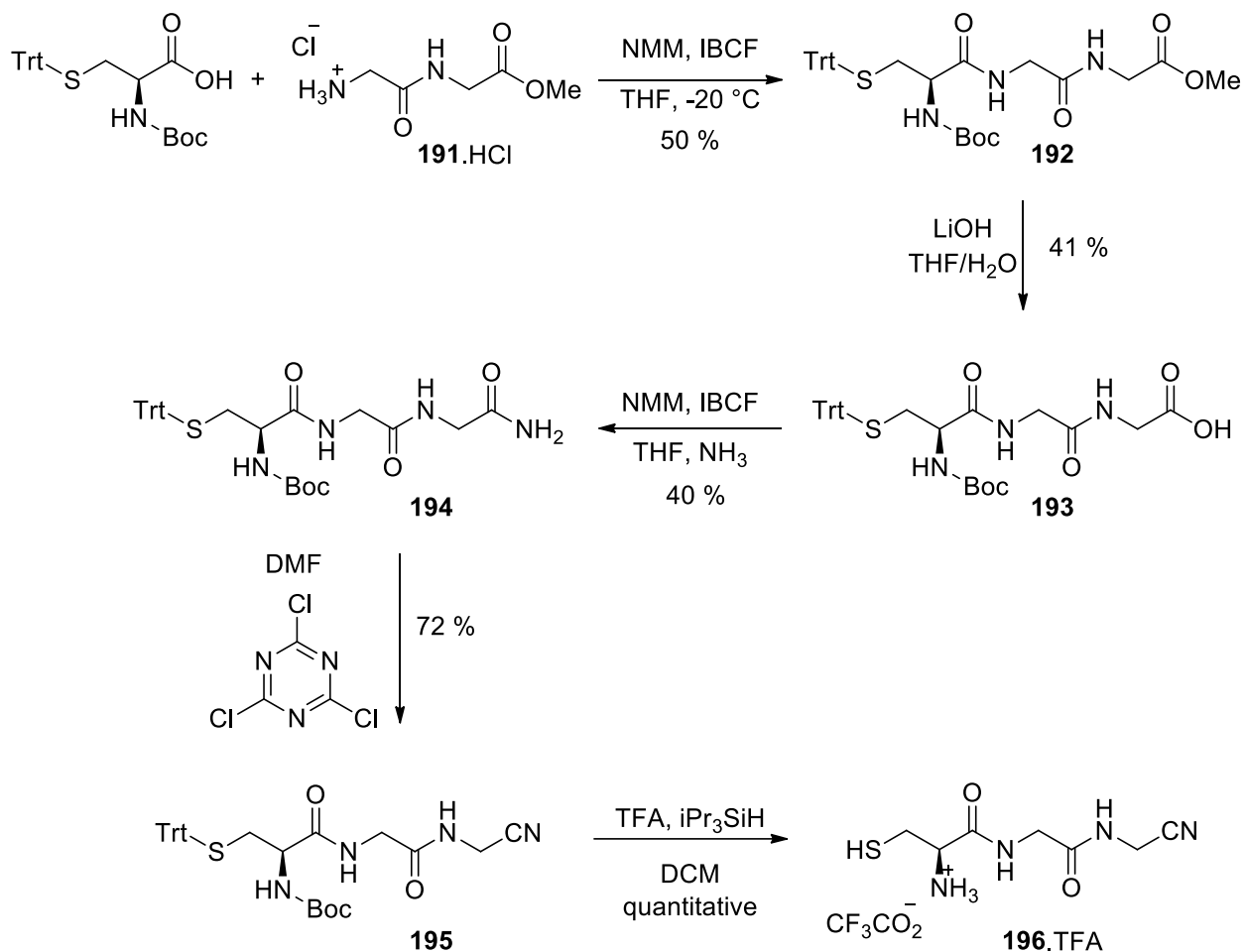
IV.2.3 Synthesis of Cys-Gly-Gly-CN

We first obtained Gly-Gly-OMe **191** by treating diglycine with thionyl chloride in methanol (*Scheme 55*).¹⁷² This ester was used to introduce the Gly-Gly moiety onto cysteine.



Scheme 55

In this synthesis, the treatment of Boc-Cys(Trt)-OH with IBCF and NMM was followed by the introduction of Gly-Gly-OMe **191** to form Boc-Cys(Trt)-Gly-Gly-OMe **192**. The newly formed ester **192** was transformed to the corresponding acid **193** via saponification reaction. The acid **193** was then activated using IBCF and NMM, and ammonia gas was introduced to afford the amide **194**. The formed amide was treated with cyanuric chloride to afford the corresponding nitrile **195**. Deprotection of Boc and Trt groups was achieved in a single step to yield the required nitrile **196** (*Scheme 56*).

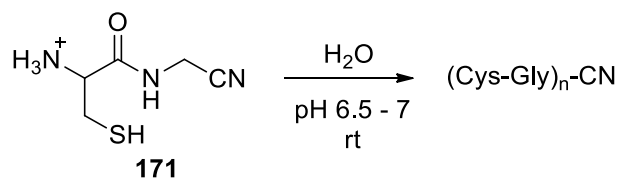


Scheme 56

IV.3 Prebiotic Formation of thiol-rich polypeptides

IV.3.1 Formation of poly(Cys-Gly)

Cys-Gly-CN **171** was found to be stable for a long period in acidic water (pH 1). When the pH was nearly neutralized (pH 6.5) by adding sodium bicarbonate, the solution started to become turbid indicating the formation of polymers (*Scheme 57*)



Scheme 57

Few hours later, a white solid started to precipitate at the bottom of the flask. After one night, the white thin layer that precipitated was recovered and dried. *Figure 76* shows the observed solution during the polymerization process.

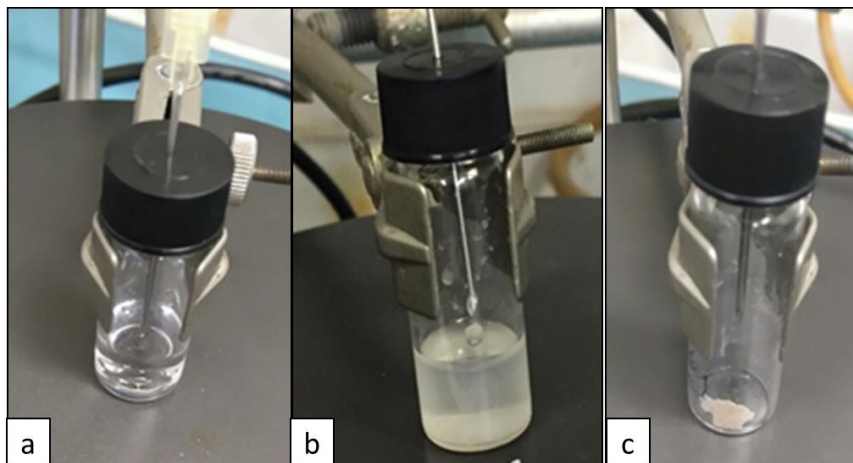


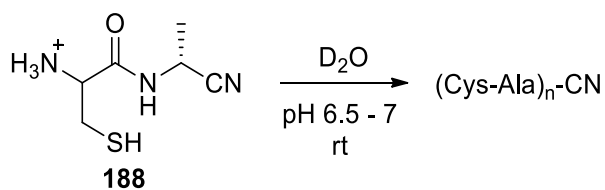
Figure 76: Polymerization process of Cys-Gly-CN under argon. a) Cys-Gly-CN in degassed water, b) turbid solution after few hours, c) the recovered solid, (Cys-Gly)_n, stored under argon

This solid was only poorly soluble in DMSO (but completely insoluble in DMF). However, a mixture of this product in DMSO was subjected for MALDI-TOF analysis.

Unfortunately, the expected masses were not observed in MALDI analysis. This might be attributed to the poor solubility of the solid in DMSO. This is one reason for which we decided to introduce amino acids other than glycine. These amino acids contain other “more organic” substituents that we expected to increase the solubility of our polymers in organic solvents and would lead to better results.

IV.3.2 Formation of poly(Cys-Ala)

After deprotection of the nitrile **185**, Cys-Ala-CN **188** was extracted as TFA salt in water. We found that it was stable in acidic water for hours. When the pH was neutralized (6.5 – 7) using sodium bicarbonate, the clear water solution started to become turbid and within few hours, a very thin layer of a white solid started to precipitate indicating the formation of poly(Cys-Ala) (*Scheme 58*).



Scheme 58

We first run this reaction in D_2O that was neutralized and left unstirred. After few hours the NMR tube showed a precipitate out of the solution indicating the polymerization of Cys-Ala-CN units. The formation of thiazoline rings was characterized via $^1\text{H-NMR}$ spectroscopy by the appearance of broad signals around 5 ppm (CH of thiazolines), and in the region from 3.3 to 3.6 (CH_2 of thiazolines) (Figure 77).

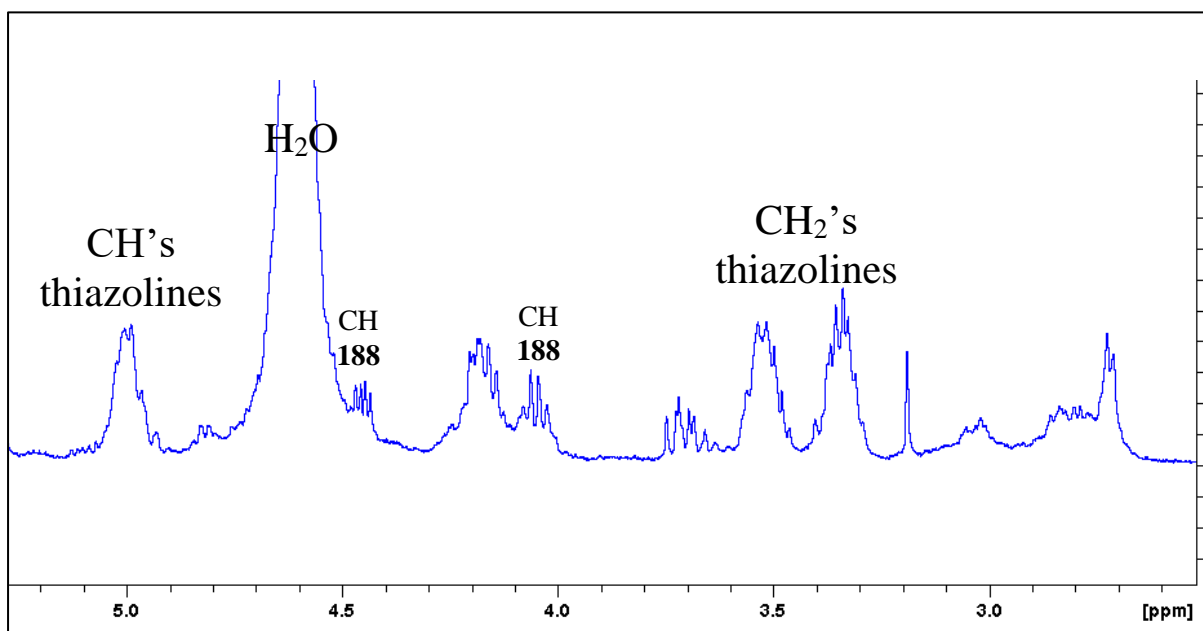


Figure 77: Observation of thiazoline rings of poly(Cys-Ala) in $^1\text{H-NMR}$

The solution was left unstirred for overnight, then the solid was recovered and subjected to MALDI-TOF analysis. In order to prepare the matrix for MALDI analysis, 10 % of acetonitrile was added to DHB solution in DMSO. This solution was mixed with our polymer (dissolved in DMSO) and this mixture was dropped on the plate and introduced to MALDI. The obtained spectrum was rather complex. It did not show the expected peptides as a series of single masses (a mass for each $(\text{Cys-Ala})_n$), but for each given peptide length, a set of

masses was observed (Figure 78). These masses were attributed to the presence of a varied number of hydrolysed and non-hydrolysed thiazoline rings in each chain.

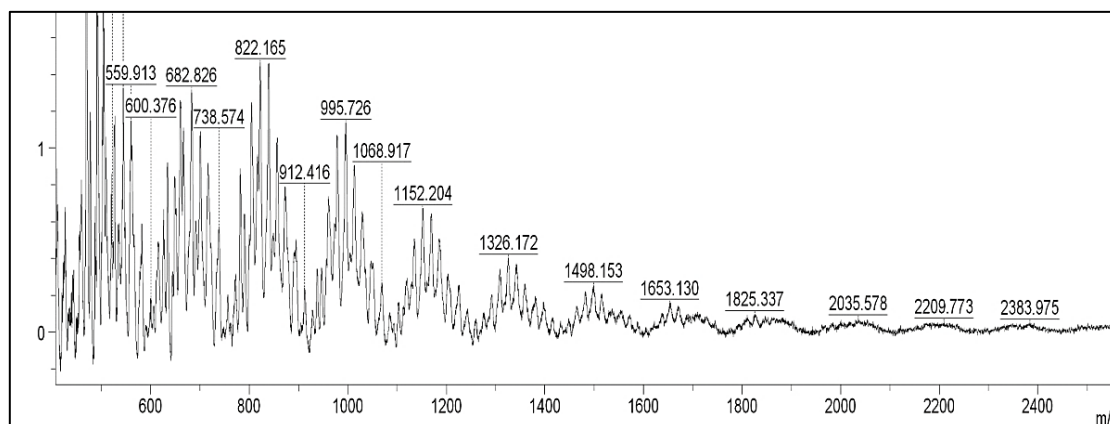


Figure 78: MALDI MS-spectrometry of poly(Cys-Ala)

For example, the masses around 1152 correspond to seven “monomers” (= Cys-Ala) ending with a nitrile (observed as $[M+Na]^+$). Focusing on this region, seven single masses were observed. The highest number of possible thiazoline rings in the case of (Cys-Met)₇-CN is six rings, which corresponded to a calculated mass 1134.4598 (observed 1134.965). With less rings, higher masses are expected, the difference between two peaks being 18.

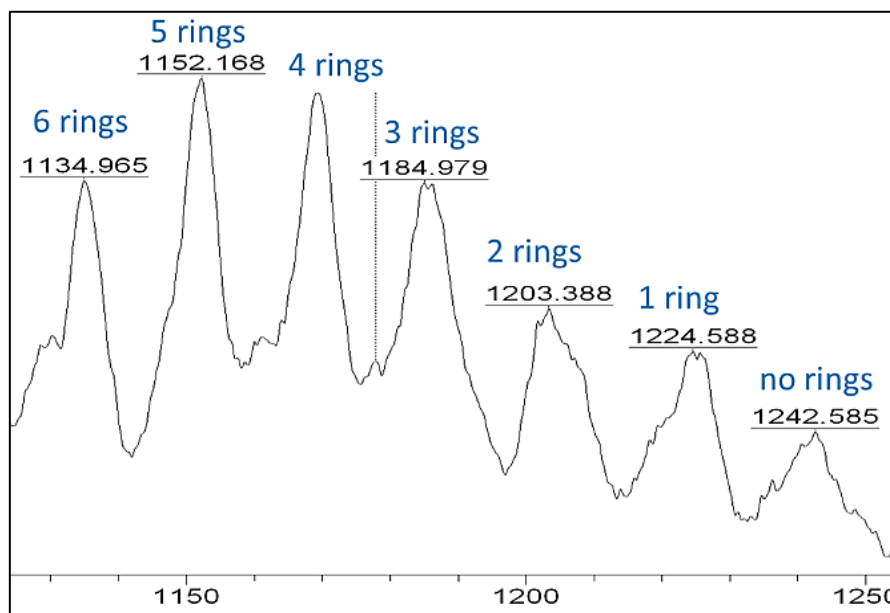


Figure 79: Peaks corresponding to (Cys-Met)₇-CN on MALDI spectrometry

For instance, taking the peak of 1184.979, this mass corresponds to a (Cys-Met)₇-CN in which three thiazoline rings were hydrolysed (and thus three rings still present). An example of such structure is presented in (Figure 79). Of course, we cannot determine the precise place of thiazoline rings in the chain. Other structures are possible, so that the observed mass must correspond to a mixture of oligomers. What is sure is that all these structures are heptamers of CysGly, *i.e.* chains containing 14 amino acids **197** (Figure 80).

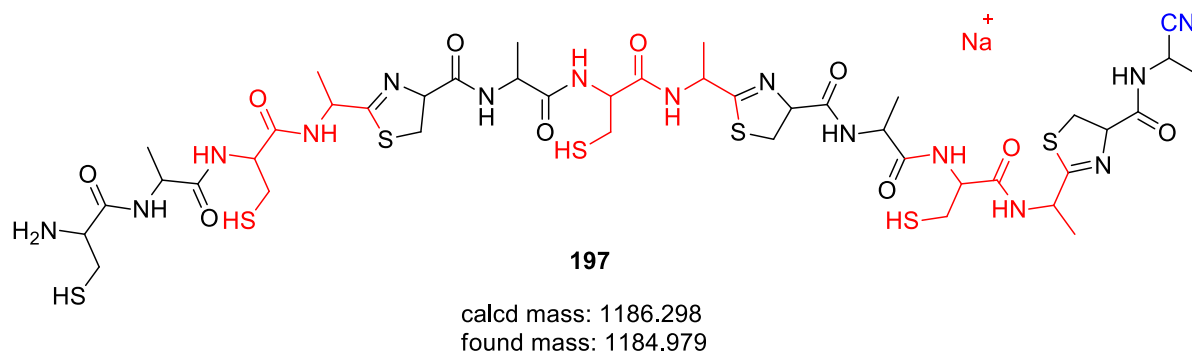


Figure 80: A possible structure of $M = 1184.979$

Some examples of the found linear polypeptides that contain thiazoline rings are shown in table 2.

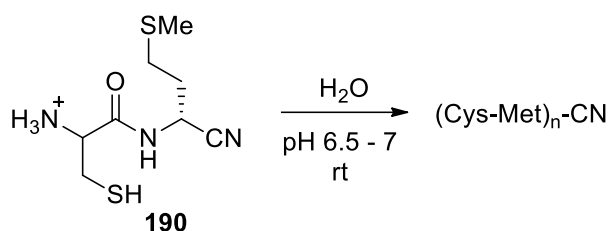
No. of monomers	No. of thiazolines	Calculated molecular wt. (M+Na) ⁺	Observed average mass (M+Na) ⁺
(Cys-Ala) ₅ -Cys-Ala-CN	5	977.2534	978.267
(Cys-Ala) ₅ -Cys-Ala-CN	4	995.2687	995.765
(Cys-Ala) ₅ -Cys-Ala-CN	3	1013.284	1013.111
(Cys-Ala) ₆ -Cys-Ala-CN	6	1133.4589	1134.965
(Cys-Ala) ₆ -Cys-Ala-CN	5	1151.4742	1152.168

Table 2

IV.3.3 Formation of poly(Cys-Met)

Cys-Met-CN **190** solution started to become turbid at neutral pH, indicating the formation of insoluble polymers (Scheme 59). After leaving the solution at room temperature without

stirring for overnight, a thin layer of white solid was observed at the bottom of the container. This solid was recovered and subjected to MALDI spectrometry (in DMSO solution)



Scheme 59

Several polypeptide chains of various length (up to 25 chains, *i.e.* 50 amino acids) were observed. Just like in the case of $(\text{Cys-Ala})_n$, we observed a set of masses for each peptide length. Here also, various numbers of thiazoline rings remained in the formed chain (*Figure 81*).

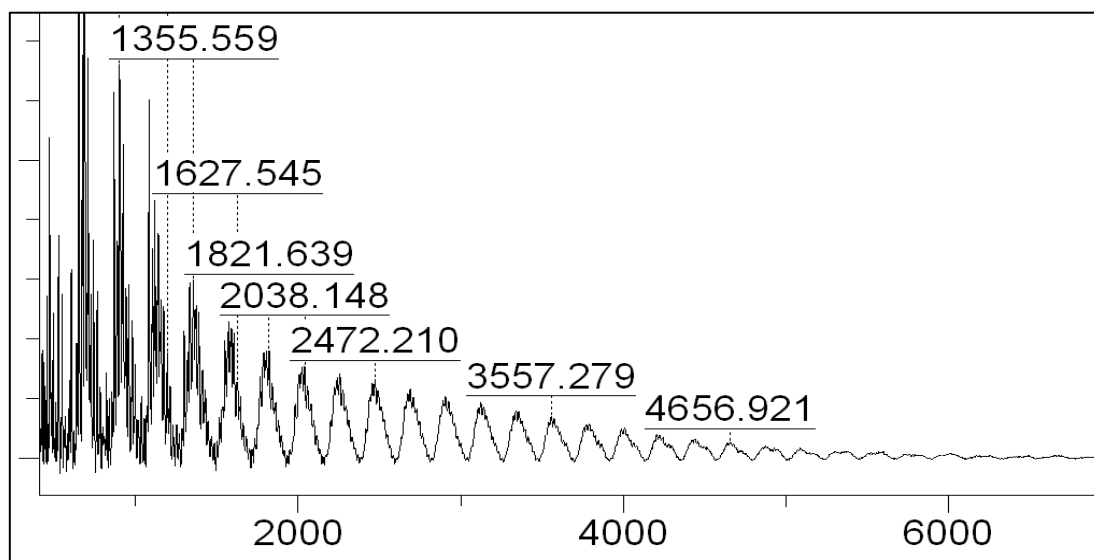


Figure 81: MALDI MS-Spectrometry for poly(Cys-Met)

Taking $(\text{Cys-Met})_{14}\text{-CN}$ as an example, several masses were detected (*Figure 81*). The observed masses started with 3068.278; this value indicates the presence of 13 thiazoline rings (the maximum number of rings possible for 14 “monomeric” Cys-Met). This mass is followed by 3091.303 (12 thiazoline rings) and so on.

It should be noticed that, contrarily to the previous example for which $[\text{M}+\text{Na}]^+$ masses were observed, in this case $[\text{M}+\text{H}]^+$ masses were detected.

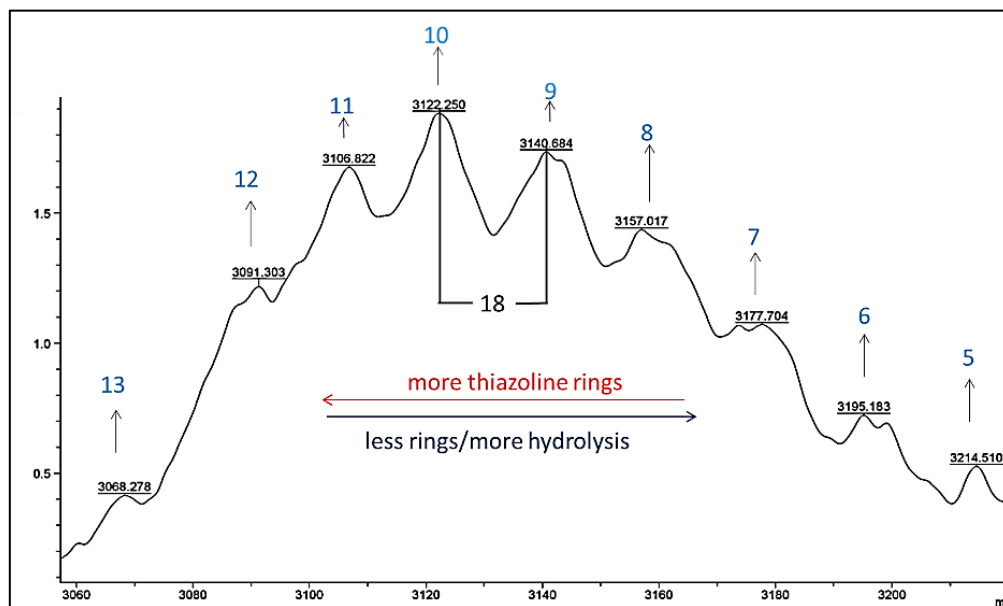


Figure 82: Peaks correspond to $(\text{Cys-Met})_{14}\text{-CN}$ on MALDI

As shown in *figure 82*, the difference between every two successive masses is clearly 18, indicating a gain of one water molecule (thiazoline hydrolysis). A structure **198** consisting of six “monomers” (= Cys-Met) containing three thiazoline rings and ending with a nitrile is presented in *figure 83* as an example.

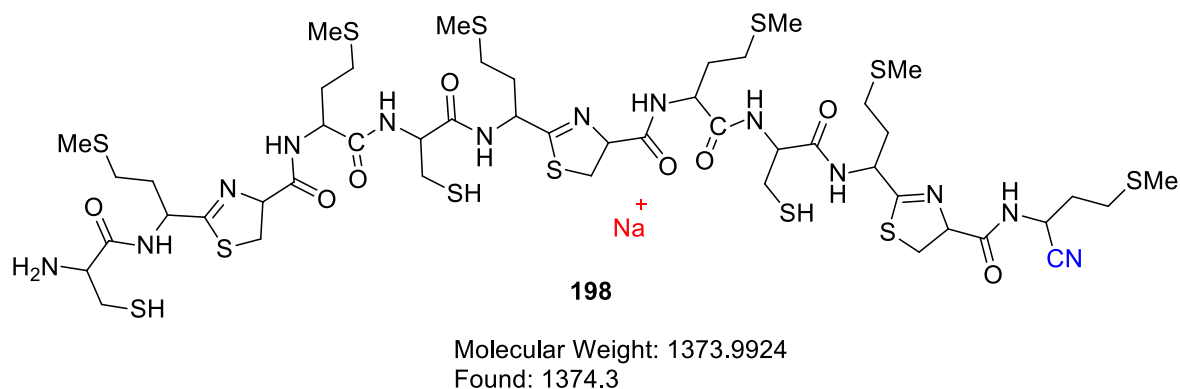


Figure 83: A possible structure for $M = 1374.3$

Indeed, the observed masses belonged to peptides that are mainly ended with a nitrile, However, chains that ends with COOH were observed sometimes. Other examples on some observed masses on MALDI are given in *table 3*.

No. of monomers	No. of thiazolines	Calculated molecular wt. (M+Na) ⁺	Observed average mass (M+Na) ⁺
(Cys-Met) ₁₃ -Cys-Met-CN	13	3068.5522	3068.278
(Cys-Met) ₁₃ -Cys-Met-CN	10	3122.5981	3122.250
(Cys-Met) ₁₃ -Cys-Met-CN	9	3140.6134	3140.684
(Cys-Met) ₁₄ -Cys-Met-CN	9	3374.9524	3375.55
(Cys-Met) ₁₅ -Cys-Met-CN	11	3573.2608	3573.852
(Cys-Met) ₁₅ -Cys-Met-CN	8	3627.3067	3626.564

Table 3

However, some odd single peaks were also observed on the MALDI. As an example, the peak at 1081.959 shown in *figure 84* does not seem to fit with the observed series of masses. We were not able to relate this mass a peptidic chain consisting of thiazoline rings and amides. Thus, another structure should be proposed in this case.

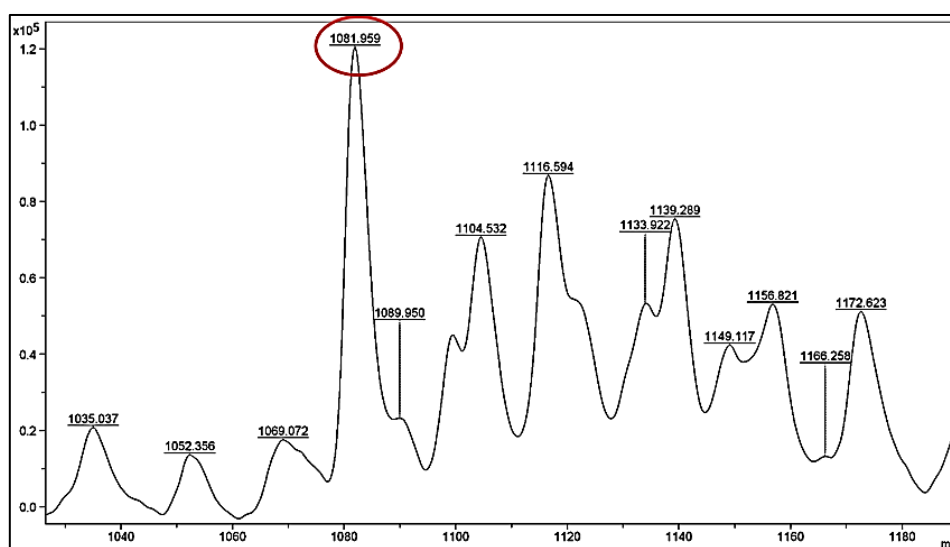


Figure 84: Unusual peak corresponds to the cyclic structure 203

Indeed, at any degree of polymerization, the obtained polymer contains both a nitrile function and an aminothiols part. Such a molecule has two pathways to take. For example, when (Cys-Met)₅-CN **199** is present in a solution, it can undertake further reaction to give a longer linear polypeptides. But, it may also cyclize if the thiol function at one end of the chain reacts intramolecularly with the nitrile at the opposite end of the chain. Such a process eventually yields a macrocycle cycle containing five thiazoline rings **200**, that agrees with the observed mass (*Figure 85*).

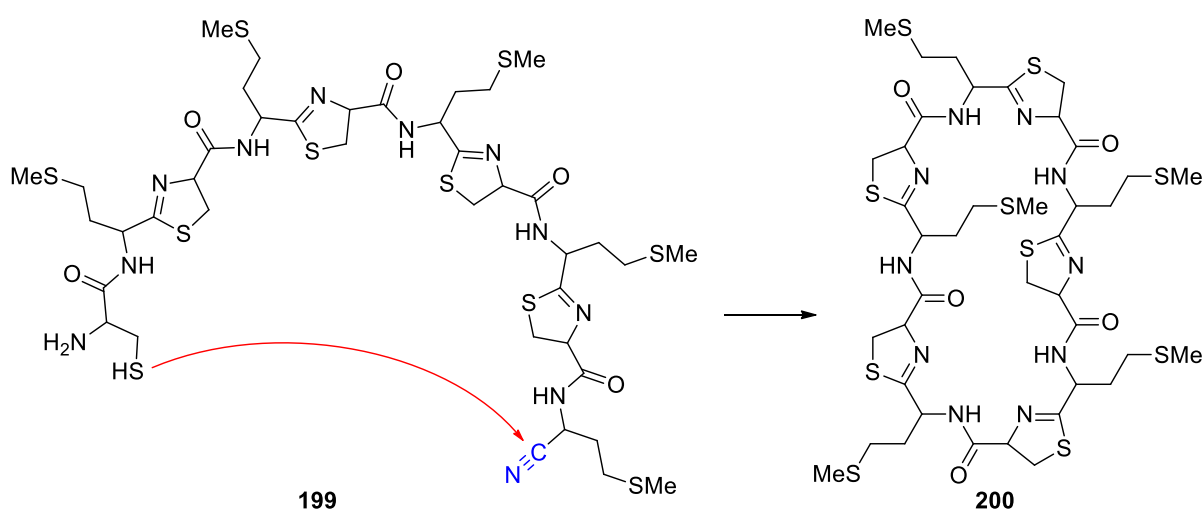


Figure 85: Formation of the cyclic structure 200

A list of macrocycles that we observed is presented in *table 4*.

No. of monomers	No. of thiazolines	Chemical formula	Calculated molecular wt. (M+H) ⁺	Observed average mass (M+H) ⁺
Cyclo(Cys-Met)₃	3	C ₂₄ H ₃₇ N ₆ O ₃ S ₆ ⁺	649.978	649.074
Cyclo(Cys-Met)₄	4	C ₃₂ H ₄₉ N ₈ O ₄ S ₈ ⁺	866.3027	865.745
Cyclo(Cys-Met)₅	5	C ₄₀ H ₆₁ N ₁₀ O ₅ S ₁₀ ⁺	1082.6264	1081.959
Cyclo(Cys-Met)₆	6	C ₄₈ H ₇₃ N ₁₂ O ₆ S ₁₂ ⁺	1298.9501	1298.332
Cyclo(Cys-Met)₇	7	C ₅₆ H ₈₅ N ₁₄ O ₇ S ₁₄ ⁺	1515.2738	1515.213
Cyclo(Cys-Met)₈	8	C ₆₄ H ₉₇ N ₁₆ O ₈ S ₁₆ ⁺	1731.5975	1731.996

Table 4

Masses presented in *table 4* are average masses. To confirm the proposed structures, we recorded their high resolution mass spectra (HRMS-SEI). As shown in *figure 86*, the exact masses agree with the proposed thiazoline-containing cyclopeptides. However, the quantities of cyclo(CysMet)₇ and cyclo(CysMet)₈ in the mixture of polymers were not sufficient to allow the measurement of their exact mass.

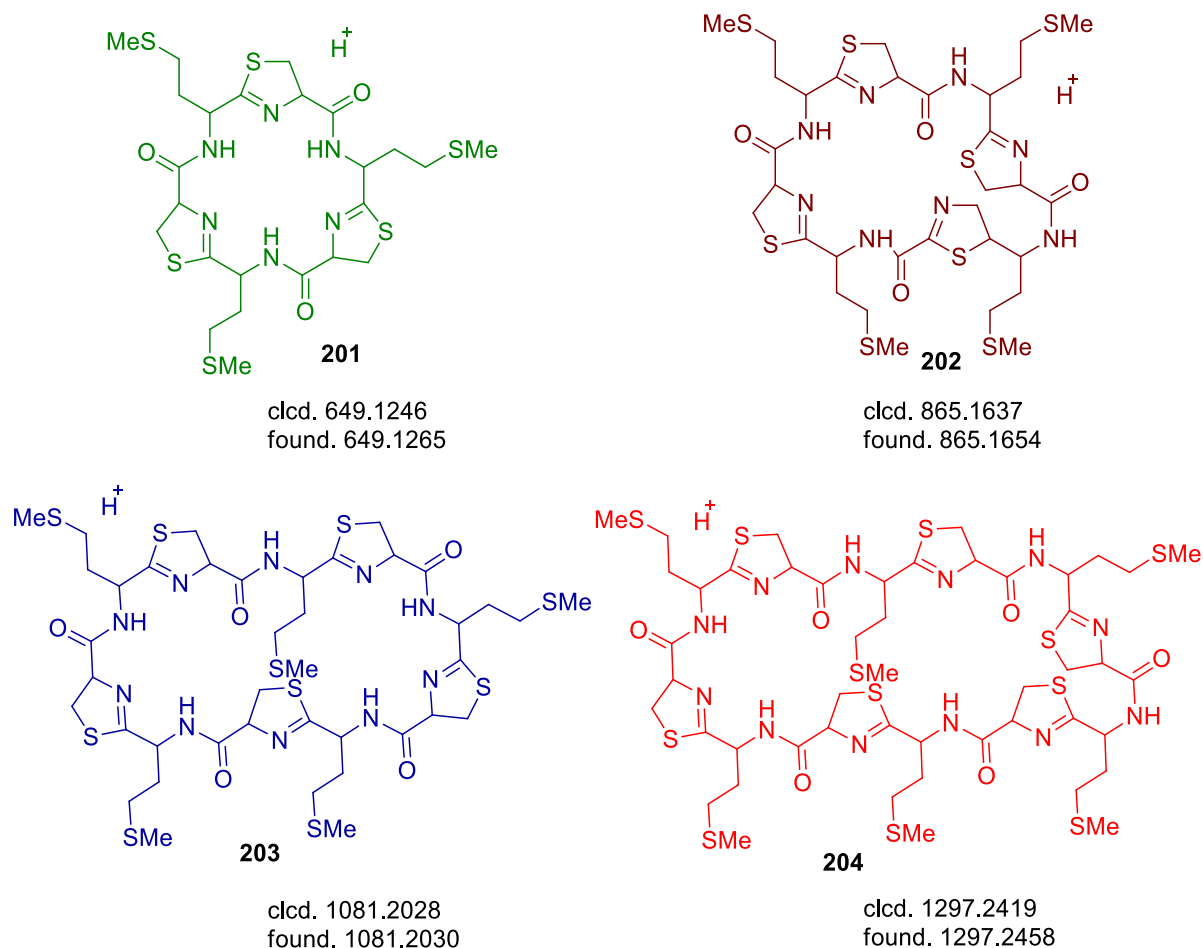
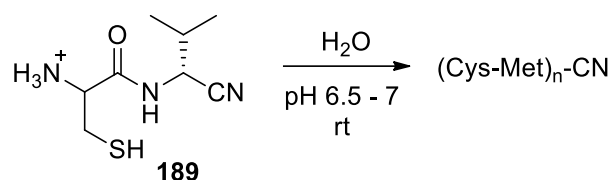


Figure 86: Some detected cyclic structures of poly(Cys-Met)

IV.3.4 Formation of poly(Cys-Val)

Cys-Val-CN **189** was found to be stable under acidic conditions and its solution started to become turbid at neutral pH, indicating the formation of insoluble polymers (*Scheme 60*).



Scheme 60

We first tried the reaction in D₂O and followed it by NMR spectroscopy. After one hour, the spectra showed the formation of the early stages of the first oligomers (dimers, trimers,...) (Figure 87). After few hours, the polymerization process continued and eventually led to a polymer that was not soluble in D₂O and precipitated at the bottom of the NMR tube.

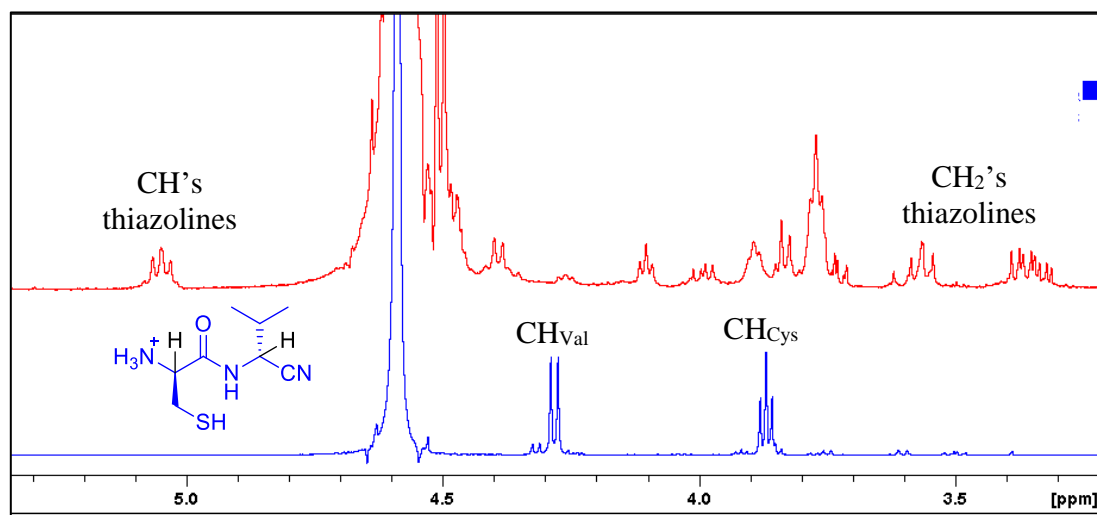


Figure 87: ¹H-NMR showing the beginning of Cys-Val-CN polymerization

We repeated this reaction in H₂O and left the solution at room temperature without stirring. The obtained solid was subjected to MALDI spectrometry.

Several polypeptide chains of various length (up to 30 chains, *i.e.* 60 amino acids) were observed. Sets of peaks were also observed in the mass spectrometry analysis indicating the presence of various numbers thiazoline rings (Figure 88).

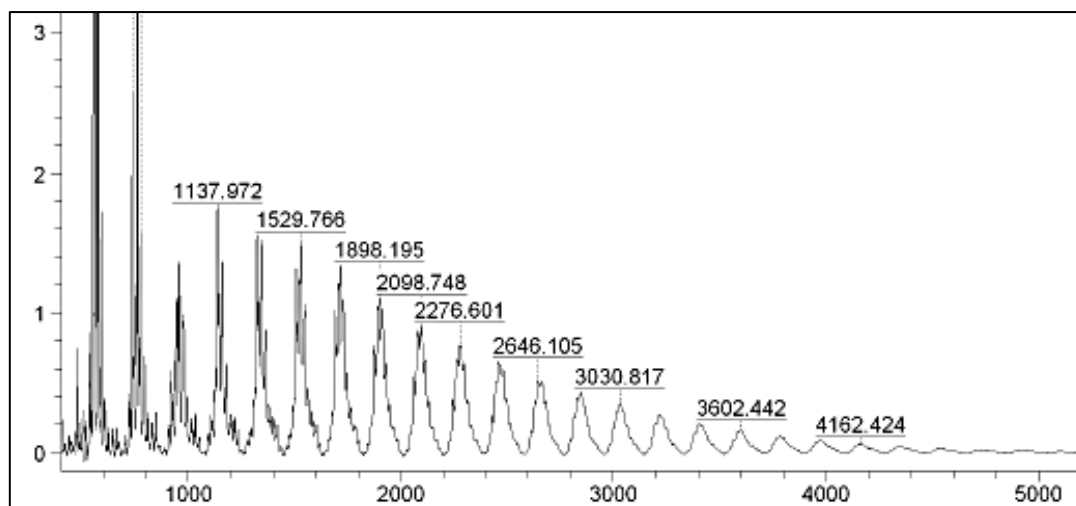


Figure 88: MALDI MS-Spectrometry of poly(Cys-Val)

We noticed that the farther we go in MALDI spectra, the less precise values we observed. That is to say, in higher masses, the values had some little deviations from the expected ones. This observation can be attributed to the formation of disulfide bridges (inter- or intramolecular). If so, it is logic to have more disulfides in the case of higher masses.

Examples of some observed values are given in *table 5*.

No. of monomers	No. of thiazolines	Calculated molecular wt. (M+Na) ⁺	Observed average mass (M+Na) ⁺
(Cys-Val) ₁₃ -Cys-Val-CN	8	2708.907	2709.719
(Cys-Val) ₁₅ -Cys-Val-CN	9	3096.251	3095.221
(Cys-Val) ₁₇ -Cys-Val-CN	12	3446.754	3447.252
(Cys-Val) ₂₁ -Cys-Val-CN	17	4165.773	4162.424
(Cys-Val) ₂₄ -Cys-Val-CN	19	4736.564	4734.501

Table 5

Similarly to what we observed in the case of poly(Cys-Met), cyclic structures were observed. These cycles containing thiazolines were identified precisely by means of high resolution mass spectrometry (HRMS-SEI) (*Figure 89*).

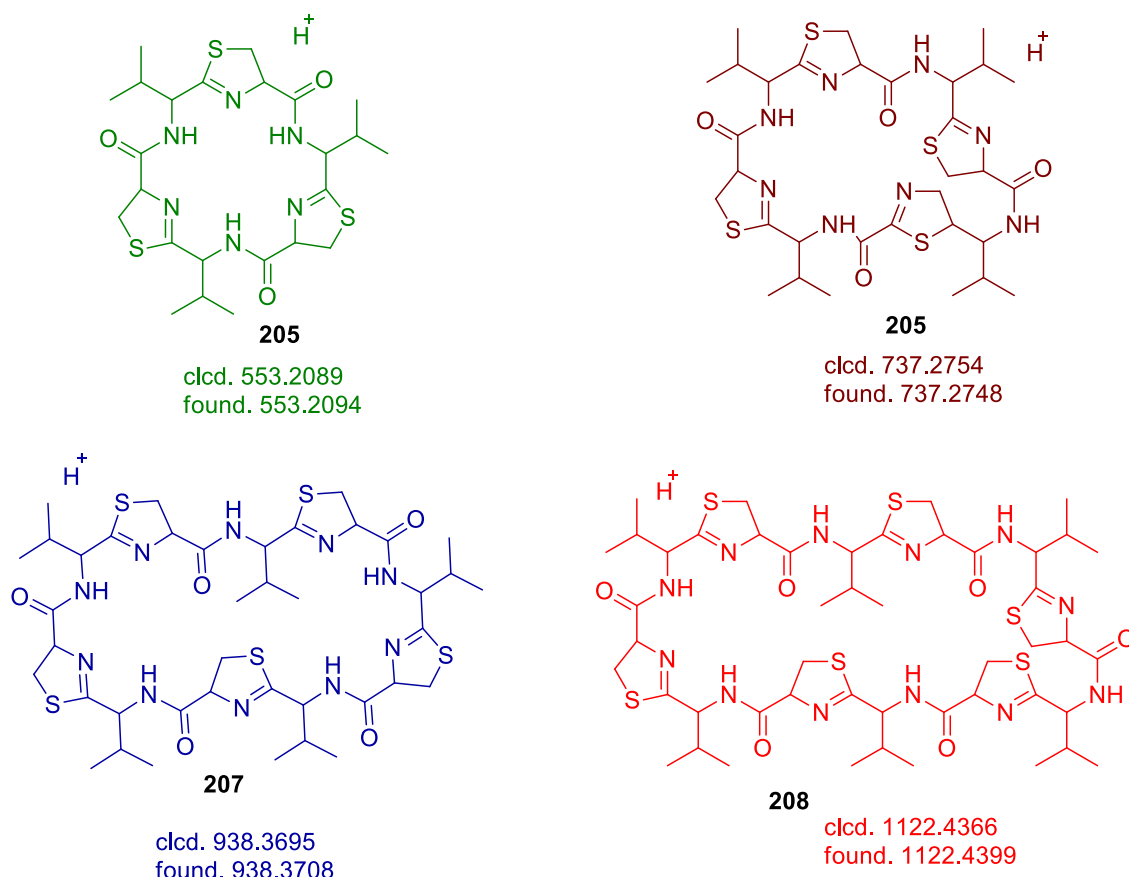


Figure 89: Some detected cyclic structures of poly(Cys-Val)

In this example, we decided to try MS/MS experiments, that were performed by Maxime Bridoux (CEA – Centre DAM – Île de France), in order to detect the expected fragmentations of these cyclic structures. For example, when the peak of the cycle **205** (553) was separated and submitted to further energy (collision energy 15), the expected fragments were observed (*Figure 90*). Once again, confirming the proposed structures. Similar spectra were obtained with other macrocycles.

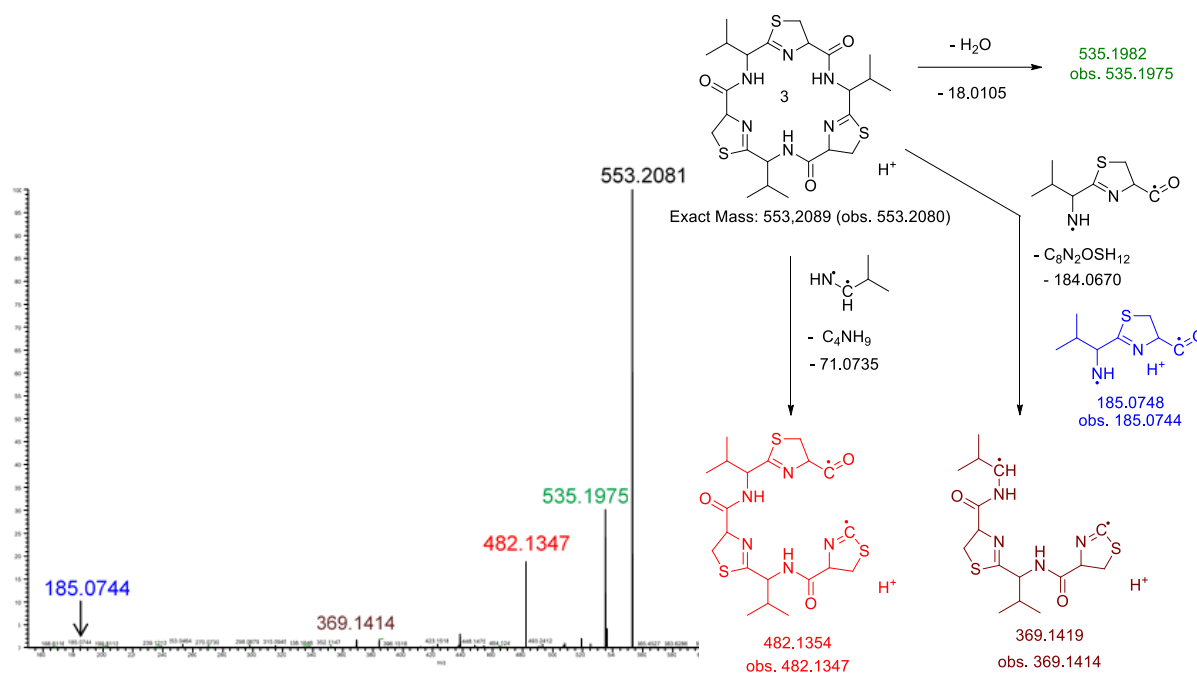
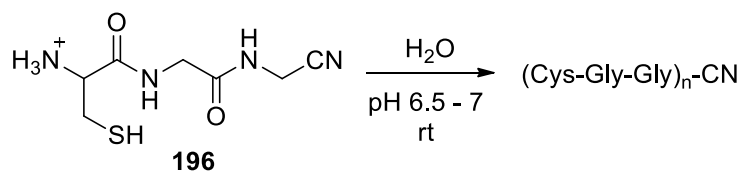


Figure 90: MS/MS of 553 (m/z), observation of the expected fragments

IV.3.5 Formation of poly(Cys-Gly-Gly)

Similarly to Cys-Gly-CN **171**, Cys-Gly-Gly-CN **196** was found to be stable in very acidic water (pH 1). After few hours at pH 6.5, poly(Cys-Gly-Gly) chains started to precipitate (*Scheme 61*). The recovered solid was dried and finally subjected to MALDI-TOF analysis



Scheme 61

However, the mass resulted from MALDI was not satisfactory and did not show the expected masses. This might be due to the very poor solubility of the obtained polymers in DMSO.

IV.4 First attempts to observe a catalytic behaviour of the obtained polymers

To study the catalytic effect of our polymers in the formation of peptide bonds, poly(Cys-Val) was used as an example. A small amount of poly(Cys-Val) was put in D₂O. DCl was added to the obtained suspension in order to promote the hydrolysis of thiazoline rings in peptide chains and the mixture was heated at 45 °C for 24 hours, before it was divided in six NMR tubes. In each tube, we added a “monomer”. We used Glycine, glycine amide, aminoacetonitrile, serine, cysteine and homocysteine. The six tubes were left at room temperature and the evolution of each mixture was followed by NMR spectroscopy.

Expected products were oligopeptides (at least dipeptides), but we did not observe any evolution product even after two months.

In these first experiments, the complete lack of catalytic effect (especially with aminoacetonitrile **39**, that we know that it forms GlyGlyCN **42** in the presence of thiols) might be due to the low solubility of long thiol-containing peptides. Another reason could be that thiazoline derived from valine were too stable, so that we had only a small amount of “free” thiol functions. Such high stability was observed in the case of the thiazoline of Val-Cys-NH₂ **92**, in which the hydrolysis step was very slow. Thus, poly(Cys-Val) was probably not the best polymer to test. Such experiments should be repeated with other polymers.

Chapter IV

Conclusion

Our research focuses on the prebiotic synthesis of peptides, especially thiol-rich ones. This research supports the metabolism first hypothesis, an alternative proposal to the RNA world hypothesis. It is related to de Duve's thioester world, and to Sutherland's cyanosulfidic protometabolism. We proposed that this world, that we proposed to name "Thiol Rich Peptide (TRP) world", had appeared in primitive ocean thanks to the presence of significant amounts of HCN that permitted the synthesis of aminonitriles from aldehydes. The chemistry of this world is based on thiols and nitriles. Thiol-containing peptides would have been the most important molecules in this world.

In the first chapter, we study the reaction of aminoacetonitrile (Gly-CN) with cysteine. This reaction leads to the formation of a thiazoline cycle, which is then hydrolysed to yield the dipeptide Gly-Cys. Similarly, homocysteine gives Gly-Hcy via a 6-membered cycle. These reactions were found to work under a wide range of pH's and temperatures. However, we chose to run most of our reactions in slightly acidic water (pH 6 at 45 °C). In this way, we were able to obtain several dipeptides (aa-Cys or aa-Hcy) by introducing different aminonitriles (the first products of the Strecker reaction) (*Figure 91*).

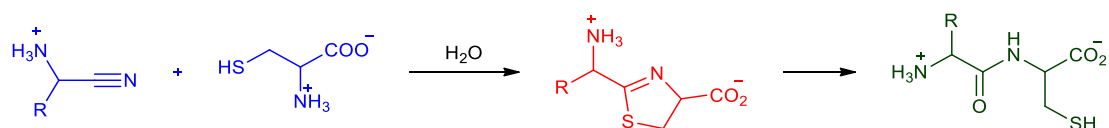


Figure 91

We found that these thiol-containing dipeptides were able to form active thioester bonds that promote the formation of longer peptide chains. For instance, any dipeptide nitrile aa₁-aa₂-CN (the simplest example being Gly-GlyCN), produced from the reaction of H₂N-aa₂-CN with a thioester of aa₁, would react with cysteine to give the tripeptide aa₁-aa₂-Cys, and with homocysteine to give aa₁-aa₂-Hcy. We have demonstrated that dipeptides like GlyCys are able to promote the formation of heterocycles that play an important role in today's biochemistry. This remarkable catalytic behaviour might have had a crucial role to promote chemical evolution processes.

In the second chapter, we study the behaviour of homocysteine nitrile (Hcy-CN) in water. We found that it cyclized to form homocysteine thiolactone (Hcy-thiolactone). This argues for the

Conclusion

existence of this thiolactone in the primitive ocean. Hcy-thiolactone showed a double reactivity, thiolactone opening by amines followed by aminothiol condensation reaction with nitriles. For example, when Hcy-thiolactone reacted with aminoacetonitrile, Hcy-Gly-CN was formed. The obtained thiol-containing dipeptide reacted again with aminoacetonitrile to form Gly-Hcy-Gly-CN (*Figure 92*).

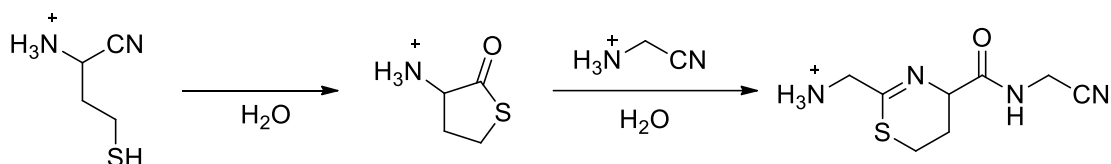


Figure 92

In the third chapter, we discuss the plausible prebiotic formations of polycysteine. We found that cysteine nitrile (Cys-CN) in water lead to the formation of a cysteine dimer, a trimer, and finally to the formation of a non-soluble polymer, polycysteine. This was also the case when Cys-CN was replaced by cysteine ethyl ester, another plausible prebiotic precursor of poly(Cys) (*Figure 93*).

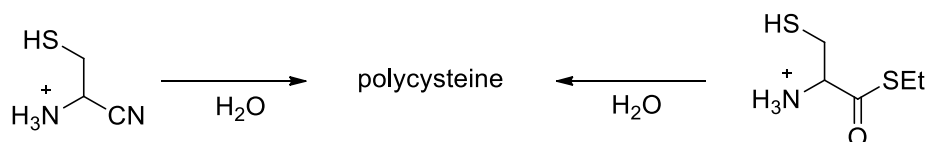


Figure 93

Finally, we discuss the behaviour of Cys-aa-CN model molecules under our prebiotic conditions in the last chapter. In this case, the corresponding polypeptides, containing up to 50 amino acid units, have been formed (*Figure 94*).

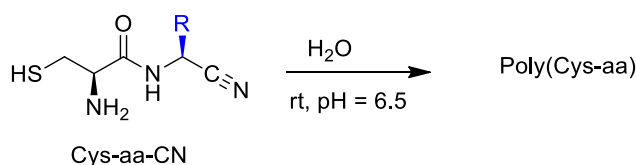


Figure 94

Conclusion

In addition to the linear polypeptides, macrocyclic polythiazolines have also been detected, that would lead to macrocyclic polypeptides after hydrolysis of thiazoline rings. We believe that such thiol-rich cycles, especially if formed from various Cys-aa-CN (an example is presented in *figure 95*), could have a catalytic role in the prebiotic ocean.

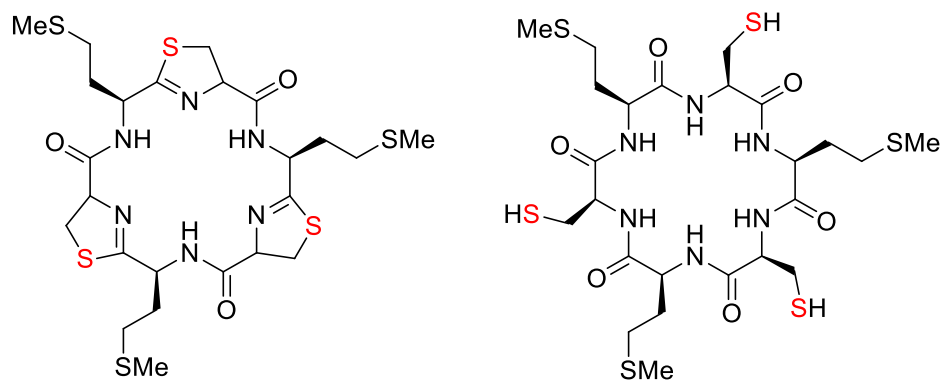


Figure 95

Perspectives

In this study, we have used the corresponding nitriles of glycine, methionine and valine (beside other aminonitriles) to react with aminothiols. However, using other aminonitriles (which may correspond to proteinogenic or non-proteinogenic amino acids) will lead to the formation of a wide range of thiol-containing dipeptides.

Similarly, introducing new amidonitriles (Cys-aa-CN) in water will give different polypeptides. In this part of our study, we have used one type of amidonitriles for each attempt. These attempts have led to the formation of polypeptides that were always separated by one type of repeating units. However, mixing two or more different amidonitriles together in water will probably lead to the formation of polypeptide sequences containing different amino acid residues with random (or slightly selective) arrangement. These random assemblies of amino acid residues would be closer to the real scenarios that gave rise to life.

Peptides with longer repeating units were obtained when CysGlyGly-CN monomers were used to obtain cysteine-rich peptides in which Cys residues were separated by two Gly residues. Further steps of our research will focus on the synthesis of various Cys-aa₁-aa₂-CN or Cys-aa₁-aa₂-...-aa_n-CN in order to obtain more complex polymers.

Even though we did not notice an obvious effect of zinc ions in our first attempts (p. 50 and p. 79), the effect of metal cations that were present in the primitive ocean cannot be totally ignored. For instance, it has been found that sodium and potassium ions have different effects on cyclic and linear peptide hydrolysis.¹⁷³ In our polymer formation experiments, replacing sodium with another metal may lead to different structures or, at least, may enhance the probability to form a given structure over the others. Beside the presence of metal cations, concentration factor of our prebiotic models in water should be taken into consideration. For instance, concentration would affect the reaction between aminothiols and aminonitriles as well as possible S-S bond forming reactions, either intra- or intermolecular. Such factor would affect, for instance, the size of the obtained cyclopeptides.

Last but not least, the reaction of aminothiols with aminonitriles can be generally used to couple any two peptides. Therefore, this chemoselective reaction can be considered as an alternative approach to chemical ligation technique in peptide synthesis.

Experimental Part

General experimental methods

The standard handling techniques under an inert atmosphere were used for sensitive air and moisture reagents. The reactions were monitored by thin layer chromatography (TLC) using commercial aluminum-backed silica gel plates (Merck, Kieselgel 60 F254). TLC spots were viewed under ultraviolet light, by heating the plate after treatment with Phosphomolybdic acid (PMA) solution or by heating the plate after treatment with Ninhydrin solution (1.5 g/100 mL n-butanol and 3 mL acetic acid). Product purifications by gravity column chromatography were performed using Macherey-Nagel silica gel 60 (70–230 mesh).

¹H-NMR and ¹³C-NMR spectra were recorded on a Bruker Avance 500 (¹H: 500 MHz, ¹³C: 125 MHz) and 400 (¹H: 400 MHz, ¹³C: 100 MHz) spectrometers.

Chemical shifts for ¹H spectra are values from tetramethylsilane in CDCl₃ (δ 0.00 ppm), MeOH peak in CD₃OD (δ 3.31 ppm) or relative to H₂O peak in D₂O (calibrated at 4.7 ppm). Chemical shifts for ¹³C spectra are values from CDCl₃ (δ 77.16 ppm), or CD₃OD (δ 49.00 ppm).

¹H-NMR spectra are reported as following: chemical shift (ppm), multiplicity (s: singlet; d: doublet; dd: doublet of doublets; t: triplet; sex: sextet; sep: septet; m: multiplet), coupling constants (Hz) and integration. Proton and carbon signal assignments were established using COSY, HSQC, and HMBC experiments.

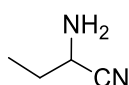
High-resolution mass spectra (HRMS) were recorded on a Waters G2-S Q-TOF mass spectrometer and LTQ Orbitrap XL (Thermo Scientific). HPLC analysis of compounds was performed on a Thermo Scientific Dionex Ultimate 3000 apparatus equipped with a Varian 380-LC DEDL detector. MS/MS experiments: Low resolution ESI analysis was performed on amazon speed (Bruker Daltonics) - IonTrap Spectrometer.

For MALDI analysis, DMSO was used as a solvent to dissolve the obtained polymers and DHB (2,5-Dihydroxybenzoic acid) was used as a matrix.

General procedure for the preparation of racemic amino nitrile:

NH₄Cl (3 equiv.) and NaCN (1 equiv.) were dissolved in distilled water (20 mL). The mixture was cooled to 0 °C for 5 minutes and then the aldehyde (propionaldehyde, isobutyraldehyde or 3-(Methylthio)propionaldehyde) (1 equiv.) was added. The mixture was left at room temperature for 4 hours. Organic layer was extracted with DCM (3 × 20 mL), dried over MgSO₄, concentrated under vacuum. Flash chromatography afforded the amino nitriles.

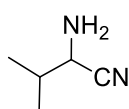
2-aminobutanenitrile (1)



Following the general procedure, NH₄Cl (2.21 g, 41.31 mmol), NaCN (675 mg, 13.77 mmol) and propionaldehyde (800 mg, 13.77 mmol). Crude product was concentrated and purified by silica column chromatography (95-50 % pentane/EA). The product was collected, and concentrated under vacuum to give **1** (580 mg, 50 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 3.58 (t, *J* = 6.9 Hz, 1H), 1.73 (s, 2H), 1.63-1.71 (m, 2H); ¹³C-NMR (125 MHz, CDCl₃) (δ, ppm): 122.18, 44.65, 28.51, 9.71.

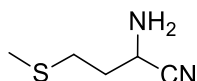
2-amino-3-methylbutanenitrile (2)¹⁷⁴



Following the general procedure, NH₄Cl (8 g, 133.3 mmol), NaCN (2.18 g, 44.4 mmol) and isobutyraldehyde (3.2 g, 44.4 mmol). Crude product was concentrated and purified by silica column chromatography (95-50 % pentane/EA), and the product was collected, concentrated under vacuum to give **2** (1.5 g, 35 %).

¹H-NMR (400 MHz, CDCl₃) (δ, ppm): 3.49 (d, *J* = 5.62 Hz, 1H), 1.89 (m, 1H), 1.05 (d, *J* = 6.60 Hz, 3H), 1.03 (d, *J* = 6.60 Hz, 3H); ¹³C-NMR (100 MHz, CDCl₃) (δ, ppm): 121.22, 49.78, 32.87, 18.78, 17.56.

2-amino-4-(methylthio)butanenitrile (3)



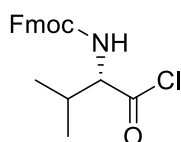
Following the general procedure, NH_4Cl (1.60 g, 3 mmol), NaCN (0.47 g, 9.6 mmol) and 3-(Methylthio)propionaldehyde (1 g, 9.6 mmol). Crude product was concentrated and purified by silica column chromatography (90-20 % pentane/EA), and the product was collected, concentrated under vacuum to give **3** (0.53 g, 45 %).

$^1\text{H-NMR}$ (400 MHz, CDCl_3) (δ , ppm): 3.92 (s, 1H), 2.63-2.77 (m, 2H), 2.13 (s, 3H), 1.98-2.08 (m, 2H), 1.64 (s, 2H); $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) (δ , ppm): 121.97, 42.12, 34.55, 30.02, 15.59.

General procedure for the preparation of the acid chlorides of Fmoc-L amino acids:

To a solution of Fmoc amino acid (1 equiv.) in dry DCM (5 mL) was added SOCl_2 (10 equiv.) gently at 0 °C. The solution was stirred for 15 minutes at rt. The mixture was concentrated under vacuum to give the product that was taken to the next step without further purification.

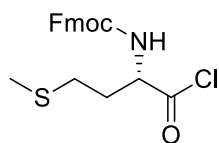
(S)-(9H-fluoren-9-yl)methyl (1-chloro-3-methyl-1-oxobutan-2-yl)carbamate (4)¹⁷⁵



Following to the general procedure, Fmoc-L-Val-OH (500 mg, 1.47 mmol) reacted with SOCl_2 (1.75 g, 14.7 mmol) to give a yellowish solid (quantitative).

$^1\text{H-NMR}$ (400 MHz, CDCl_3) (δ , ppm): 7.66 (d, $J = 7.5$ Hz, 2H), 7.51 (t, $J = 5.7$ Hz, 2H), 7.31 (t, $J = 7.4$ Hz, 2H), 7.23 (t, $J = 7.3$ Hz, 2H), 5.56 (d, $J = 8.8$ Hz, 1H), 4.34-4.49 (m, 2H), 4.13 (t, $J = 6.7$ Hz, 1H), 2.30 (q, $J = 6.2$ Hz, 1H), 0.97 (d, $J = 6.5$ Hz, 3H), 0.87 (d, $J = 6.7$ Hz, 3H); $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) (δ , ppm): 174.75, 156.29, 143.80, 143.64, 141.43, 127.91, 127.22, 125.09, 125.09, 120.16, 68.23, 67.46, 47.20, 29.95, 19.34, 17.22.

(S)-(9H-fluoren-9-yl)methyl (1-chloro-4-(methylthio)-1-oxobutan-2-yl)carbamate (5)



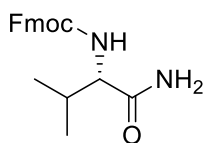
Following to the general procedure, Fmoc-L-Met-OH (540 mg, 1.45 mmol) reacted with SOCl_2 (1.73 g, 14.5 mmol) to give a yellowish solid (quantitative).

$^1\text{H-NMR}$ (400 MHz, CDCl_3) (δ , ppm): 7.73 (d, $J = 7.5$, 2H), 7.55 (d, $J = 6.7$ Hz, 2H), 7.37 (t, $J = 7.4$, 2H), 7.29 (t, $J = 7.2$ Hz, 2H), 5.56 (d, $J = 7.79$, 1H), 4.72 (d, $J = 3.6$ Hz, 1H), 4.45 (m, 2H), 4.19 (t, $J = 6.6$ Hz, 1H), 2.78-2.64 (m, 2H), 2.19-2.33 (m, 1H), 2.07 (s, 3H), 1.94-2.04 (m, 1H); **$^{13}\text{C-NMR}$ (100 MHz, CDCl_3) (δ , ppm):** 174.75, 156.29, 143.80, 143.64, 141.43, 127.91, 127.22, 125.09, 125.09, 120.16, 68.23, 67.46, 47.20, 29.95, 19.34, 17.22.

General procedure for the preparation of the amides of Fmoc-L amino acids:

Fmoc amino acid chloride was dissolved in DCM (15 mL). The mixture was cooled to 0 °C then ammonia gas was bubbled for 1-2 min. Excess ammonia and DCM were evaporated gently under vacuum. The mixture was then quenched with water and the product was extracted in EA (3 X 20 mL), dried over Na_2SO_4 and concentrated under vacuum.

(S)-(9H-fluoren-9-yl)methyl (1-amino-3-methyl-1-oxobutan-2-yl)carbamate (6)¹⁷⁶

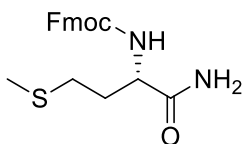


Following to the general procedure, Fmoc-L-Val-Cl **4** (1.05 g, 2.59 mmol) reacted with ammonia. Crude product was concentrated and purified by silica column chromatography (95-50 % pentane/EA), and the product was collected, concentrated under vacuum to give **6** as a white solid (0.86 g, yield = 86 %).

$^1\text{H-NMR}$ (400 MHz, CDCl_3) (δ , ppm): 7.76 (d, $J = 7.6$ Hz, 2H), 7.58 (d, $J = 7.4$ Hz, 2H), 7.40 (t, $J = 7.5$ Hz, 2H), 7.31 (t, $J = 7.5$ Hz, 2H), 5.86 (s, 1H), 5.59 (s, 1H), 5.39 (d, $J = 7.0$ Hz, 1H), 4.42 (s, 2H), 4.21 (t, $J = 6.7$ Hz, 1H), 4.03 (br, 1H), 2.07-2.21 (m, 1H), 0.98 (d, $J =$

6.4 Hz, 3H), 0.95 (d, $J = 6.4$ Hz, 3H); $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) (δ , ppm): 173.45, 143.76, 141.35, 127.75, 127.12, 125.03, 120.01, 66.97, 47.25, 30.80, 29.72, 19.19, 17.76.

(S)-(9H-fluoren-9-yl)methyl (1-amino-4-(methylthio)-1-oxobutan-2-yl)carbamate (7)¹⁷⁶



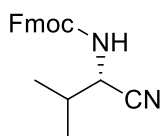
Following to the general procedure, Fmoc-L-Met-Cl **5** (1 g, 2.56 mmol) reacted with ammonia. Crude product was concentrated and purified by silica column chromatography (95-50 % pentane/EA), and the product was collected, concentrated under vacuum to give **7** as a white solid (0.83 g, yield = 95 %).

$^1\text{H-NMR}$ (500 MHz, CDCl_3) (δ , ppm): 7.8 (d, $J = 7.87$ Hz, 2H), 7.6 ppm (m, 2H), 7.4 (t, $J = 7.42$ Hz, 2H), 7.3 (t, $J = 7.42$ Hz, 2H), 4.45 (d, $J = 6.21$ Hz, 2H), 4.4 (m, 1H), 4.2 (t, $J = 6.33$ Hz, 1H), 2.66 (m, 2H), 2.1 (m, 2H), 2 (s, 3H); $^{13}\text{C-NMR}$ (500 MHz, CDCl_3) (δ , ppm): 173.26, 143.82, 141.50, 127.93, 127.26, 120.19, 67.12, 53.6, 47.37, 31.42, 30.24, 15.38

General procedure for the preparation of the nitriles of Fmoc-L amino acids:

Amide **6** or **7** (1 equiv.) was dissolved in DMF (6 mL). To this solution was added cyanuric chloride (1 equiv.). The mixture was stirred for 2 h at rt. Then, 20 mL of water was added and the aqueous phase was extracted with EA three times. The organic layer was then recovered and washed with water 5 times to get rid of remaining DMF. The resulted solution was concentrated under reduced pressure and purified by silica column chromatography.

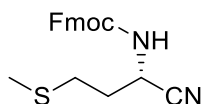
(S)-(9H-fluoren-9-yl)methyl (1-cyano-2-methylpropyl)carbamate (8)¹⁷⁷



Following to the general procedure, Fmoc-L-Val-NH₂ **6** (860 mg, 2.5 mmol) reacted with cyanuric chloride (940 mg, 5 mmol). The product was purified by silica column chromatography (95-75 % pentane/EA), collected and concentrated under vacuum to give **8** as a white solid (651 mg, 80 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.71 (d, *J* = 7.5 Hz, 2H), 7.53 (d, *J* = 7.4 Hz, 2H), 7.35 (t, *J* = 7.4 Hz, 2H), 7.26 (t, *J* = 7.4 Hz, 2H), 5.50 (d, *J* = 8.9 Hz, 1H), 4.34-4.58 (m, 3H), 4.14 (t, *J* = 6.2 Hz, 1H), 1.98 (m, 1H), 1.04 (d, *J* = 6.4 Hz, 3H), 1.00 (d, *J* = 6.5 Hz, 3H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 155.55, 143.52, 141.38, 127.90, 127.20, 125.00, 120.12, 117.91, 67.47, 49.08, 47.10, 31.82, 18.58, 18.07.

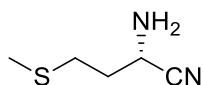
(S)-(9H-fluoren-9-yl)methyl (1-cyano-3-(methylthio)propyl)carbamate (9)¹⁷⁸



Following to the general procedure, Fmoc-L-Met-NH₂ **7** (0.86 g, 2.32 mmol) reacted with cyanuric chloride (0.48 g, 2.32 mmol). The product was purified by silica column chromatography (95-70 % pentane/EA), collected and concentrated under vacuum to give **9** as a white solid (0.69 g, 84 %).

m.p. 136.5-137.5 °C; **[α]²⁰_D** -33.33 (*c* 1.00, CHCl₃); **¹H-NMR (500 MHz, CDCl₃) (δ, ppm):** 7.76 (d, *J* = 7.28 Hz, 2H), 7.57 ppm (d, *J* = 6.24 Hz, 2H), 7.41 (t, *J* = 7.81 Hz, 2H), 7.33 (t, *J* = 7.29 Hz, 2H), 5.33 (d, *J* = 7.5 Hz, 1H), 4.84 (d, *J* = 7.5, 1H), 4.50 (d, *J* = 6.9 Hz, 2H), 4.21 (s, 1H), 2.63 (s, 2H), 2.00-2.21 (m, 2H), 2.11 (s, 3H); **¹³C-NMR (500 MHz, CDCl₃) (δ, ppm):** 155.06, 143.42, 141.38, 127.90, 127.18, 124.91, 118.05, 67.46, 47.07, 41.86, 32.17, 29.64, 15.47.

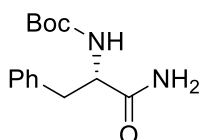
(S)-2-amino-4-(methylthio)butanenitrile (11)



Fmoc-L-Met-CN **9** (474 mg, 1.35 mmol) was dissolved in 6 mL of 2/1 mixture of DCM /Piperidine. The reaction was followed by TLC at rt. After 15 min, the mixture was concentrated and was purified by silica column chromatography (98-10 % pentane/EA), collected and concentrated under vacuum to give **11** as a yellowish oil (140 mg, 80 %).

[α]²⁰_D -24.02 (*c* 0.97, CHCl₃); **¹H-NMR (400 MHz, CDCl₃) (δ, ppm):** 3.92 (t, *J* = 7.1 Hz, 1H), 2.62-2.78 (m, 2H), 2.12 (s, 3H), 1.97-2.07 (m, 2H), 1.67 (s, 1H); **¹³C-NMR (100 MHz, CDCl₃) (δ, ppm):** 121.85, 41.98, 34.40, 29.88, 29.88, 15.44.

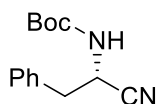
(S)-tert-butyl (1-amino-1-oxo-3-phenylpropan-2-yl)carbamate (13)¹⁷⁹



Boc-Phe-OH (1 g, 3.77 mmol) was dissolved in 40 mL of THF at -20 °C under argon. NMM (622 μ L, 5.65 mmol) and IBCF (733 μ L, 5.65 mmol) were added respectively and the reaction left to stir for 15 min. Ammonia gas was then introduced to the solution for 4 minutes and then left stirring for overnight at rt. The mixture was gently evaporated, quenched with water and extracted with DCM. Organic layer was collected, dried over Na₂SO₄ and evaporated under vacuum. The crude was submitted to silica column chromatography (95-20 % pentane/EA). The solvent was collected and concentrated under vacuum to give **13** as a white solid (quantitative).

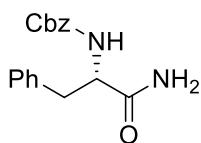
¹H-NMR (500 MHz, CDCl₃) (δ , ppm): ¹H-NMR (400 MHz, CDCl₃) δ 7.35-7.19 (m, 5H), 5.71 (s, 1H), 5.30 (s, 1H), 5.02 (s, 1H), 4.35 (s, 1H), 3.07 (d, J = 6.6 Hz, 2H), 1.41 (s, 9H).

(S)-tert-butyl (1-cyano-2-phenylethyl)carbamate (14)¹⁸⁰



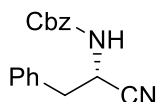
Boc-Phe-NH₂ **13** (345 mg, 1.3 mmol) was dissolved in 6 mL DMF. Cyanuric chloride (184.4 mg, 1.3 mmol) was added. The mixture was stirred for 2 h at rt. Then it was quenched with water, and extracted with EA (3 x 20 mL). Organic phase was washed with water (5 x 20 mL), dried over NaSO₄ and evaporated under vacuum to yield a white solid (95 %).

¹H-NMR (500 MHz, CDCl₃) (δ , ppm): 7.41-7.27 (m, 5H), 4.80 (d, J = 31.4 Hz, 2H), 3.11 (dd, J = 13.8, 5.8 Hz, 1H), 3.05 (dd, J = 13.8, 7.1 Hz, 1H), 1.44 (s, 9H); **¹³C-NMR (125 MHz, CDCl₃) (δ , ppm):** 133.89, 129.51, 129.02, 127.94, 121.74, 118.34, 81.45, 45.06, 39.20, 28.20.

(S)-benzyl (1-amino-1-oxo-3-phenylpropan-2-yl)carbamate (17)¹⁷⁶

Cbz-Phe-OH (250 mg, 0.84 mmol) was dissolved in 30 mL THF at -20 °C under argon. NMM (138 μ L, 1.25 mmol) and IBCF (162 μ L, 1.25 mmol) were added respectively and the reaction left to stir for 15 min. Ammonia gas was then introduced to the solution for 5 minutes and then left stirring for overnight at rt. The mixture was gently evaporated, quenched with water and extracted with DCM. Organic layer was collected, dried over Na₂SO₄ and evaporated under vacuum. The crude was submitted to silica column chromatography (0.2-1.5 % MeOH/DCM). The solvent was collected and concentrated under vacuum to give **17** as a white solid (91 %).

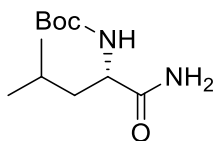
¹H-NMR (500 MHz, CDCl₃) (δ , ppm): 7.40-7.16 (m, 10), 5.61 (s, 1H), 5.30 (d, J = 3.8 Hz, 2H), 5.10 (s, 2H), 4.43 (d, J = 6.6 Hz, 1H), 3.14 (dd, J = 13.6, 6.1 Hz, 1H), 3.05 (dd, J = 13.8, 7.4 Hz, 1H).

(S)-benzyl (1-cyano-2-phenylethyl)carbamate (18)¹⁸⁰

Cbz-Phe-NH₂ **17** (220 mg, 0.74 mmol) was dissolved in 7 mL DMF. Cyanuric chloride (136 mg, 0.74 mmol) was added. The mixture was stirred for 2 h at rt. Then it was quenched with water, and extracted three times with EA. Organic phase was washed five times with water, dried over NaSO₄ and evaporated under vacuum. The crude was submitted to silica column chromatography (95-70 % pentane/EA). The solvent was collected and concentrated under vacuum to give **18** as a white solid (78 %).

[α]²⁰_D -20.9 (c 1.00, CHCl₃); **¹H-NMR (400 MHz, CDCl₃) (δ , ppm):** 7.42-7.26 (m, 10H), 5.13 (s, 2H), 4.98 (s, 1H), 4.90 (s, 1H), 3.13 (dd, J = 13.8, 5.6 Hz, 1H), 3.06 (dd, J = 13.8, 6.9 Hz, 1H); **¹³C-NMR (100 MHz, CDCl₃) (δ , ppm):** 135.50, 133.54, 129.49, 129.11, 128.66, 128.55, 128.33, 128.07, 117.99, 67.81, 43.75, 39.03, 30.93.

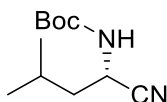
(S)-tert-butyl (1-amino-4-methyl-1-oxopentan-2-yl)carbamate (19)¹⁸¹



Boc-Leu--OH (1 g, 4.32 mmol) was dissolved in 40 mL THF at -20 °C under argon. NMM (713 μ L, 6.49 mmol) and IBCF (841 μ L, 6.49 mmol) were added respectively and the reaction left to stir for 15 min. Ammonia gas was then introduced to the solution for 5 minutes and then left stirring for overnight at rt. The mixture was gently evaporated, quenched with water and extracted with EA. Organic layer was collected, dried over Na₂SO₄ and evaporated under vacuum. The crude was submitted to silica column chromatography (95-15 % Pentane/EA). The solvent was collected and concentrated under vacuum to give **19** as a white solid (65 %).

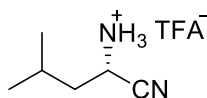
¹H-NMR (500 MHz, CDCl₃) (δ , ppm): 6.19 (s, 1H), 5.55 (s, 1H), 4.91 (d, J = 7.9 Hz, 1H), 4.14 (s, 1H), 1.78-1.61 (m, 2H), 1.50 (dt, J = 13.7, 7.2 Hz, 1H), 1.44 (s, 9H), 0.94 (t, J = 6.1 Hz, 6H); ¹³C-NMR (125 MHz, CDCl₃) (δ , ppm): 175.06, 155.81, 80.22, 52.58, 41.08, 28.32, 24.76, 22.96, 21.93.

(S)-tert-butyl (1-cyano-3-methylbutyl)carbamate (20)¹⁸⁰



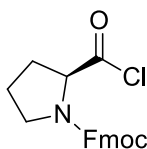
Cbz-Phe-NH₂ **19** (510 mg, 2.22 mmol) was dissolved in 7 mL DMF. Cyanuric chloride (408 mg, 2.22 mmol) was added. The mixture was stirred for 2 h at rt. Then it was quenched with water, and extracted three times with EA. Organic phase was washed five times with water, dried over NaSO₄ and evaporated under vacuum. The crude was submitted to silica column chromatography (95-80 % pentane/EA). The solvent was collected and concentrated under vacuum to give **20** as a pale yellow oil (92 %).

¹H-NMR (500 MHz, CDCl₃) (δ , ppm): 4.74 (s, 1H), 4.59 (d, J = 5.0 Hz, 1H), 1.90-1.78 (m, 1H), 1.75-1.57 (m, 2H), 1.47 (s, 9H), 0.98 (d, J = 6.6 Hz, 6H); ¹³C-NMR (125 MHz, CDCl₃) (δ , ppm): 154.27, 119.13, 81.20, 42.20, 40.87, 28.24, 24.80, 22.21, 21.85.

(S)-1-cyano-3-methylbutan-1-aminium trifluoroacetate (21)

The nitrile **20** (225 mg, 1.06 mmol) was dissolved in DCM (4 mL), *i*Pr₃SiH (870 μL, 4.24 mmol) and TFA (817 μL, 10.6 mmol) were added respectively. The mixture was stirred for 1 h and then was extracted in D₂O (1 mL) as a TFA salt. Finally, the aqueous layer was left under vacuum at room temperature to remove the remaining DCM.

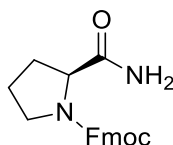
¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.48 (q, *J* = 5.3 Hz, 1H), 1.90-1.96 (m, 1H), 1.54-1.85 (m, 2H), 0.92 (d, *J* = 6.2 Hz, 6H); **¹³C-NMR** (125 MHz, D₂O) (δ, ppm): 116.10, 39.81, 38.72, 24.54, 21.54, 20.41.

(S)-(9H-fluoren-9-yl)methyl 2-(chlorocarbonyl)pyrrolidine-1-carboxylate (27)¹⁷⁵

To a solution of Fmoc-Pro-OH (1 g, 2.96 mmol) in dry DCM (10 mL) was added oxalyl chloride (COCl)₂ (500 μL, 5.93 mmol) gently at 0 °C, 1 drop of DMF was added. The solution was stirred for 2 h at rt. The mixture was evaporated gently at 35 °C to give a pale green oil that was taken to the next step without further purification.

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.77 (dd, *J* = 7.4, 4.2 Hz, 2H), 7.64-7.51 (m, 2H), 7.41 (t, *J* = 7.4 Hz, 2H), 7.36-7.28 (m, 2H), 4.71-4.33 (m, 3H), 4.30-4.14 (m, *J* = 12.1, 6.5 Hz, 1H), 3.69-3.45 (m, 2H), 2.40-2.18 (m, 2H), 2.10-1.84 (m, 2H); **¹³C-NMR** (125 MHz, CDCl₃, both conformers were observed) (δ, ppm): 174.41, 174.38, 163.74, 154.74, 153.85, 143.88, 143.81, 143.68, 143.53, 141.46, 141.35, 141.32, 141.27, 127.78, 127.76, 127.74, 127.17, 127.10, 127.09, 127.06, 125.07, 124.99, 124.79, 124.72, 120.09, 120.01, 67.93, 67.79, 67.57, 67.31, 47.20, 47.18, 47.12, 46.65, 30.91, 30.49, 29.34, 24.06, 22.96.

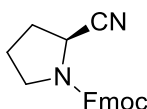
(S)-(9H-fluoren-9-yl)methyl 2-cyanopyrrolidine-1-carboxylate (28)



The obtained Fmoc-Pro-Cl in the previous step **27** was dissolved in DCM (15 mL). The mixture was cooled to 0 °C then ammonia gas was bubbled for 1 min. Excess ammonia and DCM were evaporated gently under vacuum. The mixture was then quenched with water and the product was extracted in EA (3 X 20 mL), dried over Na₂SO₄ and concentrated under vacuum. The crude was submitted to silica column chromatography (80-10 % pentane/EA, then with 1 % MeOH/DCM). The product was collected and concentrated under vacuum to give **28** as a pale yellow oil (84 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.77 (d, *J* = 7.5 Hz, 2H), 7.59 (d, *J* = 7.3 Hz, 2H), 7.40 (t, *J* = 7.4 Hz, 2H), 7.32 (t, *J* = 7.4 Hz, 2H), 6.57 (s, 1H), 5.30 (s, 1H), 4.73-4.17 (m, 4H), 3.58-3.26 (m, 2H), 2.38 (s, 1H), 2.23-1.72 (m, 3H).

(S)-(9H-fluoren-9-yl)methyl 2-cyanopyrrolidine-1-carboxylate (29)

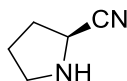


Cbz-Pro-NH₂ **28** (515 mg, 1.53 mmol) was dissolved in 10 mL DMF. Cyanuric chloride (282.3 mg, 1.53 mmol) was added. The mixture was stirred for 1 h at rt. Then it was quenched with water, and extracted three times with EA. Organic phase was washed five times with water, dried over Na₂SO₄ and evaporated under vacuum. The crude was submitted to silica column chromatography (95-70 % pentane/EA). The solvent was collected and concentrated under vacuum to give **29** as a colourless oil (81 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.77 (d, *J* = 7.5 Hz, 2H), 7.71-7.63 (m, 2H), 7.59 (t, *J* = 7.3 Hz, 2H), 7.41 (t, *J* = 7.4 Hz, 2H), 7.38-7.28 (m, 2H), 4.65-4.17 (m, 4H), 3.65-3.23 (m, 2H), 2.34-1.92 (m, 4H); **¹³C-NMR (125 MHz, CDCl₃, both conformers were observed) (δ, ppm):** 154.33, 153.78, 143.92, 143.83, 143.71, 143.62, 141.36, 127.88, 127.26, 127.19,

125.17, 125.07, 125.03, 120.10, 119.01, 118.80, 68.21, 67.86, 47.56, 47.18, 47.04, 46.45, 45.99, 31.88, 30.76, 24.72, 23.73.

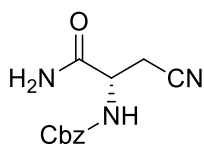
(S)-pyrrolidine-2-carbonitrile (26)



Fmoc-L-Pro-CN **29** (582 mg, 1.83 mmol) was dissolved in 10 mL of 3/1 mixture of DCM /Piperidine. The reaction was followed by TLC rt. After 15 min, the mixture was concentrated and was purified by silica column chromatography (98-10 % pentane/EA), collected and concentrated under vacuum to give **26** (quantitative).

¹H-NMR (400 MHz, CDCl₃) (δ, ppm): 3.90 (d, *J* = 6.9 Hz, 1H), 3.12-2.80 (m, 1H), 2.77-2.48 (m, 1H), 2.30-2.07 (m, 2H), 2.05-1.73 (m, 2H), 0.95-0.74 (m, *J* = 6.6 Hz, 1H); **¹³C-NMR (100 MHz, CDCl₃) (δ, ppm):** 118.78, 118.09, 51.95, 51.74, 49.92, 49.48, 29.60, 29.54, 22.38, 21.91.

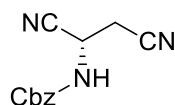
(S)-benzyl (1-amino-3-cyano-1-oxopropan-2-yl)carbamate (30)



Cbz-Asn-OH (2 g, 7.54 mmol) was dissolved in 100 mL THF. NMM (1.65 mL, 15.03 mmol) and IBCF (1.95 mL, 15.03 mmol) were added at 0 °C. NH₃ (gas) was bubbled in the mixture during 30 minutes at 0 °C. The mixture was stirred overnight at RT. The mixture was concentrated under reduced pressure, then water was added and the organic layer was extracted 3 times with EA (3 x 30 mL). The organic phase was evaporated under pressure to obtain a white solid. The product was purified using silica column with (90-25 % Pentane/EA. White solid (49 %).

¹H-NMR (CD₃OD, 500 MHz) (δ, ppm): 7.23 (m, 5H), 5.15 (q, 2H), 4.51 (dd, X of ABX, *J*_{AX} = 5.4 Hz, *J*_{BX} = 8.2 Hz, 1H), 3.02-2.84 (2 dd, AB of ABX, *J*_{AX} = 5.4 Hz, *J*_{BX} = 8.2 Hz, *J*_{AB} = 17 Hz, 2H); **¹³C-NMR (CD₃OD, 125 MHz) (δ, ppm):** 172.2, 156.9, 136.5, 128.1, 127.7, 127.5, 116.9, 66.7, 51.0, 20.0.

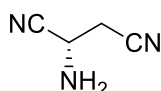
(S)-benzyl (1,2-dicyanoethyl)carbamate (31)



30 (900 mg, 3.64 mmol) was dissolved in DMF (7 mL). To this solution was added cyanuric chloride (336 mg, 1.82 mmol). The mixture was stirred for 3 h at rt. The 20 mL of water was added and the aqueous phase was extracted with EA (3 x 25 mL). The organic layer was separated, concentrated under reduced pressure to obtain **31** as a white solid (91 %).

¹H-NMR (CD₃OD, 500 MHz) (δ, ppm): 7.22 (m, 5H), 5.05 (s, 2H), 4.90 (t, *J* = 7.0 Hz, 1H), 3.00 (d, *J* = 7.0 Hz, 2H); **¹³C-NMR (CD₃OD, 125 MHz) (δ, ppm):** 155.9, 136.2, 128.2, 127.9, 127.6, 116.4, 115.2, 66.9, 39.4, 21.5.

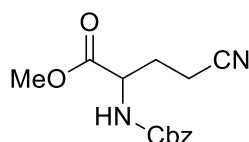
(S)-2-aminosuccinonitrile (32)



Bis-nitrile **31** (680 mg, 2.97 mmol) was dissolved in 20 mL acetonitrile. Catalytic amount of Pd/C was added, and the solution was stirred under hydrogen atmosphere for 6 h at rt. Then, the mixture was filtered on celite and washed with methanol. The filtrate was concentrated under reduced pressure to obtain **32** as a yellow oil (60 %)

¹H-NMR (CD₃OD, 400 MHz) (δ, ppm): 4.04 (t, *J* = 5.2 Hz, 1H), 2.84 (d, *J* = 5.2 Hz, 2H); **¹³C-NMR (CD₃OD, 100 MHz) (δ, ppm):** 119.6, 115.9, 40.1, 23.6.

methyl 2-(((benzyloxy)carbonyl)amino)-4-cyanobutanoate (33)



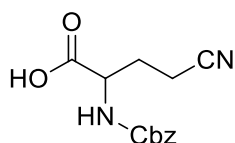
Cbz-Gln-OMe (1g, 3.41 mmol) was dissolved in 6 mL DMF. Cyanuric chloride (328 mg, 1.71 mmol) was added. After stirred for 1.5 h at rt, 10 mL water was added and the mixture was extracted 4 times with EA. The organic layer was washed 5 times with water and

Experimental part

evaporated under reduced pressure. The crude product was purified using silica column with (90-50 % Pentane/EA to give product **33** as a white solid (64 %).

¹H-NMR (400 MHz CDCl₃) (δ, ppm): 7.36 (m, 5H), 5.47 (s, 1H), 5.12 (s, 2H), 4.45 (m, 1H), 3.79 (s, 3H), 2.42 (m, 2H), 2.30-2.05 (m, 2H); **¹³C-NMR (100 MHz CDCl₃) (δ, ppm):** 171.2, 155.9, 135.9, 128.6, 128.4, 128.2, 118.6, 67.4, 52.9, 52.8, 28.7, 13.6.

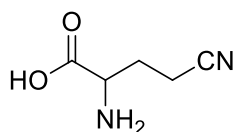
2-(((benzyloxy)carbonyl)amino)-4-cyanobutanoic acid (**34**)¹⁸²



33 (598 mg, 2.17 mmol) was dissolved in 3 mL *t*BuOH/H₂O (2 : 1). Then, LiOH (182 mg, 4.33 mmol) was added and the mixture was stirred for 1 h at 0 °C. The mixture was acidified with HCl (1 N) and extracted with EA (4 x 15 mL). The organic layer was evaporated under reduced pressure to give a yellowish oil (73 %).

¹H-NMR (400 MHz CDCl₃) (δ, ppm): 7.93 (br, 1H), 7.36 (m, 5H), 5.43 (d, 1H), 5.13 (s, 2H), 4.48 (m, 1H), 2.46 (m, 2H), 2.34-2.09 (m, 2H); **¹³C-NMR (100 MHz CDCl₃) (δ, ppm):** 174.1, 156.1, 135.7, 128.6, 128.4, 128.2, 118.5, 67.6, 52.6, 28.3, 13.6.

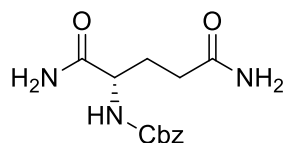
2-amino-4-cyanobutanoic acid (**35**)



The product **34** (411 mg, 3.21 mmol) was dissolved in 10 mL EA. A catalytic amount of Pd/C was added and the mixture was stirred under hydrogen atmosphere for 4 h at rt. Finally, the mixture was filtered by celite with EA to give the product **35** as white solid (60 %).

¹H-NMR (400 MHz, D₂O) (δ, ppm): 3.73 (t, *J* = 6.7 Hz, 1H), 2.63 (m, 2H), 2.16 (m, 2H); **¹³C-NMR (100 MHz, D₂O) (δ, ppm):** 173.0, 120.0, 53.3, 26.3, 13.5.

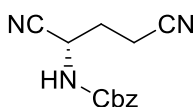
(S)-benzyl (1,5-diamino-1,5-dioxopentan-2-yl)carbamate (36)



SOCl₂ (0.52 mL, 7.14 mmol) was dissolved in 15 mL dry CH₃OH at 0 °C under argon. CbzGln (1 g, 3.57 mmol) was added to the solution, and stirred for overnight at rt. The mixture was evaporated, dry methanol was added at -10 °C and ammonia gas was bubbled during 30 min. The mixture was stirred for 48 h at rt. The white solid was separated from the obtained red solution by filtration to give **36** (50 %).

¹H-NMR (500 MHz, DMSO) (δ, ppm): 7.31 (m, 5H), 7.28, 7.01, 6.75 (s, 5H), 5.03 (s, 2H), 3.90 (m, 1H), 2.11 (m, 2H), 1.87-1.70 (m, 2H); **¹³C-NMR (125 MHz, DMSO) (δ, ppm):** 174.2, 174.1, 154.4, 137.5, 128.8, 128.3, 128.2, 65.9, 54.8, 32.0, 28.2.

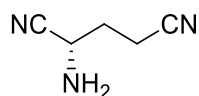
(S)-benzyl (1,3-dicyanopropyl)carbamate (37)



The product **36** (422 mg, 1.58 mmol) was dissolved in 12 mL DMF. Cyanuric chloride (292 mg, 1.58 mmol) was added to the mixture and left stirring for 3 h at rt. 10 mL water was added and the mixture was extracted with EA (5 x 10 mL). The organic phase was washed with water (5 x 10 mL), evaporated under reduced pressure to obtain **37** as a yellowish oil (40 %).

¹H-NMR (400 MHz, CDCl₃) (δ, ppm): 7.37 (m, 5H), 5.17 (s, 2H), 4.74 (m, 1H), 2.55 (m, 2H), 2.24 (m, 2H); **¹³C-NMR (100 MHz, CDCl₃) (δ, ppm):** 148.8, 135.2, 128.7, 128.4, 117.4, 116.8, 68.2, 41.8, 29.4, 13.8.

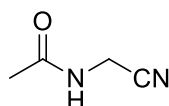
(S)-2-aminopentanedinitrile (38)



Product **37** (158 mg, 1.45 mmol) was dissolved in 6 mL EA. Catalytic amount of Pd/C was added to the mixture under hydrogen atmosphere and the mixture was stirred for 4 h at rt. The mixture was filtered by celite with EA to afford the product **38** as an orange oil (quantitative).

$[\alpha]^{20}_{\text{D}} -14.59$ (*c* 0.85, CHCl_3); $^1\text{H-NMR}$ (400 MHz, CDCl_3) (δ , ppm): 3.81 (dd, $J_{\text{AX}} = 6.2$ Hz, $J_{\text{BX}} = 8.4$ Hz, 1H), 2.63 (m, 2H), 2.12-2.03 (m, 2H), 1.76 (s, 2H); $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) (δ , ppm): 120.9, 118.4, 41.8, 30.8, 13.7.

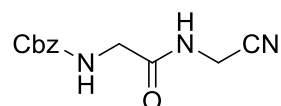
N-(cyanomethyl)acetamide (40)¹⁵⁰



Aminoacetonitrile **39** (500 mg, 5.4 mmol) was dissolved in 10 mL pyridine at 0 °C. acetic anhydride (0.76 mL, 8.1 mmol) was added to the solution and the mixture was stirred overnight at rt. Pyridine was evaporated, water was added, pH was increased to 8 by adding NaOH solution (1 N), and the mixture was extracted with EA (5 x 10 ml). The organic phase was evaporated under reduced pressure to give the product as a white solid (32 %).

$^1\text{H-NMR}$ (500 MHz, CDCl_3) (δ , ppm): 4.19 (d, $J = 5.9$ Hz, 2H), 2.07 (s, 3H); $^{13}\text{C-NMR}$ (125 MHz, CDCl_3) (δ , ppm): 169.8, 115.9, 27.5, 22.7.

benzyl (2-((cyanomethyl)amino)-2-oxoethyl)carbamate (41)¹⁵¹



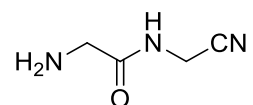
Cbz-Gly-OH (1 g, 4.78 mmol) was dissolved in dry THF at -20 °C under argon. NMM (525 μL , 4.78 mmol) and IBCF (620 μL , 4.78 mmol) were added respectively and the reaction left to stir for 20 min. Then, aminoacetonitrile hydrochloride **39** (442.3 mg, 5.74 mmol) was added and the mixture was left stirring under argon for overnight at rt. the mixture was

Experimental part

concentrated, quenched with water and extracted three times with EA. The organic phase was dried over Na_2SO_4 and evaporated. The crude product was submitted to silica column chromatography (95-30 % pentane/EA). The solvent was collected and concentrated under vacuum to give **41** as a white solid (93 %).

$^1\text{H-NMR}$ (500 MHz CDCl_3) (δ , ppm): 7.32-7.43 (m, 5H), 6.72 (br, 1H), 5.36 (br, 1H), 5.15 (s, 2H), 4.16 (d, $J = 5.8$ Hz, 2H), 3.90 (d, $J = 6.1$ Hz, 2H).

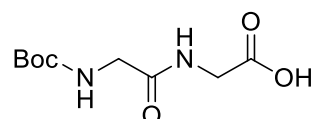
2-amino-N-(cyanomethyl)acetamide (**42**)



41 (400 mg, 3.54 mmol) was dissolved in 30 mL acetonitrile. Catalytic amount of Pd/C was added, and the solution was stirred under hydrogen atmosphere for 6 h at rt. Then, the mixture was filtered on celite and washed with methanol. The filtrate was concentrated under reduced pressure to obtain **42** as pale yellow oil (92 %).

$^1\text{H-NMR}$ (500 MHz D_2O) (δ , ppm): 4.19 (s, 2H), 3.66 (s, 2H); **$^{13}\text{C-NMR}$ (125 MHz D_2O) (δ , ppm):** 170.12, 116.82, 41.28, 27.52.

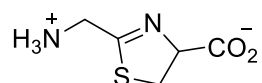
2-(2-((tert-butoxycarbonyl)amino)acetamido)acetic acid (**43**)¹⁵²



To a solution of diglycine (500 mg, 3.79 mmol) in aqueous NaOH (1 N) at 0 °C was added a solution of $(\text{Boc})_2\text{O}$ (910 mg, 4.17 mmol), in dioxane. The solution was left stirring for overnight at rt. The mixture was then concentrated and 7.5 mL NaOH (1 N) was added to the residue. Then, HCl (1 N) was added slowly to reach pH 2. The mixture was extracted three times with EA. The organic layer was collected, dried over NaSO_4 , concentrated under reduced vacuum to give **43** as a white solid (40 %).

$^1\text{H-NMR}$ (500 MHz, DMSO) (δ , ppm): 8.04 (t, $J = 5.8$ Hz, 1H), 6.98 (t, $J = 6.0$ Hz, 1H), 3.76 (d, $J = 5.8$ Hz, 2H), 3.57 (d, $J = 6.1$ Hz, 2H), 1.39 (s, 9H); **$^{13}\text{C-NMR}$ (125 MHz, DMSO) (δ , ppm):** 171.65, 170.20, 156.24, 78.54, 43.50, 41.03, 28.66.

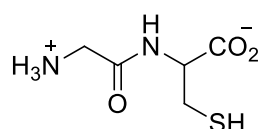
2-(ammoniomethyl)-4,5-dihydrothiazole-4-carboxylate (46)



Cysteine (3.9 mg, 0.032 mmol) and aminoacetonitrile hydrochloride (3 mg, 0.032 mmol) were dissolved in D₂O (0.7 mL) in a NMR tube under Argon.

The product was identified in the reaction mixture. **¹H-NMR (500 MHz D₂O) (δ, ppm):** 3.46 - 3.71 (AB of ABX system, 2 H), 3.83 (s, 2 H), 4.98-5.03 (dd ~t, 1 H); **¹³C-NMR (125 MHz D₂O) (δ, ppm):** 40.56, 37.52, 78.97, 166.84, 178.14; **HRMS (ESI):** calcd for C₅H₉N₂O₂S [M + H]⁺: 161.0385, found 161.0387.

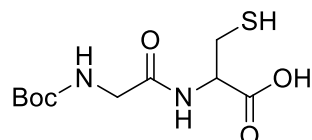
2-(2-ammonioacetamido)-3-Sulfanylpropanoate (47)



Cysteine (3.9 mg, 0.032 mmol) and aminoacetonitrile hydrochloride (3 mg, 0.032 mmol) were dissolved in D₂O (0.7 mL) in a NMR tube under Argon.

The product was identified in the reaction mixture. **¹H-NMR (500 MHz D₂O) (δ, ppm):** 2.84-2.93 (~d, 2 H), 3.84 (s, 2 H), 4.38 (t, *J* = 5.3 Hz, 1 H); **¹³C-NMR (125 MHz D₂O) (δ, ppm):** 26.09, 40.51, 56.83, 166.6, 176.06; **HRMS (ESI):** calcd for C₅H₁₁N₂O₃S [M + H]⁺: 179.0490, found 179.0492.

2-(2-((tert-butoxycarbonyl)amino)acetamido)-3-Sulfanylpropanoic acid (49)

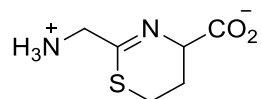


The ester **48** (85 mg, 0.28 mmol) was dissolved in 9 mL (THF/H₂O 2/1 solution). LiOH (29.1 mg, 0.69 mmol) was added at 0 °C. The mixture was stirred for 2 h at 0 °C. The mixture was extracted three times with EA. The organic layer was collected, dried over Na₂SO₄, concentrated under reduced vacuum to give **49** as a white solid (65 %).

Experimental part

¹H-NMR (400 MHz, CDCl₃) (δ, ppm): 9.39 (s, 1H), 7.41 (s, 1H), 5.72 (s, 1H), 4.87 (s, 1H), 4.07-3.61 (m, 2H), 3.15-2.90 (m, 2H), 1.58 (s, 1H), 1.45 (s, 9H); **¹³C-NMR (100 MHz, CDCl₃) (δ, ppm):** 172.22, 170.45, 156.60, 80.81, 53.76, 44.02, 28.30, 26.55.

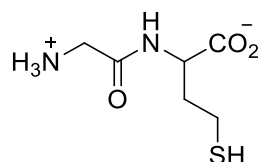
2-(ammoniomethyl)-5,6-dihydro-4H-1,3-thiazine-4-carboxylate (52)



Homocysteine (4.3 mg, 0.032 mmol) and aminoacetonitrile hydrochloride (3 mg, 0.032 mmol) were dissolved in D₂O (0.7 mL) in a NMR tube under Argon. The product was identified in the reaction mixture.

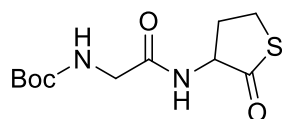
¹H-NMR (500 MHz, D₂O) (δ, ppm): δ 4.17-4.12 (m, 1H), 3.76 (t, *J* = 1.9 Hz, 2H), 3.16-3.08 (m, 1H), 3.04-2.97 (m, 1H), 2.16-2.07 (m, 2H), 1.86-1.76 (m, 1H); **¹³C-NMR (125 MHz, D₂O) (δ, ppm):** 179.71, 156.74, 60.62, 44.57, 24.37, 22.29; **HRMS (ESI):** calcd for C₆H₁₁N₂O₂S [M + H]⁺: 175.0541, found 175.0542.

2-(2-ammonioacetamido)-4-Sulfanylbutoanoate (53)



The product was identified in the reaction mixture. **¹H-NMR (500 MHz D₂O) (δ, ppm):** 4.28 (q, *J* = 4.6, 1 H), 3.75 (s, 2 H), 2.43-2.57 (m, 2 H), 1.88-2.03 (m, 2 H); **¹³C-NMR (500 MHz D₂O) (δ, ppm):** 178.36, 167.02, 54.09, 40.63, 35.58, 20.43; **HRMS (ESI):** calcd for C₆H₁₂N₂O₃S [M + H]⁺: 193.0647, found 193.0650.

tert-butyl (2-oxo-2-((2-oxotetrahydrothiophen-3-yl)amino)ethyl)carbamate (55)



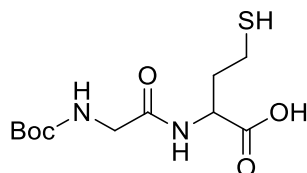
Boc-Gly-OH (500 mg, 2.85 mmol) was dissolved in 30 mL THF. NMM (0.85 mL, 7.7 mmol) and IBCF (0.55 mL, 4.28 mmol) were added at - 20 °C. After 15 min, Hcy-thiolactone

Experimental part

hydrochloride (526 mg, 3.48 mmol) was added to the mixture. The mixture was stirred overnight at rt. The mixture was concentrated under reduced pressure, then water was added and the organic layer was extracted 3 times with EA. The organic phase was evaporated under pressure. The crude product was purified using silica column with (90-50 % Pentane/EA to provide **55** as a white solid (88 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.11 (s, 1H), 5.59 (s, 1H), 4.62 (d, *J* = 5.5 Hz, 1H), 3.65-4.02 (m, 2H), 3.34-3.40 (m, 2H), 2.78 (s, 1H), 1.94-2.16 (m, 1H), 1.44 (s, 9H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 205.39, 170.42, 156.18, 80.20, 59.03, 44.12, 31.26, 28.30, 27.43; **HRMS (ESI):** calcd for C₁₁H₁₈N₂O₄SNa [M + Na]⁺: 297.0885, found 297.0890.

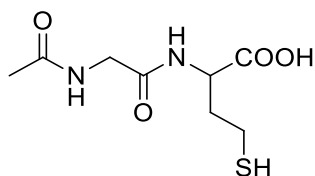
2-(2-((tert-butoxycarbonyl)amino)acetamido)-4-Sulfanylbutanoic acid (**56**)



Thiolactone **55** (382 mg, 1.4 mmol) was dissolved in 4 mL THF, 4.2 mL of NaOH (1 N) was added stepwise, and the mixture was stirred under argon for overnight at rt. The mixture was concentrated and then acidified to pH 6 with NaHSO₄ (saturated solution). The mixture was extracted four times with EA. The organic layer was recovered, dried over NaSO₄ and concentrated under reduced vacuum to afford **56** (83 %)

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 11.51 (br, 1H), 7.38 (s, 1H), 5.81 (s, 1H), 4.72 (s, 1H), 3.70-3.98 (m, 2H), 2.58 (q, *J* = 7.4 Hz, 2H), 2.18 (s, 2H), 1.60 (t, *J* = 8.1 Hz, 1H), 1.45 (s, 9 H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 174.33, 170.84, 156.61, 80.64, 51.26, 44.01, 36.12, 28.31, 20.77.

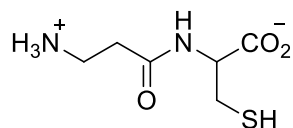
2-(2-acetamidoacetamido)-4-Sulfanylbutanoic acid (72)



In an NMR, N-acetyl- α -aminonitrile (3.2 mg, 0.033 mmol) was dissolved in 0.7 mL D₂O. Cysteine (3.96 mg, 0.033 mmol) was added. The mixture was kept at 45 °C. **72** was characterized by NMR without isolation.

¹H-NMR (500 MHz, D₂O) (δ , ppm): 4.35 (t, J = 5.4 Hz, 1H, CH), 3.89 (s, 2H, CH₂), 2.86 (dd, J = 2.6, 5.7 Hz, 2H), 2.00 (s, 3H); **¹³C-NMR (125 MHz, D₂O) (δ , ppm):** 176.1, 174.9, 170.9, 56.4, 42.6, 26.3, 21.7.

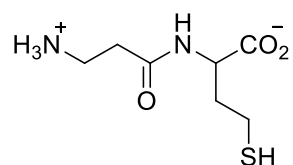
2-(3-ammoniopropanamido)-3-Sulfanylpropanoate (75)



In an NMR tube, the β -aminonitrile **73** (2 mg, 0.029 mmol) was dissolved in 0.7 mL D₂O. DCl was added to achieve pH 5.5. Cysteine (3.47 mg, 0.029 mmol) was added. The solution was left at 45 °C and the product was identified by NMR without purification.

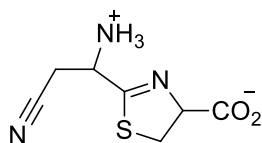
¹H-NMR (500 MHz, D₂O) (δ , ppm): 4.24 (t, J = 5.9 Hz, 1H), 3.13 (s, 4H), 2.77 (m, 2H); **¹³C-NMR (125 MHz, D₂O) (δ , ppm):** 176.5, 171.9, 56.9, 35.6, 26.1.

2-(3-ammoniopropanamido)-4-Sulfanylbutanoate (77)



77 was not isolated, but characterized by NMR. **¹H-NMR (500 MHz, D₂O) (δ , ppm):** 4.15 (dd, J = 9.4, 4.7 Hz, 1H), 3.12 (t, 2H), 2.58 (m, 2H), 2.46-2.38 (m, 2H), 1.89- 1.82 (m, 2H).

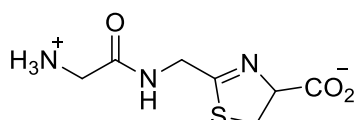
2-(1-ammonio-2-cyanoethyl)-4,5-dihydrothiazole-4-carboxylate (78)



In an NMR tube, the bis-nitrile (**32**) (7.67 mg, 0.08 mmol) was dissolved in 0.7 mL D₂O. DCl was added to achieve the pH 5.5. Then, cysteine (9.68 mg, 0.08 mmol) was added and the solution was left at 45 °C. **78** was characterized by NMR without purification.

¹H-NMR (400 MHz, D₂O) (δ, ppm): 4.95 (t, *J* = 8.7 Hz, 1H), 4.11 (t, 1H), 3.61-3.40 (2 dd, AB of ABX, *J*_{AX} = 9.7 Hz, *J*_{BX} = 8.0 Hz, *J*_{AB} = 11.2 Hz, 2H), 2.88 (m, 2H); **¹³C-NMR (100 MHz, D₂O) (δ, ppm):** 178.2, 177.3, 118.6, 79.6, 50.3, 36.3, 24.1.

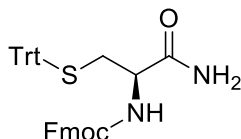
2-((2-ammonioacetamido)methyl)-4,5-dihydrothiazole-4-carboxylate (79)



In an NMR tube, GlyGly-CN (6 mg, 0.053 mmol) and cysteine (6.4 mg, 0.053 mmol) were mixed in 0.7 mL D₂O. NaHCO₃ was used to afford the pH 8. The product was identified by NMR without purification.

¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.91 (dd, *J* = 9.3, 8.6 Hz, 1H), 4.26-4.16 (m, 2H), 3.61 (dd, *J* = 11.2, 9.8 Hz, 1H), 3.50 (s, 2H), 3.38 (dd, *J* = 11.2, 8.1 Hz, 1H).

(9H-fluoren-9-yl)methyl (1-amino-1-oxo-3-(tritylthio)propan-2-yl)carbamate (87)



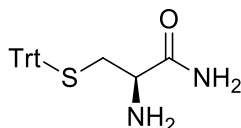
1st method: Fmoc-L-Cys(Trt)-OH (1.21 g, 2.07 mmol) was dissolved in 25 mL DCM, SOCl₂ (0.4mL, 22.1 mmol) was added at 0 °C. After stirring for 1 h at rt, the solution was evaporated and the obtained solid was dissolved in 20 mL DCM. NH₃ (gas) was bubbled in

the solution for 2 min. the mixture is then extracted three times in DCM. The organic layer was collected, dried and evaporated to give **87** (quantitative).

2nd method: Fmoc-L-Cys(Trt)-OH (2 g, 3.42 mmol) was dissolved in 50 mL DCM at 0 °C under argon. NMM (0.48 mL, 3.76 mmol) and IBCF (0.53 mL, 4.1 mmol) were added respectively and the reaction left to stir for 1 hour at 0 °C. 10 mL of ammonia solution (26 %) were added and the reaction was stirred for 15 min at 0 °C. Solvent was gently evaporated and then water was added and the product was extracted with DCM. Solvent was dried using Na₂SO₄ and evaporated under vacuum to obtain the desired product as a white solid (quantitative).

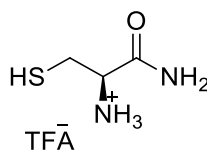
m.p. 98-104 °C; [α]_D²⁰ +8.09 (*c* 1.02, CHCl₃); **¹H-NMR (500 MHz, CDCl₃) (δ, ppm):** 7.72 (t, *J* = 6.7 Hz, 2H), 7.54 (d, 2H), 7.38 (m, 8H), 7.23 (m, 8H), 7.17 (t, 3H), 5.8 (s, 1H), 5.7 (s, 1H), 5.2 (s, 1H), 4.39 (d, *J* = 6.6 Hz, 2H), 4.16 (t, *J* = 6.6 Hz, 2H), 3.83 (m, 1H), 2.63 (m, 2H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 172.5, 156.0, 144.3, 143.6, 141.30, 129.5, 128.1, 127.7, 127.1, 127.0, 125.0, 120.0, 67.3, 66.8, 53.6, 47.1, 33.6; **HRMS (ESI):** calcd for C₃₇H₃₂N₂O₃SNa [M + Na]⁺: 607.2026, found 607.2005.

2-amino-3-(tritylthio)propanamide (**88**)



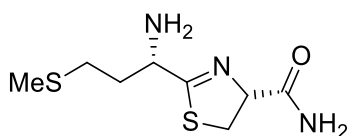
87 (895 mg, 1.53 mmol) was dissolved in 15 mL DCM, 3 mL of piperidine was added. After 15 min at rt, the mixture was evaporated and purified by silica gel chromatography with 3 % MeOH/DCM to give the product as a white solid (63 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.44 (d, *J* = 7.21 Hz, 6H), 7.29 (m, 6H), 7.21 (m, 3H), 6.81 (s, 1H), 5.21 ppm (s, 1H), 3.00 (q, *J* = 4.18 Hz, 1H), 2.72 (dd, *J* = 12.55, 3.91 Hz, 1H), 2.59 (dd, *J* = 12.82 - 8.36 Hz, 1H), 1.30 (s, 2H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 175.74, 144.59, 129.62, 128.00, 126.84, 67.08, 53.98, 37.27.

1-amino-3-Sulfanyl-1-oxopropan-2-aminium trifluoroacetate (86)

88 (1.41 g, 3.98 mmol) was dissolved in 5 mL DCM, $i\text{Pr}_3\text{SiH}$ (3.2 mL, 15.7 mmol) was added to the solution, then, TFA (3 mL, 38.6 mmol) was added. The mixture was stirred for 1 h at rt, then evaporated. The crude product was washed several times with DCM and finally evaporated to give a white solid (quantitative).

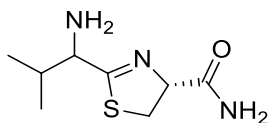
$^1\text{H-NMR}$ (500 MHz, D_2O) (δ , ppm): 4.17 (t, $J = 5.6$ Hz, 1H), 3.02 (t, $J = 6.0$ Hz, 2H); $^{13}\text{C-NMR}$ (125 MHz, D_2O) (δ , ppm): 170.0, 54.2, 24.8

(R)-2-((S)-1-amino-3-(methylthio)propyl)-4,5-dihydrothiazole-4-carboxamide (90)

Met-CN (168 mg, 1.29 mmol) was dissolved in 15 mL H_2O . Cys- NH_2 .TFA (280 mg; 1.29 mmol) was added. The pH of the solution was adjusted to 8 by adding Na_2CO_3 . The solution was stirred at 45 °C during 2.5 h. The aqueous phase was extracted three times with EA. The organic layer was dried over Na_2SO_4 , filtered, and concentrated under vacuum. After purification by silica gel chromatography (1-10 % MeOH/DCM), the thiazoline **90** was obtained as an orange oil (16 % yield).

$^1\text{H-NMR}$ (400 MHz, D_2O) (δ , ppm): 5.08 (t, $J = 8.98$ Hz, 1H), 3.95 (t, $J = 6.57$ Hz, 1H), 3.65 (t, $J = 10.82$ Hz, 1H), 3.46 (dd, $J = 11.30$; 8.20, 1H), 2.56 (t, $J = 7.08$ Hz, 2H), 2.07 (s, 3H), 1.97 (sep, $J = 7.0$, 2H); $^{13}\text{C-NMR}$ (100 MHz, D_2O) (δ , ppm): 182.99, 176.14, 77.00, 52.93, 34.89, 34.43, 29.18, 14.12; HRMS (ESI): calcd for $\text{C}_8\text{H}_{17}\text{ON}_3\text{S}_2$ $[\text{M} + \text{H}]^+$: 234.0735, found 234.0740.

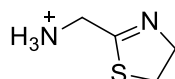
(4R)-2-(1-amino-2-methylpropyl)-4,5-dihydrothiazole-4-carboxamide (92)



Val-CN.HCl (35 mg, 0.26 mmol) was dissolved in 5 mL H₂O. Cys-NH₂.TFA (57 mg; 0.26 mmol) was added. The solution was adjusted to pH = 7 by adding Na₂CO₃. The solution was stirred at 45 °C during 24 h. The aqueous phase was extracted three times with EA. The organic layer was dried over Na₂SO₄, filtered, and concentrated under vacuum. After purification by silica gel chromatography (1-10 % MeOH/DCM), the thiazoline **92** was obtained as a yellow oil (30 % yield).

¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.40 (m, 1H), 3.75 (dd, *J* = 5.68, 2.59 Hz, 1H), 2.70–2.99 (m, 2H), 2.12 (sep, *J* = 6.60, 1H), 0.85–0.94 (m, CH₃, 6H); **¹³C-NMR** (D₂O, 125 MHz, both isomers were observed) (δ, ppm): 174.02–173.74, 169.70–169.43, 58.66–58.38, 55.65–55.63, 30.01–29.94, 25.30–25.02, 17.79–17.62, 16.85–16.65; **HRMS** (ESI): calcd for C₈H₁₇ON₃S [M + H]⁺: 202.1014, found 202.1016.

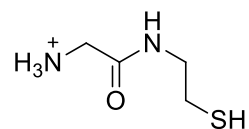
(4,5-dihydrothiazol-2-yl)methanaminium (95)



In an NMR tube, cysteamine **94** (3.1 mg, 0.027 mmol) was dissolved in 0.7 mL D₂O. Then, aminoacetonitrile **39** (2.5 mg, 0.027 mmol) was added. The mixture was left at 45 °C and followed by NMR.

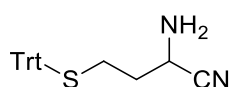
¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.16 (t, *J* = 8.5 Hz, 2H), 3.91 (m, *J* = 8.5 Hz, 2H), 3.39 (t, 2H); **¹³C-NMR** (125 MHz, D₂O) (δ, ppm): 176.2, 63.3, 33.8, 21.3.

2-((2-Sulfanylethyl)amino)-2-oxoethanaminium (96)



96 was characterized by NMR without purification. ¹H-NMR (500 MHz, D₂O) (δ, ppm): 3.76 (s, 2H), 3.39 (t, *J* = 8.1 Hz, 2H), 2.63 (t, *J* = 8.1 Hz, 2H).

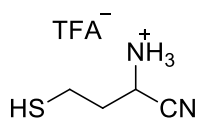
2-amino-4-(tritylthio)butanenitrile (119)



Aldehyde **117** (145 mg, 0.43 mmol) was dissolved in 5 mL THF. TMSCN (55.5 mg, 0.56 mmol) and then ZnI₂ (25.5 mg, 0.08 mmol) were added. After stirring for 15 min at rt, 1 mL of ammonia/MeOH solution was added to the mixture and left stirring for 2 h. Then, it was extracted three times with EA. The organic layer was dried over Na₂SO₄ and evaporated under reduced pressure. The crude product was purified silica gel chromatography (95-70 % pentane/EA) to give **119** (87 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.34 (d, *J* = 3.1, 1.9 Hz, 6H), 7.21 (t, *J* = 10.4, 4.9 Hz, 6H), 7.16 – 7.07 (m, 3H), 3.55 (t, *J* = 7.2 Hz, 1H), 2.31 (t, *J* = 7.2, 1.5 Hz, 2H), 1.69 – 1.40 (m, 2H), 1.33 (s, 2H); ¹³C-NMR (125 MHz, CDCl₃) (δ, ppm): 144.55, 129.58, 128.05, 126.86, 121.59, 67.07, 42.23, 34.52, 27.69.

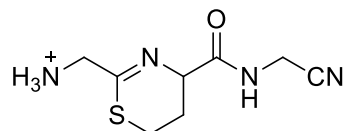
1-cyano-3-Sulfanylpropan-1-aminium trifluoroacetate (120)



119 (30 mg, 0.083 mmol) was dissolved in DCM (2 mL), *i*Pr₃SiH (69 μL, 0.33 mmol) and TFA (64 μL, 0.83 mmol) were added respectively. The mixture was stirred for 30 min at rt and then it was extracted with D₂O (1 mL). Finally, the aqueous layer was left under vacuum at room temperature to remove the remaining DCM to provide the required products **120** as a TFA salt (quantitative).

¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.56 (dd, *J* = 9.2, 5.8 Hz, 1H), 2.71-2.60 (m, 1H), 2.58-2.46 (m, 1H), 2.25-2.14 (m, 1H), 2.13-1.96 (m, 1H); **¹³C-NMR (126 MHz, D₂O) (δ, ppm):** 115.29, 40.18, 33.85, 19.07.

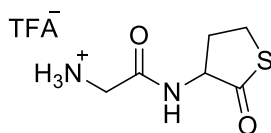
(4-((cyanomethyl)carbamoyl)-5,6-dihydro-4H-1,3-thiazin-2-yl)methanaminium (125)



Homocysteine thiolactone (10 mg, 0.065 mmol) and aminoacetonitrile hydrochloride (12 mg, 0.13 mmol) were dissolved in D₂O (0.7 mL) in a NMR tube under inert atmosphere, pH = 6. The product was identified in the reaction mixture.

¹H-NMR (500 MHz, D₂O) (δ, ppm): 1.56-1.62 (m, 1 H), 2.4-2.47 (m, 1 H), 3.03-3.09 (m, 1 H), 3.23-3.3 (m, 1 H), 4.01-4.11 (m, 1H), 3.89 (d of AB system, *J* = 16 Hz, 1 H), 3.82 (d of AB system, *J* = 16 Hz, 1 H), 4.25 (s, 2H); **¹³C-NMR (125 MHz, D₂O) (δ, ppm):** 21.69 (CH₂), 25.17, 27.54, 44.35, 59.05, 116.99, 159.09, 175.16; **HRMS (ESI):** calcd for C₈H₁₃N₄OS [M + H]⁺: 213.0810, found 213.0806.

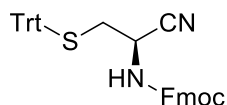
2-oxo-2-((2-oxotetrahydrothiophen-3-yl)amino)ethanaminium trifluoroacetate (133)



55 (80 mg, 0.29 mmol) was dissolved in DCM (2 mL). TIPS (0.5 mL) and TFA (0.5 mL) were added to this mixture. The mixture was left stirring for 30 min at rt. The solution was evaporated and the crude product was washed 3 times with DCM. The product was then evaporated under reduced pressure to afford **133** as TFA salt (quantitative).

¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.67, (q, *J* = 6.7 Hz, 1H), 3.79 (d, *J* = 3.0 Hz, 2H), 3.27-3.42 (m, 2H), 2.55-2.6 (m, 1H), 2.11-2.20 (m, 1H); **¹³C-NMR (125 MHz, D₂O):** 210.31, 167.22, 59.33, 40.38, 29.79, 27.47.

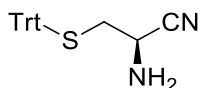
(R)-(9H-fluoren-9-yl)methyl (1-cyano-2-(tritylthio)ethyl)carbamate (155)



Fmoc-L-Cys(Trt)-NH₂ (2.13 g, 3.65 mmol) was dissolved in 6 mL DMF. Cyanuric chloride (0.67 g, 3.65 mmol) was added. The mixture was stirred for 30 min at rt. Then it was quenched with water, and organic layer was extracted with DCM (4 x 50 mL). Organic phase was washed with water (4 x 50 mL), dried, evaporated under vacuum to yield a white solid (88 %).

$[\alpha]^{20}_{\text{D}}$ -8.934 (*c* 1.00, CHCl₃); **¹H-NMR (500 MHz, CDCl₃) (δ, ppm):** 7.76 (t, *J* = 7.5 Hz, 2H), 7.54 (d, *J* = 7.5 Hz, 2H), 7.42 (m, 8H), 7.24 (m, 11H), 4.77 (d, 1H), 4.44 (d, *J* = 7.5 Hz, 2H), 4.18 (m, 1H), 4.02 (m, 1H), 2.71-2.64 (m, 2H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 154.7, 143.8, 143.4, 141.3, 129.4, 128.2, 127.9, 127.2, 127.1, 124.9, 120.1, 117.2, 67.9, 67.3, 47.0, 41.9, 34.3; **HRMS (ESI):** calcd for C₃₇H₃₀N₂O₂SNa [M + Na]⁺: 589.1920, found 589.1902.

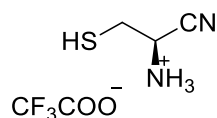
(R)-2-amino-3-(tritylthio)propanenitrile (156)



Fmoc-L-Cys(Trt)-CN (1.5 g, 2.65 mmol) was dissolved in 8 mL DCM. Piperidine (4 mL) was added. The mixture was stirred for 15 min at rt and finally the solution was evaporated to obtain a yellow solid which was then purified by column (90-40 % pentane/EA) to obtain the product as a white solid (55 %).

$[\alpha]^{20}_{\text{D}}$ +24.5 (*c* 1.00, CHCl₃); **¹H-NMR (500 MHz, CDCl₃) (δ, ppm):** 7.31-7.22 (m, 15H), 2.96 (t, *J* = 6.6 Hz, 1H), 2.60 (dd, *J* = 6.5, 2.5 Hz, 2H), 1.56 (s, 2H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 144.1, 129.5, 128.2, 127.1, 120.6, 67.5, 42.8, 37.2.

(R)-1-cyano-2-Sulfanylethanaminium trifluoroacetate (144)



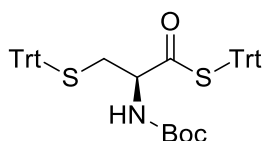
L-Cys(Trt)-CN (18 mg, 0.052 mmol) was dissolved in 1.5 mL DCM. TIPS (43 μ L, 0.21 mmol) and TFA (40 μ L, 0.52 mmol) were added respectively and the solution was stirred for 1 h at rt. The product was extracted with 1.5 mL D₂O.

¹H-NMR (500 MHz, D₂O) (δ , ppm): 4.59 (dd, X of ABX, $J_{AX} = 5.6$ Hz, $J_{BX} = 6.7$ Hz, 1H), 3.02 (2 dd, AB of ABX, $J_{AB} = 14.8$ Hz, $J_{AX} = 5.6$ Hz, $J_{BX} = 6.7$ Hz, 2H); **¹³C-NMR (125 MHz, D₂O) (δ , ppm):** 114.6, 43.8, 24.1.

General procedure for the synthesis of 157, 158, 160 and 161 esters:

L-Boc-Cys(Trt)-OH (1 equiv.) was dissolved in DCM (15 mL), the solution was cooled to 0 °C, then DMAP (0.1 equiv.) and thiol (Triphenylmethanethiol, ethanethiol, propanethiol and benzylthiol) (1.2 equiv.) were added respectively. To this mixture, a solution of DCC (1.1 equiv.) in DCM (4 mL) was added dropwise and a white precipitate started to form. The temperature was raised from 0 °C to r.t. and the mixture was stirred for 3 hours. The mixture was filtered and the filtrate was evaporated. Flash chromatography afforded the pure esters **157, 158, 160 and 161**.

(S)-S-trityl 2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanethioate (157)



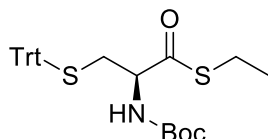
L-Boc-Cys(Trt)-OH (1 g, 2.16 mmol), DMAP (26 mg, 0.22 mmol), Triphenylmethanethiol (715 mg, 2.6 mmol) and DCC (490 mg, 2.37 mmol) were used. The crude product was purified by silica column chromatography (1 to 10 % EA/pentane) to obtain the desired product as a white solid (85 %).

¹H-NMR (500 MHz, CDCl₃) (δ , ppm): 7.36 – 7.04 (m, 30H), 4.83 (d, $J = 8.3$ Hz, 1H), 4.27 – 3.95 (m, 1H), 2.45 (dd, $J = 12.3, 6.9$ Hz, 1H), 2.27 (dd, $J = 12.4, 4.5$ Hz, 1H), 1.35 (s, 9H);

Experimental part

$^{13}\text{C-NMR}$ (125 MHz, CDCl_3) (δ , ppm): 196.88, 154.84, 144.25, 143.45, 129.91, 129.52, 128.04, 127.71, 127.07, 126.90, 115.30, 80.26, 70.43, 67.11, 59.26, 34.27, 28.30; **HRMS** (ESI): calcd for $\text{C}_{46}\text{H}_{43}\text{O}_3\text{NS}_2\text{Na}$ $[\text{M} + \text{Na}]^+$: 744.2577, found 744.2579.

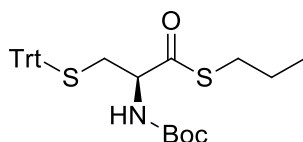
(S)-S-ethyl 2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanethioate (**158**)¹⁸³



Following the general procedure, L-Boc-Cys(Trt)-OH (1 g, 2.16 mmol), DMAP (270 mg, 0.216 mmol), ethanethiol (170 mg, 2.6 mmol) and DCC (570 mg, 2.37 mmol) were used. The crude product was purified by silica column chromatography (1-10 % EA/pentane) to obtain **158** as a white solid (98 %).

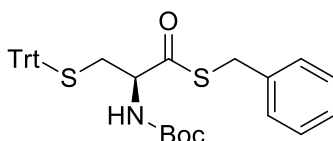
m.p. 45-50 °C; $[\alpha]_D^{20}$ -1.63 (*c* 1.00, CHCl_3); $^1\text{H-NMR}$ (500 MHz, CDCl_3) (δ , ppm): 7.4 (m, 15H), 5.1 (dd, 1H), 4.2 (s, 1H), 2.8 (q, 2H), 2.5 (dd, 2H), 1.5 (s, 9H), 1.2 (t, 3H); $^{13}\text{C-NMR}$ (125 MHz, CDCl_3) (δ , ppm): 144.28, 129.56, 128.07, 126.93, 200.19, 154.89, 59.16, 34.20, 28.36, 23.46, 14.31; **HRMS** (ESI): calcd for $\text{C}_{29}\text{H}_{33}\text{O}_3\text{NS}_2\text{Na}$ $[\text{M} + \text{Na}]^+$: 530.1794, found 530.1788.

(S)-S-propyl 2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanethioate (**160**)



Following the general procedure, L-Boc-Cys(Trt)-OH (1 g, 2.16 mmol), DMAP (270 mg, 0.216 mmol), propanethiol (197 mg, 2.6 mmol) and DCC (490 mg, 2.37 mmol) were used. The crude product was purified by silica column chromatography (1-10 % EA/pentane) to obtain **160** as a white solid (90 %).

m.p. 107.7-109.5 °C; $^1\text{H-NMR}$ (400 MHz, CDCl_3) (δ , ppm): 7.40 (d, $J = 7.6$ Hz, 6H), 7.29 (t, $J = 7.6$ Hz, 6H), 7.22 (t, $J = 7.2$ Hz, 3H), 4.92 (d, $J = 8.3$ Hz, 1H), 4.23 (dd, $J = 11.9, 7.4$ Hz, 1H), 4.60 (br, 1H), 2.81 (t, $J = 7.3$, 2H), 2.72 (dd, $J = 12.5, 7.0$ Hz, 1H), 2.52 (dd, $J = 12.5, 4.1$ Hz, 1H), 1.57 (sex, $J = 7.3$ Hz, 2H), 1.44 (s, 9H), 0.94 (t, $J = 7.4$ Hz, 3H).

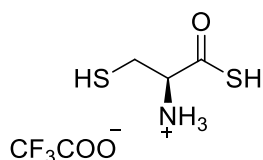
(S)-S-benzyl 2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanethioate (161)¹⁸³

Following the general procedure, L-Boc-Cys(Trt)-OH (300 mg, 0.65 mmol), DMAP (8 mg, 0.065 mmol), benzylthiol (97 mg, 0.78 mmol) and DCC (150 mg, 0.71 mmol) were used. The crude product was purified by silica column chromatography (1 to 10 % EA/Pent) to obtain the desired product as white solid (99 % yield).

¹H-NMR (400 MHz, CDCl₃) (δ, ppm): 7.60 – 7.09 (m, 20H), 5.21 (d, *J* = 8.5 Hz, 1H), 4.36 (d, *J* = 4.2 Hz, 1H), 4.18 – 3.93 (m, 2H), 2.86 (dd, *J* = 12.4, 7.3 Hz, 1H), 2.66 (dd, *J* = 12.4, 4.2 Hz, 1H), 1.52 (s, 9H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 199.65, 154.87, 144.26, 136.93, 126.93, 129.56, 80.40, 59.14, 34.22, 33.48; **HRMS (ESI):** calcd for C₃₄H₃₅O₃NS₂Na [M + Na]⁺: 592.1951, found 592.1945.

General method for the preparation of thioacid 145 and thioesters 146, 147 and 148:

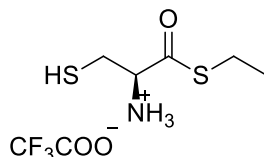
157, 158, 160 or **161** (1 equiv.) was dissolved in DCM (2 mL), *i*Pr₃SiH (8 equiv.) and TFA (20 equiv.) were added respectively. The mixture was stirred for 1 h and then it was extracted with D₂O (1.5 mL). Finally, the aqueous layer was left under vacuum at room temperature to remove the remaining DCM to provide the required products **145, 146, 147** and **148** as TFA salts.

(S)-2-Sulfanyl-1-thiocarboxyethanaminium trifluoroacetate (145)

Following the general procedure, **157** (174 mg, 0.24 mmol), *i*Pr₃SiH (502 μL, 1.92 mmol) and TFA (471 μL, 4.8 mmol) were used to give **145** as TFA salt.

¹H-NMR (500 MHz, D₂O) (δ, ppm): 3.76 (dd, *J* = 5.7, 4.5 Hz, 1H), 2.62 (dd, *J* = 15.1, 5.8 Hz, 1H), 2.54 (dd, *J* = 15.1, 4.4 Hz, 1H); **¹³C-NMR (125 MHz, D₂O) (δ, ppm):** 207.93, 61.50, 25.31.

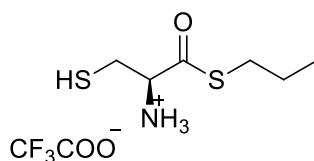
(S)-1-(ethylthio)-3-Sulfanyl-1-oxopropan-2-aminium trifluoroacetate (146)



Following the general procedure, **158** (145 mg, 0.286 mmol), *i*Pr₃SiH (470 μL, 2.29 mmol) and TFA (440 μL, 5.72 mmol) were used to give **146** as TFA salt.

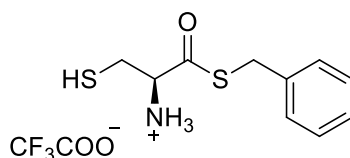
¹H-NMR (400 MHz, D₂O) (δ, ppm): 4.27 (t, 1H), 2.83-2.98 (m, 2H), 2.72-2.83 (m, 2H), 1.01 (t, 3H); **¹³C-NMR (100 MHz, D₂O) (δ, ppm):** 196.51, 59.95, 24.93, 23.66, 13.33.

(S)-3-Sulfanyl-1-oxo-1-(propylthio)propan-2-aminium trifluoroacetate (147)



Following the general procedure, **160** (30 mg, 0.058 mmol), *i*Pr₃SiH (94 μL, 0.46 mmol) and TFA (88 μL, 1.15 mmol) were used to give **147** as TFA salt.

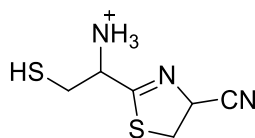
(S)-1-(benzylthio)-3-Sulfanyl-1-oxopropan-2-aminium trifluoroacetate (148)



Following the general procedure, **161** (30 mg, 0.053 mmol), *i*Pr₃SiH (89 μL, 0.43 mmol) and TFA (82 μL, 1.06 mmol) were used to give **148** as TFA salt.

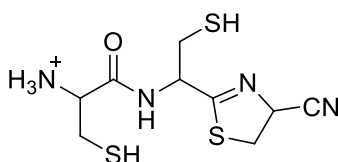
¹H-NMR (400 MHz, D₂O) (δ, ppm): 6.90 (m, 5H), 4.07 (t, 1H), 3.48 (m, 2H), 2.72-2.83 (m, 2H), 2.69 (m, 2H).

1-(4-cyano-4,5-dihydrothiazol-2-yl)-2-Sulfanylethanaminium (162)



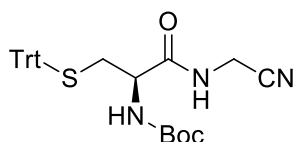
162 was characterized in the mixture without isolation. $^1\text{H-NMR}$ (500 MHz, D_2O) (δ , ppm): 5.47 (t, $J = 9.2$ Hz, 1H), 4.56 (t, $J = 5.1$ Hz, 1H), 3.82 (m, 2H), 3.10 (m, 2H); $^{13}\text{C-NMR}$ (125 MHz, D_2O) (δ , ppm): 162.9, 117.5, 63.4, 53.8, 37.5, 25.6.

1-((1-(4-cyano-4,5-dihydrothiazol-2-yl)-2-Sulfanylethyl)amino)-3-Sulfanyl-1-oxopropan-2-aminium (163)



163 was characterized in the mixture without isolation. $^1\text{H-NMR}$ (500 MHz, D_2O) (δ , ppm): 5.54 (t, $J = 7.0$ Hz, 1H), 4.65 (t, $J = 6.4$ Hz, 1H), 4.58 (t, 1H), 3.68 (m, 2H), 3.01 (m, 4H); $^{13}\text{C-NMR}$ (125 MHz, D_2O) (δ , ppm): 174.1, 163.1, 115.2, 63.4, 53.8, 44.1, 37.5, 25.6, 24.5.

(R)-tert-butyl (1-((cyanomethyl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (172)

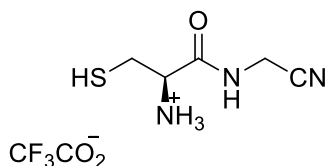


Boc-Cys(Trt)-OH (500 mg, 1.08 mmol) was dissolved in dry THF (40 mL) at -20 °C under argon. NMM (261 μL , 2.37 mmol) and IBCF (168 μL , 1.29 mmol) were added respectively and the reaction left to stir for 20 min. Then, aminoacetonitrile (110 mg, 1.18 mmol) was added and the mixture was left stirring under argon for overnight at rt. The mixture was concentrated, quenched with water and extracted three times with EA. The organic phase was dried over Na_2SO_4 and evaporated. The crude product was submitted to silica column chromatography (95-40 % pentane/EA). The solvent was collected and concentrated under vacuum to give **172** as a white solid (67 %).

Experimental part

m.p. 81.8-82.5 °C; **¹H-NMR (500 MHz, CDCl₃) (δ, ppm):** 7.48-7.39 (m, 6H), 7.31 (t, *J* = 10.3, 4.9 Hz, 6H), 7.27-7.17 (m, 3H), 6.61 (s, 1H), 4.65 (s, 1H), 4.28-3.94 (m, 2H), 3.76 (d, *J* = 5.9 Hz, 1H), 2.75 (dd, *J* = 13.0, 7.2 Hz, 1H), 2.58 (dd, *J* = 13.3, 5.3 Hz, 1H), 1.42 (s, 9H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 171.27, 155.67, 144.35, 129.57, 128.15, 126.99, 115.98, 80.62, 67.24, 60.49, 33.64, 28.36, 27.42.

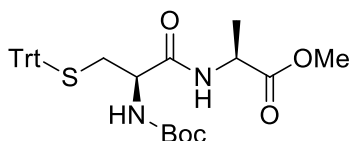
(R)-1-((cyanomethyl)amino)-3-Sulfanyl-1-oxopropan-2-aminium trifluoroacetate (171)



172 (120 mg, 0.24 mmol) was dissolved in DCM (5 mL), *i*Pr₃SiH (393 μL, 1.92 mmol) and TFA (370 μL, 4.78 mmol) were added respectively. The mixture was stirred for 1 h and then it was extracted with H₂O (1.5 mL). Finally, the aqueous layer was left under vacuum at room temperature to remove the remaining DCM to provide the required products **171** as a TFA salt (quantitative). The solution was then neutralized with NaHCO₃ and the resulted polymers were identified by MALDI.

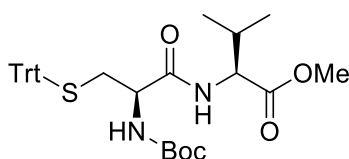
General method for the preparation of methyl esters Boc-Cys(Trt)-aa-OMe:

Boc-L-Cys(Trt)-OH (1 equiv.) was dissolved in dry THF (~ 40 mL) at -20 °C under argon. NMM (3 equiv.) and IBCF (1.5 equiv.) were added respectively and the reaction left to stir for 1 h. Then, ester **173**, **174** or **175** (1.1 equiv.) was added to the mixture and was left stirring under argon for overnight at rt. The mixture was concentrated, quenched with water and extracted three times with DCM. The organic phase was dried over Na₂SO₄ and evaporated under reduced pressure. Flash chromatography afforded the pure esters **176**, **177** and **178**.

(S)-methyl 2-((R)-2-((tert-butoxycarbonyl)amino)-3-tritylthio)propanamido)propanoate (176)

Following the general procedure, Boc-L-Cys(Trt)-OH (1.5 g, 3.24 mmol) was dissolved in THF. NMM (1.07 mL, 9.72 mmol) and IBCF (0.63 mL, 4.6 mmol) were added respectively to give the mixed anhydride. Ala-OMe (0.452 g, 3.24 mmol) was added. The crude product was submitted to silica column chromatography (95-80 % pentane/EA). The solvent was collected and concentrated under vacuum to give **176** as a white solid (80 %).

m.p. 195.4-196.5 °C; **¹H-NMR (400 MHz, CDCl₃) (δ, ppm):** 7.35 (d, *J* = 7.5 Hz, 6H), 7.23 (t, *J* = 7.5 Hz, 6H), 7.16 (dd, *J* = 13.2, 7.2 Hz, 3H), 6.44 (br, 1H), 4.69 (br, 1H), 4.42 (quin, *J* = 7.2 Hz, 1H), 3.75 (br, 1H), 3.63 (s, 3H), 2.60-2.73 (b, 1 H), 2.45 (dd, *J* = 12.9, 4.9 Hz, 1H), 1.35 (s, 9 H), 1.28 (d, *J* = 7.1 Hz, 3H); **¹³C-NMR (100 MHz, CDCl₃) (δ, ppm):** 172.84, 169.93, 155.32, 144.41, 129.60, 128.09, 126.92, 80.27, 67.22, 53.49, 52.37, 48.13, 33.65, 28.22, 18.94.

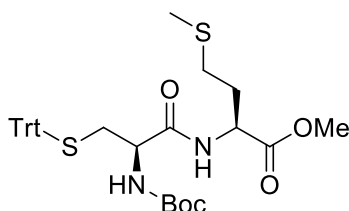
(S)-methyl 2-((R)-2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanamido)-3-methylbutanoate (177)

Following the general procedure, Boc-L-Cys(Trt)-OH (719 mg, 1.55 mmol) was dissolved. NMM (342 μL, 3.1 mmol) and IBCF (241 μL, 1.86 mmol) were added respectively to give the mixed anhydride. Val-OMe (260 mg, 1.55 mmol) was added. The crude product was submitted to silica column chromatography (95-60 % pentane/EA). The solvent was collected and concentrated under vacuum to give **177** as a white solid (96 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.43 (d, *J* = 7.4 Hz, 6H), 7.29 (t, *J* = 7.6 Hz, 6H), 7.22 (t, *J* = 7.3 Hz, 3H), 6.58 (s, 1H), 4.74 (s, 1H), 4.45 (dd, *J* = 8.8, 5.0 Hz, 1H), 3.80 (s, 1H), 3.69 (s, 3H), 2.76 (d, *J* = 4.8 Hz, 1H), 2.52 (dd, *J* = 13.0, 5.1 Hz, 1H), 2.19-2.06 (m,

1H), 1.42 (s, 9H), 0.88 (dd, $J = 12.1, 6.9$ Hz, 6H); $^{13}\text{C-NMR}$ (125 MHz, CDCl_3) (δ , ppm): 171.82, 170.40, 144.42, 129.57, 128.08, 126.88, 67.21, 57.12, 52.05, 31.35, 28.25, 18.89, 17.66; **HRMS** (ESI): calcd for $\text{C}_{33}\text{H}_{41}\text{N}_2\text{O}_5\text{S}$ [$\text{M} + \text{H}$] $^+$: 577.2731, found 577.2722.

(S)-methyl 2-((R)-2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanamido)-4-(methylthio)butanoate (178)



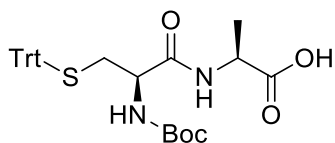
Following the general procedure, Boc-L-Cys(Trt)-OH (3 g, 6.48 mmol) was dissolved. NMM (2.14 mL, 19.44 mmol) and IBCF (1.26 mL, 9.72 mmol) were added respectively to give the mixed anhydride. Met-OMe (1.294 g, 6.48 mmol) was added. The crude product was submitted to silica column chromatography (90-80 % pentane/EA). The solvent was collected and concentrated under vacuum to give **178** as a white solid (59 %).

m.p. 121.9-123.2 °C; $[\alpha]^{20}_{\text{D}} +25.7$ (c 1.03, CHCl_3); $^1\text{H-NMR}$ (400 MHz, CDCl_3) (δ , ppm): 7.41 (d, $J = 7.7$ Hz, 6H), 7.25 (t, $J = 7.6$ Hz, 6H), 7.17(t, $J = 7.2$ Hz, 3H), 6.85 (d, $J = 7.0$ Hz, 1H), 5.07 (d, $J = 6.0$ Hz, 1H), 4.63 (dd, $J = 12.6, 7.4$ Hz, 1H), 3.86 (s, 1H), 3.64 (s, 3H), 2.78 (dd, $J = 12.6, 7.2$ Hz, 1H), 2.53 (dd, $J = 12.6, 4.9$ Hz, 1H), 2.46 (t, $J = 7.4$ Hz, 2H), 2.09-2.18 (m, 1H), 1.99 (s, 3H), 1.88-1.97 (m, 1H), 1.41 (s, 9H); $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) (δ , ppm): 171.81, 170.44, 155.29, 144.45, 129.57, 128.09, 126.91, 80.09, 67.14, 52.39, 51.51, 31.63, 29.74, 28.32, 15.34; **HRMS** (ESI): calcd for $\text{C}_{33}\text{H}_{40}\text{N}_2\text{O}_5\text{S}_2\text{Na}$ [$\text{M} + \text{Na}$] $^+$: 631.2276, found 631.2278.

General method for the preparation of acids Boc-Cys(Trt)-aa-OH:

An ester **176**, **177** or **178** (1 equiv.) was dissolved in THF/water solution (2/1). LiOH (2.5 equiv.) was added to the mixture at 0 °C. The mixture was left to stir for 2 h at rt. The solution was concentrated and water was added. Then, NaHSO_4 (sat.) was added to afford pH 2-3. The mixture was extracted three times with DCM. The organic layer was collected, dried over Na_2SO_4 and concentrated under reduced pressure. Flash chromatography afforded the pure acids **179**, **180** and **181**.

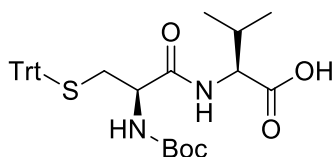
(S)-2-((R)-2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanamido)propanoic acid (179)



Following the general procedure, To **176** (930 mg, 1.69 mmol) solution, LiOH (178 mg, 4.24 mmol) was added. The crude product was submitted to silica column chromatography (90-30 % pentane/EA). The solvent was collected and concentrated under vacuum to give **179** as a white solid (85 %).

m.p. 80.7-81.2 °C; $[\alpha]^{20}_{\text{D}}$ +10.38 (*c* 1.05, CHCl₃): **¹H-NMR (400 MHz, CDCl₃) (δ, ppm):** 7.35 (d, *J* = 7.7 Hz, 6H), 7.22 (t, *J* = 7.7 Hz, 6H), 7.16 (t, *J* = 7.3, 3H), 6.59 (d, *J* = 5.5 Hz, 1H), 4.75 (br, 1H), 4.41 (quin, *J* = 6.8 Hz, 1H), 3.76 (br, 1H), 2.66 (dd, *J* = 12.8, 7.2 Hz, 1H), 2.47 (dd, *J* = 12.8, 4.7 Hz, 1H), 1.34 (s, 9 H) 1.32 (d, *J* = 7.5 Hz, 3H); **¹³C-NMR (100 MHz, CDCl₃) (δ, ppm):** 174.01, 170.50, 155.85, 144.19, 129.49, 128.19, 127.07, 80.91, 67.47, 54.00, 48.57, 33.28, 28.34, 17.54.

(S)-2-((R)-2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanamido)-3-methylbutanoic acid (180)

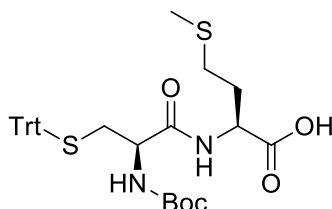


Following the general procedure, To **177** (200 mg, 0.69 mmol) solution, LiOH (73 mg, 1.73 mmol) was added. The crude product was submitted to silica column chromatography (90-10 % pentane/EA). The solvent was collected and concentrated under vacuum to give **180** as a white solid (55 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.43 (d, *J* = 7.5 Hz, 6H), 7.29 (t, *J* = 7.6 Hz, 6H), 7.24-7.18 (m, 3H), 6.67 (d, *J* = 6.9 Hz, 1H), 4.80 (s, 1H), 4.46 (dd, *J* = 8.2, 4.7 Hz, 1H), 3.81 (d, *J* = 5.5 Hz, 1H), 2.84-2.62 (m, 1H), 2.54 (dd, *J* = 13.1, 5.1 Hz, 1H), 2.29-2.12 (m, 1H), 1.41 (s, 9H), 0.91 (dd, *J* = 12.7, 6.9 Hz, 6H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 174.85, 170.95, 155.59, 129.57, 128.10, 126.90, 80.47, 67.24, 57.08, 53.62, 33.09, 31.22,

28.26, 18.93, 17.50; **HRMS** (ESI): calcd for $C_{32}H_{38}N_2O_5SNa$ $[M + Na]^+$: 585.2394, found 585.2379.

(S)-2-((R)-2-((tert-butoxycarbonyl)amino)-3-(tritylthio)propanamido)-4-(methylthio)butanoic acid (181)



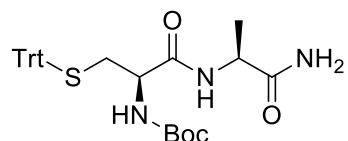
Following the general procedure, to **178** (2.4 g, 3.80 mmol) solution, LiOH (399 mg, 9.5 mmol) was added. The crude product was submitted to silica column chromatography (90-10 % pentane/EA). The solvent was collected and concentrated under vacuum to give **181** as a white solid (74 %).

m.p. 56.2-58.2 °C; $[\alpha]^{20}_D$ +6.474 (*c* 1.00, $CHCl_3$); **1H -NMR (400 MHz, $CDCl_3$) (δ , ppm):** 10.39 (br, 1H), 7.41 (d, *J* = 7.7 Hz, 6H), 7.29 (t, *J* = 7.7 Hz, 6H), 7.21 (t, *J* = 7.3 Hz, 3H), 6.85 (s, 1H), 4.86 (s, 1H), 4.62 (d, *J* = 4.4 Hz, 1H), 3.87 (dd, *J* = 12.3, 7.1 Hz, 1H), 2.77 (dd, *J* = 13.0, 7.1 Hz, 1H), 2.51 (dd, *J* = 13.0, 5.0 Hz, 1H), 2.49 (t, *J* = 7.3 Hz, 2H), 2.16-2.23 (m, 1H), 2.03 (s, 3H), 1.94-2.01 (m, 1H), 1.41 (s, 9H); **^{13}C -NMR (125 MHz, $CDCl_3$) (δ , ppm):** 174.72, 170.74, 155.59, 144.36, 129.56, 128.13, 126.95, 80.75, 67.28, 53.60, 51.81, 33.37, 31.42, 29.75, 28.29, 15.34; **HRMS** (ESI): calcd for $C_{32}H_{38}N_2O_5S_2Na$ $[M + Na]^+$: 617.2120, found 617.2120.

General method for the preparation of amides Boc-Cys(Trt)-aa-NH₂:

Acid **179**, **180** or **181** (1 equiv.) was dissolved in dry THF at -20 °C under argon. NMM (2 equiv.) and IBCF (1.5 equiv.) were added respectively and the reaction left to stir for 30 min. Then, ammonia gas was introduced to the mixture for 10 minutes and was left stirring under argon for overnight at rt. The mixture was concentrated, quenched with water and extracted three times with EA. The organic phase was dried over Na_2SO_4 and evaporated under reduced pressure. Flash chromatography afforded the pure amides **182**, **183** and **184**.

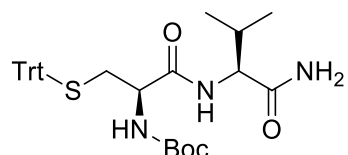
tert-butyl ((R)-1-(((S)-1-amino-1-oxopropan-2-yl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (182)



Following the general procedure, **179** (595 mg, 1.11 mmol), NMM (0.19 mL, 1.7 mmol) and IBCF (0.29 mL, 2.23 mmol) were used. The crude product was submitted to silica column chromatography (90-50 % pentane/EA). The solvent was collected and concentrated under vacuum to give **182** as a white solid (88 %).

m.p. 97.3-101.2 °C; $[\alpha]^{20}_{\text{D}}$ -8.0 (*c* 0.65, CHCl₃); **¹H-NMR (400 MHz, CDCl₃) (δ, ppm):** 7.42 (d, *J* = 7.6 Hz, 6H), 7.30 (t, *J* = 7.5 Hz, 6H), 7.24 (t, *J* = 7.2 Hz, 3H), 6.47 (s, 1H), 6.31 (d, *J* = 7.7 Hz, 1H), 5.21 (s, 1H), 4.74 (d, *J* = 3.7 Hz, 1H), 4.43 (quin, *J* = 7.3 Hz, 1H), 3.75 (d, *J* = 4.9 Hz, 1H), 2.71 (dd, *J* = 12.8, 7.4 Hz, 1H), 2.61 (dd, *J* = 12.8, 5.3 Hz, 1H), 1.42 (s, 9 H) 1.36 (d, *J* = 7.2 Hz, 3H); **¹³C-NMR (100 MHz, CDCl₃) (δ, ppm):** 174.06, 170.29, 155.89, 144.19, 129.46, 128.19, 127.07, 80.91, 67.47, 53.98, 48.57, 33.28, 28.34, 28.23, 17.55.

tert-butyl ((R)-1-(((S)-1-amino-3-methyl-1-oxobutan-2-yl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (183)

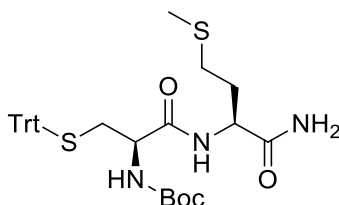


Following the general procedure, **180** (300 mg, 0.53 mmol), NMM (88 μL, 0.8 mmol) and IBCF (138 μL, 1.1 mmol) were used. The crude product was submitted to silica column chromatography (90-80 % pentane/EA). The solvent was collected and concentrated under vacuum to give **183** as a white solid (73 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.45-7.38 (m, 6H), 7.30 (dd, *J* = 10.4, 4.9 Hz, 6H), 7.26-7.18 (m, 3H), 6.53 (s, 2H), 5.62 (s, 1H), 4.85 (d, *J* = 6.0 Hz, 1H), 4.27 (dd, *J* = 8.7, 5.1 Hz, 1H), 3.78 (d, *J* = 4.7 Hz, 1H), 2.71 (dd, *J* = 12.3, 7.0 Hz, 1H), 2.60 (dd, *J* = 13.0, 5.3 Hz, 1H), 2.40-2.19 (m, 1H), 1.41 (s, 9H), 0.92 (d, *J* = 6.9 Hz, 3H), 0.86 (d, *J* = 6.9 Hz, 3H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 173.29, 170.65, 155.82, 144.26, 129.49, 128.18, 127.02,

80.80, 67.39, 57.90, 54.04, 33.07, 29.66, 28.23, 19.35, 17.10; **HRMS** (ESI): calcd for $C_{32}H_{40}N_3O_4S$ $[M + H]^+$: 562.2734, found 562.2726.

tert-butyl ((R)-1-(((S)-1-amino-4-(methylthio)-1-oxobutan-2-yl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (184)



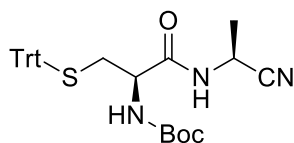
Following the general procedure, **181** (1.25 g, 2.03 mmol), NMM (355 μ L, 3.05 mmol) and IBCF (526 μ L, 4.06 mmol) were used. The crude product was submitted to silica column chromatography (90-50 % pentane/EA). The solvent was collected and concentrated under vacuum to give **184** as a white solid (81 %).

m.p. 85.3-86.9 $^{\circ}$ C; $[\alpha]^{20}_D$ -10.88 (c 1.00, $CHCl_3$); **1H -NMR (400 MHz, $CDCl_3$) (δ , ppm):** 7.41 (d, J = 7.5 Hz, 6H), 7.31 (t, J = 7.5 Hz, 6H), 7.24 (t, J = 7.2 Hz, 3H), 6.92 (s, 1H), 6.66 (s, 1H), 5.27 (s, 1H), 4.72 (s, 1H), 4.55-4.61 (m, 1 H), 3.77 (dd, J = 11.3, 5.2 Hz, 1H), 2.78 (dd, J = 12.9, 6.7 Hz, 1H), 2.58 (dd, J = 12.9, 4.9 Hz, 1H), 2.11-2.23 (m, 1H), 2.04 (s, 1H), 1.95-2.10 (m, 1H), 1.42 (s, 9 H); **^{13}C -NMR (100 MHz, $CDCl_3$) (δ , ppm):** 173.23, 170.52, 155.84, 144.20, 129.48, 128.18, 127.06, 80.85, 67.40, 60.41, 54.13, 52.24, 33.27, 28.27, 15.19; **HRMS** (ESI): calcd for $C_{32}H_{39}N_3O_4S_2Na$ $[M + Na]^+$: 616.2280, found 616.2277.

General method for the preparation of nitriles Boc-Cys(Trt)-aa-CN:

Amide (**182**, **183** and **184**) (1 equiv.) was dissolved in 6 mL DMF. Cyanuric chloride (1 equiv.) was added. The mixture was stirred for 1 h at rt, was quenched with water, and was extracted three times with EA. The organic phase was washed with five times with water, dried, evaporated under vacuum. Flash chromatography afforded the pure nitriles **185**, **186** and **187**.

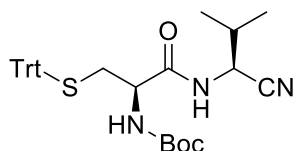
tert-butyl ((R)-1-(((S)-1-cyanoethyl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (185)



Following the general procedure, **182** (440 mg, 0.8 mmol), cyanuric chloride (152 mg, 0.8 mmol) were used. The crude product was submitted to silica column chromatography (90-50 % pentane/EA). The solvent was collected and concentrated under vacuum to give **182** as a white solid (91 %).

m.p. 202.4-203.4 °C; $[\alpha]^{20}_{\text{D}}$ -8.5 (*c* 1.00, CHCl₃); **¹H-NMR (400 MHz, CDCl₃) (δ, ppm):** 7.36 (d, *J* = 7.7 Hz, 6H), 7.23 (t, *J* = 7.5 Hz, 6H), 7.16 (t, *J* = 7.2 Hz, 3H), 6.48 (s, 1H), 4.72 (quin, *J* = 7.4 Hz, 1H), 4.61 (d, *J* = 5.4 Hz, 1H), 3.60 - 3.65 (m, 1H), 2.67 (dd, *J* = 13.3, 7.5 Hz, 1H), 2.46 (dd, *J* = 13.3, 5.0 Hz, 1H), 1.40 (d, *J* = 7.2 Hz, 3H), 1.34 (s, 9H); **¹³C-NMR (100 MHz, CDCl₃) (δ, ppm):** 170.08, 155.76, 144.26, 129.55, 128.17, 127.01, 118.73, 80.88, 77.23, 67.49, 35.75, 29.71, 28.22, 19.36.

tert-butyl ((R)-1-(((S)-1-cyano-2-methylpropyl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (186)



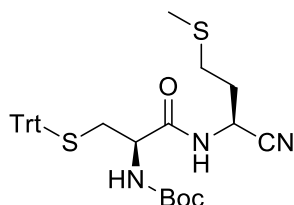
Following the general procedure, **183** (150 mg, 0.27 mmol), cyanuric chloride (49 mg, 0.27 mmol) were used. The crude product was submitted to silica column chromatography (90-80 % pentane/EA). The solvent was collected and concentrated under vacuum to give **186** as a white solid (80 %).

m.p. 178.8-180 °C; $[\alpha]^{20}_{\text{D}}$ -1.864 (*c* 3.60, CHCl₃); **¹H-NMR (500 MHz, CDCl₃) (δ, ppm):** 7.43 (d, *J* = 7.5 Hz, 6H), 7.30 (t, *J* = 7.4 Hz, 6H), 7.27-7.17 (m, *J* = 13.2, 6.1 Hz, 3H), 6.75 (s, 1H), 4.91-4.36 (m, 2H), 3.66 (d, *J* = 5.3 Hz, 1H), 2.84-2.65 (m, 1H), 2.54 (dd, *J* = 13.3, 4.8 Hz, 1H), 2.04-1.82 (m, 1H), 1.41 (s, 9H), 1.01 (t, *J* = 6.0 Hz, 9H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 170.30, 155.90, 144.29, 129.55, 128.18, 127.00, 117.40, 80.85, 67.48,

Experimental part

53.35, 46.37, 32.35, 31.66, 28.23, 18.36, 17.89; **HRMS** (ESI): calcd for C₃₂H₃₇N₃O₃SNa [M + Na]⁺: 566.2453, found 566.2463.

tert-butyl ((R)-1-(((S)-1-cyano-3-(methylthio)propyl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (**187**)



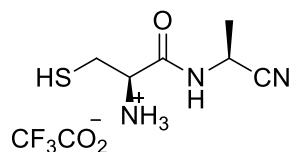
Following the general procedure, **184** (880 mg, 1.4 mmol), cyanuric chloride (263 mg, 1.4 mmol) were used. The crude product was submitted to silica column chromatography (90-70 % pentane/EA). The solvent was collected and concentrated under vacuum to give **187** as a white solid (72 %).

m.p. 126.6-128.2 °C; **¹H-NMR** (400 MHz, CDCl₃) (δ, ppm): 7.42 (d, *J* = 7.7 Hz, 6H), 7.31 (t, *J* = 7.6 Hz, 6H), 7.23 (d, *J* = 7.2 Hz, 3H), 6.73 (d, *J* = 7.5 Hz, 1H), 4.97 (d, *J* = 7.4 Hz, 1H), 4.65 (s, 1H), 3.70 (d, *J* = 5.0 Hz, 1H), 2.59 (t, *J* = 7.1 Hz, 1H), 2.53 (dd, *J* = 13.3, 5.1 Hz, 1H), 2.07 (s, 9H), 2.13-1.97 (m, 2H), 1.42 (s, 9H); **¹³C-NMR** (100 MHz, CDCl₃) (δ, ppm): 170.25, 144.23, 129.53, 128.18, 127.03, 117.67, 80.92, 67.50, 53.50, 39.48, 32.76, 32.10, 29.50, 28.23, 15.41; **HRMS** (ESI): calcd for C₃₂H₃₇N₃O₃S₂Na [M + Na]⁺: 598.2174, found 598.2180.

General method for the preparation of nitriles (Cys-aa-CN):

Nitrile (**185**, **186** or **187**) (1 equiv.) was dissolved in DCM (2 mL), *i*Pr₃SiH (8 equiv.) and TFA (20 equiv.) were added respectively. The mixture was stirred for 1 h and then it was extracted with D₂O (1.5 mL). Finally, the aqueous layer was left under vacuum at room temperature to remove the remaining DCM to provide the required products **188**, **189** and **190** as TFA salts.

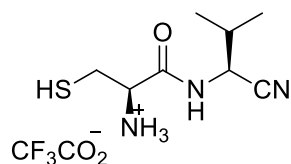
(R)-1-(((S)-1-cyanoethyl)amino)-3-Sulfanyl-1-oxopropan-2-aminium trifluoroacetate (188)



Following the general procedure, **185** (30 mg, 0.058 mmol), *i*Pr₃SiH (85 μL, 0.46 mmol) and TFA (89 μL, 1.16 mmol) were used to give **188** as TFA salt.

¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.77 (q, *J* = 7.2 Hz, 1H), 4.23-4.03 (m, 1H), 3.14-2.88 (m, 2H), 1.52 (d, *J* = 7.3 Hz, 3H); ¹³C-NMR (125 MHz, D₂O) (δ, ppm): 167.65, 119.41, 54.20, 36.78, 24.62, 16.92.

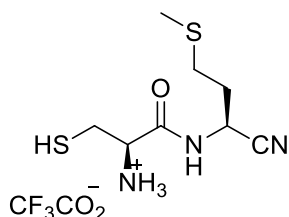
(R)-1-(((S)-1-cyano-2-methylpropyl)amino)-3-Sulfanyl-1-oxopropan-2-aminium trifluoroacetate (189)



Following the general procedure, **186** (42 mg, 0.08 mmol), *i*Pr₃SiH (127 μL, 0.62 mmol) and TFA (120 μL, 1.54 mmol) were used to give **189** as TFA salt.

¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.53 (d, *J* = 7.1 Hz, 1H), 4.10 (t, *J* = 5.8 Hz, 1H), 3.04-2.85 (m, 2H), 2.11-1.97 (m, 1H), 0.95 (d, *J* = 6.7 Hz, 3H), 0.89 (d, *J* = 6.8 Hz, 3H); ¹³C-NMR (125 MHz, D₂O) (δ, ppm): 167.81, 117.98, 54.22, 47.56, 30.40, 24.69, 24.69, 17.75, 17.45.

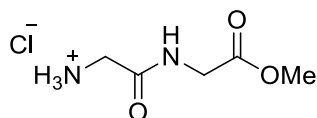
(R)-1-(((S)-1-cyano-3-(methylthio)propyl)amino)-3-Sulfanyl-1-oxopropan-2-aminium trifluoroacetate (190)



Following the general procedure, **187** (30 mg, 0.05 mmol), *i*Pr₃SiH (82 μL, 0.40 mmol) and TFA (77 μL, 1.00 mmol) were used to give **190** as TFA salt.

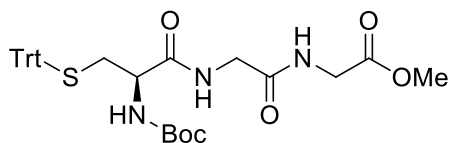
¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.81-4.77 (m, 1H), 4.03 (t, *J* = 5.7 Hz, 1H), 2.96-2.81 (m, 2H), 2.53-2.36 (m, 2H), 2.13-1.95 (m, 2H), 1.90 (s, 3H); **¹³C-NMR (125 MHz, D₂O) (δ, ppm):** 167.79, 118.17, 54.19, 39.92, 30.13, 28.58, 24.52, 13.86.

2-((2-methoxy-2-oxoethyl)amino)-2-oxoethanaminium chloride (191)¹⁷²



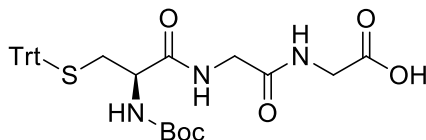
Diglycine (3.3 g, 24.98 mmol) was dissolved in MeOH (75 mL). SOCl₂ (3 mL, 37.47 mmol) was added dropwise at 0 °C and the reaction was left to stir for 3 h at rt. The mixture was concentrated, washed three times with Et₂O ether and evaporated under reduced pressure to afford the 191 as a white solid (quantitative).

¹H-NMR (500 MHz, MeOD) (δ, ppm): 3.91 (s, 1H), 3.65 (s, 1H), 3.61 (s, 1H); **¹³C-NMR (125 MHz, MeOD) (δ, ppm):** 170.17, 166.54, 51.39, 40.44, 40.11.

(R)-methyl 2,2-dimethyl-4,7,10-trioxo-6-((tritylthio)methyl)-3-oxa-5,8,11-triazatridecan-13-oate (192)¹⁷²

Boc-Cys(Trt)-OH (500 mg, 1.08 mmol) was dissolved in dry THF (30 mL) at -20 °C under argon. NMM (261 μ L, 2.37 mmol) and IBCF (168 μ L, 1.29 mmol) were added respectively and the reaction left to stir for 20 min. Then, **191** (217 mg, 1.18 mmol) was added and the mixture was left stirring under argon for overnight at rt. The mixture was concentrated, quenched with water and extracted three times with EA. The organic phase was dried over Na₂SO₄ and evaporated. The crude product was submitted to silica column chromatography (0.5-10 % MeOH/DCM). The solvent was collected and concentrated under vacuum to give **192** as a white solid (50 %).

¹H-NMR (400 MHz, CDCl₃) (δ , ppm): 7.41 (d, J = 7.7 Hz, 6H), 7.30 (t, J = 7.5 Hz, 6H), 7.24 (t, J = 12.2, 4.9 Hz, 3H), 6.97 (s, 1H), 6.50 (t, J = 5.8 Hz, 1H), 4.79 (d, J = 5.1 Hz, 1H), 3.94 (dd, J = 9.4, 4.3 Hz, 2H), 3.89 (d, J = 5.8 Hz, 2H), 3.69 (s, 3H), 2.75 (dd, J = 12.9, 7.3 Hz, 1H), 2.63 (dd, J = 13.0, 5.3 Hz, 1H), 1.40 (s, 9H); **¹³C-NMR (100 MHz, CDCl₃) (δ , ppm):** 170.89, 169.83, 168.95, 155.49, 144.19, 129.49, 128.17, 127.08, 80.92, 67.42, 52.46, 52.23, 42.96, 40.86, 28.24, 18.99.

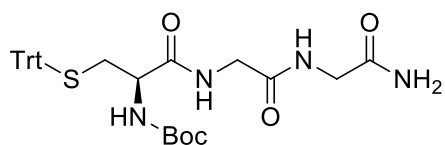
(R)-2,2-dimethyl-4,7,10-trioxo-6-((tritylthio)methyl)-3-oxa-5,8,11-triazatridecan-13-oic acid (193)¹⁷²

192 (470 mg, 0.79 mmol) was dissolved in 30 mL THF/water solution (2/1). LiOH (83.3 mg, 1.99 mmol) was added to the mixture at 0 °C. The mixture was left to stir for 2 h at rt. The solution was concentrated and water was added. Then, NaHSO₄ (sat.) was added to afford pH 2-3. The mixture was extracted three times with EA. The organic layer was collected, dried over Na₂SO₄ and concentrated under reduced pressure. The crude product was submitted to

silica column chromatography (0.5-10 % MeOH/DCM). The solvent was collected and concentrated under vacuum to give **193** as a white solid (41 %).

¹H-NMR (500 MHz, MeOD) (δ, ppm): 7.25 (d, *J* = 7.7 Hz, 6H), 7.11 (t, *J* = 7.6 Hz, 6H), 7.05 (t, *J* = 7.2 Hz, 2H), 3.90-3.75 (m, 2H), 3.75-3.54 (m, 2H), 2.67-2.25 (m, 2H), 1.29 (s, 9H); **¹³C-NMR (125 MHz, MeOD) (δ, ppm):** 175.28, 172.25, 170.39, 156.30, 144.55, 129.36, 127.76, 126.66, 79.82, 66.73, 54.13, 42.49, 42.39, 33.62, 27.52.

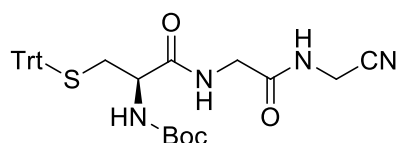
(R)-tert-butyl (1-((2-((2-amino-2-oxoethyl)amino)-2-oxoethyl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (194)



193 (400 mg, 0.69 mmol) was dissolved in dry THF (30 mL) at -20 °C under argon. NMM (114 μL, 1.04 mmol) and IBCF (108 μL, 0.83 mmol) were added respectively and the reaction left to stir for 25 min. Then, ammonia gas was bubbled in the mixture for 5 minutes. The mixture was left stirring under argon for overnight at rt. The excess ammonia was evaporated gently, and the mixture was quenched with water and extracted three times with EA. The organic phase was dried over Na₂SO₄ and evaporated. The crude product was submitted to silica column chromatography (0.5-10 % MeOH/DCM). The solvent was collected and concentrated under vacuum to give **194** as a white solid (40 %).

¹H-NMR (500 MHz, CDCl₃) (δ, ppm): 7.60 (s, 1H), 7.44 (s, 1H), 7.38 (d, *J* = 7.7 Hz, 6H), 7.27 (t, *J* = 9.3, 5.9 Hz, 6H), 7.20 (t, *J* = 7.3 Hz, 3H), 6.78 (s, 1H), 6.16 (s, 1H), 5.43-5.18 (m, 1H), 4.00-3.58 (m, 4H), 2.69-2.56 (m, 2H), 2.19 (s, 1H), 1.38 (s, 9H); **¹³C-NMR (125 MHz, CDCl₃) (δ, ppm):** 172.14, 171.85, 169.86, 156.00, 144.24, 129.50, 128.15, 127.04, 80.78, 67.27, 54.08, 50.52, 43.19, 42.55, 33.61, 28.31.

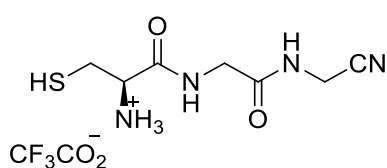
(R)-tert-butyl (1-((2-((cyanomethyl)amino)-2-oxoethyl)amino)-1-oxo-3-(tritylthio)propan-2-yl)carbamate (195)



194 (130 mg, 0.23 mmol) was dissolved in 9 mL DMF. Cyanuric chloride (41.5 mg, 0.23 mmol) was added. The mixture was stirred for 3 h at rt. Then, the mixture was quenched with water, and was extracted three times with EA. The organic phase was washed with five times with water, dried, evaporated under vacuum. The crude product was submitted to silica column chromatography (90-10 % pentane/EA). The solvent was collected and concentrated under vacuum to give **195** as a white solid (72 %).

$[\alpha]^{20}_{\text{D}} +16.4$ (*c* 3.00, CHCl_3); **$^1\text{H-NMR}$ (500 MHz, CDCl_3) (δ , ppm):** 7.61 (s, 1H), 7.40 (d, $J = 7.8$ Hz, 6H), 7.30 (t, $J = 10.3, 4.9$ Hz, 6H), 7.27-7.18 (m, 3H), 6.70 (s, 1H), 4.96 (d, $J = 4.4$ Hz, 1H), 4.02-3.74 (m, 4H), 3.58 (s, 1H), 2.71 (dd, $J = 13.1, 7.6$ Hz, 1H), 2.64 (dd, $J = 13.1, 5.3$ Hz, 1H), 1.42 (s, 9H); **$^{13}\text{C-NMR}$ (125 MHz, CDCl_3) (δ , ppm):** 171.12, 169.39, 156.16, 144.11, 129.48, 128.26, 127.20, 115.92, 81.23, 67.52, 54.39, 42.82, 33.32, 28.34, 27.18.

(R)-1-((2-((cyanomethyl)amino)-2-oxoethyl)amino)-3-Sulfanyl-1-oxopropan-2-aminium trifluoroacetate (196)



195 (62 mg, 0.11 mmol) was dissolved in DCM (2 mL), $i\text{Pr}_3\text{SiH}$ (182 μL , 0.88 mmol) and TFA (171 μL , 2.2 mmol) were added respectively. The mixture was stirred for 1 h and then it was extracted with D_2O (1 mL). Finally, the aqueous layer was left under vacuum at room temperature to remove the remaining DCM to provide the required products **196** as a TFA salt (quantitative).

Experimental part

¹H-NMR (500 MHz, D₂O) (δ, ppm): 4.01 (t, *J* = 5.6 Hz, 1H), 3.93 (s, 2H), 3.75 (s, 2H), 2.82 (d, *J* = 5.6 Hz, 2H); **¹³C-NMR (125 MHz, D₂O) (δ, ppm):** 171.04, 168.46, 116.65, 54.24, 42.00, 27.37, 24.68.

References

- (1) Follmann, H.; Brownson, C. Darwin's Warm Little Pond Revisited: From Molecules to the Origin of Life. *Naturwissenschaften* **2009**, *96* (11), 1265–1292.
- (2) Bada, J. L.; Lazcano, A. Prebiotic Soup—Revisiting the Miller Experiment. *Science* **2003**, *300* (May), 745–746.
- (3) Malik, A. H.; Ziermann, J. M.; Diogo, R. An Untold Story in Biology: The Historical Continuity of Evolutionary Ideas of Muslim Scholars from the 8th Century to Darwin's Time. *J. Biol. Educ.* **2017**, *52* (1), 3–17.
- (4) Alakbarli, F. A 13th-Century Darwin? Tusi's Views on Evolution. *Azerbaijan Int.* **2001**, *9.2*, 48–49.
- (5) Strecker, A. Ueber Die Künstliche Bildung Der Milchsäure Und Einen Neuen, Dem Glycocoll Homologen Körper. *Liebigs Ann. Chem.* **1850**, *75* (1), 27–45.
- (6) Lemmon, R. M. Chemical Evolution. *Chem. Rev.* **1970**, *70* (1), 95–109.
- (7) Benner, S. A. Defining Life. *Astrobiology* **2010**, *10* (10), 1021–1030.
- (8) Lazcano, A. What Is Life? *Chem. Biodivers.* **2008**, *5* (1), 1–15.
- (9) Oparin, A. I. The Origin of Life. *Proiskhozhdenie zhizny* **1924**.
- (10) Marchi, S.; Bottke, W. F.; Elkins-Tanton, L. T.; Bierhaus, M.; Wuennemann, K.; Morbidelli, A.; Kring, D. A. Widespread Mixing and Burial of Earth's Hadean Crust by Asteroid Impacts. *Nature* **2014**, *511* (7511), 578–582.
- (11) Patel, B. H.; Percivalle, C.; Ritson, D. J.; Duffy, C. D.; Sutherland, J. D. Common Origins of RNA, Protein and Lipid Precursors in a Cyanosulfidic Protometabolism. *Nat. Chem.* **2015**, *7* (4), 301–307.
- (12) Castresana, J.; Moreira, D. Respiratory Chains in the Last Common Ancestor of Living Organisms. *J. Mol. Evol.* **1999**, *49* (4), 453–460.
- (13) Sutherland, J. D. Opinion: Studies on the Origin of Life — the End of the Beginning. *Nat. Rev. Chem.* **2017**, *1*, 0012.
- (14) Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, 3rd ed.; Murray: London, **1861**.
- (15) Hug, L. A.; Baker, B. J.; Anantharaman, K.; Brown, C. T.; Probst, A. J.; Castelle, C. J.; Butterfield, C. N.; Hermsdorf, A. W.; Amano, Y.; Ise, K.; Suzuki, Y.; Dudek, N.; Relman, D. A.; Finstad, K. M.; Amundson, R.; Thomas, B. C.; Banfield, J. F. A New View of the Tree of Life. *Nat. Microbiol.* **2016**, *1* (5), 1–6.
- (16) Pinti, D. L. The Origin and Evolution of the Oceans. In *Advances in Astrobiology and Biogeophysics*; **2005**; Vol. I, pp 83–112.
- (17) Wilde, S. A.; Valley, J. W.; Peck, W. H.; Graham, C. M. Evidence from Detrital Zircons for the Existence of Continental Crust and Oceans on the Earth 4.4 Gyr Ago. *Nature* **2001**, *409* (6817), 175–178.
- (18) Walker, J. C. G.; Zahnle, K. J. Lunar Nodal Tide and Distance to the Moon during the Precambrian. *Nature* **1986**, *320* (6063), 600–602.
- (19) Shaw, G. H. Earth's Atmosphere – Hadean to Early Proterozoic. *Chemie der Erde* -

References

- Geochemistry* **2008**, 68 (3), 235–264.
- (20) Martin, R. S.; Mather, T. A.; Pyle, D. M. Volcanic Emissions and the Early Earth Atmosphere. *Geochim. Cosmochim. Acta* **2007**, 71 (15), 3673–3685.
- (21) Zahnle, K.; Schaefer, L.; Fegley, B. Earth's Earliest Atmospheres. *Cold Spring Harb. Perspect. Biol.* **2010**, 2 (10), 1–17.
- (22) Cleaves, H. J. Prebiotic Chemistry: What We Know, What We Don't. *Evol. Educ. Outreach* **2012**, 5 (3), 342–360.
- (23) Santosh, M.; Arai, T.; Maruyama, S. Hadean Earth and Primordial Continents: The Cradle of Prebiotic Life. *Geosci. Front.* **2017**, 8 (2), 309–327.
- (24) Anders, E.; Grevesse, N. Abundances of the Elements: Meteoritic and Solar. *Geochim. Cosmochim. Acta* **1989**, 53 (1), 197–214.
- (25) Kitadai, N.; Maruyama, S. Origins of Building Blocks of Life: A Review. *Geosci. Front.* **2018**, 9 (4), 1117–1153.
- (26) Crampton, M.; Lord, S.; Millar, R. Some Reactions of Ammonia and Primary Amines with Propanal, 2-Chloroethanal, 2, 2-Dichloroethanal and 2, 2, 2-Trichloroethanal in Acetonitrile. *J. Chem.Soc., Perkin Trans. 2* **1997**, 909–916.
- (27) Burton, A. S.; Stern, J. C.; Elsila, J. E.; Glavin, D. P.; Dworkin, J. P. Understanding Prebiotic Chemistry through the Analysis of Extraterrestrial Amino Acids and Nucleobases in Meteorites. *Chem. Soc. Rev.* **2012**, 41 (16), 5459–5472.
- (28) Öberg, K. I.; Guzmán, V. V.; Furuya, K.; Qi, C.; Aikawa, Y.; Andrews, S. M.; Loomis, R.; Wilner, D. J. The Comet-like Composition of a Protoplanetary Disk as Revealed by Complex Cyanides. *Nature* **2015**, 520 (7546), 198–201.
- (29) CHEUNG, A. C.; RANK, D. M.; TOWNES, C. H.; THORNTON, D. D.; WELCH, W. J. Detection of Water in Interstellar Regions by Its Microwave Radiation. *Nature* **1969**, 221 (5181), 626–628.
- (30) Turner, B. E.; Fourikis, N.; Morris, M.; Palmer, P.; Zuckerman, B. Microwave Detection of Interstellar HDO. *Astrophys. J.* **1975**, 198, L125.
- (31) Jacq, T.; Henkel, C.; Walmsley, C. M.; Jewell, P. R.; Baudry, A. H₂(O-18) in Hot Dense Molecular Cloud Cores. *Astron. Astrophys. (ISSN 0004-6361)* **1988**, 199, L5–L8.
- (32) Hjalmarsen, Å.; Bergman, P.; Biver, N.; Florén, H.-G.; Frisk, U.; Hasegawa, T.; Justtanont, K.; Larsson, B.; Lundin, S.; Olberg, M.; Olofsson, H.; Persson, G.; Rydbeck, G.; Sandqvist, A. Recent Astronomy Highlights from the Odin Satellite. *Adv. Sp. Res.* **2005**, 36 (6), 1031–1047.
- (33) Thaddeus, P.; Kutner, M. L.; Penzias, A. A.; Wilson, R. W.; Jefferts, K. B. Interstellar Hydrogen Sulfide. *Astrophys. J.* **1972**, 176 (c), L73.
- (34) Vastel, C.; Phillips, T. G.; Ceccarelli, C.; Pearson, J. First Detection of Doubly Deuterated Hydrogen Sulfide. *Astrophys. J.* **2003**, 593 (2), L97–L100.
- (35) Cheung, A. C.; Rank, D. M.; Townes, C. H.; Thornton, D. D.; Welch, W. J. Detection of NH₃ Molecules in the Interstellar Medium by Their Microwave Emission. *Phys. Rev. Lett.* **1968**, 21 (25), 1701–1705.
- (36) Turner, B. E.; Zuckerman, B.; Morris, M.; Palmer, P. Microwave Detection of Interstellar Deuterated Ammonia. *Astrophys. J.* **1978**, 219, L43.
- (37) Roueff, E.; Tiné, S.; Coudert, L. H.; Pineau des Forêts, G.; Falgarone, E.; Gerin, M.

References

- Detection of Doubly Deuterated Ammonia in L134N. *Energy* **2000**, 354, L63–L66.
- (38) van der Tak, F. F. S.; Schilke, P.; Müller, H. S. P.; Lis, D. C.; Phillips, T. G.; Gerin, M.; Roueff, E. Triply Deuterated Ammonia in NGC 1333. *Astron. Astrophys.* **2002**, 388 (3), L53–L56.
- (39) Wilson, T. L.; Pauls, T. The Detection of Interstellar $N-15/H_3$. *Astron. Astrophys.* **1979**, 73, L10–L12.
- (40) Despois, D.; Crovisier, J.; Bockelee-Morvan, D.; Gerard, E.; Schraml, J. Observations of Hydrogen Cyanide in Comet Halley. *Astron. Astrophys.* **1986**, 160, L11–L12.
- (41) Snyder, L. E.; Buhl, D. Observations of Radio Emission from Interstellar Hydrogen Cyanide. *Astrophys. J.* **1971**, 163 (C), L47.
- (42) Smith, I. W. M.; Talbi, D.; Herbst, E. The Production of HCN Dimer and More Complex Oligomers in Dense Interstellar Clouds. *Astron. Astrophys.* **2001**, 369 (2), 611–615.
- (43) Matthews, C. N.; Minard, R. D. Hydrogen Cyanide Polymers, Comets and the Origin of Life. *Faraday Discuss.* **2006**, 133, 393.
- (44) Hudson, R. L.; Moore, M. H. Reactions of Nitriles in Ices Relevant to Titan, Comets, and the Interstellar Medium: Formation of Cyanate Ion, Ketenimines, and Isonitriles. *Icarus* **2004**, 172 (2), 466–478.
- (45) Belloche, a; Menten, K. M.; Comito, C.; Müller, H. S. P.; Schilke, P.; Ott, J.; Thorwirth, S.; Hieret, C. Detection of Amino Acetonitrile in Sgr B2(N). *Astron. Astrophys.* **2008**, 482 (1), 179–196.
- (46) McGuire, B. A.; Burkhardt, A. M.; Kalenskii, S.; Shingledecker, C. N.; Remijan, A. J.; Herbst, E.; McCarthy, M. C. Detection of the Aromatic Molecule Benzonitrile ($c-C_6H_5CN$) in the Interstellar Medium. *Science* **2018**, 359 (6372), 202–205.
- (47) Snyder, L. E.; Buhl, D.; Zuckerman, B.; Palmer, P. Microwave Detection of Interstellar Formaldehyde. *Am. Phys. Soc.* **1969**, 22 (13), 679–681.
- (48) Fourikis, N.; Sinclair, M.; Robinson, B.; Godfrey, P.; Brown, R. Microwave Emission of the 2-11/2-12 Rotational Transition in Interstellar Acetaldehyde. *Aust. J. Phys.* **1974**, 27 (3), 425.
- (49) Irvine, W. M.; Brown, R. D.; Cragg, D. M.; Friberg, P.; Godfrey, P. D.; Kaifu, N.; Matthews, H. E.; Ohishi, M.; Suzuki, H.; Takeo, H. A New Interstellar Polyatomic Molecule - Detection of Propynal in the Cold Cloud TMC-1. *Astrophys. J.* **1988**, 335, L89.
- (50) Hollis, J. M.; Jewell, P. R.; Lovas, F. J.; Remijan, A.; Møllendal, H. Green Bank Telescope Detection of New Interstellar Aldehydes: Propenal and Propanal. *Astrophys. J.* **2004**, 610 (1), L21–L24.
- (51) Hollis, J. M.; Lovas, F. J.; Jewell, P. R. Interstellar Glycolaldehyde: The First Sugar. *Astrophys. J.* **2000**, 540 (2), L107–L110.
- (52) Pizzarello, S.; Shock, E. The Organic Composition of Carbonaceous Meteorites: The Evolutionary Story Ahead of Biochemistry. *Cold Spring Harb. Perspect. Biol.* **2010**, 2 (3), a002105–a002105.
- (53) Kvenvolden, K.; Lawless, J.; Pering, K.; Peterson, E.; Flores, J.; Ponnampereuma, C.; Kaplan, I. R.; Moore, C. Evidence for Extraterrestrial Amino-Acids and Hydrocarbons in the Murchison Meteorite. *Nature* **1970**, 228 (5275), 923–926.

References

- (54) Pizzarello, S.; Weber, A. L. Prebiotic Amino Acids as Asymmetric Catalysts. *Science* **2004**, *303* (5661), 1151–1151.
- (55) Knowles, D. J.; Wang, T.; Bowie, J. H. Radical Formation of Amino Acid Precursors in Interstellar Regions? Ser, Cys and Asp. *Org. Biomol. Chem.* **2010**, *8* (21), 4934.
- (56) Callahan, M. P.; Smith, K. E.; Cleaves, H. J.; Ruzicka, J.; Stern, J. C.; Glavin, D. P.; House, C. H.; Dworkin, J. P. Carbonaceous Meteorites Contain a Wide Range of Extraterrestrial Nucleobases. *Proc. Natl. Acad. Sci.* **2011**, *108* (34), 13995–13998.
- (57) Stoks, P. G.; Schwartz, A. W. Uracil in Carbonaceous Meteorites. *Nature* **1979**, *282* (5740), 709–710.
- (58) Martins, Z.; Botta, O.; Fogel, M. L.; Sephton, M. A.; Glavin, D. P.; Watson, J. S.; Dworkin, J. P.; Schwartz, A. W.; Ehrenfreund, P. Extraterrestrial Nucleobases in the Murchison Meteorite. *Earth Planet. Sci. Lett.* **2008**, *270* (1–2), 130–136.
- (59) Miller, S. L. A Production of Amino Acids Under Possible Primitive Earth Conditions. *Science* **1953**, *117* (3046), 528–529.
- (60) Bada, J. L. New Insights into Prebiotic Chemistry from Stanley Miller's Spark Discharge Experiments. *Chem. Soc. Rev.* **2013**, *42* (5), 2186–2196.
- (61) Parker, E. T.; Cleaves, H. J.; Callahan, M. P.; Dworkin, J. P.; Glavin, D. P.; Lazcano, A.; Bada, J. L. Prebiotic Synthesis of Methionine and Other Sulfur-Containing Organic Compounds on the Primitive Earth: A Contemporary Reassessment Based on an Unpublished 1958 Stanley Miller Experiment. *Orig. Life Evol. Biosph.* **2011**, *41* (3), 201–212.
- (62) Miller, S. L. The Atmosphere of the Primitive Earth and the Prebiotic Synthesis of Amino Acids. *Orig. Life* **1974**, *5* (1–2), 139–151.
- (63) Van Trump, J. E.; Miller, S. L. Prebiotic Synthesis of Methionine. *Science* **1972**, *178* (4063), 859–860.
- (64) Parker, E. T.; Cleaves, H. J.; Dworkin, J. P.; Glavin, D. P.; Callahan, M.; Aubrey, A.; Lazcano, A.; Bada, J. L. Primordial Synthesis of Amines and Amino Acids in a 1958 Miller H₂S-Rich Spark Discharge Experiment. *Proc. Natl. Acad. Sci.* **2011**, *108* (14), 5526–5531.
- (65) Matthews, C. N. The Origin of Proteins: Heteropolypeptides from Hydrogen Cyanide and Water. *Orig. Life* **1975**, *6* (1–2), 155–162.
- (66) Miller, S. L. Production of Some Organic Compounds under Possible Primitive Earth Conditions 1. *J. Am. Chem. Soc.* **1955**, *77* (9), 2351–2361.
- (67) Shen, C.; Yang, L.; Miller, S. L.; Oró, J. Prebiotic Synthesis of Histidine. *J. Mol. Evol.* **1990**, *31* (3), 167–174.
- (68) Oró, J.; Kimball, A. P. Synthesis of Purines under Possible Primitive Earth Conditions. I. Adenine from Hydrogen Cyanide. *Arch. Biochem. Biophys.* **1961**, *94* (2), 217–227.
- (69) Orgel, L. E. The Origin of Life—a Review of Facts and Speculations. *Trends Biochem. Sci.* **1998**, *23* (12), 491–495.
- (70) Voet, A. B.; Schwartz, A. W. Prebiotic Adenine Synthesis from HCN—Evidence for a Newly Discovered Major Pathway. *Bioorg. Chem.* **1983**, *12* (1), 8–17.
- (71) Ferris, J. P.; Orgel, L. E. Studies in Prebiotic Synthesis. I. Aminomalononitrile and 4-Amino-5-Cyanoimidazole 1,2. *J. Am. Chem. Soc.* **1966**, *88* (16), 3829–3831.
- (72) Fox, S. W.; Harada, K. Synthesis of Uracil under Conditions of a Thermal Model of

References

- Prebiological Chemistry. *Science* **1961**, *133* (3468), 1923–1924.
- (73) Sanchez, R. A.; Ferris, J. P.; Orgel, L. E. Cyanoacetylene in Prebiotic Synthesis. *Science* **1966**, *154* (3750), 784–785.
- (74) Choughuley, A. S. U.; Subbaraman, A. S.; Kazi, Z. A.; Chadha, M. S. A Possible Prebiotic Synthesis of Thymine: Uracil-Formaldehyde-Formic Acid Reaction. *Biosystems* **1977**, *9* (2–3), 73–80.
- (75) Butlerow, A. Bildung Einer Zuckerartigen Substanz Durch Synthese. *Ann. der Chemie und Pharm.* **1861**, *120* (3), 295–298.
- (76) Gabel, N. W.; Ponnampereuma, C. Model for Origin of Monosaccharides. *Nature* **1967**, *216* (5114), 453–455.
- (77) Ricardo, A. Borate Minerals Stabilize Ribose. *Science* **2004**, *303* (5655), 196–196.
- (78) Decker, P.; Schweer, H. Gas—liquid Chromatography on OV-225 of Tetroses and Aldopentoses as Their O-Methoxime and O-n-Butanxime Pertrifluoroacetyl Derivatives and of C3–C6 Alditol Pertrifluoroacetates. *J. Chromatogr. A* **1982**, *236* (2), 369–373.
- (79) Reid, C.; Orgel, L. E. Synthesis of Sugars in Potentially Prebiotic Conditions. *Nature* **1967**, *216* (5114), 455–455.
- (80) Larralde, R.; Robertson, M. P.; Miller, S. L. Rates of Decomposition of Ribose and Other Sugars: Implications for Chemical Evolution. *Proc. Natl. Acad. Sci.* **1995**, *92* (18), 8158–8160.
- (81) Ritson, D.; Sutherland, J. D. Prebiotic Synthesis of Simple Sugars by Photoredox Systems Chemistry. *Nat. Chem.* **2012**, *4* (11), 895–899.
- (82) Zubay, G.; Mui, T. Prebiotic Synthesis of Nucleotides. *Orig. life Evol. Biosph.* **2001**, *31* (1–2), 87–102.
- (83) Fahrenbach, A. C.; Giurgiu, C.; Tam, C. P.; Li, L.; Hongo, Y.; Aono, M.; Szostak, J. W. Common and Potentially Prebiotic Origin for Precursors of Nucleotide Synthesis and Activation. *J. Am. Chem. Soc.* **2017**, *139* (26), 8780–8783.
- (84) Fialho, D. M.; Clarke, K. C.; Moore, M. K.; Schuster, G. B.; Krishnamurthy, R.; Hud, N. V. Glycosylation of a Model Proto-RNA Nucleobase with Non-Ribose Sugars: Implications for the Prebiotic Synthesis of Nucleosides. *Org. Biomol. Chem.* **2018**, *16* (8), 1263–1271.
- (85) Fuller, W. D.; Sanchez, R. A.; Orgel, L. E. Studies in Prebiotic Synthesis. VII. *J. Mol. Evol.* **1972**, *1* (3), 249–257.
- (86) Sanchez, R. A.; Orgel, L. E. Studies in Prebiotic Synthesis. V. Synthesis and Photoanomerization of Pyrimidine Nucleosides. *J. Mol. Biol.* **1970**, *47* (3), 531–543.
- (87) Yamagata, Y.; Watanabe, H.; Saitoh, M.; Namba, T. Volcanic Production of Polyphosphates and Its Relevance to Prebiotic Evolution. *Nature* **1991**, *352* (6335), 516–519.
- (88) Beck, A.; Lohrmann, R.; Orgel, L. E. Phosphorylation with Inorganic Phosphates at Moderate Temperatures. *Science* **1967**, *157* (3791), 952–952.
- (89) Ferris, J. P.; Yanagawa, H.; Hagan, W. J. Prebiotic Synthesis and Reactions of Nucleosides and Nucleotides. *Adv. Sp. Res.* **1983**, *3* (9), 61–68.
- (90) Powner, M. W.; Anastasi, C.; Crowe, M. A.; Parkes, A. L.; Raftery, J.; Sutherland, J. D. On the Prebiotic Synthesis of Ribonucleotides: Photoanomerisation of Cytosine

References

- Nucleosides and Nucleotides Revisited. *ChemBioChem* **2007**, *8* (10), 1170–1179.
- (91) Powner, M. W.; Sutherland, J. D. Potentially Prebiotic Synthesis of Pyrimidine β -D-Ribonucleotides by Photoanomerization/Hydrolysis of α -D-Cytidine-2'-Phosphate. *ChemBioChem* **2008**, *9* (15), 2386–2387.
- (92) Becker, S.; Thoma, I.; Deutsch, A.; Gehrke, T.; Mayer, P.; Zipse, H.; Carell, T. Origin of Life: A High-Yielding, Strictly Regioselective Prebiotic Purine Nucleoside Formation Pathway. *Science* **2016**, *352* (6287), 833–836.
- (93) Danger, G.; Plasson, R.; Pascal, R. Pathways for the Formation and Evolution of Peptides in Prebiotic Environments. *Chem. Soc. Rev.* **2012**, *41* (16), 5416–5429.
- (94) Maciejowska, A.; Godziek, A.; Talik, E.; Sajewicz, M.; Kowalska, T. Investigation of Spontaneous Chiral Conversion and Oscillatory Peptidization of L-Methionine by Means of TLC and HPLC. *J. Liq. Chromatogr. Relat. Technol.* **2015**, *38* (11), 1164–1171.
- (95) Godziek, A.; Maciejowska, A.; Sajewicz, M.; Kowalska, T. Dynamics of Spontaneous Peptidization of L-, d- and DL-Serine in an Abiotic Solution as Investigated with Use of TLC-Densitometry and the Auxiliary Chromatographic Techniques. *J. Chromatogr. Sci.* **2016**, *54* (7), 1090–1095.
- (96) Danger, G.; Charlot, S.; Boiteau, L.; Pascal, R. Activation of Carboxyl Group with Cyanate: Peptide Bond Formation from Dicarboxylic Acids. *Amino Acids* **2012**, *42* (6), 2331–2341.
- (97) Huber, C. Peptides by Activation of Amino Acids with CO on (Ni , Fe) S Surfaces : Implications for the Origin of Life. **1998**, *281* (July), 670–672.
- (98) Bonfio, C.; Valer, L.; Scintilla, S.; Shah, S.; Evans, D. J.; Jin, L.; Szostak, J. W.; Sasselov, D. D.; Sutherland, J. D.; Mansy, S. S. UV-Light-Driven Prebiotic Synthesis of Iron-Sulfur Clusters. *Nat. Chem.* **2017**, *9* (12), 1229–1234.
- (99) Braymer, J. J.; Lill, R. Iron–sulfur Cluster Biogenesis and Trafficking in Mitochondria. *J. Biol. Chem.* **2017**, *292* (31), 12754–12763.
- (100) Lill, R. Function and Biogenesis of Iron-Sulphur Proteins. *Nature* **2009**, *460* (7257), 831–838.
- (101) Rawat, S.; Stemmler, T. L. Key Players and Their Role During Mitochondrial Iron-Sulfur Cluster Biosynthesis. *Chem. - A Eur. J.* **2011**, *17* (3), 746–753.
- (102) Huber, C. Activated Acetic Acid by Carbon Fixation on (Fe,Ni)S Under Primordial Conditions. *Science* **1997**, *276* (5310), 245–247.
- (103) Wächtershäuser, G. From Volcanic Origins of Chemoautotrophic Life to Bacteria, Archaea and Eukarya. *Philos. Trans. R. Soc. B Biol. Sci.* **2006**, *361* (1474), 1787–1806.
- (104) Edwards, M. R. From a Soup or a Seed? Pyritic Metabolic Complexes in the Origin of Life. *Trends Ecol. Evol.* **1998**, *13* (5), 178–181.
- (105) RUSSELL, M. J.; HALL, A. J. The Emergence of Life from Iron Monosulphide Bubbles at a Submarine Hydrothermal Redox and PH Front. *J. Geol. Soc. London.* **1997**, *154* (3), 377–402.
- (106) Rivas, M.; Becerra, A.; Peretó, J.; Bada, J. L.; Lazcano, A. Metalloproteins and the Pyrite-Based Origin of Life: A Critical Assessment. *Orig. Life Evol. Biosph.* **2011**, *41* (4), 347–356.

References

- (107) Mulkidjanian, A. Y. On the Origin of Life in the Zinc World: 1. Photosynthesizing, Porous Edifices Built of Hydrothermally Precipitated Zinc Sulfide as Cradles of Life on Earth. *Biol. Direct* **2009**, *4*, 1–39.
- (108) Mulkidjanian, A. Y.; Galperin, M. Y. On the Origin of Life in the Zinc World. 2. Validation of the Hypothesis on the Photosynthesizing Zinc Sulfide Edifices as Cradles of Life on Earth. *Biol. Direct* **2009**, *4*, 27.
- (109) de Duve, C. A Research Proposal on the Origin of Life. *Orig. Life Evol. Biosph.* **2003**, *33* (6), 559–574.
- (110) Fry, I. The Role of Natural Selection in the Origin of Life. *Orig. Life Evol. Biosph.* **2011**, *41* (1), 3–16.
- (111) Jakubowski, H. Aminoacyl-TRNA Synthetases and the Evolution of Coded Peptide Synthesis: The Thioester World. *FEBS Lett.* **2016**, *590* (4), 469–481.
- (112) White, H. B. Coenzymes as Fossils of an Earlier Metabolic State. *J. Mol. Evol.* **1976**, *7* (2), 101–104.
- (113) Benner, S. A.; Ellington, A. D.; Tauer, A. Modern Metabolism as a Palimpsest of the RNA World. *Proc. Natl. Acad. Sci.* **1989**, *86* (18), 7054–7058.
- (114) Neveu, M.; Kim, H.-J.; Benner, S. A. The “Strong” RNA World Hypothesis: Fifty Years Old. *Astrobiology* **2013**, *13* (4), 391–403.
- (115) Pearce, B. K. D.; Pudritz, R. E.; Semenov, D. A.; Henning, T. K. Origin of the RNA World: The Fate of Nucleobases in Warm Little Ponds. *Proc. Natl. Acad. Sci.* **2017**, *114* (43), 11327–11332.
- (116) Bernhardt, H. S. The RNA World Hypothesis: The Worst Theory of the Early Evolution of Life (except for All the Others). *Biol. Direct* **2012**, *7* (1), 23.
- (117) van der Gulik, P.; Speijer, D. How Amino Acids and Peptides Shaped the RNA World. *life* **2015**, *5* (1), 230–246.
- (118) Carter, C. What RNA World? Why a Peptide/RNA Partnership Merits Renewed Experimental Attention. *life* **2015**, *5* (1), 294–320.
- (119) Saladino, R.; Botta, G.; Pino, S.; Costanzo, G.; Di Mauro, E. Genetics First or Metabolism First? The Formamide Clue. *Chem. Soc. Rev.* **2012**, *41* (16), 5526–5565.
- (120) Palego, L.; Betti, L.; Gino, G. Sulfur Metabolism and Sulfur-Containing Amino Acids: I- Molecular Effectors. *Biochem. Pharmacol. Open Access* **2015**, *04* (01), 1–8.
- (121) Gockel, P.; Vogler, R.; Vahrenkamp, H. Solution Behavior and Zinc Complexation of Dipeptides Made up Solely from Histidine and Cysteine. *Chem. Ber.* **1996**, *129* (8), 887–895.
- (122) Kozłowski, H.; Révérend, B. D. Le; Ficheux, D.; Loucheux, C.; Sovago, I. Nickel(II) Complexes with Sulfhydryl Containing Peptides. Potentiometric and Spectroscopic Studies. *J. Inorg. Biochem.* **1987**, *29* (3), 187–197.
- (123) Xie, X. X.; Ma, Y. F.; Wang, Q. S.; Chen, Z. L.; Liao, R. R.; Pan, Y. C. Yeast CUP1 Protects HeLa Cells against Copper-Induced Stress. *Brazilian J. Med. Biol. Res.* **2015**, *48* (7), 616–621.
- (124) Sóvágó, I.; Kiss, T.; Várnagy, K.; Révérend, B. D.-L. Cobalt(II) and Zinc(II) Complexes of Cysteine Containing Dipeptides. *Polyhedron* **1988**, *7* (12), 1089–1093.
- (125) Leustek, T.; Martin, M. N.; Bick, J.; Davies, J. P. Pathways And Regulation Of Sulfur Metabolism Revealed Through Molecular And Genetic Studies. *Annu. Rev. Plant*

References

- Physiol. Plant Mol. Biol.* **2000**, *51* (1), 141–165.
- (126) Stipanuk, M. H. SULFUR AMINO ACID METABOLISM: Pathways for Production and Removal of Homocysteine and Cysteine. *Annu. Rev. Nutr.* **2004**, *24* (1), 539–577.
- (127) Heinen, W.; Lauwers, A. M. Organic Sulfur Compounds Resulting from the Interaction of Iron Sulfide, Hydrogen Sulfide and Carbon Dioxide in an Anaerobic Aqueous Environment. *Orig. Life Evol. Biosph.* **1996**, *26* (2), 131–150.
- (128) Lavergne, V.; J. Taft, R.; F. Alewood, P. Cysteine-Rich Mini-Proteins in Human Biology. *Curr. Top. Med. Chem.* **2012**, *12* (14), 1514–1533.
- (129) Margulis, L.; Chapman, M.; Guerrero, R.; Hall, J. The Last Eukaryotic Common Ancestor (LECA): Acquisition of Cytoskeletal Motility from Aerotolerant Spirochetes in the Proterozoic Eon. *Proc. Natl. Acad. Sci.* **2006**, *103* (35), 13080–13085.
- (130) Sigalevich, P.; Cohen, Y. Oxygen-Dependent Growth of the Sulfate-Reducing Bacterium *Desulfovibrio Oxycloinae* in Coculture with *Marinobacter* Sp. Strain MB in an Aerated Sulfate-Depleted Chemostat. *Appl. Environ. Microbiol.* **2000**, *66* (11), 5019–5023.
- (131) Škovierová, H.; Vidomanová, E.; Mahmood, S.; Sopková, J.; Drgová, A.; Červeňová, T.; Halašová, E.; Lehotský, J. The Molecular and Cellular Effect of Homocysteine Metabolism Imbalance on Human Health. *Int. J. Mol. Sci.* **2016**, *17* (10), 1733.
- (132) Banerjee, R.; Evande, R.; Kabil, Ö.; Ojha, S.; Taoka, S. Reaction Mechanism and Regulation of Cystathionine β -Synthase. *Biochim. Biophys. Acta - Proteins Proteomics* **2003**, *1647* (1–2), 30–35.
- (133) Baddiley, J.; Thain, E. M.; Novelli, G. D.; Lipmann, F. Structure of Coenzyme A. *Nature* **1953**, *171* (4341), 76–76.
- (134) Krebs, H. A.; Cohen, P. P. Metabolism of Alpha-Ketoglutaric Acid in Animal Tissues. *Biochem. J.* **1939**, *33* (11), 1895–1899.
- (135) Huxtable, R. J. Physiological Actions of Taurine. *Physiol. Rev.* **1992**, *72* (1), 101–163.
- (136) Dalle-Donne, I.; Rossi, R.; Giustarini, D.; Colombo, R.; Milzani, A. S-Glutathionylation in Protein Redox Regulation. *Free Radic. Biol. Med.* **2007**, *43* (6), 883–898.
- (137) Beinert, H.; Holm, R. H.; Münck, E. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science* **1997**, *277* (5326), 653–659.
- (138) Vallee, Y.; Shalayel, I.; Ly, K.-D.; Rao, K. V. R.; De Paëpe, G.; Märker, K.; Milet, A. At the Very Beginning of Life on Earth: The Thiol-Rich Peptide (TRP) World Hypothesis. *Int. J. Dev. Biol.* **2017**, *61* (8–9), 471–478.
- (139) Bourgis, F.; Roje, S.; Nuccio, M. L.; Fisher, D. B.; Tarczynski, M. C.; Li, C.; Herschbach, C.; Rennenberg, H.; Pimenta, M. J.; Shen, T. L.; Gage, D. A.; Hanson, A. D. S-Methylmethionine Plays a Major Role in Phloem Sulfur Transport and Is Synthesized by a Novel Type of Methyltransferase. *Plant Cell* **1999**, *11* (8), 1485–1498.
- (140) Cohen, H.; Salmon, A.; Tietel, Z.; Hacham, Y.; Amir, R. The Relative Contribution of Genes Operating in the S-Methylmethionine Cycle to Methionine Metabolism in *Arabidopsis* Seeds. *Plant Cell Rep.* **2017**, *36* (5), 731–743.
- (141) Ranocha, P.; McNeil, S. D.; Ziemak, M. J.; Li, C.; Tarczynski, M. C.; Hanson, A. D. The S-Methylmethionine Cycle in Angiosperms: Ubiquity, Antiquity and Activity.

References

- Plant J.* **2001**, 25 (5), 575–584.
- (142) Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. Prebiotic Systems Chemistry: New Perspectives for the Origins of Life. *Chem. Rev.* **2014**, 114 (1), 285–366.
- (143) Strecker, A. Ueber Die Künstliche Bildung Der Milchsäure Und Einen Neuen, Dem Glycocoll Homologen Körper; *Ann. der Chemie und Pharm.* **1850**, 75 (1), 27–45.
- (144) Guthrie, J. P.; Yim, J. C. H.; Wang, Q. Hydration of Nitriles: An Examination in Terms of No Barrier Theory. *J. Phys. Org. Chem.* **2014**, 27 (1), 27–37.
- (145) Krimmer, H. P.; Drauz, K.; Kleemann, A. Reaction of β -Mercapto α -Amino Acids with Nitriles. *Chemiker-Zeitung* **1987**, 111, 357–361.
- (146) Berteotti, A.; Vacondio, F.; Lodola, A.; Bassi, M.; Silva, C.; Mor, M.; Cavalli, A. Predicting the Reactivity of Nitrile-Carrying Compounds with Cysteine: A Combined Computational and Experimental Study. *ACS Med. Chem. Lett.* **2014**, 5 (5), 501–505.
- (147) Guillemin, J.-C.; Denis, J.-M.; Lasne, M.-C.; Ripoll, J.-L. Synthèse d'imeses Cycliques Non-Stabilisées Par Reactions Gaz-Solide Sous Vide et Thermolyse-Eclair (1,2). *Tetrahedron* **1988**, 44 (14), 4447–4455.
- (148) Saednya, A. Conversion of Carboxamides to Nitriles Using Trichloroacetyl Chloride/Triethylamine as a Mild Dehydrating Agent. *Synthesis* **1985**, 1985 (02), 184–185.
- (149) Galpin, I. J.; Hancock, G.; Kenner, G. W.; Morgan, B. A. The Synthesis of an Insulin Active Site Analogue. *Tetrahedron* **1983**, 39 (1), 149–158.
- (150) Palumbo Piccionello, A.; Musumeci, R.; Cocuzza, C.; Fortuna, C. G.; Guarcello, A.; Pierro, P.; Pace, A. Synthesis and Preliminary Antibacterial Evaluation of Linezolid-like 1,2,4-Oxadiazole Derivatives. *Eur. J. Med. Chem.* **2012**, 50, 441–448.
- (151) Löser, R.; Schilling, K.; Dimmig, E.; Gütschow, M. Interaction of Papain-like Cysteine Proteases with Dipeptide-Derived Nitriles. *J. Med. Chem.* **2005**, 48 (24), 7688–7707.
- (152) Choudhury, N.; Mete, S.; Kambalapalli, S.; De, P. Side-Chain Glycylglycine-Based Polymer for Simultaneous Sensing and Removal of Copper(II) from Aqueous Medium. *J. Polym. Sci. Part A Polym. Chem.* **2018**.
- (153) Shalayel, I.; Coulibaly, S.; Ly, K.; Milet, A.; Vallée, Y. The Reaction of Aminonitriles with Aminothiols: A Way to Thiol-Containing Peptides and Nitrogen Heterocycles in the Primitive Earth Ocean. *life* **2018**, 8 (4), 47.
- (154) Pinner, A.; Klein, F. Umwandlung Der Nitrile in Imide. *Berichte der Dtsch. Chem. Gesellschaft* **1877**, 10 (2), 1889–1897.
- (155) Boger, D. L.; Keim, H.; Oberhauser, B.; Schreiner, E. P.; Foster, C. A. Total Synthesis of HUN-7293. *J. Am. Chem. Soc.* **1999**, 121 (26), 6197–6205.
- (156) Xiang, Y.; Drenkard, S.; Baumann, K.; Hickey, D.; Eschenmoser, A. Chemie von Alpha-Aminonitrilen. Sondierungen Über Thermische Umwandlungen von Alpha-Aminonitrilen. *Helv. Chim. Acta* **1994**, 77 (8), 2209–2250.
- (157) Song, B. D.; Jencks, W. P. Aminolysis of Benzoyl Fluorides in Water. *J. Am. Chem. Soc.* **1989**, 111 (22), 8479–8484.
- (158) Elkholy, Y. M.; Erian, A. W. An Aminoimidazole and Its Utility in Heterocyclic Synthesis. *Heteroat. Chem.* **2003**, 14 (6), 503–508.
- (159) Almaliti, J.; Al-Hamashi, A. A.; Negmeldin, A. T.; Hanigan, C. L.; Perera, L.; Pflum,

References

- M. K. H.; Casero, R. A.; Tillekeratne, L. M. V. Largazole Analogues Embodying Radical Changes in the Depsipeptide Ring: Development of a More Selective and Highly Potent Analogue. *J. Med. Chem.* **2016**, *59* (23), 10642–10660.
- (160) Mai, K.; Patil, G. Facile Synthesis of α -Aminonitriles. *Tetrahedron Lett.* **1984**, *25* (41), 4583–4586.
- (161) Du Vigneaud, V.; Wilbur I. Patterson and; Hunt, M. Opening of the Ring of the Thiolactone of Homocysteine. *J. Biol. Chem.* **1938**, *3* (126), 217–231.
- (162) Bulaj, G.; Kortemme, T.; Goldenberg, D. P. Ionization-Reactivity Relationships for Cysteine Thiols in Polypeptides. *Biochemistry* **1998**, *37* (25), 8965–8972.
- (163) Wu, Y.; Brosh, R. M. DNA Helicase and Helicase-Nuclease Enzymes with a Conserved Iron-Sulfur Cluster. *Nucleic Acids Res.* **2012**, *40* (10), 4247–4260.
- (164) Boulanger, Y.; Goodman, C. M.; Forte, C. P.; Fesik, S. W.; Armitage, I. M. Model for Mammalian Metallothionein Structure. *Proc. Natl. Acad. Sci.* **1983**, *80* (6), 1501–1505.
- (165) Baker, D. H.; Czarnecki-Maulden, G. L. Pharmacologic Role of Cysteine in Ameliorating or Exacerbating Mineral Toxicities. *J. Nutr.* **1987**, *117* (6), 1003–1010.
- (166) Hall, A. C.; Young, B. W.; Bremner, I. Intestinal Metallothionein and the Mutual Antagonism between Copper and Zinc in the Rat. *J. Inorg. Biochem.* **1979**.
- (167) Suzuki, M.; Makimura, K.; Matsuoka, S. I. Thiol-Mediated Controlled Ring-Opening Polymerization of Cysteine-Derived β -Thiolactone and Unique Features of Product Polythioester. *Biomacromolecules* **2016**, *17* (3), 1135–1141.
- (168) Iwakura, Y.; Kurita, K.; Hayako, F. Polythiazolines. *J. Polym. Sci.* **1969**, *7*, 3075–3087.
- (169) Han, F. S.; Osajima, H.; Cheung, M.; Tokuyama, H.; Fukuyama, T. Synthesis of Chiral Cyclic Oligothiazolines: A Novel Structural Motif for a Macrocyclic Molecule. *Chem. Commun.* **2006**, No. 16, 1757.
- (170) Ruiz-Rodríguez, J.; Miguel, M.; Preciado, S.; Acosta, G. A.; Adan, J.; Bidon-Chanal, A.; Luque, F. J.; Mitjans, F.; Lavilla, R.; Albericio, F. Polythiazole Linkers as Functional Rigid Connectors: A New RGD Cyclopeptide with Enhanced Integrin Selectivity. *Chem. Sci.* **2014**, *5* (10), 3929–3935.
- (171) Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. Synthesis of Proteins by Native Chemical Ligation. *Science* **1994**, *266* (5186), 776–779.
- (172) Lu, X.; Olsen, S. K.; Capili, A. D.; Cisar, J. S.; Lima, C. D.; Tan, D. S. Designed Semisynthetic Protein Inhibitors of Ub/Ubl E1 Activating Enzymes. *J. Am. Chem. Soc.* **2010**, *132* (6), 1748–1749.
- (173) 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 Lett.* **2018**, *59* (23), 2264–2267.
- (174) McLaughlin, M.; Mohareb, R. M.; Rapoport, H. An Efficient Procedure for the Preparation of 4-Substituted 5-Aminoimidazoles. *J. Org. Chem.* **2003**, *68* (1), 50–54.
- (175) Cianci, J.; Baell, J. B.; Harvey, A. J. Microwave-Assisted, Zinc-Mediated Peptide Coupling of N-Benzyl- α,α -Disubstituted Amino Acids. *Tetrahedron Lett.* **2007**, *48* (34), 5973–5975.
- (176) Ezawa, T.; Kawashima, Y.; Noguchi, T.; Jung, S.; Imai, N. Amidation of Carboxylic

References

- Acids via the Mixed Carbonic Carboxylic Anhydrides and Its Application to Synthesis of Antidepressant (1 S ,2 R)-Tranylcypramine. *Tetrahedron: Asymmetry* **2017**, *28* (12), 1690–1699.
- (177) Sureshbabu, V. V.; Venkataramanarao, R.; Naik, S. A.; Chennakrishnareddy, G. Synthesis of Tetrazole Analogues of Amino Acids Using Fmoc Chemistry: Isolation of Amino Free Tetrazoles and Their Incorporation into Peptides. *Tetrahedron Lett.* **2007**, *48* (39), 7038–7041.
- (178) Huber, T. D.; Wang, F.; Singh, S.; Johnson, B. R.; Zhang, J.; Sunkara, M.; Van Lanen, S. G.; Morris, A. J.; Phillips, G. N.; Thorson, J. S. Functional AdoMet Isosteres Resistant to Classical AdoMet Degradation Pathways. *ACS Chem. Biol.* **2016**, *11* (9), 2484–2491.
- (179) Xie, S.; Savchenko, A. I.; Kerscher, M.; Grange, R. L.; Krenske, E. H.; Harmer, J. R.; Bauer, M. J.; Broit, N.; Watters, D. J.; Boyle, G. M.; Bernhardt, P. V.; Parsons, P. G.; Comba, P.; Gahan, L. R.; Williams, C. M. Heteroatom-Interchanged Isomers of Lissoclinamide 5: Copper(II) Complexation, Halide Binding, and Biological Activity. *European J. Org. Chem.* **2018**.
- (180) Herrera, R. P.; Sgarzani, V.; Bernardi, L.; Fini, F.; Pettersen, D.; Ricci, A. Phase Transfer Catalyzed Enantioselective Strecker Reactions of α -Amido Sulfones with Cyanohydrins. *J. Org. Chem.* **2006**, *71* (26), 9869–9872.
- (181) Ke, D.; Zhan, C.; Li, X.; Li, A.; Yao, J. A Simple Synthesis of Imide-Dipeptides. *Synlett* **2009**, *2009* (09), 1506–1510.
- (182) Beauchard, A.; Twum, E. A.; Lloyd, M. D.; Threadgill, M. D. S-2-Amino-4-Cyanobutanoic Acid (β -Cyanomethyl-l-Ala) as an Atom-Efficient Solubilising Synthone for l-Glutamine. *Tetrahedron Lett.* **2011**, *52* (41), 5311–5314.
- (183) D'Arrigo, P.; Arosio, D.; Cerioli, L.; Moscatelli, D.; Servi, S.; Viani, F.; Tessaro, D. Base Catalyzed Racemization of Amino Acid Derivatives. *Tetrahedron: Asymmetry* **2011**, *22* (8), 851–856.

Life emerged on Earth probably about 3.8 billion years ago, on a planet that was largely covered by water. This work focuses on the prebiotic synthesis of peptides, especially thiol-rich ones. We studied the reactions of aminonitriles (the first products of the Strecker reaction) with cysteine and homocysteine. These reactions lead to the formation of 5- or 6-membered rings which are then hydrolysed to give the corresponding dipeptides (aa-Cys or aa-Hcy). The obtained thiol-containing dipeptides are able to promote the formation of longer peptide chains via thioesters bonds, and to promote the formation of some heterocycles. Homocysteine nitrile cyclizes in water to form homocysteine thiolactone, which shows a double reactivity, thiolactone opening by amines followed by aminothiol condensation reaction with nitriles. Cysteine nitrile and the *S*-ethyl thioester of cysteine lead to the formation of polycysteine, while Cys-aa-CN molecules gives linear and cyclic polypeptides. Our results support the hypothesis that thiol-containing peptides would have been important molecules in the early stages of life development.

Keywords: origin of life, prebiotic chemistry, metabolism first hypothesis, thiol-rich peptides, aminonitriles, amidonitriles, cysteine, homocysteine, imidazoles, polycysteine

La vie a émergé sur Terre il y a probablement 3,8 milliards d'années, sur une planète largement recouverte d'eau. Ce travail porte sur la synthèse prébiotique de peptides, en particulier de peptides riches en thiol. Nous avons étudié les réactions des aminonitriles (les premiers produits de la réaction de Strecker) avec la cystéine et l'homocystéine. Elles conduisent à la formation de cycles à 5 ou 6 chaînons qui sont ensuite hydrolysés pour donner les dipeptides correspondants (aa-Cys ou aa-Hcy). Les dipeptides contenant un thiol obtenus sont capables de favoriser la formation de chaînes peptidiques plus longues via des liaisons thioesters et de favoriser la formation de certains hétérocycles. L'homocystéine nitrile se cyclise dans l'eau pour former l'homocystéine thiolactone, qui présente une double réactivité, la thiolactone est ouverte par des amines puis on observe une condensation de l'aminothiol ainsi formé avec les nitriles. Le nitrile de cystéine et le thioester de *S*-éthyle de la cystéine conduisent à la formation de polycystéine, tandis que les molécules de type Cys-aa-CN donnent des polypeptides linéaires et cycliques. Nos résultats soutiennent l'hypothèse que des peptides contenant des thiols auraient joué un rôle important dans les premiers stades du développement de la vie.

Mots-clés : origine de la vie, chimie prébiotique, hypothèse « métabolisme d'abord », peptides riches en thiols, aminonitriles, amidonitriles, cystéine, homocystéine, imidazoles, polycystéine